

Development of Polypyridine Metal- Dependent Switches as Artificial Regulation Sites

by

Emmanuel Oheix

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY



School of Chemistry
College of Engineering and Physical Sciences
The University of Birmingham
September 2013

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This thesis describes the design, preparation and study of synthetic peptides dimerised with polypyridine linkers. Their conformational transitions, which result from the coordination of metal-ions, are proposed to trigger the orientation of their peptide substituents, thus allowing for the design of conjugates incorporating artificial regulation sites. To test this hypothesis, different polypyridine linkers were designed, such that the substituents orientation is either dependent or independent of the polypyridine conformations. The approach developed herein includes a detailed conformational study of low-molecular weight species (model conjugates) and the preparation of conjugates mimicking a natural protein, for which dimerisation is essential for its bioactivity. For the latter conjugates, which are based on the GCN4 transcription factor, the influence of metal addition on sequence-specific DNA binding was tested using a combination of spectroscopic and electrophoretic techniques. The results presented herein indicate that metal addition can influence the interaction of the polypyridine peptide conjugates with DNA, depending on the peptides and linkers design, and that this can be partly attributed to a conformational transition of the polypyridine linker. As an extension, the potential of these conjugates as sequence-specific DNA sensors or nuclease agents was partly investigated.

Acknowledgements

This thesis both the work and the writing would not have been possible without the contribution of countless persons. I will try to not forget anybody but I apologise if I do. First, I am deeply grateful to my supervisor, Dr Anna F. A. Peacock, for her support, patience, availability, the countless advices she gave to me, and for welcoming me in her group. Thanks also go to my second supervisor, Prof. Jim Tucker, for his helpful advice and lending me some equipment. I thank all my Peacock group colleagues (past and present) and more particularly Zhong, with whom I started this PhD and shared a lot of things, Matt, who has always been a helpful and reliable colleague, and Gemma, whose culinary and administrative skills were both really appreciated. Even though their work does not appear in this thesis, I thank our collaborators Lee A Gethings, Dr Mark T. Oakley, and Prof. Roy L. Johnston, with whom I enjoyed working. I would like to thank my viva examiners for spending some time reading my thesis.

In addition, I thank the people who spend time to teach me analytical, spectroscopic or computational techniques: Dr John Zhao, Dr Ariel Mucha, and Dr Eva Hyde (electrophoresis); Graham Burns and Dr Chi Tsang (HPLC); Dr Neil Spencer (NMR); Nick May and Dr Peter Ashton (mass spectrometry); Dr Jean-Louis Duprey (circular dichroism); Dr John Wilkie (molecular dynamics). I thank all the staff of the school of chemistry, and more particularly Stuart Arkless and Norihan M Taib, for their deep support.

I was lucky enough to meet plenty of nice people within the department of chemistry and want to thank all my PhD colleagues with special thought for Bick and Bene (les darons quoi), Ornella, Mahrez, Rory, Dave, Marie, Andrea, Jack, Pete, Bawbag, Yanouk, Antoine, Tom, and Cris, with whom I spent plenty of nice time. Thanks to Michel for the funky touch and the nice moment spent. Thanks to Antoine for his vigorous vituperation never void of virtue, and for all the amazing discussion. Thanks to Antoine for his spontaneous serenade

and sobriquet soliciting my smile, and all the things we shared. Thanks to the beautiful and colourful Carlotta for her generosity. With such a quality, you don't need many others to be amazing.

I also thank my other Birmingham friends: Myrto gave me plenty of nice advice, Simon, Dani and Sofia were all nice homebased supports during my first years here. Alex was both a home and pub mate. Je n'oublie pas tout mes amis et ma famille qui m'ont supportés toutes ces années (dans les deux sens du terme) et m'excuse de ne donner que peu de nouvelles. Une pensée particulière pour papa, maman, gaby, samy et louisie qui sont la base de tout ce que j'entreprend.

Table of content

Table of content	1
List of Abbreviations	5
CHAPTER I: LITERATURE REVIEW	8
1.1 – Introduction	9
1.2 – Stabilisation of peptide secondary structure in natural peptides and their derivatives	10
1.2.1 – α -Helices	10
1.2.2 – Other secondary structures	16
1.2.3 – Examples of higher order structure	20
1.3 – DNA binding peptides and artificial derivatives	24
1.3.1 – Common structural motifs of DNA bound proteins	24
1.3.2 – Artificial dimerisation and stitchery of basic domains	28
1.3.3 – Stabilisation and destabilisation of helices to enhance or regulate DNA binding	32
1.4 – The coordination chemistry of polypyridines	35
1.4.1 – Bipy metal complexes	35
1.4.2 – Terpy metal complexes	43
1.4.3 – Impact of polypyridine substituents on coordination	46
1.5 – Summary	47
1.6 – Project aims	48
1.7 – References	49
CHAPTER II: MODEL PEPTIDE SWITCHES AS ALLOSTERIC REGULATORS ..	67
2.1 – Introduction	68
2.1.1 – Artificial regulation sites based on polypyridine	68
2.1.2 – Aims of the chapter	73
2.2 – Results and discussion	74
2.2.1 – Cysteine and glutathione dimerised through a pyridine linker	74
2.2.1.1 – Design, preparation and stability	74
2.2.1.2 – UV studies of pyr-Br ₂ , cys ₂ pyr and pyr-GS ₂	78
2.2.1.3 – Cu(II) titration of cys ₂ pyr and pyr-GS ₂ monitored by UV spectroscopy	79
2.2.2 – Polypyridine linkers Me ₂ bipy and Me ₂ terpy	82
2.2.2.1 – Design and preparation of the polypyridine linkers	82
2.2.2.2 – Zn(II) binding studies for Me ₂ bipy and Me ₂ terpy monitored by ¹ H NMR spectroscopy	85
2.2.3 – Cysteine dimerisation with polypyridine linkers	88
2.2.3.1 – Preparation of polypyridine-cysteine conjugates	88
2.2.3.2 – Cu(II)/Zn(II) titration of cys ₂ bipy and cys ₂ terpy monitored by UV spectroscopy	90

2.2.4 – Glutathione model compounds.....	95
2.2.4.1 – Design and preparation.....	95
2.2.4.2 –pH, Cu(II), and Zn(II) titrations of bipy-GS ₂ and terpy-GS ₂ monitored by UV-visible spectroscopy.....	96
2.2.4.3 – Cu(II)/Zn(II) titration of bipy-GS ₂ and terpy-GS ₂ monitored by CD.....	104
2.2.4.3 – Zn(II) titration of bipy-GS ₂ and terpy-GS ₂ monitored by NMR	109
2.3 – Summary	115
2.4 – Experimental section.....	117
2.4.1 – Reagents and equipment.....	117
2.4.2 – Synthetic procedure and characterisation.....	118
2.4.3 – Analytical procedures.....	129
2.4.3.1 – Pyridine conjugates of cysteines and glutathione (related to section 2.2.1).	129
2.4.3.2 – Polypyridine linkers (related to section 2.2.2).....	130
2.4.3.3 – Bipy and terpy conjugates of cysteine (related to section 2.2.3).....	131
2.4.3.4 – Bipy and terpy conjugates of glutathione (related to section 2.2.4).....	131
2.5 – References.....	135

CHAPTER III: PEPTIDE SWITCHES BASED ON GCN4 AS POTENTIAL METAL-DEPENDENT DNA BINDERS141

3.1 – Introduction.....	142
3.1.1 – Artificial dimerisation of transcription factors.....	142
3.1.2 – Aims of the chapter	145
3.2 – Results and discussion	146
3.2.1 – Polypyridine-GCN4bd1 peptide conjugates.....	146
3.2.1.1 – Design of GCN4bd1 peptide	146
3.2.1.2 – Synthesis of GCN4bd1 peptide and polypyridine conjugates	148
3.2.1.3 – Cu(II) addition to GCN4bd1-polypyridine conjugates monitored by UV spectroscopy	151
3.2.2 – Polypyridine GCN4bd2 short peptide dimer.....	154
3.2.2.1 – Design of GCN4bd2 peptide	154
3.2.2.2 – Preparation and characterisation of GCN4bd2 and conjugates	157
3.2.2.3 – Cu(II) and Zn(II) addition to bipy(GCN4bd2) ₂ and terpy(GCN4bd2) ₂ monitored by UV spectroscopy	158
3.2.2.4 – Cu(II) and Zn(II) addition to bipy(GCN4bd2) ₂ and terpy(GCN4bd2) ₂ monitored by circular dichroism	164
3.3 – Summary	166
3.4 – Experimental.....	167
3.4.1 – Equipment and reagents	167
3.4.2 – Synthetic procedure and characterisation.....	168
3.4.3 – Analytical procedures.....	174
3.4.3.1 – Pyr, bipy and terpy conjugates of GCN4bd1 (related to section 3.2.1).....	174
3.4.3.2 – Bipy and terpy conjugates of GCN4bd2, and (GCN4bd2) ₂ (relative to section 3.2.2).....	175

3.5 – References.....	177
CHAPTER IV: DNA BINDING OF PEPTIDE DIMER CONJUGATES BASED ON GCN4	179
4.1 – Introduction.....	180
4.1.1 – Sequence specific DNA binding by b-ZIP peptides and derivatives	180
4.1.2 – Aims of the chapter	184
4.2 – Results and discussion	184
4.2.1 – DNA binding of GCN4bd1 and related conjugates	184
4.2.1.1 – CD studies of GCN4bd1 and conjugates in the presence or absence of the DNA target site.....	184
4.2.1.2 – Bandshift assay to estimate GCN4bd1 and conjugates affinity for CRE sites	186
4.2.1.3 - Influence of Cu(II) and Zn(II) on bipy(GCN4bd1) ₂ and terpy(GCN4bd1) ₂ folding in the presence of target DNA.....	188
4.2.2 – DNA binding of GCN4bd2 and related conjugates	189
4.2.2.1 – CD study of GCN4bd and conjugates	189
4.2.2.2 – Bandshift assay for GCN4bd2 conjugates.....	198
4.2.2.3 – Cu(II) and Zn(II) binding studies of bipy(GCN4bd2) ₂ and terpy(GCN4bd2) ₂ in the presence of DNA monitored by UV spectroscopy	199
4.3 – Summary	206
4.4 – Experimental.....	208
4.4.1 – Equipment and reagents	208
4.4.2 – Synthetic procedure and characterisation.....	209
4.4.3 – Analytical procedures.....	210
4.4.3.1 – Measurements on GCN4bd1 and related dimer conjugates (relative to section 4.2.1).....	210
4.4.3.2 – Measurements on GCN4bd2 dimer, and related conjugates (relative to section 4.2.2).....	214
4.5 – References.....	217
CHAPTER V: PRELIMINARY STUDIES ON SENSING AND NUCLEASE APPLICATIONS	220
5.1 – Introduction.....	221
5.1.1 – Sequence specific DNA sensors based on polypyridine and peptides	221
5.1.2 – Polypyridine peptide conjugates as sequence specific nuclease agents	224
5.1.3 – Aims of the chapter	225
5.2 – Results and discussion	225
5.2.1 – Polypyridine-peptide conjugates as sensors	225
5.2.1.1 – Polypyridine ligand fluorescence applied to DNA sensing.....	225
5.2.1.2 – Attempts at synthesising [Ru ^{II} (bipy) ₃] ²⁺ peptide conjugates and investigating their phosphorescent properties	228
5.2.2 – Polypyridine-GCN4bd peptides as sequence-specific nuclease agents	232

5.2.2.1 – Modelling of GCN4bd peptide conjugates	232
5.2.2.2 – Cleavage studies of supercoiled plasmid containing specific site by Cu:bipy(GCN4bd2) ₂	233
5.3 – Summary	237
5.4 – Experimental	238
5.4.1 – Equipment and reagents	238
5.4.2 – Synthetic procedure and characterisation	239
5.4.3 – Analytical procedures	243
5.4.3.1 – Polypyridine ligand fluorescence (related to section 5.2.1)	243
5.4.3.2 – Nuclease activity of bipy-(GCN4bd2) ₂ (related to section 5.2.2)	243
5.5 – References	244
CHAPTER VI: CONCLUSIONS AND FUTURE WORK	249
6.1 - Summary	250
6.2 – Future work	252
6.2.1 – Preparation of new linkers and related models	252
6.2.2 – Design of new peptides and conjugates	255
6.3 - References	256
APPENDICES: EXPERIMENTAL TECHNIQUES	257
A.1 – Spectroscopic techniques	258
A.1.1 – Electronic absorption or UV-visible spectroscopy	258
A.1.2 – Circular dichroism spectroscopy	260
A.1.3 – Emission spectroscopy	261
A.1.4 – Nuclear Magnetic Resonance spectroscopy	263
A.1.5 – Mass Spectrometry	266
A.2 – Electrophoretic techniques	267
A.2.1 – Study of protein-DNA interactions	267
A.2.2 – Cleavage of supercoiled plasmid DNA	269
A.3 – Miscellaneous methods	270
A.3.1 – Solid-phase peptide synthesis (SPPS)	270
A.3.2 – Reversed-phase high performance liquid chromatography	271
A.3.3 – C H N analysis	272
A.4 - References	272

List of Abbreviations

All abbreviations found in the text are summarised below. Non-common or specific abbreviations are displayed in bold. Proteinogenic amino-acids, variables and their respective units are not included in this list, however, they were abbreviated according to IUPAC norms. Compounds are sometimes referred to by their chemical formula.

(GCN4bd2)₂ = dimer designed in this work

[G28T_S]₂Fe = conjugate complex designed by Schepartz *et al.*

A = adenine

Aib = α -methylalnine

aPP = avian Pancreatic polypeptide

Bcl = B-cell leukemia

bd = basic domain

bHLH = basic helix-loop-helix proteins

bHLHZ = basic helix-loop-helix zipper proteins

bipy = 2,2'-bipyridine

bipy(GCN4bd1)₂ = conjugate designed in this work

bipy(GCN4bd2)₂ = conjugate designed in this work

bipy-Br₂ = 5,5'-dimethyl-2,2'-bipyridine

bipy-GS₂ = 5,5'-bis(methyl-S-glutathionyl)-2,2'-bipyridine

bp = base-pair

bz = basic-zipper

bZIP = basic zipper proteins

C = cytosine

CD = circular dichroism

COSY = correlation spectroscopy

cys₂bipy = 5,5'-bis(methyl-S-cysteinyl)-2,2'-bipyridine

cys₂pyr = 2,6-bis(methyl-S-cysteinyl)pyridine

cys₂terpy = 6,6''-bis(methyl-S-cysteinyl)-2,2':6',2''-terpyridine

DIEA = diisopropylethylamine

DMF = dimethylformamide

DMSO-d₆ = deuterated dimethylsulfoxide

DNA = deoxyribonucleic acid

DQF-COSY = double quantum filtered correlation spectroscopy

DTT = dithiotreithiol

EDTA = ethylenediaminetetraacetic acid

EI = electronic impact

EMSA = electrophoretic mobility shift assay

eq. = equivalent

ESI = electrospray ionisation

EtOH = ethanol

exc. = excitation

Fmoc- = fluorenylmethyloxycarbonyl-group

FT = Fourier transform

G = guanine

G28 = peptide designed by Schepartz *et al.*

G28-bipy-G28 = conjugate designed by
Schepartz *et al.*

G28^{SS} = peptide dimer designed by
Schepartz *et al.*

GCN4bd1 = peptide designed in this work

GCN4bd2 = peptide designed in this work

GCN4br5^{SS} = conjugate designed by Kim
et al.

GCN4br6^{SS} = conjugate designed by Kim
et al.

GSH = glutathione

HBTU = (O-(benzotriazol-1-yl)-
NN,N',N'-tetramethyluronium)
hexafluorophosphate

HOHAHA = homonuclear Hartman Hahn

HOMO = highest occupied molecular
orbital

HPLC = high-performance liquid
chromatography

HRES = high-resolution electrospray

HTH = helix-turn-helix proteins

Ida = iminodiacetic acid

LUMO = lowest unoccupied molecular
orbital

MALDI = matrix-assisted laser desorption

MBHA = 4-methylbenzhydrylamine

MCS = multiple cloning site

Me₂bipy = 5,5'-dimethyl-2,2'-bipyridine

Me₂pyr = 2,6-dimethylpyridine

Me₂terpy = 6,6''-dimethyl-2,2':6',2''-
terpyridine

MeOH = methanol

MLCT = metal-to-ligand charge transfer

MPA = 3-mercaptopropionic acid

MS = mass spectrometry

NMR = nuclear magnetic resonance

NOE = nuclear Overhauser effect

NOESY = nuclear Overhauser effect
spectroscopy

pD = deuterium ion potential

pdb = protein data bank

pH = hydrogen ion potential

phen = 1,10-phenanthroline

phen' = 5-(amidoglutaryl)-1,10-
phenanthroline

phen-IA = N-iodoacetyl-5-amino-1,10-
phenanthroline

phin = 9,10-phenanthrenequinone diimine

pKa = acid dissociation constant potential

pyr(GCN4bd1)₂ = conjugate designed in
this work

pyr-Br₂ = 2,6-bis(bromomethyl)pyridine

pyr-GS₂ = 2,6-bis(methyl-S-
glutathionyl)pyridine

RNA = ribonucleic acid

ROS = reactive oxygen species

RP-HPLC = reversed-phase high
performance liquid chromatography

Rt = retention time

SPPS = solid-phase peptide synthesis

T = thymine

TBE = Tris-borate-EDTA buffer

TCEP = tris(2-carboxyethyl)phosphine

TCNQ = 7,7',8,8'-
tetracyanoquinodimethane

terpy = 2,2:6',2''-terpyridine

terpy(GCN4bd1) = impurity observed in
this work

terpy(GCN4bd1)₂ = conjugate designed in
this work

terpy(GCN4bd2)₂ = conjugate designed in this work

terpy-Br₂ = 6,6''-bis(bromomethyl)-2,2':6',2''-terpyridine

terpy-GS₂ = 6,6''-bis(methyl-S-glutathionyl)-2,2':6',2''-terpyridine

TFA = trifluoroacetic acid

TG = Tris:Glycine buffer

THF = tetrahydrofuran

TOF = time-of-flight

TPPI = time proportional phase incrementation

Tris = 2-amino-2-hydroxymethyl-propane-1,3-diol

UV = ultraviolet

vis = visible

Ac- = acetyl- group

Pbf = 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl- group

Trt = trityl- group

wt = wild type

CHAPTER I: LITERATURE REVIEW

1.1 – Introduction

Artificial regulation of biomolecular activity is a biochemist's fantasy. Biomolecular recognition or activity of protein is often dictated by their three-dimensional arrangement. In order to reach their functional folded states, proteins may use templating proteins (ca chaperones [1]), and the structure may vary with the conditions (pH,[2] temperature [3,4]) or following contact with other entities (proteins,[5] metal ions,[6] light,[7] or DNA (see section 1.3)). Artificial stabilisation, or introduction of stimuli-responsive elements able to control the structure, therefore represents an opportunity to control the activity of biomolecules. The first part of this chapter introduces the research done so far towards the stabilisation of α -helices, which are particularly relevant to biomolecular recognition and to this work, and other secondary and super-secondary structures are briefly introduced with a focus on model mimics which include polypyridine units. Next, some larger assemblies (tertiary and quaternary structures), based on stabilised secondary structure elements will be introduced.

DNA is often regarded as a genetic encyclopedia owning the code for protein transcription, in which case transcription factor proteins serve to act as a set of bookmarks, marking off the reading frame for the messenger RNA. Considering their ability to bind DNA at specific sites, transcription factors are the last links in large signalling chains.[8] Indeed, they often function as multi-protein complexes and their activity, which consists of either promoting or inhibiting transcription, is often regulated by co-factors such as amino-acids,[9] metal ions,[10] hormones,[11] and/or by post-translational modification such as phosphorylation.[12] There exists a wide variety of quaternary and tertiary protein structures which mainly relies on α -helices to bind DNA at distinct recognition sites. As has previously been done in many reviews,[13,14] we chose to categorise transcription factors based on a similar structural motif within their DNA binding domains, rather than the more general structural classification of the proteins database (SCOP).[15]

The second part of this chapter will introduce some natural examples of transcription factors, for which DNA binding is well-understood, and then focus on their miniaturisation and the preparation of peptide based artificial mimics with enhanced or regulated binding to specific DNA sequences (often based on α -helix stabilisation or turn insertion). It is worth mentioning that this can alternatively be achieved by some natural non-proteinogenic molecules[16] and their synthetic derivatives.[17] Some aromatic molecules, including polypyridine complexes, can efficiently intercalate into DNA, but their selectivity is usually based on shape rather than sequence.[18-20] This concept will be briefly introduced in Chapter 6, but will not be expanded on here.

The third part describes the coordination chemistry of 2,2'-bipyridine (**bipy**) and 2,2':6',2''-terpyridine (**terpy**). A large range of complexes containing **bipy/terpy** and transition metals have been previously reported and studied. However here the discussion will be limited to the dication from the second half of the 1st row of transition metals: Fe(II), Co(II), Ni(II), Cu(II) and Zn(II). The stoichiometry and structure of the resulting $M(II)L_n$ complexes in aqueous solution are particularly relevant to this work and will be described.

1.2 – Stabilisation of peptide secondary structure in natural peptides and their derivatives

1.2.1 – α -Helices

The α -helix is the most common peptide secondary structure for which the backbone organisation is stabilised by a repeated hydrogen bonding pattern between carboxyl oxygens from residues (i), with the amino hydrogens of others located 4 residues later in the sequence ($i+4$), affording a helix with 3.6 residues per turn (often noted as a heptad repeat of two α -helical turns) (see Figure 1.1A).[21] The propensity of a given peptide chain to form an α -helix depends largely on the primary structure, which corresponds to the nature of the residue

side-chains.[22] Among the 23 proteinogenic L-amino-acids, two have been shown to disfavour α -helix formation due to their great (glycine) or too restricted flexibility (proline). In contrast, conformational restrictions in non-proteinogenic dialkylated residues such as α -methylalanine (Aib) are compatible with α -helices, and therefore promote their formation.[23] α -Helices are among the most common peptide recognition motif, and are for example, involved in the apoptosis activation by B-cell leukemia peptides (Bcl),[24] or the inhibition of p53 transcription activation by the MDM2 oncoprotein.[25] However, short peptide fragments (less than 20 residues) rarely form α -helices,[26,27] and chemist have explored strategies to stabilise them.[28] The most straightforward approach for stabilisation of α -helices involves constraining residues at position $i \rightarrow i+3$, $i \rightarrow i+4$, or $i \rightarrow i+7$ in similar orientations, thus stabilising the 3.6 residues per turn helical motif (see Figure 1.1B). This can be achieved through diverse types of interaction/linkage, the presentation of which will start from those with the lowest bond energy.

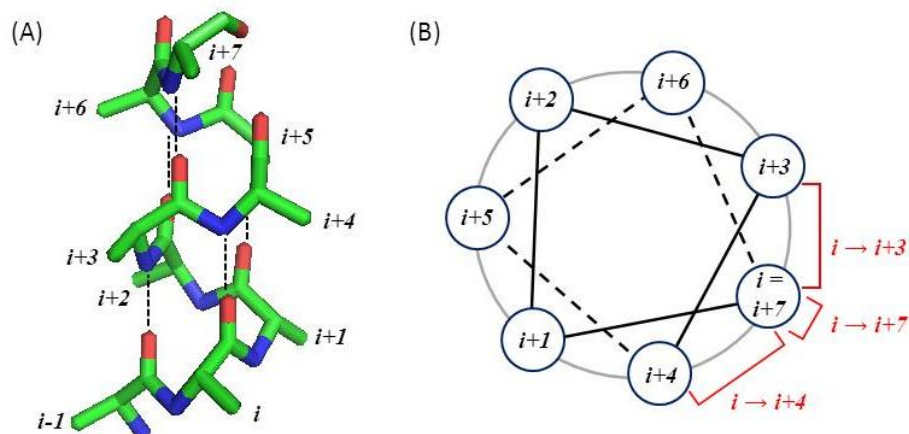


Figure 1.1 – Two representations of the α -helix and associated residue positions: (A) three-dimensional representation of a polyalanine α -helix and the $i \rightarrow i+4$ hydrogen bond pattern (green = carbon; blue = nitrogen; red = oxygen; hydrogen atoms are omitted for clarity); (B) top-down view of a helical wheel diagram displaying the proximity of i , $i+3$, $i+4$ and $i+7$ residues (backbone is displayed in black, circles represents α -carbons) and the possibility of helix stabilisation by enhancing $i \rightarrow i+3$, $i \rightarrow i+4$, or $i \rightarrow i+7$ interactions between side-chains, is indicated.

Based on the observation that a salt bridge can stabilise a short RNase A peptide mimic in a partial α -helical conformation,[27,29] Baldwin and co-workers first introduced the

glutamic acid-lysine salt pair at position $i \rightarrow i+4$ of a 17-mer polyalanine synthetic peptide (see Figure 1.2A). Interestingly, the resulting model peptide displayed higher helical content at neutral pH (~97% folded in the presence of 0.01 M NaCl and 274 K, calculation based on reference 30), associated with glutamic acid and lysine side chains being charged, compared to data recorded at lower or higher pH (~87 % in both cases).[31] By extension, similar α -helical stabilisation was obtained upon shortening,[32] or elongation of the propanoate side-chain.[33] Intermolecular electrostatic interaction with spermine [34,35] or guanidinium [36] is also an efficient strategy to stabilise the α -helix conformation of aspartate and glutamate bearing peptides (see Figure 1.2B).

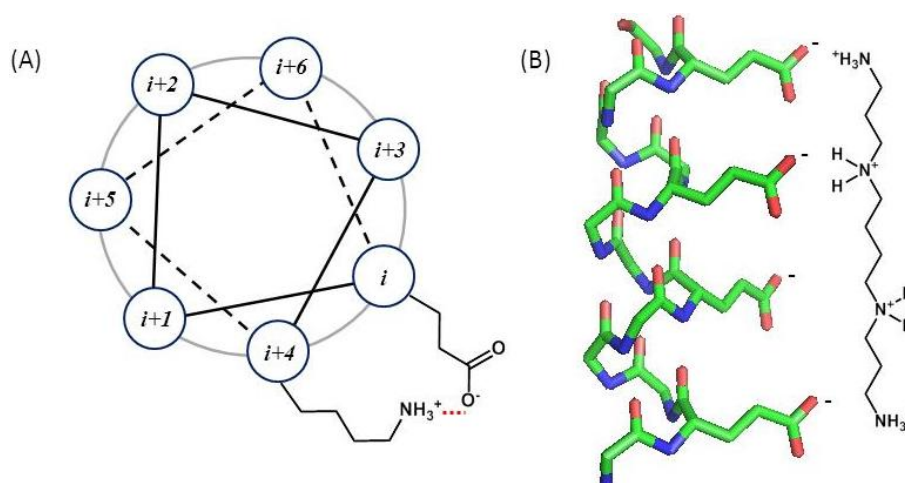


Figure 1.2 – Two examples of α -helix stabilisation at neutral pH by electrostatic interactions: (A) top-down view helical wheel diagram representing helices stabilised by intramolecular $i \rightarrow i+4$ interactions between glutamate and lysine side-chains, designed by Baldwin *et al.*; (B) side view of a α -helix bearing glutamate residues at positions i and $i+4$, stabilised by intermolecular interactions with spermine as shown by Sasaki *et al* (side-chains of other residues are omitted for clarity).

A variant of this strategy involves α -helix stabilisation with coordination bonds, by careful positioning of residues able to bind metal ions. Ghadiri and co-workers introduced a cysteine and/or histidines at positions $i \rightarrow i+4$ of synthetic peptides and reported an increase in helicity upon addition of Cd(II), Zn(II), Cu(II), Ni(II),[37] or *cis*-Ru^{III}(NH₃)₄. [38] In contrast, the helicity decreased upon addition of Pd(II) or Pt(II), which were proposed to either adopt a different geometry or to coordinate different heteroatoms.[37] The strategy was

expanded to other residues bearing side-chains with metal coordination ability, such as methionines [39] or unnatural residues with iminodiacetic acids (Ida),[40] or pyridyl side-chains,[41] allowing the binding of a larger variety of metal ions (see Figure 1.3A). Increased helical stability as a result of metal addition was used to enhance biomolecular recognition of short peptides, such as antimicrobial peptide mimics [42,43], but also in the construction of haemic models [44,45], or higher order structures (see section 1.2.3). Conversely, Futaki and co-workers have shown that metal-complexation could lead to destabilisation of a α -helical motif when Ida residues are introduced at $i \rightarrow i+2$ positions (see Figure 1.3B).[46] Synthetic peptides have also been stabilised in helical conformations using electrostatic dipole-dipole (host-guest systems),[47,48] cation- π ,[49, 50] and π - π interactions.[51]

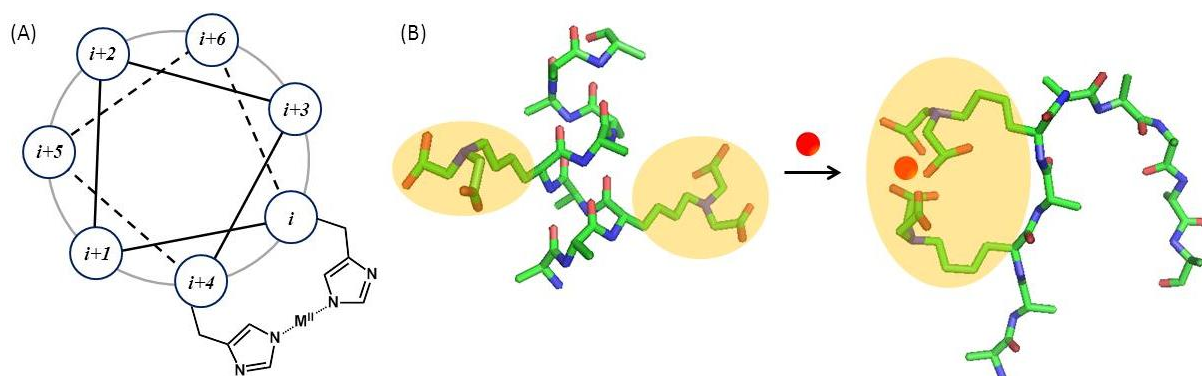


Figure 1.3 – (A) Helix stabilisation by transition metal ions (M^{II}) upon coordination to side-chains of histidines at positions i and $i+4$ of a *de novo* peptide designed by Ghadiri *et al.* (B) $Fe(II)$ (red sphere) coordination to iminodiacetic acid side-chains of Ida residues (highlighted in orange) at positions i and $i+2$ results in α -helix destabilisation (relative to Futaki *et al.* studies).

α -Helices were also stabilised using artificial covalent linkages between side-chains with increasing biological lifetimes: from cystine,[52,53] to lactam (cyclic amide) [54,55] and hydrocarbon linkage (use of non-natural residues).[56,57] Aside from the atoms directly involved, the size and the chirality of the linkage are important features to afford efficient stabilisation, and numerous cross-linkers have been evaluated, including unmetallated 5,5'-dimethyl-3,3'-bipyridine.[58] A recent report by Greenbaum and co-workers describes a cross-linker screening method for determining the most suitable one (molecule and position),

to achieve stabilisation of a decapeptide derived from the minimal active domain of calpastatin into the bioactive two-turn long α -helix motif (see Figure 1.4).[59] Unlike uncrosslinked analogues, the most helical crosslinked decapeptide studied was shown to efficiently inhibit calpain-1 protease activity, while remaining selective with respect to other enzymes from the papain protease family.

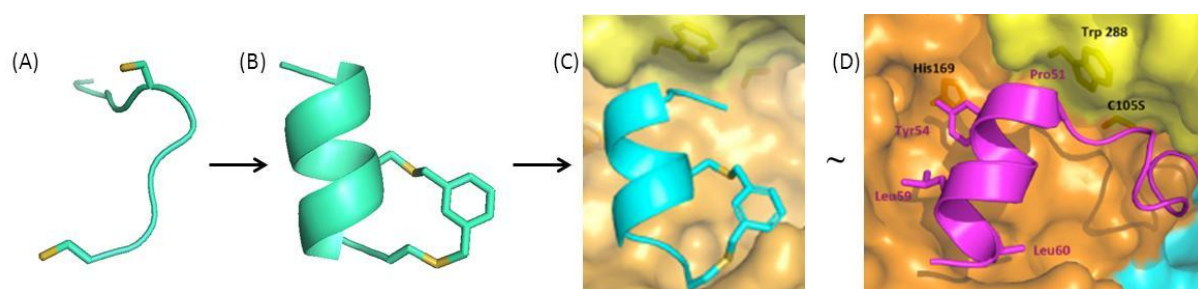


Figure 1.4 – Cartoon representation of the decapeptide prepared by Greenbaum *et al.* derived from the calpastatin fragment alone (A) before and (B) after cross-linking with 1,3-bis-(bromomethyl)benzene, and (C) when bound to the surface of the calpain protein. (D) Crystal structure of the calpastatin protein bound to calpain; adapted with permission from reference 59. Copyright 2012 American Chemical Society.

Schepartz and co-workers developed a general residue grafting strategy (see section 1.3.3) that involves identifying the nature and relative positioning of residues essential to a specific function, such as those from the *Bak* protein which is involved in apoptosis activation upon binding to Bcl-x_L, and to incorporate them onto a well folded scaffold such as the avian pancreatic polypeptide (aPP), which consists of an α -helix stabilised by hydrophobic interactions with a polyproline motif. The PPBH3-1 mutant that incorporates 6 key residues from the *Bak*₇₂₋₈₇ peptide onto the aPP scaffold, displays a higher affinity for Bcl-2 and Bcl-x_L than the *Bak*₇₂₋₈₇ peptide alone.[60,61] The same strategy was also applied, among others, to the design of helical peptides capable of inhibiting p53-hDM2 interactions.[62]

The α -helical content of short peptides can be controlled by side-chain crosslinking through stimuli-responsive units. Woolley and co-workers introduced an azobenzene photoswitchable unit as a side-chain crosslink at positions $i \rightarrow i+4$ [63] and $i \rightarrow i+7$ [64] of two similar 17 residues synthetic peptides, through alkylation of cysteine residues. Irradiation

at 380 nm resulted in 38% and 36% increase of α -helicity (respectively), consistent with the hypothesis that the irradiated *cis*- azobenzene linker better suits the $i \rightarrow i+4$ and $i \rightarrow i+7$ gap of an α -helix than the larger *trans*- azobenzene (see Figure 1.5A). Conversely, a similar peptide conjugate with an azobenzene cross-linker in a larger $i \rightarrow i+11$ positioning, displays the opposite behaviour. Irradiation at 370 nm leads to a 40 % decrease of the helical content (see Figure 1.5B).[63]

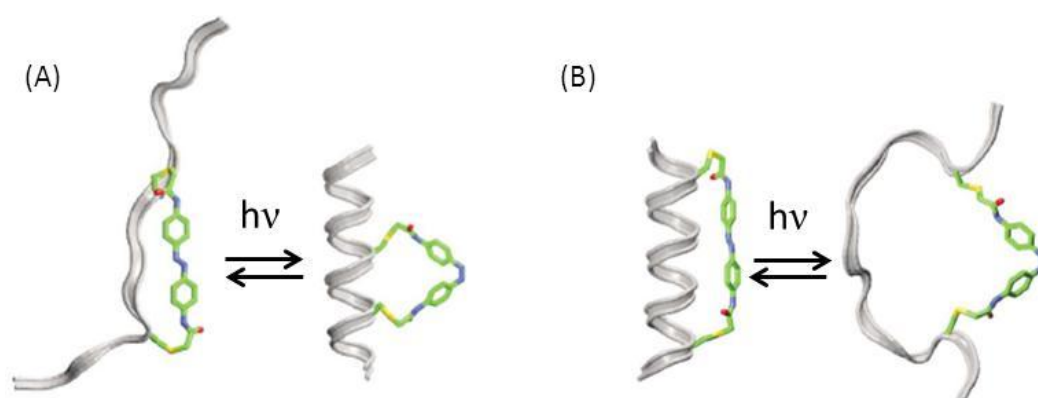


Figure 1.5 – Representation of the impact of *cis-trans*- isomerisation of the azobenzene cross-linker on the peptide secondary structure when introduced at either position $i \rightarrow i+4$ (A), or $i \rightarrow i+11$ (B)., adapted with permission from reference 63. Copyright Elsevier 2002.

As an extension, Allemann and co-workers introduced the azobenzene cross-linker ($i \rightarrow i+4$, $i \rightarrow i+7$ and $i \rightarrow i+11$) within synthetic helical peptides based on *Bak*₇₂₋₈₇ and *Bid*₉₁₋₁₁₁ recognition domains, which are both able to inhibit the anti-apoptotic activity of Bcl-x_L upon binding to its hydrophobic groove. Bcl-x_L binding constant measurements for the azobenzene modified peptides, indicate that irradiation has a higher impact on Bcl-x_L binding in cases where the dark-adapted *trans*- conformation are associated with the less helical peptide (ca $i \rightarrow i+4$ and $i \rightarrow i+7$).[65]

α -Helix stabilisation can also be achieved using other strategies, notably those involving extra linkages,[66,67] templating,[68-70] or modifications of the peptide backbone.[71] However, this introduction is limited to those strategies which have also found applications in controlling biomolecular recognition events.

1.2.2 – Other secondary structures

Despite their simplicity, the low structural constraint, and the poor residual analogies within the motif, turns are nonetheless considered a secondary structure element where the peptide chain is folding back on itself due to an inter-residue hydrogen bond between a carboxyl oxygen and amine nitrogen.[72] They are classified based on the number of residues; $i \rightarrow i \pm 1$ hydrogen bond forms a δ -turn, $i \rightarrow i \pm 2$ a γ -turn, $i \rightarrow i \pm 3$ a β -turn, $i \rightarrow i \pm 4$ a α -turn, and $i \rightarrow i \pm 5$ a π -turn. In contrast, an unstructured 6-16 residue peptide segment is defined as a loop.[73] Both loops and turns (especially β -turns) have been extensively mimicked, or stabilised,[74,75] but only the work involving polypyridine peptide conjugates will be herein presented. It is worth mentioning that careful introduction of turn mimics are often the decisive factor in stabilisation of supersecondary or tertiary structures.[76,77]

Imperiali and co-workers prepared a series of 2,2'-bipyridine (**bipy**) bearing protected amine and carboxylic acid substituents, suitable for SPPS (solid-phase peptide synthesis [78]) and subsequent insertion into the backbone of artificial peptides.[79-82] They subsequently prepared a undecapeptide incorporating two **bipy** units, separated by four glycine residues bearing a cyanoanthracene fluorophore and a quencher on the N- and C-terminus, respectively. Upon addition of Zn(II), the fluorescence of the peptide conjugate decreased, consistent with the turn-like structure hypothesised, in which the quencher and the fluorophore are positioned in close enough proximity for efficient electron transfer (see Figure 1.6A).[82] McCafferty and co-workers isolated enantiomers from an octahedral $\text{Ru}^{\text{II}}(\text{bipy})_3$ peptide conjugate derivative, composed of a peptide turn incorporating two **bipy** units within its backbone which are involved in Ru(II) coordination together with an additional unsubstituted **bipy**. [83] López and co-workers prepared a similar conjugate in which the two **bipy** units are separated by a 5-amino-3-oxapentanoic acid-glycine-proline-glycine motif. Interestingly, upon changing the chirality of one residue, that is L-proline for D-

proline, the chirality of the metal center was also altered.[84] Constable and co-workers linked 4'-(4-carboxyphenyl)-2,2',6',2''-terpyridine units at both the N- and C-terminus of a undecapeptide, bearing a glycine-proline β -turn inducer motif. Following reaction with a Fe(II) salt, the isolated conjugate was assigned to the expected cyclometallopeptide 2,2',6',2''-terpyridine (**terpy**) conjugate (see Figure 1.6B).[85]

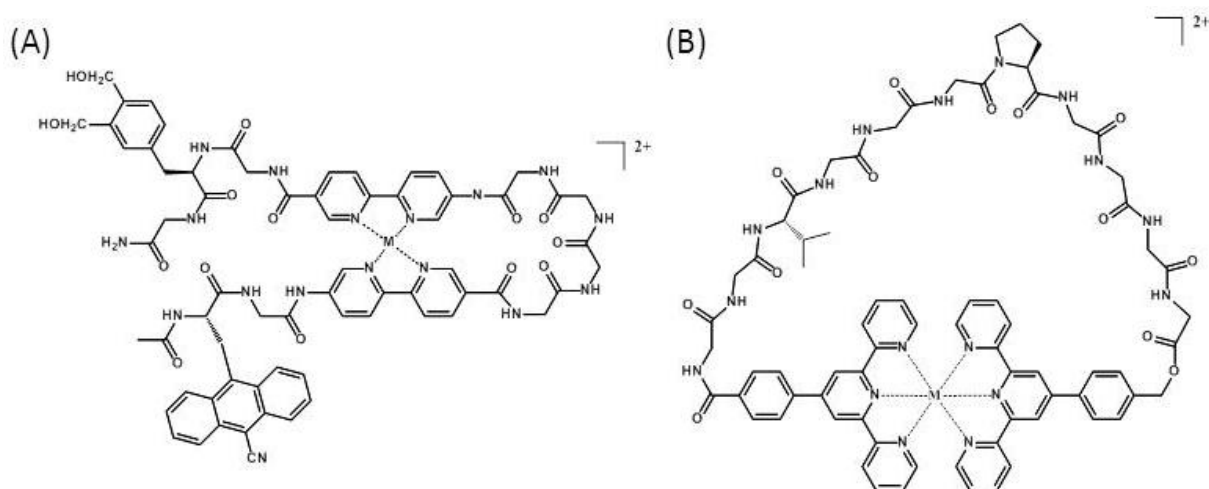


Figure 1.6 – Scheme illustrating the model peptides incorporating two **bipy** (A) or **terpy** units (B), stabilised in a turn conformation upon metal coordination and formation of ML_2^{2+} cyclometallopeptide complexes where L is a polypyridine unit, based on references 82, 85, 86.

Hairpins are long devices used to hold a person's hair in place, whereas in biochemistry it describes a peptide supersecondary structure with a similar geometry, ca both ends of a turn are connected to flanking domains with defined secondary structures.[87] Certain artificial α -hairpin (an α -helix-turn- α -helix motif), such as HN1, displayed interesting hydrolase activity,[88] which is dependent on the spatial proximity between the two helices.[89,90] Upon replacement of the turn motif by a stillbene linker, its activity can be regulated by light irradiation (see Figure 1.7).[91]

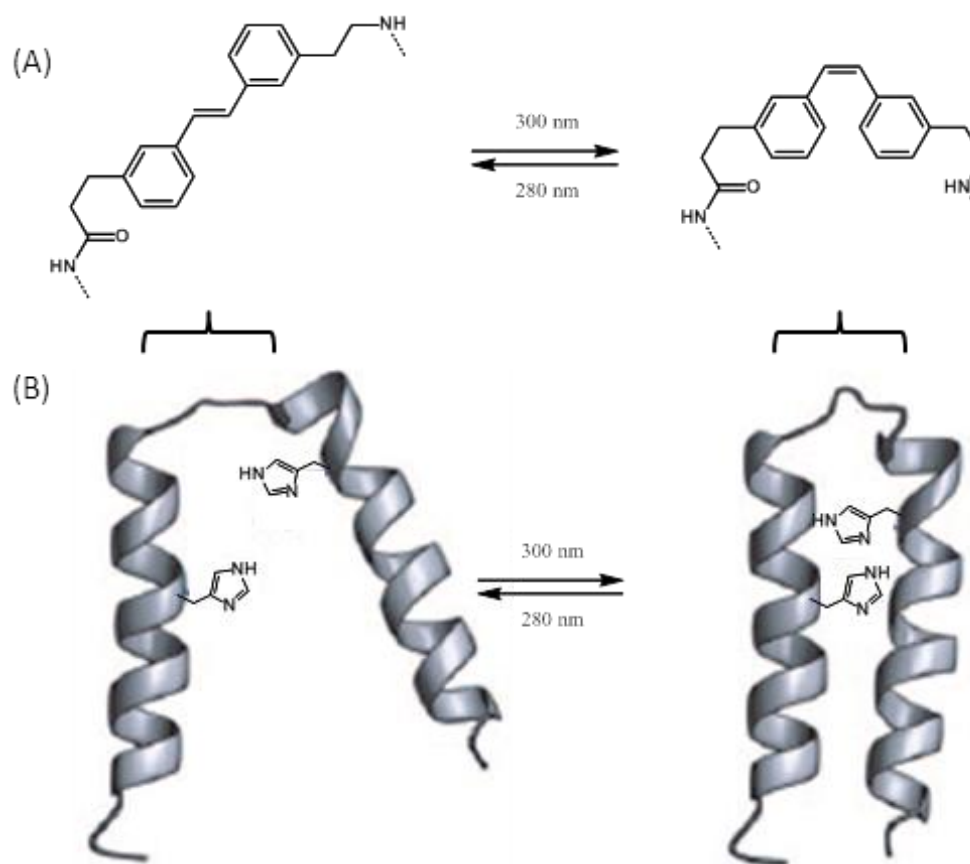


Figure 1.7 – Schemes illustrating (A) the *trans*-to-*cis* conformational transition of the stillbene linker upon light-irradiation, and (B) its impact on the supersecondary structure of the HN1 artificial hydrolase and especially the distance separating two imidazole side-chains, which are essential for the enzyme activity; adapted with permission from reference 91. Copyright 2009 Wiley-VCH.

In the β -strand, the second most encountered secondary structure motif, the peptide backbone forms linear strands in which the side-chain is positioned alternatively up and down. It is stabilised by inter-strand hydrogen bonding between carboxy and amino groups either in an anti-parallel or parallel β -sheet arrangement.[92] Despite some β -branched (valine, isoleucine) and aromatic residues (threonine, tyrosine, tryptophan, phenylalanine) being encountered more frequently in β -sheets,[93] the role of the side-chains in stabilising the arrangement can only be explained with respect to the tertiary structure.[94] Mimicking large β -sheet assemblies usually leads to the formation of insoluble aggregates in water.[95] For this reason, more soluble and shorter assemblies such as β -hairpins [96] (β -strand-turn- β -

strand motif [97]), are more attractive targets for chemist seeking to stabilise the β -sheet structural motif.[98] Herein, two strategies used to stabilise β -hairpins will be briefly introduced, with a focus on polypyridine(s) containing systems.

Promoting inter-strand side-chain interactions is one strategy by which to stabilise β -hairpins. This has been achieved through coordination bonds,[99] hydrophobic interactions,[100] or covalent bonding.[101] Imperiali and co-workers designed several artificial peptides with primary structures which have a different propensity to form β -hairpins, but all of which incorporate two **bipy** bearing residues. The peptides with a high propensity to form β -hairpins, promoted formation of 1M:2**bipy** complexes (where M corresponds to either Zn(II) or Co(II)) consistent with the closer proximity of the two **bipy** groups. In contrast, their counterparts with non β -hairpins promoting sequences, formed mainly 1M:1**bipy** complexes (see Figure 1.8A).[81] A more common strategy for β -hairpin stabilisation, involves replacing the turn with an artificial mimic to ensure efficient and stable reversal of the peptide chain, whilst ensuring the strands remain hydrogen-bonded.[102] Kelly and co-workers designed metal-dependent β -hairpin switches taking advantage of the *trans*- to *cis*- conformational transition experienced by a **bipy** unit upon Cu(II) complexation.[103] Addition of Cu(II) was shown to promote β -strand formation for some of the peptide conjugates designed, likely due to strand realignment triggered by the conformation of the **bipy** turn (see Figure 1.8B). To the best of our knowledge, this constitutes the first and so far only application of conformational switches based on **bipy** within peptide scaffolds (other applications are introduced in Chapter 3). Alternative switches triggered by light were also studied in a similar context.[104-107]

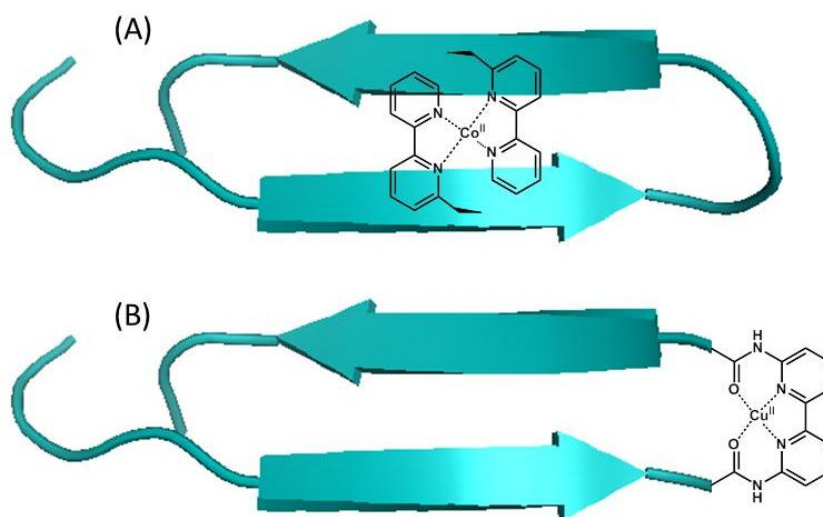


Figure 1.8 – (A) Scheme illustrating the cooperativity between secondary structure formation and Co(II) coordination in the β -hairpin prepared by Imperiali *et al.* for which each strand contains one **bipy** side-chain.[81] (B) Scheme illustrating the stabilisation of β -hairpin motif by Cu(II) complexation resulting in the conformational transition of a **bipy** unit inserted in the peptide backbone as a metal-dependent turn mimic.[103]

1.2.3 – Examples of higher order structure

In nature, certain α -helical peptides such as keratin are able to associate resulting in the formation of a left-handed supercoil of helices.[108] These structures, termed coiled coils, are stabilised by multiple inter- and intramolecular interactions (including $i \rightarrow i+4$).[109] The motif for the simplest example (a homodimeric parallel coiled coil) consists of two helical peptides based on the heptad repeat approach (peptide fragment of seven residues with positions denoted $a-b-c-d-e-f-g$ and $a'-b'-c'-d'-e'-f'-g'$). In a polar environment such as water, helix association is favoured if all residues with hydrophobic side-chains are located on the same face of the helix (e.g. a , a' , d and d' sites) thus triggering aggregation and formation of a hydrophobic core (see Figure 1.9A). Moreover, the formation of salt bridges in (e.g. in e , g and e' , g') provides electrostatic interactions and further stabilisation. For sufficient stability a coiled coil requires several repeats of the heptad motif (3-4),[110] and a slight left-handed distortion to allow for correct side-chain alignment (due to the non-integer numbers of

residues per α -helical turn: 3.6). Such peptides bearing hydrophobic and hydrophilic faces are termed amphipathic and when most of the hydrophobic residues are leucines, the dimer is called a leucine zipper.[109] Natural coiled coils assembled into either parallel or antiparallel orientations, are important di- or trimerisation domains (see section 1.3.1). However, *de novo* (from “first principles”) design allows for the preparation of coiled coils with a right handed coiling,[111] assembling directly seven helices,[112] or more if involved in coiled coil packages.[113]

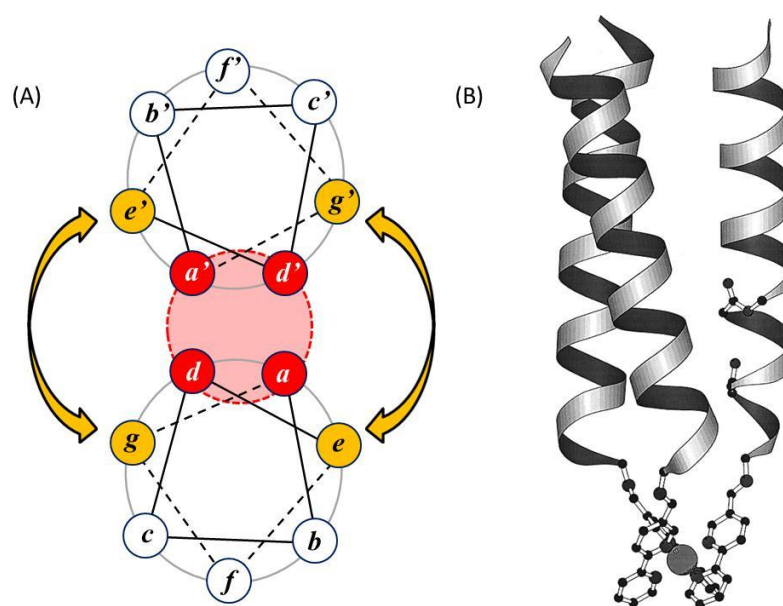


Figure 1.9 – (A) Schematic top-down view of a helical wheel diagram depicting a dimeric parallel coiled coil motif where the hydrophobic side-chains of the residues in positions *a*, *a'*, *d* and *d'* form the hydrophobic core (red), and the salt bridges formed between positions *e*, *g* and *e'*, *g'* (orange). (B) Model of a three stranded coiled coil stabilised by formation of a $\text{Ru}(\text{bipy})_3^{2+}$ complex, where each **bipy** unit is covalently bound to a helix N-termini. Reproduced with permission from reference 114. Copyright 1998 American Chemical Society.

As for α -helices, designed coiled coils may undergo structural transitions as a function of pH,[115,116] temperature,[3] or upon metal addition.[117-119] Interestingly, Liberman and co-workers prepared a 15-mer amphipathic peptide with a **bipy** coupled to the N-terminus which displayed poor helicity (35%). However, addition of 0.33 equivalents of an Fe(II) salt resulted in an important increase in the helicity (85%), associated with formation of a $[\text{Fe}^{\text{II}}(\text{bipy})_3]$ complex and a three- α -helix coiled coil.[120] Case, McLendon and co-workers

prepared similar coiled coils based on peptide-**bipy** conjugates (see Figure 1.9B) and studied their stability upon changing the metal ion,[121] or the primary structure.[114] They also designed Ru(II) analogues, which they used to measure either electron transfer towards an additional coordination sites located on the exterior of the coiled coil,[122] or the binding of host molecules to an internal cavity.[123]

Conversely, when introduced into a hydrophobic environment, such as a lipid bilayer, amphipathic helices can self-assemble and form hydrophilic pores across membranes, thus promoting the permeation and migration of molecules, hence their cytotoxic and antimicrobial activities.[124,125] It is possible to take advantage of pores created by artificial pH-dependent peptides such as GALA, that form amphipathic helices at pH lower than 6, to facilitate membrane crossing of drugs, in a virus-like fashion (see Figure 1.10A).[126,127] Based on these features, Futaki and co-workers prepared a metal-gated ion channel by introducing a 64 residue peptide into a liposome (artificial membrane).[128] The peptide was composed of three domains: (1) a domain derived from alamethicin (a channel forming peptide [129]), (2) a flexible and unstructured tetraglycine linker, and (3) an extramembrane segment mimicking the leucine zipper domain of Fos,[130] but bearing two non-conserved Ida residues in position $i \rightarrow i+2$. Fe(III) coordination at Ida destabilises α -helices in the extramembrane segments thus forming random coils, resulting in a significant increase in the pore size and the ion flux (see Figure 1.10B). A variant in which the extramembrane fragments resemble the C-terminal domain of calmodulin, was also prepared, and displayed increased ion-flux upon addition of Ca(II) associated with an increase of the hydrophobic surface of the calmodulin segments.[131]

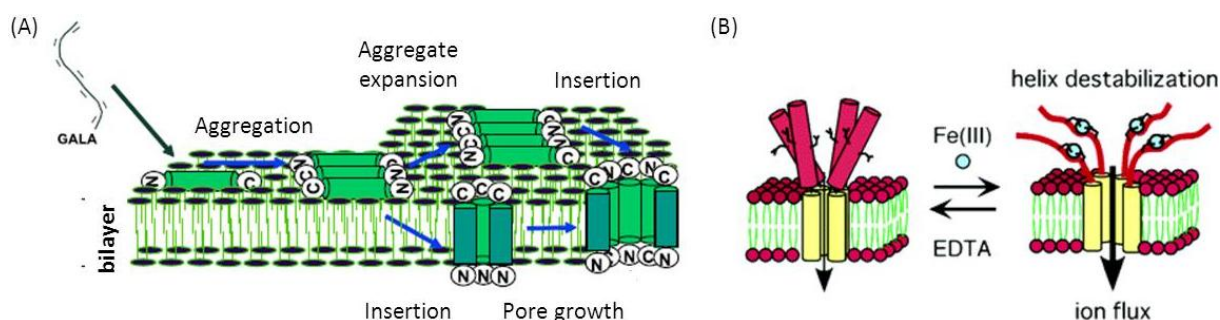


Figure 1.10 – Schemes illustrating (A) the pore formation by GALA peptides within a lipid bilayer membrane at acidic pH, for which two plausible mechanisms are displayed and (B) the metal-gated ion-channel, for which Fe(III) complexation results in destabilisation of the extramembrane helical segments (red) favoring ion flux through the pore (yellow); reproduced with permissions from references 126 and 128, respectively. Copyright 2004 Elsevier and 2006 American Chemical Society.

Finally, peptide conjugates have been assembled through coordination bonds to form even larger artificial peptide constructs. Tezcan and co-workers have prepared a variety of peptide building blocks based on cytochrome *b*₅₆₂ peptides by functionalisation of side-chains with unnatural metal chelation motifs, such as 1,10-phenanthroline (**phen**), or **terpy**.^[132] Following coordination studies, the building blocks were assembled by addition of metal ions leading to a large variety of periodic protein nanoarrays (see Figure 1.11).^[133-135]

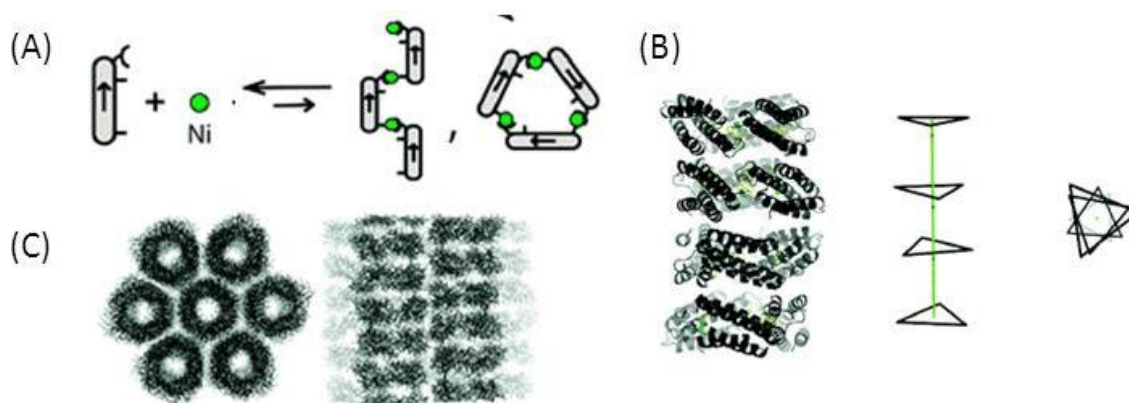


Figure 1.11 – (A) Schemes representing the peptide self-assembly in triangles upon Ni(II) coordination to the **phen** side-chain, (B) the crystal structure of the resulting asymmetric unit and cartoon side-on and top-down view, and (C) the lattice packing in top-down and side-on views; reproduced with permission from reference 134. Copyright 2009 American Chemical Society.

Similarly, Chmielewski and co-workers prepared artificial peptides based on collagen which incorporate multiple **bipy** moieties, and which can self-assemble upon metal ion coordination, promoting the formation of fibrous materials (see Figure 1.12).[136,137]

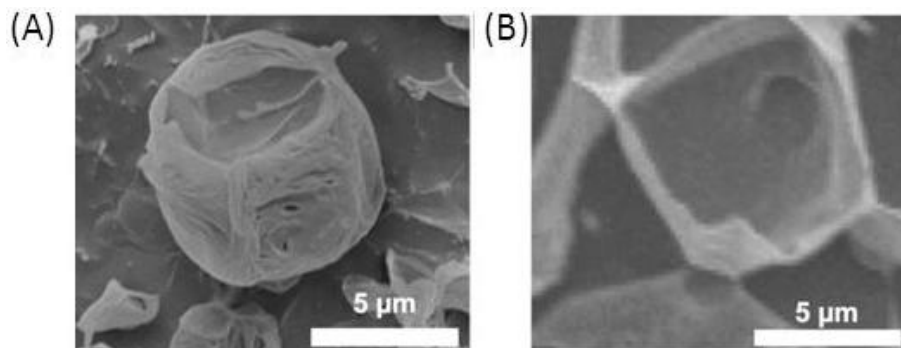


Figure 1.12 – (A) Cryo-TEM image of self-assembled collagen inspired peptides functionalised with bipy-Fe(II) prior to, or (B) after addition of EDTA; reproduced with permission from reference 137. Copyright 2013 American Chemical Society.

1.3 – DNA binding peptides and artificial derivatives

1.3.1 – Common structural motifs of DNA bound proteins

Deoxyribonucleic acid (DNA) is a polymeric chain of nucleotide motifs, which each include a phosphate, a sugar and one out of four possible aromatic nucleobases. In the most common structure adopted, B-DNA, the phosphosugar backbone forms an antiparallel right-handed double helix. This is stabilised by inter-strand interactions between complementary nucleobases (see Figure 1.13A) which occupy a plane perpendicular to the double helix orientation while pointing towards its interior.[138] The exterior of the DNA presents gaps with alternated 12 and 22 Å width, termed the major and minor grooves, respectively, which are separated by the negatively charged backbone (see Figure 1.13B).

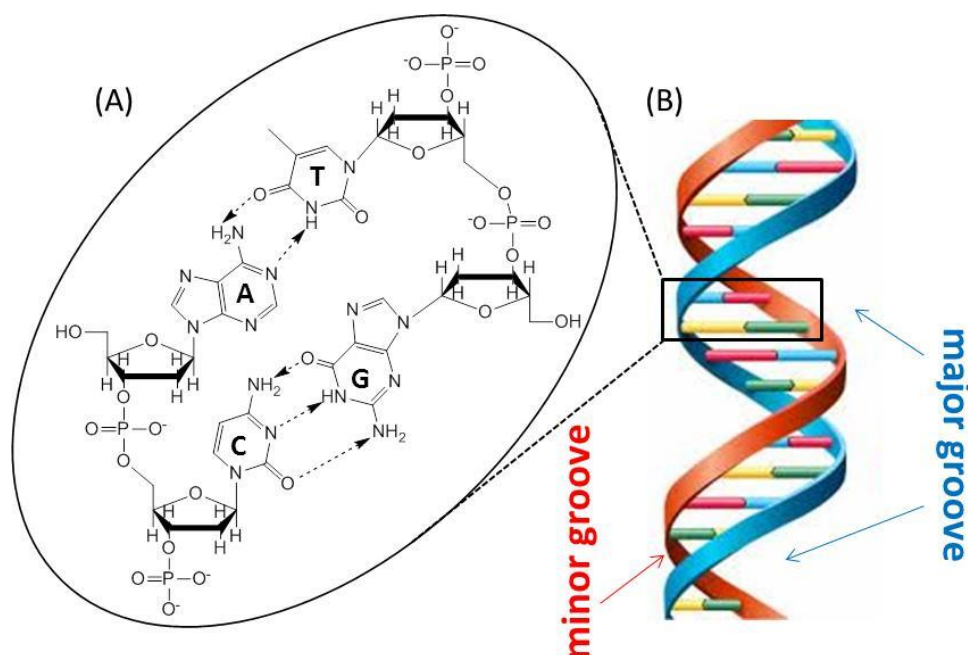


Figure 1.13 – (A) Scheme illustrating the chemical structure of the four different nucleotides, held together by specific hydrogen bonds (A-T: adenine-thymine, C-G: cytosine-guanine). (B) Structure of the B-DNA right-handed double helix with the major and minor grooves identified; adapted from reference 139.

Within the group of DNA binding proteins, the most encountered structural motif is the helix-turn-helix (HTH),^[14] which forms a triangular-like bundle including at least three α -helices. The amphipathic recognition helix (α_3) sits in the major groove of DNA and its side-chains are involved either in specific hydrogen bonds or electrostatic interactions with nucleobases and phosphates, or in hydrophobic interactions with the other helices (α_1 and α_2), which contribute to pack and position α_3 , orthogonal to them, within the major groove (see Figure 1.14A). The sharp turn connecting α_2 and α_3 consists of a highly conserved primary structure within HTH proteins, hence the name of this superclass of proteins.^[140-142] Among the simplest example, the homeodomain fold consists of one tri-helical HTH domain plus a N-terminal arm which binds to the DNA minor groove and makes specific hydrogen bonds with bases.^[143] Homeodomain proteins all bear asparagine residues at position 51, thus enhancing recognition of AT-rich sequences, however there exist a large variety of mutants which bind different recognition sites (see Figure 1.14B).^[14,144] For example, most homeodomain proteins bear a glutamine residue at position 50 and recognise 5'-TAATTG-3'

and 5'-TAATTA-3' consensus sequences, whereas a variant including a lysine at this position (Q50K) will bind DNA with a higher affinity at the 5'-TAATCC-3' site.[145]

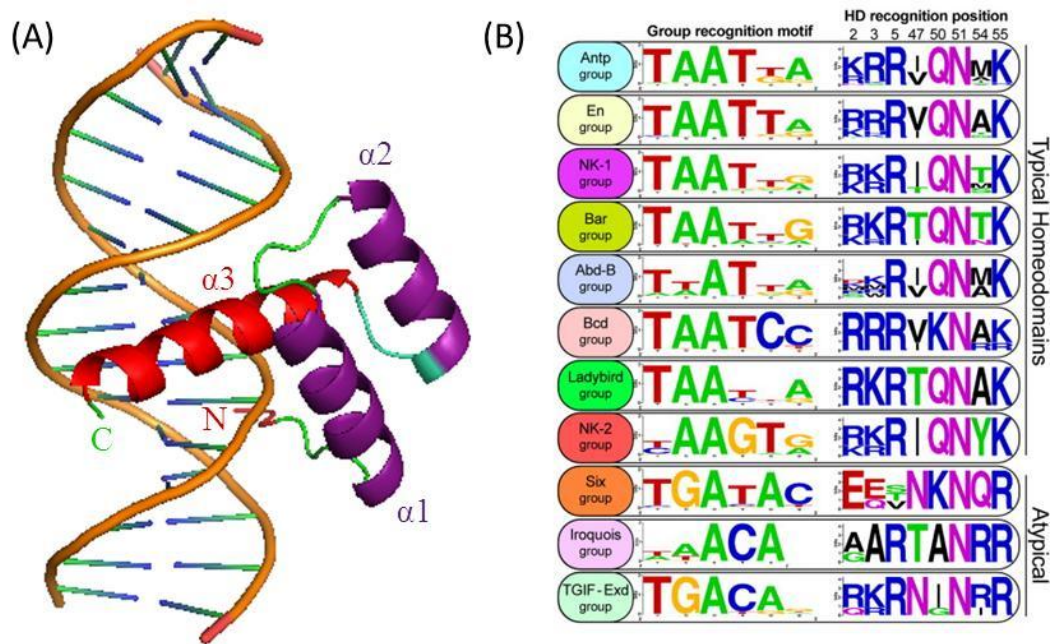


Figure 1.14 – (A) Cartoon of the engrailed homeodomain of *Drosophila* complexed with DNA bearing the recognition site 5'-TAATTA-3': secondary structure element contacting DNA is displayed in red, supporting helices in purple, and the sharp turn domain in light blue; based on pdb code 1HDD [143]. (B) Sequence logo representation of the average DNA recognition site and key peptide residues of 82 homeodomain proteins classified in specificity groups; reproduced with permission from reference 144. Copyright 2008 Elsevier.

In the basic domain (**bd**) superclass, transcription factor proteins bind DNA as dimers (either as homodimers or heterodimers). Each monomer includes an unfolded **bd** containing basic residues, which folds into an α -helix in the presence of the DNA recognition site, thus fitting in its major groove and forming direct and specific contacts with the backbone and exposed nucleobases.[146-150] This is attached to a dimerisation domain located towards the C-terminus, which may consist, depending on the subclass, of a leucine zipper (basic leucine zipper = bZIP),[109] a four helix bundle (basic helix-loop-helix = bHLH),[151] or a combination of the two (basic helix-loop-helix zipper = bHLHZ).[152] Therefore, the overall structure of a DNA bound bZIP transcription factor, such as the general control non repressed 4 (GCN4) basic zipper (**bz**) homodimer, is a helical fork made of uninterrupted α -helices (see Figure 1.15A), whereas these are interrupted by loops of variable length and sequence in

bHLH transcription factors, such as MyoD (see Figure 1.15B). In all cases, the selection of the dimerisation partner triggered by the dimerisation domains, is essential for DNA binding and bioactivity. For example, Fos proteins are unable to form homodimers, whereas Jun homodimers bind the 5'-TGAGTCAT-3' DNA recognition site (AP1), although the affinity is low compared to that of the Fos/Jun heterodimer.[153] All heterodimers composed of proteins from the Jun, Fos, ATF and JDP families, are termed AP-1 transcription factors. In a given cell, a large number of combinations are possible, each characterised by distinct DNA affinity and transactivation, thus allowing for the regulation of a large variety of processes such as proliferation, differentiation, apoptosis and transformation.[154]

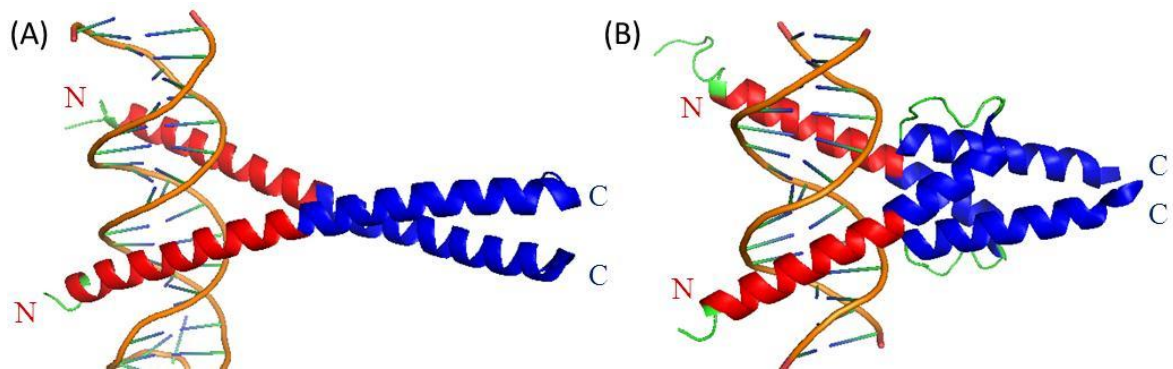


Figure 1.15 – Cartoon of (A) the GCN4^{bd} homodimer from *Saccharomyces cerevisiae* complexed with the 5'-TGACTCAT-3' DNA recognition site, and (B) the MyoD homodimer from *Mus Musculus* complexed with the 5'-CAGCTG-3' DNA site; ^{bd} are displayed in red, dimerisation domain in blue; based on pdb code 1YSA [155] and 1MDY [156].

The zinc finger is another important structural motif, in which Zn(II) co-factors are required to form the folded structure allowing for DNA binding, which can include β -strands.[157-160] Transcription factors of this family often function as multidomain proteins and regulate various functions such as DNA recognition, RNA packaging and apoptosis.[160] A distinction can be made with metal transcription sensors (e.g. NikR or MerR), for which DNA binding is also regulated by metal ions, however these regulate the transcription of proteins involved in the recruitment, transport, sequestration and regulation of the metal-ion sensed.[10,161,162] It is worth mentioning that more families of DNA binding proteins have

been identified, and that some transcription factors contact DNA through β -strands rather than helical motifs.[13]

1.3.2 – Artificial dimerisation and stitchery of basic domains

The DNA binding of basic transcription factors rely on the short **bd** sequence to form direct contact with the DNA target site, whereas slightly longer domains which allow for dimerisation and thus tight DNA binding, were assumed non-essential for the specificity. McKnight and co-workers interchanged the **bd** and zipper domains between transcription factors of the same sub-family, namely GCN4 and C\EPB, and tested their DNA selectivity. A mutant peptide bearing a GCN4**bd** and a C\EPB zipper domain was shown to selectively bind the GCN4 recognition site, whereas the reverse peptide (ca C\EPB**bd** and GCN4 zipper domain) selectively binds to C\EPB recognition site.[163] They concluded that DNA selectivity was not affected by the nature of the dimerisation domain and that the two distinct functions (dimerisation and selective DNA binding) are carried out independently by the zipper and **bd**.

Based on these results, Kim and co-workers prepared a short peptide dimer, consisting of a fully conserved **bd** and a glycine-glycine-cysteine C-terminal sequence, which forms a disulfide bond upon oxidation, thus allowing for covalent dimerisation with a flexible turn-like motif. The 2×34 residue synthetic dimer binds specifically to GCN4 recognition sites at low temperature with an affinity similar to that of the wild type protein and related **bz** peptides (see Figure 1.16).[164] This strategy was further exploited to estimate the minimum number of residues from GCN4**bd** required for sequence specific DNA binding, using a variety of disulfide bridged artificial dimers of varying length.[165]

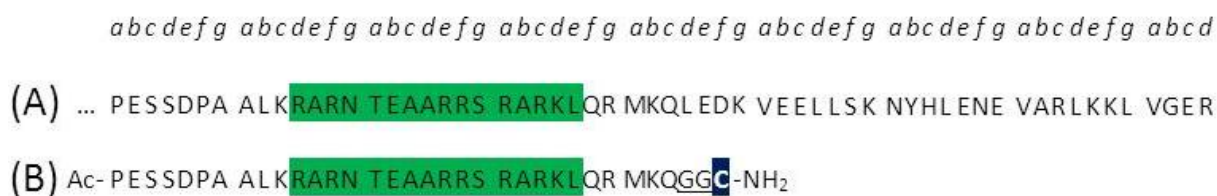


Figure 1.16 – Sequence alignment for (A) b-ZIP C-terminal domain of wild type GCN4 (residue 222-284), and (B) artificial peptide prepared by Kim et al. with 31 residues conserved from GCN4 wild type, including the whole **bd** (green), two glycines residues (underlined) for flexibility, and a cysteine residue (blue) for dimerisation; top lane indicates the relative heptad assignment based on the leucine zipper repeat; adapted from reference 164.

Related studies have shown that GCN4**bd** dimerisation with other types of covalent bonds,[166] coordination bonds,[167-169] or electrostatic dipole-dipole (host-guest) interactions,[170,171] also result in DNA binding properties similar to that of wild-type GCN4. Schepartz and co-workers prepared a series of polypyridine GCN4**bd** conjugates, which self-assemble upon Fe(II) binding thus increasing their specific DNA affinity, and found that bulky dimerisation domains promoted binding to a DNA sequence where the two half-sites of the native target DNA sequence are separated by an additional central base-pair (**bp**) (see Chapter 5).[167-169] Morii and co-workers investigated the impact of conformational restriction in the dimerisation domain on DNA binding. For this, they prepared two artificial dimers, where GCN4**bd** moieties are linked through either a (9R, 10R)-trans-9,10-dihydrophenanthrene-9,10-diol linker, or its (9S, 10S) enantiomer. Both conjugates display similar DNA affinity regardless of the linker orientation.[172] In contrast, a peptide mimicking the MyoD**bd** and dimerised through the (9R, 10R)-trans-9,10-dihydrophenanthrene-9,10-diol linker displayed higher specificity toward the MyoD recognition site (E-box) than the analogue bearing the (9S, 10S) enantiomer (Figure 1.17).[173] This is consistent with the crystal structure of the DNA-bound MyoD published the following year, displaying a right-handed orientation of the two **bd** with respect to the DNA helix.[156]

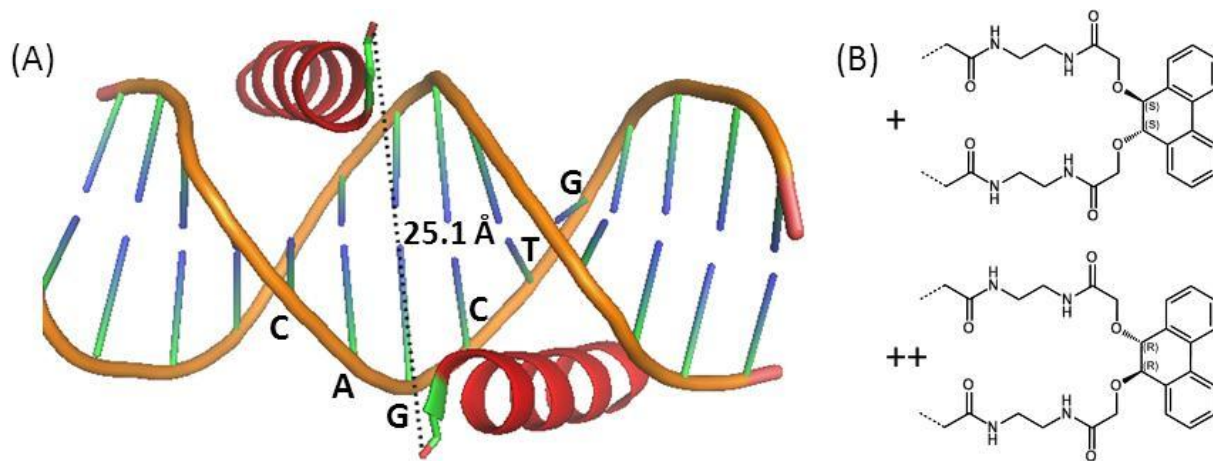


Figure 1.17 – (A) Cartoon illustrating the distance and orientation between two β -atoms of residue 123 of MyoD (equivalent to sulfur from cysteine and linker insertion position for the peptide prepared by Morii *et al.*) in the MyoD homodimer complexed with E-box; single letters indicate the DNA sequence of the recognition site; based on pdb code 1MDY [156] (B) Chemical structure of trans-9,10-dihydrophenanthrene-9,10-diol-based enantiomers used to dimerise the peptide based on the **bd** of MyoD; based on reference 173.

