warwick.ac.uk/lib-publications

Exploring biomolecules, metallodrugs, and their interactions via the use of UHR-FT-ICR Mass Spectrometry

Christopher A. Wootton

A thesis submitted for the degree of

Doctor of Philosophy

Department of Chemistry

University of Warwick

April 2016

Acknowledgements

Author information and *curriculum vite*

Education:

4.5 A-levels at NCN high pavement (Nottingham, UK). Chemistry, Physics, Biology, Mathematics.

MChem 4 year course at University of Warwick, UK. 3 Year pure Chemistry BSc course followed by 1 year Masters course including 9 month research project involving the design, synthesis, and characterisation of block-co-polymer unimolecular micelles for the targeted transportation of drugs to target sites, under the supervision of Dr. Remzi Becar.

PhD in analytical chemistry at the University of Warwick, UK under the supervision of Prof. Peter B. O'Connor and Prof. Peter J. Sadler. Main project focus was using ultra high resolution mass spectrometry (FT-ICR MS) to study the mechanism of action of organometallic and photoactivatable metal-based anticancer complexes, especially with respect to biomolecules such as peptides, proteins, and DNA.

Papers and patents:

Patents:

- Superacid supercharging ESI source ""IONISING MOLECULES AND ELECTROSPRAY IONISATION APPARATUS" Peter B. O'Connor, **Christopher A. Wootton**, Haytham Hussein, Man-ying Wong, 2016 "

Accepted papers:

- **Christopher A. Wootton**, Carlos Sanchez-Cano, Hong-Ke Liu, Mark P. Barrow, Peter J. Sadler, and Peter B. O'Connor. Binding of an organo–osmium(II) anticancer complex to guanine and cytosine on DNA revealed by electron-based dissociations in high resolution Top–Down FT-ICR mass spectrometry, Dalton Transactions, 2015, 44, 3624 - 3632, DOI: 10.1039/c4dt03819c
- Maria A. van Agthoven, Mark P. Barrow, Lionel Chiron, Marie-Aude Coutouly, David Kilgour, **Christopher A. Wootton**, Juan Wei, Andrew Soulby, Marc-André Delsuc, Christian Rolando, Peter B. O'Connor. Differentiating Fragmentation Pathways of Cholesterol by Two-Dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, JASMS, 2015, 1044-0305, DOI:10.1007/s13361-015-1226-7
- Maria A. van Agthoven, **Christopher A. Wootton**, Lionel Chiron, Marie-Aude Coutouly, Andrew J. Soulby, Juan Wei, Mark P. Barrow, Marc-André Delsuc, Christian Rolando, and Peter B O'Connor. Two-Dimensional Mass Spectrometry for Proteomics, a Comparative Study with Cytochrome C, Analytical Chemistry, 2016, DOI:10.1021/acs.analchem.5b04878
- Floris, F.; van Aghtoven, M.A.; Chiron L.; **Wootton C. A**.; Lam P.Y.; Delsuc M.-A.; Barrow P.M.; O'Connor, P.B. - Two-dimensional Fourier Transform Ion Cyclotron Resonance Mass Spectrometry of Calmodulin: a top-down and bottom-up approach. J. Am. Soc. Mass Spectrom. (2016). DOI:10.1007/s13361-016-1431-z

Submitted papers:

- "Automatic Assignment of Metal-containing Peptides in Proteomic LC-MS and MS/MS Datasets" **Christopher A. Wootton**, Yuko P. Y. Lam, Matthew Willetts, Maria A. van Agthoven, Mark P. Barrow, Peter J. Sadler, Peter B. O'Connor. Submitted to RSC Analyst, 2016.

- "Molecular "Twister" for facile sequential deuteration of methyl groups on cyclopentadienyl rings" J. J. Soldevila-Barreda, Juliusz Wolny, **Christopher A. Wootton**, Abraha Habtemariam, Guy J. Clarkson, Peter B. O'Connor, Volker Schunemann, and Peter J. Sadler. Submitted to Nature Chemistry, 2016.
- "Islet Amyloid Polypeptide as a Molecular Clock: Does Deamidation lead to Dimerization?" Lam, Yuko; **Wootton, Christopher A.**; Wei, Juan; Hands-Portman, Ian; Barrow, Mark; O'Connor, Peter. Submitted to Journal of the American Chemical Society (JACS), 2016.

Manuscripts in preparation:

- "Superacid Supercharging using CH_5^+" **Christopher A. Wootton** and Haytham E. M. Hussein, Man Ying Wong, Cookson K. C. Chiu, Lewis A. Baker, Scott Habershon, Anthony J. Stace, Mark P. Barrow, Peter B. O'Connor
- "The challenges of electron capture sequencing with an electron-quenching modification" **Christopher A. Wootton**, Adam Millet, Andrea Lopez-Clavijo, Mark Barrow, Peter J. Sadler, and Peter B. O'Connor

Conference presentations and awards:

Poster presentations and awards:

- American society of mass spectrometry annual conference (ASMS 2013) - Minneapolis, US
- East Midlands Proteomic Workshop (EMPW 2013) - Leicester, UK
- British Mass Spectrometry Society annual conference (BMSS 2013) - Eastbourne, UK
- Uppsala conference on electron-based dissociation techniques (Uppcon 2014) - Obergurgl, Austria - Awarded Poster Prize by The Analyst journal
- European Fourier Transform Mass Spectrometry conference (EFTMS 2014) - Paris, France
- Uppsala conference on electron-based dissociation techniques (Uppcon 2015) - California, US - Poster + 5 minute "lightning talk"
- Isolated Biomolecules and Biomolecular Interactions conference (IBBI 2016) - Oxford, UK - Awarded Poster Prize by the IBBI conference organisers

Oral presentations and awards:

- East Midlands Proteomic Workshop (EMPW 2014) - Leicester, UK - Awarded Oral presentation prize for best Early Career Researcher talk
- American society of mass spectrometry annual conference (ASMS 2015) - St. Louis, US - Presented in the Top-down Protein Analysis session
- British Mass Spectrometry Society annual conference (BMSS 2015) - Birmingham, UK - Presented in the Top-down Protein session
- East Midlands Proteomic Workshop (EMPW 2015) - University of Warwick, UK
- Annual Tandem Mass Spectrometry Workshop, Lake Louise, Canada - December 2015 -Awarded a competitive $2000 grant to present talk by Lake Louise conference organisers
- European Fourier Transform Mass Spectrometry conference (EFTMS 2016) - Matera, Italy - Presented in Nanodroplets, Ion sources, and clusters session
- Proteomics Methods Forum (PMF) 2016 - Warwick, UK - Presented in the Diverse Proteomics session
- British Mass Spectrometry Society annual conference (BMSS 2016) - Eastbourne, UK - Confirmed speaker in Fundamentals and Instrument Developments session

Declaration

I hereby declare that except where specifically stated/references are made to other sources, the thesis entitled 'Exploring biomolecules, metallodrugs, and their interactions via the use of UHR-FT-ICR Mass Spectrometry' is the original work of the named Author. It has been composed by myself and co-authors where stated and has not been submitted in whole or in part for any other degree, diploma, or qualification.

Christopher A. Wootton

April 2016

Abstract

The work presented herein focuses on the study of novel metallodrugs and their interaction with various possible targets and off-targets in the form of biomolecules such as peptides, proteins, DNA, and small molecules via the use of ultra-high resolution Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry.

Beyond traditional platinum(II) metallodrug such as cisplatin and oxaliplatin, new designs of metallodrugs are being conceived to attack and kill cancerous cells via new mechanisms of action in an effort to exceed the potency, selectivity, and effectiveness of metallodrug treatments. These new metallodrugs range from various activation strategies, to specific target binding, and even catalysis of cell-disrupting processes. As a result these novel drugs can have a wide array of effects on various biomolecule species the may encounter in the body.

It is the aim of this thesis to show that mass spectrometry, specifically FT-ICR Mass Spectrometry, is uniquely suited to studying the wide array of metal-based drugs, their biomolecule targets/off-targets, and the numerous reaction products produced from their interactions. Though the study of metallodrug-modified biomolecules via mass spectrometry was shown to be challenging in many cases, mass spectrometry is currently the only analytical technique viable for studying the complex systems involved to a useful level of detail. The majority of the thesis focuses on study of isolated biomolecules interacting with novel metallodrugs and the MS and tandem-MS based study of the resulting observable components.

A great range of metallodrugs and biomolecule interactions were observed. A photoactivated platinum(IV) metallodrug (*trans, trans, trans*[Pt(py)$_2$(OH)$_2$(N$_3$)$_2$]) was shown to produce a variety of platinum(II) based modifications to the peptides studied when activated with blue (463nm) visible light, with the ligand configurations varying depending on whether a histidine amino acid residue was present, allowing retention of both pyridine ligands, or not, allowing release of any of the bound ligands. Tandem MS studies using electron based dissociations showed the primary preference of

binding to be at the Histidine residues, and when not available the complexes could bind to lysing functional groups, distinctly different behaviour to previously studied Pt(II) complexes. Oxidation of peptide species was also found to be a significant product of these reactions. Tandem MS studies located the oxidation sites to methionine and tryptophan residues, the latter of which provided insights into the oxidation mechanism. The oxidation process was found to be due to a hydroxyl radical process, not a singlet oxygen mechanism. UV/vis and EPR measurements were also undertaken and supported the results found.

Studies into the interaction of a potent osmium(III) compound with isolated DNA strands showed that the metallodrug could bind to both guanine and cytosine sites along the biomolecule. Analysis of these species via CAD MS/MS proved challenging due to damage and eventual dissociation of the metal complex modifications. Whereas electron detachment dissociation enabled the elucidation of the two distinct binding locations. CAD MS/MS was found to be useful for studying hydrogen bonded/pi-stacking stabilised structures.

Further investigations into DNA MS and MS/MS lead to the study of the platinum(IV) compound interactions with DNA stands, showing a vastly improved rate of reaction for the compound when compared to the previous peptide reactions. In addition uniform retention of pyridine ligands and no oxidation was observed, providing further evidence of the biomolecule playing a key role in the activation process of the Pt(IV) drugs. MS/MS studies of platinated DNA resulted in similar findings to the osmium-based metallodrugs. Experiments with native MS of DNA showed that observation and interrogation of duplex-DNA strands is possible, even with DNA strands which are not stabilised easily in ammonium acetate solutions, and without annealing. MS, MS/MS, and MS^3 was achieved on duplex DNA strands using large concentrations of ammonium acetate and potassium chloride solutions, along with extremely carefully tuned MS source and transfer parameters.

The effects of metals on tandem-MS techniques was uniquely apparent during the study of functionalised Iridium-based metallodrug modified peptides. The iridium complexes were shown to

effectively quench many electrons used during ECD MS/MS. Though the species were still able to be studied using optimised ECD and CAD MS/MS parameters.

The reactivity of metal centres was also shown to affect their own bound ligands, as observed herein with rhodium(II) compounds. The rhodium piano-stool complexes were shown to vastly accelerate and enable room temperature activation of C-H bonds within arene ligands towards hydrogen-deuterium exchange experiments, with some compounds achieving full exchange of available groups within just 1 hour. The process was studied using FT-ICR MS to track the exchange process and observe sequential exchange for 10 different rhodium compounds. An iridium analogue was also studied, which was ineffective, displaying the Rh metal centre's unique chemistry for this reaction.

A novel method for the enhancement of the electrospray ionisation process was conceived and developed in order to achieve improvement of analyte charging, so-called "supercharging". Using a combination of standard electrospray ionisation and atmospheric chemical ionisation, CH_5^+ "superacid" ions were introduced into the ESI plume and enabled solution-additive free supercharging of analytes under a variety of conditions. The achievement of higher charge states/enhancement of charge is of uniform benefit to biomolecule characterisation and could help reduce the need for solution phase additives which can disrupt many chemical processes and biomolecule structure.

The final section of the thesis is concerned with the escalation of metallodrug-biomolecule interaction from isolated biomolecule up to full-cell proteomes via FT-ICR MS. The great array of challenges faced in previous studies is addressed and strategies for accurate and reliable studies of large metallodrug-modified systems are outlined and tested. Two major strategies are proposed, one based on liquid-chromatography mass spectrometry with modified data processing techniques. The other using a niche MS/MS technique known as two-dimensional mass spectrometry, which would enable whole proteome characterisation without chromatographic separation. Preliminary result using both approaches and future outlook are presented.

Table of Contents:

List of Figures, tables, and schemes:

Chapter 3: Photoactivatable Pt(IV) biomolecule interactions:

Chapter 3 - SI:

platinum based modification associated with each species (shown in Figure 3, main text, and

correlating tables of assignments, below).

Chapter 4: Osmium metallodrug-DNA interactions:

Chapter 5 - Further DNA-metallodrug investigations:

Chapter 5 – SI:

Chapter 6 - Studying electron-quenching Iridium metallodrug-modified biomolecules using CAD and ECD FT-ICR MS/MS:

Chapter 6 – SI:

Chapter 7 – Superacid Supercharging using CH_5^+:

Chapter 8 – The Rhodium revolver:

Chapter 9 – Scaling up to proteomics:

phosphopeptides in the proteomic sample, a comparison of this with the untreated cell

digest could reveal differences in cell signalling due to metallodrug exposure.   **Page:** 434

Chapter 10 – Conclusions:

Abbreviations

AC	Alternating Current
ACN	Acetonitrile
AI-	Activated Ion-
AI-ECD	Activated Ion-Electron Capture Dissociation
APCI	Atmospheric Pressure Chemical Ionisation
APPI	Atmospheric Pressure Photo-ionisation
CAD	Collisionally Activated Dissociation aka CID
CASI	Continuous Accumulation of Selected Ions
CDDP	Cisplatin
CEM	Chain Ejection Model
CHEF	Correlated Harmonic Excitation Field
CI	Chemical Ionisation
CID	Collisionally Induced Dissociation aka CAD
Complex 1	$trans,trans,trans$-[Pt(N$_3$)$_2$(OH)$_2$(pyridine)$_2$]
CRAFTI	Cross sectional areas by Fourier Transform Ions
CRM	Charge Residue Model
CRS	Charge Reduced Species
CytC	Cytochrome C
DC	Direct Current
DMSO	DimethylSulfoxide
DNA	Deoxyribose Nucleic Acid
DNA12	DNA strand with the sequence TAGTAATTACTA
ECD	Electron Capture Dissociation
EDD	Elecron Detachment Dissociation

EI	Electron Impact / Electron Ionisation
EID	Electron-Induced Dissociation aka Electron Ionisation Dissociation
EIEIO	Electron Ionisation of Electronically Excited Organics
EPR	Electron Paramagnetic Resonance
ESI	Electrospray Ionisation
ETD	Electron Transfer Dissociation
ETnoD	Electron Transfer no Dissociation
eV	electron Volt
ExD	Electron x Dissociation - generalised term
FFT	Fast-Fourier Transform
FID	Free Induction Decay
FM190	$trans,trans,trans$-$[Pt(N_3)_2(OH)_2(pyridine)_2]$
FRCM	Free Radical Cascade Mechanism
FT	Fourier Transform
FT-ICR	Fourier Transform-Ion Cyclotron Resonance
FTMS	Fourier Transform Mass Spectrometry
GAG	Glycosaminoglycan
gamma	frequency
gamma c	Cyclotron frequency
gamma m	magnetron motion frequency
gamme t	trapping frequency
GC	Gas Chromatography
GCMS	Gas Chromatography Mass Spectrometry
GSH	Glutathione
hECD	hot-Electron Capture Dissociation

HG	Haemoglobin
HOMO	Highest Occupied Molecular Orbital
hot-ECD	hot-Electron Capture Dissociation
HPLC	High Pressure Liquid Chromatography
HSA	Human Serum Albumin
IC50	Inhibition concentration needed for 50% effectiveness
ICR	Ion Cyclotron Resonance
IEM	Ion Ejection Model
IMS	Ion Mobility Spectrometry
IR-	Infra-Red-
Ir0	Iridium piano stool compound - see chapter 6
Ir1	Di-aldehyde functionalised Iridium piano stool compound - see chapter 6
Ir2	mono-aldehyde functionalised Iridium piano stool compound - see chapter 6
Ir3	mono-aldehyde functionalised Iridium piano stool compound - see chapter 6
IR-ECD	Infra-Red-Electron Capture Dissociation
IRMPD	Infra-Red Multi-Photon Dissociation
ISD	In-Source Dissociation
K3-Bom	[Lys]3-Bombesin peptide with the sequence PyrQKLGNQWAVGHLM-NH2
L3BB3	[Lys]3-Bombesin peptide with the sequence PyrQKLGNQWAVGHLM-NH2
LC	Liquid Chromatography
LCMS	Liquid Chromatography Mass Spectrometry
LUMO	Lowest Unoccupied Molecular Orbital
LZ	Lysozyme
m/z	mass to charge ratio
MAI	Matrix-Assisted Ionisation

MALDI	Matrix-Assisted Laser Desorption Ionisation
MD	Molecular Dynamics
MeOH	Methanol
MG	Myoglobin
MLCT	Metal-to-Ligand Charge Transfer
Mn	number weighted average
m-NBA	meta-Nitrobenzylalcohol
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MS/MS/MS	Two stage Tandem Mass spectrometry aka MS3
MS3	Two stage Tandem Mass spectrometry aka MS/MS/MS
MSn	Mass Spectrometry with n stages
Mw	Weighted average
NAD+	Nicotinamide in its oxidised form
NADH	Nicotinamide in its reduced form
nESI	nano-Electrospray Ionisation
nETD	negative-Electron Transfer Dissociation
niECD	negative ion- Electron Capture Dissociation
nLC	nano-Liquid Chromatography
NMR	Nuclear Magnetic Resonance
Os1	Os1-Cl after loss of the Chlorido ligand and counterion
Os1-Cl	Osmium piano stool complex [(η6-bip)Os(en)Cl]PF6 (Os1-Cl·PF6)
PACT	Photo-Activated Chemo-Therapy
ppm	part-per-million
PTM	Post-Translational Modification

RF	Radio Frequency
RNA	Ribose Nucleic Acid
RP	Resolving Power
SC	Super Charging
SCL	Side-Chain Loss
SIMS	Secondary Ionisation Mass Spectrometry
SORI	Sustained Off-Resonance Irradiation
SORI-CID	Sustained Off-Resonance Irradiation-Collisionally Induced Dissociation
SubP	Substance P peptide with the sequence RPKPQQFFGLM-NH2
SWIFT	Stored-Waveform Inverse Fourier Transform
Tandem-MS	Tandem Mass Spectrometry
TD	Acquisition time/time duration of transient
ToF	Time of Flight
TrxR	Thioredoxin Reductase
UV/vis	Ultra-Violet/Visible spectrometry
UVPD	Ultra-Violet Photo-Dissociation
UW-	Utah-Washington-
v/v	volume per volume
Vp-p	Voltage peak to peak

Mass Spectrometry:

Mass spectrometry is an analytical technique surrounding the analysis of molecules based on their mass; however, to do so requires the species in question to become a charged ion, and as a result mass spectrometry uniformly measures and studies different species based upon their mass to charge ratio (m/z).

The study of chemical reactions based upon mass changes dates back hundreds of years, but the study of a molecule's mass using charged ions was first successfully achieved by Sir J. J. Thompson and Francis Aston in 1913 in the UK. Thompson and Aston had constructed an instrument to deflect ions towards a photosensitive plate, they found that although the same force would be applied to each ion as it passed though the force region – the ions were affected differently (due to their different masses) and so would spread out before hitting the photosensitive plate and leaving characteristic and reproducible marks/impact patterns on the plate depending on their mass. Using this state of the art "mass spectrograph" Thompson and Aston were able to show the various characteristic ion patterns for the components of air, and proved undeniably the presence of different isotopes for the same chemical species, a feat which later won Aston the Nobel Prize in Chemistry in 1922. Instantly it was obvious that Mass Spectrometry was adept at separating mixtures of compounds and providing individualised information on each chemical species within a sample.

100 years after Sir J. J. Thompson and Francis Aston constructed and successfully tested the first mass spectrometer, modern mass spectrometry has come a long way, but is still based upon these first fundamental ideas; to transform species of interest into manipulateable ions, and to separate them based upon their mass to charge ratio in order to provide a mass-based analysis of the compounds.

In addition, moving on from photosensitive plates, there are many ways of separating and detecting ions produced for mass spectrometry analysis including, but not limited to; Time of Flight (ToF), Quadrupole-MS, Orbitrap FTMS, Ion-Trap-MS , Magnetic sector-MS, FT-Time of Flight, and FT-ICR MS, including hybrid instruments combining components of each. The content of this thesis will be focused almost exclusively on the latter; Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry, with other methods referred to as needed.

1.1 Fourier Transform Ion Cyclotron Mass Spectrometry (FT-ICR MS):

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, first developed by Comisarow and Marshall in 1974,[1,2] is based around separating charged analyte ions using a phenomenon known as the Lorentz force. The Lorentz force is experienced by charged ions while inside a magnetic field (shown in equations 1a and 1b, and depicted in Figure 1.1, below).

$$F = ma = mv^2/r = m(\omega_c r)^2/r = q\, v\, B \qquad\qquad\qquad \text{eqn. 1a}$$

$$F_{Lorentz} = q\, v \times B \qquad\qquad\qquad \text{eqn. 1b}$$
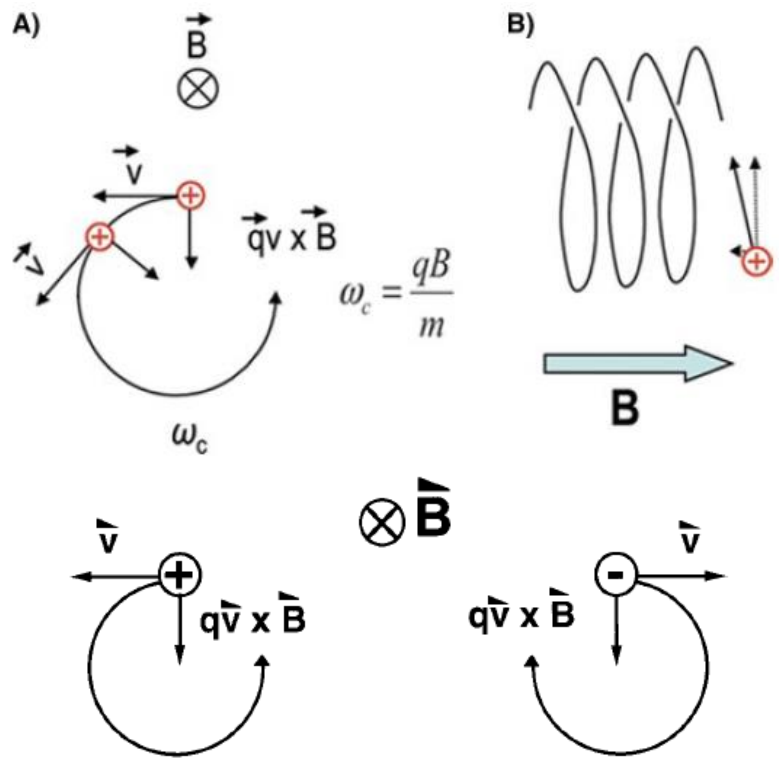
FIGURE 1. Ion cyclotron motion. The path of an ion moving in the plane of the paper is bent into a circle (see text) by the inward-directed Lorentz magnetic force produced by a magnetic field directed perpendicular to the plane of the paper. Note that positive and negative ions orbit in opposite senses. Reproduced with permission from (Marshall and Grosshans, 1991).

Figure 1.1: Top: depiction of the forces exerted on a charged species while within a magnetic field, where q is the charge of the ion, v is the perpendicular speed of the ion, and B is the magnetic field strength. Reproduced from Makarov *et. al.*[3] Bottom: a similar representation of the movement of both positive and negatively charged ions inside a magnetic field, reproduced from Marshall *et.al.*[4]

Ions orbiting inside the magnetic field will each have a unique frequency based upon their mass to charge ratio and the magnetic field strength, this frequency is dubbed the Cyclotron frequency (ω_c) and is simply given by equation 2 below;

$$\omega_c = \frac{qB}{2\pi m}$$

eqn. 2

Due to the cyclotron frequency ions can be separated based on their mass to charge ratio, have their frequency measured, and then in theory, a calibration function can be applied to convert measured frequency to accurate *m/z* values, thus providing mass spectral information on analytes.

For ions to remain inside a magnetic field and have their cyclotron frequencies measured they must be trapped. In FT-ICR MS the analyte ions are trapped in an ICR cell (Figure 1.2, below), an ICR cell consists of four major components; trapping electrodes (to keep ions confined axially), excitation electrodes, detection electrodes, and an amplifier to amplify the small magnitude signals up to a transmittable level.
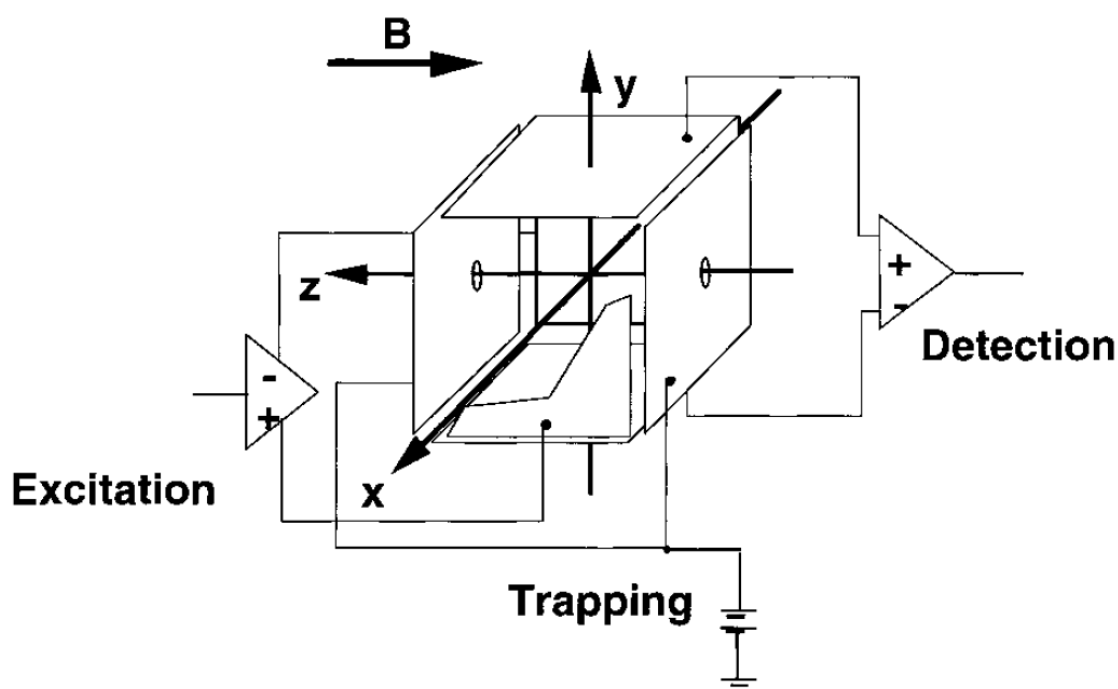
Figure 1.2: **Schematic representation of a cubic ICR cell, showing the major trapping, excitation, and detection electrodes, along with limited circuitry used during operation, reproduced from Marshall *et. al.*[5]**

As mentioned previously the magnetic field causes ions to rotate and will confine them radially within the bore of the magnet, the two trapping electrodes of the ICR cell can be provided

4

with a voltage of equal polarity to that of the ion and will thus keep the ions confined axially within the limits of the ICR cell – effectively trapping the ions in all dimensions and allowing experimentation.

In practice ions do not orbit within a magnetic field at their pure cyclotron frequency, instead they are perturbed by electric fields resulting from the trapping plates and from other ion's own electric field (the latter referred to as "Space charge").[6] The resulting cyclotron frequency is known as the reduced cyclotron frequency, given by equation 3, below;

$$\omega_{measured} = \omega_c\text{-} \omega_m \qquad\qquad \text{eqn. 3}$$

The combination of the electric and magnetic forces on the ions produces a second motion; the so-called magnetron motion of ions. The magnetron motion acts along the same axis as the cyclotron motion, but at a lower frequency (<100Hz), the equation for the magnetron motion is shown below:

$$f_m = \frac{\alpha V}{\pi a_2 B} \qquad\qquad \text{eqn. 4}$$

Where f_m is the magnetron frequency/Hz, α is a geometry factor, V is the trapping potential used, a is the distance between the trapping plates, and B is magnetic field strength/Tesla.[7]
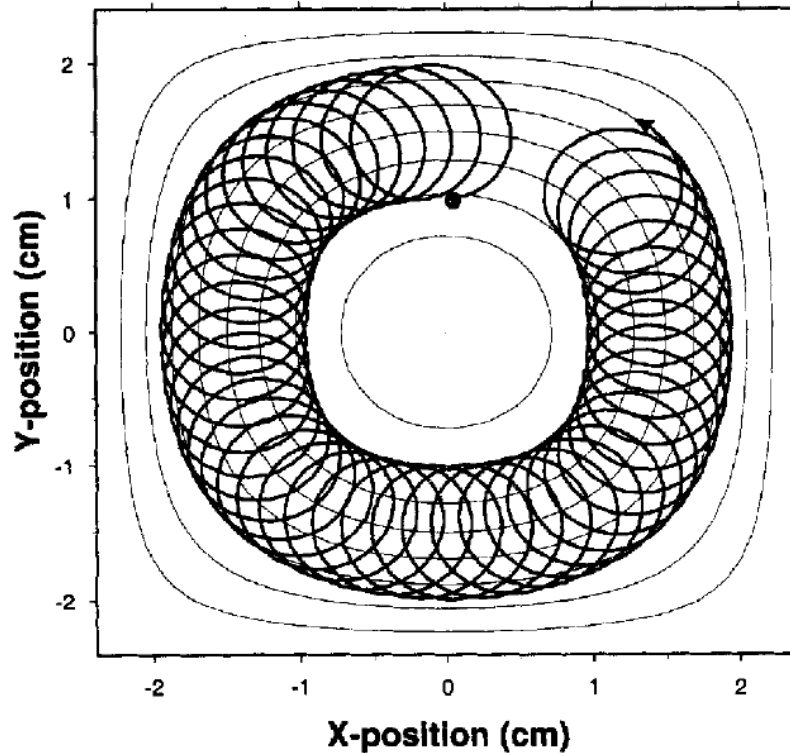
Figure 1.3: depiction of magnetron (low frequency, large (exaggerated) radius) and cyclotron (higher frequency, smaller radius) motion of an ion while trapped within a magnetic field. Reproduced from Amster _et. al._[7]

Magnetron motion arises because the electric field generated by the trapping electrodes extends over the entire cell, see Figure 1.4 below. Since the electric field at the centre of the ICR cell is not zero; ions are repelled from the centre of the cell towards the lowest energy point, while still being confined by the magnetic field. The repulsive force away from the centre of the cell acts directly against the homogeneous magnetic field confining the ions radially, as a result the magnetron motion causes expansion of the ion orbit around the z-axis.
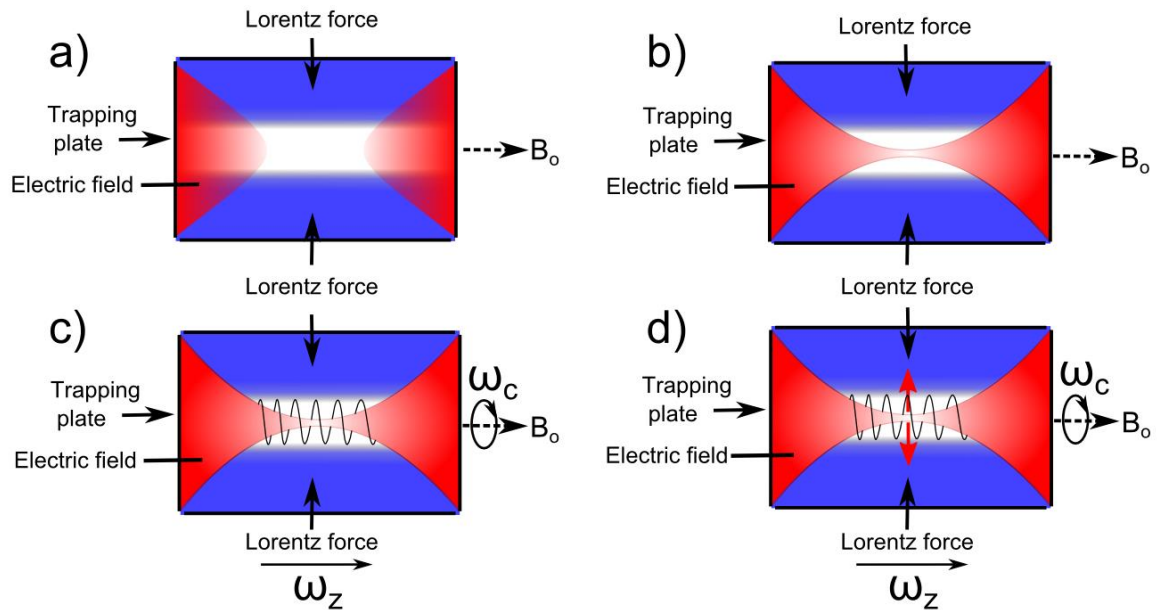
Figure 1.4: Representation of the electric fields(red) and magnetic field induced force (Lorentz force) (blue) present within an ICR cell a) the ideal fields/force induced by trapping plate voltage and magnetic field b) the resulting fields/force produced by electric field being quadratic (to a first order approximation) and non-zero at the centre of the cell[7] c) The fields within an ICR cell, and the depiction of cyclotron and trapping motion within the cell for a given ion (cyclotron radius is exaggerated for clarity) d) inclusion of the non-zero electric field exerting an outward force (red arrow) on the travelling ion, inducing magnetron motion.

The magnetron orbit of ions continues to expand with time as the ions are held within the ICR cell, at a relatively slow rate, meaning if ions were held within the ICR cell and were otherwise stable (from collisions etc.); they would slowly expand their orbit and still eventually be lost from the trap.[7] Currently the only effective way of returning ions to the centre of an ICR cell, despite magnetron motion, is the use of quadrupolar axialisation,[8,9] which has been shown to be an effective way of not only returning ions to the centre of an ICR cell, but also in converting magnetron motion to true, unperturbed cyclotron motion, and allowing re-measurement of stored ion packets,[10,9] in theory this enables an "infinite sensitivity" measurement, where the same ions are repeatedly excited, detected, then returned to the ICR cell; allowing re-measurement many, many times,

possibly down to single-ion detection using FT-ICR MS, but also allowing many summed scans of the same species, without introduction of new ion packets and subsequent usage of sample.
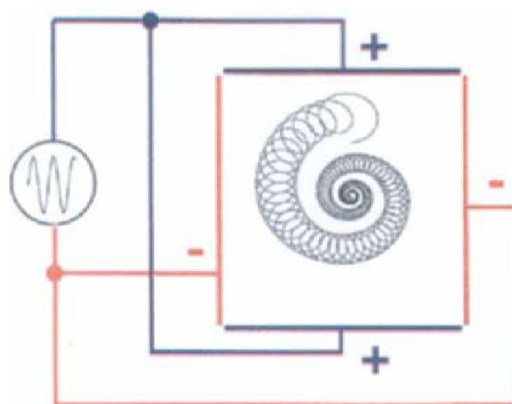
Figure 1.5: Quadrupolar axialisation - ions at an excited cyclotron orbit and/or large magnetron orbit can be returned to the centre of the ICR cell by a) injecting a pulse of collision gas (e.g. Argon) and b) causing collisions during excitation, c) following a pumping delay ions can be re-excited to be detected as normal, see below.[9] Reproduced from Amster *et.al.*[7]

Unfortunately magnetron motion has no tangible analytical use,[7] but is amplified if ions deviate from the magnetic field axis (z-axis), meaning good alignment of the ICR cell with the centre of the homogenous magnetic field region is essential to minimise magnetron motion issues. Magnetron motion effects are uniform for all ions, irrespective of mass to charge ratio, meaning using ions of known cyclotron frequencies, quantities, and trapping potentials it is possible to calculate the magnetron frequency shift in a given spectrum using equation 4, above.

Finally though the ions are confined within the electric field well created by the trapping plate electrode voltages, the ions will still effectively oscillate between the two plates creating a z axis motion of ions with a frequency ~10Hz, depending on cell geometry, plate voltages etc.[5] This frequency is based upon equation 5, below;

$$\omega_z = \sqrt{\frac{2qVta}{ma^2}}$$
eqn. 5

The combination of the cyclotron frequency, magnetron motion, and trapping motion results in a much more complicated ion motion path than the pure cyclotron frequency, shown in Figure 1.6, below:
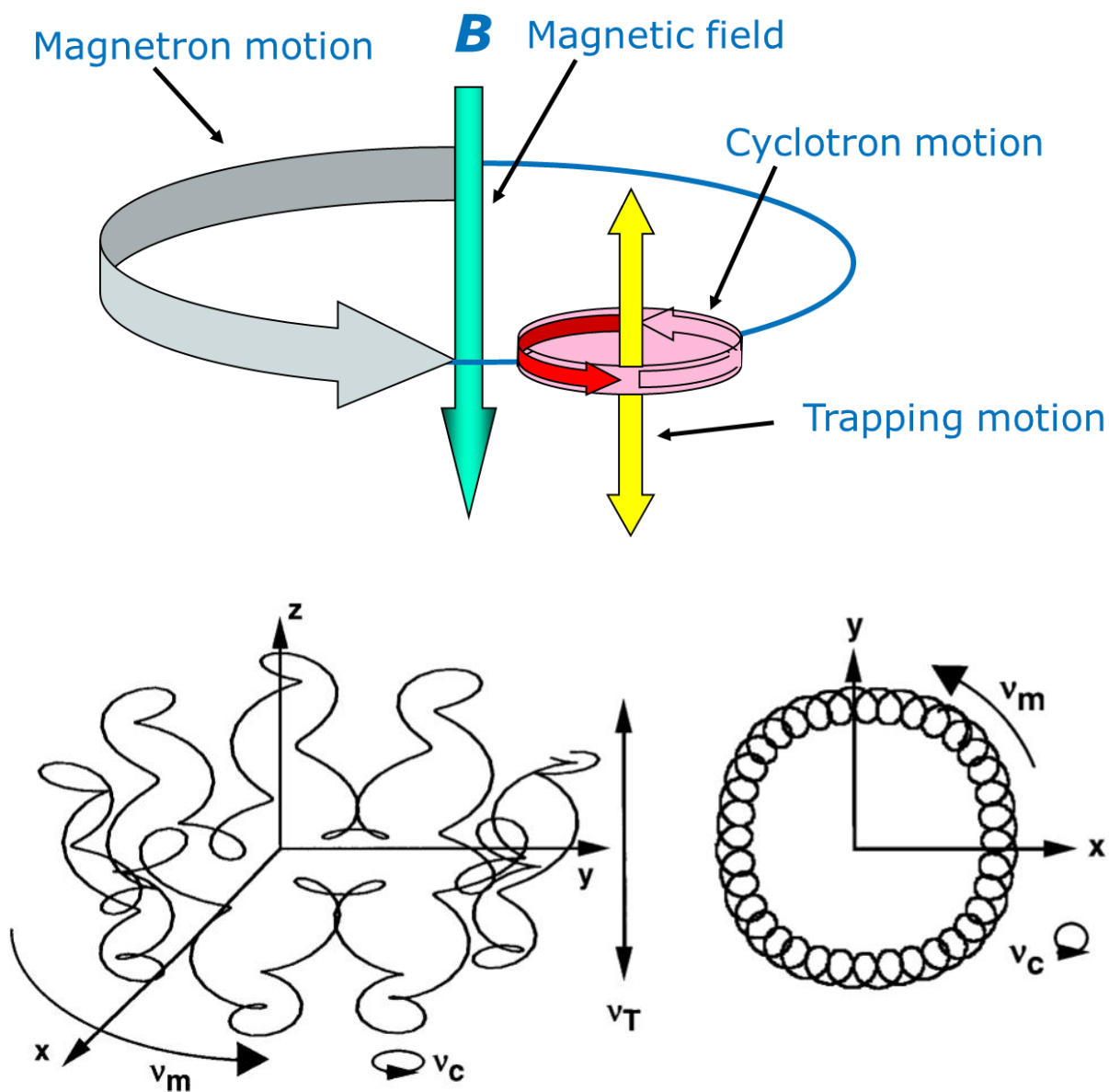
Figure 1.6 – Top: depiction of the 3 main motions of an ion trapped within an ICR held inside a magnetic field; cyclotron motion, trapping motion, and magnetron motion. Bottom: a representation of the full ion path, courtesy of Bruker Daltonics, Bremen, Germany and reproduced from Marshall *et. al.*[5]

It is worth noting that the cyclotron frequency of any ion within the magnetic field is independent of its kinetic energy, which is often very hard to control and reduce in variance, this is a key advantage of FT-ICR MS measurements over other MS platforms which often suffer from ill-defined/large ranges of ion kinetic energy, such as ToF's.[5] However the motion of ions discussed above assumes an ideal vacuum, which is unfortunately unobtainable. Collisions between ions and background gas molecules can perturb ion motion in any of the 3 principle motions discussed and can cause undesirable effects, most notably the damping of cyclotron motion, the eventual target of measurement (see below). The damping of the cyclotron motion of an ion is shown below in Figure 7, excessive collisions with background gas can quickly lead to loss of ions from the ICR cell, preventing measurement. It is for these reasons that the use of ultra-high vacuum conditions is consistently used for FT-ICR MS, despite the technical drawbacks of doing so, it is required to allow effective ion motion of trapped ions.
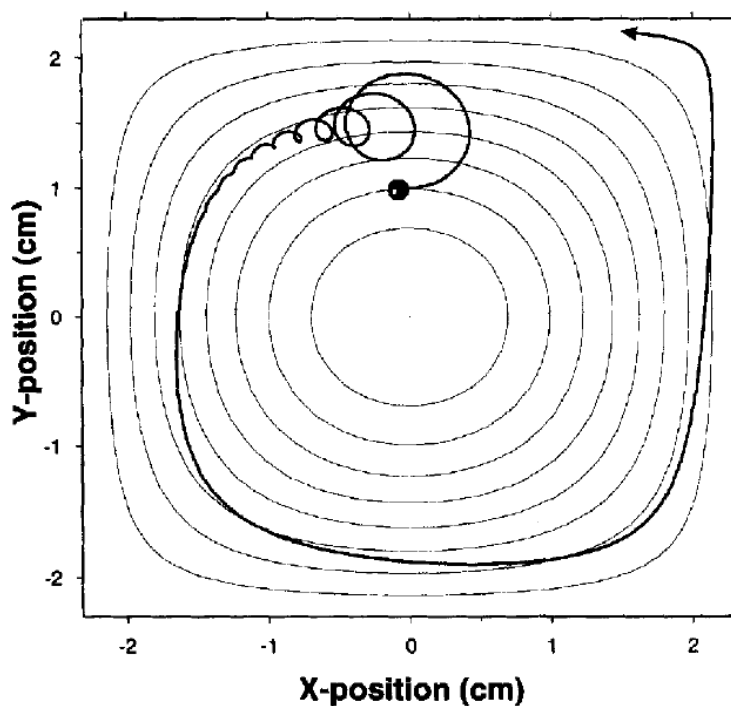
Figure 1.7: representation of ion motion inside of an ICR cell in the presence of background gas. The cyclotron motion of the ion is quickly damped though magnetron motion continues to slowly

increase, and the ion is soon lost from the trap. Pressure shown is 0.01 Torr. Reproduced from

Amster *et. al.*[7]

1.2 Excitation/detection of ions:

The Cyclotron motion of ions will occur as soon as the ions enter the magnetic field; however the radius of this initial motion is within the range of several millimetres and is incoherent (ions of the same frequency are not travelling together), which hinders effective detection of the ions inside the FT-ICR MS. As a result the ICR cell is fitted with 2 excitation electrodes, to which an RF pulse can be applied in order to excite ions. The RF pulse transfers energy to the ions trapped within the cell; exciting them to a higher radius within the ICR cell. As mentioned previously; the cyclotron frequency is independent of kinetic energy and so although the added energy will make the ions move faster, they will still move with the same frequency around the ICR cell, hence the excited radius to compensate. Figure 1.8 summarises the excitation and subsequent deletion of the ions within and ICR cell:
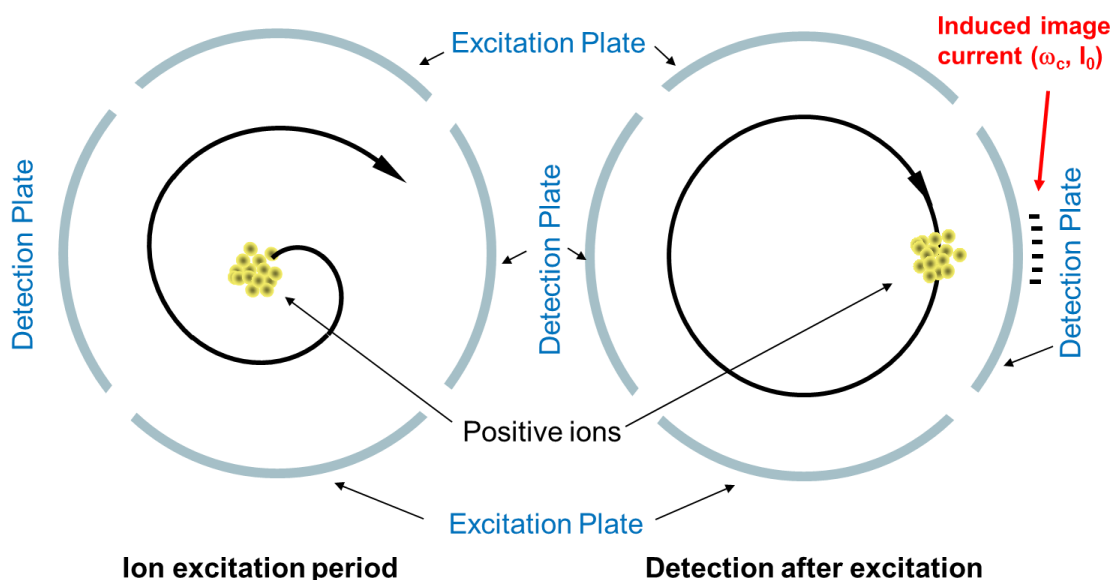
Figure 1.8 – a) excitation and b) subsequent detection of a species within an ICR cell. Courtesy of Bruker Daltonics, Bremen, Germany.

After the ions have been excited out to a detectable radius, the RF pulse is ceased, the ions continue to orbit around the ICR cell at the excited radius, as shown above. While orbiting in close proximity to the ICR cell plates the ions (for example positively charge ions) will affect a charge onto one detection electrode, the so-called "image current", however a moment later ($\omega_t=\pi$) they will have rotated around the cell to the other detection electrode and affect a charge there instead, if these two electrodes are linked electronically then there will be an observable net movement of charge from one electrode to the other and back again as the charged ions rotates around its cyclotron orbit. This movement will have the same frequency as the cyclotron frequency of the rotating within the cell and can thus be used as a measure of the ions frequency. This image current is very small (<10mV) and so the signal is amplified as soon and closely as possible to the ammeter, before being recorded on a data station, shown in Figure 1.8 above, producing a transient signal otherwise known as a free induction decay signal (FID).

Ion Detection
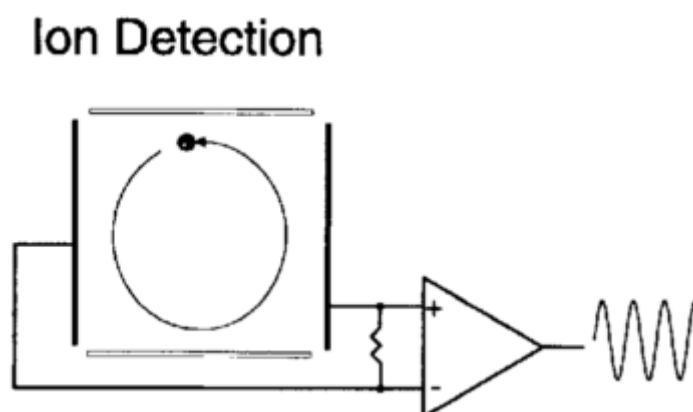
Figure 1.9: Representation of an ion cloud orbiting an ICR cell post-excitation and the current being measured by an ammeter after amplification to produce a transient (FID). Reproduced from Amster *et.al.*[7]

Though excitation of ions to a higher cyclotron orbit has the benefit of producing nice, coherent ion packets, which are readily detected, the ion motion within the cell is not completely ideal. Collisions of travelling ions inside an ICR cell with background gas can quickly damp the ion motion, pushing the ions out of their cyclotron orbit and back to the centre of the ICR cell, causing

signal damping, and can cause loss of ion coherence. To avoid this process ultra-high vacuum conditions are essential inside the ICR cell, greatly reducing the frequency of collisions, and enabling transient times on the order of several seconds routinely. However despite studies showing the ability to confine ions within an ICR cell for long periods of time (hours even to days) ions exited to a larger than natural orbit will damp and return to the centre of the cell/be knocked out of the stable confines of the trap usually within several seconds to several minutes.[11] This signal damping can continue to cause compete loss of ion signal within the ICR cell, despite the FID still being recorded, such transient "crashing" can adversely affect the resulting frequency and mass spectra badly,[12] and so fine tuning of all components within an FT-ICR MS is required to ensure long transient times and achieve the ultra-high performance FT-ICR MS is used and famous for.[11,13]

1.3 Data processing:

Though the reduced cyclotron frequency of ions can be effectively measured using the process above, it yields a transient signal, not a list of cyclotron frequencies which can be used to calculate an ion's mass to charge ratio. To obtain the consistent ion frequencies from the recorded signal a Fourier transform is required to transform the data from a pure FID to a time domain spectrum. The process is detailed below and can be found in a number of previous works,[5,7,3,14] briefly (Figure 1.10 below); the pure FID signal from the instrument (a) is fast Fourier transformed (if plotted would appear as (b) (e) and (f) for plot real, plot, and plot absolute functions respectively), the Fourier transformed real FID (c) is then zerofilled (below), cut in half to remove the redundant mirrored section of data (seen in Figure 1.10 (b)), the frequency domain spectrum (Figure 1.10 (c)) is then calculated using the acquisition rate (for data points). Finally the frequency domain spectrum is calibrated using a given calibration equation (see below) to convert the frequency axis to corresponding *m/z* and produce a mass spectrum (Figure 1.10 (d)). It is often reported that the measured FID signal is simply Fast Fourier Transformed and the plotted to retrieve the frequency spectrum, which can then be calibrated to produce the final mass spectrum.[7,3] Unfortunately this is

an oversimplification and simply plotting after the FFT would produce a combination of real and imaginary points (Figure 1.10, plot (e)), and even plotting these values as absolute numbers instead of the combination, would produce an x-axis distorted spectrum (Figure 1.10 plot (f)).
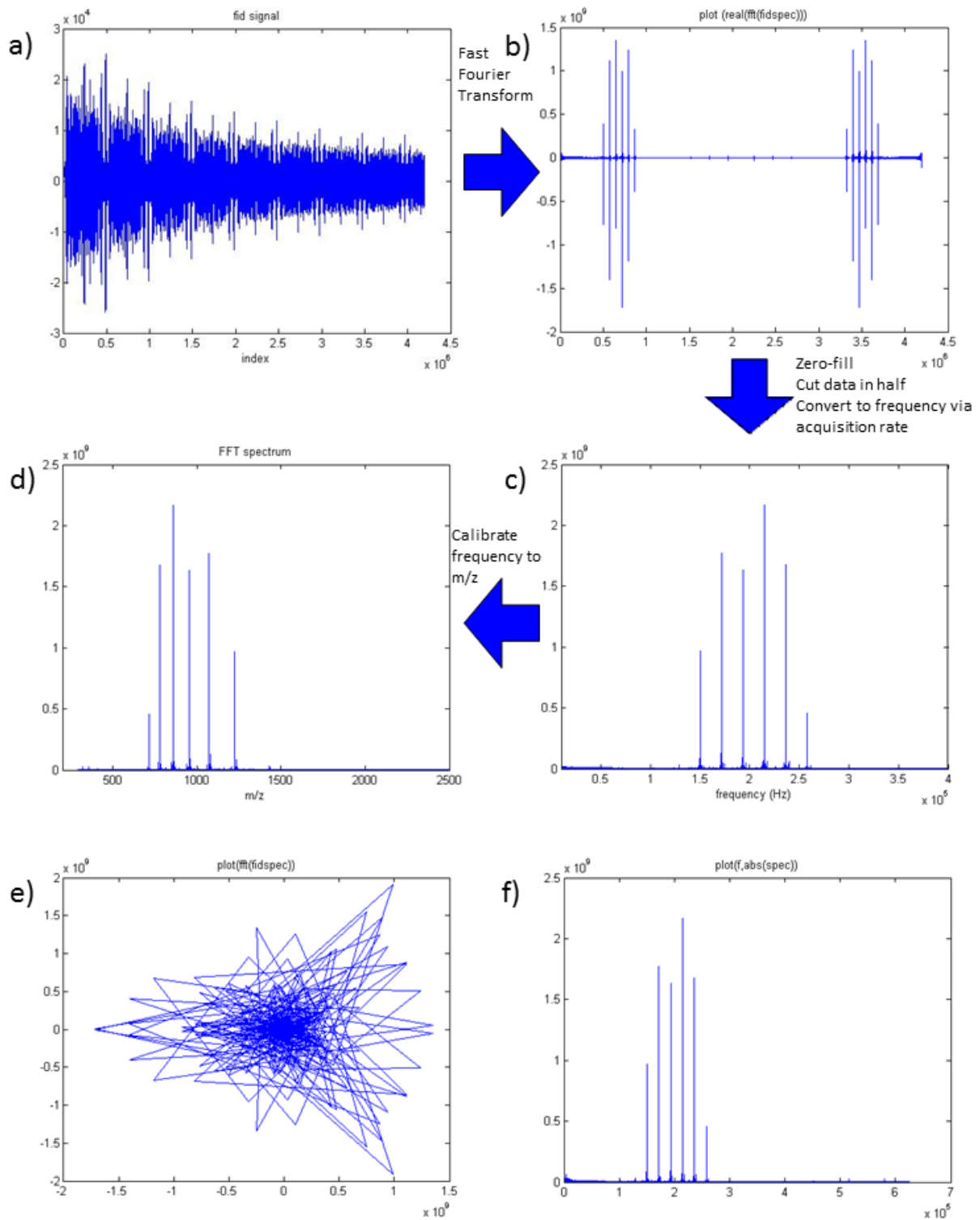
Figure 1.10 – Steps involved in the data processing of a measured FID signal (a) which can then be Fourier transformed to produce a mirrored time domain spectrum (b) followed by halving, and re-plotting as real numbers to produce a frequency spectrum (c), finally the frequency spectrum can be calibrated using a variety of equations (see below), to produce an FT-ICR MS mass spectrum (d). Should the data from the FID simply be plotted after Fourier transformation as so often recorded it would produce a complex mixture of real and imaginary numbers (e) and if it were plotted as simply absolute numbers it would produce a plot with an uncalibrated x-axis, though containing the frequency information, with the wrong separation to be informative.

1.3.2 Performance metrics – Resolving power and mass accuracy:

Though mass spectrometers are advanced analytical instruments and contain many individual components and techniques (discussed later), the two most widely used metrics to judge analytical performance achieved by an instrument/within a spectrum are the resolving power and mass accuracy of the measurement.

The resolving power of a mass spectrometer is a measure of how effectively the instrument can resolve (separate) mass spectral peaks; the higher the number, the more effective the spectrometer is at separating closely spaced peaks. Resolving power is defined by equation 6 (below) where m is the mass to charge ratio of the peak in question and Δm is the width of the peak at half its own height; Full Width Half Maximum (FWHM).

Resolving power = $\dfrac{m}{\Delta m}$ 						eqn. 6

Resolving powers of around 300-2000 are routinely achieved on quadrupole mass analysers and Ion Trap mass spectrometers, which translates into the ability to separate analytes/isotopes of ~1m/z difference (often referred to as "unit resolution"), up to ~0.25-0.3m/z for Ion trap performance. Whereas high resolution instruments are usually quoted as achieving resolving power performance

of >20,000, and Ultra-high resolving power instruments/methods (such as FT-ICR MS) are consistently above 100,000; a 12T FT-ICR MS achieves >500,000 trivially and has been shown to achieve up to ~9 million resolving power with careful tuning and developments in FT-ICR MS data processing techniques.[15,11] A peak can be considered fully/baseline resolved if there is clear space (baseline noise) between two individual peaks at their base. Currently the highest reported resolving power recorded using an FT-ICR mass spectrometer was ~47 million, achieved by Nikolaeve *et. al.*[16]

The resolving power performance of a mass spectrometer tends to change with the *m/z* range being studied, usually with a slow descent towards higher *m/z* values. Specifically Fourier Transform based MS platforms (such as the FT-ICR MS and more recent Orbitrap MS[17] platforms) show a decrease in resolving power with increasing m/z values. FT-ICR MS shows a linear decrease in resolving power of 1/x, so the resolving power at 800m/z is one half of that at 400m/z.

The FT-ICR MS resolving power increases linearly with the strength of the magnetic field used to contain ions within the ICR cell and induce cyclotron motion.[5,18] The uses of higher magnetic fields thus provides not only a way of more trivially achieving higher resolving power performance on a given FT-ICR MS, but also allows a marked increase in the achievable resolving power per second of the mass spectrometer,[18] which can be critical in time-sensitive measurements, such as chromatography-coupled mass spectrometry (discussed in section 1.9).

Ultra-high resolving power has been shown to be a critical requirement in many contemporary MS experiments, especially for the effective analysis of complex mixtures[19] (containing 10,000-100,000 analyte peaks per spectrum) and for effective, resolved tandem mass spectra of complex molecules, such as biomolecules (discussed in section 1.7.2). When analysing complex hydrocarbon-based mixtures, such as petroleum/oil based samples, analysis frequently produces mass spectra containing 10,000+ analyte peaks, within a relatively small mass range (e.g. 150-800 m/z) creating an extremely dense mass spectrum, with peaks often separated by only a few milli-Daltons (mDa). Ultra-high resolution FT-ICR MS is currently the only viable technique to

effectively analyse these mixtures in a single mass spectrum, even separating ^{13}C isotopes from ^{15}N and Sulphur isotopic components from the same/different analytes; the so-called "Fine structure" of compounds. Fine structure analysis arose from the small mass deviations in the M+1 isotope of analytes depending on the (hetero)atoms present in the specific analyte (shown in Figure 11, below). Whereas the first ("M") isotopologue of any analyte would appear as a single peak, the M+1 isotopologue would "split" at ultra-high resolving powers to separate the ^{13}C isotopologue from the ^{15}N isotopologue (assuming the analyte contained both atoms), calculating the ratio of these peaks and taking into account the natural abundance of the individual M+1 isotopes relative to the M isotope allowed the elemental composition of analytes to be calculated purely based on accurate mass measurements of analytes, and has since enabled the field of complex mixture analysis by high-performance mass spectrometry.
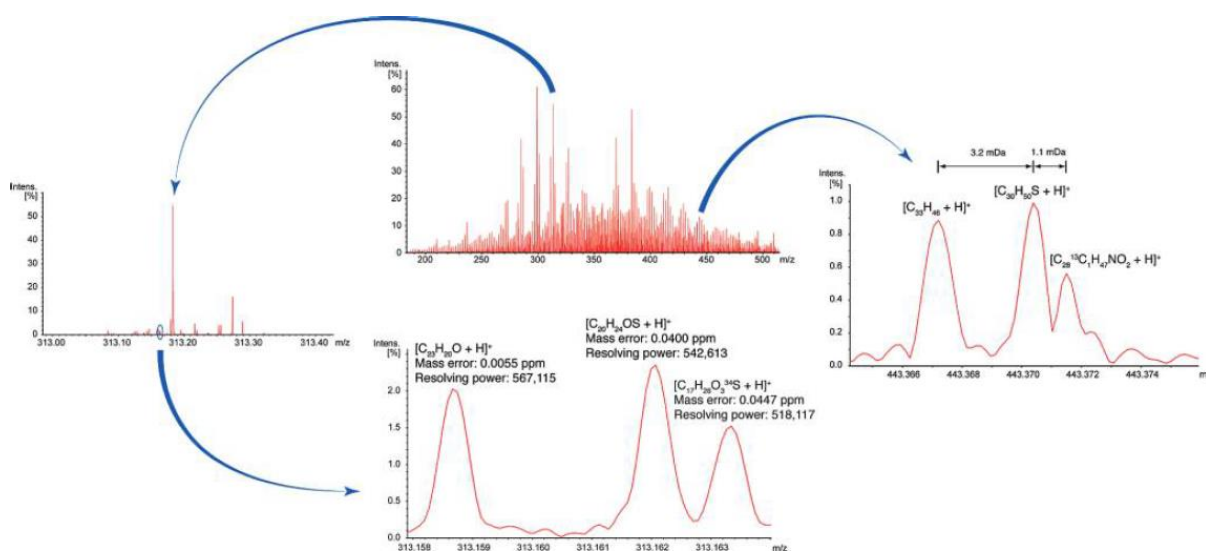
Figure 1.11: Example MS spectrum showing the complexity of petroleum based samples via FT-ICR mass spectrometry, included are insets showing effective resolution of different components isotopic fine structure and the corresponding milli-Dalton range separation needed. Reproduced from Barrow *et.al.*[19]

Particularly large molecule analysis also requires high performance MS and ultra-high resolving power values. As particularly large analytes gain multiple charges during ionisation (discussed in section 1.6.1) the *m/z* spacing between isotopes is no longer ~1*m/z*, instead it will be ~1/z, if the analyte acquires a significant number of charge carriers (e.g. z=50-100) then the spacing between individual isotopologues observed is extremely small (0.02-0.01 m/z respectively for 50 and 100 charges). As a result the effective analysis of both particularly large analytes and particularly small/complex analytes both require ultra-high resolving power performance, such as FT-ICR MS. UHR-MS is also particularly suited when studying the extremely complex spectra produced from selected fragmentation of ions during tandem mass spectrometry experiments (discussed in section 1.7).
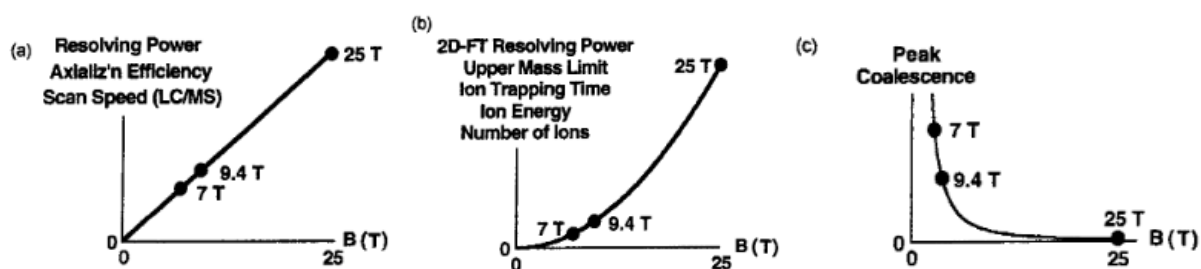
Figure 1.12: plots representing the various dependencies of FT-ICR MS performance on magnetic field strength, and thus the advantages of increasing magnetic field for FT-ICR Ms analysis. Reproduced from Marshall *et.al.*[18]

1.3.3 Mass Accuracy:

The second key parameter for mass spectrometer performance is the inherent accuracy of the mass measurement being taken, termed "mass accuracy", this refers to the deviation of the measured m/z value of a certain analyte from the theoretically calculated m/z value for that species then expressed in parts-per-million, as shown in equation 7, below:

$$Mass\ Accuracy\ (ppm) = \frac{(Measured\frac{m}{z} - Exact\frac{m}{z})}{Exact\frac{m}{z}}\ x\ 1,000,000 \qquad \text{eqn. 7}$$

Unlike many other key analytical techniques such as Nuclear Magnetic Resonance (NMR), UV/visible spectroscopy, gas/liquid chromatography, and Infra-red spectroscopy, mass spectrometry analysis enables the calculation of the exact mass of an analyte, and thus is the only common analytical chemistry technique to provide an exact error calculation for each measurement and provide numerically accurate, instead of precise measurements.

It is this fundamental capability which makes mass spectrometry able to accurately assign unknown compounds from complex mixtures,[13,19] or to pattern match spectra/masses against databases effectively for fast analysis of target compounds (e.g. in explosives detection[20] or forensic applications[21]).

The mass accuracy of mass spectrometers varies depending on the fundamental principles upon which they base their mass measurements on. For example a Quadrupole mass analyser bases mass measurement on a mass filtering principle (discussed in detail in section 1.6.2.2), allowing ions of m/z 50 to pass through and then detecting current, then m/z 51 and so forth, to produce a spectrum of ion populations/current detect at each unit m/z values, and so providing unit-resolution. The resolving power of a quadrupole mass analyser is therefore dictated by how accurately it can filter ions; if this could be done accurately to 0.5m/z then the resolving power of the quadrupole would need to be doubled, but could also then increase the mass accuracy.

The mass accuracy of FT-ICR mass spectrometers is dictated by accurate measurement of the cyclotron frequency (ω_c), which as previously discussed is negatively affected by magnetron motion (ω_m), to become the measured reduced cyclotron frequency (ω). Thus if one could decrease magnetron motion to zero, and eliminate electric field perturbations on ions of interest then perfect accuracy measurements of ions would be possible.[5] Unfortunately in practice this is not possible, as electric fields are needed to contain ions, they are produced from plates of finite dimensions, and magnetron motion is not possible to be eliminated. However the mass accuracy of FT-ICR MS is directly proportional to the square of the magnetic field (B^2) and so use of higher magnetic fields

enables much more accurate mass analysis of target compounds (Figure 1.12 above). Coupled together with careful calibrations (Section 1.3.6), FT-ICR MS is currently the most accurate mass spectrometry technique in the world, reaching part-per-billion mass accuracy on compounds in some applications.[22,19]

1.3.4 Zero-filling:

One of the key performance parameters for MS is the resolving power of the resulting mass spectrum, for FTMS this is directly proportional to the length of the transient/acquisition (T_D). Thus the opportunity to increase T_D has been consistently beneficial to FTMS (unless acquiring under time-constrained situations). Unfortunately due to collisions, inhomogeneities within ICR cells, and ineffective ion trapping, it is not always possible to obtain a non-damped/slowly damping transient signal for extreme T_D lengths.[5,7] However it is possible for all acquisitions to obtain a transient of length T_D and then mathematically extend this transient to double its previous size by including an equal number of zero's to the end of the recorded FID signal. The technique, known as Zero-filling, doubles the effective length of the transient signal and thus doubles the sampling resolution of the FT, while still recording the same T_D within the FT-ICR MS instrument. Zero-filling, thus, increases the number of points on every peak measured, but does not change the peak shape or peak resolution, but improves the accuracy of peak-centre (centroid) calculation via better fitting. Zero-filling FTMS spectra is very effective for all acquisitions, both individual long-T_D ultra-high resolution scans, and time-constrained repeating acquisitions. Zero-filling of FTMS data can be applied more than once, thus increasing the number of zero's added to 3 fold, 5 fold etc., but despite the obvious performance increase (with little/no additional effort) is usually only performed once and a maximum of twice in order to avoid reported mathematically added artefacts in the resulting mass spectrum.[7]

1.3.5 Apodisation:

Although the pursuit of higher and higher resolving powers has been a distinct driving force in FT-ICR MS over the years, not all signal processing directly benefits/increases resolving power, but instead provides other benefits crucial to not only obtaining high-quality data but also aiding the interpretation of the resulting data.

FT-ICR MS produces extremely high resolving power mass spectra, meaning very narrow peaks, however the base of FTMS peaks broaden greatly with increasing signal-to-noise ratios. Termed "spectral leakage", the FTMS peaks broaden at the base as a by-product of the Fourier Transform, and are not considered "real" peaks. Though improved ICR cell design, tuning etc., has been achieved to help reduce these unwanted effects, high signal-to-noise peaks will always continue to disturb areas of the mass spectrum, and unfortunately can completely over-shadow lower signal-to-noise peaks (Figure 1.14, below).

To avoid these negative effects a signal processing technique known as apodisation has been very effectively applied to FTMS spectra to reduce the effect of the signal at extremes of the FID. Apodisation scales the contribution of each area of a given transient using a mathematical function, thus some areas of the transient will then contribute "more" to the resulting spectrum, and some will contribute "less". Examples of apodisation are shown below both for theoretical data (Figure 1.13) and for actual FT-ICR MS data (Figure 1.14), both are remarkably similar to the unapodised spectrum, but suffer from much less "spectral leakage", i.e. broadening at peak bases and enables baseline resolution and observation of low intensity isotopes (Figure 1.14).
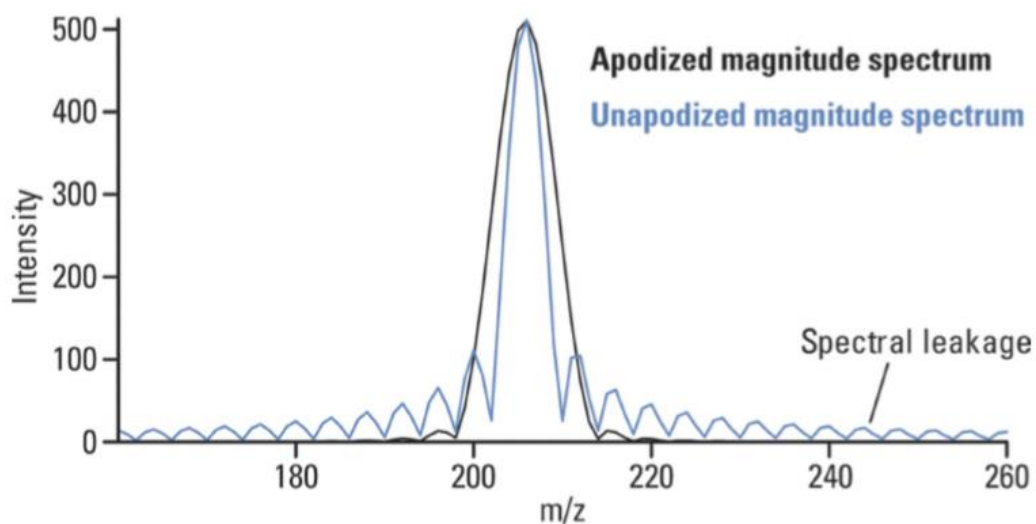
Figure 1.13: example of apodisation on an FTMS spectrum showing that apodisation can suppress side lobes and broad peak bases, which can improve dynamic range, but negatively affects resolving power. Reproduced from Makarov _et. al._[3]
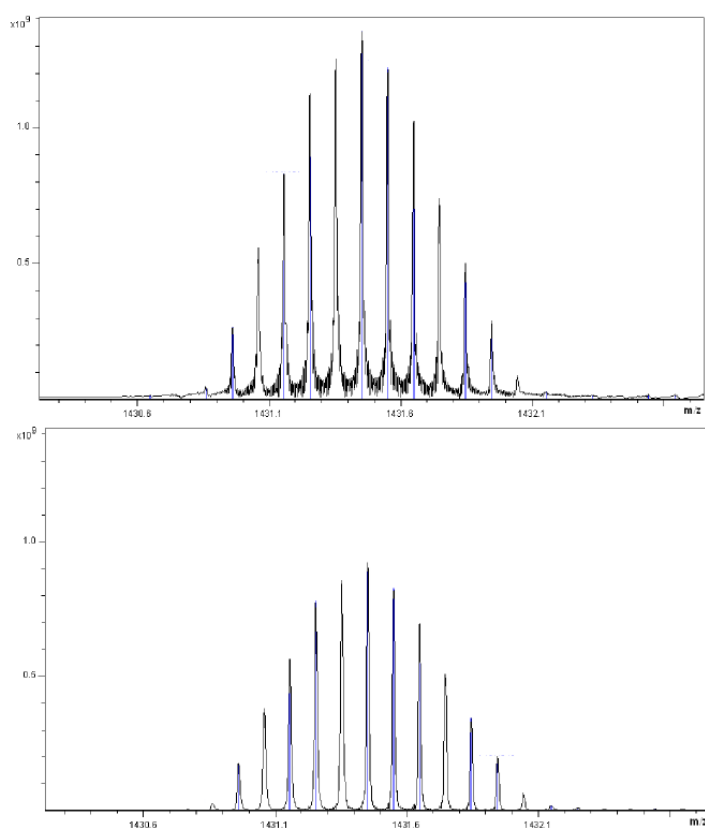
Figure 1.14 – FT-ICR MS spectrum of the 10+ charge state of lysozyme protein: un-apodised (top) and processed using sine-bell apodisation (bottom), although the resolving power drops from

165,000 to 129,000 for the peaks observed the apodisation does improve the peak bases and allow visualisation of the low signal-to-noise isotopes. Also note the decrease in intensity due to peak broadening.

Apodisation can be achieved using a variety of different mathematical functions to suit the mass spectrometrist's needs/priorities, each having a different effect on the overall spectrum and resolving power performance.

The major disadvantage to apodisation is that it decreases the resolving power of the resulting mass spectrum peaks, usually by around 20-25%, which is a significant drop in analytical performance. However due to the extreme amount of spectral leakage caused by high signal-to-noise peaks, and the distinct loss of low signal-to-noise peaks in these areas, which are often critical for accurate MS assignments, apodisation has become a standardised signal processing technique and is commonly applied to almost all contemporary FTMS data.[5] Apodisation also improves MS peak shapes, which directly improves mass accuracy via better fitting, centroids, and thus assignments.

1.3.6 Calibration:

As mentioned previously; FT-ICR MS detects and records frequencies, which can then be extracted via (amongst other operations) the Fourier transform. This is fortunate as not only are frequencies one of the easier, more reliable features to detect in analytical chemistry,[7] but it also allows FT-ICR MS to take advantage of electronics and advances centred around more common frequency-based techniques, such as nuclear magnetic resonance spectroscopy (NMR).

However, the detection of frequencies and representation of a frequency spectrum (Figure 10c, above), is not a mass spectrum. Although in theory all cyclotron frequencies detected can be converted trivially to m/z ratios using the cyclotron frequency equation (equation 2, above), we do not detect the cyclotron frequencies of ions, only the reduced cyclotron frequencies can be detected

(Equation 3, above). As a result all frequency domain spectra obtained from FT-ICR MS instruments need to be calibrated to convert them into mass spectra.

Calibration can most simply be achieved externally – i.e. running a sample of known components, spanning the analytical mass range, obtaining a frequency domain spectrum and then plotting the observed frequencies against the known masses for those compounds (Figure 1.15, below). Applying a polynomial fit to the resulting graph will provide the external calibration function needed to calibrate future, unknown, frequency domain spectra into mass spectra on-the-fly.
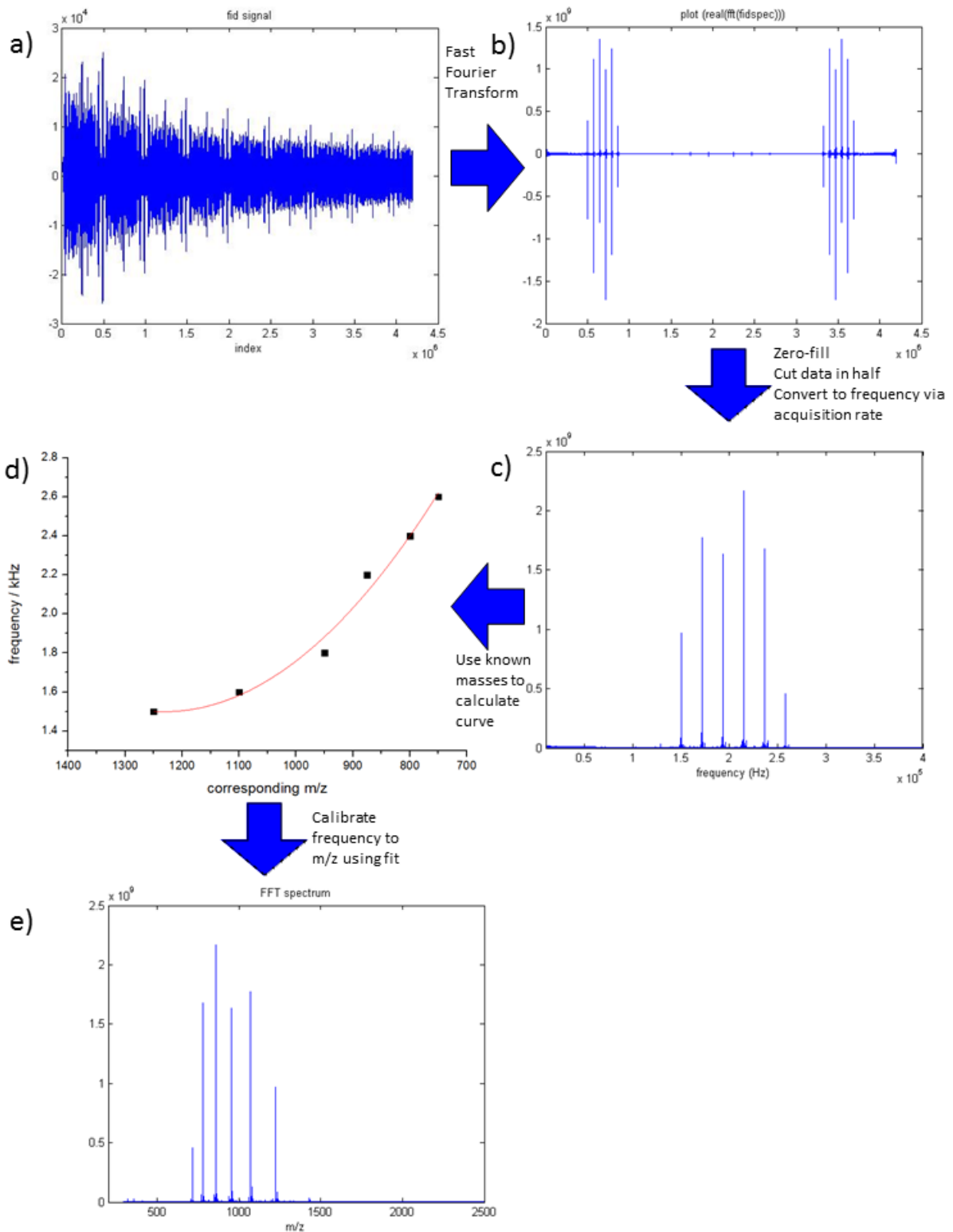
Figure 1.15: representation of external calibration procedure to produce a mass spectrum (e) from a recorded FID signal (a) by creation of a calibration curve (d) of known peaks in the frequency spectrum (c). Example shown is a ubiquitin protein FT-ICR mass spectrum.

Though this external calibration procedure seems adequate for FT-ICR MS, the mass accuracy of the resulting spectrum will suffer due to differences in magnetron motion, electric fields, and space charge differences between the calibrant spectrum and the sample spectrum.[3,7,23]

In order to avoid such deviation affecting analysis, many different calibration functions have been derived and expanded over the years, some of which are presented below in Table 1.1, from Gross *et. al.*[23]

Table 1.1 – Examples of FTMS calibration equations reproduced from Gross *et. al.*[23]

$f = \dfrac{a}{m}$	basic law of ions in a B field
$f^2 = \dfrac{a}{m^2} + \dfrac{b}{m}$	(Beauchamp-Armstrong et al., 1969)
$f^2 = \dfrac{a}{m^2} + \dfrac{b}{m} + c$	(Ledford et al., 1980)
$f_{sideband} = \dfrac{a}{m}$	(Allemann et al., 1981)
$f = \dfrac{a}{m} + c$	(Francl et al., 1983)
$\left(\dfrac{M}{Z}\right) = \dfrac{a}{f_{obsd}} + \dfrac{b}{f_{obsd}^2}$	(Ledford et al., 1984b)
$f_{estimated} = f_{measured} + c(I_{calibrant} - I_{analyte})$ $\dfrac{m}{z} = \dfrac{A}{f_{estimated}} + \dfrac{B}{f_{estimated}^2} + \dfrac{C}{f_{estimated}^3}$	(Easterling et al., 1999)
$M = \left(\dfrac{kB}{f_n + \Delta f}\right)n - n(M_c)$	(Bruce et al., 2000)
$\left(\dfrac{M}{Z}\right)_i = \dfrac{a}{f_{obsd}} + \dfrac{b}{f_{obsd}^2} + \dfrac{CI_i}{f_{obsd}^2}$	(Masselon et al., 2002)
$\dfrac{m}{z} = \dfrac{A}{v} + \dfrac{B}{v^2} + \dfrac{C}{v^3} + \dfrac{BC}{Av^4}$	(Wang et al., 1988)

The external calibration procedure described above used the Francl *et. al.* equation f=(a/(m+c) however many other, and more accurate calibration equations can be used for both external and internal calibration (described below).

Internal calibration:

As mentioned above external calibration can only go so far in achieving high mass accuracy mass-spectra, typically externally calibrated FTMS spectra are accurate to within 1-2ppm, but can be as high as 15ppm if the same electric trapping potentials and ICR cell parameters are not used for both the calibrant and sample acquisitions. To further improve mass accuracy of FTMS, the spectra acquired can be re-calibrated "internally" using known species within a sample mass spectrum, the known species may be ionised at the same time as the analyte (co-ionisation/spiked sample), leaked into the MS continuously, or may be species assignments identified from the external calibration and then used to further calibrate. Internal calibration can again use any of the calibration equations above, some instrument manufacturers use proprietary calibration equations developed and corrected in-house to improve calibration for their instruments when using their corresponding software packages. Careful internal calibration of FT-ICR MS data has been shown to allow sub-ppm and even ppb mass accuracies,[22,19] this is especially crucial when assigning unknown/complex mixture spectra purely on accurate mass to determine elemental compositions, which has been shown to be effective at assigning tens of thousands of compounds in a single, direct-infusion mass spectrum.[19,24]

1.4 Detection modes in FT-ICR MS:

1.4.1 Nyquist frequency:

Ions of different inherent mass to charge ratios have distinctly different cyclotron frequencies, but as mentioned before the relationship of cyclotron frequency and mass-to-charge ratio is inversely proportional. Due to the increase of cyclotron frequency with increasing magnetic field strength (Equation 2), ions of low/small m/z values have extremely high frequencies and orbit the ICR cell at extreme speeds. For instance a singly charged ion of 1000m/z in a 12T magnetic field will have a cyclotron frequency of 1.16MHz and a natural orbit radius of ~0.06mm (0.000061m) using equations 2+8, below:

$$r = \frac{mv_{xy}}{qB_0} = \frac{1.036427 \; x \; 10^{-8} \left(\frac{m}{z}\right) v_{xy}}{B_0}$$

Eqn. 8

Where r is the cyclotron radius in meters for an ion of velocity v_{xy}, contained within a magnetic field of B_0 Tesla.

The orbit is therefore 0.3833mm in circumference and so the ion is moving at 443.9 m/s (998.7 miles per hour). While a singly charged ion of 100 *m/z* will have a natural cyclotron frequency of 11.58MHz and an orbit of 0.1254mm, meaning the ion is moving at 1683.6 m/s (3788.1 miles an hour), which is a very large difference in speed for the two unexcited ions, though this does assume unperturbed frequency, no magnetron motion, and no space charge effects, the difference is clear as an example. It can therefore be harder to detect such fast moving ions of low m/z, despite the theoretical benefit of working in the low *m/z* domain regarding the ability to resolve peaks. The consequence of such fast moving ions is the need to be able to detect such high frequency species, especially the ability to sample the detected current fast enough to keep up with those ions of interest. Ions above the sampling rate will simply not be detected as they will be making multiple revolutions between each

sample event and the net difference between the 2 detection plates will not be detected. As a result the highest frequency (i.e. lowest m/z) scan point desired dictates the scan rate needed for a given experiment, if the user wishes to scan to lower m/z values then the signals must be sampled faster to keep up. The highest frequency desired is known as the Nyquist frequency, the sampling rate for the resulting scan is required to be a minimum of twice as fast as the Nyquist frequency to avoid aliasing effects and product an accurate representation of the detected signals.[7,4] For an acquisition containing a given number of data points (for example 4MW – 4194304 points), the Nyquist frequency also dictates the length of the resulting transient, and thus influences the maximum achievable resolving power for a given m/z in the mass spectrum. As a result for usual operation the desire is to have a scan range that encompasses the desired range (e.g. 150-3000 *m/z*) but without impacting negatively on the resolving power performance required. The reverse of this trade-off is also an effective way of increasing the achievable resolving power for a given number of data points – by scanning at higher m/z values (lower frequencies) much longer transient durations are achievable and thus resolving power performance is increased. Thus it is always important to know and record the scan range in any FTMS experiment. The use of higher m/z scans to enhance performance has been of great use in studies where the size of the resulting data may be a limiting factor and smaller end file sizes were required. This was the basis of using heterodyne detection (discussed below), and has implications in large-data applications of FTMS such as 2-Dimensional MS (discussed later).

1.4.2 Broadband mode:

As mentioned previously; ions at a low (natural) cyclotron orbit are excited by a resonant pulse to a higher orbit radius within the ICR cell to enable effective, sensitive, detection of coherent ion packets. Excitation of single ions is fairly trivial, however excitation of a wide frequency (and thus wide m/z) range to view a whole selection of ions in a single FT-ICR MS experiment is less straight forward. In order to view ions of many frequencies, many resonant frequencies must be produced to

excite ions, this is usually achieved via a frequency "chirp", see below Figure 16. The chirp as viewed in the frequency domain increases in amplitude at the low frequency cut off desired (often back-calculated via the high m/z range desired in the resulting mass spectrum using the cyclotron frequency equation) and then remains at a high amplitude at a reasonably constant (though slightly varying) level until the high frequency (low m/z) range value is reached.
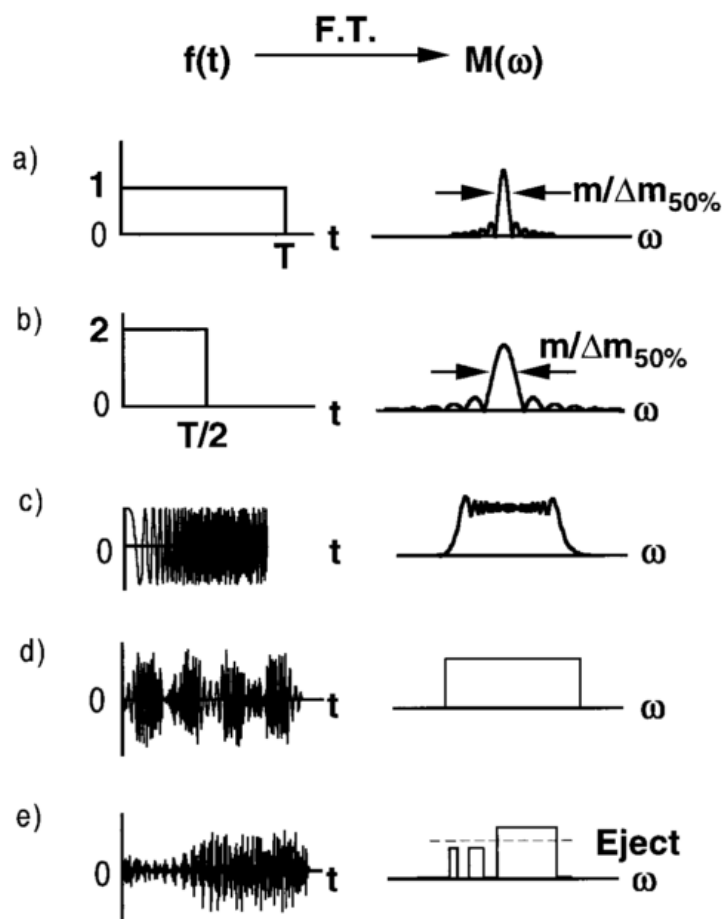
Figure 1.16: Various frequency pulses used to excite ions to a larger radius within the ICR cell and enable detection. A) and b) are rectangular pulses of different times and magnitudes, c) the standard "Chirp" excitation profile used in the vast majority of FT-ICR MS instruments and experiments, d) Stored Waveform Inverse Fourier Transform (SWIFT) excitation profile, discussed below, and finally e) SWIFT Isolation/ejection excitation profile. Reproduced from Marshall *et.al.*[25]

The so-called "broadband chirp" (Figure 1.16c) provides the benefit of a wide frequency excitation, relatively easily, and so allows the detection of a wide m/z range only limited in theory by the frequency generation electronics. However the broadband chirp also exhibits a frequency profile which is not completely flat/consistent in the region of interests, and also exhibits "horns" at the beginning and end of each pulse, over-exciting ions in these regions and producing irregular behaviour and peak shapes at the very extreme frequency values in the resulting mass spectrum.[5],[25]

1.4.3 Heterodyne mode:

Another popular detection mode is referred to as heterodyne/narrowband mode.[7,26] Heterodyne mode uses a smaller number of data points than the average broadband FT-ICR MS spectrum (e.g. 128k vs 4M data points) but over a much smaller frequency/m/z range (typically 1-50m/z), resulting in a higher number of data points per frequency/ m/z, and so much longer transient length and thus much higher resolving power for the resulting spectrum.

The principle of heterodyne mode acquisition relies on the use of a reference frequency alongside the usual recorded image current signal. A reference frequency near to the desired experimental range is chosen, e.g. 200kHz reference frequency for a measurement of ~215kHz). The image current detected will be a sinusoidal transient signal at ~215kHz, this will then be summed with a cosinusoidal reference signal at 200kHz, resulting in a composite signal. The composite signal is then processed through a low-pass filter producing a heterodyne signal containing frequencies equal to the sum and differences of the two signals, leaving the resulting transient of ~15kHz spectral width, which can be sampled at 30kHz instead of the broadband/single frequency equivalent which would have to be twice the value of the frequency, i.e. 430kHz. As a result the spectrum can be sampled at a much lower rate, while still using the same amount of physical memory, this translates to an elongation of $T_{acquisition}$ (T_D), which is directly proportional to resolving power. For the example above, a standard 4 MegaWord (MW) data set would result in a transient length of 10.2 s at a 430kHz sampling rate. Using heterodyne detection and the reference frequency stated the same 4MW data

set at 30kHz results in a T_D of 146.1 s, a performance increase of over 14-fold compared to broadband/single frequency excitation regimes. However, when using heterodyne mode the reference frequency dictates the frequency range scanned, as a result only frequencies in the range of 185-215kHz (200kHz +/-15kHz)would be measured during this heterodyne experiment. Increasing the width of the heterodyne scanning range would mean changing the difference between the reference frequency and that of the measured target, resulting in a decrease in the benefit of heterodyne-processing of the signal.

Heterodyne mode is currently less utilised than broadband mode (due to its narrow field of scanning), and less than historically where it was exceptionally useful as a way of attaining ultra-high resolving powers while using smaller file sizes, whereas contemporary FT-ICR MS is less limited by computer file size/memory requirements than previously. Though heterodyne mode continues to find niche applications and is still an excellent method for acquiring ultra-high resolution spectra.[13,26]
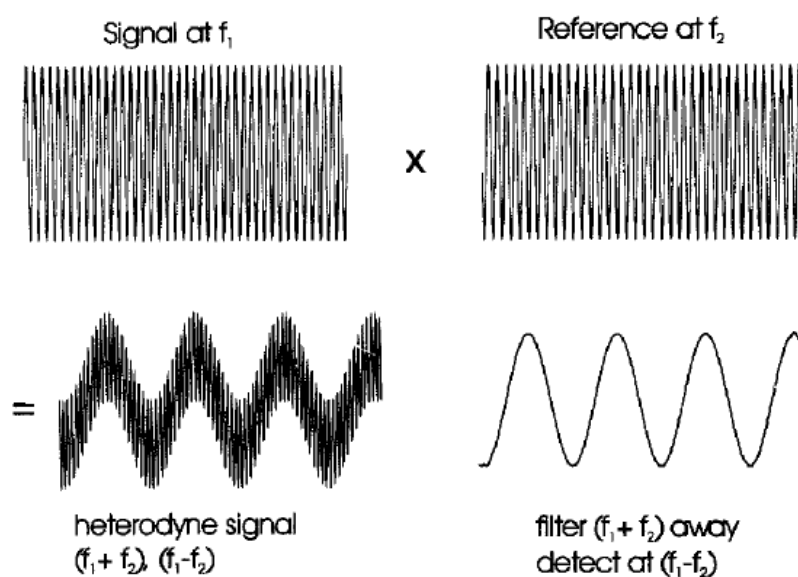
Figure 1.17: Heterodyne (narrowband) signals before (left) and after (right) removal of the reference frequency used to reduce the sampling rate and extend the available T_D for a given number of data points. Reproduced from Amster et.al.[7]

1.4.4 SWIFT:

As mentioned previously; a broadband chirp excitation profile is not flat and can produce unwanted effects at the extremes of excitation frequencies. In order to combat these short-comings Marshall *et. al.* developed another method for exciting ions inside an ICR cell.[25] Named the Stored Waveform Inverse Fourier Transform (aka SWIFT), the concept behind the method was instead of mimicking a flat excitation profile as with the chirp, SWIFT predicts the frequency profile desired (a perfect rectangular excitation, Figure 17, above), then uses an inverse Fourier transform to calculate the time domain signal required to achieve the perfect excitation profile needed (Figure 1.16d, right). The SWIFT time domain excitation pulse looks remarkably different than the standard broadband chirp, but it is also clear to see the excitation profile in the frequency domain is clearly flatter and more defined than the corresponding chirp, as a result SWIFT is a more ideal excitation profile, on instruments where it is available.

SWIFT has been used effectively to both excite ions ready for detection[25] and for isolation of ions (see section 1.6.3.2).[13,27] SWIFT is also useful for excitation of multiple chosen frequencies (/m/z values) using a "notched" excitation profile with excitations at each chosen frequency, providing a way of monitoring chosen ions of interest without exciting all ions in a given population.

1.4.5 Signal averaging:

As with many analytical techniques, although peaks may be visible using a single scan/experiment, the intensity of target peaks may be low. In order to increase the intensity of peaks within the resulting mass spectrum, without accumulating an excessive number of ions which can lead to space charge effects, FT-ICR mass spectra can be summed during a long multi-spectra acquisition experiment. Summing many transient signals prior to Fourier transformation, calibration etc., results in an increase in signal intensity for all signals (including noise), and an increase in signal to noise of persistent peaks only, meaning random (white noise) peaks will grow slowly (as $\sqrt{n}$) but remain relatively low, while persistent (real) peaks will steadily grow linearly over the course of the

experiment. This increase in signal to noise ratio is directly proportional to the square root of the number of scans summed, so the benefit of summation can plateau relatively quickly. Aside from the long experimental time required for long (100+ summed-scan) acquisitions, the only practical disadvantages to long signal averaging experiments are linked to sustainable ionisation/delivery of analyte ions to the ICR cell. Scan-to-scan variability in ion population results in a change in the space charge and magnetron frequency effects discussed earlier, resulting in very slight changes to the observed (reduced) cyclotron frequencies of ions. If the flow of ions for each scan is not consistent then a scan of unusually low ion population will produce a peak offset from the otherwise consistent and growing peak produced by the "average" ion populations otherwise recorded – leading to split peaks and/or peak broadening as multiple, very similar, frequency ion packets are summed together.

1.4.6 Ion coalescence:

As mentioned before FT-ICR MS relies on the excitation of incoherent individual ions orbiting within a magnetic field into discrete, coherent, and detectable ion packets via the application of an r.f. pulse to excite them to a higher cyclotron radius. Excessive ion populations have been shown to cause detrimental effects to this process (space charge); most notably frequency shifts and peak splitting resulting from dephasing of an individual ion packet into two or more closely spaced, but temporally separate ion clouds. Another phenomenon observed in FT-ICR MS is the exact opposite of this, i.e. closely spaced ion clouds converging and coalescing into a single detectable entity. Often occurring with large ion populations; peak coalescence can be observed most easily when two species are present in an analyte spectrum, yet only one peak is observed in the resulting mass spectrum, despite having sufficient/excessive analytical performance (resolving power) to separate the two species into individual peaks, as shown below in Figure 1.18.
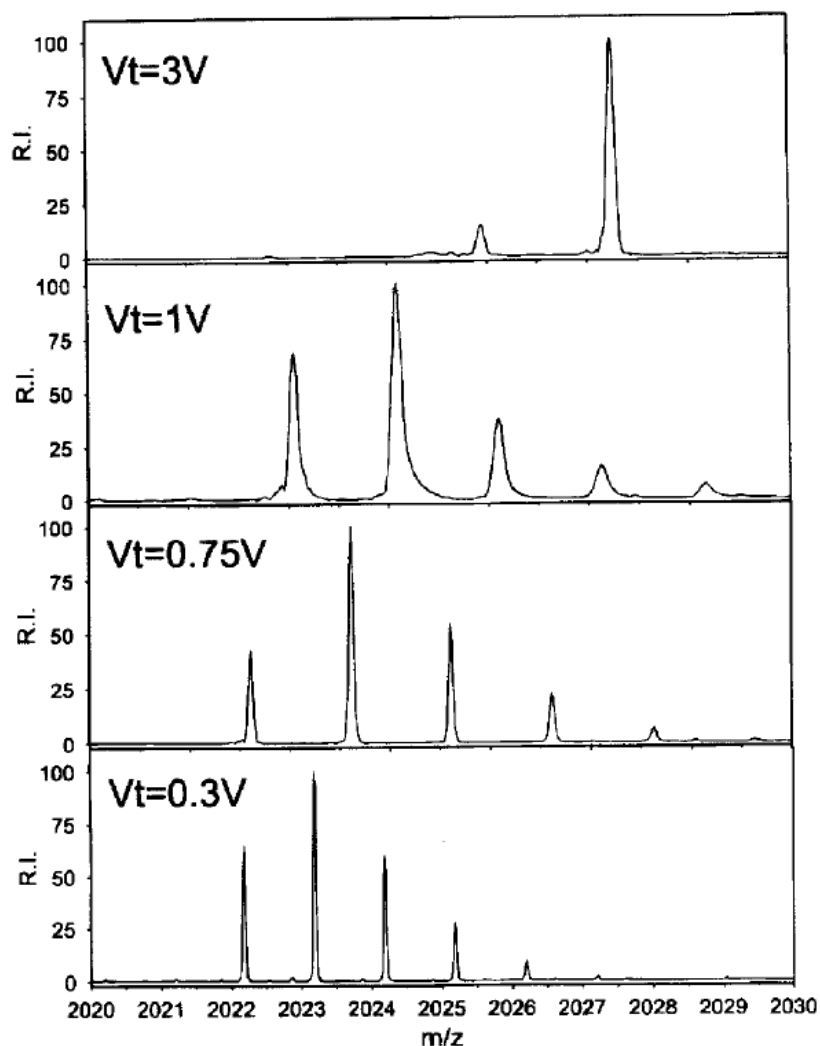
Figure 1.18: Mass spectrum of Poly(ethylene glycol) (PEG) using various trapping voltages during FT-ICR MS. The movement and eventual combination of peaks (at Vt 3V) shows the effect extreme trapping voltages can have on the resulting mass spectrum, but also shows how different ion packets can coalesce into a single peak, despite achieving the resolution to separate them. Reproduced from Amster *et. al.* [7]

The effect is enhanced when using high trapping voltages during confinement and detection of ions within the ICR cell, as a result most cell designs and FT-ICR MS experiments strive to minimise the trapping voltages needed to achieve acceptable performance. Using higher field superconducting magnets helps to reduce peak coalescence, making them more suited both for high complexity mixtures (to resolve many thousands of peaks simultaneously) and for ultra-high resolution scans of

target compounds. Despite the ability for many low field instruments to reach very long transient length (T_D) scans, higher magnetic fields are more able to effectively mitigate unfavourable motions/effects within FT-ICR MS, such as ion coalescence, magnetron motion, and space charge effects.

1.5 FT-ICR MS cells:

1.5.1 Geometry and development:

The ICR cell is the heart of the FT-ICR mass spectrometer, it is the mass analyser required in all FT-ICR MS based experiments, yet ICR cells appear in many varieties, each designed and optimised to perform particularly well in certain situations and/or combat one or more challenges during an FT-ICR MS experiment (trapping field effects, z-axis ejection, peak harmonics, etc).

The first FT-ICR MS cells were based on the cubic geometry (Figure 1.19, below), with straight forward fabrication, able to fit effectively between the poles of permanent magnets, and to a lesser degree in superconducting magnets, cubic ICR cells were used for many years on extremely successful platforms.[26,28,29] Cubic cells had typical electrode configurations – two excitation electrodes, two detection electrodes, and two trapping electrodes, all equal in size. Multiple-cell instruments were also constructed using Cubic cells for unique experiments,[5] more recently this approach has been revisited and utilised for fast acquisition using 3 linked cubic cells in a parallel acquisition FT-ICR MS experiment.[30] However Cubic cells suffered from low ion capacity and were not optimised for superconducting magnet geometries now in more common use.
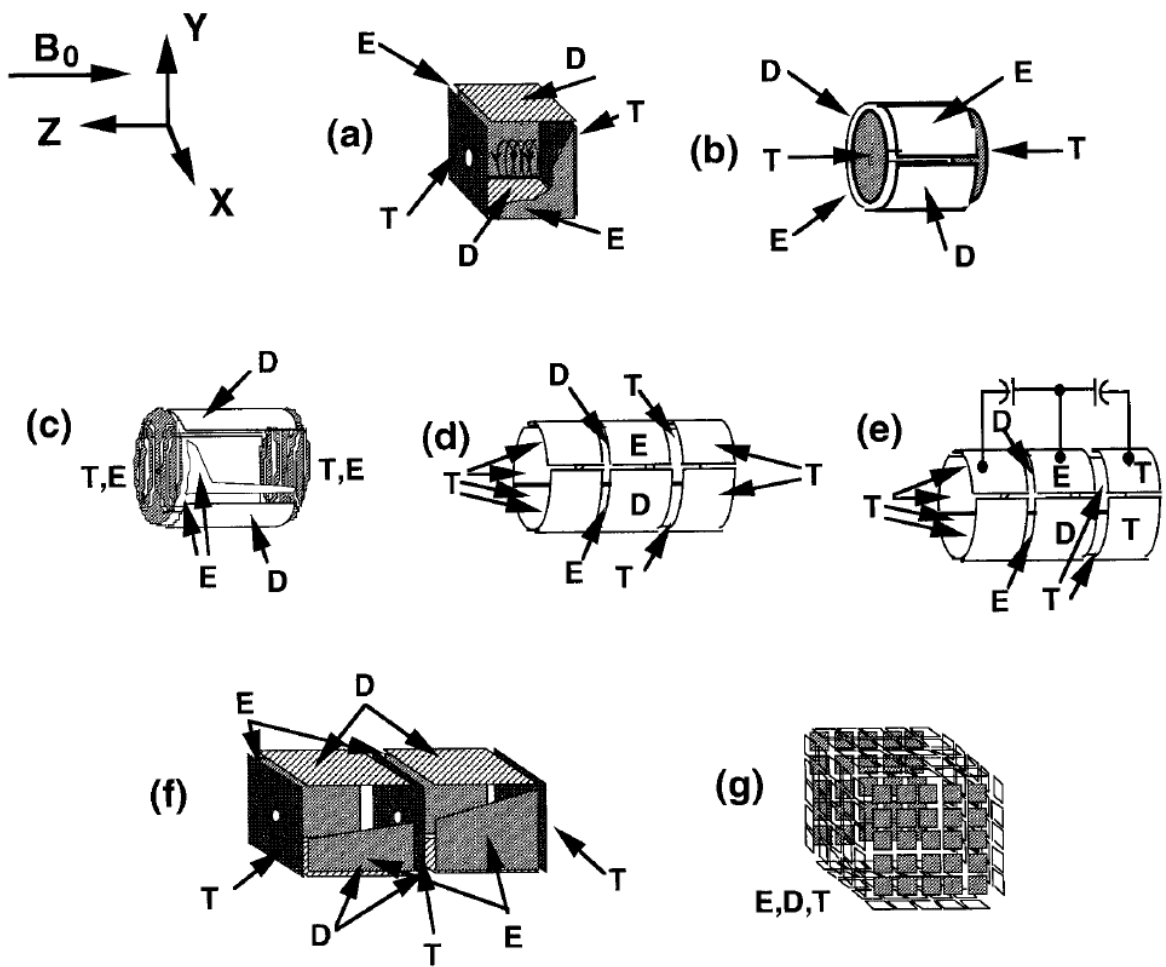
Figure 1.19: examples of various ICR cells which have been effectively used in FT-ICR MS, including the (a) Cubic (b) cylindrical (c) segmented end cap cylindrical cell aka the "infinity cell" (d) open cylindrical cell (e) open capacitively coupled cylindrical cell (f) dual cubic cell (g) "matrix shimmed" cell. E = excitation D = detection T = trapping plates. Reproduced from Marshall *et. al.* [5]

Cubic cells were succeed by cylindrical ICR cells, using rounded, quarter-circumference excitation/detection plates, with either circular end cap trapping electrodes (for closed cylindrical cells) or cylindrical trapping plates (for open cylindrical cells). The curved shape of the detection electrodes used in cylindrical cells enabled a more prolonged interaction with ion-packets as they passed the surface during detection, increasing the interaction between ions and available electrons, improving sensitivity. End-capped cylindrical cells suffer from electric field permeation throughout the ICR cell and disruption of the ion orbit away from the centre of the cell, as discussed previously.[31]

While open cylindrical cells suffer more from z-axis ejection of ions during excitation, due to resonant energy transfer into this frequency mode.[32] In order to combat these issues ever-improving ICR cell designs have been released; the capacitively coupled open-cylindrical cell was introduced by Beu *et. al.*[32] in order to capacitively couple the excitation and detection electrodes during ion excitation to eliminate z-axis ejection of ions along the magnetic field axis. In another approach to avoid the problems caused by using end cap electrodes and a cylindrical cell of finite dimensions, Caravatti *et. al.*[33] introduced the aptly named "Infinity cell", a closed-cylindrical cell with special resistively-coupled, surface-mounted trapping electrode wires which are positioned and designed to mimic an infinitely long trapping electrode geometry, and thus minimise/eliminate some axial excitation and ejection during the ion excitation event.

The Infinity cell and the capacitively coupled open cylindrical cell are both shown above in Figure 1.19. The Infinity cell has for many years been the proprietary cell of Bruker Daltonics FT-ICR MS instruments, measuring just 6cm in diameter and 6cm in length, the infinity cell has been shown (with careful tuning and operation) to achieve performance into the multiple millions of resolving power[34] and low part-per-million (ppm) to part-per-billion (ppb) accuracy in mass measurement.[34] The majority of the work presented herein was conducted on a Bruker Daltonics FT-ICR MS instrument fitted with an Infinity cell FT-ICR cell and so only the infinity cell will be shown in detail.
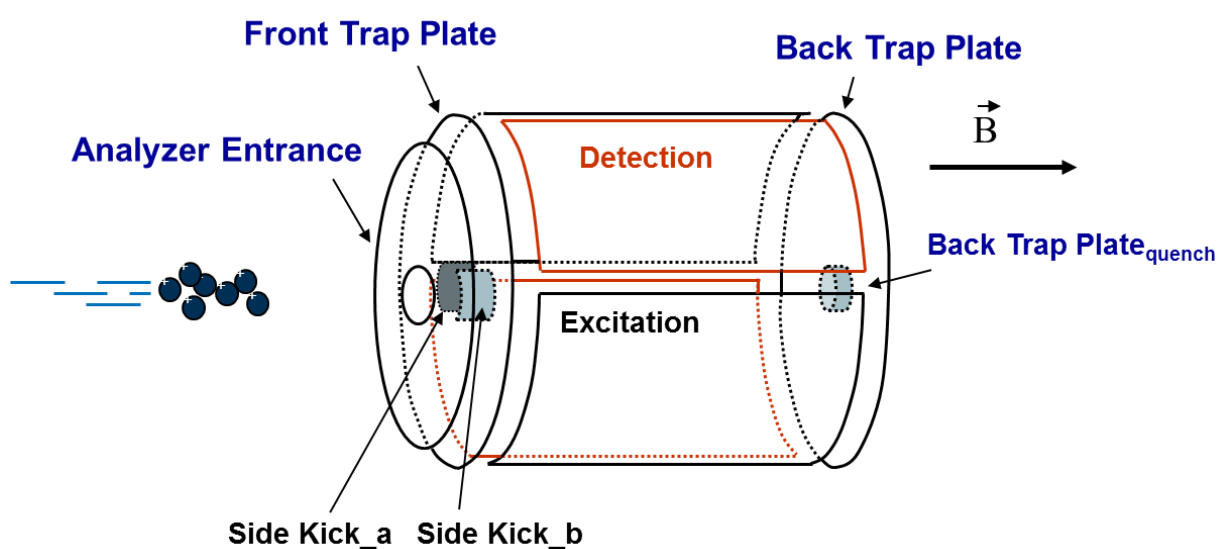
Figure 1.20: Schematic representation of the Infinity ICR cell, a closed cylindrical cell equipped with side-kick electrodes, two excitation plates, to detection plates, and two trapping plates. Courtesy of Bruker Daltonics, Bremen, Germany.
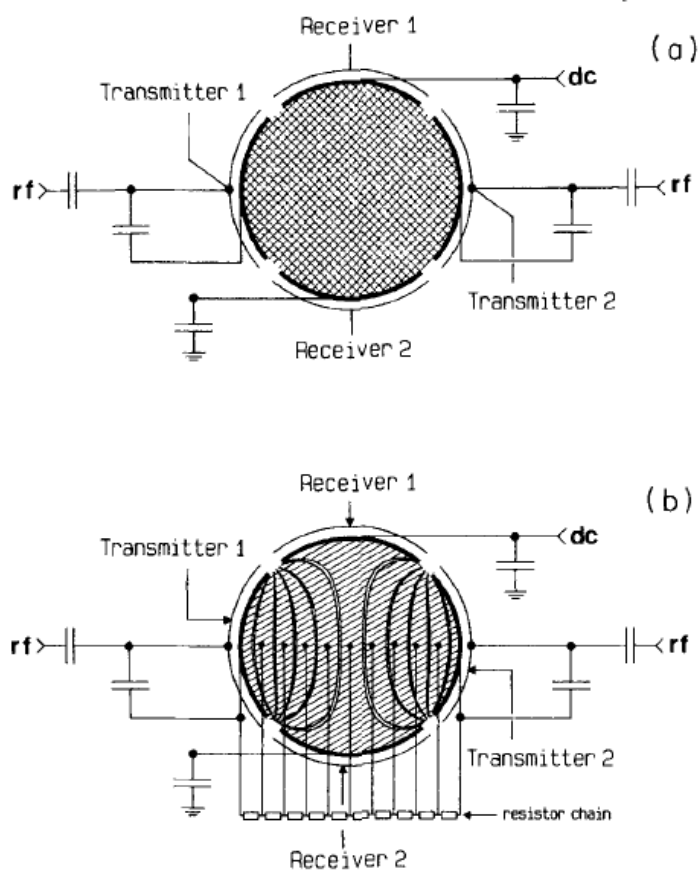
Figure 2. Wiring of the excitation rf potential. In addition to one trapping plate, the detection and excitation electrodes are indicated. The connections are essentially the same for both the resistive trapping plate (a) and the eleven-segment trapping plate (b). The rf voltage divider in (b) consists of ten 1 kΩ SMD resistors mounted directly on the rear side of each trapping plate. The capacitors shown are located outside the vacuum chamber.

Figure 1.21: Left: Schematic of the Infinity cell trapping electrodes (a), showing the resistor chains and circuitry associated with the Infinity cell design (b) which attempts to mimic an infinitely long cell in order to avoid unwanted ejection of ions along the z-axis during excitation. Reproduced from Caravatti et.al.[31] Right: Image of the Infinity cell mounted on a standard 6" stainless steel flange for connection to the end of the flight tube of an FT-ICR MS.

The Infinity cell has the benefit of virtually eliminating the unwanted ejection of ions along the z-axis during ion excitation by using the capacitor-functionalised trapping plates to mimic an infinitely long cylindrical cell. This enables higher sensitivity than the corresponding open-cylindrical cell as fewer ions are lost during excitation. The Infinity cell has been shown to require slightly more tuning than open-cylindrical cell geometries, but can achieve extremely long transients and thus resolving power performance. The main disadvantages of the Infinity cell are the need for more tuning at high performances, the closed cell geometry producing challenges for application/alignment of lasers/electron beams for further experiments, and its small size (6cmx6cm) causing a lower ion capacity than a corresponding cylindrical cell which could fit in the same magnet bore. Though the Infinity cell suffers from less ion loss and thus increased sensitivity, which reduces the latter point.

1.5.2 Side-kick electrodes:

As mentioned previously magnetron motion will increase if ions are displaced from the magnetic field axis (z-axis). Ideally with perfect alignment of all ion optics/ICR cell/ionisation source etc. all ions would be able to enter the ICR cell perfectly on axis and magnetron motion/electric/magnetic field inhomogeneities would not cause any issues during measurements. However since these prerequisites are not feasible, adjustment of ions' position as they enter the ICR cell has become a very effective technique in correcting their orbit once in the cell and has enabled much more control over ion packets and allowed more ultra-high resolution experiments. Coined "Side-kick" electrodes by its inventor Marshal *et. al.*[35] these curved electrodes are placed at the entrance to the ICR cell, with a potential difference applied between the two. As the ions approach and pass through the ICR cell inlet they experience an asymmetric field from the sidekick electrodes and are deflected slightly while moving, altering their end-position within the ICR cell, meaning they begin a different orbit than usual. The alteration of an ions orbit can be detrimental to the measurement at hand if not balanced correctly, if the ion's are pushed too far away from the principal axis (z-axis), then this increases the magnitude of magnetron motion which can be deleterious to the resulting mass

spectra. However if the ICR cell/housing is very slightly off axis with the magnetic field, then a small adjustment of the side kick electrodes can often produce much more ideal results and enable much more stable ion orbits, resulting in longer stable ion signal, increasing resolution greatly. The sidekick electrode have also been shown to have a positive effect when their polarity is inverted during detection.[35] The sidekick electrode are labelled within the infinity cell, Figure 1.20, above. The effective voltage on each side kick electrode is determined by the "side kick voltage" and the "side kick offset", the two values are used in equations 9a and b (below) to dictate experimental parameters.

Side Kick_a = Side Kick Offset – Side Kick/2   eqn. 9a

Side Kick_b = Side Kick Offset + Side Kick/2   eqn. 9b

1.6 MS components:

Though FT-ICR MS can be described purely as the trapping, excitation, and detection of ions within the ICR cell, practically there are many individual components and theory to a working FT-ICR mass spectrometer before the detection process can occur, including; ion generation, transportation, filtering, excitation/cooling, and optional isolation and/or dissociation. These concepts, techniques, and their applications are discussed herein, including an overview of the mass spectrometer used in the vast majority of the work relating to this thesis. Again, though many of these concepts extend to use within many different mass spectrometers, the focus will remain on their use in FT-ICR mass spectrometers specifically.

1.6.1 Generation of ions – ionisation sources:

As mentioned continuously; mass spectrometry requires that the analyte species are charged ions, and so all mass spectrometry studies are gas phase measurements of mass to charge ratios. Though most samples /analytes will originally be present as a neutral species prior to analysis and so all

samples must be ionised in some fashion in order to be analysed by mass spectrometry. Thus the ionisation source is the first critical component in any mass spectrometer.

There are currently many different ways to ionise species of interest and to introduce them into a mass spectrometer for analysis, each with their own unique advantages and disadvantages, and suitability for different samples. The ability to ionise such a wide variety of compounds one way or another is one of the key features which makes mass spectrometry so widely applicable to studying such a vast array of different species, often in one experiment, but also in many individual analyses, each focusing on different types of species specifically for tailored analysis.

1.6.1.1 Electron Impact/ionisation (EI):

The first widespread ionisation source and still one of the most commonly used methods of generating detectable ions today, Electron Impact (EI) ionisation uses high energy electrons, typically 70eV, to ionise analytes by displacing a bound electron on the analyte species. By losing one electron the analyte becomes a radical cation ($M^{+\cdot}$), this species can often then be detected. Unfortunately radical cations are not as stable as their precursors and coupled with the excess energy the 70eV electrons add into the molecule, the resulting ions often fragment/dissociate into one or more fragment ions before they reach the mass analyser. These fragmentations are radical driven and usually strive to produce a more stabilised radical species (e.g. from a primary radical to a secondary/tertiary radical species), and a number of neutral species, which increases the entropic driving force. Radical re-arrangements are also common and can produce even more peaks in the resulting mass spectrum, complicating analysis.

Thus the main disadvantage of EI is that the $M^{+\cdot}$ species cannot always be observed (if the analyte does not contain the structure/functional groups to stabilise the radical effectively), and that ionisation of one species produces a very large number of peaks (Figure 1.22, below).
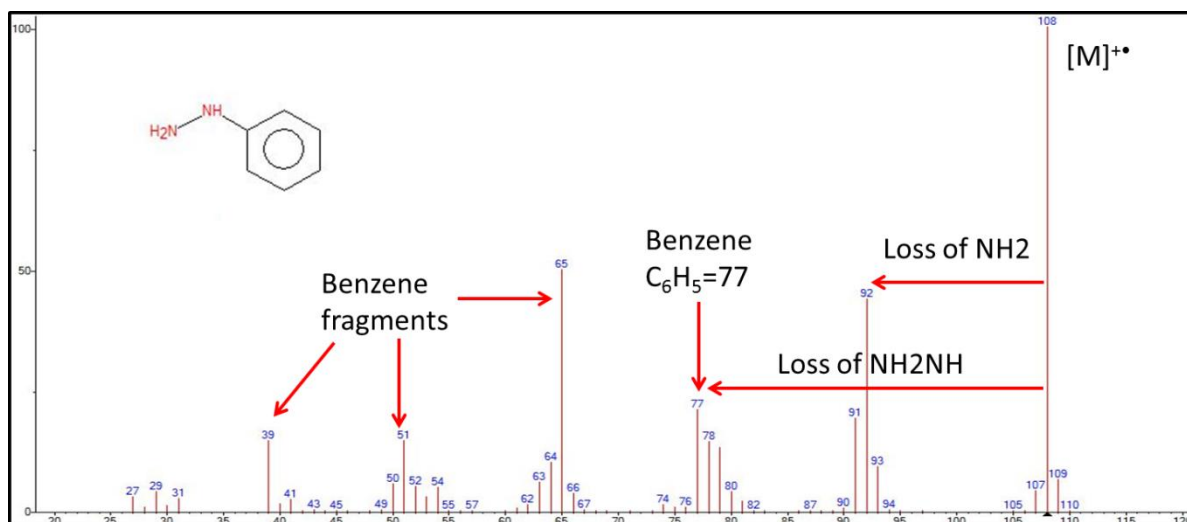
Figure 1.22: Electron-Impact (EI) generated ions of phenyl hydrazine, the molecular ion ([M]$^{+\bullet}$) is observed at high intensity showing the species is able to stabilise a radical functionality readily. Other fragments observed provide insights into the molecular structure and can be used to characterise and/or identify the species.

The main benefits of EI ionisation are the simplicity and robustness of the EI source, which is able to operate effectively with little maintenance and ionise the vast majority of small molecule analytes without tuning the electron energy away from 70eV, and that the fragments produced can be used to characterise the species they resulted from. The fragments produced by EI ionisation are formed from the radical dissociations and/or rearrangements of the molecule ionised. As a result the radical chemistry of each analyte is unique and will produce characteristic fragments for each and every analyte conceivable. The fragments produced may be as simple as $C_6H_5^{+\bullet}$ (produced by release of a phenyl group), but correlation with nearby peaks can quickly inform structure and atoms present (e.g. series of peaks separated by ~14Da often indicates saturated alkyl chain groups). A large amount of literature and experimentation has been conducted into EI fragmentation and application for a wide variety of analyte species. However EI ionisation is a gas phase ionisation technique and so is usually limited to application with volatile small molecules, as large/large biomolecules produce extremely complex EI fragmentation spectra, which yield little/no practical information due to the extreme extent of electron-induced fragmentation. The only effective EI-MS characterisation of

biomolecules was performed by Biemann *et. al.* and required extreme (complete) derivatisation of every individual functional group within the peptide analyte chosen,[36] this is of course not practical for wide-scale/routine biomolecule characterisation, which is the focus of the majority of contemporary mass spectrometry research groups. The low volatility of most biomolecules is also an extremely limiting factor in EI ionisation of biomolecules.

1.6.1.2 Chemical Ionisation (CI)

Chemical Ionisation (CI) was introduced by Victor Talrose in Russia[37] and subsequently developed extensively by Munson and Field.[38] Unlike EI ionisation, CI was designed as a proton-transfer ionisation source, instead of an electron based ionisation method. Chemical ionisation relies on the reaction of analyte species with highly reactive protonated gas ions created in a specialised chemical ionisation source. These gas-phase proton transfer reagents usually have an extremely low proton affinity, meaning they are extremely reactive Lewis acids (able of donating their extra proton). The formation of the gas-phase proton-transfer reagent is usually achieved through a glow-discharge source, using a high-energy spark to convert methane into (amongst other products) CH_5^+, an extremely reactive Lewis acid, this process is shown below in equations 10-12:

$$CH_4 + e^- \rightarrow CH_4^{+\bullet} + 2e^- \quad\quad\quad\quad [Eq.10]$$

$$CH_4 + CH_4^{+\bullet} \rightarrow CH_5^+ + CH_3^\bullet \quad\quad\quad\quad [Eq.11]$$

The CH_5^+ generated can then transfer its excess proton to analytes;

$$M + CH_5^+ \rightarrow [MH]^+ + CH_4 \quad\quad\quad\quad [Eq.12]$$

Chemical ionisation proved to be a particularly effective ionisation method, causing little to no fragmentation to analyte species, which was in stark contrast to EI at the time, providing predominantly MH^+ ions. CI proved to be effective at ionising polar, and semi-polar analyte species.[38] Though CI only produced singly charged ions, making data analysis of analytes trivial, this limited its

use to smaller analyte species, as the lack of multiply-charged ions prevented the study of larger species and particularly large biomolecules species such as peptides and proteins.

1.6.1.3 Electrospray ionisation (ESI):

Developed by John Fenn at Yale university in 1985, electrospray ionisation was not discovered from a desire to create ions for mass spectrometry, but instead as a serendipitous discovery while investigating molecular beams. Never the less John Fenn's discovery and then development of Electrospray ionisation has become an extremely effective and widely used ionisation method that is used almost ubiquitously in laboratories investigating (semi) polar compounds, complex mixtures, and most importantly; biomolecules.

Electrospray ionisation is a solution phase ionisation technique which utilises a fine needle through which the solution can be driven in order to create a fine mist/plume of sample, which while applying a high potential difference will become charged droplets and can be attracted towards a vacuum inlet. The eventual desolvation of these charged droplets liberates charged (protonated/deprotonated depending on the direction of the electric field) analyte species; ions, which can be manipulated, transferred into a mass spectrometer/analyser and detected.

The experimental arrangement for ESI is shown below in Figure 1.23. The solution is first exposed to a high potential difference between the solution inlet and the mass spectrometer inlet (typically 3000-6000 V), causing charge separation, the solution is then passed through a fine ESI needle (typically ~100um diameter, though this can vary), as the solution exits the needle it is mixed with an inert gas (usually nitrogen) to form a Taylor cone and nebulise the solution into a fine mist (ESI plume),[39] the charged droplets within the plume are then attracted towards the mass spectrometer inlet.
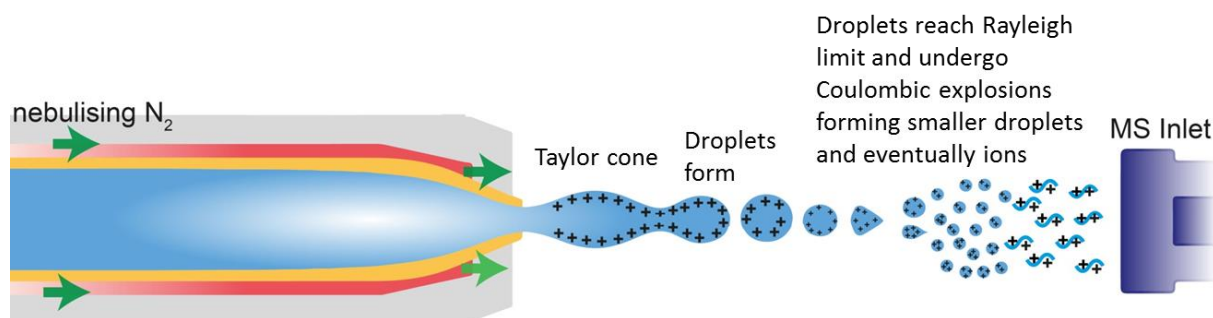
Figure 1.23: representation of the ESI process, showing the sample solution passing through the ESI needle and forming a Taylor cone, creating ESI droplets which will then desolvate, reach their Raleigh limit and undergo coulombic explosions/discharge events to form daughter ions, this process will repeat until detectable individual ions and/or stable droplets form.

As the charged droplets move through the electric field region and through the mass spectrometer inlet, desolvation of the droplets occurs. As the solvent evaporates, charge within the droplet remains relatively constant, meaning the droplets eventually approach the Raleigh limit of instability.[40] Droplets approaching the Raleigh limit will fission into two or more species (up to 14 in some cases)[41] the daughter droplets created will then continue to travel along the electric field gradient and continue to desolvate, undergoing multiple Raleigh "explosions"/fissions along the way, until eventually producing desolvated ions. The ions generated can then be detected within the mass analyser. Droplet formation can be visualised using high-resolution and high speed cameras, as shown in Figure 1.24, below, and has also been modelled computationally to show not only how droplets change shape while under various stress forces, but also how they dissociate and eject ions.[42]
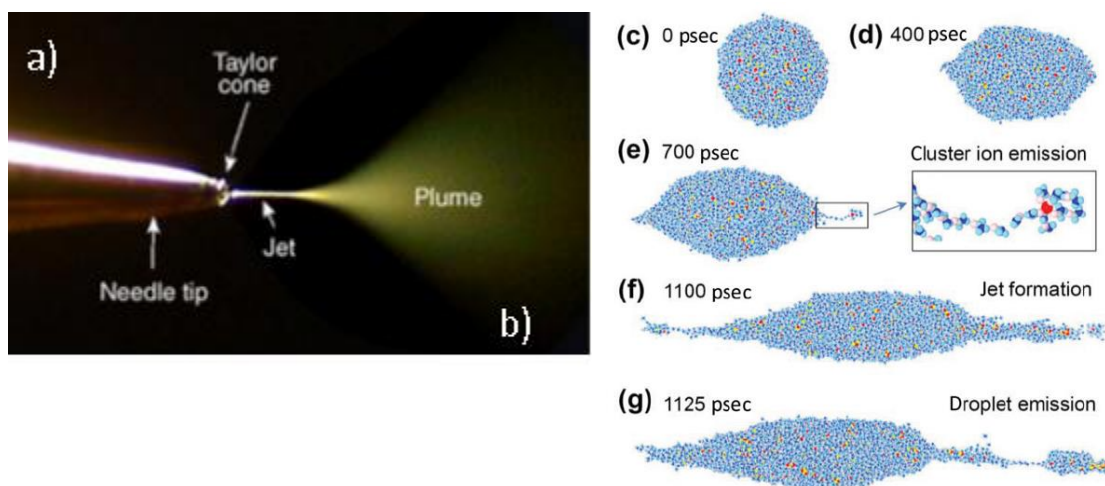
Figure 1.24: Charged droplet emission during ESI. (a) Photographs showing taylor cone formation and resulting ESI plume (b). (c)-(g) snapshots from Molecular Dynamics computational simulations showing the change in atomic configurations inside a 10nm ESI droplet of NaI (red) sequestered in formaldehyde (blue) within an electric field of 0.625V/nm. Reproduced from Meyer *et. al.*[42] and www.newobjective.com (accessed 12/08/2016)

The study of droplet desolvation occurring during ESI has been ongoing ever since the conception of ESI, unfortunately the process is still not completely understood. However a detailed study of droplet fission by Beauchamp *et. al.* was able to use Doppler Interferometry to study many and also individual droplets as they continue to desolvate within and electric field gradient. [41] Beauchamp *et. al.* showed that as the desolvation occurs ions undergo multiple "discharge events", upon which 10-25% of droplet charge is lost, whereas only 0-5% overall charge is lost. These discharge events continued until the droplets were no longer detectable. Though the fact that at least 10% of charge is lost with each event, yet some discharge events occurred with no detectable mass lost implies that highly charged species, with negligible mass compared to the droplet as a whole, are ejected during the ESI process (Figure 1.25, below).
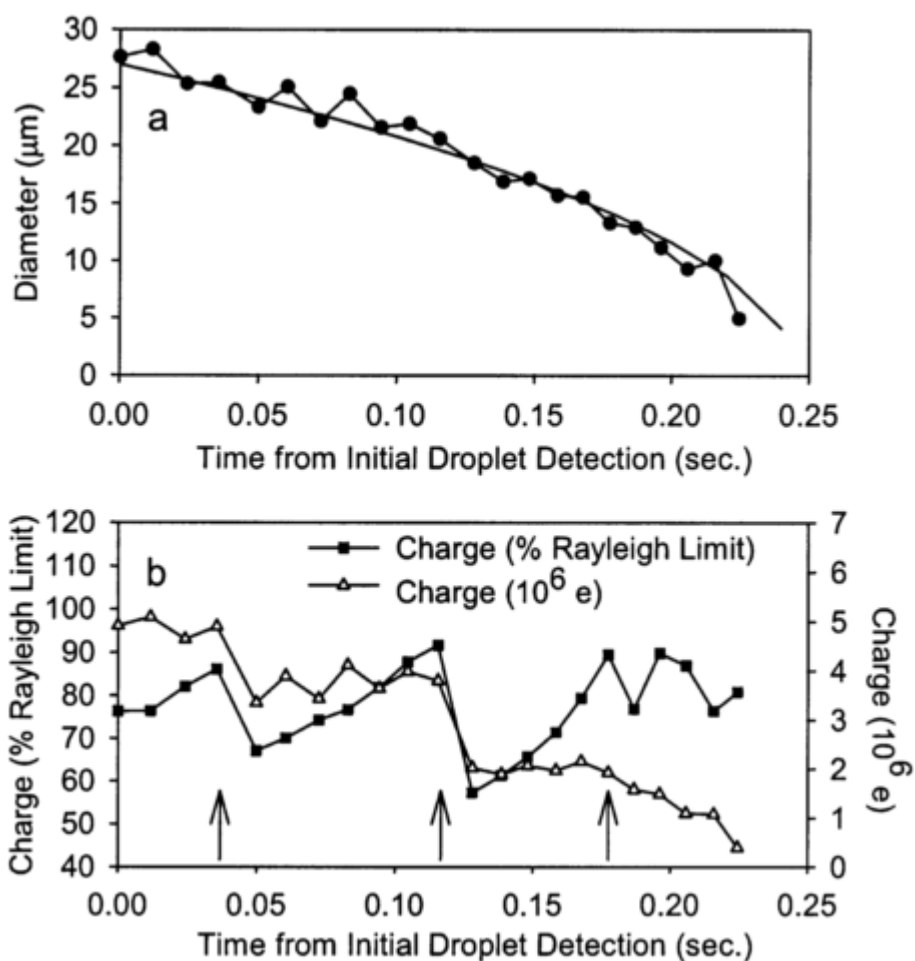
Figure 1.25: measurements from Beauchamp *et. al.*[41] showing the desolvation of a charged droplet formed from ElectroSpray Ionisation (ESI). The measurements show the continued loss of solvent, during which several "discharge events" occur when the Raleigh limit of stability is reached for a given droplet (indicated by arrows). Reproduced from Beauchamp *et.al.*[41]

The key advantages of ESI over EI and CI, mentioned above, are that a) it is an extremely "soft" ionisation technique – producing little to no modification of analytes (beyond the protonation/deprotonation), and most importantly b) it can produce multiply charged analyte species.

The ability to create multiply charged species greatly facilitated the analysis of particularly large molecules, by increasing their charge and lowering the resulting mass-to-charge ratio into a more convenient detectable and manipulateable range for most mass spectrometers. Though some mass

spectrometers are able to perform reasonably at high m/z values (>10,000 m/z), most notably Time

of Flight (ToF) mass spectrometers, the majority of mass spectrometers (including ESI-ToF's), FTMS

specifically, are at optimum performance at more moderate m/z values (~100-5000 m/z), as a result

the ability to multiply charge ions and bring them into the 100-5000 m/z region is a benefit to the

entire MS field. The impact ESI-MS has had on the field as a whole is reflected in the joint awarding

of the 2002 Nobel prize to John Fenn for his work on ESI, the other recipient Koichi Tanaka for his

work on laser-based ionisations, discussed below.

The mechanism of analyte charging and eventual ejection from the solvated-analyte charged droplet

state to free ion is still under debate, however several models have been proposed, the most widely

accepted of which are currently the Ion Ejection Model (IEM), Charge Residue Model (CRM), and the

Chain Ejection Model (CEM)[43]. Figure 1.26, below, summarises the 3 models, covered recently by

Konermann *et. al.*[43]
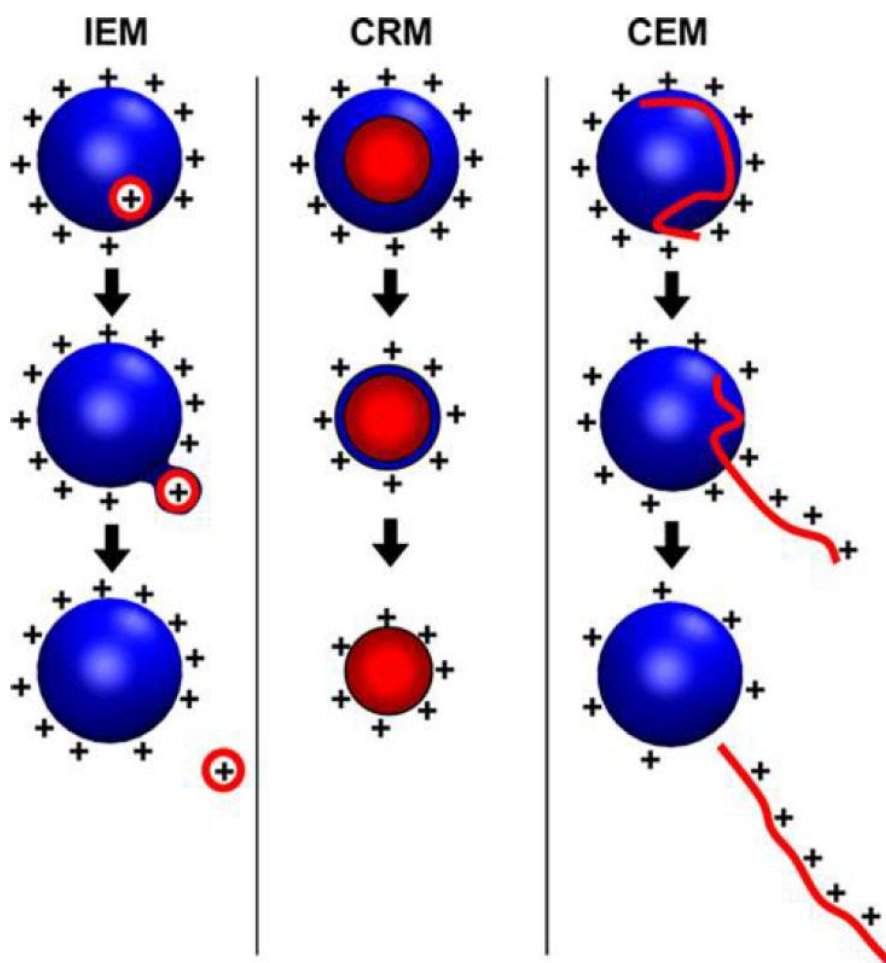
Figure 1.26: Summary of the three main models for ESI ion formation, summarised by Konermann *et. al.*[43] **From left to right; The Ion Ejection Model, the Charge Residue Model, and the Chain Ejection Model.**

Ion Evaporation Model (IEM):

The Ion Evaporation Model is mainly applicable to small molecules, usually charged via protons/salt ions (such as Na^+ or K^+) already in solution, these charged species can either reside in the bulk solution where they are stabilised via solvent interactions, much like salt ions themselves (see Figure 1.27, below), or near/at the surface of the droplets, which is where the IEM can take place. The location of the species within the droplet will thus be affected by the structure and functionality/polarity of the analyte, species with particularly non-polar groups will often be more stabilised at the surface of the droplets, and so can undergo the IEM more effectively, under the condition that they have functional groups to accommodate a fixed charge in addition to their non-polar groups.
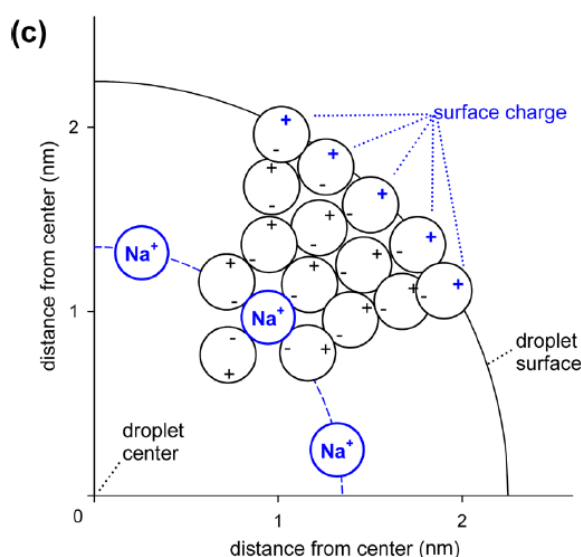
Figure 1.27: Snapshot of a Molecular Dynamics (MD) simulation conducted by Konermann *et. al.*[44] **showing the solvent structure reorganisation effect of Na^+ ions on a water droplet. Reproduced from Konermann** *et.al.*[44]

The IEM hypothesises that ions at the surface of charged droplets during ESI can be ejected, usually simultaneously with several solvent molecules, producing distinct ions from the parent droplet, which can then be analysed. This process is obviously stepwise and relies on the reorganisation of many solvent molecules near the droplet surface, several snapshots of the IEM MD simulations are shown in Figure 1.28, below. It has been shown that the electric field from a droplet near/at the Rayleigh limit of stability is repulsive enough to overcome the energy barrier of ejection(~32kJ/mol).[43] As can be seen from Figure 1.28 the timescale for the transition from solvated ion to ejected ion is ~1ns.
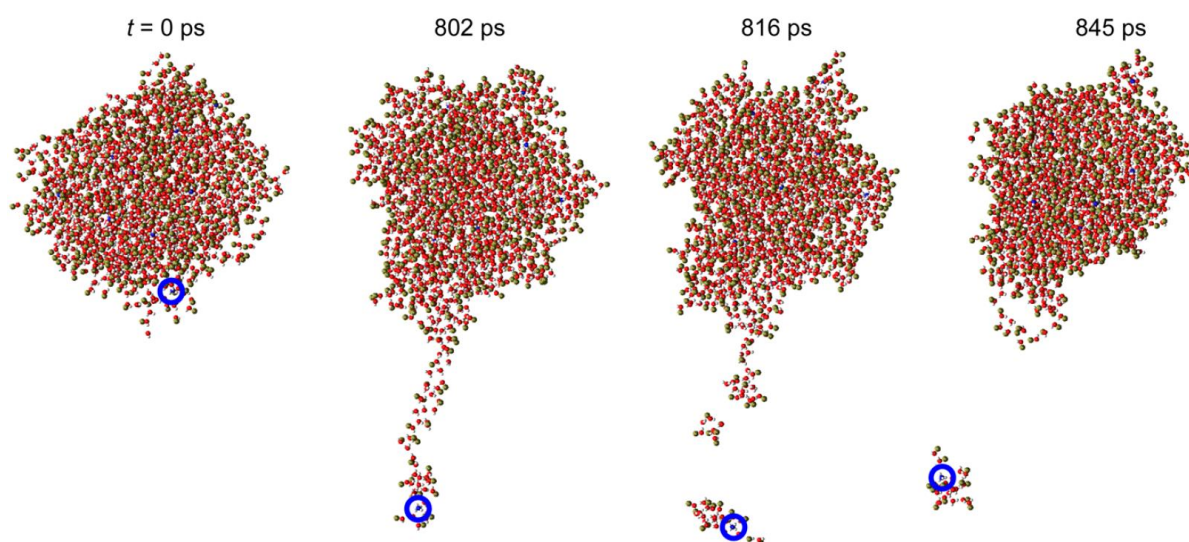
Figure 1.28: MD simulation snapshots at different times summarising the Ion Ejection Model (IEM) of ESI showing a small ammonium ion (NH_4^+) being ejected from an ESI droplet. Note that the ion is not necessarily ejected as NH_4^+ but is still associated with several solvent molecules, which will then be released during desolvation as they travel into/through the mass spectrometer. Reproduced from Konermann *et. al.* [43]

Charge Residue Model (CRM):

For large molecules such as proteins and protein complexes the energy barrier for ejection from an ESI droplet is far too high to allow the IEM shown above. Instead particularly large molecules are often believed to be ionised during ESI via the Charge Residue Model (CRM). During the CRM the

protein/large molecule remains within the droplet and it is the solvent which is ejected/removed via desolvation, until only the protein remains,[43] with the remaining surface charges being deposited on available groups on the protein's accessible surface. As a result the CRM has been mostly applied to large, folded protein charging during ESI, as unfolded proteins behave differently (see CEM below), and take considerably longer than the IEM (microseconds vs. nanoseconds) as the limiting step in the CRM is not the position of the protein inside the droplet, but the amount and nature of the solvent present.

