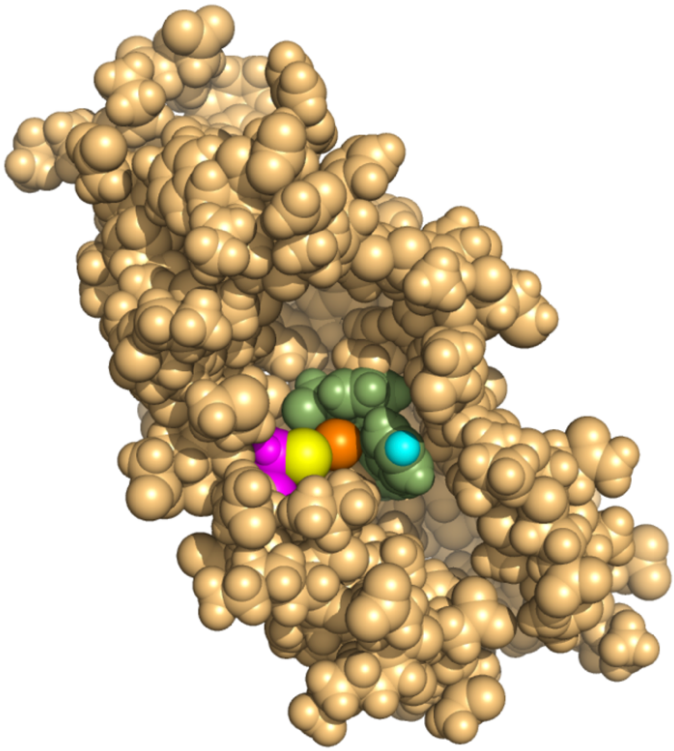
UNIVERSITY OF CAMBRIDGE

Understanding the Speciation of Ruthenium Arene Complexes as Precursors to Unnatural Enzyme Cofactors



George Samuel Biggs

Peterhouse

September 2020

This dissertation is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee

George Samuel Biggs

Abstract

Understanding the Speciation of Ruthenium Arene Complexes as Precursors to Unnatural Enzyme Cofactors

George Samuel Biggs

Ruthenium arene complexes have been extensively explored as metallo-pharmaceuticals and as small molecule catalysts. Exploring the overlap between these areas, this thesis describes a body of work aimed at quantitatively understanding the biological speciation and catalytic behaviour of ruthenium arene complexes when exposed to the many potential Lewis basic ligands provided by protein scaffolds.

Combining nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry (MS), X-ray crystallography and other biophysical characterisation techniques, the speciation of ruthenium arene bipyridine complexes with small molecule amino acids, short peptides and whole proteins has been monitored. A ^{19}F NMR spectroscopic method was developed to quantitively trace the preferred amino acid binding partners of ruthenium complexes coordinated to fluorinated ligands. Liquid chromatography mass spectrometry (LC-MS) was used to explore the ligand exchange behaviour between ruthenium complexes and protein scaffolds, particularly variants of the small proteins ubiquitin and cytochrome b_{562}. Additionally tandem MS/MS experiments were used to determine the final protein binding sites of non-fluorinated ruthenium fragments coordinated to proteins.

The resulting deep understanding of how ruthenium arene complexes coordinate to specific proteins was used to develop artificial metalloenzymes (ArMs) with direct protein-metal coordination. The catalytic capabilities of these hybrid systems was then explored. ArMs are synthetic biocatalysts that result from either the combination of an artificial metallo-cofactor being introduced into a protein scaffold or a natural metalloprotein being evolved to perform catalytic reactivity.

Taking inspiration from naturally occurring metalloproteins, this research showed that it was possible to form ArMs *via* a ligand exchange process between a ruthenium arene precursor complex and a protein scaffold, resulting in the precursor complex being activated towards catalysis. Direct protein – metal coordination enables the protein to impart both an electronic and steric contribution to catalysis and attenuate reactivity at the metal centre in ways that have not been previously studied. The four helical bundle protein cytochrome b_{562} was selected as

the protein scaffold for ArM development due to its dynamic structure and nascent haem binding site, which in the absence of haem provides a hydrophobic pocket capable of accommodating a ruthenium cofactor and catalytic substrate. Cytochrome b_{562} – ruthenium hybrids (with direct coordination) were identified that have catalytic transfer hydrogenation activity greater than a known dimeric catalyst. This demonstrates an exciting starting point to explore the evolutionary potential of these ArMs through directed evolution, hopefully enhancing catalytic activity.

Dedicated to Manaka Shibuya

'If it was not for chemistry, I'd probably be a lawyer,
and I am glad that's not the case.'

Jean-Marie Lehn – Chemistry Nobel Laureate 1987

Acknowledgements

Studying for a PhD has been a wonderful experience and privilege. There are many people who I wish to thank who have made my time in Cambridge thoroughly enjoyable and rewarding, so here goes…

First and foremost, I would like to thank my supervisors, Paul Barker and Sally Boss. They have been a constant source of help and advice, both scientifically and personally, throughout my time in Cambridge. I have learnt so much from both Paul and Sally who have supported me through both the good and the tough times in Cambridge, thank you.

Paul, Sally and Nick Bampos have always encouraged me to teach alongside my PhD, which I have found thoroughly rewarding. In particular, I have supervised four students during my PhD who have helped me out in the lab: Amzar Muzani Bin-Maarof, Alex Heavers, Jamie Klein and Manaka Shibuya. It has been particularly enjoyable working and collaborating with Jamie, whose name you will see many times throughout this thesis. Jamie has now joined the group, and his enthusiasm, commitment and intellect have helped me greatly in the last two years of my PhD. My thanks also go to Alasdair Keith who has shared the struggles of protein purification with me.

There have been a number of others who have helped me in my scientific endeavours. In particular, I would like to thank: Paul Skelton, Asha Boodhun, Roberto Canales and Dijana Matak-Vinkovic of the mass spectrometry department; Duncan Howe, Peter Grice and Andrew Mason of the NMR department; Andrew Bond for his X-Ray crystallography support; Mark Skehel, Sarah Maslen, Trevor Rutherford, Stefan Freund at the Laboratory of Molecular Biology; and the floor technicians, Steve Wilkinson, Carlos Davies, Naomi Hobbs and Kevin Judd, who have provided excellent support throughout.

I have made some great friends in Cambridge who I have shared some amazing experiences with. The members of the third-floor corridor in particular have kept me smiling through all the failed experiments and laboratory mishaps: Evan and Lily, your friendship and support has meant so much to me; Josh and Jess, thank you for all the memorable times; and Jonny and Andy, thank you for your constant friendship. I have also been lucky to live with some great people in Cambridge, in particular Will and James who have shared Geldart Street with me.

Finally, I would not be where I am today without the unconditional and unwavering love and support of my family. I would like to say thank you to: My Nana (Dr Joan Ladbury) and my Uncle John who inspired me to be a scientist; to Eleanor, Tom and Olivia, who bring me so

much happiness; to Louis, who is my best friend; and to my Mum and Dad, who are always there for me, have always been proud of me, and why all this means so much.

List of Publications

Biggs G; Klein OJ; Boss S & Barker P, Unlocking the full evolutionary potential of artificial metalloenzymes through direct metal-protein coordination. *Johnson Matthey Technol. Rev*., 2020, **64**, (4), 407–418.

Biggs G; O'Neill M; Carames Mendez P; Scrase T; Lin Y; Bin-Maroof AM, Bond A; Boss S & Barker P, Use of a fluorinated probe to quantitatively monitor amino acid binding preferences of ruthenium(II) arene complexes. *Dalt. Trans*., 2019, **48**, 6910–6920.

Table of Contents

List of Abbreviations

Ac	acetyl
Å	angstrom
ADC	antibody drug conjugate
Ala	alanine
ArM	artificial metalloenzyme
Asp	aspartic acid
Asn	asparagine
ATH	asymmetric transfer hydrogenation
Ben	benzene
Bipy	2,2'-bipyridine
Boc	tert-butyloxycarbonyl
br	broad
^{t}Bu	tert-butyl
Bz	benzoyl
CD	circular dichroism
COSY	homonuclear correlation spectroscopy
p-Cym	*para* – cymene = 1-isopropyl-4-methyl benzene
Cys	cysteine
Cyt	cytochrome
d	doublet
Da	Dalton(s)
DEA	diethanolamine
DCM	dichloromethane
dd	doublet of doublets
^{o}C	degrees Celsius
Dha	dehydroalanine
DHFR	dihydrofolate reductase
DFT	density functional theory
dmb	1,3- dimethylbenzimidazol-2-ylidene
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DTT	dithiothreitol

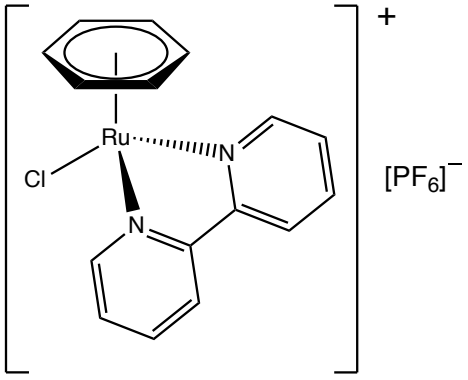
E.coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
en	ethylenediamine
EPR	electron paramagnetic resonance
Eq.	equivalent(s)
ESI-MS	electrospray ionisation mass spectrometry
Et	ethyl
et al.	et alia (and others)
FBipy	difluorobipyridine
Gln	glutamine
Glu	glutamic acid
Gly	glycine
GSH	reduced glutathione
GSSG	oxidised glutathione
h / hrs	hour(s)
HB	haemoglobin
HEWL	hen egg white lysozyme
His	histidine
HMB	hexamethylbenzene
HOMO	highest occupied molecular orbital
HR	high resolution
HSA	human serum albumin
HSQC	heteronuclear single quantum coherence
Hz	Hertz
ICP-OES	inductively coupled plasma optical emission spectroscopy
IEC	ion-exchange chromatography
im	imidazole
IMAC	immobilised metal affinity chromatography
IPA	isopropyl alcohol
IPTG	isopropyl β-D-1-thiogalactopyranoside
K	Kelvin
L	litre(s)
LB	lysogeny broth

LC-MS	liquid chromatography mass spectrometry
LFSE	ligand field stabilisation energy
LmrR	Lactoccal multidrug resistant Regulator
LUMO	lowest unoccupied molecular orbital
LR	low resolution
m	multiplet
MB	myoglobin
MaxEnt	maximum entropy
M	molar
MBP	maltose binding protein
MeCN	acetonitrile
Me	methyl
Mes	2-morpholin-4-ylethanesulfonic acid
MeOH	methanol
Met	methionine
MHz	megahertz
Min	minute(s)
MLCT	metal to ligand charge transfer
Mol	mole(s)
m/z	mass to charge ratio
NADH	nicotinamide adenine dinucleotide
NaP_i	sodium phosphate
NHC	N-heterocyclic carbene
$^nJ_{AB}$	n-bond coupling constant between nuclei A and B
NMR	nuclear magnetic resonance
OM	olefin metathesis
PAGE	poly(acrylamide) gel electrophoresis
PEG	poly(ethylene glycol) 4000
Ph	phenyl
Phos	phosphate
PMT	photomultiplier tube
PPTS	pyridinium *p*-toluenesulfonate
PTA	1,3,5-triaza 7-phosphaadamantane
ppm	parts per million

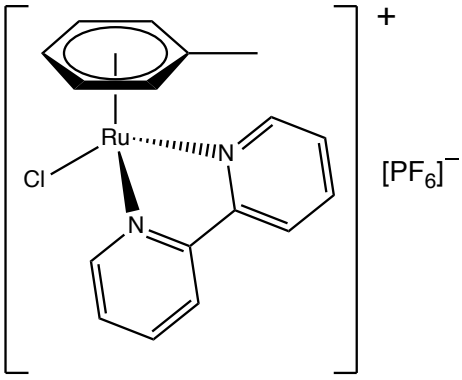
ppb	parts per billion
^{i}Pr	isopropyl
PYA	pyridylidene amide
PYR	pyridine
q	quartet
ROS	reactive oxygen species
rpm	rotations per minute
rt	room temperature
s	singlet
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
secs	seconds
TB	terrific broth
TCEP	tris(2-carboxyethyl)phosphine
Tev	tobacco etch virus
Tf	triflate
TFM	di(trifluormethyl)bipyridine
TH	transfer hydrogenation
THF	tetrahydrofuran
TrxR	thioredoxin
THP	tetrahydropyran
TOF	time of flight / turnover frequency
TON	turnover number
Tol	toluene
Tris	trisaminomethane
TY	yeast-tryptone
Ubq	ubiquitin
UPLC	ultra-performance liquid chromatography
UV-Vis	ultra-violet visible
V	Volt(s)
WT	wild-type

List of Compounds

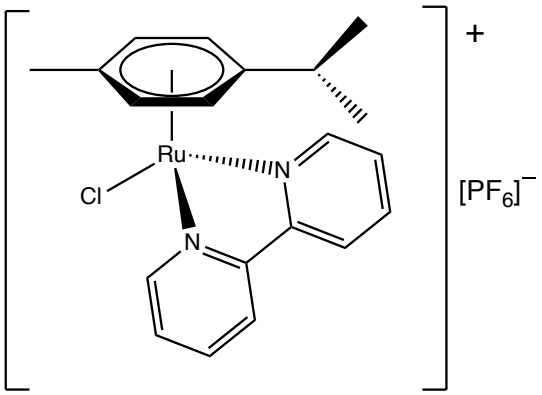
[1]



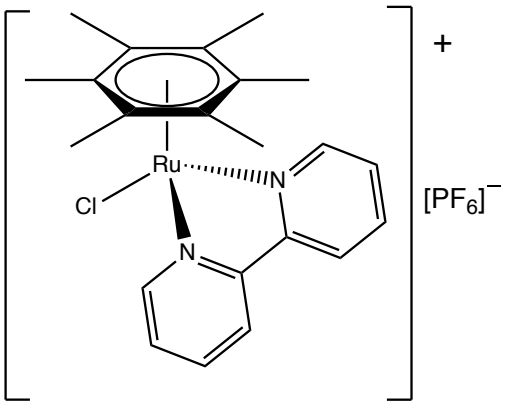
[2]



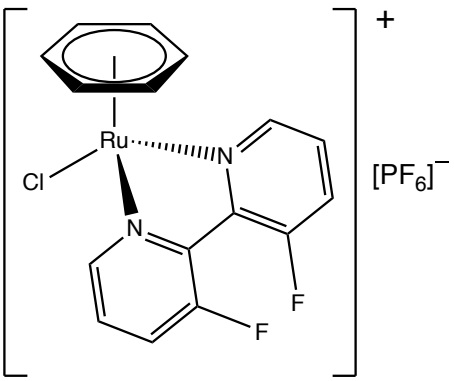
[3]



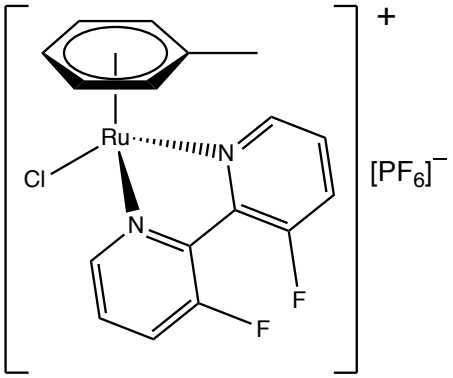
[4]



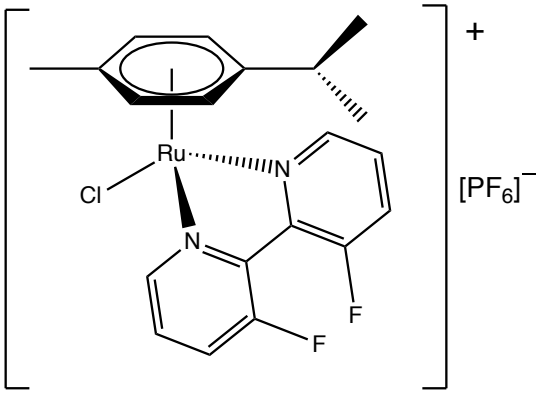
[5]



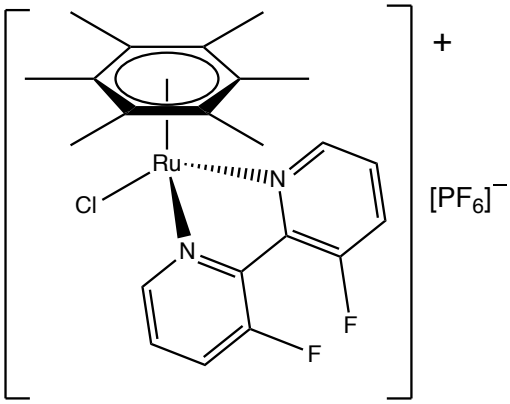
[6]



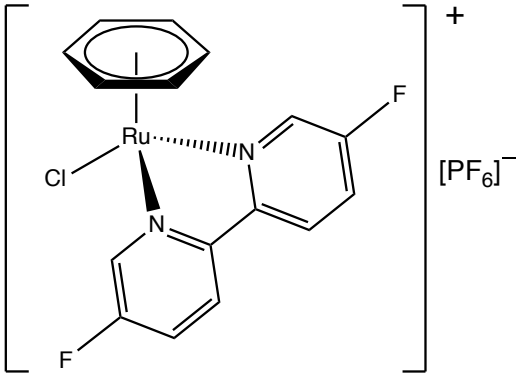
[7]



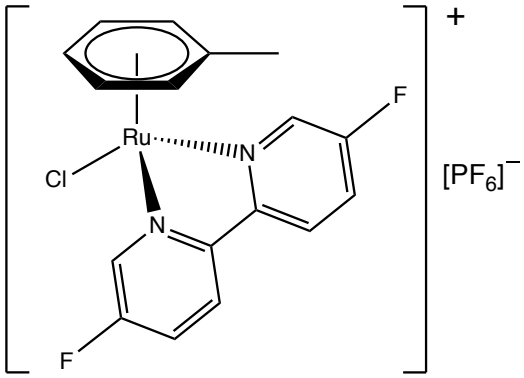
[8]



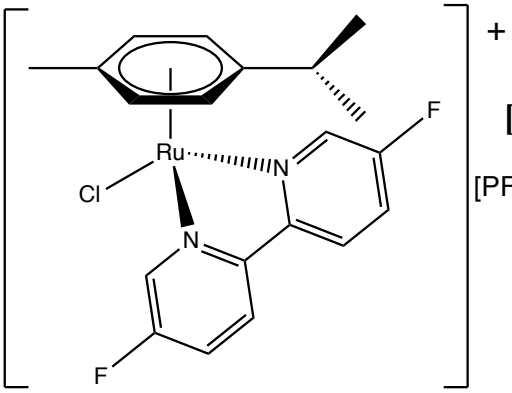
[9]



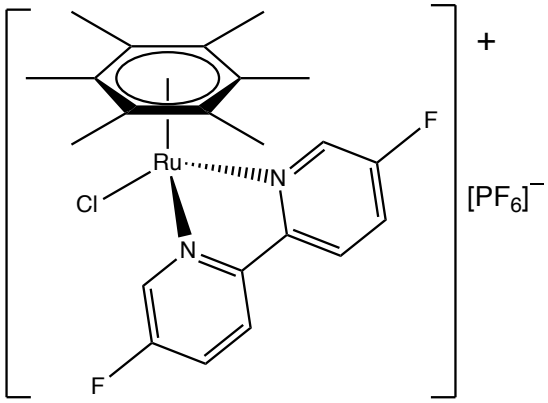
[10]



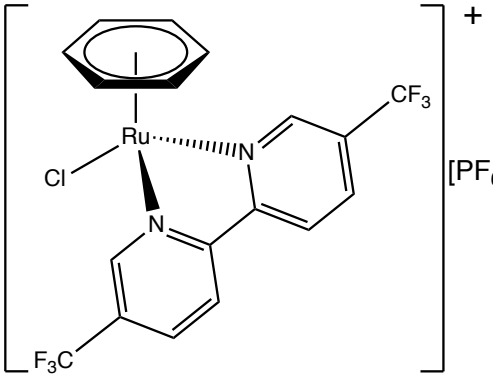
[11]



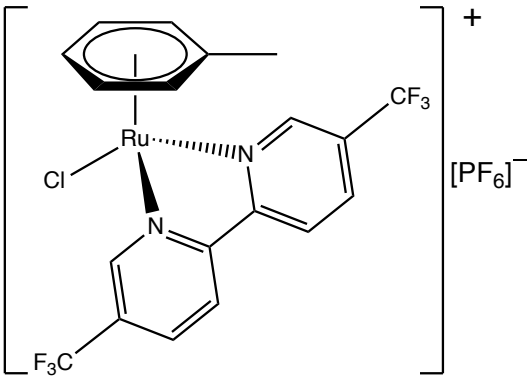
[12]



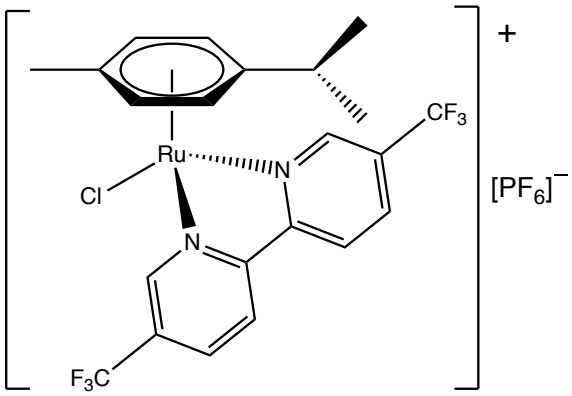
[13]



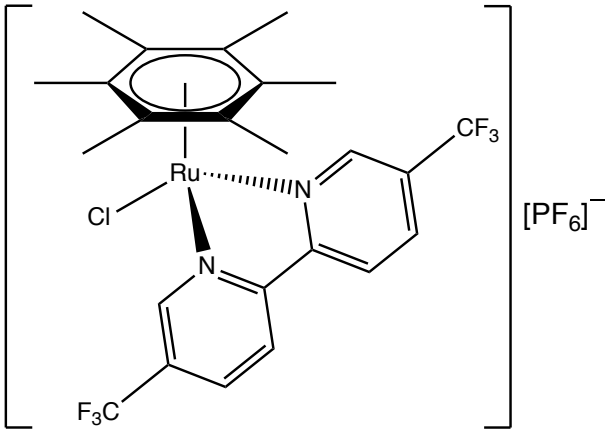
[14]



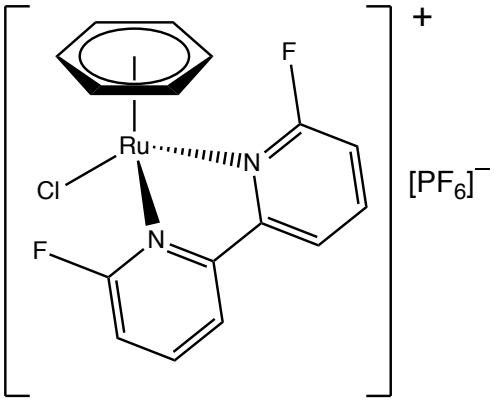
[15]



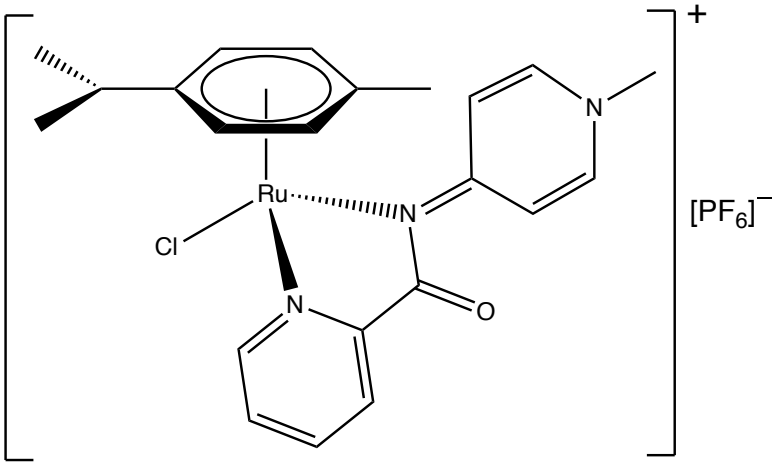
[16]


CF3
Ru
N
Cl
N
F3C
[PF6]−

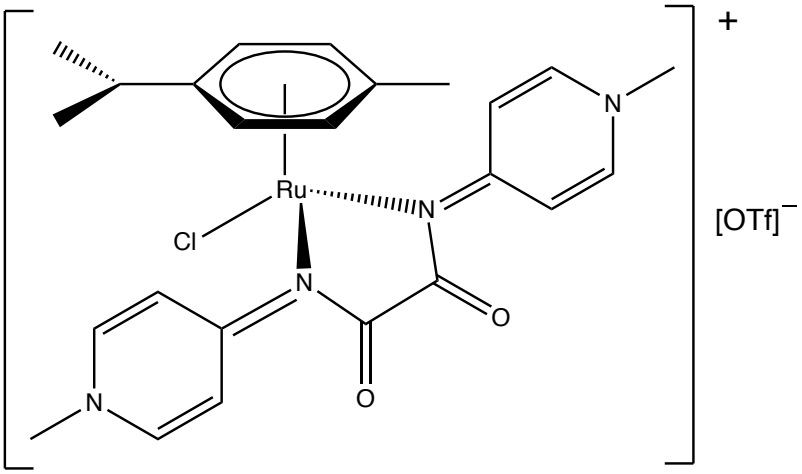
[17]


F
Ru
N
Cl
N
F
[PF6]−
+

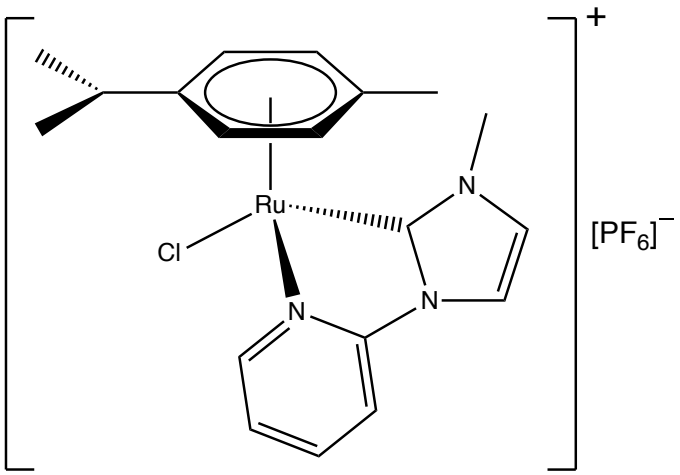
[18]


N
Ru
N
Cl
N
O
[PF6]−
+

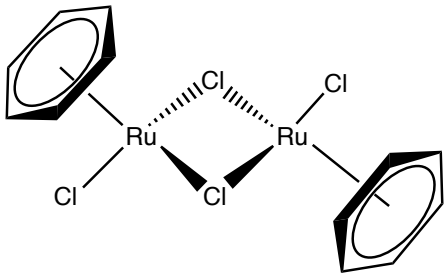
[19]


N
Ru
N
Cl
N
O
O
N
[OTf]−
+

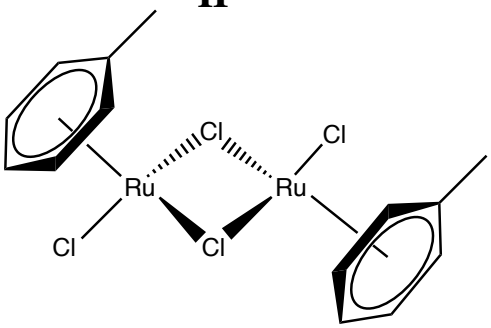
[20]



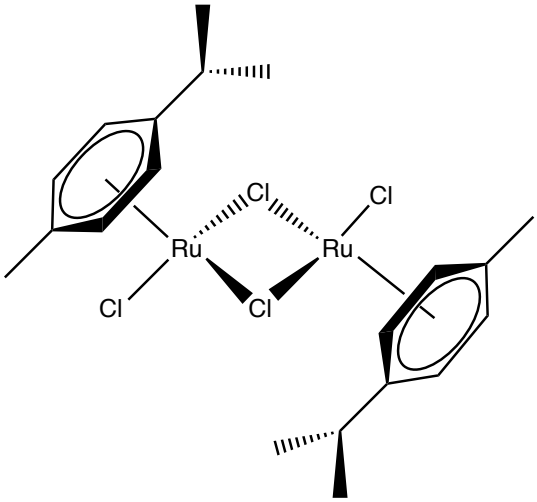
I



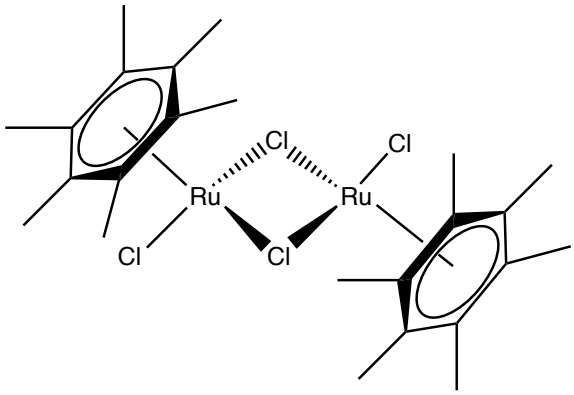
II



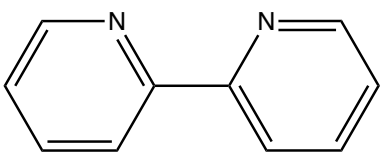
III



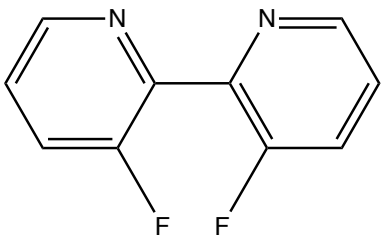
IV



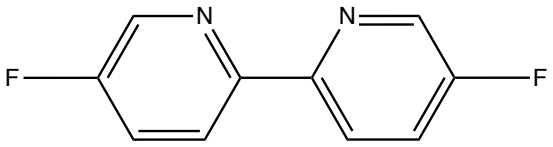
A



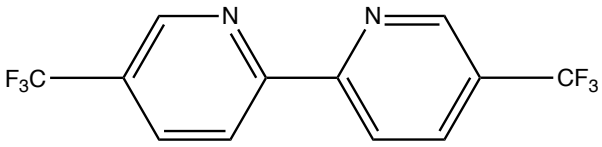
B



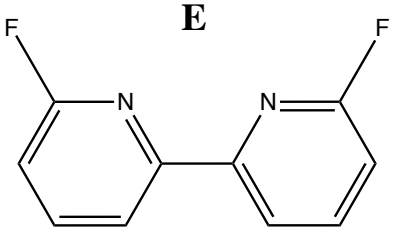
C



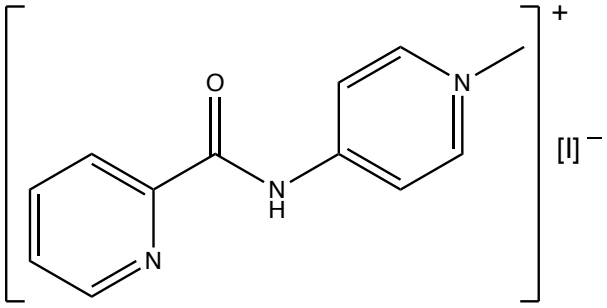
D



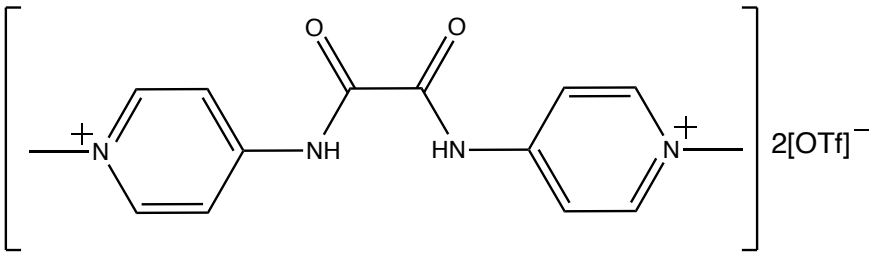
E



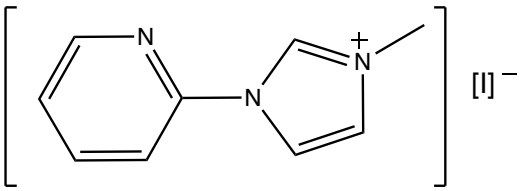
L1



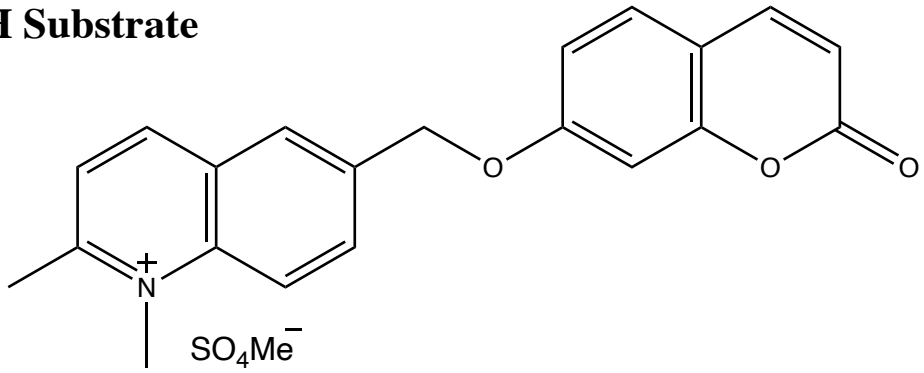
L2



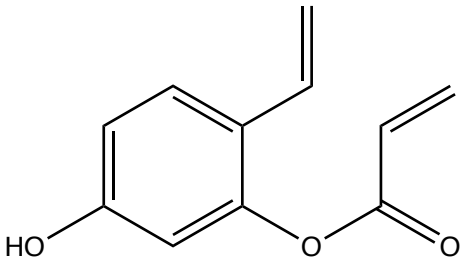
L3



TH Substrate



OM Substrate



Chapter 1

1. Introduction

1.1. Metals in Synthetic and Biological Catalysis

The production and functionalisation of a large range of fine chemicals, biomolecules and materials is underpinned by catalytic chemistry, with the resulting products being crucial for health, quality of life and the global economy. Catalysis can be broadly split into three main areas, which are all used in everyday industrial processes: (a) heterogeneous, solid catalysis, (b) homogeneous, small molecule catalysis and (c) enzymatic catalysis (including metalloenzymes).[1,2]

Unlike heterogeneous catalysis, which occurs at the interface between two phases, homogeneous catalysis is molecular in nature and occurs at a single site, resulting in greater selectivity, milder reaction conditions and greater molecular efficiency than heterogeneous alternatives. The disadvantages of homogeneous catalysts are that they are often shorter lived and less robust their heterogeneous counterparts and their separation from products is challenging. Transition metal catalysts consist of a metal ion or ions, with a number of ligands coordinated, forming a primary coordination sphere. Different ligands can control the reactivity, chemo and stereoselectivity of the catalysis, and there is an extensive body of research aiming to optimise ligands for catalytic transformations.[3] Such ligand optimisation is a difficult process, as the chemical possibilities for metal-ligand combination are vast.

Like homogeneous catalysis, enzymatic catalysis in biological systems also involves molecular species. Biological catalysts contain a combination of the 20 amino acid building blocks which fold into a 3D structure with one (or more) active site(s), which localise, bind and organise substrates ready for catalytic transformation. Many enzymes have transition metal ions at their active site and these metal ions play a crucial role in catalysis: this family of enzymes are called metalloenzymes. Scientists have long been fascinated by how proteinaceous ligands tune the properties of metal ions for function, and with technological advances, metalloenzymes can not only be studied but research has been driven towards designing and engineering metalloenzymes for the synthesis of fine chemicals. This research is focussed towards the goal that petroleum based, energy intensive and inefficient chemical synthesis can be replaced by enzymes to fulfil societal needs with enhanced efficiency.[4]

The most abundant transition metals found in metalloenzymes are Mn, Fe, Co, Cu, Zn and Mo, and biology has evolved remarkable methods of being able to accumulate, transport, store and use these six metals for catalysis. Metal ions that are required for an enzyme's activity as a catalyst are called cofactors. These cofactors increase the rate of the chemical reaction and are essential to enzymatic catalysis. Metal cofactors can be found in proteins encapsulated by ligands supplied by the protein (for example zinc in carbonic anhydrase), or with non-protein ligands coordinated (for example a porphyrin ring coordinated to iron in the case of the haem cofactor).

Biology has evolved to predominantly incorporate and use 3d metals over 4d metals for catalysis, apart from molybdenum, which will be discussed later. A number of reasons can be used to explain this preference, but most obvious is that in prebiotic chemistry the 3d metals were more naturally abundant and available for incorporation than the 4d metals (**Table 1.1**).[5] Therefore, when exploring the usefulness of 3d metals for catalytic transformations and structure, cells developed machinery that enabled them to selectively accumulate specific metals based on their electronic properties and size. Through evolution, biology has ensured that organisms will accumulate and use 3d metals and largely ignore 4d and 5d metals. Biological systems will decide whether to accumulate a metal based on a combination of abundance and usefulness, 4d and 5d metals are simply not abundant enough to be fit for purpose. In the presence of toxic 4d and 5d metals, biology has evolved systems to export or pacify them (e.g. lead and cadmium).

An interesting evolutionary question to consider is whether, if the later 4d and 5d transition metals had been as naturally abundant as 3d metals when life began, would biology have used these elements for metalloenzymes instead or as well?

Table 1.1: Abundance of a selection of transition metal elements in the Earth's crust given in mg/Kg.

Cr	**Fe**	**Co**	**Ni**	**Cu**	**Zn**
1.02×10^{2}	5.63×10^{4}	2.5×10^{1}	8.4×10^{1}	6.0×10^{1}	7.0×10^{1}
Mo	**Ru**	**Rh**	**Pd**	**Ag**	**Cd**
1.2×10^{0}	1.0×10^{-3}	1.0×10^{-3}	1.5×10^{-2}	7.5×10^{-2}	1.5×10^{-1}

1.1.1. 3d vs 4d Metals; When Biology is Given a Choice

The group 6 metals molybdenum and chromium are the exception to the trend of biological systems accumulating 3d over 4d transition metals. Molybdenum is essential for life, whereas

chromium has no significant biological role. Although chromium is more Earth abundant, the favourable properties of molybdenum led to it being selected for by biology. One useful property of molybdenum is that it is redox active under physiological conditions and can access +IV, +V and +VI oxidation states with relative ease enabling both one- and two-electron redox reactivity to occur.[6]

A key factor to consider when analysing the biological usefulness of 3d and 4d metals is their ligand exchange rates, which inform how kinetically active the metal is **Figure 1.1**.[7] Understanding the exchange kinetics of inorganic complexes was pioneered by Henry Taube in the 1950s.[8] The strength of metal-ligand coordination bonds and steric effects greatly influence the kinetic stability of a complex; if the coordination bonds are strong than the complex will be slower to exchange ligands. Another key factor in ligand exchange rates of transition metal complexes are the changes in ligand field stabilisation energy (LFSE) moving from the reactant species to the transition state.

To generalise, ligand exchange rates are generally faster for M^{2+} ions than for M^{3+} ions as there is a stronger electrostatic attraction between the positively charged metal and the ligand electrons. Additionally, the 4d and 5d transition metals are much more kinetically inert than 3d metals due to a better orbital overlap between the diffuse 4d and 5d orbitals and the ligand orbitals, leading to stronger metal-ligand bonds. This is not however, the case when comparing chromium and molybdenum, as chromium, in its most stable +III oxidation state, forms strong bonds to ligands and has very slow ligand exchange rates, as shown in **Figure 1.1**.

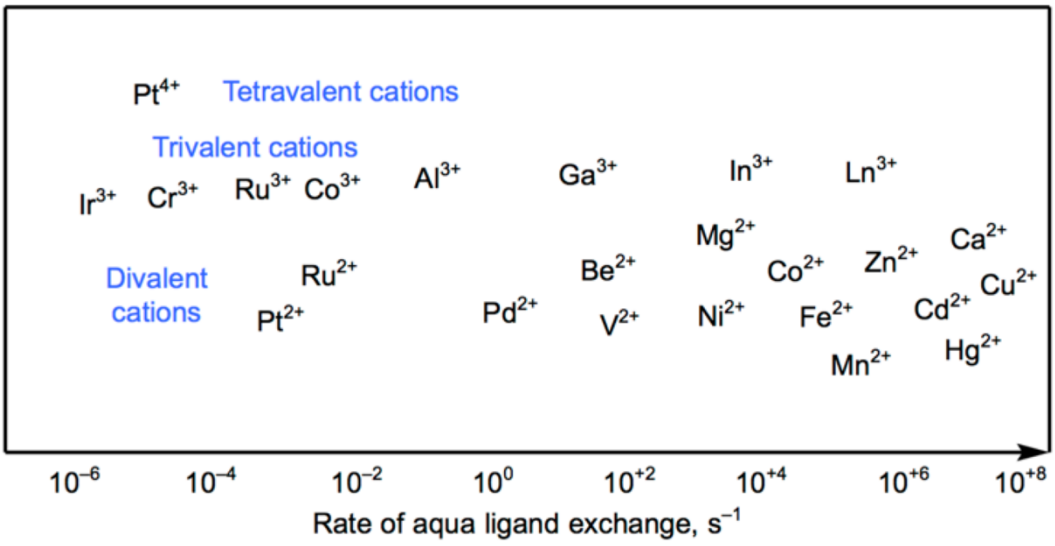


Figure 1.1: Schematic representation of relative kinetics for the aquation of a variety of metal ions. Figure developed from reference (7).

Biology has evolved unique methods of manipulating the fast ligand exchange rates of certain first row transition metal ions used in catalysis. The metal-ligand exchange rates of Fe^{2+} are greatly reduced when bound to multidentate chelating ligands, e.g. porphyrins. However, in iron porphyrins the axial ligands, which are not part of the chelate ring, maintain kinetic lability enabling ligand exchange and catalysis to occur. Haem, or iron protoporphyrin IX is a metal-containing cofactor found ubiquitously in nature.

1.1.2. Biological Chemistry of Iron

Iron is the most important transition metal to all life and performs a diverse number of catalytic roles. Iron uptake, distribution and storage is tightly regulated in cells as too much free iron leads to oxidative damage to proteins, lipids and nucleic acids, whereas insufficient iron impedes the function of many iron-utilising proteins. In aqueous chemistry at pH 7.0 the redox potential restricts iron to the +III oxidation state, which precipitates as insoluble $Fe(OH)_3$; $[Fe^{3+}]$ cannot exceed 10^{-18} M, which explains the low concentration (0.05 ppb) of iron in sea water. Nature has found ways of overcoming this chemical limitation with iron; there is an evolutionary pressure to survive, therefore the element can be extracted to remarkable extents which is highlighted by a maximum concentration of iron in marine organisms of 86,000 ppb.[9]

The ability of iron to gain and lose electrons, shuttling between Fe^{2+} and Fe^{3+} is crucial to its widespread participation in a variety of biochemical reactions. The redox potential of iron can be modulated *via* the binding of different ligands and the careful manipulation of different ligands can tune the different reactivity of iron. Iron has catalytic roles in three main forms in biology; (i) as haem **Figure 1.2 (left)**, (ii) as iron-sulphur prosthetic groups, **Figure 1.2 (middle)** and (iii) as iron-oxo clusters **Figure 1.2 (right)**.[10]

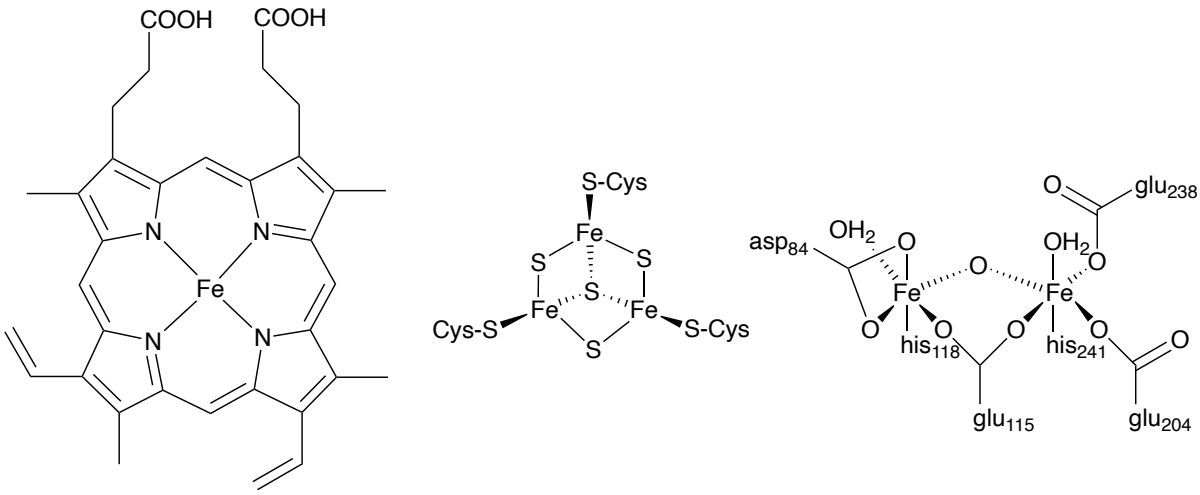


Figure 1.2: (left): The structure of haem, iron protoporphyrin IX, a cyclic tetrapyrrole with an iron centre. **(middle):** A non-planar $[Fe_3S_4]$ cluster **(right):** A carboxylate-bridged diiron unit in oxidised R2 protein of ribonucleotide reductase.

1.1.3. A Chemist's Comparison of Iron and Ruthenium

Iron has an atomic number of 26 with four stable isotopes and a ground state configuration of [Ar] $3d^6 4s^2$. In the previous section the most stable oxidation states have been identified as +II and +III, with Fe(IV) existing as a highly reactive and oxidising species. For Fe(II) and Fe(III) the balance between high spin and low spin complexes is at its most delicate, as complexes with weak field ligands (e.g. H_2O and Cl) are high spin, whereas strong field ligands (e.g. CN^- and CO) are low spin. Iron has been used in many industrial processes, most famously in the Haber-Bosch process as a heterogeneous catalyst for the fixation of hydrogen and nitrogen to produce ammonia. Furthermore, as discussed in the previous section, molecular iron catalysts, both biomimetic haem and non-haem complexes are being explored as they offer a sustainable catalytic source.

Below iron in the periodic table is ruthenium, an element first discovered by Karl Ernst Claus in 1844. Ruthenium has an atomic number of 44, with seven major isotopes, and a ground state electronic configuration of [Kr]$4d^7 5s^1$. The most stable oxidation states of ruthenium are +II, +III and +IV, and ruthenium has an extensive coordination chemistry in multiple oxidation states. In solution the most common oxidation state is +III, however, with different ligands attached the +II, and +IV can easily be preferred. Ru(VIII), the group oxidation state, can also be achieved in the tetrahedral compound RuO_4, whereas for iron, the ionisation energy required to reach group oxidation state is not offset by the enthalpic gain of bond formation.

1.2. What Happens when Biology Interacts with 4d and 5d Transition Metals?

There is a common misconception that heavy, non-essential 4d and 5d metals are more toxic than the metals used by biology. The term 'toxic heavy metals' should be avoided without careful assessment. In fact, the concentration of the essential metals, e.g. Fe is very tightly regulated and if found in excess these metals can be just as toxic as their heavier counterparts in biology. There are many different factors to consider when assessing the toxicity of metals as the biological properties differ greatly between the atoms and the coordinated compounds of that element. Of particular importance are the solubility, oxidation state and bioavailability of the elements in question, for example, Ni(II) and Zn(II) chlorides demonstrate moderate toxicity, whereas the oxide compounds are non-toxic.[11] Therefore, when referring to a metal's toxicity careful consideration of the specific compound is needed, as is the test used to measure the toxicity (e.g. an IC_{50} value).[12]

Dispelling the myth that all heavy metal compounds are toxic, the effects of introducing 4d and 5d metals into biological systems have been widely studied. Many organometallic compounds of 4d and 5d metals are ignored by cells in low concentrations. The term bio-orthogonal chemistry refers to a chemical reaction that neither interacts nor interferes with biological systems, therefore many 4d and 5d organometallic complexes can be referred to as bio-orthogonal.

1.2.1. Metals in Medicine, What has been Learnt?

Although the previous section states that not all 4d and 5d organometallic compounds are toxic, there are many 4d and 5d organometallic compounds that show high toxicity to certain cells, even at very low concentration. Metal complexes for medicinal chemistry applications have been extensively studied due to their ease of functionality and ability to bring about cell death. In particular, this field has been dominated by the development of platinum based compounds, which began when Barnett Rosenberg discovered that cisplatin (cis-diamminedichloroplatinum(II)) inhibited growth of *E.coli* cells.[13] These platinum-based compounds have found particular application in the treatment of a variety of different cancers. Cancer is a disease where certain cells begin to replicate uncontrollably; these cells invade and destroy healthy tissues and organs. Some metal complexes, often called metallodrugs e.g. cisplatin, bind to DNA and accumulate in rapidly dividing cells, leading to cell cycle arrest and stopping a cancerous tumour from forming.

In order to develop an effective anti-cancer metallodrug there are some key challenges: (i) first the drug must be transported to the cancerous region, (ii) then, preferentially, the drug specifically targets the cancerous cells and is transported into the cellular matrix, (iii) finally the drug binds to a specific cellular target that induces the phenotypic response of apoptosis. Although cisplatin is a very simple coordination compound, it triggers an apoptotic response by binding to DNA and inhibiting replication. Certain cancerous cell lines are particularly vulnerable to cisplatin (testicular, bladder and ovarian) as they do not have the cellular machinery capable of repairing the DNA damage caused by cisplatin. However, cisplatin is not a selective drug, when injected into the blood stream it is distributed all over the body and kills healthy cells as well as cancerous cells leading to aggressive side effects.[14]

Alternative metal complexes are being explored as potential drugs, particularly 4d and 5d organometallic complexes, incorporating Ru, Au, Os, Re etc.. There is extensive research, particularly in the field of antibody drug conjugates, looking to target toxic payloads (e.g. metal drugs) specifically to cancerous cells, which would reduce any side effects associated with non-selectivity. Of particular interest in the work contained in this thesis was to understand the speciation of ruthenium complexes once they enter a cell, in order to try and identify specific cellular targets.

Several ruthenium-based drugs have entered clinical trials, however, no lead compounds have made it to use in a clinical setting. One clear stumbling block to advancement in the field of ruthenium metallodrugs is the lack of understanding of the cellular speciation of these complexes. If the link between cellular binding site and apoptotic response can be made, this will stimulate a more rational design to metallodrugs. **Section 1.3** outlines the advances that have been made in the field of ruthenium metallo-pharmaceuticals, with particular focus on understanding the cellular speciation.

1.3. Ruthenium Complexes Interacting with Biology for Medicinal Application

1.3.1. Organoruthenium Anti-Cancer Chemistry

The development of ruthenium anti-cancer complexes is closely related to the worldwide clinical use of cisplatin and other platinum therapies. Interest has grown in ruthenium complexes due to their apparent difference in mechanism of action to platinum complexes; many complexes are perceived to have reduced side effects and been shown to be active against cisplatin-resistant cell lines.

In recent years, this group's research has been focused on developing a fundamental understanding of how ruthenium complexes coordinate to proteins of therapeutic relevance.[15] Previous work in the group has enhanced our understanding of the speciation of ruthenium complexes in biological mixtures. This section gives a brief history of ruthenium complexes as anticancer drugs and the major limitations and advances that need to be made for them to be used in a clinical setting.

Metal Based Drugs vs Organic Drugs

Small, organic drugs traditionally work by forming specific intermolecular interactions with a biological target thereby altering its activity. Intermolecular forces in combination can, however, be matched, and often bettered in energy by one metal-ligand coordination bond, which can range in energy between 50-350 kJ mol^{-1}. For example, the well-known anti-cancer drug methotrexate makes a number of intermolecular contacts with the target protein DHFR (dihydrofolate reductase) that contribute to an enthalpy of binding of 55.6 kJ mol^{-1}.[16] This is matched by the enthalpy of substitution of a single ruthenium-thioether bond, 57.3 kJ mol^{-1}, measured through calorimetric analysis by Taube et al..[17] Furthermore, small molecule drugs have carbon skeletons, which generally restrict the geometries of compounds to linear, trigonal and tetrahedral. Metal ions on the other hand can access coordination numbers between two and ten, allowing for greater structural diversity, which may be beneficial for drug design.[18] Finally the variable oxidation states of transition metal ions has been utilised for the development of redox-activatable drug candidates that are activated in the reducing environments of cancer cells.[19]

What Features of Ruthenium Make it an Attractive Element to Use?

Ruthenium compounds have shown excellent activity against cancer cell lines and ruthenium is a promising metal for medicinal use for the following reasons:

1. Ruthenium complexes are proposed to act through biochemical mechanisms that are different to the DNA cross-linking mechanism of platinum complexes, therefore they show activity against cisplatin resistant tumours.[20]
2. Ruthenium complexes bind to Lewis basic residues, which are plentiful in cells, to form strong metal-ligand coordination bonds.[21]
3. The large enthalpy of the metal ligand bonds means ruthenium complexes have relatively slow ligand exchange rates which means that once a complex is bound to its target its dissociation rate will be low and it will stay bound for the lifetime of the cell.

Importantly, ruthenium exchange rates are not too slow so that no ligand exchange is possible as often seen for 3rd row transition metal complexes.

4. The reactivity and physical properties of a given ruthenium complex can be tuned for purpose through incorporation of carefully designed ligands, For example, certain ligands can make complexes more air and water stable.[7,22]
5. Ruthenium complexes are bio-orthogonal, meaning they are active inside a biological environment, but are not highly regulated or interfere with native biochemical processes.
6. In terms of practicality, ruthenium is also an attractive metal to use due to its affordability (in comparison to platinum) and its good air and moisture stability.

Two ruthenium(III) complexes have been evaluated in clinical trials, NAMI-A and KP1019/NKP-1339, **Figure 1.3**. NAMI-A shows excellent selectivity against solid metastatic tumours, however, this compound did not make it past phase I clinical trials as at a high dose the patients developed blisters on hands, fingers and toes.[23] KP1019, and its sodium salt NKP-1339, showed very promising activity and limited adverse side effects throughout phase I clinical trials, however, it was concluded a that deeper understanding of the interactions of KP1019/NKP-1339 within a cell was required before clinical use.[24] One ruthenium(II) complex has also entered clinical trials, TLD1433, which contains large aromatic ligands and acts as a photosensitizer in photodynamic therapy, however this complex is not designed to form coordination bonds to biological molecules.[25]

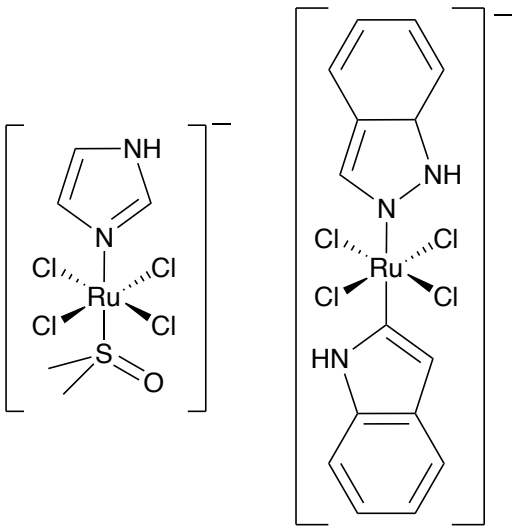


Figure 1.3: The two ruthenium anticancer complexes which have reached clinical trials, NAMI-A **(left)** and KP1019 **(right)**. The counter ions have been removed for clarity, however, NAMI-A has a sodium cation and KP1019 has a protonated indazole cation.

It has been proposed that both of these Ru(III) complexes are activated by reduction in a biological environment to a Ru(II) species, and the redox properties play an important role in the drugs' phenotypic effect.[26] Ru(II) species, especially those with an η^6-arene ligand, have also been explored as potential metallodrugs. The metal in these complexes is stabilised in this lower oxidation by the arene's bonding modes to the metal: (i) A bonding σ and π-type interaction from the filled ligand p-orbitals to empty metal d-orbitals and (ii) a π-back-bonding interaction between the filled d_{xz} and d_{yz} orbitals, and the empty p-antibonding orbitals. The ability to synthetically tune the ligands around the metal centre has led to a large number of Ru(II) compounds being screened to find the most efficacious drug, **Figure 1.4**.[27–29]

In comparison with Ru(III) complexes, the Ru(II)(η^6-arene) bioactive scaffold appears to lead to increased control over the biomolecular targets of ruthenium complexes. RAPTA-type complexes, [Ru(arene)(PTA)X_2] (PTA = 1,3,5-triaza-7-phosphaadamantane) preferentially bind to proteins[30,31] whereas smaller RAED-type complexes with the general formula [Ru(arene)(en)Cl] (en = 1,2-ethylenediamine) preferentially bind to DNA; this binding can be enhanced through extended π-systems.[32,33] Searching for cellular protein targets has been performed using an integrated proteomics-based target-response profiling approach, where proteins are incubated with cellular lysate and metallodrug-protein adducts are separated in pull down assays. This has identified a preference for a Ru(arene) pyridinecarbothioamide complex to target selectively the scaffold protein plectin which regulates keratin and tubulin networks.[34] These methods are, however, time consuming, and the link between plectin binding and phenotypic response needs further research. Although some cellular targets have been identified, most Ru(II) arene complexes are still highly promiscuous and there is little insight into their complete cellular speciation.

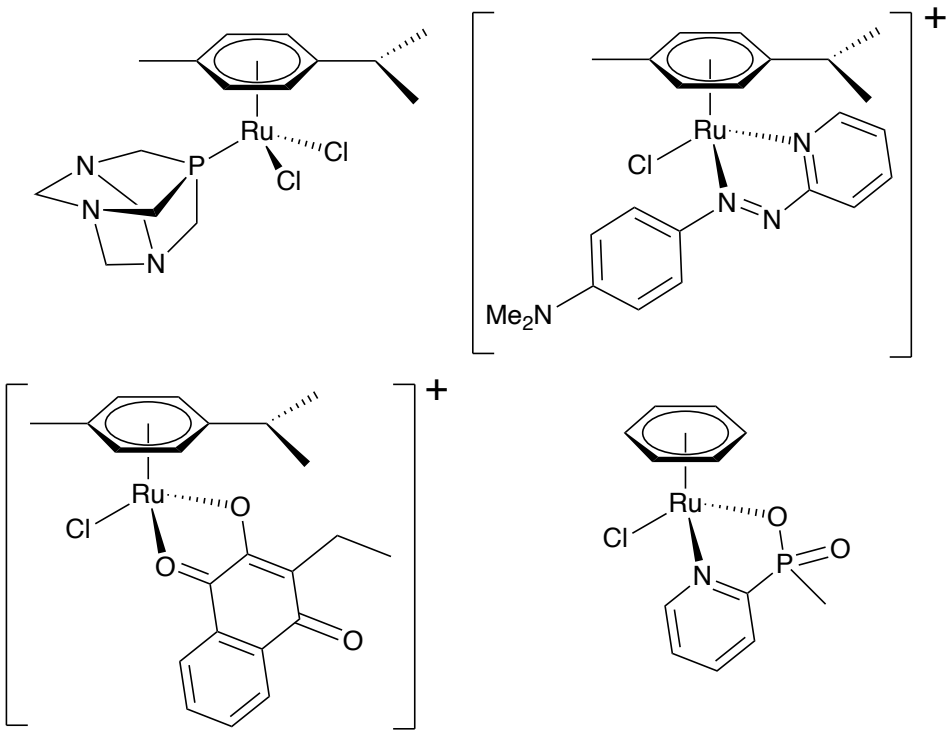


Figure 1.4: Chemical structures of a selection of Ru(II) arene compounds with anti-proliferative properties.

1.3.2. *In vivo* Ruthenium Catalysis with Therapeutic Applications

Many research groups have also postulated that organometallic catalysts at low concentrations in cellular environments can bring about an apoptotic response, either through directly perturbing the cellular redox potential or through the uncaging of an inactive prodrug in a catalytic cycle. These methods apply the well understood catalytic reactivity of platinum group metals to the field of medicinal chemistry, and potentially allow for very low concentrations of metal to be administered.

Firstly, Sadler et al. reported on the use of a Noyori-type ruthenium hydride transfer catalyst that when co-administered with a hydride source e.g. formate, could reduce NAD^+ to NADH, and interfere with redox processes within the cell that is hoped to lead to cell death.[35] This methodology, although intriguing, has some clear limitations, most obviously, that the redox state of a cell is so highly regulated that large amounts of formate are required and significant catalysis would be necessary for effect to be observed.

Secondly, metal-catalysed uncaging reactions have drawn recent interest to activate a prodrug at a desired therapeutic area.[36,37] This has shown particular promise in the field of antibody-drug conjugates (ADCs), where an antibody is connected to a drug *via* a bio-orthogonally cleavable bifunctional linker. When the drug has accumulated at the target area, a metal catalyst

is administered to cleave the linker and release the drug payload. Ruthenium complexes have been used for this purpose as shown by Meggers et al. who reported a ruthenium piano-stool complex that deprotected the organic drug doxorubicin in HeLa cells resulting in apoptosis.[38]

1.3.3. Factors Limiting Advances in Ruthenium Medicinal Chemistry

To move this field forward to be able to effectively design drugs it is necessary to determine how metal adducts are getting to cancerous cells and which metal adducts, generated inside a cell, lead to the observed therapeutic response. It is highly possible that the cellular cytotoxicity of ruthenium complexes is a result of binding to multiple cellular targets, however, through understanding the site preferences and the mechanisms of action of the metal, it might be possible to selectively target weak points in a cancerous cell.[32]

Understanding the biomolecular targets and cellular speciation of ruthenium complexes in a cellular environment remains a difficult task, but through the use of modern analytical techniques and a more systematic approach than has been used to date, significant progress can be made. The dynamic nature of the cellular concentrations and accessibility of biomolecules combined with the characteristically slow ligand exchange rates associated with Ru(II) arene complexes, makes understanding the relationship between speciation of the metal complexes and cellular response problematic. A direct read out of what the ruthenium compounds are bound to in biological environments remains challenging and has so far been restricted to methods which do not report the dynamic speciation in solution, such as mass spectrometry[39–41] and X-ray spectroscopic methods.[42–44] This research looks to address this issue, and develop new methods to track the speciation of ruthenium organometallic compounds.

1.4. Catalysis Involving the Group 8 Metals Iron and Ruthenium

1.4.1. Biological Catalysis Using Haem Enzymes

In order for a protein to harness the catalytic potential of haem, the protein must be able recognise and accommodate haem. Referring to **Figure 1.2 (left)**, the haem cofactor contains a large conjugated π-system, a metal ion and ionic functional groups; a protein must be able to accommodate all three of these functionalities. Proteins do this through non-specific π-stacking of the porphyrin, hydrophobic interactions with the propanoate groups and coordinating axial ligands to the metal, usually methionine, histidine or sometimes cysteine. There are two axial positions available for coordination to haem; some haem proteins, particularly those involved in electron transfer supply two ligands, and other haem proteins supply only one, enabling small molecules (e.g. CO and O_2) to coordinate to the metal centre.

A large group of haem proteins called monooxygenases and dioxygenases are required in biological systems for the oxidation of very resistant, inert molecules (e.g. simple aliphatics). These proteins manipulate dioxygen as an axial ligand, and form very reactive oxidising species. The cytochromes P450 family of enzymes are monooxygenases which generally catalyse the hydroxylation of inert R-H organics to R-OH. In cytochromes P450 the general mechanism is as follows (i) binding of the organic substrate RH to the Fe(III) enzyme; (ii) reduction of iron to high-spin Fe(II); (iii) O_2 binding to give a low-spin Fe(II) species or a superoxide bound low spin Fe(III) species; (iv) reduction of the Fe(II)-O_2 unit to give a short lived peroxo group which is rapidly protonated releasing one water molecule and leaving a reactive Fe(IV) oxo species with a radical delocalised over the porphyrin; (v) attack of the Fe(IV) oxo species on RH to give ROH.

Chemists have long tried to design synthetic metal complexes which mimic the roles of iron in biology. The purpose was initially to establish how natural systems have evolved their catalytic activity, and to search for inspiration into how to develop new synthetic catalysts. In the pioneering work of Collman et al. in the 1970s, a synthetic [Fe(picket-fence porphyrin)] complex was developed which mimicked dioxygen coordination to the respiratory proteins haemoglobin (Hb) and myoglobin (Mb).[45,46] This model helped confirm that dioxygen coordination to iron in Hb and Mb, occurs in an angular end-on fashion as postulated by Pauling,[47] and there is a formal change in oxidation state from Fe(II) to Fe(III) resulting in a reduction in Fe ionic radius bringing the Fe atom into the haem plane. These Fe picket-fence porphyrin complexes still inspire the development of new synthetic catalysts. Dey and co-workers report the use of a synthetic cytochrome P450 mimic, which is capable of the catalytic hydroxylation of C–H bonds to alcohols and epoxidation of alkenes using molecular O_2 in water.[48]

1.4.2. Synthetic Catalysis Using Molecular Ruthenium Complexes

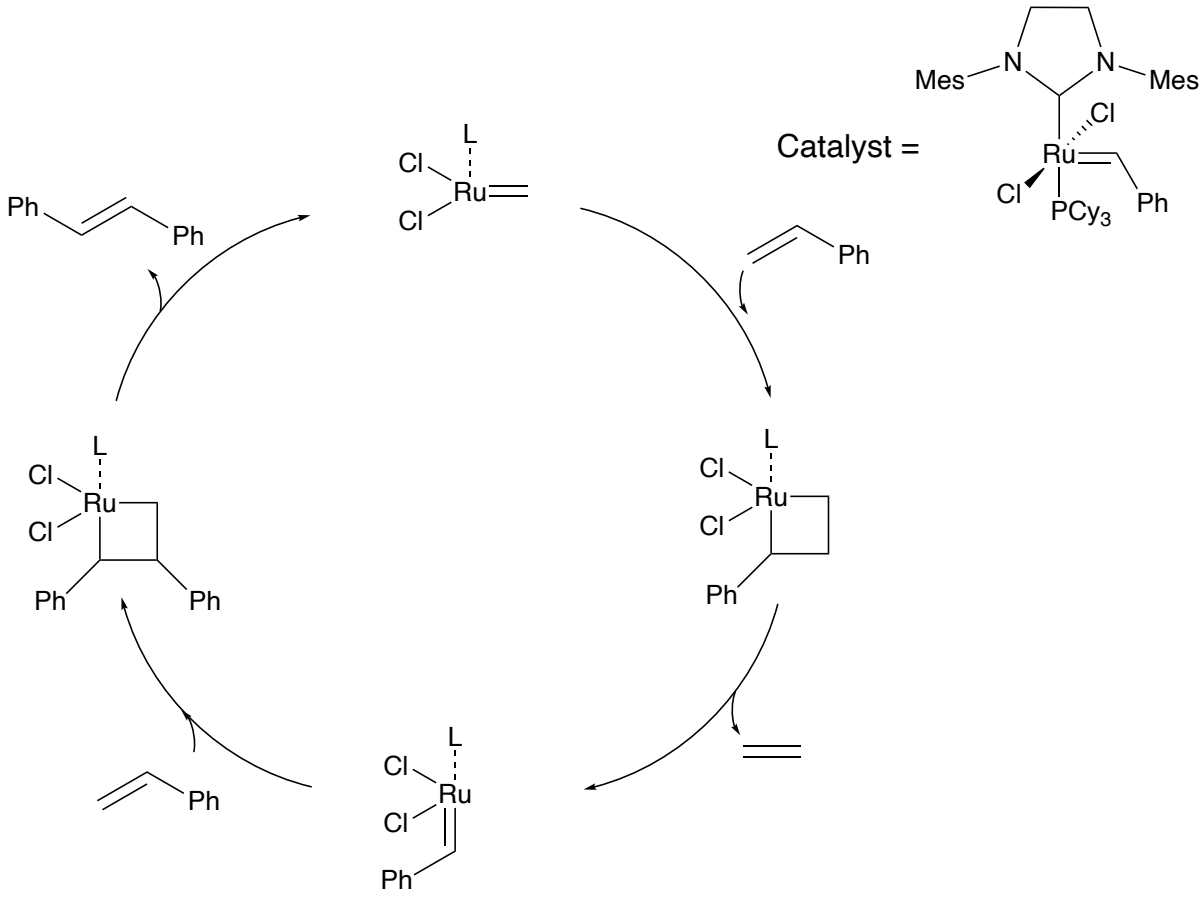
The different accessible oxidation states of ruthenium mean that organometallic compounds of the metal have been studied extensively in homogeneous small molecule catalysis. This section will focus upon some key chemical transformations catalysed by organometallic ruthenium complexes.

Olefin Metathesis

The 2005 Nobel Prize in Chemistry was awarded to Grubbs, Shrock and Chauvin for the development of the metathesis method in organic synthesis.[49] Olefin metathesis involves the

redistribution of alkene fragments *via* the breaking and reforming of C=C double bonds, **Scheme 1.1**. It is applied in the synthesis of a number of high value chemicals, including pharmaceuticals, and is used widely in polymerisation processes, *via* ROMP (ring opening metathesis polymerisation). Although a number of different transition metals have been tested as metathesis catalysts (titanium, molybdenum and tungsten), the most stable and functional group tolerant catalysts are ruthenium based. In the early 2000s, ruthenium complexes with N-heterocyclic carbene ligands were proven to be excellent metathesis catalysts, exhibiting greater metathesis activity, solvent and thermal stability than previous ruthenium phosphane complexes.[50]

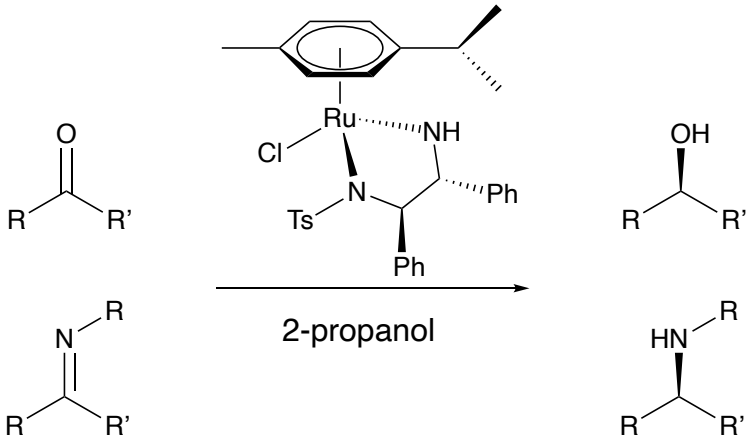
There are no known natural metalloenzymes capable of metathesis activity. It could be argued that biology has found alternative methods of forming C–C bonds, and that substrates involved in metathesis (specifically terminal alkenes) were never part of metabolism. However, the lack of biologically-catalysed metathesis reactions could also be partially understood by looking at the requirement for highly electron donating ligands on the ruthenium to stabilise the Ru(II) and Ru(IV) oxidation states present in the reaction mechanism. None of the natural, un-modified 20 amino acid ligands have comparable electronic donating properties to an N-heterocyclic carbene or a phosphane.



Scheme 1.1: Mechanism for olefin metathesis catalysed by Grubbs 2nd generation ruthenium catalyst.

Transfer Hydrogenation

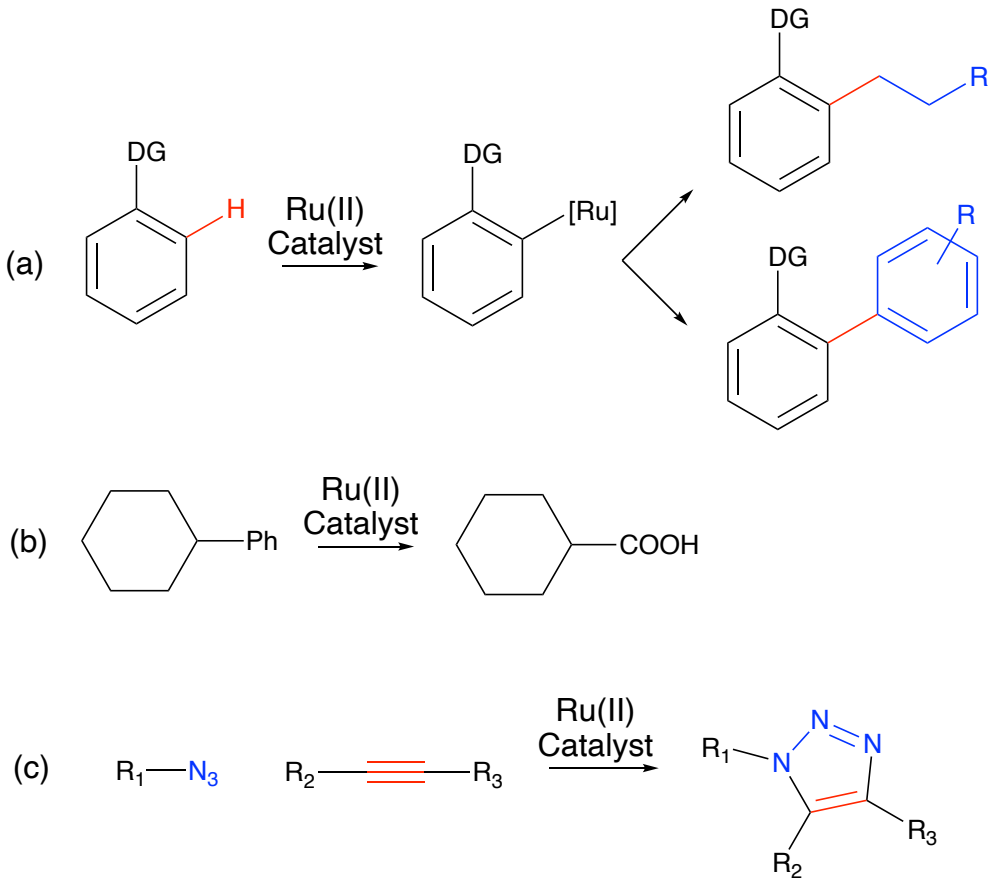
The 2001 Nobel Prize in Chemistry was part awarded to Knowles and Noyori for the development of hydrogenation reactions. In particular, Noyori and co-workers pioneered ruthenium catalysed asymmetric transfer hydrogenation. Asymmetric transfer hydrogenation (ATH) is the transfer of hydrogen atom to a molecule from a source other than gaseous H_2, usually isopropanol (IPA), with excellent stereochemical control. Ru(II)(η^6-arene) complexes bearing a substituted ethylenediamine ligand with two chiral centres, are effective asymmetric catalysts of ketones and imines to alcohols and amines respectively, **Scheme 1.2**.[51] Biological systems perform a number of hydride transfers and hydrogenations. One family of proteins that perform these transformations are dehydrogenases, often utilising the cofactor NADH.



Scheme 1.2: Asymmetric transfer hydrogenation of a ketone to an alcohol, and an imine to an amine catalysed by Noyori's ruthenium catalyst.

Ruthenium Catalysed Reactions

There are many catalytic transformations catalysed by ruthenium complexes, including C–H functionalisation,[52] oxidation reactions[53] and cycloadditions,[54] summarised by **Scheme 1.3**. All these reactions show the versatility of small-molecule ruthenium catalysts. Fundamental to this reactivity is the metal's ability to access a range of different oxidation states and coordinate varied substrates.



Scheme 1.3: (a) Ruthenium catalysed C–H functionalisation. (b) Ruthenium catalysed oxidation. (c) Ruthenium catalysed azide alkyne cycloaddition.

1.4.3. A New Approach to Catalysis – Combining Transition Metals and Enzymes

The coordination chemistry of many transition metal complexes has been explored with biological molecules, specifically proteins, to explore the catalytic potential of these protein-metal hybrids. This work has developed into a thriving area of research where these protein-metal hybrids, capable of catalysis are called artificial metalloenzymes (ArMs).

1.5. Artificial Metalloenzymes

Artificial metalloenzymes are synthetic biocatalysts that result from either the combination of an artificial metallocofactor being introduced into a protein scaffold or a natural metalloprotein with an unnatural/evolved protein scaffold. These protein-metal hybrids combine the catalytic features of natural enzymes and versatile reactivity toolbox of transition metal catalysts. As discussed, enzymes have excellent catalytic properties, they combine rapid reaction kinetics, high turnover numbers, substrate selectivity and product stereo-control all in relatively mild reaction conditions (water as a solvent and mild temperatures).