Mascareñas and co-workers inserted an azobenzene dimerisation motif which mimics a turn between GCN4**bd** peptides. Binding of the resulting conjugate to the DNA recognition site was significantly enhanced upon light-irradiation, associated with conversion to the *cis*-azobenzene motif which pre-organises the peptide moieties for DNA binding (see Chapter 4).[174] Chimeric fusion proteins incorporating artificial GCN4**bd** dimers and either cytochrome b_{562} , [175,176] or photoactive yellow protein from *Halorhodospira halophila* [177] were shown to display redox and light-dependent DNA binding, respectively. As previously mentioned, metal transcription sensor proteins are involved in metal level regulation or detoxification, where the metal ions of interest are the trigger for sensing and DNA binding. Recent efforts have been directed towards the redesign of such proteins, ca design of a MerR (Hg sensor) analogue with a single chain coiled coil,[178] or redesign of the NikR tetramerisation allosteric site to favour $[\text{UO}_2]^{2+}$ rather than Ni(II) sensing.[179]

Goddart III and co-workers came up with the concept of **bd** stitchery, which involves studying the consequence of displacing the artificial linker towards the N-terminus of the **bd**. A peptide based on Jun**bd** and dimerised with a disulfide at the N-terminus specifically

recognised a palindromic sequence, similar to the wild-type Jun DNA recognition site (5'-ATGACTCAT-3'), but with inverted half-sites and an additional central **bp** (5'-TCATCGATGA-3').[180] Moreover, an artificial Jun homodimer with a C- to N- terminus disulphide linker was shown to bind selectively to the non-palindromic sequence 5'-ATGACGATGA-3'.[181] It was concluded that upon artificial dimerisation, each **bd** peptide recognises one DNA half-site (4 **bp**), the relative positioning of which are altered when peptides are dimerised through the N- rather than the C-terminus. The two central **bp** were required for high peptide-DNA affinity, and probably account for geometrical constraints. By analogy, a GCN4**bz** dimer with reversed sequences (C→N) recognise a DNA sequence, which half-site are inverted compared to the natural protein binding site (see Figure 1.18).[182] The concept of protein stitchery was further extended to the design of artificial peptides composed of three **bd** able to recognise longer DNA sequences.[183]

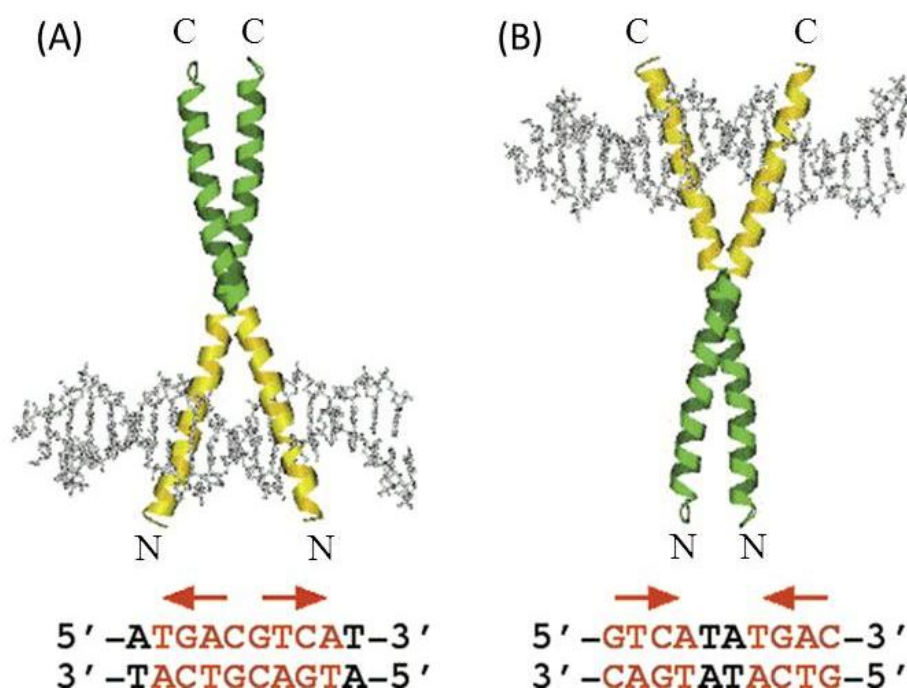


Figure 1.18 – Application of **bd** stitchery to GCN4**bz** peptides. Cartoon representation of (A) wild type GCN4**bz** binding to the CRE recognition sequence (red), or (B) GCN4**bz** with inverted sequence (N→C) binding with wild type affinity to a DNA palindromic site with reverted half-sites compared to CRE; **bd** are displayed in yellow, leucine zipper in green; reproduced with permission from reference 182. Copyright 2001 American Chemical Society.

Recently, Mascareñas and co-workers combined protein stitchery principles with the strategies previously developed for GCN4**bd** dimerisation, to design the first artificial dimer for which sequence preference is stimuli-responsive. For this, a GCN4**bd** peptide bearing a thiol group at the N-terminus and a side-chain functionalised with **terpy** towards the C-terminus was prepared. The monomer was unable to bind any of the DNA duplexes tested. However, the dimer prepared by formation of a disulfide N-terminal linkage, binds preferentially to the inverted GCN4 recognition site rather than the wild-type site, whereas, under reducing conditions, addition of Ni(II) resulted in preferential binding to the wild-type GCN4 DNA recognition site.[184]

1.3.3 – Stabilisation and destabilisation of helices to enhance or regulate DNA binding

As discussed in section 1.3.1, the **bd** is a short unfolded peptide fragment which folds into a α -helix conformation in the presence of its DNA target site. Therefore, it was proposed that stabilising a **bd** in a helical conformation would increase its affinity for the DNA target site. The residue grafting strategy developed by Schepartz and co-workers (see section 1.2.1) was applied to the design of short DNA binding peptides. In the first example, 13 residues which were identified as essential for the sequence selective DNA binding of GCN4, were grafted onto the exterior of the well-folded aPP scaffold while retaining their functional spacing, affording a 39 residue sequence (different combinations were attempted as three positions generated conflict with aPP key residues) (see Figure 1.19). One of these (PPBR2^{SR}) displayed affinity for the GCN4**bd** monomer recognition motif (5'-ATGAC-3') similar to the GCN4 dimers affinity for the full recognition sites (either 5'-ATGACGTCAT-3' or 5'-ATGACTCAT-3').[185] The DNA affinity was further increased 3-fold upon substitution of the three C-terminal residues of the peptide for alanine, as in PPBR4^{SR}. Moreover, peptide design was further refined using combinatorial mutations (screening and amplification of peptides or proteins based on a selection criteria, in this case binding to the GCN4 half-

recognition sequence) on four residues of PPBR2^{SR} located in the polyproline II domain. The most potent candidate, named p007, displays affinity for 5'-ATGAC-3' sites, 100 and 20,000 times higher than PPBR4^{SR} and the GCN4^{bd} monomer, respectively.[186]

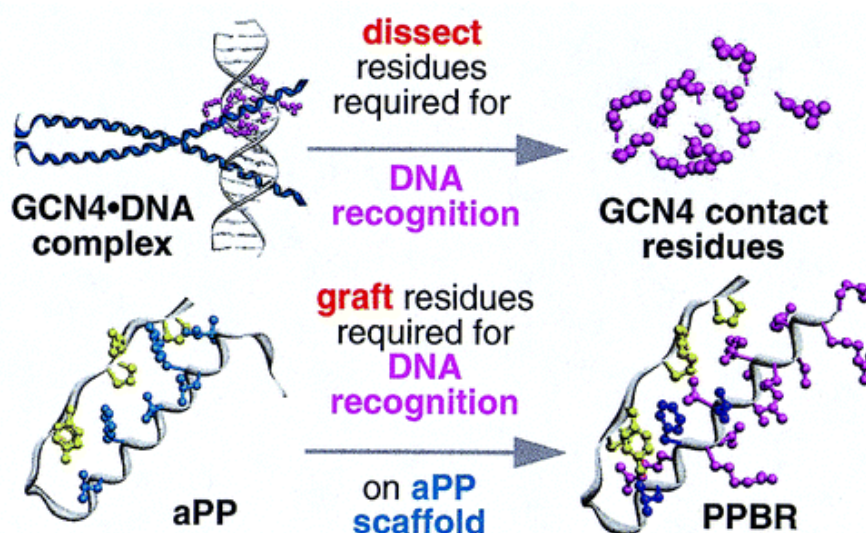


Figure 1.19 – Scheme illustrating the grafting of residues from GCN4^{bd}, essential for DNA binding (purple), onto the well-folded aPP scaffold (key residues from the helix and polyproline domains are highlighted in blue and yellow respectively) affording PPBR peptides; reproduced from reference 185. Copyright 1999 American Chemical Society.

Schepartz and co-workers further applied the residue grafting strategy to the design of short DNA binding peptides based on the engrailed homeodomain. Thirteen residues from helix 3 of Q50K engrailed mutant were grafted onto the aPP scaffold, affording the 34 residue peptide PPeng4, which displayed an affinity for the QRE recognition site (5'-TAATCC-3') 100 times higher than a peptide retaining the whole sequence of the helix 3. Despite not retaining the three residues from the N-terminal domain of Q50K responsible for discrimination of the 2 **bp** within the QRE site (underlined residues), PPeng4 binds with 32 times greater affinity to the QRE rather than the MRE site (5'-CCATCC-3').[187] Other noteworthy studies based on homodimer scaffolds used alanine,[188,189] or α -methylalanine enriched peptides [190] to increase α -helix stability and DNA binding.

The regulation of α -helical content using azobenzene photoswitches by Woolley and co-workers (see section 1.2.1) was equally applied to the design of peptides with light-responsive DNA binding or transcription activity. The helicity of a peptide mimicking the GCN4 zipper domain cross-linked at positions $f(i \rightarrow i+7)$ was first shown to increase upon light-irradiation at 370 nm, which promotes conversion of azobenzene linkers towards the *cis*-conformation.[191] Dimerisation and sequence specific DNA binding of this peptide coupled to the GCN4**bd**, were subsequently shown to increase slightly upon light-irradiation at 370 nm.[192] Further development involved the preparation of azobenzene cross-linked peptides mimicking the Fos zipper domain, which were shown to inhibit DNA binding of the Jun**bz** homodimer and, more importantly, of the bioactive Fos**bz**/Jun**bz** heterodimer (AP-1) upon light-irradiation at 365 nm (associated with conversion to the *cis*-azobenzene cross-linker and an increased helical content). This strategy relies on formation of heterodimers that lack one **bd**, hence their reduced DNA affinity (see Figure 1.20), and has been shown to efficiently reduce AP-1 transcription activity *in vivo*. [193]

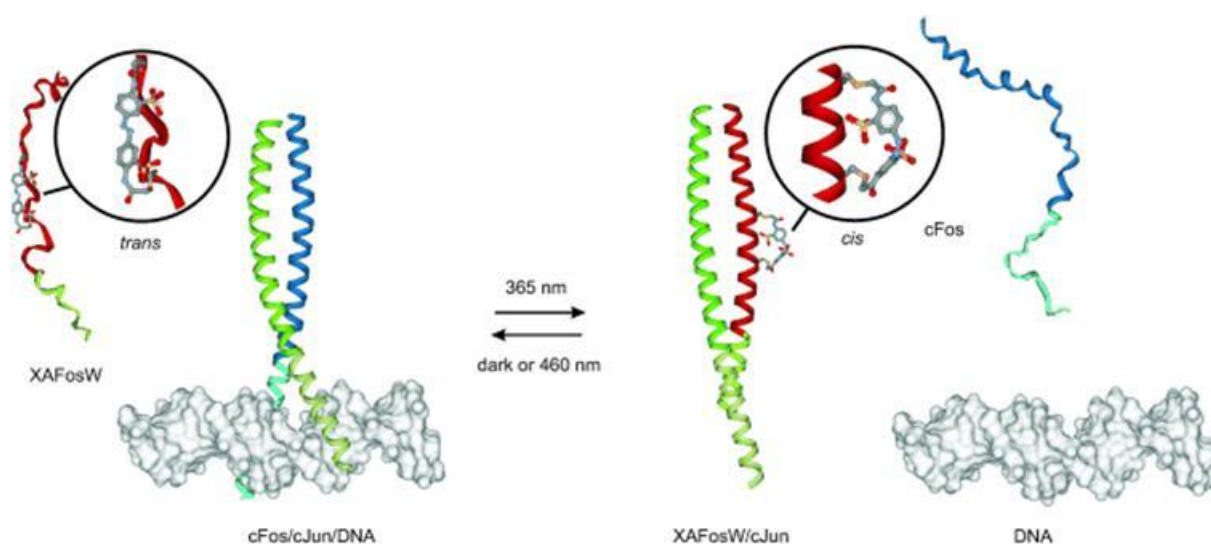


Figure 1.20 – Cartoon illustrating the indirect photoregulation of the Fos**bz**/Jun**bz** heterodimer DNA binding by the repressor XAFosW which incorporates an azobenzene cross-linker ($i \rightarrow i+7$) within a peptide mimicking the Fos leucine zipper but lacking a **bd**. Upon irradiation at 365 nm, the azobenzene converts to the *cis*-conformation associated with an α -helical conformation of XAFosW, which subsequently competes with Fos**bz** for dimerisation, thus inhibiting DNA binding; reproduced with permission from reference 193. Copyright 2010 Wiley-VCH.

The azobenzene cross-linker also allowed for sequence specific DNA binding photoregulation, when inserted in the **bd** of the MyoD homodimers,[194] or the Q50K engrailed homeodomain mutant.[195] Inouye and co-workers also investigated the photoregulated interactions between DNA and peptides mimicking different homeodomain proteins and bearing diarylethene linkers.[196] Futaki and co-workers equally applied the metal-induced helix destabilisation strategy (see section 1.2.1 and 1.2.3) to GCN4**bz** mimics, by introducing pairs of *Ida* residues at positions *c*, *e* or/and *g*, *b* ($i \rightarrow i+2$) of the GCN4 leucine zipper. They found that two *Ida* pairs were required to significantly alter the DNA binding ability of the GCN4**bz** peptides, and that Co(II) was a more potent effector than either Ni(II) or Mn(II).[46,197] Efforts have also been directed towards the design of short DNA binding peptides incorporating engineered β -hairpin motifs targeting single-stranded rather than double-stranded DNA.[198]

1.4 – The coordination chemistry of polypyridines

1.4.1 – Bipy metal complexes

In the crystal structure of the $[\text{Fe}(\text{bipy})_3]^{2+}$ complex, a low spin d^6 (spin singlet) Fe(II) occupies a distorted octahedral coordination site made of six **bipy** nitrogens,[199] and the six electrons on its d-orbitals are in a low spin arrangement (spin singlet).[200,201] The racemic mixtures of Δ and Λ enantiomers (see Figure 1.21) can be resolved by use of a chiral counteranion, thus modifying the optical rotation of the mixture (Pfeiffer effect [202]),[203] or alternatively enriched using an adsorbing clay.[204] Upon degradation of $[\text{Fe}(\text{bipy})_3]^{2+}$ (either pyrolysis or light-induced), few $[\text{Fe}(\text{bipy})_2\text{L}_2]^{2+}$ complexes have been obtained, and the Fe(II) environments were best described as distorted octahedrons where **bipy** units are positioned *cis*- to each other.[201] Even though $[\text{Fe}(\text{bipy})_2\text{L}_2]^{2+}$ complexes usually adopt high spin-states (quintet state), the heteroleptic complex $[\text{Fe}(\text{bipy})_2(\text{NCS})_2]^{2+}$ was shown to undergo a spin crossover upon raising the temperature.[201,205,206] I was unable to find a

crystal structure of $[\text{Fe}(\text{bipy})_2(\text{NCS})_2]^{2+}$ (except for a retracted paper [207]) or any other mixed-ligand Fe(II)-**bipy** complexes, however, it must be noted that the crystal structure of the analogous 1,10-phenanthroline (**phen**) containing complex, $[\text{Fe}(\text{phen})_2(\text{NCS})_2]^{2+}$, has been reported.[208] Fe(II) and **bipy** form almost exclusively the $[\text{Fe}(\text{bipy})_3]^{2+}$ complex in aqueous solution, except when a huge excess of Fe(II) is present, consistent with the high $\log\beta_3$ value and $K_3 > K_1, K_2$ (see Table 1.1).[209] In the absence of a chiral counteranion, the two enantiomers, Δ and Λ , have been shown to interconvert in solution.[210]

Table 1.1 – Summary of the formation stability constant for M(II):n**bipy** metal complexes with M = Fe, Co, Ni, Cu, Zn (measurements performed at 298 K and ionic strength $\mu = 0.1$ M) and involving either a partition procedure,[209] potentiometric measurements (either pH,[211-215] or pAg-metric [213,214]), or redox measurements.[216]

Metal ion	$\log K_1$	$\log K_2$	$\log K_3$	Ref.
Fe(II)	4.3	3.7	9.5	[209]
Co(II)	5.7	5.6	4.8	[209]
	6.1	5.4	4.6	[211]
Ni(II)	7.1	6.8	6.2	[209]
	7.1	6.9	6.5	[211]
Cu(II)	8.1	5.5	3.4	[209]
	8.0	5.6	3.5	[211]
	9.1 ^c	5.9 ^c	3.3 ^c	[212]
	8.5 ^a	5.6 ^a	3.5 ^a	[213]
	8.2 ^b	5.5 ^b	3.3 ^b	[216]
Zn(II)	5.2	4.4	3.8	[209]
	5.3	4.5	3.8	[211]
	5.1	4.8	3.2	[213]
	5.2	4.5	3.7	[214]
	5.3 ^{c,d}	4.6 ^{c,d}	4.0 ^{c,d}	[215]

Ionic strength was set to either 0^a, 0.3^b, 1.0^c M; ^d measurement performed at 303.3 K.

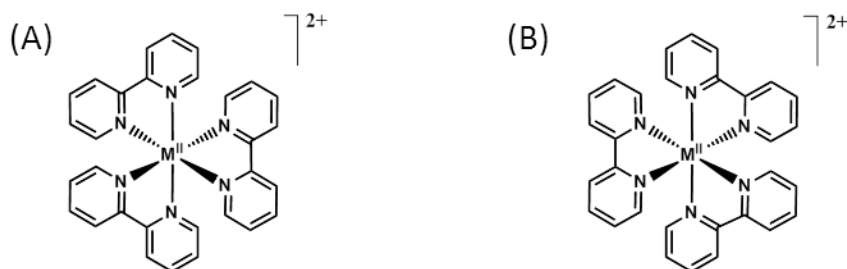


Figure 1.21 – Representation of the (A) Λ -, and (B) Δ enantiomers for the homoleptic $[M(\text{bipy})_3]^{2+}$ complex.

Similar to $[\text{Fe}(\text{bipy})_3]^{2+}$, the Co(II) cation adopts a distorted octahedral coordination geometry in $[\text{Co}(\text{bipy})_3]^{2+}$, but is d^7 high spin (spin quartet),[217] and thus $[\text{Co}(\text{bipy})_3]^{2+}$ readily oxidises to $[\text{Co}(\text{bipy})_3]^{3+}$. [218,219] As for the Fe(II) analogue, the two $[\text{Co}(\text{bipy})_3]^{2+}$ enantiomers can be separated based on the Pfeiffer effect.[203] Heteroleptic Co(II)-bipy complexes have been prepared either in solution,[220,221] or by decomposition of $[\text{Co}(\text{bipy})_3]^{2+}$ which requires less energy than for their Fe(II) analogues.[222,223] Structural investigation of $[\text{Co}(\text{bipy})_2]^{2+}$ indicates that Co(II) also adopts a distorted octahedral site and a high spin arrangement.[221,224-226] In solution, $[\text{Co}(\text{bipy})_3]^{2+}$ dissociation occurs faster than its Fe(II) analogue (see Table 1.2),[227] and complexes with different stoichiometry are more likely to be formed.[209]

Table 1.2 – Summary of the kinetic formation (k_f) and dissociation constant (k_d) for $1\text{M(II)}:n\text{bipy}$ metal complexes with $\text{M} = \text{Fe}, \text{Co}, \text{Ni}, \text{Cu}, \text{Zn}$ (calculated at 298 K). Species were quantified using spectroscopic measurement.[210,227-229] ML_2 complexes could not be isolated.

Metal ion	$\log k_1$	$\log k_3$	$\log k_{-1}$	$\log k_{-3}$	Ref.
Fe(II)	5.2	/	/	/	[228]
	5.2	5.1	/	/	[229]
	/	/	/	-3.1 ^a	[210]
Co(II)	4.8	/	/	/	[228]
	/	/	/	1.55 ^b	[227]
Ni(II)	3.2	3.3	-4.3	-2.5	[228]
Cu(II)	> 7.0	/	-0.7	/	[228]
Zn(II)	6.0	/	1.2	/	[228]

Ionic strength was set to 1.5^a, or 2^b M.

As for the Co(II) and Fe(II) analogues, the Ni(II) cation adopts a distorted octahedral environment in $[\text{Ni}(\text{bipy})_3]^{2+}$, which is associated with the high spin d^8 (triplet state).[230] The $[\text{Ni}(\text{bipy})_3]^{2+}$ cation can be resolved as salts which rapidly racemise in aqueous solution.[203,231] The structure of the complexes $[\text{Ni}(\text{bipy})_2\text{Cl}_2]$, $[\text{Ni}(\text{bipy})_2(\text{NCO})_2]$ and $[\text{Ni}(\text{bipy})_2(\text{H}_2\text{O})_2]^{2+}$ have been solved, displaying Ni(II) in a distorted octahedral environment with **bipy** units in a *cis*- positions.[232-234] Despite being more thermodynamically stable than $[\text{Fe}(\text{bipy})_3]^{2+}$, the complex $[\text{Ni}(\text{bipy})_3]^{2+}$ is more labile in solution, probably due to the fact that $K_3 < K_1, K_2$ (see Table 1.1).[209] Thus, complexes with Ni(II)-**bipy** stoichiometry of 1:1 or 1:2 can also be formed in solution, even though there exist few solid-state structural studies.[231] The kinetics of **bipy** association and dissociation of the Ni(II)-**bipy** complexes are much slower compared to other metals of interest, making it a suitable choice for studying substitution mechanism (see Table 1.2).[228]

The homoleptic complex $[\text{Cu}(\text{bipy})_3]^{2+}$ can be prepared from the specific Cu(II) salts and excess **bipy**, affording a mixture of Δ and Λ enantiomers in dynamic exchange and could not be isolated.[231] Structural study of the complex $[\text{Cu}(\text{bipy})_3]^{2+}$ indicates that Cu(II) occupies a distorted octahedral environment, due to Jahn-Teller distortions.[235] Upon reduction, $[\text{Cu}(\text{bipy})_3]^{2+}$ loses a coordinated **bipy** affording the $[\text{Cu}(\text{bipy})_2]^+$ cation, which readily oxidises in the presence of dioxygen, forming $[\text{Cu}(\text{bipy})_2]^{2+}$. [231] In $[\text{Cu}(\text{bipy})_2]^+$ complexes, Cu(I) occupies a tetrahedral environment,[231,236] however, there exists different geometry for the $[\text{Cu}(\text{bipy})_2]^{2+}$ cation.[237]

The Cu(II) coordination sphere in the structures of the $[\text{Cu}(\text{bipy})_2(\text{BF}_4)][\text{BF}_4]$ and $[\text{Cu}(\text{bipy})_2(\text{ClO}_4)][\text{ClO}_4]$ complexes are best described as elongated octahedron in which the two **bipy** units occupy *trans*- positions (the base of the xy plane of an octahedron elongated in z) and the bridging counteranions occupy the elongated axial (z) positions (see Figure 1.22A).[238] In contrast, both counteranions are weakly associated in $[\text{Cu}(\text{bipy})_2][\text{PF}_6]_2$, resulting in a tetracoordinated Cu(II) cation (see Figure 1.22B).[239] The tetragonal

arrangement is favoured due to the steric strain associated with the presence of two **bipy** units in the same plane (H6 and H6' protons, see Figure 2).[240] In the crystal structure of the cations $[\text{Cu}(\text{bipy})_2\text{Cl}]^+$ and $[\text{Cu}(\text{bipy})_2(\text{H}_2\text{O})]^{2+}$, Cu(II) adopt trigonal bipyramid coordination geometries (more or less distorted depending on the counter-anion), in which the two **bipy** are in a *cis*- arrangement (see Figure 1.22C).[241-244] Finally, the crystal structure of the complex $[\text{Cu}(\text{bipy})_2(\text{ONO}_2)]^+$, where $(\text{NO}_3)^-$ acts as a weakly coordinating dichelate, displays Cu(II) in a distorted octahedral environment (either 4+2 or 4+1+1 depending on the counteranion), where **bipy** units are in *cis*- arrangements (see Figure 1.22D).[245] To the best of my knowledge, no crystal structure of *cis*- $[\text{Cu}(\text{bipy})_2\text{L}_2]^{2+}$ complexes, where Cu(II) occupies a distorted octahedron and L represents a monodentate ligand, has been reported.

Cu(II) complexes bearing a single **bipy** unit have been isolated in the solid state. Examples are complexes of the type $[\text{Cu}(\text{bipy})\text{Cl}_2]_n$ involving one and two μ -bridging chloro-ligand(s), in which Cu(II) adopt either a distorted square based pyramid,[246,247] or an octahedral geometry, respectively.[248]

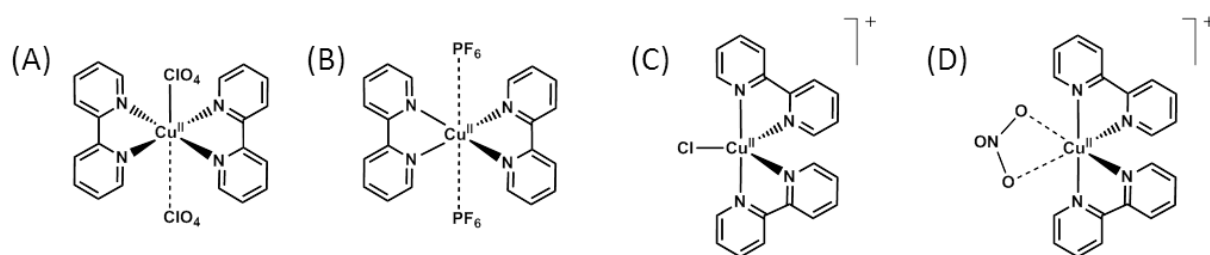


Figure 1.22 – Representation of the four main Cu(II) coordination geometries in the $[\text{Cu}^{\text{II}}(\text{bipy})_2]^{2+}$ cation where the **bipy** ligands occupy either trans- (A,B) or *cis*- positions (C,D): (A) octahedron $[\text{Cu}(\text{bipy})_2(\text{ClO}_4)]^+[\text{ClO}_4]^-$; (B) tetragonal $[\text{Cu}(\text{bipy})_2]^{2+}[\text{PF}_6]_2^-$; (C) trigonal bipyramid $[\text{Cu}(\text{bipy})_2\text{Cl}]^+$; (D) distorted octahedron $[\text{Cu}(\text{bipy})_2(\text{NO}_3)]^+$ (4+2).[237]

It is important to note that complexes from the Cu(II)-**bipy** systems are labile and undergo disproportion in solution, especially if alternative ligands are present.[249] As an example, some catalytic properties of $[\text{Cu}(\text{bipy})_3]^{2+}$ complexes might partly arise from the presence of $[\text{Cu}(\text{bipy})_2]^{2+}$, in which it disproportionates in anaerobic solutions.[231] The kinetics of association for the Cu(II)-**bipy** systems are too fast to be accurately measured,

however the complex dissociation is slow (see Table 1.2). Cu(II)-**bipy** complexes hydrolyse in aqueous solution, and the distribution of mono-, bis- and/or tris- chelated species is highly pH-dependent.[250-253] Based on absorbance measurements and continuous variation plots, Pfaulm et al. hypothesised that complexes with stoichiometry 2**bipy**:1Cu(II) predominate at pH 4, whereas 1**bipy**:1Cu(II) complexes are present at pH 12, following complex hydrolysis.[250] Cu(II)-**bipy** binding constants estimated in aqueous solution indicate that $K_1 > K_2 > K_3$ (see Table 1.1), thus favouring the formation of complexes with 1:1 and 1:2 stoichiometry. Studies conducted in EtOH/water mixtures resulted in slightly lower binding constants,[213] associated with different EPR parameters.[252,254] Species distribution diagrams of the **bipy**-Cu(II) system as a function of pH have been reported (see Figure 1.23).[212,252] When **bipy** and Cu(II) are present in equimolar amounts, complexes with the general formula $[\text{Cu}(\text{bipy})(\text{OH})]_n^{n+}$ predominate at physiological pH (see Figure 1.23A).[212,252] The complex $[\text{Cu}(\text{bipy})(\text{OH})]_2^{2+}$, which prevails at concentrations close to or greater than 50 mM, can be differentiated from its analogue $[\text{Cu}(\text{bipy})(\text{OH})]^+$, which constitutes the main species at lower concentrations, based on EPR or electronic absorption measurements (additional symmetry centre is present in the $n = 2$ species).[252] Recently, the complex $[\text{Cu}(\text{bipy})(\text{OH})_2]$, which predominates at slightly higher pH, was shown to be an efficient water-oxidation catalyst.[255]

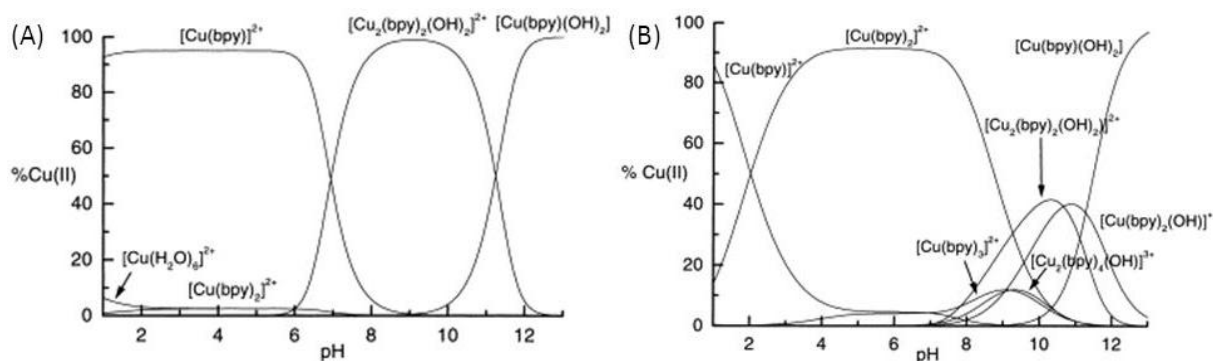


Figure 1.23 – Representation of Cu(II) species distribution diagram as a function of pH for the Cu(II)-**bipy** system (either (A) 1:1 or (B) 1:2 ratio) and $[\text{Cu}] = 5 \text{ mmol}$. Reproduced with permission from reference 252. Copyright 2000 Elsevier.

In contrast, when **bipy** is present in excess, complexes of the type $[\text{Cu}(\text{bipy})_2]^{2+}$ and $[\text{Cu}(\text{bipy})_3]^{2+}$ represent the main species at physiological pH (see Figure 1.23B). $[\text{Cu}(\text{bipy})_2]^{2+}$ species were successively proposed to adopt *trans*- (square planar Cu(II) coordination with tetragonal distortion) or *cis*- arrangement of **bipy** ligands (associated with square-based pyramid or trigonal bipyramid coordination).[256,257] EPR measurements and spectra simulation indicate that the *trans*- species is present in aqueous conditions, whereas the addition of ethanol (40 %) promotes formation of the *cis*- species (square pyramid).[252]

Cu(II) can alternatively form ternary complexes with **bipy** and other ligands in solution, the most relevant to this thesis are amino-acids and derivatives thereof.[258,259] In the solid state, the crystal structure of various Cu(II)-**bipy**-glycine and Cu(II)-**bipy**-glycolic acid complexes have been solved, in which Cu(II) adopts either distorted octahedrons,[260,261] or distorted square based pyramids (both mono- [262-264] or dinuclear complexes [265]). Solution studies of various Cu(II)-**bipy**-dipeptide ternary systems by Sigel and co-workers, indicates that the presence of **bipy** tends to promote the coordination of the deprotonated amide groups over side-chain substituents. [266-268] In all cases, the species distribution diagram indicates the presence of multiple species at physiological pH, where complexes with stoichiometry 1:1:1 (Cu(II)-**bipy**-dipeptide) generally prevail for equimolar solutions.[267]

Similarly to the Cu(II)-**bipy** system, complexes with 1Zn(II):1**bipy**, 1Zn(II):2**bipy**, or 1Zn(II):3**bipy** stoichiometry have been studied.[269,270] The Zn(II) cation adopts an octahedral environment in the crystal structure of the homoleptic complex $[\text{Zn}(\text{bipy})_3]^{2+}$. [231,270]. In the crystal structures of complexes $[\text{Zn}(\text{bipy})_2(\text{ONO})]^+$ and $[\text{Zn}(\text{bipy})_2(\text{OCO}(\text{Me}))]^+$, the Zn(II) cation adopts a distorted hexa-coordinated environment (pseudo-octahedral) where the two **bipy** are in a *cis*-arrangement (see Figure 1.22D).[271-273] However, no crystal structure of $[\text{Zn}(\text{bipy})_2]^{2+}$ cation where **bipy** are in *trans*-

arrangement have to the best of my knowledge been reported. In the complexes $[\text{Zn}(\text{bipy})\text{Cl}_2]^{2+}$ and $[\text{Zn}(\text{bipy})\text{Br}_2]^{2+}$, the Zn(II) cations adopt slightly distorted tetrahedral environments.[270]

Analysis of **bipy**-Zn(II) complexes formed in solution are complicated by the lack of spectroscopic feature for the Zn(II) cation, which is EPR silent and does not display any d-d transitions in the UV absorbance spectra. Based on the measured stability constant which indicates $K_1 > K_2 > K_3$ (see Table 1.1), it is reasonable to assume that all mono-, bis, and tris-chelates are potentially present in aqueous solution depending on the pH, the **bipy**:Zn(II) ratio and the concentration.[209] The association of the 1**bipy**:1Zn(II) complex is fast and its dissociation slow, even if this is less pronounced compared to the Cu(II) analogue (see Table 1.2).

Ishiguro and co-workers reported the species distribution diagram for the binary system **bipy**-Zn(II) in DMF indicating that a complex with 1**bipy**:1Zn(II) stoichiometry is the main species in the sub millimolar range, whereas 2**bipy**:1Zn(II) and 3**bipy**:1Zn(II) complexes predominate at higher concentrations.[274] However, none were reported (to the best of my knowledge) in water, where **bipy**-Zn(II) stability constants are higher (ca one log unit) compared to DMF.

No crystal structure have been reported for ternary systems **bipy**-Zn(II)-glycine or **bipy**-Zn(II)-glycolic acid (to the best of my knowledge). However, the ternary system **bipy**-Zn(II)-Aha (where Aha stands for hydroxamic acid) was studied in aqueous solution. The species distribution diagram for **bipy**-Zn(II)-Aha, when present in a ratio 2:1:2, indicates that the sum of the molar fraction of four different $[\text{Zn}(\text{bipy})_p(\text{Aha})_q]$ complexes (with $p = 0-1$ and $q = 1-2$) represents 90 % of the species present under these conditions.[275]

1.4.2 – Terpy metal complexes

Complexes with 1M(II):2**terpy** and 1M(II):1**terpy** stoichiometry (with M = Fe, Co, Ni, Cu) have been isolated in the solid-state (see Figure 1.24). Fe(II) adopts a distorted octahedral coordination geometry in $[\text{Fe}(\text{terpy})_2]^{2+}$.^[276,277] However, structural studies of $[\text{Fe}(\text{terpy})]^{2+}$ are consistent with a pentacoordinated Fe(II), which adopts a trigonal bipyramid coordination environment.^[278,279] The Fe(II) d-electrons in $[\text{Fe}(\text{terpy})]^{2+}$ have been shown to adopt high-spin states (quintet),^[278,279] contrasting with the low-spin state (singlet) associated with $[\text{Fe}(\text{terpy})_2]^{2+}$.^[278,280] Even though protonation of **terpy** and $[\text{Fe}(\text{terpy})_2]^{2+}$ hydrolysis can occur in aqueous solution, the $[\text{Fe}(\text{terpy})_2]^{2+}$ complex is highly stable, consistent with $K_2 > K_1$ (see Table 1.3),^[281,282] and hence represents the main species in solution (when Fe(II) is not present in large excess). Remarkably, kinetic measurements indicate that binding of the second **terpy** occurs more than 100 times faster than the first **terpy** (see Table 1.4).^[229]

Table 1.3 – Summary of the formation stability constants for M(II):n**terpy** complexes with M = Fe, Co, Ni, Cu, Zn (calculated at 298 K and ionic strength $\mu = 0.1$ M), involving either potentiometric,^[283] spectroscopic,^[229,284] or redox measurements.^[216]

Metal ion	log K_1	log K_2	Ref.
Fe(II)	7.1	13.6	[283]
Co(II)	9.5	9.1	[283]
Ni(II)	10.7	11.1	[283]
Cu(II)	$\sim 13^{\text{a,b}}$	/	[216]
	12.3	6.8	[283]
Zn(II)	6.0	/	[229,283]
	7.6^{c}	/	[284]

^a based on spectroscopic measurement performed in dioxane/water (1/1); Ionic strength were set to ^b0.3 or ^c0 M .

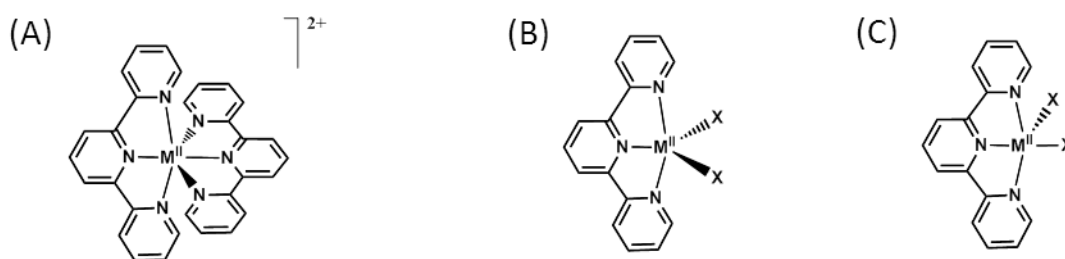


Figure 1.24 – Representation of the three main complexes formed between M(II) and **terpy**: (A) the $[M(\text{terpy})_2]^{2+}$ complex in an approximate D_{2d} geometry, and the $[M(\text{terpy})X_2]$ either in (B) a trigonal bipyramid for $M = \text{Fe}, \text{Co}, \text{Ni}, \text{Zn}$, or (C) in a square-based pyramid for $M = \text{Cu}$. X corresponds to halogenates.

Table 1.4 – Summary of the kinetic formation (k_n) and dissociation rate constant (k_{-n}) for M(II):**nterpy** metal complexes with $M = \text{Fe}, \text{Co}, \text{Ni}, \text{Cu}, \text{Zn}$ (calculated at 298 K). For the first set of data, species were quantified using spectroscopic measurement,[229] whereas the second set involved $[\text{H}^3]$ -terpyridine and radioassays.[285]

Metal ion	$\log k_1$	$\log k_2$	$\log k_{-1}$	$\log k_{-2}$	Ref.
Fe(II)	4.9	$\sim 7^a$	-2.2	-6.8	[229]
	/	/	3.8	-0.77	[285]
Co(II)	4.4	6.7^a	-4.0	-3.2	[229]
	/	/	2.0	2.8	[285]
Ni(II)	3.1	5.3	-7.6	-5.8	[229]
	/	/	-1.6	0.22	[285]
Cu(II)	$\sim 7.3^b$	/	/	/	[229]
	/	/	$> 5.1^c$	/	[285]
Zn(II)	6.1	/	0.1	/	[229]

Rate constant calculated at ^a278, ^b279.5, and ^c273.1 K.

In the solid state, $[\text{Co}(\text{terpy})_2]^{2+}$ exists as a mixture of two species, where the electronic configuration of the Co(II) d-orbitals adopt either a high (quartet) or a low-spin state (doublet). The nature of the counter-anions and the hydration have been shown to impact the spin-state distribution,[286,287] however, no crystal with an exclusive spin-state has been isolated. In all cases, Co(II) adopts a distorted octahedral geometry in $[\text{Co}(\text{terpy})_2]^{2+}$, due to Jahn-Teller effects. In contrast, the crystal structures of 1Co(II):1**terpy** complexes, such as

$[\text{Co}(\text{terpy})\text{Br}_2]^{2+}$, display Co(II) cations in trigonal bipyramidal coordination geometries (mostly high spin in this case).[288,289] In solution, **terpy** derivatives have been shown to form both 1Co(II):1**terpy** and 1Co(II):2**terpy** complexes, consistent with $K_2 \sim K_1$ (see Table 1.3), and the exact composition depends on the ratio of species present,[285,290] but also from the solvent.[291] As for Fe(II) complexes, kinetic measurements indicate that the binding of the second **terpy** motif is more than 100 times faster than the first **terpy** (see Table 1.4).[229]

As for Co(II), structural studies of complexes with 1Ni(II):2**terpy** and 1Ni(II):1**terpy** indicate that the Ni(II) cation adopts an octahedral,[292-294] or trigonal bipyramid coordination geometry (both distorted),[279,295] respectively. The stability constant calculated are again consistent with both, 1Ni(II):2**terpy** and 1Ni(II):1**terpy** species, being present in solution (consistent with $K_2 \sim K_1$, see Table 1.3), and again kinetic data indicate that coordination of the second **terpy** is more than 100 time faster than the first one (see Table 1.4).[229]

The homoleptic complex $[\text{Cu}(\text{terpy})_2]^{2+}$ is readily prepared when reacting excess **terpy** with Cu(II). In the crystal structure, the Cu(II) cation adopts an octahedral site with an orthorhombic Jahn-Teller distortion.[296] In contrast, Cu(II) adopts a distorted square-based pyramid in the crystal structure of the complex $[\text{Cu}(\text{terpy})_2\text{Cl}_2]$ (see Figure 1.24C).[297] In aqueous solution and at room temperature, spectroscopic evidence supports the formation of complexes with 1**terpy**:1Cu(II) stoichiometry,[216,298-300] even at acidic and basic pH.[250] This is consistent with the binding constants for which $K_2 < K_1$ (see Table 1.3). Kinetic measurements indicate that both association and dissociation of the 1**terpy**:1Cu(II) complex are fast (see Table 1.4). When excess **terpy** is present, the complex $[\text{Cu}(\text{terpy})_2]^{2+}$ was reported to precipitate from solution,[285] however, spectroscopic evidence and binding constants suggest it might be present in solution as a minor species (see Table 1.3).[216,299] As for **bipy**, crystal structures of ternary complexes **terpy**-Cu(II)-glycine have been reported,

where glycine acts as an axial monodentate ligand in the square-based pyramid coordination environment of Cu(II).[301]

Even though the literature seems to lack data for complexes of the type $[\text{Zn}(\text{terpy})_2]^{2+}$, the structure of analogous complexes involving substituted **terpy** have been reported, in which the Zn(II) cation adopts a distorted octahedral site.[302] The structure of the complex $[\text{Zn}(\text{terpy})\text{Cl}_2]$, where Zn(II) adopts a trigonal bipyramid, is also known.[303,304] In aqueous solution and at room temperature, potentiometric and spectroscopic studies both indicate formation of a complex with 1**terpy**:1Zn(II) stoichiometry.[229,283,284] Despite a fast association process, the decomposition of the complex $[\text{Zn}(\text{terpy})]^{2+}$ was very slow at room temperature (see Table 1.4).

1.4.3 – Impact of polypyridine substituents on coordination

The introduction of substituents on the polypyridine rings have been shown to impact the ability to bind protons or metal ions, depending on their nature and position. For example, Fe(II) coordination to **bipy** is prevented upon symmetrical dimethyl substitution at the 3,3'-positions, due to sterically disfavoured the *cis*- conformation.[305,306] Most reports deal with the coordination ability of **bipy** substituted on 6- and 3- positions which are governed by steric effects (see Figure 1.25). However, coordination studies on **phen** indicate that electronic effects might also play a minor role. For instance, substitution of **phen** at either the 4,7- or 3,8 positions slightly increases or decreases the metal ion affinity (respectively), consistent with π -delocalisation promoting metal-ion binding in the former case only.[306,307] Despite substitution at position 2- and/or 9- of **phen** (equivalent to 6,6'-for **bipy**) equally being able to favour coordination through electronic effects, the metal-ion affinity is in fact much lower than for **phen**. [308]

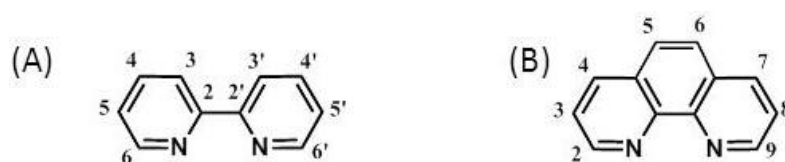


Figure 1.25 – Representation of the numbering scheme for (A) **bipy** and (B) **phen** substituent positions.

Similarly, the **bipy** metal ion affinity significantly decreases upon introduction of 6,6' as opposed to 4,4' or 5,5'- substituents.[216] Interestingly, this allows for the preparation of polypyridine metal complexes with alternative coordination geometries,[309-311] or metal:ligand stoichiometries.[312] On a side-note, 5,5'-dimethyl substituents were reported to sterically hinder the formation of dinuclear **bipy**:M(II) complexes, thus promoting the formation of mononuclear analogues.[313]

1.5 – Summary

This Chapter briefly introduces the main secondary structure of peptides and the design of peptide mimics with stimuli- responsive and stabilised conformations. Protein interactions often rely on α -helices, and chemists have prepared a large variety of short peptide fragments stabilised in this conformation, using different strategies, able to recognise their specific target site. Moreover, incorporation of stimuli-responsive elements may allow for the control of biomolecular activity by artificial conformational regulation. Examples are emerging of biomolecular recognition involving β -strands, artificial mimics of which can be stabilised by promoting interactions between two or more strands. This can be achieved by introduction of side chains which promote such interactions, or incorporation of turn mimics which stabilise the chain reversal (U-shaped). So far, most polypyridine peptide conjugates prepared are cyclometallopeptides, which may serve as large turn mimics with specific chiral orientation or photoactive properties. Taken together, these studies might allow for the control of peptide tertiary or quaternary structures thus allowing for the development of functional peptides

capable of, for example introducing pores within membranes, or for the preparation of new peptide-based materials.

Chromosomal DNA, which is packed in nucleosomes, is a biomolecular target for peptides involved in initiation and regulation of protein transcription. Unlike some of the proteins involved in DNA maintenance (copy, repair, or insertion), they often bind as multimers to symmetric recognition sites. These represent an important source of inspiration for the design of small molecules able to influence the transfer of genetic information. Miniature, highly specific and stimuli responsive transcription factors have been designed either using artificial dimerisation methods or artificial stabilisation of helices. They may ultimately allow for *in vivo* partial sequencing. Importantly, the access to alternative binding sites can be achieved through symmetry operations (stitchery), thus enlarging the scope of applications. Moreover, the identification of some natural proteins acting both as sensors and transcription regulators (in contrast to the *trans*- or *cis*- activation processes) is challenging and modulation of their activity with artificial inhibitors might help to identify their role.

The stability constants for complexes 1M(II):1**bipy** and 1M(II):1**terpy** follow the order predicted by the Irving-Williams series.[314] However this is no longer the case for complexes containing more than one polypyridine ligand.[209] Cu(II) and Zn(II) have been shown to preferentially form 1:1 complexes with equimolar amounts of **bipy** and **terpy** under physiological conditions, and are therefore of particular interest for this work.

1.6 – Project aims

The aim of this work is to prepare and study some peptide dimer conjugates, the dimerisation domains of which consist of disubstituted polypyridine linkers. As previously seen, the *cis*-to-*trans*- conformational transition experienced by polypyridines upon metal complexation, could trigger reorientation of its substituents, and might allow for switching on and off the biomolecular recognition of peptide dimers.

Chapter 2 reports the synthesis and study of low molecular weight peptide dimer polypyridine conjugates, for which conformational and coordination studies were attempted. Part of this chapter was published as: Oheix E., Spencer N., Gethings L. A., Peacock A. F. A., *Z. Anorg. Allg. Chem.*, **2013**, 639, 1370-1383, as an invited contribution.

Chapter 3 describes the preparation of larger peptides based on the **bd** of the GCN4 transcription factor, their dimerisation with the **bipy** and **terpy** linkers, and studies of the Cu(II) and Zn(II) coordination of the resulting conjugates. Part of this chapter was published as: Oheix E., Peacock A. F. A., *Chem. Eur. J.*, **2014**, doi: 10.1002/chem.201303747.

Chapter 4 describes the DNA binding studies of the GCN4**bd** polypyridine conjugates and the impact of metal addition on peptide-DNA interactions. Part of this chapter was published as: Oheix E., Peacock A. F. A., *Chem. Eur. J.*, **2014**, doi: 10.1002/chem.201303747.

Chapter 5 presents initial work towards application of GCN4**bd** polypyridine conjugates as DNA sensors or nuclease agents.

Chapter 6 concludes on the work done and develops new perspectives.

The appendices describe the experimental techniques involved in this work.

1.7 – References

- [1] Young J. C., Agashe V. R., Siegers K., Hartl F. U., *Nat. Rev. Mol. Cell Biol.*, **2004**, 5, 781-791.
- [2] Lee A. Y., Gulnik S. V., Erickson J. W., *Nat. Struct. Biol.*, **1998**, 5, 866-871.
- [3] Rabindran S. K., Haroun R. I., Clos J., Wisniewski J., Wu C., *Science*, **1993**, 259, 230-234.
- [4] Zhang S., Rich A., *Proc. Natl. Acad. Sci. USA*, **1997**, 94, 23-28.
- [5] Janin J., Chothia C., *J. Biol. Chem.*, **1990**, 265, 16027-16030.
- [6] Wilson C. J., Apiyo D., Wittung-Stafshede P., *Q. Rev. Biophys.*, **2004**, 37, 285-314.
- [7] Kiser P. D., Golczak M., Palczewski K., *Chem. Rev.*, dx.doi.org/10.1021/cr400107q

- [8] Joyce A. R., Palsson B. Ø., *Nat. Rev. Mol. Cell Biol.*, **2006**, 7, 198-210.
- [9] Avin-Wittenberg T., Galili G., *Nat. Chem. Biol.*, **2012**, 8, 23-24.
- [10] Chen P. R., He C., *Curr. Opin. Chem. Biol.*, **2008**, 12, 214-221.
- [11] Moore J. T., Collins J. L., Pearce K. H., *ChemMedChem*, **2006**, 1, 504-523.
- [12] Weigel N. L., Moore N. L., *Mol. Endocrinol.*, **2007**, 21, 2311-2319.
- [13] Garvie C. W., Wolberger C., *Mol. Cell*, **2001**, 8, 937-946.
- [14] Rohs R., Jin X., West S. M., Joshi R., Honig B., Mann R. S., *Annu. Rev. Biochem.*, **2010**, 79, 233-269.
- [15] Murzin A. G., Brenner S. E., Hubbard T., Chothia C., *J. Mol. Biol.*, **1995**, 247, 536-540.
- [16] Tse W. C., Boger D. L., *Chem. Biol.*, **2004**, 11, 1607-1617.
- [17] Dervan P. B., *Bioorg. Med. Chem.*, **2001**, 9, 2215-2235.
- [18] Jenette K. W., Lippard S. J., Vassiliades G. A., Bauer W. R., *Proc. Natl. Acad. Sci. USA*, **1974**, 71, 3839-3843.
- [19] Zeglis B. M., Pierre V. C., Barton J. K., *Chem. Commun.*, **2007**, 4565-4579.
- [20] Oleksy A., Blanco A. G., Boer R., Usón I., Aymami J., Rodger A., Hannon M. J., Coll M., *Angew. Chem. Int. Ed.*, **2006**, 45, 1227-1231.
- [21] Pauling L., Corey R. B., Branson H. R., *Proc. Natl. Acad. Sci. USA*, **1951**, 37, 205-211.
- [22] Davies D. R., *J. Mol. Biol.*, **1964**, 9, 605-609.
- [23] Paterson Y., Rumsey S. M., Benedetti E., Némethy G., Scheraga H. A., *J. Am. Chem. Soc.*, **1981**, 103, 2947-2955.
- [24] Sattler M., Liang H., Nettlesheim D., Meadows R. P., Harlan J. E., Eberstadt M., Yoon H. S., Shuker S. B., Chang B. S., Minn A. J., Thompson C. B., Fesik S W., *Science*, **1997**, 275, 983-986.
- [25] Kussie P. H., Gorina S., Marechal V., Elenbaas B., Moreau J., Levine A. J., Pavletich N. P., *Science*, **1996**, 274, 948-953.
- [26] Schegara H. A., *Pure & Appl. Chem.*, **1978**, 50, 315-324.

- [27] Bierzynski A., Kim P. S., Baldwin R. L., *Proc. Natl. Acad. Sci. USA*, **1981**, 79, 2470-2474.
- [28] Andrews M. J., Tabor A. B., *Tetrahedron*, **1999**, 55, 11711-11743.
- [29] Schoemaker K. R., Kim P. S., Brems D. N., Marqusee S., York E. J., Chaiken I. M., Stewart J. M., Baldwin R. L., *Proc. Natl. Acad. Sci. USA*, **1985**, 82, 2349-2353.
- [30] Jackson D. Y., King D. S., Chmielewski J., Singh S., Schultz P. G., *J. Am. Chem. Soc.*, **1991**, 113, 9391-9392.
- [31] Marqusee S., Baldwin R. L., *Proc. Natl. Acad. Sci. USA*, **1987**, 84, 8898-8902.
- [32] Huyghues-Despointes B. M. P., Baldwin R. L., *Biochemistry*, **1997**, 36, 1965-1970.
- [33] Cheng R. P., Wang W., Girinath P., Yang P., Ahmad R., Li J., Hart P., Kokona B., Fairman R., Kilpatrick C., Argiros A., *Biochemistry*, **2012**, 51, 7157-7172.
- [34] Tabet M., Labroo V., Sheppard P., Sasaki T., *J. Am. Chem. Soc.*, **1993**, 115, 3866-3868.
- [35] Haack T., Peczu M. W., Salvatella X., Sánchez-Quesada J., de Mendoza J., Hamilton A. D., Giralt E., *J. Am. Chem. Soc.*, **1999**, 121, 11813-11820.
- [36] Albert J. S., Goodman M. S., Hamilton A. D., *J. Am. Chem. Soc.*, **1995**, 117, 1143-1144.
- [37] Ghadiri M. R., Choi C., *J. Am. Chem. Soc.*, **1990**, 112, 1630-1632.
- [38] Ghadiri M. R., Fernholz A. K., *J. Am. Chem. Soc.*, **1990**, 112, 9633-9635.
- [39] Ma M. T., Hoang H. N., Scully C. C. G., Appleton T. G., Fairlie D. P., *J. Am. Chem. Soc.*, **2009**, 131, 4505-4512.
- [40] Ruan F., Chen Y., Hopkins P. B., *J. Am. Chem. Soc.*, **1990**, 112, 9403-9404.
- [41] Nagata Y., Kanuka S., Kinoshita T., Takizawa A., Tsujita Y., Yoshimizu H., *Biopolymers*, **1994**, 34, 701-707.
- [42] Tang J., Signarvic R. S., DeGrado W. F., Gai F., *Biochemistry*, **2007**, 46, 13856-13863.
- [43] Signarvic R. S., DeGrado W. F., *J. Am. Chem. Soc.*, **2009**, 131, 3377-3384.
- [44] Karpishin T. B., Vannelli T. A., Glover K. J., *J. Am. Chem. Soc.*, **1997**, 119, 9063-9064.

- [45] Arnold P. A., Benson D. R., Brink D. J., Hendrich M. P., Jas G. S., Kennedy M. L., Petasis D. T., Wang M., *Inorg. Chem.*, **1997**, 36, 5306-5315.
- [46] Futaki S., Kiwada T., Sugiura Y., *J. Am. Chem. Soc.*, **2004**, 126, 15762-15769.
- [47] Voyer N., Guérin B., *Tetrahedron*, **1994**, 50, 989-1010.
- [48] Matsumura S., Sakamoto S., Ueno A., Mihara H., *Chem. Eur. J.*, **2000**, 6, 1781-1788.
- [49] Olson C. A., Shi Z., Kallenbach N. R., *J. Am. Chem. Soc.*, **2001**, 123, 6451-6452.
- [50] Shi Z., Olson A., Kallenbach N. R., *J. Am. Chem. Soc.*, **2002**, 124, 3284-3291.
- [51] Albert J. S., Hamilton A. D., *Biochemistry*, **1995**, 34, 984-990.
- [52] Jackson D. Y., King D. S., Chmielewsky J., Singh S., Schultz P. G., *J. Am. Chem. Soc.*, **1991**, 113, 9391-9392.
- [53] Pellegrini M., Royo M., Chorev M., Mierke D. F., *J. Pept. Res.*, **1997**, 49, 404-414.
- [54] Osapay G., Taylor J. W., *J. Am. Chem. Soc.*, **1992**, 114, 6966-6973.
- [55] Phelan J. C., Skelton N. J., Braisted A. C., McDowell R. S., *J. Am. Chem. Soc.*, **1997**, 119, 455-460.
- [56] Schafmeister C. E., Po J., Verdine G. L., *J. Am. Chem. Soc.*, **2000**, 122, 5891-5892.
- [57] Blackwell H. E., Grubbs R. H., *Angew. Chem. Int. Ed.*, **1998**, 37, 3281-3284.
- [58] Muppidi A., Wang Z., Li X., Chen J., Lin Q., *Chem. Commun.* **2011**, 47, 9396-9398.
- [59] Jo H., Meinhardt N., Wu Y., Kulkarni S., Hu X., Low K. E., Davies P. L., DeGrado W. F., Greenbaum D. C., *J. Am. Chem. Soc.*, **2012**, 134, 17704-17713.
- [60] Chin J. W., Schepartz A., *Angew. Chem. Int. Ed.*, **2001**, 40, 3806-3809.
- [61] Gemperli A. C., Rutledge S. E., Maranda A., Schepartz A., *J. Am. Chem. Soc.*, **2005**, 127, 1596-1597.
- [62] Kritzer J. A., Zutshi R., Cheah M., Rann F. A., Webman R., Wongjirad T. M., Schepartz A., *J. Am. Chem. Soc.*, **2006**, 128, 29-31.
- [63] Flint D. G., Kumita J. R., Smart O. S., Woolley G. A., *Chem. Biol.*, **2002**, 9, 391-397.

- [64] Kumita J. R., Smart O. S., Woolley G. A., *Proc. Natl. Acad. Sci. USA*, **2000**, 97, 3803-3808.
- [65] Kneissl S., Loveridge E. J., Williams C., Crump M. P., Allemann R. K., *ChemBioChem*, **2008**, 9, 3046-3054.
- [66] Cabezas E., Satterthwait A. C., *J. Am. Chem. Soc.*, **1999**, 121, 3862-3875.
- [67] Patgiri A., Jochim A. L., Arora P. S., *Acc. Chem. Res.*, **2008**, 41, 1289-1300.
- [68] Kemp D. S., Curran T. P., Davis W. M., Boyd J. G., Muendel C., *J. Org. Chem.*, **1991**, 56, 6672-6682.
- [69] Kemp D. S., Boyd J. G., Muendel C. C., *Nature*, **1991**, 352, 451-454.
- [70] Hack V., Reuter C., Opitz R., Schmieder P., Beyermann M., Neudörfl M., Kühne R., Schmalz H., *Angew. Chem. Int. Ed.*, doi:10.1021/anie.201302014.
- [71] Orner B. P., Ernst J. T., Hamilton A. D., *J. Am. Chem. Soc.*, **2001**, 123, 5382-5383.
- [72] Smith J. A., Pease L. G., *Crit. Rev. Biochem.*, **1980**, 315-399.
- [73] Leszczynski J. F., Rose G. D., *Science*, **1986**, 234, 849-855.
- [74] Giannis A., Kolter T., *Angew. Chem. Int. Ed.*, **1993**, 32, 1244-1267.
- [75] Schneider J. P., Kelly J. W., *Chem. Rev.*, **1995**, 95, 2169-2187.
- [76] Struthers M. D., Cheng R. P., Imperiali B., *Science*, **1996**, 271, 342-345.
- [77] Eckhardt B., Grosse W., Essen L., Geyer A., *Proc. Natl. Acad. Sci. USA*, **2010**, 107, 18336-18341.
- [78] Merrifield R. B., *J. Am. Chem. Soc.*, **1963**, 85, 2149-2154.
- [79] Imperiali B., Fisher S. L., *J. Org. Chem.*, **1992**, 57, 757-759.
- [80] Imperiali B., Prins T. J., Fisher S. L., *J. Org. Chem.*, **1993**, 58, 1613-1616.
- [81] Cheng R. P., Fisher S. L., Imperiali B., *J. Am. Chem. Soc.*, **1996**, 118, 11349-11356.
- [82] Torrado A., Imperiali B., *J. Org. Chem.*, **1996**, 61, 8940-8948.
- [83] Bishop B. M., McCafferty D. G., Erickson B. W., *Tetrahedron*, **2000**, 4629-4638.

- [84] Rama G., Ardá A., Maréchal J., Gamba I., Ishida H., Jiménez-Barbèro J., Vázquez M. E., López M. E. V., *Chem. Eur. J.*, **2012**, 18, 7030-7035.
- [85] Constable E. C., Housecroft C. E., Mundwiler S., *Dalton Trans.*, **2003**, 2112-2114.
- [86] Albrecht M., Stortz P., *Chem. Soc. Rev.*, **2005**, 34, 496-506.
- [87] Venkatraman J., Shankaramma S. C., Balaram P., *Chem Rev.*, **2001**, 101, 3131-3152.
- [88] Razin J., Nillsson H., Baltzer L., *J. Am. Chem. Soc.*, **2007**, 129, 14752-14758.
- [89] Olofsson S., Johansson G., Baltzer L., *J. Chem. Soc. Perkin Trans. 2*, **1995**, 2047-2056.
- [90] Olofsson S., Baltzer J., *Folding Des.*, **1996**, 1, 347-456.
- [91] Lindgren N. J. V., Varedian M., Gogoll A., *Chem. Eur. J.*, **2009**, 15, 501-505.
- [92] Pauling L., Corey R. B., *Proc. Natl. Acad. Sci. USA*, **1951**, 37, 729-740.
- [93] Kim C. A., Berg J. M., *Nature*, **1993**, 362, 267-270.
- [94] Minor Jr D. L., Kim P. S., *Nature*, **1994**, 371, 264-267.
- [95] Richardson J. S., Richardson D. C., *Proc. Natl. Acad. Sci. USA*, **2002**, 99, 2754-2759.
- [96] Blanco F. J., Rivas G., Serrano L., *Nat. Struct. Biol.*, **1994**, 1, 584-590.
- [97] Sibanda B. L., Thornton J. M., *Nature*, **1985**, 316, 170-174.
- [98] Gellman S. H., *Curr. Opin. Chem. Biol.*, **1998**, 2, 717-725.
- [99] Platt G., Chung C., Searle M. S., *Chem. Commun.*, **2001**, 1162-1163.
- [100] Löwik D. W. P. M., Linhardt J. G., Adams P. J. H., van Hest J. C. M., *Org. Biomol. Chem.*, **2003**, 1, 1827-1829.
- [101] Celentano V., Diana D., De Rosa L., Romanelli A., Fattorusso R., D'Andrea L. D., *Chem. Commun.*, **2012**, 48, 762-764.
- [102] Diaz H., Tsang K. Y., Choo D., Espina J. R., Kelly J. W., *J. Am. Chem. Soc.*, **1993**, 115, 3790-3791.
- [103] Schneider J. P., Kelly J. W., *J. Am. Chem. Soc.*, **1995**, 117, 2533-2546.
- [104] Aemissegger A., Kräutler V., van Gunsteren W. F., Hilvert D., *J. Am. Chem. Soc.*, **2005**, 127, 2929-2936.

- [105] Kräutler V., Aemissegger A., Hünenberger P. H., Hilvert D., Hansson T., van Gunsteren W. F., *J. Am. Chem. Soc.*, **2005**, 127, 4935-4942.
- [106] Doran T. M., Anderson E. A., Latchney S. E., Opanashuk L. A., Nilsson B. L., *ACS Chem. Neurosci.*, **2012**, 3, 211-220.
- [107] Hoppmann C., Barucker C., Lorenz D., Multhaup G., Beyermann M., *ChemBioChem*, **2012**, 13, 2657-2660.
- [108] Crick F. H. C., *Nature*, **1952**, 170, 882-883.
- [109] Landschulz W. H., Johnson P. F., McKnight S. L., *Science*, **1988**, 240, 1759-1764.
- [110] Woolfson D. N., *Adv. Protein Chemistry*, **2005**, 70, 79-111.
- [111] Harbury P. H., Plecs J. J., Tidor B., Alber T., Kim P. S., *Science*, **1998**, 282, 1462-1467.
- [112] Liu J., Zheng Q., Deng Y., Cheng C., Kallenbach N. R., Lu M., *Proc. Natl. Acad. Sci.*, **2006**, 103, 15457-15462.
- [113] Moutevelis E., Woolfson D. N., *J. Mol. Biol.*, **2009**, 385, 726-732.
- [114] Zhou J., Case M. A., Wishart J. F., McLendon G. L., *J. Phys. Chem B*, **1998**, 102, 9975-9980.
- [115] Suzuki K., Doi T., Imanishi T., Kodama T., Tanaka T., *Biochemistry*, **1997**, 36, 15140-15146.
- [116] Doi T., Kurasawa M., Higashino K., Imanishi T., Mori T., Naito M., Takahashi K., Kawabe Y., Wada Y., Matsumoto A., Kodama T., *Proc. Natl. Acad. Sci. USA*, **1994**, 269, 25598-25604.
- [117] Handel T. M., Williams S. A., DeGrado W. F., *Science*, **1993**, 261, 879-885.
- [118] Kohn W. D., Kay C. M., Sykes B. D., Hodges R. S., *J. Am. Chem. Soc.*, **1998**, 120, 1124-1132.
- [119] Suzuki K., Hiroaki H., Kohda D., Nakamura H., Tanaka T., *J. Am. Chem. Soc.*, **1998**, 120, 13008-13015.
- [120] Lieberman M., Sasaki T., *J. Am. Chem. Soc.*, **1991**, 113, 1470-1471.

- [121] Gochin M., Khorosheva V., Case M. A., *J. Am. Chem. Soc.*, **2002**, 124, 11018-11028.
- [122] Case M. A., McLendon G. L., *Acc. Chem. Res.*, **2004**, 37, 754-762.
- [123] Doerr A. J., Case M. A., Pelczer I., McLendon G. L., *J. Am. Chem. Soc.*, **2004**, 126, 4192-4198.
- [124] Shai Y., *Biochim. Biophys. Acta*, **1999**, 1462, 55-70.
- [125] Zasloff M., *Nature*, **2002**, 415, 389-395.
- [126] Li W., Nicol F., Szoka Jr F., *Adv. Drug Deliv.*, **2004**, 56, 967-985.
- [127] Sasaki K., Kogure K., Chaki S., Nakamura Y., Moriguchi R., Hamada H., Danev R., Nagayama K., Futaki S., Harashima H., *Anal. Bioanal. Chem.*, **2008**, 391, 2717-2727.
- [128] Kiwada T., Sonomura K., Sugiura Y., Asami K., Futaki S., *J. Am. Chem. Soc.*, **2006**, 128, 6010-6011.
- [129] Woolley G. A., Wallace B. A., *Biochemistry*, **1993**, 32, 9819-9825.
- [130] O'Shea E. K., Rutkowski R., Stafford III W. F., Kim P. S., *Science*, **1989**, 245, 646-648.
- [131] Noshiro D., Sonomura K., Yu H. H., Imanishi M., Asami K., Futaki S., *Bioconjugate Chem.*, **2013**, 24, 188-195.
- [132] Radford R. J., Nguyen P. C., Tezcan F. A., *Inorg. Chem.*, **2010**, 49, 7106-7115.
- [133] Brodin J. D., Ambroggio X. I., Tang C., Parent K. N., Baker T. S., Tezcan F. A., *Nature Chem.*, **2012**, 4, 375-382.
- [134] Radford R. J., Tezcan F. A., *J. Am. Chem. Soc.*, **2009**, 131, 9136-9137.
- [135] Salgado E. N., Radford R. J., Tezcan F. A., *Acc. Chem. Res.*, **2010**, 43, 661-672.
- [136] Przybyla D. E., Chmielewski J., *J. Am. Chem. Soc.*, **2008**, 130, 12610-12611.
- [137] Przybyla D. E., Rubert Pérez C. M., Gleaton J., Nandwana V., Chmielewski J., *J. Am. Chem. Soc.*, **2013**, 135, 3418-3422.
- [138] Watson J. D., Crick F. H. C., *Nature*, **1953**, 171, 737-738.
- [139] www.astrochem.org/sci/Nucleobases.php, accessed on 29/08/2013.

- [140] Ohlendorf D. H., Anderson W. F., Fisher R. G., Takeda Y., Matthews B. W., *Nature*, **1982**, 298, 718-723.
- [141] Sauer R. T., Yocum R. R., Doolittle R. F., Lewis M., Pabo C. O., *Nature*, **1982**, 298, 447-451.
- [142] Aravind L., Anantharaman V., Balaji S., Babu M. M., Iyer L. M., *FEMS Microbiol. Rev.*, **2005**, 29, 231-262.
- [143] Kissinger C. R., Liu B., Martin-Blanco E., Kornberg T. B., Pabo C. O., *Cell*, **1990**, 63, 579-590.
- [144] Noyes M. B., Christensen R. G., Wakabayashi A., Stormo G. D., Brodsky M. H., Wolfe S. A., *Cell*, **2008**, 133, 1277-1289.
- [145] Tucker-Kellogg L., Rould M. A., Chambers K. A., Ades S. E., Sauer R. T., Pabo C. O., *Structure*, **1997**, 5, 1047-1054.
- [146] Weiss M. A., Ellenberger T., Wobbe C. R., Lee J. P., Harrison S. C., Struhl K., *Nature*, **1990**, 347, 575-578.
- [147] Patel L., Abate C., Curran T., *Nature*, **1990**, 347, 572-575.
- [148] O’Neil K. T., Hoess R. H., DeGrado W. F., *Science*, **1990**, 249, 774-778.
- [149] Anthony-Cahill S. J., Benfield P. A., Fairman R., Wasserman Z. R., Brenner S. L., Stafford III W. F., Altenbach C., Hubbell W. L., DeGrado W. F., *Science*, **1992**, 255, 979-983.
- [150] Johnson N. P., Lindstrom J., Baase W. A., von Hippel P. H., *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 4840-4844.
- [151] Murre C., McCaw P. S., Baltimore D., *Cell*, **1989**, 56, 777-783.
- [152] Ferré-D’Amaré A. R., Prendergast G. C., Ziff E. B., Burley S. K., *Nature*, **1993**, 363, 38-45.
- [153] Halazonetis T. D., Georgopoulos K., Greenberg M. E., Leder P., *Cell*, **1988**, 55, 917-924.

- [154] Hess J., Angel P., Schorpp-Kistner M., *J. Cell Sci.*, **2004**, 117, 5965-5973.
- [155] Ellenberger T. E., Brandl C. J., Struhl K., Harrison S. C., *Cell*, **1992**, 71, 1223-1237.
- [156] Ma P. C. M., Rould M. A., Weintraub H., Pabo C. O., *Cell*, **1994**, 77, 451-459.
- [157] Miller J., McLachlan A. D., Klug A., *EMBO J.*, **1985**, 4, 1609-1614.
- [158] Pavletich N. P., Pabo C. O., *Science*, **1991**, 252, 809-817.
- [159] Klug A., *Annu. Rev. Biochem.*, **2010**, 79, 213-231.
- [160] Laity J. H., Lee B. M., Wright P. E., *Curr. Opin. Struct. Biol.*, **2001**, 11, 39-46.
- [161] Tottey S., Harvie D. R., Robinson N. J., *Acc. Chem. Res.*, **2005**, 38, 775-783.
- [162] Ma Z., Jacobsen F. E., Giedroc D. P., *Chem. Rev.*, **2009**, 109, 4644-4681.
- [163] Agre P., Johnson P. F., McKnight S. L., *Science*, **1989**, 246, 922-926.
- [164] Talanian R. V., McKnight C. J., Kim P. S., *Science*, **1990**, 249, 769-771.
- [165] Talanian R. V., McKnight J., Rutkowski R., Kim P. S., *Biochemistry*, **1992**, 31, 6871-6875.
- [166] Pellegrini M., Ebright R. H., *J. Am. Chem. Soc.*, **1996**, 118, 5831-5835.
- [167] Cuenoud B., Schepartz A., *Science*, **1993**, 259, 510-513.
- [168] Cuenoud B., Schepartz A., *Proc. Natl. Acad. Sci. USA*, **1993**, 90, 1154-1159.
- [169] Palmer C. R., Sloan L. S., Adrian Jr. J. C., Cuenoud B., Paolella D. N., Schepartz A., *J. Am. Chem. Soc.*, **1995**, 117, 8899-8907.
- [170] Ueno M., Murakami A., Makino K., Morii T., *J. Am. Chem. Soc.*, **1993**, 115, 12575-12576.
- [171] Aizawa Y., Sugiura Y., Morii T., *Biochemistry*, **1999**, 38, 1626-1632.
- [172] Okagami M., Ueno M., Makino K., Shimomura M., Saito I., Morii T., Sugiura Y., *Bioorg. Med. Chem.*, **1995**, 3, 777-784.
- [173] Morii T., Shimomura M., Morimoto M., Saito I., *J. Am. Chem. Soc.*, **1993**, 115, 1150-1151.

- [174] Cammaño A. M., Vásquez M. E., Martínez-Costas J., Castedo L., Mascareñas J. L., *Angew. Chem. Int. Ed.*, **2000**, 39, 3104-3107.
- [175] Jones D. D., Barker P. D., *ChemBioChem*, **2004**, 5, 964-971.
- [176] Jones D. D., Barker P. D., *Angew. Chem. Int. Ed.*, **2005**, 44, 6337-6341.
- [177] Fan H. Y., Morgan S., Brechun K. E., Chen Y., Jaikaran A. S. I., Woolley G. A., *Biochemistry*, **2011**, 50, 1226-1237.
- [178] Song L., Caguiat J., Li Z., Shokes J., Scott R. A., Olliff L., Summers A. O., *J. Bacteriol.*, **2004**, 186, 1861-1868.
- [179] Wegner S. V., Boyaci H., Chen H., Jensen M. P., He C., *Angew. Chem. Int. Ed.*, **2009**, 48, 2339-2341.
- [180] Park C., Campbell J. L., Goddard III W. A., *Proc. Natl. Acad. Sci. USA*, **1992**, 89, 9094-9096.
- [181] Park C., Campbell J. L., Goddard III W. A., *Proc. Natl. Acad. Sci. USA*, **1993**, 90, 4892-4896.
- [182] Hollenbeck J. J., Gurnon D. G., Fazio G. C., Carlson J. J., Oakley M. G., *Biochemistry*, **2001**, 40, 13833-13839.
- [183] Park C., Campbell J. L., Goddard III W. A., *J. Am. Chem. Soc.*, **1995**, 117, 6287-6291.
- [184] Mosquera J., Jiménez-Balsa A., Dodero V. I., Vázquez M. E., Mascareñas J. L., *Nat. Commun.*, doi:10.1038/ncomms2825.
- [185] Zondlo N. J., Schepartz A., *J. Am. Chem. Soc.*, **1999**, 121, 6938-6939.
- [186] Chin J. W., Schepartz A., *J. Am. Chem. Soc.*, **2001**, 123, 2929-2930.
- [187] Montclare J. K., Schepartz A., *J. Am. Chem. Soc.*, **2003**, 125, 3416-3417.
- [188] Zhang M., Wu B., Zhao H., Taylor J. W., *J. Pept. Sci.*, **2002**, 8, 125-136.
- [189] Lajmi A. R., Lovrencic M. E., Wallace T. R., Thomlinson R. R., Shin J. A., *J. Am. Chem. Soc.*, **2000**, 122, 5638-5639.
- [190] Mazumber A., Maiti A., Roy K., Roy S., *ACS Chem. Biol.*, **2012**, 7, 1084-1094.

- [191] Kumita J. R., Flint D. G., Woolley G. A., Smart O. S., *Faraday Discuss.*, **2003**, 122, 89-103.
- [192] Woolley G. A., Jaikaran A. S. I., Berezovski M., Calarco J. P., Krylov S. N., Kumita J. R., *Biochemistry*, **2006**, 45, 6075-6084.
- [193] Zhang F., Timm K., Arndt K., Woolley G. A., *Angew. Chem. Int. Ed.*, **2010**, 49, 3943-3946.
- [194] Guerrero L., Smart O. S., Weston C. J., Burns D. C., Woolley G. A., Allemann R. K., *Angew. Chem. Int. Ed.*, **2005**, 44, 7778-7782.
- [195] Guerrero L., Smart O. S., Woolley G. A., Allemann R. K., *J. Am. Chem. Soc.*, **2005**, 127, 15624-15629.
- [196] Fujimoto K., Kajino M., Sakaguchi I., Inouye M., *Chem. Eur. J.*, **2012**, 18, 9834-9840.
- [197] Asuma Y., Imanishi M., Yoshimura T., Kawabata T., Futaki S., *Angew. Chem. Int. Ed.*, **2009**, 48, 6853-6856.
- [198] Stewart A. L., Park J. H., Waters M. L., *Biochemistry*, **2011**, 50, 2575-2584.
- [199] García Posse M. E., Juri M. A., Aymonino P. J., Piro O. E., Negri H. A., *Inorg. Chem.*, **1984**, 23, 948-952.
- [200] Burstall F. H., Nyholm R. S., *J. Chem. Soc.*, **1952**, 3570-3579.
- [201] König E., *Coord. Chem. Rev.*, **1968**, 3, 471-495.
- [202] Pfeiffer P., Quehl K., *Chem. Ber.*, **1932**, 65, 560-565.
- [203] Tada T., Kushi Y., Yoneda H., *Inorg. Chim. Acta*, **1982**, 64, L243-L245.
- [204] Yamagishi A., Soma M., *J. Am. Chem. Soc.*, **1981**, 103, 4640-4642.
- [205] König E., Madeja K., Watson K. J., *J. Am. Chem. Soc.*, **1968**, 90, 1146-1153.
- [206] Hudson A., Whitfield H. J., *Chem. Commun.*, **1966**, 606-607.
- [207] Liu T., Zhu J. Y., *Acta Cryst.*, **2007**, E63, m2506-m2507.
- [208] Gallois B., Real J., Hauw C., Zarembowitch J., *Inorg. Chem.*, **1990**, 29, 1152-1158.
- [209] Irving H., Mellor D. H., *J. Chem. Soc.*, **1962**, 5222-5237.

