

***Multivalent scaffolds for use as protein surface
mimetics***

Sarah Helen Hewitt

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

School of Chemistry

May 2017

Intellectual Property and Publication Statement

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 1 is adapted from a review article 'Metal Complexes as Protein Surface Mimetics', S. H. Hewitt and A. J. Wilson, *Chem. Commun.*, 2016, 52, 9745-9756. The contributions from the authors were as follows: SHH (the candidate) drafted the review, AJW edited the manuscript into its present form.

Work presented in chapter 2 formed the basis of the research article 'Protein Surface Mimetics; Understanding How Ruthenium *Tris* (bipyridines) Interact with Proteins', S. H. Hewitt, M. H. Filby, E. Hayes, L. Kuhn, A. Kalverda, M. E. Webb and A. J. Wilson, *ChemBioChem*, 2017, 18, 223-231. The contributions from the author were as follows: SHH (the candidate), MHF, EH and AJW designed the research; SHH synthesised the compounds; MHF aided by MEW performed the PPI inhibition experiments; EH and MHF performed luminescence quenching assays; SHH performed the published luminescence quenching assays; SHH, LK and AK obtained the protein NMR spectra; SHH analysed the NMR spectra; SHH drafted the manuscript and AJW edited the manuscript into its present form.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2017 The University of Leeds and Sarah Helen Hewitt

Acknowledgements

Firstly I would like to thank my supervisor, Andy Wilson, for help and support throughout my PhD, and not giving up on the project that didn't seem to work until the end. I would like to thank the technical staff, Martin Huscroft for the LCMS and analytical HPLCs, Simon Barret for the NMR, Chris Empson for the plate reader and always being willing to help set up new methods and solve problems, Dave Foggarty for the UV/Vis and fluorometer and Nasir Khan for his help with the CD. Lars Kuhn and Arnout Kalverda for the protein NMR, I couldn't have done it without you both. Thanks also to Stuart Warriner for fixing the HRMS when it inevitably broke during a time course.

Thanks to the Wilson group past and present. To Anna for being the 'lab mum' when I started my PhD, and to Valeria for giving me your desk on my first day, I kept it to the end. To Dave, thank you for being there during the struggles during my first (second and third) years, you still have no idea how much you helped me, I am forever grateful. To George for all the stories, and to Kelly for being the other vegetarian when I started. To Jenny for being my go-to biologist long after you left the group. To Silvia, for adding a Spanish flavour to the group, and the Silvia-isms. To Racquel for the trips away, and of course the 2p machines. To Jayapal, we struggled on this project together, but you always kept a smile. To Georgina, my MChem, you made a Ru(II)(bpy)_3 complex and started the array work, which has now gone so much further. To Ravi, you stood next to me at the peptide fume cupboard and chatted to me while I struggled to make my PhD work. To Irene you showed me French people are a little crazy, we started our PhDs together, and both had some scientific difficulties, well done on finishing (first). To Claire for showing me French people are slightly less crazy (at least when sober). To Kris, you must have liked us, you kept on coming back, enjoy being in the group even longer. To Sam, for all the discussions on politics and religion. To Heather, for all the and for being so Yorkshire, even if I didn't always agree with the choice of radio station. To Sonja, for being a late night buddy, doing HPLC and thank you for carrying on some of the Ru(II)(bpy)_3 work. To Stacey thank you for getting us organised. To Phil for showing us it's never too late to do a PhD. To Zsofie for joining us at the pub, I hope you enjoy your time in chemistry. To Marcela, I haven't really got to know you but you have made the group a lot more international again. To Emma S for being the new Cambridge graduate and bringing baby T to see us. To Emma C for your unrelenting drive, even when I was working late, you seemed to be working later. To Kirsten, thanks for all the protein, even though I wasn't on the PoPPI project, good luck with your PhD in York. To the current post-doc team, Tom, for showing us you can go for a pint during the day and for knowing so much about synthesis, Martin, you only just joined

the group, but are a fountain of knowledge about everything. And to Kat for the assay knowledge and of course all the wine and pub trips, they made the final few months of my PhD so much more bearable.

To my family, to my Mum for always being down the phone, ready to listen, even if you didn't understand what I was doing. To my Dad and brothers for being there if/when I go home for some light relief. And finally to Harry, even if you didn't know (or perhaps even want to know) any chemistry, you stuck by me with it, even if I worked late or wanted to come to the lab at the weekend.

Abstract

The development of ligands for protein surfaces to inhibit protein-protein interactions (PPI)s is challenging, as protein surfaces often lack the clefts and pockets associated with traditionally druggable targets like enzyme active sites. One way in which protein surfaces can be targeted is by the use of protein surface mimetics, whereby a multivalent scaffold is functionalised with many binding groups on its periphery in order to achieve high affinity protein recognition. One such scaffold is a ruthenium(II) *tris* (bipyridine)s (Ru(II)(bpy)₃).

The work in this thesis aimed to further develop these Ru(II)(bpy)₃ protein surface mimetics; gaining information as to how they interact with proteins, looking at new ways of achieving high affinity protein surface recognition and the development of new applications for these molecules. In Chapter 2 an indepth study of the binding of two Ru(II)(bpy)₃ complexes to a model protein, cytochrome *c*, is presented, looking at the thermodynamic and electrostatic contributions to binding as well as using protein NMR to elucidate the binding site. In Chapter 3 the development of dynamic combinatorial chemistry (DCC) scaffolds based on Ru(II)(bpy)₃ complexes and tetraphenyl porphyrins was explored as a potential avenue for new receptor design, enabling the development of biologically compatible DCC systems, prime for protein ligand discovery. Chapter 4 presents another avenue for using the Ru(II)(bpy)₃ complexes; using an array approach to discriminate between different protein.

Table of Contents

Intellectual Property and Publication Statement	iv
Acknowledgements	v
Abstract	vii
Table of Contents.....	viii
List of Figures	xii
List of Tables.....	xvii
List of Schemes.....	xviii
List of Abbreviations.....	xix
1 Introduction.....	1
1.1 Protein-protein interactions.....	1
1.1.1 Modulation of protein-protein interactions	1
1.1.2 Conventional methods for development of protein surface ligands.....	2
1.1.2.1 High-throughput screening.....	2
1.1.2.2 Fragments.....	2
1.1.2.3 Secondary structure mimetics.....	3
1.2 The surface mimetic approach.....	4
1.2.1 Non-metallo protein surface mimetics	5
1.2.1.1 Calixarenes	5
1.2.1.2 Resorcinarenes.....	5
1.2.1.3 Porphyrins.....	7
1.2.1.4 Dendrimers	8
1.2.1.5 Other scaffolds.....	9
1.2.2 Metallo protein surface mimetics.....	10
1.2.2.1 Advantages of using metal complexes	10
1.2.2.2 Metal coordination to protein surfaces	11
1.2.2.3 Metalloporphyrins	14
1.2.2.4 Metallo-dendrimers	16
1.2.2.5 Transition metal complexes	17
1.2.2.5.1 Meggers Ruthenium(II) complexes	17
1.2.2.5.2 Inert Group 9 complexes.....	18
1.2.2.5.3 Metal tris (bipyridines).....	19
1.3 Project aims	24

2 Biophysical and Structural Studies on Cytochrome c Recognition with Functionalised Ruthenium(II) Tris (bipyridine) Complexes	25
2.1 Cyt c and the cyt c/CCP PPI.....	25
2.2 Ru(II)(bpy)₃ complexes for binding to cyt c.....	26
2.2.1 Synthesis	27
2.3 Assay development.....	29
2.3.1 Assay development from fluorometer to plate reader.	30
2.4 Binding of different Ru(II)(bpy)₃ complexes.....	31
2.4.1 UV/Vis ascorbate reduction assay	32
2.5 Cyt c/CCP PPI inhibition	33
2.6 Binding in different conditions	34
2.6.1 Temperature.....	34
2.6.2 Ionic strength.....	36
2.6.3 Binding in different buffers.....	38
2.6.4 pH.....	39
2.6.5 Conclusions of conditions screen	41
2.7 NMR spectroscopy	41
2.7.1 Oxidised/reduced cyt c.....	42
2.7.2 ¹ H- ¹⁵ N HSQC of cyt c alone.....	43
2.7.3 ¹ H- ¹⁵ N HSQC spectrum of Ru(II)(bpy) ₃ complex 31 with cyt c.....	43
2.7.4 ¹ H- ¹⁵ N HSQC spectrum of Ru(II)(bpy) ₃ complex 26 with cyt c.....	45
2.7.5 <i>S. cerevisiae</i> cyt c binding.....	46
2.8 Conclusions	47
3 Design of multivalent Ru(II)(bpy)₃ and porphyrin scaffolds for dynamic combinatorial chemistry	48
3.1 Biologically compatible DCC.....	49
3.2 Ru(II)(bpy)₃ scaffolds	50
3.2.1 Initial scaffold	50
3.2.1.1 Synthesis.....	51
3.2.1.2 Hydrazone formation.....	51
3.2.1.3 Hydrazone exchange	52
3.2.2 Ru(II)(bpy) ₃ scaffold developments	54
3.2.2.1 Glycine hydrazide Ru(II)(bpy) ₃ complex.....	55
3.2.3 New synthetic methodology.....	56
3.2.4 New Ru(II)(bpy) ₃ hydrazone scaffolds	59
3.2.4.1 Serine- and aspartic acid- hydrazide Ru(II)(bpy) ₃ complexes 54 and 55 synthetic attempts	60

3.2.4.2	Using hydrazido-amino acids to functionalise the Ru(II)(bpy) ₃ complexes.....	61
3.3	Porphyrin scaffolds	63
3.3.1	Porphyrin scaffold design and synthesis	64
3.3.2	Hydrazone exchanges.....	66
3.3.2.1	Hydrazone exchange to 1 aldehyde.....	68
3.3.2.1.1	Catalysts	71
3.3.2.2	Is this a true dynamic equilibrium or a static mixture?.....	74
3.3.2.3	Hydrazone exchanges to 2 aldehydes	75
3.3.2.3.1	Time to reach equilibrium with 2 aldehydes	78
3.3.3	Protein incubation	79
3.4	Conclusions	80
4	Using functionalised Ru(II)(bpy)₃ complexes in a protein sensing array	82
4.1	Protein sensing arrays	82
4.2	Discrimination of proteins using a range of Ru(II)(bpy)₃ complexes.....	83
4.2.1	Statistical analysis.....	86
4.2.1.1	Linear discriminant analysis on Ru(II)(bpy) ₃ complex array with a panel of proteins	87
4.3	Using 2 fluorophores to enhance protein discrimination.....	88
4.3.1	Deciding on appropriate concentrations	89
4.3.2	Discrimination of proteins	90
4.3.2.1	LDA of 6 protein, 2 fluorophore array.....	93
4.3.3	Including more therapeutically interesting proteins	95
4.4	Conclusions	97
5	Thesis summary and future work	99
6	Experimental section	101
6.1	Synthesis	101
6.1.1	General considerations	101
6.1.2	Synthetic protocols.....	102
6.2	Biophysical analyses- General considerations chapter 2.....	150
6.2.1	Luminescence quenching assays (Chapter 2).....	151
6.2.1.1	Fluorometer.....	151
6.2.1.2	Plate reader.....	151
6.2.2	UV/Vis ascorbate reduction.....	152
6.2.3	Protein NMR	152
6.3	Hydrazone Exchange Studies (Chapter 3)	152

6.3.1	Ru(II)(bpy) ₃ Complexes.....	153
6.3.1.1	Hydrazone exchange reactions.....	153
6.3.2	Porphyrins.....	153
6.3.2.1	Hydrazone exchange reactions.....	153
6.3.2.2	Protein incubation.....	153
6.4	Biophysical analyses- General considerations chapter 4.....	154
6.4.1	Ru(II)(bpy) ₃ complexes and proteins.....	154
6.4.2	Ru(II)(bpy) ₃ complexes, FITC-NOXA B (R-A) tracer and proteins.....	154
6.4.3	Linear discriminant analysis.....	155
6.5	General considerations for Appendix IV - DCC array studies with hydrazone functionalised Ru(II)(bpy)₃ complexes.....	155
6.6	General considerations for Appendix V – Stabilisation of the p53 tetramerisation domain	155
6.6.1	Circular dichroism	156
7	Appendices.....	157
7.1	Appendix I- Debye Hückel equation	157
7.2	Appendix II- Monosubstituted Ru(II)(bpy) ₃ complex synthesis.....	160
7.2.1	5' monosubstituted bipyridine synthesis.....	160
7.2.2	4' monosubstituted bipyridine complex synthesis.....	160
7.3	Appendix III- Principle component analysis and linear discriminant analysis.....	162
7.3.1	Statistical definitions	162
7.3.2	Principal component analysis.....	163
7.3.3	Linear discriminant analysis.....	164
7.4	Appendix IV- DCC array studies with hydrazone functionalised Ru(II)(bpy) ₃ complexes	166
7.5	Appendix V –Stabilisation of the p53 tetramerisation domain.....	170
7.5.1	The p53 tetramerisation domain	170
7.5.2	Ru(II)(bpy) ₃ complexes and peptides	171
7.5.3	Circular dichroism	172
7.5.3.1	In phosphate buffer	172
7.5.3.2	In water.....	173
7.5.4	Conclusions and future work.....	174
8	References.....	176

List of Figures

Figure 1.1 Targeting enzymes and PPIs.....	1
Figure 1.2 Nutlin 3a (1), an inhibitor of the <i>hDM2/p53</i> PPI discovered by high-throughput screening,.....	2
Figure 1.3 ABT-263 (2), an inhibitor of the <i>Bcl_{xL}/BAK</i> PPI, discovered by fragment based screening,.....	3
Figure 1.4 Protein crystal structure of a hydrocarbon-constrained BID peptide bound to Mcl-1 (PDB ID: 5C3F).....	4
Figure 1.5 The surface mimetic approach. A multivalent inhibitor binds to a large area of protein surface in order to inhibit a PPI.....	4
Figure 1.6 Calixarene protein surface binders.....	5
Figure 1.7 Uchiyama's tetrameric resorcinarene compound (4) for binding to histone surfaces.....	6
Figure 1.8 Porphyrin protein surface mimetics.....	7
Figure 1.9 Protein crystal structure of tetrasulfonatophenyl porphyrin bound to Jacalin (PDB ID: 1PXD).....	8
Figure 1.10 Schematic of a dendrimer, showing the core, branches and terminal (binding) groups.....	8
Figure 1.11 1st generation PAMAM dendrimer.....	9
Figure 1.12 Anthracene protein surface mimetic scaffold.....	10
Figure 1.13 Difference between sp ³ carbon and octahedral metal scaffolds.....	11
Figure 1.14 Cu(II)-IDA for binding to protein surfaces,.....	12
Figure 1.15 <i>bis</i> (Zn(II)-(dpa)) species bound to two phospho-amino acid residues, stabilising the α -helical conformation.....	13
Figure 1.16 Metalloporphyrins for binding to <i>cyt c</i>	15
Figure 1.17 Zn(II) dendrimers that bind to <i>cyt c</i>	16
Figure 1.18 Meggers Ru(II) based protein kinase ligands.....	17
Figure 1.19 Inert group 9 protein surface ligands.....	18
Figure 1.20 Irreversible group 9 metal complex protein surface binders.....	19
Figure 1.21 Schematic of DCC based on Fe(II)(bpy) ₃ complexes.....	20
Figure 1.22 Seeberger's mannose functionalised Ru(II)(bpy) ₃ complex for ConA/GNA binding.....	21
Figure 1.23 Hamachi's carboxylate functionalised Ru(II)(bpy) ₃ complexes for <i>cyt c</i> binding and photoreduction.....	22
Figure 1.24 Ru(II)(bpy) ₃ complexes designed by the Wilson and Ohkanda groups for binding to <i>cyt c</i> (26a , 27a) and α -ChT (26b).....	23

Figure 2.1 The cyt <i>c</i> /CCP PPI structure (PDB ID 1U75). ¹⁴⁸	25
Figure 2.2 Ru(II)(bpy) ₃ complexes designed for binding to cyt <i>c</i>	26
Figure 2.3 Cartoon illustrating how the luminescence quenching assay works	29
Figure 2.4 Luminescence quenching assay	30
Figure 2.5 Binding of Ru(II)(bpy) ₃ complex 26 to cyt <i>c</i> on plate reader in 5 mM sodium phosphate, pH 7.5 a) without BSA additive, b) with 0.2 mg mL ⁻¹ BSA	31
Figure 2.6 Ascorbate (0.75 mM) reduction of cyt <i>c</i> (16 μM) in the presence of various Ru(II)(bpy) ₃ complexes (16 μM)	32
Figure 2.7 Cyt <i>c</i> /CCP PPI inhibition by Ru(II)(bpy) ₃ complex 26 ,	33
Figure 2.8 van't Hoff analysis for the binding of Ru(II)(bpy) ₃ complexes 26 and 31 measured in 5 mM sodium phosphate, 0.2 mg mL ⁻¹ BSA, pH 7.5	35
Figure 2.9 Debye-Hückel analysis for the binding of Ru(II)(bpy) ₃ complexes 26 and 31 to cyt <i>c</i> and literature values for CCP for comparison	37
Figure 2.10 Buffers	38
Figure 2.11 Effect of pH on the binding of Ru(II)(bpy) ₃ complex 26 to cyt <i>c</i>	40
Figure 2.12 NMR spectra of cyt <i>c</i> alone	43
Figure 2.13 ¹ H- ¹⁵ N HSQC NMR data of Ru(II)(bpy) ₃ complex 31 binding to cyt <i>c</i>	44
Figure 2.14 NMR spectra for the binding of Ru(II)(bpy) ₃ complex 26 to cyt <i>c</i>	45
Figure 2.15 Binding isotherms for a) Ru(II)(bpy) ₃ complex 31 and b) Ru(II)(bpy) ₃ complex 26 with cyt <i>c</i> from <i>S. cerevisiae</i>	46
Figure 3.1 Cartoon schematic of DCC around a multivalent scaffold	48
Figure 3.2 Hydrazone exchanges on the initial Ru(II)(bpy) ₃ hydrazone scaffold 46 ..	54
Figure 3.3 Difference between linear and divergent synthesis	56
Figure 3.4 Serine (54) and Aspartic/Glutamic acid (55 and 56) hydrazide Ru(II)(bpy) ₃ complexes, with the hydroxyl/carboxylate groups potentially solubilising the core DCC scaffold	59
Figure 3.5 Porphyrin scaffold	63
Figure 3.6 Aldehydes used in hydrazone exchange reactions	68
Figure 3.7 Exemplary HRMS showing hydrazone exchange on a porphyrin scaffold (79b)	69
Figure 3.8 Initial hydrazone exchange time courses on porphyrin hydrazone scaffolds 79	70
Figure 3.9 Effect of catalysis on porphyrin hydrazone exchange, using exemplary reaction with 4-carboxy benzaldehyde 44c on all 3 porphyrin hydrazone scaffolds 79 and addition of no catalyst, aniline 45 and anthranillic acid 80	72
Figure 3.10 Effect of catalyst concentration on hydrazone exchanges. E	73

Figure 3.11 Effect of catalyst concentration on serine hydrazone porphyrin scaffold 79a	74
Figure 3.12 Cartoon depicting changing distribution of species on addition of second aliquot of aldehyde	74
Figure 3.13 Addition of 2 separate batches of aldehyde, incubation with 25 eq. of aldehyde 44c for 24 hours followed by addition of a second 25 eq. of aldehyde 44c . 1 st incubation of 100 μ M hydrazone porphyrin 79 with 2.5 mM 4-carboxy benzaldehyde 44c and 10 mM aniline 45 in 10 % DMSO in 5 mM NH ₄ OAc, pH 6.75. 2 nd incubation, addition of a further 2.5 mM 4-carboxy benzaldehyde 44c to first incubation. Graphs show an average of 5 separate measurements.	75
Figure 3.14 Cartoon depicting 2 methods of performing hydrazone exchange reaction with 2 aldehydes	76
Figure 3.15 Single time point 2 aldehyde hydrazone exchanges, using 2,4-dimethoxy benzaldehyde 44b and 4-carboxy benzaldehyde 44c . HRMS taken after 24 hours incubation, pre-incubated systems incubated with the first aldehyde for 24 hours before addition of the second aldehyde.	77
Figure 3.16 Time for porphyrin hydrazone exchanges to reach equilibrium with two aldehyde	78
Figure 3.17 Cartoon depicting incubation with protein. Different distributions of porphyrin hydrazone products are obtained in the presence of a protein compared to its absence.....	79
Figure 4.1 Low affinity binding of an analyte (e.g. smell molecule) to a range of receptors, generates a fingerprint response.....	82
Figure 4.2 The Ru(II)(bpy) ₃ complexes used in the protein sensing arrays	83
Figure 4.3 Cartoon illustrating differential sensing with Ru(II)(bpy) ₃ complexes and proteins.....	84
Figure 4.4 X-ray crystal structures of the nine proteins for use in the array, showing the basic (blue) and acidic (red) amino acid residues, molecular weights (MW) and isoelectric points (pI) of the proteins.....	85
Figure 4.5 Differential luminescence responses from the different Ru(II)(bpy) ₃ complexes 26 , 29-32 , 34 , 35 and 81 (2.5 μ M) on incubation with various different proteins (10 μ M).....	85
Figure 4.6 PCA/LDA aims to take the 8 dimensional Ru(II)(bpy) ₃ complex array data down to two (or three) dimensions by taking linear combinations of the original data, thus allowing the data to be plotted on a 2-D (or 3-D) scatter graph, with the aim of getting clusters for the different analytes (proteins)	86

Figure 4.7 LDA for the Ru(II)(bpy) ₃ complex-protein array, a) 2-D analysis, b) 2 views of the 3-D analysis.....	88
Figure 4.8 Array of potential different interactions between the two fluorophores and different proteins, including interactions between the two luminescent molecules, binding on different sites on the protein, masking of binding sites for one molecule by the other molecule, docking of the two molecules together on the protein	89
Figure 4.9 Concentration test to determine appropriate concentration of FITC-NOXA B peptide to use with 2.5 μM Ru(II)(bpy) ₃ complex. Luminescence emission intensity at variable wavelengths (exc. 467 nm) upon incubation of 2.5 μM Ru(II)(bpy) ₃ complex 29 with various concentrations of FITC-NOXA B peptide.....	90
Figure 4.10 Luminescence response (exc. 467 nm, in 5 mM sodium phosphate, pH 7.5, 2 hour incubation) of the FITC-NOXA B peptide alone (a), Ru(II)(bpy) ₃ complex 26 alone (b), FITC-NOXA B with Ru(II)(bpy) ₃ complex 26 (c) and FITC NOXA B with Ru(II)(bpy) ₃ complex 30 (d), showing differences in the spectra	91
Figure 4.11 Fingerprint luminescent responses for the 6 different proteins on incubation of Ru(II)(bpy) ₃ complexes alone (a) and b)), Ru(II)(bpy) ₃ complexes with FITC-NOXA B peptide and the FITC-NOXA B peptide (c), d), e) and f)), looking at both the fluorescein (emission 520 nm) (e) and f)) and Ru(II)(bpy) ₃ channels (emission 630 nm) (c), d), e) and f)) on excitation at 467 nm where appropriate. b), d) and f) 20 hour incubation.....	92
Figure 4.12 2-D LDA for the two fluorophore array with 6 proteins, 93	
Figure 4.13 3-D LDA of 2 fluorophore array data and Ru(II)(bpy) ₃ complex array for 6 proteins (combined 2 hour and 24 hour incubation data).....	94
Figure 4.14 Confidence ellipsoids for the LDA of the 6 protein, 2 fluorophore array a) 95 % confidence ellipsoids, b) 80 % confidence ellipsoids.....	95
Figure 4.15 3-D LDA of 2 fluorophore array data after addition of Mcl-1 and hDM2 incubations.....	96
Figure 4.16 3-D LDA of 2 fluorophore array after removal of the cyt <i>c</i> data from the analysis	96
Figure 4.17 Confidence ellipsoids for the LDA of 2 fluorophore-protein array, excluding the cyt <i>c</i> data.....	97
Figure 7.1 Cartoon depiction of DCC array	166
Figure 7.2 Aldehydes for use in the DCC array.....	167
Figure 7.3 Fingerprint bar charts for the DCC arrays with the glycine hydrazide Ru(II)(bpy) ₃ complex 43 and 10 aldehydes, using 2 equivalents (a) and 6 equivalents (b) of aldehyde to hydrazide Ru(II)(bpy) ₃ complex 43	168

Figure 7.4 2-D LDA of the DCC array data with the glycine hydrazide Ru(II)(bpy) ₃ complex 43 and 10 aldehydes 44	169
Figure 7.5 X-ray crystal structure of the p53 tetramerisation domain, highlighting residue 337.....	170
Figure 7.6 Cartoon depicting hypothesised binding of the Ru(II)(bpy) ₃ complexes to the p53 tetramerisation domain, resulting in stabilisation of the tetrameric structure....	171
Figure 7.7 Ru(II)(bpy) ₃ complexes and p53 tetramerisation domain peptides.....	171
Figure 7.8 CD spectra in 5 mM sodium phosphate buffer, pH 7.5	172
Figure 7.9 CD spectra in water	174

List of Tables

Table 2.1 Deprotection conditions for the Ru(bpy) ₃ complexes	29
Table 2.2 Difference in cyt <i>c</i> binding affinity between the plate reader and fluorometer	31
Table 2.3 Binding affinities for the different Ru(bpy) ₃ complexes synthesised to cyt <i>c</i> , 5 mM sodium phosphate, 0.2 mg mL ⁻¹ BSA, pH 7.5	32
Table 2.4 Thermodynamic parameters derived from the van't Hoff analysis for the binding of the Ru(bpy) ₃ complexes 26 and 31 to cyt <i>c</i>	35
Table 2.5 Binding of Ru(bpy) ₃ complexes 26 and 31 in variable ionic strengths,	36
Table 2.6 Values derived from the Debye-Hückel approximation for Ru(bpy) ₃ complexes 26 and 31 and CCP ¹⁵⁴	37
Table 2.7 Cyt <i>c</i> binding of Ru(bpy) ₃ complex 26 in a range of different buffers	39
Table 2.8 Binding constants for Ru(bpy) ₃ complex 26 binding to oxidised and reduced cyt <i>c</i>	42
Table 3.1 Examples of biologically compatible DCC reactions ¹⁶⁶	49

List of Schemes

Scheme 2.1 Synthesis of Ru(II)(bpy) ₃ complexes 26, 29-35	28
Scheme 3.1 Hypothesised initial DCC system	50
Scheme 3.2 Synthesis of the initial Ru(II)(bpy) ₃ complex hydrazide scaffold 43	51
Scheme 3.3 Hydrazone formation on initial hydrazide Ru(II)(bpy) ₃ complex 43	51
Scheme 3.4 Radical cleavage of N-N bond in mass spectrometer	52
Scheme 3.5 Hydrazone exchange a) Reaction to be performed, b) Cartoon schematic of reaction, c) Cartoon schematic of individual hydrazone exchanges	53
Scheme 3.6 Synthesis of glycine hydrazide and hydrazone Ru(II)(bpy) ₃ complexes 49 and 50	55
Scheme 3.7 Synthesis of acid functionalised Ru(II)(bpy) ₃ complex 29	57
Scheme 3.8 Attempts at amide bond formation on intact Ru(II)(bpy) ₃ complexes 29 and 30 using peptide coupling agents a) 29 b) 30	57
Scheme 3.9 Amide bond formation on an intact Ru(II)(bpy) ₃ complex	58
Scheme 3.10 Synthesis of orthogonally protected serine and aspartic acid hydrazide Ru(II)(bpy) ₃ complexes 41j and k	60
Scheme 3.11 Attempts at hydrazide Ru(II)(bpy) ₃ complex formation on serine and aspartic acid Ru(II)(bpy) ₃ complexes 41j and k	61
Scheme 3.12 Synthesis of serine, aspartic acid and glutamic acid hydrazido-amino acids 62	61
Scheme 3.13 Attempts at amide bond formation between intact acid functionalised Ru(II)(bpy) ₃ complexes 29, 64 and 65 and the hydrazido amino acids 62	62
Scheme 3.14 Attempts at synthesising serine hydrazide Ru(II)(bpy) ₃ scaffolds 63a and 67a by first synthesising ligands 68 and 70 ,	63
Scheme 3.15 Initial attempts at making porphyrin hydrazide scaffolds 73 and 75 a) directly attached to porphyrin 73 , b) glycine hydrazide porphyrin 75	64
Scheme 3.16 First attempted synthesis of serine hydrazide porphyrin 77	65
Scheme 3.17 Synthesis of serine, aspartic acid and glutamic acid hydrazide and benzaldehyde hydrazone porphyrins 78 and 79	66
Scheme 3.18 Porphyrin hydrazone exchanges a) General reaction scheme, b) Cartoon representation of reaction, c) Cartoon representation of individual exchanges	67
Scheme 7.1 Synthesis of 5' monosubstituted bipyridine Ru(II)(bpy) ₃ complex 65 ...	160
Scheme 7.2 4' monosubstituted Ru(II)(bpy) ₃ complex synthesis using Kroehnke methodology	161
Scheme 7.3 DCC with glycine hydrazide complex scaffold	167

List of Abbreviations

α-ChT	α -chymotrypsin
ATP	adenosine triphosphate
Bak	Bcl-2 homologous Antagonist Killer protein
Bcl_{XL}	B-Cell Lymphoma 'extra large' anti-apoptotic protein
biq	2,2'-biquinoline
Boc	<i>tertiary</i> -butyl carbamate
bpy	bipyridine
BSA	bovine serum albumin
CCP	cytochrome <i>c</i> peroxidase
CD	circular dichroism
cyt	cytochrome
DA	discriminant analysis
DCC	dynamic combinatorial chemistry
DCL	dynamic combinatorial library
DIPEA	<i>diisopropylethylamine</i>
DNA	deoxyribonucleic acid
DMF	dimethylformamide
DMSO	dimethylsulfoxide
dpa	dipicolylamine
EGF	endothelial growth factor
EGFR	endothelial growth factor receptor
eq.	equivalents
ESI	electro-spray ionisation
FITC	fluorescein isothiocyanate
Fmoc	9-fluorenylmethyl oxycarbonyl
FP	fluorescence polarisation
FRET	Förster resonance energy transfer

HATU	1-[<i>Bis</i> (dimethylamino)methylene]-1H-1,2,3- triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HCTU	N,N,N',N'-Tetramethyl-O-(6-chloro-1H-benzotriazol-1- yl) uranium hexafluorophosphate
hDM2	human double minute 2
His	histidine
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSA	human serum albumin
HSQC	heteronuclear single quantum correlation
IC₅₀	Half maximal inhibitory concentration
IDA	iminodiacetate
IR	infrared (spectroscopy)
ITC	Isothermal titration calorimetry
K_a	association constant
K_d	dissociation constant
K_i	inhibition constant
LCMS	liquid chromatography mass spectrometry
LDA	linear discriminant analysis
Mcl-1	induced myeloid leukemia cell differentiation protein
MD	molecular dynamics
MW	molecular weight
MWCO	molecular weight cut off
NMR	nuclear magnetic resonance
p53	(tumour) protein 53
PAMAM	poly(amidoamine)
PCA	principal component analysis
PDB	protein data bank
PDGF	platelet derived growth factor
PET	photo-induced electron transfer

phen	phenanthroline
PNA	peanut agglutinin
PPI	protein-protein interaction
ppy	phenylpyridinato
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
QDA	quadratic discriminant analysis
RNase	ribonuclease
SAR	structure-activity relationship
SAXS	small angle x-ray scattering
SH2	<i>src</i> homology domain 2
SPR	surface plasmon resonance
STAT	signal transducer and activator of transcription
TFA	trifluoroacetic acid
TIPS	tri <i>is</i> opropyl silane
TLC	thin-layer chromatography
T_m	melting temperature
TNF-α	tumour necrosis factor α
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor

1 Introduction

This chapter is adapted from 'Metal Complexes as Protein Surface Mimetics', S. H. Hewitt and A. J. Wilson, *Chem. Commun.*, 2016, 52, 9745-9756

1.1 Protein-protein interactions

Protein-protein interactions (PPIs) are ubiquitous within biology, being used for a plethora of different processes, including signal transduction, antibody responses, the formation of complex structures, apoptosis and the control of the cell cycle.¹ As a result of their abundant nature, the modulation (inhibition and stabilisation) of PPIs is of great interest both therapeutically and for the study of biological processes.

Many studies have been performed in order to establish the topology and chemical nature of the protein-protein interface, but these have led to many discrepancies, over generic conclusions such as the hydrophobicity and charge content of the interface.² However, from these studies it has been established that there is a tendency for the interface to be mainly hydrophobic, like the protein core, but to contain a relatively larger proportion of charged and polar amino acid residues.²

1.1.1 Modulation of protein-protein interactions

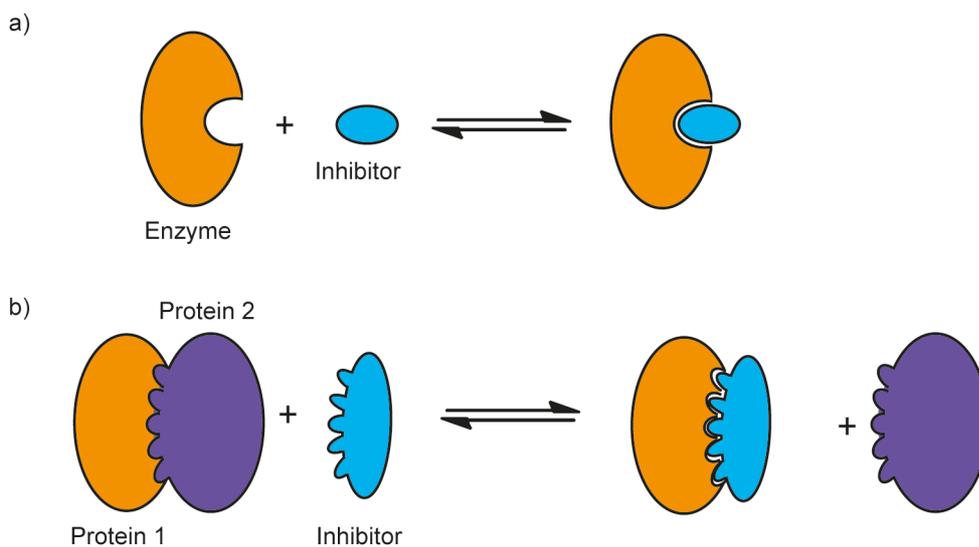


Figure 1.1 Targeting enzymes and PPIs a) Targeting a clearly defined enzyme pocket with a small molecule inhibitor, b) Targeting a PPI over a much larger surface area

The nature of protein interfaces makes targeting PPIs notoriously difficult (**Figure 1.1**); indeed they have been described as 'undruggable', or 'high-hanging fruit'.³ In order to

target a PPI, a small molecule that binds to a protein surface is required, yet protein surfaces are often large (~1500 – 3000 Å²) and flat, lacking the clefts and pockets associated with more conventional protein targets, like enzyme active sites or G protein-coupled receptors, making it difficult to discover small molecules capable of protein surface recognition.³

1.1.2 Conventional methods for development of protein surface ligands

Various different approaches have been used to target PPIs. Some successes having been found with peptides and peptidomimetics,⁴ along with fragment based approaches, high-throughput screening and virtual screening.^{5,6} However PPIs still remain a difficult target for small molecule ligands.

1.1.2.1 High-throughput screening

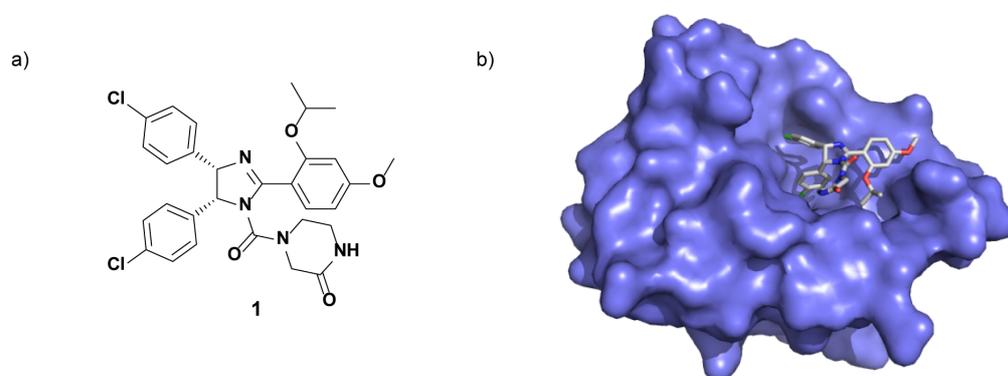


Figure 1.2 Nutlin 3a (**1**), an inhibitor of the *hDM2*/p53 PPI discovered by high-throughput screening, a) Molecular structure of Nutlin 3a (**1**), ii) Protein crystal structure of Nutlin 3a (**1**) bound to *hDM2* (PDB ID: 4HG7)

The screening of large libraries of drug-like molecules, followed by further optimization has yielded the *Nutlins* as inhibitors of the p53/*hDM2* interaction (**Figure 1.2**),⁷ the most potent of which (Nutlin 3 (**1**), IC₅₀ 0.09 μM) has been shown to inhibit tumourigenesis in human xenograft cells *in vivo*.⁸ However these high throughput approaches have yielded few results, potentially as the libraries being screened are comprised of molecules ‘designed’ for more traditional drug targets, like G-coupled receptors and enzyme active sites.³

1.1.2.2 Fragments

Fragment based approaches involve the screening of small (<300 Da), relatively polar molecules for binding to proteins which can then be further elaborated by linking together separate fragments or growing from one fragment into a larger more drug-like molecule. Fragment based approaches have been used, by the Fesik group, to develop ABT-263 (**2**),

an inhibitor of the Bcl_{xL}/BAK PPI (**Figure 1.3**),⁹ a drug candidate which entered phase II clinical trials. However there have been reported problems with the fragment based approach as applied to PPIs, as fragments required for the targeting of PPIs, are often large compared to that for the more conventional drug targets.¹⁰

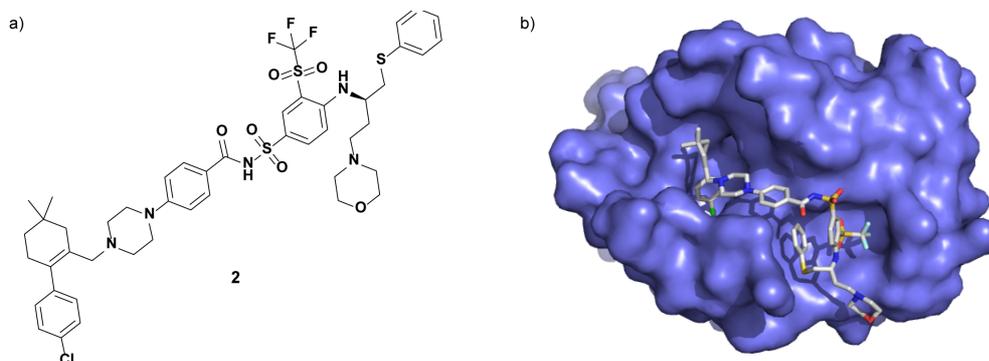


Figure 1.3 ABT-263 (**2**), an inhibitor of the Bcl_{xL}/BAK PPI, discovered by fragment based screening, a) ABT-263 molecular structure, b) Protein crystal structure of ABT-263 bound to Bcl_{xL} (PDB ID: 4LVT)

1.1.2.3 Secondary structure mimetics

Much work involves the design of mimics of protein secondary structure, mimicking β -turns,¹¹ α -helices¹² and β -sheets.¹³ These allow the mimicry of the secondary structural elements of one protein binding partner, delivering a molecule able to bind to the other partner, thus inhibiting the PPI. This has potential as a general approach to small molecule structural design, allowing the development of inhibitors for many PPIs based on similar secondary structural interactions, utilizing a single molecular scaffold.

An example of an α -helical secondary structure mimic is the use of constrained peptides (**Figure 1.4**). In this approach a sequence of α -amino acids is preorganised into a helical conformation by, for example, the addition of a hydrocarbon constraint.¹⁴ Theoretically this reduces the entropy loss on binding from that of the unconstrained peptide, potentially increasing the binding affinity,¹⁵ as there is now no barrier to the formation of the bound helical conformation. The introduction of the constraint also increases the stability of the peptide to proteolysis¹⁴ and can increase the cell permeability.¹⁶

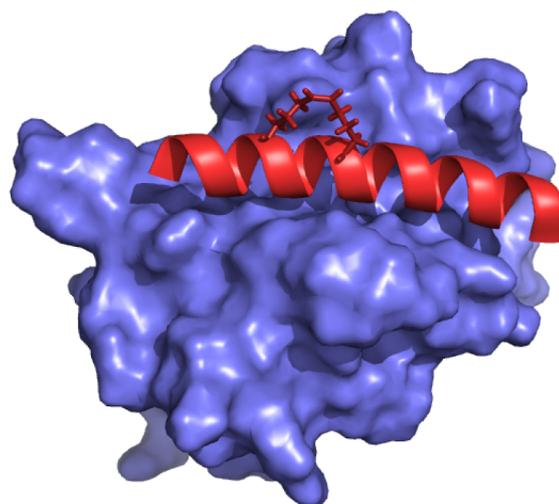


Figure 1.4 Protein crystal structure of a hydrocarbon-constrained BID peptide bound to Mcl-1 (PDB ID: 5C3F)

1.2 The surface mimetic approach

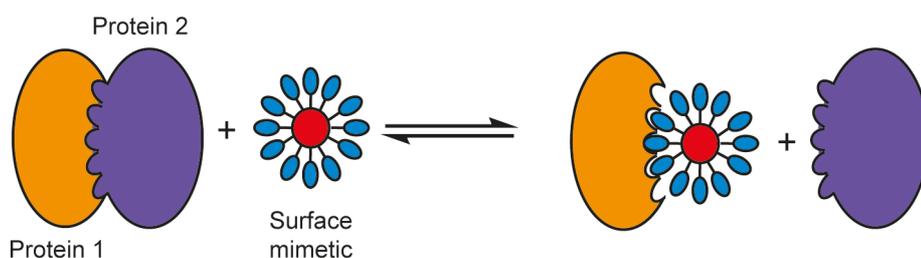


Figure 1.5 The surface mimetic approach. A multivalent inhibitor binds to a large area of protein surface in order to inhibit a PPI

The surface mimetic approach utilises a large, often supramolecular, scaffold to project binding groups over a large area of protein surface (**Figure 1.5**),¹⁷⁻¹⁹ to achieve high affinity protein binding. This allows for the recognition of large areas of protein surface, with fewer discernable features than can be recognised, for example by a secondary structure mimetic. The binding utilizes multivalency, the interaction of multiple binding groups located around a central scaffold on a host molecule (surface mimetic), towards multiple recognition sites on a receptor molecule (protein) in order to achieve high affinity binding.¹⁹ Multivalency is widely used in nature,²⁰ where it allows for an increased binding affinity by increasing the number of ligand and receptor sites, for example in signal transduction, cell membrane adherence, and immunological responses.

1.2.1 Non-metallo protein surface mimetics

Many conventional organic supramolecular scaffolds have been developed for binding to protein surfaces. These include calixarenes, porphyrins, anthracenes, cyclodextrins, resorcinarenes and dendrimers.¹⁹

1.2.1.1 Calixarenes

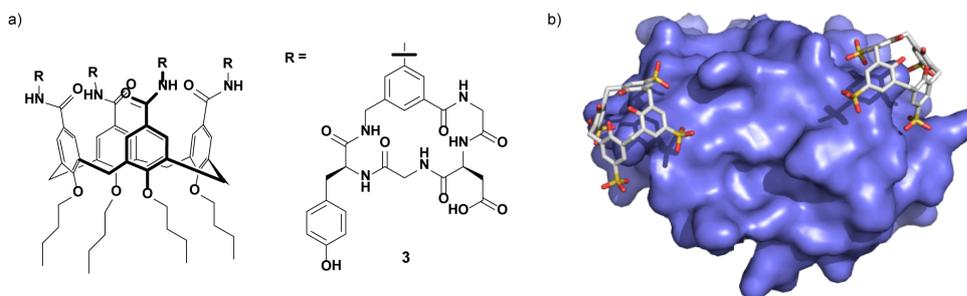


Figure 1.6 Calixarene protein surface binders a) GFB-111 (**3**), a PDGF ligand, b) Protein crystal structure of two *p*-sulfonatocalix[4]arenes bound to the surface of cyt *c* (PDB ID: 3TYI)

Calixarenes are cone-like molecules with two distinct edges that can be functionalised with recognition elements for protein surface recognition.²¹ The Hamilton group have synthesised calix[4]arene derivatives which bind to cytochrome (cyt) *c*, α -chymotrypsin (α -ChT) and platelet-derived growth factor (PDGF), acting as antibody mimics.^{22–25} GFB-111 (**3**) (**Figure 1.6a**), a PDGF binder with IC_{50} 250 nM, was shown to be functional *in vivo* (mouse).²⁵ Crowley and coworkers, more recently, solved a crystal structure of a *p*-sulfonatocalix[4]arene bound to cyt *c* at three different sites (**Figure 1.6b**), with the calixarenes acting as mediators of the PPIs required for crystallisation.²⁶

The Neri group also generated calix[4]arene derivatives, functionalised with tetrameric peptides, to bind to and inhibit the acyl transfer enzyme transglutaminase (with up to 62 % reduction in activity), by blocking the entrance of the substrate to the active site.²⁷ The Hof group used calixarenes along with readily-available dyes to form dye-displacement sensors for anti-body free reading of histones through lysine side chain recognition.²⁸

1.2.1.2 Resorcinarenes

Uchiyama and coworkers developed the use of resorcinarene scaffolds for histone surface recognition (**Figure 1.7**).^{29–32} Histones are basic proteins, containing many lysine residues, so the development of the resorcinarene compounds with many carboxylic acids aids their binding. They first developed compounds with 8 (monomeric) and 28 (tetrameric) surrounding carboxylates (**4**),²⁹ followed by a more extended scaffold with 84

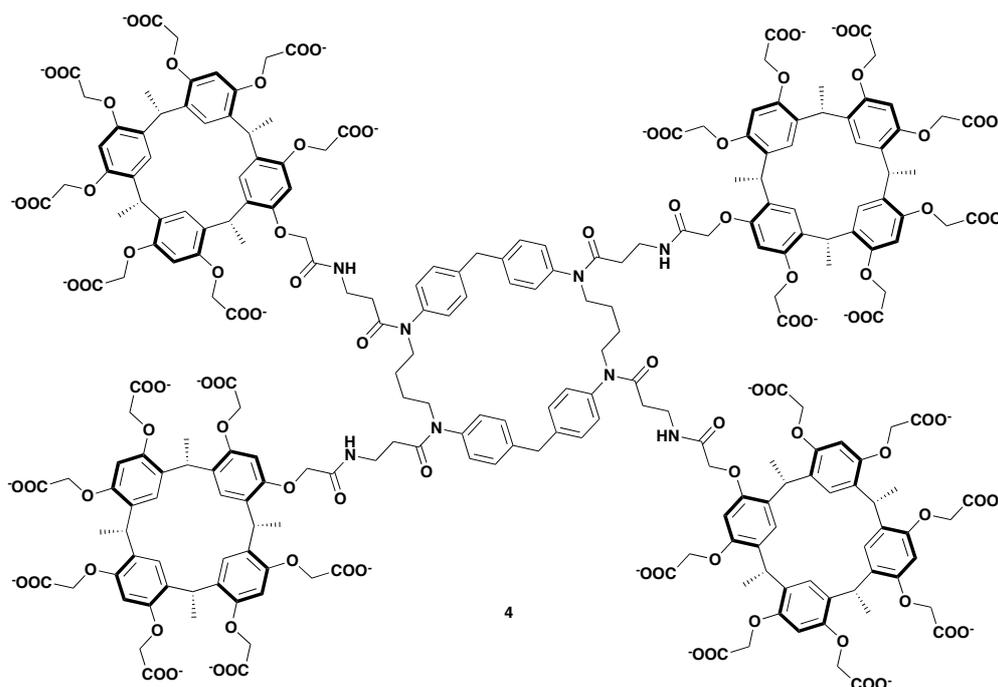


Figure 1.7 Uchiyama's tetrameric resorcinarene compound (**4**) for binding to histone surfaces

carboxylates,³⁰ which were shown to bind to mixed histone proteins with K_a 4.2×10^5 M⁻¹, 1.3×10^7 M⁻¹, and 8.4×10^7 M⁻¹ respectively by a surface plasmon resonance (SPR) assay. They were also shown to be selective for histones over lysozyme and ovalbumin, with binding increasing with increasing numbers of carboxylates.³⁰ They later, transformed their tetrameric (28 carboxylate) ligand (**4**) into one that could be used in fluorescence intensity assays, by changing one of the resorcinarene units to a fluorescent dansyl group. The fluorescence of this dansyl group showed a 5-fold increase in fluorescence intensity on saturation with histone, and indicated the dansyl moiety as being located on a non-polar region of the histone surface.³¹ The binding affinity was found to be in the region of 10^6 M⁻¹ with 1:1 binding demonstrated by this fluorescence intensity assay, and SPR. The K_a was found to be higher at lower ionic strengths, thus implicating electrostatics as a major driving force for binding. The binding to ~70 % acetylated histone was also shown to be negligible.³¹ Further modifications to the tetrameric ligand (**4**) yielded a rotaxane wheel for a 2,6-disubstituted naphthalene thread with two fluorophore (fluorescein and rhodamine) moieties on either end.³² These showed similar fluorescence intensity profiles to the dansyl substituted compounds with K_a 2.3×10^6 M⁻¹ and 9.0×10^6 M⁻¹ respectively, and large polarisation values.³² The compounds were similarly found to be selective over a variety of other proteins and were used in initial studies for FRET detection of histones.³²

The Botta group have also used dipeptide substituted resorc[4]arenes for binding to human serum albumin (HSA) and α -ChT.³³ With HSA there was a change in the circular dichroism (CD) spectrum for the resorcinarene but not for the protein showing there is an interaction but it does not affect the structural integrity of HSA. An enzyme assay showed they acted as non-competitive inhibitors of α -ChT with a non-denaturing gel showing the formation of a new species between α -ChT and the resorcinarene compounds.

1.2.1.3 Porphyrins

Since 1950, studies have investigated porphyrins binding to proteins; initially this was focussed on binding to HSA,³⁴⁻³⁶ but subsequently, porphyrins have been used for binding to a variety of different proteins, including K_v potassium channels,³⁷⁻⁴¹ VEGF,⁴² cyt *c*^{43,44} and lectins.⁴⁵⁻⁴⁷

Trauner and coworkers rationally designed a tetraphenylporphyrin-based scaffold (**5**) to target the K_v potassium channel with nanomolar affinity (**Figure 1.8a**), reducing the current through the channel.³⁸ The C_4 symmetry of the porphyrin **5** was thought to be well-suited to the tetrameric nature of the potassium channel;³⁸ however, solid state NMR has since shown that porphyrin **5** does not interact with all four subunits of the eukaryotic $K_v1.3$ channel simultaneously,^{37,48} and instead lies perpendicular to the protein, projecting one of its cationic side chains into the channel rather than interacting with the surface of the channel.³⁷ Porphyrin **5** has since been shown to block the ion conduction pathway and stabilise a closed K_v channel state upon interaction with the voltage sensor domain.⁴¹ The Nolan group expanded on this work, looking at inhibiting specific K_v1 channels, using a structure activity relationship (SAR) study to find ligands selective for the $K_v1.1$ and $K_v1.2$ channels.⁴⁹

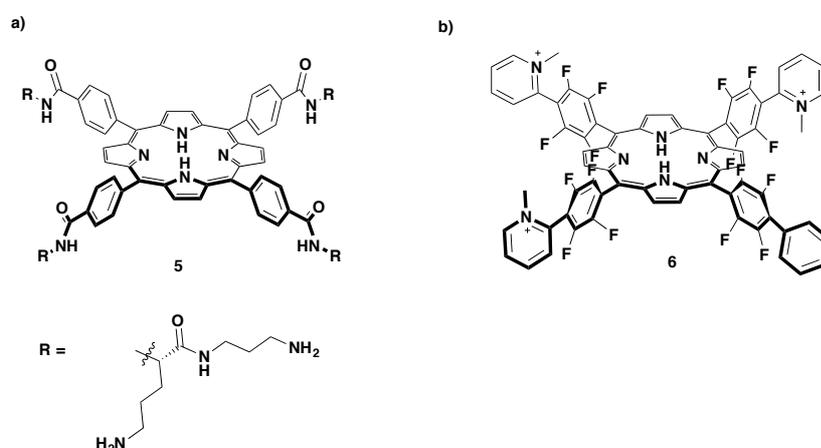


Figure 1.8 Porphyrin protein surface mimetics a) Porphyrin ligand of the K_v potassium channel, b) Porphyrin ligand of FGF

The Yayon group have studied other porphyrins (**6**) that bind to fibroblast growth factor (FGF) and vascular endothelial growth factor (VEGF) (**Figure 1.8b**),⁴² with low micromolar affinity *in vitro*, *in cellulo*, and *in vivo*. They also showed they were selective inhibitors of the VEGF/VEGF receptor PPI over the EGF/EGF receptor PPI. Further analysis led to the elucidation of higher affinity ligands with cationic porphyrins having the highest affinity *in cellulo*.⁴²

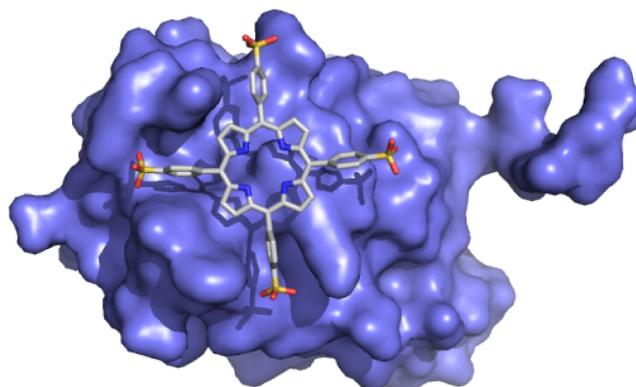


Figure 1.9 Protein crystal structure of tetrasulfonatophenyl porphyrin bound to Jacalin (PDB ID: 1PXD)

The binding of porphyrins to lectins has been extensively studied with crystal structures having been solved for tetrasulfonatophenyl porphyrin binding to Jacalin (**Figure 1.9**),⁵⁰ peanut lectin (PNA),⁵¹ and concanavalin A (ConA),⁵² showing different binding modes in each, highlighting the utility of porphyrins in protein binding.

1.2.1.4 Dendrimers

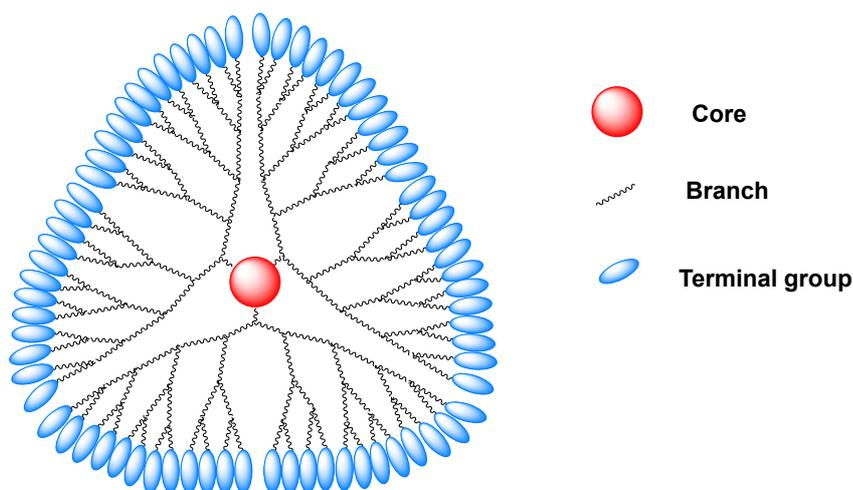


Figure 1.10 Schematic of a dendrimer, showing the core, branches and terminal (binding) groups

Dendrimers are supramolecular scaffolds of high valency, with a central core that projects a branching network of repeating units culminating in terminal functionality which can be used for binding to proteins (**Figure 1.10**).⁵³ Twyman and coworkers have designed poly anionic poly(amidoamine) (PAMAM) dendrimers (**Figure 1.11, 8**) which bind to cyt *c* and α -ChT.⁵⁴ The best ligands for both proteins, are those whose maximum addressable surface area matches the interfacial surface area of the protein. They have also shown that these dendrimers do not effect conformational changes in the structure of either protein.⁵⁵

PAMAM dendrimers have also been shown to bind to HSA in an extensive study by the Giri group.⁵⁶ They studied binding constants, NMR (¹H, STD and DOSY) and molecular dynamic (MD) simulations of 19 PAMAM dendrimers in order to gain insight into the interactions, looking at differences in core, dendrimer generation and terminal group. This allowed for an analysis of the effect of hydrogen bonding, hydrophobicity, and electrostatic interactions on the binding to HSA. The NMR and MD simulations show that the inner shell protons of the dendrimer interact more strongly with the protein, compared to the outer protons.

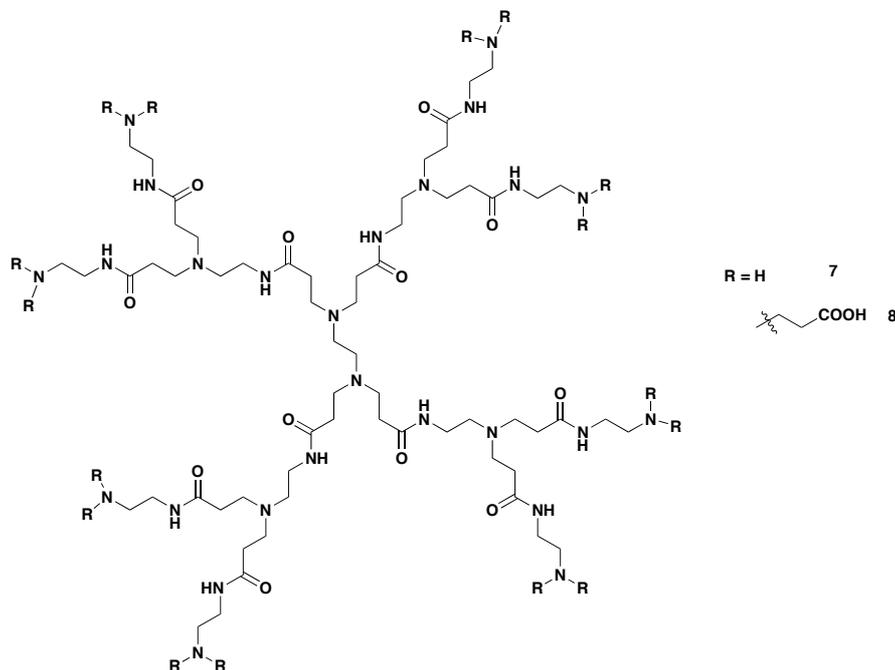


Figure 1.11 1st generation PAMAM dendrimer

1.2.1.5 Other scaffolds

The Hamilton group have also investigated anthracene receptors which bind to cyt *c* and lysozyme (**Figure 1.12**).⁵⁷ The anthracene analogues (**9**) have a hydrophobic core surrounded by carboxylic acids to enable complementarity with the cyt *c* and lysozyme

surfaces. 0.66 and 0.52 μM binding affinities (5 mM sodium phosphate, pH 7.4 buffer) were observed for cyt *c* and lysozyme respectively using a fluorescence quenching assay for cyt *c* and a Förster resonance energy transfer (FRET) assay for lysozyme. Thus implicating a different, but not further developed, surface mimetic scaffold.

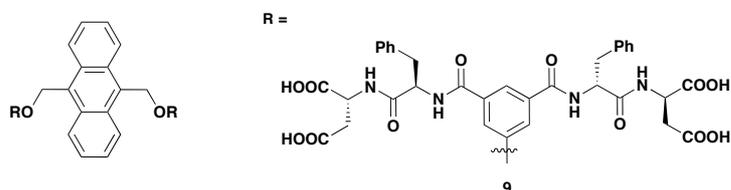


Figure 1.12 Anthracene protein surface mimetic scaffold

Bivalent cyclodextrins have been synthesised, by Breslow and coworkers, to inhibit aggregation of citrate synthase and L-lactate dehydrogenase, by binding to (and thus preventing the aggregation of) hydrophobic patches on the protein surfaces. This was demonstrated by enzyme activity assays, non-denaturing size exclusion chromatography and chemical crosslinking.⁵⁸

Kano and Ishida also developed a polyanionic β -cyclodextrin capable of binding to cyt *c*.⁵⁹ Initial binding studies by isothermal titration calorimetry (ITC) show 1:1 entropically driven, electrostatic binding, with NMR suggesting binding at the haem exposed edge. Further confirmation and proposals for other binding sites were obtained by observation of partial inhibition of ascorbate reduction, and a decrease in reduction rate of cyt *c* in the presence of cyt *c* reductase. Additional studies involved making a more complicated ternary complex with a porphyrin spanning two cyt *c* bound cyclodextrins.

1.2.2 Metallo protein surface mimetics

1.2.2.1 Advantages of using metal complexes

Metals can offer advantages over conventional organic scaffolds, including the ability to offer a wider variety of coordination numbers and geometries, thus expanding the number of globular shapes available, allowing the potential for them to fit into pockets and onto surfaces not accessible to small organic molecules.⁶⁰ Metal complexes can also exist as many more stereoisomers than organic molecules, for example an sp^3 carbon with 4 different substituents has only two possible stereoisomers whereas an octahedral metal centre with 6 different ligands can exist in up to 30 different stereoisomers (**Figure 1.13**).⁶¹ The protein binding selectivity of small molecules has been shown to correlate to both their shape and stereochemical complexity,⁶² thus highlighting how the complexity of

a metal coordination complex can potentially be used for the generation of selective protein ligands.

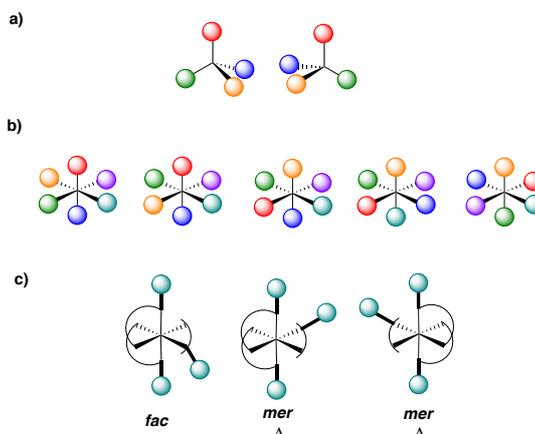


Figure 1.13 Difference between sp^3 carbon and octahedral metal scaffolds a) An sp^3 carbon with 4 substituents can exist as two possible enantiomers, b) Examples of the 30 possible stereoisomers for an octahedral centre with 6 different substituents, c) *fac*, *mer*, Δ and Λ isomers on an octahedral centre can also exist, particularly when using unsymmetrical bidentate ligands

The use of metal complexes allows for combinatorial synthesis in order to generate a wide range of metal complexes using similar reactions,⁶¹ allowing the screening of a variety of compounds more easily. The metal centre itself can be used solely as a scaffold, for forming coordinative bonds with biological macromolecules, and for its reactive capacity, thus expanding the scope of possible binding interactions.⁶¹ Ligands can play a large role in redox behaviour, biostability, absorption and delivery of the metal complex. The ligands can also be used to direct the synthesis towards particular stereoisomers, for example by utilising the *trans* effect.

The use of metal scaffolds as molecular sensors also offers advantages over conventional small molecules, as it is possible to choose a metal scaffold which can itself be visualised using its intrinsic luminescence, rather than requiring functionalization of the organic framework, which may have an effect on binding, and other molecular properties, such as solubility. For example ruthenium(II) and iridium(III) complexes are phosphorescent, allowing for their direct visualisation in both biological assays and for cellular imaging.⁶³

1.2.2.2 Metal coordination to protein surfaces

Metal-ligand interactions in water are stronger than the conventionally important protein recognition interactions such as hydrogen-bonding, electrostatics and van der Waals contacts.⁶⁴ This makes metal-ligand interactions a potentially useful tool in the

recognition of proteins, as fewer interactions will be needed to achieve high affinity, selective binding. The scope of this approach, however, is limited to certain amino acids and post-translational modifications, which are able to coordinate to a metal centre.

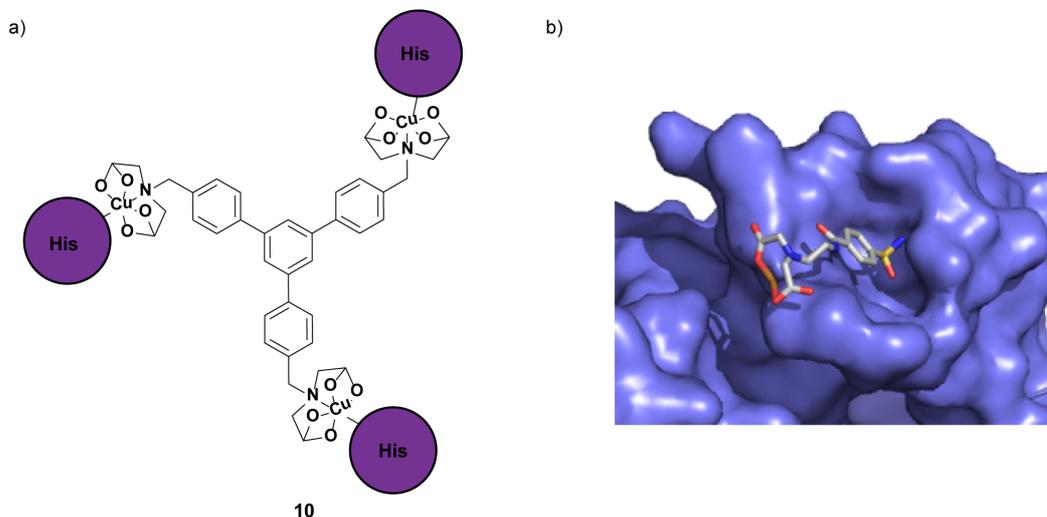


Figure 1.14 Cu(II)-IDA for binding to protein surfaces, a) Structure of Cu(II)-iminodiacetate for binding to surface exposed histidines on bovine erythrocyte carbonic anhydrase, b) Protein crystal structure of a Cu(II)-IDA bound to a histidine on human carbonic anhydrase II (PDB ID: 2FOV)

Mallik and coworkers used the knowledge that many transition metals bind to the imidazole side chains of histidines on proteins, a widely used concept in the purification of proteins by immobilized metal affinity chromatography.⁶⁵⁻⁶⁷ They used molecules with copper(II)-iminodiacetate (IDA) arms to recognise patterns of surface-exposed histidine residues (**Figure 1.14**), resulting in recognition of bovine erythrocyte carbonic anhydrase,^{68,69} after seeing that they could be used to bind to histidines on peptides in solution.^{70,71} A three Cu(II) system (**10**) was used to bind three histidine side chains on the *N*-terminus of the carbonic anhydrase, with the ligand alone showing no binding, illustrating the importance of the metal centre for recognition. The highest affinity compound (**10**, 3 μM K_d by ITC) was also found to be selective for carbonic anhydrase over chicken egg albumin, a protein with the same number of surface exposed histidine residues (six) but in different spatial orientations.

Along similar lines, Hamachi and coworkers used *bis*(Zn(II)-dipicolylamine (Dpa))) derivatives to bind histidine residues on the surface of α -helical peptides, stabilising their α -helical conformations.^{72,73} This led to the use of the *bis*(Zn(II)-Dpa) complexes in the binding of both mono- and multi- phosphorylated peptides via bidentate binding between the Zn(II) and the phosphate groups, resulting in stabilisation of the peptides (**Figure 1.15**).^{74,75} The targeting of phosphate groups on protein surfaces is of particular interest

as protein phosphorylation is used as a ubiquitous signaling mechanism within cells so binding preferentially to a phosphorylated state of a protein may allow modulation of signaling pathways.

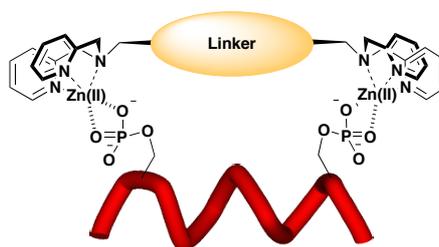


Figure 1.15 *bis*(Zn(II)-(dpa)) species bound to two phospho-amino acid residues, stabilising the α -helical conformation

The *bis*(Zn(II)-Dpa) species have been used to generate chemosensors tailoring the bridging group between the two Zn(II) centres to cause a change in fluorescence on binding, in Hamachi's case this involved using anthracene or bipyridine moieties, thus allowing for fluorescence binding assays to be developed.⁷⁴⁻⁷⁷ Using di-phosphorylated model α -helical peptides it was shown, by CD, that appropriately spaced Zn(II) centres increased the α -helical content of the peptide. They showed 10-fold selectivity for di-phosphorylated over mono-phosphorylated peptides, but little selectivity between different di-phosphorylated peptides, due to the high flexibility in the linker between the two metal binding sites.^{75,78} This has subsequently been used to inhibit the phospho-PPI between the phosphorylated CTD peptide and the Pin1 WW domain with 5.6 μM affinity.⁷⁹ A more rigid diazastilbene linker has since been used to generate a receptor for $(i, i+1)$ di-phosphorylated peptides, which gives different luminescent responses for different spatial relationships between the phospho-amino acids residues.⁸⁰ Zn(II) complexes based on these scaffolds have also been coupled to a *bis*[(4,6-difluorophenyl)pyridanto-*N,C*2'] iridium(III) picolinate to generate a phosphorescent sensor for phosphorylated peptides, with much better selectivity over ATP.⁸¹

Following on from Hamachi's work, Gunning and coworkers used Cu(II)-*bis*(Dpa) and Zn(II)-(Dpa) complexes to bind to phospho-tyrosine STAT3, inhibiting STAT3/STAT3 dimerisation.^{82,83} ITC and fluorescence polarisation data showed the Cu(II) complexes binding to a phospho-peptide (with micromolar K_d), inhibiting the phospho-peptide-protein complex, with micromolar K_i .⁸² The Cu(II) complexes were further shown to inhibit STAT3/STAT3:DNA binding, using an electrophoretic mobility shift assay, with 8.2 μM affinity for the highest affinity ligand. They also exhibited low micromolar IC_{50} in 3 different cancer cell lines but much lower inhibition in, and low cytotoxicity in healthy

NIH3T3 cells, thus showing their potential therapeutic utility.⁸² Later, they showed the use of *bis*(Zn(II)-(Dpa)) complexes as similar mimics of *src* homology domain 2 (SH2) domains showing, by fluorescence quenching experiments, that the Zn(II) complexes could bind phospho-tyrosine peptides, with $K_d \sim 10^{-7}$ M and some sequence identity discrimination.⁸³ Some of the Zn(II) compounds were shown to be cytotoxic in three types of cancer cell, but with some inconsistencies with the fluorescence binding data.⁸³

1.2.2.3 Metalloporphyrins

Considerable research was performed in the 80s and 90s on electron transfer between both metallo and non-metallo anionic porphyrins and *cyt c*.⁸⁴⁻⁹¹ With two types of porphyrins being compared: uroporphyrins and 4-carboxyporphyrins by Jameson *et al.*⁸⁹ in 1997. 4-Carboxy porphyrins were shown to have higher fluorescence quenching rates, probably due to the difference in orientation for the two porphyrins, which can be visualised by the induced CD of the porphyrins on binding to *cyt c*. Interestingly these original porphyrins do not affect the CD of *cyt c* at room temperature, indicating that there is no change in protein tertiary structure on binding. The Rodgers group also used cationic metalloporphyrins as extrinsic probes to study peptide aggregation by analysing photoinduced electron transfer (PET) from tyrosine or tryptophan residues in the protein to the metalloporphyrin.^{92,93}

Following Fisher's initial observation that tetra-carboxy phenyl porphyrin bound to *cyt c*, selectively over acetylated *cyt c*, with K_d in the region $0.05 \mu\text{M} - 5 \mu\text{M}$ using a flavodoxin competition assay,⁸⁴ the Hamilton group developed higher affinity *cyt c* tetra-phenyl porphyrin ligands.^{43,44} Tetra-phenyl porphyrin scaffolds provide a large, flat, semi-rigid molecular surface of $\sim 300 - 400 \text{ \AA}^2$ which with anionic substituents on the periphery bind to *cyt c* in a 1:1 stoichiometry.⁴³ They developed first nano-⁴³ then subnano-molar⁴⁴ receptors for *cyt c* and showed that increasing the number of carboxylates on the periphery led to significant increases in binding affinity.⁴⁴ They also showed that there were changes in affinity by altering the relative proportions of acidic and aromatic functionalities.^{43,44} The compounds were found to be selective for *cyt c* over the related proteins, *cyt c*₅₅₁ and ferredoxin.⁴⁴ Crowley and coworkers later analysed sulfonato-porphyrins binding to *cyt c* by ¹H-¹⁵N HSQC NMR, with the results corroborated by molecular docking.²⁶ They showed that a dynamic ensemble of energetically similar interactions exists, with the porphyrin being able to move over different patches on the *cyt c* surface.²⁶ The Hamilton group have also used functionalised porphyrins to design an array for protein detection.^{94,95}

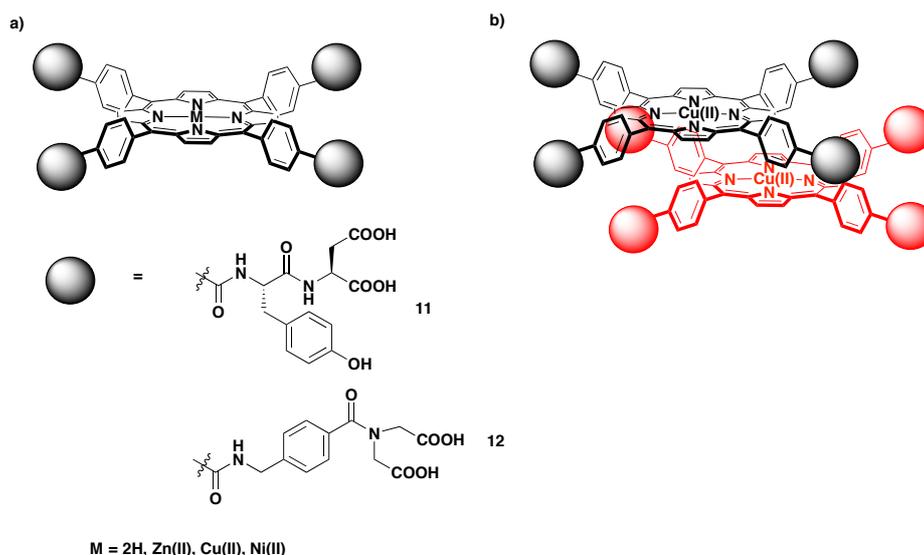


Figure 1.16 Metalloporphyrins for binding to cyt *c* a) The two best cyt *c* binders and denaturants, b) schematic of the Cu(II) porphyrin dimers

Interestingly the compounds were found to denature cyt *c*, shown by a lowered melting temperature (T_m),⁹⁶ by up to 50 °C.⁴⁴ The anionic tetra-phenyl porphyrins (**11** and **12**, M = 2H, **Figure 1.16a**) had a more profound effect compared to similarly charged molecules and porphyrins without the phenyl group; this could perhaps be seen as unsurprising as studies with different anionic polymers on their denaturing effect on cyt *c* showed that a hydrophobic portion is required for observable changes in cyt *c* stability.⁹⁷ The porphyrins do not show a denaturing profile for acetylated cyt *c* or cyt *c*₅₅₁, showing that charge complementarity is key, though higher affinity binding to cyt *c* does not necessarily make for a better denaturant. This is likely because the porphyrin is thought to destabilize the native state and stabilize the unfolded state of the protein, so some of the higher affinity ligands may bind better to the native state than the lower affinity binders, stabilising the native state comparatively. The compounds also show increased trypsin digest rates of cyt *c*, even with only 0.1 equivalents of porphyrin to protein.

After studying free base porphyrins, the Hamilton group analysed metalloporphyrins, in particular Cu(II) porphyrins.^{98,99} These are known to dimerise in aqueous solutions due to enhanced π - π stacking (**Figure 1.16b**).^{100,101} The Cu(II) porphyrins, were shown to have 60 nM affinity for cyt *c* and bind in a 2:1 porphyrin:protein ratio, with the porphyrins maintaining their dimeric nature. The Cu(II) porphyrins (unlike the monomeric Zn(II) equivalents) were shown to denature cyt *c*, first stoichiometrically⁹⁸ and then catalytically (0.1 eq.),⁹⁹ with some able to do so at room temperature. This denaturation was selective for cyt *c* over α -lactalbumin, Bcl-x_L, cyt *c*₅₅₁, myoglobin and RNase A. The denaturation was shown to unravel the α -helical nature of the protein as well as increase the tryptic digest

rates. There are significant differences between the free base, where there is some lowering of the T_m , and the Cu(II) porphyrins, which have much larger effects arising from the dimeric nature of the Cu(II) porphyrins compared to the monomeric free base porphyrins ($M = 2H$). In order to bind to the native *cyt c* structure the Cu(II) porphyrin dimers may have to dissociate, whereas to bind to the denatured *cyt c* they do not, this creates a catalytic cycle whereby the Cu(II) porphyrin catalyses the denaturation. After studying *cyt c* denaturation they decided to assess other haem proteins and showed that Cu(II) porphyrins can also be used to denature both myoglobin and haemoglobin.¹⁰²

Zn(II) and Fe(III) metalloporphyrins have also been shown to multimerise *cyt c*₇ from *Geobacter sulfureducens*, hen egg lysozyme and horse heart *cyt c* at high (millimolar) porphyrin and protein concentrations,¹⁰³ as observed by SAXS and corroborated by MD calculations.

1.2.2.4 Metallo dendrimers

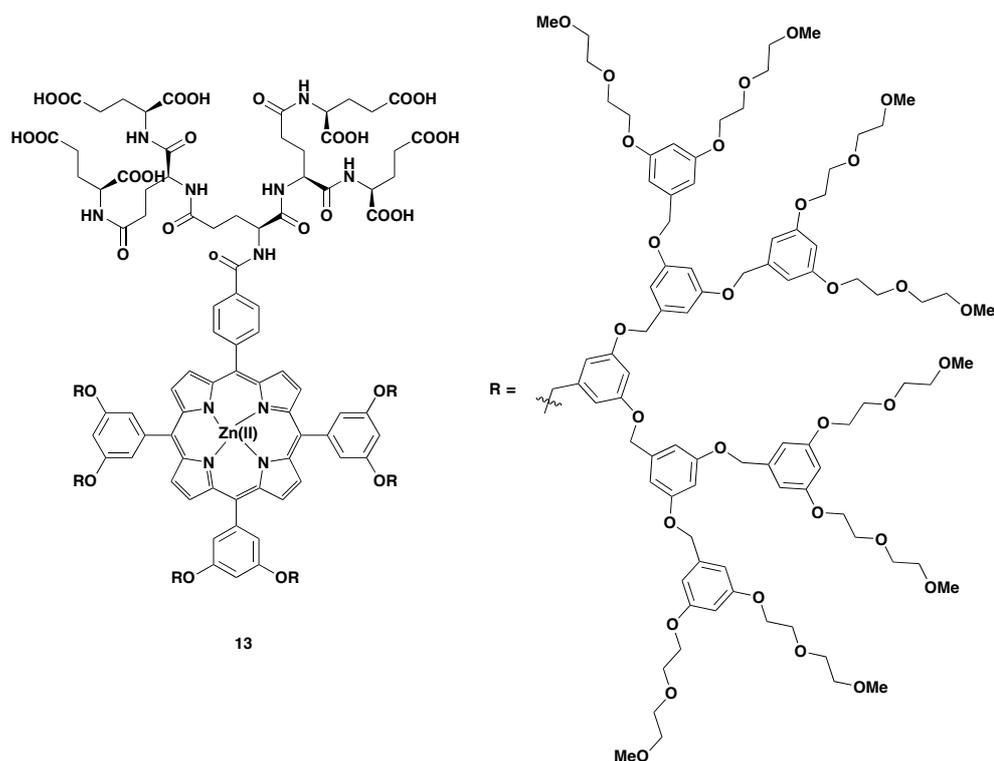


Figure 1.17 Zn(II) dendrimers that bind to *cyt c*

Zn(II) porphyrin-based dendrimers have also been developed, with the fluorescent metalloporphyrin-core being utilised for detection (**Figure 1.17**).¹⁰⁴ These large multivalent nanoscale structures have been used to bind to *cyt c*, with the *cyt c*/dendrimer complex being more stable than the native *cyt c*/*cyt b*₅ PPI, evidenced by 20 % fluorescence recovery on addition of 14 equivalents of *cyt b*₅ to the *cyt c*/dendrimer

complex. One of these original Zn(II)-porphyrin dendrimers (**13**), and subsequent generations, were shown to improve cell viability when cells are subjected to an apoptotic stimulus.¹⁰⁵ It has been hypothesised that the dendrimers trap *cyt c*, preventing it from interacting with Apaf1 to form the apoptosome, thus inhibiting apoptosis.