Transition metal catalysts are capable of catalysing a wide range of chemical transformations. These metal complexes have ligands which can be tuned for a specific reaction and many of these transformations have well understood catalytic mechanisms. Combining the advantageous features of enzymes and transition metal catalysts into ArMs has emerged as a powerful tool in synthetic biology showing promising results in the sustainable production of commodity chemicals and new-to-nature metabolites.[55] Replacing synthetic catalysts, acting on petrochemical feedstocks in non-aqueous solvents, with biocatalytic systems working in water with simple carbon neutral feedstocks is clearly highly desirable. But why engineer new enzymes, particularly using expensive and relatively scarce transition metals, when the ability to find new catalysts amongst gene products from all corners of the biological world has developed at staggering pace?[56–58] As a consequence of the latter, any target chemical can conceivably be obtained by recombining pre-existing metabolic pathways.[59]

One clear feature is orthogonality – the objective of introducing functionality into a cell that has no counterpart in the natural world, could provide reactivity that biology cannot currently catalyse – alkene metathesis for example. As there is a limit to the number of additional transformations a viable cell will perform, these orthogonal reactions may allow access to much shorter, and therefore more efficient, pathways. If not for a synthetic purpose, one could also imagine orthogonal catalytic chemistry providing a diagnostic or reporter output without interference from the host endogenous processes. For it to be truly orthogonal, it is difficult to imagine evolving a new enzyme based around metals already abundant in nature and already used as catalysts in biology. The transition metals used by nature are very carefully controlled by acquisition and regulatory networks that ensure catalytic metal ions are not free to operate outside the endogenous metabolism. Therefore, there is significant advantage in trying to introduce metals that biology currently has no evolved means of metabolising.

A key factor in the potential of ArMs is that the protein scaffold is genetically encoded, therefore, the artificial metalloenzyme can be evolved through laboratory evolution techniques and catalytic performance can be optimised. The field of ArMs has been advanced greatly through the development of directed evolution.

1.5.1. Directed Evolution of Metalloenzymes

Darwinian evolution schemes can be fast-tracked in the laboratory *via* directed evolution which has revolutionised biosynthesis and biotechnology. Built upon the ground-breaking discoveries of George Smith and Greg Winter on phage display, which enabled protein libraries of slightly

mutated proteins to be screened for target affinity,[60–62] Frances Arnold has developed and applied directed evolution to evolve natural enzymes into different catalysts.[63,64]

Directed evolution is a method where the gene for the protein is subjected to iterative rounds of mutagenesis creating a library of gene variants. This gene library must be expressed as a protein library in a manner which couples the expression products (the different protein variants) and genetic sequence information (the nucleic acid), i.e. a linkage between the genotype and phenotype. Following this, the protein library is screened for catalysis *via* an activity assay, with the active variants then selected and the DNA recovered. Multiple rounds of directed evolution can lead to an optimisation in many different features of the enzyme, for example, stability and catalytic activity. In directed evolution of ArMs with an unnatural metal cofactor, there is an additional step, in which each individual member of a protein library is exposed to the metal cofactor, **Figure 1.5**.

The choice of starting point for such a forced evolution campaign, which in the case of ArMs is the protein scaffold of choice and a metal complex, is of great importance. Since any particular enzyme follows a unique evolutionary trajectory as new mutations move it along the fitness landscape towards (potentially local) maxima, choice of the starting point may directly predetermine the result. By nature of the selection process, it is further possible, that trajectories leading to the global maximum fall beneath the cut-off limit, becoming inaccessible.

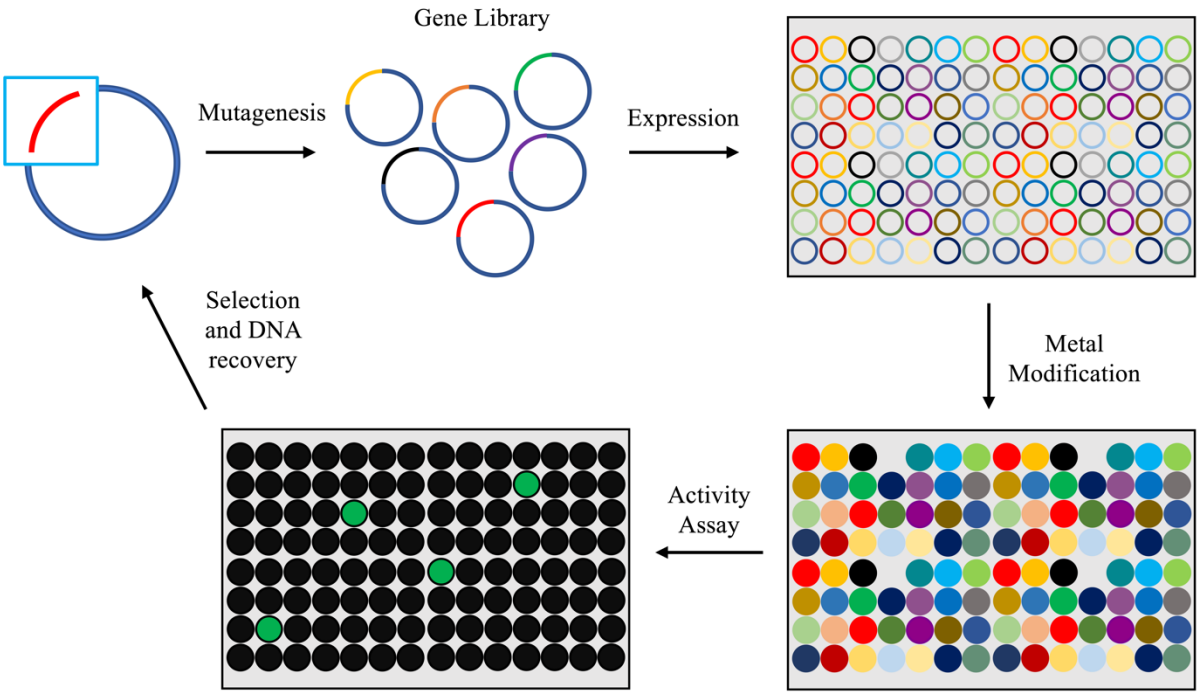


Figure 1.5: The general overview of a directed evolution campaign for artificial metalloenzymes. The Darwinian algorithm can be reproduced in the laboratory, greatly increasing the speed of evolution. Mutagenesis methods introduce mutations with various levels of randomness, depending on the method used, to the starting point gene, forming a gene library. This library can then be expressed in a manner that couples expression products and genetic sequence information to yield the different proteins. Upon addition of the metal cofactor, the artificial metalloenzymes are formed and can be selected for improved variants in regard to desired parameters (reaction rates, yield, stereoselectivity, stability etc.). The sequence information of the improved candidates is recovered and can be subjected to further rounds of directed evolution.

There are three key practical considerations in the directed evolution of ArMs: (i) Expressing protein variants with a genotype-phenotype linkage, (ii) a clean and efficient metal modification step and (iii) a quick, sensitive and robust activity assay.

First, crucial to any high-throughput directed evolution campaign is a genotype-phenotype linkage. Nature links genotype and phenotype by compartmentalising genes in cells. *In-vivo* directed evolution experiments have generally employed colony screening or cell surface display.[65] In this methodology, each well plate contains cells expressing different variants of a protein and catalytic activity is assayed. *In-vitro* techniques use the desired gene and subsequently expressed protein outside of the cell and make use of readily available *in-vitro* translation systems originally from *E.coli* or yeast cells. Upon *in-vitro* expression a linkage, often covalent, must be established between the genotype and phenotype, or by artificial compartmentalisation, i.e. within a droplet.

Second, a clean and efficient metal modification step is required to ensure that any activity measured arises from the protein-metal hybrid and not from free metal complex. Performing this step in complex biological environments is particularly challenging as the cell has many reactive species that are potentially metal poisons, (e.g. glutathione) which will disrupt protein-metal hybrid formation. The following section will explore different methodologies developed for efficient protein-metal modification.

Finally, a quick, sensitive and robust activity assay is required to test a large number of ArM variants. Fluorescence-based assays are commonly used in high throughput screening and directed evolution due to their high sensitivity, variety of potential fluorophores and ease of operation.

1.5.2. Introducing Unnatural Cofactors into Protein Scaffolds

Four successful strategies have been employed to generate artificial metalloenzymes with a transition metal cofactor in a well-defined location in a host protein: 1) Metal substitution, 2) covalent anchoring, 3) supramolecular assembly, and 4) dative anchoring. This section will give a brief historical overview of artificial metalloenzymes, and then will go through the different anchoring methods giving examples of the most successful and recent ArMs to date.

Historical Overview

The first unnatural protein-metal hybrid was reported by Akabori and co-workers in 1956 where a protein-palladium complex was prepared by adsorption of palladium chloride onto silk fibres. Asymmetric hydrogenation to yield optically-pure amino acid products was achieved.[66] In 1976 Yamamura and Kaiser reported metal substitution of an active-site zinc ion by a copper ion, in carboxypeptidase A, which resulted in the conversion of a powerful hydrolytic catalyst into an oxidase artificial metalloenzyme.[67] Two years later, Wilson and Whitesides reported localising a biotinylated rhodium catalyst, capable of asymmetric hydrogenation, into the protein avidin *via* supramolecular assembly. Biotin and avidin have a very high association which is effectively irreversible, making this a unique way of ensuring that the metal is localised in the protein.[68]

From the early 1980s to the early 2000s there are many examples of research papers involving the development of artificial metalloenzymes, mainly *via* metal substitution. The great limitation of the artificial metalloenzymes generated throughout this time was that they did not match the kinetic prowess, activity and selectivity of natural enzymes. Until the development

and application of directed evolution, the potential of artificial metalloenzymes was not fully realised.

Metal Substitution

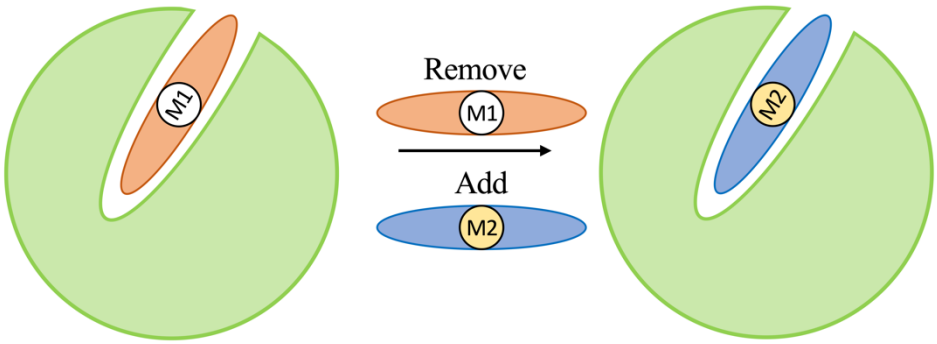
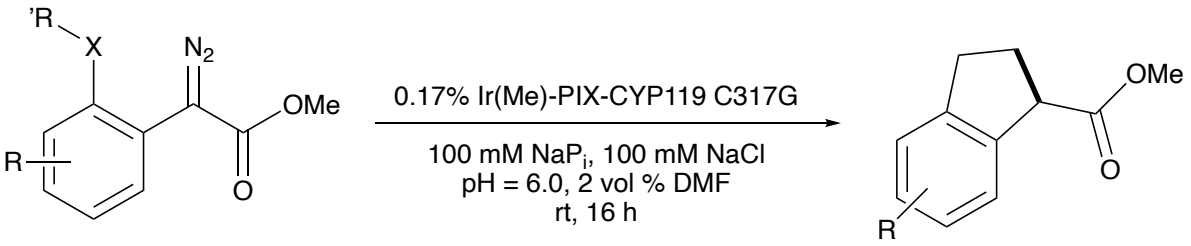


Figure 1.6: Schematic representation of metal substitution. The natural cofactor (red) can be substituted with a suitable unnatural cofactor (blue). This may include the bare metal ion or larger cofactors such as haem.

Many research groups have attempted to monitor how changing an active site metal impacts upon the metalloenzymes catalytic function, **Figure 1.6**. The most elegant method of metal substitution to date, was developed by Hartwig and co-workers, where an iridium methyl cofactor is substituted for the iron in two different haem containing proteins.

In both examples, Ir(Me) units were substituted into a porphyrin in place of iron to catalyse the functionalisation of C–H bonds to C–C bonds by carbene insertion, **Scheme 1.4**. This reactivity was initially shown using modified myoglobins, with many other metal substitutions also tested, including a Ru(CO) unit, which was catalytically inactive.[69] The reaction was then tested after replacing the iron in the variants of the cytochrome P450 enzyme CYP119. This led to a major breakthrough as through directed evolution, artificial iridium metalloenzymes were generated which showed 98% enantiomeric excess, 35,000 turnovers and 2500 hours^{-1} turnover frequency. This confirmed that these ArMs were comparable with natural enzymes, in that they exhibit fast kinetics, high productivity and high selectivity.[70]



Scheme 1.4: Ir(Me) cytochrome P450 variants, capable of catalysing the functionalisation of C–H bonds to C–C bonds *via* carbene transfer.

Covalent Anchoring

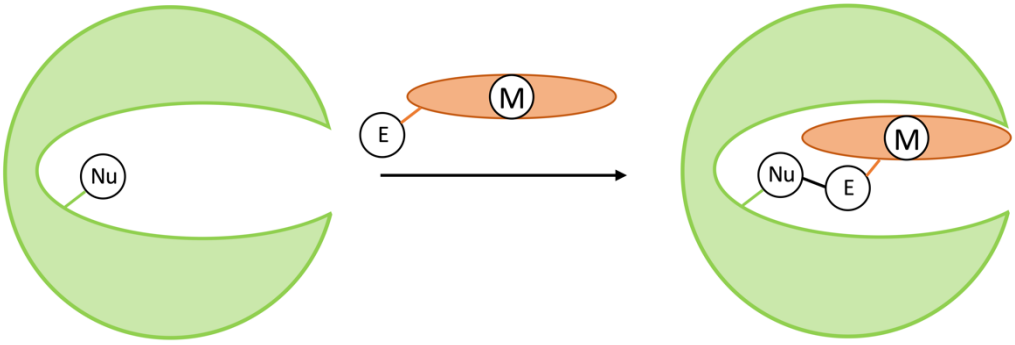
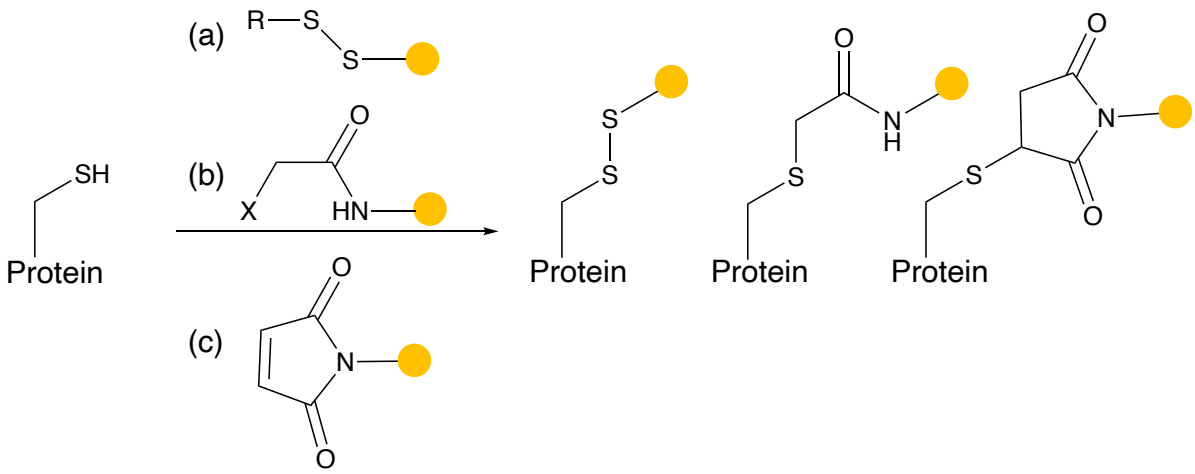


Figure 1.7: Schematic representation of covalent anchoring. The metal cofactor (red) is attached to the protein by a reaction forming a covalent bond, for instance nucleophile (Nu) attacking an electrophile (E).

The second method of localising an abiotic cofactor to a specific location is *via* a covalent linkage, **Figure 1.7**. There is extensive research interest in developing novel bioconjugation and protein modification techniques which fall into two broad categories: (a) modifying the protein template directly, *via* a nucleophilic – electrophilic reaction and (b) genetically manipulating the amino acid sequence before protein modification.[71,72]

Many direct covalent modification reactions exist for the nucleophilic amino acids cysteine, lysine and tyrosine. In particular modification of cysteine has been widely applied for generation of artificial metalloenzymes. Due to the rarity of free cysteine in proteins (most in disulphide bonds and therefore unavailable), and its high nucleophilicity, modifying cysteine is an attractive methodology, as multiple modifications do not occur. Lysine residues are common on the surface of proteins and modification is rarely regioselective. Furthermore, there are many effective and reliable methods to modify a cysteine site, with three highlighted in **Scheme 1.5.**



Scheme 1.5: Classical methods used in cysteine modification of proteins: (a) thiol exchange, (b) alkylation of α-halocarbonyl electrophiles, X = Br or Cl and (c) maleimide Michael addition.

Salmain and co-workers have modified the free Cys25 in the cysteine protease papain, using a variety of Ru(II)(η^6-arene), CpFe, CpRu, $Re(CO)_3$ and Rh(III)(η^6-arene) complexes, all functionalised with either a maleimide or chloroacetamide group.[73,74] The Rh(III)-papain conjugate showed some activity as a transfer hydrogenation catalyst, however, only a moderate enantiomeric excess of 7-10% was measured.[75]

Cysteine residues have also been mutated into a variety of proteins, *via* site directed mutagenesis, and targeted with traditional conjugation techniques. Hayashi et al. reported an ArM with a maleimide functionalised Rh(Cp)(cod) unit covalently attached to a cysteine mutant Q96C of the β-barrel protein aponitrobindin. This hybrid was found to be a catalyst in the polymerisation of phenylacetylene, preferentially yielding the trans polymer.[76]

Cysteine specific modifications can also be made prior to covalent anchoring to a metal complex, for example, cysteine residues can be selectively and simply reduced to dehydroalanine (Dha) which broadens the reactivity profile even further.[77] When Dha has been installed, N, S and P nucleophiles can be covalently added under biocompatible conditions. If these nucleophiles were on the metal ligands, the linkage from the metal to the protein back-bone could be very short, which is advantageous in localising a metal in a protein.[72,78–81]

Genetic manipulation can also be used to introduce unnatural amino acids (UAAs) into a protein scaffold, which can then be covalently attached to the metal cofactor. In 1989, Peter Schultz reported a method to site-specifically incorporate UAAs into the enzyme β-lactamase, using modified transfer RNA responding to a stop codon substituted for the codon encoding the residue of interest.[82]

The most successful ArMs involving a covalent linkage to an UAA have been reported by Jared Lewis *et al.*.[83] This method involves covalently linking an alkyne-substituted dirhodium catalyst to a genetically encoded L-4-azidophenylalanine residue through strain-promoted azide-alkyne cycloaddition.[84,85] In this study, the protein scaffold used was a β-barrel prolyl oligopeptidase from *Pyrococcus furiousus* selected for its cylindrical shape, large hydrophobic pocket for cofactor enclosure and high thermal stability.

This rhodium-POP hybrid catalyses olefin cyclopropanation, and mutagenesis was employed to improve the enantioselectivity of the reaction. The effectiveness of this methodology is that the reaction is fast and irreversible, and the conjugation could be performed in a complex mixture, for example a cell, as the alkyne is selective for the azide. Although mutagenesis has been employed for this ArM, a full directed evolution campaign was not performed. Evolving enzymes with UAAs is complicated, as the UAA could be mutated out, and expression levels for these proteins are often very low.

Supramolecular Assembly

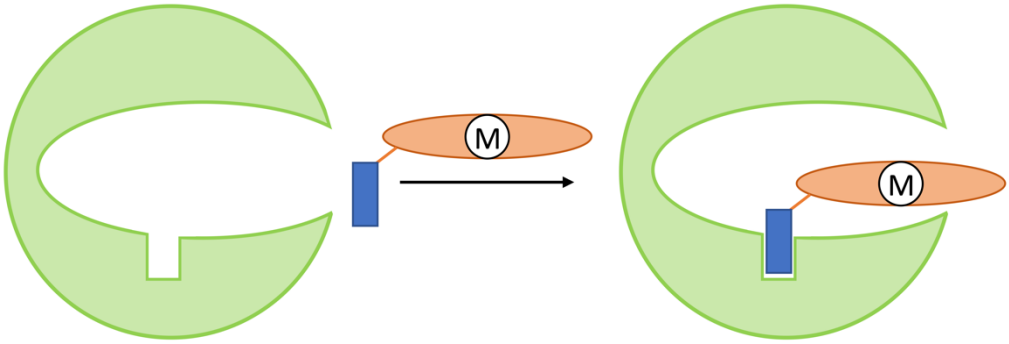


Figure 1.8: Schematic representation of supramolecular assembly**.** The metal cofactor (red) is localised by non-covalent interaction between a ligand bound recognition group (blue) and the protein.

Artificial metalloenzymes have been generated where a catalytic metal complex has been attached to a small molecule with high affinity for a protein target, **Figure 1.8.** Building on the work of Wilson and Whitesides in the 1970s,[86] Thomas Ward and co-workers have assembled ArMs based on the high supramolecular affinity of small molecule biotinylated metal catalysts for the protein streptavidin (Sav). As many as 12 different catalytic transformations have been performed by these metal-streptavidin hybrids, including ruthenium catalysed olefin metathesis,[87] ruthenium-catalysed deallylation,[88] iridium-catalysed transfer hydrogenation[89] and dirhodium-catalysed cyclopropanation,[90] all *in-vivo*.

Where Ward's work has really enhanced the field of ArMs is being able to carry out catalysis in complex cellular environments. Through molecular biology methods, the host protein, streptavidin is expressed by *E.coli* into the periplasm instead of the cytoplasm, so to avoid catalyst poisoning by glutathione. Catalyst reactivity, measured *via* TON and TOF, is often enhanced by having the streptavidin scaffold present in solution, which can form an ArM with the metal catalyst.

This strategy has also been employed in a therapeutic application of ArMs that was reported by Tanaka et al. In this example, a coumarin derivative tagged with a ruthenium metathesis catalyst was localised to a hydrophobic binding site in human serum albumin. The metalloenzyme was directed to cancerous tissue (through specific glycosylation) and a pro-drug was administered which upon metathesis induced cellular death.[91]

One key benefit to supramolecular assembly is apparent in both examples described above, and that is that the conjugation between metal and protein is robust enough to be performed in complex cellular environments. Furthermore, unlike covalent attachment, supramolecular assembly can be a reversible process, which allows for component recycling. In a recent report of Duhme-Klair et al. catalytic transfer hydrogenation is demonstrated from a siderophore–protein combination that enables strong but redox-reversible catalyst anchoring.[92]

Dative Coordination

Dative ArMs have coordination bond(s) directly from the metal to a Lewis basic amino acid residue (His, Cys, Ser, Glu, Asp, etc.) on the enzyme. There are few examples of ArMs with dative attachment presented in the literature most of which rely of 1st row transition metals, **Figure 1.9.**

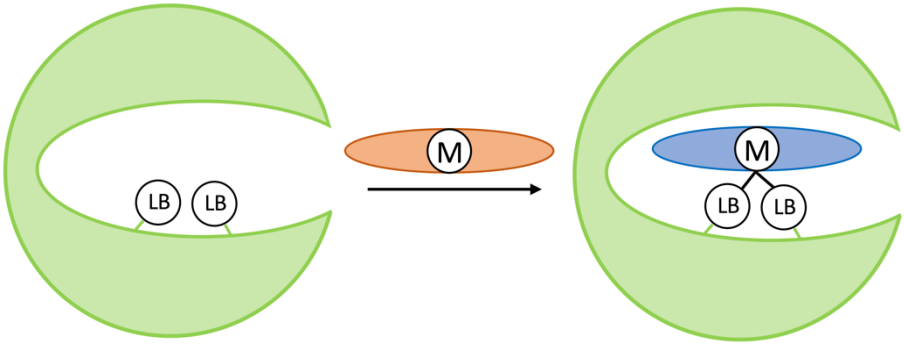


Fig 1.9: Schematic representation of cofactor attachment *via* direct coordination**.** The free metal cofactor (red) attaches to Lewis basic residues on the protein (LB) *via* ligand substitution reactions, forming a new protein-metal complex (blue).

One methodology in forming datively attached ArMs is to remove a prosthetic group e.g. haem and replace it with a metal complex which is structurally similar. This has been achieved with cobalt porphyrins, in the haem proteins myoglobin and cytochrome b_{562}.[93,94] Watanabe and co-workers have applied this method in generating artificial myoglobin metalloenzymes, where a chromium Schiff-base is attached to the same His93 as the original haem group. These cobalt metalloenzymes are capable of catalysing the enantioselective sulfoxidation of thioanisole.[95]

Unnatural amino acids that have well defined coordination sites for metals have also been used to generate dative artificial metalloenzymes. Bidentate bipyridine ligands can be genetically encoded into a protein sequence, and these ligands can strongly chelate a variety of transition metal ions including, $Fe^{2+/3+}$, Cu^{2+}, $Co^{2+/3+}$ and $Ru^{2+/3+}$.[96] Through this method, Roelfes et al. generated copper ArMs which catalyse the Friedal-Crafts alkylation of a variety of indoles with a 94% conversion and high enantioselectivity.[97] The protein scaffold used was the transcription factor Lactoccal multidrug resistant Regulator (LmrR) which, at the dimeric interface, has a hydrophobic cavity which is capable of accommodating many diverse substrates.[98]

Although not technically focussed on artificial metalloenzymes, it is important to reference the pioneering research of Harry Gray who decorated many proteins with different ruthenium fragments and monitored electron transfer processes between ruthenium and iron metal centres.[99] Ru(bipy) complexes were often used due to their well-defined photochemical and electrochemical behaviour. In one example $[Ru(bipy)_2(im)]^{2+}$ units were datively coordinated to surface histidine residues on cytochrome c, and distant Fe^{2+}-Ru^{3+} electron couplings were extracted from intramolecular electron transfer rates for Ru(histidine$^{X)}$) (where X = 33, 39, 62, 72).[100]

Assembly Driven Dative Coordination

A bottom up approach of generating artificial metalloenzymes, from metal and peptidic starting materials has also proved to be fruitful. The theory behind this is to mimic nature and to use laboratory-enhanced evolution techniques, to generate comparable-to-nature, artificial metalloenzymes. Building upon the insight of Margaret Dayhoff,[101] in early evolution, higher-order metalloproteins were formed through the metal mediated assembly of simple, short peptides followed by polypeptide fusion and diversification.[102]

In a 2018 research article, Hilvert et al. report the evolution of an active zinc metalloenzyme capable of accelerating ester cleavage with high enantiospecificity and large catalytic efficiency ($k_{cat}/K_M = 10^6\ M^{-1}s^{-1}$).[103] Starting from a homo-dimeric peptide MID1, with two

Zn(II)His_3 binding sites, Zinc(II) ions were added and originally served as a template for protein assembly. The N and C termini of the dimer subunits were fused together to make a single polypeptide chain, and a combination of computational design with mutagenesis reduced the number of zinc sites to one. Through nine rounds of directed evolution, the resultant construct MID1sc10 had > 10,000-fold higher esterase activity, than the original construct. This remarkable improvement in activity, shows the power of directed evolution when applied to a metalloenzyme whereby the increase in activity is reliant on the protein providing ligands to the metal. However, the increase in activity is mainly due to evolving a protein which is specific for a given ester substrate, in particular a highly aromatic umbelliferone type derivative.

Metal ions have also been employed to template the supramolecular assembly of four α-helical bundle components to generate larger protein constructs. Tezcan et al. reported the tetrameric assembly of cytochrome cb_{562} units, with zinc sites in the interface between two protein units.[104] Following directed evolution, *in vivo* β-lactamase activity was monitored, with certain variants displaying catalytic proficiency for ampicillin hydrolysis.

1.5.3. Critical Evaluation of Artificial Metalloenzymes to Date

Evaluating these four methods of ArM formation, a key observation is that to for directed evolution to really optimise an ArM for catalytic function and improve both the k_{cat} (the chemical turnover rate) and the K_M (simplified in many systems to the substrate binding affinity), there must be direct coordination between the metal complex and the protein scaffold, as is the case for dative anchoring. In order for directed evolution to make measurable and significant improvements to k_{cat} the protein scaffold must be able to have an electronic contribution to catalysis through direct protein-metal coordination.

The work of Hilvert et al. has highlighted the catalytic possibilities of dative ArMs, in terms of the activity improvements that can be achieved through laboratory based evolutionary optimisation. In comparison to all other artificial metalloenzymes, Hilvert et al. have shown that when the metal is datively coordinated to the protein, directed evolution can have the most significant impact on catalytic efficiency (k_{cat}/K_M). Conceivably, if the protein is exposed to different substrates, different pathways of evolution can lead to a great increase in catalytic activity for each individual substrate. Henceforth, the initial ArM formed can be evolved from a generalist, with poor selectivity and activity, to a specialist, with enhanced activity; and this engineerability highlights the potential of dative artificial metalloenzymes.

Artificial metalloenzymes generated through covalent anchoring and supramolecular assembly do not have a direct bond from the metal to the protein, therefore there is a limited electronic contribution from the protein towards catalysis. Enantioselectivity and substrate selectivity can be evolved into these ArMs, however, achieving the kinetics, and high k_{cat}/K_M values of natural enzymes through directed evolution remains challenging. In many current examples, the protein is being used as a hydrophobic pocket, where enantioselective catalysis can occur, which does not utilise the full electronic, pre-organising and 3D properties that enzymes can provide. ArMs formed *via* covalent attachment rely on a chemical reaction to occur before catalysis, and often the catalyst is at the end of a flexible linker, therefore, the metal is not necessarily localised. In the particular case of streptavidin-biotin technology, the major limitation is that it is only applicable to one bacterial family, the avidins.

ArMs generated through metal substitution have shown clear catalytic potential, as shown by the case of Hartwig's iridium CYP119 metalloenzyme. Directed evolution has greatly optimised the binding and pre-organisation of the substrate for catalysis, lowering the value for K_M, **Figure 1.10**. In this system there is no direct iridium-protein coordination; the iridium metal is coordinatively saturated by four haem nitrogens, one methyl ligand and coordination to the substrate. Therefore, the small increase in k_{cat} cannot have come through an electronic (through bond) contribution to catalysis from amino acid side chain ligands and protein fold energy but from other minor contributors as indicated above. Another limitation of such a system is that it does not allow for the metal to bind more than one substrate, an essential feature of many interesting organometallic transformations e.g. metathesis.

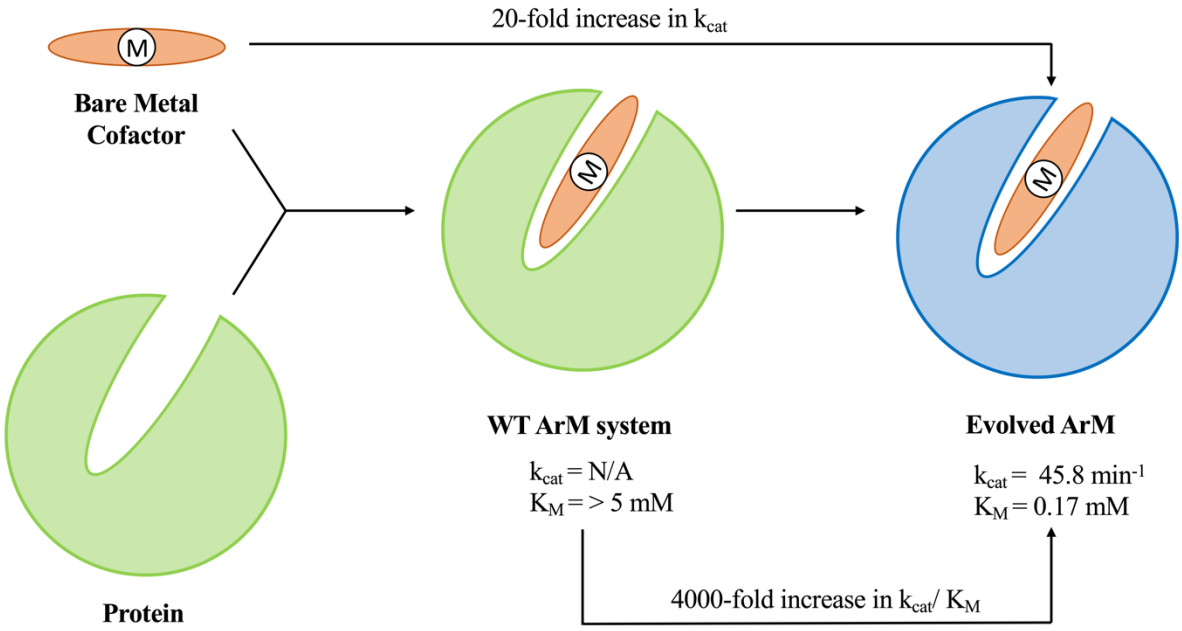


Figure 1.10: Comparison of evolution regarding bare cofactor and initial artificial metalloenzyme. The data in this figure is taken from the work from Hartwig et.al.[70] This elegant study is exemplary for the issue of using fully-substituted artificial cofactors. Whereas directed evolution was able to achieve an impressive 4000-fold increase in k_{cat}/K_M, the actual chemical turnover rate (k_{cat}) was only moderately enhanced.

1.5.4. Energetics of Dative Artificial Metalloenzymes

Upon dative coordination of a metal complex to a protein, the energetics of the active site are dependent on a combination of multiple factors. Important interactions to consider are: (i) the metal to protein bond energies, (ii) the metal to ligand bond energies, (iii) the non-covalent interactions between the protein and the metal ligands and (iv) the non-covalent interactions between the protein and the substrate. All these factors impact the energetic landscape from which catalysis occurs.

Considerations on Metal Chemistry

One possible consequence of dative coordination in ArM formation is that the metal ion will organise around available protein ligands into the lowest energetic state, the proteins ground state configuration.[105] Natural zinc finger proteins are a classic example of this, where in the absence of zinc the most stable protein confirmation is the denatured state, but upon addition of Zn(II), the metal adopts its favoured tetrahedral geometry, with four amino acid ligands, which stabilises the protein's secondary and tertiary structure.[106] These metalloenzymes often lie within thermodynamic wells therefore there is a large barrier towards catalysis.

Another possibility of dative coordination in ArM formation is a metal coordination environment pre-organised by a stable protein fold and upon metal addition the coordination site is at a higher energy than it would prefer.[105] These energised coordination environments are often highly active catalytic centres, with the metal adopting a distorted coordination environment e.g. the hydrolytic enzyme carbonic anhydrase. These energised states, referred to as entatic active sites, describe a state that is 'under tension' and catalytically poised.[107] In terms of artificial metalloenzymes these entatic states are very difficult to generate from bottom up, design-based approaches.

Considerations of Non-Covalent Protein-Substrate Interactions

With dative coordination highlighted as the optimal method of ArM formation, careful thought must be given to the selection of the starting protein scaffold to coordinate the metal complex and pre-organise the catalytic substrate for catalysis. Non-covalent interactions between the protein scaffold and catalytic substrate are vital in inferring natural metalloenzymes with their remarkable selectivity and reactivity. These interactions are crucial in placing the substrate into an optimum geometric arrangement which enhances molecular orbital overlap to the metal so that the energetic barrier to catalysis is lowered.

Given the novelty of ArMs, the lack of reliable parameters for defining transition metal bonding, and the immense complexity of the many low energy interactions that determine binding of small molecules to proteins, it is beyond current computational capabilities to predict what primary sequence and cofactors are necessary to achieve the optimal arrangement for metal catalysis. It therefore becomes important to have a malleable, promiscuous starting system that can be used to sample a large space of different structures.[108] Hence, while choosing proteins with well-defined properties and unique structures has some advantages from a design point of view, starting points that do not fold into one specific structure may be desirable, since they are not as closely constrained by any one particular energy well. Greater discussions will be made throughout this thesis into the optimum protein starting point for ArM development as coordination of metal complexes to a range of different protein scaffolds is explored.

1.5.5. Summary and Challenges of Dative Artificial Metalloenzymes

Dative attachment of unnatural metal cofactors to protein scaffolds is the most promising method to achieving artificial metalloenzymes which can match the activity of natural enzymes. There are a number of factors in support of this statement:

(i) In naturally evolved metalloenzymes, the direct coordination between the metal and the protein scaffold is integral to catalysis as the protein can attenuate metal activity through direct electronic contact as well as non-covalent contributions to catalysis, through intermolecular interactions that pre-organise substrates and stabilise transition states.

(ii) ArMs with direct coordination have a vast scope in terms of engineerability and potential evolvability of the protein-metal hybrids generated. Mutations can be made which directly alter the energetic contribution of the protein which are crucial in activating metal centres for catalysis. Once catalytic activity has been identified in dative ArMs (even low levels) then evolutionary campaigns could make remarkable improvements on activity.

(iii) Generating dative ArMs, enables the introduction of novel synthetic ligands (coordinated to the metal ion) into a protein scaffold. There is a direct linkage between the protein scaffold and synthetic ligands which are not present in biology.

(iv) Finally, selection of a dynamic, malleable protein scaffold to coordinate a desirable metal complex is vital in the successful development of artificial metalloenzymes.

An Activatable Metal Cofactor

One vision of an optimal artificial metalloenzyme would involve an exogenous metal cofactor that is catalytically inert until localised in the desired protein fold, where catalytic activity is unmasked. To our best knowledge, all current ArMs presented in the literature involve metal species that are catalytically competent without the host protein being present. The only foreseeable method to activate an inert exogenous metal cofactor to generate an artificial metalloenzyme is through direct coordination to a Lewis basic residue and subsequent metal-ligand exchange. The subsequent ligand(s) provided by the protein and the strain energy provided by the protein fold, could potentially place the metal active site into an entatic state, which is poised for catalysis. Importantly the protein in absence of the metal cofactor must retain significant stability in order to generate this entatic state.

Introducing 4d metals like ruthenium, as bare metal ion cofactors is chemically impossible, as the bare metal ion would coordinate to multiple positions; therefore, a coordinatively-saturated ruthenium complex must be exposed to a protein scaffold and a ligand exchange reaction must occur for ArM formation.

Chapter 2

2. Project Aims

2.1. Project Aims

Previous research in the group has focussed on exploring the therapeutic potential of ruthenium complexes, particularly Ru(II)(η^6-arene) complexes. Many of these complexes have been extensively studied by this group and others due to their ability to bring about cellular death at low concentrations, which has drawn comparisons to the anti-cancer drug cisplatin. The speciation of these complexes in cells, and how they bring about cellular death is poorly understood. The reactivity of Ru(II) arene complexes with proteins has drawn considerable interest as it is likely to play a crucial role in how many of these complexes induce cellular apoptosis.

Further research goals of the group are centred around the development of artificial metalloenzymes which have a direct coordination bond between the protein scaffold and a metal complex. These hybrid catalysts have properties reminiscent of both homogeneous and enzymatic catalysis and could have applications in aqueous enantioselective catalysis and synthetic biology, where they could complement metabolic pathways.

This overall aim of this project, which ties into both applications highlighted above (ruthenium therapeutics and artificial metalloenzymes), was to understand the reactivity and speciation of Ru(II) arene complexes with protein scaffolds, and explore the catalytic activity of the protein-metal hybrids generated.

Firstly, it is vital to identify the amino acid side chain residues that are capable of coordinating to a ruthenium metal centre in aqueous conditions and at biological pH. To do this, the aim was to develop novel spectroscopic approaches, particularly using solution-based NMR spectroscopy to track the speciation of bioactive Ru(II) arene complexes with free amino acids.

Understanding where and to what extent a Ru(II)(η^6-arene) complex is modifying a protein is challenging, but an extra level of complexity is being able to unmask catalytic activity upon protein coordination. At the beginning of this project, it was highlighted how controlled ligand exchange with a desirable protein scaffold could be crucial to this aim. Identifying the optimum protein scaffold and ruthenium complex to unmask catalytic activity upon coordination cannot

be predicted by simply analysing the metal complex and the protein tertiary structure. Therefore, the aim was to identify a system where the desired ligand exchange occurred and then test this system for catalytic activity.

Chapter 3

3. Experimental Methods and Protocols

3.1. General Methods

General Considerations

When necessary, all reactions were kept under an inert atmosphere or under a N_2 flow using standard Schlenk line techniques. Methanol (MeOH), acetonitrile (MeCN), dichloromethane (DCM) and other required solvents were dried through distillation, stored over suitable drying agents and purged with N_2 before use.

Chemicals and Reagents

Chemicals and solvents were purchased from chemical suppliers Sigma-Aldrich, Merck, Acros Organic and Insight Biotechnologies.

3.2. Instrumentation

3.2.1. Nuclear Magnetic Resonance Spectroscopy

All NMR data were collected at 298 K (unless stated) using Bruker Avance spectrometers with ^{1}H resonance frequencies of 400, 500, 600 MHz. All 1D ^{19}F NMR spectra were recorded using an inverse gated decoupling pulse program so that integration values can be quantified.

In proton NMR, chemical shifts (δH) are reported in parts per million (ppm), to the nearest 0.01 ppm and are referenced to the residual non-deuterated solvent peak. Coupling constants (J) are reported in Hertz (Hz) to the nearest 0.1 Hz. Data are reported in the order: (i) chemical shift, (ii) multiplicity (s = singlet; d = doublet; t = triplet; q = quartet; sep = septet; m = multiplet; or as a combination of these, e.g. dd, dt etc.), (iii) coupling constant(s), (iv) integration and (v) assignment. In carbon NMR, chemical shifts (δC) are quoted in ppm, to the nearest 0.1 ppm, and are referenced to the residual non-deuterated solvent peak

Collaborative NMR experiments were performed with Dr Stefan Freund and Dr Trevor Rutherford. Fluorine COSY and fluorinated protein NMR experiments were performed on a 600 MHz Bruker spectrometer fitted with triple-resonance cryoprobes with 5% D_2O added to each sample as a lock solvent.

All data was processed using Bruker's Topspin software, including time-dependent spectra which used Bruker's Dynamic Centre Package for kinetic analysis.

3.2.2. X-Ray Crystallography

Single-crystal X-ray diffraction data were collected at 180 K using either a Nonius Kappa CCD or Bruker D8-QUEST diffractometer, equipped with MoKα or CuKα radiation, respectively. Structures were solved using SHELXT[109] and refined on F^2 using SHELXL.[110] Data collection and analysis were conducted by Dr Andrew Bond.

3.2.3. Mass Spectrometry

Low Resolution Mass Spectrometry

Low-resolution electrospray ionisation mass spectrometry (LR-MS) was undertaken on a Micromass Quattro mass spectrometer (capillary voltage 1 – 3 kV, cone voltage 10 – 30 kV; desolvation temperature 320 K) infused manually from a Harvard syringe pump at a rate of 10 $\mu L min^{-1}$. Mass values are reported within the error limits of ±0.05 mass units.

High Resolution Mass Spectrometry

High resolution mass spectrometry (HRMS) measurements were recorded with a Xevo G2-S TOF mass spectrometer spectrometer (capillary voltage 2 kV, cone voltage 40 kV; desolvation temperature 350 °C) with automatic injection. Mass values are reported within the error limits of ±5 ppm mass units. Nitrogen was used as the desolvation gas at a total flow of 850 L/h.

Comparing Calculated and Observed Masses

Where mass spectrometry has been used for the characterisation of ruthenium complexes, the major isotopes ^{102}Ru (molecular weight 101.9034 g mol^{-1}) and ^{35}Cl (molecular weight 34.9689 g mol^{-1}) were used to compare observed and calculated masses.

3.2.4. Liquid Chromatography Mass Spectrometry

Measurement

Protein LC–MS was performed on a Xevo G2-S TOF mass spectrometer coupled to an Acquity UPLC system using an Acquity UPLC BEH300 C4 column (1.7 μm, 2.1 × 50 mm). H_2O with 0.1 % formic acid (solvent A) and 95 % acetonitrile and 5 % water with 0.1 % formic acid (solvent B), were used as the mobile phase at a flow rate of 0.2 mL/min. The gradient was programmed as follows: 95% A for 0.93 min, then a gradient to 100% B over 4.28 min, then 100% B for 1.04 minutes, then a gradient to 95% A over 1.04 min. The electrospray source

was operated with a capillary voltage of 2.0 kV, a cone voltage of 40 kV and a desolvation temperature of 350 °C . Nitrogen was used as the desolvation gas at a total flow of 850 L/h.

Analysis

The initial output from the LC-MS is a chromatogram that reports the total intensity counts (number of ions that hit the mass detector) over the course of the gradient. Combining and collecting all of these mass peaks generates an ion series for a protein or modified protein between a mass range of 200 – 2000 m/z. Total mass spectra were reconstructed from the ion series using the maximum entropy (MaxEnt) algorithm pre-installed on MassLynx software (v4.1 from Waters).

The parameters used for this MaxEnt deconvolution algorithm are vital to achieve accurate zero charge mass spectra. Throughout this thesis, the MaxEnt procedure splits the deconvoluted spectra into 0.25 Da channels, and uses a width at half peak height of 0.75 Da. There is a large error associated with the MaxEnt procedure, however a 0.25 Da/channel aims to minimise this error, so that observed and calculated masses are ±3 Da apart. Importantly, in all experiments, the modifications theoretically possible within a given reaction mixture differ by > 100 Da, which allows for assignment of products to be given with some assurance.

3.2.5. Tandem MS/MS Experiments

Protein samples (30 µM) were prepared for mass spectrometric analysis by hand, solution samples were digested with trypsin (Promega, UK) at a 1:50 ratio overnight at 37 °C. The resulting peptides were diluted in 2% v/v formic acid, 2% v/v acetonitrile. The digests were analysed by nano-scale capillary LC-MS/MS using an Ultimate U3000 HPLC (ThermoScientific Dionex, San Jose, USA) to deliver a flow of approximately 300 nL/min. A C18 Acclaim PepMap100 5 µm, 100 µm x 20 mm nanoViper (ThermoScientific Dionex, San Jose, USA), trapped the peptides prior to separation on a C18 T3 1.8 µm, 75 µm x 250 mm analytical UPLC column (Waters, UK). Peptides were eluted with a 30 minute gradient of acetonitrile (2% to 40%). The analytical column outlet was directly interfaced *via* a nano-flow electrospray ionisation source, with a quadrupole Orbitrap mass spectrometer (Q-Exactive HFX, ThermoScientific, USA). MS data were acquired in data-dependent mode using a top 10 method, where ions with a precursor charge state of +1 were excluded. High-resolution full scans (R = 60,000, m/z 300-1800) were recorded in the Orbitrap followed by higher energy collision dissociation (HCD) (26 % Normalised Collision Energy) of the 10 most intense MS peaks. The fragment ion spectra were acquired at a resolution of 15 000 and dynamic exclusion

window of 20 s was applied.

LC-MS/MS data were then searched against an in-house database using the Mascot search engine programme (Matrix Science, UK). Database search parameters were set with a precursor tolerance of 30 ppm and a fragment ion mass tolerance of 0.1 Da. One missed enzyme cleavage was allowed and variable modifications for ruthenium fragments allowed on the protein N-terminus, Cys, His, Lys, Asn, Gln, Arg and Trp were included. MS/MS data were validated using the Scaffold programme (Proteome Software Inc., USA).[111] All data were additionally interrogated manually.

3.2.6. UV-Vis Spectrometry

All UV-vis spectra were recorded using Cary 400 or Cary 100 UV-Vis spectrophotometers. Standard 9/9/B quartz cuvettes with a 10 mm pathlength were used for recording absolute spectra. Data were recorded at 0.5 nm intervals using a spectral bandwidth of 2 nm, with full slit height. For all absolute spectra, a baseline of appropriate buffer was recorded and subtracted from all sample data.

UV-Vis spectroscopy is used throughout this thesis to measure protein concentration, *via* the Beer Lambert law, **Equation 3.1**.

$$A = \varepsilon c l \qquad (3.1)$$

3.2.7. Circular Dichroism

CD spectra were recorded in 1 mm pathlength quartz cuvettes using a Chirascan CD spectrometer. CD data were recorded from 300 – 180 nm with sampling at 0.5 nm intervals. Samples were made to an expected concentration of approximately 5 μM in low buffer concentration.

Temperature melt experiments were performed measuring ellipticity (θ) with a fixed λ = 222 nm across a temperature range of 25 – 90 °C with a linear increase of 1 °C / min.

3.2.8. Elemental Analysis

Elemental microanalytical data were obtained from the University of Cambridge, Department of Chemistry microanalytical service. C,H,N analysis was carried out using an Exeter Analytical CE-440 Elemental Analyser by combustion of the sample under a pure O2 atmosphere at 975 ºC.

ICP-OES analyses were carried out on a Thermo Scientific iCAP 7400 Duo. Double-deionised water was used for all analysis. The Ruthenium Specpure plasma standard (ruthenium chloride, 1004 ± 5 µg/mL in 10% v/v hydrochloric acid) was diluted with HNO_3 to 2% v/v HNO_3 and calibrants were prepared at concentrations of 1 − 1000 ppb. Samples were prepared through dilution in 2% v/v HNO_3.

3.2.9. Fluorescence Measurements

All fluorescence and UV-Vis spectroscopy in a 96-well plate format was performed on a SpectraMax i3x plate reader (Molecular Devices). Unless stated differently, all measurements monitoring transfer hydrogenation were taken with λ_{Ex} = 370 nm, λ_{Em} = 460 nm and all measurements monitoring olefin metathesis were taken with λ_{Ex} = 325 nm, λ_{Em} = 460 nm. All measurements used a bottom read, PMT = high, bandwidth - 9 nm excitation and 15 nm emission and catalytic time course experiments used 5 secs orbital shaking at the start and 1 secs between each measurement.

3.3. Synthetic Protocols and Characterisation

3.3.1. Synthesis of the Fluorinated Bipyridines B – E

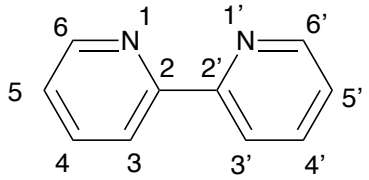
Method 1 Fluorinated Bipyridine Synthesis

The starting bromopyridine (1 mmol), $Pd(OAc)_2$ (0.025 mol Eq.), K_2CO_3 (1 mol Eq.) and poly(ethylene glycol) (M_w 4000, 5.0 – 10g) were combined in a nitrogen purged flask. This mixture was heated to 120 °C and the temperature maintained for 48 hours with stirring. The mixture was cooled to 80 °C and 15 mL of warm water was added. Once at room temperature a further 10 mL of water was added and the suspension exhaustively extracted with ethyl acetate. The combined extracts were washed with saturated aqueous $Na_2S_2O_3$ and three times with brine. The organic layer was then dried over $MgSO_4$ and solvent removed in vacuo. The product was then purified *via* sublimation.

Method 2 Fluorinated Bipyridine Synthesis[112]

The starting bromopyridine (1.0 mmol), $Pd(OAc)_2$ (0.025 mmol), indium (0.5 mmol) and lithium chloride (1.5 mmol) in dry DMF (2 mL) was stirred at 100 °C for 1 h under a nitrogen atmosphere. The reaction mixture was quenched with $NaHCO_3$ (satd aq). The aqueous layer was extracted with ethyl acetate (320 mL) and the combined organic phase was washed with water (20 mL) and brine (20 mL), dried with $MgSO_4$, and filtered. The residue was purified by silica gel column chromatography (EtOAc–hexane = 1:2).

Labelling of the Bipyridine



Characterisation of 3,3'-difluorobipyridine (B)

Synthesised using Method 1 with the starting material 2-bromo-3-fluoropyridine. Yield: 34 % (32.6 mg). Appearance: White solid

^{1}H NMR (400.13 MHz, $CDCl_3$): δ (ppm) 8.61 (d, $^{3}J_{HH}$ = 5.0 Hz, 2H, 6,6'-position), 7.54 (m, 2H, 4,4'-position,), 7.52 (m, 2H, 5,5'-position).**^{13}C{^{1}H} NMR** (100.57 MHz, $CDCl_3$): δ (ppm) 157.9 (dd, $^{1}J_{CF}$ = 265.1 Hz, $^{4}J_{CF}$ = 3.8 Hz, 3,3'-position), 146.1 (overlapping doublets, $^{4}J_{CF}$ = 2.7 Hz, $^{4}J_{CF}$ = 3.8 Hz, 6,6'-position), 142.3 (dd, $^{2}J_{CF}$ = 8.3 Hz, $^{3}J_{CF}$ = 2.3 Hz, 2,2'-position), 125.5 (overlapping doublets, $^{3}J_{CF}$ = 2.6 Hz, $^{3}J_{CF}$ = 2.1 Hz, 5,5'-position), 124.2 (d, $^{2}J_{CF}$ = 6.7 Hz, 4,4'-position). **^{19}F{^{1}H} NMR** (376.50 MHz, $CDCl_3$): δ (ppm) -121.9 (s).

Characterisation of 5,5'-difluorobipyridine (C)

Synthesised using Method 2 with the starting material 2-bromo-5-fluoropyridine. Yield: 78 % (74.9 mg). Appearance: White solid

^{1}H NMR (400.13 MHz, $CDCl_3$): δ (ppm) 8.50 (d, $^{4}J_{HH}$ = 2.8 Hz, 2H, 6,6'-position), 8.38 (dd, $^{3}J_{HH}$ = 8.7 Hz, $^{4}J_{HF}$ = 4.5 Hz, 2H, 3,3'-position), 7.52 (m, $^{3}J_{HH}$ = 8.7 Hz, $^{4}J_{HH}$ = 2.8 Hz, 2H, 4,4'-position). **^{13}C{^{1}H} NMR** (100.57 MHz, $CDCl_3$): δ (ppm) 160.0 (d, $^{1}J_{CF}$ = 258 Hz, 5,5'-position), 151.7 (d, $^{4}J_{CF}$ = 4 Hz, 2,2'-position), 137.4 (d, $^{2}J_{CF}$ = 24 Hz, 6,6'-position), 123.8 (d, $^{2}J_{CF}$ = 18 Hz, 4,4'-position), 122.3 (d, $^{3}J_{CF}$ = 5 Hz, 3,3'-position). **^{19}F{^{1}H} NMR** (376.50 MHz, $CDCl_3$): δ (ppm) -127.4 (s).

Characterisation of 5,5'-di(trifluoromethyl)bipyridine (D)

Synthesised using Method 2 with the starting material 2-bromo-5-(trifluoromethyl)pyridine. Yield: 65 % (94.9 mg). Appearance: White solid

^{1}H NMR (400.13 MHz, $CDCl_3$): δ (ppm) 8.99 (d, $^{4}J_{HH}$ = 1.8 Hz, 2H, 6,6'-position), 8.65 (d, $^{3}J_{HH}$ = 8.4 Hz, 2H, 3,3'-position), 8.12 (dd, $^{3}J_{HH}$ = 8.4 Hz, $^{4}J_{HH}$ = 1.8 Hz, 2H, 4,4'-position). **^{13}C{^{1}H} NMR** (100.57 MHz, $CDCl_3$): δ (ppm) 157.8 (s, 2,2'-position), 146.5 (q, $^{3}J_{CF}$ = 3.9 Hz, 6,6'-position), 134.5 (q, $^{3}J_{CF}$ = 3.9 Hz, 4,4'-position), 127.3 (q, $^{2}J_{CF}$ = 33 Hz, 5,5'-position),

123.7 (q, $^{1}J_{CF}$ = 273 Hz, $\underline{C}F_3$,), 121.4 (s, 3,3'-position). **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, $CDCl_3$): δ (ppm) -62.4 (s).

Characterisation of 6,6'-difluorobipyridine (E)

Synthesised using Method 1 with the starting material 2-bromo-6-fluoropyridine. Yield: 5 % (4.8 mg). Appearance: White solid

^{1}H NMR (400.13 MHz, $CDCl_3$): δ (ppm) 8.26 (dd, $^{3}J_{HH}$ = 7.8 Hz, $^{4}J_{HH}$ = 2.4 Hz, 2H, 3,3'-position), 7.92 (overlapping dd, $^{3}J_{HH}$ = 7.8 Hz, 2H, 4,4'-position), 6.97 (dd, $^{3}J_{HH}$ = 7.8 Hz, $^{4}J_{HH}$ = 2.4 Hz, 2H, 5,5'-position). Lack of sample restricted analysis to ^{1}H and $^{19}F\{^{1}H\}$ NMR. **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, $CDCl_3$): δ (ppm) -67.1 (s).

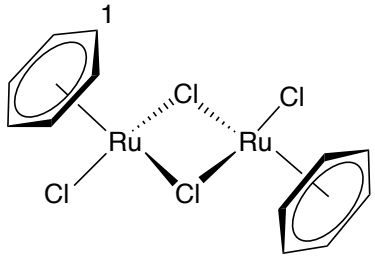
3.3.2. Synthesis of the Ruthenium Dimers I – IV

Synthetic Procedure for Dimers I – III [113]

Ruthenium trichloride (9.6 mmol) was dissolved in ethanol (80 mL). The appropriate hexacyclic diene (38.4 mmol) was added slowly and the mixture was refluxed for 16 hours. The suspension was filtered and after being washed with EtOH (10 mL), Et_2O (10 mL) and was dried in vacuo.

Characterisation of $[Ru(benzene)Cl_2]_2$

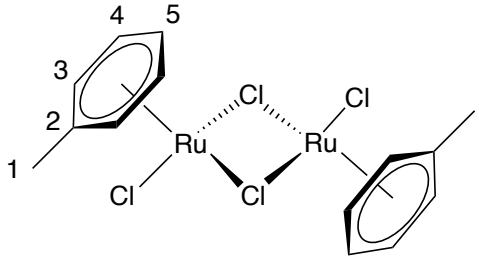
Synthesised from the starting diene 1,4-cyclohexadiene. Yield: 74 % (1.78 g). Appearance: Red solid



^{1}H NMR (400.13 MHz, D_2O): 5.99 (s, 6H, PhH). **$^{13}C\{^{1}H\}$ NMR** (100.57 MHz, D_2O): 80.5 (s, PhH).

Characterisation of $[Ru(toluene)Cl_2]_2$

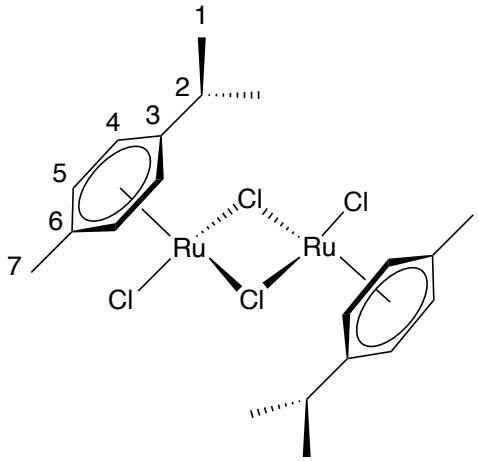
Synthesised from the starting diene 1-methyl-1,4-cyclohexadiene. Yield: 82 % (2.08 g). Appearance: Red solid.



^{1}H NMR (400.13 MHz, D_2O): 5.99 (overlapping dd, $^{3}J_{HH}$ = 5.7 Hz, 2H, 4-position), 5.71 (d, $^{3}J_{HH}$ = 5.7 Hz, 2H, 3-position), 5.60 (t, $^{3}J_{HH}$ = 5.7 Hz, 1H, 5-position), 2.14 (s, 3H, 1-position). **^{13}C{^{1}H} NMR** (100.57 MHz, $CDCl_3$): δ (ppm) 105.7 (s, 2-position), 89.5 (s, 4-position), 84.9 (s, 3-position), 82.2 (s, 5-position), 18.6 (s, 1-position).

Characterisation of [Ru(*p*-cymene)Cl_2]$_2$

Synthesised from the starting diene α - terpinene. Yield: 48 % (1.41 g). Appearance: Red solid



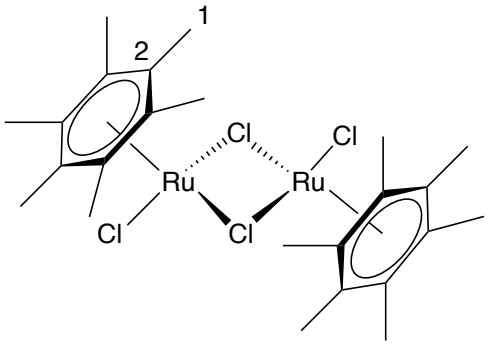
^{1}H NMR (400.13 MHz, D_2O): 5.82 (d, $^{3}J_{HH}$ = 6.0 Hz, 2H, 5-position), 5.77 (d, $^{3}J_{HH}$ = 5.7 Hz, 2H, 4-position), 2.83 (sept, $^{3}J_{HH}$ = 7.0 Hz, 1H, 2-position), 2.09 (s, 3H, 7-position), 1.19 (d, $^{3}J_{HH}$ = 7.0 Hz, 6H, 1-position). **^{13}C{^{1}H} NMR** (100.57 MHz, $CDCl_3$): δ (ppm) 106.4 (s, 5-position), 100.1 (s, 4-position), 86.3 (s, 6-position), 85.5 (s, 3-position), 30.0 (s, 2-position), 21.5 (s, 7-position), 17.9 (s, 1-position).

Synthetic Procedure for Dimer IV [114]

A mixture of [Ru(*p*-cymene)Cl_2]$_2$ (0.21 mmol) and hexamethylbenzene (0.21mmol) was refluxed in diglyme (18 ml) with stirring under dry nitrogen atmosphere for around 9 hr. The solution was cooled to room temperature and the red brown product was filtered off, washed with hexane (510 ml) to remove excess hexamethylbenzene and *p*-cymene dimer and finally with Et_2O. The compound was recrystallised from CH_3Cl/Et_2O

Characterisation of [Ru(hexamethylbenzene)Cl_2]$_2$

Yield: 42 % (58.9 mg). Appearance: Red solid



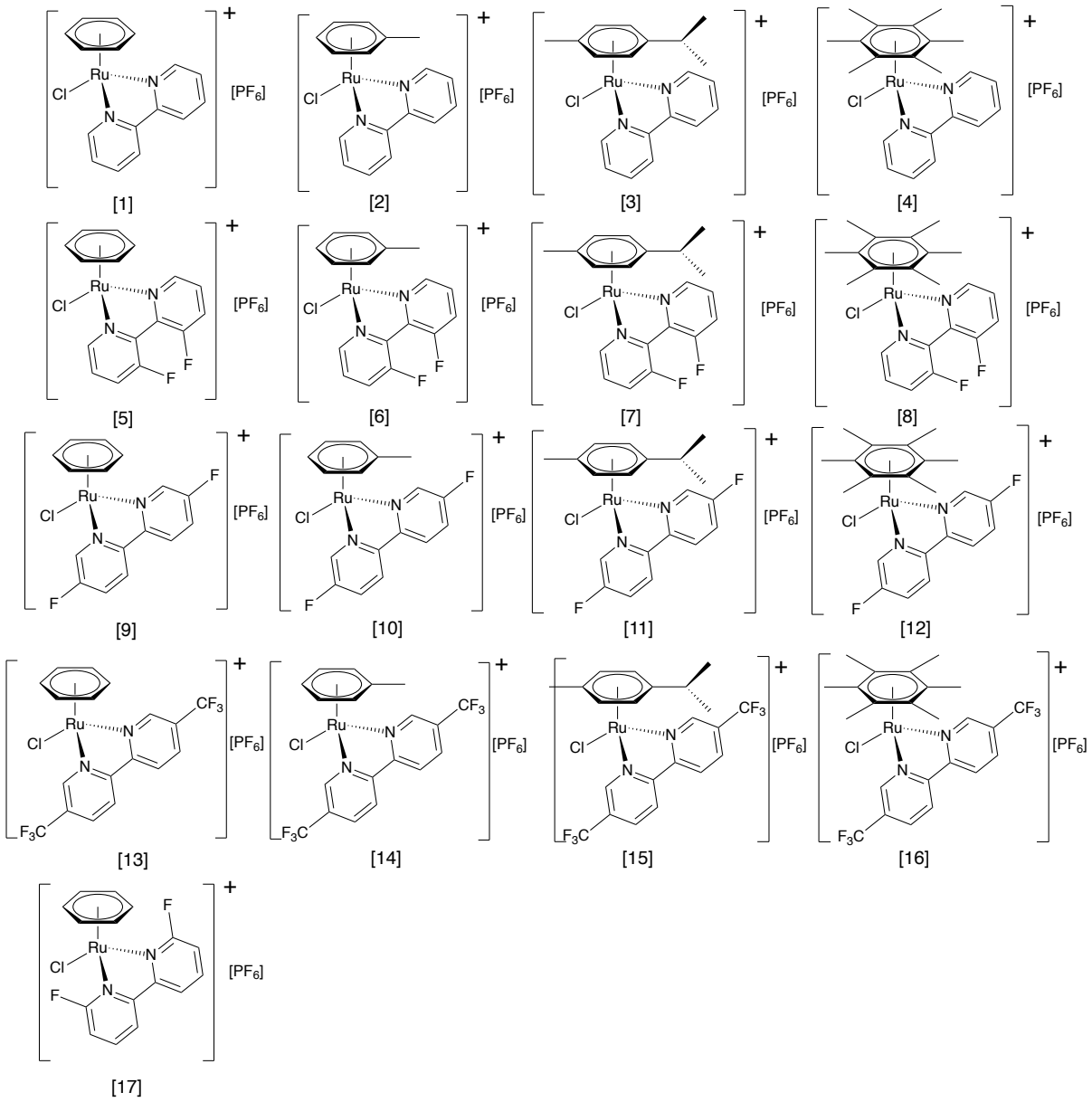
^{1}H NMR (400.13 MHz, D_2O): 2.02 (s, 18H, 1-position), **^{13}C{^{1}H} NMR** (100.57 MHz, $CDCl_3$): δ (ppm) 96.1 (s, 2-position), 15.3 (s, 1-position)

3.3.3. Synthesis of the Ru(II)(η^6-arene)(bipyridine) complexes [1] – [17]

Synthetic Procedure for Complexes [1] – [17]

The appropriate ruthenium arene dimer (I – IV) (0.06 mmol) and appropriate bipyridine (A – E) (0.12 mmol) were added to a nitrogen purged flask. Freshly distilled MeOH (25 mL) was added and the reaction was stirred for 24 hours at room temperature. The contents were filtered under gravity to remove excess ruthenium and the solution was reduced to approximately 5mL in vacuo. NH_4PF_6 (115 mg, 0.72 mmol) was added and the mixture was shaken and left at -10 °C for a further 24 hours. The product was collected by gravity filtration and washed with Et_2O (10 mL).

Structures of Complexes [1] – [17]



Characterisation of [Ru(η^6-benzene)(bipyridine)Cl][PF_6] – [1]

Synthesised from the starting dimer I and 2,2'-bipyridine. Yield: 74 % (44.0 mg). Appearance: Yellow needles.

Elemental: Anal. Calcd for $C_{16}H_{14}ClF_6N_2PRu$: C, 37.26; H, 2.74; N, 5.43. Found: C, 37.07; H, 2.57; N, 5.28. **LRMS (ESI$^+$):** m/z 371.02 $[M - PF_6]^+$ (m_{calc} = 370.99). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.63 (d, $^3J_{HH}$ = 5.6 Hz, 2H, 6,6'-position), 8.64 (d, $^3J_{HH}$ = 8.1 Hz, 2H, 3,3'-position), 8.29 (overlapping dd, $^3J_{HH}$ = 8.1 Hz, $^3J_{HH}$ = 7.9 Hz, 2H, 4,4'-position), 7.79 (overlapping dd, $^3J_{HH}$ = 7.9 Hz, $^3J_{HH}$ = 5.6 Hz, 2H, 5,5'-position), 6.25 (s, 6H, PhH). **^{13}C{^{1}H}**

NMR (100.57 MHz, d^6-DMSO): 156.0 (2,2'-position), 154.5 (6,6'-position), 140.0 (4,4'-position) 127.4 (5,5'-position), 123.7 (3,3'-position), 87.0 (PhH, s).

Characterisation of [Ru(η^6-toluene)(bipyridine)Cl][PF_6] – [2]

Synthesised from the starting dimer II and 2,2'-bipyridine. Yield: 61 % (37.2 mg). Appearance: Yellow solid.

Elemental: Anal. Calcd for $C_{17}H_{16}ClF_6N_2PRu$: C, 38.54; H, 3.04; N, 5.29. Found: C, 38.71; H, 2.61; N 4.85. **LRMS (ESI$^+$):** m/z 385.00 $[M - PF_6]^+$ (m_{calc} = 385.00). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.55 (d, $^3J_{HH}$ = 5.4 Hz, 2H, 6,6'-position), 8.63 (d, $^3J_{HH}$ = 7.9 Hz, 2H, 3,3'-position), 8.28 (overlapping dd, $^3J_{HH}$ = 8.4 Hz, $^3J_{HH}$ = 7.9 Hz, 2H, 4,4'-position), 7.78 (overlapping dd, $^3J_{HH}$ = 8.4 Hz, $^3J_{HH}$ = 5.4 Hz, 2H, 5,5'-position), 6.30 (overlapping dd, $^3J_{HH}$ = 6.2 Hz, 2H, 4-Tol-position), 5.96 (d, $^3J_{HH}$ = 6.2 Hz, 2H, 3-Tol-position), 5.83 (t, $^3J_{HH}$ = 6.2 Hz, 1H, 5-Tol-position), 2.23 (s, 3H, 1-Tol-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 156.3 (2,2'-position), 155.0 (6,6'-position), 140.3 (4,4'-position) 127.8 (5,5'-position), 124.1 (3,3'-position), 106.4 (s, 2-Tol-position), 91.0 (s, 4-Tol-position), 83.3 (s, 3-Tol-position), 80.2 (s, 5-Tol-position), 19.2 (s, 1-Tol-position).

Characterisation of [Ru(η^6-*p*-cymene)(bipyridine)Cl][PF_6] – [3]

Synthesised from the starting dimer III and 2,2'-bipyridine. Yield: 64 % (42.1 mg). Appearance: Yellow solid.

Elemental: Anal. Calcd for $C_{20}H_{22}ClF_6N_2PRu$: C, 42.00; H, 3.88; N, 4.90. Found: C, 41.89; H, 3.86; N, 4.78 **HRMS (ESI$^+$):** m/z 427.0625 $[M - PF_6]^+$ (m_{calc} = 427.0515). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.54 (d, $^3J_{HH}$ = 5.8 Hz, 2H, 6,6'-position), 8.64 (d, $^3J_{HH}$ = 8.0 Hz, 2H, 3,3'-position), 8.29 (overlapping dd, $^3J_{HH}$ = 8.6 Hz, $^3J_{HH}$ = 8.0 Hz, 2H, 4,4'-position), 7.79 (overlapping dd, $^3J_{HH}$ = 8.0 Hz, $^3J_{HH}$ = 6.5 Hz, 2H, 5,5'-position), 6.21 (d, $^3J_{HH}$ = 6.1 Hz, 2H, 5-cym-position), 5.98 (d, $^3J_{HH}$ = 6.1 Hz, 2H, 4-cym-position), 2.57 (sept, $^3J_{HH}$ = 6.9 Hz, 1H, 2-cym-position), 2.18 (s, 3H, 7-cym-position), 0.94 (d, $^3J_{HH}$ = 6.9 Hz, 6H, 1-cym-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 155.7 (2,2'-position), 154.3 (6,6'-position), 139.9 (4,4'-position) 127.5 (5,5'-position), 123.8 (3,3'-position), 104.1 (3-cym position), 104.0 (6-cym position), 86.7 (5-cym-position), 83.9 (4-cym-position), 30.3 (2-cym-position), 21.6 (1-cym-position), 18.3 (7-cym-position).

Characterisation of [Ru(η^6-hexamethylbenzene)(bipyridine)Cl][PF_6] – [4]

Synthesised from the starting dimer IV and 2,2'-bipyridine. Yield: 54 % (37.3 mg). Appearance: Orange solid

Elemental: Anal. Calcd for $C_{22}H_{26}ClF_6N_2PRu$: C, 44.04; H, 4.37; N, 4.67. Found: C, 44.18; H, 4.03; N 4.17. **HRMS (ESI$^+$):** m/z 454.0855 $[M - PF_6]^+$ (m_{calc} = 454.0828). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 8.91 (d, $^3J_{HH}$ = 5.7 Hz, 2H, 6,6'-position), 8.60 (d, $^3J_{HH}$ = 8.0 Hz, 2H, 3,3'-position), 8.24 (overlapping dd, $^3J_{HH}$ = 8.6 Hz, $^3J_{HH}$ = 8.0 Hz, 2H, 4,4'-position), 7.79 (overlapping dd, $^3J_{HH}$ = 8.6 Hz, $^3J_{HH}$ = 6.7 Hz, 2H, 5,5'-position), 2.02 (s, 2-hmb-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 154.9 (2,2'-position), 153.6 (6,6'-position), 139.5 (4,4'-position) 127.7 (5,5'-position), 123.4 (3,3'-position), 95.4 (s, 2-hmb-position), 15.1 (s, 1-hmb-position).

Characterisation of [Ru(η^6-benzene)(3,3'-difluorobipyridine)Cl][PF_6] – [5]

Synthesised from the starting dimer I and starting bipyridine B. Yield: 41 % (26.0 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange needles.

Elemental: Anal. Calcd for $C_{16}H_{12}ClF_8N_2PRu$: C, 34.83; H, 2.19; N, 5.08. Found: C, 35.01; H, 2.22; N, 4.95. **LRMS (ESI$^+$):** m/z 407.00 $[M - PF_6]^+$ (m_{calc} = 406.97). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.64 (d, $^3J_{HH}$ = 5.3 Hz, 2H, 6,6'-position), 8.34 (dd, $^3J_{HH}$ = 8.0 Hz, $^3J_{HH}$ = 5.3 Hz, 2H, 5,5'-position), 7.97 (d, $^3J_{HH}$ = 8.0 Hz, 2H, 4,4'-position), 6.29 (s, 6H, PhH). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 3,3'-bipy peak not resolvable, 153.9 (s, 6,6'-position), 141.1 (dd, $^2J_{CF}$ = 8.2 Hz, $^3J_{CF}$ = 6.4 Hz, 2,2'-position), 130.0 – 129.0 (complex m, 5,5'-position and 4,4'-position), 87.9 (s, PhH). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6), -103.9 (s, 3,3'-position)

Characterisation of [Ru(η^6-toluene)(3,3'-difluorobipyridine)Cl][PF_6] – [6]

Synthesised from the starting dimer II and starting bipyridine B. Yield: 74 % (48.2 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange needles.

Elemental: Anal. Calcd for $C_{17}H_{14}ClF_8N_2PRu$: C, 36.09; H, 2.49; N, 4.95. Found: C, 35.86; H, 2.38; N, 4.89. **LRMS (ESI$^+$):** m/z 421.10 $[M - PF_6]^+$ (m_{calc} = 420.99). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.54 (d, $^3J_{HH}$ = 5.0 Hz, 2H, 6,6'-bipy-position), 8.32 (dd, $^3J_{HF}$ = 8.0 Hz, $^3J_{HH}$ = 5.0 Hz, 2H, 4,4'-bipy-position), 7.95 (m, $^3J_{HH}$ = 5.0 Hz, $^3J_{HH}$ = 5.0 Hz, $^4J_{HF}$ = 3.1 Hz, 2H, 5,5'-bipy-position), 6.30 (overlapping dd, $^3J_{HH}$ = 6.0 Hz, 2H, 4-Tol-position), 5.96 (d, $^3J_{HH}$ = 6.0 Hz, 2H, 3-Tol-position), 5.85 (t, $^3J_{HH}$ = 6.0 Hz, 1H, 5-Tol-position), 2.24 (s, 3H, 1-

Tol-position). **$^{13}C\{^{1}H\}$ NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 157.1 (d, $^{1}J_{CF}$ = 264 Hz, 1H, 3,3'-bipy-position), 153.2 (s, 6,6'-position), 140.8 (dd, $^{2}J_{CF}$ = 8.7 Hz, $^{3}J_{CF}$ = 6.4 Hz, 2,2'-position,), 129.3 (s, 5,5'-position), 129.0 (d, $^{2}J_{CF}$ = 12.5 Hz, 4,4'-position), 106.8 (s, 2-Tol-position), 91.2 (s, 4-Tol-position), 82.9 (s, 3-Tol-position), 80.2 (s, 5-Tol-position), 18.7 (s, 1-Tol-position). **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^{1}J_{PF}$ = 711 Hz, PF_6), -108.1 (s, 3,3'-position).