Chain Ejection Model (CEM):

The CRM model has been shown to effectively explain the ionisation of large folded protein structures during ESI, however many proteins are analysed via ESI in denaturing conditions (such as organic solvent and/or acidic/basic conditions) which disrupts their tertiary structures, and often provide more effective ionisation and/or higher charge states observable in the resulting mass spectrum. Also the CRM does not effectively explain the charging of similarly unstructured non-bio polymer species. For these situations the most accurate model currently accepted is the Chain Ejection Model (CEM). The CEM suggests that disordered/extended chain structures (such as polymers/biopolymers) are less stable in the centre of the bulk solution due to the unfolding of their structure exposes hydrophobic regions usually concealed within the tertiary structure away from external solvent species. Due to this change the chains migrate towards the surface of the droplets. Once at the surface the terminus of the protein/polymer can be "ejected" from the droplet while the rest of the chain is still tethered to the surface. Over time the rest of the chain will be ejected from the droplet and become charged due to exposure and proximity to surface charges. Eventually the entire chain will be released from the droplet and become a free ion, possibly with some solvent/buffer adduction (as observed in the IEM). The CEM process is summarised in Figure 1.29, below:
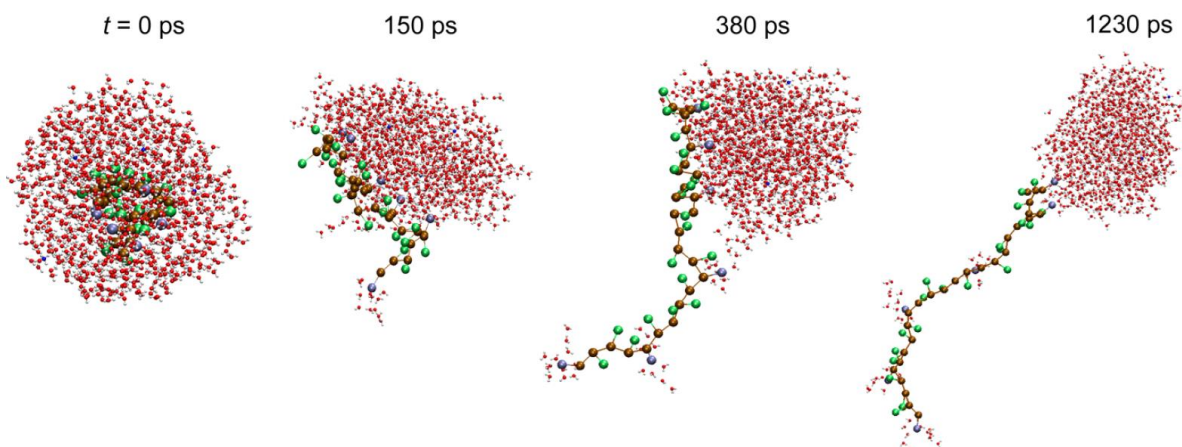
Figure 1.29: MD simulation snapshots of the Chain Ejection Model for ESI showing a long chain biopolymer being slowly ejected from an ESI droplet. Reproduced from Konermann _et. al._[43]

The CRM and CEM help explain the differences in protein ionisation observed under different conditions, as the solvent conditions can dictate the protein conformation in solution and thus will alter the ionisation mechanism the species undergoes during ESI. It has also been shown[43,45] that the maximum droplet size the CEM model is applicable is ~10nm, meaning that above this limit there may be limited ionisation of unstructured/chain molecules before Rayleigh/Coulombic explosions reduce the size of ESI droplets during the desolvation process detailed above. It has also been suggested[43] that the different amounts of charge which proteins can achieve (varying number of protons acquired/released during ESI), see below, could be due to the same protein being ejected from droplets of varying diameter and thus varying amounts of charge, as larger droplets are able to accommodate higher numbers of charge at their surface. This suggests that larger ESI droplets could produce higher charge state protein analytes, though this assumption does not take into account many other factors, some of which are described below.
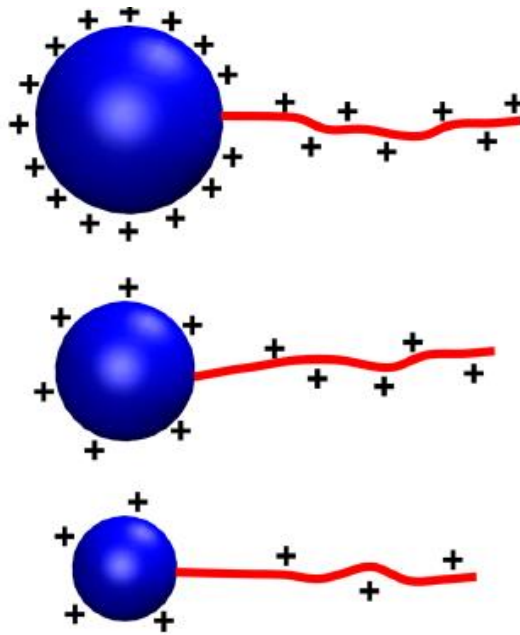
Figure 1.30: Summary of the sequential charging process from the CEM for ESI showing long chain polymers/biopolymers acquiring multiple charges upon existing different sized droplets and acquiring a different degree of charge. Reproduced from Konermann *et. al.*[43,45]

Protein ESI:

As mentioned above, ESI of large molecules can produce multiply charged species, one of the main benefits of the ionisation method, however the ability of say a protein to accommodate said charges will dictate the amount of charge acquired and thus influence the species observed via the mass spectrometer.

Peptides and proteins contain multiple functional groups capable of accommodating additional charge, most notably side chain residues of Arginine, Lysine, Histidine, a free amine N-terminus, and even backbone nitrogen atoms to a degree. Additionally (particularly acidic) proteins can be deprotonated and be observed in ESI via deprotonation of acidic amino acid side chains, such as Glutamic acid and Aspartic acid.

The ability for a particular charge-capable site within a protein to become charged is dependent on the conditions (solvent, temperature etc.), pH, and conformation of the protein itself. If a protein

has a large number of basic amino acids, able to be protonated readily in theory, but is folded into a tightly packed conformation with said amino acid groups concealed/protected from the protonating solvent, then they will remain uncharged during ESI and a protein may appear at a low charge state (if only a small number of sites can be accessed). Though many currently believe that charging of proteins during ESI may be more significantly affected by solvent-accessible surface area over purely number of accessible basic/acidic sites.These ESI-inhibiting conformational effects are usually overcome in protein MS via use of denaturing solvents (such as water-methanol or water acetonitrile (50-50 v/v), usually with the addition of an acid (e.g. formic/acetic acid) to decrease pH and aid in positively charging available groups.[42] The denaturisation of proteins promotes a more open structure and exposes more of the protein backbone/side chain residues, which can then be charged if possible, enhancing charging of analytes. However this approach, though commonly used, is not suitable if the study at hand has any focus on protein structure or conformation, which would need to be retained during the ESI process.[46] Use of denaturing/harsh solvent conditions for protein analysis has also been an issue for the study of fragile protein modifications, especially if their binding is pH dependant and/or easily displaced by acid/base/solvent molecules, which is the case with many metal-based modifications.[47]

Since proteins can be present in different conformations and attain different levels of charge during ESI, it is often observed that in ESI-MS of a protein-containing solution, the protein of interest is observed in the mass spectrum at multiple different charge states, e.g. Charge state 8^+, 9^+, and 10^+ for the lysozyme protein shown in Figure 1.31, below;
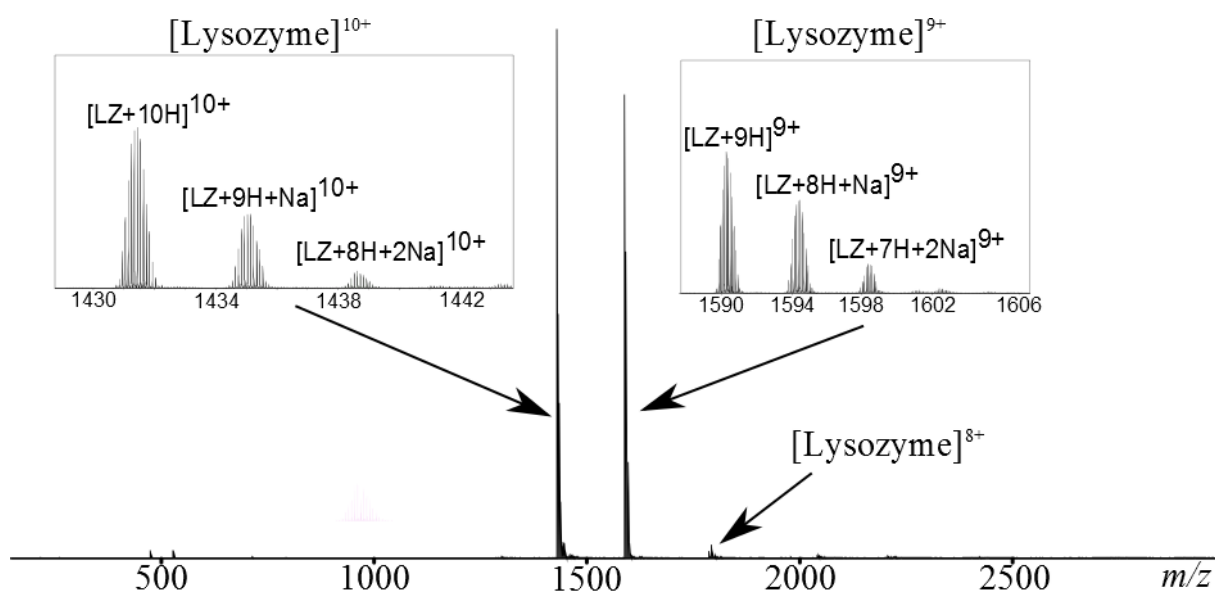
Figure 1.31: ESI-FT-ICR MS of Lysozyme protein in positive ionisation mode showing multiply charged lysozyme species in the 8+, 9+, and 10+ charge states. Inset: Isotopic distributions of the 10+ and 9+ charge states showing the protonated and Sodiated species present at each charge state.

The 8^+, 9^+, and 10^+ charge states refer to the species [Lysozyme+8H]$^{8+}$, [Lysozyme+9H]$^{9+}$, and [Lysozyme+10H]$^{10+}$, respectively. The observation of different charge states has the benefit of producing "higher" charge state species for extended experiments (see MS/MS below), at a lower m/z ratio (for higher FTMS performance), and the range of states can provide indirect information regarding protein structural conformations. However the production of a range of protein charge states, instead of a single charge state, "dilutes"/distributes the analyte signal over many, many peaks, and reduces the sensitivity of the MS measurement. Thus the ability to use techniques/conditions to control the charge states observed can be very beneficial to effective protein ESI-MS.[48,49]

In addition to the protonation/deprotonation achieved by positive/negative mode ESI (respectively), other charge carrying atoms can be adducted/removed from an analyte to achieve an observable charged species. The most notable of these species being free/solvated Sodium (Na$^+$), which has been a notorious adduct in ESI since its inception. Free Sodium can be found ubiquitously in all

experimental environments, from solvents, to glassware, and plastic-ware, down to fingerprints from the sample preparer. Sodium/salt adducts are usually expressed/observed close to the protonated/deprotonated peak and in a relatively salt-clean sample will be minor compared to the protonated/deprotonated peak of interest, (see Figure 1.30 zoom in, above), though if the sample contains too much salt/has been handled badly this balance can change and the salt-containing peaks can easily dominate a given spectrum causing the purely protonated/deprotonated peak to become unobservable, which can lead to false-positive/negative results for protein identification based purely on searches for protonated/deprotonated species. Though some researchers have found salt adducts to be particularly useful in the effective study of certain compounds, particularly small molecules which are not amenable to protonation/deprotonation (due to lack of functional groups/stabilising conformation), but can accommodate certain salt ions.[50]

The use of salt/buffer solutions to stabilise protein conformations, and certain protein-protein interactions, which can be of great biological importance, can create challenging ESI—MS conditions. The so-called Native-MS method, is growing continuously as researchers continue to study protein conformations and delicate interactions using mass spectrometry, but the use of salts/buffers and non-denaturing conditions (no organic solvent, no acids, and low ESI source temperatures) inevitably causes large numbers of salt/buffer adduct peaks from adducted/sequestered salt ions, distributes ion signals, and lowers observed charge states due to either in-effective ESI desolvation and/or densely folded protein conformations limiting availability of basic/acidic amino acid side chains and reducing solvent-accessible surface area.[51–53]

A nano-scale application of ESI has been developed over the years to provide a lower-sample consumption while retaining reliable and abundant ion signal, thus an increased sensitivity. The so-called nano-ESI (nESI) is usually achieved using a much smaller diameter orifice emitter than standard ESI (~10-100um) and thus smaller droplets are formed after the Taylor cone emission during the electrospray process.[54] The generation of smaller ESI droplets enables more effective desolvation down-stream, and tends to result in a more effective/efficient conversion of analyte to ions, resulting in a more sensitive ionisation source.
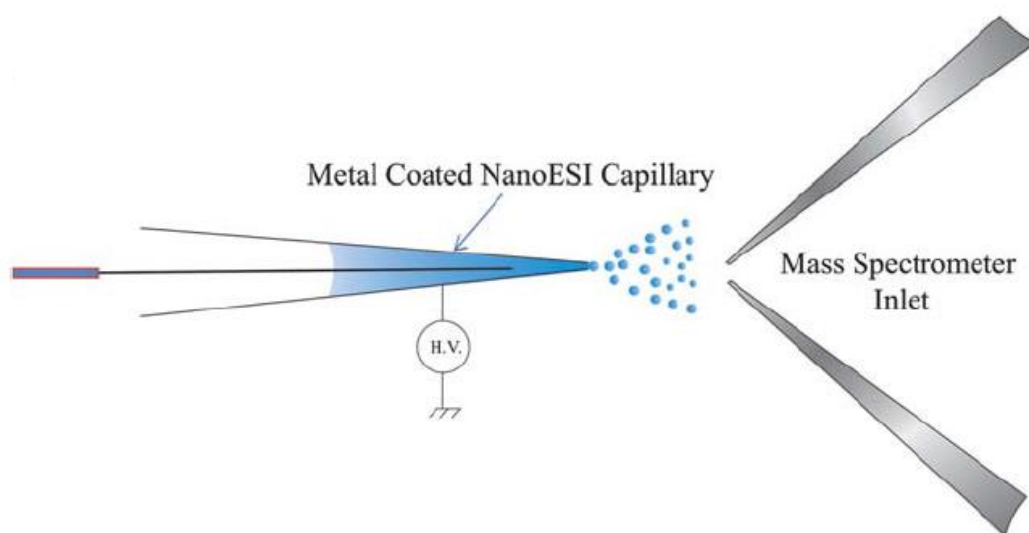
Figure 1.32: nano-ESI source layout, unlike standard ESI there is no syringe pump to drive the sample through the emitter, nor any nebulisation gas needed to assist in generation of the ion plume. Sample is drawn out from the emitter tip to the MS inlet via a potential difference applied between the two at either the liquid sample (via a wire, shown) or on the emitter itself (requires metal/conductive coating). Reproduced from Hiraoka *et.al.*[55]

Though largely similar to standard ESI, nESI has a few key features and differences to ESI beyond the emitter orifice diameter. Due to the low flow rates used (nano-litres/minute, hence nESI), nESI does not require an additional inert nebulising gas such as nitrogen to achieve an effective ESI plume. Though it has been clearly calculated and shown that flow rate does not affect the length of a Taylor

cone,[41] the creation of smaller droplets initially eliminates the need for further nebulisation. Due to the creation of smaller highly charged droplets earlier in the ESI process, nESI emitters are placed closer to the MS inlet than ESI emitters, as they do not require nearly as much space to allow the resulting plume to expand and produce ions. As a result of their closer proximity to the MS inlet, nESI typically uses much lower voltages to achieve effective ionisation (500-1800V for nESI vs 3000-6000V for ESI). nESI can also be achieved at lower source temperatures (due to less desolvation needed vs the large droplets formed in ESI), this is particularly advantageous for achieving Native MS of peptides and proteins, which may denature if heated to higher temperatures.

nESI tends to use pulled Boro-silicate glass/metal/silica emitter tips, which can often be created in-house via use of a heated capillary puller (also used for many electrochemistry experiments), or en masse from external companies, making nESI tips much more accessible and disposable than ESI emitters. Unfortunately nESI tips are generally not re-useable (unlike ESI emitters) and so need to be disposed of after use, however this does limit cross-over contamination, and nESI tips can be effective made/purchased for £0.05/£1 respectively, and so are economically viable for disposable use. The nESI emitter tips are then either coated with a conducting material (e.g. gold sputtered), or are bound to an internal conducting wire, to provide electrical contact with the analyte solutions and complete the ESI "circuit".

1.6.1.5 Matrix Assisted Laser Desorption Ionisation (MALDI):

MALDI, invented by Karas & Hillenkamp *et.al.*[56] while also developed by Koichi Tanaka (winner of the Nobel Prize for said work), along with ESI, are the two most-commonly used soft-ionisation techniques used in contemporary mass spectrometry. Unlike ESI, MALDI uses a solid sample, and is ionised within a chamber that is under vacuum, not atmospheric pressure (AP) (though MALDI has been achieved at ~1mbar by O'Connor *et.al.*[57,58] and at ~AP by Laiko *et.al.*[59]) MALDI relies on the laser-activation of matrix molecules which are in close proximity to analyte ions of interest, the laser-induced plume created just above the point of impact contains protonated/deprotonated

matrix ions and neutral analyte ions.[60] The protonated/deprotonated matrix ions can then undergo

proton-transfer reactions to protonate/deprotonate the analyte molecules into (mostly) singly-

charged ions, which are then extracted and transferred into the mass analyser for detection etc. as

shown in Figure 1.33 below:
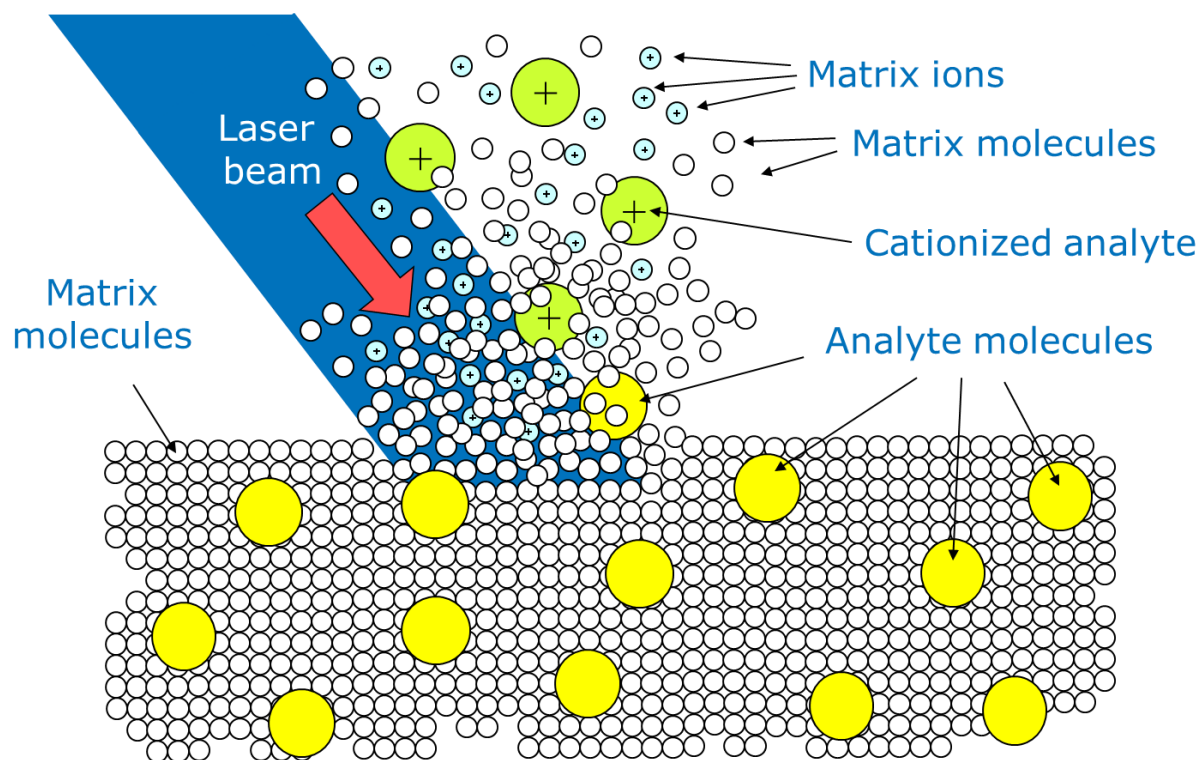
Figure 1.33: MALDI process involving the ablation of a surface with a UV laser, activating and

ionising MALDI matrix molecules into ions, which can then undergo proton transfer reactions with

nearby analyte molecules to produce detectable analyte ions, usually in the 1+ charge state.

Courtesy of Bruker Daltonics, Bremen, Germany.

Common MALDI matrices include aromatic species (to absorb the UV laser radiation used)

functionalised with acidic/basic groups (for proton-transfer reaction capability), some commonly

used matrices are shown in Figure 1.34, below:
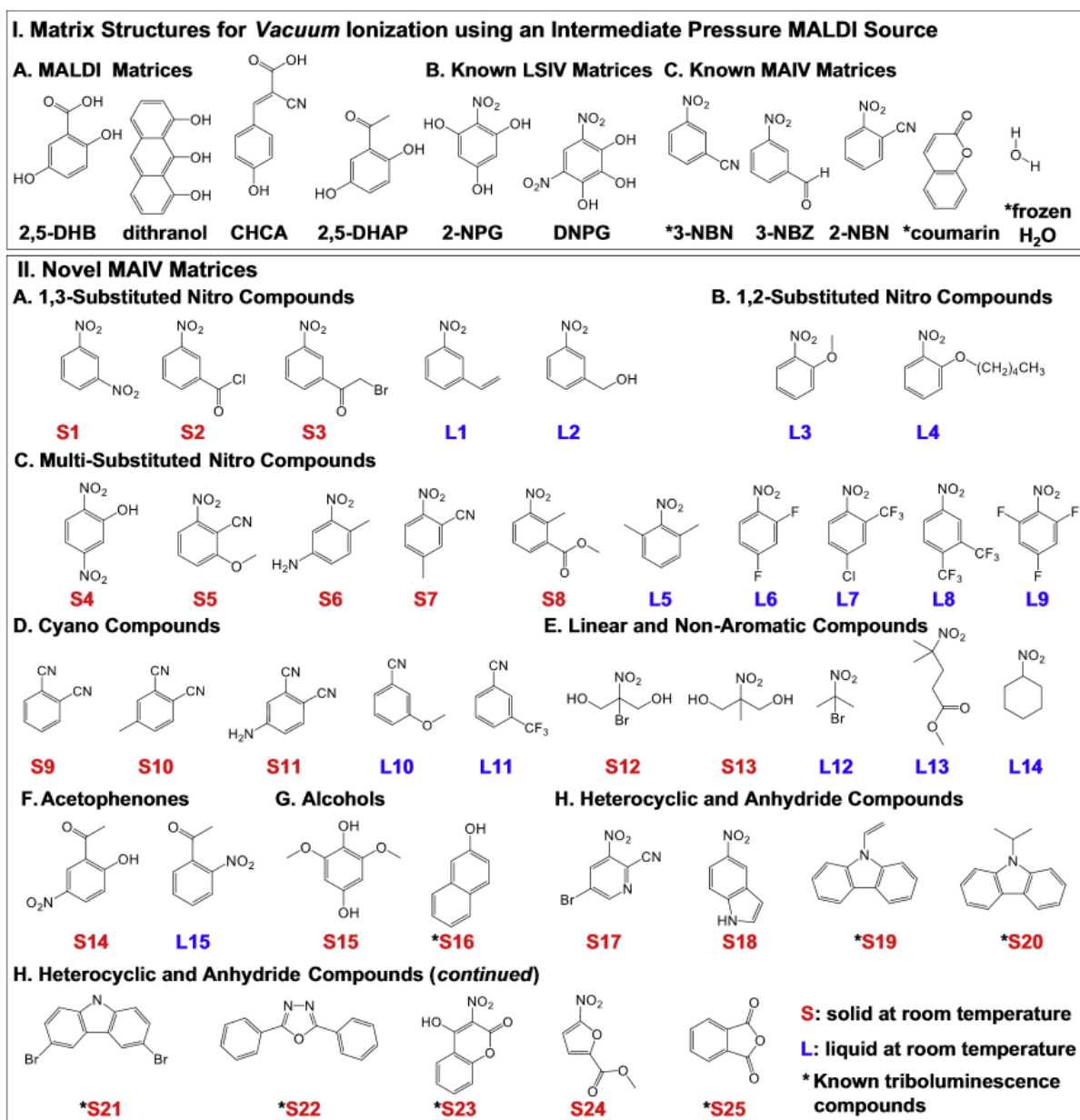
Figure 1.34: Common compounds often used as a matrix for MALDI-MS, note that MALDI matrices usually contain a benzene ring to assist in absorption of the UV laser photons. Matricies labelled with an asterix have been shown to be effective matricies for MAI-MS (below) as well as MALDI-MS. Reproduced from Trimpin _et.al._[61]

In order to achieve effective MALDI generated analyte ions, effective mixing of the MALDI matrix ions and analyte molecules is essential, and so mixing of the two solutions either in a solution phase, or in a multiple-layer solution deposition process is usually used. The sample is "spotted" onto an inert surface (usually a stainless steel MALDI "target") and then allowed to dry, this can then be

covered using a solution of the MALDI matrix, which after drying will form the solid MALDI "spot", able to be ionised and analysed. The optimisation of MALDI sample preparation and creation of effective layers onto targets can be a long process and can yield varying results, but has been simplified by the use of vapour/spray deposition devices, which can automatically coat targets with matrix solution after manual transfer and drying of the analyte sample.

Unlike ESI, which is a continuous process in terms of ion generation, MALDI generates ions in pulses, which are dictated by the laser pulses used to activate the matrix. Thus MALDI ions plumes can be sampled individually, transferred into the mass analyser and provide mass spectra from a single ion cloud, without the waste of sample consumed/ions generated outside of the sampling period like in ESI/EI. The sampling of individual ion clouds via MALDI-MS has allowed for very low detection limits/high sensitivity experiments, often down to atto-moles, and even so-called "single-ion detection" experiments using MALDI-ToF-MS.

The main disadvantages of MALDI are a) MALDI, for the vast majority of cases, will only produce singly charged ions, which can limit the range of effective/optimal detection for most instruments, b) ions created are much "hotter" than those created by say ESI, meaning the ions have more energy acquired during the ionisation process and thus usually require more collisional cooling and/or will be harder to handle/stabilise in other parts of the MS instrument (e.g. in effective trapping within the ICR cell), and c) the activation of the matrix molecules causes ionisation of both matrix molecules and other species, producing a high number of non-analyte peaks in the resulting mass spectrum, often referred to as a high "matrix background". The matrix-associated peaks are generally singly charged peaks of low m/z (50-500m/z), and so can cause significant problems when attempting to analyse molecules at low m/z and/or small molecules.[60] Despite these disadvantages MALDI-MS is a thriving ionisation technique, which has considerable application in analysis of small molecules, complex mixtures (simplified by producing only singly charged species), biomolecules, and especially polymer samples, as the creation of singly charged ions at high m/z (3000-300,000 m/z, away from

the matrix background) is ideal for showing the varying length of polymer synthesis reaction products, and has been effective at calculating molecular weights, statistical averages for products (M_n and M_w), dispersities of products (PDI), and even functional end groups present on reaction products.[62]

1.6.1.6 Niche ionisation methods:

Other, less commonly used ionisation techniques have found application in the analysis of particularly specialised/difficult samples that may be more compatible with certain sample types than ESI, MALDI etc. (e.g. analysis of solid samples). These ionisation methods are not utilised for any analysis within this thesis and are included very briefly, not comprehensively, simply for reference.

Atmospheric-Pressure Photo Ionisation (APPI) is based upon spraying a solution phase sample, usually at a high flow rate (compared to ESI), through a small orifice, ESI like probe, with a nebulising gas, creating a Taylor cone and plume (as with ESI). The resulting plume is then exposed to a high energy (usually krypton) UV lamp, which excites analyte and solvent species into an electronically excited state, causing electron transfer and fragmentation reactions, and generating detectable ions.[19,63] The high energy lamp can be altered for the effective analysis of different compounds if needed, the krypton APPI lamps usually emit photons with an energy of ~10.6eV, which have been shown to effectively ionise a range of compounds, including hydrocarbon/petroleum mixtures.[19] APPI has found particular application for analysis of complex mixtures due to its ability to ionise both polar and non-polar species within petroleum/oil derived samples.[19] However the production of both protonated and radical species for many analytes can complicate spectra, and may cause some fragmentation of analytes/solvents, also complicating data analysis and hindering accurate assignment of components within the solution. The ionisation setup for APPI is shown below in Figure 1.35, next to the more commonly used atmospheric pressure Chemical Ionisation, discussed previously:

63
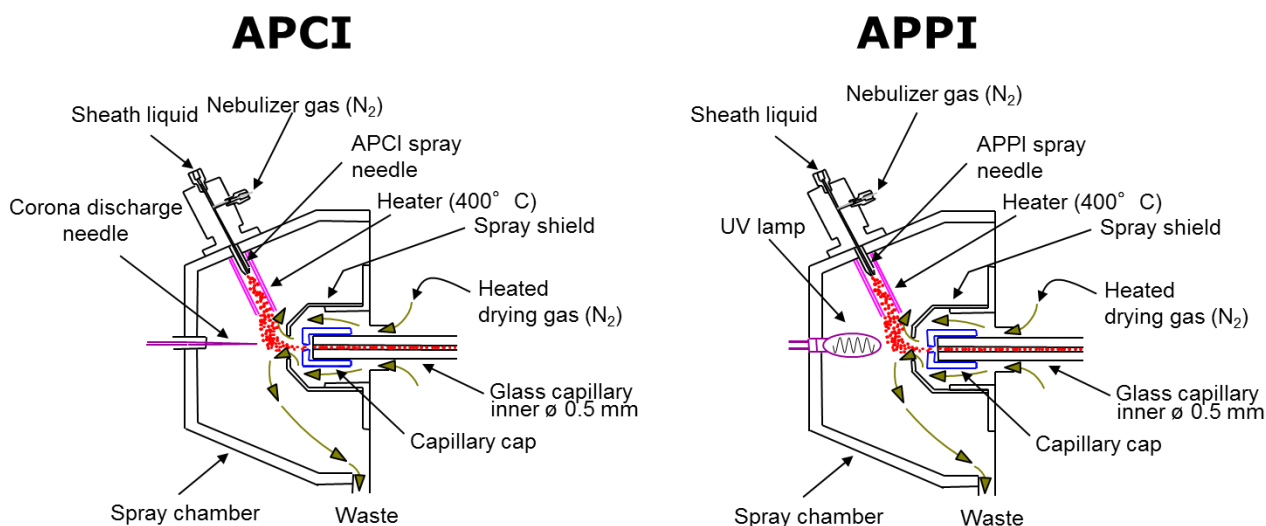
Figure 1.35: Ionisation sources for Atmospheric-Pressure Chemical Ionisation (APCI) and Atmospheric-Pressure Photo-Ionisation (APPI). Courtesy of Bruker Daltonics, Bremen, Germany.

Secondary Ion Mass Spectrometry (SIMS) is a long standing technique which utilises a ~150nm charged beam of ions impacting a surface to cause secondary ionisation of species within the target region, the once surface-bound ions can then be attracted to the MS instrument and detected, in a similar fashion to MALDI generated ions (above). SIMS is well known for its spatial resolution – the ability to ionise a very well defined and small area at a time, this inherent accuracy has allowed SIMS to be widely applied in the field of Mass Spectrometry Imaging (MSi). MSi involves the ionisation of many individual locations on a surface and the correlation of the resulting mass spectrum with each point, allowing a sampling of an entire region to provide "chemical maps" to show distribution of many species. SIMS with high spatial resolution has enabled MSi of sub-cellular components and analysis thereof.[64]

Matrix-Assisted Ionisation (MAI) is a recent addition to the atmospheric pressure ionisation techniques, and was again a fortunate discovery during routine MALDI experiments. MAI is a very similar process to MALDI (described above) but does not utilise a matrix-activating laser to ionise species, instead MAI relies on the matrix itself and the vacuum of the instrument inlet to ionise analytes of interest and pull them into the MS for analysis. This seemingly unlikely process has been

proven to not only be effective at ionising a range of analytes very easily, but also with reduced

sample preparation compared to MALDI-MS, and reduced tuning of source conditions (compared to

ESI and nESI-MS), making MAI a very attractive choice for routine/solid samples and non-specialised

personal. MAI has been shown to ionise small molecules, polymers, peptides, and even

proteins.[61,65,66,67] Most interestingly; unlike MALDI, from which the technique is derived, MAI can

easily achieve multiple charge states on peptides, proteins, and large polymers, which is in stark

contrast to MALDI's well-known 1+ charge state preference. The multiple charge states produced by

MAI are still a topic of debate, but suggest a distinctly different mechanism of ionisation to

traditional MALDI. It is currently believed that the matrix used for MAI must have a large

"triboelectric effect" where shattering analyte containing matrix crystals causes charge separation

with sparks and allows excess charge to be transferred to analytes, causing ionisation.

1.6.2 Movement of ions:

The range of ionisation sources available to contemporary mass spectrometry allows an extensive range of analytes to be transformed into detectable ions readily, from ESI and CI for small and/or non-polar molecules, ESI/nESI/MALDI for peptides, proteins, and polymers, to APCI and APPI for complex hydrocarbon/petroleum mixtures to name just a few examples of use, with many species ionisable via many different methods. However many of these sources operate at atmospheric pressure (ESI, nESI, APCI, APPI etc.), and many instrument configurations require the ionisation source to be a distance from the ICR cell/mass analyser in order to operate multiple pumping regions needed to achieve the ultra-high vacuum (UHV) needed for accurate mass measurement in an ICR cell (and other mass analysers).

1.6.2.1 Ion funnels and ion skimmers:

When ions are generated they tend to be created in large, distributed clouds, e.g. an ESI plume, or MALDI ion plume. These spatially distributed ions are difficult to transport and handle within a mass spectrometer, and ideally would be more confined and coherent so that they can be moved as tight ion packets. Older mass spectrometers utilised simple cone-shaped conductive metal orifices, known as skimmers. As ions exit confined space the ion plume expands rapidly, a process known as supersonic adiabatic exspansion,[68,69] these ions are thus disperse and again difficult to handle. Skimmers are designed to repel ions which deviate from the axis of ion transport and to only allow the centre of the expanded plume to pass through the small (~≤1mm) orifice. The voltage applied to the skimmer repels ions outside of this centre and increases the energy of those passing through. Thus the Skimmer should be considered as a sampling ion optic rather than a confining one. Skimmers were effective at producing confined ion clouds and were also useful in reducing the orifice between pumping regions (allowing separation and low vacuum pressure to be reached).[70] Increasing skimmer voltages was also an effective means of causing ion-skimmer collisions, which were shown to be effective at causing desolvation of large, solvated ESI ions, and clusters, which did

not produce individual ions as desired during the ionisation process.[71] In order to advance towards this Smith *et. al.*[72] developed the ion funnel; an arrangement of capacitively-coupled metal plates/rings held at alternating voltages to attract and direct ions from a wide starting area and towards a confined exit (hence funnel). Unlike the skimmers, ion funnels were focusing optics, not sampling optics, retaining more ions than traditional optics. The ion funnel development was able to improve the transmission of generated ions ~10 fold, and thus boost the sensitivity of mass spectrometers.[72] Although Ion funnels are effective at focusing ions, they are not suited to pure transportation as they do not attract ions towards the target orifice. Ions are usually accelerated through the ion funnel barrier via use of a repulsive plate/voltage on-axis with the funnel plates.
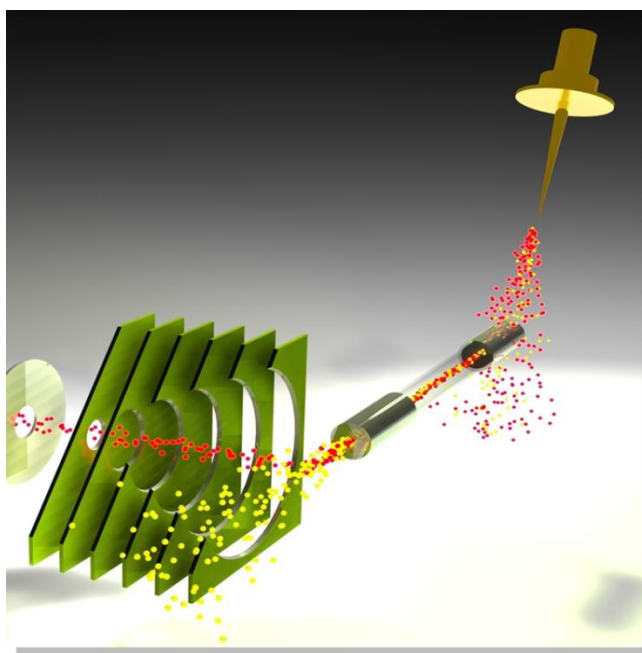
Figure 1.36: representation of an orthogonal ESI-capillary-ion funnel source for MS. The figure shows that although both neutral (yellow) and ionised (red) species travel through the heated capillary, even after desolvation, the orthogonal injection geometry ensures that neutral species are diverted away from the directing ion funnel plates and allows only the ions to be collected/directed further into the mass spectrometer for detection, greatly reducing background noise and/or interference from neutral molecules. Courtesy of Bruker Daltonics, Bremen, Germany.

1.6.2.2 Multipoles:

Unfortunately, although skimmers and ion funnels can attract and focus ions into other regions inside a mass spectrometer, they are not adept at transporting or storing generated ions, nor filtering them for analysis. Instead charged, alternating-polarity rods are used as effective "ion pipes" to confine ions radially and guide them through vacuum regions inside mass spectrometers. Taking a quadrupole (4 charged rods) as an example; by attaining the same polarity as the ions in question the rods can repel ions away from two "sides" of the ion guide and prevent radial ejection, the polarity is then switched a moment later to repel the ions from the other two sides and attract them in the opposite direction. Thus the ions are constantly repelled and attracted, but kept to within the centre of the guide and are effectively radially confined. The frequency of the polarity switching and the amplitude (often given in terms of voltage peak-to-peak (V_{p-p})) used can be tuned to stabilise the transmission of different m/z ranges, to achieve the desired range for analysis.

The stabilisation of ion transmission through multipoles is dictated and controlled via the Mathieu stability diagram, below, and the desired range of transmission can be effectively defined by tuning/varying values of **a** and **q** as shown in Figure 1.36:
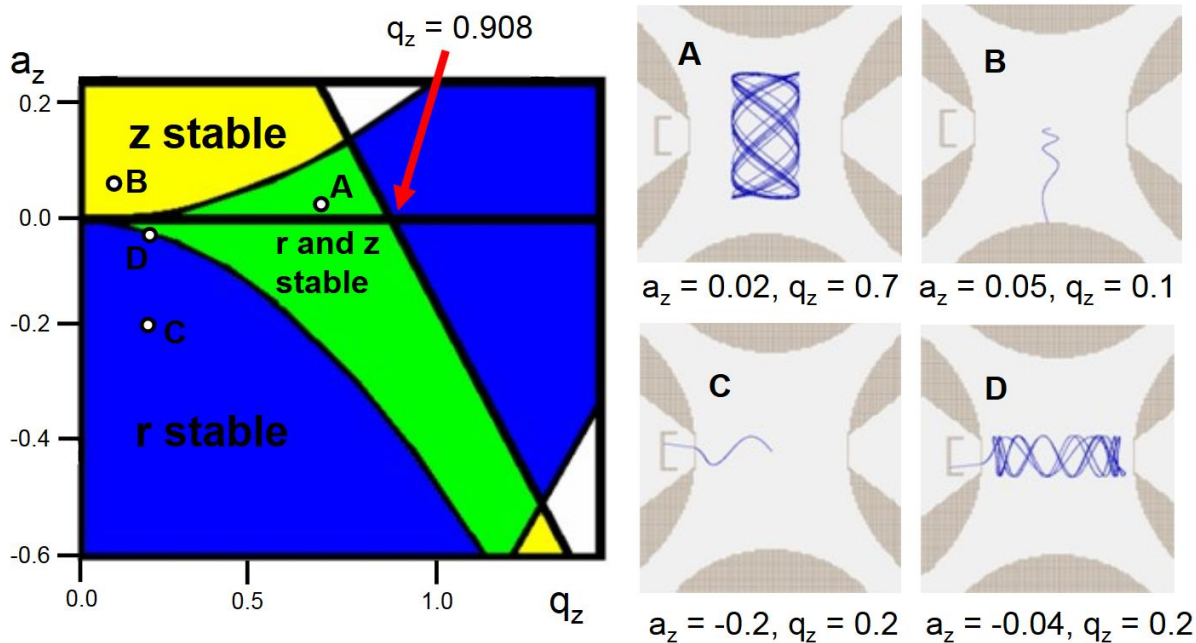
Figure 1.37: Mathieu stability diagram for a given ion in a quadrupole (left) and typical ion trajectories for the different labelled areas (right).

Evidently achieving transmission of all desirable ions (above a certain low mass cut off – q_z=0.908) seems possible by simply setting the value of V_{dc} to 0, and then using different values of **q** to alter the range. However, should one want to only transmit a smaller range of ions then the user can simply set the centre of the stable region to the peaks/range of interest (using **q**) and then increase the value of **a** to narrow the stability region enough to achieve only the desired transmission range.

The use of m/z specific stability regions is the basis of quadrupole mass filtering, and thus the quadrupole mass analyser, simply, a certain m/z scan range is defined (e.g. 50-500 *m/z*) then the values of **a** and **u** are tuned to stabilise the first value in the scan range (50 *m/z* in the example), ions are allowed to transmit through the multipole for a set amount of time (usually microseconds to milliseconds), the transmitted ions are then detected via a charge detector such as an electron multiplier device, then the next *m/z* value (51 *m/z*) is stabilised and the process is repeated for the entire range of interest. Once the entire range has been sampled at each desired m/z value the mass

spectrum can be produced by simply correlating the magnitude of current detected at each time to the correct m/z stabilised at that time, thus creating a mass spectrum of counts/intensity vs. *m/z*.

Stability areas (upper part) for ions with different masses (m₁< m₂ < m₃) as a function of U and V.

The red line representing the intersection of the scan function with the stability diagram is directly related to spectral peak width. If for example, the slope of the scan line is reduced, this line will lengthen – the spectral peak width will increase and spectral resolution will decrease.

The area under the curve represented here is analogous to spectral peak area. If the slope of the scan line is decreased, for example, the spectral peak area will increase and instrument sensitivity will increase.

Figure 1.38: Basis of quadrupole analyser/transport operation, by altering the voltages and frequencies applied to the quadrupole the region of stability changes, and thus the species able to traverse the entirety of the optic also changes. Once carefully calibrated the ions of a certain m/z can be transmitted with voltages corresponding to their m/z and thus if they successfully move through the optic, the m/z can be deduced. Reproduced from Chrome-academy (http://www.chromacademy.com/lms/sco36/Fundamental_LC-MS_Mass_Analysers.pdf, accessed 28/03/2016).

Beyond quadrupoles, hexapoles and octupoles have also been used for ion transport and filtration, however it has been shown that although possible hexapoles and octapoles are not as efficient at

filtering ions as quadrupoles and are more suited to purely ion transport. Higher order multipoles have been shown to be more effective at pure ion transmission than quadrupoles due to a more distinct energy barrier as ions approach the poles' electric confinement region.

Multipoles for long distance ion transmission, e.g. into the bore of a large superconducting magnet to an ICR cell, were shown to be a critical improvement over previous electrostatic designs which simply relied on acceleration of ions, "steering" of the ion cloud using static lenses and then "catching" of the ions at the other end (the ICR cell), as the optics were notoriously difficult to tune and accelerated ions to such high energy that upon capture they still retained a significant energy and were still "hot", which hinders formation of coherent ions packets, and thus ICR performance.

1.6.3 In cell isolations/manipulations:

Before FT-ICR MS was effectively interfaced with quadrupole and multipole technology mass spectrometrist's were still able to conduct isolation and ion manipulation experiments within the ICR cell due to the ability to use excitation fields within the trap. The ICR cell adds an enormous amount of flexibility to mass spectrometry analysis as it enables the control over ion packets and the ability to manipulate them in a variety of ways.

In-cell isolations were conceived from the same methodology as excitation/detection pulses, if ions are excited too much they will hit the ICR cell plates, become neutralised and will then be pumped out of the cell as neutral gas molecules. As a result one can simply excite unwanted ions to an extreme degree, forcing their ejection from the ICR cell, ions of interest can then be excited to a lower (normal) radius and be detected as usual.

1.6.3.1 Sweep, shots, and resonant ejections:

The application of a large/extreme pulse which is resonant to unwanted ions will result in ejection of unwanted ions (as described above). The application of these pulses is usually presented in two ways; either a broadband excitation of every undesired species with a small window/notch centred on the ion of interest (i.e. the species to remain post-excitation), this excitation profile will excite a broad range of unwanted ions and ideally leave the ion(s) of interest, and is known as a sweep excitation. The inverse of this procedure is also very common; the application of a small, define, pulse centred on an unwanted ion/multiple ions, causing excitation of small windows/notches within the mass spectrum, known as a "shot" excitation, multiple shots are easily programmable and applied to an individual ion population, the only limiting factor being the overall time, as each shot excitation waveform is applied individually, so 10 shots will take ~10x the time as one shot ejection.

The main disadvantages of using sweep/shots ejections in an ICR cell are the same as previously mentioned for a standard chirp broadband excitation procedure; a non-ideal excitation waveform.

The sweep and shots ejections use the same profile as the chirp excitation, meaning the beginning and end of the profile is distorted and the central section is non-linear, although the ejection profile is effective at exciting the species of interest, it may also eject species close to the boundary edges. In a usual broadband excitation slight deviations in excitation profile may have little effect if they are a distance from ions of interest, however isolation experiments usually require defined isolations and the ability to separate species which are close in *m/z* value, meaning a well-define excitation pulse is needed. As a result resonant Sweep and Shots isolations are used usually for ions which are separated from other species by a reasonably large *m/z* range (for an isolation, e.g. 10-20 *m/z*) and so will not be affected by the "horns" distorting the excitation pulse.

1.6.3.2 SWIFT:

As mentioned previously SWIFT can be used as an accurate means of exciting ions of interest using an excitation profile that is better defined and closer to the ideal rectangular excitation desired (Figure 16 above). Though SWIFT is an excellent tool for the excitation of ions for detection,[73] it has also been shown to be very effective at high-resolution isolations inside the ICR cell.[74–76] Back calculating a SWIFT pulse from a desired ejection profile is a more accurate means of creating a notched isolation pulse, either leaving the ion(s) of interest by ejecting everything else from the ICR cell, or by ejecting individual species by creating a smaller ejection region containing the unwanted ions. Thus SWIFT can be used to replace both sweep and shots resonant ejections. Due to the more ideal ejection pulse, SWIFT has been shown to achieve isolation resolving powers in excess of 29,000-39,000.[25]

In addition to SWIFT other In-Cell Isolations (ICI's) have been developed to provide targeted/multi-stage Isolation experiments for mass spectrometry. Correlated Harmonic Excitation Fields (CHEF) was a method developed by de Koning *et. al.* in 1997.[77] CHEF (single isolation) and Multi-CHEF (multiple CHEF isolation events) ejection of ions are achieved by on resonant excitation of unwanted species and off-resonant excitations of the species of interest for controlled (correlated) amounts to avoid the target gaining energy through off-resonant excitations and being subsequently ejected from the ICR cell, resulting in ion loss at smaller isolation window sizes (a key limiting factor in quadrupole isolations). Inadvertent energy gain during in cell isolations can be high enough to trigger ion-molecule reactions between the isolated ion and background gas in the ICR cell.[77] Single/sweep excitation profiles, even if tailored to avoid excitation of certain (target) species, will cause unwanted excitation of ions. CHEF avoids this unwanted excitation by calculating the unwanted excitation of the target species and adjusting the $T_{excitation}$ and/or the excitation field magnitude ($/V_{p-p}$) to excite and de-excite the retained species an integer number of times. As a result the desired ions will be excited and de-excited an even number of times and thus be returned to the centre of the ICR cell after the full excitation (/isolation) pulse. For convenience either $T_{excitation}$ or excitation field magnitude can be set and the other varied for the intended target species. CHEF Isolation pulses can be applied for many species, modifying the waveform used with the addition of each target. Examples of these waveforms and the resulting pulses are shown below in Figure 1.39:
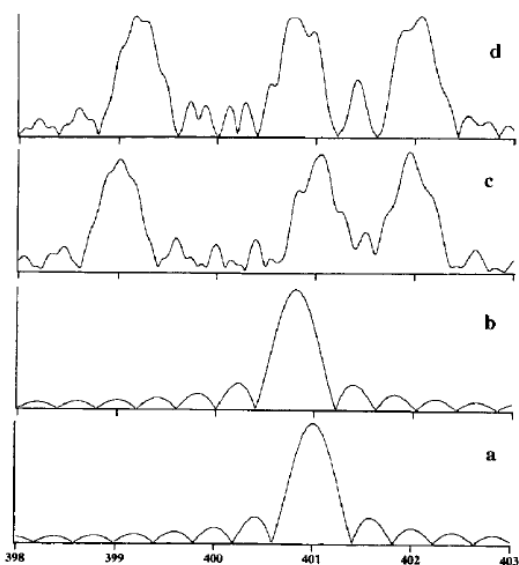
Fig. 1. Computed excitation waveforms corresponding with (multiple) single-frequency excitation fields in a 4.7 T magnetic field. (a) Single-frequency excitation field for the ejection of ions with m/z 401, with T_{exc} of 5500 μs. (b) Correlated single-frequency excitation field for the ejection of ions with m/z 401, with T_{exc} of 5500 μs, for the selection of ions with m/z 400.00000. (c) Multiple consecutive single-frequency excitation fields for the ejection of ions with m/z 399, 401 and 402, with T_{exc} of 5500 μs and total excitation duration of 16 500 μs. (d) Correlated multiple consecutive single-frequency excitation fields for the ejection of ions with m/z 399, 401 and 402, with T_{exc} of 5500 μs and total excitation duration of 16 500 μs, for the selection of ions with m/z 400.00000 (see Table 1).
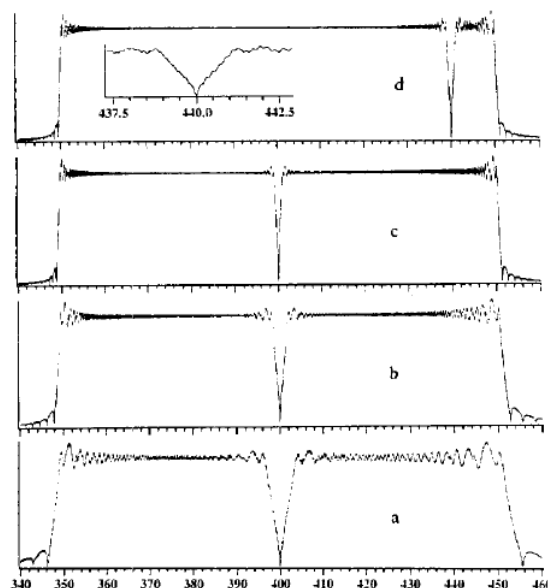
Fig. 3. Computed excitation waveforms corresponding with correlated swept-frequency excitation fields in a 4.7 T magnetic field over a bandwidth from m/z 350 to 450. (a) Selection of ions with m/z 400.00000 with T_{exc} of 500 μs. (b) Selection of ions with m/z 400.00000 with T_{exc} of 1000 μs. (c) Selection of ions with m/z 400.00000 with T_{exc} of 2000 μs. (d) Selection of ions with m/z 440.00000 with T_{exc} of 2,000 μs (see Table 2).

Figure 1.39: Examples of CHEF waveforms (left) and CHEF pulses (right) for the isolation/ejection of species from an ICR cell using CHEF. Reproduced from de Koning *et. al.*[77]

The advantages of Multi-CHEF are centred around its ability to eject/isolate a varying number of species (by application of more and more modifications to the CHEF pulse), the extraordinary isolation resolving power (>60,000, *c.f.* SWIFT = ~29,000), and the zero ion-loss during isolation, which follows through from theory to application. The only two disadvantages to Multi-CHEF ICI are that the isolation of many species/high resolution ICI requires the use of long CHEF pulses, which increases duty cycle and can result in ion loss of higher pressure FT-ICR MS instruments (though this shouldn't be an issue for contemporary FT-ICR MS). CHEF pulses can also slightly alter the ion position of isolated species retained within the ICR cell post-activation, though not documented this effect can alter tuning parameters needed for particularly sensitive and difficult species during in-cell MS/MS (see Chapters 3 and 6 for particularly challenging examples). Regardless Multi-CHEF is a very

75

versatile and often under-used technique widely available on commercial and bespoke FT-ICR MS systems, boasting ultra-high resolving power isolations (e.g. shown below in Figure 1.40), and zero ion loss,[77] Multi CHEF can even be used to isolate individual isotopes of species, which is especially useful for overlapping isotopic distributions, an unfortunately common occurrence in MS and MS/MS experiments.
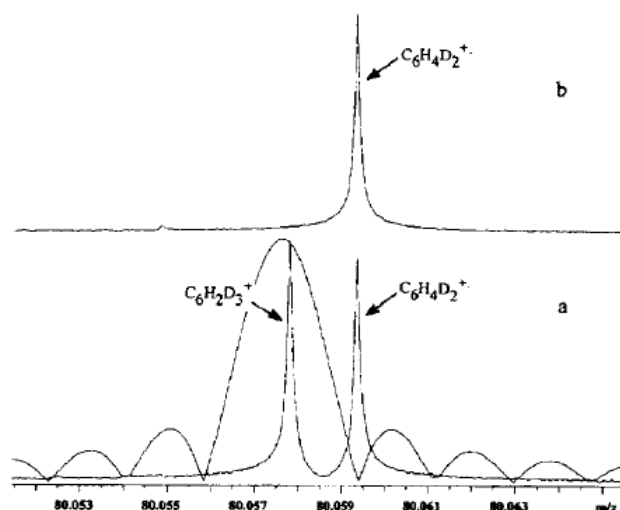
Fig. 2. Ejection of $C_6H_2D_3^+$ ions in the presence of $C_6H_4D_2^{+\cdot}$ ions. (a) FT-ICR mass spectrum with the signals for the $C_6H_2D_3^+$ and $C_6H_4D_2^{+\cdot}$ ions, with a mass resolution of about 500 000. Superimposed the recorded waveform is shown, corresponding to the actual correlated excitation field transmitted to the FT-ICR cell to eject the $C_6H_2D_3^+$ ions, with T_{exc} of 50 000 μs and a field amplitude of 0.43 $V_{p \cdot p}$. (b) FT-ICR mass spectrum recorded after ejection of the $C_6H_2D_3^+$ ions.

Figure 1.40: Example of UHR-CHEF isolation performed on particularly close peaks separated by just 1.5mDa. Reproduced from de Koning *et. al.*[77]

1.7 Tandem Mass Spectrometry (MS/MS):

As shown above in many cases Mass Spectrometry is a powerful tool for the analysis of both simple and complex mixtures, providing mass information on analytes present, and able to perform interesting ion manipulation experiments, all to a very high degree of accuracy and performance. However this analysis is, for the most part (and for all soft-ionisation analysis), based on the overall mass of a particular analyte and although fine structure analysis has been extremely useful in providing elemental compositions for (mostly small molecule) species, the analysis of larger, particularly biomolecule species is hindered by the different combinations of elements possible as mass increases, and the overall mass providing less and less chemical information as the size/mass of the analyte increases. However as shown above, techniques such as Electron Impact ionisation (EI) have been able to provide both overall mass and structural information for decades based upon the fragmentation of analytes produced, with the fragments providing insights into functional groups present, branching structures, and stability of remaining fragments.[78] However fragmentation of analytes during ionisation limits MS analysis to pure/pre-separated compounds in order to avoid overlapping fragmentation peaks produced by different analytes. To this end selected fragmentation of species separated inside the mass spectrometer (using multipole filtration and/or in-cell-isolations, both described above), followed by individual analysis of the parent ion (species isolated) and fragment ions has become widely applied and known as tandem-mass spectrometry (MS/MS).

The standard MS/MS experiment is based upon the observation of ions during a normal MS experiment, followed by selected isolation of a/many target compound(s) by a mass filtering multipole (usually a quadrupole), after the species of interest has been successfully isolated it can be selectively fragmented using a range of techniques (discussed below) to produce an MS/MS spectrum, analysis of which can provide detailed chemical and structural information of the isolated ion. Following this the next species of interest can then be isolated and the process can be repeated

until all observable species have been fully characterised by both MS and MS/MS techniques. This process is summarised in Figure 1.41, below:
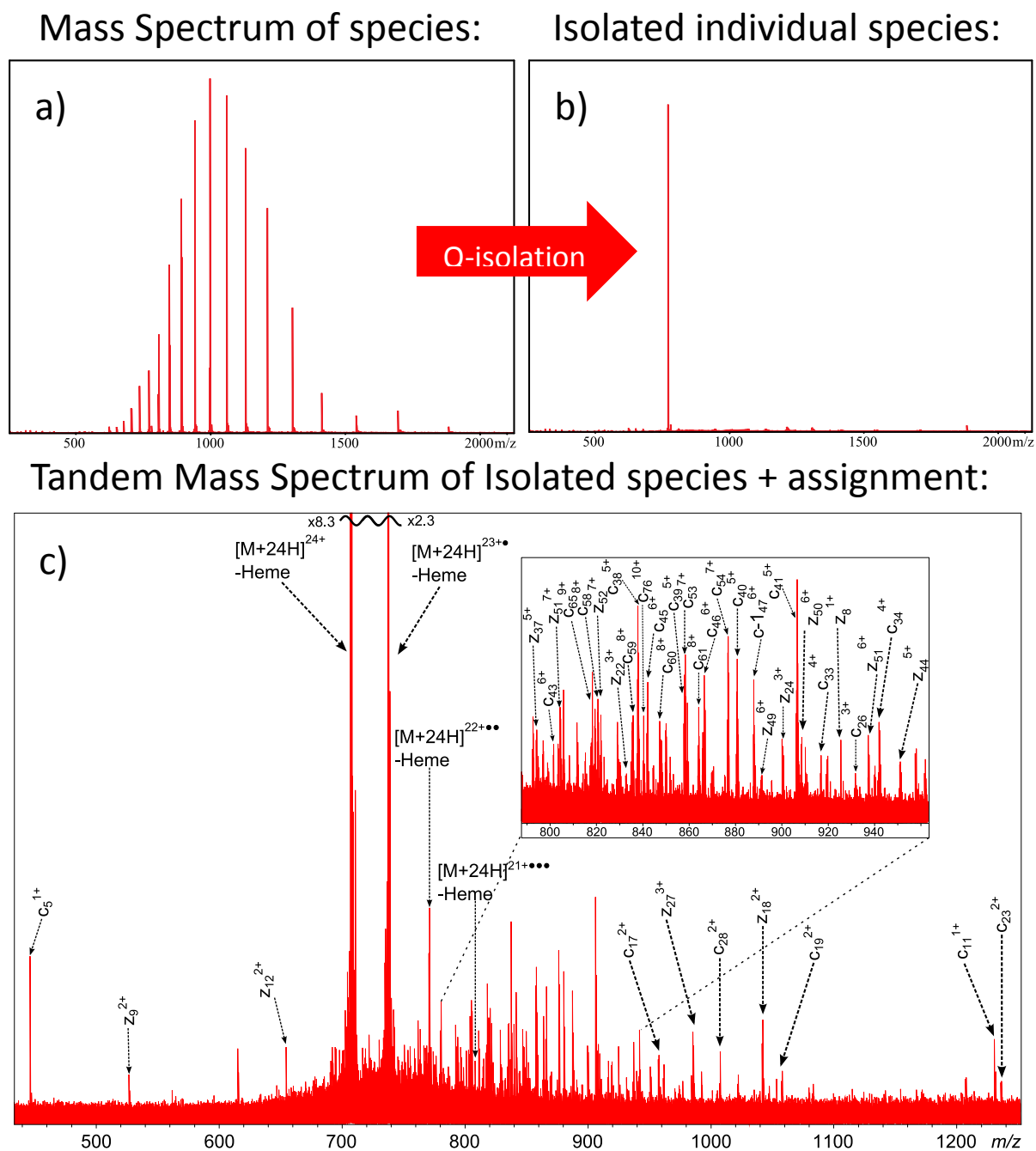
Figure 1.41: Standard acquisitions taken during an MS/MS experiment. The full mass spectrum of all species observed in a sample (a) followed by isolation of a particular species of interest (e.g. via a quadrupole mass filter) (b), and finally fragmentation of the isolated species using an MS/MS technique – electron capture dissociation (c) subsequent analysis of fragments reveals information

of species structure and composition. Examples shown are from Myoglobin protein (16.9 kDa) spectra.

MS/MS techniques vary greatly and can be based on photons, electrons, and collisions, to name a few, applications, specifically to biomolecule characterisation, have made a huge impact on the field and allowed detailed characterisation of countless compounds. Many different MS/MS techniques are discussed below, mainly with a focus on biomolecule characterisation to inform results recorded later in this thesis, though many techniques have been shown to be very effective for characterising a range of molecular species.

1.7.1 Collisionally Activated Dissociation / Collisionally Induced Dissociation (CAD/CID):

The discovery of Collosionally Activated Dissociation (CAD), also known as Collisionally Induced Dissociation (CID) was accidental, but opened up the entire field of tandem mass spectrometry. CAD was first achieved by Keith Jennings in 1968, and was first noticed during routine scans of small molecules, Jennings noticed that there were additional peaks outside the target compound, eventually finding that these were caused due to collisions induced by background gas which was present inside the MS from a leaking vacuum seal. From this serendipitous discovery and eventual application, CAD can now be induced voluntarily by intentionally colliding species of interest with background gas molecules and causing fragmentation, this is usually achieved by using an opposite polarity voltage to accelerate the ions into a stationary cloud of "collision gas", which is usually an inert gas such as Argon/Neon/He/N_2. Forcing ions of interest to collide with neutral gas molecules will cause and energy transfer and "heat up" the ion of interest (and the gas molecule, but these are continuously added/pumped away). A single collision is rarely enough to cause any dissociation, unless extreme voltages are used, however each ion will undergo many collisions as it enters and passes through a collision cell (trapping multipole), with each collision increasing the internal (mostly vibrational) energy of the ion slightly more, causing a "slow heating" of ions. Eventually the deposited energy will exceed the bond strength of a bond within the analyte and the species will

dissociate, causing fragments, the contents of the collision cell, both parent ions and fragment ions can then be transferred to the mass analyser, whatever configuration that may be, and be detected, providing an MS/MS spectrum (Figure 1.42 below). Energy transferred to the ion of interest can be calculated using equation 13 below:

$$Maximum\ possible\ energy\ transfer = \frac{m_{neutral}}{m_{neutral}+m_{ion}} * E_{lab} \qquad \text{eqn. 13}$$

Where $m_{neutral}$ is the mass of the neutral species used as a collision gas, m_{ion} is the mass of the charged species being accelerated and dissociated, and E_{lab} is the energy used to accelerate the ions into the collision gas. Each CAD collision typically deposits <0.1eV of energy into the accelerated species, making it a very slow-heating, multiple collision method to dissociate ions.
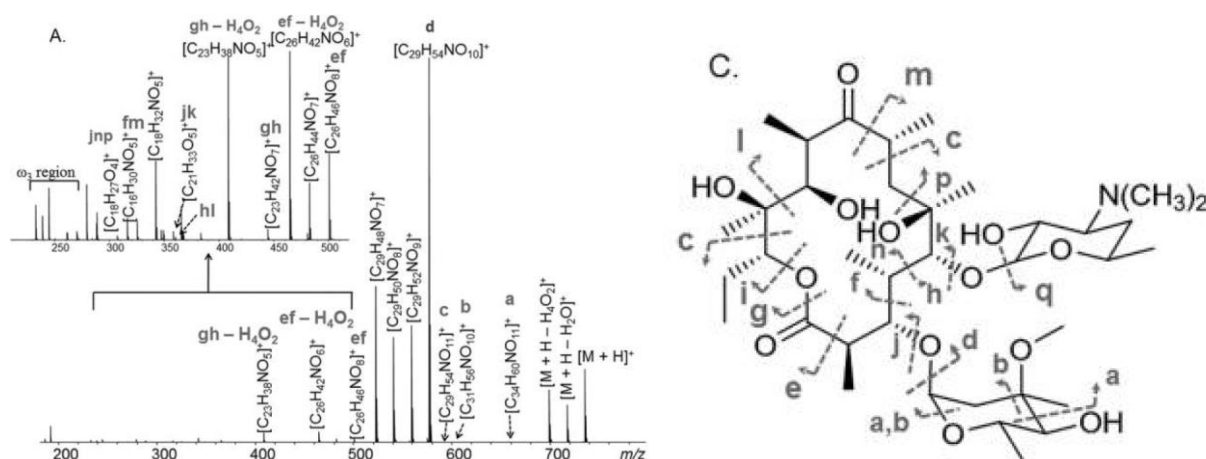
Figure 1.42: CAD MS/MS spectrum of Erythromycin A (a), and the assigned fragments mapped onto the molecular structure (c). Reproduced from Wills _et.al._[79]

As mentioned, CAD slowly adds energy to an ion, over the course of many, many collisions, each depositing a small amount of vibrational/kinetic energy into the ion, eventually causing bond dissociation. Due to this many collision/slow heating mechanism, the ions are brought up to a dissociation energy, and the bond to break in the molecule will be the lowest energy bond, as that will be the energy reached first during the slow heating process, known as an ergodic dissociation.[80] Because of this ergodic mechanism, the CAD fragmentation of ions is not random, but is dictated by

bond energetics/stability and can a) provide information of bond strengths within the system b) be predicted based on bond energetics. For small molecule CAD MS/MS the fragments produced differ greatly from the radical-directed fragmentation caused by EI ionisation (above), but can still offer a great deal of information dictated by functional groups etc. and has been shown to be an effective tool for both singly charged and multiply charged small molecule analytes. [81–83] Small molecule CAD MS/MS has also been shown to differ greatly depending on the charge-carrying cation adducted to ESI generated small molecules, protonated species fragment very differently to sodiated, lithiated, lotassiated, and cesiated species, which can form different fragmentation patterns and provide complementary (different but supportive) information about the target species.[50]

CAD can also be achieved within an ICR cell, and relies on the acceleration of ions via excitation to collide with gas molecules pulsed into the ICR cell. Unfortunately in-cell CAD MS/MS raises the pressure inside the ICR cell dramatically and requires a pumping delay to reduce this pressure before excitation and detection.

1.7.2 MS/MS of peptides and proteins:

Biomolecule CAD MS/MS was a huge step forward in biomolecule characterisation and is currently the most widely used fragmentation technique for the characterisation of peptides, proteins, DNA, RNA, metabolites, and everything in-between.[82,83] Unlike small molecules the majority of biomolecules are polymeric species (biopolymers) and so are constructed from repeating units, often with varying functional groups, but with a consistent/repeating chain/backbone. For peptides and proteins these repeating units are amino acids, naturally zwitterionic species that have a core amine-carbon-carboxylic acid structure, differentiated only by the carbon-linked side chain group (known as the residue/side chain). The individual amino acids link/bond via condensation reactions to form amide linked biopolymers, i.e. peptides and proteins, which can then be further expanded upon/modified/bound to other species post-transcription within a cell. During MS/MS of peptides and proteins the most commonly observed fragmentations occur along the amide-linked backbone, with CAD MS/MS causing fragmentation at the amide bond, the weakest bond available due to the

ergodic fragmentation mechanism. These amide bond dissociations cause corresponding fragments,

known as b or y ions (for an amide bond cleavage), with different bond cleavage points producing

different fragment ions. The nomenclature for peptide/protein MS/MS ions/fragmentation has been

changed iteratively over the years, however most contemporary MS/MS groups use the Roepstorff

/Zubarev nomenclature, summarised below:
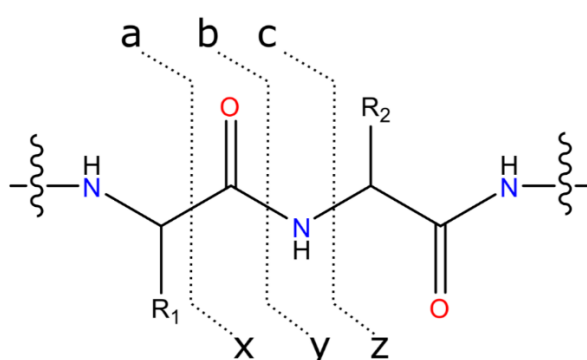
Figure 1.43: Summary of the Roepstorff nomenclature for peptide/protein MS/MS cleavage of

backbone bonds. Different MS/MS techniques can cause cleavage of different bonds (see below).

The nomenclature is based upon which bonds are cleaved during the poly(aminoacid) MS/MS

fragmentation and which part of the molecule resides in the remaining fragment. For instance

cleavage of the C=carbonyl carbon to nitrogen bond in Figure 1.43 will produce a b-ion towards the

N-terminus and a y-ion towards the C-terminus, though both may not be observed in the resulting

MS/MS spectrum if they do not possess a charge. If the bond mentioned was cleaved between the

2nd and 3rd amino acid in a tri-peptide (e.g. R-G-K) then this would produce a b_2-ion as the bond was

cleaved at amino acid position 2 (numbering from the left for a/b/c ions), and is a b-type ion. The

other side of this fragmentation would produce a y_1-fragment ion, as the bond was cleaved at

position 1 (numbering from the right for x/y/z ions) and is a y-type ion. This nomenclature can thus

be easily extended to any size of poly(amino acid) species by simply numbering from the two termini

and deciphering which fragments have been generated. The fragments can then be summarised in a

fragmentation map.
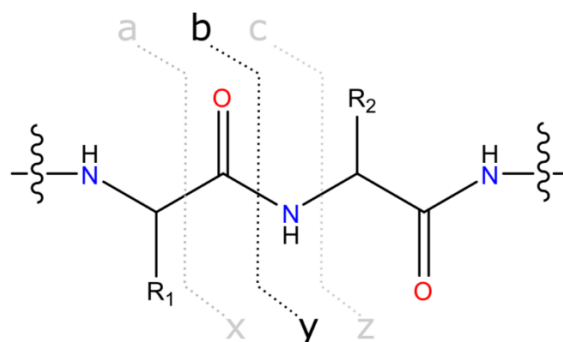
Figure 1.44: representation of peptide backbone showing the bonds cleaved during CAD MS/MS and the type of fragments produced (highlighted).

The main fragmentation pathway for CAD MS/MS of poly(amino acids) is to form b/y ions, however these fragmentations/ions can also cause small neutral molecules losses (such as loss of H_2O, CO_2, CO, NH_3, etc.).[83] Extra small molecule/side chain loss pathways increase the complexity of MS/MS spectra, though can provide limited information on the amino acids still present within the fragment (as only certain amino acids produce each type of neutral loss). CAD MS/MS can also cause fragmentation not just in one location in a peptide/protein, but subsequent collisions after the initial fragmentation can cause fragmentation elsewhere, e.g. further down the poly(aminoacid) backbone, known as secondary/internal fragments[83]), these fragments can be very difficult to assign accurately as the combinations of amino acids lost from either end of a chain to form an internal fragment increase dramatically with increasing chain length, thus often are not assigned and become non-sequence informative fragments, which only serve to increase spectral complexity and reduce ion signal in pathways that are more useful for sequence analysis.[83]
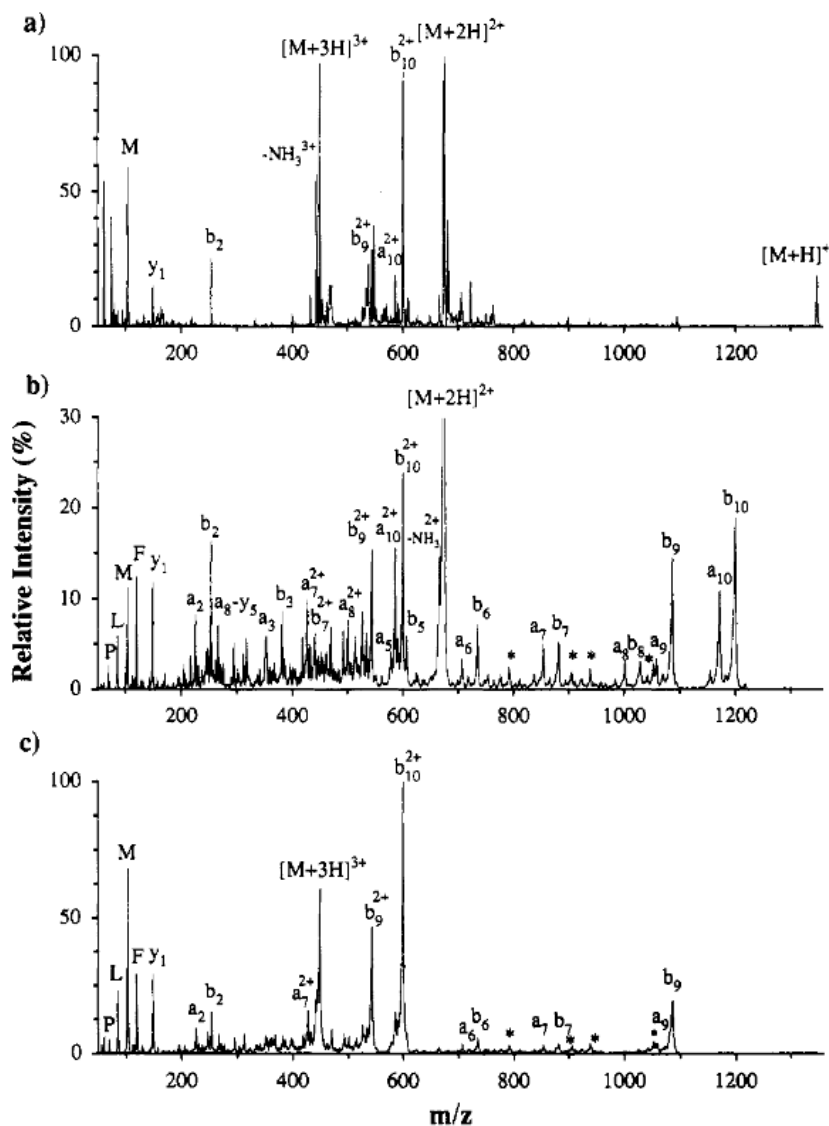
Figure 2. Ion spray mass spectrum of substance P (RPKPQQFFGLM) and fragment ion spectra produced by CID of different protonated forms. Peaks labeled with asterisks are second-generation fragments derived from $(b_{10}')^{2+}$ and $(b_e')^{2+}$ intermediates (see Figure 3): (a) ion spray mass spectrum; (b) doubly-protonated precursor, $(M + 2H)^{2+}$; (c) triply-protonated precursor, $(M + 3H)^{3+}$.

Figure 1.45: Example CAD MS/MS spectra of the peptide Substance P (RPKPQQFFGLM) in the 1+ (a) 2+ (b) and 3+ (c) charge states, showing the range of ions types detected and the variation of the observable fragments with different charge states. Reproduced from Boyd et.al.[84]

Example CAD MS/MS spectra for the peptide Substance P (SubP) in the 1+, 2+, and 3+ charge states ($[SubP+H]^+$, $[SubP+2H]^{2+}$, $[SubP+3H]^{3+}$) are shown above in Figure 1.45 from Boyd *et.al.*[84] using a ToF system. b/y fragments and many neutral losses, increase of the collision voltage used would result

in many secondary/internal fragments and possibly sequence scrambling. Sequence scrambling is a phenomenon in MS/MS where the usual sequence of amino acids/monomers within a certain peptide/protein/biomolecule re-arrange after fragmentation to form more thermodynamically stable structures, but consequently form species with a different monomer sequence than before. Unfortunately this is a purely detrimental effect on MS/MS quality and deleterious to data interpretation, and has no analytical use. However, the origin of sequence scrambling is usually the ions' attempt to stabilise an otherwise less stable structure (such as a given b/y ion) by rearranging into a more stable structure, where the charge/dissociated bond is on a group more able to stabilise the charge due to the electron negativity of the groups around it, e.g. a positive charge resting on a benzyl or tertiary carbon instead of a primary carbon/benzene ring. Sequence scrambling can have a seriously negative effect on efforts to characterise a biomolecule purely from MS/MS data (so-called "*de novo* sequencing"), as it produces false positives for other sections of sequence. Sequence scrambling can also impede efforts to pattern match/assign spectra automatically by including extra non-sequence peaks which will not match with the theoretical MS/MS spectra for a given sequence. Sequence scrambling, internal fragments, and excessive side-chain losses can usually be minimised/eliminated by careful instrument tuning (specifically the voltage/energy used for CAD) and delicate handling of ions within the MS during MS/MS and detection.
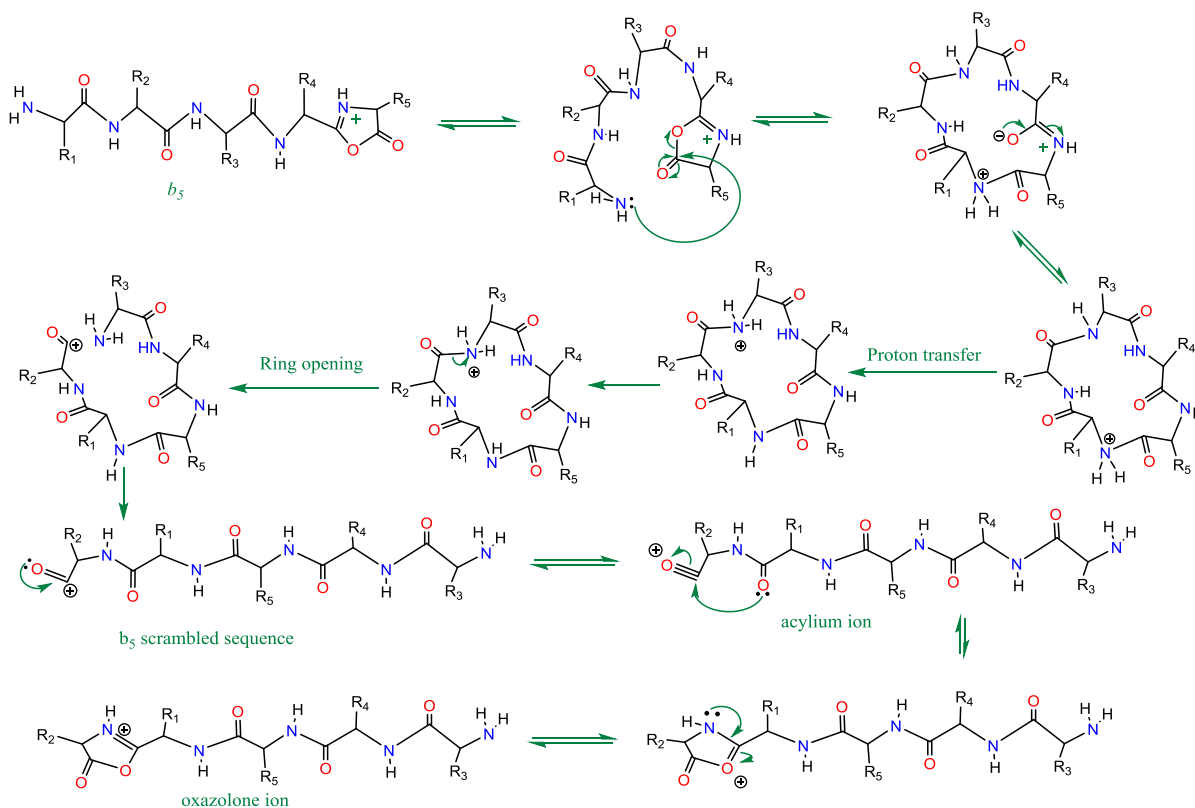
Figure 1.46: subsequent reactions to b-ion fragments produced during CAD MS/MS which can cause rearrangement (scrambling) of various functional groups within a previously linear peptide/protein structure, most importantly the scrambling of the original amino acid sequence.[85]

CAD MS/MS has been very effective at sequencing standard peptides, and has had limited success at large peptides/small proteins, but tends to offer limited sequence/cleavage coverage across the backbone sequence, with fragmentation often limited to the termini regions.[86] Despite this CAD MS/MS is very fast (little increase to duty cycle/scan speed of instrument), and can be used on both singly charged and multiply charged species, it has been very effective at identifying unknown peptides in biological samples by pattern/sequence matching them against the predicted proteome generated from the genome of the host/source organism.[83]

Unfortunately CAD MS/MS has also been known to dissociate many post-translational modifications (PTM's) added after the transcription from the host genome, meaning that the key differences between modified and unmodified species are often lost when using ergodic fragmentations such as

CAD MS/MS, often due to many modifications inherently low bond strength compared to the peptide backbone.[87,88]

1.7.3 In-Source Dissociation (ISD)/Nozzle-Skimmer Dissociation:

In Source Dissociation is another form of CAD which can be implemented on almost any modern mass spectrometer and causes dissociations by using high skimmer voltages to induce collisions between the incoming ions and the skimmer optic.[71] ISD has been effectively used to "pre-heat" ions before other types of MS/MS (see AI-MS/MS later), but its main use has come in the form of assisting in the desolvation/de-clustering of species as they enter the mass spectrometer.[89,71] At high concentrations many solvents/salt/small molecules (e.g. ACN/CsI/arginine respectively) can form clusters and create high m/z species (hundreds up to thousands of m/z) and cause unexpected peaks in the resulting spectrum (though these can be used to calibrate MS, see calibration below). ISD has been shown to be an effective way of breaking up clusters and buffer/solvent adducts on large molecules/proteins, by inducing low energy cleavages that will not break peptide/protein backbone bonds or small molecule bonds, but can dissociate non-covalent interactions and buffer adduction. ISD has been shown to be particularly useful for the analysis of protein complexes as a method of dissociating the native complex before subsequent MS/MS of each subunit.[90,91]

1.7.4 Infra-Red Multi-Photon Dissociation (IRMPD):

Use of laser-generated photons has proven to be another useful way of transferring energy to analytes of interest, eventually causing dissociation for MS/MS analysis. 10.6um, i.e. Infra-Red photons have been shown to be absorbed effectively by both peptide linkage bonds, and a variety of small molecules, and like CAD is a multi-stage process, thus Infra-Red Multi-Photon Dissociation (IRMPD) was developed.[92] IRMPD usually uses a 10.6um continuous-wave CO_2 Laser to irradiate trapped ions of interest, usually in an ICR cell, though examples have been shown of use in Ion Trap mass spectrometers.[93,94] IRMPD of ions inside ICR cells usually occurs after trapping/cooling stages and just before excitation/detection occurs.
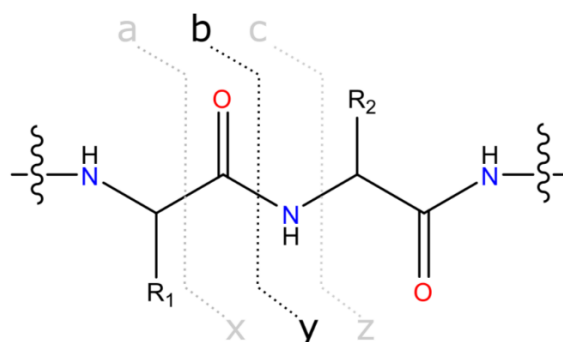
Figure 1.47: Bonds cleaved during peptide/protein IRMPD MS/MS usually include the amide (CONH) bond, highlighted, causing the production of b/y ions and a series of neutral, small-molecule losses such as H_2O, NH_3, CO, CO_2.

IRMPD of biomolecules such as peptides and proteins tend to produce the same types of fragments as CAD, though the tuning of the two techniques differs from tuning of a CAD acceleration voltage to the altering the laser power and laser pulse length (irradiation time) for IRMPD. Though other factors such as choice of collision gas, alignment of laser, and position of the ions inside the ICR cell, can affect the two techniques differently, they can otherwise be used to generate the same fragmentation data for peptides and proteins, though this changes for other analytes such as DNA/RNA, discussed later.

1.7.5 Ultra-Violet Photo-Dissociation (UVPD):

UVPD is extremely similar to IRMPD; selected ions are irradiated with laser-generated photons in order to excite molecules and cause dissociation. Though instead of exciting the ions purely vibrationally, as with IRMPD, UVPD excites ions vibrationally and electronically using a pulsed UV laser, ideally a 193nm Excimer laser, as shown by Broadbelt *et. al.,*[95] though Nd:YAG lasers set at 216nm have been shown to be viable for UVPD if the molecule is tagged with a chromophore capable of absorbing the particular UV wavelength.[96,97] The excited species generated by UVPD have been shown to dissociate into many different fragmentation channels, producing b/y ions, c/z ions,

and even a/x ions, making UVPD spectra very complex, but provide a great degree of fragmentation information in one spectrum.[95]
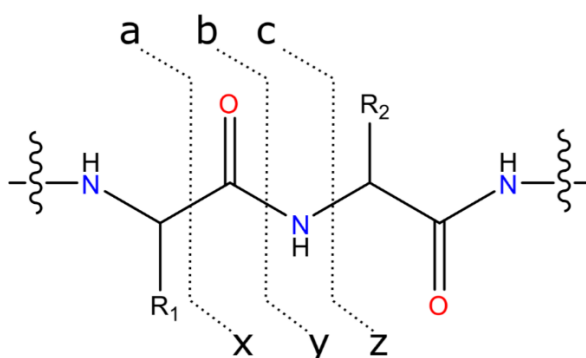
Figure 1.48: Fragments produced by UVPD MS/MS of poly(amino acids). UVPD causes extensive backbone and side chain fragmentation, producing fragments from cleavage of all 3 backbone bonds.

UVPD can be conducted in the ICR cell[4] in much the same way as IRMPD, prior to any excite/detect pulses, but has also been implemented in accumulation cells and non-linear instrument configurations such as Orbitraps.[95] Nevertheless UVPD is entering an increased resurgence in popularity within the MS field and has been shown to provide detailed information both for Bottom-Up (peptide) studies and Top Down (full protein) MS/MS studies, and including metal modified biomolecules, as shown in a recent article centred around the ability for various MS/MS techniques to characterise Platinated oligonucleotides.[95]

1.7.6 Electron Capture Dissociation (ECD):

Developed in the late 1990's by Zubarev *et. al.*,[98] Electron Capture dissociation was discovered during an attempt to achieve Ultra-Violet Photo-Dissociation (UVPD), the UV laser used struck a metal plate within an ICR cell and caused photo-emission of electrons, which were subsequently captured by trapped analyte ions, causing Electron Capture Dissociation (ECD) and producing detectable fragments. The experiment was then repeated with a direct electron source and the resulting electrons caused effective dissociation of isolated biomolecules.[98] ECD is based on the emission of relatively low energy (0.1-3eV, usually ~1eV for biomolecules) electrons into an area containing trapped ions of interest, the electrons are then captured by positively charged analytes, creating a radical-cation species (known as the charged-reduced species (CRS)). Once captured the radical-cation can dissociate into several fragmentation pathways, for ECD MS/MS of poly(amino acids), this generally results in breakage of the N-Cα bond (NH-CHR bond), producing c/z• fragment ions:
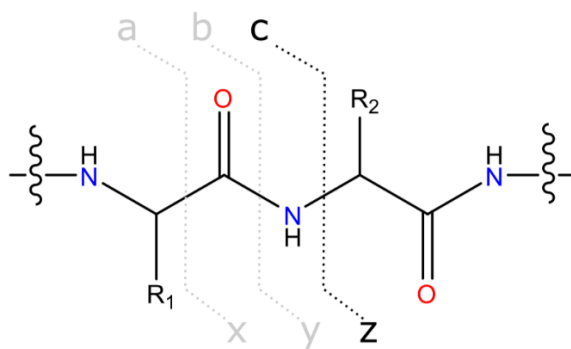
Figure 1.49: Fragments produced by ECD MS/MS of Poly(amino acids). ECD typically causes cleavage of the N-Cα bond and production of c/z• fragments, though many fragmentation channels exist and are influenced by various factors and/or functional groups (see main text).

The radical dissociation mechanism responsible for the creation of c/z ions from the charged reduced species has been the point of debate for many years, with many mechanisms proposed to explain the wide array of fragments, side chain losses, and neutral molecule losses resulting from

ECD MS/MS. However recently the most accepted mechanisms for ECD dissociations are the Utah-Washington mechanism (UW mechanism) which explains the majority of what could be considered "standard fragmentations" and the Free Radical Cascade Mechanism (FRCM)[99] which covers many other specialised circumstances.

Fragment	Observed mass/ Da	Exact mass / Da	Error (ppm)
c2	271.18770	271.187699	0.004
c4	496.33543	496.335425	0.010
c5	624.39409	624.394003	0.139
c6	752.45256	752.452580	-0.027
a6	708.43884	708.438940	-0.141
a7	855.50728	855.507350	-0.082
c7	899.52095	899.520994	-0.049
c8	1046.58930	1046.589408	-0.103
a9	1059.59724	1059.597220	0.019
z9·	1078.56298	1078.562810	0.158
c9	1103.61086	1103.610872	-0.011
c10	1216.69476	1216.694936	-0.145
$[M+H]^{+\cdot}$	1347.73546	1347.735421	0.029

Figure 1.50: ECD MS/MS spectrum of Substance P in the 2+ charge state ([SubP+2H]$^{2+}$) (top). The parent ion at ~674 m/z is shown to capture an ECD electron to create the charge reduced species (CRS) observed at the far right hand side of the spectrum, with assignable fragments (middle) produced labelled and then summarised on the fragmentation map (bottom)

An example ECD MS/MS spectrum of the peptide Substance P in the 2+ charge state ([SubP+2H]$^{2+}$) is above in Figure 50, The ECD MS/MS spectrum shown is very different to that of the same peptide under CAD MS/MS shown previously in Figure 1.45, producing not only different fragment types (c/z instead of b/y), but also different cleavage sites, the key differences in the fragmentation mechanism (discussed below) allow CAD/IRMPD and ECD to provide complementary data to each other.[71,98,100,101] ECD has also been shown to have a preference for capture at di-sulphide bond functional groups, which have previously been shown to have a detrimental effect on CAD MS/MS due to their impressive stabilisation of protein structure and ability to create cyclic regions within the biomolecules.[102,103] The ability for ECD to cleave disulphide bonds preferentially, then be able to further induce electron capture on the now reduced-protein structure has enabled MS/MS characterisation of not just large biomolecules but also heavily disulphide-linked biomolecules.[71,102,103]

1.7.6.1 Utah-Washington ECD mechanism:

The Utah-Washington mechanism to explain ECD MS/MS fragmentation is a combination of two previously released mechanisms; the Utah mechanism[104] and the Washington mechanism.[105] While each individual mechanism focused on different steps in the dissociation mechanism (Utah mainly focusing on the electron capture steps while the Washington on the subsequent dissociation), the combined UW mechanism is currently the most widely accepted mechanism for ECD MS/MS cleavage of biomolecules.

The UW mechanism postulates that incoming electrons are captured directly into coulomb-stabilised OCN amide π* orbitals via Rydberg orbitals (at Rydberg numbers 3-6) which are formed around charge sites.[106,107] Coulombic stabilisation of the ONC π* orbital lowers the energy barrier of electron capture from ~2.5 eV down to an approachable level for the incoming electron (~0.8-3eV), thus allowing effective electron capture. Coulombic stabilisation of the π* orbital is believed to be caused by electrostatic interactions between the π* orbital and nearby positively charged groups, such as $-NH_3^+$ or $-NH_2Na^+$, as suggested by Simons and Sobczyk & Simons[103,108,109] The coulomb stabilisation requires close proximity, and for the decrease in energy barrier needed the stabilising group would need to be within ~6 Å and can be described using equation 14:[109]

$$C = 14.4(eV\text{Å}) \sum_{j=1}^{N} \frac{Z_j}{R_j(\text{Å})}$$
eqn. 14

Where Z_j is the jth charge; R_j is the distance between the jth charge and the OCN amide groups; and 14.4(eV Å) is the result of the Coulomb constant multiplied by the square of the electron charge.

The radical can be captured at either a disulphide bond, causing reduction and cleavage of the bond, and generally stabilisation of the resulting radical, or at the OCN group as described above, this will create an aminoketyl radical in the excited state (see Figure 1.51, below).
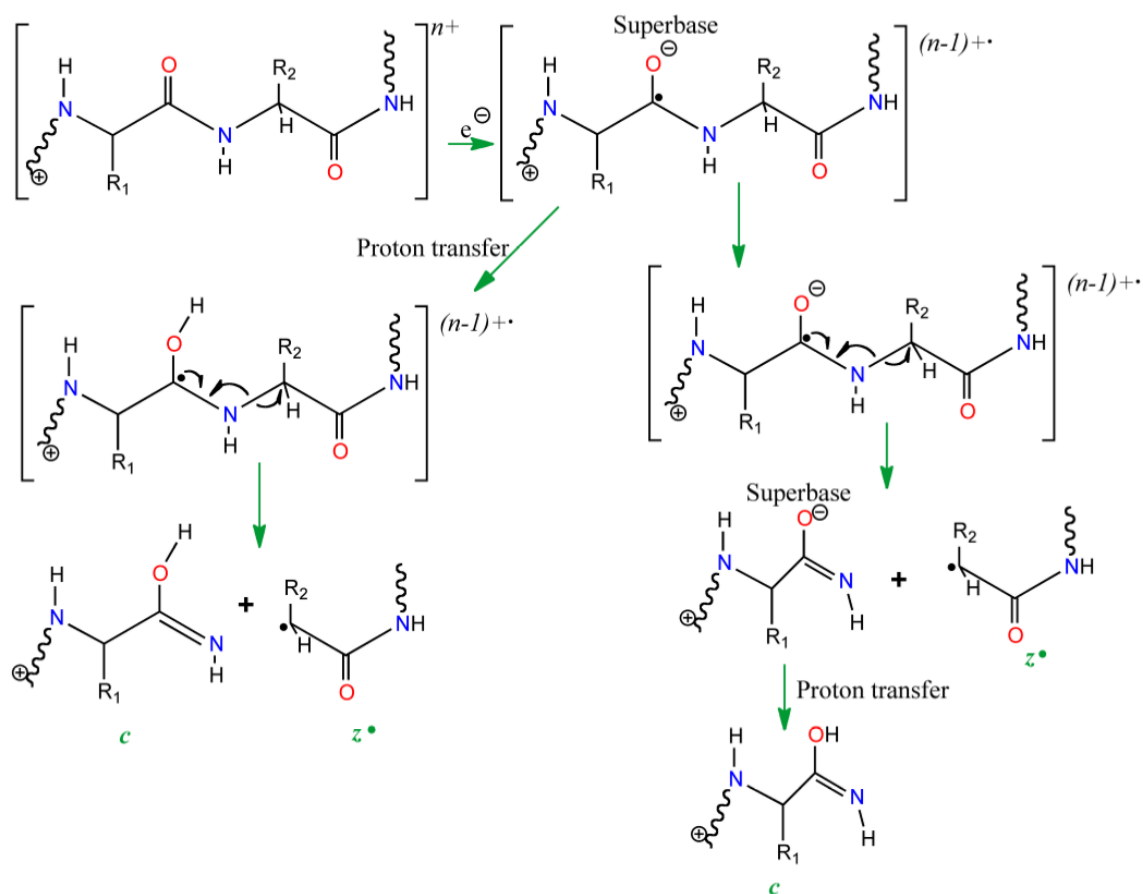
Figure 1.51: representation of the Utah-Washington ECD MS/MS dissociation mechanism, showing the two pathways to dissociation ions can follow after electron capture at the backbone carbonyl group. Reproduced from Lopez-Clavijo et.al.[85]

The newly formed aminoketyl radical is inherently unstable, but can be partially stabilised by charged groups up to 15 Å away,[110] and is incredibly basic, the so called "superbase" species has a proton affinity of >1200kJ/mol[111,85] and rapidly abstracts a proton from a charged site nearby, thus forming the fragile charged reduced species (CRS). The CRS can then undergo typical radical fragmentation to form c and/or z• ions from cleavage of the N-Cα bond, as shown above in Figure 51. The enol-imidate intermediate is also a superbase species with a proton affinity of >1400kJ/mol[111,112,85] and can easily abstract an available proton to form a c ion. Results suggest that proton transfer reactions involving c/z• species occur after N-Cα bond cleavage, not before.[113] The UW mechanism also explains many experimental results showing that low-energy activation of the CRS species can produce an array of c/z• fragments, remarkably similar to the usual ECD spectrum of

the given species, providing evidence that the CRS is a delicate, hydrogen-bonded species containing

many pairs of fragment species which although have been effectively cleaved via ECD are still held

together and can be broken apart using post-ECD activation[114,115] as shown in Figure 1.52, structure

IV, below, from the Utah mechanism.[103,108]
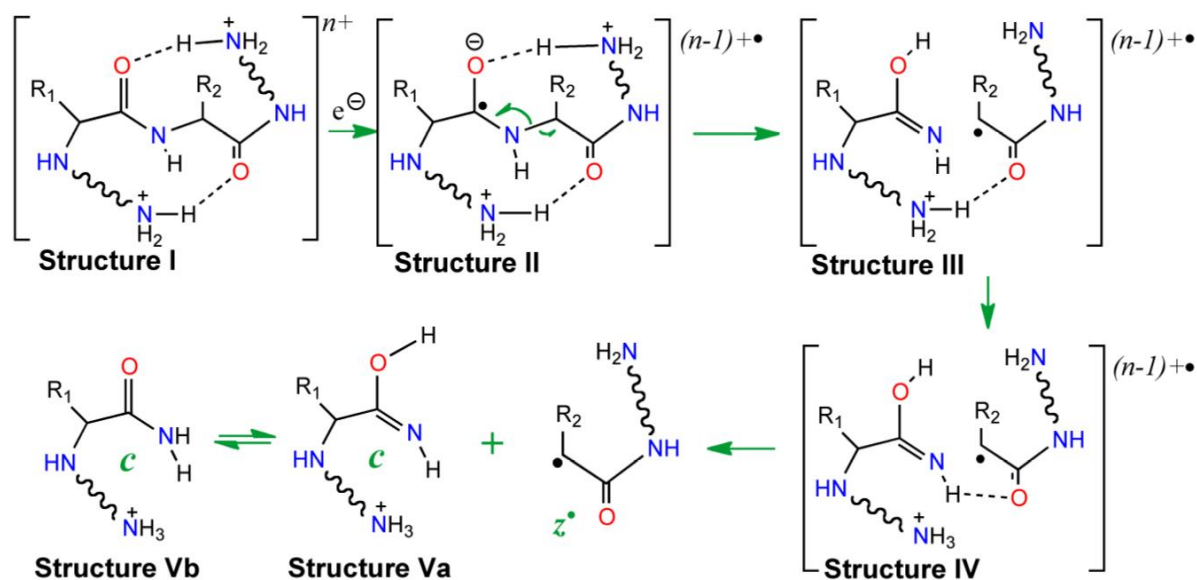
Figure 1.52: ECD-induced dissociation of a dipeptide according to the Utah mechanism.

Reproduced from Lopez-Clavijo *et.al.*[85]

1.7.6.2 Free Radical Cascade Mechanism (FRCM):

The Free Radical Cascade Mechanism (FRCM) was proposed by O'Connor *et.al.* and is centred

around uncovering the possible mechanisms that allow observable ECD-induced fragmentation of

cyclic biomolecule species. Peptide/protein species, if fragmented by ECD MS/MS, generally follow

the rule that one electron captured equates to one bond cleaved, due to the products formed post-

bond cleavage and stabilisation of the radical post-fragmentation. The perfect example of this rule is

the proline effect experienced by users of ECD MS/MS,[112] where the ECD-induced cleavage at the

proline residues in peptides and proteins causes no observable fragment ions. The proline effect is

caused by the fact that although ECD can effectively cleave the N-Cα bond within the proline

residue, the alkyl functionality of proline is cyclic and so the amino acid may be fragmented in one

96

location, but is effectively tethered to the two sections of the poly(amino acid) backbone via the alkyl group, and so does not produce individual fragment ions to be detected. This process is summarised in Figure 1.53, below:
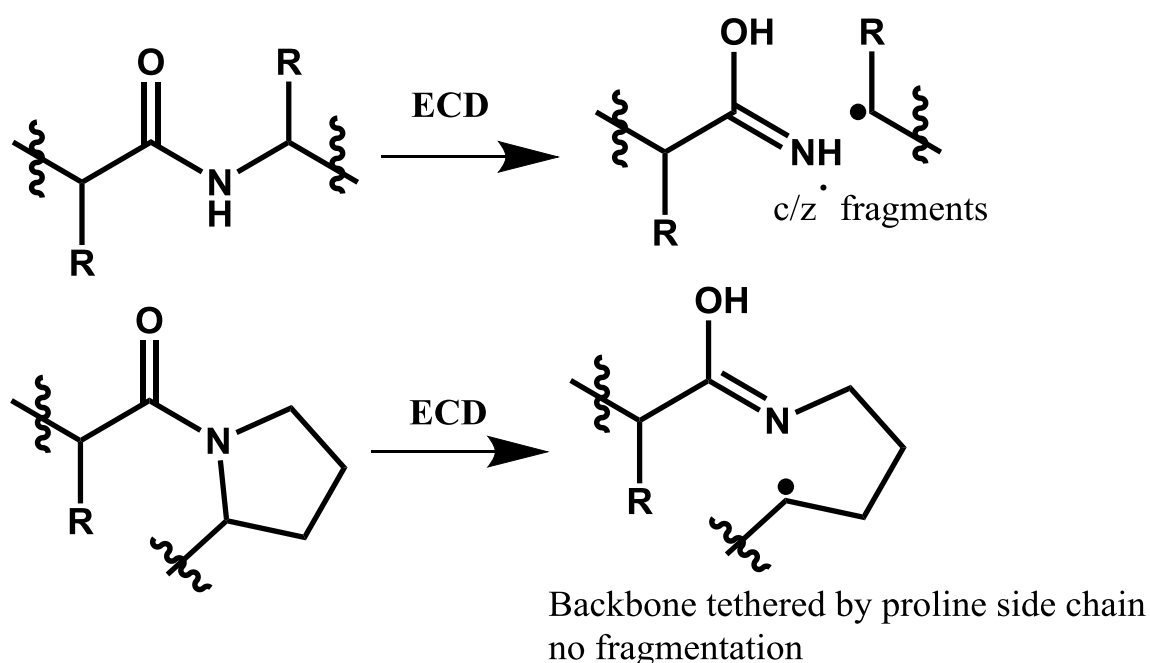
Figure 1.53: Representation of the reaction products formed after ECD MS/MS cleavage of normal amino acids (top) and proline (bottom), explaining the proline effect where no fragmentation at the proline amino acid sites in peptides and proteins can be cleaved to produce observable fragments in ECD MS/MS.

1.7.6.3 Cyclic peptides and the proline effect:

The proline effect is an unfortunate limitation of electron-based dissociations of poly(amino acid) species as it reduces the number of possible cleavage sites in a biomolecule sequence. It is worth noting CAD MS/MS does not suffer from this Proline effect as CAD cleaves the amid bond within proline, which is not tethered to the backbone, and so can cause fragmentation. There is also a "Proline effect" for CAD MS/MS where is was noticed that CAD fragmentation was enhanced at amino acids neighbouring proline residues, however this is a minor effect and not of great

importance for the results presented herein. For the extent of the thesis work presented the Proline effect mentioned will be specific to the electron-based MS/MS proline effect discussed.

Extending the same initial idea of the election-based MS/MS proline effect, ECD/Electron Transfer Dissociation (ETD) MS/MS of cyclic peptide species should also be similarly limited by the tethering of fragmented ions. If a cyclic peptide species was fragmented at one location, the ion would still be held together at another point in the cyclic structure, and so no smaller fragments could be produced/detected beyond small molecule/side chain losses, as the amino acid backbone will remain intact. However it was observed by O'Connor *et.al.*[99] that ECD MS/MS of some known cyclic peptide species (shown below in Figure 1.54) as the $[M+2H]^{2+}$ ions, showed both the capture of the ECD electron to produce a CRS (as expected) and side chain losses (also expected) O'Connor *et.al.* also produced an array of smaller fragments resulting from cleavage of the amino acid backbone, which couldn't be possible using the currently held ECD mechanisms (discussed above).
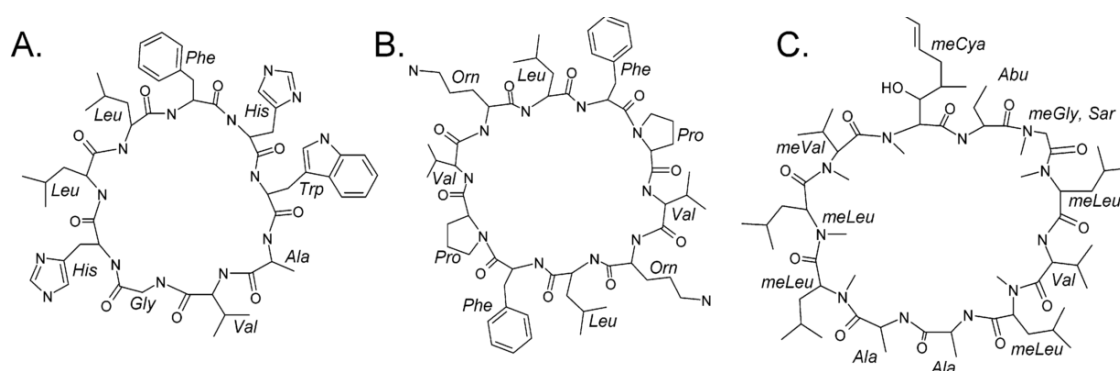
Figure 1.54: Cyclic peptide species studied by O'Connor *et.al.*[99] to show that cyclic peptide species can fragment under ECD MS/MS to produce unique sequence-informative fragments, despite the cyclic structure tethering fragments together. A) cyclo-LLFHWAVGH, B) gramicidin S, and C) cyclosporine A. Reproduced from O'Connor *et.al.*[99]

The ECD MS/MS spectra of the cyclic peptides produced an array of fragments shown in the spectra below, Figure 1.55:
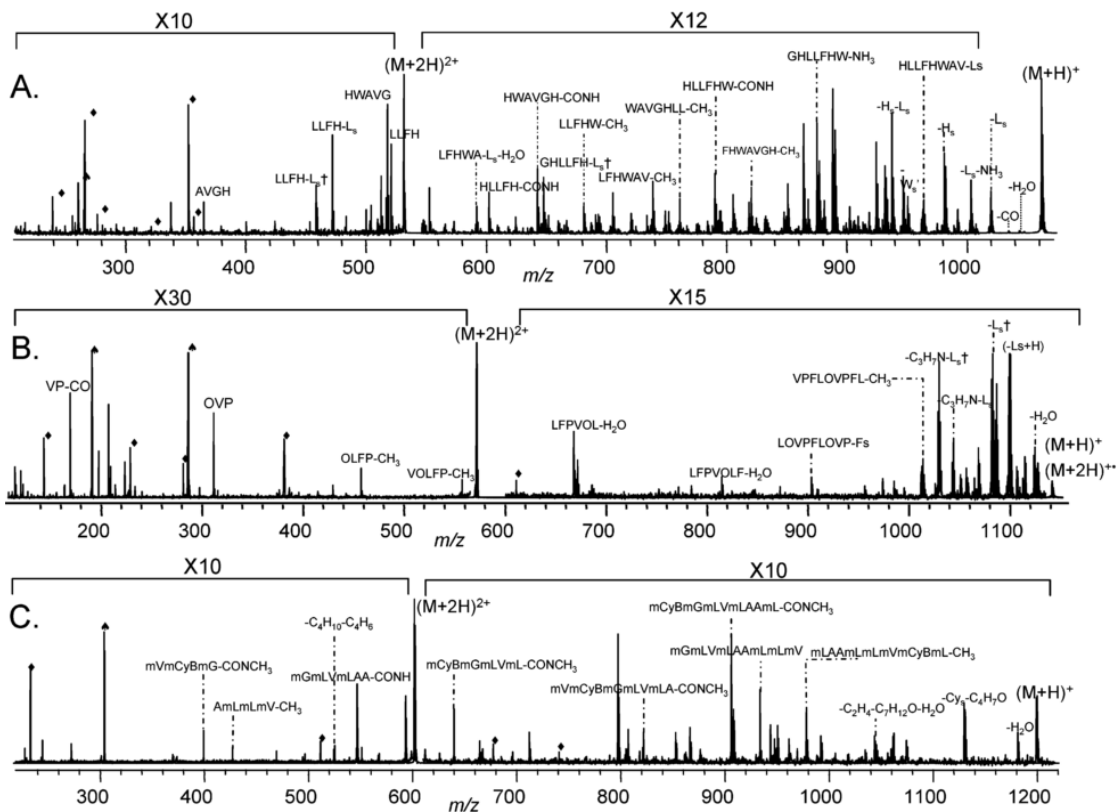
Figure 1.55: ECD MS/MS spectra of the [M+2H]2+ ions for A) cyclo-LLFHWAVGH, B) gramicidin S, and C) cyclosporine A. Reproduced from O'Connor et.al.[99]

ECD MS/MS of the cyclic peptides produced enhanced degrees of side chain losses compared to the expected results for linear counterparts. This enhancement was likely due to the lack of backbone cleavages and stabilisation of the radical by movement to the Cα, which commonly results in loss of the amino acid side chain, a corresponding entropy increase accompanies these reactions. The most interesting observation was the recorded sequence-informative fragments produced during ECD MS/MS. These fragments could only be produced by multiple cleavages of the peptide backbone by a single ECD electron, which did not follow previously discussed rules/mechanisms. As a result O'Connor *et.al.* believed that the electron could cause dissociations from the already cleaved position in the cyclic peptide via a radical-induced cyclic rearrangement reaction, shown below in Figure 56:
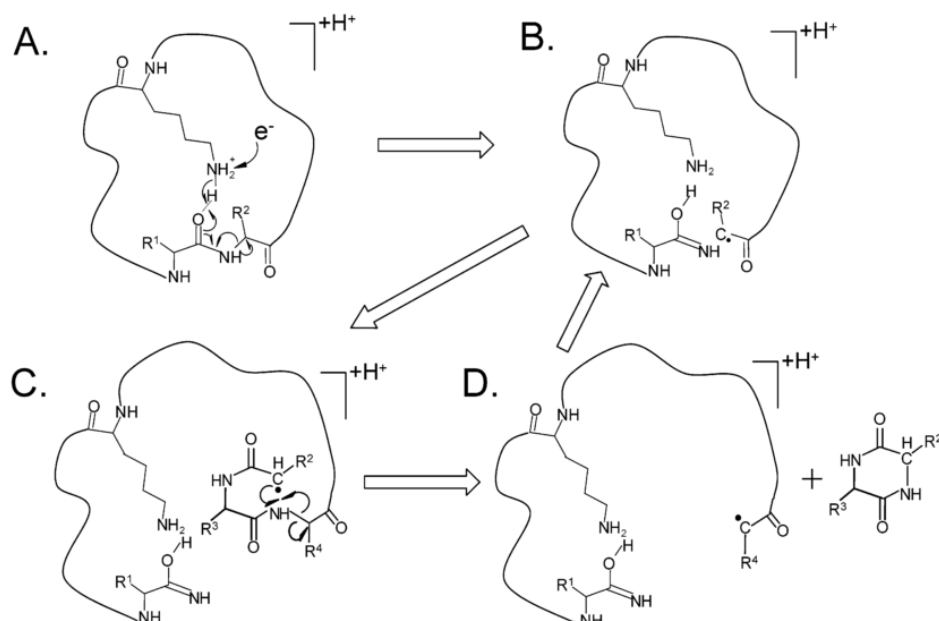
Figure 3. Nonergodic cleavage from electron capture dissociation (A → B) initiates an α-carbon radical which can propagate along a peptide backbone by free radical rearrangements (B → C → D), cleaving the N−Cα bond and forming another α-carbon radical.

Figure 1.56: The radical-induced cyclic rearrangement reaction proposed by O'Connor *et.al.*[99] to explain the sequence-informative fragments produced during ECD MS/MS of cyclic peptides. Reproduced from O'Connor *et.al.*[99]

The release of pairs (or more) of cyclic oligo(amino acids) via a radical rearrangement helped assign a large number of peaks in the ECD MS/MS spectra produced. The Authors also believed that the radical site could propagate along the poly(amino acid) backbone and thus move to different sites prior to cyclisation, rearrangement, and release of other amino acid combinations, according to Figure 1.56, above.

The Free Radical Cascade Mechanism (FRCM) is therefore a niche ECD mechanism which is of use to non-standard and/or cyclic ion MS/MS spectra, and helps explain why radical mechanism MS/MS techniques can not only cause enhanced side chain losses, but also fragment cyclic species in multiple locations to produce detectable and, with the FRCM, assignable species to inform the user of the biomolecule sequence and structure.

1.7.6.4 ECD in biomolecule analysis:

The use of ECD has, for the most part, been confined to use in ultra-high vacuum regions of an MS and usually requiring use of an FT-ICR MS, this is to avoid damage to the electron emitting cathode needed to generate and release low energy electrons, though efforts are underway to create ECD MS/MS devices for use in other instruments, namely the Chimera branched-RF Ion Trap[116] and the electromagnetostatic ECD cell.[117,118] The benefits of ECD MS/MS are numerous and have been studied extensively in ECD's relatively short lifetime.[85,101,112,119,120] ECD has been described as a non-ergodic process,[98,100] meaning that unlike CAD and IRMPD, bonds are broken before the energy of the reaction is evenly distributed throughout the molecule, and so is not a "slow heating" or threshold type dissociation, though recent studies have suggested this process to not be strictly non-ergodic.[105] This also means that if the species of interest contains weakly-bound molecules/interactions, these may be preserved in the fragment ions detected, which is a huge benefit over threshold techniques such as CAD and IRMPD. The ability to retain delicate/fragile modifications/interactions has been an exceptional feature of ECD MS/MS and has enabled the study of fragile PTM's such as carboxylation , sulfation, O-glycosylation, N-glycosylation, disulphide bonds, phosphorylation.[101,85,115,121] Beyond PTM's ECD has also enabled the study of non-covalent interactions,[122,123] protein folding,[124] and metallated biomolecules.[86,125,126,127,128]

Once captured the ECD induced CRS dissociation is dictated by radical chemistry, unlike CAD/IRMPD/threshold based MS/MS, meaning that the rearrangement and dissociation of biomolecules, and other species, is usually driven by the stabilisation of the unstable radical. Charge location, electronegativity, and tertiary structure are thus now large factors in fragmentation, whereas radical stabilising groups (such as electron quenching functionalities) can have a detrimental effect on ECD MS/MS due to the stabilisation of the radical in a non-dissociative manner.

ECD of transition metal complex-containing species is thus a very interesting area to consider as the interaction of electrons with metal d-orbitals and more complex electronic configurations experienced with heavy metals can complicate and alter the ECD MS/MS process. Metal-induced side chain loss during ECD has already been documented by Li *et. al.,*[86] while studying platinum complex association with biomolecules. They showed that not only could ECD electrons interact with the platinum complex modification, but also dissociate it if the radical were sequestered on the platinum centre itself, this could also cause unusual fragmentation of the methionine residue previously bound to the complex, as shown in Figure 57 below. It has been suggested that this and various other fragmentation behaviour could be used to as a "fingerprint" to effectively characterise platinum containing peptides and proteins.[126]
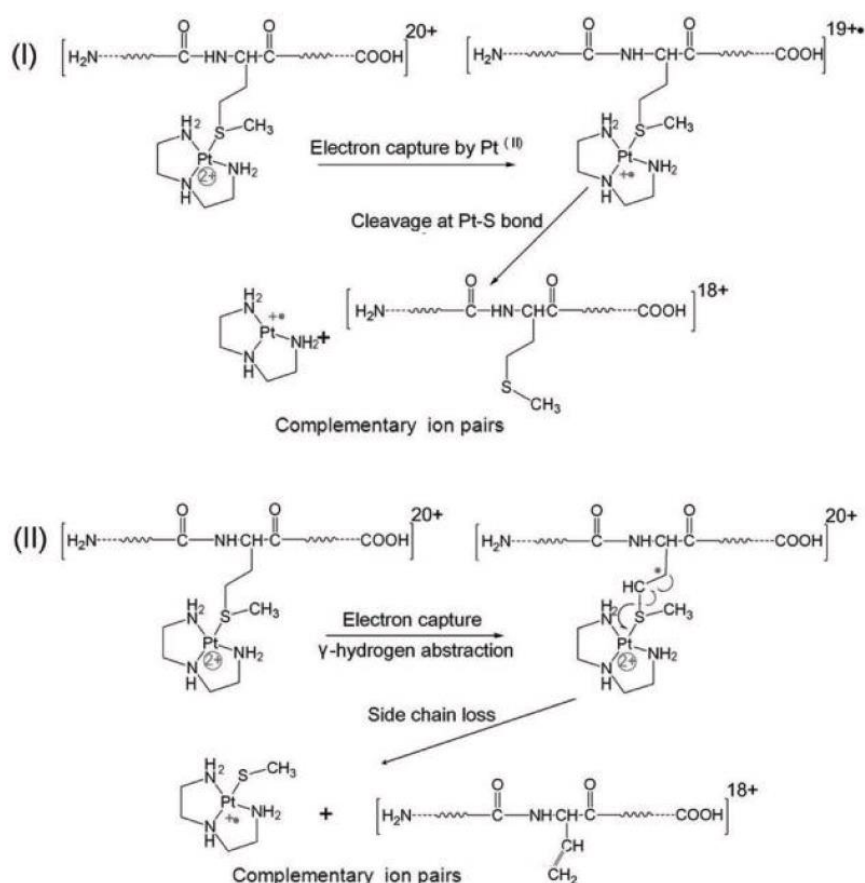
Figure 1.57: Platinum complex loss and Platinum-induced methionine side chain loss due to electron capture at / electron transfer to platinum metal centre during ECD MS/Ms analysis of platinated protein species. Reproduced from Li *et.al.*[86]

1.7.7 Hot-ECD (hECD):

ECD utilises rather low energy electrons, typically ~0.8-3eV, which can cause extensive biomolecule fragmentation, as discussed above, ECD also causes an array of amino-acid side-chain losses, some of which can be used diagnostically.[85,122] However ECD Side Chain Losses (SCL's) are not only distinctly different from those observed for CAD, which tend to be neutral molecules, they are also more diagnostic, and less abundant than those observed in CAD/IRMPD. In certain cases SCL's are characteristic and can be used to confirm/test the presence of certain amino acid groups present within a given biomolecule,[129] to this end it can sometimes be advantageous to attempt to increase the abundance/yield of SCL's in ECD MS/MS spectra in order to aid this type of analysis.

hECD typically uses electron energies slightly higher than that of normal ECD (~5-7eV)[85,122,130] and the transition into hECD can be subtle, however extra SCL peaks appear and ECD efficiency can lower, as the electron energy moves away from the optimum for normal ECD and increased loss through extra SCL pathways. Due to hECD's ability to cleave unique side chains during dissociation, hECD has been used as an effective tool at differentiating between aspartic acid (45Da Side chain loss and iso-aspartic acid (57Da Side chain loss)[131], which is a key assignment of deamidation at these residues and could be used to study the progression of Alzheimer's patients.[131]

1.7.8 Electron Induced Dissociation (EID):

ECD proved to be a remarkable technique, able to achieve far greater sequence/cleavage coverage of proteins than CAD, and could retain fragile PTM's effectively. However for the analysis of some peptides and small molecules, ECD was not a viable technique due to the requirement that a parent ion would need to be at least a 2+ charge in order to produce 1+ detectable fragments after the electron capture event. This charge-limitation proved difficult for small molecule MS/MS, which often only observes 1+ ions, and a distinct proportion of peptide MS/MS in which many peptides only present as 1+ ions, or only produce a very limited intensity 2+ or more ion. In order to combat this lack of electron-based dissociation for singly charged species Budnik *et. al.*[132] and Fung *et. al.*[133]

investigated the effect of tuning electron energy during ECD to find a viable region for electron-based dissociation of 1+ ions. High energy electrons (70eV) were known from EI ionisation to interact with small molecules, and induce fragmentation, however lower (medium) energy electrons were found to fragment singly charged ions. Using electrons ~9-18eV,[85] electron induced dissociation (EID) could be performed. For singly charged small molecules this produced a range of fragments, akin to EI, but available to selected MS/MS of selected ions, which allowed soft-ionisation followed by MS/MS of various analytes. For Singly charged biomolecules EID produced c/z• and a/x type fragments. EID has also been shown to cause a series of cross-ring cleavages and enable detailed characterisation of small and cyclic species.[79]
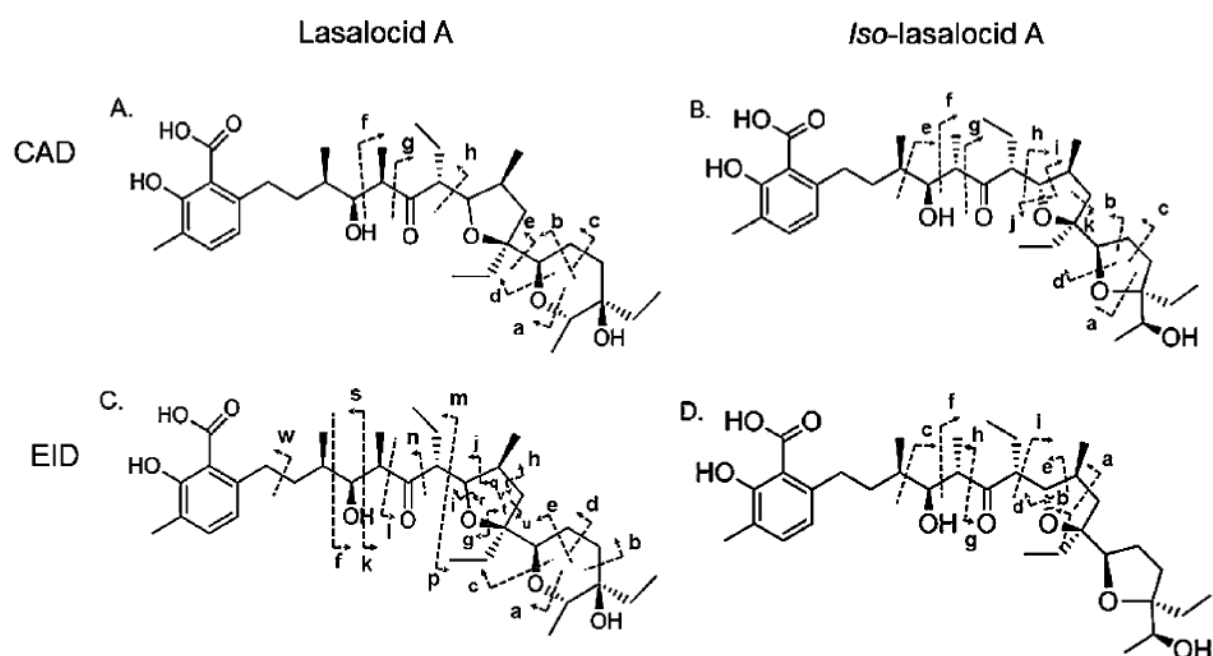
Figure 1.58: EID and CAD MS/MS fragmentation of polyketide natural products showing the complementary MS/MS fragmentation data provided by the different techniques. Reproduced from Wills *et. al.*[79]
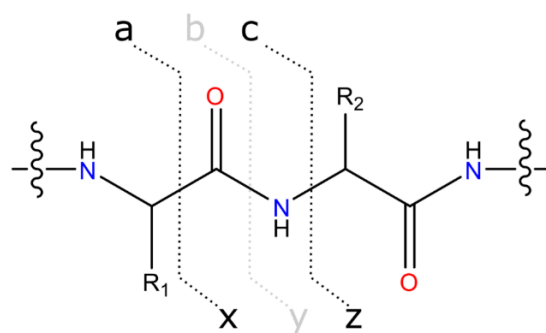
Figure 1.59: Fragments produced by EID MS/MS of poly(amino acids). EID produced c/z• and a/x fragments, highlighted above.

Recently EID has been shown to be an effective and viable technique for electron-based MS/MS of larger, multiply charged biomolecules, such as proteins. Li *et. al.* have shown that EID actually provides complementary data to both ECD and CAD MS/MS, particularly for very large biomolecules, where sequence/cleavage coverage during MS/MS can be extremely limited. Particular uses of EID and various ExD fragmentation techniques have also been reviewed recently by Qi *et.al.*[134]

1.7.9 Electron Detachment Dissociation (EDD):

Many analytes, and biomolecules, specifically acidic peptides/proteins and RNA/DNA species, do not ionise well in positive mode ESI, if at all, though are readily ionised via negative mode ESI. Though the effectiveness of ECD had be proven on a range of biomolecules and expanded the effective range of MS/MS to larger proteins, it was still an exclusively positive-mode technique, and could not work on negatively charged ions. In order to create a viable electron based dissociation technique for negatively charged ions Budnik and Zubarev *et. al.*[135] investigated the use of electrons on negatively charged analytes trapped in an ICR cell. They showed that using medium energy electrons (15-20eV) could induce electron detachment on the target analyte, causing an $[M-xH]x-$ species to become an $[M-xH]^{(x-1)\bullet-}$ radical species, which would then undergo radical directed fragmentation, in a similar way to ECD, but for negatively charged analytes.[85] EDD was extremely effective at dissociating oligonucleotides[136] and a variety of other biomolecules, and most

importantly presented many of the same benefits as ECD; including PTM retention and complementary MS/MS fragmentation to that obtained via CAD and IRMPD.[13685] EDD has been used to effectively sequence even large poly(nucleotides) up to 61 base pairs (bp), as shown by Breuker *et. al.*[136] presenting an effective way of sequencing biologically relevant poly(nucleotides) via MS/MS. EDD of negatively charged peptides causes cleavage of the N-Cα bond to form a and z ions, though a ions are also commonly detected, whereas EDD of oligonucleotides has been shown to cleave the O-C bonds either side of the phosphorous atom and form d & w ions, both are shown in Figure 1.61, below:
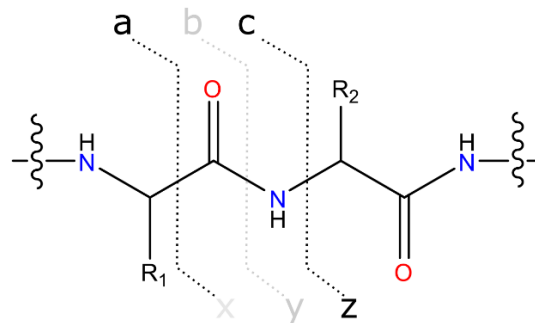
Figure 1.60: Peptide and protein backbone bond cleavages under EDD MS/MS, predominantly c, c˙, z, z˙, a, and a˙ fragments are formed and detectable.
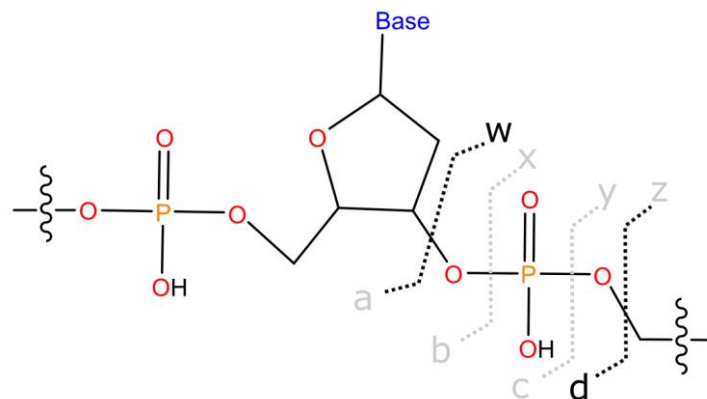
Figure 1.61: Structure of DNA/RNA species, unlike peptides and proteins, when fragmented via EDD MS/MS produce mainly d and w type ions, centred around fragmentation at the backbone phosphate groups.

EDD has also been shown to be critical in the full MS/MS characterisation of glycosaminoglycans (GAG's); Amster *et. al.* have shown[137] the sulfonation of individual sugars within the backbone chain (also known as the glycan code)[138] are lost using standard CAD MS/MS, but can be deciphered using carefully tuned EDD which causes the cross-ring cleavages in the sugar units needed to accurately assign PTM locations, shown in Figure 1.62 below:
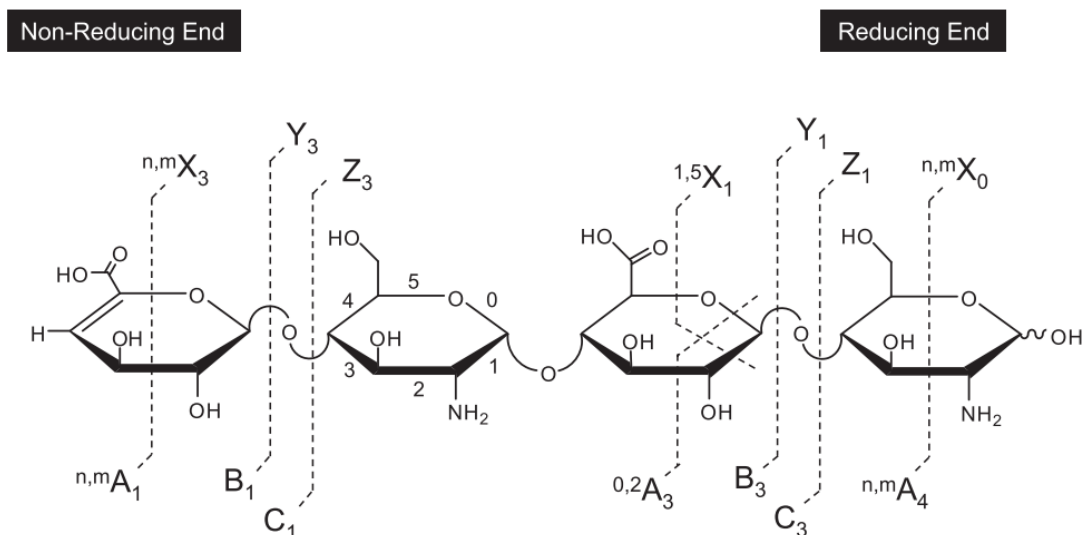
Figure 1.62: Fragmentation nomenclature used for the MS/MS of Glycosaminoglycans (GAG's). Reproduced from Wolff and Amster.[137]
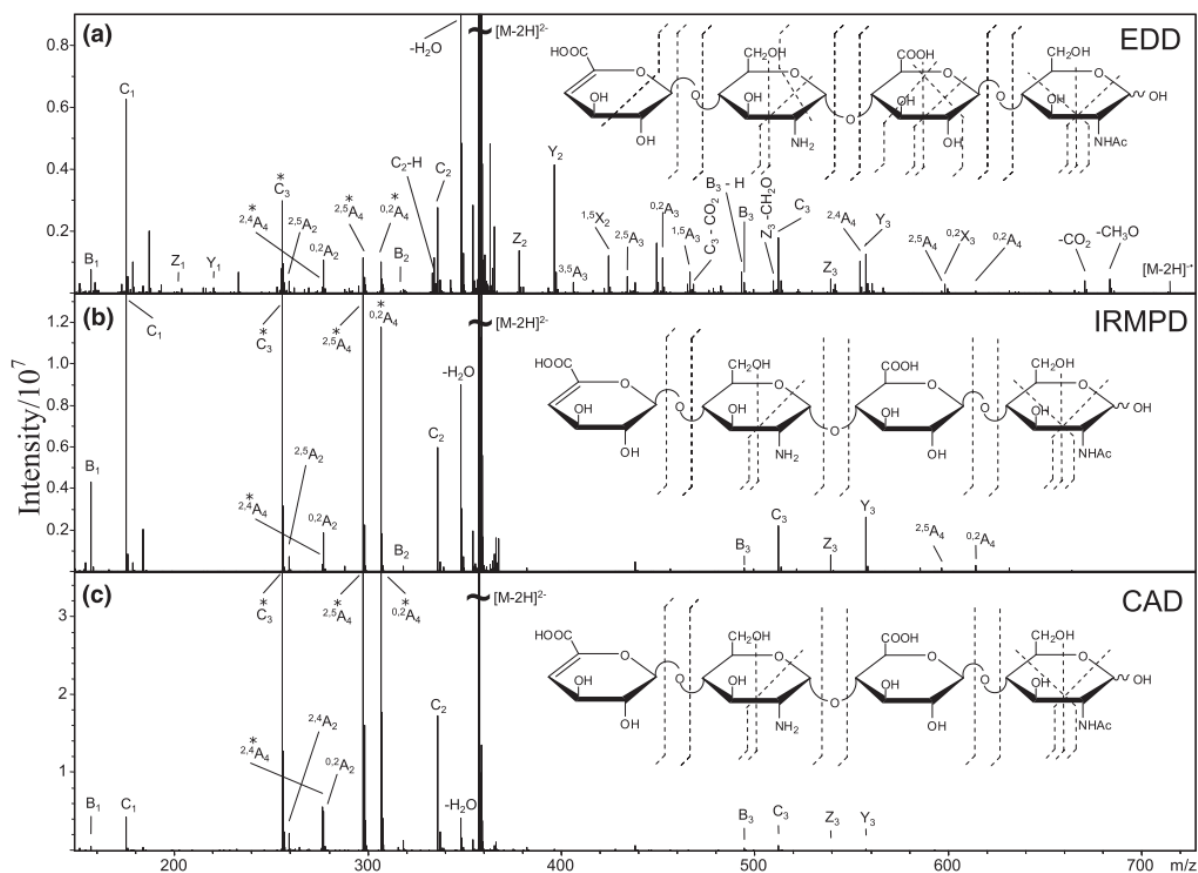
Figure 1.63: EDD, IRMPD, and CAD MS/MS of GAG's showing the complementary data available with each technique, but also the superiority of EDD for the effective fragmentation and critical cross-ring cleavages needed to effectively characterise GAG polymers. Reproduced from Wolff and Amster.[137]

Negative ion Electron Capture Dissociation (niECD):

Proposed by Hakansson *et.al.* in 2011 niECD was an alternative way to dissociate negatively charged ions to EDD.[139] Believed to be a pure electron capture event, like ECD, niECD requires that electrons overcome a repulsive coulombic barrier in order to allow the negatively charged electrons to approach the negatively charged analyte of interest. However if this can be overcome then electron capture can occur, causing electron transfer to the analyte $[M-nH]^{n-}$, creation of a radical species, in this case a charge increased species $[M-nH]^{n-\bullet}$, and eventually fragmentation.

Hakansson et. al. showed that by using extremely long irradiation times (~10-20 seconds) and electron energies of ~5.5eV niECD was possible,[139] producing effective fragmentation of oligo nucleotides and negatively charged peptide analytes.

1.7.10 Electron Transfer Dissociation (ETD):

The dramatic increase in MS/MS performance (both for sequence/cleavage coverage, and diagnostic ions) offered by ECD and the ability to probe structural information using radical directed dissociations was a very attractive prospect, unfortunately it was limited only to FT-ICR mass spectrometers. In an attempt to access electron-based dissociations on other instrument platforms Syka *et. al.* investigated the use of another way of creating charge reduced radical cation species; by using a carrier molecule/ion to transfer the electron to the analyte species instead of an electron-generating cathode.[140] Syka *et. al.* showed that by using negative anion radicals, generated by a negative chemical ionisation (nCI) source, electron transfer could be achieved by trapping both the positively charged species of interest and the radical anion. Dubbed Electron Transfer Dissociation (ETD), this process enabled effective electron based dissociations on any trapping mass spectrometer, and although possessing distinct differences and sometimes disadvantages over ECD, ETD allowed many different types of mass spectrometers to access these fragmentations and improve MS/MS analysis. ETD produces charged reduced species (CRS's) and c/z ions, similarly to ECD, shown below in Figure 1.64. The general process for ETD is also shown below in Figure 65, the choice of radical anion reagent can vary from instrument to instrument, but many are based on extended ring structures such as anthracene, which when reacted with CH_5^+ ions from the CI source can produce a range of ion peaks, including active ETD reagent ions such as $C_{14}H_9^-$ and $C_{14}H_{11}^{-}$.[140]
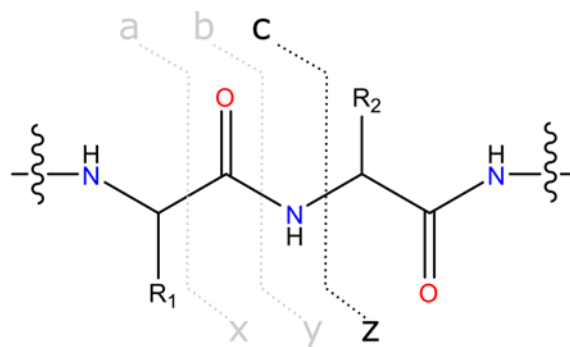
Figure 1.64: Ions produced during ETD of peptides and proteins, c and z ions are most frequently observed as product ions.
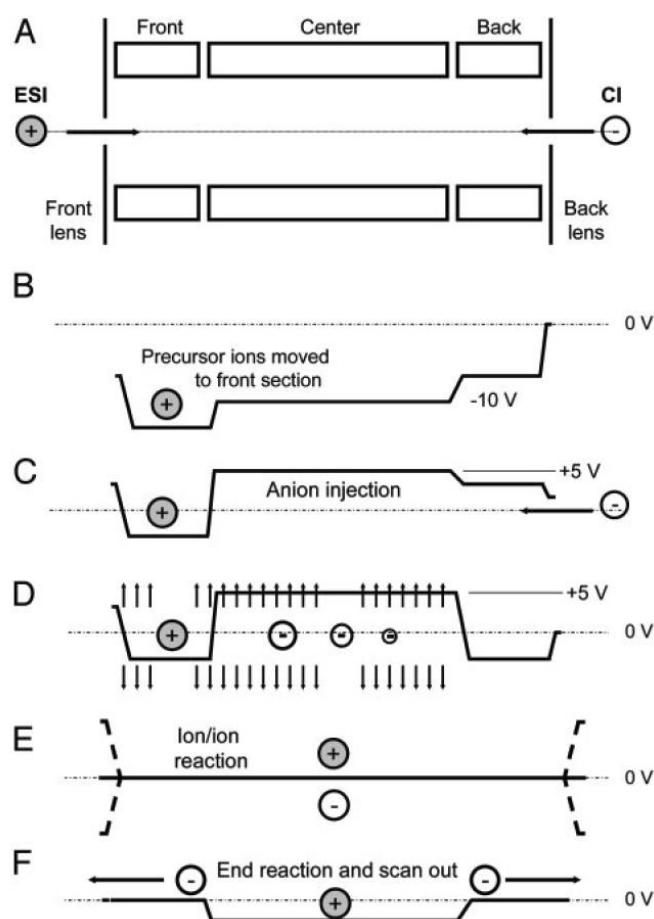
Figure 1.65: general procedure for achieving Electron Transfer Dissociation (ETD) within a multi-pole-coupled mass spectrometer (A). Ions are introduced into the multi-pole trap and confined using voltage barriers (B) the ETD reagent generated via a separate CI source are then introduced separately into the trap (C) Both ions of interest and ETD reagent ions are then co-confined within

the multi-pole trap (D) both ion types are then allowed to react for a certain (variable) time to allow interaction and electron transfer (E) the ETD reagent is then ejected, leaving the ions of interest and their daughter fragment ions produced due to radical directed fragmentation (ETD) all of the remaining ions can then be transferred to the mass analyser to be detect and produce an ETD MS/MS spectrum. Reproduced from Syka *et. al.*[140]

ETD has been successfully applied to the study of peptides,[141] proteins,[142] protein complexes,[143] antibodies,[89] PTM's,[144] and structural protein studies.[119] ETD has also been widely applied to larger scale proteomics studied to provide complementary data to CAD and allow PTM analysis on a proteome wide scale.[145] ETD often requires post-electron transfer activation (via CAD and/or IRMPD) to add additional energy to the ETD charge reduced species and break non-covalent bond in order to liberate more fragments and increase fragment ion yield, increasing sequence coverage of many species, especially larger proteins.[146]
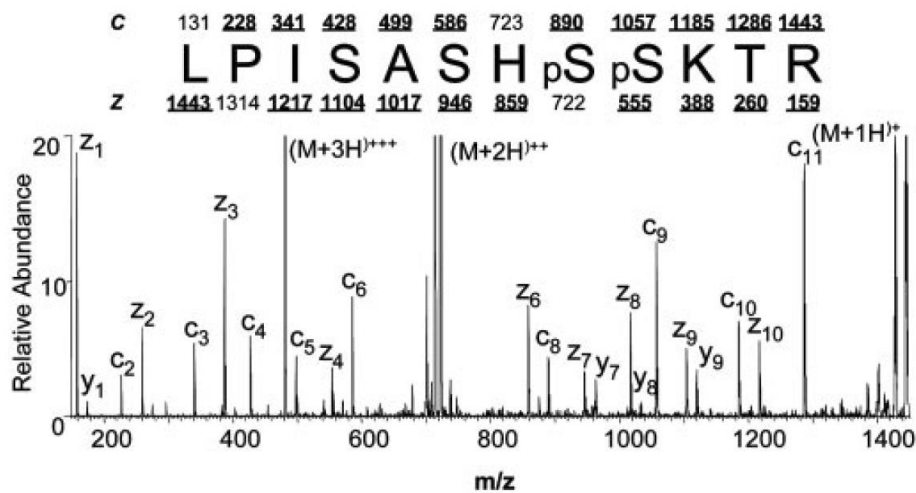
Figure 1.66: ETD MS/MS spectrum of a phosphorylated peptide (shown inset), observed c and z ions are labelled, ETD provided complete cleavage coverage of the phosphopeptide, proving ETD is a viable technique to sequence post-translationally modified biomolecules while still retaining the associated fragile modifications.

Recently various applications of Electron Transfer without dissociation (ETnoD) reactions have shown promise as a way of either lowering charge states in situations favouring low-charge state analysis, such as polymer mixtures, or study of high charge state species on lower resolution instruments seeking to reduce the charge state and achieve isotopic resolution.[147] The benefit of ETnoD being that ions can always be supplementally activated post-ETnoD to observe c/z ions if needed thereafter.

1.7.11 Multi-stage Tandem MS, MS/MS/MS, MS^3, or MS^n:

Ions can also be re-isolated in the ICR cell and IRMPD/UVPD/ECD/ExD (among other in-cell MS/MS techniques) can be used to perform tandem MS of the fragments produced (MS/MS/MS or MS^3), allowing further interrogation of fragment species.[5,148] This process can also be repeated further for MS^n experiments[73], though with each stage of MS/MS target fragment intensity will decrease (as no new ions are entering the cell) and so practically this places a limit on the number of stages of MS/MS. It is worth noting that MS^n experiments can also be performed outside of an ICR cell via movement in and out of different optic regions of a mass spectrometer, e.g. multiple stages of CAD MS/MS via continued exit and entry of fragment ions, though the same intensity limitations still apply.