1.2.2.5 Transition metal complexes

1.2.2.5.1 Meggers Ruthenium(II) complexes

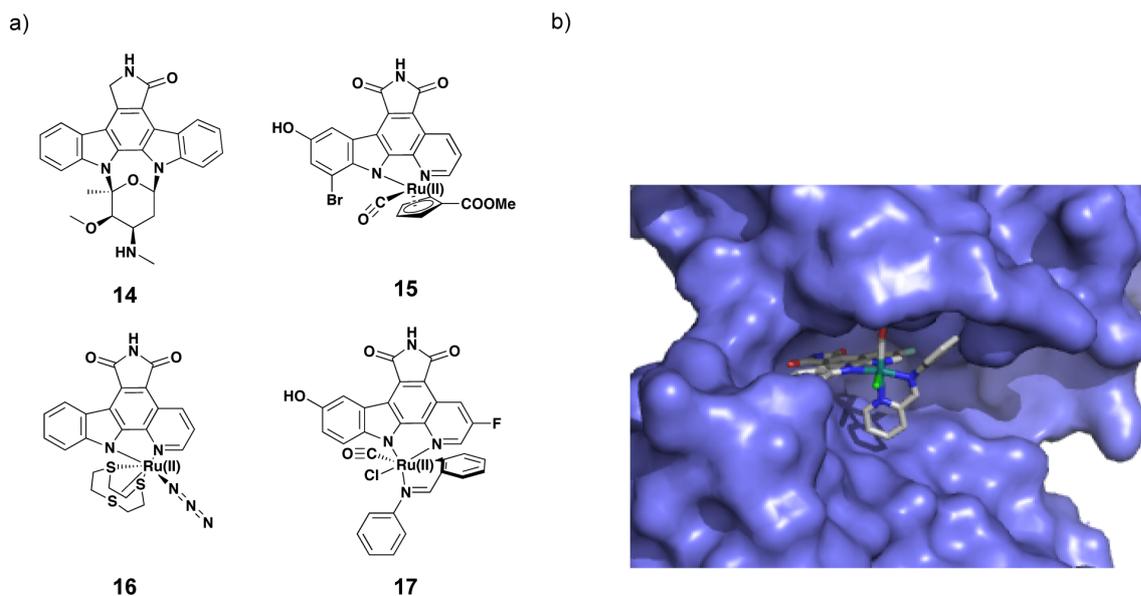


Figure 1.18 Meggers Ru(II) based protein kinase ligands a) Structures of staurosporine (**14**), a pan kinase inhibitor and Ru(II) complexes based on it for binding to GSK-3 β (**15**), MSK-1 (**16**) and PAK-1 (**17**), b) Protein crystal structure of a Ru(II) complex bound to PAK-1 (PDB ID: 3FXZ)

The use of metals as scaffolds for protein binding molecules has been pursued by many groups, with the Meggers group being a front-runner. They have mainly used ruthenium(II) complexes (**Figure 1.18**), but more recently have branched out to using rhodium(III),^{106,107} iridium(III),^{108,109} osmium(II)¹¹⁰ and platinum(II)¹¹¹ for the inhibition of many different protein kinases including Pim1,^{112,113} glycogen synthetases kinase 3 β (GSK3 β),¹¹⁴ MSK1,¹¹² BRAF kinase,¹¹⁵ and PAK1.¹¹⁶ X-ray crystal structures have been solved for many of these compounds bound to their target kinases (**Figure 1.18b**), showing the metals acting in purely structural capacities.^{113,117} The majority of these have been adenosine triphosphate (ATP) mimics, being based on staurosporine (**14**), a widely studied organic ATP mimic, but non-ATP mimics have been studied more recently,¹¹⁸ as have inhibitors of other nucleotide binding proteins including the human repair enzyme 7,8-dihydro-8-oxoguanosine triphosphatase,¹¹⁹ and the lipid kinase PI3K.¹²⁰

However these studies have all looked at ‘small’ metal complexes fitting into defined pockets (ATP binding sites) on the protein rather than the protein surface; in order to bind to protein surfaces it is more appropriate to consider larger supramolecular scaffolds.

1.2.2.5.2 Inert Group 9 complexes

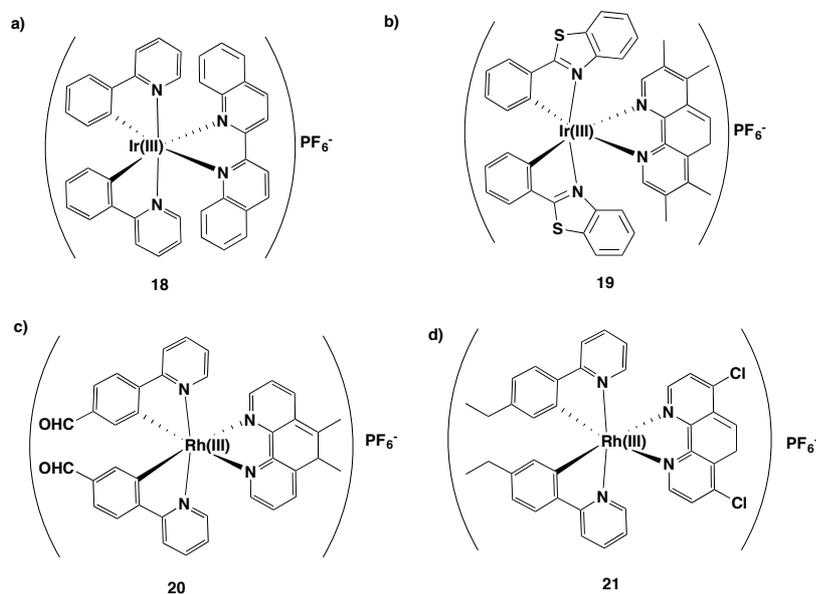


Figure 1.19 Inert group 9 protein surface ligands a) TNF- α ligand, b) Improved TNF- α ligand, c) STAT3 ligand d) *hDM2* ligand

The Leung group have worked on multiple Ir(III) and Rh(III) compounds, capable of binding to protein surfaces. They first developed cyclometalated Ir(III) complexes capable of binding to tumour necrosis factor- α (TNF- α).¹²¹ The Ir(III) complex developed utilises the aromatic bidentate ligands 2-phenylpyridinato (ppy) and 2,2'-biquinoline (biq) (**Figure 1.19a**), in order to target a hydrophobic binding site on the TNF- α dimer, preventing active trimer formation. Both enantiomers of the Ir(III) complex were found, by ELISA, to have an IC₅₀ in the region of 20 μ M, comparable to that of SPD304,¹²² one of the strongest inhibitors of TNF- α . Structure-activity relationships have since been performed, using 22 Ir(III) complexes with ligands of different shapes and sizes in order to generate low micromolar inhibitors (**Figure 1.19b**) (seen in an *in cellulo* inhibition of TNF- α induced NF- κ B luciferase assay in HEP G2 cells).¹²³ They also looked at the effect of stereochemistry, comparing the Δ and Λ isomers, showing that the Λ isomers had increased cellular activity (3.4 μ M versus 9.9 μ M IC₅₀ in the cellular assay) and binding affinity (30 versus 57 μ M IC₅₀ in an *in vitro* assay).¹²³

They have synthesised Ir(III) and Rh(III) compounds capable of binding to, and preventing dimerization and phosphorylation of STAT3 (**Figure 1.19c**).¹²⁴ The most

potent Rh(III) compound (**20**) was found to have anti-tumour activity in a mouse xenograft tumour model and was found to bind to the SH2 domain of STAT3 with an IC₅₀ of 4.8 μM. STAT3 pull-down assays showed an inhibition of STAT3 dimerisation and Western blotting showed an inhibition of STAT3 phosphorylation. The group have also screened a series of iridium(III) complexes as inhibitors of the p53/hDM2 interaction (**Figure 1.19d**).¹²⁵ One compound (**21**) was shown to be a 16 μM inhibitor in a fluorescence anisotropy competition assay. Subsequent cellular analysis confirmed the induction of p21 (a downstream target of p53) and apoptosis.

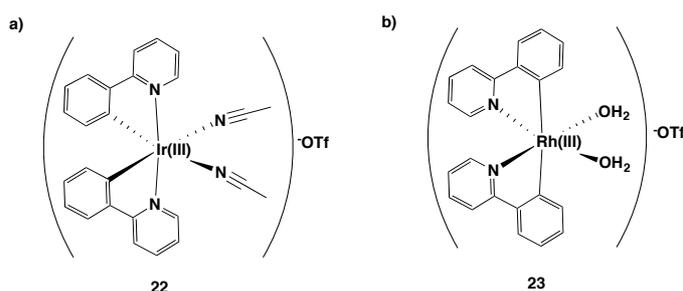


Figure 1.20 Irreversible group 9 metal complex protein surface binders a) BRD4, b) A β ₁₋₄₀

More recently they have developed an Ir(III) based irreversible inhibitor (**22, Figure 1.20a**) of the interaction between bromodomain-containing protein 4 (BRD4) and an acetylated histone peptide.¹²⁶ They initially screened 27 compounds and found a compound capable of modulating the interaction between BRD4 and chromatin *in vitro* and *in vivo*. The compound was found to bind to histidine residues, with the loss of acetonitrile ligands, and was found to be selective over the other histidine containing proteins STAT3 and caspase-6. They have also developed Ir(III) and Rh(III) complexes (**Figure 1.20b**) that inhibit the aggregation of A β ₁₋₄₀,¹²⁷ a peptide implicated in neurodegeneration in Alzheimer's disease. The compounds bind to histidine residues in the peptide, replacing the water ligands with these histidines, allowing further interactions of the hydrophobic coligands with hydrophobic residues at the N-terminus of the peptide. The compounds can be used as luminescent probes for the A β ₁₋₄₀ peptide and so offer an approach to the study of this important peptide.

1.2.2.5.3 Metal tris (bipyridines)

In the 1950s and 60s Dwyer and coworkers showed that simple bipyridine (bpy) and phenanthroline (phen) ruthenium(II) complexes show bacteriostatic and bacteriocidal activities and also inhibit tumour growth, thus showing the potential biological utility of these kinds of complex.^{128,129} Sasaki *et al.* further reported, in 1993, a saccharide

substituted Fe(II)(bpy)₃ complex capable of binding to lectins,¹³⁰ thus opening the idea of using metal *tris* (bipyridines) for binding to proteins.¹³⁰

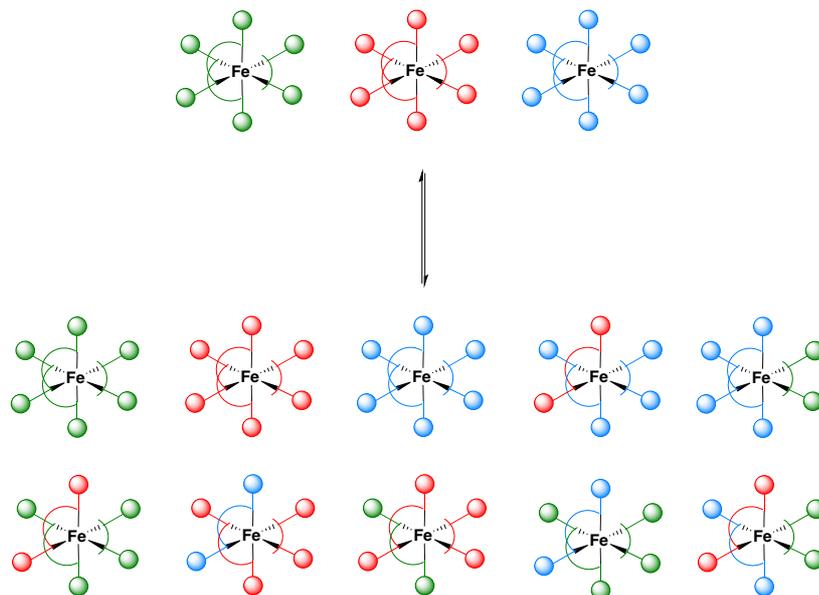


Figure 1.21 Schematic of DCC based on Fe(II)(bpy)₃ complexes

Fe(II)(bpy)₃ complexes are relatively dynamic in aqueous solution, this allows for the use of dynamic combinatorial chemistry (DCC) around the Fe(II) core (**Figure 1.21**). This has been used by the Sasaki and de Mendoza groups in order to generate lectin binding complexes.^{131,132} Sasaki and coworkers generated an Fe(II)(bpy)₃ complex with mono-GalNAc substituted bpy ligands, which altered its stereochemical configuration in solution resulting in the enrichment of higher affinity compounds for various different lectins.¹³¹ De Mendoza and co-workers used bipyridines functionalised with 3 different sugars, complexed them to Fe(II) then incubated the Fe(II)(bpy)₃ complexes with the mannose-binding lectin, Con A, this, as predicted, resulted in the enrichment of the mannose functionalised complex (detected by LCMS).¹³²

While the labile nature of the Fe(II)(bpy)₃ complexes can be useful for the generation and selection of high affinity protein binders, the inert nature of Ru(II)(bpy)₃ is of great interest, as decomplexation will not occur in biological media in dilute solution.¹³³⁻¹³⁰ Kaboyashi and coworkers,^{133,134} generated a series of glyco-functionalised Fe(bpy)₃ and Ru(II)(bpy)₃ compounds, and showed that the ruthenium(II) glycoclusters had high lectin affinity and increased luminescence on lectin binding. The Seeberger group have since developed sugar functionalised Ru(II)(bpy)₃ complexes (**24**) that bind to the concavalin A (ConA) and *galanthus nivalis agglutinin* (GNA) (**Figure 1.22**), by an electron transfer assay.¹³⁵ This was followed by a study using digital logic analysis to determine the best

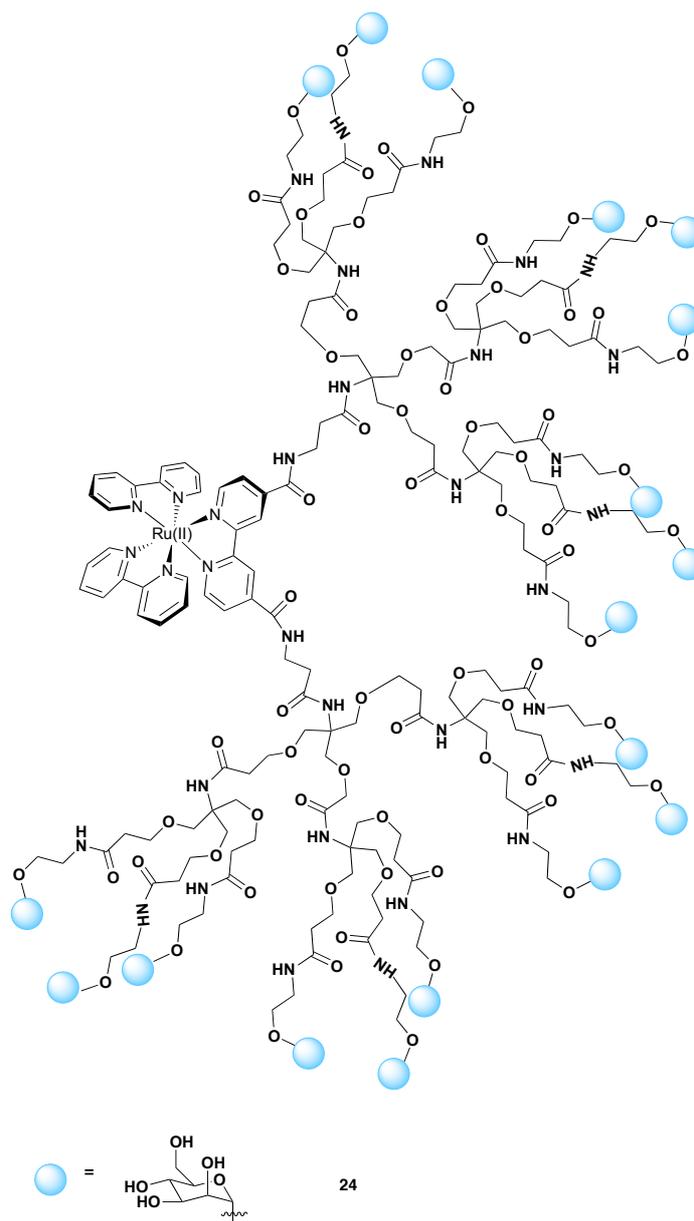


Figure 1.22 Seeberger's mannose functionalised Ru(II)(bpy)₃ complex for ConA/GNA binding

lectin binders for further study, by assessing the increase in luminescence output of the Ru(II)glycodendrimers with various different lectins.¹³⁶ The Ru(II)(bpy)₃ complexes with surface bound lectins have also been used as luminescent sensors for measuring monosaccharide and oligosaccharide concentrations, by using the displacement of the Ru(II)glycodendrimers from the lectin surface by the sugars.¹³⁷ They then further functionalised their Ru(II)(bpy)₃ complex scaffold, by addition of adamantane units, and adding mannose functionalised β -cyclodextrin to encapsulate the adamantane units, thus making a highly mannose functionalised subunit which binds to high density ConA with 0.14 μM K_d , as determined by SPR.¹³⁸ The Okada group have also looked at sugar functionalised Ru(II)(bpy)₃ complexes and have shown that galactose functionalised

Ru(II)(bpy)₃ complexes bind to *peanut agglutinin* with 6.1 μM K_d and glucose functionalised Ru(II)(bpy)₃ complexes bind to ConA with 18 μM K_d, by both fluorescence emission and polarisation assays.¹³⁹

Electron transfer experiments between *cyt c* and Ru(II)(bpy)₃ complexes (as well as Ru(II)(phen)₃, Os(II)(bpy)₃ and Os(II)(phen)₃ complexes) were initially reported by Cho in the 1980s.¹⁴⁰ Subsequently Hamachi and coworkers reported carboxylate functionalised Ru(II)(bpy)₃ derivatives (**25**, **Figure 1.23**) that could bind to, and photoreduce *cyt c* (pI = 10.0) selectively over other proteins (myoglobin, horseradish peroxidase and *cyt b*₅₆₂) with lower pIs (7.0, 8.0 and 5.0 respectively).¹⁴¹ The Ru(II)(bpy)₃ complexes **25** were found to bind to *cyt c*, by an ultrafiltration binding assay. The compound with the highest number of carboxylic acids (18) was shown to bind an order of magnitude better than an unfunctionalised Ru(II)(bpy)₃ complex. The Ru(II)(bpy)₃ complexes were capable of photoreducing *cyt c* with the most effective being an asymmetric Ru(II)(bpy)₃ complex with 12 carboxylates rather than the fully functionalised Ru(II)(bpy)₃ complex with 18 carboxylates.¹⁴¹

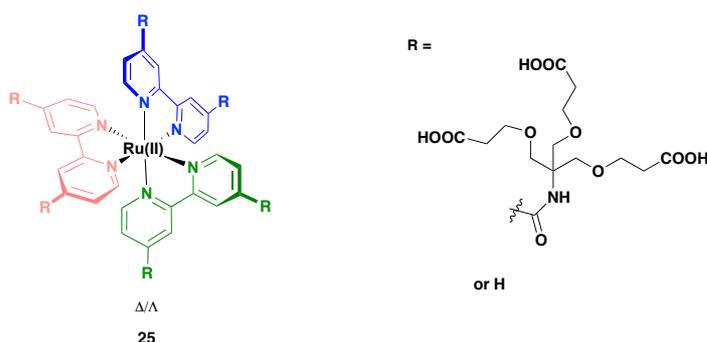


Figure 1.23 Hamachi's carboxylate functionalised Ru(II)(bpy)₃ complexes for *cyt c* binding and photoreduction

Following on from Hamachi's initial observations both the Ohkanda and Wilson groups further established selective binding of Ru(II)(bpy)₃ complexes to *cyt c* and α-ChT (**Figure 1.24**). The Wilson group developed both mono- (5') (**27a**) and di- (4,4') (**26a**) substituted bpy moieties which, when complexed to ruthenium(II), show 1.6 nM binding affinity (**26a**, 5 mM sodium phosphate, pH 7.4) for *cyt c* by a luminescence quenching assay.^{142,143} As with Hamilton's porphyrins,⁴³ negatively charged substituents (based on aspartic acid moieties) show good binding in luminescence quenching assays.¹⁴² Negative cooperativity was observed, with increasing numbers of carboxylates, the binding affinity per carboxylate decreases. Geometrically the *mer* isomers of the 5' mono-substituted Ru(II)(bpy)₃ complexes (**27a**) showed ~10 fold better binding affinity (25 vs. 172 nM for

the Δ isomers, 5 mM sodium phosphate, pH 7.4) than the *fac* isomers for cyt *c*, but the Δ and Λ isomers showed little difference in their binding affinities (25 vs 29 nM for the *mer* isomers).¹⁴³ Further analysis by a functional ascorbate assay showed that both the (4,4') disubstituted and 5' monosubstituted bipyridines slow the rate of reduction of cyt *c*, probably as a result of blocking the approach of the reducing agent to the haem group, which is surrounded by basic amino acid residues.¹⁴² The absence of binding to 60 % acetylated cyt *c* confirms that the charge complementarity is key to binding.^{142,143} The binding of the Ru(II)(bpy)₃ complex **26a**, similarly to Hamilton's porphyrins,⁹⁸ lowers the melting temperature of cyt *c* by 25 °C and shows an increased rate of proteolytic degradation in both stoichiometric and substoichiometric quantities of the Ru(II)(bpy)₃ complex.¹⁴⁴ At higher temperatures there is a change in binding stoichiometry observed with a change from 1:1 binding to 2:1 (protein:Ru(II)(bpy)₃ complex) binding being observed on changing from 25 to 70 °C. *In cellulo* work has also been performed with the 4,4'-disubstituted Ru(II)(bpy)₃ complexes showing 95 % efficiency of transfection into HEK-293T cells at 10 μ M concentration.¹⁴⁵ The Ru(II)(bpy)₃ complexes appear to localise in the lysosomes and in several instances are non-cytotoxic.¹⁴⁵

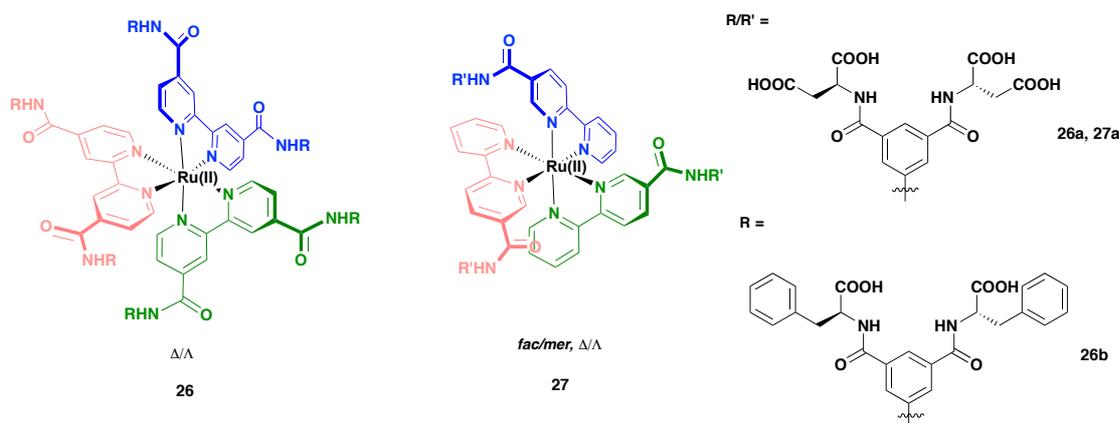


Figure 1.24 Ru(II)(bpy)₃ complexes designed by the Wilson and Ohkanda groups for binding to cyt *c* (**26a, 27a**) and α -ChT (**26b**)

Ohkanda and co-workers have shown that similar dendritic Ru(II)(bpy)₃ complexes (**26b, Figure 1.24**) bind to α -ChT in a mixed 1:1 and 1:2 (Ru(II)(bpy)₃ complex: α -ChT) stoichiometry with 130 and 430 nM K_d s (5 mM phosphate, pH 7.4) for the first and second equilibrium step respectively, inhibiting the enzyme by non-competitive inhibition.¹⁴⁶ They later synthesised heteroleptic Ru(II)(bpy)₃ complexes, containing multiple bipyridine ligands, for binding to both α -ChT and cyt *c*, with submicromolar affinity.¹⁴⁷ Molecular dynamics showed that 3 isophthalic arms interact with α -ChT, and 4 isophthalic

arms interact with cyt *c*.¹⁴⁷ *In cellulo* work also showed the Ru(II)(bpy)₃ complexes being able to enter cells.¹⁴⁷

1.3 Project aims

As described previously Ru(II)(bpy)₃ complexes have been rationally designed for binding to protein surfaces, including to lectins, cyt *c* and α -ChT, however little is known of how these Ru(II)(bpy)₃ complexes interact with proteins; the scope of these Ru(II)(bpy)₃ complexes for binding to protein surfaces and if we can use methods other than rational design to obtain novel and unsymmetrical Ru(II)(bpy)₃ complexes for binding to proteins.

The aims of this project were to further develop the Ru(II)(bpy)₃ scaffold in order to discover its utility for protein surface recognition. This involved-

- Furthering the understanding of the binding of previously designed Ru(II)(bpy)₃ to the surface of cyt *c*, by elucidating more information about the binding mode and the interactions involved in binding, as well as the location of the binding sites. This allowed a more complete understanding of how Ru(II)(bpy)₃ complexes bind to proteins, to allow for further rational design for binding to other, more therapeutically interesting, protein surfaces.
- The development of a suitable scaffold for dynamic combinatorial chemistry in order to generate higher affinity, and less symmetrical Ru(II)(bpy)₃ complex, and porphyrin binders for different proteins.
- The design of a protein sensing array using Ru(II)(bpy)₃ complexes to discriminate between a variety of different protein surfaces, to give another potential protein surface binding application of these Ru(II)(bpy)₃ complexes.

2 Biophysical and Structural Studies on Cytochrome *c* Recognition with Functionalised Ruthenium(II) Tris (bipyridine) Complexes

Work in this chapter has been reported in the manuscript ‘Protein Surface Mimetics; Understanding How Ruthenium *Tris* (bipyridines) Interact with Proteins’, S. H. Hewitt, M. H. Filby, E. Hayes, L. Kuhn, A. Kalverda, M. E. Webb and A. J. Wilson, *ChemBioChem*, 2017, 18, 223–231.

Cytochrome (cyt) *c* is a highly abundant, well-studied protein, which can be bought in gram quantities; this makes it an ideal protein for the study of how different molecular scaffolds can be used to bind to protein surfaces. In this work, cyt *c* has been used as a model protein for understanding how Ru(II)(bpy)₃ complexes bind to the surface of proteins. The cyt *c* binding of a range of Ru(II)(bpy)₃ complexes has been studied, followed by a more in depth study of the binding of two of these Ru(II)(bpy)₃ complexes to cyt *c* aimed at understanding how these Ru(II)(bpy)₃ complexes interact with cyt *c*, if they mimic the cyt *c*/cyt *c* peroxidase (CCP) PPI, and also to elucidate the binding site of the Ru(II)(bpy)₃ complexes on cyt *c*. Increased knowledge of how Ru(II)(bpy)₃ complexes bind to protein surfaces should allow for the design and use of novel Ru(II)(bpy)₃ complexes for binding to other, more therapeutically interesting, protein surfaces.

2.1 Cyt *c* and the cyt *c*/CCP PPI

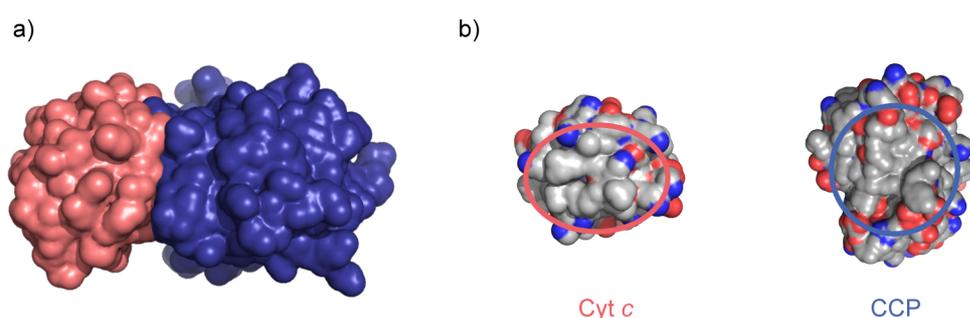


Figure 2.1 The cyt *c*/CCP PPI structure (PDB ID 1U75).¹⁴⁸ a) the cyt *c*/CCP PPI, b) the binding interface between cyt *c* and CCP showing the ring of basic amino acids (blue) on cyt *c* (red ring) and acidic amino acids (red) on CCP (blue ring) surrounding a central hydrophobic core on both proteins

Cyt *c* is a small, basic haem protein ordinarily located in the inner mitochondrial membrane. It is used within the electron transport chain of respiration, shuttling

electrons between cyt *c* reductase and cyt *c* oxidase. Cyt *c* is also implicated in apoptosis where, upon appropriate cellular signalling, it is released from the mitochondrion,¹⁴⁹ binds to apoptosis protease activating factor-1 (Apaf-1), generating a signalling cascade which results in the cleavage of many key cellular proteins.^{150,151}

Another of cyt *c*'s native protein partners is cyt *c* peroxidase (CCP), which is found in the mitochondrial intermembrane space of plants and takes electrons from reduced cyt *c* to convert hydrogen peroxide into water. The cyt *c*/CCP protein-protein interaction (PPI) has been very well studied,¹⁵² it was the first non-antibody PPI to have its X-ray crystal structure solved (**Figure 2.1a**).¹⁵³ The cyt *c*/CCP binding interface is located close to the haem exposed edges on both cyt *c* and CCP,¹⁵³ this allows it to perform its function, transferring electrons between the two proteins. The binding is based on a central hydrophobic core of amino acids on both proteins, surrounded by basic amino acid residues on cyt *c* and acidic amino acid residues on CCP (**Figure 2.1b**).¹⁵³ This allows for charge-charge complementarity between the two protein surfaces. The binding is very ionic strength dependent, indicating a large electrostatic driving force to binding, showing the importance of this charge-charge complementarity.¹⁵⁴ The binding interaction is entropy-driven, and is even enthalpically unfavourable.¹⁵⁵

2.2 Ru(II)(bpy)₃ complexes for binding to cyt *c*

Different multivalent scaffolds have been used to design molecules to bind to cyt *c*, many of these were discussed in Chapter 1. A defining feature of the design of these molecules is the addition of multiple carboxylic acid moieties to bind to the basic surface of cyt *c*.

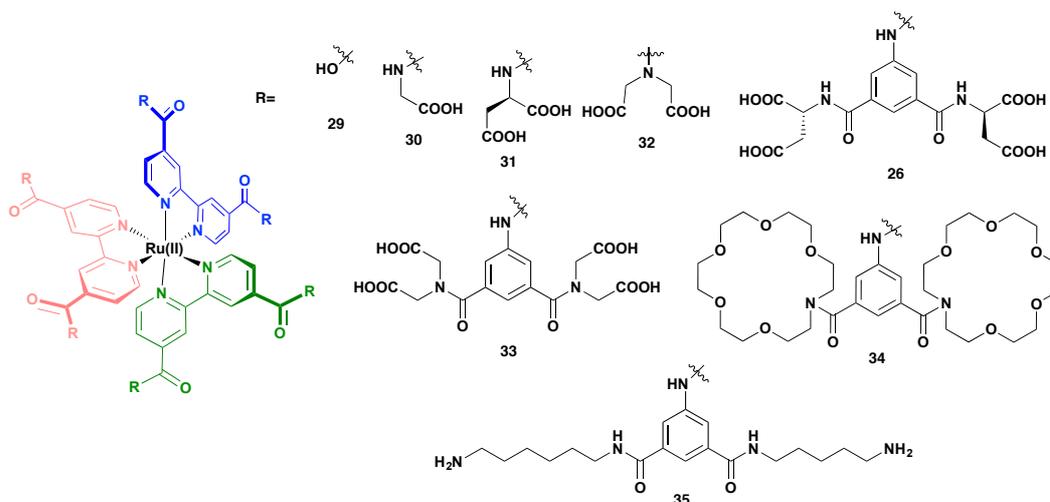


Figure 2.2 Ru(II)(bpy)₃ complexes designed for binding to cyt *c*

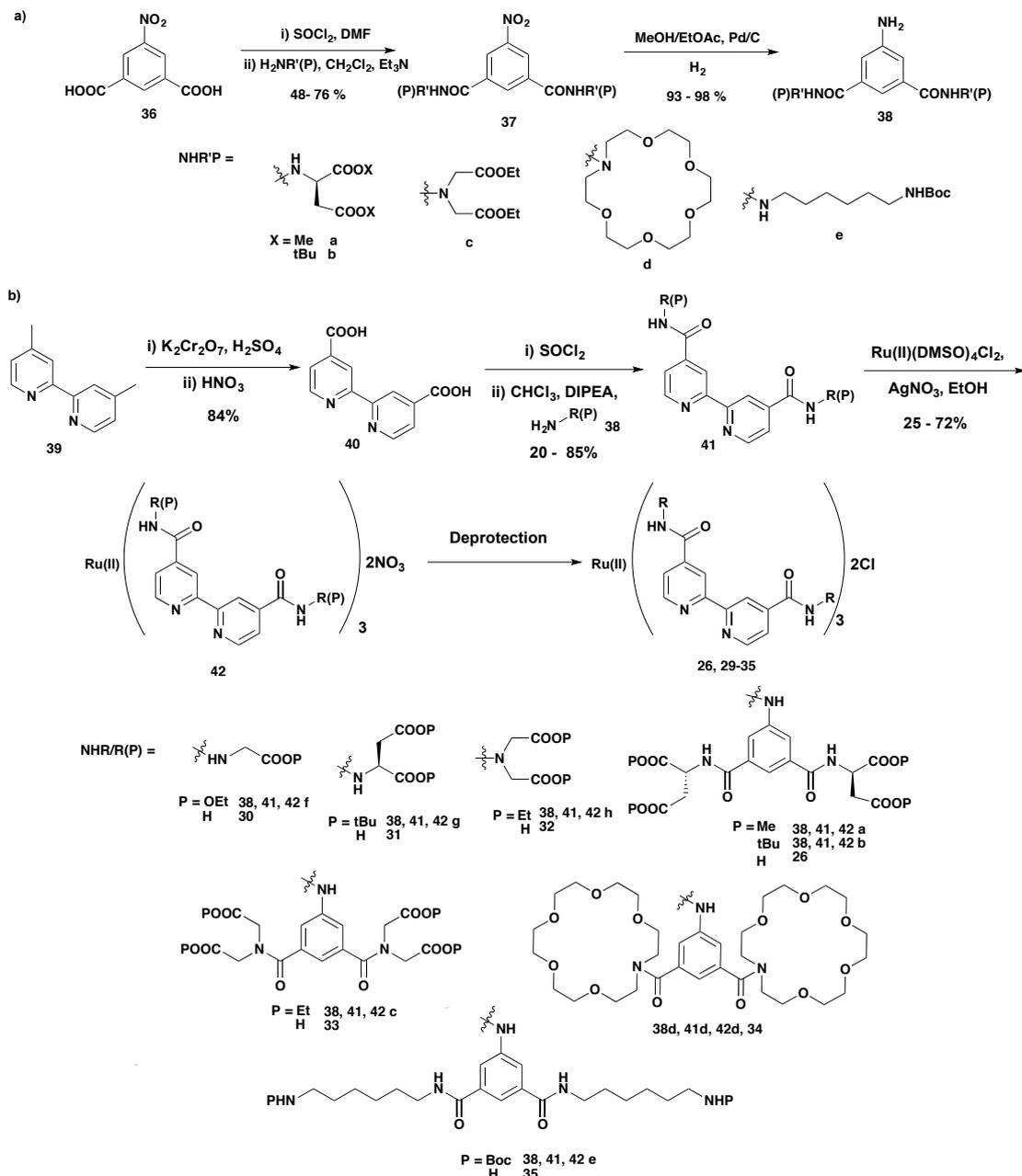
It was decided to synthesise the Ru(II)(bpy)₃ complexes **26** and **29 - 35** (**Figure 2.2**) in order to test their binding to cyt *c* and to understand their protein binding interactions. Ru(II)(bpy)₃ complexes **26**, **29**, **31**, **34** and **35** had previously been synthesised by the Wilson group for binding to cyt *c*.¹⁴² The Ru(II)(bpy)₃ complexes **26**, **29** and **31** were functionalised with multiple carboxylic acid moieties in order to bind to the basic amino acid residues on cyt *c*. Ru(II)(bpy)₃ complex **34** possesses crown ether moieties was designed to potentially chelate the lysine residues on cyt *c*, however no binding was detected, and Ru(II)(bpy)₃ complex **35** was designed as a negative control, being amine functionalised it should not bind to the basic cyt *c*. In addition to the previously designed Ru(II)(bpy)₃ complexes, Ru(II)(bpy)₃ complexes **30** and **32** have been synthesised in order to further expand the number of Ru(II)(bpy)₃ complexes (Ru(II)(bpy)₃ complex **32** was synthesised by Georgina Pleasance) to see if the spatial location of the carboxylates affects binding affinity. The synthesis of Ru(II)(bpy)₃ complex **33** was also attempted, however the secondary amide bond was found to cleave during the ruthenium(II) complexation reaction. The initial thought behind the synthesis of Ru(II)(bpy)₃ complexes **32** and **33** was to systems with a similar cyt *c* binding profile to the previously designed Ru(II)(bpy)₃ complexes **26** and **31** that existed as enantiomers rather than diastereomers, potentially permitting attempts at protein crystal growing trials with the Ru(II)(bpy)₃ complexes.

2.2.1 Synthesis

The synthesis of these Ru(II)(bpy)₃ complexes was achieved as described previously (**Scheme 2.1**),¹⁴² with formation of the protected ligands **41** by amide bond formation on 4,4'-dicarboxylic acid-2,2'-bipyridine **40**, *via* a particularly water sensitive diacid chloride. For the larger Ru(II)(bpy)₃ complexes (**26**, **33**, **34** and **35**), the functionalised anilines **38** for the amide bond formation were synthesised as in **Scheme 2.1a**, by amide bond formation, *via* the diacid chloride, on 5-nitro isophthalic acid **36**, followed by hydrogenation of the nitro group to the aniline **38**. These hydrogenation reactions could not be performed on large scale, as there was no reaction or incomplete reaction when performed on more than 1 g scale.

The bipyridine ligands **41** synthesised were then complexed to ruthenium(II) using Wilkinson's reagent (Ru(II)(DMSO)₄Cl₂)¹⁵⁶ to yield the protected Ru(II)(bpy)₃ complexes **42**. These protected Ru(II)(bpy)₃ complexes **42** were purified by conventional silica or alumina column chromatography. This purification was changed from the previously reported ion exchange Sephadex columns,^{142,143} as Ru(II)(bpy)₃ complexes with large

hydrophobic protecting groups precipitated on these columns due to the aqueous NaCl eluents used.



Scheme 2.1 Synthesis of Ru(II)(bpy)₃ complexes **26, 29-35** a) Synthesis of isophthalimide anilines **38** for Ru(II)(bpy)₃ complexes **26, 31 - 35**, b) Synthesis of final Ru(II)(bpy)₃ complexes, R(P) = protected R group

These protected Ru(II)(bpy)₃ complexes **42** were subsequently deprotected in an appropriate manner, dependent on their protecting groups (**Table 2.1**). It should be noted that the anilide bonds in Ru(II)(bpy)₃ complex **26** are activated towards hydrolysis in both acidic and basic conditions, meaning mild conditions were required for this deprotection, with both acid (*tert*-butyl ester) and base (methyl ester) labile protecting groups being

attempted. A methyl ester strategy with only 50 equivalents (~2 per acid) of lithium hydroxide and a short (1 hour) reaction time was eventually found to be sufficient for deprotection, without cleavage of the anilide bond.

Table 2.1 Deprotection conditions for the Ru(II)(bpy)₃ complexes

Complex	Protecting group	Deprotection method
30	Et	1 M NaOH + EtOH
31	^t Bu	TFA/H ₂ O
32	Et	1 M NaOH + EtOH
26	^t Bu	TFA/H ₂ O/TIPS
26	Me	50 eq. LiOH in H ₂ O
33	Et	Degraded earlier in synthesis
34	None	n/a
35	Boc	1 M HCl in dioxane

2.3 Assay development

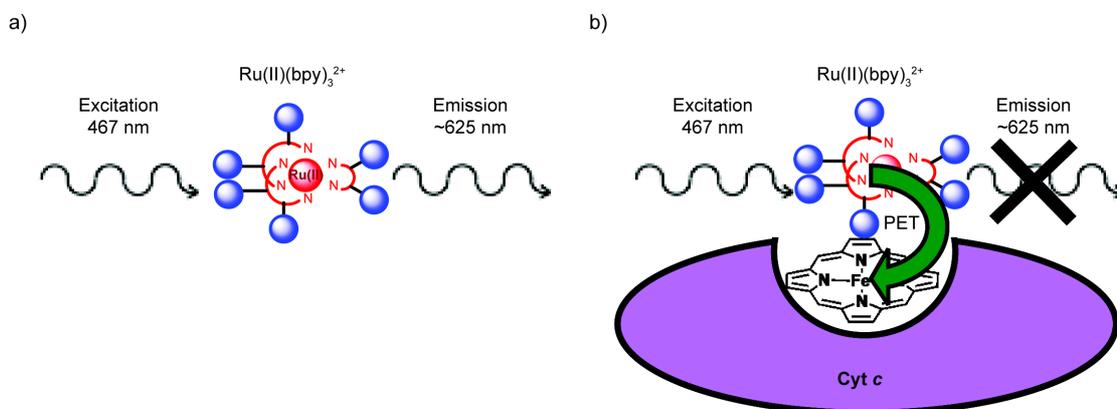


Figure 2.3 Cartoon illustrating how the luminescence quenching assay works a) The Ru(II)(bpy)₃ complex on excitation with light at 467 nm consequently emits light at ~625 nm, b) When bound to cyt *c*, the emission is quenched by photoinduced electron transfer (PET) to the haem group of cyt *c*

The detection of binding between the Ru(II)(bpy)₃ complexes and cyt *c* was achieved by a, previously reported, luminescence quenching assay (**Figure 2.3**).^{142,143} The Ru(II)(bpy)₃ complexes are luminescent, when excited by light at 467 nm, they consequently emit light at ~625 nm, however when bound to cyt *c* this luminescence is quenched by photoinduced electron transfer (PET) to the iron in the haem group of cyt *c*.

Titrating variable concentrations of cyt *c* into a fixed concentration of Ru(II)(bpy)₃ complex allows a binding isotherm to be established (**Figure 2.4**).

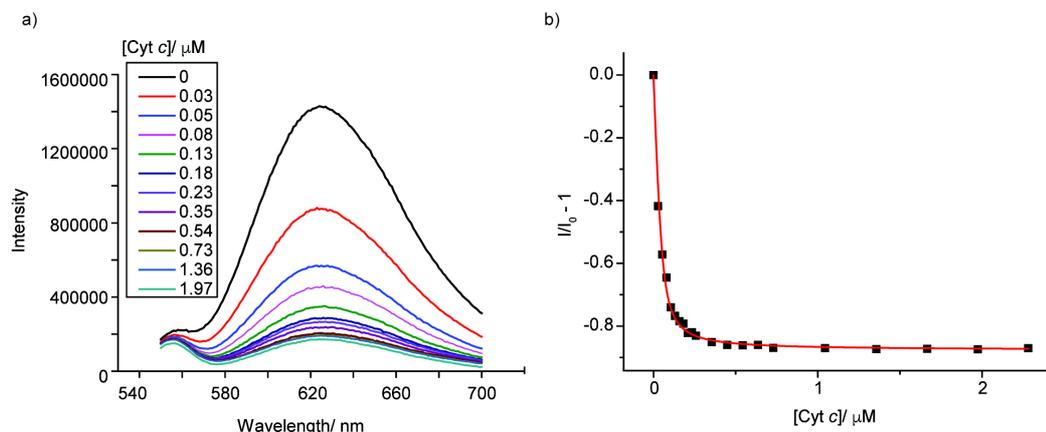


Figure 2.4 Luminescence quenching assay a) Graph showing the quenching of the luminescence of complex **26** upon addition of increasing concentrations of cyt *c* b) Fitting to 1:1 binding isotherm, in 5 mM sodium phosphate, pH 7.5

2.3.1 Assay development from fluorometer to plate reader.

Previously the luminescence quenching assays were performed on a fluorometer,¹⁴² however this does not allow for high-throughput screening, taking over 2 hours to measure the binding of one Ru(II)(bpy)₃ complex to cyt *c* in triplicate, therefore it was decided to develop the assay for use on a fluorescence plate reader. The use of a fluorometer allows for the use of quartz cuvettes, which have very little interaction with the components of the assay, as compared to the plastic plates used in conventional fluorescence plate reader assays, onto which binding of proteins and molecules is often observed.

Using the same conditions as for the fluorometer on a plate reader, did not produce an agreement in binding between the two machines, and data often did not fit a 1:1 binding isotherm (**Figure 2.5a**), indicating interactions between the plate and cyt *c* or the Ru(II)(bpy)₃ complex. This showed an additive to disrupt the interaction between the plate and the assay components was required.

Two commonly used buffer additives to disrupt these interactions are the detergent Tween20 and the blocking agent bovine serum albumin (BSA). Assays with these different additives were first attempted with a sample of Ru(II)(bpy)₃ complex **26**, with a small amount of impurity from the cleavage of an anilide bond in deprotection. Upon attempting the assay with the Tween20 and BSA additives on both the plate reader and fluorometer (**Table 2.2**), there was still a large discrepancy between the fluorometer and plate reader

when using Tween20, but there was little difference with the addition of BSA. Therefore BSA was added for all plate reader binding experiments. This decision was validated using a resynthesized clean sample of Ru(II)(bpy)₃ complex **26**, where the K_d values were 40.0 ± 4.5 nM and 17.5 ± 3.3 nM (50 nM Ru(II)(bpy)₃ complex **26**, 5 mM sodium phosphate, 0.2 mg mL⁻¹ BSA, pH 7.5) for the plate reader and fluorometer respectively, whereas without BSA the binding curve on the plate reader did not fit to the 1:1 binding isotherm (**Figure 2.5**), but gave a K_d of 10.5 ± 0.4 nM on the fluorometer, confirming the previously obtained data.

Table 2.2 Difference in cyt *c* binding affinity between the plate reader and fluorometer, all 100 nM Ru(II)(bpy)₃ complex **26**, in 5 mM sodium phosphate, pH 7.4

Buffer additive	K _d fluorometer/ nM	K _d plate reader/ nM
No additive	7.39 ± 3.29	95.27 ± 10.78
0.05 % Tween 20	9.23 ± 1.79	113.22 ± 30.27
0.2 mg mL ⁻¹ BSA	35.74 ± 7.19	49.62 ± 27.63

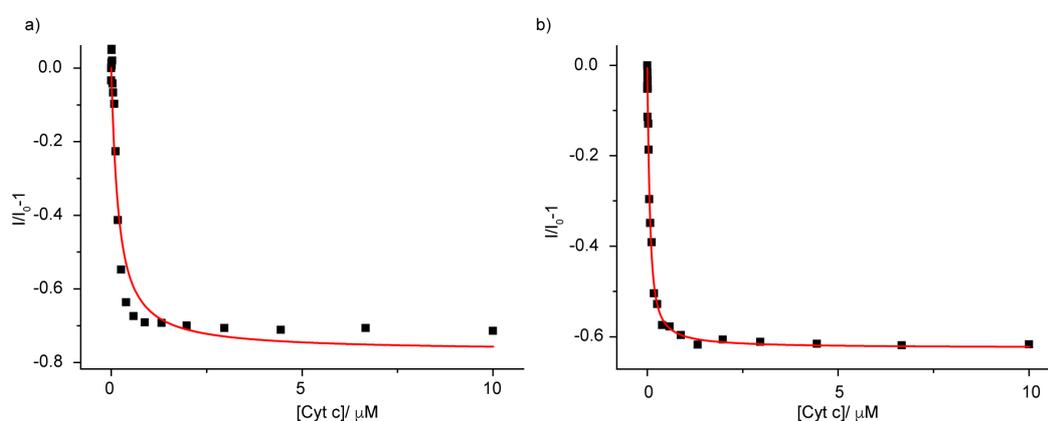


Figure 2.5 Binding of Ru(II)(bpy)₃ complex **26** to cyt *c* on plate reader in 5 mM sodium phosphate, pH 7.5 a) without BSA additive, b) with 0.2 mg mL⁻¹ BSA

2.4 Binding of different Ru(II)(bpy)₃ complexes

Having established an assay regime for more high-throughput binding cyt *c* detection, the binding affinities of the different Ru(II)(bpy)₃ complexes synthesised were assessed (**Table 2.3**). As can be seen Ru(II)(bpy)₃ complex **26** with 24 carboxylic acids has the highest binding affinity, the two Ru(II)(bpy)₃ complexes **31** and **32** with 12 carboxylic acids have lower affinity, but similar affinities to each other, the two Ru(II)(bpy)₃ complexes **29** and **30** with 6 carboxylic acids bind with much lower affinity, and the two negative controls **34** and **35** don't bind. This is as would be expected for binding to the

basic amino acid residues present on *cyt c*. The similarity between Ru(II)(bpy)₃ complexes **31** and **32** and complexes **29** and **30** show that the binding interaction is more dependent on the overall global charge located around the Ru(II)(bpy)₃ complex rather than small changes in location of the charge.

Table 2.3 Binding affinities for the different Ru(II)(bpy)₃ complexes synthesised to *cyt c*, 5 mM sodium phosphate, 0.2 mg mL⁻¹ BSA, pH 7.5

Ru(II)(bpy) ₃ complex	K _d / μM
29	114 ± 20
30	65.9 ± 7.0
31	2.58 ± 0.72
32	1.73 ± 0.50
26	0.0429 ± 0.0031
34	>100
35	>>100

2.4.1 UV/Vis ascorbate reduction assay

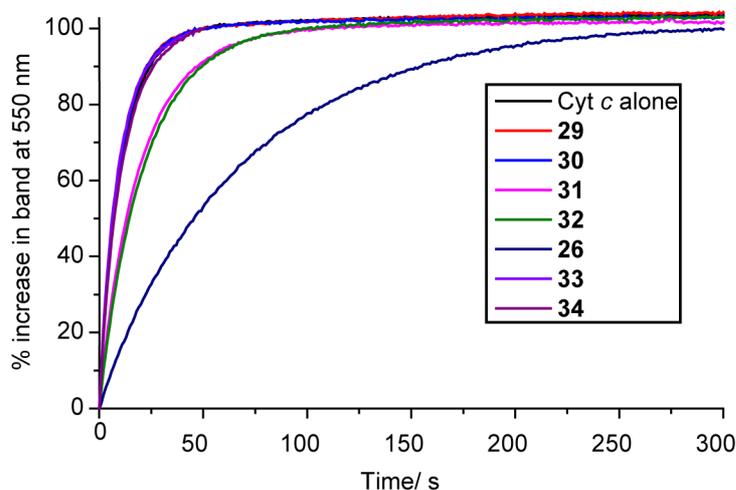


Figure 2.6 Ascorbate (0.75 mM) reduction of *cyt c* (16 μM) in the presence of various Ru(II)(bpy)₃ complexes (16 μM), measured by change in absorbance at 550 nm over time, 5 mM sodium phosphate, pH 7.4

An orthogonal assay to detect binding of these Ru(II)(bpy)₃ complexes to *cyt c* is to monitor the rate of ascorbate reduction of *cyt c* and see how the Ru(II)(bpy)₃ complexes affect this rate. In an oxidising environment the iron in the haem of *cyt c* exists as Fe(III), however it can be reduced to Fe(II) by addition of ascorbate. This reduction of the iron is accompanied with the appearance of a UV/Vis absorbance peak at 550 nm. The rate of

reduction can thus be seen using a UV/Vis spectrometer by measuring the absorption at 550 nm over time. The binding of a molecule to the surface of *cyt c*, close to the haem group, will block the approach of ascorbate and thus reduce the rate of ascorbate reduction.

This assay was performed for all the Ru(II)(bpy)₃ complexes (**Figure 2.6**) and, as can be seen, this mirrors the luminescence quenching assay results, whereby the most acidic Ru(II)(bpy)₃ complex **26** reduces the rate of *cyt c* reduction the most, and the Ru(II)(bpy)₃ complexes **31** and **32** reduce the rate less, but by similar amounts. The other Ru(II)(bpy)₃ complexes don't reduce the rate of ascorbate reduction, indicating that they aren't binding over the haem exposed edge or are binding with low affinity, as expected.

2.5 *Cyt c*/CCP PPI inhibition

The work presented in this section was performed by Dr. Maria Filby, aided by Dr. Michael Webb.

Previously these Ru(II)(bpy)₃ complexes had been hypothesised to bind at the haem exposed edge of *cyt c*, due to the location of the basic amino acid residues and the reduced rate of *cyt c* reduction by ascorbate. The haem-exposed edge is the location of the CCP binding site, therefore the Ru(II)(bpy)₃ complexes were hypothesised to inhibit the PPI.

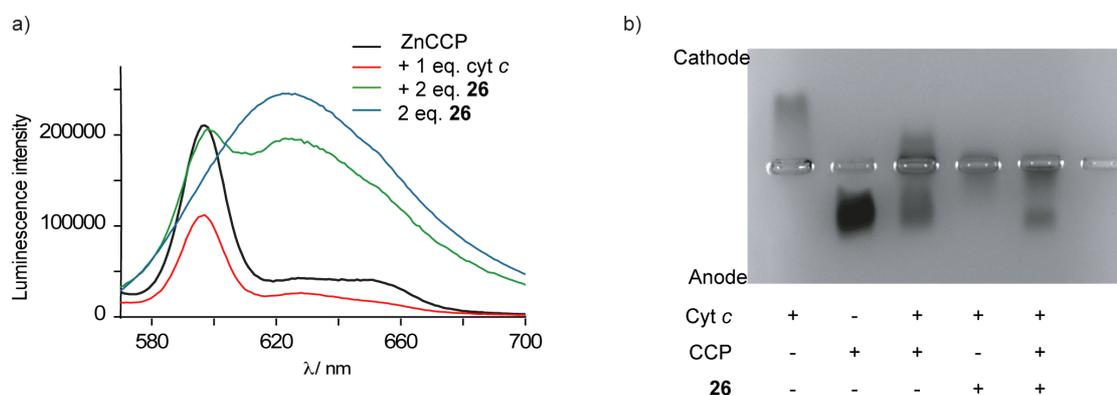


Figure 2.7 *Cyt c*/CCP PPI inhibition by Ru(II)(bpy)₃ complex **26**, a) Fluorescence spectrum of zinc substituted CCP alone (purple), with 1 eq. *cyt c* (red), with 1 eq. *cyt c* and 2 eq. Ru(II)(bpy)₃ complex **26** (green) and 2 eq. of Ru(II)(bpy)₃ complex **26** alone, b) Native agarose gel, showing the migration of *cyt c*, CCP and Ru(II)(bpy)₃ complexes

In order to test this hypothesis the fluorescence spectra of Zn-substituted CCP was taken alone, and when bound to *cyt c* (**Figure 2.7a**). Upon addition of *cyt c*, the fluorescence band of CCP at ~580 nm was quenched. On addition of Ru(II)(bpy)₃ complex **26** to this *cyt c*/CCP complex the fluorescence band at 580 nm was restored, along with

giving the appearance of the Ru(II)(bpy)₃ complex **26** luminescence emission band at ~635 nm. This luminescence band at ~635 nm is however quenched compared to the Ru(II)(bpy)₃ complex **26** alone, showing it as bound to cyt *c*. This indicated that the Ru(II)(bpy)₃ complex **26** displaced cyt *c* from CCP, thus inhibiting the PPI.

A native agarose gel was also performed, to corroborate these findings. Cyt *c* and CCP migrate towards the cathode and anode respectively, but when they complex together this migration was retarded. The presence of Ru(II)(bpy)₃ complex **26** retarded the migration of cyt *c* to the cathode, but the CCP is less affected (Lane 5). Thus, again, indicating the Ru(II)(bpy)₃ complex **26** binds to cyt *c*, inhibiting the cyt *c*/CCP interaction.

2.6 Binding in different conditions

Previously protein surface mimetics, including Ru(II)(bpy)₃ complexes, have been designed for binding to cyt *c* with charge-charge complementarity in mind, designing scaffolds functionalised with carboxylic acids in order to bind to the basic amino acid residues on cyt *c*. However little information has been obtained as to how these molecules bind to cyt *c*, and how this compares to its native protein partners, and thus if they are indeed mimicking the recognition of a native PPI, cyt *c*/CCP.

To find out more information as to how these molecules interact with cyt *c*, binding of the Ru(II)(bpy)₃ complexes **26** and **31** was tested in a variety of different conditions, using the previously described luminescence quenching. The effects of different conditions on cyt *c* binding of these Ru(II)(bpy)₃ complexes were compared to the binding of CCP to cyt *c*, allowing an understanding of whether the binding of the Ru(II)(bpy)₃ complex to cyt *c* indeed mimics this native PPI, to be developed.

2.6.1 Temperature

Binding of Ru(II)(bpy)₃ complexes **26** and **31**, using the luminescence quenching assay, was tested at different temperatures. This allows a van't Hoff analysis to be performed (Eq. 2.3) to determine the thermodynamic parameters of binding. The van't Hoff equation is derived from the Gibbs free energy definition (Eq. 2.1) and the Gibbs free energy isotherm equations (Eq. 2.2), where ΔH and ΔS are assumed to be temperature independent.

$$\Delta G = \Delta H - T\Delta S \quad \text{Eq. 2.1}$$

$$\Delta G = -RT\ln K \quad \text{Eq. 2.2}$$

$$\ln K = \frac{-\Delta H}{RT} + \frac{\Delta S}{R} \quad \text{Eq. 2.3}$$

ΔG = change in Gibb's free energy, ΔH = change in enthalpy, ΔS = change in entropy, T = temperature (in Kelvin), R = gas constant, K = binding constant

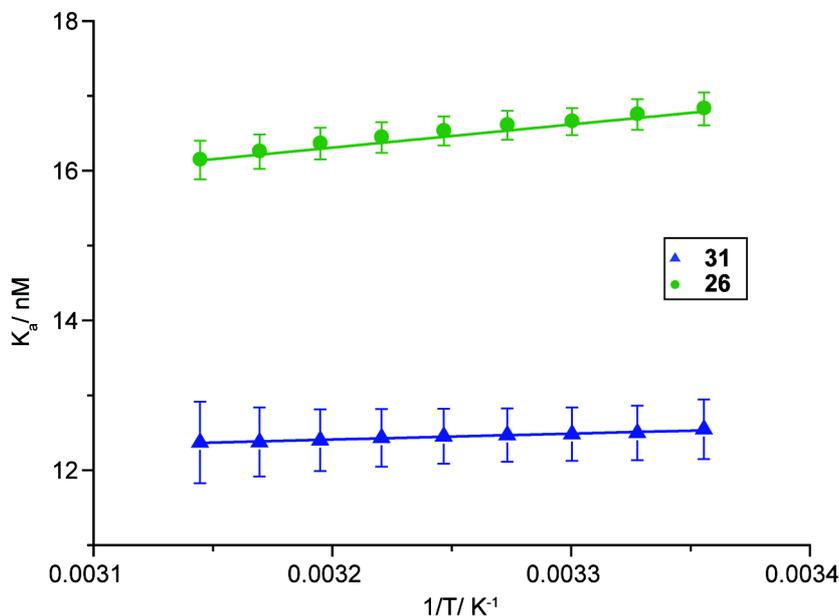


Figure 2.8 van't Hoff analysis for the binding of Ru(II)(bpy)₃ complexes **26** and **31** measured in 5 mM sodium phosphate, 0.2 mg mL⁻¹ BSA, pH 7.5

The binding of the two Ru(II)(bpy)₃ complexes do fit the van't Hoff equation (**Figure 2.8**), allowing the thermodynamic parameters (ΔH , ΔS and ΔG) to be derived (**Table 2.4**). The thermodynamic parameters for the two Ru(II)(bpy)₃ complexes were then compared to the native PPI to see if and how the Ru(II)(bpy)₃ complex/cyt *c* interaction mimics the native cyt *c*/CCP PPI.

Table 2.4 Thermodynamic parameters derived from the van't Hoff analysis for the binding of the Ru(II)(bpy)₃ complexes **26** and **31** to cyt *c* in 5 mM sodium phosphate, 0.2 mg mL⁻¹ BSA, pH 7.5, errors derived from triplicate experiments, and literature values for the cyt *c*/CCP interaction in similar conditions (10 mM potassium phosphate, pH 6.0, by ITC at 25 °C)¹⁵⁵

	31	26	CCP¹⁵⁵
ΔH / kJ mol ⁻¹	-6.6 ± 0.4	-26.3 ± 3.0	9.4 ± 0.8
T ΔS (25 °C) / kJ mol ⁻¹	24.5 ± 0.4	16.0 ± 3.0	38.4 ± 0.9
ΔG (25 °C) / kJ mol ⁻¹	-31.0 ± 0.4	-42.3 ± 0.0	-27.9 ± 1.0

The binding of Ru(II)(bpy)₃ complex **31** to cyt *c* is predominantly entropically driven whereas the binding of Ru(II)(bpy)₃ complex **26** is both entropically and enthalpically driven. In comparison, the binding of the native cyt *c*/CCP PPI is entropically controlled and is even mildly enthalpically unfavourable. Thus, Ru(II)(bpy)₃ complex **31**, with fewer carboxylates, is a closer match to cyt *c*'s interaction with its endogenous protein partner, CCP, showing this Ru(II)(bpy)₃ complex is indeed acting as a protein surface mimetic. The enhanced binding of Ru(II)(bpy)₃ complex **26**, is likely derived from the additional carboxylic acids forming increased numbers of salt bridges with the basic amino acids on the cyt *c* surface, increasing the enthalpic contribution to binding.

2.6.2 Ionic strength

To further understand the electrostatic contribution to binding, the binding affinities of Ru(II)(bpy)₃ complexes **26** and **31** were determined at different ionic strengths (I). The binding of both Ru(II)(bpy)₃ complexes **26** and **31** is highly dependent upon ionic strength (**Table 2.5**) with binding affinity decreasing with increasing ionic strength, suggesting electrostatic contributions dominate the binding of both of these Ru(II)(bpy)₃ complexes to cyt *c*, as would be expected. The binding of CCP to cyt *c* is also highly dependent on ionic strength,¹⁵⁷ indicating again that the Ru(II)(bpy)₃ complex/cyt *c* interaction is mimicking that of the native cyt *c*/CCP PPI.

Table 2.5 Binding of Ru(II)(bpy)₃ complexes **26** and **31** in variable ionic strengths, 5 mM sodium phosphate, 0.2 mg mL⁻¹ BSA, pH 7.5, variable concentration NaCl, n.d. = not determined

Ionic strength/ mM	31 K _d / μM	26 K _d / nM
8.39	2.88 ± 0.46	25.3 ± 2.4
13.39	4.25 ± 0.47	64.8 ± 13.7
18.39	10.30 ± 1.61	196.5 ± 59.2
28.39	20.23 ± 0.16	426.5 ± 59.8
48.39	n.d.	2040.9 ± 152.6

The K_d values were fit to the Debye-Hückel relationship (Eq. 2.4), (the theory behind and derivation of this is presented in the Appendix I) (**Figure 2.9**) in this case using a Güntelberg approximation (Eq. 2.5), which is valid up to I = 100 mM.

$$\log K_d = \log K_d^0 - 0.509 Z_1 Z_2 \mu \quad \text{Eq. 2.4}$$

$$\mu \approx \sqrt{\frac{I}{1+\sqrt{I}}} \quad \text{Eq. 2.5}$$

K_d = dissociation constant, K_d^0 = theoretical dissociation constant at 0 ionic strength, Z_1 and Z_2 = charges on species 1 and 2, μ = a function of the ionic strength (empirically derived), I = ionic strength

From this relationship the parameters K_d^0 and Z_1Z_2 can be established, providing an estimate of the binding affinity at $I = 0$ and the product of the interacting positive and negative charges, respectively. The data were consistent with the Güntelberg approximation for both Ru(II)(bpy)₃ complexes (**Figure 2.9**), giving a linear relationship. The calculated values of K_d^0 (**Table 2.6**) show high affinity binding for Ru(II)(bpy)₃ complex **26** and weaker binding for Ru(II)(bpy)₃ complex **31** at zero ionic strength, with the value for CCP being between these values.

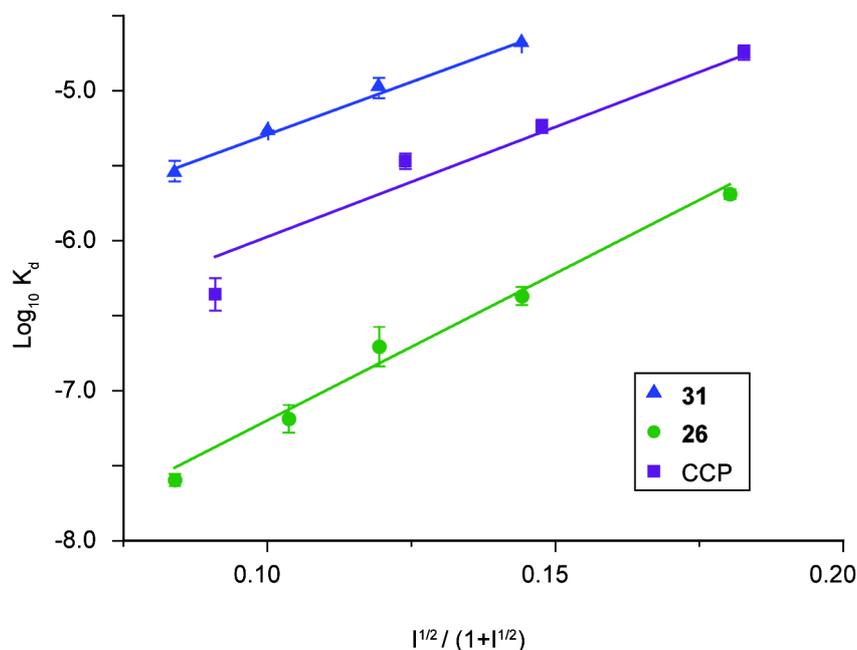


Figure 2.9 Debye-Hückel analysis for the binding of Ru(II)(bpy)₃ complexes **26** and **31** to cyt *c* and literature values for CCP for comparison, measured in 5 mM sodium phosphate, variable concentration NaCl, 0.2 mg mL⁻¹ BSA, pH 7.5

Table 2.6 Values derived from the Debye-Hückel approximation for Ru(II)(bpy)₃ complexes **26** and **31** and CCP¹⁵⁴

	31	26	CCP¹⁵⁴
K_d^0 / nM	253 ± 5	1.11 ± 0.21	40.7 ± 23.0
Z_1Z_2	25.9 ± 1.9	35.6 ± 1.3	28.8 ± 4.8

The Z_1Z_2 value provides an indication of the charges involved in the interaction, with Ru(II)(bpy)₃ complex **26** having a larger value than Ru(II)(bpy)₃ complex **31** and CCP. Using these data, the charge on the Ru(II)(bpy)₃ complex interacting with cyt *c* can be estimated. Assuming cyt *c* has the same charge in all cases (calculated to be ~ 6 at pH 7.5),¹⁵⁸ the charge on Ru(II)(bpy)₃ complex **26** and **31** and CCP can be calculated to be 5.9, 4.3 and 4.8, respectively. Ru(II)(bpy)₃ complex **31** and CCP have relatively similar charges, indicating similar interactions with cyt *c*, further showing that its binding to cyt *c* is mimicking that of the native PPI. Ru(II)(bpy)₃ complex **26** has a larger charge, indicating increased electrostatic interactions with cyt *c*, showing that it is possible to increase the interactions in a native PPI, in order to gain high affinity protein surface ligands. Accounting for the crudeness of the Debye-Hückel approximation where small (~ 3 Å), evenly dispersed charges are assumed (even when using the Güntelberg extension, which extends past the single point charges used in the first (Debye-Hückel) approximation), the data indicate that perhaps not all the carboxylates are deprotonated under the assay conditions (pH 7.5) and/or that a limited number of the carboxylates are needed for the protein surface recognition, even fewer than the 4 isophthalate arms as found by the Ohkanda group using heteroleptic complexes.¹⁴⁷

2.6.3 Binding in different buffers

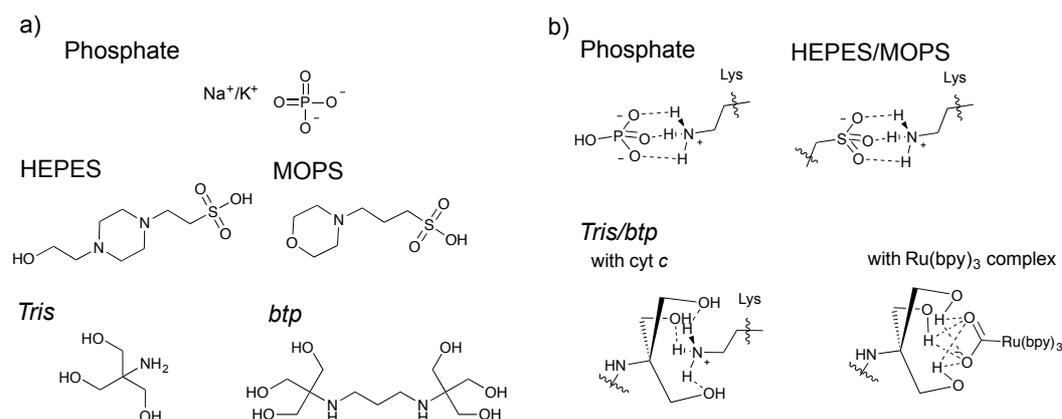


Figure 2.10 Buffers i) Structures of the buffer components, ii) Plausible interactions of the buffer components with the lysine residues on cyt *c*, and *Tris/btp* with carboxylates on the Ru(II)(bpy)₃ complex **26**

The binding in different buffers can also give a further indication as to what interactions are important for binding, as for binding to occur, negatively charged anions must be displaced from the positively charged surface of cyt *c*, and positively charged cations must be displaced from the negatively charged surfaces of the Ru(II)(bpy)₃

complexes. The binding affinity of Ru(II)(bpy)₃ complex **26** to cyt *c* was tested in a range of different buffers (**Table 2.7**), with different structures and potential interactions with cyt *c* and Ru(II)(bpy)₃ complex **26** possible (**Figure 2.10**).

Table 2.7 Cyt *c* binding affinities of Ru(II)(bpy)₃ complex **26** in a range of different buffers, all buffers at 5 mM, pH 7.5 with 0.2 mg mL⁻¹ BSA

Buffer	K _a / nM
Sodium phosphate	42.9 ± 3.1
Potassium phosphate	26.2 ± 3.1
MOPS	35.2 ± 3.1
HEPES	31.2 ± 3.1
Tris	106.3 ± 32.6
Bis-tris propane (btp)	133.5 ± 37.4

There is little difference in the binding affinity between Ru(II)(bpy)₃ complex **26** and cyt *c* in sodium and potassium phosphate buffer, indicating the interactions of the cationic buffer components with the Ru(II)(bpy)₃ complex **26** are not significant. There also is not a significant difference between the binding in the phosphate buffers and the sulfonic acid buffers (MOPS and HEPES), suggesting that the nature of the anion, and the hydrophobicity of the buffer are not significant. This reinforces the conclusions of the Debye-Hückel analysis, with the interaction being dominated by electrostatic contributions. For the *tris* buffers (*btp* and *tris*) a small decrease in binding affinity is seen compared to the other buffers; this is likely due to different interactions between the buffer and cyt *c*, and Ru(II)(bpy)₃ complex **26** and its chloride counterion. Both the ammonium and hydroxyl functionalities of the *tris* buffers may interact with cyt *c* and the Ru(II)(bpy)₃ complex **26** with the potential for chelating hydrogen bonding interactions (**Figure 2.10b**), this would diminish binding affinity, by masking both the carboxylic acid functionality on the Ru(II)(bpy)₃ complex as well as the lysine residues on cyt *c*.

2.6.4 pH

Cyt *c* is a stable protein that does not unfold over a wide range of pHs, however its ionisation state is affected by pH (**Figure 2.11b**),¹⁵⁹ hence the pH of the solution is expected to affect the interaction of cyt *c* with the Ru(II)(bpy)₃ complexes, especially as the interaction is driven by electrostatic interactions. The effect of pH on the binding

affinity of Ru(II)(bpy)₃ complex **26** to cyt *c* was investigated. *btp* was used for this study as it allows for a broad pH range (pH 6.5 - 9.5) to be used in the same buffer.

The binding affinity follows an inverted bell shaped profile (**Figure 2.11a**), which maps well onto the ionisation state of cyt *c* (**Figure 2.11b**).¹⁵⁹ The binding affinity is relatively constant between pH 7.0 and pH 8.5, with decreased binding affinity observed at pH 6.5 and pH 9.0. This again indicates electrostatics as a major contributor to the binding between the Ru(II)(bpy)₃ complex and cyt *c*.

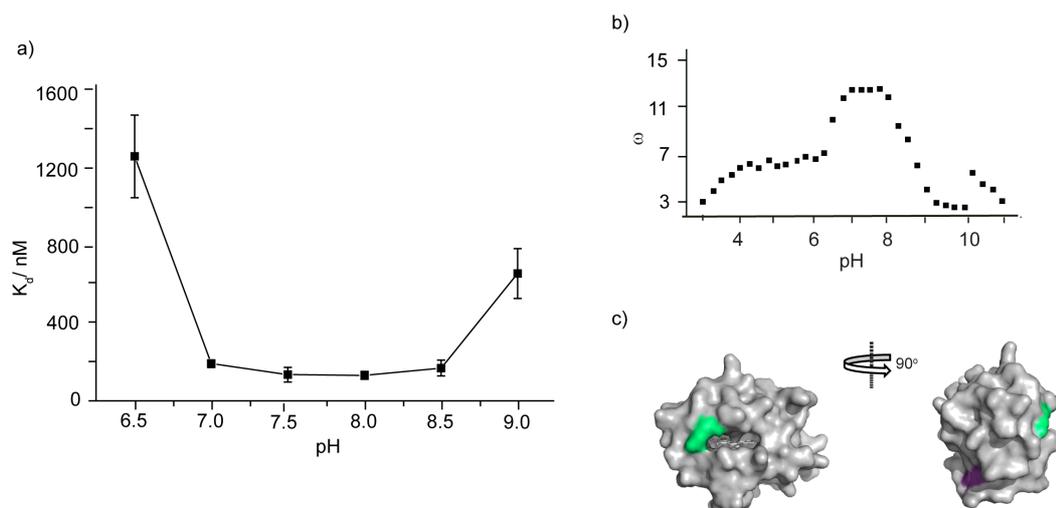


Figure 2.11 Effect of pH on the binding of Ru(II)(bpy)₃ complex **26** to cyt *c*. a) Binding affinity over the range pH 6.5 – 9.0 (5 mM *btp*, 0.2 mg mL⁻¹ BSA, variable pH), b) the electrostatic interaction factor (ω) of cyt *c* over a range of pHs (base limb of titration curve)¹⁵⁹, c) Cyt *c* structure (PDB ID 1HRC)¹⁶⁰ with residues that become protonated/deprotonated at pH 6.5 and 9.0, His-33 (pink) and Lys-79 (green) respectively

From these data, it is also possible to glean information as to a potential binding site, as at specific pHs different amino acids become protonated or deprotonated, and thus are changing their role in the binding interaction. The amino acid residues that become protonated/deprotonated at pH 6.5 and 9.0 are His-33 and Lys-79 respectively (**Figure 2.11c**).¹⁵⁹ Lys-79 is located on the haem-exposed edge, where binding of Ru(II)(bpy)₃ complex **26** is hypothesised to occur, whereas His-33 is on the distal face of cyt *c*. There are a number of different reasons as to the decrease in binding at this pH despite hypothesised binding at a different site: i) Ru(II)(bpy)₃ complex **26** binds to different or multiple sites on cyt *c*, ii) the protonation of His-33 causes a small conformational change in the cyt *c* structure, affecting the binding interactions on the haem-exposed edge, iii) the protonation state of Ru(II)(bpy)₃ complex **26** is changed at pH 6.5, affecting its binding interaction. Upon closer inspection of the pH/K_d and pH/ionisation state relationships,

there is a slight discrepancy with the cyt *c* ionisation state dropping at pH 8.0 rather than pH 8.5, where the binding affinity diminishes, whereas this difference does not exist at the lower pH, with both binding affinity and ionisation state decreasing at pH 6.5. This implies that the binding of Ru(II)(bpy)₃ complex **26** masks Lys-79 and increases its pK_a, whereas the His-33 protonation state is not affected by the binding of Ru(II)(bpy)₃ complex **26**. Therefore the loss in binding affinity observed at pH 6.5 is likely to arise from changes in ionisation state of the Ru(II)(bpy)₃ complex **26**, rather than that of His-33.

2.6.5 Conclusions of conditions screen

From the data obtained from screening the cyt *c* binding under different conditions it is possible to draw some conclusions as to how the Ru(II)(bpy)₃ complexes **26** and **31** bind to cyt *c*. The binding for both Ru(II)(bpy)₃ complexes is electrostatically driven, with the binding of Ru(II)(bpy)₃ complex **31** being entropically favourable and that of Ru(II)(bpy)₃ complex **26** being both entropically and enthalpically favourable. Compared to the native cyt *c*/CCP interaction, an entropy controlled, electrostatic interaction, the binding of Ru(II)(bpy)₃ complex **31** acts as a closer mimic. Increased enthalpic contributions for the binding of Ru(II)(bpy)₃ complex **26** arise from increased electrostatic interactions due to an increased number of carboxylates, allowing for further interactions over that in the native PPI. This shows it may be possible to use information on the binding interface of a known PPI, and enhance the interactions present when designing a molecular ligand in order to gain a high affinity protein surface mimetic.

2.7 NMR spectroscopy

From the UV/Vis ascorbate reduction data (2.4.1) it can be established that the Ru(II)(bpy)₃ complexes bind close to the haem group on cyt *c*, as otherwise there would be no reduction in the rate of cyt *c* reduction. Similarly, the pH profile data indicate binding at the haem-exposed edge, close to Lys-79. However, detailed information as to the binding site of these Ru(II)(bpy)₃ complexes had not been previously established. In order to study where these Ru(II)(bpy)₃ complexes are binding and to gain a fuller picture of the binding interaction protein NMR spectroscopy was performed. In this case a ¹H-¹⁵N HSQC was used, where the backbone amide bond N-H cross-correlations are monitored, and the spectrum of cyt *c* with and without the Ru(II)(bpy)₃ complex present compared. The shifting of cross-peaks on addition of the Ru(II)(bpy)₃ complex suggest binding close to the amino acid residue corresponding to that cross-peak, whereas if the cross-peak does not shift it indicates the Ru(II)(bpy)₃ complex binds at a distal site to that residue.