Characterisation of [Ru(η^6-*p*-cymene)(3,3'-difluorobipyridine)Cl][PF_6] – [7]

Synthesised from the starting dimer III and starting bipyridine B. Yield: 59 % (41.3 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange needles.

Elemental: Anal. Calcd for $C_{20}H_{20}ClF_8N_2PRu$: C, 39.52; H, 3.32; N, 4.61. Found: C, 38.50; H, 3.31; N, 4.79. **LRMS (ESI$^+$):** m/z 463.06 $[M - PF_6]^+$ (m_{calc} = 463.03). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.52 (d, $^{3}J_{HH}$ = 5.0 Hz, 2H, 6,6'-position), 8.34 (dd, $^{3}J_{HF}$ = 8.3 Hz, $^{3}J_{HH}$ = 5.0 Hz, 2H, 4,4'-position), 7.97 (5,5'-position, 2H, m, $^{3}J_{HH}$ = 5.0 Hz, $^{3}J_{HH}$ = 5.0 Hz, $^{4}J_{HF}$ = 3.1 Hz), 6.23 (d, $^{3}J_{HH}$ = 6.4 Hz, 2H, 5-cym-position), 6.00 (d, $^{3}J_{HH}$ = 6.4 Hz, 2H, 4-cym-position), 2.64 (sept, $^{3}J_{HH}$ = 7.0 Hz, 1H, 2-cym-position), 2.17 (s, 3H, 7-cym-position), 1.00 (d, $^{3}J_{HH}$ = 7.0 Hz, 6H, 1-cym-position). **$^{13}C\{^{1}H\}$ NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 3,3'-bipy peak in baseline, 153.3 (s, 6,6'-position), 140.5 (dd, $^{2}J_{CF}$ = 6.8 Hz, $^{3}J_{CF}$ = 4.0 Hz, 2,2'-position), 129.4 (s, 5,5'-position), 128.7 (d, $^{2}J_{CF}$ = 12.4 Hz, 4,4'-position), 105.1 (s, 3-cym-position), 104.5 (s, 6-cym-position), 86.7 (s, 5-cym-position), 84.2 (s, 4-cym-position), 30.4 (s, 2-cym-position), 21.7 (s, 1-cym-position), 18.2 (s, 7-cym-position). **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^{1}J_{PF}$ = 711 Hz, PF_6), -103.9 (s, 3,3'-position).

Characterisation of [Ru(η^6-hexamethylbenzene)(3,3'-difluorobipyridine)Cl][PF_6] – [8]

Synthesised from the starting dimer IV and starting bipyridine B. Yield: 55 % (40.2 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange/Red needles.

Elemental: Anal. Calcd for $C_{22}H_{24}ClF_8N_2PRu$: C, 41.55; H, 3.80; N, 4.41. Found: C, 41.45; H, 3.73; N, 4.32. **LRMS (ESI$^+$):** m/z 490.98 $[M - PF_6]^+$ (m_{calc} = 491.06). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 8.87 (d, $^{3}J_{HH}$ = 5.0 Hz, 2H, 6,6'-position), 8.27 (m, $^{3}J_{HF}$ = 8.0 Hz $^{3}J_{HH}$ = 5.0 Hz, 2H, 4,4'-position), 7.94 (m, $^{3}J_{HH}$ = 5.0 Hz, $^{3}J_{HH}$ = 5.0 Hz, $^{4}J_{HF}$ = 3.0 Hz, 2H, 5,5'-position), 2.02 (s, 18H). **$^{13}C\{^{1}H\}$ NMR** (100.57 MHz, d^6-DMSO): 159.2 (d, $^{1}J_{CF}$ = 265 Hz, 1H, 3,3'-position), 151.1 (s, 6,6'-position), 140.6 (dd, $^{2}J_{CF}$ = 8.0 Hz, $^{3}J_{CF}$ = 6.0 Hz, 2,2'-position), 129.6 (s, 5,5'-position), 128.7 (d, $^{2}J_{CF}$ = 11.0 Hz 4,4'-position), 96.1 (s, 2-hmb-

position), 15.0 (s, 1-hmb-position). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6,), -104.2 (s, 3,3'-position).

Characterisation of [Ru(η^6-benzene)(5,5'-difluorobipyridine)Cl][PF_6] – [9]

Synthesised from the starting dimer I and starting bipyridine C. Yield: 66 % (41.8 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange needles

Elemental: Anal. Calcd for $C_{16}H_{12}ClF_8N_2PRu$: C, 34.83; H, 2.19; N, 5.08. Found: C, 35.03; H, 2.39; N, 4.85. **LRMS (ESI$^+$):** m/z 407.01 $[M - PF_6]^+$ (m_{calc} = 406.97). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.78 (s, 2H, 6,6'-position), 8.71 (dd, $^3J_{HH}$ = 9.6 Hz, $^3J_{HF}$ = 4.6 Hz, 2H, 4,4'-position), 8.36 (overlapping dd, $^3J_{HH}$ = 9.6 Hz, $^4J_{HF}$ = 7.8 Hz, 2H, 3,3'-position), 6.30 (s, 6H, PhH). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 159.4 (d, $^1J_{CF}$ = 256 Hz, 5,5'-position), 150.7 (s, 2,2'-position), 144.7 (d, $^2J_{CF}$ = 33 Hz, 6,6'-position,), 127.4 (d, $^2J_{CF}$ = 19.0 Hz, 4,4'-position), 125.7 (d, $^3J_{CF}$ = 8.0 Hz, 3,3'-position), 87.2 (s, PhH). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6), -119.4 (s, 5,5'-position).

Characterisation of [Ru(η^6-toluene)(5,5'-difluorobipyridine)Cl][PF_6] – [10]

Synthesised from the starting dimer II and starting bipyridine C. Yield: 54 % (35.1 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange plates

Elemental: Anal. Calcd for $C_{17}H_{14}ClF_8N_2PRu$: C, 36.09; H, 2.49; N, 4.95. Found: C, 35.75; H, 2.46; N, 4.73. **LRMS (ESI$^+$):** m/z 420.88 $[M - PF_6]^+$ (m_{calc} = 420.99). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.68 (overlapping dd, $^3J_{HF}$ = 3.2 Hz, $^4J_{HH}$ = 2.5 Hz, 2H, 6,6'-bipy-position), 8.71 (dd, $^3J_{HH}$ = 9.0 Hz, $^4J_{HF}$ = 4.8 Hz, 2H, 3,3'-bipy-position), 8.35 (ddd, $^3J_{HH}$ = 9.0 Hz, $^3J_{HF}$ = 8.0 Hz, $^4J_{HH}$ = 2.5 Hz, 2H, 4,4'-bipy-position), 6.41 (overlapping dd, 2H, $^3J_{HH}$ = 6.0 Hz, 4-Tol-position), 6.02 (d, $^3J_{HH}$ = 6 Hz, 2H, 3-Tol-position), 5.83 (t, $^3J_{HH}$ = 6.0 Hz, 1H, 5-Tol-position), 2.26 (s, 3H, 1-Tol-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 159.1 (d, $^1J_{CF}$ = 256 Hz, 5,5'-bipy-position), 150.7 (s, 2,2'-bipy-position) 144.6 (d, $^2J_{CF}$ = 33 Hz, 6,6'-bipy-position), 127.3 (d, $^2J_{CF}$ = 19 Hz, 4,4'-bipy-position), 125.2 (d, $^3J_{CF}$ = 7.0 Hz, 3,3'-bipy-position), 107.2 (s, 2-Tol-position), 90.9 (s, 4-Tol-position), 82.6 (s, 3-Tol-position), 79.9 (s, 5-Tol-position), 18.9 (s, 1-Tol-position). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6), -123.6 (s, 5,5'-position).

Characterisation of [Ru(η^6-*p*-cymene)(5,5'-difluorobipyridine)Cl][PF_6] – [11]

Synthesised from the starting dimer III and starting bipyridine C. Yield: 59 % (41.3 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange needles

Elemental: Anal. Calcd for $C_{20}H_{20}ClF_8N_2PRu$: C, 39.52; H, 3.32; N, 4.61. Found: C, 39.73; H, 2.99; N, 4.12. **HRMS (ESI$^+$):** m/z 463.0370 $[M - PF_6]^+$ (m_{calc} = 463.0327). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.66 (overlapping dd, $^3J_{HF}$ = 3.2 Hz, $^4J_{HH}$ = 2.5 Hz, 2H, 6,6'-bipy-position), 8.72 (dd, $^3J_{HH}$ = 9.0 Hz, $^4J_{HF}$ = 4.8 Hz, 2H, 3,3'-bipy-position), 8.36 (td, $^3J_{HH}$ = 9.0 Hz, $^3J_{HF}$ = 8.0 Hz, $^4J_{HH}$ = 2.5 Hz, 2H, 4,4'-bipy-position), 6.34 (d, $^3J_{HH}$ = 6.4 Hz, 2H, 5-cym-position), 6.06 (d, $^3J_{HH}$ = 6.4 Hz, 2H, 4-cym-position), 2.60 (sept, $^3J_{HH}$ = 7.0 Hz, 1H, 2-cym-position), 2.21 (s, 3H, 7-cym-position), 0.94 (d, $^3J_{HH}$ = 7.0 Hz, 6H, 1-cym-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 159.2 (d, $^1J_{CF}$ = 256 Hz, 5,5'-bipy-position), 150.6 (s, 2,2'-bipy-position), 144.5 (d, $^2J_{CF}$ = 32.5 Hz, 6,6'-bipy-position), 127.4 (d, $^2J_{CF}$ = 18.3 Hz, 4,4'-bipy-position), 124.2 (d, $^3J_{CF}$ = 8.0 Hz, 3,3'-bipy-position,), 105.1 (s, 3-cym-position), 104.5 (s, 6-cym-position), 86.8 (s, 5-cym-position), 83.5 (s, 4-cym-position), 30.4 (s, 2-cym-position), 21.8 (s, 1-cym-position), 18.4 (s, 7-cym-position). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 d, $^1J_{PF}$ = 711 Hz PF_6), -123.1 (s, 5,5'-position).

Characterisation of [Ru(η^6-hexamethylbenzene)(5,5'-difluorobipyridine)Cl][PF$_6$] – [12]

Synthesised from the starting dimer IV and starting bipyridine C. Yield: 59 % (43.1 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Red needles

Elemental: Anal. Calcd for $C_{22}H_{24}ClF_8N_2PRu$: C, 41.55; H, 3.80; N, 4.41. Found: C, 41.63; H, 3.46; N, 3.86. **HRMS (ESI$^+$):** m/z 491.0676 $[M - PF_6]^+$ (m_{calc} = 491.0640). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 8.89 (6,6'-bipy-position, 2H, t, $^3J_{HF}$ = 3.2 Hz, $^4J_{HH}$ = 2.3 Hz), 8.72 (3,3'-bipy-position, 2H, overlapping dd, $^3J_{HH}$ = 9.0 Hz, $^4J_{HF}$ = 4.8 Hz), 8.34 (4,4'-bipy-position, 2H, td, $^3J_{HH}$ = 9.0 Hz, $^3J_{HF}$ = 8.0 Hz, $^4J_{HH}$ = 2.3 Hz), 2.06 (18H, s, 2-hmb-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): δ (ppm) 159.5 (d, $^1J_{CF}$ = 257 Hz, 5,5'-position), 150.6 (s, 2,2'-position), 141.8 (d, $^2J_{CF}$ = 32 Hz, 6,6'-position), 127.3 (d, $^2J_{CF}$ = 19.5 Hz, 4,4'-position), 125.3 (d, $^3J_{CF}$ = 8 Hz, 3,3'-position,), 95.8 (s, 2-hmb-position) 15.0 (s, 1-hmb-position). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6), -122.7 (s, 5,5'-position).

Characterisation of [Ru(η^6-benzene)(5,5'-di(trifluoromethyl)bipyridine)Cl][PF$_6$] – [13]

Synthesised from the starting dimer I and starting bipyridine D. Yield: 60 % (45.0 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange prisms.

Elemental: Anal. Calcd for $C_{18}H_{12}ClF_{12}N_2PRu$: C, 33.17; H,1.86; N, 4.30. Found: C, 33.15; H, 1.80; N, 4.21. **LRMS (ESI$^+$):** m/z 507.00 $[M - PF_6]^+$ (m_{calc} = 506.96). ^{1}H NMR (400.13

MHz, d^6-DMSO): δ (ppm) 9.95 (s, 2H, 6,6'-position), 9.04 (d, $^3J_{HH}$ = 8.7 Hz, 2H, 3,3'-position), 8.86 (d, $^3J_{HH}$ = 8.7 Hz, 2H 4,4'-position), 6.35 (s, 6H, PhH). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 156.6 (s, 2,2'-position), 152.7 (q, $^3J_{CF}$ = 4.1 Hz, 6,6'-position,), 137.8 (q, $^3J_{CF}$ = 3.1 Hz , 4,4'-position,), 128.7 (q, $^2J_{CF}$ = 34 Hz, 5,5'-position,), 125.5 (s, 3,3'-position) 122.1 (q, $^1J_{CF}$ = 278 Hz , $\underline{C}F_3$,), 87.5 (s, PhH). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6,), -60.1 (s, CF_3 groups).

Characterisation of [Ru(η^6-toluene)(5,5'-di(trifluoromethyl)bipyridine)Cl][PF$_6$] – [14]

Synthesised from the starting dimer II and starting bipyridine D. Yield: 61 % (76.6 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange needles.

Elemental: Anal. Calcd for $C_{19}H_{14}ClF_{12}N_2PRu$: C, 34.28; H, 2.12; N, 4.21. Found: C, 34.34; H, 2.08; N, 4.15. **LRMS (ESI$^+$):** m/z 520.97 $[M - PF_6]^+$ (m_{calc} = 520.98). ^{1}H NMR (400.13 MHz, d^6-DMSO): δ (ppm) 9.84 (s, 2H, 6,6'-position), 9.01 (d, $^3J_{HH}$ = 8.5 Hz, 2H, 3,3'-position), 8.83 (d, $^3J_{HH}$ = 8.5 Hz, 2H 4,4'-position), 6.46 (overlapping dd, 2H, $^3J_{HH}$ = 6.0 Hz, 4-Tol-position), 6.11 (d, $^3J_{HH}$ = 6.0 Hz, 2H, 3-Tol-position), 5.84 (t, $^3J_{HH}$ = 6.0 Hz, 1H, 5-Tol-position), 2.25 (s, 3H, 1-Tol-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 156.7 (s, 2,2'-position), 152.4 (q, $^3J_{CF}$ = 4.2 Hz, 6,6'-position,), 137.7 (q, $^3J_{CF}$ = 3.5 Hz , 4,4'-position,), 128.7 (q, $^2J_{CF}$ = 35 Hz, 5,5'-position,), 125.4 (s, 3,3'-position) 122.1 (q, $^1J_{CF}$ = 275 Hz , $\underline{C}F_3$,), 108.5 (s, 2-Tol-position), 92.0 (s, 4-Tol-position), 81.9 (s, 3-Tol-position), 79.7 (s, 5-Tol-position), 19.0 (s, 1-Tol-position). **^{19}F{^{1}H} NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^1J_{PF}$ = 711 Hz, PF_6,), -60.7 (s, CF_3 groups).

Characterisation of [Ru(η^6-*p*-cymene)(5,5'-di(trifluoromethyl)bipyridine)Cl][PF$_6$] – [15]

Synthesised from the starting dimer III and starting bipyridine E. Yield: 62 % (81.4 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Yellow prisms.

Elemental: Anal. Calcd for $C_{22}H_{20}ClF_{12}N_2PRu$: C, 37.33; H, 2.12; N, 4.21. Found: C, 37.59; H, 2.71; N, 3.59. **HRMS (ESI$^+$):** m/z 563.0639 $[M - PF_6]^+$ (m_{calc} = 563.0263). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 9.83 (s, 2H, 6,6'-position), 9.04 (d, $^3J_{HH}$ = 8.5 Hz, 2H, 3,3'-position), 8.85 (d, $^3J_{HH}$ = 8.5 Hz, 2H, 4,4'-position), 6.45 (d, $^3J_{HH}$ = 6.1 Hz, 2H, 5-cym-position), 6.17 (d, $^3J_{HH}$ = 6.1 Hz, 2H, 4-cym-position), 2.62 (sept, $^3J_{HH}$ = 6.9 Hz, 1H, 2-cym-position), 2.20 (s, 3H, 7-cym-position), 0.97 (d, $^3J_{HH}$ = 6.9 Hz, 6H, 1-cym-position). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 156.4 (s, 2,2'-position), 152.3 (q, $^3J_{CF}$ = 4.2 Hz, 6,6'-position,), 137.8 (q, $^3J_{CF}$ = 3.3 Hz , 4,4'-position,), 128.9 (q, $^2J_{CF}$ = 34 Hz, 5,5'-position,), 125.7

(s, 3,3'-position) 122.0 (q, $^{1}J_{CF}$ = 273 Hz , $\underline{C}F_3$,), 105.7 (s, 3-cym-position), 104.9 (s, 6-cym-position), 87.2 (s, 5-cym-position), 83.3 (s, 4-cym-position), 30.4 (s, 2-cym-position), 21.8 (s, 1-cym-position), 18.4 (s, 7-cym-position). **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^{1}J_{PF}$ = 711 Hz, PF_6,), -60.4 (s, CF_3 groups).

Characterisation of [Ru(η^6-hexamethylbenzene)(5,5'-di(trifluoromethyl)bipyridine)Cl][PF_6] – [16]

Synthesised from the starting dimer IV and starting bipyridine E. Yield: 45 % (38.0 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Orange Needles.

Elemental: Anal. Calcd for $C_{22}H_{20}ClF_{12}N_2PRu$: C, 39.17; H, 3.29; N, 3.81. Found: C, 38.94; H, 2.93; N, 3.50. **HRMS (ESI$^+$):** m/z 591.0573 $[M - PF_6]^+$ (m_{calc} = 591.0576). **^{1}H NMR (**400.13 MHz, d^6-DMSO): δ (ppm) 9.16 (s, 2H, 6,6'-position), 9.08 (d, $^{3}J_{HH}$ = 8.4 Hz, 2H, 3,3'-position), 8.88 (d, $^{3}J_{HH}$ = 8.4 Hz, 2H 4,4'-position), 2.08 (18H, s, 2-hmb-position). **$^{13}C\{^{1}H\}$ NMR** (100.57 MHz, d^6-DMSO): 156.8 (s, 2,2'-position), 149.8 (q, $^{3}J_{CF}$ = 4.3 Hz, 6,6'-position), 137.7 (q, $^{3}J_{CF}$ = 2.9 Hz , 4,4'-position,), 129.1 (q, $^{2}J_{CF}$ = 34 Hz, 5,5'-position,), 126.1 (s, 3,3'-position) 122.6 (q, $^{1}J_{CF}$ = 274 Hz , $\underline{C}F_3$,), 96.7 (s, 2-hmb-position) 15.4 (s, 1-hmb-position). **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^{1}J_{PF}$ = 711 Hz, PF_6,), -60.8 (s, CF_3 groups).

Characterisation of [Ru(η^6-hexamethylbenzene)(6,6'-difluorobipyridine)Cl][PF_6] – [17]

Synthesised from the starting dimer I and starting bipyridine E. A synthesis at 10-fold lower molar concentration was used for this complex due to lack of starting bipyridine. Yield: 55 % (3.5 mg). Recrystallised *via* vapour diffusion Et_2O:Acetone. Appearance: Yellow prisms.

Elemental: Anal. Calcd for $C_{18}H_{12}ClF_{12}N_2PRu$: C, 34.83; H, 2.19; N, 5.08. Found: C, 34.61; H, 2.33; N, 4.84. **LRMS (ESI$^+$):** m/z 407.02 $[M - PF_6]^+$ (m_{calc} = 406.97). **^{1}H NMR** (400.13 MHz, d^6-DMSO): δ (ppm) 8.63 (d, $^{3}J_{HH}$ = 7.9 Hz, 2H, 3,3'-position), 8.52 (dd, $^{3}J_{HH}$ = 8.5 Hz, $^{3}J_{HH}$ = 7.9 Hz, 2H, 4,4'-position), 7.91 (d, $^{3}J_{HH}$ = 8.5 Hz, 2H, 5,5'-position), 6.32 (s, 6H, PhH). Lack of sample restricted analysis to ^{1}H and ^{19}F NMR analysis. **$^{19}F\{^{1}H\}$ NMR** (376.50 MHz, d^6-DMSO): δ (ppm) -69.8 (d, $^{1}J_{PF}$ = 711 Hz, PF_6), -49.7 (s, 6,6'-position)

3.3.4. Synthesis of Catalytic Substrates for Chapter 6

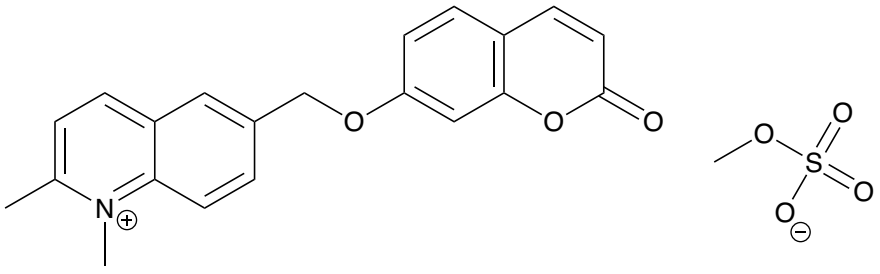
Fellow group member Jamie Klein performed all synthesis of catalytic substrates, including synthetic optimisation and characterisation required for this chapter.

Synthesis of this substrate was reported by Ward et al.,[89] however, the method was modified by Jamie Klein and a much improved yield was reported.

Transfer Hydrogenation Substrate (TH substrate)

(2-methyl-6-quinolinyl)methanol (100 mg, 0.58 mmol, 1 eq.), methanesulfonic anhydride (110 mg, 0.63 mmol, 1.1 eq.) and DMAP (7.3 mg, 0.06 mmol, 0.1 eq.) were weighed into a dry Schlenk tube and dried under vacuum for 1 h. The system was put under a N_2 atmosphere CH_2Cl_2(5 mL), CH_3CN (1 mL) and anhydrous pyridine (47 μL, 0.58 mmol. 1 eq.) were added and the mixture stirred at room temperature for 30 min., after which the solvent was removed in vacuo. Umbelliferone (113 mg, 0.70 mmol, 1.2 eq.) and caesium carbonate (331 mg, 1.02 mmol, 1.75 eq.) were dried in vacuo before addition of CH_3CN (10 mL). After stirring for 20 min., the mesylated quinoline was taken up in CH_3CN (10 mL) and small portions were added slowly to the umbelliferone mixture. After complete addition, the mixture was stirred at room temperature overnight. The mixture was filtered, the solvent was removed in vacuo and the residue taken up in EtOAc (20mL). The mixture was washed with 1N $NaOH_{(aq)}$ (3 x 20 mL), after which the combined aqueous phase was extracted with EtOAc (2 x 20 mL). The product (142 mg, 0.45 mmol, 77% yield) was obtained as a white solid after flash chromatography (EtOAc:Hexane 2:1).

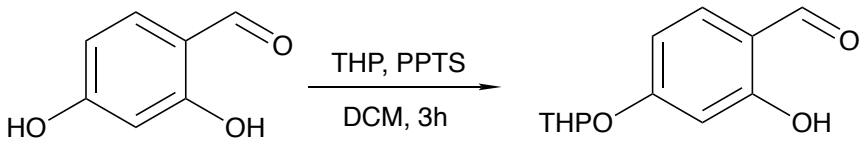
The coupled product (132 mg, 0.42 mmol, 1 eq.) was dissolved is dry toluene (30 mL) under a N_2 atmosphere. Dimethyl sulphate (395 μL, 4.2 mmol, 10 eq.) was added and the mixture stirred for 24 h at 100 °C. The product was obtained by filtration and washing with Et_2O and hexane as an off-white solid (57 mg, 0.13 mmol, 31% yield).



^{1}H NMR (400.13 MHz, d^6-DMSO): δ (ppm) 9.11 (d, J = 8.6 Hz, 1H), 8.64 (d, J = 9.2 Hz, 1H), 8.47 (s, 1H), 8.29 (dd, J = 9.2, 1.9 Hz ,1H), 8.13 (d, J = 8.6 Hz, 1H), 8.02 (d, J = 9.5 Hz, 1H), 7.69 (d, J = 8.6 Hz, 1H), 7.16 (d, J = 2.2 Hz, 1H), 7.11 (dd, J = 8.6, 2.4 Hz, 1H), 6.33 (d, J = 9.5 Hz, 1H), 5.57 (m, 2H), 4.45 (s, 3H), 3.37 (s, 3H) 3.08 (s, 3H).

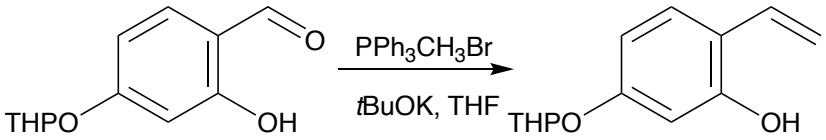
Olefin Metathesis Substrate (OM substrate)

Step 1: Protection to OM Intermediate 1:



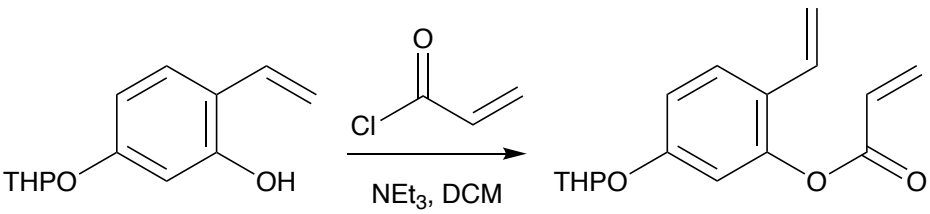
3,4-dihydro-2H-pyran (DHP) (3.02 g, 3.30 mL, 36.0 mmol) was added to a solution of 2,4-dihydroxybenzaldehyde (4.14 g, 30.0 mmol) and pyridinium *p*-toluenesulfonate (754 mg, 3.00 mmol) in dry DCM (60 cm^3). The mixture was stirred for 3 h at room temperature under N_2 atmosphere. The mixture was quenched with saturated $NaHCO_{3\ (aq.)}$ (90 cm^3), the aqueous phase extracted with EtOAc (2 x 90 cm^3) and the combined organic phases dried with $MgSO_4$. The solvent was removed under reduced pressure to yield the product as colourless crystals (6.0 g, 90%), which was directly used for subsequent steps.

Step 2: Wittig reaction to OM Intermediate 2:



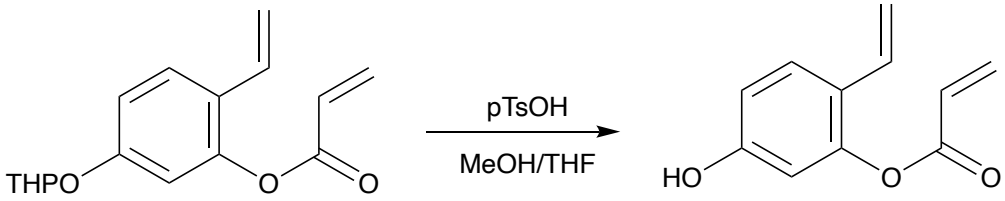
Under a N_2 atmosphere, potassium tert-butoxide (3.0 g, 27.0 mmol) was added to a suspension of methyltriphenyl-phosphonium bromide (4.8 g, 13.4 mmol) in dry THF (50 cm^3). The mixture was stirred at room temperature for 1 h, then cooled to -78 °C. OM Intermediate 1 (2.0 g, 9.0 mmol) was dissolved in dry THF (30 cm^3) then added dropwise to the reaction mixture. The mixture was warmed to room temperature and stirred for a further 30 min. The reaction was quenched with saturated $NH_4Cl_{(aq.)}$ (100 cm^3) and the aqueous layer extracted with EtOAc (2 x 50 cm^3). The combined organic phases were washed with brine and dried over Na_2SO_4. The solvent was reduced under vacuum and the product (1.3 g, 5.7 mmol, 63%) was obtained as colourless solid after purification by flash chromatography (hexane:EtOAc, 3:1).

Step 3: Nucleophilic addition of acyl-chloride and alcohol to OM Intermediate 3:



Triethylamine (1.4 cm^3, 10 mmol) was added to a solution of OM Intermediate 2 (1.1 g, 5.0 mmol) in dry DCM (60 cm^3) under N_2 atmosphere. Acrolyl chloride (0.6 cm^3, 7.5 mmol) was slowly added and the mixture stirred for 30 min. at room temperature. The solvent was removed under reduced pressure and the resulting oil was taken straight to further reaction due to limited stability.

Deprotection to OM Substrate:



p-Toluenesulfonic acid monohydrate (77 mg, 0.4 mmol) was added to a mixture of OM Intermediate 3 (~1.1 g, 4.1 mmol) in MeOH/THF 1:1 (100 cm^3). After stirring at 50 °C for 1 h, the mixture was concentrated under reduced pressure and the residue taken up in DCM (80 cm^3). The mixture was washed with brine and dried over $MgSO_4$. The product (0.46 g, 2.4 mmol) was obtained as a colourless oil after purification *via* flash chromatography (cyclohexane/EtOAc 5:2). The product was rapidly stored as a ~0.2 M solution in DMSO (12 cm^3) at -20°C.

^{1}H NMR (400 MHz, d_6-DMSO) δ = 9.90 (s, 1H), 7.51 (d, $^3J_{HH}$ = 8.6 Hz, 1H), 6.73 (dd, $^3J_{HH}$ = 8.6, 2.4 Hz, 1H), 6.61-6.50 (m, 3H), 6.43 (dd, $^3J_{HH}$ = 17.3, 10.2 Hz, 1H), 6.16 (dd, $^3J_{HH}$ = 10.3, $^2J_{HH}$ = 1.3 Hz, 1H), 5.64 (dd, $^3J_{HH}$ = 17.6, $^2J_{HH}$ = 1.1 Hz, 1H), 5.14 (dd, $^3J_{HH}$ = 11.2, $^2J_{HH}$ = 1.1 Hz, 1H).

3.3.5. Synthesis of Ruthenium Complexes for Chapter 7

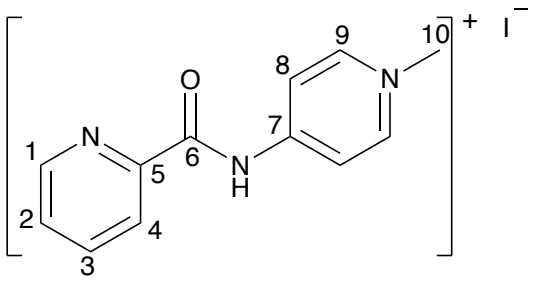
Synthetic Procedure and Characterisation of L1 and [18]

Synthesis of ligand previously reported by Albrecht et al,[115] however, the ruthenium complex is novel.

Picolinic acid (1.23 g, 10 mmol) was refluxed at 85 °C in $SOCl_2$ (10 mL, 0.14 mol) for 3 h under an N_2 atmosphere. The excess of $SOCl_2$ was removed under reduced pressure in a well-ventilated fume hood. The residue was dissolved in dry THF (40 mL) and transferred *via* cannula to a solution of the 4-amino pyridine (0.94 g, 10 mmol) in THF (50 mL) with the presence of NEt_3 (2.5 mL, 18 mmol). The mixture was refluxed for at 65 °C 48 h, filtered, and

the solvent was removed under reduced pressure. The brown solid residue was purified by column chromatography (SiO_2) in (4 : 1, EtOAc : Hexane) to yield a white solid. (1.40 g, 65 %).

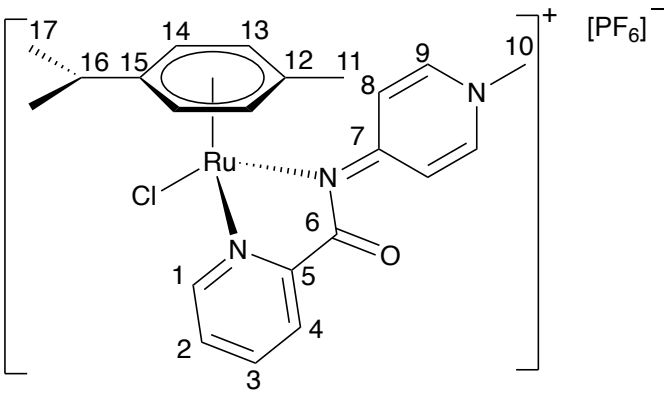
The purified white solid (398 mg, 2.0 mmol) was dissolved in MeCN (15 mL) in a Schlenk tube. MeI (187 µL, 3.0 mmol) was added, and the reaction mixture was stirred for 12 h at 80 °C. The solution was cooled to room temperature and concentrated to 5 mL. Et_2O (50 mL) was added, and the formed yellow precipitate was filtered and washed with cold Et_2O to yield the product L1 (578 mg, 85 %).



^{1}H NMR (400.13 MHz, d^6-DMSO): δ (ppm) 12.01 (s, 1H, NH), 8.83 (d, $^3J_{HH}$ = 4.7 Hz, 1H, H1), 8.80 (d, , $^3J_{HH}$ = 6.6 Hz Hz, 2H, H9), 8.50 (d, $^3J_{HH}$ = 6.6 Hz, 2H, H8), 8.24 (d, $^3J_{HH}$ = 7.8 Hz, 1H, H4), 8.15 (td, $^3J_{HH}$ = 7.8 Hz, $^4J_{HH}$ = 1.7 Hz 1H, H3), 7.80 (ddd, $^3J_{HH}$ = 7.8 Hz, $^3J_{HH}$ = 4.7 Hz, $^3J_{HH}$ = 1.4 Hz 1H, H2), 4.20 (s, 3H, H10). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 164.7 (C6), 151.4 (C7), 148.8 (C5), 148.2 (C1), 145.9 (C9), 138.5 (C3), 128.2 (C2), 123.5 (C4), 115.8 (C8), 46.5 (C10).

Synthetic Procedure and Characterisation for Complex [18]

Ligand L1 (86 mg, 0.25 mmol), [Ru(*p*-cym)Cl_2]$_2$ (91 mg, 0.15 mmol), $AgPF_6$ (126 mg, 0.5 mmol), and K_2CO_3 (69 mg, 0.50 mmol) were stirred in dry MeCN (10 mL) under N_2 atmosphere for 3 h. The volatiles were removed under reduced pressure, and dry CH_2Cl_2 (20 mL) was added. The mixture was filtered through a short pad of Celite and washed with more CH_2Cl_2 (20 mL). All filtrates were combined and evaporated to dryness to obtain the Ru complex [18]. The crude product was purified by recrystallization from CH_2Cl_2/pentane (35 mg, 37 %)



Elemental: Anal. Calcd for $C_{22}H_{25}ClF_6N_3OPRu$: C, 42.01; H, 4.01; N, 6.68. Found: C, 41.97; H, 4.10; N, 6.75. **HRMS (ESI$^+$):** m/z 484.0721 $[M - PF_6]^+$ (m_{calc} = 484.0730). **^{1}H NMR (**400.13 MHz, d^6-DMSO): δ (ppm) 9.36 (d, $^3J_{HH}$ = 5.2 Hz, 1H, H1) 8.59 (d, $^3J_{HH}$ = 6.8 Hz, 2H, H9), 8.28 (d, $^3J_{HH}$ = 6.8 Hz, 2H, H8), 8.19 (td, $^3J_{HH}$ = 7.8 Hz, 1H, H4), 8.00 (d, $^3J_{HH}$ = 7.8 Hz 1H, H3), 7.81 (ddd, $^3J_{HH}$ = 7.8 Hz, $^3J_{HH}$ = 4.7 Hz, $^3J_{HH}$ = 1.4 Hz, 1H, H2), 5.87 (d, $^3J_{HH}$ = 5.8 Hz, 1H, H13), 5.72 (d, $^3J_{HH}$ = 5.8 Hz, 1H, H14), 5.60 (d, $^3J_{HH}$ = 5.8 Hz, 1H, H13), 5.49 (d, $^3J_{HH}$ = 5.8 Hz, 1H, H14), 4.18 (s, 3H, H10), 2.45 (sept, $^3J_{HH}$ = 5.8, 1H, H16), 2.21 (s, 3H, H11), 1.02 (d, $^3J_{HH}$ = 6.9 Hz, H17), 0.95 (d, $^3J_{HH}$ = 6.9 Hz, H17). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 168.3 (C6), 165.4 (C7), 155.4 (C1), 152.4 (C9), 144.3 (C5), 139.8 (C4), 128.7 (C2), 125.9 (C3), 123.4 (C8), 102.7 (C15), 101.9 (C12), 85.9 (C13), 84.8 (C13), 83.1 (C14), 82.8 (C14), 46.0 (C10), 30.4 (C16), 21.9 (C17), 21.5 (C17), 18.3 (C11).

Synthetic Procedure and Characterisation of L2 and [19]

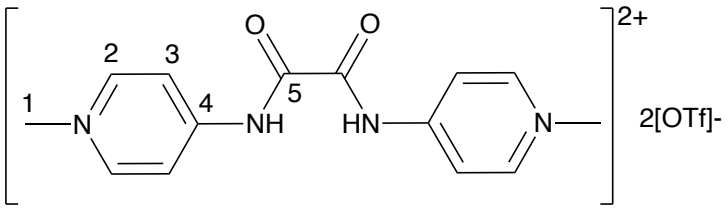
This ligand and complexes are novel, however, build upon a synthesis developed by Albrecht et al.[116]

Oxalyl chloride (1.29 mL, 15 mmol) and freshly distilled CH_2Cl_2 (20 mL) were added into a nitrogen purged Schlenk flask. 4-aminopyridine (2.82 g, 30 mmol) was added followed by the slow addition of freshly distilled NEt_3 (5 mL, 36 mmol) under a nitrogen stream. The reaction mixture was stirred at room temperature overnight and then and a brown solid collected on the filter pad. Continual washes with H_2O (3 x 25 mL) yielded an off-white solid which was dried on a high vacuum for 12 hrs (1.66 g, 46 %).

The off white solid (0.94 g, 3.88 mmol) was dissolved in freshly distilled CH_2Cl_2 (30 mL) and stirred at 0 °C for 10 mins. Methyl trifluoromethanesulphonate (2.12 mL, 19.9 mmol) was added slowly to the reaction mixture and the reaction was stirred for a following 30 mins at 0

°C. The reaction mixture was stirred for a following 2 hrs at room temperature and then filtered and collected (1.84 g, 83 %).

In some purifications, counter-ion exchange was performed here from $[CF_3SO_3]^-$ to $[PF_6]^-$ through dissolving the product in a minimum amount of MeOH and adding 5 Eq. of NH_4PF_6.

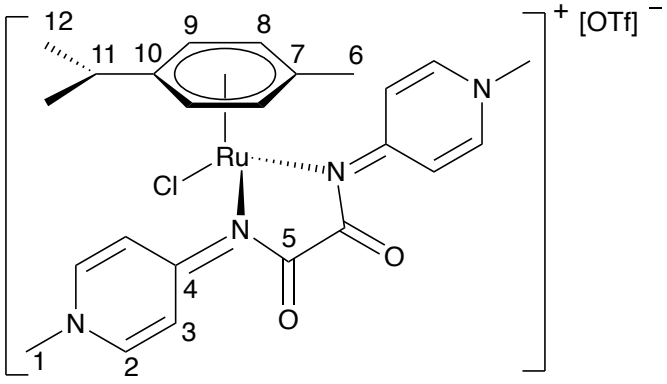


HRMS (ESI$^+$): m/z 136.0627 $[M - 2(OTf)]^{2+}$ (m_{calc} = 136.0631). **^{1}H NMR (**400.13 MHz, d^6-DMSO): δ (ppm) 12.18 (s, 1H, NH), 8.82 (d, $^3J_{HH}$ = 7.5 Hz, 1H, H2), 8.39 (d, $^3J_{HH}$ = 7.5 Hz, 1H, H3), 4.22 (s, 6H, H1). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 159.4 (C5), 151.1 (C4), 146.7 (C2), 116.9 (C3), 47.7 (C1).

Synthetic Procedure and Characterisation for Complex [19]

L2 (162 mg, 0.33 mmol), Ag_2O (250 mg, 1.09 mmol), $[RuCl_2(p\text{-cymene})]_2$ (97.9 mg, 0.16 mmol) and NMe_4Cl (104 mg, 0.95 mmol) were dissolved in CH_3CN (10 ml) and stirred at room temperature for 18 hrs in the absence of light. The resulting mixture was filtered through a short pad of Celite, concentrated in vacuo, and precipitated from Et_2O (50 ml). The resulting solid was dissolved in the minimum amount of MeOH (ca. 2 ml) and filtered through and concentrated to 1 ml, and addition of Et_2O (50 ml) afforded the product as a yellow solid (101.2 mg, 49 %).

Crystallisation quality crystals were achieved through a vapour diffusion method (DCM/MeCN : Pentane).



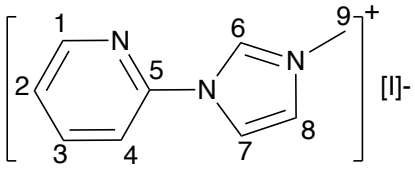
HRMS (ESI$^+$): m/z 541.0919 $[M - OTf]^+$ (m_{calc} = 541.0944). **^{1}H NMR (**400.13 MHz, d^6-DMSO): δ (ppm) 8.50 (d, $^3J_{HH}$ = 7.5 Hz, 4H, H2), 8.41 (d, $^3J_{HH}$ = 7.5 Hz, 4H, H3), 5.32 (d, J = 6.0 Hz, 2H, H9), 5.23 (d, J = 6.0 Hz, 2H, H8), 4.12 (s, 6H, H1), 2.38 (sept, 1H, H11), 2.22 (s, 3H, H6), 1.03 (d, J = 6.0 Hz, 6H, H12).

Synthetic Procedure and Characterisation of L3 and [20]

Synthesis of ligand and ruthenium complex previously reported by Hartinger et al.[117] Crystal structure report by Albrecht et al.[118]

A mixture of imidazole (1.00 g, 15 mmol) and 2-bromopyridine (3.80 g, 24 mmol) was suspended in a solution of degassed dimethyl sulfoxide (30 mL) and stirred for 5 min. Potassium phosphate (8.50 g, 40 mmol), l-proline 0.50g, 4 mmol), imidazole (1.00g , 15 mmol) and copper(I) iodide (0.40 g, 2 mmol) were added to the mixture and kept at 110 °C for 48 h. The crude product was filtered and washed with CH_2Cl_2 (3 × 10 ml), resulting in a brown filtrate which was collected. The organic phase was washed with water in an extraction funnel until the aqueous layer turned clear. Anhydrous sodium sulphate was used to dry the organic phase. The solvent was removed under reduced pressure and dried under reduced pressure to afford a brown solid.

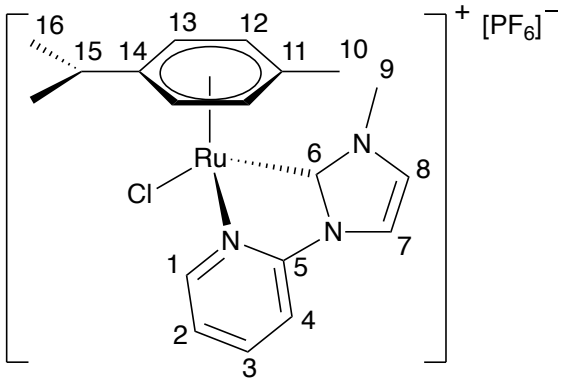
A mixture of the brown solid (0.50 g, 2.21 mmol) and iodomethane (0.55 ml, 8.83 mmol) in tetrahydrofuran (25 ml) was refluxed for 24 hrs to afford the product as an off white solid. The solvent was removed under reduced pressure, and the crude product was purified by trituration with acetone, affording the pure product (0.84 g, 2.9 mmol, 87%).



^{1}H NMR (400.13 MHz, d^6-DMSO): δ (ppm) 10.03 (s, 1H, H6), 8.65 (d, J = 4.4 Hz, 1H, H1), 8.51 (t, J = 1.9 Hz, 1H, H8), 8.23 (t , J = 8.1 Hz, 1H, H3), 8.00 (d, J = 8.1 Hz, 1H, H4), 7.97 (t, J = 1.9 Hz, 1H, H7), 7.65 (overlapping dd, J = 8.1 Hz, J = 4.4 Hz, 1H, H2), 3.98 (s, 3H, H9). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 149.1 (C1), 146.3 (C5), 140.6 (C3), 135.5 (C6), 125.2 (C8), 124.7 (C7), 118.9 (C2), 114.1 (C4), 36.5 (C9).

Synthetic Procedure and Characterisation for Complex [20]

L3 (149 mg, 0.52 mmol) and silver hexafluorophosphate (52 mg, 0.21mmol) were added to dry CH_3CN (6 ml) and MeOH (1 ml) in a dry Schlenk flask under N2. The flask was sealed, and the mixture was stirred in darkness at 55 °C for 3 h. Ag_2O (144 mg , 0.62 mmol) was added and the mixture stirred in darkness for a further 24 hrs at 65 °C. A solution of $[RuCl^2(\eta^6\text{-}p\text{-cymene})]_2$ (129 mg , 0.21 mmol) in dry 1,2-dichloroethane (3 ml) and dichloromethane (9 ml) was added and stirred at 70 °C in darkness overnight. The resultant suspension was filtered over Celite and the filtrate collected and reduced to dryness. The orange solid was then washed with toluene (3x10 ml) followed by pentane (10 ml). Purification was performed *via* column chromatography over silica (MeOH:CH_2Cl_2 1:25). The solvent of the collected fractions was evaporated, and the crude compound was dissolved in a minimal amount of MeOH and precipitated by addition of diethyl ether. The precipitate was dried in vacuo to afford the $Ru(\eta^6\text{-}p\text{-cymene})$ complex (45 mg, 0.097 mmol, 23%).



HRMS (ESI$^+$): m/z 461.1156 $[M - PF_6]^+$ (m_{calc} = 461.1172).**^{1}H NMR (**400.13 MHz, d^6-DMSO): δ (ppm) 9.33 (d, J = 5.7 Hz, 1H, H1), 8.41 (d, J = 1.8 Hz, 1H, H7), 8.34 (dd, J = 8.0 Hz, 1H, H3), 8.15 (d, J = 8.0 Hz, 1H, H4), 7.79 (d, J = 1.8 Hz, 1H, H8), 7.50 (dd, J = 8.0 Hz, J = 5.7 Hz 1H, H2), 6.39 (d, J = 5.9 Hz, 1H, H13), 6.36 (d, J = 5.9 Hz, 1H, H12), 6.16 (d, J = 5.9 Hz, 1H, H13), 5.73 (d, J = 5.9 Hz, 1H, H12), 4.10 (s, 3H, H9), 2.34 (sept, J = 6.8 Hz, 1H, H15), 2.10 (s, 3H, H10), 0.84 (d, J = 6.8 Hz, 3H, H16), 0.83 (d, J = 6.8 Hz, 3H, H16). **^{13}C{^{1}H} NMR** (100.57 MHz, d^6-DMSO): 184.6 (C6), 156.3 (C1), 151.9 (C5), 142.1 (C3), 127.1 (C8), 123.6 (C2), 116.9 (C7), 112.9 (C4), 99.5 (C14), 93.2 (C11), 91.3 (C13), 91.2 (C12), 87.0 (C13), 81.7 (C12), 38.3 (C9), 31.1 (C10), 22.6 (C15), 19.1 (C16).

3.4. Protein Production

C2Am was kindly donated by the research group of Dr Gonçalo Bernardes, and used without further purification.

A plasmid for expression of Ubq K63C was kindly donated by the group of Dr Gonçalo Bernardes.

A plasmid for expression of Tev protease was kindly donated by the group of Prof. Jane Clarke.

3.4.1. Buffers

All buffers were prepared with ultrapure water, and filter-sterilised by passing through a 0.2 μm cellulose acetate membrane (Sartorius Stedium Biotech). The pH of each buffer was checked by using a PHM210 pH-meter (Radiometer Analytical) following a two-point calibration with standard solutions (pH 4.005 and 10.012, Hannah Instruments). If necessary, the pH was adjusted using the appropriate set of acids and bases. The following buffers were used in protein production and incubations:

- Sodium phosphate (NaP_i)
- Trisaminomethane (Tris)
- Diethanolamine (DEA)
- 2-morpholin-4-ylethanesulfonic acid (MES)
- Sodium dodecyl sulfate (SDS)

3.4.2. Reagents Used

Ampicillin stock solutions were prepared at a concentration of 100 mg/mL in water, filter-sterilised, aliquoted and stored at −20 °C.

Kanamycin stock solutions were prepared to a concentration of 50 mg/mL in water, filter-sterilised, aliquoted and stored at −20 °C.

IPTG (isopropyl β-D-1-thiogalactopyranoside) stock solutions were prepared at a concentration of either 0.4 M or 1 M in water, filter-sterilised, aliquoted and stored at −20 °C.

DTT (dithiothreitol) stock solutions were prepared at a concentration of either 0.1 M OR 1 M in water, filter-sterilised, aliquoted and stored at −20 °C. Stock solutions were used to reduce disulphide bonds at final concentrations ranging from 1–5 mM.

TCEP (tris(2-carboxyethyl)phosphine) stock solutions were prepared at a concentration of either 0.1 M OR 1 M in water, filter-sterilised, aliquoted and stored at −20 °C. Stock solutions were used to reduce disulphide bonds at final concentrations ranging from 1–5 mM.

LB (lysogeny broth) medium was prepared by dissolving pre-prepared capsules which constitute 10 g of tryptone, 5 g of yeast extract and 10 g of NaCl in 1 L of water, and autoclaving for 20 min at 120 °C.

2×TY (double yeast-tryptone) medium was prepared by dissolving pre-prepared capsules which constitute 6 g of tryptone, 10 g of yeast extract and 5 g of NaCl (added as capsules, MP Biomedicals) in 1 L of water, and autoclaving for 20 min at 120 °C.

TB (terrific broth) modified medium was prepared by dissolving 12 g of tryptone, 24 g of yeast extract, 8 mL of glycerol, 2.31 g of K_2HPO_4, 12.54 g of K_2HPO_4, 3.75 g of aspartic acid and 0.5 g of $MgSO_4$ and dissolving in 1 L of sterilised H_2O. The media was autoclaved for 20 min at 120 °C.

3.4.3. SDS-Polyacrylamide Gel Electrophoresis

SDS-PAGE analysis was performed using the NuPAGE mini-gel system with 4-12 % bis-tris gels. Samples were treated with SDS sample buffer with reducing agents. Gels were run in MES running buffer, and stained using 2.5 % (w/v) Coomassie blue in methanol-acetic acid.

3.4.4. Protein Expression and Purification

Method 1: Ubiquitin K63C Expression and Purification

2.5 µL of ubiquitin plasmid #86589 (Addgene) was transformed into 50 µL of super-competent Bl21-Gold(DE3) cells by electroporation. Following this 1 mL of 2 x YT media was added to the cuvette which was incubated at 37 ºC for 45 minutes. 200 µL of cells was spread onto ampicillin plates and incubated at 37 ºC overnight to allow for colonies to grow. A single colony from this culture was transferred to 20 mL of LB media with ampicillin (100 µg/mL) and incubated overnight at 37 ºC and 180 rpm.

This culture was used to inoculate 2 L of LB medium containing ampicillin (100 µg/mL) and cells were grown to OD600 of 0.6 (37 ºC, 180 rpm). Then expression of ubiquitin K63C was induced with 1mM IPTG and protein was expressed at 18 ºC overnight. Then, the cells were harvested by centrifugation at 4 ºC and 4000 g for 30 minutes. Supernatant was discarded and cells were resuspended in 50 mM NaP_i buffer (pH 7.4) containing 250 mM of NaCl and, 25

mM imidazole and 1 mM TCEP and a protease inhibitor cocktail. Cells were lysed by sonication and cell fragments were collected and removed by centrifugation. The supernatant was purified by IMAC chromatography with a linear gradient of imidazole from 25 mM to 500 mM. Samples containing ubiquitin K63C were combined and cleavage of the fusion partner maltose binding protein (MBP) was conducted by incubating a 1:20 molar ratio of TEV protease at 25 °C 24 hrs. Following cleavage, the protein was purified by IMAC chromatography to remove MBP and TEV protease. Further purification was often required (confirmed by SDS-gel electrophoreses) and this was performed through a Mono S cation exchange (high performance) in 20 mM MES buffer pH = 5.0 developed with a linear gradient of 0 – 200 mM KC1. The flow-through was pooled, concentrated and buffer exchange to 50 mM NaP_i buffer pH 8 was performed. The Beer-Lambert law was applied with an extinction coefficient of $\varepsilon_{280} = 1490\ M^{-1}\ cm^{-1}$ to measure protein concentration.

Method 2: Cytochrome *b*$_{562}$ Expression and Purification

1 - 2 µL of the required cytochrome plasmid was transformed into 50 µL of super-competent Bl21-Gold(DE3) cells by electroporation. Following this 1 mL of 2 x YT media was added to the cuvette which was incubated at 37 ºC for 45 minutes. 200 µL of cells was spread onto ampicillin resistant plates and incubated at 37 ºC overnight to allow for colonies to grow. A single colony from this culture was transferred to 20 mL of 2 x YT media with ampicillin (100 µg/mL) and incubated overnight at 37 ºC and 180 rpm to give a pre-culture.

10 mL of the pre-culture was inoculated into a 1 L culture 2 x YT media with ampicillin (100 µg/mL) in a 2 L flask and this was shaken at 200 rpm at 37 ºC, conditions considered to be microaerobic. Innoculae were grown to an optical density at 600 nm of about 1 before induction with 0.1 mM IPTG, and the cultures continued into stationary phase before harvest by centrifugation at 18 h after induction. Then, the cells were harvested by centrifugation at 4 ºC and 4000 g for 30 minutes. Following cell growth, cytochrome *b*$_{562}$ is found in both the periplasmic fractions and the culture medium, however, for practical reasons only cytochrome *b*$_{562}$ proteins in the periplasm was harvested, therefore supernatant was discarded.

Protein was harvested from E. coli cells by fractionation of the pelleted cells into spheroplasts and soluble periplasm contents using the lysozyme/EDTA method. The apo- and holo-cytochromes in these fractions were separated on a Q-Sepharose anion exchanger in DEA buffer pH = 8.5 developed with a linear gradient of 0 – 200 mM KC1. Further purification was often required (confirmed by SDS-gel electrophoreses) and this was performed on a Superdex

S75 gel filtration column in 50 mM NaP_i buffer with 1 mM TCEP. The Beer-Lambert law was applied with an extinction coefficient of $\varepsilon_{280} = 2980\ M^{-1}cm^{-1}$ o measure protein concentration.

Method 3: Tev Protease Expressions and Purification

1 - 2 μL of the plasmid pMHTDelta238 (pBR322 origin with T5 promoter and Kanamycin resistance) was transformed into 50 μL of super-competent Bl21-Gold(DE3) cells by electroporation. Following this 1 mL of LB media was added to the cuvette which was incubated at 37 ºC for 45 minutes. 100 μL of cells was spread onto 2 x TY kanamycin resistant plates and incubated at 37 ºC overnight to allow for colonies to grow. 10 mL of pre-culture with 10 μL kanamycin stock was pre-warmed to 37 ºC and inoculated with a generous scrape from the overnight colony and incubated at 37 ºC for 4 hrs. 1 L of media with 1 mL of kanamycin stock was pre-warmed to 37 ºC and inoculated with 10 mL of pre-culture, and this was shaken (200 rpm at 37 ºC). Innoculae were grown to an optical density at 600 nm of about 0.6 before induction with 0.1 mM IPTG and protein expression was carried out for 3 hr (200 rpm at 37 ºC). Then, the cells were harvested by centrifugation at 8 ºC and 5000 rpm for 15 mins. The cell pellet was re-suspended in 20 mM Tris buffer (pH 7.5, NaCl 125 mM) and sonicated on ice (15 s on, 45 s off, 5 min) and centrifuged at 8 ºC and 18,000 rpm for 45 mins. The supernatant was purified by IMAC chromatography with a linear gradient of imidazole from 20 mM to 500 mM. This purification process also cleaved the maltose binding protein from the His7-Tev product. The protein solution and 5 mM EDTA was then dialysed with sterilised H_2O to remove imidazole and then concentrated. The Beer-Lambert law was applied with an extinction coefficient of $\varepsilon_{280} = 32{,}400\ M^{-1}\ cm^{-1}$ to measure protein concentration.

Method 4: Preparation of purchased cysteine mutants

Papain (102 mg crude powder) was dissolved in 50 mM NaP_i buffer (7 mL) by vortexing for several minutes. The solution mixture was filtered through a 0.22 μM filter cartridge and flow through collected. DTT (6.1 mg, 0.04 mmol) was added and the mixture incubated at 37 °C for 30 min. The resulting reduced protein solution (1 mL) was filtered through a de-salting column pre-equilibrated with phosphate buffer (2 mL, 7 kDa MWCO, GE Healthcare). The sample was eluted with 1.5 mL phosphate buffer, aliquoted and flash frozen. The Beer-Lambert law was applied with an extinction coefficient of $\varepsilon = 52{,}770\ M^{-1}\ cm^{-1}$ to measure protein concentration.

An identical procedure was used to prepare the human serum albumin stock solutions. The Beer-Lambert law was applied with an extinction coefficient of $\varepsilon_{280} = 34{,}445\ M^{-1}\ cm^{-1}$ to measure protein concentration.

3.4.5. Protein Sequences

Ubq K63C

SAQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSD
YNIQCESTLHLVLRLRGG

M_w – 8567 Da

Cytochrome b_{562}

ADLEDNMETLNDNLKVIEKADNAAQVKDALTKMRAAALDAQKATPPKLEDKSPDS
PEMKDFRHGFDILVGQIDDALKLANEGKVKEAQAAAEQLKTTRNAYHQKYR

WT M_w – 11780 Da

R98C M_w – 11727 Da

L10C H102M M_w – 11764 Da

Human serum albumin

DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCVAD
ESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNL
PRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC
QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFP
KAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK
PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARR
HPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL
FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAE
DYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETF
TFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADD
KETCFAEEGKKLVAASQAALGL

M_w – 66472 Da

M_w with 17 Disulphide Bridges – 66438

C2Am S95C

GSPGISGGGGGILDSMVEKLGKLQYSLDYDFQNNQLLVGIIQAAELPALDMGGTSDP
YVKVFLLPDKKKKFETKVHRKTLNPVFNEQFTFKVPYCELGGKTLVMAVYDFDRFS
KHDIIGEFKVPMNTVDFGHVTEEWRDLQSAEK

M_w – 16222 Da

Tev Protease

AIAHHHHHHHGESLFKGPRDYNPISSSICHLTNESDGHTTSLYGIGFGPFIITNKHLFRR
NNGTLLVQSLHGVFKVKDTTTLQQHLVDGRDMIIIRMPKDFPPFPQKLKFREPQREE
RICLVTTNFQTKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVG
IHSASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMVK
PEEPFQPVKEATQLMNE

M_w – 28136 Da

3.5. Incubation Protocols

3.5.1. Incubations with Small Molecules

Deuterated Phosphate Buffer

10 mM deuterated phosphate buffer solution was prepared using potassium dihydrogen orthophosphate in D_2O and adjusted to different pD values with 1N NaOD. All pD values are corrected pH meter readings. Covington et al. stated that "the operational pH of buffer solutions in heavy water at 25 °C, measured with the glass electrode can be converted into a pD value by adding 0.41 (molar scale)".[119]

pK_a Measurements

The different buffers in a pD range of 4.7 to 11.2 were incubated with Complexes [1] – [4], [10] and [11] for 2 hours at 298 K and ^1H and $^{19}F\{^1H\}$ NMR spectra recorded.

Biomolecule interactions

The incubations between ruthenium complexes and amino acids were performed in either a 1.5 or 3 molar excess of amino acid / glutathione. Stock complex of each ruthenium complex were prepared in DMF and diluted with a solution of amino acid in D_2O phosphate buffer pD = 7.2, to a final concentration of DMF 1 % v/v. $^{19}F\{^1H\}$ NMR spectra were recorded at 2, 4, 8 and 24 hours, with mass spectra recorded at 8 and 24 hours.

3.5.2. Incubations with Proteins

To an eppendorf tube with NaP_i (50 mM, pH 8.0) and DMF (10% of total volume), an aliquot of a stock solution of protein (final concentration either 10 μM, 50 μM or 200 μM) was added. Afterwards, a solution of the ruthenium complex (1 to 20 eq.) in DMF was added and the resulting mixture was vortexed for 10 seconds. The reaction was mixed for up to 24 h, at 37

°C. At desired time points, an aliquot was taken and diluted to a final concentration of 5 μM with water and analysed by LC-MS.

For modifications with N-ethyl maleimide, from a stock 10 mM solution in DMF, 1 Eq. of reactant was added and the mixture vortexed for 10 secs, followed by incubation for 30 mins at 37 °C. Modification was confirmed by LC-MS.

3.6. Catalytic Experiments

Sample Preparation

Ruthenium-protein hybrids prepared and tested for catalytic activity in Chapter 6 were purified by anion exchange chromatography.

Upon modification, identified by LC-MS, protein – ruthenium hybrids were purified by anion exchange chromatography. The incubation was buffer exchanged into DEA (20 mM, pH 8.5), thus removing some of the excess ruthenium complex, and concentrated to the required volume for anion exchange. The anion exchange was performed with a linear gradient of 0 – 200 mM KCl. and protein components analysed by LC-MS.

3.6.1. Transfer Hydrogenation Assay

Ruthenium Dimers and Complexes

Ruthenium compounds were dissolved in DMF to yield 5 mM stock solutions. These were diluted 1:5 with water fresh before each experiment. 80 μL of water, 10 μL of ruthenium solution, 10 μL of buffer (500 mM sodium phosphate, 1 M formate which upon 1 in 10 dilution gave a pH 8.0) and 1 μL substrate stock solution (100 mM in DMSO) were pipetted into a clear-bottom 96 well plate (final concentrations 100 μM ruthenium compound, 100 mM formate, 1 mM substrate). The wells were sealed with tape and the fluorescence measured over 16 h.

Protein-Ruthenium Hybrids

90 μL of each protein sample (in water) of known concentration was pipetted into a clear-bottom 96 well plate. 10 μL of buffer (500 mM sodium phosphate, 1 M formate which upon 1 in 10 dilution gave a pH 8.0) and 1 μL of substrate stock solution (100 mM in DMSO) was added to each well (final concentrations 100 mM formate, 1 mM substrate). The wells were sealed with tape and the fluorescence was measured over 16 h.

3.6.2. Olefin Metathesis Assay

Grubbs 2nd Generation Catalyst

A stock solution of Grubbs 2nd generation catalyst was dissolved in DCM to yield 5 mM stock solutions. These were diluted 1:5 with water fresh before each experiment. 80 µL of water, 10 µL of ruthenium solution, 10 µL of buffer (500 mM sodium phosphate, which upon 1 in 10 dilution gave a pH 7.0) and 1 µL substrate stock solution (100 mM in DMSO) were pipetted into a clear-bottom 96 well plate (final concentrations Grubbs 2nd generation catalyst, 100 mM formate, 1 mM substrate). The metathesis substrate was dissolved in DMSO immediately before use and stored as a solid at -20 °C. The wells were sealed with tape and the fluorescence measured over 16 h.

Protein-Ruthenium Hybrids

90 µL of each protein sample (in water) of known concentration was pipetted into a clear-bottom 96 well plate. 10 µL of buffer (500 mM sodium phosphate, which upon 1 in 10 dilution gave a pH 7.0) and 1 µL substrate stock solution (100 mM in DMSO) were pipetted into a clear-bottom 96 well plate (final concentrations Grubbs 2nd generation catalyst, 100 mM formate, 1 mM substrate). The metathesis substrate was dissolved in DMSO immediately before use and stored as a solid at -20 °C. The wells were sealed with tape and the fluorescence measured over 16 h.

Chapter 4

4. A Fluorinated Probe for Quantifying Ruthenium – Amino Acid Coordination

4.1. Introduction

This chapter outlines the development of a quantitative method to characterise the coordination of Ru(II) arene complexes to amino acids and simple peptides. As discussed, the ability to quantify the coordination of Ru(II) arene complexes to preferred biological binding partners could further understanding in two key areas of research: (i) ruthenium artificial metalloenzymes and (ii) ruthenium metallo-pharmaceuticals.

Ru(II) arene complexes have been extensively studied due to their potential as anti-cancer and anti-microbial complexes.[27,120] Currently, the major limitation to the use of these complexes as pharmaceuticals is understanding their speciation within a complex biological environment. Ru(II) arene complexes have labile ligands which can exchange with a number of different biological nucleophiles, and identifying these biological nucleophiles is fundamental to understanding their cellular response. Proteins are known targets for a number of Ru(II) arene anti-cancer complexes,[30,31] they contain a wide variety of possible metal binding sites that can form strong coordination bonds to ruthenium. By inhibiting a specific protein's function *via* metal coordination, it is possible to induce a cascade of cellular events that result in cell apoptosis.

Through incorporating fluorinated ligands to ruthenium in simple complexes, the aim was to enable the use of ^{19}F NMR spectroscopy as a direct reporter of the solution-based behaviour of the metal coordination environment. Throughout this chapter, the quantitative nature of using ^{19}F NMR spectroscopy to monitor the speciation of Ru(II) arene complexes will be emphasised; individual species can be identified and quantified in complex biological environments *via* this method.

Understanding how to localise and quantify the coordination of Ru(II) arene complexes to a specific protein site is important in the development of datively coordinated ArMs. As discussed, one future goal of this research, is to develop a ruthenium ArM where the ruthenium

cofactor is localised to a specific amino acid through a dative coordination bond. This requires analytical methods to confirm that coordination to the desired side chain residue has occurred, and that there is no unbound complex.

4.1.1. Methods of Probing Ruthenium-Protein Coordination

As discussed in Chapter 1, previous attempts to understand the speciation of Ru(II) arene complexes with proteins have focussed on using mass spectrometry and X-ray crystallographic methods. Although these methods are useful, they have limitations in that neither are quantitative, and neither are directly relevant to the context of in situ biological activity, since crystallography probes the sample in the solid phase, and mass spectrometry is a gas phase method.

NMR spectroscopy is a quantitative, sensitive, direct reporter of behaviour in solution. However, using 1-Dimensional ^{1}H NMR to follow the speciation of ruthenium complexes has proved testing for anything other than the simplest examples. The large number of observable proton signals together with their chemical similarity and hence proximity in the spectra severely restricts the usefulness of ^{1}H NMR in this context. Instead ^{15}N and ^{13}C are commonly used to examine the structure and dynamics of proteins *in vivo*.[121]

4.1.2. Fluorine NMR

^{19}F NMR spectroscopy can be used as a sensitive probe of specific metabolic signals when a fluorine atom is incorporated into a biomolecule as a reporter.[122,123] Using ^{19}F NMR spectroscopy to probe interactions of biological molecules is effective because: ^{19}F is nearly totally absent from biology, it is small in size (van der Waals radius of 1.47 Å, hydrogen = 1.20 Å), it is monovalent (same as hydrogen) and it has favourable NMR spectroscopic characteristics (spin 1/2, 100 % natural abundance, high gyromagnetic ratio and strong dipolar coupling allowing for ^{19}F-^{19}F and ^{1}H-^{19}F nuclear Overhauser effects to be measured).[124] The work of Dalvit et al. using ligand-based fluorine NMR screening methods to rapidly identify biologically active fluorinated organic compounds has demonstrated the usefulness of being able to detect a ^{19}F NMR signal from a small molecule reporter in complex environments.[125,126]

4.2. Prior Work in the Group

Building upon recent research in the group exploring the reactivity of Ru(II) arene bipyridine complexes, Dr Michael O'Neill synthesised complex [9], **Figure 4.1**, and probed its reactivity with amino acids using ^{19}F NMR spectroscopy.

Using the Ru(II) arene fragment as a starting point, there are three further coordination sites, two of which were taken up by the bipyridine ligands in this work. Bipyridine ligands are known to form good coordination bonds to ruthenium centres that are stable to dissociation upon coordination of the ruthenium centre to a protein. The final ligand was chosen as a chloride ligand, which acts as a labile ligand that can undergo ligand exchange. Furthermore, ruthenium bipyridine complexes are relatively simple to synthesise and purify *via* a salt metathesis reaction. There is an extensive literature, pioneered by Prof. Harry Gray, exploring electron transfer processes in proteins, using ruthenium bipyridine complexes.[100] Using a screening approach where all amino acids with nucleophilic side chain residues were incubated with complex [9], it was identified that different ruthenium-amino acid adducts, gave unique ^{19}F NMR spectra, with different chemical shifts and peak morphologies.[127]

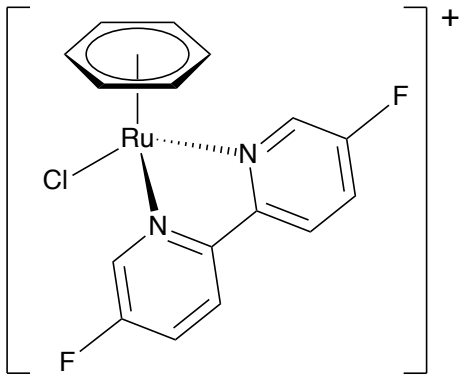
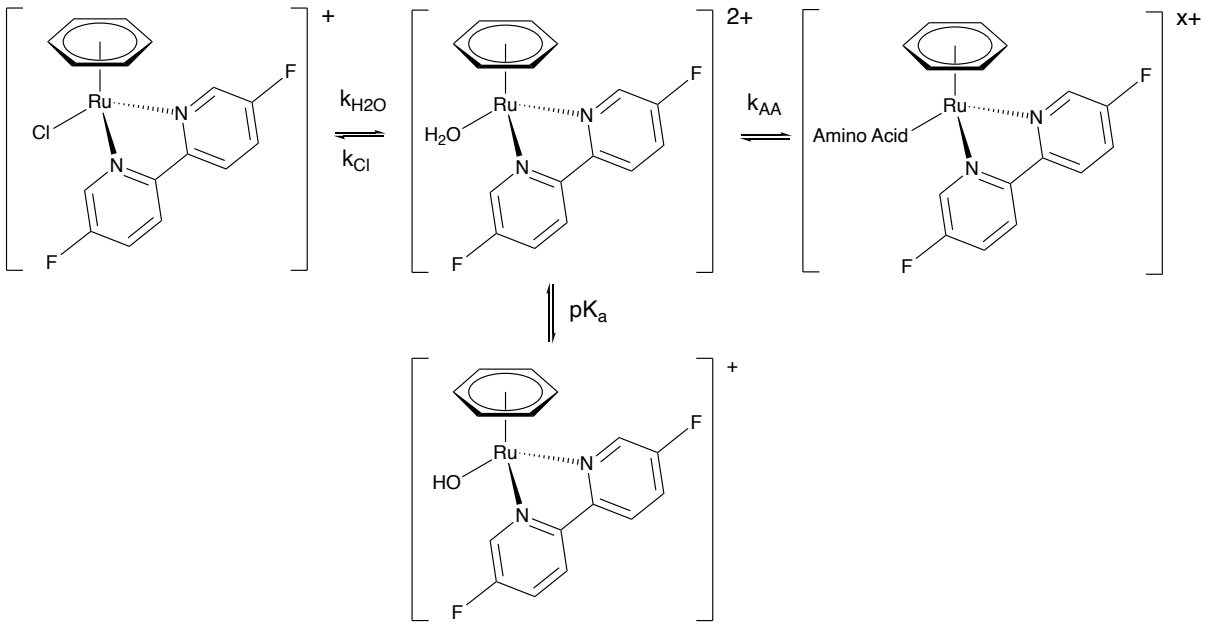


Figure 4.1: Structure of complex [9]. [PF_6] counter ion removed for clarity.

The hydrolysis of the chloride ligand in these Ru(II) arene complexes plays an important role in their reactivity, however, monitoring this reaction in solution in a quantitative manner is challenging. In an aqueous mixture of complex [9] with a Lewis basic amino acid, the following complexes are present in solution [9]-Cl, [9]-H_2O, [9]-OH and [9]-Amino acid, **Scheme 4.1**. ^{19}F chemical shift values for the chloro, aquo and hydroxide species as well as a number of amino acid bound species were reported. Through a number of competition studies, it was found that the cysteine thiol and the carboxylate ligand formed the most stable complexes when coordinated to ruthenium.



Scheme 4.1: The equilibrium that exists upon exposure of [9] to a nucleophilic amino acid in water.

4.3. Results and Discussion

The work of previous group members identified that a ^{19}F NMR spectroscopic method could give an alternative solution-based approach to understanding the behaviour of biologically active Ru(II) arene complexes. To build upon these findings, a number of key areas were identified to research:

1. Can the stability and speciation of Ru(II) arene complexes be quantified in phosphate buffer using ^{19}F NMR spectroscopy?
2. Can the stability and sensitivity of the Ru(II) arene fluorinated probe be improved by changing the position of fluorination, fluorine substituent and arene ligand?
3. How does changing the arene ligand effect the concentration of different ruthenium-amino acid adducts in competition experiments?
4. Can quantitative parameters, e.g. rates of exchange of Ru(II) arene complexes with biomolecules be measured using ^{19}F NMR spectroscopy?

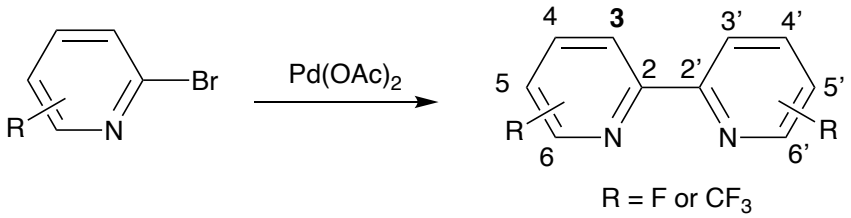
4.3.1. Synthesis of Ru(II) Arene Bipyridine Complexes

A suite of novel Ru(II) arene bipyridine complexes have been synthesised and characterised using a variety of techniques. In the following section, all Ru(II) arene bipyridine complexes synthesised throughout this PhD have been collated and the synthetic challenges and advances

summarised. For complete synthetic protocols and characterisation of these Ru(II) arene bipyridine complexes, refer to **Section 3.3.3**.

Synthesis of Functionalised Bipyridines

A number of literature preparations are available for the preparation of functionalised bipyridines,[128] however, after achieving relatively poor yields using copper catalysed, Ullman-type coupling reactions, the fluorinated bipyridines were initially synthesised using **Method 1.**[129] Ultimately this method continued to generate poor yields of the products so an alternative method, **Method 2,** was identified, based on the work of Lee et al.[112] Both methods are outlined in **Section 3.3.1**, and both used palladium catalysed homocoupling of fluorinated bromopyridines, **Scheme 4.2,** however the solvent systems and additives differed significantly.



Scheme 4.2: Palladium catalysed homocoupling of fluorinated bromopyridines, along with the favoured labelling system used for the bipyridine ligands.

The fluorinated bipyridines 3,3'-difluorobipyridine (3,3'-FBipy) – B, 5,5'-difluorobipyridine (5,5'-FBipy) – C, 5,5'-di(trifluoromethyl) bipyridine, (5,5'-TFMBipy) – D, 6,6'-difluorobipyridine (6,6'-FBipy)– E, were synthesised and characterised by NMR spectroscopy, ESI-MS, X-ray crystallography and, where possible, elemental analysis.

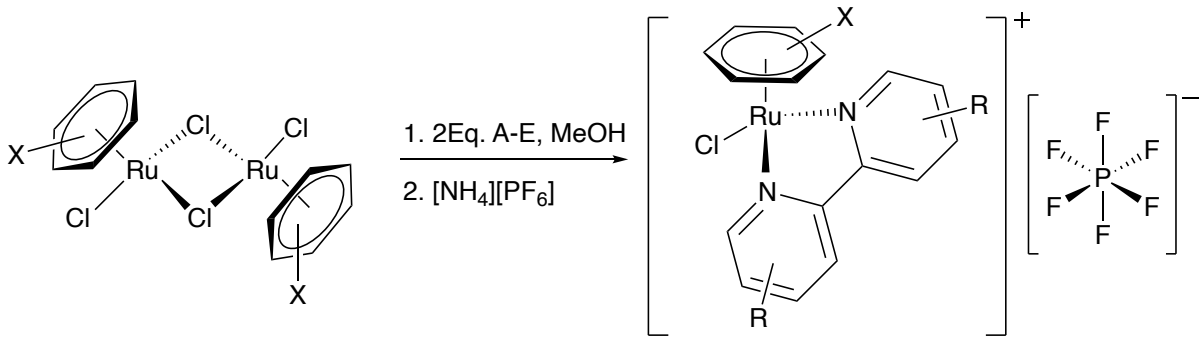
Synthesis of the Ru(II) Arene Dimers

At the beginning of this research project, there was a cost benefit to synthesising the ruthenium dimers using hydrated $RuCl_3$ with the appropriate hexacyclic diene to give a chloro-bridged arene dimer.[113] Hexacyclic dienes were used to form the dimers, $[Ru(\eta^6\text{-benzene})Cl_2]_2$ – **I**, $[Ru(\eta^6\text{-tolyl})Cl_2]_2$ – **II** and $[Ru(\eta^6\text{-}$ *p*$\text{-cymene})Cl_2]_2$ - **III**. $[Ru(\eta^6\text{-hexamethylbenzene})Cl_2]_2$ – **IV**, was synthesised through an arene exchange reaction. As the cost of $RuCl_3$ increased, purchasing dimers I and III became logical.