- [210] Basolo F., Hayes J. C., Neumann H. M., *J. Am. Chem. Soc.*, **1954**, 76, 3807-3809.
- [211] Anderegg G., *Helv. Chim. Acta*, **1963**, 46, 2397-2410.
- [212] Fábíán I., *Inorg. Chem.*, **1989**, 28, 3805-3807.
- [213] Cabani S., Moretti G., Scrocco E., *J. Chem. Soc.*, **1962**, 88-93.
- [214] Cabani S., Landucci M., *J. Chem. Soc.*, **1962**, 278-282.
- [215] Davies R. L., Dunning K. W., *J. Chem. Soc.*, **1965**, 4168-4185.
- [216] James B. R., Williams R. J. P., *J. Chem. Soc.*, **1961**, 2007-2019.
- [217] Szalda D. J., Creutz C., Mahajan D., Sutin N., *Inorg. Chem.*, **1983**, 22, 2372-2379.
- [218] Ruminski R. R., Petersen J. D., *Inorg. Chim. Acta*, **1984**, 88, 63-66.
- [219] DeSimone R. E., Drago R. S., *J. Am. Chem. Soc.*, **1970**, 92, 2343-2352.
- [220] Wheeler S. H., Zingheim S. C., Nathan L. C., *J. Inorg. Nucl. Chem.*, **1978**, 40, 779-783.
- [221] Li J., Huang L., Cai Y., Yu L., Zhou Z., *J. Mol. Struct.*, **2011**, 994, 70-74.
- [222] Lee R. H., Griswold E., Kleinberg J., *Inorg. Chem.*, **1964**, 3, 1278-1283.
- [223] Dhar S. K., Basolo F., *J. Inorg. Nucl. Chem.*, **1963**, 25, 37-44.
- [224] Kramer T., Strahle J., *Z. Naturforsch.*, **1986**, 41, 692-696.
- [225] De Munno G., Nicolò F., Julve M., *Acta Cryst.*, **1993**, C49, 1049-1052.
- [226] Chuy C., Falvello L. R., Libby E., Santa-María J. C., Tomás M., *Inorg. Chem.*, **1997**, 36, 2004-2009.
- [227] Davies R., Green M., Sykes A. G., *J. Chem. Soc., Dalton Trans.*, **1972**, 1171-1176.
- [228] Holyer R., Hubbard C., Kettle S., Wilkins R., *Inorg. Chem.*, **1965**, 4, 929-935.
- [229] Holyer R. H., Hubbard C. D., Kettle S. F. A., Wilkins R. G., *Inorg. Chem.*, **1966**, 5, 622-625.
- [230] Wada A., Sakabe N., Tanaka J., *Acta Cryst. B*, **1976**, B32, 1121-1127.
- [231] Constable E. C., *Adv. Inorg. Chem.*, **1989**, 34, 1-63.
- [232] Urtiaga M. K., Pizarro J. L., *Acta Cryst.*, **1994**, C50, 56-68.
- [233] Hauptmann R., *Acta Cryst.*, **1999**, C55, 1087-1090.

- [234] Fontaine F. G., *Acta Cryst.*, **2001**, E57, m270-m271.
- [235] Anderson O. P., *J. Chem. Soc. Dalton Trans.*, **1972**, 2597-2601.
- [236] Bardwell D. A., Jeffery J. C., Otter C. A., Ward M. D., *Polyhedron*, **1996**, 15, 191-194.
- [237] Hathaway B. J., *Coord. Chem. Rev.*, **1983**, 52, 87-169.
- [238] Foley J., Kennefick D., Phelan D., Tyagi S., Hathaway B., *J. Chem. Soc., Dalton Trans.*, **1983**, 2333-2338.
- [239] Foley J., Tyagi S., Hathaway B. J., *J. Chem. Soc., Dalton Trans.*, **1984**, 1-5.
- [240] McKenzie E. D., *Coord. Chem. Rev.*, **1971**, 6, 187-216.
- [241] Harrisson W. D., Hathaway B. J., Kennedy D., *Acta Cryst.*, **1979**, B35, 2301-2306.
- [242] Harrisson W. D., Hathaway B. J., *Acta Cryst.*, **1979**, B35, 2910-2913.
- [243] Harrisson D. W., Kennedy D. M., Power M., Sheahan R., Hathaway B. J., *J. Chem. Soc., Dalton Trans.*, **1981**, 1556-1564.
- [244] O'Sullivan C., Murphy G., Murphy B., Hathaway B., *J. Chem. Soc., Dalton Trans.*, **1999**, 1835-1844.
- [245] Fereday R. J., Hodgson P., Tyagi S., Hathaway B. J., *J. Chem. Soc. Dalton Trans.*, **1981**, 2070-2077.
- [246] Hernández-Molina M., González-Platas J., Ruiz-Pérez C., Lloret F., Julve M., *Inorg. Chim. Acta*, **1999**, 284, 258-265.
- [247] Perlepes S. P., Huffman J. C., Christou G., *Polyhedron*, **1991**, 10, 2301-2308.
- [248] Garland M. T., Grandjean D., *Acta Cryst.*, **1988**, C44, 1209-1212.
- [249] Barclay G. A., Kennard C. H. L., *Nature*, **1961**, 4801, 424-425.
- [250] Pflaum R. T., Brandt W. W., *J. Am. Chem. Soc.*, **1954**, 76, 6215-6219.
- [251] James B. R., Williams R. J. P., *J. Chem. Soc.*, **1961**, 2007-2019.
- [252] Garribba E., Micera G., Sanna D., Strinna-Erre L., *Inorg. Chim. Acta*, **2000**, 299, 253-261.

- [253] Sanna D., Buglyó P., Tomaz A. I., Pessoa J. C., Borović S., Micera G., Garribba E., *Dalton Trans.*, **2012**, 41, 12824-12838.
- [254] Marov I. N., Belyaeva V. K., Smirnova E. B., Dolmanova I. F., *Inorg. Chem.*, **1978**, 17, 1667-1669.
- [255] Barnett S. M., Goldberg K. I., Mayer J. M., *Nat. Chem.*, **2012**, 4, 498-502.
- [256] Noack M., *J. Chem. Phys.*, **1968**, 48, 2689-2699.
- [257] Sigel H., *Inorg. Chim. Acta*, **1972**, 6, 195-196.
- [258] Lim M. C., Sinn E., Martin R. B., *Inorg. Chem.*, **1976**, 15, 807-811.
- [259] Sigel H., Martin R. B., *Chem. Rev.*, **1982**, 82, 385-426.
- [260] Castiñeiras A., Balboa S., Carballo R., Niclós J., *Z. Anorg. Allg. Chem.*, **2002**, 628, 2353-2359.
- [261] Carballo R., Castiñeiras A., Balboa S., Covelo B., Niclós J., *Polyhedron*, **2002**, 21, 2811-2818.
- [262] Antolini L., Menabue L., *Inorg. Chem.*, **1984**, 23, 1418-1422.
- [263] Ghattas W., Giorgi M., Mekmouche Y., Tanaka T., Rockenbauer A., Réglier M., Hitomi Y., Simaan A. J., *Inorg Chem.*, **2008**, 47, 4627-4638.
- [264] Biswas C., Drew M. G. B., Estrader M., Ghosh A., *Dalton Trans.*, **2009**, 5015-5022.
- [265] Antolini L., Menabue L., Pellacani G. C., Saladini M., Sola M., *J. Chem. Soc. Dalton Trans.*, **1984**, 2319-2323.
- [266] Sigel H., *Angew. Chem. Int. Ed.*, **1968**, 7, 137-138.
- [267] Sigel H., *Inorg. Chem.*, **1975**, 14, 1535-1540.
- [268] Sigel H., Naumann C. F., Prijs B., McCormick B., Falk M. C., *Inorg. Chem.*, **1977**, 16, 790-796.
- [269] Ghosh S. N., *Inorg. Nucl. Chem. Lett.*, **1969**, 5, 841-843.
- [270] Eom G. H., Park H. M., Hyun M. Y., Jang S. P., Kim C., Lee J. H., Lee S. J., Kim S., Kim Y., *Polyhedron*, **2011**, 30, 1555-1564.

- [271] Walsh A., Walsh B., Murphy B., Hathaway B. J., *Acta Cryst.*, **1981**, B37, 1512-1520.
- [272] Chen X., Xu Z., *Polyhedron*, **1994**, 13, 3329-3332.
- [273] Rodrigues B. L., *Acta Cryst.*, **2004**, E60, m1169-m1171.
- [274] Ishiguro S., Nagy L., Ohtaki H., *Bull. Chem. Soc. Jpn.*, **1987**, 60, 2865-2869.
- [275] Farkas E., Enyedy E. A., Micera G., Garribba E., *Polyhedron*, **2000**, 19, 1727-1736.
- [276] Baker A. T., Goodwin H. A., *Aust. J. Chem.*, **1985**, 38, 207-214.
- [277] Machan C. W., Adelhardt M., Sarjeant A. A., Stern C. L., Sutter J., Meyer K., Mirkin C. A., *J. Am. Chem. Soc.*, **2012**, 134, 16921-16924.
- [278] Reiff W. M., Erickson N. E., Baker Jr W. A., *Inorg. Chem.*, **1969**, 8, 2019-2021.
- [279] Judge J. S., Reiff W. M., Intille G. M., Ballway P., Baker Jr W. A., *J. Inorg. Nucl. Chem.*, **1967**, 29, 1711-1716.
- [280] Hathcock D. J., Stone K., Madden J., Slattery S. J., *Inorg. Chim. Acta*, **1998**, 282, 131-135.
- [281] Farina R., Hogg R., Wilkins R. G., *Inorg. Chem.*, **1968**, 7, 170-172.
- [282] Bullock J. I., Simpson P. W. G., *J. Chem. Soc., Faraday Trans.*, **1981**, 77, 1991-1997.
- [283] Bazzicalupi C., Bencini A., Bianchi A., Danesi A., Faggi E., Giorgi C., Santarelli S., Valtancoli B., *Coord. Chem. Rev.*, **2008**, 252, 1052-1068.
- [284] Hamilton J. M., Anhorn M. J., Oscarson K. A., Reibenspies J. H., Hancock R. D., *Inorg. Chem.*, **2011**, 50, 2764-2770.
- [285] Hogg R., Wilkins R. G., *J. Chem. Soc.*, **1962**, 341-350.
- [286] Maslen E. N., Raston C. L., White A. H. J., *J. Chem. Soc., Dalton Trans.*, **1974**, 1803-1807.
- [287] Henke W., Kremer S., *Inorg. Chim. Acta*, **1982**, 65, L115-L117.
- [288] Goldschmied E., Stephenson N. C., *Acta Cryst.*, **1970**, B26, 1867-1875.
- [289] Kepert D. L., Kucharski E. S., White A. H., *J. Chem. Soc., Dalton Trans.*, **1980**, 1932-1938.

- [290] Prasad J., Peterson N. C., *Inorg. Chem.*, **1969**, 8, 1622-1625.
- [291] Constable E. C., Harris K., Housecroft C. E., Neuburger M., Zampese J. A., *Dalton Trans.*, **2011**, 40, 11441-11450.
- [292] Arriortua M. I., Rojo T., Amigo J. M., Germain G., Declercq J. P., *Bull. Soc. Chim. Belg.*, **1982**, 91, 337-338.
- [293] Baker A. T., Craig D. C., Rae A. D., *Aust. J. Chem.*, **1995**, 48, 1373-1378.
- [294] Calatayud M. L., Sletten J., Julve M., Castro I., *J. Mol. Struct.*, **2005**, 741, 121-128.
- [295] Jones G. D., Martin J. L., McFarland C., Allen O. R., Hall R. E., Haley A. D., Brandon R. J., Konovalova T., Desrochers P. J., Pulay P., Vicic D. A., *J. Am. Chem. Soc.*, **2006**, 128, 13175-13183.
- [296] Allman R., Henke W., Reinen D., *Inorg. Chem.*, **1978**, 17, 378-382.
- [297] Henke W., Kremer S., Reinen D., *Inorg. Chem.*, **1983**, 22, 2858-2863.
- [298] Morgan G., Burstall F. H., *J. Chem. Soc.*, **1937**, 1649-1655.
- [299] Sanna D., Buglyó P., Tomaz A. I., Pessoa J. C., Borovic S., Micera G., Garribba E., *Dalton Trans.*, **2012**, 41, 12824-12838.
- [300] Constable E. C., *Adv. Inorg. Chem.*, **1987**, 30, 69-121.
- [301] Zhou W., Wang X., Hu M., Guo Z., *J. Inorg. Biol. Chem.*, **2013**, 121, 114-120.
- [302] Constable E. C., Housecroft C. E., Murray N. S., Zampese J. A., *Polyhedron*, **2013**, 54, 110-118.
- [303] Corbridge D. E. C., Cox E. G., *J. Chem. Soc.*, **1956**, 594-603.
- [304] Vlasse M., Rojo T., Beltran-Porter D., *Acta Cryst.*, **1983**, C39, 560-563.
- [305] Cagle Jr F. W., Smith G. F., *J. Am. Chem. Soc.*, **1947**, 69, 1860-1862.
- [306] Brandt W. W., Dwyer F. P., Gyarfás E. C., *Chem. Rev.*, **1954**, 54, 959-1017.
- [307] Brandt W. W., Smith G. F., *Anal. Chem.*, **1949**, 21, 1313-1319.
- [308] Irving H., Mellor D. H., *J. Chem. Soc.*, **1962**, 5237-5245.

- [309] Nellas R. B., Fronczek F. R., Newkome G. R., Watkins S. F., Xia Y., *Acta Cryst.*, **2006**, E62, m1522-m1523.
- [310] Higgins S. J., *Coord. Chem. Rev.*, **1996**, 152, 175-249.
- [311] Ward M. D., Couchman S. M., Jeffery J. C., *Acta Cryst.*, **1998**, C54, 1820-1823.
- [312] Newkome G. R., Pantaleo D. C., Puckett W. E., Ziefle P. L., Deutsch W. A., *J. Inorg. Nucl. Chem.*, **1981**, 43, 1529-1531.
- [313] van Albada G. A., Mohamadou A., Mutikainen I., Turpeinen U., Reedijk J., *Eur. J. Inorg. Chem.*, **2004**, 3733-3742.
- [314] Irving H., Williams R. J. P., *J. Chem. Soc.*, **1953**, 3192-3210

CHAPTER II: MODEL PEPTIDE SWITCHES AS ALLOSTERIC REGULATORS

2.1 – Introduction

2.1.1 – Artificial regulation sites based on polypyridine

Bipy is one of the most widely used metal-chelating units.[1] The lone pair on the nitrogen atoms are basic and allow for metal ion coordination. Moreover, binding of a single metal to both nitrogen atoms results in the formation of a 5-membered chelate ring, with high-stability. Therefore, most divalent metal ions bind **bipy** through both nitrogen atoms, thus stabilising a *cis*- arrangement of the two pyridine rings, with respect to the axial bond. In the absence of a metal ion, the inter-ring conjugation of π orbitals is believed to be less important and non-planar conformations are allowed.[2,3] Repulsion between hydrogens in position 3- strongly destabilise the *cis*- arrangement, however, the *trans*- arrangement represents the energetic minima in the absence of metal ion coordination.[4] Experimental evidence such as dipolar momentum,[2] and X-ray analysis[5] are consistent with **bipy** adopting the *trans*- conformation in aprotic solvents and in the solid-state. As a result, metal-complexation of **bipy** results in a conformational transition involving the rotation of one of the pyridine units by up to 180°, with respect to the second pyridine.

Rebek et al. first took advantage of this metal-induced conformational transition as a regulation site, by incorporating the **bipy** unit into a crown ether.[6] In the resulting **bipy**-crown ether, metal ion binding at the **bipy** regulation site impacts the geometry and the size of the polyether ring, thus altering its affinity for alkali metal ions (see Figure 2.1). This concept was subsequently applied to regulate the binding of a large variety of substrates, including metal ions, nucleotides, amino-acids, and fullerenes.[7] **Bipy** conformational transitions have also been exploited in the construction of molecular machines.[8]

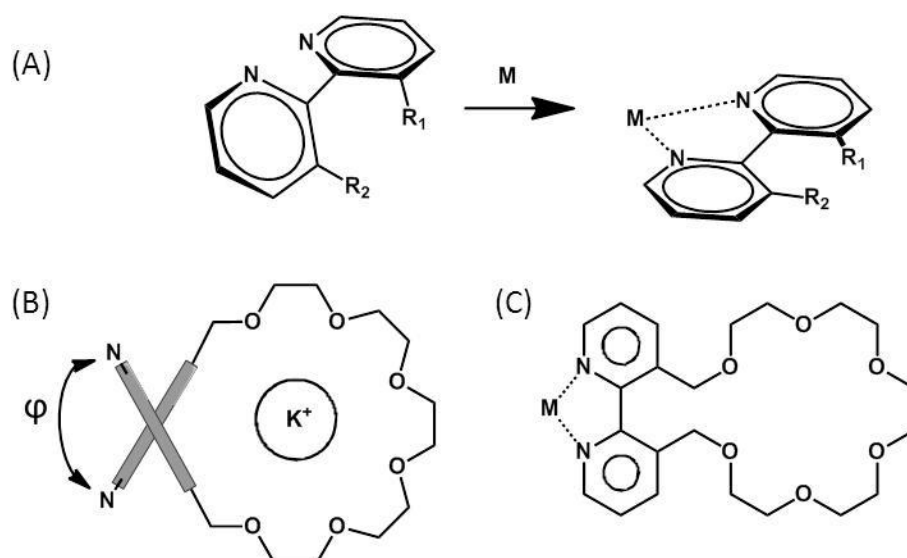


Figure 2.1 – Scheme illustrating the regulation of alkali metal ion binding through allosteric mechanism developed by Rebek. Transition metal coordination at the **bipy** (allosteric site) regulates the size of the ether crown (active site), and therefore its affinity for alkali metal ions. (A) the metal ion regulation process, (B) a side view of the device (**bipy** is metal-unbound), and (C) 2D projection of the device (metal-bound to **bipy**); adapted from reference 6.

Even though **terpy** was first isolated in the 1930's,[9] it only became increasingly used in the 1980's. In a similar fashion to **bipy**, it preferentially adopts a *trans*- arrangement of the pyridine rings in the absence of metal ions, and is able to form very stable metal-chelate complexes. **Terpy** is able to coordinate to metal ions with different coordination modes, involving two or three pyridines, and resulting in either a *cis-cis*- or a *cis-trans*- conformation.[10, 11] Based on this, molecular tweezers (acyclic molecules bearing a cavity where a host molecule can fit) containing **terpy** units, were designed by Lehn and co-workers. These exploit the conformational transition on metal ion coordination to regulate binding of an aromatic substrate, such as 7,7',8,8'-tetracyanoquinodimethane (TCNQ) (see Figure 2.2).[12]

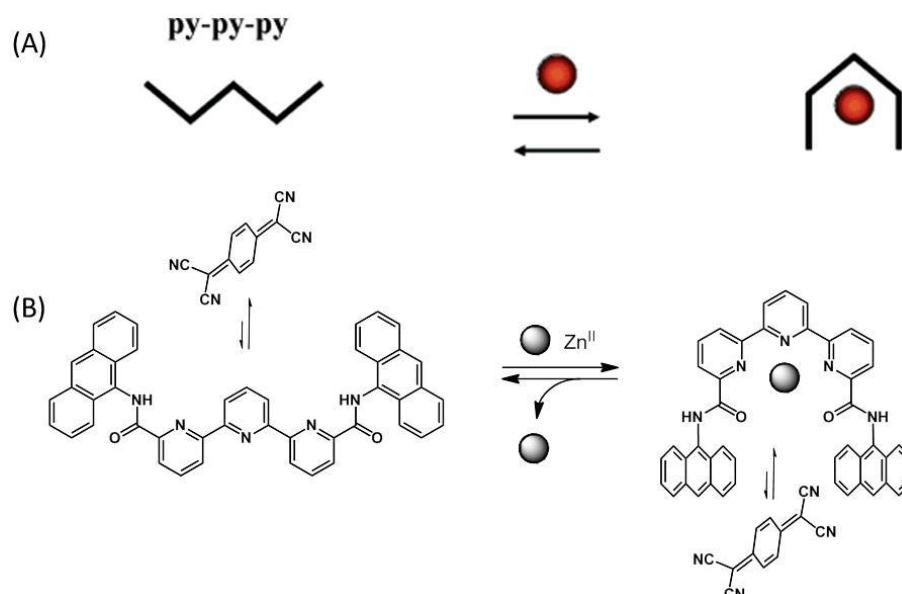


Figure 2.2 – Scheme illustrating Lehn's tweezer. Zinc binding at the **terpy** (allosteric site) allows for pre-organisation of anthracene moieties for π stacking interaction with an aromatic guest molecule such as TCNQ (active site). (A) Schematic representation and (B) ChemDraw scheme of the terpyridine conformational change, and repositioning of anthracene units, on metal ion coordination; adapted with permission from reference 12. Copyright 2004 American Chemical Society.

The polypyridine regulation ability is linked to the rotation of pyridines around their axial bond, and is often described by the value of the dihedral angle N-C-C-N or ϕ (see Figure 2.1B). In order for metal ion coordination to result in a significant structural change, essential for a regulation trigger, the two constituents which make up the active site need to be attached to the polypyridine through different pyridine rings. There are various locations at which these constituents can be attached to the pyridine ring. Successful regulation processes involving repositioning of the active sites substituents at position 6,6'-,[13] 4,4'-,[14] and 3,3'-,[6,15] of **bipy**, have been reported. In contrast, 5,5'- positioning constitutes an exception, because the metal ion induced *cis-trans* conformational transition does not result in modification of the relative position of the two active site constituents, however the relative orientation of the substituents is altered by as much as the dihedral angle ϕ . In contrast, the 5,5'-disubstituted **bipy** unit has been exploited as a regulation site, by taking advantage of the different π conjugation between the metal bound- and free- species. This allows for electron

transfer between two active sites, thus altering the global electronic transition within the supramolecular building blocks (see Figure 2.3).[16,17]

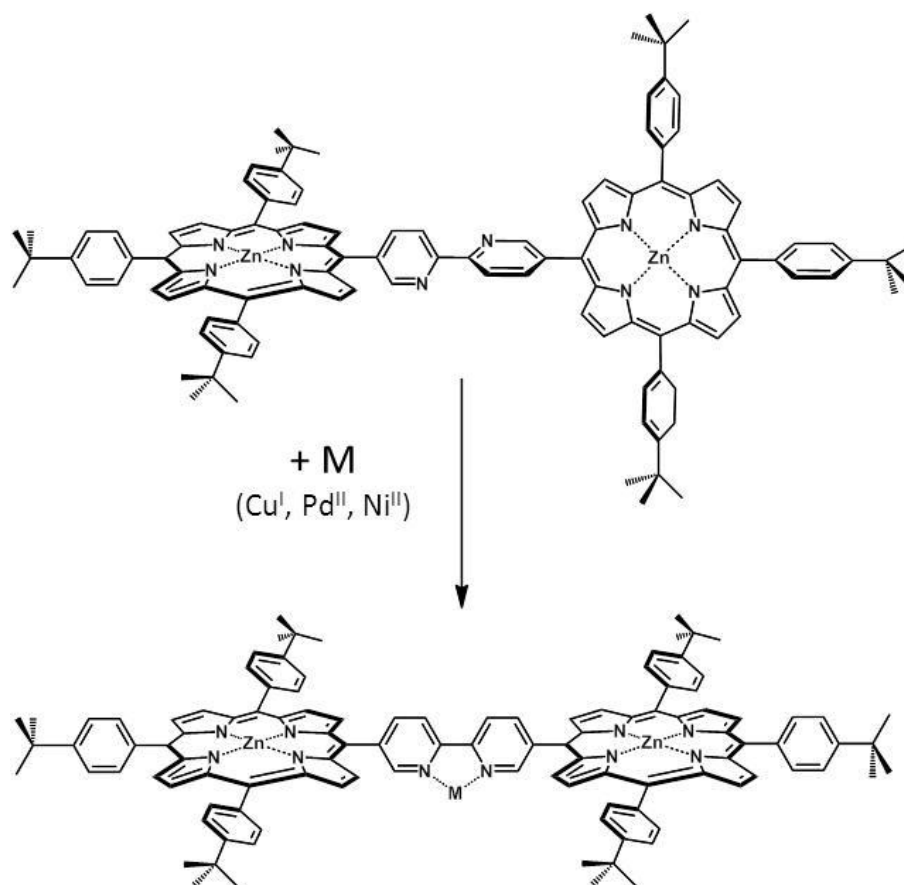


Figure 2.3 – Scheme depicting the supramolecular building block developed by Drain *et al.*: Metal binding at **bipy** results in the alignment of the porphyrin ring in the same plane, thus facilitating π molecular orbital overlap and electronic transfer; adapted from reference 16.

Alternatively, building blocks attached to **bipy** or **terpy** can be self-assembled upon metal coordination and formation of 2:1 polypyridine metal complexes. This property was exploited for the preparation of material displaying metal-dependent sol-gel transitions.[18] or for crosslinking of nanoparticles.[19]

The conformational transition resulting from **bipy** complexation has also been exploited to regulate the secondary structure of biomolecules. Sugimoto and co-workers inserted a 4,4'-disubstituted **bipy** within the DNA backbone of an oligonucleotide containing guanine tetrameric repeats. In the absence of divalent metal ion, the artificial construct adopted an antiparallel G-quadruplex in aqueous solution at pH 6. However, addition of divalent cations

such as Ni(II), Co(II), or Zn(II), results in a transition to a parallel G-quadruplex, and in formation of G-wires of various lengths.[14].

Kelly and co-workers prepared several short peptides with a **bipy** unit introduced directly into the peptide backbone.[20] The peptide sequence and the pH were found to influence the conformational state of the **bipy**, as a result the free ligand was only able to adopt the *trans*- conformation under alkaline conditions. At pH 9.5 they demonstrated that Cu(II) coordination led to a structural reorganisation of the **bipy** linker to the *cis*-conformation, resulting in a secondary structure transition from random coil to β -sheet (see Figure 2.4). However, the **bipy** linker could not be used as a switch at a physiologically relevant pH. This highlights the importance for a better understanding of the factors (c.a. protonation of polypyridine, substituent nature and positioning) which govern the conformational state of **bipy** and related polypyridine units, if they are to be exploited as potential artificial allosteric regulation sites.

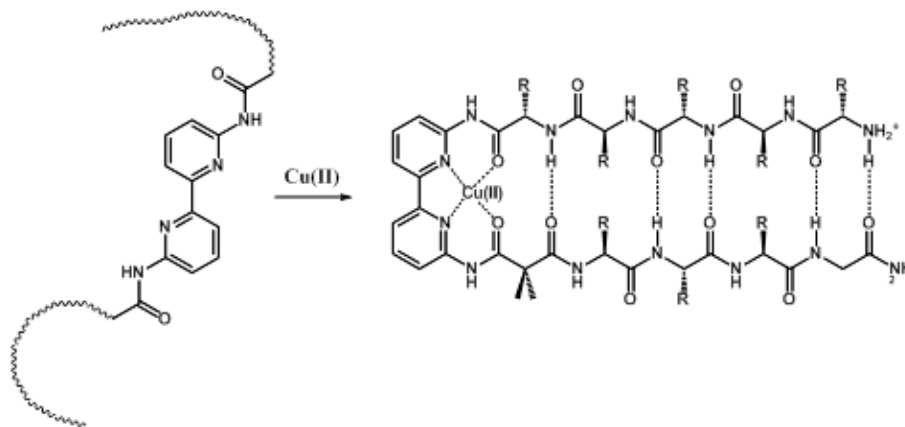


Figure 2.4 – Scheme showing the peptide with incorporated **bipy** prepared by Kelly et al. Cu(II) addition results in a *trans*-to-*cis* conformational transition of the **bipy** linker, thus inducing a random coil to β -sheet transition; adapted from reference 20.

As for inorganic materials, the ability of **bipy** and **terpy** to form assemblies following complexation and formation of 2:1 or 3:1 polypyridine metal ion complexes, were exploited for the formation of a large variety of peptide,[21,22,23] or DNA scaffolds.[24,25,26] Among the dication of first-row metal transitions (second half only), Fe(II) has been shown to form

the most stable **3bipy**:1M(II) and **2terpy**:1M(II) complexes in aqueous solution (though Ni(II) has the highest overall stability, it was shown to disproportionate in solution). In contrast, Zn(II) and Cu(II) represent the most suitable candidates for the predominant formation of 1:1 complexes with equimolar amounts of **bipy** and **terpy** at physiological pH (see Chapter 1).

Several strategies allow one to efficiently incorporate non-natural molecules within proteins (for reviews see [27-30]). An attractive approach for incorporation of non-natural molecules such as **bipy** within synthetic peptides is the preparation of synthetic amino-acids, with suitable protection schemes for incorporation within solid-phase peptide synthesis.[20,31] Alternatively, **bipy**, **terpy** or **phen** derivatives were often incorporated into peptides or proteins through coupling to amino-acid side chains. [32-34] Cysteine is a particularly attractive residue for coupling with, because (1) its nucleophilic thiol side-chain is unique among natural amino-acids, (2) cysteine has a low natural abundance, making it a suitable candidate for selective mutation,[35] and (3) cysteine can easily be introduced into peptides by site-directed mutagenesis. [27] Similarly, there exists a plethora of methods for the chemical attachment to cysteine.[35]

2.1.2 – Aims of the chapter

The aim is to prepare peptide polypyridine conjugates, for which metal binding at the polypyridine (allosteric site) would result in the spatial realignment of peptide fragments, which could be used to regulate a biomolecular recognition process. For this, an investigation of the polypyridine metal-coordination ability, as well as the conformational transition resulting from metal-binding, was envisioned. One can reason that small peptide dimer conjugates would allow to more readily study and interpret the chemistry at the metal polypyridine site. This chapter therefore reports the design, preparation, and characterisation of model compounds, where cysteine and glutathione (GSH) are both dimerised through

Me₂pyr, **Me₂bipy**, and **Me₂terpy** linkers (see Figure 2.5). Cu(II) and Zn(II) binding studies of the resulting conjugates were subsequently monitored by UV-visible, CD, or NMR spectroscopy.

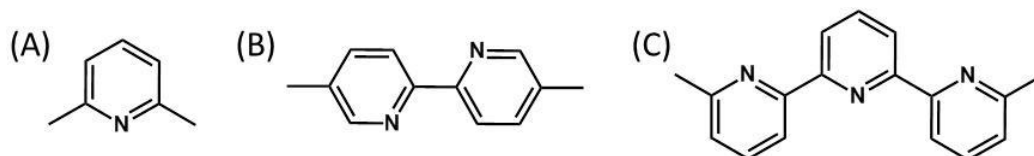


Figure 2.5 – Scheme representing linkers relative to this work: (A) 2,6-dimethylpyridine (**Me₂pyr**), (B) 5,5'-dimethyl-2,2'-bipyridine (**Me₂bipy**), and (C) 6,6''-dimethyl-2,2':6',2''-terpyridine (**Me₂terpy**).

Cysteine and GSH were first dimerised through a commercially available pyridine linker, affording highly water-soluble conjugates with low Cu(II) and Zn(II) affinity. Polypyridine linkers, bearing **bipy** or **terpy** chelation motif were then prepared, and their interactions with Zn(II) in the organic solvent DMSO-d₆ were investigated by NMR. These linkers were further used to dimerise cysteine and GSH, affording conjugates with contrasted solubility under physiological conditions. Cu(II) and Zn(II) interactions with GSH conjugates were studied in more details.

2.2 – Results and discussion

2.2.1 – Cysteine and glutathione dimerised through a pyridine linker

2.2.1.1 – Design, preparation and stability

Initial studies involved testing the proposed amino-acid/peptide dimerisation through a single pyridinyl linker with limited metal ion coordination ability, then studying the stability of the resulting conjugates in solution, and finally their interactions with metal ions. For the coupling, alkylation of cysteine was used, as it involves simple and well described preparation steps.[36,37] At pH 8, the thiol group of cysteine can act as a nucleophile (pK_a = 8.35), and react with electrophilic alkyl- derivatives, to form thioethers. These conditions are also

favourable to thiol oxidation, which is the kinetic determinant step for disulphide formation. Therefore, formation of peptide dimers, as side-products, is expected.

Coupling of two equivalents of cysteine with commercially available water soluble 2,6-bis(bromomethyl)pyridine (**pyr-Br₂**), was first attempted following the procedure reported by Prakash *et al.* to dimerise cysteine-glycine dipeptides through a 4,4'-dimethylazobenzene linker.[38] Phosphate buffer, used to regulate the pH of the reaction mixture, was changed for Tris.HCl, thus facilitating monitoring by ES-TOF and MALDI-TOF spectrometry.[39] The amount of dimethylformamide (DMF) was also reduced (2% instead of 50%) in order to facilitate both purification by analytical HPLC, and on-plate crystallisation prior to MALDI-TOF characterisation. This amount was previously reported to be low enough to ensure homogeneity of the medium.[40]

Cysteine was reacted with **pyr-Br₂**, affording the desired **cys₂pyr** conjugate (Figure 2.6), which was purified by preparative HPLC, using a water/acetonitrile gradient for elution (without additional acid). After lyophilisation of isolated fractions, the resulting white solid was characterised by ¹H and ¹³C NMR, ES-TOF, MALDI-TOF, UV-visible spectroscopy, elemental analysis, and the purity confirmed by analytical HPLC.

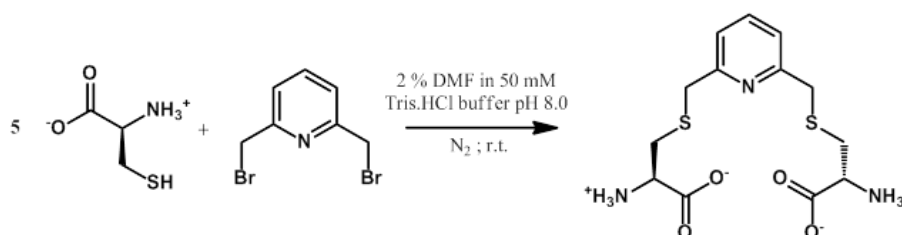


Figure 2.6 – Synthetic scheme for preparation of **cys₂pyr**.

The ¹H NMR spectrum recorded of a solution of **cys₂pyr** in D₂O (pD~6) displays 5 signals attributed to **cys₂pyr** (see Figure 2.7A). Following acquisition, the sample was stored at room temperature for 15 days, and a new ¹H NMR spectrum recorded (see Figure 2.7B). Similarities with the previous spectrum suggest that **cys₂pyr** is stable under these conditions.

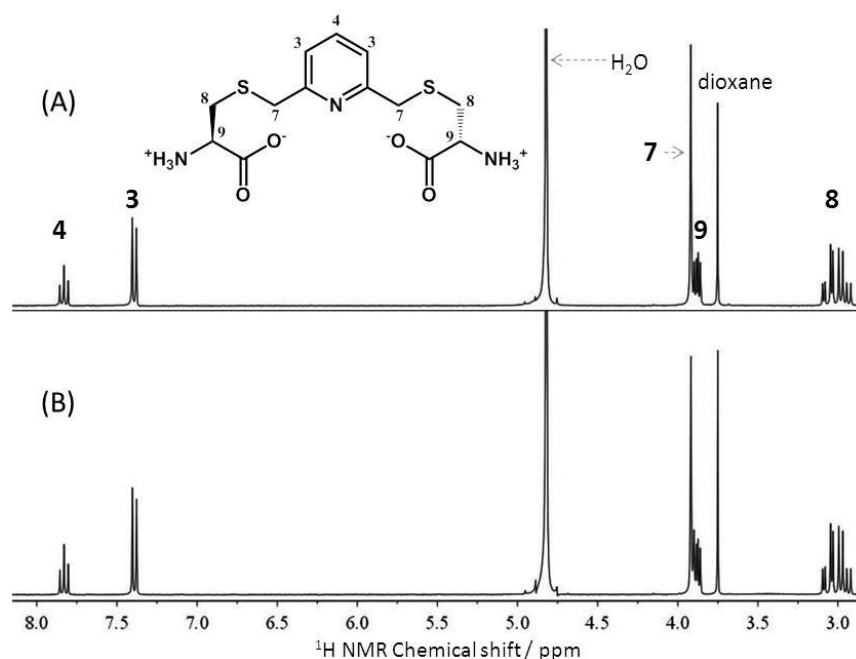


Figure 2.7 – ^1H NMR spectra of **cys₂pyr** in D_2O at pH ~ 6 recorded at 300 MHz, directly (A) and 15 days (B) after preparation, dioxane was added as an internal reference.

In order to test our dimerisation procedure with peptides, **pyr-Br₂** was reacted with the tripeptide GSH following a similar procedure, affording **pyr-GS₂**. During the reaction, the concentration of reduced GSH was monitored by the Ellman's assay (see appendix). After 14 hours, the pH had reduced significantly, and NaOH was added to restart the reaction. The plot of absorbance at 412 nm corresponding to Ellman's test recorded at different reaction times, indicates a decrease in the reduced GSH concentration, which plateaus at ca 30 hours (see Figure 2.8).

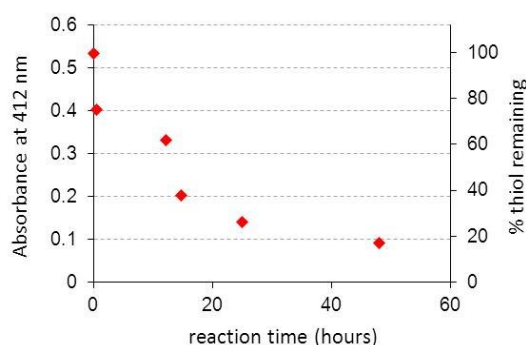


Figure 2.8 – Graph monitoring the available thiol concentration during **pyr-GS₂** synthesis. This was estimated by Ellman's test at different reaction times (0 hours corresponds to GSH addition). The absorbance at 412 nm relative to Ellman's test was plotted as the primary y-axis (left), and can be normalised to the percentage remaining thiol (secondary y-axis, right). The pH was adjusted with NaOH after 14 hours, after which it remained constant (ca 7.5).

Pyr-GS₂ was also purified by preparative HPLC. In contrast to **cys₂pyr**, no efficient separation could be obtained using water/acetonitrile mixtures as eluents, and required addition of 0.05% trifluoroacetic acid (TFA) to the water/acetonitrile eluents. The resulting colorless gel was successfully characterised by ¹H and ¹³C NMR, ES-TOF, MALDI-TOF, UV spectroscopy, and analytical HPLC. However, elemental analysis did not indicate the expected composition, suggesting that the compound was recovered as a mixture of salts. The best match of analytical data was found for **pyr-GS₂**·3TFA·3.5H₂O (see section 2.4.2), which is consistent with the HPLC conditions employed.

¹H NMR spectra recorded of a **pyr-GS₂** solution in D₂O (pD ~ 1 due to TFA salts present) displays 9 signals attributed to **pyr-GS₂** (see Figure 2.9A). The sample containing **pyr-GS₂** solution was stored at room temperature for 14 days, and a new ¹H NMR spectrum recorded. Again, the second acquisition was found to be very similar to the original (see Figure 2.9B).

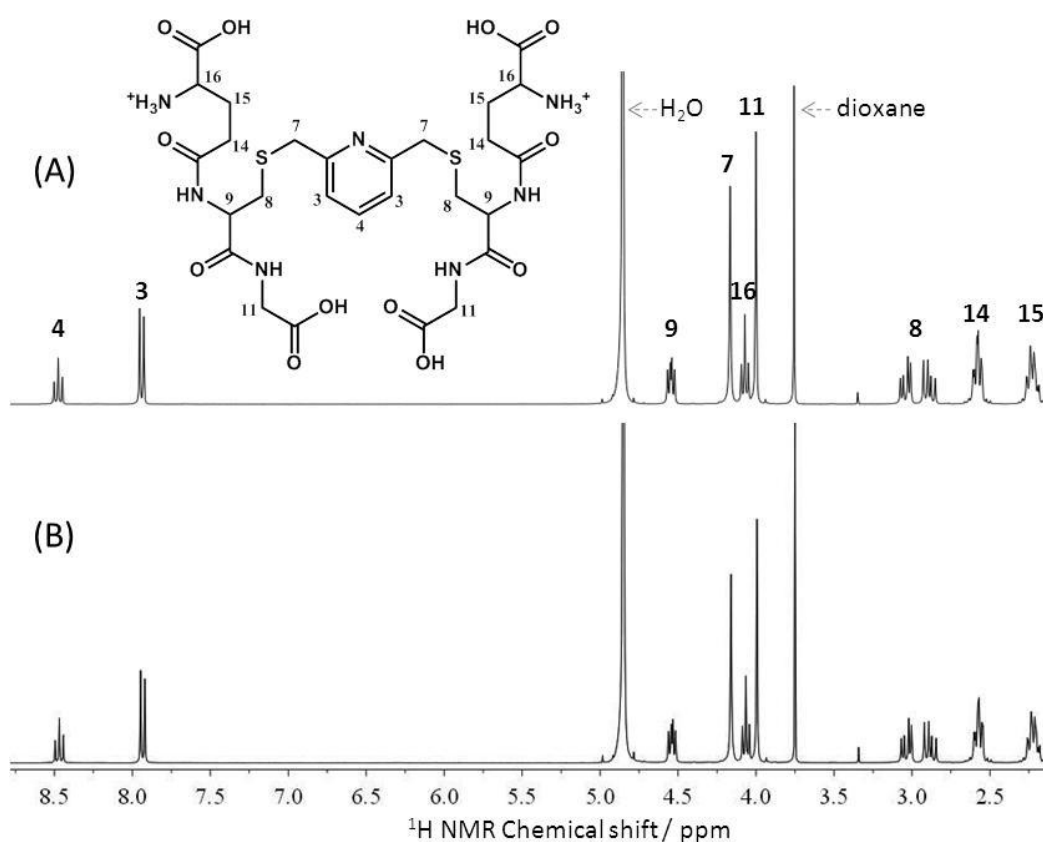


Figure 2.9 – ¹H NMR of **pyr-GS₂** recorded in D₂O at 300 MHz, directly (A) and 14 days after preparation (B), dioxane was added as a reference.

The stability of **pyr-GS₂** in aqueous solution was also monitored by analytical HPLC, and the same single absorption peak was detected after samples were stored in solution for 31 days (see Figure 2.10).

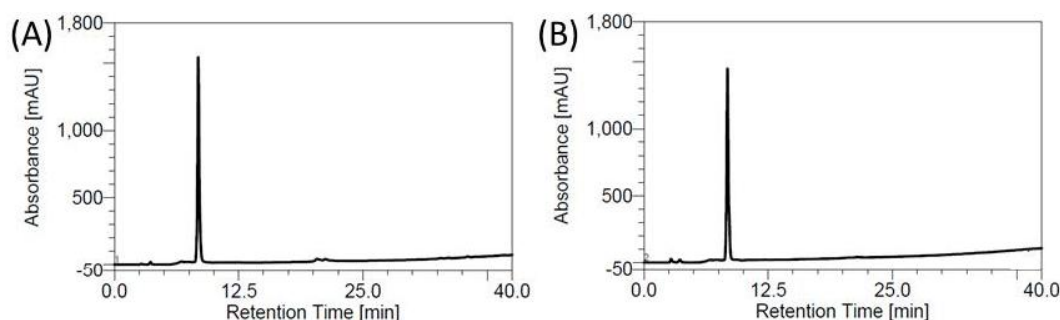


Figure 2.10 – Analytical C18-HPLC of **pyr-GS₂** aqueous solution recorded directly (A) and 31 days (B) after preparation. Solutions were eluted (monitoring at 220 nm) using a linear gradient from 0 to 100 % acetonitrile in water (+0.05 % TFA) over 40 min (flow 1 mL/min).

2.2.1.2 – UV studies of **pyr-Br₂**, **cys₂pyr** and **pyr-GS₂**

UV spectra of **pyr-Br₂** solutions buffered at pH 8.0 display one band centred at 275 nm attributed to the $\pi \rightarrow \pi^*$ transition of the pyridine ring (see Figure 2.11A). Another peak can be observed at higher energies but the maximum is in the far-UV domain. UV spectra of **cys₂pyr** solutions buffered at pH 8.0, display a similar profile, with the maximum of the $\pi \rightarrow \pi^*$ transition centred at 272 nm (see Figure 2.11B).

Extinction coefficients relative to **pyr-Br₂** and **cys₂pyr** maxima were determined based on triplicate absorbance measurements (independent solutions), in 20 mM phosphate buffer pH 8. Absorbance measurements relative to different stock solutions were averaged, and plotted versus concentration (either **pyr-Br₂** or **cys₂pyr**) (see Figure 2.11A and B). Linear regression was performed, affording values of extinction coefficients for the absorbance maxima of $\epsilon_{275 \text{ nm}} 5,190 \pm 390 \text{ M}^{-1} \text{ cm}^{-1}$ (**pyr-Br₂**) and $\epsilon_{272 \text{ nm}} 5,530 \pm 170 \text{ M}^{-1} \text{ cm}^{-1}$ (**cys₂pyr**) (see Figure 2.11C and D). Within error, the absorbance maxima of **cys₂pyr** and **pyr-Br₂** have similar extinction coefficients.

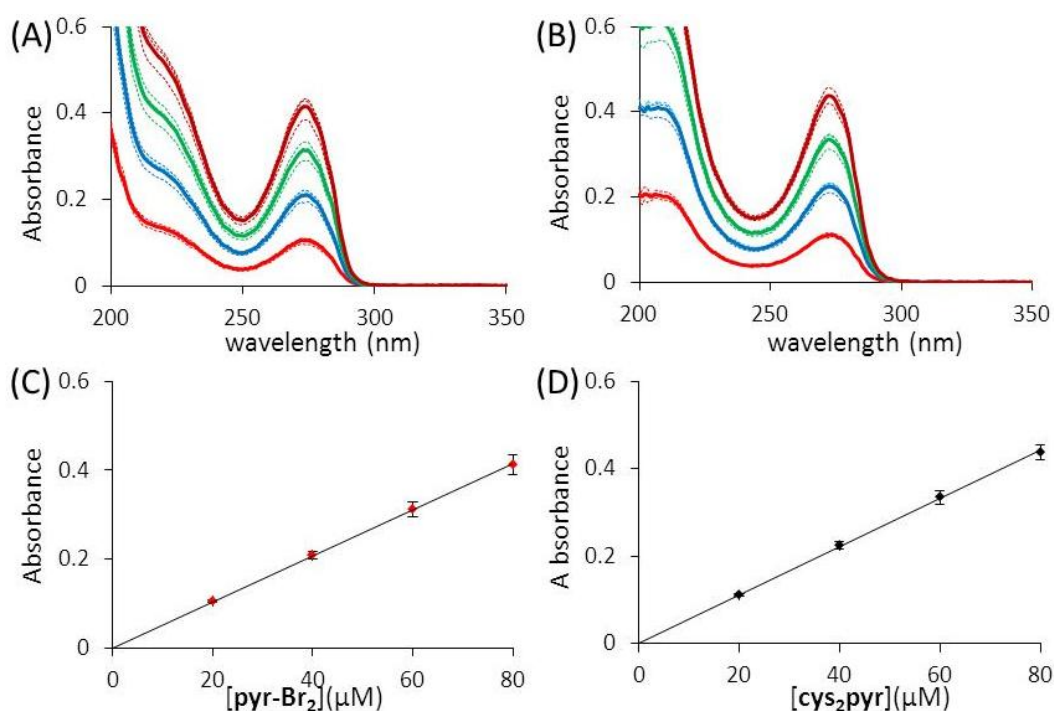


Figure 2.11 – Determination of extinction coefficient for **cys₂pyr** and **pyr-Br₂**. Top: UV spectra for **pyr-Br₂** (A) and **cys₂pyr** (B) independent solutions (---) and average spectra (—) recorded in phosphate buffer pH 8 at 20, 40, 60 and 80 μM. Bottom: plot of average absorbance at 275 (C) or 272 nm (D) as a function of the concentration of **pyr-Br₂** (C) and **cys₂pyr** (D), error bars correspond to the standard deviation.

UV spectra of **pyr-GS₂** solutions buffered at pH 8.0, display a similar profile, with the maximum of the $\pi \rightarrow \pi^*$ transition centred at 274 nm, however, the extinction coefficient seems significantly lower, ca $\epsilon_{274 \text{ nm}} 3,190 \pm 70 \text{ M}^{-1} \text{ cm}^{-1}$ (average of two independent experiments). Despite analytical HPLC showing a single peak, the elemental analysis was not consistent with pure **pyr-GS₂**, rather a hydrated TFA salt. Therefore, one can conclude that mass is not a reliable way to estimate the amount of **pyr-GS₂** in solution, and by extension, of other peptide conjugates purified using TFA containing eluents.

2.2.1.3 – Cu(II) titration of cys₂pyr and pyr-GS₂ monitored by UV spectroscopy

Cys₂pyr and **pyr-GS₂** might possess similar metal ion affinities due to their similar 2,6-bis(thiomethyl)pyridine motif. If that is the case, it might be possible to correlate spectroscopic changes during a metal titration, with the starting concentration. For this, Cu(II) titrations were performed on solutions containing 40 μM of **cys₂pyr** or **pyr-GS₂** for which

concentration were based on mass, and monitored the pyridinyl $\pi \rightarrow \pi^*$ absorbance. Blank solutions containing 80 μM L-glycine were used as controls, in order to account for Cu(II) coordination occurring through the 2,6-bis(thiomethyl)pyridine motif, only.

CuCl₂ addition to solutions containing 40 μM of **cys₂pyr** or **pyr-GS₂** buffered at pH 8.0, results in the steady increase of the absorbance at 240 nm (see Figure 2.12A and B). Similarly, an increase in the absorbance centred at 240 nm was observed upon addition of CuCl₂ to solution containing buffer only (see Figure 2.12C). In contrast, Cu(II) addition to a solution containing 80 μM L-glycine buffered at pH 8.0, resulted in the appearance of a new peak initially centred at 235 nm, but which shifted to lower wavelength at higher Cu(II) concentration (see Figure 2.12D). However, ZnCl₂ addition to solutions containing **cys₂pyr** or **pyr-GS₂** buffered at pH 8.0, did not result in any changes to the UV spectra recorded (data not shown).

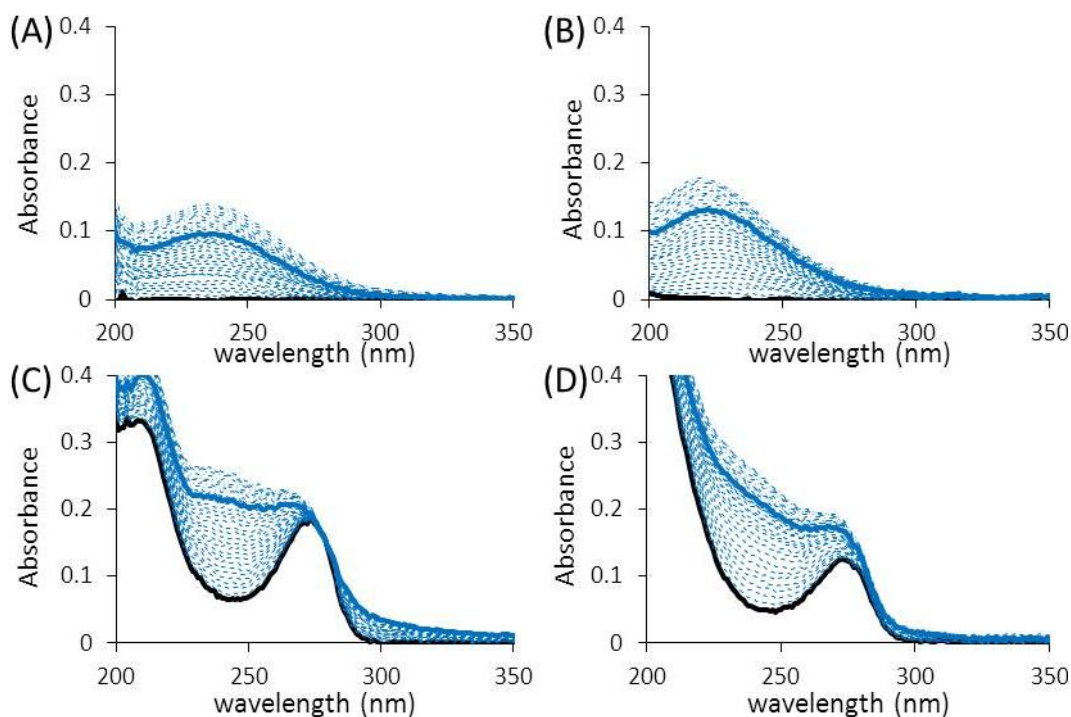


Figure 2.12 – UV spectra recorded upon CuCl₂ addition to either blank solution (A), 80 μM L-glycine (B), 40 μM **cys₂pyr** (C), or 40 μM **pyr-GS₂** (D) (based on mass). All solutions were buffered at pH 8.0 with 20 mM phosphate. (—) 0 eq., (—) 1 eq., (····) between 0 and 1 eq., and (---) more than 1 eq. CuCl₂ added.

The absorbance at 240 nm was plotted as a function of the Cu(II) concentration (see Figure 2.13). Plots are consistent with either a two stage increase (**cys₂pyr**, **pyr-GS₂** and L-glycine), or a one stage increase (blank) of the absorbance. A titration curve for the blank was fitted by linear regression using a one stage model indicating a linear coefficient of 2.2 mM^{-1} (see Figure 2.13A). The remaining three curves were fit by linear regression to two stage models, and equivalent points were calculated by extrapolation. For the conjugates, the increase of the absorbance at 240 nm was estimated to reach transition points at $C = 38.4 \text{ }\mu\text{M}$ (**cys₂pyr**) and $24.4 \text{ }\mu\text{M}$ (**pyr-GS₂**) of Cu(II) added (see Figure 2.13C and D). The first stage of absorbance growth are characterised by different linear coefficient for solutions containing **cys₂pyr** ($0 - 38.4 \text{ }\mu\text{M}$, 3.8 mM^{-1}) and **pyr-GS₂** ($0 - 24.4 \text{ }\mu\text{M}$, 5.4 mM^{-1}). In contrast, linear coefficients relative to the second stage of absorbance growth are identical to that recorded for the blank solution, ca **cys₂pyr** ($38.4 - 60 \text{ }\mu\text{M}$, 2.2 mM^{-1}) and **pyr-GS₂** ($24.4 - 60 \text{ }\mu\text{M}$, 2.2 mM^{-1}). This contrasts with the fitting of the $A_{240\text{nm}}$ plot for the L-glycine solution, where the equivalence point is at $27.4 \text{ }\mu\text{M}$ Cu(II) and linear coefficients of $0.003 \text{ }\mu\text{M}^{-1}$ ($0 - 27.4 \text{ }\mu\text{M}$), and $0.0014 \text{ }\mu\text{M}^{-1}$ ($27.4 - 60 \text{ }\mu\text{M}$) were measured.

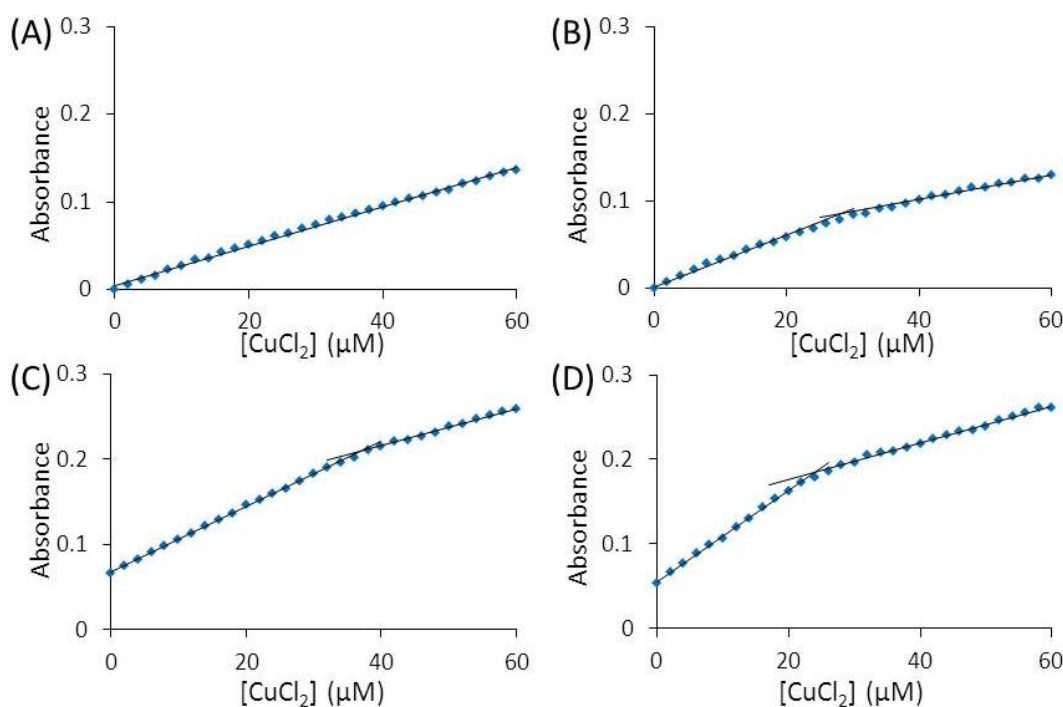


Figure 2.13 – Plot of the evolution of absorbance at 240 nm as a function of CuCl₂ concentration. (A) buffer blank, (B) 80 μM L-glycine, (C) 40 μM **cys₂pyr**, (D) 40 μM **pyr-GS₂**. Slopes are fitted to one or two linear functions (—).

These titrations indicate that both **cys₂pyr** and **pyr-GS₂** seems to interact weakly with Cu(II), however they behave very differently than L-glycine and 2,6-methylpyridine (data not shown), suggesting that they coordinate Cu(II) through the pyridinyl nitrogen and/or sulphurs from the thioether linkages. For both **cys₂pyr** and **pyr-GS₂**, Cu(II) addition past the equivalence point is consistent with Cu(II) addition to buffer. For the cysteine conjugate, the equivalence point determined is consistent with the expected 1:1 **cys₂pyr**:Cu(II) ratio. However, this is not the case for the GSH conjugate, consistent with the previous conclusion concerning the presence of a hydrated TFA salt, and therefore unreliable concentration based on mass. Indeed, if one assume that both **cys₂pyr** and **pyr-GS₂** bind Cu(II) in a 1:1 ratio, the equivalence value measured upon Cu(II) addition would correspond to the actual concentration of **cys₂pyr** and **pyr-GS₂**. If the UV spectra of **cys₂pyr** and **pyr-GS₂** are corrected to account for this, two spectra which are superimposable in the $\pi \rightarrow \pi^*$ band region are obtained. In summary, polypyridine conjugates containing peptide substituents are likely to be isolated as hydrated TFA salts, and the molecular weight cannot be reliably used. Rather, the absorbance of $\pi \rightarrow \pi^*$ transition should be used to determine their concentration in solution.

2.2.2 – Polypyridine linkers Me₂bipy and Me₂terpy

2.2.2.1 – Design and preparation of the polypyridine linkers

Our aim was to study a polypyridine linker that can act as a regulation site capable of controlling the relative orientation of the peptide moieties. In the case of **bipy** and **terpy**, protonation of pyridinyl nitrogen(s) are known to result in conformational transitions (**bipy**: *trans* \rightarrow *cis*, **terpy**: *trans-trans* \rightarrow *trans-cis* \rightarrow *cis-cis*),[11,41] therefore the cation H⁺ alone might be able to regulate the peptide moieties' orientation. **Bipy** and **terpy** have different Brønsted basicity (**bipy** 4.44, **terpy** 2.59 - 4.16).[41] Both basicity and conformation of the

protonated state might change upon substitution of **bipy** and **terpy**, especially upon attachment of bulky peptide arms. The stability constant for the 1:1 complexes that Cu(II) and Zn(II) form with **terpy** are greater by roughly 4 and 1.6 log unit, respectively, compare to those formed with **bipy** (see Chapter 1). This is particularly important when undertaking work under biologically relevant conditions, where high concentration of species able to effectively compete for metal ion coordination are present. Working with both **bipy** and **terpy** linkers, for which the substituent position would be different, would give a broader perspective, and would increase the odds to observe a *trans*-to-*cis* conformational rearrangement upon binding to a transition metal ion, at physiological pH (unlike Kelly and co-workers [20]).

Two linker units form the focus of this work: 5,5'-dimethyl-2,2'-bipyridine (**Me₂bipy**) and 6,6''-dimethyl-2,2',6',2''-terpyridine (**Me₂terpy**). **Me₂bipy** has previously been used to dimerise peptides through cysteine alkylation.[42][43] A *trans*-to-*cis* conformational transition within the **Me₂bipy** linker is expected to result in realignment of peptide substituents, while the distance separating them should remain unchanged (see Figure 2.14A). The reorientation could be compensated for by the free- rotation of the dimethylsulfide group connecting the **bipy** unit to the peptide backbone (see Figure 2.14C), and therefore this linker can be thought of as allosteric-ineffective.[4] In contrast, *trans*-to-*cis* conformational transitions of the **Me₂terpy** linker are expected to modify the distance separating the two peptide substituents (see Figure 2.14B). Distances separating the two methyl carbon substituents, in the *cis*- or *trans*- conformation, were estimated from reported crystal structures of analogous compounds: 5,5'-dimethyl-2,2'-bipyridine,[44] 6,6''-bis(trimethylsilylethynyl)-2,2':6',2''-terpyridine,[45] catena-poly[(5,5'-dimethyl-2,2'-bipyridine- κ^2 N,N')cadmium(II)]-di- μ -iodido[46] and [[6,6''-dicarboxylic acid bisanthracen-9-ylamide-(2,2':6',2''-terpyridine- κ^3 N,N',N'')]]zinc(II)]-di-trifluoroacetic acid.[12]

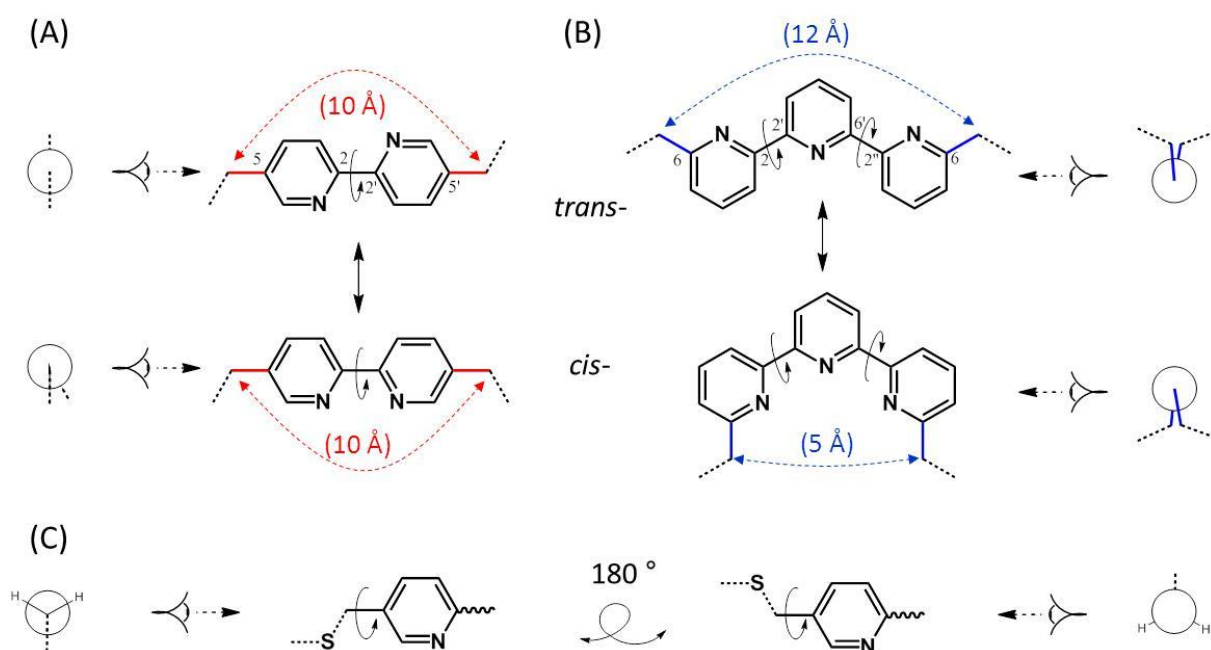


Figure 2.14 – Schematic illustrating the impact of conformational transition in (A) **Me₂bipy** and (B) **Me₂terpy** on the inter-substituent distance and relative orientation. (C) Illustration of the flexibility of the thioether linkage of (2-polypyridine) substituted at position 5-. Distances displayed are estimated based on reported structures for analogous compounds.[12,44-46]

The dimethyl- compounds were either purchased (**Me₂bipy**) or prepared following previously reported procedures (**Me₂terpy**).[47,48] NMR was used to estimate their interaction with Zn(II), and the resulting conformational transitions (see section 2.2.2.2). However, the synthetic strategy chosen for peptide dimerisation required the preparation of the di-bromomethyl polypyridine derivatives. Therefore, 5,5'-bis(bromomethyl)-2,2'-bipyridine (**bipy-Br₂**), and 6,6''-bis(bromomethyl)-2,2':6',2''-terpyridine (**terpy-Br₂**) were prepared following previously reported procedures,[48,49] and were characterised by ¹H and ¹³C NMR, UV, as well as HRES-TOF spectroscopy. Interestingly, the UV profile recorded for these species display $\pi \rightarrow \pi^*$ transitions at slightly higher wavelength (5-15 nm) compared to previously reported values for **bipy** and **terpy**. [41] However, the substitution pattern of the polypyridine ring can be confirmed through the coupling pattern observed in the ¹H NMR spectra.

2.2.2.2 – Zn(II) binding studies for Me₂bipy and Me₂terpy monitored by ¹H NMR spectroscopy.

¹H NMR was used to study the interaction of Zn(II) with the Me₂bipy and Me₂terpy polypyridine linkers. The ¹H NMR spectrum of 25 mM Me₂bipy recorded in DMSO-d₆ displays 3 signals in the aromatic region, which are centred at 8.49 (d), 8.24 (d) and 7.73 ppm (dd). These peaks were attributed to H⁶, H³, and H⁴ respectively, as determined by 2D COSY NMR. Addition of aliquots of ZnCl₂ solution in DMSO-d₆, resulted in a shift of all three peaks toward higher frequencies, 8.60, 8.50, and 8.10 ppm (see Figure 2.15A). Though the H⁶ proton does not shift significantly a plot of H³ and H⁴ chemical shift as a function of Zn(II) equivalence, is consistent with a 1:1 binding ratio (see Figure 2.15B). The stoichiometry of the main complex formed between ZnCl₂ and Me₂bipy was confirmed to be 1:1 by performing a jobplot analysis (see Figure 2.15C).

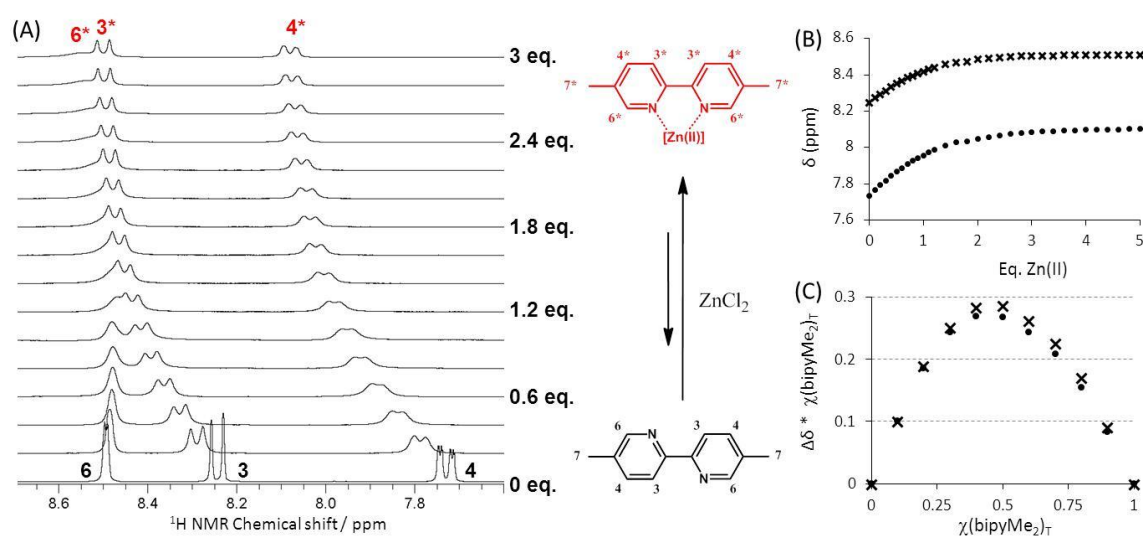


Figure 2.15 – Stack of NMR spectra (A) and plot of the chemical shift evolution of Me₂bipy aromatic resonances as a function of Zn(II) equivalence (B), relative to the Zn(II) titration experiment monitored by ¹H NMR at 300 MHz in DMSO-d₆. (C) Jobplot based on ¹H NMR spectra at 300 MHz in DMSO-d₆ recorded for mixtures of Me₂bipy and Zn(II) (C_{total} constant and equal to 25 mM). The molar fraction of Me₂bipy complexed with Zn(II) ($\chi(\text{Zn:Me}_2\text{bipy})$), calculated from the variation of chemical shift for the different resonances) is plotted versus the molar fraction represented by Me₂bipy (bound and free) among the different species present in solution ($\chi(\text{Me}_2\text{bipy})_T$). (●) H⁴; (X) H³.

The ^1H NMR spectrum of a solution of **Me₂terpy** in DMSO- d_6 (23.2 mM, 0.75 mL) displays four signals in the aromatic region, which are centred at 8.42 ($\text{H}^3 + \text{H}^{3a}$), 8.07 (H^4), 7.88 (H^{4a}) and 7.35 ppm (H^{5a}). Aliquots of a concentrated ZnCl_2 stock solution in DMSO- d_6 (232 mM) were added, resulting in a gradual decrease in intensity of these four resonances, and the gradual appearance of four new peaks, 8.81 (H^3), 8.54 (H^4), 8.52 (H^{3a}), 8.14 (H^{4a}) and 7.64 ppm (H^{5a}), consistent with slow exchange on the NMR-timescale (see Figure 2.16).[12] All the original four resonances have fully disappeared on addition of one equivalent of ZnCl_2 , indicating a binding ratio of 1:1 **Me₂terpy**: ZnCl_2 .

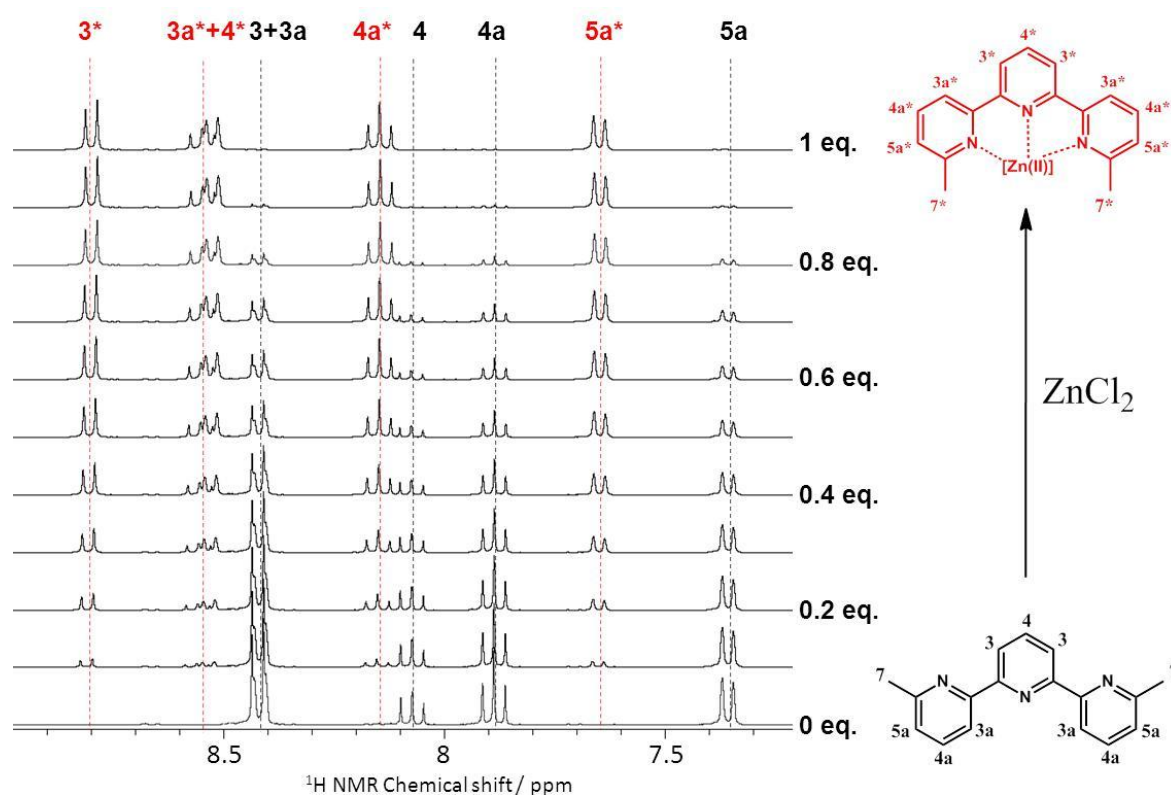


Figure 2.16 – ^1H NMR Zn titration of a 23.2 mM **Me₂terpy** solution in DMSO- d_6 , recorded at 300 MHz.

NOESY spectra of **Me₂terpy** can provide information on the conformation of the terpyridine unit through inter-ring coupling, in particular the $\text{H}^3 \leftrightarrow \text{H}^{3a}$ cross-peak. However, data acquisition is complicated by overlapping of H^3 and H^{3a} resonances in polar solvents (CD_3CN and DMSO- d_6). Unfortunately, complexes formed between **Me₂terpy** and ZnCl_2 have poor solubility in solvents of low or intermediate polarity (e.g. CD_3CN and CDCl_3). The

^1H NMR spectrum of **Me₂terpy** in CDCl_3 , displays two distinct doublets attributed to H^3 and H^{3a} , at 8.46 and 8.41 ppm (assigned based on COSY NMR), respectively. In the NOESY spectra, no cross-peaks were observed between resonances at 8.46 and 8.41 ppm (see Figure 2.17A). In contrast, the NOESY spectrum recorded of **Me₂terpy** in the presence of one equivalent of ZnCl_2 in DMSO-d_6 , displays a cross-peak between signals at 8.81 (H^{3*}) and 8.52 ppm (H^{3a*}) (attribution based on COSY spectrum) (see Figure 2.17B).

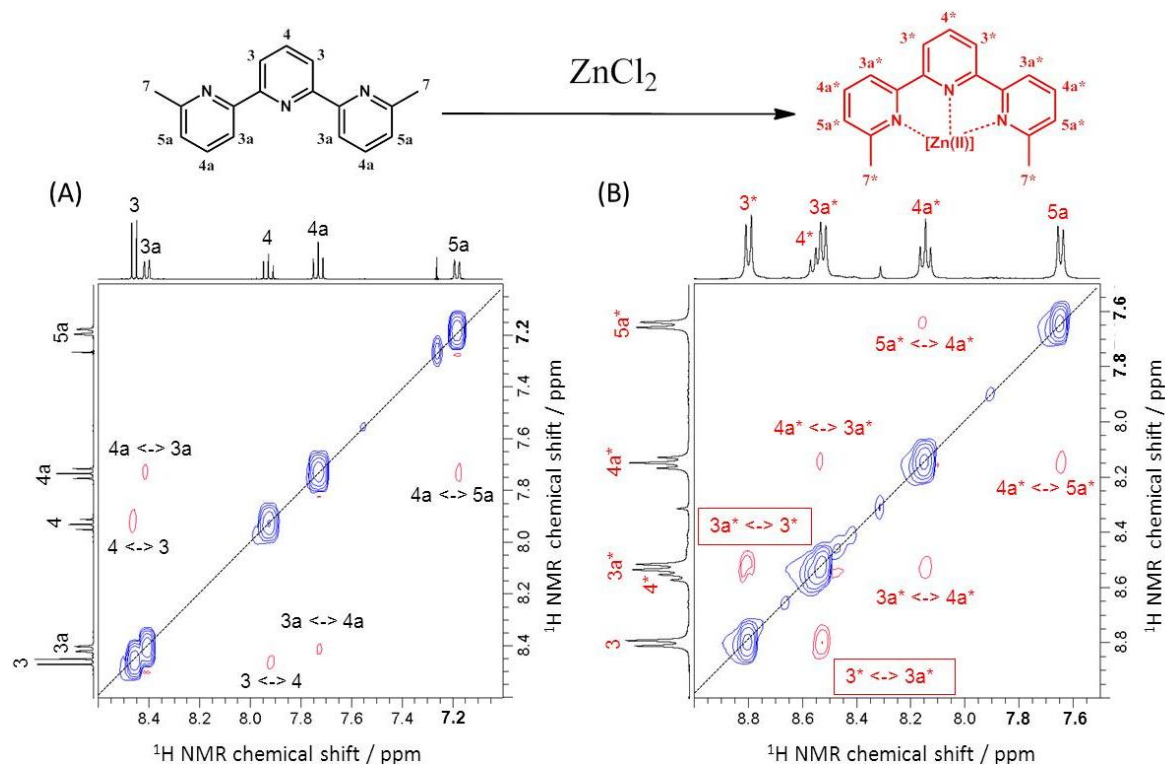


Figure 2.17 – NOESY spectra of **Me₂terpy** recorded in (A) the absence (CDCl_3) and (B) presence of 1 eq. ZnCl_2 (DMSO-d_6), at 400 MHz ($T = 298\text{ K}$). Positive levels (—), negative levels (—).

^1H NMR can also provide an indication on the conformation of polypyridine in the absence and presence of metal ions. Previous NMR studies conducted on **bipy** and derivatives thereof, concluded that H^4 and H^5 resonances shift towards higher frequency, upon *trans*-to-*cis* transition, whereas H^6 and H^3 can display contrasting shifts, depending on the **bipy** substitution, the nature of the cation bound (monovalent/divalent), and the solvent.[3,50-52] This is in line with our NMR measurement on **Me₂bipy** recorded in the absence and presence of ZnCl_2 . Similarly, broadening of the peak upon Zn(II) addition is consistent with the rapid kinetics previously reported for **bipy**- Zn(II) complexation (see Chapter 1).[53]

Concerning **Me₂terpy**, our results suggest that all aromatic signals also move to higher chemical shifts upon ZnCl₂ addition, and proton signals relative to the central pyridine (H³ and H⁴) experience higher shift than their counterpart (H^{3a}, H^{4a} and H^{5a}). This is in stark contrast to two recent reports, which indicate that H^{3a} resonances move to lower chemical shift upon *trans*-to-*cis* conformational transition.[11,12] However, the ¹H NMR spectra of **Me₂terpy** recorded in the absence and presence of ZnCl₂ are similar to the NMR spectra of **terpy** and its Pt(II) complex, reported by Cini *et al.*[54] Moreover, NOESY spectra would be consistent with **Me₂terpy** adopting a *trans-trans*- conformation in the absence of metal ions, where protons H³ and H^{3a} point in opposite directions and are separated by distances of ~4.5 Å,[45] associated with weak coupling in the NOESY. In contrast, the NOESY spectra of **Me₂terpy** in the presence of 1 equivalent of ZnCl₂ in DMSO-d₆ displays cross-peaks between protons H³ and H^{3a}, consistent with the expected *cis*- conformation of the **Me₂terpy**:ZnCl₂ complex, where H³ and H^{3a} would be separated by approximately ~2.3 Å.[12] Our two NOESY experiments cannot be directly compared as they were recorded in different solvents (as a consequence of resonance overlapping and solubility issues previously described), and overlapping of resonances H^{3a*} and H^{4a*}. On a side-note, the ¹H NMR spectra of **Me₂terpy** recorded in the presence of ZnCl₂ display sharper peaks compared to those recorded for **Me₂bipy**, which is consistent with the tenfold greater kinetics of dissociation for **bipy**-Zn(II) compared to **terpy**-Zn(II) complexes (see Chapter 1).[55]

2.2.3 – Cysteine dimerisation with polypyridine linkers

2.2.3.1 – Preparation of polypyridine-cysteine conjugates

Our synthetic strategy is to dimerise peptides through the cysteine side-chain: initial studies included dimerising two equivalents of cysteine in order to establish optimal conditions, and to evaluate if the metal binding properties of the linker unit is retained. The preparation of **cys₂bipy** and **cys₂terpy** were first attempted following the procedure used to prepare **cys₂pyr** and **pyr-GS₂** (section 2.2.1.1), however little cysteine-polypyridine

conjugates, if any, could be obtained, most likely due to the low solubility of **bipy-Br₂** and **terpy-Br₂** in polar solvents such as DMF and water.