Ordinarily when performing protein NMR spectroscopy, a protein labelled with NMR-active ^{15}N and/or ^{13}C isotopes is required, dependent on the experiment being performed. The Astbury Centre had recently acquired a new 950 MHz NMR spectrometer so it was decided to see if this machine could detect naturally abundant levels of ^{15}N in an HSQC experiment, both as a test for the instrument and to see if it was possible to detect binding of the Ru(II)(bpy) $_3$ complexes to cyt *c*. Cyt *c* is a good test for this capability as it can be bought in large (gram) quantities, and studied in high (millimolar) concentration solutions, as is required for the detection of the low natural abundance of ^{15}N . NMR spectra were obtained with the help of Dr. Lars Kuhn and Dr. Arnout Kalverda.

2.7.1 Oxidised/reduced cyt *c*

One problem that could be observed in the NMR spectrum of cyt *c* is due to the presence of the haem iron. In the oxidised Fe(III) state present in normal, oxygen-rich, conditions the Fe(III) has a d^5 high spin electronic configuration; this has unpaired electrons, thus making it paramagnetic. Paramagnetism generates a local magnetic field, which leads to interference in the NMR spectrum. In the case of a protein NMR spectra this leads to line broadening of the resonances for atoms close to the paramagnetic atom. As binding was expected to be on the haem exposed edge, by the haem iron, it was thought that this may complicate the NMR spectrum in the region of interest. Therefore the Fe(III) was reduced to Fe(II) in order to run the NMR spectroscopy experiment. The simplest way to do this is to add sodium ascorbate to the buffer.

As a different buffer and oxidation state of cyt *c* was going to be used in the NMR spectroscopy experiment, the binding of Ru(II)(bpy) $_3$ complex **26** to cyt *c* under the reducing conditions was tested. A comparison was also made between chemically oxidised cyt *c* and reduced cyt *c* in the same buffer in order to see the effect cyt *c* reduction has on the binding affinity. Irreversibly, chemically oxidised cyt *c* was obtained by addition of $\text{K}_3\text{Fe}(\text{CN})_6$.

Table 2.8 Binding constants for Ru(II)(bpy) $_3$ complex **26** binding to oxidised and reduced cyt *c* (5 mM sodium phosphate, 0.1 mM sodium ascorbate, 0.2 mg mL $^{-1}$, pH 7.5), and the native PPI for comparison (3.5 mM potassium phosphate, pH 7.5)

	K_d / nM	CCP 154 K_d / nM
Oxidised, Fe$^{3+}$	49.6 \pm 13.3	440 \pm 110
Reduced, Fe$^{2+}$	92.4 \pm 5.5	690 \pm 200

The luminescence quenching assay was used to test the binding of Ru(II)(bpy)₃ complex **26** to the oxidised and reduced cyt *c*. In this case, a lower concentration sodium ascorbate buffer than was to be used in the NMR spectroscopy (0.1 mM rather than 2 mM) was employed due to the concentration of oxidised cyt *c* stock being relatively low (0.508 mM), as a result of dilution during dialysis. As it is known that at high ionic strength the binding affinity is relatively weak it was decided to assess this binding in a lower ionic strength buffer to allow for the use of this cyt *c* stock. As can be seen in **Table 2.8** the binding affinity is similar between both the oxidised and reduced cyt *c*, as it is with the native PPI. This means that using the reduced cyt *c* to obtain the NMR spectrum is a valid approach.

2.7.2 ¹H-¹⁵N HSQC of cyt *c* alone

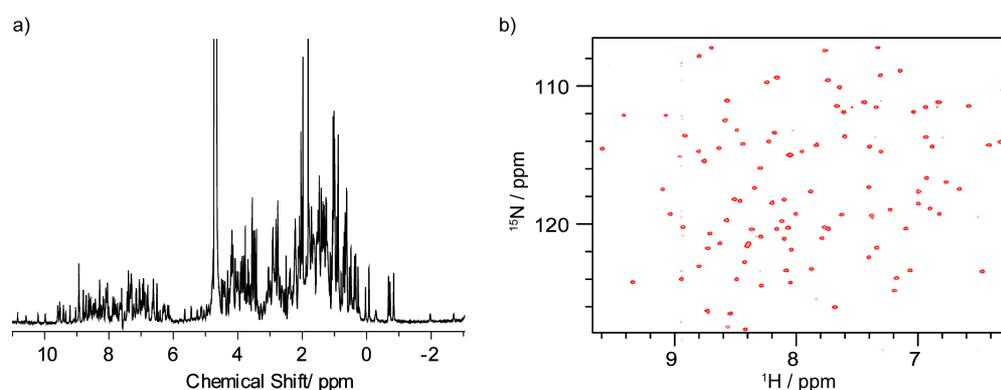


Figure 2.12 NMR spectra of cyt *c* alone a) 1-D ¹H spectrum, b) Natural abundance ¹H-¹⁵N HSQC spectrum, both 2 mM cyt *c*, in 5 mM sodium phosphate, 2 mM sodium ascorbate, 10 % D₂O, pH 7.3 buffer

Following a promising ¹H 1-D spectrum (**Figure 2.12a**) of 2 mM cyt *c* in the 5 mM sodium phosphate, 2 mM sodium ascorbate, 10 % D₂O, pH 7.3 buffer, a ¹H-¹⁵N HSQC was obtained. This gave clearly defined cross-peaks (**Figure 2.12b**) which could be assigned using a previously reported HSQC assignment.¹⁶¹ This shows the utility of the 950 MHz NMR spectrometer, showing that it is possible to obtain good, assignable ¹H-¹⁵N HSQC spectra using natural abundance ¹⁵N.

2.7.3 ¹H-¹⁵N HSQC spectrum of Ru(II)(bpy)₃ complex **31** with cyt *c*

Having obtained a good spectrum for cyt *c* alone, a ¹H-¹⁵N HSQC spectrum of 1 mM cyt *c* with 0.5 mM Ru(II)(bpy)₃ complex **31** was obtained (**Figure 2.13**). This spectrum could similarly be assigned, with some cross-peaks having stayed in the same place, others having shift changes ranging from 0.015 - 0.05 ppm and others disappearing. This

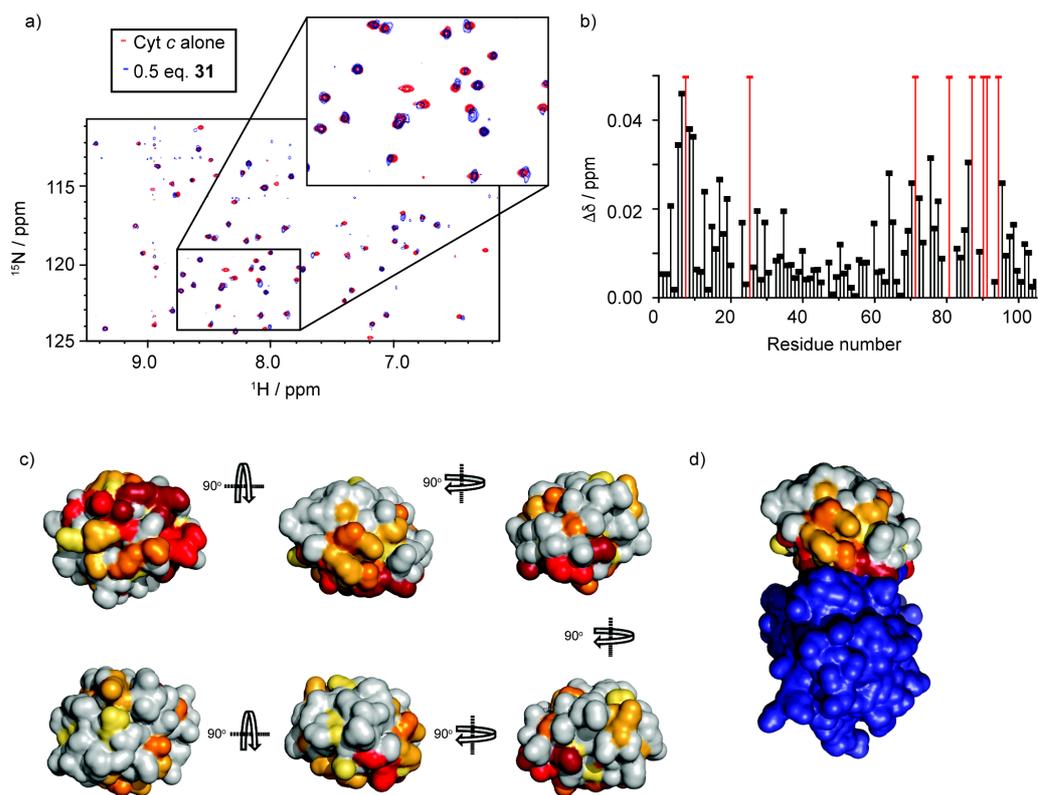


Figure 2.13 ^1H - ^{15}N HSQC NMR data of $\text{Ru}(\text{II})(\text{bpy})_3$ complex **31** binding to *cyt c*. a) Region of the overlaid ^1H - ^{15}N HSQC spectrum of *cyt c* (red) and *cyt c* with 0.5 eq $\text{Ru}(\text{II})(\text{bpy})_3$ complex **31** (blue), Inset shows zoom in of part of the spectrum, showing some cross-peaks staying the same, some having shifted and one disappearing. b) ^1H - ^{15}N chemical shift differences ($\Delta\delta$) for the different amino acid residues with and without $\text{Ru}(\text{II})(\text{bpy})_3$ complex **31**. Gaps are for prolines, unassigned amino acids, red peaks are amino acids for which the signal disappears (arbitrarily $\Delta\delta$ shown to be 0.05) - due to significant line-broadening of NH cross-peaks - on addition of $\text{Ru}(\text{II})(\text{bpy})_3$ complex **31**. c) Chemical shift perturbation map of *cyt c*, molecular surface of *cyt c* generated from PyMol (PDB ID 1U75),¹⁶⁰ with colouring corresponding to the extent of chemical shift changes ($\Delta\delta$) on addition of the $\text{Ru}(\text{II})(\text{bpy})_3$ complex **31**. Amino acid residues with ^1H - ^{15}N resonances that disappear in dark red, exhibit large chemical shift changes ($\Delta\delta > 0.03$) in red, moderate changes ($\Delta\delta > 0.02$) in orange, small changes ($\Delta\delta > 0.015$) in yellow-orange and very small chemical shift changes ($\Delta\delta > 0.01$) in yellow. d) perturbation map of *cyt c* (as in c), corresponding to the top central structure) in complex with CCP (purple), (PDB ID 1U75)¹⁴⁸

indicates the presence of protein-ligand interactions. These chemical shift changes were mapped onto the structure of *cyt c* where they indicate binding to one side of the haem group of *cyt c*, with the opposite face having very few amino acid residues with sizeable shifts in their HSQC cross-peaks (**Figure 2.13c**), indicating a binding site to one side of the haem exposed edge. The binding site identified here is in a similar location to that of carboxylate functionalised porphyrins, as determined by the Crowley group.²⁶ Mapping these amino acid residues onto the *cyt c*/CCP PPI structure (**Figure 2.13d**) it can be seen

that the amino acid residues whose cross-peaks have shifted are in and around the cyt *c*/CCP PPI interface, indicating Ru(II)(bpy)₃ complex **31** binds in the same region as CCP, and so indeed could be mimicking this PPI.

2.7.4 ¹H-¹⁵N HSQC spectrum of Ru(II)(bpy)₃ complex **26** with cyt *c*

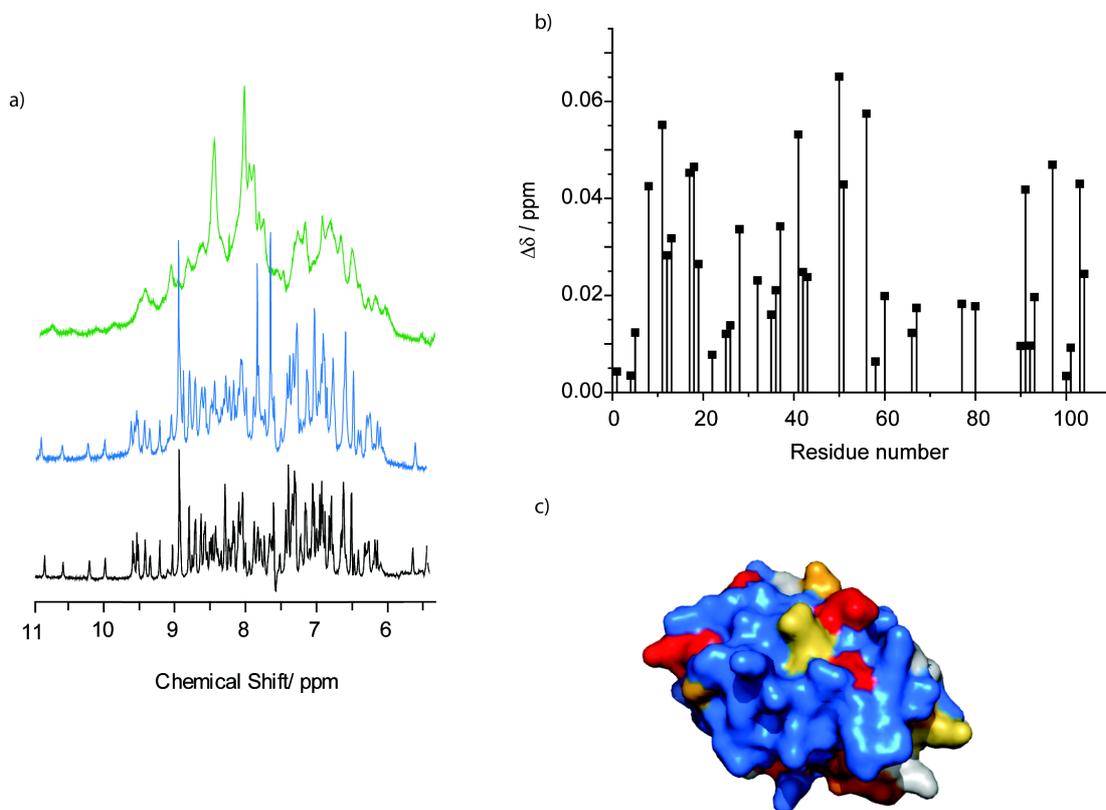


Figure 2.14 NMR spectra for the binding of Ru(II)(bpy)₃ complex **26** to cyt *c* a) 1-D ¹H NMR spectrum of cyt *c* alone (black), cyt *c* with 0.5 eq Ru(II)(bpy)₃ complex **31** (blue) and with 0.5 eq Ru(II)(bpy)₃ complex **26** (green), b) ¹H-¹⁵N chemical shift differences ($\Delta\delta$) for the different amino acid residues with and without Ru(II)(bpy)₃ complex **26**. Gaps are for prolines, unassigned amino acids, and signals that disappear, c) Chemical shift perturbation map of cyt *c* with Ru(II)(bpy)₃ complex **26**, molecular surface of cyt *c* generated as in **Figure 2.13**, except showing amino acid residues whose cross-peaks disappear in blue. This view is of the binding site of Ru(II)(bpy)₃ complex **31**

Attempts were also made to obtain spectra with both 0.5 and 1 equivalents of Ru(II)(bpy)₃ complex **26**. On obtaining a 1-D ¹H NMR spectrum (**Figure 2.14a**) (green) it could be seen that the peaks have broadened compared to both the cyt *c* alone (black) and with 0.5 equivalents of Ru(II)(bpy)₃ complex **31** (blue). This could indicate the formation of a larger species, for example many Ru(II)(bpy)₃ complexes **26** binding to cyt *c* or the Ru(II)(bpy)₃ complex(es) **26** causing oligomerisation of cyt *c*. This is perhaps unsurprising as carboxylate functionalised porphyrins have been shown to have multiple binding sites

on cyt *c*, and to multimerise cyt *c* at high (millimolar) concentrations.¹⁰³ Upon attempting a ¹H-¹⁵N HSQC of cyt *c* with 0.5 equivalents of Ru(II)(bpy)₃ complex **26**, only 41 % of the cross-peaks (**Figure 2.14b**) present in the cyt *c* alone were present in the spectrum, compared to 93 % with 0.5 eq. of Ru(II)(bpy)₃ complex **31**. This means that detailed information as to the binding site could not be gleaned, for example a view of the Ru(II)(bpy)₃ complex **31** binding site is shown in **Figure 2.14c**, with the shifts for Ru(II)(bpy)₃ complex **26** shown as varying colours dependent on shift difference, and blue indicating signals that have disappeared. This again indicates the formation of larger species, or potentially that the structure of the cyt *c* is disrupted, however previous studies from the Wilson group have shown that Ru(II)(bpy)₃ complex **26** does destabilise cyt *c* but does not change the structure at 20 °C, the temperature at which the NMR spectroscopy was run.¹⁴⁴

2.7.5 *S. cerevisiae* cyt *c* binding

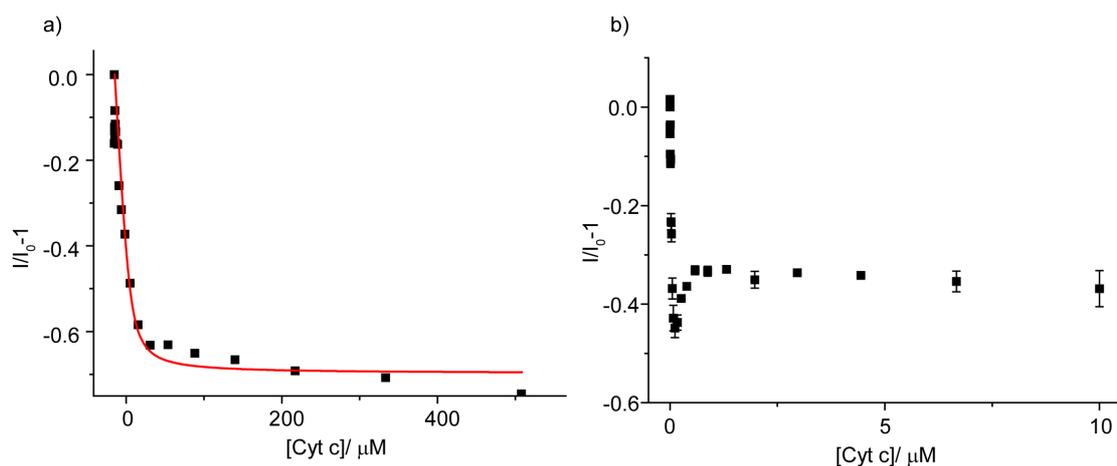


Figure 2.15 Binding isotherms for a) Ru(II)(bpy)₃ complex **31** and b) Ru(II)(bpy)₃ complex **26** with cyt *c* from *S. cerevisiae*

Further evidence for multiple binding sites for Ru(II)(bpy)₃ complex **26** on cyt *c* has been obtained by looking at the binding of Ru(II)(bpy)₃ complex **26** to cyt *c* from a different species. So far all the binding discussed has been with horse heart cyt *c*, with data fitting to a 1:1 binding isotherm. On testing the binding of Ru(II)(bpy)₃ complex **31** with *S. cerevisiae* (yeast) cyt *c*, a similar binding isotherm (**Figure 2.15a**) with similar K_d (2.81 ± 0.68 μM compared to 2.58 ± 0.72 μM with horse heart cyt *c*) was obtained, as would be expected, especially for such an evolutionarily conserved protein. However, looking at the binding of Ru(II)(bpy)₃ complex **26** to the *S. cerevisiae* cyt *c*, a different shaped binding isotherm was obtained (**Figure 2.15b**). This can be hypothesised to be due to a second binding event, and is consistent with the NMR data with such broadening of peaks all over the protein.

2.8 Conclusions

The synthesis and cyt *c* binding of 7 different Ru(II)(bpy)₃ complexes has been established. A luminescence quenching assay was used to show that increasing numbers of carboxylates correlate with increased cyt *c* binding affinity. A UV/Vis cyt *c* reduction assay corroborated these result. The binding of two of these Ru(II)(bpy)₃ complexes **26** and **31** was assessed in varying conditions, the effect of these differing conditions was compared to that with one of cyt *c*'s native protein partners, CCP, for which Ru(II)(bpy)₃ complex **26** had been shown to inhibit the PPI. This showed that the smaller Ru(II)(bpy)₃ complex **31** bound in a similar manner to CCP with binding being an electrostatic, entropy driven process, whereas the binding of the larger Ru(II)(bpy)₃ complex **26** was found to be both entropy and enthalpy driven with the increased enthalpic contributions arising from increased numbers of electrostatic interactions. Natural abundance ¹H-¹⁵N HSQC NMR spectra were obtained for both cyt *c* alone and Ru(II)(bpy)₃ complex **31** bound to cyt *c*, showing binding occurring at the cyt *c*/CCP binding interface, and indicating that the binding of Ru(II)(bpy)₃ complex **31** to cyt *c* does indeed mimic that of CCP. This shows that using known information about a native PPI it is possible to mimic the PPI in order to achieve ligands for one of the protein partners. The enhanced binding of Ru(II)(bpy)₃ complex **26**, shows it is possible to enhance the known binding interactions present in the native PPI to achieve high affinity ligands for one protein partner. This knowledge could be used for the design of new, high affinity, ligands for known PPIs of therapeutic interest.

3 Design of multivalent $Ru(II)(bpy)_3$ and porphyrin scaffolds for dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCC) is an approach whereby a thermodynamic equilibrium of many competing reversible chemical reactions is generated, which can be exploited for receptor discovery. The equilibrium is established, in the presence and absence of a template, in order to determine which compounds out of the dynamic combinatorial library (DCL) generated are amplified and therefore bind to that template (**Figure 3.1**).^{162–165} If a protein template is used, ligands for that protein template may be generated, with the highest affinity ligands being amplified, in the presence of the protein, compared to in its absence. This allows for the selection of high affinity ligands from an array of potential candidates.

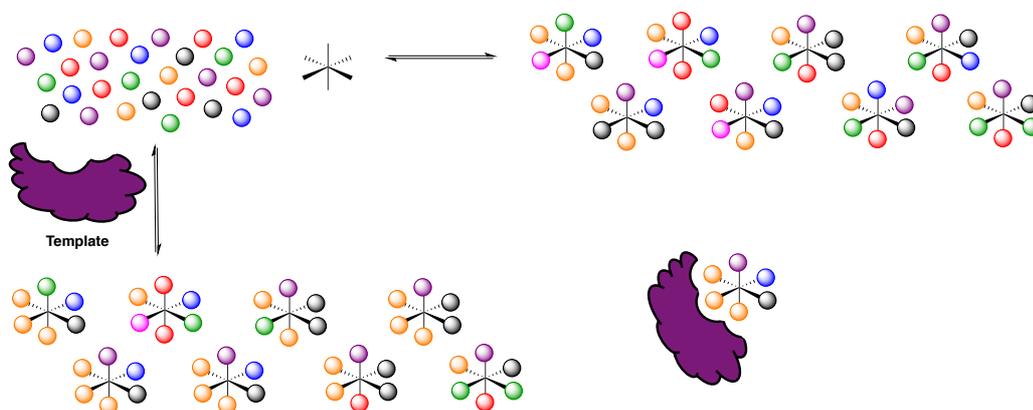


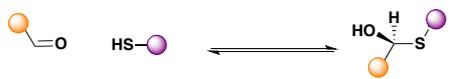
Figure 3.1 Cartoon schematic of DCC around a multivalent scaffold. A DCC scaffold is incubated with different groups that can reversibly covalently attach around the scaffold under thermodynamic equilibrium conditions. This thermodynamic equilibrium is reached in the absence and presence of a template protein; the molecules that bind to this template become enriched in the presence of the template compared to without, as they become more thermodynamically stable.

A major limitation in the identification of protein surface mimetics, is the generation of large multivalent compounds with differing binding groups projected around the surface. DCC could allow the reversible sampling of many different groups for their protein binding affinity around the protein surface mimetic scaffold, thus allowing the development of structurally and compositionally diverse protein surface mimetics. To this end, the design of $Ru(II)(bpy)_3$ complex and tetraphenyl-porphyrin DCC scaffolds has been attempted, allowing for the potential use of two different protein surface mimetic scaffolds in biologically relevant DCLs.

3.1 Biologically compatible DCC

As the aim of using these multivalent dynamic systems is to generate multivalent protein ligands, the system must be biologically compatible. For DCC to be biologically compatible, the reactions must be performed in an aqueous environment, at neutral pH, at an appropriate temperature and must reach equilibrium within a reasonable time frame.¹⁶⁶

Table 3.1 Examples of biologically compatible DCC reactions¹⁶⁶

Reaction name	Reaction	Protein examples
Imine formation		SARS coronavirus SARS-CoVM ^{pro}
Hydrazone formation		γ -amino butyric acid transporter 1 ¹⁶⁷
Acyl hydrazone formation		Glutathione S-transferase ¹⁶⁸
Hemithioacetal formation		β -galactosidase ¹⁶⁹
Thioether formation		Glutathione S-transferase ¹⁷⁰
Disulfide formation		Concanavalin A ¹⁷¹
Boronate ester formation		2-oxoglutarate dependent oxygenases ¹⁷²
Metal ligand coordination	$M + L \rightleftharpoons M-L$ (e.g. M-L = Fe(II)-bpy and Re-S)	Lectins ¹³² (Fe(II)-bpy)

Examples of biologically compatible DCC reactions, are shown in **Table 3.1**. Many of these use the same types of linkages with imines and various sulfur containing bonds being common.

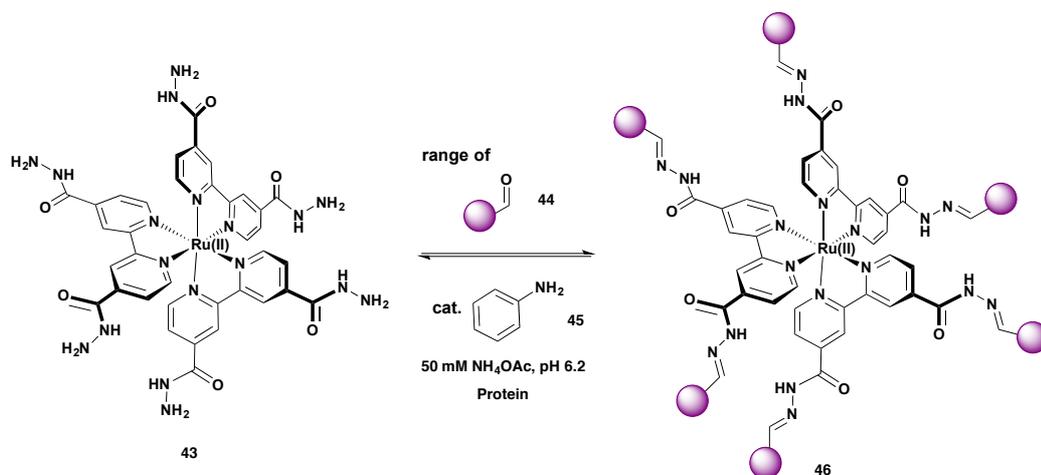
DCLs have also been designed where the system is first equilibrated in a non-biologically relevant context, and the mixture of products obtained tested against the biological molecule of interest. The procedure can then be repeated with different components missing, to determine which components of the mixture contribute to

binding. This allows a wider range of reaction types to be used, but involves many more DCL experiments, and is not a truly adaptive system.

Previously, in the Wilson group, an Fe(II)(bpy)₃ system, akin to that used by the Sasaki and de Mendoza groups described in the Chapter 1,^{131,132} was developed,¹⁷³ but this methodology did not prove to be reproducible. Therefore a different type of reversible reaction, was attempted around a Ru(II)(bpy)₃ core. The multivalent nature of the scaffolds mean that the use of thiol-based DCC would be problematic due to the potential for polymerisation of the scaffold by disulfide bond formation. Acyl hydrazones were chosen as they do not require a fixation reaction at the end of the DCL generation, as required, for example, for imine formation. Acyl hydrazides are also facile to synthesise from methyl esters, for which Ru(II)(bpy)₃ functionalised species have previously been synthesised.¹⁴²

3.2 Ru(II)(bpy)₃ scaffolds

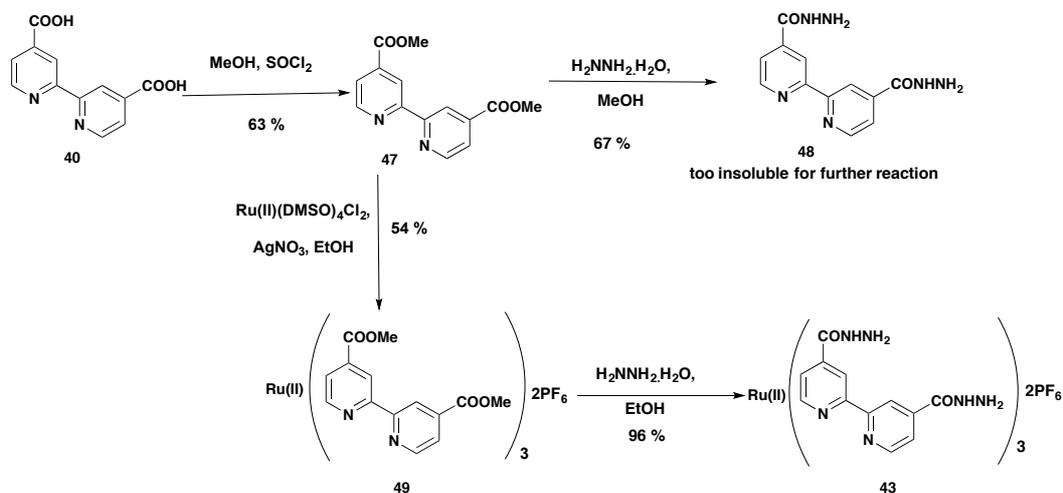
3.2.1 Initial scaffold



Scheme 3.1 Hypothesised initial DCC system

A hydrazide Ru(II)(bpy)₃ scaffold **43**, initially directly attached to the Ru(II)(bpy)₃ core was chosen as a good starting point. This would allow for 6 different hydrazones to form around the central Ru(II)(bpy)₃ complex core (**Scheme 3.1**). It was hypothesised that conditions previously established for hydrazone exchange in biological media by the Greaney group could be used.¹⁶⁸ Here the hydrazone exchange is performed in a biological buffer with aniline **45** acting as a nucleophilic catalyst, catalysing the hydrazone exchange at pH 6.2, as opposed to the acidic conditions generally required. This would allow the DCL to be incubated with a protein in order to select for a protein ligand.

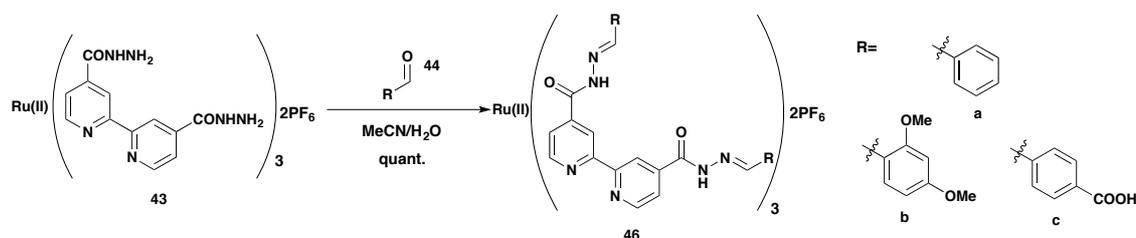
3.2.1.1 Synthesis



Scheme 3.2 Synthesis of the initial Ru(II)(bpy)₃ complex hydrazide scaffold **43**

Hydrazides can readily be formed by the reaction of hydrazine monohydrate with a methyl or ethyl ester. Initially this was attempted on the methyl ester ligand **47** (**Scheme 3.2** top), however the hydrazide ligand **48** formed was only soluble in very strong acids, preventing further reaction to form the hydrazide Ru(II)(bpy)₃ complex **43**. The hydrazide Ru(II)(bpy)₃ complex **43**, however, could easily be formed from the methyl ester Ru(II)(bpy)₃ complex **49** (**Scheme 3.2** bottom), by refluxing it with hydrazine monohydrate in methanol. This gave the hydrazide Ru(II)(bpy)₃ complex **43** as a methanol insoluble, but water-soluble solid, allowing it to be separated from the excess hydrazine. The hydrazide Ru(II)(bpy)₃ complex **43** was found to be oxygen sensitive, degrading over the period of a few hours, thus this compound could not be stored. Degradation of hydrazides with atmospheric oxygen has previously been reported, with first row transition metal cations increasing the rate of this reaction.¹⁷⁴

3.2.1.2 Hydrazone formation



Scheme 3.3 Hydrazone formation on initial hydrazide Ru(II)(bpy)₃ complex **43**

Various hydrazone Ru(II)(bpy)₃ complexes could be formed from the hydrazide Ru(II)(bpy)₃ complex **43** (**Scheme 3.3**), with almost immediate reaction on addition of the appropriate aldehyde **44**. Benzaldehyde derivatives were used as they lack α-protons and hence cannot form enamines, which could subsequently react with the excess aldehyde present. These hydrazone Ru(II)(bpy)₃ complexes **46** were stable, allowing the compound to be stored and used for analysis of the hydrazone exchange required for the generation of a DCL.

3.2.1.3 Hydrazone exchange

Prior to the generation of a complex DCL it was necessary to show that hydrazone exchange occurs. High resolution mass spectrometry (HRMS) was used to show this exchange, as it requires very little material and measurements can be taken at various time points from the same reaction mixture. In following these species by mass spectrometry an assumption that all species ionise equally so that their proportion in the mass spectrum is the same as that in the solution is made. In the case of the Ru(II)(bpy)₃ complexes, we can -to some extent- assume that the species ionise in similar proportions to their presence in the solution, as they exist as 2+ species and so are already ionised. The ionisation is therefore much less dependent on the ionisation susceptibility of different functional groups, as would be required of an organic molecule.

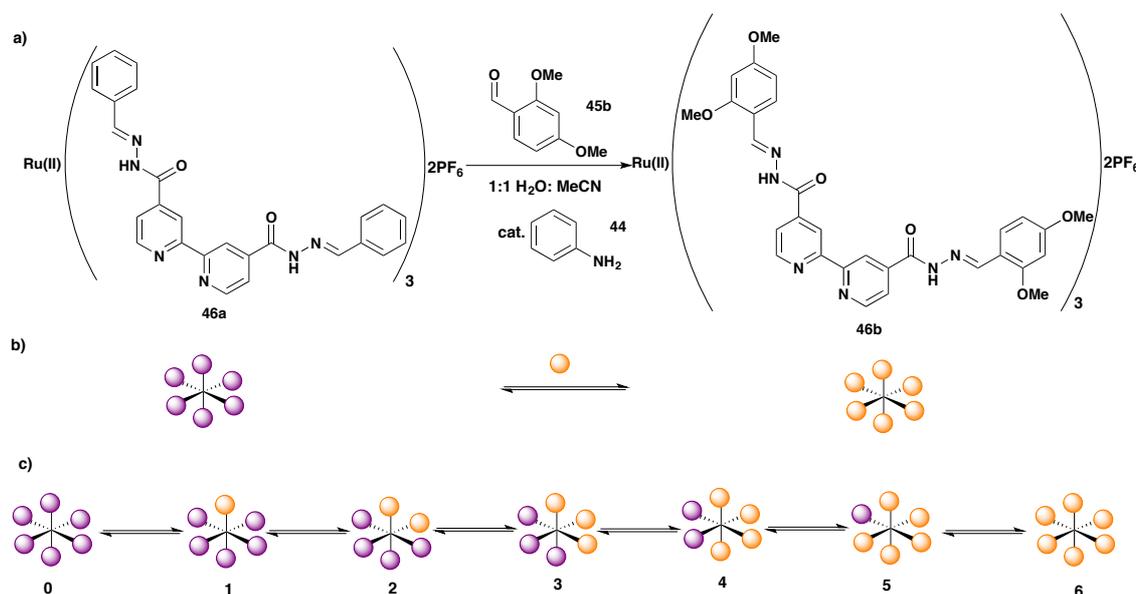


Scheme 3.4 Radical cleavage of N-N bond in mass spectrometer

Obtaining the mass spectrum of the various different hydrazone Ru(II)(bpy)₃ complexes **46** formed revealed the M²⁺ ion peak as expected, but also revealed peaks for radical degradation of the Ru(II)(bpy)₃ complex in the mass spectrometer, through cleavage of the N-N bond (**Scheme 3.4**). This ties in well with the instability of the hydrazide Ru(II)(bpy)₃ complex in oxygen as this also relies on the breaking of the N-N bond using an oxygen single electron. Knowledge of this, however, allows all peaks in the mass spectrum can be assigned.

A benzaldehyde hydrazone Ru(II)(bpy)₃ complex **46a** was incubated with 100 equivalents (eq.) (16.7 eq. per hydrazone) of 2,4-dimethoxy benzaldehyde **45b**, in order to see if hydrazone exchange could occur and on what time scale any hydrazone exchange occurs (**Scheme 3.5**). For initial tests an aniline **44** catalyst was used, as this was observed to catalyse the hydrazone exchange, by Greaney *et al.*¹⁶⁸ This exchange was performed in

1:1 acetonitrile:water, as this gave good solubility of all species involved, and provided a test to see if the hydrazone exchange could indeed occur without the addition of an acid.



Scheme 3.5 Hydrazone exchange a) Reaction to be performed, b) Cartoon schematic of reaction, c) Cartoon schematic of individual hydrazone exchanges

As can be seen in **Figure 3.2a**, there was a change in the mass spectrum from time 0 and after 4 days, with many hydrazone exchanges having occurred. However complete hydrazone exchange, as would be expected for the addition of 100 eq. (over 15 equivalents per hydrazone) was not observed. Following this hydrazone exchange over a time course (**Figure 3.2b**) it can be seen that the initial species **46a** (0) decreased over time and the species from one hydrazone exchange (1) at first increased, and then decreased. The species from subsequent hydrazone exchanges (2, 3, 4 etc.) increased over time at decreasing rates for later hydrazone exchanges. However, the system did not reach thermodynamic equilibrium after 32 hours, as this would result in the proportion of each species remaining constant (flat-lining), which was not observed. In terms of a biological system, the use of such a system would take too long, and the protein would potentially not be stable for this length of time. Incubation of the same system without a catalyst showed no hydrazone exchange occurring in the same time period, showing the requirement for the aniline **44** catalyst for hydrazone exchange to occur.

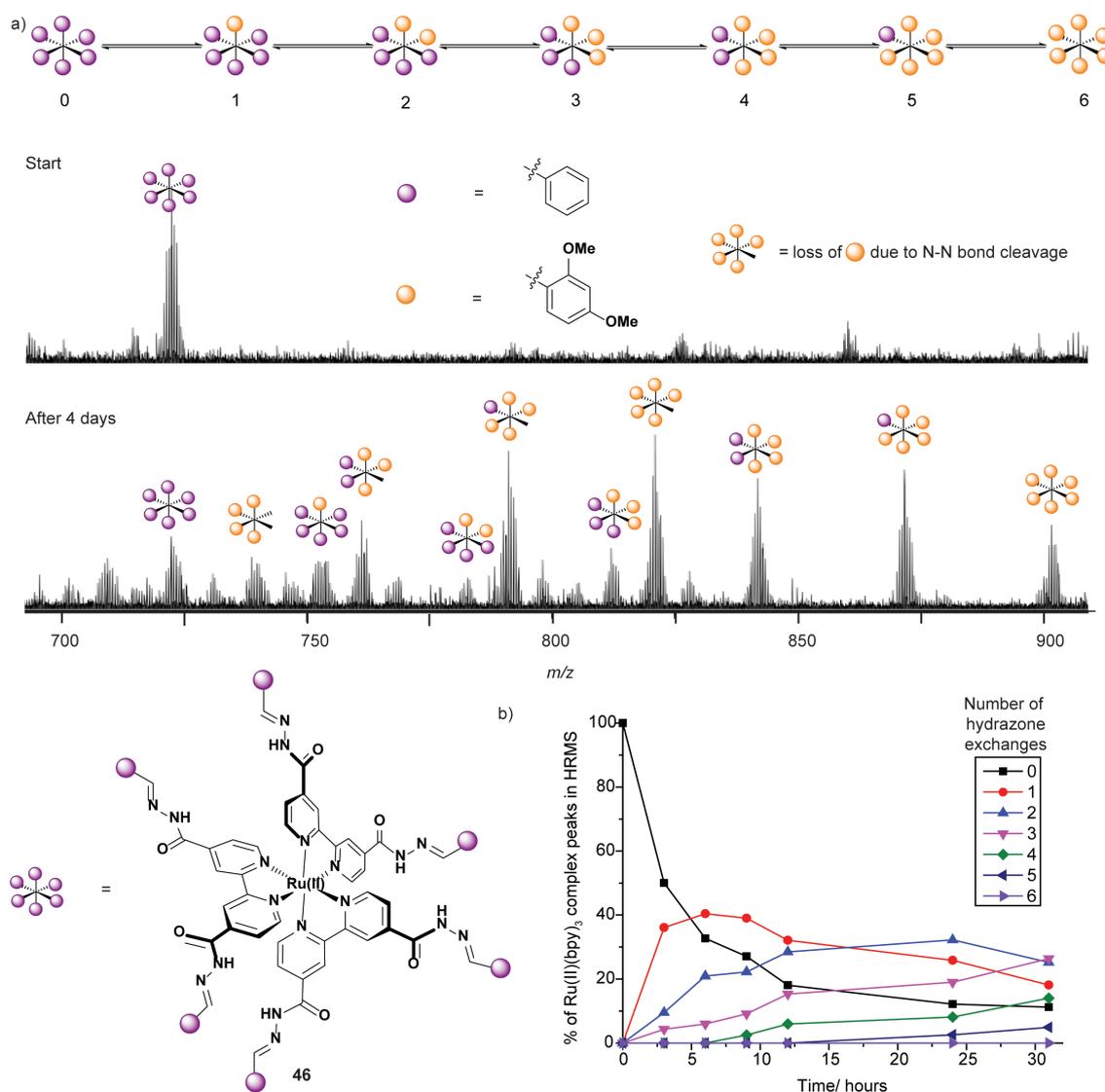
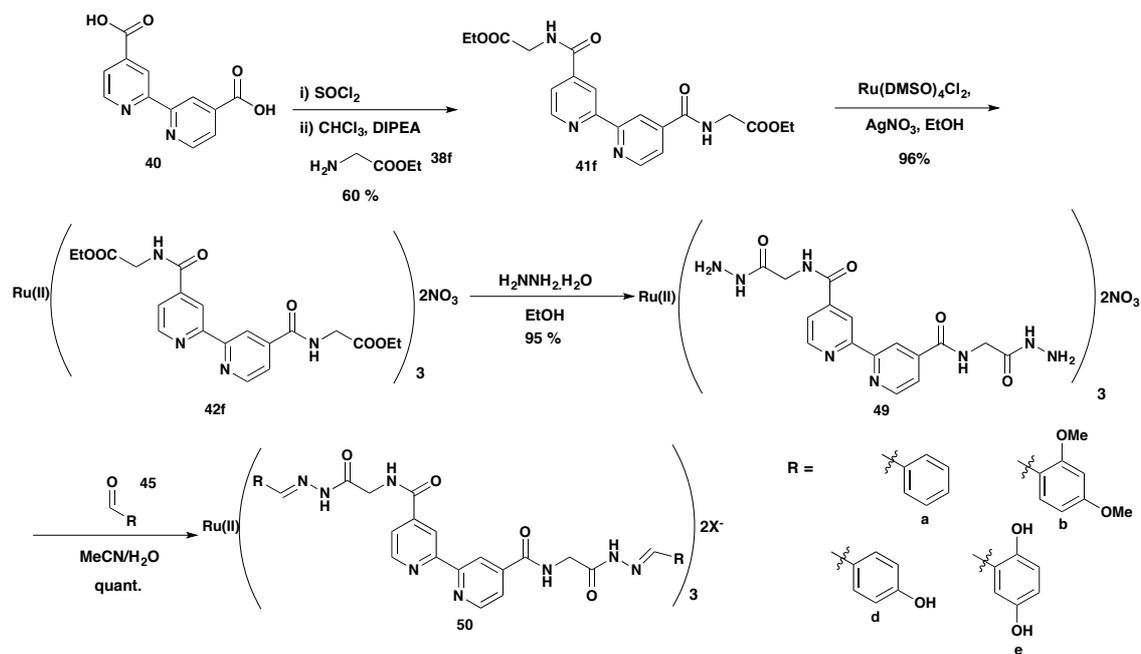


Figure 3.2 Hydrazone exchanges on the initial Ru(II)(bpy)₃ hydrazone scaffold **46**. a) Mass spectra depicting the starting benzaldehyde hydrazone Ru(II)(bpy)₃ species **46a** (100 μ M) (top) and the species present after 4 days (bottom) of incubation with 2,4-dimethoxy benzaldehyde **44b** (10 mM) with aniline **45** catalyst (10 mM) in 1:1 MeCN:H₂O. b) Time course following the species present at different time points

3.2.2 Ru(II)(bpy)₃ scaffold developments

The initial hydrazide Ru(II)(bpy)₃ complex scaffold **43** demonstrated that hydrazone exchange can be performed around a Ru(II)(bpy)₃ core. However, scaffold **43** was not well optimised, for two reasons: i) the hydrazone Ru(II)(bpy)₃ complexes **46** formed were not soluble in biologically relevant media (requiring high DMSO concentrations to solubilise in water), ii) the degradation by N-N bond cleavage in the mass spectrometer complicated any analysis of the kinetics of hydrazone exchange.

3.2.2.1 Glycine hydrazide Ru(II)(bpy)₃ complex



Scheme 3.6 Synthesis of glycine hydrazide and hydrazone Ru(II)(bpy)₃ complexes **49** and **50**

The N-N bond cleavage in the mass spectrometer was hypothesised to be due to stabilisation of the radical formed by the bipyridine π -system attached to the ruthenium(II) centre. It was thought that this interaction could be broken by introducing an sp^3 centre between the hydrazide and the Ru(II)(bpy)₃ core. In order to achieve this, an ethyl glycine complex **42f** was synthesised (**Scheme 3.6**), which was used to make a new hydrazide Ru(II)(bpy)₃ complex **49**. This glycine hydrazide Ru(II)(bpy)₃ complex **49** was seemingly more stable to oxygen than the initial Ru(II)(bpy)₃ hydrazide complex **43**, and could similarly be used to form hydrazone Ru(II)(bpy)₃ complexes **50**. In the mass spectrum less N-N bond cleavage was observed than for the initial Ru(II)(bpy)₃ hydrazone complexes **46**, although it was still observable.

Various glycine hydrazone Ru(II)(bpy)₃ complexes **50** could again be formed, however these were less soluble than the hydrazones of the initial hydrazide Ru(II)(bpy)₃ complex scaffold **43**. These hydrazone Ru(II)(bpy)₃ complexes **50** were only soluble in DMSO and DMF, while the initial Ru(II)(bpy)₃ complex scaffold hydrazones **46** were soluble in water/acetonitrile mixtures. Upon attempts to dilute DMSO stocks of the new hydrazone Ru(II)(bpy)₃ complexes **50** into water and various buffers in the concentrations required for HRMS analysis, precipitation was observed until 50 % DMSO was used, even when using more hydrophilic aldehydes, like 4-hydroxy benzaldehyde **45d** and 2,5-dihydroxy benzaldehyde **45e** in an attempt to aid this solubilisation. Therefore this system also could

not be used in a biologically relevant setting, so it was decided to look at new potential Ru(II)(bpy)₃ complex scaffolds with solubilising groups attached to the core scaffold.

3.2.3 New synthetic methodology

As new biologically compatible, stable hydrazone Ru(II)(bpy)₃ complex scaffolds were required for DCC, it was decided to see if a new, more divergent synthesis could be exploited.

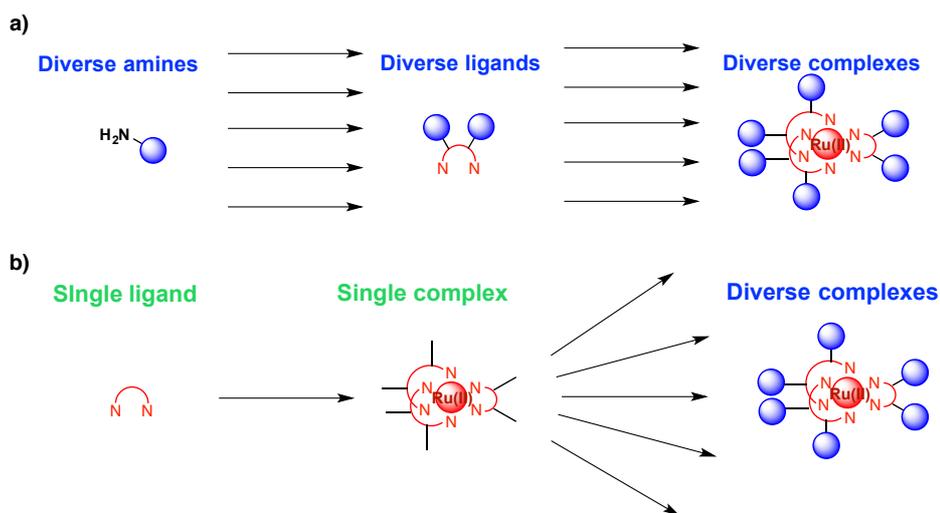
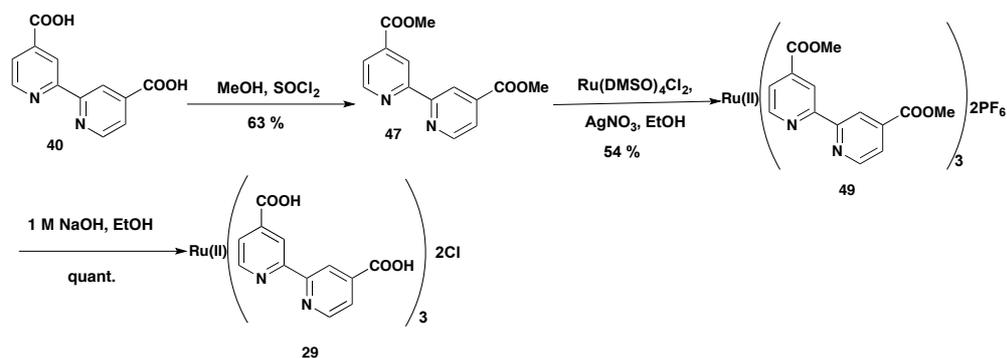


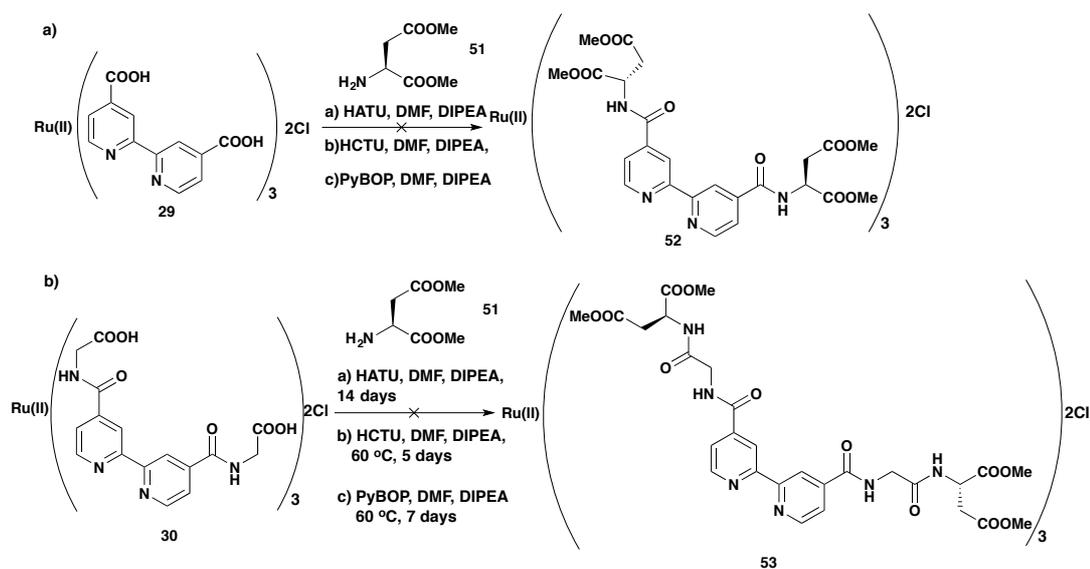
Figure 3.3 Difference between linear and divergent synthesis a) Linear synthesis, b) Divergent synthesis

The syntheses previously described, in Chapter 2, involve the synthesis of separate ligands and protected Ru(II)(bpy)₃ complexes for each of the functionalised Ru(II)(bpy)₃ complexes required (**Figure 3.3a**). However, especially for the synthesis of a wide range of Ru(II)(bpy)₃ complexes, it would be useful to use a more divergent route, whereby a single Ru(II)(bpy)₃ complex precursor could be synthesised and then functionalised (**Figure 3.3b**).

To this end, an acid functionalised Ru(II)(bpy)₃ complex **29** was synthesised (**Scheme 3.7**),¹⁴² with the objective of forming amide bonds on its periphery. This acid Ru(II)(bpy)₃ complex **29** could be readily synthesised in gram quantities. The carboxylic acid functionality on Ru(II)(bpy)₃ complex **29** was hypothesised to be more reactive than that of the diacid ligand **41**, due to both an increase in solubility and as the carboxylic acid is more susceptible to attack due to the Ru(II)(bpy)₃ core withdrawing electron density. The activation of a carbonyl in this position can be seen in the activation of amide bonds in this position to hydrolysis.

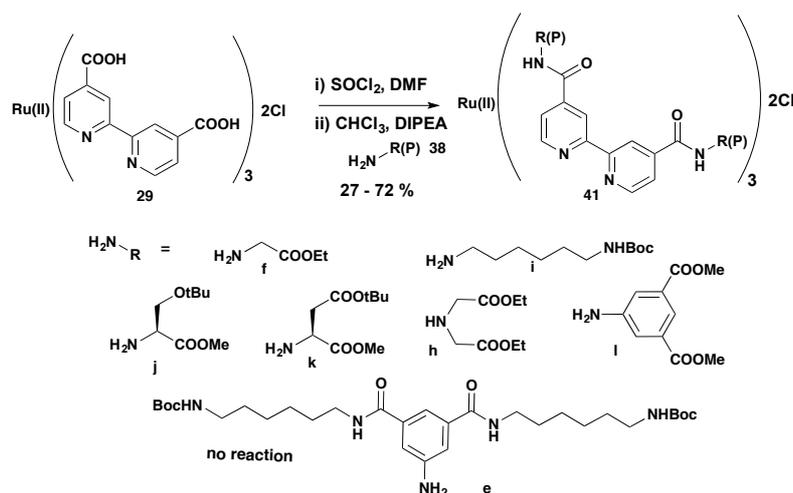


Scheme 3.7 Synthesis of acid functionalised Ru(II)(bpy)₃ complex **29**



Scheme 3.8 Attempts at amide bond formation on intact Ru(II)(bpy)₃ complexes **29** and **30** using peptide coupling agents a) **29** b) **30**

Attempts to form amide bonds on the acid Ru(II)(bpy)₃ complex **29** using standard peptide coupling agents (**Scheme 3.8a**), including HATU, HCTU and PyBOP did not prove fruitful, even with week-long reaction times and heating to 60 °C. Similar attempts were also made with a deprotected glycine substituted Ru(II)(bpy)₃ complex **30** (**Scheme 3.8b**), as it was thought that this would be more peptide-like, so that the acid functionality may react more readily with the peptide coupling agents. However, these also did not prove fruitful. These reactions were also difficult to follow: it was found that all the Ru(II)(bpy)₃ complexes, except the expected final Ru(II)(bpy)₃ complexes **52** and **53** stuck to the baseline of the TLC plate. Similarly LCMS did not prove helpful as the starting (and presumably intermediate) Ru(II)(bpy)₃ complexes **29** and **30** did not ionise well, and appeared on the solvent front of the HPLC trace.



Scheme 3.9 Amide bond formation on an intact Ru(II)(bpy)₃ complex

Due to the problems encountered using the peptide coupling agents it was decided to attempt to form an acid chloride Ru(II)(bpy)₃ complex from the acid Ru(II)(bpy)₃ complex **29** and use this to form amide bonds. This route (**Scheme 3.9**) worked well, giving the fully-functionalised Ru(II)(bpy)₃ complexes **41**. This reaction proceeded with various different primary amines **38f**, **i**, **j** and **k** and even with a secondary amine **38h**, and a small aniline **38l**, however, reaction with a highly functionalised aniline **38e** did not lead to the desired product, presumably due to sterics. Ru(II)(bpy)₃ complexes formed in this manner also proved to be easier to purify; so long as the amine **38** could be removed by acid washes, only the fully functionalised Ru(II)(bpy)₃ complex **41** moved off the baseline on the TLC plate in 10 % methanol in dichloromethane. This is compared to the products from ruthenium(II) complexation from the previous methodology where the Ru(II)(bpy)₃ complexes **41** needed to be separated from the ligands **42** which often had similar R_fs. This meant that the column purification was much quicker, so removal of acid-sensitive protecting groups as seen with some of the larger Ru(II)(bpy)₃ complexes **41c** and **e** on silica was not observed.

There were some disadvantages to this new synthetic route; the yields were much lower than that for ligand formation prior to ruthenium(II) complexation, presumably due to the requirement for 6 amide bond formations on the Ru(II)(bpy)₃ complex **29** to form the desired product as opposed to 2 amide bond formations on the bpy ligand **40**. However, the advantage of being able to form small quantities of many Ru(II)(bpy)₃ complexes much more quickly could allow the acceleration of testing Ru(II)(bpy)₃ complexes for a variety of different applications.

Synthetic methods for the synthesis of both 4' and 5' monosubstituted Ru(II)(bpy)₃ complexes have also been studied, these are presented in Appendix II. The synthesis of these lower functionality Ru(II)(bpy)₃ complexes offer the potential advantage of less complicated analytical chemistry in the application of these Ru(II)(bpy)₃ complexes for DCC.

3.2.4 New Ru(II)(bpy)₃ hydrazone scaffolds

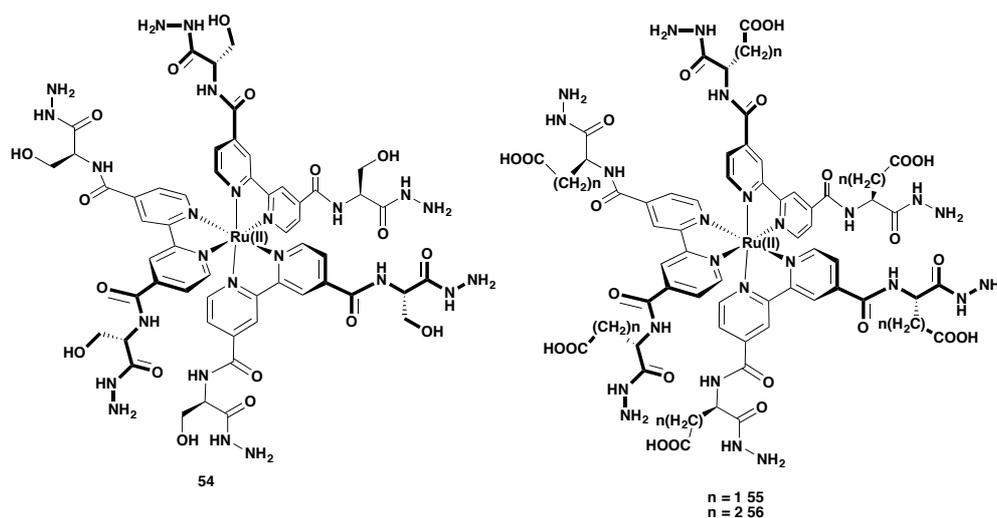
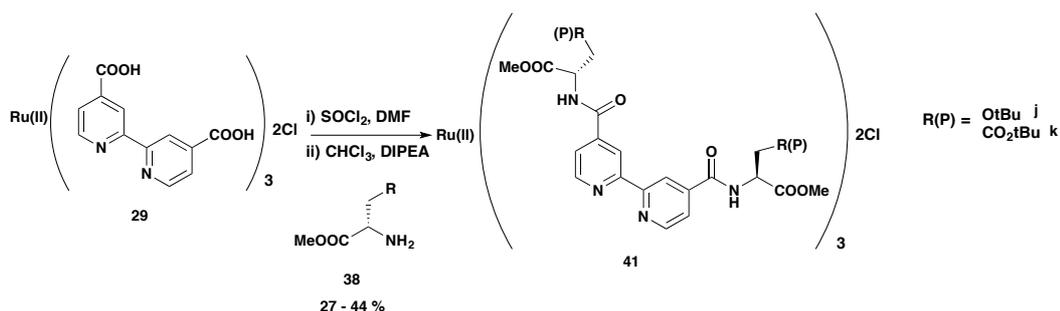


Figure 3.4 Serine (**54**) and Aspartic/Glutamic acid (**55** and **56**) hydrazide Ru(II)(bpy)₃ complexes, with the hydroxyl/carboxylate groups potentially solubilising the core DCC scaffold

With a new synthetic methodology in hand, it was decided to attempt to improve the solubility of the previously synthesised glycine hydrazone Ru(II)(bpy)₃ complexes **50**, by increasing the solubility of the core Ru(II)(bpy)₃ complex scaffold. It was decided to change the glycine linker in **50** to serine (**54**), aspartic acid (**55**) and glutamic acid (**56**) (**Figure 3.4**), introducing hydroxyl or carboxylic acid groups to potentially solubilise the whole Ru(II)(bpy)₃ complex scaffold.

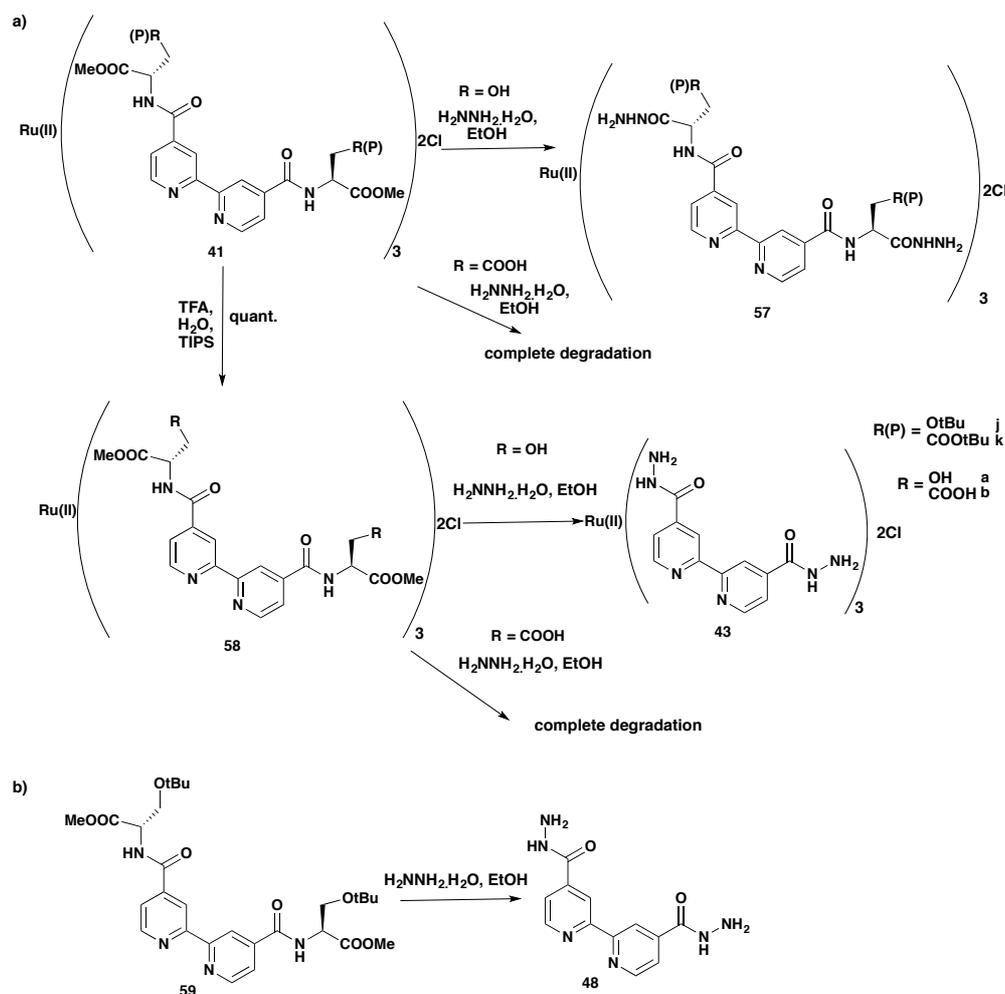
The addition of the acid groups also had the potential advantage of facilitating the hydrazone exchange by neighbouring group participation. So far nucleophilic catalysts (aniline) for catalysing the hydrazone exchange at neutral pH, have been discussed, however hydrazone exchange can also be catalysed by acid, and indeed *ortho*-carboxy phenylhydrazine has been shown to increase the rate of hydrazone formation, compared to phenylhydrazine.¹⁷⁵ In the aspartic acid and glutamic acid hydrazide Ru(II)(bpy)₃ complexes **55** and **56** an acid functionality is in close proximity to the hydrazone so this could facilitate the hydrazone exchange, this could be envisaged to work in a manner akin to general acid catalysis in an enzyme active site.

3.2.4.1 Serine- and aspartic acid- hydrazide Ru(II)(bpy)₃ complexes **54** and **55** synthetic attempts



Scheme 3.10 Synthesis of orthogonally protected serine and aspartic acid hydrazide Ru(II)(bpy)₃ complexes **41j** and **k**

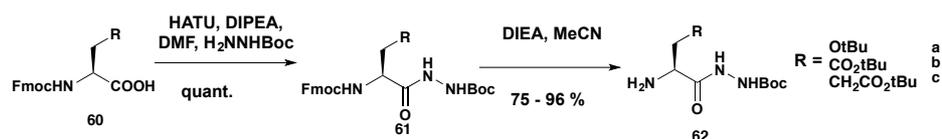
Initially orthogonally protected serine and aspartic acid functionalised Ru(II)(bpy)₃ complexes **41j** and **k** were synthesised (**Scheme 3.10**), using the new synthetic methodology. From these Ru(II)(bpy)₃ complexes there are two further reactions required to synthesise the desired hydrazide Ru(II)(bpy)₃ complexes **54** and **55**, these are: i) *tert*-butyl deprotection of the hydroxyl/carboxylic acid side chains and ii) hydrazide formation on the methyl esters. These were attempted in both orders (**Scheme 3.11a**). The *tert*-butyl deprotection proceeded without fault for both the Ru(II)(bpy)₃ complexes **41j** and **k**, however on reaction of these species with hydrazine monohydrate, rather than obtaining the desired hydrazide Ru(II)(bpy)₃ complexes **54** and **55**, amide bond cleavage (with **58a**) to give the initial hydrazide Ru(II)(bpy)₃ complex **43** or complete degradation (with **58b**) were observed. Attempts at first making the hydrazide functionality, yielded less amide bond cleavage on the serine Ru(II)(bpy)₃ complex **41j**, however, as there was still observable amide bond cleavage a clean sample of the Ru(II)(bpy)₃ complex **57j** could not be obtained. With the aspartic acid Ru(II)(bpy)₃ complex **41k** the reaction with hydrazine monohydrate, again yielded degradation. Similar attempts at the hydrazide formation reaction (**Scheme 3.11b**) on the orthogonally protected serine ligand **59** yielded the highly insoluble hydrazide ligand **48**, by amide bond cleavage.



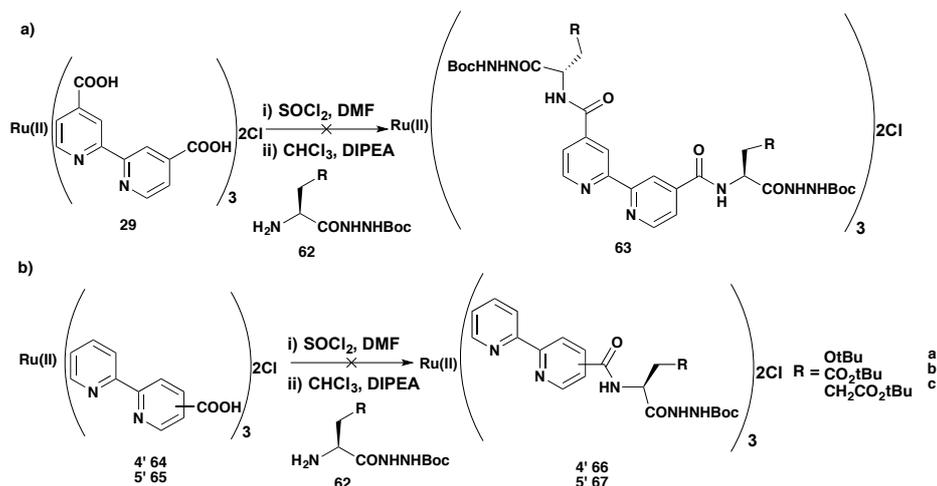
Scheme 3.11 Attempts at hydrazide Ru(II)(bpy)_3 complex formation on serine and aspartic acid Ru(II)(bpy)_3 complexes **41j** and **k**

3.2.4.2 Using hydrazido-amino acids to functionalise the Ru(II)(bpy)_3 complexes

Seemingly the problem with the first attempts at synthesising the serine and aspartic acid hydrazide complexes **54** and **55** was the hydrazide formation reaction. Therefore it was decided to add an amino acid already possessing the hydrazide functionality onto the Ru(II)(bpy)_3 complex **29**. These amino acids **62** could readily be synthesised from Fmoc-amino acids (**Scheme 3.12**).



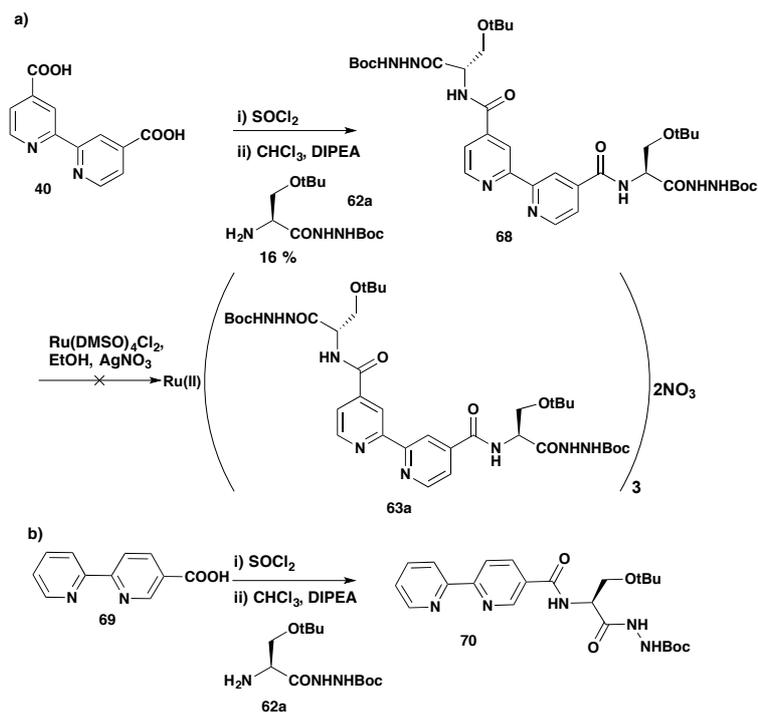
Scheme 3.12 Synthesis of serine, aspartic acid and glutamic acid hydrazido-amino acids **62**



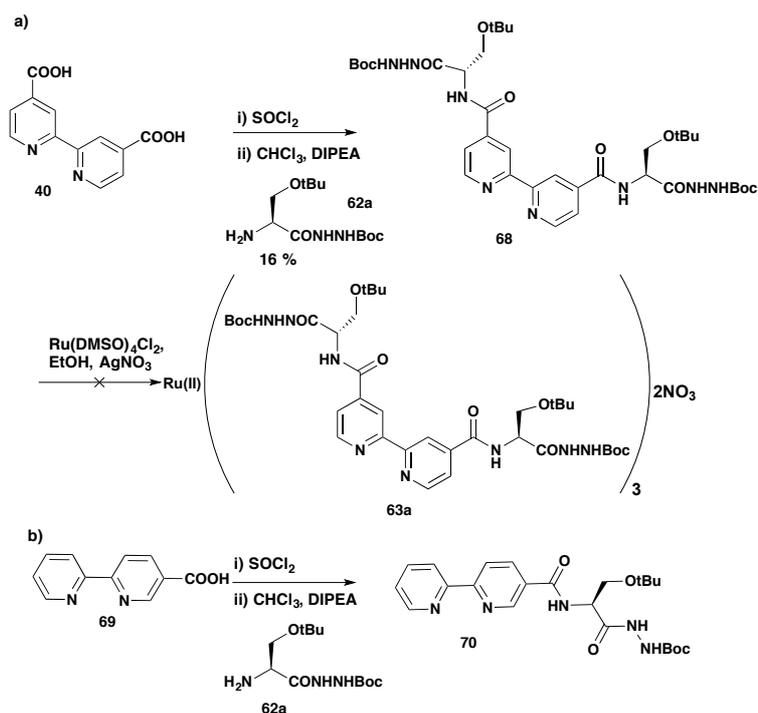
Scheme 3.13 Attempts at amide bond formation between intact acid functionalised Ru(II)(bpy)₃ complexes **29**, **64** and **65** and the hydrazido amino acids **62** a) disubstituted Ru(II)(bpy)₃ complex **63**, b) monosubstituted Ru(II)(bpy)₃ complexes **66** and **67**

After synthesising these hydrazido-amino acids **62** it was attempted to add them onto the intact acid functionalised Ru(II)(bpy)₃ complex core **29** (**Scheme 3.13a**). Attempts were made with all three hydrazido-amino acids **62**, but the reactions did not prove fruitful, with the aqueous layers staying red, and brown organic phases on aqueous work up. Further analysis of the organic phase revealed no discernable mass peaks in the mass spectrum, and a brown smudge by TLC, where a single red spot would be expected. Similar attempts were made with the 4' and 5' monosubstituted Ru(II)(bpy)₃ complexes **64** and **65** (synthesis of these acid functionalised monosubstituted Ru(II)(bpy)₃ complexes **64** and **65** is described in Appendix II) but again this did not yield the desired products **66** and **67** (**Scheme 3.13**).

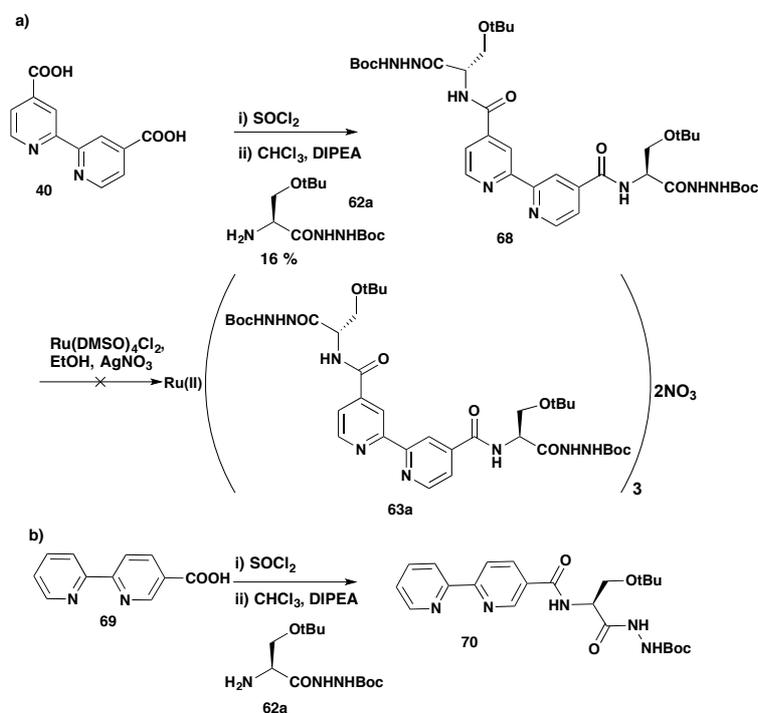
An attempt to add the serine hydrazido-amino acid **62a** to the diacid ligand **40** (



Scheme 3.14a), did give the desired ligand **68**, but with much lower yield than usually observed for similar reactions on the ligand **40**. A ruthenium(II) complexation reaction of the ligand **68** was attempted, however the ligand **68** seemingly degraded in the ruthenium(II) complexation reaction. From this it was thought that perhaps higher temperatures were causing the degradation so the amide bond formation on the intact $\text{Ru}(\text{II})(\text{bpy})_3$ complex **29** was attempted on ice, however this gave similar results to those obtained at room temperature, followed by reflux. A low yield was also observed in attempts to synthesise the 5' serine hydrazide ligand **70** (



Scheme 3.14b), with not enough ligand **70** being formed for attempts at ruthenium(II) complexation.



Scheme 3.14 Attempts at synthesising serine hydrazone $\text{Ru}(\text{II})(\text{bpy})_3$ scaffolds **63a** and **67a** by first synthesising ligands **68** and **70**, a) disubstituted ligand **68**, b) 5' monosubstituted ligand **70**

3.3 Porphyrin scaffolds

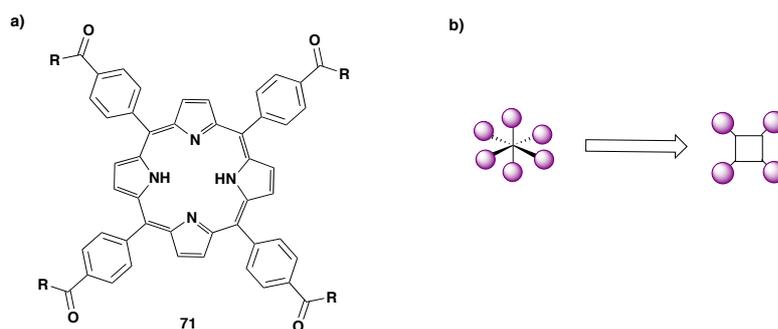


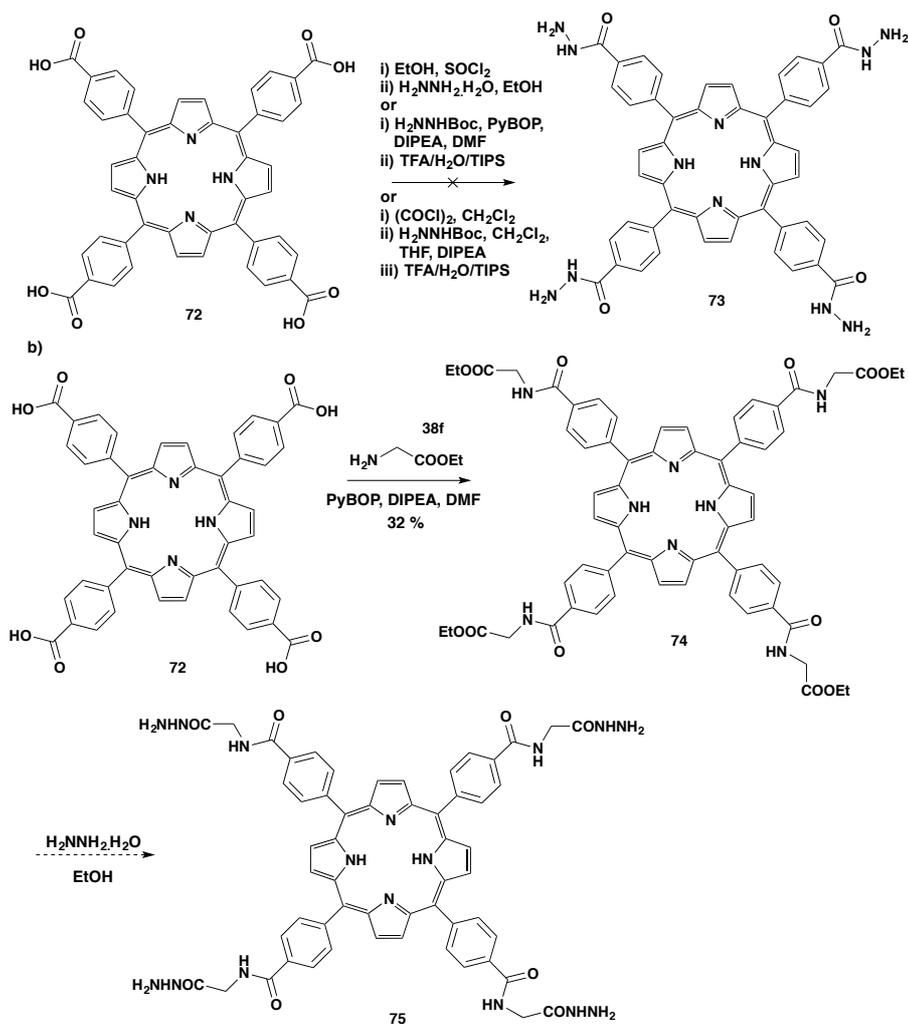
Figure 3.5 Porphyrin scaffold a) Tetra-phenyl porphyrin **71** b) Changing from 6 exchangeable groups on the $\text{Ru}(\text{II})(\text{bpy})_3$ complexes to 4 on the porphyrin scaffold

Due to the degradation problems experienced in the synthesis of the hydrazone $\text{Ru}(\text{II})(\text{bpy})_3$ complexes **54**, **55** and **56**, a different multivalent hydrazone scaffold was explored. To this end a tetra-phenyl porphyrin scaffold **71** (**Figure 3.5a**) was chosen. Using this scaffold several advantages can be envisaged: i) there are only 4 exchangeable

groups as opposed to the 6 with the disubstituted Ru(II)(bpy)₃ complexes (**Figure 3.5b**), potentially making the analysis less challenging, ii) the porphyrins primarily ionise as +1 rather than +2 species separating the peaks in the mass spectrum, and iii) there is no multi-isotopic metal ion, decreasing the isotopic pattern window in the mass spectrum, potentially increasing signal to noise.