Synthesis of Ru(II) Arene Bipyridine Complexes

Synthesis of a suite of Ru(II) arene bipyridine piano stool complexes was performed by reacting 1 Eq. of **I** – **IV** with 2 Eq. of A-E. Upon addition of $[NH_4][PF_6]$ the complexes [1] – [17] were isolated as their $[PF_6]$ salts, **Scheme 4.3, Table 4.1.** Crude products were purified

via a number of crystallisation techniques, and characterised by multinuclear NMR spectroscopy, ESI-MS, X-ray crystallography and, where possible, elemental analysis, see **Section 3.3.3**. The crystal structures of all complexes [1] – [17] are shown in **Figure 4.2,** with key bond lengths and angles summarised in **Table 4.2,** and complete crystallographic parameters given in **Appendix A**.



Scheme 4.3: Synthetic procedure to complexes [1] – [17].

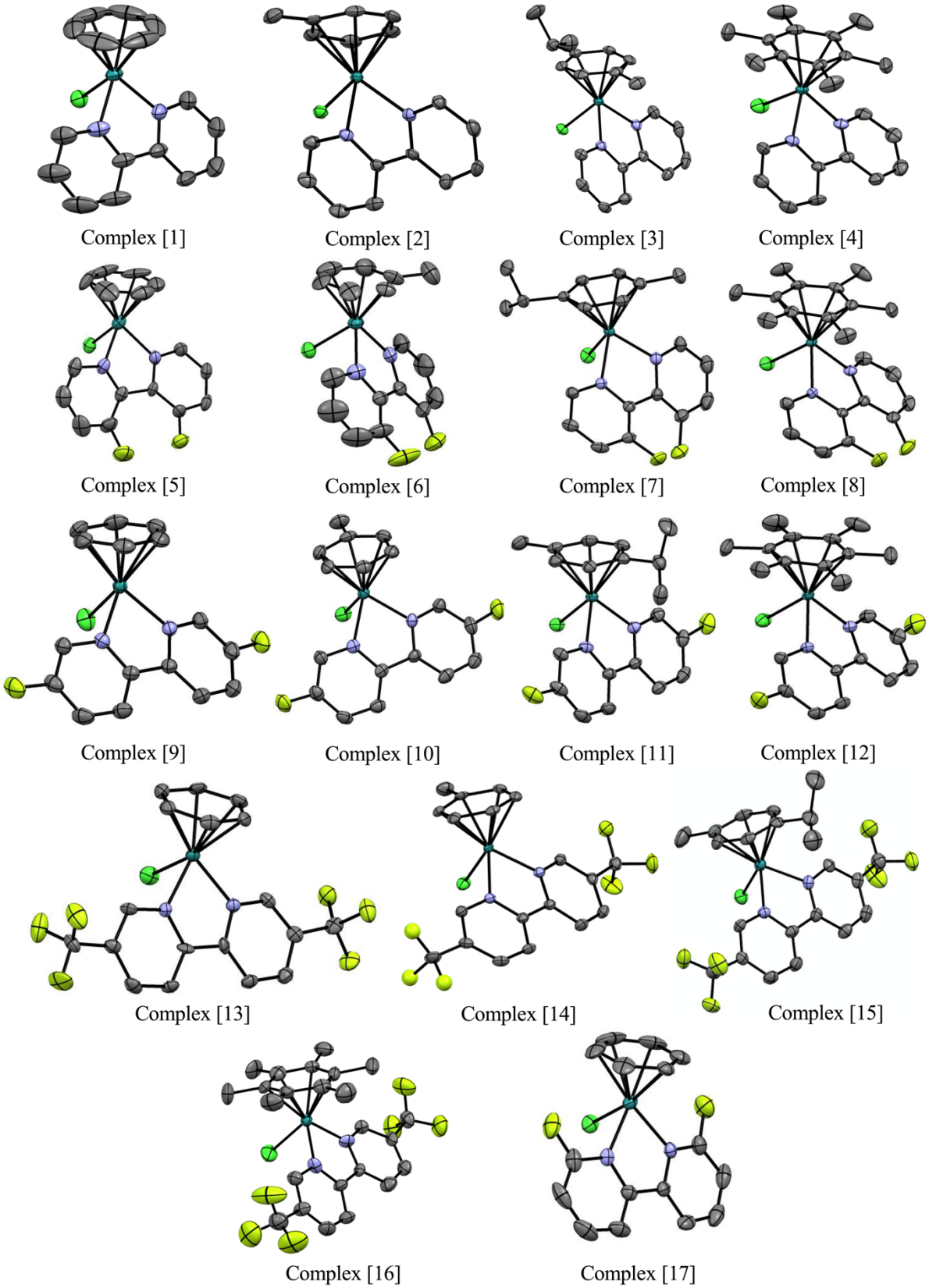


Figure 4.2: Molecular structures of the complexes [1] – [17] . Anionic salt and hydrogen atoms removed for clarity and thermal ellipsoids set to 30 %. Complex [14], one CF_3 group is rotationally disordered, therefore modelled as three separate components with a common isotropic displacement parameter.

Table 4.1: Summary of the starting materials and how they combine to form complexes [1] – [17].

	Benzene – I	**Tolyl – II**	***p*-Cymene – III**	**Hexamethyl-benzene – IV**
Bipy – A	[1]	[2]	[3]	[4]
3,3'-FBipy – B	[5]	[6]	[7]	[8]
5,5'-FBipy – C	[9]	[10]	[11]	[12]
5,5'-TFMBipy – D	[13]	[14]	[15]	[16]
6,6'-FBipy – E	[17]	Not synthesised	Not synthesised	Not synthesised

Table 4.2: Key bond lengths for complexes [1] – [17].

	Ru–Cl / Å	**Ru–N / Å**	**Ru–Centroid / Å**
[1]	2.3671(9)	2.071(2)	1.674
[2]	2.3967(9)	2.079(3)	1.691
[3]	2.3988(10)	2.085(4)	1.680
[4]	2.3990(14)	2.091(4)	1.707
[5]	2.3970(20)	2.071(7)	1.683
[6]	2.3680(20)	2.071(6)	1.697
[7]	2.4066(11)	2.078(3)	1.699
[8]	2.3972(12)	2.083(5)	1.708
[9]	2.3754(7)	2.083(2)	1.681
[10]	2.3772(12)	2.086(4)	1.688
[11]	2.3898(7)	2.090(2)	1.687
[12]	2.3809(6)	2.105(2)	1.703
[13]	2.3910(20)	2.079(1)	1.684
[14]	2.3892(8)	2.089(3)	1.692
[15]	2.3908(12)	2.091(3)	1.687
[16]	2.3850(7)	2.085(17)	1.723
[17]	2.3855(13)	2.122(4)	1.680

Analysis of Key Bond Length and Angles

In complexes [1]-[17] the key ligands that have been functionalised are the bipyridine and the arene. The arene ligand is capable of donating electron density to the metal through σ and π type interactions. The bipyridine ligand is a σ-donor ligand through the N sp^2 lone pair directed away from the ring skeleton. Both ligands are also capable of accepting electron density via π-

backdonation. The electron-donating properties of the different arene ligands vary; on moving from benzene → toluene → *p*-cymene → hexamethylbenzene the electron donation to the metal would be expected to become more significant because of the increased inductive effect of the organic substituents on the aromatic ring.

Analysis of the Ru–N, Ru–Cl and Ru–arene bond lengths in complexes [1]-[4], shows a general trend (in some cases statistically significant) that these bond lengths increase moving from the benzene complex [1] to the hexamethylbenzene complex [4]. In particular, the average Ru–N bond length increases from 2.071(2) Å for the benzene complex [1] to 2.091(4) Å for the hexamethylbenzene complex [4]. Moving from complex [1] to complex [4] there is increased arene electron donation, therefore less electron donation is required from the bipyridine, thus weakening the sigma bonding component and increasing measured bond lengths. This general trend in Ru–N bond lengths is mirrored when changing the bipyridine ligand to 5,5-difluorobipyridine, as shown by increase in Ru–N bond length for complexes [9] to [12], which between complex [9] and [12] is statistically significant. The Ru–N bond lengths do not follow the same trend for complexes with the 5,5-di(trifluoromethyl)bipyridine ligand. In this series of complexes the *p*-cymene complex [15], has a longer Ru–N bond length (2.091(3) Å) than the hexamethylbenzene complex [16] (2.085(17)), however this is not a large enough difference considering the error to be determined statistically significant.

Functionalising the bipyridine ligand with increasingly electron withdrawing substituents moving from fluorine to trifluoromethyl leads to slight variations in Ru–Cl bond lengths. In particular, for the benzene complexes, there is a significant, yet slight, increase in Ru–Cl bond length from 2.3671(9) Å for complex [1] to 2.3754(7) Å for complex [9] to 2.3910(20) Å for complex [13]. However, these trends in Ru–Cl bond lengths as affected by bipyridine functionalisation, are not reliably corroborated when the arene ligands are functionalised. Additionally, Ru–N and Ru–arene bond lengths do not always follow a predictable trend upon variation of the bipyridine ligands. Although complete structural analysis has been performed for these complexes the challenges in correlating structural information to solution-based behaviour have been highlighted throughout this research. There is precedence in the literature for bond length variations correlating with lability, however, from extensive experience of the solution-based behaviour of this family of complexes such comparisons cannot be relied upon absolutely. Other factors, such as sterics, mechanistic features of ligand exchange and subtle

difference between solid- and solution-state structures, will no doubt also affect the lability of ligands.

4.3.2. Quantitative Analysis of Speciation in Aqueous Phosphate Buffer

Before exposing these fluorinated ruthenium complexes to amino acids, the speciation in deuterated aqueous phosphate buffer was explored. With a number of titratable groups involved in these systems, examining the ligand exchange behaviour of these complexes around conditions of pH relevant to biological conditions requires strong buffering. Phosphate buffer was chosen for its relevance to biological conditions. However, it was identified that at high phosphate concentration, and in the absence of other possible ligands, phosphate anions can compete as ligands to the ruthenium. Phosphate has been shown to be a poor ligand for Ru(II) arene complexes by Sadler et al. who identified that the structurally related [Ru(η^6-arene)(en)Cl]$^+$ binds to the phosphate in the nucleobase 5-GMP before being displaced by the guanine N7 over time.[130]

The ^{19}F NMR spectrum of the 5,5'-FBipy complex [10] in 10 mM phosphate buffer, **Figure 4.3 (Left),** identified three different fluorine environments, which were attributed to three different molecular species, [10]-OD_2, [10]-Cl and [10]-Phosphate. The presence of [10]-Phosphate, was confirmed by ESI-MS, **Figure 4.3 (Right)**, which showed higher ion counts for this species at a higher phosphate concentration.

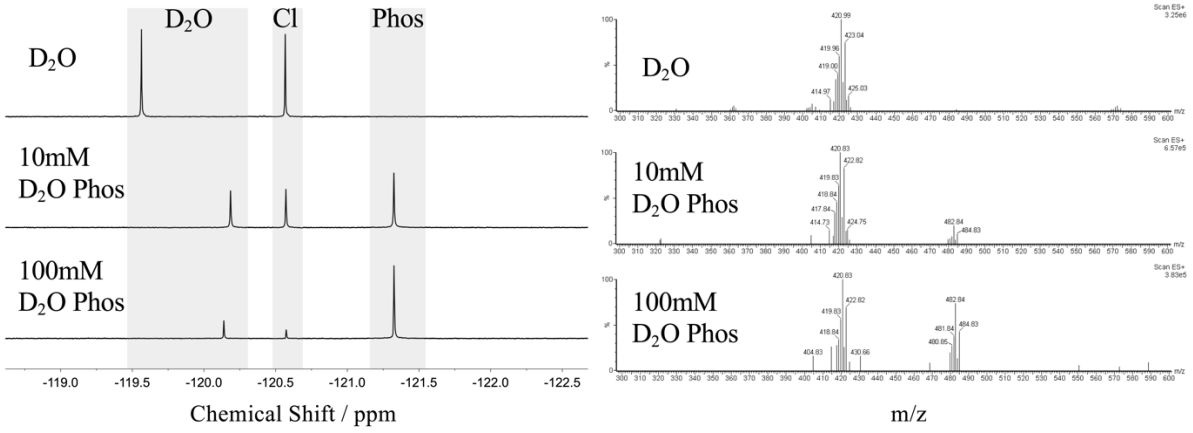


Figure 4.3: (Left): ^{19}F{^{1}H} NMR spectra of complex [10] incubated in D_2O and buffered D_2O of different phosphate concentration (2 mM Ru, pD = 7.2 (when buffered), 2 hr, 310 K). **(Right):** ESI-MS spectra recorded of the solution mixture when complex [10] is incubated in D_2O and buffered D_2O, expected masses quoted in **Appendix B, Table B.1.**

Identifying and quantifying the presence of these three different complexes over time was possible using ^{19}F NMR spectroscopy. **Figure 4.4** plots the integration of the chloride, water and phosphate peaks of complex [9] in 10 mM phosphate buffer. Upon placing complex [9]-

Cl in phosphate buffer there is a rapid formation of the aquo-bound species [9]-OD_2, which is a more labile ligand as there is a weaker ionic contribution to bonding in the neutral D_2O ligand. [9]-OD_2 will undergo ligand exchange to form a phosphate bound species, [9]-Phos. Being able to accurately quantify this solution behaviour provided a foundation upon which more complex biological environments could be understood.

The hydrolysis of [9]-Cl is consistent with a 1st order process where k = 0.000836 ± 0.000005 s^{-1}. This is presumably following pseudo 1st order kinetics, where the water is in large excess. After approximately 4500 s, the reaction reaches an apparent equilibrium where the concentration of the three species remains relatively constant.

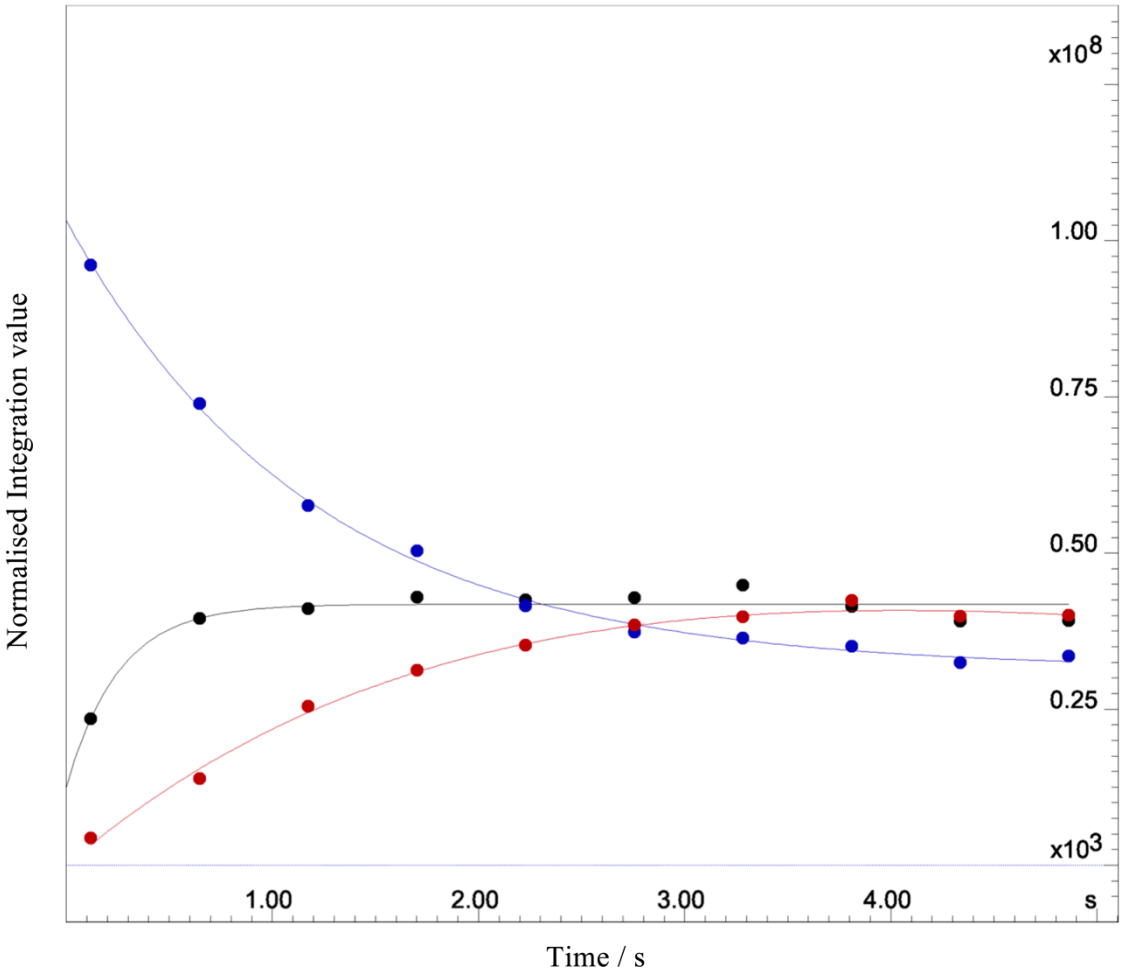


Figure 4.4: Plot of the integration values of the chloride, aquo and phosphate peaks of complex [9] in 10 mM phosphate buffer (2 mM Ru, pD = 7.2, 1 hr, 310 K). Blue trace - [9]-Cl, black trace - [9]- D_2O, red trace - [9]-Phos.

Through observing the reactions of [9]-[12] in phosphate buffer, it is clear that the benzene and toluene complexes are not indefinitely stable. If left in phosphate buffer, the solution will change colour from a pale yellow to blue which intensifies over time, **Figure 4.5**. It has not

been possible to isolate and identify any new species by NMR, mass spectrometry or X-ray crystallography, however it is possible that some form of ruthenium phosphate dimers or clusters are forming, which could involve oxidation of the metal in the complex to Ru(III).[131–133]

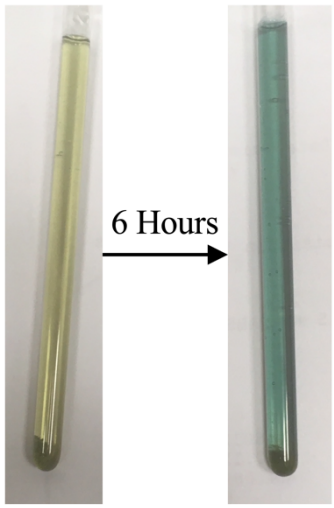
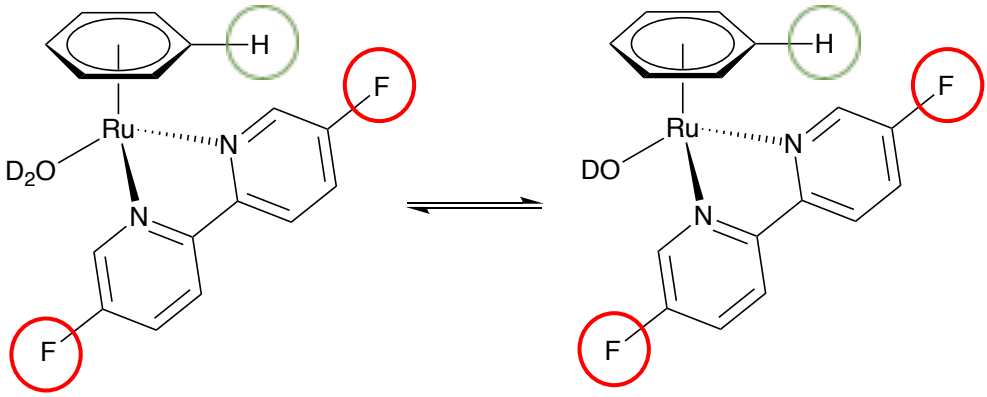


Figure 4.5: Pictures of complex [9] incubated in 10mM D_2O phosphate buffer after 0 hrs and after 6 hrs at 310 K.

pH Dependency of the Ruthenium Aquo Complex

In deuterated aqueous solution, an equilibrium exists between the Ru-[OD_2] and Ru-[OD] species, **Scheme 4.4**.



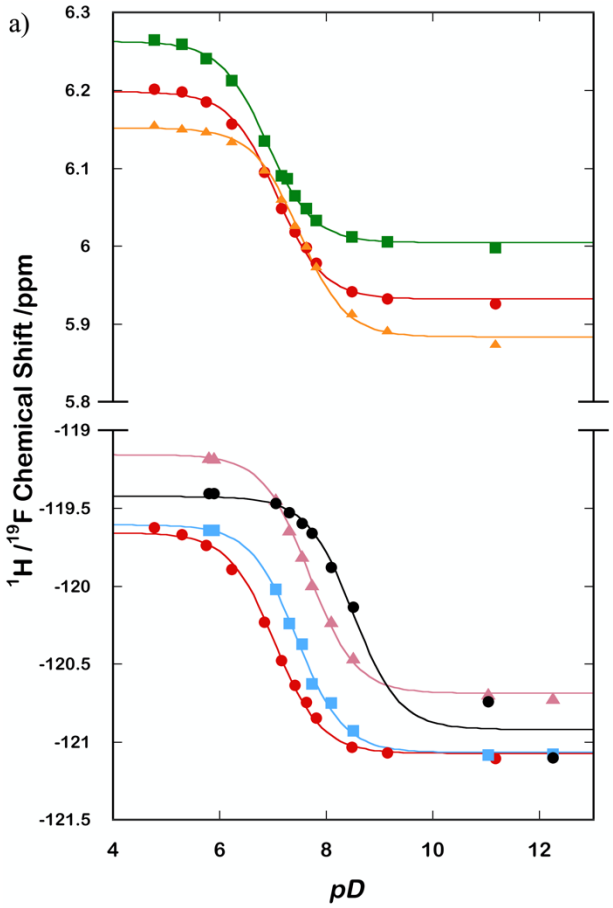
Scheme 4.4: Equilibrium of complex [9] with a deuterated aquo ligand and deuteroxide ligand.

These species are in rapid exchange on an NMR timescale, thus yielding a single peak which is an average of the exchanging species, [Ru-OD_2] and [Ru-OD]. The chemical shift of the fluorine atoms (red circles in **Scheme 4.4**) for the aquo/hydroxide species are pH dependant, **Figure 4.6 (b),** therefore a pK_a for the deprotonation of the bound aquo ligand can be measured, by plotting δ against the pD, **Figure 4.6 (a)** and fitting the graph to **Equation 4.1.**

$$\delta = \delta_{max} + \left(\frac{\delta_{max} - \delta_{min}}{1 + 10^{(pD - pK_a)}}\right) \quad (4.1)$$

The same pK_a can also be measured, using ^1H NMR, for the complexes where the arene ligand is benzene, as it is possible to identify the peaks which correspond to the six arene protons (green circles in **Scheme 4.4**) for the aquo/hydroxide species.

The effect on the pK_a of changing the arene ligand in the 5,5'-FBipy complexes [9]-[12], and changing the bipyridine ligand in complexes [1], [9] and [13] has been measured, **Figure 4.6 (b)**. The pK_a values presented are measured from the pD values without a correction factor applied.



b)

Complex	pKa from ^{19}F data
[9]	7.03
[10]	7.45
[11]	7.68
[12]	8.50

Complex	pKa from ^1H data
[1]	7.50
[9]	7.07
[13]	6.88

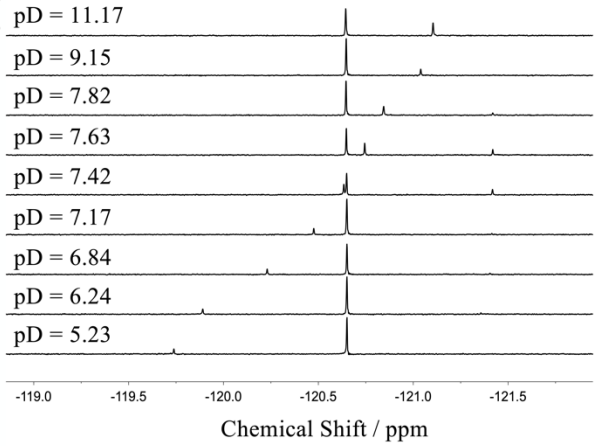


Figure 4.6: (a): pH dependence of the chemical shifts of the aquo coordinated ruthenium species and determination of the pK_a in the following complexes [9] - Red circles; [10] - Blue squares; [11] - Pink triangles; [12] - black circles; [13] - Green squares; [1] - Orange triangles. ^{19}F chemical shifts of the 5,5'-difluorobipyridyl fluorine's or ^1H chemical shift of the coordinated benzene ligand were plotted against pD. **(b):** The pK_a values of a series of ruthenium complexes measured using $^{19}F\{^1H\}$ NMR and ^1H NMR. **(c):** A series of $^{19}F\{^1H\}$ decoupled spectra of complex [1] in D_2O phosphate buffer at differing pD values.

The observed differences in pK_a are entirely consistent with the changes in electron density at the metal due to the subtleties of the electron withdrawing effects of fluorines on the bipyridine

ligand and the electron donating effects of the different arenes.[134] In other words, the trend observed going from complex [9] – [12] shows that the increasing electron donating capability of the η^6-arene ligand to the ruthenium centre leads to a higher measured pK_a value; an increased electron density on the metal lowers its Lewis acidity. The electron withdrawing capabilities of the trifluoromethyl group on complex [13] significantly increases the Lewis acidity, reflected in the lowest pK_a (6.88) measured for any of these complexes. For complex [9], the pK_a has been measured using both ^1H and ^{19}F NMR spectroscopy, and the values are consistent, 7.03 and 7.07 respectively. These data have also identified that these species have pK_a values close to the working pD of the reactions.

Comparing the analytical techniques for monitoring the speciation of Ru(II) arene complexes, NMR can be used to measure both kinetic parameters and dissociation constants, but this is not possible for traditional mass spectrometry, X-ray crystallography and UV spectroscopy methods. Furthermore, comparing ^1H and ^{19}F NMR spectroscopy, when many species are present in solution the ^1H NMR spectra become very complex, but when fluorine atoms are present, the spectra are dramatically simplified. For the fluorinated complexes, if the aquo species is present at any time in solution, then a direct read out of the apparent pH can be measured, which is particularly useful.

4.3.3. Amino Acid Coordination

Reaction of these fluorinated ruthenium complexes with all naturally occurring amino acids gave a series of characteristic ^{19}F NMR signals for each Ru-amino acid adduct. Throughout this study, four amino acids were selected which had the most reactive side chain ligands. All four amino acids have been protected at the amino-group, and the carboxylic acid group has been esterified, **Figure 4.7**. This is to ensure that there is a focus on side chain – metal coordination, to mimic reactivity with a polypeptide, where the amino and carboxylate groups form peptide bonds.

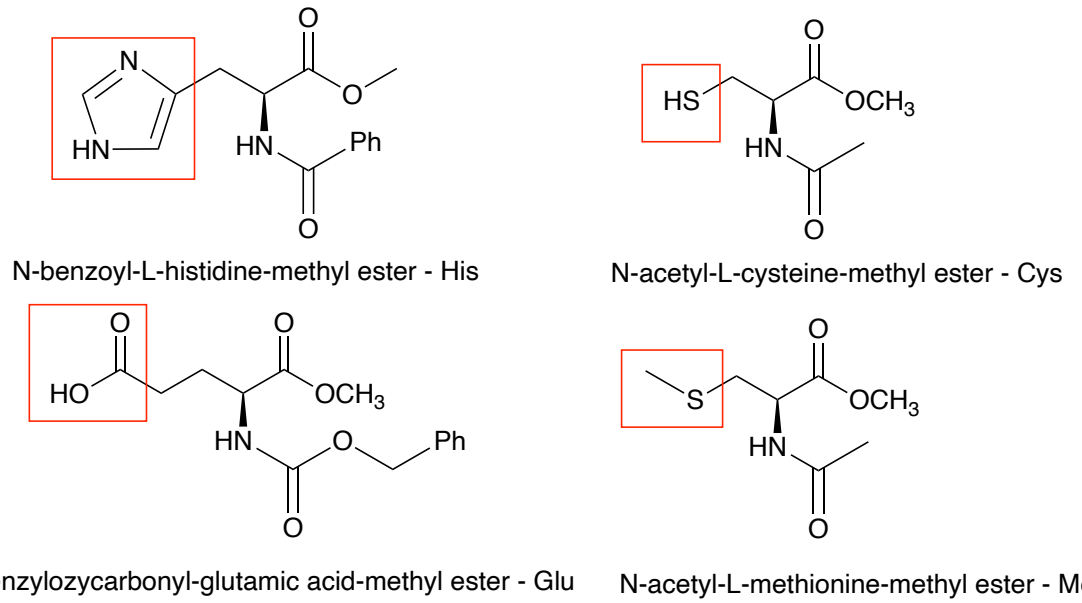


Figure 4.7: Structures of four protected amino acids, with their nucleophilic side chains highlighted.

Identification of Amino Acid Peaks

As shown in **Section 4.3.2**, the chemical shifts for all chloro, aquo and phosphate bound species were identified for all ruthenium complexes , therefore, through incubating the complexes with amino acids and monitoring the ^{19}F NMR spectra over time it was possible to identify any new peaks corresponding to amino acid coordination. The presence of any Ru-amino acid adducts were confirmed by ESI-MS, **Appendix B, Table B.2.**

Effect of Position of Fluorination – 3,3' vs 5,5'

The fluorine atoms on the bipyridyl ring are enantiotopic when an achiral ligand is coordinated to the metal e.g. Cl or H_2O, and once a chiral amino acid ligand is coordinated they become diastereotopic and hence magnetically inequivalent.[135] Therefore, despite the symmetry of the starting molecule, the amino acid coordinated adducts of these complexes do not always give rise to single resonances in the proton decoupled ^{19}F NMR spectra. This is clearly observed when exploring the coordination of Cys to the 5,5'-difluorinated complexes [9] – [12]. ^{19}F NMR spectra show that, a Cys coordinated species, gives rise to two peaks, which are very clearly resolved in complexes [10], [11] and [12], **Figure 4.8, (Left).**

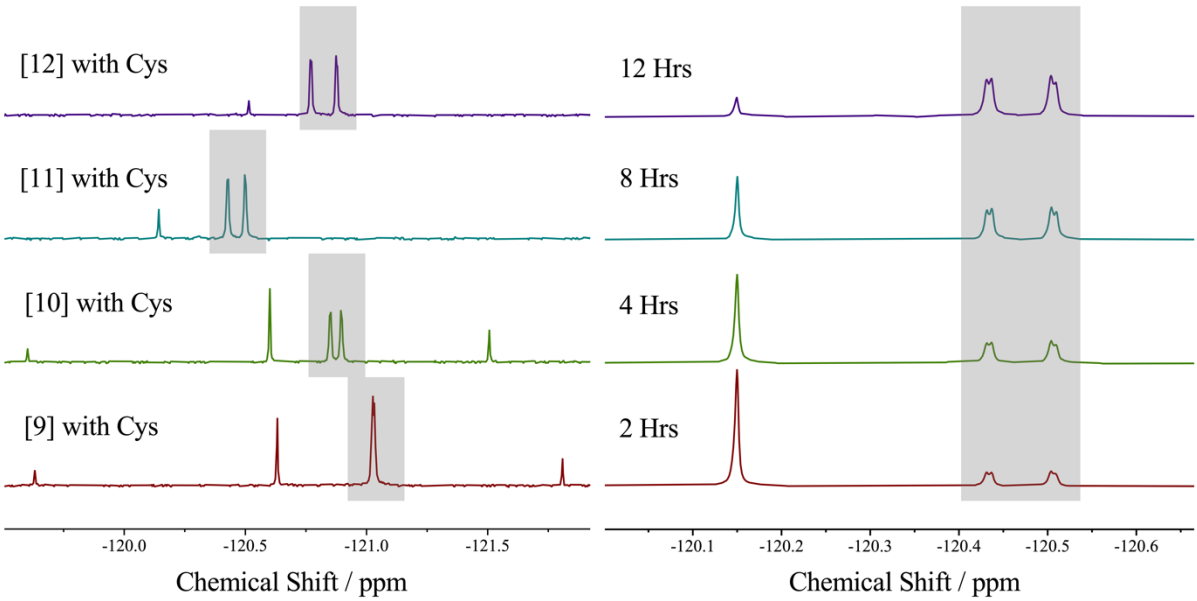


Figure 4.8: (Left): $^{19}F\{^{1}H\}$ NMR spectra of complexes [9], [10], [11] and [12] incubated with N-acetyl cysteine methyl ester (2 mM Ru, 3 eq. amino acid, 12 hr, 310 K). The signals highlighted in the grey box are assigned to Ru-[Amino Acid] adducts **(Right):** Time course $^{19}F\{^{1}H\}$ NMR spectra of complex [11] with N-acetyl cysteine methyl ester (2 mM Ru, 3 eq. amino acid, 310 K).

Looking in more detail at the coordination of Cys to complex [11] over time, **Figure 4.8 (Right),** the peaks for each fluorine are in fact doublets. That these signals originate from the same species is clearly evidenced by the $^{8}J_{F\text{-}F}$ coupling of 1.49 Hz observed in the high resolution ^{19}F NMR spectra and a clear cross-peak in a proton decoupled ^{19}F-^{19}F COSY experiment, **Figure 4.9.**

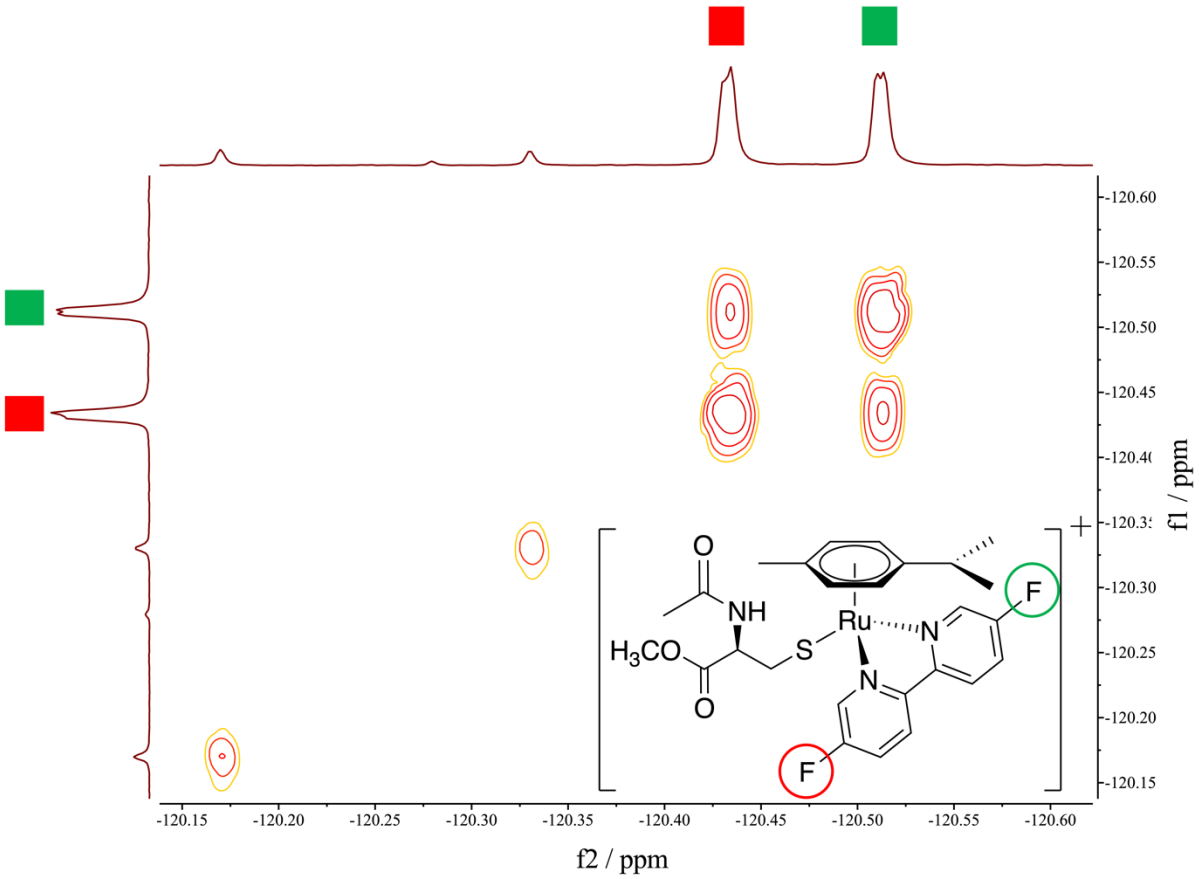


Figure 4.9: (Insert): Structure of [11]-Cys with the magnetically inequivalent fluorine atoms highlighted. **(Spectrum)** $^{19}F\{^{1}H\}$ - $^{19}F\{^{1}H\}$ COSY spectrum from the incubation of complex [11] with N-acetyl cysteine methyl ester (2 mM Ru, 3 eq. amino acid, 12 hr, 310 K).

The origin of these Ru-Cys peaks have been explored in great detail. Variable temperature analyses are consistent with there being free rotation around the Ru–S bond at room temperature, **Appendix C, Figure C.1**. Complex [9] has also been incubated with the achiral substrates glycine, cysteamine, 3-mercaptopropianoate, which all give rise to a singlet Ru-adduct peak, indicating that the chiral environment is crucial to peak complexity, **Appendix C, Figure C.2**.

In comparison, when the 3,3'-difluorinated derivative, complex [7], is incubated with Cys, only one singlet peak is attributed to a Ru-Cys adduct, [7]-Cys, **Figure 4.10**. Although, the fluorine atoms are diastereotopic, this does not manifest itself in the spectra, as the fluorine atoms resonate at the same frequency and they do not couple. A combination of the proximity of the fluorine atoms in space and F-F electronic repulsion forces the fluorine atoms as far apart as possible, slightly twisting the bipyridine ligand, which is observed in the crystal structures of 3,3'-FBipy complexes [5] – [8]. The metal-ligand coordination sphere is flexible enough in solution for there to be fluxionality between the two non-planar pyridyl rings such that the two

fluorine environments appear equivalent and we observe only one ^{19}F NMR signal, the achiral average.

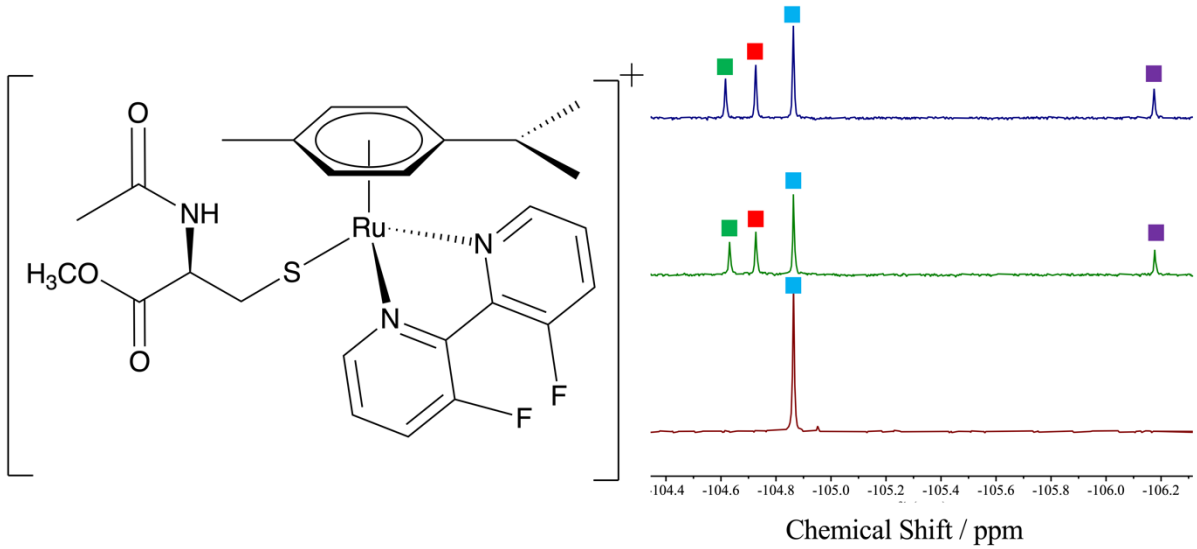


Figure 4.10: (Left): Structure of complex [7]-Cys. **(Right):** Time course $^{19}F\{^{1}H\}$ NMR spectra of complex [7] with N-acetyl cysteine methyl ester (2 mM Ru, 1 eq. amino acid, 310 K, 0 hr = bottom spectrum, 4 hr = middle spectrum and 8 hr = top spectrum). Green squares [7]-D_2O, red squares [7]-Cys, blue squares [7]-Cl, purple squares [7]-Phos.

The peaks observed when the 5,5'- and 3,3'- FBipy derivatives are incubated with the protected amino acids, Cys, Glu, His and Met are presented in **Figure 4.11.** For the 5,5'-FBipy, when a chiral amino acid coordinates to the metal centre, two peaks are observed, one for each fluorine atom. These fluorine atoms are capable of long-range coupling, therefore these peaks can appear as doublets, e.g. for Cys and His. There is extra complexity for histidine coordination as there are two sets of doublets, with a different integration ratio, which is interpreted as due to coordination at either the δ-N or the ε-N on the imidazole ring.

The subtleties in peak width, coupling constants and peak separation, particularly with 5,5' fluorinated complexes, give each different Ru-amino acid a unique fingerprint which can help in identifying coordination partners for Ru(II) arene complexes. Furthermore, these unique fingerprints could be invaluable for identifying coordination environments in the context of more complex peptide and protein environments.

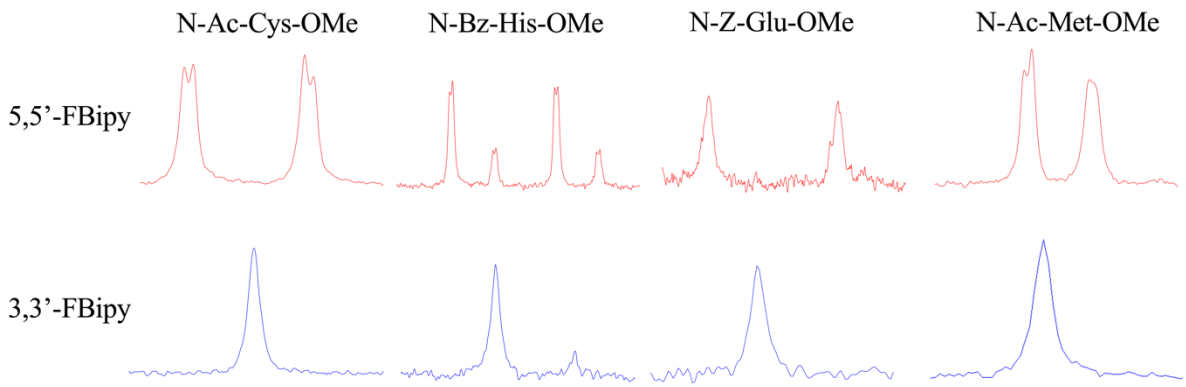


Figure 4.11: Comparison of ^{19}F NMR peaks observed for the amino acid adducts of complex [7] and complex [11].

Effect of Fluorinated Substituent Group F vs CF_3

The purpose of changing the substituent group from a fluorine to a trifluoromethyl group was to increase the number of equivalent fluorine atoms in the ruthenium complex. By increasing the number of fluorine atoms from two to six the ^{19}F NMR signal intensity should increase, enabling experiments to be performed at lower complex concentrations. Complexes [13] – [16] were synthesised and incubated with different amino acids, however, the overlap in the ^{19}F NMR spectra from peaks assigned to Ru-OD_2 and Ru-Cl made it more difficult to determine those peaks attributed to ruthenium amino acid coordination. The extra bond between the fluorine atoms and the metal, result in the fluorine being less sensitive to changes in metal coordination environment (**Appendix C, Figure C.3**).

4.3.4. Competition Experiments

With an understanding of the nature and ^{19}F NMR chemical shifts of the different Ru-amino acid adduct peaks, competition experiments were performed to identify and quantify preferred amino acid coordination sites. To do this, all four 5,5'-FBipy complexes, [9]-[12] were incubated with a slight excess of each amino acid individually and collectively and temporally monitored by ^{19}F NMR spectroscopy, **Figures 4.12** to **4.15.**

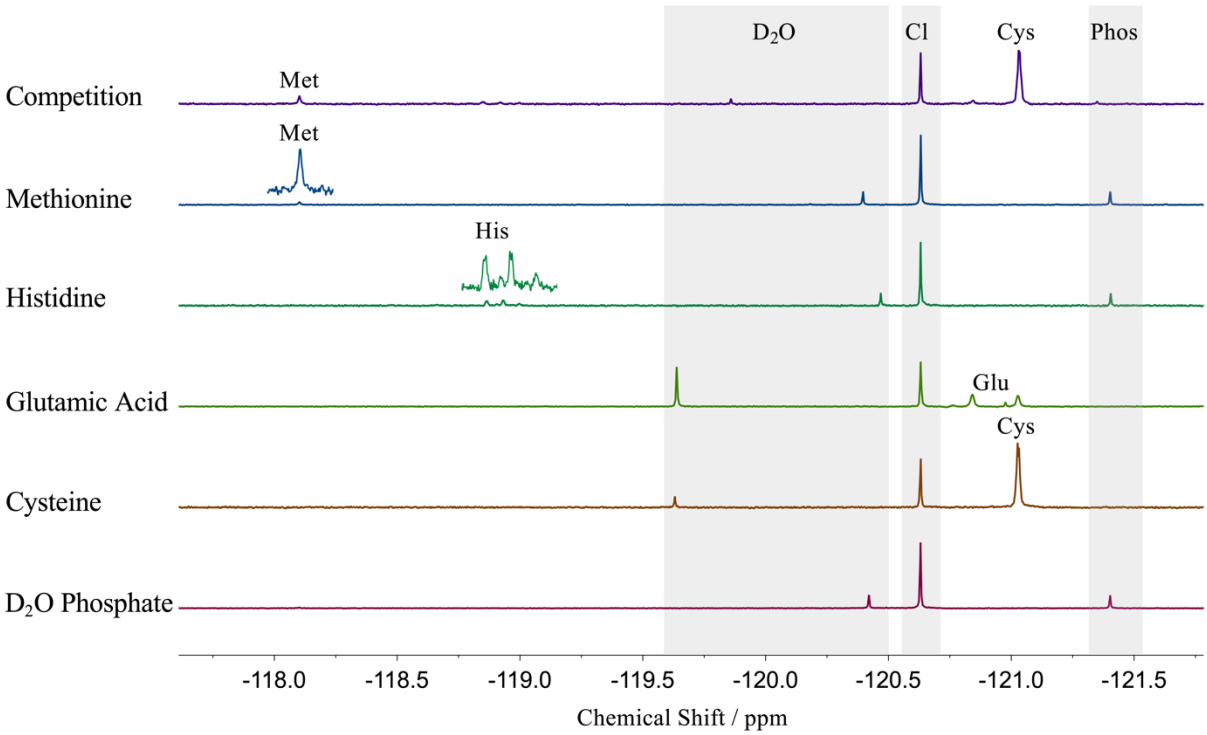


Figure 4.12: A series of $^{19}F\{^{1}H\}$ NMR spectra of complex [9] incubated with the protected amino acids, N-Ac-Cys-OMe, N-Z-Glu-OMe, N-Bz-His-OMe, N-Ac-Met-OMe, and a mixture of all amino acids together, (2 mM Ru, 3 eq. each amino acid, 24 hr, 310 K).

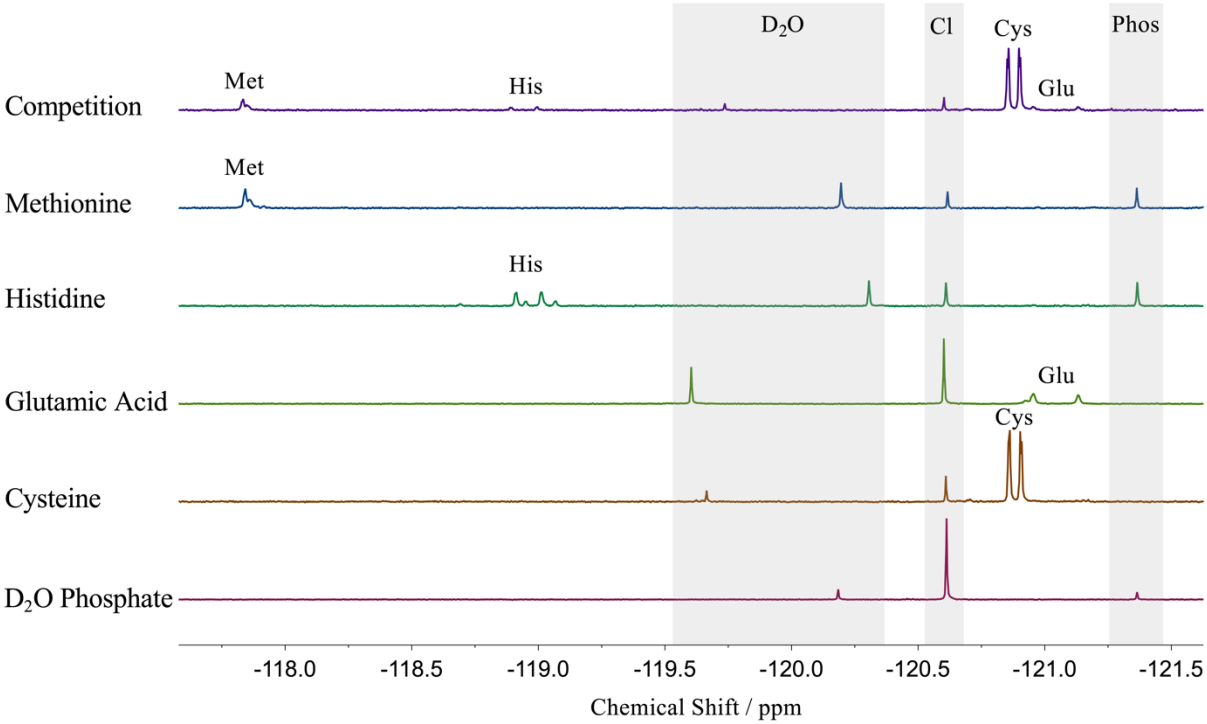


Figure 4.13: A series of $^{19}F\{^{1}H\}$ NMR spectra of complex [10] incubated with the protected amino acids, N-Ac-Cys-OMe, N-Z-Glu-OMe, N-Bz-His-OMe, N-Ac-Met-OMe, and a mixture of all amino acids together, (2 mM Ru, 3 eq. each amino acid, 24 hr, 310 K).

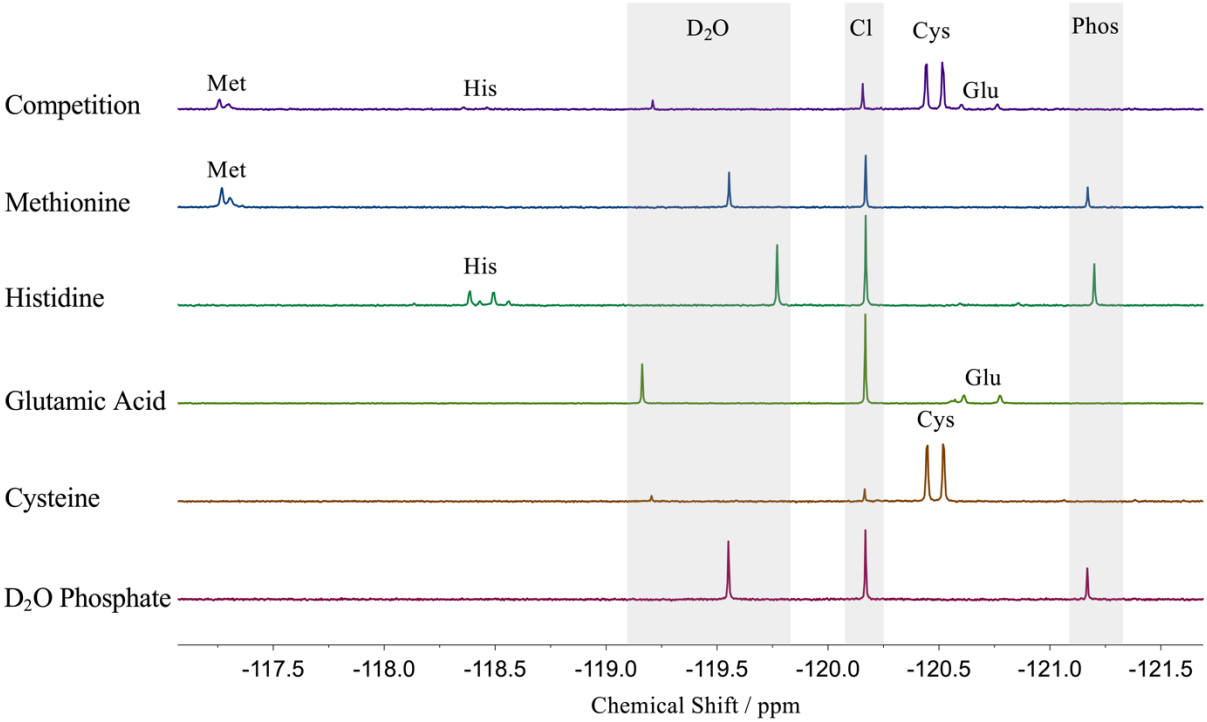


Figure 4.14: A series of $^{19}F\{^{1}H\}$ NMR spectra of complex [11] incubated with the protected amino acids, N-Ac-Cys-OMe, N-Z-Glu-OMe, N-Bz-His-OMe, N-Ac-Met-OMe, and a mixture of all amino acids together, (2 mM Ru, 3 eq. each amino acid, 24 hr, 310 K).

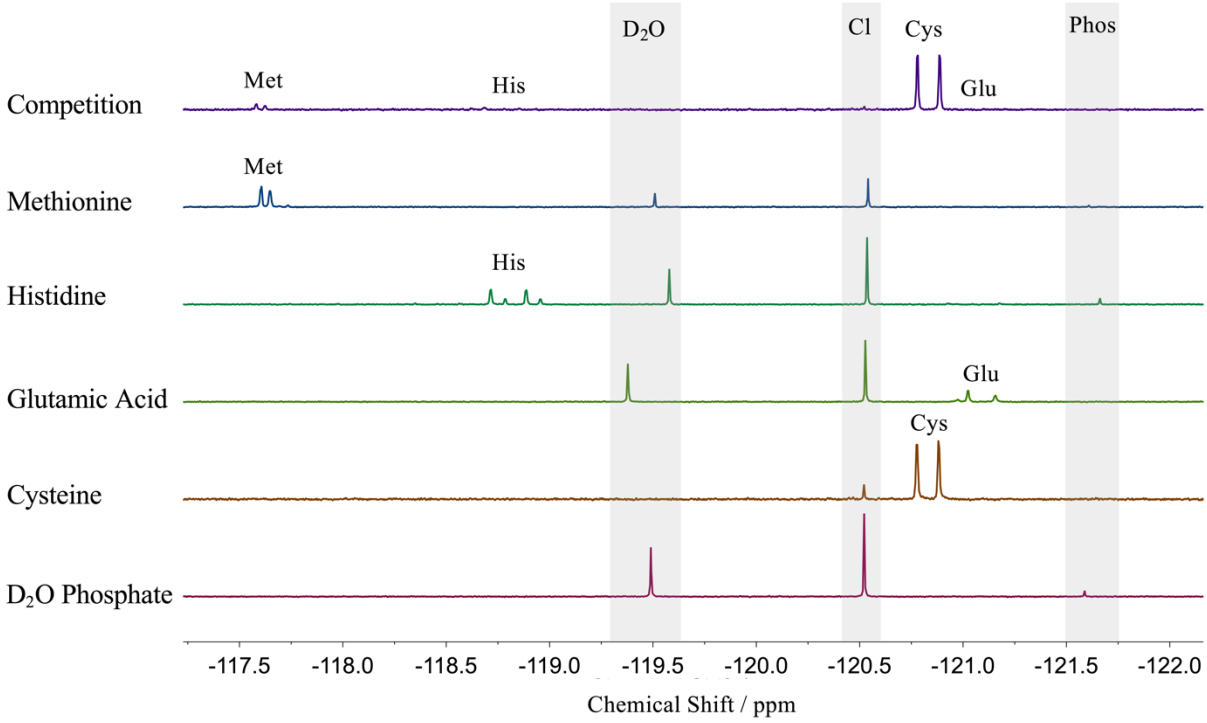


Figure 4.15: A series of $^{19}F\{^{1}H\}$ NMR spectra of complex [12] incubated with the protected amino acids, N-Ac-Cys-OMe, N-Z-Glu-OMe, N-Bz-His-OMe, N-Ac-Met-OMe, and a mixture of all amino acids together, (2 mM Ru, 3 eq. each amino acid, 24 hr, 310 K).

Some key findings from these experiments are summarised as follows.

In all competition experiments, preference for cysteine coordination is absolute. Although this is not chemically surprising, (Ru(II) is a soft metal centre and the sulphur thiol is a soft ligand) solution based ^{19}F NMR spectroscopy has been used to confirm this for these complexes. What is apparent, is that cysteine coordination, almost completely out-competes coordination to other soft nucleophiles (e.g. the histidine imidazole and the methionine thioether), therefore making it an attractive residue for localising a ruthenium cofactor in a protein.

Increasing the electron donating capability of the arene ligand from benzene → hexamethylbenzene, drives the equilibrium towards the Cys bound adduct and at a faster rate. To quantify this, rates of formation of the Cys bound adducts with different arene ligands were measured, through integration of the relevant peaks in the ^{19}F NMR spectra. The formation of the [Ru]-Cys adduct peak was adequately described by a 1st order process, with the *p*-cymene and hexamethylbenzene complexes [11] and [12] being faster than the benzene and tolyl complexes [9] and [10], **Table 4.2.**

Table 4.2: A comparison of the observed rates for the complexes [9] – [12] for the formation of a ruthenium cysteine adduct (2mM Ru, 3 eq. amino acid, 310 K)**.**

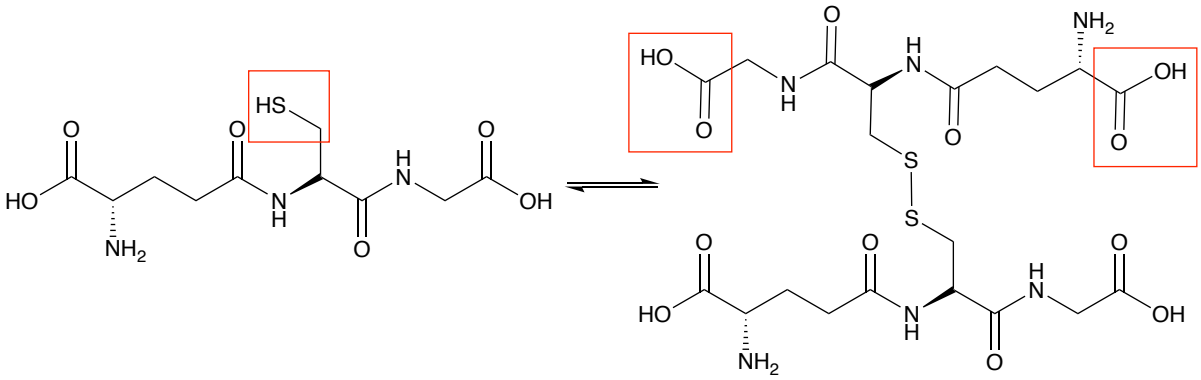
Complex	k_{obs} / x 10^{-4} s^{-1}
[9]	2.17 ± 0.03
[10]	2.08 ± 0.05
[11]	2.76 ± 0.04
[12]	3.71 ± 0.08

As previously discussed, there is an initial equilibrium in aqueous solution between the [Ru]-Cl, [Ru]-OH_2 and [Ru]-OH species. The rate of formation of a [Ru]-Cys adduct is dependent on the formation of the [Ru]-OH_2 species, a theory that has been well studied, in contrast to the mechanism observed for cisplatin and Ru(II) anticancer complexes.[136,137] The increased electron density of the ruthenium in the *p*-cymene and hexamethylbenzene, complexes [11] and [12], drives the equilibrium more towards a [Ru]-OH_2 complex, contributing to an increased rate of formation of the [Ru]-Cys adduct.

This theory is also supported by the measured pK_a data in **Section 4.3.2.** The aquo complexes of [11] and [12] have much higher pK_a values therefore at an initial pD of 7.2 the equilibrium will lie heavily towards the [Ru]-OH_2 complex.

4.3.5. Binding to Glutathione

In a cellular environment the most abundant biological thiol is the tripeptide glutathione which is at millimolar concentration in the cytoplasm of many different cells. Glutathione, equilibrates between reduced monomer (GSH, containing a thiol) and oxidised dimer (GSSG, containing a S-S bridge), **Scheme 4.5**. One role of glutathione in cells is to maintain the intracellular redox potential, and also stop the accumulation of toxic metabolites. Glutathione also plays an important role in iron-sulphur cluster assembly.[138]



Scheme 4.5: The equilibrium between the reduced monomeric glutathione, GSH, and the oxidised dimeric form, GSSG.

Complexes [9] – [12], were incubated with different ratios of reduced (GSH) and oxidised (GSSG) glutathione, and the ^{19}F NMR spectra were recorded, **Figure 4.16.** The ^{19}F NMR spectra show quantifiably that binding to the carboxylate groups is preferred in the absence of free thiolate, but that binding to the cysteine thiolate group is markedly preferred in thermodynamic terms when available.

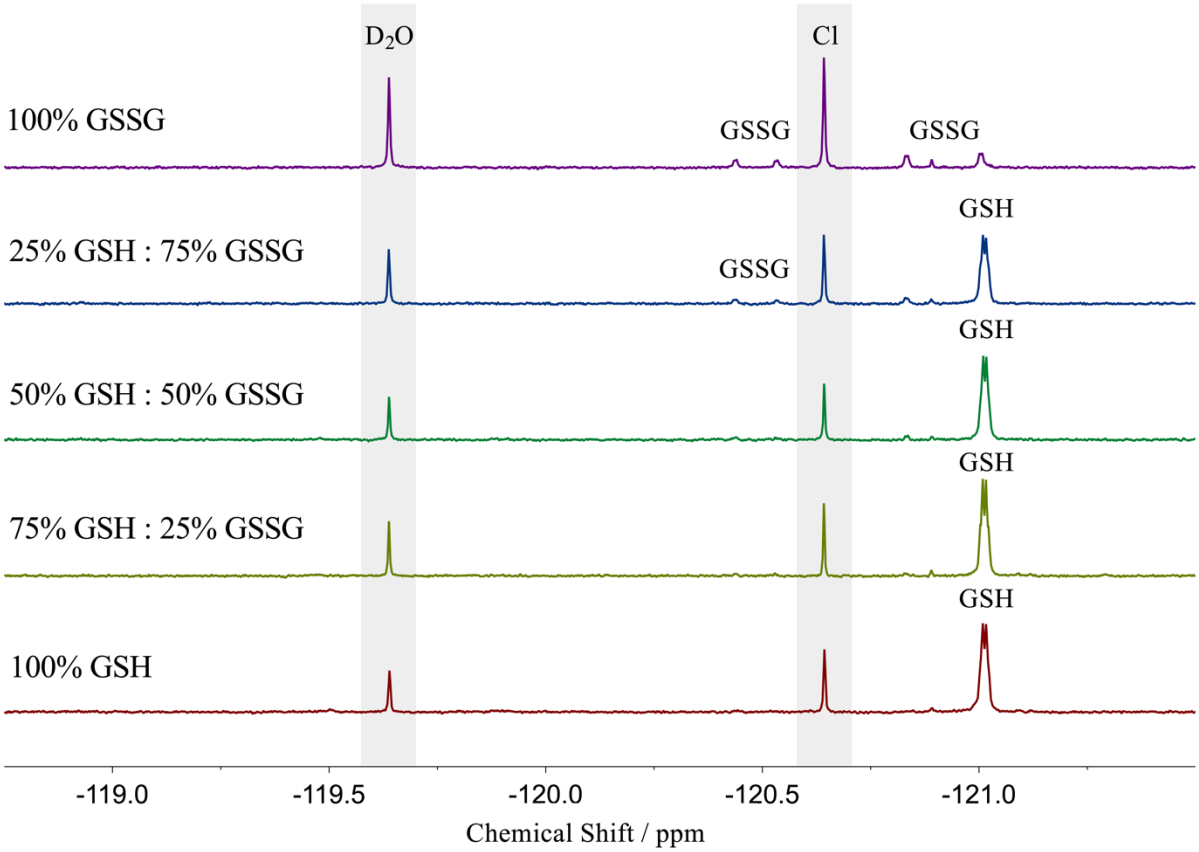
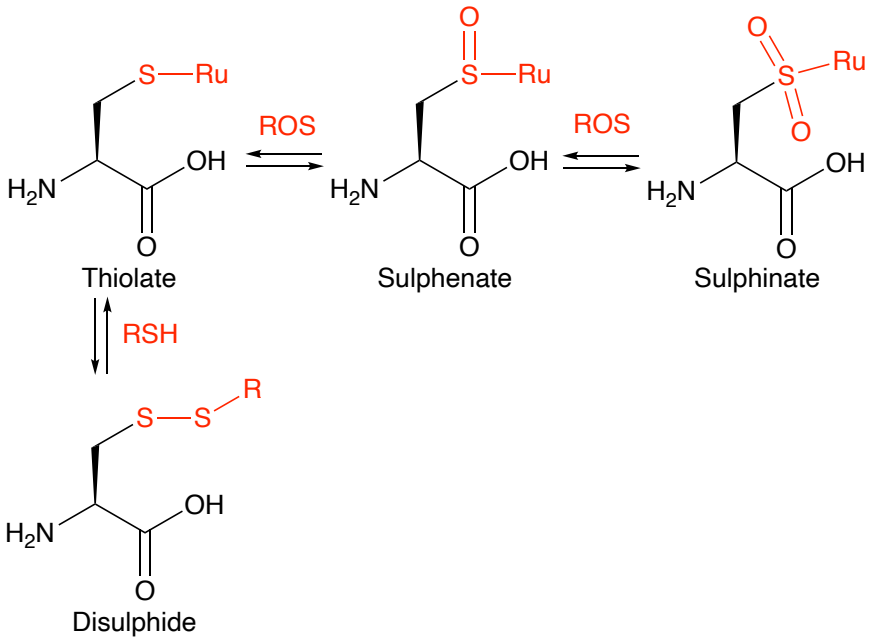


Figure 4.16: A series of $^{19}F\{^{1}H\}$ NMR spectra of complex [12] incubated with defined mixtures of reduced and oxidised glutathione (2 mM Ru, 3 eq. glutathione, 24 hr, 310 K).

4.3.6. An Aside – Exploring Cysteine Oxidation

Alongside incubating Ru(II) complexes with small molecules, Ru(II) complexes were also incubated with different proteins, which will be discussed in detail in the following chapters. In some reactions where Ru(II) arene complexes were coordinating to proteins the measured and expected masses differed by 16 mass units, which was hypothesised to be due to the cysteine thiol being oxidised after ruthenium coordination. The literature precedence for this reactivity was reported by Sadler et al. who suggested that the free thiol of Cys34 in human serum albumin was being oxidised to a sulfenate by molecular oxygen after ruthenium coordination, as the sulphur atom is more electrophilic upon metal coordination.[139] Sadler et al. also reported oxidation of ruthenium-thiolate adducts to ruthenium-sulfenate adducts on a small molecule scale by LC-ESI-MS.[21]

Oxidation of ruthenium-cysteine adducts was explored using both ESI-MS and ^{19}F NMR spectroscopy. Many different oxidation states of cysteine are capable of coordinating to a ruthenium centre, summarised in **Scheme 4.6.**



Scheme 4.6: Different biological cysteine oxidation states, with potential coordination of ruthenium to the sulphur atom.

Via ESI-MS, coordination of ruthenium to cysteine was monitored upon addition of the reactive oxygen species (ROS), H_2O_2, **Figure 4.17.** After addition of 1 Eq. of H_2O_2 to an incubation of complex [12] with 1 Eq. of Cys, an oxidation product at a mass of 649.27 was observed. This corresponds to [Ru]-Cys + [O], where the thiolate ligand has been oxidised to a sulfenate ligand.

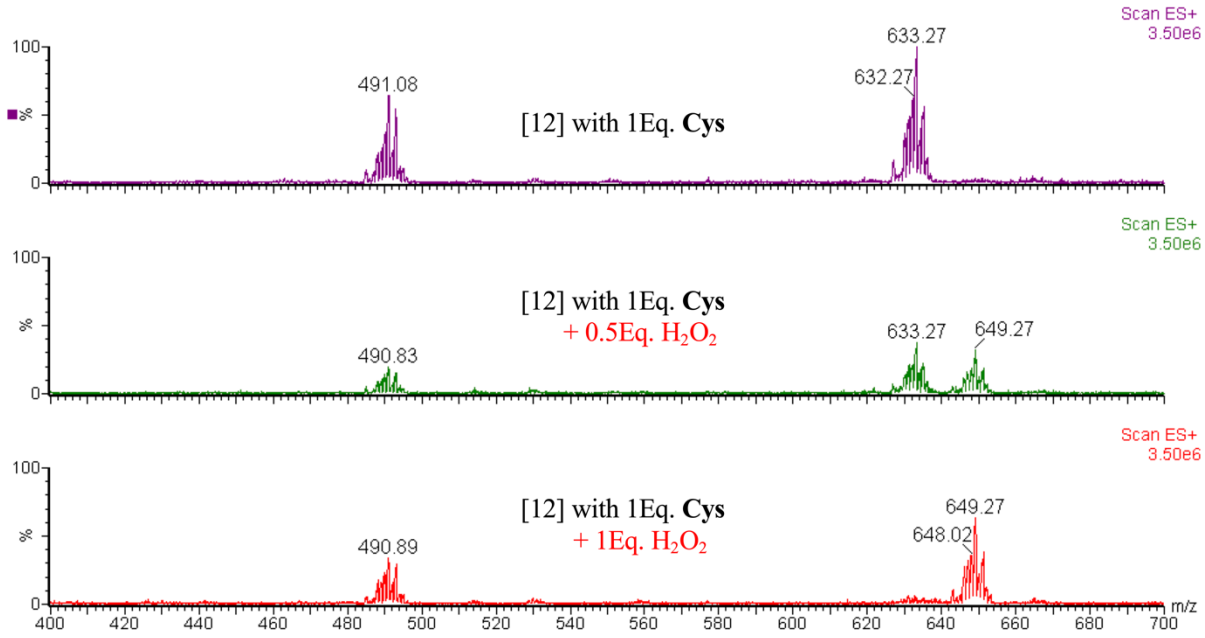


Figure 4.17: ESI-MS spectra recorded of the solution mixture when complex [12]-Cys is incubated with initially 1 Eq. of N-acetyl cysteine methyl ester, followed by different equivalents of the oxidant H_2O_2. Experimental and calculated values given in **Appendix B, Table B.3.**

Oxidation of the thiolate ligand with 0.5 Eq. of H_2O_2 leads to dramatic changes to the ^{19}F NMR spectrum of an incubation of N-acetyl cysteine methyl ester with complex [12]. The two doublets observed upon formation of a [Ru]-Cys adduct, are split even further to give two additional sets of doublets of unequal intensity, analogous to a histidine bound species, **Figure 4.18**. On addition of 1 Eq. of H_2O_2 the doublets corresponding to ruthenium thiolate coordination are no longer observed and there is an increase in peak intensity of the Ru-sulphenate species.

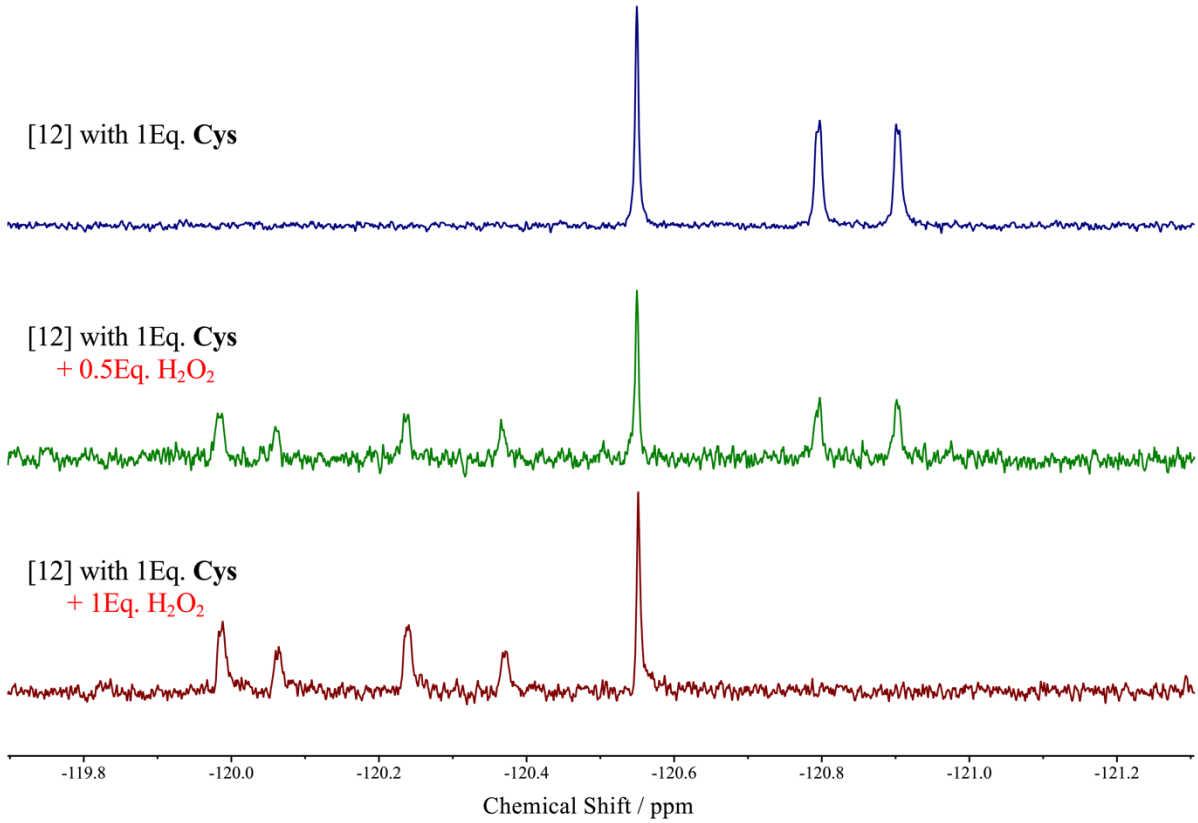


Figure 4.18: $^{19}F\{^{1}H\}$ NMR spectra when the incubation of complex [12] with 1Eq. of N-acetyl cysteine methyl ester is spike with different number of equivalents of the oxidant H_2O_2.

The ^{19}F NMR peaks corresponding to a Ru-sulphenate species are downfield shifted compared to the Ru-thiolate species. In this ^{19}F NMR spectroscopy method, negatively charge ligands (e.g. chloride, phosphate, thiolate) are upfield of neutral ligands (methionine and histidine). Upon thiolate oxidation the negative charge on the sulphur will be delocalised across the S=O bond weakening the ionic contribution to bonding. The current hypothesis for the Ru-sulphenate species giving rise to two sets of doublets is that oxidation of the sulphur introduces a chiral centre; ruthenium coordination to each enantiomer therefore gives rise to a set of doublets of equal intensity, **Figure 4.19,** from the resulting diastereomers**.**

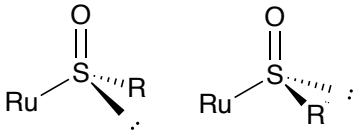


Figure 4.19: Different enantiomeric forms of the sulfenate ligand bound to ruthenium.