Bipy-Br₂ and **terpy-Br₂** were reacted with cysteine at room temperature in different solvent mixtures such as buffer/acetonitrile, buffer/acetone, buffer/THF (where buffer corresponds to 100 mM Tris.HCl pH 8.0), again few conjugates were formed. However, **cys₂bipy** and **cys₂terpy** were efficiently prepared in buffer/acetonitrile mixtures when heated at 338 K in a microwave reactor for 3 hours. Analytical HPLC run on the crude mixtures indicate the presence of either a few (**cys₂bipy**), or numerous side-products (**cys₂terpy**) (see Figure 2.18). Cysteine alkylation reactions were reported to better proceed in the presence of the reducing agent tris(2-carboxyethyl)phosphine (TCEP).[56] However, under our experimental conditions (338 K), TCEP forms an adducts with **Me₂terpy** as identified by MALDI-TOF.

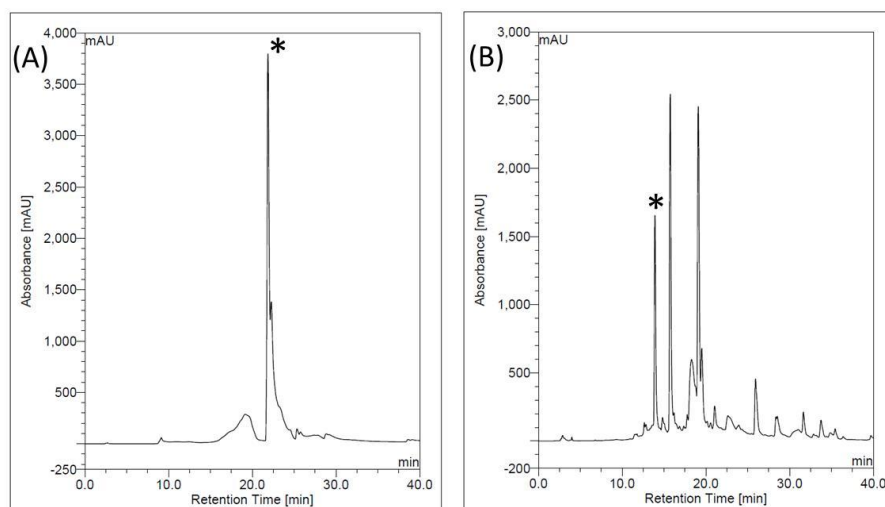


Figure 2.18 – Analytical C18 HPLC traces of the crude mixture for the synthesis of **cys₂bipy** (A) and **cys₂terpy** (B). (A) Mixture was eluted with 0 to 20 % gradient of acetonitrile in water over 40 min ($\lambda = 280$ nm) ; (B) mixture was eluted with 0 to 100 % gradient of acetonitrile in water over 40 min ($\lambda = 290$ nm). (*) marks the top of the peak attributed to the expected compound (assigned based on mass spectrometry).

Cys₂bipy and **cys₂terpy** were purified by preparative HPLC eluted with water/acetonitrile gradients (no TFA), and dried *in vacuo*. However, addition of acid (HCl) was required to re-solubilise the resulting solids and subsequently characterise them. The ¹H

NMR spectrum of **cys₂bipy** and **cys₂terpy** solutions were recorded at acidic pD only. Raising the pD of the NMR solution (millimolar range) to 7, resulted in precipitation of **cys₂bipy** and **cys₂terpy** conjugates from the solution. After drying *in vacuo*, conjugates were recovered as hydrated salts, for which elemental analysis was not reproducible, in part due to sample hydration. The solubility was sufficient at physiological pH to allow for UV studies (micromolar range).

2.2.3.2 – Cu(II)/Zn(II) titration of **cys₂bipy** and **cys₂terpy** monitored by UV spectroscopy

Previous reports indicate that the $\pi \rightarrow \pi^*$ transition of **bipy** and **terpy** shifts to lower energy upon metal coordination. This was proposed to result from π^* orbital stabilisation,[57,58] and subsequent reduction of the energetic separation between the π and π^* level. In the absorbance spectra of **terpy**, the lower energy $\pi \rightarrow \pi^*$ band separates into two sub-bands, upon metal binding. This is explained by the fact that the three pyridine rings are not equivalent. The resulting lower energy band with weaker absorbance is attributed to the central pyridinyl ring.[41] The shift of $\pi \rightarrow \pi^*$ band was previously used to estimate the metal-ligand affinity.[59]

Despite having established that the theoretical molecular weight differs from the experimental molecular weight due to recovery as a salt, it remained to determine the relationship between these taking advantage of the UV chromophore for the $\pi \rightarrow \pi^*_1$ transition. For this, the impact of Cu(II) and Zn(II) addition on the UV spectra of **cys₂bipy** and **cys₂terpy** for which concentrations were based on mass was first studied. The UV spectrum of **cys₂bipy** contains two bands with maxima at 296 nm ($\epsilon_{296 \text{ nm}} 14,397 \text{ M}^{-1} \text{ cm}^{-1}$) and 245 nm ($\epsilon_{245 \text{ nm}} 10,240 \text{ M}^{-1} \text{ cm}^{-1}$) assigned as $\pi \rightarrow \pi^*_1$ and $\pi \rightarrow \pi^*_2$, respectively. The UV spectrum recorded for **cys₂terpy** also displays a band at 298 nm ($\epsilon_{298 \text{ nm}} 16,857 \text{ M}^{-1} \text{ cm}^{-1}$) assigned as $\pi \rightarrow \pi^*_1$, however the $\pi \rightarrow \pi^*_2$ around 225 nm overlaps with a more intense band towards the far-UV domain. ZnCl₂ addition to solutions containing 50 μM **cys₂bipy** or 70 μM **cys₂terpy**

(concentration based on mass) at pH 8, results in a redshifts of the $\pi \rightarrow \pi^*$ band centred at 296 nm (**cys₂bipy**) to 313 and 320 nm, and 298 nm (**cys₂terpy**) to 328 and 340 nm (see Figure 2.19). Isosbestic points at 304 (**cys₂bipy**) and 315 nm (**cys₂terpy**) are consistent with two-state transitions.

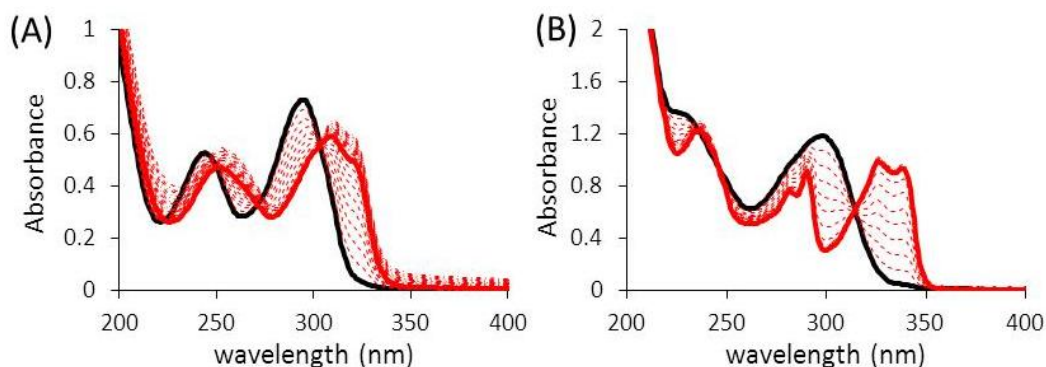


Figure 2.19 – UV spectra for the Zn(II) titration of model switches. ZnCl₂ was added to solutions containing either 50 μM of **cys₂bipy** (A) or 70 μM of **cys₂terpy** (B) in 20 mM phosphate buffer pH 8. (—) 0 eq., (—) 1 eq., (···) between 0 and 1 eq., (---) more than 1 equivalent ZnCl₂ added.

The absorbance at 320 (**cys₂bipy**) and 340 nm (**cys₂terpy**), was plotted as a function of the Zn(II) concentration, and fitted by linear regression (two stage): equivalent points of 35.5 (**cys₂bipy**) and 52.3 μM (**cys₂terpy**) were estimated (see Figure 2.20).

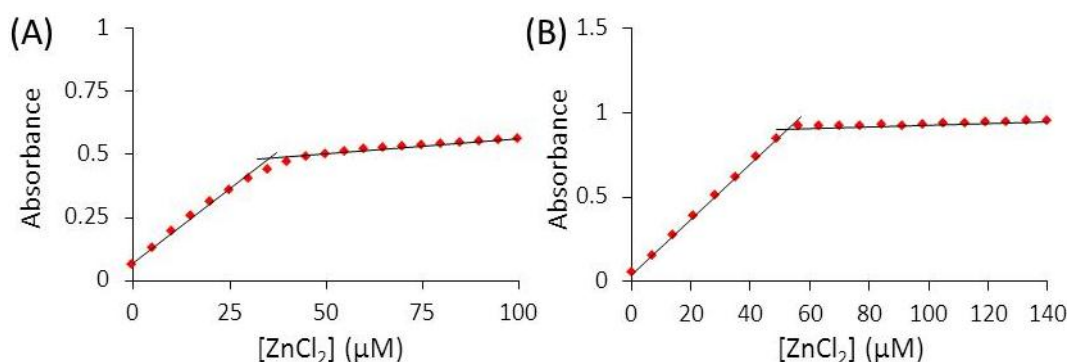


Figure 2.20 – Plot of absorbance for ZnCl₂ titration of (A) **cys₂bipy** (50 μM) monitored at 320 nm or (B) **cys₂terpy** (70 μM) monitored at 340 nm. Line (—) represents linear regression fit.

Cys₂bipy and **cys₂terpy** titrations with CuCl₂ were similarly monitored by absorption spectroscopy. CuCl₂ addition to solutions containing 50 μM **cys₂bipy** and 70 μM **cys₂terpy** (concentration based on mass) at pH 8, result in a redshift of the $\pi \rightarrow \pi^*$ band centred at 296

nm (**cys₂bipy**) to 318 and 328 nm, and at 298 nm (**cys₂terpy**) to 324 and 335 nm (see Figure 2.21A and B). Isosbestic points at 308 nm (**cys₂bipy**) and 311 nm (**cys₂terpy**) are consistent with a two-state transition.

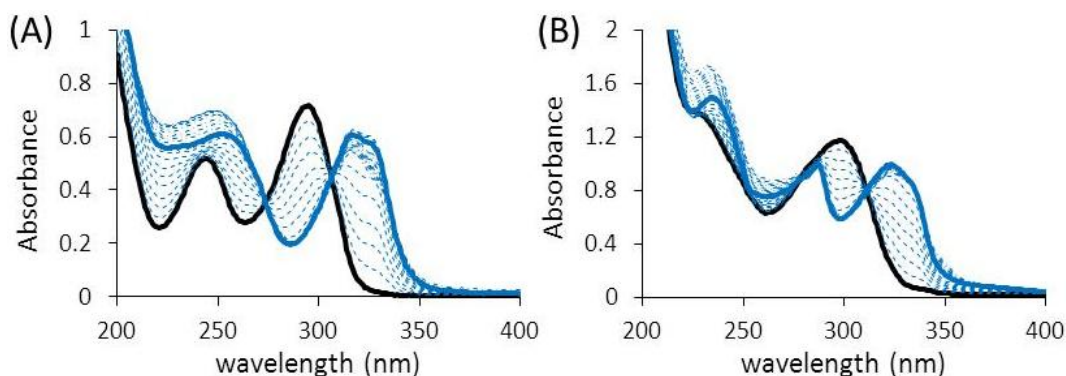


Figure 2.21 – UV spectra for the metal titration of model switches. CuCl_2 was added to solutions containing either 50 μM of **cys₂bipy** (A) or 70 μM of **cys₂terpy** (B) in 20 mM phosphate buffer pH 8. (—) 0 eq., (---) between 0 and 1 eq., (—) 1 eq., (— · —) more than 1 eq. CuCl_2 added.

The absorbance at 328 (**cys₂bipy**) and 335 nm (**cys₂terpy**), was plotted as a function of the Cu(II) concentration (see Figure 2.22C and D), and an equivalent points of 57.4 μM was estimated for **cys₂terpy**. In contrast, an equivalent point for Cu(II) titration of **cys₂bipy** could not be estimated, due to the absence of a plateau.

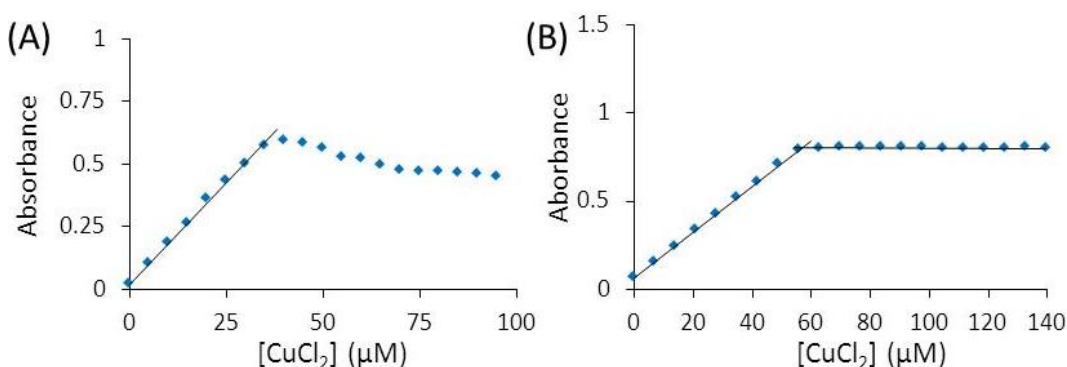
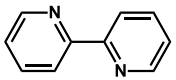
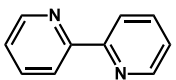
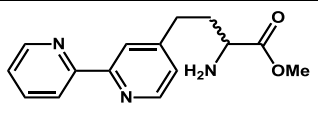
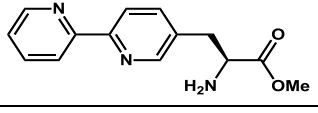
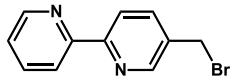
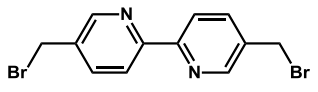
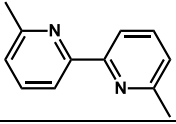
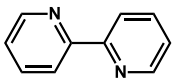
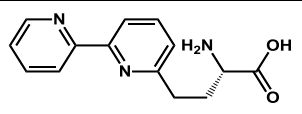


Figure 2.22 – Plot of absorbance for CuCl_2 titration of (A) **cys₂bipy** (50 μM) monitored at 328 nm, or (B) **cys₂terpy** (70 μM) monitored at 335 nm. Line (—) represents linear regression fit.

The conjugates **cys₂bipy** and **cys₂terpy** were previously prepared as salts. Therefore, solution preparation based on mass most likely results in concentration overestimation. Absorbance properties of metal-free **bipy** and **terpy** derivatives were previously reported, and

Table 2.1 contains a summary of the $\pi \rightarrow \pi^*$ characteristics for some **bipy** derivatives reported in the literature.

Table 2.1 – Range of λ_{\max} ($\epsilon_{\lambda_{\max}}$) values for $\pi \rightarrow \pi^*$ bands of some **bipy** derivatives.

compound	$\pi \rightarrow \pi^*_1$	$\pi \rightarrow \pi^*_2$	solvent	Ref.
	283 (10,200)	244 (6,600)	EtOH	[41]
	281 (13,580)	235 (10,448)	MeOH	[60]
	282 (13,300)	238 (10,600)	MeOH	[61]
	286 (17,000)	240 (12,600)	MeOH	[61]
	289 (26,700)	244 (18,600)	MeOH	[61]
	298 (13,000)	256 (6,900)	EtOH	this work
	289 (19,196)	237 (11162)	MeOH	[60]
	281 (13,800)	233 (10,600)	Water (pH 12)	[41]
	284 (13,420)	238 (8,950)	Water (free zwitterion)	[62]

Absorbance profiles of **cys₂bipy** and **cys₂terpy** display $\pi \rightarrow \pi^*$ maxima at higher wavelengths than previously reported profiles for **bipy** and **terpy** derivatives (see Table 2.1 and [63] respectively). However, they are in line with **bipy-Br₂** and **terpy-Br₂** profiles recorded in EtOH (see section 2.4.2 and Table 2.1), and consistent with *trans*- rather than non-planar *cis*- arrangement of pyridine rings for the metal-free conjugates.[60] Determined extinction coefficients are primarily influenced by the purity of the compounds, but also differ

slightly, depending on the substituents (nature, number and position), and the solvent used. Therefore, it was not possible to define **cys₂bipy** and **cys₂terpy** $\pi \rightarrow \pi^*$ extinction coefficient based on previously reported values, despite previous reports of **bipy** or **terpy** insertion within macromolecules used such methods.[31,64]

Redshift of **cys₂bipy** and **cys₂terpy** $\pi \rightarrow \pi^*$ bands upon Cu(II) and Zn(II) addition are consistent with both species binding to both metal ions. As for **pyr-GS₂**, the $\pi \rightarrow \pi^*$ extinction coefficient of **cys₂bipy** and **cys₂terpy** can be estimated based on their coordination properties, and assuming a 1:1 binding ratio. Previous studies of the Cu(II) and Zn(II) coordination of **terpy** are consistent with a 1:1 ratio.[59,63] However, small fractions of 2:1 complexes are expected for **bipy**-Cu(II) and **bipy**-Zn(II) at high concentration and when the ligand is present in large excess, despite 1:1 complexes being largely predominant.[22,65] The assumed 1:1 ratio is supported by the Zn(II) titration of **Me₂bipy** and **Me₂terpy** (see section 2.2.2.2).

The total concentration of **cys₂bipy** and **cys₂terpy** for the titration was corrected such as it is equal to the concentration of metal at equivalence, ca 35.5 μM (**cys₂bipy**), and 55.3 μM (**cys₂terpy**). The concentration of **cys₂bipy** and **cys₂terpy** were then used to recalculate extinction coefficient resulting in $\epsilon_{296\text{ nm}}$ 20,000 $\text{M}^{-1}\text{ cm}^{-1}$ (**bipy** conjugate) and $\epsilon_{298\text{ nm}}$ 21,000 $\text{M}^{-1}\text{ cm}^{-1}$ (**terpy** conjugate). These were subsequently used for quantification of all analogous **bipy** and **terpy** peptide dimer conjugates used in this work.

The plot of absorbance at 320 nm relative to Zn(II) equivalents added of **cys₂bipy** was fitted to a 1:1 binding ratio using a non-linear regression (see Figure 2.23), based on a previously reported method.[31,59]. A $\log K_M$ of 6.05 ± 0.06 was calculated for Zn(II):**cys₂bipy**, taking into account the competitive metal ion binding of the phosphate buffer employed in these experiments, for which the Zn(II) binding constant is known (see section 2.4.3.4).[66] In contrast, analogous plots for the Cu(II) and Zn(II) titrations of **cys₂terpy**, do not display sufficient dissociation of the metal complexes and are therefore, not suitable for binding constant determination (see Figures 2.20B and 2.22B). To overcome this problem,

these titrations needed to be carried out either at a lower concentration, or in the presence of a species able to compete for metal-binding, for which the K_M is known.[33] Finally, the plot of absorbance at 328 nm relative to Cu(II) titration of **cys₂bipy** suggests that more than two species are in equilibrium, and therefore cannot be fit to a 1:1 model (see Figure 2.22A).

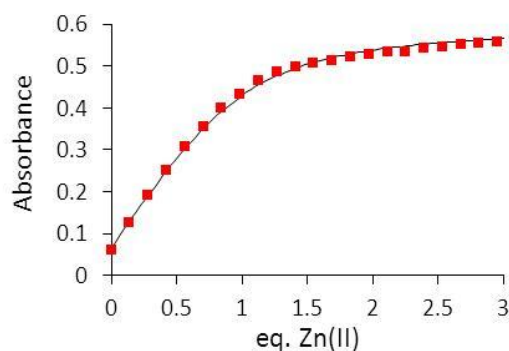


Figure 2.23 – Zn(II) titration of **cys₂bipy**, plot of absorbance at 320 nm as a function of Zn(II) added (■), and its best fit calculated by non-linear regression to a 1:1 model (—).

2.2.4 – Glutathione model compounds

2.2.4.1 – Design and preparation

The prime objective of this chapter is the study of small polypyridine peptide conjugates coordination to Cu(II) and Zn(II), how this impacts on the polypyridine linker, and the resulting impact on peptide orientation. Therefore, the preparation of conjugates bearing slightly larger peptide moieties, but containing the same **bipy** and **terpy** linkers, and their comparison to **cys₂bipy** and **cys₂terpy**, was envisioned. Larger peptide units might compete for metal binding and increase the solubility of conjugates under physiological conditions, thus allowing for a conformational study by NMR.

Following the procedure used to prepare **cys₂bipy** and **cys₂terpy** (less heating is required as products are more soluble in the reaction mixture), analogous **bipy-GS₂** and **terpy-GS₂** conjugates were obtained in high yields, purified by HPLC, and recovered as TFA salts. The following sections report the study of Cu(II) and Zn(II) interaction with **bipy-GS₂**

and **terpy-GS₂**, for which concentrations are based on their UV profile and the extinction coefficients calculated for their dicysteine analogues.

2.2.4.2 –pH, Cu(II), and Zn(II) titrations of bipy-GS₂ and terpy-GS₂ monitored by UV-visible spectroscopy

Nakamoto first studied the pH dependence of **bipy** and **terpy** in water, and demonstrated that at low pH, free **bipy** and **terpy** display similar absorption profiles to those of the metal complexes, consistent with the *cis*- and *cis-cis*- conformation. However, the $\pi \rightarrow \pi^*$ bands shift to higher energy on increasing the pH, and resemble those recorded in organic solvents, consistent with the *trans*- and *trans-trans*- conformations.[41,59] It was hypothesised that deprotonation of the pyridinyl ring on increasing the pH, resulted in *cis*-to-*trans* conformational transitions of the **bipy** and **terpy**, and has been supported more recently by theoretical studies.[67-69] Protonation of the pyridinyl nitrogen is believed to induce delocalisation of π electrons over two rings instead of only one for unprotonated analogues.[3,67]

Similar to the UV spectra of unsubstituted **bipy** and **terpy**, the absorbance of **bipy-GS₂** and **terpy-GS₂** in aqueous solution is sensitive to pH variation. The UV spectra of a 5 μ M solution (quantification based on UV profile) of **bipy-GS₂** recorded between pH 6 and 10, display two transitions with λ_{max} at 295 ($\epsilon_{295 \text{ nm}}$ 19,600 $\text{M}^{-1} \text{cm}^{-1}$) and 245 nm ($\epsilon_{245 \text{ nm}}$ 15,200 $\text{M}^{-1} \text{cm}^{-1}$), assigned as $\pi \rightarrow \pi^*_1$ and $\pi \rightarrow \pi^*_2$ transitions (see Figure 2.24A).[70] The analogous **terpy-GS₂** spectrum displays a peak with λ_{max} 297 nm ($\epsilon_{297 \text{ nm}}$ 20,300 $\text{M}^{-1} \text{cm}^{-1}$) attributed to $\pi \rightarrow \pi^*_1$, however, the $\pi \rightarrow \pi^*_2$ transition which occurs around 221 nm, overlaps with that for the peptide bond (Figures 2.24B). Upon acidification by addition of concentrated HCl, these bands decrease in intensity whilst new bands appear at lower energy. A plot of absorbance as a function of pH allows for an approximation of the associated pK_a values (see Figure 2.24C and D).

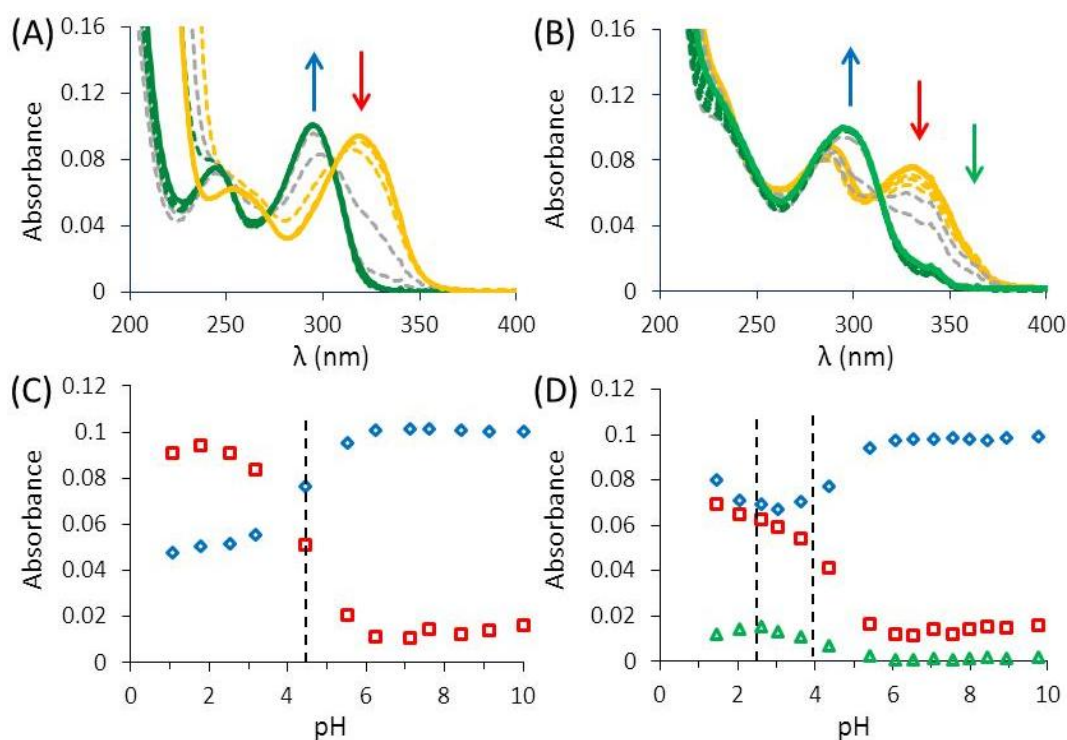


Figure 2.24 – pH Titration of model compounds monitored by UV spectroscopy. UV spectra of 5 μ M **bipy-GS₂** (A) and **terpy-GS₂** (B) recorded at pH values between 1.5 and 10, (—) acidic, (---) intermediary, (—) basic. Plot of the absorbance maxima as a function of pH for (C) **bipy-GS₂** (295 nm, \diamond ; and 319 nm, \square) and (D) **terpy-GS₂** (295 nm, \diamond ; 340 nm, \square ; and 367 nm, \triangle); dotted line represents approximation of pK_a values.

The UV spectra recorded for **bipy-GS₂** and **terpy-GS₂** at pH 8 are very similar to those recorded for **cys₂bipy** and **cys₂terpy**, under similar conditions (see section 2.2.3.2). Plots of absorbance at 295, 319 nm (**bipy-GS₂**) and 295, 340, 367 nm (**terpy-GS₂**) as a function of pH, are consistent with pK_a values within 0.5 log unit from the pK_a values reported for **bipy** (4.44) and **terpy** (2.59, 4.16).[41] More importantly, absorbance of the maxima does not change between 6-9. Therefore, slight variation of the pH solution in this range would not have important consequences on the UV spectra and the conformation of the linker unit, and that at physiological pH our linkers adopt the *trans*- conformation.

The shift of the $\pi \rightarrow \pi^*$ band was used to estimate the Cu(II) and Zn(II) binding constants to **bipy-GS₂** and **terpy-GS₂**, similarly to Zn(II) titration for **cys₂bipy** (see Figure 2.23). Titrations were performed at physiological pH (7.4), low micromolar concentrations,

and where necessary a competitor was introduced. Aliquots of a stock solution of ZnCl_2 were titrated into a $5\ \mu\text{M}$ solution of either **bipy-GS₂** or **terpy-GS₂** in 20 mM phosphate buffer pH 7.4. This resulted in the steady decrease in the absorbance at 295 (**bipy-GS₂**) and 297 nm (**terpy-GS₂**), and an increase in the absorbance at 308 and 320 nm (**bipy-GS₂**), and 330 and 340 nm (**terpy-GS₂**), respectively (see Figures 2.25A and B). Between 0 and 1 equivalent of Zn(II) added, the observation of isosbestic points at 303 (**bipy-GS₂**) and 313 nm (**terpy-GS₂**) are consistent with the clean formation of the Zn(II) complex.

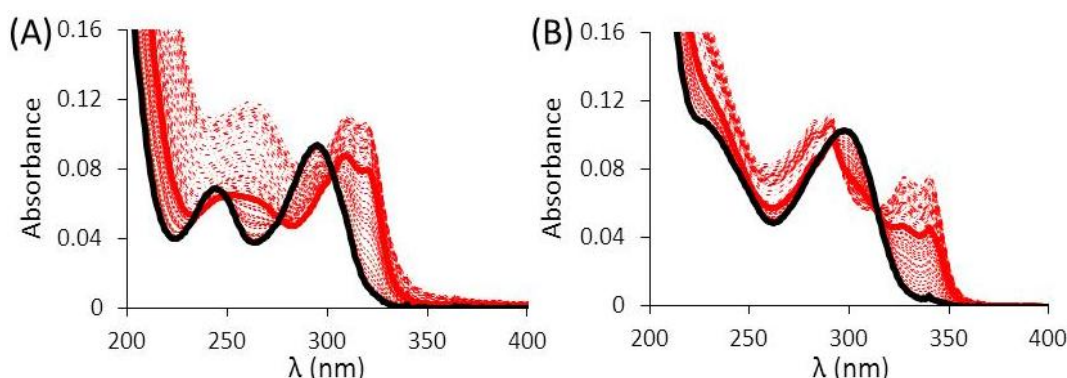


Figure 2.25 – UV spectra for the Zn(II) titration of model switches. ZnCl_2 was added to solutions containing $5\ \mu\text{M}$ of either **bipy-GS₂** (A) or **terpy-GS₂** (B) in 20 mM phosphate buffer pH 7.4. (—) 0 eq., (—) 1 eq., (···) between 0 and 1 eq., (— · —) more than 1 eq. ZnCl_2 added.

A plot of the absorbance as a function of Zn(II) concentration indicates the formation of a 1:1 complex between Zn(II) and both the model ligands (see Figure 2.26). Upon non-linear fitting to a 1:1 binding model, extinction coefficients at 320 nm for **bipy-GS₂** and Zn(II):bipy-GS_2 were determined to be $1,180\ \text{M}^{-1}\ \text{cm}^{-1}$ and $22,700\ \text{M}^{-1}\ \text{cm}^{-1}$, respectively. The extinction coefficients at 340 nm for **terpy-GS₂** and Zn(II):terpy-GS_2 were estimated to be $800\ \text{M}^{-1}\ \text{cm}^{-1}$ and $19,850\ \text{M}^{-1}\ \text{cm}^{-1}$, respectively. Formation constants, $\log K_M$, were calculated to be 6.86 ± 0.05 for Zn(II):bipy-GS_2 and 6.22 ± 0.04 for Zn(II):terpy-GS_2 , see Table 2.2.

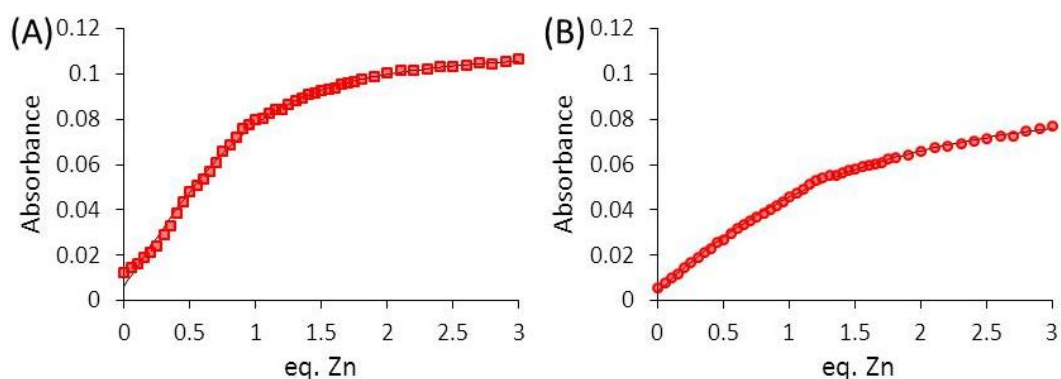


Figure 2.26 – Plot of absorbance for the Zn(II) titration of (A) **bipy-GS₂** monitored at 320 nm (■), or (B) **terpy-GS₂** monitored at 340 nm (●) versus the equivalence of Zn(II). Line represents best fit for a 1:1 Zn(II):model switch binding ratio.

Table 2.2 – Summary of data obtained for Zn(II) coordination to model switches.

	λ (nm)	ϵ_{ML} (M ⁻¹ cm ⁻¹)	K_{app} (M ⁻¹)	K_M (M ⁻¹)	R^2
Bipy-GS₂	320	$2.27 \pm 0.02 \text{ E} +04$	$1.21 \pm 0.75 \text{ E} +06$	$7.30 \pm 0.89 \text{ E} +06$	0.9965
Terpy-GS₂	340	$1.99 \pm 0.03 \text{ E} +04$	$2.77 \pm 0.29 \text{ E} +05$	$1.67 \pm 0.17 \text{ E} +06$	0.9983

The analogous titration performed with CuCl₂, resulted in a decrease in the absorbance at 295 nm (**bipy-GS₂**) and 297 nm (**terpy-GS₂**), accompanied by an increase in the absorbance at 317 and 328 nm (**bipy-GS₂**) and at 335 and 347 nm (**terpy-GS₂**), assigned to the formation of Cu(II):**bipy-GS₂** and Cu(II):**terpy-GS₂**, respectively. Between 0 and 1 equivalent of Cu(II) added, the isosbestic points at 305 (**bipy-GS₂**) and 315 nm (**terpy-GS₂**) are again consistent with a single equilibrium (see Figure 2.27A and B).

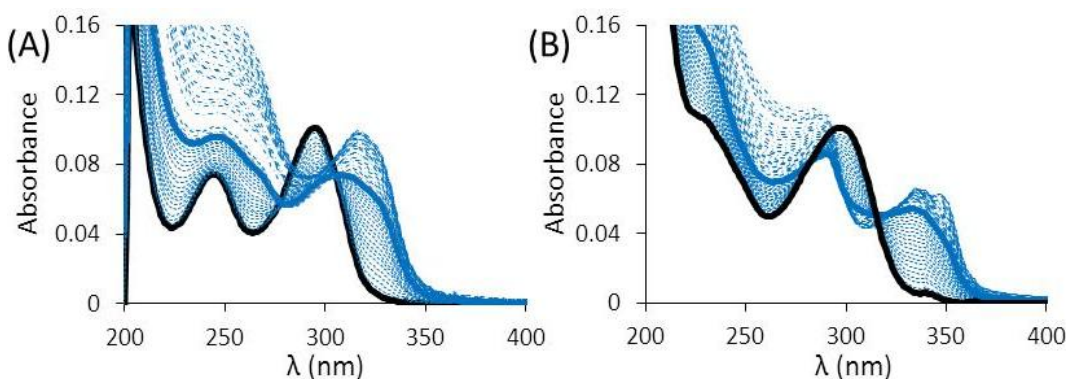


Figure 2.27 – UV spectra for the Cu(II) titration of model switches. CuCl₂ was added to solutions containing 5 μM of either **bipy-GS₂** (A) or **terpy-GS₂** (B) in 20 mM phosphate buffer pH 7.4. (—) 0 eq., (—) 1 eq., (····) between 0 and 1 eq., (---) more than 1 eq. CuCl₂ added. For (A) buffer concentration was 100 mM and 20 mM L-glycine was added as competitor.

Monitoring the absorbance at 328 nm (**bipy-GS₂**) and 335 nm (**terpy-GS₂**) and plotting these as a function of CuCl₂ equivalence, is also consistent with the formation of 1:1 complexes (see Figure 2.28). Non-linear fitting results in extinction coefficient estimations of $\epsilon_{317\text{ nm}} 19,500\text{ M}^{-1}\text{ cm}^{-1}$ (Cu(II):**bipy-GS₂**), and $\epsilon_{335\text{ nm}} 13,400\text{ M}^{-1}\text{ cm}^{-1}$, $\epsilon_{347\text{ nm}} 15,300\text{ M}^{-1}\text{ cm}^{-1}$ (Cu(II):**terpy-GS₂**). Formation constants, $\log K_M$, were estimated to be 12.53 ± 0.04 for Cu(II):**bipy-GS₂** and two slightly different values of 8.04 ± 0.04 (A_{335}) and 7.14 ± 0.04 (A_{347}) were obtained for Cu(II):**terpy-GS₂**, depending on the wavelength monitored, see Table 2.3.

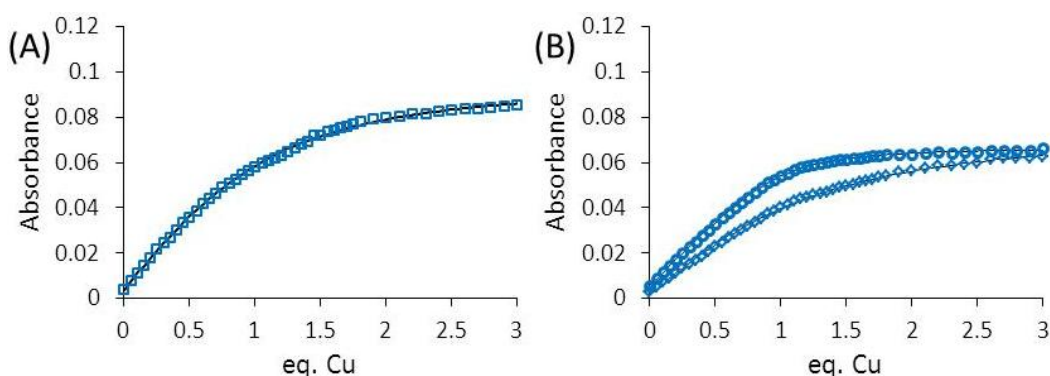


Figure 2.28 – Plot of absorbance for the Cu(II) titration of (A) **bipy-GS₂** monitored at 328 nm (\square), or (B) **terpy-GS₂** monitored at 335 (\circ) and 347 nm (\diamond), versus the equivalence of Cu(II). Line represents best fit for a 1:1 Cu(II):model switch binding ratio.

Table 2.3 – Summary of data obtained for Cu(II) coordination to model switches.

	λ (nm)	ϵ_{ML} ($\text{M}^{-1}\text{ cm}^{-1}$)	K_{app} (M^{-1})	K_M (M^{-1})	R^2
Bipy-GS₂	328	$1.95 \pm 0.01\text{ E} +04$	$6.73 \pm 0.66\text{ E} +05$	$3.39 \pm 0.34\text{ E} +12^a$	0.9987
Terpy-GS₂	335	$1.34 \pm 0.00\text{ E} +04$	$3.33 \pm 0.31\text{ E} +06$	$1.09 \pm 0.10\text{ E} +08$	0.9997
	347	$1.53 \pm 0.02\text{ E} +04$	$4.21 \pm 0.42\text{ E} +05$	$1.38 \pm 0.14\text{ E} +07$	0.9988

^a Titration performed with L-glycine (20 mM) as competitor in addition to phosphate buffer.

Cu(II) and Zn(II) titrations of **bipy-GS₂** and **terpy-GS₂** solutions, for which concentrations were based on UV profiles and using extinction coefficients determined for **cys₂bipy** and **cys₂terpy** respectively, all indicate 1:1 (conjugate:metal ion) binding ratio. These results demonstrate that the $\pi \rightarrow \pi^*$ transition is not sensitive to replacing the cysteine amino-acid with GSH, and that 1:1 complexes are formed under these experimental conditions.

The binding constants reported in Tables 2.2 and 2.3 are in good agreement with those reported previously for related ligands.[31,33,54,59,63,71] However, the affinity of **bipy-GS₂** for Zn(II) (measured at C = 5 μ M and pH 7.4) seems slightly higher than that for the analogous **cys₂bipy** (measured at C= 35.5 μ M and pH 8), but experimental conditions were different. Our model compounds, **bipy-GS₂** and **terpy-GS₂**, display higher affinity for Cu(II) than for Zn(II), consistent with the Irving-Williams series and previous reports,[59] and **bipy-GS₂** displays a higher affinity for both Cu(II) and Zn(II) than **terpy-GS₂**. The latter observation is consistent with lowering of the binding constant for **terpy-GS₂** resulting from the strain introduced by substitutions at position 6- and 6''-, as previously reported for polypyridine.[31,72] In the case of **terpy-GS₂**, fitting the data for the two $\pi \rightarrow \pi^*$ transitions as a function of Cu(II) concentration lead to different affinities, related by a factor of 10 (see Table 2.3). One can postulate that these could be due to different Cu(II) coordination environments (possibly arising from partial deprotonation of water bound to Cu(II)[73]), where two different Cu(II):**terpy-GS₂** complexes contribute differently to the absorbance at 335 and 347 nm.

The addition of CuCl₂ to solutions of **bipy-GS₂** and **terpy-GS₂** buffered at pH 7.4 was also accompanied by weaker transitions between 400-900 nm. Aliquots of a stock solution of CuCl₂ titrated into a more concentrated 350 μ M solution of **bipy-GS₂**, resulted in an increase in the absorbance at 622 nm up to 1.5 equivalents of Cu(II). Further addition of CuCl₂ led to only a small increase in the absorbance at 622 nm and an increase at 440 nm, consistent with addition of CuCl₂ to the blank buffered solution (see Figure 2.29A and C). An analogous titration of CuCl₂ into a 100 μ M solution of **terpy-GS₂** buffered at pH 7.4, resulted in an increase in the absorbance at 675 nm up to 2.0 equivalents of Cu(II), which blue-shifted slightly to 670 nm on addition of between 1.0 and 1.5 equivalents of Cu(II). No further changes were observed upon addition of between 2.0 and 3.0 equivalents of CuCl₂ (see Figure 2.29B and D).

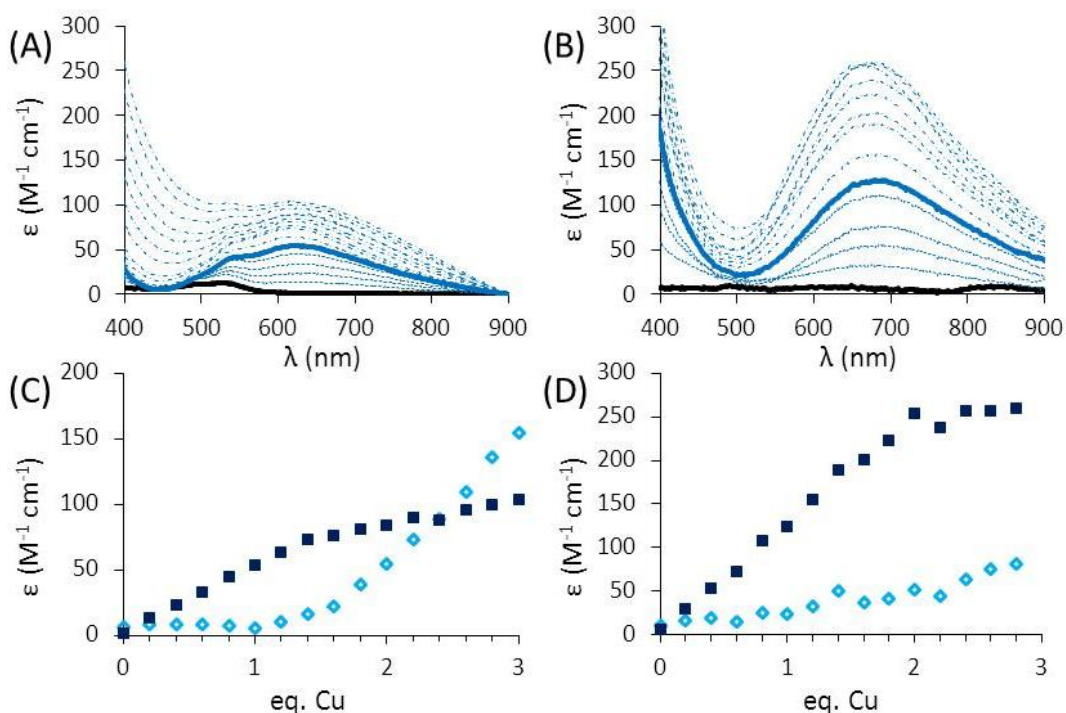


Figure 2.29 – UV spectra for the Cu(II) titration of model switches. CuCl_2 was added to solutions containing either 350 μM **bipy-GS₂** (A) or 100 μM **terpy-GS₂** (B) buffered at pH 7.4. Plot of absorbance monitored at 622 (■) and 440 nm (◇) for **bipy-GS₂** (C) and 670 (■) and 500 nm (◇) for **terpy-GS₂** (D), versus the equivalence of Cu(II).

Monitoring of the d-d band by visible spectroscopy (400-900 nm range), which appears upon coordination of Cu(II), indicates the final formation of a 1.5Cu(II):1**bipy-GS₂** and 2Cu(II):1**terpy-GS₂** complex (see Figures 2.29C and D). The d-d band for Cu(II) coordination to **bipy-GS₂** in a 1Cu(II):1**bipy-GS₂** ratio, (**bipy-GS₂**: $\lambda_{\text{max}} = 620 \text{ nm}$, $\epsilon_{620} = 51 \text{ M}^{-1} \text{ cm}^{-1}$) is consistent with previous reports for the formation of $[\text{Cu}^{\text{II}}(\text{bipy})(\text{OH})_2]$ ($\lambda_{\text{max}} = 620 \text{ nm}$, $\epsilon_{620} = 49 \text{ M}^{-1} \text{ cm}^{-1}$).^[65] In contrast, the d-d band for the Cu(II) coordination to **terpy-GS₂** in a 1:1 ratio (**terpy-GS₂**: $\lambda_{\text{max}} = 675 \text{ nm}$, $\epsilon_{675} = 131 \text{ M}^{-1} \text{ cm}^{-1}$) has an absorption coefficient higher than the values reported for the $[\text{Cu}^{\text{II}}(\text{terpy})]^{2+}$ cation at pH 4 ($\lambda_{\text{max}} = 680 \text{ nm}$, $\epsilon_{680} = 75 \text{ M}^{-1} \text{ cm}^{-1}$ ^[74,75]), but much lower compared to the $\pi\text{S} \rightarrow \text{Cu(II)}$ charge transfer bands characteristic of blue copper proteins ($\lambda_{\text{max}} = 625 \text{ nm}$, $\epsilon_{625} = 5,000 \text{ M}^{-1} \text{ cm}^{-1}$ ^[76]). However, calorimetric and EPR studies on $[\text{Cu}^{\text{II}}(\text{terpy})]^{2+}$ are consistent with a partial deprotonation of Cu(II)-bound water at physiological pH ($\text{p}K_a = 8.11$), resulting in a

tetrahedral distortion in the plane of $[\text{Cu}^{\text{II}}(\text{terpy})]^{2+}$.^[73] This distortion could be consistent with a larger absorptivity of the ligand field.^[77]

Small d-d band shifts are observed upon addition of between 1-1.5 eq. Cu(II) (**bipy-GS₂**) or 1-2 eq. Cu(II) (**terpy-GS₂**), consistent with only minor changes to the Cu ion coordination environment. Similarly the lack of contribution from the $\pi \rightarrow \pi^*$ on formation of the 1.5Cu(II):1**bipy-GS₂** and 2Cu(II):1**terpy-GS₂** complexes, is consistent with no significant change to the polypyridine coordination chemistry with respect to the analogous 1:1 complexes.

It is possible that the Cu(II) coordinates to the **bipy** or **terpy** ligand, and groups from the GSH units ca N- or C-termini, amino acid side chains or amide bonds, on formation of the initial 1:1 complexes.^[78-82] The formation of the 1.5Cu:1**bipy-GS₂** complex could involve the formation of a new intermolecular Cu(II) coordination site between two $[\text{Cu}^{\text{II}}(\text{bipy-GS}_2)]$ complexes (see Figure 2.30A). Whereas on formation of the 2Cu:1**terpy-GS₂** complex, the second Cu(II) could be coordinated exclusively by the GSH units (see Figure 2.30B).

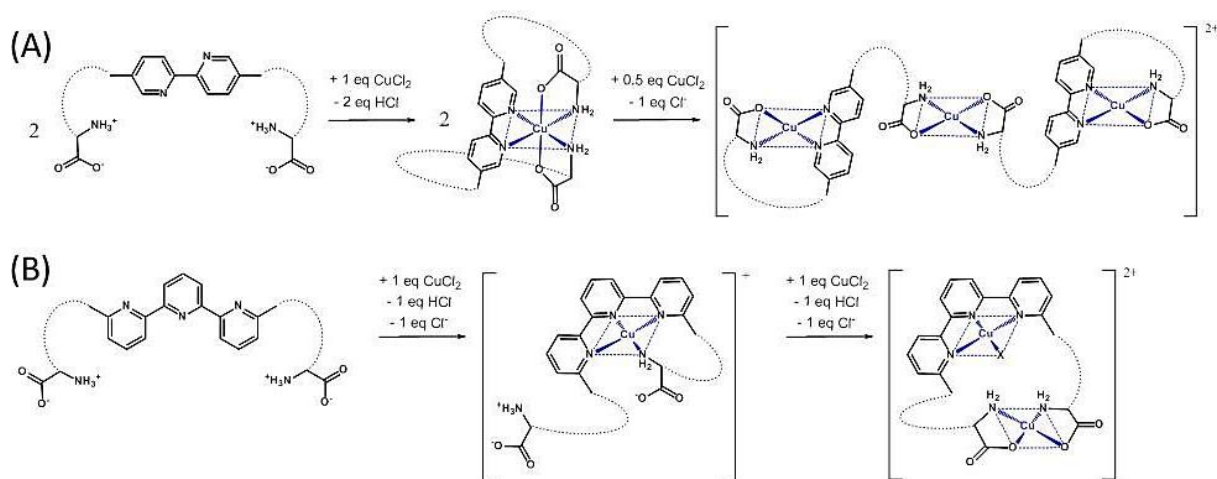


Figure 2.30 – Scheme depicting the hypothesised successive formation of complexes upon Cu(II) addition to either **bipy-GS₂** (A) or **terpy-GS₂** (B). X represents exogenous water, hydroxide, or chloride ligands.

One can expect that the formation of complexes where Cu(II) is surrounded by a glycine-like chelating motif would have a d-d transition in the region 600-650 nm

([Cu^{II}(glyglyNH₂-H)₂]: $\lambda_{\text{max}} = 650 \text{ nm}$, $\epsilon_{650} = 69 \text{ M}^{-1} \text{ cm}^{-1}$ [83], where glyglyNH₂ stands for amidated diglycine), and would therefore be undistinguishable from the d-d for Cu-polypyridine and mixed Cu-glycyl-polypyridine complexes ([Cu^{II}(**bipy**)(glyNH₂): $\lambda_{\text{max}} = 618 \text{ nm}$, $\epsilon_{618} = 60 \text{ M}^{-1} \text{ cm}^{-1}$), [84] where glyNH₂ stands for amidated glycine).

2.2.4.3 – Cu(II)/Zn(II) titration of **bipy**-GS₂ and **terpy**-GS₂ monitored by CD

The CD spectra of 350 μM solutions of **bipy**-GS₂ and **terpy**-GS₂, recorded from 400 to 200 nm, did not display any notable signal. However, the addition of increasing concentrations of ZnCl₂ to the solution of **bipy**-GS₂ at pH 7.4 led to the appearance of new positive transitions centred at 220, 241, 310 and 320 nm, as well as a negative transition at 266 nm, with isosbestic points at 251 and 285 nm (Figure 2.31A). In contrast, addition of ZnCl₂ to **terpy**-GS₂ resulted in negative transitions at 228, 329, and 340 nm and positive transitions at 283 and 290 nm, with isosbestic points at 270 and 298 nm (Figure 2.31B). In both cases, addition of excess ethylenediamine tetraacetic acid (EDTA), 20 equivalence with respect to metal, resulted in CD spectra which were in good agreement with those of **bipy**-GS₂ and **terpy**-GS₂ recorded in the absence of Zn(II).

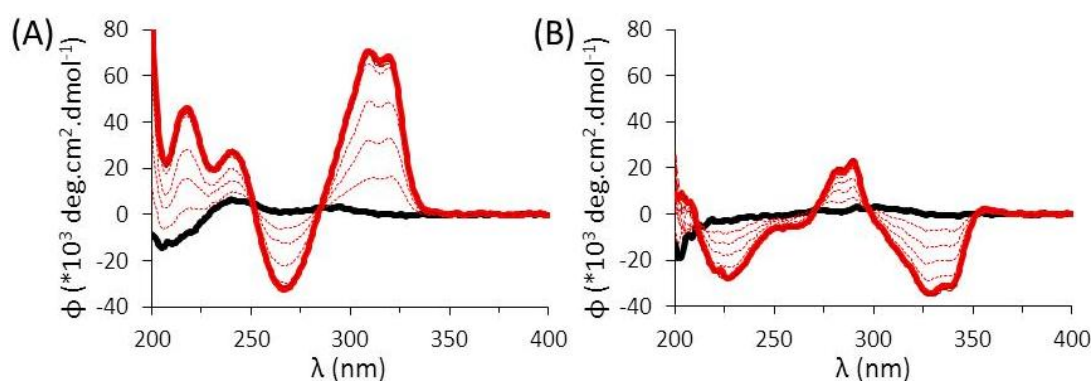


Figure 2.31 – CD spectra for Zn(II) titration of model switches. ZnCl₂ titration into 350 μM of either (A) **bipy**-GS₂ or (B) **terpy**-GS₂ in 10 mM phosphate buffer pH 7.4. (—) 0 equivalent, (....) between 0 and 1 eq., (—) 1 eq., (---) more than 1 eq. ZnCl₂.

Plots of molar ellipticity for the maximum as a function of ZnCl_2 equivalence reaches a plateau at ca. 0.9 (**bipy-GS₂**) and 1.0 equivalents (**terpy-GS₂**) of Zn(II) per model switch (Figure 2.32).

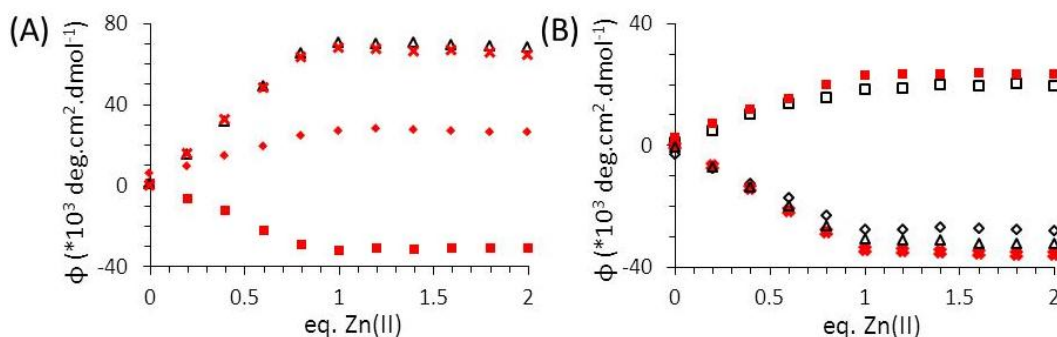


Figure 2.32 – Plot of the molar ellipticity as a function of the equivalence of Zn(II) for model switches titration: (A) **bipy-GS₂** monitored at 241 (♦), 266 (■), 310 (Δ) and 320 nm (X); (B) **terpy-GS₂** monitored at 228 (◇), 283 (□), 290 (■), 329 (X) and 340 nm (Δ).

Similarly, the addition of CuCl_2 to **bipy-GS₂** (see Figure 2.33A) resulted in the appearance of positive transitions at 247 and 320 nm, and a negative transition centred at 278 nm. The peak intensities increase up to 1.0 equivalent of Cu(II) , with clear isosbestic points at 265 and 292 nm. Further addition of CuCl_2 resulted in a gradual shift of the positive transitions toward 251 and 310 nm, and the negative transition toward 282 nm. All signals decreased in intensity, reaching a minimum on addition of 1.5 equivalents of CuCl_2 . No further spectral changes occur on addition of up to 3.0 equivalents CuCl_2 (see Figure 2.33A). The titration of increasing concentrations of CuCl_2 into a 350 μM solution of **terpy-GS₂** at pH 7.4 resulted in the appearance of a negative transition centred at 215 nm. Two overlapping negative transitions at 335 and 347 nm appear on addition of between 1.0 and 2.0 equivalent of CuCl_2 . No further spectral changes occur on addition of up to 3.0 equivalents CuCl_2 (see Figure 2.33B). In both cases, addition of excess EDTA (20 eq. with respect to metal) resulted in CD spectra which were in good agreement with those of **bipy-GS₂** and **terpy-GS₂** recorded in the absence of Cu(II) .

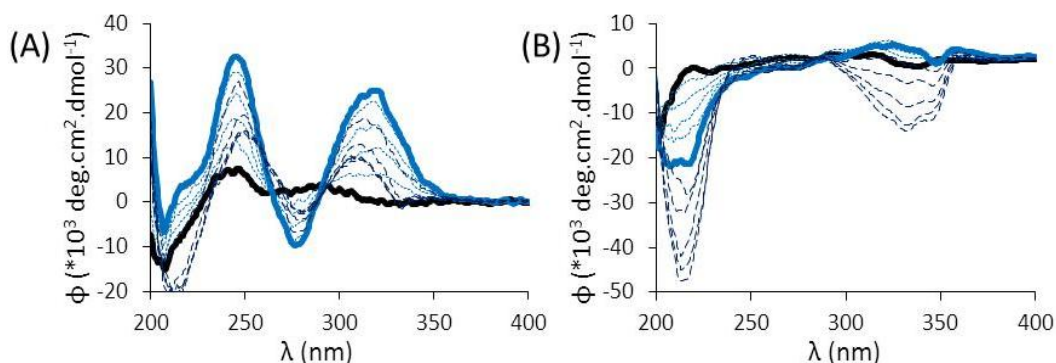


Figure 2.33 – CD spectra for Cu(II) titration of model switches. CuCl₂ titration into 350 μM of either A) **bipy-GS₂** or B) **terpy-GS₂** in 10 mM phosphate buffer pH 7.4. (—) 0 eq, (---) between 0 and 1 eq., (—) 1 eq., (—) between 1 and 2 eq. CuCl₂ added.

Plots of molar ellipticity as a function of CuCl₂ concentration reach plateaus at ca 1.5 equivalents (**bipy-GS₂**), or 2.0 equivalents (**terpy-GS₂**) of Cu(II) per model switch (see Figure 2.34).

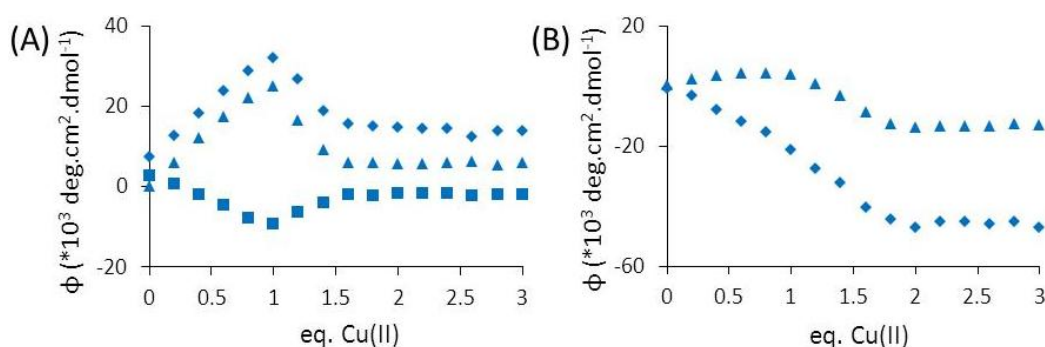


Figure 2.34 – Plot of the molar ellipticity as a function of the equivalence of Cu(II) for model switches titration: A) **bipy-GS₂** monitored at 247 (♦), 278 (■) and 320 nm (▲); B) **terpy-GS₂** monitored at 215 (♦) and 334 nm (■).

The relatively featureless CD spectra of the model switches are altered dramatically upon metal coordination. Titrations of **bipy-GS₂** and **terpy-GS₂**, display chiral induced CD signals relative to the **bipy** and **terpy** $\pi \rightarrow \pi^*$ bands (250–400 nm range). In contrast, similar Cu(II)/Zn(II) titration of 350 μM solution of **pyr-GS₂** did not result in intensity increase of $\pi \rightarrow \pi^*$ band. The cotton effect induced upon Zn(II) addition is opposite for **terpy-GS₂** and **bipy-GS₂** (see Figure 2.31). This indicates that the lower energy transition for [Zn(**terpy-GS₂**)X_n]^{m+} might arise from $\pi \rightarrow \pi^*$ electronic absorptions involving molecular orbitals

composed mainly from atomic orbitals from atoms composing the central pyridine ring, as previously suggested.[41] In contrast, the second lower energy transition (centred at 287 nm) for $[\text{Zn}^{\text{II}}(\text{terpy-GS}_2)\text{X}_n]^{\text{m}+}$ could involve orbitals comprising contributions mainly from the external pyridines, and is therefore more similar to the lower energetic absorption band for $[\text{Zn}^{\text{II}}(\text{bipy-GS}_2)\text{X}_n]^{\text{m}+}$. These observations are consistent with UV studies and formation of a 1:1 complex with the Zn(II) coordinated to the intended polypyridine chelate of both **bipy-GS₂** and **terpy-GS₂**.

Metal complexation of **bipy** or **terpy** results in polarisation of the $\pi \rightarrow \pi^*$ transitions, and the appearance of induced CD signals.[85] Octahedral complexes bearing a single bidentate π -conjugated ligand, such as $\text{Co}(\text{phen})(\text{en})_2$ (**en** = ethylenediamine and **phen** = 1,10-phenanthroline), usually display weak $\pi \rightarrow \pi^*$ signals in the CD, compared to $\text{M}^*(\text{phen})_2$ or $\text{M}^*(\text{phen})_3$ complexes,[86,87] for which the metal ions are chiral centres (see Figure 2.35A). As a significant exception, a previously reported **Me₂bipy** bearing acidic peptide substituents displays more intense CD signal for the 1:1 compared to the 1:3 Fe(II):**Me₂bipy** complex.[88] In this work, slow addition of FeCl_2 , which is expected to form octahedral $\text{Fe}(\text{bipy})_3$ complexes with **bipy** at physiological pH,[88-90] to a 350 μM solution of **bipy-GS₂** resulted in precipitation of the conjugate, and no chiral induced $\pi \rightarrow \pi^*$ were observed. Hence, the **3bipy-GS₂:1M** complexes might have low solubility under these experimental conditions. The intensity of CD $\pi \rightarrow \pi^*$ induced signals for **bipy-GS₂** and **terpy-GS₂** metal complexes could be explained by additional coordination through uncapped terminal amines, and sensing of the chirality at $\text{C}\alpha$ from peptide units. To the best of our knowledge, no CD study was reported for complexes of the type $\text{Co}(\text{phen})(\text{en})_2$, where **en**-like units are covalently attached to **phen** or **bipy** through one or several chiral center(s), similarly to **bipy-GS₂**. If one envision CD spectra of metal complexes including one **bipy** and two **en** ligands, differing by presence of chiral centers or covalent bonding between the different ligands (compare Figures 2.35B,

C, and F), it is reasonable to assume that the intensity of the CD induced $\pi \rightarrow \pi^*$ signals, could follow the order $M^*(\text{bipy})(\text{en})_2 < M^*(\text{bipy})(\text{en}^*)_2 < M^*[\kappa^6\text{-(bipy-en)}^{**}]$.

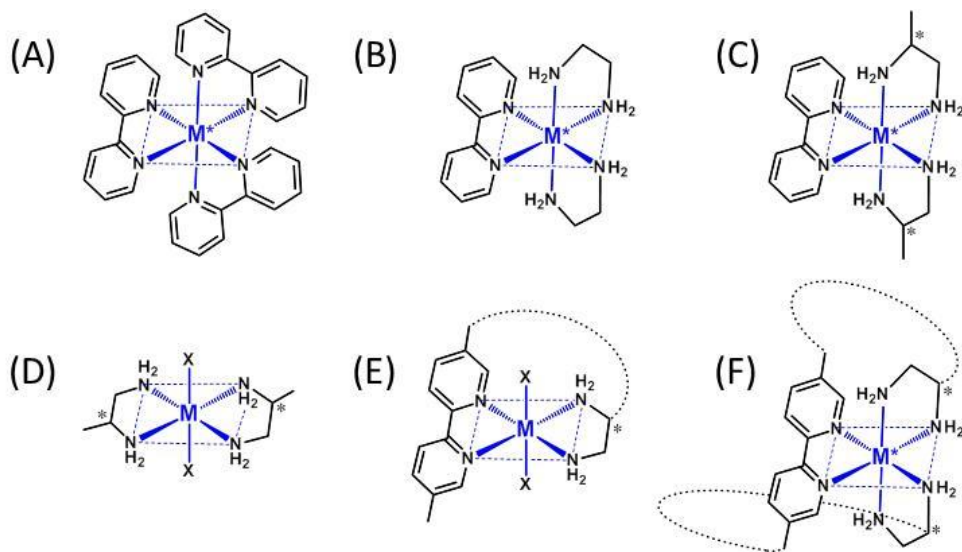


Figure 2.35 – Scheme depicting octahedral mixed complexes with **bipy** and **en** mixed ligand. X represents exogenous water or hydroxide, * indicate chiral centres, and (---) depict covalent bonding. (A) $M^*(\text{bipy})_3$, (B) $M^*(\text{bipy})(\text{en})_2$, (C) $M^*(\text{bipy})(\text{en}^*)_2$, (D) $M(\text{en}^*)_2\text{X}_2$, (E) $M[\kappa^4\text{-(bipy-en)}^*]\text{X}_2$, and (F) $M^*[\kappa^6\text{-(bipy-en}_2\text{)}^{**}]$. For (C), (D) and (F), additional isomers arise from the relative positions of the **en** substituents.

The shift in the **bipy-GS₂** and **terpy-GS₂** $\pi \rightarrow \pi^*$ band (200-400 nm range) in the UV spectra upon addition of Cu(II), is also consistent with a ca. 1:1 ratio in all cases (see Figure 2.28). However, CD spectra suggest that the complexation of Cu(II) is more complicated, and involves the formation of two different Cu(II) complexes with differing contributions to the metal-bound $\pi \rightarrow \pi^*$ bands. This is not obvious in the UV spectra, but is observed for Cu(II) complexation to both **bipy-GS₂** and **terpy-GS₂** by CD, due to exciton effects.[91-94] The CD titration of Cu(II) into **bipy-GS₂**, results in chiral induced signals relative to the **bipy** $\pi \rightarrow \pi^*$ band (250-400 nm range) up to one equivalent consistent with the formation of a 1:1 complex involving coordination through the pyridine units. However, the CD titration indicates that this is followed by the formation of a 1.5Cu(II):1**bipy-GS₂** complex, as a result of a reduction in these induced CD signals. In contrast, the analogous CD titration with **terpy-GS₂** did not result in the formation of induced CD signals relative to the **bipy** $\pi \rightarrow \pi^*$ band up to one

equivalent of Cu(II). However, it is consistent with the initial formation of a 1:1 complex, followed by a 2Cu(II):1**terpy**-GS₂ complex (see Figure 2.34).

The successive events observed by CD upon Cu(II) addition to **bipy**-GS₂ or **terpy**-GS₂, are consistent with the growing of d-d bands monitored by UV-vis (see Figure 2.29), and can be explained by the successive formation of different Cu(II) complexes hypothesised (see Figure 2.30). Following the proposed order of Cu(II) complexation events, addition of Cu(II) would first result in the formation of a complex type Cu^{II}[κ⁶-(**bipy**-en₂)**], followed by reorganisation into Cu^{II}[κ⁴-(**bipy**-en)*X₂] and Cu^{II}(en*)₂X₂, hence the decrease in signal intensities observed (see Figures 2.35D, E, and F). The weak exciton coupling for Cu^{II}[κ⁶-(**bipy**-en₂)**] compared to Zn^{II}[κ⁶-(**bipy**-en₂)**] might result from the Jahn-Teller distortion.[95-97] It is more difficult to interpret observations for **terpy**-GS₂, as the different π→π* contribution and chiral induced signals of **terpy** have been less extensively studied and reported in the literature.

2.2.4.3 – Zn(II) titration of **bipy**-GS₂ and **terpy**-GS₂ monitored by NMR

The ¹H NMR spectrum of a 5 mM solution of **bipy**-GS₂ in D₂O at pD 1, displays two singlets at 8.83 and 8.42 ppm in the aromatic region, integrating to two and four protons, respectively. In contrast, one singlet at 8.59 ppm and two overlapping doublets at 8.04 and 8.00 (AB pattern), each integrating to 2 protons, are observed on raising the pD to 7.4 (see Figure 2.36A). A titration of ZnCl₂ into a 5 mM solution of **bipy**-GS₂ in D₂O buffered at pD 7.4, results in the broadening and decrease in intensity of the peaks at 8.59, 8.04 and 8.00 ppm, and the appearance of new peaks at 8.66 (broadened, H⁶), 8.42 (doublet, H³) and 8.20 (doublet, H⁴) ppm on addition of 1.0 equivalent of ZnCl₂ (see Figure 2.36B). A plot of peak integration of the overlapping doublets (8.04 and 8.00 ppm) as a function of equivalence of ZnCl₂ indicates a 1:2 Zn:**bipy**-GS₂ ratio. Similarly a plot of peak integration for the resulting

doublet for the Zn-**bipy-GS₂** adduct at 8.20 ppm, also plateaus at 0.5 equivalence ZnCl₂ consistent with a 1:2 Zn:**bipy-GS₂** ratio (see Figure 2.36C). After 0.5 equivalents of ZnCl₂ have been added the peaks at 8.59, 8.04 and 8.00 ppm appear to have been replaced with broad new peaks at 8.42 and 8.20 ppm. Upon addition of between 0.5 and 1.0 eq. ZnCl₂ these peaks sharpen into doublets and a broad peak attributed to H⁶ appears at higher frequency (8.66 ppm). Only very small changes are observed on addition of between 1.0 and 2.0 equivalents of ZnCl₂ (see Figure 2.36B). Addition of excess EDTA (20 eq. with respect to ZnCl₂) resulted in a ¹H NMR spectrum which is in good agreement with that of **bipy-GS₂** recorded in the absence of ZnCl₂ (see Figure 2.36B).

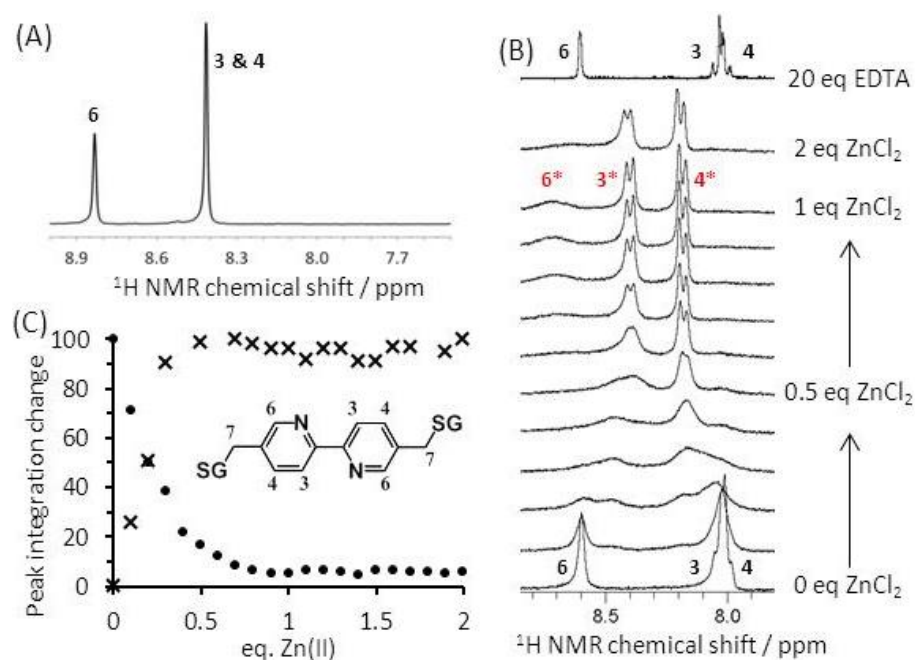


Figure 2.36 – (A) ¹H NMR spectra of 5 mM **bipy-GS₂** recorded at pD ~ 1, (B) ¹H NMR Zn(II) titration of 5 mM **bipy-GS₂** in solution buffered at pD 7.4. (C) Plot of change in percentage peak integration as a function of equivalence of ZnCl₂ for 8.04–8.00 ppm resonances for H³ and H⁴ of the free model switch, **bipy-GS₂** (●), as well as for the new peak at 8.20 ppm for H⁴ of the Zn-**bipy-GS₂** adduct (X). All spectra were recorded at 300 MHz and 293 K.

At acidic pD (ca. 1) the ¹H NMR spectrum of **terpy-GS₂** recorded in D₂O displays 5 aromatic signals (all doublets of doublets) at 8.68 (2H^{3a,b}), 8.67 (2H^{4a,b}), 8.59 (2H³), 8.47 (1H⁴) and 8.17 ppm (2H^{5a,b}), which were assigned using COSY and NOESY NMR (see

Figure 2.37). The COSY spectrum displays cross-peaks between $H^3 \leftrightarrow H^4$ and $H^{4a} \leftrightarrow H^{5a}$ (n.b. H^{3a} and H^{4a} are too close, so the cross-coupling overlaps with the diagonal peaks). In contrast, NOESY NMR recorded under similar conditions, displays an additional $H^{3a} \leftrightarrow H^3$ inter-ring coupling (see Figure 2.37B).

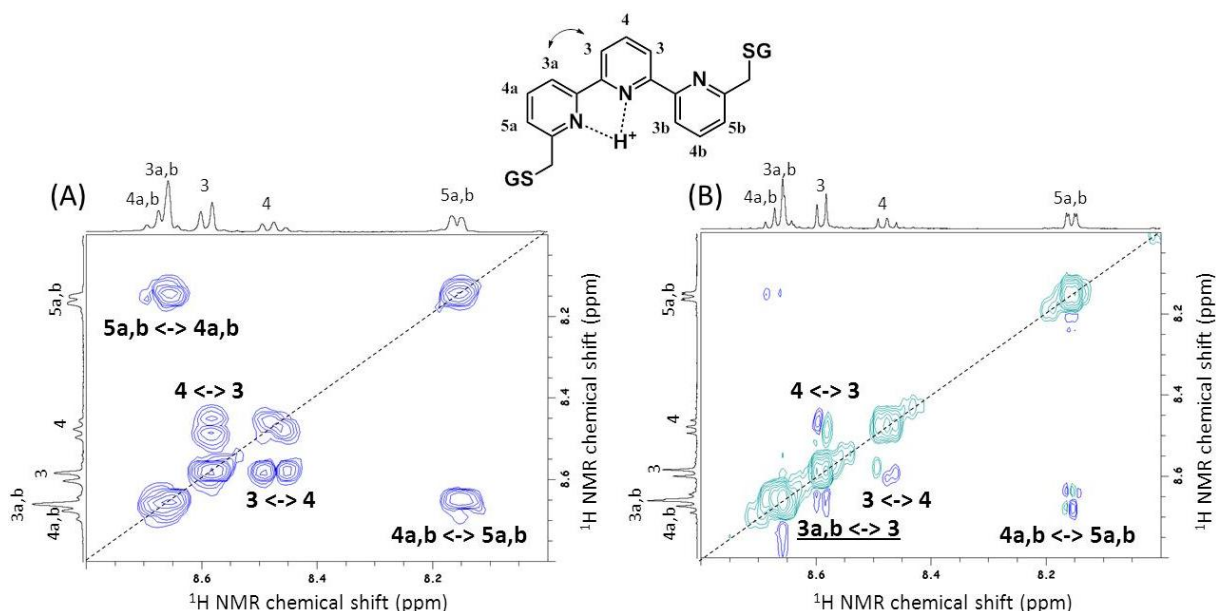


Figure 2.37 – 2D NMR of **terpy-GS₂** 9 mM solution in D₂O (pD ~1, 300 K). (A) COSY spectrum recorded at 400 MHz; (B) NOESY spectrum recorded at 500 MHz, 450 ms mixing time.

On raising the pD to 7.4, three aromatic resonances are observed at 8.08 ($1H^4$, $2H^3$, $2H^{3a}$), 7.96 ($2H^{4a}$) and 7.55 ppm ($2H^{5a}$). A titration of ZnCl₂ into a 5 mM solution of **terpy-GS₂** in D₂O buffered at pD 7.4, resulted in the decrease in intensity of the peaks at 8.08, 7.96 and 7.55 ppm, and the appearance of new broad peaks at 8.31, 8.13 and 7.65 ppm. This is accompanied by a decrease in the intensity of the singlet at 3.99 ppm assigned to the CH₂-pyridinyl group. A plot of the peak integration for the singlet at 3.99 ppm, as a function of ZnCl₂ concentration (Figure 2.38C), is consistent with formation of a 1:1 complex between Zn(II) and **terpy-GS₂**. Addition of excess EDTA (20 eq. with respect to ZnCl₂) resulted in a ¹H NMR spectrum which is in good agreement with that of **terpy-GS₂** recorded in the absence of ZnCl₂ (see Figure 2.38B).