1.8 A working hybrid-FT-ICR Mass Spectrometer:

Figure 1.67, below, shows the Bruker SolariX 12 Tesla FT-ICR Mass Spectrometer used for the majority of the work presented herein. The SolariX can be equipped with a range of ionisation sources, including but not limited to, ESI, nESI, MALDI, CI, APCI, APPI, and MAI. The SolariX is fitted with many features and techniques discussed above, most notably, an Infinity cell,[33] Sidekick electrodes,[35] a dual ion funnel-orthogonal AP-ionisation front end, and a split-octupole-quadrupole-hexapole-hexapole arrangement for transmission and effective MS/MS of target ions before the ICR cell, and an array of MS/MS techniques thereafter.
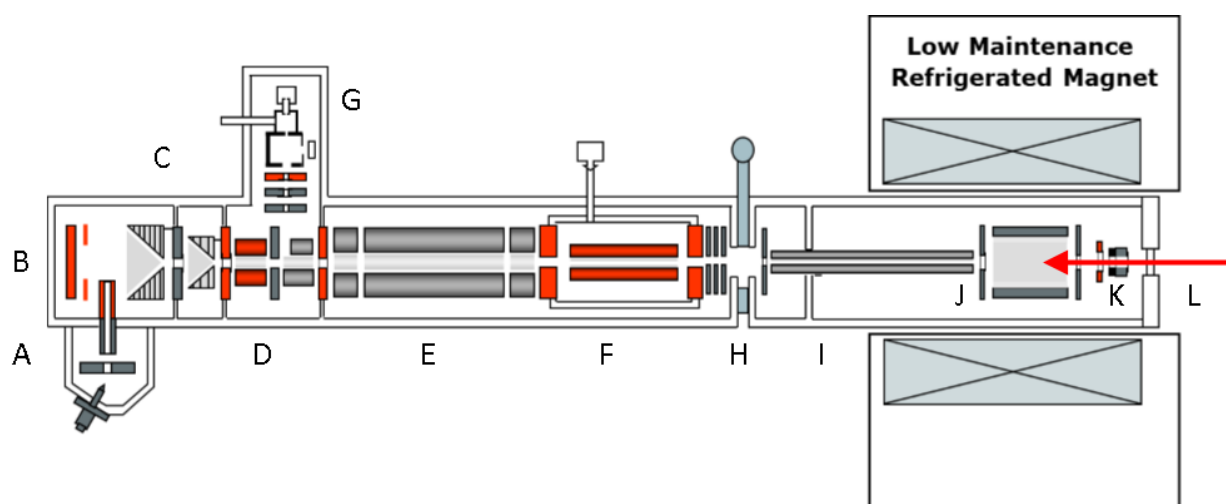
Figure 1.67: Schematic representation of the Bruker 12T SolariX instrument used for the majority of the work presented within the thesis. Corresponding labels are referred to in the main text. Courtesy of Bruker Daltonics, Bremen, Germany.

Ions generated via the AP-capillary entrance (A) of SolariX are orthogonally introduced into the vacuum region of the source front end and are then accelerated towards the ion funnels via a deflector plate (B), held at an equal polarity to the incoming ions, this allows neutrals to be taken away by vacuum, while ions are transmitted further into the FT-ICR MS. Ions generated via MALDI are generated in vacuum just under a MALDI extraction lens which will then accelerate ions towards the deflector plate, repelling them towards the ion funnels to be focused and transmitted.

Ions are focused via the dual-ion funnels (C) into a split octupole region (D) (spit to enable

orthogonal injection of ions generated by the internal nCI source (see the ETD MS/MS section,

above)) and then passed through a skimmer into the quadrupole (E).

Spray Shield
+/-4000V
(3900V)

Capillary
+/- 5000V
(4400V)

Red : typical tuning values for positive ion mode

Skimmer 2
+/-100V
(+10V)

Split Octupole: 2.5 or 5MHz, 600Vpp
(300Vpp)

CapExit
+/- 400V
(+350V)

Trap/Extract
+/- 100V
(+20V)
Fast pulsing
Rise time
<100us for 10V

Deflector
+/- 400V
(+330V)

IF1
+/-200V
(+150V)

Skimmer 1
+/-200V
(+10V)

IF2
+/-100V
(+8V)

DC
+/-100V
(+4.5V)

Funnel: 1.2MHz, 200Vpp
(150Vpp)
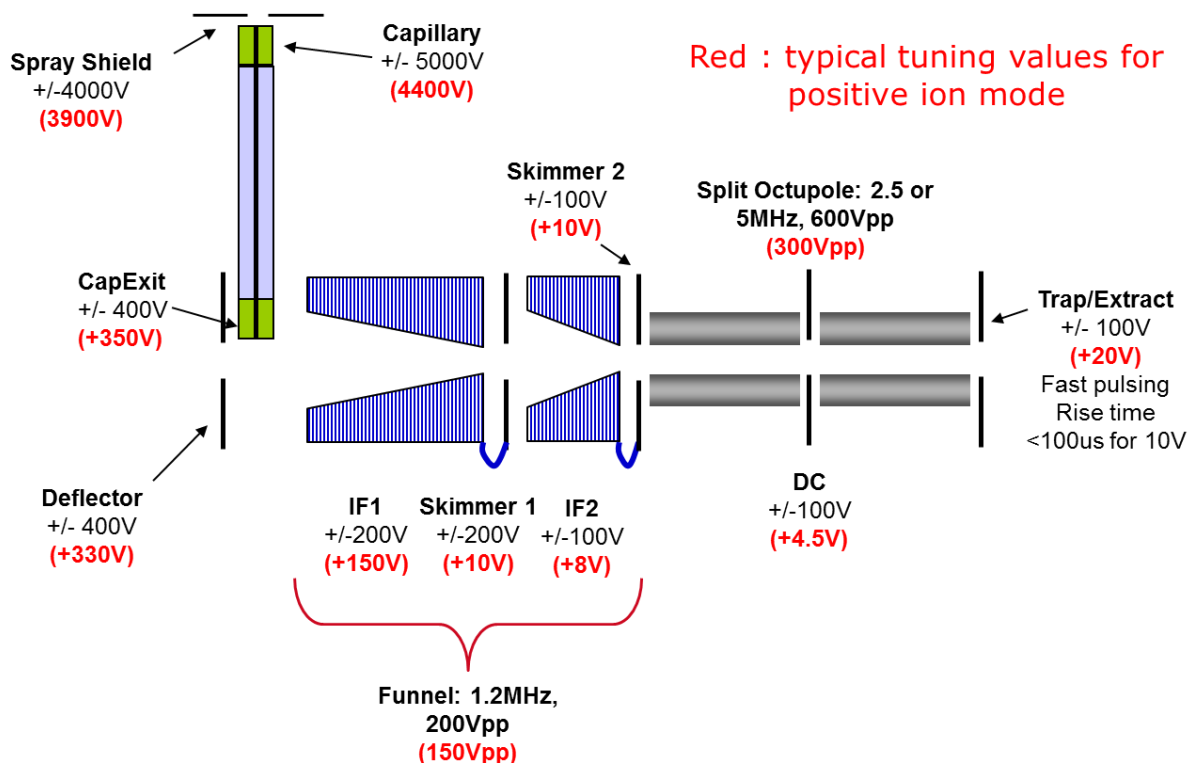
Figure 1.68: Schematic of source region of SolariX with typical tuning parameters in red. Courtesy of Bruker Daltonics, Bremen, Germany.

The quadrupole can either transmit ions or filter ions depending on the parameters used

and the mode desired, for standard MS experiments all ions are transmitted through the quadrupole

(ideally, in practice there is a small discrimination of transmittable m/z range due to the frequency of

the DC current switching and the amplitude applied to the rods). Ions are then passed into the

hexapole collision cell (F), which is populated with an adjustable low pressure of Argon gas, which

aids in ion cooling and accumulation/storage. During CAD MS/MS experiments, the quadrupole can

be re-tuned to filter ions of interest, which can then be accelerated into the collision cell containing

the Argon gas, causing CAD. Ions (regardless of transmission) can be continuously accumulated in

the hexapole collision cell in order to increase the ion population eventually transmitted to the ICR

cell for detection, thus providing larger packets of ions and enhanced detected signals. The Continuous Accumulation of Selected Ions (CASI) is a Bruker proprietary mechanic for accumulation of quadrupole-filtered ions within the hexapole collision cell.

Ions accumulated in the hexapole collision cell can also be subjected to ion-ion reactions, such as ETD and ETnoD (above) via generation of electron transfer reagents from the nCI source (G). ETD type reactions (including ETnoD) are dual-trapping experiments requiring confinement of both positively charged and negatively charged species, and the resulting fragments. Some instruments have also been modified to perform laser-based dissociations within accumulations cells such as hexapoles.[95] The SolariX is also equipped with a beam valve (H) used to isolate the source/accumulation regions of the instrument, and the ultra-high vacuum region, the beam valve of course needs to be open during normal operation, but closed when not in use/ for maintenance.

Once the ions have been accumulated/dissociated as desired they are transferred through hexapole transfer optics (I) to the ICR cell (J), they pass through two entrance lenses and through the front cell trapping plate and into the homogenous magnetic field region, at which point they are confined by raising the front trapping plate to an equal polarity of ions and a similar magnitude to the back trapping plate (usually 0.35-1V for the infinity cell).

Once in the infinity cell ions will precess according to their natural cyclotron frequency, perturbed by the electric fields, and can be excited using a broadband/narrowband excitation and/or be isolated/dissociated using in-cell-isolations/ in-cell MS/MS, described above. Equipped with a hollow-electron-emitting cathode (K) and continuous-wave 25W 10.6um IR laser (L), the SolariX is able to achieve a range of electron based dissociations (including, but not limited to ECD, hot-ECD, EID, EDD, ni-ECD, EIEIO, and EI).[134] The ICR cell is also equipped with a pulsed gas valve enabling in-cell CAD, SORI-CAD, and even CRAFTI-ion mobility spectrometry inside the ICR cell with careful operation.[149,150]

1.9 Liquid/Gas Chromatography-Mass Spectrometry:

Chromatography has been used for decades to separate mixtures of chemical species into either less complex samples or ideally pure, isolated species which can then be analysed using various analytical techniques. Chromatography is based on exploiting the different interactions (retentions) of compounds with a stationary phase while passing over a mobile phase which will slowly partition components out of the stationary phase and move them through the chromatograph, as a result compounds can be separated based on their interaction with the stationary phase and mobile phase combination (phase partitioning).

Gas Chromatography (GC), the older of the pair mentioned, has been used for many years to separate volatile chemical mixtures capable of easily transitioning to the gas phase for separation. GC tends to utilise very long (10-50m) stationary phases (columns), usually functionalised silica, and use inert gas as a mobile phase, e.g. Helium. Gas chromatography is a very fast chromatographic technique and often separates compounds on the minutes time scale. The column is usually heated in a temperature-controlled oven to force the elution of different analytes at certain times within the run. As a result a "heat ramp" can be used to start from a relatively low temperature (~50°C) which will cause elution of extremely volatile species that have limited interaction with the stationary phase, and can be ramped up to high temperatures (~300°C) to cause elution of species with a very high interaction with the stationary phase, and with all species in between eluting at some point during the ramp. Many detectors for GC have been used over the years, though the contemporary detector is known as a Flame Ionisation Detection. Flame ionisation detectors work on the principal that all species exiting the GC column are exposed to high-temperature flame, which will ionise combustible analytes, the ions move between the cathode (flame source) and towards an anode (detector plate) the current generated by movement of ions between the two electrode is then recorded and produces an intensity trace directly proportional to the ion concentration

produced via combustion.[151] This does not necessarily provide a uniform response for each chemical species, as obviously some species are more easily ionised via combustion than others.
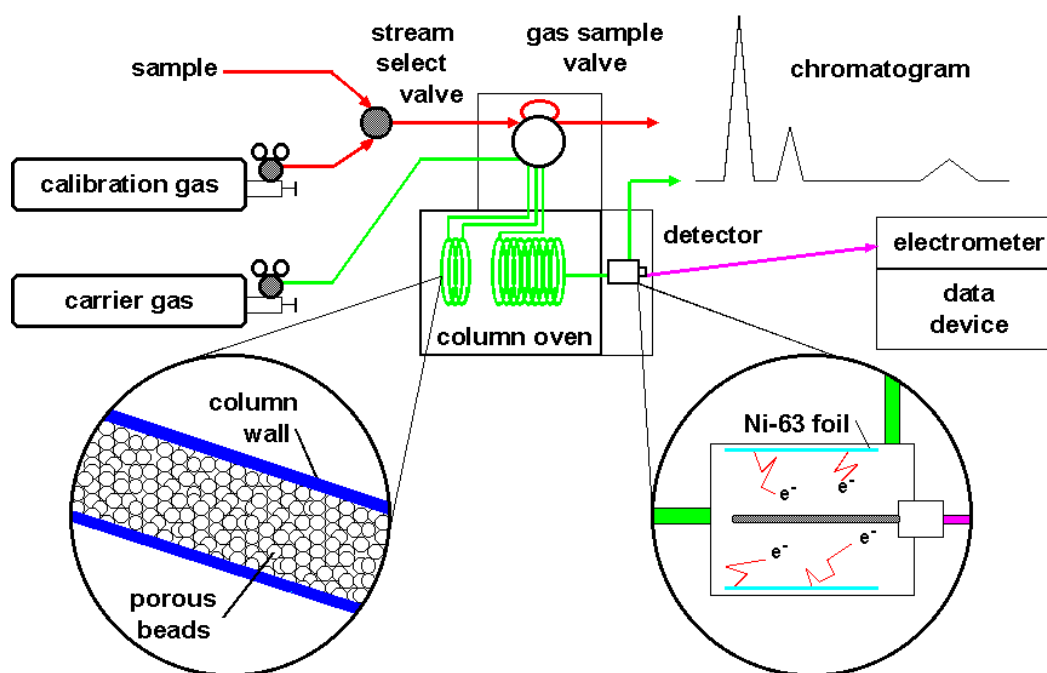
Figure 1.69: Schematic representation of a Gas Chromatograph (GC) showing the various components and functionalities, the presented detector is an electron capture detector rather then the FID described above . Image taken from http://www.esrl.noaa.gov/gmd/hats/insitu/insitu.html accessed 15/04/2016.

Though flame ionisation detectors have been shown to produce very accurate detection of chemical species and allow quantitation of a great array of compounds, the detector is fundamentally a binary detection method; it will provide a signal response based upon the amount of analyte passing through it and undergoing effective combustion ionisation, but cannot provide any further information beyond abundance. Thus the coupling of GC to a detector which can provide a more in depth analysis and provide analyte-specific information, such as mass spectrometry, was a perfect partnership. GC-MS was first successfully achieved by McLafferty *et. al.*[152] using an in-house built ToF-MS at the Dow chemical company, and has since been used worldwide for the analysis of both known and unknown chemical species. As mentioned previously, GC is a fast chromatographic

117

separation technique, despite the use of very long GC columns, thus the MS coupled to the

chromatograph would ideally have a short duty cycle and thus fast scan rate in order to keep up with

the various eluting compounds and provide a good chromatographic resolution from the resulting

detection. Below is an example chromatogram from a GC-EI-Q-MS, showing various hexane

extracted alkane hydrocarbon species separated effectively on a GC column and detected by EI-

quadrupole-MS:
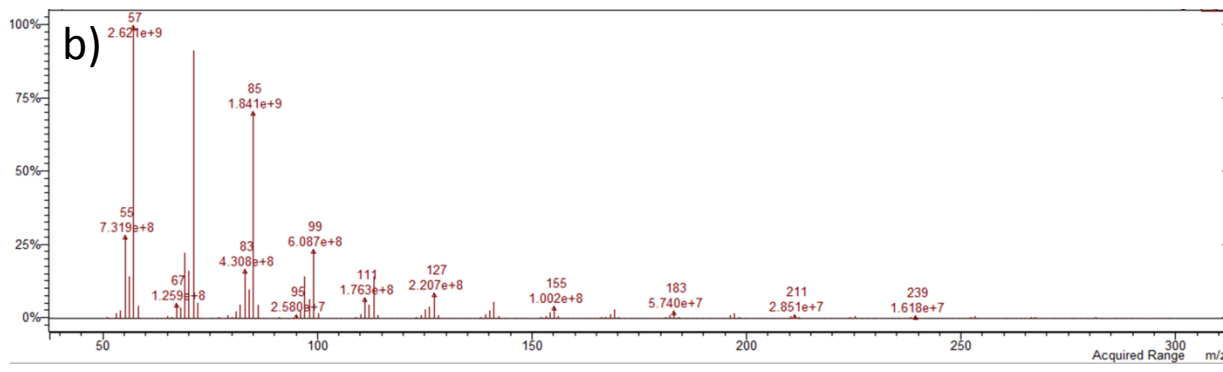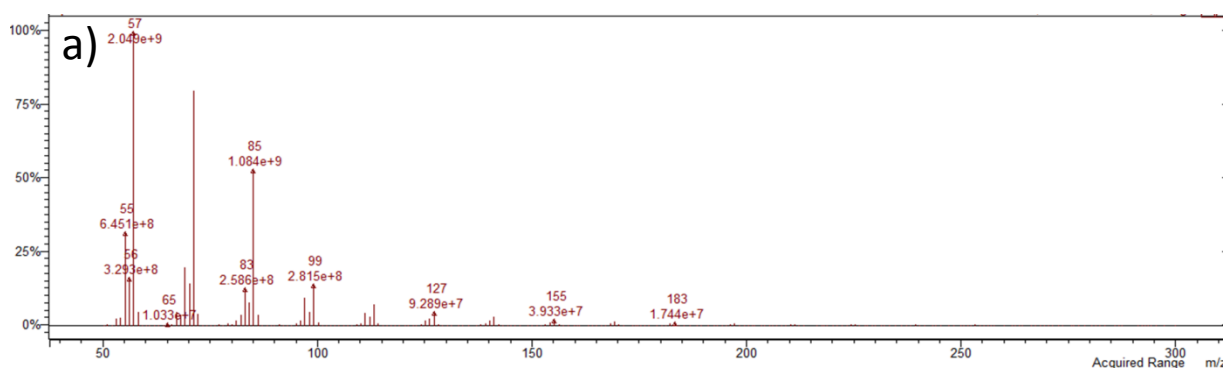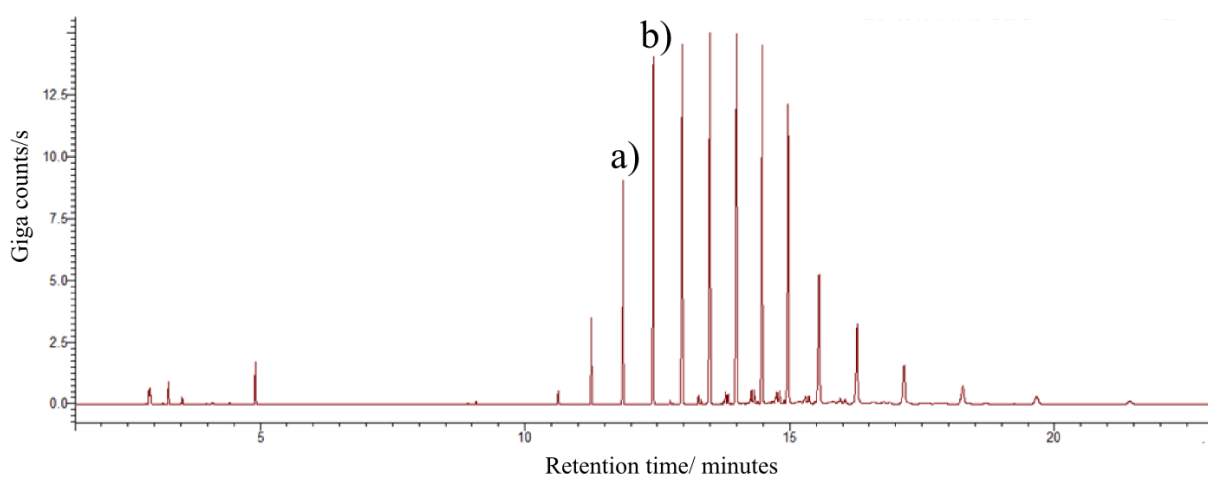
Figure 1.70: GC-EI-Q-MS chromatogram of hexane-extracted saturated hydrocarbons showing a broad distribution of components separated by GC and characterised by EI-quadrupole MS, each species will have its own corresponding EI mass spectrum enabling characterisation/pattern search for identification. EI-Q-MS spectra are shown for chromatogram peaks labelled a) and b) showing distinctly different peak distributions, yet common fragment peaks in areas.

GCMS has been shown to be very effective at the analysis of complex mixtures,[153] environmental samples,[154] pharmaceutical samples,[155] pesticides,[156] and even forensic samples,[20] where the GC can separate many species and allow the analysis of target compounds/all detectable compounds in one chromatographic-MS run. GC-MS can also be interfaced with various gas-phase ionisation sources and is not limited purely to EI.[157]

Liquid Chromatography (LC) as its name suggests is a slower, liquid based separation technique, utilising shorter chromatographic columns (~10-30cm), over longer periods (~30-120 minutes), and usually at a set temperature instead using combinations of different solvents to promote the elution of different chemical species. LC runs usually use a varying mixture of two or more solvent systems to alter the retention of compounds within the column, which in this case is usually functionalised silica beads with silyl-alkane functionality, though a great range of stationary phases are available. The most simple and widely used solvent system for LC separation is the acidifies water/Acetonitrile (ACN) system, in which a high percentage of acidified water (~95%) is used at the beginning of the chromatographic run, leaving a low percentage of acidified ACN (~5%), as the run progresses the proportion of ACN will increase at a given rate, causing elution of various compounds which prefer interaction with the higher ACN solvent system rather than the stationary phase. This simple elution methodology has been extremely effective at separating a wide range of compounds for LC, including small molecules,[158] metal complexes,[159] polymers,[154] and biomolecules[160–163] to name but a few. Although the Majority of LC analysis utilise UV/visible

detectors, like FID mentioned above, these detectors simply provide an abundance response for UV/visible interaction at a set wavelength, which can often be less than optimum for the target analytes, and so provides no further/structural data beyond the fact the compound has a chromophore which can interact with the given wavelength. As a result, and just as with GC above, the coupling of LC to MS was a perfect partnership, and with the advent of Electrospray, became easier than ever, as the simple connection of the LC outlet and tuning of the LC flowrate to be ESI-compatible is enough to couple an LC to an ESI-MS. LCMS has been effective at analysing a huge variety of solvent-soluble components, and is currently the leading technique for the analysis of complex biological mixtures, especially proteins and peptides for proteomic analysis.[164,163,165] An example of an LCMS layout and resulting separation of a simple peptide mixture is shown below in Figure 1.71, including the resulting chromatogram and example mass spectra for different time points, showing the elution of different peptides:
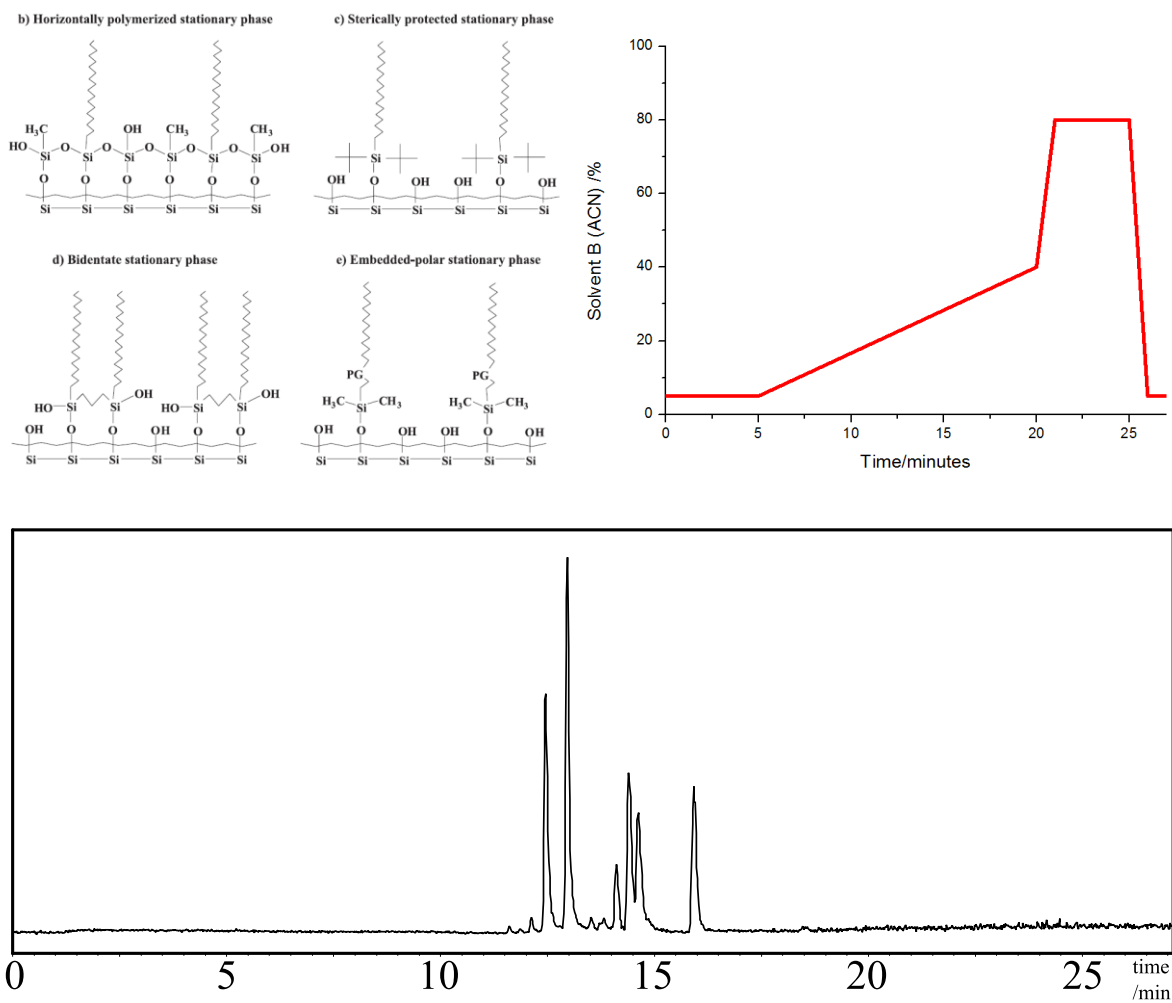
Figure 1.71: Clockwise from top-left: Various C18 bonded-phases often used in LCMS columns to separate peptides. The organic solvent gradient used to gradually elute compounds from the LC column (organic solvent proportion shown, other proportion is usually water/acidified water). The resulting LC-MS chromatogram showing a range of peptide species all eluting separately at different retention times and with different concentrations (proportional to peak area, and MS ionisation/transmission/detection efficiency).

References:

(1)   Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *26* (4), 489.

(2)   Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25* (2), 282.

(3)   Scigelova, M.; Hornshaw, M.; Giannakopulos, A.; Makarov, A. *Mol. Cell. Proteomics* **2011**, *10* (7), M111.009431.

(4)   Marshall, A. *Mass Spectrom. Rev.* **1998**, *17*, 1.

(5)   Marshall, a G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17* (1), 1.

(6)   Dunbar, R. C.; Chen, J. H.; Hays, J. D. *Int. J. Mass Spectrom. Ion Process.* **1984**, *57* (1), 39.

(7)   Amster, I. *J. mass Spectrom.* **1996**, *31*, 1325.

(8)   Pitsenberger, C. C.; Easterling, M. L.; Amster, I. J. *Anal. Chem.* **1996**, *68* (21), 3732.

(9)   Hendrickson, C. L.; Laude, D. a. *Anal. Chem.* **1995**, *67* (10), 1717.

(10)   Pitsenberger, C. C.; Easterling, M. L.; Amster, I. J. *Anal. Chem.* **1996**, *68* (21), 3732.

(11)   Qi, Y.; Witt, M.; Jertz, R.; Baykut, G.; Barrow, M. P.; Nikolaev, E. N.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2012**, *26* (17), 2021.

(12)   Aizikov, K.; Mathur, R.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (2), 247.

(13)   Shi, S. D.; Hendrickson, C. L.; Marshall, a G. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (20), 11532.

(14)   Qi, Y.; Barrow, M. P.; Li, H.; Meier, J. E.; Orden, S. L. Van; Thompson, C. J.; Connor, P. B. O. **2012**.

(15)   Qi, Y.; Barrow, M. P.; Orden, S. L. Van; Thompson, C. J.; Li, H.; Perez-hurtado, P.; O'Connor, P. B. **2011**, 8477.

(16)   Nikolaev, E. N.; Boldin, I. a; Jertz, R.; Baykut, G. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (7), 1125.

(17)   Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *J. Mass Spectrom.* **2005**, *40* (4), 430.

(18)   Marshall, A.; Guan, S. *Rapid Commun. Mass …* **1996**, *10*, 1819.

(19)   Barrow, M. P.; Witt, M.; Headley, J. V.; Peru, K. M. **2010**, *82* (9), 3727.

(20)   Collin, O. L.; Zimmermann, C. M.; Jackson, G. P. *Int. J. Mass Spectrom.* **2009**, *279* (2-3), 93.

(21)   Perry, R. H.; Cooks, R. G.; Noll, R. J. *Mass Spectrom. Rev.* **2008**, *27*, 661.

(22)   Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. *Anal. Chem.* **2008**, *80* (11), 3985.

(23)   Zhang, L.-K.; Rempel, D.; Pramanik, B. N.; Gross, M. L. *Mass Spectrom. Rev.* **2005**, *24* (2), 286.

(24)   Cooper, H. J.; Marshall, A. G. **2001**, No. 25, 5710.

(25)   Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *J. Am. Chem. Soc.* **1985**, *107* (26), 7893.

(26)   Marshall, A. G. *Int. J. Mass Spectrom.* **2000**, *200* (1-3), 331.

(27)   Heeren, R. M. a; Kleinnijenhuis, a J.; McDonnell, L. a; Mize, T. H. *Anal. Bioanal. Chem.* **2004**, *378* (4), 1048.

(28)   Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246* (1), 211.

(29)   Leach, F. E.; Norheim, R.; Anderson, G.; Pasa-Tolic, L. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (12), 2069.

(30)   Park, S.-G.; Anderson, G. A.; Navare, A. T.; Bruce, J. E. *Anal. Chem.* **2016**, *88*, 1162.

(31)   Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26* (October 1990), 514.

(32)   Beu, S. C.; Laude, D. a. *Anal. Chem.* **1992**, *64* (2), 177.

(33)   Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26* (October 1990), 514.

(34) Wei, J.; Antzutkin, O. N.; Filippov, A. V.; Iuga, D.; Lam, P. Y.; Barrow, M. P.; Dupree, R.; Brown, S. P.; O'Connor, P. B. *Biochemistry* **2016**, *55* (14), 2065.

(35) Kim, S.; Choi, M. C.; Kim, S.; Hur, M.; Kim, H. S.; Yoo, J. S.; Blakney, G. T.; Hendrickson, C. L.; Marshall, A. G. **2007**, *79* (10), 3575.

(36) Biemann, K. *Int. J. Mass Spectrom. Ion Phys.* **1982**, *45*, 183.

(37) Nikolaev, E. *J. mass Spectrom.* **1998**, *501* (33), 499.

(38) Munson M. S. B.; Field F. H. *J. Am. Chem. Soc.* **1966**, *88* (12), 2621.

(39) Taylor, G. *Proc. R. Soc. London A Math. Phys. Eng. Sci.* **1964**, *280* (1382), 383.

(40) Rayleigh, Lord. *Proc. R. Soc.* **1882**, *14*, 184.

(41) Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. *J. Phys. Chem. A* **2002**, *106* (42), 9957.

(42) Meyer, T.; Gabelica, V.; Grubmu ller, H.; Orozco, M. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3* (4), 408.

(43) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2.

(44) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2.

(45) Konermann, L.; Rodriguez, A. D.; Liu, J. *Anal. Chem.* **2012**, *84* (15), 6798.

(46) Konermann, L.; Douglas, D. J. *Biochemistry* **1997**, *36* (40), 12296.

(47) Wootton, C. a.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. *Dalt. Trans.* **2015**, *44* (8), 3624.

(48) Cassou, C. a; Sterling, H. J.; Susa, A. C.; Williams, E. R. *Anal. Chem.* **2013**, *85* (1), 138.

(49) Cassou, C. a; Williams, E. R. *Analyst* **2014**, *139* (19), 4810.

(50) Wei, J.; Bristow, A. W. T.; O'connor, P. B. *J. Am. Soc. Mass Spectrom.* **2014**, *26* (1), 166.

(51)   Yin, S.; Loo, J. a. *Int. J. Mass Spectrom.* **2011**, *300* (2-3), 118.

(52)   Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. a. *Anal. Chem.* **2014**, *86* (1), 317.

(53)   Marty, M. T.; Zhang, H.; Cui, W.; Blankenship, R. E.; Gross, M. L.; Sligar, S. G. *Anal. Chem.* **2012**, *84* (21), 8957.

(54)   Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68* (1), 1.

(55)   Mandal, M. K.; Yoshimura, K.; Saha, S.; Ninomiya, S.; Rahman, M. O.; Yu, Z.; Chen, L. C.; Shida, Y.; Takeda, S.; Nonami, H.; Hiraoka, K. *Analyst* **2012**, *137* (20), 4658.

(56)   Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60* (20), 2299.

(57)   O'Connor, P. B.; Budnik, B. A.; Ivleva, V. B.; Kaur, P.; Moyer, S. C.; Pittman, J. L.; Costello, C. E. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (1), 128.

(58)   Budnik, B. A.; Moyer, S. C.; Pittman, J. L.; Ivleva, V. B.; Sommer, U.; Costello, C. E.; O'Connor, P. B. *Int. J. Mass Spectrom.* **2004**, *234* (1-3), 203.

(59)   Laiko, V. V.; Baldwin, M. A.; Burlingame, A. L. *Anal. Chem.* **2000**, *72* (4), 652.

(60)   Hillenkamp, F.; Karas, M. **2007**.

(61)   Lutomski, C. a.; El-Baba, T. J.; Woodall, D. W.; Foley, C. D.; Kumar, R.; Manly, C. D.; Wang, B.; Liu, C.-W.; Harless, B. M.; Imperial, L. F.; Inutan, E. D.; Trimpin, S. *Int. J. Mass Spectrom.* **2015**, *377*, 532.

(62)   Montaudo, G.; Samperi, F.; Montaudo, M. S. *Prog. Polym. Sci.* **2006**, *31* (3), 277.

(63)   Griffiths, M. T.; Da Campo, R.; O'Connor, P. B.; Barrow, M. P. *Anal. Chem.* **2014**, *86* (1), 527.

(64)   Smith, D. F.; Robinson, E. W.; Tolmachev, A. V; Heeren, R. M. a; Paša-Tolić, L. *Anal. Chem.* **2011**, *83* (24), 9552.

(65)   Wang, B.; Tisdale, E.; Trimpin, S.; Wilkins, C. L. **2014**.

(66)   Trimpin, S.; Thawoos, S.; Foley, C. D.; Woodall, D. W.; Li, J.; Inutan, E. D.; Stemmer, P. M.

         Methods **2016**, 1.

(67)   Inutan, E. D.; Trimpin, S. *Mol. Cell. Proteomics* **2013**, *12* (3), 792.

(68)   Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Craig, M.; Meng, C. K. A. I.; Mann, M.;

         Whitehouse, C. M. **2013**, *246* (4926), 64.

(69)   Marshall, A. G.; Guan, S. **1995**, *363*, 397.

(70)   Jebanathirajah, J. a; Pittman, J. L.; Thomson, B. a; Budnik, B. a; Kaur, P.; Rape, M.; Kirschner,

         M.; Costello, C. E.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 1985.

(71)   McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X.; Infusini, G.; Jiang, H.; Kong, X.; Begley, T. P.

         FEBS J. **2007**, *274* (24), 6256.

(72)   Kelly, R. T.; Tolmachev, A. V; Page, J. S.; Tang, K.; Smith, R. D.; Box, P. O. **2010**, 294.

(73)   Kaiser, N. K.; Mckenna, A. M.; Savory, J. J.; Hendrickson, C. L.; Marshall, A. G. **2013**.

(74)   Kaczorowska, M. a; Hotze, A. C. G.; Hannon, M. J.; Cooper, H. J. *J. Am. Soc. Mass Spectrom.*

         2010, *21* (2), 300.

(75)   McLafferty, F. W. *Int. J. Mass Spectrom.* **2001**, *212* (1-3), 81.

(76)   Marshall, A. G.; Hendrickson, C. L. **2002**, *215*, 59.

(77)   Koning, L. J. De; Nibbering, N. M. M.; Ordenb, S. L. Van; Laukienb, F. H. *Int. J. Mass Spectrom.*

         Ion Process. **1997**, *166*.

(78)   McLafferty, F. W.; Tureček, F. *Interpretation of Mass Spectra*, 4th ed.; USB, 1993.

(79)   Wills, R. H.; Tosin, M.; Connor, P. B. O. *Anal. Chem.* **2012**, *84*, 8863.

(80)   Scigelova, M.; Hornshaw, M.; Giannakopulos, A.; Makarov, A. *J. Mass spectrome* **2011**, *10* (7),

         M111.009431.

(81)   Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* **2005**, *24* (2), 135.

(82)   Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* **2003**, *22* (3), 158.

(83)   Seidler, J.; Zinn, N.; Boehm, M. E.; Lehmann, W. D. *Proteomics* **2010**, *10* (4), 634.

(84)   Tang, X. J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65* (20), 2824.

(85)   Lopez-Clavijo, A. F.; Duque, C.; Wootton, C. A.; Warwick, O'Connotr P. B.

(86)   Li, H.; Lin, T.; Orden, S. Van. *Anal. Chem.* **2011**, 9507.

(87)   Xu, F.; Xu, Q.; Dong, X.; Guy, M.; Guner, H.; Hacker, T. a.; Ge, Y. *Int. J. Mass Spectrom.* **2011**, *305* (2-3), 95.

(88)   Zubarev, R. a. *Curr. Opin. Biotechnol.* **2004**, *15* (1), 12.

(89)   Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. *Mol. Cell. Proteomics* **2012**, No. August, 1.

(90)   van Duijn, E.; Simmons, D. A.; van den Heuvel, R. H. H.; Bakkes, P. J.; van Heerikhuizen, H.; Heeren, R. M. A.; Robinson, C. V.; van der Vies, S. M.; Heck, A. J. R. *J. Am. Chem. Soc.* **2006**, *128* (14), 4694.

(91)   Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. *Chem. Biol.* **2006**, *13* (6), 597.

(92)   Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66* (18), 2809.

(93)   Sarkar, P. K.; Prajapati, P. K.; Shukla, V. J.; Ravishankar, B.; Choudhary, A. K. *Indian J. Exp. Biol.* **2009**, *47* (12), 987.

(94)   Stephenson, J. L.; Booth, M. M.; Shalosky, J. A.; Eyler, J. R.; Yost, R. A. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (10), 886.

(95)   Xu, Z.; Shaw, J. B.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (2), 265.

(96)   Gabelica, V.; Tabarin, T.; Antoine, R.; Rosu, F.; Compagnon, I.; Broyer, M.; De Pauw, E.; Dugourd, P. *Anal. Chem.* **2006**, *78* (18), 6564.

(97)   Larraillet, V.; Antoine, R.; Dugourd, P.; Lemoine, J. *Anal. Chem.* **2009**, *81* (20), 8410.

(98)   Zubarev, R. *J. Am. Chem. Soc.* **1998**, *7863* (16), 3265.

(99)   Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125* (29), 8949.

(100) Cerda, B. A.; Horn, D. M.; Breuker, K.; Carpenter, B. K.; Mclafferty, F. W. **1999**, *338*, 335.

(101) Kruger, N. a.; Zubarev, R. a.; Carpenter, B. K.; Kelleher, N. L.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *182-183*, 1.

(102) Ganisl, B.; Breuker, K. *ChemistryOpen* **2012**, *1* (6), 260.

(103) Simons, J. *Chem. Phys. Lett.* **2010**, *484* (4-6), 81.

(104) Simons, J.; Ledvina, A. R. *Int. J. Mass Spectrom.* **2012**, *330-332*, 85.

(105) Turecek, F. *J. Am. Chem. Soc.* **2003**, *125* (19), 5954.

(106) Simons, J.; Ledvina, A. R. *Int. J. Mass Spectrom.* **2012**, *330-332*, 85.

(107) Liu, J.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2012**, *330-332*, 174.

(108) Simons, J. *J. Am. Chem. Soc.* **2010**, *132* (20), 7074.

(109) Sobczyk, M.; Simons, J. *Int. J. Mass Spectrom.* **2006**, *253* (3), 274.

(110) Syrstad, E. a; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (2), 208.

(111) Locke, S. J.; Leslie, A. D.; Melanson, J. E.; Pinto, D. M. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1525.

(112) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, Ü. a; Tsybin, Y. O. *Chem. Soc. Rev.* **2013**, *42*

(12), 5014.

(113) Li, X.; Cournoyer, J. J.; Lin, C.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (10), 1514.

(114) Sleno, L.; Volmer, D. a. *J. Mass Spectrom.* **2004**, *39* (10), 1091.

(115) Tsybin, Y. O.; He, H.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2007**, *79* (20), 7596.

(116) Baba, T.; Campbell, J. L.; Le Blanc, J. C. Y.; Hager, J. W.; Thomson, B. A. *Anal. Chem.* **2015**, *87* (1), 785.

(117) Voinov, V. G.; Beckman, J. S.; Deinzer, M. L.; Barofsky, D. F. *Rapid Commun. Mass Spectrom.* **2009**, *23* (18), 3028.

(118) Voinov, V. G.; Deinzer, M. L.; Barofsky, D. F. *Rapid Commun. Mass Spectrom.* **2008**, *22* (19), 3087.

(119) Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. *Biochem. Biophys. Res. Commun.* **2014**, *445* (4), 683.

(120) Bakhtiar, R.; Guan, Z. *Biochem. Biophys. Res. Commun.* **2005**, *334* (1), 1.

(121) Cooper, H. J.; Hudgins, R. R.; Marshall, A. G. **2004**, *234*, 23.

(122) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, Ü. a; Tsybin, Y. O. *Chem. Soc. Rev.* **2013**, *42* (12), 5014.

(123) Qi, Y.; Liu, Z.; Li, H.; Sadler, P. J.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2013**, *27* (17), 2028.

(124) Soulby, A. J.; Heal, J. W.; Barrow, M. P.; Roemer, R. a.; O'Connor, P. B. *Protein Sci.* **2015**, *24* (5), 850.

(125) Li, H.; Zhao, Y.; Phillips, H. I. a; Qi, Y.; Lin, T.-Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**,

83 (13), 5369.

(126) Li, H.; Snelling, J. R.; Barrow, M. P.; Scrivens, J. H.; Sadler, P. J.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1217.

(127) Hartinger, C. G.; Tsybin, Y. O.; Fuchser, J.; Dyson, P. J. *Inorg. Chem.* **2008**, *47* (1), 17.

(128) Groessl, M.; Tsybin, Y. O.; Hartinger, C. G.; Keppler, B. K.; Dyson, P. J. *J. Biol. Inorg. Chem.* **2010**, *15* (5), 677.

(129) Cooper, H. J.; Hudgins, R. R.; Håkansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (3), 241.

(130) Hurtado, P. P.; O'Connor, P. B. **2012**, No. February, 609.

(131) Sargaeva, N. P.; Lin, C.; O'Connor, P. B. *Anal. Chem.* **2009**, *81* (23), 9778.

(132) Haselmann, K. F.; Budnik, B. A.; Zubarev, R. A. **2000**, *2246*, 2242.

(133) Yao, C.; Fung, Y. M. E.; Hayakawa, S.; Hashimoto, M.; Matsubara, H. **2009**, 7347.

(134) Qi, Y.; Volmer, D. A. *Analyst* **2016**, 794.

(135) Budnik, B.; Haselmann, K.; Zubarev, R. *Chem. Phys. Lett.* **2001**, *342* (July), 299.

(136) Taucher, M.; Breuker, K. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 918.

(137) Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (2), 234.

(138) Jones, C. J.; Larive, C. K. *Nat. Chem. Biol.* **2011**, *7* (11), 758.

(139) Yoo, H. J.; Wang, N.; Zhuang, S.; Song, H.; Håkansson, K. *J. Am. Chem. Soc.* **2011**, *133* (42), 16790.

(140) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528.

(141) Gunawardena, H. P.; Gorenstein, L.; Erickson, D. E.; Xia, Y.; McLuckey, S. a. *Int. J. Mass Spectrom.* **2007**, *265* (2-3), 130.

(142) Skinner, O. S.; Catherman, A. D.; Early, B. P.; Thomas, P. M.; Compton, P. D.; Kelleher, N. L. *Anal. Chem.* **2014**, 2.

(143) Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 343.

(144) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. **2007**, *104* (7).

(145) Ahlf, D. R.; Compton, P. D.; Tran, J. C.; Early, B. P.; Thomas, P. M.; Kelleher, N. L. *J. Proteome Res.* **2012**, *11* (8), 4308.

(146) Lermyte, F.; Sobott, F. *Proteomics* **2015**, *15* (16), 2813.

(147) Hunt, D. F.; Shabanowitz, J.; Bai, D. L. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (7), 1256.

(148) Zhao, C.; Xie, B.; Chan, S.-Y.; Costello, C. E.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (1), 138.

(149) Yang, F.; Jones, C. a.; Dearden, D. V. *Int. J. Mass Spectrom.* **2015**, *378*, 143.

(150) Yang, F.; Voelkel, J. E.; Dearden, D. V. *Anal. Chem.* **2012**, *84* (11), 4851.

(151) McWILLIAM, I. G.; DEWAR, R. A. *Nature* **1958**, *181* (4611), 760.

(152) Gohlke, R. S.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom. 4* (5), 367.

(153) Barrow, M. P.; Peru, K. M.; Headley, J. V. *Anal. Chem.* **2014**, *86* (16), 8281.

(154) Te, U. **2010**, 9666.

(155) Rocha, S. M.; Caldeira, M.; Carrola, J.; Santos, M.; Cruz, N.; Duarte, I. F. *J. Chromatogr. A* **2012**, *1252*, 155.

(156) Fernandes, V. C.; Vera, J. L.; Domingues, V. F.; Silva, L. M. S.; Mateus, N.; Delerue-Matos, C. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (12), 2187.

(157) Portolés, T.; Mol, J. G. J.; Sancho, J. V; Hernández, F. *Anal. Chem.* **2012**, *84* (22), 9802.

(158) Zheng, X.; Kang, A.; Dai, C.; Liang, Y.; Xie, T.; Xie, L.; Peng, Y.; Wang, G.; Hao, H. *Anal. Chem.* **2012**, *84* (22), 10044.

(159) Kodali, P.; Chitta, K. R.; Landero Figueroa, J. a; Caruso, J. a; Adeoye, O. *Metallomics* **2012**, *4* (10), 1077.

(160) Pasa-Tolić, L.; Masselon, C.; Barry, R. C.; Shen, Y.; Smith, R. D. *Biotechniques* **2004**, *37* (4), 621.

(161) Shen, Y.; Zhao, R.; Belov, M. E.; Conrads, T. P.; Anderson, G. a; Tang, K.; Pasa-Tolić, L.; Veenstra, T. D.; Lipton, M. S.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **2001**, *73* (8), 1766.

(162) Tipton, J. D.; Tran, J. C.; Catherman, A. D.; Ahlf, D. R.; Durbin, K. R.; Lee, J. E.; Kellie, J. F.; Kelleher, N. L.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2012**, *84* (5), 2111.

(163) Tian, Z.; Zhao, R.; Tolić, N.; Moore, R. J.; Stenoien, D. L.; Robinson, E. W.; Smith, R. D.; Paša-Tolić, L. *Proteomics* **2010**, *10* (20), 3610.

(164) Mann, M. **2006**, *7* (December), 952.

(165) López-Ferrer, D.; Petritis, K.; Robinson, E. W.; Hixson, K. K.; Tian, Z.; Lee, J. H.; Lee, S.-W.; Tolić, N.; Weitz, K. K.; Belov, M. E.; Smith, R. D.; Pasa-Tolić, L. *Mol. Cell. Proteomics* **2011**, *10* (2), M110.001479.

Chapter 2: Introduction to Anti-cancer metallodrugs

2.1 Cancer:

Cancer is a generic term for a wide array of conditions all connected with the same fundamental symptom – uncontrolled cell proliferation (growth and replication). Excessive growth and replication of cells is detrimental to individual cell function and is unsustainable for the organism as a whole and so is usually carefully controlled and limited by cellular proteins and biological mechanisms. Cancer arises from damage/alteration/coding problems within the DNA of the cell itself. This can occur in a number of ways, including, but not limited to chemical-induced damage (e.g. by carcinogens), radiation induced damage/degradation of DNA, or problems arising during DNA replication/transcription.[1] A more recently recognised cause of cancer is via virus infection, the human papilloma virus currently causes ~70% of cervical cancer cases in women. Damage to host cell DNA can cause the cell to proliferate at a much accelerated rate than usual, demanding more resources from the host organism, and creating asymmetric growth in that area of the host, often resulting a large collection of cancerous cells, known as a tumour. Cancerous cells can often migrate from their initial creation point and move to other areas of the host, where they will then continue to proliferate and create secondary tumours, a process known as metastasis. The spread of cancerous tumours throughout a host organism can put a huge amount of strain on biological processes both from the extreme resource demand the cancerous cells present due to accelerated growth and replication, but also due to the physical stress of increased mass in delicate internal areas which can disrupt normal biological function. Finally cancer can cause a great change in cellular function as the cells previously tasked to specific function (e.g. liver cell function of metabolising toxins) will no longer take place effectively as the previously functioning cell is now damaged, its usual cellular processes interrupted, and proliferating at an uncontrolled rate, putting strain on it and the cells around it.
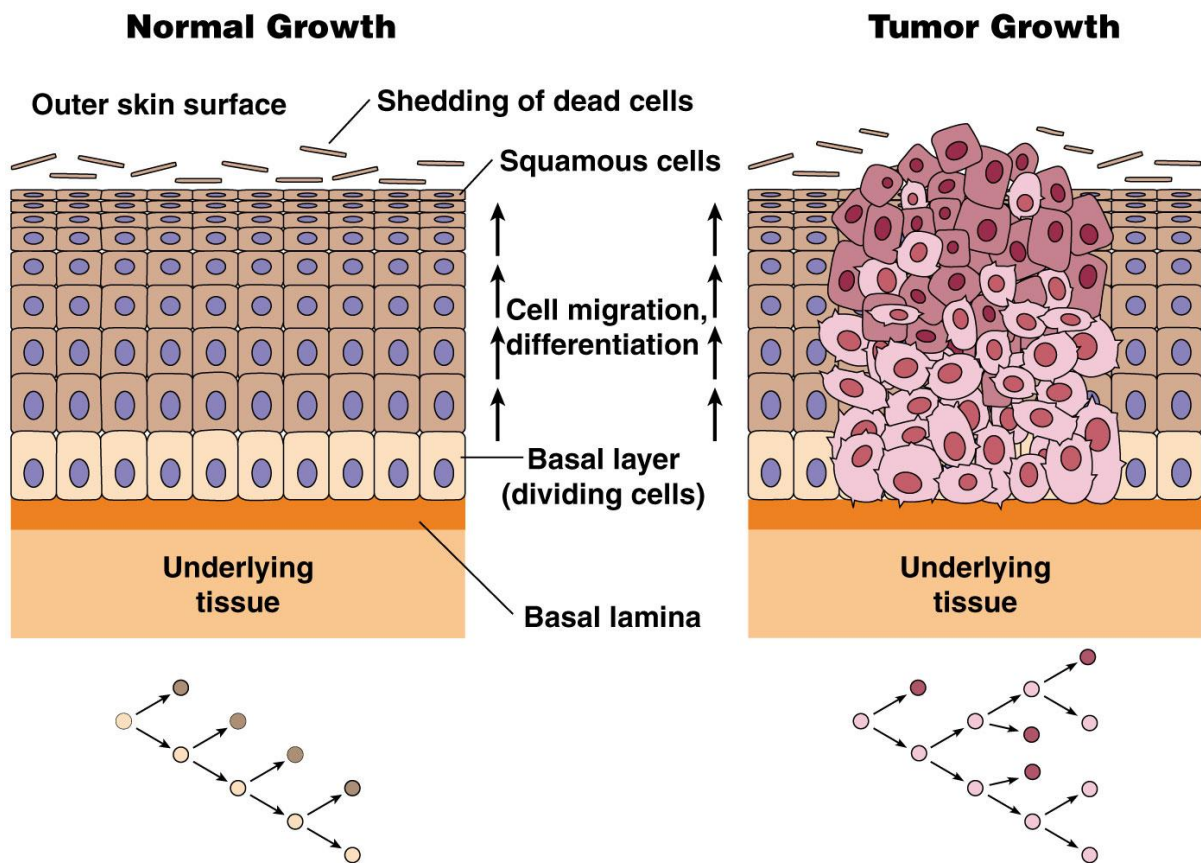
Figure 2.1: Normal (left) and uncontrolled (cancerous) cell growth (right) after DNA damage. Large growth of cancerous cells (tumour) can put pressure on nearby cells and increase stress and/or impair function. Sections of tumour may break off and relocate to other areas of the host (metastasis). Reproduced from http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-24/CB24.html, (accessed 23/02/2016).

Unfortunately Cancer is an inevitable cellular event, and is coded directly into our genome; every time a cell divides and replicates DNA there is a chance of causing DNA damage, or the creation of damaged DNA,[2] as a result due to the countless times host cells divide during a lifetime even if this is an extremely rare event, it will happen eventually. Though external factors such as exposure to carcinogens, repeated damage to cells/tissue forcing increased replication, and exposure to DNA damaging radiation (such as UV/gamma radiation) can lead to increased risk/acceleration of cancer development.[3] The development of cancer, risk factors, and especially its biological causes are a areas of intense and widespread research. Though these phenomena are not completely

understood, great progress is being made.[4,5] There are a number of genetic factors also affecting

cancer development and risk, as cancer is predominantly a DNA-related condition, hosts with certain

DNA features, mismatches, or sequences may be particularly susceptible to certain external factors,

or may have an accelerated time frame for cancer development than others.
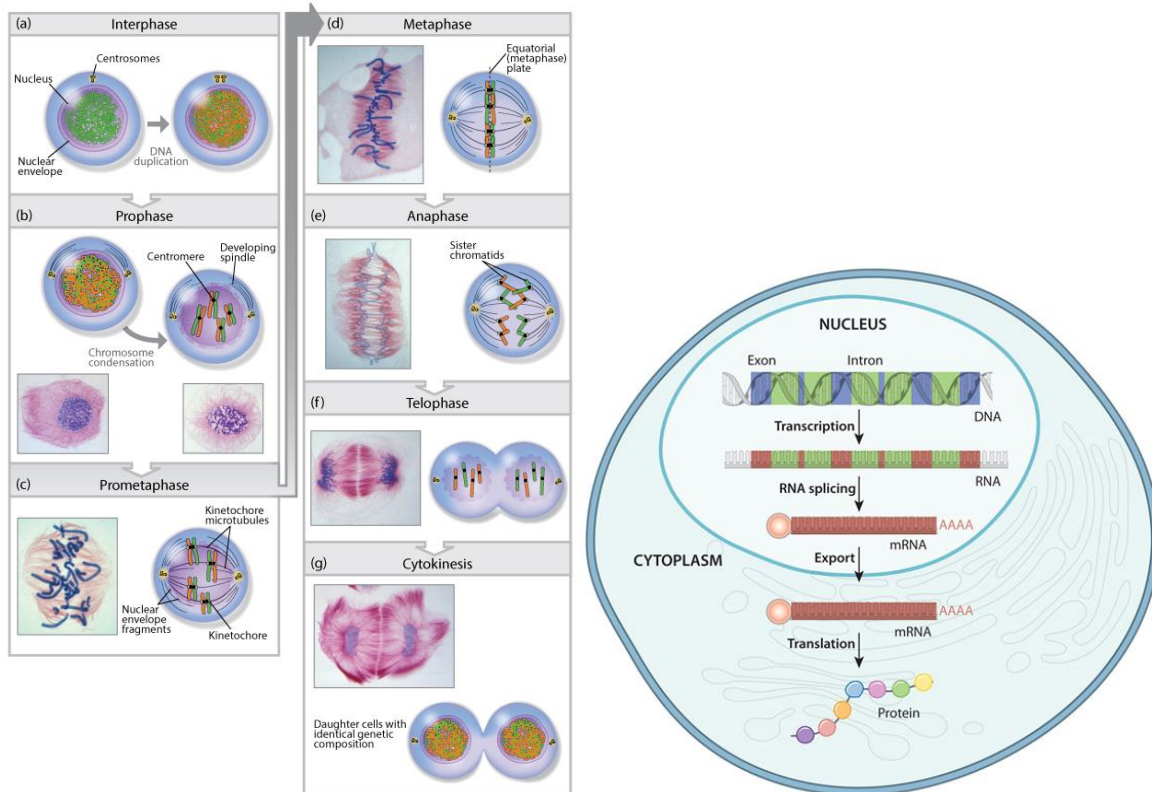
Figure 2.2: representations of the central dogma of cell division (mitosis) (left) and gene

expression (right) showing how cells divide and use DNA/RNA to produce cellular components,

such as peptides and proteins. Both reproduced from http://www.nature.com/scitable (accessed

12/08/2016).

Cancer seems to be a condition common to all eukaryotes, occurring in many of the same ways as

humans, and similarly accelerated due to exposure to carcinogens and/or radiation. As a result many

animal models, especially mice, are used for both the study of native cancer, induced cancer, and

cancer graphs (xenografts) from human to secondary host (such as mice).[6]

2.2: Treatment:

Cancer, being intrinsically DNA-related, has been historically treated medically and chemically by DNA-targeting reagents, which ideally will attack cancerous DNA, while leaving normal, non-cancerous DNA unharmed.[7] The ratio of cancerous to non-cancerous cells targeted is known as a drug's selectivity.[8] These DNA-targeting reagents come in a variety of forms, using small molecules,[1] macromolecules,[9] and even heavy/transition metal complexes.[10]

The concept behind most metal-based anti-cancer drugs is to bind to target cancerous DNA, within cancer cells, and induce a chemical change, either covalent bonding, crosslinking, or deformation of structure.[11–13] After the DNA has been modified the DNA will no longer be "readable" by the transcription proteins and will not be transcribed, thus preventing further cancerous DNA replication and thus replication of the cancerous cells. Without the ability to replicate the cancerous cells will live out their particularly shortened lifespan (due to excessive resource consumption) and die off. Though this is the ideal situation, there is very little difference between a cancerous cell and a non-cancerous/healthy cell in a specific organism; they have very similar cell function and contents (though increasingly there are studies into the subtle differences[1]),and for the most part they have extremely similar DNA encoding, aside from the often small areas causing damage and triggering excessive proliferation.[9] As a result although creating anti-cancer drugs with great potency and DNA binding can be readily achieved,[14] creating compounds with good selectivity is very challenging, and so many anti-cancer compounds which attack and kill cancerous cells also attack and kill healthy host cells, producing many of the extensive side effects experienced during chemotherapy.[15]

2.3 Metal-based chemotherapy:

Above all others, metal-based chemotherapy has become the dominant method for treating cancer by attacking cancerous DNA and preventing proliferation.[11,13,16] Specifically it is the Platinum based chemotherapeutics which have made the largest impact on cancer therapy in the last few decades. So much so that over 50% of all cancer treatments worldwide now contain a Pt(II) metal-based drug (metallodrug). [17,18]

The first FDA approved and commercialised platinum drug was Cisplatin (Cis(diaminodichlorido)platinum (CDDP)).[18] Cisplatin was shown to attack cancerous DNA by binding to exposed Guanine nucleobases, causing deformation of DNA tertiary structure, preventing proper transcription and eventually preventing proliferation. Cisplatin was also shown to produce intra and inter-strand crosslinks within DNA, with the latter preventing transcription very effectively,[18] example shown in Figure 2.3 below:
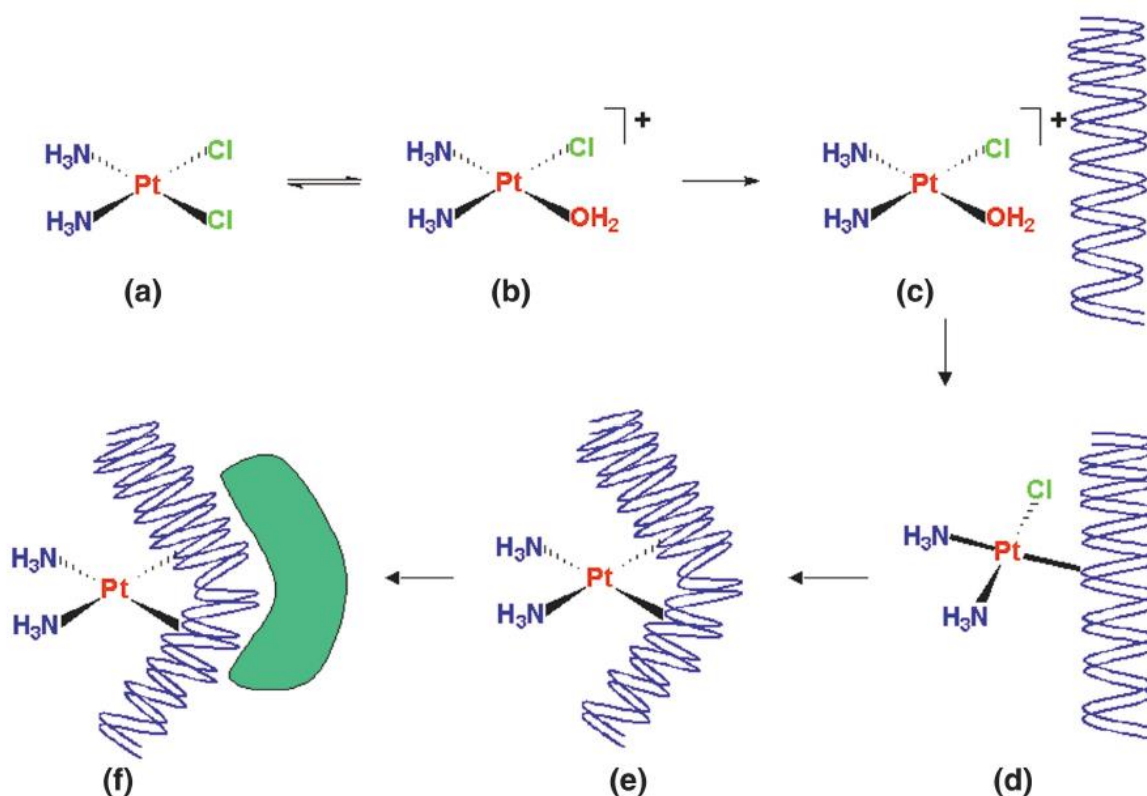
Fig. 1 The sequence of events in cisplatin binding to DNA.

Figure 2.3: Cisplatin mechanism of action against DNA. Reproduced from Hambley *et al.*[19]

Cisplatin has proven to be extremely effective over the years at combatting cancers, especially testicular, cervical, and bladder cancers. Cisplatin/Metal based anti-cancer agents exhibit many advantages over their purely organic counterparts:

- High potencies – Cisplatin possesses an IC_{50} of 1.2 µM (+/- 0.2)[20]

- High selectivity for cancer cells

- Tuneable activity – due to different ligand configurations

- Low toxicity compared to many small organic molecules

- Easy to trace – as a transition metal that is not usually present

However extensive use of Cisplatin has resulted in the observation of an array of detrimental side effects to patients, including liver failure, kidney damage, hair loss, dramatic weight loss, and many more.[21–23] The widespread use of Cisplatin has also caused the observation and rise in Cisplatin resistant strains of cancer.[24] As the body is subjected to the toxic treatment it will attempt to adapt

and remove the active complexes from the body in order to protect its cells, healthy or cancerous, unfortunately the body/cells will protect both. Though Cisplatin was a major medical and chemical breakthrough in anti-cancer research its current side effects are great, and its effectiveness is decreasing relative to more contemporary treatments. However Cisplatin is still the most widely used anti-cancer drug available and is used to treat an array of cancers.[15,22] Though many other types of Pt(II) drugs are currently used/available, including Oxaliplatin, Strataplatin, and PicoPlatin.[11,25,26] The analogous compound transplatin, an isomer of Cisplatin was unfortunately found to be much less active than its cis counterpart.[22]

Studies into the movement, transformation and eventual mechanism of action of Cisplatin and other platinum(II) drugs have shown that one or more of the chlorido ligands can be displaced while inside the body/cell, producing a much more active aqua-Pt(II) complex, this aqua/di-aqua species can then bind to cancerous DNA to cause deformations and prevent proliferation/induce cell death.[22,27] However it has also been shown by a variety of studies that only a very small percentage of the administered Pt(II) drug actually makes it to its DNA target.[15,28] Unfortunately while DNA functional groups such as guanine nitrogen groups present very reactive and viable ligands for Cisplatin, there are a plethora of other molecules and functional groups within cells, blood, and extracellular fluid which will also readily bind to Cisplatin and any reactive metal species. The binding of Cisplatin to cellular/blood/extra cellular components is so great the it has been shown that 90-98% of Cisplatin becomes protein bound within just one day of injection. [15,28] It is currently believed that many of the side effects experienced by patients may be due to these deactivation reactions with non-target/off-target compounds.
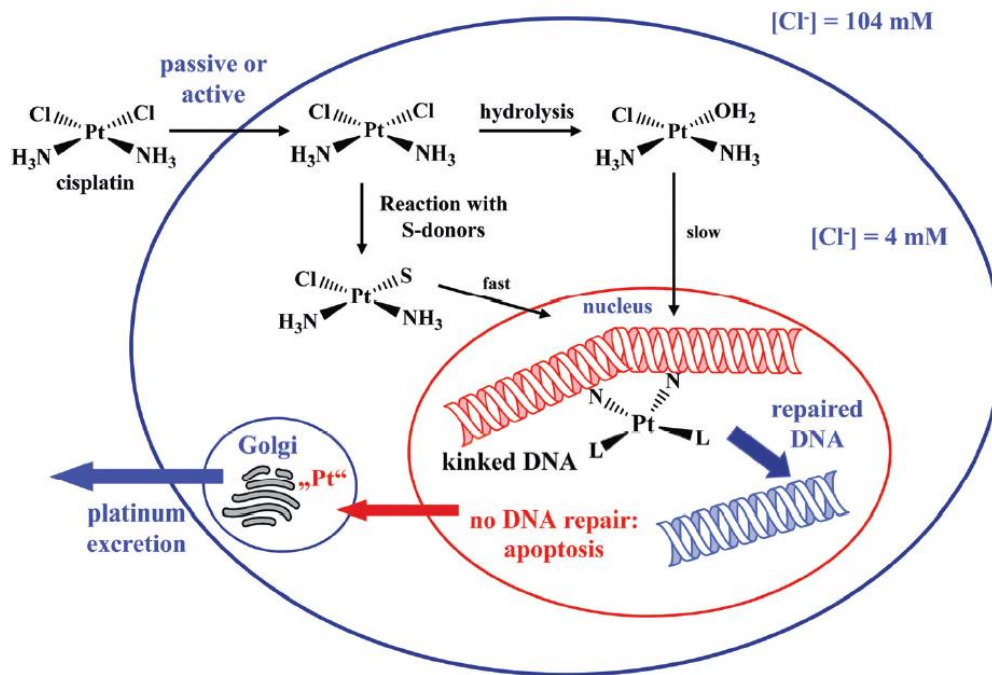
Figure 2.4: Summary of Cisplatin transformations in vivo, and subsequent apoptosis-inducing DNA binding and secretion. Reproduced from Bugarcic *et al.*[29]

Cisplatin has been shown to bind very effectively to available sulphur groups, such as those on Methionine and cysteine,[30,31] both in proteins and in GSH, which is present at milli-molar concentrations in some cells.[2] This great array of deactivating potential ligands detract from the therapeutic effects of Cisplatin and can create damaging reaction products, or at the very least prevent/alter the function of species the Pt(II) drug will bind to. Although Cisplatin has been shown to bind to various proteins including Human Serum Albumin (HSA),[21] Cytochrome c,[32] Calmodulin,[30,33,34] and many others,[23,35–38] some believe that not all of these interactions are purely deactivating and that some may be transient and actually aid in the transportation of metal-based compounds in and out of cells and cellular membranes, perhaps much in the same way that chaperone proteins have been shown to assist in the uptake of essential metal ions.[39]

Regardless platination/metalation of biomolecules has been shown to have both localised and significant effects on protein structure,[40–42] including crosslinking of proteins which usually function independently,[42] and disruption of structure reinforcing disulphide bonds.[43]
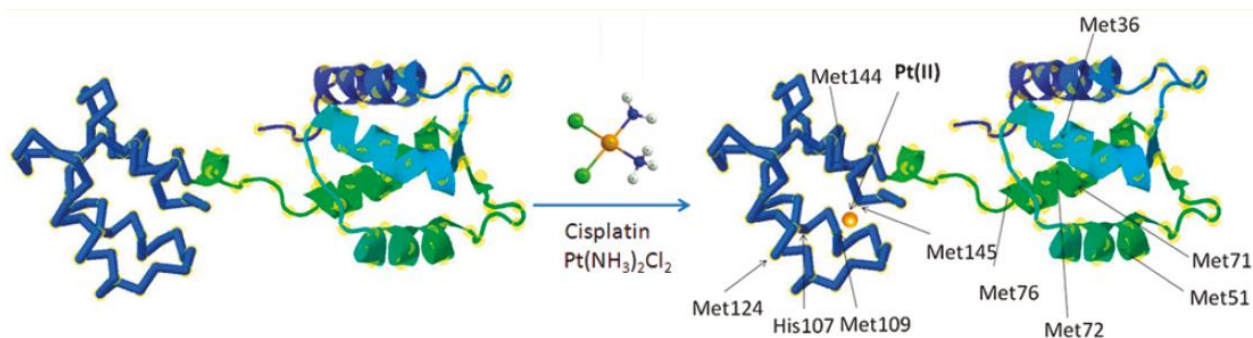
Figure 2.5: Cisplatin-induced protein-protein crosslink between two Calmodulin proteins, causing large structural changes and possible deactivation/limitation of protein function. Reproduced from Li *et al.*[30]

Due to these and many other phenomena, the activity, behaviour, and mechanism of action of metallodrugs presents an interesting and unique area of study. More importantly the study of these interesting and chemically complex drugs varies drastically with the co-ordinating ligands bound to the metal centre. With each metal complex's properties able to be altered differently with varying ligands, metallodrugs can be chemically tuned to increase selectivity, reactivity/potency, activation, and affinity to a wide array of biological targets. For these reasons study of the effects and targets of each drug variant has the potential to greatly inform future drug design, the so-called intelligent drug design method.[10,44] To this end research has moved beyond square planar Pt(II) complexes such as Cisplatin, and onto more complex, chemically tuned compounds, each not only with their own activity and potency, but also with specific tactics and targets to attack cancer cell components in different ways, in order to bypass resistance, cellular self defence mechanisms, cancer cell environmental conditions, or simply to reduce off-target reactions and minimise potential side effects.

2.4 Novel metallodrug approaches:

Below are a series of new, and very varied categories of anti-cancer metallodrugs, each with their own target/method of attacking cancer cells and either causing cell death (cytotoxic compounds), or prevent them from growing/replicating (anti-proliferative compounds).

2.4.1 Piano-stool complexes:

Classical square planar complexes such as Cisplatin and octahedral complexes of other transition metals have been synthesised for years in co-ordination chemistry and studied for their interesting properties and unique metal to ligand bonds. However using large aromatic ligands to bind to metals using more than one pair of electrons has been shown to fix/constrain those binding points, and can enhance activity of others, these so called piano-stool or half-sandwich compounds often employ 5 or 6 membered aromatic hydrocarbon rings (benzene-like or cyclopentadienyl-like rings) to bind to metals in an η^6 fashion, i.e. using 6 electrons. An example of an η^6 arene bound to a metal centre with a bidentate ligand and a single mono dentate ligand is shown below in Figure 2.6:
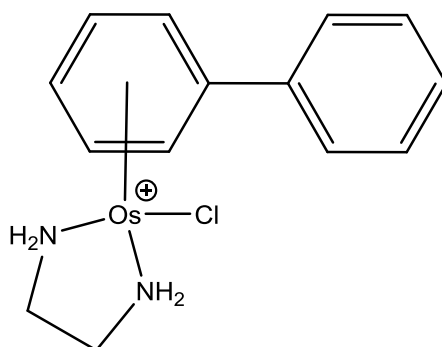
Figure 2.6: [(η^6 –biphenyl)Os(diaminoethyl)Cl]$^+$ piano stool complex proven to be a very active anti-cancer metallodrug.[45,46] The complex shows a classic η^6 –arene interaction exploited by organometallic chemists to tune the reactivity of the complex and target ligand sties.

The use of an η^6-arene and a bidentate ligand has been shown to greatly enhance activity at the mono-dentate site, enabling use as catalysts and active anti-cancer agents.[14,47,48,49]

Study into the use of piano stool complexes as anti-cancer agents show that they can boast impressive potencies, often well above Cisplatin, for instance the compound shown above has been shown to be 49x more potent than Cisplatin towards A2780 ovarian cancer cell lines,[50] the piano stool complex is clearly not only more potent, but may possess a different target/mechanism of action inside cancer cells.[50]

Though piano-stool complexes tend to be more active, and often more selective,[45–47] than Cisplatin towards attacking cancerous cells, they are still often transformed in vivo from the initial complex to a more active species, such as an aqua species,[15,22] or an iodido species to a chlorido species,[51] in a similar fashion to Cisplatin. Some believe/have tried to exploit this transformation as a mechanism of activation, by making the initial injected compound relatively inactive (a pro-drug), until it reaches a cancer cell, which has several distinctly different conditions to normal healthy cells, such as lower pH, change in temperature, and changes in NaCl and GSH concentrations. Chemists can exploit these differences in cell conditions to make intrinsic triggers for pro-drug to become active anti-cancer metallodrugs and increase selectivity of complexes.[10,13,39]

2.4.2 Functionalised piano-stool complexes:

Although many organometallic piano-stool compounds have been shown to boast impressive anti-cancer activity/potency, the selectivity of an anti-cancer drug is one of the key factors in increasing effectiveness and reducing side effects during treatment. In an effort to increase piano-stool compounds selectivity (and perhaps potency in the process) some researchers have attempted to functionalise the ligands within the complex with reactive groups in an effort to target potential binding partners (peptide/protein sequences/DNA/RNA etc.), and to tune the anti-cancer activity of the complex as a whole.[52]

The functionalisation of these complexes has produced a wide array of different metallodrugs, and many metallodrug-biomolecule conjugates. The biomolecule bound to the given drug is believed to function as a directing group, and enhance the accumulation of the drug in certain areas of the

body/cancer cells, and thus not only increase selectivity, but also increase the amount/proportion of the drug in the target area of the cell, making it more effective/potent.[53] Drug-antibody conjugates are currently of great interest due to anti-bodies' inherent selectivity and biological stability and compatibility. These functionalised complexes are also able to be bound to nano-particle frameworks in order to take advantage of their various benefits and architectures.[54] An example of a functionalised metallodrug is below in Figure 2.6, showing the classic piano-stool geometry, but with a functionalised mono-dentate ligand, able to bond to a target (bio)molecule before administration, or after administration if needed:
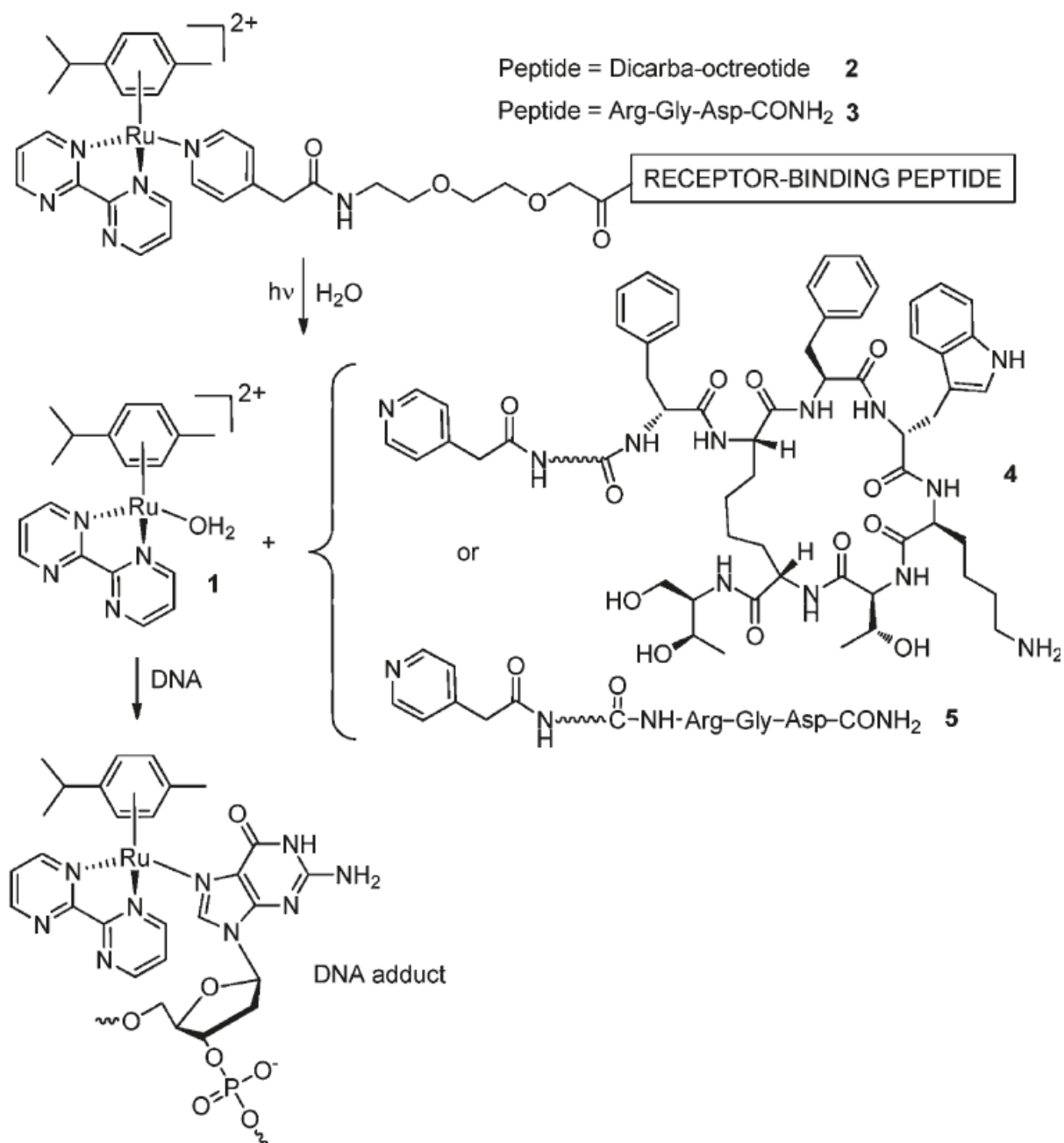
Figure 2.7: a peptide-functionalised ruthenium metallodrug for increased selectivity for binding at a peptide-derrived target within cancer cells. Reproduced from Barragan *et al.*[55]

2.4.3 Photoactivatable complexes:

A much more recent direction of research in anti-cancer drug development is the development of compounds which can be triggered/activated via an external stimulus, such as temperature, pH, or even light.[56,13] The idea behind these drugs is to be able to administer a nearly inactive pro-drug, which will not bind/attack any molecule/cell, and then be able to activate these molecules directly at the cancer/tumour site when and where they are needed. As mentioned before this is sometimes achievable by using the subtle differences in cancer cell environment compared to normal cells/blood (temperature, pH, etc.) to activate drugs passively. However a more powerful method of activation would be to have full control over when and where the compound activates.

To this end a range of photo-sensitive/photo-activatable complexes have been developed in recent years which are inert while in darkness/in the body, but if irradiated with certain wavelengths of light will become activated in some fashion, become active species, and then be able to attack nearby cancerous cells, creating an effective and controlled chemotherapy treatment.

The distinction between the activation/mechanism of these compounds has dictated their classification. Photo-sensitizers are compounds such as tetra-pyrrole, they function as energy transfer molecules; taking photons of light and re-distributing the energy to nearby species. Photo-sensitizers for cancer therapy usually function by absorbing a convenient wavelength of light (that has a good transmission through skin/tissue, but is not harmful e.g. visible light) and then transferring the absorbed energy to nearby cellular components, usually oxygen molecules in order to activate them and create singlet oxygen. Singlet oxygen is a much more reactive form of molecular oxygen which can bind to nearby cellular components, damage them, and generally increase the concentration of Reactive Oxygen Species (ROS) within target cells and induce cell death and/or apoptosis.[2] As a result photosensitisers should be unreactive themselves and usually have no mechanism of action against cancerous or healthy cells, but are simply a means to activate cellular oxygen using more convenient wavelengths of light. The main disadvantage of using

photosensitising molecules for anti-cancer therapy is that they rely on use of cellular oxygen for their action, yet cancerous cells are notoriously hypoxic (i.e. have very low levels of available oxygen), and so are hindered by the lack of active components to transfer the absorbed energy to effectively.[1,57] However there are strategies to circumvent this obviously limiting factor in the use of photosensitisers. The main tactic currently being to co-administer the photosensitising agent with a supply of oxygen to then activate where needed, thus eliminating the short supply of reagent. This co-administration has been achieved via the use of lipid nano-droplets sequestering both the photosensitiser in the lipid monolayer and the oxygen supply within a core of Perfluorocarbon, which can then be released with a wavelength of light and thus be available for photosensitising chemotherapy soon after.[58]
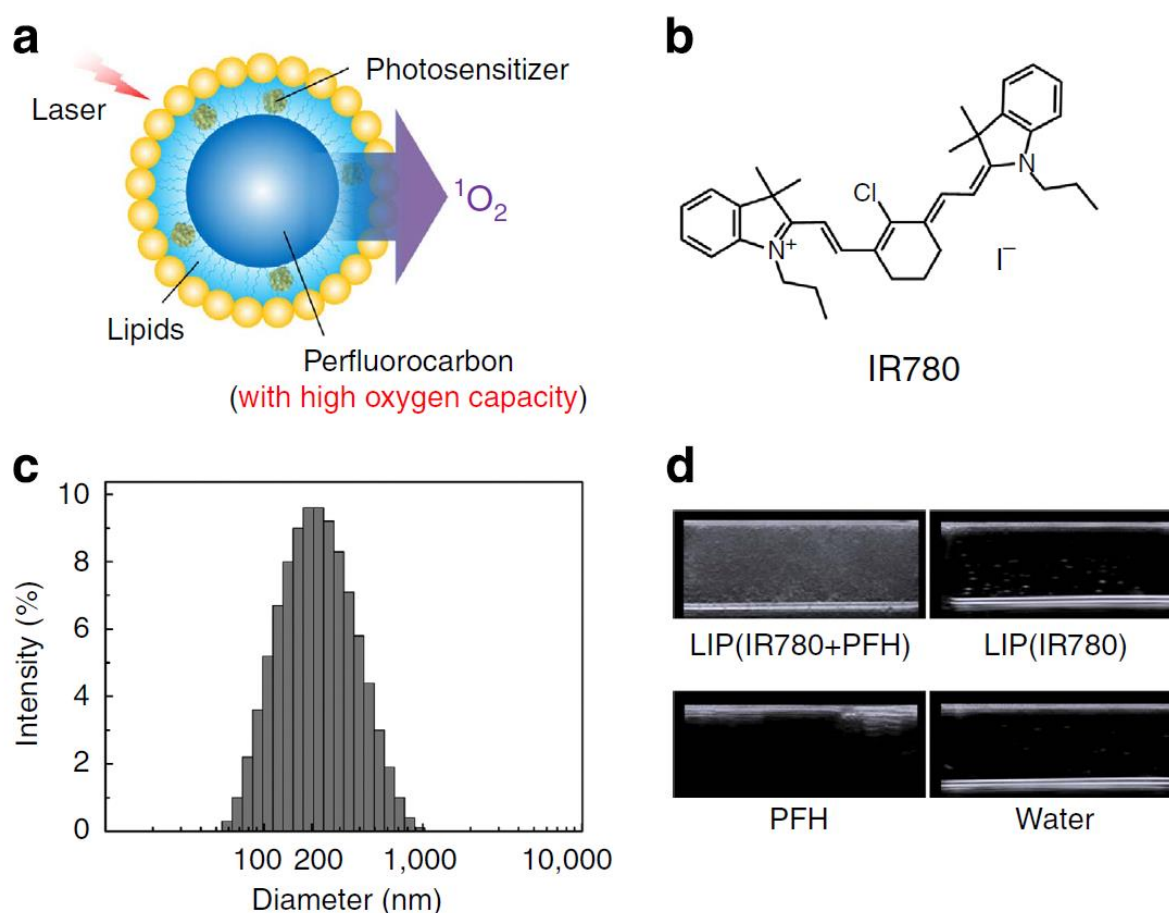
Figure 2.8: representation of a) the Oxy-PDT agent comprised of the lipid nano-droplets containing the photosensitiser and oxygen-rich perfluorocarbon core. b) The active photosensitiser IR780. c)

The size distribution measurements via dynamic light scattering (DLS) of the Oxy-PDT agent. d) Ultrasound images of the Oxy-PDT agent and other solutions in test tubes showing uptake of IR780 and PFH. Reproduced from Cheng *et al.*[58]

Due to the disadvantages of photosensitising agents, directly photo-activatable compounds have been developed and utilised as effective anti-cancer agents.[56] Termed Photo-Activated Chemo Therapy (PACT),[56] these compounds are usually activated by higher-energy wavelengths (such as blue/UV light), resulting in metal-ligand bond cleavage, liberating bound ligands (which may or may not be anti-cancer agents themselves) and creating a metal complex of different geometry and identity, which is usually a much more active and potent anti-cancer agent than the inert pro-drug.[56,59] The activation of these compounds is believed to be due to photon absorption in the Metal to Ligand Charge Transfer (MLCT) band of the pro-drug complex, causing excitation of electrons into anti-bonding orbitals, and weakening/cleavage of metal-ligand bonds, as shown in the molecular orbital diagram below in Figure 2.9:
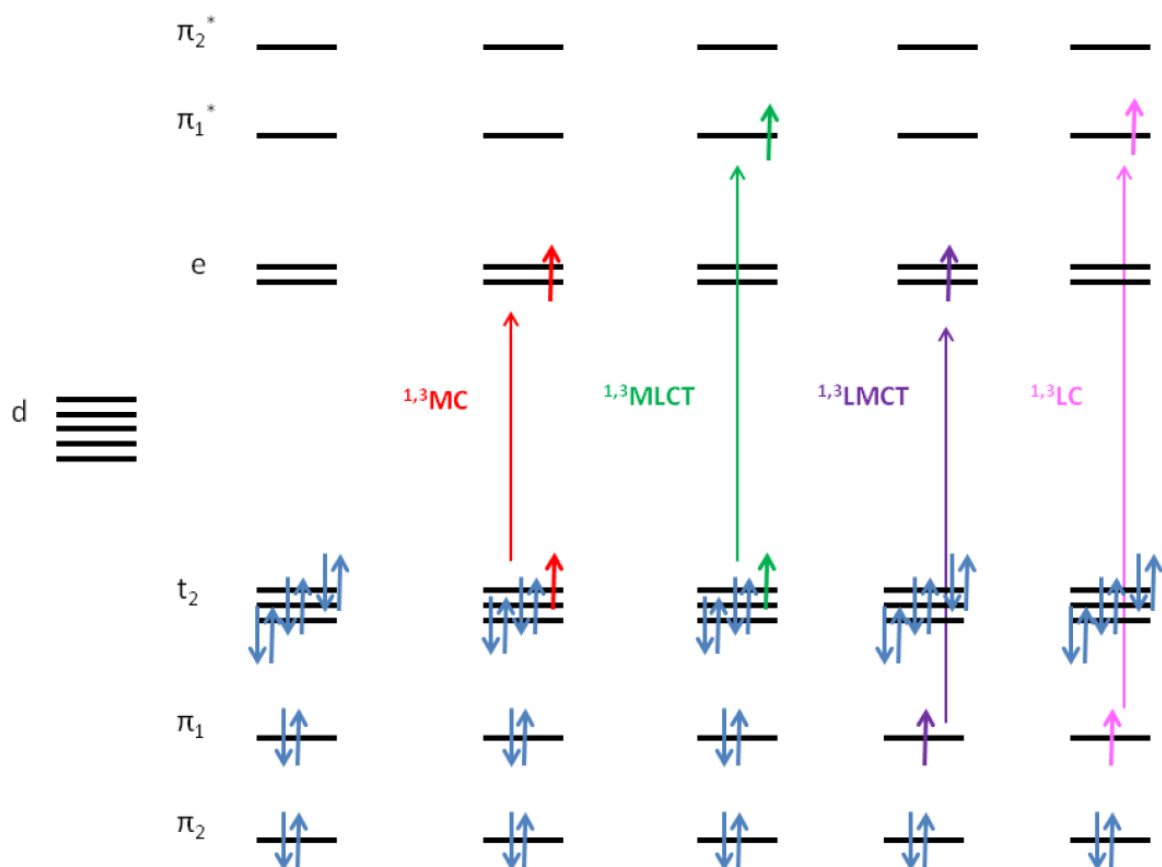
Figure 2.9: MO energy diagram showing the various transitions available to d[6] metal complexes with an octahedral geometry (such as the Pt(IV) compounds shown within). Blue arrows represent electrons with its associated spin in the ground state, coloured arrows represent electrons associated with a given transition. In the singlet state the electrons are spin up (coloured arrow pointing up) whereas in the triplet state they are spin down (coloured arrow pointing down) Reproduced from Shaili *et al.*[60]

Currently there are an array of photoactivatable compounds synthesised and tested for their anti-cancer activity based on a wide variety of metals including platinum,[8,61] ruthenium,[55,56] and copper[56]. Though many options exists, certain complexes have exhibited difficulty becoming truly effective PACT agents, for example copper PACT complexes have been shown to exhibit a distinct increase in activity upon light-activation, a key factor in PACT strategy, but have also exhibited high activity/toxicity when not irradiated (high dark-toxicity), limiting their use as viable PACT agents.[56]

As briefly mentioned, the main target of photoactivated metallodrugs is to activate an inert metal complex to create an active anti-cancer metallodrug, however the ligands released from the complex may also help towards the anti-cancer effect desired. Even small molecules, if released in a high enough number, and in a targeted area may be able to bind to nearby biomolecules/disrupt equilibria/bind to metabolites or active target and generally disrupt cell function. The ability to disrupt target cancer cell function via both the active metallodrug, but also the released ligands is an interesting method of creating a multi-modal attack on cancer cells using varying mechanisms of action, and may be a way of combatting resistant strains of cancer, which may have become resistant to a certain metal-based treatment, but not to the ligand-release mechanism which has been effectively "co-administered" in one drug. PACT agents have also been synthesised with active anti-cancer agents tagged onto ligands of the inert complex,[62] which upon release will begin to exert their own anti-cancer mechanism of action of the target cells. By using effective, proven organic-based anti-cancer agents in this way the PACT agent is able to exert both an effective metal-based therapy (from the activated complex) and an effective organic/small molecule therapeutic (from the released ligands), all in a controlled, targeted fashion.[56,61,63,64] The targeted release of active species is hoped to combat the more "global" side effects of current anti-cancer treatments, such as drugs attacking various parts of the body, far away from the actual tumour site, especially those that are naturally more similar to cancerous cells, such as hair growth cells which divide rapidly to achieve fast hair growth, a common side effect of cancer treatment.

One of the leading PACT agents currently in development is *trans, trans, trans*-$[Pt(N_3)_2(OH)_2(pyridine)_2]$, codenamed FM190.[63] FM190 is based on an octahedral platinum(IV) geometry and is inert in darkness, possessing a full co-ordination sphere and non-labile ligands, it is an ideal basis for an inert pro-drug. Upon irradiation with blue-visible light the Pt(IV) pro-drug is activated, via Pt-azide bond activation,[63] stimulating the creation of Pt(II) based species, and the release of azide radical species.[54,64] Boasting very low activity in the dark and a potent IC_{50} activity of 8.4μM when irradiated,[63] FM190 has been shown to be an effective PACT agent on a variety of

cancer cell lines, including Cisplatin-resistant strains of cancer, which is a very impressive characteristic as the active species of both compounds are believed to be Pt(II) species. FM190 has actually been shown to be 19x more potent than Cisplatin when administered under identical conditions. [63]

FM190 contains several interesting features and developments over previous Pt(IV) based compounds, including the trans-pyridine ligands, azide release, low wavelength light-activation, and the inclusion of hydroxyl ligand release. Pt(IV) based compounds with azide functionality were shown to be photo-sensitive and be viable anti-cancer agents in 2006 by Sadler *et al.*[65] However the previous compounds were based on ligands used in the widely applicable Cisplatin type molecules, such as NH_3 ligands, the inclusion of trans-pyridine ligands was interesting as the cis-geometry for the amine-containing ligands has consistently been shown to be more effective and cytotoxic against cancer cells, the clinically unviable compound trans-platin was the perfect example of this phenomenon.[61] The generation of Pt(II) species with trans amine ligand geometries which are extremely potent anti-cancer agents was a distinct change and proved to be interesting both chemically and pharmacogically.[66,67] FM190 is activated using a relatively long (lower energy) wavelength of light, which is of great advantage to previous PACT agents and platinum(IV) species as it both allows the use of safe, non-damaging wavelengths of light, i.e. visible light (compared to the UV light required previously,[66] but also that longer wavelengths of light show greater transmission/permeability through human skin and tissue than the short wavelengths, enabling external activation of the drug within the body using lights placed on the skin, These transmissions are summarised in Figure 2.10, below:
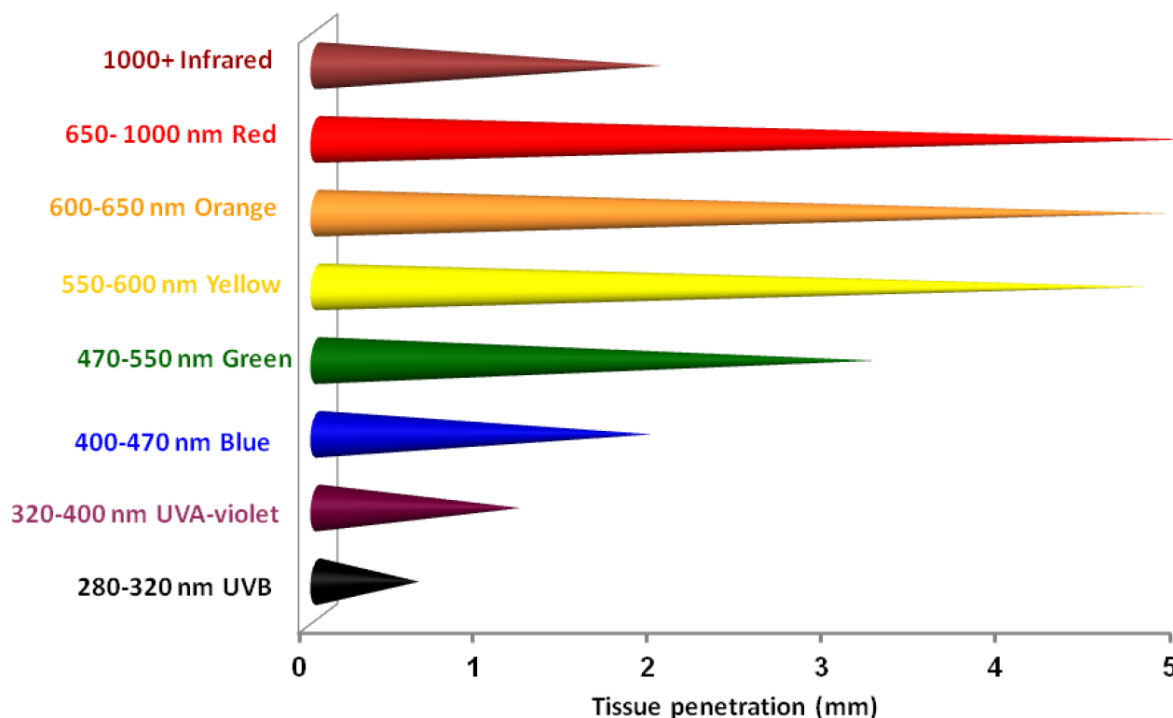
Figure 2.10: Light penetration measurements through human tissue, adapted from Shaili _et al._[68]

LMCT absorption at 285nm
Activated with UVA light
365nm

LMCT absorption at 285nm
Activated with UVA light
256nm

LMCT absorption at 410nm
Activated with blue visible light
463nm
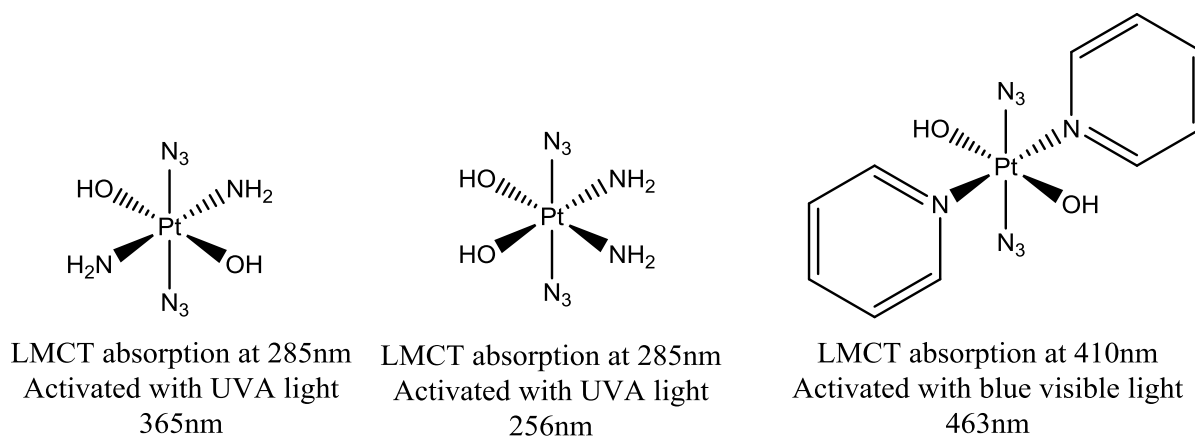
Figure 2.11: Some Pt(IV) photoactivatable metallodrugs based on the di-amine, di-hydroxido, di-azide configuration. Below are the azide-Pt Ligand-to-metal Charge-Transfer (LMCT) band values from UV/visible spectroscopy and theoretical calculations, and the wavelength of light they can be successfully activated by, values taken from MacKay _et al._[59] and Farrer _et al._[63]

FM190 also releases azide radicals during the photoactivation process, extensively studied by Butler _et al._[54,64] by amongst other techniques, Electron Paramagnetic Resonance (EPR), to trap the

produced radicals and display characteristic patterns for azide radicals.[64] Butler *et al.* also showed that the potency of FM190 against cancer cells when irradiated dropped dramatically when additional radical trapping agent (free tryptophan) was added during photoactivation, clearly showing that the radicals released during activation have a distinct contribution to the anti-cancer activity of the complex. It is also worth noting that the hydroxyl radicals released during activation would also have significant biological consequences as HO• molecules are utilised by cells for signalling purposes, and have been shown to cause extensive biomolecule damage and even cell death if present in high enough concentrations.[69–71] Though Butler *et al.* did not show any significant trapping of •OH radicals, they have been shown to be released in this and similar systems.

2.4.4 Catalytic metallodrugs:

The use of active metal complexes against cancerous cells is well practiced (as discussed above), and they have also been used as extremely effective catalysts for decades, especially in the conversion of small molecules (dehydrogenation etc.) and polymerisation reactions (for example the Grubb's metathesis catalysts[72] etc.). Therefore these active (often piano-stool) transition metal complexes have the potential to act as effective "micro-reactors" within the body and act as catalysts to create anti-cancer compounds within the cells themselves.[44,73,74] Catalytic metallodrugs are based on doing just that; remaining inside/outside target cells and using cellular components to create active anti-cancer compounds (usually small molecules) to disrupt the equilibria inside cancerous cells (eventually causing cell death) or create cell damaging molecules (e.g. toxins) to damage cellular components directly, all without being used up/deactivated themselves. Effective catalytic metallodrugs are able to convert target cellular components, present inside the cell, to their products a certain number of times before becoming inactive (the turnover frequency/number), the higher this number, the longer the complex can survive the catalytic cycle (Figure 2.11, below) and continue to exert its cytotoxic effects. Effective catalytic metallodrugs have been synthesised based

on rhodium,[44] iridium,[74] and ruthenium,[44] though many more transition metals should be viable based on other organometallic catalysts available. The catalytic cycle for a novel Rhodium metallodrug synthesised by Soldevila-Barreda *et al.*[44] is shown below in Figure 2.12, the metallodrug catalyses the conversion of NAD^+ to NADH, a key biological transformation used in many processes such as energy metabolism, mitochondrial function, antioxidation/generation of ROS, gene expression, ageing, and cell death.[75] The catalytic metallodrug-induced increase in reductive stress (in contrast to oxidative stress, ROS, earlier) and disruption of these equilibria can eventually induce cell death, but not apoptosis, and so can effectively kill cancerous cells, via different mechanisms of action than many other apoptosis-inducing compounds.[44]
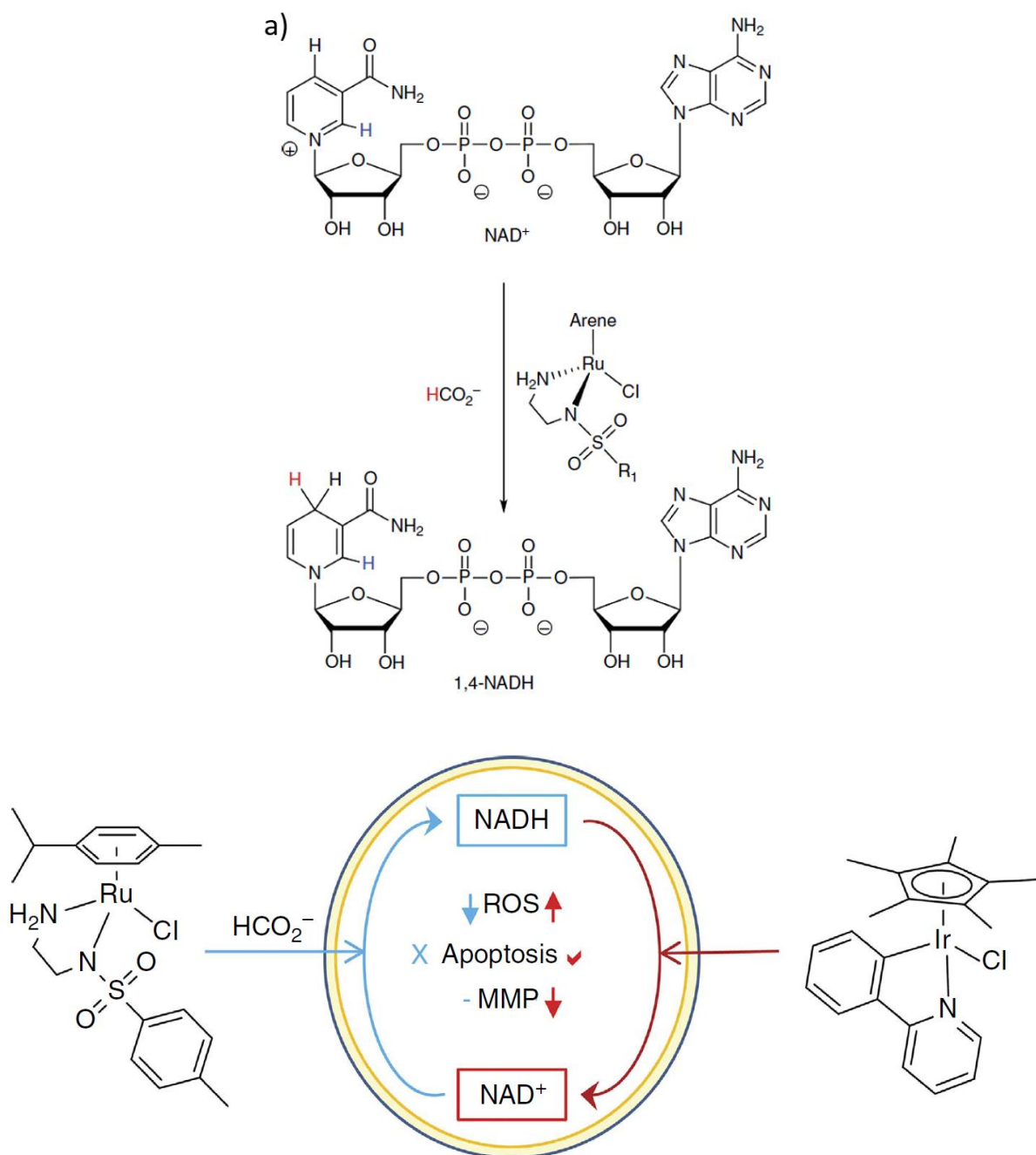
Figure 2.12: Catalysed reduction of NAD$^+$ to NADH by a ruthenium based complex within cells (a) and the comparison to two, one ruthenium, one iridium, complexes and their different mechanisms of action and effects on cancerous cells (b), despite both compounds affecting the same NAD$^+$/NADH cycle they have very different influences and biological consequences. Reproduced from Soldevila-Barreda et al.[44]

2.4.5 Multi-nuclear compounds

Transition metal complex have been observed in a wide variety of geometries and forms, with multi-dentate ligands, counter-ions, tethered complexes, and labile active positions. Also available are complexes containing more than one metal centre, often referred to as multi-nuclear complexes. These compounds have been observed in many configurations, including compounds containing "linker ligands" binding to both metal centres, metal helicates (compounds containing many long multi-dentate ligands to bind many metal centres[76]), and dimer-type complexes containing metal-to-metal bonds, which ironically is usually a common step in the synthesis of many monomeric transition metal complexes. Effective multi-nuclear anti-cancer complexes have been synthesised and tested against cancerous cells/tumors,[77] with a notable example being based on a gold metal centre, and shown in Figure 2.13 below:
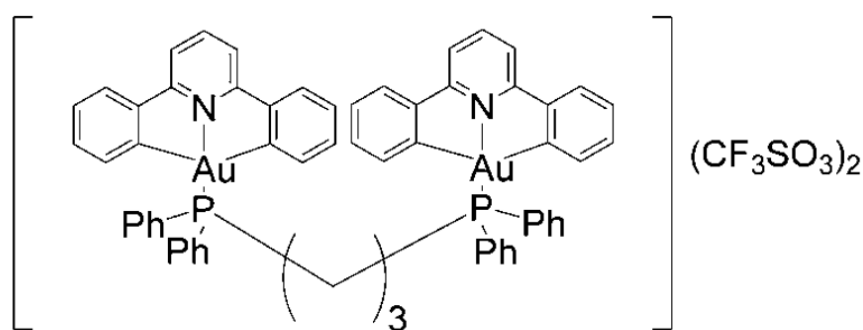
Chemical structure of [(C^N^C)$_2$Au$_2$(μ-dppp)](CF$_3$SO$_3$)$_2$ (**Au3**).

Figure 2.13: Example of a di-nuclear gold complex able to target redox-active sites via binding to neighbouring cysteine residues. Reproduced from Wai-Yin-Sun *et al.*[77]

The compound shown was designed specifically to target redox active proteins, especially selenoenzymes such as Thioredoxin Reductase (TrxR), which contains a disulphide bond between 2 neighbouring cysteine residues, this disulphide bond is an active site and can be turned on/off by reduction/oxidation of the disulphide bridge.[77] The gold-based metallodrug is designed to bind strongly to the two available sulphur/selenium atoms when the disulphide bond is reduced, permanently deactivating the active site and preventing protein function. This deactivation has been

shown to be an effective way of attacking cancerous cells and increasing ROS stress to induce/trigger cell death.[78,79] Since very few proteins contain disulphide bonds between cysteines directly next to each other in an amino acid sequence, the drug has shown relatively high selectivity, especially compared to its mono-nuclear predecessors[79,77] and gold compounds with similar targets.[78]

2.5 References:

(1)   Shannon, A. M.; Bouchier-Hayes, D. J.; Condron, C. M.; Toomey, D. *Cancer Treat. Rev.* **2003**, *29*, 297–307.

(2)   Circu, M. L.; Aw, T. Y. *Free Radic. Biol. Med.* **2010**, *48*, 749–762.

(3)   Mandelker, L. *Oxidative Stress* 1–17.

(4)   Moreno, P. M. D.; Pêgo, A. P. *Front. Chem.* **2014**, *2*, 87.

(5)   Fabris, D. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1–13.

(6)   Kozlowski, J. M.; Fidler, I. J.; Campbell, D.; Xu, Z. L.; Kaighn, M. E.; Hart, I. R. *Cancer Res* **1984**, *44*, 3522–3529.

(7)   De Clercq, E. *Met. Based. Drugs* **1997**, *4*, 173–192.

(8)   Bednarski, P. J.; Mackay, F. S.; Sadler, P. J. *Anticancer. Agents Med. Chem.* **2007**, *7*, 75–93.

(9)   Saonere, J. a; Polytechnic, G. *J. Med. Genet. Genomics* **2011**, *3*, 77–83.

(10)   Reedijk, J. *Eur. J. Inorg. Chem.* **2009**, *2009*, 1303–1312.

(11)   Dyson, P. J.; Sava, G. *Dalton Trans.* **2006**, 1929–1933.

(12)   Pizarro, A. M.; Sadler, P. J. *Biochimie* **2009**, *91*, 1198–1211.

(13)   Sava, G.; Bergamo, A.; Dyson, P. J. *Dalton Trans.* **2011**, *40*, 9069–9075.

(14)   Peacock, A. F. A.; Habtemariam, A.; Ferna, R.; Walland, V.; Fabbiani, F. P. A.; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J. **2006**, *128*, 1739–1748.

(15)   Centerwall, C.; Tacka, K. *Mol. Pharmacol.* **2006**, 348–355.

(16)   Hartinger, C. G.; Metzler-Nolte, N.; Dyson, P. J. *Organometallics* **2012**, *31*, 5677–5685.

(17)   Orvig, C.; Abrams, M. J. *Chem. Rev.* **1999**, *99*, 2201–2204.

(18)   Wong, E.; Giandomenico, C. M. *Chem. Rev.* **1999**, *99*, 2451–2466.

(19)   Hambley, T. W. *J. Chem. Soc. Dalt. Trans.* **2001**, 2711–2718.

(20)   Hearn, J. M.; Romero-Canelon, I.; Qamar, B.; Liu, Z.; Hands-Portman, I.; Sadler, P. J. *ACS Chem. Biol.* **2013**, *8*, 2345.

(21)   Ivanov, a. I. *J. Biol. Chem.* **1998**, *273*, 14721–14730.

(22)   Alderden, R. A.; Hall, M. D.; Hambley, T. W. *J. Chem. Educ.* **2006**, *83*.

(23)   Hartinger, C. G.; Tsybin, Y. O.; Fuchser, J.; Dyson, P. J. *Inorg. Chem.* **2008**, *47*, 17–19.

(24)   Sirichanchuen, B.; Pengsuparp, T.; Chanvorachote, P. *Mol. Cell. Biochem.* **2012**, *364*, 11–18.

(25)   Osinsky, S. P.; Levitin, I. Y.; Sigan, A. L.; Bubnovskaya, L. N.; Ganusevich, I. I.; Campanella, L.; Wardman, P. *Russ. Chem. Bull.* **2003**, *52*, 2636–2645.

(26)   Nh, P. N. O. H.; Ronconi, L.; Sadler, P. J. **2011**, *2*, 262–268.

(27)   Pillaire, M. J.; Hoffmann, J. S.; Defais, M.; Villani, G. *Biochimie* **1995**, *77*, 803–807.

(28)   Zayed, A.; Shoeib, T.; Taylor, S. E.; Jones, G. D. D.; Thomas, A. L.; Wood, J. P.; Reid, H. J.; Sharp, B. L. *Int. J. Mass Spectrom.* **2011**, *307*, 70–78.

(29)   Bugarčić, Ž. D.; Bogojeski, J.; Petrović, B.; Hochreuther, S.; van Eldik, R. *Dalt. Trans.* **2012**, *41*, 12329.

(30)   Li, H.; Zhao, Y.; Phillips, H. I. a; Qi, Y.; Lin, T.-Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 5369–5376.

(31)   Zhao, T.; King, F. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1141–1147.

(32)   Zhang, N.; Du, Y.; Cui, M.; Xing, J.; Liu, Z.; Liu, S. *Anal. Chem.* **2012**, *84*, 6206–6212.

(33)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, 9507–9515.

(34)   Li, H.; Wells, S. a; Jimenez-Roldan, J. E.; Römer, R. a; Zhao, Y.; Sadler, P. J.; O'Connor, P. B. *Protein Sci.* **2012**, *21*, 1269–1279.

(35)   Moreno-Gordaliza, E.; Cañas, B.; Palacios, M. a; Gómez-Gómez, M. M. *Anal. Chem.* **2009**, *81*, 3507–3516.

(36)   Gibson, D.; Costello, C. E. *Eur. J. Mass Spectrom.* **1999**, *510*, 501–510.

(37)   Weidt, S. K.; Mackay, C. L.; Langridge-Smith, P. R. R.; Sadler, P. J. *Chem. Commun. (Camb).* **2007**, *2*, 1719–1721.

(38)   Khalaila, I.; Allardyce, C. S.; Verma, C. S.; Dyson, P. J. *Chembiochem* **2005**, *6*, 1788–1795.

(39)   Johansson, K.; Ito, M.; Schophuizen, C. M. S.; Mathew Thengumtharayil, S.; Heuser, V. D.; Zhang, J.; Shimoji, M.; Vahter, M.; Ang, W. H.; Dyson, P. J.; Shibata, A.; Shuto, S.; Ito, Y.; Abe, H.; Morgenstern, R. *Mol. Pharm.* **2011**, *8*, 1698–1708.

(40)   Qi, Y.; Liu, Z.; Li, H.; Sadler, P. J.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 2028–2032.

(41)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; Connor, P. B. O. 1–25.

(42)   Li, H.; Zhao, Y.; Phillips, H. I. A.; Qi, Y.; Lin, T.; Sadler, P. J.; Connor, P. B. O. **2011**, 5369–5376.

(43)   Moreno-Gordaliza, E.; Cañas, B.; Palacios, M. a; Gómez-Gómez, M. M. *Analyst* **2010**, *135*, 1288–1298.

(44)   Soldevila-Barreda, J. J.; Romero-Canelón, I.; Habtemariam, A.; Sadler, P. J. *Nat. Commun.* **2015**, *6*, 6582.

(45)   Peacock, A. F. A.; Habtemariam, A.; Moggach, S. A.; Prescimone, A.; Parsons, S.; Sadler, P. J.; Uni, V.; Road, W. M.; Eh, E. **2007**, *46*, 2966–2967.

(46)   Peacock, A. F. a; Melchart, M.; Deeth, R. J.; Habtemariam, A.; Parsons, S.; Sadler, P. J.

         Chemistry **2007**, *13*, 2601–2613.

(47)   Peacock, A. F. a; Habtemariam, A.; Fernández, R.; Walland, V.; Fabbiani, F. P. a; Parsons, S.;

         Aird, R. E.; Jodrell, D. I.; Sadler, P. J. *J. Am. Chem. Soc.* **2006**, *128*, 1739–1748.

(48)   Morris, R. E.; Aird, R. E.; Murdoch, P. D. S.; Chen, H.; Cummings, J.; Hughes, N. D.; Parsons, S.;

         Parkin, a; Boyd, G.; Jodrell, D. I.; Sadler, P. J. *J. Med. Chem.* **2001**, *44*, 3616–3621.

(49)   Ang, W. H.; Casini, A.; Sava, G.; Dyson, P. J. *J. Organomet. Chem.* **2011**, *696*, 989–998.

(50)   Hearn, J. M.; Romero-Canelón, I.; Munro, A. F.; Fu, Y.; Pizarro, A. M.; Garnett, M. J.;

         McDermott, U.; Carragher, N. O.; Sadler, P. J. *Proc. Natl. Acad. Sci.* **2015**, *112*, E3800–E3805.

(51)   van Rijt, S. H.; Romero-Canelón, I.; Fu, Y.; Shnyder, S. D.; Sadler, P. J. *Metallomics* **2014**, *6*,

         1014.

(52)   Millett, A. J.; Habtemariam, A.; Romero-Canelón, I.; Clarkson, G. J.; Sadler, P. J.

         Organometallics **2015**, *34*, 2683–2694.

(53)   Paula, L.; Salassa, L.; Betanzos-lara, S.; Habtemariam, A.; Moreno, V.; Sadler, P. J.; March, V.

         2011, 14098–14108.

(54)   Butler, J. S.; Sadler, P. J. *Curr. Opin. Chem. Biol.* **2013**, *17*, 175–188.

(55)   Barragán, F.; López-Senín, P.; Salassa, L.; Betanzos-Lara, S.; Habtemariam, A.; Moreno, V.;

         Sadler, P. J.; Marchán, V. *J. Am. Chem. Soc.* **2011**, *133*, 14098–14108.

(56)   Lovejoy, K. S.; Lippard, S. J. *Dalton Trans.* **2009**, 10651–10659.

(57)   Song, X.; Liu, X.; Chi, W.; Liu, Y.; Wei, L.; Wang, X.; Yu, J. *Cancer Chemother. Pharmacol.* **2006**,

         58, 776–784.

(58)   Cheng, Y.; Cheng, H.; Jiang, C.; Qiu, X.; Wang, K.; Huan, W.; Yuan, A.; Wu, J.; Hu, Y. *Nat.*

Commun. **2015**, *6*, 8785.

(59)   Mackay, F. S.; Woods, J. a; Moseley, H.; Ferguson, J.; Dawson, A.; Parsons, S.; Sadler, P. J.

Chemistry **2006**, *12*, 3155–3161.

(60)   Shaili, E. Photoactivatable Platinum (IV) Anticancer Complexes, University of Warwick, 2013.

(61)   Mackay, F. S.; Woods, J. a; Heringová, P.; Kaspárková, J.; Pizarro, A. M.; Moggach, S. a;

Parsons, S.; Brabec, V.; Sadler, P. J. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20743–20748.

(62)   Smith, N. a; Sadler, P. J. *Phil. Trans. R. Soc. A* **2013**, *371*, 20120519.

(63)   Farrer, N. J.; Woods, J. a; Salassa, L.; Zhao, Y.; Robinson, K. S.; Clarkson, G.; Mackay, F. S.;

Sadler, P. J. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 8905–8908.

(64)   Butler, J. S.; Woods, J. a; Farrer, N. J.; Newton, M. E.; Sadler, P. J. *J. Am. Chem. Soc.* **2012**, *134*,

16508–16511.

(65)   Bednarski, P. J.; Grünert, R.; Zielzki, M.; Wellner, A.; Mackay, F. S.; Sadler, P. J. *Chem. Biol.*

2006, *13*, 61–67.

(66)   Ronconi, L.; Sadler, P. J. *Dalt. Trans.* **2011**, *40*, 262–268.

(67)   Farrer, N. J.; Woods, J. a; Munk, V. P.; Mackay, F. S.; Sadler, P. J. *Chem. Res. Toxicol.* **2010**, *23*,

413–421.

(68)   Shaili, E. *Sci. Prog.* **2014**, *97*, 20–40.

(69)   Thornalley, P. J.; Rabbani, N. *Iron Metab.* **2011**, 205–228.

(70)   Monnier, V. M.; Nemet, I.; Sell, D. R.; Weiss, M. F. 25–50.

(71)   Nowicka, A. M.; Kowalczyk, A.; Sek, S.; Stojek, Z. **2013**.

(72)   Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. *Org. Lett.* **1999**, *1*, 953–956.

(73)   Joseph, J.; Nagashri, K.; Janaki, G. B. *Eur. J. Med. Chem.* **2012**, *49*, 151–163.

(74)   Liu, Z.; Sadler, P. J. *Acc. Chem. Res.* **2014**, *47*, 1174–1185.

(75)   Ying, W. *Antioxid. Redox Signal.* **2007**, *10*, 179–206.

(76)   Kaczorowska, M. a; Hotze, A. C. G.; Hannon, M. J.; Cooper, H. J. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 300–309.

(77)   Sun, R. W.-Y.; Lok, C.-N.; Fong, T. T.-H.; Li, C. K.-L.; Yang, Z. F.; Zou, T.; Siu, A. F.-M.; Che, C.-M. *Chem. Sci.* **2013**, *4*, 1979.

(78)   Zou, T.; Ching, A.; Lum, T.; Lok, C.-N.; Zhang, J.-J.; Che, C.-M. *Chem. Soc. Rev* **2015**, *44*, 8786–8801.

(79)   He, L.; Chen, T.; You, Y.; Hu, H.; Zheng, W.; Kwong, W. L.; Zou, T.; Che, C. M. *Angew. Chemie - Int. Ed.* **2014**, *53*, 12532–12536.

Chapter 3:

Multi-targeted Photoactivatable Platinum Anticancer Complexes

The effective targets of most platinum based anti-cancer agents are the damaged strands of DNA within cancerous cells, however it has been shown that the vast majority of these active species never react with cancerous DNA but instead are believed to react with biomolecules along the way. Some believe the reaction of metallodrugs with non-target biomolecules to be purely deactivating, however others believe that biomolecule reactions can not only provide crucial benefits in the effective transportation of drugs through membranes and into target areas, but also that reactions with biomolecules can trigger key drug transformations in solution to create more active species and as a result better chemotherapy treatments.

The chapter ahead investigates a novel next-generation anti-cancer compound (FM190) which can be activated by certain wavelengths of light to produce extremely potent, but not fully understood active species which effectively bind to biomolecules. The activation and subsequent reactions of FM190 towards several model biomolecules of interest are studied and explained. Reactions were monitored mainly using UHR-FT-ICR Mass Spectrometry, but also UV/visible spectroscopy, and Electron Paramagnetic Resonance spectroscopy (EPR) to monitor radical formation. The novel design and activation of Pt(IV) compounds makes them a very attractive area of study for chemotherapy, but also for analysis. Almost immediately it was apparent that the Pt(IV) compounds behaved very differently to previous Pt(II) metal-based drugs and posed unique analytical challenges, especially with relation to MS and subsequent MS/MS analysis. These challenges, the fundamental causes behind them, and the effect on MS data are discussed in detail,

including investigations into the unique mechanisms of dissociation for biomolecules involving platinum-based modifications.

Although the results presented within this chapter are predominantly based on FT-ICR MS and MS/MS results from the thesis Author, this was a collaborative project and some elements were conducted by collaborating researchers, namely the synthesis of the FM190 compound was conducted by Evyenia Shaili, a proportion of the UV/visible spectroscopy was conducted by Evyenia Shaili and Carlos Sanchez-Cano, and the EPR spectroscopy was carried out and analysed by Carlos Sanchez-Cano.

A manuscript entitled "Multi-targeted Photoactivatable Platinum Anticancer Complexes" by Christopher A. Wootton, Andrea F. Lopez-Clavijo, Carlos Sanchez-Cano, Evyenia Shaili, Mark P. Barrow, Peter J. Sadler, and Peter B. O'Connor, has been prepared for submission to the Journal of the American Chemical Society (JACS) based directly on the contents and findings of this chapter.

Abstract

Octahedral platinum(IV) complexes such as *trans,trans,trans*-[Pt(N$_3$)$_2$(OH)$_2$(pyridine)$_2$] are stable in the dark, but potently cytotoxic to a range of cancer cells when activated by UVA or visible light, and active in vivo. Photoactivation can lead to reduction and unusual Pt(II) lesions on DNA, but radical chemistry is also initiated in the excited state. Here we show that peptides and proteins are also potential targets. High resolution FT-ICRMS studies show that reactions of *trans,trans,trans*-[Pt(N$_3$)$_2$(OH)$_2$(pyridine)$_2$] activated by visible light with two neuropeptides Substance P, RPKPQQFFGLM-NH$_2$ (Sub P) and [Lys][3]-Bombesin, Pyr-QKLGNQWAVGHLM-NH$_2$ (K[3]-Bom) give rise to unexpected products, both oxidized and platinated peptides. The non-ergodic character of ECD proved crucial in MSMS analysis, enabling retention of the Pt complex during fragmentation. Sub P undergoes one-O oxidation of Met[11], whereas K[3]-Bom not only undergoes one-O oxidation of its C-terminal Met to the sulfoxide and two-O oxidation of Trp[7] to N-formylkynurenine, the latter assignable to hydroxyl radical pathways. In EPR experiments on photoactivated 1 using the spin-trap DEPMPO, azidyl radicals were detected in the presence of Sub P but not K[3]-Bom which induced oxidation of the spin trap, characteristic of hydroxyl radical formation. Platination by {Pt(py)$_2$(N$_3$)}[+] occurs preferentially at His[11] of K[3]-Bom. Platination by Pt[II] was observed at Lys[3] and Arg[1] of Sub P. This work reveals the multi-targeting nature of excited state platinum anticancer complexes. Not only can they target DNA but also peptides (and proteins) by radical coordination mechanisms, in a manner that that depends strongly on the amino acid sequence of the peptide.

Introduction

Transition metal complexes have been explored and utilised as pharmaceuticals for decades, showing great promise as anti-cancer,[1] anti-malarial,[2] anti-tuberculosis,[3] and anti-viral[4] agents. However it is the platinum based chemotherapeutics which have been of particular interest due to their widespread success in combatting many different forms of cancer.[5] Many of these metallodrug complexes exert their therapeutic effect by binding to the DNA in cancerous cells, causing a conformational change in the double helix and eventually leading to apoptosis;[6] this is usually achieved via the reactive metal centre binding to preferred potential ligands on certain residues within the DNA.[7]

Similarly peptides and proteins possess a range of potential ligands for these reactive complexes and can often form strong bonds with metallodrugs in the body,[8] such that 90-98% of some of the most widely used anticancer metallodrugs in the blood plasma become protein bound within just 1 day after intravenous injection,[9] although some believe protein binding may be a critical transport step in the overall mechanism of action of the metallodrug.[10,11] Differences in mechanism of action and ambiguity regarding how these drugs effectively combat cancerous cells has encouraged research into investigating exactly how and where these metallodrugs are binding to certain biomolecules, such as proteins.[8 12] Mass spectrometry has established itself as a key technique for investigating these interactions and has been utilised to effectively study many metallodrugs including platinum based drugs Cisplatin,[13,14,15] Transplatin,[16] Oxaliplatin,[17] many ruthenium based metallodrugs,[10,18] and recently iridium and osmium based metallodrugs,[19,20] all of which are injected into the body and rely on selectivity alone to direct these active drugs to cancerous cells.

Moving on from more traditional transition metal compounds, new photoactivatable metal based chemotherapy agents are currently being investigated to improve on the promising

clinical and medical record of platinum complexes. Photoactivation of metal complexes allows temporal and spatial control over the activation of these very active anti-cancer metallodrugs, theoretically reducing the more global side effects on the body from the platination of susceptible but non-target cells often seen during treatment with more traditional platinum drugs, such as Cisplatin.[21] These novel anti-cancer compounds are injected into the body as an inert prodrug, usually as a six co-ordinate octahedral low-spin d^6 complex with non-labile ligands. The Pt(IV) prodrug can then be activated at the tumour site via exposure to certain wavelengths of light;[21,22] this exposure triggers photo-activated ligand release creating a highly active Pt(II) d^4 species which will then react with nearby biological components in a similar fashion to other metallodrugs, by binding selectively to preferred ligands found on many biomolecules, such as DNA[23] and possibly proteins.

An important requirement for any effective drug to be used in Photo Activated Chemotherapy (PACT)[24] is to exhibit little to no cytotoxicity in the dark, yet possess significant cytotoxicity to cancer cells upon irradiation.[23,24] Dark toxicity has been a limitation for some cobalt and copper photoactivatable complexes as they have exhibited a high dark toxicity.[25] In addition photoactivatable Pt(IV) complexes do not require oxygen to carry out their cytotoxic effect allowing them to function in the hypoxic conditions often seen in many cancerous tumours,[26,27] a distinct advantage over other more conventional photosensitisers used in photodynamic therapy (e.g. those based on tetrapyrrole)[28] and also over front line reagents such as Carboplatin and Cisplatin whose activity can be greatly reduced in hypoxic Ovarian cancer cell treatments.[29]

Herein a study is presented using the photoactivatable platinum complex *trans, trans, trans*-[Pt(N₃)₂(OH)₂(pyridine)₂] (Complex **1**), a novel anti-cancer prodrug (precursor to an active drug species) which, upon irradiation with certain wavelengths of visible light, will release ligands creating an active metallodrug which can bind to nearby targets, such as the DNA in cancerous cells, leading to apoptosis[6] (Scheme 3.1). It is important when investigating such photo-activated processes that both the fate of the metal complex and that of the released ligands is investigated as it has been shown that release of these reactive radical species plays a significant role not only in the potent cytotoxicity observed when using these complexes[30] but also in their activity against Cisplatin resistant strains of cancer.[21] By studying the interactions of these new drugs with model biomolecules, it is possible to investigate these unique mechanisms of action and help explain the increased cytotoxicity, effectiveness against Cisplatin resistant cancer strains and other phenomena observed during the cytotoxic investigations of the metallodrug.
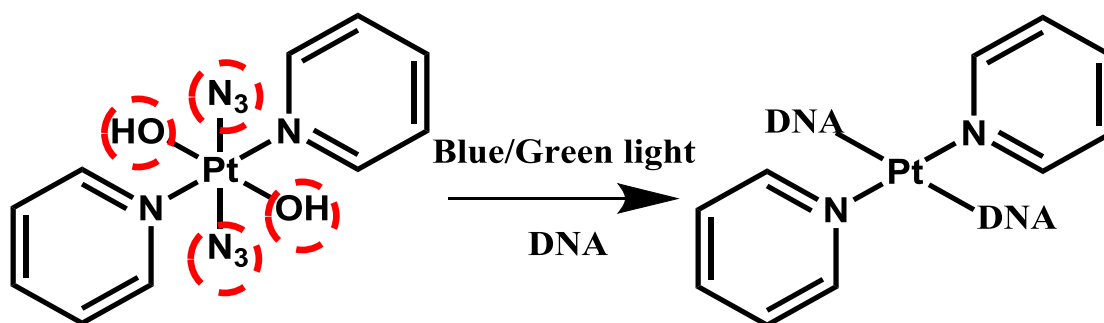
Scheme 3.1 : (Left) *Trans, trans, trans*-[Pt(N₃)₂(OH)₂(py)₂] (Complex 1), the photo-activatable prodrug used in this study, along with (Right) an illustrative photo-activation pathway leading to DNA binding (an inter-strand cross-link).[23]

Top-down tandem Mass Spectrometry (MS/MS) analysis has been successful in studying an array of peptides and proteins while retaining vital post-translational modifications (PTM's).[31,32,33] Electron capture dissociation (ECD) has been shown to be particularly effective for PTM analysis due to its non-ergodic character[34] and its ability to fragment

species without loss of fragile PTM's.[35] Electron-based dissociations have proven to be particularly useful when studying metallodrug interactions where the top down approach has allowed the characterisation of entire protein sequences and the unambiguous determination of multiple binding sites for metallodrugs to various proteins[15,36,37] without further chemical modifications such as digestion, which can disrupt/dissociate modifications, particularly organometallic complexes.

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) not only offers ultra-high resolving power and unparalleled mass accuracy,[38,39,40,41] but also a large array of MS/MS fragmentation techniques including Collisionally Activated Dissociation (CAD), Electron Capture Dissociation, Infra-Red Multi Photon Dissociation (IRMPD), Electron Transfer Dissociation (ETD), and Electron Detachment Dissociation (EDD) making it ideally suited to investigating biomolecules via MS/MS.[42,43,20] Though previous transition metal based metallodrugs have often shown a preference for certain (often sulphur containing) residues in biomolecules;[14,19] the binding sites of new metallodrugs cannot be easily predicted based purely on these arguments as the nature of the remaining ligands greatly influences the rates and targets of drug binding.

Other analytical techniques have been employed to study metallodrug-biomolecule interactions, such as X-ray crystallography[44] and NMR spectroscopy,[45] however these techniques are often limited by large sample requirements and/or suffer greatly when studying mixtures of compounds. FT-ICR MS offers a much more sensitive alternative which can provide accurate sequence information over a large mass range.[37,46,47] By studying and understanding these interactions to a high accuracy on model biomolecules we will provide unambiguous identity and binding site information which could be implemented in future larger scale investigations on complex biological mixtures where the top down approach may not yet be available/viable.

Materials and methods

Substance P (RPKPQQFFGLM-NH$_2$), [Lys]^3-bombesin (Pyr-QKLGNQWAVGHLM-NH$_2$), and formic acid were purchased from Sigma Aldrich (Sigma Aldrich Company Ltd., Dorset, UK). Low concentration Agilent tuning mix was purchased from Agilent Technologies (Santa Clara, CA) and used as received. Complex **1** was synthesised and characterised as described elsewhere.[21] EPR tubes were purchased from Wilmad Labglass. The spin trap 5-(Diethoxyphosphoryl)-5-methyl-1-pyrroline-N-oxide (DEPMPO) was obtained from Enzo Life Sciences in high purity and used without any further purification. Ultra-pure water was obtained from a Milli-Q UV III system (Milli-Q, Hertfordshire, UK).

Reactions of peptides with complex 1

Aliquots of aqueous solutions of Substance P (1 mM) and [Lys]^3-Bombesin (1 mM) were prepared and mixed with an aqueous solution of complex **1** (250 μM) to give solutions of 0.5:1, 1:1 and 2:1 drug:peptide mol ratio. The samples were then irradiated under 463 nm (blue visible) light at 298 K for various times, before being diluted to MS concentrations (~1 μM) and frozen at -80 ˚C ready for MS analysis. Freshly prepared samples were compared to those frozen for one to several weeks and showed no observable variation in the mass spectra obtained.

FT-ICR Mass spectrometry

Nano-electrospray (nESI) Mass spectrometry was performed on a Bruker SolariX Fourier Transform Ion Cyclotron Resonance Mass spectrometer (FT-ICR MS) fitted with a 12 Tesla actively shielded magnet (Bruker Daltonics, Bremen, Germany). Aqueous peptide samples (5 μM) were spiked with 0.3% formic acid to aid ionisation during nESI. Solutions containing Complex **1** (including reaction mixtures) were analysed via nESI in Milli-Q water with no added acid.

For ECD MS/MS analysis; the species of interest were isolated in the first quadrupole, externally accumulated in the collision cell for 0.1-7 seconds and then transferred to the Infinity Cell© for Electron Capture Dissociation (ECD) fragmentation and detection. Ions in the Infinity Cell© were irradiated with 1.3 eV -1.6 eV electrons from a 1.5 A hollow cathode dispenser for 50-600 ms prior to detection.

For CAD MS/MS analysis, species of interest were isolated in the first quadrupole and accelerated into argon collision gas contained within the hexapole collision cell for Collisionally Activated Dissociation (CAD) at 10-30 V and continuously accumulated for 0.1- 4 seconds before transfer to the Infinity Cell© for detection.

MS/MS spectra were internally calibrated using the minimal number of unmodified (peptide spectra) or modified (Pt adduct spectra) c/z ions and the charge reduced species $[M+nH]^{n-1+\cdot}$ where possible (species used for calibration are marked). A dual-spray nESI experiment was also conducted using the $[K^3\text{-Bom}+2H]^{2+}$ ion and ions from Agilent Tune mix, utilising an in-cell isolation (Multi-CHEF)[48] and identical ECD parameters to validate the internal calibration of the MS/MS spectra. Similar standard deviations were found for the ECD fragmentation spectra when calibrated with either fragment ions or with tune-mix peaks (see Supporting Information Figure S3.1 and Table S3.1). This dual-spray Multi CHEF ECD approach was also used for some Complex 1 + peptide reaction product ECD spectra to improve the internal calibration (spectra and calibration peaks marked accordingly).

Electron Paramagnetic Resonance (EPR) measurements

EPR spectra were recorded at ambient temperature on a Bruker EMX (X-band) spectrometer fitted with a cylindrical Tm110 mode cavity (Bruker 4103TM). Samples were contained in a quartz capillary (I.D. 1.0 mm; O.D. 1.2 mm; Wilmad Labglass) sealed with T-Blu Tac®, placed inside larger quartz tubes (O.D. 2.0 mm) to achieve easy and accurate positioning of

the sample inside the resonator. Typical key EPR spectrometer settings were modulation amplitude 2.0 G, microwave power 0.63 mW, 1.0×10^5 receiver gain, conversion time 81.92 ms, time constant 81.92 ms, sweep width 200 G, and a repeated number of 10 X-scans with a resolution in Y of 5 or 9. Spin-trapping experiments were performed on aqueous solutions of the complex with excess of spin trap (1 mM complex, 6 mM DEPMPO) in the presence or absence of peptides (1 mM). A visible blue light emitting diode (LED) (l= 463 nm, 64 mW cm^{-2}) was used as the source of irradiation and placed at a distance of 8.5 cm from the tube in the EPR cavity. Irradiations lasted up to 2 h, each slice corresponding to 14 min of irradiation (10 scans). The TM110 EPR cavity used is equipped with a grid on one side allowing optical access (approximately 50%) transmission. The refractive index of quartz is approximately 1.55; hence at normal incidence approximately 5% of the incident light is reflected at an air quartz interface. EPR spectra were analysed and simulated using the EASYSPIN software.[49]

UV/vis Spectroscopy

Aqueous solutions of Complex **1** (60 uM), SubP(120 μM), K3-Bom (120 μM), Complex **1** (60 μM)+SubP (120 μM), and Complex 1 (60 μM)+K^3-Bom (120 μM) were analysed via UV/vis spectroscopy using a Cary 300 scan UV-visible spectrophotometer (Agilent, California, US). Solutions were analysed in Quartz cuvettes (0.5 mL volume, 1 cm path length), in the region of 200-800 nm at a rate of 10 nm/s (600 nm/minute), average time = 0.1 s, data interval 1 nm.

Results and Discussion

Aqueous solutions of Peptide (either Substance P or [Lys]^3-Bombesin were mixed with an aqueous solution of Complex **1** (in varying ratios of 1:1, 1:2, and 2:1 drug:peptide respectively). These solutions were then irradiated for 30-120 minutes with Blue-visible light (463 nm) to activate Complex **1** and initiate the photoreaction. Samples were immediately diluted after the set irradiation period with water and analysed via nESI-FT-ICR Mass Spectrometry. All sample preparation, incubations, and mass spectrometry analysis were conducted in darkness or with very low levels of red light to avoid activation of the complex, control analyses using bright red light showed no activation of the complex over several hours. Experiments conducted even in very low levels of white light were shown to alter the compositions of samples, as shown by continuous nESI-FT-ICR MS scanning over several hours, with continued exposure causing further photoactivation of Complex **1**.

nESI FT-ICR mass spectra of ~1 μM aqueous Complex **1**+SubP and Complex **1**+ K^3-Bom reaction mixtures after 1 and 2 hours (respectively) of irradiation with blue visible light are shown in Figures 3.1a and 3.1b, respectively. Bombesin is a known tumour marker peptide for small cell lung carcinoma, neuroblastoma, and gastric cancer. Bombesin has also been shown to increase the proliferation of some pancreatic cancer cell lines.[50] Substance P is a neuromodulator peptide involved in pain responses and inflammatory disease (amongst other processes), elevated levels of Substance P have also been found in several types of cancer cells.[51]
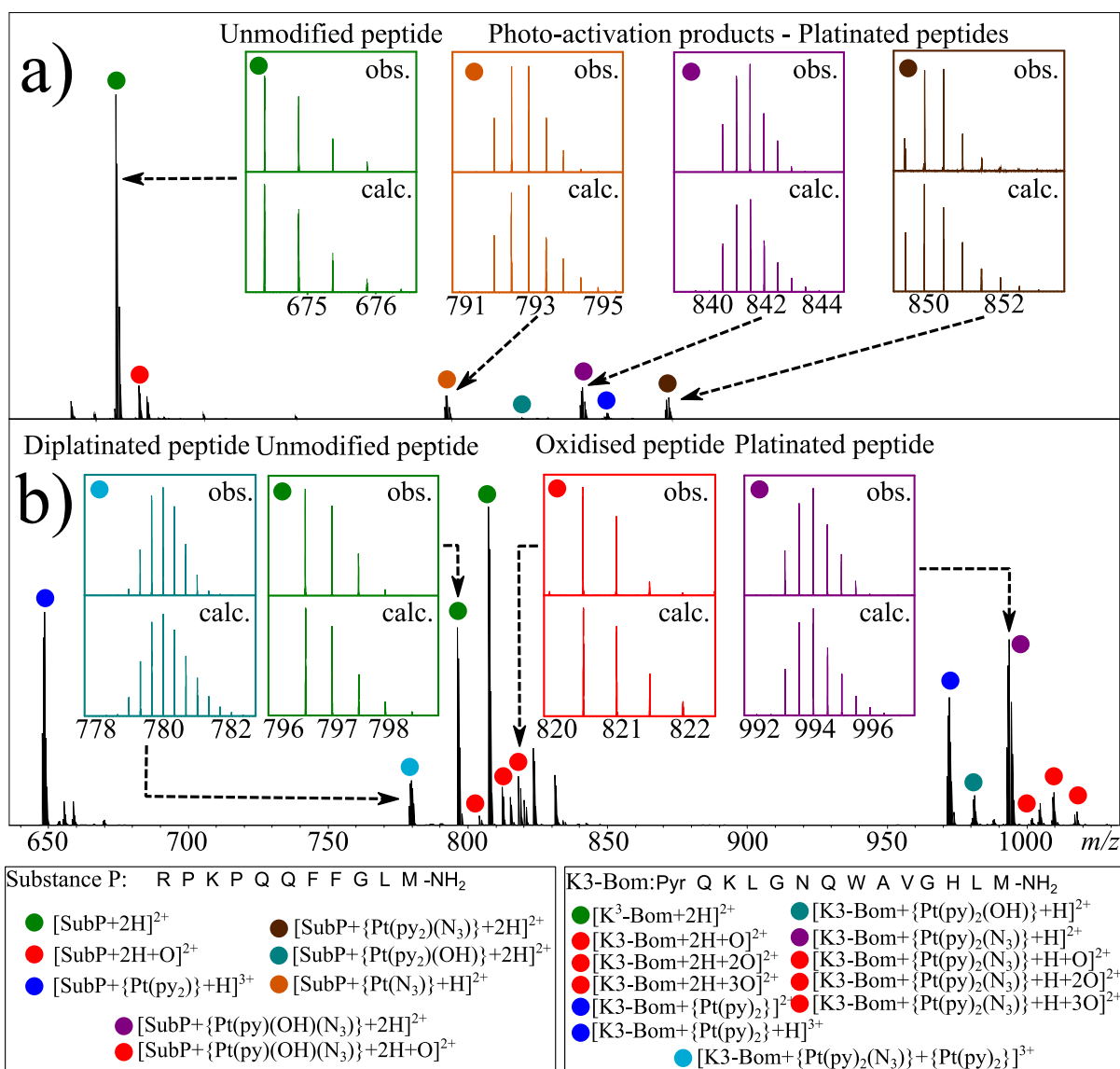
Figure 3.1: nESI FT-ICR mass spectra of ~1 μM aqueous Complex 1+SubP (a) and Complex 1+ K³-Bom (b) reaction mixtures (0.5:1 drug:peptide ratio) after 1 and 2 hours of irradiation (respectively) with blue visible light (463nm). Inset, various isotopic patterns for the observed (obs.) and calculated (calc.) species assigned from the mass spectra – showing the platinum influence on observed isotopic distributions. Green filled circles indicate unmodified peptide species, Red indicate oxidised species (both peptide and platinated peptides), other colours indicate platinated peptide species with different Pt(II) based modifications.