4.4. Conclusions

This chapter has outlined a spectroscopic approach, using ^{19}F NMR spectroscopy, to monitor the speciation of a suite of Ru(II) arene bipyridine complexes with amino acid side chain residues.

The ^{19}F NMR spectra provide quantitative, solution-based information on the reactions of these complexes with amino acid Lewis bases which could not be gathered using traditional ^{1}H NMR, ESI-MS and crystallographic methods. The introduction of a fluorinated substituent on a ruthenium ligand, enables the reactivity of specific metal species in a complex mixture of biological nucleophiles to be explored. The quantitative nature of this ^{19}F NMR spectroscopic method presented enables a deeper understanding of the preferred biological nucleophiles for Ru(II) arene complexes. Furthermore, rate constants and pK_a values can be measured from the integration and chemical shifts of carefully assigned peaks.

In the Ru(II) arene complexes presented, there are two key ligands, the η^6-arene and the fluorinated bipyridine ligand. The electronic properties of both these ligands have been varied and the effect this has on the speciation to biological nucleophiles has been monitored. The position of fluorination around the bipyridine ligand has also been varied and the ^{19}F NMR spectra monitored upon amino acid coordination to the ruthenium centre. When the fluorine atoms are in the 5,5'-position on the bipyridine ring, they are remarkably sensitive to changes in coordination at the ruthenium centre. This results in a unique fingerprint that leads to unequivocal assignment of the ^{19}F NMR spectra of Ru-amino acid adducts.

For all Ru(II) arene bipyridine complexes synthesised, the cysteine thiol was the preferred coordination partner in solution for Ru(II)(η^6-arene)(bipyridine) complexes, confirmed by ^{19}F NMR spectroscopy. Furthermore, from incubations with the tripeptide glutathione, cysteine was confirmed as the preferred coordination site for the Ru(II) complexes explored.

Understanding the speciation of Ru(II) arene complexes when exposed to biological nucleophiles was a key goal for this research. Introducing fluorinated ligands as reporters has proven to be a quantitative and information rich method of understanding how Ru(II) arene complexes coordinate to amino acids. In terms of sensitivity, it has been possible to detect metal-amino acid complexes at relatively low adduct concentrations, however this often requires longer collection times.

With a long-term goal of developing an artificial ruthenium metalloenzyme with direct protein coordination, being able to confirm and quantify in solution that Ru-cysteine adduct formation

has occurred is very useful. It was hypothesised that a cysteine residue could act as an initial anchoring residue before subsequent ligand exchange to generate a protein-metal hybrid with the potential of catalytic activity. This ^{19}F NMR spectroscopic method has been used in future chapters in characterizing Ru-protein adducts, where multiple ligands are available on the same protein molecule.

Finally, in the development of Ru(II) arene anticancer complexes, understanding the fundamental solution behaviour will contribute to deciphering the mechanism of action which brings about apoptotic response in cells. As researchers attempt to identify new protein targets for Ru(II) arene complexation, this study provides an alternative insight into how Ru-target interactions could be characterised.

Chapter 5

5. Coordination of Ruthenium Fragments to Proteins

5.1. Introduction

The results presented in Chapter 4 increased our understanding of how Ru(II) arene complexes bind to simple amino acids and short peptides. This enabled a strategy for the controlled coordination of different Ru(II) arene complexes to protein scaffolds to be developed. The purpose was two-fold: (i) to understand where and to what extent different Ru(II) arene complexes coordinate to a variety of proteins and (ii) to assess the usefulness of ^{19}F NMR spectroscopy to characterise Ru-protein adducts. As discussed in Chapter 4, cysteine was identified as the preferred amino acid binding site for a suite of Ru(II) arene bipyridine complexes. Therefore, the aim was to confirm, through a range of analytical techniques, whether this preference was translated to free cysteine residues in the context of folded proteins.

5.1.1. The Chemistry of Cysteine

Cysteine (Cys) is one of the least abundant of the twenty common amino acids in proteins, however, due to the nucleophilic sulfhydryl side chain and redox potential, Cys plays vital structural, catalytic and regulatory roles in functionally diverse proteins.[140]

A further consequence of the high nucleophilicity of the cysteine thiol side chain is that there are few proteins with free surface cysteines. Buried cysteines are either often in unique catalytic environments or linked through a disulphide bond to another cysteine residue. Therefore, in order to coordinate to a ruthenium complex, the Cys in a folded protein needs to be accessible and in the reduced, thiol form.

When coordinating a ruthenium complex to a protein it is desirable for the protein-metal conjugation to be selective for a specific amino acid. The unique properties of Cys, discussed above, make it a good target for selective ruthenium protein modification.[141]

5.1.2. Ruthenium Coordination to Cysteine Containing Proteins

This chapter deals with two categories of Cys containing proteins: (i) naturally occurring proteins with free Cys sites and (ii) proteins with engineered Cys sites. Site-directed mutagenesis is a method enabling specific modifications of a DNA sequence and, in terms of protein engineering, these modifications manifest themselves as mutation of the protein's primary structure. Through mutagenesis, Cys residues can be incorporated into a wide variety of proteins, however, it is noteworthy that any mutation can affect the structure, stability or function of the protein.

5.1.3. Engineered Cysteine Mutants

Ubiquitin

Ubiquitin (Ubq) is a small (8.6 kDa, 76 residue) protein expressed in all eukaryotic cells. Ubq is rich in lysine residues and performs a number of functions through covalent attachment to other proteins, thus altering the stability, localisation and activity of the target protein.[142] Poly-ubiquitination is a process where a number of Ubq units are covalently linked, and these Ubq chains target unwanted or damaged proteins to the proteasome for degradation. In this study, a lysine residue at position 63 has been mutated to a cysteine. Lysine residues are hydrophilic, and often found on the surface of proteins to enhance solubility, therefore the cysteine residue is solvent exposed.

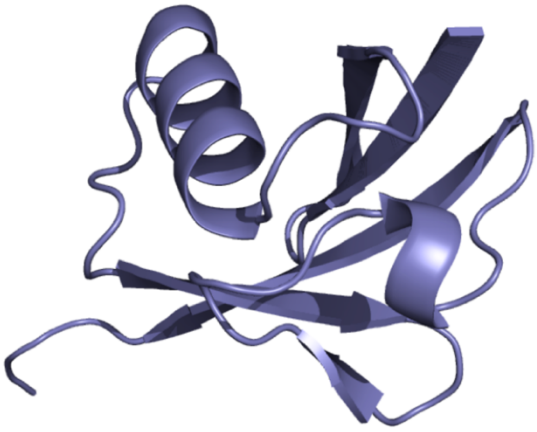


Figure 5.1: Structure of Ubq (Protein Data Bank: 1UBQ) determined by X-Ray Crystallography.[143]

Cytochrome b_{562}

Cytochrome b_{562} (Cyt b_{562}) is a periplasmic haem-binding protein (11.8 kDa, 106 residues) from *Escherichia coli* and is the only member of the family of 4-α-helical bundle cytochromes to not have a covalently bound haem cofactor. **Figure 5.2** shows a structure, determined in

solution by NMR spectroscopy, of apocytochrome b_{562} where the four helices are apparent, although there are significant distortions. In the absence of haem, many of the residues that comprise the nascent haem binding pocket are accessible to solvent and small molecules.[144] The Barker group has engineered many cysteine mutants of Cyt b_{562}, where the cysteine residues are directed into the region where haem coordinates.[145] The purpose of this study was to convert this unique cytochrome to a c-type cytochrome with covalent linkage (Cys-haem) through a thioether bond.[146]

Although the availability of the NMR structure of the apocytochrome *b* is invaluable, the highly dynamic nature of this protein when haem is not coordinated means one must be careful when predicting the position and orientation of the amino acid side chains. With cysteines introduced at different sites (L10C, R98C) the differential activity of these residues could be explored. There are other Lewis basic positions located in the hydrophobic region of the wild-type (WT) protein, including His63 and His102, which have the potential to coordinate to ruthenium.

Figure 5.2: Structure of apo-Cyt b_{562} (Protein Data Bank: 1APC) determined by NMR spectroscopy.

C_2Am

The C_2A domain of the synaptic vesicle protein synaptotagmin I (C_2Am: 16.2 kDa, 145 residues) plays a fundamental role in Ca^{2+} sensing and regulating exocytosis. The presence of Ca^{2+} triggers a cellular response to fuse the synaptic vesicle, which is full of neurotransmitters, to the presynaptic membrane. The C_2A domain is rich in aspartic acid residues which bind Ca^{2+}.[147] Previous studies, have introduced a single cysteine mutant at position 95, to attach a

radionuclide or fluorescent probe, in order to study cell apoptosis *in vivo*.[148] **Figure 5.3,** shows the β-bulge structure of this C_2 domain.[149]

Figure 5.3: Structure of C2Am (Protein Data Bank: 1RSY) determined by X-Ray crystallography.

5.1.4. Naturally Occurring Proteins with Free Cysteine Sites

Human serum albumin (HSA)

HSA is the most abundant protein in human plasma at a concentration of 640 μM and plays an important role as a transporter of hormones, fatty acid and metals (66.5 kDa, 585 residues), **Figure 5.4**. To perform its role, HSA has a number of potential ligand binding sites, including a free thiol, Cys34. This protein is thought to play a vital role in the transport of metal-based drugs to tumours, and it has even been suggested that adducts of Ru-HSA are the active forms of administered pro-drugs *in-vivo*.[150,151]

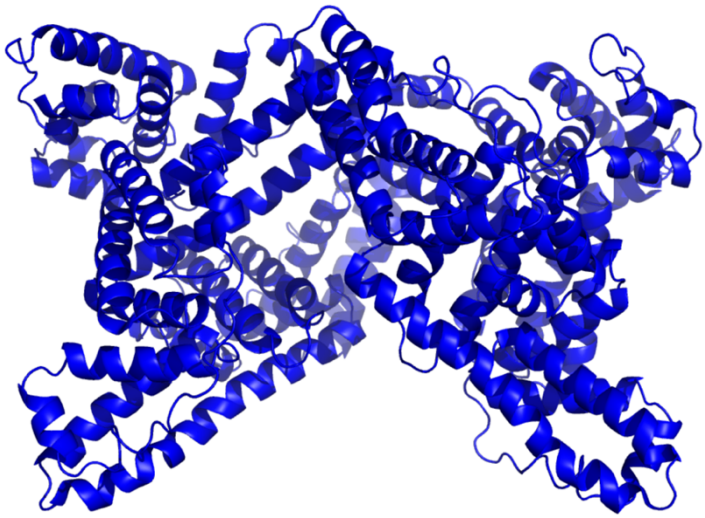


Figure 5.4: Structure of Human Serum Albumin (Protein Data Bank: 1AO6) determined by X-Ray crystallography.[152]

Papain

Papain is a cysteine protease from the latex of the papaya fruit (23.3 kDa, 212 residues). The active site contains the catalytic dyad, which is formed from the single free cysteine, Cys25, and His159, **Figure 5.5**. The cysteine nucleophilicity is enhanced as the histidine deprotonates the thiol side chain, enabling it to catalyse the hydrolysis of the peptide bond. Papain has a further 6 cysteine residues which are involved in disulphide bridges.

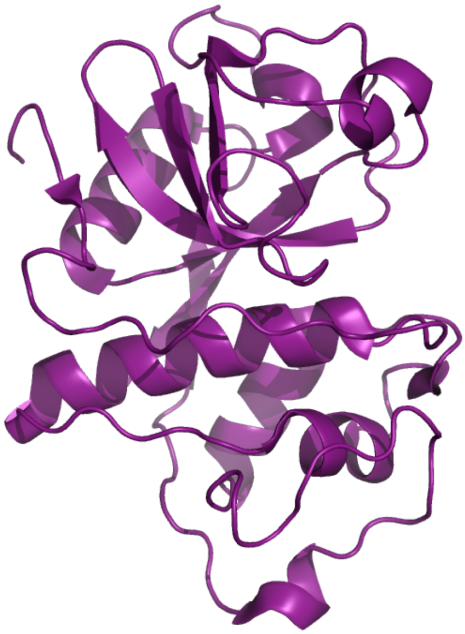


Figure 5.5: Structure of Papain (Protein Data Bank: 9PAP) determined by X-Ray crystallography.[153]

Tobacco Etch Virus Protease

Tev protease is a cysteine protease from the tobacco etch virus. The entire viral genome is translated as a single polypeptide which is cleaved into functional units by three different proteases, with Tev protease itself being one. The enzyme is specific for the amino acid sequence ENLYFQ\S, where \ represents the cleavage site, and is often used to cleave engineered fusion proteins and solubility enhancing proteins (e.g. the maltose binding protein) from the protein of interest. Tev protease has a surface Cys130, the active-site Cys151, and a disulphide bridge. The active residue Cys151 has been mutated to an alanine in the crystal structure, shown by **Figure 5.6.**

Figure 5.6: Structure of Tev Protease C151A (Protein Data Bank: 1Q31) determined by X-Ray crystallography.[154] The protein is crystallised as a dimer, however, in solution acts as a monomer, therefore the monomeric protein is shown here.

These seven different proteins were explored due to their variety in structure and function, and the presence of free cysteine residues which could be excellent targets for ruthenium bioconjugation.

5.2. Results and Discussion

5.2.1. Experimental Strategy

Throughout this chapter the ruthenium complexes incubated with different proteins either had the η^6-*p*-cymene arene ligand (complexes [3], [11] and [15]) or the η^6-hexamethylbenzene arene ligand (complexes [4], [12] and [16]).

To initially gauge the reactivity of each individual cysteine containing protein, the proteins were incubated with a 20-fold excess of the non-fluorinated *p*-cymene complex [3], **Section 5.2.2**.

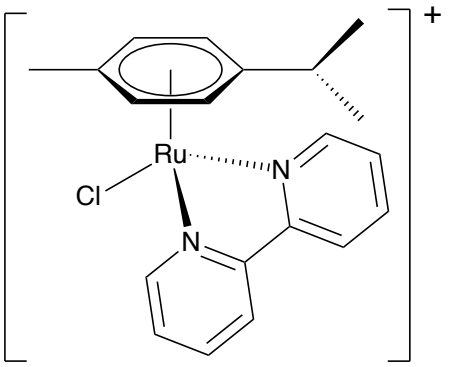


Figure 5.7: Structure of complex [3]. Anion omitted for clarity.

Two proteins, Ubq K63C and Cyt b_{562} L10C were then incubated with a suite of *p*-cymene and hexamethylbenzene bipyridine complexes; both the non-fluorinated complexes [3] and [4] and the fluorinated complexes [11], [12], [15] and [16], summarised in **Section 5.2.3** (for Ubq K63C) and **Section 5.2.4** (for Cyt b_{562} L10C).

An incubation involved mixing known concentrations of protein, ruthenium complex (from a stock solution in DMF) and sodium phosphate buffer in a temperature-controlled shaker. Incubations were periodically analysed by liquid chromatography mass spectrometry (LC-MS) (full details in **Section 3.2.4**). LC-MS is the most efficient method of separating and confirming the presence of ruthenium-protein adducts. The instrument used in this research is routinely capable of detecting proteins in samples with concentrations as low as 1 μM. A typical methodology in going from the chromatogram to the ion series and then to the deconvoluted mass spectra is demonstrated in **Appendix B, Figure B.1**.

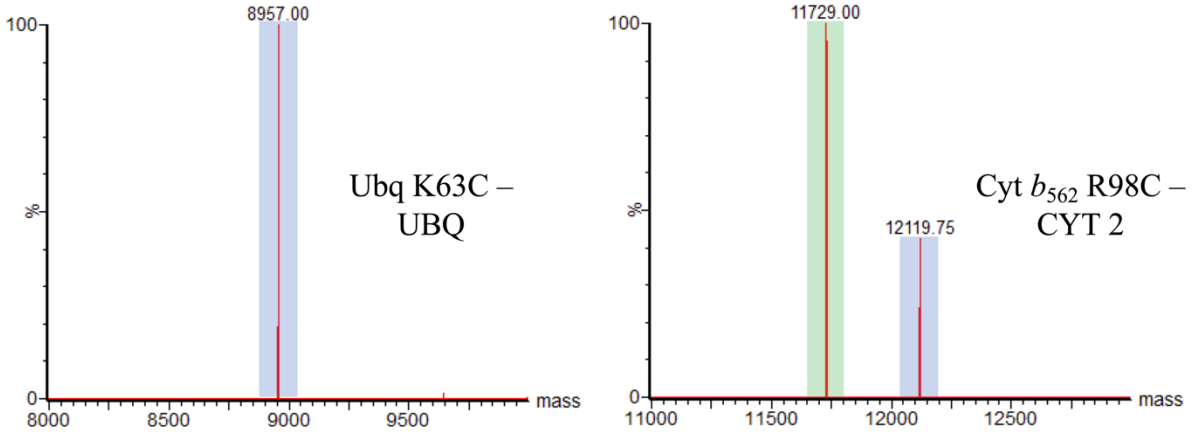
Critical Evaluation of Mass Spectra

To monitor the extent of protein modification a C4 reverse-phase column was used to separate small molecules from proteins, then the protein components were analysed *via* mass spectrometry under denaturing conditions. For significantly different ions, mass spectrometry is not considered to be a quantitative method where the peak integrals of ion counts correlate with concentration, however, throughout this research, peak heights of modified and un-modified protein have been compared. This is possible through a maximum entropy algorithm which deconvolutes the detected ion series of the protein components to give a so called 'zero charge' spectrum. This allows a comparison of the intensities of molecular weight peaks, with

a good signal to noise ratio. Further detail of the measurement and analysis procedure is given in **Section 3.2.4**.

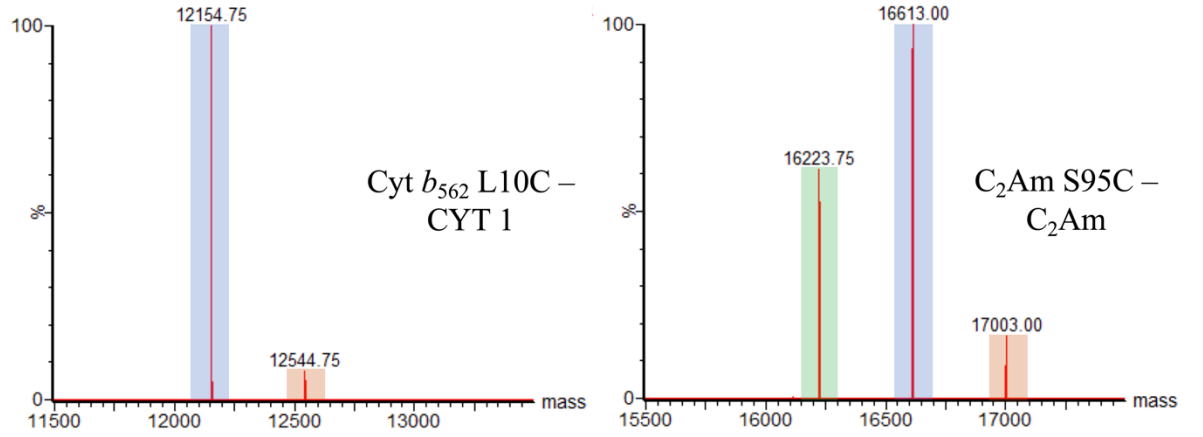
5.2.2. Reactivity of Ru(II) Arene Complexes with Different Proteins

After optimizing conditions for reaction of these ruthenium complexes with cysteine containing proteins, mass spectra can quickly be compared to reveal different reactivities. In these incubations a 20-fold excess of the ruthenium complex was used and modification was monitored at a number of different time points, with spectra chosen after 1 hr to give a clear representation of the reactivity of the different proteins. **Figure 5.8** shows results from incubation of complex [3] with six different proteins.



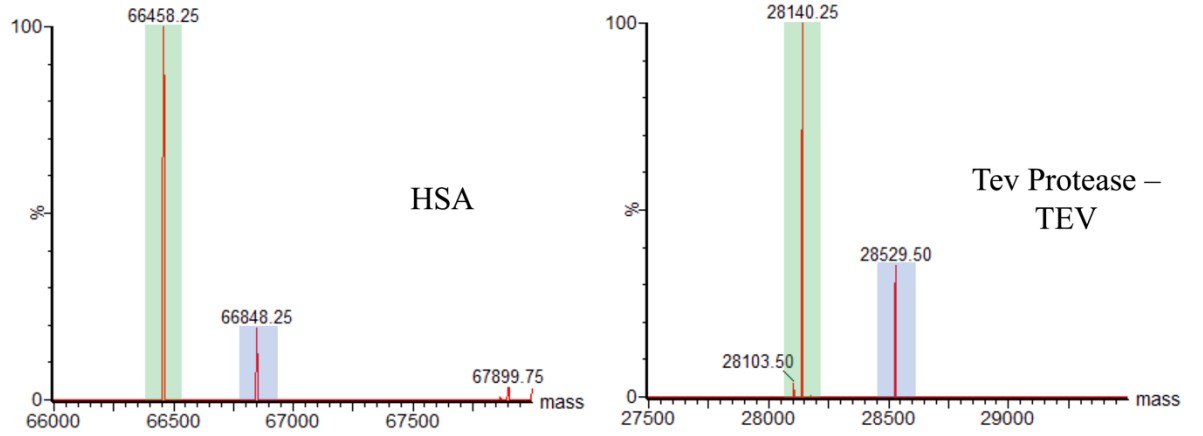
M_{obs}	M_{calc}	Adduct Type
8957	8959	UBQ + [Ru(Cym)(Bipy)]

M_{obs}	M_{calc}	Adduct Type
11729	11727	CYT 2
12120	12119	CYT 2 + [Ru(Cym)(Bipy)]



M_{obs}	M_{calc}	Adduct Type
12155	12156	CYT 1 + [Ru(Cym)(Bipy)]
12545	12548	CYT 1 + 2 [Ru(Cym)(Bipy)]

M_{obs}	M_{calc}	Adduct Type
16224	16222	C_2Am
16613	16614	C_2Am + [Ru(Cym)(Bipy)]
17003	17006	C_2Am + 2[Ru(Cym)(Bipy)]



M_{obs}	M_{calc}	Adduct Type
66458	66455	HSA + [O]
66848	66487	HSA + [O] + [Ru(Cym)(Bipy)]

M_{obs}	M_{calc}	Adduct Type
28140	28139	TEV
28429	28531	TEV + [Ru(Cym)(Bipy)]

Figure 5.8: Mass spectra of samples from incubations of complex [3] with six different cysteine containing proteins (50 µM protein, 20 Eq. [3], 310 K, 1 hr).

The cysteine mutant of ubiquitin is very reactive to ruthenium modification. This single cysteine site is on the surface of the protein and complete modification occurs within 1 hr. The two cytochrome b_{562} cysteine mutants differ in reactivity quite markedly towards the ruthenium complex. This protein is highly dynamic and complete modification of residue Cys10 occurs after 1 hr, alongside modification at a slower rate at another residue. In a series of time course experiments varying the number of equivalents of complex [3], the reactivity of the different cytochrome cysteine mutants has been explored further. From initial analysis of the NMR structure, it is difficult to predict whether or not the Cys at position 98 would be much less reactive, however, with 20 Eq. of complex [3] this variant takes over 12 hrs to be completely modified. Complete modification of the L10C mutant can be achieved with 2 Eq. of complex [3] incubated for 4 – 6 hrs.

The final cysteine mutant C_2Am S95C shows presence of two species, one with a single ruthenium addition to the protein and another with an additional modification at a secondary amino acid residue. For C_2Am, the presence of multiple coordination sites is not surprising, as this protein domain is rich in acidic side chain residues, which, as shown in Chapter 4, can coordinate to Ru(II) arene complexes.

Papain has one free cysteine residue which is at the heart of the protease active site. The mass spectrum of complex [3] with papain is not shown in **Figure 5.8**, as no modification is observed at all which is probably because access to this cysteine residue is restricted. The other cysteine protease explored was Tev protease, which shows moderate reactivity with complex [3]. It is unlikely that this is modification at the active site cysteine as this Cys residue is buried, so conjugation to an alternative surface exposed residue is more likely.

Finally, HSA initially shows little reactivity with the ruthenium complex, however, when incubated for a 24-hour period a number of different ruthenium modifications are observed. Due to the increased reactivity of a number of residues on HSA it makes it difficult to localise a ruthenium cofactor to a specific location on this protein.

Confirming Cysteine Coordination using N-Ethyl Maleimide

After identifying that Ubq K63C and Cyt b_{562} L10C were the most reactive cysteine containing proteins, the next step was to confirm cysteine as the site of ruthenium conjugation. **Figure 5.9 (Top)** demonstrates, how the organic molecule N-ethyl maleimide (NEM) is reactive towards thiols, therefore modifying the cysteine residue, which can be detected *via* LC-MS. If, however, the free cysteine site is blocked by coordination to a metal complex then addition of the

maleimide to the reaction mixture will not further change the mass of the modified protein, **Figure 5.9 (bottom)**. Example spectra of how N-ethyl maleimide is used throughout this thesis to confirm the presence of a free cysteine residue is shown in **Appendix B, Figure B.2**.

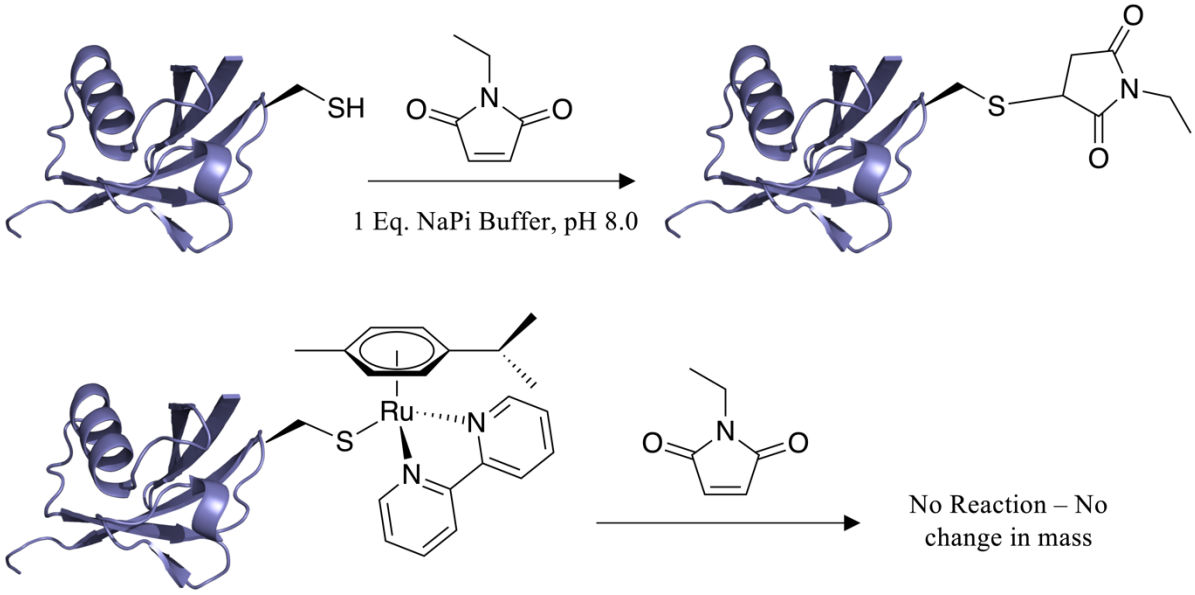
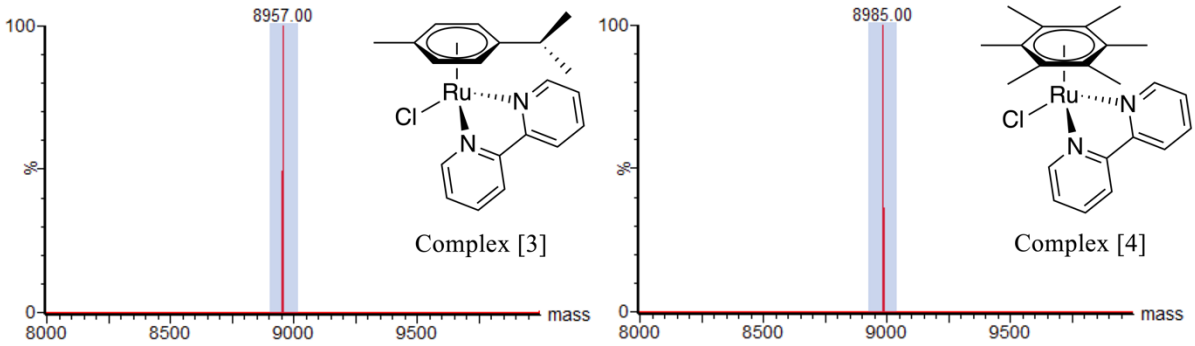


Figure 5.9: (Top): How N-ethyl maleimide reacts with free accessible thiol residues on proteins. **(Bottom):** When the cysteine thiol is coordinated to ruthenium no change in mass is observed *via* LC-MS upon reaction with the maleimide

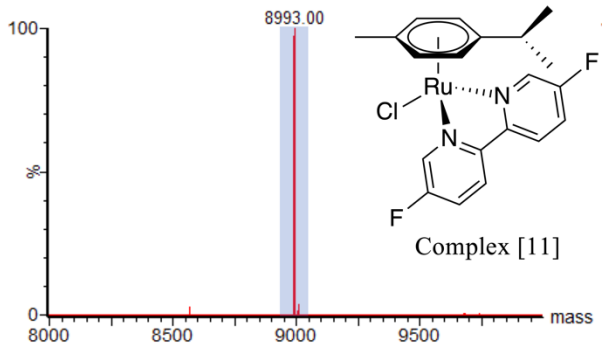
5.2.3. Reactivity of Ubiquitin with a Suite of Ruthenium Complexes

Six ruthenium complexes were incubated with Ubq K63C (8565 Da). The η^6-arene ligand was varied from *p*-cymene to hexamethylbenzene, and the bipyridine ligand from bipyridine (Bipy) to 5,5'-difluorobipyridine (5,5'-FBipy) to 5,5'-di(trifluoromethyl)bipyridine (5,5'-TFMBipy). These modifications gave changes in the electronic and steric properties of the ligands and the mass spectra after 1 hr of incubation are summarised in **Figure 5.10**.

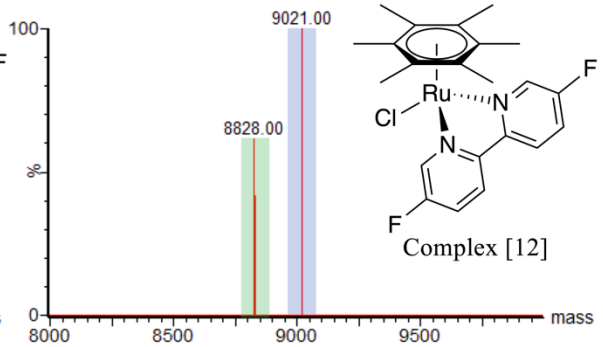


M_{obs}	M_{calc}	Adduct Type
8957	8959	UBQ + [Ru(Cym)(Bipy)]

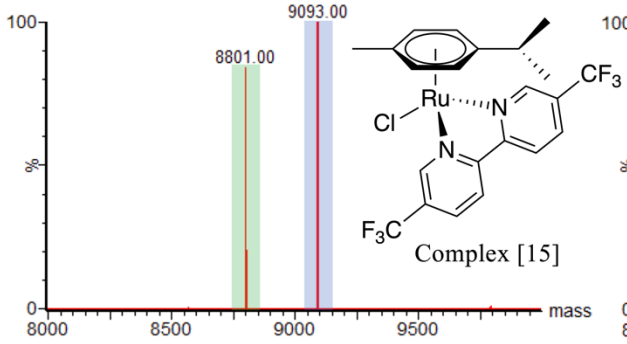
M_{obs}	M_{calc}	Adduct Type
8985	8994	UBQ + [Ru(HMB)(Bipy)]



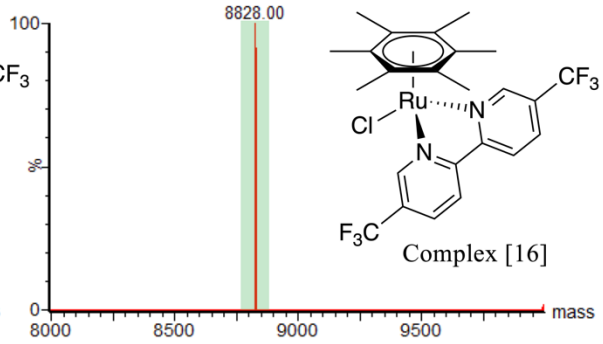
M_{obs}	M_{calc}	Adduct Type
8993	8994	UBQ + [Ru(Cym)(5,5'-FBipy)]



M_{obs}	M_{calc}	Adduct Type
8828	8830	UBQ + [Ru(HMB)]
9021	9023	UBQ + [Ru(HMB)(5,5'-FBipy)]



M_{obs}	M_{calc}	Adduct Type
8801	8802	UBQ + [Ru(Cym)]
9093	9094	UBQ + [Ru(Cym)(5,5'-TFMBipy)]



M_{obs}	M_{calc}	Adduct Type
8828	8830	UBQ + [Ru(Cym)(5,5'-FBipy)]

Figure 5.10: Mass spectra from incubations of complexes [3], [4], [11], [12], [15] and [16] with Ubq K63C (50 µM protein, 20 Eq. Ru, 310 K, 1 Hr).The structures of the singly charged complexes (PF_6 removed) added at the start of the incubation are given.

In all cases, complete conversion to modified protein is observed after 1 hour with 20 Eq. of ruthenium complex, i.e. no unmodified Ubq K63C of mass 8565 Da is observed. When the fluorinated complexes [12], [15] and [16] are incubated with Ubq K63C, ruthenium-protein adducts are detected that have masses consistent with the dissociation of the fluorinated bipyridine ligand (either 5,5'-FBipy or 5,5'-TFMBipy) from the metal.

From an incubation of complex [11] with Ubq K63C the major mass peak observed is at 8993 Da which is consistent with the adduct Ubq + [Ru(Cym)(5,5'-FBipy)]. Note here, that the ruthenium complex has exchanged its chloride ligand to coordinate to the protein. On incubation of Complex [16] with Ubq K63C the only significant mass peak observed is at 8828 Da. This mass is consistent with a Ubq + [Ru(HMB)] adduct where the 5,5'-TFMBipy ligand has dissociated. Data from an incubation of complex [12] with Ubq K63C demonstrate that dissociation of the 5,5'-FBipy ligand can also occur, when η^6-HMB is coordinated. Additionally, 5,5'-TFMBipy dissociation can occur when the *p*-cymene ligand is bound, as evidenced from mass spectra of complex [15] with Ubq K63C.

Observing bipyridine dissociation from the metal upon coordination to a protein is an unexpected and key result. Bipyridine dissociation leaves two free coordination sites at the ruthenium centre, and alongside chloride exchange there are three free ruthenium coordination sites available. The observed masses reveal that no solvent molecules or buffer salt ions are coordinated to the metal centre, therefore, for the metal to become coordinatively saturated, the protein must provide three ligands. Determining the mechanism and identifying these amino acid ligands is challenging, particularly when the fluorinated probe is no longer coordinated to the metal centre. All masses observed rule out any dimerization or higher order aggregation that could potentially arise from the metal centre cross-linking two protein species.

The differential reactivity of the varied complexes shows that dissociation of bidentate bipyridine ligands is increased by two key factors: (i) changing the η^6-arene ligand from *p*-cymene to hexamethylbenzene and (ii) changing the bipyridine ligand from Bipy to 5,5'-FBipy to 5,5'-TFMBipy. As previously discussed in Chapter 4, the hexamethylbenzene ligand is a more electron donating ligand than the *p*-cymene ligand, resulting in a more electron rich metal. When (η^6-HMB) is coordinated to ruthenium there is a lesser requirement for electron density from the bipyridine ligands leading to weaker Ru–N bonds and a more labile bipyridine ligand. Following the same logic, introducing electron withdrawing substituents on the bipyridine ring with increasing strength H→F→CF_3 reduces the electron donation to the metal from the nitrogen lone pairs and weakens the Ru–N bonds. This hypothesis does not, however, consider the π contribution to bonding, as all three bipyridine ligands are capable of π-accepting electron density. The logic here would be that the more electron-donating arene ligand, coupled with the more inductively electron-withdrawing bipyridine ligands should increase the Ru–N π-back-bonding and lead to a strengthening of the Ru–Bipy bonds. The data from incubation of

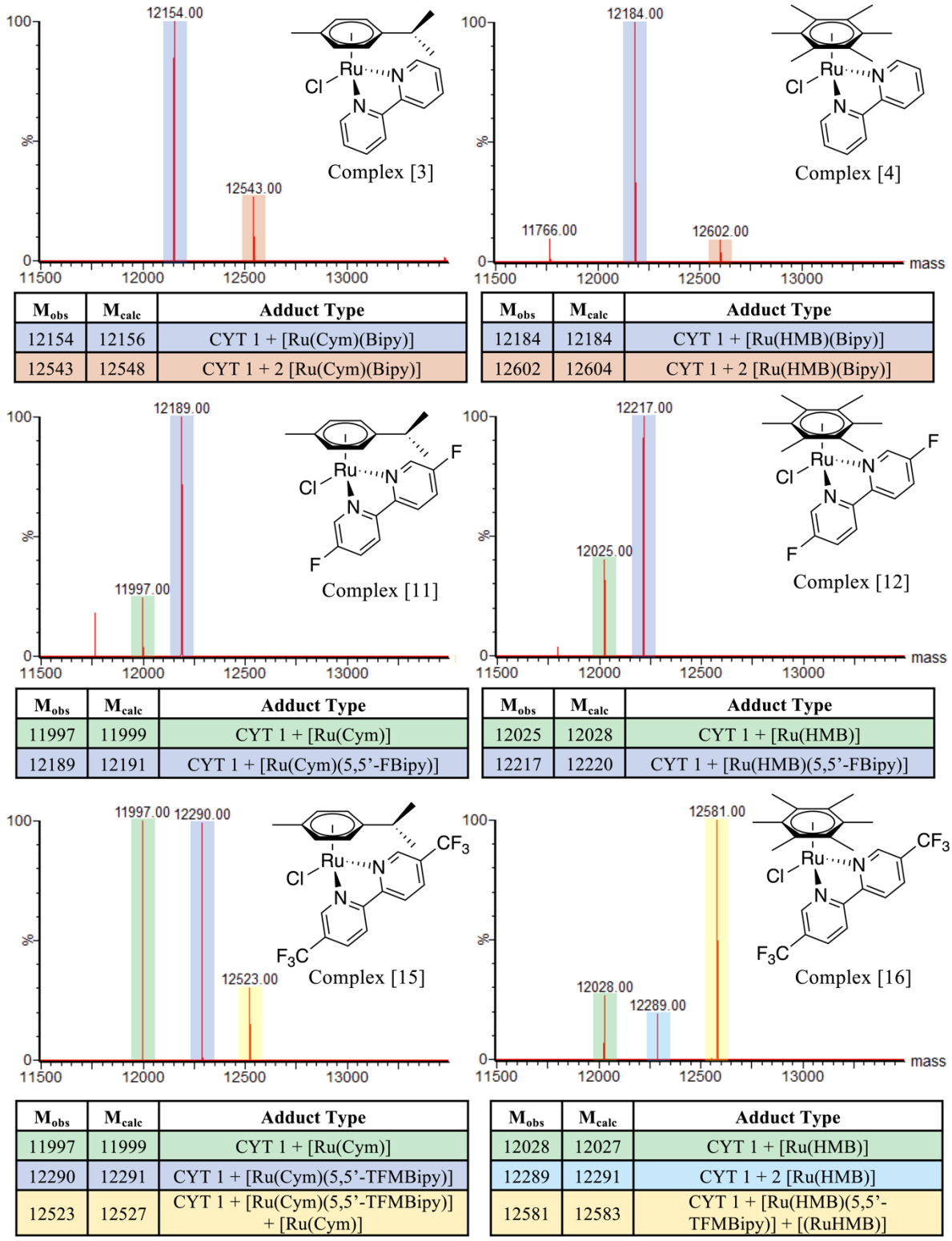
the above complexes with Ubq K63C suggest that the lability of these complexes is dominated by σ rather than π interactions.

Attempts to Crystallise Ubq-Ruthenium Adducts

Attempts have been made to identify the other amino acids ligands coordinated to the ruthenium centre upon bipyridine dissociation *via* X-ray crystallography. Many Ubq-Ru adducts were isolated and purified *via* cation-exchange chromatography. Attempts to crystallise these hybrids were performed using a Mosquito robot to dispense nanolitre samples into a variety of solvent conditions in 96 well plates. Although initial promising results were observed, **Figure 5.11**, no diffraction quality crystals have been generated yet. This is most likely because ubiquitin contains a high surface density of lysine residues, which will inhibit packing conditions and impede crystallisation.[155]

Figure 5.11: Initial crystals of a Ubq-Ru hybrid.

5.2.4. Reactivity of Cytochrome b_{562} L10C with a Suite of Ruthenium Complexes



M_{obs}	M_{calc}	Adduct Type
12154	12156	CYT 1 + [Ru(Cym)(Bipy)]
12543	12548	CYT 1 + 2 [Ru(Cym)(Bipy)]

M_{obs}	M_{calc}	Adduct Type
12184	12184	CYT 1 + [Ru(HMB)(Bipy)]
12602	12604	CYT 1 + 2 [Ru(HMB)(Bipy)]

M_{obs}	M_{calc}	Adduct Type
11997	11999	CYT 1 + [Ru(Cym)]
12189	12191	CYT 1 + [Ru(Cym)(5,5'-FBipy)]

M_{obs}	M_{calc}	Adduct Type
12025	12028	CYT 1 + [Ru(HMB)]
12217	12220	CYT 1 + [Ru(HMB)(5,5'-FBipy)]

M_{obs}	M_{calc}	Adduct Type
11997	11999	CYT 1 + [Ru(Cym)]
12290	12291	CYT 1 + [Ru(Cym)(5,5'-TFMBipy)]
12523	12527	CYT 1 + [Ru(Cym)(5,5'-TFMBipy)] + [Ru(Cym)]

M_{obs}	M_{calc}	Adduct Type
12028	12027	CYT 1 + [Ru(HMB)]
12289	12291	CYT 1 + 2 [Ru(HMB)]
12581	12583	CYT 1 + [Ru(HMB)(5,5'-TFMBipy)] + [(RuHMB)]

Figure 5.12: Mass spectra from incubations of complexes [3], [4], [11], [12], [15] and [16] with Cyt b_{562} L10C (50 µM protein, 20 Eq. Ru, 310 K, 1 Hr). The structures of the singly charged complexes (PF_6 removed) added at the start of the incubation are given.

From incubations of Cyt b_{562} L10C with the non-fluorinated bipyridine complexes [3] and [4], the major species are observed at 12154 Da and 12184 Da, respectively, consistent with one addition of the relevant [Ru(η^6-arene)(Bipy)] fragment in both cases. Masses consistent with two additions of the [Ru(η^6-arene)(Bipy)] fragment are also observed in the mass spectra for both [3] and [4] at 12543 Da and 12602 Da, respectively.

Incubations of the 5,5'-FBipy complexes [11] and [12] with Cyt b_{562} L10C gave major species at 12189 Da and 12217 Da, which, similar to the non-fluorinated bipyridine complexes, are consistent with coordination of the relevant [Ru(η^6-arene)(5,5'-FBipy)] fragment. The minor species in both incubations at 11997 Da and 12025 Da, can be explained by 5,5'-FBipy dissociation, as seen following reaction of complexes [11] and [12] with Ubq K63C. Cyt b_{562} L10C is a dynamic and flexible protein, therefore upon 5,5'-FBipy dissociation, amino acid ligands from other helices on the protein are possibly satisfying the ruthenium coordination environment.

The increased lability of the 5,5'-TFMBipy ligand, compared to 5,5'-FBipy, is clear when complex [16] is incubated with Cyt b_{562} L10C; the major species in this incubation**,** with a mass of 12581 Da, is consistent with Cyt b_{562} L10C with two ruthenium modifications. The mass, which is + 816 greater than the apoprotein is consistent with addition of one fragment of [Ru((η^6-HMB)(5,5'-TFMBipy)] and one of [Ru((η^6-HMB)].

Although direct injection LC-MS is a powerful tool for identifying the different adducts that form upon incubations of different ruthenium complexes with Cyt $b_{562,}$ information on the location of coordination and protein stability is not accessible. Therefore, further characterisation techniques, in particular, NMR, UV-visible spectroscopy and circular dichroism were required in order to probe coordination sites and the influence of metal coordination on the protein chemistry

5.2.5. Characterising Cytochrome b_{562} L10C – Ruthenium Adducts

To further characterise some of these Ru-protein adducts it was necessary to complete the incubations on a larger scale and purify the protein adducts. Once pure samples were obtained, ^{19}F NMR, UV-Vis spectroscopy and circular dichroism experiments were performed. Initially, size exclusion chromatography was used to separate unreacted metal complex and protein based on size. On further analysis of the protein fractions eluted from the column, (by LC-MS and UV-Vis) it was concluded that excess metal complex was non-coordinatively associated with Cyt b_{562}. The hydrophobic ligands can form non-covalent interactions with the

hydrophobic regions of this protein without metal coordination occurring. Changing purification method to anion exchange chromatography with a salt gradient, successfully removes all excess small molecules, and non-coordinatively bound metal complex, leaving just coordinatively bound protein-ruthenium adducts.

A series of experiments were performed to purify a number of Cyt b_{562} – Ruthenium adducts at a higher concentration. To do this, 20 Eq. of the ruthenium complex were incubated with 2 mL of Cyt b_{562} L10C at a concentration of 200 μM for 1 hr. The solution was then buffer exchanged into a high pH buffer and purified *via* anion exchange chromatography, **Figure 5.13**.

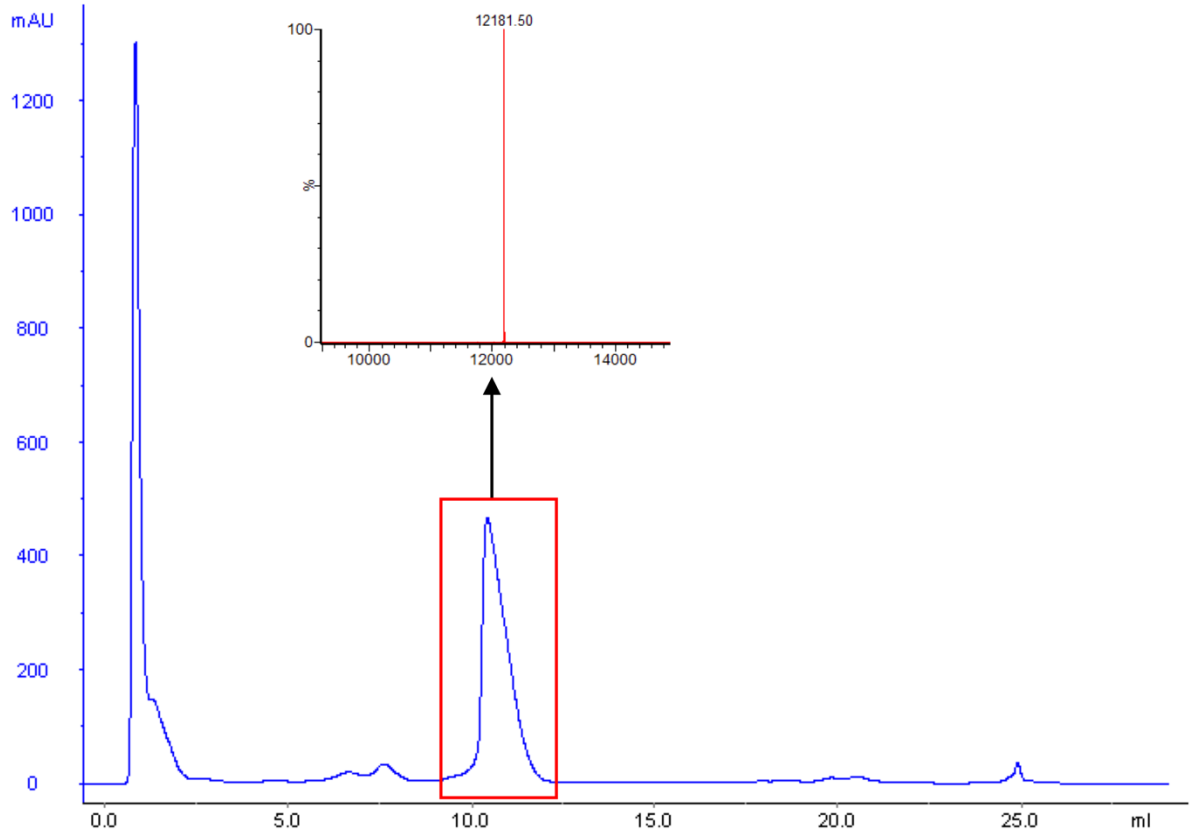


Figure 5.13: (Blue) The UV trace at 280 nm from an anion exchange purification of 20 Eq. of complex [4] with Cyt b_{562} L10C. The first peak to come off the column is excess complex [4] which is positively charged therefore has no affinity to the positively charged resin. The major peak (highlighted in red) corresponds to the adduct Cyt b_{562} L10C + [Ru(HMB)(Bipy)], as shown by the deconvoluted mass spectrum (insert) M_{obs} = 12182, M_{calc} = 12184.

Following the incubation of complex [4] with Cyt b_{562} L10C, larger scale incubations were performed with the fluorinated complexes [11], [12], [15] and [16]. In these experiments the major adducts were isolated *via* ion exchange with high levels of control, and the mass spectra demonstrate a high level of purity in the samples, **Figure 5.14**. The masses of these isolated

adducts are shown in **Table 5.2**. Cyt – [11], Cyt – [12] and Cyt – [15] all correspond to a single [Ru] addition with the arene ligand and fluorinated bipyridine ligand remaining coordinated to the metal centre. Cyt – [16] corresponds to an adduct with two ruthenium additions, one addition of [Ru(HMB)(5,5'-TFMBipy)] and one addition of [Ru(HMB)] where the 5,5'-TFMBipy has dissociated from the metal centre.

Table 5.2: Modifications and names of Ru-Cyt b_{562} L10C adducts purified by ion exchange.

Name	M_{obs}	M_{calc}	Modification
Cyt – [11]	12193	12191	Cyt b_{562} L10C + [Ru(*p*-cym)(5,5'-FBipy)]
Cyt – [15]	12293	12291	Cyt b_{562} L10C + [Ru(*p*-cym)(5,5'-TFMBipy)]
Cyt – [12]	12221	12220	Cyt b_{562} L10C + [Ru(HMB)(5,5'-FBipy)]
Cyt – [16]	12582	12583	Cyt b_{562} L10C + [Ru(HMB)(5,5'-TFMBipy)] + [Ru(HMB)]

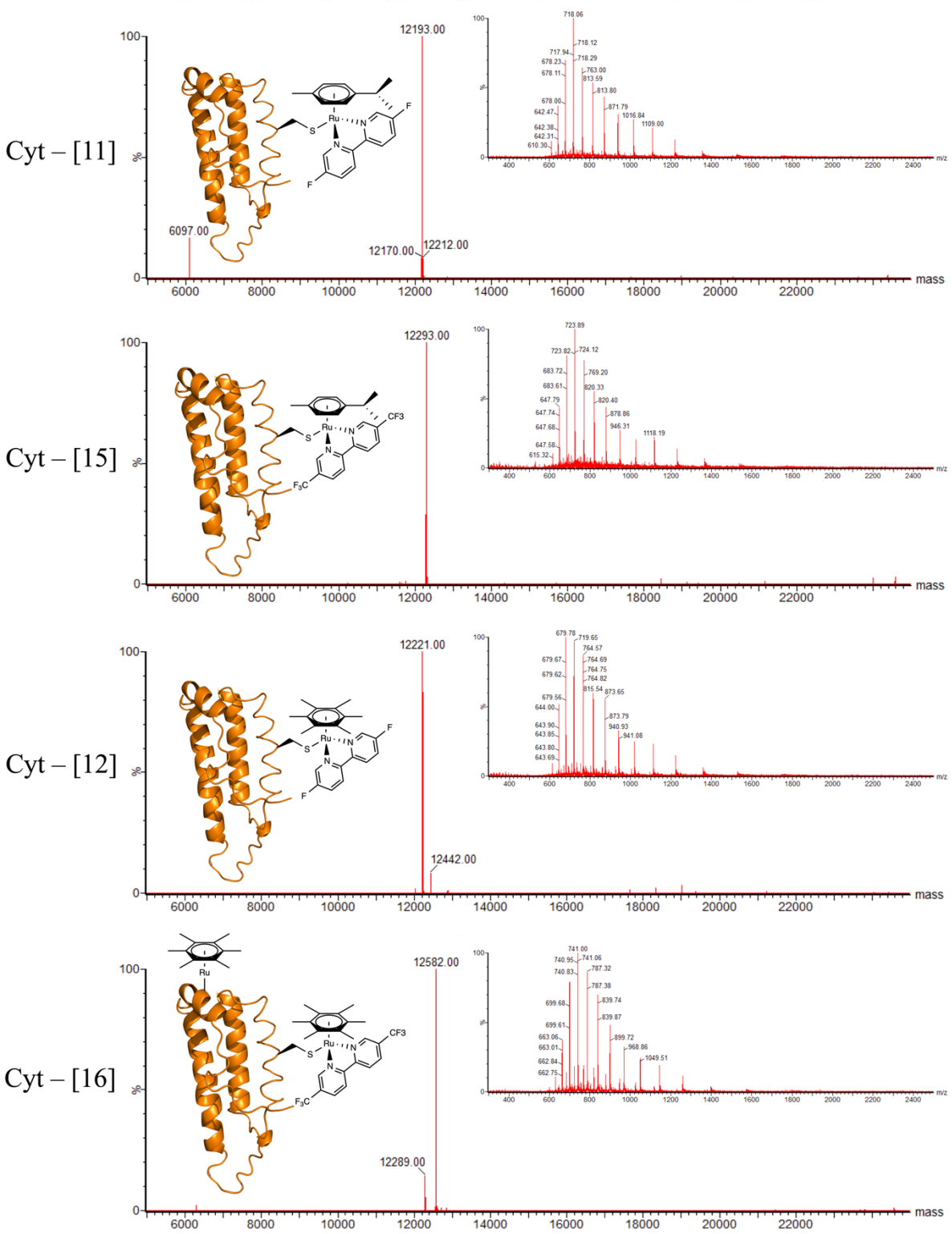


Figure 5.14: Mass spectra from purified Ru-Cyt b_{562} L10C adducts. Insert top right: The ion series of these species. Insert left: Cartoon representation of ruthenium modifications on the protein, shown *via* ChemDraw structures.

^{19}F NMR of Cytochrome b_{562} L10C – Ruthenium Adducts

This section of work was done collaboratively with the NMR spectroscopy department in the Medical Research Council Laboratory of Molecular Biology (MRC-LMB). In particular, Dr Stefan Freund and Dr Trevor Rutherford helped design and perform a number of ^{19}F NMR experiments.

The Cyt b_{562}-Ru adducts isolated Cyt – [11], Cyt – [12], Cyt – [15] and Cyt – [16] all had a fluorinated ligand present, therefore ^{19}F NMR experiments could be used to identify the amino acid coordinated to the metal centre. Optimum T_1 relaxation times were measured using an inversion recovery sequence to determine the T_1 null point, and $T_R = 1.5.T_1$ was used.

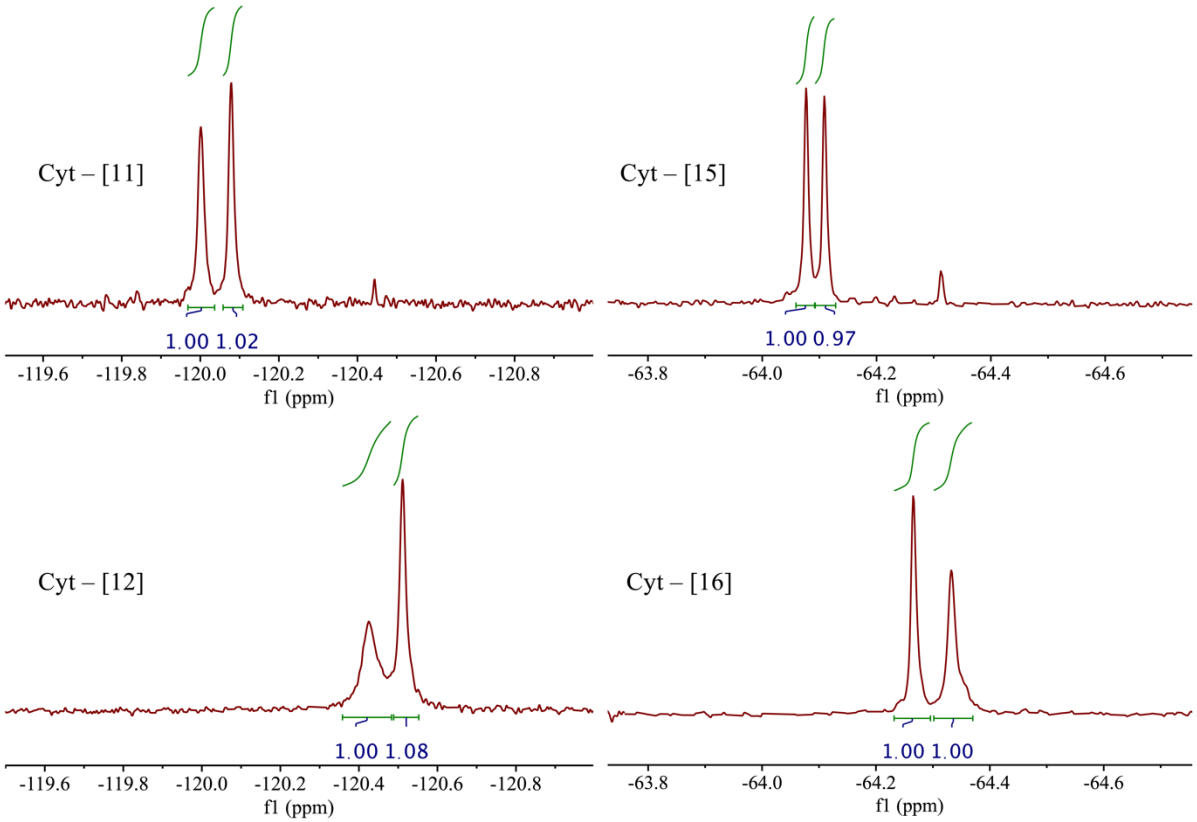


Figure 5.15: ^{19}F{^{1}H} NMR spectra of four different Ru-cytochrome b_{562} L10C adducts. Adducts were placed in 90% Buffer : 10% D_2O and a 256-scan experiment was performed.

In all spectra the Cyt b_{562}-Ru adduct gives rise to two peaks of equal integrals, one peak from each fluorine atom on the bipyridine ring. The chemical shifts are consistent with Ru-cysteine coordination, when compared with the diagnostic spectra in Chapter 4. Additionally, a lack of reaction with NEM confirms that the cysteine is blocked. Taken together, these results confirm that there is a Ru(η^6-arene)(fluorinated bipyridine) unit coordinated to the cysteine 10 position in all four of these adducts.

In contrast to ruthenium coordination to the thiol of N-acetyl cysteine methyl ester, described in Chapter 4, the ^{19}F NMR signals from the ligand in the spectra after protein coordination have different peak geometries, but the same integral values. Therefore, in the Cyt b_{562}-Ru adducts the two fluorine atoms on the bipyridyl ring have different relaxation properties, which suggests they are subtly affected by the local environment they are in. A simple model of Cyt – [11], **Figure 5.16**, shows how the ruthenium fragment [Ru(*p*-Cym)(5,5'-FBipy)]$^{2+}$ could be accommodated into the hydrophobic region of Cyt b_{562} L10C, normally used for haem binding, and coordinate to the cysteine residue. One fluorine atom, shown in turquoise, is solvent exposed, whereas the fluorine on the opposite side of the bipyridyl ring would be buried from solvent. This leads to the hypothesis that the exact local environments of nearby amino acids are determining the relaxation properties of the two fluorine atoms, thus leading to different peak shapes.

As discussed Cyt – [11], Cyt – [12] and Cyt – [15] all contain only one site of modification by a ruthenium, whereas Cyt – [16] has two sites of modification, one where the bipyridine has dissociated from the metal. The NMR chemical shifts of Cyt – [16] suggests that the fluorinated fragment [Ru(HMB)(5,5'-TFMBipy)] is coordinated to a cysteine residue. The other site of modification in Cyt – [16] (a [Ru(HMB)] fragment) must therefore be coordinated elsewhere.

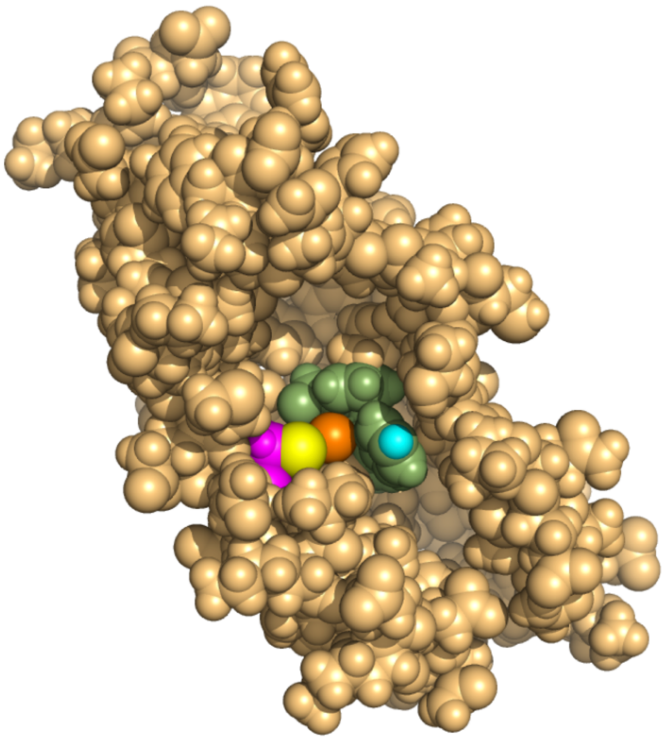


Figure 5.16: A PyMOL generated figure of Cyt – [11]. This is not an energy minimised structure, and uses an NMR structure of apo-Cyt b_{562} and elemental Van der Vaals radii. Yellow – sulphur, orange – ruthenium, turquoise – fluorine, green – organic component of [11].

UV-Vis Spectroscopy of Cytochrome b_{562} L10C – Ruthenium Adducts

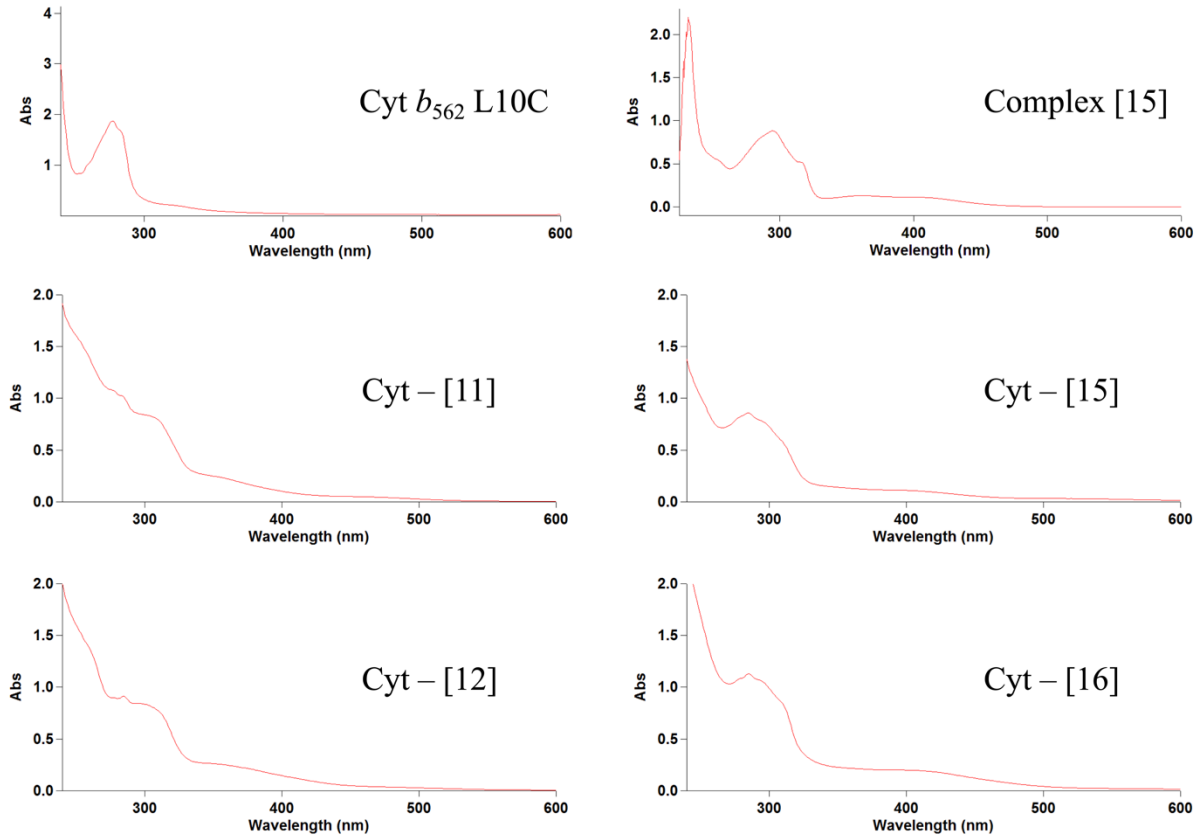


Figure 5.17: UV-visible spectra recorded of Cyt b_{562} L10C and Cyt b_{562}-Ru adducts purified *via* ion-exchange chromatography. Spectra recorded between 200 – 600 nm. Free protein concentration = 600 μM, complex [15] concentration = 50 μM and the concentration of the adducts is unknown due to the overlapping transitions at 280 nm.

As can be seen from the UV spectra in **Figure 5.17**, the free protein shows absorption bands in the range 235 – 290 nm, which are due to the π - π* electronic transitions present in the tyrosine and phenylalanine residues. The Ru(η^6-arene)(bipy) complexes all absorb strongly in the region of 265 – 330 nm, which is due to π - π* transitions in the bipyridine and arene ligands.[156] These ligand absorption bands are evident in the Ru-cytochrome adducts, particularly between 300 – 330 nm. The overlap in the bands at 280 nm makes quantification of the concentration of these adducts particularly difficult, as the extinction coefficient for each ruthenium fragment is different. What is of particular interest in **Figure 5.17** is the subtle differences in the UV-vis spectra of the adducts, particularly with distinct bipyridine ligands coordinated. It is, however, challenging to extrapolate the changes in the spectra to an understanding of the ligand exchange behaviour of Ru(II) arene complexes with protein scaffolds.

The Ru(η^6-arene)(bipyridine) complexes also absorb weakly in the region of 360 – 440 nm. These bands are particularly evident for the 5,5'-TFM adducts Cyt – [15] and Cyt – [16] and arise from metal to ligand charge transfer (MLCT). The photochemistry of ruthenium tris(bipyridine) complexes has been studied in detail, and these bands have been assigned to originate from a d-electron from the ruthenium being excited into a bipyridine anti-bonding orbital.[157]

Circular Dichroism of Cytochrome b_{562} L10C – Ruthenium Adducts

The purpose of measuring the CD spectra of these Ru-Cyt adducts was to monitor whether the coordination of the ruthenium complexes changes the secondary structure of the protein, in particular whether the protein retains a folded alpha helical structure upon ruthenium coordination. The signature spectral features of helical proteins are a positive band at 190 nm and negative bands at 222 nm and 208 nm. Evident in **Figure 5.18**, the protein-ruthenium hybrids retain their secondary helical structure upon ruthenium coordination, with the characteristic bands at 222 nm and 208 nm clearly shown.

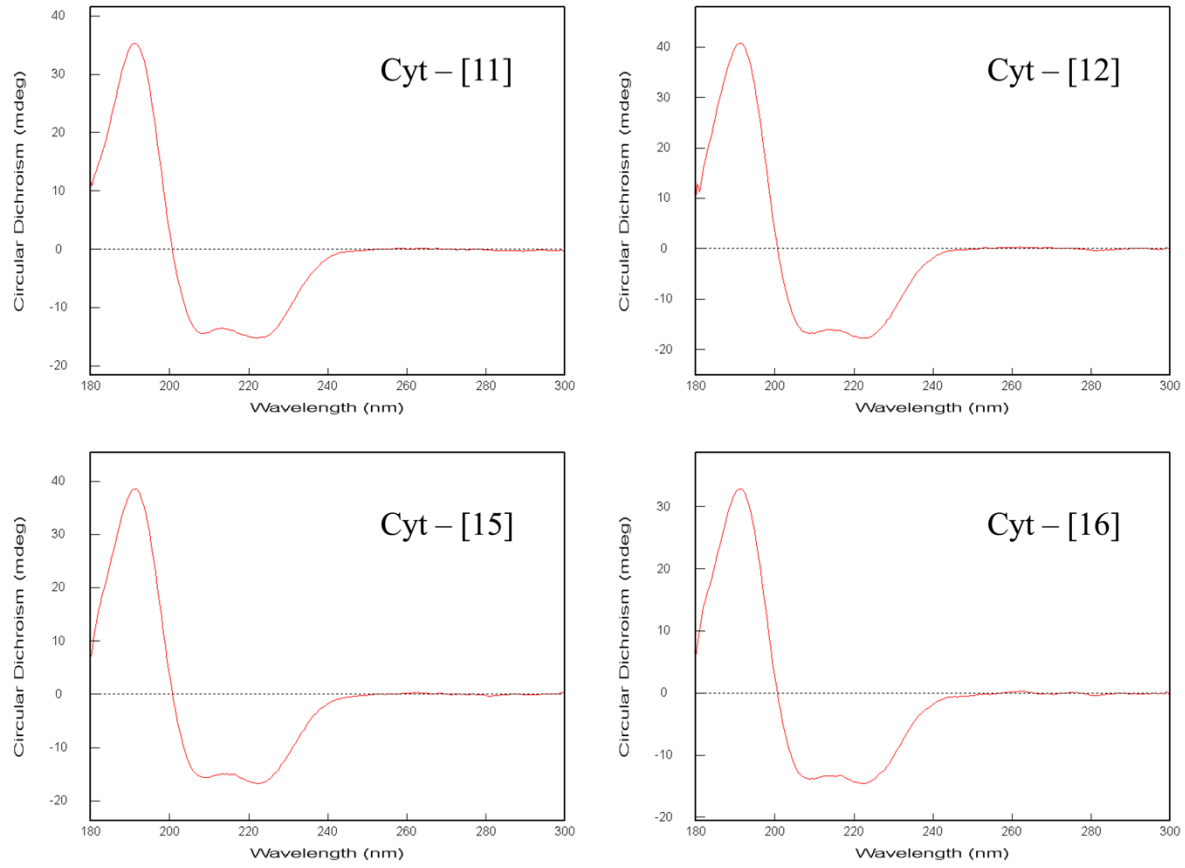


Figure 5.18: CD spectra of the Cyt b_{562} L10C – ruthenium adducts, Cyt – [11], Cyt – [12], Cyt – [15] and Cyt – [16] (5 μM adduct concentration, 10 mM sodium phosphate buffer).

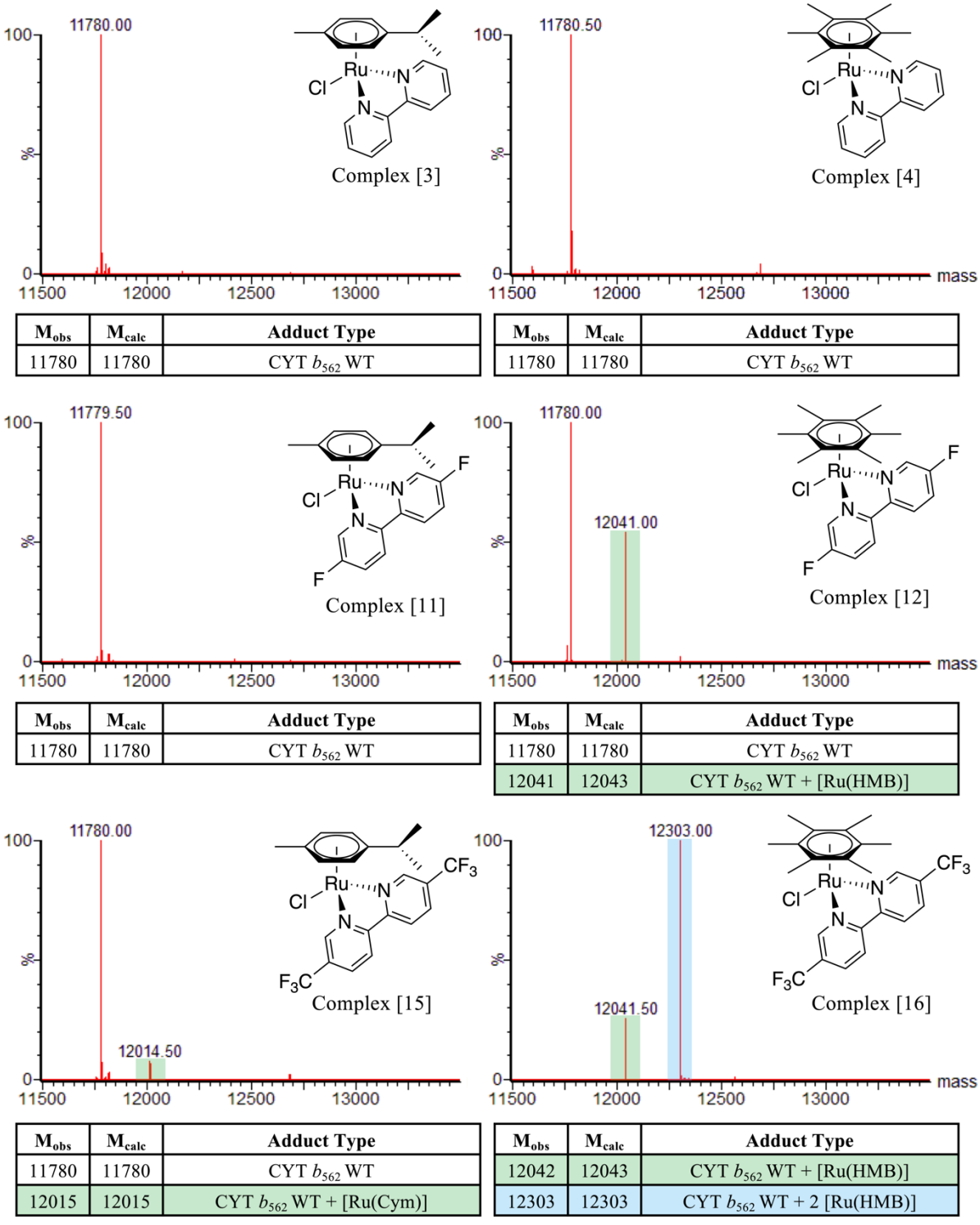
5.2.6. Reactivity of Cytochrome b_{562} WT with a Suite of Ruthenium Complexes

This section of work was done collaboratively with the mass spectrometry department in the Medical Research Council Laboratory of Molecular Biology (MRC-LMB). In particular, Dr Mark Skehel and Sarah Maslen helped design and perform a number of the tandem MS/MS experiments.

The results shown in the previous two sections are focussed on the reactivity and coordination of ruthenium fragments to the cysteine residue in the Cyt b_{562} L10C mutant. In this section, the reactivity of the same Ru(II) arene complexes with Cyt b_{562} wild-type (WT) is shown. In these incubations, shown in **Figure 5.19**, there is no free cysteine for the ruthenium centre to coordinate to. Attempts were made to characterise the Cyt b_{562} WT – Ru adducts formed using a combination of LC-MS and tandem mass spectrometry (MS/MS) methods.

In stark contrast to the reactivity of complexes [3] and [4] with Cyt b_{562} L10C, **Figure 5.12**, there is no modification of Cyt b_{562} WT by the bipy complexes [3] and [4], further emphasising their preference to coordinate to cysteine residues. Additionally, the fluorinated cymene complexes [11] and [15] also show very limited coordination to Cyt b_{562} WT. However, the fluorinated hexamethylbenzene complexes [12] and [16] show increased reactivity towards Cyt b_{562} WT, demonstrating how subtle modulation of the electronic and steric properties of the metal-bound ligands can greatly influence the reactivity with proteins.

In the incubation of complex [16] with Cyt b_{562} WT, the mass spectrum shows two major species at 12303 Da and 12042 Da which correlate to the adducts Cyt b_{562} WT + 2 [Ru(HMB)] and Cyt b_{562} WT + [Ru(HMB)] respectively. Upon ruthenium coordination to the protein scaffold the trifluoromethylbipyridine ligand dissociates from the metal centre. Attempts were therefore made to isolate these species and identify the coordination environment of the [Ru(HMB)] fragment.