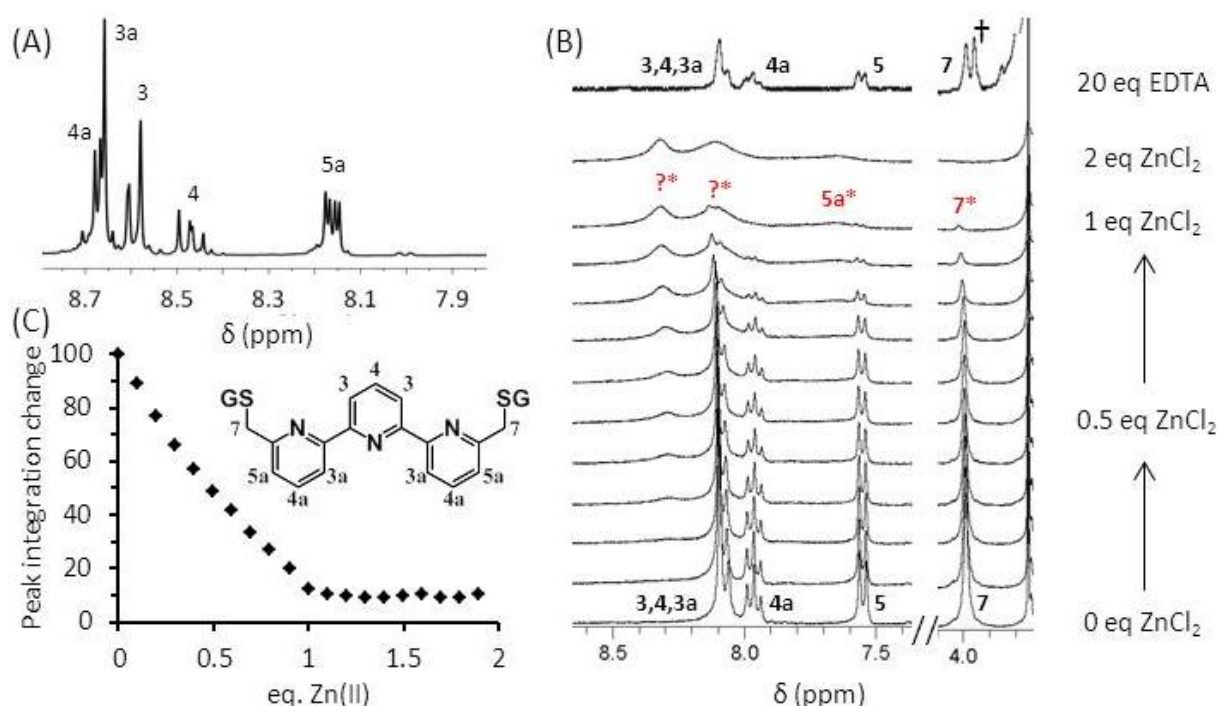


Figure 2.38 – (A) ^1H NMR spectra of 9 mM **terpy-GS₂** recorded at pH ~ 1 (500 MHz, 298 K), (B) ^1H NMR Zn(II) titration of 5 mM **terpy-GS₂** in solution buffered at pH 7.4 (300 MHz, 293 K). (C) Plot of change in percentage peak integration as a function of equivalence of ZnCl₂ for **terpy-GS₂** methyl singlet (♦) at 3.99 ppm. The peak at 3.96 ppm labelled with †, was attributed to the ^{13}C satellite relative to the -N-CH₂-COOH signal of EDTA centred at 3.72 ppm.

The ^1H NMR spectrum of **bipy-GS₂**, recorded under acidic conditions is very different from that recorded at neutral pH. A single resonance, attributed to H³ and H⁴ of **bipy-GS₂**, is observed at pH 1, however, an AB pattern where the two overlapping doublets are located at a lower chemical shift, is observed at pH 7.4. Both spectra are different from those of **Me₂bipy** recorded in DMSO-*d*₆ (see Figure 2.15A), probably due to the strong second order coupling. Addition of ZnCl₂ to a 5 mM solution of **bipy-GS₂** at pH 7.4 results in new broad peaks which indicates that **bipy-GS₂** and the complex(es) it forms with ZnCl₂ are in slow/intermediary exchange(s) on the NMR time-scale. A plot of peak integration as a function of Zn(II) equivalence is consistent with either the formation of a 2**bipy-GS₂**:1Zn complex, or an equilibrium involving 3**bipy-GS₂**:1Zn, 2**bipy-GS₂**:1Zn, and 1**bipy-GS₂**:1Zn complexes. The spectrum at 0.5 equivalents Zn(II) does not display any signal assigned to H⁶, most likely due to signal broadening as a result of H⁶ proximity with the chloride ligands. However a broad resonance assigned to H⁶ reappears on addition of more ZnCl₂ (between 0.5

and 1 equivalent) and sharpens in the presence of excess Zn(II) (5 and 10 equivalents), which may be consistent with conversion of the 2**bipy**-GS₂:1Zn(II) complex into a 1**bipy**-GS₂:1Zn(II) complex.[51,98]

Theoretical studies suggest that even though monoprotonated **bipy** and bidentate metal complexes of **bipy** have energy minima with similar conformations (*cis*), flexibility around the axial bond of **bipy** is much higher in the monoprotonated **bipy** compared to the metal complexes.[99] In fact, the difference in potential energy separating *cis*- and *trans*-conformations has been reported to be comparable for the monoprotonated and free **bipy**. [67,68,99] This could account for the similarity of resonances assigned as H³ and H⁴ in spectra of **bipy**-GS₂ recorded at acidic and neutral pD, which in turn differ from those recorded for the Zn(II) complex (see Figure 2.39).

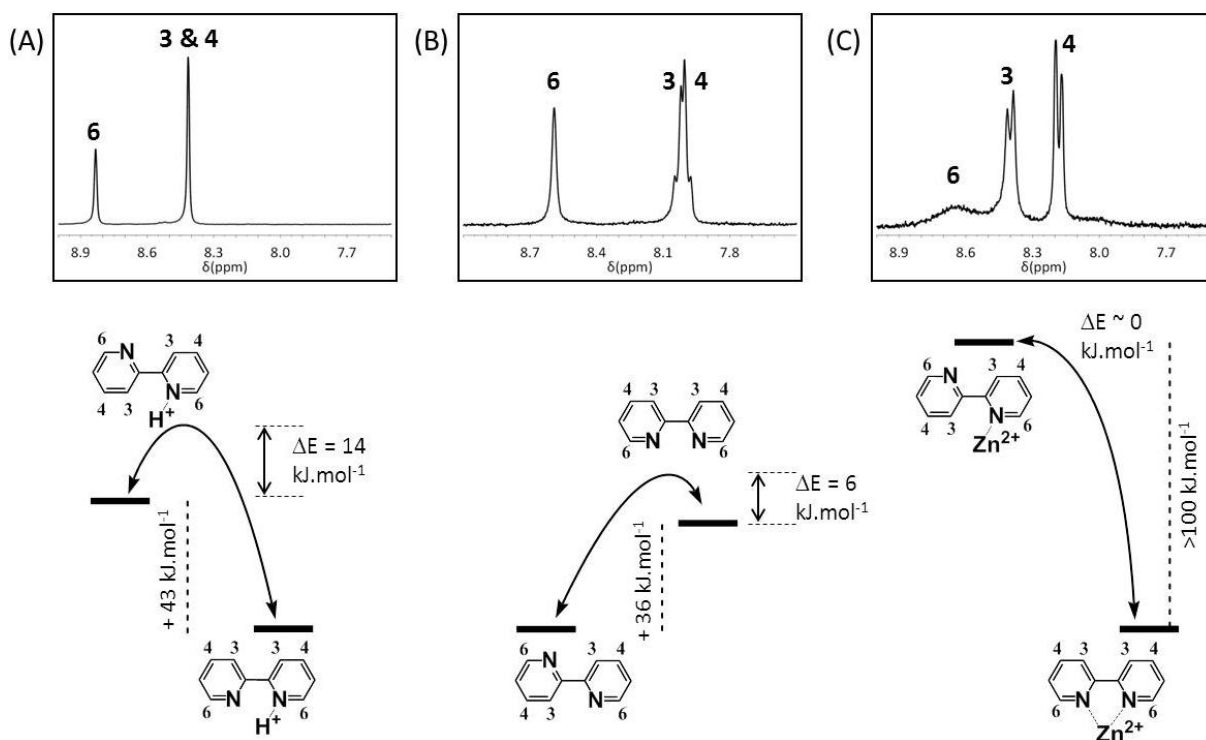


Figure 2.39 – Illustration of the relationship between energy profiles and ¹H NMR spectra for **bipy**-GS₂. ¹H NMR spectra of **bipy**-GS₂ are more greatly influenced by conformational flexibility rather than the conformation of the lowest energy structure. Representation of energy profiles regulating the population of the different conformers of the (A) monoprotonated **bipy**-GS₂, (B) **bipy**-GS₂ and (C) when complexed to Zn(II), alongside ¹H NMR spectra (D₂O, 300 MHz, 293 K). Energy profiles and activation barriers are based on *ab initio* reports for the *cis-trans* interconversion of **bipy**. [67,69,99] The impact of substitutions at position 5 and 5' on the rotation of the **bipy** can be ignored.[4]

NMR spectra of **bipy** or derivatives where H^3 and H^4 resonances overlap have previously been recorded in aqueous solution at both acidic[3,51] and physiological pH,[51,52] however, this is not exclusively the case.[3] Interpretation of **bipy** conformation based on NMR chemical shift can therefore lead to contradictory results. For example, theoretical studies suggest that H^3 are deshielded in the *trans*- (cation free) **bipy** due to the close proximity with the nitrogen on the second ring.[4] Cation binding to **bipy** is expected to result in deshielding of aromatic resonances. In contrast, H^3 can display only a moderate deshielding,[3] or shielding[100] upon *trans*-to-*cis* conformational transition, as deshielding effects on H^3 from the proximal nitrogen are lost.[50,51] Therefore an attractive method by which to assign the **bipy-GS₂** conformation is by monitoring intra-ring coupling by NOESY NMR,[22] however, this is only possible for an asymmetric **bipy** for which H^3 peaks are inequivalent, and so cannot be applied to **bipy-GS₂**.

In contrast, NOESY can be applied to **terpy-GS₂**. At pD 1, the cross-peak observed between H^3 and H^{3a} in the NOESY spectrum is consistent with at least half of the terpyridine adopting a *cis*- conformation (see Figure 2.37). The 1H NMR profile of **terpy-GS₂** at pD 1 is different for those recorded for **Me₂terpy** in DMSO- d_6 , both in the absence or presence of Zn(II). The 1H NMR spectrum of **terpy-GS₂** recorded at pD 7.4 is different from that recorded at acidic pD, and is not suitable for determination of intra-ring coupling as the resonances for H^3 and H^{3a} overlap. Upon raising the pD from 1 to 7.4, all aromatic signals move to lower frequency, and proton signals relative to external pyridine (H^{3a} , H^{4a} and H^{5a}) experience higher shielding than their counterpart (H^3 and H^4), consistent with a transition from mixed *cis-trans*- conformation to a *trans-trans*- conformation.[11] The 1H NMR profile of **terpy-GS₂** at pD 7.4 resembles that recorded for **Me₂terpy** in DMSO- d_6 , with additional strong coupling (compare Figure 2.16 and 2.38). Addition of $ZnCl_2$ lead to formation of broad resonances for a 1**terpy-GS₂**:1Zn complex in slow/intermediary exchange on the NMR timescale. Signals are generally broad and difficult to assign, but integration of the three

signals indicates that the spectrum is different from that recorded for **terpy-GS₂** both at acidic and neutral pD, and could be consistent with the *cis-cis* conformation of a terpyridine metal complex.[63,54]

Titration of Zn(II) into a 5 mM solution of **terpy-GS₂** monitored by ¹H NMR, is consistent with a 1:1 binding ratio (see Figure 2.38C). In contrast, the titration of Zn(II) into a 5 mM solution of **bipy-GS₂** indicates the formation of a **2bipy-GS₂:1Zn(II)** complex (see Figure 2.36C).[51,98] This is in stark contrast to the analogous titrations (CD, UV) recorded under more dilute (14 times and 1000 times respectively) and biologically relevant conditions, suggesting that the formation of a 2:1 complex with **bipy-GS₂** only occurs at high concentrations, and that the 1:1 [Zn^{II}(**bipy-GS₂**)X_n]^{m+} species dominates under more dilute conditions.

2.3 – Summary

This chapter reports the dimerisation of the single cysteine amino acid as well as a short cysteine containing tripeptide through (poly)pyridine linkers, and the study of the resulting species in terms of conformation, and Cu(II)/Zn(II) affinity. Both cysteine and GSH were dimerised through a 2,6-pyridinyl linker, affording **cys₂pyr** and **pyr-GS₂** conjugates, respectively. After purification, **cys₂pyr** could be obtained as a pure solid; however, **pyr-GS₂** was recovered as a salt. Both conjugates only interact weakly with Cu(II) and not at all with Zn(II). Cu(II) titrations indicate a binding ratio of 1:1 **cys₂pyr**:Cu(II), and 1:0.6 **pyr-GS₂**:Cu(II). The latter ratio was in fact demonstrated to correspond to 1:1 **pyr-GS₂**:Cu(II), when the mass of salt was taken into consideration and the experimental molecular weight corrected accordingly.

Following the pyridine work, **Me₂bipy** (allosteric ineffective) and **Me₂terpy** (allosteric effective) linkers were prepared. Zn(II) titrations monitored by ¹H NMR, indicate formation of complexes with 1:1 Zn(II):linker stoichiometry. Additionally, 2D NMR studies indicate

that the **Me₂terpy** linker undergoes *trans*-to-*cis* conformational transitions upon Zn(II) coordination.

As this was the proposed strategy by which to dimerise larger peptides or protein fragments, cysteine was dimerised through **Me₂bipy** and **Me₂terpy** linkers, affording **cys₂bipy** and **cys₂terpy**. These conjugates both displayed limited solubility at physiological pH, incompatible with NMR studies, but could be solubilised upon addition of acid.

Following this, analogous conjugates based on GSH, **bipy-GS₂** and **terpy-GS₂**, were prepared and displayed higher solubility at physiological pH than their cysteine analogues. Extinction coefficients estimated for **cys₂bipy** and **cys₂terpy** $\pi \rightarrow \pi^*$ bands, were used for quantification of **bipy-GS₂** and **terpy-GS₂**, respectively. Their Cu(II) and Zn(II) coordination ability were evaluated in different concentration ranges and monitored by UV-visible, CD and NMR (Zn(II) only) spectroscopy. Titrations indicated that **terpy-GS₂** binds Cu(II) and Zn(II) in a 1:1 ratio. In contrast, Zn(II) titration of 5 mM **bipy-GS₂** monitored by ¹H NMR indicate either the formation of a **2bipy-GS₂:1Zn(II)** complex or an equilibrium between different species, whereas Cu(II)/Zn(II) titration of solution containing 350 μ M and 5 μ M **bipy-GS₂** are in good agreement with formation of **1bipy-GS₂:1Zn(II)** complex under dilute conditions. Unlike for Zn(II), Cu(II) titrations monitored by CD indicate formation of 1:1 complexes followed by **1.5Cu:1bipy-GS₂** and **2:1terpy-GS₂** complexes. Moreover, monitoring of the Cu(II) d-d transitions also agrees with the formation of **1.5Cu:1bipy-GS₂** and **2Cu:1terpy-GS₂** complexes. The different ratio are interpreted as a result of the reorganisation of Cu(II) coordination spheres.

Cu(II) and Zn(II) titration at physiological pH were performed at low concentration and monitored by UV spectroscopy, indicating high affinity of the conjugates for both metal ions ($\log K_M > 6$). Cu(II) binding constants were found to be higher than those obtained for Zn(II), and the sterically less strained conjugates (5,5'-substitutions) form more stable complexes with both metal ions. The following chapter will focus on the preparation of conjugates where

these polypyridine regulation sites are used to dimerise larger peptide moieties, for which their spatial alignment is important for biomolecular recognition.

2.4 – Experimental section

2.4.1 – Reagents and equipment

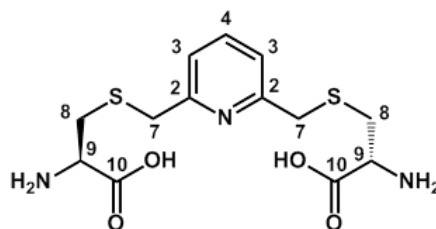
NaHCO₃, Na₂SO₄, NaNO₂, NaOH, KOH, mono- and dihydrogen potassium salts, tris base, ethylene diamine tetra acetic acid (EDTA), chloroform (CHCl₃), methanol (CH₃OH), ethanol, tetrahydrofuran (THF), toluene, dichloromethane (CH₂Cl₂), acetonitrile, dioxane, water (HPLC grade), hydrobromic acid 48 %w, HCl 32 %w, ammonia 35 %w, bromine, were all obtained from Fisher Scientific. 2,6-Bis(bromomethyl)pyridine, 5,5'-dimethyl-2,2'-bipyridine, N-butyllithium 1.6 M in hexane, diethylether, copper chloride (CuCl₂), L-glycine, and cysteine were obtained from Sigma-Aldrich. 2-Bromopicoline, 2,6-dibromopyridine, tetrakis-(triphenylphosphine) palladium, trifluoroacetic acid (TFA), tributyltin chloride, zinc chloride (ZnCl₂) were obtained from BOC. N-Bromosuccinimide was obtained from Alfa-Aesar. 5,5'-Dithiobis-(2-nitrobenzoic acid) and glutathione (> 97 % pure by HPLC, as a sum of enantiomers) were obtained from Fluka. TCEP.HCl was obtained from Thermo Scientific. Deuterated solvents (CDCl₃, DMSO-d₆ and D₂O) were obtained from Cambridge Isotope Laboratory Inc. Dimethylformamide (DMF) was obtained from AGTC Bioproducts.

HRES and ES-TOF MS were recorded on a Microwaters LCT TOF spectrometer equipped with a 3000 V capillary voltage, and a cone voltage of 35 V. GC-MS were recorded on a Waters GCT Premier Micromass equipped with an EI probe. Analytical RP-HPLC traces were recorded on C18 column using a 0 to 100 % gradient acetonitrile in water over 40 minutes (either containing 0.05% TFA or not) and the absorption was monitored at 220 nm. The compound purity is reported as a percentage of its peak integral over the total integration for all peaks present between 0 and 40 minutes. UV-visible spectra were recorded in a 1 cm pathlength quartz cuvette at 298 K on either a CARY50 (spectrometer(a)) or Shimadzu 1800

(spectrometer(b)). Wavelengths (λ) are given in nm, and extinction coefficient (ϵ) in $\text{M}^{-1} \text{cm}^{-1}$. CD spectra were recorded in 1 mm pathlength quartz cuvettes at 298 K on a Jasco J-715 spectropolarimeter. The observed ellipticities in millidegrees were converted into molar ellipticity, $[\Phi]$, and are reported in units of $\text{deg dmol}^{-1} \text{cm}^2$. All ^1H and ^{13}C NMR spectra were collected on either a Bruker DRX500 (500 MHz ^1H and 125 MHz ^{13}C , Gradient NOESY, and Phase cycle ROESY, $T = 300 \text{ K}$), AVIII400 (400 MHz ^1H , 100 MHz ^{13}C , DQF-COSY, and Gradient NOESY, $T = 298 \text{ K}$) or AVIII300 (300 MHz ^1H , $T = 293 \text{ K}$) spectrometer equipped with a 5 mm probe. Solvent residual signal, DMSO quintet ($\delta = 2.50 \text{ ppm}$), CHCl_3 singlet ($\delta = 7.26 \text{ ppm}$), were used as an internal reference. As an exception, dioxane was added to the tube and used as the internal standard ($\delta = 3.75 \text{ ppm}$), for all experiments recorded in D_2O . For pD measurement, either pH paper (when sign \sim is used), or a Jenway 3510 pH meter were used. Chemical shifts (δ) are given in parts per million (ppm) to higher frequency compared to the methyl signal of either trimethylsilane (organic solvent), or the sodium salt of 3-(trimethylsilyl)propanesulfonic acid (D_2O) at 0 ppm.[101] Data were processed using Bruker Topspin version 2.1 (300 and 400 MHz) or 1.3 (500 MHz) and Mestrenova Lite version 5.2.5.

2.4.2 – Synthetic procedure and characterisation

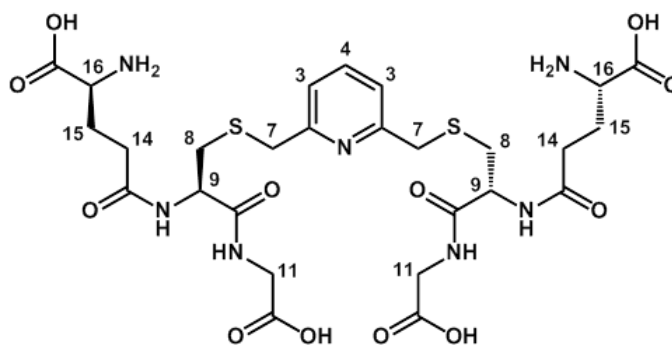
Synthesis of cys_2pyr : 2,6-bis(methyl-S-cysteinyl)pyridine



To a degassed solution of cysteine (20.4 mM, 245 mL, 5 mmol) in 50 mM aqueous Tris.HCl buffer pH 8.0, was added 2,6-bis(bromomethyl)pyridine in DMF (0.4 M, 2.5 mL, 1 mmol) dropwise. A white solid initially precipitated and was redissolved on addition of DMF (2.5 mL) and stirring for 4 hours under N_2 bubbling at room temperature. The solvent was

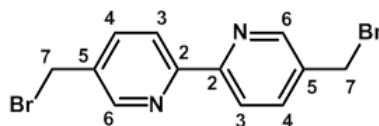
evaporated *in vacuo* at 322 K to yield a white solid. The crude product was purified by preparative RP-HPLC (C18 Phenomenex, using water/acetonitrile methods, monitoring absorbance at 210 and 256 nm) twice: (1) between 2 and 3 % acetonitrile in water (isocratic sequence, but composition of the eluent was slightly varied depending on the quantity injected), and (2) 0 to 6 % gradient acetonitrile in water over 10 min, followed by 6 % isocratic acetonitrile in water for an additional 15 minutes. The solvent was evaporated *in vacuo* to yield pure 2,6-bis(methyl-S-cysteinyl)pyridine, as a white solid (112 mg, 32 %). ^1H NMR (300 MHz, D_2O , dioxane reference, pD ~ 6): 7.83 (t, $^3J_{\text{H4-H3}} = 7.8$ Hz, 1H, H^4), 7.39 (d, $^3J_{\text{H3-H4}} = 7.8$ Hz, 2 H, H^3), 3.92 (s, 4 H, H^7), 3.88 (dd, $^3J_{\text{H9-H8b}} = 7.6$ Hz, $^3J_{\text{H9-H8a}} = 4.3$ Hz, 2 H, H^9), 3.07 (dd, $^2J_{\text{H8a-H8b}} = 14.9$ Hz, $^3J_{\text{H8a-H9}} = 4.3$ Hz, 2 H, H^{8a}), 2.96 (dd, $^2J_{\text{H8b-H8a}} = 14.9$ Hz, $^3J_{\text{H8b-H9}} = 7.6$ Hz, 2 H, H^{8b}). ^{13}C NMR ^1H decoupled (100 MHz ; D_2O ; dioxane reference, pD ~ 6): 157.9 (C^2), 139.8 (C^4), 123.4 (C^3), 54.3 (C^9), 37.1 (C^7), 32.5 (C^8). HRES-TOF (water), calculated exact mass for $\text{C}_{13}\text{H}_{19}\text{N}_3\text{O}_4\text{NaS}_2$: 368.0715; measured: 368.0721. ES-TOF: $m/z = 368.2$ $[\text{M}+\text{Na}]^+$ (50 %), $m/z = 346.2$ $[\text{M}+\text{H}]^+$ (100 %). MALDI-TOF: $m/z = 378.8$ $[\text{M}+\text{CH}_3\text{OH}+\text{H}]^+$ (23 %), $m/z = 367.8$ $[\text{M}+\text{Na}]^+$ (30 %), $m/z = 345.8$ $[\text{M}+\text{H}]^+$ (100 %). RP-HPLC (no TFA): 100%. Elemental analysis: % calculated for $\text{C}_{13}\text{H}_{19}\text{N}_3\text{O}_4\text{S}_2$: C 45.20, H 5.54, N 12.16; % measured: C 45.10, H 5.38, N 12.14. UV (water): 272 (5,530).

Synthesis of pyr-GS₂: 2,6-bis(methyl-S-glutathionyl)pyridine

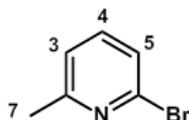


2,6-Bis(bromomethyl)pyridine (75.8 mg, 0.286 mmol) was added to 40.4 mL of 60 mM aqueous Tris.HCl buffer pH 8.0 solution. The suspension was degassed by N_2 bubbling

directly into the solution, glutathione (58.8 mM, 0.573 mmol, 9.75 mL) aqueous solution was added, and the suspension was stirred at room temperature. The concentration of glutathione was monitored by Ellman's test during the reaction. After 14 hours, the pH dropped down to 3.7 and was increased to 7.4 by adding NaOH solution (1 M, 0.5 mmol, 500 μ L). The reaction was stopped after 48 hours and the crude mixture was dried under vacuum, affording a colorless gel. The crude product was purified by preparative RP-HPLC (C18 Phenomenex, monitoring absorbance at 210 and 260 nm) in twice : (1) 0 to 25 % gradient of acetonitrile in water over 40 min (both containing 0.05 % TFA), and (2) 0 to 5 % gradient of acetonitrile in water over 20 min, followed by 5 % of acetonitrile in water isocratic for an additional 20 min (all eluents contained 0.05 % TFA). The solvent was evaporated *in vacuo* to yield 2,6-bis(methyl-S-glutathionyl)pyridine as a salt, and as a colorless gel (160 mg, 78 %). ^1H NMR (300 MHz, D_2O , dioxane reference, pD \sim 1): 8.47 (t, $^3J_{\text{H4-H3}} = 8.0$ Hz, 1 H, H^4), 7.93 (d, $^3J_{\text{H3-H4}} = 8.0$ Hz, 2 H, H^3), 4.54 (dd, $^3J_{\text{H9-H8a}} = 5.4$ Hz, $^3J_{\text{H9-H8b}} = 8.3$ Hz, 2 H, H^9) ; 4.16 (s, 4 H, H^7), 4.07 (t, $^3J_{\text{H16-H15}} = 6.6$ Hz, 4 H, H^{16}), 3.99 (s, 4 H, H^{11}), 3.04 (dd, $^2J_{\text{H8a-H8b}} = 14.2$ Hz, $^3J_{\text{H8a-H9}} = 5.4$ Hz, 2 H, H^{8a}), 2.88 (dd, $^2J_{\text{H8b-H8a}} = 14.2$ Hz, $^3J_{\text{H8b-H9}} = 8.3$ Hz, 2 H, H^{8b}), 2.57 (m, 4 H, H^{14}), 2.22 (m, 4 H, H^{15}). ^{13}C NMR ^1H decoupled (100 MHz ; D_2O ; dioxane reference, pD \sim 1): 174.9 (C^{13}), 173.4 (C^{12}), 172.8 (C^{10}), 172.2 (C^{17}), 154.5 (C^2), 148.0 (C^4), 126.6 (C^3), 53.3 (C^9), 52.9 (C^{16}), 41.7 (C^{11}), 33.4 (C^8), 33.1 (C^7), 31.5 (C^{14}), 26.1 (C^{15}). HRES-TOF (water, negative mode) calculated exact mass for $\text{C}_{27}\text{H}_{40}\text{N}_7\text{O}_{12}\text{S}_2$: 718.2176; measured: 718.2184. ES-TOF: $m/z = 716.1$ $[\text{M-H}]^-$ (100 %). MALDI-TOF: $m/z = 740.3$ $[\text{M}+\text{Na}]^+$ (87 %), $m/z = 718.2$ $[\text{M}+\text{H}]^+$ (100 %), $m/z = 565.3$ (11 %), $m/z = 413.6$ $[\text{M}-\text{C}_{10}\text{H}_{16}\text{N}_3\text{O}_6\text{S}+\text{H}]^+$ (33 %), $m/z = 309.0$ $[\text{C}_{10}\text{H}_{17}\text{N}_3\text{O}_6\text{S} + \text{H}]^+$ (13 %). RP-HPLC (no TFA): 100%. Elemental analysis : % calculated for $\text{C}_{27}\text{H}_{39}\text{N}_7\text{O}_{12}\text{S}_2$ (**pyr-GS₂**): C 45.18, H 5.48, N 13.66; % measured: C 38.82, H 4.55, N 8.83; best match found: $\text{C}_{33}\text{H}_{49}\text{N}_7\text{O}_{43/2}\text{S}_2\text{F}_9$ (**pyr-GS₂**, 3TFA, 3.5H₂O): C 35.30, H 4.40, N 8.73. UV: 274 (based on **pyr-GS₂**: 3,190; based on **pyr-GS₂**, 3TFA, 3.5H₂O: 4,976).

Synthesis of bipy-Br₂: 5,5'-dibromomethyl-2,2'-bipyridine [48,49]

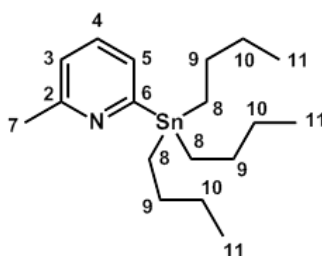
5,5'-dimethyl-2,2'-bipyridine (0.389 g, 2.11 mmol), N-bromosuccinimide (0.756 g, 4.23 mmol) and azobisisobutyronitrile (10 mg, 0.06 mmol) were dissolved in 20 mL dichloromethane, refluxed using a 500 W halogen lamp, and the reaction progress monitored by TLC (SiO₂, eluent: CH₂Cl₂/CH₃OH (9/1)). After 4 hours, more N-bromosuccinimide (0.375 g, 2.10 mmol) was added and the reaction was refluxed for a further 2 hours. The mixture was allowed to cool to room temperature and the solution extracted with 0.1M aqueous NaHCO₃ (5x 20 mL). The organic layer was dried over Na₂SO₄, filtered and concentrated *in vacuo*. The solid was re-dissolved in 15 mL of a 50/50 mix CHCl₃/CH₃OH, and the resulting solution stored in the freezer overnight (253 K). A white solid precipitate was collected by filtration, air-dried, and recrystallized from CHCl₃, to afford white crystals (0.219 g, 30 %). ¹H NMR (300 MHz, CDCl₃): 8.69 (d, ³J_{H6-H4} = 2.1 Hz, 2H, H⁶), 8.41 (d, ³J_{H3-H4} = 8.2 Hz, 2H, H³), 7.86 (dd, ³J_{H4-H6} = 2.3 Hz, ³J_{H4-H3} = 8.2 Hz, 2H, H⁴), 4.54 (s, 4H, H⁷); ¹³C NMR ¹H decoupled (100 MHz, CDCl₃): 155.4 (C²), 149.4 (C⁶), 137.9 (C⁴), 134.1 (C⁵), 121.4 (C³), 29.6 (C⁷); HRES-TOF (CH₂Cl₂) calculated mass for C₁₂H₁₀N₂⁷⁹Br₂Na: 362.9108; measured: 362.9126. ES-TOF : m/z = 362.9 ([M+Na]⁺, 100 %). UV (ethanol): 256 (6,900), 298 (13,000).

Synthesis of 2-bromo-6-methyl pyridine[47]

2-bromopyridine (10.8 g, 115 mmol) in 40 mL of hydrobromic acid (48 %) was cooled to 253 K in an ethanol bath. Bromine (14.4 mL, 280 mmol) was added dropwise, and the suspension was stirred for 90 min at 253 K. 30 mL of an aqueous solution of NaNO₂ (8.9 M,

268 mmol) was added dropwise, and the solution was allowed to warm to room temperature over 2 hours with stirring. The mixture was re-cooled to 253 K, and 110 mL of a cool NaOH aqueous solution (16.5 M, 1.81 mol) added slowly, while maintaining the temperature below 283 K. The mixture was allowed to warm to room temperature over the course of one more hour with continuous stirring. The mixture was then extracted with ethyl acetate and the organic layer dried over Na₂SO₄, filtered and concentrated *in vacuo*. The dark oil was then purified by Kugelrohr distillation to yield a colourless oil (9.906 g, 50 %). ¹H NMR (300 MHz, CDCl₃): 7.43 (t, ³J_{H4-H5} = ³J_{H4-H3} = 7.7 Hz, 1H, H⁴), 7.29 (d, ³J_{H5-H4} = 7.9 Hz, 1H, H⁵), 7.10 (d, ³J_{H3-H4} = 7.4, 1H, H³), 2.54 (s, 3H, H⁷); ¹³C NMR ¹H decoupled (100 MHz, CDCl₃): 160.2 (C⁶), 141.5 (C²), 138.7 (C⁴), 125.2 (C⁵), 122.2 (C³), 24.3 (C⁷); HRES-TOF (CH₂Cl₂) calculated exact mass for C₆H₆⁷⁹BrN: 170.9684; measured: 170.9686. EI-TOF: m/z = 171.0 ([M⁺], 50 %), 92.0 ([M-Br]⁺, 100 %), 65.0 ([M-C₂H₃Br]⁺, 75 %).

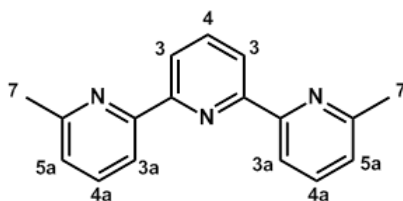
Synthesis of 2-tributylstannyl-6-methyl pyridine[47]



A solution of 2-bromo-6-methylpyridine (4.63 g, 27 mmol) in dry THF (20 mL) was cooled to 213 K, 27 mL of n-butyllithium in hexane (1.1 M, 30.1 mmol) added dropwise, and the solution stirred for 2 hours at 213 K. Tributyltin chloride (8.9 mL, 32.8 mmol) was slowly added and the solution allowed to return to room temperature over 20 minutes with continuous stirring. Water (30 mL) was added to the reaction mixture, and phases were separated. The organic phase was washed with water (3x 30 mL), and the combined aqueous layers washed with diethylether (4x 30 mL). The combined organic phases were dried over Na₂SO₄, filtered and concentrated *in vacuo* to afford a black oil (9.534 g, 92 %). The crude

product was determined to be ca. 95% pure by GC analysis and so used directly in the following synthesis. ^1H NMR (300 MHz, CDCl_3): 7.36 (t, $^3J_{\text{H4-H5}} = ^3J_{\text{H4-H3}} = 7.5$ Hz, 1 H^4), 7.17 (d, $^3J_{\text{H4-H5}} = 7.3$ Hz, 1H, H^5), 6.96 (d, $^3J_{\text{H4-H3}} = 7.9$, 1H, H^3), 2.54 (s, 3 H^7), 1.56 (m, 6H, H^9), 1.33 (m, 6H, H^{10}), 1.09 (m, 6H, H^8), 0.88 (t, $^3J_{\text{H11-H10}} = 7.2$, 9H, H^{11}); ^{13}C NMR ^1H decoupled (100 MHz, CDCl_3): 173.2 (C^6), 158.7 (C^2), 133.3 (C^4), 129.5 (C^5), 121.6 (C^3), 29.2 (C^9), 27.5 (C^{10}), 25.1 (C^7), 13.8 (C^{11}), 10.0 (C^8); HRES-TOF (CH_2Cl_2): calculated exact mass for $\text{C}_{18}\text{H}_{33}\text{N}^{120}\text{Sn}$: 384.1713; measured: 384.1715. EI-TOF: $m/z = 326.0$ ($[\text{M}-\text{C}_4\text{H}_9]^+$, 52 %), 268.0 ($[\text{M}-\text{C}_8\text{H}_{19}]^+$, 54 %), 211.9 ($[\text{M}-\text{C}_{12}\text{H}_{26}]^+$, 100 %), 177.0 ($[\text{M}-\text{C}_{14}\text{H}_{24}]^+$, 14 %), 120.9 ($[\text{M}-\text{C}_{18}\text{H}_{32}]^+$, 37 %), 93.1 ($[\text{M}-\text{C}_{12}\text{H}_{26}\text{Sn}]^+$, 25 %).

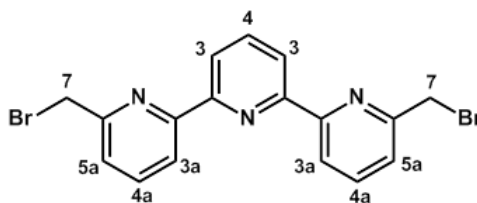
Synthesis of Me_2terpy : 6,6''-dimethyl-2,2':6',2''-terpyridine[48]



A solution of 2,6-dibromopyridine (1.39 g, 5.9 mmol), 2-tributylstannyl-6-methylpyridine (5.02 g, 13.1 mmol) and tetrakis(triphenylphosphine)palladium(0) (0.38 g, 0.33 mmol) was refluxed in 40 mL degassed toluene for 5 days under a nitrogen atmosphere. Extra tetrakis(triphenylphosphine)palladium(0) (0.38 mg, 0.33 mmol) was added and the reaction refluxed for a further day under a nitrogen atmosphere. The crude mixture was then concentrated *in vacuo*, and dichloromethane (50 mL) and 6M hydrochloric acid (10 mL) added to form a dark brown slurry. The aqueous layer was washed with dichloromethane (2x 50 mL) and the combined organic layers were washed with 6M hydrochloric acid (3x 10 mL). The combined aqueous layers were filtered and cooled in ice. Ammonia was slowly added until a light brown solid precipitated out. The resulting solid was filtered, air dried, redissolved in dichloromethane, dried over Na_2SO_4 , and concentrated *in vacuo*. The crude product was purified on a chromatography column (Al_2O_3 , hexane/DCM gradient) to afford

the pure 6,6''-dimethyl-2,2':6',2''-terpyridine as a white solid (0.63 g, 41 %). ^1H NMR (300 MHz, CDCl_3): 8.46 (d, $^3J_{\text{H}3-\text{H}4} = 7.8$ Hz, 2H, H^3), 8.41 (d, $^3J_{\text{H}3\text{a}-\text{H}4\text{a}} = 7.8$ Hz, 2H, $\text{H}^{3\text{a}}$), 7.93 (t, $^3J_{\text{H}4-\text{H}3} = 7.8$ Hz, 1H, H^4), 7.73 (t, $^3J_{\text{H}4\text{a}-\text{H}3\text{a}} \sim ^3J_{\text{H}4\text{a}-\text{H}5\text{a}} = 7.7$ Hz, 2H, $\text{H}^{4\text{a}}$), 7.19 (d, $^3J_{\text{H}5\text{a}-\text{H}4\text{a}} = 7.6$ Hz, 2H, $\text{H}^{5\text{a}}$), 2.65 (s, 6H, H^7); ^{13}C NMR ^1H decoupled (100 MHz, CDCl_3): 158.0 ($\text{C}^{6\text{a}}$), 155.9 ($\text{C}^{2/2\text{a}}$), 155.8 ($\text{C}^{2/2\text{a}}$), 137.8 (C^4), 137.1 ($\text{C}^{4\text{a}}$), 123.3 ($\text{C}^{5\text{a}}$), 121.0 (C^3), 118.3 ($\text{C}^{3\text{a}}$), 24.8 (C^7); HRES-TOF (CH_2Cl_2): calculated mass for $\text{C}_{17}\text{H}_{15}\text{N}_3\text{Na}$: 284.1164; measured: 284.1152. ES-TOF: $m/z = 284.0$ [$\text{M}+\text{Na}$] $^+$ (100 %).

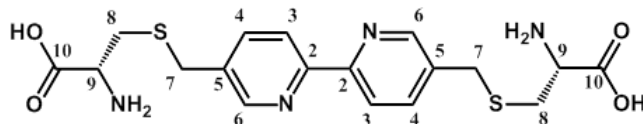
Synthesis of terpy-Br₂: 6,6''-dibromomethyl-2,2':6',2''-terpyridine[48,49]



A solution of 6,6'-dimethyl-2,2':6',2''-terpyridine (0.44 g, 1.7mmol), N-bromosuccinimide (0.77 g, 4.3mmol) and azobisisobutyronitrile (10 mg, 0.06 mmol) in dichloromethane (30 mL) was refluxed using a 500 W halogen lamp. Progress of the reaction was monitored by TLC (SiO_2 , eluent: $\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ (9/1)). After 15 hours, further N-bromosuccinimide (0.3 g, 1.7mmol) was added and the reaction refluxed for a further 17 hours. The mixture was extracted with 0.1 M NaHCO_3 (5x 50 mL), the organic layers dried over Na_2SO_4 and concentrated *in vacuo*. Addition of a 50/50 mixture of $\text{CHCl}_3/\text{CH}_3\text{OH}$ (15 mL) and storage at 253 K overnight resulted in the formation of a white precipitate. This was collected by filtration, air-dried, and recrystallized from CHCl_3 to afford white crystals (0.41 g, 58 %). ^1H NMR (300 MHz, CDCl_3): 8.53 (d, $^3J_{\text{H}3-\text{H}4} = 7.8$ Hz, 2H, H^3), 8.52 (d, $^3J_{\text{H}3\text{a}-\text{H}4\text{a}} = 7.8$ Hz, 2H, $\text{H}^{3\text{a}}$), 7.96 (t, $^3J_{\text{H}4-\text{H}3} = 7.8$ Hz, 1H, H^4), 7.86 (t, $^3J_{\text{H}4\text{a}-\text{H}3\text{a}} \sim ^3J_{\text{H}4\text{a}-\text{H}5\text{a}} = 7.8$ Hz, 2H, $\text{H}^{4\text{a}}$), 7.50 (d, $^3J_{\text{H}5\text{a}-\text{H}4\text{a}} = 7.7$ Hz, 2H, $\text{H}^{5\text{a}}$), 4.66 (s, 4H, H^7). ^{13}C NMR ^1H decoupled (100 MHz, CDCl_3): 156.4 ($\text{C}^{2/2\text{a}}$), 156.1 ($\text{C}^{2/2\text{a}}$), 155.1 ($\text{C}^{6\text{a}}$), 138.0 (C^4), 137.0 ($\text{C}^{4\text{a}}$), 123.6 ($\text{C}^{5\text{a}}$), 121.6 ($\text{C}^{3\text{a}}$), 120.4 (C^3), 34.3 (C^7); HRES-TOF (CH_2Cl_2): calculated exact mass for

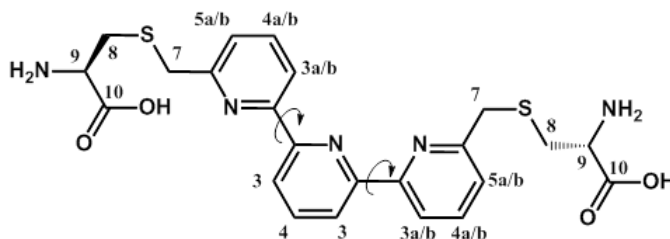
$C_{17}H_{13}N_3^{79}Br^{81}BrNa$: 441.9353; measured: 441.9355 ($[M+Na]^+$). ES-TOF: $m/z = 440.0$ ($[M+Na]^+$ (100 %)). UV (EtOH): 212 (28,900), 248 (12,800), 290 (15,900).

Synthesis of cys₂bipy: 5,5'-bis(methyl-S-cysteinyl)-2,2'-bipyridine



In a microwave-vessel containing a suspension of 5,5'-bis(bromomethyl)-2,2'-bipyridine (23 mg, 67.3 μ mol) in 3 mL of acetonitrile, was added 100 mM aqueous tris.HCl buffer pH 8.0 (2 mL), and a solution of cysteine (135 mM, 1 mL, 0.135 mmol) in the same buffer. The suspension was bubbled with N_2 for 10 minutes and was then stirred for 3 hours at 338 K and 100 W in a microwave reactor. The solvent was evaporated *in vacuo* at 323 K to yield a pink gel. Concentrated acid was added in order to solubilise the product, and pH was neutralised using a solution of NaOH. The crude product was purified by preparative RP-HPLC (C18 Phenomenex, monitoring absorbance at 210 and 290 nm) using 0 to 20 % gradient acetonitrile in water over 30 min. The solvent was evaporated *in vacuo* to yield pure 5,5'-bis(methyl-S-cysteinyl)-2,2'-bipyridine, as a pink solid (19 mg, 67 %). 1H NMR (300 MHz ; D_2O ; dioxane reference ; pD~ 1) : 8.86 (br. s, 2 H, H^6), 8.44 (br. s, 4 H, $H^{3/4}$), 4.17 (br. m, 2 H, H^9), 4.06 (br. s, 4 H, H^7), 3.08 (br. m, 2 H, H^8). ^{13}C NMR 1H decoupled (100 MHz ; D_2O ; dioxane reference, pD~1) : 171.9 (C^{10}), 146.7 (C^6), 144.0 (C^4), 138.8 (C^5), 124.4 (C^3), 53.3 (C^9), 32.6 (C^7), 31.6 (C^8). HRES-TOF (water): calculated exact mass for $C_{18}H_{23}N_4O_4S_2$: 423.1161; measured: 423.1172. ES-TOF: $m/z = 423.3$ $[M+H]^+$ (100 %), $m/z = 192.2$ (42 %). MALDI-TOF: $m/z = 444.5$ $[M+Na]^+$ (100 %), $m/z = 422.5$ $[M+H]^+$ (29 %). RP-HPLC (no TFA): 100%. UV: 245 (10,240), 296 (14,397).

Synthesis of cys₂terpy: 6,6''-bis(methyl-S-cysteinyl)-2,2':6',2''-terpyridine

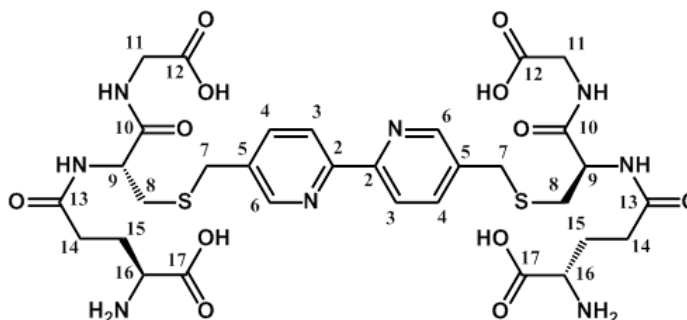


In a microwave-vessel containing a suspension of 6,6''-bis(bromomethyl)-2,2':6',2''-terpyridine (25.5 mg, 60.8 μ mol) in 3 mL of acetonitrile, was added 100 mM aqueous Tris.HCl buffer pH 8.0 (2 mL), and a solution of cysteine (122 mM, 1 mL, 0.122 mmol) in the same buffer. The suspension was bubbled with N₂ for 10 minutes and was then stirred for 3.5 hours at 338 K and 100 W in a microwave reactor. The solvent was evaporated *in vacuo* at 323 K to yield a colorless gel. Concentrated acid was added in order to solubilise the product, and pH was neutralised using a solution of NaOH. The crude product was purified by preparative RP-HPLC (C18 Phenomenex, monitoring absorbance at 210 and 290 nm) using 0 to 60 % gradient acetonitrile in water over 25 min. The solvent was evaporated *in vacuo* to yield pure 6,6''-bis(methyl-S-cysteinyl)-2,2':6',2''-terpyridine, as a white solid (7 mg, 23 %).

¹H NMR (300 MHz ; D₂O ; dioxane reference ; pD~1): 8.76 (dd, ³J_{H4a-H3a} = 8.2 Hz ; ³J_{H4a-H5a} = 7.0 Hz, 2 H, H^{4a}), 8.72 (dd, ³J_{H3a-H4a} = 8.3 Hz ; ⁴J_{H3a-H5a} = 2.0 Hz, 2 H, H^{3a}), 8.62 (d, ³J_{H3-H4} = 7.9 Hz, 1 H, H³), 8.46 (dd, ³J_{H4-H3} = 7.4 Hz ; ³J_{H4-H3} = 8.6 Hz, 1 H, H⁴), 8.24 (dd, ³J_{H5a-H4a} = 7.0 Hz ; ⁴J_{H5a-H3a} = 1.9 Hz, 2 H, H^{5a}), 4.53 (s, 4 H, H⁷), 4.43 (dd, ³J_{H9-H8b} = 6.8 Hz, ³J_{H9-H8a} = 4.6 Hz, 2 H, H⁹), 3.31 (dd, ²J_{H8a-H8b} = 15.0 Hz, ³J_{H8a-H9} = 4.6 Hz, 2 H, H^{8a}), 3.21 (dd, ²J_{H8b-H8a} = 15.0 Hz, ³J_{H8b-H9} = 6.9 Hz, 2 H, H^{8b}). ¹³C NMR ¹H decoupled (125 MHz ; D₂O ; dioxane reference, pD~1): 171.9 (C¹⁰), 155.0 (C^{6a}), 148.4 (C^{2/2a}), 148.0 (C^{4a}), 148.0 (C^{2/2a}), 142.2 (C⁴), 128.83 (C^{5a}), 126.4 (C³), 124.6 (C^{3a}), 53.3 (C⁹), 33.2 (C⁷), 31.7 (C⁸). HRES⁻-TOF (water, negative mode): calculated exact mass for C₂₃H₂₄N₅O₄S₂: 498.1270; measured: 498.1263. ES-TOF: m/z = 498.2 [M-H]⁻ (100 %) ; m/z = 411.2 (5 %). MALDI-TOF: m/z = 520.0 [M+Na]⁺

(22 %), $m/z = 498.2$ $[M+H]^+$ (100 %), $m/z = 379.8$ $[M-C_3H_6NO_2S+H]^+$ (65 %). RP-HPLC (no TFA): 84%. UV: 298 (16,857).

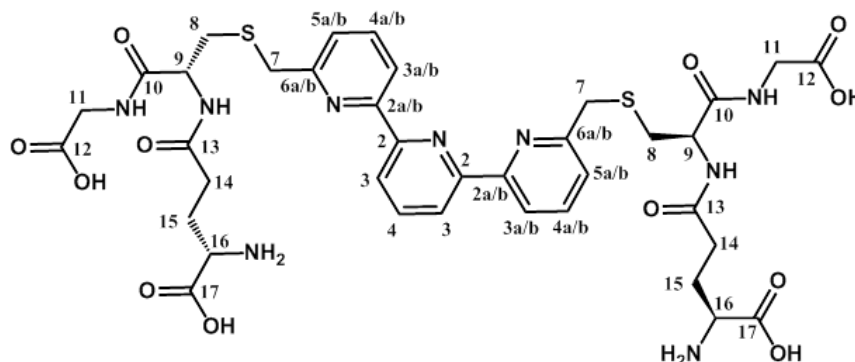
Synthesis of bipy-GS₂: 5,5'-bis(methyl-S-glutathionyl)-2,2'-bipyridine



A solution of glutathione (30.3 mM, 5 mL, 0.152 mmol) in 100 mM Tris.HCl buffer pH 8.0 was added to a solution of 5,5'-bis(bromomethyl)-2,2'-bipyridine (15.3 mM, 76.6 μ mol, 5 mL) in acetonitrile. The resulting suspension was degassed with $N_{2(g)}$ for 10 minutes and then stirred for 11 hours at room temperature. The solvent was evaporated *in vacuo* at 323 K to yield a pink gel. Deionised water (ca. 5 mL) and few drops of HCl (35 % w) were added to the gel resulting in complete solubilisation. The pH was neutralised on addition of NaOH (1 M) solution. The product was purified by preparative RP-HPLC (C18 Phenomenex, monitoring absorbance at 210 and 290 nm) using 0 to 15% gradient acetonitrile in water (containing 0.05% TFA) over 30 min. The solvent was evaporated *in vacuo* to yield pure 5,5'-bis(methyl-S-glutathionyl)-2,2'-bipyridine as a pink solid (42 mg, 68 %). 1H NMR (300 MHz, D_2O , pD \sim 1): 8.83 (s, 2H, H^6), 8.42 (s, 4H, H^3 and H^4), 4.53 (dd, $^3J_{H9-H8a} = 5.3$ Hz, $^3J_{H9-H8b} = 8.4$ Hz, 2H, H^9), 4.02 (s, 4H, H^7), 4.00 (m, 2H, H^{16}), 3.99 (s, 4H, H^{11}), 3.04 (dd, $^2J_{H8a-H8b} = 14.2$ Hz, $^3J_{H8a-H9} = 5.4$ Hz, 2H, H^{8a}), 2.86 (dd, $^2J_{H8b-H8a} = 14.2$ Hz, $^3J_{H8b-H9} = 8.5$ Hz, 2H, H^{8b}), 2.56 (m, 4H, H^{14}), 2.19 (m, 4H, H^{15}); ^{13}C NMR 1H decoupled (100 MHz, D_2O , pD \sim 1): 175.0 (C^{13}), 173.5 (C^{12}), 173.1 (C^{10}), 172.7 (C^{17}), 146.6 (C^6), 144.0 (C^4), 139.4 (C^5), 124.3 (C^3), 53.4 (C^9), 53.2 (C^{16}), 41.7 (C^{11}), 33.1 (C^8), 32.8 (C^7), 31.6 (C^{14}), 26.2 (C^{15}); HRES-TOF

(water): calculated exact mass for $C_{32}H_{43}N_8O_{12}S_2$: 795.2442; measured: 795.2476. ES-TOF: $m/z = 817.2 [M+Na]^+$ (45 %), $m/z = 795.5 [M+H]^+$ (100 %). RP-HPLC (0.05% TFA): 100%.

Synthesis of terpy-GS₂: 6,6''-bis(methyl-S-glutathionyl)-2,2':6',2''-terpyridine



To a suspension of 6,6''-bis(bromomethyl)-2,2':6',2''-terpyridine (2.52 mM, 25.2 μ mol, 10 mL) in acetonitrile, was added 100 mM aqueous Tris.HCl buffer pH 8.0 (5 mL) and a solution of glutathione (20.1 mM, 5 mL, 0.100 mmol) in the same buffer. The suspension was degassed with $N_{2(g)}$ for 10 min, heated to 318 K and stirred at this temperature for 8 hours. The solvent was evaporated *in vacuo* at 323 K to yield a colourless gel. Deionised water (ca. 5 mL) and few drops of HCl (35 % w) were added to the gel resulting in complete solubilisation. The pH was neutralised on addition of few drops of NaOH (1 M) solution. The product was purified by preparative RP-HPLC (C18 Phenomenex, monitoring absorbance at 210 nm) using 0 to 30% gradient acetonitrile in water (containing 0.05 % TFA) over 25 min. The solvent was evaporated *in vacuo* to yield pure 6,6''-bis(methyl-S-glutathionyl)-2,2':6',2''-terpyridine as a white solid (17 mg, 76 %). 1H NMR (300 MHz, D_2O , pD \sim 1): 8.68 (dd, $^3J_{H4a-H3a} = 8.1$ Hz, $^3J_{H4a-H5a} = 6.2$ Hz, 2H, H^{4a}), 8.67 (dd, $^3J_{H3a-H4a} = 8.1$ Hz, $^4J_{H3a-H5a} = 2.7$ Hz, 2H, H^{3a}), 8.59 (d, $^3J_{H3-H4} = 7.5$ Hz, 1H, H^3), 8.47 (dd, $^3J_{H4-H3} = 7.1$ Hz, $^3J_{H4-H3} = 8.8$ Hz, 1H, H^4), 8.17 (dd, $^3J_{H5a-H4a} = 6.2$ Hz, $^4J_{H5a-H3a} = 2.5$ Hz, 2H, H^{5a}), 4.56 (dd, $^3J_{H9-H8b} = 8.2$ Hz, $^3J_{H9-H8a} = 5.5$ Hz, 2H, H^9), 4.38 (s, 4H, H^7), 4.01 (t, $^3J_{H16-H15} = 6.6$ Hz, 2H, H^{16}), 3.89 (s, 4H, H^{13}), 3.15 (dd, $^2J_{H8a-H8b} = 14.2$ Hz, $^3J_{H8a-H9} = 5.6$ Hz, 2H, H^{8a}), 2.94 (dd, $^2J_{H8b-H8a} = 14.2$ Hz, $^3J_{H8b-H9} = 8.3$ Hz, 2H, H^{8b}), 2.51 (dd, $^3J_{H14-H15a} = 7.0$ Hz, $^3J_{H14-H15b} = 7.9$ Hz, 4H, H^{14}),

2.14 (m, 4H, H¹⁵). ¹³C NMR ¹H decoupled (125 MHz, D₂O, pD ~ 1): 174.8 (C¹³), 173.3 (C¹²), 172.7 (C¹⁰), 172.1 (C¹⁷), 155.8 (C^{6a}), 148.4 (C^{2/2a}), 148.1 (C^{2/2a}), 147.7 (C^{4a}), 142.5 (C⁴), 128.6 (C^{5a}), 126.2 (C³), 124.3 (C^{3a}), 53.3 (C⁹), 52.8 (C¹⁶), 41.6 (C¹¹), 33.9 (C⁸), 33.6 (C⁷), 31.5 (C¹⁴), 26.0 (C¹⁵); HRES⁺-TOF (water): calculated exact mass for C₃₇H₄₅N₉O₁₂S₂Na: 894.2527; measured: 894.2498. ES-TOF: m/z = 894.7 [M+Na]⁺ (100%), 872.7 [M+H]⁺ (40%), 589.5 [M-C₁₀H₁₆N₃O₆S+Na]⁺ (17%), 567.5 [M-C₁₀H₁₆N₃O₆S+H]⁺ (19%). RP-HPLC (no TFA): 100%.

2.4.3 – Analytical procedures

2.4.3.1 – Pyridine conjugates of cysteines and glutathione (related to section 2.2.1)

For monitoring of **pyr-GS₂** formation by Ellman's test, 20 µL of 50 mM 5,5'-dithiobis-(2-nitrobenzoic acid) in 100 mM phosphate buffer pH 8 was mixed with 600 µL of 100 mM phosphate buffer pH 8 in a 1 mL cuvette, and a blank spectrum was recorded on spectrometer(a). To the cuvette, 2 µL of either 11.8 mM glutathione stock solution or reaction mixture were then added, and a new spectrum recorded after 10 min equilibration. All measurements were run in triplicate. The absorbance at 412 nm was converted to thiol concentration knowing the extinction coefficient of the 5-thiolato-2-nitrobenzoate anion ($\epsilon_{412\text{ nm}} 14,150\text{ M}^{-1}\text{ cm}^{-1}$). [102]

For the determination of extinction coefficient, 3 distinct solutions of **cys₂pyr** and **pyr-Br₂** were prepared based on mass. 2 mM stock solutions were prepared in deionised water based on mass. Sonification and gentle heating were required to ensure substrates were fully dissolved. Prior to each measurement, spectrometer(a) was blanked with 2.6 mL of 23 mM phosphate buffer pH 8.0 solution in a 3 mL cuvette. Then, 30, 60, 90, or 120 µL of stock solution were added, as well as $V = 400 - X$ µL of water to adjust the total volume to 3 mL (X

equal volume of stock solution), and a spectrum recorded. Linear regression on the absorbance versus concentration plot was performed in excel 2007.

For the Cu(II) titration of pyridine linkers monitored by UV, aliquots of an aqueous 6 mM stock solution of CuCl₂ were titrated into 3 mL of solutions that all contained 20 mM potassium phosphate buffer pH 8, and either: nothing else (blank), 40 μM **cys₂pyr** or **pyr-GS₂**, or 80 μM L-glycine. Spectra were recorded on spectrometer(a), after 30 seconds equilibration. Linear fitting were performed with excel 2007.

2.4.3.2 – Polypyridine linkers (related to section 2.2.2)

Titration were performed by addition of aliquots of a 250 (**Me₂bipy**) or 232 mM (**Me₂terpy**) stock solution of ZnCl₂ in DMSO-d₆ to 0.75 mL of the dimethyl-polypyridine linker solution, 25 mM (**Me₂bipy**) or 23.2 mM (**Me₂terpy**), in the same solvent. NMR spectra were recorded on an AVIII 300 MHz spectrometer. All solution concentrations were based on mass. 2D NMR of **Me₂terpy** in either CDCl₃ or DMSO-d₆, were recorded on AVIII400: DQF-COSY and Gradient NOESY spectra (400 ms mixing time).

For the continuous variation method, a 25 mM stock solution of ZnCl₂ (**S_{Zn}**) and **Me₂bipy** (**S_{bipy}**) in DMSO-d₆ were prepared. From these stocks, 10 NMR tubes containing 1 mL of solution were prepared. Each tube contained a different volume of the **S_{Zn}** stock (from 0 to 0.9 mL, with a 0.1 mL step), and a volume of **S_{bipy}** equal to 1-V(**S_{Zn}**) mL. For each solution, the fraction of complex present in solution, χ (Zn:**Me₂bipy**), can be calculated from the ratio of chemical shift variation over the total variation, using equation (2.1).[103]

$$\chi (\text{Zn: Me}_2\text{bipy}) = \chi (\text{Me}_2\text{bipy})_T \times \frac{\delta_{obs} - \delta_0}{\delta_{\infty} - \delta_0} \quad (2.1)$$

$\chi (\text{Me}_2\text{bipy})_T$ represents the molar fraction of **Me₂bipy** (complexed and free) over the total concentration of species present (total **Me₂bipy** and ZnCl₂). δ_{obs} corresponds to the observed chemical shift, δ_0 the chemical shift of the free ligand (**Me₂bipy**), and δ_{∞} that for the

solution with the highest $\text{ZnCl}_2\text{:Me}_2\text{bipy}$ ratio (chemical shift at $\chi(\text{Me}_2\text{bipy})_{\text{T}} = 0.1$ was used here).

2.4.3.3 – Bipy and terpy conjugates of cysteine (related to section 2.2.3)

For the metal titration of cysteine conjugates monitored by UV spectroscopy, aliquots of aqueous 15 (**cys₂bipy**) or 21 mM (**cys₂terpy**) stock solution of either CuCl_2 or ZnCl_2 , were titrated into 3 mL of solutions either 50 (**cys₂bipy**) or 70 μM (**cys₂terpy**) cysteine polypyridine conjugates in 20 mM potassium phosphate buffer pH 8, and the spectra recorded on spectrometer(a), after 30 seconds equilibration. Linear fitting were performed with excel 2007. Additionally, K_{app} and K_{M} values for Zn(II) titration of **cys₂bipy** were calculated by non-linear fitting, using the same method than for glutathione conjugates, as described in the following section.

2.4.3.4 – Bipy and terpy conjugates of glutathione (related to section 2.2.4)

The concentrations of all **bipy-GS₂** or **terpy-GS₂** stock solutions were calculated after recording the UV profile in phosphate buffer (pH was either 7.4 or 8), and assuming extinction coefficients of 20,000 (± 200) $\text{M}^{-1} \text{cm}^{-1}$ (**bipy-GS₂**), or 21,000 (± 210) $\text{M}^{-1} \text{cm}^{-1}$ (**terpy-GS₂**). For the pH titration of model switches monitored by UV spectroscopy, aliquots of HCl, NaOH (**terpy-GS₂**) or KOH (**bipy-GS₂**) solutions of various concentrations (0.01, 0.1, 1M) were added to cuvettes containing 3 mL of a 5 μM solution of the model switch. The solution was allowed to equilibrate for 10 min prior to recording the pH on a Jenway 3510 pH meter, and recording a UV spectrum on spectrometer(b).

For the metal titrations of model switches monitored by UV (range 200-400 nm), aliquots of an aqueous 750 (± 32) μM stock solution of either CuCl_2 or ZnCl_2 , were titrated into 3.00 (± 0.03) mL of a 5.00 (± 0.05) μM solution of model switches in 20.0 (± 0.3) mM potassium phosphate buffer pH 7.4, and the spectra recorded after 3 min equilibration on spectrometer(b). Non-linear fitting was performed with Kaleidagraph software version 4.0.

K_{app} values were calculated by fitting data for the absorbance maximum of the metal complexes as a function of Cu(II)/Zn(II) concentration, to equations (2.2) and (2.3):

$$A = [L] \varepsilon_L b + [C](\varepsilon_C - \varepsilon_L) b \quad (2.2)$$

$$[C] = \frac{1 + K_{app} \times ([L] + [M]) - \sqrt{(1 + K_{app} \times ([L] + [M]))^2 - 4 (K_{app})^2 [L][M]}}{2 K_{app}} \quad (2.3)$$

b corresponds to the cuvette pathlength, $[C]$ corresponds to the concentration of the 1:1 complex, $[L]$ the total **bipy-GS₂** or **terpy-GS₂** and $[M]$ the total CuCl₂/ZnCl₂ concentration at each point, ε_L and ε_C are the extinction coefficient of the ligand and complex at the wavelength of interest, respectively.

In order to ensure an accurate estimation, measurements were performed at concentrations close to the apparent dissociation constants, such that:

$$\frac{[L]}{50} < \frac{1}{K_{app}} < [L]$$

The K_M values were corrected by accounting for the contribution from phosphate metal ion binding, based on values reported in the literature,[66] see equation (2.4):

$$K_{app} = K_M \times \left(1 - \frac{K_{PM}[P]}{1 + (K_{PM}[P])} \right) \quad (2.4)$$

This is based on the assumption that the concentration of free phosphate, $[P]$, is equal to the total amount of phosphate in solution. K_M and K_{PM} corresponds to the estimated binding constant of the metal ion to the ligand of interest and the phosphate anion, respectively. When L-glycine was added as a competitor, K_M was estimated using equation (2.5), based on the reported Cu(II) binding constant for L-glycine.[104] The contribution from the phosphate anion was regarded as negligible, when 20.0 (\pm 0.4) mM L-glycine was present.

$$K_{app} = K_M \times \left(1 - \frac{K_{Gly} [Gly]^2}{1 + (K_{Gly}[Gly]^2)} \right) \quad (2.5)$$

The absolute errors (Δa) indicated for solution concentrations arise from the highest values obtained upon calculation of the error accumulations which account for each solution preparation (error on mass and volume) and dilution step (error on concentration and volumes), following the general equation (2.6). The absolute errors on volume measurements cumulate the errors for each pipetting step, based on the general equation (2.7). The errors for the ligand concentrations ($[\Delta L]$) and extinction coefficients arise from the errors on the related equivalences (volumes and concentrations) calculated using the same method, and considering the cumulated errors after each aliquot addition. For the non-linear fittings, the errors on K_{app} cumulate the errors on calculation (given by the software) with the experimental errors introduced in the non-linear fitting. In the latter case, the relative error, $\Delta K_{app}(\text{exp})/K_{app}$, were maximised upon replacement of the parameter $[L]$ by $[L-\Delta L]$ in the fitting equations (2.2) and (2.3), and the list of $[M]$ values by $[M+\Delta M]$ (with ΔM , the cumulated errors for each addition step) affording a value of 0.088. The relative error $\Delta K_M/K_M$ arises either from the errors calculated for K_{app} , $[P]$ and reported for K_{PM} (0.089 for titrations performed in the absence of L-glycine), or from the errors calculated for K_{app} , $[Gly]$ and reported for K_{Gly} (0.092 for titrations performed in the presence of L-glycine). On a side-note, the errors reported in the tables for the metal complexes extinction coefficients are absolute calculation errors (associated with the fitting only). The errors were calculated by uncertainty analysis on sum/difference (2.6) and product/quotient (2.7), [105] where b and c are random independent variables from which a is calculated:

$$\text{For } a = b \times c \text{ or } a = b \div c \quad \frac{\Delta a}{a} = \sqrt{\left(\frac{\Delta b}{b}\right)^2 + \left(\frac{\Delta c}{c}\right)^2} \quad (2.6)$$

$$\text{For } a = b + c \text{ or } a = b - c \quad \Delta a = \sqrt{(\Delta b)^2 + (\Delta c)^2} \quad (2.7)$$

For the metal titrations of model switches monitored by visible spectroscopy (range 400-900 nm), aliquots of aqueous stock solution either 21 (**bipy-GS₂**) or 3 mM (**terpy-GS₂**) of CuCl₂, were titrated into 600 μ L of either 350 (**bipy-GS₂**) or 100 μ M (**terpy-GS₂**) solution of model switches in 20 mM potassium phosphate buffer pH 7.4, and the spectra recorded after 3 min equilibration on spectrometer(b).

Metal-titrations of model compounds monitored by circular dichroism were performed by addition of aliquots of a 21 mM stock solution of CuCl₂ or ZnCl₂, into a 350 μ M solution of **bipy-GS₂**/**terpy-GS₂** in 10 mM phosphate buffer pH 7.4. Spectra did not change over two minutes, and therefore samples were allowed to equilibrate for two minutes prior to measurement. Reported spectra are an average of 5 scans recorded between 200 and 400 nm at 200 nm min⁻¹ (0.5 nm pitch) and the buffer blank subsequently subtracted (except for Cu titration of **terpy-GS₂**). After two (Zn) or three (Cu) equivalents of metal ion were added, 21 or 31.5 μ L respectively of a 0.2 M EDTA stock solution was added to assess reversibility (20 equivalents of EDTA per metal ion).

For the preparation of NMR samples, stock solutions of **bipy-GS₂**/**terpy-GS₂** and phosphate buffer pD 7.4 in D₂O were premixed, and the pD adjusted with a 0.1 M solution of NaOH in D₂O, using the Jenway 3510 pH meter. D₂O was added to adjust the concentrations of **bipy-GS₂**/**terpy-GS₂** and phosphates to 5 and 50 mM, respectively, and the pD was checked, prior to transfer to the NMR tube. Titrations were performed by addition of aliquots of a 0.1 M stock solution of ZnCl₂ in D₂O to a 5 mM solution of polypyridyl-conjugate in 50 mM phosphate buffer pD 7.4 in D₂O, and either 4 mM dioxane (δ = 3.75 ppm) or 1 mM acetone (δ = 2.22 ppm) was used as an internal reference. Addition of two equivalents ZnCl₂ resulted in no more than a 10% increase to the total volume. 0.8 mL of a 0.25 M solution of EDTA in D₂O (pH adjusted to 8), and 100 μ L D₂O, were then added (20 equivalents EDTA vs. Zn) to the NMR sample, resulting in a dilution of the sample by two. Changes in peak integrations (see Figure 2.38) are reported relative to the internal dioxane reference.

2D NMR of **terpy-GS₂** in D₂O at acidic pD (dioxane internal reference), were first recorded on AVIII400: DQF-COSY and Gradient NOESY spectra (400 ms mixing time). Additionally, Gradient NOESY (States-TPPI, 450 ms mixing time) and Phase cycle ROESY (States-TPPI, 450 ms mixing time, 10 kHz spin lock field and an offset of 10 kHz in spin lock period to minimise HOHAHA effects) experiments were recorded on DRX500 with partial presaturation of water signal, both displaying the previously mentioned inter-ring coupling.

2.5 – References

-
- [1] Kaes C., Katz A., Hosseini M. W., *Chem. Rev.*, **2000**, 100, 3553-3590.
- [2] Fielding P. E., Le Fèvre R. J. W., *J. Chem. Soc.*, **1951**, 1811-1814.
- [3] Düggeli M., Christen T., von Zelewsky A., *Chem. Eur. J.*, **2005**, 11, 185-194.
- [4] Zahn S., Reckien W., Kirchner B., Staats H., Matthey J., Lützen A., *Chem. Eur. J.*, **2009**, 15, 2572-2580.
- [5] Merritt Jr. L. L., Schroeder E. D., *Acta Cryst.*, **1956**, 9, 801-804.
- [6] Rebek, Jr J., Trend J. E., Wattlely R. V., Chakravorti S., *J. Am. Chem. Soc.*, **1979**, 101, 4333-4337.
- [7] Kovbasyuk L., Krämer R., *Chem. Rev.*, **2004**, 104, 3161-3187.
- [8] Ross Kelly T., Bowyer M. C., Vijaya Bhaskar K., Bebbington D., Garcia A., Lang F., Kim M. H., Jette M. P., *J. Am. Chem. Soc.*, **1994**, 116, 3657-3658.
- [9] Morgan S. G., Burstall F. H., *J. Chem. Soc.*, **1932**, 20-30.
- [10] Chotalia R., Constable E. C., Hannon M. J., Tocher D. A., *J. Chem. Soc. Dalton Trans.*, **1995**, 3571-3580.
- [11] Hoang T. N. Y., Lathion T., Guénée L., Terazzi E., Piguet C., *Inorg. Chem.*, **2012**, 51, 8567-8575.
- [12] Petitjean A., Khouri R. G., Kyritsakas N., Lehn J., *J. Am. Chem. Soc.*, **2004**, 126, 6637-6647.

- [13] Haberhauer G., *Angew. Chem. Int. Ed.*, **2008**, 47, 3635-3638.
- [14] Miyoshi D., Karimata H., Wang Z.-M., Koumoto K., Sugimoto N., *J. Am. Chem. Soc.*, **2007**, 129, 5919-5925.
- [15] Costero A. M., Gil S., Parra M., Huguet N., Allouni Z., Lakhmiri R., Atlamsani A., *Eur. J. Org. Chem.*, **2008**, 1079-1084.
- [16] Cheng K. F., Drain C. M., Grohman K., *Inorg. Chem.*, **2003**, 42, 2075-2083.
- [17] Wang B., Wasielewski M. R., *J. Am. Chem. Soc.*, **1997**, 119, 12-21.
- [18] Kawano S., Fujita N., Shinkai S., *J. Am. Chem. Soc.*, **2004**, 126, 8592-8593.
- [19] Winter A., Hager M. D., Newkome G. R., Schubert U. S., *Adv.Mater.*, **2011**, 23, 5728-5748.
- [20] Schneider J. P., Kelly J. W., *J. Am. Chem. Soc.*, **1995**, 117, 2533-2546.
- [21] Lieberman M., Sasaki T., *J. Am. Chem. Soc.*, **1991**, 113, 1470-1471.
- [22] Rama G., Ardá A., Maréchal J., Gamba I., Ishida H., Jiménez-Barbèro J., Vázquez M. E., López M. E. V., *Chem. Eur. J.*, **2012**, 18, 7030-7035.
- [23] Radford R. J., Nguyen P. C., Tezcan F. A., *Inorg. Chem.*, **2010**, 49, 7106-7115.
- [24] Göritz M., Krämer R., *J. Am. Chem. Soc.*, **2005**, 127, 18016-18017.
- [25] Choi J. S., Kang C. W., Jung K., Yang J. W., Kim Y., Han H., *J. Am. Chem. Soc.*, **2004**, 126, 8606-8607.
- [26] Weizman H., Tor Y., *J. Am. Chem. Soc.*, **2001**, 123, 3375-3376.
- [27] Qi D., Tann C.-M., Haring D., Distefano M. D., *Chem. Rev.*, **2001**, 101, 3081-3111.
- [28] Hackenberger C. P. R., Schwarzer D., *Angew. Chem. Int. Ed.*, **2008**, 47, 10030-10074.
- [29] Sletten E., Bertozzi C. R., *Angew. Chem. Int. Ed.*, **2009**, 48, 6974-6998.
- [30] Lim R. K. V., Lin Q., *Chem. Commun.*, **2010**, 46, 1589-1600.
- [31] Cheng R. P., Fisher S. L., Imperiali B., *J. Am. Chem. Soc.*, **1996**, 118, 11349-11356.
- [32] Cuenoud B., Schepartz A., *Tetrahedron Lett.*, **1991**, 28, 3325-3328.
- [33] Radford R. J., Nguyen P. C., Tezcan F. A., *Inorg. Chem.*, **2010**, 49, 7106-7115.

- [34] Chen C. H., Sigman D. S., *Science*, **1987**, 237, 1197-1201.
- [35] Chalker J. M., Bernardes G. J. L., Lin Y. A., Davis B. G., *Chem. Asian J.*, **2009**, 4, 630-640.
- [36] Woolley G. A., *Acc. Chem. Res.*, **2005**, 38, 486-493.
- [37] Jo H., Meinhardt N., Wu Y., Kulkarni S., Hu X., Low K. E., Davies P. L., DeGrado W. F., Greenbaum D. C., *J. Am. Chem. Soc.*, **2012**, 134, 17704-17713.
- [38] Prakash H., Shodai A., Yasui H., Sakurai H., Hirota S., *Inorg. Chem.*, **2008**, 47, 5045-5047.
- [39] Kromidas S., More practical problem solving in HPLC, John Wiley & Sons, Weinheim, **2008**, p 197, ISBN 3-527-31113-0.
- [40] Chalker J. M., Lin Y. A., Boutureira O., Davis B. G., *Chem. Commun.*, **2009**, 3714-3716.
- [41] Nakamoto K., *J. Phys. Chem.*, **1960**, 64, 1420-1425.
- [42] Palmer C. R., Sloan L. S., Adrian Jr. J. C., Cuenoud B., Paoletta D. N., Schepartz A., *J. Am. Chem. Soc.*, **1995**, 117, 8899-8907.
- [43] Rau H. K., DeJonge N., Haehnel W., *Proc. Natl Acad. Sci. USA*, **1998**, 95, 11526-11531.
- [44] Khoshtarkib Z., Ebadi A., Ahmadi R., Alizadeh R., *Acta Crystallogr. E*, **2009**, E65, o1586.
- [45] Khan M. S., Muna Al-Mandhary M. R., Al-Suti M. K., Hisahm A. K., Raithby P. R., Ahrens B., Mahon M. F., Male L., Marseglia E. A., Tedesco E., Friend R. H., Köhler A., Feeder N., Teat S. J., *J. Chem. Soc., Dalton Trans.*, **2002**, 1358-1368.
- [46] Ahmadi R., Kalateh K., Amani V., *Acta Crystallogr. E*, **2010**, E66, m562.
- [47] Schubert U. S., Eschbaumer C., Heller M., *Org. Lett.*, **2000**, 2, 3373-3376.
- [48] Galaup C., Couchet J.-M., Bedel S., Tisnès P., Picard C., *J. Org. Chem.*, **2005**, 70, 2274-2284.
- [49] Bedel S., Ulrich G., Picard C., *Tetrahedron Lett.*, **2002**, 43, 1697-1700.
- [50] Castellano S., Günther H., Ebersole S., *J. Phys. Chem.*, **1965**, 69, 4166-4176.