An unexpectedly large number of products were observed for both reactions. Many spectral peaks can be assigned to adducts of platinum complexes to the peptide along with various bound ligands, as a result of the photoactivation process. For the Complex **1**+Substance P mixture, only mono-platinated peptide adducts were observed containing the modifications

{Pt(py)(OH)(N$_3$)}, {Pt(py)$_2$(N$_3$)}, {Pt(N$_3$),},and {Pt(py)$_2$}with reasonable intensity, along with a {Pt(py)$_2$(OH)}-modified species at a very low intensity.

The spectrum from the 1:2 reaction of Complex 1+ K^3-Bom showed an even larger number of reaction products, including both mono- and di-platinated adducts, were observed including K^3-Bom adducted to {Pt(py)$_2$}, {Pt(py)$_2$(N$_3$)}, {Pt(py)$_2$(OH)}, and both {Pt(py)$_2$} and {Pt(py)$_2$(N$_3$)} simultaneously.

Interestingly, for both peptides, oxidation with the addition of one oxygen (by exact mass) to the unmodified and platinated Substance P species and addition of one, two, or three oxygen atoms to the unmodified and platinated K^3-Bom species was observed.

A total of 8 products peaks were assigned for the reactions of Complex **1** and Substance P, and 10 for Complex **1** and K^3-Bom, not including sodiated adducts thereof. This study has revealed new photochemical reaction pathways which have not been anticipated on the basis of previous work on this photoactivatable metallo-prodrug.[23] Each new species is discussed below, along with the ECD MS/MS performed.

Oxidised products

Unexpected oxidation of both peptide and platinated peptide species was observed in the full nESI mass spectra of products from both peptides. These species were not observed in the mass spectra of the peptides without the addition of the metallodrug, or when either peptide was irradiated with blue visible light for the same period of time, and so are produced as a result of the photoactivation of the Pt(IV) metallodrug in the presence of biomolecules.

ECD MS/MS of the individually isolated oxidised peptide species (i.e. [SubP+O+2H]$^{2+}$, [K^3-Bom +O+2H]$^{2+}$, [K^3-Bom +O+2H]$^{2+}$, and [K^3-Bom +O+2H]$^{2+}$) are summarised in fragmentation maps (Figure 3.2, below), showing oxidation at the methionine residue for the Substance P+O species and oxidation of the methionine and tryptophan residues within the

[Lys]^3-Bombesin peptide for the K^3-Bom +O, K^3-Bom +2O, and K^3-Bom +3O species. Fully

assigned ECD MS/MS spectra are included in the supporting information (Figure SF2) along

with lists of all assigned species (SI Tables ST3-6).
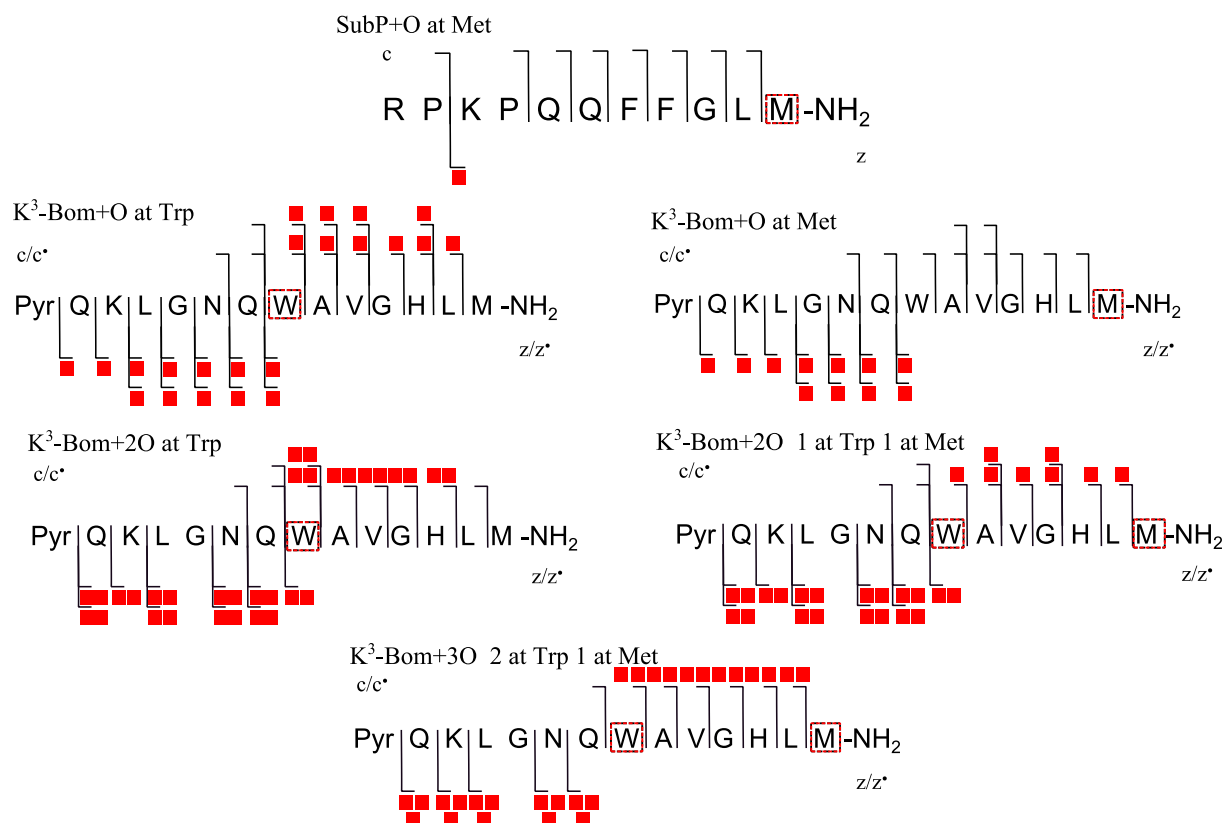
Figure 3.2: Oxidised peptide species observed and fragmented by ECD MS/MS, red squares indicate a modification observed on the particular fragment. As a result, the modifications can be readily located to single amino acid residues (modified residues are marked). Fully annotated tandem mass spectra can be found in the supporting information (Figure SF2).

The ECD MS/MS spectra of the mono- and di-oxidised K^3-Bom species ([K^3-Bom+2H +O]$^{2+}$

and [K^3-Bom+2H+2O]$^{2+}$) showed a mixture of products due to the 2 different reaction sites

of the oxidation. The mono-oxidised K^3-Bom species showed 2 product fragment

distributions; one relating to the oxidation occurring at the methionine residue and one

occurring at the tryptophan residue. The di-oxidised K^3-Bom species also showed two

product distributions; one relating to the oxidation of both the methionine and tryptophan

once each, and the other product showing oxidation of the tryptophan residue twice. These

product distributions were clearly separated using high resolution mass spectrometry and ECD MS/MS, which produces far fewer side chain/neutral losses (such as H_2O, CO, CO_2, and NH_3) which commonly populate CAD/CID MS/MS spectra and can mask multiple product distributions.[52] Distinct products were identified via fragments such as the c12 and c12+O in the $[K^3\text{-Bom} +2H+O]^{2+}$ ECD MS/MS spectrum (Figure 3.2) which clearly conflict, with each stating a different composition for the $[Pyr]^1$ to $[His]^{12}$ region of the species. Since ECD is known to preserve such modifications, only when a mixture of components is considered do the fragmentation patterns become clear and both product MS/MS fragments can be assigned accurately. Unfortunately the two oxidation products for Trp and Met oxidised K^3-Bom have exactly the same mass (K^3-Bom +15.994915Da) and so cannot be resolved/separated purely via MS, even with ultra-high resolving power, but would be distinguishable using MS/MS techniques. The species are also unlikely to separate definitively via liquid chromatography due to the oxidation producing such a small change in conformation, size, and hydrophobicity.

The terminal methionine residue present in both peptides have been shown previously to possess the potential to be oxidised once (to the sulfoxide)[53] or twice (to the sulfone).[54] Despite the multiple stages of oxidation observed as a result the reaction, the methionine residues present were shown via ECD MS/MS to be oxidised only once to the sulfoxide moiety.

The tryptophan residue present in the K^3-Bom peptide was shown to become oxidised up to two times during the photoactivation process (Figure 3.2). Tryptophan oxidation however can cause a multitude of reaction products,[55] causing different effects in structure and chemistry of the tryptophan residue side chain. Even more interestingly these products can be used diagnostically, as each product is characteristic of the process in which they were produced.[55] Work by Finley et al.[56] showed a range of products can be formed upon Trp oxidation, the 5

structures are summarised in Scheme 3.2 (below) along with corresponding calculated exact mass changes for each reaction.
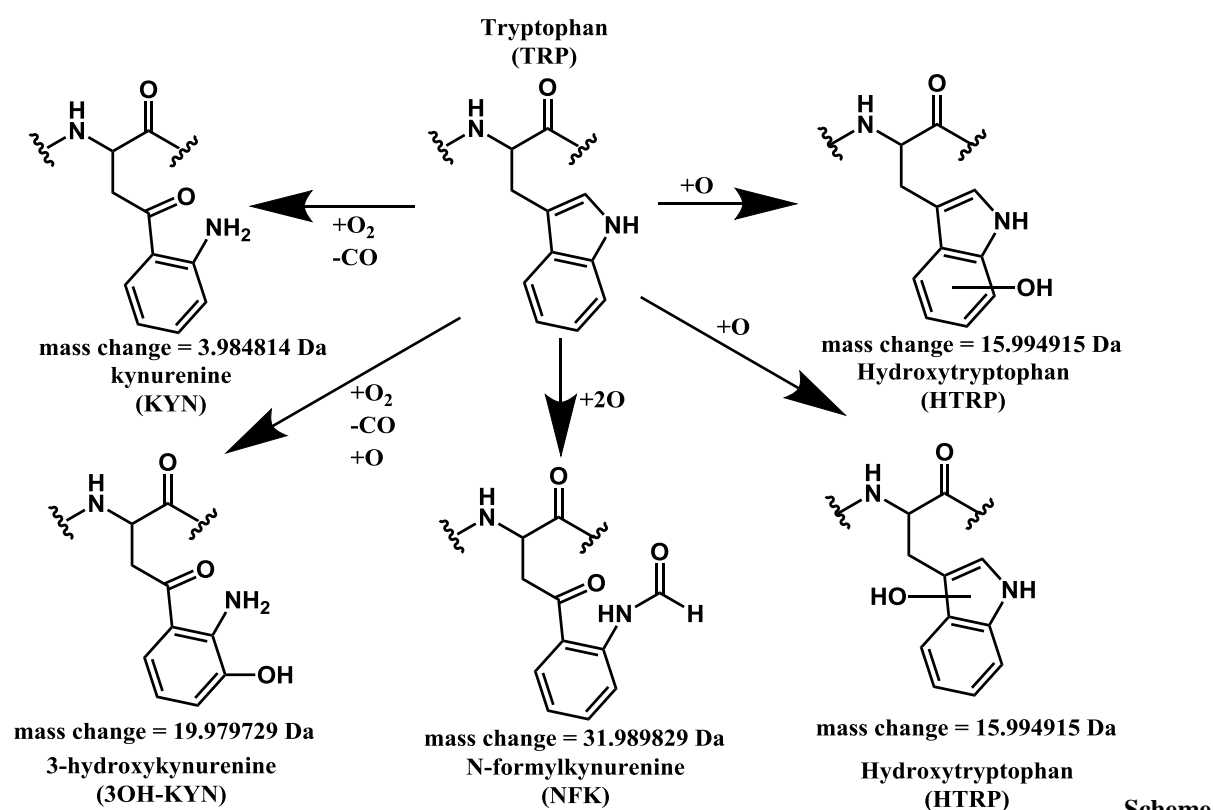
Scheme 3.2 -Possible oxidation products arising from singlet oxygen (KYN and 3OH-KYN) and radical oxidation (HTRP and NFK) of a tryptophan residue in a biomolecule[56,57] along with corresponding expected mass changes for the modification.

Furthermore work by Plowman *et. al.*[55] has shown that the nature of the oxidising species influences which Trp oxidation products are produced; hydroxy-tryptophan and N-formyl kynurenine (NFK) reaction products are formed during the mono and di-oxidation of tryptophan residues when the oxidation is achieved via hydroxyl radical induced oxidation, while kynurenine (KYN) and 3-hydroxy-kynurenine (3OH-KYN) species are produced when the tryptophan oxidation is achieved via singlet oxygen induced oxidation. These studies show distinct chemical structure changes upon oxidation of tryptophan residues. However, these reaction products also have different mass changes compared to the original reagents, as a result accurate-mass mass spectrometry can be used to determine not only the products of the oxidation reaction, but also provide insights into the mechanism of the oxidation itself,

despite the products residing within a mixture of many reaction products, which would have to be carefully separated (e.g. by High Pressure Liquid Chromatography (HPLC)) in order to achieve individual accurate H-NMR spectra. Scheme 2 summarises the findings of Finley *et al.* and Taylor *et al.,*[56] and also shows the calculated corresponding accurate mass changes due to each chemical reaction. From both the UHR-FT-ICR MS spectra of the reaction products (Figure 1a-b) and the corresponding MS/MS spectra (Figure S2a-d) the tryptophan oxidation products observed during this study are clearly the hydroxy-tryptophan (HTRP) and N-formyl kynurenine (NFK) moieties for the mono and di-oxidised species respectively showing the corresponding +15.994915 Da and 31.989829 Da mass shift in the mass spectrum and for tryptophan containing fragments in all MS/MS spectra (Figure S3.2a-d).

The HTRP and NFK modified species are both the result of a radical based oxidation reaction involving hydroxyl radicals (as shown by Plowman *et. al.*),[55] indicating that the release of the hydroxyl radicals from Complex **1** can then react with nearby biomolecules and cause oxidation of both methionine and tryptophan amino acid residues present.

Although both methionine and tryptophan oxidation have been observed previously, the metallodrug-induced biomolecule oxidation observed here was of particular interest as this process had not previously been observed for these Pt(IV) compounds and could present a method of inducing oxidative stress in target cancer cells during photoactivatable chemotherapy (PACT).[58] The release of reactive oxygen species (ROS) in cancer cells has been shown to trigger apoptosis[59] and so photoactivated release of these species can be used as a targeted method of action against cancer independent of the metal centre's cytotoxic/anti-proliferative characteristics which will carry out their own mechanism of action against the target cells upon binding to targets such as cancerous DNA. A platinum-independent MoA such as release of these ROS species may also help explain Complex 1's continued effectiveness against platinum (namely cisplatin) resistant strains of cancer cells. [21]

Due to the metal-independent nature of the ROS MoA, cancer cells which have gained resistance to platinum-based cancer therapy (for instance due to under-dosing or prolonged exposure/treatment)[60] will still be susceptible to an ROS based MoA, and so Complex **1** can continue to be an effective chemotherapy metallodrug. Unlike PDT treatments, since Complex **1** does not require an external source of oxygen to cause release of ROS, Complex **1** can trigger ROS-induced apoptosis in hypoxic cancer cells, such as those from ovarian cancer, which have been shown to be particularly resistant to oxygen dependent treatments such as PDT, Cisplatin, and Carboplatin,[29] and circumvents cancer cells "hypoxia-protection" towards anti-cancer therapy.[61]

Interestingly, oxidation of tryptophan to the N-formylkynurenine species is also a critical step in the biological production of NAD^+. By using Complex **1** to increase not only ROS concentration, but specifically to oxidise (free) tryptophan into NFK, production of oxidised tryptophan can push the equilibrium of this reaction to produce more NAD^+. Artificially altering the NAD^+/NADH balance inside target cancer cells has been recently shown[62] to be a new and effective mechanism of action to combat cancer cells via the use of catalytic metallodrugs to change the concentration of NAD^+/NADH inside cells and trigger cell death. Therefore Complex **1** could be a possible non-catalytic way of triggering cell death by a similar mechanism via production of NFK, meaning Complex 1 also has potential as a "kynurenergic" therapeutic compound.[63]

The effect of free Trp on the photoactivation of Complex **1** was previously reported[30] and showed a distinct change in the cytotoxicity of the complex towards target cancer cells. The observed drop in activity was at the time attributed to azide radical quenching by the free tryptophan residues, shown via EPR and further interrogated via isotope labelling of the azide ligands, followed by changes in the EPR results, correlating with the hypothesis of azide radical quenching. No reactions of the released azide radicals were shown to interact with

tryptophan residues in the K^3-Bom peptide studied. Though the released azide radicals are effectively quenched by free tryptophan, bound tryptophan appeared to be less susceptible to free azide radicals. Observation of a drop in cytotoxicity due to additional free Trp would correlate well with the results observed here for Trp oxidation and the hypothesised impact on NAD^+ production. If additional free Trp was added to a target cell environment, OH radicals produced during photoactivation of Complex **1** could be quenched, causing Trp oxidation to the hydroxyl tryptophan and NKF species (as shown in this study), in addition to released azide radical quenching observed previously.[30]

UHR-MS is shown here to be an effective means of not only observing these modifications, but also elucidating and identifying the mechanisms leading to their production, which can facilitate further study and design of ROS releasing metallodrugs for treatment of platinum resistant and/or hypoxic cancer cells.

Photoactivation of Pt(IV) complex in the presence of histidine-containing biomolecules:

In addition to biomolecule oxidation, an array of platinum-based modifications were also observed following photoactivation of Complex **1** in the presence of biomolecules, discussed above and shown in Figure 3.1.

[Lys]^3-Bombesin (PyrKLGNQWAVGHLM-NH$_2$ aka K^3-Bom) contains a terminal methionine residue, a single histidine residue and a single lysine residue, all of which have the potential to donate electron density to a Pt metal centre (via a Sulphur/Nitrogen lone pair) and form a stable dative covalent bond. Extensive previous studies on Pt(II) complexes, such as Cisplatin, Carboplatin, and Oxaliplatin, have shown an almost uniform primary preference for methionine sulphur binding,[14,37,64–66] often with a smaller secondary preference for histidine residues. Sulphur being a large, soft ligand has been shown to participate in very

strong bonds to platinum complexes, causing deformation of structure and intra-strand crosslinks,[37] strong enough to survive harsh pH digestion conditions and high-energy gas phase dissociations in tandem MS experiments.[14]

Electron capture Dissociation (ECD) MS/MS spectra of platinated K^3-Bom species are summarised and shown in Figure 3.3. Platinum containing species (in both MS and MS/MS scans) were readily identified due to platinum's characteristic isotopic pattern affecting the overall pattern observed (shown in Figure 1, inset). Full assignment lists and ECD MS/MS spectra can be found in the supporting information ST7-9 and SF3 respectively. The mono-dentate Pt modifications (namely $\{Pt(py)_2(N_3)\}^+$ and $\{Pt(py)_2(OH)\}^+$) were shown to bind unambiguously to the His^{12} residue of K^3-Bom, despite the methionine sulphur availability at the biomolecule terminus.
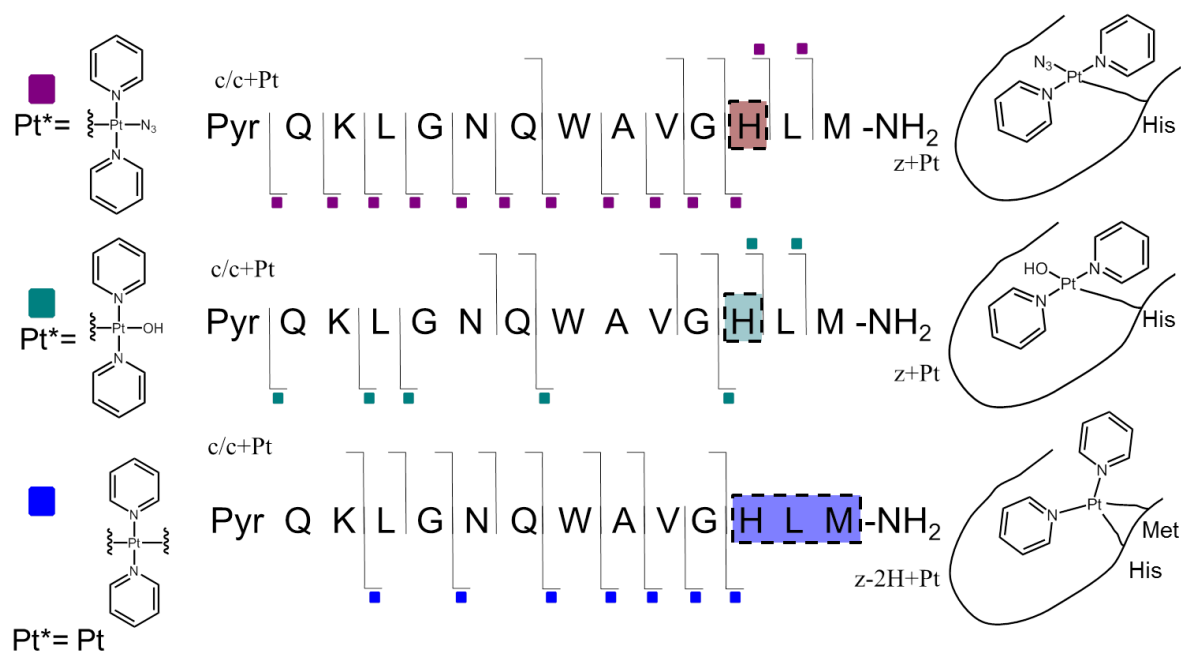
Figure 3.3: Platinated K^3-Bom peptide species observed and fragmented by ECD MS/MS, coloured squares indicate a modification observed on the particular fragment. As a result the modifications can be readily located to single amino acid residues. Fully annotated spectra can be found in the supporting information (Figure SF3.3).

Uniform retention of pyridine ligands during the photoactivation of Complex **1** correlates well with previous ^{1}H-NMR experiments.[21,27] However retention of one azide ligand (in the case of the $\{Pt(py)_2(N_3)\}^+$ modification) was not observed previously, azide release was observed via UV/Vis, EPR, and NMR experiments, however these techniques suffer when a range of products are formed, the production of $\{Pt(py)_2\}^{2+}$ modified species may mask azide retention and/or unreacted Complex **1** may impede spectroscopy data analysis. However, clear peaks for various platinated species can be observed simultaneously in the UHR-MS data presented, showing release of 1-2 hydroxyl ligands and 1-2 azide ligands.

ECD MS/MS analysis of the $[K^3\text{-Bom}+Pt(py)_2]^{2+}$ species yielded no sequence informative fragments, only ligand and/or side chain loss peaks, which is usually attributed to extended cyclic structures which can disrupt usual MS/MS fragmentation.[67] However the presence of metal centres present unique challenges to MS and MS/MS analysis due to their unique behaviour compared to unmodified peptide/protein MS/MS, here lack of backbone fragmentation is attributed to both positive charges present on the ion originating from the Pt^{2+} centre, without additional protons along the poly(amino acid) backbone, the normal electron capture dissociation mechanism can be disrupted/impeded.[68,67,69,70] and electron capture at the Pt centre would dominate, causing release of bound ligands and side chain losses from nearby amino acids (discussed below). ECD MS/MS of the $[K^3\text{-Bom} +Pt(py)_2+H]^{3+}$ species produced abundant backbone fragmentation (Figure 3d) and enabled location of the binding site to be determined as the His[12] and Met[14] residues. The available proton allows for usual ECD fragmentation along the peptide backbone, however at least two electrons can still be captured at the Pt centre causing the previously observed ligand and side chain losses described (discussed below).

Photoactivation of Pt(IV) complex in the presence of non-histidine-containing biomolecules:

Photoactivation of Complex **1** in the presence of a non-histidine containing peptide (in this case Substance P – RPKPQQFFGLM-NH$_2$) was shown to produce a dramatically different range of reaction products (Figure 3.1a). Previously (with K^3-Bom, Figure 3.1b) uniform retention of the pyridine ligands was observed and histidine binding was the main pathway for all adducts. However Substance P contains no histidine residue, but still possesses a terminal methionine residue and a lysine residue at position 3, both of which are able to bind to a Pt metal centre, though previous studies indicate Pt compounds would show a heavy preference for the Met residue.[14,71,72]
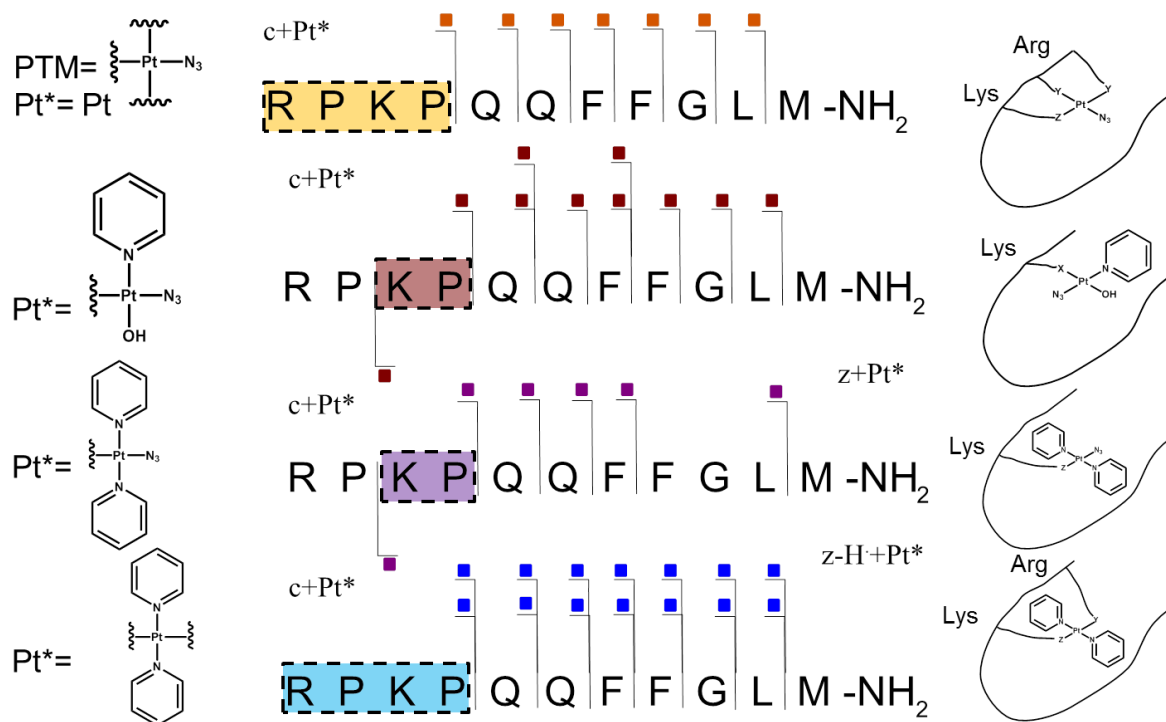
Figure 3.4: Platinated SubP peptide species observed and fragmented by ECD MS/MS, coloured squares indicate a modification observed on the particular fragment. Fully annotated spectra can be found in the supporting information (Figure SF4).

185

Interestingly the Substance P Pt modifications observed did not show uniform retention of the pyridine ligands, instead the main adduct observed was Substance P modified with the $\{Pt(py)(OH)(N_3)\}$ moiety. The presence of this adduct (and the dominant intensity of it) indicates that not only can any of the bound ligands be released during the photoactivation process, but that the products of photoactivation depend directly on the ligands available in the vicinity of the complex during the activation, and so products are biomolecule dependant as are binding sites (discussed below).

The $[SubP+\{Pt(py)(OH)(N_3)\}+2H]^{2+}$, $[SubP+\{Pt(py)_2(N_3)\}+H]^{2+}$, $[SubP+\{Pt(N_3)+H]^{2+}$, and $[SubP+\{Pt(py)_2+H\}]^{3+}$ species were shown (via ECD MS/MS) to exhibit the platinum modification within the N-terminal 4 amino acids of the peptide (RPKP) (Figure 3.4, above), annotated MS/MS spectra can be found in the supporting information (SI Figure SF3.4) along with lists of assignments (SI Tables ST3.10-12). ECD induced fragmentation is limited by proline residues,[73] although the N-Cα bond of proline is cleaved during ECD MS/MS, the overall structure is held together via the cyclic alkyl side chain group of the amino acid residue, the so called "proline effect".[74] Unfortunately the proline effect limited fragmentation towards the N-terminal region of substance P (RPKP) due to the 2 proline residues (P2 and P4), however observation of the z_9-H+$\{Pt(py)(OH)(N_3)\}$ fragment showed the location of the modification to be between the Lys[3] and Pro[4]. The $[SubP+\{Pt(py)_2(OH)\}+H]^{2+}$ species was observed at very low intensity at all irradiation times, as a result reliable MS/MS fragmentation of this species was not possible.

It is worth noting that the Complex **1**+ K[3]-Bom reaction mixture results shown were produced via irradiation for 2 hours since all products formed were present in shorter irradiation time samples (30, 60, 90 minutes) and seemed to increase in intensity with increasing irradiation time. The Complex **1** + SubP reaction mixture products were observed in varying intensities at different time points throughout the reaction; with

$[SubP+Pt(py)(OH)(N_3)+2H]^{2+}$ remaining the predominant product at shorter irradiation times. Increasing irradiation of the Complex **1** + SubP mixture encouraged further ligand release, decreasing intensity of multi-ligand complexes and increasing the intensity of the $[SubP+Pt(N_3)+H]^{2+}$ species over time, indicating that photoactivation is not a single transition from a Pt(IV) to Pt(II) species, but can continue after the Pt(II) species is formed, releasing further ligands and altering the product intensities observed. As a result all possible light was eliminated during sample preparation and mass spectrometry analysis of these reaction samples, to avoid further ligand release, which was observed even under low levels of distant white light. Low levels of red light were shown not to activate Complex **1** over several hours, as a result low levels of red light were used during analysis. Lack of activation under low levels of red light used during analysis and high levels used in control spectra, correlate well with previous results[21] (also with UV/vis experiments conducted herein, see below) showing Complex **1** absorbs light in the blue region effectively, while also expressing a smaller absorbance to green light, which can activate the sample more slowly, yet showing no absorbance in the red light region of the visible spectrum.

Pt induced ligand loss and Pt associated effects during electron dissociations:

As mentioned previously, the ECD MS/MS spectrum of each Pt adduct showed a series of high intensity, singly-charged peaks close to the charge reduced species (SI Figures SF3.3a-d and SF3.4a-d). These peaks did not match any calculated usual c or z ion, nor any side chain loss commonly observed during ECD,[75] but the highest m/z peak did correspond to the mass of the peptide+Pt (e.g. $[\text{Substance P+Pt}]^{+}$ for the $[\text{Substance P+Pt(py)}_2(N_3)+H]^{2+}$ species), indicating an ECD-induced loss of ligands from Pt. The peak spacings were 15.02374 m/z and 31.97186 m/z. This corresponds to two separate losses of a terminal methyl group $\cdot CH_3$ (-0.16 ppm) and loss of $\cdot CH_3\text{-}SH$ (46.9956 m/z from spectrum, -0.25 ppm) (SI Figure SF3.7). Considering the sequences of Substance P (RPKPQQFFGLM-NH$_2$) and K^3-Bom

(PyrQKLGNQWAVGHLM-NH$_2$) the only sulphur atom present is within the side chain methionine. The observed side chain loss differs from the usual side chain losses commonly observed for methionine which is C$_3$H$_6$S or C$_2$H$_4$S,[75] suggesting that the Pt centre has affected the usual fragmentation of the peptide under ECD.[76] For this to occur, Pt would have to be in close proximity to the methionine residue, as seen in previous work by Li *et al*.[36] yet all current data points clearly to the binding of Pt to the histidine (His[11]) and lysine (Lys[3]) residues, depending on the amino acid sequence. Methionine sulphur is known to be a strong ligand for Pt(II),[14,37] and the side chain losses here indicate a localised interaction. As a result this localised reaction must occur after the electron capture-induced loss of the Pt-bound ligands, the methionine side chain can then coordinate, to vacant sites on the Pt centre and account for the side chain losses observed. A proposed mechanism for the interaction of a methionine residue and the subsequent loss of the terminal methyl and CH$_3$-SH species from the peptide is shown in Scheme 3.3. This scheme correlates with the side chain losses observed during this study, with the predicted interaction with methionine found in several reported Pt peptide MS/MS studies,[16,37,77,] Pt- centred capture of an electron during ECD,[36] and commonly observed in Pt-ligand binding studies in solution.[78,79]
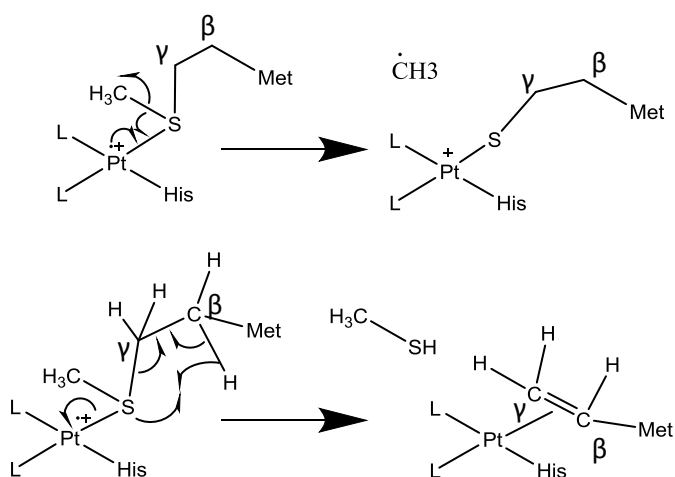
Scheme 3.3: Platinum centred side chain losses from methionine commonly observed during presented ECD MS/MS studies of platinated peptides.

Scheme 3.4 (Supporting information (SF3.7)) shows a proposed mechanism for the loss of the platinum complex from peptides during electron capture, often observed for singly bound Pt modifications in this study (single binding site on peptide).

ECD MS/MS of the diplatinated K^3-Bom species produced a similar array of side chain losses as observed in mono-platinated species, however a more extensive number and variety of unusual sidechain losses and fragmentation channels were observed (SI Figure SF3.10). Unfortunately as observed in the ECD MS/MS of the $[K^3\text{-Bom}+Pt(py)_2]^{2+}$ species, without an additional proton, ECD MS/MS was unable to provide enough sequence informative fragments to assign binding positions of the two platinum complexes, though based on the data observed for the previous modifications we would tentatively expect these modifications to be present on the His^{12} and Met^{14} residues (for $\{Pt(py)_2\}$) and the Lys^3 (for $\{Pt(py)_2(N_3)\}^+$).

<u>Slow heating MS/MS of platinated species:</u>

It should be noted that more commonly used Collisionally Activated Dissociation (CAD) and Infra-Red Multi Photon Dissociation (IRMPD) MS/MS techniques were attempted on the observed platinum-peptide adducts, however these ergodic (so called "slow heating") fragmentation methods break bonds in the lowest energy pathways in a molecule;[80,81] for

peptides, these usually involve the amide bonds along the backbone. However, CAD and IRMPD on the Pt adducts resulted in a series of peaks corresponding to the loss of each platinum bound ligand present, loss of the platinum modification (unmodified peptide), and a series of unassignable Pt and non-Pt containing peaks due to sequence scrambling commonly observed during the study of peptides and (gas-phase) labile post translational modifications such as metallation.[82,83] In contrast, ligand loss was observed due electron capture at the metal centre during ECD MS/MS of the Pt adducts studied here, electrons could also be captured at points along the backbone according to the accepted ECD MS/MS mechanism,[68] allowing normal ECD fragmentation and production of interpretable c/z ions.

It should also be noted that although the Pt complex showed the potential to bind to multiple groups present on the peptides studied; no crosslinking was observed, despite the low drug-to-peptide mol ratio (0.5:1) which might have promoted this possibility. The lack of crosslinking observed could be due to the short irradiation (reaction) times used (compared to the ~24 hour times used for Cisplatin crosslinking studies)[37] and the low concentrations of reagents, reducing the rate of tri-molecular reactions, though this has not been a limitation previously.[77] Square-planar Pt(II) has previously shown an ability to act as an effective crosslinking agent, even for proteins such as calmodulin.[77] Although lack of crosslinking could also be attributed to the mechanism of Pt binding to the biomolecules; with Cisplatin-type complexes relying on a solvent/biomolecule displacement of bound ligands in order to form biomolecule-Pt bonds, while the Pt based modifications observed herein were produced from a much faster and distinctly different photoactivation process. Due to the lack of a slow solvent-assisted ligand displacement reaction mechanism, the fast photo-induced reaction may be not as susceptible to further solvent/biomolecule displacement of bound ligands, limiting second/third binding of external species to the platinum complex. Lack of solvent assisted ligand dissociation correlated well with results found in this study, as no solvent

adducts/solvent displaced ligand complex species were observed and no reactions showed evidence of being displaced by solvent/biomolecule ligands post photoactivation without additional white/blue/green light, even after several hours in solution. A lack of crosslinking potential, while still expressing potent cytotoxic activity against cancer cells, suggests photoactivated Complex **1** may inhibit cancerous DNA in a different fashion to Pt(II)/Cisplatin reagents, which are believed to disrupt cancerous DNA/RNA most significantly via intra-strand DNA cross-linkages, preventing transcription of individual DNA strands,[6] which warrants future investigation.

Electron paramagnetic resonance:

EPR has previously been used to effectively monitor the radicals released during the photoactivation of Pt(IV) metallodrug complexes.[30] The FT-ICR MS data presented here has shown the effect that the (now identified) hydroxyl radicals can have upon nearby biomolecules and the extent of oxidation which can be achieved. However previous work has shown the release of azide radicals during this photoactivation process.[3084] These findings were mirrored in the work presented here which shows Pt(II) based modifications with only a single or no azido ligands remaining bound to the Pt(II) centre post-photoactivation, clearly indicating 1-2 azide radicals are released in the process. One would expect such a reactive radical species to react with the nearby biomolecules and cause modifications, however no azide modification of the peptides or platinated peptides was observed at any point during this study. In order to monitor the release of the azide radicals, and to explore the possibility that the peptides themselves could be influencing the release of the azide radical species, EPR was employed to study the photoactivation of Complex **1**, Complex **1**+Substance P, and Complex **1**+[Lys][3]-Bombesin, the results are shown in Figure 3.5, below (EPR experimental values calculated can be found in Supporting information tables ST3.14a-c).
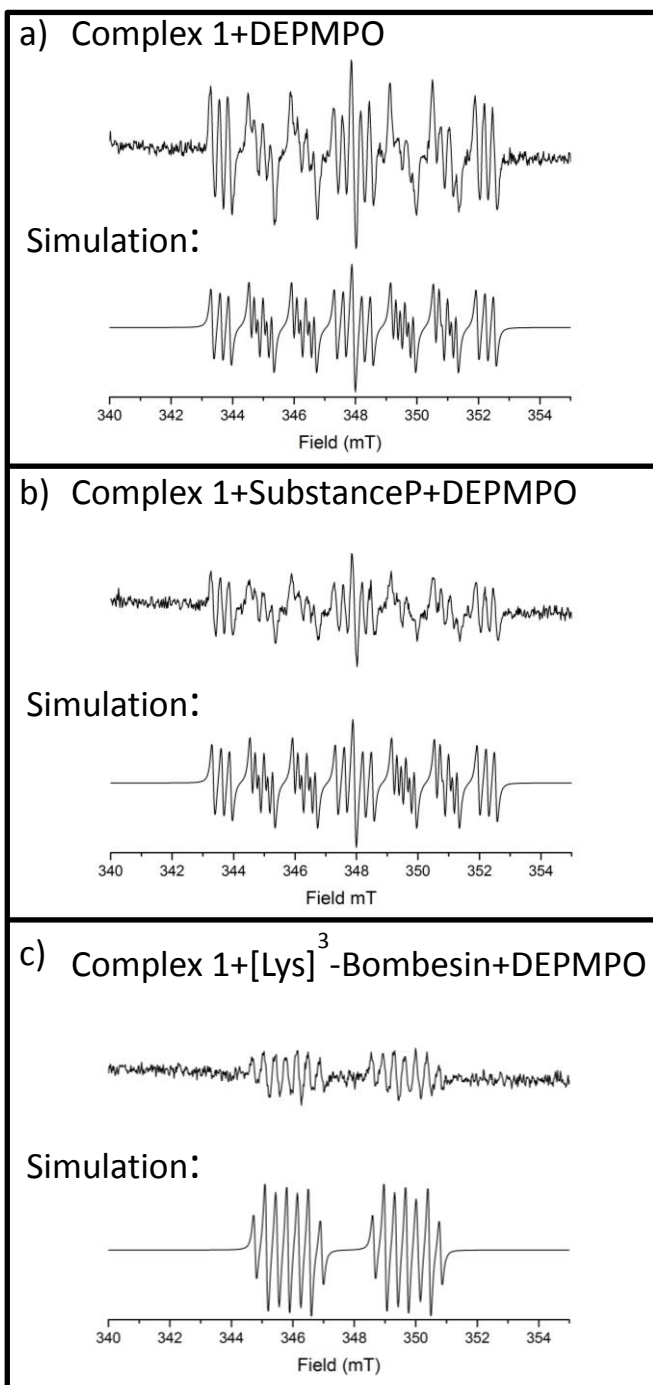
Figure 3.5 – EPR scans for photoactivation (via blue visible light) of (a) Complex 1 (b) Complex 1 + Substance P and (c) Complex 1 + [Lys]³-Bombesin - showing spin trap oxidation. Simulations of corresponding patterns are included below their corresponding observed signal.

EPR of the samples was conducted using the DEPMPO spin trap, specifically to be able to differentiate between azide and hydroxyl radicals, should both be produced simultaneously.[85] EPR using the DEPMPO produces more complicated splitting patterns than the more commonly used DMPO trap, however Figure 3.5(a) shows clearly that the azide radicals are being effectively quenched from the reaction, trapped by the DEPMPO and producing an EPR spectrum similar to Butler *et al.* for this system.[30] The EPR spectrum of Complex **1**+Substance P shows a similar EPR spectrum to Figure 3.5(a), with a slightly lower intensity for the same time points (SI Figure SF3.9), indicating that Substance P does not appear to be interfering with the production of azide radicals during photoactivation, however it does suggest fewer are being trapped by the DEPMPO spin trap, this could arise from the stabilisation of azido-Pt(II) complexes by the substance P peptide, as shown in the FT-ICR MS there are multiple azido-containing species detected, as a result fewer would be released, trapped, and detected.

The EPR spectrum of the Complex **1**+[Lys]^3-bombesin sample was dramatically different to the Complex irradiated without a peptide present. The EPR splitting pattern does not resemble azide or hydroxyl radical quenching by DEPMPO. Instead it was found that the pattern matched that of oxidised DEPMPO (Scheme 3.5, below).[85] The large extent of oxidation occurring during Complex **1** photoactivation in the presence of K^3-Bom (detected via FT-ICR MS) appears to spread to oxidation of the DEPMPO also (shown here in the EPR spectra). The FT-ICR MS and MS/MS spectra above show the clear evidence for production of hydroxyl radicals reacting with biomolecules when available, but no azide modification. The EPR spectra are able to monitor the azide radicals released which do not appear to react with the peptide Substance P and create a new radical species (which would have produced a different EPR splitting pattern to that of Complex **1** alone).
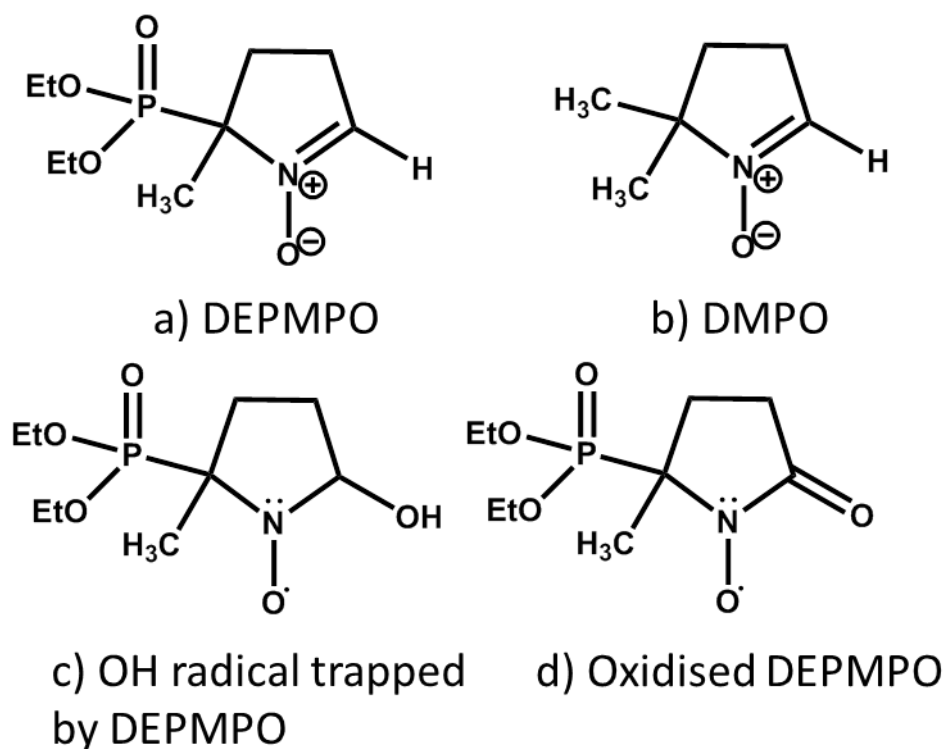
Scheme 3.5: a) DEPMPO spin trap used during the EPR studies of Complex 1 photoactivation, b) Usual DMPO spin trap commonly used for EPR, c) DEPMPO+trapped OH radical structure detected by EPR, [85] d) Oxidised DEPMPO caused by high concentration of OH radicals produced via photoactivation of Complex 1. [85]

UV/Visible spectroscopy:

Previously,[23] it has been shown that Complex **1** produces a very pronounced charge transfer band when studied by UV/visible spectrometry, this band corresponds to Pt-azide charge transfer and can be used to monitor the extent of the bond photoactivation. Photoactivation of Complex **1** was again studied, also in the presence of the Substance P /[Lys]^3-Bombesin peptide in order to see if UV/vis spectroscopy could detect any influence the biomolecule's presence was having upon the photoactivation process. Complex **1** and both peptides were all scanned individually for baseline measurements, the two peptides showing minimal UV/vis absorbance in the region studied. Complex **1** was then photoactivated in the presence of either peptide and the mixture was scanned at various time points. The resulting UV/vis traces can be found in the supporting information (SI Figure SF3.10a-c). Substance P has an

unusually low UV/vis absorbance for a peptide (despite phenylalanine residues), and so produced a photoactivation profile almost identical to that of Complex **1** alone (SF3.10a and SF3.10b). The observation of a similar spectrum to that of Complex **1** itself suggests Substane P is not having a significant effect upon the rate of photoactivation of the Pt-Azide bond. K^3-Bom effectively absorbs UV/visible light and produces a characteristic UV/vis spectrum (SI figure SF10.9c). Photoactivation of Complex **1** in the presence of K^3-Bom showed steady production of a new peak shifted ~50nm from the original Pt-azide band, which increased with irradiation time, indicating new photoproducts produced during the reaction, most likely the azide-containing Platinated K^3-Bom species observed previously in the FT-ICR MS spectra (above).

<u>The effect of free tryptophan:</u>

Addition of free tryptophan during the photoactivation process of Complex **1** was previously shown to have a dramatic effect on the cytotoxicity measurements of the metallodrug, this was believed to be due to tryptophan's ability to trap the radicals released during the photoactivation process.[30] Very little additional tryptophan was needed to cause a dramatic effect on the cytotoxicity measured, in fact only 1/8th of the concentration of free Trp (with respect to Complex **1**) was needed to see a 70% change in cytotoxicity, which presented the opportunity to use free Trp as a molecular switch for activity of Complex **1** during treatment. Due to this dramatic effect during biological tests it was decided to investigate the effect of additional free tryptophan on the products and product ratio's observed for the Complex **1**+peptide reactions studied here.

Reaction mixtures were prepared as for previous experiments with the exception that additional free tryptophan was added before the photoactivation step, end concentration of tryptophan was 1/8 the concentration of Complex **1** (end concentration 31.25 µM, 3.1 nMoles Trp added).

The SubstanceP+Complex **1**+Trp reaction mixture mass spectrum showed the same reaction products as previously observed, no different/previously observed species were/were not detected as a result of free Trp addition. Product ion intensities were broadly similar to the previous SubstanceP+Complex **1** sample, though oxidised substance P and oxidised Substance P+platinum adducts were observed in a lower intensity (29% and 95% lower respectively), indicating that the Trp may be reducing the available ROS (namely the hydroxyl radicals) and limiting biomolecule oxidation. The K^3-Bom + Complex **1**+Trp reaction mixture showed a much greater difference in intensities; every oxidised/Platinated species (aside from the $[K^3$-Bom+O+2H$]^{2+}$ and $[K^3$-Bom+3O+2H$]^{2+}$ species) were reduced

in intensity by 32-98% compared to the previous reaction mixture without additional tryptophan. The drastic differences in product ion intensities indicated radical trapping agents could play a large role in the formation of certain platinated adducts in solution, although intensities may have varied between the two reaction mixtures; all previously observed products were also observed in the added tryptophan reaction mixtures. Graphs showing the relative ion intensities of the individual reaction products are shown in Figure SF3.5, Tables showing the relative intensity changes are shown in the supporting information, Tables ST3.13a and b.

Previous results from Butler *et al.*[30] showed a 70% drop in the number of radicals using the same 1/8th concentration of free tryptophan. For the SubstanceP+Complex **1**+Trp reaction mixture an average relative intensity drop of 17% was observed over all reaction products. For the K^3-Bom +Complex **1**+Trp reaction mixture an average relative intensity drop of 67% was observed, which is very similar to previous results by Butler *et al.* using NMR and EPR studies.[30] As many peptides and proteins in cells are likely to contain at least one tryptophan and/or histidine residue; most of the Complex **1**+biomolecule products may resemble those observed in the Complex **1**+ K^3-Bom mass spectrum, showing uniform retention of both pyridine ligands, and also showing a similar drastic effect to the addition of free tryptophan. Substance P on the other hand contains neither tryptophan or histidine residues and so we are able to show here the differences in susceptibility to this effect, which appears to be much less extreme in the case of a non-His/Trp containing peptide, which is most likely due to the nature of the platinum complexes formed, the sequence dependent binding, and the ligands released as a result. Though no products involving azide radical reactions/damage were observed in any the mass spectra obtained during this study, they may also be affected by the addition of free tryptophan (as previously shown the EPR and NMR studies).

Conclusions:

Photoactivatable platinum(IV) complexes have been shown to possess potent anticancer potential and provide a new level of spatial and temporal control over the drug activation process in photo activated chemotherapy (PACT). Here a novel Pt(IV) PACT agent, *trans, trans, trans*-[Pt(N$_3$)$_2$(OH)$_2$(py)$_2$] (Complex **1**), was activated in the presence of model biomolecules (Substance P and [Lys]^3-Bombesin peptides) in order to gain insights into the species produced during the photoactivation process and the species which would be produced if activated in the body.

When histidine moieties were available (i.e. when activated in the presence of [Lys]^3-bombesin), complete retention of both pyridine ligands was observed (correlating with previous ^{1}H-NMR studies), however when this residue was not available (i.e. when activated in the presence of Substance P) a very different range of products was observed, many of which did not retain both, or even one of the pyridine ligands. The discovery of the multiple possible platinum modifications due to the photoactivation process has helped provide limited explanation for the interesting biological data observed for these metallodrugs, most notably the obvious differences in chemistry and mechanisms of action observed for this compound in comparison with previous generations of platinum(II) based anti-cancer metallodrugs, such as Cisplatin.

Other interesting features observed included the large extent of oxidation caused by irradiation of this compound in the presence of biomolecules. Release of multiple hydroxyl ligands was seen to cause multiple stages of oxidation to nearby peptides, both with and without the platinum complex modification. For Substance P; one incidence of oxidation was observed at the methionine residue. However for [Lys]^3-bombesin, up to three stages of oxidation were observed; high resolution MS/MS experiments showed oxidation at the

methionine residue (x1) and the tryptophan residue (x2), using the mass differences observed and previous oxidation studies of biomolecules[56] it was confirmed that the hydroxy-tryptophan (HTRP) and N-formylkynurenine (NFK) moieties were formed. The preference for the tryptophan site, an electron acceptor and redox active amino acid,[86] and the formation of the HTRP and NFK moieties both re-enforce the conclusion that this was a radical based oxidation process, not a singlet-oxygen process as previously believed for some Pt(IV) photoactivated compounds.[84]

Though addition of free tryptophan had been previously shown to cut the number of radicals by 70% by EPR and NMR studies of in cell reactions,[30] a much more modest reduction in product ion intensities of 17% was observed for a SubstanceP+ Complex **1**+tryptophan reaction mixture of comparable reagent ratios. A much greater reduction of 67% was observed for the K^3-Bom + Complex **1**+tryptophan reaction mixture, and showed decreases in relative intensity of up to 98% for certain reaction products, indicating that radical trapping agents can have a dramatic effect on the products observed in photoactivatable processes involving homolytic bond cleavage.

The multitude of reactions which occur following the photoactivation of Pt(IV) compounds have been shown to produce a range of products, varying greatly, and producing a mixture of different species which would be challenging to separate and study with commonly used analytical techniques. UHR-MS and MS/MS experiments offer a unique and highly accurate method of studying these species and can provide unambiguous determination of modifications and modification binding site, despite the wide variety of products produced. Study into the fundamental mechanism of action of these new and interesting compounds is key to informing and developing improved anti-cancer strategies such as photoactivated chemotherapy therapy, and UHR-MS is becoming increasingly desirable to use when attempting to study such complex problems.
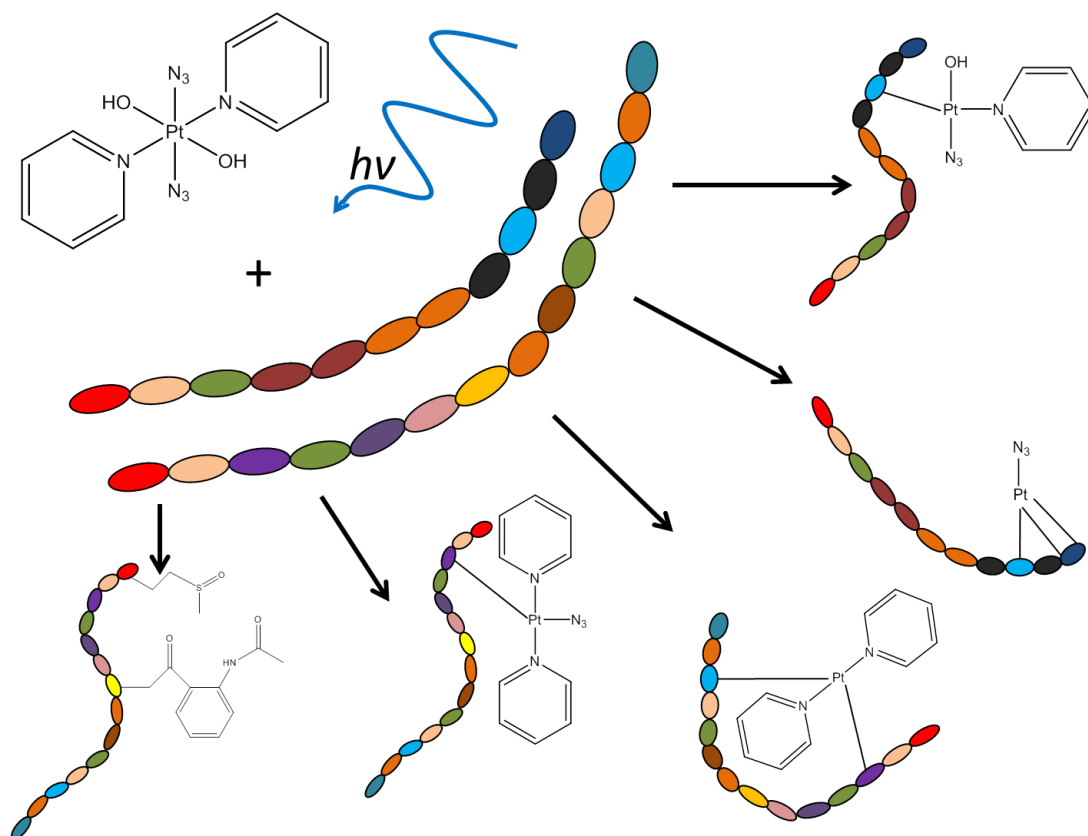
Figure 3.6: Summarising figure of the reaction products observed during the study of Complex 1 with the model peptides

References:

(1)   Alderden, R. A.; Hall, M. D.; Hambley, T. W. *J. Chem. Educ.* **2006**, *83* (5).

(2)   Biot, C.; Castro, W.; Navarro, M.; Botte, C. Y. *Dalt. Trans.* **2012**, *41*, 6335–6349.

(3)   Joseph, J.; Nagashri, K.; Janaki, G. B. *Eur. J. Med. Chem.* **2012**, *49*, 151–163.

(4)   De Clercq, E. *Met. Based. Drugs* **1997**, *4* (3), 173–192.

(5)   Wong, E.; Giandomenico, C. M. *Chem. Rev.* **1999**, *99* (9), 2451–2466.

(6)   Pillaire, M. J.; Hoffmann, J. S.; Defais, M.; Villani, G. *Biochimie* **1995**, *77* (10), 803–807.

(7)   Burstyn, J. N.; Heiger-Bernays, W. J.; Cohen, S. M.; Lippard, S. J. *Nucleic Acids Res.*

2000, *28* (21), 4237–4243.

(8)   Timerbaev, A. R.; Hartinger, C. G.; Aleksenko, S. S.; Keppler, B. K. *Chem. Rev.* **2006**, *106* (6), 2224–2248.

(9)   Deconti, R. C.; Toftness, B. R.; Lange, R. C.; Creasey, W. A. **1973**, No. Ii, 1310–1315.

(10)   Casini, A.; Gabbiani, C.; Michelucci, E.; Pieraccini, G.; Moneti, G.; Dyson, P. J.; Messori, L. *J. Biol. Inorg. Chem.* **2009**, *14* (5), 761–770.

(11)   Orvig, C.; Abrams, M. J. *Chem. Rev.* **1999**, *99* (9), 2201–2204.

(12)   Dyson, P. J.; Sava, G. *Dalton Trans.* **2006**, No. 16, 1929–1933.

(13)   Allardyce, C. S.; Dyson, P. J.; Coffey, J.; Johnson, N. *Rapid Commun. Mass Spectrom.* **2002**, *16* (10), 933–935.

(14)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 9507–9515.

(15)   Weidt, S. K.; Mackay, C. L.; Langridge-Smith, P. R. R.; Sadler, P. J. **2007**, *2* (c), 1719–1721.

(16)   Hartinger, C. G.; Tsybin, Y. O.; Fuchser, J.; Dyson, P. J. *Inorg. Chem.* **2008**, *47* (1), 17–19.

(17)   Egger, A. E.; Hartinger, C. G.; Ben Hamidane, H.; Tsybin, Y. O.; Keppler, B. K.; Dyson, P. J. *Inorg. Chem.* **2008**, *47* (22), 10626–10633.

(18)   Groessl, M.; Tsybin, Y. O.; Hartinger, C. G.; Keppler, B. K.; Dyson, P. J. *J. Biol. Inorg. Chem.* **2010**, *15* (5), 677–688.

(19)   Qi, Y.; Liu, Z.; Li, H.; Sadler, P. J.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2013**, *27* (17), 2028–2032.

(20)   Wootton, C. a.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor, P. B. *Dalt. Trans.* **2015**, *44* (8), 3624–3632.

(21)   Farrer, N. J.; Woods, J. a; Salassa, L.; Zhao, Y.; Robinson, K. S.; Clarkson, G.; Mackay, F. S.; Sadler, P. J. *Angew. Chem. Int. Ed. Engl.* **2010**, *49* (47), 8905–8908.

(22)   Westendorf, A. F.; Woods, J. a; Korpis, K.; Farrer, N. J.; Salassa, L.; Robinson, K.; Appleyard, V.; Murray, K.; Grünert, R.; Thompson, A. M.; Sadler, P. J.; Bednarski, P. J. *Mol. Cancer Ther.* **2012**, *11* (9), 1894–1904.

(23)   Pracharova, J.; Zerzankova, L.; Stepankova, J.; Novakova, O.; Farrer, N. J.; Sadler, P. J.; Brabec, V.; Kasparkova, J. *Chem. Res. Toxicol.* **2012**, *25* (5), 1099–1111.

(24)   Farrer, N. J.; Sadler, P. J. *Aust. J. Chem.* **2008**, *61* (9), 669.

(25)   Funston, A. M.; Cullinane, C.; Ghiggino, K. P.; McFadyen, W. D.; Stylli, S. S.; Tregloan, P. a. *Aust. J. Chem.* **2005**, *58* (3), 206.

(26)   Bednarski, P. J.; Mackay, F. S.; Sadler, P. J. *Anticancer. Agents Med. Chem.* **2007**, *7* (1), 75–93.

(27)   Mackay, F. S.; Woods, J. a; Heringová, P.; Kaspárková, J.; Pizarro, A. M.; Moggach, S. a; Parsons, S.; Brabec, V.; Sadler, P. J. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (52), 20743–20748.

(28)   Paper, F. **2008**, No. 05, 202–207.

(29)   Song, X.; Liu, X.; Chi, W.; Liu, Y.; Wei, L.; Wang, X.; Yu, J. *Cancer Chemother. Pharmacol.* **2006**, *58* (6), 776–784.

(30)   Butler, J. S.; Woods, J. a; Farrer, N. J.; Newton, M. E.; Sadler, P. J. *J. Am. Chem. Soc.* **2012**, *134* (40), 16508–16511.

(31)   Ryan, C. M.; Souda, P.; Bassilian, S.; Ujwal, R.; Zhang, J.; Abramson, J.; Ping, P.; Durazo, A.; Bowie, J. U.; Hasan, S. S.; Baniulis, D.; Cramer, W. a; Faull, K. F.; Whitelegge, J. P. *Mol. Cell. Proteomics* **2010**, *9* (5), 791–803.

(32)   Yin, S.; Loo, J. a. *Int. J. Mass Spectrom.* **2011**, *300* (2-3), 118–122.

(33)   Xu, F.; Xu, Q.; Dong, X.; Guy, M.; Guner, H.; Hacker, T. a.; Ge, Y. *Int. J. Mass Spectrom.* **2011**, *305* (2-3), 95–102.

(34)   Zubarev, R. *J. Am. Chem. Soc.* **1998**, *7863* (16), 3265–3266.

(35)   Zubarev, R. a. *Curr. Opin. Biotechnol.* **2004**, *15* (1), 12–16.

(36)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 9507–9515.

(37)   Li, H.; Zhao, Y.; Phillips, H. I. a; Qi, Y.; Lin, T.-Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83* (13), 5369–5376.

(38)   Amster, I. *J. Mass Spectrom.* **1996**, *31* (September), 1325–1337.

(39)   Marshall, A. *Mass Spectrom. Rev.* **1998**, *17*, 1–35.

(40)   Shi, S. D.; Hendrickson, C. L.; Marshall, a G. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (20), 11532–11537.

(41)   Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. *Anal. Chem.* **2008**, *80* (11), 3985–3990.

(42)   McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X.; Infusini, G.; Jiang, H.; Kong, X.; Begley, T. P. *FEBS J.* **2007**, *274* (24), 6256–6268.

(43)   Tolmachev, A. V; Robinson, E. W.; Wu, S.; Paša-Tolić, L.; Smith, R. D. *Int. J. Mass Spectrom.* **2009**, *281* (1-3), 32–38.

(44) Calderone, V.; Casini, A.; Mangani, S.; Messori, L.; Orioli, P. L. *Angew. Chem. Int. Ed. Engl.* **2006**, *45* (8), 1267–1269.

(45) Ivanov, a. I. *J. Biol. Chem.* **1998**, *273* (24), 14721–14730.

(46) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *J. Am. Chem. Soc.* **1985**, *107* (26), 7893–7897.

(47) Kaddis, C. S.; Lomeli, S. H.; Yin, S.; Berhane, B.; Apostol, M. I.; Kickhoefer, V. a; Rome, L. H.; Loo, J. a. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (7), 1206–1216.

(48) Koning, L. J. De; Nibbering, N. M. M.; Ordenb, S. L. Van; Laukienb, F. H. *Int. J. Mass Spectrom. Ion Process.* **1997**, *166*.

(49) Stoll, S.; Schweiger, A. *J. Magn. Reson.* **2006**, *178* (1), 42–55.

(50) B. Ohlsson J. Axelson, N. F. *Scand. J. Gastroenterol.* **1999**, *34* (12), 1224–1229.

(51) Singh, D.; Joshi, D. D.; Hameed, M.; Qian, J.; Gascón, P.; Maloof, P. B.; Mosenthal, A.; Rameshwar, P. *Proc. Natl. Acad. Sci.* **2000**, *97* (1), 388–393.

(52) Sleno, L.; Volmer, D. a. *J. Mass Spectrom.* **2004**, *39* (10), 1091–1112.

(53) Stadtman, E. R.; Van Remmen, H.; Richardson, A.; Wehr, N. B.; Levine, R. L. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1703* (2), 135–140.

(54) Skvortsov, a N.; Zavodnik, V. E.; Stash, a I.; Bel'skii, V. K.; Skvortsov, N. K. *Russ. J. Org. Chem.* **2003**, *39* (2), 170–175.

(55) Plowman, J. E.; Deb-Choudhury, S.; Grosvenor, A. J.; Dyer, J. M. *Photochem. Photobiol. Sci.* **2013**, *12* (11), 1960–1967.

(56) Finley, E. L.; Dillon, J.; Crouch, R. K.; Schey, K. L. *Protein Sci.* **1998**, *7* (11), 2391–2397.

(57)   Taylor, S. W.; Fahy, E.; Murray, J.; Capaldi, R. a; Ghosh, S. S. *J. Biol. Chem.* **2003**, *278* (22), 19587–19590.

(58)   Lovejoy, K. S.; Lippard, S. J. *Dalton Trans.* **2009**, No. 48, 10651–10659.

(59)   Circu, M. L.; Aw, T. Y. *Free Radic. Biol. Med.* **2010**, *48* (6), 749–762.

(60)   Sirichanchuen, B.; Pengsuparp, T.; Chanvorachote, P. *Mol. Cell. Biochem.* **2012**, *364* (1-2), 11–18.

(61)   Shannon, A. M.; Bouchier-Hayes, D. J.; Condron, C. M.; Toomey, D. *Cancer Treat. Rev.* **2003**, *29* (4), 297–307.

(62)   Soldevila-Barreda, J. J.; Romero-Canelón, I.; Habtemariam, A.; Sadler, P. J. *Nat. Commun.* **2015**, *6*, 6582.

(63)   Schwarcz, R. *Curr. Opin. Pharmacol.* **2004**, *4* (1), 12–17.

(64)   Gibson, D.; Costello, C. E. *Eur. J. Mass Spectrom.* **1999**, *510*, 501–510.

(65)   Zhao, T.; King, F. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 1141–1147.

(66)   Hartinger, C. G.; Ang, W. H.; Casini, A.; Messori, L.; Keppler, B. K.; Dyson, P. J. *J. Anal. At. Spectrom.* **2007**, *22* (8), 960.

(67)   Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125* (29), 8949–8958.

(68)   Syrstad, E. a; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (2), 208–224.

(69)   Turecek, F. *J. Am. Chem. Soc.* **2003**, *125* (19), 5954–5963.

(70)   Ganisl, B.; Breuker, K. *ChemistryOpen* **2012**, *1* (6), 260–268.

(71)   Li, H.; Zhao, Y.; Phillips, H. I. A.; Qi, Y.; Lin, T.; Sadler, P. J.; Connor, P. B. O. **2011**,

5369–5376.

(72) Zhang, N.; Du, Y.; Cui, M.; Xing, J.; Liu, Z.; Liu, S. *Anal. Chem.* **2012**, *84* (14), 6206–6212.

(73) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, Ü. a; Tsybin, Y. O. *Chem. Soc. Rev.* **2013**, *42* (12), 5014–5030.

(74) Pan, J.; Han, J.; Borchers, C. H. *Int. J. Mass Spectrom.* **2012**, *325-327*, 130–138.

(75) Cooper, H. J.; Hudgins, R. R.; Håkansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (3), 241–249.

(76) Li, H.; Snelling, J. R.; Barrow, M. P.; Scrivens, J. H.; Sadler, P. J.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1217–1227.

(77) Li, H.; Wells, S. a; Jimenez-Roldan, J. E.; Römer, R. a; Zhao, Y.; Sadler, P. J.; O'Connor, P. B. *Protein Sci.* **2012**, *21* (9), 1269–1279.

(78) Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. *Org. Lett.* **1999**, *1* (6), 953–956.

(79) Black, M.; Mais, R. H. B.; Owston, P. G. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **1969**, *25* (9), 1753–1759.

(80) Cooper, H. J. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 1932–1940.

(81) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(82) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *J. Am. Chem. Soc.* **2008**, *130* (52), 17774–17789.

(83) Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128* (32), 10364–10365.

(84)   Ronconi, L.; Sadler, P. J. *Dalt. Trans.* **2011**, *40*, 262–268.

(85)   Clément, J.-L.; Gilbert, B. C.; Rockenbauer, A.; Tordo, P. *J. Chem. Soc. Perkin Trans. 2* **2001**, No. 9, 1463–1470.

(86)   Shih, C.; Museth, A. K.; Abrahamsson, M.; Blanco-Rodriguez, A. M.; Di Bilio, A. J.; Sudhamsu, J.; Crane, B. R.; Ronayne, K. L.; Towrie, M.; Vlcek, A.; Richards, J. H.; Winkler, J. R.; Gray, H. B. *Science* **2008**, *320* (5884), 1760–1762.
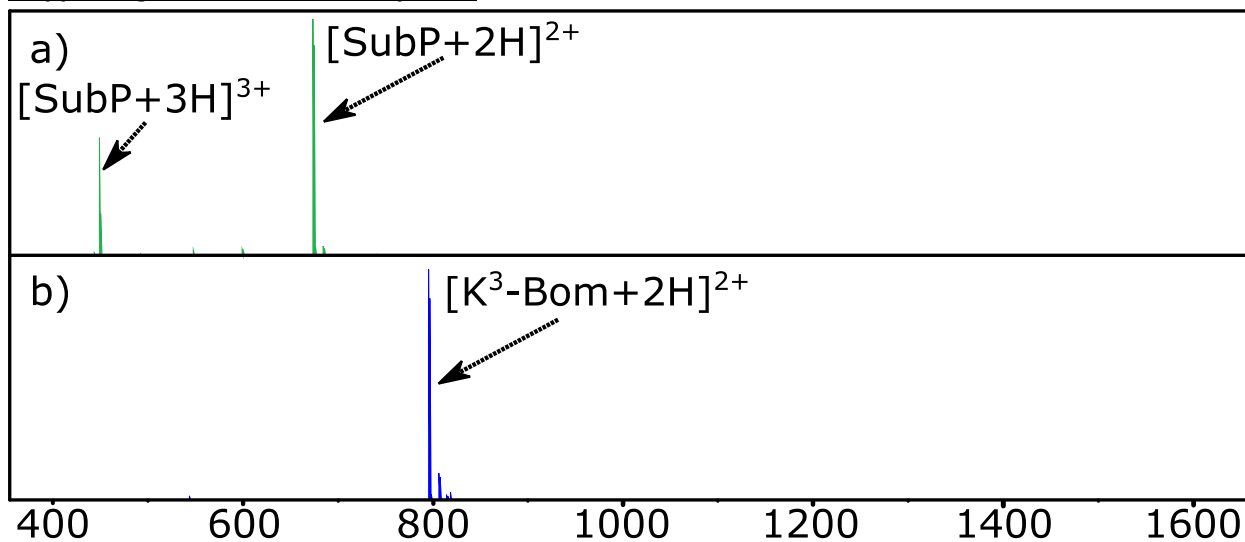
SI Figure 3.1 (SF3.1) – nESI-FT-ICR MS spectra of Substance P (a) and K^3-Bom (b)
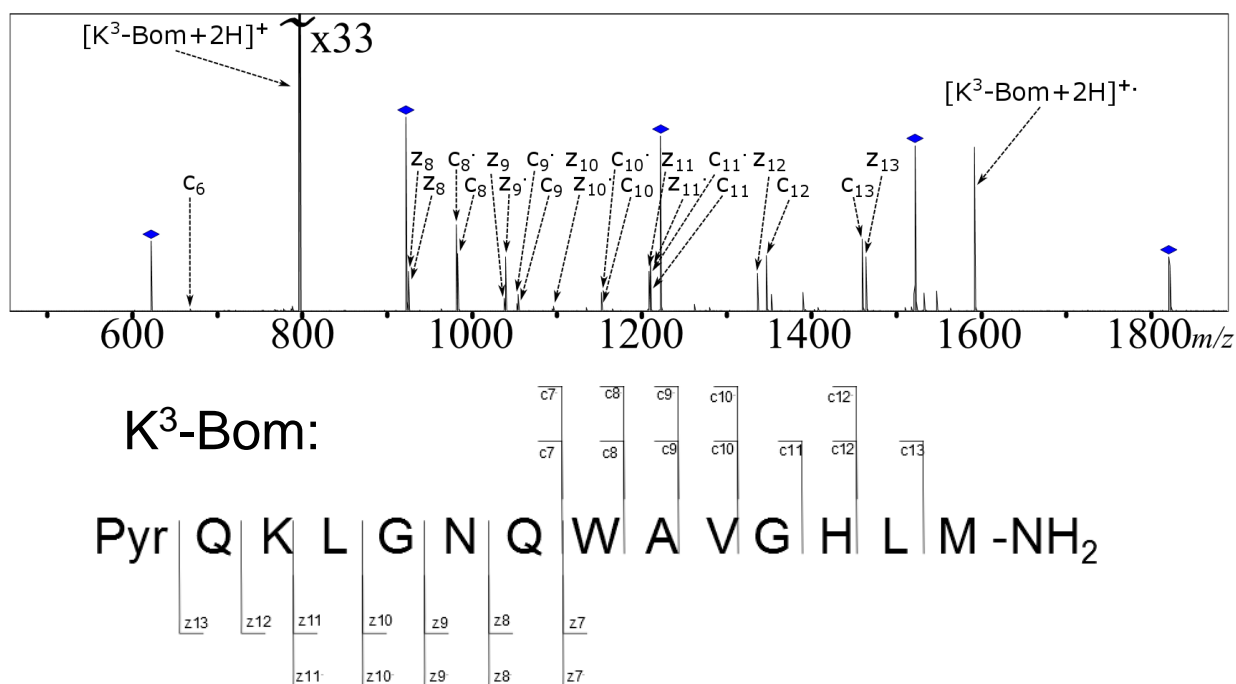
SI Figure 3.2 (SF3.2) – ECD MS/MS spectrum of K^3-Bom, peptide was co-isolated with 5 calibrant ions in order to improve calibration and compare the use of peptide MS/MS fragments and unrelated ions when internally calibrating MS/MS spectra.
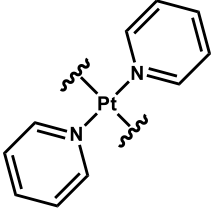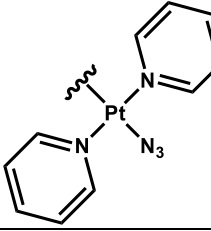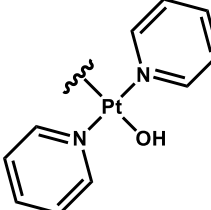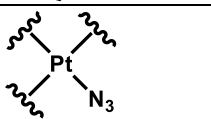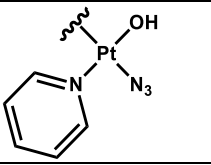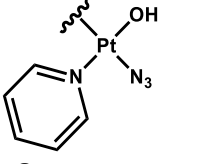
SI Table 1 (ST3.1) – Resulting assignments and associated Mass errors (in parts per million – ppm) for each calibration method used during the analysis of the K^3-Bom ECD MS/MS spectrum shown above.

Using co-isolated HP mix calibrant ions:				Using K^3-Bom fragment ions:			
Assign ment	Theoretical m/z	Observed m/z	Error/ ppm	Assign ment	Theoretical m/z	Observed m/z	Error/ ppm
c6	669.367290	669.36780	0.76	c6	669.367290	669.36729	0.00
z7	796.404850	796.40441	-0.55	z7	796.404850	796.40386	-1.24
c7	797.425870	797.42642	0.69	c7	797.425870	797.42587	0.00
z8	924.463430	924.46324	-0.21	z8	924.463430	924.46270	-0.79
z8·	925.471255	925.47090	-0.38	z8·	925.471255	925.47035	-0.98
c8·	982.497355	982.49757	0.22	c8·	982.497355	982.49703	-0.33
c8	983.505180	983.50543	0.25	c8	983.505180	983.50489	-0.29
z9	1038.506360	1038.50604	-0.31	z9	1038.506360	1038.50552	-0.81
z9·	1039.514185	1039.51385	-0.32	z9·	1039.514185	1039.51333	-0.82
c9·	1053.534465	1053.53491	0.42	c9·	1053.534465	1053.53439	-0.07
c9	1054.542290	1054.54262	0.31	c9	1054.542290	1054.54211	-0.17
z10	1095.527820	1095.52701	-0.74	z10	1095.527820	1095.52652	-1.19
z10·	1096.535645	1096.53490	-0.68	z10.	1096.535645	1096.53440	-1.14
c10·	1152.602875	1152.60309	0.19	c10.	1152.602875	1152.60262	-0.22
c10	1153.610700	1153.61108	0.33	c10	1153.610700	1153.61061	-0.08
z11	1208.611880	1208.61158	-0.25	z11	1208.611880	1208.61114	-0.61
z11·	1209.619705	1209.61878	-0.76	z11·	1209.619705	1209.61834	-1.13
c11·	1209.624335	1209.62488	0.45	c11·	1209.624335	1209.62444	0.09
c11	1210.632160	1210.63259	0.36	c11	1210.632160	1210.63215	-0.01
z12	1336.706840	1336.70647	-0.28	z12	1336.706840	1336.70612	-0.54
c12	1347.691070	1347.69166	0.44	c12	1347.691070	1347.69133	0.19
c13	1460.775130	1460.77538	0.17	c13	1460.775130	1460.77515	0.01
z13	1464.765420	1464.76540	-0.01	z13	1464.765420	1464.76520	-0.15
CRS-H	1591.815620	1591.81624	0.39	CRS-H	1591.815620	1591.81617	0.35
	Average error		0.40		Average error		0.47
	Standard deviation		0.44		Standard deviation		0.51

Fragment	Observed mass/ Da	Exact mass / Da	Error (ppm)
c2	271.18770	271.187699	0.004
c4	496.33543	496.335425	0.010
c5	624.39409	624.394003	0.139
c6	752.45256	752.452580	-0.027
a6	708.43884	708.438940	-0.141
a7	855.50728	855.507350	-0.082
c7	899.52095	899.520994	-0.049
c8	1046.58930	1046.589408	-0.103
a9	1059.59724	1059.597220	0.019
z9·	1078.56298	1078.562810	0.158
c9	1103.61086	1103.610872	-0.011
c10	1216.69476	1216.694936	-0.145
[M+H]⁺·	1347.73546	1347.735421	0.029

SI Figure 3.3 (SF3.3) ECD MS/MS of unmodified Substance P (top) followed by list of assignments (middle) and corresponding fragmentation map (bottom)

211

SI table 3.2 (ST3.2a): Species observed within the nESI-FT-ICR Mass Spectra of the Complex 1+SubP solution post-photoactivation.

Species	Elemental composition	Modification to peptide observed
[SubP+2H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SH_2$	N/A
[SubP+H]$^+$	$C_{63}H_{98}N_{18}O_{13}SH$	N/A
[SubP+O+2H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SH_2O$	+O
[SubP+{Pt(py)$_2$}+H]$^{3+}$	$C_{63}H_{98}N_{18}O_{13}SPtC_{10}H_{10}N_2H$	
[SubP+{Pt(py)$_2$(N$_3$)}+H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SPtC_{10}H_{10}N_5H$	
[SubP+{Pt(py)$_2$(OH)}+H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SPtC_{10}H_{10}N_2OHH$	
[SubP+{Pt(N$_3$)}+H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SPtN_3H$	
[SubP+{Pt(py)(OH)(N$_3$)}+2H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SPtC_5H_5NOHN_3H_2$	
[SubP+{Pt(py)(OH)(N$_3$)}+O+2H]$^{2+}$	$C_{63}H_{98}N_{18}O_{13}SPtC_5H_5NOHN_3H_2O$	+O

(ST3.2b): Species observed within the nESI-FT-ICR Mass Spectra of the Complex 1+K^3-Bom solution post-photoactivation.

Species	Elemental composition	Modification to peptide observed
[K^3-Bom+H]$^+$	$C_{71}H_{110}N_{22}O_{18}SH$	N/A
[K^3-Bom+2H]$^{2+}$	$C_{71}H_{110}N_{22}O_{18}SH_2$	N/A
[K^3-Bom+O+2H]$^{2+}$	$C_{71}H_{110}N_{22}O_{18}SH_2O$	+O

[K³-Bom+2O+2H]²⁺	$C_{71}H_{110}N_{22}O_{18}SH_2O_2$	+2O
[K³-Bom+3O+2H]²⁺	$C_{71}H_{110}N_{22}O_{18}SH_2O_3$	+3O
[K³-Bom+{Pt(py)₂}]²⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2$	
[K³-Bom+{Pt(py)₂}+H]³⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2H$	
[K³-Bom+{Pt(py)₂(OH)}+H]²⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2OHH$	
[K³-Bom+{Pt(py)₂(N3)}+H]²⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2N_3H$	
[K³-Bom+{Pt(py)₂(N3)}+O+H]²⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2N_3OH$	+O
[K³-Bom+{Pt(py)₂(N3)}+2O+H]²⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2N_3O_2H$	+2O
[K³-Bom+{Pt(py)₂(N3)}+3O+H]²⁺	$C_{71}H_{110}N_{22}O_{18}SHPtC_{10}H_{10}N_2N_3O_3H$	+3O
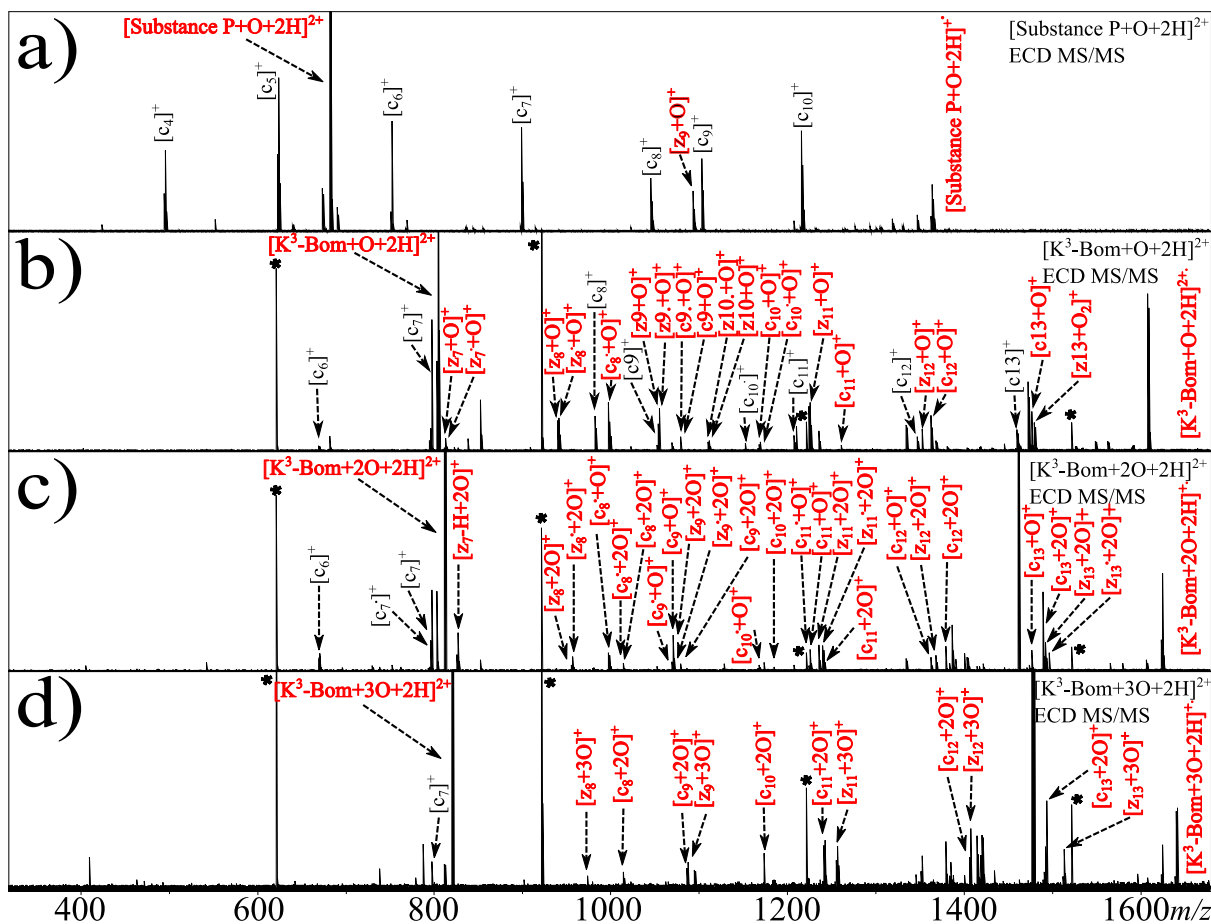
SI Figure 3.4 (SF3.4): ECD MS/MS spectra of Oxidised peptides individually isolated and dissociated using ECD within the FT-ICR MS; a) Substance P+O b) K^3-Bom+O C) K^3-Bom+2O d) K^3-Bom+3O. Red labels indicate modified fragments. Assignment table for each species can be found below.

SI Tables 3.3-6: assignment lists for individual ECD MS/MS spectra for the isolated oxidised peptide species; ST3) Substance P+O, ST4) K^3-Bom+O, ST5) K^3-Bom+2O, ST6) K^3-Bom+3O. Marked species were used for internal calibration, in tables with no marked species; co-isolated calibrant ions were used for calibration instead of MS/MS fragments, as discussed in the experimental section.

SI Table 3.3 (ST3.3): ECD MS/MS assignments for the SubP+O species:

Fragment	Exact mass	Observed mass	Error/ppm
[c2]+	271.187690	271.18770	0.04
[c4]+	496.335410	496.33543	0.04
[c5]+	624.393990	624.39400	0.02
[c6]+	752.452570	752.45257	0.00
[c7]+	899.520980	899.52094	-0.04
[c8]+	1046.589390	1046.58935	-0.04
[c9]+	1103.610850	1103.61064	-0.19
[c10]+	1216.694910	1216.69471	-0.16
[SubP+O+2H]+·	1364.738145	1364.73817	0.02
[z9]+	1094.557725	1094.55765	-0.07
		Absolute average	0.06
		Standard deviation	0.08

SI Table 3.4 (ST3.4): ECD MS/MS assignments for the K^3-Bom+O species:

Fragment	Exact mass	Measured mass	Error/ppm
[c6]+	669.367290	669.36794	0.97
[c7]+	797.425870	797.42638	0.64
[c9]+	1070.537205	1070.53756	0.33
[c10]+	1169.605615	1169.60610	0.41
[c11]+	1226.627075	1226.62770	0.51
[c12]+	1363.685985	1363.68616	0.13
[c13]+	1476.770045	1476.77013	0.06
[K^3-Bom+O+2H-H]+·	1607.810535	1607.80981	-0.45
[c8·]+	998.492270	998.49250	0.23
[c9·]+	1069.529380	1069.52972	0.32
[c10·]+	1168.597790	1168.59847	0.58
[z7]+	812.399765	812.40030	0.66
[z8]+	940.458345	940.45818	-0.18
[z9]+	1054.501275	1054.50101	-0.25
[z10]+	1111.522735	1111.52226	-0.43
[z11]+	1224.606795	1224.60649	-0.25
[z12]+	1352.701755	1352.70124	-0.38
[z13]+	1480.760335	1480.75951	-0.56
[z7·]+	813.407590	813.40772	0.16
[z8·]+	941.466170	941.46654	0.39
[z9·]+	1055.509100	1055.50884	-0.25
[z10·]+	1112.530560	1112.53004	-0.47
		Absolute average	0.39
		standard deviation	0.43

SI Table 3.5 (ST3.5): ECD MS/MS assignments for the K^3-Bom+2O species:

Fragment	Exact mass	Observed Mass	Error/ppm
[c6]$^+$	669.36729	669.36795	0.99
[c7.]$^+$	796.41805	796.41856	0.65
[c7]$^+$	797.42587	797.42640	0.66
[z7-H+2O]$^+$	827.38740	827.38688	-0.63
[z8+2O]$^+$	956.45326	956.45340	0.15
[z8.+2O]$^+$	957.46109	957.46118	0.10
[c8.+O]$^+$	998.49227	998.49253	0.26
[c8.+2O]$^+$	1014.48719	1014.48791	0.71
[c8+2O]$^+$	1015.49501	1015.49604	1.01
[c9.+O]$^+$	1069.52938	1069.53030	0.86
[c9+O]$^+$	1070.53721	1070.53746	0.24
[z9+2O]$^+$	1070.49619	1070.49576	-0.40
[z9.+2O]$^+$	1071.50402	1071.50399	-0.02
[c9+2O]$^+$	1086.53212	1086.53442	2.12
[c10.+O]$^+$	1168.59779	1168.59867	0.75
[c10+2O]$^+$	1185.60053	1185.60047	-0.05
[c11.+O]$^+$	1225.61925	1225.61937	0.10
[c11+O]$^+$	1226.62708	1226.62731	0.19
[z11+2O]$^+$	1240.60171	1240.60117	-0.44
[z11.+2O]$^+$	1241.60954	1241.60878	-0.61
[c11+2O]$^+$	1242.62199	1242.62276	0.62
[c12+O]$^+$	1363.68599	1363.68585	-0.10
[z12+2O]$^+$	1368.69667	1368.69605	-0.45
[c12+2O]$^+$	1379.68090	1379.68053	-0.27
[c13+O]$^+$	1476.77005	1476.77044	0.27
[c13+2O]$^+$	1492.76496	1492.76449	-0.31
[z13+2O]$^+$	1496.75525	1496.75359	-1.11
[z13.+2O]$^+$	1497.76308	1497.76135	-1.15
		Absolute average	0.54
		Standard deviation	0.69

SI Table 3.6 (ST3.6): ECD MS/MS assignments for the K^3-Bom+3O species:

Fragment	Exact Mass	Observed Mass	Error/ppm
[c7]$^+$	797.42587	797.42619	0.40
[c9+2O]$^+$	1086.53212	1086.53357	1.33
[c11+2O]$^+$	1242.62199	1242.62120	-0.64
[c12+2O]$^+$	1379.68090	1379.68204	0.83
[c13+2O]$^+$	1492.76496	1492.76538	0.28
[K^3-Bom+3O+2H-H]$^{+.}$	1639.80037	1639.80098	0.38
[c8.+2O]$^+$	1014.48719	1014.48766	0.47

[c10·+2O]⁺ → $[c_{10}·+2O]^+$	1184.59271	1184.59291	0.17
$[c_{13}·+2O]^+$	1491.75714	1491.75668	-0.31
$[z_{11}+3O]^+$	1256.59663	1256.59580	-0.66
$[z_{12}+3O]^+$	1384.69159	1384.69032	-0.91
$[z_{13}+3O]^+$	1512.75017	1512.74992	-0.16
$[z_8·+3O]^+$	973.45600	973.45558	-0.43
$[z_9·+3O]^+$	1087.49893	1087.49961	0.63
		Absolute average	0.54
		Standard deviation	0.62
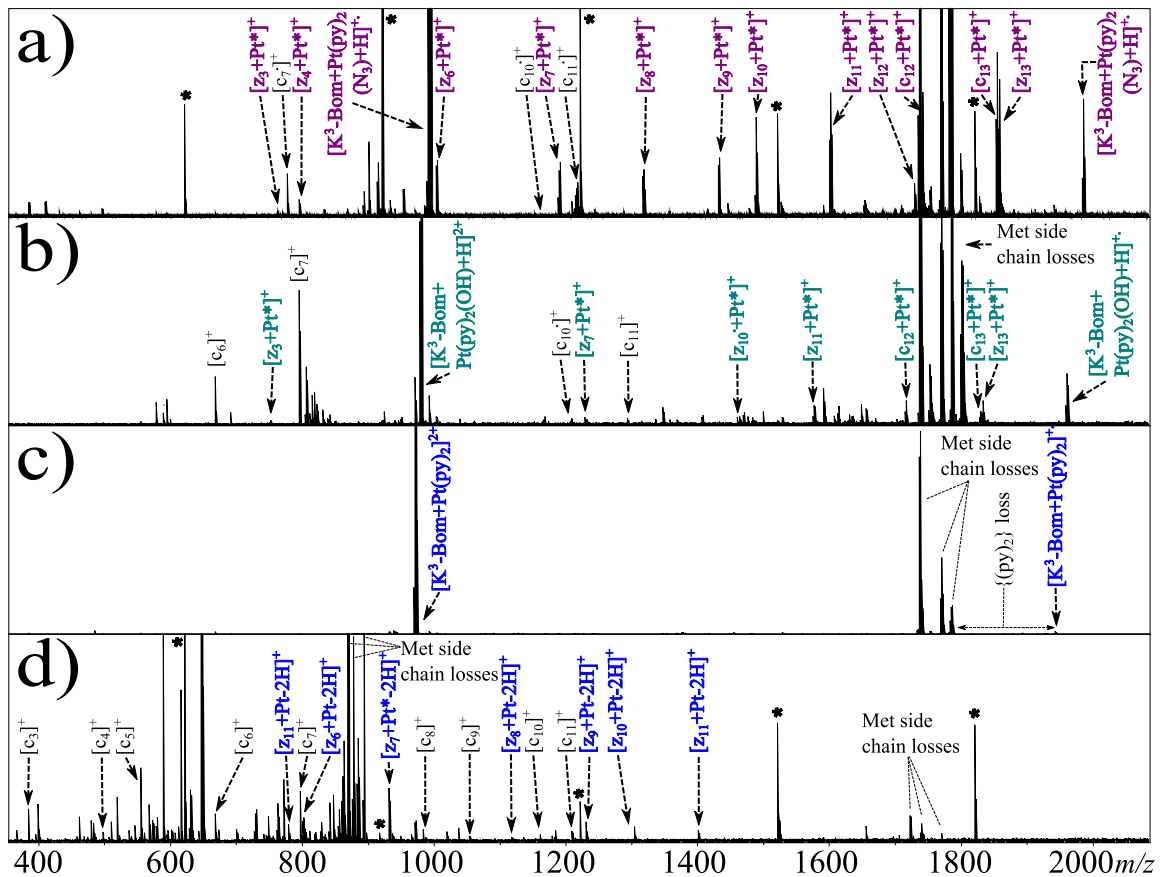
SI Figure 3.5 (SF3.5): ECD MS/MS spectra of Platinated K^3-bom species individually isolated and dissociated using ECD within the FT-ICR MS. Coloured labels indicate modified fragments. Assignment table for each species can be found below. Pt* indicates the platinum based modification associated with each species (shown in Figure 3, main text, and correlating tables of assignments, below).

SI Table 3.7 (ST3.7): ECD MS/MS assignments for the $[K^3\text{-Bom}+Pt(py)_2(N_3)+H]^{2+}$ species (calibrated using co-isolated calibrant ions):

Fragment	Exact mass	Observed mass	Error/ppm
$[c7\cdot]^+$	796.418045	796.41805	0.01
$[c10\cdot]^+$	1152.602875	1152.60367	0.69
$[z7+ Pt(py)_2(N_3)]^+$	1189.453306	1189.45339	0.07
$[c11\cdot]^+$	1209.624335	1209.62400	-0.28
$[z8+Pt(py)_2(N_3)]^+$	1317.511886	1317.51124	-0.49
$[z9+Pt(py)_2(N_3)]^+$	1431.554816	1431.55358	-0.86
$[z10+Pt(py)_2(N_3)]^+$	1488.576276	1488.57688	0.41
$[z11+Pt(py)_2(N_3)]^+$	1601.660336	1601.66124	0.56
$[z12+Pt(py)_2(N_3)]^+$	1729.755296	1729.75603	0.42
$[c13+Pt(py)_2(N_3)]^+$	1853.823586	1853.82508	0.81
$[z13+Pt(py)_2(N_3)]^+$	1857.813876	1857.81385	-0.01
		Absolute average	-0.18
		standard deviation	0.50

SI Table 3.8 (ST3.8): ECD MS/MS assignments for the $[K^3\text{-Bom}+Pt(py)_2(OH)+H]^{2+}$ species (calibrated using co-isolated calibrant ions):

Fragment	Exact mass	Observed Mass	Error/ppm
$[c6]^+$	669.367290	669.36729	0.00
$[z3\cdot+Pt(py)_2(OH)]^+$	752.248358	752.24755	-1.07
$[c7]^+$	797.425870	797.42614	0.34
$[c10\cdot]^+$	1152.602875	1152.60293	0.05
$[z7\cdot+Pt(py)_2(OH)]^+$	1165.454648	1165.45505	0.34
$[c11]^+$	1210.632160	1210.63215	-0.01
$[z10\cdot+Pt(py)_2(OH)]^+$	1464.577618	1464.57681	-0.55
$[z11+Pt(py)_2(OH)]^+$	1576.653853	1576.65538	0.97
$[c12+Pt(py)_2(OH)]^+$	1715.733043	1715.73301	-0.02
$[c13+Pt(py)_2(OH)]^+$	1828.817103	1828.81907	1.08
$[z13+Pt(py)_2(OH)]^+$	1832.807393	1832.80709	-0.17
$[K^3\text{-Bom}+Pt(py)_2(OH)+H]^+$	1959.857593	1959.85800	0.21
		Absolute average	0.40
		Standard deviation	0.56

SI Table 3.9 (ST3.9): ECD MS/MS assignments for the $[K^3\text{-Bom}+Pt(py)_2+H]^{3+}$ species:

Fragment	Exact mass	Observed Mass	Error/ppm
$[c3]^+$	385.21880	385.21880	0.00
$[c4]^+$	498.30290	498.30280	-0.26
$[c5]^+$	555.32440	555.32430	-0.05
$[c6]^+$	669.36730	669.36740	0.12
$[c7]^+$	797.42590	797.42600	0.16
$[c8]^+$	983.50520	983.50480	-0.43
$[c9]^+$	1054.54200	1054.54200	-0.36
$[c10]^+$	1153.61100	1153.61100	0.33
$[c11]^+$	1210.63200	1210.63200	0.17
$[z6\text{-H-H}+Pt]^+$	802.27310	802.27200	-1.40
$[z7\text{-H-H}+Pt]^+$	988.35240	988.35220	-0.21
$[z8\text{-H-H}+Pt]^+$	1116.4110	1116.41100	-0.29
$[z9\text{-H-H}+Pt]^+$	1230.4540	1230.45300	-0.43
$[z10\text{-H-H}+Pt]^+$	1287.4750	1287.47500	-0.63
$[z11\text{-H-H}+Pt]^+$	1400.55900	1400.55900	-0.65
$[z11\text{-2H}++Pt(py)_2]^{2+}$	779.82560	779.82470	-1.12
		Absolute average:	0.41
		Standard deviation:	0.46
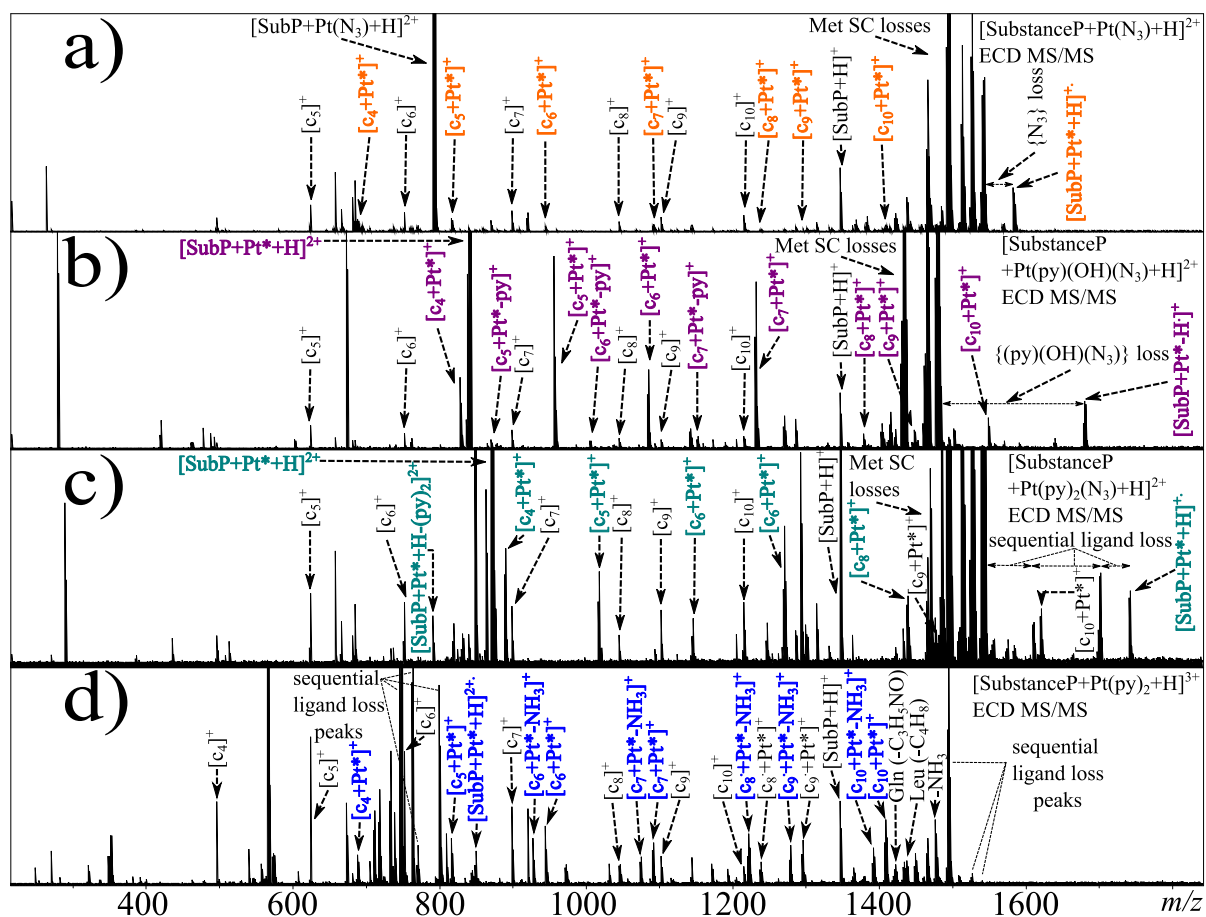
SI Figure 3.6 (SF3.6): ECD MS/MS spectra of Platinated Substance P species individually isolated and dissociated using ECD within the FT-ICR MS. Coloured labels indicate modified fragments. Assignment table for each species can be found below. Pt* indicates the platinum based modification associated with each species (shown in Figure 4, main text, and correlating tables of assignments, below).

SI Table 3.10 (ST3.10): ECD MS/MS assignments for the $[SubP+Pt(py)_2+H]^{3+}$ species:

Fragment	Exact mass	Observed mass	Error/ppm
$[c4\cdot+Pt]^+$	688.282420	688.28242	0.00
$[c5\cdot+Pt]^+$	816.341000	816.34120	0.24
$[c6\cdot+Pt]^+$	944.399580	944.39917	-0.43
$[c7\cdot+Pt]^+$	1091.467990	1091.46803	0.04
$[c8\cdot+Pt]^+$	1238.536400	1238.53627	-0.10
$[c9\cdot+Pt]^+$	1295.557860	1295.55822	0.28
$[c10\cdot+Pt]^+$	1408.641920	1408.64168	-0.17
$[c5\cdot+Pt-NH_3]^+$	799.314455	799.31421	-0.31
$[c6\cdot+Pt-NH_3]^+$	927.373035	927.37283	-0.22
$[c7\cdot+Pt-NH_3]^+$	1074.441445	1074.44128	-0.15
$[c8\cdot+Pt-NH_3]^+$	1221.509855	1221.50967	-0.15
$[c9\cdot+Pt-NH_3]^+$	1278.531315	1278.53071	-0.47
$[c10\cdot+Pt-NH_3]^+$	1391.615375	1391.61515	-0.16
$[SubP+Pt]^+\cdot-H\cdot$	1539.681866	1539.68187	0.00

[SubP+Pt]$^{+\cdot}$-H$^{\cdot}$-Me	1524.658939	1524.65993	0.65
[SubP+Pt]$^{+\cdot}$-MeS	1493.694694	1493.69432	-0.25
[SubP+Pt]$^{+\cdot}$-MeS-NH$_3$	1476.668145	1476.66816	0.01
[SubP+Pt]$^{+\cdot}$-MeS-Leu C$_4$H$_8$	1437.632094	1437.63193	-0.11
[SubP+Pt]$^{+\cdot}$-MeS-Glu C$_3$H$_5$NO	1422.65758	1422.65785	0.19
		Absolute average	0.21
		Standard deviation	0.26

SI Table 3.11 (ST3.11): ECD MS/MS assignments for the [SubP+Pt(py)$_2$(N$_3$) +H]$^{2+}$ species:

Fragment	Exact mass	Observed mass	Error/ppm
[c4+Pt(py)$_2$(N$_3$)]$^+$	889.383866	889.38374	-0.14
[c5+Pt(py)$_2$(N$_3$)]$^+$	1017.442446	1017.44264	0.19
[c6+Pt(py)$_2$(N$_3$)]$^+$	1145.501026	1145.50134	0.27
[c7+Pt(py)$_2$(N$_3$)]$^+$	1292.569436	1292.57007	0.49
[z9-H$^{\cdot}$+Pt(py)$_2$(N$_3$)]$^+$	1470.603986	1470.60669	1.84
[c10+Pt(py)$_2$(N$_3$)]$^+$	1609.743366	1609.74471	0.83
[SubP+Pt(py)$_2$(N$_3$)+H]$^{+\cdot}$	1740.783856	1740.78386	0.00
[SubP+Pt]$^+$	1540.690813	1540.69154	0.47
[SubP+Pt-Me]$^+$	1525.667338	1525.66709	-0.16
[SubP+Pt-S]$^+$	1493.695268	1493.69489	-0.25
		Absolute average	0.47
		Standard deviation	0.59

SI Table 3.12 (ST3.12): ECD MS/MS assignments for the [SubP+Pt(N$_3$)+H]$^{2+}$ species:

Fragment	Exact mass	Observed mass	Error/ppm
[c4$^{\cdot}$+Pt]$^+$	688.28242	688.28242	0.00
[c5$^{\cdot}$+Pt]$^+$	816.34100	816.34078	-0.27
[c6dot+Pt]$^+$	944.39958	944.39958	0.00
[c7$^{\cdot}$+Pt]$^+$	1091.46799	1091.46827	0.26
[c8$^{\cdot}$+Pt]$^+$	1238.53640	1238.53767	1.03
[c9$^{\cdot}$+Pt]$^+$	1295.55786	1295.55875	0.69
[c10$^{\cdot}$+Pt]$^+$	1408.64192	1408.64181	-0.08
[SubP+Pt(N$_3$)+H]$^{+\cdot}$	1582.69837	1582.69848	0.07
		Absolute average	0.30
		Standard deviation	0.41
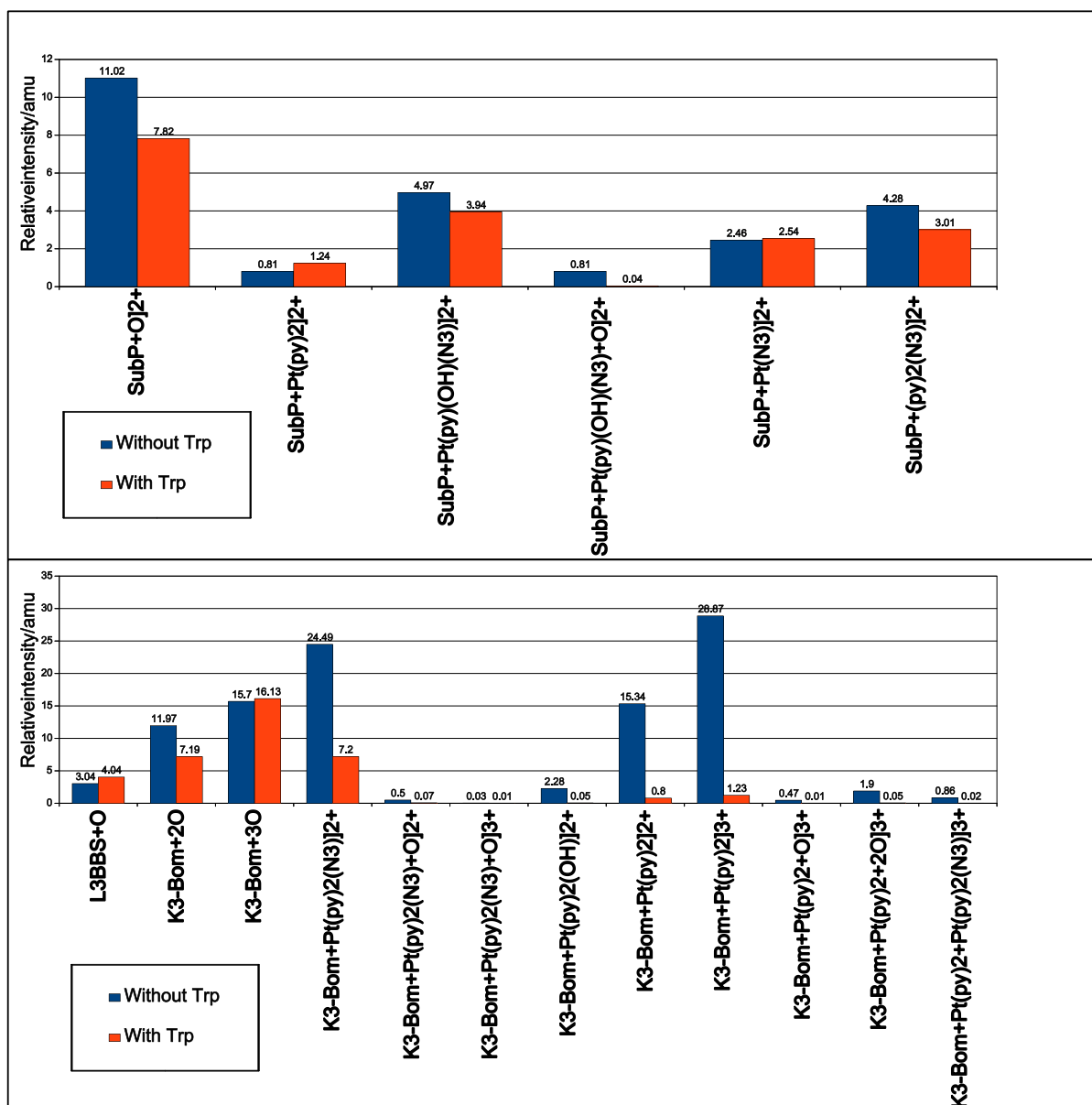
SI Figure 3.7 (SF3.7): Bar charts showing the relative intensity changes between the Complex 1+peptide and Complex 1+peptide+free Trp reaction mixtures for both Substance P (top) and (bottom).

SI Table 3.13a (ST3.13a): Relative intensity measurements of assigned species in the irradiated Complex1+Substance P reaction mixture FT-ICR MS with and without additional tryptophan.

Species	Relative Intensity/%	
	Without additional Trp	With additional Trp
$[SubP+2H+O]^{2+}$	11.02	7.82
$[SubP+Pt(py)_2]^{2+}$	0.81	1.24
$[SubP+Pt(py)(OH)(N_3)+H]^{2+}$	4.97	3.94
$[SubP+Pt(py)(OH)(N_3)+H+O]^{2+}$	0.81	0.04
$[SubP+Pt(N)_3+H]^{2+}$	2.46	2.54
$[SubP+Pt(py)_2(N)_3+H]^{2+}$	4.28	3.01

SI Table 3.13b (ST3.13b): Relative intensity measurements of assigned species in the irradiated Complex1+K^3-Bom reaction mixture FT-ICR MS with and without additional tryptophan.

Species	Relative Intensity/%	
	Without additional Trp	With additional Trp
$[K^3\text{-Bom}+2H+O]^{2+}$	3.04	4.04
$[K^3\text{-Bom}+2H+2O]^{2+}$	11.97	7.19
$[K^3\text{-Bom}+2H+3O]^{2+}$	15.7	16.13
$[K^3\text{-Bom}+Pt(py)_2(N_3)+H]^{2+}$	24.40	7.20
$[K^3\text{-Bom}+Pt(py)_2(N_3)+H+O]^{2+}$	0.50	0.07
$[K^3\text{-Bom}+Pt(py)_2(N_3)+2H+O]^{3+}$	0.03	0.01
$[K^3\text{-Bom}+Pt(py)_2(OH)+H]^{2+}$	2.28	0.05
$[K^3\text{-Bom}+Pt(py)_2]^{2+}$	15.34	0.80
$[K^3\text{-Bom}+Pt(py)_2+H]^{3+}$	28.87	1.23
$[K^3\text{-Bom}+Pt(py)_2+H+O]^{3+}$	0.47	0.01
$[K^3\text{-Bom}+Pt(py)_2+H+2O]^{3+}$	1.90	0.05
$[K^3\text{-Bom}+Pt(py)_2+Py(py)_2(N_3)]^{3+}$	0.86	0.02
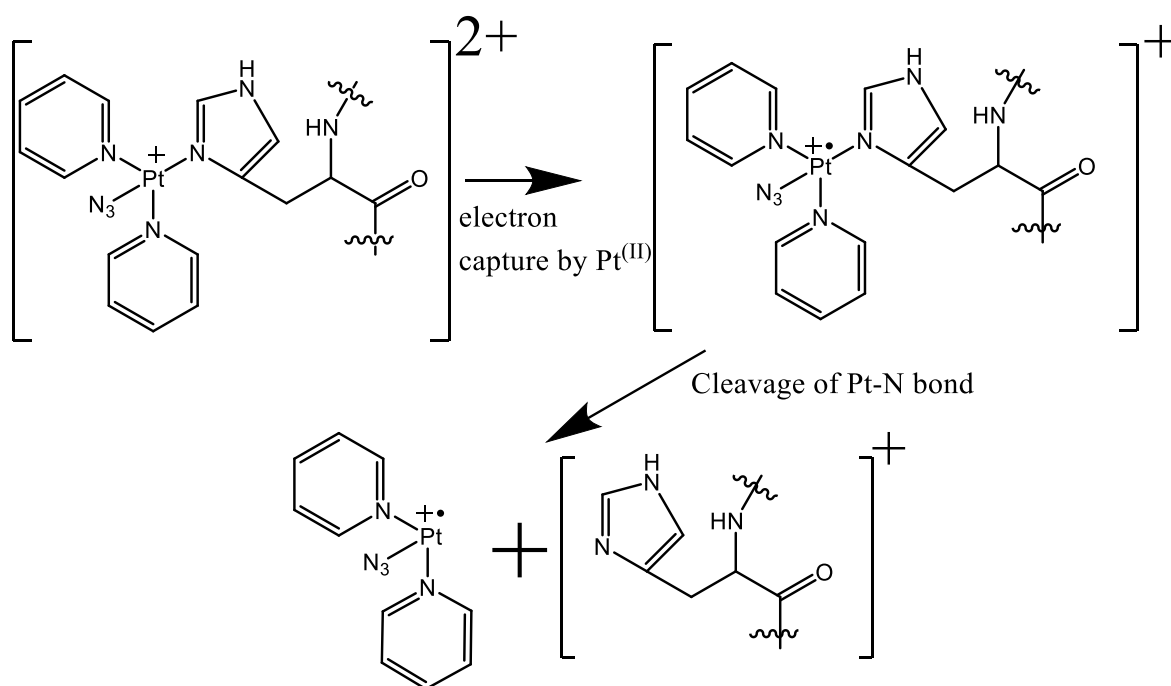
SI Figure 3.8 (SF3.8)- Proposed mechanism for the loss from the peptide of the Pt(II) modification during ECD, ligands shown vary with the nature of the Pt adduct.

SI Table 3.14a-c (ST3.14a-c): EPR results for irradiation of complex one in blue visible light (a), Complex 1+SubstanceP (b), and Complex 1+[Lys][3]-Bombesin (c). Included are previously published values acquired for the system, along with references for said values.

Complex 1+DEPMPO (irradiation blue light)

	g	a^N_{NO}	a^P	a^H_β	a^N_α
Experimental	2.012	13.89	46.16	13	2.79
Published[1]	-	13.93	46.05	12.39	2.80

Complex 1+Substance P+DEPMPO (irradiation blue light)

	g	a^N_{NO}	a^P	a^H_β	a^N_α
Experimental	2.012	13.89	46.16	13	2.79
Published[1]	-	13.93	46.05	12.39	2.80

Complex 1+[Lys][3]-Bombesin+DEPMPO (irradiation blue light)

	g	a^N_{NO}	a^P	a^H_β
Experimental	2.012	7.08	38.75	3.76
Published[2/3]	-	7.8/7.14	41.2/38.69	4.1/3.69

Complex 1
+DEPMPO 42 min

Complex 1+SubP
+DEPMPO 42 min

Complex 1
+DEPMPO 28 min

Complex 1+SubP
+DEPMPO 28 min

Complex 1
+DEPMPO 14 min

Complex 1+SubP
+DEPMPO 14 min

Field mT

Field mT

Complex $\mathbf{1}$+K^3-Bom
+DEPMPO 42 min

Complex $\mathbf{1}$+K^3-Bom
+DEPMPO 28 min

Complex $\mathbf{1}$+K^3-Bom
+DEPMPO 14 min
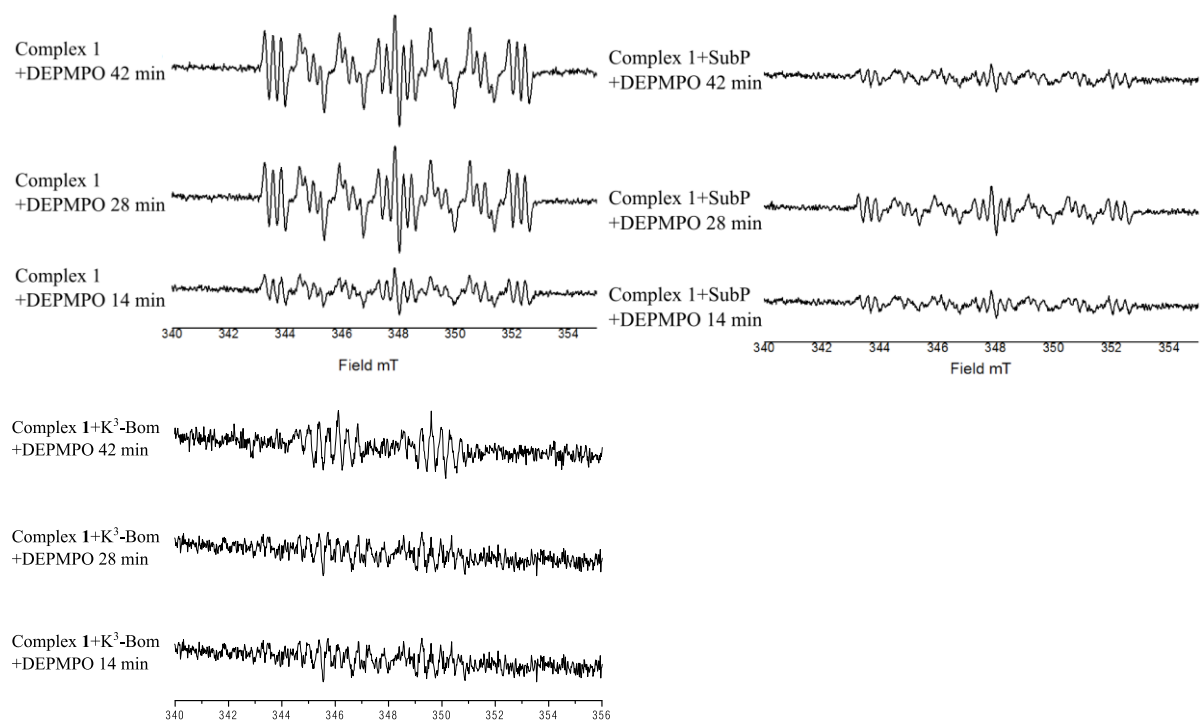
SI Figure 10a-c (SF3.10a-c): UV/Vis spectra of Complex 1 (a), Complex 1 + Substance P (b), and Complex 1 + [Lys]^3-Bombesin (c).
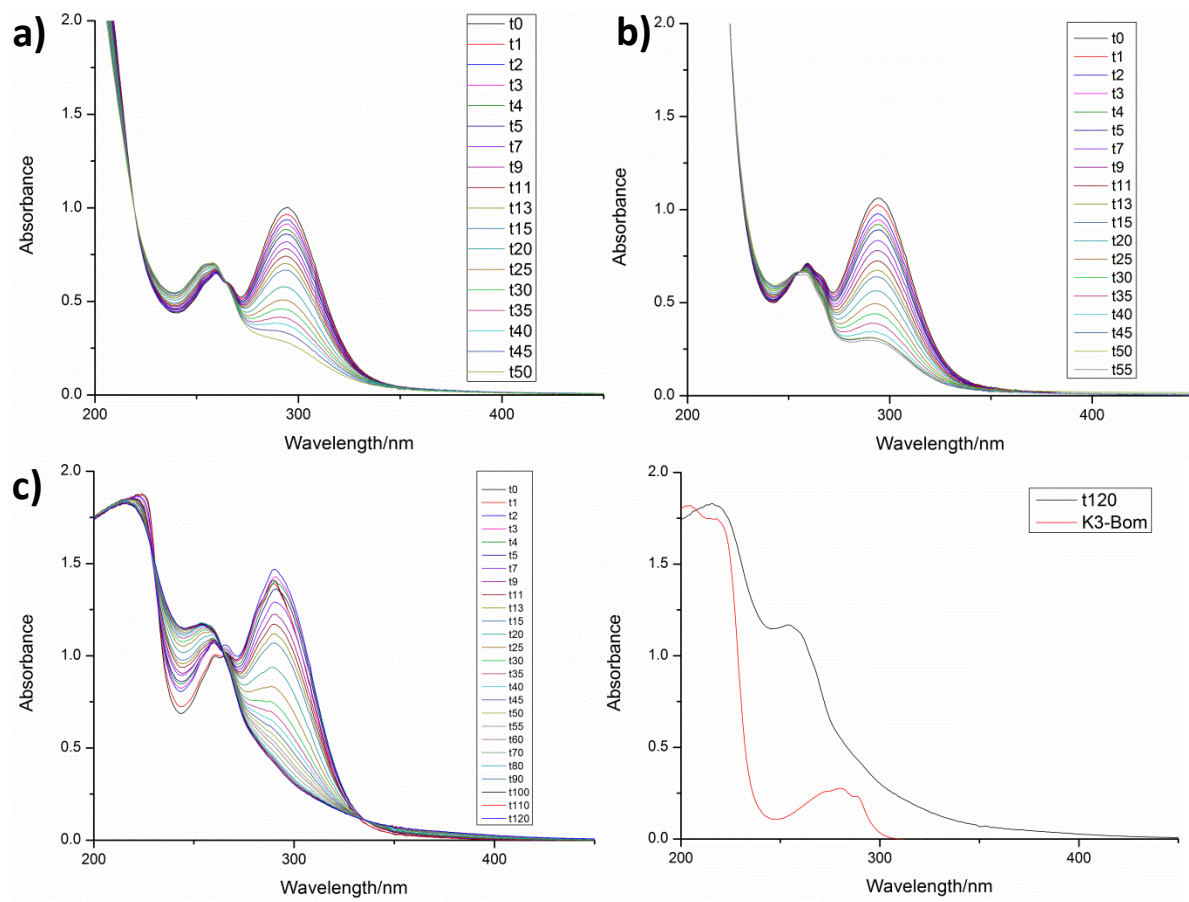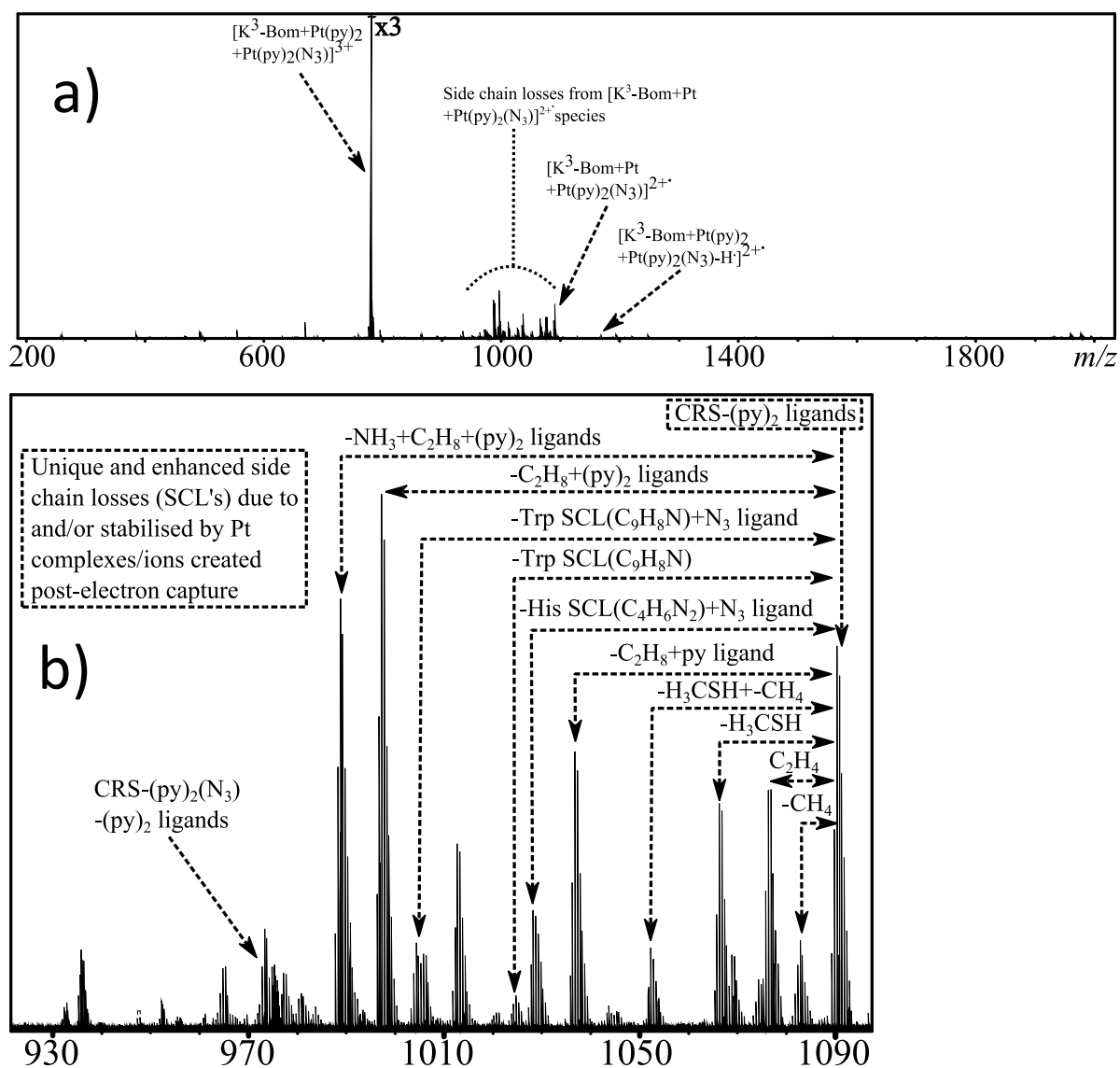
SI Figure 3.11 (SF3.11): ECD MS/MS spectra of di-Platinated K^3-bom species a) and zoom-in of side chain loss (SCL) region (b) resulting from electron capture at a platinum centre, causing ligand loss (to create $[K^3\text{-Bom+Pt+Pt(py)}_2(N_3)]^{2+\cdot}$ from the $[K^3\text{-Bom+Pt(py)}_2+\text{Pt(py)}_2(N_3)]^{2+\cdot}$ species) along with unique side chain losses. Unfortunately no sequence informative fragments were observed from the ECD MS/MS spectrum, highlighting the need for an additional proton to allow standard dissociation.

Chapter 4:

Osmium metallodrug-DNA interactions studied by FT-ICR MS and MS/MS

Transition metal complexes offer a range of geometries, reactivities, and activities against

many forms of cancer and have been proven to be effective at treating a wide array of conditions.

With the increased use of platinum anti-cancer agents a marked increase in platinum-resistant

cancer cases has developed, often due to non-ideal dosing regiments and treatments. After

cancerous cells become resistant to platinum-based therapies chances of survival diminish greatly.

For these and other reasons investigations into non-platinum metallodrugs is becoming a thriving

field, with many metals not just matching platinum drug effectiveness but often exceeding them.

Potent piano-stool compounds such as those based on ruthenium, rhodium, iridium, and even

osmium have been shown to be extremely effective, and selective against cancerous cells, and

importantly, platinum-resistant cancerous cells. Many of these heavy metal drugs are now in/being

considered for clinical trials and may soon become available for widespread use.

With new metals, geometries and functional groups come new mechanisms of action, and

these different mechanisms of action are believed to be the key not only to these metallodrugs

increased potencies, but also to their ability to circumvent metal-resistance in cancer cells and attack

cellular components in different ways. For this reason a promising osmium-based piano-stool

complex (referred to as Os1) was chosen to investigate the interaction of osmium compounds with

model DNA strands (namely 10-mer and 12-mer oligonucleotides) and use a variety of MS/MS

techniques to fragment the resulting reaction products effectively while retaining these unique

biomolecule modifications. The use of specialised electron-based dissociation Electron Detachment

Dissociation (EDD) was shown to be challenging to achieve, but produced extensive fragmentation

information, retained the osmium modifications effectively, and enabled effective characterisation

of multiple osmium-DNA interaction products. This was also compared to more commonly used

techniques such a CAD and IRMPD, which showed limited effectiveness, but still provided key information for some species.

The work presented in chapter is mostly based directly on the publication by C. A. Wootton *et al.* "Binding of Organo-Osmium(II) anticancer Compounds to Guanine and Cytosine on DNA Revealed by Electron-based Dissociations in High-Resolution FT-ICR Mass Spectrometry", Dalton Transactions, 44, 8, 2015, 3624-3632. DOI: 10.1039/C4DT03819C. The collaborators on the project offered the Osmium compounds, the purified oligonucleotides, and showed preliminary data for the reactions of the drug with the DNA. However all data presented herein (outside of the HPLC purification) was produced, analysed and written by the thesis Author. The extensive DNA purification was conducted by Carlos Sanchez-Cano.

Binding of an Organo-osmium(II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry

Christopher A. Wootton,[1] Carlos Sanchez-Cano,[1] Hong-Ke Liu,[1,2] Mark P. Barrow,[1] Peter J. Sadler,[1] * Peter B. O'Connor[1] *

1. Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom

2. Jiangsu Collaborative Innovation Center of Biomedical Functional Materials, School of Chemistry and Materials Science, Nanjing Normal University, Wenyuan Road 1, Nanjing 210023, PR China

p.oconnor@warwick.ac.uk; p.j.sadler@warwick.ac.uk

CORRESPONDING AUTHORS

Peter B. O'Connor, Peter J. Sadler, Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK

PBO'C: phone: +44 (0)24 76151008; fax: +44 (0)24 76151009;

email: p.oconnor@warwick.ac.uk;

PJS: phone: +44 (0)24 765 23818; fax: +44 (024) 765 23819;

Email: p.j.sadler@warwick.ac.uk

Abstract

The Os^{II} arene anticancer complex $[(\eta^6\text{-bip})Os(en)Cl]^+$ (**Os1-Cl**; where bip = biphenyl, and en = ethylenediamine) binds strongly to DNA. Here we investigate reactions between **Os1-Cl** and the self-complementary 12-mer oligonucleotide 5′-TAGTAATTACTA-3′ (DNA12) using ultra high resolution Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry(FT-ICR MS). Identification of the specific sites of DNA osmiation with $\{(\eta^6\text{-bip})Os(en)\}^{2+}$ was made possible by the use of Electron Detachment Dissociation (EDD) which produced a wide range of assignable osmiated MS/MS fragments. In contrast, the more commonly used CAD and IRMPD techniques produced fragments which lose the bound osmium. These studies reveal that not only is guanine G_3 a strong binding site for $\{(\eta^6\text{-bip})Os(en)\}^{2+}$ but, unexpectedly, so too is cytosine C_{10}. Interestingly, the G_3/C_{10} di-osmiated adduct of DNA12 also formed readily but did not undergo such facile fragmentation by EDD, perhaps due to folding induced by van der Waal's interactions of the bound osmium arene species. These new insights into osmium arene DNA adducts should prove valuable for the design of new organometallic drugs and contribute to understanding the lack of cross resistance of this organometallic anticancer complex with cisplatin.

Introduction

Metal based therapeutics are currently at the forefront of anticancer therapy, with compounds such as cisplatin being used extensively in chemotherapy treatments.[1] Potent cytotoxic and anti-proliferative compounds have also been developed based on a range of transition metals, including ruthenium[2] , iridium[3], vanadium[4], cobalt[5], osmium[6], as well as platinum[7–9]. Though many of these compounds have been shown to exhibit different and interesting mechanisms of action (MoA), many metal-based drugs are believed to achieve their cytotoxic/ anti-proliferative effect via binding to the strands of DNA within the nuclei of cancer cells.[10] This binding of metallodrugs to DNA residues often causes a change in the conformation/structure of the DNA strands, and the deformation of tertiary structure can then trigger the cytotoxic and/or anti-proliferative effects, e.g. via apoptosis.[10,11]

It is therefore of much interest to study how these new metallodrugs can bind to DNA, which nucleotide residues are preferred binding partners, whether certain nucleotide sequences are targeted selectively, and to characterise the coordination sphere of the bound metal. Many metallodrugs change their composition during their journey to DNA, often multiple times, depending on various environmental factors (e.g. pH changes, chloride concentration, small/large biomolecule binding) and so the metallodrug injected into a system may not be the active species eventually causing the desired anticancer effects.[12] Due to the multiple activation pathways in which injected pro-drugs become active metallodrugs, and many pathways of deactivation of these compounds, it has often been observed that a variety of reaction products are observed when metal based compounds encounter biomolecules.[13]

Many analytical techniques have been utilised in order to study metal complex-DNA interactions, often ^{1}H-NMR or x-ray crystallography, with both techniques providing powerful structural information and sometimes identifying the location of DNA modification.[14,15] However, the mixture of products produced pose problems for both NMR and x-ray crystallography, and although separation of reaction products, e.g. via High Pressure Liquid Chromatography (HPLC), can assist in the acquisition of purer samples for analysis, this invariably ends in further dilution of samples and can often affect equilibria in dynamic systems, which may render separations ineffective via decomposition, hydrolysis etc.

To address these issues, other analytical techniques are being explored which offer superior sensitivity (and reduced sample requirements) and can handle mixtures of adducts more effectively. Mass spectrometry is increasingly being utilised to study biological and complex chemical systems, due to its inherent high sensitivity and ability to cope with extremely complex samples (over 100,000 species have been detected in a single mass spectrum[16]). With the advent of tandem mass spectrometry (MS/MS), the ability to fragment and interrogate ionised species in the gas phase has paved the way for in-depth chemical and structural analysis of complex systems[17], with biomolecules being the main focus for contemporary tandem mass spectrometry studies. An extensive array of ion dissociation

techniques have been developed including Collisionally Activated Dissociation (CAD)[18], Electron Capture Dissociation (ECD),[19] Electron Transfer Dissociation (ETD),[20] Infra-Red Multi Photon Dissociation (IRMPD),[21] Electron Induced Dissociation (EID),[22] Ultra-Violet Photo Dissociation (UVPD),[23] and Electron Detachment Dissociation (EDD).[24] There are now many established methods for interrogating and studying biomolecules and their modifications via MS/MS analysis, each suited to a variety of different analytes, and many able to offer complementary data to provide increased confidence in analysis.

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) uses electric and magnetic fields to trap ions of interest within a Penning trap during MS/MS experiments and detection,[25,26] allowing ultra-high resolving powers of 500,000-10,000,000+ with the latest commercial setups.[27] Combining the flexibility of the ICR cell with quadrupole isolation and collisional activation, FT-ICR MS offers the largest array of fragmentation techniques and the highest mass accuracy of any mass spectrometer, providing the highest possible confidence in spectral assignment, with mass errors into the sub part per million (ppm) and even ppb (part per billion) ranges. Importantly, FT-ICR MS is uniquely and perfectly suited to Top Down fragmentation studies.[28]

Here we study the metallation of a DNA 12-mer oligonucleotide by the organometallic Os^{II} arene anticancer complex, **Os1-Cl**·PF_6 (Fig. 4.1). We use Ultra High Resolution Mass Spectrometry (UHR-MS) together with MS/MS analysis to locate the osmium binding sites and to identify the coordinated ligands. In particular we use a range of fragmentation techniques to gain new insights into the nature of the osmium binding sites on DNA.
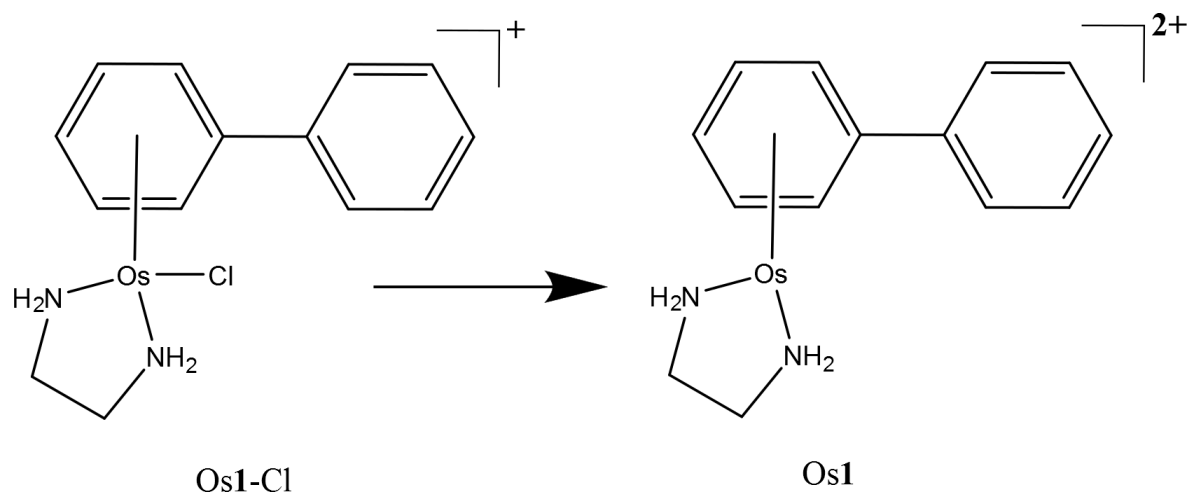
Scheme 4.1: The Os^{II} arene complex Os1-Cl studied in this work and the detected species/modification observed via ESI-MS (Os1).

Experimental

[(η^6-bip)Os(en)Cl]PF$_6$ (**Os1-Cl**·PF$_6$) was synthesised and characterised as described previously.[29]

The 12-mer DNA oligonucleotide **DNA12** (defined here for MS purposes as fully protonated, i.e. each of the 11 phosphates is protonated)

5′-TAGTAATTACTA-3′ (**DNA12**)

 DNA12 (T$_m$ 28°C) was purchased desalted from DNA Technology (Denmark). Further purification was performed by HPLC, using an Agilent 1200 series liquid chromatography system. The oligonucleotide was separated from shorter oligonucleotides (present as minor impurities in the commercial sample) using a PL-SAX ion exchange column (1000 Å pore size; 8 µM particle size; 150 mm length, 4.6 mm diameter; Polymer Laboratories, Amherst, USA); buffer A: 7% acetonitrile + 0.1 M triethylammonium acetate (TEAA), pH 7; buffer B: 7% acetonitrile + 0.1M TEAA + 1 M sodium acetate, pH 7; temp 40°C; flow rate 1.5 ml/min; gradient: 0-5 min 0% B, 5-45 min 0-40% B, 45-55 min 40-100% B, 55-65 min 100% B, 65-75 min 0% B (see Supporting Information Figure S1: HPLC traces of commercial sample and inset; the resulting purified fraction). The solvent was removed on a freeze-drier, and a

reverse phase chromatography PLRP-S column (100 Å pore size; 8 µM particle size; 300 mm length, 7.5 mm diameter; Polymer Laboratories, US) was used to remove sodium acetate; buffer A: 50 mM TEAA, pH 7; buffer B: 70% acetonitrile + 50 mM TEAA, pH 7; temp 40°C; flow rate 2 ml/min; gradient: 0-5 min 5% B, 5-35 min 5-100% B, 35-45 min 100% B, 45-45 min 0% B. The solvent was removed again on a freeze-drier.

Aqueous solutions of the oligonucleotide (250 µM) and osmium complex **Os1-Cl**·PF$_6$ (250 µM) were mixed in 0.5:1 and 1:1 Os:oligonucleotide ratios. Under these conditions of low ionic strength, DNA12 would be present largely as a single strand. Osmium-oligonucleotide samples were then incubated in the dark at 37°C for 3 hours before dilution with purified (Milli-Q) water to mass spec concentrations (ca. 0.5- 1 µM total concentration) for FT-ICR MS analysis.