M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT b_{562} WT

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT b_{562} WT

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT b_{562} WT

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT b_{562} WT
12041	12043	CYT b_{562} WT + [Ru(HMB)]

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT b_{562} WT
12015	12015	CYT b_{562} WT + [Ru(Cym)]

M_{obs}	M_{calc}	Adduct Type
12042	12043	CYT b_{562} WT + [Ru(HMB)]
12303	12303	CYT b_{562} WT + 2 [Ru(HMB)]

Figure 5.19: Mass spectra from incubations of complexes [3], [4], [11], [12], [15] and [16] with Cyt b_{562} WT (50 µM protein, 20 Eq. Ru, 310 K, 1 Hr). The structures of the singly charged complexes (PF_6 removed) added at the start of the incubation are given.

Where does the [Ru(HMB)] Fragment Coordinate to Cytochrome b_{562} WT?

Initial attempts to purify the doubly modified adduct Cyt b_{562} WT + 2 [Ru(HMB)] (12303 Da) *via* anion exchange proved challenging due to the small overall negative charge of the protein

adduct, which reduces its affinity for the positively charged resin. However, through reducing the number of equivalents of complex [16] from 20 Eq. to 2 Eq., isolation of the adduct Cyt b_{562} WT + [Ru(HMB)] (12042 Da) was achieved. After 1 hr the reaction mixture was purified by anion exchange chromatography and four species eluted from the column, identified by Peaks 1 – 4 on the chromatogram, **Figure 5.20 (top)**. Peak 1 corresponds to unreacted ruthenium complex [16], and Peak 4 corresponds to unreacted Cyt b_{562} WT. The species present in Peak 2 and Peak 3 were analysed by mass spectrometry and found to have the same mass of 12042 Da which is a consistent mass to the adduct Cyt b_{562} WT + [Ru(HMB)], where complex [16] has undergone both chloride and TFMBipy dissociation to coordinate to the protein. Analogous purifications were performed when complex [12] was incubated with Cyt b_{562} WT, and two species of mass 12042 Da were isolated (in Chapter 6 called Hybrid 5 and 6) with different retention times in the anion exchange separation. This section will focus on the differences between the species isolated from an incubation of Cyt b_{562} WT with complex [16].

These two species, Cyt WT – [16] – P2 and Cyt WT – [16] – P3 (in Chapter 6 called Hybrid 7 and Hybrid 8) not only have different retention times, but have a different charge state distribution in the ion series observed in the mass spectra **Figure 5.20 (middle and bottom, insert)**. Analysing the LC-MS, the Cyt b_{562} WT – Ru adduct which elutes off the column first, peak 2, has a major ion of 1095.65 Da corresponding to a +11 charge state under denaturing conditions. The second adduct to elute off the column, peak 3, has a major ion of 861.15 Da, corresponding to a +14 charge state. The influence of tertiary structure on the charge state distribution of proteins is well explored with most data supporting the theory that a shift in charge state distribution towards lower m/z values is induced by protein unfolding. Two major hypothesises have been put forward to interpret this shift: (i) lower solvent accessibility of acidic and basic residues in a more folded structure inhibits ionisation of these residues and (ii) unfolded confirmations can stabilise higher charge states by increasing the distance between two charges of the same polarity.[158]

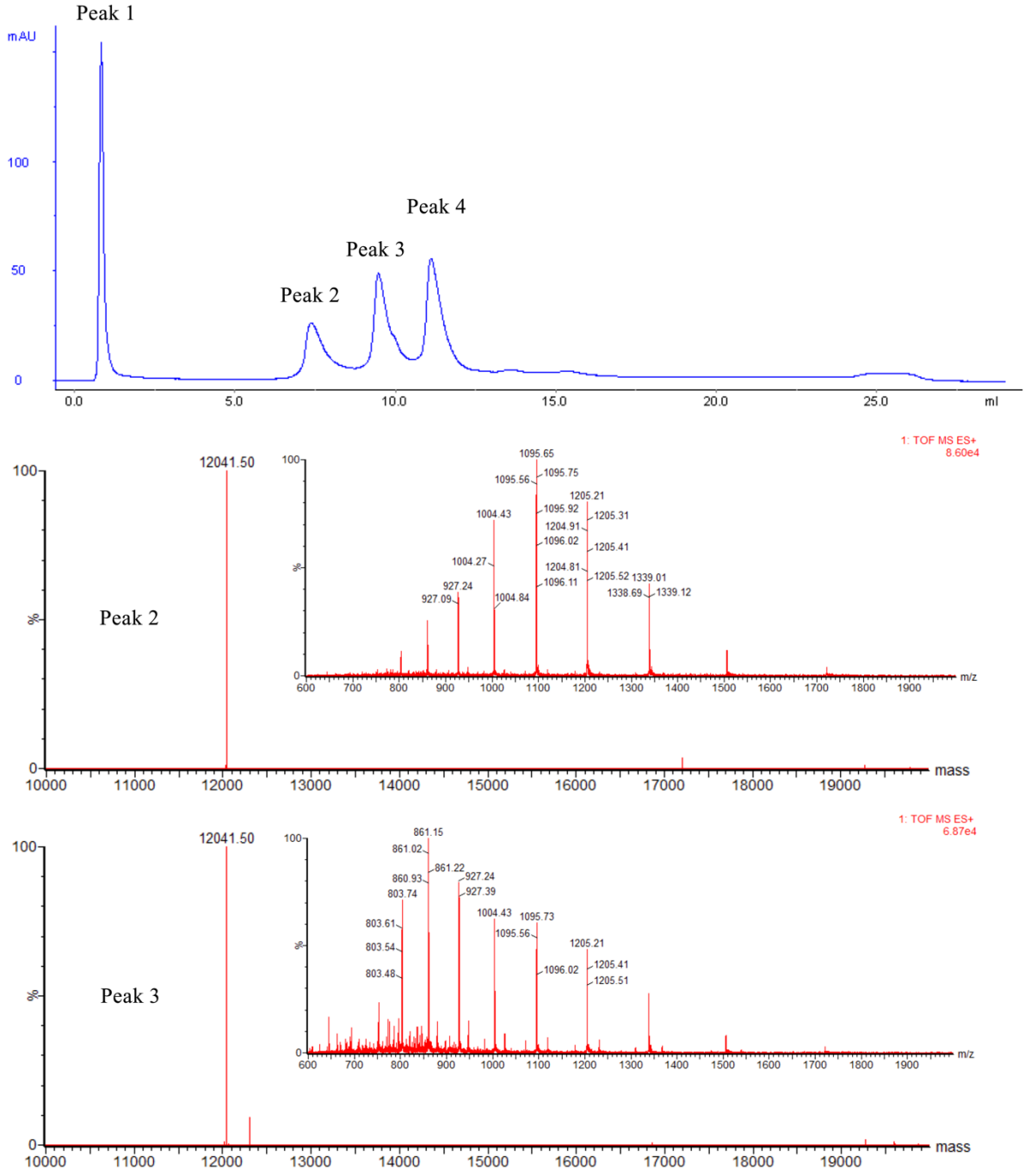


Figure 5.20: (Top): An ion exchange chromatogram, from an incubation of complex [16] with Cyt b_{562} WT. Blue trace – UV absorbance, brown trace – conductivity (S/m), green trace - % salt buffer. (**Middle and Bottom):** Deconvoluted mass spectra and ion series (insert) of the different species that come off the ion exchange column,

The differences in charge state distribution observed between Cyt WT – [16] – P2 and Cyt WT – [16] – P3 suggests that the two protein adducts have different folding properties. To explore this further, CD spectroscopy was used to measure the melting temperatures for the following species, Cyt WT, Cyt WT – [16] – P2, Cyt WT – [16] – P3 and Cyt L10C – [4], **Figure 5.21**. Similar to the ruthenium modified cysteine mutants shown in **Figure 5.18**, the raw CD spectra

between 185 – 300 nm showed little variation in peak shape upon metal coordination, with the negative bands at 208 nm and 222 nm retained. The intensity of the peak at 222 nm for the four different adducts was monitored whilst raising the temperature incrementally from 25 to 90 °C. Fitting these data to a standard logistic function enabled the melting temperatures to be measured.

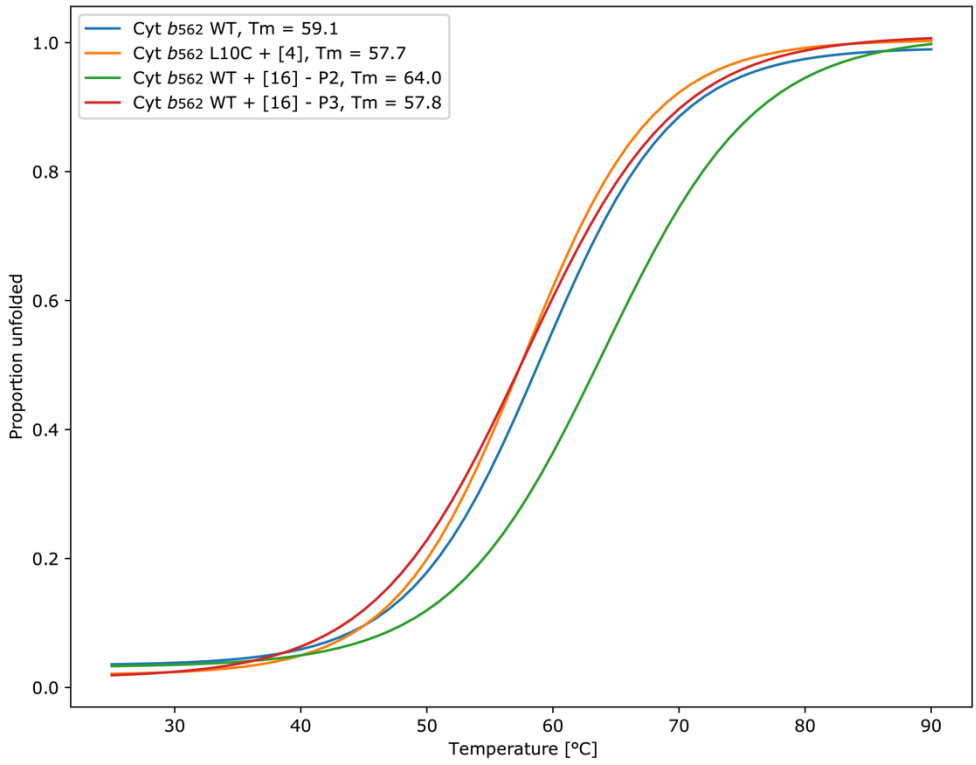


Figure 5.21: Normalised measure of proportion of unfolded protein as a function of temperature for Cyt b_{562} WT and three different Cyt b_{562} – Ru hybrids. The melting temperatures, T_m, i.e. the temperature where 50 % of the protein is unfolded, are given for the four different protein species.

Upon coordination of the [Ru(HMB)] fragment to Cyt WT the protein species which elutes from the column first, Cyt WT – [16] – P2, has an increased melting temperature, 64.0 °C, compared to 59.1 °C for the WT. This indicates that for this adduct, ruthenium coordination is stabilising the helical fold of the protein. The protein species which elutes off the column second, Cyt WT – [16] – P3 has melting temperature of 57.8 °C which is similar to the WT protein. This suggests that in this species, the addition of the [Ru(HMB)] unit does not stabilise the fold of the protein. Similarly, the addition of a [Ru(HMB)(Bipy)] fragment to Cyt b_{562} L10C does not increase the stability of the protein fold, as this species has a melting temperature slightly lower than the WT protein of 57.7 °C.

The initial hypothesis was that Cyt WT – [16] – P2 and Cyt WT – [16] – P3 had different retention times, charge state distributions and melting temperatures because the [Ru(HMB)] fragment coordinated to different amino acid residues. To test this hypothesis and to determine the coordination sites of the [Ru(HMB)] fragment a tryptic digest procedure and mass spectrometry analysis was performed on Cyt WT – [16] – P2 and Cyt WT – [16] – P3. Unlike the [Ru(Cym)(Bipy)] fragment the [Ru(HMB)] fragment is likely to be coordinated to more than one amino acid residue in order to be coordinatively saturated. The [Ru(HMB)] fragment has three vacant coordination sites.

Upon direct infusion of the digested sample into a mass spectrometer, masses of peptides were detected between 300 – 2000 m/z. **Figure 5.22** focusses upon the 900 – 1200 m/z region, with selected masses and assignments shown in **Table 5.3**.

Table 5.3: Assignment of peptidic fragments modified and un-modified by [Ru(HMB)] fragments.

Mass	Charge	m/z	Assignment
1058.55	+1	1058.55	EAQAAAEQLK – Peptide 86-95
1995.81	+2	998.91	ADLEDNMETLNDNLK – Peptide 1-15 + [Ru(HMB)]
1901.89	+2	951.95	HGFDILVGQIDDALK – Peptide 63-77 + [Ru(HMB)]

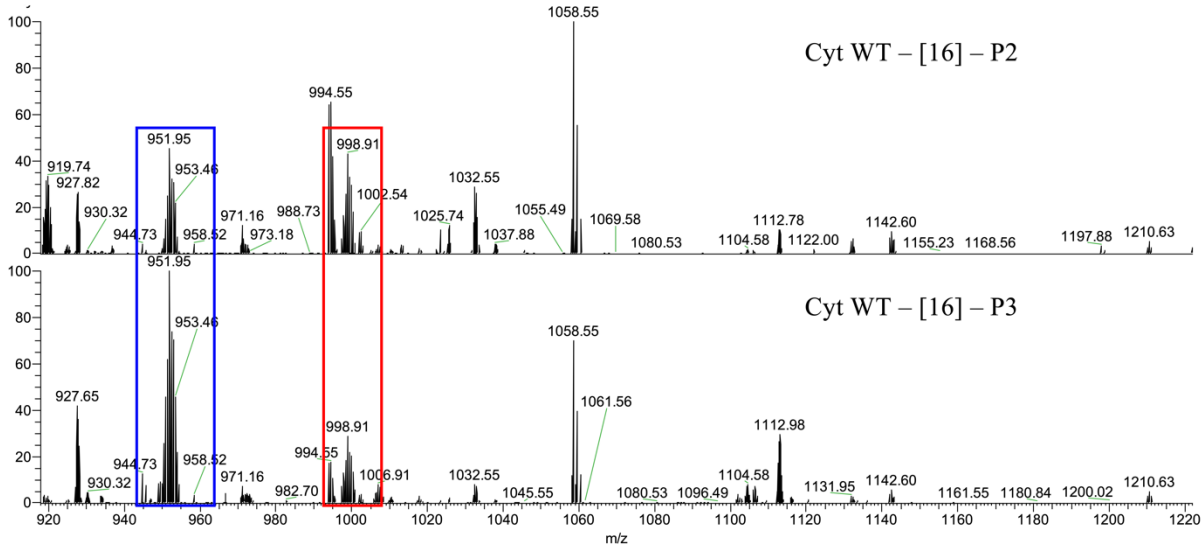


Figure 5.22: An ESI-MS spectrum of the adducts observed when the Cyt b_{562} – Ru species Cyt WT – [16] – P2 and Cyt WT – [16] – P3 have undergone tryptic digestion. Two ruthenium modified peptides are highlighted with the masses and assignments given in **Table 5.3**.

In these spectra two peptides containing the coordinated [Ru(HMB)] fragment can be identified. This was aided by the unique isotopic pattern of ruthenium, which can be seen in the highlighted sections of **Figure 5.22**. Both Cyt WT – [16] – P2 and Cyt WT – [16] – P3 had

similar mass spectra upon tryptic digest, therefore it was concluded that they were coordinating to the same peptides.

Tandem mass spectrometry experiments were performed on both ruthenium modified peptides (Peptide 1-15 + [Ru(HMB)] and Peptide 63-77 + [Ru(HMB)]) for each Cyt b_{562} WT – [Ru(HMB)] species (Cyt WT – [16] – P2 and Cyt WT – [16] – P3), with the spectra shown in **Figures 5.23 to 5.25.** A detailed explanation of how the tandem MS/MS data was analysed is given in **Section 3.2.5**.

MS/MS of Peptide 63-77 + [Ru(HMB)] – Coordination to Histidine 63

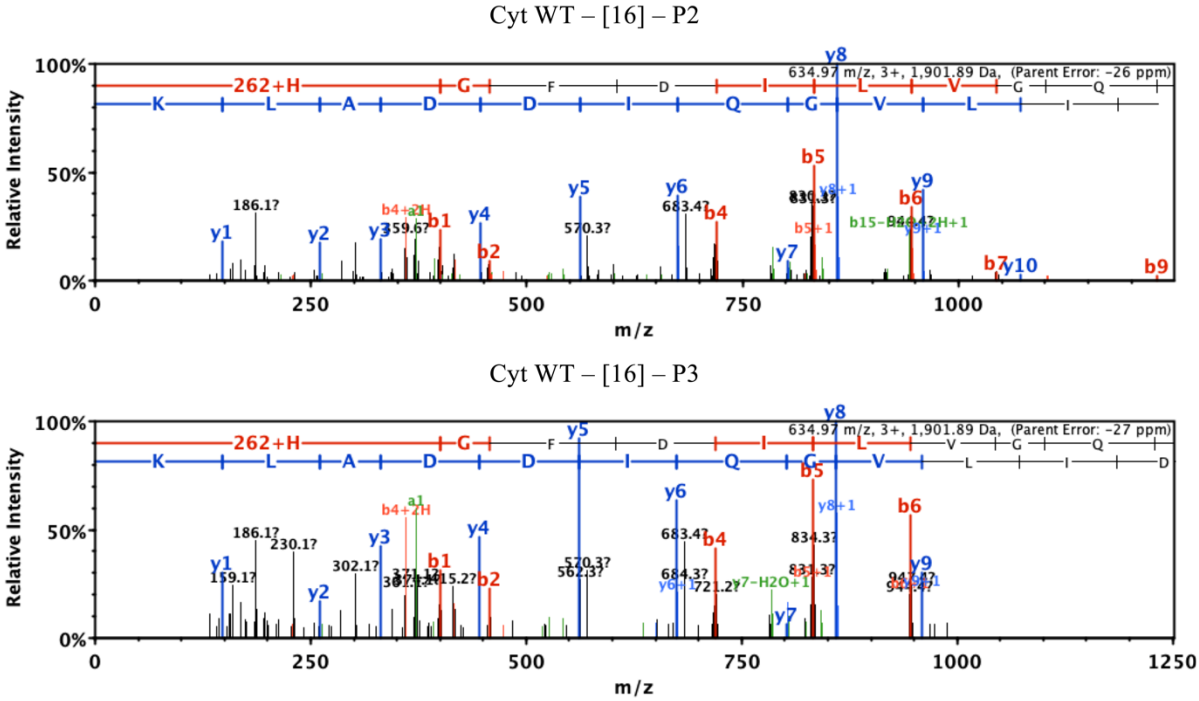


Figure 5.23: ESI-MS/MS spectra for CID of the ruthenium-modified peptide 63-77 with an actual mass of 1901.89 for two species Cyt WT – [16] – P2 and Cyt WT – [16] – P3. The presence of ions with a mass consistent to b1 + [Ru(HMB)] indicates histidine coordination.

MS/MS of Peptide 1-15 + [Ru(HMB)] – Coordination to Alanine 1

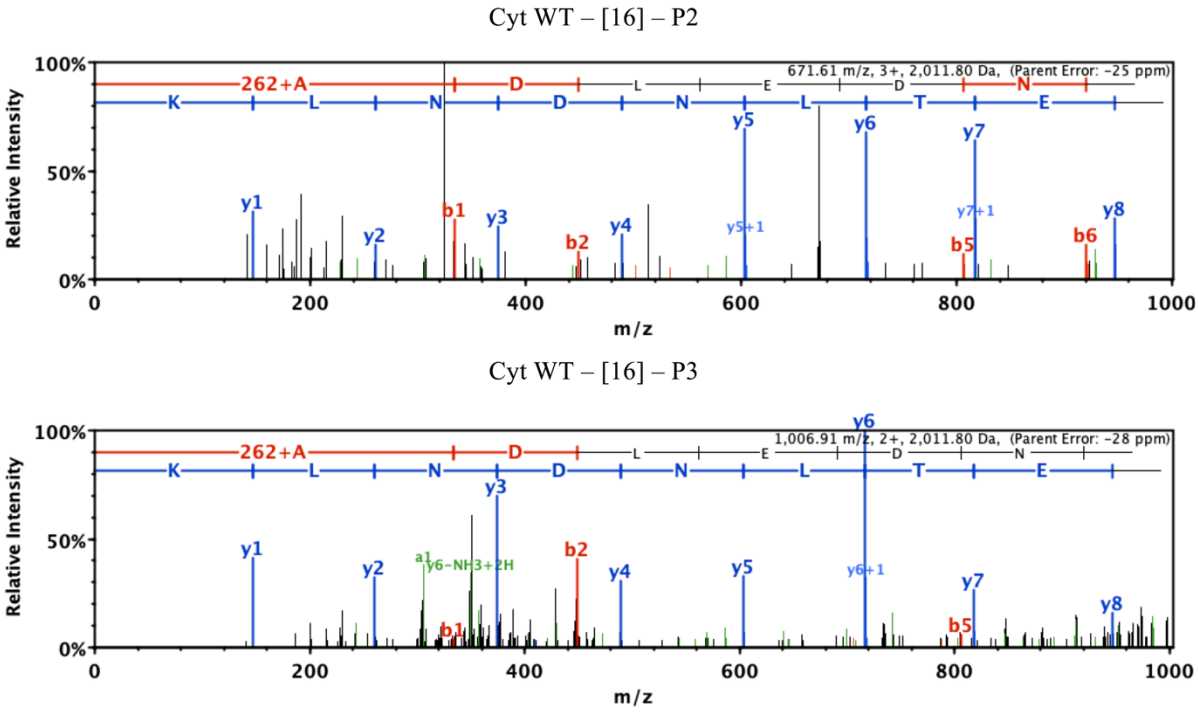


Figure 5.24: ESI-MS/MS spectra for CID of the ruthenium-modified peptide 1-15 with an actual mass of 2011.80 for two species Cyt WT – [16] – P2 and Cyt WT – [16] – P3. The presence of ions with a mass consistent to b1 + [Ru(HMB)] indicates alanine coordination.

MS/MS of Peptide 1-15 + [Ru(HMB)] – Coordination to Asparagine 6

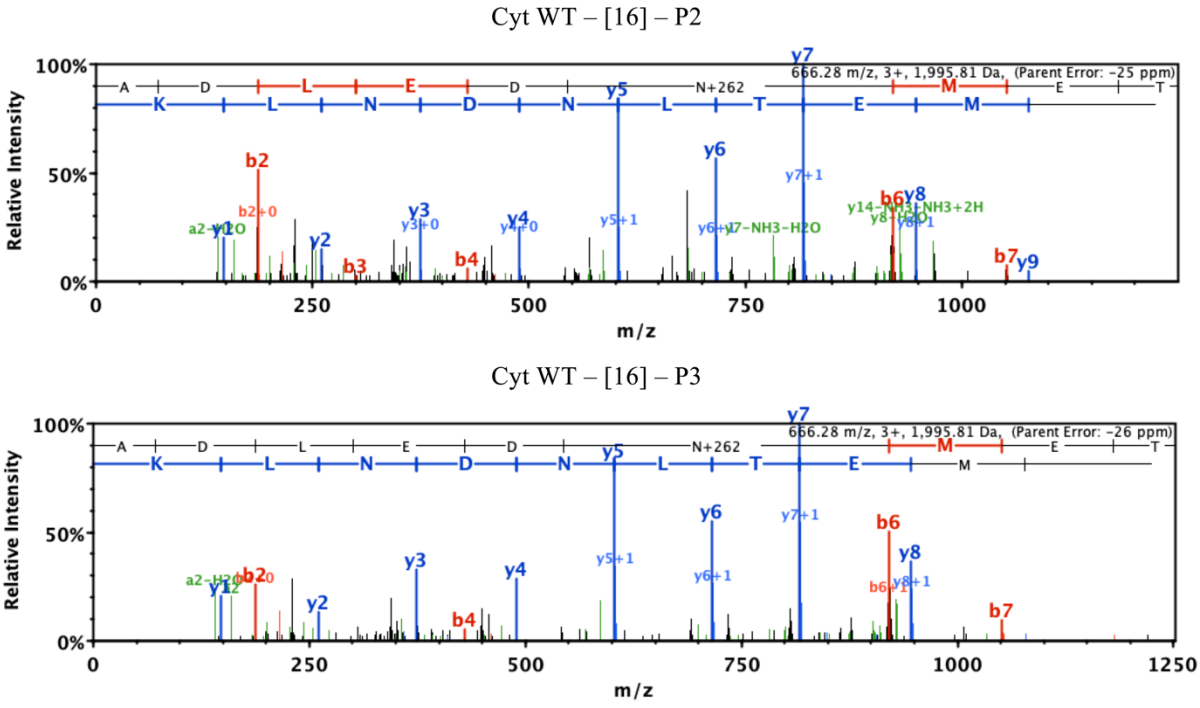


Figure 5.25: ESI-MS/MS spectra for CID of the ruthenium-modified peptide 1-15 with an actual mass of 1995.81 for two species Cyt WT – [16] – P2 and Cyt WT – [16] – P3. The presence of ions with a mass consistent to b6 + [Ru(HMB)] indicates asparagine coordination.

Discussion of Tandem Mass Spectrometry Results

This tandem MS/MS data suggests that when the [Ru(HMB)] fragment coordinates to Cyt b_{562} WT in both Cyt WT – [16] – P2 and Cyt WT – [16] – P3 the metal centre is coordinated to the following residues:

(i) Histidine 63 (on peptide 63-77)

(ii) Alanine 1 (on peptide 1-15)

(iii) Asparagine 6 (on peptide 1-15).

If all three of these residues were coordinated to the [Ru(HMB] fragment through side chain coordination then the ruthenium metal centre would be coordinatively saturated with the following ligands: A histidine imidazole nitrogen, the N-terminal amino group and either the amide oxygen or amide nitrogen from the asparagine, **Figure 5.26.** It is also possible that coordination occurs to the backbone carbonyls that make up the protein backbone. It is unlikely that the backbone nitrogen atoms will coordinate to the metal centre due to their relatively high pKas and delocalised lone pair.

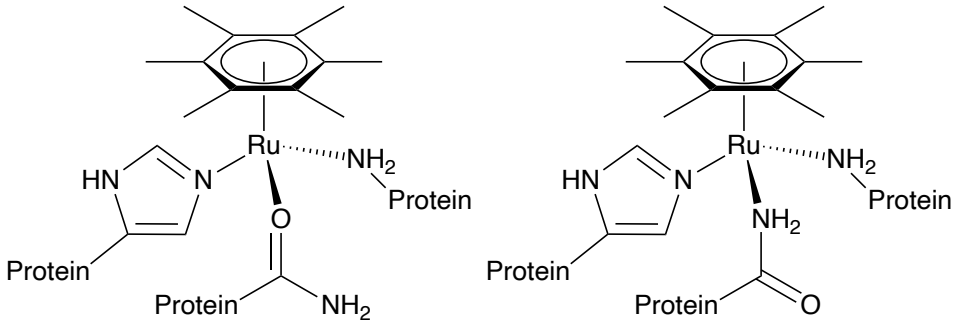


Figure 5.26: Two coordination environments possible when the [Ru(HMB)] fragment is coordinated to the Ala1, Asn6 and His63. Asn6 could either coordinate through the carbonyl or amino ligand.

Throughout the ^{19}F amino acid NMR experiments previously presented, it was found that the amide ligand (i.e. asparagine) and amino ligand (i.e. N-terminal alanine) were very weak ligands to ruthenium when free in solution. Additionally, these amide and amino ligands are reported to be weak ligands in natural systems, and incubations of the free amino acids (asparagine and alanine) with Ru(II) arene complexes were poorly yielding. The histidine imidazole, however, was found to be a satisfactory ligand for ruthenium coordination in solution. It is therefore feasible that when complex [16] is exposed to Cyt b_{562} WT that there is initial coordination to the histidine imidazole, followed by bipyridine dissociation, to give two

free coordination sites, which, due to the chelate effect, are then occupied by the weak ligands provided by Ala1 and Asn6.

These data disprove the initial hypothesis that the Cyt b_{562} WT – [Ru(HMB)] adducts with different retention times ((Cyt WT – [16] – P2 and Cyt WT – [16] – P3) are coordinated to different residues. Determining why these two species, which appear to be coordinated the exactly the same residues, have different retention times, charge state distributions and melting temperatures is challenging.

This could be due to a number of potential reasons, such as the [Ru(HMB)] fragments being in different conformations within the protein, i.e. the arene ligand being above or below the ruthenium in the protein. The protein being in a different conformation should also be considered, as these adducts have clear differences in protein stability as indicated by the temperature melt studies. Another potential reason, which was eluded to previously, is that the ruthenium is coordinated to different parts of the same residues, for example the amide nitrogen or the carbonyl backbone of Asn6. These subtle changes could potentially bring about large differences in the fold of the protein-metal hybrid. In order to further investigate this, structural based studies are underway within the research group involving X-ray crystallography and NMR spectroscopy.

5.2.7. Discussion on the Dynamics of Cytochrome b_{562}

Cyt b_{562} is a four helical bundle protein (α1-α2-loop-α3-α4) that has evolved to coordinate the haem cofactor non-covalently, with His102 and Met7 providing axial ligands to iron. There are three key structures of Cyt b_{562} reported: (i) a crystal structure of the holo-protein,[159] (PDB: 256B) (ii) an NMR solution structure of the holo-protein with the iron oxidised,[160] (PDB: 1QPU) and (iii) an NMR solution structure of the apoprotein (PDB: 1APC).[144] The NMR structure of the holo-protein is essentially identical to what is observed in the crystal structure, with the biggest differences coming between residues 50-57 which is a portion of the loop between α2 and α3. This region is highly disordered in the crystal structure indicating a number of alternative conformations in equilibrium.[160]

In order to understand the coordination of the [Ru(HMB)] fragment to Cyt b_{562}, an awareness of how haem effects the structure and folding kinetics of Cyt b_{562} is crucial. The haem cofactor and [Ru(HMB)] have different structures. How the protein coordinates the metal as well as how it forms non-covalent interactions with the metal ligands (porphyrin or arene) are likely to be very different; an appreciation of how this particular protein fold (the four helical bundle)

accommodates haem is useful. Many haem proteins fold independently of the haem cofactor to form an apoprotein, with the nascent potential for haem cofactor-binding whereas others require the cofactor to be present for productive and correct folding.

The folding kinetics of denatured apo- and holo-Cyt b_{562} were explored by Garcia et al. through a series of detailed solution NMR experiments. It was found that high concentrations of denaturant were required to fully denature the holo-protein to the denatured apoprotein and non-coordinated haem.[145] To reform the holo-protein, the apoprotein will slowly refold, then the protein fold will recognise the porphyrin ring through non-covalent interactions before the haem iron coordinates. Cyt b_{562} is therefore an example of haem protein which folds independently of the cofactor. The non-covalent interactions are vital in Cyt b_{562} coordinating haem, for example the orientation of haem entering the hydrophobic pocket is crucial in haem coordination. The non-covalent interactions observed between ruthenium complexes and Cyt b_{562} are therefore vital in the coordination environment described above.

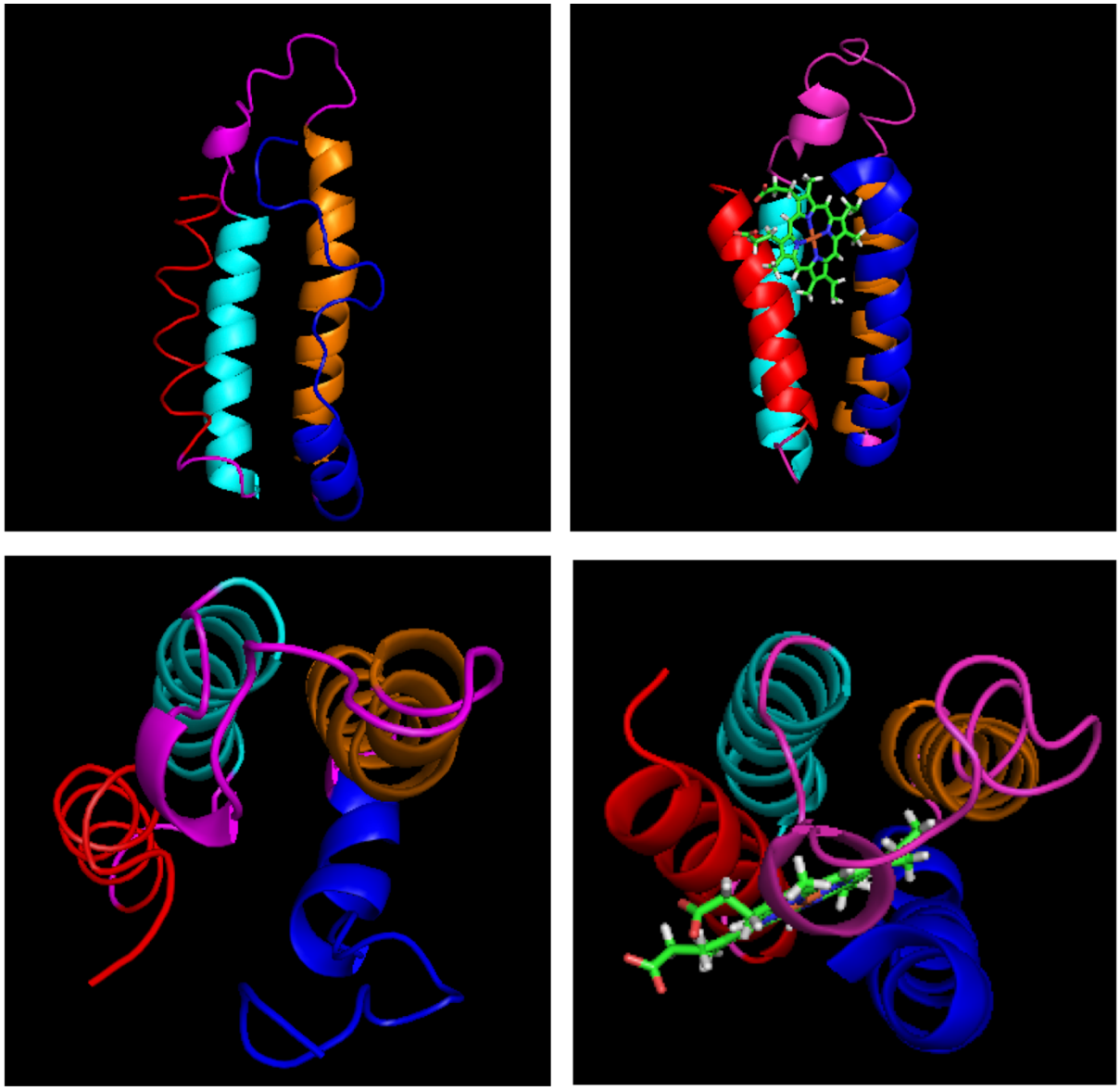


Figure 5.27: A PyMol generated cartoon of apo- and holo-Cyt b_{562} from a side on view (Top) and a view from the top (bottom). Each helix is colour coded, α1 – red, α2 – cyan, α3 – orange and α4 – blue, with the loops coloured magenta.

Upon haem coordination to the folded apoprotein the protein becomes more organised and tightly packed. Comparing the NMR structures of apo- and holo-Cyt b_{562}, **Figure 5.27**, α1 and α4 (beyond Gln93), become significantly more ordered. There is also significant flexibility in the loop between α2 and α3 with this region becoming much more ordered moving from the apoprotein to the oxidised holoprotein. The residues of the loop leading into α3 also become substantially more ordered, and a helical turn becomes apparent in this dynamic region.

Upon coordination of the [Ru(HMB)] fragment to Cyt b_{562} WT, α1 and α3 are crosslinked through coordination to Ala1, Asn6 and His63. **Figure 5.28** shows the apoprotein structure with these three residues highlighted.

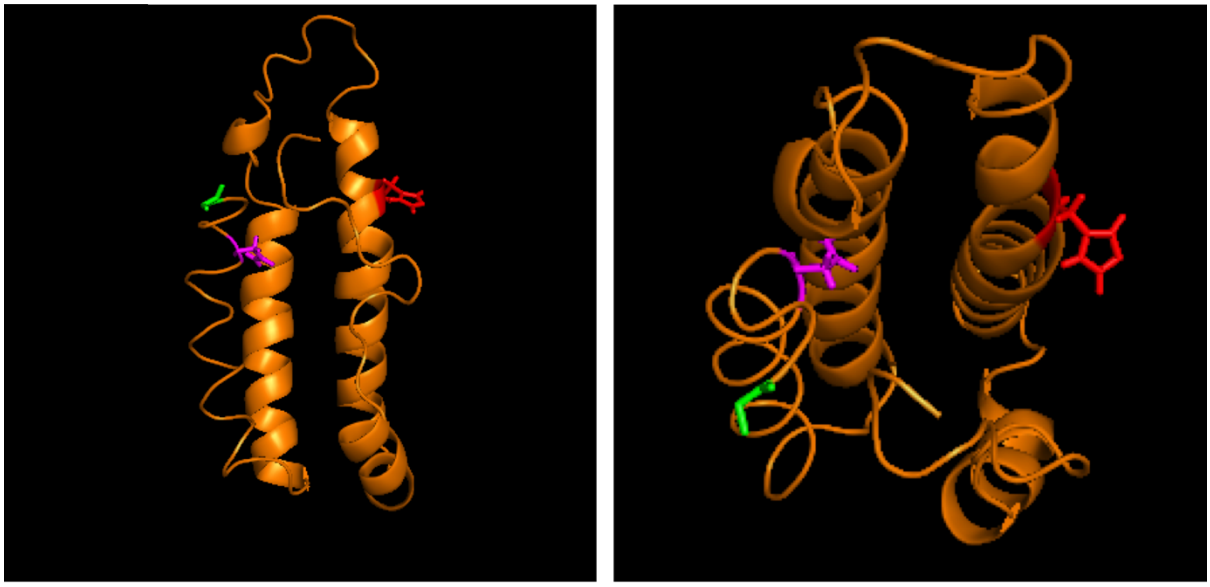


Figure 5.28: A PyMol generated cartoon of apo Cyt b_{562} from a side on view (Left) and a view from the top (Right). Residues for ruthenium coordination are highlighted: Ala1 – Green, Asn6 – Magenta and His63- Red.

Evidently from **Figure 5.28**, in order for the [Ru(HMB)] fragment to coordinate to all three of these residues, there must be a significant conformational change of the apoprotein to bring these three residues closer. The apoprotein structure indicates that His63 is orientated away from the central hydrophobic pocket, where haem binds. The flexibility in the loop between α2 and α3 could be essential for the conformational change required for a [Ru(HMB)] fragment to coordinate.

An Initial Comparison of Haem and [Ru(HMB)] Coordination to Cytochrome b_{562} WT

The coordination of the iron cofactor haem and the ruthenium fragment [Ru(HMB)] to apo - Cyt b_{562} differs greatly. Haem coordination to Met7 and His102 crosslinks α1 and α4, whereas ruthenium coordination to Ala1, Asn6 and His63 crosslinks α1 and α3. The 3d metal Fe and the 4d metal Ru have very different ligand exchange rates and electronic requirements. Furthermore, haem and [Ru(HMB)] coordination will have different steric implications and form very different non-covalent interactions with the protein; haem is a planar cofactor, whereas the Ru(II) arene cofactor adopts a pseudo-octahedral geometry capable of forming directional bonds to the protein scaffold, and this will determine how the apoprotein folds upon cofactor uptake. With His102 available for coordination, it is intriguing that the [Ru(HMB)] fragment coordinates to the His63 residue instead. One might predict that, like haem iron, the [Ru(HMB)] fragment would coordinate to His102 as this residue points towards the

hydrophobic pocket, but the clear preference for His63 emphasises the difficulty in predicting and designing metal coordination sites for exogenous metal complexes.

5.3. Conclusions

Understanding the speciation of Ru(II) arene complexes with proteins has formed a large part of this research. In particular, the reactivity of Ru(II) arene complexes with three proteins has been explored: (i) ubiquitin K63C; (ii) cytochrome b_{562} L10C and (iii) cytochrome b_{562} wild-type. When a cysteine is present and accessible, Ru(II) arene complexes will preferentially coordinate to the thiol. This has been confirmed by ^{19}F NMR, LC-MS, and experiments involving blocking the cysteine with the small molecule N-ethyl maleimide.

The coordination of the fluorinated Ru(II)(η^6-arene)(5,5'-FBipy) and Ru(II)(η^6-arene)(5,5'-TFMBipy) complexes to Cyt b_{562} L10C has been explored by ^{19}F NMR spectroscopy. The deep understanding of the reactivity of fluorinated ruthenium complexes with protected amino acids and of their ^{19}F NMR spectra, as outlined in Chapter 4, enabled this spectroscopic approach for probing the protein adducts formed. The ^{19}F NMR spectra are very sensitive and reveal, through the intricacies of protein-metal coordination, how the diastereotopic fluorine atoms on the bipyridine ligand have different peak geometries and peak widths, indicative of different relaxation properties. This demonstrates the differential effect that placing an organometallic complex within a protein scaffold can have on the electronic environment on the atoms which form the organometallic cofactor. These results confirm that the metal fragments have been incorporated into the structured environment of the protein. If the ruthenium fragment was coordinated on the surface of the protein, it is likely that the fluorine atoms would have the same relaxation properties, therefore identical relaxation properties and peak shapes.

This chapter has outlined a systematic study of the coordination behaviour of Ru(II)arene complexes with two variants of four helical bundle protein cytochrome b_{562} (Cyt b_{562} L10C and Cyt b_{562} WT) and revealed that minor changes to the small molecule complex led to drastically different conjugates being formed. Of particular interest, was the observation that even bidentate ligands can be exchanged for protein derived ligands. Considerations on the origin of these results allowed for carefully designed metal precursor complexes, containing labile ligands that maintain complex integrity in aqueous solution but upon contact with apocytochrome b_{562} variants undergo bipyridine exchange.

This was most clearly evidenced when [Ru(η^6-HMB)(5,5'-TFMBipy)Cl]$^+$ is incubated with Cyt b_{562} WT, resulting in a species where a [Ru(HMB)] fragment is coordinated to three

residues Ala1, Asn6 and His63 and crosslinks two of the four helices that form this protein (α1 and α3). Based on results presented in Chapter 4, we hypothesise that the complex undergoes an initial chloride exchange to an aquo ligand in aqueous buffer, followed by subsequent coordination to the imidazole of His63, which is a more thermodynamically favourable ligand than the potential ligands of Ala1 and Asn6. The complex then undergoes bipyridine dissociation with Ala1 and Asn6 providing the necessary donor ligands to satisfy the metal.

Bipyridine dissociation is also observed, but to a much lesser extent when Ru(II)(η^6-arene)(fluorinated bipyridine) complexes are incubated with the cysteine mutant Cyt b_{562} L10C. As previously discussed, if a thiol is present and accessible this will be the preferred binding site for a Ru(II) (η^6-arene)(bipyridine) as the bond enthalpy of forming a Ru–S bond is thermodynamically favourable. Based on the mass spectra presented throughout this chapter (particularly for the [Ru(η^6-HMB)(fluorinated bipyridine)Cl]$^+$ complexes [12] and [16]) we hypothesise that, bipyridine dissociation is much more favourable when initial coordination occurs at His63 instead of Cys10 (which is only present in Cyt b_{562} L10C).

One factor that could be a potential driving force for bipyridine dissociation and coordination to a potentially weaker ligand set is the protein fold energy which places the cofactor under steric pressure within the protein. Upon the predicted initial coordination, to His63 on α3, or Cys10 on α1, the cofactor is under different steric pressures within the protein, with it being more thermodynamically favourable to undergo bipyridine dissociation when the initial ligand is the histidine imidazole. Another potential that could favour bipyridine dissociation at the histidine is the electronic contributions from either the imidazole or the thiol. The strength of the Ru–bipyridine bonds will be greatly influenced by the other metal bound ligands and the different electronic contributions of imidazole vs thiol coordination could

To generate an artificial metalloenzyme based on coordination between a ruthenium cofactor and a protein scaffold, Cyt b_{562} has emerged as a potentially excellent candidate for a starting protein structure. The dynamic nature of the protein, the presence of a hydrophobic pocket and the distinct reactivity with different ruthenium organometallic fragments is promising. This dynamic four helical bundle protein is capable of coordinating a variety of different ruthenium fragments and has the conformational flexibility to provide ligands to saturate the ruthenium centre even when the dissociation of apparently strong ligands occurs. The Cyt b_{562} – Ru adducts presented within this chapter have the potential to be catalytically active.

For the majority of transition metal catalysed reactions, the metal centre must be able to undergo ligand exchange, often in order to coordinate catalytic substrates. The Cyt b_{562} – Ru adducts generated throughout this chapter are all coordinatively saturated, so to be catalytic they must undergo ligand exchange. The Cyt b_{562} WT – [Ru(HMB)] adducts coordinated to Ala1, Asn6 and His63 have two relatively labile ligands to ruthenium (from Ala1 and Asn6) which could potentially undergo ligand dissociation to coordinate a catalytic substrate and initiate a catalytic cycle.

Chapter 6

6. Cytochrome b_{562} – Ruthenium Hybrids for Catalysis

Preface: The work reported in this section was in collaboration with fellow PhD student Jamie Klein. I carried out all protein production, synthesis of ruthenium compounds, protein-ruthenium incubations (including purification and characterisation), whereas Jamie Klein carried out the organic synthesis of the assay substrates and prepared all catalytic experiments that were carried out on the plate reader.

6.1. Introduction

Chapter 1 outlined in detail the benefits and opportunities of developing ArMs with direct protein–metal coordination bonds. In particular, these systems have the advantage of the protein having a direct structural and electronic influence on the reactivity of the metal centre and, coupled with directed evolution, hybrid catalysts can be generated with significant improvements in catalytic efficiency.

This introduction will revisit four key considerations in developing dative ArMs for a given catalytic transformation, and introduce how these factors have identified cytochrome b_{562} – ruthenium adducts as potential starting points to search for catalysis. These four key considerations are: (i) the protein starting point, (ii) a controlled ligand exchange reaction activating a metal cofactor for catalysis, (iii) protein-metal bonds – realising an entatic state and (iv) the possible reactions that the chosen metal cofactor could catalyse.

6.1.1. The Protein Starting Point

A number of protein scaffolds have been used in the formation of ArMs and research groups in the field have used different parameters to select their protein scaffold. One of the key features of the ArMs reported is a hydrophobic pocket capable of accommodating an artificial cofactor and the catalytic substrate. This is most clearly evidenced in the work of Roelfes et al. and Lewis et al. who use the scaffolds LmrR and prolyl oligoptidase respectively; both these

proteins were chosen because they have a large central hydrophobic cavity where catalysis occurs.

Another key feature in the selection of the starting protein scaffold is how evolvable the protein is. In the development of an ArM the protein scaffold will undergo directed evolution in order to introduce mutations which could potentially bring about catalytic rate enhancements. Predicting whether or not a protein can withstand laboratory evolution based on inspection of the structure is challenging, if not impossible. In the pioneering work of Tawfik and co-workers, the concept of catalytic promiscuity is explored, i.e. whether an existing enzyme evolved for a given reaction and substrate can be evolved to also perform another function, with native activity retained.[161] From this study Tawfik highlights two key features from natural enzymes which can help predict the evolvability of the desired protein scaffold; robustness and protein dynamism (plasticity).[162] Robustness refers to the ability of a protein to withstand multiple mutations without significant loss of structure. Second, protein dynamism and flexibility are vital to enable the metal complex and substrate to be positioned for catalysis. Molecular dynamics and NMR studies can be used to assess the dynamism of a protein, and have shown that mutations in flexible loop regions can be crucial in rate enhancements.[163,164]

Instead of selecting a naturally existing protein scaffold many groups have used computational design to build up desired protein scaffolds. De novo design of metalloenzymes aims to use computational methods to predict what primary sequence and cofactors are necessary to achieve the optimal protein arrangement for metalloenzyme catalysis. Degrado and co-workers have pioneered the design of a number of synthetic proteins which directly coordinate bare metal atoms or metal cofactors.[165,166] The design of ArMs from scratch is particularly challenging due to the lack of reliable parameters for defining transition metal bonding and the immense complexity of the many low energy interactions that determine binding of small molecules to proteins. However, the designed protein structures that have shown particularly promising catalytic activity, with a designed metal binding site incorporated, are four helical bundle proteins.[108]

Assessing the optimal features for a starting protein scaffold, from both a natural metalloenzyme and a de novo design viewpoint, the natural four helical bundle protein apocytochrome b_{562} can be highlighted as an excellent potential starting point. First, without the haem present, there is a relatively large, nascent, hydrophobic pocket between the four helices. Second, as explored by the Barker group, the protein is robust and can withstand multiple mutations without the loss of structure. Third, as discussed in Chapter 5, the

apoprotein is dynamic, particularly in α1 and α4 and a flexible loop region between α2 and α3, evidenced through NMR structural studies.[160] The hypothesis is that this loop region and the nascent cofactor binding pocket are of significant importance to accommodating ruthenium complexes.

Apocytochrome b_{562} is smaller in comparison to other proteins used as scaffolds for artificial metalloenzymes, which could limit the formation of defined metal and substrate binding sites. However, the fewer residues could prove to be beneficial for an evolutionary campaign as a greater proportion of evolutionary space can be explored in comparison to larger proteins. The Barker group has also engineered apocytochrome b_{562} with additional helices at both the N and C termini which means the protein can me made bigger.

6.1.2. Controlled Ligand Exchange Activating a Metal Cofactor for Catalysis

As discussed in Chapter 1, a hugely desirable system for artificial metalloenzymes is to have both the protein and initial metal complex catalytically inactive and stable, but upon incorporation of the metal complex into the protein scaffold, activity is unmasked. This ensures that the only catalysis being measured is from the artificial metalloenzyme and not the free metal complex. We hypothesised that the optimum method to unmask catalytic activity from an initially inactive species, is through a controlled ligand exchange reaction with a protein scaffold. Chapter 5 has outlined a series of controlled ligand exchange reactions of Ru(II) arene complexes with variants of apocytochrome b_{562}. Of particular interest are the ligand exchange reactions which result in protein-metal hybrids with multiple protein-metal coordination bonds of varying strength. For catalysis to occur within the artificial metalloenzyme, further ligand exchange must also occur with a catalytic substrate. Having protein ligands which satisfy the metal in the absence of a catalytic substrate (i.e. ligands that can come on and off) is highly advantageous.

6.1.3. Protein-Metal Bonds – Realising an Entatic State

In small molecule transition metal catalysts, the ligands arrange around the metal centre to maximise bonding interactions. In the context of dative artificial metalloenzymes at least one of the ligands to the metal is provided by the protein scaffold. These peptidic ligands are not only influenced by the coordination bond(s) to the metal, but also by the intramolecular interactions that make up the protein fold. The immediate complex can therefore be distorted into a suboptimal geometry that is not a global energy minimum on the coordination energy landscape. The whole protein can be seen as a secondary coordination sphere where the protein

fold energy is used to poise the metal centre for reactivity. These energised, or entatic states have a reactivity that is not easily realised in conventional, synthetic metal catalysts.

From the perspective of the protein scaffold, the formation of an entatic state requires the peptide to be at least partially folded before binding the metal; the more defined the fold, the greater the ability of the fold to energise the metal complex. As such, fully folded rigid proteins, that impart a distorted, unfavoured geometry on a metal centre are optimal for adopting an entatic or energised state. This is in contrast to the desirable dynamic system for the evolutionary process. A potential compromise can be struck by using a starting scaffold that is partially folded as the apoprotein and upon cofactor binding rigidifies to a completely folded form, hence the attractiveness of apocytochrome b_{562} as a starting scaffold. The initial folding energy can be used to poise the metal in an activated state, while the folding process occurring during cofactor binding allows for the system to adapt during directed evolution. Once the ArM becomes more specialised after rounds of evolution, the apoprotein will probably approach a more fully folded form, yielding an ArM after cofactor addition that is less promiscuous but contains a more energised and active metal centre.

6.1.4. The Potential Reactions that the Chosen Metal Cofactor Can Catalyse

Throughout this research there has been a focus on the reactivity and speciation of Ru(II) arene bipyridine complexes with proteins. Ru(II) arene complexes have shown a wide application in a number of catalytic transformations, including transfer hydrogenation and C–H bond activation. Their broad reactivity scope and compatibility in aqueous biological systems has made them an attractive starting point to explore catalytic activity with ArMs.

Fluorescence-based assays are excellent methods for fast online monitoring of catalytic activity in directed evolution experiments. These assays are applicable for use in high-throughput screening, including micro-fluidics assays, due to the high sensitivity of fluorescence spectroscopy. In comparison to chromatographic screening, fluorogenic screens allow for an increased number of experiments, with high temporal resolution, to be performed.

In this research, transfer hydrogenation and olefin metathesis have been highlighted as potential reactions to monitor in aqueous systems with ruthenium-protein hybrid molecules. These reactions are well studied, catalysed by a variety of different ruthenium catalysts and have shown compatibility within aqueous systems, although water is not necessarily the optimal solvent. Furthermore, as this chapter will illustrate, a number of different fluorescence-based assays have been developed to monitor these catalytic reactions.

In the generation of ArMs with direct protein-metal coordination, the ligand exchange process has the potential to change the electronics of the metal centre. The cytochrome b_{562} – ruthenium adducts in this research are unlikely to be a specialist catalyst for a particular substrate or reaction, however, if initial activity is observed, they offer an excellent initial starting point for evolutionary campaigns.

6.2. Experimental Goals

Throughout this chapter, the aim was to generate a suite of cytochrome b_{562} – ruthenium hybrids and test them for transfer hydrogenation and metathesis activity using fluorescence-based assays. This required the development of a method to accurately quantify the concentration of the cytochrome b_{562} – ruthenium hybrids and the incorporation of fluorescence-based assays for transfer hydrogenation and metathesis.

6.3. Results and Discussion

6.3.1. Quantifying Cytochrome b_{562} – Ruthenium Adducts

Chapter 5 described how different protein-ruthenium hybrids can be purified as individual species using ion exchange chromatography. Having these hybrids as individual species, for example a species with a single [Ru(HMB)] addition, rather than a mixture of species, is crucial in identifying catalytically active species. For catalytic testing, 8 different cytochrome b_{562} – ruthenium adducts have been obtained as individual species from purifying large scale incubations of different ruthenium complexes with either the Cyt b_{562} L10C mutant or Cyt b_{562} WT, **Table 6.1.**

Table 6.1: Cytochrome b_{562} – ruthenium adducts - the major products from a variety of protein-ruthenium incubations purified *via* anion exchange chromatography.

Incubation	Species Isolated	M_{obs}	M_{calc}	Hybrid Number
Cyt b_{562} L10C H102M				
Cyt b_{562} L10C + 20 Eq. complex [3]	Cyt b_{562} L10C + [Ru(Cym)(Bipy)]	12154	12156	1
Cyt b_{562} L10C + 20 Eq. complex [11]	Cyt b_{562} L10C + [Ru(Cym)(5,5'-FBipy)]	12190	12191	2
Cyt b_{562} L10C + 2 Eq. complex [15]	Unable to Isolate Individual Species			
Cyt b_{562} L10C + 20 Eq. complex [4]	Cyt b_{562} L10C + [Ru(HMB)(Bipy)]	12182	12184	3
Cyt b_{562} L10C + 20 Eq. complex [12]	Cyt b_{562} L10C + [Ru(HMB)(5,5'-FBipy)]	12217	12220	4
Cyt b_{562} L10C + 2 Eq. complex [16]	Unable to Isolate Individual Species			
Cyt b_{562} WT				
Cyt b_{562} WT + 20 Eq. of complex [12]	Cyt b_{562} WT + [Ru(HMB)]	12043	12043	5
	Cyt b_{562} WT + [Ru(HMB)]	12043	12043	6
Cyt b_{562} WT + 2 Eq. of complex [16]	Cyt b_{562} WT + [Ru(HMB)]	12042	12043	7
	Cyt b_{562} WT + [Ru(HMB)]	12042	12043	8

Hybrids 1 – 4 arise from incubations of the desired ruthenium complex with the cysteine mutant Cyt b_{562} L10C. From a combination of ^{19}F NMR spectroscopy and competition experiments with N-ethyl maleimide the cysteine residue is confirmed to be coordinating to ruthenium. In all four of these species the bipyridine ligand has remained coordinated to the metal centre.

Hybrids 5 – 8 are species which correspond to the modification of Cyt b_{562} WT with a [Ru(HMB)] fragment. Hybrids 7 and 8 have been discussed in detail in Chapter 5 and result from an incubation of Cyt b_{562} WT + 2 Eq. of complex [16]. Upon purification, two hybrids with the same mass are isolated, Cyt WT – [16] – P2 = hybrid 7 and Cyt WT – [16] – P3 = hybrid 8. Hybrids 5 and 6 appear to be analogous to the species purified from an incubation of

Cyt b_{562} WT + 2 Eq. of complex [16], however the protein was incubated with 20 equivalents of complex [12] instead of 2 equivalents of complex [16].

In order to standardise catalytic experiments, it is essential to know the concentration of the protein-ruthenium adduct. Typically, measuring approximate protein concentration is straight forward through the Beer-Lambert law which correlates UV absorbance at 280 nm, with protein concentration, through the extinction coefficient ε_{280}. However, in all samples the ruthenium fragment coordinated to the protein contributes significantly to the absorbance at 280 nm, therefore a direct concentration read out *via* UV spectroscopy is not possible. Furthermore, depending on the exogenous ligands, ligands provided by the protein and the non-covalent interactions of the metal complex, each ruthenium fragment has a different extinction coefficient, which is challenging to measure in this context.

Quantifying ArM concentration with an unnatural metal cofactor has previously been performed through a variety of methods. Lewis et al. standardised the concentration of all rhodium-prolyl oligopeptidase ArMs generated by multiplication of the ESI-MS peak intensities of the ArM and scaffold in these mixtures.[85] Ward et al. measured the concentration of all biotin-streptavidin variants assuming complete 1:1 conversion and therefore assumed the protein concentration before cofactor attachment was equivalent to final ArM concentration. These methods outlined above, do not give accurate and consistent quantification of ArM concentration, therefore alternative methods were explored.

Throughout this project, cytochrome b_{562} – ruthenium adducts have been quantified through ICP-OES (inductively coupled plasma optical emission spectroscopy). This method takes a liquid sample of known dilution (in HNO_3) and runs it through a plasma which vaporises the atoms and ions so that each element's characteristic wavelengths are emitted. The intensity of the different wavelengths corresponds to the concentration of each individual element.

ICP-OES was used to measure the concentration of ruthenium in a protein-ruthenium hybrid and by knowing the number of ruthenium modifications on the protein (through LC-MS) the concentration of the protein can be determined.

6.3.2. A Transfer Hydrogenation Assay

Self-immolative chemistry involves molecules that link, through covalent bonds, 'trigger' moieties with functional substrates of interest, such as drugs or fluorophores. This chemistry has found applications within the fields of high throughput screening, prodrugs, biosensing and biomolecular imaging.[167] Upon activation of the trigger moiety by a specific analyte or

chemical modification, there is a cascade reaction that causes the release of the fluorophore, **Figure 6.1**.

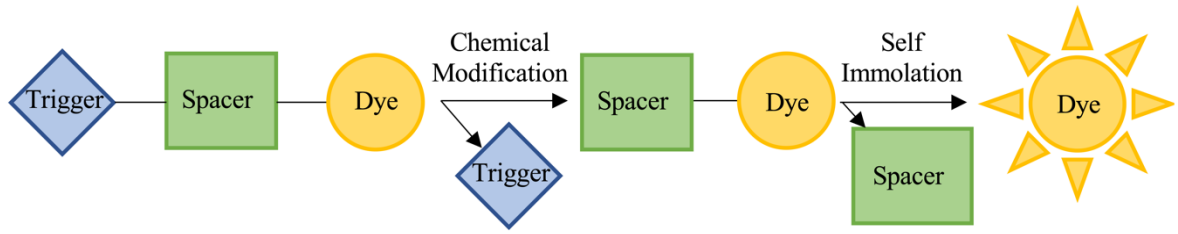
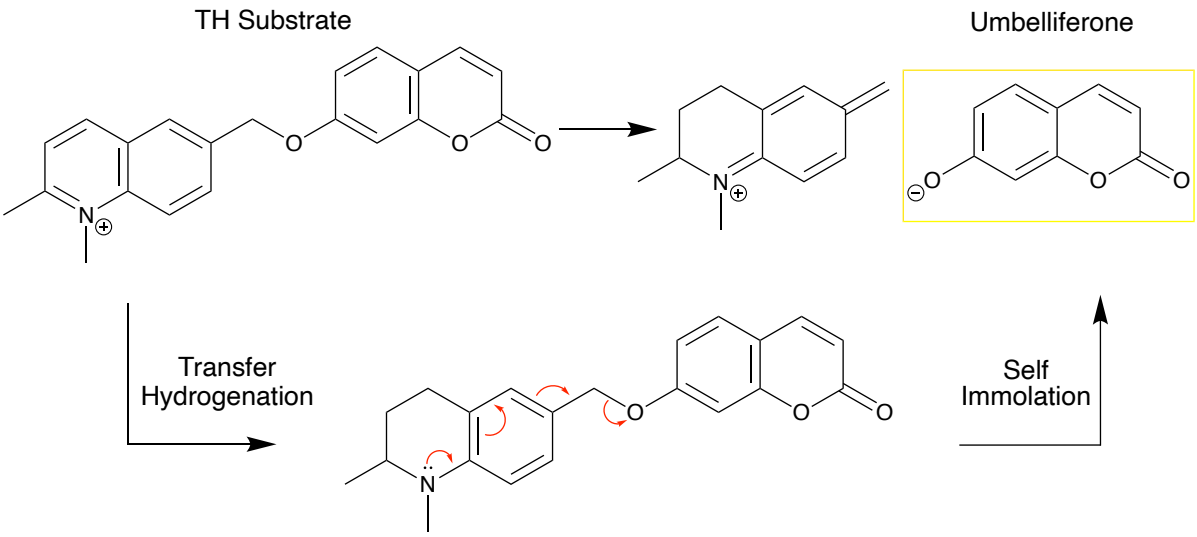


Figure 6.1: Schematic of self-immolative molecules that release fluorescent dyes.

Recently, there have been a raft of studies which look to design novel, self-immolative molecules which are activated by different chemical modifications. In a particular study, a quinolinium substrate has been designed which, upon C=N reduction, may undergo self-immolation to release a fluorogenic umbelliferone derivative, **Scheme 6.1**.[89]



Scheme 6.1: Reduction of the TH substrate which undergoes self-immolation to release a fluorogenic umbelliferone molecule which is highlighted in yellow.

Asymmetric transfer hydrogenation of imines[168] and quinolines[169] has been reported using a chiral Ru(II)(η^6-arene) catalyst. The general mechanism for ruthenium catalysed transfer hydrogenation of imines involves initial coordination of formate to the metal centre, followed by β-hydride elimination to form a ruthenium hydride species. This hydride is transferred onto the C=N, carbon electrophilic centre, and the N picks up a H^+ from the aqueous solvent resulting in reduction of the imine to an amine.[170] These ruthenium catalysts have chiral ligands coordinated to the metal centre in order to control the enantioselectivity of the reaction, however, using an ArM means that the secondary coordination sphere provided by the protein should dictate the stereochemistry of product.

Synthesis of the TH substrate was performed using a modified preparation and reported in **Section 3.3.4**. All catalytic experiments in this section were performed in 96-well plates with transfer hydrogenation activity measured by fluorescence output using a SpectraMax i3x plate reader, as described in **Section 3.2.9**.

The fluorescent properties of umbelliferone have been extensively studied.[171] Umbelliferone exhibits an intense blue fluorescence band at 460 nm over a wide range of pH and excitation wavelengths. As shown by **Figure 6.2**, the emission at 460 nm is greatest with an λ_{ex} = 325 nm at pH 7.0, and with λ_{ex} = 370 nm at pH 8.0. This can be attributed to the deprotonation of the phenolic proton at the 7-position moving to slightly basic conditions. The anionic form of umbelliferone absorbs more effectively at an λ_{ex} = 370 nm.

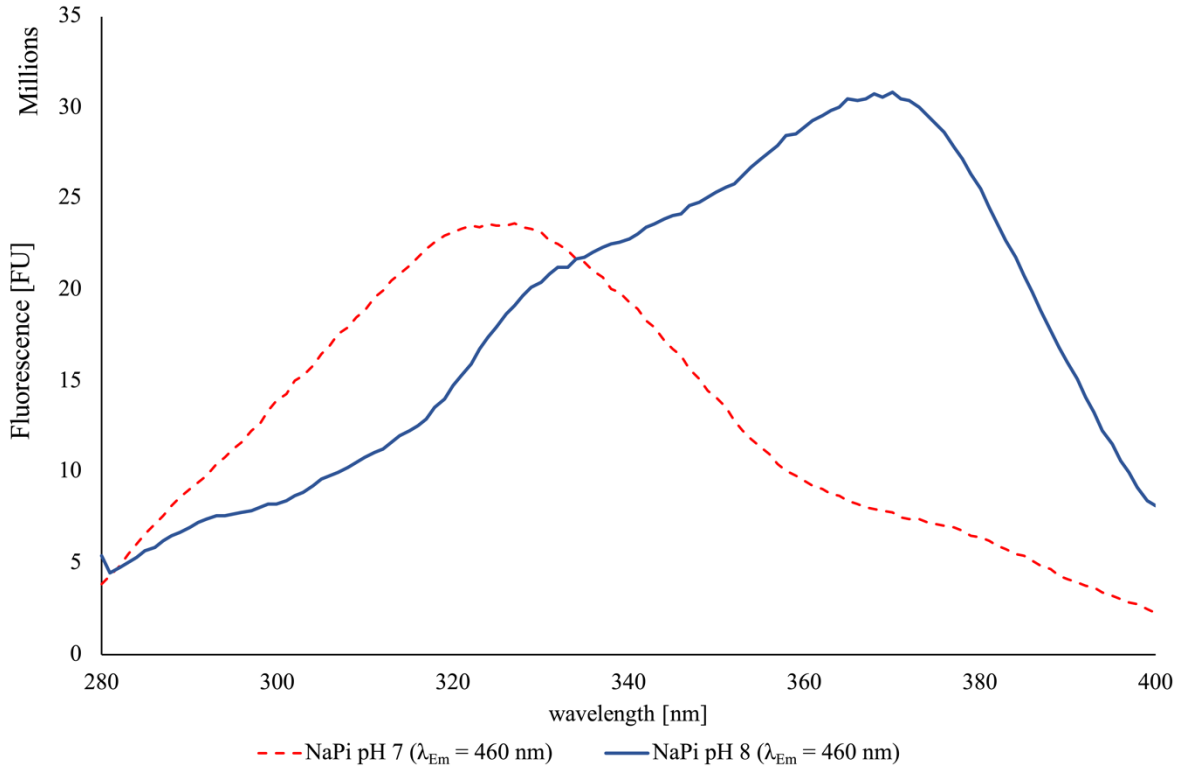


Figure 6.2: Excitation spectra of umbelliferone in different pH sodium phosphate (NaP_i) buffer between 280 – 400 nm with a fixed λ_{em} of 460 nm.

In the transfer hydrogenation reaction mixture, the coupled quinolinium – umbelliferone substrate is in excess concentration, at 1 mM. As shown by **Figure 6.3**, this substrate absorbs very strongly between 285 nm to 360 nm, which is attributed to strong quinolinium absorbance. This strong absorbance filters the vast majority of the light between these wavelengths meaning that the excitation spectrum of umbelliferone changes in the presence of the substrate.

Importantly, at pH 8.0 the optimum λ_{ex} = 370 nm, where the substrate does not absorb, allowing for quantitative measurement of the fluorescence intensity.

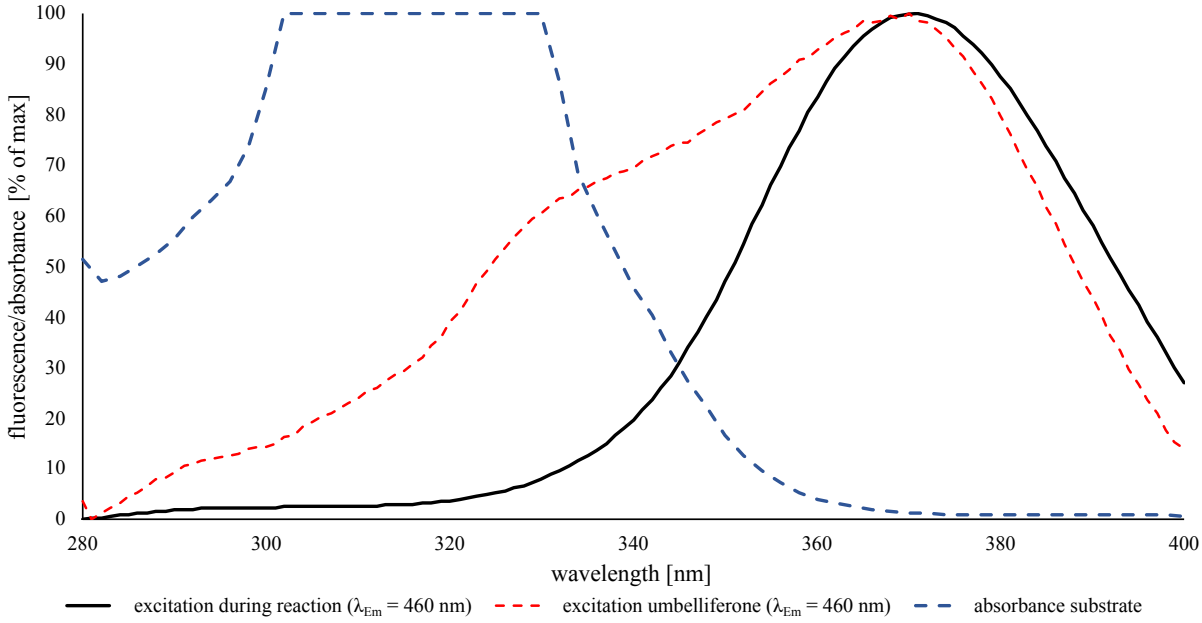


Figure 6.3: A combined graph of the excitation spectra of umbelliferone in 50 mM pH 8.0 NaP_i buffer with and without the presence of the starting transfer hydrogenation substrate and the absorbance spectra of the transfer hydrogenation substrate at 1 mM measured using the same instrument conditions. As evidenced by the flat line between 300 nm and 330 nm the absorbance measured is the maximum the machine under those conditions can measure.

Calibration curves for umbelliferone at pH 7.0 and at pH 8.0 in 50 mM sodium phosphate buffer have been measured, **Figure 6.4**. 100 mM formate has been added to the pH 8.0 calibration curve to more accurately represent the conditions under which transfer hydrogenation will occur.

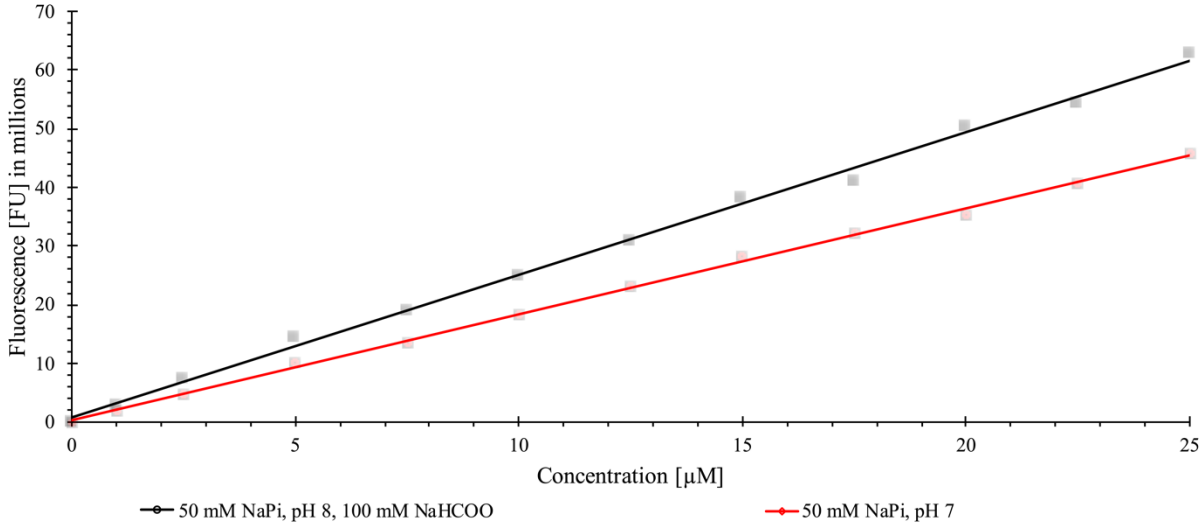


Figure 6.4: Umbelliferone calibration curves between 1 and 25 µM in two different buffer conditions.

6.3.3. Testing for Transfer Hydrogenation Activity

The Cyt b_{562} – ruthenium hybrids 1-8, the bipyridine complexes [3], [4], [11], [12], [15], [16] and dimer III were tested for transfer hydrogenation activity by monitoring the fluorescence intensity over a 16 hr period with an λ_{ex} = 370 nm and an λ_{em} = 460 nm. The rate of transfer hydrogenation was measured by taking the steady state rate and normalising to the ruthenium concentration of the sample, **Figure 6.5**. Hybrids 1-4 showed no transfer hydrogenation activity so are removed from **Figure 6.5**.

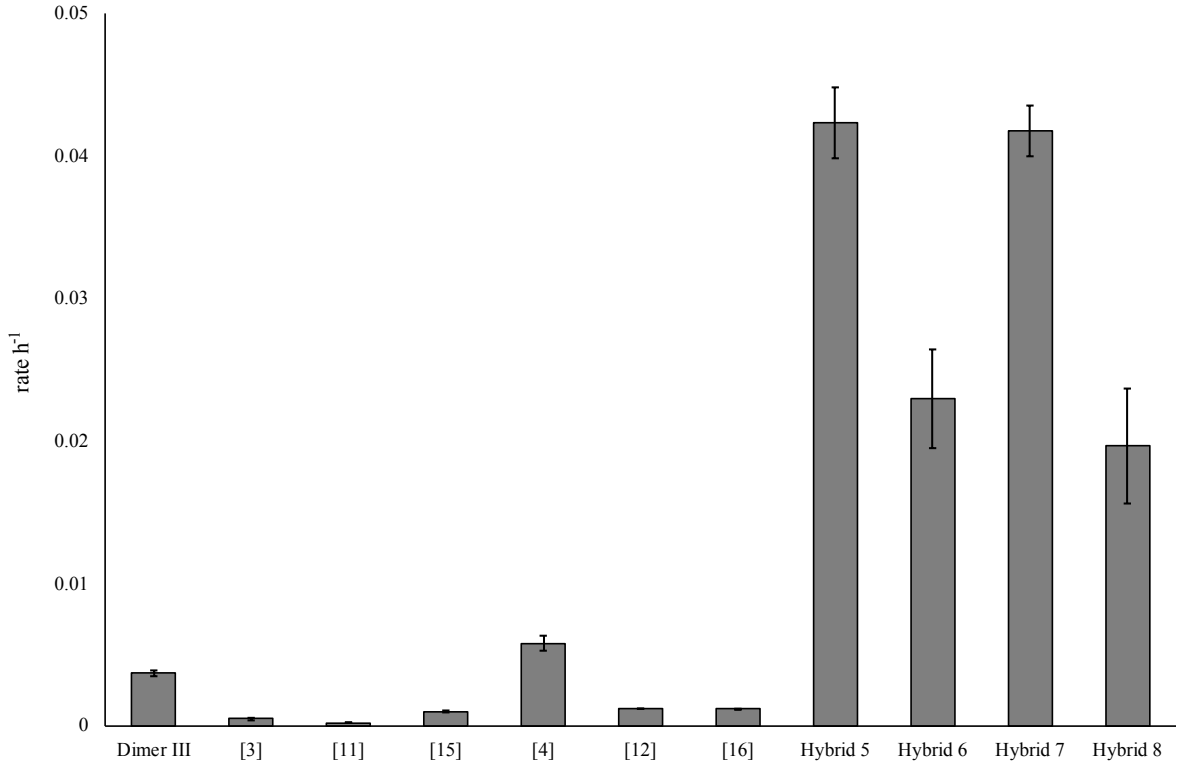


Figure 6.5: A graph to show the rate of transfer hydrogenation for the Cyt b_{562} – Ru hybrids 5-8, the Ru-cymene complexes, [3], [11] and [15], Ru-hexamethylbenzene complexes, [4], [12] and [16], the dimeric complex III. Concentrations of the protein-metal hybrids ranged from 20 μM to 70 μM, with a concentration of 100 μM used for all ruthenium complexes. Error bars are one standard error of the mean. For the ruthenium complexes these errors are calculated from 3 independent measurements. Errors for hybrids 5 and 6 are calculated from 5 independent repeats and errors for hybrids 7 and 8 are calculated from 4 independent repeats.