- [51] Jezowska-Trzebiatowska B., Kozłowski H., Latos-Grazynski L., Kowalik T., *Chem. Phys. Lett.*, **1975**, 30, 355-357.
- [52] Shen W. Z., Trötscher-Kaus G., Lippert B., *Dalton Trans.*, **2009**, 8203-8214.
- [53] Holyer R., Hubbard C., Kettle S., Wilkins R., *Inorg. Chem.*, **1965**, 4, 929-935.
- [54] Cini R., Donati A., Giannettoni R., *Inorg. Chim. Acta*, **2001**, 315, 73-80.
- [55] Holyer R. H., Hubbard C. D., Kettle S. F. A., Wilkins R. G., *Inorg. Chem.*, **1966**, 5, 622-625.
- [56] Burns D. C., Zhang F., Woolley G. A., *Nat. Protoc.*, **2007**, 2, 251-258.
- [57] Fan J., Autschbach J., Ziegler T., *Inorg. Chem.*, **2010**, 49, 1355–1362.
- [58] Niezborala C., Hache F., *J. Phys. Chem. A*, **2007**, 111, 7732-7735.
- [59] Hamilton J. M., Anhorn J. M., Oscarson K. A., Reibenspies J. H., Hancock R. D., *Inorg. Chem.*, **2011**, 50, 2764-2770.
- [60] Blanchet-Boiteux C., Friant-Michel P., Marsura A., Regnouf-de-Vains J.-B., Ruiz-Lopez M., *J. Mol. Struct. : Theochem.*, **2007**, 811, 169-174.
- [61] Imperiali B., Prins T. J., Fisher S. L., *J. Org. Chem.*, **1993**, 58, 1613-1616.
- [62] Imperiali B., Fisher S. L., *J. Org. Chem.*, **1992**, 57, 757-759.
- [63] Bazzicalupi C., Bencini A., Bianchi A., Danesi A., Faggi E., Giorgi C., Santarelli S., Valtancoli B., *Coord. Chem. Rev.*, **2008**, 252, 1052-1068.
- [64] Bezer S., Rapireddy S., Skorik Y. A., Ly D. H., Achim C., *Inorg. Chem.*, **2011**, 50, 11929-11937.
- [65] Garribba E., Micera G., Sanna D., Strinna-Erre L., *Inorg. Chim. Acta*, **2000**, 299, 253-261.
- [66] Banerjee D., Kaden T. A., Sigel H., *Inorg. Chem.*, **1981**, 8, 2586-2590.
- [67] Howard S. T., *J. Am. Chem. Soc.*, **1996**, 118, 10269-10274.
- [68] Göller A., Grummt U., *Chem. Phys. Lett.*, **2000**, 321, 399-405.
- [69] Göller A., Grummt U., *Chem. Phys. Lett.*, **2002**, 354, 233-242.

- [70] Braterman P. S., Song J. I., Peacock R. D., *Inorg. Chem.*, **1992**, 31, 555-559.
- [71] Fábián I., *Inorg. Chem.*, **1989**, 28, 3805-3807.
- [72] Yagi M., Wada Y., Takemura K., Ikemoto A., Kaneshima T., Seki K., *Chem. Phys. Lett.*, **1994**, 227, 261-266.
- [73] Sanna D., Buglyó P., Tomaz A. I., Pessoa J. C., Borovic S., Micera G., Garribba E., *Dalton Trans.*, **2012**, 41, 12824-12838.
- [74] Pfaulm R. T., Brandt W. W., *J. Am. Chem. Soc.*, **1955**, 76, 6216-6219.
- [75] Hogg R., Wilkins R. G., *J. Chem. Soc.*, **1962**, 341-350.
- [76] Solomon E. I., Szilagyí R. K., George D. S., Basumallick L., *Chem. Rev.*, **2004**, 104, 419-458.
- [77] Livoreil A., Sauvage J., Armaroli N., Balzani V., Flamigni L., Ventura B., *J. Am. Chem. Soc.*, **1997**, 119, 12114-12124.
- [78] Sigel H., Naumann C. F., Prijs B., Mc Cormick D. B., Falk M. C., *Inorg. Chem.*, **1977**, 16, 790-796.
- [79] Sigel H., *Inorg. Chem.*, **1975**, 14, 1535-1540.
- [80] Pagenkopf G. K., Margerum D. W., *J. Am. Chem. Soc.*, **1968**, 90, 6963-6967.
- [81] Kim M. K., Martell A. E., *J. Am. Chem. Soc.*, **1966**, 88, 914-918.
- [82] Nagy N. V., Szabó-Plánka T., Rockenbauer A., Peintler G., Nagypál I., Korecz L., *J. Am. Chem. Soc.*, **2003**, 125, 5227-5235.
- [83] Sóvágó I., Sanna D., Dessi A., Várgany K., Micera G., *J. Inorg. Biochem.*, **1996**, 63, 99-117.
- [84] Lim M. C., Sinn E., Bruce Martin R., *Inorg. Chem.*, **1976**, 15, 807-811.
- [85] Rudolph M., Autschbach J., *J. Phys. Chem. A*, **2011**, 115, 2635-2649.
- [86] Hidaka J., Douglas B. E., *Inorg. Chem.*, **1964**, 3, 1180-1184.
- [87] Fan J., Ziegler T., *Chirality*, **2011**, 23, 155-166.
- [88] Geotti-Bianchini P., Darbre T., Reymond J., *Org. Biomol. Chem.*, **2013**, 11, 344-352.

- [89] Ford-Smith M. H., Sutin N., *J. Am. Chem. Soc.*, **1961**, 83, 1830-1834.
- [90] Lieberman M., Sasaki T., *J. Am. Chem. Soc.*, **1991**, 113, 1470-1471.
- [91] Harada N., Uda H., *J. Chem. Soc., Chem. Commun.*, **1982**, 230-232.
- [92] Harada N., Tomoyoshi A., Uda H., *J. Chem. Soc., Chem. Commun.*, **1982**, 232-233.
- [93] Canary J. W., Mortezaei S., Liang J., *Chem. Comm.*, **2010**, 46, 5850-5860.
- [94] Bosnich B., *Acc. Chem. Res.*, **1969**, 2, 266-273.
- [95] Funck L. L., Ortolano T. R., *Inorg. Chem.*, **1968**, 7, 567-573.
- [96] Jensen H. P., Larsen E., *Acta Chem. Scand.*, **1971**, 25, 1439-1451.
- [97] Murakami T., Hatano M., *Inorg. Chem.*, **1975**, 14, 999-1001.
- [98] Jezowska-Trzebiatowska B., Formicka-Kozłowska G., Kozłowski H., Latos-Grazynski L., Kowalik T., *Chem. Phys. Lett.*, **1975**, 30, 358-362.
- [99] Grummt U., Erhardt S., *J. Mol. Struc-Theochem.*, **2004**, 685, 133-137.
- [100] Campá C., Camps J., Font J., de March P., *J. Org. Chem.*, **1987**, 52, 521-525.
- [101] Gottlieb H. E., Kotlyar V., Nudelman A., *J. Org. Chem.*, **1997**, 62, 7512-7515.
- [102] Collier H. B., *Anal. Biochem.*, **1973**, 56, 310-311.
- [103] Wachter H. N., Fried V., *J. Chem. Edu.*, **1974**, 51, 798-799.
- [104] Martell E. M., Smith R. M., Critical stability constants, volume 1: amino acids, Plenum Press, New York, **1974**, p. 2, ISBN 978-0306352119.
- [105] Yates P. C., *J. Chem. Ed.*, **2001**, 78, 770-771.

CHAPTER III: PEPTIDE SWITCHES BASED ON
GCN4 AS POTENTIAL METAL-DEPENDENT DNA
BINDERS

3.1 – Introduction

3.1.1 – Artificial dimerisation of transcription factors

Transcription factors are proteins whose functions are to regulate genes accessibility for RNA polymerase (see Figure 3.1). Transcription factor activity, achieved by binding to specific sequences of DNA, is regulated by allosteric mechanisms and binding of effector molecules.[1] Cell events resulting in binding of the effector molecule to a specific domain of the transcription factor protein, can result in structural preorganisation of a different domain for binding to the DNA recognition site, which can, in turn, activate or deactivate genes thus regulating transcription. Sensing cell events and signalling to the genomic machine is part of a wider signalling process, called mechanotransduction.[2]

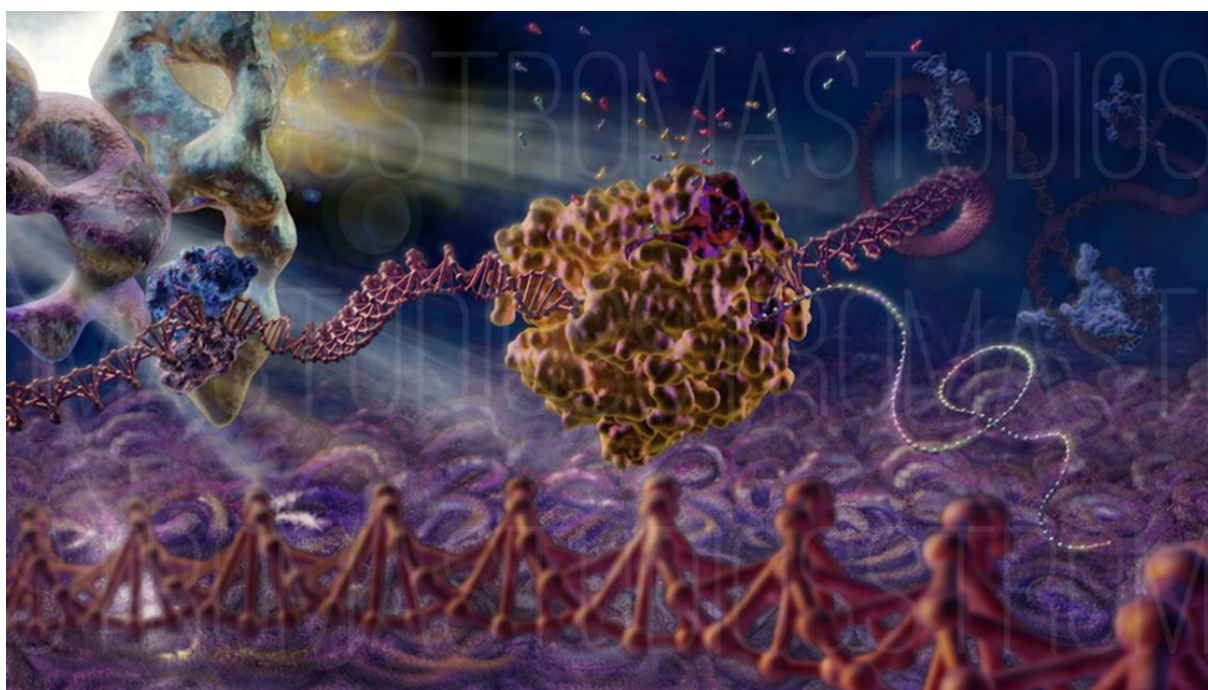


Figure 3.1 – Illustration showing the RNA polymerase II (yellow) during synthesis of the RNA (silver gray) templated on DNA (red). Several transcription factors (blue) are also represented along the DNA. Illustration was obtained from Stroma Studio.[3]

Transcription factors are usually classified, based on the DNA site they target and the binding mode (see Chapter 1).[4] GCN4 is a bacterial transcription factor involved in the

regulation of amino-acid synthesis. It contains 281 amino-acids and binds DNA recognition sites bearing the consensus sequence 5'-TGACTCA-3' or 5'-TGACGTCA-3' (AP1 and CRE, respectively [4,5]), as a homodimer. However, the minimum sequence required for sequence selective binding to DNA involves 60 amino-acids located towards the C-terminus of GCN4. This consists of 2 subdomains, the first is responsible for protein dimerisation (zipper domain),[6,7] whereas the second involves interaction with the DNA recognition site through specific contact (**bd**).[8,9] Both functions are indispensable to ensure strong DNA binding, however recent studies have shown that the zipper domain could be conveniently replaced by a simple covalent bond.[10]

Kim and co-workers prepared two synthetic peptides based on GCN4, and compared their specific DNA binding activity *in vitro*. The first one, GCN4-bZIP1, contains a fully conserved wild-type DNA binding domain (**bd** and zipper subdomain), and the second, GCN4-br1, contains a conserved **bd** and a glycine-glycine-cysteine motif toward its C-terminus. In the latter, dimerisation is achieved by formation of an intermolecular disulphide bond between unique cysteine residues, whereas glycine residues were added to increase the flexibility of the linkage (see Figure 3.2).[10] GCN4-bZIP1 and dimerised GCN4-br1 were both shown to bind DNA tightly. However, DNA binding of GCN4br1, but not GCN4-bZIP1, was reversed upon addition of the reducing agent dithiotreitol (DTT).[10] This work demonstrated how carefully designed chemical linkers, much less cumbersome than the 30 amino-acid zipper domain, could perform dimerisation, while retaining sequence-specific DNA binding *in vitro*. Following this pioneering work, different linkers have been used to artificially dimerise peptides derived from GCN4,[11-14] or related transcription factors.[15,16]

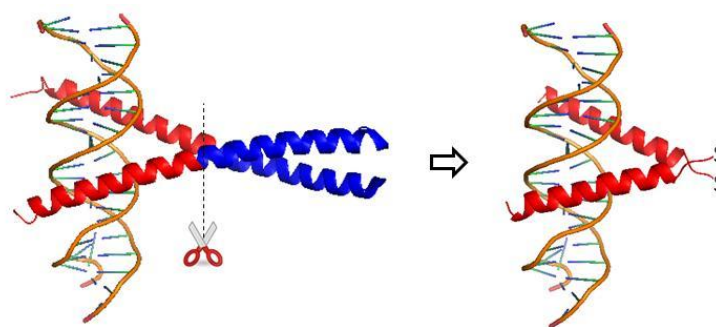


Figure 3.2 – Scheme illustrating the replacement of the zipper domain (blue) of GCN4, by a simple sulphur-sulphur bond between cysteines, introduced at the C-terminus of GCN4br1 (red). Kim and co-workers have shown that the resulting peptide dimer displays sequence-specificity similar to peptides with conserved zipper and **bd** domains; based on reference 10, and prepared using pdb file 1YSA.[9]

Schepartz and co-workers prepared a large variety of **terpy**-GCN4**bd** conjugates, where **terpy** units bearing a single thiol group are attached to the sulphur side-chains of cysteine bearing GCN4**bd** peptides with dithiol linkages. The conjugates displayed poor DNA affinity in the absence of metals or in the presence of Cu(II), which was expected to form only 1:1 complexes (**terpy**-GCN4**bd**)Cu(II). However addition of Fe(II), which forms 2:1 complexes (**terpy**-GCN4**bd**)₂Fe(II), led to dimerisation of two **terpy** units, each coupled to a single GCN4**bd**, and resulted in a significant increase in DNA binding.[13]

In a different approach, Mascareñas and co-workers successfully dimerised a synthetic analogue of GCN4**bd**, with an azobenzene linker by cysteine alkylation. The resulting GCN4**bd**-azobenzene-GCN4**bd** conjugate adopts the relaxed *trans*- conformation, associated with a 12-13 Å separation between the two peptide moieties, and poorly binds CRE DNA. Upon irradiation at 365 nm, the azobenzene linker switches to the *cis*-conformation, associated with a shorter 8-9 Å gap between peptide moieties and a higher affinity for CRE DNA (see Figure 3.3).[14] However, the process is not reversible. This constitutes the first report of DNA binding regulation of a GCN4**bd** derivative, by a conformational transition of the linker.

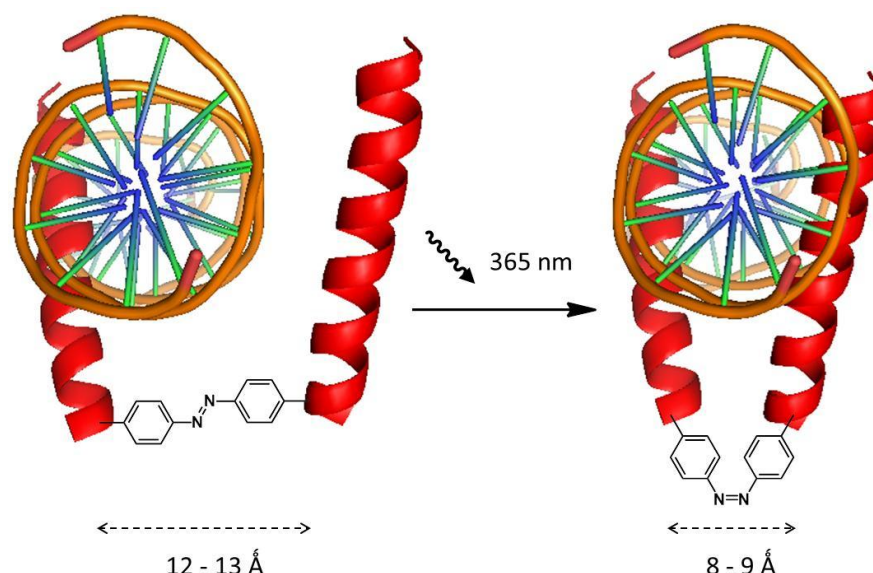


Figure 3.3 – Scheme illustrating DNA binding of GCN4bd dimerised with an azobenzene linker. Mascareñas and co-workers have shown that the sequence-specific DNA binding ability of this conjugate was significantly increased upon irradiation at 365 nm, as a result of *trans*-to-*cis* conformational transition of the azobenzene linker; based on reference 14, adapted from pdb file 1YSA.[9]

The design of artificial peptides which mimic the DNA binding of transcription factors, but which are regulated by alternative mechanisms, represent a developing area for biochemists.[17,18] The main interest is to artificially regulate the synthesis of proteins, and therefore, to take control of a cell's activity. Artificial transcription could allow for control of the cell defence, in the case of a deficient natural immune response, associated with protein mis-regulation (proteostasis deficiency), such as Alzheimer disease.[19]

3.1.2 – Aims of the chapter

As previously shown, metal coordination to polypyridine linkers results in a conformational transition that can alter the relative positioning of substituents. The aim is to take advantage of this repositioning in the context of sequence-selective DNA binding by artificial analogues of transcription factors (for review of the field see reference 20,21). For this, artificial peptides bearing the primary sequence of a transcription factor which binds

DNA as a homodimer (e.g. GCN4), could be artificially dimerised with the previously studied polypyridine linkers. The aim of this chapter is to design and prepare such peptide dimer conjugates, and to study their affinity for Cu(II) and Zn(II). DNA binding studies will be reported in the following Chapter 4.

Our interest was driven towards GCN4 due to the large volume of reports on its DNA binding characteristics as well as various mutagenesis studies. Based on these reports, a first peptide **GCN4bd1** was designed, which contains a high number of conserved residues from **GCN4bd**, in order to ensure strong DNA binding. A more minimalist peptide, **GCN4bd2**, was also designed. The polypyridine conjugates were prepared, purified, and their interaction with Cu(II) and Zn(II) analysed by UV and CD spectroscopy.

3.2 – Results and discussion

3.2.1 – Polypyridine-GCN4bd1 peptide conjugates

3.2.1.1 – Design of GCN4bd1 peptide

Aside from **terpy**-GCN4bd conjugates, Schepartz and co-workers prepared a **G28-bipy-G28** conjugate, in which two GCN4bd peptides are dimerised with a 5,5'-dimethyl-2,2'-bipyridine linker (**Me₂bipy**), by cysteine alkylation.[22] DNA binding properties of **G28-bipy-G28** were studied by band shift assays, however, the impact of metal coordination at the **bipy** linker on DNA binding was not investigated. The preparation and study of a similar **bipy** GCN4bd peptide dimer conjugate of similar size would be a good starting point. Therefore, a 29 amino-acid peptide based on GCN4bd, and named **GCN4bd1**, was designed (see Figure 3.4).



Figure 3.4 – Sequence alignment for wild type GCN4 (residue 224-263), **G28** peptide [22] and **GCN4bd1** (this work). Letter coding (top) represents residues positioning in α -helical heptad repeat. Domain highlighted in green is considered essential for binding to the CRE site, as it contains all residues believed to make contact with DNA. The nature of DNA contacts for each residues is indicated with the following terminology: b, direct to base ; w, via water to base ; p, direct to phosphate ; x, via water to phosphate.[23] Underlined residues are believed to be important for the thermal stability of the GCN4:DNA complex.[24] Cysteine residues highlighted in dark blue in **G28** and **GCN4bd1** are essential for dimerisation.

The **GCN4bd1** primary structure includes all residues from GCN4 which are believed to make contact with the DNA target site, in order to achieve strong DNA binding.[23] A cysteine residue, which is essential for dimerisation, was introduced either at position 253 (**G28**), or position 256 (**GCN4bd1**) relative to wild-type GCN4. In contrast to **G28**, a glycine-glycine motif separates cysteine from the GCN4 conserved residues in **GCN4bd1**, in order to bring sufficient flexibility in the dimerisation region,[14,22] and an additional glycine residue was added at the C-terminus to prevent epimerisation upon attachment of the first residue to the solid support during peptide synthesis.[25] Moreover, the leucine at position 253 was conserved in **GCN4bd1** in order to better position the cysteine residue (*vide infra*) and to promote hydrophobic interactions between peptides. The N- and C-terminus of the peptide were both capped to avoid potential interactions resulting from carboxylic- and amino-groups. Remarkably, **GCN4bd1** does not contain any aromatic residues, or other natural chromophore. Therefore, polypyridine-**GCN4bd1** conjugates require quantification based on the $\pi \rightarrow \pi^*$ transition of the polypyridine linkers.

Structural [9,23] and spectroscopic evidence [26] suggests that, in the presence of the DNA target site, GCN4 basic zipper peptides (**GCN4bz**) form uninterrupted α -helices. The C-

terminal region of DNA-bound **GCN4bd1** dimer can be interpreted as a dimeric parallel coiled coil motif, where the introduced cysteine (residue 256) would occupy a *g* position in the α -helix heptad terminology (see Figure 3.5).

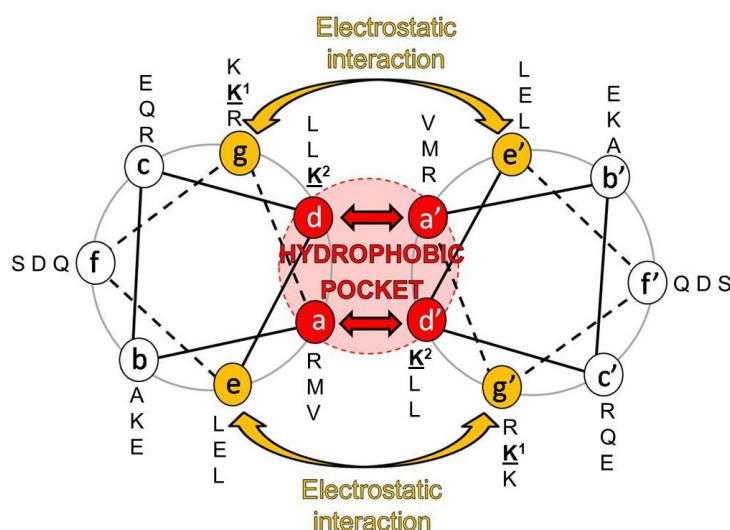


Figure 3.5 – Helical wheel diagram representing a dimeric coiled coil and relative positioning of residues (*a-g*) describing the α -helices heptad motifs. Positions of residues 243 to 263 (inside to outside starting from *a*) from wild type GCN4 are indicated. Positions occupied by cysteines in the **GCN4bd** analogues designed (this work) were underlined and marked as follow: **GCN4bd1** (¹), **GCN4bd2** (²); based on reference 4.

The C-terminal region of **GCN4bd1** contains three glycine residues which are known to disfavour helix formation. Therefore, it is more likely that this short section constitutes a highly flexible domain, incompatible with the coiled coil representation.

3.2.1.2 – Synthesis of **GCN4bd1** peptide and polypyridine conjugates

GCN4bd1 peptide was prepared by Fmoc solid-phase peptide synthesis on rink amide MBHA resin using a CEM Liberty 1 synthesiser, as described in the experimental section (see section 3.3). After cleavage from the solid support, the peptide was purified by HPLC and characterised by ESI, MALDI-TOF mass spectrometry (MS) and analytical HPLC.

As for the model compounds, the dimerisation of **GCN4bd1** with **Me₂pyr** by alkylation of cysteine, was first attempted. The reaction progress was monitored by the Ellman's assay (see Appendices). A plot of absorbance at 412 nm as a function of reaction time, indicates that after 20 hours 75 % of the thiol groups have been consumed, and that the reaction is near-completion (see Figure 3.6). However, this does not indicate whether thiol groups were converted to thioether or disulphide linkages.

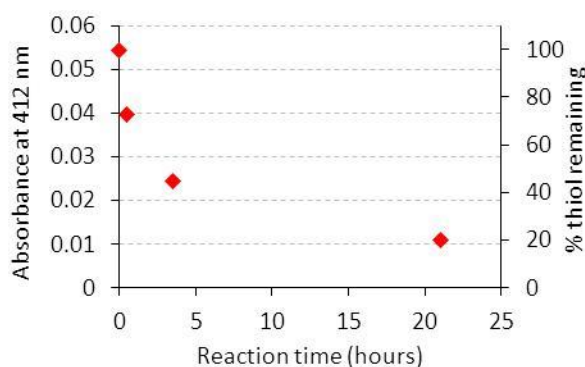


Figure 3.6 – Plot monitoring the available thiol concentration during **pyr(GCN4bd1)₂** synthesis, based on the Ellman's assay. Absorbance at 412 nm (primary Y axis, left), and remaining thiol percentage (secondary Y axis, right) were estimated at different reaction times.

GCN4bd1 peptide dimerised with the **Me₂pyr** linker has a very similar retention time to the oxidised dimer (**GCN4bd1**)₂, under our HPLC elution conditions (C18 column, water/acetonitrile/TFA gradients) (see Figure 3.7B). Therefore, DTT was added to the reaction mixture, in order to reduce the disulphide bond from (**GCN4bd1**)₂, thus reforming **GCN4bd1** starting material (see Figure 3.7A and C). HPLC of the reaction mixture recorded after reduction displays enhanced peak resolution (see Figure 3.7D), thus allowing for purification and qualitative estimation of **GCN4bd1** conversion to **pyr(GCN4bd1)₂**.

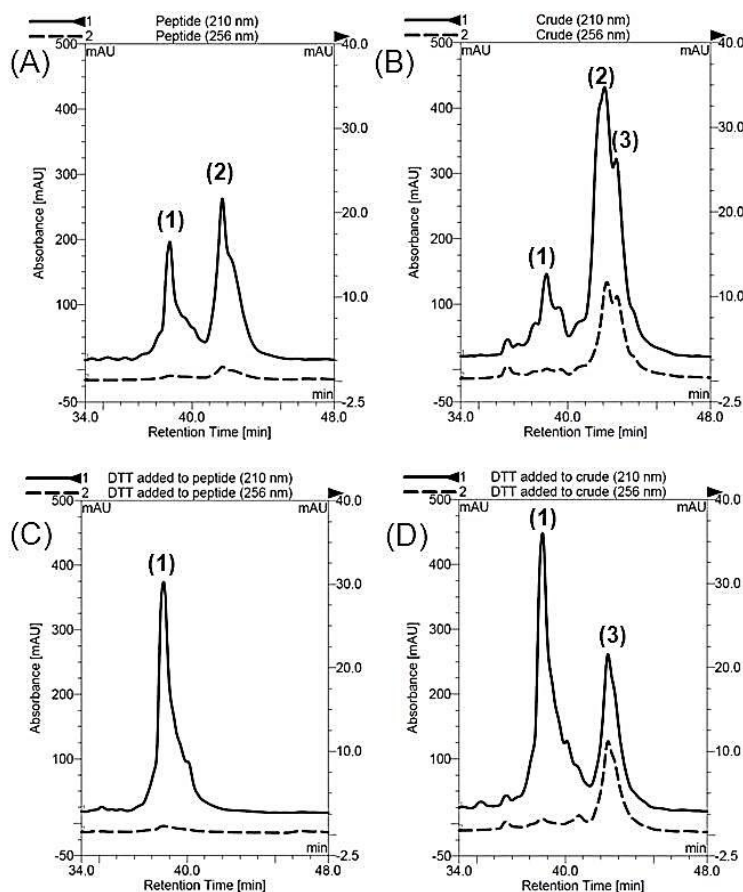


Figure 3.7 – Analytical C18 RP-HPLC profile for a solution containing: (left) **GCN4bd1** partly oxidised before (A) or after DTT addition (C); (right) crude reaction mixture for **pyr(GCN4bd1)₂** preparation before (B) or after DTT addition (D). All solutions were eluted with gradients of 0 to 30 % acetonitrile in water (+0.05 % TFA constant) over 60 min and were monitored at 210 nm (—) and 256 nm (---). Peaks observed were assigned to: (1) **GCN4bd1** monomer, (2) (**GCN4bd1**)₂ oxidised dimer and (3) the product **pyr(GCN4bd1)₂** (based on electrospray spectra recorded for collected fractions).

Bipy(GCN4bd1)₂ and **terpy(GCN4bd1)₂** were similarly prepared by reaction of excess **GCN4bd1** with **bipy-Br₂** or **terpy-Br₂**. However, the preparation of the conjugate **terpy(GCN4bd1)₂** required heating of the reaction mixture, the HPLC analysis of which indicated the formation of numerous side-products (the retention times were not reproducible). The purification of the conjugate was attempted using HPLC and various elution sequences (involving both gradients and isocratic mixtures of water/acetonitrile/TFA), but the conjugate could not be efficiently purified. Electrospray and analytical HPLC recorded are both consistent with a mixture of the expected compound **terpy(GCN4bd1)₂** and the analogous monomer, **terpy(GCN4bd1)**, in a ~70:30 ratio (see Figure 3.8). Despite this, the mixture was used in the subsequent analysis described in Chapter 3 and 4.

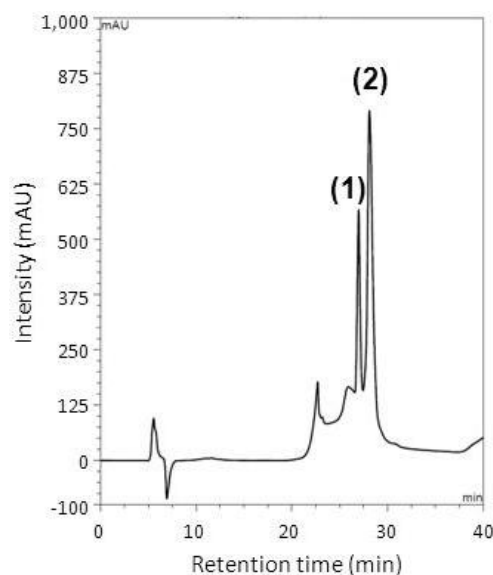


Figure 3.8 – Analytical C18 RP-HPLC profile for **terpy(GCN4bd1)₂** after rounds of HPLC purification. Solution was eluted with 0 to 50 % gradient of acetonitrile in water (+0.05 % TFA constant) over 40 min (absorbance monitored at 210 nm). Main peaks were assigned as (1) **terpy(GCN4bd1)** monomer and (2) **terpy(GCN4bd1)₂** dimer by ESI-MS.

Conjugates (**pyr(GCN4bd1)₂**, **bipy(GCN4bd1)₂**) and mixture relative to (**terpy(GCN4bd1)₂**), were characterised by ESI-MS, even though the isotopic distribution of the high charge peaks (+7 - +14) could not be resolved, and analytical HPLC. Attempts to characterise the conjugates by MALDI-TOF led to either no or a weak signal at the expected masses.

3.2.1.3 – Cu(II) addition to GCN4bd1-polypyridine conjugates monitored by UV spectroscopy

Here, the UV spectra of **pyr(GCN4bd1)₂**, **bipy(GCN4bd1)₂** and **terpy(GCN4bd1)₂** are reported both in the absence and presence of Cu(II). The UV spectrum of a 10 μ M solution of **pyr(GCN4bd1)₂** at pH 8 displays one transition with λ_{max} at 276 nm ($\epsilon_{276 \text{ nm}}$ 5,500 $\text{M}^{-1} \text{cm}^{-1}$), assigned as a $\pi \rightarrow \pi^*$ transition. Addition of 1 equivalent of Cu(II), resulted in a slight increase of the absorbance at 276 nm ($\epsilon_{276 \text{ nm}}$ 6,200 $\text{M}^{-1} \text{cm}^{-1}$) and 255 nm ($\epsilon_{255 \text{ nm}}$: 3,600 \rightarrow 5,600 $\text{M}^{-1} \text{cm}^{-1}$) (see Figure 3.9).

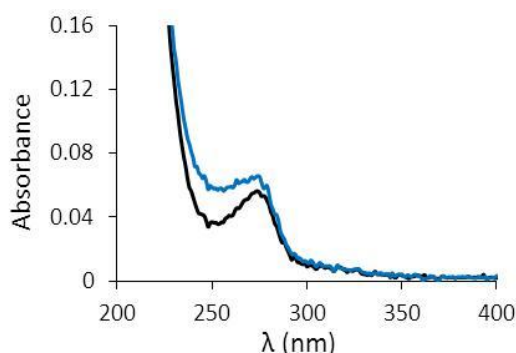


Figure 3.9 – UV spectra of a 10 μM solution of **pyr(GCN4bd1)₂** in 20 mM phosphate buffer pH 8, before (—) and after (—) addition of 1 eq. CuCl_2 .

The UV spectrum of a 4.5 μM solution of **bipy(GCN4bd1)₂** recorded in solution buffered at pH 8, displays two transitions with λ_{max} at 257 ($\epsilon_{257 \text{ nm}}$ 13,500 $\text{M}^{-1} \text{cm}^{-1}$) and 310 nm ($\epsilon_{310 \text{ nm}}$ 20,100 $\text{M}^{-1} \text{cm}^{-1}$), assigned as $\pi \rightarrow \pi^*_1$ and $\pi \rightarrow \pi^*_2$ transitions. Gradual addition of Cu(II) resulted in the steady decrease of the absorbance at 310 nm (N.B. the absorbance at 310 nm is similar in the presence of 0 and 1 equivalent of CuCl_2 , but lower between these two values), and an increase in the absorbance at 313 and 328 nm ($\epsilon_{313 \text{ nm}}$: 19,700 \rightarrow 21,100; $\epsilon_{328 \text{ nm}}$: 12,200 \rightarrow 17,100 $\text{M}^{-1} \text{cm}^{-1}$) (see Figure 3.10A). The transition is characterised by the absence of a clear isosbestic point. The UV spectrum of a 2.1 μM solution of **bipy(GCN4bd1)₂** recorded in the presence of 0.3 mM EDTA, displays a profile similar to that recorded in buffer pH 8, but both $\pi \rightarrow \pi^*$ maxima are slightly shifted towards lower wavelength (243 and 304 nm) (see Figure 3.10B).

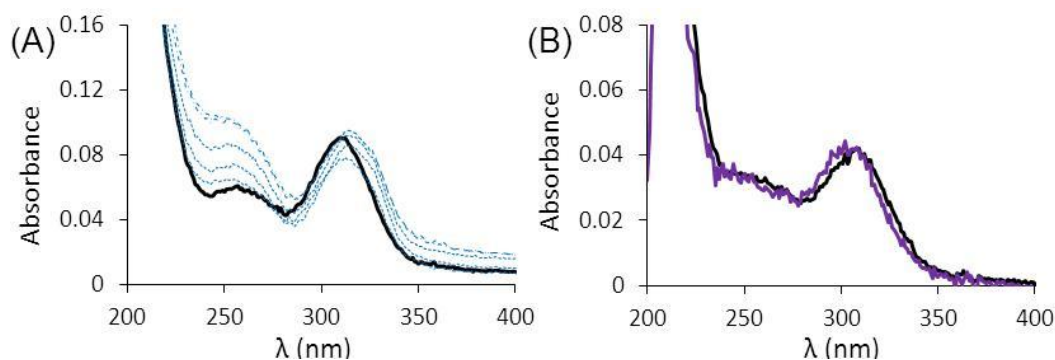


Figure 3.10 – (A) Cu(II) titration of a 4.5 μM solution of **bipy(GCN4bd1)₂** in 20 mM phosphate buffer pH 8, monitored by UV spectroscopy (0.3 eq. step): (—) 0 eq., (····) between 0 and 1 eq., (— · —) more than 1 eq. CuCl_2 added. (B) UV profile of 2.1 μM **bipy(GCN4bd1)₂** recorded in either 20 mM phosphate pH 8 (—), or 0.3 mM EDTA solution (—).

The UV spectrum of a 1.1 μM solution of **terpy(GCN4bd1)₂** recorded in solution buffered at pH 8, displays one transition with λ_{max} at 299 nm ($\epsilon_{299 \text{ nm}}$ 21,000 $\text{M}^{-1} \text{cm}^{-1}$), assigned as $\pi \rightarrow \pi^*_1$ transition. Addition of 1 equivalent Cu(II) resulted in the steady decrease of the absorbance at 299 nm ($\epsilon_{299 \text{ nm}}$ 16,000 $\text{M}^{-1} \text{cm}^{-1}$), and an increase of the absorbance at 348 nm ($\epsilon_{348 \text{ nm}}$: 4,500 \rightarrow 6,700 $\text{M}^{-1} \text{cm}^{-1}$) (see Figure 3.11).

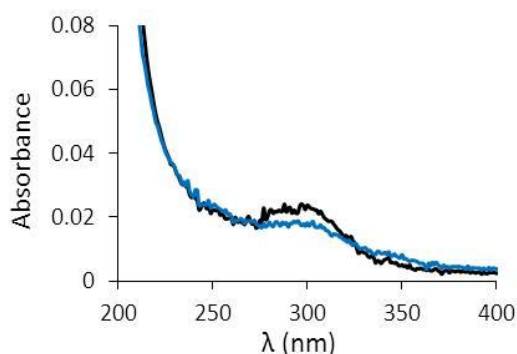


Figure 3.11 – UV spectrum of a 1.1 μM **terpy(GCN4bd1)₂** solution in 20 mM phosphate buffer pH 8, before (—) and after (—) addition of 1 eq. Cu(II).

The absorbance profile of **pyr(GCN4bd1)₂** is very similar to that recorded for **cys₂pyr** and **pyr-GS₂** under similar conditions, both in the absence and presence of CuCl₂ (compare Figures 3.12C, D and 3.9). In contrast, the UV spectrum recorded for **bipy(GCN4bd1)₂**, contains bands shifted 12-15 nm towards higher wavelength compared to **cys₂bipy** and **bipy-GS₂**. However, the UV spectrum of **bipy(GCN4bd1)₂** recorded in the presence of EDTA, displays maxima at slightly lower wavelength, suggesting that our **bipy(GCN4bd1)₂** stock solution may be contaminated with low concentrations of metal ions. The addition of one equivalent of Cu(II) results in only a minor shift, but the resulting spectrum is similar to that recorded for **cys₂bipy** and **bipy-GS₂** in the presence of Cu(II), suggesting formation of a similar **bipy-Cu** species (see Figure 3.25A, 3.32A and 3.10). The spectra of **terpy(GCN4bd1)₂** recorded in the absence of CuCl₂ is similar to that recorded for **cys₂terpy** and **terpy-GS₂** (compare Figure 3.25B, 3.32B and 3.11). Upon addition of CuCl₂, the UV spectrum displays a new band at 348 nm, consistent with the spectra recorded for **cys₂terpy-**

Cu and **terpy-GS₂-Cu**, but less intense, suggesting **terpy(GCN4bd1)₂** binds Cu(II) less tightly than **terpy-GS₂** and **cys₂terpy**.

UV analysis of impure **bipy(GCN4bd1)₂** and **terpy(GCN4bd1)₂** are not conclusive, due to the small shift of the $\pi \rightarrow \pi^*$ band upon Cu(II) addition (for which high binding would be expected based on model compounds studies). It remains unclear if these conjugate linkers are undergoing a significant conformational rearrangement upon metal-addition.

3.2.2 – Polypyridine GCN4bd2 short peptide dimer

3.2.2.1 – Design of GCN4bd2 peptide

Experiments described in Chapter 4, indicate that DNA binding of **GCN4bd1** polypyridine conjugates is not influenced by metal ion coordination. In order to understand this result, one needs to consider the global flexibility within the conjugates, which might buffer conformational change occurring at the linker, preventing peptide reorientation. Another explanation would be that peptide moiety reorientation would occur, but would not significantly influence DNA binding. Therefore, a second shorter peptide, **GCN4bd2**, was designed in order to try to address these issues. Analogous polypyridine conjugates bear DNA binding domains proximal to the linker and highly sensitive to its conformation (unlike **GCN4bd1**). Glycine introduced between cysteine and the conserved GCN4bd (as for **GCN4bd1**) brings unnecessary flexibility and so have been removed in our second design.

Structures of **GCN4bz** homodimers bound to AP1 [9] and CRE [23] DNA target sites have been reported (pdb code 1YSA and 2DGC, respectively), and allow for a qualitative estimation of inter-strand distances. As can be seen in Figure 3.12, distances separating the γ -atoms increase on approaching the DNA, but remain minimal between residues located within the core of the homodimer, which correspond to position *a* and *d* of the homodimer coiled coil (see Figure 3.5). The ideal substitution position corresponds to a residue for which this

distance would match that previously estimated for **bipy** and **terpy** linkers (see Chapter 2), allowing us to identify the most suitable site cysteine introduction.

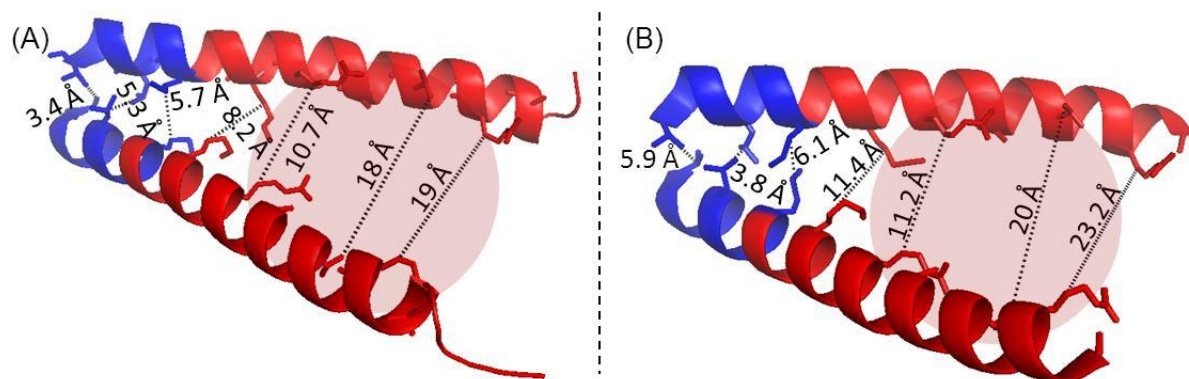


Figure 3.12 – Illustration of peptide separation in a GCN4**bz** homodimer (residue 229 to 257, shown in red (**bd**), zipper domain shown in blue) complexed with AP1 (A), and CRE (B) DNA (double helix was replaced by brown circles for clarity). Distances correspond to the separation between γ -atoms from residues located in positions *a* and *d* from α -helical heptad repeats; adapted from pdb files 1YSA [9] and 2DGC,[23] respectively.

In order to retain sequence-specific DNA binding, the majority of residues believed to make direct contact with DNA need to be conserved. Additional studies by Kim and co-workers estimate the impact of residue truncation within the GCN4**bd** on DNA binding properties of relative artificial dimers.[24] Even though crystal structure studies of GCN4**bz**:CRE complex indicate that Lysine246 makes water mediated contact with base G-3 of the palindromic CRE site,[23] Kim et al. report that it is not essential for sequence-specific binding. Indeed, artificially dimerised **GCN4br6^{SS}**, which lacks residue Lysine246, was shown to bind selectively to the CRE target site.[24] Moreover it has been suggested that the uncharged side-chain of leucine247, can clash with the DNA backbone in the GCN4:CRE complex crystal structure, thus destabilising the complex.[23] Residues 248 to 252, which precede the zipper domain (QRMKQ), are not believed to make contact with DNA, but were suggested to be important for the thermal stability of the GCN4:DNA complex [24]. In order to give the keys of the GCN4**bd** orientation to polypyridine linkers, residues relative to the hinge area needs to be suppressed. Based on these requirements, the peptide **GCN4bd2** was designed (see Figure 3.13).

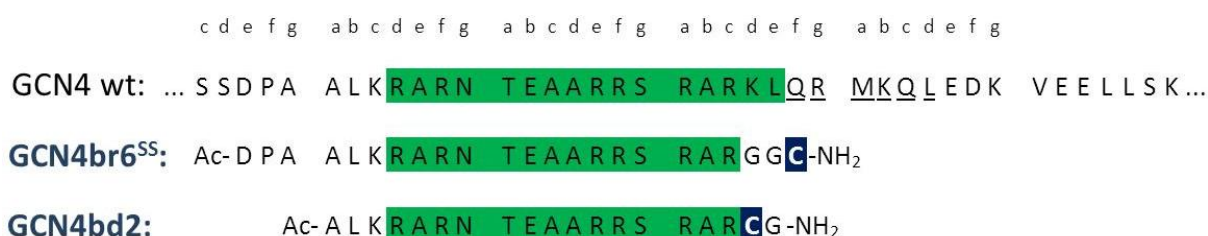


Figure 3.13 – Sequence alignment for wild type GCN4 (residue 224-263), **GCN4br6^{SS}** (shortest artificial GCN4bd dimer reported to selectively bind CRE sites[24]), and **GCN4bd2** (this work). Domains highlighted in green contain all residues believed to make contact with DNA.[23] Underlined residue are believed to be important for thermal stability of the GCN4:DNA complexes [24]. Cysteine residues (dark blue) are essential for dimerisation.

GCN4bd2 contains 17 conserved residues from GCN4 (corresponding to the domain located between residue 229 and 245) including the GCN4bd, but lacks the flexible GCN4 linker domain. The location of cysteine (equivalent to lysine246 of native GCN4) is such that the two thiols from the side chains should be positioned ca. 5.7 Å apart, a more suitable distance for the *cis*- form of the **Me₂terpy** linker unit, in the folded GCN4:AP1 complex (measurements based on crystal structure pdb code 1YSA [9]), or 7.1 Å apart, in the GCN4:CRE complex, suitable for either a **Me₂bipy** (*cis*- or *trans*-) or for the *cis*- form of **Me₂terpy** (measurements based on crystal structure pdb code 2DGC).[23] Different distances in the two peptide dimer DNA complexes may indicate that peptide moieties retain the same specific contacts to DNA in both complexes, but the presence of an extra central base pair in the CRE site, results in a slightly larger distance separating the two peptides (see Figure 3.14).

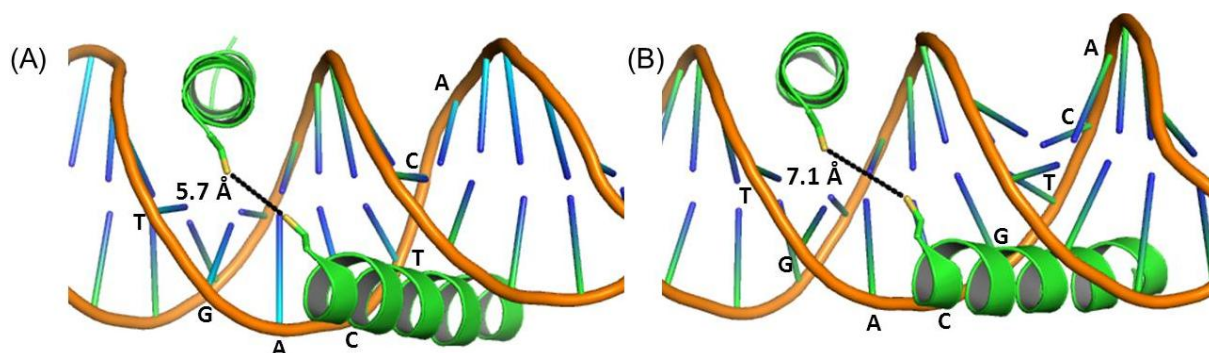


Figure 3.14 – Cartoon illustrating distances separating two sulphur atoms from a theoretical cysteine, introduced as the best aligned rotamer, in place of residue lysine246 of the GCN4 homodimer complexed with AP1 (A) or CRE duplex (B). DNA base from the semi-palindromic (A) and palindromic sites (B) are marked with one-letter symbol on one strand; adapted from pdb files 1YSA [9] and 2DGC,[23] respectively.

3.2.2.2 – Preparation and characterisation of GCN4bd2 and conjugates

GCN4bd2 was prepared by Fmoc solid-phase peptide synthesis on rink amide MBHA resin using a CEM Liberty 1 synthesiser. After cleavage from the solid support, the peptide was purified by HPLC and characterised by ESI, MALDI-TOF mass spectrometry and analytical HPLC.

In order to estimate the propensity of peptide oxidation under the coupling conditions, a solution containing **GCN4bd2** was incubated at pH 8.0 and room temperature for 16 hours, and analytical HPLC was recorded indicating only partial oxidation (see Figure 3.15A). In contrast, similar incubation performed in the presence of 0.5 equivalent of **bipy-Br₂** lead to near total conversion to **bipy(GCN4bd2)₂**, as indicated by HPLC and ESI-MS (see Figure 3.15B and C).

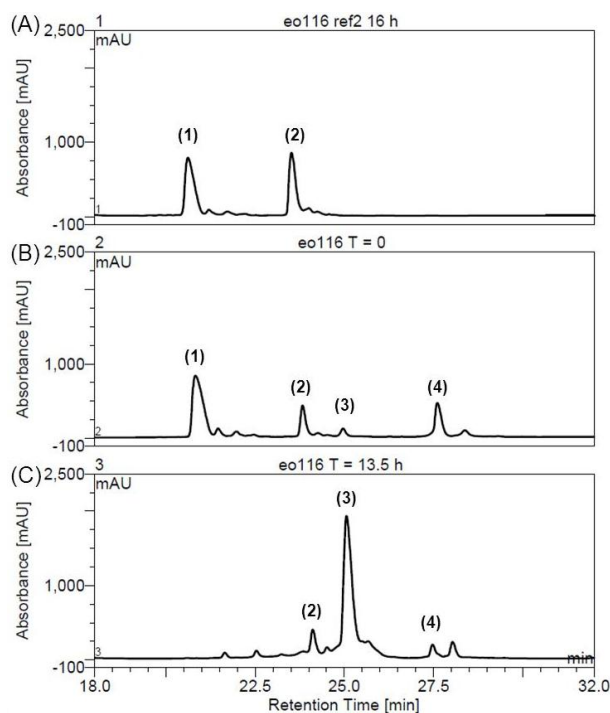


Figure 3.15 – Analytical C18 RP-HPLC profile for a 0.48 mM solution of **GCN4bd2** in acetonitrile and 85 mM Tris.HCl buffer pH 8.0 (1:1) incubated at room temperature in the absence (A), or presence of 0.5 eq. of **bipy-Br₂** (B and C). Incubation times were either 16 hours (A), two minutes (B), or 13.5 hours (C). 100 μ L of solution was injected and eluted with a 0 to 30 % gradient of acetonitrile in water (+0.05 % TFA constant) over 40 min (absorbance monitored at 210 nm). Peaks observed were assigned to: (1) **GCN4bd2**, (2) **(GCN4bd2)₂** oxidised dimer, (3) **bipy(GCN4bd2)₂** (based on ESI-MS data recorded on collected fractions). No fraction relative to peak (4) was collected, and it might corresponds to an intermediary species, such as the monomer **bipy(GCN4bd2)**.

Peptide conjugates **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** were prepared, following procedures similar to those previously described for the preparation of **GCN4bd1** dimer conjugates. Dimerisation due to disulphide bond formation at pH 8.0 was slower for **GCN4bd2**, than previously observed for **GCN4bd1**. In contrast, alkylation of cysteine by reaction with dibromo- linker was faster and cleaner (see Figure 3.15 and section 3.2.1.2). Unlike **GCN4bd1**, the **GCN4bd2** peptide dimer conjugates and oxidised dimer have sufficiently different retention times under these HPLC conditions, such that they could be readily be separated, without need for additional reducing agents (see Figure 3.15C). They were obtained in high purity, as indicated by analytical HPLC.

For comparison purposes and for use as a control dimer, the oxidised dimer (**GCN4bd2**)₂, where the two peptides are linked with a disulphide bond, was also prepared by incubation of **GCN4bd2** at pH 8.0 and 313 K overnight. Conversion was monitored by analytical HPLC and was found to be greater than 97 %. ESI-MS of the same mixture gave the expected mass, and the stock solution was used directly, without further purification. ESI-MS spectra recorded for (**GCN4bd2**)₂, **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂**, all display the expected charge envelopes, with appropriate isotopic distributions.

3.2.2.3 – Cu(II) and Zn(II) addition to **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** monitored by UV spectroscopy

The UV spectrum of a $5 \pm 0.6 \mu\text{M}$ solution of **bipy(GCN4bd2)₂** in 20 mM phosphate buffer pH 7.4 displays two transitions with λ_{max} at 299 ($\epsilon_{299 \text{ nm}} 20,000 \text{ M}^{-1} \text{ cm}^{-1}$) and 246 nm ($\epsilon_{246 \text{ nm}} 16,900 \text{ M}^{-1} \text{ cm}^{-1}$), assigned as $\pi \rightarrow \pi^*_1$ and $\pi \rightarrow \pi^*_2$ transitions.[27] The analogous spectrum of **terpy(GCN4bd2)₂** shows a peak with λ_{max} 298 nm ($\epsilon_{298 \text{ nm}} 21,000 \text{ M}^{-1} \text{ cm}^{-1}$) attributed to $\pi \rightarrow \pi^*_1$, however, the $\pi \rightarrow \pi^*_2$ transition overlaps with that for the peptide bond. These bands shift to lower energy upon Zn(II) and Cu(II) complexation, allowing estimation

of an apparent binding constant.[28] Aliquots of a stock solution of ZnCl_2 were titrated into a $5\ \mu\text{M}$ solution of **bipy(GCN4bd2)₂** or **terpy(GCN4bd2)₂** in 20 mM phosphate buffer pH 7.4, up to 3 equivalent of ZnCl_2 . This resulted in the steady decrease in the absorbance at 299 (**bipy(GCN4bd2)₂**) and 298 nm (**terpy(GCN4bd2)₂**), and an increase in the absorbance at 313 and 320 nm (**bipy(GCN4bd2)₂**), or at 328 and 341 nm (**terpy(GCN4bd2)₂**), respectively (see Figure 3.16). The isosbestic point at 308 (**bipy(GCN4bd2)₂**) and 318 nm (**terpy(GCN4bd2)₂**) are consistent with the clean conversion to a metal bound complex.

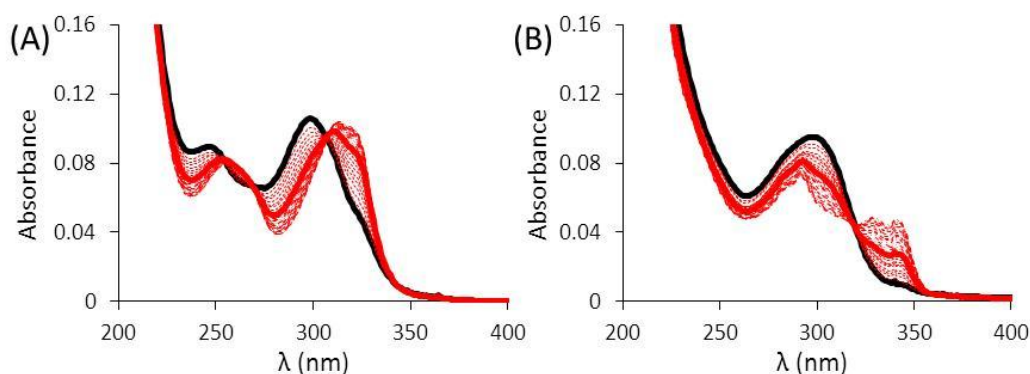


Figure 3.16 – UV spectra for the Zn(II) titration of peptide conjugates. ZnCl_2 was added to solutions containing $5\ \mu\text{M}$ of either **bipy(GCN4bd2)₂** (A) or **terpy(GCN4bd2)₂** (B) in 20 mM phosphate buffer pH 7.4. (—) 0 eq., (—) 1 eq., (····) between 0 and 1 eq., or (— · —) more than 1 eq. ZnCl_2 added.

A plot of the absorbance of the new transitions versus Zn(II) concentration was fitted to a 1:1 binding model (see Figure 3.17), and the extinction coefficients of the peptide dimers and the resulting Zn(II) complexes determined and reported in Table 3.1. Taking into account the competitive metal ion binding of the phosphate buffer employed in these experiments,[29] binding constants, $\log K_M$, were calculated to be 6.92 ± 0.05 for formation of the $\text{Zn-bipy(GCN4bd2)_2}$ complex, and 5.67 ± 0.05 for the analogous complex with **terpy(GCN4bd2)₂**.

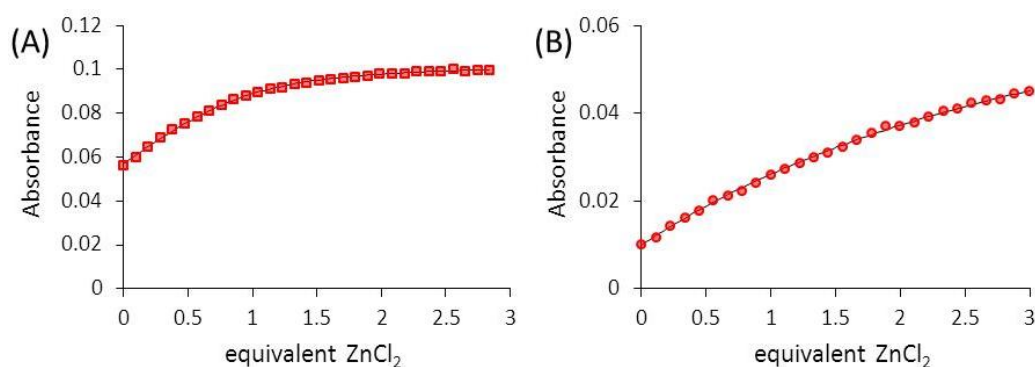


Figure 3.17 – Plot of absorbance monitored at 320 nm (■) for **bipy(GCN4bd2)₂** (A) and 341 nm (●) for **terpy(GCN4bd2)₂** (B) versus the equivalence of ZnCl_2 . (—) Represents best fit for formation of a 1:1 metal:peptide dimer conjugate complex.

Table 3.1 – Summary of the UV data obtained for Zn(II) coordination to polypyridyl peptide conjugates, **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂**

Pep. conjugate	λ (nm)	ϵ_{ML} ($\text{M}^{-1} \text{cm}^{-1}$)	K_{app} (M^{-1})	K_{M} (M^{-1})	R^2
bipy(GCN4bd2)₂	320	$1.95 \pm 0.01 \text{ E} +04$	$1.38 \pm 0.16 \text{ E} +06$	$8.29 \pm 0.95 \text{ E} +06$	0.9986
terpy(GCN4bd2)₂	341	$1.89 \pm 0.06 \text{ E} +04$	$7.74 \pm 0.86 \text{ E} +04$	$4.66 \pm 0.52 \text{ E} +05$	0.9989

The addition of increasing aliquots of CuCl_2 into a $5 \pm 0.6 \text{ } \mu\text{M}$ solution of **bipy(GCN4bd2)₂** or **terpy(GCN4bd2)₂** in 20 mM phosphate buffer pH 7.4, resulted in a decrease in the absorbance at 299 nm (**bipy(GCN4bd2)₂**) and 298 nm (**terpy(GCN4bd2)₂**), and an increase in the absorbance at 317 and 328 nm (**bipy(GCN4bd2)₂**) and at 335 and 348 nm (**terpy(GCN4bd2)₂**). Again the isosbestic points at 309 (**bipy(GCN4bd2)₂**) and 320 nm (**terpy(GCN4bd2)₂**) are consistent with clean formation of the complex (see Figure 3.18).

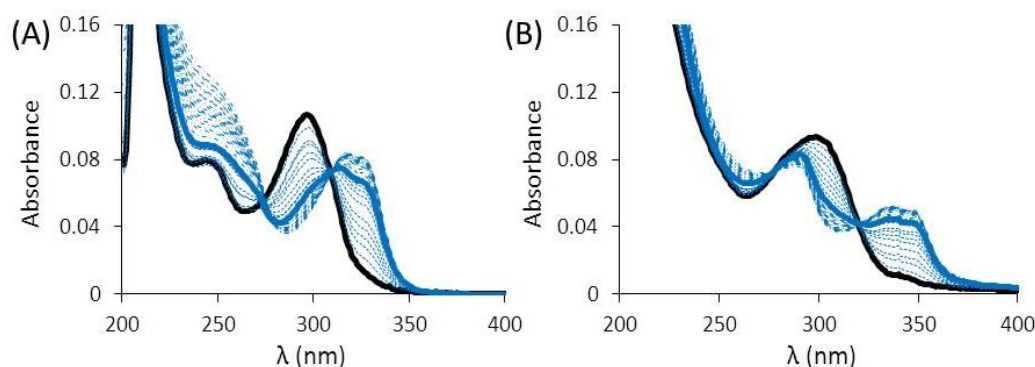


Figure 3.18 – UV spectra for the Cu(II) titration of peptide conjugates. Cu(II) was added to solutions containing $5 \text{ } \mu\text{M}$ of either **bipy(GCN4bd2)₂** (A) or **terpy(GCN4bd2)₂** (B) in 20 mM phosphate buffer pH 7.4. (—) 0 eq., (—) 1 eq., (·····) between 0 and 1 eq., (---) more than 1 eq. CuCl_2 added. For (A) buffer concentration was 100 mM and 20 mM L-glycine was added as competitor.

The plot of absorbance as a function of increasing Cu(II) concentration was fitted to a 1:1 binding equation, from which extinction coefficients were determined (see Figure 3.19 and Table 3.2). Titration of **bipy(GCN4bd2)₂** with Cu(II) had to be performed in the presence of a competitor (20 mM L-glycine), in order to accurately estimate the Cu(II) affinity, see Table 3.2. Formation constants, $\log K_M$, were determined to be 13.17 ± 0.07 for the Cu-**bipy(GCN4bd2)₂** complex, and 7.87 ± 0.05 (A_{335}), 7.94 ± 0.06 (A_{348}) for the Cu-**terpy(GCN4bd2)₂** complex.

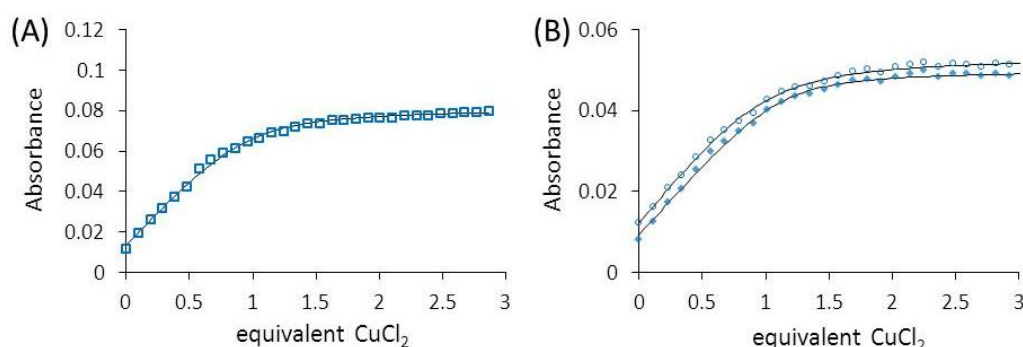


Figure 3.19 – Plot of absorbance monitored at 328 nm (\square) for **bipy(GCN4bd2)₂** (A), 335 nm (\circ) and 348 nm (\blacklozenge) for **terpy(GCN4bd2)₂** (B) versus the equivalence of CuCl₂. (—) Represents best fit for formation of a 1:1 metal:peptide dimer conjugate binding ratio.

Table 3.2 – Summary of the UV data obtained for Cu(II) coordination to polypyridyl peptide conjugates, **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂**.

Pep. conjugate	λ (nm)	ϵ_{ML} (M ⁻¹ cm ⁻¹)	K_{app} (M ⁻¹)	K_M (M ⁻¹)	R^2
bipy(GCN4bd2)₂	328	$1.55 \pm 0.01 \text{ E} +04$	$2.94 \pm 0.47 \text{ E} +06$	$1.48 \pm 0.24 \text{ E} +13^a$	0.9966
	335	$1.21 \pm 0.01 \text{ E} +04$	$2.25 \pm 0.30 \text{ E} +06$	$7.37 \pm 0.97 \text{ E} +07$	0.9973
terpy(GCN4bd2)₂	348	$1.14 \pm 0.01 \text{ E} +04$	$2.67 \pm 0.36 \text{ E} +06$	$8.72 \pm 1.18 \text{ E} +07$	0.9972

^a Titration performed in the presence of 20 mM L-glycine as competitor.

The continuous variation method was used in order to confirm the formation of 1:1 complexes.[30] The total concentration of peptide conjugate plus metal ion was kept constant and equal to 10 μ M, and absorbance spectra were recorded for mixtures with different peptide conjugate/metal ratios. Absorbance maxima for metal complexes, before and after subtraction

of the free ligand contribution,[30] was plotted versus the total fraction of peptide conjugate in solution χ (**PJ**)_T. Corrected jobplots indicate that in all four cases the main complex is formed in a 1:1 peptide conjugate:metal ion ratio (see Figure 3.20).

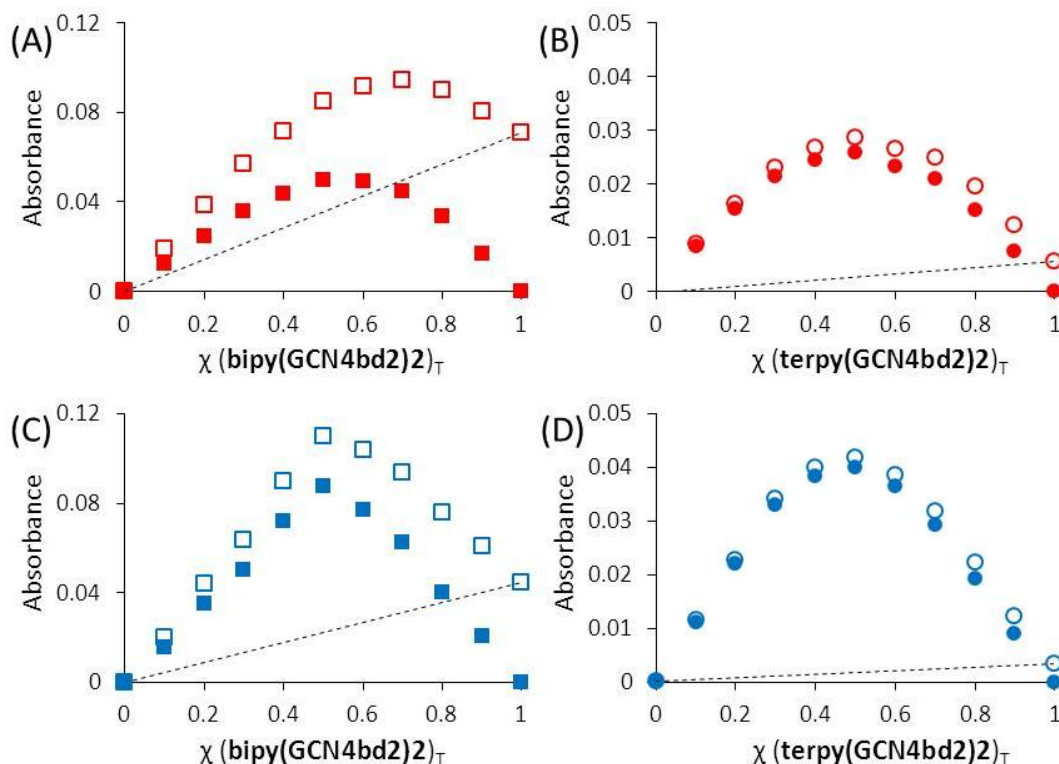


Figure 3.20 – Jobplot based on UV spectra for mixtures of metal and peptide conjugates (C_{total} constant and equal to 10 μM), recorded in 20 mM phosphate buffer pH 7.4: (A) ZnCl₂ with **bipy**(GCN4bd2)₂, (B) ZnCl₂ with **terpy**(GCN4bd2)₂, (C) CuCl₂ with **bipy**(GCN4bd2)₂ and (D) CuCl₂ with **terpy**(GCN4bd2)₂. Empty symbols represent plot of absorbance (monitored at 320 (□), 328 nm (□) for **bipy**(GCN4bd2)₂ and 341 (○), 348 nm (○) for **terpy**(GCN4bd2)₂ versus the molar fraction of peptide conjugate. Filled symbol represent the absorbance corrected for contribution from the metal-free peptide conjugate, represented by the dotted line (---).

In contrast to the UV spectra recorded for **bipy**(GCN4bd1)₂ and **terpy**(GCN4bd1)₂, significant $\pi \rightarrow \pi^*$ redshift are observed for **bipy**(GCN4bd2)₂ and **terpy**(GCN4bd2)₂ upon addition of CuCl₂ or ZnCl₂ consistent with previously described *trans*-to-*cis* conformational transitions (see Chapter 2). For the latter only, UV spectra in the absence or presence of metal ion are similar to those recorded for **bipy**-GS₂ and **terpy**-GS₂, respectively (see Chapter 2). Similarly, a plot of absorbance maxima relative to the new $\pi \rightarrow \pi^*$ transition were used to estimate the affinity of **bipy**(GCN4bd2)₂ and **terpy**(GCN4bd2)₂ for Cu(II) and Zn(II), by

fitting to 1:1 binding models (supported by the jobplot, see Figure 3.20). Affinity constants calculated for **bipy(GCN4bd2)₂:Zn(II)** and **terpy(GCN4bd2)₂:Cu(II)** polypyridine conjugates, are very similar to those previously reported for analogous glutathione conjugates (see table 3.3 and 3.4 and Figure 3.21). However, unlike for **terpy-GS₂**, fitting of absorbance data upon addition of CuCl₂ to **terpy(GCN4bd2)₂** for both the 335 and 348 $\pi \rightarrow \pi^*$ transition, results in similar binding constants (within error), suggesting formation of a single Cu(II):**terpy(GCN4bd2)₂** species.

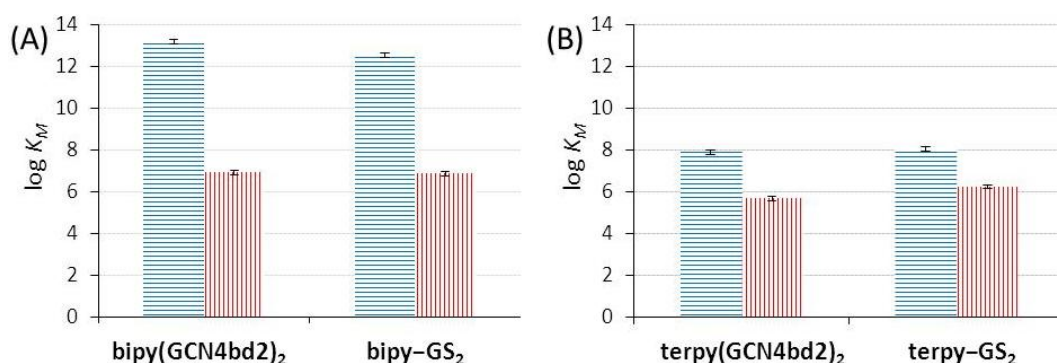


Figure 3.21 – Histogram representing the log K_M values calculated for **GCN4bd2** and **GS** polypyridine conjugates for Cu(II) (left, ■, horizontal stripes) and Zn(II) (right, ■, vertical stripes). (A) Comparison of Cu(II) and Zn(II) affinities for **bipy(GCN4bd2)₂** and **bipy-GS₂** (based on absorbance at 328) (B) comparison of Cu(II) and Zn(II) affinities for **terpy(GCN4bd2)₂** and **terpy-GS₂** (based on absorbance at 335 nm). Error bars represent the absolute error relative to calculation.

In contrast, the Cu(II) affinity of **bipy(GCN4bd2)₂** is ten times greater than that for the analogous **bipy-GS₂**, and the Zn(II) affinity of **terpy(GCN4bd2)₂** is three times lower than that for **terpy-GS₂**. These differences illustrate the influence of extra-ligands from the peptide backbone that can complement the polypyridyl-M(II) coordination sphere, thus increasing the stability of **bipy(GCN4bd2)₂:Cu(II)** compared with **bipy-GS₂:Cu(II)**. Alternatively, these extra ligands may compete with the **terpy** linker, in order to form separate and independent Zn(II) binding sites, thus decreasing the Zn(II) affinity of the **terpy** moiety in **terpy(GCN4bd2)₂:Zn(II)** compared to **terpy-GS₂:Zn(II)**.

3.2.2.4 – Cu(II) and Zn(II) addition to $\text{bipy}(\text{GCN4bd2})_2$ and $\text{terpy}(\text{GCN4bd2})_2$ monitored by circular dichroism

The CD spectrum of a 5 μM solution of $(\text{GCN4bd2})_2$ in 10 mM phosphate buffer pH 7.4, displays one transition with an intense minima around 200 nm, and a shoulder at 220 nm. The CD spectra recorded for $\text{bipy}(\text{GCN4bd2})_2$ and $\text{terpy}(\text{GCN4bd2})_2$ under similar conditions display similar transitions (see Figure 3.22). Upon addition of either CuCl_2 or ZnCl_2 , to $(\text{GCN4bd2})_2$, $\text{bipy}(\text{GCN4bd2})_2$, or $\text{terpy}(\text{GCN4bd2})_2$ solutions, no significant changes are observed in the recorded CD spectra (see Figure 3.22). Similarly no changes were observed upon addition of EDTA to solutions containing peptide conjugates and metals.

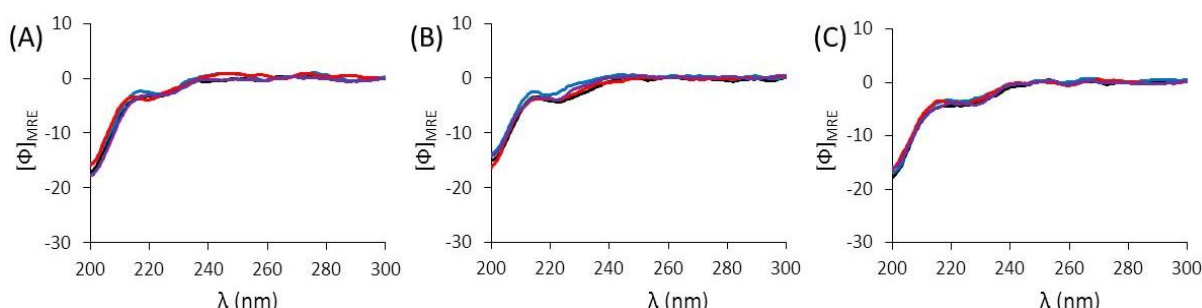


Figure 3.22 – CD spectra of 5 μM peptide dimer, $(\text{GCN4bd2})_2$ (A), peptide dimer conjugates $\text{bipy}(\text{GCN4bd2})_2$ (B) and $\text{terpy}(\text{GCN4bd2})_2$ (C) before (—) or after addition of two eq. of metal ion, either CuCl_2 (—) or ZnCl_2 (—), recorded in 10 mM phosphate buffer pH 7.4. Reversibility was monitored by addition of excess EDTA (10 eq. to metal ion (—)). $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$.

The addition of ZnCl_2 to $\text{bipy}(\text{GCN4bd2})_2$ was repeated at higher concentration of 100 μM . Similarly, addition of two equivalents of ZnCl_2 does result in significant changes to the CD profile of $\text{bipy}(\text{GCN4bd2})_2$. A closer look at the region from 220–380 nm, in which $\pi \rightarrow \pi^*$ contribution could be expected (see Chapter 2), shows weak transitions (range: $-5,000$ to $5,000 \text{ deg cm}^2 \text{ dmol}^{-1}$) with minima at 277 and 322 nm, and negative maxima at 277 and 322 nm. Upon addition of ZnCl_2 , minima at 277 nm decreases in intensity, whilst minima at 322 nm increases in negative ellipticity (see Figure 3.23).

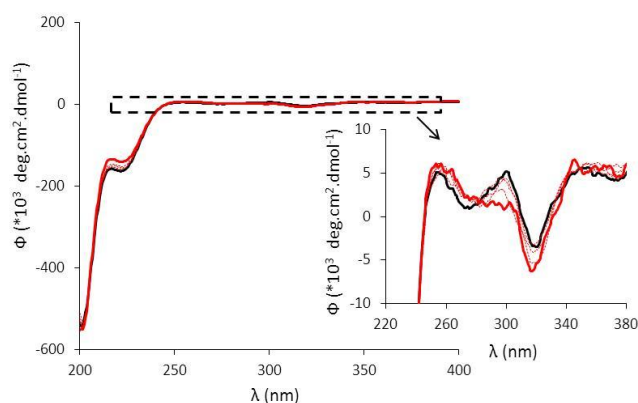


Figure 3.23 – Zn(II) titration of 100 μM **bipy(GCN4bd)₂** solution containing 10 mM phosphate buffer pH 7.4, monitored by CD. (—) 0 eq., (■■■) 0.2, 0.4, 0.6 eq., or (—) 2 eq. ZnCl_2 . $[\Phi]$ represents the molar ellipticity in $\times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. Insert on the right represent a zoom on the region indicated with a dotted square on the full spectra.

CD spectra recorded for **(GCN4bd2)₂**, **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂**, in the absence or presence of two equivalent of either CuCl_2 or ZnCl_2 , are consistent with peptide moieties that do not adopt a defined secondary structure (random coil), as indicated by the negative transition at 200 nm attributed to the $n \rightarrow \pi^*$ transition of the peptide bond. For CD spectra of **bipy(GCN4bd2)₂** recorded at a higher concentration (100 μM), weak CD signals ($-5,000 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$) could be distinguished in the polypyridine $\pi \rightarrow \pi^*$ transition region (250-350 nm). Upon addition of ZnCl_2 there is a slight shift of the signal, but no significant change in ellipticity is observed. This contrasts with addition of ZnCl_2 to **bipy-GS₂**, which resulted in a significant increase in both negative and positive ellipticity ($-30,000$ to $70,000 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$). N-terminus capped **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** do not retain metal-induced chirality of polypyridine $\pi \rightarrow \pi^*$ transition, even though the number of chiral center in the conjugate has significantly increased (from 4 chiral center in the glutathione dimer conjugates, to 38 for **GCN4bd2** dimer conjugates). This supports the hypothesis that the uncapped N-terminus of glutathione participates in the formation of hybrid complexes with **bipy/terpy:Zn** and **bipy/terpy:Cu** complexes in the glutathione based model compounds, as suggested in Chapter 2.

3.3 – Summary

Two new peptides based on GCN4**bd** were designed **GCN4bd1** and **GCN4bd2**, differing in the number of conserved residues and the flexibility of the sequence directly preceding the linker attachment. Polypyridine conjugates **pyr(GCN4bd1)₂**, **bipy(GCN4bd1)₂**, and **terpy(GCN4bd1)₂** of the longer **GCN4bd1** peptide were prepared. Only small quantities of the latter two could be obtained, and they proved challenging to purify. UV spectra recorded for **pyr(GCN4bd1)₂**, in the absence and presence of CuCl₂ resemble those recorded for the smaller glutathione analogue. In contrast, similar experiments performed with **bipy(GCN4bd1)₂**, and **terpy(GCN4bd1)₂** indicate reduced CuCl₂ binding and partial *trans*-to-*cis* conformational transition.