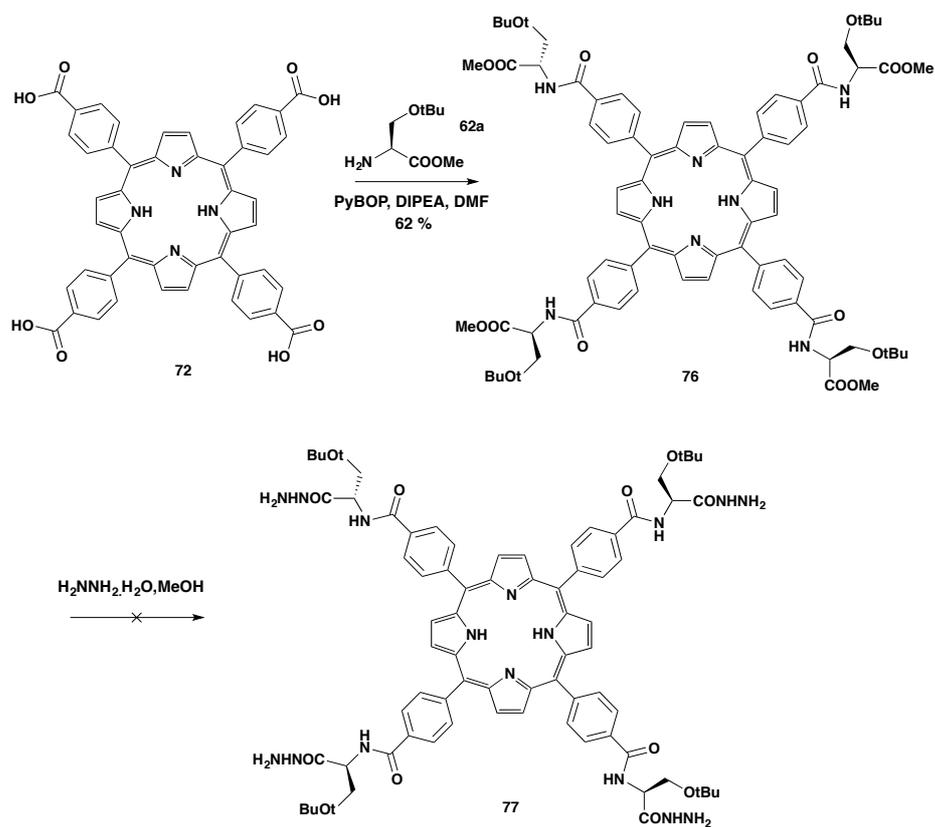
3.3.1 Porphyrin scaffold design and synthesis

Attempts were first made to synthesise a scaffold analogous to the initial hydrazide Ru(II)(bpy)₃ complex scaffold **43** with the hydrazide directly attached to the tetra-phenyl porphyrin **73** (**Scheme 3.15a**). This synthesis was attempted *via* several routes, however determining what happened in the reactions was difficult due to insolubility. Due to these solubility problems and a hypothesised potential for radical degradation of the N-N bond in the mass spectrometer it was decided not to pursue this scaffold any further, and to again incorporate an sp³ centre between the central tetraphenyl porphyrin core and the hydrazide moiety.



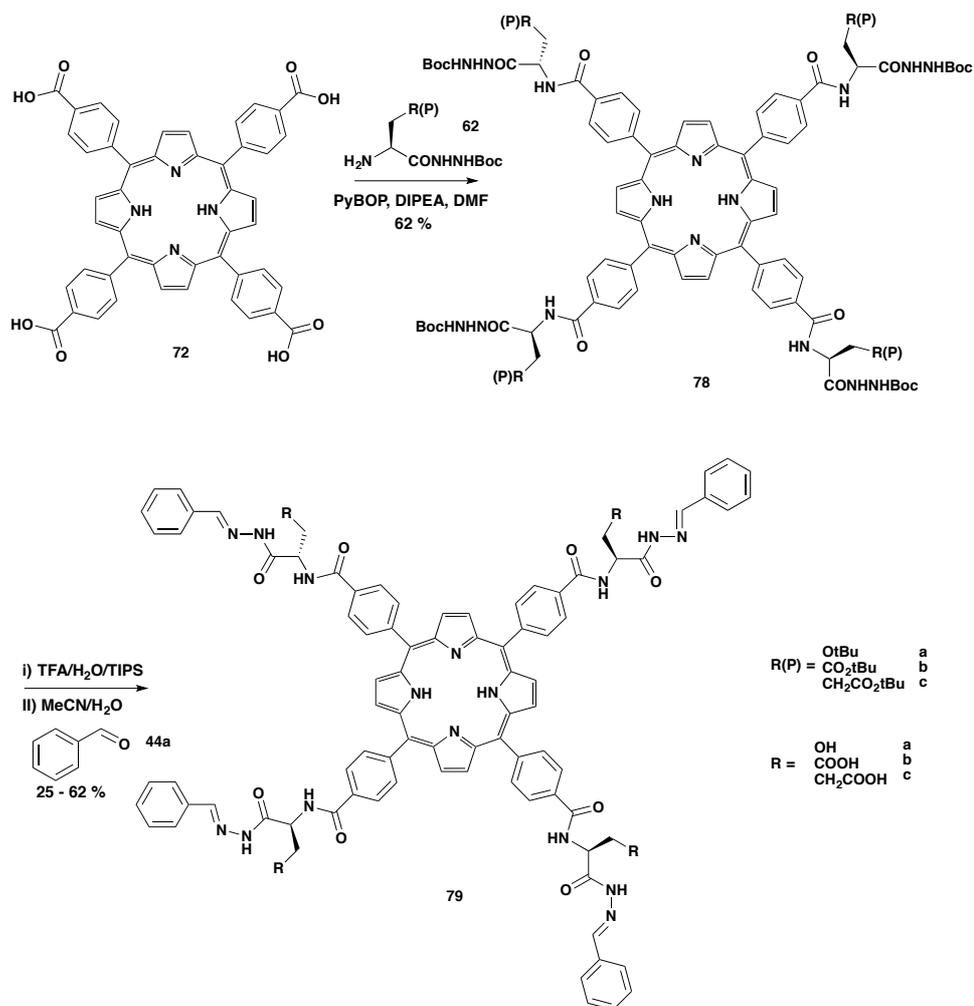
Scheme 3.15 Initial attempts at making porphyrin hydrazide scaffolds **73** and **75** a) directly attached to porphyrin **73**, b) glycine hydrazide porphyrin **75**

At first, a glycine hydrazide porphyrin scaffold **73**, analogous to the glycine hydrazide Ru(II)(bpy)₃ complex **49**, was chosen. To this end, an ethyl glycine porphyrin **74** was synthesised (**Scheme 3.15b**), and subjected to hydrazide formation conditions. This gave a very insoluble purple solid, which was hypothesised to be the glycine hydrazide porphyrin **75** but could not be characterised due to its insolubility. The insolubility of the glycine hydrazide porphyrin **75**, is perhaps unsurprising given that the deprotected glycine porphyrin is also very insoluble, and hydrazides seem to mirror the solubility of their corresponding acid.



Scheme 3.16 First attempted synthesis of serine hydrazide porphyrin **77**

With the solubility issues with the glycine hydrazide porphyrin **75**, it was decided to attempt to use a serine in place of the glycine to, as with the Ru(II)(bpy)₃ complexes, increase the scaffold solubility. A serine methyl ester amino acid **62a** was first attached to the porphyrin core (**Scheme 3.16**), however this, as with the Ru(II)(bpy)₃ complexes, did not fair well in the hydrazide formation reaction, therefore it was decided to use the hydrazido-amino acids **62** previously synthesised.

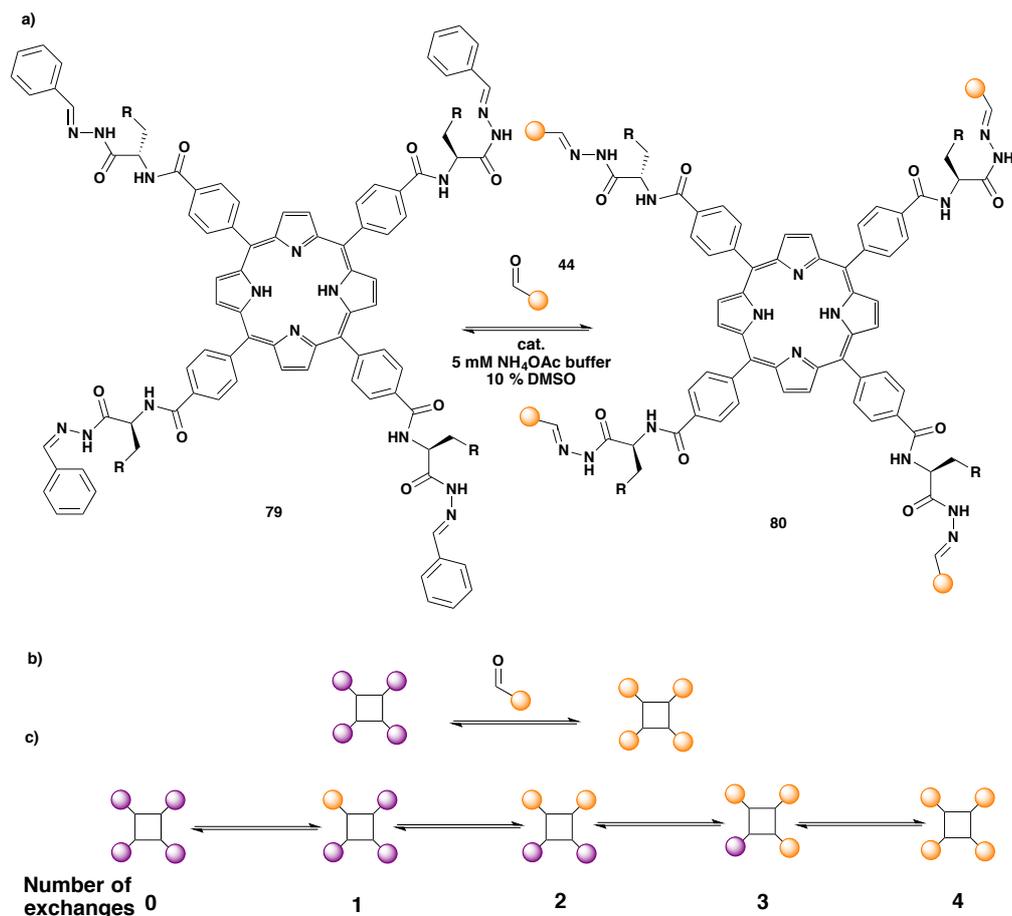


Scheme 3.17 Synthesis of serine, aspartic acid and glutamic acid hydrazide and benzaldehyde hydrazone porphyrins **78** and **79**

Amide bonds were formed between the serine, aspartic acid and glutamic acid hydrazido amino acids **62**, and tetra-carboxy phenyl porphyrin **72** (**Scheme 3.17**), using PyBOP as a coupling agent. The hydrazide Boc and side chain *tert*-butyl protecting groups could readily be removed with TFA to yield the desired hydrazide porphyrins **78**. These hydrazide porphyrins were used to form the benzaldehyde hydrazone porphyrins **79**. These benzaldehyde hydrazone porphyrins **79** were then used in all further DCC analyses.

3.3.2 Hydrazone exchanges

With the three benzaldehyde hydrazone porphyrins **79** in hand it was first decided to see if it was possible to exchange the benzaldehyde moiety with other aldehydes **44** (**Scheme 3.18**). The benzaldehyde hydrazone porphyrins **79** were soluble in 10 % DMSO in aqueous solutions, thus allowing attempts at hydrazone exchange to occur in biologically relevant media, in comparison to the attempts with the initial hydrazone Ru(II)(bpy)₃ complexes **46** which required acetonitrile/water mixtures.



Scheme 3.18 Porphyrin hydrazone exchanges a) General reaction scheme, b) Cartoon representation of reaction, c) Cartoon representation of individual exchanges

As with the Ru(II)(bpy)_3 complexes, HRMS was used to follow the hydrazone exchange reactions. The argument for the validity of this method for the Ru(II)(bpy)_3 complexes was that as the Ru(II)(bpy)_3 complexes are already ionised as a 2+ species, there will be little difference in the ionisation of the different hydrazone Ru(II)(bpy)_3 complexes. This argument is not valid for the neutral porphyrins. However, much of the porphyrin ionisation is due to protonation of the central porphyrin core, but protonation of the peripheral groups will now also be more important. This means that the mass spectrum may not give the actual proportions of species in solution, however they will be related. Therefore, if the relative proportions of the peaks in the mass spectrum remain constant, the DCC system has reached a point where the proportions of each species is constant, a static point or a dynamic equilibrium.

3.3.2.1 Hydrazone exchange to 1 aldehyde

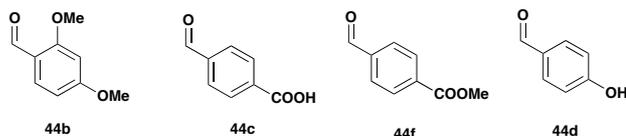


Figure 3.6 Aldehydes used in hydrazone exchange reactions

Hydrazone exchanges were performed with the three different porphyrin scaffolds **79**, with three different aldehydes (**Figure 3.6**): 2,4-dimethoxy benzaldehyde **44b**, 4-carboxy benzaldehyde **44c** and 4-methyl ester benzaldehyde **44f**. It should be noted that attempts were also made with 4-hydroxy benzaldehyde **44d**, however the mass peaks were harder to discern in this case. Hydrazone exchanges were initially performed without a catalyst with the aldehyde added in 50 equivalents compared to the hydrazone porphyrin **79** (12.5 equivalents per hydrazone). These hydrazone exchanges were followed over time (**Figure 3.7b**), in order to establish i) if hydrazone exchange occurred, ii) if equilibrium was reached and iii) the timescale for the establishment of equilibrium.

Considering one of these systems, (aspartic acid hydrazone porphyrin **79b** with 4-carboxy benzaldehyde **44c**), hydrazone exchange was observed within 24 hours, and the hydrazone exchanges could be visualised by HRMS (**Figure 3.7a**), with clearly discerned mass peaks for each of the subsequent hydrazone exchanges being observed. There were some extra peaks in the mass spectrum; these corresponded to sodiated and ammoniated species for each of the subsequent hydrazone exchanges, as well as peaks for N-N bond cleavage. Looking at just the protonated peaks, the progression of these hydrazone exchanges could be followed over time (**Figure 3.7b**) with successive hydrazone exchanges being observed. The initial species **79b** (0) decreased over time, with the species with one hydrazone exchange (1) to 4-carboxy benzaldehyde **44c** at first increasing then decreasing. The products of successive hydrazone exchanges (2, 3 and 4) then followed in succession.

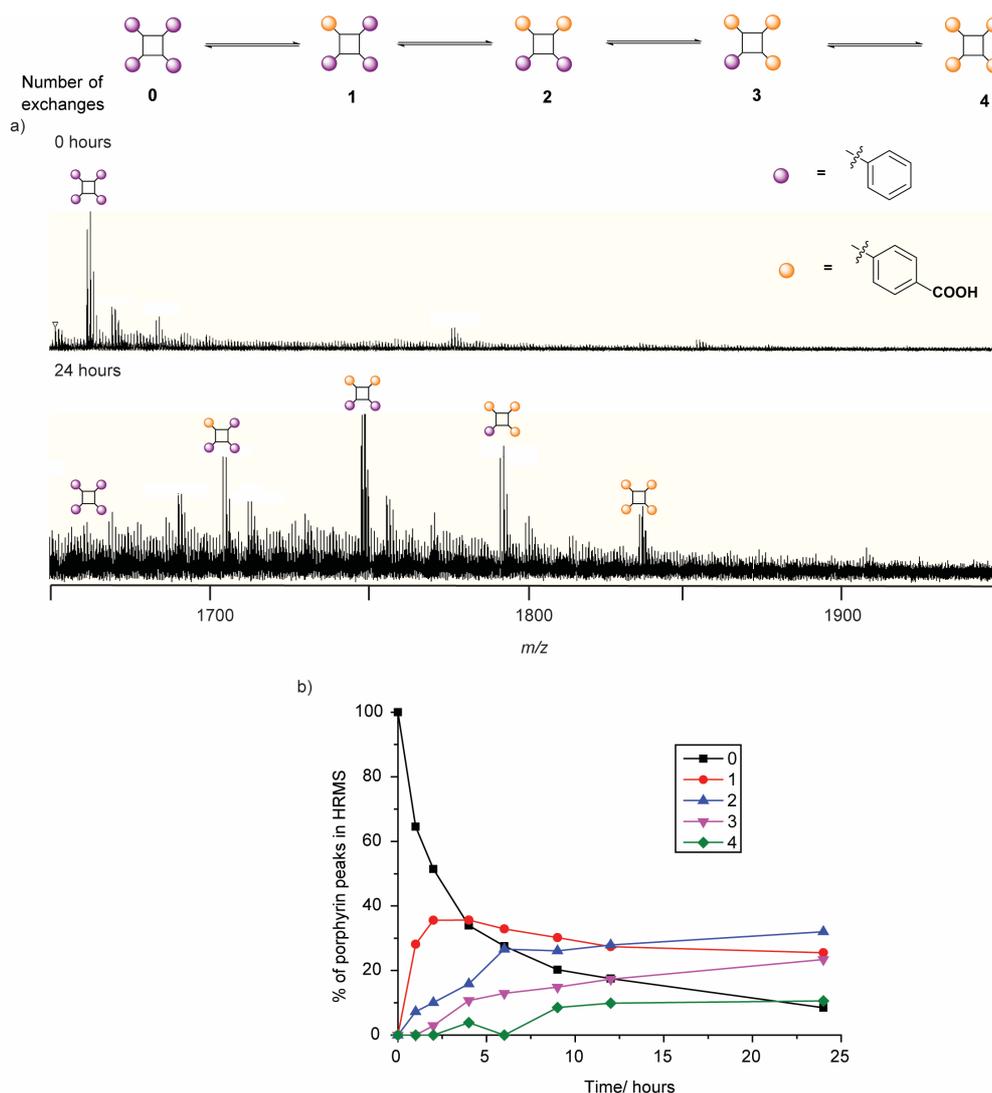


Figure 3.7 Exemplary HRMS showing hydrazone exchange on a porphyrin scaffold (**79b**). Conditions : 100 μ M **79b**, 5 mM **44c**, 10 % DMSO, 5 mM NH_4OAc , pH 6.75. a) HRMS trace in porphyrin +1 region at start of incubation (top) and after 24 hour incubation (bottom). b) Following the species present over a 24 hour time course.

These hydrazone exchanges were repeated for all three porphyrin hydrazone scaffolds **79** with the three different aldehydes (**44b**, **c** and **f**) (**Figure 3.8**) with hydrazone exchanges occurring over a 24 hour period. This shows that all 3 porphyrin scaffolds **79** could be compatible for biological DCC, with hydrazone exchange occurring in biologically relevant buffer, at ambient temperature.

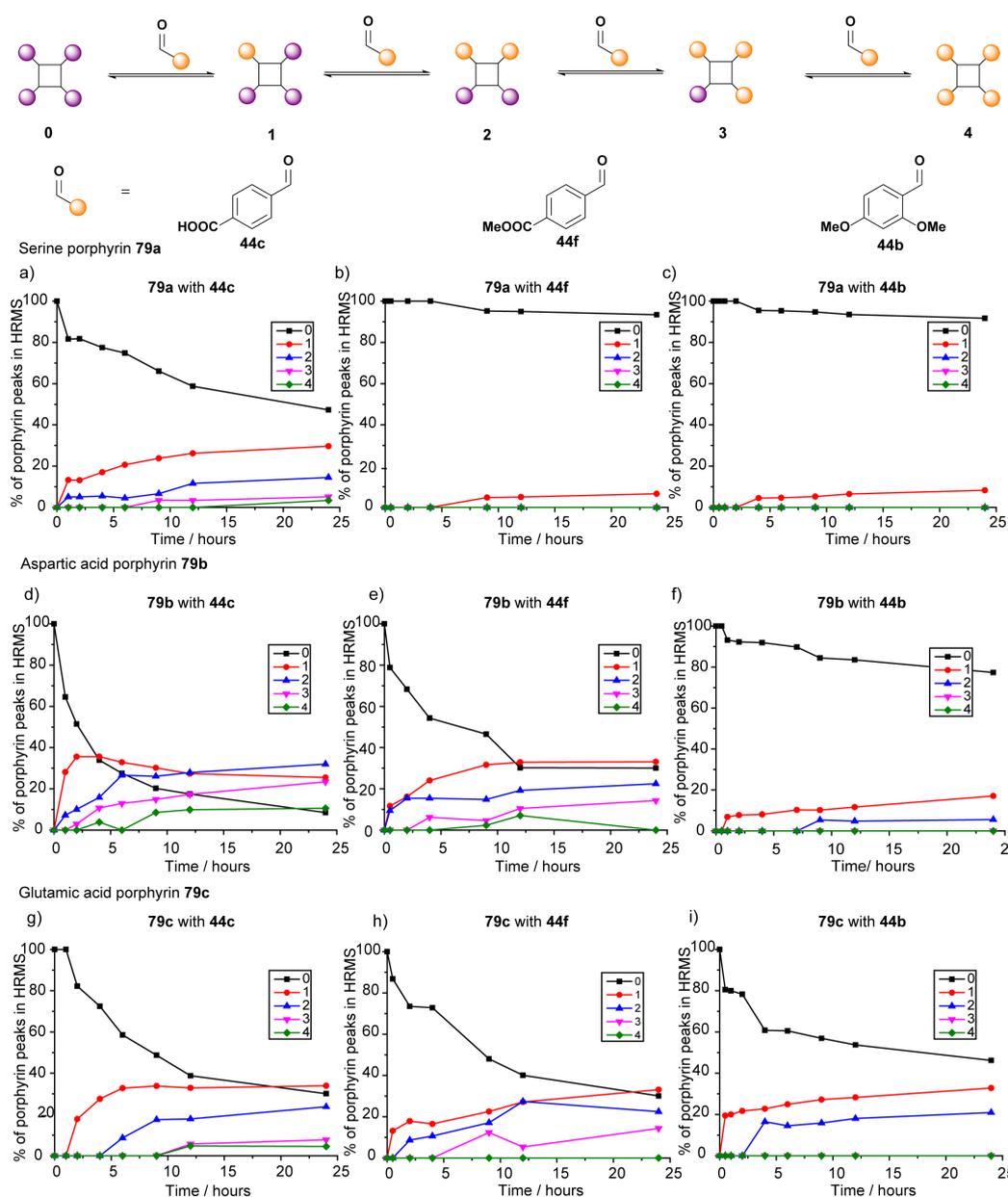


Figure 3.8 Initial hydrazone exchange time courses on porphyrin hydrazone scaffolds **79**. The separate lines on the graphs correspond to different numbers of benzaldehyde moieties having exchanged for a new aldehyde. 100 μM **79**, 5 mM **44**, 10 % DMSO in 5 mM NH_4OAc , pH 6.75. a), b) and c) Serine hydrazone porphyrin **79a**, d), e) and f) Aspartic acid hydrazone porphyrin **79b**, g), h) and i) Glutamic acid hydrazone porphyrin **79c** incubated with 4-carboxy benzaldehyde **44c** (a), d) and g)), 4-methyl ester benzaldehyde **44f** (b), e) and h)) and 2,4-dimethoxy benzaldehyde **44b** ((c), f) and i)) for 24 hours

Differing rates of hydrazone exchange were observed between the three porphyrin scaffolds and with the different aldehydes. Generally the glutamic acid and aspartic acid hydrazone porphyrins **79b** and **c** (**Figure 3.8d, e, f, g, h** and **i**) showed faster exchange rates than the serine hydrazone porphyrin **79a** (**Figure 3.8a, b, and c**). This supports the hypothesis that the acid functionality may catalyse the hydrazone exchanges. 4-Methyl

ester benzaldehyde **44f** and 4-carboxy benzaldehyde **44c** also seem to exchange at a faster rate than 2,4-dimethoxy benzaldehyde **44b**, this is likely due to the electron-withdrawing nature of the *para* carbonyl activating the aldehyde/imine to nucleophilic attack, while the electron-donating *ortho* and *para* methoxy groups deactivate the aldehyde/imine to attack. The 4-carboxy benzaldehyde **44c** also showed faster rates of hydrazone exchange than the 4-methyl ester benzaldehyde **44f**.

However in almost all cases equilibration was not being reached within the 24 hour period. Therefore, the addition of a catalyst to the DCC systems to increase the rate of reaching equilibration, was investigated, in the hope to allow equilibration at a timescale more compatible with biological applications.

3.3.2.1.1 Catalysts

As equilibration was not reached within 24 hours, the DCC systems were not completely suitable for protein templation. Different groups have reported the use of nucleophilic catalysts for use in hydrazone exchange reactions in biologically relevant media.^{168,176,177} Aniline **45** was the initial catalyst reported,¹⁶⁸ and was the catalyst investigated with the Ru(II)(bpy)₃ complexes **46** however, subsequently, different catalysts have been shown to catalyse hydrazone exchanges with faster rates of hydrazone formation and exchange in two different studies.^{176,177}

To look at the effects of catalysis on the hydrazone exchanges, and the rate of reaching equilibrium, aniline **45** and anthranilic acid **80** were chosen, as they are cheap, relatively soluble and are less likely to degrade in DMSO compared to the phenylene diamines reported by Distefano *et al.*¹⁷⁷ Hydrazone exchange reactions to exchange the benzaldehyde moiety for 4-carboxy benzaldehyde **44c** with the three porphyrin scaffolds **79** (Figure 3.9) were carried with both these catalysts.

As can be seen the two catalysts did increase the rate of equilibration in all cases. Comparing the aniline **45** and anthranilic acid **80** catalysts it seems the incubation with aniline **45** may have increased the rate slightly more than that of anthranilic acid **80**, therefore aniline **45** was chosen for use in all further studies.

Similar attempts to establish the effects of catalysts on the hydrazone exchanges with 2,4-dimethoxy benzaldehyde **44b** and 4-methyl ester benzaldehyde **44f** again showed the catalysts increasing the rate of reaching equilibrium. Both catalysts, however, could not be explored with these aldehydes, due to the formation of insoluble imines between the catalysts and the benzaldehyde derivatives. 2,4-Dimethoxy benzaldehyde **44b**

precipitated with anthranillic acid **80** and 4-methyl ester benzaldehyde **44f** precipitated with aniline **45**.

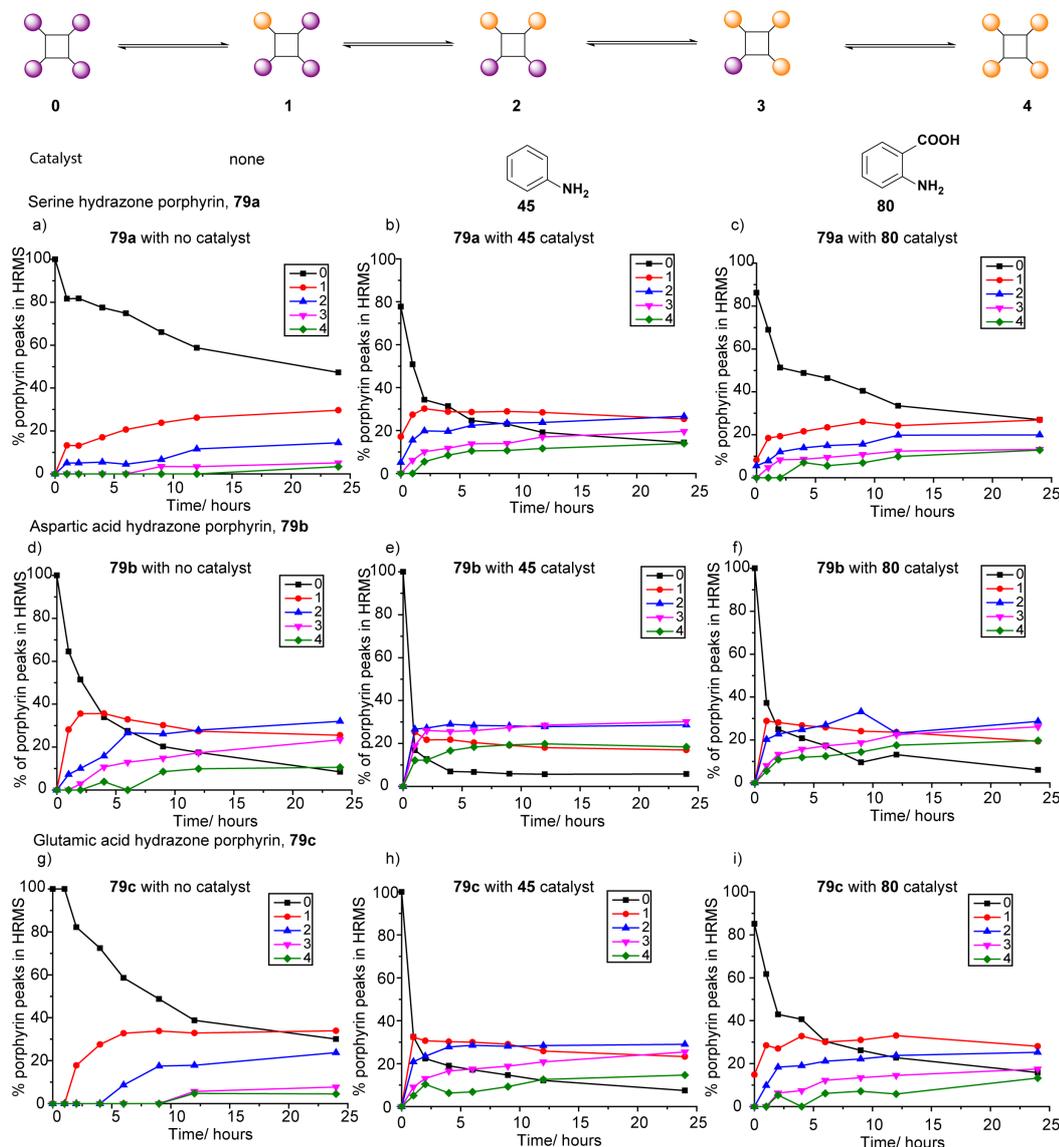


Figure 3.9 Effect of catalysis on porphyrin hydrazone exchange, using exemplary reaction with 4-carboxy benzaldehyde **44c** on all 3 porphyrin hydrazone scaffolds **79** and addition of no catalyst, aniline **45** and anthranillic acid **80**. 100 μ M hydrazone porphyrin **79**, 5 mM 4-carboxy benzaldehyde **44c**, 10 mM catalyst, 10 % DMSO in 5 mM NH_4OAc , pH 6.75 a), b) and c) Serine hydrazone porphyrin **79a**, d), e) and f) Aspartic acid hydrazone porphyrin **79b** and g), h) and i) Glutamic acid porphyrin **79c**, a), d) and g), no catalyst, b), e) and h) aniline **45** catalyst, c), f) and i) anthranillic acid **80** catalyst

Another important factor when considering catalysis is the catalyst loading. Greaney *et al.* used very high aniline **45** loadings compared to both the aldehydes (2000 eq.) and hydrazides (500 eq. per hydrazone). These catalyst loadings are much higher than can be used in this set up, due to solubility, however different concentrations of catalyst to

aldehyde should have an effect on the rate, especially when the catalyst concentrations are similar to that of the aldehyde. Attempts were made at different aniline **45** loadings, on exchange reactions with 2,4-dimethoxy benzaldehyde **44b**, for all three hydrazone porphyrin scaffolds **79**, with exemplary results for the aspartic acid hydrazone porphyrin scaffold **79b** shown in **Figure 3.10**. 2,4-dimethoxy benzaldehyde **44b** was chosen for this study as it displayed the slowest rates of reaching equilibrium in the initial, no catalyst studies. The rate of reaching equilibrium was faster with increasing catalyst concentrations, as would be expected, however so long as there was catalyst present, equilibrium was reached within 12 hours, a suitable time scale for performing DCC in the presence of a protein. Similar results were obtained for both the serine and glutamic acid hydrazone porphyrins **79a** and **c**, with increasing initial rates being particularly apparent for the serine hydrazone porphyrin **79a**. From this serine hydrazone porphyrin **79a** data (**Figure 3.11**) it was seen that under 10 mM aniline **45** concentration hydrazone exchanges occur at a slow rate, with a large rate increase observed at 10 mM aniline, therefore it was decided to use this catalyst concentration for further studies.

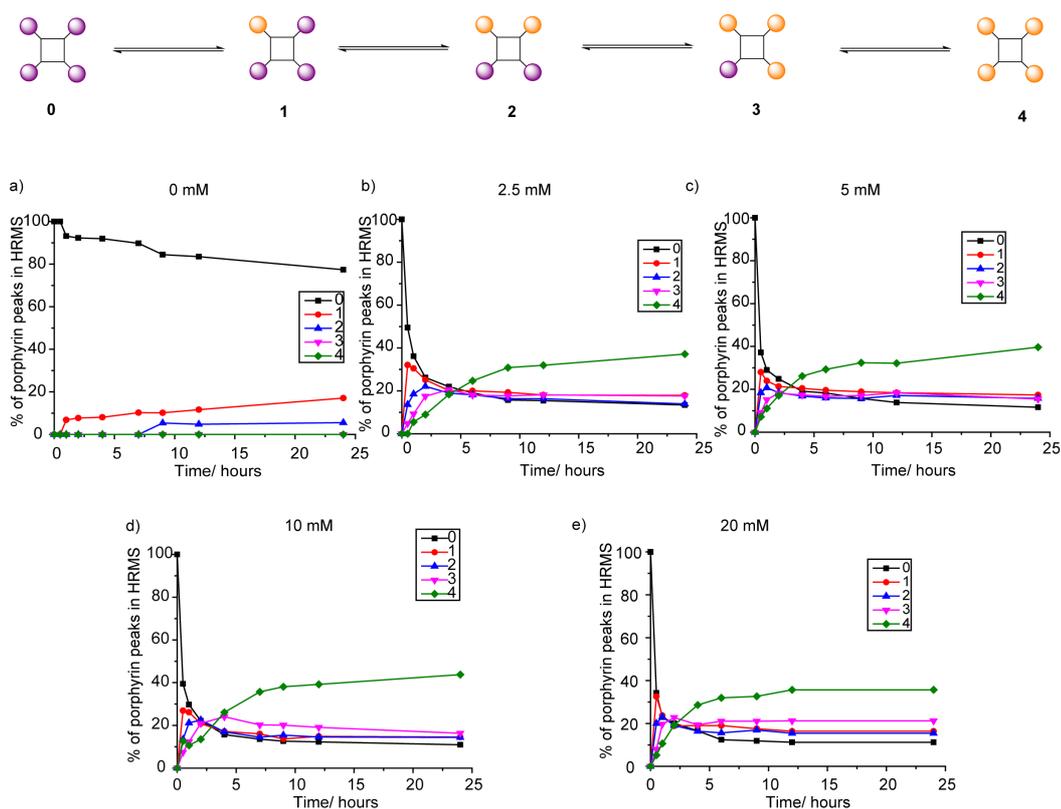


Figure 3.10 Effect of catalyst concentration on hydrazone exchanges. Exemplary data with Aspartic acid hydrazone porphyrin scaffold **79b** (100 μ M) with 2,4-dimethoxy benzaldehyde **44b** (5 mM) and varying concentrations of aniline **45** in 10 % DMSO in 5 mM NH_4OAc , pH 6.75

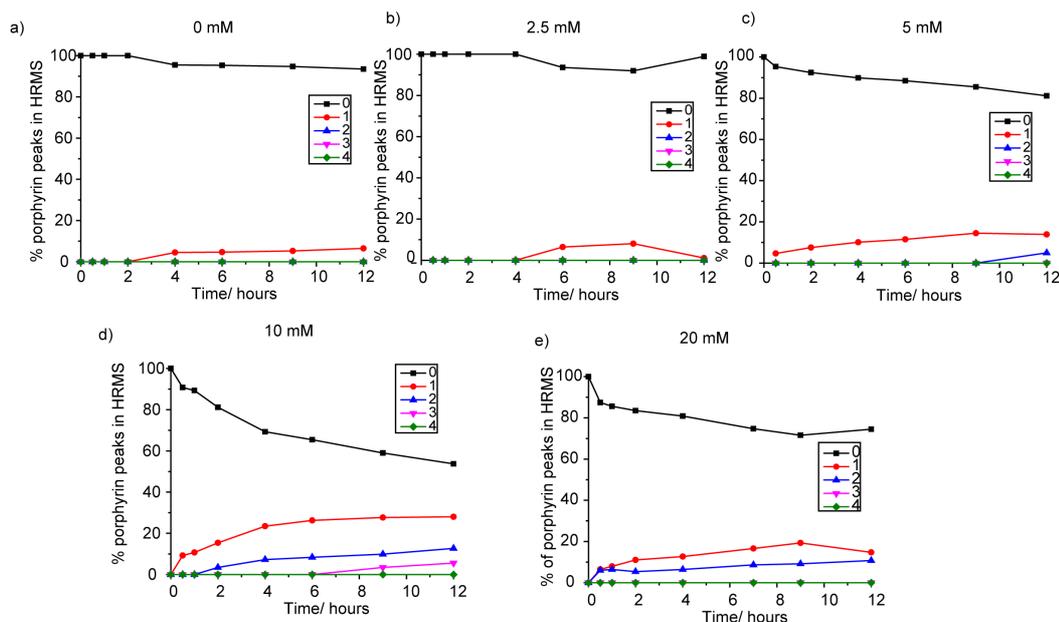


Figure 3.11 Effect of catalyst concentration on serine hydrazone porphyrin scaffold **79a**. Serine hydrazone porphyrin scaffold **79a** (100 μ M) with 2,4-dimethoxy benzaldehyde **44b** (5 mM) and varying concentrations of aniline **45** in 10 % DMSO in 5 mM NH_4OAc , pH 6.75

3.3.2.2 Is this a true dynamic equilibrium or a static mixture?

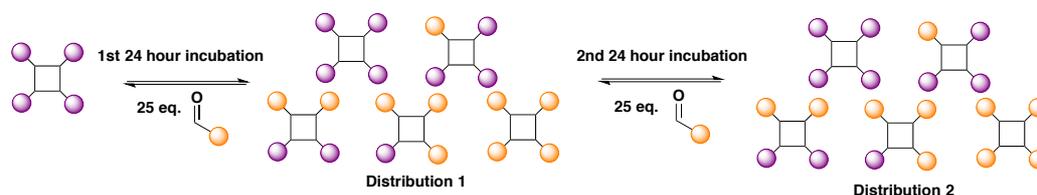


Figure 3.12 Cartoon depicting changing distribution of species on addition of second aliquot of aldehyde

Having shown that it was possible to exchange the benzaldehyde moiety for another aldehyde, it was necessary to show that the system did not reach a static point and was indeed at a dynamic equilibrium. This was achieved by taking the established mixture from incubation with 25 equivalents of an aldehyde for 24 hours then adding a second 25 equivalents of the same aldehyde to show if it was possible to perturb the equilibrium (**Figure 3.12**). This was attempted by considering just 2 time points, looking at the distributions following the first 24 hours incubation then 24 hours after the second incubation. This experiment was attempted with 4-carboxy benzaldehyde **44c**, using an aniline catalyst for all three hydrazone porphyrin scaffolds **79**. If the equilibrium had been perturbed, and the distribution of species changes on second incubation, then the system was indeed in a dynamic equilibrium rather than being static.

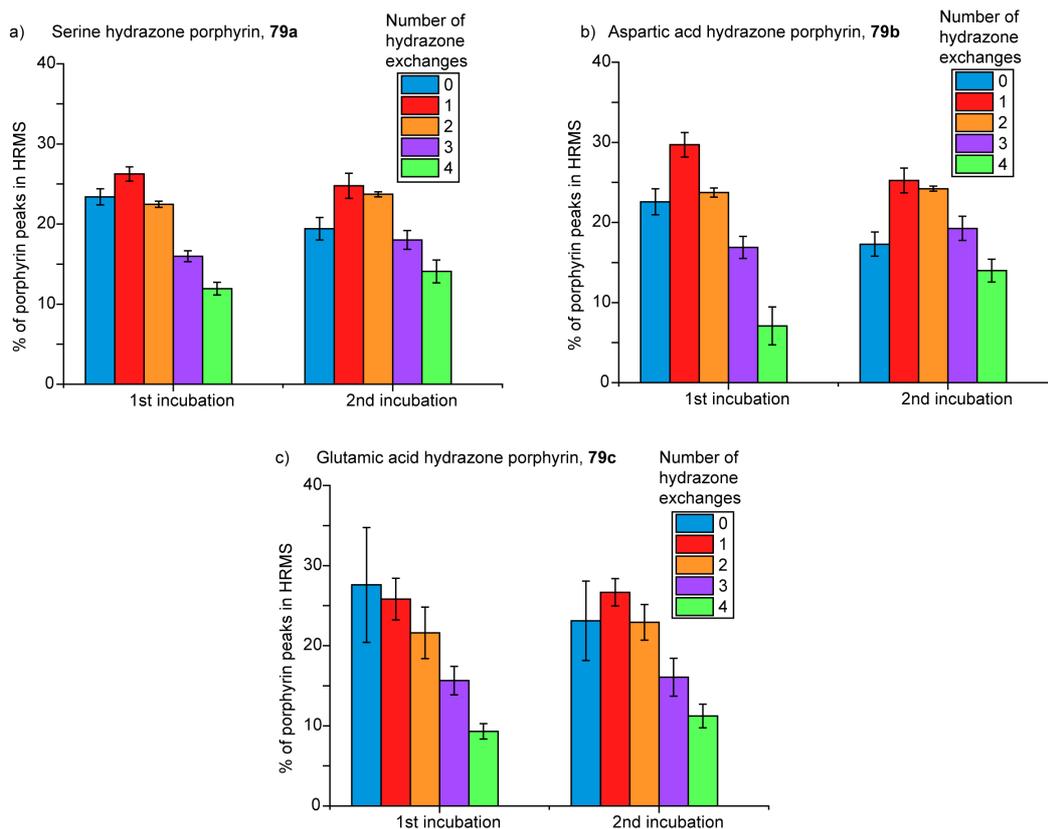


Figure 3.13 Addition of 2 separate batches of aldehyde, incubation with 25 eq. of aldehyde **44c** for 24 hours followed by addition of a second 25 eq. of aldehyde **44c**. 1st incubation of 100 μ M hydrazone porphyrin **79** with 2.5 mM 4-carboxy benzaldehyde **44c** and 10 mM aniline **45** in 10 % DMSO in 5 mM NH_4OAc , pH 6.75. 2nd incubation, addition of a further 2.5 mM 4-carboxy benzaldehyde **44c** to first incubation. Graphs show an average of 5 separate measurements. a) Serine hydrazone porphyrin **79a**, b) Aspartic acid hydrazone porphyrin **79b**, c) Glutamic acid hydrazone porphyrin **79c**

The addition of the second 25 equivalents of aldehyde **44c** does indeed change the distribution for all three porphyrin scaffolds **79** (Figure 3.13). With all three hydrazone porphyrin scaffolds **79** there is a shift to more hydrazone exchanges having occurred, with the initial species (0) decreasing and the final species (4) increasing, along with changes in the proportion of intermediary species (1, 2 and 3). As it was possible to perturb the species present, it indicates that the system is at an equilibrium rather than being a static mixture.

3.3.2.3 Hydrazone exchanges to 2 aldehydes

Having shown that a dynamic equilibrium was indeed generated it was decided to incubate the system with 2 different aldehydes, to look at the thermodynamic mixture obtained. Initially this was attempted using the single time point method described above, looking at both mixtures that had been preincubated with one aldehyde followed by

addition of the other aldehyde (**Figure 3.14b**) as well as ones where a mixture of the two aldehydes was present from the start (**Figure 3.14a**). For these studies an aniline **45** catalyst with 4-carboxy benzaldehyde **44c** and 2,4-dimethoxy benzaldehyde **44b** were used with all three hydrazone porphyrin scaffolds **79**.

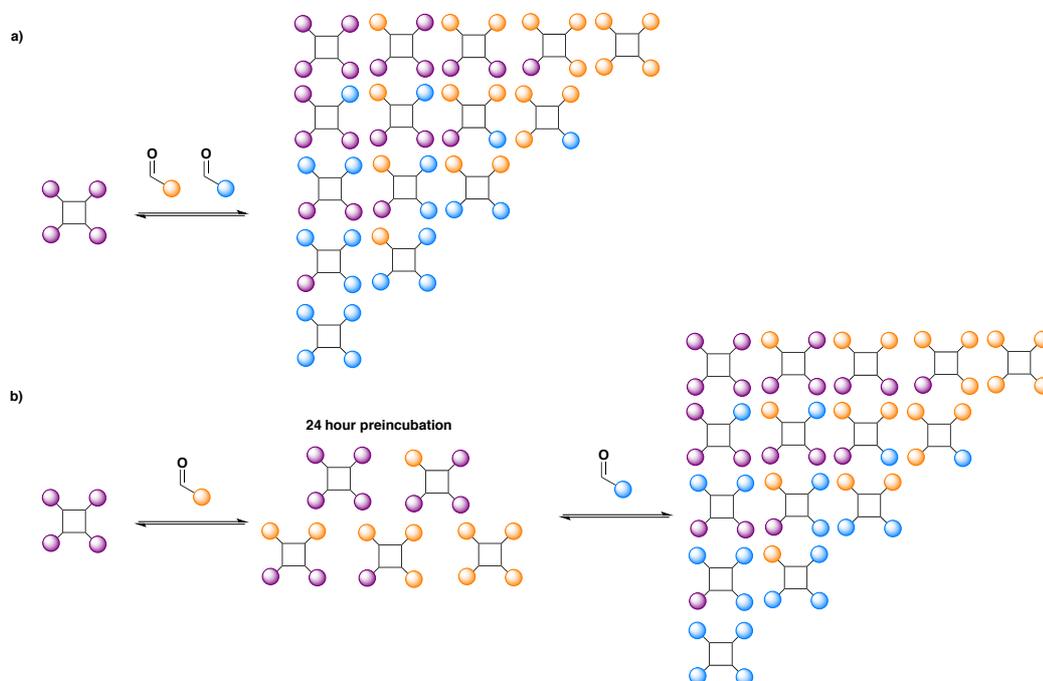


Figure 3.14 Cartoon depicting 2 methods of performing hydrazone exchange reaction with 2 aldehydes a) Mixing both aldehydes with the scaffold from the start, b) Pre-incubation with one aldehyde followed by addition of the second aldehyde

In all cases the system does give a mixture of different hydrazones (**Figure 3.15**) as hydrazone exchange occurs with both aldehyde moieties. In the case where one aldehyde has been preincubated with the hydrazone porphyrin **79** followed by addition of the other aldehyde this again showed the establishment of an equilibrium which can be perturbed.

With all three hydrazone porphyrin scaffolds **79**, the pre-incubation with 2,4-dimethoxy benzaldehyde **44b** followed by addition of 4-carboxy benzaldehyde **44c** (**Figure 3.15b, e and h**) gave a similar distribution of products to that without pre-incubation and direct mixing of the two aldehydes **44b** and **c** from the start (**Figure 3.15c, f, and i**). This indicates that it is possible to reach the same equilibrium from the two different starting points, indicating these systems are at a thermodynamic equilibrium. There was, however, a bigger difference between the pre-incubation with 4-carboxy benzaldehyde **44c** (**Figure 3.15a, d and g**) and the direct mixing (**Figure 3.15c, f and i**), this could indicate the system had not yet reached an equilibrium in this set of conditions, and may indicate a longer

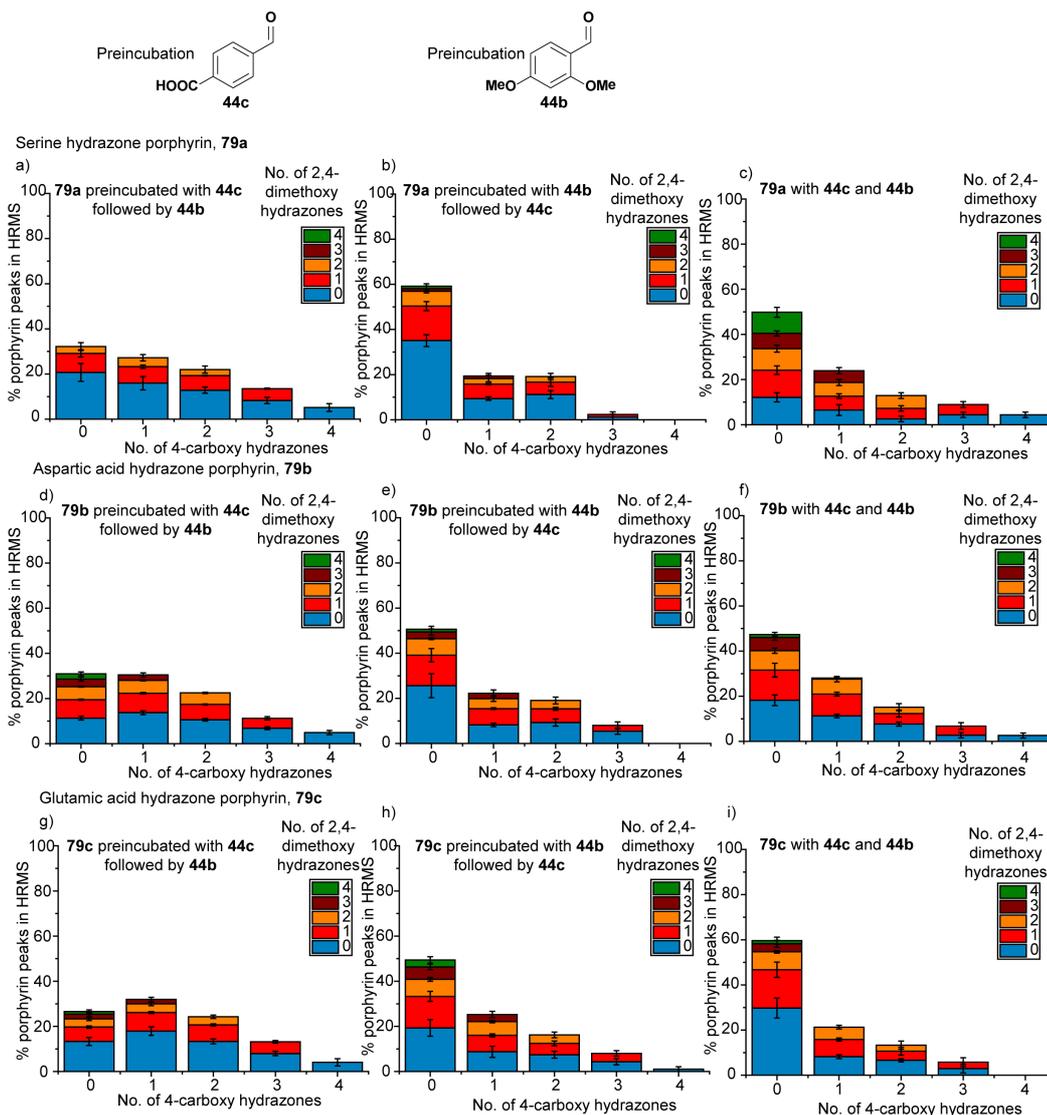


Figure 3.15 Single time point 2 aldehyde hydrazone exchanges, using 2,4-dimethoxy benzaldehyde **44b** and 4-carboxy benzaldehyde **44c**. HRMS taken after 24 hours incubation, pre-incubated systems incubated with the first aldehyde for 24 hours before addition of the second aldehyde. a), b) and c) Serine hydrazone porphyrin **79a**, d), e) and f) Aspartic acid hydrazone porphyrin **79b**, g), h) and i) Glutamic acid hydrazone porphyrin **79c**, a), d) and g) Pre-incubation with 4-carboxy benzaldehyde **44c**, b), e) and h) Pre-incubation with 2,4-dimethoxy benzaldehyde **44b**, c), f) and i) Direct mixing of both aldehydes Conditions: 100 μ M hydrazone porphyrin **79**, 2.5 mM each aldehyde (**44b** and **c**), 10 mM aniline **45**, 10 % DMSO, 5 mM NH_4OAc , pH 6.75. Graphs show an average of 5 separate incubations.

time was needed for equilibration to occur. However, some of this discrepancy may come from the method of analysis, for many of these species, especially with incubations with two aldehydes, discerning the signal for the species present above the noise in the mass spectrum could be difficult. This may be leading to the larger discrepancy in this case, as, especially in this system, after the total 48 hour incubation the signal above noise in the

mass spectrum is small. The small signal to noise ratio means that species which have relatively low actual abundance were shown as being present at a much higher percentage than if the signal to noise ratio was much larger.

3.3.2.3.1 Time to reach equilibrium with 2 aldehydes

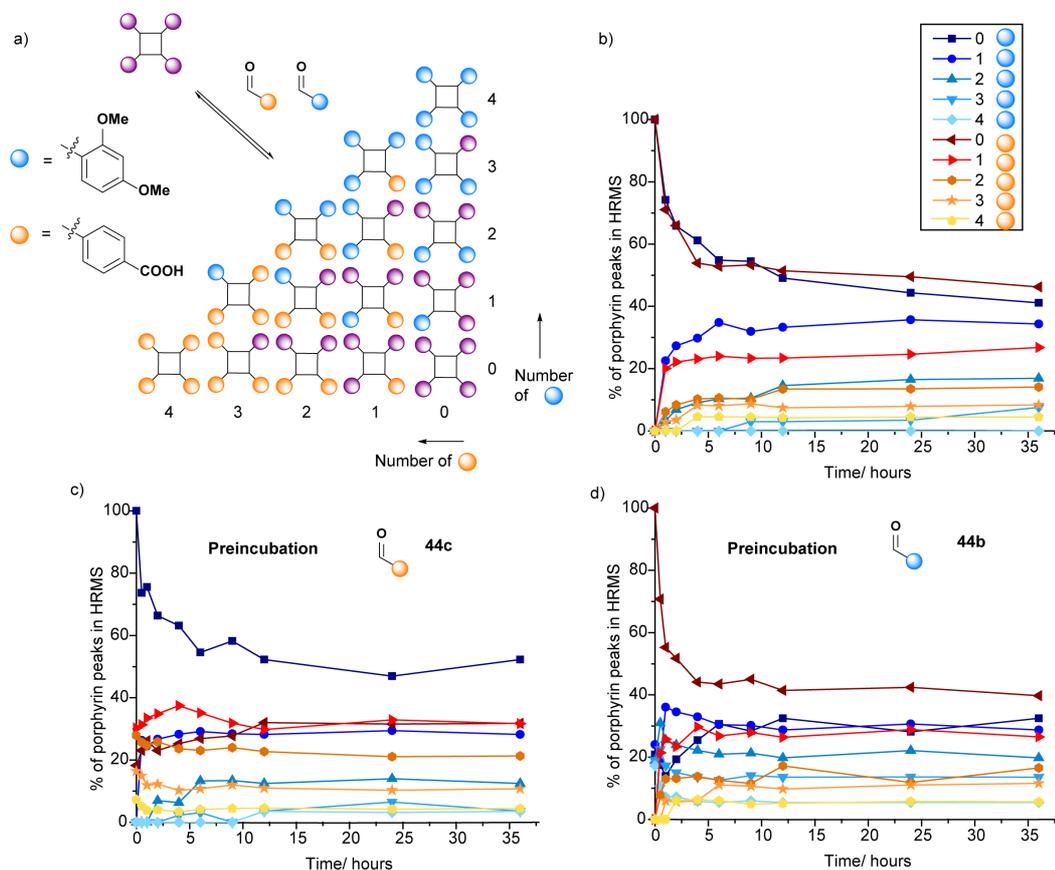


Figure 3.16 Time for porphyrin hydrazone exchanges to reach equilibrium with two aldehydes, a) Cartoon depicting reaction, b), c) and d) Time course showing the distribution of species with the number of species with each aldehyde functionality over time for the aspartic acid hydrazone porphyrin scaffold **79b** after a) Direct mixing of 4-carboxy benzaldehyde and 2,4-dimethoxy benzaldehyde b) Pre-incubation with 4-carboxy benzaldehyde **44c** and c) Pre-incubation with 2,4-dimethoxy benzaldehyde **44b**. Conditions: 100 μ M hydrazone porphyrin **79b**, 5 mM each aldehyde **44b** and **44c**, 10 mM aniline **45**, 10 % DMSO, 5 mM NH_4OAc , pH 6.75

Having shown that a mixture of different hydrazone functionalised porphyrins do form on incubation with two aldehydes, the rate of forming this equilibrium was investigated, as in the case of pre-incubation with 4-carboxy benzaldehyde **44c**, the rate of reaching the dynamic equilibrium could be different to that with just one aldehyde. This was done by following the hydrazone exchanges with 2 aldehydes over time, with exemplary data for the aspartic acid hydrozone porphyrin **79b** shown in **Figure 3.16**.

In both the pre-incubation and the direct mixing cases an equilibrium was seemingly reached within 12 hours (**Figure 3.16**). This rate is similar to that with a single aldehyde, and is a reasonable time for protein templation. Again, the species obtained from pre-incubation with 2,4-dimethoxy benzaldehyde **44b** and the species obtained from direct mixing of both aldehydes were similar, however the species present after pre-incubation with 4-carboxy benzaldehyde **44c** was somewhat different, though the system was seemingly equilibrating. It could be that in this case the species present reach a kinetic trap, with it being difficult to exchange so many 4-carboxy benzaldehydes **44c**, or it could again be due to the low signal to noise ratio present in the mass spectrum. However, using the other data obtained, it was possible to see that a dynamic equilibrium is formed from the direct mixing systems, and therefore the system is primed for incubation with a protein.

3.3.3 Protein incubation

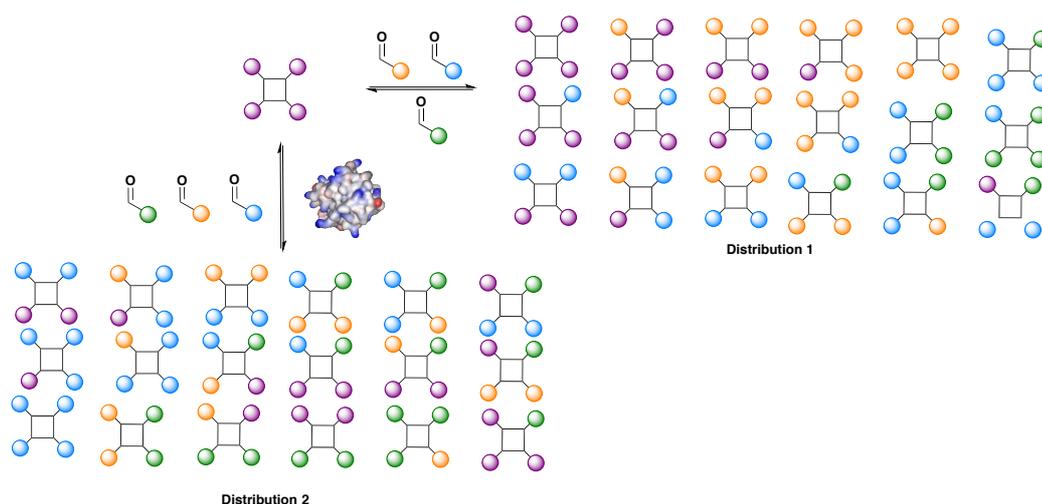


Figure 3.17 Cartoon depicting incubation with protein. Different distributions of porphyrin hydrazone products are obtained in the presence of a protein compared to its absence.

After showing that hydrazone exchange to generate a dynamic equilibrium on the porphyrin hydrazone scaffolds **79** was possible, incubation with a protein template was attempted. In this case the protein should perturb the equilibrium to enrich compounds which bind to it, as they are now more thermodynamically stable in its presence compared to its absence (**Figure 3.17**). As a first proof of principle experiment, cytochrome (cyt) *c* was chosen as the protein. Being a basic protein, it theoretically should enrich porphyrin hydrazones possessing acidic functionality

Incubation of the three hydrazone porphyrin scaffolds **79** with 2 or 3 additional aldehydes **44** (4-carboxy benzaldehyde, 2,4-dimethoxy benzaldehyde and 3-methyl 2-

carboxaldehyde pyridine) and an aniline **45** catalyst with no protein, with BSA and with cyt *c* was performed. The incubation with BSA acts as a control to see if the DCL is generating generic protein ligands or more specific cyt *c* ligands.

Initially a DCL (similar to that in the hydrazone exchange studies) was generated after a 24 hour incubation without the protein template at which point the protein was added and again incubated for 24 hours. After 24 hours the hydrazone exchange was quenched by addition of ammonium hydroxide. Then the separation of protein and the porphyrin DCL was attempted, to allow for analysis. Initially this was attempted using protein concentrators (MWCO 5 kDa) which theoretically should retain the protein, and not the porphyrin DCL. However, the porphyrins stayed in the protein concentrator, even without any protein present. It was then attempted to precipitate the protein from the DCL mixture, using ice-cold ethanol and methanol/chloroform mixtures. These gave some precipitation but mass spectra of the solutions left still showed protein present and did not allow detection of the porphyrin DCL. Analytical HPLC was also attempted on samples with no protein present, however due to the complex mixture of the DCL present the separation of the species requires long separation times, so is not high throughput or compatible with having protein still present in the sample.

3.4 Conclusions

Hydrazone exchange has been shown to be possible on multivalent scaffolds. Initially on a Ru(II)(bpy)₃ scaffold, then on a tetraphenyl porphyrin scaffold. Hydrazone exchange, as required for setting up a DCL, was shown to be possible around a Ru(II)(bpy)₃ complex hydrazone scaffold **46** in water/acetonitrile mixtures. However, attempts at making this system more biologically compatible did not prove fruitful with solubility and stability/degradation issues. In the process of this study a new synthetic method for 4,4'-disubstituted Ru(II)(bpy)₃ complexes was developed, which could prove useful for the development of new Ru(II)(bpy)₃ complexes for a variety of different applications.

Subsequently, three porphyrin hydrazone scaffolds **79** have been generated which are more soluble in biologically compatible media (10 % DMSO in aqueous solutions). Hydrazone exchange reactions were performed on these porphyrin scaffolds using several aldehydes, following the reactions over time in order to gain insight into the rates of reaching equilibration. The effects of nucleophilic catalysis using aniline **45** and anthranilic acid **80** have been studied, showing addition of these catalysts lead to equilibration within 12 hours, giving a system capable of reaching equilibrium in a timeframe suitable for templation with biological molecules. The generation of

thermodynamic mixtures after incubation with 2 aldehydes was shown, with equilibration again occurring within 12 hours, giving a system prime for incubation with different protein.

Incubation of all three hydrazone porphyrin scaffolds **79** in a DCL with cyt *c* was attempted, however separation of the protein and porphyrin DCL proved difficult, and a new method of separating the two will need to be found in order for these incubations to be analysed.

4 Using functionalised $Ru(II)(bpy)_3$ complexes in a protein sensing array

A potential application for functionalised $Ru(II)(bpy)_3$ complexes is as part of an array for the discrimination of different proteins. This works in a similar way to a mammalian nose or tongue, where a relatively small number of receptors can detect a large number of smells, with each smell having a fingerprint-like response often as a result of low affinity binding to many different smell receptors (**Figure 4.1**), thus allowing discrimination between the different smells. Here the luminescent responses of a range of functionalised $Ru(II)(bpy)_3$ complexes when incubated with different proteins were used to give a fingerprint response for the different proteins, in a 'chemical nose/tongue' approach.

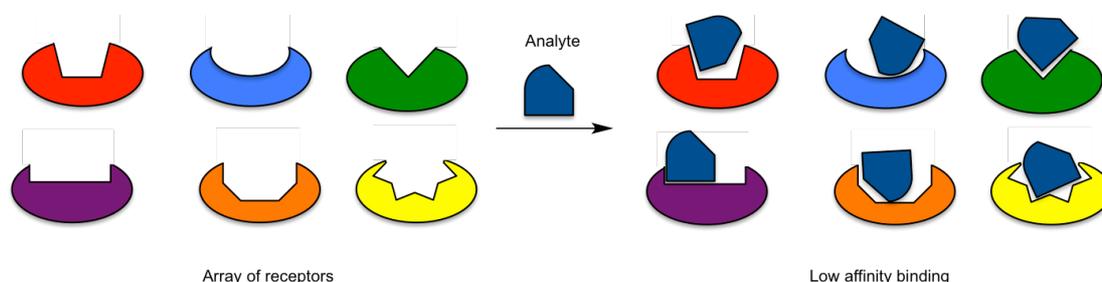


Figure 4.1 Low affinity binding of an analyte (e.g. smell molecule) to a range of receptors, generates a fingerprint response

4.1 Protein sensing arrays

Sensing arrays have been widely reported for the discrimination of different metal ions, anions and various different small molecules.¹⁷⁸ Arrays for the sensing of large biomolecules, and cells, however, have been less widely reported, but present an interesting target due to the potential for use in multiple applications, including disease diagnosis,^{179,180} and proteome sensing.¹⁸⁰⁻¹⁸² Traditional approaches for the discrimination of proteins are relatively limited in scope, relying on enzymes and antibody responses, for example using enzyme-linked immunosorbent assays (ELISA), which are cumbersome, expensive and can be plagued with false results, as illustrated by the fact that a study of antibodies for post-translational histone modifications found that 20 % of commercially available antibodies failed specificity tests.¹⁸³ Therefore it is interesting to see if synthetic molecules can be used as a cheaper, more high throughput and more specific platform for the discrimination of proteins.

In an initial, proof-of-concept experiment, the Hamilton group used functionalized porphyrins to discriminate between cyt *c*, myoglobin, cyt *c*₅₅₁ and ferredoxin to generate fingerprint-like responses for each of the different proteins.⁹⁵ They then furthered their study by using a statistical analysis technique, linear discrimination analysis (LDA), to show significant discrimination of cyt *c*, lysozyme, ferredoxin and α -lactalbumin.^{94,95}

The Anslyn group subsequently described discrimination of proteins and glycoproteins and some 'within-protein-class' discrimination of proteins,¹⁸⁴ using combinatorially synthesised tripeptide and boronic acid functionalised hexasubstituted benzenes. They have also used Zn(II)-(dpa) complexes to discriminate between different classes of MAP kinases.¹⁸⁵ The Rotello group have used fluorescent polymers,¹⁸¹ fluorescent polymer appended gold nanoparticle sensors,^{179,180} and fluorescent protein-appended gold nanoparticle^{180,186} for the detection of a range of proteins,¹⁸⁰⁻¹⁸² and for the detection of metastatic cells.^{179,180} These approaches allowed for quantification of protein concentration¹⁸⁰ and for protein detection within human serum.¹⁸² More recently, array approaches have also been developed for the detection of histone modifications;^{28,187} the Hof group have used different functionalized calixarenes with displaceable indicator dyes to discriminate various different histone post-translational modifications.²⁸

4.2 Discrimination of proteins using a range of Ru(II)(bpy)₃ complexes

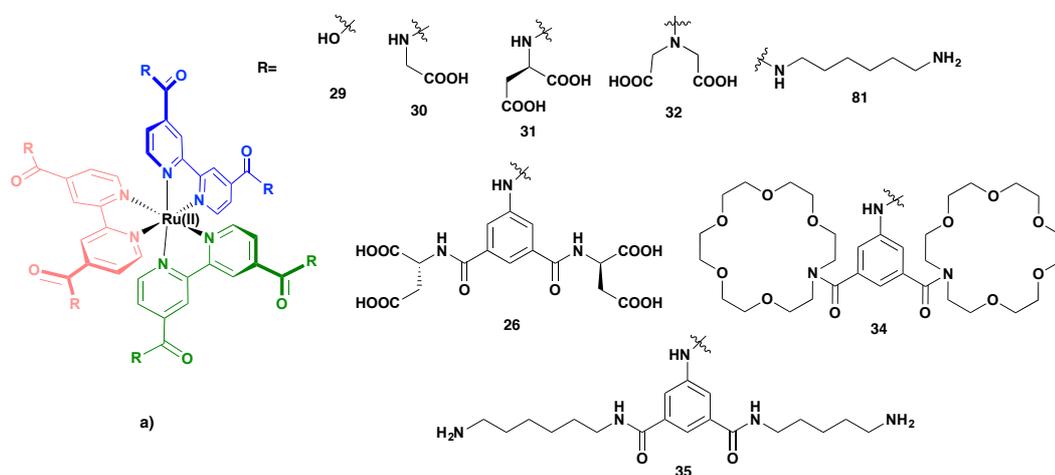


Figure 4.2 The Ru(II)(bpy)₃ complexes used in the protein sensing arrays

Initially, similar to the Hamilton group studies with functionalised porphyrins,^{94,95} a range of proteins were incubated with a range of Ru(II)(bpy)₃ complexes (**Figure 4.2**). The Ru(II)(bpy)₃ complexes possess different groups around the periphery and present these groups in different spatial orientations, therefore they should interact with the different protein surfaces in different ways, for example binding with varying affinities, repulsion,

aggregation, disaggregation and affecting the interaction of the Ru(II)(bpy)₃ complex with the well plate (**Figure 4.3**). All these different interactions can lead to changes in the luminescence intensity of the functionalised Ru(II)(bpy)₃ complexes, with the potential for quenching or enhancement of their luminescence, which can readily be detected.

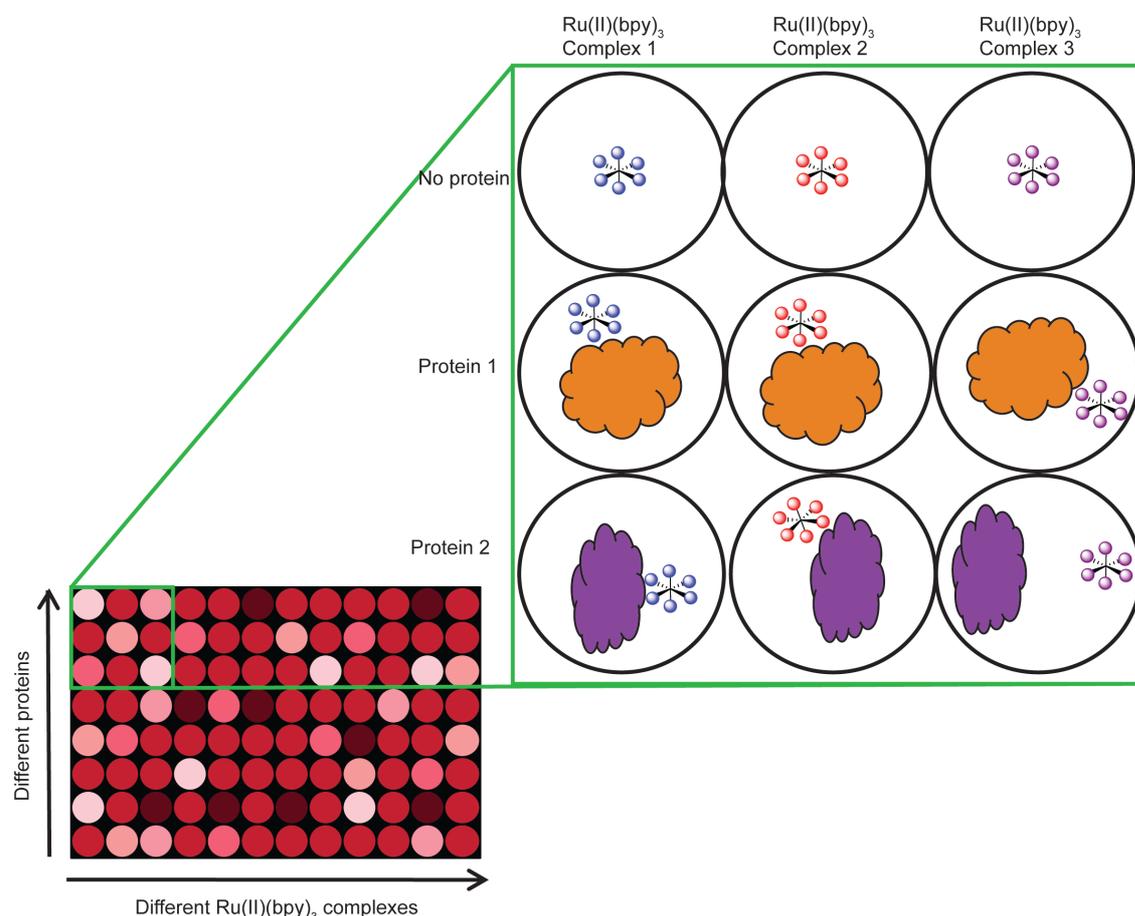


Figure 4.3 Cartoon illustrating differential sensing with Ru(II)(bpy)₃ complexes and proteins. The various combinations of Ru(II)(bpy)₃ complexes and proteins are incubated in a 96 or 384 well plate, with differential interactions between the Ru(II)(bpy)₃ complexes and proteins, for example, different Ru(II)(bpy)₃ complexes binding to the same site or to distal sites on the protein or some Ru(II)(bpy)₃ complexes binding and others not, or even being repelled. These differential interactions lead to differences in the luminescence spectra of the Ru(II)(bpy)₃ complexes.

Eight different Ru(II)(bpy)₃ complexes (**26**, **29 – 32**, **34**, **35** and **81**) (**Figure 4.2**) were incubated with nine different proteins (lysozyme, α -chymotrypsin (α -ChT), horse cyt *c*, papain, ribonuclease (RNase) A, bovine serum albumin (BSA), yeast cyt *c*, *hDM2* and Mcl-1) (**Figure 4.4**). The luminescence intensity of each of the Ru(II)(bpy)₃ complexes with each of the proteins was measured and compared to its luminescence intensity without any protein present. The comparison between the presence and absence of protein is

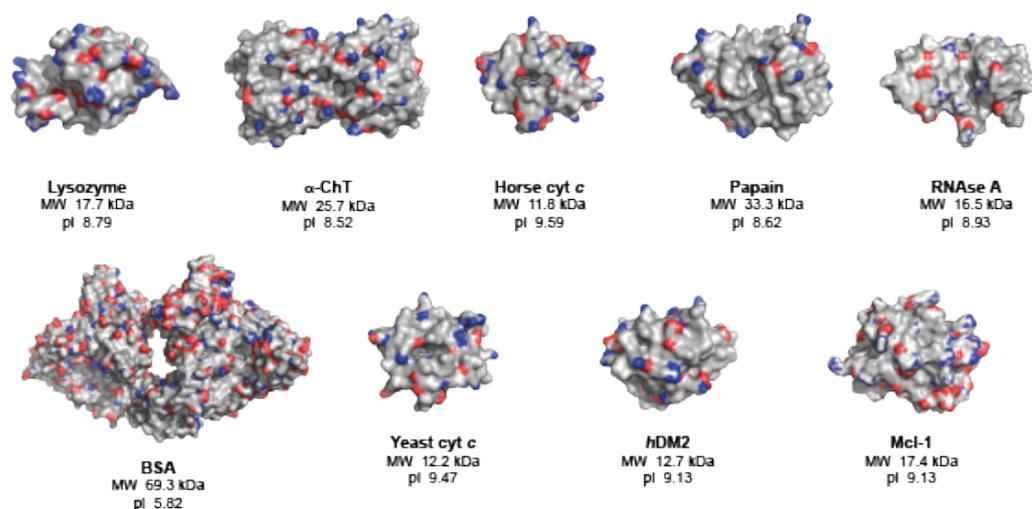


Figure 4.4 X-ray crystal structures of the nine proteins for use in the array, showing the basic (blue) and acidic (red) amino acid residues, molecular weights (MW) and isoelectric points (pI) of the proteins; lysozyme (PDB ID: 1LYZ), α -chymotrypsin (α -ChT) (PDB ID: 4CHA), horse cyt *c* (1HRC), papain (PDB ID: 9PAP), RNase A (PDB ID: 5RSA), bovine serum albumin (BSA) (PDB ID: 3V03), yeast cyt *c* (PDB ID: 2YCC), *hDM2* (PDB ID: 4GH7) and Mcl-1 (PDB ID: 5C3F)

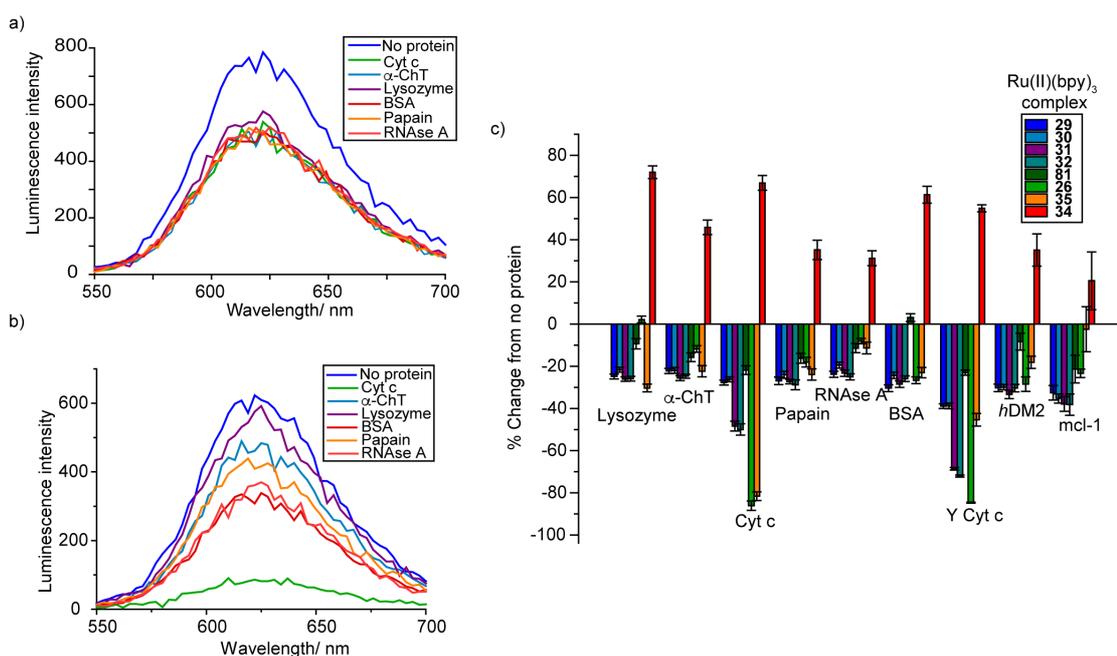


Figure 4.5 Differential luminescence responses from the different Ru(II)(bpy)₃ complexes **26**, **29-32**, **34**, **35** and **81** (2.5 μ M) on incubation with various different proteins (10 μ M), (5 mM sodium phosphate, pH 7.5, exc. 467 nm). a) and b) Illustrative luminescence intensity over variable wavelengths for Ru(II)(bpy)₃ complex **30** (a) and **26** (b), obtained using plate reader monochromators. c) Fingerprint changes in luminescence responses from all the different Ru(II)(bpy)₃ complexes with the different proteins, illustrated as percentage differences from incubation with and without protein, responses measured using plate reader fixed filters (exc. 467 nm, em. 625 nm)

necessary as the magnitude of the intensities of the Ru(II)(bpy)₃ complexes varies between plates.