A key finding from these data is that a Cyt b_{562} – Ru hybrid has been generated from a ligand exchange reaction between a Ru(arene)(bipyridine) complex and Cyt b_{562} WT that has an activity that is 35 times faster than the starting ruthenium complex. Additionally, hybrids 5-8 all show increased catalytic activity than the known transfer hydrogenation catalyst dimer III, $[Ru(Cym)Cl_2]_2$. In particular hybrids 5 and 7 are more than 5 times faster than dimer III. These data highlight how catalytic activity can be unmasked through a controlled ligand exchange

reaction with an appropriate protein scaffold. This is the desirable reactivity, discussed in detail in **Section 6.1.2.** The key to this activity is taking a stable complex and activating it into a metastable catalytically active species with weak ligands held by the protein fold.

These data clearly display the increased activity of the Ru(arene) unit with direct coordination to a protein scaffold, compared to free dimeric complex which is likely to undergo hydrolysis within aqueous solution. The increased activity of these protein-ruthenium hybrids could be due to the propensity for side reactions and catalyst decomposition being lowered once the Ru(arene) unit is in the hydrophobic protein pocket of Cyt b_{562}. Furthermore, the hydrophobic pocket in the protein could not only provide a localisation site for the Ru(arene) unit, but also be a suitable binding site for the organic, hydrophobic substrate. Bringing together the catalytic substrate and metal centre into close proximity within the protein scaffold will enhance catalysis. Important to consider is that Cyt b_{562} WT is not evolved to bind either the Ru(arene) unit nor the large hydrophobic substrate, highlighting the potential for a directed evolution campaign in order to optimise catalysis. For transfer hydrogenation to occur *via* the well-studied catalytic mechanism, the Ru centre must initially undergo a ligand exchange reaction with formate to form a Ru-hydride species. In all of the active hybrids, we hypothesise that the metal centre has weak coordination bonds to Asn6 and Ala. Therefore, these ligands have a propensity to undergo ligand exchange with formate in order to form the hydride species, thus initiating the catalytic cycle.

Hybrid pairs 5/7 and 6/8 have the same catalytic activity (within error), while the activity is significantly different between the pairs of hybrids 5/6 and 7/8. A number of pieces of evidence support the hypothesis that hybrids 5/7 are structurally the same adduct (Cyt WT + [Ru(HMB)]), as are hybrids 6/8. Hybrids 5/7 and hybrids 6/8 have the same retention times and charge state distributions. Additionally, measuring the same activity for these hybrids further supports this hypothesis. The differing activity between hybrids 5/6 and hybrids 7/8 further supports the hypothesis that these species are structurally different, with a potentially different fold or conformation around the metal centre. This is potentially exciting as its gives more than one starting point for an evolutionary campaign from one protein-metal complex combination. Additionally, different starting metal complexes can be used to give the same hybrid species upon only slight variations of reaction conditions.

The thermal stability experiments using CD were previously particularly enlightening to demonstrate that hybrids 5 and 6 were different species. Additional CD temperature melt experiments are consistent with hybrids 5/7 and hybrids 6/8 being the same species, **Figure**

6.6. Hybrids 5 and 7 had similar melting temperatures of 63.4 and 64.0 °C respectively, and hybrids 6 and 8 also had similar melting temperatures of 59.4 and 57.8 °C.

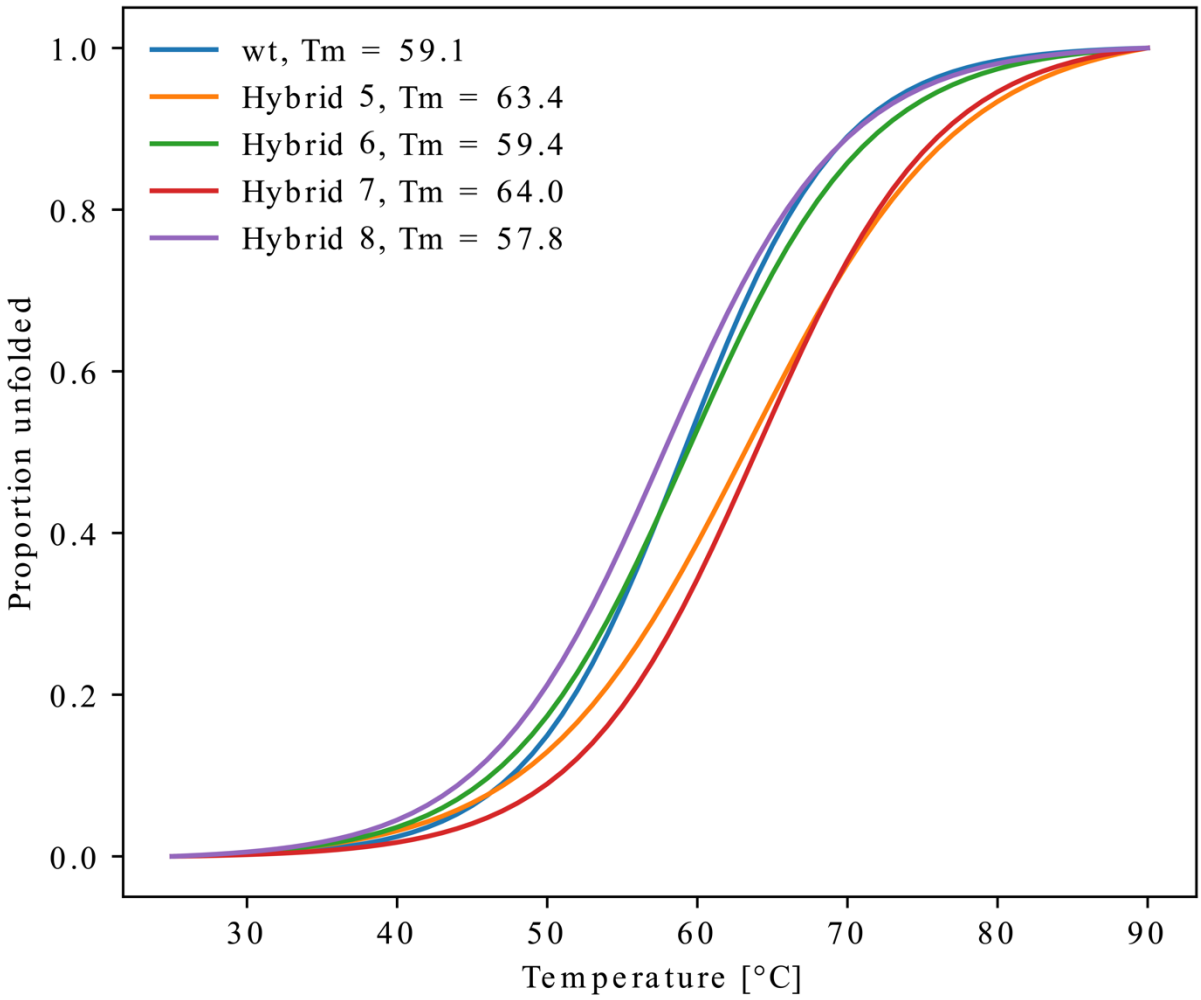


Figure 6.6: Normalised measure of proportion of unfolded protein as a function of temperature for Cyt b_{562} WT and hybrids 5-8. The melting temperatures, T_m, i.e. the temperature where 50% of the protein is unfolded, are given for the five different protein species.

6.3.4. Direct Observation of Active Catalyst Formation

The catalytic protein-metal hybrids presented so far have all undergone a purification process after incubation with the metal complexes in order to isolate individual species and remove excess ruthenium complex. As these hybrids display much increased activity than the free ruthenium complexes, the rate increase upon hybrid formation should be detectable by monitoring the fluorescence as the hybrid forms in the presence of the transfer hydrogenation substrate. Referring back to the speciation reactions, 2 Eq. of complex [16] were incubated for 2 hrs with Cyt b_{562} WT before purification. Here, in subsequent reactions, 0.5, 1 and 2 Eq. of complex [16] were incubated with 100 μM of Cyt b_{562} WT in the presence of 1 mM substrate and the fluorescence was monitored over time, **Figure 6.7**.

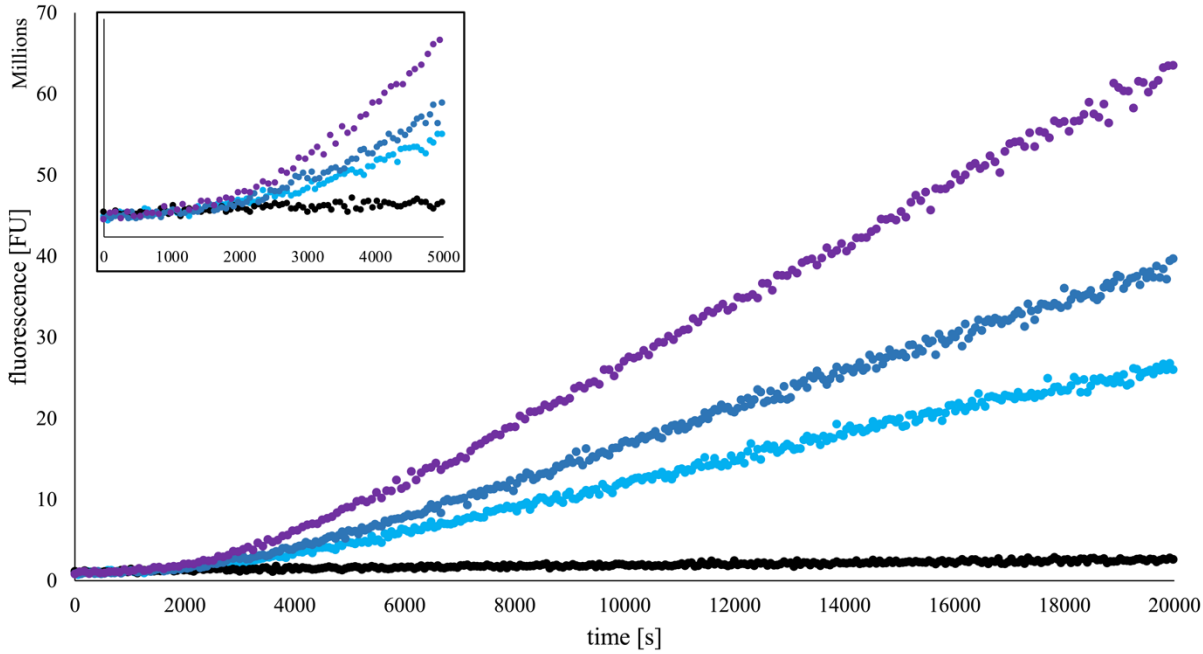


Figure 6.7: A graph showing the evolution of fluorescence activity upon hybrid formation in comparison to free complex [16] (black trace). The transfer hydrogenation substrate is combined with Cyt b_{562} WT and complex [16] at the start of the experiment. Varying equivalents of complex [6] are added 0.5 Eq. (light blue trace), 1 Eq. (dark blue trace) and 2 Eq. (purple trace). The insert is zoomed in to the first part of the experiment to demonstrate the lag period which occurs whilst the hybrid is forming.

As shown by **Figure 6.7**, there is an initial lag in the rate of transfer hydrogenation. In this lag period, the rate of reaction matches the rate of the free ruthenium complex. As the Cyt b_{562} WT – [Ru(HMB)] hybrids begin to form, there is an increase in the rate of transfer hydrogenation. Assuming full conversion with 1 Eq. and 50 % conversion with 0.5 Eq. to the species Cyt b_{562} WT – [Ru(HMB)] (with a mass of 12042 Da), the normalised rates of transfer hydrogenation within the linear section of the graph corroborate with the rates observed for the purified hybrids 7-8 shown in **Figure 6.5**. When 2 Eq. of complex [16] is used, the observed rate is faster than the purified hybrids, which is hypothesised to be due to the presence of an even more active double addition product. This double addition product, Cyt b_{562} WT + 2 [Ru(HMB)] (12402 Da) confirmed by LC-MS, has activity arising from two different locations of [Ru(HMB)], with the second location likely to be a weakly-bound metal fragment.

6.3.5. Discussion of Transfer Hydrogenation Activity

Catalysis which starts with an inert complex which is catalytically activated upon protein-metal coordination is a particularly attractive approach in the development of ArMs. This concept is most clearly exemplified by the vitamin B12 cobalt cofactor, which upon direct coordination

to the protein methionine synthase, methyl transfer activity is unmasked with great control and substrate specificity.[172] Most literature examples of ArMs involving unnatural metal complexes to date, start with catalytically proficient complexes that upon incorporation into a protein scaffold show increased catalytic efficiency (k_{cat}/K_M) or enhanced stereoselectivity.

This increased catalytic efficiency is often as a result of the proximity of the metal and substrate within a hydrophobic protein pocket. In many literature examples, directed evolution has been used to evolve this hydrophobic pocket which increases substrate binding affinity which, under simplifying assumptions, is characterised by a decrease in K_M. In these systems, the chemical turnover rate (in many systems characterised by the initial k_{cat}) does not increase significantly from the rate observed from free metal complex. This is in contrast to the data presented above, where there is a large rate enhancement between the starting metal complex and the protein-metal hybrid generated upon coordination.

In hybrids 5/7 and 6/8 there are protein-metal coordination bonds from the [Ru(arene)] fragment to Cyt b_{562} WT, that have a direct influence on the transfer hydrogenation turnover rate, k_{cat} at the metal centre. The differential activity between samples is strong evidence that the protein environment is influencing catalytic activity in different ways. Directed evolution has the potential to test a large number of variants of the protein matrix which will each have differing electronic contributions to the metal centre, that could be beneficial for catalysis.

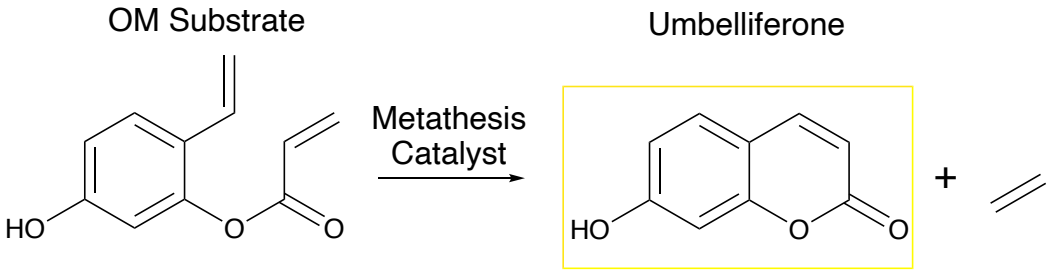
6.3.6. Measuring the Turnover of the Reaction

The rate enhancement in transfer hydrogenation of hybrids 5/7 over the free complexes [12] and [16] highlights an excellent evolutionary starting point. Knowing the concentration of the hybrids in solution, as well as an accurate umbelliferone calibration curve allows for determination of the turnover number of the reaction. After 7 days at 37 °C, a TON of 1.32 was measured for hybrid 7, however throughout the reaction, the rate was diminishing back towards the background rate of complex [16]. It was confirmed by LC-MS that approximately over 70% of the [Ru(HMB)] had dissociated from the hybrid, thus explaining the drop off in rate of transfer hydrogenation. Although a relatively modest TON was measured at this concentration, it is important to observe that the enzyme is acting catalytically and to reiterate that this protein is not evolved for binding of the [Ru(HMB)] or the catalytic substrate.

6.3.7. An Olefin Metathesis Assay

In **Section 1.4.2**, the mechanism of olefin metathesis (OM) is discussed, and it is highlighted how catalysis with small molecule ruthenium catalysts has been optimised through ligand

development, in particular the introduction of N-heterocyclic carbene ligands. Performing metathesis in polar protic solvents like water is particularly challenging as the catalysts are often vulnerable to poisoning with coordinating solvent ligands.[173] No metathesis reactivity has been observed in natural systems, however Ward et al. have demonstrated that metathesis can be introduced *in-vivo* using a biotinylated 2rd generation Grubbs-Hoveyda catalyst which assembles with a streptavidin scaffold in the periplasm of *E.coli* cells.[87] One could envisage the usefulness in synthetic biology of having an *in-vivo* artificial metathase complementing a metabolic or synthetic pathway. Ward et al. have developed a number of pro-fluorescent substrates that yield a fluorescent product upon ring closing metathesis (RCM),[174] **Scheme 6.3** shows the pro-fluorescent OM substrate used in this study.



Scheme 6.3: The pro-fluorescent OM substrate which upon RCM releases the fluorescent derivative umbelliferone.

Synthesis of an olefin metathesis (OM) substrate was performed using a modified synthesis as reported in **Section 3.3.4**, and catalytic activity was initially tested using the known Grubbs 2nd generation catalyst.

Historically, in order to minimise presence of the nucleophilic hydroxide ion and degeneration of Grubbs 2nd generation catalyst, aqueous metathesis reactions are performed in buffer at low pH. To test the viability of the reaction, fluorescence was monitored over 18 hours with the OM substrate and Grubbs 2nd generation catalyst in 50 mM sodium acetate buffer, pH 4.0. An λ_{ex} = 325 nm and λ_{em} = 460 nm was used and, unlike the TH substrate, the OM substrate does not strongly absorb at 325 nm allowing for quantitative monitoring at that wavelength. The strong activity of Grubbs 2nd generation catalyst at pH 4.0 highlighted the potential to monitor this reaction in aqueous conditions, however to test the protein-metal hybrids, a pH of reaction closer to biological pH was preferable.

6.3.8. Metathesis Activity of Cytochrome b_{562} – Ruthenium Adducts

The metathesis activity of Grubbs 2nd generation catalyst, complexes [12] and [16] and hybrids 7 and 8 was tested in two different buffer conditions: sodium acetate buffer pH 4.0 and sodium

phosphate buffer pH 7.0. The activity of Grubbs catalyst was majorly reduced moving to more neutral pH, however a small signal was observed above the blank. Preliminary data shows that hybrids 7 and 8, also showed low levels of metathesis activity comparable to Grubbs catalyst at pH 7.0, but showed no activity at pH 4.0, **Figure 6.8**.

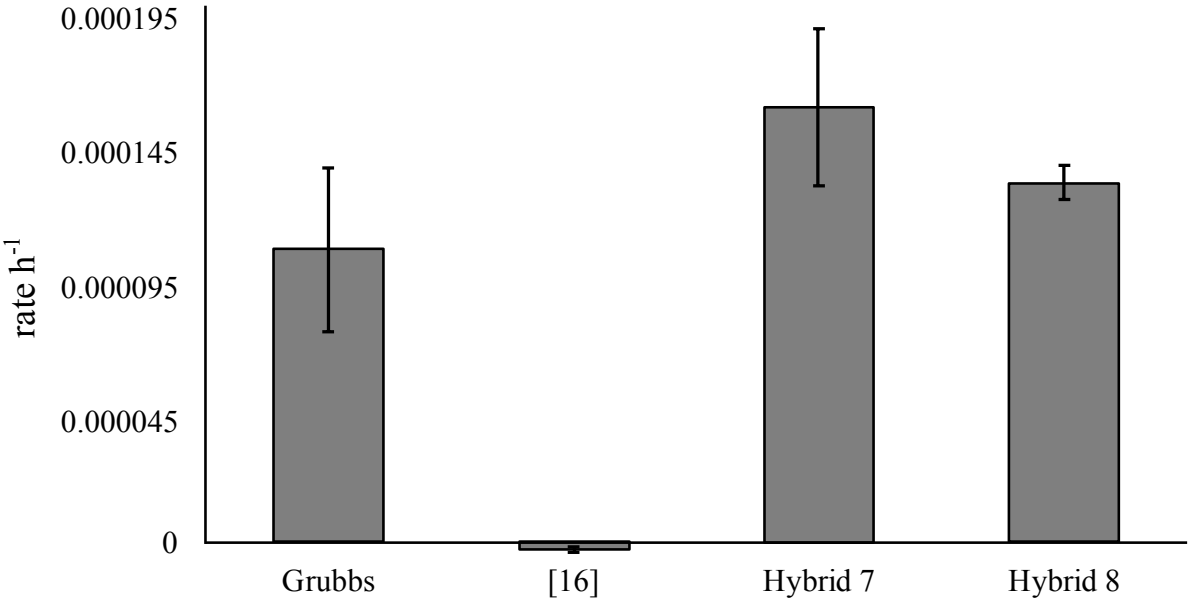


Figure 6.8: A graph to show the rate of olefin metathesis for the Cyt b_{562} – Ru hybrids 7 and 8, the Ru-cymene complex, [16] and Grubbs 2nd generation catalyst. The rate has been measured and normalised by the ruthenium concentration. Error bars are one standard error of the mean. Errors are calculated from 3 independent measurements.

That any form of metathesis activity is observed from these Cyt b_{562} – Ru hybrid is promising. Following the classical Chauvin mechanism for metathesis, there is a [2+2] cycloaddition of an alkene double bond to a Ru-alkylidene species to form a metallacyclobutane, which is a Ru(IV) species. Ruthenium complexes are often stabilised in the +II oxidation state by bonding modes to the arene. The metathesis activity observed from the [Ru(HMB)]-protein adduct emphasises the potential catalysis that can occur when you can localise a ruthenium cofactor and substrate within a protein hydrophobic pocket. Additionally, the protein provides a scaffold that has the ability to rearrange and stabilise the ruthenium centre at a higher oxidation state, enabling metathesis to occur.

6.4. Conclusions

A series of cytochrome b_{562} – ruthenium(II) arene adducts were purified *via* anion exchange chromatography. These species varied in which ligands coordinate to the metal, either those provided by the initial complex or by the protein scaffold. Importantly an ICP-OES method

was developed to quantify the concentration of the adducts generated, so that any potential catalytic experiments could be normalised to protein-ruthenium hybrid concentration.

Transfer hydrogenation activity and metathesis activity of all Cyt b_{562} – Ru(II) arene adducts was explored using fluorescence-based assays. Transfer hydrogenation activity was observed in hybrids where [Ru(HMB)] fragments have coordinated to Cyt b_{562} WT. In these active hybrids transfer hydrogenation activity has been unmasked from the relatively inert Ru(II) arene complexes [12] and [16] upon protein coordination. Through a combination of ion-exchange chromatography, LC-MS and CD, it was confirmed that the same adducts, hybrids 5 and 7 and hybrids 6 and 8 could be generated from incubations of Cyt b_{562} WT with different Ru(HMB) complexes. In particular, hybrids 5/7 were particularly active transfer hydrogenation catalysts in comparison to the known dimeric catalyst III and the Ru(II) arene bipyridine complexes. Hybrids 5/7 were over 5 times more active than the dimer and up to 35 times more active than the Ru(II) arene complexes.

The enhancement in the rate of transfer hydrogenation demonstrated in this chapter has shown the ability to activate relatively inert metal complexes towards catalysis through cooperative ligand loss and coordination to a protein scaffold in a controlled manner. However, the hybrids generated still have limited activity as highlighted by the relatively modest TON. This is not necessarily too surprising as the protein is not evolved to bind either a [Ru(HMB)] unit or the catalytic substrate. The hybrids generated are, however, a fantastic starting point to begin an evolutionary campaign to evolve hybrids into very efficient catalysts. The two key reasons for this are the evolvability of the protein scaffold and the direct coordination between the metal and the scaffold.

The optimum protein scaffold must be able to accommodate metal coordination and substrate association and have potential evolvability (robustness and plasticity) to enhance catalysis at the metal centre. The four helical bundle protein, apocytochrome b_{562} has been identified as an excellent starting point due to its hydrophobic cavity, dynamic nature and ability to withstand mutations.

In more detail, mutations could be made to the hydrophobic pocket which accommodates the metal complex and substrate that could place the metal cofactor into an entatic, sub-optimal state, poised for catalysis, and improve the binding and pre-organisation of the hydrophobic, fluorogenic substrate. Further, of particular interest could be the effect on activity of mutations to dynamic regions of the protein, for example the loop between α2 and α3 and, separately,

α4. Predicting changes and designing mutations to bring about enhanced activity is however challenging, if not impossible, highlighting the importance of being able to screen a large library of mutants of apocytochrome b_{562}.

Direct protein-metal coordination enables directed evolutionary campaigns to attenuate and significantly vary the electronic environment experienced by the metal. In the active example identified above, there are three protein-metal coordination environments (Ala1, Asp6 and His63) including two comparatively weak interactions, enabling substrate and co-catalyst coordination. Varying the electronic contribution of these three ligands to the ruthenium centre through genetic mutation, could enhance the turnover number (k_{cat}) associated with that distinct metal coordination environment.

A concept that is central to the development of ArMs with direct protein-unnatural metal coordination is catalytic promiscuity. One way to define catalytic promiscuity is whether the same initial protein-metal hybrids can be used as evolutionary starting points to achieve different catalytic transformations. Observing preliminary results of both metathesis and transfer hydrogenation activity from the same protein-metal hybrids highlights this exciting concept.

Chapter 7

7. Developing a Ruthenium Fragment Library for Catalytic Screening

7.1. Introduction

Building upon the understanding of the aqueous and biological speciation of Ru(II) arene complexes with bipyridine ligands, this chapter outlines the synthesis and biological reactivity of novel Ru(II) arene complexes with alternative bidentate ligands to bipyridine.

Many transition metal catalysed transformations involve the metal accessing a variety of oxidation states, acting as either single electron donors/acceptors or as two electron donors/acceptors as seen in oxidative addition and reductive elimination chemistry. For example, in ruthenium catalysed olefin metathesis the metal centre shuttles between Ru(II) and Ru(IV). Ligands that can both donate electron density and withdraw electron density from the metal centre (referred to as push-pull ligands) are ubiquitous in organometallic catalysis, by virtue of their ability to stabilise both high and low metal oxidation states.

Ligands which can act as strong electron donating ligands are capable of stabilising high oxidation states. A common example is phosphanes (PR_3), which can modulate activity depending on the substituent R groups. A further example is N-heterocyclic carbene (NHC) ligands which will be discussed in detail in the following section. The majority of phosphanes and NHCs are not stable as free species in aqueous solution, due to the high Lewis basicity of the lone pair on the phosphorous or the carbon, however they could have great potential for catalysis when coordinated to a metal within a biological setting. In order to realise this potential, and explore their effect on catalysis in biological systems, the Lewis basic lone pair must be involved in bonding to a metal centre before the metal species is exposed to an aqueous biological environment.

Many metalloenzymes are capable of performing multi-electron reactions under physiological conditions. In natural metalloenzymes, Fe, Cu and Mo are the most frequently employed redox active metal centres. These metal centres have specific biological ligands (the primary coordination sphere) and are optimally positioned within an active site to tune both reductive

and oxidative chemistry.[175] This is clearly evidenced by differing redox potentials of the $Fe^{III}Fe^{III/II}$ [2Fe2S] clusters found in ferrodoxins (-460 to -300 mV versus the standard hydrogen electrode (SHE)) and Rieske proteins (-100 to 490 mV versus SHE). In ferrodoxins the [2Fe2S] cluster is anchored through four Cys residues, whereas in Rieske proteins the cluster has two Cys ligands and two His ligands.[4] These direct changes in the primary coordination sphere, as well as changes in the protein fold and secondary coordination sphere, tune the redox properties of metal cofactors to ensure enzymatic function.

In many literature studies, the redox properties of Ru(II) arene complexes have been investigated and, under physiological conditions, the +II oxidation state is stabilised by bonding interactions of the arene ligand, making oxidation to higher oxidation states challenging.[176] This made it somewhat surprising that olefin metathesis was measured (shown in Chapter 6) for a Ru(II) arene fragment coordinated to cytochrome b_{562}, as oxidation to a Ru(IV) species is required. As discussed in Chapter 6, it was suggested that the localisation of the ruthenium metal and the catalytic substrate within the protein scaffold enabled catalysis to occur, with the protein capable of stabilising the ruthenium centre in the +IV oxidation state. An exciting question we asked was, if the arene ligand was replaced (in the protein-metal hybrid) with a ligand more capable of stabilising a metal in a higher oxidation state, would an increase in catalytic activity be observed?

Therefore, in order to expand the catalytic potential of the fragments within the library under study, this chapter will explore the reactivity and biological speciation (particularly with Cyt b_{562}) of a number of novel ruthenium complexes with abiological, synthetic ligands capable of stabilising higher oxidation states.

7.1.1. N-Heterocyclic Carbene Complexes in Biology

A carbene is a molecule containing a neutral carbon atom with a valence of two and two unshared electrons. Metal carbene complexes are often classified into either Fischer-type (singlet) carbenes where the carbene carbon atom is electrophilic and usually possesses adjacent π-donor groups stabilising the empty p-orbital or Schrock-type (triplet) carbene complexes which are nucleophilic and carry alkyl groups or hydrogen atoms on the carbene carbon. A third class of carbene complexes are N-heterocyclic carbenes (NHCs), which are defined as having a carbene carbon and at least one nitrogen atom within the ring structure. NHCs exhibit a singlet ground-state electronic configuration where the HOMO is the sp^2-hybridised orbital with the carbon lone pair, and the LUMO is the unoccupied p-orbital on the

carbon. The neighbouring nitrogen atoms (which are σ-electron withdrawing and π-electron donating) stabilise this singlet configuration through inductively lowering the energy of the occupied σ-orbital and mesomerically by donating electron density into the empty p-orbital, **Figure 7.1**.[177]

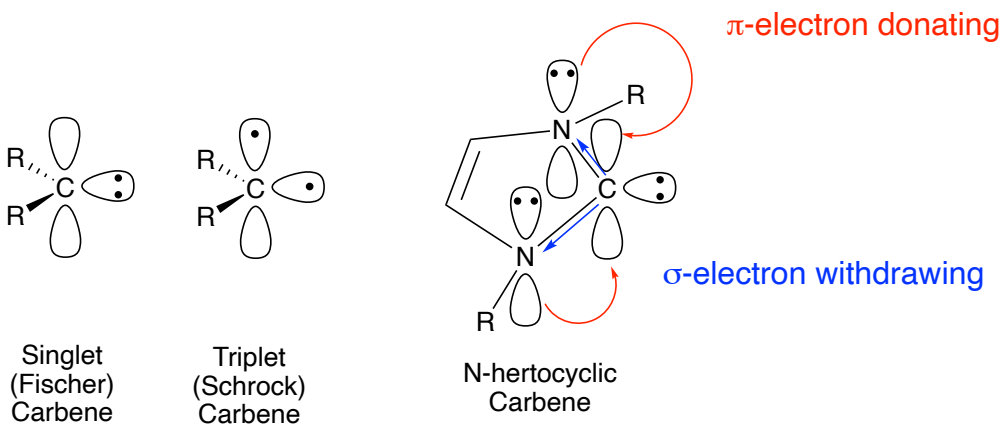


Figure 7.1: Simplified structures of three different carbene classes, with the electronic stabilisation shown on the structure of the N-heterocyclic carbene.

The strong σ-donation of NHCs is the most important component of the bonding to transition metals, however, the contribution of π -back-bonding into the carbene p-orbital and π -donation from the carbene p-orbital can amount for up to 20 % of the overall bond energy in some metal-NHC complexes.[177] However, in comparison to Fischer and Schrock carbenes which are depicted with double bonds to the metal, NHCs coordinated to metal centres are drawn with a single bond which best reflects the experimentally observed potential for rotation around the M–C bond. In 1991, Arduengo and co-workers reported NHCs with extraordinary stability (they referred to them as 'bottle-able'), ease of synthesis and applicability.[178] NHC ligands have attractive features for use in small molecule catalysis; they can stabilise high oxidation states through increased electron donation and the electronics can be altered by changing the nature of the azole ring. Referencing Grubbs 2nd generation catalyst, this metal complex has two types of carbene ligands present; first, the NHC ligand plays an important role as a strong σ–donating ligand helping stabilise the metal when in the +IV oxidation state and encouraging phosphine dissociation to allow for alkene association, second the Ru(C=C(H)(Ph)) carbene ligand initiates the metathesis catalytic cycle.

The use of NHC ligands in biocatalysis is a virtually unexplored area, and modification of natural metalloenzymes using an NHC ligand has very minimal literature precedent. Within the field of ArMs, NHC ligands have been introduced in metal complexes which also have ligands that form supramolecular interactions with the host protein. Most notable is the work

of Ward et al. and Tanaka et al. who have assembled ArMs based on the supramolecular assembly of Grubbs 2nd generation catalyst within their chosen protein scaffold.[87,91]

There is, however, a body of research which has explored the antimicrobial and anticancer activity of metal-carbene complexes. Gold NHC complexes in particular have shown meaningful application in the medicinal anticancer field, with the inhibition of thioredoxin (TrxR) being a possible mechanism for their activity.[179,180] Recent research by Hartinger et al. has explored the speciation and biological activity of a range of ruthenium-carbene complexes, including [Ru(Cym)(dmb)Cl_2] (dmb = 1,3- dimethylbenzimidazol-2-ylidene), **Figure 7.2**, which inhibits TxrR at μM concentrations.

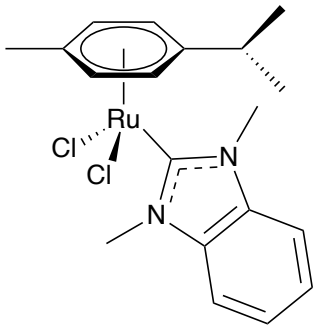


Figure 7.2: Structure of [Ru(Cym)(dmb)Cl_2]

In a recent structural study the reactivity and speciation of [Ru(Cym)(dmb)Cl_2] with the model protein hen-egg white lysozyme (HEWL) was explored.[181] Upon saturating a crystal of HEWL with a solution of the complex [Ru(Cym)(dmb)Cl_2] and analysing the resultant structure, different sites of metal coordination were identified. One site of coordination indicated that, upon protein-metal coordination, the *p*-cymene ligand dissociates from the ruthenium centre and a [Ru(dmb)$Cl_2(OH_x)$] fragment coordinates to His15 and Arg14, **Figure 7.3**. Detailed EPR studies have explored the impact of arene ligand exchange on the oxidation state of the Ru centres and identified that this ruthenium centre has been oxidised to Ru(III) which is paramagnetic and EPR active. This highlights the importance of the arene ligand in stabilising the ruthenium centre in the +II oxidation state as, once ligand exchange occurs, the metal centre is susceptible to oxidation. This methodology, of controlled ligand exchange upon direct coordination identifies a potential way to introduce redox-active Ru centres into protein scaffolds.

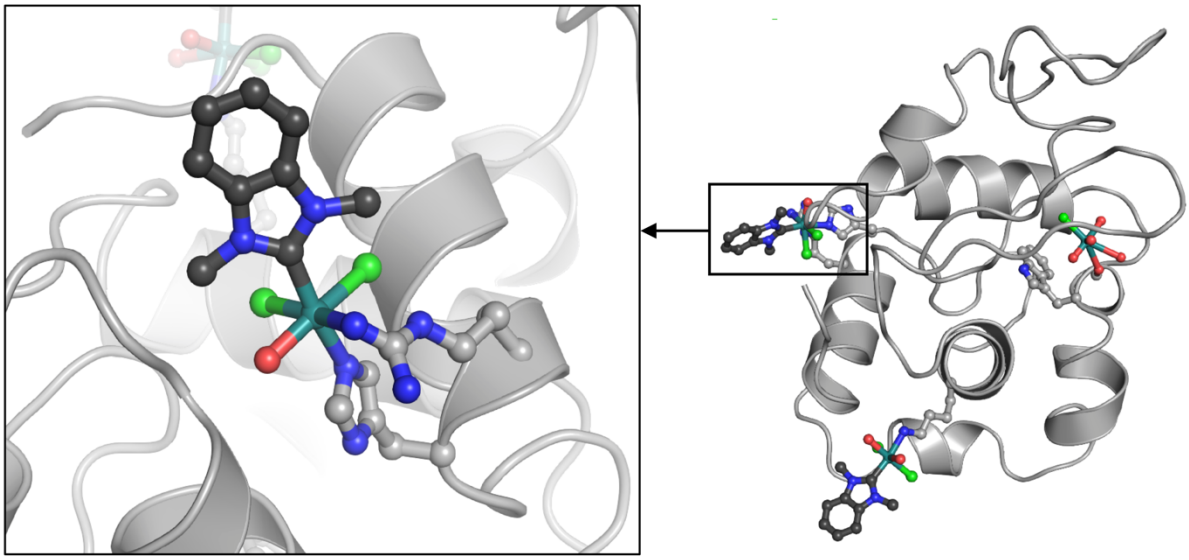


Figure 7.3: Coordination sites identified after reaction of [Ru(Cym)(dmb)Cl_2] with HEWL. Site 1 is highlighted and shows axial coordination of His15 and the dmb ligand and equatorial coordination of Arg14, two Cl ligands and an OH_x ligand. Site 2 has been identified as monodentate coordination of a Ru(dmb)$(OH_x)_2Cl_2$ fragment to Lys33. Site 3 was identified after crystal soaking for one month and involves the weak coordination of a [Ru$(OH_x)_4Cl_2$] fragment to the carbonyl backbone on Ala107. This figure was supplied by Dr Matthew Sullivan.

7.1.2. Pyridylidene amide (PYA) Complexes a New Class of Push and Pull ligands.

For the purpose of redox catalysis, ligands that can both accept and donate electron density based on the oxidation state of the metal are particularly useful. Pyridylidene amide (PYA) ligands are a particularly attractive class as they are characterised by two different resonance structures: (i) a zwitterionic structure with a π–basic N-donor site (**Figure 7.4 (left)**) and (ii) a neutral structure with a π–acidic N-donor site (**Figure 7.4 (right)**). Electron donation is dependent on many factors including solvent polarity, the spectator ligands and metal oxidation state. When coordinated to a redox active metal centre this ligand system has the potential to promote both oxidative addition and reductive elimination in its zwitterionic π–basic and π–acidic forms, respectively.

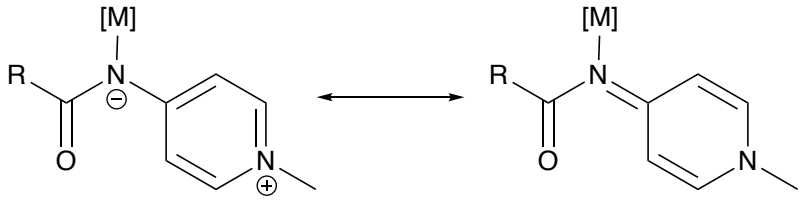


Figure 7.4: Resonance structure of the PYA ligand. (Left) The zwitterionic π–basic N-donor site which stabilises high-valent metal species. (Right) The neutral π–acidic N-donor site which stabilises low-valent metal species.

These ligands have been utilised most effectively by Albrecht et al. who have developed ruthenium complexes incorporating PYA ligands for the oxidative cleavage of olefins using $NaIO_4$ as a sacrificial oxidant, in water/organic solvent mixtures. In this study, they postulate a catalytic cycle involving the formation of a bis-PYA stabilised high valent ruthenium (IV) dioxo species [Ru(bis-PYA)(=O)$_2$] which activates and oxidises the olefin.[182]

7.2. Experimental Goals

The aim of this chapter was to synthesise and characterise a series of Ru(II) (*p*-cymene) complexes with different bidentate ligands, based around the NHC, PYA or pyridine (PYR) framework. Upon synthesis, the coordination and ligand exchange properties of these complexes were explored with cytochrome b_{562} .

7.3. Results and Discussion

7.3.1. Synthesis and Characterisation

Three different Ru(II) (*p*-cymene) complexes were identified as synthetic targets. Complex [18] is novel and has a mixed PYA-PYR framework,[115] complex [19] has a bis-PYA framework and was developed as a novel complex, however has been recently reported by Albrecht et al,[182]. and complex [20] has a mixed NHC-PYR framework and has been reported by Hartinger et al, **Figure 7.5**.[183] These synthetic targets were chosen as they all contained bidentate ligands, and could be synthesised readily from the common starting material dimer III. Furthermore, they provided some functional group variety to explore the ligand exchange properties of PYA and NHC ligands. The synthesis of complexes [18], [19] and [20] are all developed from various literature procedures and reported in **Section 3.3.5.**

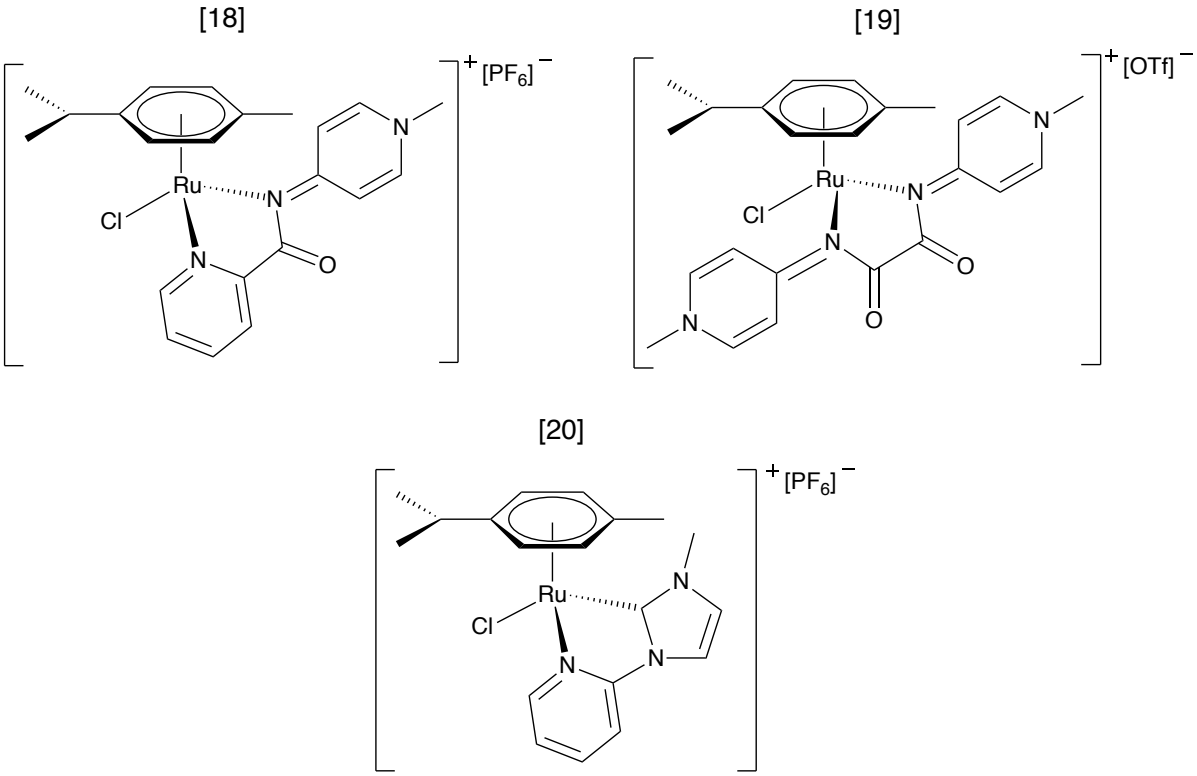


Figure 7.5: Structures of complexes [18], [19] and [20].

Structure of Complex [19]

Complex [19] was purified with a triflate counter ion and diffraction quality crystals were obtained *via* vapour diffusion methods (Et_2O:MeCN). During the process of writing this thesis the structure of complex [19] with the PF_6 counter ion was reported by Albrecht et al.[182] The structure of complex [19] derived from the synthetic procedure outlined in **Section 3.3.5** was also determined as part of this work and is shown in **Figure 7.6**.

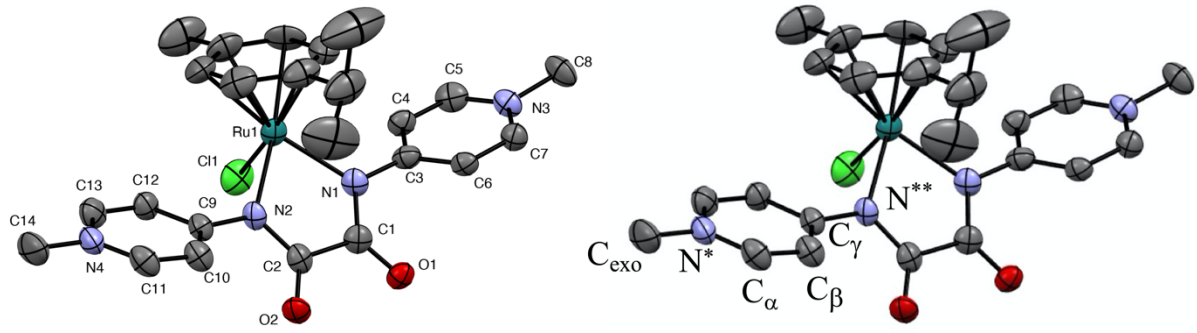


Figure 7.6: Crystal structure of complex [19] showing displacement ellipsoids at 50 % probability. The counterions and hydrogen atoms have been omitted for clarity. **(Left)** Individual labelling of the atoms, **(Right)** Labelling of the atoms in, and adjacent to the pyridyl ring.

Table 7.1: Selected bond lengths (Å) for complex [19]. The C_{α}, C_{β}, C_{γ}, C_{exo}, N^{*} and N^{**} labelling scheme for the pyridyl ring has been introduced in **Figure 7.6**.

Bond	Bond Length Å
Ru1–N1	2.107(7)
Ru1–N2	2.114(8)
N1–C3	1.402(12)
N2–C9	1.385(11)
N1–C1	1.344(12)
N2–C2	1.366(12)
Average Pyridyl Ring Bond Lengths	
C_{β}–C_{γ}	1.394(13)
C_{α}–C_{β}	1.366(14)
C_{α}–N*	1.350(12)
N*–C_{exo}	1.477(12)

By analysing the bond lengths of complex [19] the charge localisation around the pyridyl ring can be explored. The average C_{β}–C_{γ} bond length (1.394(13) Å) is longer than the average C_{α}–C_{β} bond length (1.366(13) Å), however this is not statistically significant. This difference in bond length is corroborated by the published structure from Albrecht et al., who report statistically different bond lengths between the C_{β}–C_{γ} bond (1.405(3) Å) and the C_{α}–C_{β} bond (1.371(3) Å). These data support there being double bond localisation between the α and β carbons, as indicated in the neutral resonance structure, **Figure 7.4** (right).[182] However, the bond lengths from the C_{γ} to the exocyclic nitrogen's (N1–C_{γ} = 1.402(12) and N2–C_{γ} = 1.385(11)) are not particularly short as would be expected for an exocyclic imine, hence highlighting that solid state information is not enough to support the distribution of electron density fitting the neutral resonance model.

7.3.2. Effect of Solvent Polarity on Complex [19]

^{1}H NMR spectroscopy was used to probe the electronic flexibility of the PYA ligand by modifying the polarity of the solvent environment, **Table 7.2**. An adequate approximation of solvent polarity is given by the dielectric constant (ε) or relative permittivity, which simply measures the solvents ability to insulate charges from one other. A higher ε means a higher solvent polarity and greater ability to stabilise charges.

Table 7.2: Selected ^{1}H NMR shifts (ppm) of complex [19] in different solvents. H_α and H_β were identified from NOE spectroscopy, as the H_α gave a clear cross peak with the C_{exo} hydrogens found at 4.1 ppm.

Solvent	ε	H_α	H_β	Δδ [ppm]
d_2-DCM	8.93	8.733	8.029	0.704
d_6-acetone	21.01	8.736	8.487	0.249
d_4-MeOH	32.04	8.562	8.468	0.094
d_3-MeCN	35.94	8.544	8.181	0.363
d_6-DMSO	46.45	8.417	8.515	-0.098

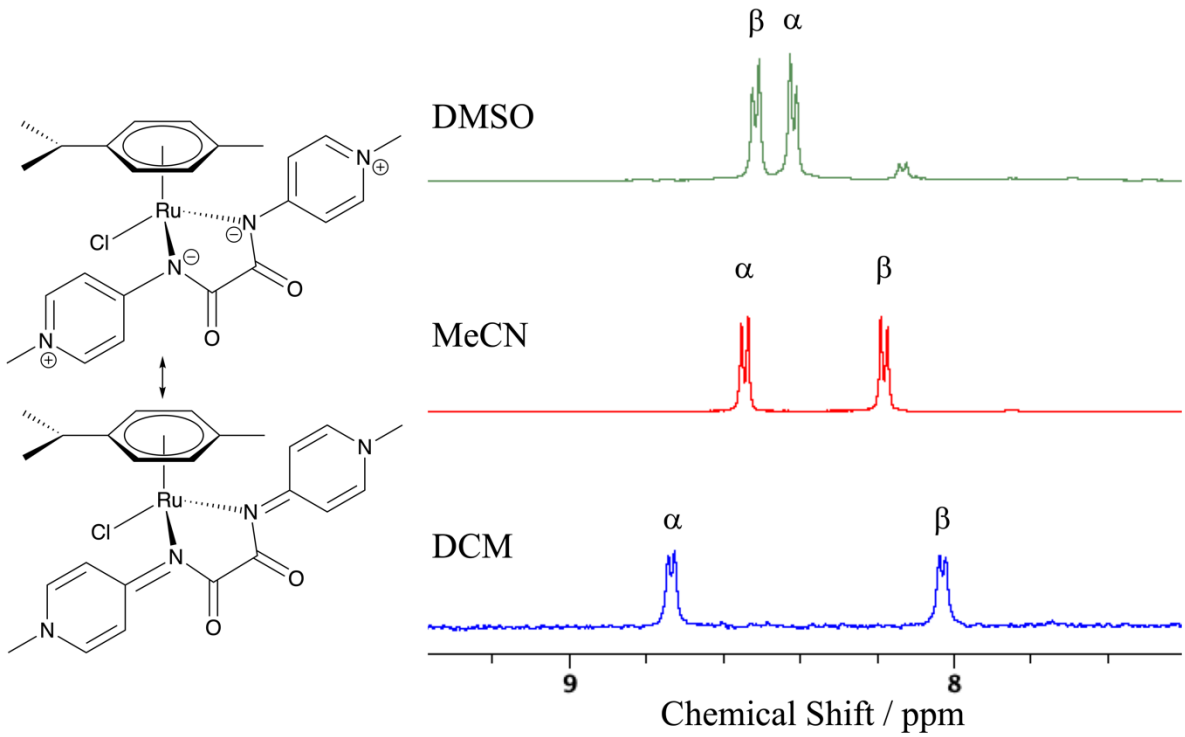


Figure 7.7: Section of the ^{1}H NMR spectra of complex [19] in d_6-DMSO, d_4-MeCN and d_2-DCM showing how changing solvent polarity impacts the chemical shift of the pyridyl protons H_α and H_β. Again, H_α is assigned through a NOE peak to the hydrogen atoms on C_{exo}. The small peak in the DMSO spectra at 8.1 ppm is a minor contaminant.

In non-polar solvents, H_β is upfield of H_α and there is a larger Δδ suggesting a higher contribution of the neutral diene-type resonance structure, (**Figure 7.7** (bottom structure)). At higher solvent polarities, the zwitterionic complex is stabilised to a greater extent, therefore the H_α and H_β have similar chemical shifts, even to an extent where H_β shifts downfield of H_α, confirmed through NOE spectroscopy (**Figure 7.7** (top structure)). Attempts have previously been made to linearly correlate the dielectric constant with the Δδ in different solvents.[116] From our experience, the plot can be fitted differently depending on which solvent

systems are selected, as indicated by **Figure 7.8**. For example, if MeCN is not considered, then the corresponding plot is consistent with an $\varepsilon^{-1/2}$ relationship, (where ε is the dielectric constant) often found for phenomena dominated by electrostatic effects in solution. However, if MeCN is considered the trend could be argued to be linear as reported by Albrecht et al.[116]

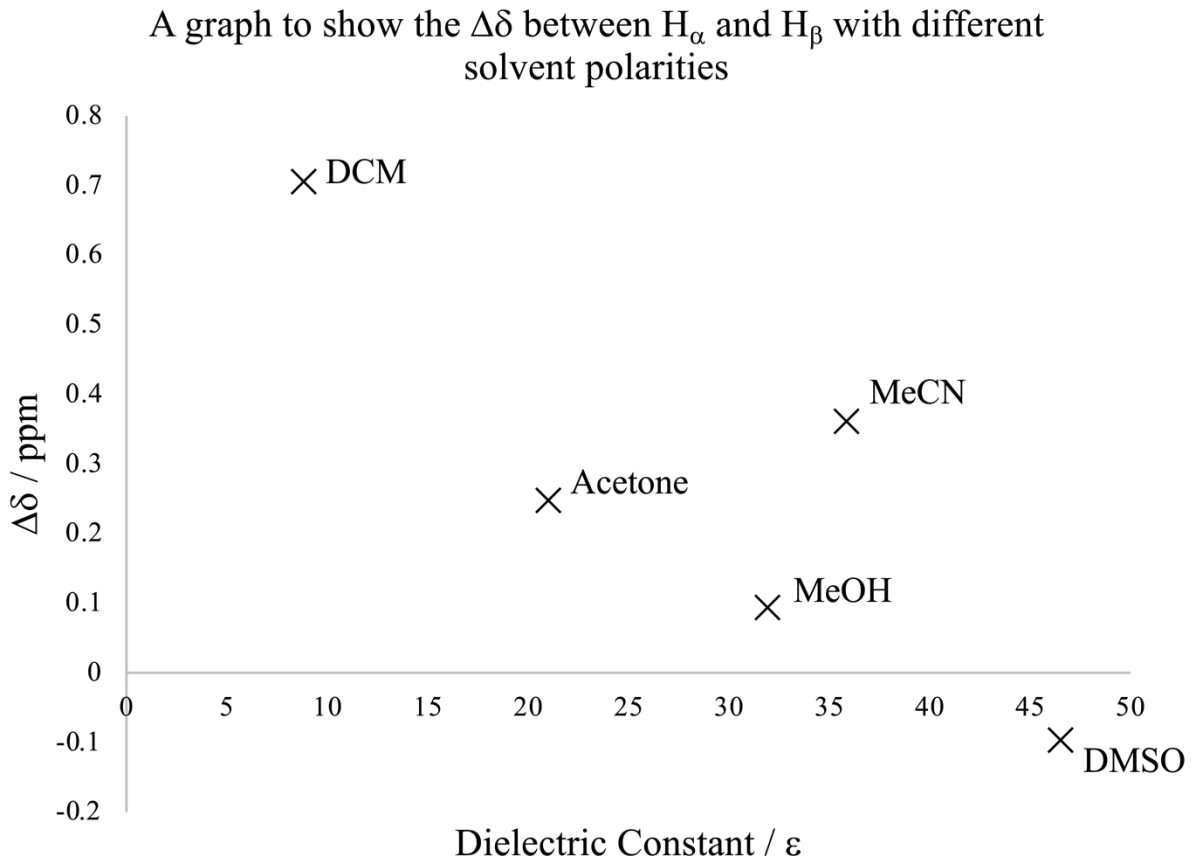
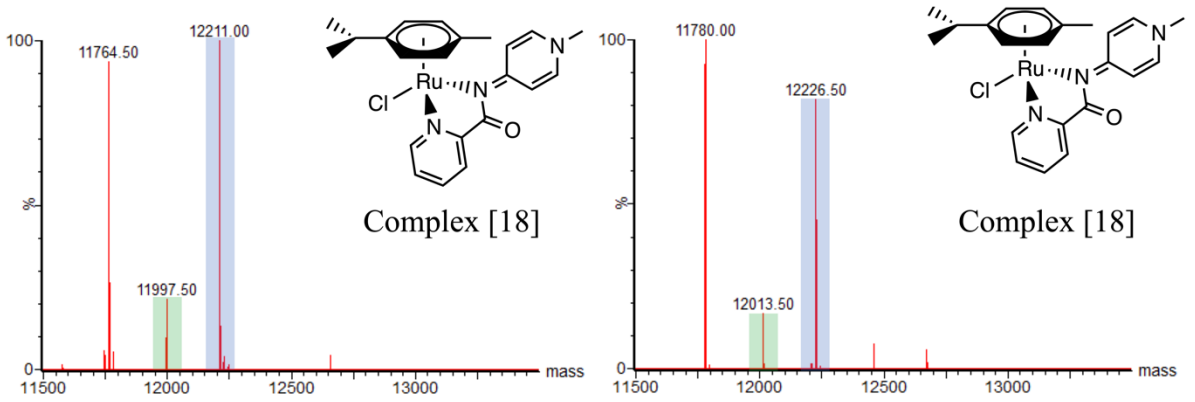


Figure 7.8: A plot of dielectric constants of a range of different deuterated solvents against the difference in chemical shift between pyridyl protons H_α and H_β in complex [19].

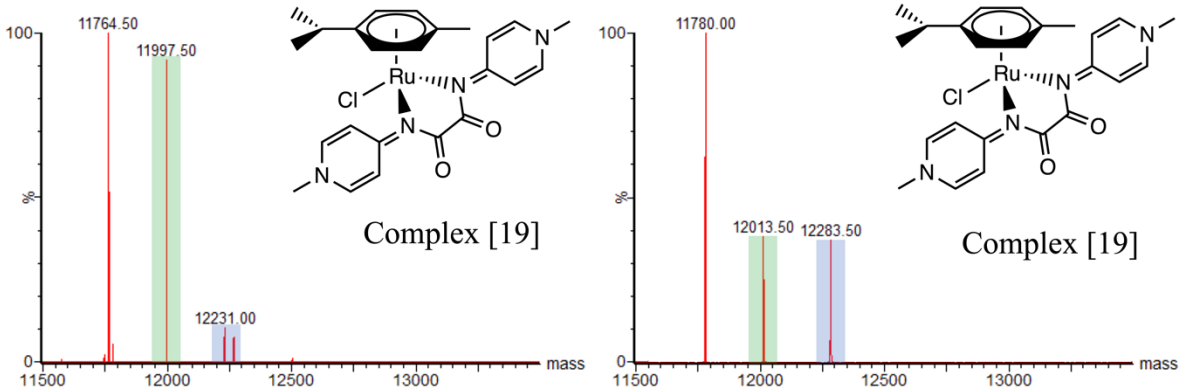
7.3.3. Incubations of Complexes [18], [19] and [20] with Cytochrome *b*$_{562}$

A range of experiments were performed varying the number of equivalents of ruthenium complexes [18], [19] and [20] and concentration of the two Cyt b_{562} variants, the cysteine mutant, Cyt b_{562} L10C, and the wild-type, Cyt b_{562} WT. Samples were taken after 1, 3 and 6 hours and the extent of modification was monitored *via* LC-MS, **Figure 7.9**. Building upon an understanding of the reactivity of the bipyridine complexes with the cysteine mutant Cyt b_{562} L10C, fewer ruthenium equivalents (two) were used in order to focus on the reactivity of the more favourable Ru(II) binding site, Cys10.



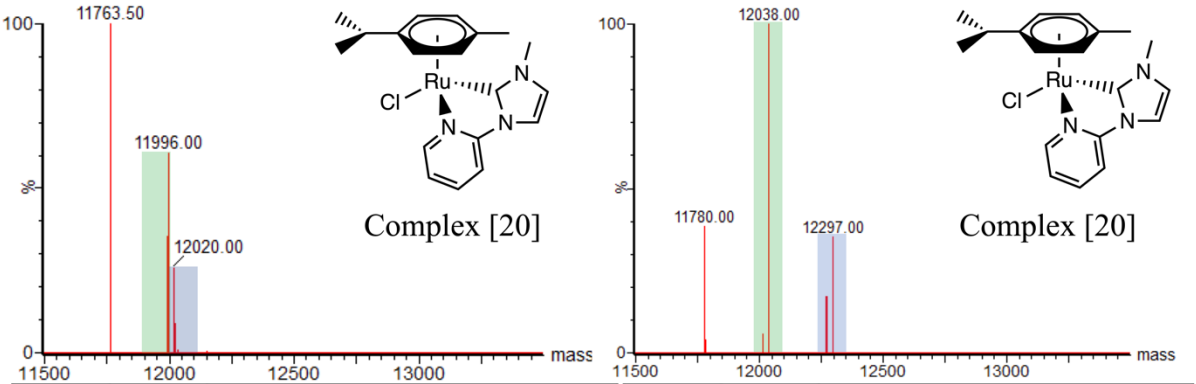
M_{obs}	M_{calc}	Adduct Type
11765	11764	CYT 1
11998	11999	CYT 1 + [Ru(Cym)]
12211	12213	CYT 1 + [Ru(Cym)(PYR-PYA)]

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT WT
12014	12015	CYT WT + [Ru(Cym)]
12227	12229	CYT WT + [Ru(Cym)(PYR-PYA)]



M_{obs}	M_{calc}	Adduct Type
11765	11764	CYT 1
11998	11999	CYT 1 + [Ru(Cym)]
12231	12234	CYT 1 + 2 [Ru(Cym)]

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT WT
12014	12015	CYT WT + [Ru(Cym)]
12284	12286	CYT WT + [Ru(Cym)(bis-PYA)]



M_{obs}	M_{calc}	Adduct Type
11764	11764	CYT 1
11996	11999	CYT 1 + [Ru(Cym)]
12020	12025	CYT 1 + [Ru(PYR-NHC)]

M_{obs}	M_{calc}	Adduct Type
11780	11780	CYT WT
12038	12041	CYT WT + [Ru(PYR-NHC)]
12297	12302	CYT WT + 2 [Ru(PYR-NHC)]

Figure 7.9: Mass spectra from incubations of complexes [18], [19], [20] with Cyt b_{562} L10C – referred to as Cyt 1 **(Left)** (50 µM protein, 2 Eq. Ru, 310 K, 6 Hr) and Cyt b_{562} WT **(Right)** (50 µM protein, 20 Eq. Ru, 310 K, 6 Hr).

7.3.4. Discussion of Reactivity of Complexes [18], [19] and [20] with Cytochrome b_{562}

Complex [18] has a bidentate N,N-coordinating PYR-PYA framework. Incubation of 2 Eq. of complex [18] with Cyt b_{562} shows reactivity with Cys10 occurring to a moderate extent (approximately 50 % modification after 6 hrs) with the most abundant peak at 12211 Da, consistent with modification of Cyt b_{562} L10C with a [Ru(Cym)(PYR-PYA)] fragment. There is also a peak at 11998 Da, consistent with modification of a [Ru(Cym)] fragment to a much lesser extent. A similar spectrum, and pattern of modifications, is observed when incubating 20 Eq. of complex [18] with Cyt b_{562} WT. In previous studies, the reactivity of Ru(II)(η^6-arene)(bipyridine) complexes with Cyt b_{562} WT is very slow, unless complex [16] was used and bipyridine dissociation occurs upon coordination. The major species in this incubation at 12226 Da is consistent with modification of Cyt b_{562} WT with a [Ru(Cym)(PYR-PYA)] fragment. Although modification is still slow, the adduct retains coordination of the PYR-PYA fragment which could be a potential avenue of interest to test for catalysis.

In comparison to complex [18], modification of the cytochrome variants occurs to an even lesser extent with complex [19], which has a bidentate N,N-coordination bis-PYA framework. With 2 Eq. of complex [19] modification at the Cys10 site occurs moderately with the most abundant peak at 11998 Da consistent with modification of a [Ru(Cym)] fragment and dissociation of the bis-PYA ligand. Incubation of Cyt b_{562} WT with 20 Eq. of complex [19] leads to two peaks of similar peak intensity at 12014 Da and 12284 Da, which is consistent with modification of [Ru(Cym)] and [Ru(Cym)(PYA-PYA)] respectively. The major peak, however, is the un-modified protein.

From the incubations of complex [18] and [19] with the cytochrome b_{562} variants there is no sign of *p*-cymene dissociation, therefore, these bidentate N,N-coordinating ligands do not appear to be acting as strong σ-donating ligands, which have been shown to promote arene dissociation. Upon coordination of complex [18] to both Cyt b_{562} L10C and Cyt b_{562} WT the major metal modified species for both these incubations retains coordination of the PYR-PYA ligand. There is some minor species that have undergone PYR-PYA dissociation upon coordination to both cytochrome variants. For complex [19] dissociation of the bidentate bis-PYA ligand occurs more readily, with almost complete dissociation upon coordination to Cyt b_{562} L10C and approximately 50 % dissociation upon coordination to Cyt b_{562} WT. This, however, does not compare with the rapid dissociation of the N,N-coordinating 5,5'-TFMBipy ligand in complex [16] observed in Chapter 5, highlighting the controlled speciation we have with this species.

In comparison to complexes [18] and [19] the bidentate N,C-coordinating complex [20] will react with Cyt b_{562} WT almost to completion, with no un-modified protein left. The key observation is that upon reaction of complex [20] with Cyt b_{562} WT the major peak observed at 12038 Da is consistent with coordination of a [Ru(PYR-NHC)] fragment where *p*-cymene dissociation has occurred. This reactivity draws clear resemblance to the example outlined above, where a [Ru(NHC)] fragment is coordinated to HEWL (**Figure 7.3**). The presence of the NHC ligand at the Ru centre upon coordination to the protein appears to labilise the π-bound *p*-cymene ligand. In other, related work, structural analyses of bond lengths trans to an NHC ligand have revealed an elongation, suggesting a weakening of the bond.[181,184] Dissociation of the *p*-cymene ligand which stabilises the Ru centre in the +II oxidation state, could open up the potential for redox catalysis allowing access to the +III oxidation state and altering the reactivity of the metal complex *in situ*.

Unfortunately attempts to purify these protein – ruthenium carbene adducts *via* ion exchange chromatography have proved unsuccessful to date, however this remains an active area of research.

7.4. Conclusions

Throughout chapters 4 and 5 the bipyridine ligand has been modified with different fluorinated substituents to act as a direct reporter of the coordination environment at the metal and also to modulate the ligand exchange behaviour with protein scaffolds in an attempt to develop artificial metalloenzymes. Transfer hydrogenation and olefin metathesis activity was measured in Chapter 6 for species which had undergone bipyridine dissociation upon coordination to Cyt b_{562} WT, leaving a [Ru(arene)] fragment coordinated to the protein.

This chapter aimed to explore changing the bipyridine ligand to alternative bidentate ligands which can encourage arene dissociation upon coordination to Cyt b_{562}. The bonding of the η^6-arene ligand to the ruthenium centre stabilises the +II oxidation state of the metal, potentially limiting the redox catalysis that can occur within a protein scaffold. It was therefore hypothesised that more electron donating ligands, including N-heterocyclic carbene (NHC) and pyridylidene amide (PYA) ligands, could promote arene dissociation upon protein coordination, and henceforth stabilise higher oxidation states in redox catalytic cycles.

Three different ruthenium complexes with alternative bidentate ligands (NHC-PYR, bis-PYA, PYA-PYR) were successfully synthesised and characterised through a combination of NMR spectroscopy, mass spectrometry and X-ray crystallography. The [Ru(Cym)(bis-PYA)Cl]

complex demonstrated some intriguing electron donating/withdrawing properties which could be probed through variation of the solvent polarity.

These three species were then incubated with both the cysteine mutant Cyt b_{562} L10C and Cyt b_{562} WT. Gratifyingly the [Ru(Cym)(NHC-PYR)Cl] complex undergoes arene dissociation upon coordination to both Cyt b_{562} variants, particularly the wild-type, resulting in a protein-metal hybrid with a [Ru(NHC-PYR)] fragment coordinated to the protein scaffold. These hybrids have four potential ligands provided by the protein scaffold which resembles an exciting evolutionary starting point. Combining this reactivity, with that observed with the bipyridine complexes, highlights the control in ligand exchange that can be achieved through modulating the electron donating properties of the metal ligands. This thesis has highlighted how a number of different protein-metal hybrids can be accessed incorporating a range of distinct metal fragments, and the subtle biochemistry demonstrated in this chapter further adds to this reactivity toolbox.

Chapter 8

8. Concluding Remarks and Future Work

8.1. Restatement of Project Aims

This overall aim of this project was to understand the fundamental reactivity and speciation of Ru(II) arene complexes with proteins. Within this, the aim was to develop novel spectroscopic approaches and use traditional analytical techniques to understand the ligand exchange behaviour of different ruthenium complexes coordinating to proteins. The information gathered on the ligand exchange behaviour of Ru(II) arene complexes can now hopefully be applied to further explore the catalytic potential of ruthenium artificial metalloenzymes with direct protein-metal coordination.

8.2. Conclusions

8.2.1. Understanding the Speciation and Reactivity of Ru(II) Arene Complexes

Throughout this thesis a number of different analytical methods have been used to shed light on the reactivity of Ru(II) arene complexes with proteins. Proteins have a number of potential Lewis basic residues which will form coordination bonds to ruthenium, it is therefore challenging to predict which residue, how fast and to what extent metal modification will occur.

A suite of Ru(II)(arene)(bipyridine) complexes [1]-[17] were synthesised and characterised. Their ligand exchange behaviour has been explored in aqueous buffer, with small molecule amino acid complexes and with proteins through the powerful combination of NMR spectroscopy and mass spectrometry. Fluorinated ligands were incorporated into complexes [5]-[17], and due to the magnetic sensitivity of the fluorine atoms to changes in coordination at the metal centre, the ^{19}F NMR spectra show for the first time the varied, dynamic behaviour of organoruthenium compounds when exposed to simple biomolecules in complex mixtures. Incorporating an NMR-active heteroatom into ruthenium organometallic complexes provides a quantitative, diagnostic ‘fingerprint’ to track solution-phase behaviour and allow for unambiguous assignment of any given amino acid adduct.

Using ^{19}F NMR spectroscopy, coordination of these Ru(II) arene complexes to the Lewis basic side chains of the small molecules His, Cys, Glu and Met was confirmed, with a clear thermodynamic preference for all complexes for the cysteine thiol. Extending this finding,

chosen Ru(II)(arene)(bipyridine) complexes were incubated with seven different cysteine-containing proteins and the extent of modification was monitored using LC-MS. Proteins with naturally occurring free cysteines and genetically introduced cysteines were purified with each protein displaying differential reactivity with the ruthenium complexes. Having differential speciation activity with different proteins is not overly surprising, however being able to monitor and manipulate this activity is challenging.

The cysteine mutants of Ubq K63C and Cyt b_{562} L10C displayed the fastest, most efficient metal modification. The binding preference was confirmed to be the cysteine thiol through ^{19}F NMR spectroscopy and incubations with the cysteine modifying small molecule N-ethylmaleimide. The ^{19}F NMR spectra of purified Cyt b_{562}-Ru adducts are extremely sensitive to the coordination site and also indicate how the protein environment imparts varying electronic and steric pressures on different parts of the metal complex through non-covalent interactions, emphasised by distinct relaxation properties of fluorine substituents across the bipyridine ring. These distinct relaxation properties could potentially be explained by more buried and solvent exposed fluorine positionings, thus highlighting the ability of the scaffold to create asymmetric conditions for an unnatural cofactor.

8.2.2. Controlled Ligand Exchange

Throughout this thesis, the concept of controlled ligand exchange upon protein-metal coordination is crucial. For a [Ru(II)(arene)(bipyridine)Cl]$^+$ complex to coordinate to a cysteine residue, the proposed dominant mechanism, supported by ^{19}F NMR, is initial exchange of the chloride ligand to a water/hydroxide ligand, followed by exchange of the water/hydroxide for the Lewis basic protein ligand, which is driven by the formation of a strong coordination bond. Through varying the electronic contribution of the arene and bipyridine ligands, it was observed that some species will undergo a further ligand exchange process, whereby the bipyridine ligand will dissociate and the protein will provide the necessary ligands to saturate the metal coordination sphere.

This reactivity is exemplified upon incubation of complex [16] (which has three ligands, a chloride, an η^6-hexamethylbenzene and a 5,5'-di(trifluoromethyl)bipyridine) with Cyt b_{562} WT where the major species observed is a Cyt b_{562} WT – [Ru(η^6-hexamethylbenzene)] adduct, whereby the bipyridine and chloride ligands have undergone a ligand exchange process to ligands provided by residues His63, Ala1 and Asn6. In this reaction, we hypothesise that initial coordination to the protein occurs at the His63; following this, the steric and conformational

pressure on the complex results in bipyridine dissociation with Asn and Ala providing the required ligands to satisfy the metal complex. In comparison, upon incubation of complex [16] to the cysteine mutant Cyt b_{562} L10C, the major species retain coordination of the fluorinated bipyridine ligand. Upon coordination to the cysteine thiol, dissociation of the fluorinated bipyridine ligand is observed but to a much lesser extent than when initial coordination occurs at His63. This could be due to the different electronic contributions of the ligand and/or the different steric pressures of the metal complex within the protein.

By changing the bipyridine ligand to the strongly electron donating, bidentate pyridine-carbene ligand, coordination to the protein can promote arene dissociation from the metal centre. As demonstrated from the above examples, the strength of metal-ligand coordination bonds in the complex and the steric and electronic pressures that the protein imparts on the metal complex upon coordination are crucial in controlling ligand dissociation.

Understanding this fundamental ligand exchange behaviour of the metal complex in the development of ArMs with direct protein-metal coordination is imperative. The concept of an inert complex being activated towards catalysis upon protein coordination (similar to vitamin B12 and methionine synthase) is attractive as it ensures that reactivity can only occur within the specific protein environment evolved for rapid turnover and pristine enantioselectivity. This metal complex must be stable for long enough in aqueous conditions, but when the desired protein coordination site is found, ligand exchange must occur and catalytic activity can then be unmasked.

8.2.3. Developing an Artificial Metalloenzyme with Direct Coordination

To develop an ArM which can deliver un-paralleled catalytic activity for a desired reaction, there are many crucial factors to consider including: (i) The starting metal pre-cursor complex, (ii) the starting protein scaffold, (iii) forming protein-metal coordination bonds in a controlled ligand exchange reaction, (iv) non-covalent interactions between the protein and the metal complex and (v) the catalytic substrate interactions to both the metal complex and the protein scaffold.

The molten globule, apocytochrome b_{562} has potential as a starting point for ArM formation and subsequent catalysis. It has a dynamic, four α-helical fold capable of accommodating different ruthenium fragments within a central hydrophobic core by providing multiple ligands to the metal centre from different parts of the protein. Additionally, it is evolvable, i.e. it is

robust enough to withstand mutation and malleable enough for mutations to impact catalysis, therefore it is a good candidate for subsequent directed evolution campaigns.

Throughout this research, it has been identified that cytochrome b_{562} – Ru(II) arene hybrids are a promising starting point to explore the catalytic transfer hydrogenation of imines. One particular Cyt b_{562} WT – [Ru(η^6-hexamethylbenzene)] hybrid showed approximately 5 times faster activity than a known dimeric ruthenium catalyst and up to 35 times faster than the starting ruthenium cofactor. The enhanced activity over the free metal complex highlights the potential to activate a metal complex for catalytic activity through a controlled ligand exchange reaction with a protein. Preliminary results have also shown the potential of cytochrome b_{562} – Ru(II) arene hybrids to perform olefin metathesis, reactivity thought to be beyond the capability of these hybrids. Tentatively, this demonstrates that localising a metal cofactor and substrate within a hydrophobic pocket of a protein can access activity not realised by small molecule catalysis.

What we believe to be crucial for directed evolution of ArMs to make significant improvements in catalysis is direct protein-metal coordination. Directed evolution can significantly improve how the substrate binds and interacts with the metal centre (lowering K_M), but with direct protein-metal coordination directed evolution has the potential to drastically change the ligands coordinated to the metal and vary the proteins electronic contribution to catalysis, thus improve the chemical turnover rate (k_{cat}).

8.2.4. Concluding Remarks

The work presented in this thesis has further developed the understanding of how and where ruthenium complexes form coordination bonds to protein scaffolds. The legacy of this work and the Ru(II) arene fragment stems from the requirement to understand the cellular speciation of Ru(II) arene complexes in order to develop and design new ruthenium complexes that could be relevant within a clinical setting for the treatment of cancer. Although the methodologies described here for monitoring biological speciation are highly transferable to this application, recent advances have been made in developing artificial metalloenzymes with direct protein-ruthenium coordination bonds. Direct coordination between metal ions and a protein scaffold is essential in order to deliver artificial metalloenzymes where the metal and protein are truly interdependent. This represents a method to access entatic states which deliver enhanced reactivity, efficiency and selectivity that cannot easily be replicated in conventional, synthetic metal catalysis. This novel class of hybrid enzymatic catalysts have the potential to greatly

impact chemical production on this planet; hopefully the work presented in this thesis will underpin future advancements made in artificial metalloenzymes with direct protein-metal coordination.