FT-ICR mass spectrometry analysis. All samples were analysed via nano-electrospray ionisation (nESI) for increased sensitivity and lower sample consumption compared to traditional electrospray ionisation (ESI).[30] All experiments were carried out on a SolariX FT-ICR Mass Spectrometer, fitted with a 12 tesla actively shielded magnet (Bruker Daltonik GmbH, Bremen, Germany), and samples were sprayed in negative ion mode unless otherwise stated. For MS experiments, ions were accumulated for 0.01 s in the hexapole-based collision cell before transfer to the infinity cell[31] for detection.

For Collisionally Activated Dissociation (CAD) MS/MS experiments, ions of interest were isolated in the front end quadrupole using an isolation window of 3-10 *m/z* to isolate deprotonated species only. Ions were then accelerated into argon collision gas at 8 V (for unmodified DNA species) and 10-12 V (for Os-modified species). Ions were continuously accumulated for 0.1-2 s before transmission and detection.

For Infra-Red Multi Photon Dissociation (IRMPD) experiments, ions of interest were isolated, accumulated for 0.1-2 s, then transferred and trapped in the infinity cell. Trapped ions were then subjected to infra-red photons prior to detection. IR photons were produced

from a continuous wave, 25 W, CO_2 laser (Synrad Inc., Washington, USA) held at 60% power output and pulsed into the infinity cell for 30-100 ms.

For Electron Detachment Dissociation MS/MS experiments, ions of interest were isolated in the front end quadrupole, accumulated in the hexapole for 0.7-3 s, then transferred and trapped in the infinity cell. Trapped ions were then irradiated for 0.8 s with 20.2 eV electrons produced from a 1.5 A indirectly-heated hollow cathode dispenser via an extraction lens held at 18 V.

All spectra were internally calibrated using a quadratic calibration function[32] and then manually interpreted and assigned via Data Analysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany). All fragments in the low to sub-ppm range were assigned; tables of assignments can be found in the supporting information for each MS/MS spectrum assigned here.

Results and discussion

Electrospray ionisation (including nESI) forms ions via the addition or removal of protons during the desolvation process occurring after the emission of sample droplets from the electrospray needle tip and while passing through the electric fields in the source region of the mass spectrometer.[33] Due to the high density of phosphate groups in oligonucleotides, negative mode electrospray produces abundant $[M-nH]^{n-}$ ions in the resulting mass spectrum, with the phosphate groups stabilising the negative charges. This stabilisation is much greater than that seen in most peptide and protein mass spectra. As a result, DNA and RNA ESI/nESI spectra show charge states higher (per monomer) than commonly observed for poly(amino acid) species, even without the addition of additives such as bases to aid deprotonation.

Figure 1A shows the full mass spectrum of the 12-mer DNA produced by nESI. Isotopic simulations (Figure 4.1A inset) match the expected pattern for the 12-mer DNA (sequence inset) and show no modification of the DNA before reaction. CAD MS/MS of the 7- charge-state species is shown in Figure 4.2a; although very stable in solution, the high charge states (per nucleotide) of the DNA precursor make the gas phase species more fragile, causing it to

dissociate under lower than average collision energies. The CAD MS/MS spectrum showed many abundant sequence fragments, most notably **w** ions and **a**-base ions.[34] The nomenclature devised by McLuckey *et al.*[34] based on the McCloskey *et.al.* system[35] for assignment of oligonucleotide MS/MS is summarised in Figure 4.2a (inset, with assignments listed in Table S4.1 in SI) and was used for all assignments. Nucleotide base loss is a common fragmentation channel for CAD MS/MS, especially for **a** ions, which was also observed during these experiments. The overall sequence coverage was excellent, with 100% cleavage coverage for the unmodified species produced during one experimental run (Figure 4.2a). It is worth noting that, due to the fragility of DNA ions in the gas phase, careful tuning of transmission and extraction voltages was essential to preserve ions of interest as they move through the mass spectrometer to avoid unwanted fragmentation of precursor ions before and after MS and MS/MS experiments. Due to the large extent of fragmentation (even at low voltages) numerous internal fragments are often observed in DNA CAD MS/MS spectra.[36] These species are most often produced via fragmentation in two separate areas of the molecule, creating peaks not easily identified using a "linear" fragmentation model, though are still often predictable to some extent.[37]

Figure 4.1b shows the full mass spectrum of a 1:1 solution of **Os1-Cl**·PF_6:DNA12 (1 µM) after 3 h of reaction. The unreacted DNA peaks were readily observable as before, as were a series of peaks of osmiated DNA. Osmium-containing species are easily identifiable by UHR-MS due to the characteristic isotope pattern for osmium isotopes (Figure 4.1b inset). It was immediately apparent from the full MS spectrum that the osmium complex had lost the monodentate chloride ligand, and was bound as $\{(\eta^6\text{-bip})Os(en)\}^{2+}$. Since the oligonucleotides were analysed in negative ion mode and the **Os1** modification is doubly charged, two additional protons have to be removed from the precursor to achieve the same charge state as the corresponding unmodified species, as detected by comparison with the theoretical isotope simulation shown in Figure 1b (inset) depicting the simulation and observed spectrum for [**DNA12**+**Os1**-9H]$^{7-}$. Additionally two distinct Os containing oligonucleotide isotopic distributions were observed, one for the mono-osmiated DNA 12mer

([**DNA12**+**Os1**-(n+2)H]$^{n-}$) and one for the di-osmiated DNA 12 mer ([**DNA12**+2**Os1**-(n+4)H]$^{n-}$).
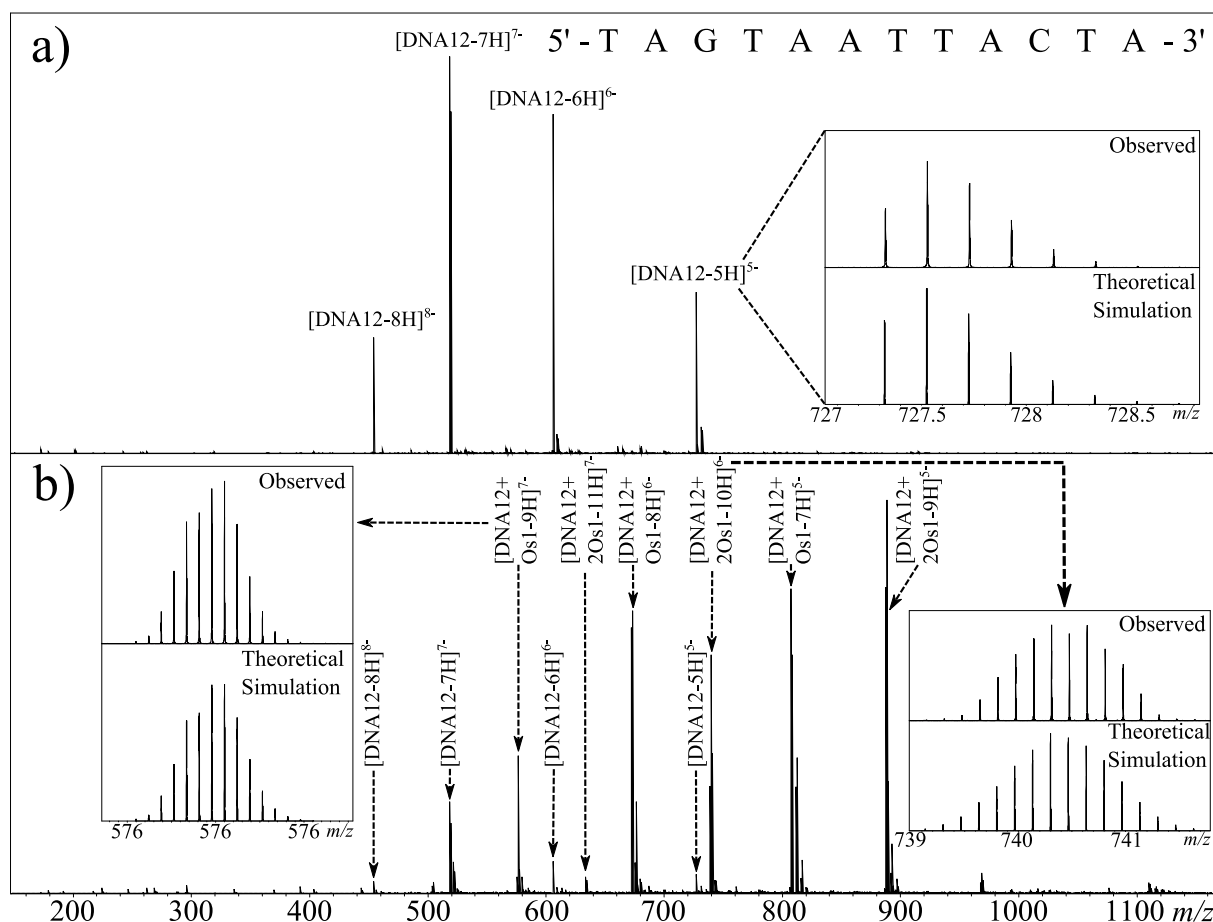
Figure 4.1: nESI Mass spectrum of (a) DNA12 and (b) DNA12+Os1 reaction mixture after 3 h incubation at 37°C (b). Inset: nucleotide sequence of the oligonucleotide 12-mer and observed vs. Simulated isotope patterns for unreacted DNA12 (top right), DNA12+Os1 (bottom left), and DNA12+2Os1 (bottom right).

The usual binding site for Ru^{II} and Os^{II} organometallic complexes within DNA is N7 of guanine, due to its high basicity and availability, even in the DNA duplex.[38] However, the DNA 12mer used during this study contained only one guanine residue. This suggested that in the present case a nucleotide other than guanine could also be a target for this metallodrug.

The CAD MS/MS spectrum for the [**DNA12**+**Os1**-8H]$^{6-}$ ion is shown in Figure 4.2b. Again, as with the corresponding unmodified DNA, CAD produced a spectrum with a large number of fragments, some osmiated, some not, and a range of assignable fragments (Figure 4.2b and Table S4.2 in SI). Analysis of the MS/MS data showed Os-containing fragments could be used to assign the binding location to the A_2-G_3 region of the oligonucleotide. CAD is based upon multiple collisions of the species of interest with background gas (in this case Ar),

238

producing a "slow heating effect" of ions until the weakest bonds within the molecule break, causing the dissociation.[18] Unfortunately, due to the slow heating caused by CAD, biomolecule modifications are often dissociated along with the biomolecule backbone;[39] this effect was also observed for the CAD spectra of osmiated DNA in this study, as shown by the observation of [**DNA12**-xH]$^{x-}$ species in the CAD spectra and further fragmentation. There are clear peaks showing the loss of the en ligand, followed by the loss of the whole complex during CAD, producing the unmodified species ([**DNA12**-xH]$^{x-}$), which then fragments again under the multiple collisions in CAD and produces a string of unmodified fragments. These fragments from the unmodified species can lead to problems during data analysis, such as creating fragments which contradict the location of the modification sites, and can add extra fragment peaks to already dense MS/MS spectra. Unfortunately the loss of the DNA modification under CAD meant that unmodified fragments could not be used to locate the modification, instead only the Os-containing fragments were used during this analysis. However, despite this limitation the modification was narrowed down to the A_2-G_3-T_4 region for the DNA 12-mer from the CAD MS/MS spectrum.

(a)

[a6-B]²⁻   [w3]¹⁻

[DNA12-7H]⁷⁻ CAD 8V

a/a-BH

5' - T | A | G | T | A | A | T | T | A | C | T | A - 3'

w
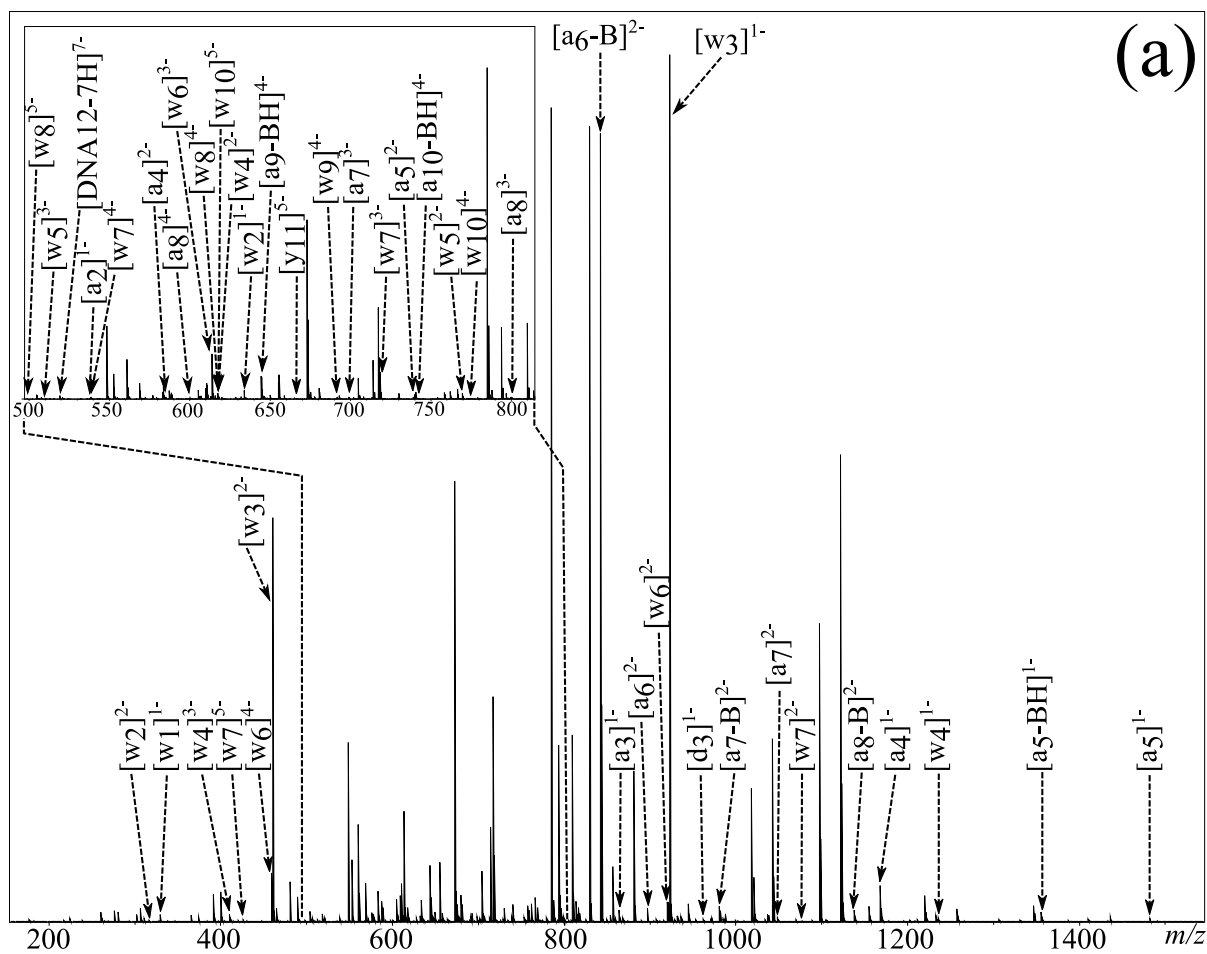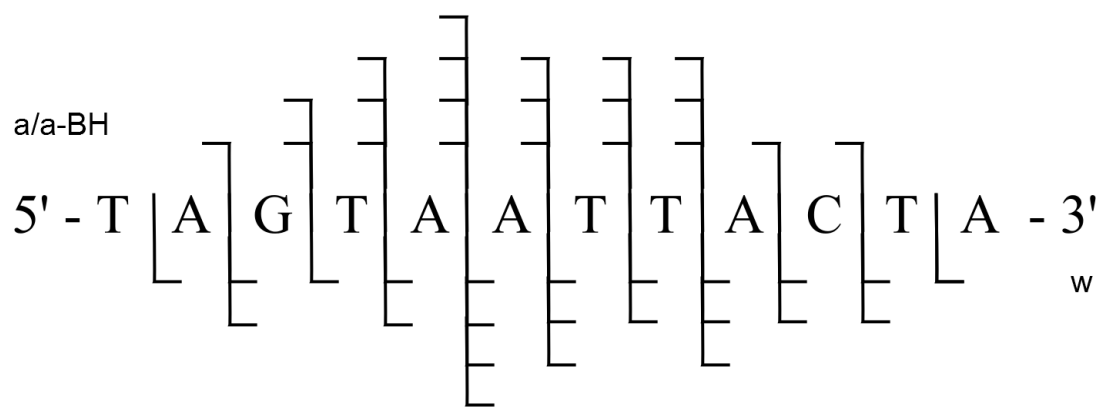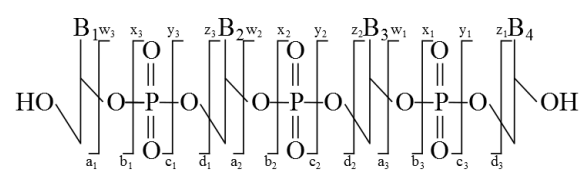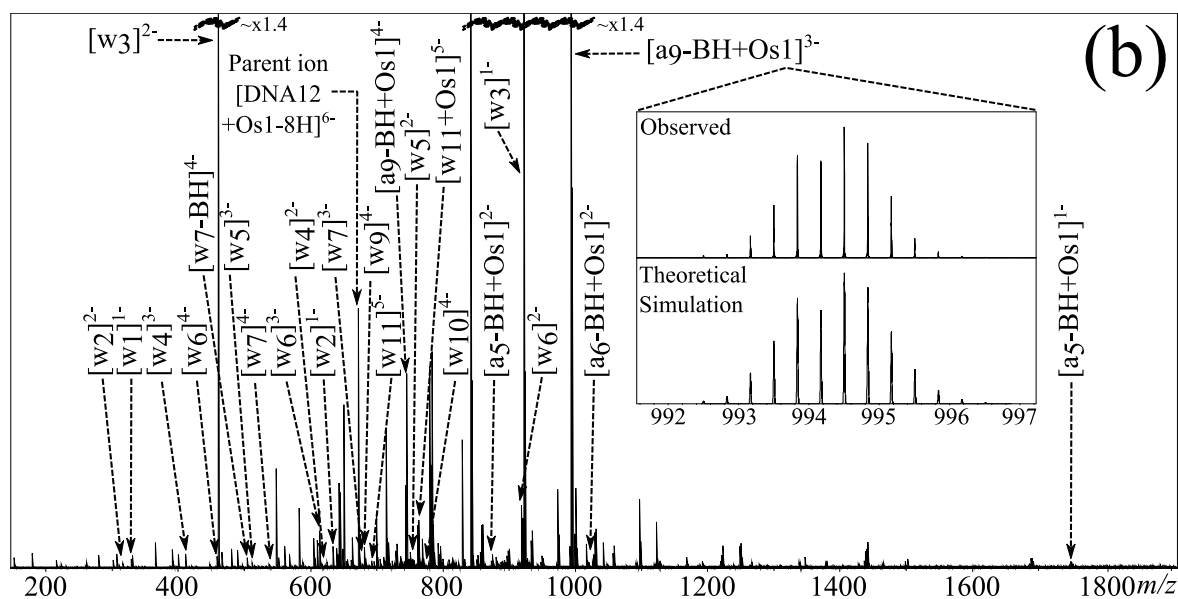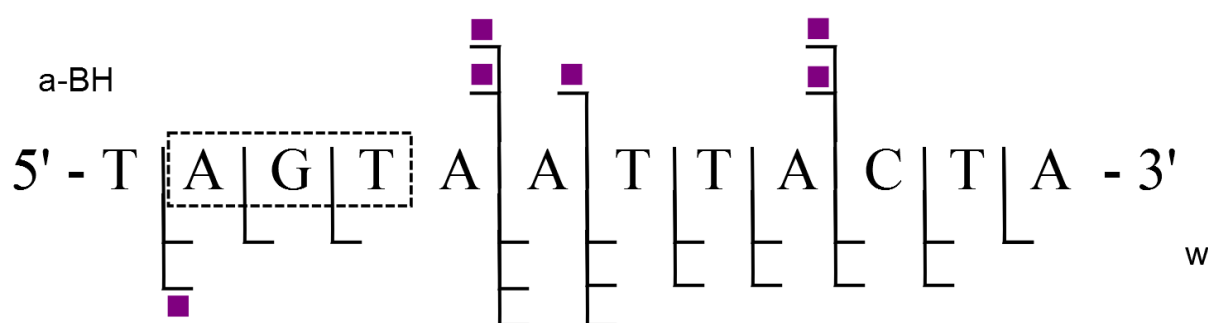
Figure 4.2: FT-ICR CAD MS/MS spectra for [DNA12-7H]^7- (a) and [DNA12+Os1-9H]^7- (b) species, along with corresponding fragmentation maps; osmiated fragments are indicated with shaded squares. Inset: Theoretical simulation of the [a_9-BH+Os1-5H]^3- fragment and the observed isotopic pattern during CAD MS/MS.

In order to identify the osmium binding sites unequivocally, further fragmentation techniques were investigated to explore their ability to fragment the backbone, while maintaining the osmium modification. It has been recently shown by Xu *et al.*[40] that although different slow heating MS/MS techniques produce very similar MS/MS spectra of poly(amino acids) (CAD vs. IRMPD), if tuned correctly, these dissociations show a much greater difference for oligonucleotides; with changes in the proportion of fragments exhibiting base loss, internal fragmentation, and proportions of different types of ions (e.g. a & w). IRMPD was conducted on the same isolated osmiated DNA species as the CAD shown in Figure 4.2b. Unfortunately, although the laser-based fragmentation produced many fragments, IRMPD dissociated the osmium modification to an even greater extent than CAD, despite varying the

laser power and pulse length. Due to the extensive secondary fragmentation and modification loss; IRMPD was shown to be unsuitable for studying the osmiated DNA species here.

In view of the problems associated with dissociations caused by slow heating when studying biomolecule modifications, many reported studies now utilise electron-based fragmentations, which can cleave backbone bonds while preserving modifications/adducts. By transferring or creating radical sites on/at the backbone, fragmentation is achieved via radical reactions/rearrangements.[19,41] These electron-based fragmentations have been effective in studies of modifications in positive ion mode via electron capture, both for natural modifications[42] and for metal/metallodrug binding to peptides/proteins.[13,43] However for negative ions, capture of electrons, such as those from Electron Capture Dissociation (ECD), is less likely due to the negative nature of the analyte, causing electrostatic repulsion. Instead other electron-based techniques such as Electron Detachment Dissociation (EDD) have been developed to allow electron-based fragmentation of negative ions.[24] EDD is a dissociative technique unique to FT-ICR MS, which causes backbone cleavage by irradiating oligonucleotides or other biomolecules, with medium energy electrons (ca. 14-24 eV).[44,45] The incoming electrons interact with the negative ions and cause emission/release of an electron from the analyte, creating a radical site, which, like other electron based fragmentations, can then cause dissociations along the backbone. A general pathway for this process is shown in Scheme 4.2.

$$[N]^{n-} \rightarrow [N]^{n-} + e^-_{fast} \rightarrow [N]^{(n-1)-\cdot} + \text{Fragments} + 2e^-$$

Scheme 4.2: A general EDD reaction pathway, adapted from Zubarev *et al.*[24]

EDD requires very fine tuning of the electron energy and other cathode parameters to cause effective dissociation, but once tuned can result in extremely high quality fragmentation patterns. Figure 4.3a shows the EDD MS/MS spectrum created by irradiation of the [**DNA12**-7H]$^{7-}$ species with 20.2 eV electrons. Due to the removal of electrons from the original ion, so called "Charge Reduced Species" (CRS) for the [**DNA12**-7H]$^{6-}$ and [**DNA12**-7H]$^{5-}$ ions are observed as high intensity peaks in the spectrum. Along with the CRS peaks,

an extensive series of **w** and **d** ions can be observed and assigned, summarised in the oligonucleotide fragmentation map in Figure 4.3a (and Table S4.3 in SI).

a)

Parent ion [DNA12-7H]$^{7-}$

1st charge reduced ion [DNA12-7H]$^{6-}$

2nd charge reduced ion [DNA12-7H]$^{5-}$

~x10   ~x1.5   ~x1.5

[d2+H]$^{2-}$
[d1+H]$^{1-}$
[w1]$^{1-}$
[w6]$^{4-}$
[w3]$^{2-}$
[d3+H]$^{2-}$
[w5]$^{3-}$
[d7+H]$^{4-}$ [w7]$^{4-}$
[d7+H]$^{4-}$
[w9]$^{5-}$
[w6]$^{3-}$
[w10]$^{5-}$
[d6+H]$^{3-}$ [w4]$^{2-}$
[d8+H]$^{4-}$
[d4+H]$^{2-}$
[d2+H]$^{1-}$
[w11]$^{5-}$
[w9]$^{4-}$
[w7]$^{3-}$
[d9+H]$^{4-}$
[d7+H]$^{3-}$
[w5]$^{2-}$
[d10+H]$^{4-}$
[d11+H]$^{4-}$
[w6]$^{2-}$
[w3]$^{1-}$

200   300   400   500   600   700   800   900   1000   *m/z*

[DNA12-7H]$^{7-}$ EDD:

d+H

5'-T A G T A A T T A C T A -3'

w

244
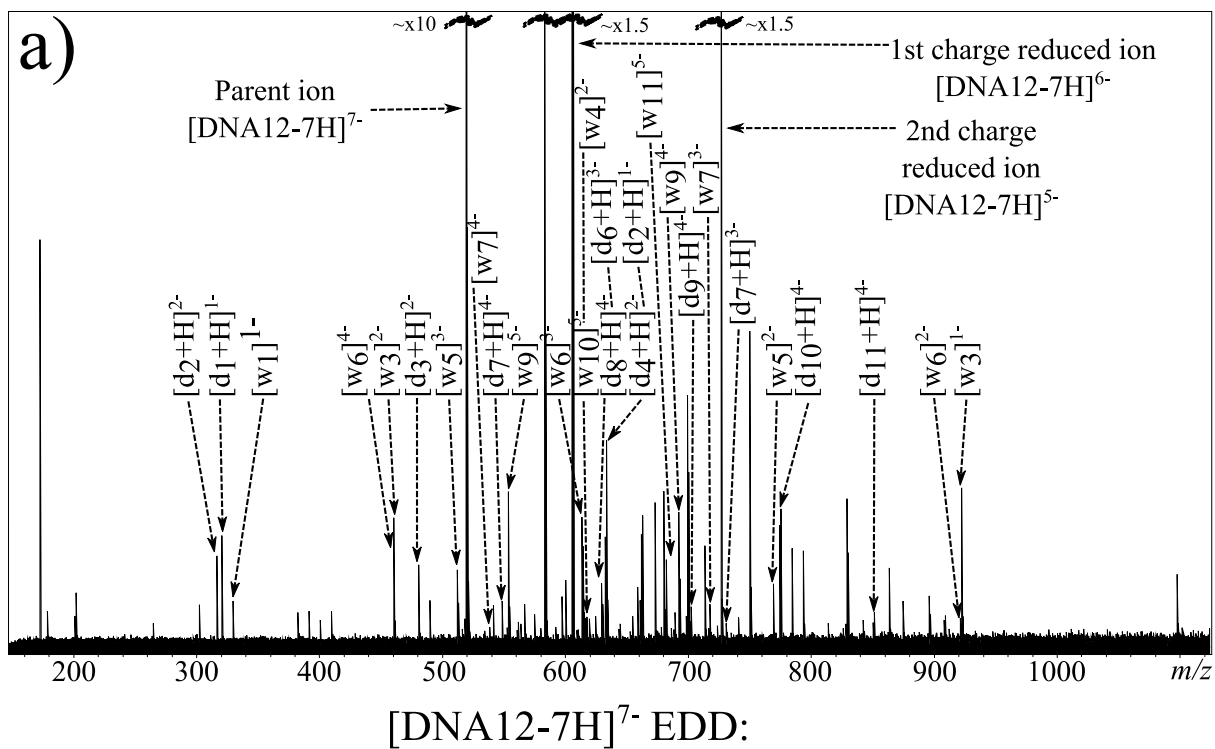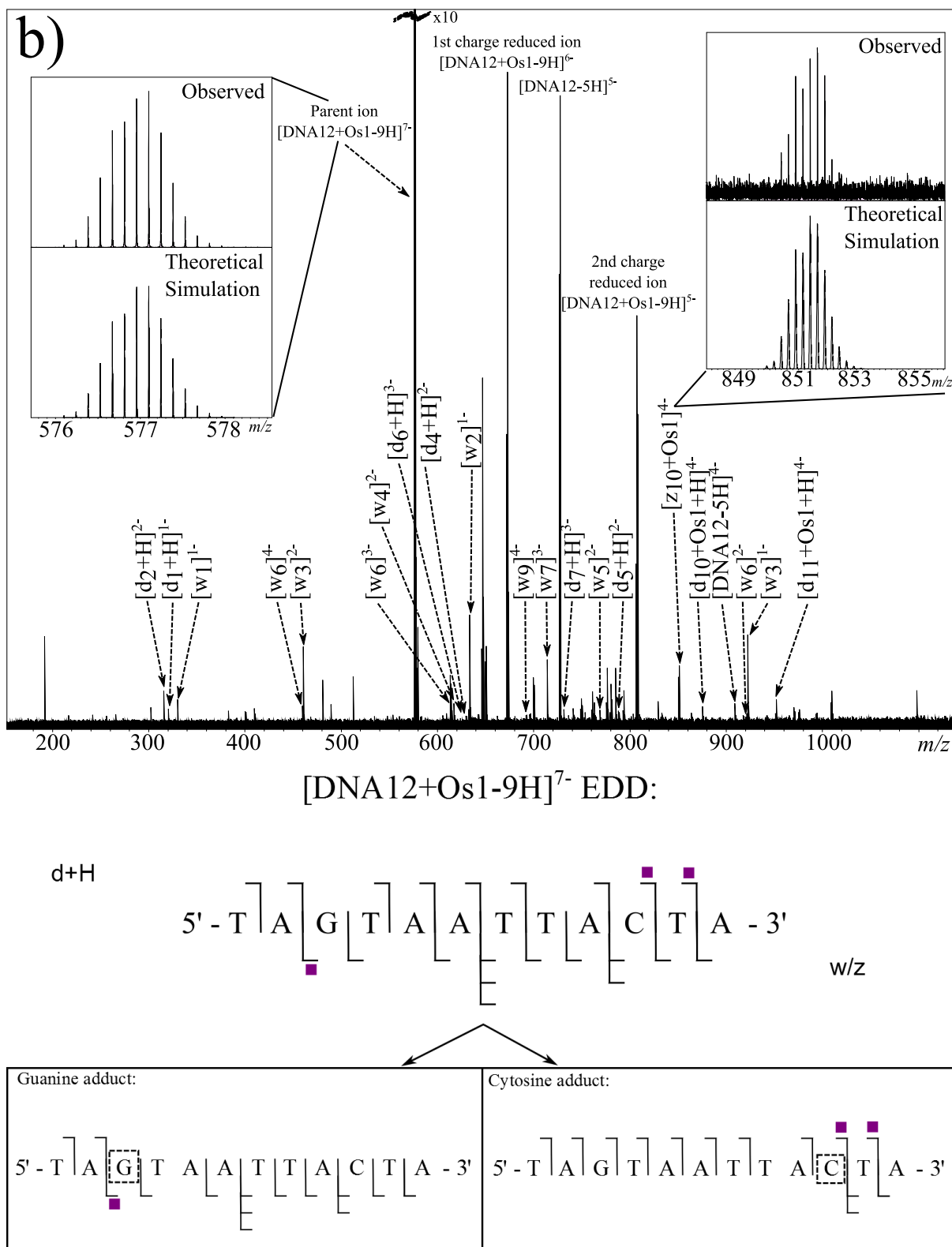
Figure 4.3: Electron Detachment Dissociation FT-ICR MS/MS spectra of (a) [DNA12-7H]$^{7-}$ and (b) [DNA12+Os1-9H]$^{7-}$ ions, along with corresponding fragmentation maps. Shaded squares indicate presence and number of osmium complex modifications bound to observed fragments. Inset; comparisons of selected theoretical simulations and observed species.

EDD was then carried out on the [**DNA12**+**Os1**-9H]$^{9-}$ species using the same parameters as for the unmodified species. The resulting MS/MS spectrum is shown in Figure 4.3b. Again, as with the previous EDD MS/MS spectrum, very intense peaks for the CRS were produced, as were a string of lower intensity fragments. Assigned fragments are summarised in the fragmentation map in Figure 4.3b (and Table S4.4 in SI). EDD produced 100% cleavage coverage of the isolated species. The resulting spectrum also contained far fewer internal fragment ions, mainly consisting of **w** and **d** ions, meaning that a larger proportion of spectral peaks corresponded to sequence informative fragments. The fragments assigned indicate two different binding locations for the bound {(η^6-bip)Os(en)} complex. Fragmentation analysis clearly shows the T_1-A_2 and T_4-A_{12} regions are free from osmium modification, while the G_3-A_{12} region is modified by the osmium modification, which clearly indicated the osmium modification was bound to the Guanine (G_3) nucleotide. Analysis also showed a second product in which the T_1-A_9 region was free from osmium modification, however the T_1-C_{10} and T_1-T_{11} regions were **Os1** modified, clearly showing the osmium complex was bound to cytosine (C_{10}) nucleotide. Unlike CAD MS/MS scans shown above, EDD cleaved the A_2-G_3 bond and narrowed down the exact site of the modification to one nucleotide. A peak corresponding to the unmodified oligonucleotide was produced during the EDD experiment; [**DNA12**-5H]$^{5-}$ at m/z~727 indicating that cleavage of the metallodrug from the DNA is only a minor fragmentation channel during EDD for this species. However, unlike CAD, further fragmentation of the **DNA12** ion was not observed, a CRS at sufficient intensity for the **DNA12** ion was not present, indicating that after dissociation of the metallodrug, the species did not undergo enough further electron capture to cause further dissociation and affect the peak intensities of fragments. The lack of further fragmentation of the unmodified [**DNA12**-xH]$^{x-}$ ions indicated that use of unmodified fragments for fragment analysis was viable and allowed the determination of the two reaction products in two discrete locations. The loss of the biphenyl ligand from the metallodrug complex was also observed, indicating the influence of Os^{II} on the electron dissociation process, though MS/MS analysis was not hindered in this instance.

MS/MS experiments were then conducted on the di-osmiated species at $m/z\sim740$ [**DNA12**+2**Os1**-10H]$^{6-}$. Effective CAD was achieved at a slightly higher value of 12 V, which could indicate a stabilising effect from the osmium complex in the gas phase. Figure 4a shows the CAD MS/MS spectrum obtained. The fragments assigned are summarised in the fragmentation map in Figure 4.4a (and Table S4.5 in SI). The osmiated fragments were more numerous than for the mono-osmiated **DNA12** MS/MS described earlier, and indicated binding to the A_2-G_3 region (correlating with the previous CAD result) and to the terminal 3 nucleotides at the 3′ end of **DNA12** (C_{10}-T_{11}-A_{12}). The results correlate well with the CAD and EDD results obtained for the mono-osmiated species, and are consistent with discrete G and C binding, with CAD showing poor fragmentation directly next to the osmiated nucleotide. The lack of fragmentation adjacent to the osmiated nucleotide via slow heating dissociation might indicate interactions between the osmium complex and its neighbouring nucleotides/bases; for instance, the biphenyl ligand might interact with the A_2, hindering dissociation of the A_2-G_3 bond, and likewise for the thymine residue neighbouring the osmiated cytosine at position 10. Since CAD MS/MS of the unmodified species produced such extensive fragmentation, analysis of the differences compared to the osmiated product may be indicative of interactions with neighbouring groups.
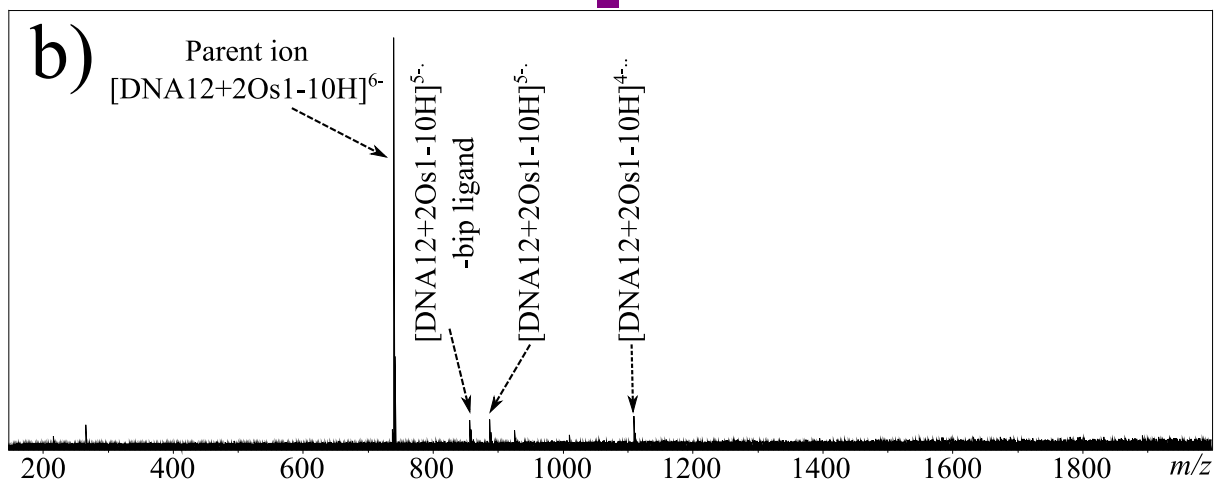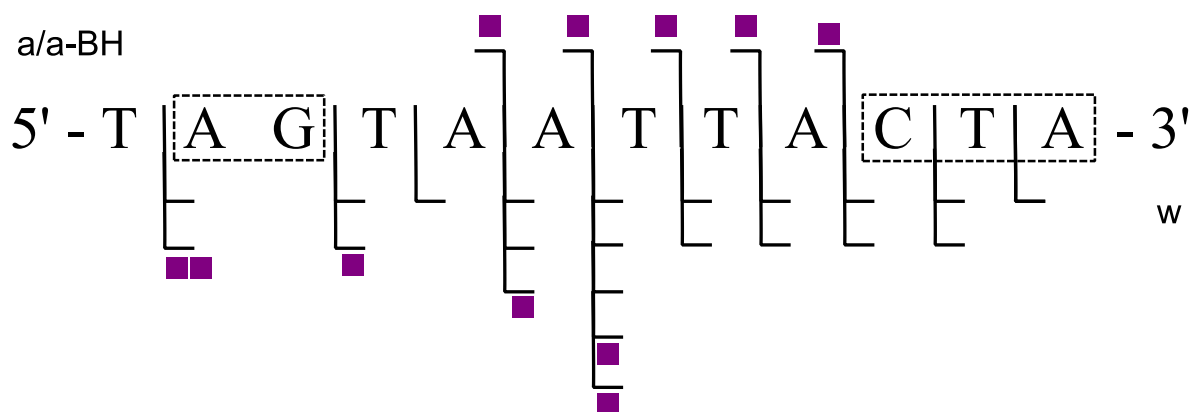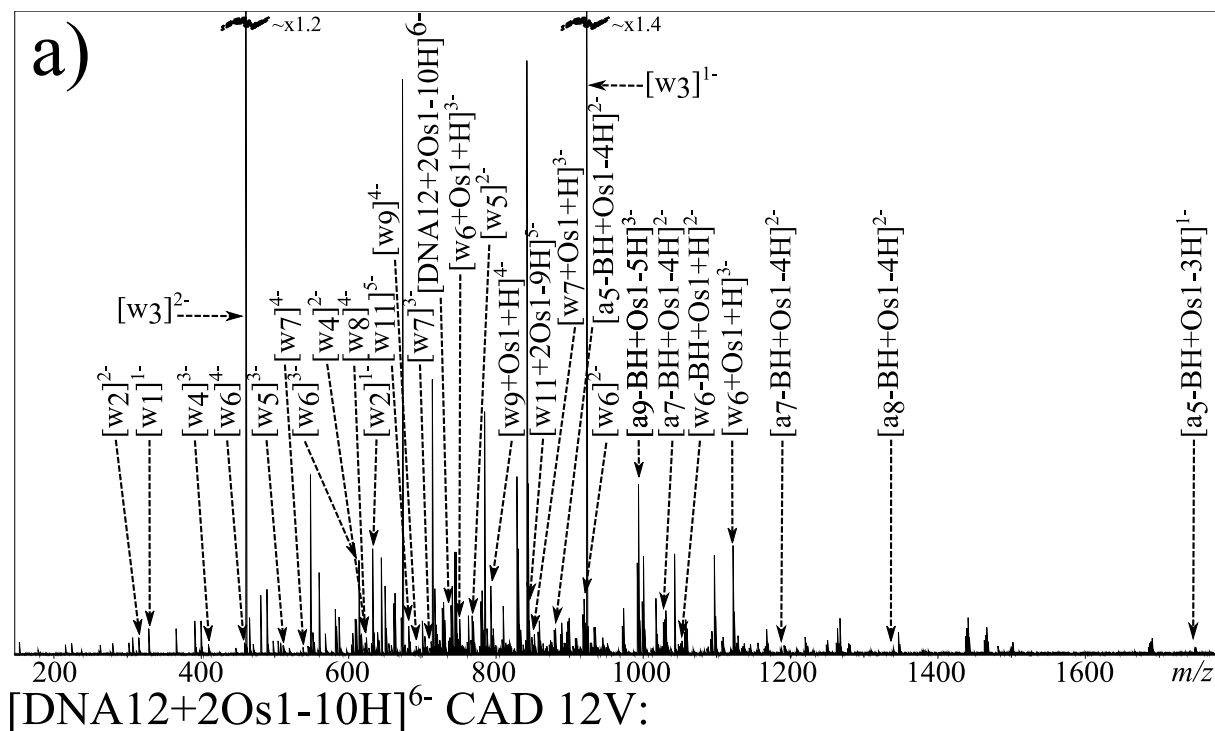
Figure 4.4: FT-ICR MS/MS spectra of the [DNA12+2Os1-10H]⁶⁻ species fragmented by (a) CAD, (b) EDD, along with the corresponding fragmentation map for CAD (centre), no backbone dissociation was observed via EDD for the (possibly cyclic) species presumably due to non-covalent interactions preventing gas phase fragmentation. Shaded squares indicate presence and number of osmium complex modifications bound to observed fragments.

EDD MS/MS was also attempted on the [**DNA12**+2**Os1**-10H]$^{6-}$ species using the same EDD parameters as before. The resulting spectrum, shown in Figure 4.4b, shows consistent electron capture events as before, with the observation of the [**DNA12**+2**Os1**-10H]$^{5-}$ and [**DNA12**+2**Os1**-10H]$^{4-}$ species. Loss of the bip ligand was also observed as before. Unfortunately very little fragmentation of the DNA backbone was observed during EDD MS/MS of the di-osmiated species, which was unexpected compared to the 100% sequence coverage observed for the unmodified species and mono-osmiated oligonucleotide described above. Lack of fragmentation during electron based dissociation such as this is usually the result of electron quenching/capture (e.g. by electron traps such as metal ions) with no resulting fragmentation thereafter, otherwise it is usually the indication of cyclisation by either covalent or non-covalent interactions, with dissociation of the backbone producing two fragments which would stay bound together by the interaction elsewhere in the macrocycle, hindering sequence informative fragmentation. Since the osmium complex {(η^6-bip)Os(en)}$^{2+}$ was shown to bind to different nucleotides during the EDD MS/MS of the mono-osmiated species above, and the mass of the precursor clearly showed retention of all but the chloride ligand of **Os1-Cl**, it can be concluded that the osmium modifications at discrete locations at either end of the oligonucleotide (as shown by the EDD MS/MS) hinder the electron based fragmentation when both binding sites are occupied simultaneously (as shown by the EDD here). These observations could be explained by the osmium arene complexes causing folding of the oligonucleotide, most likely involving π-π stacking of two biphenyl ligands, one from each osmium complex, meaning when the oligonucleotide backbone fragments, the resulting fragments remain bound non-covalently, in a similar fashion to a poly(amino acid) CRS during electron based dissociations.[46]

Conclusions

The organometallic half-sandwich 'piano-stool' osmium(II) complex **Os1-Cl**, is thought to exert its anticancer activity partly through binding to DNA.[28] The clinical anticancer drug cisplatin also has DNA as a target site, causing deformation of the DNA structure and, eventually, leading to apoptosis. However, since **Os1-Cl** is not cross-resistant with cisplatin,

the mode of interaction with DNA would be expected to be different. Indeed **Os1-Cl** is monofunctional and cannot readily crosslink DNA bases, in contrast to bifunctional cisplatin. The base specificity of **Os1-Cl** may also be different from cisplatin. Prior to the current work previous studies had indicated that guanine N7 is a preferred binding site for **Os1**-Cl on DNA as it is for cisplatin.

In this work, $[(\eta^6\text{-biphenyl})Os(en)Cl]^+$ was reacted with a 12mer DNA oligonucleotide containing single guanine and cytosine sites under conditions (low ionic strength) in which it was single-stranded. Using ultra-high resolution-FT-ICR MS up to two osmium complexes were shown to bind to **DNA12**. Each reaction product was fragmented via tandem mass spectrometry (MS/MS) and the identity and binding location of the modifications were elucidated.

$\{(\eta^6\text{-biphenyl})Os(en)\}^{2+}$ was shown to bind to both the guanine and cytosine residues, both independently in the mono-osmiated $[\text{oligonucleotide}+[(\eta^6\text{-biphenyl})Os(en)]-(n+2)H]^{n-}$ species, and in the di-osmiated $[\text{oligonucleotide}+2[(\eta^6\text{-biphenyl})Os(en)]-(n+4)H]^{n-}$ species.

Though Collision Activated Dissociation (CAD) produced extensive fragmentation and excellent sequence coverage with the unmodified oligonucleotide, the metallodrug modification dissociated from the mono-osmiated oligonucleotides. With the retention of the complex in only a fraction of the fragments, CAD was able to narrow the location of the binding site of the modification to the A_2-G_3-T_4 region of the sequence. However Electron Detachment Dissociation (EDD) allowed unambiguous identification of both guanine (G_3) and cytosine (C_{10}) binding from assignment of fragmentation patterns. Fragmentation of the di-osmiated **DNA12** species proved more difficult, since the two bound osmium complexes appeared to stabilise the oligonucleotide toward EDD fragmentation. Such stabilisation toward EDD fragmentation observed here for **DNA12**-2**Os1** species may result from folding of **DNA12** due to π-π stacking interactions between the two biphenyl ligands from separate guanine- and cytosine-bound osmium complexes (as previously characterised by NMR for related guanine bound Ru arene-oligonucleotide adducts)[46], creating a cyclic gas phase species and hindering fragmentation via using electron detachment. Collisional dissociation

methods were able to dissociate the species due to a multiple collision/slow heating mechanism providing results in agreement to those for the mono-osmiated species studied.

This work demonstrates the flexibility and aptitude of FT-ICR Mass Spectrometry for Top-Down fragmentation of DNA-metallodrug reaction products, and shows how a multi-platform fragmentation approach is often needed to fully characterise metallodrug binding to biomolecules of interest. FT-ICR MS provides the highest resolving power and mass accuracy performance of any mass spectrometry platform available, and offers the highest possible reliability and confidence in mass spectral assignment, coupled with the largest range of fragmentation techniques. FT-ICR MS is therefore uniquely suited to Top-Down fragmentation of metallodrug-biomolecule adducts.

Acknowledgements

References:

(1)   Dyson, P. J.; Sava, G. *Dalton Trans.* **2006**, No. 16, 1929–1933.

(2)   Betanzos-Lara, S.; Salassa, L.; Habtemariam, A.; Novakova, O.; Pizarro, A. M.; Clarkson, G. J.; Liskova, B.; Brabec, V.; Sadler, P. J. *Organometallics* **2012**, *31* (9), 3466–3479.

(3)   Hearn, J. M.; Romero-Canelon, I.; Qamar, B.; Liu, Z.; Hands-Portman, I.; Sadler, P. J. *ACS Chem. Biol.* **2013**, *8* (10), 2345.

(4)   Strianese, M.; Basile, A.; Mazzone, A.; Morello, S.; Turco, M. C.; Pellecchia, C. *J. Cell. Physiol.* **2013**, *228* (11), 2202–2209.

(5)   Funston, A. M.; Cullinane, C.; Ghiggino, K. P.; McFadyen, W. D.; Stylli, S. S.; Tregloan, P. a. *Aust. J. Chem.* **2005**, *58* (3), 206.

(6)   Peacock, A. F. a; Habtemariam, A.; Fernández, R.; Walland, V.; Fabbiani, F. P. a; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J. *J. Am. Chem. Soc.* **2006**, *128* (5), 1739–1748.

(7)   Mackay, F. S.; Woods, J. a; Moseley, H.; Ferguson, J.; Dawson, A.; Parsons, S.; Sadler, P. J. *Chemistry* **2006**, *12* (11), 3155–3161.

(8)   Alderden, R. A.; Hall, M. D.; Hambley, T. W. *J. Chem. Educ.* **2006**, *83* (5).

(9)   Wong, E.; Giandomenico, C. M. *Chem. Rev.* **1999**, *99* (9), 2451–2466.

(10)   Reedijk, J. *Eur. J. Inorg. Chem.* **2009**, *2009* (10), 1303–1312.

(11)   Pizarro, A. M.; Sadler, P. J. *Biochimie* **2009**, *91* (10), 1198–1211.

(12)   Centerwall, C.; Tacka, K. *Mol. Pharmacol.* **2006**, 348–355.

(13)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 9507–9515.

(14)   Farrer, N. J.; Gierth, P.; Sadler, P. J. *Chemistry* **2011**, *17* (43), 12059–12066.

(15)   Calderone, V.; Casini, A.; Mangani, S.; Messori, L.; Orioli, P. L. *Angew. Chem. Int. Ed. Engl.* **2006**, *45* (8), 1267–1269.

(16)   Cho, Y.; Ahmed, A.; Islam, A.; Kim, S. *Mass Spectrom. Rev.* **2014**, 1–16.

(17)   McLafferty, F. W. *Int. J. Mass Spectrom.* **2001**, *212* (1-3), 81–87.

(18)   Jennings, K. R. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1*, 227–235.

(19)   Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (16), 3265–3266.

(20)   Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.

(21)   Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66* (18), 2809–2815.

(22)   Wolff, J. J.; Laremore, T. N.; Aslam, H.; Linhardt, R. J.; Amster, I. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (10), 1449–1458.

(23)   Bowers, W. D.; Delbert, S.; Hunter, R. L.; Mciver, R. T. *J. Am. Chem. Soc.* **1984**, No. 3, 7288–7289.

(24)   Budnik, B.; Haselmann, K.; Zubarev, R. *Chem. Phys. Lett.* **2001**, *342* (July), 299–302.

(25)   Amster, I. *J. Mass Spectrom.* **1996**, *31*, 1325–1337.

(26)   Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.

(27)   Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. a. *Anal. Chem.* **2014**, *86* (1), 317–320.

(28)   McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X.; Infusini, G.; Jiang, H.; Kong, X.; Begley, T. P. *FEBS J.* **2007**, *274* (24), 6256–6268.

(29) Peacock, A. F. A.; Habtemariam, A.; Ferna, R.; Walland, V.; Fabbiani, F. P. A.; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J. **2006**, *128*, 1739–1748.

(30) Quenzer, T. L.; Emmett, M. R.; Hendrickson, C. L.; Kelly, P. H.; Marshall, a G. *Anal. Chem.* **2001**, *73* (8), 1721–1725.

(31) Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26* (October 1990), 514–518.

(32) Zhang, L.-K.; Rempel, D.; Pramanik, B. N.; Gross, M. L. *Mass Spectrom. Rev.* **2005**, *24* (2), 286–309.

(33) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. *Anal. Chem.* **2013**, *85*, 2–9.

(34) Wu, J.; McLuckey, S. a. *Int. J. Mass Spectrom.* **2004**, *237* (2-3), 197–241.

(35) Phillips, D. R.; McCloskey, J. A. *Int. J. Mass Spectrom. Ion Process.* **1993**, *128* (1-2), 61–82.

(36) Tromp, J.; Schürch, S. *Rapid Commun. Mass Spectrom.* **2006**, 2348–2354.

(37) Nyakas, A.; Blum, L. C.; Stucki, S. R.; Reymond, J.-L.; Schürch, S. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (2), 249–256.

(38) Reedijk, J. *Chem. Rev.* **1999**, *99* (9), 2499–2510.

(39) Qi, Y.; Liu, Z.; Li, H.; Sadler, P. J.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2013**, *27* (17), 2028–2032.

(40) Xu, Z.; Shaw, J. B.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (2), 265–273.

(41) Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125* (29), 8949–8958.

(42) Zhang, J.; Guy, M. J.; Norman, H. S.; Chen, Y.-C.; Xu, Q.; Dong, X.; Guner, H.; Wang, S.; Kohmoto, T.; Young, K. H.; Moss, R. L.; Ge, Y. *J. Proteome Res.* **2011**, *10* (9), 4054–4065.

(43)   Li, H.; Zhao, Y.; Phillips, H. I. a; Qi, Y.; Lin, T.-Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83* (13), 5369–5376.

(44)   Taucher, M.; Breuker, K. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 918–929.

(45)   Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (2), 234–244.

(46)   Anusiewicz, I.; Berdys-Kochanska, J.; Simons, J. *J. Phys. Chem. A* **2005**, *109* (26), 5801–5813.

Binding of an Organo-osmium (II) Anticancer Complex to Guanine and Cytosine on DNA Revealed by Electron-based Dissociations in High Resolution Top-Down FT-ICR Mass Spectrometry

Christopher A. Wootton, Carlos Sanchez-Cano, Hongke Liu, Mark P. Barrow, Peter J. Sadler,*

Peter B. O'Connor*

Supporting Information

Table S4.1. Fragment assignments for Figure 4.2A: CAD MS/MS of the [**DNA12**-7H]$^{7-}$ species. Highlighted species (red) were used for internal calibration.

Assignment	Observed m/z	Exact mass/ m/z	Mass error/ppm
[w3]3-	307.04607	307.046068	0.01
[w2]2-	316.54956	316.549554	0.02
[w1]1-	330.06062	330.060346	0.83
[w4]3-	411.39860	411.398603	-0.01
[w7]5-	430.66642	430.666187	0.54
[w6]4-	460.32002	460.320152	-0.29
[w3]2-	461.07251	461.072740	-0.50
[w8]5-	493.27759	493.277708	-0.24
[w5]3-	512.74725	512.747282	-0.06
[DNA12-7H]7-	519.22994	519.230058	-0.23
[w7]4-	538.58456	538.584553	0.01
[a2]1-	536.12963	536.130584	-1.78
[a4]2-	584.11046	584.110931	-0.81
[a8]4-	600.10311	600.103648	-0.90
[w6]3-	614.09573	614.095961	-0.38
[w4]2-	617.60162	617.601542	0.13
[w10]5-	619.89740	619.897419	-0.03
[w8]4-	616.84874	616.848954	-0.35
[a9-BH]4-	644.60404	644.604562	-0.81
[w2]1-	634.10696	634.106383	0.91
[y11]5-	666.51578	666.515674	0.16
[w9]4-	692.86094	692.860463	0.69
[a7]3-	699.12472	699.125277	-0.80
[w7]3-	718.44837	718.448496	-0.17
[a5]2-	740.63912	740.639733	-0.83
[a10-BH]4-	722.86938	722.868963	0.58
[w5]2-	769.62440	769.624561	-0.21
[w10]4-	775.12451	775.123593	1.18
[a8]3-	800.47396	800.473956	0.005
[a6-B]2-	829.64287	829.641559	1.58
[a3]1-	865.18203	865.183104	-1.24
[a6]2-	897.16734	897.168536	-1.33
[w6]2-	921.64754	921.647579	-0.04
[w3]1-	923.15106	923.152755	-1.84
[d3]1-	963.15877	963.158902	-0.14
[a7-B]2-	986.16924	986.170361	-1.14
[a7]2-	1049.19067	1049.191554	-0.84
[w7]2-	1078.17740	1078.176382	0.94
[a8-B]2-	1138.19271	1138.193380	-0.59
[a4]1-	1169.22763	1169.229141	-1.29
[w4]1-	1236.21034	1236.210360	-0.02

$[DNA12\text{-}7H]^{7-}$ CAD 8V

$$5' - T \mid A \mid G \mid T \mid A \mid A \mid T \mid T \mid A \mid C \mid T \mid A - 3'$$

a-type fragments:

$[a_2]^{1-}$, $[a_3]^{1-}$, $[d_3]^{1-}$, $[a_4]^{1-}$, $[a_4]^{2-}$, $[a_4\text{-GH}]^{2-}$, $[a_5]^{2-}$, $[a_5\text{-TH}]^{1-}$, $[a_5\text{-TH}]^{2-}$, $[a_5\text{-AH}]^{1-}$, $[a_6]^{2-}$, $[a_6\text{-A}]^{2-}$, $[a_7]^{2-}$, $[a_7]^{3-}$, $[a_7\text{-B}]^{2-}$, $[a_8]^{3-}$, $[a_8]^{4-}$, $[a_8\text{-A}]^{2-}$, $[a_9\text{-TH}]^{4-}$, $[a_{10}\text{-CH}]^{4-}$

w-type fragments:

$[w_1]^{1-}$, $[w_2]^{1-}$, $[w_2]^{2-}$, $[w_3]^{1-}$, $[w_3]^{2-}$, $[w_4]^{1-}$, $[w_4]^{2-}$, $[w_4]^{3-}$, $[w_5]^{2-}$, $[w_5]^{3-}$, $[w_6]^{2-}$, $[w_6]^{3-}$, $[w_6]^{4-}$, $[w_7]^{2-}$, $[w_7]^{3-}$, $[w_7]^{4-}$, $[w_7]^{5-}$, $[w_8]^{4-}$, $[w_8]^{5-}$, $[w_9]^{4-}$, $[w_{10}]^{4-}$, $[w_{10}]^{5-}$, $[y_{11}]^{5-}$
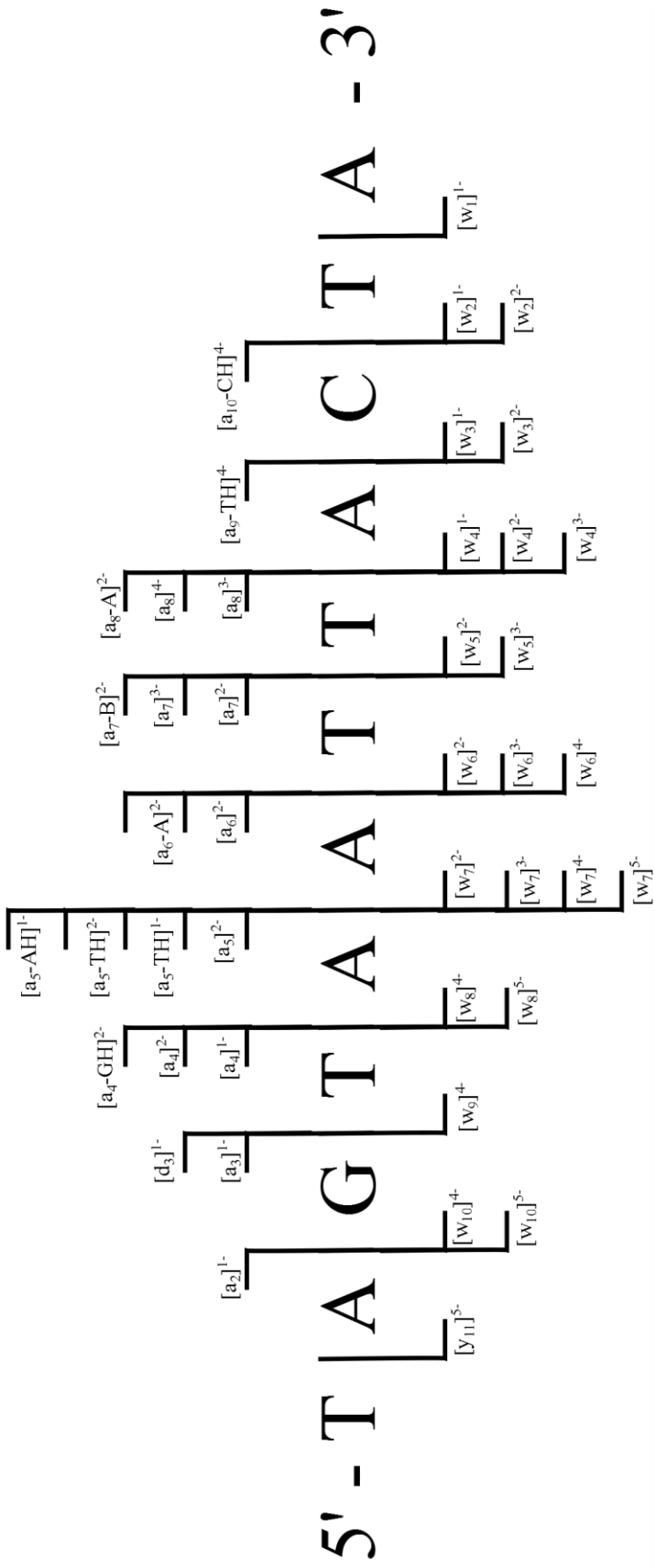
Table S4.2. Fragment assignments for Figure 4.2B: CAD MS/MS of the [**DNA12+Os1**-8H]$^{6-}$ species. Highlighted species (red) were used for internal calibration. Note: the highest intensity isotope was used for fragments containing the osmium modification.

Assignment	Observed m/z	Exact mass/ m/z	Mass error/ppm
[w2]2-	316.54955	316.549554	-0.01
[w1]1-	330.06062	330.060346	0.83
[w4]3-	411.39854	411.398603	-0.15
[w6]4-	460.32003	460.320152	-0.26
[w3]2-	461.07265	461.072740	-0.19
[w7-BH]4-	504.82084	504.821065	-0.45
[w5]3-	512.74730	512.747282	0.04
[w7]4-	538.58431	538.584553	-0.45
[w6]3-	614.09598	614.095961	0.03
[w4]2-	617.60194	617.601542	0.64
[w2]1-	634.10680	634.106383	0.66
[w11]5-	682.50854	682.508940	-0.59
[w9]4-	692.86044	692.860463	-0.03
[w7]3-	718.44849	718.448496	-0.01
[a9-BH+Os1]4-	745.62715	745.627830	-0.91
[w11+Os1]5-	763.12695	763.126400	0.72
[w5]2-	769.62518	769.624561	0.80
[w10]4-	775.12384	775.123593	0.32
[a5-BH+Os1]2-	875.15856	875.158910	-0.40
[w6]2-	921.64732	921.647579	-0.28
[w3]1-	923.15207	923.152755	-0.74
[a9-BH+Os1]3-	994.50451	994.506200	-1.70
[a6-BH+Os1]2-	1031.68664	1031.687800	-1.12
[a5-BH+Os1]1-	1749.32253	1749.322510	0.01

[DNA12+Os1-8H]$^{6-}$ CAD 10V:

5' - T | A | G | T | A | A | T | T | A | C | T | A - 3'

[a$_5$-BH +Os1]$^{2-}$
[a$_5$-BH +Os1]$^{1-}$

[a$_6$-BH +Os1]$^{2-}$

[a$_9$-BH +Os1]$^{4-}$
[a$_9$-BH +Os1]$^{3-}$

[w$_{11}$]$^{5-}$
[w$_{11}$+Os1]$^{5-}$

[w$_{10}$]$^{4-}$

[w$_9$]$^{4-}$

[w$_7$]$^{3-}$
[w$_7$]$^{4-}$
[w$_7$-BH]$^{4-}$

[w$_6$]$^{2-}$
[w$_6$]$^{3-}$
[w$_6$]$^{4-}$

[w$_5$]$^{2-}$
[w$_5$]$^{3-}$

[w$_4$]$^{2-}$
[w$_4$]$^{3-}$

[w$_3$]$^{1-}$
[w$_3$]$^{2-}$

[w$_2$]$^{1-}$
[w$_2$]$^{2-}$

[w$_1$]$^{1-}$
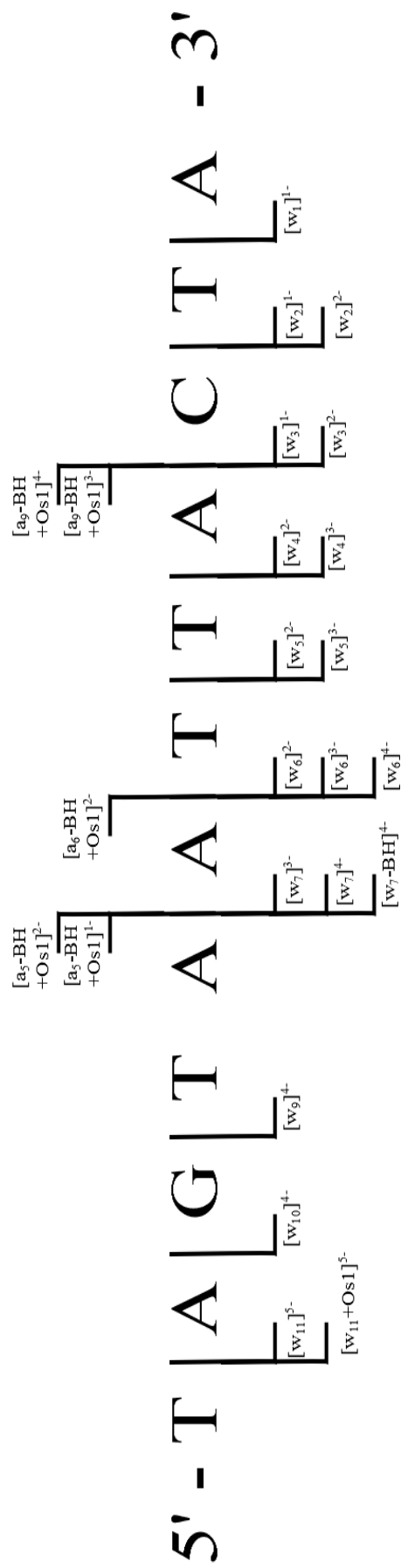
Table S4.3. Fragment assignments for Figure 4.3A: EDD MS/MS of the [**DNA12**-7H]$^{7-}$ species. Highlighted species (red) were used for internal calibration.

Assignment	Observed m/z	Exact mass/ m/z	Mass error/ppm
[d2+H]2-	316.54955	316.549555	-0.02
[d1+H]1-	321.04900	321.048777	0.69
[w1]1-	330.06060	330.060346	0.77
[w6]4-	460.31995	460.320152	-0.44
[w3]2-	461.07274	461.072740	0.00
[w5]3-	512.74716	512.747282	-0.24
[w7]4-	538.58460	538.584553	0.09
[w9]5-	554.08685	554.086915	-0.12
[w6]3-	614.09572	614.095961	-0.39
[w4]2-	617.60167	617.601542	0.21
[w10]5-	619.89676	619.897419	-1.06
[d8+H]4-	624.59780	624.597600	0.32
[d6+H]3-	630.43500	630.435200	-0.32
[d4+H]2-	633.09880	633.098834	-0.05
[d2+H]1-	634.10650	634.106382	0.19
[w11]5-	682.50894	682.508940	0.00
[w9]4-	692.86046	692.860463	0.00
[d9+H]4-	702.86181	702.862001	-0.27
[w7]3-	718.44892	718.448496	0.59
[d7+H]3-	731.78376	731.783879	-0.16
[w5]2-	769.62471	769.624561	0.19
[d10+H]4-	775.12384	775.123594	0.32
[d11+H]4-	851.13487	851.135103	-0.27
[w6]2-	921.64801	921.647579	0.47
[w3]1-	923.15276	923.152755	0.01

[DNA12-7H]$^{7-}$ EDD:

5' - T | A | G | T | A | A | T | A | C | T | A - 3'

$[d_1+H]^{1-}$
$[d_2+H]^{1-}$
$[d_2+H]^{2-}$
$[d_3+H]^{2-}$
$[d_4+H]^{2-}$
$[d_6+H]^{3-}$
$[d_7+H]^{3-}$
$[d_7+H]^{4-}$
$[d_8+H]^{4-}$
$[d_9+H]^{4-}$
$[d_{10}+H]^{4-}$
$[d_{11}+H]^{4-}$

$[w_1]^{1-}$
$[w_3]^{2-}$
$[w_3]^{3-}$
$[w_4]^{2-}$
$[w_5]^{2-}$
$[w_5]^{3-}$
$[w_6]^{2-}$
$[w_6]^{3-}$
$[w_6]^{4-}$
$[w_7]^{3-}$
$[w_7]^{4-}$
$[w_9]^{4-}$
$[w_9]^{5-}$
$[w_{10}]^{5-}$
$[w_{11}]^{5-}$
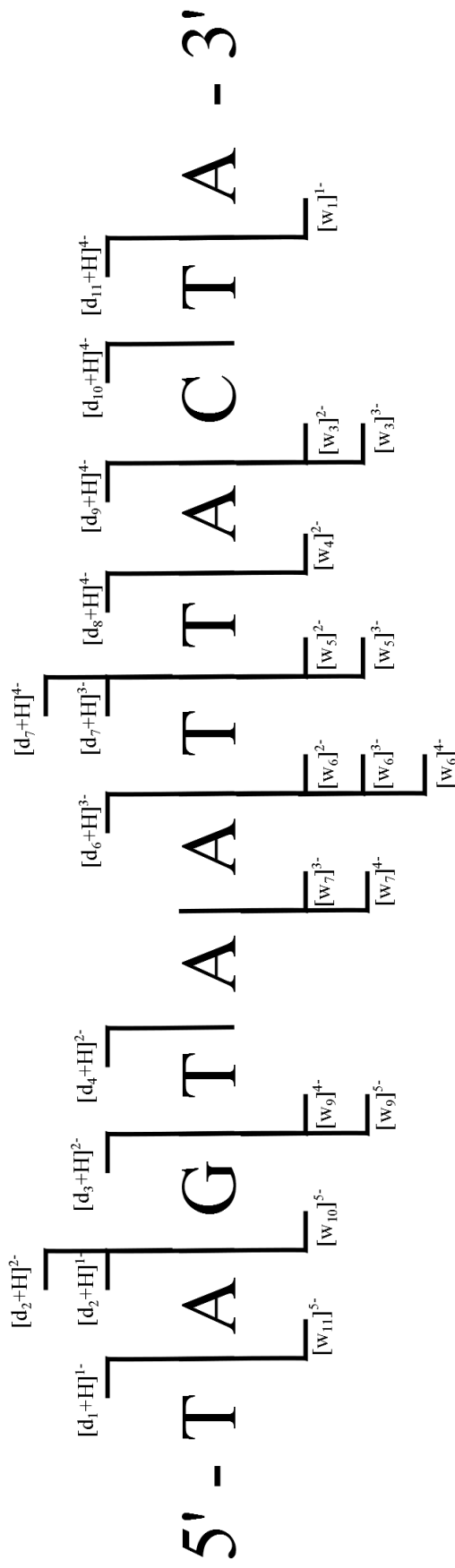
Table S4.4. Fragment assignments for Figure 4.3A: EDD MS/MS of the [**DNA12**+**Os1**-9H]$^{7-}$ species. Highlighted species (red) were used for internal calibration. Note: the highest intensity isotope was used for fragments containing the osmium modification.

Assignment	Observed m/z	Exact mass/ m/z	Mass error/ppm
[d2+H]2-	316.54955	316.549555	-0.02
[d1+H]1-	321.04908	321.048777	0.94
[w1]1-	330.06057	330.060346	0.68
[w6]4-	460.32029	460.320152	0.30
[w3]2-	461.07274	461.072740	0.00
[w6]3-	614.09596	614.095961	0.00
[w4]2-	617.60160	617.601542	0.09
[d6+H]3-	630.43488	630.435200	-0.51
[d4+H]2-	633.09931	633.098834	0.75
[w2]1-	634.10669	634.106383	0.48
[w9]4-	692.86083	692.860463	0.53
[d9+2H]4-	703.11513	703.113821	1.86
[w7]3-	718.44839	718.448496	-0.15
[d7+H]3-	731.78402	731.7838793	0.19
[w5]2-	769.62436	769.624561	-0.26
[d5+H]2-	789.62774	789.627636	0.13
[z10+Os1]4-	851.65110	851.651640	-0.63
[d10+Os1+H]4-	876.14725	876.147330	-0.09
[w6]2-	921.64677	921.647579	-0.88
[w3]1-	923.15289	923.152755	0.15
[d11+Os1+H]4- HP	952.40935	952.409350	0.00

[DNA12+Os1-9H]$^{7-}$ EDD:

5'- T A G T A A T T A C T A -3'
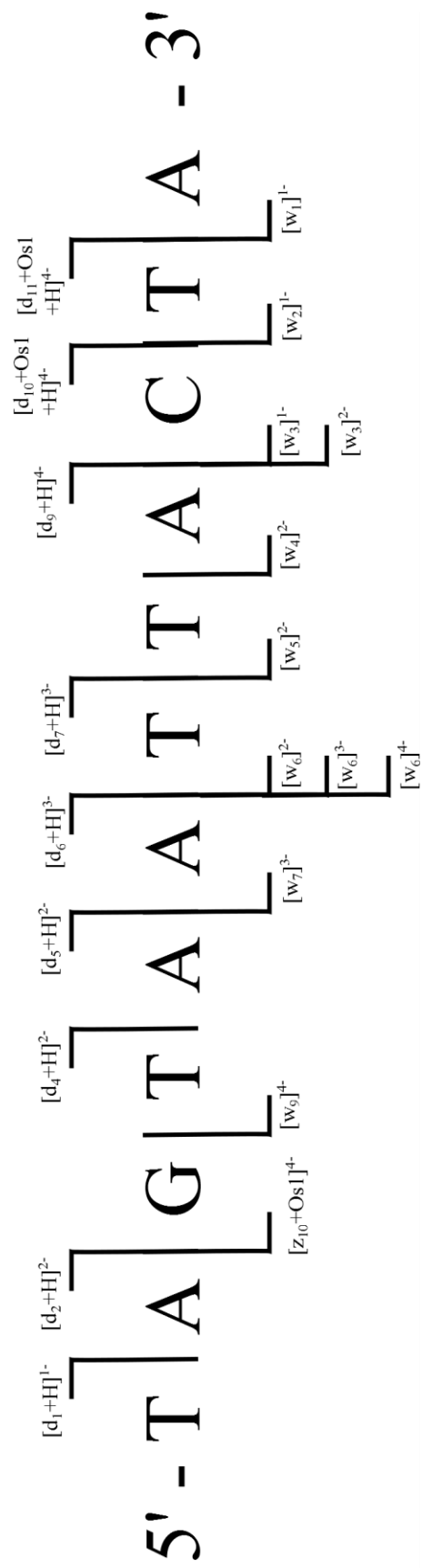
Table S4.5. Fragment assignments for Figure 4.4A: CAD MS/MS of the [**DNA12**+2**Os1**-10H]$^{6-}$ species. Highlighted species were used for internal calibration. Note: the highest intensity isotope was used for fragments containing the osmium modification.

Assignment	Observed m/z	Exact mass/ m/z	Mass error/ppm
[w2]2-	316.54955	316.549554	-0.01
[w1]1-	330.06062	330.060346	0.83
[w4]3-	411.39864	411.398603	0.09
[w6]4-	460.32015	460.320152	0.00
[w3]2-	461.07277	461.072740	0.07
[w5]3-	512.74745	512.747282	0.33
[w7]4-	538.58470	538.584553	0.27
[w6]3-	614.09589	614.095961	-0.12
[w8]4-	616.84921	616.848954	0.42
[w4]2-	617.60154	617.601542	0.00
[w2]1-	634.10670	634.106383	0.50
[w11]5-	682.50881	682.508940	-0.19
[w9]4-	692.86033	692.860463	-0.19
[w7]3-	718.44879	718.448496	0.41
[w6+Os1+H]3-	748.79373	748.794030	-0.40
[w5]2-	769.62505	769.624561	0.64
[w9+Os1+H]4-	793.63323	793.632200	1.30
[w11+2Os1-9H]5-	843.54589	843.544500	1.65
[w7+Os1+H]3-	853.14599	853.146620	-0.74
[a5-BH+Os1-4H]2-	875.15842	875.158910	-0.56
[w6]2-	921.64778	921.647579	0.22
[w3]1-	923.15296	923.152755	0.22
[a9-BH+Os1-5H]3-	994.50590	994.506200	-0.30
[a6-BH+Os1-4H]2-	1031.68728	1031.687800	-0.50
[w6+Os1+H]2-	1123.69411	1123.694680	-0.51
[a7-BH+Os1-4H]2-	1188.21684	1188.216690	0.13
[a8-BH+Os1-4H]2-	1340.23851	1340.239810	-0.97
[a5-BH+Os1-3H]1-	1749.32253	1749.322510	0.01

$[DNA12+2Os1-10H]^{6-}$ CAD 12V:

5' - T | A | G | T | A | A | T | T | A | C | T | A - 3'

$[a_5\text{-BH}+Os1]^{2-}$ $[a_6\text{-BH}+Os1]^{2-}$ $[a_7\text{-BH}+Os1]^{2-}$ $[a_8\text{-BH}+Os1]^{2-}$ $[a_9\text{-BH}+Os1]^{3-}$

$[w_{11}]^{5-}$ $[w_9]^{4-}$ $[w_8]^{4-}$ $[w_7]^{3-}$ $[w_6]^{2-}$ $[w_5]^{2-}$ $[w_4]^{2-}$ $[w_3]^{1-}$ $[w_2]^{1-}$ $[w_1]^{1-}$

$[w_{11}+2Os1]^{5-}$ $[w_9+Os1+H]^{4-}$ $[w_7]^{4-}$ $[w_6]^{3-}$ $[w_5]^{3-}$ $[w_4]^{3-}$ $[w_3]^{2-}$ $[w_2]^{2-}$

$[w_7+Os1]^{3-}$ $[w_6]^{4-}$

$[w_6\text{-BH}+Os1]^{2-}$

$[w_6+Os1]^{3-}$
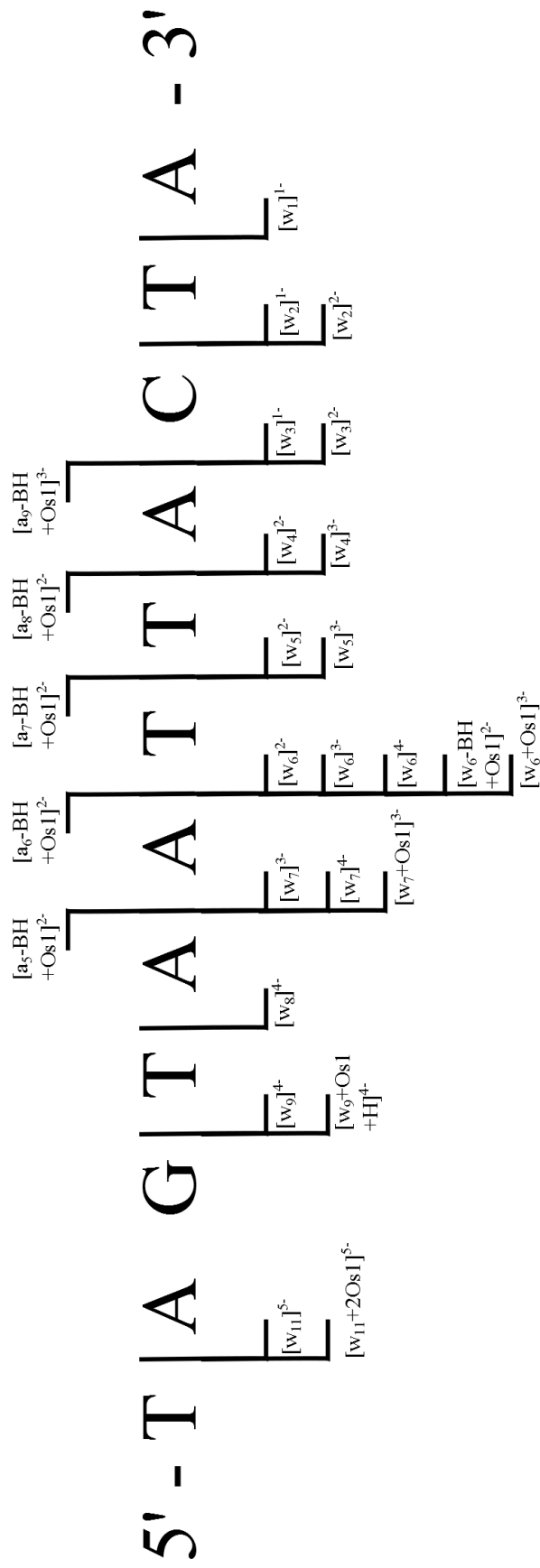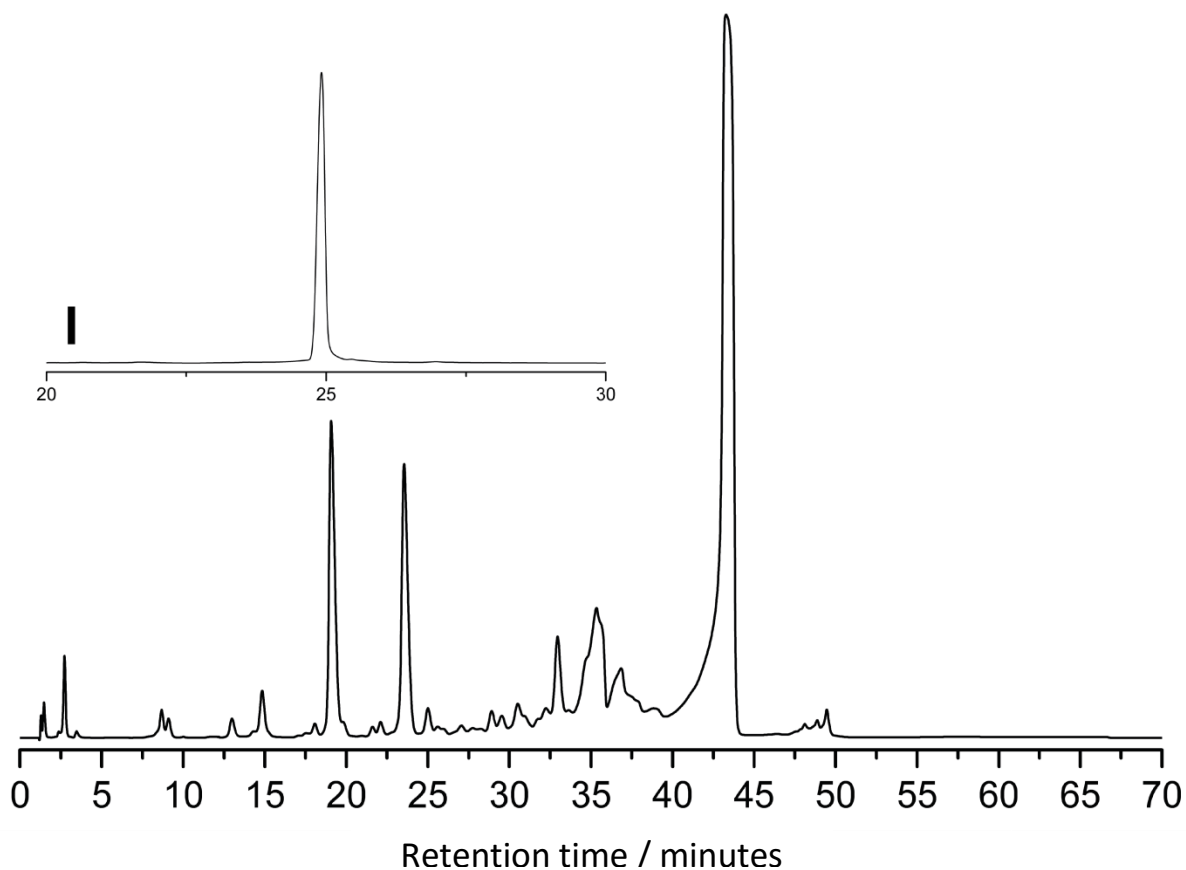
Figure S4.1. HPLC trace of the commercially-synthesised **DNA12** sample, showing multiple shorter oligonucleotides resulting from truncated synthesis and the DNA 12mer (highest peak) which was isolated and then re-analysed by HPLC; the resulting trace is inset, showing only one remaining product, the DNA 12mer (confirmed by UHR-FT-ICR MS, Figure 4.1A in main text).
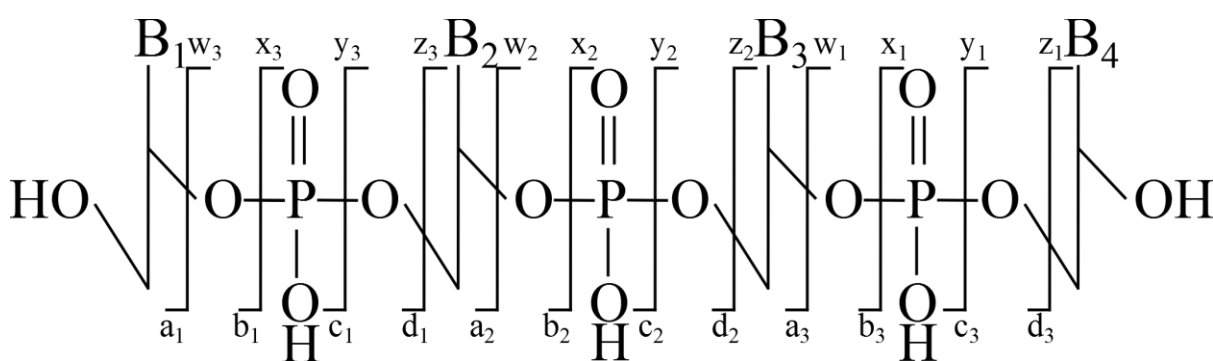
Figure S4.2. The nomenclature devised by McLuckey *et al.*[34] based on the McCloskey *et.al.* system[35] for assignment of oligonucleotide MS/MS fragments.

Chapter 5:

Further study into metallodrug-DNA interactions by tandem mass spectrometry

<u>Introduction:</u>

The study of metallodrug-biomolecules interactions can be greatly facilitated by the use of mass spectrometry and especially tandem mass spectrometry. UHR-FT-ICR MS and MS/MS were effective at studying the wide range of platinum metallodrug-modified peptides in Chapter 3. Unfortunately >90% of common metallodrugs do end up protein-bound within hours of introduction into the body.[1] FT-ICR MS/MS was also shown to be well suited to the study of osmium metallodrug-DNA interactions in Chapter 4. As a result it was logical to expand the study of photoactivatable platinum(IV) complexes into observing and interrogating their interactions with model strands of DNA, which is the focus of Chapter 5 herein.

Cancerous DNA is the idealised target for many anti-cancer metallodrugs, including platinum(II) drugs such as Cisplatin and Oxaliplatin, and photoactivatable platinum(IV) compounds such as FM190 shown in Chapter 3, and below in Scheme 5.1. As mentioned previously by binding to cancerous DNA strands preferentially, metallodrugs can cause deformations in DNA tertiary structure and at the least prevent reading/replication of DNA, or at most induce apoptosis/cell death.
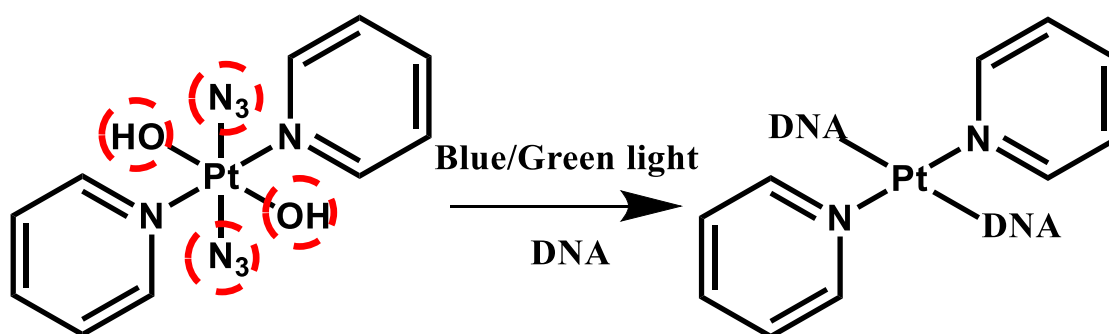
Scheme 5.1 : (Left) *Trans, trans, trans*-[Pt(N$_3$)$_2$(OH)$_2$(py)$_2$] (**FM190/Complex 1**), the photo-activatable prodrug used in this study, along with (Right) an illustrative photo-activation pathway leading to DNA binding (an inter-strand cross-link).

The main target of further investigation into metallodrug-DNA interactions was focused around the photo-induced reaction of FM190 with the model DNA strand studied in Chapter 4; DNA12, and the interrogation of subsequent reaction products by a range of MS/MS techniques within the FT-ICR Mass Spectrometer. With the wide array of platinum modifications observed for the two peptides in Chapter 3 showing release of any of the bound ligands during photo-activation and unexpected oxidation of nearby amino acid residues, reactions with DNA were expected to produce a similarly wide array of species, with various modifications.

The study of DNA by Mass spectrometry can be challenging due to oligonucleotides' extreme affinity for salts such as sodium and potassium, which in solution are key to stabilising their tertiary structure, and their complex fragmentation pathways. As a result of these challenges higher levels of sample clean up and desalting procedures are usually required for study of DNA/RNA species compared to similar sized poly(amino acids). In addition DNA monomer units (nucleotides) have a much higher mass than amino acids, meaning relatively short oligonucleotides quickly achieve large masses and become more difficult to fully characterise with MS and MS/MS. As a comparison poly(amino acids) (such as proteins) of 100+ kDa such as the 147kDa protein complex ADH-tetramer studied by Li *et al.*[2] translating to 1496 amino acids, however currently large oligonucleotides such as those studied by Breuker *et al.* equate to single strands of ~61 nucleotides, but this returns a mass of ~19.5 kDa.[3]

Another challenge to DNA study via MS is the retention of tertiary structure. In solution DNA is most often viewed as a double-helical structure of two coiled oligonucleotide stands , stabilised by inter-base hydrogen bonding interactions, but also by high levels of cations, such as Na^+ and K^+, which interact with the vast number of negatively charged phosphate groups present along the DNA backbone at physiological pH. Of course for ideal analysis by ESI/nESI-MS these salt ions must be removed from solution to enable the detection of purely protonated/deprotonated species. Though this is challenging, it is possible to achieve an acceptable level of salt concentration to allow effective ionisation and observation of purely deprotonated DNA species, as shown in Chapter 4. However the question will continue to remain; is this representative of biological conditions? DNA is most commonly found in the duplex form. Although it has been argued, quite accurately, that the reason most anti-cancer metallodrugs selectively bind to cancerous DNA and kill target cells is not some long-sought after cancer DNA sequence specificity, but simply that the compounds bind most effectively to DNA when it is both uncoiled and single stranded, so that the most preferential binding sites, the nucleobases (usually Guanine nucleobases for many Pt(II) compounds etc.), are fully exposed and able to co-ordinate to the metal centres. Because of this, the study of metallodrug-DNA interactions on single stranded DNA could be considered a reasonable representation of how a given metallodrug would interact with replicating DNA, and since cancerous cells replicate far more than usual cells this would be a common occurrence and available target. However it also suggests that the reactivity (and possible binding sites) of the metallodrug towards DNA could vary greatly between the single stranded, duplex helical, and folded-protein bound duplex DNA structures available within cell nuclei.
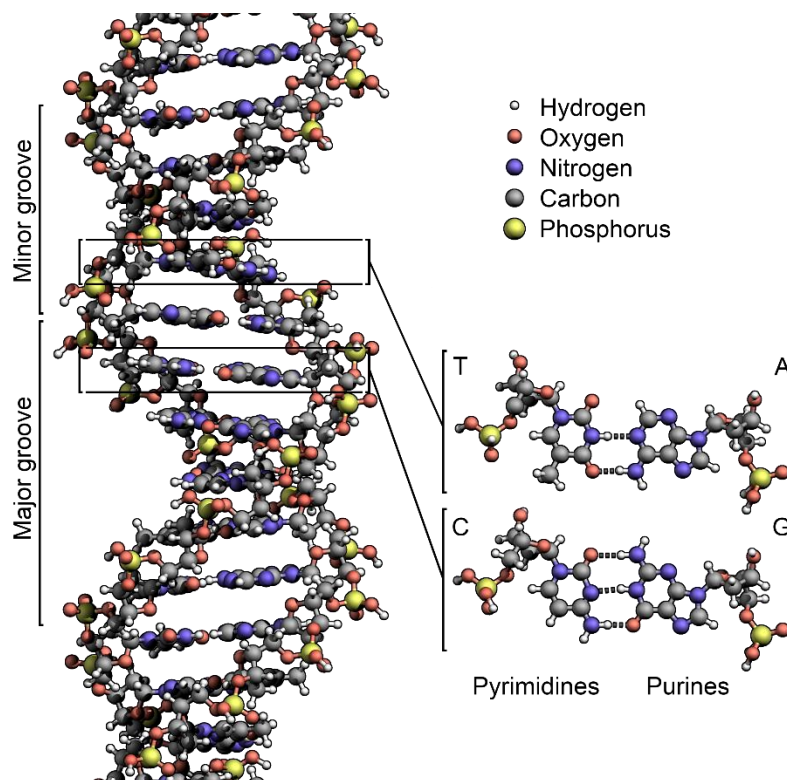
Figure 5.1: Representation of duplex DNA structure in a double-helical structure, showing both atomic detail and major and minor groove features of the structure. Because of its shape the major groove is particularly accessible to other entities. Reproduced from http://www.richardwheeler.net (accessed 12/08/2016).

The retention of DNA duplex structure is similar to that of Native MS of peptides, proteins, and protein complexes, often requiring less than ideal ESI/nESI solvents, but often achievable with extreme care. To this end an investigation of the ability of nESI-MS to retain DNA-DNA interactions and produce gas phase ions for duplex DNA, to be then observed by MS, and interrogated for sequence informative information via MS^n, was conducted using the model DNA strand DNA 12 studied in Chapter 4. The ability to react metallodrug with duplex DNA could inform us about the varying mechanisms of action and DNA bond-disruptions which often accompany binding of metallodrugs to DNA strands. The opportunity to retain duplex DNA structure during nESI-MS could also progress into the ability to study DNA-protein interactions and provide further information on

metallodrug impact on nuclear proteins and their ability to interact with DNA, another mechanism of action to exploit/disrupt in target cancer cells.

Materials and Methods:

DNA12 was isolated and purified as described in the previous Chapter for Os1-DNA reactions, to recapitulate:

The 12-mer DNA oligonucleotide DNA12 (defined here for MS purposes as fully protonated, i.e. each of the 11 phosphates is protonated)

<div align="center">

5'-TAGTAATTACTA-3' (**DNA12**)

</div>

DNA12 (Tm 28°C) was purchased desalted from DNA Technology (Denmark). Further purification was performed by HPLC, using an Agilent 1200 series liquid chromatography system. The oligonucleotide was separated from shorter oligonucleotides (present as minor impurities in the commercial sample) using a PL-SAX ion exchange column (1000 Å pore size; 8 μM particle size; 150 mm length, 4.6 mm diameter; Polymer Laboratories, Amherst, USA); buffer A: 7% acetonitrile + 0.1 M triethylammonium acetate (TEAA), pH 7; buffer B: 7% acetonitrile + 0.1M TEAA + 1 M sodium acetate, pH 7; temp 40°C; flow rate 1.5 ml/min; gradient: 0-5 min 0% B, 5-45 min 0-40% B, 45-55 min 40-100% B, 55-65 min 100% B, 65-75 min 0% B (see Supporting Information Figure S1: HPLC traces of commercial sample and inset; the resulting purified fraction). The solvent was removed on a freeze-drier, and a reverse phase chromatography PLRP-S column (100 Å pore size; 8 μM particle size; 300 mm length, 7.5 mm diameter; Polymer Laboratories, US) was used to remove sodium acetate; buffer A: 50 mM TEAA, pH 7; buffer B: 70% acetonitrile + 50 mM TEAA, pH 7; temp 40oC; flow rate 2 ml/min; gradient: 0-5 min 5% B, 5-35 min 5-100% B, 35-45 min 100% B, 45-45 min 0% B. The solvent was removed again on a freeze-drier.

Reaction mixtures of FM190+DNA12 were created at 1-1 and 0.5-1 (metallodrug:biomolecule) ratios at using aqueous stock solutions of FM190 (250 μM) and DNA12 (1 mM). samples were irradiated

with 463 nm blue visible light using an LED lamp (23 mW/cm-2) for set time intervals and then diluted with pure water prior to nESI-FT-ICR MS analysis, as before the use of acids/organic solvents/disrupting buffers was strictly avoided, as was any activating light (white/blue/green) before and during nESI-FT-ICR MS analysis. All sample preparation and analysis was conducted in darkness and/or very low levels of red light to avoid unwanted activation of the photoactivatable Pt(IV) metallodrug.

All FM190+DNA12 samples were analysed by nESI-FT-ICR MS operated in the negative mode, in a similar fashion to the osmiated DNA in Chapter 4. CAD, IRMPD, and EDD MS/MS techniques were applied to observed adducts, selected informative results are shown and discussed below.

Native (duplex) DNA experiments are detailed in the text below and were operated in both positive and negative mode nESI, from buffered aqueous conditions.

Photoactivatable Pt(IV) + DNA12 reactions:

FM190 was shown to react with model peptides in Chapter 2 to an observable degree when activated with 463 nm blue visible light for just 1-2 hours. This photo-activation-induced binding was much faster than previous Pt(II) agents such as Cisplatin which rely on ~24 hour solution based displacement reactions in order to create active aqua-Pt(II) species and allow subsequent Pt complex-biomolecule binding.[4–6] For the FM190+DNA12 reactions required solutions of the metallodrug+DNA were irradiated as for the peptides as previously, but for shorter periods of time i.e. t = 30, 60, 90, 120 minutes, longer irradiation times were attempted, but as discussed below were not required. The photo-activation produced an array of products, many of which resemble those previously observed during the reactions with peptides in Chapter 3. The mass spectra for each irradiation time are shown below in Figure 5.2, with different species marked according to the key below:
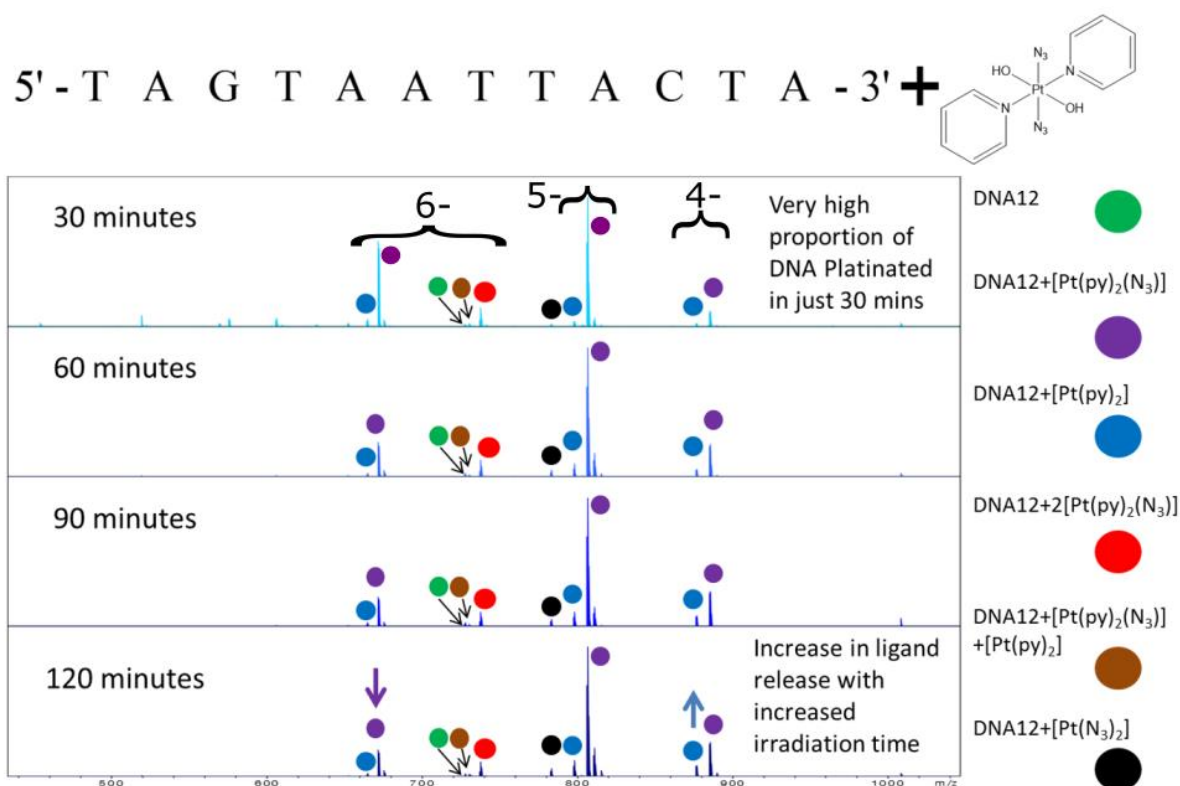
Figure 5.2: nESI-FT-ICR MS spectra of the DNA 12-mer +FM190 metallodrug (both shown inset) mixtures after varying irradiation times with blue visible light (463 nm). Each different product is marked with a coloured circle, with key shown on the right.
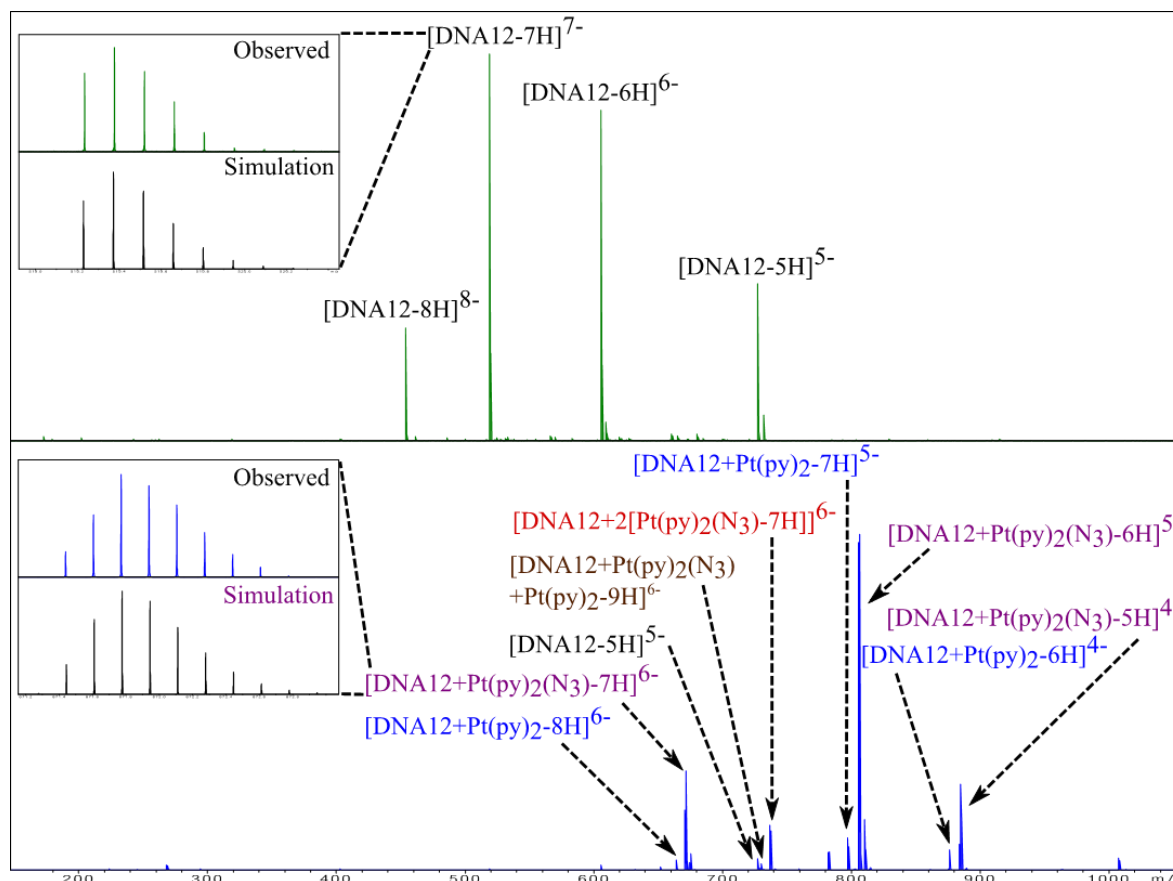
Figure 5.3: nESI-FT ICR MS of the DNA 12-mer+FM190 mixture after 60 minutes irradiation with blue visible light (463 nm). Each product is marked accordingly. Isotopic simulations presented in the inset show the effect of the platinum metal centre on the overall DNA isotopic pattern. The difference is clear, however less drastic than for the osmium metallodrug shown in Chapter 3.

Immediately surprising was the extremely fast rate of reaction observed, with only a very small proportion of the unmodified DNA12 species observable after just 30 minutes. Indeed the FM190 compound was much more reactive towards DNA species than the peptides studied in Chapter 3, reacting to near completion within 30 minutes, and only minor changes from then on, in contrast to the hours for peptides species observed before. Similarly to the [Lys]³-Bombesin peptide all of the

275

Pt-bound reaction products had retained the trans-pyridine groups on the platinum complex. This uniform retention of pyridine functional groups, if in a trans configuration, is in direct difference to the ineffective trans-platin compound, whose NH_3 amine groups bound to the Pt centre hindered DNA interaction due to a trans configuration, leading to the possibility of either a) the pyridine groups are able to re-orientate during the photo-activation/biomolecule binding steps and move out of the *trans, trans, trans* configuration, or b) the photo-activated compound retains its stereochemistry and is still able to bind to oligonucleotides, which could cause different distortions of tertiary structure than previous Pt(II) compounds such as Cis- and trans-platin.

Also surprising was the absolute lack of biomolecule oxidation. Previously in Chapter 3, unexpected but clearly observable levels of oxidation were observed in both platinum-modified and unmodified peptide species, which were then studied using tandem mass spectrometry and found to be oxidised at the tryptophan and methionine amino acid residues. However both the platinum-modified and unmodified DNA12 species observed here were present without any oxidation products, despite the ability for nucleobases to be oxidised in the presence of either hydroxide radicals[7] or singlet oxygen.[8] The lack of oxidation products shows a clear indication that, as with the peptide species above, the activation and resulting platinated reaction products are influenced by the identity, structure and functional groups of nearby biomolecules and in this case the DNA12 species did not enable the stabilisation/reactions leading to oxidation of oligonucleotides through release of hydroxide radicals.

As mentioned above the reaction between DNA12 and FM190 occurred very quickly, with a near complete platination of DNA12 species within just 30 minutes irradiation with 463 nm blue visible light. Also, similar to the two peptides studied in Chapter 3, after a certain time point longer irradiations simply caused further activation of the platinum complex and release of further ligands, as seen in Figure 5.2 above, the DNA12+Pt(py)$_2$(N$_3$) species decreased in relative intensity with further irradiation while the DNA12+Pt(py)$_2$ species increased, due to further activation of the Pt-azide bond and further release of Azide species. Despite these differences, fragility post-ionisation,

and extreme light sensitivity, as with the peptides, each reaction product was effectively isolated and fragmented using a range of dissociation techniques in order to characterise the platinum complex modifications and their location within the DNA strand.

MS/MS of platinated reaction products:

Each observable platinated-DNA reaction product was isolated and fragmented using CAD and EDD MS/MS, in a similar fashion to the osmiated DNA in Chapter 4. General CAD collision energies/voltages were 2-12V, EDD was achieved through the use of a 1.5A heater electron-dispensing cathode emitting 20.6 eV electrons for 1.0 seconds using a lens held at 18V.

Unfortunately, as observed for the platinated peptides in Chapter 3, CAD MS/MS of the singly platinated species caused dissociation of platinum-bound ligands, followed by dissociation of the platinum metal centre from the DNA strand. The main fragmentation channels seemed to be loss of the pyridine ligands (if available) forming high intensity peaks, followed by subsequent loss of other Pt-bound ligands, and eventually the metal. It is desirable to use CAD/IRMPD for fragmentation of oligonucleotides, both for the ease of operation and the increased sequence coverage they can provide for these species via MS/MS, as recently shown by Shaw *et al.*,[9] but also in Chapter 4. Regardless EDD MS/MS, shown in Chapter 4 as being a viable method for retention of metal complex-based modifications, can also be adopted for fragmentation of the platinated DNA species. Successful EDD MS/MS can require very fine tuning of both the electron cathode parameters, but also the ICR cell parameters in order to optimise the overlap of parent ion with the incoming electron beam in a way that produces the maximum backbone fragmentation of the target species and enough fragments to be detected. The CAD MS/MS of the $[DNA12+Pt(py)_2(N_3)-6H]^{5-}$ species is shown below in Figure 5.4:
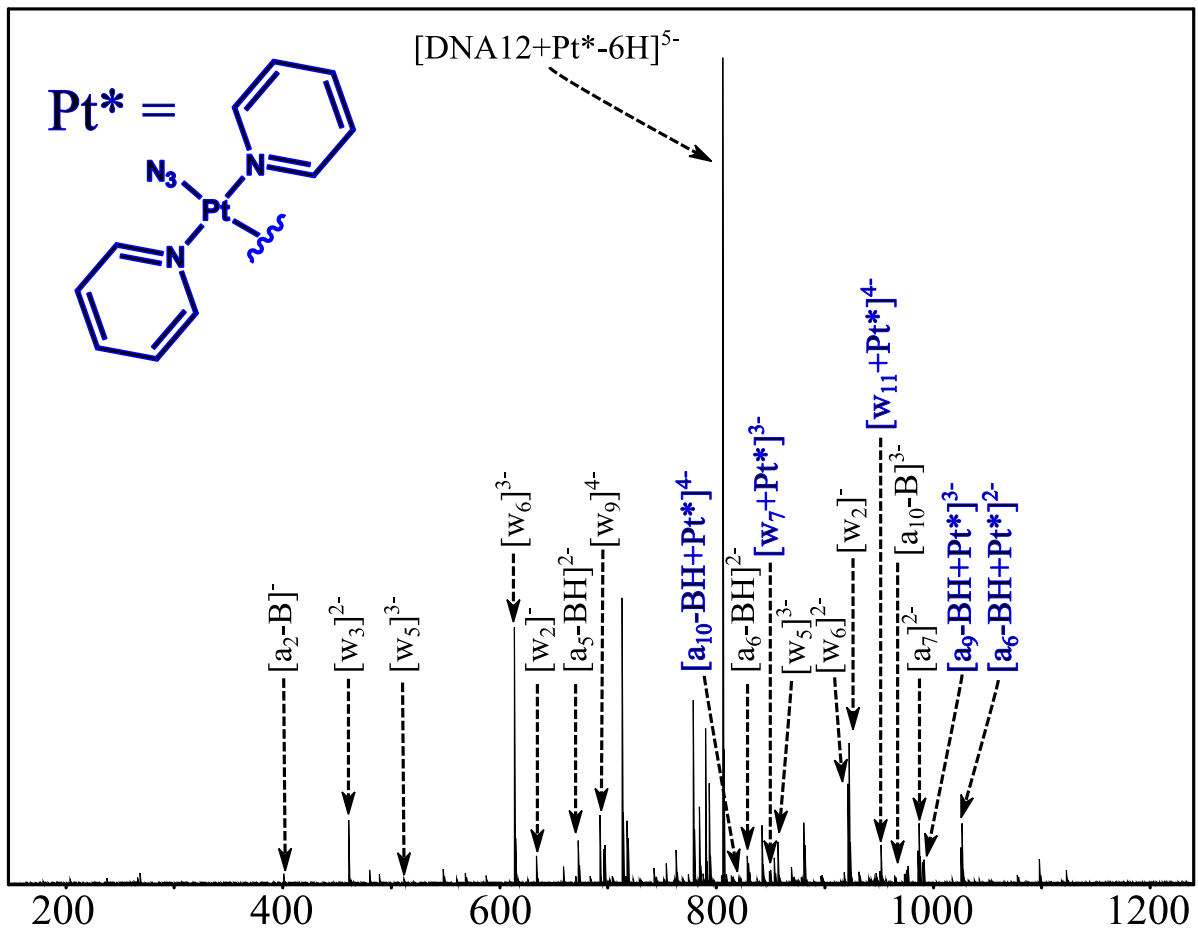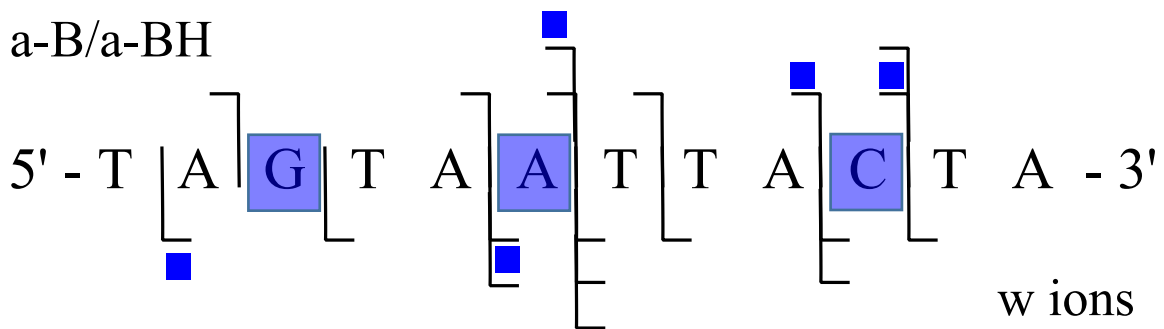
Figure 5.4: nESI-FT-ICR CAD MS/MS of the [DNA12+Pt(py)$_2$(N$_3$)-6H]$^{5-}$ species (top). The resulting a and w ions assigned are summarised in the fragmentation map (bottom) shaded squares indicate fragments containing a metal modification. Highlighted regions indicate possible binding site.

From the analysis of CAD MS/MS spectrum the Pt(py)$_2$(N$_3$) modification could have been bound to any or all of three different positions, the G^3, A^6, and/or C^{10}. Due to the CAD-induced loss of the modification secondary fragmentation afterwards complicated the spectrum and analysis, causing

conflicting fragments, in a similar fashion to the osmiated DNA studied in Chapter 4. As a result EDD

MS/MS was employed to provide complementary fragmentation data, as shown in Figure 5.5,

below:
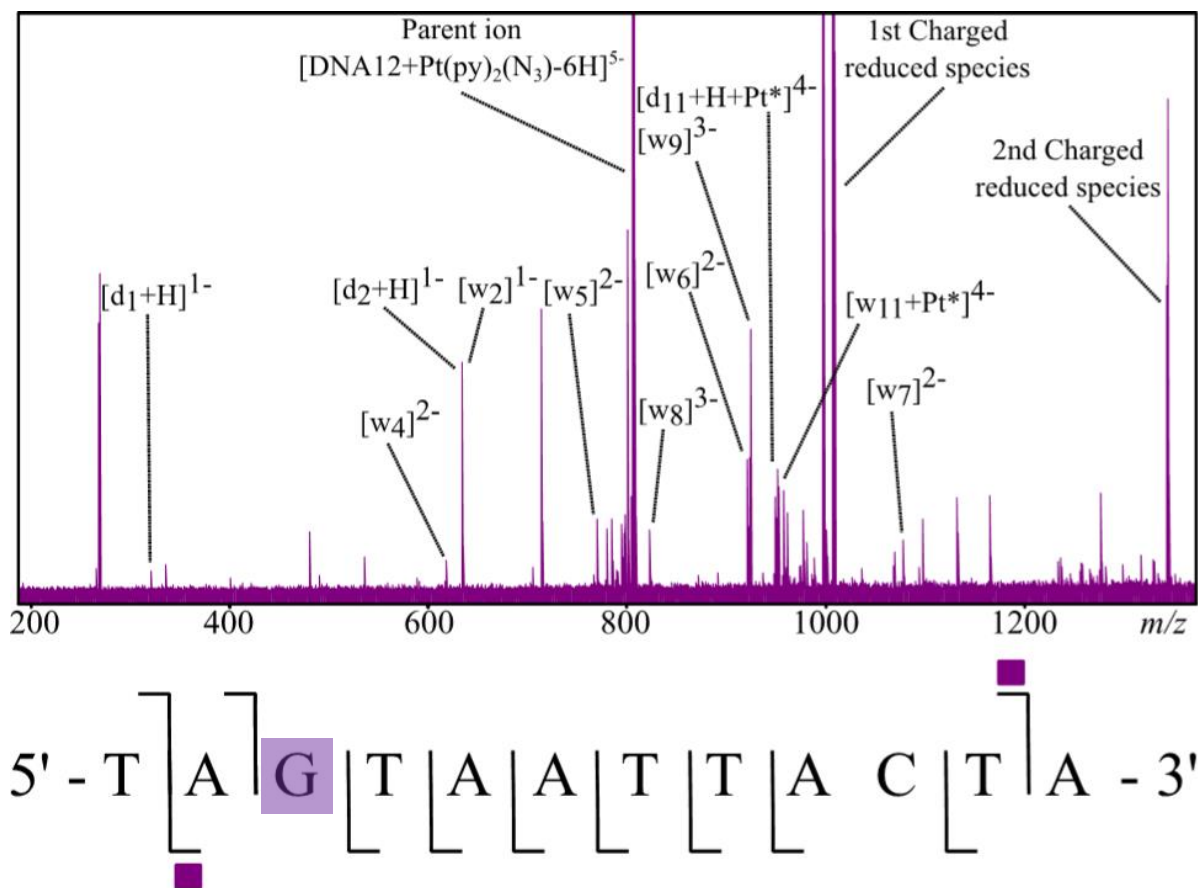
Figure 5.5: nESI-FT-ICR EDD MS/MS of the [DNA12+Pt(py)$_2$(N$_3$)-6H]$^{5-}$ species (top) using 20.6 eV

electrons. The resulting d and w ions assigned are summarised in the fragmentation map (bottom)

shaded squares indicate fragments containing a metal modification.

Analysis of the EDD MS/MS spectrum for the [DNA12+Pt(py)$_2$(N$_3$)-6H]$^{5-}$ species clearly shows that the

platinum complex modification seemed to be binding to the guanine[3] residue of the DNA 12-mer.

The binding of the platinum complex to the guanine residue is in line with various Pt(II) based

metallodrugs such as Cisplatin and Transplatin, which have both been shown to have a preference

for binding at the exposed guanine nucleobases in DNA.[9–11] The interesting observation particular to

the products observed in the nESI-FT-ICR MS in Figure 5.3, above, was the two di-platinated species

([DNA12+2(Pt(py)$_2$(N$_3$))-9H]$^{6-}$ and [DNA12+Pt(py)$_2$(N$_3$)+Pt(py)2-9H]$^{6-}$), due to the singular guanine nucleobase, the observation of diplatinated species suggested that the platinum complex modifications could bind to residues other than the guanine sites, which could be a distinct difference in behaviour compared to many Pt(II) based compounds. EDD MS/MS of the diplatinated [DNA12+2(Pt(py)$_2$(N$_3$))-7H]$^{6-}$ species is shown below in Figure 5.6:
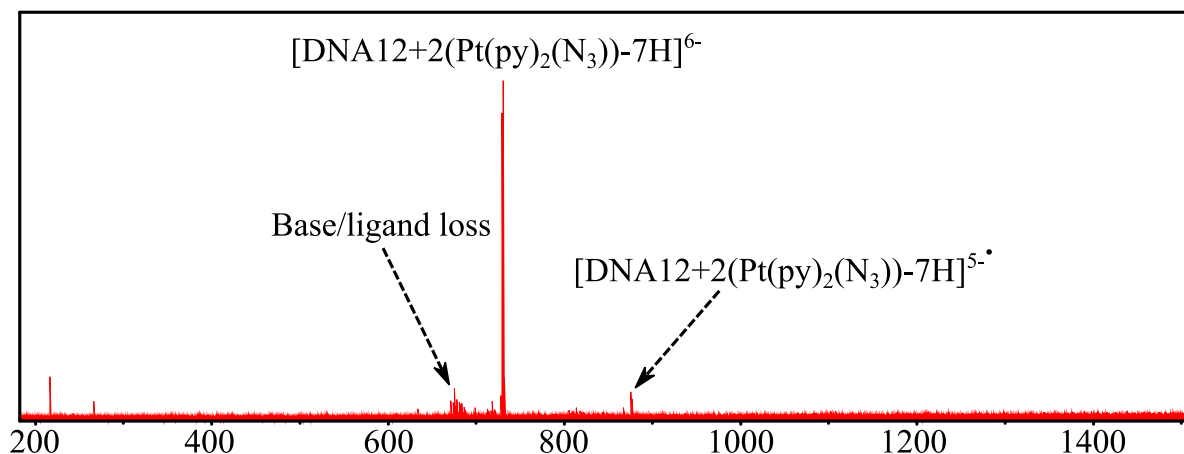
Figure 5.6: EDD MS/MS nESI-FT-ICR MS spectra of the diplatinated [DNA12+2(Pt(py)$_2$(N$_3$))-7H]$^{6-}$ species. Charge reduction, ligand and base loss was observed, but no sequence informative fragments could be assigned. Most likely due to the platinum centre disruption of normal EDD MS/MS.

Unfortunately the EDD MS/MS proved ineffective in providing significant sequence informative fragments of the di-Platinated oligonucleotide. This could be due to the effect of the platinum metal centres influencing the electron-based dissociation. Peaks were observed for the charge-reduced species, showing that electrons were captured effectively by the ions. Peaks were also observed showing base loss and ligand loss, but no significant backbone fragmentation was observed despite this. The platinum metal centres could be particularly effective at stabilising the electron deficient radical site created by EDD MS/MS, either via the metal centre, or via loss of a platinum-bound ligand.

CAD MS/MS was also attempted on the diplatinated $[DNA12+2(Pt(py)_2(N_3))-9H]^{6-}$ species, shown below in Figure 5.7. The CAD MS/MS of the diplatinated oligonucleotide showed an array of platinated and non-platinated fragments, which is in stark contrast to the CAD MS/MS of the mono-platinated reaction product shown above. The assignable fragments are included in the Figure below and suggest binding of the metal complexes both at the guanine site (as seen above) and at either the C^{10} or T^{11} sites. Unfortunately CAD MS/MS also causes dissociation of Pt-bound ligands and entire Pt-based modifications, which caused conflicting fragments. However when using the unmodified fragments as well as the platinated fragments it could be suggested that the second $Pt(py)_2(N_3)$ modification is binding at the C^{10} site, in a similar fashion to the osmium drug studied in Chapter 4. However this is unlikely based on previous reports of Pt(II) species and their preferential binding locations in oligonucleotides and larger DNA construct, and so would require more study and experiment to confirm the binding location, hopefully using dissociations which either fully retain the Pt modifications or which can cleave the C^{10}-T^{11} bond to allow confident single-nucleotide assignment of the binding location. Analysis of both EDD and CAD MS/MS spectra of other observed reaction products is ongoing. Digestion of the oligonucleotides into smaller, more approachable species is possible, but would likely cause unwanted side reactions/effect towards the metal-based modifications causing dissociation or possible chemical transformation and negatively affect experimental validity.
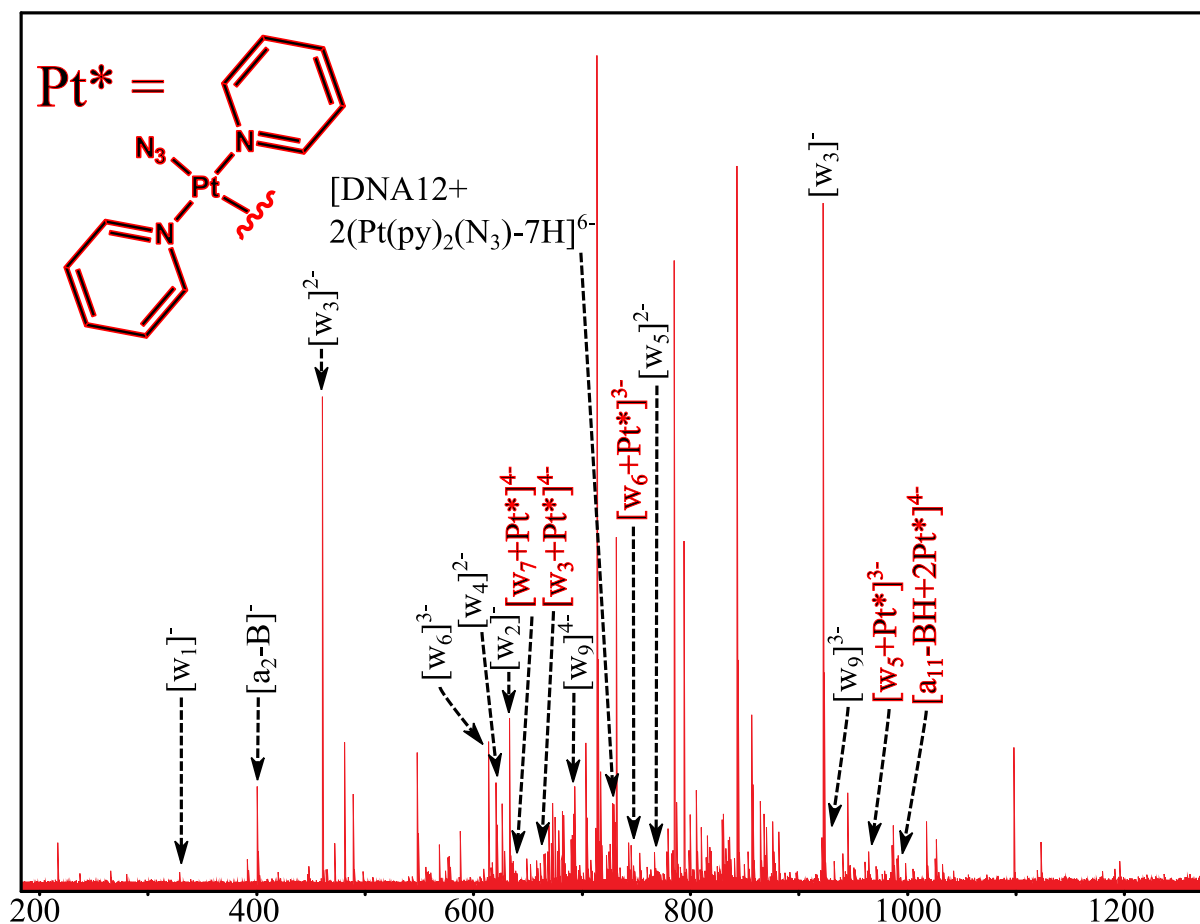
Figure 5.7: CAD MS/MS nESI-FT-ICR MS spectra of the diplatinated [DNA12+2(Pt(py)$_2$(N$_3$))-9H]$^{6-}$ species (top) with corresponding fragmentation map summarising assignable fragments (bottom). Coloured markers indicate fragments retaining a platinum-based modification, Pt* = Pt(py)$_2$(N$_3$)

Moving towards Duplex DNA-MS:

Although the analysis and characterisation of metallodrug-DNA interactions using isolated DNA strands was effective and provided unique insights into the mechanisms of action for both the osmium and photoactivatable platinum compounds studied herein, all studies involved reactions with single stranded DNA, which was then observed in the resulting nESI-FT ICR MS spectrum. DNA usually is present within the nucleus of cells in a duplex helical form, with two complementary strands coiled around each other via hydrogen bonds between nucleotide base groups, thus the reactive bases can often be shielded from the metallodrug and possible binding to some degree. In order to study the effects of metallodrugs on representative DNA structures and provide insights into mechanisms which were similar to in vivo activity; the possibility of studying duplex DNA by nESI-FT-ICR MS was investigated.

For standard peptides and proteins Native MS (described in Chapter 1: MS introduction) attempts to stabilise protein structure during ionisation, transport and detection within the mass spectrometer, attempting to retain structural features through the transition between solution phase solubilised protein to detectable gas phase ions. This is usually achieved by avoiding denaturing conditions such as organic solvents, acid/bases, and low desolvation temperatures etc. during ESI. Also the use of aqueous buffers instead of pure solvents is usually required, which are not optimal for ESI-MS of proteins, and can thus hinder effective ionisation, high concentration Ammonium Acetate (AA) (10-200 mM) has become a particularly common buffer choice for Native MS as it is reasonably compatible with ESI and does not contain high amounts of additional sodium/potassium which would drastically complicate resulting spectra and hinder detection/ MS/MS. As mentioned previously Native MS usually produces lower intensity protein peaks, large amounts of buffer adducts, and charge state ions than that of aqueous ESI-MS of the same analyte, again reducing detected intensity of ions and most importantly reducing MS/MS efficiency drastically. It is for these,

and other reasons, that Native MS is particularly challenging, firstly to achieve, and secondly to work with in subsequent experiments/ MS/MS.

Native MS of DNA structures is very different to that of poly(amino acids) such as proteins. Protein structure can (for the most part) be stabilised by non-ideal buffers such as AA or ammonium bicarbonate (ABC), however DNA strands are mainly stabilised by salt ions such as Na^+, K^+, and mostly Mg^{2+} which are usually disastrous additives for ESI-MS of biomolecules. Salt adduction to biomolecules is one of the key issues and hindrances when attempting to achieve effective ESI-MS. Salt adduction spreads out a single analyte distribution over many different *m/z's* and detection channels, effectively "diluting" the detectable signal and reducing sensitivity, also salt adduction can help protein structure survive during MS/MS and hinder fragmentation, or at the very least alter the fragmentation channels available. As a result of the challenges salt adduction to biomolecules induce, extensive sample preparation to remove salt species from solution and samples is ubiquitous throughout the biological MS field. DNA/RNA in particular are extremely challenging to study if not extensively purified and desalted to eliminate such problems, as observed in the previous sets of results the DNA 10 and 12-mers studies were purified and desalted multiple times, by various chromatography and size-exclusion techniques in order to produce the very clear and nearly salt-free nESI-FT-ICR MS spectra shown herein.

To begin the investigation into Native MS of DNA by FT-ICR MS, a sample of the DNA 12-mer (used during the previous studies) was analysed both in aqueous conditions (as previously studied as a control) and using high (with respect to MS analysis) concentration of ammonium acetate (200 mM), as would be expected for particularly challenging proteins for Native MS. The samples were analysed using nESI-FT-ICR MS as previously, however Native MS samples were operated using extremely low ESI source temperatures, so as not to induce unfolding/dissociation during heat-aided desolvation. The usual source temperature for nESI would be 180-200°C, for native MS samples this was reduced to 20-40°C. Usually Native MS samples have been reported using source temperatures ~80-100°C,

however as DNA is particularly sensitive to thermal transitions a lower temperature was chosen. The results of the different solutions for nESI-FT-ICR MS of the DNA 12-mer are shown below in Figure 5.8.
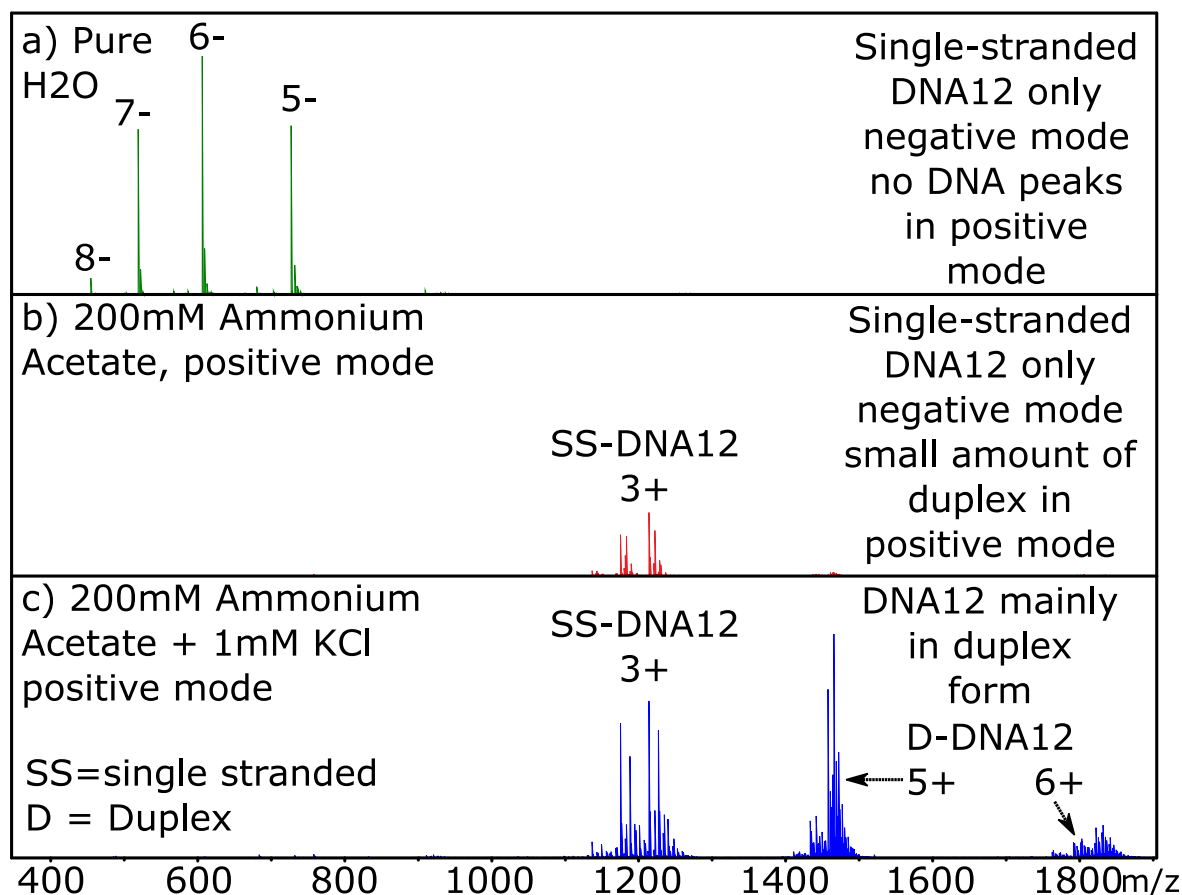
Figure 5.8: nESI-FT-ICR MS spectra of the DNA12 (TAGTAATTACTA) in water (a, top) 200 mM ammonium acetate (b, middle), and a 200mM ammonium acetate + 1mM KCl solution (c, bottom). Though particularly unusual, the addition of KCl to the buffer solution proved effective in stabilising the DNA duplex, but required the species to be detected in positive nESI mode, instead of negative mode as for the previous samples.

Unfortunately even at 200 mM concentrations, the AA buffer solution was unable to effectively stabilise the structure of the DNA duplex and produced a spectrum containing only the single stranded DNA species (albeit with more buffer adducts than previously). Similar challenges in studying the structural aspects of DNA were recently observed by Gabelica *et al.* when studying DNA

quadruplexes.[12] DNA quadruplexes rely on a coiled duplex structure around central cations such as

potassium (K$^+$), but without these they fall apart. It was shown by Gabellica *et al.* that observation of

the quadruplexes was challenging by MS, but by the unusual tactic of actually adding salt species

(KCl) to the ESI solution they were able to observe the fully potassiated ([DNA+2K]$^+$ and

[DNA+3K]$^+$)species, achieving native-like representation of the tertiary structure.
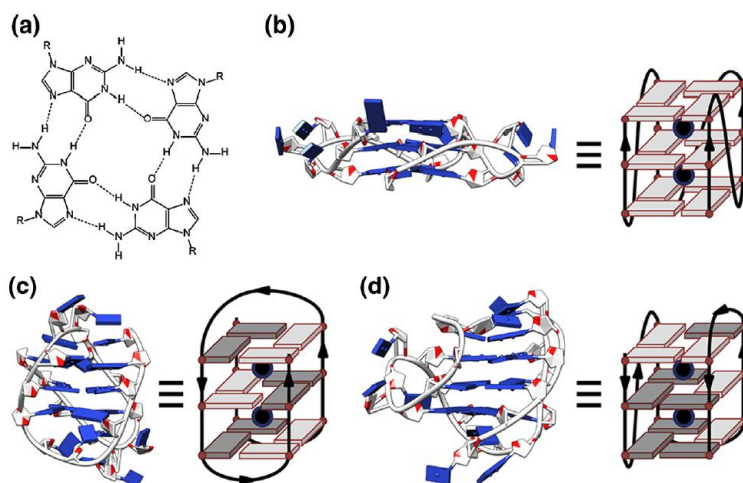
Figure 1. **(a)** A guanine quartet (G-quartet). **(b)** X-ray crystal structure of the parallel G-quadruplex form of d(A(GGGTTA)$_3$GGG) in K$^+$ conditions [12] (PDB ID: 1KF1). **(c)** NMR structure of the antiparallel G-quadruplex form of d(A(GGGTTA)$_3$GGG) in Na$^+$ solution [13] (PDB ID: 143D). **(d)** NMR structure of the hybrid G-quadruplex form of d(TT(GGGTTA)$_3$GGGA) in K$^+$ solution [15] (PDB ID: 2GKU). The schemes underline strand directionality, and base-sugar conformations (light grey for *anti*, dark grey for *syn*)

Figure 5.9: Structural features of DNA G-quadruplexes shown by Gabelica *et al.* Reproduced from

Gabelica *et al.*[12]
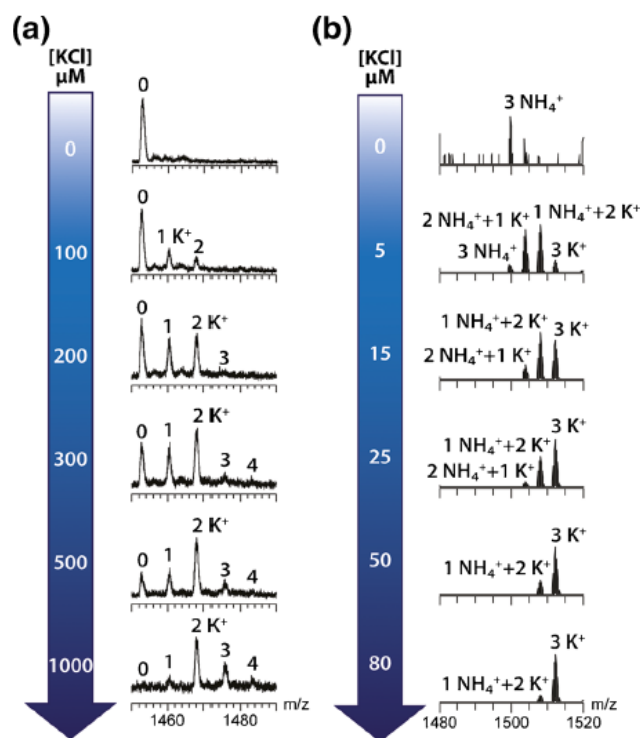
Figure 6. KCl titration of **(a)** 5 μM 23TAG and **(b)** 20 μM TG$_4$T single strand into 100 mM TMAA. The annotations indicate the number of adducts on the 5- charge state. In titration **(a)**, the single strand was allowed to react for 4 h at room temperature before the mass spectra were recorded using the LCT mass spectrometer. In titration **(b)**, the solutions were prepared with 100 μM in TG$_4$T single strand for 24 h before being diluted to 20 μM, and the mass spectra were recorded on the Exactive mass spectrometer

Figure 5.10: Results of the potassium-titration experiment performed by Gabelica *et al.* to show the uptake of K+ ions into the DNA quadruplex-G structure, note the particular preference for 2K$^+$ ions in (a) and 3K$^+$ ions in (b) due to the different tertiary structures present. Reproduced from Gabelica *et al.*[12]

Addition of KCl to ESI samples is usually particularly unfavourable, for reasons discussed above, though for the DNA quadruplexes was effective. However the DNA quadruplexes could accommodate up to 2 potassium ions only, putting a hard upper limit on the possible adducts observed, which is particularly convenient compared to if the same tactic would be used for say a protein. Unfortunately the DNA 12-mer studied within this thesis does not have this low upper limit to salt/buffer adduction, containing 12 phosphate groups along its backbone, and many bases with

accommodating lone pairs for cation adduction, the DNA 12-mer is very capable of many, many buffer and salt adduct products. Despite this the 200 mM ammonium acetate DNA 12-mer sample was spiked with 1 mM KCl, as a "high" concentration from KCl titration experiments performed by Gabelica *et al.* and shown above in Figure 5.10a.[12] The resulting negative-mode nESI-FT-ICR MS spectrum did not show the desired duplex DNA species, nor did it show any single stranded DNA species, believed to be due to the disruption of KCl to the nESI process and/or blocking of the narrow nESI tip orifice by KCl aggregates. However despite this the nESI and transmission/detection voltages were reversed and the sample was analysed using positive mode nESI-FT-ICR MS to test for heavily potassiated DNA species, which may have become cations in solution due to the high concentrations of cations available. The resulting nESI-FT-ICR MS spectrum produced clearly observable single stranded and even duplex DNA species. The resulting "Native DNA MS" spectrum is shown above in Figure 5.8c.

Unfortunately although observable, the DNA species had expectedly acquired a significant amount of buffer and salt adducts. Careful simulations of each possible combination of salt adduct ($[DNA12+(H)_w(Na)_x+(K)_y+(AA)_z]^{(w+x+y)+}$) showed that indeed there was extensive potassium and sodium salt adduction, but also that the DNA duplex dimer was also observable. Again simulations for the Duplex DNA species were required to distinguish correct isotopic patterns from the various adduct peaks, using $[DNA12+(H)_w(Na)_x+(K)_y+(AA)_z]^{(w+x+y)+}$. Surprisingly the fully protonated duplex-DNA species was also observed, without potassium/salt adduction, despite the inability to observe the duplex-DNA 12-mer in aqueous conditions without additional salt, leading to the conclusion that adducted cations such as Na^+ or K^+ could move/dissociate during the ESI process, or could be displaced by the high concentration of H^+ induced during desolvation etc. Regardless the species were observed, and assigned (see Figure 5.11), and showing that the observation of duplex DNA is possible by nESI-FT-ICR MS.

The next stage of characterisation after observation, as for all biomolecular MS, is tandem MS. Unfortunately as observed above the salt adducts and the protonated duplex DNA species are very close in mass and not easily separable, however they need not be separated unless overlapping with other reaction products (as would be the case for metallodrug reactions) or unless observing/investigating the dissociation differences between the various adducts and the protonated species. The entire observable envelope for the protonated and various salt-adducted duplex DNA was isolated in the quadrupole region of the FT-ICR, and then fragmented using low-energy CAD MS/MS, fragments were then accumulated in the hexapole region and then transferred to the ICR cell for detection. The resulting MS/MS spectrum showed the DNA duplex dissociating easily into single stranded DNA 12-mer ions as observed before. The lowest energy fragmentation channel appeared to be (predictably) the dissociation of the intra-strand hydrogen bonds to form individual oligonucleotide ions post CAD.