For each of the proteins a fingerprint-like response (**Figure 4.5**) with different luminescence responses from each of the Ru(II)(bpy)₃ complexes on addition of different proteins was obtained. Most of the Ru(II)(bpy)₃ complexes show varying luminescence quenching on incubation, and hence interaction with, the proteins. Ru(II)(bpy)₃ complex **34** shows varying degrees of enhanced luminescence intensity on interaction with each of the different proteins. This indicates that this may be a method in which it is possible to discriminate between different proteins.

4.2.1 Statistical analysis

From the bar chart in **Figure 4.5c**, it can be seen that there are differences in luminescence response to the individual proteins with the different Ru(II)(bpy)₃ complexes, however the discrimination of the proteins, by eye, is not trivial, especially if the number of Ru(II)(bpy)₃ complexes or proteins were to be increased. Therefore it is necessary to simplify the data from 8 dimensional data (arising from the 8 Ru(II)(bpy)₃ complexes) to a more manageable number of dimensions, 2 or 3 dimensions which can readily be readily visualised on a scatter graph.

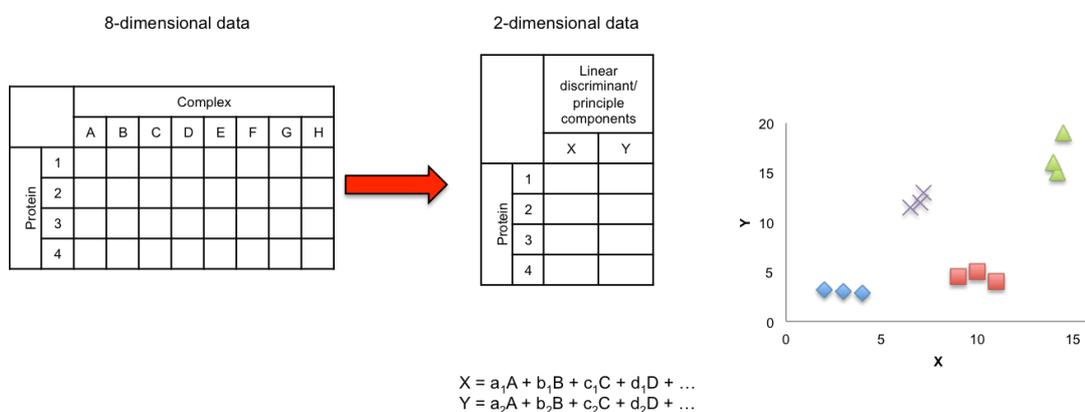


Figure 4.6 PCA/LDA aims to take the 8 dimensional Ru(II)(bpy)₃ complex array data down to two (or three) dimensions by taking linear combinations of the original data, thus allowing the data to be plotted on a 2-D (or 3-D) scatter graph, with the aim of getting clusters for the different analytes (proteins)

Two different statistical techniques, principal component analysis (PCA) and discriminant analysis (DA), have been widely used in the literature for reducing the dimensionality of array data.¹⁸⁸ These two techniques generate score plots using combinations of the original data on each of the axes, in order to take the n-dimensional

data down to 2 or 3 dimensions, allowing the data to be plotted graphically (**Figure 4.6**).¹⁸⁸ Both techniques reduce the data using matrix techniques, finding eigenvectors and eigenvalues to describe new axes and the level of discrimination respectively. Matrix derivations and discussion of the mathematics behind the techniques are discussed in more detail in Appendix III.

DA is a supervised statistical method,¹⁸⁹ meaning it uses a training set of data, and looks for the best way of organising the data so as to increase the discrimination between classes while decreasing the variation within classes. PCA is unsupervised and looks at finding the maximum variation between all the data, irrespective of the data classes.¹⁹⁰ This means that DA looks for clustering of classes whereas PCA spreads data points out as much as possible.¹⁸⁸ As the ultimate aim with this analysis is to see if it is possible to cluster all the data from each protein, and separate that cluster from that of the other proteins, discriminant analysis has been used.

In this case linear discriminant analysis (LDA) was used, taking linear combinations of the original components as described earlier. This makes an assumption that all the covariance matrices for all the different classes are equivalent, i.e. how much the data for a single protein/Ru(II)(bpy)₃ complex combination compares with another protein/Ru(II)(bpy)₃ complex combination within the same replicate is the same across different replicates for the same protein. This means that it is assumed that the luminescence responses for a specific Ru(II)(bpy)₃ complex with a protein should be the same across replicates. Quadratic discriminant analysis (QDA) does not make this assumption, finding quadratic surfaces rather than linear combinations of the original data.¹⁸⁸ However, QDA requires many more replicates to be performed as it requires more parameters to be derived from the data set.¹⁹¹

4.2.1.1 Linear discriminant analysis on Ru(II)(bpy)₃ complex array with a panel of proteins

Initially a 2-D LDA was performed (**Figure 4.7a**) on the array data obtained. This showed clear and distinct clusters for the two cyt *c* proteins (horse heart and yeast cyt *c*). This is a promising result, as cyt *c* is a highly evolutionarily conserved protein so separation of cyt *c* from two different species could be seen as difficult to achieve. Clustering was also seen for the other proteins, but these clusters are all located within the same region. The use of 3-D LDA was then explored, in the hope of discriminating between these clusters. The 3-D LDA (**Figure 4.7b**) again shows distinct clusters for the cyt *c* proteins, and clustering of the other proteins. However, the clustering of the other

proteins is again in one region, with little discrimination between them, therefore array was changed in order to increase the discrimination between these proteins.

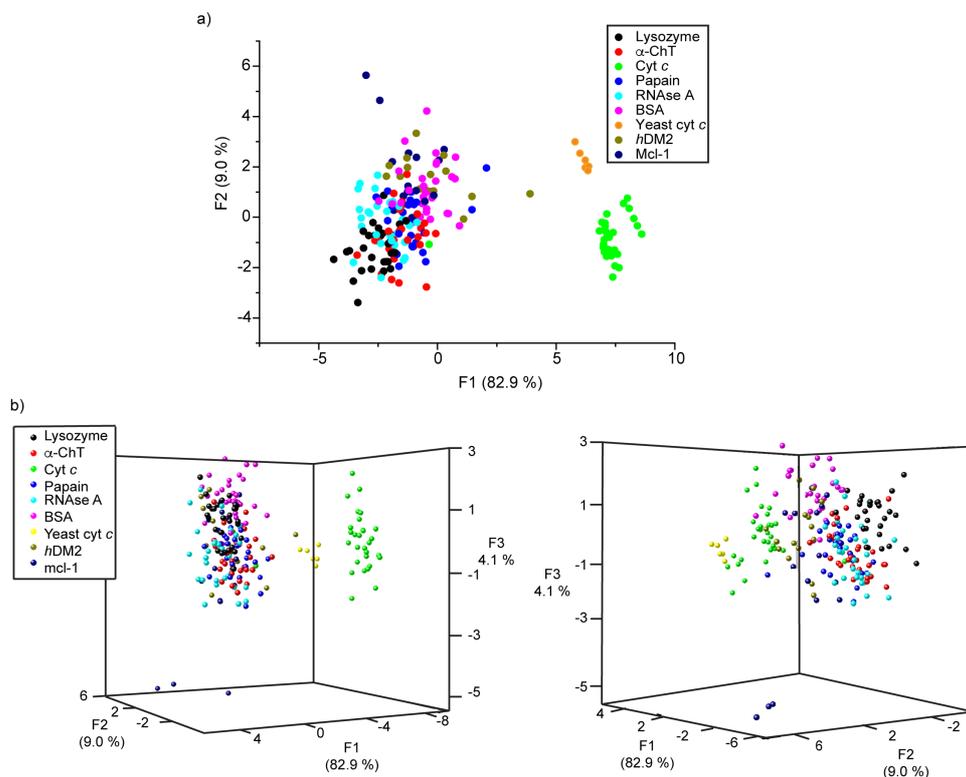


Figure 4.7 LDA for the Ru(II)(bpy)₃ complex-protein array, a) 2-D analysis, b) 2 views of the 3-D analysis

4.3 Using 2 fluorophores to enhance protein discrimination

As all the proteins could not be discriminated using just the Ru(II)(bpy)₃ complexes the use of a second fluorophore to produce better discrimination was investigated. With two different fluorophores present a much more complicated network of interactions becomes possible leading to changes in the luminescence of the two separate fluorophores (**Figure 4.8**).

Array studies with molecules possessing multiple fluorophores have been reported by the Margulies group, where they synthesised complicated molecules in order to detect a variety of different analytes, including a range of pharmaceuticals in urine samples,¹⁹² and more recently different aggregation states of β -amyloid peptides.¹⁹³ The molecules used in the Margulies group studies are large and require considerable synthetic effort, however it was thought that it may be possible to achieve the same goal of using multiple interacting

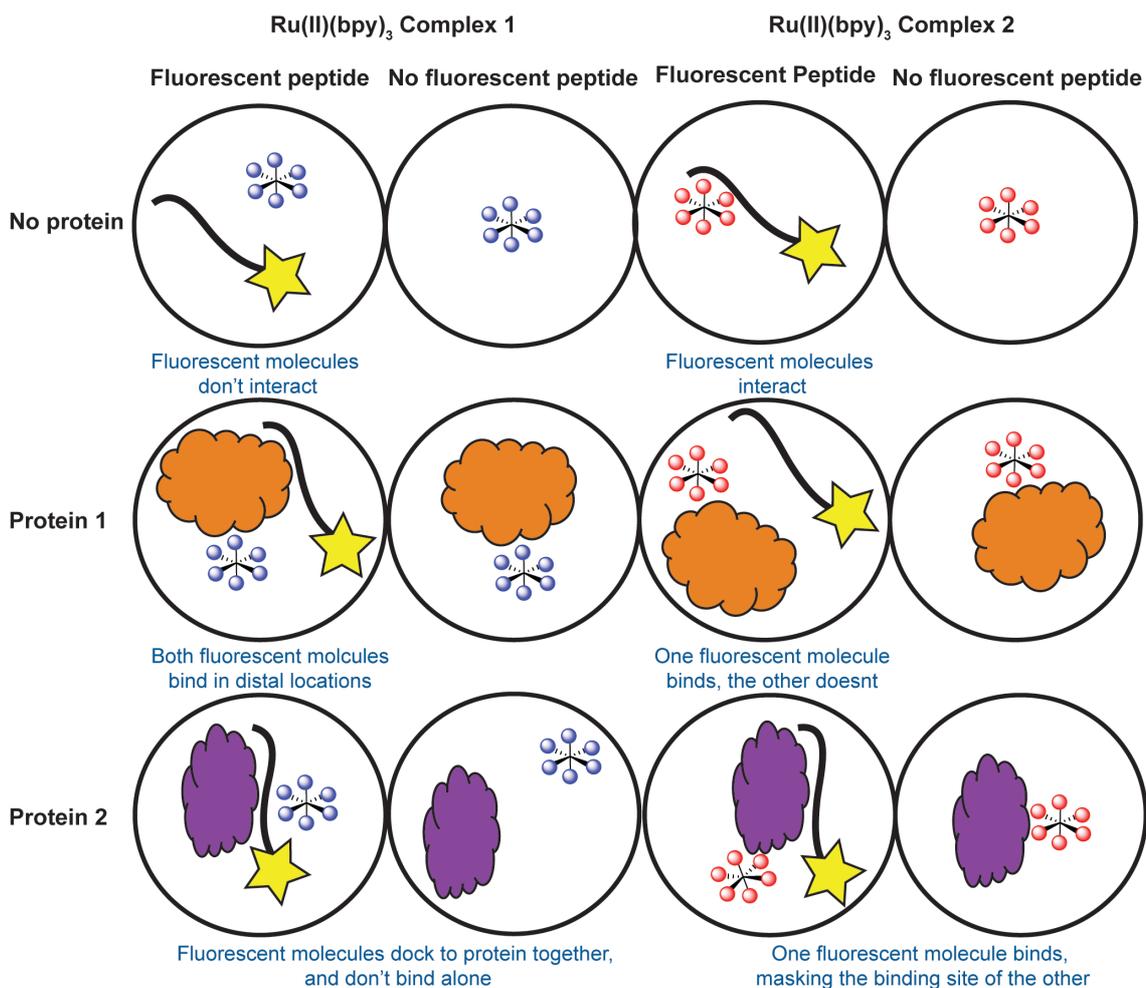


Figure 4.8 Array of potential different interactions between the two fluorophores and different proteins, including interactions between the two luminescent molecules, binding on different sites on the protein, masking of binding sites for one molecule by the other molecule, docking of the two molecules together on the protein

fluorophores by using different luminescent molecules already synthesised within the Wilson group. To this end the Ru(II)(bpy)₃ complexes described previously and a FITC-NOXA B peptide (sequence FITC-(Ga)AAQLARIGDKVNLQKLLN-NH₂), which had been synthesised by Dr. Katherine Horner, were used. Fluorescein is a logical fluorophore choice in this case; although its maximum excitation wavelength 494 nm, is higher than that used for the Ru(II)(bpy)₃ complexes (467 nm), 467 nm is still within its excitation spectrum and can therefore be used, allowing simultaneous visualisation of the luminescence of both the Ru(II)(bpy)₃ complexes (emission 625 nm) and the fluorescent peptide (emission 520 nm).

4.3.1 Deciding on appropriate concentrations

Due to the differences in quantum yield of the Ru(II)(bpy)₃ complexes and the fluorescein moiety, appropriate concentrations of the species to be ascertained, which i)

allowed the luminescence of both species to be visualised, ii) didn't have overlap of the emission bands for the two fluorophores, iii) didn't require high concentrations of protein and iv) would be responsive to protein recognition.

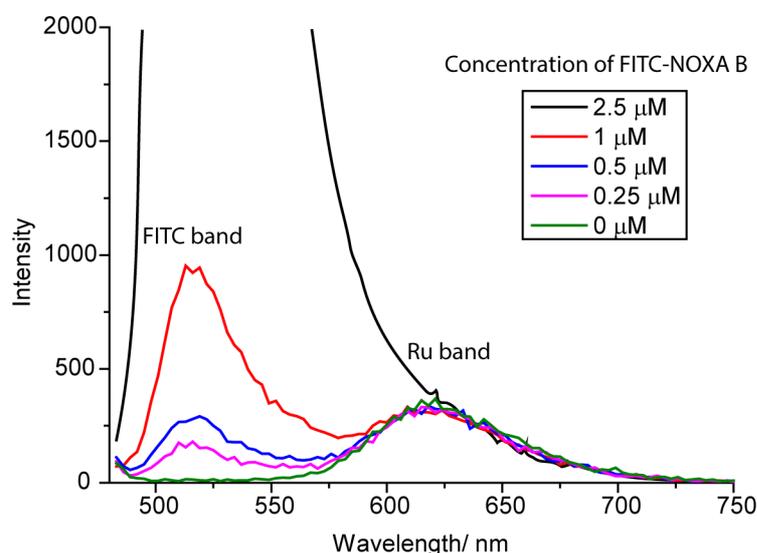


Figure 4.9 Concentration test to determine appropriate concentration of FITC-NOXA B peptide to use with 2.5 μM $\text{Ru}(\text{II})(\text{bpy})_3$ complex. Luminescence emission intensity at variable wavelengths (exc. 467 nm) upon incubation of 2.5 μM $\text{Ru}(\text{II})(\text{bpy})_3$ complex **29** with various concentrations of FITC-NOXA B peptide in 5 mM sodium phosphate, pH 7.5

2.5 μM $\text{Ru}(\text{II})(\text{bpy})_3$ complex solution and 10 μM protein were the concentrations selected for further study as these concentrations worked well with the $\text{Ru}(\text{II})(\text{bpy})_3$ array, showing specific responses to different proteins, even if they could not be discriminated. Therefore it was only necessary to decide on a compatible concentration of the FITC-NOXA B peptide. The fluorescence of various concentrations of the FITC-peptide with 2.5 μM of $\text{Ru}(\text{II})(\text{bpy})_3$ complexes **29** was characterised (**Figure 4.9**). From this, it was decided to use 0.5 μM FITC-NOXA B in future array studies as it had similar luminescence intensity to the $\text{Ru}(\text{II})(\text{bpy})_3$ complexes at 2.5 μM , with excitation at 467 nm, and allowed the maxima for both emission bands to be visualised.

4.3.2 Discrimination of proteins

After deciding on the concentrations to use, an array with 6 different proteins (cyt *c*, BSA, RNase A, papain, α -ChT and lysozyme) was performed, and the luminescence of the $\text{Ru}(\text{II})(\text{bpy})_3$ complexes and the FITC-NOXA B peptide recorded and compared to that without the protein present. The incubation with the different proteins, again gave rise to differential responses from the $\text{Ru}(\text{II})(\text{bpy})_3$ complexes (**Figure 4.10**), as well as the FITC-

NOXA B peptides, giving a fingerprint-like response for the different proteins (**Figure 4.11**).

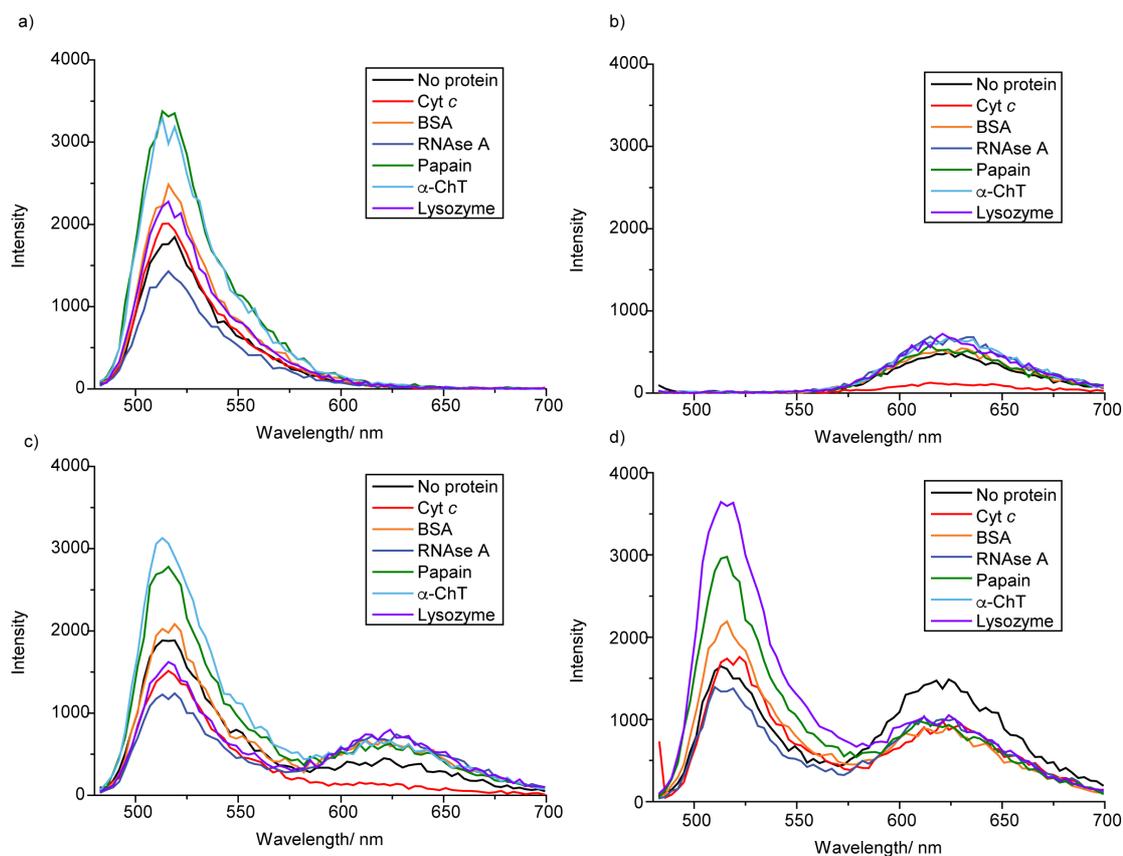


Figure 4.10 Luminescence response (exc. 467 nm, in 5 mM sodium phosphate, pH 7.5, 2 hour incubation) of the FITC-NOXA B peptide alone (a), Ru(II)(bpy)₃ complex **26** alone (b), FITC-NOXA B with Ru(II)(bpy)₃ complex **26** (c) and FITC-NOXA B with Ru(II)(bpy)₃ complex **30** (d), showing differences in the spectra.

Generally the luminescence intensities of the different Ru(II)(bpy)₃ complexes was quenched (by variable amounts) on addition of the various different proteins, as in the Ru(II)(bpy)₃ complex-protein array discussed earlier. In the majority of cases the fluorescence intensity of the FITC-NOXA B peptide was enhanced. This is as would be expected for any binding to a protein surface as fluorescein is a solvatochromic fluorophore, changing its fluorescence dependent on the solvent (or local environment). Differences in its spectral properties dependent on polarity,¹⁹⁴ the state of hydrogen bonding in its local environment,¹⁹⁵ and pH.¹⁹⁶ Both hydrogen bonding and the polarity will be affected by binding to a protein surface.

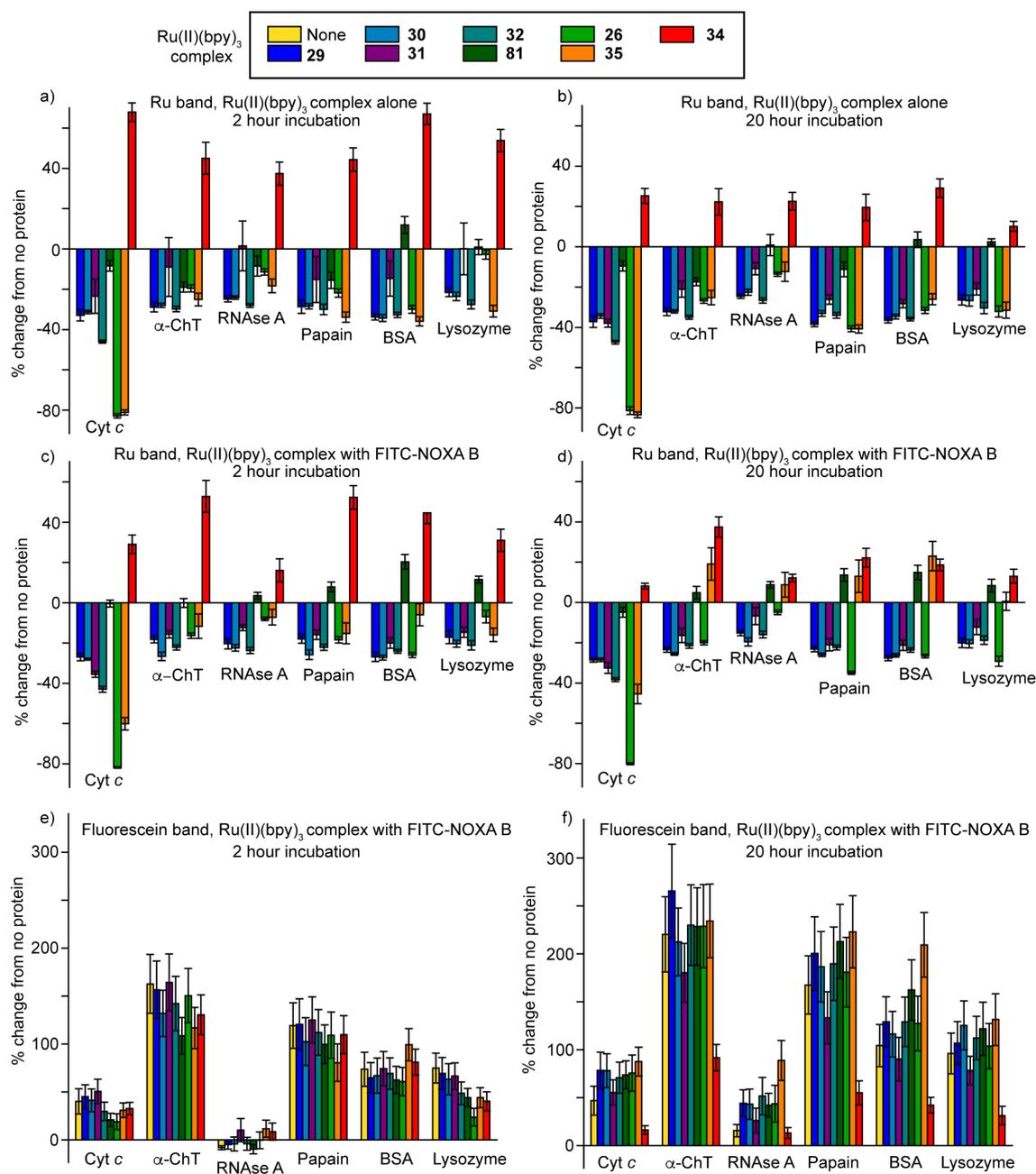


Figure 4.11 Fingerprint luminescent responses for the 6 different proteins on incubation of Ru(II)(bpy)₃ complexes alone (a) and b)), Ru(II)(bpy)₃ complexes with FITC-NOXA B peptide and the FITC-NOXA B peptide (c), d), e) and f)), looking at both the fluorescein (emission 520 nm) (e) and f)) and Ru(II)(bpy)₃ channels (emission 630 nm) (c), d), e) and f)) on excitation at 467 nm where appropriate. In 5 mM sodium phosphate, pH 7.5 buffer, 2.5 μM Ru(II)(bpy)₃ complex, 0.5 μM FITC-NOXA B peptide and 10 μM protein where appropriate, readings using fixed filters. a), c) and e) 2 hour incubation, b), d) and f) 20 hour incubation

Another factor which was identified as important for this array was incubation time. Data was collected after 2 hour and 20 hour incubations (Figure 4.11a, c and e and b, d and f respectively) of the same plate, with subtly different responses obtained after these two different incubations. The data obtained after both these incubations was

reproducible, indicating that after both these incubations, the system was indeed at equilibrium, however there was a different equilibrium present at these times. These differences might be expected for the two proteases used in the arrays (α -ChT and papain), as these can potentially degrade both themselves and the FITC-NOXA B peptide at different rates, however differences are observable for all the proteins, this shows that the system is more complicated. For example, it could be that incubation for long periods of time with the Ru(II)(bpy)₃ complexes and the FITC-NOXA B leads to changes in protein structure, indeed it has been shown that the Ru(II)(bpy)₃ complex **26** destabilises *cyt c*, lowering its melting temperature.¹⁴⁴

4.3.2.1 LDA of 6 protein, 2 fluorophore array

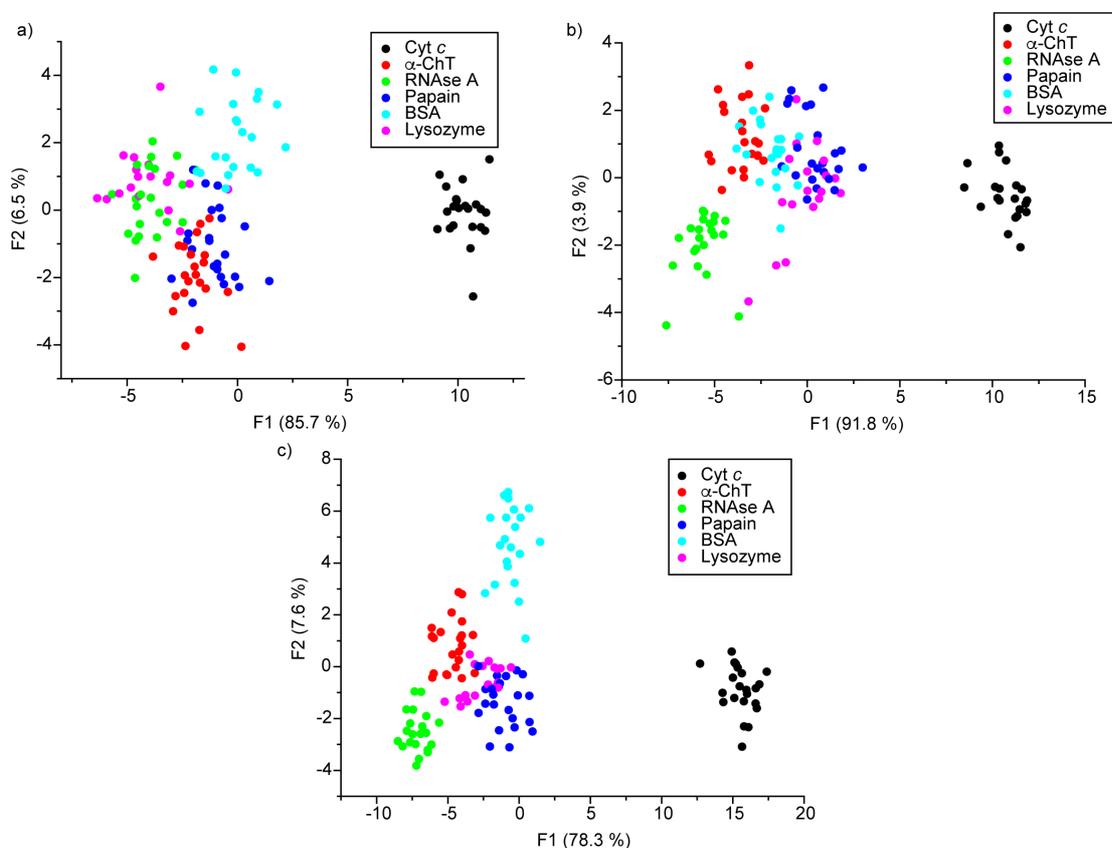


Figure 4.12 2-D LDA for the two fluorophore array with 6 proteins, a) 2 hour incubation data, b) 20 hour incubation data, c) Combined 2 and 20 hour incubation data

The data from these arrays were analysed using LDA. In two dimensions (**Figure 4.12**), looking at the 2 discriminants which define the most between class variation, it is seen that all the proteins are within distinct clusters, for the 2 hour incubation, 20 hour incubation and the combined data. In all three cases the data for the different proteins cluster together, with the *cyt c* always being in a distinct separate cluster. After 2 hours the BSA cluster is separate from the other protein clusters, and after 20 hours the RNase A

cluster is separate from the other proteins. Combining the 2 hour and 20 hour incubation data shows separation of both of these proteins, however there is still significant overlap between α -ChT and lysozyme and papain and lysozyme. To separate these protein clusters further 3-D LDA was used (**Figure 4.13**), assigning the 3rd linear discriminant (F3) as the z axis.

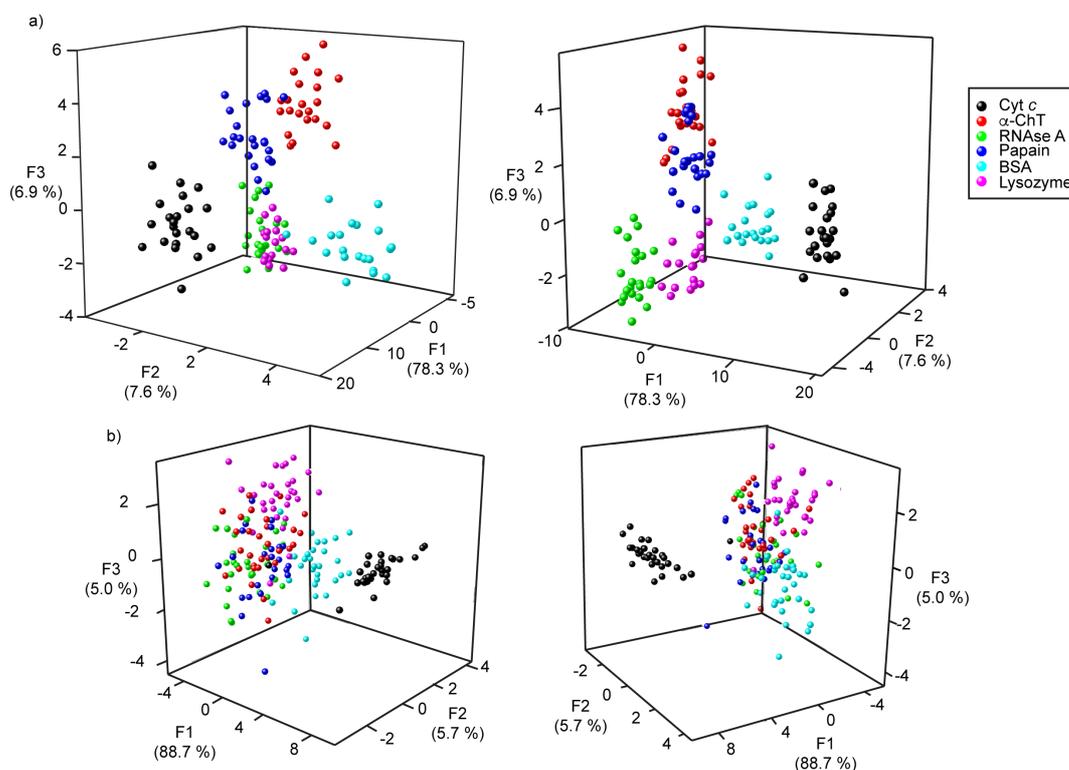


Figure 4.13 3-D LDA of 2 fluorophore array data and Ru(II)(bpy)₃ complex array for 6 proteins (combined 2 hour and 24 hour incubation data). a) 2 fluorophore array showing separation of the clusters for the 6 proteins, b) LDA of Ru(II)(bpy)₃ complex array with the same 6 proteins, doesn't show separation of the clusters for the same 6 proteins

The 3-D LDA (**Figure 4.13a**), again, shows the clearly defined clusters and now shows separation of the clusters for all the proteins, thus showing that using this technique it is possible to discriminate these 6 proteins. Compared to the original Ru(II)(bpy)₃ array (**Figure 4.13b**), where there is only discrimination of cyt *c* from the other proteins, there is a clear and distinct difference with the addition of this second fluorophore, allowing for much better discrimination of the different proteins. Looking at some subsets of the data obtained for the two fluorophore array, for example only considering the 2 hour or the 20 hour incubation readings or only considering the Ru(II)(bpy)₃ or FITC bands, do not allow this level of discrimination, showing a requirement to use all the data obtained.

It is possible to plot confidence ellipsoids (**Figure 4.14**), to see the confidence of discrimination between the different proteins. For this 6 protein array it is seen that at the 95 % confidence level (**Figure 4.14a**) the ellipsoids are clearly distinct but have some overlap, this means that at the 95 % confidence level the proteins are not completely distinguished from each other. However, at the 80 % confidence level (**Figure 4.14b**), the ellipsoids are distinct from each other, showing significant discrimination between the proteins at this level.

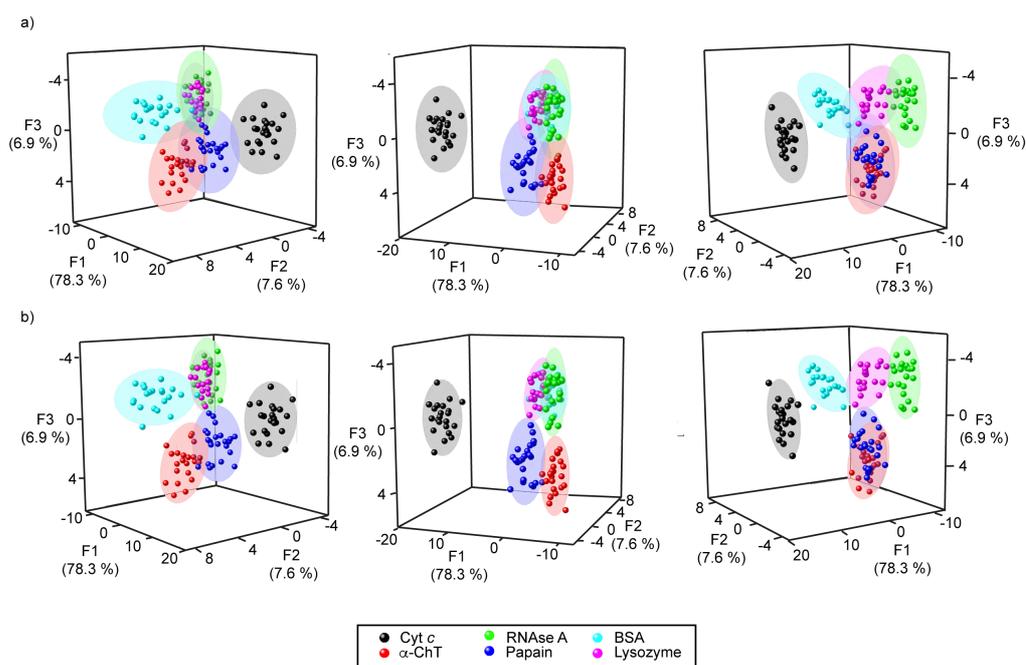


Figure 4.14 Confidence ellipsoids for the LDA of the 6 protein, 2 fluorophore array a) 95 % confidence ellipsoids, b) 80 % confidence ellipsoids

4.3.3 Including more therapeutically interesting proteins

Having had success in discriminating commercially available proteins, more therapeutically interesting proteins were added to the array, namely Mcl-1 and *hDM2*, which had been prepared by Kirstin Spence, so as to demonstrate the diagnostic potential of the approach. The protein incubations were performed in the same manner as for the original array.

A 3-D LDA was performed with the data obtained for all the proteins, again showing clearly defined clusters of each of the proteins, however there was considerable overlap of the clusters for some of the proteins, with all the data for 5 of the proteins (BSA, α -ChT, papain, RNase A and lysozyme) existing on a straight line parallel to the z (F3) axis. Using other discriminants (the fourth and fifth discriminants) to replace the first, second or both discriminants, still did not separate these clusters.

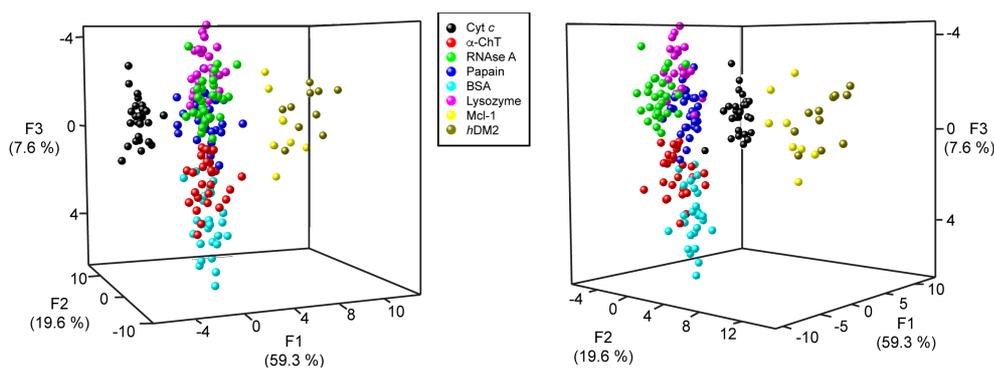


Figure 4.15 3-D LDA of 2 fluorophore array data after addition of Mcl-1 and *hDM2* incubations

From all the LDA performed, *cyt c* was always well separated from all the other proteins; this could mean that the separation of *cyt c* from all the other data dominates in the LDA, therefore it was hypothesised that it might be possible to separate the other proteins if *cyt c* was removed from the analysis. Upon removing the *cyt c* data from the analysis (**Figure 4.16**) it was observed that the clusters were much more effectively separated, demonstrating that it may be possible to discriminate the different proteins.

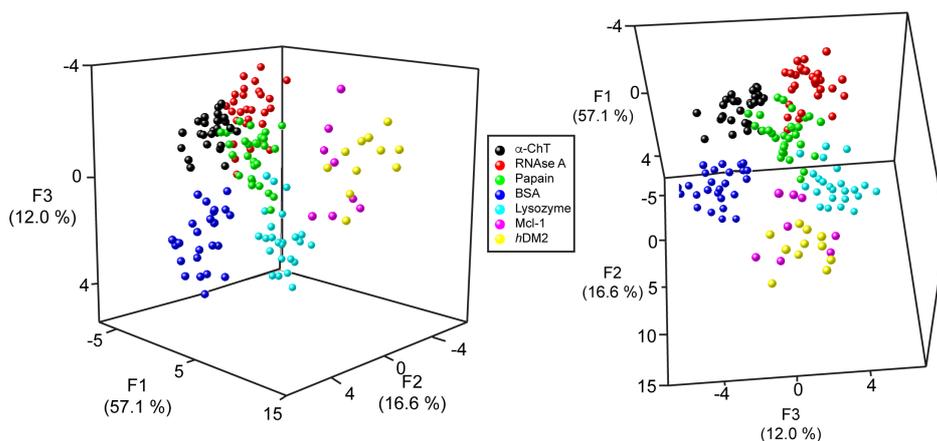


Figure 4.16 3-D LDA of 2 fluorophore array after removal of the *cyt c* data from the analysis

Looking at the 80 % and 95 % confidence ellipsoids (**Figure 4.17**) for these data shows ellipsoids for Mcl-1 and *hDM2* distinct from the other proteins, but overlapping each other. The ellipsoids for both of these proteins are quite large as many fewer replicates were performed with these data compared to with the other proteins (7 replicates for Mcl-1 and 9 for *hDM2*, compared to ~30 for the other proteins). More replicates for both of these proteins may distinguish these two proteins more readily. The other protein confidence ellipsoids are much smaller but show some overlap at both the 80 % and 95 % confidence levels, showing that while the clusters are distinct from each other, complete discrimination of the proteins is not achieved. However the small amounts of overlap

between the different protein ellipsoids mean that if the luminescence data were to be obtained for an unidentified protein, it would likely be attributed to the correct protein.

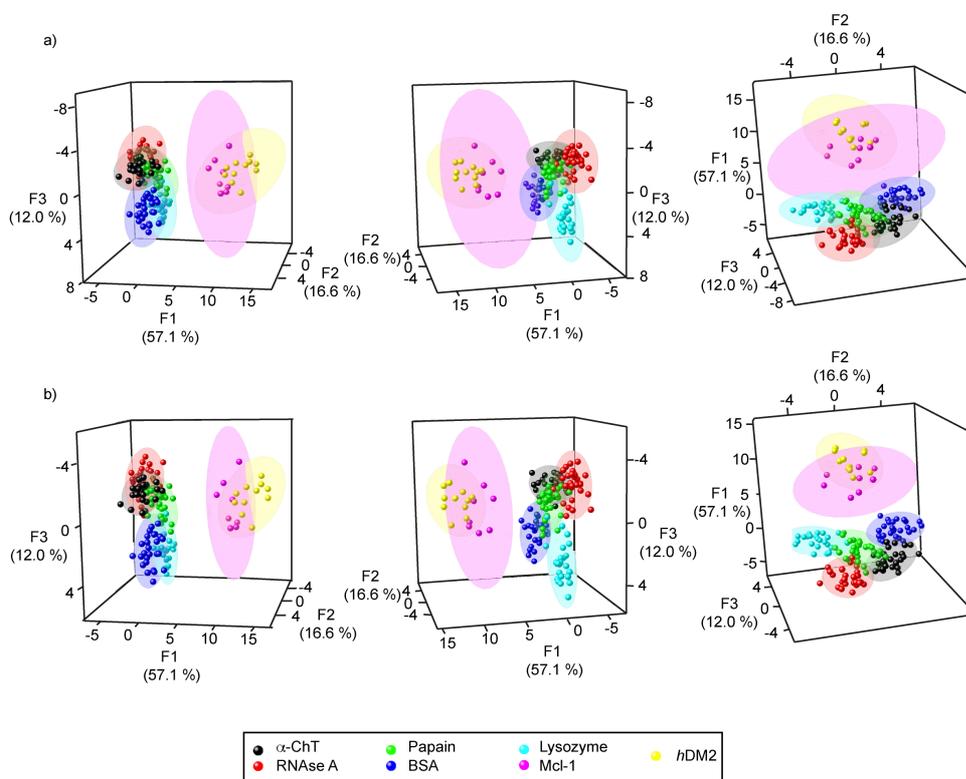


Figure 4.17 Confidence ellipsoids for the LDA of 2 fluorophore-protein array, excluding the *cyt c* data a) 95 % confidence ellipsoids, b) 80 % confidence ellipsoids

4.4 Conclusions

Attempts were made to discriminate various different proteins using the luminescence responses of 8 different Ru(II)(bpy)₃ complexes on incubation with the proteins, giving a fingerprint-like response for the different proteins. However, on performing a LDA on the data, yeast *cyt c* and horse heart *cyt c* could readily be discriminated from each other and the other proteins, but all the other proteins clustered within the same region, and could not be discriminated from each other, even in 3-D. To increase the potential for discrimination the arrays were incubated with a second fluorophore, a FITC-NOXA B peptide. This generated a more complicated fingerprint response and allowed for the discrimination of 6 different commercially available proteins, when using 3-D LDA. The addition of the more therapeutically interesting proteins, Mcl-1 and hDM2 decreased this discrimination, however if the *cyt c* data, which seemed to dominate the LDA separation in all cases, was removed some discrimination could again be achieved. Therefore, a platform has been generated for the discrimination of different proteins using two fluorophores,

which does not necessarily require the lengthy synthesis of molecules akin to those used by the Margulies group.^{193,197-199}

5 Thesis summary and future work

This thesis has focussed on the development of multivalent scaffolds (namely Ru(II)(bpy)₃ complexes and tetraphenyl-porphyrins) for their use as protein surface mimetics. The focus has centred on understanding how they interact with proteins, designing new platforms for the development of high affinity ligands and using them for new applications in protein detection and discrimination.

In Chapter 2, the utility of the Ru(II)(bpy)₃ scaffold to effectively mimic a native protein-protein interaction (PPI) was established. One Ru(II)(bpy)₃ complex was shown to effectively mimic the cytochrome (cyt) *c*/cyt *c* peroxidase PPI, binding to cyt *c* with the same binding site, similar thermodynamic and electrostatic binding profile as cyt *c* peroxidase. A larger, more functionalised Ru(II)(bpy)₃ complex was shown to enhance the interactions of this native PPI, with increased electrostatic interactions with cyt *c*, leading to an increase in the enthalpic contribution to binding. This showed how it is possible to enhance the interactions of a native PPI using synthetic molecules, in order to generate high affinity ligands for protein surfaces. Using this knowledge the rational design of new protein surface mimetics, based on both the Ru(II)(bpy)₃ scaffold and other multivalent scaffolds for binding to new, more therapeutically interesting protein targets could be established. The new synthetic methodology for the synthesis of these Ru(II)(bpy)₃ complexes, presented in Chapter 3, could aid in this endeavour, allowing for a much more expedient synthesis of a wide range of Ru(II)(bpy)₃ complexes for testing against different proteins.

Hydrazone-based dynamic combinatorial chemistry around three porphyrin scaffolds was presented in Chapter 3, showing that it is possible to generate a dynamic equilibrium in biologically relevant media. This system is prime for incubation with protein and for the generation of high affinity protein ligands. A proof-of-principle incubation with cyt *c* did not prove fruitful as the protein could not be separated from the (quenched) porphyrin dynamic combinatorial library, therefore a method for separation of the porphyrin DCL and protein needs to be established. This will firstly show if it is indeed possible to generate ligands for the model protein cyt *c*, and then allow the incubation with other, harder to target, proteins to see if it is possible to generate high affinity ligands for other protein surfaces.

In Chapter 4, an array approach for the discrimination of different proteins by incubation with two fluorophores (Ru(II)(bpy)₃ complexes and a FITC-labelled NOXA B

peptide) was presented. This allowed the discrimination of a range of different commercially available proteins, and showed some discrimination with the more therapeutically interesting proteins, Mcl-1 and hDM2. This array could be expanded to include additional therapeutically interesting proteins, to see if it is possible to discriminate a range of proteins of therapeutic interest. The use of further FITC labelled peptides, or the addition of a third fluorophore may aid in this, and lead to a tool for quick discrimination and later identification of a wide range of proteins.

Another potential platform for the discrimination of different proteins is to combine the ideas presented in Chapter 3 and 4, using the composition of a dynamic combinatorial library generated in the presence of different proteins, to discriminate between the proteins. Initial experiments in this vain are presented in Appendix IV, using a hydrazide functionalised Ru(II)(bpy)₃ complex in the presence of a range of aldehydes to discriminate between three different proteins. This approach requires more repeats to establish if it is a method capable of discriminating the different proteins. Subsequently the discrimination of a wider range of proteins will need to be established, and potential investigations into the use of different dynamic combinatorial chemistry scaffolds for this application, for example the porphyrin scaffolds presented in Chapter 3.

An additional avenue for future study with the Ru(II)(bpy)₃ complexes is to use them as PPI stabilisers. Initial results (presented in Appendix V) suggest that two of the synthesised Ru(II)(bpy)₃ complexes can stabilise the quaternary structure of a mutant variant (R337H) of the p53 tetramerisation domain. Further work on this showing that the p53 tetramerisation domain is indeed stabilised, for example by proteolysis experiments and showing binding affinity and binding site will show if these Ru(II)(bpy)₃ complexes can be used as PPI stabilisers as well as PPI inhibitors.

6 Experimental section

6.1 Synthesis

6.1.1 General considerations

Reagents and solvents were purchased from major suppliers and used without further purification. Anhydrous chloroform, dichloromethane, and methanol were obtained from the in-house solvent purification system, from Innovative Technology Inc. PureSolv®, other solvents used were of HPLC grade. Water for aqueous solutions was deionised.

Thin layer (silica) chromatography was performed using Merck Kiesegel 60 F₂₅₄ 0.25 mm precoated aluminium plates. Product spots were visualised by colour and under UV light (254 nm and 365 nm). Flash column chromatography was performed using silica gel 60 (0.043 – 0.063 mm VWR or Sigma Aldrich) or alumina (Brockman I from Sigma Aldrich), unless otherwise stated silica gel was used and pressure was applied by means of head bellows.

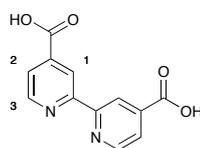
¹H NMR spectra were obtained on Bruker DPX 300 (300 MHz), Avance 500 (500 MHz) or DRX500 (500 MHz) spectrometers and referenced to either residual non-deuterated solvent peaks or tetramethylsilane. ¹³C spectra were recorded on a Bruker DPX 300 (75 MHz) Bruker or an Avance 500 (126 MHz) and referenced to the solvent peak. ¹H spectra are reported as follows: ¹H NMR (spectrometer frequency, solvent) δ ppm to 2 d.p. (multiplicity, *J* coupling constant in Hertz, number of protons, assignments). Chemical shifts (δ) are quoted in ppm with signal splitting recorded as singlet (s), doublet (d), triplet (t), quartet (q), quintet (qu.) multiplet (m) and broad (br.). Coupling constants (*J*) are measured to the nearest 0.1 Hz. Similarly, ¹³C spectra are reported as follows: δ (spectrometer frequency, solvent) δ ppm to one decimal place. Assignments of spectra were assisted by the results of DEPT, COSY, HMQC and HMBC experiments. ¹³C NMR spectra were obtained for all novel, and most literature, small molecules and ligands. Where possible ¹³C spectra were obtained for porphyrin molecules, however due to the large molecular size and thus broadness of peaks in the ¹H NMR, ¹³C NMR spectra were not able to be obtained for the Ru(II)(bpy)₃ complexes and some porphyrins. For the larger protected Ru(II)(bpy)₃ complexes it was found necessary to use DMSO-d₆ for discernment of peaks in the ¹H NMR spectra, despite solubility in CDCl₃ and MeOD.

Infrared spectra were recorded on a Perkin Elmer Fourier-Transfer spectrometer. Spectra were analysed neat and structurally important absorptions are quoted. Absorption maxima (ν_{\max}) are quoted in wavenumbers (cm^{-1}).

HPLC LCMS were recorded on a Bruker HCT ultra under electrospray ionisation (ESI) conditions. High resolution mass spectra were recorded on a Bruker Daltonics microTOF Premier Mass Spectrometer, under positive ESI conditions unless otherwise stated.

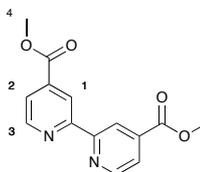
6.1.2 Synthetic protocols

2,2'-Bipyridine-4,4'-dicarboxylic acid, **40**^{142,200,201}



Potassium dichromate (9.10 g, 30.9 mmol) was added to a stirred solution of 4,4'-dimethyl-2,2'-bipyridine, **39** (2.00 g, 10.9 mmol) in concentrated sulfuric acid (50 mL) at 70°C over 2 hours, keeping the temperature between 70 and 80 °C. After the addition, the hot solution was poured onto ice (200 g) and the ice allowed to melt. The off-white precipitate formed was isolated by vacuum filtration. The solid was redissolved in 50 % nitric acid (50 mL) and the solution heated under reflux for 2 hours. The resulting mixture was cooled and poured onto ice (200 g) and the white precipitate isolated by vacuum filtration to yield the product as a white solid (2.23 g, 9.13 mmol, 84 %); ¹H NMR (300 MHz, TFA-*d*_x) δ ppm 8.36 (d, *J* = 5.5 Hz, 2 H, H3), 8.93 (d, *J* = 5.5 Hz, 2 H, H2), 9.03 (s, 2 H, H1), 11.42 (s, 2 H, COOH); IR (solid state, cm^{-1}) 2435, 1707; ESI-MS *m/z* found 245.0558 [M+H]⁺, [C₁₂H₉N₂O₄]⁺ requires 245.0557

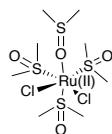
4,4'-Dimethylester-2,2'-bipyridine, **47**¹⁴²



Thionyl chloride (20 mL) was added dropwise to a stirred solution of 2,2'-bipyridine-4,4'-dicarboxylic acid, **40** (3.00 g, 12.3 mmol) in anhydrous methanol (200 mL) under a nitrogen atmosphere and the resulting solution heated under reflux for 20 hours. The reaction was then quenched with saturated sodium hydrogen carbonate solution (100

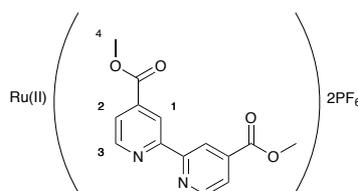
mL), and the mixture extracted with chloroform (3 × 100 mL). The combined organic phases were dried (sodium sulfate), and the solvent removed to yield the crude product as a pale orange solid, this was recrystallized in chloroform to yield the product as pale pink crystals (2.10 g, 7.72 mmol, 63 %); ¹H NMR (300 MHz, CDCl₃) δ ppm 4.02 (s, 6 H, H4), 7.93 (d, *J* = 4.9 Hz 2 H, H3), 8.89 (d, *J* = 4.9 Hz, 2 H, H2), 8.98 (s, 2 H, H1); ¹³C NMR (126 MHz, CDCl₃) δ ppm 52.8, 120.6, 123.3, 138.6, 150.2, 156.5, 165.6; IR (solid state, cm⁻¹) 1726 (C=O ester); ESI-MS *m/z* found 273.0870 [M+H]⁺, [C₁₄H₁₃N₂O₄]⁺ requires 273.0870

Ru(II)(DMSO)₄Cl₂¹⁵⁶



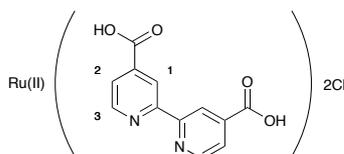
Ru(III)Cl₃.xH₂O (1.00 g) in anhydrous dimethylsulfoxide (5 mL) was degassed for 30 minutes. The solution was then heated under reflux under a nitrogen atmosphere for 30 minutes. Half of the solvent was then boiled off and the mixture cooled to room temperature. The reaction mixture was then diluted with acetone (10 mL) and the yellow precipitate isolated. To yield the product as a yellow solid (1.89 g, 3.91 mmol); ESI-MS *m/z* found 501.9317 [M+NH₄]⁺, [C₈H₂₈NO₄S₄Ru]⁺ requires 501.9322. (Poor solubility of this compound in all but DMSO meant NMR spectra could not be obtained)

Tris (4,4'-dimethylester-2,2'-bipyridine) ruthenium(II) dihexafluorophosphate, **49**¹⁴²



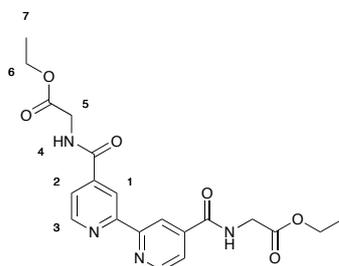
4,4'-Dimethylester-2,2'-bipyridine, **40** (382 mg, 1.40 mmol), Ru(II)(DMSO)₄Cl₂ (150 mg, 0.310 mmol), silver nitrate (105 mg, 0.620 mmol) and ethanol (20 mL) were heated under reflux for 7 days. The red solution was filtered and the filtrate concentrated. The red solid was redissolved in water (15 mL) and an excess of ammonium hexafluorophosphate added. The resultant red precipitate was isolated (203 mg, 0.168 mmol, 54 %); ¹H NMR (300 MHz, Acetone) δ ppm 2.88 (s, 18 H, H4), 7.98 (d, *J* = 5.9 Hz, 6 H, H2), 8.40 (d, *J* = 5.9 Hz, 6 H, H3), 9.38 (s, 6 H, H1); IR (solid state, cm⁻¹) 1726 (C=O ester); ESI-MS *m/z* found 459.0720 [M]²⁺, [C₄₂H₃₆N₆O₁₂Ru]²⁺ requires 459.0717

Tris (2,2'-bipyridine-4,4'-dicarboxylic acid) ruthenium(II) dichloride, 29¹⁴²



Tris (4,4'-dimethylester-2,2'-bipyridine) ruthenium(II) dihexafluorophosphate, **49** (1.00 g, 0.828 mmol) in ethanol (20 mL) and 1 M sodium hydroxide solution (20 mL) was stirred for 1 hour. The solution was neutralized with 1 M hydrochloric acid and concentrated. The salt was removed by dialysis (MWCO 0.1 - 0.5 kDa) against pure water to yield the product as a red solid (702 mg, 0.776 mmol, 94 %); ¹H NMR (300 MHz, D₂O) δ ppm 7.52 (d, *J* = 5.7 Hz, 6 H, H2), 7.73 (d, *J* = 5.7 Hz, 6 H, H3), 8.67 - 8.78 (br. s, 6 H, H1); IR (solid state, cm⁻¹) 3305 (O-H), 1600 (C=O acid); ESI-MS *m/z* found 417.0246 [M]²⁺, [C₃₆H₂₄N₆O₁₂Ru]²⁺ requires 417.0248

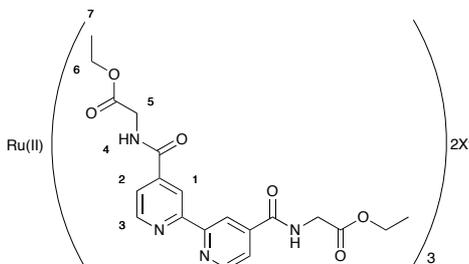
Ethyl 2-[(2-{4- [(2-ethoxy-2-oxoethyl) carbamoyl] pyridin-2-yl} pyridin-4-yl) formamido] acetate, 41f



2,2'-Bipyridine-4,4'-carboxylic acid, **40** (1.50 g, 6.14 mmol) and thionyl chloride (20 mL) were heated under reflux for 16 hours, the solvent was then removed *in vacuo*, and the dry acid chloride flushed with nitrogen and used immediately. To the dry acid chloride was added anhydrous chloroform (40 mL), ethyl glycine hydrochloride salt (1.89 g, 13.5 mmol) and triethylamine (1.88 mL, 13.5 mmol) and the reaction mixture heated under reflux, under a nitrogen atmosphere, for 16 hours. The reaction was then cooled and concentrated to yield the crude product as a pink solid, which was purified by flash column chromatography (5 % methanol in dichloromethane) to yield the product as a beige solid (1.52 g, 3.65 mmol, 60 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.37 (t, *J* = 7.2 Hz, 4 H, H7), 4.24 - 4.40 (m, 8 H, H5 and H6), 7.93 (d, *J* = 4.8 Hz, 2 H, H2), 8.82 (s, 2 H, H1), 8.91 (d, *J* = 4.8 Hz, 2 H, H3); ¹³C NMR (126 MHz, DMSO-*d*₆) δ ppm 14.1, 41.4, 60.6, 118.2, 121.9,

142.1, 150.2, 155.6, 165.1, 169.5; IR (solid state, cm^{-1}) 3303 (N-H), 1741 (C=O ester), 1648 (C=O amide); ESI-MS m/z found 415.16203 $[\text{M}+\text{H}]^+$, $[\text{C}_{20}\text{H}_{23}\text{N}_4\text{O}_6]^+$ requires 415.1612

Tris (ethyl 2-[(2-{4-[(2-ethoxy-2-oxoethyl)carbamoyl]pyridin-2-yl}pyridin-4-yl)formamido]acetate) ruthenium(II) diX, 42f



Method 1

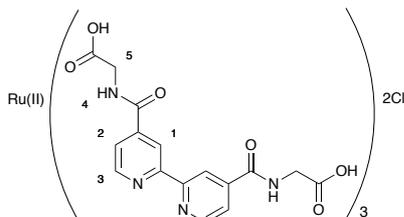
Ethyl 2- [(2-{4- [(2-ethoxy-2- oxoethyl) carbamoyl] pyridin-2-yl} pyridin-4-yl) formamido] acetate, **41f** (410 mg, 0.991 mmol), Ru(II)(DMSO)₄Cl₂ (150 mg, 0.310 mmol) and silver nitrate (105 mg, 0.620 mmol) in ethanol (20 mL) were heated under reflux for 7 days. The resulting solution was cooled to room temperature, filtered through celite, the and the celite pad washed thoroughly with dichloromethane until no more red compound could be seen. The red filtrate concentrated to yield the crude product as a red solid. This was purified by flash column chromatography (5 % methanol in dichloromethane) to yield the dinitrate salt as a red solid (426 mg, 0.290 mmol, 94 %).

Method 2

Tris (2,2'-bipyridine-4,4'-dicarboxylic acid) ruthenium(II) dichloride, **29** (50 mg, 0.055 mmol) in thionyl chloride (10 mL) and dimethylformamide (1 drop) was heated under reflux for 18 hours. The reaction mixture was then concentrated *in vacuo* and flushed with nitrogen to yield the acid chloride as a red solid. The acid chloride was resuspended in anhydrous chloroform (30 mL) and ethyl glycine hydrochloride salt (94 mg, 0.50 mmol) and diisopropylethylamine (0.17 mL, 0.99 mmol) were added. The reaction mixture was heated under reflux for 18 hours. The reaction mixture was then cooled to room temperature and washed with saturated sodium hydrogen carbonate solution (30 mL), 1 M hydrochloric acid (30 mL) and brine (30 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a red solid. This was purified by flash column chromatography (10 % methanol in dichloromethane) to yield the dichloride salt as a red solid (56 mg, 0.039 mmol, 72 %).

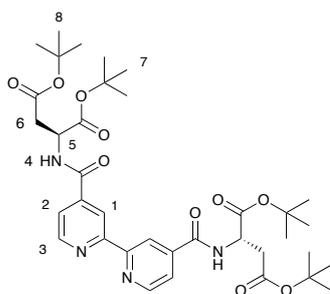
$^1\text{H NMR}$ (500 MHz, CDCl_3) δ ppm 1.16 (t, $J = 6.8$ Hz, 18 H, H7), 3.97 - 4.18 (m, 48 H, H5 + H6), 7.71 (br. s, 6 H, H2), 9.12 (s, 6 H, H1), 9.34 (br. s, 6 H, H3); IR (solid state, cm^{-1}) 3256, (N-H), 1734 (C=O ester), 1664 (C=O amide); ESI-MS m/z found 672.1840 $[\text{M}]^{2+}$, $[\text{C}_{60}\text{H}_{66}\text{N}_{12}\text{O}_{18}\text{Ru}]^{2+}$ requires 672.1834

Tris (2- [(2- {4-[(carboxymethyl) carbamoyl] pyridin-2-yl} pyridin-4-yl) formamido]acetic acid) ruthenium(II) dichloride, 30



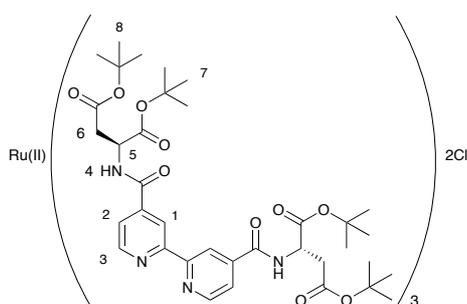
Tris (ethyl 2- [(2- {4- [(2- ethoxy-2- oxoethyl) carbamoyl] pyridin-2-yl} pyridin-4-yl) formamido]acetate) ruthenium(II) dinitrate, **42f** (200 mg, 0.136 mmol) was dissolved in ethanol (5 mL), and water (5 mL), 1 M sodium hydroxide solution (5 mL) was added, and the resulting mixture stirred for 18 hours. The reaction mixture was then neutralized with 1 M hydrochloric acid and concentrated to yield the product as a red solid in a mixture with sodium chloride. This mixture was redissolved in a minimum amount of water and dialysed (MWCO 0.1 – 0.5 kDa) against pure water to yield the product as a dark red solid (173 mg, 0.133 mmol, 98 %); $^1\text{H NMR}$ (500 MHz, D_2O) δ ppm 3.95 (s, 12 H, H5), 7.73 (dd, $J = 5.9, 1.6$ Hz, 6 H, H2), 7.94 (d, $J = 5.9$ Hz, 6 H, H1), 8.99 (app. s, 6 H, H3); IR (solid state, cm^{-1}) 3223 (N-H), 3251 (O-H), 1644 (C=O acid), 1585 (C=O amide); ESI-MS m/z found 588.0885 $[\text{M}]^{2+}$, $[\text{C}_{48}\text{H}_{42}\text{N}_{12}\text{O}_{18}\text{Ru}]^{2+}$ requires 588.0819

(2S,2'R)-tetra-*tert*-butyl-2,2'-([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanediyl) disuccinate, 41g



2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (100 mg, 0.400 mmol), triethylamine (1 drop) and thionyl chloride (10 mL) were heated under reflux for 16 hours. The mixture was cooled to room temperature and the solvent removed *in vacuo* to yield the acid chloride as an orange-red solid. The dry acid chloride was then redissolved in anhydrous chloroform (20 mL) and added dropwise to a stirred solution of di-*tert* butyl L-aspartic acid hydrogen chloride salt (253 mg, 0.901 mmol) and triethylamine (0.25 mL, 1.80 mmol) in anhydrous chloroform at 0 °C, under a nitrogen atmosphere. The reaction mixture was warmed to room temperature and heated under reflux for 48 hours. The mixture was cooled to room temperature and the solvent removed to yield the crude product as a brown oil. This was purified by flash column chromatography (3 % - 6 % methanol in chloroform) to yield the product as a yellow solid (262 mg, 0.375 mmol, 91 %); ¹H NMR (300 MHz, CDCl₃) δ ppm 1.49 (s, 18 H, H7/H8), 1.52 (s, 18 H, H7/H8), 2.91 (dd, *J* = 17.2, 4.3 Hz, 2 H, H6), 3.04 (m, *J* = 17.2, 4.3 Hz, 2H, H6'), 4.92 (dt, *J* = 7.5, 4.3 Hz, 2 H, H5), 7.46 (d, *J* = 7.5 Hz, 2 H, H4), 7.77 (dd, *J* = 5.0, 1.7 Hz, 2 H, H2), 8.78 (app. s, 2 H, H1), 8.83 (d, *J* = 5.0 Hz, 2 H, H3); ¹³C NMR (126 MHz, CDCl₃) δ ppm 28.0, 28.1, 37.5, 49.7, 81.9, 82.8, 118.0, 121.8, 142.3, 150.1, 156.3, 165.1, 169.5, 170.2; IR (solid state, cm⁻¹) 3346 (N-H), 1723 (C=O ester), 1650 (C=O amide); ESI-HRMS: found *m/z* 699.3615 [M+H]⁺, [C₃₆H₅₁N₄O₁₀]⁺ requires 699.3599

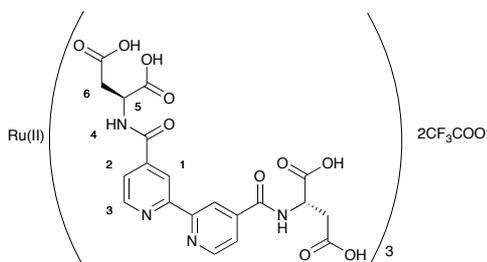
Tris ((2*S*,2'*R*)- tetra- *tert*-butyl- 2,2'- ([[2,2'-bipyridine]- 4,4'- dicarbonyl] bis (azanediyl)) disuccinate) ruthenium(II) dichloride, **42g**



(2*S*,2'*R*)-tetra-*tert*-butyl-2,2'-([[2,2'-bipyridine]-4,4'-dicarbonyl]bis(azanediyl)) disuccinate, **41g** (300 mg, 0.429 mmol), Ru(II)(DMSO)₄Cl₂ (65 mg, 0.134 mmol), silver nitrate (46 mg, 0.268 mmol) and ethanol (20 mL) were heated under reflux for 7 days. After which time the reaction mixture was filtered hot and concentrated. The red solid was then dissolved in a minimum amount of ethanol and loaded onto an SP Sephadex column and eluted with 1:1 acetone: 0.1 M sodium chloride solution and all the red fractions collected and concentrated. The combined red fractions were redissolved in acetone and filtered to remove sodium chloride, and this was repeated until no more white salt was visible in the concentrated sample. The Ru(II)(bpy)₃ complex was then purified by flash

chromatography (1 – 3 % methanol in chloroform) and the red fractions collected. These were concentrated, redissolved in chloroform and extracted into water, until the organic phase was no longer red. The combined aqueous phases were concentrated to yield the product as a red solid (77 mg, 0.034 mmol, 25 %); ¹H NMR (300 MHz, Acetone) δ ppm 2.10 (app. s, 108 H, H7 + H8), 2.81 - 3.07 (m, 12 H, H6), 4.79 - 5.07 (m, 6 H, H5), 7.89 (br. m, 6 H, H3), 8.37 (br. m, 6 H, H2), 8.82 - 8.98 (br. m, 6 H, H1); IR (solid state, cm⁻¹) 3310 (N-H), 1725 (C=O ester), 1710 (C=O ester), 1673 (C=O amide); ESI-HRMS: found *m/z* 1098.4854 [M]²⁺, [C₁₀₈H₁₅₀N₁₂O₃₀Ru]²⁺ requires 1098.4822; λ_{max} (MeOH): 306 nM (ε/ dm³ mol⁻¹ cm⁻¹ 240 723 981)

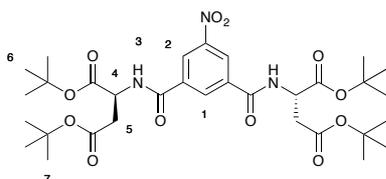
Tris ((2S,2'R)-2,2'-([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azanediy))disuccinic acid) ruthenium(II) ditrifluoroacetate, 31



Tris ((2S,2'R)-tetra-*tert*-butyl-2,2'-([2,2'-bipyridine]-4,4'-dicarbonyl)bis(azane-diyl))disuccinate) ruthenium(II) dichloride, **42g** (68 mg, 0.030 mmol), trifluoroacetic acid (4.5 mL) and water (0.5 mL) were stirred for 3 days. The reaction mixture was then concentrated *in vacuo* to yield the product as a red-black solid (57 mg, 0.029 mmol, 98 %); ¹H NMR (500 MHz, D₂O) δ ppm 2.67 (br. s, 6 H, H6), 2.78 (br. s, 6 H, H6'), 4.61 (br. s, 6 H, H5), 7.70 (br. s, 6 H, H3), 7.90 (br. s, 6 H, H2), 8.97 (br. s, 6 H, H1); IR (solid state, cm⁻¹) 3181 (O-H), 1648 (C=O acid, amide); ESI-HRMS: found *m/z* 762.1081 [M]²⁺, [C₆₀H₅₄N₁₂O₃₀Ru]²⁺ requires 762.1056

(2S,2'S)-tetra-*tert*-butyl 2,2'-((5-nitroisophthaloyl)bis(azanediy))disuccinate,

37b¹⁴³

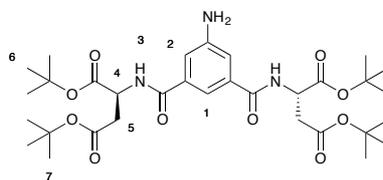


5-nitroisophthalic acid (1.00 g, 4.74 mmol) and dimethylformamide (1 drop) in thionyl chloride (5 mL) were heated under reflux for 4 hours. The reaction mixture was

concentrated *in vacuo* to yield the acid chloride as a white solid. The acid chloride, under a nitrogen atmosphere, was redissolved in anhydrous dichloromethane (50 mL) and *tert*-butyl L-aspartic acid.Hydrogen chloride salt (2.94 g, 10.4 mmol) and triethylamine (1.45 mL, 10.4 mmol) were added. The reaction mixture was stirred at room temperature, under a nitrogen atmosphere, for 16 hours. The resulting solution was then washed with 1 M hydrochloric acid (50 mL), saturated sodium hydrogen carbonate solution (50 mL) and brine (50 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a beige solid, this was purified by flash column chromatography (3:7 ethyl acetate:dichloromethane) to yield the product as a white solid (1.50 g, 2.25 mmol, 48 %); ¹H NMR (300 MHz, CDCl₃) δ ppm 1.38 (s, 18 H, H6/H7), 1.42 (s, 18 H, H6/H7), 2.78 (dd, *J* = 17.2, 4.3 Hz, 2 H, H4), 2.94 (dd, *J* = 17.2, 3.9 Hz, 2 H, H4'), 4.70 - 4.92 (m, 2 H, H5), 7.35 (m, 2 H, H3), 8.52 (app. s, 1 H, H1), 8.70 (d, *J* = 1.5 Hz, 2 H, H2); ¹³C NMR (126 MHz, CDCl₃) δ ppm 27.9, 28.1, 37.3, 50.0, 82.0, 83.0, 124.7, 131.4, 136.4, 148.5, 163.8, 169.4, 170.3; IR (solid state, cm⁻¹) 3371 (N-H), 1734 (C=O ester), 1656 (C=O amide); ESI-MS *m/z* found 688.3052 [M+Na]⁺, [C₃₂H₄₈N₃O₁₂Na]⁺ requires 688.5052

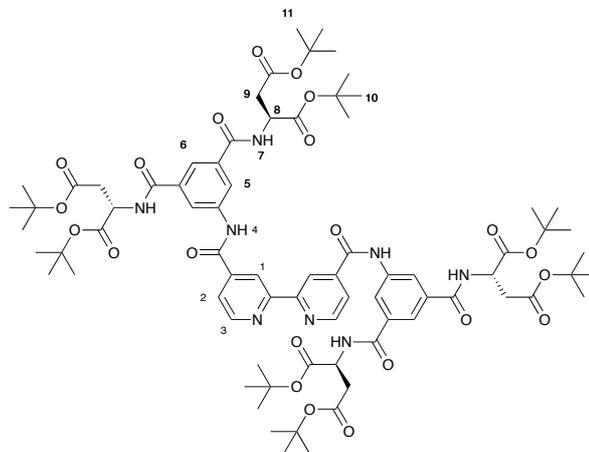
(2*S*,2'*S*)-tetra-*tert*-butyl 2,2'-((5-aminoisophthaloyl)*bis*(azanediyl))disuccinate,

38b¹⁴³



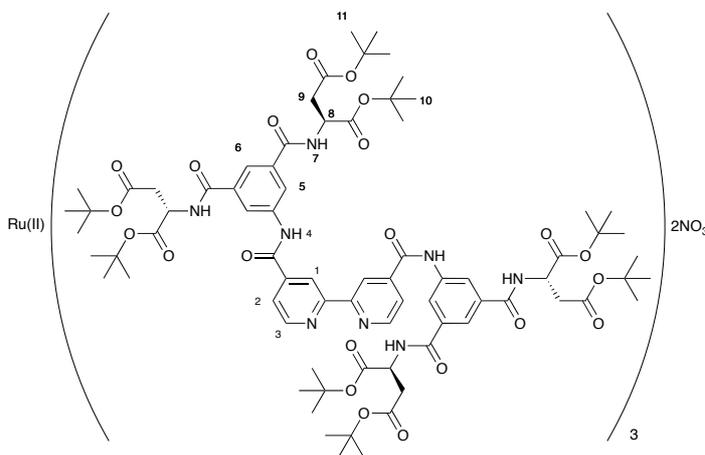
(2*S*,2'*S*)-tetra-*tert*-butyl 2,2'-((5-nitroisophthaloyl)*bis*(azanediyl))disuccinate, **37b** (1.00 g, 1.65 mmol) was dissolved in methanol (20 mL) and ethyl acetate (20 mL). The reaction mixture was degassed and palladium on activated charcoal (11 mg) was added and the solution degassed again. The solution was then put under a hydrogen atmosphere and stirred for 16 hours. The solution was then filtered twice and concentrated to yield the product as a flocculent cream solid (981 mg, 1.54 mmol, 94 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.48 (s, 18 H, H6/H7), 1.53 (s, 18 H, H6/H7), 2.88 (dd, *J* = 17.0, 4.4 Hz, 1 H, H5), 2.99 (dd, *J* = 17.0, 4.5 Hz, 1 H, H5'), 4.83 - 4.99 (m, 1 H, H4), 7.33 (m, 2 H, H3), 7.54 (s, 2 H, H2), 7.77 (s, 1 H, H1); ¹³C NMR (126 MHz, CDCl₃) δ ppm 27.9, 28.0, 37.5, 49.7, 81.6, 82.4, 114.9, 116.6, 135.5, 147.3, 166.6, 169.9, 170.3; IR (solid state, cm⁻¹) 3341 (N-H), 1725 (C=O ester), 1671 (C=O ester), 1648 (C=O amide); ESI-MS *m/z* found 636.3500 [M+H]⁺, [C₃₂H₅₀N₃O₁₀]⁺ requires 636.3491

1,4-Di-*tert*-butyl (2*S*)-2-[(3-[[[(2*S*)-1,4-bis(*tert*-butoxy)-1,4-dioxobutan-2-yl] carbamoyl] -5- [2- (4- [[3,5-bis ([[[(2*S*)-1,4-bis (*tert*-butoxy)-1,4- dioxobutan-2-yl] carbamoyl]) phenyl] carbamoyl] pyridin-2-yl) pyridine-4-amido] phenyl) formamido]butanedioate, 41b



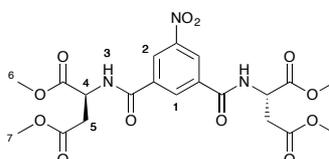
2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (128 mg, 0.524 mmol) was heated under reflux in thionyl chloride (10 mL) for 18 hours, and the solvent removed *in vacuo*. The resultant acid chloride was flushed with nitrogen and used immediately. The acid chloride was redissolved in anhydrous chloroform (15 mL) and heated to reflux, under a nitrogen atmosphere. A solution of (2*S*,2'*S*)-tetra-*tert*-butyl 2,2'-((5-aminoisophthaloyl) bis(azanediyl)) disuccinate (600 mg, 0.943 mmol), **38b** and anhydrous triethylamine (0.16 mL, 1.15 mmol) in anhydrous chloroform (35 mL), under a nitrogen atmosphere, was added dropwise to the refluxing solution and the resulting mixture heated under reflux for 18 hours. The solvent was then removed to yield the crude product as a pink solid, which was purified by flash column chromatography (1 – 5 % methanol in dichloromethane) to yield the product as a beige solid (110 mg, 0.0743 mmol, 14 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 1.42 (s, 36 H, H10/H11), 1.44 (s, 36 H, H10/H11), 2.72 (dd, *J* = 16.2, 7.3 Hz, 4 H, H9), 2.85 (dd, *J* = 16.2, 7.3 Hz, 4 H, H9'), 4.77 (q, *J* = 7.3 Hz, 4 H, H8), 8.05 (d, *J* = 6.5 Hz, 2 H, H3), 8.08 (s, 2 H, H6) 8.44 (s, 4 H, H5), 8.90 (d, *J* = 6.5 Hz, 2 H, H2), 9.01 (s, 2 H, H1), 11.07 (app. s, 2 H, H4); ¹³C NMR (126 MHz, DMSO-*d*₆) δ ppm 27.5, 27.7, 37.0, 78.7, 81.1, 104.5, 118.6, 121.9, 122.4, 122.6, 135.0, 138.7, 142.9, 150.3, 155.5, 164.1, 165.8, 169.2, 169.7; IR (solid state, cm⁻¹) 3343 (N-H), 1726 (C=O ester), 1671 (C=O ester), 1656 (C=O amide); ESI-MS *m/z* found 1479.7185 [M+H]⁺, [C₇₆H₁₀₃N₈O₂₂]⁺ requires 1479.7181