Peptide conjugates **bipy(GCN4bd2)₂**, and **terpy(GCN4bd2)₂**, based on the shorter peptide **GCN4bd2**, were prepared in larger quantities, and in higher purity. UV analysis of **bipy(GCN4bd2)₂**, and **terpy(GCN4bd2)₂** recorded in the absence and presence of both CuCl₂ and ZnCl₂, respectively, are consistent with the conformational transition and metal binding observed for **bipy-GS₂** and **terpy-GS₂**, respectively. Small differences were observed between glutathione and **GCN4bd2** conjugates, as illustrated by slightly different binding constant estimated for **bipy**:Cu and **terpy**:Zn systems. Moreover, metal ion addition (CuCl₂ or ZnCl₂) to **bipy(GCN4bd2)₂**, and **terpy(GCN4bd2)₂** did not result in significant induced CD signals for polypyridine $\pi \rightarrow \pi^*$ transitions, as was previously observed for **bipy-GS₂** and **terpy-GS₂**, even though **GCN4bd2** contains 9 times more chiral center than glutathione. It is possible that capping of N- and C-termini of **GCN4bd2** reduce their metal-ions affinity, thus preventing the formation of the chiral complexes of the type $[M^*(\kappa 6\text{-bipy-gly})^{**}]$ and $[M^*(\kappa 4\text{-terpy-gly})^{**}]$ (see Chapter 2). Despite respective **GCN4bd2** and glutathione polypyridine conjugates bind Cu(II) and Zn(II) in similar ratio and with similar affinity, the resulting complexes are different.

This chapter successfully illustrates that polypyridine can be used to dimerise large peptides (19 or 29 amino acids) and in the case of the former was shown to retain similar Cu(II)/Zn(II) binding affinity to the linker domain.

3.4 – Experimental

3.4.1 – Equipment and reagents

Sodium hydroxide, mono- and dihydrogen potassium salts, Tris base, ethylene diamine tetraacetic acid (EDTA), acetonitrile, water (HPLC grade), HCl 32 %w, acetic anhydride were all obtained from Fisher Scientific. Copper chloride (CuCl_2), and L-glycine were obtained from Sigma-Aldrich. Trifluoroacetic acid (TFA), zinc chloride (ZnCl_2) were obtained from BOC. All peptide synthesis reagents, dithiotreithiol, and diisopropylethylamine (DIEA) were obtained from AGTC bioproducts. 5,5'-Dithiobis-(2-nitrobenzoic acid) was obtained from Fluka. **Pyr-Br₂**, **bipy-Br₂** and **terpy-Br₂** were prepared as described in Chapter 2.

ES-TOF MS were recorded on a Microwaters LCT TOF spectrometer equipped with a 3000 V capillary voltage, and a cone voltage of 35 V. Analytical RP-HPLC spectra were recorded on C18 column using a 0 to 100 % gradient acetonitrile in water over 40 minutes (with 0.05% TFA constant) and the absorption was monitored at 220 nm. The compound purity is reported as a percentage of its peak integral over the total integration for all peaks present between 0 and 40 minutes, at the exception of the signal attributed to buffer salts which was neglected. UV measurements were recorded in a 1 cm pathlength quartz cuvette at 298 K, either on a CARY50 (spectrometer(a)), or Shimadzu 1800 spectrometer (spectrometer(b)). CD spectra were recorded in 1 mm pathlength quartz cuvettes at 298 K on a Jasco J-715 spectropolarimeter. Cartoon figures based on GCN4 were all prepared using pymol version 1.1eval from files deposited in the protein data bank (pdb). Non-linear fitting were performed using Kaleidagraph software version 4.0 and was similar to that described in Chapter 2.

3.4.2 – Synthetic procedure and characterisation

This section describes the synthesis and characterisation of peptides and peptide conjugates relevant to this chapter. Peptide conjugate preparation was adapted from previous work on polypyridine peptide conjugates.[22] Peptides formula and molecular weight were calculated using an online calculator.[31]

Synthesis of GCN4bd1 :

GCN4bd1 (Ac-ALKRARN TEAARRS RARKLQR MKQLGGC G-NH₂) was prepared by Fmoc solid-phase peptide synthesis[25] on a CEM Liberty 1 automated synthesiser. DMF was the main solvent, and all steps were carried out under a N₂ atmosphere. The scale selected was 0.25 mmol. Rink amide MBHA resin (0.65 mmol/g, 0.385 g) was used as the solid support, and a protected peptide chain was prepared by successive steps consisting of Fmoc deprotection (20% piperidine in DMF), and amino-acid coupling (25 W, 348 K, 330 sec, 5 equivalent of Fmoc protected amino-acids, DIEA and HBTU were used as activators). Resin was washed with DMF in between each step. Fmoc-L-Arg(Pbf)-OH and Fmoc-L-Cys(Trt)-OH were double coupled to ensure high yield. After all coupling and N-terminal deprotection cycles were carried out (including deprotection of the last residue), the resin was transferred from the synthesiser reaction vessel to a round bottom flask. The peptide was then capped (acetylated), by stirring the resin in a mixture containing DIEA/acetic anhydride/DMF (20/20/60) for 20 min. The resin was successively washed with 5x10 mL of DMF, 5x10 mL CH₂Cl₂, and 2x10 mL of diethylether. After drying of the resin in a dessicator overnight, the peptide was simultaneously cleaved from the support and the amino acid side-chains deprotected, by stirring the resin in a 10 mL cleavage mixture containing TFA / thioanisole / ethanedithiol / anisole (90 / 5 / 3 / 2), for 2 hours. The resin was removed by filtration and washed with TFA. The filtrate volume was reduced to ~5-7 mL by blowing air over the top of

the solution, and the peptide was precipitated by addition of cold diethylether (~10 times the volume of filtrate). After storing in the freezer overnight, the supernatant was removed, and diethylether added. The solution was allowed to stand in the freezer for 1 hour, after which the supernatant was removed once more. The remaining diethylether was removed *in vacuo*, affording a white solid/gel, a small fraction of which was analysed by MALDI-TOF and analytical HPLC indicating incomplete deprotection. Therefore, the whole deprotection step was repeated. The peptide was then purified by preparative RP-HPLC (C18 Phenomenex, all eluents contained 0.05 % TFA, absorbance monitored at 210 and 220 nm) using a 5 to 18 % gradient of acetonitrile in water over 30 min, directly followed by isocratic elution with 18 % acetonitrile in water for a further 15 minutes ($R_t = 38.4$ min). Solvents were removed *in vacuo*, and the peptide freeze-dried, affording **GCN4bd1** as salts (167 mg, 20 %). ES-TOF (water) calculated for $C_{135}H_{247}N_{57}O_{37}S_2$: 3324; found: 3323; $m/z = 1,108.7 [M+3H]^{3+}$ (18 %), $m/z = 831.7 [M+4H]^{4+}$ (59 %), $m/z = 665.6 [M+5H]^{5+}$ (100 %), $m/z = 554.8 [M+6H]^{6+}$ (41 %). MALDI-TOF: $m/z = 3324.9 [M+H]^+$ (100 %). RP-HPLC: 100%.

Synthesis of pyr(GCN4bd1)₂: 2,6-bis(methyl-S-GCN4bd1)pyridine

Di(bromomethyl)pyridine (0.4 mg, 1.51 μ mol) was added to 4.026 mL of 62 mM aqueous Tris.HCl buffer pH 8.0 solution, degassed by N₂ bubbling directly into the suspension, and an aqueous solution of **GCN4bd1** (6.20 mM, 6.04 μ mol, 974 μ L) added (water was degassed prior to peptide dissolution and the concentration was then determined by the Ellman's test). The suspension was stirred at room temperature and the concentration of free thiol in solution monitored by the Ellman's test. After 21 hours, fresh dithiotreithiol solution (0.32 M, 0.32 mmol, 1 mL) was added to the reaction mixture, stirred for 1 hour, and then purified by preparative RP-HPLC (C18 Phenomenex, monitoring at 210 and 260 nm, all eluents contained 0.05 % TFA) using a 0 to 19 % gradient of acetonitrile in water over 30 minutes, directly followed by isocratic elution with 19 % acetonitrile in water for a further 15

minutes. Peptide ($R_t = 35.8$ min) and product ($R_t = 37.6$ min) were isolated. Solvents were removed *in vacuo*, and the peptide conjugate freeze-dried, affording **pyr(GCN4bd1)₂** as a salt (7 mg, 69%). ES-TOF (water) calculated for $C_{277}H_{499}N_{115}O_{64}S_4$: 6753; found: 6753; $m/z = 965.8$ $[M+7H]^{7+}$ (12 %), $m/z = 845.2$ $[M+8H]^{8+}$ (28 %), $m/z = 751.5$ $[M+9H]^{9+}$ (55 %), $m/z = 676.5$ $[M+10H]^{10+}$ (83 %), $m/z = 615.0$ $[M+11H]^{11+}$ (100 %), $m/z = 563.9$ $[M+12H]^{12+}$ (79 %), $m/z = 520.6$ $[M+13H]^{13+}$ (43 %), $m/z = 483.5$ $[M+14H]^{14+}$ (12 %). RP-HPLC: 100%.

Synthesis of **bipy(GCN4bd1)₂**: 5,5'-bis-(methyl-S-GCN4bd1)-2,2'-bipyridine

Acetonitrile (80 μ L) and water (180 μ L) were added to a 5,5'-bis-(bromomethyl)-2,2'-bipyridine solution in acetonitrile (0.64 μ mol, 0.64 mM, 1.00 mL). The suspension was degassed with N_2 bubbling, and **GCN4bd1** solution (2.6 μ mol, 2.89 mM, 900 μ L) in 100 mM Tris.HCl buffer pH 8 was added (buffer was degassed prior to peptide dissolution and concentration was then checked by Ellman's test), and the suspension stirred at room temperature. After 15 hours, fresh dithiotreithiol solution (0.13 mmol, 0.32 M, 0.4 mL) was added to the reaction mixture, stirred for 1 hour, and then purified by preparative RP-HPLC (C18 Phenomenex, absorbance monitored at 210 and 260 nm, all eluents contained 0.05 % TFA) using a 0 to 19 % gradient of acetonitrile in water over 30 minutes, directly followed by isocratic elution with 19 % acetonitrile in water for a further 15 minutes. Peptide ($R_t = 35.8$ min) and product ($R_t = 38.0$ min) were isolated. Solvents were removed *in vacuo*, and the peptide conjugates freeze-dried, affording **bipy(GCN4bd1)₂** as a salt (3 mg, 14%). ES-TOF (water) calculated for $C_{282}H_{502}N_{116}O_{74}S_4$: 6830; found: 6831; $m/z = 976.7$ $[M+7H]^{7+}$ (22 %), $m/z = 854.9$ $[M+8H]^{8+}$ (53 %), $m/z = 760.0$ $[M+9H]^{9+}$ (75 %), $m/z = 684.1$ $[M+10H]^{10+}$ (76 %), $m/z = 622.0$ $[M+11H]^{11+}$ (90 %), $m/z = 570.2$ $[M+12H]^{12+}$ (100 %), $m/z = 526.4$ $[M+13H]^{13+}$ (63 %), $m/z = 488.9$ $[M+14H]^{14+}$ (20 %). RP-HPLC: 100%.

Synthesis of $\text{terpy}(\text{GCN4bd1})_2$: 6,6''-bis(methyl-S-GCN4bd1)-2,2':6',2''-terpyridine

6,6''-bis-(bromomethyl)-2,2':6',2''-terpyridine (0.34 mg, 0.63 μmol) was added to a flask containing 5 mL of degassed acetonitrile. 100 mM aqueous Tris.HCl buffer pH 8.0 solution (2.5 mL), and water (2.104 mL) were then added. The suspension was degassed for 20 min, by N_2 bubbling directly into the stirring suspension. Then, **GCN4bd1** (6.31 mM, 2.5 μmol , 396 μL) aqueous solution was added (water was degassed prior to peptide dissolution and concentration was then checked by Ellman's test). The flask equipped with a condenser was heated at ~ 323 K under N_2 for 2 days. Analytical HPLC of an aliquot did not show any sign of coupling. More 6,6''-bis-(bromomethyl)-2,2':6',2''-terpyridine (0.08 mg, 0.20 μmol) was added, and the reaction was heated at 343 K under N_2 for a further 5 days. Again an aliquot was analysed by analytical HPLC and ES-TOF, indicating partial formation of the expected product. The crude product was purified by RP-HPLC (C18 Phenomenex, absorbance monitored at 210 nm, all eluents contained TFA) in twice: (1) 0 to 30 % gradient of acetonitrile in water over 40 minutes on preparative column ($R_t = 37.9$ min), (2) 0 to 35 % gradient of acetonitrile in water over 40 minutes on semi-preparative column ($R_t = 29.7$ min). A mixture of $\sim 70:30$ expected compound **$\text{terpy}(\text{GCN4bd1})_2$** and related monomer **$\text{terpy}(\text{GCN4bd1})$** was obtained. Solvents were removed *in vacuo*, and the mixture freeze-dried, affording **$\text{terpy}(\text{GCN4bd1})_2$: $\text{terpy}(\text{GCN4bd1})$** in a 70:30 mixture, as a red gel (0.18 mg, 3 %). ES-TOF (water) calculated for $\text{C}_{287}\text{H}_{505}\text{N}_{117}\text{O}_{74}\text{S}_4$: 6907; found: 6907; $m/z = 1152.1$ $[\text{M}+6\text{H}]^{6+}$ (20 %), $m/z = 987.7$ $[\text{M}+7\text{H}]^{7+}$ (31 %), $m/z = 864.3$ $[\text{M}+8\text{H}]^{8+}$ (60 %), $m/z = 768.4$ $[\text{M}+9\text{H}]^{9+}$ (73 %), $m/z = 691.7$ $[\text{M}+10\text{H}]^{10+}$ (74 %), $m/z = 628.9$ $[\text{M}+11\text{H}]^{11+}$ (100 %), $m/z = 576.6$ $[\text{M}+12\text{H}]^{12+}$ (98 %), $m/z = 532.2$ $[\text{M}+13\text{H}]^{13+}$ (63 %), $m/z = 494.4$ $[\text{M}+14\text{H}]^{14+}$ (16 %). RP-HPLC: 87%.

Synthesis of GCN4bd2

GCN4bd2 (Ac-ALKRARN TEAARRS RARCG-NH₂) was prepared by Fmoc solid-phase peptide synthesis [25] on a CEM Liberty 1 automated synthesiser, as previously reported for **GCN4bd1**. Peptide was purified by preparative RP-HPLC (C18 Phenomenex, monitoring at 210 and 220 nm, all eluents solution contained 0.05 % TFA) using a 0 to 30 % gradient of acetonitrile in water over 30 minutes, directly followed by isocratic elution with 30 % acetonitrile in water over 7 more minutes (R_t = 34.0 min). The collected fraction was analysed by MALDI-TOF and analytical HPLC indicating incomplete deprotection. The whole deprotection and purification steps were repeated, but the composition of the cleavage mixture was changed for: TFA / triethylsilane / ethanedithiol (90 / 5 / 5),[32] and the HPLC purification involved a 0 to 15 % gradient of acetonitrile in water over 20 minutes, directly followed by isocratic elution with 15 % acetonitrile in water (R_t = 27.0 min). Solvents were removed *in vacuo*, and the peptide freeze-dried, affording **GCN4bd2** as a salt (63 mg, 12 %). ES-TOF (water), calculated for C₈₆H₁₅₈N₄₀O₂₅S: 2184; found: 2184; m/z = 1,092.7 [M+2H]²⁺ (35 %), m/z = 728.7 [M+3H]³⁺ (90 %), m/z = 546.8 [M+4H]⁴⁺ (100 %). MALDI-TOF: m/z = 2185.4 [M+H]⁺ (100 %), m/z = 2501.3 (52 %), m/z = 2818.2 (25 %). RP-HPLC: 99%.

Synthesis of (GCN4bd2)₂

To an aqueous solution of **GCN4bd2** (2.70 mM, 0.080 μmol, 29.6 μL), were added: 100 mM phosphate buffer pH 7.4 (20 μL) and water (150.4 μL) and the mixed solution was incubated at 310 K for 12 hours. The oxidation was monitored by analytical HPLC. The solution was used without the need for further purification. ES-TOF (water) calculated for C₁₇₂H₃₁₄N₈₀O₅₀S₂: 4367; found: 4367; m/z = 874.3 [M+5H]⁵⁺ (9 %), m/z = 728.7 [M+6H]⁶⁺ (30 %), m/z = 624.8 [M+7H]⁷⁺ (63 %), m/z = 546.8 [M+8H]⁸⁺ (100 %), m/z = 486.1 [M+9H]⁹⁺ (69 %), m/z = 437.6 [M+10H]¹⁰⁺ (55 %). RP-HPLC: 97%.

Synthesis of bipy(GCN4bd2)₂: 5,5'-bis-(methyl-S-GCN4bd2)-2,2'-bipyridine

To a 25 mL round-bottom flask containing 100 mM Tris.HCl buffer pH 8 (4 mL), and acetonitrile (2.98 mL) degassed mixture, were successively added: **GCN4bd2** aqueous solution (5.92 mM, 4.44 μ mol, 0.750 mL), and 5,5'-bis(bromomethyl)-2,2'-bipyridine (1.46 mM, 2.22 μ mol, 1.52 mL) dissolved in acetonitrile. The suspension was degassed by N₂ bubbling directly into the solution, and the suspension was stirred at room temperature. Progress of the reaction were monitored by analytical HPLC, indicating the reaction was quasi complete after 14 hours. The reaction mixture was dried on rotavapor, redissolved in water and purified by preparative RP-HPLC (C18 Phenomenex column, absorbance monitored at 210 nm, all eluents contained 0.05 % TFA) using a 0 to 21 % gradient of acetonitrile in water over 50 minutes (Rt = 45.7 min). Solvents were removed *in vacuo*, affording **bipy(GCN4bd2)₂** together with salts (5.6 mg, 56 %). ES-TOF (water), calculated for C₁₈₄H₃₂₄N₈₂O₅₀S₂: 4549; found: 4550; m/z = 910.8 [M+5H]⁵⁺ (41 %), m/z = 759.2 [M+6H]⁶⁺ (52 %), m/z = 650.9 [M+7H]⁷⁺ (79 %), m/z = 569.6 [M+8H]⁸⁺ (100 %), m/z = 504.4 [M+9H]⁹⁺ (96 %), m/z = 455.9 [M+10H]¹⁰⁺ (25 %). RP-HPLC: 100%.

Synthesis of terpy(GCN4bd2)₂: 6,6''-bis(methyl-S-GCN4bd2)-2,2':6',2''-terpyridine

In a 25 mL round bottom flask, were successively added: 100 mM Tris.HCl pH 8 (0.9 mL), **GCN4bd2** aqueous solution (5.92 mM, 0.592 μ mol, 100 μ L), acetonitrile (876 μ L), 6,6''-bis(bromomethyl)-2,2',6',2''-terpyridine (1.19 mM, 0.148 μ mol, 124 μ L) solubilised in acetonitrile. The suspension was degassed by N₂ bubbling directly into the solution, and then stirred at room temperature. Progress of the reaction was monitored by analytical HPLC, indicating that after 16 hours, the reaction was not proceeding further. The reaction mixture was dried on rotavapor, redissolved in water and purified by semi-preparative RP-HPLC (C18 Phenomenex column, absorbance monitored at 210 and 290 nm, all eluents contained 0.05 % TFA) using a 0 to 23 % gradient of acetonitrile in water over 50 min (Rt = 45.1 min).

Solvents were removed *in vacuo*, affording a colorless gel (0.15 mg, Y = 22 %). ES-TOF (water), calculated for $C_{189}H_{327}N_{83}O_{50}S_2$: 4626; found: 4626; $m/z = 926.1 [M+5H]^{5+}$ (16 %), $m/z = 771.7 [M+6H]^{6+}$ (30 %), $m/z = 661.8 [M+7H]^{7+}$ (62 %), $m/z = 579.2 [M+8H]^{8+}$ (100 %), $m/z = 514.9 [M+9H]^{9+}$ (71 %), $m/z = 463.6 [M+10H]^{10+}$ (26 %). RP-HPLC: 100%.

3.4.3 – Analytical procedures

3.4.3.1 – Pyr, bipy and terpy conjugates of GCN4bd1 (related to section 3.2.1)

For monitoring of **pyr(GCN4bd1)₂** formation by the Ellman's test,[33,34] 20 μ L of a 50 mM 5,5'-dithiobis-(2-nitrobenzoic acid) in 100 mM phosphate buffer pH 8.0 solution, was premixed with 600 μ L of 100 mM phosphate buffer pH 8.0 in a 1,000 μ L cuvette, and a blank spectrum recorded on spectrometer(a). To the cuvette, 2 μ L of either **GCN4bd1** stock solution, or reaction mixture was added, and a new spectrum was recorded after 15 min equilibration. All measurements were performed in triplicate and averaged. Spectra recorded for **GCN4bd1** stock solution was corrected to account for dilution upon reaction mixture preparation.

UV spectra of **pyr(GCN4bd1)₂**, **bipy(GCN4bd1)₂** and **terpy(GCN4bd1)₂** were obtained by respective addition of 30 μ L of 210 μ M, 10 μ L of 280 μ M and 8 μ L of 85 μ M of conjugate aqueous stock solutions to cuvettes containing 600 μ L of 20 mM phosphate buffer pH 8.0, on spectrometer(a). For **pyr(GCN4bd1)₂** and **terpy(GCN4bd1)₂**, 1 equivalent of $CuCl_2$ was added directly using either 1.5 μ L of 4.2 mM (for **pyr(GCN4bd1)₂** spectrum), or 1.4 μ L of 0.5 mM (for **terpy(GCN4bd1)₂** spectrum) $CuCl_2$ stock solution. Solutions were allowed to equilibrate for 15 min prior to recording spectra. For the **bipy(GCN4bd1)₂** Cu(II) titration, aliquots of a 0.28 mM $CuCl_2$ solution were added to the 600 μ L solution containing 4.5 μ M **bipy(GCN4bd1)₂** (0.1 equivalent per step). Equilibration time between $CuCl_2$ addition and recording spectra was three hours after the first addition; it was reduced to 10

min for the following 4 additions, and 2 minutes for the remainder of the titration. A plot of absorbance at 313 nm versus equivalence of CuCl_2 could not be fitted using non-linear regression due to the strong contribution of free **bipy(GCN4bd1)₂** at this wavelength.

In order to check for contamination of **bipy(GCN4bd1)₂** stock solution with metal ions, blank spectra were recorded for two cuvettes containing 600 μL of either 20 mM phosphate pH 8.0 solution, or 0.3 mM EDTA solution (pH was not recorded). It is important to note that the spectrum recorded for the EDTA solution was saturated below 220 nm. To each cuvette, 2 μL of stock solution containing 620 μM **bipy(GCN4bd1)₂**, was added and new spectra (blank subtracted) were recorded.

3.4.3.2 – Bipy and terpy conjugates of GCN4bd2, and (GCN4bd2)₂ (relative to section 3.2.2)

For metal ion titrations of **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** monitored by UV spectroscopy, aliquots of aqueous 300 (\pm 13) μM stock solutions of CuCl_2 or ZnCl_2 , were titrated into 600 μL of between 4.40 and 5.60 (\pm 0.05) μM solution of polypyridyl-peptide dimer conjugate in 20.0 (\pm 0.3) mM potassium phosphate buffer pH 7.4. Spectra were recorded on spectrometer(b) after 3 min equilibration. When the apparent Cu(II) binding constant was too high to be accurately estimated, L-glycine was added as a competitor ligand. 20.0 (\pm 0.4) mM L-glycine was present for the reported titration of **bipy(GCN4bd2)₂**. K_{app} values were calculated by fitting data for the absorbance maximum of the metal complexes as a function of Cu(II)/Zn(II) concentration, following the method described in Chapter 2.

For the continuous variation methods monitored by UV spectroscopy, solutions of varying ratio of **bipy(GCN4bd2)₂**/**terpy(GCN4bd2)₂** were prepared from 120 μM stock solutions, so that the total concentration of substrate $C_{\text{PJ}} + C_{\text{M}}$ was equal to 10 μM (where C_{PJ} and C_{M} represent the concentration of peptide dimer conjugate and metal ion, respectively). UV-visible spectra were recorded and the absorbance at λ_{max} , for the metal complexes $\pi \rightarrow \pi^*$

transition, plotted vs. the fraction of ligand. Due to a contribution at this wavelength from the free-ligand (**bipy(GCN4bd)**₂ and **terpy(GCN4bd)**₂), the absorbance values were corrected to account for this.[30] The fraction of peptide conjugates, χ **bipy(GCN4bd)**₂ and χ **terpy(GCN4bd)**₂, are defined as the ratio $C_p/(C_p+C_M)$.

For CD spectra of 5 μ M (**GCN4bd2**)₂, **bipy(GCN4bd2)**₂, or **terpy(GCN4bd2)**₂, a 300 μ L blank solution containing 10 mM phosphate buffer pH 7.4 was first recorded. 7.5 μ L of a 200 μ M peptide dimer or peptide dimer conjugate solution was added, and the data collected after 15 minute equilibration. Two equivalents of CuCl₂ or ZnCl₂ (1.5 mM, 3 nmol, 2 μ L) was added and the solution allowed to equilibrate for 15 min prior to recording spectrum. Ten equivalents of EDTA per metal ion (10 mM, 30 nmol, 3 μ L, pH 6-8) were added and the CD spectra recorded, in order to investigate reversibility. Spectra are an average of 20 scans recorded between 185 and 400 nm at 500 nm min⁻¹ (0.5 nm pitch). Difference spectra were obtained by subtracting CD spectra of buffer alone (blank), from that of mixtures containing both peptide and buffer, in order to observe the CD contribution from the peptide component only. The observed ellipticities in millidegrees were converted into residual molar ellipticity, Φ_{RME} , reported in units of deg cm² dmol⁻¹ res⁻¹, thus allowing estimation of secondary structure. Finally, spectra were blanked a second time by subtracting from each data point, the average Φ_{RME} measured between 400 and 300 nm for the same spectra (a region where no intense peak are expected).

The CD spectrum of a 100 μ M **bipy(GCN4bd2)**₂ solution was recorded directly in 100 μ M **bipy(GCN4bd2)**₂ and 10 mM phosphate pH 7.4. Aliquots of a 3 mM solution of ZnCl₂ were added to the solution, and spectra recorded after 15 minute equilibration. The measured ellipticity was converted into molar ellipticity, Φ , reported in units of deg cm² dmol⁻¹, thus allowing intensity comparison of the polypyridine $\pi \rightarrow \pi^*$ chiral induced signal, with **bipy-GS**₂ model compounds (see Chapter 2).

3.5 – References

- [1] Calkhoven C. F., Ab G., *Biochem J.*, **1996**, 317, 329-342.
- [2] Jaalouk D. E., Lammerding J., *Nat. Rev. Mol. Cell. Biol.*, **2009**, 10, 63-73.
- [3] <http://www.stromastudios.com/portfolio/transcription.html>, accessed on 23/05/2013.
- [4] Hurst H. C., Protein Profiles Transcription Factors 1:bZIP Proteins, Academic Press, London, **1994**, ISSN 1070-3667.
- [5] Sellers J. W., Vincent A. C., Struhl K. C., *Mol. Cell Biol.*, **1990**, 10, 5077-5086.
- [6] Landschulz W., Johnson P., McKnight S., *Science*, **1988**, 240, 1759-1764.
- [7] O' Shea E. K., Rutkowski R., Kim P. S., *Science*, **1989**, 243, 538-542.
- [8] Hope, I. A.; Struhl, K. *Cell* **1986**, 46, 885-894.
- [9] Ellenberger T. E., Brandl C. J., Struhl K., Harrison S. C., *Cell*, **1992**, 71, 1223-1237.
- [10] Talanian R. V., McKnight C. J., Kim P. S., *Science*, **1990**, 249, 769-771.
- [11] Ueno M., Murakami A., Makimo K., Morii T., *J. Am. Chem. Soc.*, **1993**, 115, 12575-12576.
- [12] Morii T., Saimei Y., Okagami M., Makino K., Sugiura Y., *J. Am. Chem. Soc.*, **1997**, 119, 3649-3655.
- [13] Cuenoud B., Schepartz A., *Science*, **1993**, 259, 510-513.
- [14] Cammaño A. M., Vásquez M. E., Martínez-Costas J., Castedo L., Mascareñas J. L., *Angew. Chem. Int. Ed.*, **2000**, 39, 3104-3107.
- [15] Morii T., Simomura M., Morimoto S., Saito I., *J. Am. Chem. Soc.*, **1993**, 115, 1150-1151.
- [16] Park C., Campbell J. L., Goddard III W. A., *Proc. Natl. Acad. Sci. USA*, **1993**, 90, 4892-4896.
- [17] Weber W., Fussenegger M., *Nat. Rev. Genet.*, **2012**, 13, 21-35.
- [18] Hörner M., Weber W., *FEBS Lett.*, **2012**, 586, 2084-2096.
- [19] Powers E. T., Morimoto R. I., Dillin A., Kelly J. W., Balch W. E., *Annu. Rev. Biochem.*, **2009**, 78, 959-991.

- [20] Vázquez M. E., Cammaño A. M., Mascareñas J. L., *Chem Soc. Rev.*, **2003**, 32, 338-349.
- [21] Vezerle D., Carrette L. L. G., Maddar A., *Drug Discov. Today: Technologies*, **2010**, 7, e115-e122.
- [22] Palmer C. R., Sloan L. S., Adrian Jr J. C., Cuenoud B., Paoletta D. N., Schepartz A., *J. Am. Chem. Soc.*, **1995**, 117, 8899-8907.
- [23] Keller W., König P., Richmond T. J., *J. Mol. Biol.*, **1995**, 254, 657-667.
- [24] Talanian R. V., McKnight J., Rutkowski R., Kim P. S., *Biochemistry*, **1992**, 31, 6871-6875.
- [25] Chan, W. C., and White, P. D., *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press: New York, **2000**, ISBN 978-0199637249.
- [26] Weiss M. A., Ellenberger T., Wobbe C. R., Lee J. P., Harrison S. C., Struhl K., *Nature*, **1990**, 347, 575-578.
- [27] Braterman P. S., Song J.-I., Peacock R. D., *Inorg. Chem.*, **1992**, 31, 555-559.
- [28] Cheng R. P., Fisher S. L., Imperiali B., *J. Am. Chem. Soc.*, **1996**, 118, 11349-11356.
- [29] Banerjee D., Kaden T. A., Sigel H., *Inorg. Chem.*, **1981**, 8, 2586-2590.
- [30] Bruneau E., Lavabre D., Levy G., Micheau J. C., *J. Chem. Ed.*, **1992**, 69, 833-837.
- [31] https://www.genscript.com/ssl-bin/site2/peptide_calculation.cgi, accessed on 22-04-2013.
- [32] http://www3.appliedbiosystems.com/cms/groups/psm_marketing/documents/generaldocuments/cms_040654.pdf, accessed on 22/01/2013.
- [33] Ellman G. L., *Arch. Biochem. Biophys.*, **1959**, 82, 70-77.
- [34] Collier H. B., *Anal. Biochem.*, **1973**, 56, 310-311.

CHAPTER IV: DNA BINDING OF PEPTIDE DIMER CONJUGATES BASED ON GCN4

4.1 – Introduction

4.1.1 – Sequence specific DNA binding by b-ZIP peptides and derivatives

In the mid 1980's, the GCN4 protein (previously AAS3 [1]) was shown to bind 5'-TGACTC-3' promoter sites (similar to the AP1 binding sites: 5'-TGAG/CTCA-3' [2,3]) from genes such as *HIS4*, thus participating in regulation of the amino-acids biosynthesis in *cerevisae*. [4-6] It was later shown that, *in vitro*, GCN4 also binds the consensus sequence 5'-TGACGTC-3' [7] (similar to the CRE site [2]), differing from AP1 by addition of a central bp. Various N-terminal truncations of wild-type GCN4, and subsequent affinity studies by bandshift assay (see Appendices) indicated that a C-terminal domain, consisting of 60 residues, is essential for binding to AP1 sites (see Figure 4.1). [8]

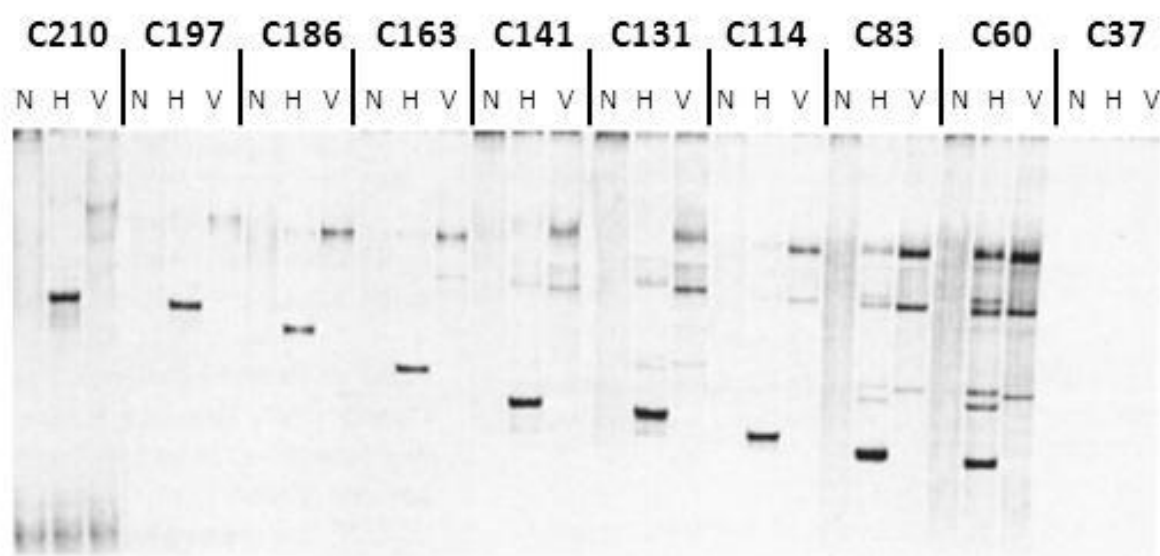


Figure 4.1 – Radiogram for bandshift assay performed on GCN4 derivatives with varying N-terminal deletion (the number of conserved GCN4wt residues are indicated at the top of the lane) by Struhl and co-workers. Peptides are products from cloned *pSP64* gene expression bearing ^{35}S -methionine label, and were incubates with either: (N) no DNA, (H) excess specific DNA (TaqI digested pUC8-*His3*), or (V) excess non-specific DNA (pUC9), prior to resolution on denaturing gel. Fast-migrating free protein is out of frame except for GCN4-C210 lane. The intense bands in the H lanes were attributed to specific GCN4-DNA complexes, and others to non-specific binding; adapted with permission from reference 8. Copyright 1986 Elsevier.

The 60 residue C-terminal domain was later shown to consist of 2 subdomains with distinct functions: (A) the **bd**, which makes specific contact to DNA,[9] and (B) the zipper domain responsible for dimerisation.[10,11] CD studies of a synthetic peptide containing the fully conserved basic and zipper domains of GCN4 (GCN4**bz**) demonstrated that peptide α -helicity increases in the presence of the AP1 and CRE DNA target sites (Figure 4.2A).[12] Considering precedent CD studies on leucine zipper peptides,[7] it was suggested that the **bd** undergoes a coil-to-helix structural transition in the presence of the DNA binding site. More recently, bandshift assays and CD studies performed on GCN4**bz** peptides, indicated similar affinity for and induced peptide folding in the presence of DNA containing the half-CRE site (containing half the palindromic CRE site) than for CRE and the natural AP1 binding site (Figures 4.2B and C).[13]

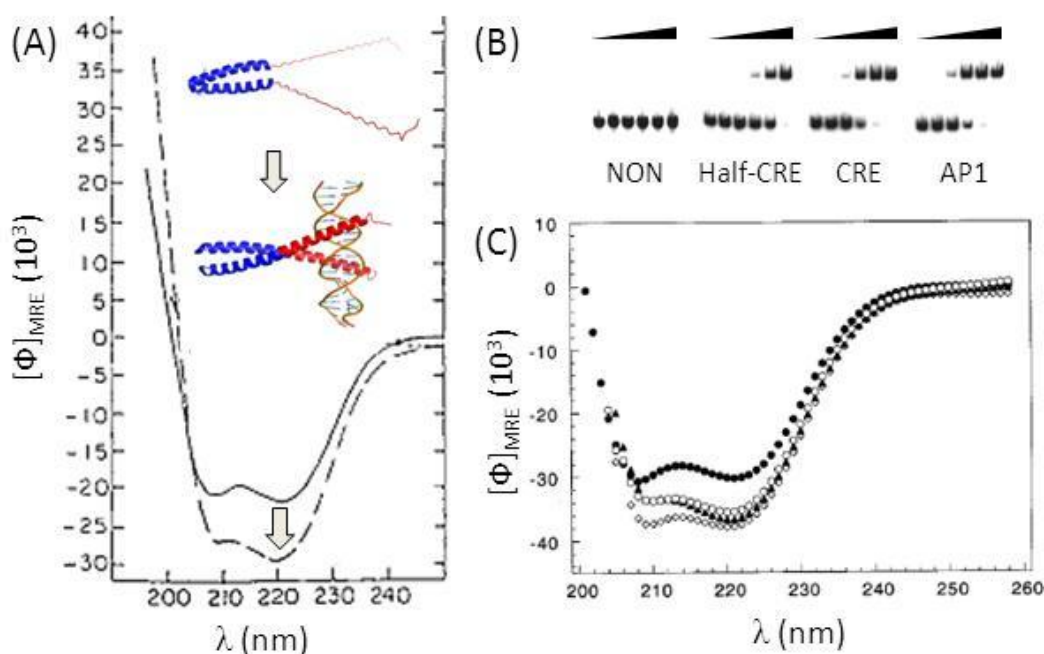


Figure 4.2 – (A) CD spectra for GCN4**bz** peptide recorded in the absence (—) or presence (---) of CRE DNA (0.5 eq.). (B) Radiogram of bandshift assay displaying sequence specificity of GCN4**bz**: lower band corresponds to free DNA, and the upper to the peptide/DNA complex; Increasing concentrations of peptide were incubated with different DNA sites labelled at the 5'-end with ^{32}P (poly[dI.dC]) was present in all mixtures as a competitor). (C) CD spectra for GCN4**bz** peptide in the absence (\bullet), or presence of 0.5 equivalent DNA, containing the NON (\diamond), Half-CRE (\circ), or AP1 target site (\blacktriangle); $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$; adapted with permissions from references 12 and 13, respectively. Copyright 1990 Nature Publishing Group and 2000 American Chemical Society.

As previously mentioned, Kim *et al.* prepared an artificial GCN4bd dimer, lacking a zipper domain, but dimerised through oxidation of an introduced cysteine (**GCN4brI^{SS}**). DNase footprint studies conducted at 277 K indicate that both **GCN4brI^{SS}** and GCN4bz peptides selectively protect AP1 consensus sites from enzymatic cleavage (see Figure 4.3A). However, DNase footprinting conducted at higher temperatures indicated that, unlike GCN4bz, sequence-specific DNA binding of **GCN4brI^{SS}** is lost at 297 K.[14] CD studies conducted at 298 K are consistent with **GCN4brI^{SS}** alone adopting a random coil secondary structure, which folds into an α -helix in the presence of the AP1 DNA target site (see Figure 4.3B).

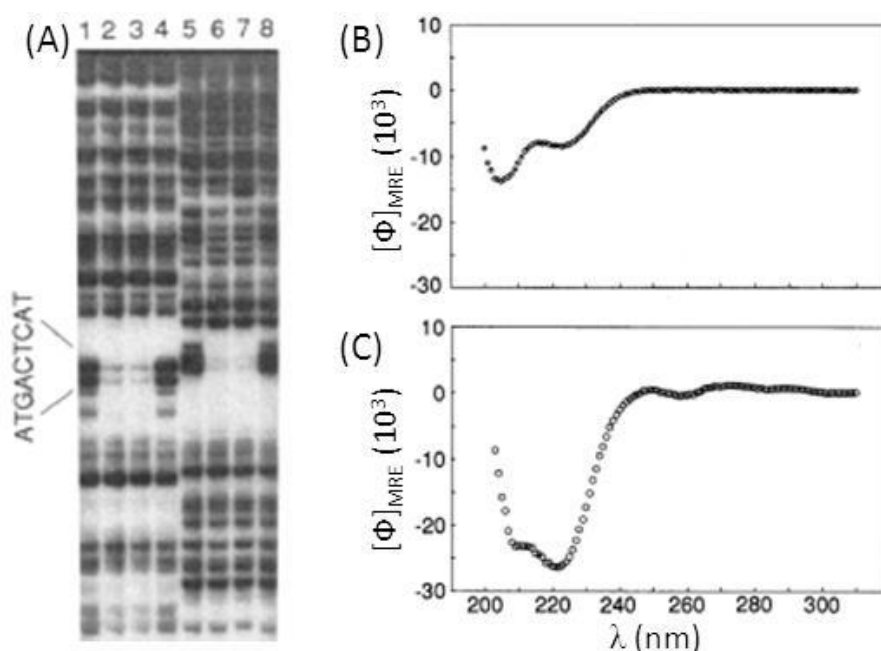


Figure 4.3 – DNA binding studies of (**GCN4brI^{SS}**) performed at 277 K. Left (A): Relevant section of radiogram relative to DNaseI footprinting experiment run for each single strand (1-4 and 5-8) of 251bp DNA duplex containing AP1 site (³²P labelled on each 5'-end). Enzymatic cleavage were performed on solutions containing DNA only (lanes 1,4,5,8), or in presence of either **GCN4brI^{SS}** (lanes 2,6), or GCN4bz peptide (3,7). Right: CD spectra of **GCN4brI^{SS}** alone, (B) or in the presence of 1 eq. AP1 (C); $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$; adapted with permission from reference 14. Copyright 1990 AAAS.

It was later shown that double-stranded DNA can serve as a template to induce the folding of short peptides containing alanine and lysine residues into amphipatic α -helices

(positive residues located on the same side of the helix), without specific peptide-DNA contacts.[15] Therefore, it is necessary to compare the folding of **GCN4bd** derivatives in the presence of different oligonucleotides containing specific or non-specific sites, in order to assess sequence specificity. Sequence-specific DNA binding properties of artificially dimerised peptides based on **GCN4bd** (see section 4.1) are normally studied by electrophoresis and/or CD, and involve oligonucleotides containing specific (AP1, CRE) as well as non-specific sequences (NON and half-CRE).[16,17] Notably, **[G₂₈T_S]₂Fe** conjugates prepared by Schepartz and co-workers were shown by CD and bandshift assays to discriminate between CRE and AP1. It was proposed that artificial dimerisation with a bulky linker reduces the conjugate's affinity for AP1, without affecting its affinity for the larger CRE site (see Figure 4.4).[18] Even though GCN4 peptides have been shown to bind AP1 and CRE with similar affinity, CRE has been shown to better accommodate the bulk resulting from the large polypyridine linker used to artificially dimerised GCN4.[18] These structural features are likely the reason why CRE target sites are generally used in previous DNA binding studies of artificially dimerised GCN4bd.[17,19,20]

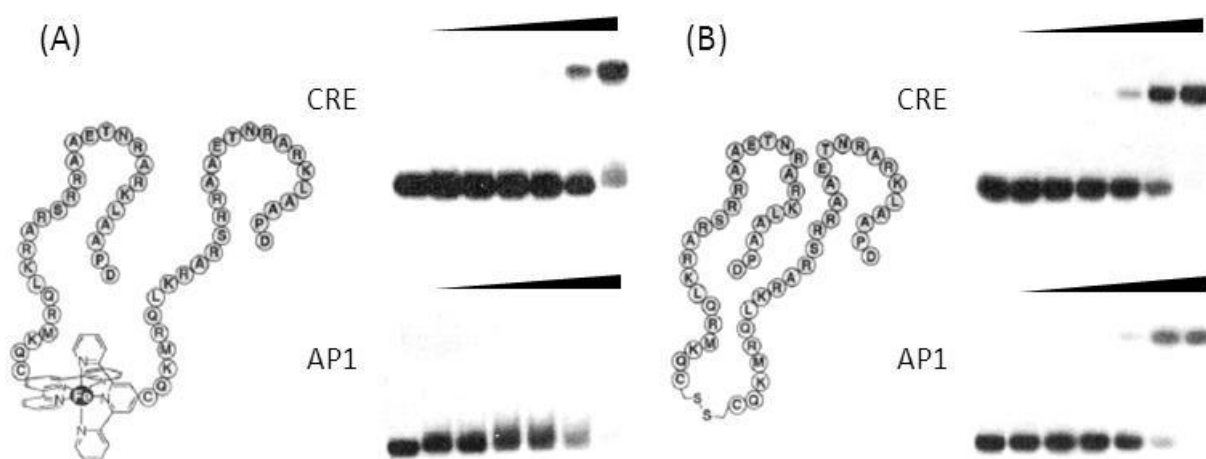


Figure 4.4 – Bandshift assays display a strong preference of conjugate **[G₂₈T_S]₂Fe** for CRE rather than AP1 sites (A), as opposed to **G₂₈^{SS}** able to bind both sequences with similar affinity (B). Gel radiograms correspond to electro-elution of solutions containing increasing amounts of **[G₂₈T_S]₂Fe** or **G₂₈^{SS}**, together with constant amounts of DNA, 5'-end labelled with ³²P and containing either the CRE or AP1 site; adapted with permission from reference 21. Copyright 1995 American Chemical Society.

4.1.2 – Aims of the chapter

In the preceding chapter, the preparation of peptide analogues of GCN4**bd**, and their dimerisation with several polypyridine linkers was reported. Metal coordination and the influence on linker conformation was also investigated and compared to shorter analogues. The work described in this chapter was aimed to evaluate the ability of conjugates bearing GCN4**bd** moieties to bind target and non-target DNA, and how it is influenced by the addition of metal ions, using CD measurement and bandshift assay. The metal affinity of the conjugates in the presence of DNA bearing either a specific or non-specific site was also studied, so as to evaluate the cooperativity between metal ion and sequence specific DNA binding.

4.2 – Results and discussion

4.2.1 – DNA binding of GCN4**bd1** and related conjugates

4.2.1.1 – CD studies of GCN4**bd1** and conjugates in the presence or absence of the DNA target site

The CD profiles of solutions containing 10 μ M **GCN4bd1** were recorded in the presence of excess tris(2-carboxyethyl)phosphine (TCEP), a suitable reagent for disulphide bonds reduction,[22] in order to prevent formation of the dimer. In the absence of DNA, **GCN4bd1** alone displays weak ellipticity at 222 nm (Φ_{222} -7,514 deg $\text{dmol}^{-1} \text{cm}^2 \text{res}^{-1}$), which is consistent with a random coil structure. In the presence of non-specific DNA, the CD signal relative to that of **GCN4bd1** displays a slightly higher negative ellipticity value at 222 nm (Φ_{222} -13,366 deg $\text{dmol}^{-1} \text{cm}^2 \text{res}^{-1}$). However, in the presence of the target CRE DNA site, the CD profile displays an important increase of ellipticity at 222 nm (Φ_{222} -26,140 deg $\text{dmol}^{-1} \text{cm}^2 \text{res}^{-1}$), and the spectrum is consistent with that of an α -helix (see Figure 4.5A).

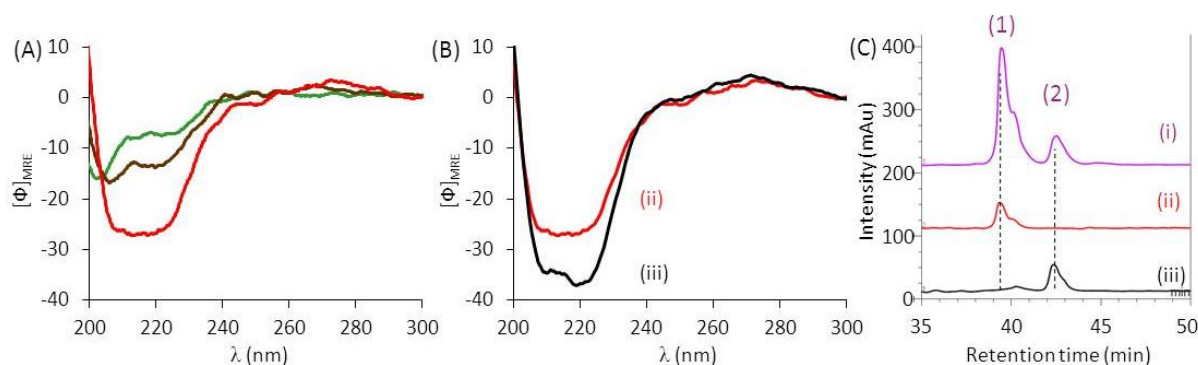


Figure 4.5 – (A) CD profile of 10 μM **GCN4bd1** alone (—), in the presence of 0.5 eq. DNA bearing either NON (—), or the CRE sites (—) (1 mM TCEP present, contribution from DNA was subtracted). (B) Impact of oxidation on the CD spectra of 10 μM **GCN4bd1** in the presence of 0.5 eq. of CRE DNA; solutions contained either 1 mM TCEP (ii, —), or 100 μM DTT (iii, —) as reducing agent (solution iii was incubated for 12 hours at 298 K). (C) Analytical C18 RP-HPLC for solutions containing **GCN4bd1** partially oxidised (i), or relative to CD measurements in B (ii, iii). Peaks observed in (i) were assigned to: (1) **GCN4bd1** monomer, and (2) (**GCN4bd1**)₂ oxidised dimer (based on electrospray mass spectra recorded for collected fractions). $[\Phi]_{MRE}$ represents the molar residual ellipticity in $\times 10^3$ deg cm² dmol⁻¹ res⁻¹.

Unlike TCEP, DTT rapidly oxidises in air losing its ability to reduce peptide bonds.[23]

As a result, the CD spectrum recorded of a solution containing 10 μM **GCN4bd1**, 5 μM CRE, and 100 μM DTT changes over time. After incubation at 298 K for 12 hours, the ellipticity at 222 nm has significantly increased (Φ_{222} -26,492 \rightarrow -36,316 deg dmol⁻¹ cm² res⁻¹), and HPLC indicates total oxidation and formation of (**GCN4bd1**)₂ (see Figures 4.5B and C).

CD spectra of solutions containing 5 μM peptide conjugates **pyr(GCN4bd1)**₂, **bipy(GCN4bd1)**₂, and **terpy(GCN4bd1)**₂ were also recorded. As mentioned in Chapter 3, **terpy(GCN4bd1)**₂ but also to some extent the **bipy(GCN4bd1)**₂ used for these experiments are not pure (despite using HPLC), and related results are to be treated qualitatively rather than quantitatively. The 3 conjugates display low negative ellipticity at 222 nm in the absence of DNA (**pyr(GCN4bd1)**₂ Φ_{222} -5,219; **bipy(GCN4bd1)**₂ Φ_{222} -3,560; **terpy(GCN4bd1)**₂ Φ_{222} -5,598 deg dmol⁻¹ cm² res⁻¹), and in the presence of NON DNA (**pyr(GCN4bd1)**₂ Φ_{222} -8,295; **bipy(GCN4bd1)**₂ Φ_{222} -5,393; **terpy(GCN4bd1)**₂ Φ_{222} -6,058 deg dmol⁻¹ cm² res⁻¹). In contrast, the ellipticity at 208 and 222 nm of the 3 conjugates increased significantly in the presence of DNA containing the CRE site (**pyr(GCN4bd1)**₂ Φ_{222} -17,009; **bipy(GCN4bd1)**₂

Φ_{222} -11,329; **terpy(GCN4bd1)₂** Φ_{222} -11,987 deg dmol⁻¹ cm² res⁻¹), suggesting partial formation of α -helices (see Figure 4.6).

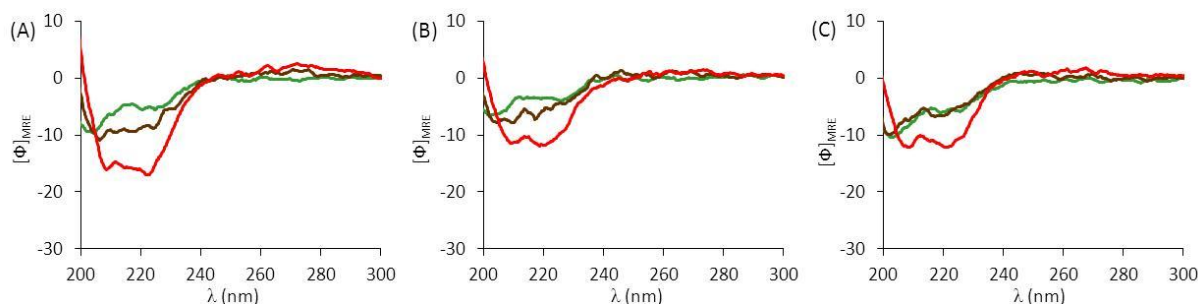


Figure 4.6 – CD spectra of 5 μ M **GCN4bd1** dimer conjugates in 10 mM phosphate buffer pH 7.4, recorded in the absence (—) or presence of 5 μ M DNA containing either NON (—), or CRE sites (—) (contribution from DNA was subtracted). $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3$ deg cm² dmol⁻¹ res⁻¹. (A) **pyr(GCN4bd1)₂**, (B) **bipy(GCN4bd1)₂** and (C) **terpy(GCN4bd1)₂**.

The folding of the **GCN4bd1** peptide and related conjugates in the presence of DNA containing the CRE target site, is consistent with the sequence-selective folding previously described for **GCN4bd** derivatives.[14,18] Notably, peptide oxidation and subsequent dimerisation with a disulphide linker results in a further increase of the helicity in the presence of CRE DNA. Dimerisation through larger linkers, such as **Me₂pyr**, **Me₂bipy**, and **Me₂terpy** results in similar sequence selectivity, however the conjugates are notably less α -helical than either **GCN4bd1** or the oxidised dimer.

4.2.1.2 – Bandshift assay to estimate **GCN4bd1** and conjugates affinity for CRE sites

Further investigations on the DNA binding of **GCN4bd1** conjugates were envisioned, including the calculation of their affinity for CRE DNA using bandshift assays, by analogy with previous reports on **GCN4bd** synthetic dimers. The affinity can be calculated from relative quantities of bound versus free peptide or DNA, and therefore requires partial dissociation of the peptide/DNA complex, which is achieved by working at concentrations close to or below the dissociation constant. **GCN4bz** homodimers and artificial **GCN4bd** dimers were previously estimated to bind CRE sites with nanomolar affinities by bandshift assay.[7,13] Therefore, related dissociation constant estimation involves use of a sensitive

detection method, such as radioactive labelling. Consequently, oligonucleotides containing CRE sites were labelled at the 5'-end with ^{32}P -ATP, affording DNA strands with activity of ~ 20 Ci/mmol (based on the assumption that the labelling reaction reached completion and calculated using the activity indicated by the manufacturer at day 0). DNA solutions containing as low as 10 nM ^{32}P -DNA (~ 4 nCi) were eluted on polyacrylamide gels and detected on a phosphorimaging screen, visualised with a phosphorimager.

Similarly, solutions containing 10 nM ^{32}P -labelled CRE DNA in binding buffer and increasing amount of **pyr(GCN4bd1)₂** were loaded on a polyacrylamide gel and subsequently electrophoresed. Upon visualisation, the radiogram indicates that the five solutions containing 0-4 nM **pyr(GCN4bd1)₂** all eluted as single bands with similar retention times, which was attributed to the free DNA (see Figure 4.7). However, electro-elution of the sixth and seventh solution, containing 6.5 and 10 nM **pyr(GCN4bd1)₂**, resulted in formation of an additional slower-migrating band, attributed to the **pyr-(GCN4bd1)₂**-DNA complex. The intensity of the fast-migrating band decreases as the concentration of **pyr(GCN4bd1)₂** increases, whereas the slow-migrating band increases in intensity. As a result, electro-elution profiles of solutions containing 15-100 nM **pyr(GCN4bd1)₂** all display an intense slow-migrating band and only a faint fast-migrating band. Solutions containing more than 100 nM display bands at the wells latitude, indicating precipitation of DNA under these conditions of high excess of peptide.

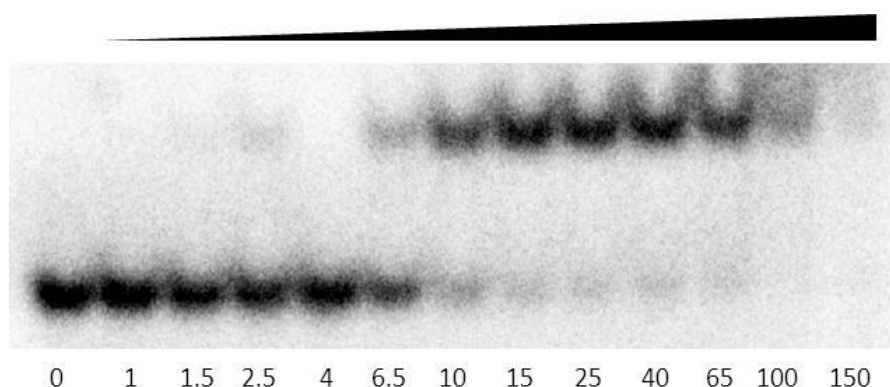


Figure 4.7 – Radiogram recorded after electro-elution of solutions containing ~ 10 nM of 5'-end labelled CRE oligonucleotides (^{32}P , 4 nCi), and increasing amounts of **pyr(GCN4bd1)₂** (concentrations are indicated in nM below the radiogram) in binding buffer. When the **pyr(GCN4bd1)₂** concentration was higher than 100 nM, ^{32}P DNA precipitated in the loading well (out of frame).

As expected, addition of **pyr(GCN4bd1)₂** to solutions containing DNA bearing the CRE site, resulted in band retardation, characteristic of complex formation and consistent with peptide folding observed in CD. Relative intensity indicates that 50% DNA present is peptide bound, when the **pyr(GCN4bd1)₂** concentration is between 6.5 and 10 nM. This concentration correspond to near equimolar amounts of **pyr(GCN4bd1)₂** and DNA, and might indicate formation of a 1:1 complex. One can conclude that the dissociation constant for **pyr(GCN4bd1)₂/CRE** would be close to or lower than 10 nM, even though it cannot be accurately estimated from this experiment (see section 4.4.3.1).

Solutions of **bipy(GCN4bd1)₂** and **terpy(GCN4bd1)₂** were not available of sufficient purity for such quantitative estimation and analogous experiments were not performed or are not described.

4.2.1.3 - Influence of Cu(II) and Zn(II) on bipy(GCN4bd1)₂ and terpy(GCN4bd1)₂ folding in the presence of target DNA

It remained to evaluate if sequence selective folding of **GCN4bd1** polypyridine conjugates was sensitive to the presence of Cu(II) and Zn(II) ions, and specifically whether these could regulate the conformation of the polypyridine linker domain in such a way that could be exploited to control DNA binding. For this, the CD spectra of solutions containing CRE DNA and peptide conjugates (either **bipy(GCN4bd1)₂**, or **terpy(GCN4bd1)₂** both 5 μ M), before and after addition of Cu(II) or Zn(II) were compared. In both case, Cu(II) and Zn(II) addition does not result in significant changes in the CD profile of the polypyridine-peptide dimer conjugates (see Figure 4.8). It must be noted that **bipy(GCN4bd1)₂** stock solution used in this experiment was not pure, and resulting spectra poorly compare with that displayed in Figure 4.6.

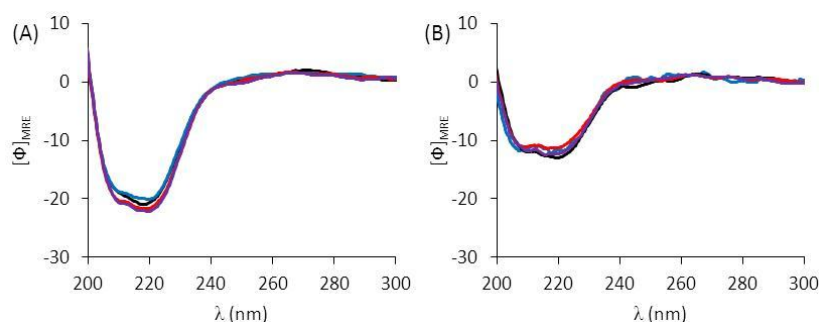


Figure 4.8 – CD spectra of (A) **bipy(GCN4bd1)₂**, or (B) **terpy(GCN4bd1)₂** in the presence of 1 eq. CRE DNA in 10 mM phosphate buffer pH 7.4, in the absence (—) or presence of either CuCl₂ (—) or ZnCl₂ (—). Reversibility was monitored by addition of excess EDTA (10 equivalents to metal ion (—)). $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$.

Previous UV measurements recorded at micromolar concentrations were consistent with conjugates **bipy(GCN4bd1)₂** and **terpy(GCN4bd1)₂** binding Cu(II) at the polypyridine linker sites (see Chapter 3). In contrast, it seems that addition of Cu(II) or Zn(II) are unable to alter folding of the peptide moieties, and by extension, the DNA binding ability of either **bipy(GCN4bd1)₂** or **terpy(GCN4bd1)₂** under these experimental conditions.

4.2.2 – DNA binding of GCN4bd2 and related conjugates

4.2.2.1 – CD study of GCN4bd and conjugates

Cu(II) and Zn(II) addition were previously shown to not affect the CD spectra of solutions containing **(GCN4bd2)₂**, **bipy(GCN4bd2)₂**, and **terpy(GCN4bd2)₂** (see Chapter 3). In this section, the CD profile of these conjugates recorded in the presence of equimolar amounts of DNA containing various sites (NON, CRE, Half-CRE, or AP1), and for which the contribution of the DNA has been subtracted, are reported and discussed. Ellipticity values at 222 nm, which are directly indicative of the peptide helicity,[24] are summarised in Table 4.1.

Table 4.1 – Summary of residual molar ellipticities values at 222 nm (displayed as $\times 10^3$ deg $\text{cm}^2 \text{dmol}^{-1} \text{res}^{-1}$) for peptide conjugates in the absence and presence of 1 equivalent of DNA bearing different consensus sequences. Two equivalents of metal ion (CuCl_2 or ZnCl_2) were then added to each solution, followed by addition of 10 equivalents EDTA.

Pep. conjugate	metals	no DNA	NON	CRE	Half-CRE	AP1
(GCN4bd2)₂	no	-3.39	-5.58	-14.58	-9.69	-11.31
	CuCl_2	-2.77	-4.86	-14.20	-10.41	-9.89
	ZnCl_2	-3.60	-4.21	-13.42	-10.18	-9.79
	EDTA	-3.13	-5.35	-13.71	-9.37	-8.94
bipy(GCN4bd2)₂	no	-4.40	-4.52	-13.14	-8.05	-9.44
	Cu	-2.90	-4.77	-17.00	-9.47	-8.86
	Zn	-4.08	-4.41	-22.30	-11.58	-10.86
	EDTA	-3.98	-3.92	-10.97	-8.46	-8.11
terpy(GCN4bd2)₂	no	-4.23	-4.05	-9.92	-6.07	-4.89
	Cu	-3.43	-6.67	-19.97	-16.90	-22.57
	Zn	-3.75	-4.93	-15.37	-11.04	-11.26
	EDTA	-4.05	-4.08	-9.54	-5.40	-3.38

The CD spectra of solutions containing 5 μM of NON and either **(GCN4bd2)₂**, **bipy(GCN4bd2)₂**, or **terpy(GCN4bd2)₂** all display two negative minima of different intensity located around 200 and 222 nm, which are not characteristic of any particular peptide secondary structure. These are similar to CD spectra recorded in the absence of DNA, and ellipticity recorded at 222 nm have low values (**(GCN4bd2)₂** Φ_{222} -5,583; **bipy(GCN4bd2)₂** Φ_{222} -4,523; **terpy(GCN4bd2)₂** Φ_{222} -4,046 deg $\text{dmol}^{-1} \text{cm}^2 \text{res}^{-1}$), and spectra are largely unaffected by addition of two equivalents of CuCl_2 or ZnCl_2 , and followed by 10 equivalents EDTA (see Figure 4.9). An exception being addition of two equivalents of CuCl_2 to a solution containing 5 μM of NON and **terpy(GCN4bd2)₂**, which resulted in a small increase of the negative ellipticity at 222 nm (Φ_{222} -4,046 \rightarrow -6,670 deg $\text{dmol}^{-1} \text{cm}^2 \text{res}^{-1}$).

¹), which is fully reversed upon addition of 10 equivalent EDTA (Φ_{222} -6,670 \rightarrow -4,075 deg dmol⁻¹ cm² res⁻¹) (see Figure 4.9C).

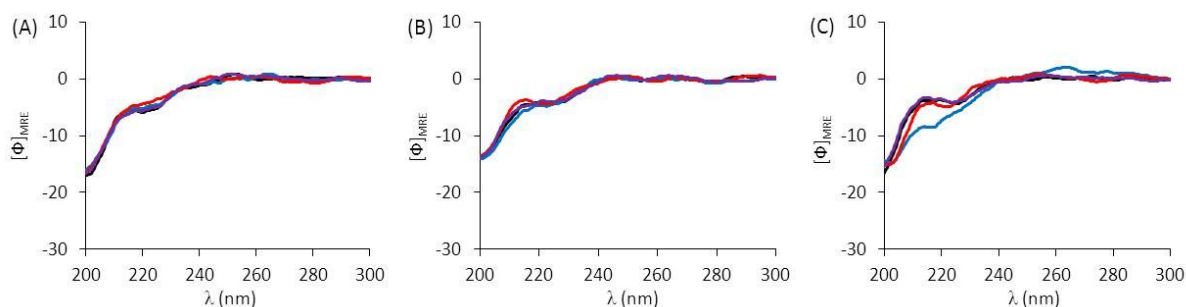


Figure 4.9 – CD spectra of 5 μ M **GCN4bd2** dimer and conjugates in the presence of 5 μ M NON DNA in 10 mM phosphate buffer pH 7.4. $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3$ deg cm² dmol⁻¹ res⁻¹. A) (**GCN4bd2**)₂ oxidised dimer, B) **bipy(GCN4bd2)**₂ and C) **terpy(GCN4bd2)**₂ DNA, recorded in the absence (—) and presence of 2 equivalents of either CuCl₂ (—) or ZnCl₂ (—). Reversibility was monitored by addition of excess EDTA (10 equivalents to metal ion (—)).

In contrast, more can be said about CD spectra of conjugates recorded in the presence of DNA containing the CRE site. The spectra of 5 μ M (**GCN4bd2**)₂ and **bipy(GCN4bd2)**₂ recorded in the presence of CRE DNA displays two negative bands centred at 208 and 220 nm of similar intensities, suggesting partial folding into α -helices (see Figures 4.10A and B). In contrast, the spectrum of 5 μ M **terpy(GCN4bd2)**₂ recorded in the presence of CRE DNA displays a more intense band around 204 nm, and a weaker band centred at 220 nm, which are consistent with **terpy(GCN4bd2)**₂ being more folded in the presence of CRE DNA than in the presence of NON, but less helical compared to either the disulphide or **bipy** analogues (see Figure 4.10C). Upon addition of two equivalents of CuCl₂ or ZnCl₂, the CD profile of (**GCN4bd2**)₂ remains largely unaffected (see Figure 4.10A). In contrast, the helicity of **bipy(GCN4bd2)**₂ in the presence of CRE DNA, increases by 29% upon addition of two equivalents of CuCl₂ (Φ_{222} -13,139 \rightarrow -16,997 deg dmol⁻¹ cm² res⁻¹) and 70% upon addition of two equivalents of ZnCl₂ (Φ_{222} -13,139 \rightarrow -22,301 deg dmol⁻¹ cm² res⁻¹) (see Figure 4.10B). Moreover, the peptide helicity of **terpy(GCN4bd2)**₂ in the presence of CRE DNA doubles upon addition of two equivalents of CuCl₂ (Φ_{222} -9,925 \rightarrow -19,974 deg dmol⁻¹ cm²

res⁻¹), but only increases by a modest 55% upon addition of two equivalents of ZnCl₂ (Φ_{222} - 9,925 → -15,370 deg dmol⁻¹ cm² res⁻¹) (see Figure 4.10C).

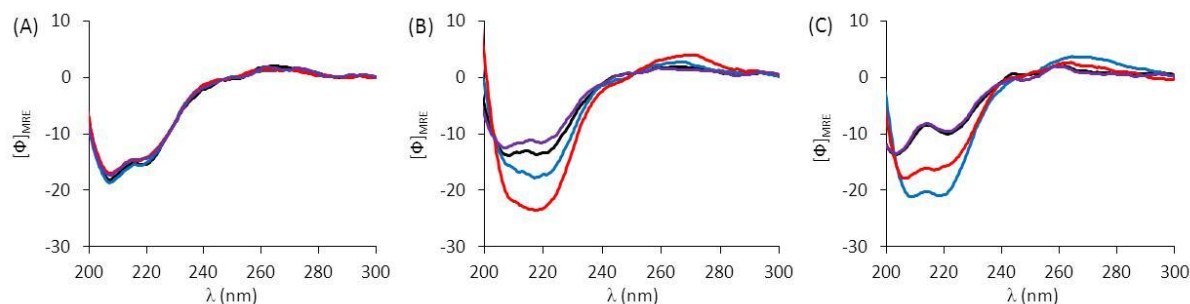


Figure 4.10 – CD spectra of 5 μ M **GCN4bd2** dimer and conjugates in the presence of 5 μ M CRE DNA in 10 mM phosphate buffer pH 7.4. A) (**GCN4bd2**)₂ oxidised dimer, B) **bipy(GCN4bd2)**₂ and C) **terpy(GCN4bd2)**₂ DNA. Spectra recorded in the absence (—) or presence of 2 eq. either CuCl₂ (—) or ZnCl₂ (—). Reversibility was monitored by addition of excess EDTA (10 eq. to metal ion (—)). [Φ]_{MRE} represents the molar residual ellipticity in $\times 10^3$ deg cm² dmol⁻¹ res⁻¹.

In order to assess reversibility, excess EDTA (10 equivalents per metal ion) was added to solutions of the metal-peptide conjugates in the presence of CRE DNA, resulting in a decrease in the negative ellipticity at 222 nm for both **bipy(GCN4bd2)**₂ (Cu Φ_{222} -16,997 → -10,965; Zn Φ_{222} -22,302 → -10,967 deg dmol⁻¹ cm² res⁻¹) and **terpy(GCN4bd2)**₂ (Cu Φ_{222} -19,974 → -9,538; Zn Φ_{222} -15,370 → -7,998 deg dmol⁻¹ cm² res⁻¹), whereas the spectra of (**GCN4bd2**)₂ were largely unaffected (see Figure 4.10).

At this stage, it was necessary to evaluate whether an increase of α -helicity upon metal addition was consistent with metal-binding at the polypyridine linkers, and if peptide helicity was optimum when two equivalents of metal ion were present. For this, solutions containing 10 μ M CRE and 10 μ M of either **bipy(GCN4bd2)**₂ or **terpy(GCN4bd2)**₂ were titrated with Cu(II) and Zn(II), and CD spectra were recorded for each 0.2 equivalence step (see Figure 4.11).

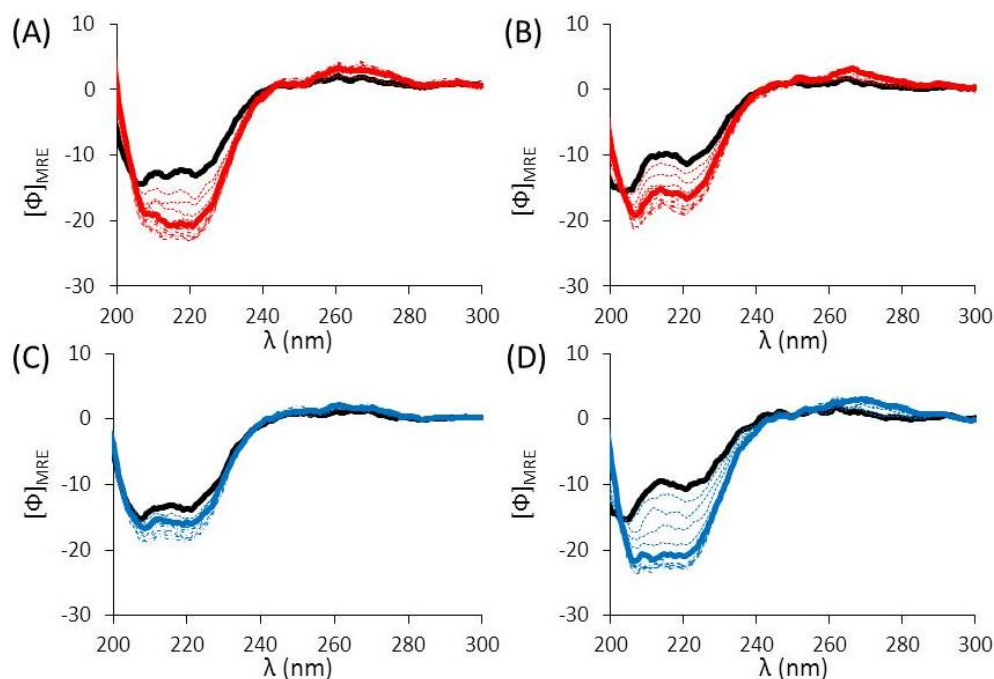


Figure 4.11 – CD spectra for metal ion titration of 10 μM peptide dimer conjugates and CRE DNA in 10 mM phosphate buffer pH 7.4. (A) Zn(II) titration of **bipy(GCN4bd2)₂**, (B) Zn(II) titration of **terpy(GCN4bd2)₂**, (C) Cu(II) titration of **bipy(GCN4bd2)₂**, and (D) Cu(II) titration of **terpy(GCN4bd2)₂**; (—) 0 eq. metal ion, (....) between 0 and 1 eq., (—) 1 eq., (---) more than 1 eq. ZnCl₂ added; (....) between 0 and 1 eq., (—) 1 eq., (---) more than 1 eq. CuCl₂ added. $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$.

Molar ellipticities at 222 nm were plotted as a function of metal ion equivalence. Out of the four plots, three are consistent with the formation of 1:1 complexes, and the ellipticity at 222 nm reaching a maximum when 2 equivalents of metal ion (either CuCl₂ or ZnCl₂) have been added. The exception could be the Cu(II)-**bipy(GCN4bd2)₂**-CRE system where the less intense change is less conclusive (Figure 4.12). Data were fitted to 1:1 models to determine equilibrium constant values relative to metal-induced folding (K_{mf}). Metal-induced folding constants, $\log K_{mf}$, were determined to be 5.87 ± 0.21 for the Zn-**bipy(GCN4bd2)₂**-CRE complex, 6.72 ± 0.21 for the Cu-**terpy(GCN4bd2)₂**-CRE complex, and 5.88 ± 0.14 for the Zn-**terpy(GCN4bd2)₂**-CRE complex. However, the data for the Cu(II)-**bipy(GCN4bd2)₂**-CRE could be fitted to neither 1:1, nor 1:2 binding model.

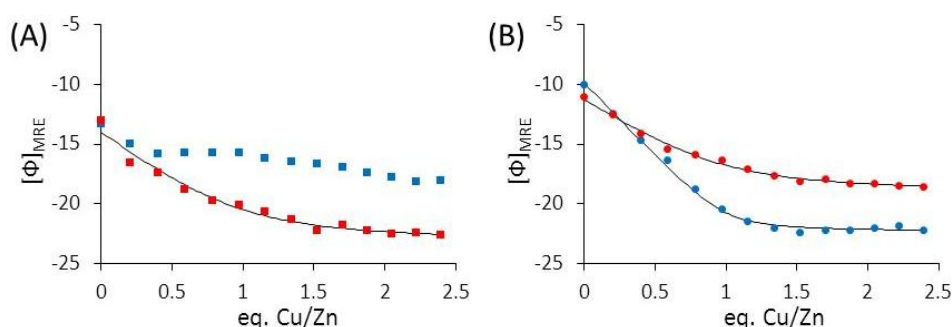


Figure 4.12 – Plot of residual molar ellipticities at 222 nm versus the metal ion equivalence, relative to CuCl₂ (■ and ●) and ZnCl₂ (■ and ●) titration of solution containing 10 μM CRE DNA and either (A) **bipy(GCN4bd2)₂**, or (B) **terpy(GCN4bd2)₂**. (—) represents best fit for a 1:1 metal:peptide dimer conjugate binding ratio. Data for the Cu titration **bipy(GCN4bd2)₂/CRE** could not be fitted.

Table 4.2 – Summary of the CD data obtained for Zn(II)/Cu(II) titration of polypyridine peptide dimer conjugates, **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂**.

Pep. conjugate	Metal	Φ_{222} (deg cm ² dmol ⁻¹ res ⁻¹)	K_{mf} (M ⁻¹)	R ²
bipy(GCN4bd2)₂	Zn	$-2.34 \pm 0.07 \text{ E } +04$	$7.33 \pm 4.64 \text{ E } +05$	0.9738
	Cu ^a	-	-	-
terpy(GCN4bd2)₂	Zn	$-1.92 \pm 0.03 \text{ E } +04$	$7.55 \pm 2.78 \text{ E } +05$	0.9928
	Cu	$-2.24 \pm 0.02 \text{ E } +04$	$5.20 \pm 3.24 \text{ E } +06$	0.9960

^a data could not be fitted to a 1:1 and 1:2 models.

As an extension to this work, the impact of metal addition to conjugates complexed with oligonucleotides similar to CRE, namely AP1 and half-CRE, was investigated. In the absence of Cu(II) or Zn(II), the CD spectra recorded for the three conjugates in the presence of half-CRE or AP1 are similar to that recorded in the presence of CRE: **(GCN4bd2)₂** and **bipy(GCN4bd2)₂** display more helical profiles compared to **terpy(GCN4bd2)₂**, which displayed lower ellipticity at 222 nm (see Figure 4.13). As for previous measurements, the addition of either CuCl₂ or ZnCl₂ does not alter the CD spectrum of **(GCN4bd2)₂**. Moreover, the addition of metal ions does not significantly alter the CD spectra of **bipy(GCN4bd2)₂** in the presence of either AP1 or half-CRE DNA, despite being sensitive to the presence of these metal ions in the presence of CRE DNA. In contrast, the helicity of **terpy(GCN4bd2)₂** in the presence of half-CRE and AP1, is highly sensitive to the presence of metal ions. For both half-CRE and AP1 the addition of Zn(II) is consistent with a more α -helical peptide (half-

CRE Φ_{222} -6,649 \rightarrow -11,039; AP1 Φ_{222} -6,328 \rightarrow -11,264 deg dmol⁻¹ cm² res⁻¹), of similar α -helicity as the **bipy**(GCN4bd2)₂ analogue. However the addition of Cu(II) is accompanied by a very significant increase in the α -helical signal at 222 nm (half-CRE Φ_{222} -6,074 \rightarrow -16,901; AP1 Φ_{222} -4,896 \rightarrow -22,572 deg dmol⁻¹ cm² res⁻¹) (Figure 4.13). As for spectra recorded in the presence of CRE, all changes observed upon addition of two equivalent CuCl₂ or ZnCl₂ were fully reversed when excess EDTA was added to the solution.