The individual chains could then be re-isolated, fragmented, and characterised using in-cell isolations such as Multi-CHEF and in-cell dissociation techniques such as IRMPD and ECD MS/MS. In this way any modification of the DNA strand, say in future reactions with metallodrugs, would be elucidated through the isolation of individual species and the choice of dissociation techniques could be optimised in order to retain the modifications (e.g. ECD would be required for the osmium and platinum species studied in previous chapters). The Species produced by CAD MS/MS were then dissociated using ECD MS/MS for MS^3 analysis of the duplex DNA. The species were not re-isolated during this experiment using Multi-CHEF, though this is easily achievable, this initial proof-of-concept experiment was to show the CAD-ECD MS^3 analysis was possible and could provide sequence-informative data. As mentioned in Chapter 4, electron-based dissociation of oligonucleotides can be achieved by Electron Detachment Dissociation (EDD) MS/MS, however this is a negative mode-exclusive technique, and so ECD was adopted for characterisation in the second stage of MS/MS here. Surprisingly ECD MS/MS of the DNA species was achieved by tuning of the 1.5 A ECD cathode to emit 1.8 eV electron for 0.1 seconds via a lens voltage of 15V. The resulting MS^3

spectrum is shown below in the summarising Figure 5.11, along with the **d**, **w**, and **z** ions identified from the ECD MS[3].
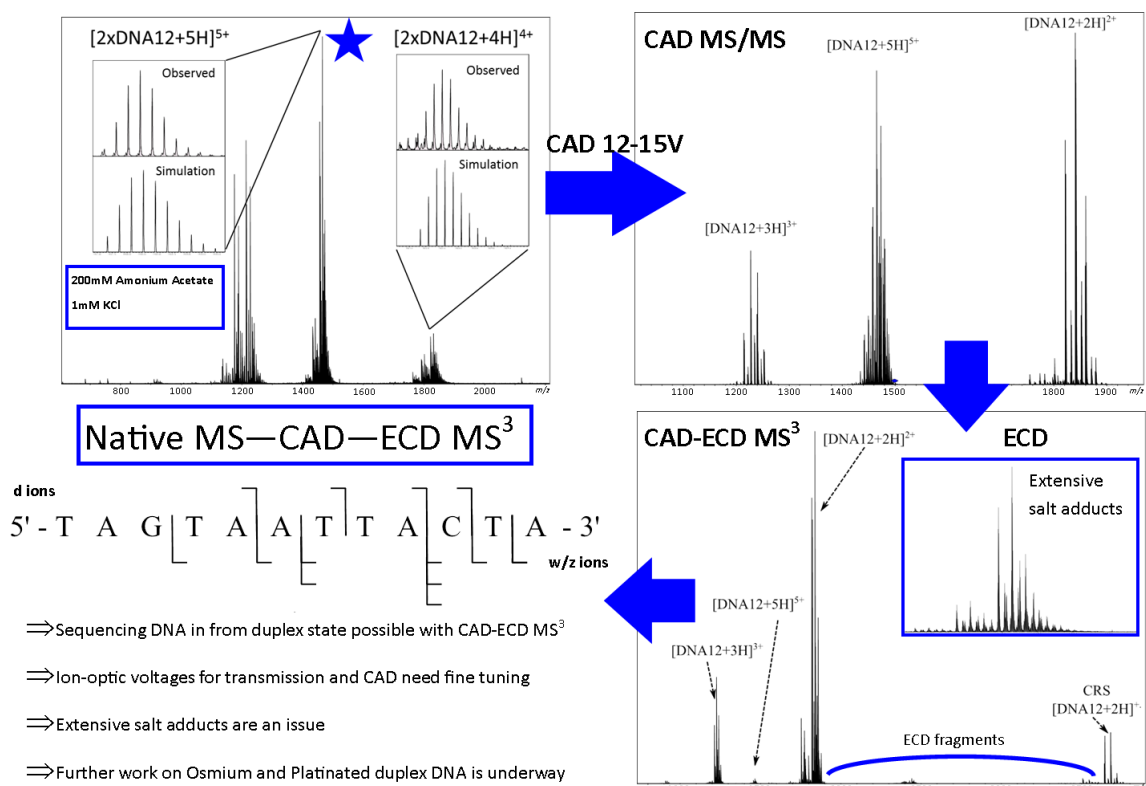
Figure 5.11: Summarising figure for the native Duplex-DNA nESI-FT-ICR MS3 experiments conducted. Clockwise from top left: nESI-mass spectrum of the DNA12+200mM AA+ 1mM KCl aqueous solution, CAD MS/MS of the quadrupole-isolated DNA12-Duplex 5+ charge state, CAD-ECD MS[3] of the isolated duplex species, fragmentation map showing the SS-DNA peaks observed in the CAD-ECD MS[3] experiments, along with comments on the results obtained.

The observation of **z** ions as well as **d** and **w** ions (usually observed in EDD) was slightly unusual, however the dissociation mechanisms of ECD and EDD differ greatly, not least because of the addition of electrons compared to the removal of electrons as the initial reaction step (respectively), so the similarities could be considered more surprising than the slight differences. Regardless CAD-ECD-MS[3] of positively charged salt-stabilised DNA duplexes in the gas phase was possible using nESI-FT-ICR MS, and should be the next target in studying metallodrug-DNA interactions in order to

provide more representative scenarios for targets metallodrugs should encounter during their mechanism of action.

To that end preliminary experiments into the binding of the osmium drug studied in Chapter 4 (Os1) with the Native DNA sample were attempted, using the same 3 hour incubation time and conditions as those mentioned previously, also longer reaction times of 12 hours were attempted. Unfortunately the osmium drug was shown to not bind to the duplex DNA within the timescale tested. This is believed either to be due to the osmium drug binding to the nucleobase, which is tightly involved in intra-strand hydrogen bond when in the duplex (as mentioned previously) or due to the 1 mM KCl solution used, as the Os1 drug relies on the dissociation of the Chlorido ligand in order to form the active complex and bind to target biomolecules, addition of KCl will shift the equilibrium of this reaction towards the inert Chlorido-bound Os1-Cl complex and prevent binding (according to le Chatelier's principle). The first possibility, intra-stand hydrogen bonding, is particularly relevant as this is believed to be what "protects" "normal" DNA from attack by metallodrugs, for the most part, during treatment, as it is the cancerous DNA which is more often replicating, and thus exposing its nucleobases to the possibility of binding, while normal, healthy cells are not so exposed, this is also believed to be the basis of the selectivity of most classical metallodrugs to bind to cancerous cellular DNA instead of the DNA of healthy cells. The equilibrium disruption is unfortunate, but may be addressed in the future by using non-Chloride containing ionic potassium salts – such as KOH. It is also worth noting that the same equilibrium issue would not be expected for compounds such as the platinum(IV) drug studied in Chapter 3 – FM190, as it contains no Chlorido ligand functionality, however, the extensive salt adducts would cause issues with the small m/z differences in reaction products observed (e.g. $[Pt(py)_2(N_3)]^+$ vs. $[Pt(py)_2(OH)]^+$ modifications), and this would have to be addressed. These reactions were not fully tested/optimised during the timespan of this thesis, but lay the basis for such investigations in the near future, which are already underway.

References:

(1)   Centerwall, C.; Tacka, K. *Mol. Pharmacol.* **2006**, 348–355.

(2)   Li, H.; Wongkongkathep, P.; Van Orden, S. L.; Ogorzalek Loo, R. R.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 2060–2068.

(3)   Taucher, M.; Breuker, K. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 918–929.

(4)   Li, H.; Zhao, Y.; Phillips, H. I. a; Qi, Y.; Lin, T.-Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 5369–5376.

(5)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, 9507–9515.

(6)   Li, H.; Snelling, J. R.; Barrow, M. P.; Scrivens, J. H.; Sadler, P. J.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1217–1227.

(7)   Nowicka, A. M.; Kowalczyk, A.; Sek, S.; Stojek, Z. **2013**.

(8)   Armstrong, D. *Oxidative Stress in Applied Basic Research and Clinical Practice*.

(9)   Xu, Z.; Shaw, J. B.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 265–273.

(10)   Nyakas, A.; Stucki, S. R.; Schürch, S. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 875–887.

(11)   Egger, A. E.; Hartinger, C. G.; Ben Hamidane, H.; Tsybin, Y. O.; Keppler, B. K.; Dyson, P. J. *Inorg. Chem.* **2008**, *47*, 10626–10633.

(12)   Marchand, A.; Gabelica, V. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1146–1154.

Assignment	Observed m/z	exact m/z	error/ppm
$[w3]^{2-}$	461.07274	461.07274	0.00
$[w5]^{3-}$	512.74713	512.74728	-0.30
$[w6]^{3-}$	614.09565	614.09596	-0.51
$[w2]^{1-}$	634.10643	634.10638	0.07
$[a5-BH]^{2-}$	673.11184	673.11276	-1.36
$[w9]^{4-}$	692.86039	692.86046	-0.11
$[a10-BH+Pt*]^{4-}$	821.38135	821.37979	1.90
$[a6-BH]^{2-}$	829.64091	829.64156	-0.78
$[w7+Pt*]^{3-}$	849.46440	849.46484	-0.51
$[w6]^{2-}$	921.64760	921.64758	0.02
$[w3]^{1-}$	923.15299	923.15276	0.25
$[w11+Pt*]^{4-}$	951.65001	951.65025	-0.25
$[a9-BH+Pt*]^{3-}$	990.82420	990.82485	-0.65
$[a6-BH+Pt*]^{2-}$	1026.16522	1026.16607	-0.83
		Absolute average error	0.54
		Standard deviation	0.72

SI table 5.1: fragmentation assignments for the CAD MS/MS spectrum of the $[DNA12+Pt(py)_2(N_3)-6H]^{5-}$ species. $Pt* = Pt(py)_2(N_3)$

Assignment	Exact mass	Observed mass	error/ppm
[w1]⁻	330.060346	330.06035	0.01
[a2-B]⁻	401.076085	401.07521	-2.18
[w3]²⁻	461.0727395	461.07274	0.00
[w6]³⁻	614.09596	614.09606	0.16
[w4]²⁻	617.601542	617.6014	-0.23
[w2]⁻	634.106383	634.10653	0.23
[w7+Pt*]⁴⁻	636.846808	636.84644	-0.58
[w3+Pt*]²⁻	657.59725	657.5972	-0.08
[w9]⁴⁻	692.86046	692.86044	-0.03
[w6+Pt*]³⁻	745.112301	745.11204	-0.35
[w5]²⁻	769.62456	769.62458	0.03
[w3]⁻	923.152755	923.15276	0.01
[w9]³⁻	924.14971	924.14852	-1.29
[w5+Pt*]²⁻	966.149071	966.14749	-1.64
[a11-BH+2Pt*]⁴⁻	992.1587068	992.16016	1.46
		Absolute average	0.55
		Standard deviation	0.84

SI table 2: fragmentation assignments for the CAD MS/MS spectrum of the [DNA12+2(Pt(py)$_2$(N$_3$))-6H]$^{5-}$ species. Pt* = Pt(py)$_2$(N$_3$)

Chapter 6:

Studying electron-quenching-Iridium metallodrug-modified biomolecules using CAD and ECD FT-ICR MS/MS.

Square-planar platinum compounds such as Cisplatin and Oxaliplatin have been shown to bind to an array of biomolecule targets, such as cancerous DNA, and off-target species, such as peptides and proteins, via loss of their bound ligands in vivo and creation of more active species (such as the aqua-modified cisplatin species) which can then bind to biomolecules. Piano-stool geometry transition metal complexes (such as Os1 shown in Chapter 4) have also been shown to effectively bind to target biomolecules via a reactive site in the piano-stool geometry, but again only after the chlorido ligand dissociation/displacement in solution.

This Chapter focuses on interesting new iridium piano-stool complexes, based upon complexes which have been shown to bind effectively to peptides and proteins previously by Qi *et.al.*, via the loss of a chlorido ligand. The complexes presented herein though have been modified to include reactive functional groups on the bidentate ligands in the piano-stool geometry, allowing the possibility of not only the standard metal-centred binding to biomolecular species (as with previous compounds) but also new reactivities and products via reaction with the ligand aldehyde functional groups.

These added functionalities were not only shown to be effective at binding to poly(amino acids) by MS, but the reaction products produced also posed interesting challenges to subsequent MS/MS experiments, causing changes in fragmentation mechanisms, fixed charge-induced distortion of fragmentation patterns, and interestingly stabilisation/"quenching" of ECD electrons by the iridium centre while covalently bound causing a large charge-reduced species peaks and little dissociation. However when the ECD electron was quenched by dative-covalent bound iridium modifications, the

reduced iridium centre dissociated from the biomolecule species, hindering but not preventing

MS/MS characterisation. These new functionalities were shown to not only provide new binding

mechanisms but were also found to provide a new way of targeting specific amino acid residues

which the non-functionalised iridium complex (studied by Qi *et.al.*) did not bind to. Furthermore the

aldehyde functionality on the bidentate ligand of the Ir complex was also shown to alter the binding

preference of the metal centre (in comparison to the un-functionalised complex), as the Ir

complexes presented herein formed metal-biomolecule bonds at histidine residues, not methionine

residues as expected from previous studies. Overall these new methods of functionalising piano-

stool complexes may be able to allow more tuning of metallodrug reactivity and allow targeted

binding in more ways than previously observed, and may present similarly interesting mechanisms

of action against cancer cell components in vivo.

Although the results presented within this chapter are predominantly based on FT-ICR MS and

MS/MS results from the thesis Author, this was a collaborative project and some elements were

conducted by collaborating researchers; namely Adam Millet provided the Ir1,Ir2, and Ir3

compounds and was responsible for the synthesis and purification of said compounds. Subsequent

NMR and other non-MS data associated with the compounds originated from Adam Millet and/or

his laboratory. All biomolecule-metallodrug reactions, analysis, MS/MS characterisation etc.

presented were conducted by the Author.

Studying electron-quenching-Iridium metallodrug-modified biomolecules using CAD and ECD FT-ICR MS/MS.

Christopher A. Wootton, Adam Millet, Andrea Lopez-Clavijo, Mark Barrow, Peter J. Sadler, and Peter B. O'Connor

Department of Chemistry, University of Warwick, Coventry, UK.

Abstract:

Transition metal complexes represent a key force in developing effective anti-cancer treatments, but with platinum drug resistance cases increasing each year, new transition metals have been investigated and synthesised to offer increased potencies, selectivities, and possible circumvention of resistance mechanisms inside resistant cancer cells. Iridium piano-stool complexes have been shown to offer many of these properties. Herein we present 3 novel iridium(III) anti-cancer complexes and show, using MS and MS/MS techniques, how functionalising the bidentate ligand within the metallodrug complex can offer new binding mechanisms and alter the preferred binding location of the complexes to model biomolecules compared to the unfunctionalised complex previously studied. The presence of the Iridium complex modification within target ions was shown to have a drastic effect on both CAD and ECD MS/MS behaviour, requirements, and fragmentation channels. CAD MS/MS was shown to only be effective when studying the covalently-linked condensation products. In contrast, ECD MS/MS, though hindered by the electron-quenching iridium complex modifications, was found to be suitable for studying many species observed and locating modification binding sites and identities often to within a single amino acid residue.

Introduction:

Platinum based metallodrugs are currently used in over 50% of all anti-cancer chemotherapy, they have shown to be effective against a wide variety of cancerous cells and have been studied extensively as a result, yielding insights into their mechanisms of action,[1] transformations in solution,[2] and binding preferences with respect to peptides,[3] proteins,[4] and oligonucleotides.[5] Unfortunately many traditional Pt[(II)] based metallodrugs, such as Cisplatin, Transplatin, and Oxaliplatin have been shown to cause a wide array of side effects and be less and less effective during long term courses of treatment where cancerous cells often develop resistance to the Pt[(II)] species and treatment becomes ineffective.[6]

Other transition metals have shown great promise as chemotherapy agents, boasting potencies and selectivities exceeding that of Cisplatin.[7] More interestingly these new transition metal complexes have been shown to have varying mechanisms of action against cancerous cell components, while some are able to bind to cancerous DNA, causing deformations and eventually apoptosis (in a comparable fashion to Cisplatin), other metallodrugs have been shown to disrupt other vital processes in cells in order to kill target cancer cells, such as metallodrug binding to mitochondria,[7] or disruption of the NAD^+/NADH balance inside cells.[8] Exploring these new mechanisms of action against cancerous cells has become the new focus of the next generation of anti-cancer compound synthesis efforts in order to create compounds which can circumvent previous resistances and find more effective ways of killing cancerous cells.

Unlike the square planar Pt[(II)] compounds described above, other d-block metals readily adopt an octahedral ligand geometry. Piano-stool octahedral anti-cancer complexes have been shown to be particularly potent with a wide variety of metals utilised including; ruthenium,[9]

rhodium[10], iridium,[11] vanadium,[12] and recently osmium compounds[13,7] have all been shown to express effective mechanisms of action against cancerous cells.
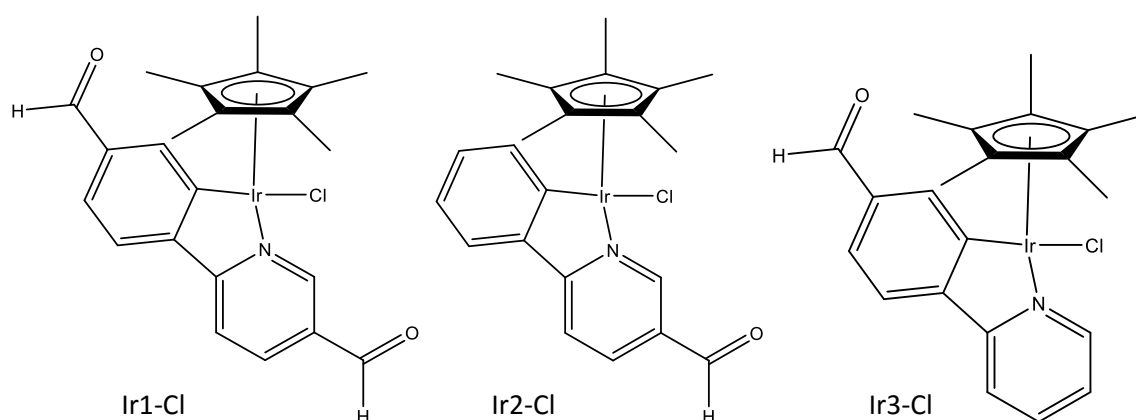
Figure 6.1: Iridium based piano-stool complexes used in the study – all compounds possess aldehyde functionalised bi-dentate ligands.

Herein a biomolecule binding study is presented exploring 3 iridium based piano-stool complexes which have been shown to possess high cytotoxicity/anti-proliferative effects against cancerous cells (Figure 6.1, above).[11] Methylated-cyclopentadienyl iridium compounds have been synthesised and studied using MS and MS/MS by Qi *et al.* previously.[14] However, unlike many metal-based chemotherapeutic agents previously described, the iridium compounds presented were modified to possess functional groups on the bidentate ligands within the complexes, which allows the possibility of biomolecules of interest to bind not only to the active metal centre (the mechanism of the vast majority of metal complexes and metallodrugs), but also for species to react with the ligand functional groups and produce new covalent bonds in addition to the dative bonds formed at the metal centre. The enhanced reactivity of the ligand groups produced interesting new reaction pathways and both the ligand-bound and metal centre bound species showed interesting challenges to both MS and especially MS/MS analysis, with the iridium complex having a

drastic effect on both threshold based dissociations (such as CAD and IRMPD) and electron based dissociations (such as ECD).

Materials and methods:

Substance P, [Lys][3]-Bombesin, and formic acid were purchased from Sigma Aldrich (St. Louis, MO). LCMS grade acetonitrile (ACN) was obtained from VWR (Radnor, PA) and used without further purification. Low concentration tuning mix was purchased from Agilent Technologies (Santa Clara, CA) and used as received. The iridium based Metallodrugs Ir2 and Ir3 were synthesised and characterised using previous methods.[11] The di-aldehyde Iridium drug Ir1 synthetic procedure is presented below:

2-(4'-formylphenyl)-5-pyridinecarboxaldehyde (ligand 1) 4-formylphenylboronic acid (250 mg, 1.67 mmol) and 6-bromo-3-pyridinecarboxaldehyde (222 mg, 1.19 mmol) were dissolved in 1:1:1 mixture of tetrahydrofuran:water:1M sodium carbonate aqueous solution and stirred at ambient temperature under nitrogen for 1 h. Tetrakis(triphenylphosphine)palladium (28 mg, 0.02 mmol) was added and the reaction mixture was refluxed at 100°C for 24 h. The reaction mixture was cooled to ambient temperature, extracted with dichloromethane, the organic layer was washed with saturated sodium hydrogen carbonate solution and brine, then dried with magnesium sulphate and concentrated to dryness, yielding the crude material. Purification by column chromatography was performed using 1:1 chloroform:ethyl acetate (Rf = 0.59), yielding an off-white solid (152 mg, 60 %). **^{1}H NMR** (400 MHz, dmso-d_6): δ 10.17 (s, 1H), 10.11 (s, 1H), 9.21 (d, 1H, J = 2.0 Hz), 8.41 (m, 2H), 8.38 (dd, 1H, J = 8.3, 2.0 Hz), 8.32 (d, 1H, J = 8.5 Hz), 8.07 (m, 2H). **^{13}C NMR** (100 MHz, dmso-d_6, DEPT135): δ 192.89 (C*H*O), 192.05 (C*H*O), 151.63, 137.49, 129.99, 127.93, 121.54. **ESI-MS** (MeCN) m/z = 212.1 [M+H]$^+$.

[(η⁵-Cp*)Ir(2-(4'-formylphenyl)-5-pyridinecarboxaldehyde)Cl] (Ir1) 2-(4'-formylphenyl)-5-pyridinecarboxaldehyde (42 mg, 0.202 mmol) was dissolved in dichloromethane (30 mL) followed by the addition of sodium acetate (33 mg, 0.40 mmol) and the reaction mixture was stirred at ambient temperature under nitrogen for 30 min. [(Cp*)IrCl$_2$]$_2$ (80 mg, 0.101 mmol) was added and the reaction mixture was heated under reflux for 46 h. The crude product was recrystallised from chloroform/hexane at 273 K to yield a red crystalline solid (65 mg, 57 %). **¹H NMR** (400 MHz, dmso-d$_6$): δ 10.17 (s, 1H, C*H*O), 10.13 (s, 1H, C*H*O), 9.20 (d, 1H, J = 1.5 Hz), 8.46 (d, 1H, J = 8.3 Hz), 8.34 (dd, 1H, J = 8.3, 2.0 Hz), 8.26 (d, 1H, J = 1.5 Hz), 8.17 (d, 1H, J = 8.0 Hz), 7.55 (dd, 1H, J = 8.0, 1.5 Hz), 1.67 (s, 15H). **ESI-MS** (MeCN) m/z = 538.1 [M-Cl]⁺. Slow evaporation of a chloroform/hexane mixture at 273K led to crystals suitable for X-ray diffraction.
A crystal structure for the Ir1 complex is also included in the supporting information (Figure SI1).

Reaction of Peptides with Iridium metallodrugs:

Aqueous solutions of Substance P (1 mM) and Bombesin (1 mM) were prepared and mixed with a solution of an iridium metallodrug in ACN (250μM) to give solutions of 0.5:1 and 1:1 (drug:peptide) ratio. The samples were then placed in an incubator (Genlab, Cheshire, UK) at 37 ˚C for 1-4 days before being diluted 50 fold with ACN to MS concentrations (~2.5 μM summed concentration for all reaction products) and frozen at -80 ˚C ready for MS analysis. Freshly prepared samples were compared to those frozen for 1 week to several months and showed no observable variation in the mass spectra obtained. Frozen samples were thawed and placed on a multi vortexer (Grant Bio, Cambridge, UK) prior to MS analysis.

FT-ICR Mass spectrometry:

Nano-electrospray (nESI) Mass spectrometry was performed on a Bruker SolariX Fourier Transform Ion Cyclotron Resonance Mass spectrometer (FT-ICR MS) fitted with an Apollo II Ion source and a 12 Tesla actively shielded magnet (Bruker Daltonics, Bremen, Germany). Aqueous Peptide samples (5 µM) were spiked with 0.3% formic acid to aid in effective ionisation during nESI. Iridium containing samples were sprayed in ACN with no added acid.

For CAD MS/MS analysis, species of interest were isolated in the first quadrupole and accelerated into the hexapole collision cell for Collisionally Activated Dissociation (CAD) at 10-30 V and continuously accumulated for 0.1-4 seconds before transfer to the Infinity Cell© for detection.

For ECD MS/MS analysis, the species of interest were isolated in the first quadrupole, externally accumulated in the collision cell for 0.1-7 seconds and then transferred to the Infinity Cell© for Electron Capture Dissociation (ECD) fragmentation and detection. Ions in the Infinity Cell© were irradiated with 1.3 eV (peptides)-2.5 eV (Ir adducts) electrons from a 1.5 A hollow cathode dispenser for 50-1200 ms prior to detection.

MS/MS spectra were internally calibrated using the minimal number of unmodified (peptide spectra) or modified (Ir adduct spectra) b/c ions and the charge reduced species $[M+nH]^{n-1+\cdot}$ where possible (species used for calibration are marked).

Results and discussion:

The iridium metallodrugs were shown to react readily with Substance P (SubP) and [Lysine]³-Bombesin (L3BBS); the resulting spectra (below, Figure 6.2, and supporting information Figures SI6.2 and SI6.3) show few species; unreacted iridium drug+solvent adducts (minus the Cl ligand, lost during the ESI process), unreacted peptide species, and iridium modified peptides. Iridium containing species were readily assigned as such due to iridium's characteristic isotope pattern (Figure 6.2 inset). The Ir drug+Substance P mass spectra all showed a single iridium containing species as a result of the reaction with substance P, this species corresponded to the mass of the iridium drug+peptide minus the mass of a water molecule (hereafter referred to as the condensation product), since the amidated Substance P used in this study contains no amino acid group which would readily lose a water molecule, even if the species were additionally collisionally activated (which it was not), the reaction product could have resulted from a condensation reaction with one of the aldehyde functional groups contained within the iridium ligand.

The Ir drug+L3BBS mass spectra showed the same reaction products as the Ir drug+Substance P reaction; unmodified drug+adducts, unmodified peptide species, and a condensation product ([Ir drug+L3BBS-H_2O]$^{2+}$), the spectra also showed an additional reaction product; that of a non-condensation product; [Ir drug+L3BBS]$^{2+}$, due to the observation of both the Ir drug without the Cl ligand and water/solvent molecule adduction ([$IrC_{10}H_{15}C_{10}H_6N_2$]$^+$ and [$IrC_{10}H_{15}C_{10}H_6N_2+H_2O$]$^+$ respectively for Ir1-containing spectra, Figure 6.2 inset), the identity of the iridium metallodrug modification for the L3BBS species was ambiguous, even with accurate mass-mass spectrometry (readily achieved on the FT-ICR MS used), the 2 species have exactly the same mass and only varied by the location of the H_2O. Both Collisionally Activated Dissociation (CAD) and Electron Capture Dissociation (ECD) were conducted on all iridium-modified species in order to elucidate not only the

identity of the Iridium metallodrug once bound to poly(amino acids) but also to locate the preferential binding locations of these novel therapeutic agents. It is noteworthy that after a 24 hour incubation with the iridium drugs the Substance P peptide had a near-complete reaction with the iridium drug and produced a single reaction product. Although reaction with L3BBS produced 2 distinct reaction products for each iridium drug, the extent of reaction was much lower (as observed by the lower relative intensity in the Ir-L3BBS spectra, Figure 6.2).
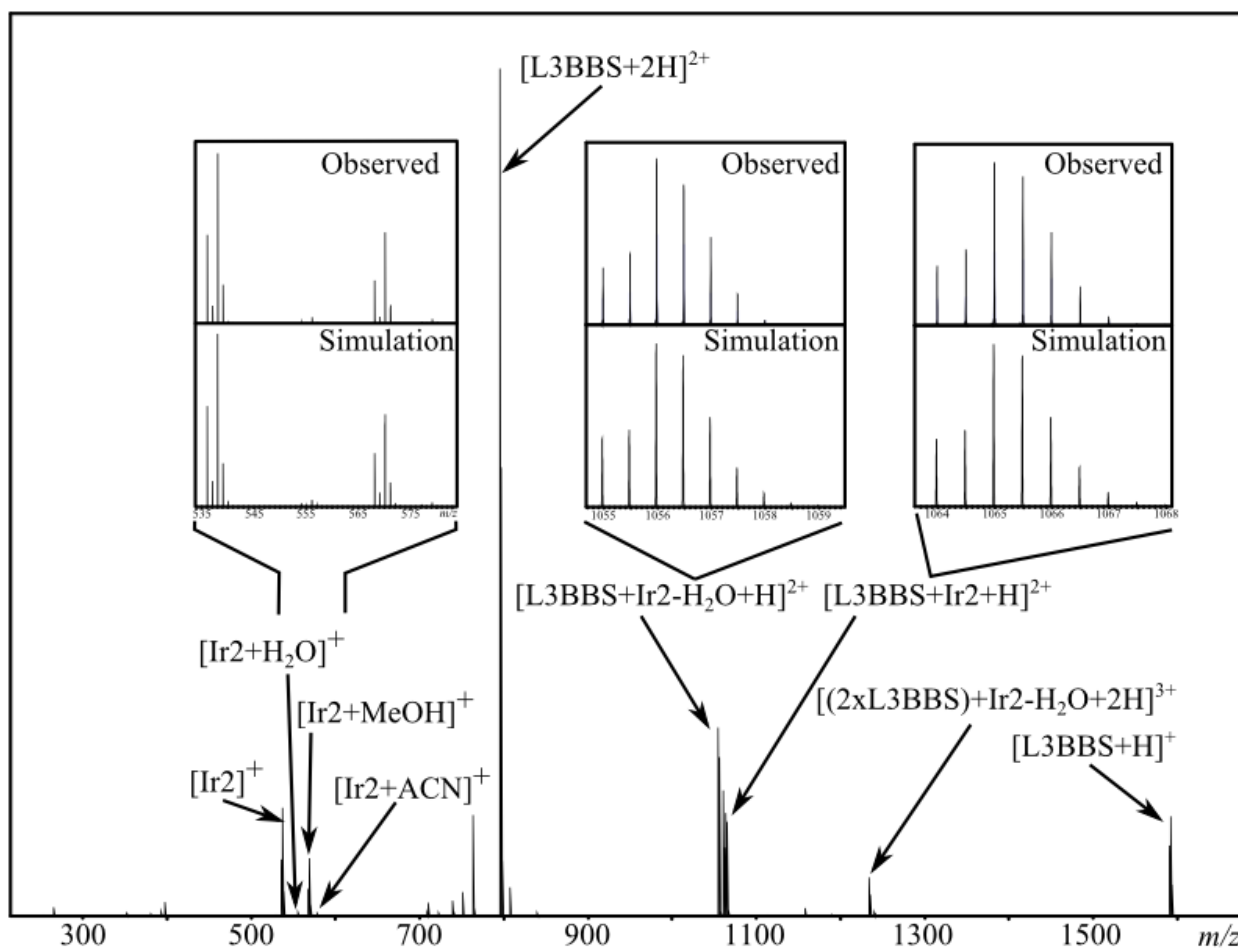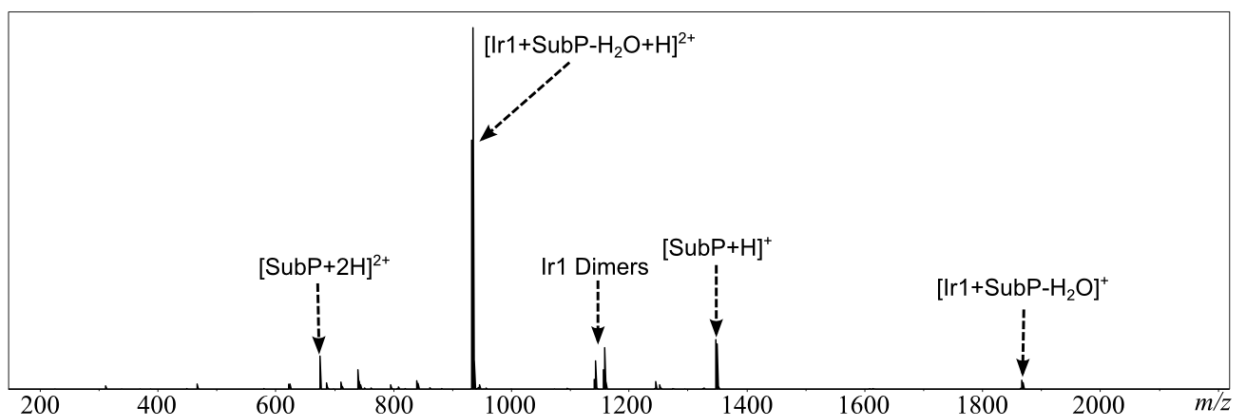
Figure 6.2: nESI-FT-ICR Mass Spectra of Iridium metallodrug+peptide mixtures, reacted for 24 hours at 37°C prior to analysis. In the SubstanceP+Ir1 metallodrug (top) only the condensation reaction product ([SubP+Ir1-H$_2$O+H]$^{2+}$) is observed outside of reagent peaks. In the Lys3-Bombesin+Ir2 metallodrug spectrum 2 Iridiated peptide species are observed (inset) both the condensation product ([L3BBS+Ir1-H$_2$O+H]$^{2+}$) and the non-condensation (metal bound) product

([L3BBS+Ir2+H]²⁺). Inset: Observed Iridiated peptide/Iridium containing peaks, showing the influence of the transition metal on the observed isotopic distribution.

Tandem MS of Iridium modified species:

Previous studies of piano-stool iridium compounds binding to poly(amino acids) by Qi *et al.* showed that the non-aldehyde functionalised analogues of the metallodrug bind to methionine residues in target proteins, and could be effectively studied using both Collisionally Activated Dissociation (CAD) and Electron Capture Dissociation (ECD).[14] Many peptide and protein modifications, be they natural Post Translational Modifications (PTM's) such as phosphorylation, synthetic modifications, or metal-based modifications, are easily lost using slow-heating/threshold based dissociation techniques (such as CAD), but can be retained using electron based dissociation techniques such as Electron Capture/Transfer Dissociation (ECD and ETD respectively), with metal complex-based modifications being particularly susceptible to loss during CAD.[15][13] However, standard CAD and ECD MS/MS of the reaction products shown in Figure 2 provided interesting results and non-standard dissociation mechanisms compared to the unmodified peptide species, each of which are detailed below.

The irdium metallodrug-Substance P reactions showed only one iridiated-peptide reaction product ([Irx+SubP-H$_2$O+H]$^{2+}$). The subsequent ECD MS/MS spectrum of the quadrupole-isolated species is shown below in Figure 6.3.
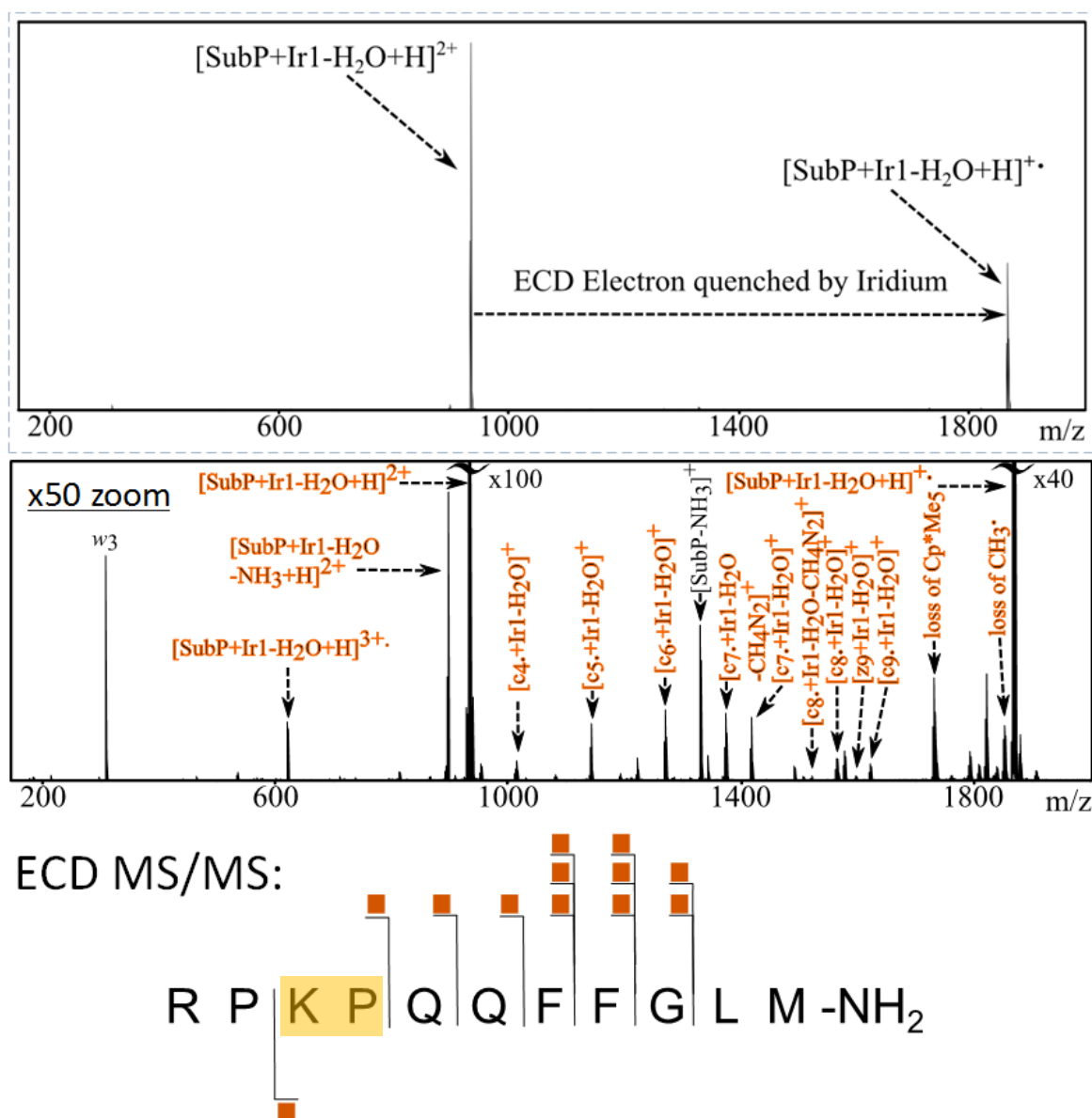
Figure 6.3: ECD MS/MS spectrum of SubstanceP+Ir1-H₂O condensation product. The available aldehyde group on the bidentate ligand binds to a lysine residue via a condensation reaction to form a covalently bound imine. ECD MS/MS of the imine-bonded peptide-Ir complex species leads to electron quenching by the Ir centre and little to no dissociation (top) though low intensity fragments (<1% of precursor) can be observed with high number of summed acquisitions (bottom), so much so that even the suppressed harmonic signals from the precursor can be observed (w_3).

Interestingly not only did the iridium modified species require very different electron energies and pulse lengths to achieve effective ECD MS/MS (compared to the corresponding unmodified peptide ions), but the parent ion-charge reduced species intensity ratio was drastically different. For instance, Substance P ECD MS/MS shows very effective electron capture, resulting in a charge reduced species (CRS, [SubstanceP+2H]$^{+\cdot}$)~5-10% intensity of the parent ion ([Substance P+2H]$^{2+}$) and good intensity ECD fragments (~10-25% of parent ion intensity).[16] For the iridium modified Substance P the intensity of the CRS compared to the 2+ parent ion was 41%, which is unusually high for peptide ECD MS/MS, the resulting ECD fragments were also of very low intensity (≤1%) compared to unmodified Substance P ECD MS/MS, indicating the iridium-based modification is having a drastic effect upon the ECD process.

The iridium metal centre within the metallodrug modifications was in the 3+ charge state (i.e. Ir^{III}), with 2 charges neutralised via bonds to ligands (Cp* and the NC$^-$ ligand), leaving the iridium ccomplex in an overall 1+ charge state, which explained the addition of a single proton to produce the [Peptide+Irx+H]$^{2+}$ species detected. The Ir$^+$ complex can be easily reduced to Ir^0 by electron capture during ECD MS/MS to produce a charge reduced species in which the electron has been effectively quenched by the iridium centre, rather than a more typical CRS where the electron has caused backbone/sidechain dissociation, with the individual fragments remaining adducted together via hydrogen bonds/other covalent/backbone bonds.[17,18]

The quenching of electrons during electron-based dissociations has been observed before when studying covalently-bound heme-containing biomolecules, such as Cytochrome c, and when studying species with particularly high histidine content, the conjugated pi-systems were believed to be able to stabilise incoming electrons and the resulting radical species generated, preventing/subduing ECD/ETD fragmentation during MS/MS.[19] Europium-

adducted biomolecules and some other heavy metal adducts are also believed to be so-called "electron sinks" and quench electrons,[20] however these effect have been observed only with heptadentate ligands surrounding the metal centre or with bare metal centres as adducts/charge carriers and not as specific biomolecule modifications as metal complexes.

Despite the partial quenching of ECD electrons by the iridium centre, low intensity fragments were observed in all ECD MS/MS spectra of Ir-containing reaction products (Figure 6.3 and in SI). These fragments are summarised in the corresponding fragmentation maps (Figure 6.3 and in SI). For the iridium complex-modified Substance P species there is a clear preference for lysine binding, accompanied by a water loss during the process, indicating a condensation reaction between the terminal primary amine on the lysine side chain and the aldehyde group present on the bidentate ligand within the iridium complex. The condensation reaction between the complex and the biomolecule results in the formation of an imine, a strong covalent bond which surprisingly does not involve the metal centre directly. Also surprisingly there seems to be no evidence of methionine binding between the terminal methionine residue of either peptide and the Iridium centre, which is in contrast to previous studies by Qi *et al.,*[14] who showed the non-aldehyde functionalised iridium[(III)] complex analogue binds strongly to methionine residues in the calcium-binding protein Calmodulin; the presence of the aldehyde functional group(s) seem(s) to be affecting both the mechanism of binding (additional binding via the ligand) and the preferential location of said binding (lysine over methionine).

ECD MS/MS of the Irx+L3BBS condensation products (x=1, 2, or 3) (Figure 4, below and in SI) also showed an unusually high CRS relative intensity compared to the parent ion (~250%), showing severe quenching of ECD electrons. ECD of the unmodified L3BBS was much more routine, with a CRS intensity ~10-20% compared to that of the precursor. ECD MS/MS analysis of the iridium complex-modified peptides was hindered greatly by the

iridium-induced electron stabilisation/quenching, limited fragments were assigned showing

iridium metallodrug binding towards the N-terminus of the peptide (see below).
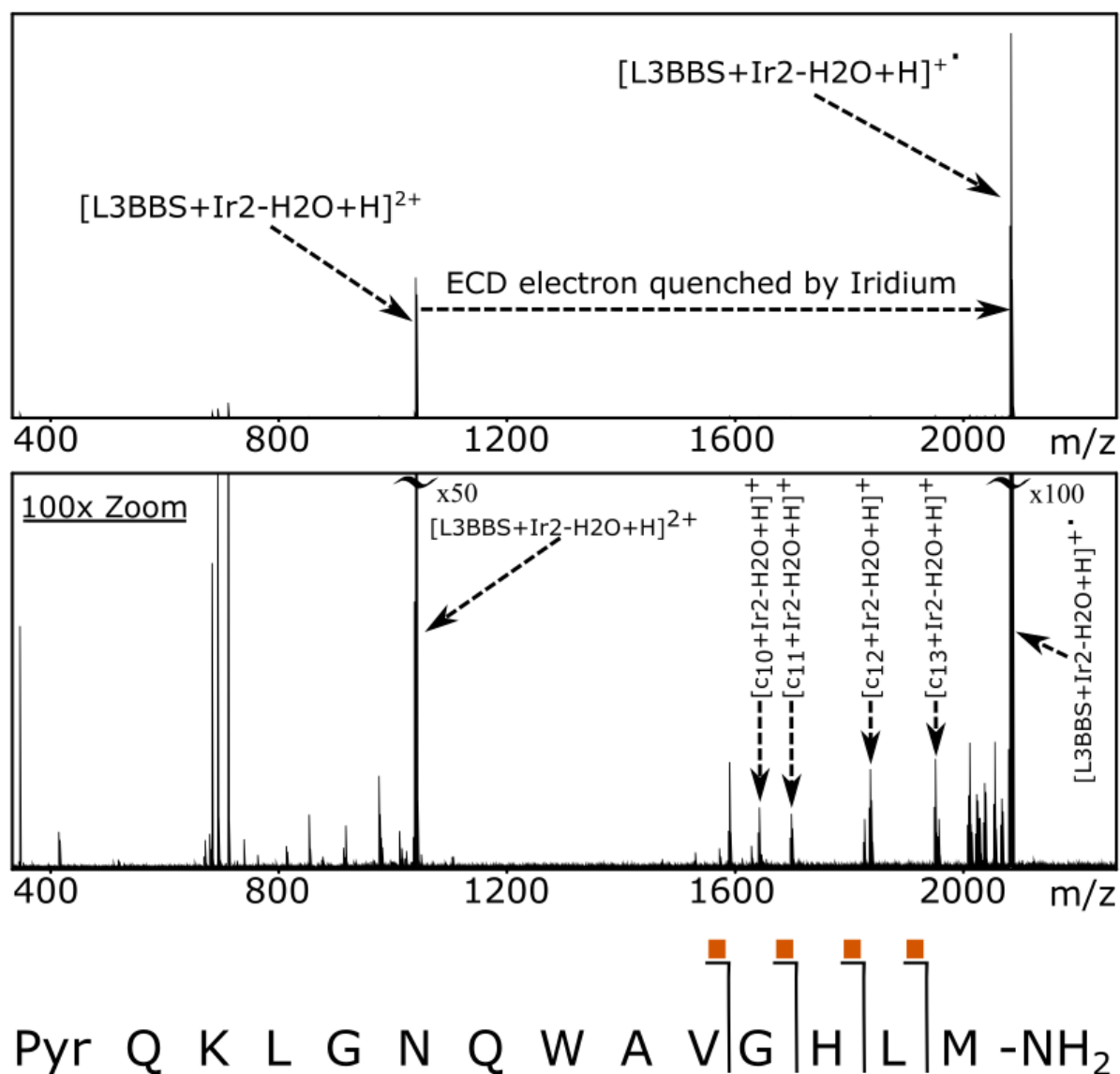
Figure 6.4: ECD MS/MS of the condensation reaction product between Ir2 and L3BBS ([L3BBS+Ir2-H$_2$O+H]$^{2+}$). The spectrum shows the iridium complex modification is effectively quenching the ECD electrons producing electron capture, but little dissociation (top). Zoom in of the spectrum shows when the small proportion of electrons are captured along the backbone, low intensity fragments can be observed

(middle), fragments can then be assigned (fragmentation map, bottom). Orange squares indicate fragments containing the Ir modification.
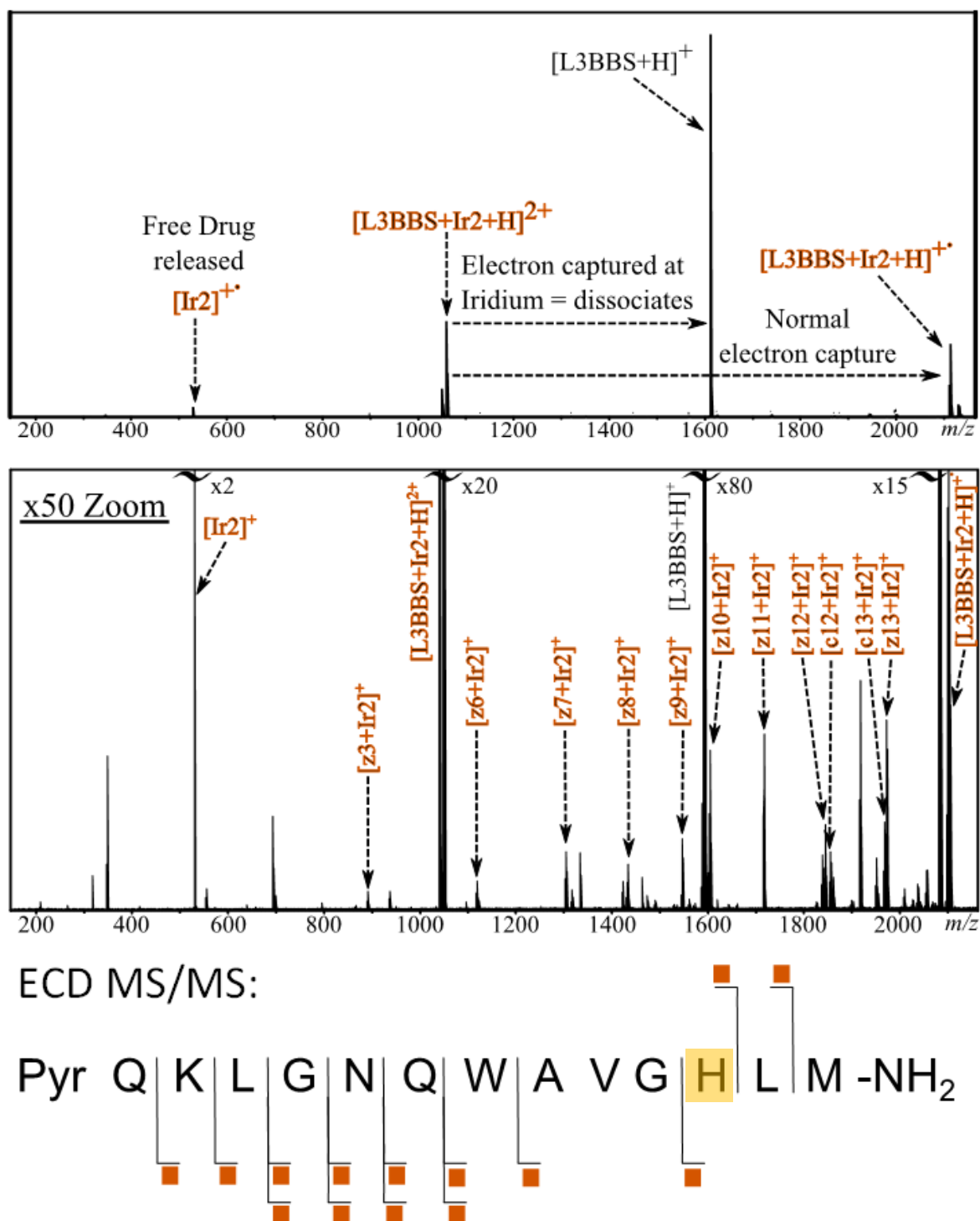
Figure 6.5: ECD MS/MS of the iridiated K³-Bombesin non-condensation product [L3BBS+Ir2+H]²⁺ (top) and zoom in of ECD fragments produced (middle) by the small

proportion of electrons captured along the peptide backbone and not quenched at the iridium centre, which caused dissociation of the complex and an abundant unmodified peptide peak. Resulting ECD MS/MS fragmentation is summarised in the fragmentation map (bottom), orange squares indicate the iridium modification is present on a given fragment.

ECD MS/MS of the non-condensation product produced ($[Ir_x+L3BBS+H]^{2+}$) are shown in Figure 6.5 and in the SI. The effect of the iridium complex bound to the peptide was shown to have a drastic effect on the ECD spectrum, with the main fragmentation pathway appearing to be electron capture at the iridium centre (as seen in the condensation product ECD spectrum), unfortunately the reduction of the iridium centre by captured electrons caused the complex to dissociate from the biomolecule and produce a large peak corresponding to unmodified $[L3BBS+H]^+$. Fortunately being a 1+ ion, the $[L3BBS+H]^+$ species produced was not able to undergo further ECD MS/MS fragmentation and produce detectable fragments due to the charge being neutralized upon electron capture. Though most of the precursor ions fragmented via loss of the iridium complex modification, low intensity fragment ions were observed and were used to locate the modification binding site. All 3 Iridium complexes were shown to preferentially bind to the [histidine][11] residue within L3BBS in a non-condensation (i.e. metal-centred) fashion via a dative covalent bond to the Ir centre. The detection of histidine-centred binding via the Ir metal centre showed a change in the binding preference of these functionalised iridium compounds in contrast to the unfunctionalised ligand iridium(III) compounds studied previously[14] which showed a preference for methionine residues and to the condensation products described above (preference for lysine residues). The change in binding behaviour of these functionalised iridium complexes compared to other transition metal complexes, which often show a distinct preference for sulphur containing residues (such as methionine), especially those based upon platinum, show that the behaviour of

transition metal complexes, even when based upon the same metal centre, can change dramatically with even small changes to ligand composition and reactivity.

The formation of the covalent imine bond suggested that the condensation Iridium modification should be able to survive (even partially) a threshold based dissociation technique. CAD MS/MS was conducted on the condensation and non-condensation reaction products observed in the L3BBS spectrum. As expected CAD of the non-condensation product resulted in loss of the Iridium complex modifications at low CAD energies/voltages (~5-8V) whereas backbone fragmentation would not occur to a useful extent without the use of high voltages (15-20V). CAD MS/MS of the L3BBS condensation products on the other hand was very successful (Figure 6, below). Though CAD caused some loss of the bidentate ligand from the CRS (shown in Figure 6.6 and in SI tables of assignments), this was a minor fragmentation pathway. The vast majority of peaks observed were readily assigned to unmodified/Iridium complex-modified b/y fragments and side chain loss ions therefrom. The CAD MS/MS provided correlating and complementary data to the ECD MS/MS of the condensation product, summarised in Figure 6.5 fragmentation maps, the complexes are clearly shown to bind preferentially to the Lysine residue, via loss of a water molecule.
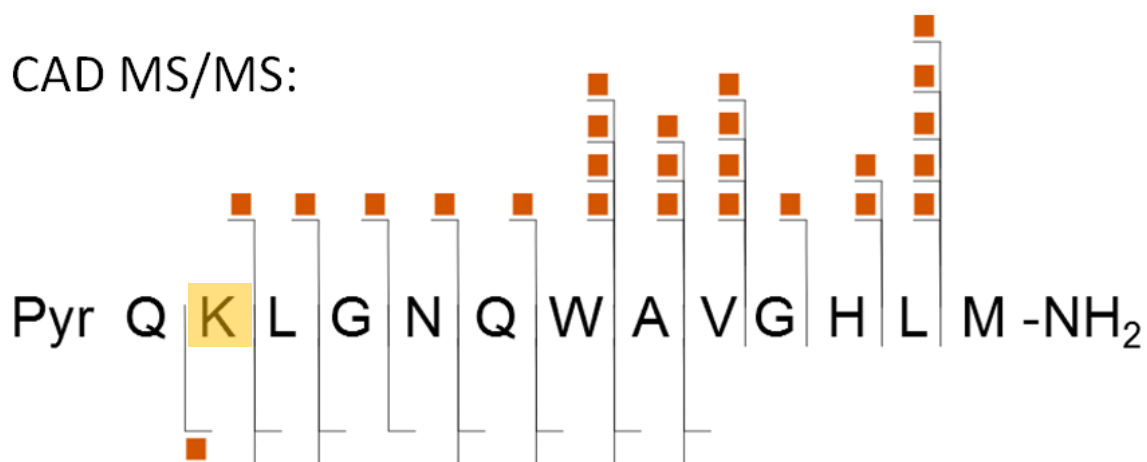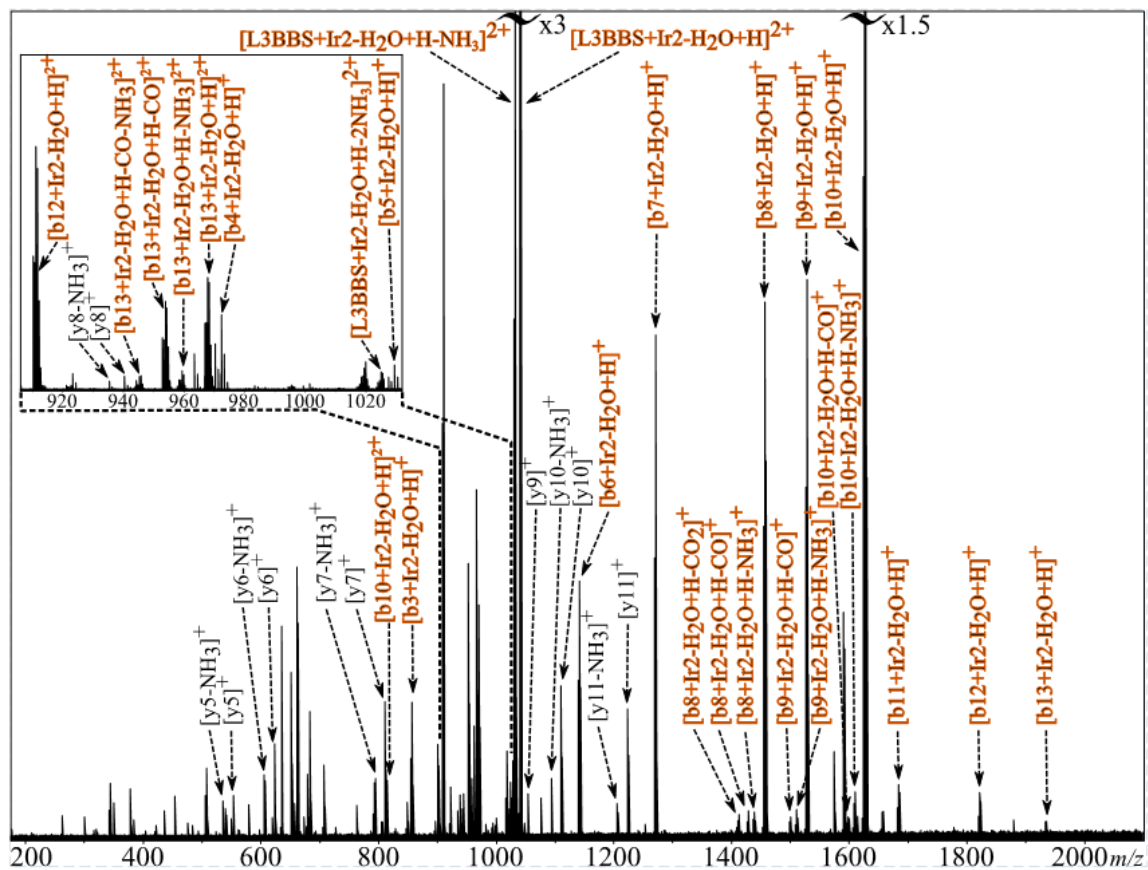
Figure 6.6: CAD MS/MS spectrum of iridium-modified peptide (Ir2+L3BBS) via a covalent imine bond caused by a condensation reaction, allowing the usually fragile metal modification to survive CAD MS/MS. Orange squares indicate a fragment modified with the Iridium complex. Binding of the complex was readily located to the Lysine[3] residue.

Although the Ir1 complex contained 2 aldehyde functional groups within the bidentate ligand, no cross-linked, di-condensation product was observed (i.e. a product formed from 2 condensation reactions, one at each aldehyde group), even during subsequent experiments using higher ratios of Ir1 to peptide, nor with longer reaction times (up to 7 days). This lack of further reaction could indicate a difference in reactivity between the aldehyde groups. Due to the different compositions of the aromatic rings within the bi-dentate ligand, a difference in reactivity could be possible (due to one being pyridine-based and one phenyl-based). Though comparisons with Ir2-peptide and Ir3-peptide spectra showed that regardless of which ring the aldehyde group is positioned, only one condensation reaction occurs, indicating that another factor inhibits future binding once the first imine bond is formed – most likely steric hindrance of the peptide biomolecule coiling around/blocking further access to the areas of the metal complex post-Imine formation.

It is worth noting that a low intensity peak was detected in the mass spectrum of the Ir1-L3BBS reaction mixture corresponding to a possible cross-linked Ir1-peptide species ($[Ir1+2xL3BBS-H_2O+2H]^{3+}$). Though it should also be noted that a peptide dimer peak ($[2xL3BBS+3H]^{3+}$) was observed – resulting from 2 peptide ions held together via non-covalent interactions (such as hydrogen bonds). Since both a peptide dimer and an Ir1-modified species containing 2 peptides were both observed the identity of the Ir1 modified species was ambiguous, possibilities include a crosslinked peptide species (resulting from a condensation product (imine bound to one peptide) and a metal centre-bound peptide) or an L3BBS dimer modified at one lysine residue (as described above). CAD and ECD MS/MS were carried out on the possible dimer species. CAD MS/MS predictably caused dissociation of the dimer at low CAD voltages (as discussed above), and little information beyond other than typical lysine modified fragments (as observed in the CAD MS/MS of the individual condensation products). ECD MS/MS was hindered greatly by the low intensity of the

species and by the loss of the Iridium complex forming the most popular fragmentation channel. Very few sequence informative fragments were assigned, but those that were (summarised in the SI) showed correlating information to that of the ECD MS/MS of individual modified peptides – i.e. that of a metal complex bound via an imine linkage to the lysine residue and the complex bound via a metal centred dative bond from the histidine residue, indicating the species was in fact an iridium-cross-linked reaction product. Similar iridium-cross-linked products were also observed in the Ir2-L3BBS and Ir3-L3BBS mass spectra, again re-enforcing that only one aldehyde group and one metal centred bond are required for crosslinking using these iridium(III) complexes. The Substance P+Ir1/Ir2/Ir3 spectra did not show any evidence of iridium-cross-linking, despite the methionine residue availability and the presence of two reactive aldehyde groups present on the Ir1 complex.

The various different binding modes of these functionalised iridium-based modifications are summarised below in Figure 6.7:
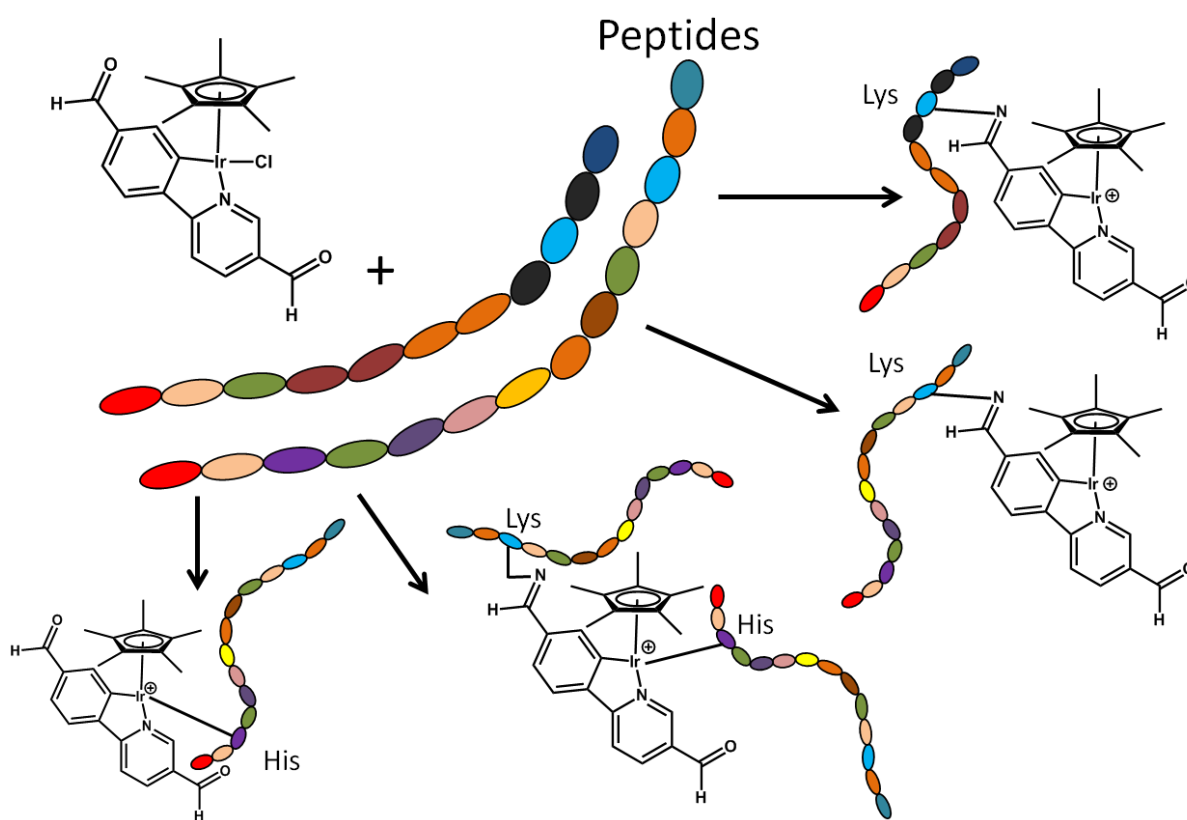
Figure 6.7: Summarising figure showing the range of reaction products observed during the reaction of functionalised iridium complexes with the model peptides

Conclusions:

Iridium piano-stool complexes have been shown to be effective anti-cancer metallodrugs, with potencies and selectivities exceeding that of widely used compounds such as Cisplatin. However studying more exotic metal-based modifications poses interesting challenges for both MS and especially MS/MS experiments. A study has been presented exploring how aldehyde-functionalised iridium(III) complexes can bind to model poly(amino acids), and the unusual effects these modifications have upon usually routine MS and MS/MS analysis.

The aldehyde-functionalised compounds were shown to bind to peptide via a dative covalent bond between the iridium metal centre and histidine residues of biomolecules when available. CAD MS/MS of dative covalent bonded species resulted in loss of the metal based modification and subsequent secondary fragmentation of the unmodified peptide produced, losing all sequence informative fragments related to modification location. ECD MS/MS of the dative covalent bonded species was challenging due to the electron quenching effect of the iridium based modification, but did produce fragments revealing the modification locations. The functionalised iridium compounds were also shown to bind to lysine residues within biomolecules via a condensation reaction, creating a covalent imine based linkage. CAD MS/MS of the covalently bonded imine species was very effective, with the iridium modification remaining on subsequent fragments and enabling location of the modification site.

References:

(1)   Reedijk, J. *Chem. Rev.* **1999**, *99* (9), 2499–2510.

(2)   Li, H.; Wells, S. a; Jimenez-Roldan, J. E.; Römer, R. a; Zhao, Y.; Sadler, P. J.; O'Connor, P. B. *Protein Sci.* **2012**, *21* (9), 1269–1279.

(3)   Li, H.; Snelling, J. R.; Barrow, M. P.; Scrivens, J. H.; Sadler, P. J.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1217–1227.

(4)   Li, H.; Lin, T.; Orden, S. L. Van; Zhao, Y.; Barrow, M. P.; Pizarro, A. M.; Qi, Y.; Sadler, P. J.; O'Connor, P. B. *Anal. Chem.* **2011**, *83*, 9507–9515.

(5)   Xu, Z.; Shaw, J. B.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (2), 265–273.

(6)   Sirichanchuen, B.; Pengsuparp, T.; Chanvorachote, P. *Mol. Cell. Biochem.* **2012**, *364* (1-2), 11–18.

(7)   van Rijt, S. H.; Romero-Canelón, I.; Fu, Y.; Shnyder, S. D.; Sadler, P. J. *Metallomics* **2014**, *6* (5), 1014.

(8)   Soldevila-Barreda, J. J.; Romero-Canelón, I.; Habtemariam, A.; Sadler, P. J. *Nat. Commun.* **2015**, *6*, 6582.

(9)   Ang, W. H.; Casini, A.; Sava, G.; Dyson, P. J. *J. Organomet. Chem.* **2011**, *696* (5), 989–998.

(10)   Lovejoy, K. S.; Lippard, S. J. *Dalton Trans.* **2009**, No. 48, 10651–10659.

(11)   Millett, A. J.; Habtemariam, A.; Romero-Canelón, I.; Clarkson, G. J.; Sadler, P. J. *Organometallics* **2015**, *34* (11), 2683–2694.

(12)   Strianese, M.; Basile, A.; Mazzone, A.; Morello, S.; Turco, M. C.; Pellecchia, C. *J. Cell. Physiol.* **2013**, *228* (11), 2202–2209.

(13)   Wootton, C. a.; Sanchez-Cano, C.; Liu, H.-K.; Barrow, M. P.; Sadler, P. J.; O'Connor,

P. B. *Dalt. Trans.* **2015**, *44* (8), 3624–3632.

(14)   Qi, Y.; Liu, Z.; Li, H.; Sadler, P. J.; O'Connor, P. B. *Rapid Commun. Mass Spectrom.* **2013**, *27* (17), 2028–2032.

(15)   Sweet, S. M. M.; Bailey, C. M.; Cunningham, D. L.; Heath, J. K.; Cooper, H. J. *Mol. Cell. Proteomics* **2009**, *8* (5), 904–912.

(16)   Lopez-Clavijo, A. F.; Barrow, M. P.; Rabbani, N.; Thornalley, P. J.; O'Connor, P. B. *Anal. Chem.* **2012**, *84* (24), 10568–10575.

(17)   Simons, J. *J. Am. Chem. Soc.* **2010**, *132* (20), 7074–7085.

(18)   Simons, J. *Chem. Phys. Lett.* **2010**, *484* (4-6), 81–95.

(19)   Chung, T. W.; Moss, C. L.; Wyer, J. A.; Ehlerding, A.; Holm, A. I. S.; Zettergren, H.; Nielsen, S. B.; Hvelplund, P.; Chamot-rooke, J.; Bythell, B. *J. Am. Chem. Soc.* **2010**, *132* (7), 10728–10740.

(20)   Mosely, J.; Murray, B.; Parker, D. *Eur. J. Mass Spectrom.* **2009**, *15* (2), 145–155.
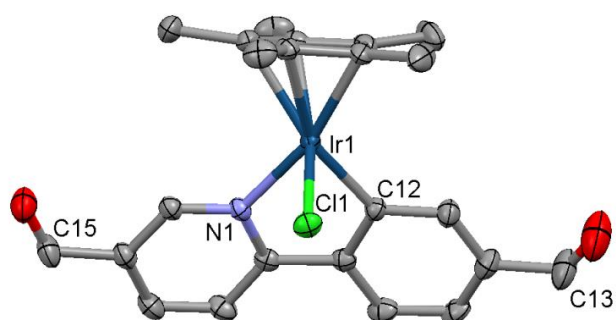
	2			2
Formula	C$_{23}$H$_{23}$ClIrNO$_2$		Ir-C (Cp* ring)	2.155(2)
MW	573.07			2.160(2)
Crystal Colour	Brown			2.166(2)
Cryst size (mm)	0.25 x 0.20 x 0.15			2.251(2
λ (Å)	0.71073			2.259(2
Temp(K)	150		Ir-C (centroid)	1.826
Cryst system	Monoclinic		Ir-C	2.032(2)
Space group	P2(1)/c		Ir-N	2.0789(19)
a (Å)	17.3889(2)		Ir-Cl	2.4051(6)
b (Å)	7.66780(10)			
c (Å)	16.5920(2)		C-Ir-N	77.86(9)
α (°)	90		C-Ir-Cl	87.08(6)
β (°)	113.945(2)		N-Ir-Cl	84.91(6)
γ (°)	90			
Vol (Å^3)	2021.89(4)			
Z	4			
$R(Fo^2)$	0.0195			
$Rw(Fo^2)$	0.0502			
GOF	1.088			

SI Figure 6.1: X-ray crystal structure of the Ir1 dialdehyde functionalised Iridium(III) piano-stool complex and associated information.
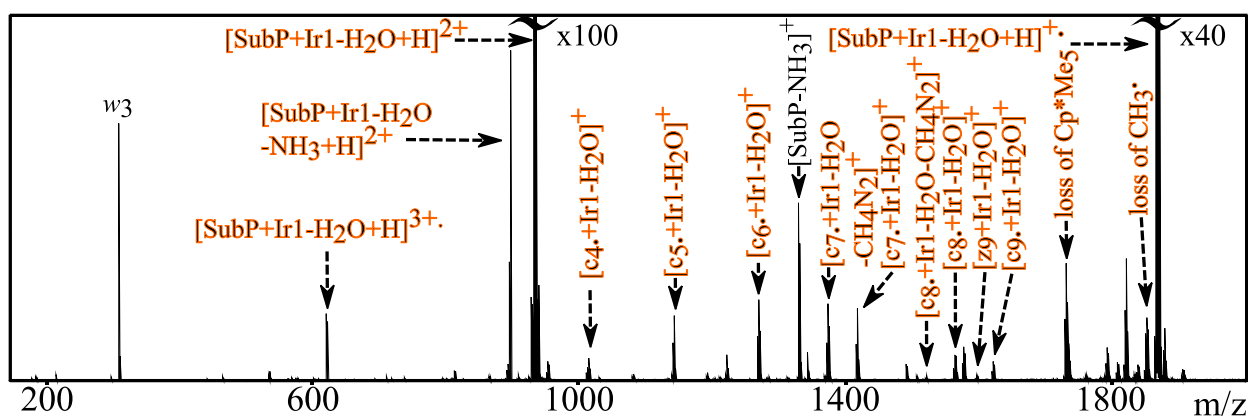
Fragment	Observed mass	Exact mass	Error (ppm)
$[C4\bullet]^+$	1013.45049	1013.4506	-0.11
$[C5\bullet]^+$	1141.50907	1141.50919	-0.11
$[C6\bullet]^+$	1269.56765	1269.56763	0.02
$[C7\bullet]^+$	1416.63606	1416.63635	-0.20
$[C8\bullet]^+$	1563.70447	1563.70452	-0.03
$[Z9-H\bullet]^+$	1595.67789	1595.67722	0.42
$[C9\bullet]^+$	1620.72593	1620.72693	-0.62
$[M+H]^{+\bullet}$	1865.858328	1865.8583	0.02
$[Sub\ P-NH3]^+$	1330.70897	1330.70886	0.08
$[C7\bullet-CH_4N_2]^+$	1372.62245	1372.598902	-0.19
$[C8\bullet-CH_4N_2]^+$	1519.69438	1519.667072	-0.33
$[M+H-CpMe_5]^+$	1730.73857	1730.740386	-1.05
$[M+H-Me]^+$	1850.83652	1850.834835	0.91
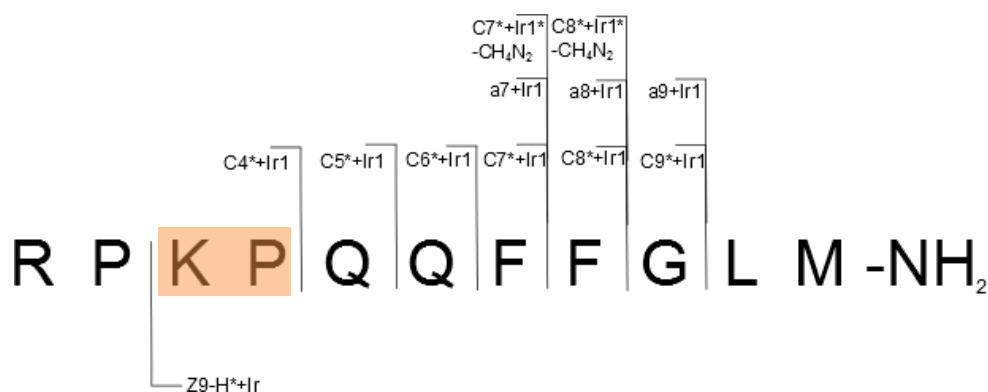
SI Figure 6.2: ECD MS/Ms spectrum, assignments and fragmentation map for the analysis of the [SubstanceP+Ir1-H2O+H]$^{2+}$ species

Fragment	Measured	Theoretical	ppm
[z3•+Ir2]$^+$	891.33643	891.33657	-0.16
[z6•+Ir2]$^+$	1118.46426	1118.464091	0.15
[z7+Ir2]$^+$	1303.53558	1303.53558	0.00
[z7•+Ir2]$^+$	1304.54389	1304.543401	0.37
[z8+Ir2]$^+$	1431.59435	1431.59416	0.13
[z8•+Ir2]$^+$	1432.60186	1432.601981	-0.08
[z9+Ir2]$^+$	1545.63665	1545.63709	-0.28
[z9•+Ir2]$^+$	1546.64474	1546.644911	-0.11
[z10+Ir2]$^+$	1602.65836	1602.65855	-0.12
[z10•+Ir2]$^+$	1603.66588	1603.666371	-0.31
[z11+Ir2]$^+$	1715.7425	1715.74261	-0.06
[z12+Ir2]$^+$	1843.83686	1843.83757	-0.39
[c12+Ir2]$^+$	1854.82214	1854.821796	0.19
[c13+Ir2]$^+$	1967.90497	1967.905856	-0.45
		Absolute average	0.20
		Standard deviation	0.23
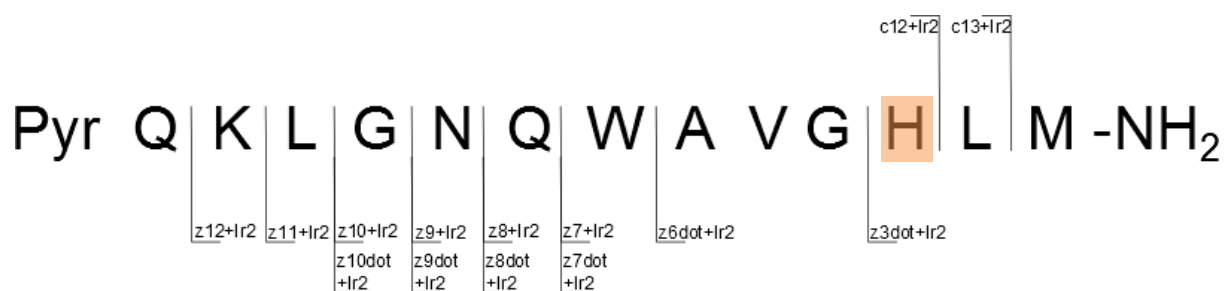
SI Figure 6.3: Assignments and fragmentation map for the ECD MS/MS analysis of the [L3BBS+Ir2+H]$^{2+}$ species.

Assignment	measured	theoretical	ppm
[b3+(Ir2-H$_2$O)-H]$^+$	857.313	857.313005	-0.01
[b4+(Ir2-H$_2$O)-H]$^+$	970.39702	970.397065	-0.05
[b5+(Ir2-H$_2$O)-H]$^+$	1027.41868	1027.418525	0.15
[b6+(Ir2-H$_2$O)-H]$^+$	1141.46129	1141.461455	-0.14
[b7+(Ir2-H$_2$O)-H]$^+$	1269.52003	1269.520035	0.00
[b8+(Ir2-H$_2$O)-H]$^+$	1455.59958	1455.599345	0.16
[b9+(Ir2-H$_2$O)-H]$^+$	1526.63602	1526.636455	-0.28
[b10+(Ir2-H$_2$O)-H]$^+$	1625.70466	1625.704865	-0.13
[b11+(Ir2-H$_2$O)-H]$^+$	1682.72557	1682.726325	-0.45
[b12+(Ir2-H$_2$O)-H]$^+$	1819.78479	1819.785235	-0.24
[b13+(Ir2-H$_2$O)-H]$^+$	1932.86903	1932.869295	-0.14
[b8+(Ir2-H$_2$O)-H+-NH$_3$]$^+$	1438.57279	1438.5728	-0.01
[b9+(Ir2-H$_2$O)-H+-NH$_3$]$^+$	1509.60922	1509.60991	-0.46
[b10+(Ir2-H$_2$O)-H+-NH$_3$]$^+$	1608.67813	1608.67832	-0.12
[b8+(Ir2-H$_2$O)-H+-CO]$^+$	1427.60488	1427.60443	0.32
[b9+(Ir2-H$_2$O)-H+-CO]$^+$	1498.64119	1498.64154	-0.23
[b10+(Ir2-H$_2$O)-H+-CO]$^+$	1597.71107	1597.70995	0.70
[b8+(Ir2-H$_2$O)-H+-CO$_2$]$^+$	1411.60967	1411.609516	0.11
[y5]$^+$	555.30719	555.30715	0.07
[y6]$^+$	626.3443	626.34426	0.06
[y7]$^+$	812.42356	812.42357	-0.01
[y8]$^+$	940.48218	940.48215	0.03
[y9]$^+$	1054.52498	1054.52508	-0.09
[y10]$^+$	1111.54644	1111.54654	-0.09
[y11]$^+$	1224.63048	1224.6306	-0.10
[y5-NH$_3$]$^+$	538.2806	538.280605	-0.01
[y6-NH$_3$]$^+$	609.31776	609.317715	0.07
[y7-NH$_3$]$^+$	795.39692	795.397025	-0.13
[y8-NH$_3$]$^+$	923.45555	923.455605	-0.06
[y10-NH$_3$]$^+$	1094.51992	1094.519995	-0.07
[y11-NH$_3$]$^+$	1207.60419	1207.604055	0.11
[b10+Ir2-H$_2$O-H]$^{+2}$	813.35599	813.3560725	-0.10
[b12+Ir2-H$_2$O-H]$^{+2}$	910.39616	910.3962575	-0.11
[b13+Ir2-H$_2$O-H]$^{+2}$	966.93832	966.9382875	0.03
[b13+Ir2-H$_2$O -H-NH$_3$]$^{+2}$	958.42511	958.425013	0.10
[b13+Ir2-H$_2$O-H-CO]$^{+2}$	952.94074	952.94083	-0.09
[b13+Ir2-H$_2$O -H-CO-NH$_3$]$^{+2}$	944.42731	944.4275555	-0.26
[L3BBS+(Ir2-H$_2$O)+ H+]2+	1040.97159	1040.971808	-0.21
[(L3BBS+(Ir2-H$_2$O)+ H+-NH$_3$]$^{+2}$	1032.45839	1032.458533	-0.14
[L3BBS+(Ir2-H$_2$O)+ H+ -2NH$_3$]$^{+2}$	1023.94508	1023.945259	-0.17
[y12+(Ir2-H$_2$O)-H]$^{+2}$	921.4265	921.426775	-0.30

CAD 25V		Absolute average =	0.15
		Standard deviation =	0.20

Si Figure 6.4: Assignments and fragmentation map for the CAD MS/MS analysis of the [L3BBS+Ir2-H$_2$O+H]$^{2+}$ species

Fragment	Expected mass	Observed mass	ppm
$[b3-CO+Ir1-H_2O]^+$	857.313005	857.31314	0.16
$[b3+Ir1-H_2O]^+$	885.30792	885.30792	0.00
$[b4+Ir1-H_2O]^+$	998.39198	998.3919	-0.08
$[b5+Ir1-H_2O]^+$	1055.41344	1055.41309	-0.33
$[b6+Ir1-H_2O]^+$	1169.45637	1169.45623	-0.12
$[b7-CO_2+Ir1-H_2O]^+$	1253.52512	1253.52463	-0.39
$[b7-CO+Ir1-H_2O]^+$	1269.520035	1269.51992	-0.09
$[b7-H_2O+Ir1-H_2O]^+$	1279.504385	1279.50412	-0.21
$[b7+Ir1-H_2O]^+$	1297.51495	1297.51479	-0.12
$[b8-CO_2+Ir1-H_2O]^+$	1439.60443	1439.60517	0.51
$[b8-CO+Ir1-H_2O]^+$	1455.599345	1455.59912	-0.15
$[b8-NH_3+Ir1-H_2O]^+$	1466.567715	1466.56765	-0.04
$[b8+Ir1-H_2O]^+$	1483.59426	1483.59429	0.02
$[b9-CO_2+Ir1-H_2O]^+$	1510.64154	1510.64096	-0.38
$[b9-CO+Ir1-H_2O]^+$	1526.636455	1526.63648	0.02
$[b9-NH_3+Ir1-H_2O]^+$	1537.604825	1537.60477	-0.04
$[b9+Ir1-H_2O]^+$	1554.63137	1554.63148	0.07
$[b10-CO_2+Ir1-H_2O]^+$	1609.70995	1609.71108	0.70
$[b10-CO+Ir1-H_2O]^+$	1625.704865	1625.70489	0.02
$[b10-H2O+Ir1-H_2O]^+$	1635.689215	1635.6895	0.17
$[b10-NH_3+Ir1-H_2O]^+$	1636.673235	1636.67331	0.05
$[b10+Ir1-H_2O]^+$	1653.69978	1653.7002	0.25
$[b11-CO+Ir1-H_2O]^+$	1682.726325	1682.72685	0.31
$[b11+Ir1-H_2O]^+$	1710.72124	1710.72187	0.37
$[b12-CO_2+Ir1-H_2O]^+$	1803.79032	1803.79105	0.40
$[b12-NH_3+Ir1-H_2O]^+$	1830.753605	1830.75434	0.40
$[b12+Ir1-H_2O]^+$	1847.78015	1847.78068	0.29
$[b13+Ir1-H_2O]^+$	1960.86421	1960.86419	-0.01
$[b12+Ir1-H_2O]^{2+}$	924.393715	924.39371	-0.01
$[M+H-NH_3]^{2+}$	1046.455993	1046.45595	-0.04
$[a12+Ir1-H_2O]^+$	1803.78978	1803.79159	1.00
$[y10]^+$	1111.54654	1111.54641	-0.12
$[y10-NH_3]^+$	1094.519995	1094.51987	-0.11
$[y11]^+$	1224.6306	1224.63046	-0.11
$[y11-NH_3]^+$	1207.604055	1207.60385	-0.17
Average error	0.06	RMS	0.09
Standard deviation	0.29	RMSD	0.30
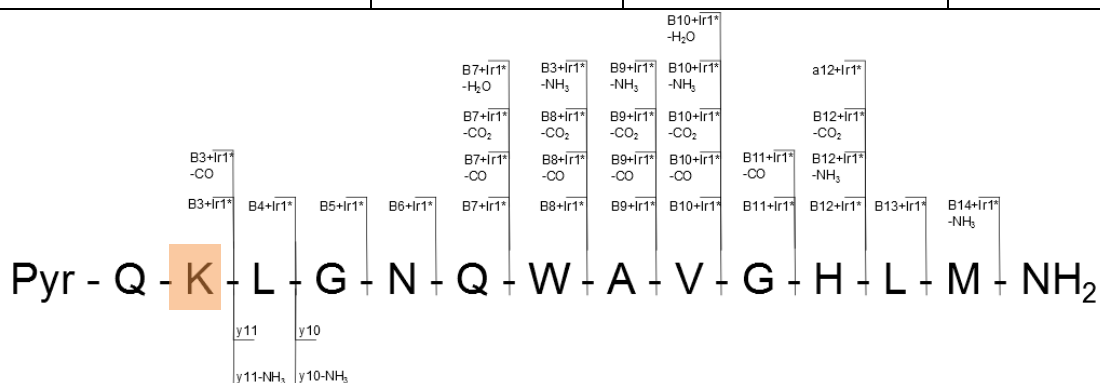
Si Figure 6.5: Assignments and fragmentation map for the CAD MS/MS analysis of the [L3BBS+Ir1-H$_2$O+H]$^{2+}$ species

Ir1-L3BBS 0.5-1X I1 ISO 1056-5 IRECD 002	Observed m/z	Theoretical m/z	Error/ppm
[z8]$^+$	924.46343	924.46346	-0.03
[z8•]$^+$	925.471255	925.47126	-0.01
[z9]$^+$	1038.50636	1038.50595	0.39
[z9•]$^+$	1039.514185	1039.51411	0.07
[z10•]$^+$	1096.535645	1096.5355	0.13
[z11]$^+$	1208.61188	1208.6119	-0.02
[z11•]$^+$	1209.619705	1209.62037	-0.55
[c8• -H+ Ir1-H2O]$^+$	1499.61298	1499.61322	-0.16
[c10• -H+ Ir1-H2O]$^+$	1669.7185	1669.71832	0.11
[z12-H$^+$ +Ir1-H2O]$^+$	1853.822465	1853.82123	0.67
[M+Ir1-H2O-NH$_3$]$^{+•}$	2108.9307	2108.93071	-0.005
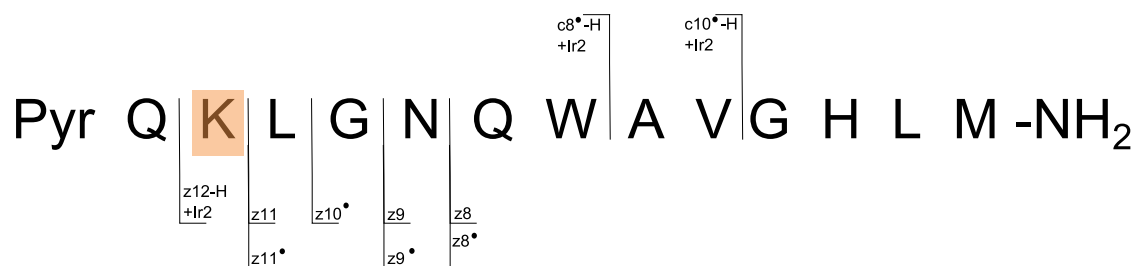
SI Figure 6.6: Assignments and fragmentation map for the IRECD MS/MS analysis of the [L3BBS+Ir1-H$_2$O+H]$^{2+}$ species

Chapter 7:

Super-acid supercharging using CH_5^+

The effective characterisation of any molecule by MS requires effective ionisation, which can be particularly challenging for some analytes and/or conditions. Particularly for biomolecule characterisation the challenge not only lies in ionising large biomolecules, but transferring enough charge to each species to observe and interrogate high charge-state ions. Subsequent MS/MS experiments on biomolecules of interest are heavily affected by the charge state of the parent ion, with higher charge states always being desirable for their increased MS/MS cleavage efficiency, increased sequence/cleavage coverage of fragmentation, and increased sensitivity on all mass analyser platforms. It is therefore of ubiquitous benefit to make progress into achieving these high charge states for effective biomolecule and modified-biomolecule characterisation. One thriving field of research is the investigation of solution-phase additives to alter the properties of sample solutions, usually via changes in surface tension, to affect the resulting droplets produced during electrospray ionisation-MS. The changes to droplet characteristics have been shown to alter the charge states of solvated biomolecules, sometimes producing higher charge states. This process of solution-phase "supercharging" of analytes has shown great promise over the years and has been improved upon greatly with a range of additives now available, the majority of which achieve their effect via increasing the surface tension of the ESI droplets formed, thus enhancing charge of droplets and the resulting protein analytes solvated within.

Though solution-phase additives have been shown to be very effective at enhancing analyte (and particularly protein) charge states, resulting in corresponding benefits to MS and MS/MS analysis, they are unfortunately not consistent benefits for all samples, each supercharging agent seems to be effective for certain samples, but not others, they also require optimisation for the amount and solvent conditions they are used with, and they require extra sample preparation steps,

which although not extensive can have negative side effects with some species. As a result it has been an area of great interest to not only find new species/ways of supercharging analytes, but also to see if supercharging may be achievable without solution phase additives, and to achieve enhanced charging through instrumentation improvements instead of chemical ones.

The content of this chapter focuses on enhancement of analyte charging through a novel ion source, a marriage of electrospray ionisation and atmospheric pressure chemical ionisation, which utilises one of the strongest acid species known (CH_5^+) to charge droplets produced during ESI and achieve supercharging without solution phase additives and without modification of analytes. The source and subsequent supercharging of a range of analytes in various conditions are presented, along with a detailed discussion on the underlying mechanism of the supercharging concerned, most notably the repulsive force between the positively-charged ESI droplets formed and the incoming CH_5^+ ions.

Although the results presented within this chapter are predominantly based on FT-ICR MS and results from the thesis Author, this was a collaborative project and some elements were conducted by collaborating researchers, namely; Haytham Hussein and Cookson Chiu assisted with the design of the supercharging source at stages of development, and creation of various Figures related to the source and mass spectra. Lewis Baker, Scott Habershon, and Anthony Stace were collaborators in the investigation of the calculation of electrostatic forces between ESI droplets and CH_5^+ ions.

The work presented within Chapter 7 was used to file a patent for the supercharging of analytes using CH_5^+ and ESI-MS, the details of which are below:

"IONISING MOLECULES AND ELECTROSPRAY IONISATION APPARATUS"

Peter B. O'Connor, **Christopher A. Wootton**, Haytham Hussein, Man-ying Wong, 2016

Super-acid supercharging using CH$_5^+$

Christopher A. Wootton[1][†] and Haytham E. M. Hussein[1][†], Man Ying Wong[1], Cookson K.
C. Chiu[1], Lewis A. Baker[1], Scott Habershon[1], Anthony J. Stace[2], Mark P. Barrow[1], Peter
B. O'Connor[1*]

† These authors contributed equally to the work presented

1. Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry
   CV4 7AL, United Kingdom

2. Department of Physical and theoretical Chemistry, School of Chemistry,
   University of Nottingham, University Park, Nottingham, NG7 2RD, United
   Kingdom

p.oconnor@warwick.ac.uk

CORRESPONDING AUTHORS
Peter B. O'Connor, Department of Chemistry, University of Warwick, Gibbet Hill Road,
Coventry CV4 7AL, UK
PBO'C: phone: +44 (0)24 76151008; fax: +44 (0)24 76151009;
email: p.oconnor@warwick.ac.uk;

Abstract:

Supercharging of protein ions can be achieved by reaction with the super-acid CH_5^+. Analyte ions are supercharged in the gas phase immediately before entering the mass spectrometer inlet via reaction with super-acid CH_5^+ ions produced adjacent to the ESI plume; the method enabled enhanced charging of proteins, producing charge states previously unobserved for the non-supercharged samples. Supercharging of Native MS protein samples also produced a 500% increase in intensity for the most abundant charge states observed, while retaining native MS conditions. Super-acid supercharging was also shown to increase the intensity of singly charged small molecules and small molecule dimers by 22% and 291% respectively, and was shown to preserve the analyte from modification, as no methyl adduction was observed. The experiment was implemented by interfacing a standard electrospray ion source with a custom-built glow discharge chamber at the FT-ICR mass spectrometer inlet. The location of CH_5^+ creation and interaction with the ESI plume was shown to be critical to achieving ESI stability and effective supercharging.

Electrostatic repulsions between CH_5^+ and the charged droplets created during electrospray ionization cause a Coulombic repulsion force, hindering interaction between the two species. Taking a purely Coulombic force approach this interaction is shown to be always repulsive and will tend towards infinity as an incoming CH_5^+ ion approaches a charged droplet, preventing proton transfer, and the positive supercharging effect. A modified electrostatic model was then applied to take into account the polarizable nature of solvent molecules and their ability to re-orientate to cause ion-induced dipole polarizabilities creating an attractive force between the charged droplet and incoming

330

CH_5^+ ions, at very small distances and lowering the charge repulsion "barrier". The modified electrostatic model showed an attractive force was possible at low charge densities. However this attractive force dissipated quickly at higher charge densities and the resulting force calculations approached that of a purely mean-field model. Though the attractive force is believed to be a contributing factor in enabling CH_5^+ ions to overcome the coulombic barrier and allow transfer of additional protons to ESI droplets, allowing the supercharging effects observed in experiment, it is clear that other forces (such as turbulent motion and high-velocity CH_5^+ ions) are likely to be the deciding factors in enabling the supercharging process.
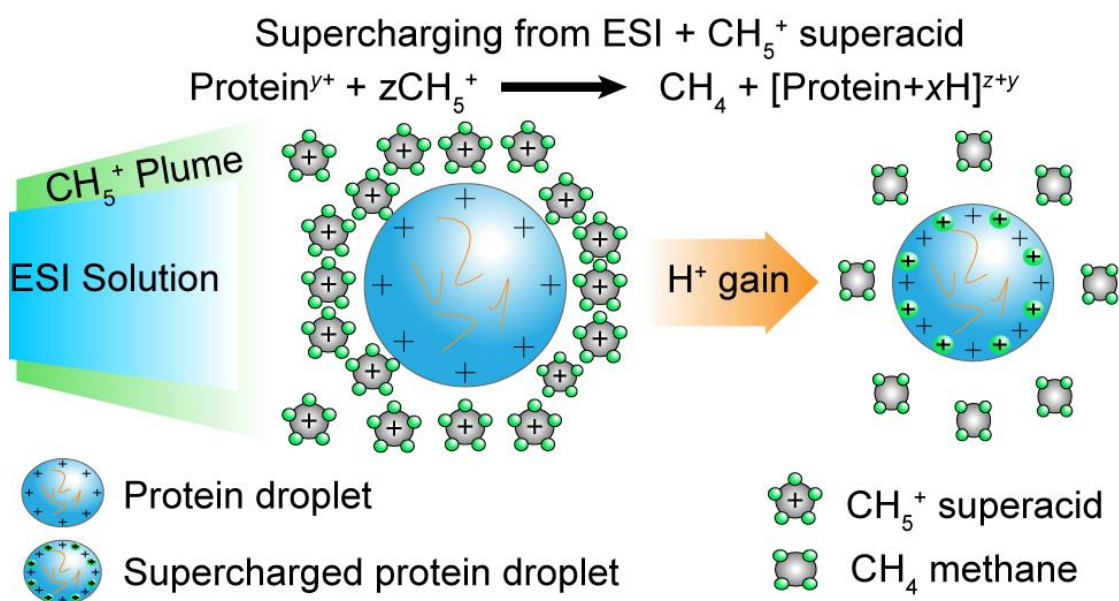
Introduction:

The CH_5^+ cation is considered a "super-acid" as it has an extremely low proton affinity, even lower than that of sulfuric acid (531.4 vs. 699.6 kJ/mol, respectively).[1] CH_5^+ was first introduced as an ionization reagent by Talrose *et al.*[2,3] and subsequently utilized

extensively by Munson and Field, 1966,[4,5,6,7] a method later referred to as "Chemical ionization" (CI) which was capable of producing even-electron $[M+H]^+$ ions from analytes, rather than the fragile and unstable $[M]^{+\bullet}$ ions produced by the widely used electron ionization (EI).[8] CH_5^+ can be readily produced by reaction with high-energy electrons, as shown in equations 1&2.

$$CH_4 + e^- \rightarrow CH_4^{+\bullet} + 2e^- \hspace{4cm} \text{[Eq.1]}$$

$$CH_4 + CH_4^{+\bullet} \rightarrow CH_5^+ + CH_3^{\bullet} \hspace{3.5cm} \text{[Eq.2]}$$

CH_5^+ ions are the primary, stable reaction products from equations 1&2 and are thus available for further reaction with target analytes. CH_5^+ ions act as extremely strong BrØnsted acids (the so called "super-acid") which readily transfer protons to nearby species within range, producing a protonated analyte species $(M+H)^+$ as shown in Equation 3.[9]

$$CH_5^+ + M \rightarrow [MH]^+ + CH_4 \text{ (proton transfer)} \hspace{2cm} \text{[Eq.3]}$$
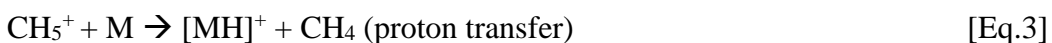
Although chemical ionization greatly assisted in the analysis of many polar and non-polar chemicals, it was not a viable ionization method for the study/interrogation of large biomolecules as it involves a gas-phase proton-transfer reaction, meaning CI is of limited use when attempting to ionize non-volatile molecules (such as peptides and proteins). CI also generally only produced singly charged products[10] which limited its use to smaller species of interest. With the advent of electrospray ionization (ESI and nanoESI), mass

spectrometry analysis of large biomolecules, such as multiply charged protein ions became viable.[11,12]

The introduction of the supercharged biomolecules concept by Iavarone *et al.*, 2001, showed that enhanced charging of analytes can be achieved, allowing observation of higher charge states for analytes via ESI than previously observed.[13,14] Addition of supercharging reagents such as glycerol, sulfolane, or m-Nitro benzyl alcohol increases the maximum observed charge states, with increased signal intensity of biomolecules and lower mass-to-charge ratios, allowing for increased achievable resolving powers for FTMS based instruments (such as FT-ICR MS and Orbitrap platforms).[13,15–18,19] Accessing higher charge states has also been shown to dramatically improve fragmentation efficiency and cleavage coverage during MS/MS experiments,[20] increasing with higher charge-states of selected precursor ions,[21] especially for electron capture dissociation and electron transfer dissociation MS/MS.[18] Enhancing the charge state enables a higher degree of fragmentation, thus more molecular sequence information, which is critical for effective identification of unknown proteins and peptides via MS/MS fragmentation and the analysis of any proteomic-based sample.[22,23] Studies of post-translational modifications could also be facilitated by this increase in fragmentation efficiency, especially in the ever-growing field of Top-Down MS/MS.[24,25]

Other supercharging methods and reagents have been explored; in late 2007, Kjeldsen *et al.* introduced supercharging reagents (m-NBA) to the LC mobile phase of LC-ESI-MS system.[21] Sterling *et al.*[26] studied the effect of DMSO as supercharging reagent and the relationship between the supercharging reagents, analyte solution and protein

conformation.[26–28] In May 2012, Miladinovic *et al.* demonstrated the introduction of a supercharging reagent into the ESI Taylor cone using a dual-sprayer ESI microchip.[18] Recently, Teo and Donald showed that solution additives, ethylene carbonate, propylene carbonate, and *o*-nitroanisole could be used to enhance the charging of proteins and form significantly higher protonation states.[29]

Despite the exciting efforts to increase the observed charge state of protein analytes, the mechanism of supercharging remains unclear.[18,17,30,31,32] Iavarone *et al.* outlined that the surface tension of the electrospray droplet is a major factor in determining the analyte charge,[16] adding high surface tension solvents with low vapor pressure to the analyte will increase the surface charge density of the droplet due to the significant increase in the concentration of the supercharging reagent in the ESI droplets while evaporation occurs; as droplets undergo evaporation and reach the Rayleigh limit, Equation 4, the electrospray droplet will undergo Rayleigh fission and, due to increasing surface tension, increase charge density leading to an enhanced protein charge state.[15,16]

$$Z_R e = 8 \pi (\epsilon_0 \gamma R^3)^{1/2} \qquad \text{[Eq. 4]}$$

From Equation 4, the Rayleigh charge limit theory indicates the direct proportionality between the surface tension (γ) and the charge availability of the droplet.[33]

The investigation of ion supercharging efficiency displayed appears to be protein dependent and the supercharging process may depend on a conformational change of the analyte proteins as an unfolded protein could accommodate more charges than the folded conformer, as proposed by Williams *et al.*[27,28] Other research by Loo *et al.* provided claims to the contrary and pointed out the role of proton transfer reactivity in supercharging from aqueous solution.[31] Overall, many factors such as solvent surface

tension, gas phase and solution phase basicity of both analyte and solvent, charge

competition between molecules of analyte and solvent play a major role in the extent of

charging and ion charge distribution.[17,26–28]

The work presented herein shows supercharging, through a direct reaction between

methane super-acid, CH_5^+ ions (BrØnsted acid), with the sprayed solution droplets of the

analyte at atmospheric pressure during the ESI process. Due to their inherent strong

protonation capability, and extremely low proton affinity, CH_5^+ "super-acid ions" are

ideal species for exploring the combination of ESI and gas-phase supercharging via APCI

generated reagent ions. Enhanced droplet charge translates to increased protonation of

charge sites in proteins and increased intensity and/or higher observed charge states of

analyte species. Potentially, supercharging MS analytes with superacid CH_5^+ ions will aid

the analysis of slightly polar compounds, large and small biomolecules (such as peptides

and proteins) and enable increased cleavage coverage during subsequent MS/MS

experiments which can benefit Top-down protein analysis and the entire Bottom-Up

proteomics field as it is possible to readily hyphenate this concept to liquid

chromatography. Computational models are utilized and explored to help understand how

this proton transfer between the superacid CH_5^+ and charged droplets may occur.

Materials, design, and methods:

Materials and chemical reagents

Bovine heart Cytochrome c, chicken egg-white Lysozyme, and Equine heart Myoglobin were obtained from Sigma-Aldrich, UK. The samples were prepared by dissolving the proteins in a solvent system that consists of either purified water 18 MΩ cm^{-1} only, 50/50 %, v/v water/methanol (for denatured protein studies), or aqueous solutions of 200 mM ammonium acetate (for native MS studies). Methane reagent gas was purchased from CK special gases (UK). L-(+)-ascorbic acid ($C_6H_8O_6$), was purchased from Alfa Aesar (Heysham, Lancashire, UK). High purity grade N_2 (99.99 % BOC, UK) was used as the nebulizer gas for ESI and APCI-ESI experiments. Ultrapure water (18.2 MΩ cm^{-1}) was obtained from a Direct-Q water purification unit (Millipore, Watford, UK) and LC-MS grade methanol (VWR Chemicals, UK) were used in the sample preparation. All reagents were used as received.

ESI-APCI Supercharging Ion Source:

The experimental setup is shown in Figure 7.1. The ion source was constructed by interfacing a modified Analytica ESI source with a home-built super-acid glow discharge chamber to produce CH_5^+ ions.
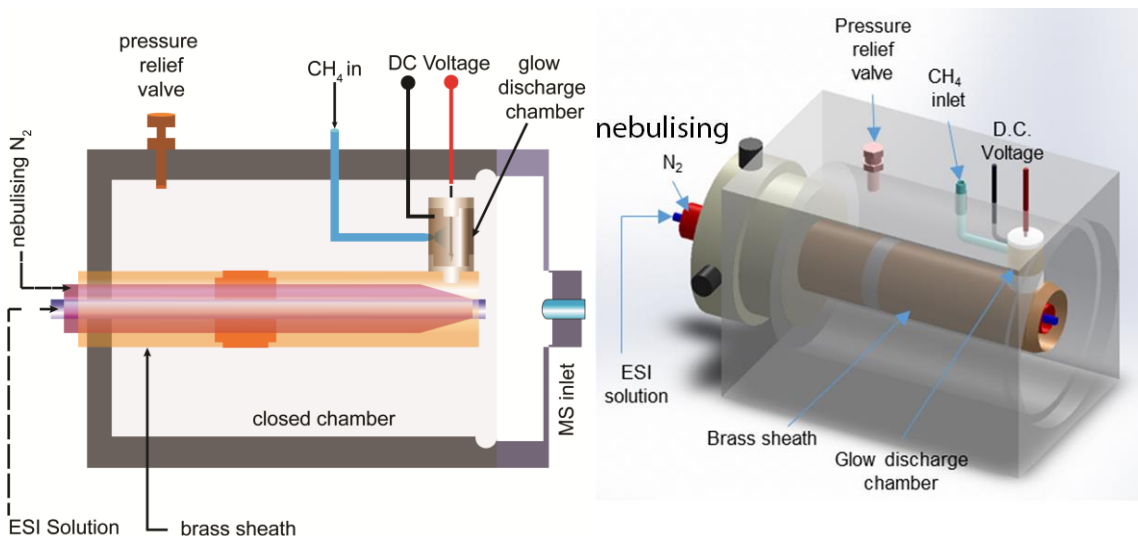
Figure 7.1: Schematic of the ESI-APCI supercharging source (left) and 3D representation (right).

The super-acid CH_5^+ is produced by ionizing a stream of methane gas flowing through a potential difference region between a needle and the exit of a custom brass glow-discharge chamber to the ESI needle tip. The glow-discharge chamber consists of a tailor-made threaded brass cylinder isolated from both the discharge needle and the ESI sheath via insulating Teflon caps. The tapered shape along the inside of the chamber and the position of the discharge needle ensures the flow of methane is focused into the exit of the chamber, where the electrical discharge occurs, improving the efficiency of CH_5^+ production, which occurs just 6 mm from the ESI needle tip. The glow discharge chamber was connected to a tailor made brass sheath (length 58 mm, O.D. 9.55 mm, I.D. 5.44 mm), fitted with a 15 cm ESI stainless steel capillary (I.D. 0.20 mm, O.D. 0.22 mm,

Bruker Daltonik, Bremen, Germany), housed inside a stainless steel tube (length 68 mm, O.D 1.25 mm., I.D 1.20 mm.). The distance between the tip of the ESI needle and the mass spectrometer capillary inlet was 1.7 cm. Methane was introduced from the top of the source and connected to the glow discharge chamber using PTFE tubing. A pressure relief valve (1/4 inch, VRV 250, Generant, Butler, NJ, USA) is used to regulate the pressure within the source chamber (~5 psi), and to purge O_2 from the chamber prior to operation in order to avoid igniting the methane gas.

Analyte solution was delivered to the ESI needle via a 250 μL syringe connected to a PTFE transfer line and driven via a syringe pump at a rate of 100-200 μL/hour. The ESI needle was biased from +4000 to +6000 V to achieve stable ESI. A voltage difference of +3500 V was applied to the glow discharge system via an external high-voltage power supply (Applied Kilovolts, KS20/43, Exelis, West Sussex, UK), input power was provided via a 24 V power supply (EL301R, TTI, Huntingdon, UK). A cylindrical mesh surrounded the ESI assembly and was held at positive potential to focus the ions, allowing the improved interaction between the generated super-acid CH_5^+ and the sample solution in the gas phase and preventing charge build up on the source walls. Finally, a home-built Aluminum flange connects the super-charging source to the inlet of the FT-ICR mass spectrometer.

The effect of super-acid CH_5^+ as a supercharging reagent depends on geometry, pressure, flow rate, and the potential difference between the ESI needle and mass spectrometer inlet. Different pressures/flow rates of methane did not show appreciable differences on the supercharging effect past 0.3 bar above atmosphere, at which point ESI and instrument performance began to degrade due to high source pressure. The distance

between the tip of the ESI needle and origin of super-acid CH_5^+ generation played a major role in the supercharging process.

FT-Ion Cyclotron Resonance Mass Spectrometry:

All experiments were conducted on a 12 Tesla SolariX Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR MS) (Bruker Daltonik GmbH, Bremen, Germany), all mass spectra were collected in positive ion mode. Supercharging experiments were conducted using the modified APCI-ESI source described above, varying the ESI high voltage, methane gas flow, and glow-discharge high voltage. Control spectra conducted with methane flowing through, but without the glow discharge external high voltage yielded the same mass spectral results as the normal ESI control. The FT-ICR instrument was tuned to scan 147 to 3000 m/z and 10-20 spectra were signal-averaged per run. Ions were externally accumulated in a hexapole collision cell for 0.1-0.2 s before being transferred to the infinity cell[34] for excitation and detection.

Results and Discussion:
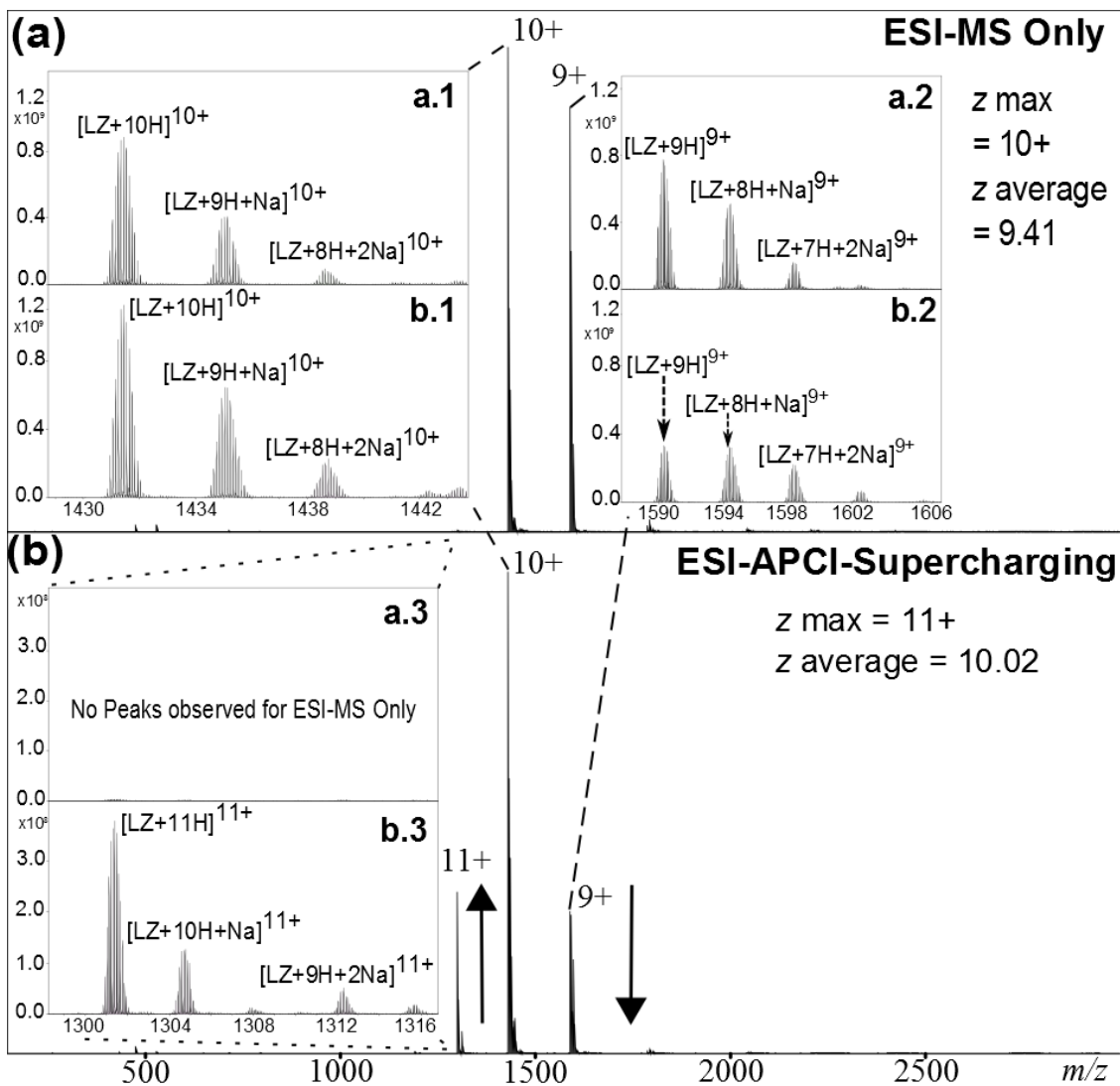
Figure 7.2: a) Normal ESI-mass spectrum of 1 μM aqueous Lysozyme (LZ) solution, showing 2 main charge states; 8⁺, 9⁺, and a minor charge state of 10⁺. b) Supercharged ESI-APCI-mass spectrum of the same 1 uM aqueous Lysozyme solution, showing increased analyte charging, main charge states observed now being 9⁺, 10⁺, and 11⁺.

Aqueous protein solution analysis:

Figure 7.2a shows the ESI-FT-ICR MS spectrum of 1 μM Lysozyme dissolved in water, a very narrow range of charge states was observed (8^+, 9^+, and 10^+), correlating with Konermann *et al.*[35]

Figure 7.2b shows the same protein-containing solution ionized with the supercharging ESI-APCI source active. The supercharged lysozyme spectrum shows the 11^+ charge state of the protein appearing in good intensity and a decrease of lower charge states (8^+ and 9^+), focusing ion populations into the higher charge states. The weighted average charge state for each spectrum was calculated via equation 5 (below) where N is the number of observed charge states ith, q_i is the charge value and W_i is the signal intensity.[14,36] The calculations show a distinct increase in average charge from 9.41 (ESI-MS) to 10.02 (ESI-APCI-MS). The incorporation of APCI derived CH_5^+ ions has clearly enhanced charging of the protein analyte species.

$$q_{average} = \sum_i^N q_i\, W_i \, / \sum_i^N W_i \qquad\qquad \text{[Eq. 5]}$$

Figure 7.2 insets (a.1-3 and b.1-3) show the protonated lysozyme species from both the ESI-MS and the ESI- APCI-MS spectra, along with the sodium adducts. The comparison of the 2 spectra shows that no change/modification of the analyte species has been observed (e.g. via methylation, CH_5^+ adduction, or reaction with any other intermediates from equations 1-5 (above)), the ESI-APCI-MS ionization method is shown to produce a similar distribution of protonated and sodium-adduct peaks, only with the benefit of enhanced charge states.

Denatured protein supercharging:

Figure 7.3a and 7.3b show the ESI and ESI-APCI-mass spectra of denatured myoglobin (a 16.9 kDa oxygen transport protein)[18] electrosprayed from a 50:50 v/v water:methanol solution. Denatured Myoglobin shows a range of charge states upon ESI-MS ranging from 10^+ to 17^+ (Figure 7.3a). The denatured myoglobin ionized via the supercharging ESI-APCI source attained much higher charge states (see inset), previously unobtainable without supercharging. Denaturation of Myoglobin also causes loss of the non-covalently-bound Heme group, however ESI-MS does not show a detectable Heme species released. The ESI-APCI source shows increased signal of the $[Heme+H]^+$ ion released upon Myoglobin denaturation, allowing it to be detected. The normal ESI-MS spectrum of denatured Myoglobin showed a weighted average charge of 14.5+, a base charge state of 15^+, and a max charge state of 18^+. The supercharged ESI-APCI-MS spectrum of Myoglobin showed an average charge of 16.7+, a base charge of 17^+, and a maximum charge of 25^+. The high abundance peaks observed in Figure 7.3 were assigned to the apo form of Myoglobin, produced due to denaturing conditions, producing the bimodal charge state distribution typical of denatured protein ESI-MS. Low levels of holo-Myoglobin were also visible, along the baseline.
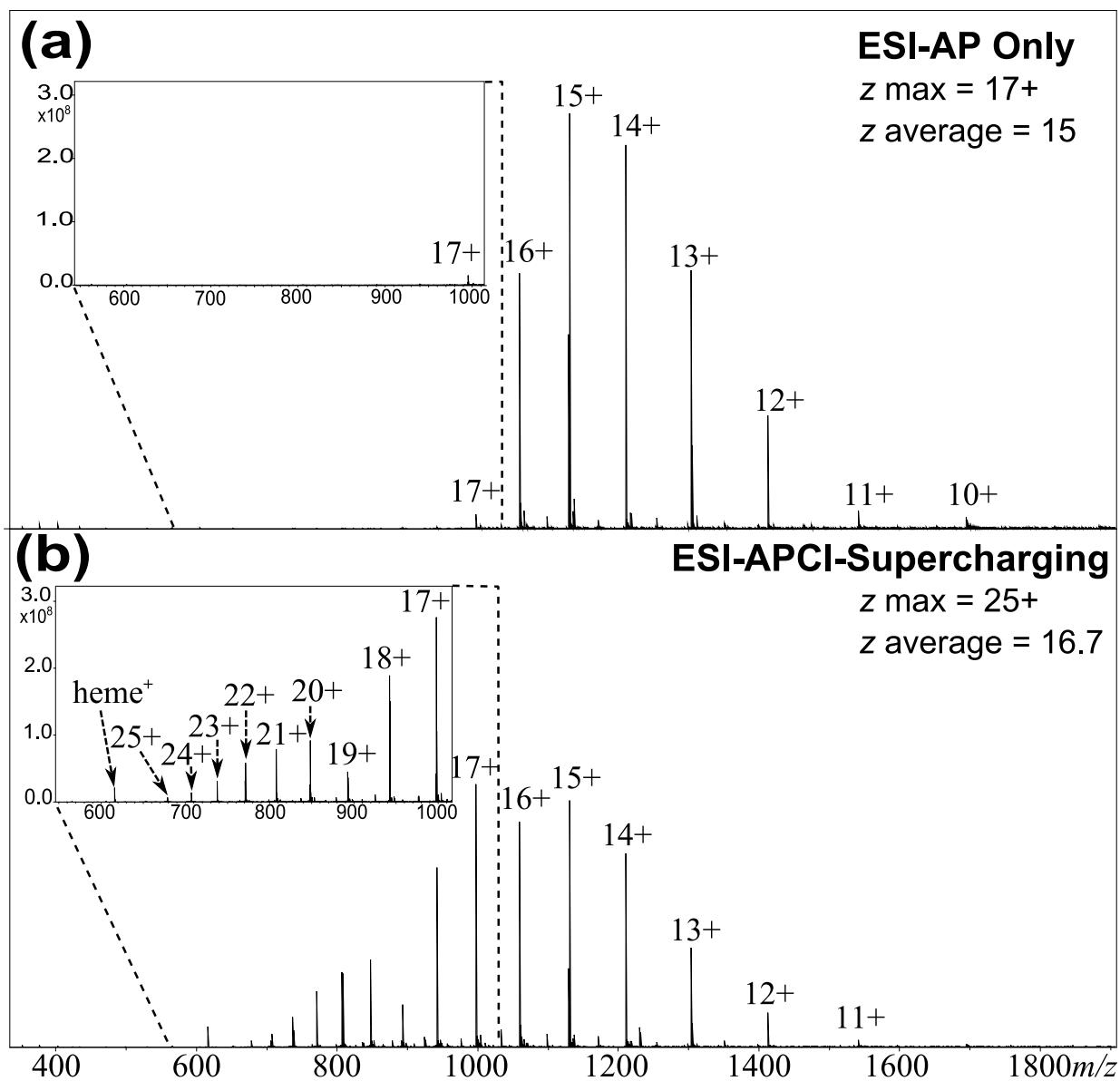
Figure 7.3: ESI-MS of 1 μM Myoglobin in a water-ACN solution (a) and supercharged ESI-APCI-MS spectrum of the same Myoglobin-water-ACN solution, showing enhanced charging of analyte species.

The ability to access these increased charge states for protein analytes can have a drastic effect on the quality of MS/MS spectra and sequence information produced.[18,20,37,38] McLafferty *et al.*[20] showed that each individual charge state of proteins

343

can produce different MS/MS fragments during electron capture dissociation (ECD), and that higher charge states provided much more sequence information than lower ones, this effect has also been observed for the more commonly used collision based dissociation (CAD/CID).

Though other chemical-based supercharging methods have been developed, and successfully implemented, in many studies,[15,18,36,31] the ESI-APCI source presented here represents a new method of achieving these charge states without the addition of chemical additives to the sample solution (process shown in Figure 7.4), and thus avoiding unwanted side reactions/changes in chemical equilibria which can occur upon addition of new species to solution chemical mixtures. In addition, use of the ESI-APCI source requires no further sample preparation steps, as each sample is simply introduced in the same way as a normal ESI-MS sample.
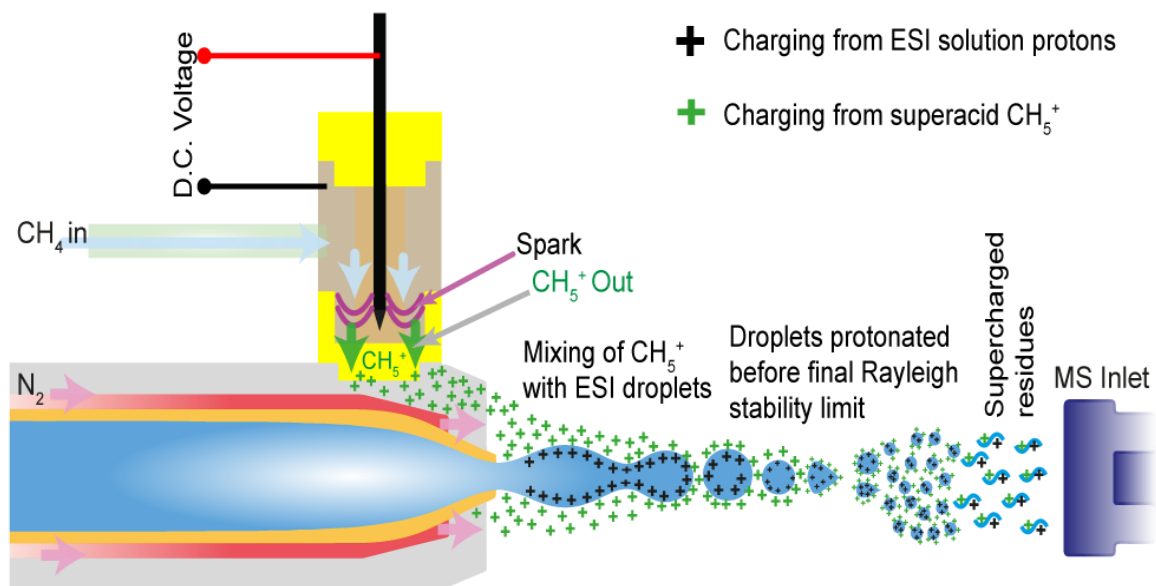
Figure 7.4: Representation of the ESI-APCI supercharging processes – CH₄ gas was ionized via a glow discharge chamber in close proximity to the ESI needle tip,

enabling the newly created CH$_5^+$ ions to interact with ESI plume and droplets and enhance charging of analytes.

Supercharged-Native MS of proteins:

There is an increasing interest in attempting to retain protein structure during ionization and mass spectrometry analysis, called Native-MS.[25,39] Usually involving the use of stabilizing buffers and possibly salt-containing solutions, Native MS attempts to keep proteins/analytes folded in their solution conformation as much as possible during analysis, thus charge state distributions,[40] surface labelling strategies (such as HDX[41] and FPOP[42]), and careful MS/MS experiments[40,43] can then inform and interrogate tertiary + quaternary protein structure via mass spectrometry.

Unfortunately the use of high-concentration buffer solutions and/or salt solutions are detrimental to effective ESI, desolvation, and protein MS analysis as a whole. As a result Native MS usually produces low signal to noise spectra (due to ineffective desolvation), with protein species observed with many adducted salt/buffer ions (decreasing signal to noise further as analyte ions are distributed into many signal channels), and very low charge states compared to water/denatured conditions. The latter of these disadvantages decreases observed intensity of ions (as all MS techniques are inherently more sensitive for higher charge state ions), and dramatically impedes MS/MS fragmentation characterization and subsequent sequence coverage.[25] Therefore effective characterization of proteins in their native state is particularly challenging, and would benefit greatly from the enhanced charging effects supercharging MS can offer. As a result Native MS of Lysozyme in 200 mM ammonium acetate buffer was conducted

using the Supercharging ESI-APCI source. The results, shown in Figure 7.5, show that

Native MS of the Lysozyme protein presents a narrow charge state distribution, at lower

charge states than the water solution above (Figure 7.2a). Lysozyme in the 7^+, 8^+, and 9^+

charge states were observed at relatively low signal to noise ratio.
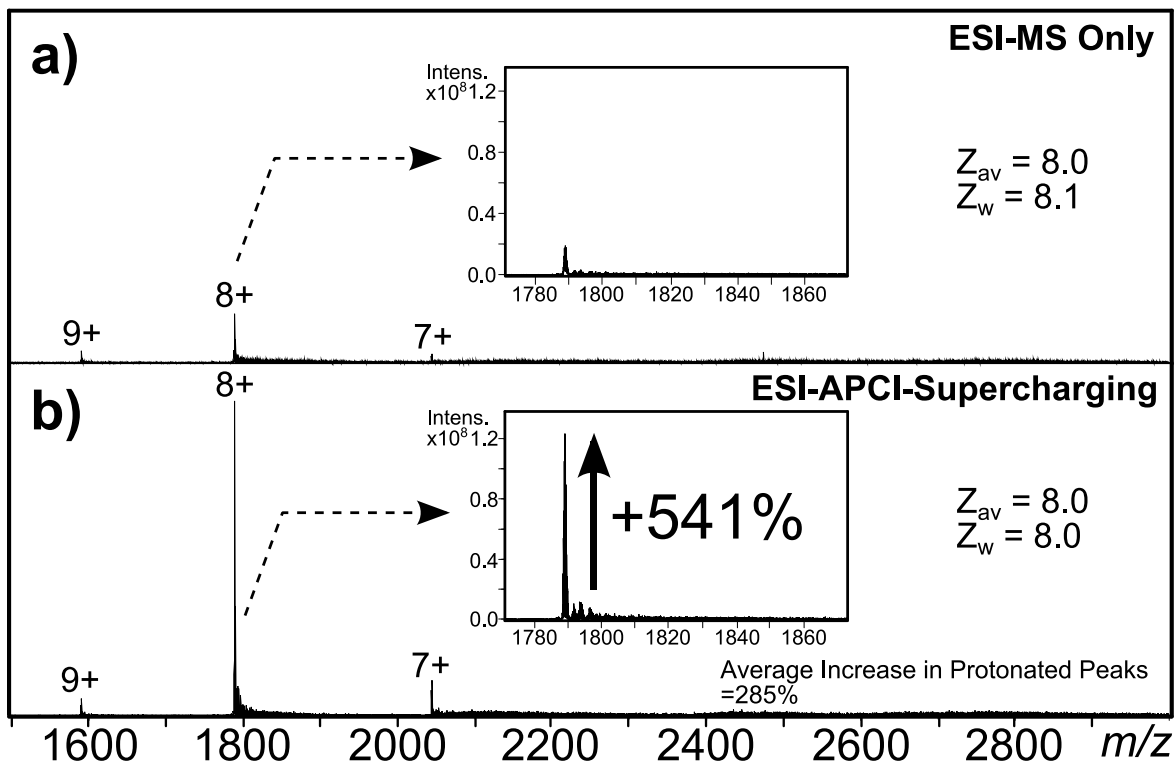
Figure 7.5 – Native ESI-MS of Lysozyme a) and supercharged Native-MS of the same sample (b), showing a vast increase in signal of lysozyme peaks due to enhanced charging.

ESI-APCI of the same solution of Lysozyme (aqueous 200 mM ammonium acetate

buffer), shows the effect of enhanced supercharging on the sample; the observed intensity

of the Lysozyme peaks has increased by on average 285%, and >500% increase in signal

intensity for the 8^+ isotope distribution, a dramatic increase over ESI-MS. The same

charge states are observed in both ESI-MS and supercharging ESI-APCI-MS, despite the

obvious enhanced charging, suggesting that, in its native conformation, the available protonation sites of lysozyme have been fully protonated, with multiple charge states being observed due to subtly different conformations, as has been suggested previously by Barren *et al.*[40]

The large increase in both intensity and signal to noise for the native lysozyme species is of great benefit for both detection and accurate characterization of the protein species by MS, and should be readily applicable for other protein species in the future, and aid in subsequent Top-Down MS/MS analysis.

<u>Small molecule supercharging:</u>

In addition to biomolecule supercharging, small molecule ionization can vary greatly depending on the chemical nature of the groups within the analyte, with some (particularly non-polar) species being particularly difficult to ionize effectively via ESI.[44] As a result, the effect of the ESI-APCI source on small molecule ionization was investigated using ascorbic acid. Ascorbic acid ($C_6H_8O_6$) is a small (176.032088Da) molecule, a form of vitamin C, believed to provide anti-oxidant activity in the body.[33] Figure 7.6a and 7.6b show the ESI-MS and ESI-APCI-MS spectra of Ascorbic acid. The main species observed in the ESI-MS of Ascorbic acid were the protonated and sodiated forms of Ascorbic acid ($[C_6H_8O_6+H]^+$ and $[C_6H_8O_6+Na]^+$), along with a high intensity peaks for the protonated and sodiated Ascorbic acid dimer ($[C_6H_8O_6\ C_6H_8O_6+H]^+$ and $[C_6H_8O_6\ C_6H_8O_6+Na]^+$).
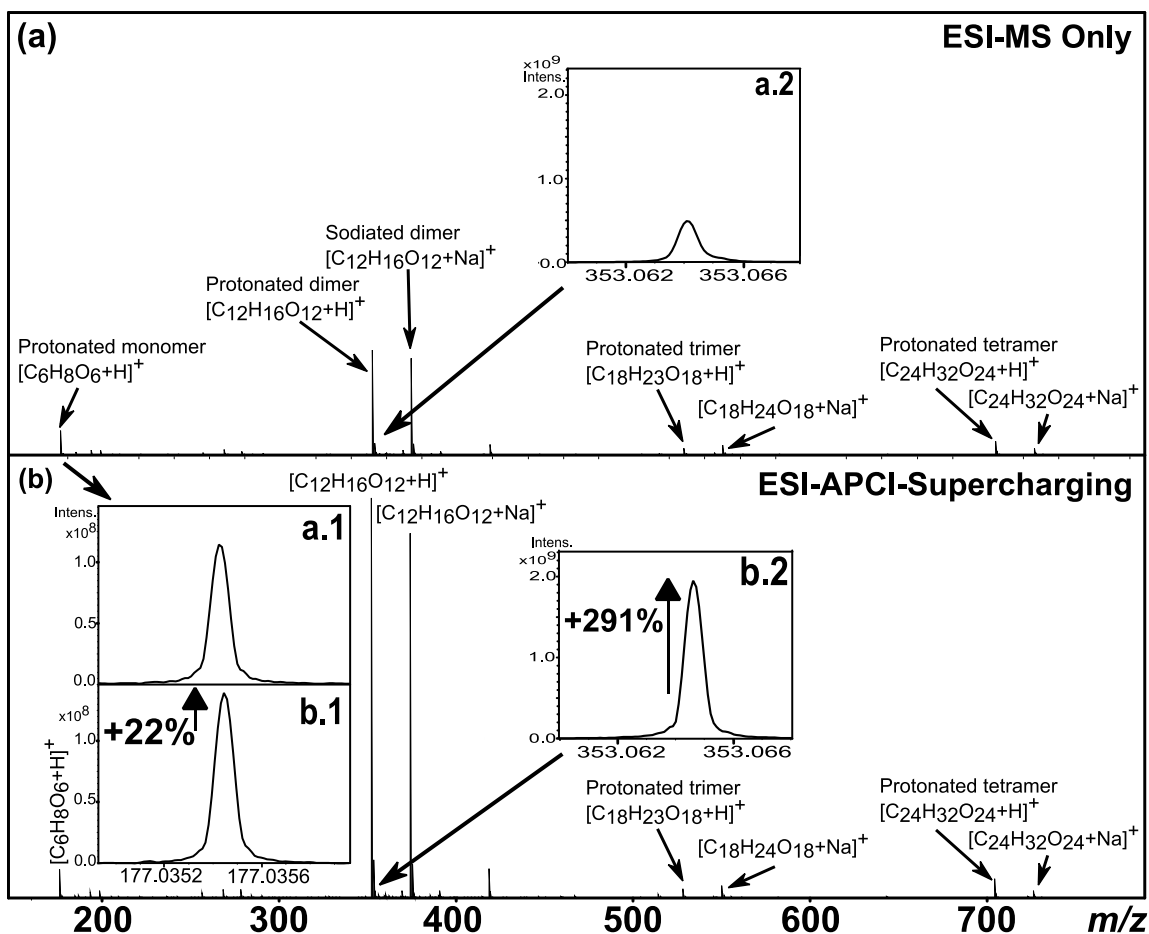
Figure 7.6: ESI-MS of the Ascorbic acid solution (a) and ESI-APCI-MS spectrum of the same Ascorbic acid solution, showing increased signal intensity due to enhanced charging of protonated Ascorbic acid (22% increase) and the Ascorbic acid dimer (291% increase).

The ESI-APCI-MS spectrum is shown in figure 7.5b, though no peak corresponding to doubly charged Ascorbic acid $[C_6H_8O_6+2H]^{2+}$ was observed, the singly protonated ascorbic acid species was shown to increase in intensity by 22%, with respect to ESI only (Figure 7.6b, inset a.1 and b.1), this increase is believed to be a result from the increased proton transfer from the CH_5^+ reagent gas to the droplet. Figure 7.6b (inset a.2 and b.2)

348

also shows a much larger increase in observed intensity for the ascorbic acid dimer species of 291%, which is a substantial increase over traditional electrospray.

It appears that not only can CH_5^+ assist in attaining higher charge states for biomolecules, but also increase the ionization efficiency of small molecules, on average peak intensities increased by 103% (and up to 291% for some species), this could lead to improved limits of detection (LOD) for small molecule species and enable more sensitive MS detection of key compounds during analysis where low LOD's are crucial, e.g. trace organic analysis, forensic testing, trace heavy metal quantitation etc.

Combinations of charge-enhancing techniques:

Since the Superacid-supercharging ESI-APCI Ion source is not a solution-phase supercharging reagent, it can be readily combined with current solution-phase procedures for enhancing analyte charge and supercharging reagents. The Superacid supercharging ion source was used to ionize myoglobin protein samples containing denatured, acidic, and/or chemically supercharged solvent systems, in order to explore the potential benefits of combining multiple techniques and any possibility of enhancing the charging of analytes even further than previously observed. The results for the supercharging combination experiments are shown below in Figure 7.7:
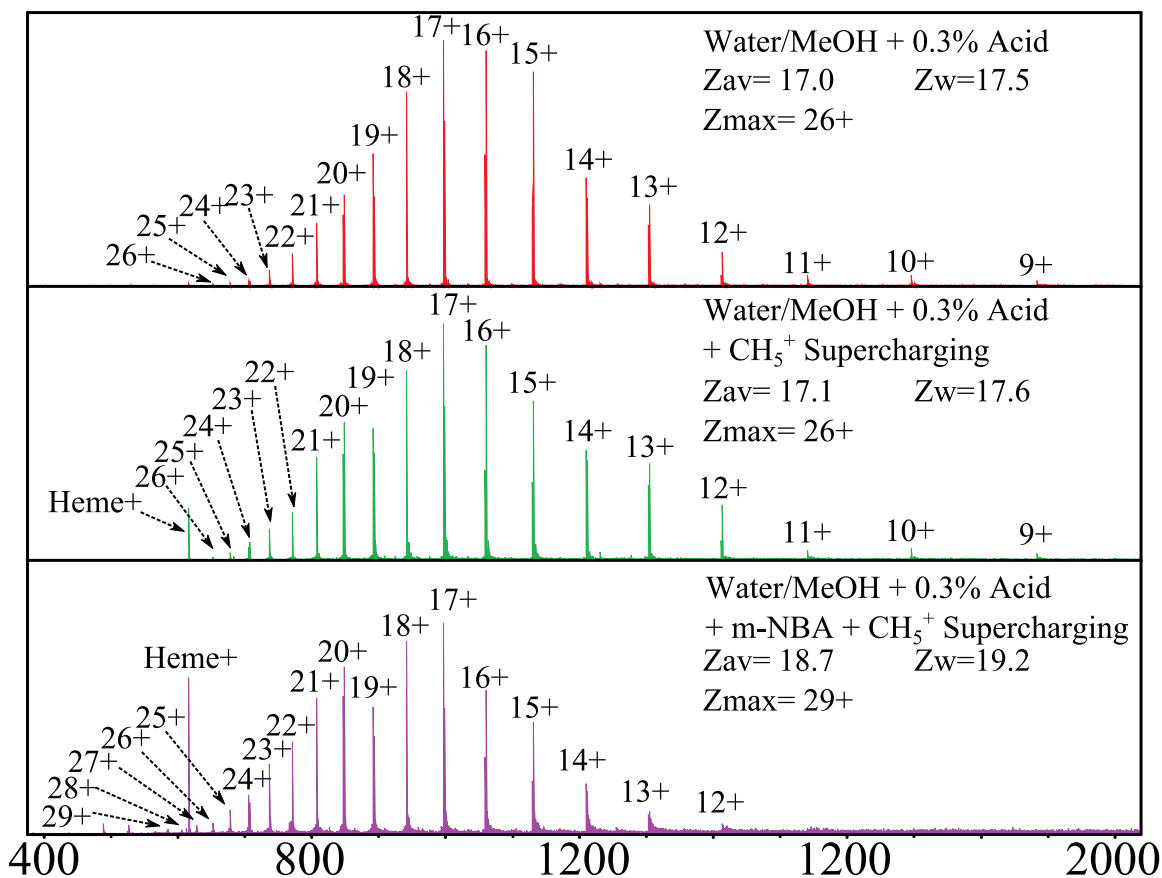
Figure 7.7: Denatured and/or supercharged myoglobin protein spectra produced by acidic denatured conditions (top), followed Superacid supercharging (middle), and a combination of denatured acidic conditions with m-NBA solution phase supercharging and Superacid-supercharging (bottom) showing the possibility of combining enhanced charging techniques to achieve even higher charge states than individual techniques.

The results presented in Figure 7.7 (above) show marginal increases to the charge states of acidically denatured myoglobin samples by the ESI-APCI supercharging. However this could be further enhanced by the use of acidic denaturing conditions and both solution phase supercharging reagents (m-NBA) and the Superacid-supercharging ion source, producing appreciable increase in both average charge (17.0+ increased to 18.7+)

and maximum observed charge state (26+ increased to 29+). The results produced show promise for the combined used of various techniques for increased supercharging of analyte species and observation of even higher charge states when needed.

CH_5^+-droplet interaction and proton transfer calculations:

The interaction of CH_5^+ with positively-charged ESI droplets involves a repulsive coulombic force between these two positively charged entities. If the magnitude of repulsive force between the incoming CH_5^+ ions and the charged ESI droplet is too high then the CH_5^+ ion will be deflected before reaching a close enough distance to allow H^+ transfer to the droplet/analyte species contained within. Since CH_5^+ ions were clearly shown to enhance the charging of analyte species within the ESI droplets (as shown above) this repulsive force must have been overcome at some point during the electrospray process.

Beauchamp *et al.* [45] have shown through careful measurements of ESI droplet size and average charge that ESI droplets undergo many discharge events during the ESI process, splitting up into many daughter ions with each discharge event. These discharge events occur as the droplet solvent evaporates, while the charge remains, until the droplet eventually reaches the Rayleigh limit of stability and discharges. The exact point where the charged droplet fissions is dependent on the solvent used, methanol and acetonitrile were shown to fission at 100% to 120% of the Rayleigh limit, while water was shown to fission below/at the calculated Rayleigh limit (85-100%).

The electrostatic interactions between the charged ESI droplet and a CH_5^+ ion was evaluated computationally using two approaches; first a mean field calculation of the coulombic forces between the charged droplet and a CH_5^+ ion at varying distances and

varying charge, second using coulomb and Particle-particle dielectric electrostatic model (PPDEM) theory to model the same interaction.

Mean field calculations. The force between the CH_5^+ ions and ESI droplets was calculated using Coulomb's law, detailed measurements from Beauchamp *et al.* focused on droplets in the 5-20 µm range, with no discharge/fission events occurring in a droplet above 20 µm, as a result 20 µm was used as the droplet diameter in all calculations as a starting diameter condition, though the Rayleigh limit scales with diameter of droplet, as a result the calculations would be applicable regardless of chosen diameter.