Tris (1,4-di-*tert*-butyl(2*S*)-2-[(3-[[[(2*S*)-1,4-bis(*tert*-butoxy)-1,4-dioxobutan-2-yl] carbamoyl] -5-[2- (4- {3,5-bis ([[[(2*S*)-1,4-bis(*tert*-butoxy)-1,4-dioxobutan-2-yl] carbamoyl]) phenyl] carbamoyl] pyridin-2-yl) pyridine-4-amido]phenyl) formamido] butanedioate) ruthenium(II) dinitrate, 42b



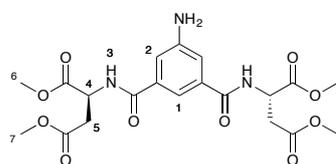
1,4-Di-*tert*- butyl (2*S*)-2- [(3- [[[(2*S*)- 1,4- bis (*tert*- butoxy)- 1,4- dioxobutan-2-yl] carbamoyl]-5-[2-(4-[[3,5-bis([[[(2*S*)-1,4-bis(*tert*-butoxy)-1,4-dioxobutan-2-yl] carbamoyl]) phenyl] carbamoyl] pyridin-2-yl)pyridine-4-amido]phenyl)formamido]butanedioate, **41b** (400 mg, 0.270 mmol), Ru(II)(DMSO)₄Cl₂ (42 mg, 0.0872 mmol), and silver nitrate (92 mg, 0.540 mmol) in ethanol (30 mL) were heated under reflux for 6 days. The solution was then filtered through celite and the filtrate concentrated. The resulting red solid was purified by flash column chromatography (5 % methanol in dichloromethane) to yield the product as a red solid (50 mg, 0.0107 mmol, 12 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 1.39 (br. s, 108 H, H10/H11), 1.40 - 1.52 (br. s, 108 H, H10/H11), 2.63 - 2.76 (br. m, 12 H, H9), 2.77 - 2.92 (br. m, 12 H, H9'), 4.76 (q, *J* = 7.7 Hz, 12 H, H8), 8.03 - 8.15 (m, 24 H, H7 + H2 + H3), 8.35 - 8.53 (m, 18 H, H5, H6), 8.76 - 8.94 (m, 12 H, H1 + H4); IR (solid state, cm⁻¹) 3274 (N-H), 1729 (C=O ester), 1666 (C=O amide); ESI-MS *m/z* found 2270.5083 [M]²⁺, [C₂₂₈H₃₀₆N₂₄O₆₆Ru]²⁺ requires 2269.0185

1,4-Dimethyl (2*S*)-2-[(3-[[[(2*S*)-1,4-dimethoxy-1,4-dioxobutan-2-yl] carbamoyl]-5-nitrophenyl)formamido]butanedioate, 37a¹⁴²



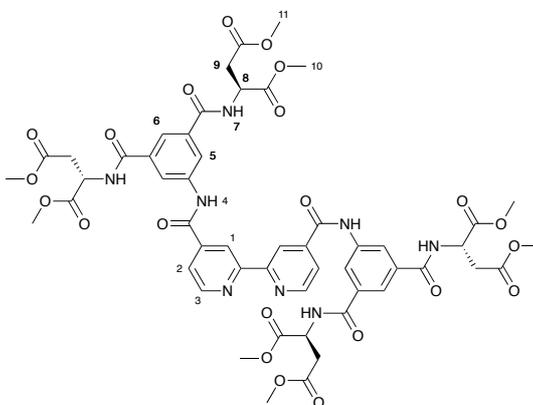
5-nitroisophthalic acid (2.00 g, 9.47 mmol) and dimethylformamide (1 drop) in thionyl chloride (20 mL) were heated under reflux for 18 hours. The solvent was removed *in vacuo* to yield the acid chloride as a white solid. The acid chloride was redissolved in anhydrous dichloromethane (40 mL), dimethyl L-aspartic acid hydrochloride salt (4.12 g, 20.8 mmol) and triethylamine (2.92 mL, 20.8 mmol) were added and the reaction mixture stirred for 16 hours. The resulting solution was quenched with saturated sodium hydrogen carbonate solution (50 mL). The organic phase was then washed with 1 M hydrochloric acid (50 mL) and brine (50 mL), dried (sodium sulfate) and concentrated to yield the crude product as a beige solid, this was purified by flash column chromatography (3:7 ethyl acetate:dichloromethane) to yield the product as a white solid (2.54 g, 5.11 mmol, 54 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 3.03 (dd, *J* = 17.4, 6.1 Hz, 2 H, H5), 3.18 (dd, *J* = 17.4, 4.8 Hz, 2 H, H5'), 3.74 (s, 6 H, H6/H7), 3.84 (s, 6 H, H6/H7), 5.06 - 5.18 (m, 2 H, H4), 7.65 (d, *J* = 7.9 Hz, 3 H, H3), 8.60 (t, *J* = 1.5 Hz, 1 H, H1), 8.79 (d, *J* = 1.5 Hz, 2 H, H2); ¹³C NMR (126 MHz, CDCl₃) δ ppm 35.9, 49.4, 52.3, 53.1, 125.2, 131.4, 135.8, 148.5, 164.0, 171.0, 171.5; IR (solid state, cm⁻¹) 3386 (N-H), 1747 (C=O ester), 1727 (C=O ester), 1672 (C=O amide); ESI-MS *m/z* found 498.1368 [M+H]⁺, [C₂₀H₂₄N₃O₁₂]⁺ requires 498.1354

1, 4- dimethyl (2S)-2- [(3- amino-5-[(2S)- 1,4-dimethoxy- 1,4- dioxobutan-2-yl]carbamoyl]phenyl)formamido]butanedioate, 38a¹⁴²



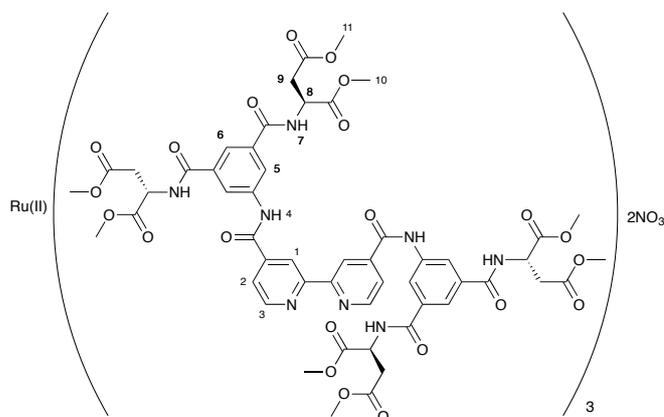
1,4-dimethyl (2S)-2-[(3-[(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl] carbamoyl]-5-nitrophenyl)formamido]butanedioate, **37a** (1.00 g, 2.14 mmol) in methanol (30 mL) and ethyl acetate (30 mL) was degassed and palladium on activated charcoal (20 mg) added. The solution was degassed again and put under a hydrogen atmosphere. The mixture was stirred for 18 hours, then filtered twice and concentrated to yield the product as a cream solid (1.07 g, 2.29 mmol, quant.); ¹H NMR (500 MHz, CDCl₃) δ ppm 3.00 (dd, *J* = 17.4, 4.9 Hz, 2 H, H5), 3.10 (dd, *J* = 17.4, 6.5 Hz, 2 H, H5'), 3.71 (s, 6 H, H6/H7), 3.80 (s, 6 H, H6/H7), 5.06 (dt, *J* = 6.5, 4.9 Hz, 2 H, H4), 7.18 (s, 2 H, H2), 7.45 (s, 1 H, H1), 7.46 - 7.48 (m, 2 H, NH₂); ¹³C NMR (126 MHz, CDCl₃) δ ppm 36.1, 49.1, 52.2, 52.9, 115.0, 116.6, 135.1, 147.4, 166.7, 171.3, 171.6; IR (solid state, cm⁻¹) 3360 (N-H), 1727 (C=O ester), 1643 (C=O amide); ESI-MS *m/z* found 468.1624 [M+H]⁺, [C₂₀H₂₆N₃O₁₀]⁺ requires 468.1613

1,4-Dimethyl (2S)-2-({3-[2-(4-{{3,5- bis ({{(2S)-1,4- dimethoxy-1,4- dioxobutan-2-yl] carbamoyl}) phenyl] carbamoyl} pyridin-2-yl)pyridine-4-amido]-5-{{(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl]carbamoyl}phenyl}formamido)butanedioate **41a¹⁴²**



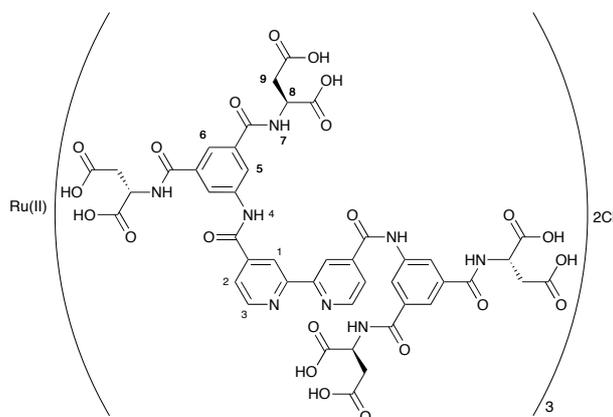
2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (471 mg, 1.93 mmol) in thionyl chloride (15 mL) was heated under reflux for 18 hours. The solvent was then removed *in vacuo*, and the resulting acid chloride flushed with nitrogen and used immediately. The acid chloride was redissolved in anhydrous chloroform (10 mL) and heated to reflux, under a nitrogen atmosphere. A solution of 1,4-dimethyl (2S)-2-[(3-amino-5-{{(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl]carbamoyl}phenyl}formamido]butanedioate, **38a** (1.62 g, 3.47 mmol) and anhydrous triethylamine (0.59 mL, 4.3 mmol) in anhydrous chloroform (20 mL), kept under a nitrogen atmosphere was added dropwise to the refluxing solution and the solution then heated under reflux for 16 hours. The reaction mixture was concentrated and the resulting pink crude product purified by flash column chromatography (5 % methanol in chloroform) to yield the product as a beige solid (1.41 g, 1.23 mmol, 64 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 2.90 (dd, *J* = 16.4, 8.0 Hz, 4 H, H₉), 3.01 (dd, *J* = 16.4, 8.0 Hz, 4 H, H_{9'}), 3.64 (s, 12 H, H₁₀/H₁₁), 3.67 (s, 12 H, H₁₀/H₁₁), 4.91 (app. q, *J* = 8.0 Hz, 4 H, H₈), 8.06 (dd, *J* = 5.1, 1.3 Hz, 2 H, H₂), 8.12 (app. s, 2 H, H₆), 8.46 (d, *J* = 1.3 Hz, 4 H, H₅), 8.98 – 9.04 (m, 4 H, H₁ + H₃), 9.15 (d, *J* = 7.6 Hz, 4 H, H₇), 11.09 (s, 2 H, H₄); ¹³C NMR (126 MHz, DMSO-*d*₆) δ ppm 35.3, 49.4, 51.7, 52.3, 118.6, 122.1, 122.4, 122.8, 134.6, 138.8, 142.9, 150.3, 155.6, 164.1, 165.8, 170.5, 171.1; IR (solid state, cm⁻¹) 3306 (N-H), 3011 (N-H), 1731 (C=O ester), 1651 (C=O amide), 1598 (C=O amide); ESI-MS *m/z* found 572.1751 [M+H]²⁺, [C₅₂H₅₆N₈O₂₂]²⁺ requires 572.1749

Tris (1,4- dimethyl (2S)-2-({3-[2-(4-{{3,5-bis({[(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl]carbamoyl}) phenyl] carbamoyl} pyridin-2-yl) pyridine-4-amido]-5- {{(2S)-1,4-dimethoxy-1,4- dioxobutan-2-yl] carbamoyl} phenyl} formamido) butanedioate) ruthenium(II) dinitrate, 42b¹⁴²



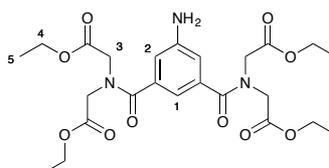
1,4-Dimethyl (2S)-2- ({3-[2-(4-{{3,5-bis ({{(2S)-1,4-dimethoxy -1,4- dioxobutan-2-yl] carbamoyl}) phenyl] carbamoyl} pyridin-2-yl) pyridine-4-amido]-5-{{(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl]carbamoyl}phenyl}formamido)butanedioate, **41a** (1.00 g, 0.875 mmol), Ru(II)(DMSO)₄Cl₂ (132 mg, 0.273 mmol) and silver nitrate (93 mg, 0.547 mmol) in ethanol (20 mL) were heated under reflux for 7 days. The reaction mixture was then filtered and the red filtrate concentrated to yield the crude product as a red solid. This was purified by flash column chromatography (5 % - 10 % ethanol in dichloromethane) to yield the product as a red solid (625 mg, 0.171 mmol, 63 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 2.89 (dd, *J* = 16.3, 6.7 Hz, 12 H, H9), 2.99 (dd, *J* = 16.3, 6.7 Hz, 12 H, H9'), 3.43 (br. s, 12 H, H7), 3.64 (s, 36 H, H10/H11), 3.67 (s, 36 H, H10/H11), 4.89 (q, *J* = 6.7 Hz, 7 H, H8), 8.06 - 8.16 (m, 6 H, H6), 8.18 (d, *J* = 6.5 Hz, 6 H, H2) 8.45 (s, 12 H, H5), 9.15 (d, *J* = 6.5 Hz, 6 H, H3), 9.58 (br. s, 6 H, H1), 11.23 (app. s, 6 H, H4); IR (solid state, cm⁻¹) 3293 (N-H), 2953 (N-H), 1731 (C=O ester), 1656 (C=O amide), 1599 (C=O amide); ESI-MS *m/z* found 1764.9596 [M+H]²⁺, [C₁₅₆H₁₆₂N₂₄O₆₆Ru]²⁺ requires 1764.4551

Tris ((2S)-2-({3-[2-(4-{{3,5-bis({[(1S)-1,2-dicarboxyethyl] carbamoyl}) phenyl] carbamoyl}pyridin-2-yl)pyridine-4-amido]-5-{{(1S)-1,2-dicarboxyethyl}carbamoyl} phenyl}formamido)butanedioic acid) ruthenium(II) dichloride, 26



Tris (1,4-dimethyl (2S)-2-({3-[2-(4-{{3,5-bis({[(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl]carbamoyl})phenyl]carbamoyl}pyridin-2-yl) pyridine-4-amido]-5-{{(2S)-1,4-dimethoxy-1,4-dioxobutan-2-yl] carbamoyl} phenyl} formamido) butanedioate) ruthenium(II) dinitrate, 42b (15 mg, 0.0041 mmol), lithium hydroxide (5 mg, 0.021 mmol), tetrahydrofuran (2 mL) and water (2 mL) were stirred for 1 hour. The solution was then neutralized by addition of 1 M hydrochloric acid. The red solution was concentrated, and dialysed (MWCO 0.1 – 0.5 kDa) against pure water, to yield the product as a red solid (11 mg, 0.0032 mmol, 82 %); ¹H NMR (500 MHz, D₂O) δ ppm 2.68 - 2.78 (m, 24 H, H9), 4.71 (m, 12 H, H8), 7.94 (d, *J* = 5.4 Hz, 6 H, H2), 7.99 - 8.04 (br. s, 6 H, H6), 8.09 (br. d, *J* = 5.4 Hz, 6 H, H3), 8.11 - 8.21 (br. s, 12 H, H5), 8.93 - 9.31 (br. s, 6 H, H1); IR (solid state, cm⁻¹) 3255 (N-H), 2549 (O-H acid), 1625 (C=O acid), 1601 (C=O amide); ESI-HRMS found *m/z* 1596.211 [M]²⁺, [C₁₃₂H₁₁₄N₂₄O₆₆Ru]²⁺ requires 1596.2672

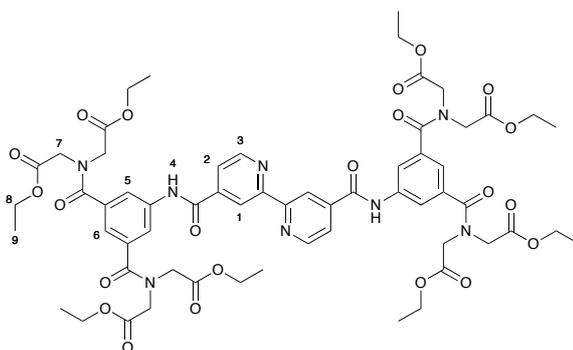
Ethyl 2-(1- {3-amino-5- [bis (2-ethoxy-2-oxoethyl) carbamoyl] phenyl}-N- (2-ethoxy-2-oxoethyl)formamido)acetate, 38c



Ethyl 2-(1- {3-[bis (2- ethoxy-2- oxoethyl) carbamoyl] -5-nitrophenyl} -N-(2- ethoxy-2-oxoethyl)formamido)acetate, **37b** (synthesised by Georgina Pleasance, 974 mg, 1.76 mmol) was dissolved in methanol (20 mL) and ethyl acetate (20 mL) and the solution

degassed. Palladium on activated charcoal (20 mg) was added and the solution degassed again. The solution was put under a hydrogen atmosphere and stirred for 18 hours. The reaction mixture was then filtered twice and concentrated to yield the product as a sticky pale yellow solid (823 mg, 1.57 mmol, 89 %); ^1H NMR (500 MHz, CDCl_3) δ ppm 1.23 (t, $J = 7.1$ Hz, 9 H, H5), 1.28 (t, $J = 7.1$ Hz, 9 H, H5'), 4.07 (s, 8 H, H3), 4.13 (q, $J = 7.1$ Hz, 4 H, H4), 4.22 (q, $J = 7.1$ Hz, 4 H, H4'), 6.75 (s, 1 H, H1), 6.79 (s, 2 H, H2); ^{13}C NMR (126 MHz, CDCl_3) δ ppm 14.1, 14.2, 47.4, 51.6, 61.4, 61.7, 113.8, 114.5, 136.4, 147.5, 168.8, 169.0, 171.5; IR (solid state, cm^{-1}) 3361 (N-H), 1734 (C=O ester), 1642 (C=O amide); ESI-MS m/z found 524.2214 $[\text{M}+\text{H}]^+$, $[\text{C}_{24}\text{H}_{34}\text{N}_3\text{O}_{10}]^+$ requires 524.2239

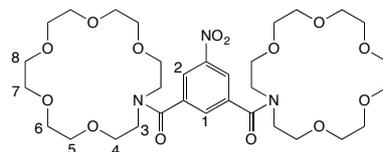
2-(1-{3-[Bis(2-ethoxy-2-oxoethyl)carbamoyl]-5-{2-[4-({3,5-bis[bis(2-ethoxy-2-oxoethyl)carbamoyl]phenyl}carbamoyl)pyridin-2-yl]pyridine-4-amido}phenyl}-N-(2-ethoxy-2-oxoethyl)formamido)-1-(ethylperoxy)ethylidene, 41c



2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (181 mg, 0.743 mmol) in thionyl chloride (10 mL) was heated under reflux for 18 hours. The solvent was then removed *in vacuo* and the yellow acid chloride flushed with nitrogen and used immediately. The acid chloride was redissolved in anhydrous chloroform (10 mL) and heated to reflux under a nitrogen atmosphere. A solution of ethyl 2-(1-{3-amino-5-[bis(2-ethoxy-2-oxoethyl) carbamoyl] phenyl} -N-(2-ethoxy-2-oxoethyl) formamido) acetate, **38c** (702 mg, 1.34 mmol) and anhydrous triethylamine (0.23 mL, 1.63 mmol) in anhydrous chloroform (15 mL) was added dropwise to the refluxing solution and the resulting solution heated under reflux for 24 hours. The reaction mixture was then concentrated to yield the product as a pale pink solid (1.05 g, 0.836 mmol, quant.). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ ppm 1.20 (t, $J = 7.2$ Hz, 12 H, H9), 1.22 (t, $J = 7.0$ Hz, 12 H, H9'), 4.08 (s, 8 H, H7), 4.14 (q, $J = 7.0$ Hz, 8 H, H8), 4.12 (q, $J = 7.2$ Hz, 8 H, H8'), 4.20 (s, 8 H, H7'), 6.24 (s, 4 H, H5), 6.55 (s, 2 H, H6), 7.89 (d, $J = 5.0$ Hz, 2 H, H2), 8.85 (s, 2 H, H1), 8.88 (d, $J = 5.0$ Hz, 2 H, H3); ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ ppm 13.9, 14.1, 47.7, 48.1, 60.5, 61.1, 112.4, 118.5, 119.5, 122.4, 135.9, 139.2, 142.8,

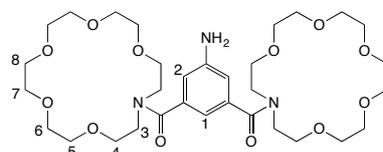
150.3, 155.5, 164.2, 168.6, 169.0, 170.1; IR (solid state, cm^{-1}) 1737 (C=O ester), 1648 (C=O amide); ESI-MS m/z found 1255.4672 $[\text{M}]^+$, $[\text{C}_{60}\text{H}_{71}\text{N}_8\text{O}_{22}]^+$ requires 1255.4682

(5-Nitro-1,3-phenylene)bis((1,4,7,10,13-pentaoxa-16-azacyclooctadecan-16-yl)methanone), 37d¹⁴²



5-nitroisophthalic acid (300 mg, 1.42 mmol), thionyl chloride (5 mL) and dimethylformamide (1 drop) were heated under reflux for 20 hours. The solvent was then removed *in vacuo* to yield the acid chloride as a pale yellow solid. The acid chloride was redissolved in anhydrous dichloromethane (15 mL) and 1-aza-18-crown-6 (823 mg, 3.12 mmol) and triethylamine (0.44 mL, 3.1 mmol) added and the resulting mixture stirred, under a nitrogen atmosphere, for 18 hours. The reaction mixture was then concentrated to yield the crude product as a brown solid. This was purified by flash column chromatography (100:40:8 chloroform:acetone:ethanol), to yield the product as a brown solid (760 mg, 1.09 mmol, 76 %); ^1H NMR (300 MHz, CDCl_3) δ ppm 3.49 - 3.74 (m, 48 H, H3 - H8), 7.87 (t, $J = 1.4$ Hz, 1 H, H1), 8.42 (d, $J = 1.4$ Hz, 2 H, H2); IR (solid state, cm^{-1}) 1632 (C=O amide), 1538 (NO_2), 1472 (NO_2); ESI-MS m/z found 702.3447 $[\text{M}+\text{H}]^+$, $[\text{C}_{32}\text{H}_{52}\text{N}_3\text{O}_{14}]^+$ requires 702.3444

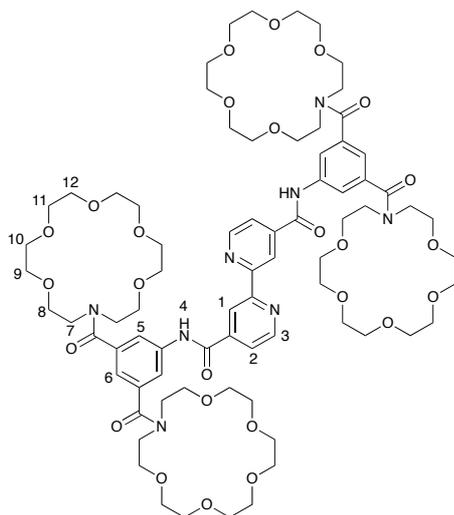
(5-Amino-1,3-phenylene)bis((1,4,7,10,13-pentaoxa-16-azacyclooctadecan-16-yl)methanone), 38d¹⁴²



(5-Nitro-1,3-phenylene)bis((1,4,7,10,13-pentaoxa-16-azacyclooctadecan-16-yl)methanone), **37d** (650 mg, 0.926 mmol) was dissolved in methanol (10 mL) and the solution degassed. Palladium on activated charcoal (67 mg) was added and the solution degassed again, and then placed under a hydrogen atmosphere and stirred for 16 hours. The solution was filtered twice and concentrated to yield the aniline as a yellow-orange oil (380 mg, 0.566 mmol, 61 %); ^1H NMR (300 MHz, CDCl_3) δ ppm 3.49 - 3.93 (m, 48 H, H3 -

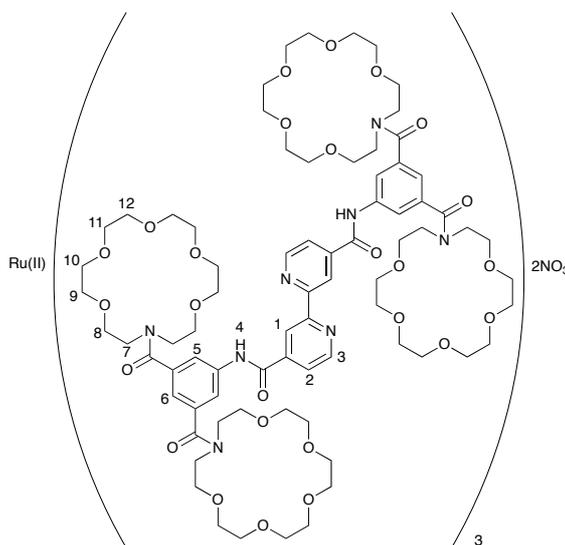
H8), 6.73 (s, 2 H, H2), 7.28 (s, 1 H, H1); IR (solid state, cm^{-1}) 3350 (N-H), 1622 (C=O amide); ESI-MS m/z found 694.3531 $[\text{M}+\text{Na}]^+$, $[\text{C}_{32}\text{H}_{53}\text{N}_3\text{O}_{12}\text{Na}]^+$ requires 694.3521

N4,N4'-bis(3,5-di(1,4,7,10,13-pentaoxa-16-azacyclooctadecane-16-carbonyl)phenyl-[2,2'-bipyridine]-4,4'-dicarboxamide, 41d¹⁴²



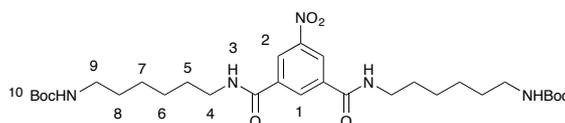
2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (50 mg, 0.205 mmol), thionyl chloride (3 mL, 15.4 mmol) and triethylamine (1 drop) were heated under reflux for 18 hours. The reaction mixture was concentrated *in vacuo* to yield the acid chloride as an orange solid. The acid chloride was flushed with nitrogen and used immediately. The acid chloride was redissolved in anhydrous chloroform (5 mL) and added dropwise to a stirred solution of (5-amino-1,3-phenylene)*bis*((1,4,7,10,13-pentaoxa-16-azacyclooctadecan-16-yl)methanone, **38d** (303 mg, 0.451 mmol) in anhydrous chloroform (5 mL) at 0 °C under a nitrogen atmosphere. The solution was then heated under reflux for 48 hours. The reaction mixture was concentrated to yield the crude product as a red-brown solid, which was purified by flash alumina column chromatography (100:40:8 chloroform:acetone:ethanol) to yield the product as a brown solid (278 mg, 0.179 mmol, 87 %); ^1H NMR (300 MHz, CDCl_3) δ ppm 3.40 - 3.62 (m, 80 H, H8 - H12), 3.63 (br. s, 16 H, H7), 7.25 (s, 2 H, H6), 7.89 (d, $J = 4.5$ Hz, 2 H, H3), 7.97 (br. s, 4 H, H5), 8.81 (d, $J = 4.5$ Hz, 2 H, H2), 8.93 (s, 2 H, H1); IR (solid state, cm^{-1}) 3490 (N-H), 1678 (C=O amide), 1623 (C=O amide)

Tris (N4,N4'-bis (3,5-di (1,4,7,10,13-pentaoxa-16-azacyclooctadecane-16-carbonyl) phenyl-[2,2'-bipyridine]-4,4'-dicarboxamide) ruthenium(II) dinitrate, 42d¹⁴²



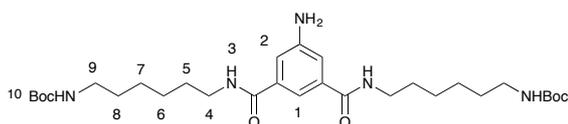
N4,N4'-bis(3,5-di(1,4,7,10,13-pentaoxa-16-azacyclooctadecane-16-carbonyl)phenyl-[2,2'-bipyridine]-4,4'-dicarboxamide, 41d (294 mg, 0.190 mmol), Ru(II)(DMSO)₄Cl₂ (28.7 mg, 0.0592 mmol), silver nitrate (20.1 mg, 0.118 mmol) and ethanol (20 mL) were heated under reflux for 7 days. The resulting red solution was then filtered and the filtrate concentrated *in vacuo*. The resulting red solid was purified by flash alumina column chromatography (1 – 10 % methanol in chloroform) and the red fractions collected. These were then redissolved in water (20 mL) and washed with ether (2 × 20 mL), then extracted with chloroform (3 × 40 mL), and the combined chloroform phases concentrated to yield the product as a red solid (94 mg, 0.019 mmol, 33 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 3.44 - 3.94 (m, 288 H, H7 – H12), 6.71 - 6.80 (m, 18 H, H6 + H4 + H3), 7.39 - 7.51 (m, 18 H, H5 + H2), 8.06 - 8.13 (m, 6 H, H1); IR (solid state, cm⁻¹) 3423 (N-H), 1672 (C=O amide), 1627 (C=O amide) (MS was not able to be obtained, presumably due to the multiple ionisation states present due to chelation of many different ions)

Di-tert- butyl (((5- nitroisophthaloyl) bis(azanediyl)) bis (hexane -6,1- diyl)) dicarbamate, 37e^{142,143}



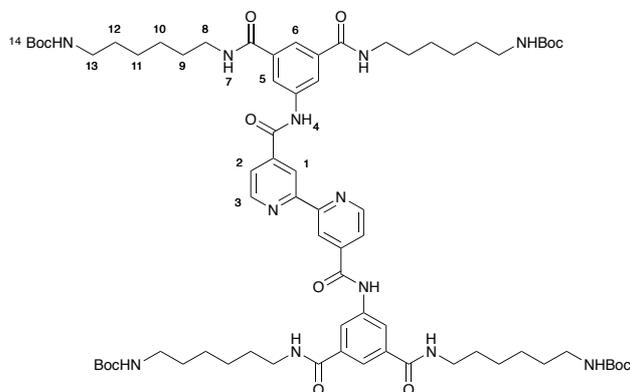
5-nitroisophthalic acid (200 mg, 0.950 mmol), thionyl chloride (3 mL, 41.4 mmol) and dimethylformamide (1 drop) were heated under reflux for 5 hours. The solvent was then removed *in vacuo* to yield the acid chloride as a pale yellow solid, which was reacted immediately. The acid chloride was redissolved in anhydrous dichloromethane (15 mL) and *N*-*tert*-butyl-6-aminohexanoate (0.47 mL, 2.09 mmol) and triethylamine (0.29 mL, 2.09 mmol) added and the resulting solution stirred for 18 hours. After which time the mixture was washed with 1 M hydrochloric acid (50 mL), saturated sodium hydrogen carbonate solution (50 mL) and brine (50 mL), and dried (sodium sulfate). The solution was concentrated to yield the crude product as a yellow oil. The crude product was purified by flash column chromatography (6:1 to 3:1 ethyl acetate:dichloromethane) to yield the product as a cream solid, (334 mg, 0.550 mmol, 58 %); ^1H NMR (300 MHz, CDCl_3) δ ppm 1.25 - 1.56 (m, 26 H, H6 + H7 + H10), 1.57 - 1.69 (m, 8 H, H5 + H8), 3.16 (d, $J = 6.2$ Hz, 4 H, H9), 3.46 (q, $J = 5.9$ Hz, 4 H, H4), 4.85 (br. s, 2 H, NH_{Boc}), 7.28 (br. s, 2 H, H3), 8.71 (s, 1 H, H1), 8.89 (s, 2 H, H2); ^{13}C NMR (75 MHz, CDCl_3) δ ppm 25.6, 25.9, 28.4, 29.1, 29.9, 39.9, 79.2, 124.8, 131.0, 136.6, 148.4, 156.6, 164.6; IR (solid state, cm^{-1}) 3342 (N-H), 1674 (C=O carbamate), 1645 (C=O amide), 1580 (NO_2), 1517 (NO_2); ESI-MS m/z found 608.3672 $[\text{M}+\text{H}]^+$, $[\text{C}_{30}\text{H}_{50}\text{N}_5\text{O}_8]^+$ requires 608.3659

Di-*tert*-butyl (((5- aminoisophthaloyl) bis (azanediyl)) bis (hexane-6,1-diyl)) dicarbamate, 38e^{142,143}



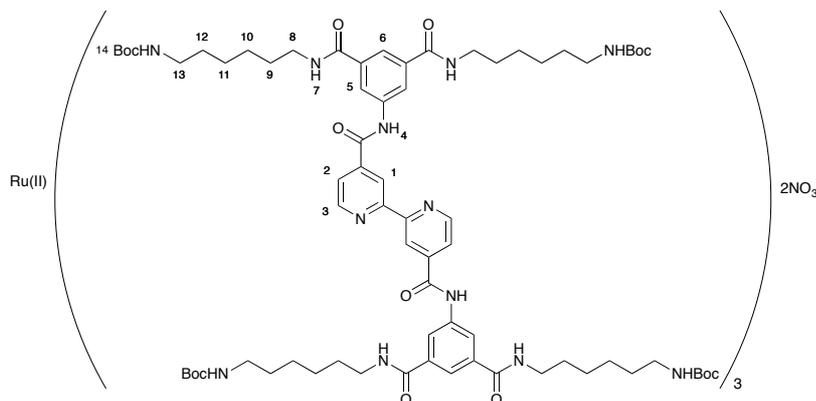
Di-*tert*-butyl(((5-nitroisophthaloyl)bis(azanediyl))bis(hexane-6,1-diyl))dicarbamate, **37e** (1.30 g, 2.14 mmol) in methanol (20 mL) and ethyl acetate (20 mL) was degassed and palladium on activated charcoal added (156 mg) and the solution degassed again. The reaction mixture was then placed under a hydrogen atmosphere and stirred for 18 hours. The reaction mixture was then filtered twice and concentrated to yield the product as a cream solid (1.148 g, 1.99 mmol, 91 %); ^1H NMR (300 MHz, CDCl_3) δ ppm 1.30- 1.43 (m, 4 H, H5), 1.40 (s, 18 H, H10), 1.43- 1.52 (m, 4 H, H7), 1.46 - 1.55 (m, 4 H, H6), 1.56 - 1.72 (m, 4 H, H8), 3.16 (d, $J = 6.0$ Hz, 4 H, H9), 3.46 (app. q, $J = 6.0$ Hz, 4 H, H4), 8.71 (s, 1 H, H1), 8.89 (s, 2 H, H2); ^{13}C NMR (75 MHz, CDCl_3) δ ppm 26.0, 26.3, 28.4, 29.3, 29.9, 39.8, 40.1, 79.1, 114.8, 116.8, 135.6, 147.1, 156.4, 167.3; IR (solid state, cm^{-1}) 3355 (N-H), 1685 (C=O amide); ESI-MS m/z found 578.3921 $[\text{M}+\text{H}]^+$, $[\text{C}_{30}\text{H}_{52}\text{N}_5\text{O}_6]^+$ requires 578.3912

Tert-butyl N-{6-[(3-{2-[4-({3,5-*bis*[(6-[[*tert*-butoxy) carbonyl] amino} hexyl) carbamoyl] phenyl} carbamoyl) pyridin-2-yl] pyridine-4-amido)-5-[(6-[[*tert*-butoxy) carbonyl]amino}hexyl)carbamoyl]phenyl)formamido]hexyl}carbamate, **41e** ^{142,143}



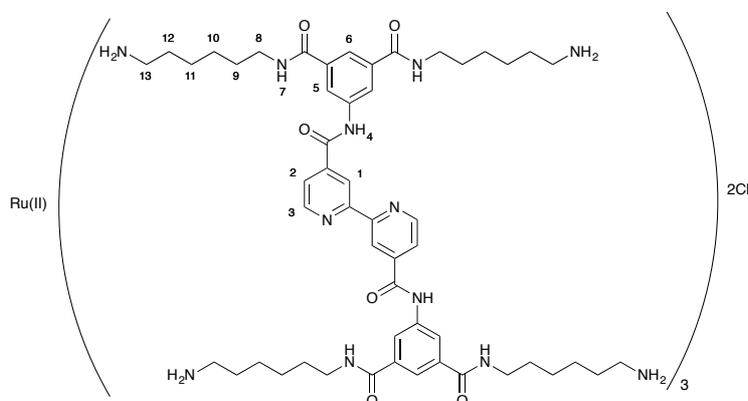
2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (100 mg, 0.410 mmol) and thionyl chloride (10 mL) were heated under reflux for 18 hours. The reaction mixture was concentrated *in vacuo* and the acid chloride flushed with nitrogen and reacted immediately. The acid chloride was redissolved in anhydrous chloroform (10 mL) and heated to reflux, under a nitrogen atmosphere. A solution of di-*tert*-butyl(((5-aminoisophthaloyl)*bis*(azanediyl))*bis*(hexane-6,1-diyl))dicarbamate, **38e** (426 mg, 0.737 mmol) and anhydrous triethylamine (0.13 mL, 0.90 mmol) in anhydrous chloroform (25 mL) under a nitrogen atmosphere was added dropwise to the refluxing solution and the resulting mixture heated under reflux for 18 hours. The reaction mixture was then concentrated to yield the crude product as a pink solid, which was purified by flash column chromatography (5 – 20 % methanol in dichloromethane) to yield the product as a beige solid (220 mg, 0.161 mmol, 39 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 1.26 - 1.35 (m, 16 H, H10 + H11), 1.40 (s + m, 44 H, H14 + H9/H12), 1.55 (quin., *J* = 7.1 Hz, 8 H, H9/H12), 2.92 (q, *J* = 6.2 Hz, 8 H, H8/H13), 3.29 (q, *J* = 7.1 Hz, 8 H, H8/H13), 6.78 (t, *J* = 5.5 Hz, 4 H, H7), 8.03 (s, 2 H, H1), 8.05 (d, *J* = 5.2 Hz, 2 H, H2), 8.40 (s, 2 H, H6), 8.55 (t, *J* = 5.2 Hz, 4 H, H3), 8.95 - 9.04 (m, 4 H, H5), 10.99 (s, 2 H, H4); ESI-MS *m/z* found [M+H]⁺, [C₇₂H₁₀₇N₁₂O₁₄]⁺ requires 1363.8030

Tris (tert-butyl N-{6-[(3-{2-[4-({3,5-bis[(6-{{(tert-butoxy)carbonyl}amino}hexyl) carbamoyl] phenyl} carbamoyl) pyridin-2-yl] pyridine-4-amido}-5- [(6-{{(tert-butoxy) carbonyl} amino} hexyl) carbamoyl] phenyl) formamido] hexyl} carbamate)) ruthenium(II) dinitrate, 42e^{142,143}



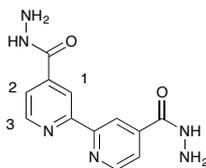
Tert-butyl N- {6- [(3- {2- [4-({3,5-bis [(6- {(tert- butoxy) carbonyl] amino} hexyl) carbamoyl] phenyl} carbamoyl) pyridin- 2- yl] pyridine-4- amido} -5-[(6- {(tert- butoxy)carbonyl]amino}hexyl)carbamoyl]phenyl)formamido]hexyl}carbamate, **41e** (172 mg, 0.126 mmol), Ru(II)(DMSO)₄Cl₂ (19 mg, 0.0394 mmol) and silver nitrate (13 mg, 0.0788 mmol) in ethanol (20 mL) were heated under reflux for 7 days. The reaction mixture was then filtered and the filtrate evaporated to yield the product as a red solid (223 mg, 0.0517 mmol, quant.); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 1.22 - 1.34 (m, 24 H, H10), 1.36 (s, 108 H, H14), 1.54 (m, 24 H, H8), 2.56 (s, 24 H, H11), 2.91 (q, *J* = 6.5 Hz, 24 H, H9), 3.29 (d, *J* = 5.0 Hz, 24 H, H12), 5.77 (s, 24 H, H8/H13), 6.76 (br. s, 24 H, H8/H13), 8.06 (s, 6 H, H6), 8.08 - 8.12 (m, 6 H, H2), 8.12 - 8.22 (m, 6 H, H3), 8.39 (s, 12 H, H5), 8.53 (br. s, 6 H, H1), 9.49 (br. s, 12 H, H7), 11.06 (br. s, 6 H, H4); IR (solid state, cm⁻¹) 3254 (broad) (N-H), 1558 (broad) (C=O amide/carbamate); ESI-MS *m/z* found 2096.1500 [M]²⁺, [C₂₁₆H₃₁₈N₃₆O₄₂Ru]²⁺ requires 2095.1449

Tris (N1,N3-bis (6-aminohexyl)-5-{2-[4-({3,5-bis [(6-aminohexyl) carbamoyl] phenyl} carbamoyl) pyridin-2-yl] pyridine-4-amido} benzene-1,3-dicarboxamide) ruthenium(II) dichloride, 35^{142,143}



Tris (tert-butyl N-{6-[(3-{2-[4-({3,5-bis[(6-{{tert-butoxy}carbonyl} amino} hexyl) carbamoyl] phenyl} carbamoyl) pyridin-2-yl] pyridine-4-amido}-5-[(6-{{tert-butoxy}carbonyl}amino}hexyl)carbamoyl]phenyl)formamido]hexyl}carbamate)) ruthenium(II) dinitrate, 42e (50 mg, 0.012 mmol) in 1 M hydrogen chloride in dioxane (5 mL) was stirred for 18 hours. The solvent was then evaporated and the red solid redissolved in water (10 mL). The solution was then neutralized with 1 M sodium hydroxide solution and the solvent removed. The resulting solid was dialysed (MWCO 0.1 – 0.5 kDa) against pure water to yield the product as a red solid (22 mg, 0.0072 mmol, 60 %); ¹H NMR (500 MHz, D₂O) δ ppm 1.51 (br. m, 96, H9 – H12), 2.86 (br. m, 48 H, H8 + H13), 7.20 – 8.70 (br. m, 36 H, H1 – H6); IR (solid state, cm⁻¹) 3385 (N-H), 3263 (N-H), 3043 (N-H), 1671 (C=O amide), 1641 (C=O amide) (The many possible ionisation states present prevented mass spectrum analysis)

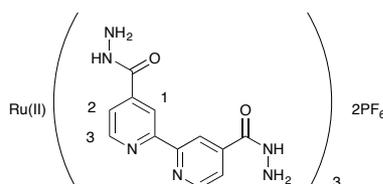
[2,2'-bipyridine]-4,4'-dicarbohydrazide, 48



4,4'-Dimethylester-2,2'-bipyridine, **47** (100 mg, 0.367 mmol) and hydrazine monohydrate (0.036 mL, 0.73 mmol) in methanol (5 mL) were heated under reflux for 6 hours. The solution was cooled to room temperature, filtered and the precipitate washed with ethanol to yield the product as a very insoluble white powder (66 mg, 0.24 mmol, 67

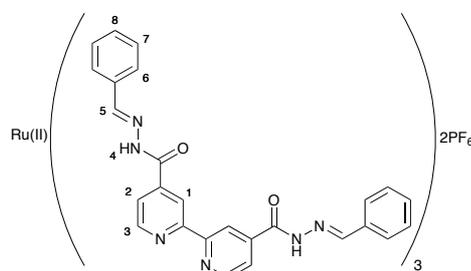
); ^1H NMR (500 MHz, TFA- d_x) δ ppm 8.53 (d, $J = 5.3$ Hz, 2 H, H2), 9.26 (d, $J = 5.3$ Hz, 2 H, H3), 9.33 (s, 2 H, H1); ^{13}C NMR (126 MHz, TFA- d_x) δ ppm 122.1, 125.9, 143.0, 147.0, 147.7, 163.7; IR (solid state, cm^{-1}): 3292 (N-H), 3065 (N-H), 1611 (C=O amide) (Poor solubility limited the ability to obtain mass spectra)

Tris ([2,2'-bipyridine]-4,4'-dicarbohydrazide) dihexafluorophosphate, 43



Tris (4,4'-dimethylester-2,2'-bipyridine) ruthenium(II) dihexafluorophosphate, **49** (128 mg, 0.0990 mmol) and hydrazine monohydrate (0.053 mL, 1.08 mmol) in methanol (10 mL) were heated under reflux for 5 hours. The hot solution was diluted with ethanol and filtered. The precipitate was redissolved in water, filtered and the filtrate evaporated to yield the product as a red solid (120 mg, 0.093 mmol, 94%); ^1H NMR (300 MHz, D_2O) δ ppm 7.48 - 7.58 (m, 6 H, H2), 7.79 (d, $J = 5.9$ Hz, 6 H, H3), 8.78 (br. s, 6 H, H1); IR (solid state, cm^{-1}) 3614 (N-H), 3062 (N-H), 1654 (C=O amide); ESI-MS m/z found 459.1026 $[\text{M}]^{2+}$, $[\text{C}_{36}\text{H}_{36}\text{N}_{18}\text{O}_6\text{Ru}]^{2+}$ requires 459.1054 (this Ru(II)(bpy) $_3$ complex is not stable to O_2 and was therefore reacted on immediately)

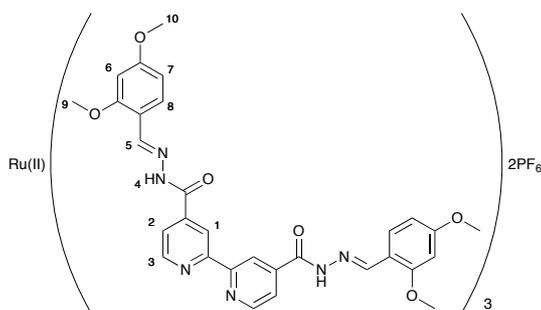
Tris (N'- [(1E)- phenylmethylidene] -2-(4-{N'- [(1E)- phenylmethylidene] hydrazinecarbonyl} pyridin-2-yl) pyridine-4-carbohydrazide) ruthenium(II) dihexafluorophosphate, 46a



To *tris* ([2,2'-bipyridine]-4,4'-dicarbohydrazide) ruthenium(II) dihexafluorophosphate, **43** (94 mg, 0.0778 mmol) in degassed acetonitrile (10 mL) and water (10 mL) under a nitrogen atmosphere was added benzaldehyde (148 mg, 0.891 mmol) and the resulting solution stirred for 30 minutes. The solution was then concentrated *in vacuo*, and the resulting red solid suspended in chloroform (20 mL) and filtered. The precipitate was

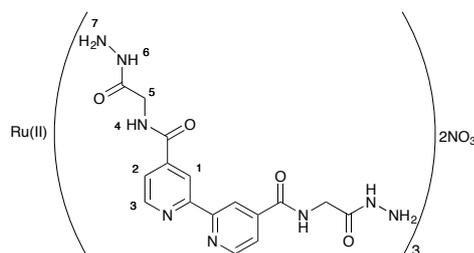
redissolved in acetonitrile and filtered, the filtrate was concentrated to yield the product as a red solid (112 mg, 0.0713, 92 %); $^1\text{H NMR}$ (500 MHz, CD_3CN) δ ppm 3.64 - 3.72 (br. m, 2 H, H4), 3.72 - 3.80 (br. m, 4 H, H4'), 6.36 (s, 6 H, H1), 7.53 (t, $J = 8.0$ Hz, 8 H, H7), 7.66 (t, $J = 7.1$ Hz, 4 H, H7'), 7.77 (ddd, $J = 7.2, 5.0, 1.6$ Hz, 6 H, H2), 8.04 (d, $J = 7.1$ Hz, 8 H, H6), 8.09 (td, $J = 8.0, 1.8$ Hz, 4 H, H8), 8.12 (d, $J = 8.0$ Hz, 4 H, H6'), 8.16 (t, $J = 7.1$ Hz, 8 H, H8'), 8.66 (d, $J = 7.2$ Hz, 6 H, H3), 8.74 (d, $J = 6.0$ Hz, 2 H, H5), 8.84 (d, $J = 4.6$ Hz, 4 H, H5') (*cis* and *trans* isomers of hydrazone seen in 1:2 ratio); IR (solid state, cm^{-1}) 3368 (N-H), 1723 (C=O amide), 1654 (C=N); ESI-MS m/z found 723.2003 $[\text{M}]^{2+}$, $[\text{C}_{78}\text{H}_{60}\text{N}_{18}\text{O}_6\text{Ru}]^{2+}$ requires 723.1993

***Tris* (N'- [(1E)- (2,4- dimethoxyphenyl) methyldene] -2- (4-{N'- [(1E)- (2,4- dimethoxyphenyl) methyldene] hydrazinecarbonyl} pyridin-2-yl) pyridine-4- carbohydrazide) dihexafluorophosphate, 46b**



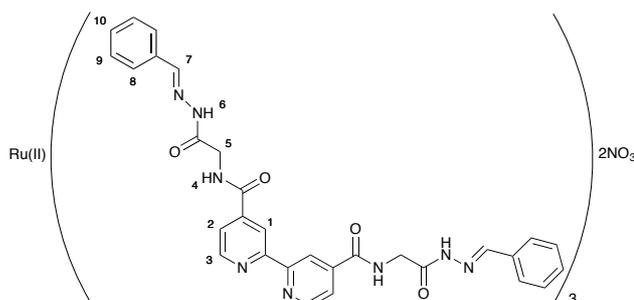
To *tris* ([2,2'-bipyridine]-4,4'-dicarbohydrazide) ruthenium(II) dihexafluorophosphate, **43** (94 mg, 0.0778 mmol) in degassed acetonitrile (10 mL) and water (10 mL) under a nitrogen atmosphere was added 2,4-dimethoxy benzaldehyde (148 mg, 0.891 mmol) and stirred for 16 hours. The solution was then concentrated *in vacuo*, and the resulting red solid suspended in chloroform (20 mL) and filtered. The precipitate was redissolved in acetonitrile and filtered, the filtrate was concentrated to yield the product as a red solid (84 mg, 0.044, 56 %). $^1\text{H NMR}$ (500 MHz, CD_3CN) δ ppm 3.94 (s, 12 H, H9/H10), 4.01 (s, 12 H, H9/H10), 4.02 - 4.03 (m, 6 H, H9'/H10'), 4.05 (m, 6 H, H9'/H10'), 6.40 - 6.45 (m, 6 H, H4), 6.67 (d, $J = 2.3$ Hz, 6 H, H1), 6.73 (td, $J = 6.5, 2.3$ Hz, 6 H, H8), 7.55 (app. d, $J = 9.4$ Hz, 6 H, H2), 7.63 (dd, $J = 6.5, 3.7$ Hz, 6 H, H7), 7.74 (d, $J = 3.7$ Hz, 6 H, H6), 7.86 (d, $J = 9.4$ Hz, 6 H, H3), 8.87 (app. s, 6 H, H5) (*cis* and *trans* hydrazone isomers seen on OMe in 1:2 ratio); IR (solid state, cm^{-1}) 1665 (C=O amide), 1594 (C=N); ESI-MS m/z found 903.2673 $[\text{M}]^{2+}$, $[\text{C}_{90}\text{H}_{84}\text{N}_{18}\text{O}_{18}\text{Ru}]^{2+}$ requires 903.2627

Tris (N-[(hydrazinecarbonyl) methyl]-2- (4- {[(hydrazinecarbonyl) methyl] carbamoyl} pyridin-2-yl)pyridine-4-carboxamide) ruthenium(II) dinitrate, 49



Tris (ethyl 2-[(2-{4-[(2-ethoxy-2-oxoethyl)carbamoyl]pyridin-2-yl}pyridin-4-yl) formamido] acetate) ruthenium(II) dinitrate, **42f** (50 mg, 0.0340 mmol) and hydrazine monohydrate (0.02 mL, 0.408 mmol) in ethanol (5 mL) were heated under reflux, under a nitrogen atmosphere, for 18 hours, then filtered. The red precipitate was redissolved in water, filtered and concentrated to yield the product as a red solid (45 mg, 0.034 mmol, 95 %); ¹H NMR (500 MHz, D₂O) δ ppm 4.08 (s, 12 H, H5), 7.73 (d, *J* = 5.4 Hz, 6 H, H2), 7.94 (d, *J* = 5.4 Hz, 6 H, H3) 8.97 (s, 6 H, H1); IR (solid state, cm⁻¹) 3246 (N-H), 3065 (N-H), 1733 (C=O amide), 1650 (C=O amide); ESI-MS found *m/z* 630.1641 [M]²⁺, [C₄₈H₅₄N₂₄O₁₂Ru]²⁺ requires 630.169

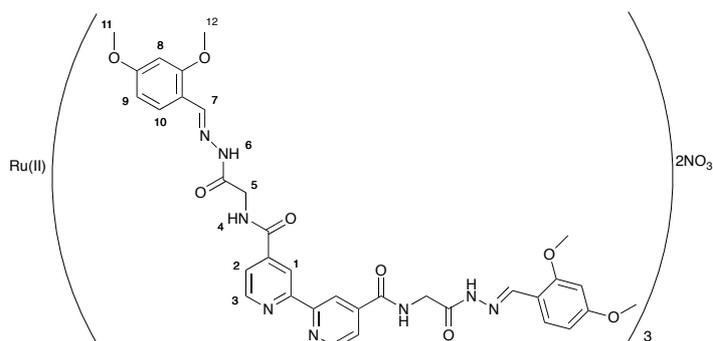
Tris (N-({N'-[(1E)-phenylmethylidene]hydrazinecarbonyl}methyl)-2-{4-[(N'-[(1E)-phenylmethylidene]hydrazinecarbonyl}methyl)carbamoyl]pyridin-2-yl}pyridine-4-carboxamide) dinitrate, 50a



To *tris* (N- [(hydrazinecarbonyl) methyl]-2- (4- {[(hydrazinecarbonyl) methyl] carbamoyl} pyridin-2-yl)pyridine-4-carboxamide) ruthenium(II) dinitrate, **49** (50 mg, 0.036 mmol) in degassed acetonitrile (10 mL) and water (10 mL) under a nitrogen atmosphere was added benzaldehyde (148 mg, 0.891 mmol) and the resulting solution stirred for 30 minutes. The solution was then concentrated *in vacuo*, and the resulting red solid suspended in chloroform (20 mL) and filtered, to yield the product as a red precipitate (55 mg, 0.029 mmol, 80 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 4.14 (m, 8 H, H5), 4.56 (br. s, 4 H, H5'), 7.45 (br. s, 8 H, H9), 7.72 (br. s, 4 H, H9'), 7.92 (br. s, 12 H, H1 +

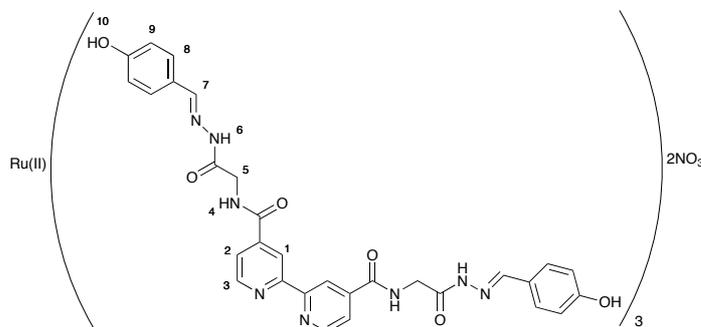
H2), 8.00 - 8.14 (m, 8 H, H8), 8.25 (br. s, 4 H, H8'), 9.39 (br. s, 6 H, H3), 9.50 (br. s, 4 H, H10), 9.70 (d, $J = 23.8$ Hz, 2 H, H10'), 11.62 (br. s, 4 H, H7), 11.65 - 11.71 (m, 2 H, H7') (*cis* and *trans* hydrazone isomers observed in 1:2 ratio); IR (solid state, cm^{-1}) 3218 (N-H), 1655 (C=O amide), 1541 (C=N); ESI-MS found m/z 894.2639 $[\text{M}]^{2+}$, $[\text{C}_{90}\text{H}_{78}\text{N}_{24}\text{O}_{12}\text{Ru}]^{2+}$ requires 894.2640

***Tris* (N-({N'-[(1E)-2,4 dimethoxy phenylmethylidene] hydrazinecarbonyl} methyl)-2-{4-[(N'-[(1E)-2,4 dimethoxy phenylmethylidene] hydrazinecarbonyl} methyl)carbamoyl] pyridin-2-yl}pyridine-4-carboxamide) dinitrate, 50b**



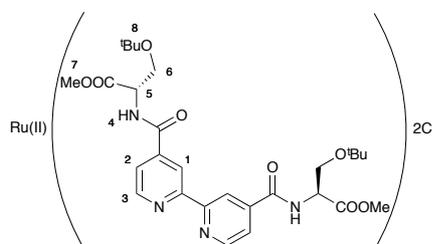
To *tris* (N-[(hydrazone carbonyl) methyl]-2-(4-[(hydrazone carbonyl) methyl] carbamoyl]pyridin-2-yl)pyridine-4-carboxamide) ruthenium(II) dinitrate (50 mg, 0.036 mmol) in degassed acetonitrile (10 mL) and water (10 mL) under a nitrogen atmosphere was added 2,4-dimethoxy benzaldehyde (46 mg, 0.43 mmol) and the resulting solution stirred for 30 minutes. The solution was then concentrated *in vacuo*, and the resulting red solid suspended in dichloromethane (20 mL) and filtered, to yield the product as a red solid (60 mg, 0.026, 73 %); ^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ ppm 3.84 (m, 36 H, H11 + H12), 4.00 - 4.19 (m, 4 H, H5), 4.52 (br. s, 8 H, H5'), 6.48 - 6.74 (m, 12 H, H8 + H9), 7.56 - 7.82 (m, 6 H, H2), 7.91 (br. s, 6 H, H1), 8.00 - 8.17 (br. s, 6 H, H3), 8.27 (br. s, 4 H, H9), 8.41 - 8.56 (m, 2 H, H9'), 9.26 - 9.53 (m, 6 H, H10), 11.28 - 11.57 (m, 6 H, H7) (*cis* and *trans* hydrazone isomers seen in 1:2 ratio); IR (solid state, cm^{-1}) 3215 (N-H), 1659 (C=O amide), 1600 (C=N); ESI-MS found m/z 1074.3298 $[\text{M}]^{2+}$, $[\text{C}_{102}\text{H}_{102}\text{N}_{24}\text{O}_{24}\text{Ru}]^{2+}$ requires 1074.3271

Tris (N-({N'-[(1E)-4 hydroxy phenylmethylidene]hydrazinecarbonyl)methyl)-2-{4-[[{N'-[(1E)-4 hydroxy phenyl methylidene] hydrazine carbonyl} methyl] carbamoyl} pyridin-2-yl}pyridine-4-carboxamide) dinitrate, 50d



To *tris* (N-[(hydrazine carbonyl) methyl]-2-(4-[(hydrazine carbonyl) methyl] carbamoyl} pyridin-2-yl)pyridine-4-carboxamide) ruthenium(II) dinitrate (10 mg, 0.0072 mmol) in degassed acetonitrile (5 mL) and water (5 mL) under a nitrogen atmosphere was added 4-hydroxy benzaldehyde (20 mg, 0.16 mmol) and the resulting solution stirred for 30 minutes. The solution was then concentrated *in vacuo*, and the resulting red solid suspended in dichloromethane (20 mL) and filtered, to yield the product as a red solid (16 mg, 0.080 mmol, quant.); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 4.04 (br. s, 2 H, H5), 4.48 (br. s, 10 H, H5'), 6.80 (br. s, 10 H, H9), 6.93 (br. s, 2 H, H9'), 7.45 (br. s, 10 H, H8), 7.74 (br. s, 2 H, H8'), 7.88 (br. s, 6 H, H2), 8.10 (br. s, 6 H, H3), 9.28 (br. s, 6 H, H1), 9.75 (br. s, 1 H, H7), 9.90 (br. s, 5 H, H7'), 11.37 (br. s, 6 H, H4) (*cis* and *trans* hydrazone isomers seen in a 1:5 ratio); IR (solid state, cm⁻¹) 3213 (N-H), 3075 (O-H), 1654 (C=O amide), 1603 (C=N); ESI-MS found *m/z* 972.2506 [M]²⁺, [C₉₀H₇₈N₂₄O₁₈Ru]²⁺ requires 1074.3271

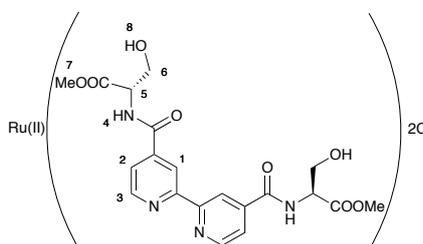
Tris (methyl (2S)-3-(*tert*-butoxy)- 2-[[2-(4-[[{(2S)-3-(*tert*-butoxy)- 1-methoxy- 1-oxopropan-2-yl] carbamoyl} pyridin-2-yl) pyridin-4-yl]formamido} propanoate) ruthenium(II) dichloride, 41j



Tris (2,2'-bipyridine-4,4'-dicarboxylic acid) ruthenium(II) dichloride, **29** (120 mg, 0.133 mmol) was heated under reflux in thionyl chloride (20 mL) and dimethylformamide (1 drop) for 6 hours. The reaction mixture was concentrated *in vacuo* and the acid chloride flushed with nitrogen and used immediately. The resulting acid chloride was resuspended

in anhydrous chloroform (30 mL) and heated to reflux. *O*-*tert*-Butyl-L-serine methyl ester hydrochloride salt (252 mg, 1.19 mmol) and anhydrous diisopropylethylamine (0.41 mL, 2.4 mmol) were added to the refluxing solution and the reaction mixture heated under reflux for 18 hours under a nitrogen atmosphere. The reaction mixture was cooled to room temperature and washed with saturated sodium hydrogen carbonate solution (30 mL), 1 M hydrochloric acid (30 mL) and brine (30 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a red solid. This was purified by flash column chromatography (10 % methanol in dichloromethane) to yield the product as a red solid (108 mg, 0.0584 mmol, 44 %); ¹H NMR (500 MHz, CDCl₃) δ ppm, 1.17 (s, 54 H, H8), 3.74 (s, 18 H, H7), 3.80 (dd, *J* = 9.2, 4.4 Hz, 6 H, H6'), 3.93 (dd, *J* = 9.2, 4.4 Hz, 6 H, H6), 4.86 (dt, *J* = 7.4, 4.4 Hz, 6 H, H5), 7.82 (dd, *J* = 12.3, 5.4 Hz, 6 H, H2), 7.92 (d, *J* = 5.4 Hz, 6 H, H1), 8.78 (d, *J* = 7.4 Hz, 6 H, H4), 9.36 (d, *J* = 12.3 Hz, 6 H, H3); IR (solid state, cm⁻¹) 3243 (N-H), 1739 (C=O ester), 1662 (C=O amide); ESI-MS *m/z* found 888.3430 [M]²⁺, [C₈₄H₁₁₄N₁₂O₂₄Ru]²⁺ requires 888.3556

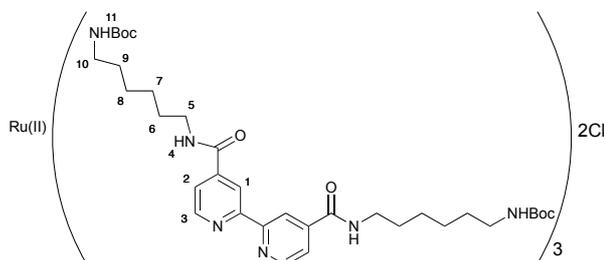
Tris (methyl (2S)-3-(hydroxy)-2-[[2-(4-[[[(2S)-3-(hydroxy)-1-methoxy-1-oxopropan-2-yl]carbamoyl] pyridin-2-yl] pyridin-4-yl] formamido] propanoate) ruthenium(II) ditrifluoroacetate, 58a



Tris (methyl (2S)-3-(*tert*-butoxy)-2-[[2-(4-[[[(2S)-3-(*tert*-butoxy)-1-methoxy-1-oxopropan-2-yl] carbamoyl] pyridin-2-yl] pyridin-4-yl] formamido] propanoate) ruthenium(II) dinitrate, 41j (25 mg, 0.015 mmol) in trifluoroacetic acid (4.5 mL), water (0.4 mL) and triisopropylsilane (0.1 mL) was stirred for 6 hours. The reaction mixture was then concentrated to yield the product as a red solid (24 mg, 0.014 mmol, 96 %); ¹H NMR (500 MHz, MeOD) δ ppm 2.70 (s, 18 H, H7), 4.00 (dd, *J* = 11.5, 3.8 Hz, 6 H, H6'), 4.06 (dd, *J* = 11.5, 5.5 Hz, 6 H, H6), 4.76 - 4.87 (m, 6 H, H5), 7.95 (d, *J* = 5.8 Hz, 6 H, H2), 8.09 (d, *J* = 5.8 Hz, 6 H, H3), 9.27 (s, 6 H, H1); IR (solid state, cm⁻¹) 3290 (O-H), 3071 (N-H), 1733 (C=O ester), 1656 (C=O amide); ESI-HRMS found *m/z* 720.1714 [M]²⁺, [C₆₀H₆₆N₁₂O₂₄Ru]²⁺ requires 720.1678

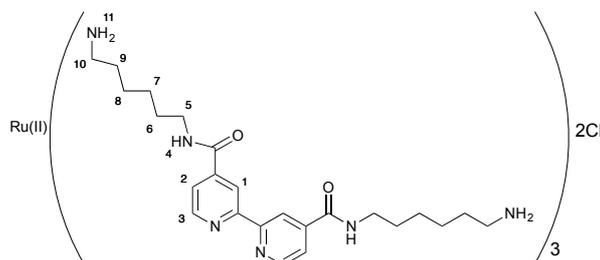
Tris (N4, N4'- bis (6-boc aminohexyl) -2,2'- bipyridine- 4,4'- dicarboxamide)

ruthenium(II) dichloride, 41i



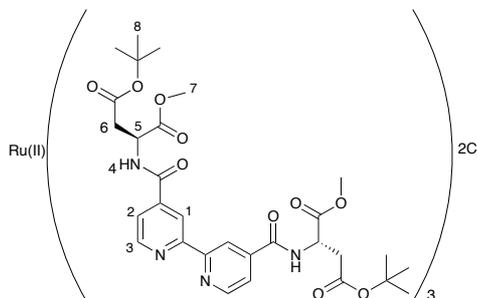
Tris (2,2'-bipyridine-4,4'-dicarboxylic acid) ruthenium(II) dichloride, **29** (114 mg, 0.125 mmol) was heated under reflux in thionyl chloride (30 mL) and dimethylformamide (1 drop) for 6 hours. The solvent was removed *in vacuo* and the resulting red acid chloride flushed with nitrogen and used immediately. The acid chloride was resuspended in anhydrous chloroform (30 mL) and heated to reflux, under a nitrogen atmosphere. *N*-Boc-1,6-diamino hexane (0.25 mL, 1.1 mmol) and anhydrous diisopropylethylamine (0.39 mL, 2.3 mmol) were added, and the resulting solution heated under reflux for 16 hours. The reaction mixture was then allowed to cool to room temperature, and the reaction mixture quenched with saturated sodium hydrogen carbonate solution (30 mL). The aqueous layer was removed and the organic phase washed with 1 M hydrochloric acid (30 mL) and brine (30 mL). The organic phase was dried (sodium sulfate) and concentrated in order to yield the crude product as a red solid. This was purified by flash column chromatography (10 % methanol in dichloromethane) to yield the product as a red solid (95 mg, 0.045 mmol, 36 %); ^1H NMR (500 MHz, CDCl_3) δ ppm 1.34 (br. d, $J = 5.2$ Hz, 12 H, H7/H8), 1.39 (s, 54 H, H11), 1.44 - 3.47 (br. s, 12 H, H7/H8), 1.53 (m, 12 H, H6/H9), 1.69 (br. s, 24 H, H6/H9 + H5/H10), 3.09 (br. s, 12 H, H5/H10), 4.77 (br. s, 6 H, NH_2), 7.66 (br. s, 6 H, H2), 8.06 (br. s, 6 H, H1), 8.97 (br. s, 6 H, H3), 10.02 (br. s, 6 H, H4); IR (solid state, cm^{-1}) 3291 (N-H), 1657 (C=O amide); ESI-HRMS found m/z 1011.5489 $[\text{M}]^{2+}$, $[\text{C}_{102}\text{H}_{156}\text{N}_{18}\text{O}_{18}\text{Ru}]^{2+}$ requires 1011.5444

Tris (N4,N4'-bis (6-aminohexyl) -2,2'- bipyridine-4,4'- dicarboxamide) ruthenium(II) dichloride, 81



Tris (N4,N4'-bis(6-Boc aminoethyl)-2,2'-bipyridine-4,4'-dicarboxamide) ruthenium(II) dichloride (20 mg, 0.0095 mmol) was stirred in 1 M hydrogen chloride in dioxane (5 mL) and water (0.5 mL) for 2 hours. The resulting mixture was concentrated and redissolved in water (10 mL). The solution was neutralised by addition of 1 M sodium hydroxide solution. The neutral solution was concentrated to ~2 mL and the resulting solution dialysed (MWCO 0.1 - 0.5 kDa) against pure water to yield the product as a red solid (14 mg, 0.0094 mmol, 98 %); ¹H NMR (500 MHz, D₂O) δ ppm 1.37 (br. s, 24 H, H7 + H8), 1.62 (br. s, 24 H, H6 + H9), 2.94 (t, *J* = 7.5 Hz, 12 H, H5/H10), 3.38 (t, *J* = 6.8 Hz, 12 H, H5/H10), 7.67 (d, *J* = 5.8 Hz, 6 H, H3), 7.90 (d, *J* = 5.8 Hz, 6 H, H2), 8.89 (s, 6 H, H1); IR (solid state, cm⁻¹) 3386 (N-H), 3255 (N-H), 1717 (C=O amide); ESI-MS *m/z* found 711.3884 [M]²⁺, [C₇₂H₁₀₈N₁₈O₆Ru]²⁺ requires 711.3871

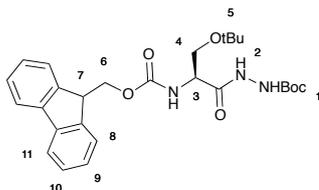
Tris (4-*tert*-butyl 1-methyl (2S)-2-[(4'-[(2S)-4-(*tert*-butoxy)-1-methoxy-1,4-dioxobutan-2-yl] carbamoyl]- [2,2'-bipyridin]-4-yl) formamido] butanedioate) ruthenium(II) dichloride, 41k



Tris (2,2'-bipyridine-4,4'-dicarboxylic acid) ruthenium(II) dichloride (50 mg, 0.055 mmol) was heated under reflux in thionyl chloride (20 mL) and dimethylformamide (1 drop) for 6 hours. The reaction mixture was concentrated *in vacuo* and the acid chloride flushed with nitrogen. The resulting acid chloride was resuspended in anhydrous

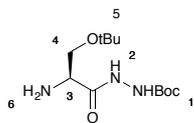
chloroform (30 mL) and H-Asp(OtBu) OMe. hydrogen chloride salt (199 mg, 0.497 mmol) and anhydrous diisopropylethylamine (0.17 mL, 0.99 mmol) were added and the reaction mixture heated under reflux under a nitrogen atmosphere for 18 hours. The resulting solution was then cooled to room temperature and washed with saturated sodium hydrogen carbonate solution (50 mL), 1 M hydrochloric acid (50 mL) and brine (50 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a red solid. This was purified by flash column chromatography (10 % methanol in dichloromethane) to yield the product as a red solid (30 mg, 0.015 mmol, 27%); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.36 - 1.52 (m, 60 H, H6 + H8), 1.59 - 1.71 (m, 6 H, H6'), 3.72 (s, 18 H, H7), 5.07 (m, 6 H, H5), 7.74 (br. s, 6 H, H3), 7.98 (br. s, 6 H, H2), 9.24 (br. s, 6 H, H1), 9.87 - 10.23 (m, 6 H, H4); IR (solid state, cm⁻¹) 3055 (N-H), 1723 (C=O ester), 1666 (C=O amide) ESI-MS *m/z* found 972.3447 [M]²⁺, [C₉₀H₁₁₄N₁₂O₃₀Ru] requires 972.3404

***Tert*-butyl 2- (2- (((9H-fluoren-9-yl) methoxy) carbonylamino) -3- *tert*-butoxypropanoyl) hydrazine carboxylate, 61a**



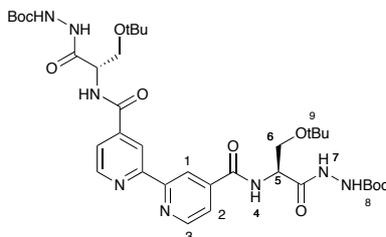
Fmoc-Ser (O^tBu) COOH (5.00 g, 13.0 mmol), *tert*-butyl carbazate (5.17 g, 39.1 mmol), HATU (5.44 g, 14.3 mmol) and diisopropylethylamine (9.02 mL, 52.0 mmol) were stirred in anhydrous dimethylformamide (20 mL) under a nitrogen atmosphere for 18 hours. The solution was then diluted with ethyl acetate (300 mL) and washed with saturated sodium hydrogen carbonate solution (200 mL), 1 M hydrochloric acid (200 mL) and brine (3 × 500 mL) to yield the crude product as an off-white solid. This was purified by flash column chromatography (20 % ethyl acetate in dichloromethane) to yield the product as a white solid (4.32 g, 8.68 mmol, 67 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.24 (s, 9 H, H1/H5), 1.50 (s, 9 H, H1/H5), 3.48 (m, 1 H, H2+H3), 3.74 (q, *J* = 6.5 Hz, 2 H, H4), 4.25 (t, *J* = 7.0 Hz, 1 H, H7), 4.44 (d, *J* = 7.0 Hz, 2 H, H6), 5.70 (br. s, 1 H, NH), 6.50 (br. s, 1 H, NH), 7.34 (t, *J* = 7.6 Hz, 2 H, Fmoc), 7.43 (t, *J* = 7.6 Hz, 2 H, Fmoc), 7.62 (dd, *J* = 7.6, 3.0 Hz, 2 H, Fmoc), 7.79 (d, *J* = 7.6 Hz, 2 H, Fmoc), 8.42 (br. s, 1 H, NH); ¹³C NMR (75 MHz, DMSO-*d*₆) δ ppm 27.2, 28.0, 46.6, 54.2, 61.4, 61.8, 65.8, 72.9, 120.0, 127.0, 127.6, 125.4, 140.7, 143.7, 143.8, 155.0, 155.8; IR (solid state, cm⁻¹) 3297 (N-H), 3256 (N-H), 1714 (C=O carbamate), 1688 (C=O amide); ESI-MS *m/z* found 520.2422 [M+Na]⁺, [C₂₇H₃₅N₃O₆Na]⁺ requires 520.2424

***Tert*-butyl 2-(2-amino-3-*tert*-butoxypropanoyl)hydrazinecarboxylate, 62a**



Tert-butyl 2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-*tert*-butoxypropanoyl)hydrazine carboxylate, **61a** (1.00 g, 2.00 mmol) in 20 % diethylamine in acetonitrile was stirred for 16 hours. The reaction mixture was concentrated, and redissolved in a minimal amount of ethyl acetate and precipitated with hexane. The suspension was filtered through a celite pad and washed with hexane. The celite pad was then washed with dichloromethane and methanol, and this filtrate concentrated to yield the product as an off-white waxy solid (490 mg, 1.78 mmol, 88 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.18 - 1.28 (m, 9 H, H1/H5), 1.50 (s, 9 H, H1/H5), 1.66 - 1.99 (m, 4 H, H6+H2+ NHBoc), 3.55 (dd, *J* = 4.7, 3.1 Hz, 1 H, H3), 3.59 - 3.68 (m, 2 H, H4/H4'); ¹³C NMR (101 MHz, MeOD) δ ppm 26.3, 27.2, 42.1, 63.2, 73.4, 80.6, 156.3. 172.5; IR (solid state, cm⁻¹) 3368 (N-H), 3242 (N-H), 1720 (C=O carbamate), 1692 (C=O amide); ESI-MS *m/z* found 276.1923 [M+H]⁺, [C₁₂H₂₆N₃O₄]⁺ requires 276.1923

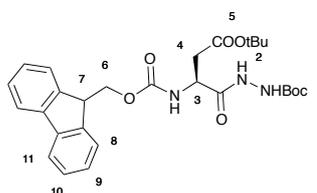
N₄,N_{4'}- bis [(1S) -2- (*tert*- butoxy) -1- {N'- [(*tert*- butoxy) carbonyl] hydrazinecarbonyl} ethyl]-[2,2'-bipyridine]-4,4'-dicarboxamide, 68



2,2'-Bipyridine-4,4'-dicarboxylic acid, **40** (241 mg, 0.988 mmol) was heated under reflux in thionyl chloride (30 mL) for 18 hours, and the reaction mixture then concentrated *in vacuo*. The resultant acid chloride was flushed with nitrogen and used immediately. The acid chloride was redissolved in anhydrous chloroform (30 mL) and heated to reflux, under a nitrogen atmosphere. *Tert*-butyl 2-(2-amino-3-*tert*-butoxypropanoyl) hydrazinecarboxylate, **62a** (598 mg, 2.17 mmol) and anhydrous diisopropylethylamine (0.76 mL, 4.34 mmol) was added to the refluxing solution and the mixture heated under reflux for 18 hours. The solution was cooled to room temperature and washed with 1 M hydrochloric acid (50 mL) and brine (50 mL). The organic phase was dried (sodium sulfate), and concentrated to yield the crude product as a pink solid. This

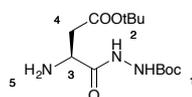
was purified by flash column chromatography (5% methanol in dichloromethane) to yield the product as a beige solid (122 mg, 0.178 mmol, 16 %); ^1H NMR (500 MHz, CDCl_3) δ ppm 1.30 (s, 18 H, H8/H9), 1.48 (s, 18 H, H8/H9), 3.67 (t, $J = 8.5$ Hz, 2 H, H6), 3.92 (dd, $J = 8.5$, 4.8 Hz, 2 H, H6'), 4.85 (br. s, 2 H, H5), 6.96 - 7.12 (br. s, 2 H, NH_{Boc}), 7.67 (d, $J = 4.4$ Hz, 2 H, H3), 7.78 - 7.91 (m, 2 H, H7), 8.53 (br. s, 2 H, H1), 8.70 (d, $J = 4.4$ Hz, 2 H, H2), 9.12 (br. s, 2 H, H4); IR (solid state, cm^{-1}) 3279 (N-H), 1702 (C=O carbamate), 1651 (C=O amide); ESI-MS m/z found 759.3905 $[\text{M}+\text{H}]^+$, $[\text{C}_{36}\text{H}_{55}\text{N}_8\text{O}_{10}]^+$ requires 759.4041

(S)-tert-butyl 2-(2-(((9H-fluoren-9-yl) methoxy) carbonylamino)-4-tert-butoxy-4-oxobutanoyl)hydrazinecarboxylate, 61b