8.3. Future Work

Future work from this research is being directed towards developing artificial metalloenzymes with direct-protein coordination. As this research moves more into the realms of protein engineering and synthetic catalytic chemistry to optimise ArM activity, an appreciation of the metal-ligand coordination environment and ligand exchange behaviour of metal complexes must remain, particularly an understanding of how metal complexes enhance catalytic activity. Time and thought must be given into deciphering the catalytic mechanism of an artificial metalloenzyme and an understanding of this mechanism must also be tested through rational design of protein scaffolds alongside laboratory evolution.

8.3.1. Characterising Cytochrome b_{562} – Ruthenium Coordination

Characterising protein-ruthenium hybrids will continue to be crucial in the development of ArMs as identifying the optimal protein-metal coordination environment for catalysis is essential. Mass spectrometry methods have been ubiquitously used throughout this thesis to confirm protein-ruthenium modification, and in particular tandem MS/MS has proved particularly enlightening. Other structural biology techniques could prove informative in characterising Cytochrome b_{562} – ruthenium adducts, including X-ray crystallography, cryogenic electron microscopy (cryoEM), electron paramagnetic resonance (EPR) and isotopically labelled NMR spectroscopy.

8.3.2. Developing a Ruthenium Fragment Library Coordinated to Cytochrome b_{562}

Throughout this thesis, a number of different fragments have been coordinated to variants of cytochrome b_{562} including: [Ru(arene)(bipyridine)], [Ru(arene)(bis-PYA)], [Ru(arene)(PYR-PYA)], [Ru(arene)(PYR-Carbene)], [Ru(arene)] and [Ru(PYR-Carbene)]. Expanding this library of fragments, for example [Ru(phosphane)] fragments, could widen the scope of potential reactions catalysed by Cyt b_{562} – ruthenium metalloenzymes. Within this research, a number of synthetic advancements have been made in the synthesis, purification and characterisation of Ru(II) arene complexes. Continuing these synthetic advancements could identify novel ruthenium complexes which could have unpredictable and potentially useful ligand exchange properties with protein scaffolds, particularly Cyt b_{562}.

8.3.3. Directed Evolution of a Cytochrome b_{562} – Ruthenium Artificial Metalloenzyme

Following initial transfer hydrogenation activity observed for cytochrome b_{562} WT – [Ru(HMB)] hybrids with direct protein-metal coordination, directed evolutionary campaigns have the potential to evolve this protein scaffold to enhance the catalytic activity of this protein-metal hybrid. This thesis has highlighted three key considerations for a directed evolutionary campaign for ArMs which are: (i) Expressing protein variants with a genotype-phenotype linkage, (ii) a clean and efficient metal modification step and (iii) a quick, sensitive and robust activity assay. This thesis has demonstrated controlled ligand exchange methods that lead to clean and efficient metal modification. Further, fluorescence-based assays have been developed that are sensitive and robust.

For ArMs where an unnatural metal cofactor has been anchored to a protein scaffold, all examples of directed evolution campaigns have been performed *in-vivo* which links the genotype and phenotype through compartmentalization within a cell. Cells expressing different protein variants are separated, typically through colony picking into 96 well plates, and following metal modification, catalytic activity is screened in a plate assay using a fluorogenic or chromogenic substrate. Upon identification of active variants, the cells are lysed and the gene is recovered for continual rounds of evolution.

The major limitations of *in-vivo* methods in the directed evolution of ArMs are: (i) catalytic activity must be compatible with cellular conditions, (ii) the metal modification step must be able to withstand complex biological environments, which is particularly challenging as the cell has many potential catalyst poisons, e.g. glutathione and (iii) low transformation efficiency limits the library size which restricts the sequence space that can be explored in each round of directed evolution.

Considering these limitations, future work will focus on directed evolution of cytochrome b_{562} – ruthenium ArMs through *in-vitro* methods involving compartmentalisation in water in oil emulsions (microfluidics) and covalent linkages between the genotype and phenotype. Upon generation of a gene library, protein expression will be performed in droplets using modern *in-vitro* translation (IVTT) kits. Within the droplet, both the expressed protein variant and gene will be covalently linked to a polyacrylamide hydrogel bead maintaining a covalent genotype-phenotype linkage. At this stage, the droplets can be broken up and protein-metal modification can occur before the beads (with gene and protein variant still covalently attached) are re-encapsulated within a droplet with a fluorogenic substrate. Droplets are sorted based on

fluorescence intensity (FADS), with the genes of active variants recovered and subjected to iterative rounds of directed evolution.

These *in-vitro* translation on beads and droplet compartmentalization methods are particularly attractive techniques for the directed evolution of ArMs. Possibly most important is being able to perform the metal modification step outside of a cellular environment. Additionally, much larger protein libraries can be generated in comparison to *in-vivo* and well-plate systems allowing for a greater sequence space to be explored.

Bibliography

1 G. Swiegers, *Mechanical Catalysis*, John Wiley & Sons, New York, 2008.

2 J. C. Lewis, *ACS Catal.*, 2013, **3**, 2954–2975.

3 R. H. Crabtree, *New J. Chem.*, 2011, **35**, 18–23.

4 A. Hemschemeier and T. Happe, *Nat. Rev. Chem.*, 2018, **2**, 231–243.

5 W. M. Haynes, Ed., *CRC Handbook of Chemistry and Physics*, CRC Press, 97th edn., 2016.

6 R. R. Mendel and T. Kruse, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2012, **1823**, 1568–1579.

7 J. Reedijk, *Platin. Met. Rev.*, 2008, **52**, 2–11.

8 H. Taube, *Chem. Rev.*, 1952, **50**, 69–126.

9 J. J. R. Frausto da Silva and R. J. P. Williams, *The biological chemistry of the elements*, Oxford University Press, 2nd edn., 2001.

10 T. A. Rouault, *Nat. Chem. Biol.*, 2006, **2**, 406–414.

11 K. S. Egorova and V. P. Ananikov, *Organometallics*, 2017, **36**, 4071–4090.

12 Y. He, Q. Zhu, M. Chen, Q. Huang, W. Wang, Q. Li, Y. Huang and W. Di, *Oncotarget*, 2016, **7**, 70803–70821.

13 B. Rosenberg, L. Van Camp and T. Krigas, *Nature*, 1965, **205**, 698–699.

14 A. Abu-Surrah and M. Kettunen, *Curr. Med. Chem.*, 2006, **13**, 1337–1357.

15 T. G. Scrase, S. M. Page, P. D. Barker and S. R. Boss, *Dalt. Trans.*, 2014, **43**, 8158–8161.

16 S. Subramanian and B. T. Kaufmant, *Biochemistry*, 1978, **75**, 3201–3205.

17 J. F. Wishart, H. Taube, K. J. Breslauer and S. S. Isied, *Inorg. Chem.*, 1984, **23**, 2997–3001.

18 A. L. Noffke, A. Habtemariam, A. M. Pizarro, P. J. Sadler, A. Louisa Noffke, Abraha Habtemariam, A. M. Pizarro and P. J. Sadler, *Chem. Commun.*, 2012, **48**, 5219–46.

19 N. Graf and S. J. Lippard, *Adv. Drug Deliv. Rev.*, 2012, **64**, 993–1004.

20 M. J. Clarke, *Coord. Chem. Rev.*, 2002, **232**, 69–93.

21 F. Wang, J. Xu, K. Wu, S. K. Weidt, C. L. Mackay, P. R. R. Langridge-Smith and P. J. Sadler, *Dalt. Trans.*, 2013, **42**, 3188–3195.

22 C. Scolaro, A. B. Chaplin, C. G. Hartinger, A. Bergamo, M. Cocchietto, B. K. Keppler, G. Sava and P. J. Dyson, *Dalt. Trans.*, 2007, **2**, 5065–72.

23 J. M. Rademaker-Lakhai, D. Van Den Bongard, D. Pluim, J. H. Beijnen and J. H. M. Schellens, *Clin. Cancer Res.*, 2004, **10**, 3717–3727.

24 R. Trondl, P. Heffeter, C. R. Kowol, M. A. Jakupec, W. Berger, B. K. Keppler, J. C. Bendell, A. Ogden, D. D. Von Hoff, U. Jaehde, R. Stoika, A. Zaichenko and W. Berger, *Chem. Sci.*, 2014, **5**, 2925–2932.

25 S. Monro, K. L. Colón, H. Yin, J. Roque, P. Konda, S. Gujar, R. P. Thummel, L. Lilge, C. G. Cameron and S. A. McFarland, *Chem. Rev.*, 2019, **119**, 797–828.

26 E. Reisner, V. B. Arion, B. K. Keppler and A. J. L. Pombeiro, *Inorganica Chim. Acta*, 2008, **361**, 1569–1583.

27 P. Zhang and P. J. Sadler, *J. Organomet. Chem.*, 2017, **839**, 5–14.

28 B. S. Murray, M. V. Babak, C. G. Hartinger and P. J. Dyson, *Coord. Chem. Rev.*, 2016, **306**, 86–114.

29 J. M. Cross, N. Gallagher, J. H. Gill, M. Jain, A. W. McNeillis, K. L. Rockley, F. H. Tscherny, N. J. Wirszycz, D. S. Yufit and J. W. Walton, *Dalt. Trans.*, 2016, **45**, 12807–12813.

30 M. V. Babak, S. M. Meier, K. V. M. Huber, J. Reynisson, A. A. Legin, M. A. Jakupec, A. Roller, A. Stukalov, M. Gridling, K. L. Bennett, J. Colinge, W. Berger, P. J. Dyson, G. Superti-Furga, B. K. Keppler and C. G. Hartinger, *Chem. Sci.*, 2015, **6**, 2449–2456.

31 B. Wu, M. S. Ong, M. Groessl, Z. Adhireksan, C. G. Hartinger, P. J. Dyson and C. A. Davey, *Chem. Eur. J.*, 2011, **17**, 3562–6.

32 Z. Adhireksan, G. E. Davey, P. Campomanes, M. Groessl, C. M. Clavel, H. Yu, A. A. Nazarov, C. H. F. Yeo, W. H. Ang, P. Dröge, U. Rothlisberger, P. J. Dyson and C. A. Davey, *Nat. Commun.*, 2014, **5**, 1387–1407.

33 Z. Ma, G. Palermo, Z. Adhireksan, B. S. Murray, T. von Erlach, P. J. Dyson, U. Rothlisberger and C. A. Davey, *Angew. Chemie - Int. Ed.*, 2016, **55**, 7441–7444.

34 S. M. Meier, D. Kreutz, L. Winter, M. H. M. Klose, K. Cseh, T. Weiss, A. Bileck, B. Alte, J. C. Mader, S. Jana, A. Chatterjee, A. Bhattacharyya, M. Hejl, M. A. Jakupec, P. Heffeter, W. Berger, C. G. Hartinger, B. K. Keppler, G. Wiche and C. Gerner, *Angew. Chemie - Int. Ed.*, 2017, **56**, 8267–8271.

35 J. J. Soldevila-Barreda, I. Romero-Canelón, A. Habtemariam and P. J. Sadler, *Nat. Commun.*, 2015, **6**, 6582.

36 B. L. Oliveira, B. J. Stenton, V. B. Unnikrishnan, C. R. de Almeida, J. Conde, M. Negrão, F. S. S. Schneider, C. Cordeiro, M. G. Ferreira, G. F. Caramori, J. B. Domingos, R. Fior and G. J. L. Bernardes, *J. Am. Chem. Soc.*, 2020, **142**, 10869–10880.

37 S. Davies, B. J. Stenton and G. J. L. Bernardes, *Chimia (Aarau).*, 2018, **72**, 771–776.

38 T. Völker, F. Dempwolff, P. L. Graumann and E. Meggers, *Angew. Chemie Int. Ed.*, 2014, **53**, 10536–10540.

39 C. G. Hartinger, M. Groessl, S. M. Meier, A. Casini and P. J. Dyson, *Chem. Soc. Rev.*, 2013, **42**, 6186.

40 A. Casini, G. Mastrobuoni, W. H. H. Ang, C. Gabbiani, G. Pieraccini, G. Moneti, P. J. J. Dyson and L. Messori, *ChemMedChem*, 2007, **2**, 631–635.

41 C. Artner, H. U. Holtkamp, W. Kandioller, C. G. Hartinger, S. M. Meier-Menches, B. K. Keppler, A. A. Nazarov, C. H. Yeo, W. H. Ang and P. Droge, *Chem. Commun.*, 2017, **5**, 3462.

42 I. W. McNae, K. Fishburne, A. Habtemariam, T. M. Hunter, M. Melchart, F. Wang, M. D. Walkinshaw, P. J. Sadler, G. Boyd, D. I. Jodrell and P. J. Sadler, *Chem. Commun.*, 2004, **47**, 1786–1787.

43 M. P. Sullivan, M. Groessl, S. M. Meier, R. L. Kingston, D. C. Goldstone, C. G. Hartinger, T. Sohnel, S. M. F. Jamieson, C. G. Hartinger, W. Weigand, J. Colinge, W. Berger, P. J. Dyson, G. Superti-Furga, B. K. Keppler and C. G. Hartinger, *Chem. Commun.*, 2017, **53**, 4246–4249.

44 C. A. Smith, A. J. Sutherland-Smith, F. Kratz, E. N. Baker and B. H. Keppler, *J. Biol.*

Inorg. Chem., 1996, **1**, 424–431.

45 J. P. Collman, R. R. Gagne, C. A. Reed, T. R. Halbert, W. T. Robinson and G. Lang, *J. Am. Chem. Soc.*, 1975, **97**, 1427–1439.

46 J. P. Collman, *Acc. Chem. Res.*, 1977, **10**, 265–272.

47 L. Pauling, *Nature*, 1964, **203**, 182–183.

48 M. Mukherjee and A. Dey, *ACS Cent. Sci.*, 2019, **5**, 671–682.

49 R. H. Grubbs, *Angew. Chemie - Int. Ed.*, 2006, **45**, 3760–3765.

50 O. M. Ogba, N. C. Warner, D. J. O'Leary and R. H. Grubbs, *Chem. Soc. Rev.*, 2018, **47**, 4510–4544.

51 R. Noyori and S. Hashiguchi, *Acc. Chem. Res.*, 1997, **30**, 97–102.

52 P. B. Arockiam, C. Bruneau and P. H. Dixneuf, *Chem. Rev.*, 2012, **112**, 5879–5918.

53 P. H. J. Carlsen, T. Katsuki, V. S. Martin and K. B. Sharpless, *J. Org. Chem.*, 1981, **46**, 3936–3938.

54 B. C. Boren, S. Narayan, L. K. Rasmussen, L. Zhang, H. Zhao, Z. Lin, G. Jia and V. V Fokin, *J. Am. Chem. Soc.*, 2008, **130**, 8923–8930.

55 M. Jeschek, S. Panke and T. R. Ward, *Trends Biotechnol.*, 2018, **36**, 60–72.

56 J. D. Tyzack, A. J. M. Ribeiro, N. Borkakoti and J. M. Thornton, *ACS Synth. Biol.*, 2019, **8**, 2494–2506.

57 A. Bhushan, P. J. Egli, E. E. Peters, M. F. Freeman and J. Piel, *Nat. Chem.*, 2019, **11**, 931–939.

58 M. C. Wilson, T. Mori, C. Rückert, A. R. Uria, M. J. Helf, K. Takada, C. Gernert, U. A. E. Steffens, N. Heycke, S. Schmitt, C. Rinke, E. J. N. Helfrich, A. O. Brachmann, C. Gurgui, T. Wakimoto, M. Kracht, M. Crüsemann, U. Hentschel, I. Abe, S. Matsunaga, J. Kalinowski, H. Takeyama and J. Piel, *Nature*, 2014, **506**, 58–62.

59 A. Casini, F. Y. Chang, R. Eluere, A. M. King, E. M. Young, Q. M. Dudley, A. Karim, K. Pratt, C. Bristol, A. Forget, A. Ghodasara, R. Warden-Rothman, R. Gan, A. Cristofaro, A. E. Borujeni, M. H. Ryu, J. Li, Y. C. Kwon, H. Wang, E. Tatsis, C. Rodriguez-Lopez, S. O'Connor, M. H. Medema, M. A. Fischbach, M. C. Jewett, C. Voigt and D. B. Gordon, *J. Am. Chem. Soc.*, 2018, **140**, 4302–4316.

60 G. P. Smith, *Science*, 1985, **228**, 1315–1317.

61 G. P. Smith and V. A. Petrenko, *Chem. Rev.*, 1997, **97**, 391–410.

62 T. Clackson, H. R. Hoogenboom, A. D. Griffiths and G. Winter, *Nature*, 1991, **352**, 624–628.

63 F. H. Arnold, *Angew. Chemie - Int. Ed.*, 2018, **57**, 4143–4148.

64 F. H. Arnold, *Acc. Chem. Res.*, 1998, **31**, 125–131.

65 H. Leemhuis, V. Stein, A. D. Griffiths and F. Hollfelder, *Curr. Opin. Struct. Biol.*, 2005, **15**, 472–478.

66 S. Akabori, S. Sakurai, Y. Izumi and Y. Fujii, *Nature*, 1956, **178**, 323–324.

67 K. Yamamura and E. T. Kaiser, *J. Chem. Soc. Chem. Commun.*, 1976, 830–831.

68 M. E. Wilson and G. M. Whitesides, *J. Am. Chem. Soc.*, 1978, **100**, 306–307.

69 H. M. Key, P. Dydio, D. S. Clark and J. F. Hartwig, *Nature*, 2016, **534**, 534–537.

70 P. Dydio, H. M. Key, A. Nazarenko, J. Y.-E. Rha, V. Seyedkazemi, D. S. Clark and J. F. Hartwig, *Science*, 2016, **354**, 102–106.

71 J. N. DeGruyter, L. R. Malins and P. S. Baran, *Biochemistry*, 2017, **56**, 3863–3873.

72 E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.*, 2019, **3**, 147–171.

73 P. Haquette, B. Talbi, S. Canaguier, S. Dagorne, C. Fosse, A. Martel, G. Jaouen and M. Salmain, *Tetrahedron Lett.*, 2008, **49**, 4670–4673.

74 P. Haquette, M. Salmain, K. Svedlunq, A. Martel, B. Rudolf, J. Zakrzewski, S. Cordier, T. Roisnel, C. Fosse and G. Jaouen, *ChemBioChem*, 2007, **8**, 224–231.

75 N. Madern, N. Queyriaux, A. Chevalley, M. Ghasemi, O. Nicolotti, I. Ciofini, G. F. Mangiatordi and M. Salmain, *J. Mol. Catal. B Enzym.*, 2015, **122**, 314–322.

76 A. Onoda, K. Fukumoto, M. Arlt, M. Bocola, U. Schwaneberg and T. Hayashi, *Chem. Commun.*, 2012, **48**, 9756–9758.

77 J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, M. Fernández-González, G. J. L. Bernardes, L. Griffin, H. Hailu, C. J. Schofield and B. G. Davis, *Chem. Sci.*, 2011, **2**, 1666–1676.

78 A. M. Freedy, M. J. Matos, O. Boutureira, F. Corzana, A. Guerreiro, P. Akkapeddi, V. J. Somovilla, T. Rodrigues, K. Nicholls, B. Xie, G. Jiménez-Osés, K. M. Brindle, A. A. Neves and G. J. L. Bernardes, *J. Am. Chem. Soc.*, 2017, **139**, 18365–18375.

79 C. D. Spicer and B. G. Davis, *Nat. Commun.*, 2014, **5**, 4740.

80 K. A. Chambers, N. S. Abularrage, C. J. Hill, I. H. Khan and R. A. Scheck, *Angew. Chemie Int. Ed.*, 2020, **59**, 7350–7355.

81 J. M. Chalker, L. Lercher, N. R. Rose, C. J. Schofield and B. G. Davis, *Angew. Chemie Int. Ed.*, 2012, **51**, 1835–1839.

82 C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989, **244**, 182–188.

83 J. C. Lewis, *Curr. Opin. Chem. Biol.*, 2015, **25**, 27–35.

84 H. Yang, P. Srivastava, C. Zhang and J. C. Lewis, *ChemBioChem*, 2014, **15**, 223–227.

85 P. Srivastava, H. Yang, K. Ellis-Guardiola and J. C. Lewis, *Nat. Commun.*, 2015, **6**, 7789.

86 M. E. Wilson and G. M. Whitesides, *J. Am. Chem. Soc.*, 1978, 100, 306–307.

87 M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, *Nature*, 2016, **537**, 661–665.

88 T. Heinisch, F. Schwizer, B. Garabedian, E. Csibra, M. Jeschek, J. Vallapurackal, V. B. Pinheiro, P. Marlière, S. Panke and T. R. Ward, *Chem. Sci.*, 2018, **9**, 5383–5388.

89 J. Zhao, J. G. Rebelein, H. Mallin, C. Trindler, M. M. Pellizzoni and T. R. Ward, *J. Am. Chem. Soc.*, 2018, **140**, 13171–13175.

90 A. D. Liang, J. Serrano-Plana, R. L. Peterson and T. R. Ward, *Acc. Chem. Res.*, 2019, **52**, 585–595.

91 S. Eda, I. Nasibullin, K. Vong, N. Kudo, M. Yoshida, A. Kurbangalieva and K. Tanaka, *Nat. Catal.*, 2019, **2**, 780–792.

92 D. J. Raines, J. E. Clarke, E. V. Blagova, E. J. Dodson, K. S. Wilson and A. K. Duhme-Klair, *Nat. Catal.*, 2018, **1**, 680–688.

93 D. J. Sommer, M. D. Vaughn, B. C. Clark, J. Tomlin, A. Roy and G. Ghirlanda, *Biochim. Biophys. Acta - Bioenerg.*, 2016, **1857**, 598–603.

94 D. Joseph Sommer, M. David Vaughn and G. Ghirlanda, *Chem. Commun.*, 2014, **50**, 15852–15855.

95 M. Ohashi, T. Koshiyama, T. Ueno, M. Yanase, H. Fujii and Y. Watanabe, *Angew. Chemie - Int. Ed.*, 2003, **42**, 1005–1008.

96 J. Xie, W. Liu and P. G. Schultz, *Angew. Chemie - Int. Ed.*, 2007, **46**, 9239–9242.

97 I. Drienovská, A. Rioz-Martínez, A. Draksharapu and G. Roelfes, *Chem. Sci.*, 2015, **6**, 770–776.

98 J. Bos, W. R. Browne, A. J. M. Driessen and G. Roelfes, *J. Am. Chem. Soc.*, 2015, **137**, 9796–9799.

99 H. B. Gray and J. R. Winkler, *Annu. Rev. Biochem.*, 1996, **65**, 537–561.

100 D. S. Wuttke, M. J. Bjerrum, J. R. Winkler and H. B. Gray, *Science*, 1992, **256**, 1007–1009.

101 R. V. Eck and M. O. Dayhoff, *Science*, 1966, **152**, 363–366.

102 A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone and W. F. DeGrado, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 6298–6305.

103 S. Studer, D. A. Hansen, Z. L. Pianowski, P. R. E. Mittl, A. Debon, S. L. Guffy, B. S. Der, B. Kuhlman and D. Hilvert, *Science*, 2018, **362**, 1285–1288.

104 W. J. Song and F. A. Tezcan, *Science*, 2014, **346**, 1525–1528.

105 P. D. Barker, *Curr. Opin. Struct. Biol.*, 2003, **13**, 490–499.

106 C. A. Blasie and J. M. Berg, *Biochemistry*, 2002, **41**, 15068–15073.

107 B. L. Vallee and R. J. Williams, *Proc. Natl. Acad. Sci. U. S. A.*, 1968, **59**, 498–505.

108 R. B. Leveson-Gower, C. Mayer and G. Roelfes, *Nat. Rev. Chem.*, 2019, 3, 687–705.

109 G. M. Sheldrick, *Acta Crystallogr. Sect. A*, 2015, **71**, 3–8.

110 G. M. Sheldrick, *Acta Crystallogr. Sect. C*, 2015, **71**, 3–8.

111 A. Keller, A. I. Nesvizhskii, E. Kolker and R. Aebersold, *Anal. Chem.*, 2002, **74**, 5383–5392.

112 K. Lee and P. H. Lee, *Tetrahedron Lett.*, 2008, **49**, 4302–4305.

113 G. Winkhaus and H. Singer, *J. Organomet. Chem.*, 1967, **7**, 487–491.

114 R. Lalrempuia and M. R. Kollipara, *Polyhedron*, 2003, **22**, 3155–3160.

115 M. Navarro, V. Rosar, T. Montini, B. Milani and M. Albrecht, *Organometallics*, 2018, **37**, 3619–3630.

116 K. F. Donnelly, C. Segarra, L. X. Shao, R. Suen, H. Müller-Bunz and M. Albrecht, *Organometallics*, 2015, **34**, 4076–4084.

117 S. Movassaghi, S. Singh, A. Mansur, K. K. H. Tong, M. Hanif, H. U. Holtkamp, T. Söhnel, S. M. F. Jamieson and C. G. Hartinger, *Organometallics*, 2018, **37**, 1575–1584.

118 V. Leigh, W. Ghattas, R. Lalrempuia, H. Müller-Bunz, M. T. Pryce and M. Albrecht, *Inorg. Chem.*, 2013, **52**, 5395–5402.

119 A. K. Covington, M. Paabo, R. A. Robinson and R. G. Bates, *Anal. Chem.*, 1968, **40**, 700–706.

120 S. M. Meier-Menches, C. Gerner, W. Berger, C. G. Hartinger and B. K. Keppler, *Chem. Soc. Rev.*, 2018, **47**, 909–928.

121 K. H. Gardner and L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.*, 1998, **27**, 357–406.

122 H. Chen, S. Viel, F. Ziarelli and L. Peng, *Chem. Soc. Rev.*, 2013, **42**, 7971–82.

123 Y. Suzuki, *Chem. Biol.*, 2014, **9**, 1242–1250.

124 H. Chen, S. Viel, F. Ziarelli and L. Peng, *Chem. Soc. Rev.*, 2013, **42**, 7971–7982.

125 C. Dalvit and A. Vulpetti, *J. Med. Chem.*, 2019, **62**, 2218–2244.

126 C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2007, **51**, 243–271.

127 Michael O'Neill, University of Cambridge, 2016.

128 M. Furue, K. Maruyama, T. Oguni, M. Naiki and M. Kamachi, *Inorg. Chem*, 1992, **31**, 3792–3795.

129 L. Wang, Y. Zhang, L. Liu and Y. Wang, *J. Org. Chem.*, 2006, **71**, 1284–1287.

130 H. Chen, J. A. Parkinson, R. E. Morris and P. J. Sadler, *J. Am. Chem. Soc.*, 2003, **125**, 173–186.

131 D. Rose and G. Wilkinson, *J. Chem. Soc. A*, 1970, 1791–1795.

132 J. D. Gilbert, D. Rose and G. Wilkinson, *J. Chem. Soc. A*, 1970, 2765–2769.

133 E. E. Mercer and P. E. Dumas, *Inorg. Chem.*, 1971, **10**, 2755–2759.

134 A. K. Renfrew, A. D. Phillips, E. Tapavicza, R. Scopelliti, U. Rothlisberger and P. J. Dyson, *Organometallics*, 2009, **28**, 5061–5071.

135 S. J. Pike, M. De Poli, W. Zawodny, J. Raftery, S. J. Webb and J. Clayden, *Org. Biomol. Chem.*, 2013, **11**, 3168–3176.

136 A. L. Pinto and S. J. Lippard, *BBA - Rev. Cancer*, 1985, **780**, 167–180.

137 F. Wang, H. Chen, S. Parsons, I. D. H. Oswald, J. E. Davidson and P. J. Sadler, *Chem. - A Eur. J.*, 2003, **9**, 5810–5820.

138 W. Qi, J. Li, C. Y. Chain, G. A. Pasquevich, A. F. Pasquevich and J. A. Cowan, *Chem. Commun.*, 2013, **49**, 6313–6315.

139 W. Hu, Q. Luo, M. Xiaoyan, K. Wu, J. Liu, Y. Chen, S. Xiong, J. Wang, P. J. Sadler and F. Wang, *Chem. - A Eur. J.*, 2009, **15**, 6586–6594.

140 S. M. Marino and V. N. Gladyshev, *J. Mol. Biol.*, 2010, **404**, 902–916.

141 J. M. Chalker, G. J. L. Bernardes, Y. A. Lin and B. G. Davis, *Chem. - An Asian J.*, 2009, **4**, 630–640.

142 C. M. Pickart and M. J. Eddins, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2004, 1695, 55–72.

143 S. Vijay-kumar, C. E. Bugg and W. J. Cook, *J. Mol. Biol.*, 1987, **194**, 531–544.

144 Y. Feng, S. G. Sligar and A. J. Wand, *Nat. Struct. Biol.*, 1994, **1**, 30–35.

145 P. Garcia, M. Bruix, M. Rico, S. Ciofi-Baffoni, L. Banci, M. C. R. Shastry, H. Roder, T. de Lumley Woodyear, C. M. Johnson, A. R. Fersht and P. D. Barker, *J. Mol. Biol.*, 2005, **346**, 331–344.

146 P. D. Barker, E. P. Nerou, S. M. V. Freund and I. M. Feamley, *Biochemistry*, 1995, **34**, 15191–15203.

147 S. Martens, M. M. Kozlov and H. T. McMahon, *Science*, 2007, **316**, 1205–1208.

148 I. S. Alam, A. A. Neves, T. H. Witney, J. Boren and K. M. Brindle, *Bioconjug. Chem.*, 2010, **21**, 884–891.

149 R. B. Sutton, B. A. Davletov, A. M. Berghuis, T. C. Sudhof and S. R. Sprang, *Cell*, 1995, **80**, 929–938.

150 Y. Zhang, A. Ho, J. Yue, L. Kong, Z. Zhou, X. Wu, F. Yang and H. Liang, *Eur. J. Med. Chem.*, 2014, **86**, 449–455.

151 M. Liu, Z. J. Lim, Y. Y. Gwee, A. Levina and P. A. Lay, *Angew. Chemie - Int. Ed.*, 2010, **49**, 1661–1664.

152 S. Sugio, A. Kashima, S. Mochizuki, M. Noda and K. Kobayashi, *Protein Eng.*, 1999, **12**, 439–446.

153 I. G. Kamphuis, K. H. Kalk, M. B. A. Swarte and J. Drenth, *J. Mol. Biol.*, 1984, **179**, 233–256.

154 C. M. Nunn, M. Jeeves, M. J. Cliff, G. T. Urquhart, R. R. George, L. H. Chao, Y. Tscuchia and S. Djordjevic, *J. Mol. Biol.*, 2005, **350**, 145–155.

155 P. J. Loll, P. Xu, J. T. Schmidt and S. L. Melideo, *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, 2014, **70**, 1434–1442.

156 F. Felix, J. Ferguson, H. U. Guedel and A. Ludi, *J. Am. Chem. Soc.*, 1980, **102**, 4096–4102.

157 B. Durham, J. V. Caspar, J. K. Nagle and T. J. Meyer, *J. Am. Chem. Soc.*, 1982, **104**, 4803–4810.

158 R. Grandori, *J. Mass Spectrom.*, 2003, **38**, 11–15.

159 F. Lederer, A. Glatigny, P. H. Bethge, H. D. Bellamy and F. S. Mathews, *J. Mol. Biol.*, 1981, **148**, 427–448.

160 F. Arnesano, L. Banci, I. Bertini, J. Faraone-Mennella, A. Rosato, P. D. Barker and A. R. Fersht, *Biochemistry*, 1999, **38**, 8657–8670.

161 A. Aharoni, L. Gaidukov, O. Khersonsky, S. M. Gould, C. Roodveldt and D. S. Tawfik, *Nat. Genet.*, 2005, **37**, 73–76.

162 R. B. Leveson-Gower, C. Mayer and G. Roelfes, *Nat. Rev. Chem.*, 2019, **3**, 687–705.

163 D. S. Tawfik, *Science*, 2006, **311**, 475–476.

164 X. Garrabou, T. Beck and D. Hilvert, *Angew. Chemie - Int. Ed.*, 2015, **54**, 5609–5612.

165 W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri and A. Lombardi, *Annu. Rev. Biochem.*, 1999, **68**, 779–819.

166 A. Lombardi, F. Pirro, O. Maglio, M. Chino and W. F. DeGrado, *Acc. Chem. Res.*,

2019, **52**, 1148–1159.

167 J. Yan, S. Lee, A. Zhang and J. Yoon, *Chem. Soc. Rev.*, 2018, **47**, 6900–6916.

168 N. Uematsu, A. Fujii, S. Hashiguchi, T. Ikariya and R. Noyori, *J. Am. Chem. Soc.*, 1996, **118**, 4916–4917.

169 T. Wang, L. G. Zhuo, Z. Li, F. Chen, Z. Ding, Y. He, Q. H. Fan, J. Xiang, Z. X. Yu and A. S. C. Chan, *J. Am. Chem. Soc.*, 2011, **133**, 9878–9891.

170 S. E. Clapham, A. Hadzovic and R. H. Morris, *Coord. Chem. Rev.*, 2004, **248**, 2201–2237.

171 D. W. Fink and W. R. Koehler, *Anal. Chem.*, 1970, **42**, 990–993.

172 C. L. Drennan, S. Huang, J. T. Drummond, R. G. Matthews and M. L. Lidwig, *Science*, 1994, **266**, 1669–1674.

173 J. Steinreiber and T. R. Ward, *Coord. Chem. Rev.*, 2008, **252**, 751–766.

174 R. Reuter and T. R. Ward, *Beilstein J. Org. Chem.*, 2015, **11**, 1886–1892.

175 J. W. Jurss, R. S. Khnayzer, J. A. Panetier, K. A. El Roz, E. M. Nichols, M. Head-Gordon, J. R. Long, F. N. Castellano and C. J. Chang, *Chem. Sci.*, 2015, **6**, 4954–4972.

176 T. Bugarcic, A. Habtemariam, R. J. Deeth, F. P. A. Fabbiani, S. Parsons and P. J. Sadler, *Inorg. Chem.*, 2009, **48**, 9444–9453.

177 M. N. Hopkinson, C. Richter, M. Schedler and F. Glorius, *Nature*, 2014, **510**, 485–496.

178 A. J. Arduengo, R. L. Harlow and M. Kline, *J. Am. Chem. Soc.*, 1991, **113**, 361–363.

179 H. G. Raubenheimer and S. Cronje, *Chem. Soc. Rev.*, 2008, **37**, 1998–2011.

180 R. Rubbiani, E. Schuh, A. Meyer, J. Lemke, J. Wimberg, N. Metzler-Nolte, F. Meyer, F. Mohr and I. Ott, *Medchemcomm*, 2013, **4**, 942–948.

181 M. P. Sullivan, M. K. Nieuwoudt, G. A. Bowmaker, N. Y. S. Lam, D. Truong, D. C. Goldstone and C. G. Hartinger, *Chem. Commun.*, 2018, **54**, 6120–6123.

182 K. Salzmann, C. Segarra and M. Albrecht, *Angew. Chemie - Int. Ed.*, 2020, **59**, 8932–8936.

183 S. Movassaghi, S. Singh, A. Mansur, K. K. H. Tong, M. Hanif, H. U. Holtkamp, T. Söhnel, S. M. F. Jamieson and C. G. Hartinger, *Organometallics*, 2018, **37**, 1575–

1584.

184 M. Hollering, M. Albrecht and F. E. Kühn, *Organometallics*, 2016, **35**, 2980–2986.

Appendix A – X-ray Crystallographic Structural Analysis

Table A.1: X-ray crystallographic data for complexes [1] to [4].

Complex	[1]		[2]		[3]		[4]	
CCDC No.	N/A		N/A		N/A		N/A	
Empirical formula	$[C_{16}H_{14}ClN_2Ru]^+(PF_6^-)$		$[C_{17}H_{16}ClN_2Ru]^+(PF_6^-)$		$[C_{20}H_{22}ClN_2Ru]^+(PF_6^-)$		$[C_{22}H_{26}ClN_2Ru]^+(PF_6^-)$	
Formula weight	515.78		529.81		571.88		599.94	
Temperature	180(2) K		180(2) K		180(2) K		180(2) K	
Wavelength	0.7107 Å		1.54178 Å		1.54178 Å		1.54178 Å	
Crystal system	orthorhombic		monoclinic		orthorhombic		orthorhombic	
Space group	$Cmc2_1$		$P\ 2_1/c$		$P\ 2_12_12_1$		$Pna2_1$	
Unit cell dimensions	11.8079(2)	α = 90 °	14.4572(4) Å	α = 90 °	11.6945(2) Å	α = 90 °	18.1054(5) Å	α = 90 °
	10.8639(2)	β = 90 °	8.3310(3) Å	β = 109.350(2) °	12.2925(2) Å	β = 90 °	7.6932(2) Å	β = 90 °
	14.2305(3)	γ = 90 °	16.4324(5) Å	γ = 90 °	15.2581(3) Å	γ = 90 °	16.6857(5) Å	γ = 90 °
Volume	1825.49(6) $Å^3$		1867.37(10) $Å^3$		2193.42(7) $Å^3$		2324.13(11) $Å^3$	
Z	4		4		4		4	
Density (calculated)	1.877 mg/m^3		1.885 mg/m^3		1.732 mg/m^3		1.715 mg/m^3	
Absorption coefficient	1.155 mm^{-1}		9.531 mm^{-1}		8.164 mm^{-1}		7.736 mm^{-1}	
F(000)	1016		1048		1144		1208	
Crystal size	0.21 x 0.18 x 0.14 mm^3		0.08. x 0.05. x 0.02 mm^3		0.20 x 0.05 x 0.05 mm^3		0.12 x 0.10 x 0.04 mm^3	
Theta range	1.00 to 32.03 °		3.24 to 66.76 °		4.62 to 66.63°		4.89 to 66.70 °	
Index ranges	-16<=h<=17, -11<=k<=16, -21<=l<=19		-17<=h<17=, -9<=k<=9, -18<=l<=19		-13<=h<=13, -13<=k<=14, -18<=l<=14		-21<=h<=21, -9<=k<=-9, -19<=l<=19	
Reflections collected	10174		21414		15177		21012	
Independent reflections	3120 [R(int) = 0.030]		3295 [R(int) = 0.077]		3853 [R(int) = 0.0406]		4073 [R(int) = 0.065]	
Completeness to θ(max)			99.8 %		99.9 %		99.9 %	
Data / restraints / param.	3120 / 1 / 133		3295 / 0 / 254		3853 / 0 / 283		4073 / 1 / 305	
Goodness-of-fit on F^2	1.036		1.068		1.096		1.042	
R indices [I>2sigma(I)]	R1 = 0.028	wR2 = 0.066	R1 = 0.032	wR2 = 0.061	R1 = 0.022	wR2 = 0.051	R1 = 0.027	wR2 = 0.056
R indices (all data)	R1 = 0.031	wR2 = 0.067	R1 = 0.046	wR2 = 0.064	R1 = -0.025	wR2 = 0.052	R1 = 0.032	wR2 = 0.057
Largest diff. peak and hole			0.52 and -0.42 e.$Å^{-3}$		0.26 to -0.41e.$Å^{-3}$		0.033 to -0.36 e.$Å^{-3}$	
Flack parameter	-0.04(4)				-0.022(6)		0.020(13)	

Table A.2: X-ray crystallographic data for complexes [5] to [8].

Complex	[5]		[6]		[7]		[8]	
CCDC No.	1867665		1867666		1867667		1867668	
Empirical formula	$[C_{16}H_{12}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{17}H_{14}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{20}H_{20}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{22}H_{24}ClF_2N_2Ru]^+(PF_6^-)$	
Formula weight	551.77		565.79		607.87		635.92	
Temperature	180(2) K		180(2) K		180(2) K		180(2) K	
Wavelength	0.7107 Å		1.5418 Å		1.5418 Å		1.5418 Å	
Crystal system	Triclinic		Orthorhombic		Orthorhombic		Orthorhombic	
Space group	P–1		$Cmc2_1$		$Pca2_1$		$Pna2_1$	
Unit cell dimensions	a = 8.0630(5) Å	α = 104.378(3) °	a = 11.8675(3) Å	α = 90°	a = 12.3683(3) Å	α = 90 °	a = 18.4489(5) Å	α = 90 °
	b = 8.3719(5) Å	β = 95.336(3) °	b = 11.3437(3) Å	β = 90 °	b = 13.3931(3) Å	β = 90 °	b = 7.8290(2) Å	β = 90 °
	c = 13.8505(11) Å	γ = 96.409(3) °	c = 14.2301(4) Å	γ = 90 °	c = 13.3561(3) Å	γ = 90 °	c = 16.0806(4) Å	γ = 90 °
Volume	892.90(11) $Å^3$		1915.68(9) $Å^3$		2212.44(9) $Å^3$		2322.62(10) $Å^3$	
Z	2		4		4		4	
Density (calculated)	2.052 mg/m^3		1.962 mg/m^3		1.825 mg/m^3		1.819 mg/m^3	
Absorption coefficient	1.203 mm^{-1}		9.498 mm^{-1}		8.273 mm^{-1}		7.912 mm^{-1}	
F(000)	540		1112		1208		1272	
Crystal size	0.10 x 0.04 x 0.01 mm^3		0.12 x 0.10 x 0.02 mm^3		0.25 x 0.06 x 0.02 mm^3		0.40 x 0.08 x 0.03 mm^3	
Theta range	3.53 to 25.14°		5.39 to 70.42°		3.30 to 66.78°		4.79 to 66.93°	
Index ranges	-9<=h<=9, -9<=k<=9, -14<=l<=16		-14<=h<=14, -12<=k<=13, -17<=l<=17		-14<=h<=14, -15<=k<=15, -13<=l<=15		-21<=h<=21, -9<=k<=9, -19<=l<=16	
Reflections collected	7589		14279		27010		14417	
Independent reflections	3118 [R(int) = 0.105]		1909 [R(int) = 0.042]		3787 [R(int) = 0.043]		3794 [R(int) = 0.042]	
Completeness to θ(max)	97.6 %		99.9 %		99.9 %		99.5 %	
Data / restraints / param.	3118 / 0 / 262		1909 / 43 / 164		3787 / 1 / 301		3794 / 1 / 322	
Goodness-of-fit on F^2	1.10		1.03		1.05		1.04	
R indices [I>2sigma(I)]	R1 = 0.072	wR2 = 0.110	R1 = 0.033	wR2 = 0.087	R1 = 0.019	wR2 = 0.044	R1 = 0.025	wR2 = 0.054
R indices (all data)	R1= 0.124	wR2 = 0.128	R1 = 0.035	wR2 = 0.088	R1 = 0.022	wR2 = 0.045	R1 = 0.030	wR2 =0.056
Largest diff. peak and hole	0.73 and -0.76 $e.Å^{-3}$		0.71 and -0.48 $e.Å^{-3}$		0.25 and -0.31 $e.Å^{-3}$		0.45 and -0.30 $e.Å^{-3}$	
Flack parameter			0.003(17)		0.017(6)		0.017(6)	

Table A.3: X-ray crystallographic data for complexes [9] to [12]

Complex	[9]		[10]		[11]		[12]	
CCDC No.	1867661		1867662		1867663		1867664	
Empirical formula	$[C_{16}H_{12}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{17}H_{14}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{20}H_{20}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{22}H_{24}ClF_2N_2Ru]^+(PF_6^-)$	
Formula weight	551.77		565.79		607.87		635.92	
Temperature	180(2) K		180(2) K		180(2) K		180(2) K	
Wavelength	0.7107 Å		1.5418 Å		0.7107 Å		1.5418 Å	
Crystal system	Orthorhombic		Triclinic		Orthorhombic		Triclinic	
Space group	Pbcm		P–1		$Pca2_1$		P–1	
Unit cell dimensions	a = 8.18600(10) Å	α = 90°	a = 7.1872(3) Å	α = 99.964(2)°	a = 12.0020(2) Å	α = 90°	a = 8.4825(3) Å	α = 103.1690(11)°
	b = 12.7339(2) Å	β = 90°	b = 11.6858(4) Å	β = 97.263(2)°	b = 13.7944(2) Å	β = 90°	b = 11.8360(4) Å	β = 94.0286(11)°
	c = 17.9660(3) Å	γ = 90°	c = 11.9743(4) Å	γ = 101.058(2)°	c = 13.2815(2) Å	γ = 90°	c = 12.4837(4) Å	γ = 103.4380(10)°
Volume	1872.77(5) $Å^3$		958.79(6) $Å^3$		2198.89(6) $Å^3$		1176.89(7) $Å^3$	
Z	4		2		4		2	
Density (calculated)	1.957 mg/m^3		1.960 mg/m^3		1.836 mg/m^3		1.795 mg/m^3	
Absorption coefficient	1.147 mm^{-1}		9.488 mm^{-1}		0.986 mm^{-1}		7.808 mm^{-1}	
F(000)	1080		556		1208		636	
Crystal size	0.30 x 0.23 x 0.18 mm^3		0.04 x 0.04 x 0.03 mm^3		0.32 x 0.30 x 0.18 mm^3		0.18 x 0.12 x 0.06 mm^3	
Theta range	3.73 to 32.02°		3.80 to 66.73°		3.70 to 33.70°		3.67 to 67.18°	
Index ranges	-12<=h<=12, -17<=k<=19, -26<=l<=26		-8<=h<=8, -13<=k<=13, -14<=l<=14		-18<=h<=18, -21<=k<=21, -20<=l<=20		-10<=h<=10, -13<=k<=14, -14<=l<=14	
Reflections collected	16324		12013		24942		12913	
Independent reflections	3323 [R(int) = 0.035]		3393 [R(int) = 0.049]		8467 [R(int) = 0.031]		4156 [R(int) = 0.025]	
Completeness to θ(max)	99.3 %		99.6 %		99.8 %		98.8 %	
Data / restraints / param.	3323 / 0 / 135		3393 / 0 / 275		8467 / 1 / 301		4156 / 34 / 360	
Goodness-of-fit on F^2	1.09		1.05		1.00		1.08	
R indices [I>2sigma(I)]	R1 = 0.032	R1 = 0.040	R1 = 0.040	wR2 = 0.091	R1 = 0.027	wR2 = 0.069	R1 = 0.024	wR2 = 0.059
R indices (all data)	R1 = 0.039	R1 = 0.054	R1 = 0.054	wR2 = 0.096	R1 = 0.032	wR2 = 0.071	R1 = 0.025	wR2 = 0.060
Largest diff. peak and hole	0.87 and -0.78 $e.Å^{-3}$		1.38 and -0.59 $e.Å^{-3}$		0.74 and -0.83 $e.Å^{-3}$		0.45 and -0.30 $e.Å^{-3}$	
Flack parameter					-0.030(14)			

Table A.4: X-ray crystallographic data for complexes [13] to [16].

Complex	[13]		[14]		[15]		[16]	
CCDC No.	1867670							
Empirical formula	$[C_{18}H_{12}ClF_6N_2Ru]^+(PF_6^-)$		$[C_{19}H_{14}ClF_6N_2Ru]^+(PF_6^-)$		$[C_{22}H_{20}ClF_6N_2Ru]^+(PF_6^-)$		$[C_{24}H_{24}ClFN_2PRu]^+(PF_6^-)$ 0.425(H_2O)	
Formula weight	651.79		665.81		707.89		743.66	
Temperature	180(2) K		180(2) K		180(2) K		180(2) K	
Wavelength	0.7107 Å		1.5418 Å		0.7107 Å		1.5418 Å	
Crystal system	Orthorhombic		monoclinic		Orthorhombic		orthorhombic	
Space group	$Pca2_1$		P $2_1/n$		Pbca		$Pca2_1$	
Unit cell dimensions	a = 16.0681(5) Å	α = 90°	9.5903(3) Å	α = 90 °	a = 15.8815(4)	α = 90°	22.9874(9) Å	α = 90 °
	b = 8.0666(3) Å	β = 90°	13.5886(5) Å	β = 95.963(2) °	b = 17.5702(4)	β = 90°	8.9895(5) Å	β = 90 °
	c = 15.9972(5) Å	γ = 90°	17.1719(6) Å	γ = 90 °	c = 18.4954(7)	γ = 90°	27.0327(11) Å	γ = 90 °
Volume	2073.48(12) $Å^3$		2225.71(13) $Å^3$		5161.0(3)		5586.2(4) $Å^3$	
Z	4		4		8		8	
Density (calculated)	2.088 mg/m^3		1.987 mg/m^3		1.822 mg/m^3		1.768 mg/m^3	
Absorption coefficient	1.080 mm^{-1}		8.565 mm^{-1}		0.876 mm^{-1}		6.909 mm^{-1}	
F(000)	1272		1304		2800		2962	
Crystal size	0.10 x 0.07 x 0.05 mm^3		0.10 x 0.10 x 0.10 mm^3		0.22 x 0.18 x 0.08 mm^3		0.15 x 0.12 x 0.05 mm^3	
Theta range	4.39 to 27.50°		4.16 to 66.94°		3.58 to 25.02°		4.18 to 54.28°	
Index ranges	-20<=h<=16, -10<=k<=10, -20<=l<=20		-11<=h<=11, -16<=k<=16, -20<=l<=20		-18<=h<=18, -20<=k<=20, -21<=l<=21		-24<=h<=24, -9<=k<=9, -28<=l<=28	
Reflections collected	8822		26162		22559		40731	
Independent reflections	3778 [R(int) = 0.057]		3967 [R(int) = 0.055]		4531 [R(int) = 0.104]		6815 [R(int) = 0.19]	
Completeness to θ(max)	97.6 %		99.8 %		99.5%		99.9 %	
Data / restraints / param.	3778 / 1 / 316		3967 / 25 / 331		4531 / 0 / 355		6815 / 289 / 761	
Goodness-of-fit on F^2	1.07		1.02		1.008		1.03	
R indices [I>2sigma(I)]	R1 = 0.041	wR2 = 0.074	R1 = 0.029	wR2 = 0.063	R1 = 0.052	wR2 = 0.0741	R1 = 0.066	wR2 = 0.13
R indices (all data)	R1= 0.061	wR2 = 0.082	R1 = 0.036	wR2 = 0.066	R1 = 0.110	wR2 = 0.0833	R1 = 0.098	wR2 = 0.14
Largest diff. peak and hole	0.59 and -0.64 e.$Å^{-3}$		1.01 and -0.83 e.$Å^{-3}$		0.73 and -0.44 e.$Å^{-3}$		0.44 to -0.50 e.$Å^{-3}$	
Flack parameter	-0.04(4)							

Table A.5: X-ray crystallographic data for complexes [17] and [19].

Complex	[17]		[19]	
CCDC No.	1867669		N/A	
Empirical formula	$[C_{16}H_{12}ClF_2N_2Ru]^+(PF_6^-)$		$[C_{24}H_{28}ClN_4O_2Ru]^+ (CF_3SO_3)^-$	
Formula weight	551.77		690.09	
Temperature	180(2) K		180(2) K	
Wavelength	0.7107 Å		1.54178 Å	
Crystal system	Monoclinic		triclinic	
Space group	$P2_1/n$		P -1	
Unit cell dimensions	a = 11.8901(4) Å	α = 90°	8.8457(4) Å	α = 100.884(4)°
	b = 11.9766(4) Å	β = 101.146(2)°	11.3399(6) Å	β = 101.826(4)°
	c = 12.8399(5) Å	γ = 90°	14.6021(8) Å	γ = 98.315(4)°
Volume	1793.95(11) $Å^3$		1382.01(13) $Å^3$	
Z	4		2	
Density (calculated)	2.043 mg/m^3		1.658 mg/m^3	
Absorption coefficient	1.198 mm^{-1}		6.756 mm^{-1}	
F(000)	1080		700	
Crystal size	0.10 x 0.05 x 0.03 mm^3		0.05 x 0.02 x 0.02 mm^3	
Theta range	3.66 to 27.50°		3.17 to 66.66°	
Index ranges	-15<=h<=15, -15<=k<=15, -16<=l<=16		-10<=h<=7, -11<=k<=13, -16<=l<=17	
Reflections collected	11426		7365	
Independent reflections	4031 [R(int) = 0.086]		4493 [R(int) = 0.063]	
Completeness to θ(max)	97.7%		91.7 %	
Data / restraints / param.	4031 / 0 / 262		4493 / 121 / 407	
Goodness-of-fit on F^2	1.03		1.04	
R indices [I>2sigma(I)]	R1 = 0.051	wR2 = 0.085	R1 = 0.073	wR2 = 0.192
R indices (all data)	R1 = 0.105	wR2 = 0.103	R1 = 0.116	wR2 = 0.162
Largest diff. peak and hole	0.65 and -0.99 $e.Å^{-3}$		1.24 to -0.69$e.Å^{-3}$	
Flack parameter				

Appendix B – Mass Spectrometry Supporting Information

Table B.1: LR-ESI-MS spectra recorded of the solution mixture when complexes [9]-[12] are incubated in 10 mM phosphate buffered D_2O. Samples are taken directly from the solution mixture and diluted before direct injection into the Micromass Quattro mass spectrometer.

Species	M_{obs}	M_{calc}	Molecular Formula / Assignment
[9]-Cl	406.82	406.97	$C_{16}H_{12}ClF_2N_2Ru$
[9]-Phoshate	468.77	468.97	$C_{16}H_{14}F_2N_2O_4PRu$
[10]-Cl	420.83	420.99	$C_{17}H_{14}ClF_2N_2Ru$
[10]-Phoshate	482.84	482.99	$C_{17}H_{16}ClF_2N_2O_4PRu$
[11]-Cl	462.85	463.03	$C_{20}H_{20}ClF_2N_2Ru$
[11]-Phoshate	524.87	525.03	$C_{20}H_{22}ClF_2N_2O_4PRu$
[12]-Cl	491.08	491.06	$C_{22}H_{24}ClF_2N_2Ru$
[12]-Phoshate	552.96	553.06	$C_{22}H_{26}ClF_2N_2O_4PRu$

Table B.2: ESI-MS data from the incubations of complexes [1] – [4] with the protected amino acids, N-Ac-Cys-OMe, N-Z-Glu-OMe, N-Bz-His-OMe, N-Ac-Met-OMe and reduced glutathione (2 mM Ru, 3 eq. amino acid, 24 hr, 310 K). Samples are diluted in D_2O diluted before direct injection into the Micromass Quattro mass spectrometer, therefore the titratable groups present remain deuterated.

Species	M_{obs}	M_{calc}	Molecular Formula / Assignment
[9]-Cys	549.09	549.05	$C_{22}H_{21}DF_2N_3O_3RuS$
[9]-Glu	667.15	667.11	$C_{30}H_{27}DF_2N_3O_6Ru$
[9]-His	323.68	323.33	$C_{30}H_{25}D_2F_2N_5O_3Ru$
[9]-Met	289.10	288.82	$C_{24}H_{26}DF_2N_3O_3RuS$
[9]-GSH	683.13	683.11	$C_{26}H_{23}D_5F_2N_5O_6RuS$
[10]-Cys	563.01	563.06	$C_{23}H_{23}DF_2N_3O_3RuS$
[10]-Glu	681.07	681.12	$C_{31}H_{29}DF_2N_3O_6Ru$
[10]-His	330.79	330.34	$C_{31}H_{27}D_2F_2N_5O_3Ru$
[10]-Met	296.14	295.83	$C_{25}H_{28}DF_2N_3O_3RuS$
[10]-GSH	697.08	697.12	$C_{27}H_{25}D_5F_2N_5O_6RuS$
[11]-Cys	605.00	605.11	$C_{26}H_{29}DF_2N_3O_3RuS$
[11]-Glu	722.96	723.17	$C_{34}H_{35}DF_2N_3O_6Ru$
[11]-His	350.39	351.60	$C_{34}H_{33}D_2F_2N_5O_3Ru$
[11]-Met	317.30	317.08	$C_{28}H_{34}DF_2N_3O_3RuS$
[11]-GSH	740.31	740.17	$C_{27}H_{30}D_6F_2N_5O_6RuS$
[12]-Cys	633.27	633.14	$C_{28}H_{33}DF_2N_3O_3RuS$
[12]-Glu	751.16	751.20	$C_{36}H_{39}DF_2N_3O_6Ru$
[12]-His	365.71	365.61	$C_{36}H_{37}D_2F_2N_5O_3Ru$
[12]-Met	331.28	331.09	$C_{30}H_{38}DF_2N_3O_3RuS$
[12]-GSH	768.24	768.20	$C_{32}H_{35}D_6F_2N_5O_6RuS$

Table B.3: ESI-MS data from the incubations of complexes [12] with the protected amino acids N-Ac-Cys-OMe upon the addition of 0.5 Eq. of H_2O_2. Samples are diluted in D_2O diluted before direct injection into the Micromass Quattro mass spectrometer, therefore the titratable groups present remain deuterated.

Species	M_{obs}	M_{calc}	Molecular Formula / Assignment
[12]-Cl	491.08	491.06	$C_{22}H_{24}ClF_2N_2Ru$
[12]-Cys	633.27	633.14	$C_{28}H_{33}DF_2N_3O_3RuS$
[12]-Cys + [O]	649.27	649.13	$C_{28}H_{33}DF_2N_3O_4RuS$

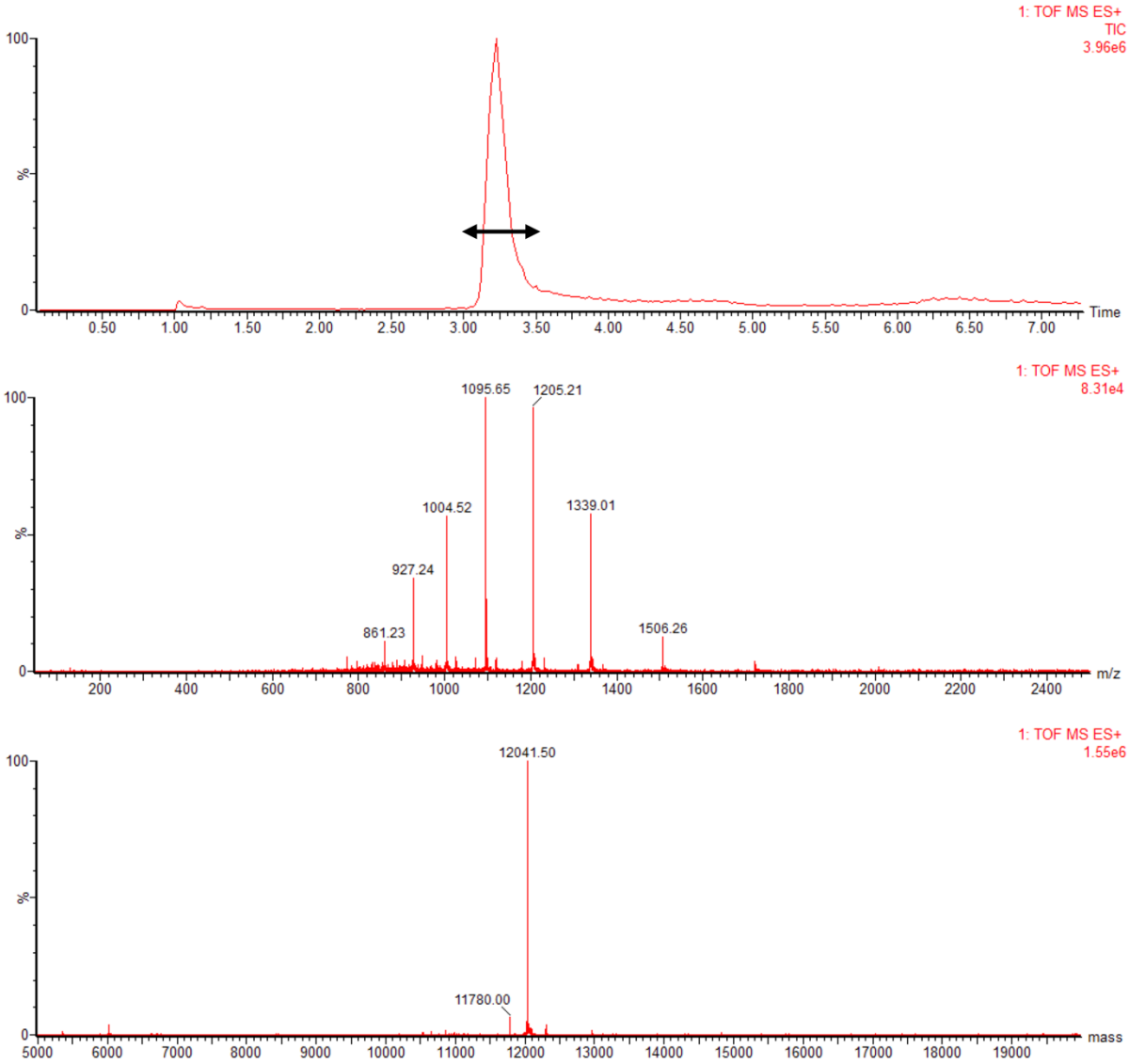


Figure B.1: A typical analysis of a metal protein conjugation reaction followed by LC–MS. The total ion chromatogram **(Top)**, combined ion series **(Middle)** and deconvoluted spectra **(Bottom)** are shown for the product of the reaction. Identical analyses were carried out for all the conjugation reactions performed in this work. Conversion rates were determined by integration of the peak area in liquid chromatogram, with the area integrated shown by the arrow in the above figure.

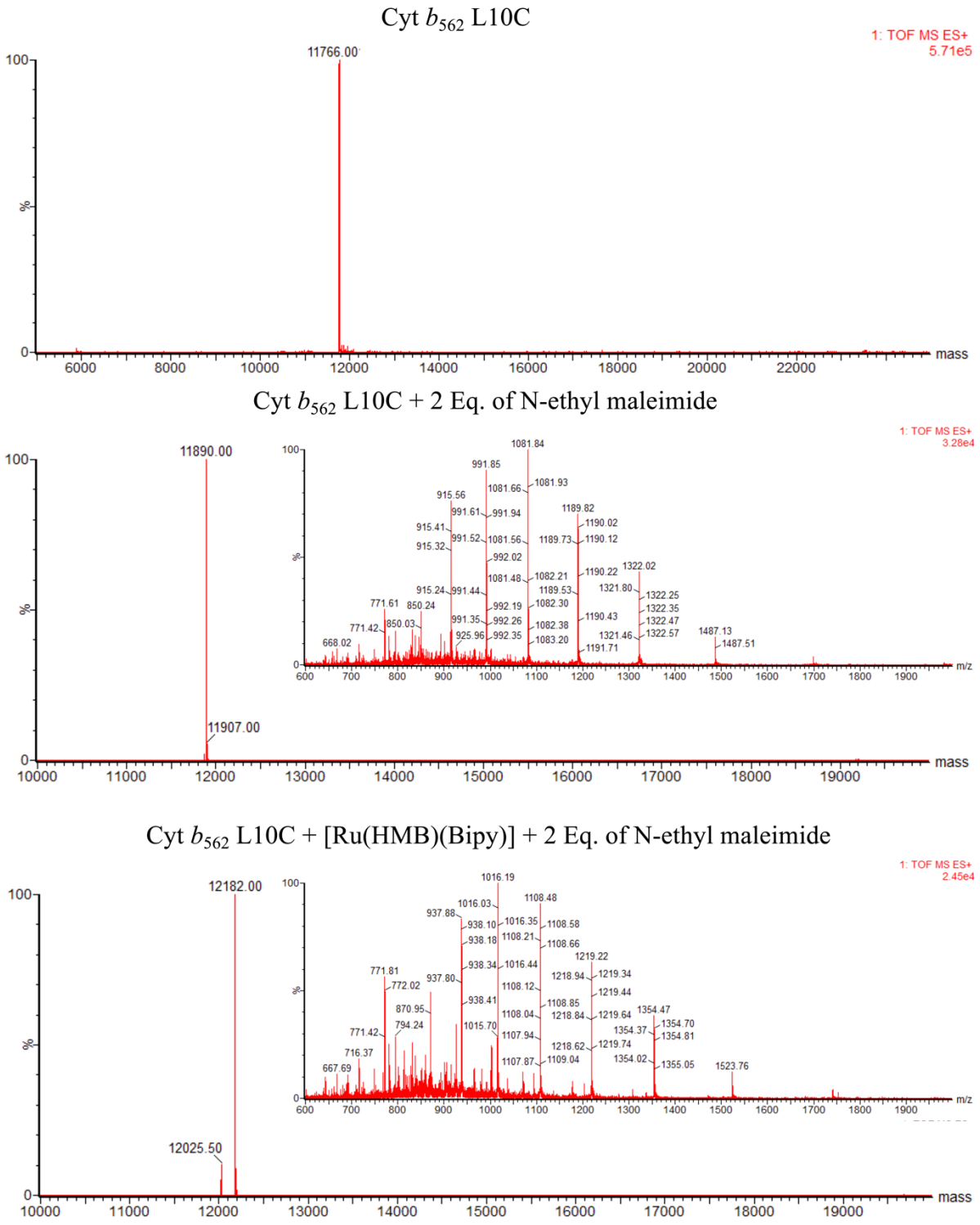


Figure B.2: (Top) Deconvoluted LC-MS spectra of Cyt b_{562} L10C (M_{obs} = 11766 Da, M_{calc} = 11764 Da)**. (Middle)** Deconvoluted and ion series spectra of an incubation of Cyt b_{562} L10C and N-ethyl maleimide (20 uM Prot, 2 Eq. NEM, 15 mins). The cysteine residue is modified by the maleimide – Cyt b_{562} L10C-(Mal) (M_{obs} = 11890 Da, M_{calc} = 11889 Da). **(Bottom)** Deconvoluted and ion series spectra of an incubation of Cyt b_{562} L10C + [Ru(HMB)(Bipy)] (M_{obs} = 12182 Da, M_{calc} = 12184 Da) and N-ethyl maleimide (20 uM Prot, 2 Eq. NEM, 15 mins). Upon addition of the maleimide, there is no change in mass to the ruthenium modified protein, indicating that the cysteine residue is blocked.

Appendix C – Nuclear Magnetic Resonance Supporting Information

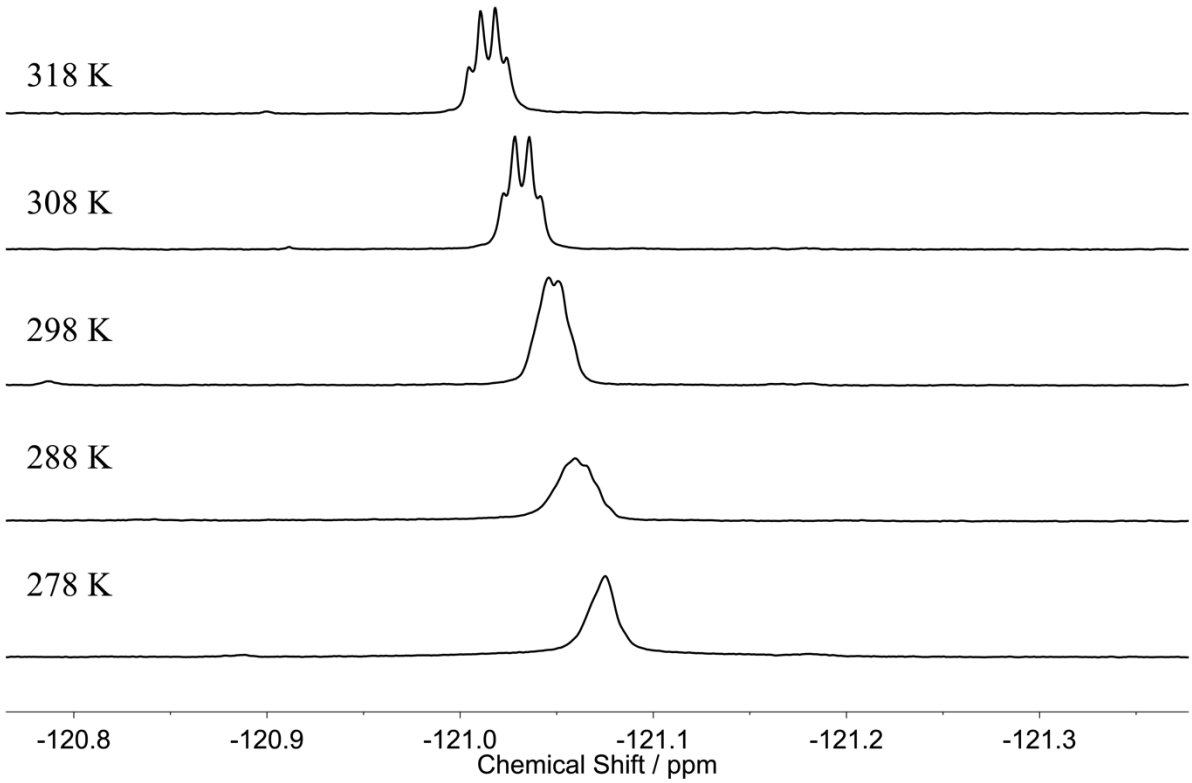


Figure C.1: The $^{19}F\{^{1}H\}$ NMR peaks of [Ru(η^6-benzene)(5,5'-difluorobipyridine)(N-ac-cysteine-OMe)]$^+$ when complex [1] is incubated with N-acetyl cysteine methyl ester (2 mM Ru, 3 eq amino acid) at a temperature range 278 K – 328 K.

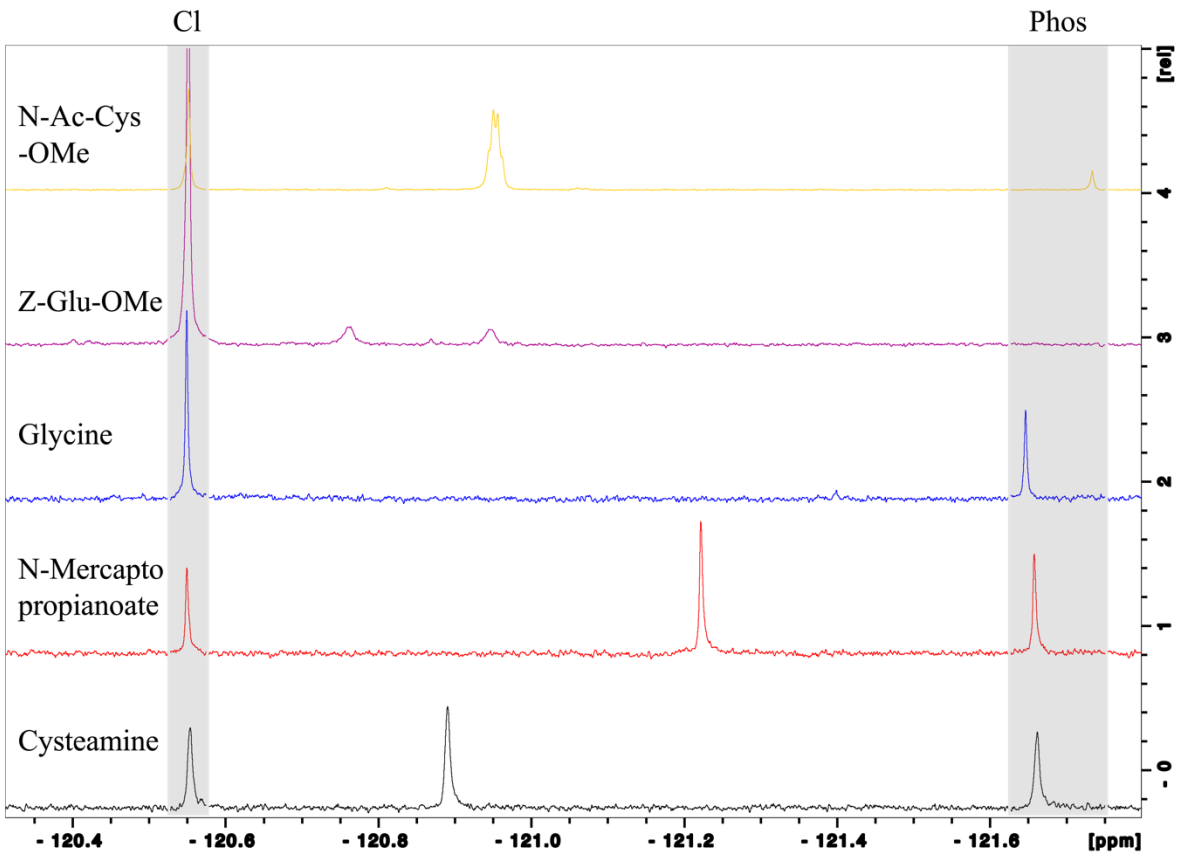


Figure C.2: The $^{19}F\{^{1}H\}$ NMR peaks observed when complex [9] is incubated with both chiral amino acids and achiral small molecules (2 mM Ru, 3 eq. AA, 8 hours, 310 K). The achiral species cysteamine, N-mercaptopropianoate and glycine give rise to singlet adduct peaks, whereas the chiral amino acids N-Ac-Cys-OMe (protected cysteine) and Z-Glu-OMe (protected glutamic acid) give rise to two peaks.

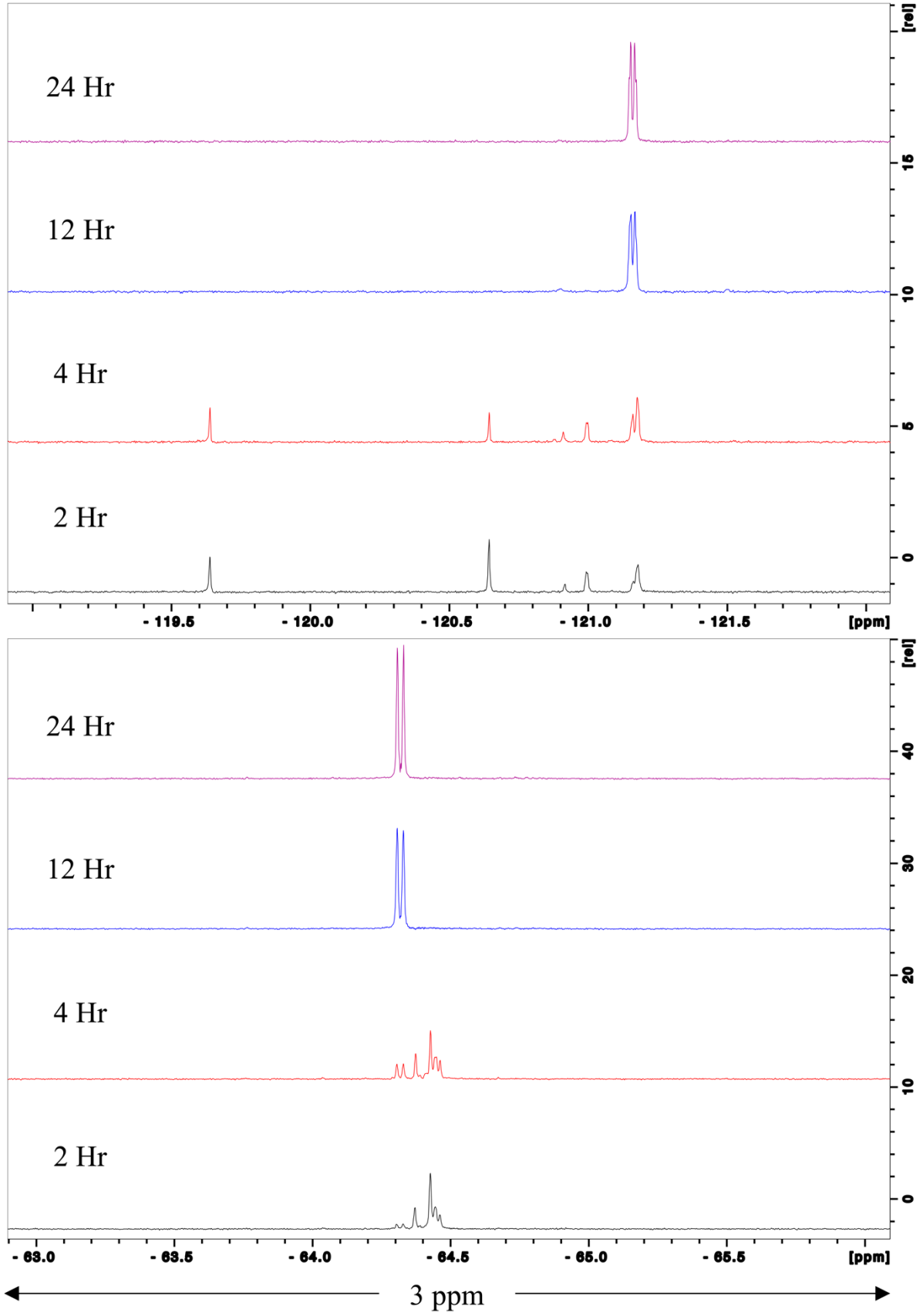


Figure C.3: A comparison between the incubations of complex [9] (Top) and [13] (Bottom) with N-Ac-Cys-OMe. This indicates the reduced chemical shift sensitivity changing the substituent on the bipyridyl ring from a fluorine to a trifluoromethyl group.