  For the calculation, a spherical 20 µm droplet was randomly populated with a certain number of charges upon its surface (a varying percentage of the Rayleigh limit for a 20 µm droplet; Rayleigh limit of charge for 20 µm droplet of water is ~12.6 Me). Once the droplet was populated with point charges a CH_5^+ ion was introduced, at distance x, the force between every individual charge on the droplet surface and the CH_5^+ ion was then calculated and summed for distance x. The distance x was varied and each summed value plotted for the specific number of charges, the number of surface charges was then altered and the process repeated, the summation of these calculations is shown in Figure 7.8.
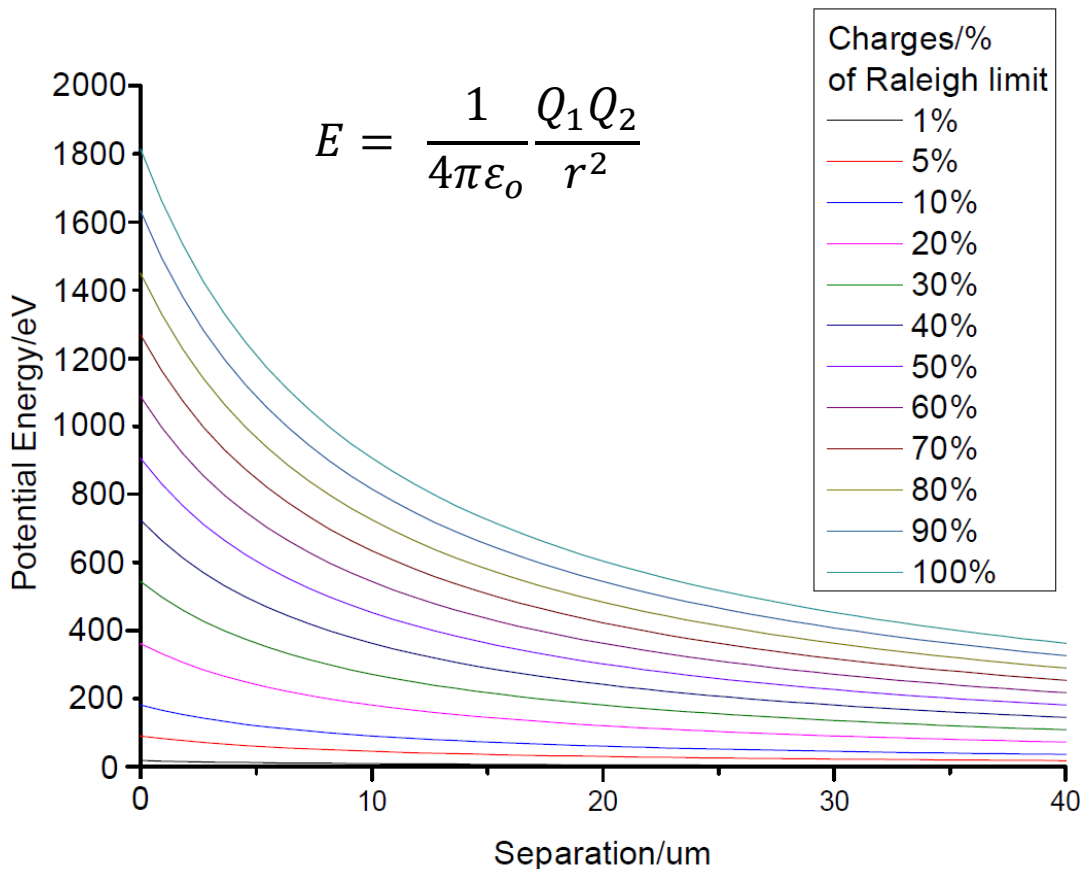
Figure 7.8 - Mean field calculation results of the repulsive force between a CH₅⁺ ion and a 20 μm charged ESI droplet at a varying distance and T=300K, each curve represents calculations for a varying number of charges (as a percentage of the Rayleigh limit for the 20 μm droplet: 12.6 Me).

Particle-particle dielectric electrostatic model (PPDEM). As mentioned previously the mean-field calculations are based on pure-coulombic interactions, only mitigated by the dielectric constant of the solvent. However as a charged species approaches a charged droplet of polarizable solvent, the individual solvent molecules will re-orientate in order to minimize their repulsive interaction with the incoming charged species, and thus present their other polarity (δ-) outward towards the incoming charge – causing a more

attractive force and reducing the energy barrier of approach. If enough of the droplet molecules are affected by the ion-induced dipole rearrangement then the overall barrier could be lowered enough to become attractive.[46,47] As a result additional calculations were conducted using a modified Coulombic repulsion model, taking into account ion-induced dipole, induced quadrupole, and induced octapole interactions using equation 6, below, from Stace *et al.*[47]

$$U1 = K\frac{Q_1 Q_2}{h} - \frac{KQ_1^2}{2} \sum_{m=0}^{\infty} \frac{(\varepsilon_2-1)m}{(\varepsilon_2-1)m+1} \frac{a_2^{2m+1}}{h^{2m+2}} \qquad \text{[eq. 6]}$$

Where the static interaction energy (U1) between a point charge (Q_1) and a sphere of radius a_2, charge (Q_2), and dielectric constant ε_2, separated by distance h for m= 1-3, and K = $1/4\pi\varepsilon_0$ ($\approx 9 \times 10^9$ V m C^{-1})

Equation 6 can be evaluated analytically to produce equation 7, below, and was then used to model the same 20 μm diameter charged droplet of water at varying proportions of its Rayleigh limit as conducted in the mean-field calculations.

$$U1 = K\frac{Q_1 Q_2}{h} - K\frac{Q_1^2}{2} \frac{a_2^3}{h^2(h^2-a_2^2)} \qquad \text{[eq. 7]}$$

The initial results of the Coulombic force calculations between the ESI droplet and CH$_5^+$ ion are shown below in Figure 7.8. Although the force between the species increases as the separation distance decreases, as before for the mean field calculations, the PPDEM shows clearly that at very close distances (<~20 nm) there can be an overall attractive force between the CH$_5^+$ ion and the charged ESI droplet, allowing close interaction between the two species and as a result – proton transfer.
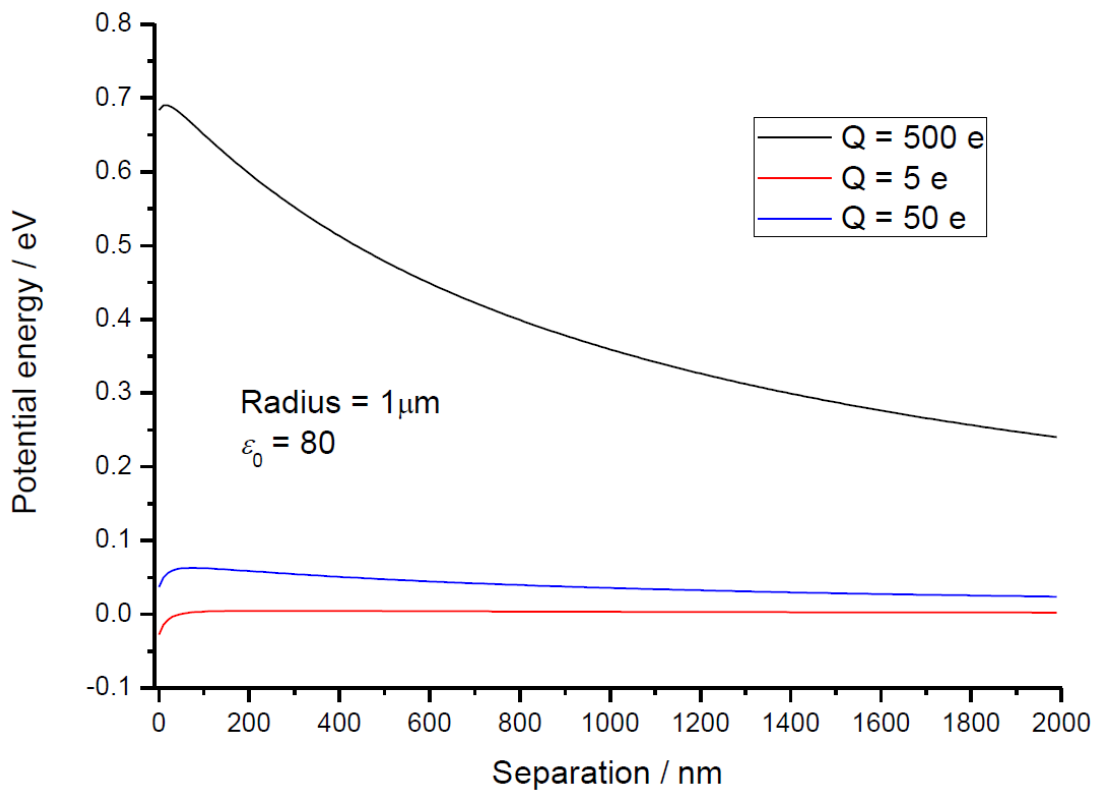
Figure 7.9 – Initial results from the PPDEM calculations for a 1 µm droplet with a very low number of surface charges, showing that although initially repulsive, due to dielectric effects within the droplet, at close distances CH_5^+ ions can be attracted to positively charged electrospray droplets, enabling proton transfer and supercharging.

The plot shows Potential energy / eV on the y-axis (0 to 2000) versus Separation / nm on the x-axis (0 to 40000), with annotations:

Raleigh limit 12.6 Me
$a_1 = 10\ \mu m$
$a_2 = 1.1 \times 10^{-4}\ \mu m$
$k_1 = 80$
$k_2 = 2$

100% RI

10% RI
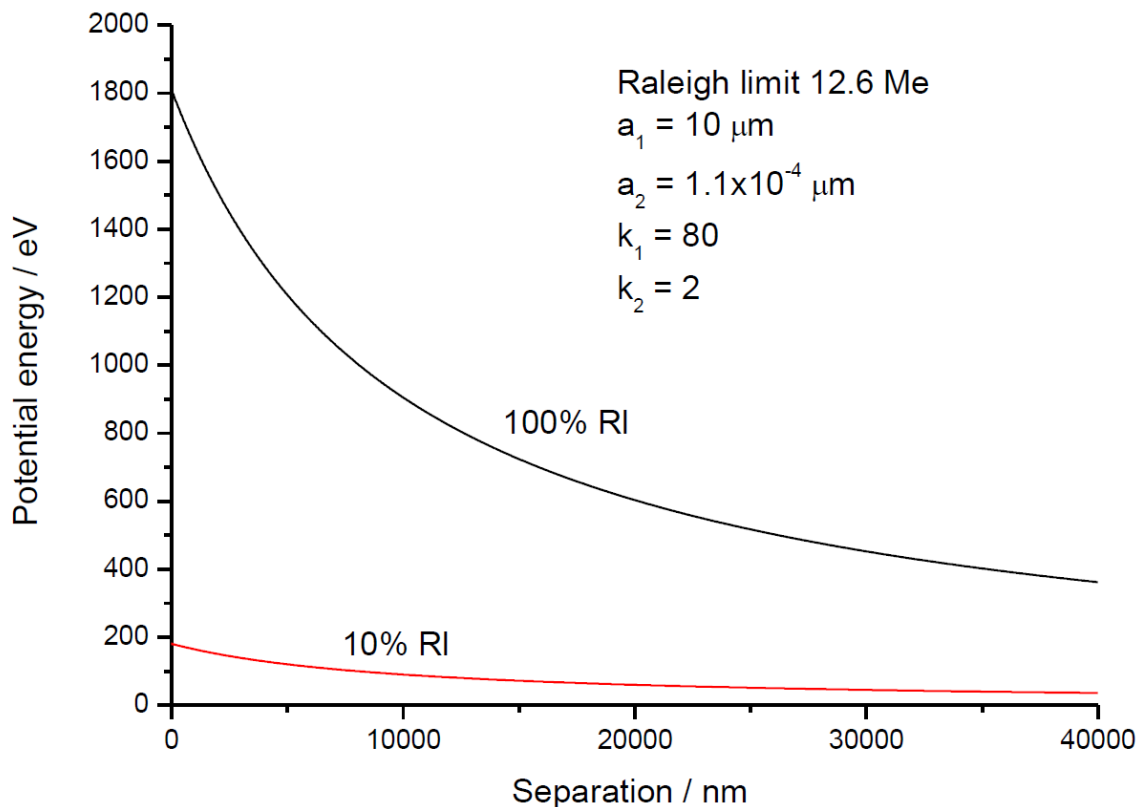
Figure 7.10 – Final results from the PPDEM calculations showing that at high numbers of surface charge the slightly attractive force previously observed becomes unfortunately negligible and the calculations approach the mean-field approximation.

As before, the force between the CH_5^+ ion and a 20 µm positively charged aqueous ESI droplet was calculated at varying distances, and for various numbers of surface charges, the results of which are shown above (Figure 7.10). Although the initial results shown in Figure 9 could potentially explain a drop in the activation barrier of CH_5^+ interacting with the droplet, when the surface charge density increase to the levels calculated in the 20µm droplet, the charge-induced reorientation of solvent molecules based on the variables chosen seemed to be negligible in reducing the magnitude of force. As a result for higher

charge densities the PPDEM model values seemed to closely approach the pure-electrostatic model discussed above.

Energetics and evaluation:

The energy of the CH_5^+ ion was then estimated for temperature of 300K (assuming a purely thermal process), taking the energy of the system to be $3/2k_BT$ where k_B is Boltzman's constant and T is temperature, this simplistic model does not take into account chaotic effects such as Brownian motion, turbulent forces, and/or solvation effects. However, using this Boltzman distribution of energies (shown below in Figure 7.11) the probability of a CH_5^+ ion overcoming the repulsive force from the charged droplet could then be calculated for each charge density.
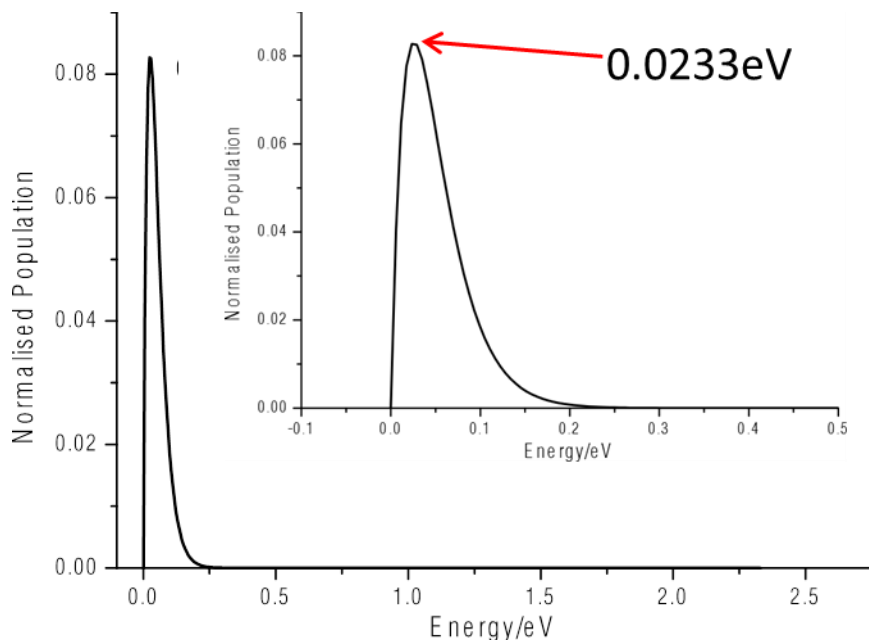
Figure 7.11: Results from the calculation Boltzmann energy curve for CH_5^+ ions at room temperature, assuming the energy of the system to be equal to $3/2k_BT$. The main proportion of species are shown inset, along with the corresponding energy of the most populated energy state.

The resulting CH_5^+ energy calculations showed the most populated state for CH_5^+ at RTP corresponded to an energy of 0.0233 eV, which is substantially lower than even the 1% Rayleigh limit charge density calculations performed during the electrostatic models (~18 eV).

The models proposed do not take into account various processes which may assist the interaction of CH_5^+ with the charged droplet (such as Brownian motion/solvent shielding of charge), the substantial difference between the theoretical calculations of available energy and repulsive interaction indicate that the interaction process must be enabled by other factors. It is entirely possible that the electrostatic repulsive forces, though large with respect to the energy of stationary CH_5^+ ions (as calculated above), become much lower, or even negligible when considering other obvious processes such as the turbulence force provided from the inert nebulizing gas and the electric field actively attracting both species to a common, intersecting path into the mass spectrometer inlet. It is also likely that CH_5^+ ions are not able to travel to the ESI droplets without side-reactions occurring. With the large amount of water/solvent vapour present within the ESI-APCI ionisation the reactive CH_5^+ ions would quickly transfer the bound proton to a more stable species, such as H_2O, forming H_3O^+ ions. The H_3O^+ ions could well then travel further and interact with the ESI droplets to cause the observed supercharging, however ambiguity still remains in the mechanism at this time. Further working into investigating this ionisation mechanism is currently underway.

The super-acid supercharging experimental results and the calculations performed lead to two possible mechanisms for the supercharging effect, either:

a) the addition of CH_5^+ ions accelerates charge build up at each ESI droplet's surface due to proton transfer from the CH_5^+ to the surface - up until the Rayleigh limit is reached, at which point the droplet will discharge into smaller daughter droplets and continue to evaporate, as with normal ESI. As a result CH_5^+ creates a faster and more efficient means of causing coulombic explosions within ESI droplets, leading to increased charging of analytes.

And/or b) protons transferred from the CH_5^+ ions are sequestered into the ESI droplets, lowering the pH of the bulk solution within the droplet which will then increase the addition of protons to analyte molecules during the ESI process, which will increase the charge states observed.

Further experiments focused on probing the supercharging mechanism of the ESI-APCI via the use of various reagent gases through the glow-discharge source to produce proton-donors with varying proton affinities and study the resulting effect on the resulting ESI-MS analysis. Due to the extreme ability for CH_5^+ to donate the additional proton to nearly any species, it is possible that the CH_5^+ ions do not directly interact with the ESI droplets in question. As there would be a substantial concentration of water vapour available, is is likely the CH_5^+ ions will first protonate/ionize water vapour molecules/clusters which could then interact with the positively charged ESI droplets. The additional solvent system could screen the repulsive forces more effectively, and promote a different mechanism of reaction. Further experimentation using gases which would produce slightly less effective proton-donors may be able to avoid donation to stray solvent, avoid additional intermediates, and directly interact with/penetrate ESI droplets and directly

increase surface charge/protonate analyte species. These experiments are currently underway. Examples of possible reagent gases are shown below in Table 7.1:

Table 7.1: Examples of CI reagent gases used in previous studies possessing a range of proton affinities (based on the neutral species), which could be suitable for probing the mechanism of superacid-supercharging.

Reagent gas	Reactant ions	Neutral form reactant ions	Proton affinity of neutral product	Analyte ions
H_2	H_3^+	H_2	424	$[M+H]^+$, $[M-H]^+$
CH_4	CH_5^+ ($C_2H_5^+$ and ($C_3H_5^+$)	CH_4	522	$[M+H]+$ ($[M+C_2H_5]+$ and $[M+C_3H_5]^+$ possible with pure-CI)
$i\text{-}C_4H_{10}$	$t\text{-}C_4H_9^+$	$i\text{-}C_4H_8$	820	$[M+H]_+$ ($[M+C_4H_9]+$, eventually $[M+C_3H_3]^+$, $[M+C_3H_5]^+$, and $[M+C_3H_7]^+$)
NH_3	NH_4^+	NH_3	854	$[M+H]^+$ and $[M+NH_4]^+$

Conclusions:

Reaction of the super-acid CH_5^+ with droplets formed during electrospray ionization (ESI) mass spectrometry was achieved via implementation of a methane glow discharge chamber near the tip of an ESI needle in an enclosed reaction chamber. Addition of CH_5^+

ions to the Taylor cone during ESI-MS resulted in an increase in the observable charge states of protein analytes, causing a distinct "supercharging" effect. Average and base peak charge state increases were observed for both native-like protein samples (protein dissolved in pure water) and for denatured protein samples (protein dissolved in methanol-water solutions), with the greatest increases observed for the denatured protein sample.

Application of the ESI-APCI source to native protein MS provided an increase to protein peak signal intensity of over 500% compared to standard ESI MS.

Super-acid supercharging was also used to study the effect on a small molecule sample – Ascorbic acid. Peak intensity for the protonated Ascorbic acid peak was shown to increase by 22% when exposed to CH_5^+ ions. Peak intensity of the protonated Ascorbic acid dimer species was shown to increase by 291% when exposed to CH_5^+ ions, a considerable enhancement compared to traditional ESI alone, with potential increase to sensitivity and limits of detection for small molecules of interest.

Theoretical calculations were performed to evaluate the repulsive force between the positively charged droplets formed during ESI-MS and the incoming superacid CH_5^+ ions. Using two different modelling approaches (mean field coulombic repulsion and Particle-particle dielectric electrostatic models (PPDEM)), it was found that the CH_5^+ ions were not likely to be able to transfer additional protons the charged ESI droplets when at/near the Rayleigh limit, but was possible at (much) lower charge density levels. Though the mean field approach predicted an ever-increasing repulsion between the CH_5^+ ion and the charged droplet the PPDEM calculations showed that at small distances (<~20 nm) and low charge densities the repulsive force between the two species can

become attractive and can lower the coulombic barrier and facilitate the proton transfer process.

Though many potentially reactive species are produced during methane ionization, no modification (methylation or otherwise) of protein/small molecule species was observed during the supercharging process. This method is widely applicable to any mass spectrometer using an electrospray source interface, and can easily be activated/deactivated within seconds if desired.

The ability to supercharge analytes in the gas phase enables attainment of high charge state biomolecule species, without the addition of solution-phase chemical reagents. Accessing higher charge states enables more comprehensive tandem MS (MS/MS) cleavage coverage during fragmentation, which is of particular interest to both Top-Down protein analysis and any bottom-up proteomic style studies of poly-amino acids, enabling higher confidence in peptide/protein assignments.

Moreover gas-phase supercharging provides an alternative/complementary technique to enhance charging of particularly challenging/ large biomolecules in order to bring them into a more moderate m/z range to enable study of these species via MS/MS more easily. Due to the use of the standard ESI needle set-up, there is little reason why this methodology cannot be hyphenated to Liquid-Chromatography (LC) systems for the online supercharging of species of interest.

Acknowledgments

<u>References:</u>

(1)   White, E. T. *Science.* **1999**, *284*, 135–137.

(2)   Talrose, V.; Lyubimova, A. K. *Reports Sov. Acad. Sci.* **1952**, *86*, 909.

(3)   Nikolaev, E. *J. mass Spectrom.* **1998**, *501* (33), 499–501.

(4)   Munson M. S. B.; Field F. H. *J. Am. Chem. Soc.* **1966**, *88* (12), 2621–2630.

(5)   Field, F. H.; Munson, M. S. B. *J. Am. Chem. Soc.* **1965**, *87* (15), 3289–3294.

(6)   Munson, B. *Int. J. Mass Spectrom.* **2000**, *200*, 243–251.

(7)   Field, F. H. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 277–283.

(8)   McLafferty, F. W.; Tureček, F. *Interpretation of Mass Spectra*, 4th ed.; USB, 1993.

(9)   Chai, R.; Harrison, A. G. *Anal. Chem.* **1981**, *53* (1), 34–37.

(10)   Meng, C. K.; Mann, M.; Fenn, J. B. *Zeitschrift Phys. D Atoms, Mol. Clust.* **1988**, *10* (2-3), 361–368.

(11)   Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Craig, M.; Meng, C. K. A. I.; Mann, M.; Whitehouse, C. M. **2013**, *246* (4926), 64–71.

(12)   Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68* (1), 1–8.

(13)   Iavarone, A. T.; Jurchen, J. C.; Williams, E. R. **2000**, *0305* (00).

(14)   Iavarone, A. T.; Jurchen, J. C.; Williams, E. R. **2001**, *73* (7), 1455–1460.

(15)   Iavarone, A. T.; Williams, E. R. *Int. J. Mass Spectrom.* **2002**, *219* (1), 63–72.

(16)   Iavarone, A. T.; Williams, E. R. **2003**, No. 16, 2319–2327.

(17)   Lomeli, S. H.; Yin, S.; Ogorzalek Loo, R. R.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (4), 593–596.

(18)   Miladinović, S. M.; Fornelli, L.; Lu, Y.; Piech, K. M.; Girault, H. H.; Tsybin, Y. O. *Anal. Chem.* **2012**, *84* (11), 4647–4651.

(19)   Amster, I. *J. Mass Spectrom.* **1996**, *31*, 1325–1337.

(20)   Horn, D. M.; Breuker, K.; Frank, A. J.; Mclafferty, F. W.; V, C. U.; York, N.; August, R. V. **2001**, No. d, 9792–9799.

(21)   Kjeldsen, F.; Giessing, A. M. B.; Ingrell, C. R.; Jensen, O. N. **2007**, *79* (24), 9243–9252.

(22)   Shen, Y.; Tolić, N.; Masselon, C.; Pasa-Tolić, L.; Camp, D. G.; Lipton, M. S.; Anderson, G. a; Smith, R. D. *Anal. Bioanal. Chem.* **2004**, *378* (4), 1037–1045.

(23)   Cooper, H. J.; Akbarzadeh, S.; Heath, J. K.; Zeller, M. *J. Proteome Res.* **2005**, *4* (5), 1538–1544.

(24)   Tipton, J. D.; Tran, J. C.; Catherman, A. D.; Ahlf, D. R.; Durbin, K. R.; Lee, J. E.; Kellie, J. F.; Kelleher, N. L.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2012**, *84* (5), 2111–2117.

(25)   Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. *Biochem. Biophys. Res.*

Commun. **2014**, *445* (4), 683–693.

(26) Sterling, H. J.; Daly, M. P.; Feld, G. K.; Thoren, K. L.; Kintzer, A. F.; Krantz, B. A.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (10), 1762–1774.

(27) Sterling, H. J.; Prell, J. S.; Cassou, C. a; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (7), 1178–1186.

(28) Sterling, H. J.; Kintzer, A. F.; Feld, G. K.; Cassou, C. a; Krantz, B. a; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (2), 191–200.

(29) Teo, C. A.; Donald, W. A. **2014**.

(30) Yin, S.; Loo, J. a. *Int. J. Mass Spectrom.* **2011**, *300* (2-3), 118–122.

(31) Lomeli, S. H.; Peng, I. X.; Yin, S.; Loo, R. R. O.; Loo, J. a. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (1), 127–131.

(32) Sterling, H. J.; Kintzer, A. F.; Feld, G. K.; Cassou, C. A.; Bryan, A. **2011**, No. 510.

(33) Rayleigh, Lord. *Philos. Mag. Ser. 5* **1882**, *14* (87), 184–186.

(34) Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26* (October 1990), 514–518.

(35) Konermann, L.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (12), 1248–1254.

(36) Fisher, C. M.; Kharlamova, A.; Mcluckey, S. A. **2014**.

(37) Zubarev, R.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (16), 3265–3266.

(38) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, O. S.;

365

Wurm, F. M.; Hartmer, R. **2011**, 8919–8927.

(39)   Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. a. *Anal. Chem.* **2014**, *86* (1), 317–320.

(40)   Dickinson, E. R.; Jurneczko, E.; Pacholarz, K. J.; Clarke, D. J.; Reeves, M.; Ball, K. L.; Hupp, T.; Campopiano, D.; Nikolova, P. V.; Barran, P. E. *Anal. Chem.* **2015**, *87* (6), 3231–3238.

(41)   Pan, J.; Han, J.; Borchers, C. H. *Int. J. Mass Spectrom.* **2012**, *325-327*, 130–138.

(42)   Chen, J.; Rempel, D. L.; Gau, B. C.; Gross, M. L. *October* **2012**, *141* (45), 18724–18731.

(43)   Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, Ü. a; Tsybin, Y. O. *Chem. Soc. Rev.* **2013**, *42* (12), 5014–5030.

(44)   Barrow, M. P.; Witt, M.; Headley, J. V.; Peru, K. M. **2010**, *82* (9), 3727–3735.

(45)   Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. *J. Phys. Chem. A* **2002**, *106* (42), 9957–9967.

(46)   Raggi, G.; Stace, A. J.; Bichoutskaia, E. *Phys. Chem. Chem. Phys.* **2013**, *15* (46), 20115–20119.

(47)   Chen, X.; Bichoutskaia, E.; Stace, A. J. *J. Phys. Chem. A* **2013**, *117* (19), 3877–3886.

Chapter 8:

The Rhodium Revolver: how transition metal complexes can facilitate Hydrogen-Deuterium Exchange in bound ligands

Hydrocarbon C-H bonds are particularly weak bases, usually requiring strong acids to achieve deprotonation, and especially stable and chemically inert as a result. However it was recently observed that certain rhodium-based compounds, designed as metallodrugs, might be able to exchange hydrogen atoms on bound arene ligands to other species and possibly deuterium atoms if available.

This chapter of the Thesis is centred around a study of these remarkably enabling rhodium complexes and their ability to facilitate the exchange of protons on particularly stable bound arene ligands with deuterium atoms from deuterated solvent. Resembling the Hydrogen-Deuterium eXchange (HDX) experiments used frequently in mass spectrometry for the study of protein structure, HDX with the rhodium complexes showed not only the surprising ease and speed of the HDX with bound ligands, but also the steric and proximity requirements to allow the metal centre to facilitate the exchange.

Although the MS-based results presented within this chapter are produced and interpreted solely by the thesis Author, this was a collaborative project with the FT-ICR MS analysis providing a proportion of the evidence and data. As a result some elements were conducted by collaborating researchers, namely Joan J. Soldevila-Barreda synthesised the complexes studied and characterised them by

standard NMR etc. prior to MS analysis, Juliusz A. Wolny performed the molecular modelling calculations for the HDX reactions, and Guy J. Clarkson performed the X-ray diffraction experiments. The Author's contribution to this work was centred around performing and monitoring the HDX reactions by FT-ICR MS and chemically altering bound mono-dentate ligands/conditions to study the effect on subsequent HDX.

This chapter is based predominantly on results of the following paper, with a large amount of additional nESI-FT-ICR MS data collected for the study, not all of which was used in the final manuscript:

"Molecular "Twister" for facile sequential deuteration of methyl groups on cyclopentadienyl rings"

J. J. Soldevila-Barreda, Juliusz Wolny, **Christopher A. Wootton**, Abraha Habtemariam, Guy J. Clarkson, Peter B. O'Connor, Volker Schunemann, and Peter J. Sadler. Submitted to Nature Chemistry, 2016.

8.1 Introduction:

Transition metal complexes have shown a wide variety of activities and have been able to "activate" many different groups in chemistry, e.g. oxidation of alcohol groups to ketones and polymerisation catalysis via transition metal intermediated such as the Grubb's metathesis catalysis.[1] However carbon-hydrogen bonds in pure hydrocarbon structures still remain particularly inert to most reaction conditions, they are exceptionally poor acids, rarely relinquishing protons, and to do so often require extremely strong bases, such as caesium cyclohexylamide. However C-H bond activation is of particular interest in chemistry, and a very sought after goal in heterogeneous catalysis.

Transition-metal compounds have been at the forefront of C-H bond activation for many years, and have been successful in activating bonds from many species, usually by lowering activation energies for reactions, or by stabilising usually energetically-unfavoured intermediates, making them accessible. However this C-H bond activation has almost ubiquitously been applied to external species outside of the metal-complex, or available solvent molecules after binding to the metal complex/centre. It is particularly rare for these species to activate bonds within the complex itself and not external solvated species/solvents.

Therefore it was particularly interesting when a series of rhodium compounds, synthesised by J..J. Soldevila-Barreda, appeared to lose arene-bound hydrogen atoms from methyl groups bound to the cyclopentadienyl ring of the piano stool complexes (shown below in Figure 1). The H-NMR signal for the methyl protons decreased steadily over time while in deuterated NMR solvent, believed to be due to the deuteration of the bound ligands by the solvent. However the methyl protons of the Cp* rings are particularly stable and very unlikely to be deprotonated, both due to the extremely low acidity of C-H hydrocarbon bonds, and due to the fact the methylated-cyclopentadienyl ligands are bound via an η_6 interaction with the metal centre and are already negatively charged, meaning it is unlikely they would be easily deprotonated beyond this state.
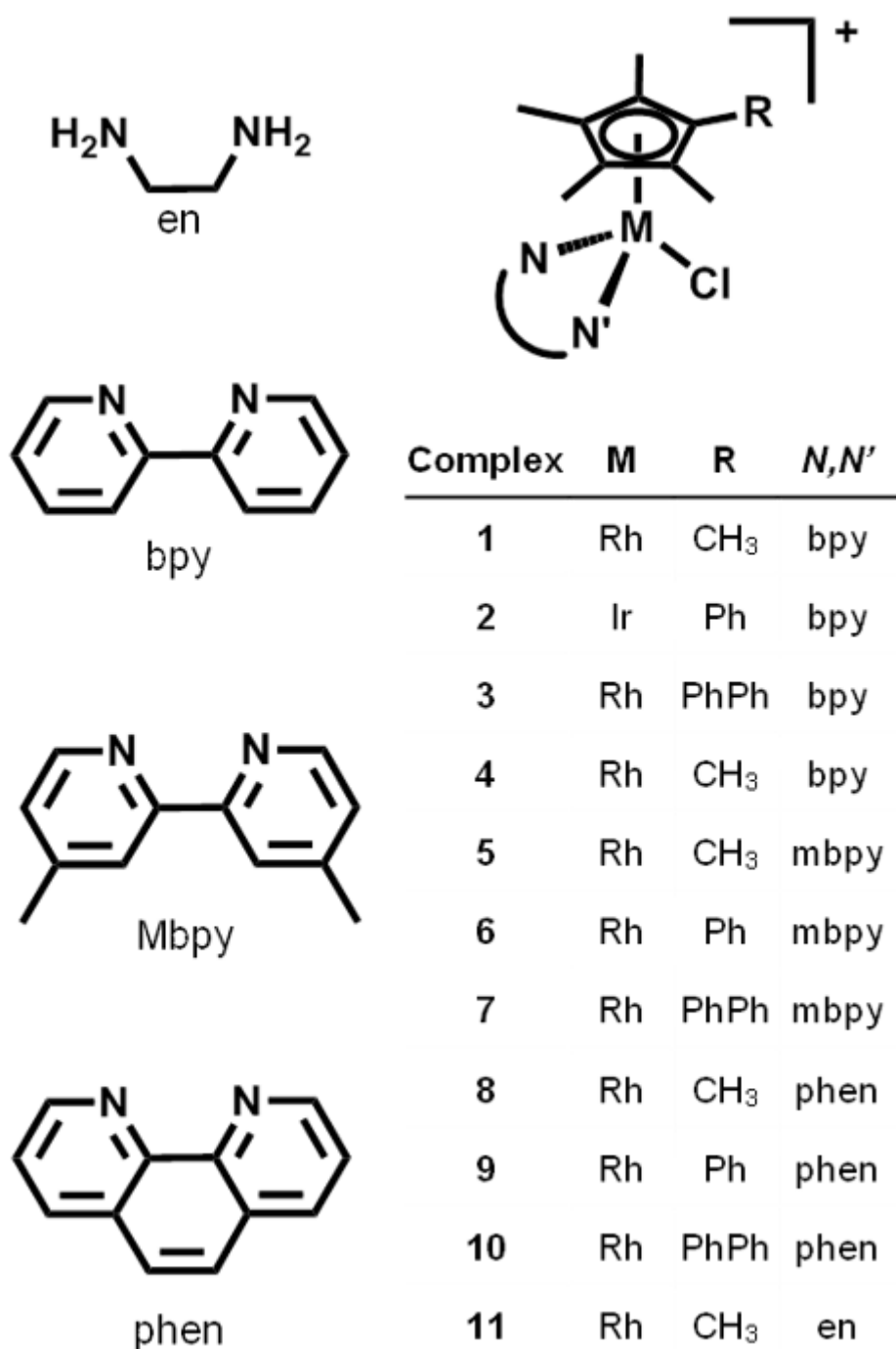
Figure 8.1: The range of metal complexes synthesised and studied during the course of the project. All are rhodium complexes with varying ligands, aside from one iridium compound as a control/test for Ir metal centres, but was ineffective at achieving the same HDX reactions. Reproduced from Soldevila-Barreda *et.al.*[2]

8.2 C-H bond activation and subsequent HDX-MS:

Since the ^{1}H-NMR signals for Cp* protons was decreasing it was believed to be due to exchange with deuterium from the deuterated NMR solvent, low resolution quadrupole MS confirmed a mass shift to higher m/z but suffered with overlapping isotopes and lack of isotopic resolution. nESI-FT-ICR MS was then employed to study the observation of this exchange, but also the rate and rate accelerants associated with the reaction.

8.2.1 nESI-FT-ICR mass spectrometry:

Samples were analysed via nano-electrospray ionisation (nESI)-Fourier Transform Ion Cyclotron Resonance Mass Spectrometery (FT-ICR MS). All experiments were carried out on a Bruker SolariX FT-ICR Mass Spectrometer, fitted with a 12 Tesla actively shielded magnet (Bruker Daltonik GmbH, Bremen, Germany). Samples were diluted 1000-fold from NMR samples/reaction mixtures with appropriate deuterated/non-deuterated solvents (to ~1 µM), an aliquot of each sample (10-20 µL) was ionised from ~1 µm glass nESI capillaries using a capillary voltage of 900-1200 V, ions were accumulated for 0.01 s in the hexapole-based collision cell before transfer to the ICR cell for detection. Ions of m/z 147-3000 were excited using a frequency sweep excitation and detected for 3.2 s transient length (8 mega-word data points), producing a resolving power of ~850,000 (at 400 m/z) for all spectra. Mass spectra were externally calibrated using a quadratic calibration function/internally using single point calibrations and then manually interpreted and assigned via Data Analysis v4.2 (Bruker Daltonik GmbH, Bremen, Germany).

The majority of samples had been pre-deuterated via exposure to deuterated NMR solvents (D[6]-MeOH) for 72 hours prior to analysis. These samples were then diluted 1000-fold for nESI-FT-ICR MS analysis, the dilution produced a sample of unusually high concentration for nESI-MS analysis, however it accounted for the large variation between NMR samples and due to the low volume consumption of nESI (~2-5 uL/hour) and low sample complexity these high concentrations were not an issue and caused no contamination/carryover with the source or MS.

Selected rhodium samples were reacted with Silver Nitrate (AgNO$_3$, 0.95 molar equivalents) to transform the chlorido complexes into the Aqua species and form silver chloride salt, which then precipitates out of solution and drove the reaction to ~completion within minutes. Samples were then filtered through Celite to remove the silver chloride salt formed prior to nESI-FT-ICR MS analysis. Non-deuterated (all H-containing) samples were dissolved in deuterated solvent and then immediately analysed via nESI-FT-ICR MS, then re-analysed every 10 minutes for the first 80 minutes, then hourly from then on until 14 hours at which point the sample were then left in deuterated solvents overnight and re-sampled at 24, 48, and 72 hours reaction time.

Results and discussion:

8.3 Deuteration kinetics:

The rhodium complex **1** [Cp*)-Rh-(Bipy)Cl]$^+$ was selected for kinetic HDX measurements, the compound was transformed to the aqua species via silver nitrate (as discussed above), filtered through celite and then analysed via nESI-FT-ICR MS at various time points. The exchange of hydrogens on the cyclopentadienyl methyl groups was surprisingly fast, the first time point was taken at 20 minutes (due to filtration time) and a number of hydrogen atoms had already exchanged to deuterium atoms. Continuing to take scans at 10 minute intervals it was found that the fully deuterated species (D$_{15}$ for the 15 H's, now D's, on the Cp* ligand) was detectable within 60 minutes and the reaction reached completion/hit equilibrium within 80 minutes. These results are summarised in Figure 8.2 below:
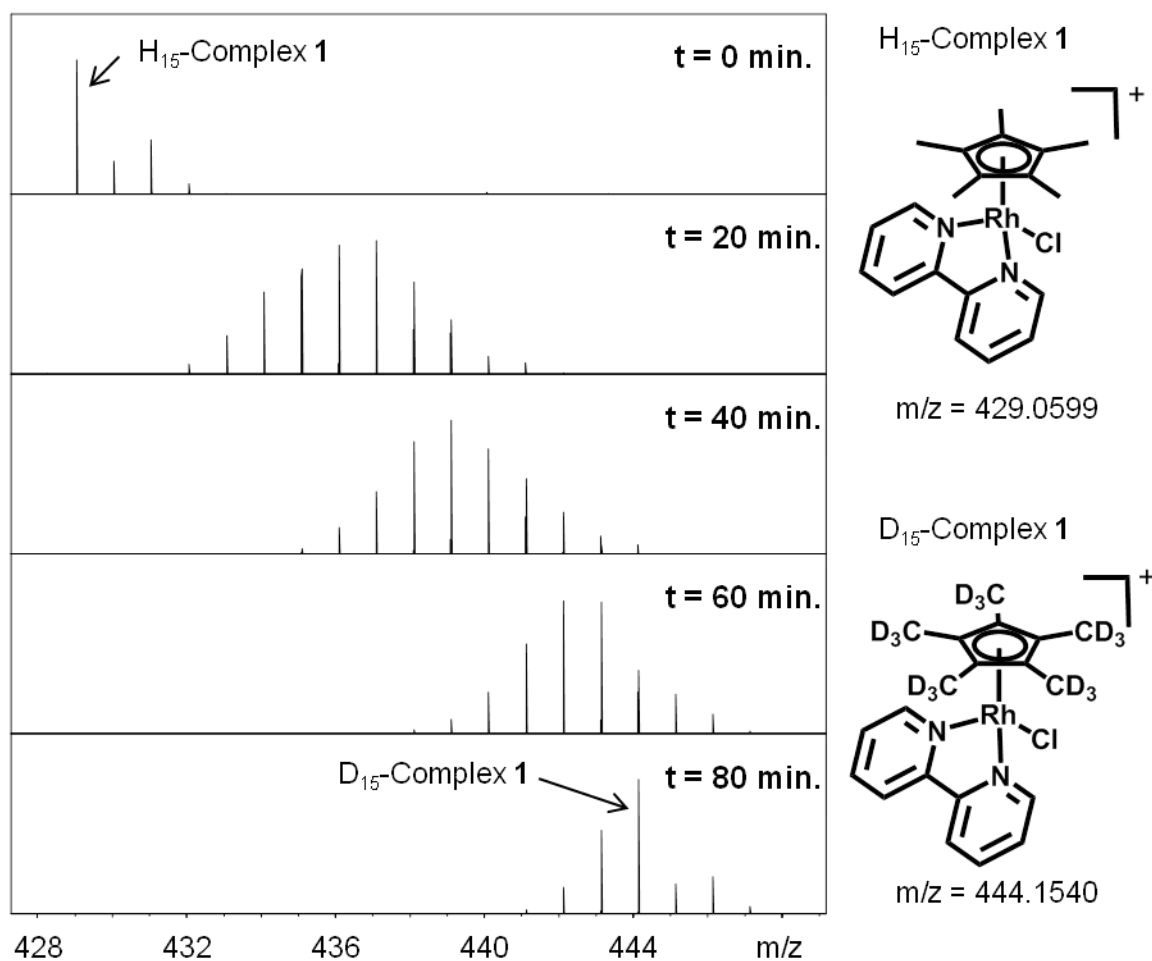
Figure 8.2: nESI-FT-ICR MS spectra of the deuteration (HDX) of Complex 1 (shown inset), at various time points, creating the D₁₅-Complex within 60 minutes and reaching ~equilibrium within 80 minutes, subsequent time point spectra (hourly up to 14 hours, then at 24, 48, 72 hours and 7 days showed only incremental increase in D₁₅ and little decrease in D<₁₅ species).

Though HDX on small molecules/singly charged species is readily achieved on most mass spectrometers, and does not require the 900,000+ resolving power (at 400 m/z) achieved here during nESI-FT-ICR MS analysis, the UHR-MS provided allowed observation of the individual rhodium species' isotopic fine structure, allowing a "true" representation of the HDX reaction, without overlap between isotopologues, and the disruption of the chloride isotope interfering with the M+2 isotopologue. The isotopic fine structure of the rhodium species is shown below in Figure 8.3, along with the 3 mD

difference between certain species, and a representation of the same mass spectrum, but at lower resolving power performance achievable by standard MS instruments.
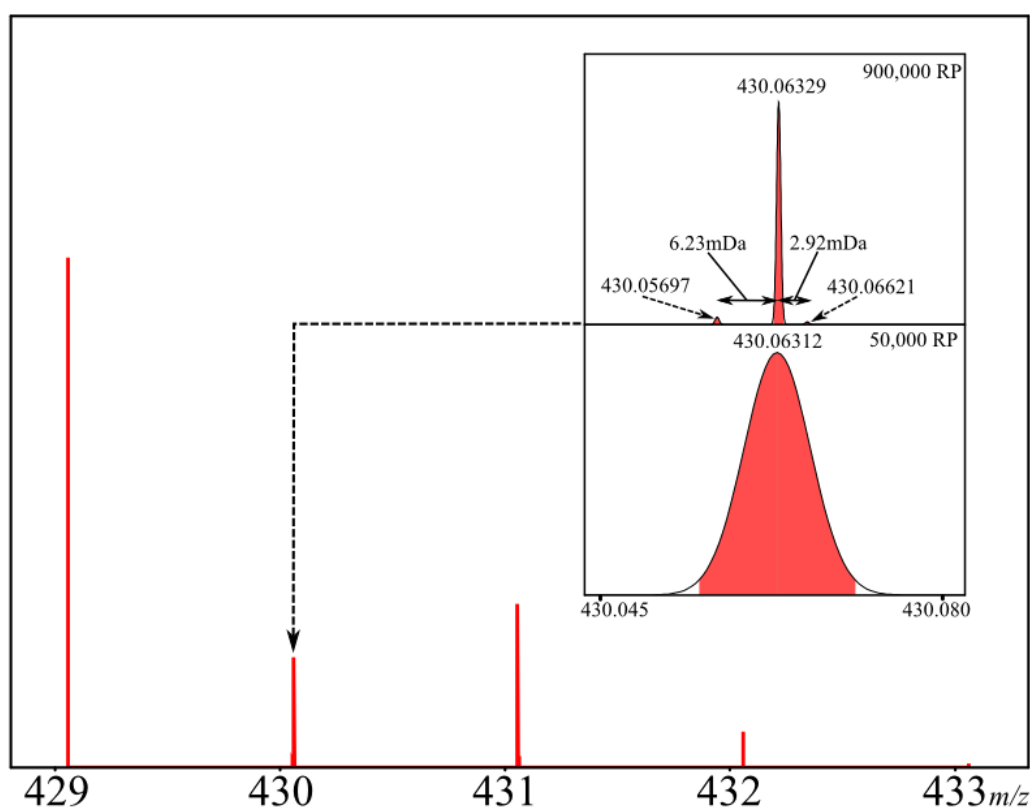
Figure 8.3: A rhodium metallodrug complex (Complex 1, below) simulated mass spectrum of the $[RhC_{10}H_{15}C_{10}N_2H_8Cl]^+$ ion produced during ESI/nESI MS analysis. Inset: a zoom-in of the M+1 isotopologue, at two different resolving powers standard performance for a high-resolution ToF-MS 50,000 resolving power (bottom) and standard performance for the 3.3 s transient acquired using the Bruker Solarix UHR-FT-ICR MS used for all spectra within this chapter 900,000 at 400 m/z (top). The higher performance allows observation and assignment of the isotopic fine structure, providing additional elemental composition information and reducing overlap between isotopes.

The lower resolving power simulations show that the apparent HDX patterns are distorted due to the inability of isotopic fine structure separation and although the HDX is still visible is it not as representative of each H exchanging sequentially and could give false conclusions about the mechanism of HDX with the rhodium complex. If every other isotopologue was seen to be at artificially

374

high intensity due to the chloride isotope, one would be led to believe that the H's were exchanging in pairs, however the UHR allows the isotopologues to be separated and see that the process is indeed sequential and the extent to which it occurs for each compound species (discussed later). These important accurate observations of HDX, time scale, extent of exchange, and sequential reaction helped inform subsequent molecular modelling and insights into the mechanism of HDX for these species conducted by the collaborator Juliusz A. Wolny, the findings of which are summarised in Figure 8.4, below:
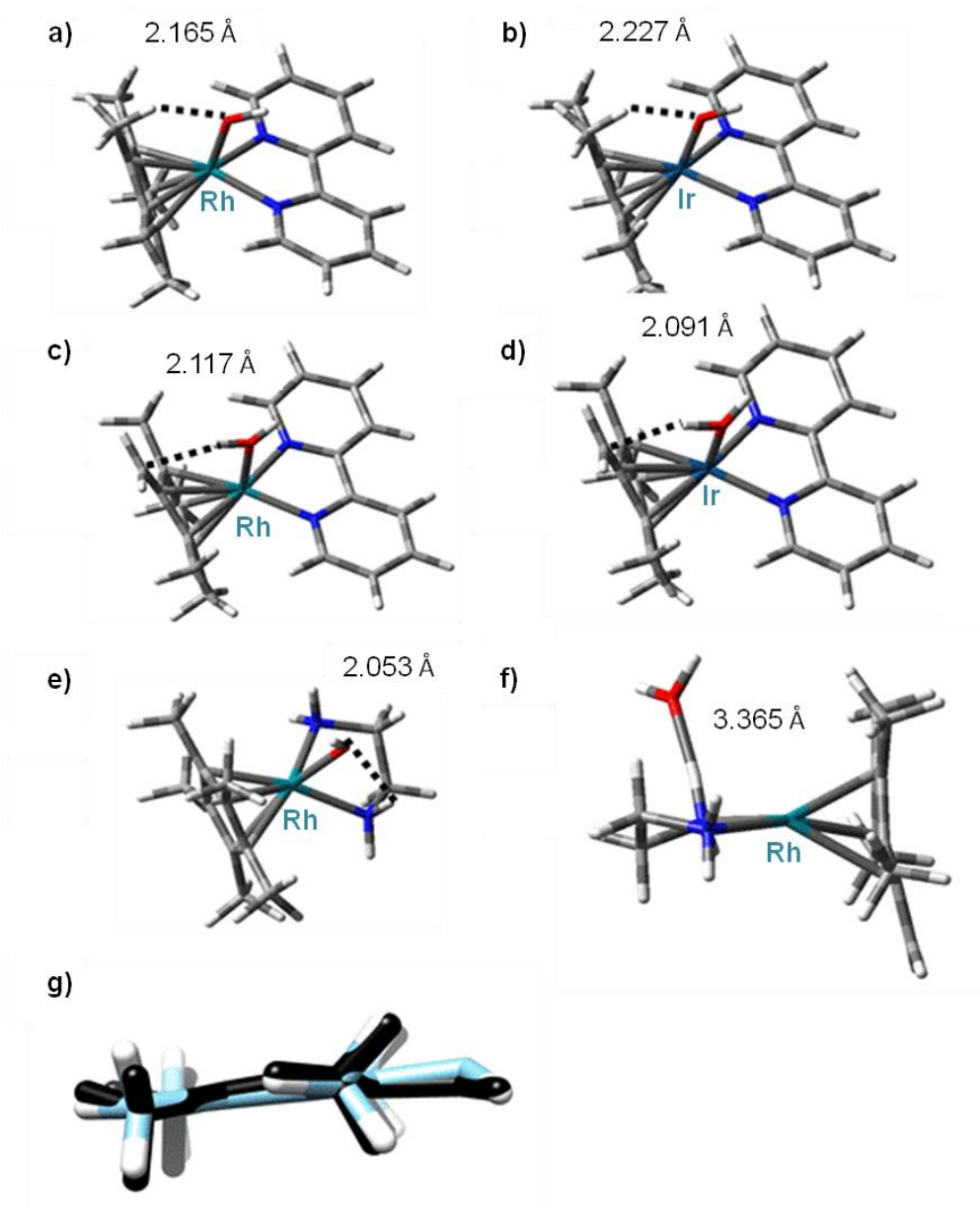
Figure 8.4: Key structures from DFT calculations produced by Juliusz A. Wolny and Volker Schunemann. Molecules with optimised geometry modelled with DFT (CAM-B3LYP/CEP-31G) The important weak interactions are shown. a) Rh(Cp^x)(bpy)OH b) Ir(Cp^x)(bpy)OH c) Rh(Me$_4$Cp=CH$_2$)(bpy)·OH$_2$ d) Rh(Me$_4$Cp=CH$_2$) (bpy)·OH$_2$ e) Rh(Cp^x)(en)OH f) Rh(Me$_4$Cp=CH$_2$) (en)·OH$_2$ g) superimposed structures of Me$_4$Cp=CH$_2$ dianion in the optimised structure of Rh(Me$_4$Cp=CH$_2$)(bpy)·OH$_2$ (in black) and the optimised structure of free dianion (in blue).

The molecular modelling simulations, informed by the NMR and FT-ICR MS experiments were able to calculate that the HDX reaction was most likely to occur through a deprotonated Cp*Me$_x$ intermediate containing a Cp=CH$_2$ bond which will then be deuterated to Cp-CH$_2$D, and so on until each and every methyl is deuterated. The deprotonation is believed to be facilitated by the Rh metal centre when it has an available co-ordination site which can be aquated/hydroxylated, the OH$_2$/OH/OD$_2$/OD ligand can be the proton/deuterium acceptor in the reaction. Since there is only one available co-ordination site in the piano stool complex, due to the bidentate ligand occupying 2 fixed sites, the HDX of methyl protons across the Cp* ring relies on the free rotation of the arene above the rhodium centre, with each methyl exchanging when orientated with the available/aquated/hydroxylated site, thus the rhodium revolver.

The HDX was shown to be accelerated when the complex was aquated (hence the AgNO$_3$ addition), though non-aquated samples were also analysed (see below) and the same deuteration was observed, only slower (hours compared to minutes). The belief that the Cl ligand must dissociate in order to allow deuteration was tested through the use of excess sodium chloride salt solubilised within the deuterated solvent which would, in theory, encourage the Cl ligand to remain bound to the Rh metal centre through shifting the equilibrium.

All complexes synthesised, with and without additional 1 mM NaCl, were dissolved in deuterated solvent an allowed to react/remain for 72 hours, then analysed with nESI-FT-ICR MS. Zoom-ins of the resulting spectra are shown below, along with a table of assignments showing the observed species

within each mass spectrum. It should be noted that during ionisation of charged/neutral metal complexes, loss of the Cl ligand is common during ESI/nESI to produce a positively charged species, loss of this ligand can often be accompanied via loss of an available proton/deuterium, producing an [M-HCl]$^+$ and/or [M-Cl]$^+$ peak. Since the study was of deuteration it can be hard to differentiate between a [M-Cl]$^+$ species with say 14 deuterium atoms and 1 hydrogen atom and a [M-DCl]$^+$ species with 15 deuterium atoms. This ambiguity was settled by observation of the 2+ ion in the spectrum (see assignment table) and the corresponding [M+Cl]$^+$ peak if available. The correct number of deuterium atoms was decipherable in every nESI-MS spectrum by one or both of these methods and the assigned species are summarised in the table below.

Complex 1: [Rh(Cp*)(Bipy)Cl]$^+$

Complex 2: [Ir(Cp*)(Bipy)Cl]⁺

No Salt

$[M(H_{15})-HCl]^+$ $[M(H_{15})-HCl+H_2O]^+$

with Salt

$[M(H_{15})Cl]^+$ $[M(H_{15})Cl+H_2O]^+$
$[M(H_{15})-HCl]^+$ $[M(H_{15})-HCl+H_2O]^+$

$[M(H_{15})-HCl+CD_2HOH]^+$

$[M(H_{15})-HCl+CD_3OH]^+$
$[M(H_{15})-HCl+CD_2HOH]^+$
$[M(H_{15})-HCl+CDH_2OH]^+$
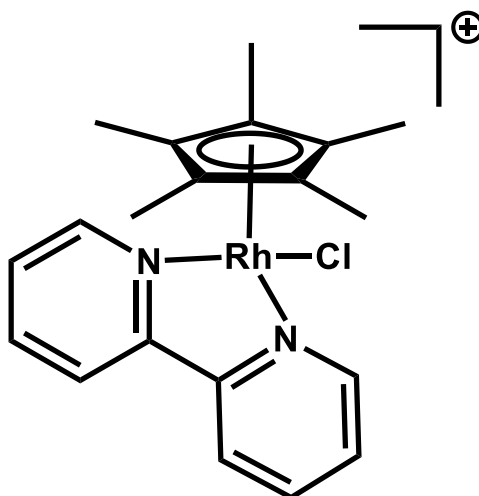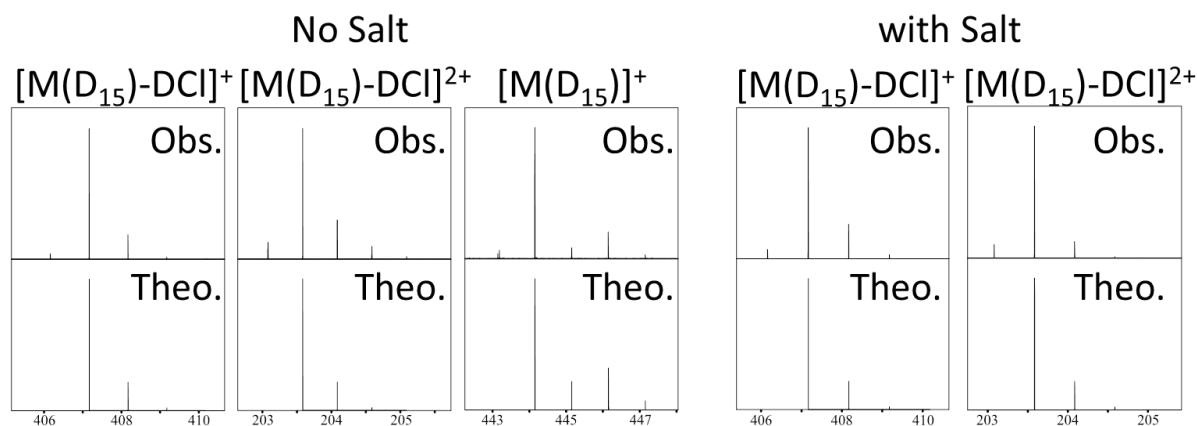
Complex 3: [Rh(Cp*-biphenyl)(Bipy)Cl]⁺
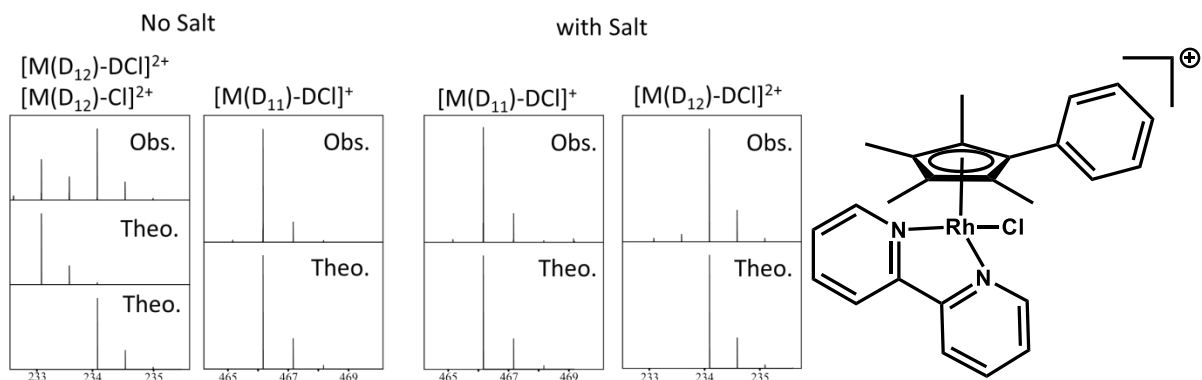
Complex 4: [Rh(Cp*)(Bipy)Cl]⁺
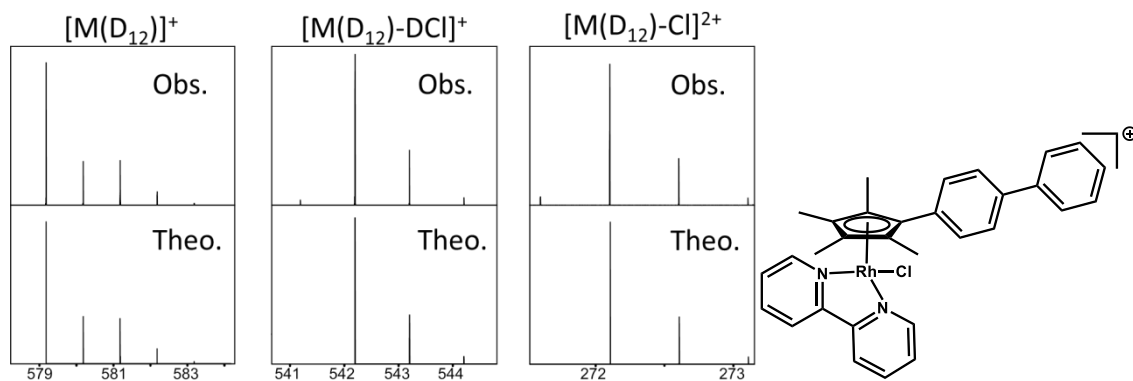
Complex 5: [Rh(Cp*)(Me-Bipy)Cl]⁺
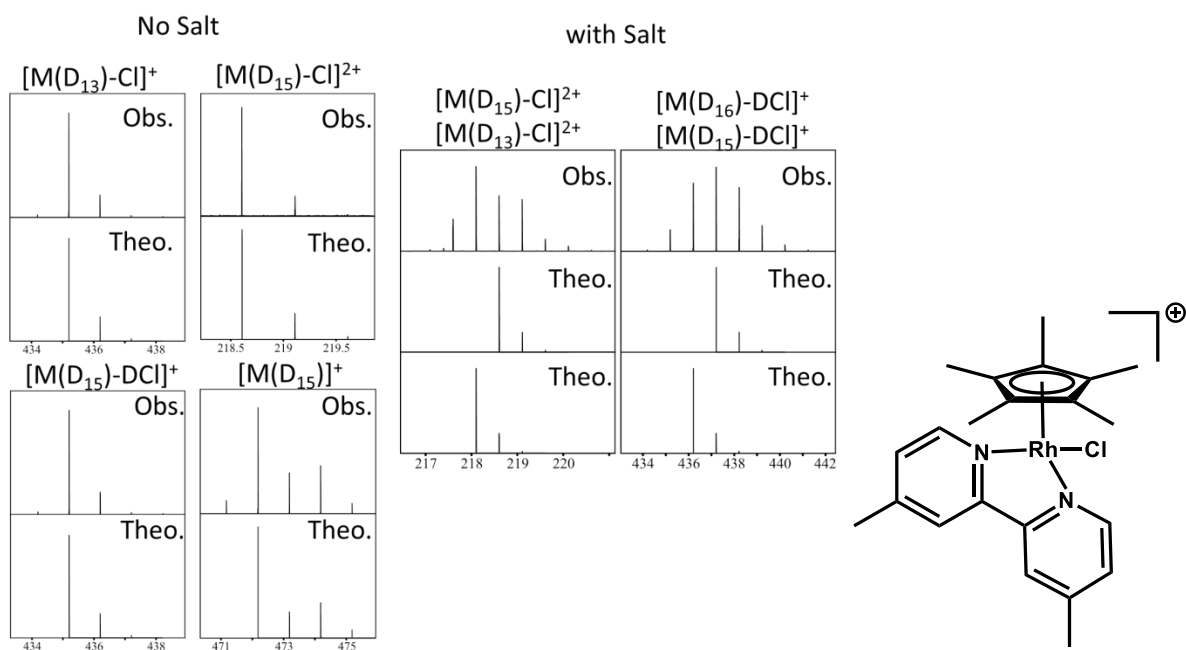
Complex 6: [Rh(Cp*-phenyl)(Me-Bipy)Cl]⁺

[M(D$_{12}$)-DCl]⁺
[M(D$_{13}$)-DCl]⁺

[M(D$_{13}$)-DCl]²⁺
[M(D$_{14}$)-DCl]²⁺
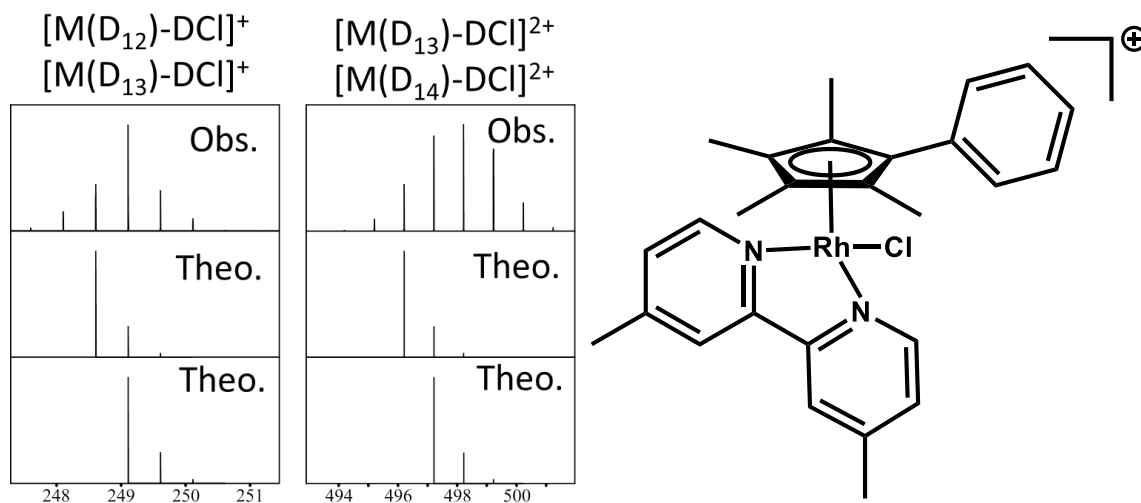
Complex 7: [Rh(Cp*-biphenyl)(Me-Bipy)Cl]⁺

No Salt

[M(D$_{12}$)-DCl]²⁺ [M(D$_{16}$)-DCl]⁺
[M(D$_{12}$)-Cl]²⁺   [M(D$_{17}$)-DCl]⁺

with Salt

[M(D$_{16}$)-DCl]²⁺
[M(D$_{16}$)-Cl]²⁺
[M(D$_{12}$)-DCl]⁺   [M(D$_{17}$)-Cl]²⁺

Complex 8: [Rh(Cp*)(Phen)Cl]⁺

[M(D$_{15}$)-DCl]²⁺   [M(D$_{15}$)-DCl]⁺
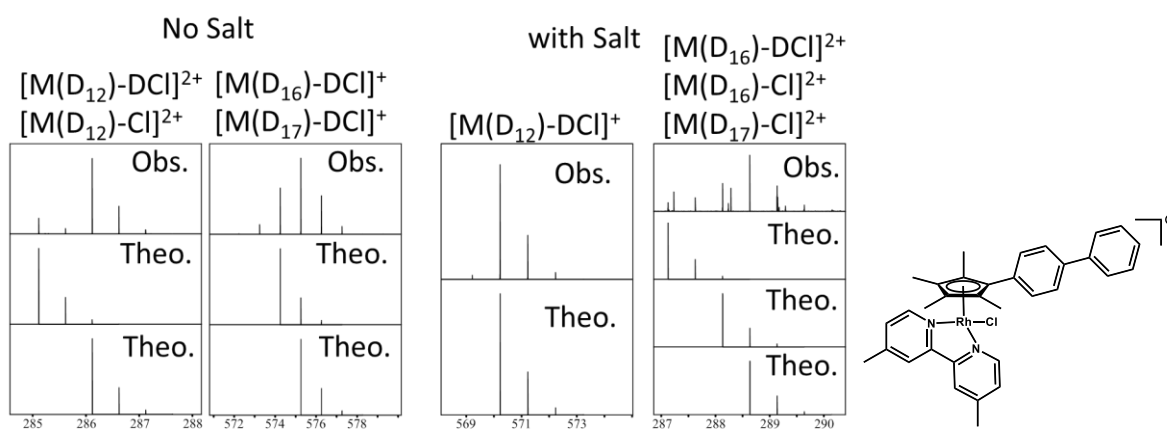
[M(D$_{15}$)-DCl]²⁺
[M(D$_{15}$)-Cl]²⁺   [M(D$_{15}$)-DCl]⁺
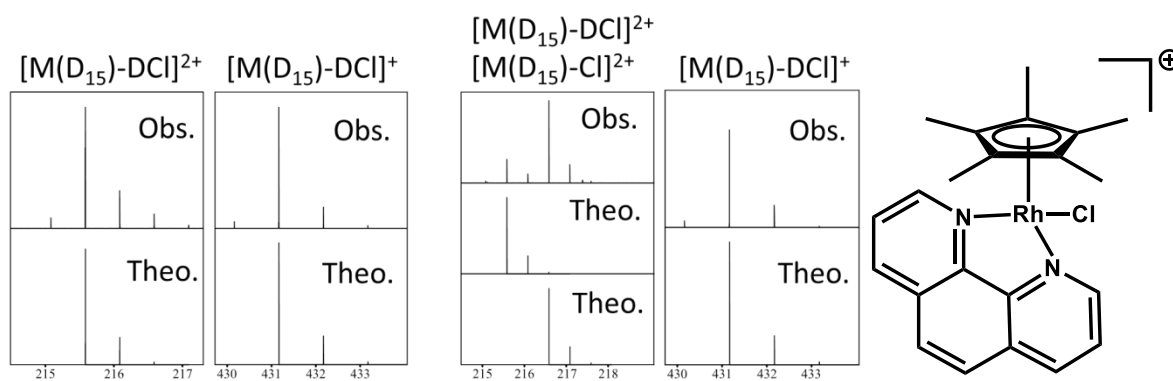
Complex 9: [Rh(Cp*-phenyl)(Phen)Cl]⁺
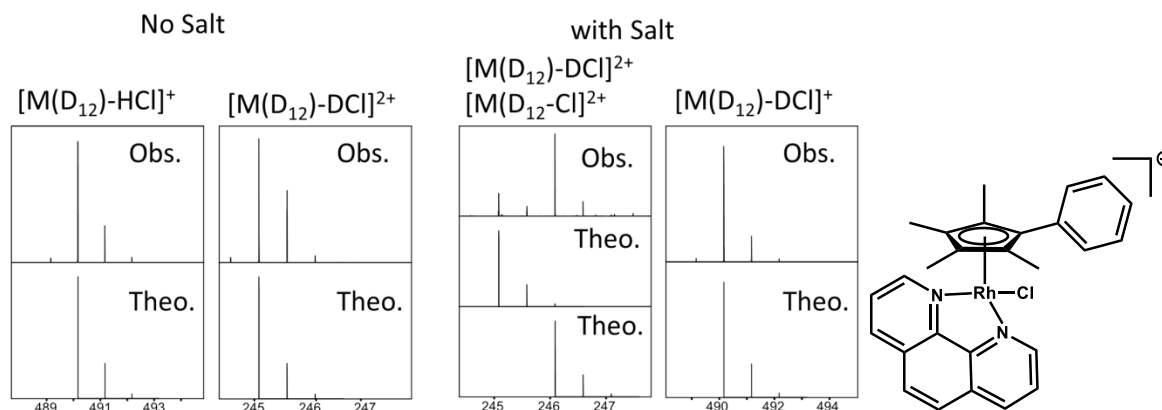
Complex 10: [Rh(Cp*-biphenyl)(Phen)Cl]⁺
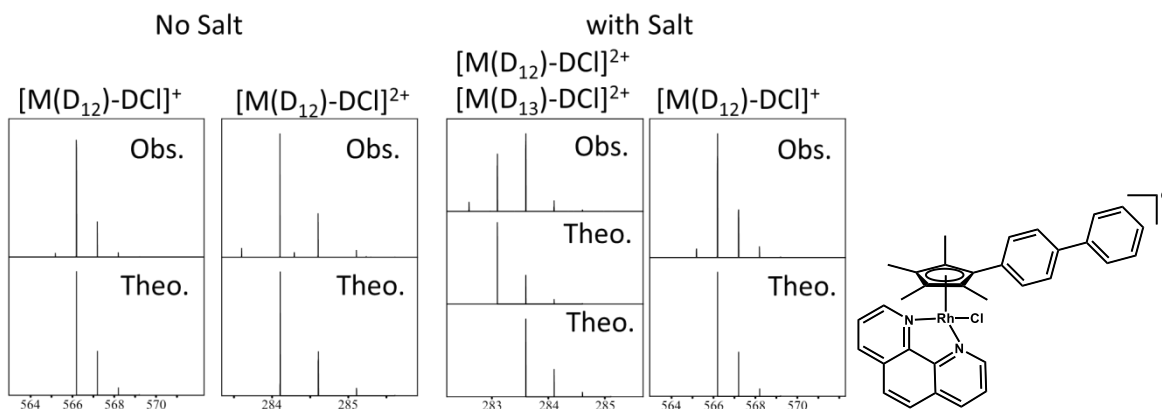
Figure 8.5: Structures, Mass spectra, and theoretical simulations of assignments for the 11 compounds studied for metallodrug HDX, all samples were analysed after 72 hours in deuterated solvent and both with and without excess (1mM) NaCl added, those species for which one set of spectra are provided indicate no added salt spectra, as the additional salt spectrum did not produce detectable target analyte signal.

#	Complex	Salt? y/n	Assignment	Assignment composition	#D's	Observed Mass	Exact mass
1	Rh-Cp-B	n	[M-DCl]+	$[RhC_{10}D_{14}N_2C_{10}H_8]^+$	15	407.16752	407.17113
		n	[M-DCl]2+	$[RhC_{10}D_{14}N_2C_{10}H_8]^{2+}$	15	203.58376	203.58529
		n	[M]+	$[RhC_{10}D_{14}N_2C_{10}H_8DCl]^+$	15	444.15045	444.15408
		y	[M-DCl]+	$[RhC_{10}D_{14}N_2C_{10}H_8]^+$	15	407.16779	407.17113
		y	[M-DCl]2+	$[RhC_{10}D_{14}N_2C_{10}H_8]^{2+}$	15	203.58388	203.58529
2	Ir-Cp-B	n	[M-HCl]+	$[IrC_{10}H_{14}N_2C_{10}H_8]^+$	0	481.134	481.13834
		n	[M-HCl+H2O]+	$[IrC_{10}H_{14}N_2C_{10}H_8H_2O]^+$	0	499.14451	499.14891
		n	[M-HCl+CD2HOH]+	$[IrC_{10}H_{15}N_2C_{10}H_8CD_2HOD]^+$	0	516.17871	516.18494
		y	[M-Cl]+	$[IrC_{10}H_{15}N_2C_{10}H_8]^+$	0	482.14189	482.14617
		y	[M-HCl]+	$[IrC_{10}H_{14}N_2C_{10}H_8]^+$	0	481.13608	481.13834
		y	[M-Cl+H2O]+	$[IrC_{10}H_{15}N_2C_{10}H_8H_2O]^+$	0	500.15306	500.15673
		y	[M-HCl+H2O]+	$[IrC_{10}H_{14}N_2C_{10}H_8H_2O]^+$	0	499.14747	499.14891
		y	[M-HCl+CD3OD]+	$[IrC_{10}H_{15}N_2C_{10}H_8CD_3OD]^+$	0	518.18476	518.19749
		y	[M-HCl+CD3OH]+	$[IrC_{10}H_{15}N_2C_{10}H_8CD_2HOD]^+$	0	517.18719	517.19121
		y	[M-HCl+CD2HOH]+	$[IrC_{10}H_{15}N_2C_{10}H_8CDH_2OD]^+$	0	516.18203	516.18494
3	Rh-Cpph-B	n	[M-DCl]+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8]^+$	12	466.16389	466.16795
		n	[M-DCl]2+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8]^{2+}$	12	233.08253	233.08370
		n	[M-Cl]2+	$[RhC_9D_{12}C_6H_5C_{10}N_2H_8]^{2+}$	12	234.08788	234.09075
		y	[M-DCl]+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8]^+$	12	466.16372	466.16795
		y	[M-DCl]2+	$[RhC_9D_{12}C_6H_5C_{10}N_2H_8]^{2+}$	12	234.08783	234.09075
4	Rh-CpBph-B	n	[M]+	$[RhC_9D_{12}C_6H_4C_6H_5C_{10}N_2H_8Cl]^+$	12	579.182203	579.18220
		n	[M-DCl]+	$[RhC_9D_{11}C_6H_4C_6H_6C_{10}N_2H_8]^+$	12	542.198697	542.19925
		n	[M-Cl]2+	$[RhC_9D_{12}C_6H_4C_6H_5C_{10}N_2H_8]^{2+}$	12	272.104581	272.10640
		y	[M-Cl]2+	$[RhC_9D_{12}C_6H_4C_6H_5C_{10}N_2H_8]^{2+}$	12	272.10281	272.10640
5	Rh-Cp-M	n	[M-Cl]+	$[RhC_{10}H_2D_{13}C_{10}N_2H_6CH_3CH_3]^+$	13	435.1999	435.20398
		n	[M-Cl]2+	$[RhC_{10}D_{15}C_{10}N_2H_6CH_3CH_3]^{2+}$	15	218.60586	218.60799
		n	[M-DCl]+	$[RhC_{10}D_{14}C_{10}N_2H_6CH_3CH_3]^+$	15	435.1999	435.20243
		n	[M]+	$[RhC_{10}D_{15}C_{10}N_2H_6CH_3CH_3Cl]^+$	15	472.18399	472.18538
		y	[M-Cl]2+	$[RhC_{10}D_{15}C_{10}N_2H_6CH_3CH_3]^{2+}$	15	218.60621	218.60799
		y	[M-Cl]2+	$[RhC_{10}HD_{14}C_{10}N_2H_6CH_3CH_3]^{2+}$	14	218.10324	218.10485
		y	[M-DCl]+	$[RhC_{10}HD_{14}C_{10}N_2H_6CH_3CH_3]^+$	15	436.2059	436.21025
		y	[M-DCl]+	$[RhC_{10}D_{15}C_{10}N_2H_6CH_3CH_3]^+$	16	437.21201	437.21653
6	Rh-Cpph-M	n	[M-DCl]+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_6CH_3CH_3]^+$	12	494.19515	494.19925
		n	[M-DCl]2+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	12	247.09842	247.09935
		n	[M-Cl]2+	$[RhC_9D_{12}C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	12	248.10367	248.10640
		y	[M-DCl]+	$[RhC_9D_{12}C_6H_5C_{10}N_2H_6CH_3CH_3]^+$	13	496.20739	496.21335
		y	[M-DCl]+	$[RhC_9D_{12}C_6H_4DC_{10}N_2H_6CH_3CH_3]^+$	14	497.21337	497.21963
		y	[M-DCl]2+	$[RhC_9D_{12}C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	13	248.10457	248.10640
		y	[M-Cl]2+	$[RhC_9D_{12}C_6H_3D_2C_{10}N_2H_6CH_3CH_3]^{2+}$	14	249.11046	249.11268

7	Rh-CpBph-M	n	[M-DCl]2+	$[RhC_9D_{11}C_6H_4C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	12	285.11472	285.11500
		n	[M-Cl]2+	$[RhC_9D_{12}C_6H_4C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	12	286.11926	286.12205
		n	[M-DCl]+	$[RhC_9D_{11}C_6D_4C_6H_5C_{10}N_2H_6CH_3CH_3]^{+}$	16	574.24752	574.25566
		n	[M-DCl]+	$[RhC_9D_{11}C_6D_4C_6H_5C_{10}N_2H_5DCH_3CH_3]^{+}$	17	275.25584	575.26193
		y	[M-DCl]+	$[RhC_9D_{11}C_6H_4C_6H_5C_{10}N_2H_6CH_3CH_3]^{+}$	12	570.22592	570.23055
		y	[M-DCl]2+	$[RhC_9D_{11}C_6D_4C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	16	287.1265	287.12755
		y	[M-Cl]2+	$[RhC_9D_{12}C_6D_4C_6H_5C_{10}N_2H_6CH_3CH_3]^{2+}$	16	288.13143	288.13460
		y	[M-Cl]2+	$[RhC_9D_{12}C_6D_4C_6H_4DC_{10}N_2H_6CH_3CH_3]^{2+}$	17	288.63456	288.63774
8	Rh-Cp-P	n	[M-DCl]+	$[RhC_{10}D_{14}C_{10}N_2H_8C_2]^{+}$	15	431.16777	431.17113
		n	[M-DCl]2+	$[RhC_{10}D_{14}C_{10}N_2H_8C_2]^{2+}$	15	215.58405	215.58529
		y	[M-DCl]+	$[RhC_{10}D_{14}C_{10}N_2H_8C_2]^{2+}$	15	431.16717	431.17113
		y	[M-DCl]2+	$[RhC_{10}D_{14}C_{10}N_2H_8C_2]^{2+}$	15	215.58379	215.58529
		y	[M-Cl]2+	$[RhC_{10}D_{15}C_{10}N_2H_8C_2]^{2+}$	15	216.5894	216.59234
9	Rh-Cpph-P	n	[M-DCl]+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8C_2]^{+}$	12	490.16394	490.16795
		n	[M-DCl]2+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8C_2]^{2+}$	12	245.08273	245.08370
		y	[M-DCl]+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8C_2]^{+}$	12	490.16433	490.16795
		y	[M-DCl]2+	$[RhC_9D_{11}C_6H_5C_{10}N_2H_8C_2]^{2+}$	12	245.08285	245.08370
		y	[M-Cl]2+	$[RhC_9D_{12}C_6H_5C_{10}N_2H_8C_2]^{2+}$	12	246.08799	246.08705
10	Rh-CpBph-P	n	[M-DCl]+	$[RhC_9D_{11}C_6H_4C_6H_5C_{10}N_2H_8C_2]^{+}$	12	516.19428	566.19925
		n	[M-DCl]2+	$[RhC_9D_{12}C_6H_4C_6H_5C_{10}N_2H_8C_2]^{2+}$	12	284.10353	284.10640
		y	[M-DCl]+	$[RhC_9D_{11}C_6H_4C_6H_5C_{10}N_2H_8C_2]^{+}$	12	566.19467	566.19925
		y	[M-DCl]2+	$[RhC_9D_{11}C_6H_4C_6H_5C_{10}N_2H_8C_2]^{2+}$	12	283.09894	283.09935
		y	[M-DCl]2+	$[RhC_9D_{11}C_6H_3DC_6H_5C_{10}N_2H_8C_2]^{2+}$	13	283.60129	283.60249

Table 8.1: Assignments of all identifiable species in the metallodrug samples for complexes 1-11 after 72 hours in deuterated solvent. Please refer to above figure for structure of each species.

8.4 Stability tests:

Deuterated rhodium compounds which were shown by nESI-FT-ICR MS to be fully deuterated (D_{15}, or higher for different arene ligands), were then re-dissolved in pure protic solvents (H_6-MeOH and ACN) to test the stability of the deuterium back-exchange and see if the exchange was purely equilibrium based (due to excess of deuterated solvent), or was based on kinetic isotope effects favouring the deuterated species as a more stable composition.

Deuterated compounds were dissolved in LCMS grade MeOH/ACN (60/40 v/v) to match their previous solvent composition (above) and were sampled at 1 day, 2 days, 3 days, and 7 days to observe the protonation/stability. Surprisingly the compounds, once deuterated, did not seem to exchange bound deuterium atoms for the available hydrogens atoms in the solvent. Complex 1, i.e. Rh-Cp*-Bp species

(as the [M(D$_{15}$)]+ ion, [RhC$_{10}$D$_{14}$N$_2$C$_{10}$H$_8$DCl]$^+$) is shown below at t0 and at 7 days in protic solvents, showing very little back exchange with hydrogens:
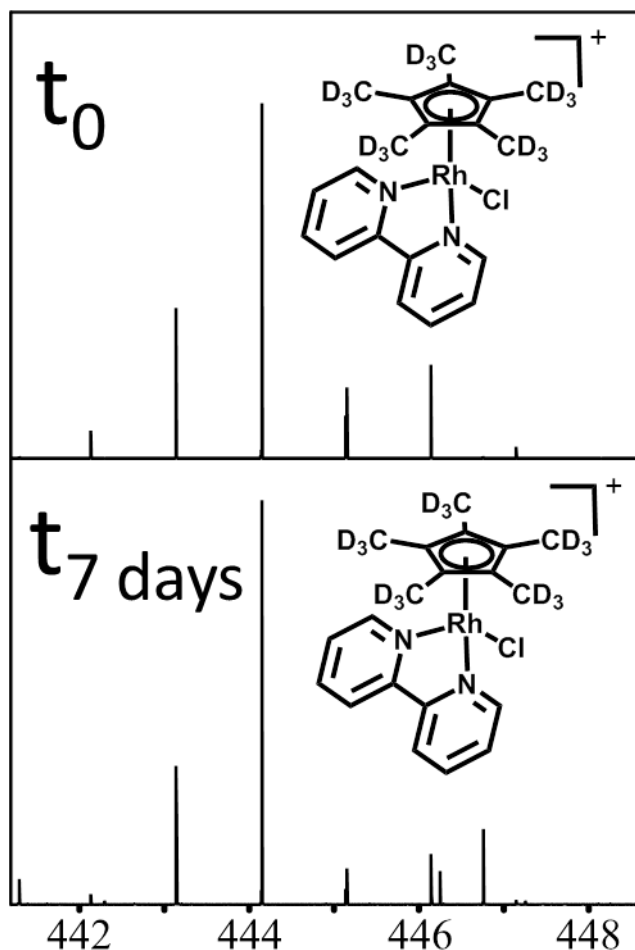
Figure 8.6 : nESI-FT-ICR MS spectra of the Rhodium compound Rh-(Cp*)-bipy), aka Complex 1, above (as the [M(D$_{15}$)]+ ion, [RhC$_{10}$D$_{14}$N$_2$C$_{10}$H$_8$DCl]$^+$), analysed immediately after deuteration and after 7 days in pure protic solvent (H$_6$-MeOH/ACN), showing no observable back-exchange with protic solvent, and remains deuterated.

The observation of such rapid deuteration in deuterated solvents, but little back exchange in protic solvent, even after 7 days, is a clear indication that the deuterated compounds are either much more energetically stable than the protonated analogues or are much more stable towards de-protonation/de-deuteration to achieve the Cp*=CH2/Cp*=CD2 intermediate shown in Figure 8.4.

8.5 Future outlook/prospects:

The facile deuteration of usually very inactive methyl groups is chemically very interesting and may open avenues in the future for other C-H bond activation, which is of exceptional interest in the field of catalysis.[2,3] Unfortunately this deuteration does nothing to enhance the metallodrug mechanism of action against cancerous cells, and does not have a practical use in the study of metal-biomolecule interactions, or could it? This section briefly discusses applications of this serendipitous discovery towards anti-cancer research and metallodrug-biomolecule interactions.

The main technique to exploit the facile deuteration is for accurate, easy, and stable isotopic labelling of rhodium compounds. It has long been the technique of chemists to isotopically label species in order to study changes in subsequent analytical experiments such as NMR, IR etc. However isotopic labelling has also been exploited by mass spectrometrists, commonly for proteomic quantitation/labelling, such as SILAC[4] and its high-resolution neutron labelled relative NeuCode.[5] These isotopic labels have assisted in quantitation, mass tagging specific amino acid residues, and mechanistic investigations of reactions during MS/MS.[5,4,6] Hydrogen-deuterium Exchange has also been used to study protein structure, with solvent-accessible amino acids exchanging hydrogen atoms for deuterium atoms available from solvent, either in the solution phase,[7] or in the gas/droplet phase during ESI-MS.[8]

However these approaches would not be assisted with the use of the presented rhodium compounds, even if they were shown to selectively target specific residues/functional groups the protonated species are equally as adept as a mass tag as the deuterated species.

One key feature of many transition metals is the unique and abnormal isotopic distribution the transition metal centre possesses, and thus the influence it would have when attached to other species, such as biomolecules. This influence and the characteristic patterns are evident in previous chapters for the different metals studied. The impact of these characteristic patterns is detrimental to mass spectrometry analysis, as they "spread out" the MS signal for a given species over many

isotopic detection channels, for example the Osmium Drug OS1 (chapter 4), contains one osmium

metal centre, which has 7 natural isotopes, instead of the signal for the Os1 drug being directed into

one detection channel (possible if all species had one observable isotope), the number of ions for

the Os1 compound are split into multiple ion clouds, differentiated by mass, this occurs for common

elements (such as CHNOS) but the unusual and extreme patterns from transition metals can in fact

decrease the overall intensity and signal to noise ratio of metal-bound species, compared to their

unmodified counterparts due to this "dilution" of signal into more masses.

The only beneficial consequence of these unusual metal isotope distributions is that they can quickly

highlight the presence of a given transition metal and can be used to identify metal bound species

(as done so in previous chapters). Unfortunately not all transition metals have a distinguishable

metal isotope pattern, and rhodium happens to be one of those species. Rhodium has one natural,

stable, observable isotope at 102.905504 Da, thus the detection and characterisation of Rhodium-

containing compounds is not assisted by the rhodium metal centre's presence. The use of either a

fully protonated (e.g. H_{15} from above) or fully deuterated (e.g. D_{15}) does not assist in this endeavour

either.

However since it has been shown that the deuterated rhodium compounds above are not only easily

converted from the protonated to deuterated species with little effort and within ~1 hour, they are

also stable for several days (at least 7-day stability shown above). These two factors allow for a way

of inducing a custom isotopic pattern for a given Rhodium drug sample – by mixing the two species.

If the rhodium Cmplex 1 H^{15} analogue was mixed with Complex 1 D_{15} analogue in a 50/50 molar ratio

then in subsequent MS spectra one would expect to observe 2 sets of peak distributions ~15Da apart

(H_{15}=15.117375Da, D_{15}=30.211527 => Δ=15.094152Da), and at ~equal intensity (assuming 100% D-

labelling and no kinetic isotope effects during ionisation, transmission, excitation, and detection

during FT-ICR MS). These species could then be easily identified as rhodium-containing based purely

on this mass difference.

The ability to detect rhodium species based on mass differences would also assist in the identification of rhodium complex-bound biomolecules as the species would also appear with ~15Da difference, allowing fast and effective identification even before MS/MS and without relying on the comparison of peaks to often unreliable/inappropriate biomolecule identification databases. This tactic for identification of modified biomolecules is discussed at length in Chapter 9, and was adapted into a specialist method for searching for transition metal isotope patterns. The deuterated labelling of rhodium compounds however allows the circumvention of the disadvantage some metals like rhodium have which is the lack of a characteristic Isotope pattern. In fact the ability to create custom-isotopic mixtures using different molar ratios of Dx:Hx complex allows for truly characteristic isotopic patterns. For example 50:50 molar ratio mixtures of deuterated complex with small numbers of deuterium atoms may overlap with Bromine containing species as Br also expresses a ~50:50 intensity ratio for its stable isotopes. However mixture of two unique masses, or indeed many isolated rhodium complexes, each with a different degree of deuterium labelling could be used to create custom isotopic patterns which could then be accurately searched for using the approaches discussed in Chapter 9.

The power of this simple idea expands to being able to track the mixed-rhodium complex species through tissue and cell areas via the use of either imaging mass spectrometry or using proteomic assignments, filtered for metal-containing species, to see where the metallodrug moves to in vivo (via imaging MS), and what it binds to/reacts with (by proteomics). This and other approaches are already under investigation in order to assess their potential for the tracking of these and other compounds for accurate analysis.

The remarkable stability of the deuterated products against back-exchange to their protonated form suggests an additional factor, outside of purely kinetic arguments, influencing the increased stability of the deuterated form over the protonated form. This factor could be structural, or subtle steric

factors, but warrants further investigation requiring molecular modelling calculations to elucidate these important influences.

8.6 References:

(1)   Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. *Org. Lett.* **1999**, *1* (6), 953.

(2)   Soldevila-Barreda, J. J.; Wolny, J.; Wootton, C. A.; Habtemariam, A.; Clarkson, G. J.; O'Connor, P. B.; Schunemann, V.; Sadler, P. J. *Submitt. to Nat. Chem.* **2016**.

(3)   Soldevila-Barreda, J. J.; Romero-Canelón, I.; Habtemariam, A.; Sadler, P. J. *Nat. Commun.* **2015**, *6*, 6582.

(4)   Ong, S.; Mann, M. *Nat. Protoc.* **2006**, *1* (6), 2650.

(5)   Merrill, A. E.; Hebert, A. S.; MacGilvray, M. E.; Rose, C. M.; Bailey, D. J.; Bradley, J. C.; Wood, W. W.; El Masri, M.; Westphall, M. S.; Gasch, A. P.; Coon, J. J. *Mol. Cell. Proteomics* **2014**, *13* (9), 2503.

(6)   Mann, M. *Nat. Perspect.* **2006**, *7*, 952.

(7)   Pan, J.; Han, J.; Borchers, C. H. *Int. J. Mass Spectrom.* **2012**, *325-327*, 130.

(8)   Somogyi, Á.; Harrison, A. G.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (12), 2055.

Chapter 9:

Scaling up – Moving from isolated biomolecule-metallodrug interactions to proteome-wide characterisation

Progress made in previous chapters/studies was conducted under the premise that it would (aside from provide interesting mechanistic investigations) be able to inform larger scale "proteomic" style investigations of target cells, i.e. human cancer cells, in order to provide MS derived insights into metallodrug effects and possible mechanisms of action against cellular components.

This chapter is focused on methodologies and subsequent results of such scaling-up studies and the challenges associated with them when attempting to incorporate metallodrugs and metallodrug-modified biomolecules, eventually attempting to create effective and accurate ways of studying these species.

The work presented herein was a collaborative effort and contains contributions from other researchers, though the MS-based acquisition of data was done by/with the Thesis Author, other parties did contribute to many steps and developments along the way, these are detailed below:

- The nLC-HR-ToF data was acquired by the Author at Bruker Daltonics UK, Coventry with the gracious assistance of Julia Smith

- Implementation of the Author-designed SNAP-LC concept into DA compatible VB code was conducted in collaboration with Matthew Willets, Bruker Daltonics, Billerica, USA

- Set-up, maintenance and operation of the EASY nLC-II was carried out both by the Author and by Yuko P. Y. Lam

- Cancer cell growth, treatment, and harvesting was conducted by Dr. Isolda Romero-Canelon

- Protein/Cancer cell lysis, digestion, and desalting was conducted graciously by Yuko P.Y. Lam, with some assistance from the Author

- Implementation of the 2D-FTMS pulse-program onto the SolariX FT-ICR MS was conducted by Maria A. van Agthoven, shown in her previous publications (referenced where appropriate). Author-created instrument parameters were used in conjunction with this standard 2D-MS pulse program to acquire data

- Processing and denoising of the raw 2D-FT-ICR MS data into a readable format, along with subsequent individual MS scan extraction was conducted by Maria A. van Agthoven

- Implementation of the 3D plotting of processed 2D-FTMS data was conducted using the code and interface created by Alice Lynch

The first section of this work, based around LCMS analysis of metallated biomolecules has since been used in the following submitted paper:

"Automatic Assignment of Metal-containing Peptides in Proteomic LC-MS and MS/MS Datasets"

Christopher A. Wootton, Yuko P. Y. Lam, Matthew Willetts, Maria A. van Agthoven, Mark P. Barrow, Peter J. Sadler, * Peter B. O'Connor. Submitted to RSC Analyst, 2016.

The study of isolated biomolecules and modified biomolecule species is essential in understanding the species itself and the various Post-Translations Modifications (PTM's) species they can be decorated with in vivo. Study of these species is also essential to understand the fundamental fragmentation patterns and mechanisms associated with each MS/MS technique and provide an accurate, detailed analysis of MS/MS spectra. The addition of modifications can complicate this further, be they natural, biological, synthetic, or as shown above organometallic. The unique modifications the body can impose on biomolecules are not only impressive, but vast on range, each imposing different requirements on sample preparation and MS tuning, instrumentation, and tandem MS application. Even common PTM's can vary from small molecule addition such as glycation[1] and phosphorylation,[2,3] to addition of protein subunits onto already large biomolecules/macromolecules e.g. ubiquitination[4–6] (addition of entire ubiquitin proteins of >7KDa) and Sumoylation[7] (addition of a range of SUMO proteins each ~12kDa). Metallodrug modified biomolecules are just as demanding, and as shown in previous chapters produce large arrays of modifications, require careful handling and controlled conditions, and each have unique effects on the resulting MS/MS spectra acquired. Unfortunately regardless of the MS/MS technique metal-bound biomolecules seem to be consistently challenging to study, even using a range of dissociation techniques. Moreover the vast majority of metallodrug modifications are solvent-sensitive, with various ligands able to be displaced if using incorrect/reactive solvents (such as methanol for the iridium metallodrugs discussed previously). Even if the metallodrug-modified species are stable in an MS-compatible solvent many other challenges need be addressed for effective MS and MS/MS analysis, including (but not limited to):

- Use of extremely soft ionisation techniques to retain non-covalent/fragile interactions
- Delicate tuning of MS optics to avoid excess energy deposition into fragile/high charge state ions resulting in loss of modification

- Use of compatible solvent system, for the biomolecule, the bound/non-covalent modification, and the metal-bound ligands

- Aversion of acids and denaturing conditions which can disrupt equilibria and biomolecule/modification interactions

- Ergodic MS/MS (e.g. CAD / IRMPD) induced loss of modifications

- Metal-induced/enhanced sequence scrambling during MS/MS

- Metal-enhanced side chain/small molecule losses during MS/MS

- Extreme radical dissociation mechanism changes in electron-based dissociation techniques such as ECD, EDD, EID, hECD, and ETD

- Alteration of predicted masses and data analysis procedure to account for changes in dissociation mechanisms

- Partial loss of modifications during MS/MS causing creation of multiple products, which complicate spectra and data analysis.

- Modification induced complexity of spectra by increasing original number of species by orders of magnitude

- Increased complexity of isotopic pattern due to many metals having non-standard and varying ranges of isotopes and distributions; effects on resulting biomolecule-metal species patterns not only complicate spectra but leads to erroneous automatic data interpretation.

Despite these challenges isolated biomolecule-metallodrug interactions can be effectively characterised using a range of MS and MS/MS techniques, and so with care scaling up from these isolated interactions to multiple/many biomolecules is possible in theory.

9.2.1 – LC-MS of metal-modified biomolecules:

Many contemporary mass spectrometry groups have shown great progress in attempting to effectively characterise cell-wide protein and peptide species, often referred to as the study of Proteomics. Due to the tens of thousands of different poly(amino acid) species present inside any

cell, the field of proteomics deals with extremely complex mixtures of biomolecules, including species not based upon poly(amino acid) structures. These complex mixtures are almost ubiquitously studied using some form of liquid-chromatography (LC),[8–10] which although comes in a wide array of forms, is a solution phase separation technique and is thus easily compatible with electrospray ionisation.[11] The coupling of LC and MS has been shown to effectively characterise a range of proteomes and their components, from relatively simple proteomes such as yeast (~6000 proteins)[12,13] up to and including complex human proteome samples.[14,15] To fully realise the potential effects metallodrugs have on their human cancer cell targets and to effectively utilise the data and mechanisms developed in the study of isolated biomolecules, the next logical step for the study of novel metallodrug compounds and biomolecule interactions was to scale up to proteome wide investigations to hopefully observe the effect these new compounds have on each and every cellular peptide, protein, nucleotide and eventually even metabolite.

Though, as mentioned above there are a range of challenges and disadvantages to attempting to study metallodrug-modified biomolecules via Mass spectrometry, and these require specialised strategies to achieve accurate and truly representative data on the biomolecule mixtures in question. The "standard" proteomic methods such as use of acidified water/Acetonitrile (ACN) gradients for effective LC separations and harsh instrument tuning parameters to achieve good desolvation and adequate sequence coverage in subsequent MS/MS experiments are both not tolerable when studying metallodrug-modified species. Furthermore, the standard automated methods for MS and MS/MS peak-picking, analysis and database searching of species against genome-derived biomolecule predictions will produce extremely erroneous data when studying metal-modified species as these modifications are not yet known in many cases, will not be included in database search algorithms, and most importantly will be matched against the wrong isotope peaks from the MS and MS/MS due to the transition metal isotope pattern influences, see Figure 9.1, below:
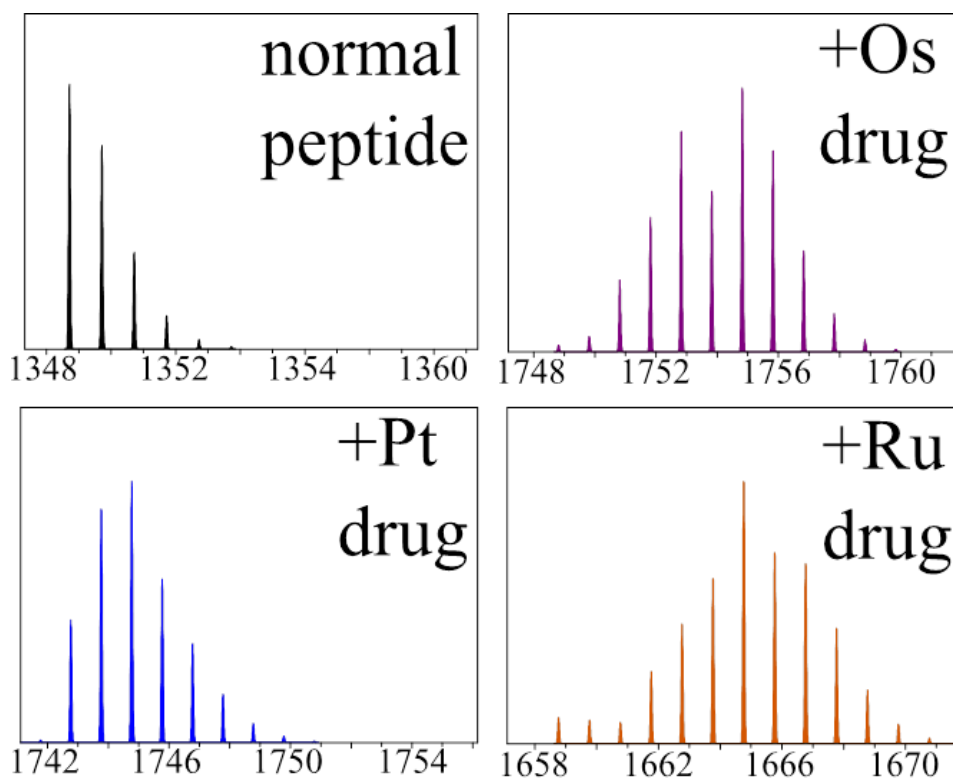
Figure 9.1: Typical Isotopic distributions for a small peptide, Substance P (top left) and then the same peptide modified with an array of transition metal modifications, showing the drastic effect on isotopic patterns some metals can induce.

Though the standard peak picking analysis could pick one or several peaks from the above distribution, only one would produce an accurate assignment for the modified biomolecule species. The consistent extraction of the correct mono-isotopic peaks for metal-modified species is more difficult than for simple CNOHS-constructed compounds such as unmodified peptides and proteins, as the number and intensity of individual isotopologues varies with each transition metal attached. However there have been numerous attempts to accurately pick and assign unmodified biomolecules in the past,[16] many of which centred around assigning one species to a group of peaks, i.e. the multiple isotopes for a given species. One particular advancement of this concept from Bruker Daltonics was the Sophisticated Numerical Annotation Procedure, aka SNAP.[17] The SNAP algorithm works by using a prediction of the isotopic pattern and matching it the to the observed peaks in a mass spectrum using a least-squares fit (Figure 9.2, below), any matched distribution are

then treated as one assignment, producing a much more accurate list of assignments and avoiding

multiple species assignments for species in the same isotopic pattern, which is an unfortunately

common occurrence as many assignable peptides in a genome are only a few Daltons apart.

Obviously since SNAP is based on matching isotopic patterns with observed peaks, knowledge of the

species is required, but unknown, as a result predictions have to be made to assign matching

distributions, namely the elemental composition of target species. Since peptide and proteins are all

based on amino acids, they are composed of a relatively limited number of different elements,

namely CHNOS. Moreover the proportion of the individual elements in poly(amino acids) has been

calculated and used in various studies by simply taking the human genome-predicted library of

peptides and proteins and calculating the number of each amino acid, then calculating the

composition of each and thus the contribution of each element to the "average biomolecule".

Named "averagine" the proportion of each element is thus set, and scalable for any mass of

biomolecule. The ratio calculated and used for the averagine proportions were C= 4.9384, N=

1.3577, O= 1.4773, S= 0.0417, and H= 7.7583.[16,18,17] Thus the composition of a poly(amino acid) of

say 1000 Da can be predicted by simply applying the averagine composition ratios. Obviously the

averagine predictions are more accurate with increasing mass of biomolecule, as small deviations of

"uncommon" elements (such as sulphur) in small peptides can cause large deviations in the accuracy

of averagine predictions. The Averagine assumptions are also slightly less accurate for non-standard

proteins with a significantly atypical amino acid proportion distribution, e.g. collagen which is often

>30% proline residues.

Figure 1: A least square fit of a MALDI-TOF spectrum of the peptide Bombesin. Squared sum of the difference between measurement and theory (z-axis) as function of the monoisotopic mass (x-axis) and the mass resolution (y-axis).

Spectrum
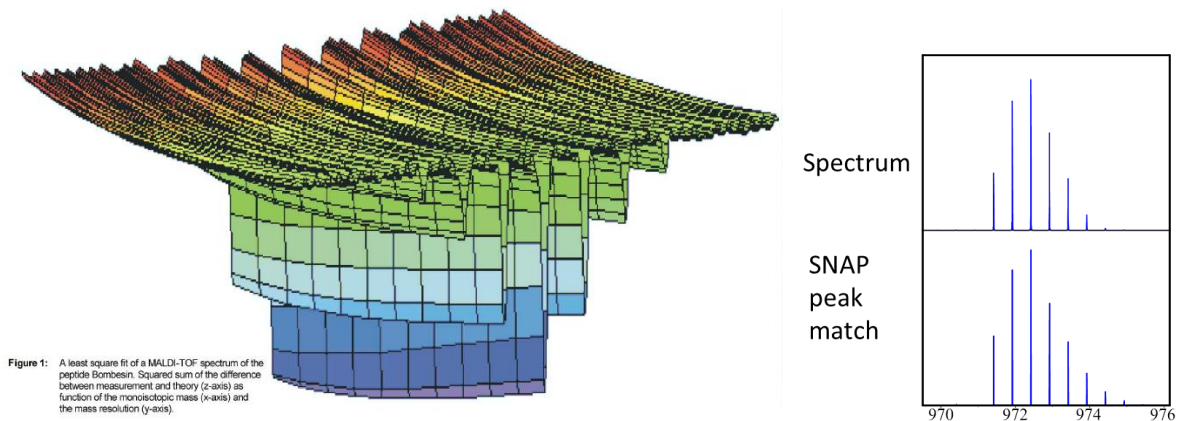
SNAP
peak
match

$$I\big(M,\ R,\ tpoint\big)_{theory}$$

$$\sum (M,\ R)^2 = \sum \big(I(t_{point})_{measured} - I\big(M,\ R,\ Tpoint\big)_{theory}\big)^2$$

Figure 9.2: Basis of SNAP annotation procedure from Bruker Daltonics,[17] showing (from top left) the least squares fit of the target isotope pattern for a modified peptide species (based on Bombesin), an example match of isotopic pattern for a biomolecule with a platinum metal centre attached showing SNAP to work with different elements, and the equations from Koster *et.al.*[17] which form the least squares fit SNAP uses.

The key and subtle benefit of the SNAP peak assignment method is that the elemental composition predictions can be modified to include additional species of interest, such as metals, and thus the SNAP algorithm can be modified to search and assign only metal-modified poly(amino acid) species. By inclusion of a "constant unit" being the metal, and retention of the "repeating unit" being the averagine ratio, the isotopic patterns of only metal bound biomolecules can be extracted and assigned accurately, which would be otherwise impossible using standard peak picking techniques. In addition species with one, two, or more bound metal modifications can be differentiated and extracted, further increasing the amount of information and accuracy of assigned species. An

example of the SNAP extraction of metal-containing biomolecules from a mass spectrum of both

unmodified and metal-modified peptides is shown below in Figure 9.3:
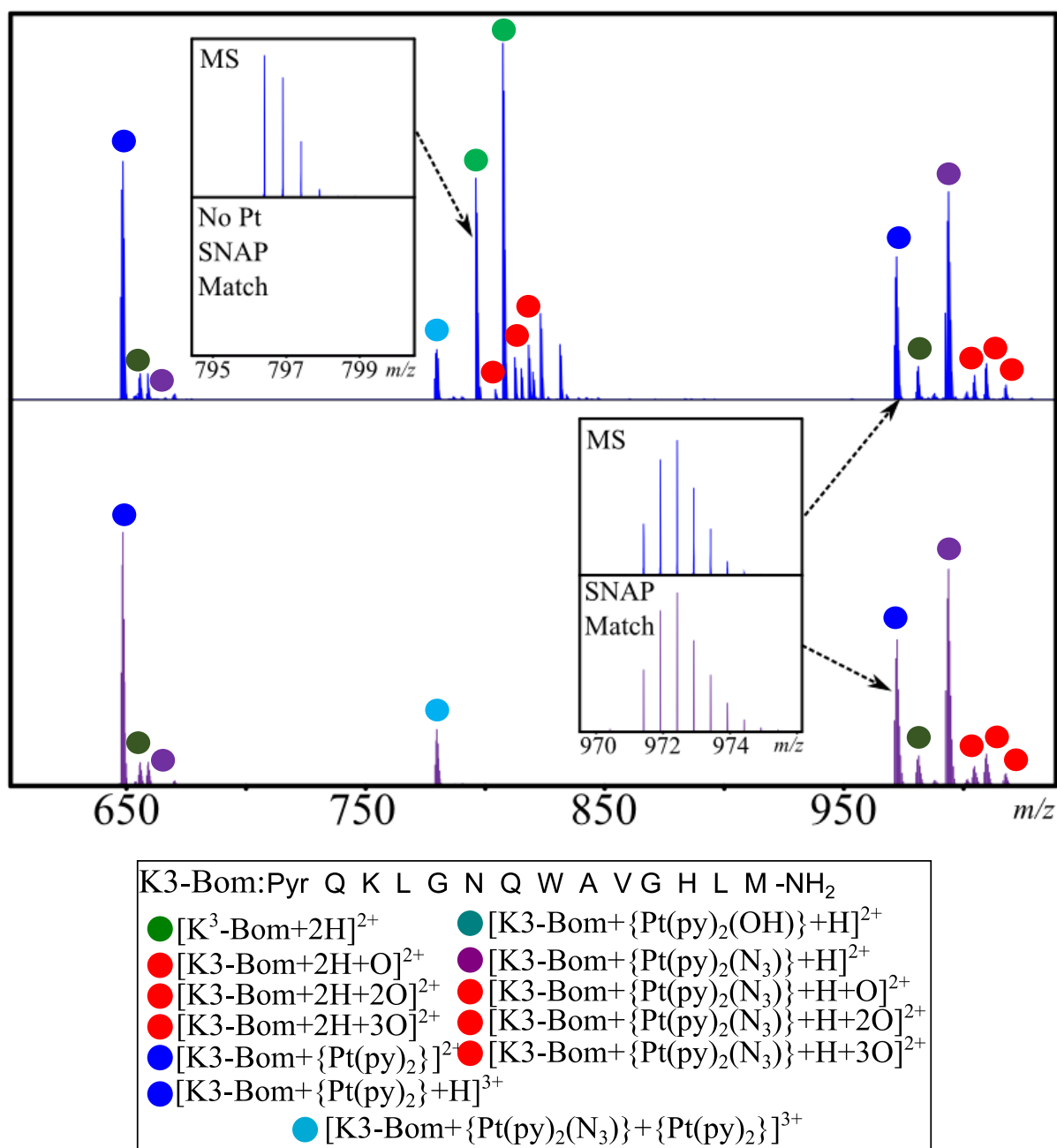
Figure 9.3: examples of modified SNAP peak picking to filter a given mass spectrum (top) for only

Platinum containing biomolecules, then output the matched isotopic patterns (middle) species are

marked according to the key presented and discussed in chapter 3 and included (bottom)Examples

inset are the SNAP algorithm ignoring normal peptide distributions (which do not include the

Platinum constant unit) and matching Pt containing species accurately with a predicted isotopic pattern.

The application of SNAP to mixtures of biomolecules is not only limited to assigning which peaks are metal-containing, and determination of how many metal species are bound to individual species, but also to filter/search for which metal/uncommon element may be bound to unknown species. SNAP could be used to search for any species with an uncommon isotopic distribution, which covers the majority of transition metal species, but also encompasses common halides etc. such as chlorine, bromine, and non-metals such as selenium. Since SNAP is based upon the matching of a theoretical isotopic distribution to the observed peaks in a spectrum, and no match will be perfect (due to sampling issues, natural deviations due to substance used, and detection limits), a statistical match score is applied to each match and this threshold can be adjusted in order to account for these small deviations. The match threshold is extremely useful, especially if the species contains many isotopes, some of which may be low abundance (e.g. <1% relative to the base peak) and frequently not detected. It was found that the highest match factor (0.95) was required for accurate assignment of metallodrug modified species in order to avoid assigning unmodified biomolecules simultaneously.

Previously however, SNAP assignment of peptides and proteins had been limited to individual mass spectra, be they MS or MS/MS spectra. The content of the first section of scaling up to proteome analysis of metallodrug effects was the application of SNAP to full liquid chromatography-MS experimental runs to analyse which spectra contained metal/uncommon isotope species, and then extract them into a SNAP-chromatogram of target peaks. The SNAP'd peaks could then be extracted, peak picked (more accurately using a metal-corrected SNAP assignment) and then searched/assigned either manually (for small chemical studies) or databases (for full proteome analysis).

The concept behind the SNAP-LC procedure was to load each chromatogram from an LC-MS run, apply the SNAP peak picking algorithm, specifying which metal to search for (e.g. Pt), to generate a peak picked spectrum, then store those peaks and move to the next peak, in theory then one could re-plot the LC chromatogram and see which LC peaks contained metal species.

In practice this means applying the modified SNAP parameter to each spectrum, retrieving the SNAP'd peaks, summing their intensities, storing these values into an array and then moving onto the next data point until all spectra had been analysed, at which point the array could be plotted simply as a new chromatogram using the spectrum number (rescaled to be retention time) and the summed intensity to produce a total ion chromatogram (TIC) but only for metal-bound species.

The automated procedure was achieved through the scripting functions within Data Analysis (Bruker Daltonics, Bremen, Germany), which is based upon Visual Basic (VB) coding language. The first attempts at this script were promising, but worked only for a limited number of spectra, regardless of the run chosen, it was eventually discovered that the SNAP-LC script would load each spectrum individually and remain loaded, quickly filling computer memory and preventing further calculations. A work around was then conceived that would output summed intensities into an external .csv file after each SNAP calculation/extraction, then delete the SNAP extractions before moving onto the next data point in the run via a loop. Although this work around deleted individual SNAP'd spectral assignment post summation, it worked effectively on entire LCMS runs (several thousand spectra) in one automated operation. The successful script (after several rescaling issue changes etc.) is shown below (Figure 9.4), annotated in figure caption for convenience.

```vbscript
SetLocale(1033) ' Localise for US (. instead , accepted as decimal
separator for calculations)
 Dim startSpectrum, NumSpectra, ProcessFullChrom, spectraToPro-
cess, firstMass, lastMass
'****************************************************
*
'*****SNAP each Spectrum and Export version 2
'****************************************************
*
ProcessFullChrom = False
startSpectrum = 1  ' specify spectrum number for method to start
from
NumSpectra = 4000    'in spectra
firstMass = 300 'lowest mass for SNAP
lastMass = 3000    'highest mass for SNAP
'****************************************************
*
'****************************************************
*
 spectraToProcess = 0
Call Main()
Call Form.Close()
   Sub Main
   Call ClearPreviousData
   spectraToProcess = NumSpectra+ startSpectrum
   If ProcessFullChrom then
     spectrToProcess = Analysis.chromatograms(1).size
     startSpectrum = 1
   End If
   Call CreateSpectra(startSpectrum, NumSpectra)
   Call AddChrom
   Analysis.Compounds.Clear
End Sub
 Function CreateSpectra(startSpectrum, NumSpectra)
     Dim spec, x, intensity, rt, rtinsec, count
   Dim results()
   ReDim results(NumSpectra,1)
   count = 0
       Set spec = analysis.spectra
     For x = startSpectrum To (NumSpectra+ startSpectrum)
     spec.add x, daProfileOnly
     spec.masslistfind firstMass,lastMass
     spec(1).ExportMassList "", daCSV
     rt = CDbl(ExtractDecimal(spec(1).name))
     rtinsec = rt*60
     results(count,0) = rtinsec
     results(count,1) = ReadIntensityfromText
     Analysis.compounds.clear
     Delete_File (Analysis.Path & "\Analysis.csv")
     count = count +1
   Next
     Call WriteIntensities(results)
 End Function
Function ExtractDecimal(str)
   Dim re, newstr, result
   Set re = New RegExp
     With re
```

```vbscript
       .Pattern = "(\d+\.\d+)"
       Set   newstr = .Execute(str)
    End With
    If newstr.count <> 0 Then
       result = newstr.Item(0).submatches.Item(0)
    End If
       ExtractDecimal = result
   End Function
 Function WriteIntensities(results)
     Dim objFSO, objTextFile, spec, y, sum, i, rt, rtinsec, count
     Const ForAppending = 8
     Set objFSO = CreateObject("Scripting.FileSystemObject")
   Set objTextFile = objFSO.OpenTextFile(Analysis.Path & "\" &
"intensities.csv", ForAppending, True)
     For count = 0 To UBound(results)
     objTextFile.WriteLine(count & "," & results(count,0) & "," &
results(count,0) & "," &  results(count,1))
   Next
       objTextFile.Close
   End Function
 Sub ClearPreviousData
     Analysis.Compounds.Clear
   Call Delete_File(Analysis.Path & "\Intensities.csv")
   Call Delete_File(Analysis.Path & "\Analysis.csv")
   End Sub
 Sub Delete_File (FileName)
     Dim fso
   Set fso = CreateObject("Scripting.FileSystemObject")
   If fso.FileExists(FileName) Then
     fso.DeleteFile(FileName)
   End If
 End Sub
 Sub AddChrom
   Dim Chrom
   Set Chrom = CreateObject
("DataAnalysis.ImportedDataChromatogramDefinition")
   Chrom.FileName = Analysis.Path & "\Intensities.csv"
   Analysis.Chromatograms.AddChromatogram Chrom
 End Sub
 Function ReadIntensityfromText
   Dim objFSO, objTextFile, strLine, intensity, line
   intensity = 0
   Dim path
   path = Analysis.Path
   Set objFSO = CreateObject("Scripting.FileSystemObject")
   Set objTextFile = objFSO.OpenTextFile(Analysis.Path & "\" &
"analysis.csv")
     strLine = objTextFile.ReadLine
       do while not objTextFile.AtEndOfStream
       strLine = objTextFile.ReadLine()
       line = Split(strLine, ",")
       intensity = intensity + line(4)
   loop
       objTextFile.close
   ReadIntensityfromText = intensity
 End Function
```

Figure 9.4: SNAP-LC script Version 2.0 – brief summary, from top left to bottom right: declaration of variables (black), selection of LC data points to search and masses for SNAP (red), clear previous data and selection of first LC data point, load first spectrum and clear analysis (blue), Declare variables for SNAP LC then apply SNAP using inputted parameters followed by export of results to .csv file (underlined) then clear results for next spectrum (Black), write intensity results to new separate .csv file outside of loop (orange), clear data to save RAM (purple), finally create new chromatogram from .csv file and re-import into Data Analysis for plotting.

The SNAP-LC code was then applied to a simple experiment (non-LC) showing 2 samples administered to the FT-ICR MS at separate time points, one an unmodified peptide, and the other a mixture of iridium-modified and unmodified species to see if SNAP-LC was correctly calculating only the contributions of the metal and metal-bound peaks, and not simply summing all peaks in a given spectrum. The results of which are shown below in Figure 9.5:
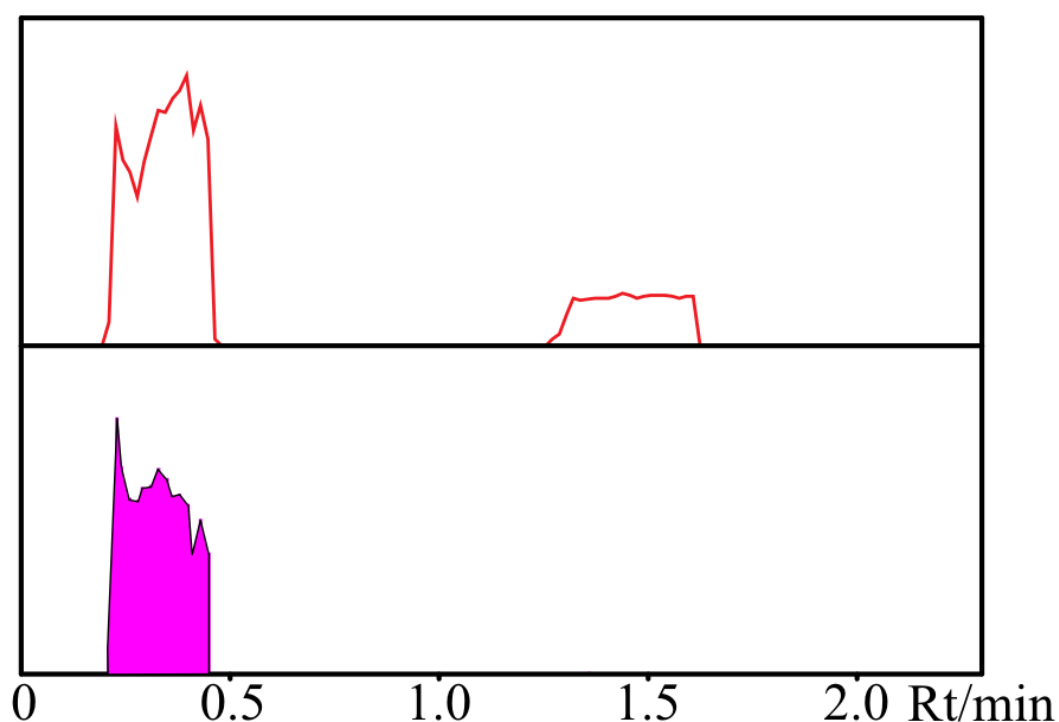
Figure 9.5: Initial results using SNAP-LC analysis on two ESI-MS samples one metal containing (Ir-peptide mixture, left), and one unmodified peptide with no metal (right). SNAP-LC was able to distinguish, purely from isotopic patterns, which sample contained iridium species.

The SNAP-LC search for Iridium modified species was effective at ignoring unmodified distributions, and only summing the Iridium-modified biomolecule peaks, resulting in a lower intensity for the chromatographic peak containing both unmodified and Ir-modified species, which is an important ability for analysis of co-eluting mixtures (i.e. all proteomic samples).

The SNAP-LC analysis method was then applied to larger peptide mixtures modified with various metal-based drugs. Standard proteins (namely Lysozyme, bovine serum albumin (BSA) and myoglobin) were reduced, alkylated, and then trypsin digested using very standard procedures (detailed below in 9.31 - cancer cell preparation) to produce peptide mixtures of varying complexity. The peptide mixtures were then incubated with various metallodrugs mentioned previously in the thesis (namely Ir1 and Ir2, Os1, and FM190) using the same incubation/activation times and conditions as for their reactions with previous peptide and biomolecules. The resulting metal modified peptide mixtures were analysed using nESI-FT-ICR MS to ensure metal-peptide reaction products had been formed, examples of these spectra are shown below in Figure 9.6, along with the corresponding SNAP'd metal distributions plotted underneath. Though SNAP for full MS and MS/MS experiments is effective at extracting and recording metal-modified species, like most isotope and peak picking algorithms it suffers with overlapping isotopic patterns and often "misses" metallated peaks if their pattern is overlapped with another species, be it metal-modified or not. This further enhances the importance of SNAP for LCMS analysis as the LC separation assists in reducing the number of species, and more importantly the number of overlapping species in resulting mass spectra, enabling SNAP to work more to its full potential and suffer less from these close distributions.
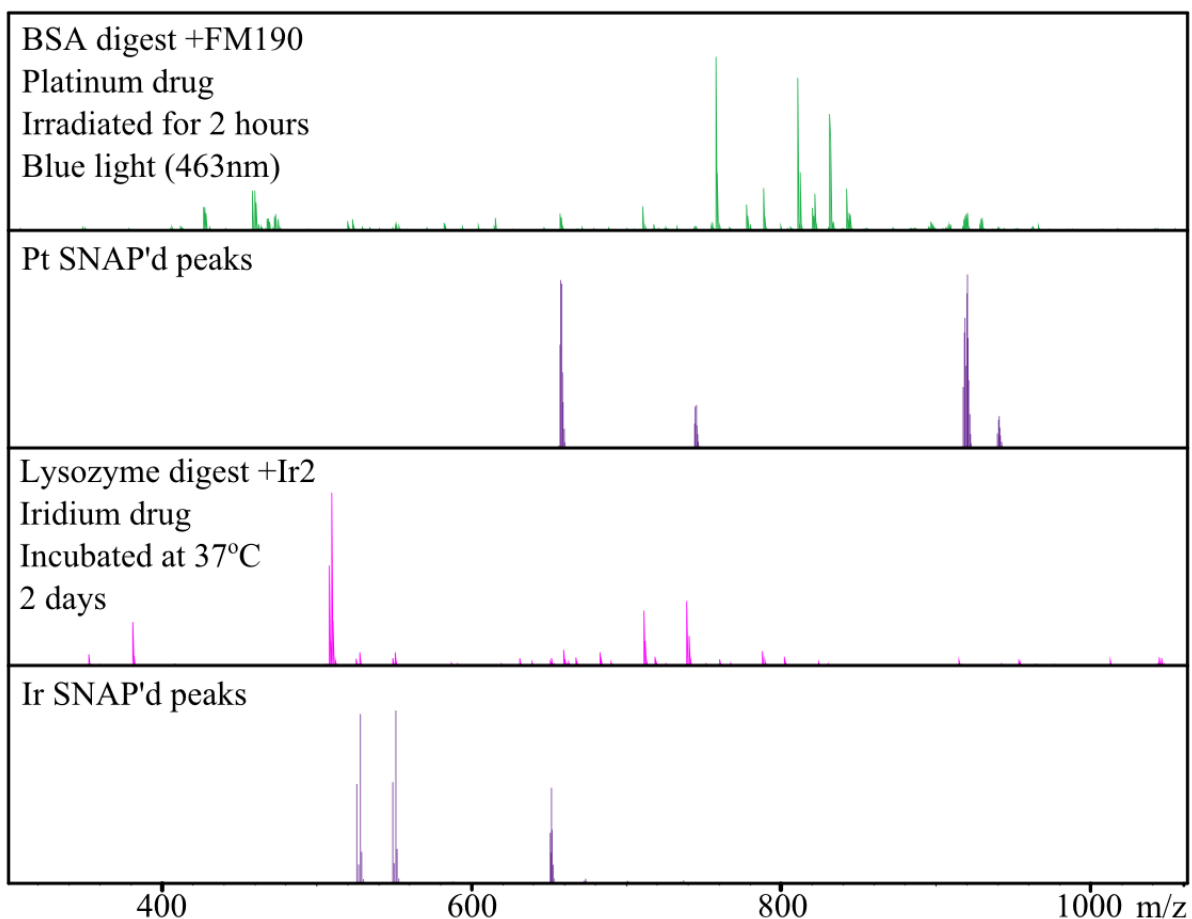
Figure 9.6: Example protein digest peptide mixtures reacted with previously studied metallodrugs nESI-FT-ICR MS spectra, followed by SNAP-LC metal isotope searching to highlight corresponding metal-containing species.

The peptide-metallodrug samples were then analysed by nLC-ToF-MS analysis using a Dionex Ultimate 3000 nLC, 25cm C18 reverse phase column, and a captive-spray micro-spray ESI source attached to a Maxis Impact II, High-Resolution Time of Flight mass spectrometer (HR-ToF-MS), (Bruker Daltonics, Bremen, Germany) operated in the positive ion mode. The ToF MS instrument provides fast scan times and reasonable resolution (40,000- 50,000 at 400 m/z for the Maxis Impact II) MS and MS/MS analysis.

Unfortunately the majority of LC-MS based instruments are usually configured to acquire MS and MS/MS spectra, peak pick each spectrum individually, and then delete the original (profile) data, replacing it with stick-plot type spectra of only peak-picked points. The process, known as

Centroiding,[19,20] assists in the reduction of the extremely large data files acquired during

chromatography-mass spectrometry runs (up to dozens of Giga-bytes) into manageable sizes

(~megabytes), and much faster display and processing speeds post-acquisition. Unfortunately it was

quickly discovered that accurate SNAP assignment of species is not feasible using centroided data

and requires full profile data to be successful. As a result spectra acquired in standard centroid-

mode were unusable for SNAP-LC post-acquisition analysis. A re-acquired profile nLC-HR-ToF-MS

chromatogram, MS, and MS/MS spectra shown below in Figure 9.7 for the Lysozyme digest+Ir2

metallodrug, example CAD MS/MS spectra are shown for a given time point. 2 precursors were

isolated and fragmented separately and dissociated according to CAD energies on the pre-set rolling

dissociation energy curve (CAD MS/MS was the only available MS/MS technique for the Maxis

Impact II).
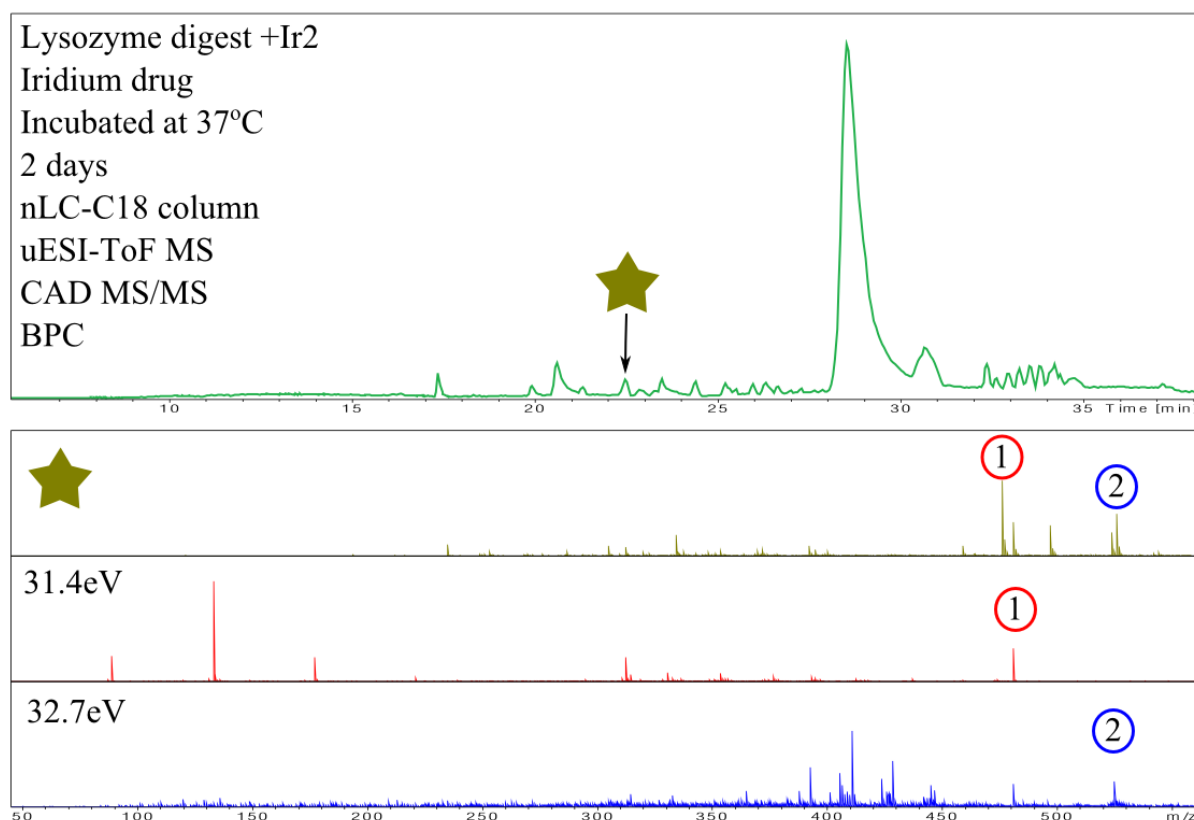
Figures 9.7a-d: Example nLC-HR-ToF base peak chromatogram (BPC) chromatogram (top) followed

by a mass spectrum from the time point marked (star) and 2 CAD MS/MS spectra from precursors

isolated and fragmented during the automatic LC MS/MS run (marked). The large peak in the BPC is the free Ir2 drug (-Cl).

The nLC-HR-ToF effectively separated, detected, and fragmented a series of peptides and metal-modified species. The SNAP-LC analysis of the resulting chromatograms though was not as effective as that for the FT-ICR data, though successful in some limited cases, SNAP-LC was shown to be more effective on FT-ICR MS and MS/MS data, this is currently believed to be due to the higher baseline noise level on the ToF and the difference in typical detected intensities between the two techniques, another possible hypothesis is the resolving power requirements and maximum # of assignments (see below). Though with continued improvements the SNAP-LC method could extend to LC-ToF runs, this is future work and the focus of the thesis will remain on FT-ICR MS characterisation of target metal-modified species.

One stipulation for accurate SNAP assignment of peaks containing specific isotopic pattern is that the individual isotopes need to be very clearly isotopically resolved in order to generate reliable assignments and matches. For the majority of cases HR-ToF instrument are effective at resolving the isotopic patterns of peptides, small proteins, and even medium sized (~20kDa) proteins with some of the latest instruments.[21] However separation of complex MS/MS spectra and even MS spectra of higher molecular weight species is more achievable using FT-ICR MS, additionally FT-ICR MS has access to the range of MS/MS techniques to effectively study the unique metal based modifications presented in this Thesis (particularly ECD and EDD).

nLC-FT-ICR MS analysis of peptide-metallodrug reaction mixtures described above was achieved using an EASY nLC II (Proxeon, Denmark), equipped with a 10 cm C18 reverse phase capillary column (10 cm x 75 um, 2 μm particle size) and 2 cm C18 trap column (2 cm x 75 μm, 2 μm particle size), coupled to a captive spray micro-spray ESI source (Bruker Daltonics, Bremen, Germany, attached to the SolariX 12 T FT-ICR mass spectrometer described earlier. FT-ICR MS has a much slower scan rate (/longer "duty cycle") compared to ToF-MS, mostly due to the fact that the resolving power for FT-

ICR MS is directly proportional to the acquisition time/duration of signal transient (discussed in the

introduction section), and so typical high resolution scans for an FT-ICR MS instrument are on the

order of 1-3 seconds, not including additional accumulation times for externally generated ions etc.

In summary for a high resolution (500,000-1,000,000 RP) scan on a 12 T FT-ICR MS instrument the

duty cycle is ~2-5 seconds, whereas many ToF mass spectrometers can achieve several if not dozens

of scans per second.[21] However the duty cycle of FT-ICR MS can be greatly reduced by acquiring

shorter transients, of say 350 ms, reducing the resolving power from ~500,000 to ~125,000 (at 400

m/z) but allows the scan rate to increase to around 1Hz, which is the standard benchmark for LCMS

scan rate.[8,22] The nLC-FT-ICR MS data shown herein was acquired using 1MW data points, equating

to ~350 ms transients and the ~125,000 resolving power (at 400 m/z) stated above. Though higher (4

MW, 500,000 RP) transient acquisitions were also attempted, the shorter scan time provided the

improved LCMS chromatogram quality, and would enable more MS and MS/MS scans for the many

species eluting from the peptide mixtures. An example of nLC-FT-ICR MS analysis of Cytochrome c

digest peptides is shown below in Figure 9.8, showing the ability of the EASY nLC II system to resolve

peptide peaks effectively despite using such a short (10 cm) C18 nLC column, and despite a fast

gradient (0% ACN to 100% ACN in 25 minutes):
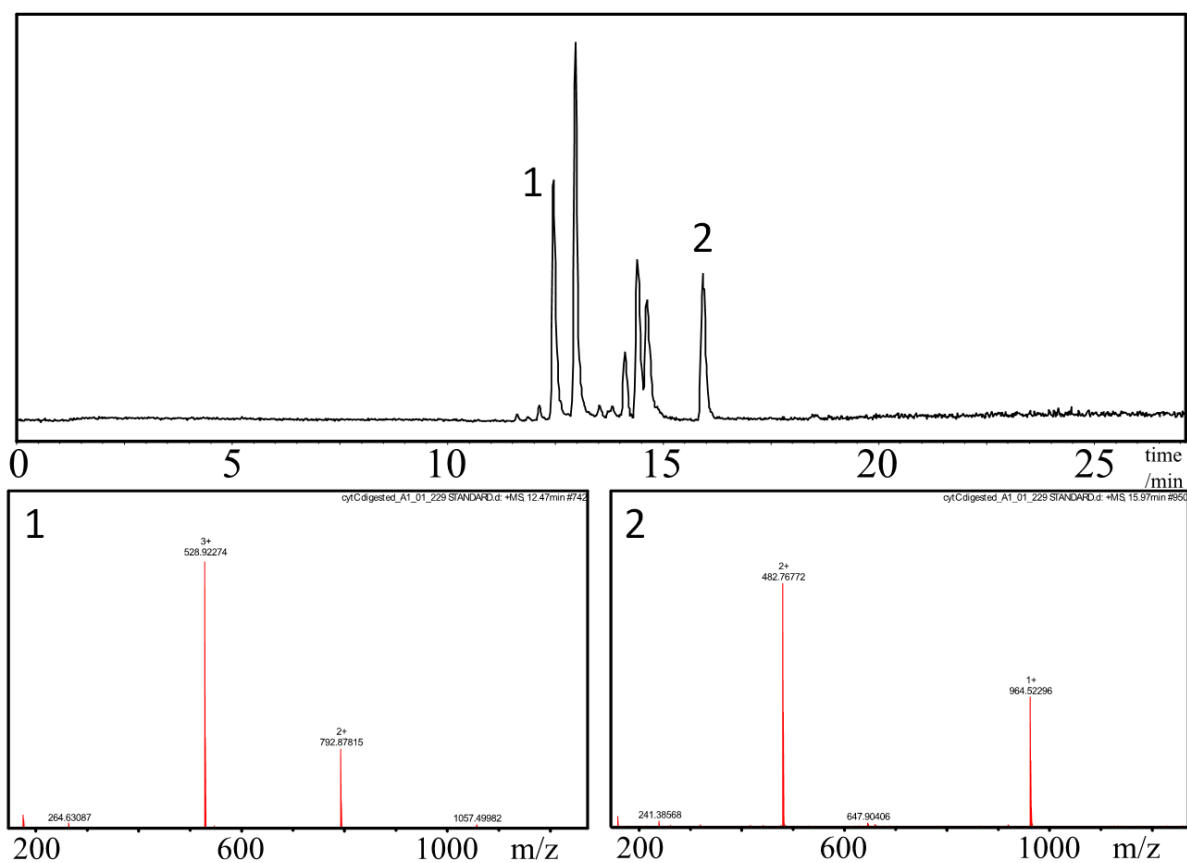
Figure 9.8: nLC-FT-ICR MS analysis of unmodified cytochrome c digest peptides via a 10cm C18 column, 2 µm particles. Effective separation of all peptides was achieved using a particularly steep gradient and short run time (see main text). Below: typical peptides observed and resolving power for annotated peaks. RP= ~125,000 at 400m/z.

The nLC-FT-ICR MS analysis was then attempted on an iridium metallodrug-peptide mixture, shown below in Figure 9.9; separation was achieved using the shallow gradient shown inset. The nLC performance was not as consistent as with the unmodified peptides, believed to be partly due to the acidified solvent conditions (0.1% formic acid in both the water and ACN solvents), the acidic solvents can cause dissociation of the iridium metallodrugs, which was observed in the extreme peak towards the end of the nLC run, identified as the free iridium metallodrug (shown inset). Despite the nESI FT-ICR MS spectrum of the same mixture (which showed little free Ir drug), the acidified solvent system used has disrupted the binding of the metallodrug to the peptides, again re-enforcing the

need for particular solvent systems and aversion of acidic solvents when studying metallodrug-

biomolecule interactions.
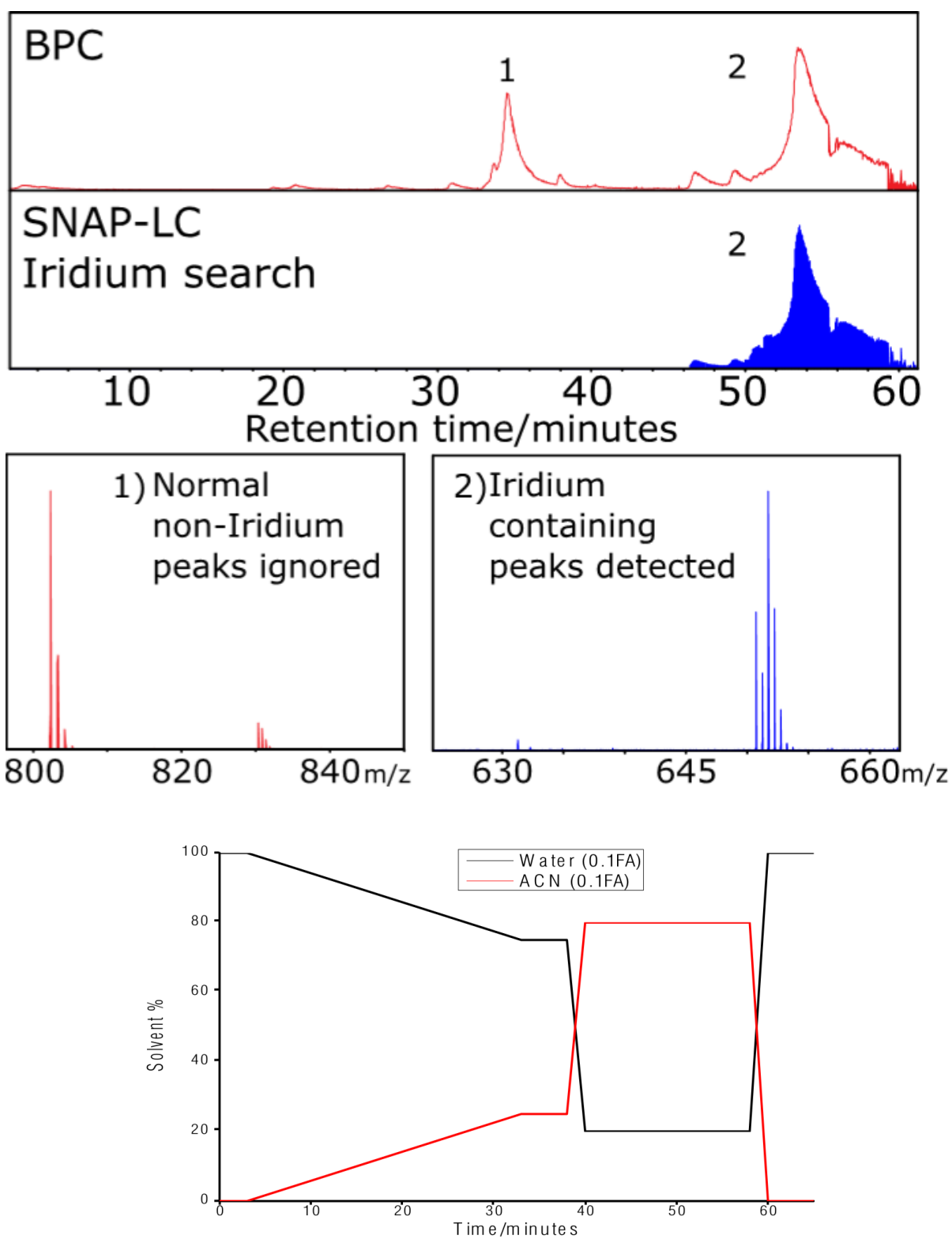
Figure 9.9: nLC-FT-ICR MS of Lysozyme digest peptides reacted with iridium metallodrug (IR2 – see previous section) producing a base peak chromatogram (BPC) of species (top) followed by SNAP-LC analysis to search for iridium isotope patterns (Blue). Examples of non-Ir containing species not counted in SNAP-LC (1) and species with the Iridium isotope pattern being successfully picked up (2). Finally, the nLC gradient used for separation on the 10 cm, C18, 2 μm particle nLC column.