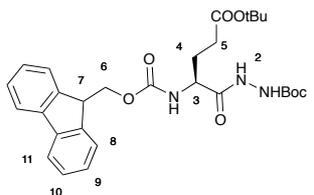
Fmoc- Asp(O^tBu)COOH (1.50 g, 3.65 mmol), *tert*-butyl carbazate (1.45 g, 10.9 mmol), HATU (1.53 g, 4.00 mmol) and diisopropylethylamine (1.27 mL, 7.30 mmol) in anhydrous dimethylformamide (10 mL) were stirred for 16 hours under a nitrogen atmosphere. The solution was then diluted with ethyl acetate (100 mL) and washed with saturated sodium hydrogen carbonate solution (100 mL), 1 M hydrochloric acid (100 mL), brine (3 × 200 mL) and ammonium hydroxide solution (100 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the product as a white solid (2.02 g, 3.85 mmol, quant.); ^1H NMR (500 MHz, CDCl_3) δ ppm 1.48 (s, 9 H, H1/H5), 1.49 (s, 9 H, H1/H5), 2.69 (m, 1 H, H4), 2.91 (m, 1 H, H4'), 4.25 (t, $J = 7.5$ Hz, 1 H, H7), 4.41 - 4.54 (m, 1 H, H3) 4.57 - 4.68 (m, 1 H, NH) 5.98 (d, $J = 7.5$ Hz, 1 H, H6), 6.40 (br. s, 1 H, NH), 7.34 (t, $J = 7.5$ Hz, 2 H, Fmoc) 7.43 (t, $J = 7.5$ Hz, 2 H, Fmoc), 7.61 (dd, $J = 7.5$, 3.7 Hz, 2 H, Fmoc), 7.79 (d, $J = 7.5$ Hz, 2 H, Fmoc), 8.22 (br. s, 1 H, NH); ^{13}C NMR (75 MHz, DMSO-d_6) δ ppm 27.4, 27.7, 28.0, 31.3, 46.6, 52.3, 65.7, 79.7, 120.1, 125.3, 127.1, 127.6, 140.7, 143.8, 155.1, 155.8, 171.0, 171.6; IR (solid state, cm^{-1}) 3285 (N-H), 1695 (C=O amide); ESI-MS m/z found 548.2265 $[\text{M}+\text{Na}]^+$, $[\text{C}_{28}\text{H}_{35}\text{N}_3\text{O}_7\text{Na}]^+$ requires 548.2373

(S)-tert-butyl 2-(2-amino-4-tert-butoxy-4-oxobutanoyl) hydrazinecarboxylate 62b



(S)-*tert*-butyl 2-(2-(((9H-fluoren-9-yl) methoxy) carbonyl amino)-4-*tert*-butoxy-4-oxobutanoyl) hydrazine carboxylate, **61b** (1.85 g, 3.52 mmol) in 20 % diethylamine in acetonitrile (50 mL) was stirred for 16 hours. The reaction solution was then concentrated and the resulting residue redissolved in a minimal amount of ethyl acetate, and precipitated by addition of hexane. The slurry was then filtered through celite. The celite pad was then washed with dichloromethane and methanol and this filtrate concentrated to yield the product as an off-white waxy solid (1.03 g, 3.40 mmol, 96%); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.48 (s, 9 H, H1/H2), 1.50 (s, 9 H, H1/H2), 2.59 (dd, *J* = 16.7, 8.1 Hz, 1 H, H4), 2.82 (dd, *J* = 16.7, 3.7 Hz, 1 H, H4'), 3.16 (br. s, 1 H, H3), 3.91 (br. s, 2 H, H6), 6.38 (br. s, 1 H, H2); ¹³C NMR (101 MHz, MeOD) δ ppm 23.5, 25.8, 27.0, 36.7, 76.7, 81.1, 155.4, 170.4, 172.3; IR (solid state, cm⁻¹) 3274 (N-H), 1707 (C=O ester); ESI-MS *m/z* found 326.1685 [M+Na]⁺, [C₁₃H₂₅N₃O₅Na]⁺ requires 326.1686

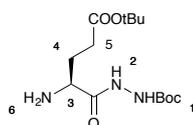
(S)-*tert*-butyl 2-(2-(((9H-fluoren-9-yl) methoxy) carbonylamino)-5-*tert*-butoxy-5-oxopentanoyl) hydrazinecarboxylate, **61c**



Fmoc Glu(O^tBu)COOH (3.00 g, 7.05 mmol), HATU (2.95 g, 7.76 mmol), *tert*-butyl carbazate (2.80 g, 21.2 mmol) and diisopropylethylamine (2.46 mL, 14.1 mmol) in anhydrous dimethylformamide (20 mL) were stirred under a nitrogen atmosphere for 18 hours. The solution was then diluted with ethyl acetate (100 mL), and washed with saturated sodium hydrogen carbonate solution (100 mL), 1 M hydrochloric acid (100 mL), brine (3 × 200 mL) and ammonium hydroxide solution (100 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the product as a white solid (3.85 g, 7.13 mmol, 92 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.47 (s, 9 H, H1/^tBu), 1.48 (s, 9 H, H1/^tBu), 2.01 (dq, *J* = 14.0, 8.0 Hz, 1 H, H4), 2.15 (m, 1 H, H4'), 2.46 (m, 2 H, H5), 4.22 (t, *J* = 6.8 Hz, 1 H, H7), 4.32 (m, 1 H, H3), 4.40 (d, *J* = 6.8 Hz, 2 H, H6), 5.92 (br. s, 1 H, NH), 6.63 (br. s, 1 H, NH), 7.32 (t, *J* = 7.2 Hz, 2 H, Fmoc), 7.41 (t, *J* = 7.2 Hz, 2 H, Fmoc), 7.61 (dd, *J* = 7.2, 4.0 Hz, 2 H, Fmoc), 7.78 (d, *J* = 7.2 Hz, 2 H, Fmoc), 8.45 (m, 1 H, NH); ¹³C NMR (75 MHz, DMSO-*d*₆) δ ppm 27.7, 28.0, 31.3, 46.6, 52.3, 65.7, 79.1, 79.7, 120.0, 125.3, 127.0, 127.6, 140.7, 143.7, 143.9, 155.1, 155.8, 170.3, 171.0; IR (solid state, cm⁻¹) 3276 (N-H), 1692 (C=O amide); ESI-MS *m/z* found 562.2529 [M+Na]⁺, [C₂₉H₃₇N₃O₇Na]⁺ requires 562.2529

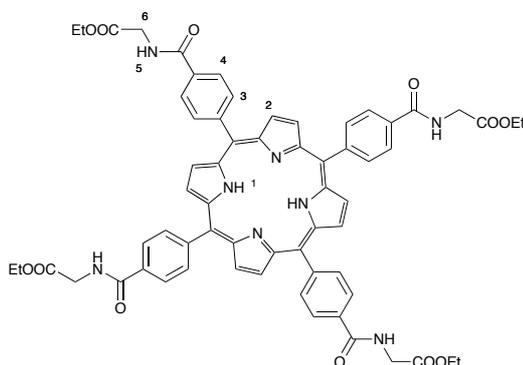
(S)-tert-butyl 2-(2-((9H-fluoren-9-yl) methoxy) carbonylamino) -5- tert- butoxy-5-oxopentanoyl) hydrazinecarboxylate,

62c



(S)-tert-butyl 2-(2-(((9H-fluoren-9-yl) methoxy) carbonylamino) -5- tert- butoxy-5-oxopentanoyl) hydrazinecarboxylate, **61c** (650 mg, 1.20 mmol) in 20 % diethylamine in acetonitrile (50 mL) was stirred for 16 hours. The reaction solution was then concentrated and the resulting residue redissolved in a minimal amount of ethyl acetate, and precipitated by addition of hexane. The slurry was then filtered through celite. The celite pad was then washed with dichloromethane and methanol and this filtrate concentrated to yield the product as an off-white waxy solid (257 mg, 0.811 mmol, 68 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.45 (s, 9 H, H1/^tBu), 1.48 (s, 9 H, H1/^tBu), 1.85 (dq, *J* = 14.0, 7.0 Hz, 1 H, H4), 2.10 (dq, *J* = 14.0, 7.0 Hz, 1 H, H4'), 2.41 (app. t, *J* = 7.0 Hz, 2 H, H5), 3.51 (app. t, *J* = 7.0 Hz, 1 H, H3); ¹³C NMR (75 MHz, DMSO-*d*₆) δ ppm 27.7, 28.0, 30.6, 31.3, 48.5, 52.7, 79.4, 155.2, 172.2, 174.4; IR (solid state, cm⁻¹) 3293 (N-H), 1716 (C=O ester), 1595 (C=O amide); ESI-MS *m/z* found 318.2036 [M+H]⁺, [C₁₄H₂₇N₃O₅]⁺ requires 318.2029

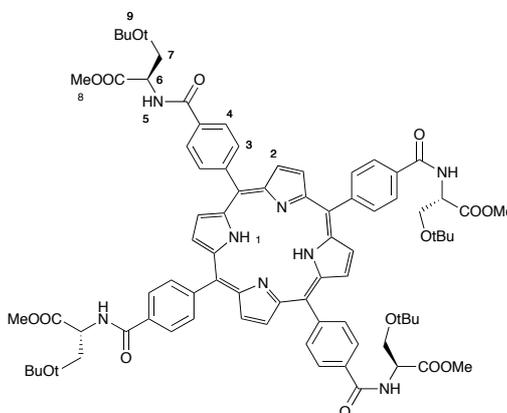
Ethyl 2-({4-[7,12,17-tris ((4-[(2-ethoxy-2-oxoethyl) carbamoyl] phenyl))-21,22,23,24-tetraazapentacyclo [16.2.1.1^{3,6}.1^{8,11}.1^{13,16}] tetracos-1,3,5,7,9,11,13,15,17,19-decaen-2-yl]phenyl}formamido)acetate, **74**



Tetracarboxyphenyl porphyrin (100 mg, 0.126 mmol), PyBOP (394 mg, 0.757 mmol), diisopropylethylamine (0.26 mL, 1.51 mmol) and ethyl glycine hydrochloride salt (106 mg, 0.757 mmol) in anhydrous dimethylformamide (5 mL) were stirred for 16 hours. Methylisocyanate polystyrene resin (0.340g, 0.063 mmol, 200-400 mesh 1.8 mmol/g) was

then added and the reaction mixture stirred for a further 3 hours. The solution was then diluted with dichloromethane (50 mL) and washed successively with saturated sodium hydrogen carbonate solution (100 mL), 1 M hydrochloric acid (100 mL) and brine (100 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a purple solid. This was purified by flash column chromatography (3:7 ethyl acetate:dichloromethane) to yield the product as a purple solid (45 mg, 0.0397 mmol, 32 %); ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm 1.31 (t, *J* = 6.8 Hz, 12 H, Et CH₃), 4.19 (d, *J* = 5.9 Hz, 8 H, H₆), 4.24 (q, *J* = 6.8 Hz, 8 H, Et CH₂), 8.34 (d, *J* = 8.5 Hz, 8 H, H₄), 8.38 (d, *J* = 8.5 Hz, 8 H, H₃), 8.82 - 8.94 (s, 4 H, H₂), 9.34 (t, *J* = 5.9 Hz, 4 H, H₅); IR (solid state, cm⁻¹) 3272 (N-H), 1759 (C=O ester), 1638 (C=O amide); ESI-HRMS found *m/z* 1132.4295 [M]⁺, [C₆₄H₆₀N₈O₁₂]⁺ requires 1132.4331

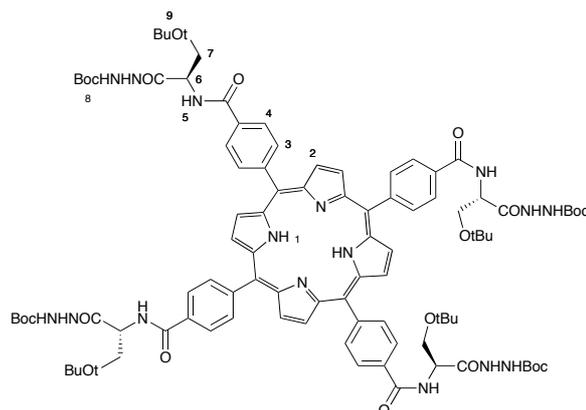
Methyl (2S)-2-({4-[7,12-bis(4-{{(2R)-3-(tert-butoxy)-1-methoxy-1-oxopropan-2-yl} carbamoyl} phenyl)-17-(4-{{(2S)-3-(tert-butoxy)-1-methoxy-1-oxopropan-2-yl} carbamoyl} phenyl)-21,22,23,24-tetraazapentacyclo [16.2.1.1^{3,6}.1^{8,11}.1^{13,16}] tetracos-1,3 (24), 4, 6, 8, 10, 12, 14, 16 (22), 17,19-undecaen-2-yl] phenyl] formamido)-3-(tert-butoxy) propanoate, 76



Tetracarboxy phenyl porphyrin (50 mg, 0.067 mmol), PyBOP (197 mg, 0.402 mmol), *O*-tert-butyl-L-serine methyl ester hydrochloride (85 mg, 0.402 mmol), and diisopropylethylamine (0.14 mL, 0.40 mmol) in anhydrous dimethylformamide (5 mL) were stirred under a nitrogen atmosphere for 18 hours. The reaction mixture was then dissolved in ethyl acetate (50 mL) and washed with saturated sodium hydrogen carbonate solution (50 mL), 1 M hydrochloric acid (50 mL) and brine (3 × 100 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the product as a purple solid (97 mg, 0.076 mmol, quant.); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.18 (s, 36 H, H₉), 3.77 -

3.82 (m, 8 H, H8 + H7), 3.98 (dd, $J = 9.2, 2.7$ Hz, 4 H, H7'), 5.05 (dt, $J = 8.3, 2.7$ Hz, 4 H, H6), 7.27 (d, $J = 8.3$ Hz, 4 H, H5), 8.17 (d, $J = 8.0$ Hz, 8 H, H4), 8.22 - 8.29 (d, $J = 8.0$ Hz, 8 H, H3), 8.78 (s, 8 H, H2); ^{13}C NMR (75 MHz, CDCl_3) δ ppm 27.5, 46.3, 52.6, 62.2, 73.7, 119.4, 125.7, 128.5, 131.2, 133.6, 134.7, 145.5, 167.2, 171.2; IR (solid state, cm^{-1}) 3314 (N-H), 1741 (C=O ester), 1656 (C=O amide); ESI-MS m/z found 1419.6575 $[\text{M}+\text{H}]^+$, $[\text{C}_{80}\text{H}_{90}\text{N}_8\text{O}_{16}]^+$ requires 1419.6553

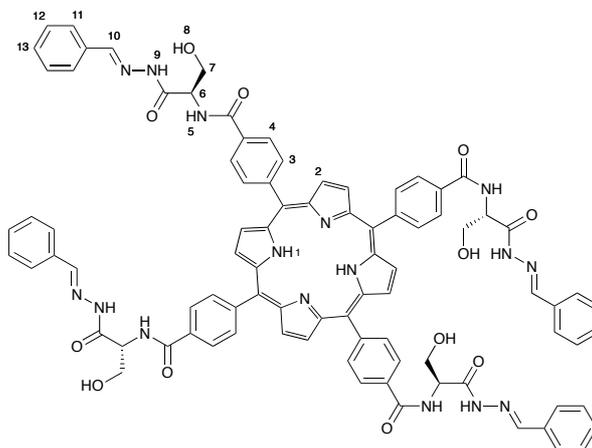
4-[7,12-Bis(4-[[[(1R)-2-(*tert*-butoxy)-1-{N'-[(*tert*-butoxy) carbonyl] hydrazine carbonyl] ethyl] carbamoyl] phenyl)-17-(4-[[[(1S)-2-(*tert*-butoxy)-1-{N'-[(*tert*-butoxy) carbonyl] hydrazinecarbonyl] ethyl] carbamoyl] phenyl)-21,22,23,24-tetraazapentacyclo [16.2.1.1³,6.1⁸,11.1¹³,16] tetracos-1,3(24), 4,6,8,10,12,14,16(22), 17,19-undecaen-2-yl]-N-[(1S)-2-(*tert*-butoxy)-1-{N'-[(*tert*-butoxy)carbonyl] hydrazine carbonyl] ethyl]benzamide, 78a



Tetracarboxyphenyl porphyrin (100 mg, 0.126 mmol), PyBOP (395 mg, 0.759 mmol), *tert*-butyl 2-(2-amino-3-*tert*-butoxypropanoyl)hydrazinecarboxylate, **62a** (209 mg, 0.759 mmol) and diisopropylethylamine (0.26 mL, 1.5 mmol) in anhydrous dimethylformamide (5 mL) were stirred under a nitrogen atmosphere for 18 hours. The reaction mixture was then diluted with ethyl acetate (50 mL) and washed with saturated sodium hydrogen carbonate solution (50 mL), 1 M hydrochloric acid (50 mL) and brine (3 x 50 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a purple solid. This was purified by flash column chromatography (1:1 ethyl acetate:dichloromethane then ethyl acetate) to yield the product as a purple solid (142 mg, 0.078 mmol, 62%); ^1H NMR (500 MHz, CDCl_3) δ ppm 1.29 (s, 36 H, H8/H9), 1.45 (m, 36 H, H8/H9), 3.51 - 3.63 (m, 8 H, H7 + H7'), 4.07 (dt, $J = 8.3, 4.7$ Hz, 4 H, H6), 4.83 (br. s, 4 H, NH), 6.40 - 6.57 (m, 4 H, NH), 8.15 (d, $J = 7.5$ Hz, 8 H, H4), 8.21 (d, $J = 7.5$ Hz, 8 H, H3),

8.40 (d, $J = 8.3$ Hz, 4 H, H5), 8.73 (br. s, 8 H, H2); ^{13}C NMR (101 MHz, MeOD) δ ppm 26.8, 27.5, 38.4, 53.3, 61.6, 74.0, 80.8, 106.6, 119.3, 125.8, 133.4, 134.4, 145.2, 156.2, 168.4, 170.8; IR (solid state, cm^{-1}) 3251 (N-H), 1696 (C=O carbamate), 1645 (C=O amide); ESI-MS m/z found 1820.9402 $[\text{M}+2\text{H}]^+$, $[\text{C}_{96}\text{H}_{124}\text{N}_{16}\text{O}_{20}]^+$ requires 1820.9178

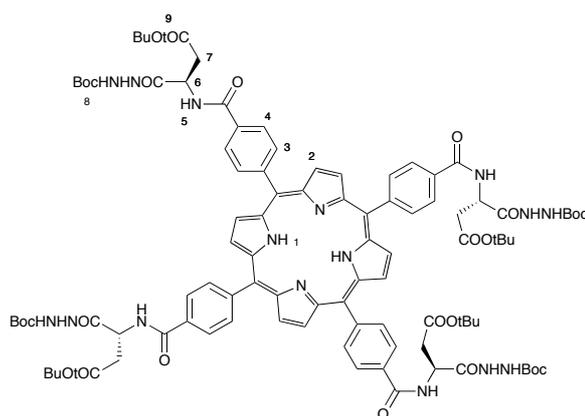
4- [7,12-Bis (4- {[(1R) -2- hydroxy-1- {N'- [(1E)- phenylmethylidene] hydrazinecarbonyl} ethyl] carbamoyl} phenyl)-17-(4-{[(1S)-2-hydroxy-1-{N'-[(1E)- phenylmethylidene] hydrazinecarbonyl} ethyl] carbamoyl} phenyl)-21,22,23,24-tetraazapentacyclo [16.2.1.1^{3,6}.1^{8,11}.1^{13,16}] tetracos-1,3(24),4,6,8,10,12,14,16(22), 17,19- undecaen-2-yl] -N-[(1S)-2- hydroxy-1- {N'- [(1E)-phenylmethylidene] hydrazinecarbonyl} ethyl]benzamide, **79a**



4-[7,12-bis(4-[[[1R)-2-(*tert*-butoxy)-1-{N'-[[*tert*-butoxy]carbonyl]hydrazinecarbonyl} ethyl] carbamoyl} phenyl)-17-(4- [[[1S)-2- (*tert*-butoxy) -1-{N'-[[*tert*-butoxy] carbonyl] hydrazine carbonyl} ethyl] carbamoyl]phenyl)- 21,22,23,24-tetraazapentacyclo [16.2.1. 1^{3,6}.1^{8,11}.1^{13,16}] tetracos- 1,3(24), 4,6,8,10,12,14,16(22), 17,19-undecaen-2-yl]-N-[(1S)-2- (*tert*-butoxy)-1-{N'-[[*tert*-butoxy]carbonyl]hydrazinecarbonyl}ethyl]benzamide, **78a** (142 mg, 0.0780 mmol) in trifluoroacetic acid (4.5 mL), water (0.25 mL) and triisopropylsilane (0.25 mL) was stirred for 6 hours. The solution was then concentrated to yield the deprotected hydrazide as a green solid (complete deprotection was confirmed by HRMS). The green solid was then redissolved in water (2 mL) and acetonitrile (2 mL) and benzaldehyde (2 drops) was added, the mixture was then stirred for 30 minutes. The precipitate was isolated and washed with water and acetonitrile to yield the product as a dark purple solid (30 mg, 0.019 mmol, 25 %); ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ ppm 1.67 (br. s, 4 H, H8), 3.94 (br. s, 8 H, H7), 4.76 (br. s, 4 H, NH), 5.63 (br. s, 4 H, H6), 7.51 (m, 8 H,

H4), 7.63 (m, 8 H, H12), 7.75(m, 4 H, H11), 7.95 (br. s, 4 H, H11'), 8.12 (br. s, 4 H, H2), 8.39 (br. s, 8 H, H3), 8.84 (m, 2 H, H13), 8.91 (br. s, 10 H, H12, H13'), 11.56 (br. s, 4 H, H10), 11.73 (br. s, 4 H, H10') (*cis* and *trans* hydrazone isomers observed in 1:1 ratio); IR (solid state, cm⁻¹) 3212 (N-H), 1633 (C=O amide), 1608 (C=N); ESI-MS *m/z* found 1547.5750 [M+H]⁺, [C₈₈H₇₅N₁₆O₁₂]⁺ requires 1547.5340

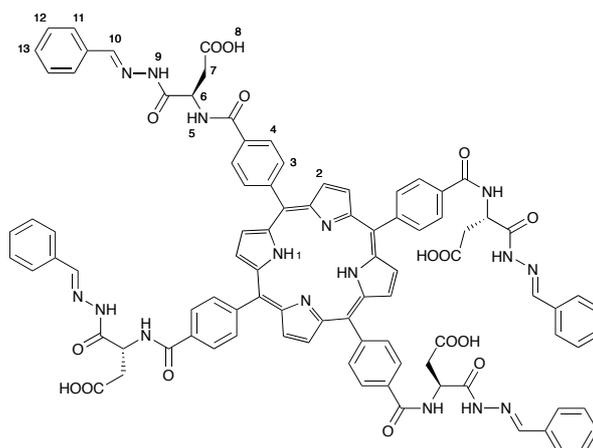
Tert-butyl (3S)-3-((4- [7,12-bis (4-((1R)-3- (tert-butoxy)-1- {N'- [(tert-butoxy)carbonyl] hydrazinecarbonyl}-3-oxopropyl] carbamoyl}phenyl)-17-(4- [[(1S) -3- (tert-butoxy) -1-{N'-[(tert-butoxy) carbonyl] hydrazinecarbonyl} -3-oxopropyl] carbamoyl} phenyl) -21,22,23,24- tetraazapentacyclo [16.2.1.1.^{3,6}.1.^{8,11}.1.^{13,16}]tetracos-1,3(24),4,6,8,10, 12,14,16(22),17,19-undecaen-2-yl] phenyl} formamido)- 3- {N'-[(tert-butoxy)carbonyl] hydrazinecarbonyl} propanoate, 78b



Tetracarboxyphenyl porphyrin (43 mg, 0.055 mmol), PyBOP (172 mg, 0.330 mmol), (*S*)-*tert*-butyl 2-(2-amino-4-*tert*-butoxy-4-oxobutanoyl)hydrazinecarboxylate, **62b** (100 mg, 0.330 mmol) and anhydrous diisopropylethylamine (0.11 mL, 0.66 mmol) in anhydrous dimethylformamide (5 mL) were stirred under a nitrogen atmosphere for 18 hours. The reaction mixture was then diluted with ethyl acetate (50 mL) and washed with saturated sodium hydrogen carbonate solution (50 mL), 1 M hydrochloric acid (50 mL) and brine (3 × 50 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a purple solid. This was purified by flash column chromatography (1:1 ethyl acetate:dichloromethane then ethyl acetate) to yield the product as a purple solid (103 mg, 0.0533 mmol, 97 %); ¹H NMR (500 MHz, MeOD) δ ppm 1.53 (s, 36 H, H8/H9), 1.55 (s, 36 H, H8/H9), 2.95 (dd, *J* = 17.4, 8.0 Hz, 4 H, H7), 3.10 (dd, *J* = 17.4, 4.8 Hz, 4 H, H7'), 5.30 (br. s, 4 H, H6), 5.45 (d, *J* = 1.4 Hz, 4 H, NH hydrazone), 8.21

(br. s, 8 H, H2), 8.27 (d, $J = 6.7$ Hz, 8 H, H3), 8.73 (br. s, 8 H, H5), 9.02 (d, $J = 6.7$ Hz, 8 H, H4); ^{13}C NMR (101 MHz, MeOD + 10% CDCl_3) δ ppm 27.4, 27.6, 37.2, 49.3, 80.9, 81.5, 119.3, 125.9, 133.2, 134.4, 145.4, 149.5, 156.3, 168.4, 170.2, 171.3, 185.1; IR (solid state, cm^{-1}) 3275 (N-H), 1723 (C=O ester), 1711 (C=O carbamate), 1647 (C=O amide); ESI-MS m/z found 1931.9891 $[\text{M}+\text{H}]^+$, $[\text{C}_{100}\text{H}_{123}\text{N}_{16}\text{O}_{24}]^+$ requires 1931.8896

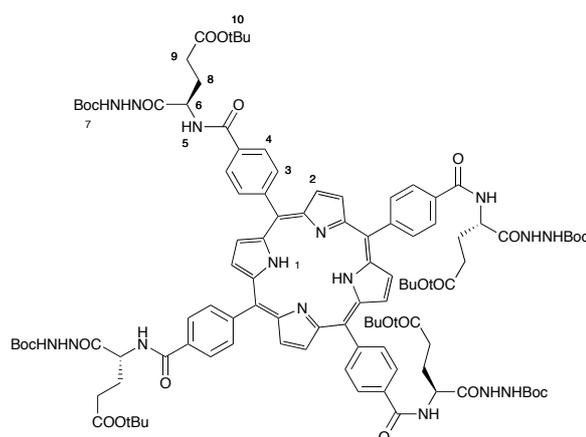
3(S) -3- ({4- [7,12-bis (4- {[(1R) -2- carboxy-1- {N'- [(1E) -phenylmethylidene] hydrazinecarbonyl} ethyl] carbamoyl} phenyl)-17- (4- {[(1S)-2-carboxy-1- {N'-[(1E)-phenylmethylidene] hydrazinecarbonyl}ethyl] carbamoyl} phenyl)-21,22,23,24-tetraazapentacyclo[16.2.1.1^{3,6}.1^{8,11}.1^{13,16}]tetracos-1,3(24),4,6,8,10,12,14,16(22), 17,19- undecaen-2-yl] phenyl} formamido) -3- {N'- [(1E)- phenylmethylidene] hydrazinecarbonyl} propanoic acid, 79b



Tert-butyl (3S)-3-({4-[7,12-bis(4-{[(1R)-3-(*tert*-butoxy)-1-{N'-[(*tert*-butoxy) carbonyl] hydrazinecarbonyl}-3-oxopropyl]carbamoyl}phenyl)-17-(4-{[(1S)-3-(*tert*-butoxy)-1-{N'-[(*tert*-butoxy)carbonyl]hydrazinecarbonyl}-3-oxopropyl]carbamoyl} phenyl)-21,22,23,24-tetraaza pentacyclo[16.2.1.1^{3,6}.1^{8,11}.1^{13,16}]tetracos-1,3(24), 4,6,8,10,12,14,16(22),17,19-undecaen-2-yl] phenyl} formamido) -3-{N'- [(*tert*-butoxy) carbonyl] hydrazinecarbonyl} propanoate (50 mg, 0.026 mmol) in trifluoroacetic acid (4.5 mL), water (0.25 mL) and triisopropylsilane (0.25 mL) was stirred for 6 hours. The solution was then concentrated to yield deprotected hydrazide as a green solid (complete deprotection was confirmed by HRMS). The green solid was then redissolved in water (2 mL) and acetonitrile (2 mL) and benzaldehyde (2 drops) added, the mixture was then stirred for 30 minutes. The precipitate was isolated and washed with water and acetonitrile to yield the product as a dark purple solid (16 mg, 0.0096 mmol, 37 %); ^1H NMR (500 MHz, DMSO-d_6) δ ppm -2.98 -

-2.85 (br. s, 2 H, H1), 2.78 - 3.06 (m, 8 H, H7 + H7'), 5.06 (m, 4 H, H6), 5.87 (m, 1 H, NH), 7.46 (d, $J = 6.7$ Hz, 8 H, H3), 7.73 (d, $J = 6.8$ Hz, 4 H, H11), 7.76 (d, $J = 6.8$ Hz, 4 H, H11'), 8.06 (m, 4 H, H2), 8.34 (m, 16 H, H2, H4), 8.85 (br. s, 8 H, H12 + H12'), 9.08 (m, 2 H, H13), 9.14 - 9.23 (m, 2 H, H13'), 11.54 (br. s, 2 H, H10), 11.67 (br. s, 2 H, H10'), 12.46 (br. s, 4 H, OH) (*cis* and *trans* isomers of hydrazone present in 1:1 ratio); IR (solid state, cm^{-1}) 3270 (N-H), 1714 (C=O acid), 1643 (C=O amide), 1607 (C=N); ESI-MS m/z found 1659.5058 $[\text{M}+\text{H}]^+$, $[\text{C}_{92}\text{H}_{75}\text{N}_{16}\text{O}_{16}]^+$ requires 1659.5547

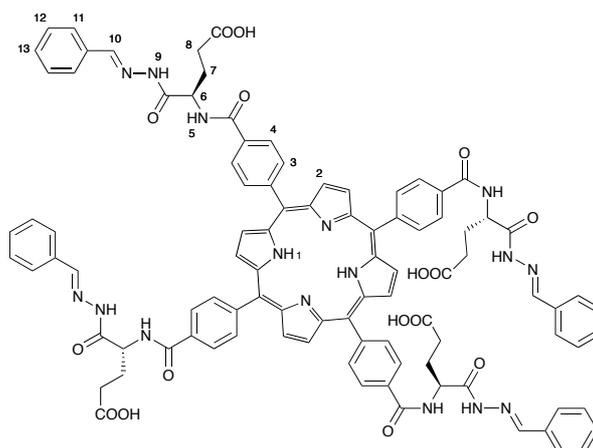
***Tert*-butyl (4S) -4- ((4- [7,12- bis (4- (((1R)-4- (*tert*-butoxy) -1- {N'- [(*tert*-butoxy)carbonyl] hydrazinecarbonyl)-4-oxobutyl]carbamoyl}phenyl)-17-(4-(((1S)-4- (*tert*-butoxy) -1- {N'- [(*tert*-butoxy) carbonyl] hydrazinecarbonyl} -4- oxobutyl] carbamoyl} phenyl) -21,22,23,24- tetraazapentacyclo [16.2.1.1^{3,6}.1^{8,11}.1^{13,16}] tetracos-1,3(24),4,6,8,10,12,14,16(22), 17,19-undecaen-2-yl] phenyl) formamido)-4-{N'-[(*tert*-butoxy)carbonyl] hydrazine carbonyl} butanoate, 78c**



Tetracarboxy phenyl porphyrin (42 mg, 0.053 mmol), PyBOP (164 mg, 0.315 mmol), (*S*)-*tert*-butyl 2-(2-amino-5-*tert*-butoxy-5-oxopentanoyl)hydrazinecarboxylate, **62c** (100 mg, 0.315 mmol) and diisopropylethylamine (0.11 mL, 0.63 mmol) in anhydrous dimethylformamide (5 mL) were stirred under nitrogen for 18 hours. The reaction mixture was then diluted with ethyl acetate (50 mL) and washed with saturated sodium hydrogen carbonate solution (50 mL), 1 M hydrochloric acid (50 mL) and brine (3 x 50 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a purple solid. This was purified by flash column chromatography (1:1 ethyl acetate:dichloromethane then ethyl acetate) to yield the product as a purple solid (65 mg, 0.033 mmol, 62 %); ^1H NMR (500 MHz, MeOD) δ ppm 1.51 (s, 36 H, H7/H10), 1.54 (s, 36 H, H7/H10), 2.34 (m, 8 H, H8), 2.63 (m, 8 H, H9), 3.35 (br. s, 4 H, NH hyd), 3.39 (s, 4 H, NH

hyd), 4.92 (br. s, 4 H, H6), 8.00 (br. s, 8 H, H3), 8.18 (br. s, 8 H, H4), 8.61 (m, 8 H, H2); ¹³C NMR (101 MHz, MeOD + 10% CDCl₃) δ ppm 27.1, 27.5, 27.7, 31.7, 51.9, 81.0, 119.4, 125.9, 133.2, 134.5, 145.4, 149.5, 156.2, 167.5, 168.4, 170.3, 200.0; IR (solid state, cm⁻¹) 3275 (N-H), 1723 (C=O ester), 1643 (C=O carbamate), 1608 (C=O amide); ESI-MS *m/z* found 1988.0341 [M+H]⁺, [C₁₀₄H₁₃₁N₁₆O₂₄]⁺ requires 1987.9522

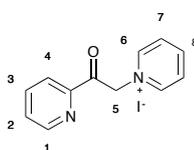
(4S)-4- ((4- [7,12-bis (4-(((1R)- 3-carboxy-1- {N'- [(1E)- phenylmethylidene] hydrazine carbonyl)propyl]carbamoyle)phenyl)-17-(4-(((1S)-3-carboxy-1-{N'-[(1E)- phenyl methylidene]hydrazinecarbonyl)propyl]carbamoyle)phenyl)-21,22,23,24-tetraaza pentacyclo[16.2.1.1.3³,6.1⁸,11.1¹³,16]tetracos-1,3(24),4,6,8,10,12,14,16(22), 17,19-undecaen-2-yl] phenyl} formamido) -4-{N'- [(1E)- phenylmethylidene] hydrazinecarbonyl} butanoic acid, 79c



Tert-butyl (4S) -4- ((4- [7, 12-bis (4-(((1R)-4- (*tert*- butoxy)-1- {N'-[(*tert*- butoxy) carbonyl] hydrazinecarbonyl)-4-oxobutyl] carbamoyle) phenyl) -17- (4- (((1S)-4- (*tert*- butoxy)-1-{N'-[(*tert*-butoxy)carbonyl]hydrazinecarbonyl)-4-oxobutyl]carbamoyle)phenyl)-21,22,23,24-tetraazapentacyclo[16.2.1.1.3³,6.1⁸,11.1¹³,16]tetracos-1,3(24),4,6,8,10,12,14, 16 (22), 17,19-undecaen-2-yl] phenyl} formamido)-4-{N'-[(*tert*-butoxy)carbonyl] hydrazine carbonyl} butanoate, **78b** (30 mg, 0.015 mmol) in trifluoroacetic acid (4.5 mL), water (0.25 mL) and trisopropylsilane (0.25 mL) was stirred for 6 hours. The solution was then concentrated to yield the deprotected hydrazide porphyrin as a green solid (complete deprotection was confirmed by HRMS). The green solid was then redissolved in water (2 mL) and acetonitrile (2 mL) and benzaldehyde (2 drops) was added, the mixture was then stirred for 30 minutes. The precipitate was isolated and washed with water and acetonitrile to yield the product as a dark purple solid (25 mg, 0.0093 mmol, 62 %); ¹H

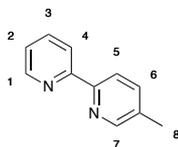
NMR (500 MHz, DMSO- d_6) δ ppm -2.94 (br. s, 2 H, H1), 0.79 (m, 4 H, H7), 1.36 (m, 4 H, H7'), 2.07 (m, 4 H, H8), 4.59 (br. s, 4 H, H6), 5.48 (br. s, 4 H, NH), 7.45 (br. s, 8 H, H3), 7.56 - 7.67 (m, 4 H, H12), 7.73 (m, 4 H, H12'), 7.82 (m, 4 H, H11), 7.87 - 7.98 (m, 4 H, H11'), 8.09 (br. s, 4 H, H2), 8.36 (br. s, 8 H, H4), 8.83 (br. s, 8 H, H3), 8.94 (m, 2 H, H13), 9.04 (br. s, 2 H, H13'), 11.52 (m, 2 H, H10), 11.68 (m, 2 H, H10'), 12.33 (br. s, 4 H, OH) (*cis* and *trans* isomers of hydrazone observed in 1:1 ratio); IR (solid state, cm^{-1}) 3283 (N-H), 1633 (C=O acid), 1607 (C=O amide), 1529 (C=N); ESI-MS m/z found 1715.5801 $[\text{M}+\text{H}]^+$, $[\text{C}_{96}\text{H}_{83}\text{N}_{16}\text{O}_{16}]^+$ requires 1715.6173

1-[2-oxo-2-(pyridin-2-yl)ethyl]pyridin-1-ium iodide, **83**¹⁴³



Pyridine (11.5 mL, 0.142 mmol), 2-acetyl pyridine (5 mL, 0.044 mmol) and iodine (3.39 g, 0.0133 mmol) were stirred at 80 °C for 4 hours. The solution was then filtered and the precipitate washed with pyridine. The grey precipitate was then boiled in ethanol with activated charcoal and hot filtered. The filtrate was concentrated to yield the product as a golden solid (6.56 g, 0.0201 mmol, 46 %); ^1H NMR (500 MHz, CDCl_3) δ ppm 7.03 (s, 2 H, H5), 7.64 - 7.69 (m, 1 H, H4), 7.97 (td, $J = 7.9, 1.6$ Hz, 1 H, H3), 8.12 (d, $J = 7.9$ Hz, 1 H, H1), 8.19 (dd, $J = 5.9, 4.8$ Hz, 2 H, H7), 8.61 (t, $J = 7.9$ Hz, 1 H, H2), 8.82 (t, $J = 4.8$ Hz, 1 H, H8), 9.26 (d, $J = 5.9$ Hz, 2 H, H6); ^{13}C NMR (126 MHz, DMSO- d_6) δ ppm 48.6, 66.3, 122.0, 127.7, 129.1, 138.1, 146.3, 149.5, 150.4, 191.4; IR (solid state, cm^{-1}) 1709 (C=O); ESI-MS m/z found 199.0899 $[\text{M}]^+$, $[\text{C}_{12}\text{H}_{11}\text{N}_2\text{O}]^+$ requires 199.0871

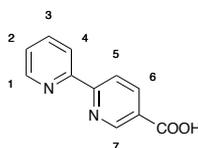
5-methyl-2-(pyridin-2-yl)pyridine, **85**¹⁴³



1-[2-oxo-2-(pyridin-2-yl)ethyl]pyridin-1-ium iodide, **83** (5.00 g, 15.3 mmol), methacrolein (1.26 mL, 15.3 mmol) and ammonium acetate (4.72 g, 61.2 mmol) in methanol (50 mL) were stirred at 65 °C for 16 hours. The solution was diluted with dichloromethane (50 mL) and washed with water; the aqueous phase was extracted with dichloromethane (2 \times 50 mL). The combined organic phases were washed with brine (50

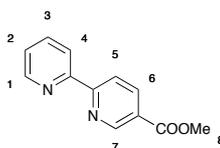
mL), dried (sodium sulfate) and concentrated to yield the crude product as a brown oil. This was purified by flash column chromatography (3:7 ethyl acetate:dichloromethane) to yield the product as a pale yellow oil (1.80 g, 10.6 mmol, 69 %); ^1H NMR (500 MHz, CDCl_3) δ ppm 2.37 (s, 3 H, H8), 7.27 (ddd, $J = 7.5, 4.7, 1.2$ Hz, 1 H, H2), 7.61 (dd, $J = 7.5, 1.2$ Hz, 1 H, H4), 7.78 (td, $J = 7.5, 1.7$ Hz, 1 H, H3), 8.27 (d, $J = 8.0$ Hz, 1 H, H5), 8.34 (d, $J = 8.0$ Hz, 1 H, H6), 8.50 (s, 1 H, H7), 8.66 (d, $J = 4.7$ Hz, 1 H, H1); ^{13}C NMR (126 MHz, CDCl_3) δ ppm 21.2, 121.4, 122.1, 123.7, 124.8, 127.1, 148.4, 148.9, 149.1, 155.8, 156.2; IR (solid state, cm^{-1}) 1589 (aromatic C=C), 1575 (aromatic C=C), 1558 (aromatic C=C); ESI-HRMS found m/z 171.0920 $[\text{M}+\text{H}]^+$, $[\text{C}_{11}\text{H}_{11}\text{N}_2]^+$ requires 171.0922

6-(Pyridin-2-yl)pyridine-3-carboxylic acid, **69**¹⁴³



5-Methyl-2-(pyridin-2-yl)pyridine, **85** (750 mg, 4.41 mmol) and potassium permanganate (2.79 g, 17.6 mmol) in water were heated under reflux for 16 hours. The solution was then cooled to room temperature and filtered through celite. The filtrate was made basic by addition of 1 M sodium hydroxide solution, and washed with dichloromethane (100 mL). The aqueous phase was acidified by addition of acetic acid, and extracted with dichloromethane (5×100 mL). The combined organic phases were dried (sodium sulfate) and concentrated, to yield the product as a pale pink solid (547 mg, 3.21 mmol, 73 %); ^1H NMR (500 MHz, MeOD) δ ppm 7.51 (ddd, $J = 7.8, 5.0, 0.9$ Hz, 1 H, H2), 8.00 (td, $J = 7.8, 1.7$ Hz, 1 H, H3), 8.43 - 8.51 (m, 3 H, H1+H4+H5), 8.72 (d, $J = 4.4$ Hz, 1 H, H6), 9.25 (s, 1 H, H7); ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ ppm 120.2, 121.2, 124.9, 127.1, 137.5, 138.1, 149.5, 150.1, 154.3, 158.2, 166.2; IR (solid state, cm^{-1}) 2544 (O-H), 1680 (C=O acid); ESI-HRMS found m/z 201.0661 $[\text{M}+\text{H}]^+$, $[\text{C}_{11}\text{H}_8\text{N}_2\text{O}_2]^+$ requires 201.0664

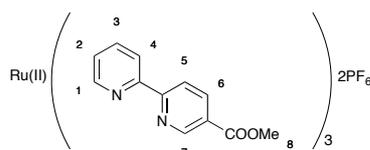
Methyl 6-(pyridin-2-yl)pyridine-3-carboxylate, **86**



6-(Pyridin-2-yl)pyridine-3-carboxylic acid, **86** (500 mg, 2.50 mmol), anhydrous methanol (30 mL) and thionyl chloride (2.50 mL) were heated under reflux for 16 hours

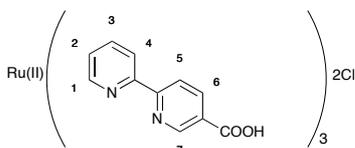
under a nitrogen atmosphere. The reaction mixture was then cooled to room temperature and quenched with saturated sodium hydrogen carbonate solution (100 mL). The mixture was extracted with dichloromethane (100 mL) and the organic phase washed with 1 M hydrochloric acid (50 mL) and brine (50 mL), dried (sodium sulfate) and concentrated to yield the product as a pale pink solid (425 mg, 1.99 mmol, 79 %); ^1H NMR (500 MHz, CDCl_3) δ ppm 3.98 (s, 3H, H8), 7.40 (dd, $J = 7.4, 5.2$ Hz, 1 H, H2), 7.89 (td, $J = 7.4, 1.4$ Hz, 1 H, H3), 8.44 (dd, $J = 8.3, 2.0$ Hz, 1 H, H6), 8.52 (d, $J = 7.4$ Hz, 1 H, H4), 8.55 (d, $J = 8.3$ Hz, 1 H, H5), 8.75 (br. s, 1 H, H1), 9.30 (br. s, 1 H, H7); ^{13}C NMR (126 MHz, CDCl_3) δ ppm 52.4, 117.2, 120.6, 124.5, 125.7, 137.1, 138.1, 149.4, 150.5, 155.1, 159.5, 165.9; IR (solid state, cm^{-1}) 1716 (C=O ester); ESI-HRMS found m/z 215.0821 $[\text{M}+\text{H}]^+$, $[\text{C}_{12}\text{H}_{10}\text{N}_2\text{O}_2]^+$ requires 215.0821

***Tris* (methyl 6- (pyridin-2-yl) pyridine -3- carboxylate) ruthenium(II) dihexafluorophosphate, 87**



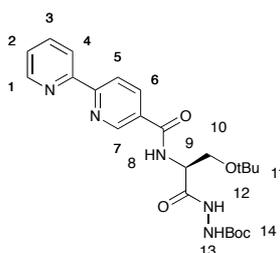
Methyl 6-(pyridin-2-yl)pyridine-3-carboxylate, **86** (250 mg, 1.17 mmol), $\text{Ru(II)(DMSO)}_4\text{Cl}_2$ (177 mg, 0.364 mmol), and silver nitrate (171 mg, 0.728 mmol) in ethanol (20 mL) were heated under reflux for 6 days. The reaction mixture was then filtered through celite, and the celite pad washed thoroughly with dichloromethane. The filtrate was then concentrated and the resulting red solid redissolved in water. An excess of ammonium hexafluorophosphate was added and the red precipitate isolated to yield the product as a red solid (308 mg, 0.298 mmol, 82 %); ^1H NMR (500 MHz, DMSO-d_6) δ ppm 3.18 (s, 4.5 H, H8), 3.80 (s, 4.5 H, H8), 7.60 (m, 3 H, H2), 7.82 (m, 3 H, H3), 7.05 (m, 3 H, H6), 8.24 (m, 3 H, H4), 8.55 (m, 3 H, H5), 8.95 (m, 6 H, H1+H7) (*fac* and *mer* isomers seen on Me in 1:1 ratio); IR (solid state, cm^{-1}) 1717 (C=O ester); ESI-MS m/z found 372.0635 $[\text{M}]^{2+}$, $[\text{C}_{36}\text{H}_{30}\text{N}_6\text{O}_6\text{Ru}]^{2+}$ requires 372.0635

***Tris* (6-(pyridin-2-yl)pyridine-3-carboxylic acid) ruthenium(II) dichloride, 88**



Tris (methyl 6-(pyridin-2-yl) pyridine-3-carboxylate) ruthenium(II) dihexafluorophosphate, **87** (400 mg, 0.387 mmol) was stirred in 1 M sodium hydroxide solution and ethanol for 2 hours. The reaction mixture was then concentrated and redissolved in water. The resulting solution was neutralized by addition of 1 M hydrochloric acid and concentrated. The salt was removed by dissolving the red solid in repeatedly in a minimal amount of methanol and filtering the precipitated salt, until no salt was visible, to yield the product as a red solid (298 mg, 0.387 mmol, quant.); ¹H NMR (500 MHz, D₂O) δ ppm 7.24 - 7.41 (m, 3 H, H2), 7.73 - 7.83 (m, 3 H, H3), 7.88 (s, 3 H, H6), 7.95 - 8.06 (m, 3 H, H4), 8.17 (br. ss, 3 H, H5), 8.25 - 8.39 (m, 3 H, H1), 8.48 - 8.57 (m, 3 H, H7); IR (solid state, cm⁻¹) 3392 (O-H), 1614 (C=O acid); ESI-MS *m/z* found 351.0410 [M]²⁺, [C₃₃H₂₄N₆O₆Ru]²⁺ requires 351.0400

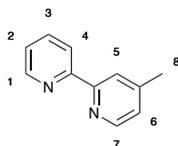
N-[(1S)-2-(*tert*-butoxy)-1-{N'-[(*tert*-butoxy)carbonyl]hydrazinecarbonyl}ethyl]-[2,2'-bipyridine]-5-carboxamide, **70**



6-(pyridin-2-yl)pyridine-3-carboxylic acid (200 mg, 1.00 mmol) in thionyl chloride (10 mL) was heated under reflux for 16 hours. The solvent was then removed *in vacuo* and the acid chloride flushed with nitrogen and used immediately. The acid chloride was redissolved in anhydrous chloroform (20 mL) under a nitrogen atmosphere and *tert*-butyl 2-(2-amino-3-*tert*-butoxypropanoyl)hydrazinecarboxylate, **62a** (303 mg, 1.10 mmol) and anhydrous diisopropylethylamine (0.19 mL, 1.10 mmol) were added. The reaction mixture was then heated under reflux for 16 hours. The reaction mixture was then washed with saturated sodium hydrogen carbonate solution (50 mL), 1 M hydrochloric acid (50 mL) and brine (50 mL). The organic phase was dried (sodium sulfate) and concentrated to yield the crude product as a pale pink solid. This was purified by flash column chromatography (10 % methanol in dichloromethane) to yield the product as a pink solid (30 mg, 0.066 mmol, 7 %); ¹H NMR (500 MHz, CDCl₃) δ ppm 1.30 (s, 9 H, H11/H14), 1.50 (s, 9 H, H11/H14), 2.06 (m, 1 H, H10), 2.36 (t, *J* = 8.5 Hz, 1 H, H10'), 4.00 (dd, *J* = 8.5, 4.5 Hz, 1 H, H9), 4.72 (br. s, 1 H, NH), 6.62 (br. s, 1 H, NH), 7.43 (m, 2 H, H5 + H6), 7.93 (t, *J* = 6.9 Hz, 1 H, H4), 8.29 (d, *J* = 8.2 Hz, 1 H, H1), 8.51 (dd, *J* = 14.5, 8.2 Hz, 1 H, H2), 8.77 (dd, *J* =

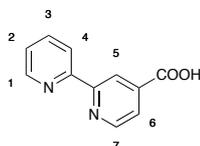
14.5, 6.9 Hz, 1 H, H3), 9.16 (s, 1 H, H7); IR (solid state, cm^{-1}) 3273 (N-H), 1699 (C=O amide), 1651 (C=O amide); ESI-MS m/z found 458.2404 $[\text{M}+\text{H}]^+$, $[\text{C}_{23}\text{H}_{32}\text{N}_5\text{O}_5]^+$ requires 458.2403

4-Methyl-2-(pyridin-2-yl)pyridine, **89**



1-[2-Oxo-2-(pyridin-2-yl)ethyl]pyridin-1-ium iodide, **83** (4.04 g, 12.4 mmol), crotonaldehyde (0.97 mL, 12.4 mmol) and ammonium acetate (3.82 g, 49.6 mmol) in methanol (50 mL) were stirred at 65 °C for 16 hours. The solution was concentrated and the orange solid diluted with petroleum ether (100 mL) and washed with water (100 mL), and the aqueous phase extracted with petroleum ether (2×50 mL). The combined organic phases were washed with brine (50 mL), dried (sodium sulfate) and concentrated to yield the crude product as a yellow solid (501 mg, 2.95 mmol, 24 %), the crude product was used without further purification; ^1H NMR (500 MHz, CDCl_3) δ ppm 2.35 (s, 3 H, H8), 7.04 (d, $J = 4.7$ Hz, 1 H, H6), 7.18 - 7.24 (m, 1 H, H2), 7.72 (td, $J = 7.9, 1.8$ Hz, 1 H, H3), 8.15 (s, 1 H, H5), 8.31 (d, $J = 7.9$ Hz, 1 H, H4), 8.45 (d, $J = 4.9$ Hz, 1 H, H1), 8.59 (d, $J = 4.7$ Hz, 1 H, H7); ^{13}C NMR (126 MHz, CDCl_3) δ ppm 21.2 121.3, 121.9, 123.7, 124.7, 137.0, 148.1, 149.1, 149.1, 156.0, 156.4; IR (solid state, cm^{-1}) 1744 (aromatic C=C), 1602 (aromatic C=C), 1582 (aromatic C=C); ESI-HRMS found m/z 193.0789 $[\text{M}+\text{Na}]^+$, $[\text{C}_{11}\text{H}_{10}\text{N}_2\text{Na}]^+$ requires 193.0742

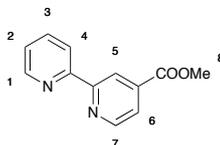
2-(Pyridin-2-yl)pyridine-4-carboxylic acid, **90**



5-Methyl-2-(pyridin-2-yl)pyridine, **89** (450 mg, 2.64 mmol) and potassium permanganate (1.67 g, 10.6 mmol) in water (50 mL) were heated under reflux for 16 hours. The solution was then cooled to room temperature, filtered through celite and the filtrate concentrated, to yield the product as a white solid (415 mg, 2.07 mmol, 79 %); ^1H NMR (500 MHz, MeOD) δ ppm 7.38 (t, $J = 6.9$ Hz, 1 H, H2), 7.84 (d, $J = 5.0$ Hz, 1 H, H6), 7.87 (t, $J = 6.9$ Hz, 2 H, H3), 8.30 (d, $J = 6.9$ Hz, 2 H, H4), 8.59 (d, $J = 6.9$ Hz, 1 H, H1), 8.73 (d, $J = 5.0$ Hz, 1 H, H7), 8.76 (s, 1 H, H5); ^{13}C NMR (126 MHz, D_2O) δ ppm 120.7, 122.0, 123.2,

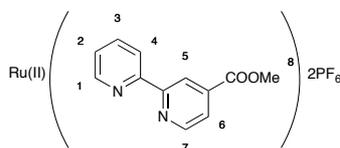
124.5, 138.2, 145.9, 148.7, 154.2, 155.2, 162.3, 172.5; IR (solid state, cm^{-1}) 2460 (O-H), 1709 (C=O acid); ESI-HRMS found m/z 200.0660 $[\text{M}]^+$, $[\text{C}_{11}\text{H}_8\text{N}_2\text{O}_2]^+$ requires 200.0586

Methyl 2-(pyridin-2-yl)pyridine-4-carboxylate, **91**



6-(Pyridin-2-yl)pyridine-3-carboxylic acid, **90** (2.20g, contaminated with salts), methanol (100 mL) and thionyl chloride (15 mL) were heated under reflux for 16 hours. The reaction mixture was then cooled to room temperature and quenched with saturated sodium hydrogen carbonate solution (100 mL) and extracted with dichloromethane (100 mL). The organic phase was washed with 1 M hydrochloric acid (50 mL) and brine (50 mL), dried (sodium sulfate) and concentrated. The resulting beige solid was redissolved in a minimum amount of methanol, and precipitated with water. The resulting solution was filtered and the precipitate washed with water, to yield the product as a beige solid (310 mg, 1.45 mmol); ^1H NMR (500 MHz, CDCl_3) δ ppm 4.00 (s, 3 H, H8), 7.28 (dd, $J = 7.7, 4.5$ Hz, 1 H, H2), 7.74 - 7.83 (m, 2 H, H7 and H3), 8.35 (d, $J = 7.7$ Hz, 1 H, H4), 8.65 (d, $J = 4.5$ Hz, 1 H, H1), 8.76 (d, $J = 5.0$ Hz, 1 H, H6), 8.87 (s, 1 H, H5); ^{13}C NMR (126 MHz, CDCl_3) δ ppm 52.7, 120.4, 121.3, 122.8, 124.2, 137.0, 138.5, 149.4, 150.0, 155.4, 157.4, 165.8; IR (solid state, cm^{-1}) 1720 (C=O ester); ESI-HRMS found m/z 215.0817 $[\text{M}+\text{H}]^+$, $[\text{C}_{12}\text{H}_{10}\text{N}_2\text{O}_2]^+$ requires 215.0821

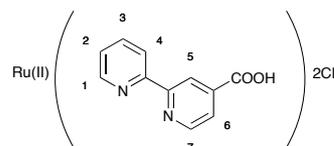
Tris (methyl 2-(pyridin-2-yl) pyridine -4- carboxylate) ruthenium(II) dihexafluorophosphate, **92**



Methyl 2-(pyridin-2-yl)pyridine-4-carboxylate, **91** (200 mg, 0.934 mmol), $\text{Ru(II)(DMSO)}_4\text{Cl}_2$ (141 mg, 0.292 mmol), silver nitrate (99 mg, 0.58 mmol) and ethanol (25 mL) were heated under reflux for 7 days. The resulting mixture was cooled and filtered through celite. The celite pad was then washed with dichloromethane until all of the red compound had washed through. The filtrate was then concentrated and the resulting red solid redissolved in water. An excess of ammonium hexafluorophosphate

was added, and a red precipitate formed. The mixture was filtered to yield the product as a red solid (256 mg, 0.248 mmol, 85 %); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ ppm 3.90 (s, 9 H, H8), 7.52 (s, 3 H, H5), 7.92 (m, 6 H, H7+H3), 8.00 (d, $J = 6.9$ Hz, 3 H, H4), 8.15 (d, $J = 17.1$ Hz, 3 H, H1), 8.35 (m, 3 H, H2), 8.75 (dd, $J = 17.1, 8.5$ Hz, 3 H, H3); IR (solid state, cm^{-1}) 1721 (C=O ester); ESI-MS m/z found 372.0636 $[\text{M}]^{2+}$, $[\text{C}_{36}\text{H}_{30}\text{N}_6\text{O}_6]^{2+}$ requires 372.0635

Tris (2-(pyridin-2-yl)pyridine-4-carboxylic acid) ruthenium dichloride, 64



Tris (methyl 2-(pyridin-2-yl)pyridine-4-carboxylate) ruthenium(II) dihexafluorophosphate (100 mg, 0.0977 mmol), in ethanol (2.5 mL) and 10% sodium hydroxide solution (2.5 mL) were stirred for 2 hours. The resulting solution was neutralized with 1 M hydrochloric acid and concentrated. The resulting solid was dissolved in methanol and filtered. The red filtrate was concentrated and the process repeated, until no more salt was visible. This gave the product as a red solid (56 mg, 0.072 mmol, 74 %). $^1\text{H NMR}$ (500 MHz, D_2O) δ ppm 7.37 (t, $J = 5.7$ Hz, 3 H, H1), 7.65 (d, $J = 4.2$ Hz, 3 H, H6), 7.72 - 7.80 (m, 3 H, H4), 7.87 (td, $J = 5.7, 3.3$ Hz, 3 H, H2), 8.05 (dt, $J = 8.5, 5.7$ Hz, 3 H, H3), 8.57 (d, $J = 7.7$ Hz, 3 H, H7), 8.80 (s, 3 H, H5); IR (solid state, cm^{-1}) 3223 (O-H), 1663 (C=O acid) ESI-MS m/z found 351.0376 $[\text{M}]^{2+}$, $[\text{C}_{33}\text{H}_{24}\text{N}_6\text{O}_6\text{Ru}]^{2+}$ requires 351.0400

6.2 Biophysical analyses- General considerations chapter 2

All stocks for luminescence intensity assays were made up in 5 mM sodium phosphate, pH 7.5 buffer. Ru(II)(bpy)_3 complex stocks were made up to 2 mM. Horse heart and yeast cyt *c* was obtained from Sigma Aldrich, and used without further purification. Cyt *c* stocks were made up to ~ 1 mM, and the concentration accurately determined using the molar extinction coefficient at 550 nm of $2.95 \times 10^4 \text{ mol}^{-1} \text{ dm}^3 \text{ cm}^{-1}$ for horse heart cyt *c*²⁰² and $2.11 \times 10^4 \text{ mol}^{-1} \text{ dm}^3 \text{ cm}^{-1}$ for yeast cyt *c*²⁰² after reduction by addition of one microspatula of sodium dithionite. Assays with chemically oxidized cyt *c* in ascorbate containing buffer used cyt *c* oxidized with $\text{K}_3\text{Fe}(\text{CN})_6$ followed by dialysis into 5 mM sodium phosphate, 2 mM sodium ascorbate, pH 7.5 buffer, to remove the excess $\text{K}_3\text{Fe}(\text{CN})_6$. The concentration of oxidized cyt *c* was determined by using the molar extinction coefficient at 410 nm of $1.061 \times 10^5 \text{ mol}^{-1} \text{ dm}^3 \text{ cm}^{-1}$.²⁰³ All other buffers used were at 5 mM concentration, 0.2 mg mL^{-1} BSA, pH 7.5, unless otherwise stated.

6.2.1 Luminescence quenching assays (Chapter 2)

In all assays the Ru(II)(bpy)₃ complex concentration was kept constant, with the concentration of cyt *c* being varied through the assay, as described below. Results obtained were fitted, using OriginPro 9, to a 1:1 binding isotherm (Eq. 6.1).

$$I = \frac{m[(a+b+K) - \sqrt{(a+b+K)^2 - 4ab}]}{2a} \quad \text{Eq. 6.1}$$

Where *I* = change in relative luminescence intensity (*I*/*I*₀), *m* = maximum value of *I*, *a* = concentration of Ru(II)(bpy)₃ complex, *K* = dissociation constant, *b* = concentration of protein added

6.2.1.1 Fluorometer

Assays were measured on a Jobin-Yvon Spex Fluorolog-3 fluorometer. Measurements were taken in a 4 mL quartz cuvette with excitation at 467 nm and emission measured over the range 575 – 675 nm, with 10 nm slit widths on both excitation and emission. Measurements were taken on triplicate titrations.

1. 2.5 mL of the Ru(II)(bpy)₃ complex at an appropriate concentration was put in the cuvette
2. The emission spectrum was taken
3. 10 μL of a solution with an appropriate cyt *c* concentration with the same concentration of Ru(II)(bpy)₃ complex (so as to keep the complex concentration constant) was added
4. The emission spectrum was taken
5. This was repeated ~15 times
6. The peak maxima were taken, and these values used (as *I*)

6.2.1.2 Plate reader

Assays were scanned using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader, with excitation at 467 nm, and emission at 630 nm fixed wavelength. A 2/3 dilution regime in a 384 well plate (Optiplate) was used, with each result measured in triplicate:

1. 25 μL of buffer was added to all wells
2. 50 μL of cyt *c* solution was added to the first well of a row
3. Serial dilution, taking 50 μL from the first well and transferring to the second well mixing, and repeating up to the 23rd well, leaving the last well without any protein

4. 25 μL of Ru(II)(bpy)_3 complex solution was added to each well
5. The plate was incubated for 45 minutes before reading (excitation 467 nm, emission 630 nm)

6.2.2 UV/Vis ascorbate reduction

A 500 μL quartz cuvette was charged with 400 μL of a 20 μM solution of *cyt c* and 20 μM Ru(II)(bpy)_3 complex in 5 mM sodium phosphate buffer, pH 7.5. 100 μL of 10 mM ascorbate in 5 mM sodium phosphate, pH 7.5 was added, and the spectrophotometer started immediately. Absorbance at 550 nm was monitored every 0.5 seconds for 5 minutes. The absorbances were normalised (against $A_{550_{\text{max}}}-A_{550_{\text{min}}}$) and plotted against time.

6.2.3 Protein NMR

Sensitivity enhanced ^1H - ^{15}N HSQC NMR correlation spectra of ligand-bound and unbound forms of horse heart *cyt c*, purchased from Sigma Aldrich, were carried out at natural abundance using a 950 MHz Bruker Ascend™ Aeon spectrometer operating at a proton (^1H) resonance frequency of 950.13 MHz equipped with a Bruker TCI triple-resonance cryo-probe. NMR acquisitions were carried out in 5 mM sodium phosphate, 2 mM sodium ascorbate, pH 7.25 buffer. For *cyt c* alone, spectra were taken at 2 mM protein concentration, with a total volume of 600 μL . With Ru(II)(bpy)_3 complex **31** and **26**, 1 mM *cyt c* and 0.5 mM Ru(II)(bpy)_3 complex were used, to a total volume of 600 μL . Spectra were analysed using the CcpNmr Analysis software package and the chemical shift perturbations were calculated as the square root of the sum of the isotope weighted shift differences squared (Eq. 6.2),

$$\Delta\delta = \sqrt{(\Delta\delta_N)^2 + (\gamma_H/\gamma_N)^2 (\Delta\delta_H)^2} \quad \text{Eq. 6.2}$$

Where $\Delta\delta$ is the overall change in chemical shift, $\Delta\delta_N$ is the change in the nitrogen dimension and $\Delta\delta_H$ is the change in the proton dimension, respectively. The change in the proton dimension is scaled by the ratio of the gyromagnetic ratio of ^{15}N (γ_N) and ^1H (γ_H) to account for the larger chemical shift range of ^{15}N .

6.3 Hydrazone Exchange Studies (Chapter 3)

All hydrazone exchange reactions were carried out in HPLC vials and were followed using high resolution mass spectrometry, using a Bruker Daltonics micrOTOF Premier Mass Spectrometer, using 10 μL injections and summing the masses over the range 1.0 to 3.0 minutes. The intensity of the maximum peak for each of the successive hydrazone

exchanges was taken and the percentage associated with each of the Ru(II)(bpy)₃ or porphyrin peaks calculated.

6.3.1 Ru(II)(bpy)₃ Complexes

Ru(II)(bpy)₃ hydrazone complex stocks were made up at 2 mM concentration in 1:1 acetonitrile:water. Aniline and aldehyde (2,4-dimethoxy benzaldehyde) stocks were made to 1 M concentration in acetonitrile.

6.3.1.1 Hydrazone exchange reactions

In hydrazone exchange reactions 1 mL of 100 μM Ru(II)(bpy)₃ hydrazone complex **46a**, 10 mM aldehyde and 10 mM aniline in 1:1 acetonitrile water were incubated in an HPLC vial and mass spectra were obtained at appropriate time periods.

6.3.2 Porphyrins

Hydrazone functionalised porphyrins **79** were made up to 5 mM concentration in DMSO and were stored in plastic Eppendorf tubes. 1 M stocks of catalyst (aniline and anthranilic acid) and aldehydes (2,4-dimethoxy benzaldehyde, 4-carboxy benzaldehyde, 4-methyl ester benzaldehyde) were made up in DMSO. 4-hydroxy benzaldehyde stock were made up to 0.5 M in DMSO.

6.3.2.1 Hydrazone exchange reactions

In all exchange reactions the hydrazone functionalised porphyrin **79** in DMSO was added to a solution of the aldehyde and catalyst to give a total concentration of 10 % DMSO in 5 mM ammonium acetate buffer, pH 6.75, to a final porphyrin concentration of 100 μM and stated concentrations of other components. For time courses, mass spectra were obtained at appropriate time points (usually 0.5, 1, 2, 4, 6, 9, 12 and 24 hours). For measurements at single time points, mass spectra were obtained after 24 hours incubation. For pre-incubated samples, the pre-incubated mixture was left for 24 hours, and a mass spectrum obtained, prior to addition of further components.

6.3.2.2 Protein incubation

A 200 μL solution containing 100 μM benzaldehyde hydrazone porphyrin **79**, 2.5 mM of any aldehyde, 10 mM aniline in 10 % DMSO in 5 mM NH₄OAc, pH 6.75 buffer was incubated for 24 hours. At which point 20 μL of 1 mM protein stock in 5 mM sodium phosphate, pH 7.5 buffer or 5 mM sodium phosphate, pH 7.5 buffer alone (for no protein comparison) was added, and incubated for a further 24 hours. The reaction was then quenched by addition of 20 μL of NH₄OH solution.

6.4 Biophysical analyses- General considerations chapter 4

All arrays were performed in 5 mM sodium phosphate, pH 7.5 buffer. All protein stocks, other than Mcl-1 and hDM2, were made up from freeze-dried protein, purchased from major suppliers, into 5 mM sodium phosphate, pH 7.5 at ~ 1 mM concentration. Accurate concentrations were determined by UV/Vis using extinction coefficients (at 280 nm) of 36,²⁰⁴ 50,²⁰⁵ 43.8,²⁰⁶ 57.6²⁰⁷ and 9.9²⁰⁷ mol⁻¹ dm³ cm⁻¹ for lysozyme, α -chymotrypsin, BSA, papain and ribonuclease A respectively and for cyt c, as for chapter 2.

All arrays were performed in 384 Optiplate well plates and were scanned using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader.

6.4.1 Ru(II)(bpy)₃ complexes and proteins

To each well was added 20 μ L of 5 μ M Ru(II)(bpy)₃ complex solution and 20 μ L of 20 μ M protein solution or 5 mM sodium phosphate, pH 7.5 buffer. The plate was incubated for 45 minutes before scanning using fixed wavelengths, excitation 467 nm, emission 630 nm, and using monochromators, excitation 467 nm, emission range 500 – 800 nm, 3 nm step, 100 flashes. The peak maxima/ intensities were taken. The values without protein were averaged over the triplicate wells, and the percentage difference for each of the other wells containing the same Ru(II)(bpy)₃ complex calculated. Each of these results was used for statistical analysis.

6.4.2 Ru(II)(bpy)₃ complexes, FITC-NOXA B (R-A) tracer and proteins

To each well was added 20 μ L of 7.5 μ M Ru(II)(bpy)₃ complex solution (or 5 mM sodium phosphate buffer, pH 7.5), 20 μ L of 30 μ M protein solution (or 5 mM sodium phosphate, pH 7.5 buffer) and 20 μ L of 1.5 μ M FITC-NOXA-B (R-A) peptide (or 5 mM sodium phosphate, pH 7.5 buffer). On each plate wells without protein were run in quadruplicate. The plate was incubated for 2 and 20 hours prior to scanning using fixed wavelengths, excitation 467 nm, emission 630 nm, and using monochromators, excitation 467 nm, emission range 480 - 750 nm, 3 nm step, 100 flashes. The peak maxima/intensities for both luminescence bands were taken (emission 520 nm and 630 nm). The values obtained for wells with no protein present were averaged over the quadruplicate wells, all other data used was from each well individually. The percentage difference for each well for that with protein to without protein was calculated and used for statistical analysis.

6.4.3 Linear discriminant analysis

Microsoft Excel was used to calculate the percentage differences from no protein for each of the individual wells. Linear discriminant analysis was carried out using XLstat software, then plotted using OriginPro 9. Confidence ellipses were obtained using OriginPro 9.4.

6.5 General considerations for Appendix IV - DCC array studies with hydrazone functionalised Ru(II)(bpy)₃ complexes

The glycine hydrazone Ru(II)(bpy)₃ complex **49** was made up to 10 mM concentration in 50 mM ammonium acetate, pH 6.2 buffer. Aniline stocks were made up to 1 M in DMSO, and aldehyde (4-carboxy benzaldehyde, 2,4-dimethoxy benzaldehyde, 4-hydroxy benzaldehyde, 2,4-dihydroxy benzaldehyde, furfural, 4-nitro benzaldehyde, *tert*-butoxy benzaldehyde, benzaldehyde, 3-pyridine carboxaldehyde, 4-chloro benzaldehyde) stocks were made up to 100 mM in DMSO. Protein stocks were made up to ~1 mM in 5 mM sodium phosphate, pH 7.5 and concentrations measured as for Chapter 4.

DCC arrays were set up in 96 well plates (Optiplat) in triplicate. In each well 250 μ L of 50 μ M glycine hydrazone complex, 10 mM aniline, 300 μ M each aldehyde and 50 μ M protein in 10 % DMSO in 50 mM ammonium acetate, pH 6.2 buffer was incubated for 18 hours at room temperature. At which point the reaction was quenched by addition of 4 μ L of 1 M NaOH. The contents of each well were transferred to a vivaspin 500 protein concentrator MWCO 5 kDa, and centrifuged at 13 000 g for 22 minutes. 4 μ L of 1 M NaOH followed by 200 μ L of NH₄OAc, pH 6.2 buffer was added to the concentrate and the concentrator centrifuged at 8 000 g for 40 minutes. The flow through was analysed by analytical HPLC (Rapid 5-95 Methanol+TFA gradient using an Ascentis Express C18 column) at 280 nm, and the peak integrations for the major peaks taken and used for analysis.

6.6 General considerations for Appendix V - Stabilisation of the p53 tetramerisation domain

The two peptides were obtained from ProteoGenix. Peptide stocks were made up to 1 mg mL⁻¹ in water, and neutralised to pH 7.0 by addition of 10 mM hydrochloric acid or 10 mM ammonium hydroxide solution, or in 5 mM sodium phosphate, pH 7.5. The accurate peptide concentration was determined by UV/Vis absorption at 280 nm (ϵ = 1490 M⁻¹cm⁻¹ for both peptides). Ru(II)(bpy)₃ complex stocks were made up to 1 mM in water, neutralised to pH 7.0 by addition of 10 mM hydrochloric acid or 10 mM ammonium hydroxide solution, or in 5 mM sodium phosphate, pH 7.5 buffer.

6.6.1 Circular dichroism

Circular dichroism was performed on an Applied Photophysics ChiraScan Apparatus using the associated Software. The range 180-260 nm was scanned using point time 1 s, 1 nm per point, step = 1 nm, 5 nm bandwidth and path length 10 mm. All samples were run in either 5 mM sodium phosphate, pH 7.5 buffer or water. 100 μ M peptide concentrations were used for all spectra, with 0, 100 or 400 μ M Ru(bpy)₃ complex concentration.

For temperature ramp experiments, the temperature was ramped in steps of either 1 °C or 2.5 °C over the range 20 – 90 °C, with a temperature equilibration time of 4 minutes between the temperature increase and obtaining the CD spectra. A final spectrum was obtained after the chamber had cooled to 20 °C. The ellipticity at 222 nm (corresponding to helical content) was used to calculate the fraction folded (Eq. 6.3).

$$\alpha = \frac{\theta(T) - \theta_f}{\theta_u - \theta_f} \quad \text{Eq. 6.3}$$

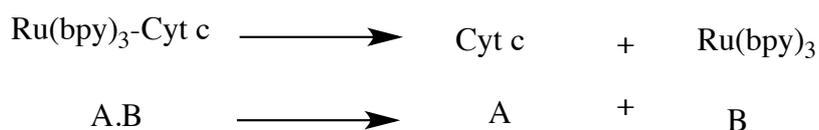
α = fraction folded, $\theta(T)$ = ellipticity at temperature T, θ_f = θ at 20 °C and θ_u = θ at 90 °C

7 Appendices

7.1 Appendix I- Debye Hückel equation

The Debye-Hückel limiting law derives from the idea that near any ion in solution, counter ions will cluster, and thus lower the chemical potential of that original ion, making it act non-ideally and lowering its electrostatic interactions.²⁰⁸ This will affect interactions within the solution, for example the binding between the Ru(bpy)₃ complex and cyt *c*, both of which are charged and thus be dependent on factors including the concentration, size and charge of the ions.

In order to see how this can provide information on the binding between the two species, it is first necessary to understand the fundamentals of chemical reaction thermodynamics. Considering the reaction equation for the dissociation of the Ru(II)(bpy)₃-cyt *c* interaction-



Using standard thermodynamic notation, the Gibbs free energy (ΔG) can be expressed in terms of the chemical potentials (μ) of each of the species

$$\Delta G = \mu_A + \mu_B - \mu_{A.B} \quad \text{Eq. 7.1}$$

The chemical potential of a species *i* can be expressed in terms of the standard chemical potential (μ^0) of that species and the activity (*a*), the effective concentration of the species in a mixture

$$\mu_i = \mu_i^0 + RT \ln a_i \quad \text{Eq. 7.2}$$

The activity can further be defined in terms of the molality (*m*) (the concentration of solute in a solvent, in water this is the molarity) and an activity coefficient (γ), which depends on the composition, molality and temperature of the solution

$$a_i = \gamma_i \frac{m_i}{m_i^0} \quad \text{Eq. 7.3}$$

Combining Eq. 6.2 and Eq. 6.3

$$\mu_i = \mu_i^0 + RT \ln \frac{m_i}{m_i^0} + RT \ln \gamma_i \quad \text{Eq. 7.4}$$

If we assume that the molality is the ideal molality ($m_i = m_i^o$), then

$$\mu_i = \mu_i^o + RT \ln \gamma_i \quad \text{Eq. 7.5}$$

This equation can then be put back into Eq. 6.1

$$\Delta G = \mu_A^o + \mu_B^o - \mu_C^o + RT \ln \gamma_A + RT \ln \gamma_B - RT \ln \gamma_C \quad \text{Eq. 7.6}$$

As $\Delta G^o = \mu_A^o + \mu_B^o - \mu_{A,B}^o \quad \text{Eq. 7.7}$

Then $\Delta G = \Delta G^o + RT \ln \gamma_A + RT \ln \gamma_B - RT \ln \gamma_{A,B} \quad \text{Eq. 7.8}$

Simplifying $\Delta G = \Delta G^o + RT \ln \frac{\gamma_A \gamma_B}{\gamma_{A,B}} \quad \text{Eq. 7.9}$

Inputting the Gibbs free energy isotherm-

$$\Delta G = -RT \ln K \quad \text{Eq. 7.10}$$

Gives $-RT \ln K = -RT \ln K^o + RT \ln \frac{\gamma_A \gamma_B}{\gamma_{A,B}} \quad \text{Eq. 7.11}$

Simplifying this gives $\ln K = \ln K^o - \ln \frac{\gamma_A \gamma_B}{\gamma_{A,B}} \quad \text{Eq. 7.12}$

In an ideal (infinitely dilute) solution or gas, the γ values are 1, thus making

$$\ln K = \ln K^o \quad \text{Eq. 7.13}$$

An ionic solution, however is not ideal and $\gamma \neq 1$, therefore we need to compute values for γ , this is done using the Debye-Hückel approximation,^{208,209} where it is assumed that the activities (a) or activity coefficients (γ) are assumed to be dependent on the ionic strength of the solution (I) and not on the composition of the solution.

The Debye-Hückel approximation- $\log_{10} \gamma = -|z_+ z_-| A(I)^{1/2} \quad \text{Eq. 7.14}$

Where, z_+ and z_- are the charges on the two species, A is a constant which is empirically derived, and in aqueous solutions, at 25 °C is 0.509.

A second approximation is the Güntelberg approximation which is empirically derived and valid to higher ionic strengths (<100 mM), and assigns some size to the ions (as opposed to the Debye-Hückel equation, which assumes point charges).²⁰⁹

Güntelberg extension- $\log_{10} \gamma_i = \frac{A |z_i|^2 \sqrt{I}}{1 + \sqrt{I}} \quad \text{Eq. 7.15}$

And
$$\log_{10}\gamma_{ij} = \frac{A|z_i z_j|\sqrt{I}}{1+\sqrt{I}} \quad \text{Eq. 7.16}$$

Converting to ln
$$\ln\gamma_i = \frac{A}{\log_{10}e} \frac{|z_i^2|\sqrt{I}}{1+\sqrt{I}} \quad \text{Eq. 7.17}$$

$$\ln\gamma_{ij} = \frac{A}{\log_{10}e} \frac{|z_i z_j|\sqrt{I}}{1+\sqrt{I}} \quad \text{Eq. 7.18}$$

Inputting Eq. and Eq. into Eq. 6.13 for the dissociation reaction

$$\ln K = \ln K^o - \frac{\frac{A}{\log_{10}e} \frac{|z_A^2|\sqrt{I}}{1+\sqrt{I}} \times \frac{A}{\log_{10}e} \frac{|z_B^2|\sqrt{I}}{1+\sqrt{I}}}{\frac{A}{\log_{10}e} \frac{|z_A z_B|\sqrt{I}}{1+\sqrt{I}}} \quad \text{Eq. 7.19}$$

Simplifying
$$\ln K = \ln K - \frac{ART}{\log_{10}e} \left(\frac{|z_A z_B|\sqrt{I}}{1+\sqrt{I}} \right) \quad \text{Eq. 7.20}$$

Changing from ln to log₁₀ and simplifying-

$$\log_{10}K = \log_{10}K^o - A \left(\frac{|z_A z_B|\sqrt{I}}{1+\sqrt{I}} \right) \quad \text{Eq. 7.21}$$

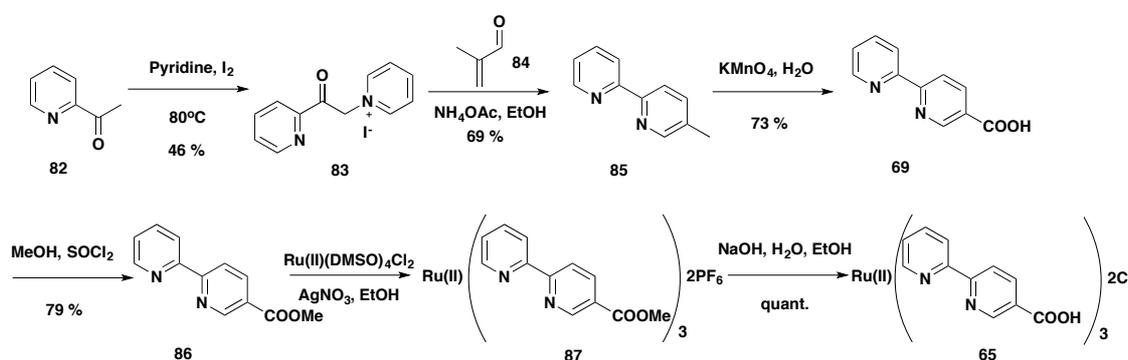
At 298 K in water A = -0.509, therefore

$$\log_{10}K = \log_{10}K^o - 0.509 \left(\frac{|z_A z_B|\sqrt{I}}{1+\sqrt{I}} \right) \quad \text{Eq. 7.22}$$

7.2 Appendix II- Monosubstituted Ru(II)(bpy)₃ complex synthesis

The synthesis of monosubstituted Ru(II)(bpy)₃ complexes was attempted for the use in DCC studies as there would be fewer exchangeable groups compared to the disubstituted Ru(II)(bpy)₃ complexes (3 versus 6). This would decrease the number of peaks present in the mass spectrum, and also potentially increase the rate of reaching equilibrium. The use of monosubstituted bipyridine ligands, however, does lead to more stereoisomers of the Ru(II)(bpy)₃ complexes formed, as *fac* and *mer* isomers as well as Δ and Λ isomers possible, will be formed

7.2.1 5' monosubstituted bipyridine synthesis

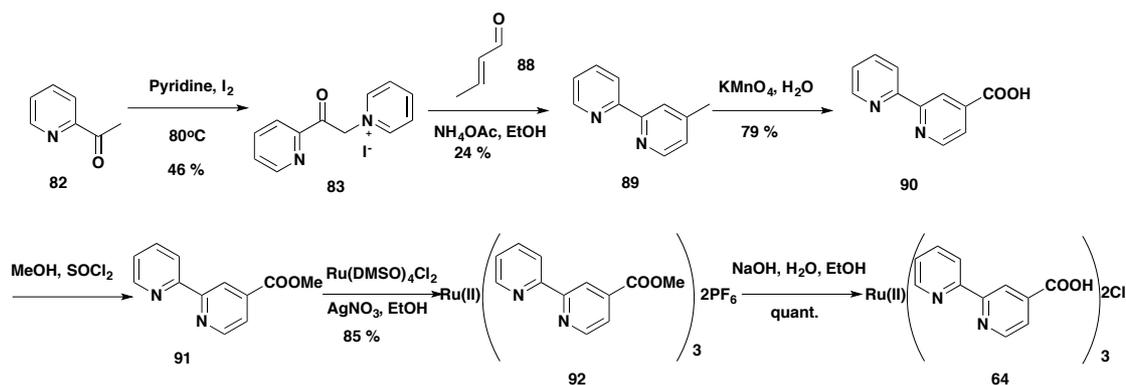


Scheme 7.1 Synthesis of 5' monosubstituted bipyridine Ru(II)(bpy)₃ complex **65**

The Wilson group have previously reported the use of 5' monosubstituted Ru(II)(bpy)₃ complexes for binding to cyt *c*.¹⁴³ It was therefore decided to synthesise a 5' acid complex, *via* a similar synthetic route (**Scheme 7.1**). The synthesis uses a Kroehnke method^{210,211} to access the bipyridine by reaction of the Kroehnke reagent **83** with methacrolein **84** to give a 5' methyl substituted bipyridine **85**. This can be oxidised to the corresponding carboxylic acid **69** with potassium permanganate, followed by methyl ester formation to form **86**, ruthenium(II) complexation to give **87** and subsequent deprotection, to yield the 5' acid Ru(II)(bpy)₃ complex **65**.

7.2.2 4' monosubstituted bipyridine complex synthesis

While the 5' bipyridine Ru(II)(bpy)₃ complexes have been used in the Wilson group previously,¹⁴³ 4' bipyridines have not been reported by the group. However, the 4' substituted Ru(II)(bpy)₃ complexes offer advantages over the corresponding 5' substituted Ru(II)(bpy)₃ complexes, including higher quantum yields of the luminescence,¹⁴⁵ making them much more compatible with luminescence assays and *in cellulo* visualisation.



Scheme 7.2 4' monosubstituted Ru(II)(bpy)₃ complex synthesis using Kroehnke methodology

After various attempts at palladium catalysed cross coupling, using Suzuki and Negishi couplings yielded little of the bipyridine ligand **89**, the Kroehnke method was again used, despite difficulties found in purification of the methyl ligand **89**. The same Kroehnke reagent **83** as in the 5' substituted bipyridine synthesis was used, and reacted with crotonaldehyde **88**. In the 5' substituted 2,2'-bipyridine synthesis the methyl ligand **85** could easily be purified by a filtration silica column, however the 4' substituted methyl ligand **89** stuck to the column and could not be eluted. Thus, in this case, the acid formation and subsequent methyl ester formation were performed crude, with the methyl ester ligand **91** purified prior to ruthenium(II) complexation. Ruthenium(II) complexation and subsequent deprotection afforded the 4' monosubstituted acid Ru(II)(bpy)₃ complex **64**.

7.3 Appendix III- Principle component analysis and linear discriminant analysis

A qualitative description of both principle component analysis (PCA) and linear discriminant analysis was presented in Chapter 4. This appendix presents some of the mathematics behind these techniques, using the Ru(II)(bpy)₃ complex-protein array as an example, where appropriate.

The aim of both of these analyses is to generate equations which give linear combinations of the data original data, e.g.

$$X = a_1A + b_1B + \dots + h_1H \quad \text{Eq. 7.23}$$

Where A, B, ... , H are the luminescence responses from the different Ru(II)(bpy)₃ complexes and a₁, b₁, ... , h₁ are linear components used to generate the first principle component/linear discriminate, these are derived using the analysis. This equation can also be written in a vector format.

$$X = \mathbf{w}^T \mathbf{a} \quad \text{Eq.7.24}$$

$$\text{Where } \mathbf{w} = \begin{pmatrix} a_1 \\ b_1 \\ \vdots \\ h_1 \end{pmatrix}, \text{ and } \mathbf{w}^T \text{ its transpose and } \mathbf{a} = \begin{pmatrix} A \\ B \\ \vdots \\ H \end{pmatrix}$$

These equations/ vectors project the data onto different axes, in the hope to have the maximum possible variance (in PCA) or the maximum discrimination of classes (in LDA).

7.3.1 Statistical definitions

Before using the more complicated mathematics it is necessary to define some statistical variables.

$$\text{Mean} \quad \mu_A = \frac{1}{n} (a_1 + a_2 + \dots + a_n) \quad \text{Eq. 7.25}$$

$$\text{Variance} \quad \text{Var}(A) = \frac{1}{n-1} ((a_1 - \mu_A)^2 + (a_2 - \mu_A)^2 + \dots + (a_n - \mu_A)^2) \quad \text{Eq. 7.26}$$

$$\text{Covariance} \quad \text{Cov}(A, B) = \frac{1}{n-1} ((a_1 - \mu_A)(b_1 - \mu_B) + \dots + (a_n - \mu_A)(b_n - \mu_B)) \quad \text{Eq. 7.27}$$

Note (by definition) Cov(A,B) = Cov(B,A)

7.3.2 Principal component analysis

Principle component analysis looks at addressing the variance in the data, without taking into account the classes.

Let the vectors **a**, **b**, **c** ... **x** etc. be each of the individual protein replicates

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_8 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_8 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_8 \end{pmatrix} \text{ etc.}$$

Where a_1 = the response of Ru(II)(bpy)₃ complex 1 in protein replicate a, a_2 = the response of Ru(II)(bpy)₃ complex 2 in protein replicate a, etc. for the 8 Ru(II)(bpy)₃ complexes and for all the protein replicates (b, c, etc.).

The mean of these vectors can be calculated to give a mean vector, $\boldsymbol{\mu}$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_8 \end{pmatrix}$$

Where μ_1 = the mean luminescence response of Ru(II)(bpy)₃ complex 1 in all the protein replicates, μ_2 = the mean luminescence response of Ru(II)(bpy)₃ complex 1 in all the protein replicates, etc.

Using the mean vector it is possible to recentre the vectors **a**, **b**, **c**, etc. around 0, by subtracting the mean vector, $\boldsymbol{\mu}$

$$\mathbf{a} - \boldsymbol{\mu}, \mathbf{b} - \boldsymbol{\mu}, \mathbf{c} - \boldsymbol{\mu} \text{ etc.}$$

These new vectors can be combined to give a matrix, X, which contains all the data, centred about 0.

$$X = \begin{pmatrix} a_1 - \mu_1 & b_1 - \mu_1 & \cdots & x_1 - \mu_1 \\ a_2 - \mu_2 & b_2 - \mu_2 & \cdots & x_2 - \mu_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_8 - \mu_8 & b_8 - \mu_8 & \cdots & x_8 - \mu_8 \end{pmatrix}$$

Using this data matrix, X, it is possible to define the covariance matrix, C

$$C = \frac{1}{x-1} XX^T \quad \text{Eq. 7.28}$$

Where x is the number of protein replicates, and X^T is the transpose of matrix X

This covariance matrix, C , is symmetric ($C = C^T$) by definition.

As C is symmetric it has real eigenvalues (and corresponding eigenvectors) which fulfil the equation-

$$C\mathbf{w}_i = \lambda_i\mathbf{w}_i \quad \text{Eq. 7.29}$$

Where λ_i = eigenvalue, and \mathbf{w}_i = the corresponding eigenvector

The solutions to this equation will give x eigenvectors and eigenvalues.

These eigenvalues can be put into magnitude order ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \geq \lambda_x \geq 0$). It may be that a small number (2 or 3) of the eigenvalues are much larger than the other eigenvalues, this means that these take the majority of the variation in the original data set. The corresponding (orthonormal) eigenvectors are the vectors \mathbf{w} in Eq. 7.24 with the variables in the vector being the linear components (a_1, b_1, \dots, h_1) in Eq. 7.23. These can be computed computationally.

7.3.3 Linear discriminant analysis

As seen above principle component analysis does not take into account the classes of the data, instead it keeps the spread in the data, this means that it may not be looking at components in the correct plane to discriminate different classes. This issue is addressed with linear discriminant analysis. Here the original data is project onto a multidimensional surface to increase the discrimination of the original data.

As with PCA, let the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \dots \mathbf{x}$ etc. be each of the individual protein replicates

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_8 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_8 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_8 \end{pmatrix} \text{ etc.}$$

However now, these samples $\mathbf{a}, \mathbf{b}, \mathbf{c}$ also have class labels y_a, y_b, y_c etc.

This means that we can define separate mean vectors for each class (i), $\boldsymbol{\mu}_i$, and we can take the total number of samples in class i , as M_i . Such that-

$$M = \sum_{i=1}^C M_i \quad \text{Eq. 7.30}$$

Where C is the total number of classes

From these parameters it is possible to define two scatter matrices, the within class scatter matrix (S_w) and the between class scatter matrix (S_B).

$$S_w = \sum_{i=1}^C \sum_{j=1}^{M_i} (\mathbf{y}_j - \boldsymbol{\mu}_i)(\mathbf{y}_j - \boldsymbol{\mu}_i)^T \quad \text{Eq. 7.31}$$

$$S_B = \sum_{i=1}^C (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T \quad \text{Eq. 7.32}$$

Where $\boldsymbol{\mu}$ is the mean of the whole data set- $\boldsymbol{\mu} = \frac{1}{C} \sum_{i=1}^C \boldsymbol{\mu}_i$ Eq. 7.33

These scatter matrices can be transformed onto new planes (W) using the equations-

$$\tilde{S}_w = W^T S_w W \quad \text{Eq. 7.34}$$

$$\tilde{S}_B = W^T S_B W \quad \text{Eq. 7.35}$$

The aim of LDA is to find an optimal W (W^*) such that $\det S_b / \det S_w$ is maximised, i.e. finding the minimum within class scatter to the maximum between class scatter.

$$W^* = \operatorname{argmax} \left\{ \frac{W^T S_B W}{W^T S_w W} \right\} \quad \text{Eq. 7.35}$$

This is found by finding a matrix W^* whose columns are the eigenvectors (\mathbf{w}_i) corresponding to the largest eigenvalues of the following equation-

$$(S_B - \lambda_i S_w) \mathbf{w}_i = 0 \quad \text{Eq. 7.36}$$

These eigenvectors and eigenvalues can again be found computationally

7.4 Appendix IV- DCC array studies with hydrazone functionalised $Ru(II)(bpy)_3$ complexes

The arrays discussed in Chapter 4, expand on previously reported luminescence arrays, using new types and combinations of molecules. However, recently the Waters group reported a study combining DCC and arrays, in order to distinguish different histone post-translational modification states.¹⁸⁷ This brings together the ideas presented in Chapter 3 and Chapter 4. In Chapter 3 the development of hydrazone-based DCC was discussed; this is a combinatorial approach whereby a range of potential candidates is sampled in order to find ligands for a particular template (like a protein). The discussion of DCC presented in Chapter 3 was for the generation of high affinity protein ligands, by sampling a wide range of different potential candidates. However when using this approach, it could be that potent protein ligands are not generated, for example, if there is not a potent ligand within the range of potential candidates. Even if high affinity ligands are not generated, it could be that many low affinity ligands are present in the range of potential candidates. Therefore it could be that instead these low affinity ligands could be detected, and the ratios of the candidates present analysed as an array (**Figure 7.1**).

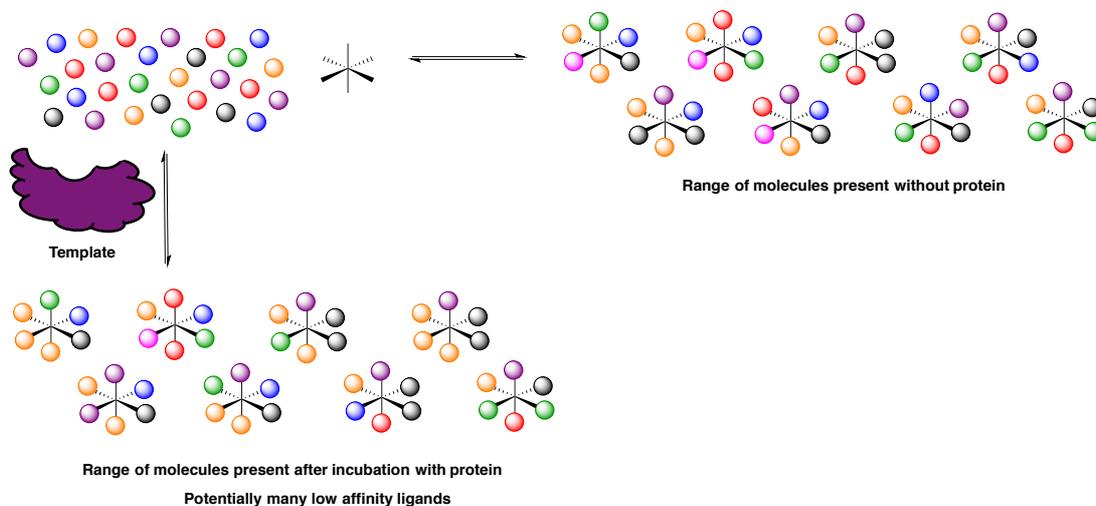
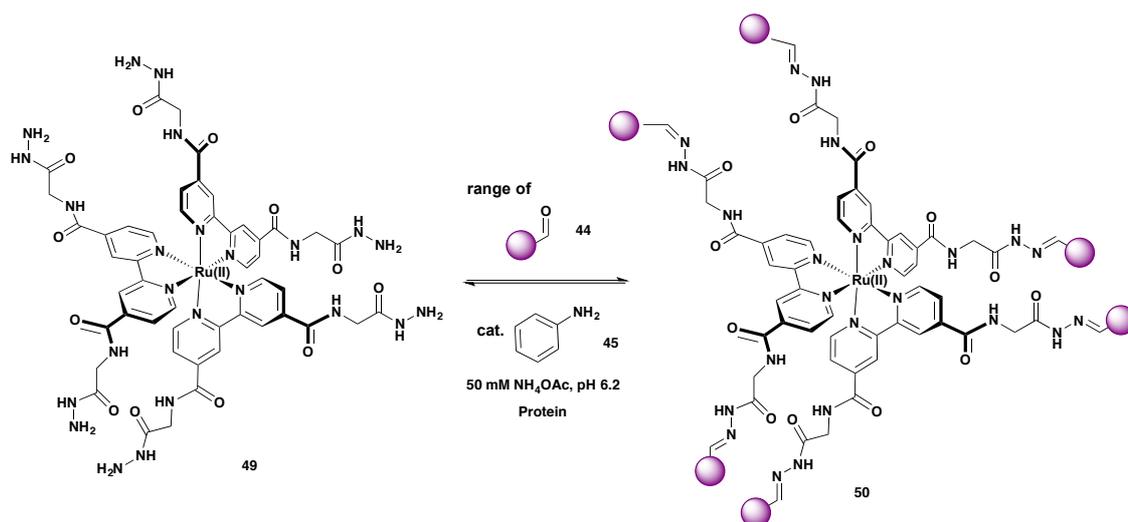


Figure 7.1 Cartoon depiction of DCC array

The glycine hydrazide $Ru(II)(bpy)_3$ complex **49** was used (**Scheme 7.3**) in these studies; as discussed in Chapter 3 the hydrazones of this $Ru(II)(bpy)_3$ complex **50** were insoluble in DMSO/aqueous mixtures, and hence the hydrazone exchange could not be studied. However, for the array study, slightly lower concentrations were used and the starting point for these experiments made use of the water soluble hydrazide $Ru(II)(bpy)_3$ complex, rather than the insoluble hydrazone $Ru(II)(bpy)_3$ complexes **50**.



Scheme 7.3 DCC with glycine hydrazide complex scaffold

Firstly, incubations were performed with 2 and 3 different aldehydes (4-carboxy benzaldehyde **44c**, 2,4-dimethoxy benzaldehyde **44b** and benzaldehyde **44a**), with the hydrazide Ru(II)(bpy)₃ complex **49**, aniline **45** and 3 different proteins (BSA, lysozyme and cyt *c*). After 16 hours incubation, attempts to separate the Ru(II)(bpy)₃ complexes formed from the proteins using protein concentrators (MWCO 5 kDa) were made, however due to the insolubility of the Ru(II)(bpy)₃ hydrazone complexes **46** they precipitated, staying with the protein, rather than filtering through to the flow through with the aldehydes and aniline. A similar problem was found to happen when protein incubations with the porphyrin DCLs were performed, as discussed in Chapter 3.

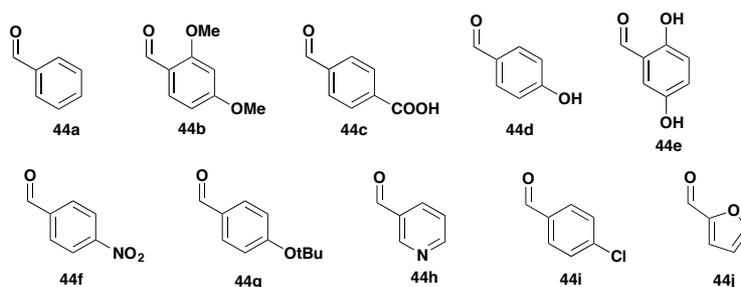


Figure 7.2 Aldehydes for use in the DCC array

As the aim of these experiments was for the discrimination between proteins and not the generation of high affinity protein ligands, analysis of the flow through from the protein concentrator would yield information as to the aldehydes present, and therefore those not used to form hydrazones. As just the aldehydes were to be analysed, rather than the complex mixture of Ru(II)(bpy)₃ hydrazone complexes **50**, it was possible to use a

wider range of aldehydes, and thus sample a much wider range of Ru(II)(bpy)₃ hydrazone complexes. To this end 10 different aldehydes (**Figure 7.2**) were used.

This experiment was performed with both 2 equivalents and 6 equivalents of each aldehyde **44** per hydrazide Ru(II)(bpy)₃ complex **49**, and the aldehyde mixture in the protein concentrator flow through analysed by analytical HPLC. The peaks present at various different retention times were analysed, and the percentage of the total area used. This gave a fingerprint bar chart for the proteins (**Figure 7.3**) similar to those previously described for the luminescence arrays. Upon inspection of the bar charts very little discrimination between the proteins, was observed that can be discerned by eye with either the 2 or the 6 aldehyde equivalents data.

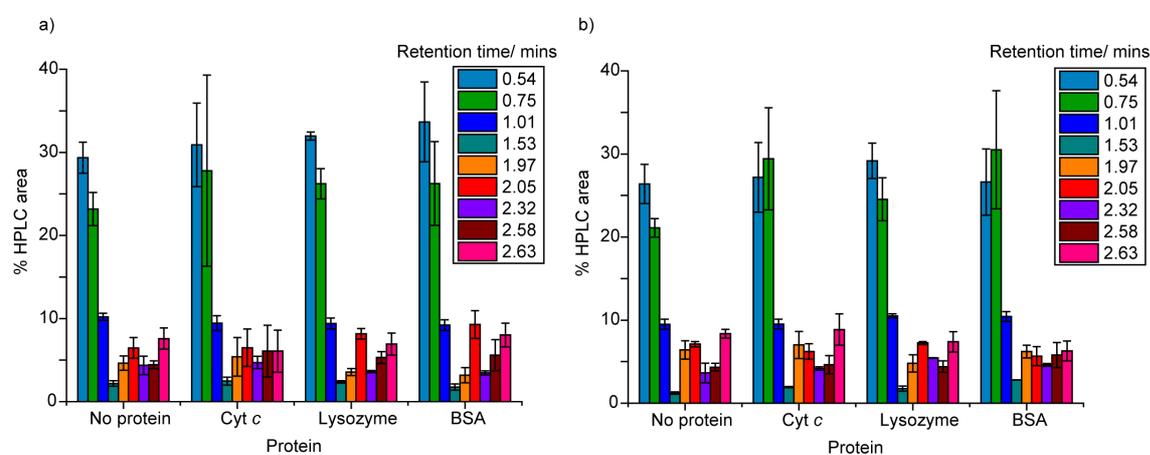


Figure 7.3 Fingerprint bar charts for the DCC arrays with the glycine hydrazide Ru(II)(bpy)₃ complex **43** and 10 aldehydes, using 2 equivalents (a) and 6 equivalents (b) of aldehyde to hydrazide Ru(II)(bpy)₃ complex **43**. Each bar represents the relative area under a peak at a particular retention time in the HPLC trace

To see if protein discrimination could be achieved using this technique, LDA (**Figure 7.4**) was attempted, even though the fingerprints looked similar. This shows some clustering of the data points, and especially in the case of using 6 equivalents of aldehyde **44** shows that the data for the different proteins may be clustering separately, indicating that there is potential to use this technique for the discrimination of different proteins. However, more data would need to be obtained from many more replicates in order to determine if this is statistically significant, as well as controls without the Ru(II)(bpy)₃ hydrazide complex **49** present to see if it is the aldehydes reacting with the protein that cause these differences rather than the hydrazone-based DCC around the Ru(II)(bpy)₃ complex **49**.

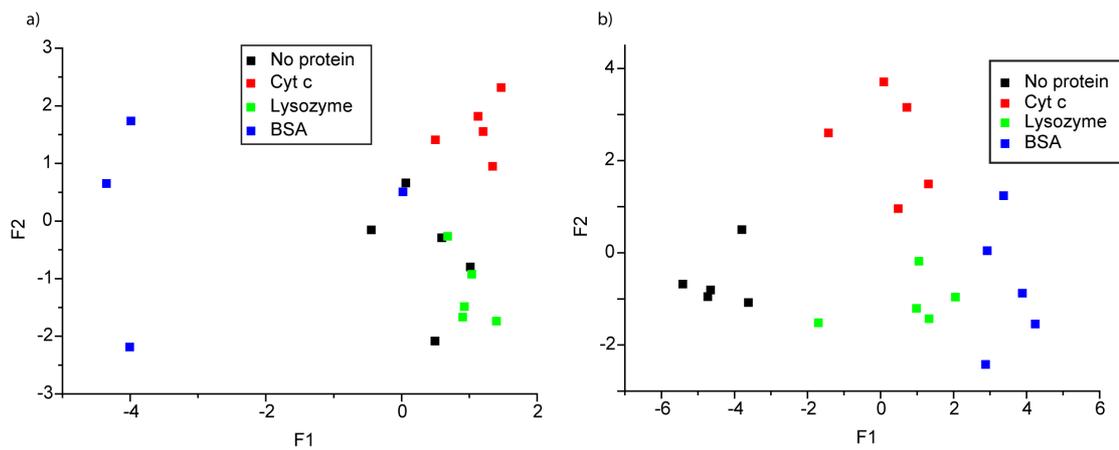


Figure 7.4 2-D LDA of the DCC array data with the glycine hydrazide Ru(II)(bpy)₃ complex **43** and 10 aldehydes **44**, using 2 equivalents (a) and 6 equivalents (b) of aldehyde to hydrazone Ru(II)(bpy)₃ complex **43**

7.5 Appendix V – Stabilisation of the p53 tetramerisation domain

Thus far the Ru(II)(bpy)₃ complexes presented have been studied as protein surface mimetics for use in PPI inhibition and protein sensing. However, another potential use of protein surface mimetics is PPI stabilisation. Previously guanidine functionalised calixarenes and peptides, and amine functionalised peptides have been used to stabilise the p53 tetramerisation domain.²¹²⁻²¹⁵ It was thought that instead Ru(II)(bpy)₃ complexes may be capable of performing the same role.

7.5.1 The p53 tetramerisation domain

p53 is a tumour suppressor protein, which functions as a homotetramer.²¹⁶ It is an essential transcription factor for the cell cycle and apoptosis.²¹⁷ p53 mutations have been implicated in many human cancers, making it an interesting anti-cancer target.²¹⁷

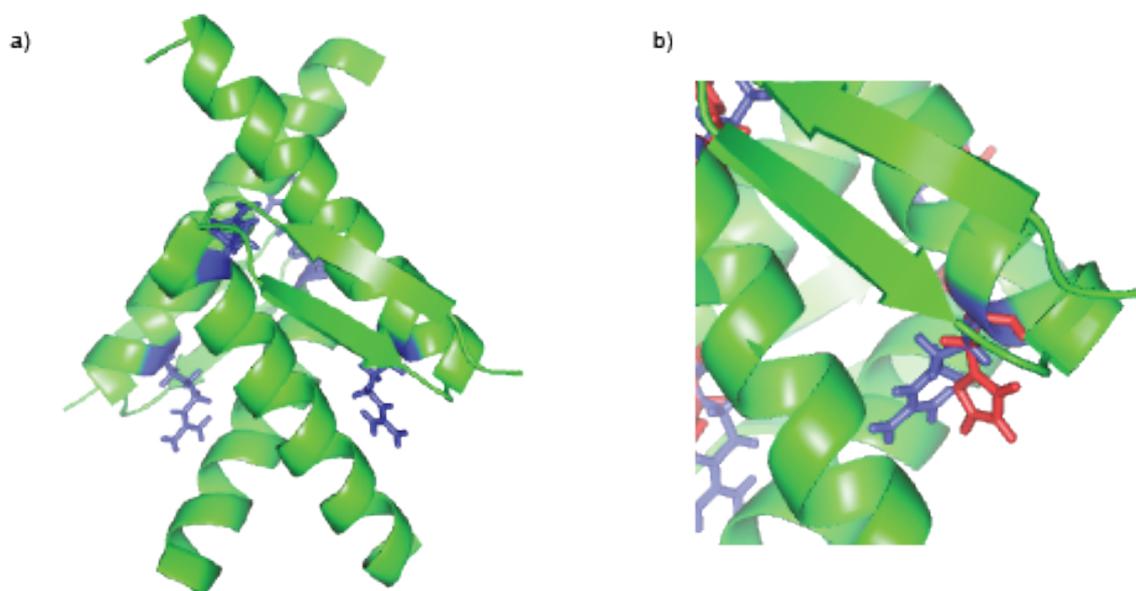


Figure 7.5 X-ray crystal structure of the p53 tetramerisation domain, highlighting residue 337. a) Full structure, Arginine-337 shown in blue, b) Zoom in showing residue 337 in the wild-type protein (arginine, blue) and in the R337H mutant (red).

The tetramerisation of p53 is controlled by a short, 30-mer, region of the p53 protein, the tetramerisation domain.²¹⁸ Each monomer within the tetramerisation domain consists of an α -helix and a β -strand, which pack together to form the tetramer (**Figure 7.5a**). The surface of this tetramerisation domain of p53 is rich in acidic amino acid residues.

Alanine scanning mutation studies performed by the Fersht group showed which of the residues in the tetramerisation domain sequence were important for the tetramerisation,²¹⁸ implicating nine hydrophobic residues present in the core of the

tetramer. One of the amino acids implicated in this study was Arg-337; the R337A mutation resulted in a decreased melting temperature (T_m) of 39.2 °C compared to the wild-type protein. A histidine mutation in this position is the most frequently inherited mutation affecting p53 tetramerisation,²¹⁹ and results in increased chances of tumour growth in organisms possessing this mutation.²²⁰ This R337H mutant of the p53 tetramerisation domain was chosen as a starting point to see if it was possible to stabilise the quaternary structure of p53, using Ru(II)(bpy)₃ complexes (**Figure 7.6**).

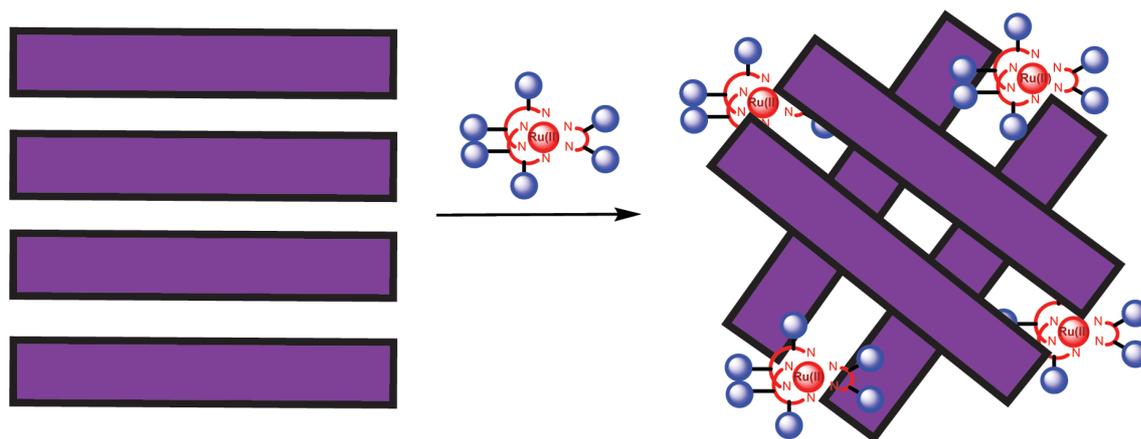


Figure 7.6 Cartoon depicting hypothesised binding of the Ru(II)(bpy)₃ complexes to the p53 tetramerisation domain, resulting in stabilisation of the tetrameric structure

7.5.2 Ru(II)(bpy)₃ complexes and peptides

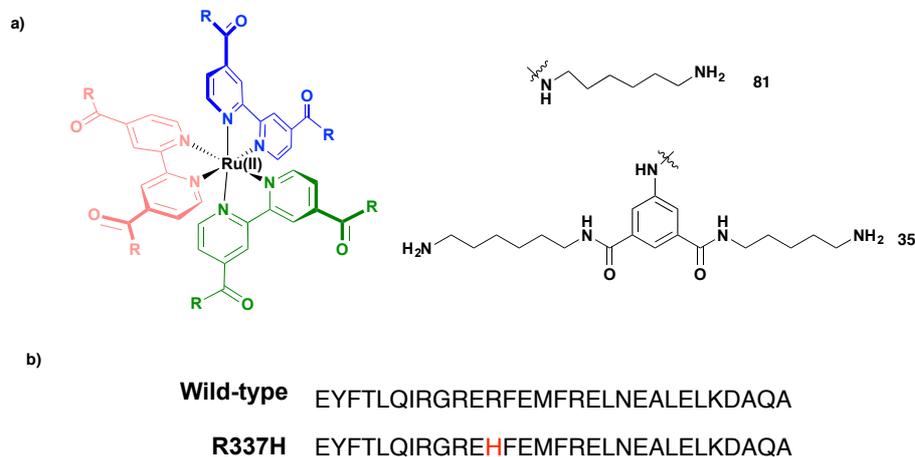


Figure 7.7 Ru(II)(bpy)₃ complexes and p53 tetramerisation domain peptides. a) Ru(II)(bpy)₃ complexes, b) Wild-type and R337H p53 peptides, showing the R337H mutation in red

A central 30-mer section of the p53 protein has been shown to be responsible for its tetramerisation,²¹⁸ therefore, it was decided to use synthetic peptides (**Figure 7.7b**) rather than protein to form the p53 tetramers. Peptides have also been used in a previous study of calixarene stabilisers of the p53 tetramerisation domain.²¹³ The use of peptides allowed

the purchase of two peptides, a wild-type (WT) and one possessing the R337H mutation, to allow a study of structural and binding differences between the two peptides with and without the Ru(II)(bpy)₃ complexes present. Using peptides allowed us to obtain the p53 tetramerisation domain quickly without the need to obtain a transcription vector or spend time optimising the protein preparation. Two amine functionalised Ru(II)(bpy)₃ complexes (**35** and **81**) (**Figure 7.7a**) previously synthesised were hypothesised to bind to the p53 tetramerisation domain, by analogy to lysine functionalised peptides which are known to bind, to the acidic surface of the p53 tetramerisation domain.²²¹

7.5.3 Circular dichroism

Circular dichroism can be used to probe the secondary structure of peptides and proteins, showing their α -helical, β -sheet and random coil content. Therefore this technique was used to probe how structured the p53 peptides are. In order to establish if it is possible to use the Ru(II)(bpy)₃ complexes to stabilise the p53 tetramerisation domain peptide structures, thermal unfolding of the structured peptides was performed, and the ellipticity at 222 nm (corresponding to α -helical content) followed and converted to fraction unfolded (with 1 corresponding to the minimum magnitude ellipticity and 0 corresponding to the maximum magnitude ellipticity)

7.5.3.1 In phosphate buffer

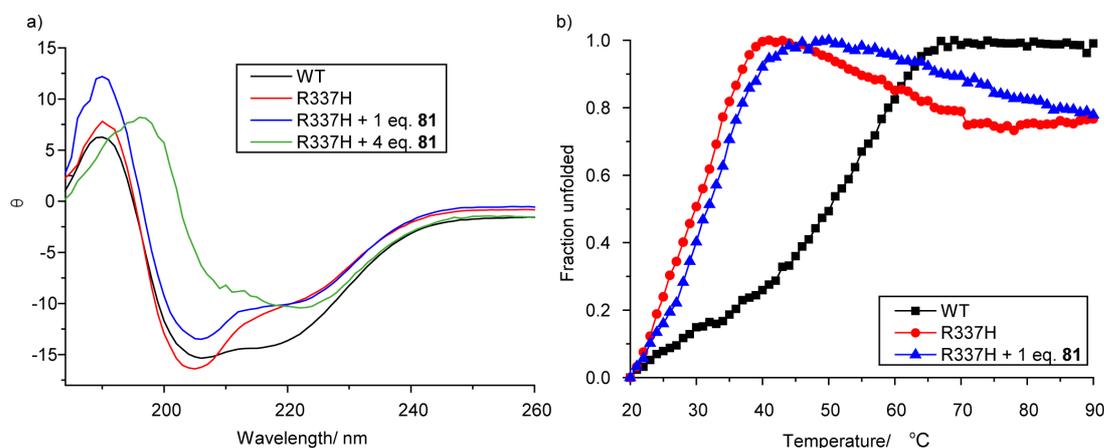


Figure 7.8 CD spectra in 5 mM sodium phosphate buffer, pH 7.5 a) CD spectra at 20 °C, b) Thermal melt, looking at ellipticity at 222 nm, corresponding to α -helical content. 100 μ M peptide and 100 μ M Ru(II)(bpy)₃ complex where appropriate

Initially the CD spectra were recorded in 5 mM sodium phosphate, pH 7.5 buffer, to see if it may be possible to stabilise the p53 tetramerisation domain in a low concentration buffer. CD spectra of the two peptides alone at 20 °C showed significant α -helical content for both (**Figure 7.8a**), however there are differences in the spectrum, indicating

differences in secondary structure between the two peptides. Addition of 1 equivalent of Ru(II)(bpy)₃ complex **81** to the R337H peptide shows a difference in the CD spectrum, indicating some change in secondary structure. Addition of 4 equivalent of Ru(II)(bpy)₃ complex **81** to the R337H peptide, again showed differences in the spectrum, however much of this is likely to arise from the absorbance of the Ru(II)(bpy)₃ complex **81** which was found to distort the spectra at these concentrations.

Looking at the thermal melt profile (**Figure 7.8b**) it can clearly be seen that the structure of the R337H peptide is destabilised compared to the wild type peptide, reducing the melting temperature (T_m) from ~52 °C to ~29 °C. On addition of 1 eq. of Ru(II)(bpy)₃ complex **81** to the R337H peptide, there is a very small shift in the thermal melt profile, however the T_m only shifts by ~1 °C which is not particularly significant.

7.5.3.2 *In water*

The CD spectra in 5 mM sodium phosphate, pH 7.5 show that both the p53 peptides are structured, however the addition of Ru(II)(bpy)₃ complex **81** shows little change in the T_m of the R337H peptide. Most previous work on synthetic stabilisers of the p53 tetramerisation domain had been performed in water,^{212,214,221,222} therefore attempts at the CD thermal melt study were also performed in water.

The CD spectra and thermal melt profiles of the wild type and R337H p53 peptides in water look similar to that in 5 mM phosphate buffer (**Figure 7.9**), with the wild-type peptide structure being stabilised by ~10 °C in water compared to phosphate buffer. Addition of 1 eq. of Ru(II)(bpy)₃ complex **81** (**Figure 7.9a** and **c**) looks to have increased the T_m of the R337H peptide by ~10 °C indicating stabilisation of the structure while the T_m of the wild-type peptide has been decrease by a similar amount, indicating destabilisation the structure. The potentially different responses of the Ru(II)(bpy)₃ complex to the two peptides indicate that it may be binding to the p53 tetramer in the region of the mutation, disrupting the interactions causing tetramerisation in the case of the wild-type peptide, and enhancing those interactions in the case of the R337H mutant peptide. Spectra with Ru(II)(bpy)₃ complex **35** indicate similar results (**Figure 7.9b** and **d**), however the data is a lot noisier, and should be repeated.

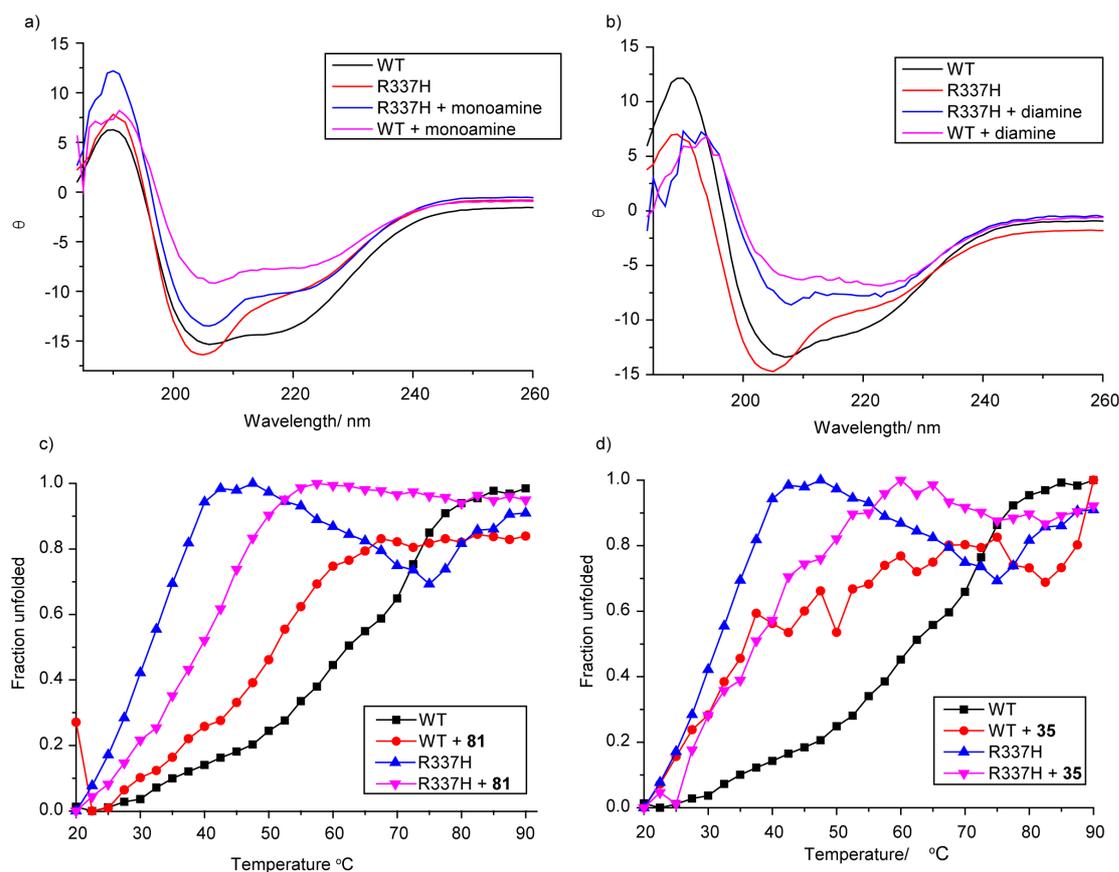


Figure 7.9 CD spectra in water a) CD spectra with Ru(II)(bpy)₃ complex **81** at 20 °C, b) CD spectra with Ru(II)(bpy)₃ complex **35** at 20 °C, c) Thermal melt profile with Ru(II)(bpy)₃ complex **81**, d) Thermal melt profile with Ru(II)(bpy)₃ complex **35**. All spectra taken in water, neutralised by addition of 1 mM HCl or 1 mM NH₄OH, with 100 μM peptide and 100 μM Ru(II)(bpy)₃ complex where appropriate

7.5.4 Conclusions and future work

Preliminary data, using CD thermal melt profiles, has been obtained, indicating that the Ru(II)(bpy)₃ complexes **35** and **81** can stabilise structure of the R337H mutant peptide of the p53 tetramerisation domain. These studies also indicate that these Ru(II)(bpy)₃ complexes also destabilise the structure of the wild-type peptide of the p53 tetramerisation domain. This indicates that the Ru(II)(bpy)₃ complexes can be used as PPI stabilisers, as well as PPI inhibitors.

In order to establish if the effects seen are valid, it is necessary to repeat the CD data, particularly for Ru(II)(bpy)₃ complex **35** in order to generate better thermal melt profiles. If the data does prove reproducible, further experiments will need to be performed in order to show: i) that it is indeed tetramers that do form, ii) an orthogonal experiment to show the stabilisation, iii) the stoichiometry of binding, iv) the binding affinity and v) the location of the binding site. Potential platforms for determining these factors include

native mass spectrometry, size exclusion chromatography, proteolytic digestion, luminescence anisotropy and protein NMR. Further work could also be performed to further destabilise the p53 tetramerisation domain, so that it is not structured at 20 °C, this would enable us to see if it is possible to induce protein structure on an unstructured peptide by addition of synthetic molecules.

8 References

- 1 S. P. Chumakov, Y. E. Kravchenko and P. M. Chumakov, *Mol. Biol.*, 2012, **46**, 627–638.
- 2 O. Keskin, A. Gursoy, B. Ma and R. Nussinov, *Chem. Rev.*, 2008, **108**, 1225–44.
- 3 J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001–9.
- 4 L. R. Whitby and D. L. Boger, *Acc. Chem. Res.*, 2012, **45**, 1698–709.
- 5 T. Geppert, S. Bauer, J. A. Hiss, E. Conrad, M. Reutlinger, P. Schneider, M. Weisel, B. Pfeiffer, K.-H. Altmann, Z. Waibler and G. Schneider, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 258–61.
- 6 T. Maurer, L. S. Garrenton, A. Oh, K. Pitts, D. J. Anderson, N. J. Skelton, B. P. Fauber, B. Pan, S. Malek, D. Stokoe, M. J. C. Ludlam, K. K. Bowman, J. Wu, A. M. Giannetti, M. A. Starovasnik, I. Mellman, P. K. Jackson, J. Rudolph, W. Wang and G. Fang, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 5299–304.
- 7 L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. A. Liu, *Science (80-.)*, 2004, **303**, 844–8.
- 8 A. Alimonti, C. Nardella, Z. Chen, J. G. Clohessy, A. Carracedo, L. C. Trotman, K. Cheng, S. Varmeh, S. C. Kozma, G. Thomas, E. Rosivatz, R. Woscholski, F. Cognetti, H. I. Scher and P. P. Pandolfi, *J. Clin. Invest.*, 2010, **120**, 681–693.
- 9 A. M. Petros, J. Dinges, D. J. Augeri, S. a Baumeister, D. a Betebenner, M. G. Bures, S. W. Elmore, P. J. Hajduk, M. K. Joseph, S. K. Landis, D. G. Nettlesheim, S. H. Rosenberg, W. Shen, S. Thomas, X. Wang, I. Zanze, H. Zhang and S. W. Fesik, *J. Med. Chem.*, 2006, **49**, 656–63.
- 10 I. Van Molle, A. Thomann, D. L. Buckley, E. C. So, S. Lang, C. M. Crews and A. Ciulli, *Chem. Biol.*, 2012, **19**, 1300–12.
- 11 K. Burgess, *Acc. Chem. Res.*, 2001, **34**, 826–35.
- 12 V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161–173.
- 13 H. Lingard, J. T. Han, A. L. Thompson, I. K. H. Leung, R. T. W. Scott, S. Thompson and A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 3650–3.
- 14 D. J. Yeo, S. L. Warriner and A. J. Wilson, *Chem. Commun.*, 2013, **49**, 9131–9133.
- 15 J. A. Miles, D. J. Yeo, P. Rowell, S. Rodriguez-Marin, C. M. Pask, S. L. Warriner, T. A. Edwards and A. J. Wilson, .
- 16 C. J. Brown, S. T. Quah, J. Jong, A. M. Goh, P. C. Chiam, K. H. Khoo, M. L. Choong, M. A. Lee, L. Yurlova, K. Zolghadr, T. L. Joseph, C. S. Verma and D. P. Lane, *ACS Chem. Biol.*, 2013, **8**, 506–512.
- 17 S. Fletcher and A. D. Hamilton, *Curr. Opin. Chem. Biol.*, 2005, **9**, 632–8.

- 18 A. J. Wilson, *Chem. Soc. Rev.*, 2009, **38**, 3289–300.
- 19 V. Martos, P. Castreño, J. Valero and J. de Mendoza, *Curr. Opin. Chem. Biol.*, 2008, **12**, 698–706.
- 20 M. Mammen, S.-K. Choi and G. M. Whitesides, *Angew. Chem. Int. Ed. Engl.*, 1998, **37**, 2754–2794.
- 21 S. B. Nimse and T. Kim, *Chem. Soc. Rev.*, 2013, **42**, 366–86.
- 22 Y. Hamuro, M. C. Calama, H. S. Park and A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.*, 1997, **36**, 2680–2683.
- 23 Q. Lin and A. Hamilton, *Comptes Rendus Chim.*, 2002, **5**, 441–450.
- 24 H. S. Park, Q. Lin and A. D. Hamilton, *J. Am. Chem. Soc.*, 1999, **121**, 8–13.
- 25 M. A. Blaskovich, Q. Lin, F. L. Delarue, J. Sun, H. S. Park, D. Coppola, A. D. Hamilton and S. M. Sebti, *Nat. Biotechnol.*, 2000, **18**, 1065–70.
- 26 P. B. Crowley, P. Ganji and H. Ibrahim, *ChemBioChem*, 2008, **9**, 1029–33.
- 27 S. Francese, A. Cozzolino, I. Caputo, C. Esposito, M. Martino, C. Gaeta, F. Troisi and P. Neri, *Tetrahedron Lett.*, 2005, **46**, 1611–1615.
- 28 S. A. Minaker, K. D. Daze, M. C. F. Ma and F. Hof, *J. Am. Chem. Soc.*, 2012, **134**, 11674–80.
- 29 O. Hayashida and M. Uchiyama, *Tetrahedron Lett.*, 2006, **47**, 4091–4094.
- 30 O. Hayashida and M. Uchiyama, *J. Org. Chem.*, 2007, **72**, 610–6.
- 31 O. Hayashida, N. Ogawa and M. Uchiyama, *J. Am. Chem. Soc.*, 2007, **129**, 13698–13705.
- 32 O. Hayashida and M. Uchiyama, *Org. Biomol. Chem.*, 2008, **6**, 3166–3170.
- 33 I. D. Acquarica, A. Cerreto, G. D. Monache, F. Subrizi, A. Boffi, A. Tafi, S. Forli, B. Botta and P. A. Moro, 2011, 4396–4407.
- 34 M. Rosenfeld and D. M. Surgenor, *J. Biol. Chem.*, 1949, **329**, 663–677.
- 35 G. H. Beaven, S. H. Chen, A. d'Albis and W. B. Gratzer, *Eur. J. Biochem.*, 1974, **41**, 539–546.
- 36 J. Davila and A. Harriman, *J. Am. Chem. Soc.*, 1990, **112**, 2686–2690.
- 37 C. Ader, R. Schneider, S. Hornig, P. Velisetty, E. M. Wilson, A. Lange, K. Giller, I. Ohmert, M.-F. Martin-Eauclaire, D. Trauner, S. Becker, O. Pongs and M. Baldus, *Nat. Struct. Mol. Biol.*, 2008, **15**, 605–612.
- 38 S. N. Gradl, J. P. Felix, E. Y. Isacoff, M. L. Garcia and D. Trauner, *J. Am. Chem. Soc.*, 2003, **125**, 12668–9.
- 39 V. Martos, S. C. Bell, E. Santos, E. Y. Isacoff, D. Trauner and J. De Mendoza, *Chem. Biol.*, 2009, **106**, 1–5.

- 40 S. N. Gradl, J. P. Felix, E. Y. Isacoff, M. L. Garcia and D. Trauner, *Ion Channels*, 2003, **1**, 12668–12669.
- 41 S. Hornig, I. Ohmert, D. Trauner, C. Ader, M. Baldus and O. Pongs, *Channels*, 2013, **7**, 473–482.
- 42 D. Aviezer, S. Cotton, M. David, A. Segev, N. Khaselev, N. Galili, Z. Gross and A. Yaron, *Cancer Res.*, 2000, **60**, 2973–2980.
- 43 R. K. Jain and A. D. Hamilton, *Org. Lett.*, 2000, **2**, 1721–3.
- 44 T. Aya and A. D. Hamilton, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 2651–2654.
- 45 S. S. Komath, K. Bhanu, B. G. Maiya and M. J. Swamy, *Biosci. Rep.*, 2000, **20**, 265–276.
- 46 K. Bhanu, S. S. Komath, B. G. Maiya and M. J. Swamy, *Curr. Sci.*, 1997, **73**, 598–602.
- 47 R. Kenoth, D. R. Reddy, B. G. Maiya and M. J. Swamy, *Eur. J. Biochem.*, 2001, **268**, 5541–5549.
- 48 C. Ader, R. Schneider, S. Hornig, P. Velisetty, V. Vardanyan, K. Giller, I. Ohmert, S. Becker, O. Pongs and M. Baldus, *EMBO J.*, 2009, **28**, 2825–2834.
- 49 D. Daly, A. Al-Sabi, G. K. K. Kinsella, K. Nolan and J. O. O. Dolly, *Chem. Commun.*, 2015, **51**, 1066–1069.
- 50 M. Goel, P. Anuradha, K. J. Kaur, B. G. Maiya, M. J. Swamy and D. M. Salunke, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2004, **60**, 281–288.
- 51 M. Goel, R. S. Damai, D. K. Sethi, K. J. Kaur, B. G. Maiya, M. J. Swamy and D. M. Salunke, *Biochemistry*, 2005, **44**, 5588–5596.
- 52 J. Matsumoto, S. Sato, T. Shiragami and M. Yasuda, *Dye. Pigment.*, 2015, **120**, 147–154.
- 53 S. Mignani, S. El Kazzouli, M. M. Bousmina and J.-P. Majoral, *Chem. Rev.*, 2014, **114**, 1327–42.
- 54 F. Chiba, T. Hu, L. J. Twyman and M. Wagstaff, *Chem. Commun.*, 2008, **4351–4353**, 4351–4353.
- 55 F. Chiba, G. Mann and L. J. Twyman, *Org. Biomol. Chem.*, 2010, **8**, 5056–5058.
- 56 J. Simpson, Y. Liu, W. A. Goddard, J. Giri and M. S. Diallo, *ACS Nanotechnol.*, 2011, **5**, 3456–3468.
- 57 A. J. Wilson, J. Hong, S. Fletcher and A. D. Hamilton, *Org. Biomol. Chem.*, 2007, **5**, 276–85.
- 58 D. K. Leung, Z. Yang and R. Breslow, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 5050–3.
- 59 K. Kano and Y. Ishida, *Angew. Chem. Int. Ed. Engl.*, 2007, **46**, 727–730.
- 60 M. Dörr and E. Meggers, *Curr. Opin. Chem. Biol.*, 2014, **19**, 76–81.
- 61 E. Meggers, *Chem. Commun.*, 2009, 1001–10.

- 62 P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler, S. L. Schreiber and A. Paul, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **107**, 18787–18792.
- 63 M. P. Coogan and V. Fernández-Moreira, *Chem. Commun.*, 2014, **50**, 384–99.
- 64 P. A. Frey and W. W. Cleland, *Bioorg. Chem.*, 1998, **26**, 175–192.
- 65 J. Porath, J. Carlsson, I. Olsson and G. Belfrage, *Nature*, 1975, **258**.
- 66 M. C. Smith, T. C. Furman, T. D. Ingolia and C. Pidgeon, *J. Biol. Chem.*, 1988, **263**, 7211–7215.
- 67 S. V Wegner and J. P. Spatz, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 7593–7596.
- 68 B. C. Roy, M. A. Fazal, S. Sun and S. Mallik, *Chem. Commun.*, 2000, **1**, 547–548.
- 69 M. A. Fazal, B. C. Roy, S. Sun, S. Mallik and K. R. Rodgers, *J. Am. Chem. Soc.*, 2001, **123**, 6283–90.
- 70 S. Sun, M. Abul Fazal, B. C. Roy and S. Mallik, *Org. Lett.*, 2000, **2**, 911–4.
- 71 S. Sun, M. A. Fazal, B. C. Roy, B. Chandra and S. Mallik, *Inorg. Chem.*, 2002, **41**, 1584–90.
- 72 Y. Mito-oka, S. Tsukiji, T. Hiraoka and N. Kasagi, *Tetrahedron Lett.*, 2001, **42**, 7059–7062.
- 73 A. Ojida, Y. Miyahara, T. Kohira and I. Hamachi, *Biopolymers*, 2004, **76**, 177–84.
- 74 A. Ojida, Y. Mito-Oka, M.-A. Inoue and I. Hamachi, *J. Am. Chem. Soc.*, 2002, **124**, 6256–8.
- 75 A. Ojida, M. Inoue, Y. Mito-Oka and I. Hamachi, *J. Am. Chem. Soc.*, 2003, **125**, 10184–5.
- 76 A. Ojida, Y. Mito-Oka, K. Sada and I. Hamachi, *J. Am. Chem. Soc.*, 2004, **126**, 2454–2463.
- 77 A. Ojida and I. Hamachi, *Bull. Chem. Soc. Jpn.*, 2006, **79**, 35–46.
- 78 T. Anai, E. Nakata, Y. Koshi, A. Ojida and I. Hamachi, *J. Am. Chem. Soc.*, 2007, **129**, 6233–6239.
- 79 A. Ojida, M. Inoue, Y. Mito-oka, H. Tsutsumi, K. Sada and I. Hamachi, *J. Am. Chem. Soc.*, 2006, **128**, 2052–8.
- 80 Y. Ishida, M. Inoue, T. Inoue, A. Ojida and I. Hamachi, *Chem. Commun.*, 2009, 2848–2850.
- 81 J. H. Kang, H. J. Kim, T.-H. Kwon and J.-I. Hong, *J. Org. Chem.*, 2014.
- 82 J. A. Drewry, S. Fletcher, P. Yue, D. Marushchak, W. Zhao, S. Sharmeen, X. Zhang, A. D. Schimmer, C. Gradinaru, J. Turkson and P. T. Gunning, *Chem. Commun.*, 2010, **46**, 892–4.

- 83 J. A. Drewry, E. Duodu, A. Mazouchi, P. Spagnuolo, S. Burger, C. C. Gradinaru, P. Ayers, A. D. Schimmer and P. T. Gunning, *Inorg. Chem.*, 2012, **51**, 8284–91.
- 84 K. K. Clark-Ferris and J. Fisher, *J. Am. Chem. Soc.*, 1985, **107**, 5007–5008.
- 85 K. C. Cho, C. M. Che, K. M. Ng and C. L. Choy, *J. Am. Chem. Soc.*, 1986, **108**, 2814–2818.
- 86 K. C. Cho, K. M. Ng, C. L. Choy and C. M. Che, *Chem. Phys. Lett.*, 1986, 129, 521–525.
- 87 J. S. Zhou, E. S. V. Granada, N. B. Leontis and M. A. J. Rodgers, *J. Am. Chem. Soc.*, 1990, **112**, 5074–5080.
- 88 J. S. Zhou and M. A. J. Rodgers, *J. Am. Chem. Soc.*, 1991, **113**, 6237–6243.
- 89 R. W. Larsen, D. H. Omdal, R. Jasuja, S. L. Niu and D. M. Jameson, *J. Phys. Chem. B*, 1997, **101**, 8012–8020.
- 90 M. Aoudia and M. A. J. Rodgers, *J. Am. Chem. Soc.*, 1997, **119**, 12859–12868.
- 91 J. C. Croney, M. K. Helms, D. M. Jameson and R. W. Larsen, *J. Phys. Chem. B*, 2000, **104**, 973–977.
- 92 M. Aoudia, A. B. Guliaev, N. B. Leontis and M. A. J. Rodgers, *Biophys. Chem.*, 2000, **83**, 121–140.
- 93 M. Aoudia and M. A. J. Rodgers, *Langmuir*, 2005, **21**, 10355–10361.
- 94 H. Zhou, L. Baldini, J. Hong, A. J. Wilson and A. D. Hamilton, *J. Am. Chem. Soc.*, 2006, **128**, 2421–5.
- 95 L. Baldini, A. J. Wilson, J. Hong and A. D. Hamilton, *J. Am. Chem. Soc.*, 2004, **126**, 5656–7.
- 96 R. K. Jain and A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.*, 2002, **41**, 641–643.
- 97 E. Sedlák, D. Fedunová, V. Veselá, D. Sedláková and M. Antalík, *Biomacromolecules*, 2009, **10**, 2533–2538.
- 98 A. J. Wilson, K. Groves, R. K. Jain, H. S. Park and A. D. Hamilton, *J. Am. Chem. Soc.*, 2003, **125**, 4420–1.
- 99 K. Groves, A. J. Wilson and A. D. Hamilton, *J. Am. Chem. Soc.*, 2004, **126**, 12833–42.
- 100 R. F. Pasternack, L. Francesconi, D. O. N. Raff and E. Spiro, *Inorg. Chem.*, 1973, **12**, 2606–2611.
- 101 R. F. Pasternack, P. R. Huber, B. P. G. Engasser, L. Francesconi, E. Gibbs, P. Fasella and G. Cerio Venturo, *J. Am. Chem. Soc.*, 1971, **669**, 4511–4517.
- 102 S. Fletcher and A. D. Hamilton, *New J. Chem.*, 2007, **31**, 623.
- 103 O. Kokhan, N. Ponomarenko, P. R. Pokkuluri, M. Schiffer and D. M. Tiede, *Biochemistry*, 2014, **53**, 5070–5079.
- 104 D. Paul, H. Miyake, S. Shinoda and H. Tsukube, *Chem. - A Eur. J.*, 2006, **12**, 1328–38.

- 105 H. Azuma, Y. Yoshida, D. Paul, S. Shinoda, H. Tsukube and T. Nagasaki, *Org. Biomol. Chem.*, 2009, **7**, 1700–1704.
- 106 S. Dieckmann, R. Riedel, K. Harms and E. Meggers, *Eur. J. Inorg. Chem.*, 2012, **2012**, 813–821.
- 107 S. Mollin, S. Blanck, K. Harms and E. Meggers, *Inorganica Chim. Acta*, 2012, **393**, 261–268.
- 108 L. Feng, Y. Geisselbrecht, S. Blanck, A. Wilbuer, G. E. Atilla-Gokcumen, P. Filippakopoulos, K. Kräling, M. A. Celik, K. Harms, J. Maksimoska, R. Marmorstein, G. Frenking, S. Knapp, L.-O. Essen and E. Meggers, *J. Am. Chem. Soc.*, 2011, **133**, 5976–86.
- 109 A. Wilbuer, D. H. Vlecken, D. J. Schmitz, K. Kräling, K. Harms, C. P. Bagowski and E. Meggers, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 3839–42.
- 110 J. Maksimoska, D. S. Williams, G. E. Atilla-Gokcumen, K. S. M. Smalley, P. J. Carroll, R. D. Webster, P. Filippakopoulos, S. Knapp, M. Herlyn and E. Meggers, *Chem. - A Eur. J.*, 2008, **14**, 4816–4822.
- 111 D. S. Williams, P. J. Carroll and E. Meggers, *Inorg. Chem.*, 2007, **46**, 2944–6.
- 112 H. Bregman, P. J. Carroll and E. Meggers, *J. Am. Chem. Soc.*, 2006, **128**, 877–84.
- 113 J. Debreczeni and A. Bullock, *Angew. Chem. Int. Ed. Engl.*, 2006, **45**, 1580–1585.
- 114 G. E. Atilla-Gokcumen, L. Di Costanzo and E. Meggers, *J. Biol. Inorg. Chem.*, 2011, **16**, 45–50.
- 115 P. Xie, C. Streu, J. Qin, H. Bregman, N. Pagano, E. Meggers and R. Marmorstein, *Biochemistry*, 2009, **48**, 5187–5198.
- 116 J. Maksimoska, L. Feng, K. Harms, C. Yi, J. Kissil, R. Marmorstein and E. Meggers, *J. Am. Chem. Soc.*, 2008, **130**, 15764–5.
- 117 G. Atilla-Gokcumen and N. Pagano, *ChemBioChem*, 2008, **9**, 2933–2936.
- 118 K. Wähler, K. Kräling, H. Steuber and E. Meggers, *ChemistryOpen*, 2013, **2**, 180–5.
- 119 M. Streib, K. Kräling, K. Richter, X. Xie, H. Steuber and E. Meggers, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 305–9.
- 120 J. Xie, X. Chen, Z. Huang and T. Zuo, *J. Mol. Model.*, 2015, **21**, 140.
- 121 C.-H. Leung, H.-J. Zhong, D. S.-H. Chan and D.-L. Ma, *Coord. Chem. Rev.*, 2013, **257**, 1764–1776.
- 122 M. M. He, A. S. Smith, J. D. Oslob, W. M. Flanagan, A. C. Braisted, A. Whitty, M. T. Cancilla, J. Wang, A. A. Lugovskoy, J. C. Yoburn, A. D. Fung, G. Farrington, J. K. Eldredge, E. S. Day, L. A. Cruz, T. G. Cachero, S. K. Miller, J. E. Friedman, I. C. Choong and B. C. Cunningham, *Science (80-.)*, 2005, **310**, 1022–5.
- 123 T. Kang, Z. Mao, C. Ng, M. Wang, W. Wang, C. Wang, S. M. Lee, Y. Wang, C. Leung and D. Ma, *J. Med. Chem.*, 2016.

- 124 D.-L. Ma, L.-J. Liu, K.-H. Leung, Y.-T. Chen, H.-J. Zhong, D. S.-H. Chan, H.-M. D. Wang and C.-H. Leung, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 9178–9182.
- 125 L.-J. Liu, B. He, J. A. Miles, W. Wang, Z. Mao, W. I. Che, J.-J. Lu, X.-P. Chen, A. J. Wilson, D.-L. Ma and C.-H. Leung, *Oncotarget*, 2016.
- 126 C.-H. Leung, M. Dik-Lung, H.-J. Zhong, L. Lu, C. Wong, C. Peng, S.-C. Yan, Z. Cai and H.-M. Wang, *Chem. Sci.*, 2015, **6**, 5400–5408.
- 127 B. Y.-W. Man, H.-M. Chan, C.-H. Leung, D. S.-H. Chan, L.-P. Bai, Z.-H. Jiang, H.-W. Li and D.-L. Ma, *Chem. Sci.*, 2011, **2**, 917.
- 128 F. P. Dwyer, E. C. Gyarfas, W. P. Rogers and J. H. Koch, *Nature*, 1952, **170**, 190–191.
- 129 F. P. Dwyer, E. Mayhew, E. M. F. Roe and A. Shulman, *Br. J. Cancer*, 1965, **19**, 195.
- 130 S. Sakai and T. Sasaki, *J. Am. Chem. Soc.*, 1994, **116**, 8295–8296.
- 131 S. S. Y. Shigemasa and S. T, *Bull. Chem. Soc. Jpn.*, 1999, **72**, 1313–1319.
- 132 P. Reeh and J. De Mendoza, *Chem. - A Eur. J.*, 2013, **19**, 5259–62.
- 133 T. Hasegawa, T. Yonemura, K. Matsuura and K. Kobayashi, *Tetrahedron Lett.*, 2001, **42**, 3989–3992.
- 134 T. Hasegawa, T. Yonemura, K. Matsuura and K. Kobayashi, *Bioconjug. Chem.*, 2003, **14**, 728–737.
- 135 R. Kikkeri, I. García-Rubio and P. H. Seeberger, *Chem. Commun.*, 2009, **2**, 235–237.
- 136 R. Kikkeri, D. Grünstein and P. H. Seeberger, *J. Am. Chem. Soc.*, 2010, **132**, 10230–10232.
- 137 R. Kikkeri, F. Kamena, T. Gupta, L. H. Hossain, S. Boonyarattanakalin, G. Gorodyska, E. Beurer, G. Coullerez, M. Textor and P. H. Seeberger, *Langmuir*, 2010, **26**, 1520–1523.
- 138 D. Grünstein, M. Maglinao, R. Kikkeri, M. Collot, K. Barylyuk, B. Lepenies, F. Kamena, R. Zenobi and P. H. Seeberger, *J. Am. Chem. Soc.*, 2011, **133**, 13957–13966.
- 139 T. Okada, T. Makino and N. Minoura, *Bioconjug. Chem.*, 2009, **20**, 1296–1298.
- 140 K. C. Cho, C. M. Che, F. C. Cheng and C. L. Choy, *J. Am. Chem. Soc.*, 1984, **106**, 6843–6844.
- 141 H. Takashima, S. Shinkai and I. Hamachi, *Chem. Commun.*, 1999, 2345–2346.
- 142 J. Muldoon, A. E. Ashcroft and A. J. Wilson, *Chem. - A Eur. J.*, 2010, **16**, 100–3.
- 143 M. H. Filby, J. Muldoon, S. Dabb, N. C. Fletcher, A. E. Ashcroft and A. J. Wilson, *Chem. Commun.*, 2011, **47**, 559–61.
- 144 A. J. Wilson, J. R. Ault, M. H. Filby, H. I. A. Philips, A. E. Ashcroft and N. C. Fletcher, *Org. Biomol. Chem.*, 2013, **11**, 2206–12.
- 145 S. J. Turrell, M. H. Filby, A. Whitehouse and A. J. Wilson, *Bioorg. Med. Chem. Lett.*,

- 2012, **22**, 985–8.
- 146 J. Ohkanda, R. Satoh and N. Kato, *Chem. Commun.*, 2009, 6949–51.
- 147 Y. Yamaguchi, N. Kato, H. Azuma, T. Nagasaki and J. Ohkanda, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 2354–8.
- 148 S. A. Kang, P. J. Marjavaara and B. R. Crane, *J. Am. Chem. Soc.*, 2004, **126**, 10836–10837.
- 149 X. Liu, C. N. Kim, J. Yang, R. Jemmerson and X. Wang, *Cell*, 1996, **86**, 147–57.
- 150 H. Zou, W. J. Henzel, X. Liu, a Lutschg and X. Wang, *Cell*, 1997, **90**, 405–13.
- 151 P. Li, D. Nijhawan, I. Budihardjo, S. M. Srinivasula, M. Ahmad, E. S. Alnemri and X. Wang, *Cell*, 1997, **91**, 479–89.
- 152 A. N. Volkov, P. Nicholls and J. A. R. Worrall, *Biochim. Biophys. Acta*, 2011, **1807**, 1482–503.
- 153 H. Pelletier and J. Kraut, *Science (80-.)*, 1992, **258**, 1748–55.
- 154 L. B. Vitello and J. E. Erman, *Arch. Biochem. Biophys.*, 1987, **258**, 621–629.
- 155 J. E. Erman, G. C. Kresheck, L. B. Vitello and M. A. Miller, *Biochemistry*, 1997, **36**, 4054–4060.
- 156 A. Spencer, I. Evans, A. Spencer and G. Wilkinson, *J.C.S. Dalt.*, 1973, 204–209.
- 157 J. E. Erman and L. B. Vitello, *J. Biol. Chem.*, 1980, **255**, 6224–6227.
- 158 H. B. Bull and K. Breese, *Biochem. Biophys. Res. Commun.*, 1966, **24**, 74–78.
- 159 R. W. Shaw and C. R. Hartzell, *Biochemistry*, 1976, **15**, 1909–1914.
- 160 G. W. Bushell, G. V Louie and G. D. Brayer, *J. Mol. Biol.*, 1990, **214**, 585–595.
- 161 W. Liu, J. Rumbley, S. W. Englander and A. J. Wand, *Protein Sci.*, 2003, **12**, 2104–2108.
- 162 J.-M. Lehn, *Chem. - A Eur. J.*, 1999, **5**, 2455–2463.
- 163 S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders and J. F. Stoddart, *Angew. Chem. Int. Ed. Engl.*, 2002, **41**, 898–952.
- 164 A. D. Corbett, J. D. Cheeseman, R. J. Kazlauskas and J. L. Gleason, *Angew. Chem. Int. Ed. Engl.*, 2004, **43**, 2432–6.
- 165 S. Ladame, *Org. Biomol. Chem.*, 2008, **6**, 219–26.
- 166 M. Mondal and A. K. H. Hirsch, *Chem. Soc. Rev.*, 2014, **44**, 2455–2488.
- 167 M. Sindelar, T. A. Lutz, M. Petrera and K. T. Wanner, *J. Med. Chem.*, 2013, **56**, 1323–1340.
- 168 V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano and M. F. Greaney,

Nat. Chem., 2010, **2**, 490–7.

- 169 R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero and O. Ramström, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 589–593.
- 170 B. Shi, R. Stevenson, D. J. Campopiano and M. F. Greaney, *J. Am. Chem. Soc.*, 2006, **128**, 8459–8467.
- 171 O. Ramström and J. Lehn, *ChemBioChem*, 2000, **1**, 41–48.
- 172 I. K. H. Leung, M. Demetriades, A. P. Hardy, C. Lejeune, T. J. Smart, A. Szöllössi, A. Kawamura, C. J. Schofield and T. D. W. Claridge, *J. Med. Chem.*, 2013, **56**, 547–555.
- 173 H. I. A. Phillips, A. V. Chernikov, N. C. Fletcher, A. E. Ashcroft, J. R. Ault, M. H. Filby and A. J. Wilson, *Chem. - A Eur. J.*, 2012, **18**, 13733–13742.
- 174 A. M. Timperio, S. Rinalducci and L. Zolla, *Bioorg. Chem.*, 2005, **33**, 459–69.
- 175 E. T. Kool, D.-H. Park and P. Crisalli, .
- 176 P. Crisalli and E. T. Kool, *J. Org. Chem.*, 2013, **78**, 1184–1189.
- 177 M. Rashidian, M. M. Mahmoodi, R. Shah, J. K. Dozier, C. R. Wagner and M. D. Distefano, *Bioconjug. Chem.*, 2013, **24**, 333–342.
- 178 L. You, D. Zha and E. V Anslyn, *Chem. Rev.*, 2015, **115**, 7840–7892.
- 179 A. Bajaj, O. R. Miranda, I.-B. Kim, R. L. Phillips, D. J. Jerry, U. H. F. Bunz and V. M. Rotello, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10912–10916.
- 180 C.-C. You, O. R. Miranda, B. Gider, P. S. Ghosh, I.-B. Kim, B. Erdogan, S. A. Krovi, U. H. F. Bunz and V. M. Rotello, *Nat. Nanotechnol.*, 2007, **2**, 318–323.
- 181 O. R. Miranda, C.-C. You, R. Phillips, I.-B. Kim, P. S. Ghosh, U. H. F. Bunz and V. M. Rotello, *J. Am. Chem. Soc.*, 2007, **129**, 9856–9857.
- 182 M. De, S. Rana, H. Akpinar, O. R. Miranda, R. R. Arvizo, U. H. F. Bunz and V. M. Rotello, *Nat. Chem.*, 2009, **1**, 461–465.
- 183 T. A. Egelhofer, A. Minoda, S. Klugman, K. Lee, P. Kolasinska-Zwierz, A. A. Alekseyenko, M.-S. Cheung, D. S. Day, S. Gadel, A. A. Gorchakov, T. Gu, P. V. Kharchenko, S. Kuan, I. Latorre, D. Linder-Basso, Y. Luu, Q. Ngo, M. Perry, A. Rechtsteiner, N. C. Riddle, Y. B. Schwartz, G. a Shanower, A. Vielle, J. Ahringer, S. C. R. Elgin, M. I. Kuroda, V. Pirrotta, B. Ren, S. Strome, P. J. Park, G. H. Karpen, R. D. Hawkins and J. D. Lieb, *Nat. Struct. Mol. Biol.*, 2011, **18**, 91–3.
- 184 A. T. Wright, M. J. Griffin, Z. Zhong, S. C. McCleskey, E. V. Anslyn and J. T. McDevitt, *Angew. Chem. Int. Ed. Engl.*, 2005, **44**, 6375–6378.
- 185 D. Zamora-Olivares, T. S. Kaoud, K. N. Dalby and E. V Anslyn, *J. Am. Chem. Soc.*, 2013, **135**, 14814–20.
- 186 S. Rana, A. K. Singla, A. Bajaj, S. Gokhan Elci, O. R. Miranda, R. Mout, B. Yan, F. R. Jirik and V. M. Rotello, *ACS Nano*, 2012, **6**, 8233–8240.
- 187 B. Peacor, C. Ramsay and M. Waters, *Chem. Sci.*, 2017, **8**, 1422–1428.

- 188 S. Stewart, M. A. Ivy and E. V Anslyn, *Chem. Soc. Rev.*, 2014, **43**, 70–84.
- 189 D. Coomans and D. L. Massart, *Chapter 8 Hard Modelling in Supervised Pattern Recognition*, 1992.
- 190 C. Chatfield and A. Collins, *Introduction to Multivariate Analysis*, 1980.
- 191 J. H. Friedman, *J. Am. Stat. Assoc.*, 1989, **84**, 165–175.
- 192 B. Rout, L. Unger, G. Armony, M. A. Iron and D. Margulies, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 12477–12481.
- 193 J. Hatai, L. Motiei and D. Margulies, *J. Am. Chem. Soc.*, 2017, **139**, 2136–2139.
- 194 A. Song, T. Wu, M. Zhang and T. Shen, *Chem. Lett.*, 1998, **27**, 907–908.
- 195 N. Klonisl, A. H. A. Clayton, E. W. Voss and W. H. Sawyer, *Photochem. Photobiol.*, 1998, **67**, 500–51.
- 196 E. Martin, A. Pardo, M. S. Guijarro and J. I. Fernandez-Alonso, *J. Mol. Struct.*, 1986, **142**, 197–200.
- 197 D. Margulies, C. E. Felder, G. Melman and A. Shanzer, *J. Am. Chem. Soc.*, 2007, **129**, 347–354.
- 198 Y. Nissinkorn, N. Lahav-Mankovski, A. Rabinkov, S. Albeck, L. Motiei and D. Margulies, *Chem. - A Eur. J.*, 2015, **21**, 15873.
- 199 T. Sarkar, K. Selvakumar, L. Motiei and D. Margulies, *Nat. Commun.*, 2016, **7**, 11374.
- 200 M. Zhou, G. P. Robertson and J. Roovers, *Inorg. Chem.*, 2005, **44**, 8317–25.
- 201 P. G. Hoertz, A. Staniszewski, A. Marton, G. T. Higgins, C. D. Incarvito, A. L. Rheingold and G. J. Meyer, *J. Am. Chem. Soc.*, 2006, **128**, 8234–45.
- 202 B. F. Van Gelder and E. C. Slater, *Biochim. Biophys. Acta*, 1962, **58**, 593–595.
- 203 E. Margoliash and N. Frohwirt, *J. Biol. Chem.*, 1959, **71**, 570–572.
- 204 R. C. C. Davies, A. Neuberger and B. M. Wilson, *Biochim. Biophys. Acta*, 1969, **178**, 294–305.
- 205 M. Volini and P. Tobias, *J. Biol. Chem.*, 1969, **244**, 5105–5109.
- 206 ThermoScientific, *Tech Tip #6*, 2003.
- 207 C. N. Pace, F. Vajdos, L. Fee, G. Grimsley and T. Gray, *Protein Sci.*, 1995, **4**, 2411–2423.
- 208 P. Atkins and J. de Paula, *Atkin's Physical Chemistry*, 2009.
- 209 P. Brezonik and W. Arnold, *Water Chemistry: An Introduction to the Chemistry of Natural and Engineered Aquatic Systems*, Oxford University Press, 2011.
- 210 W. Zecher and F. Krohnke, *Chem. Berichte-Recueil*, 1961, **94**, 690–697.

- 211 F. Krohnke, W. Zecher, J. Curtze, D. Drechsler, K. Pflgar, K. E. Schnalke and W. Weis, *Angew. Chem. Int. Ed. Engl.*, 1962, **1**, 626–632.
- 212 S. Gordo, V. Martos, E. Santos, M. Mené Ndez, C. Bo, E. Giralt and J. De Mendoza, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 16426–16431.
- 213 R. Kamada, W. Yoshino, T. Nomura, Y. Chuman, T. Imagawa, T. Suzuki and K. Sakaguchi, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 4412–4415.
- 214 X. Salvatella, M. Martinell, M. Gairí, M. G. Mateu, M. Feliz, A. D. Hamilton, J. De Mendoza and E. Giralt, *Angew. Chem. Int. Ed. Engl.*, 2004, **43**, 196–198.
- 215 R. Gabizon, T. Brandt, S. Sukenik, N. Lahav, M. Lebendiker, D. E. Shalev, D. Veprintsev and A. Friedler, *PLoS One*, 2012, **7**.
- 216 P. Chène and P. Che, *Oncogene*, 2001, **393**, 2611–2617.
- 217 A. J. Levine and M. Oren, *Nat. Rev. Cancer*, 2009, **9**, 749–58.
- 218 M. G. Mateu and A. R. Fersht, *EMBO J.*, 1998, **17**, 2748–2758.
- 219 M. I. W. Achatz, M. Olivier, F. Le Calvez, G. Martel-Planche, A. Lopes, B. M. Rossi, P. Ashton-Prolla, R. Giugliani, E. I. Palmero, F. R. Vargas, J. C. C. Da Rocha, A. L. Vettore and P. Hainaut, *Cancer Lett.*, 2007, **245**, 96–102.
- 220 B. C. Figueiredo, R. Sandrini, G. P. Zambetti, R. M. Pereira, C. Cheng, W. Liu, L. Lacerda, M. A. Pianovski, E. Michalkiewicz, J. Jenkins, C. Rodriguez-Galindo, M. J. Mastellaro, S. Vianna, F. Watanabe, F. Sandrini, S. B. I Arram, P. Boffetta and R. C. Ribeiro, *J Med Genet*, 2006, **43**, 91–96.
- 221 M. Martinell, X. Salvatella, J. Fernández-Carneado, S. Gordo, M. Feliz, M. Menéndez and E. Giralt, *ChemBioChem*, 2006, **7**, 1105–1113.
- 222 R. Gabizon, T. Brandt, S. Sukenik, N. Lahav, M. Lebendiker, D. E. Shalev, D. Veprintsev and A. Friedler, *PLoS One*, 2012, **7**, e38060.