Regardless of the solvent and dissociation issues observed, a range of peptides, and iridium-containing isotope patterns were recorded in the LCMS run, and provided a good trial for SNAP-LC analysis. The SNAP-LC analysis of the iridiated species shown above in Figure 9.9 (inset 2), iridium has a unique isotope pattern, shown above, and often extends peptide isotopic patterns when bound, unfortunately this modification usually results in 2 separate peak distributions being picked out with "standard" peak picking algorithms for peptide analysis (including standard unmodified SNAP), one being the first two (lower intensity) peaks and the other being the latter 3 peaks. However SNAP-LC with the iridium metal modification included accurately assigns one mass to the iridiated peaks which would enable accurate assignment in database/post-processing searches.

Automatic CAD MS/MS was attempted on the same iridiated Lysozyme digest sample, using a constant CAD collision energy of 18V for all species and an additional 1.5s "MS/MS boost" (ion accumulation time) when a desired precursor MS/MS target was located, the iridiated peptide shown above in Figure 9.9 was successfully detected and fragmented automatically, resulting in the CAD MS/MS spectrum shown in Figure 9.10, below, and can be used to locate the iridium based modification on the digest peptide.
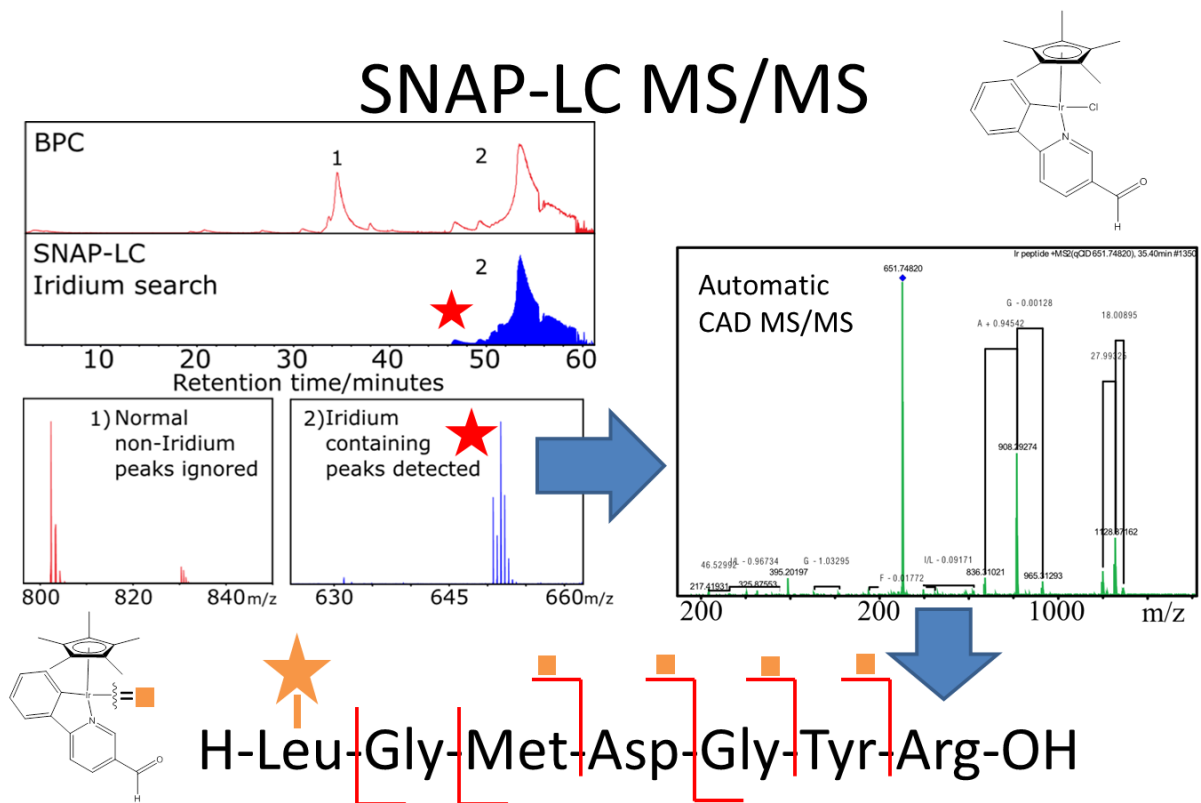
Figure 9.10: Summary of the SNAP-LC results with auto-CAD MS/MS of an Iridiated peptide detected during the nLC-FT-ICR MS analysis of the iridium-lysozyme digest peptide mixture. A series of iridiated and non-Iridiated peaks can be observed, with sequence, fragmentation map and derived location of the modification (based on MS/MS information and results from Chapter 6) shown inset.

Looking to the periodic table of elements, there are a great number of elements whose isotopic pattern deviates from the averagine predicted pattern enough to be considered viable candidates for SNAP-LC MS analysis, these potential "targets" for SNAP-LC MS searches are summarised in Figure 9.11, below:

SNAP-LC
Available target species

Periodic table (SNAP-LC available target species) — elements H, He, Li, Be, B, C, N, O, F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, As, Se, Br, Kr, Rb, Sr, Y, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Cd, In, Sn, Sb, Te, I, Xe, Cs, Ba, Lu, Hf, Ta, W, Re, Os, Ir, Pt, Au, Hg, Tl, Pb, Bi, Po, At, Rn, Fr, Ra, Lr, Rf, Db, Sg, Bh, Hs, Mt, Uun, Uuu, Uub, Uuq.

*Lanthanide series: La, Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb

**Actinide series: Ac, Th, Pa, U, Np, Pu, Am, Cm, Bk, Cf, Es, Fm, Md, No

Periodic Table of the Elements
Natural Isotope Distribution

Key:
46 Pd — Atomic Number, Symbol
Isotopic Distribution
Most Abundant Isotope

NOTE: The isotopic distribution data is from the 1997 report of the IUPAC Subcommittee for Isotopic Abundance Measurements by K.J.R. Rosman and P.D.P. Taylor *Pure Appl. Chem.* **1999**, *71*, 1593-1607.

© 2002 – 2005 Dr. R. M. Whittal, Mass Spectrometry Facility, Department of Chemistry, University of Alberta

Figure 9.11: The periodic table of elements including commonly observable isotopic patterns, all species with multiple isotopologues in distinctly different ratios to CHNOS are potential targets for SNAP-LC. Reproduced from http://zwz.cz/f/allgood788/KNLS74SW/COMPUTER-periodic-table-isotope-distribution (accessed 16/04/2016).

9.2.3 - Current challenges for SNAP-LC:

Currently the largest challenges for SNAP-LC analysis are the LC related challenges of retaining metallodrug modifications, and the fundamental limitation of the SNAP procedure, i.e. assigning only peaks and not noise.

Addressing point 1, retention of metal-based modifications, it is possible to run LC and nLC gradients without acidified solvents, though this tends to reduce LC chromatographic performance due to the presence of un-protonated silyl groups within the C18 reverse phase columns reducing the polarity of the column, and thus the quality of separation of polar analytes. nLC-FT-ICR MS analysis of metal-modified species was attempted and successful and enabled the retention of iridium based metallodrug modifications. It is also worth noting that the nLC-ToF MS analysis of metallodrug-peptide mixtures were all collected using pure, non-acidified solvents, following 2 days of purging the nLC system with non-acidified solvents to remove any residual acid from the system.

Unfortunately although LCMS can be achieved without acidified solvents, it is not commonly achievable without the use of potentially disruptive organic solvents. LC is based upon the phase fractionation of species between the mobile phase and the stationary phase, and thus species are "removed"/eluted from LC columns by using solvents individual species have a stronger preference for than the stationary phase, for peptides and proteins this is (for the majority of cases) a water/organic solvent mixture, usually ACN or methanol. Though species such as the iridium metallodrugs above are reasonably stable in water/ACN mixtures, other species (especially platinum based metallodrugs) can be particularly vulnerable to certain solvents and their electron-donating groups (such as nitrogen lone-pairs for Pt), and so can easily displace the metal modification away

412

from the peptides in question, or displace metal-bound ligands and change the nature of the metal modifications themselves, resulting in an inaccurate representation of the biomolecule mixture. The choice of phases and solvents will continue to be a challenge for LCMS analysis of particularly fragile/sensitive species and modifications, though great strides are being made by particular groups to retain modifications, usually for accurate Top-Down MS/MS analysis of proteins and/or protein complexes, such as the Kelleher[14], Pasa-Tolic[23], and Chamot-Rooke groups. Use of monolithic columns[23] for fast and effective separation of proteins and peptides, with various solvents has already allowed the effective analysis of many species, and even proteomes using the Top-Down approach,[23,24] eventually this may be modified and applicable to systems which are more metallodrug-friendly.

Addressing the second challenge for SNAP-LC, the effective assignment of peaks in LCMS data analysis is a field-wide issue and is ever under investigation and improvement. The particular issues with SNAP-LC of species on a larger scale are not uncommon ones, but are mainly focused around the effective analysis of ion-generated peaks and not baseline-noise. SNAP-LC has the ability to only assign peaks above a certain intensity level (threshold), like many other peak picking methods. The adjustment of this level dictates which peaks are chosen for consideration during peak picking, and thus will dictate how accurate the algorithm can assign target peaks. Setting a high SNAP intensity thresholds will miss potential assignments, while setting particularly low thresholds could result in erroneous peak assignments if noise is included in pattern matches, but could also cause issues with SNAP assignment limits (only 20,000 peaks can be assigned in each spectrum), causing the algorithm to crash. The first issue can be tuned over several baseline runs, and tuned with correct controls, in much the same way as standard LCMS proteomic analysis. The limit of assignments issue should be readily solvable by achieving a reasonable level of separation in LC, as there should not be a point in which thousands of co-eluting peaks are present and causing >20,000 peak assignments. However if the threshold of SNAP assignments is set too low then baseline noise will interfere with assignments and thus will trigger issue 2, as a result they are both solvable, but related.

Although the use of SNAP-LC on fully optimised proteomic nLC-FT-ICR MS runs was not achieved in the timescale of the thesis work, it is a promising avenue of research into metallodrug-biomolecule interactions and has already been included into plans of future study and application.

The effective LC-based retention of metal modifications is difficult to achieve as slight changes in solvent systems can not only affect the equilibria of binding interactions, as discussed and observed above in LCMS application. Incompatible solvent choices can also cause dissociation of non-covalent or dative-covalent bonds and loss of bound ligands, which induces inaccuracies in analysis. Furthermore certain (particularly transition metal) species can react with LC column packing (silyl/hydrocarbon groups) and cause disintegration of packing material/functional groups, contaminating LC-MS runs and irreparably damaging LC columns.

Despite the advantages of LC-MS using FT-ICR MS, and even though it is capable of studying a wide variety of metal compounds and metallodrug-biomolecule conjugates (see above), there are notable challenges and disadvantages. From a purely MS position these challenges are dominated by 1) the slow duty cycle of high-resolution scans using FTMS instruments, 2) The lower fragmentation efficiency of the more advanced, and required, electron-based MS/MS techniques for metal-biomolecule MS/MS (ECD/EDD/ExD), and 3) the effective isolation and tandem MS fragmentation of each species/precursor as they exit the chromatograph.

In order to combat these disadvantages, and the various solvent system/column interaction/time based issues associated with LC-MS, another technique was investigated to explore its potential in studying metallodrug-proteome interactions; 2-Dimentional Mass Spectrometry (2DMS).

2DMS or 2D-FT-ICR MS, is a purely MS based technique which attempts to exploit the modulation of ions within an ICR cell during MS/MS fragmentation in order to fragment all precursors within a sample, and analyse the fragments of each species, correlating them to the precursor, without a traditional quadrupole/in-cell isolation such as SWIFT/CHEF (see introduction).

2D-FT-ICR MS was first introduced by Pfandler in 1987,[25] and was developed by Ross, Marshall *et.al.*[26] and more recently by Van Agthoven *et.al.*[27] The basis behind 2DMS is the same as that for 2D

NMR; to attempt to create a coherence or correlation between two measurable factors using a sinusoidal modulation to enable a Fourier transform in a second dimension, for 2D-NMR these are the chemical shift and j-coupling values, for 2D-FT-ICR the correlation is not as obvious but has been adapted to use MS/MS fragmentation and use of a modulating excitation/de-excitation pulse to achieve the needed variability. The consequence of this is that each different parent ion species can be modulated using a pre-MS/MS modulation pulse (see Figure 9.12, below), then fragmented using a radius-dependent in-cell MS/MS technique (such as IRMPD, ECD, ExD, or other laser MS/MS methods), creating a mixture of remaining parent ions and MS/MS fragment ions.

Gaumann *et.al.*[28] achieved modulation by incrementally increasing the delay period (t_1) between two identical excitation events. The first event was used to excite ions radially outward from the centre of the ICR cell. Then if t_1 is exactly an odd integer multiple of half a period of the ICR frequency of a given parent ion that ion will be de-excited by the second pulse back to the centre of the ICR cell before MS/MS and thus not be fragmented by the MS/MS method achieved at high radius only (e.g. In-cell CAD, ECD, or ExD).[26] However if t_1 is exactly an integer multiple of one period of the parent ion ICR frequency then the second pulse will excite those parent ions out to twice the previous ICR radius achieved at the end of the first pulse and thus will be fragmented by the high-radius-dependent fragmentation. t_1 is then incremented with each scan, which will alter the phase at which the ions were excited at ($=t_1\omega_p$) and thus the radius of parent ions in successive scans will be modulated at the frequency ω_p as a function of t_1. The fragments of those precursor ions within the fragmentation zone will also be modulated in the same way as their parent, and thus are correlated, there is coherence and thus the second dimension of mass spectrometry is accessible.
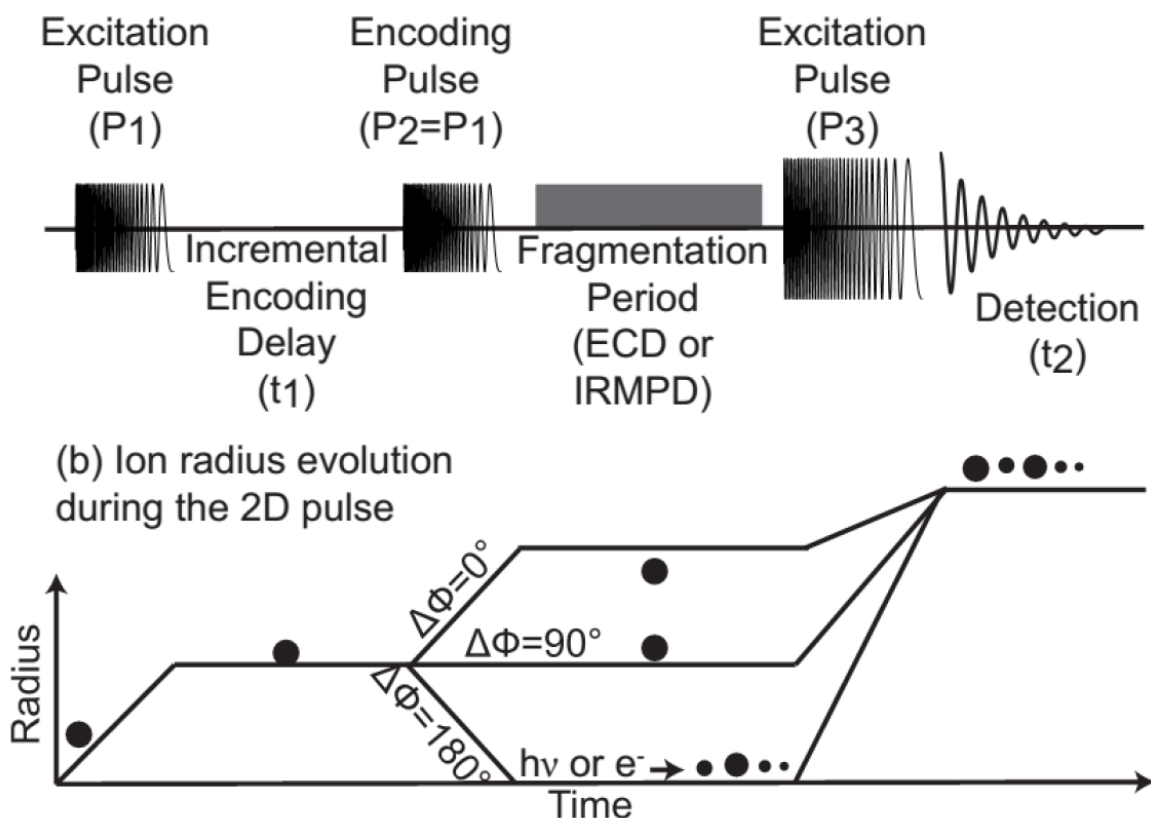
Figure 9.12: (A) Pulse sequence of a typical 2DMS experiment, the IRMPD/ECD pulse can be replaced with an ExD pulse. etc. (B) the subsequent effects of the 2DMS pulses depending on ion radius. Reproduced from Van Agthoven et.al.[29]

Within a modern hybrid-FT-ICR MS instrument, the ECD/ExD MS/MS fragmentations are usually achieved via a hollow electron-dispensing cathode mounted on the rear of the ICR cell, which is the case for the Bruker SolariX 12T FT-ICR MS used herein. The cylindrical geometry of electron production is preferred in order to leave a hole along the z axis to allow laser-based MS/MS within the cell. The different MS/MS techniques will thus have different target zones within the ICR cell, with the laser dissociations passing through the centre of the ExD cathode and irradiating the centre of the ICR cell, and the ExD electrons irradiating a circular "donut-shaped" path along the z axis. These 2 mutually exclusive "fragmentation zones" are depicted below (Figure 9.13), along with the 2DMS theory and pulses needed to achieve correlated, coherent MS/MS for successful 2DMS:
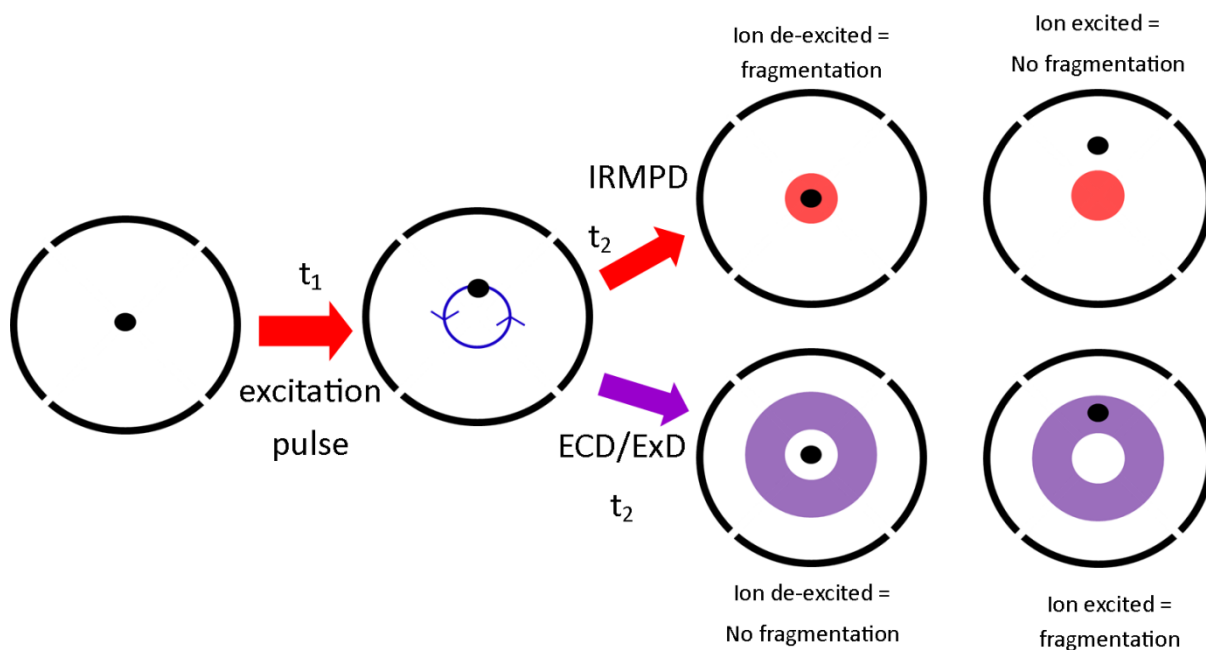
Figure 9.13: Representation of ion location during a typical 2D-FT-ICR MS experiment following the pulse sequence described above (Figure 9.12). Note: the ions in question will be rotating around the centre axis of the ICR cell at their reduced cyclotron frequency, as always, the stationary representation of ion location is purely for convenience/understanding.

Obviously for ECD 2DMS the modulation-correlation relationship is as described above; those ions who'll possess an ICR frequency period which is an integer multiple of one of t_1 will be excited by the second pulse and gain an increased radius, moving into the ECD MS/MS fragmentation zone. Whereas ions with a frequency that corresponds to t_1 being an integer multiple of half a period of the ICR frequency for those ions will be de-excited to the centre of the ICR cell and will not be fragmented by ExD MS/MS. For IRMPD 2DMS the correlation is simply reversed, those excited will not be fragmented and those de-excited are fragmented, the second pulse lengths are adjusted practically to achieve the correct amount of time within each zone and avoid loss of correlation, these values can be found in recent 2DMS studies achieved on the 12T SolariX instrument.[29] An in depth study of the optimisation of 2DMS pulse sequencing has also been carried out.[30]

The 2DMS file produced will be a series of transient signals all recorded with different incremental delays, say 4096 iterations, resulting in 4096 scans, and 4096 individual transients, and creating very large data files (>5-30Gb, depending on individual transient data point values). The data will then need to be processed, in a similar fashion to the standard FT-ICR MS data, requiring Fourier transformation (but in 2 dimensions this time), halving, zerofilling, and conversion to frequency using the acquisition rate, before calibration to create an m/z scale for both directions. The resulting 2D mass spectrum will contain a series of horizontal lines (4096 in the case above) being the individual scans, with varying increments. The Horizontal lines provide the individual MS/MS spectra acquired, the diagonal line across the spectrum will provide a reflection of the precursors which have been fragmented and has been dubbed the "auto-correlation line". Any lines with a horizontal offset to the auto correlation line show a consistent fragmentation from precursors, but no charge change, and so reveal neutral losses from, precursor ions, referred to as the neutral loss line. Vertical lines taken from the 2D spectrum will show all of the precursors to produce a fragment at the line chosen. The nomenclature and axis are summarised in Figure 9.14 below:
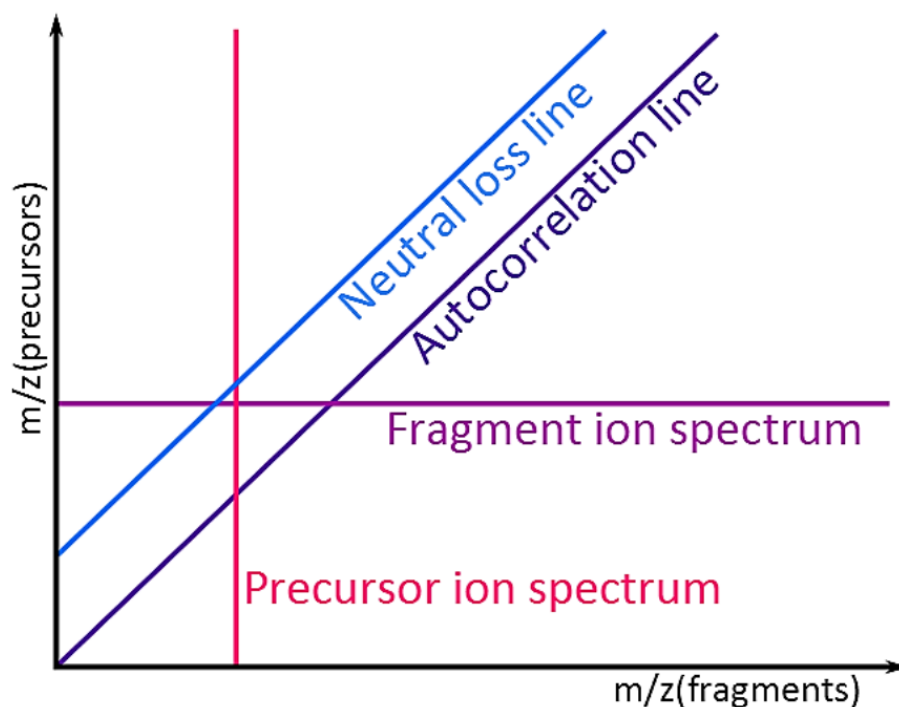
Figure 9.14: Axis, informative lines, and nomenclature for the interpretation of 2D mass spectra.

An example full 2DMS spectrum of cholesterol is below from Van Agthoven *et.al.,*[29] showing the various species produced from APPI of cholesterol, and then the IRMPD induced fragmentation of each and every species via 2D-IRMPD-FT-ICR MS is shown below in Figure 9.15:
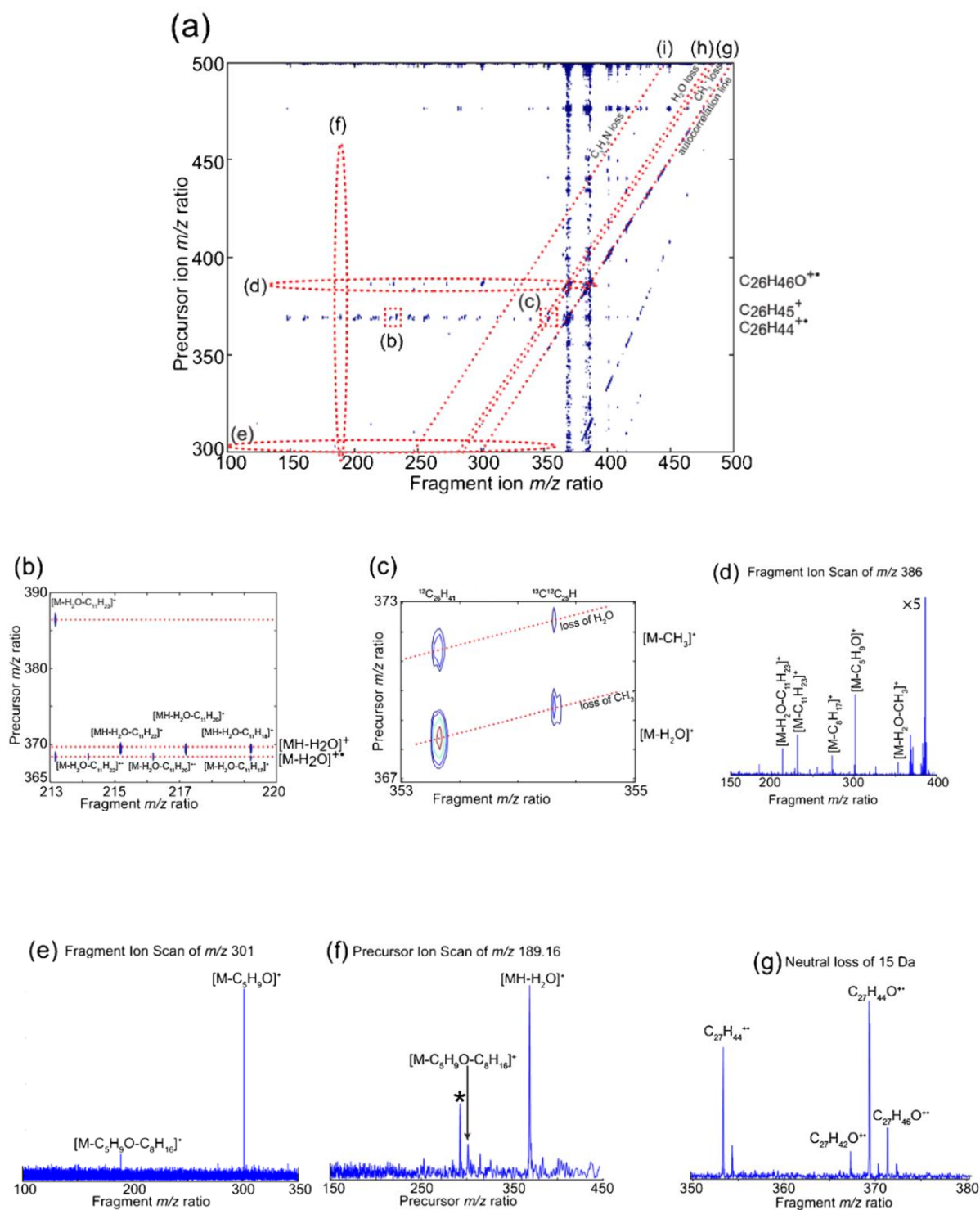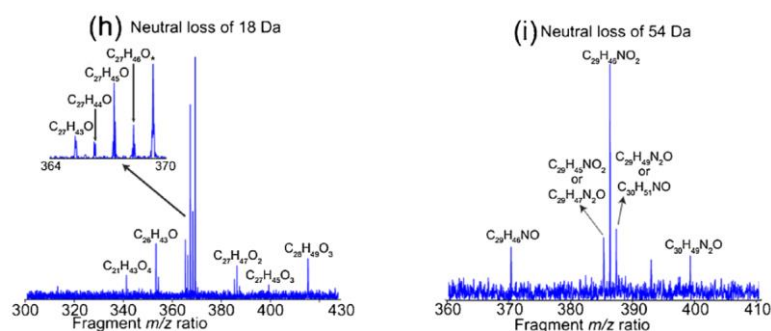
Figure 9.15: 2D-IRMPD-FT-ICR MS spectrum of Cholesterol ions produced by APPI (a) and various lines from the 2D plot, including fragment ion scans (b, c, d, e), precursor ions scan (f) and neutral loss lines (g, h, i). Reproduced from Van Agthoven *et.al.*[29]

Because 2DMS does not suffer from the reduced isolation efficiency associated with narrow-window quadrupole isolations, and does not require the individual tuning of In-cell isolations for each precursor, 2DMS is able to separate the fragmentation patters of extremely closely spaced peaks without loss of intensity, and was shown to be able to differentiate the different fragments produced by the Cholesterol-H_2O radical cation and protonated cation species, separated by 1m/z, which would be otherwise challenging via quadrupole isolation without loss of sensitivity.

2DMS has been successfully applied to such small molecules such as cholesterol,[29] but also isolated poly(amino acids),[27] and recently Van Agthoven, Wootton *et.al.* showed successful 2DMS of tryptic digest-derived mixtures of peptides. [31] With the recent advancements in 2DMS calibration and resolution,[31] 2DMS is finally becoming a viable technique for the study of biomolecules, and is even able to exceed standard MS/MS performance, as shown in Van Agthoven, Wootton *et.al.*[31] where fragments from peptide precursors just 0.0156 m/z apart were effectively separated in 2D space using 2DMS and were able to be characterised effectively. The separation of the 2DMS peaks and resulting fragmentation maps from the tryptic digest species shown below in Figure 9.16:
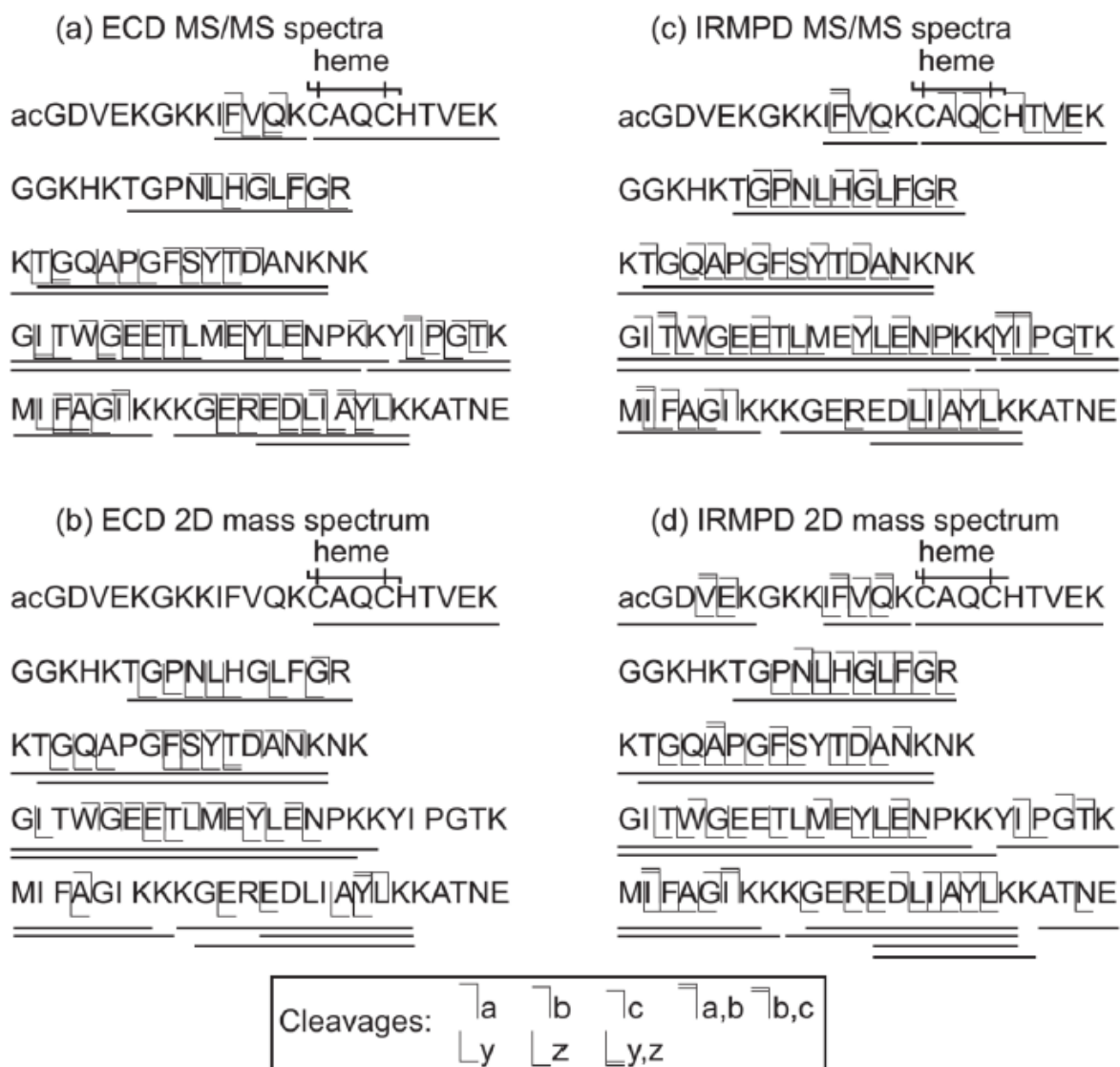
Figure 9.16: Resulting fragmentation of Cytochrome c peptides analysed by 2D-FT-ICR MS using ECD and IRMPD as fragmentation techniques. Reproduced from van Agthoven *et.al.*[31]

The ability to perform 2DMS using a variety of in-cell MS/MS techniques (such as ECD and IRMPD) allows complementary data to be acquired in a similar fashion to CAD and ETD/ECD for standard MS/MS, and the use of electron-based in-cell MS/MS provides the opportunity to study metallated biomolecules without losing fragile modifications to threshold based techniques such as CAD.

Most importantly for the study of metallodrug-biomolecule interactions is that 2DMS operates via direct infusion, using any continuous/pulsed ion source available, with the only requirement being

that the production of ions needs to be kept as consistent as possible in order to avoid scan-to-scan variation of peaks and ion count, which will produce differences between scans. The scan-to-scan variation in 2DMS is a more pronounced problem than for normal 1D MS due to the second Fourier transform, any differences between the sequential spectra will be correlated with the fragmentation of that particular precursor ion, and thus will appear as a falsely generated noise-peak. The ability to use standard nESI, and most importantly the solvent and scan conditions needed for effective metallodrug-biomolecule interaction study makes 2DMS a very attractive technique. nESI-2DMS would not suffer from the acid/solvent induced dissociation of metal modifications as observed in the LC-MS results presented above, 2DMS is also not hindered by the need to keep to a short duty cycle in the same way as LCMS, so long as there is a continuous spray of ions; 2DMS can continue to work effectively, enabling the use of high-resolution scans, long ion accumulation times, and complex/long MS/MS procedures (such as long laser/electron irradiation times).

Of particular interest was the chance to use 2D-FT-ICR MS to attempt to study the proteomic changes inside a cancer cell before and after treatment with an active anti-cancer metallodrug. As another method to "scale up" to full proteomic analysis, 2DMS was attempted on a cancer cell line of interest and an osmium based metallodrug. The anti-cancer complex codenamed FY26 (shown below), has been repeatedly shown to be extremely active (up to 49x more active than Cisplatin against 809 cancer cell lines.[32] FY26 Is also effective against ovarian cancer cells, which are currently inaccessible via photo-activated metal complexes (see PACT section of introduction). As a result FY26 (structure below in Figure 9.17) was reacted with the cancer cell line A2780, an ovarian cancer cell line studied extensively by Sadler *et.al.*[33–37] and shown to be effectively combatted using FY26. The Aim of the 2DMS analysis was to attempt to see differences in the proteome before and after treatment, and to attempt to see any osmium-modified biomolecules remaining.
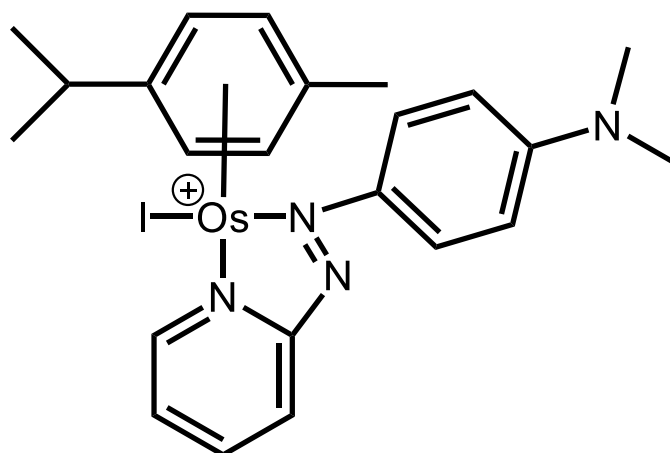
Figure 9.17: Structure of the FY26 Metallodrug used in the proteomic study of A2780 cancer cell components.

9.31: A2780 Cancer cell preparation and digestion method:

A2780 cancer cell lines were grown and treated in 2 stages, first after the desired number of cells were available (5 million for each sample), one dish were allowed to grow in media for 2 days, unhindered and provided for (by the media), to provide a control, the other was allowed to grow for one day, they were then treated with FY26 at IC_{50} concentration (100 µM)[32] the treated cells were then left for another 24 hours to allow the drug to take effect. This process is summarised in Figure 9.18, below:
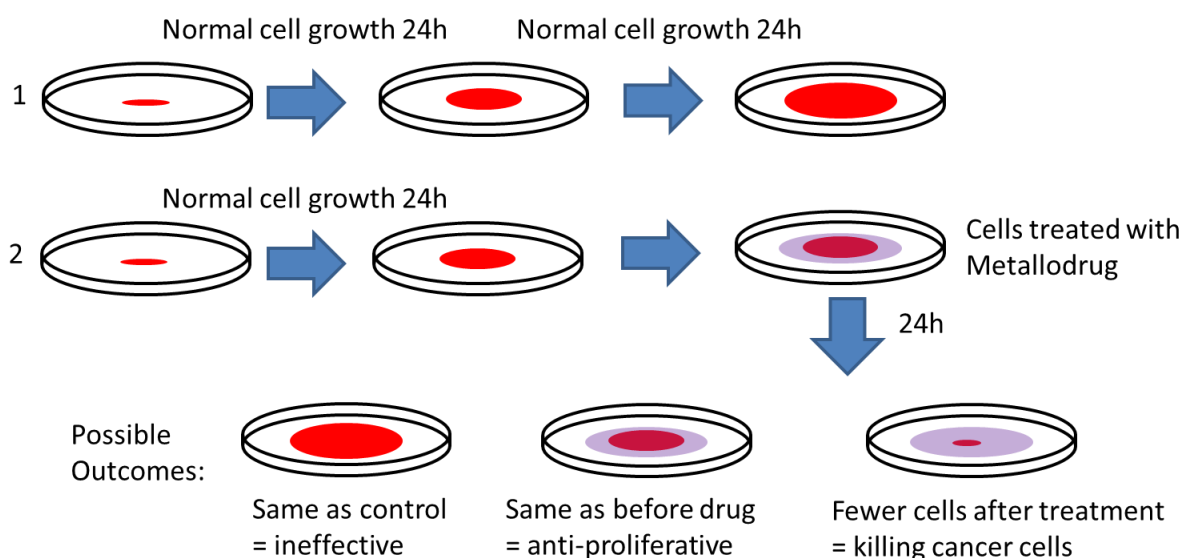
Figure 9.18: A2780 Cancer cell growth and treatment with FY26 summary, including usual outcomes from analysis via cell counting.

Cancer cells for both control and treated samples were detached from the plates using trypsin +EDTA solution, washed with Phosphate Buffer solution (PBS), then stored in PBS prior to digestion.

Cells were lysed using a variety of methods to attempt to study the differences/impact of lysis and the harsh chemicals often required on the fragile osmium-based modifications.

Lysis buffers/conditions tested used a constant "base" solution (below) and 4 different lysis formulations added to the base. Sub-milimeter glass beads (purchased from Sigma Aldrich) were sought after as a method to rupture the cell membrane without the use of harsh buffer solutions in an attempt to help retain metal based modifications. Glass beads were cleaned with ethanol before and after use twice in order to ensure no bacterial growth and to wash off residue.

Buffer base (50mL total):

- 50mM TRIS-HCl (pH 7.5 – adjusted with conc. Hydrochloric acid)

- 2.5% Glycerol (1250uL)

- 50mM DTT (0.3856g/50mL)

- 5mM EDTA (0.07306g/50mL)

Lysis formulation variants:

- Standard SDS lysis – 2% SDS added (0.0202g/mL)

- Triton X-114 0.5% + glass beads

- RLT lysis solution (2mL) + glass beads

- Only glass beads (+ same base buffer)

Cell pellets were mixed with the desired base buffer+lysis formulation (2 mL total) and sonicated in a cooled water bath (~4 °C) for 5 minutes, the samples were then placed in a 4 °C refrigerator for 2 minutes, this cycle of sonication-refrigeration was then repeated 3 times. The now lysed cell samples

consisted of insoluble membrane material (pelleted in the Eppendorf) and solubilised peptide and

protein lysate, the supernatant was extracted (~2 mL) and centrifuged at 6000 rpm for 10 minutes to

ensure insoluble material, and remaining glass beads, were effectively separated. The supernatant

was extracted again (~2 mL). SDS, RLT, and purely glass bead extracted biomolecule samples were

then stored at 4 °C, while Triton X-114 samples were shaken and re-centrifuged, followed by

extraction of the supernatant, until the solution was clear upon shaking (showing the majority of

Triton X-114 was removed, lest it would become opaque). All samples were then filtered using a

3kDa Molecular Weight Cut-Off (MWCO) filter (Amicron, Millipore) to remove buffer molecules and

salt (glycerol, DTT, EDTA, remaining lysis formulation <3kDa, etc.). The MWCO filter was filled with

450uL of solution for each filtration and centrifuged at 14,000 rpm for 25 minutes per cycle. The

remaining >3kDa peptides and proteins of interest were then extracted from the MWCO filter using

Urea (6M) to create a protein suspension which was sonicated for 5 minutes total (3 seconds

sonication, 3 seconds rest, repeat) and then centrifuged to create a protein pellet. The resulting

protein pellet was then quantified using a standard BSA calibration curve and Coomassie blue

solution. Various BSA quantities (in mg/mL) were used to create a colour above and below that of

the sample, the solutions were then analysed via a UV/Visible micro-plate reader to determine

accurate absorbances of the known BSA quantities, and the unknown sample protein quantities,

enabling reasonable quantitation of total protein content.

After quantitation extracted protein samples were adjusted to 1 µg/mL using aqueous Ammonium

BiCarbonate solution (ABC) (100 mM) and digested using a standard trypsin digestion procedure to

yield peptides. For the digestion, the protein extract (1 µg/ mL, 50ug) was mixed with DTT (50 mM in

100 mM ABC, 3 mL) for reduction of protein disulphide bonds and heating to 60 °C for 30 minutes,

IodoAcetAmide (IAA) (100 mM in 100 mM ABC, 3 mL) was then added in the dark for 1 hour to

alkylate the now free cysteine SH groups and prevent re-formation of disulphide bonds (which

inhibit effective digestion and MS/MS analysis), after the reduction and alkylation the protein extract

was digested using Trypsin solution (1 mg/mL in 100 mM ABC, 2 mL) and incubated at 37 °C

overnight for 16 hours. The resulting peptide and protein solution was then filtered through a 3kDa MWCO filter again to remove undigested protein and remaining Trypsin which could damage/block nESI tips/C18 nano-chromatography columns, all peptides under 3kDa could pass through the filter.

The filtrate was then desalted using C18 disposable Solid-Phase Extraction (SPE) cartridges. The C18 cartridges (SOLA, 1 mL capacity, 10 mg bed, Thermo scientific) were wetted with methanol (1 column volume, 1 mL), washed with wash solution (80% ACN 19.9% H_2O, 0.1% TFA) (1 column volume, 1 mL), and then conditioned (99.9% H_2O, 0.1% TFA) (2 column volumes, 2 mL). The sample (1/10th bed mass = 1 mg, 1 mg/mL) was then loaded onto the prepared C18 cartridges, the loaded sample was then washed with acidified water (99.9% H_2O, 0.1% acetic acid, 2 column volumes, 2 mL). The washed sample was then eluted from the C18 cartridge using an acidified water-organic solvent mixture (80% ACN, 19.9% H_2O, 0.1% acetic acid, 3 column volumes, 3 mL). 80% ACN was chosen as it was found to elute the target peptides in isolated protein digestions (lysozyme, BSA, etc.) without eluting remaining SDS buffer etc. (soluble in high ACN %'s). The eluted peptide sample (3mL) was then dried to solid powder via vacuum centrifugation overnight. The dried peptide powder was then reconstituted to 0.5mg/mL in water. Finally MS samples were created by diluting the stock peptide digest solution ~100 fold with water/ACN (final solvent ratio was 80% H_2O, 20% ACN).

A2780 Cancer cell digest control and treated samples were then analysed using nESI-FT-ICR MS to assess protein digest success and peptide recovery, example spectrum and zoom-in shown for the control in Figure 9.19 (below), the UHR-FT-ICR MS scans quickly showed the many thousands upon thousands of peaks able to be observed in a single FT-ICR MS spectrum and the complexity of the human cell lysate digests.
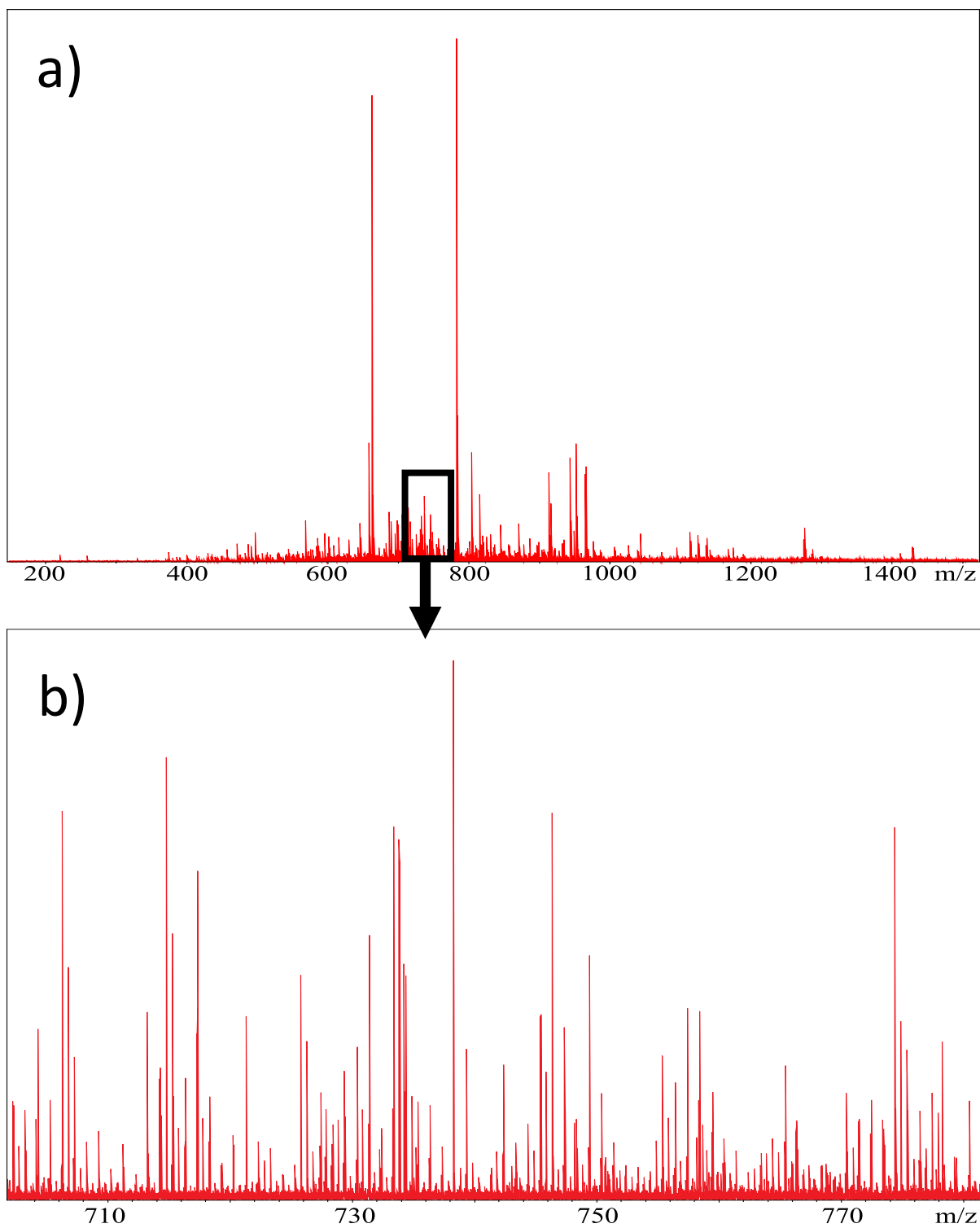
Figure 9.19: 1D-nESI-FT-ICR MS of A2780 cancer cell digest peptide mixture full mass spectrum (top, (a)) and zoom-in region (marked and in (b)) showing the high sample complexity and the ~dozens of peptide peaks per m/z.

Since the 1D-nESI-FT-ICR MS analysis was possible and showed abundant peptide recovery, subsequent 2D-FT-ICR MS analysis would be viable. Both the treated and untreated A2780 digest samples were analysed using both 2D-ECD and 2D IRMPD analysis. ECD parameters were as follows; pulse length = 0.2 s, bias = 1.2 V, lens = 18 V and IRMPD parameters; pulse length = 0.1 s, Laser power 70%, 25 W laser. Each 2D-MS was acquired using 256k data point transients and 4096 increments of t_1, creating a 2D-MS spectrum of 4096 lines with 256k data points per line, each shown below:
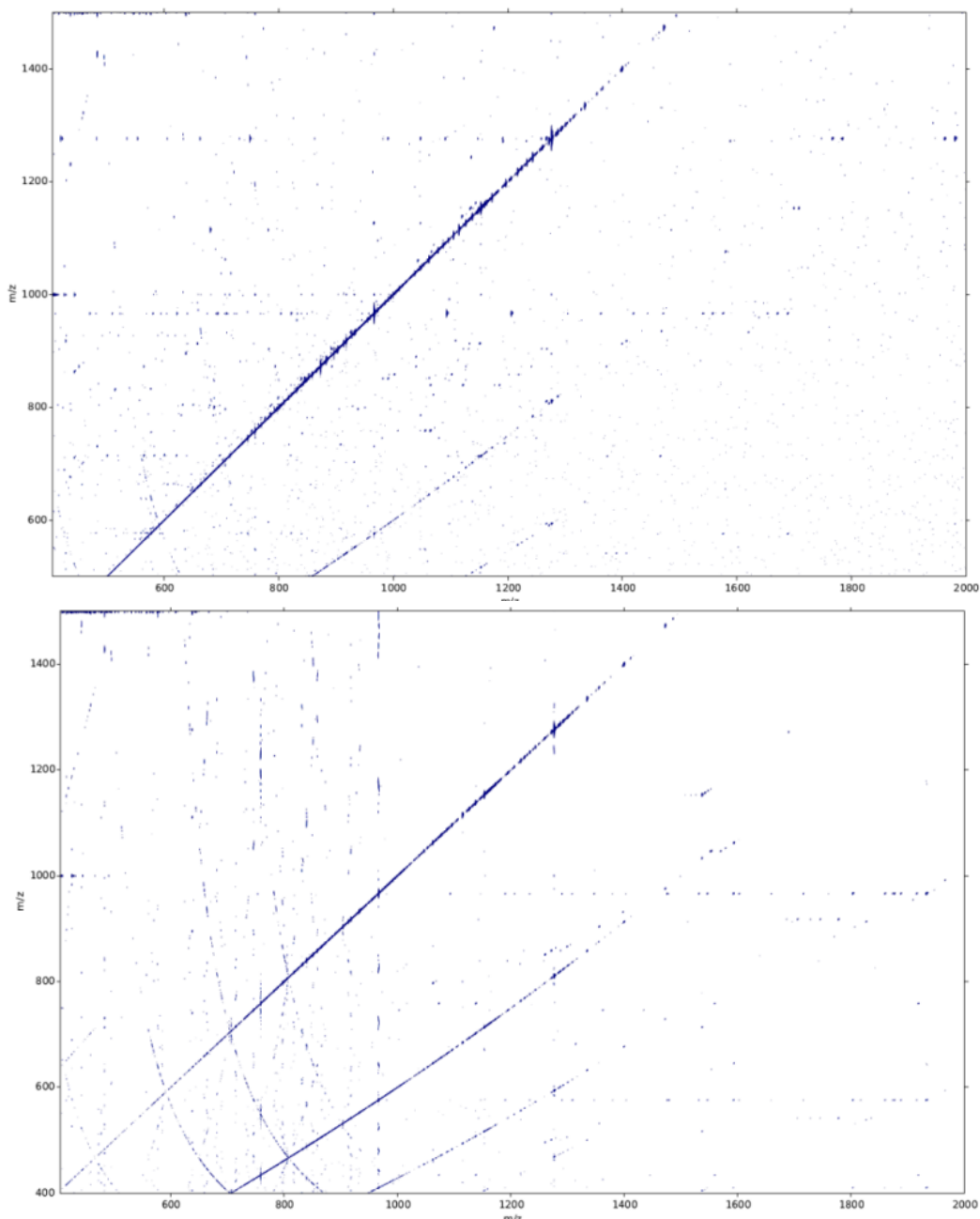
Figure 9.20: 2D-nESI-FT-ICR MS spectra of A2780 cancer cell control trypsin-digest derived peptide samples dissociated using IRMPD (top) and ECD (bottom). ECD parameters were as follows; pulse length = 0.2 s, bias = 1.2 V, lens = 18 V and IRMPD parameters; pulse length = 0.1 s, Laser power 70%, 25 W laser.

The resulting 2D-MS spectra show a huge number of parent ion and fragment ion peaks, all viewed in a single 2D-MS spectrum. The Autocorrelation line, easily visible on the diagonal is extracted for

each spectrum and presented below in Figure 9.21a-d, showing similar complexity to that in the 1D-MS above.
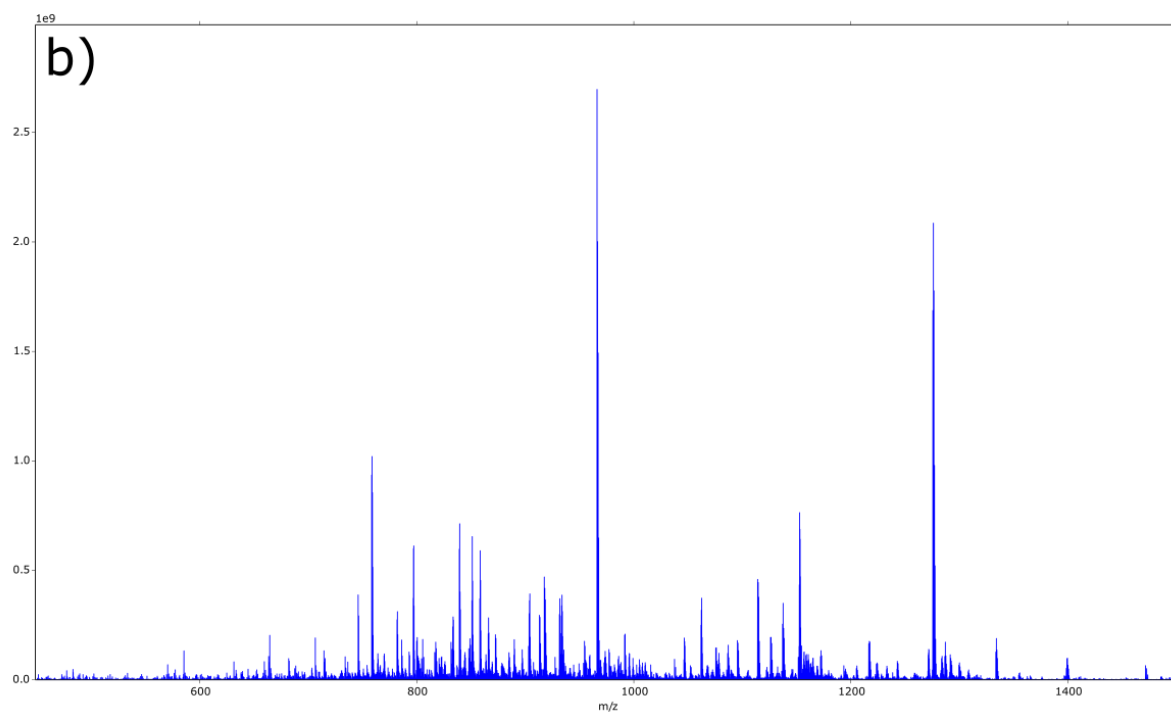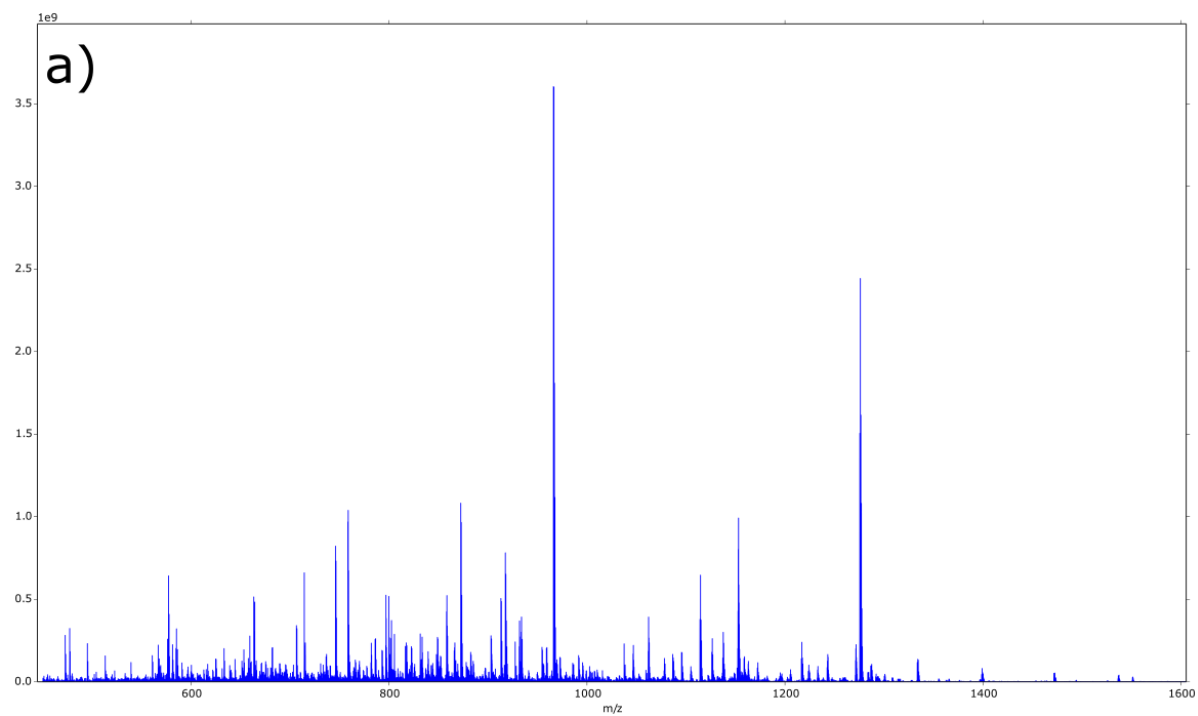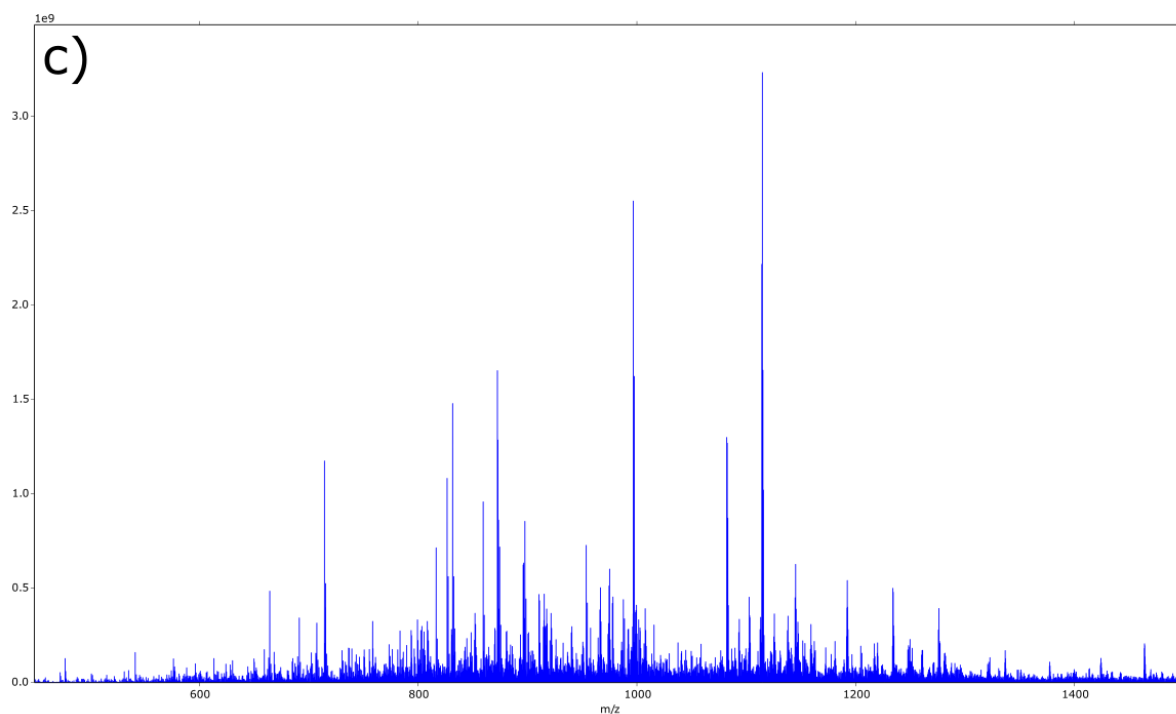
Figure 9.21: Auto-correlation lines for the A2780 cancer cell digests control sample using IRMPD (a), control sample using ECD (b), FY26-treated sample using IRMPD (c)

As mentioned previously every precursor along the Auto-correlation line will also possess its own fragment ion scan, showing the MS/MS fragmentation spectrum of that precursor by the MS/MS method chosen. Examples of extracted horizontal lines are shown below in Figure 9.22:
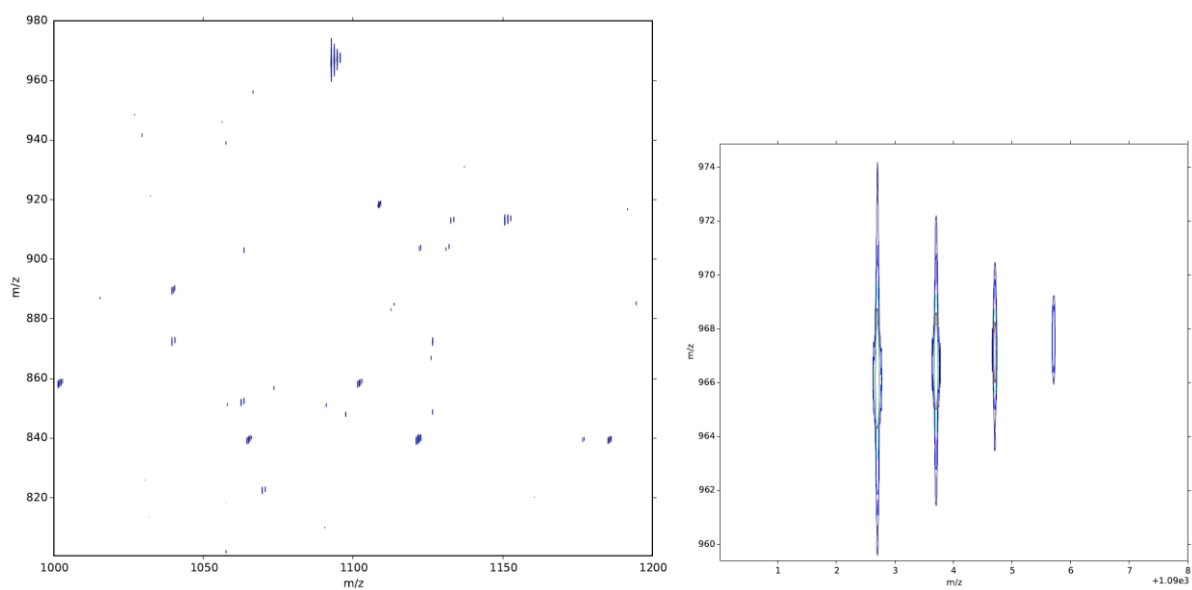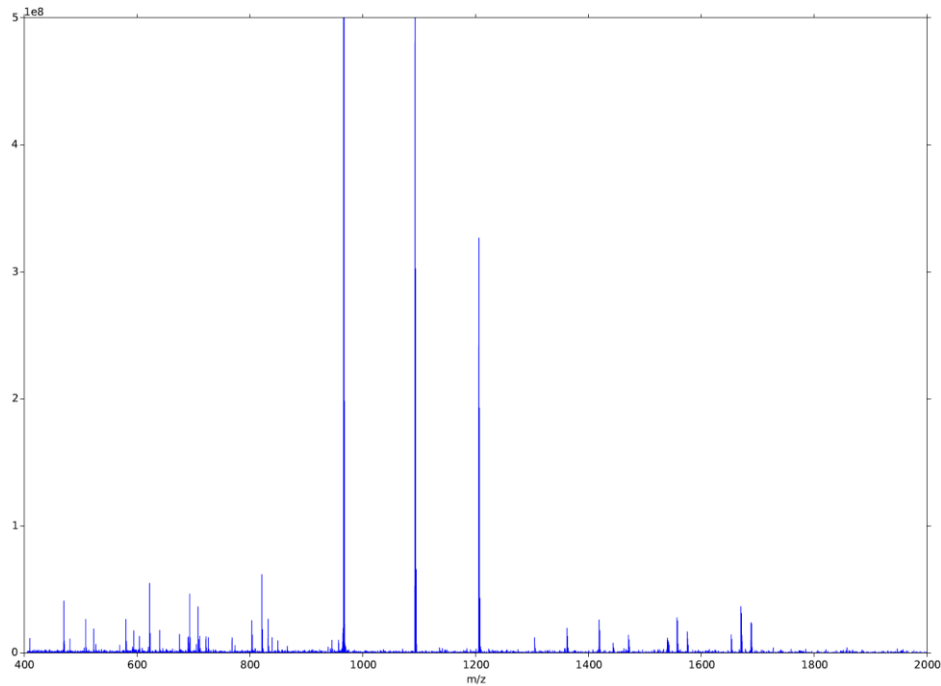
Figure 9.22: Example extracted lines from the 2D-nESI-IRMPD-FT-ICR MS analysis of A2780 control cancer cell digest proteomic sample, from top left; zoom in of small region, showing many isotopic distributions (fragments), a zoom in of an isotopic distribution featured in previous region, and finally an extracted horizontal (fragment/ MS/MS) line of a peptide precursor (966 m/z) fragmented by IRMPD MS/MS.

Neutral loss lines may also be extracted, and can inform the analysis of not only the functional groups/amino acids present in the peptide precursor, but can also be used to track certain PTM's, such as phosphorylation. IRMPD/CAD of phosphorylated peptides produces neutral losses of the phosphate group expressed at -80Da and -98Da losses from the precursor ion.[38–41] Thus neutral loss lines of Auto-correlation line -80 Da/e and -98 Da/e can be extracted from the 2D mass spectra above to instantly show all peptide species containing the phosphorylation modification and thus those peptide could be readily searched using standard database search algorithms to identify all phosphopeptides in a given sample. Phosphopeptides are particularly useful targets for analysis of proteome changes as they have been linked extensively to cell-signalling pathways and triggered

433

changes within the cell.[42,43] Examples of -80 Da and -98 Da neutral losses from the Auto-correlation

line for each charge state are shown below in Figure 9.23:
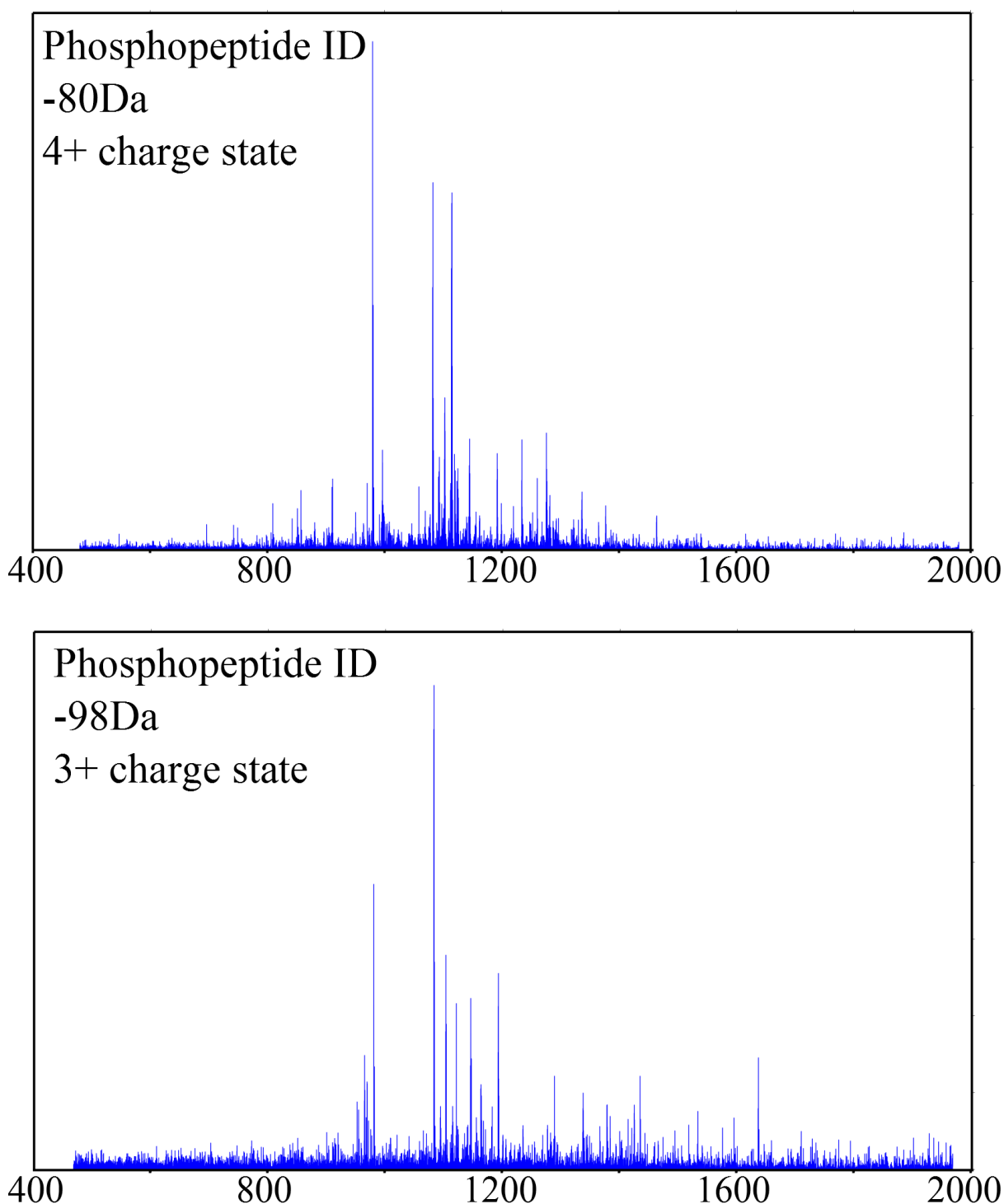
Figure 9.23: Example extracted neutral loss scans from the A2780+FY26 treated cancer cell digest

samples, showing the characteristic losses for phosphopeptides under IRMPD fragmentation (-

80Da and -98Da), the resulting peptides could then be easily identified via accurate mass and/or

MS/MS fragmentation to readily reveal the identity of all fragmented phosphopeptides in the proteomic sample, a comparison of this with the untreated cell digest could reveal differences in cell signalling due to metallodrug exposure.

Precursor ion scans revealing the precursors responsible for a given fragment ion can also be extracted if needed, this has the potential in future to be a useful marker for charged-metallodrug modifications which are easily lost in CAD/IRMPD (e.g. the Os1 and metal-bound iridium drugs shown in Chapters 4 and 6 respectively). Metallodrugs which dissociate and produce a detectable peak of their own could be extracted by a precursor ion scan, instantly identifying all species modified by the metal complex in that fashion, in a similar way to the phosphorylated peptide method discussed above. ECD-induced side chain losses could also be used to effectively identify modified amino acids (e.g. oxidised methionine produced by FM190) and to identify different metallodrug modifications showing unique side-chain/bound ligand losses under ECD MS/MS, as shown for FM190-induced modifications, and for Ir1, Ir2, Ir3 ECD MS/MS, and for Os1-DNA EDD MS/MS in previous chapters. The unique, but usually challenging metal-induced side chain losses/unique fragmentation pathways under ECD/EDD/ExD fragmentation may finally be exploited as characteristic identifiers for metal-modified species using 2D-MS.

Effective mathematical peak-picking, calibration, and automated interpretation of 2D-MS data is underway, with the goal of automatically picking and assigning peptide/modified peptide species via standard proteomic tools (such as MASCOT searches etc.), in a similar way to LC-MS data. Advancements will enable interpretation 2D-MS proteomic data and characterisation of the biomolecule species observed, these will be applied to the results presented herein at the earliest opportunity. The possibility exists for future advancement in 2D-peak picking to apply a similar isotope searching method as used in the SNAP-LC MS methodology (above), to search 2D-MS data and be able to detect metal-bound/target-bound species and reveal metal-interactions and changes inside cancer cells both for natural and administered metals/target elements.

References:

(1)   Lopez-Clavijo, A. F.; Barrow, M. P.; Rabbani, N.; Thornalley, P. J.; O'Connor, P. B. *Anal. Chem.* **2012**, *84* (24), 10568.

(2)   Creese, A. J.; Cooper, H. J. *Am. Soc. Mass Spectrom.* **2008**, *19*, 1263.

(3)   Chalmers, M. J.; Håkansson, K.; Johnson, R.; Smith, R.; Shen, J.; Emmett, M. R.; Marshall, A. G. *Proteomics* **2004**, *4* (4), 970.

(4)   Park, S.; Ntai, I.; Thomas, P.; Konishcheva, E.; Kelleher, N. L.; Statsuk, A. V. *Biochemistry* **2012**, *51* (42), 8327.

(5)   Chu, F.; Nusinow, D. a; Chalkley, R. J.; Plath, K.; Panning, B.; Burlingame, A. L. *Mol. Cell. Proteomics* **2006**, *5* (1), 194.

(6)   Cannon, J.; Nakasone, M.; Fushman, D.; Fenselau, C. *Anal. Chem.* **2012**, *84* (22), 10121.

(7)   Olsen, J. V; Mann, M.; Shevchenko, A.; Tomas, H.; Havlis, J. **2007**, *1* (6), 2856.

(8)   Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. *Biochem. Biophys. Res. Commun.* **2014**, *445* (4), 683.

(9)   Pasa-Tolić, L.; Masselon, C.; Barry, R. C.; Shen, Y.; Smith, R. D. *Biotechniques* **2004**, *37* (4), 621.

(10)   Li, L.; Masselon, C. D.; Anderson, G. a; Pasa-Tolić, L.; Lee, S. W.; Shen, Y.; Zhao, R.; Lipton, M. S.; Conrads, T. P.; Tolić, N.; Smith, R. D. *Anal. Chem.* **2001**, *73* (14), 3312.

(11)   Shen, Y.; Zhao, R.; Berger, S. J.; Anderson, G. a; Rodriguez, N.; Smith, R. D. *Anal. Chem.* **2002**, *74* (16), 4235.

(12)   Synowsky, S. a; van den Heuvel, R. H. H.; Mohammed, S.; Pijnappel, P. W. W. M.; Heck, A. J. R. *Mol. Cell. Proteomics* **2006**, *5* (9), 1581.

(13)   Xie, F.; Smith, R. D.; Shen, Y. *J. Chromatogr. A* **2012**, *1261*, 78.

(14)   Ahlf, D. R.; Compton, P. D.; Tran, J. C.; Early, B. P.; Thomas, P. M.; Kelleher, N. L. *J. Proteome Res.* **2012**, *11* (8), 4308.

(15)   Lee, S.-W.; Berger, S. J.; Martinović, S.; Pasa-Tolić, L.; Anderson, G. a; Shen, Y.; Zhao, R.; Smith, R. D. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (9), 5942.

(16)   Kaur, P.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (3), 459.

(17)   Koster, C.; Holle, A. In *American Society of Mass Spectrometry Annual conference*; ASMS, 1999; pp 1–2.

(18)   Marshall, A. G. *Int. J. Mass Spectrom.* **2000**, *200* (1-3), 331.

(19)   Murray, K. K.; Boyd, R. K.; Eberlin, M. N.; Langley, G. J.; Li, L.; Naito, Y. **2013**, No. June.

(20)   Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Craig, M.; Meng, C. K. A. I.; Mann, M.; Whitehouse, C. M. **2012**, *246* (4926), 64.

(21)   Hartmer, R.; Stoermer, C.; Main, L.; Wunderlich, D.; Ingendoh, A.; Albers, C.; Hebler, R.; Harder, A.; Schmit, P. *Bruker Tech. notes* **2010**, *#40* (#270354), 1.

(22)   Shen, Y.; Tolić, N.; Masselon, C.; Pasa-Tolić, L.; Camp, D. G.; Lipton, M. S.; Anderson, G. a; Smith, R. D. *Anal. Bioanal. Chem.* **2004**, *378* (4), 1037.

(23)   Tian, Z.; Zhao, R.; Tolić, N.; Moore, R. J.; Stenoien, D. L.; Robinson, E. W.; Smith, R. D.; Paša-Tolić, L. *Proteomics* **2010**, *10* (20), 3610.

(24)   Smith, R. D.; Tang, K.; Shen, Y. *Mol. Biosyst.* **2006**, *2* (5), 221.

(25)   Pfandler, P.; Bodenhausen, G.; Rapin, J.; Houriet, R.; Gaumann, T. *Chem. Phys. Lett.* **1987**, *138* (2,3), 195.

(26)   Ross, W.; Guan, S.; Peter, B.; Tom, L.; Mushll, S. **1993**, No. 22.

(27)   Agthoven, M. A. Van; Chiron, L.; Coutouly, M.; Delsuc, M.; Rolando, C.; Strasbourg, U. De.

2012.

(28)   Bensimon, M.; Zhao, G.; Gaumann, T. *Chem. Phys. Lett.* **1989**, *157* (1,2), 97.

(29)   van Agthoven, M. A.; Barrow, M. P.; Chiron, L.; Coutouly, M.-A.; Kilgour, D.; Wootton, C. A.; Wei, J.; Soulby, A.; Delsuc, M.-A.; Rolando, C.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (12), 2105.

(30)   Van Agthoven, M. A.; Chiron, L.; Coutouly, M. A.; Sehgal, A. A.; Pelupessy, P.; Delsuc, M. A.; Rolando, C. *Int. J. Mass Spectrom.* **2014**, *370*, 114.

(31)   Agthoven, M. A. Van; Wootton, C. A.; Chiron, L.; Coutouly, M.; Soulby, A.; Wei, J.; Barrow, M. P.; Delsuc, M.; Rolando, C.; Connor, P. B. O. *Anal. Chem.* **2016**, *Under revi*.

(32)   Hearn, J. M.; Romero-Canelón, I.; Munro, A. F.; Fu, Y.; Pizarro, A. M.; Garnett, M. J.; McDermott, U.; Carragher, N. O.; Sadler, P. J. *Proc. Natl. Acad. Sci.* **2015**, *112* (29), E3800.

(33)   Farrer, N. J.; Woods, J. a; Munk, V. P.; Mackay, F. S.; Sadler, P. J. *Chem. Res. Toxicol.* **2010**, *23* (2), 413.

(34)   Barry, N. P. E.; Edafe, F.; Therrien, B. *Dalton Trans.* **2011**, *40* (27), 7172.

(35)   Peacock, A. F. A.; Habtemariam, A.; Ferna, R.; Walland, V.; Fabbiani, F. P. A.; Parsons, S.; Aird, R. E.; Jodrell, D. I.; Sadler, P. J. **2006**, *128*, 1739.

(36)   Millett, A. J.; Habtemariam, A.; Romero-Canelón, I.; Clarkson, G. J.; Sadler, P. J. *Organometallics* **2015**, *34* (11), 2683.

(37)   Butler, J. S.; Sadler, P. J. *Curr. Opin. Chem. Biol.* **2013**, *17* (2), 175.

(38)   Wu, S.; Yang, F.; Zhao, R.; Tolić, N.; Robinson, E. W.; Camp, D. G.; Smith, R. D.; Pasa-Tolić, L. *Anal. Chem.* **2009**, *81* (11), 4210.

(39)   Larraillet, V.; Antoine, R.; Dugourd, P.; Lemoine, J. *Anal. Chem.* **2009**, *81* (20), 8410.

(40)   McLafferty, F. W. *Int. J. Mass Spectrom.* **2001**, *212* (1-3), 81.

(41)   Ryan, C. M.; Souda, P.; Bassilian, S.; Ujwal, R.; Zhang, J.; Abramson, J.; Ping, P.; Durazo, A.;

        Bowie, J. U.; Hasan, S. S.; Baniulis, D.; Cramer, W. a; Faull, K. F.; Whitelegge, J. P. *Mol. Cell.*

        Proteomics **2010**, *9* (5), 791.

(42)   Schmidt, A.; Gehlenborg, N.; Bodenmiller, B.; Mueller, L. N.; Campbell, D.; Mueller, M.;

        Aebersold, R.; Domon, B. *Mol. Cell. Proteomics* **2008**, *7* (11), 2138.

(43)   Soulby, A. J.; Heal, J. W.; Barrow, M. P.; Roemer, R. a; O'Connor, P. B. *Protein Sci.* **2015**, *24*

        (5), 850.

Chapter 10:

Conclusions

Over the course of this thesis work, many different systems have been tackled and studied using a range of Mass Spectrometry-based techniques, all with the overriding goal of being able to study metallodrugs and their interactions with biomolecules. This chapter introduces, summarises, and outlines the study, advancements, and conclusions of each of the results chapters discussed previously and what has been gained through these studies for the wider goal and use of this work.

Chapter 3: Multi-targeted Photoactivatable Platinum Anticancer Complexes
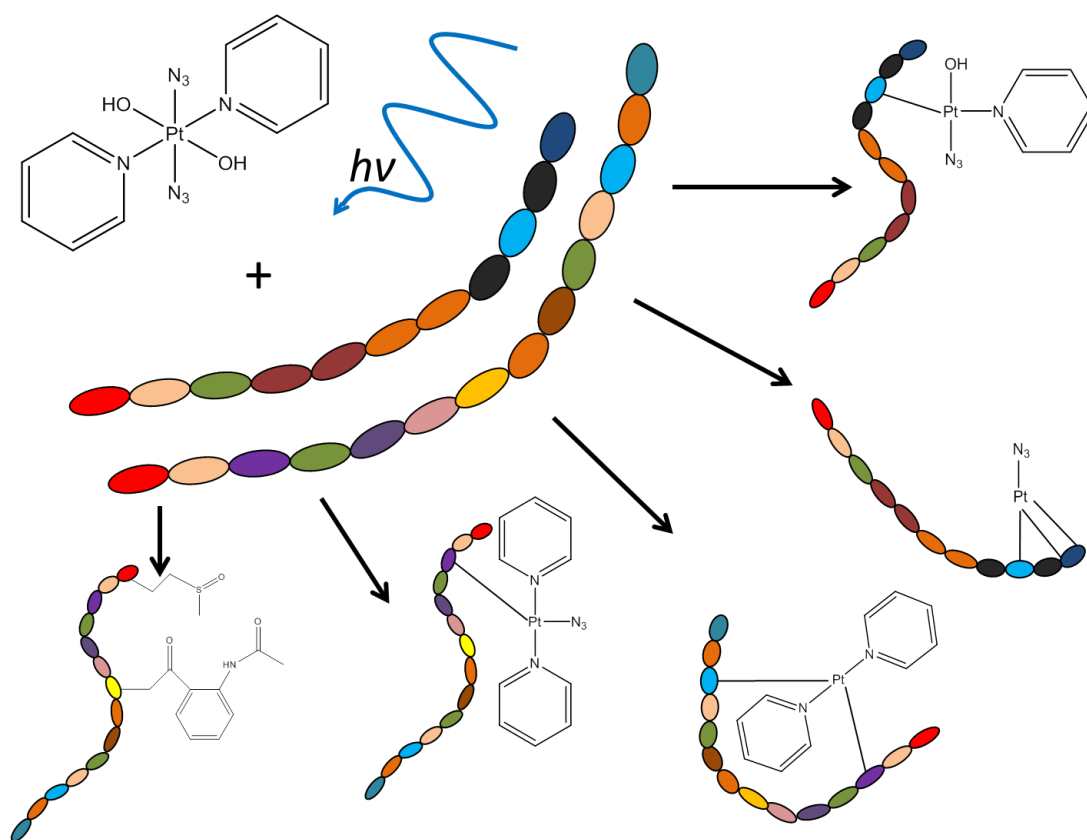
The first of the results chapters focused on the study of a novel photoactivatable platinum(IV) complex (codenamed FM190) and its reaction products under activation with blue visible light (463nm) while in the presence of model peptide biomolecules (Substance P and [Lys]³-Bombesin). The reactions caused an unexpectedly large number of reaction products, modifying the peptide species with both platinum metal complex modifications and oxidation products. The Pt(II) modifications were shown to be influenced by the sequence of the biomolecule present during the photo activation process, with [Lys]³-Bombesin+FM190 solutions producing more expected bi-pyridine functionalised modifications, but a major proportion also unexpectedly retaining one azido group functionality, despite the belief that both azido groups were released during Pt(IV) complex photo activation. Reactions of Substance P and FM190 with blue visible light produced an array of unexpected reaction products, many of which not retaining the two pyridine ligands on the Pt(II) centre, and many retaining azido functionalities, showing that the biomolecule binding to the Pt centre once activated/during activation has an impact on the photo activation process and the stable reaction products formed.

High levels of biomolecule oxidation were also induced by activation of the Pt(IV) compounds, representing a metal-independent mechanism of attacking target biomolecules/cells. This oxidation was then located using MS/MS techniques within the FT-ICR MS, and was found to be targeting the methionine and tryptophan residues within the peptides. Methionine oxidation is routine and expected with many oxidation mechanisms. However, it was found that tryptophan oxidation in particular could be utilised as an indicator for the mechanism of oxidation. Different reaction products can be formed from the oxidation of tryptophan residues within poly(amino acids) and the products formed inform us of the species causing the oxidation – either singlet oxygen or hydroxide radicals. It was found that due to the products formed that the biomolecule oxidation was being formed by hydroxide radicals, which could be released during photo activation. This extensive, yet previously unobserved, oxidation (up to 3 oxidations in the 11/14 amino acid peptides) was not only an interesting additional mechanism of action against cancerous cell components, but could also be

a way of circumventing platinum resistance in cancerous cells, and a metal-independent way of increasing ROS and inducing apoptosis/cell death when and where desired using photo activation. Surprisingly, release of the azido ligands during photoactivation did not result in any azide-modified biomolecules, despite the expectation that such reactive species could easily modify various functional groups/areas within the peptides.

MS/MS analysis of the platinum-containing species proved challenging, with more common CAD and IRMPD MS/MS techniques causing dissociation of the platinum complex modification. Unfortunately unlike the fragmentation these MS/MS techniques cause to fragile PTM's such as phosphorylation, slow-heating MS/MS of the platinum modifications caused partial damage to the complex modifications, meaning dissociation of bound ligands, but not usually the Pt metal centre. The dissociation of ligands caused the creation of an extremely reactive platinum centre which was then believed to bind to any viable functional group along the peptide structure rapidly, causing cyclic and complex species, which were subsequently dissociated during the ergodic dissociation mechanism, causing complex sequence-scrambled MS/MS data and no sequence informative fragments. Thus it was quickly obvious electron-based MS/MS dissociation techniques were required. ECD was employed to fragment and characterise every available reaction product in the two reaction mixtures. ECD MS/MS showed that in the [Lys][3]- Bombesin the metal complexes preferred to bind to the available histidine[11] residue regardless of the structure/ligand combination of the Pt(II) modification. When FM190 was activated in aqueous solutions of Substance P, no histidine amino acid residue was present and the Pt(II) modifications were shown to preferentially bind to the lysine[3] amino acid residue, despite the availability of a terminal methionine residue, which is in stark contrast to all previous platinum-based metallodrugs studied, again representing key differences between traditional Pt(II) solution activation and the Pt(IV) photo activation mechanisms of action against biomolecules.

FM190 activity against cancerous cells was previously shown to be drastically affected by addition of free radical trapping agents (free tryptophan) during photoactivation. This control over the photoactivation process was believed to be utilisable as a "switch" for activity. The addition of free Trp during photoactivation of FM190 in the presence of the model biomolecules was also studied. The radical trapping agents had a drastic subduing effect on the abundances of products formed (up to 98% reduction in some species), with some Pt modified species affected more than others, and the oxidation products affected substantially less than the platinum-modified species. Indeed Trp could be used to affect the products formed, but only certain species were subdued, which could be correlated with the reduced biological activity and highlight particularly critical/active modification species key to killing cancerous cells.

Platinum modifications bound to the biomolecules studied were shown to have drastic and unique effects on ECD fragmentation mechanisms, which for the most part complicated and hindered analysis of subsequent ECD MS/MS spectra and sequencing efforts. However after elucidation of certain fragmentation mechanisms it was also apparent that the unique losses observed during ECD MS/MS were driven from electrons captured/transferred to the platinum metal centres, causing ligand dissociation, and so could be used to identify platinum bound ligands and provide modification information. In addition the platinum modifications caused previously unobserved metal-enhanced ECD amino acid side chain losses, which provided information of amino acid groups within close proximity to the modifications present.

Chapter 4: Organo-Osmium complex-DNA interactions by FT-ICR MS/MS
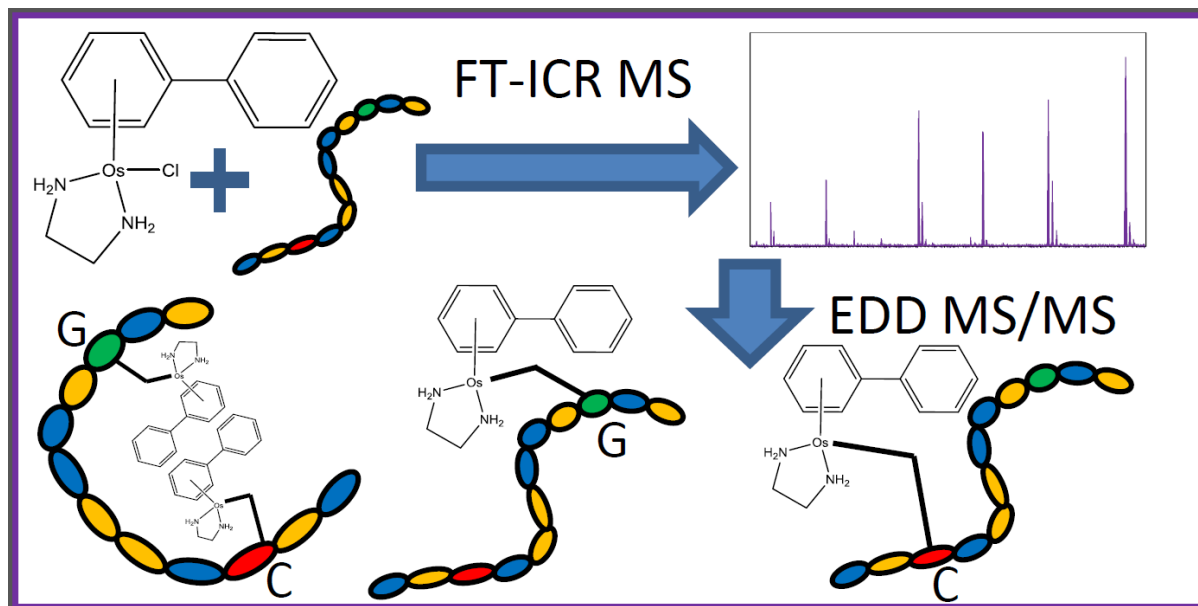
It has been shown that upon administration of the most common anticancer metallodrugs over 90% of injected compound becomes protein bound within just 24 hours in vivo. The majority of this peptide/protein metallation is believed to be deactivating and preventing the compounds reaching their target; cancerous DNA. Should the metallodrug be able to reach DNA it will bind and cause deformations in tertiary structure, preventing replication and/or inducing apoptosis/cell death. These metallodrug-DNA interactions have been well studied and characterised for traditional Pt(II) compounds such as Cisplatin, Transplatin, and Oxaliplatin. Platinum(II) compounds are now well known to bind preferentially to guanine residues, and can form inter-chain bonds via two guanine nucleotides, and even intra-strand crosslinks in a similar fashion. However novel metallodrugs such as those based on different transition metals to Pt(II) and with different geometries than the square-planar Cisplatin analogues have been shown to offer impressive potencies even in the face of cisplatin resistant cell lines, these drugs are believe to use different mechanisms of action against cancerous cell components, but these mechanisms are not fully understood, limiting drug development.

Although great proportions of many drugs react with peptides/proteins before reaching DNA, study of metallodrug-biomolecule interactions was extended to interactions with DNA (oligonucleotides) in order to study how these novel drugs will interact with their targets, if they should survive the journey. A novel osmium-based piano stool compound (Os1) was reacted with an extensively purified and isolated DNA 12-mer. The resulting reaction products were analysed by nESI-FT-ICR MS in the negative mode. Osmiated DNA was immediately assignable due to the osmium metal impact on the DNA isotopic distribution. It was found that the Os1 drug could bind to the DNA species two times producing the mono and di-osmium drug bound species. Both modified and unmodified DNA 12-mer species were then isolated in the FT-ICR MS instrument and fragmented using a variety of MS/MS techniques, namely CAD, IRMPD, and EDD. Unlike poly(amino acids) DNA can acquire extremely high charge states (per monomer unit) and can thus be very fragile when within the mass spectrometer. This fragility in the gas phase, though in striking contrast to DNA stability in the body/solution, creates challenging situations for effective transmission and analysis. DNA ions also fragment readily using collision based dissociations such as CAD MS/MS, producing a wide array of fragments (47 assigned fragments for the 12-mer) and excellent sequence coverage (100% of available backbone cleavage locations). IRMPD produces similarly informative spectra. However these slow heating MS/MS techniques also caused extensive loss of the osmium-based modifications during MS/MS of the osmiated DNA species. Loss of the modification not only hinders location of the modification using the few remaining osmium-containing fragments, but also causes conflicting fragment productions, some with the remaining modification, some without. Electron detachment dissociation was then used to assist in the study. EDD requires very fine tuning of many instrument parameters, both relating to the electrons produced, and the ICR cell containing the target ions, but once carefully optimised can provide extremely informative data. EDD MS/MS enabled the effective fragmentation of the osmiated DNA species, without the loss of osmium modification followed by secondary fragmentation which hindered previous CAD and IRMPD MS/MS analysis. EDD also allowed the elucidation of two separate species within the same MS/MS spectrum; one a DNA 12-

mer with the osmium metallodrug bound at the guanine, and the other with the Osmium drug

bound at the cytosine, a previously unobserved target for osmium compounds. EDD analysis of the

di-osmiated DNA species proved ineffective, producing MS/MS spectra characteristic of electron-

based MS/MS of cyclic species. This cyclic-biomolecule behaviour was believed to be due to the two

osmium metallodrug bound at opposing ends of the biomolecule, but interacting in a pi-pi stacking

arrangement via the biphenyl rings attached to each drug, the conclusion was supported by H-NMR.

CAD MS/MS was effective at breaking these pi-pi interactions first, and then breaking apart the DNA

backbone subsequently. Although CAD MS/MS of the di-osmiated DNA suffered similar problem as

in the mono-osmiated species, it was able to provide enough osmium location information to

support the previous findings that the osmium modifications were binding to the guanine and

cytosine sites simultaneously. Thus it was obvious that a multi-platform MS/MS approach was

needed to study these metallated DNA species in the gas phase, but was possible using FT-ICR MS.

FT-ICR MS was also able to effectively resolve the complex MS/Ms fragmentation patterns produced

during DNA fragmentation, with mass differences between certain assignments as little as <8mDa.

Chapter 5: Further DNA-metallodrugs investigations
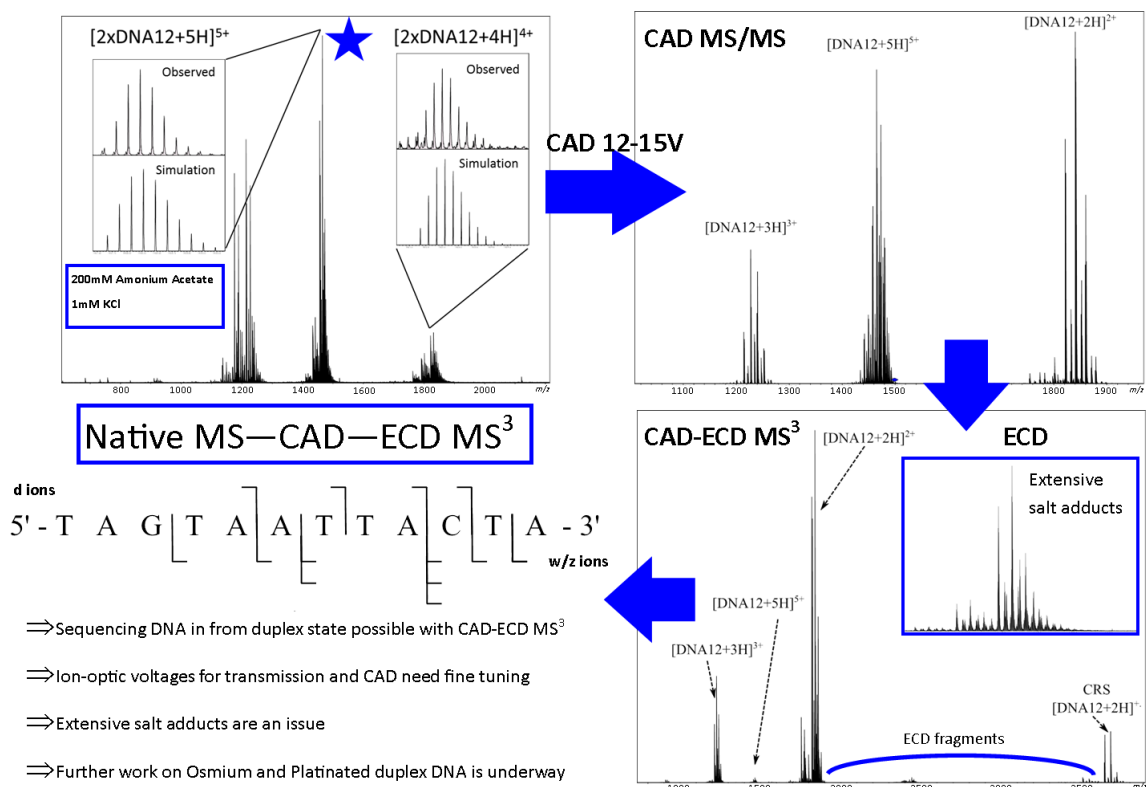
Following the success of the osmium-DNA study using CAD, IRMPD, and EDD MS/MS a study of DNA interacting with the platinum compound FM190 (from Chapter 3) was undertaken. The photoactivatable platinum compound was activated using blue visible light (463 nm) while in the presence of the DNA 12-mer used in Chapter 4 (TAGTAATTACTA). The FM190 compound was shown to bind to the DNA species extremely quickly, even more quickly than with the peptides. In less than 30 minutes over 95% of the DNA had been modified with platinum species, producing a range of modifications similar to those observed in the FM190-peptide reactions. The species observed all retained the bi-pyridine functionalities as expected, but varied in the remaining ligands. Each species was then isolated and fragmented using CAD, IMRPD, and EDD MS/MS, in a similar fashion to the osmiated DNA. The FM190 compound was shown to bind to guanine nucleotide bases in a mono-dentate fashion. While those Pt-based modifications studied with two biomolecule binding points were believed to be bound at the Guanine nucleobase and the Cytosine/Adenine bases. Though CAD

MS/MS caused partial dissociation of the platinum-based modifications, in a similar fashion to both chapters 3 and 4, EDD MS/MS was also found to be hindered/complicated by the platinum metal-centre bound to the oligonucleotides.

The vast increase in reactivity for FM190 to bind to DNA compared to peptides and proteins was in-line with previous data and the ideal mechanism of action, with the platinum drug attacking cancerous DNA preferentially over peptide and protein biomolecules. However the expected and famous platinum-induced inter-strand DNA crosslinks were conspicuous by their absence. The lack of crosslinked species was also accompanied by lack of biomolecule oxidation, which was rife in the FM190-peptide reactions observed in Chapter 3. The mechanism of activation/biomolecule binding with DNA species was clearly very different to that of the platinum compound when activated in the presence of peptides.

Although the analysis and characterisation of metallodrug-DNA interactions using isolated DNA strands was effective and provided unique insights into the mechanisms of action for both the osmium and photoactivatable platinum compounds studied herein, it was the reaction with single stranded DNA. DNA usually is present within the nucleus of cells in a duplex helical form, with two complementary strands coiled around each other via hydrogen bonds between nucleotide base groups, thus the reactive bases are often shielded to some degree from nearby metallodrugs, the notable exception to which is the N^7 of guanine. In order to study the effects of metallodrugs on representative DNA structures and provide insights into mechanisms which were similar to in vivo activity the possibility of studying duplex DNA by nESI-FT-ICR MS was investigated. This "native DNA" experiment was unfortunately not as simple as native MS analysis of many proteins, where the use of non-ideal ESI buffers (with respect to ionisation/charging efficiency) can stabilise protein structure enough to provide a representative structure in the resulting protein ion to determine structural features via subsequent MS/MS analysis. Usually the use of high-concentration ammonium acetate (e.g.10-200 mM) buffer, low source temperatures and carefully tuned instrument optics can allow

protein structure to survive ionisation and transmission/detection during ESI/nESI-MS analysis. However these tactics were ineffective at producing duplex DNA ions. Work by Gabellica *et al.* had shown that stabilisation of DNA-G-quadruplexes was possible through the use of high-concentration aqueous potassium solutions (1-2mM) to produce ions charged purely by potassium during positive mode ESI, mainly $[DNA+2K]^{2+}$ species. This approach was applied to the ionisation and study of the DNA 12-mer. Using nESI buffers containing both 200 mM ammonium acetate and 1 mM KCl salt, it was shown that the production and observation of DNA duplex ions was possible by nESI-FT-ICR MS. These DNA duplexes were present at high intensity, but also with many, many salt adduct peaks due to the added KCl. Despite the spectral complexity caused by many adduct distributions, duplex-DNA species were isolated using the quadrupole in the hybrid FT-ICR MS and then fragmented using CAD to break apart the duplex strands selectively. Low-energy CAD dissociated the duplex without causing extensive backbone fragmentation. MS/MS analysis was then performed on the resulting single-stranded DNA species using ECD MS/MS, providing MS^3 analysis for duplex DNA characterisation. MS^3 via CAD and ECD was shown to provide sequence informative fragmentation of the single stranded DNA from the duplex architecture. Thus this methodology would be a viable technique for studying metallodrug-DNA interactions in the native/duplex state and provide a more accurate reflection of reactions in vivo. Additional experiments using magnesium ions should be able to assist in such studies in future investigations as they are the ions utilised by cells to stabilise DNA tertiary structure, work to this end in ongoing.

Chapter 6: Studying electron-quenching iridium compounds
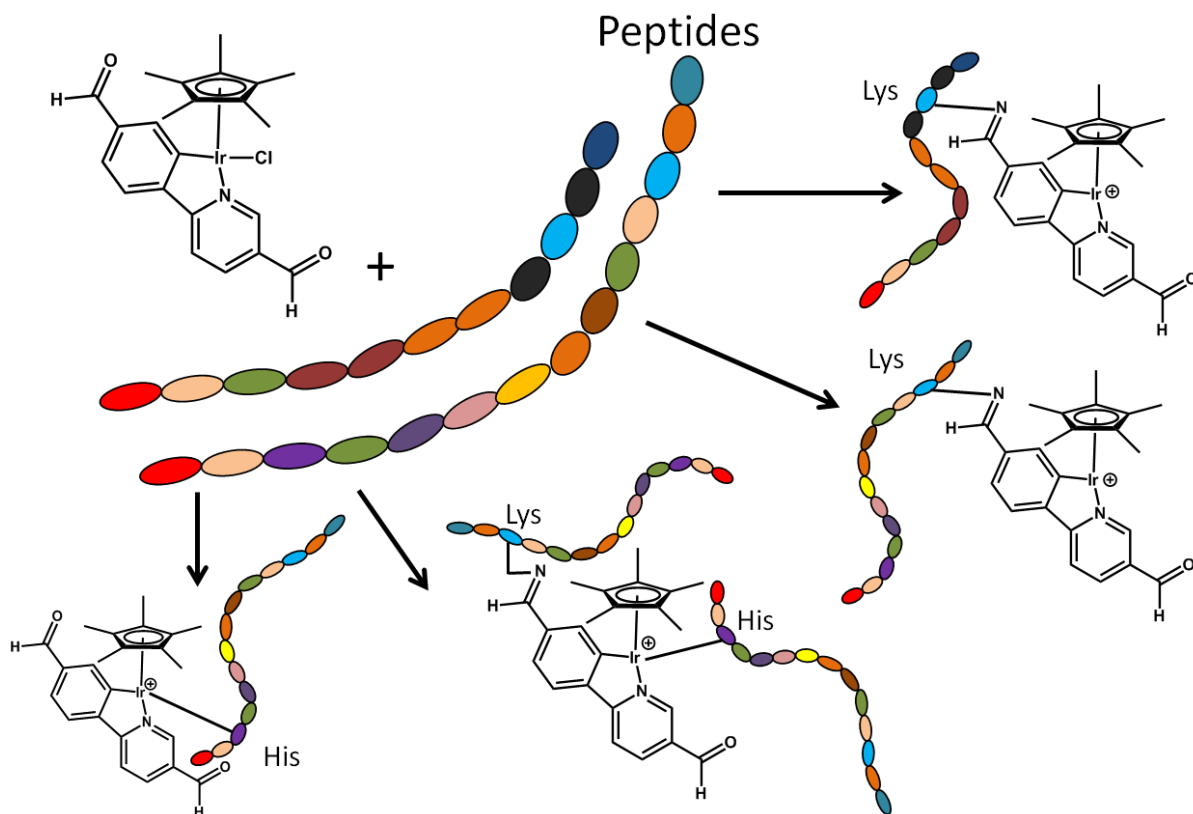
Metallodrug-biomolecule interactions are predominantly based around reaction of biomolecule functional groups (such as methionine sulphur/histidine nitrogen lone pairs) binding directly to available metal centres in a dative covalent bond fashion. Such is the case for both traditional Pt(II) compounds, photoactivatable Pt(IV) compounds such as FM190 (Chapter 3) and the osmium metallodrug (Os1) studied in Chapter 5. The reactivity and preferential binding site for the metal centre can be tuned by variation of the ligands used and their configuration/orientation, but for the majority of cases it is the metal centre where the reaction/binding takes place.

However the metal-bound ligands also present an opportunity for functionalisation and can be modified to include reactive groups/sites of their own. Functionalised small molecules have been shown to be able to target specific residues in peptides and proteins, and to bind preferentially to specific sites in DNA strands. Thus functionalising metallodrug ligands could be a way of directing

these compounds to specific targets, and so enhancing their selectivity and enabling binding to specific target sites.

A series of aldehyde-functionalised ligands were used to create a range of Iridium(III) based piano-stool metallodrug compounds with either 0, 1 (2 isomers), or 2 aldehyde functional groups. These aldehyde functionalities present the possibility for condensation reactions with primary amine groups available from peptides/proteins, and so can form imine linkages between the metallodrug and target biomolecules.

Two model peptides were chosen for the initial study; Substance P and [Lys]^3-Bombesin. Indeed upon incubation with any of the 3 aldehyde-functionalised iridium metallodrugs, a condensation reaction occurred between the aldehyde functional group and an amino acid residue. Reactions with substance P followed by nESI-FT-ICR MS analysis showed near complete reaction within just 24 hour incubation at 37°C. Reactions of the Ir compounds with [Lys]^3-Bombesin produced both a condensation product and a non-condensation product (metal complex bound to biomolecule without water loss and imine formation). Subsequent MS/MS analysis of the iridiated peptides proved unusually challenging, even for metallodrug modified biomolecules. ECD MS/MS, usually being the first choice for such analysis, was shown to be drastically hindered due to quenching/stabilisation of captured electron by the covalently bound iridium modification in the condensation product, producing charge reduced species (parent ion after subsequent electron capture) of extremely high intensities compared to what is observed in the unmodified peptide ECD MS/MS; 150-250% relative intensity to parent ion against the usual 10-20%. This stabilisation or quenching of ECD electrons hindered effective ECD MS/MS and subsequent analysis greatly, but was shown to be possible using very carefully tuned ECD and ICR cell parameters, the slightest deviation from which would cause large drops in already low (<1%) fragment intensity and/or loss of certain fragments, limiting sequence/cleavage coverage. Despite these challenges the modified biomolecules were successfully fragmented using ECD MS/MS and the imine-bound iridium

modifications were located on the Lysine[3] residues of both substance P and [Lys][3]-Bombesin. ECD

MS/MS analysis of the metal-bound iridiated biomolecules showed similar challenges of iridium ECD

electron quenching, however upon electron capture and without the tethering imine bond, the

capture/transfer of the ECD electron to the Ir centre caused dissociation of the iridium

modifications, hindering analysis. However the use of carefully tuned ECD and ICR cell parameters

enabled effective ECD fragmentation, while retaining enough of the Iridiated fragments to accurately

locate the Ir binding location to the histidine[11] residue in [Lys][3]-Bombesin. CAD MS/MS was also

attempted on the iridiated biomolecules. CAD MS/MS caused dissociation of the metal-bound

iridium modification long before fragmentation of the peptide backbone could occur, and so

produced no sequence informative fragments to lead to the modification's location. CAD MS/MS of

the imine-bound iridiated peptides proved extremely fruitful, with the imine linkage tethering the

iridium complex to the biomolecule, and thus allowing normal CAD MS/MS trivially, effective

fragmentation of the peptide backbone and location of the imine bound species was successful. The

presence of the fixed charge site at the iridium centre was shown to alter the fragments produced

and observed in both CAD and ECD MS/MS. In addition the iridium modification was shown to

produce unique CAD and ECD MS/MS fragments/side chain losses, showing an influence on the

fragmentation mechanisms beyond electron quenching. These unique fragmentation channels could

be used to help identify iridium ligands present, given a high enough mass accuracy, which FT-ICR

MS can offer. Careful calibration was also shown to be critical in assigning these novel iridium

modifications and fragmentation channels, with deviations of less than an electron in mass being

unacceptable, examples of the level of discrimination needed were investigated. Reactions of the

iridium metallodrugs were also carried out on model proteins (lysozyme, myoglobin, and BSA) and

showed similar condensation product formation as with peptide reactions. Information gathered

within this chapter later informed experiments in Chapter 9 and enabled advances in scaling up

metallodrug-biomolecule reactions to proteome level analysis.

Chapter 7: Superacid Supercharging of analytes using CH_5^+
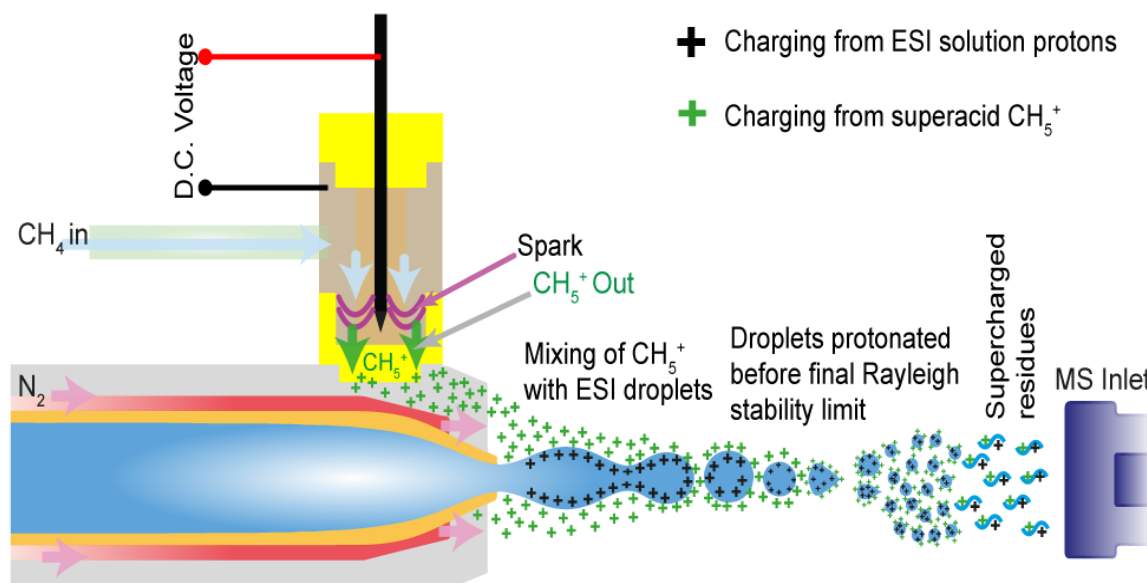
The analysis of all species by mass spectrometry relies on the effective ionisation of analytes of interest. In particular electrospray ionisation has become the ionisation method of choice for the majority of biomolecule analysis. Due to its ability to ionise peptides, proteins, and various other biomolecules into multiply charged species, instead of purely singly charged ions as with MALDI and CI, ESI and its variants have allowed the analysis of extremely large species by lowering the mass to charge ratio (through high charge states) into a more convenient and easily detectable region. Another key benefit of higher charge states for biomolecule MS is the increased efficiency and sequence coverage gained from conducting tandem MS on high charge state precursors compared to their low charge state equivalents. The increased cleavage coverage and fragment ion abundance greatly assists in characterising these species, which becomes more challenging as the parent ion mass increases.

Thus it is of ubiquitous benefit to biomolecule analysis to explore the possibilities of enhancing the charge state magnitude during the ionisation process. Great strides towards this goal have already been achieved, mostly via the use of solution phase additives which alter the surface tension of droplets formed during ESI. By increasing the surface tension of droplets, it was shown that the

charge density of the droplets could be increased and the resulting protein ions formed were of higher charge states than before. These so-called "supercharging agents" have greatly assisted biomolecule characterisation, both offline through direct-infusion ESI, and online through LC-ESI-MS experiments. The increased availability of higher order charge states increases fragmentation efficiency and sequence coverage during MS/MS experiments and since all MS detectors are inherently more sensitive for higher charge states, these techniques can also increase the sensitivity of measurements.

Though Chemical supercharging agents have been investigated for several years now, and have shown great improvements to MS and MS/MS analysis, they are not commonly used techniques as individual additives do not effectively supercharge all analytes and can have negative side effects, such as contamination, decrease in ESI spray stability, and disruption of delicate non-covalent/fragile interactions/PTM's. As a result there is still a desire to find new ways to access these higher levels of charging in a more direct, commonly implementable fashion, with or without the use of solution-phase supercharging additives such as sulfolane or m-NBA.

Chemical ionisation utilises CH_5^+ ions created by passing methane gas through an electrical discharge region. These CH_5^+ ions are incredibly acidic and have extremely low proton affinities, so much so they have been termed "superacid" ions, and are able to transfer protons to even non-polar entities in order to create ions. Unfortunately chemical ionisation is notoriously reliable for producing only singly charged ions, despite the use of such potent acidic reagents such as CH_5^+. If the proton transfer power of CH_5^+ could be combined with the multiple-charging capability of ESI, then it might be possible to enhance the ionisation and charging of biomolecule analytes into higher charge states.

To this end, the novel ESI-APCI supercharging source was developed, over many design iterations, the eventual course was able to effectively introduce CH_5^+ ions created from a glow-discharge chamber orthogonal to the ESI needle and with proper optimisation could direct them towards the

ESI droplets formed during nebulisation. The effect of the CH_5^+ ion s interacting with the ESI droplets was the enhancement of charge of protein analytes in the positive mode. Great increases in the charge states for a range of protein-containing solutions was achieved, for a range of solvent conditions, including aqueous samples, denatured protein samples, and even high concentration ammonium acetate for so called "native MS" protein samples. All showed enhanced charging due to the supercharging source via CH_5^+ ions. The benefits of using the supercharging source against solution phase additives induced; no additional sample preparation steps to optimise the volume/% of solution phase additive to add, the use of buffer/solution conditions is not affected by the supercharging reagent, increased charge states for proteins and peptides, increased intensity for small molecules, easily interfaced with existing LC systems, without addition/change of elution buffers, and enhanced MS/MS fragmentation efficiency and sequence coverage to name a few.

The ESI-APCI source was shown to be most effective on denatured biomolecule samples, most likely due to the open protein structure exposing possible protonation sites, and allowed observation of many more charge states for the myoglobin protein (maximum charge state increase from 17+ to 25+). Native MS samples benefitted from the supercharging mainly via increased intensity of the currently observed charge states, with large increases of over 500% measured, but no new charge states were observed, this is believed to be due to the folded tertiary structure of the proteins chosen with only a limited number of charge sites available. Small molecule supercharging showed intensity increases of nearly 300%.

Theoretical modelling calculations were then undertaken to investigate the possible factors and mechanisms involved in the interaction of CH_5^+ ions and charged ESI droplets. The interaction of the two positively charges entities instantly suggests a repulsive force between the two, but the MS data clearly showed an interaction. A pure electrostatic model was first investigated which consisted of randomly populating a sphere of a given diameter with a number of positive charges up to the Rayleigh limit for the droplet size. A point charge (CH_5^+ ion) was then introduced and the

electrostatic interaction between the CH_5^+ ion and every individual charge on the surface of the droplet was then calculated and summed together. The charge was then moved closer and the process repeated and iterated until the two were touching. The electrostatic model gave a predictable constantly increasing repulsive force, tending towards infinity at the point of contact. Though simplistic the model shows that there should be a high repulsive barrier to interaction between the two entities. A particle-particle dielectric model was then undertaken to account for the orientation of water molecules within the ESI droplet at low separation distances, and showed that at low charge density the repulsive barrier can revert and become attractive as the CH_5^+ ion moves close enough to the droplet to re-orientate the water molecules, repelling the δ+ hydrogen atoms and exposing the CH_5^+ ion to the δ- oxygen lone pairs, creating an attractive force. This model was promising and helped explain the MS data acquired, though it was quickly shown that at higher charge densities this effect was negligible and although the particle-particle model showed a slower increase in repulsion between the 2 entities, the repulsive force still increased dramatically at shorter separation distances. As a result it was believed that the interaction of CH_5^+ ions with ESI droplets must occur at lower charge densities and well before the droplet reaches its Rayleigh limit.

The supercharging ESI-APCI source was successful in assisting protein ionisation and could assist in subsequent MS/MS fragmentation and characterisation of biomolecules, and possibly metallodrug-modified biomolecules in the future. The use of supercharging instrumentation over the solution phase additives offers potential advantages for metallodrug modified species, especially the lack of changes to the solution conditions, as it has been shown many times, both within this thesis and without, that the addition of other solvents can easily disrupt/alter the metallodrug-biomolecule interactions, displacing modifications/altering equilibria and eliminating the chance of truly representative analysis. The supercharging ESI-APCI source could be a path to assisting characterisation without such disruptions.

The source has already found additional use in further investigations into CH_5^+-assisted ionisation and supercharging for other projects, and has been shown to drastically increase the fragmentation efficiency for protein MS/MS. Detailed tandem mass spectrometry techniques and analysis have already been undertaken to investigate the effect CH_5^+ ions and other supercharging agents have upon protein structure; the work is ongoing. The source also holds the potential to readily be coupled to liquid chromatography and assist in the ionisation and supercharging of eluted biomolecules, potentially enhancing the characterisation of entire proteomes by increased sensitivity and MS/MS efficiency already confirmed for protein MS and MS/MS respectively, this could also help overcome various issues associated with true top-down proteomics, such as low intensity precursor peaks, and lack of fragmentation efficiency.

Chapter 8: The Rhodium revolver
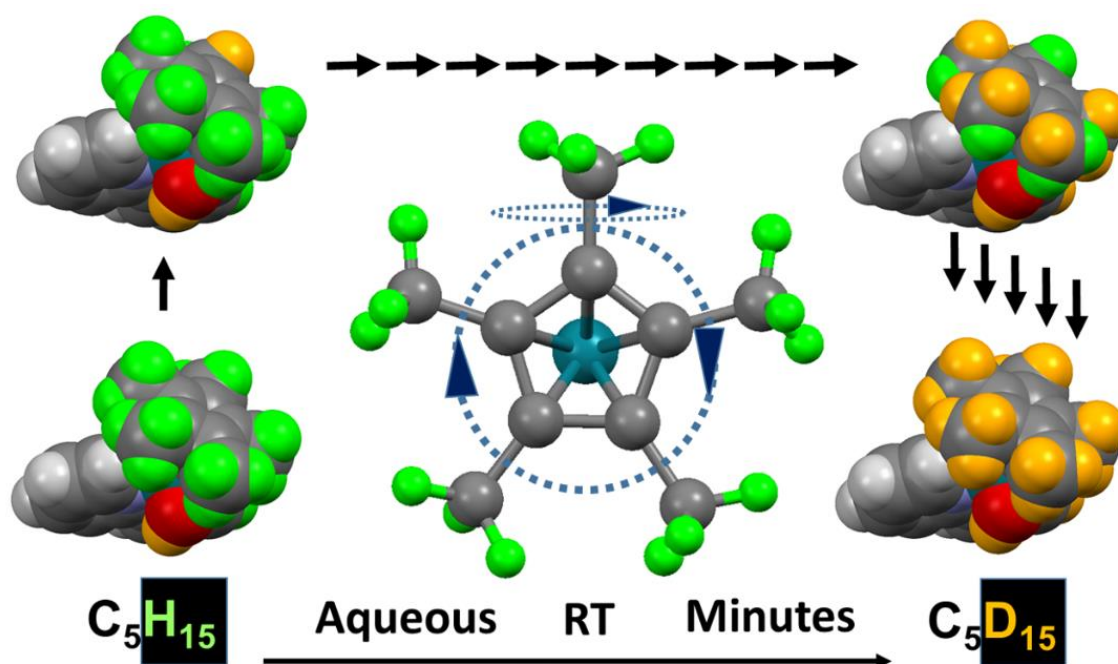

Novel metallodrug compounds are often shown to be very reactive towards biomolecules (as shown in previous chapters), however they can also be used to catalyse unique and seemingly unlikely reactions by utilising the unique chemistry of the metal centre and ligand combinations. As was the case for the rhodium-based compounds studied in Chapter 8, all of the species were shown to be

able to exchange the hydrogen atoms bound to the methylated cyclopentadienyl ring of the piano-stool complexes with deuterium atoms from deuterated solvent species. A simplistic view of this reaction is that it is synonymous with widely used Hydrogen-Deuterium Exchange (HDX) reactions in protein analysis MS and MS/MS. However, unlike protein/biomolecule HDX, the exchange occurring within the rhodium compounds was centred around the deprotonation of methyl groups and a thermodynamically beneficial substitution with deuterium atoms, a particularly unique and previously unobserved phenomenon. These reactions were then followed using a variety of techniques, including nESI-FT-ICR MS, to show that the exchange of the hydrogen atoms was sequential and complete exchange of all methyl-bound hydrogen atoms could occur in less than 1 hour at room temperature and pressure, a remarkably fast reaction for such unreactive hydrocarbon groups. This study quickly expanded to similar rhodium piano stool complexes, 10 in total, with varying ligand configurations and sizes, extending the aromatic ligand system of the piano-stool complex, and varying bidentate ligand identity to probe then extent of HDX, and the reactive area accessible from the metal centre, which was believed to be critical for this exchange process. It was shown that all rhodium piano-stool complexes were able to undergo HDX reactions while in deuterated solvents, to varying degrees depending on the ligand structure, with groups close to the rhodium centre being substituted effectively, while those further away remaining unsubstituted. Molecular modelling of the reactions by collaborators showed that the process was most likely proceeding through a deprotonated cyclopentadienyl ligand intermediate, causing a $C=CH_2$ bond between the ring and deprotonated methyl group. This $C=CH_2$ bond could then react with deuterated solvent bound to the rhodium centre and cause a deuterium atom to re-saturate the hydro carbon into a CH_2D group and so on until complete deuteration of the groups was achieved. This process was attempted on an analogous iridium compound with identical ligands to that of the first Rhodium compound studied ([(Cp*)-Ir-(Bipy)Cl]. Iridium being directly below rhodium in the periodic table seemed as though it may be able to reproduce this effect, but to no avail, the

compound remained fully hydrogenated despite days in deuterated solvent, truly this appeared to be a rhodium-specific reaction.

Though not directly related to their interaction with biomolecules, the rhodium revolver reactions serendipitously provided a method of effectively tagging rhodium species with deuterium atoms easily and quickly. The deuterated rhodium compounds were shown to remain deuterated once placed in pure H_2O, with no back-exchange to hydrogenated species even after 7 days. The HDX products were obviously more thermodynamically stabilised than the original complexes. This stable isotope labelling strategy enables the possibility of exploiting these effects via MS analysis and tracing of the compounds , possibly in vivo. By mixing a known proportion of H_{15} and D_{15} (all deuterated) Rhodium complex, a characteristic pair of isotopic distributions would be observed during ESI/nESI-MS, separated by ~15 Da. This mass difference could then be tracked using isotopic search algorithms (see Chapter 10) and used to track the distribution of the drug in vivo and possibly to help identify biomolecule targets. Usually this sort of isotopic labelling is not required for analysis of transition metal-containing species, as many already possess a unique isotopic pattern, however certain elements (such as rhodium) express only one stable isotope and so are not amenable to this analysis and tracking. The rhodium revolver HDX strategy provides and alternative, effective, and fast way of inducing unique and traceable isotopic distributions using just deuterated solvent. nESI-MS of rhodium-based metallodrugs was also shown to be affected by metal complex-specific phenomena during ESI/nESI based ionisation, namely loss of bound negatively charged ligands during the ionisation process. The loss of (negatively charged) ligands during ionisation often resulted in the increase of charge of the analyte species, enabling detection, the driving force of which was attributed to the electric field imposed to enable ESI/nESI. Such affects were documented and followed the trends observed in Chapter 6 for iridium-based compounds. The same trends were not observed for the square planar Pt(II) complex modifications observed in Chapter 3, showing again the influence of inorganic/organometallic chemistry of metal centres on the process.

Chapter 9: Scaling up to proteome wide investigations
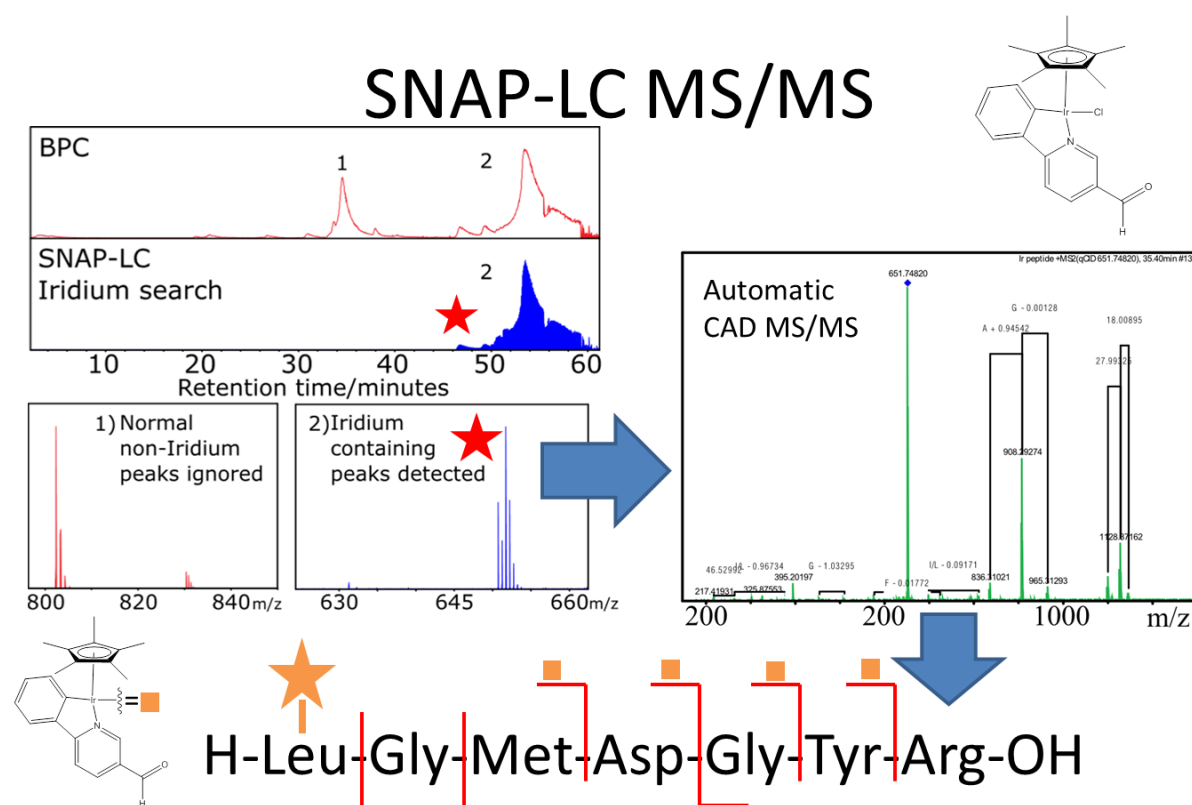
Chapters 3-6 were focused on the detailed characterisation of isolated biomolecule and metallodrug interactions in order to study the novel compounds being developed to combat cancer. These in depth studies provide the detailed chemical information about metallodrug behaviour and allow knowledge and tactics for effective MS and MS/MS characterisation, which can often be much more challenging than unmodified biomolecule MS analysis. However this has all been done with the end goal of scaling up these isolated biomolecule studies into effective and accurate strategies to tackling proteome-wide characterisation of biomolecules in cancerous cells and to provide insights into cellular changes due to metallodrug administration. This proteome-wide information would be able to provide the information needed on the impact of individual metallodrugs, to see which pathways the drug is interrupting/amplifying and to see how the cancerous cells respond to these foreign agents and finally provide the critical data needed for better intelligent drug design. Unfortunately although MS-based proteome characterisation has been well established in recent

year and is now able to provide proteome information readily, the addition of metallodrug modification greatly complicates the task and introduces challenges in not only performing the analysis but mainly in retaining the unique metal-based modifications and providing accurate and reliable analysis which is able to compensate for these non-standard entities. Without these alterations to strategies the usual, and readily available, techniques, methods, and subsequent data analysis would be possible, easily achievable, but would drastically misrepresent the sample and would not provide the accurate information needed to the level MS is known to be able to achieve with the correct care and strategy.

The final chapter of the thesis brings together the information gathered in the detailed MS and MS/MS studies from previous chapters and applies them to the idea of moving on to larger, biological scale systems, the methods available, how they may be adapted for metallodrug-containing samples, and what complementary information these tools can provide. The first section deals with the challenge associated with identifying which peaks have been modified by metal-based compounds, and then accurately assigning a mass to them, which is unfortunately impossible using the standard automated peak picking associated with LCMS for proteomics. The second section explores the use of a new MS/MS technique to characterise complex mixtures of biomolecules without using chromatographic separation, 2-Dimensional Mass Spectrometry, offering not only a different way to analyse the complex proteomic mixtures, but also to avoid the disruption of metallodrug-modifications which occur during chromatographic separations and to avoid the time-scale limitations of LCMS analysis on FTMS instruments.

The main barrier preventing the use of standard proteomic strategies for the characterisation of metallodrug modified proteomes are the use of organic solvents and acids which can disrupt metallodrug-biomolecule interactions, and the unique, but complex isotopic patterns of transition metals influencing biomolecule MS signals, preventing standard peak picking algorithms from assigning accurate masses to detected peaks.

The first strategy to accurately achieving proteomic analysis of metallodrug-modified proteomes was based on LC-MS for separation and tandem mass spectrometry characterisation, in a similar fashion to standard proteomic analysis, with the main goal of combatting the metal element influence on isotopic distributions. Every individual element has a number of stable observed isotopes, the combinations of which influence the overall distributions of isotopologues observed in MS spectra, these are predictable and often very easily calculated. However transition metals often have many more isotopes, in varying ratios, than the standard CHNOS elements observed in poly(amino acid) MS and MS/MS. Standard biomolecule and proteomic peak picking algorithms are designed to either pick all peaks observed in a spectrum for database searches, or are designed to analyse peak distributions, assume CHNOS compositions, and assign monoisotopic masses for more accurate assignments when searching databases thereafter. The more complex isotopic patterns of transition metals are not compatible with either of these approaches, with the first leading to many assignments for each isotopic distributions (for example for 6/7 observable osmium isotopes broadening the usual isotopic pattern), and the second leading to fewer, but still multiple assignments at the wrong m/z values as it will ignore isotopologues which increase in abundance with increasing m/z (the reverse of usual CHNOS species), and/or would pick out many peaks each as a new species due to the expectation and bias of CHNOS predictions. Though they may seem simple data analysis issues, the accurate peak picking and assignment of species is the cornerstone of effective LCMS characterisation of proteomes, without it the data collected will at the very least be not fully interpretable while at most (and much more likely) lead to extremely erroneous outcomes, completely misrepresenting both the data and the proteome, which can have large negative effects in drug development processes if trusted. In order to be able to assign accurate masses to metal-containing peaks, but also to identify them in the first place, these isotope distributions needed to be taken into account.

Despite all of the challenges faced in the study of the metallodrugs and the metallodrug-modified biomolecules studied over the course of the Chapters, one of the few advantages of metal-

containing species was that the unique and characteristic isotopic patterns immediately revealed the presence of the metal centre. This influence could then be exploited to assign metal or non-metal containing peaks, and if this could be done over an entire LCMS chromatographic run then all of the metal-containing LC peaks could be revealed. If these isotopes could then be taken into account when assigning these species then accurate masses could then be assigned autonomously, circumventing the assignment-accuracy issues of metallo-proteomics.

To this end isotopic prediction algorithms have been modified to include metal isotopic patterns and generate metal-matched peak spectra by Bruker Daltonics. These algorithms were originally designed to accurately assign unmodified biomolecules, but were easily modified for metal species. The use of these peak picking algorithms was shown to be extremely effective at discriminating metal-containing peaks for normal MS scans. However LCMS runs contain many thousands of individual MS spectra. The goal was to apply the algorithm to each spectrum in turn and plot the presence of metal containing species in each LCMS data point, once iterated through the thousands of spectra contained within the run all metal-containing LC peaks were revealed to high accuracy. This analysis was specific to the type and number of transition metal centres desired, with custom isotope distributions able to be programmed in as desired. The SNAP-LC algorithm developed was altered through several designs to achieve accurate assignment of metal containing peaks only, and to enable desktop computers to be able to process the many thousands of spectra without memory issues. In addition the algorithm was attempted on both nLC-UHR- FT-ICR MS data acquired and nLC-HR-ToF-MS data; it was shown that although viable on both platforms, the isotopic searches required high-resolution separation between isotopic peaks, and suffered with lower resolution data where individual isotopes may not be completely resolved from one another. Regardless there were no issues when using UHR data, and attempts are already underway to use modified algorithms to cope with lower resolution data if needed.
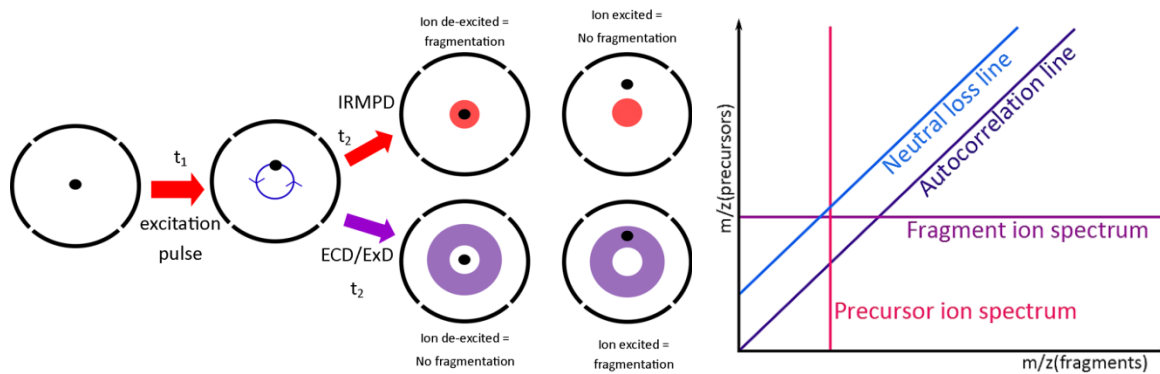
The second strategy, of using 2DMS to study metallodrug-modified proteomes was then investigated by application to A2780 cancer cell line trypsin digests. The A2780 cells were grown with or without treatment with a potent osmium piano-stool metallodrug, and were then lysed, protein extracted, purified, digested and the resulting peptides desalted before nESI analysis followed by nESI-2D-ECD/IRMPD-FT-ICR MS characterisation. 2DMS has already been shown in a number of studies to be an effective analytical method for the assignment and characterisation of biomolecules, the application shown herein was a larger scale application, requiring high resolution isolation of precursors not available using convention quadrupole isolations. Many thousands of precursor peaks were observable in the nESI-FT-ICR Ms of the cancer cell digest samples, all of which were then characterised using 2DMS via ECD and IRMPD analysis separately. The resulting 2DMS spectra were extremely complex, with each of the thousands of precursor species fragmented producing many more fragments separated in 2-dimensional space in the 2DMS plot. 2DMS was shown to be a viable tool in producing effective MS/MS spectra of all observable species, which can then be searched and matched against proteomic databases, using tools produced for traditional LCMS based proteomics, to assign the biomolecules observed and highlight proteome wide changes due to metallodrug administration. The detailed analysis of this data is ongoing, but rapidly improving. The main benefits for metallodrug-modified proteomic analysis brought by 2DMS is the long-awaited ability to use metal complex-friendly solvent systems, which will enable retention of the metal complex once bound to the biomolecules, unlike LCMS where they were shown to readily dissociate and/or relocate. In addition 2DMS enables careful, and often lengthy MS/MS experiments which are not

compatible with the duty-cycle limitations required by LCMS. The rapid scan rate needed for effective LCMS (~1 Hz) is a limitation for UHR-MS instruments which often require several seconds per scan to achieve UHR. ExD MS/MS experiments, as shown throughout this thesis to be critical for effective metallodrug-modified biomolecule characterisation, can also take additional time to apply (~0.1-1 s) and reduce scan to scan rates even more, with standard dissociation techniques (CAD) not suitable for metallo-proteomics, this can hinder the acquisition of acceptable data. 2DMS circumvents these issues by eliminating the need for LC and LC-induced scan rate requirements, and using standard nESI ionisation for sensitive and prolonged acquisitions.

Future applications of both the SNAP-LC and 2DMS techniques to characterisation of metallodrug modified proteomes are underway and are expected to be applied to many systems, utilising the data and knowledge of metallodrugs gained from the initial detailed studies in Chapters 3-6 to inform better tuning and analysis of the resulting data. It is the hope of the Author that these techniques will eventually mature into viable, effective, and most importantly accurate methods for the analysis of proteomes, without the erroneous assumptions made in many proteomic studies. Constant improvements in MS and computational instrumentation are being developed and will only help improve the acquisition of data and analysis of the resulting data, eventually detailed and accurate proteome wide characterisation of metallodrug-biomolecule interactions will be readily achievable and will help to inform better drug design, synthesis, and application to help combat many diseases and conditions including and beyond cancer.