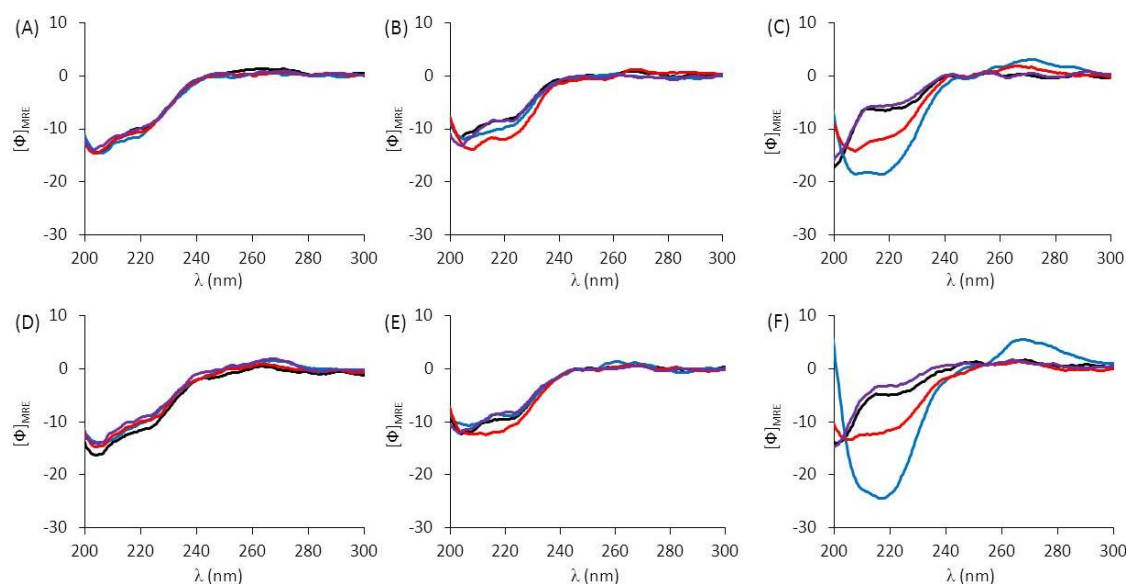


Figure 4.13 – CD spectra of 5 μ M **GCN4bd2** dimer and conjugates in the presence of 5 μ M DNA containing either half-CRE (A, B, C), or AP1 sites (D, E, F) in 10 mM phosphate buffer pH 7.4. (**GCN4bd2**)₂ oxidised dimer (A, D), **bipy**(GCN4bd2)₂ (B, E), **terpy**(GCN4bd2)₂ (C, F). Spectra recorded in the absence (—) and presence of 2 eq. of either CuCl₂ (—) or ZnCl₂ (—). Reversibility was monitored by addition of excess EDTA (10 eq. to metal ion (—)). $[\Phi]_{\text{MRE}}$ represents the molar residual ellipticity in $\times 10^3$ deg cm² dmol⁻¹ res⁻¹.

CD spectra of (**GCN4bd2**)₂ recorded in the presence or absence of specific or non-specific DNA, are all unaffected by Cu(II) or Zn(II) addition. This is consistent with the design which lacks a metal ion chelating polypyridine linker unit. However, CD spectra indicate that both **bipy**(GCN4bd2)₂ and **terpy**(GCN4bd2)₂ in the presence of the CRE site experience an increase in the negative ellipticity at 222 nm upon metal ion addition, which in turn is highly dependent on the nature of the metal ion, i.e. Cu(II) vs. Zn(II).

Stepwise addition of metal ion to solutions containing CRE and equimolar amounts of **bipy**(GCN4bd2)₂ or **terpy**(GCN4bd2)₂, indicates that the increase of ellipticity at 222 nm plateaus after addition of between 1 and 2 equivalents of metal ion, and is consistent with

formation of 1:1:1 ternary complexes (metal ion:peptide conjugate/CRE), except for Cu(II) addition to **bipy(GCN4bd2)₂/CRE** that might result from the successive formation of different complexes (see Figure 4.12). The intensity of the change in molar ellipticity at 222 nm suggests a higher affinity for DNA containing the CRE target site for Cu(II):**terpy(GCN4bd2)₂** and Zn(II):**bipy(GCN4bd2)₂** complexes, and a lower affinity for the Zn(II):**terpy(GCN4bd2)₂** and Cu(II):**bipy(GCN4bd2)₂** complexes (see Figure 4.10 and 5.11). The molar ellipticity at 222 nm in the presence of CRE DNA for Cu(II):**terpy(GCN4bd2)₂** and Zn(II):**bipy(GCN4bd2)₂** resembles those reported for CD studies of related GCN4bd artificial dimers bound to CRE DNA.[17,20] These changes are accompanied by the appearance of a positive band at 268 nm in the difference CD spectra, appearing upon addition of peptide conjugates to CRE DNA and growing in intensity upon metal addition. This can be assigned to the bending of the DNA which has previously been reported to occur upon binding of **bz** peptides (including some miniature GCN4 peptides).[12,25,26]

Upon addition of EDTA, all increase of helicity resulting from metal addition to **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** is lost, and spectra resemble those previously recorded, consistent with fully reversible processes. In some cases, EDTA addition resulted in an even greater decrease of the negative ellipticity at 222 nm, which could result from a low level of contamination from trace metal ions.

Folding constants obtained from CD titrations display a larger margin of error compared to metal-binding constants previously estimated based on UV measurements. These constants cannot be described as metal-polypyridine, nor peptide-DNA binding constants since the shifts monitored account for peptide folding, which represent indirect consequences of the metal-binding at the polypyridine linker sites. However, trends observed (such as Cu binding tighter than Zn) and the order of magnitude (compare K_{app} with K_{mf}) are consistent with those determined when monitoring by UV spectroscopy, *vide infra* (see section 4.2.2.3).

A different titration experiment was performed in order to estimate the peptide-DNA binding constants using CD spectroscopy. For this, peptide conjugates **bipy(GCN4bd2)₂** or **terpy(GCN4bd2)₂** were added stepwise to solutions containing CRE DNA, both in the absence or presence of excess metal ion. However, plots of ellipticity versus peptide conjugates equivalence, plateaued only when a high excess of peptide conjugates were present, and this coincided with precipitation (data not shown). The absence of plateau at one equivalent might arise from simultaneous specific and non-specific binding, the latter being important in the CD relevant concentration range, even though it does not result in significant peptide folding.

The helicity of conjugate **terpy(GCN4bd2)₂** in the presence of half-CRE and AP1 was shown to increase upon addition of Cu(II) and Zn(II), in much the same fashion as was previously observed in the presence of CRE DNA. Helicity is greater in the presence of AP1 DNA compared to CRE where both peptide strands are proposed to bind in an α -helical fashion and with a high affinity, in contrast to half-CRE in which only one binding site exists for one of the peptide strands to bind with high affinity. Unlike native GCN4, it appears that the Cu(II):**terpy(GCN4bd2)₂** complex is able to discriminate between AP1 and CRE sites, that differ by a single central **bp**. In contrast, metal regulated folding of **bipy(GCN4bd2)₂** is unable to discriminate between AP1 and half-CRE sites which respectively lacks one central **bp**, and half of the binding sites compared to CRE, respectively. These results are in perfect agreement with the original design which estimates that both 5,5'-disubstituted **bipy**, in both the *cis*- and *trans*- conformation, and the *cis-cis*- conformation of 6,6'-disubstituted **terpy** would result in the correct spatial alignment of the peptide substituents for binding to the CRE DNA target site. In contrast, only the *cis-cis*- conformation of 6,6'-disubstituted **terpy** was foreseen to satisfy the shorter distance between peptides when the dimers are bound to the AP1 DNA (see Chapter 3).

4.2.2.2 – Bandshift assay for GCN4bd2 conjugates

The study of the metal-dependent behaviour of **bipy(GCN4bd2)₂** was further studied, using bandshift assays. For this, solutions containing 20 nM of 5'-end labelled CRE and increasing amounts of **bipy(GCN4bd2)₂**, either in the absence or presence of unlabelled CRE (480 or 980 nM), were loaded on polyacrylamide gels (10 or 7%). Following electro-elution with different buffer (either 1× or 0.5× Tris-glycine or TG) and phosphorimaging, none of the radiograms obtained displayed a significant bandshift, even when DNA and peptide conjugates were present at micromolar concentrations (see Figure 4.14).

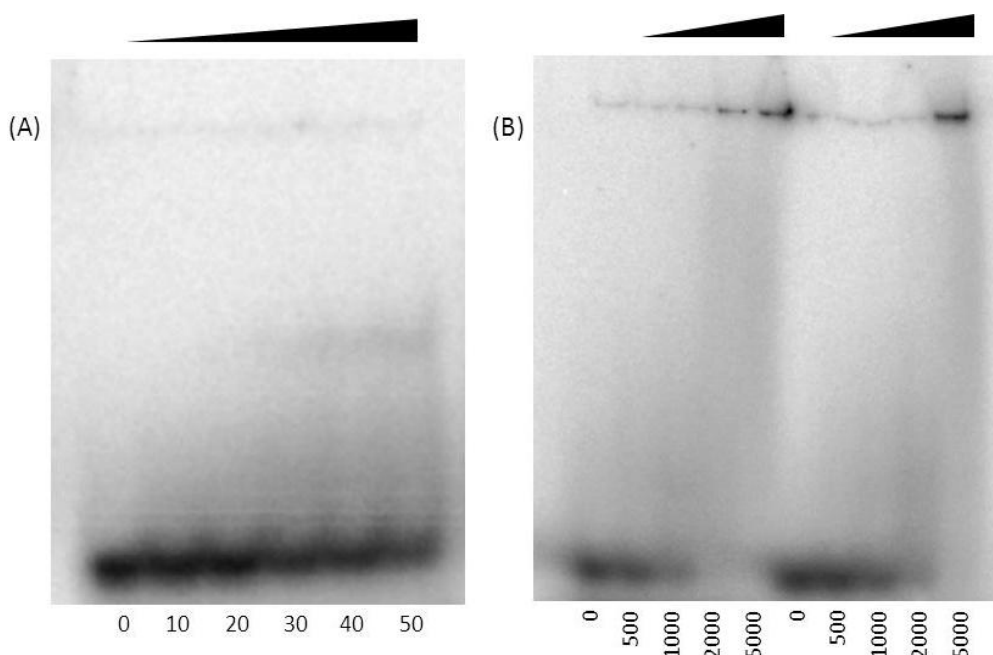


Figure 4.14 – Radiogram obtained from electro-elution at 277 K of solutions containing CRE (labelled and unlabelled), and increasing amounts of **bipy(GCN4bd2)₂** (concentration are indicated in nM below radiogram): large excess of peptide conjugates over DNA results in precipitation in wells. (A) All solutions contained 20 nM of 5'-end labelled CRE (³²P) and were loaded on 10% polyacrylamide gel in 1×TG buffer pH 8.7, eluted at 120 V. (B) Solutions containing 20 nM of 5'-end labelled CRE (³²P), unlabelled CRE (5 left slot: 480 nM; 5 right slot: 980 nM), and were loaded on 7% polyacrylamide gel in 0.5×TG buffer pH 8.7, eluted at 100 V.

Electrophoretic measurements were performed on **bipy(GCN4bd2)₂** and CRE DNA only, as conjugate **terpy(GCN4bd2)₂** and oligonucleotides containing the AP1 site were not available at the time of the experiment. No significant bandshift which could be assigned as the **bipy(GCN4bd2)₂**/CRE complex were observed, even when DNA and dimer conjugates were present at high concentrations, similar to those used for CD measurements. One can

conclude that even though CD measurements indicate interactions between **bipy(GCN4bd2)₂** and CRE DNA, the resulting complex is not stable under the electrophoretic conditions used, consistent with our peptide design that lacks the hinge area (KLQRMKQ), which is important for the complex stability.[16] By analogy, the peptide dimer **GCN4br5^{SS}** was shown, by Kim and co-workers, to bind sequence specifically to the CRE DNA site by comparative CD melting studies, even though footprinting studies indicated it does not protect this site from enzymatic cleavage.[16] For instance, the shortest artificial GCN4bd dimer reported displaying bandshift in the presence of specific DNA, has 23 residues conserved from wild type GCN4 (six more than this work).[27] As a potential alternative, a recent study used fluorescence anisotropy to estimate the peptide-DNA affinity of an artificial derivative of the homeodomain transcription factor to oligonucleotides labelled with fluorescent dyes and bearing various target sites.[28]

4.2.2.3 – Cu(II) and Zn(II) binding studies of bipy(GCN4bd2)₂ and terpy(GCN4bd2)₂ in the presence of DNA monitored by UV spectroscopy

Conversely, an investigation of the impact of DNA on the metal affinity of linkers from **GCN4bd2** conjugates was attempted in order to prove coordination was indeed still occurring at the polypyridine linker when in the presence of DNA (to rule out metal ion binding to DNA playing a role). Therefore, Cu/Zn addition to solutions containing **bipy(GCN4bd2)₂** or **terpy(GCN4bd2)₂** and one equivalent of DNA bearing either CRE or NON sites, were monitored by UV spectroscopy, and the metal-binding constants for the polypyridine peptide dimer conjugates in the presence of DNA were estimated, following methods analogous to those used in the absence of DNA (see Chapter 3). However, the band due to the $\pi \rightarrow \pi^*$ transition overlaps with a gap resulting from DNA signal subtraction (DNA nucleobases absorb light at 260 nm resulting in intense bands at 5 μ M, which saturate the detector), thus preventing an accurate quantification of the conjugates. Therefore, the concentration of

peptide conjugates for these titrations were calculated based on the stock solution concentration. The UV spectra of solutions containing 5 μM of conjugate (either **bipy(GCN4bd2)₂** or **terpy(GCN4bd2)₂**) and NON in 20 mM phosphate buffer pH 7.4 display transitions with λ_{max} at 298 (**bipy(GCN4bd2)₂**) or 296 nm (**terpy(GCN4bd2)₂**), assigned as $\pi \rightarrow \pi^*_1$. Aliquots of a stock solution of ZnCl_2 were titrated into a solution containing 5 μM **bipy(GCN4bd2)₂**/NON or **terpy(GCN4bd2)₂**/NON in 20 mM phosphate buffer pH 7.4, up to 3 equivalent of ZnCl_2 . This resulted in the steady decrease in the absorbance at 298 (**bipy(GCN4bd2)₂**) and 296 nm (**terpy(GCN4bd2)₂**), and an increase in the absorbance at 308 and 320 nm (**bipy(GCN4bd2)₂**), or at 328 and 341 nm (**terpy(GCN4bd2)₂**) (see Figures 4.15A and B). In contrast, absorbance for the new peaks were centred at 316 and 328 nm (**bipy(GCN4bd2)₂**) or at 335 and 348 nm (**terpy(GCN4bd2)₂**), upon addition of CuCl_2 into similar solutions (see Figures 4.15C and D). Isosbestic points at 303 (Zn:**bipy(GCN4bd2)₂**), 314 (Zn:**terpy(GCN4bd2)₂**), 308 (Cu:**bipy(GCN4bd2)₂**), and 319 nm (Cu:**terpy(GCN4bd2)₂**) are again consistent with clean formation of the complex.

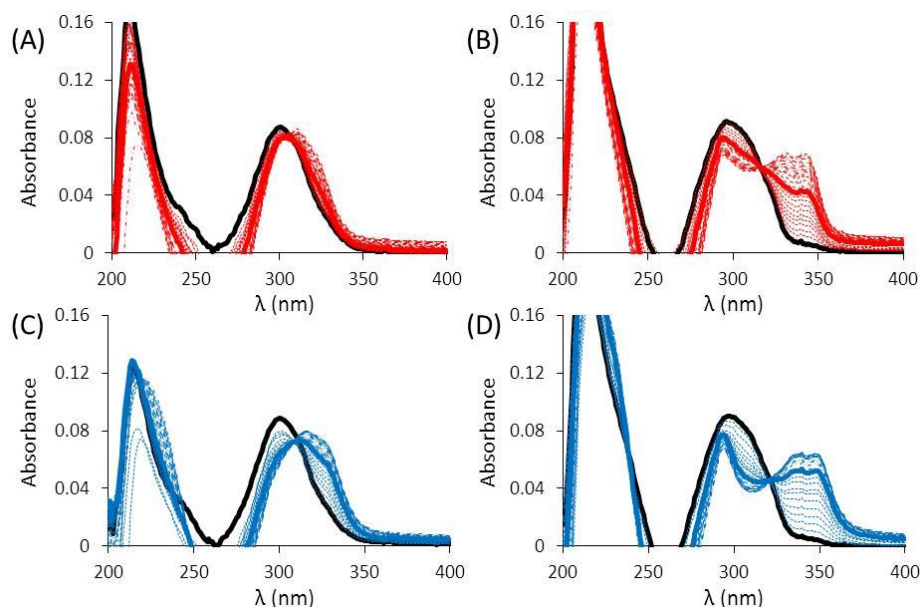


Figure 4.15 – UV spectra for the metal ion titration of solutions containing 5 μM peptide conjugates and NON DNA in 20 mM phosphate buffer pH 7.4 (DNA contribution has been subtracted). ZnCl_2 titration of (A) **bipy(GCN4bd2)₂**/NON, (B) **terpy(GCN4bd2)₂**/NON. CuCl_2 titration of (C) **bipy(GCN4bd2)₂**/NON, (D) **terpy(GCN4bd2)₂**/NON. (—) 0 eq. metal added, (■■■) between 0 and 1 eq., (—) 1 eq., (—●—) more than 1 eq. ZnCl_2 added, (■■■) between 0 and 1 eq., (—) 1 eq., (—●—) more than 1 eq. CuCl_2 added.

Plots of absorbance of the new transitions versus Cu/Zn equivalence were fitted to 1:1 binding equations (see Figure 4.16), and the extinction coefficients of the peptide dimers and the resulting Cu(II)/Zn(II) complexes determined and reported in Table 4.3. Taking into account the competitive metal ion binding of the phosphate buffer employed in these experiments,[29] binding constants in the presence of NON, $\log K_M$, were calculated to be 5.63 ± 0.09 for formation of the Zn-**bipy**(GCN4bd2)₂ complex, and 6.16 ± 0.06 for the analogous complex with **terpy**(GCN4bd2)₂. Titration of **bipy**(GCN4bd2)₂ with Cu(II) had to be performed in the presence of a competitor (10 mM L-glycine), in order to accurately estimate the Cu(II) affinity. Formation constants in the presence of NON, $\log K_M$, were determined to be 11.97 ± 0.06 for the Cu-**bipy**(GCN4bd2)₂ complex, and 8.04 ± 0.09 for the Cu-**terpy**(GCN4bd2)₂ complex.

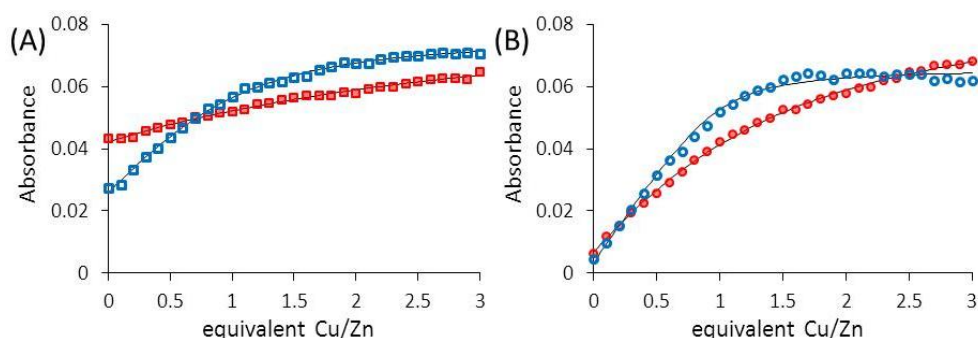


Figure 4.16 – Plot of absorbance versus the equivalence of metal ion, either ZnCl₂ (□ and ○) or CuCl₂ (□ and ○). Absorbance was monitored at (A) 320 (□, values were subtracted with absorbance at 400 nm), 328 nm (□) for **bipy**(GCN4bd2)₂/NON, and (B) 341 (○), 348 nm (○) for **terpy**(GCN4bd2)₂/NON. Cu titration of **bipy**(GCN4bd2)₂/NON was performed in the presence of 10 mM L-glycine. (—) Represents best fit for a 1:1 metal:peptide dimer conjugate binding ratio.

Table 4.3 – Summary of the UV data obtained for Cu/Zn coordination to polypyridyl peptide conjugates, **bipy**(GCN4bd2)₂ and **terpy**(GCN4bd2)₂ in the presence of 1 eq. NON DNA.

Pep. conjugate	Metal	λ (nm)	ϵ_{ML} (M ⁻¹ cm ⁻¹)	K_{app} (M ⁻¹)	K_M (M ⁻¹)	R ²
bipy (GCN4bd2) ₂	Zn	320	$1.74 \pm 0.08 \text{ E} +04^a$	$7.20 \pm 1.58 \text{ E} +04$	$4.34 \pm 0.95 \text{ E} +05$	0.9939
	Cu	328	$1.54 \pm 0.01 \text{ E} +04$	$7.34 \pm 1.10 \text{ E} +05$	$9.24 \pm 1.41 \text{ E} +11^b$	0.9979
terpy (GCN4bd2) ₂	Zn	340	$1.80 \pm 0.03 \text{ E} +04$	$2.43 \pm 0.35 \text{ E} +05$	$1.46 \pm 0.21 \text{ E} +06$	0.9987
	Cu	348	$1.32 \pm 0.01 \text{ E} +04$	$3.40 \pm 0.74 \text{ E} +06$	$1.11 \pm 0.25 \text{ E} +08$	0.9941

^aAbsorbance value at 320 nm corrected by subtraction of absorbance recorded at 400 nm.

^bTitration performed in the presence of 10 mM L-glycine as competitor.

The UV spectra of solutions containing 5 μM of conjugate (either **bipy(GCN4bd2)₂** or **terpy(GCN4bd2)₂**) and CRE DNA in 20 mM phosphate buffer pH 7.4 display transitions with λ_{max} at 299 ($\epsilon_{299 \text{ nm}}$ 17,624 $\text{M}^{-1} \text{cm}^{-1}$) (**bipy(GCN4bd2)₂**) or 296 nm ($\epsilon_{296 \text{ nm}}$ 19,134 $\text{M}^{-1} \text{cm}^{-1}$) (**terpy(GCN4bd2)₂**), assigned as $\pi \rightarrow \pi^*_1$. However, the $\pi \rightarrow \pi^*_2$ transition overlaps with a gap resulting from DNA signal subtraction. Aliquots of a stock solution of ZnCl_2 were titrated into solution containing 5 μM **bipy(GCN4bd2)₂**/CRE or **terpy(GCN4bd2)₂**/CRE in 20 mM phosphate buffer pH 7.4. This resulted in the steady decrease in the absorbance at 299 (**bipy(GCN4bd2)₂**) and 296 nm (**terpy(GCN4bd2)₂**), and an increase in the absorbance at 311 and 320 nm (**bipy(GCN4bd2)₂**), or at 328 and 341 nm (**terpy(GCN4bd2)₂**), respectively (see Figures 4.17A and B). In contrast, absorbance for the new peaks were centred at 316 and 328 nm (**bipy(GCN4bd2)₂**) or at 335 and 348 nm (**terpy(GCN4bd2)₂**), upon addition of CuCl_2 into similar solutions (see Figures 4.17C and D). Isosbestic points at 304 (Zn:**bipy(GCN4bd2)₂**), 315 (Zn:**terpy(GCN4bd2)₂**), 310 (Cu:**bipy(GCN4bd2)₂**), and 319 nm (Cu:**terpy(GCN4bd2)₂**) are consistent with clean formation of the complex.

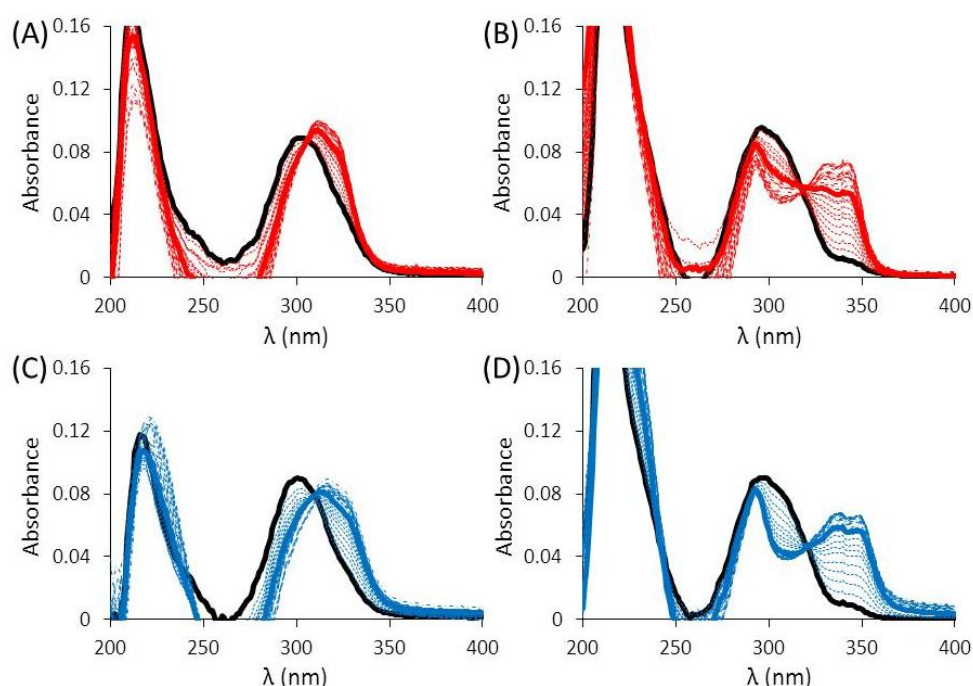


Figure 4.17 – UV spectra for the metal ion titration of solutions containing 5 μM peptide conjugates and CRE duplex DNA in 20 mM phosphate buffer pH 7.4 (contribution of DNA was subtracted). ZnCl_2 titration of (A) **bipy(GCN4bd2)₂**/CRE, (B) **terpy(GCN4bd2)₂**/CRE and CuCl_2 titration of (C) **bipy(GCN4bd2)₂**/CRE, (D) **terpy(GCN4bd2)₂**/CRE. (—) 0 eq. metal added, (---) between 0 and 1 eq., (—) 1 eq., (—•—) more than 1 eq. ZnCl_2 added, (---) between 0 and 1 eq., (—) 1 eq., (—•—) more than 1 eq. CuCl_2 added.

Plots of absorbance of the new transitions versus Cu/Zn equivalence were fitted to a 1:1 binding equation (see Figure 4.18), and the extinction coefficients of the peptide dimers and the resulting Cu(II)/Zn(II) complexes determined and reported in Table 4.4. Taking into account the competitive metal ion binding of the phosphate buffer employed in these experiments,[29] binding constants, $\log K_M$, were calculated to be 6.89 ± 0.11 for formation of the Zn-**bipy**(GCN4bd2)₂/CRE complex, and 6.62 ± 0.06 for the analogous complex with **terpy**(GCN4bd2)₂. Titration of **bipy**(GCN4bd2)₂/CRE or **terpy**(GCN4bd2)₂/CRE with Cu(II) had to be performed in the presence of 10 and 0.2 mM L-glycine, respectively, in order to accurately estimate the Cu(II) affinity. Formation constants, $\log K_M$, were determined to be 12.26 ± 0.08 for the Cu-**bipy**(GCN4bd2)₂/CRE complex, and 9.41 ± 0.09 for the Cu-**terpy**(GCN4bd2)₂/CRE complex.

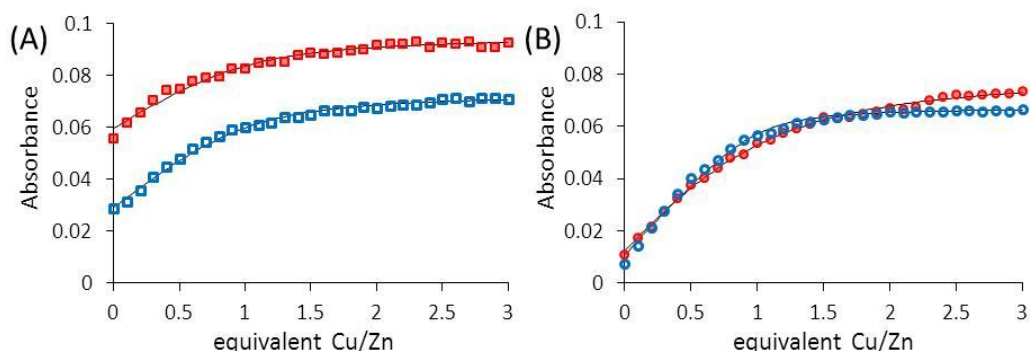


Figure 4.18 – Plot of absorbance versus the equivalence of metal ion, either ZnCl₂ (□ and ○) or CuCl₂ (□ and ○). Absorbance was monitored at 320 (□), 328 nm (□) for **bipy**(GCN4bd2)₂/CRE (A) and 341 (○), 348 nm (○) for **terpy**(GCN4bd2)₂/CRE (B). CuCl₂ titration of **bipy**(GCN4bd2)₂/CRE and **terpy**(GCN4bd2)₂/CRE were performed in the presence of 10 and 0.2 mM L-glycine, respectively. (—) represents best fit for a 1:1 metal:peptide dimer conjugate binding ratio.

Table 4.4 – Summary of the UV data obtained for Cu/Zn coordination to polypyridyl peptide conjugates, **bipy**(GCN4bd2)₂ and **terpy**(GCN4bd2)₂ in the presence of 1 eq. CRE DNA.

Pep. conjugate	Metal	λ (nm)	ϵ_{ML} (M ⁻¹ cm ⁻¹)	K_{app} (M ⁻¹)	K_M (M ⁻¹)	R ²
bipy (GCN4bd2) ₂	Zn	320	$1.90 \pm 0.02 \text{ E } +04$	$1.29 \pm 0.36 \text{ E } +06$	$7.77 \pm 2.19 \text{ E } +06$	0.9904
	Cu	328	$1.47 \pm 0.01 \text{ E } +04$	$1.45 \pm 0.27 \text{ E } +06$	$1.82 \pm 0.35 \text{ E } +12^a$	0.9945
terpy (GCN4bd2) ₂	Zn	340	$1.62 \pm 0.02 \text{ E } +04$	$6.94 \pm 1.06 \text{ E } +05$	$4.18 \pm 0.64 \text{ E } +06$	0.9976
	Cu	348	$1.35 \pm 0.07 \text{ E } +04$	$4.80 \pm 1.07 \text{ E } +06$	$2.57 \pm 0.58 \text{ E } +09^b$	0.9950

Titration performed in the presence of 10^a and 0.2^b mM L-glycine as competitor.

CuCl₂ and ZnCl₂ titrations monitored by UV were carried out in the presence of one equivalent of duplex DNA (5 μ M) containing either the CRE or NON site, so as to investigate metal binding to the polypyridine linkers under these conditions and to evaluate if metal binding to these linkers and DNA binding of the peptide substituents is cooperative. If one considers a thermodynamic cycle linking the four main species (peptide, peptide/metal, peptide/DNA, and peptide/metal/DNA) in equilibrium, any change in the metal-binding affinity of the peptide dimer conjugates is due to the peptide dimer DNA affinity which could alter the polypyridyl linkers conformation.[30,31] The affinity of **terpy(GCN4bd2)₂** in the presence of NON DNA is slightly greater for Zn(II) and almost unchanged for Cu(II) compared to the affinity determined in the absence of DNA. However, the affinity for Cu(II) and Zn(II) is enhanced 20 and 10 times, respectively, in the presence of DNA containing the CRE target site (see Figure 4.19). Therefore, one can expect the affinity of the peptide conjugates for the CRE DNA site to be increased by the same amount upon respective addition of Cu(II) or Zn(II). Therefore, the DNA binding regulation by the **terpy** linker in the **terpy(GCN4bd2)₂** conjugate appears smaller compared to that reported for **GCN4bd** dimerised with azobenzene (see section 4.1.1).[17] This result is consistent with the reversible and irreversible control of DNA binding afforded by the **terpy** and azobenzene conjugates, respectively.

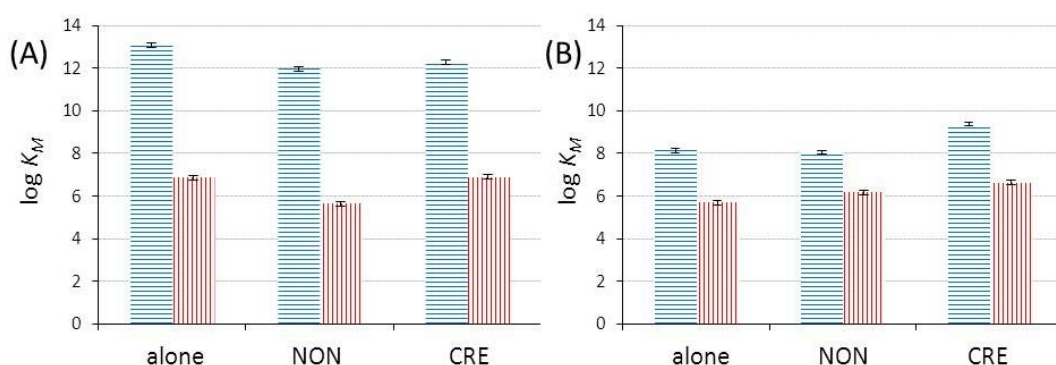


Figure 4.19 – Histogram representing the log K_M values calculated for **GCN4bd2** polypyridine conjugates in the absence or presence of DNA bearing either NON or CRE site. Comparison of calculated Cu(II) and Zn(II) affinities for (A) **bipy(GCN4bd2)₂**, or (B) **terpy(GCN4bd2)₂**. Cu(II) (horizontal stripes, ■) and Zn(II) (vertical stripes, ■); error bars displayed stand for absolute errors associated with calculation.

The Zn(II) affinity of **bipy**(GCN4**b2**)₂ decreases very slightly in the presence of NON DNA, and remains largely unchanged in the presence of DNA containing the CRE target site. The Cu(II) affinity of **bipy**(GCN4**bd2**)₂ decreases in the presence of DNA, regardless of the presence of the target site (see Figure 4.19). These results are consistent with DNA binding not promoting the more favourable *cis*- conformation for metal ion chelation on binding to DNA, consistent with a design in which both the *cis*- and *trans*- conformation are appropriate for DNA binding. Instead the DNA acts like a competitive ligand for the metal ions, either directly (metal binding to DNA), or indirectly (peptide donor atoms which contribute to the metal ion coordination sphere are no longer available when bound to DNA). The latter is less relevant to **terpy** which provides more of the donor atoms for metal ion coordination than **bipy**. Even though peptide binding to target DNA is enhanced in the presence of metal ions (due to a combination of electrostatics and conformational changes), binding to target DNA decreases the metal ion affinity (most significantly for Cu(II)).

Both peptide dimers display a higher affinity for Cu(II), which is known to favour a square planar (**bipy**)[32] or square-pyramidal (**terpy**)[33] coordination geometry, over Zn(II), which commonly adopts a tetrahedral (**bipy**)[34,35] or trigonal bipyramid (**terpy**) [33] geometry. However, peptide folding in the presence of CRE DNA, an indication of DNA binding, is greatest for **bipy**(GCN4**bd2**)₂ bound to Zn(II), for which it has a lower affinity. Whereas for **terpy**(GCN4**bd2**)₂ this is the case when bound to Cu(II), for which it has a higher affinity than Zn(II). Based on the determined binding constants, one can predict that under conditions relevant to CD measurement (Figure 4.10) ca. 99 % of Cu(II) and 87 % of Zn(II) will be bound to **terpy**(GCN4**bd2**)₂ in the presence of CRE DNA. The slightly higher occupancy of the metal binding site in the presence of two equivalents Cu(II), compared to two equivalents Zn(II), does not account for differences of peptide helicity observed in the CD spectra.

Metal ion affinity and the ability to regulate DNA binding are therefore not mutually inclusive. A parallel can be drawn with two natural transcription factors involved in metal ion regulation, through a repressive process. NmtR and SmtB adopt a similar quaternary structure, but are capable of sensing different metal ions. This is achieved due to the metal ion coordination geometry offered by the allosteric site, but is not dependent on the metal ion affinity. Robinson and co-workers report that, within a cyanobacterial cytosol, the regulatory metalloproteins *Synechococcus* PCC7942 SmtB functions in response to Zn(II) coordination but not Co(II), even though *in vitro* studies have shown that Zn(II) and Co(II) binds to the same tetrahedral allosteric site.[36] In contrast, the Ni(II) and Co(II) responsive metalloprotein NmtR from *Mycobacterium tuberculosis* regulating the *nmt* operator, contains an allosteric metal-binding site on helix $\alpha 5$ of the winged helix-turn-helix motif, similarly to SmtB, but involving two additional ligands for potential metal ion coordination in an octahedral fashion (the two additional ligands require reorganisation to bind octahedral metal ions). NmtR binds Co(II) with lower affinity compared to SmtB, but this results in a structural transition which in turn inhibits DNA binding. Moreover, Zn(II) binds to NmtR with greater affinity than Co(II), but is unable to reduce DNA binding.[37,38] These examples illustrate that metal ion allosteric regulation is largely dependent on preferred metal ion coordination geometry and how that might alter protein structure, rather than metal ion affinity. Therefore our examples above need to be considered not in terms of enhanced metal ion affinity and resulting DNA affinity, but rather in terms of preferred coordination geometry and resulting DNA affinity.

4.3 – Summary

The DNA binding studies by CD and bandshift assays of polypyridine conjugates bearing GCN4**bd** mimics, prepared in Chapter 3, is reported. CD studies are consistent with all conjugates adopting random coil secondary structures, both in the absence and presence of non-specific DNA. However, some folding of the peptide conjugates into α -helices is

promoted in the presence of DNA containing the CRE target site. For the longer design, **GCN4bd1** conjugates dimerised through metal chelation motif were shown by CD to interact in a similar fashion with the target CRE site, before and after Cu(II) or Zn(II) addition. Bandshift assays run on **pyr(GCN4bd1)₂** and CRE are consistent with the binding constant previously estimated for similar GCN4bd dimers. In contrast, the shorter **GCNbd2** conjugates/CRE complexes are not stable under our electrophoretic conditions, preventing binding constant estimation by this technique. Similar metal addition did increase the helical content of peptide moieties for the second generation conjugates bearing the shorter peptide **GCN4bd2** and the very same linkers, in the presence of target DNA sites only.

This is consistent with the hypothesis that reducing the flexibility of conjugates with careful design increases the influence of linker conformation on the DNA binding ability. These results are also consistent with our polypyridine linker design, for which metal ion coordination will promote the *cis*- or *cis-cis*- conformations, triggering the alignment of peptide moieties capable of sequence selectively binding to DNA. In comparison, Cu(II) or Zn(II) addition to solutions containing the disulphide bridged dimer (**GCN4bd2**)₂ and the CRE site do not affect the CD spectrum. Aside from the control, two conjugates **bipy(GCN4bd2)₂** which contains an allosteric ineffective **Me₂bipy** unit, and **terpy(GCN4bd2)₂** which contains the allosteric effective **Me₂terpy** unit, are studied. In the presence of CRE DNA, the former displayed a greater increase of helicity upon Zn(II) rather than Cu(II) coordination, whereas the opposite trend is observed for the latter.

The Cu(II) and Zn(II) affinity of **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** in the absence and presence of DNA containing the CRE site was estimated. Both peptide dimers display a higher affinity for Cu(II) rather than Zn(II), and for **bipy(GCN4bd2)₂** the binding constant either remains unchanged or decreases slightly in the presence of DNA. However, this is not the case for **terpy(GCN4bd2)₂**, which binds more tightly to both Cu(II) and Zn(II) in presence of CRE DNA, consistent with the *cis*- conformation required for both DNA

binding and metal ion coordination. Intriguingly, a higher metal ion affinity does not correlate with enhanced binding to CRE DNA as determined by CD, probably due to the allosteric site and metal ion coordination geometry.

Complementary CD studies performed in the presence of DNA containing the AP1 or half-CRE target site (both of which have previously been shown to interact with GCN4 *in vitro*, but with slightly different structural requirements compared to CRE), display the most significant CD change, consistent with DNA binding, for **terpy(GCN4bd2)₂** in the presence of Zn(II) and to an even greater extent with Cu(II). Significantly the enhanced AP1 binding of metallated **terpy(GCN4bd2)₂** is consistent with a *cis-cis*- terpyridine conformation which positions the two peptide moieties at an optimal distance for binding to the AP1 DNA target site. In contrast, similar metal addition did not result in significant changes in the CD profile of **bipy(GCN4bd2)₂** recorded in the presence of AP1, where the inter-peptide distance is too large for binding to the AP1 target site. Metal ion coordination and peptide realignment to promote DNA binding are fully reversible upon addition of excess EDTA.

As an extension of this work, the potential of these polypyridine peptide conjugates as sequence-selective DNA sensors or nuclease agents, was partly investigated. The preliminary results will be presented in the following chapter.

4.4 – Experimental

4.4.1 – Equipment and reagents

Mono- and dihydrogen potassium phosphate salts, sodium chloride, potassium chloride, Tris Base, DNA grade water, glycine (electrophoresis buffer only), polyacrylamide stock solution 40%, ammonium persulfate, tetramethylethylenediamine, ethylene diamine tetra acetic acid (EDTA) were all obtained from Fisher Scientific. Copper chloride (CuCl₂), glycerol, IGEPAL CA-630, L-glycine and bovine serum albumin were obtained from Sigma-Aldrich, trifluoroacetic acid (TFA) and zinc chloride (ZnCl₂) from BOC. The QIAquick

Nucleotide removal kit and the loading dye, were obtained from QIAGEN; polynucleotide kinase T4 enzyme and kinase 10x buffer were from promega. ^{32}P -ATP was obtained from Perkin-Elmer.

UV measurements were recorded in 1 cm pathlength quartz cuvette at 298 K, on Shimadzu 1800 spectrometer. CD spectra were recorded in 1 mm pathlength quartz cuvettes at 298 K on a Jasco J-715 spectropolarimeter. In all cases, difference spectra were obtained by subtracting the CD spectrum of buffer alone or in the presence of DNA (blank), from that of a mixture containing both peptide and DNA, in order to observe the CD contribution from the peptide component only. The observed ellipticities in millidegrees were converted into residual molar ellipticity, Φ_{RME} , reported in units of $\text{deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$, thus allowing estimation of secondary structure. Finally, spectra were baselined by subtracting from each data point, the average Φ_{RME} measured between 400 and 300 nm for the same spectra (a region where no intense peak are expected). Radioactivity is reported in Curie unit (Ci). Electrophoresis were run in a web scientific big runner tank, gels were visualised using a Biorad molecular imager FX, and processed with Quantity One software version 4.6.8, following exposure on a Kodak FX screen 20 x 25 cm. Figures based on GCN4 were all obtained using pymol version 1.1e and file deposited in the protein data bank (pdb). Non-linear fitting were performed using Kaleidagraph software version 4.0, as described in Chapter 2.

4.4.2 – Synthetic procedure and characterisation

Preparation of peptide and peptide conjugates were reported in the preceding chapter 3. Oligonucleotides were provided by Professor James Tucker and co-workers following preparation on an Applied Biosystems 394 DNA/RNA synthesizer. These were subsequently purified using Phenomenex Clarity Oligo-RP semi-preparative HPLC column eluted with a gradient of 0.1 M triethylammonium-acetate aqueous solution pH 7 and acetonitrile, following the methodology established by the Tucker group.[39] Following purification, oligonucleotides (single strand) were characterised by electrospray and analytical HPLC.

Table 4.5 – Summary of electrospray data and theoretical extinction coefficients for the different single-strand oligonucleotides used in this study.

Name	Sequence (5'- to 3'-)	$\epsilon_{260} (\text{M}^{-1} \text{cm}^{-1})^a$	$\text{MW}_{\text{theor.}}^b$	$\text{MW}_{\text{exp.}}$
NON1	TGGAGTATGCGTCGATTCGT	1.924 E +05	6,179.1	6,180.0
NON2	ACGAATCGACGCATACTCCA	1.972 E +05	6,055.0	6,056.0
CRE1	ACGAGATGACGTCATCTCCA	1.976 E +05	6,086.0	6,086.0
CRE2	TGGAGATGACGTCATCTCGT	1.930 E +05	6,148.1	6,149.0
Half-CRE1	TGGAGATGACGTTGTCTCGT	1.912 E +05	6,179.1	6,179.0
Half-CRE2	ACGAGACAACGTCATCTCCA	1.970 E +05	6,055.0	6,055.0
AP1-A	TGGAGATGACTCATCTCGTG	1.920 E +05	6,148.1	6,148.0
AP1-B	CACGAGATGAGTCATCTCCA	1.966 E +05	6,086.0	6,086.0

^aTheoretical absorbance values were obtained using online calculator,[40] and were used for oligonucleotides quantification. ^bMass obtained using online calculator.[41]

Oligonucleotide concentrations were estimated from the absorbance value at 260 nm, using theoretical molar absorptivity values (see Table 4.5). Duplex DNA were prepared by annealing two complementary oligonucleotide strands (50 μM in 10 mM Tris.HCl buffer pH 7 and 0.1 M NaCl) at 75 °C, and subsequently allowing the annealed DNA to return to room temperature overnight.

4.4.3 – Analytical procedures

4.4.3.1 – Measurements on GCN4bd1 and related dimer conjugates (relative to section 4.2.1)

For CD spectra of solutions containing 10 μM **GCN4bd1**, or 5 μM **pyr(GCN4bd1)₂**, **bipy(GCN4bd1)₂**, **terpy(GCN4bd2)₂**, recorded in the presence or absence of 5 μM DNA duplex bearing either NON, or CRE consensus sites, a 300 μL blank solution containing 10 mM phosphate buffer pH 7.4, reducing agents (**GCN4bd1** only), and 5 μM duplex DNA (if present) was first recorded. Equimolar amounts of peptide dimer or peptide dimer conjugate

solution was added, and the data collected after 15 minute equilibration (**GCN4bd1**: 600 μM , 3 nmol, 5 μL ; **pyr(GCN4bd1)₂**: 200 μM , 1.5 nmol, 7.5 μL ; **bipy(GCN4bd1)₂**: 300 μM , 1.5 nmol, 5 μL ; **terpy(GCN4bd1)₂**: 73.8 μM , 1.5 nmol, 20.3 μL). CD spectra are an average of 10 scans recorded between 180 and 400 nm at 200 nm min⁻¹ (0.2 nm pitch). For **GCN4bd1**, 1 mM TCEP (Figure 4.5B, and B(ii)) or alternatively 100 μM DTT (Figure 4.5B(iii)) were present in the blank solution in order to prevent sulphur oxidation and dimer formation. For metal addition to **terpy(GCN4bd1)₂**, one equivalents of CuCl₂ or ZnCl₂ (1 mM, 1.5 nmol, 1.5 μL) was added and the solution allowed to equilibrate for 15 min prior to recording spectra. Ten equivalents of EDTA per metal ion (10 mM, 30 nmol, 3 μL , pH 6-8) was added and the CD spectra recorded, in order to investigate reversibility. However, analogous experiments of **bipy(GCN4bd1)₂** (Figure 4.7) followed a similar method, but using different stock solutions (**bipy(GCN4bd1)₂**: 200 μM , 1.5 nmol, 7.5 μL ; Cu/Zn: 3 mM, 1.5 nmol, 2 μL) and spectra were instead an average of 20 scans recorded between 185 and 400 nm at 500 nm min⁻¹ (0.5 nm pitch).

!Caution the following steps need to be carried out in a designated area for handling of radioactive material!

Oligonucleotide CRE1 was labelled at the 5'-end using ³²P-ATP by preparing a mixture containing CRE1 stock solution (195 μM , 2.048 nmol, 10.5 μL), ultrapure water (1.5 μL), 10x solution of polynucleotide T4 kinase (10 unit μL^{-1} , 20 unit, 2 μL), and 10x kinase buffer (2 μL). To this solution was added ³²P labelled ATP (10mCi mL⁻¹, 3000 Ci mmol⁻¹, 13.33 pmol, 4 μL) behind a protective shield. Following mixing and centrifugation, the reaction mixture was incubated at 310 K for 1 hour. In order to deactivate the enzyme the temperature was then raised to 363 K for 1 min, followed by cooling on ice. The labelled oligonucleotide was then purified using the QIAquick nucleotide removal kit from QIAGEN (containing disposable columns and eluting solutions) following provider indications. Following purification, an

aqueous solution containing ^{32}P 5'-end labelled oligonucleotide CRE1 (59.1 μM , 2.068 nmol, 35 μL , 40 μCi) was obtained. The CRE duplex was prepared by mixing ^{32}P labelled CRE1 (59.1 μM , 591 pmol, 10 μL , 11.4 μCi), with unlabelled CRE2 (195 μM , 585 pmol, 3 μL), 10x annealing buffer containing 100 mM Tris.HCl pH 7 and 1 M NaCl (1.5 μL), and ultrapure water (0.5 μL). Following mixing and centrifugation, the mixture was heated to 353 K for 1 min, and allowed to cool down to room temperature over 4 hours. A ^{32}P 5'-end labelled DNA duplex stock solution (39 μM , 585 pmol, 15 μL , 11.4 μCi) was obtained and further diluted with ultrapure water (nb: activity displayed are indicative only and are calculated based on manufacturer calibration for day 0: 3000 Ci mmol^{-1}).

For the bandshift assay experiment, aliquots for 15 reaction mixtures (20 μL each) were prepared, containing ^{32}P duplex CRE (0.1 μM , 0.2 pmol, 2 μL , 3.9 nCi), 2 \times binding buffer (10 μL), and different amounts of **pyr(GCN4bd1)₂** (see Table 4.6). The 2 \times binding buffer was prepared following previous reports [21] and contained 20 mM phosphate buffer pH 7.4, 200 mM NaCl, 5 mM KCl, 10% glycerol, 0.2% IGEPAL CA-630, and 80 $\mu\text{g mL}^{-1}$ bovine serum albumin. Volumes of binding solutions were adjusted to 20 μL with ultrapure water and solutions were mixed and centrifuged, prior to incubation at 277 K for 1 hour.

Table 4.6 – Summary of **pyr(GCN4bd1)₂** quantities present in each binding reaction, and the volumes of stock solutions used.

Lane n°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
conc ^a (nM)	0	1	1.5	2.5	4	6.5	10	15	25	40	65	100	150	250	400
V1 ^b (μL)	-	1	1.5	2.5	4	6.5	-	-	-	-	-	-	-	-	-
V2 ^c (μL)	-	-	-	-	-	-	1	1.5	2.5	4	6.5	-	-	-	-
V3 ^d (μL)	-	-	-	-	-	-	-	-	-	-	-	1	1.5	2.5	4

^a Concentration of **pyr(GCN4bd1)₂** in each binding solution. ^{b,c,d} Volume of **pyr(GCN4bd1)₂** stock solution either 20^b, 200^c, or 2000 nM^d added to each binding solution.

In a coldroom (277 K), binding solutions were loaded on a 10% polyacrylamide (19:1 crosslinking) gel cast in running buffer. The gel was electro-eluted at 120 V for 3.5 hours in 1× TG buffer (20 mM Tris, 153 mM glycine, pH 8.7) at 277 K, and subsequently exposed for 18 hours on a Kodak-FX screen. It was then visualised using a molecular imager FX and processed.

If one considers the simple case of a protein binding a particular DNA site as a monomer, the dissociation constant is calculated following equation 4.1-4.3:

$$K_D = \frac{[P]_F [D]_F}{[PD]} \quad (4.1)$$

$$[P]_T = [P]_F + [PD] \quad (4.2)$$

$$[D]_T = [D]_F + [PD] \quad (4.3)$$

K_D represents the protein-DNA dissociation constant, $[P]_F$ and $[P]_T$ the DNA unbound and total concentration of the protein, $[D]_F$ and $[D]_T$ the protein unbound and total concentration of the DNA respectively, and $[PD]$ the concentration of the protein-DNA complex. When $[D]_F \ll K_D$, then $[P]_F \gg [PD]$ and we can write $[P]_T = [P]_F$ thus obtaining equation 4.4:

$$K_D = \frac{[P]_T [D]_F}{[PD]} \quad (4.4)$$

Therefore, the protein-DNA dissociation constant can be calculated experimentally if the proportion of free DNA over protein-DNA complex, and the total protein concentration, are known. When the concentration of free DNA is equal to the concentration of protein-DNA complex, the total protein concentration is equal to the protein-DNA dissociation constant.[42]

4.4.3.2 – Measurements on GCN4bd2 dimer, and related conjugates (relative to section 4.2.2)

For CD spectra of solutions containing 5 μM (GCN4bd2)₂, bipy(GCN4bd2)₂, or terpy(GCN4bd2)₂, and 5 μM NON, CRE, Half-CRE, or AP1 DNA, 300 μL blank solutions containing 10 mM phosphate buffer pH 7.4 and 5 μM DNA duplex were first recorded. 7.5 μL of 200 μM peptide dimer or peptide dimer conjugate solutions were added, and the data collected after 15 minute equilibration. Two equivalents of CuCl₂ or ZnCl₂ (1.5 mM, 3 nmol, 2 μL) was added and the solution allowed to equilibrate for 15 min prior to recording spectra. Ten equivalents of EDTA per metal ion (10 mM, 30 nmol, 3 μL , pH 6-8) were added and the CD spectra recorded, in order to investigate reversibility. Spectra are an average of 20 scans recorded between 185 and 400 nm at 500 nm min⁻¹ (0.5 nm pitch).

Metal titrations of peptide/CRE DNA complexes were performed at a higher concentration (10 μM bipy(GCN4bd2)₂/terpy(GCN4bd2)₂, 10 μM DNA). To a 1 mm pathlength cuvette, 60 (\pm 1) μL of a 50.0 μM solution of CRE DNA, 30 (\pm 1) μL of 100 mM phosphate buffer pH 7.4 and 195 (\pm 1) μL deionised water, were successively added. A blank spectrum was recorded, and 15.0 (\pm 0.2) μL of bipy(GCN4bd2)₂ or terpy(GCN4bd2)₂ solution (200 \pm 2 μM) was added. Samples were allowed 15 min equilibration prior to recording spectra. Aliquots of aqueous stock solutions of CuCl₂ or ZnCl₂ (300 \pm 13 μM) were then titrated into the solution containing 10.0 (\pm 0.2) μM of the peptide/CRE DNA complex, and the CD recorded after 7 min equilibration. Spectra are an average of 8 scans recorded between 190 and 300 nm at 200 nm min⁻¹ (0.5 nm pitch). The residual molar ellipticity values were corrected to account for dilution. K_{mf} values were calculated by fitting data for the ellipticity at 222 nm of the ternary complexes as a function of Cu(II)/Zn(II) concentration, following the method described for the K_{app} in Chapter 2.

!Caution the following steps need to be carried out in designated area for handling of radioactive material!

Bandshift assays run on **bipy(GCN4bd2)₂** (Figure 4.14) involved steps similar to the one previously described with little adjustments (see section 4.4.3.1). The reaction mixture for labelling step contained CRE1 stock solution (60 μM , 300 pmol, 5 μL), 10 \times solution of polynucleotide T4 kinase (10 unit μL^{-1} , 10 unit, 1 μL), 10 \times kinase buffer (1 μL), and ^{32}P labelled ATP (10 mCi mL^{-1} , 3000 Ci mmol^{-1} , 10 pmol, 3 μL). Following labelling, purification, and annealing as previously described, a 5'-end labelled DNA duplex (^{32}P) stock solution (2 μM , 90 pmol, 45 μL , 9 μCi) was obtained. The 6 + 10 binding solutions (20 μL each) contained ^{32}P duplex CRE (0.2 μM , 0.4 pmol, 2 μL , 40 nCi), and 2 \times binding buffer (10 μL) which contained 40 mM phosphate buffer pH 7.4, 100 mM NaCl, 5 mM KCl, 10% glycerol, 0.1% IGEPAL CA-630. Increasing amounts of **bipy(GCN4bd2)₂** and unlabelled CRE duplex (Figure 4.14B only) were also added to the mixture as in Table 4.7 or 5.8. Volumes of binding solutions were adjusted to 20 μL with ultrapure water and solutions were mixed and centrifuged, prior to incubation at 277 K for 30 minutes.

Table 4.7 – Summary of **bipy(GCN4bd2)₂** quantities present in each binding reaction, and volumes of stock solutions used, relative to gel in Figure 4.14A.

Lane n°	1	2	3	4	5	6
C _P ^a (nM)	0	10	20	30	40	50
V _P ^b (μL)	-	1	2	3	4	5

^a Concentration of **bipy(GCN4bd2)₂** in each binding solution. ^b Volume of **bipy(GCN4bd2)₂** 200 nM stock solution added to each binding solution.

Binding solutions relative to Figure 4.14A were loaded on a 10% polyacrylamide (19:1 reticulation) gel cast in running buffer. The gel was electro-eluted at 120 V for 4.5 hours in 1 \times TG buffer (20 mM Tris, 153 mM glycine, pH 8.7) at 277 K, and subsequently exposed for 16 hours on a Kodak-FX screen.

Table 4.8 – Summary of **bipy(GCN4bd2)₂** and unlabelled CRE quantities present in each binding reactions, and volumes of stock solutions used, relative to gel in Figure 4.14B.

Lane n°	1	2	3	4	5	6	7	8	9	10
C_P^a (μM)	0	0.5	1	2	5	0	0.5	1	2	5
C_D^b (μM)	0.48	0.48	0.48	0.48	0.48	0.98	0.98	0.98	0.98	0.98
V_{P1}^c (μL)	-	5	-	-	-	-	5	-	-	-
V_{P2}^d (μL)	-	-	1	2	5	-	-	1	2	5
V_{D1}^e (μL)	1	1	1	1	1	-	-	-	-	-
V_{D2}^f (μL)	-	-	-	-	-	1	1	1	1	1

Concentration of **bipy(GCN4bd2)₂**^a, or unlabelled CRE DNA^b in each binding solution. Volume of stock solution either 2^c, or 20 μM^d **bipy(GCN4bd2)₂** and volume of stock solution either 9.6^e, or 19.6 μM^f unlabelled CRE DNA added to each binding solution.

In contrast, binding solutions relative to Figure 4.14B were loaded on a 7% polyacrylamide (37.5:1 cross-linking) gel cast in running buffer. The gel was electro-eluted at 100 V for 5 hours in 0.5× TG buffer (10 mM Tris, 76 mM glycine, pH 8.7) at 277 K, and subsequently exposed for 14 hours on a Kodak-FX screen. Both gels were then visualised and processed.

For metal ion titrations of **bipy(GCN4bd2)₂** and **terpy(GCN4bd2)₂** in the presence of DNA monitored by UV spectroscopy, aliquots of aqueous 300 (± 13) μM stock solutions of CuCl₂ or ZnCl₂, were titrated into 600 μL of a solution containing 5.00 (± 0.19) μM of polypyridyl-peptide dimer conjugate and double stranded oligonucleotide in 20.0 (± 0.3) mM potassium phosphate buffer pH 7.4. Spectra were recorded after 3 min equilibration. When the apparent Cu(II) binding constant was too high to be accurately estimated, L-glycine was added as a competitor ligand. L-glycine was present for the reported titration of **bipy(GCN4bd2)₂** in the presence of both DNA sites (10.0 ± 0.2 mM), and of **terpy(GCN4bd2)₂** in the presence of CRE only (200 ± 4 μM). *K_{app}* values were calculated by

fitting data for the absorbance maximum of the metal complexes as a function of Cu(II)/Zn(II) concentration, and K_M were calculated following the method described in Chapter 2.

4.5 – References

-
- [1] Lucchini G., Hinnebusch A. G., Chen C., Fink G. R., *Mol. Cell. Biol.*, **1984**, 4, 1326-1333.
- [2] Hurst H. C., Protein Profiles Transcription Factors 1:bZIP Proteins, Academic Press, London, **1994**, ISSN 1070-3667.
- [3] Angel P., Karin M., *Biochim Biophys. Acta*, **1991**, 1072, 129-157.
- [4] Hope I. A., Struhl K., *Cell*, **1985**, 43, 177-188.
- [5] Arndt, K.; Fink, G. R. *Proc. Natl Acad. Sci. USA*, **1986**, 83, 8516-8520.
- [6] Hinnebusch A. G., *Microbiol. Rev.*, **1988**, 52, 248-273.
- [7] Sellers J. W., Vincent A. C., Struhl K., *Mol. Cell. Biol.*, **1990**, 10, 5077-5086.
- [8] Hope I. A., Struhl K., *Cell*, **1986**, 46, 885-894.
- [9] Agre P., Johnson P., McKnight S., *Science*, **1989**, 246, 922-926.
- [10] Landschulz W., Johnson P., McKnight S., *Science*, **1988**, 240, 1759-1764.
- [11] O' Shea E. K., Rutkowski R., Kim P. S., *Science*, **1989**, 243, 538-542.
- [12] Weiss M. A., Ellenberger T., Wobbe C. R., Lee J. P., Harrison S. C., Struhl K., *Nature*, **1990**, 347, 575-578.
- [13] Hollenbeck J. J., Oakley M. G., *Biochemistry*, **2000**, 39, 6380-6389.
- [14] Talanian R. V., McKnight C. J., Kim P. S., *Science*, **1990**, 249, 769-771.
- [15] Johnson N. P., Lindstrom J., Baase W. A., von Hippel P. H., *Proc Natl. Acad. Sci. USA*, **1994**, 91, 4840-4844.
- [16] Talanian R. V., McKnight J., Rutkowski R., Kim P. S., *Biochemistry*, **1992**, 31, 6871-6875.
- [17] Cammaño A. M., Vásquez M. E., Martínez-Costas J., Castedo L., Mascareñas J. L., *Angew. Chem. Int. Ed.*, **2000**, 39, 3104-3107.

- [18] Cuenoud B., Schepartz A., *Proc. Natl. Acad. Sci. USA*, **1993**, 90, 1154-1159.
- [19] Jiménez-Balsa A., Pazos E., Martínez-Albardonedo B., Mascareñas J. L., Vázquez M. E., *Angew. Chem. Int. Ed.*, **2012**, 51, 8825-8829.
- [20] Morii T., Saimei Y., Okagami M., Makino K., Sugiura Y., *J. Am. Chem. Soc.*, **1997**, 119, 3649-3655.
- [21] Palmer C. R., Sloan L. S., Adrian Jr J. C., Cuenoud B., Paoletta D. N., Schepartz A., *J. Am. Chem. Soc.*, **1995**, 117, 8899-8907.
- [22] Dmitrenko O., Thorpe C., Bach R. D., *J. Org. Chem.*, **2007**, 72, 8298-8307.
- [23] Burns J. A., Butler J. C., Moran J., Whitesides G. M., *J. Org. Chem.*, **1991**, 56, 2648-2650.
- [24] Woody R. W., Koslowski A., *Biophys. Chem.*, **2002**, 101-102, 535-551.
- [25] Conte M. R., Lane A. N., Bloomberg G., *Nucleic Acids Res.*, **1997**, 25, 3808-3815.
- [26] Dragan A. I., Liu Y., Makeyeva E. N., Privalov P. L., *Nucleic Acids Res.*, **2004**, 32, 5192-5197.
- [27] Ueno M., Murakami A., Makimo K., Morii T., *J. Am. Chem. Soc.*, **1993**, 115, 12575-12576.
- [28] Guerrero L., Smart O. S., Woolley G. A., Allemann R. K., *J. Am. Chem. Soc.*, **2005**, 127, 15624-15629.
- [29] Banerjee D., Kaden T. A., Sigel H., *Inorg. Chem.*, **1981**, 8, 2586-2590.
- [30] Grosseohme N. E., Giedroc D. P., Spectroscopic methods of analysis: Methods and protocols, Methods in molecular biology, Chapter 8: Illuminating allostery in metal sensing transcriptional regulators, 875, Springer science+business media, New York, **2012**, p167. DOI 10.1007/978-1-61779-806-1_8.
- [31] Hoersey I., Krishnan-Ghosh Y., Balasubramanian S., *Chem. Commun.*, **2002**, 1950-1951.
- [32] Garribba E., Micera G., Sanna D., Strinna-Erre L., *Inorg. Chim. Acta*, **2000**, 299, 253-261.

- [33] Bertrand H., Monchaud D., De Cian A., Guillot R., Mergny J.-L., Teulade-Fichou M.-P., *Org. Biomol. Chem.*, **2007**, 5, 2555-2559.
- [34] Rama G., Ardá A., Maréchal J., Gamba I., Ishida H., Jiménez-Barbèro J., Vázquez M. E., López M. E. V., *Chem. Eur. J.*, **2012**, 18, 7030-7035.
- [35] Khalighi A., Ahmadi R., Amani V., Khavasi H. R., *Acta Crystallogr. E*, **2008**, E64, m1211-m1212.
- [36] VanZile M. L., Cosper N. J., Scott R. A., Giedroc D. P., *Biochemistry*, **2000**, 39, 11818-11829.
- [37] Tottey S., Harvie D. R., Robinson N. J., *Acc. Chem. Res.*, **2005**, 38, 775-783.
- [38] Cavet J. S., Meng W., Pennella M. A., Appelhoff R. J., Giedroc D. P., Robinson N. J., *J. Biol. Chem.*, **2002**, 277, 38441-38448.
- [39] Nguyen H. V., Zhao Z., Sallustrau A., Horswell S. L., Male L., Mulas A., Tucker J. H. R., *Chem. Commun.*, **2012**, 48, 12165-12167.
- [40] <http://biophysics.idtdna.com/UVSpectrum.html>, accessed on 10/10/2012.
- [41] <http://www.basic.northwestern.edu/biotools/oligocalc.html>, accessed on 10/10/2012.
- [42] Fairall L., Buttinelli M., Panetta G., DNA-Protein Interactions, Chapter 5: Bandshift, gel retardation or electrophoretic mobility shift assays, Oxford University Press, New York, **2000**, ISBN 978-0-19-963691-4.

CHAPTER V: PRELIMINARY STUDIES ON SENSING AND NUCLEASE APPLICATIONS

5.1 – Introduction

5.1.1 – Sequence specific DNA sensors based on polypyridine and peptides

Following the hypothesis by Lerman that aromatic compounds can intercalate DNA through non-covalent interactions,[1] Lippard and co-workers showed that the complex $[\text{Pt}^{\text{II}}(\text{terpy})(\text{SCH}_2\text{CH}_2\text{OH})]^+$ interacted with DNA in a similar fashion, and that this led to an alteration in its photophysical properties.[2] The concepts of π interactions with nucleotide **bp** of DNA and shape recognition were later generalised to a broad variety of square-planar and octahedral metal complexes bearing aromatic ligands, from which new types of probes, diagnostic, and therapeutics agents were developed.[3] The design of complexes capable of sequence selective DNA intercalation based on matching shape and functionalities is challenging, though a few notable success have been reported.[4,5] An alternative strategy consists of covalently attaching metal complexes with moieties able to sequence selectively bind DNA such as urea,[6] or peptides.[7]

Barton and co-workers reported a conjugate made of a $[\text{Rh}^{\text{III}}(\text{phin})_2(\text{phen}')]^+$ complex (**phin** = 9,10-phenanthrenequinone diimine; **phen'** = 5-(amidoglutaryl)-1,10-phenanthroline) and a 13 residue peptide, derived from the α_3 -helix from the phage P22 repressor, which, upon photoactivation (see section 5.1.2), specifically cleaves DNA at the adenosine nucleotide from 5'-CCA-3' consensus sites (see Figure 5.1A). A single mutation of a glutamate residue for lysine, glutamine, aspartate, or glutamate methyl ester, abolished sequence specific binding, as indicated by the loss of cleavage specificity.[8] Also relevant to this study are the work by Ogawa and co-workers who prepared a conjugate with $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{phen-IA})]^{2+}$ (**phen-IA** = N-iodoacetyl-5-amino-1,10-phenanthroline) and GCN4**bz** peptide moieties, however the conjugate was unable to induce sequence specific photocleavage (see Figure 5.1B).[9] In both studies, the photophysical properties of the polyaromatic metal complexes peptide conjugates, and how these are affected by sequence-selective DNA binding were not investigated.

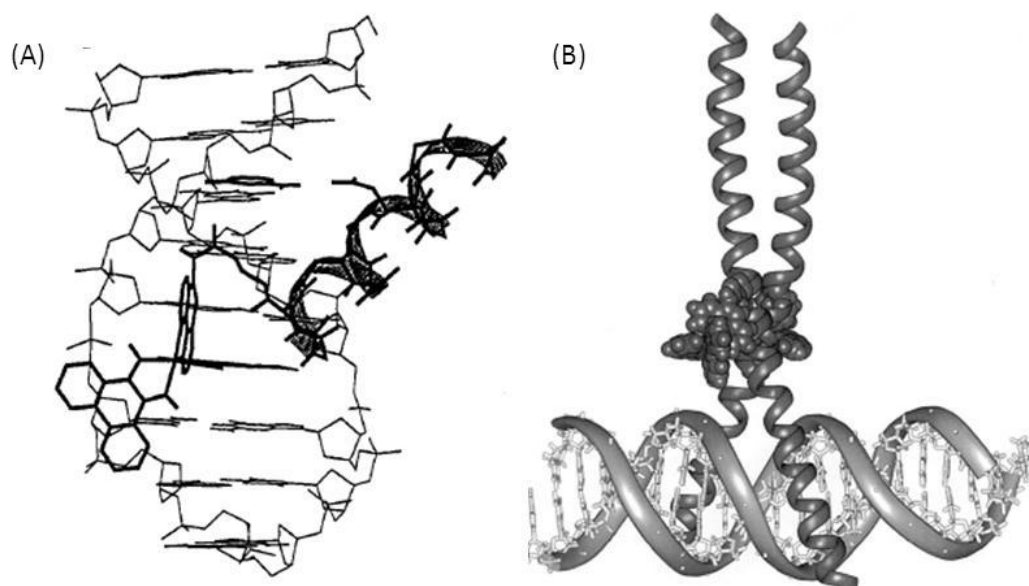


Figure 5.1 – (A) Model for the $[\text{Rh}^{\text{III}}(\text{phin})_2(\text{phen}')]]$ peptide conjugate binding to the 5'-CCA-3' DNA recognition site. The side-chain from the glutamate residue is shown forming a hydrogen bond with the 4- amino group from cytosine, reproduced with permission from reference 8. Copyright 1994 American Chemical Society. (B) Model for the $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{phen-IA})]^{2+}$ -GCN4bz conjugate binding to the AP1 DNA recognition site, reproduced with permission from reference 9. Copyright 2000 Elsevier.

Many reports describe the luminescent properties of polypyridine metal complexes,[5,10-12] and peptide based DNA sensors (including non-aromatic metal complexes peptide conjugates).[13,14] To the best of our knowledge, there is only few reports where the luminescent output of polypyridine peptide conjugates allow to discriminate between different DNA sequences,[15,16] moreover the effects described are modest.

Excitation of metal-free **bipy** and **terpy** ligands through their lower energy $\pi \rightarrow \pi^*$ ($\lambda = 298 \text{ nm}$) results in contrasted fluorescence properties. In aprotic solvents, the very low fluorescence quantum yield observed for **bipy** (5×10^{-4} in ethanol, 298 K) was explained by efficient intersystem crossing to local triplet states.[17] In contrast, these deactivation processes are less efficient for **terpy**, consistent with the higher fluorescence quantum yield (0.09 in ethanol, 300 K).[18] As for absorbance properties, the emission properties of **bipy** and **terpy** in aqueous media are largely pH-dependent. In short both ligands were previously shown to display fluorescence in basic media similar to those recorded in aprotic media (**bipy**:

Φ_F n.d., **terpy**: Φ_F 0.11). In contrast, emission profiles recorded at lower pH, associated with monoprotonated **bipy** and diprotonated **terpy** display more intense fluorescence ($[\text{bipyH}]^+$: Φ_F n.d., $[\text{terpyH}_2]^{2+}$: Φ_F 0.61) with the emission peak centred at 340 (**bipy**), or 345 and 350 nm (**terpy**).[18,19]

Metal-ion complexation can again have contrasted effects on the fluorescence spectra of polypyridine ligands. For example, Zn(II) complexation of **bipy** and **terpy** results in fluorescence enhancement, and the resulting complexes have higher fluorescence quantum yields than the mono- and diprotonated species, respectively ($[\text{bipyZn}^{\text{II}}]^{2+}$: Φ_F 0.36, $[\text{terpyZn}^{\text{II}}]^{2+}$: $\Phi_F > 0.65$).[20-22] In contrast, coordination of Cu(II) results in quasi-total quenching of the fluorescence.[22] Previous studies by Pu and co-workers have shown that, unlike the free-ligand, the complex **terpy**:Cu(II) was able to efficiently sense the Zn(II) cation among other transition metal ions by displacement of Cu(II), (see Figure 5.2A and B),[22] but also histidine among other amino-acids and imidazole derivatives without Cu(II) displacement (see Figure 5.2C).[23] In the latter case, it was hypothesised that histidine was able to coordinate the metal from **terpy**:Cu(II), through vacant orbitals, completing its coordination sphere.

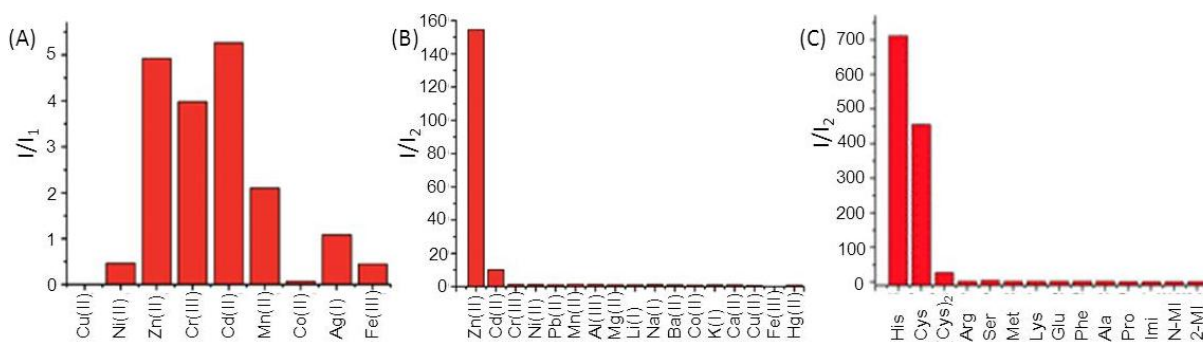


Figure 5.2 – Bar charts displaying relative fluorescence intensity ($\lambda_{\text{exc.}} = 298$ nm) in the presence of metal ions and amino-acids, by **terpy**:Cu(II) complexes. (A) Relative fluorescence intensity ($\lambda_{\text{em.}} = 338$ nm) measured upon addition of 2 eq. metal-ions to 20 μM **terpy** (I_1). (B) Relative fluorescence intensity ($\lambda_{\text{em.}} = 351$ nm) upon addition of 50 eq. metal-ion to 20 μM **terpy**:Cu(II) solution (I_2). (C) Relative fluorescence intensity ($\lambda_{\text{em.}} = 352$ nm) upon addition of 5 eq. amino-acids or imidazole derivatives to a solution containing 20 μM **terpy**:Cu(II) solution (I_2); (Cys)₂ stands for cystine, N-MI for 1-methyl-1H-imidazole, and 2-MI for 2-methyl-1H-imidazole; adapted with permissions from references 22 and 23. Copyright 2012 Royal Society of Chemistry.

5.1.2 – Polypyridine peptide conjugates as sequence specific nuclease agents

Following the discovery and study of the DNA cleavage ability of bleomycin metal complexes in the presence of oxygen,[24,25] Sigman and co-workers introduced the first artificial DNA cleavage agent based on Cu(I)-(phen) complexes.[26] The complex binds DNA through the minor groove and produces reactive oxygen species (ROS) through Fenton-type reactions between O₂ and Cu(I), which are able to abstract protons from the sugar backbone leading to phosphodiester bond cleavage (see Figure 5.3A).[27] Many metal complexes with similar nucleolytic activity were subsequently studied,[28] including Ru(II) and Rh(III) complexes the excited states of which were shown to present suitable redox properties for oxidation by oxygen and ROS production.[3] For most of the Cu(I/II) and Fe(II/III) complexes, the presence of a reducing agent is required for *in situ* reduction from Cu(II) and Fe(III) to Cu(I) and Fe(II), a pre-requisite for DNA cleavage.[27] However, some complexes are able to cleave DNA in the absence of a reducing agent.[29,30]

Similarly, the discovery of restriction enzyme functionality[31] has driven considerable interest towards the preparation of artificial molecules able to catalyse DNA phosphoester bond hydrolysis (see Figure 5.3B).[32] More efficient hydrolytic agents based on dinuclear metal complexes have appeared over the decades.[33,34] Various **bipy** or **terpy** complexes of Cu(I/II)[35-37] and Zn(II)[38,39] have been shown to cleave DNA through oxidative and hydrolytic pathways, respectively.

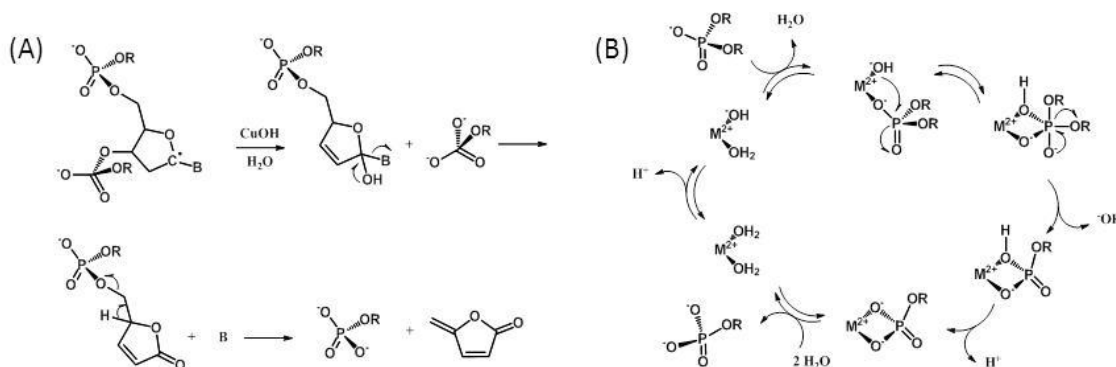


Figure 5.3 – Schemes representing either (A) a proposed mechanism for oxidative DNA cleavage following the H-1' abstraction by Cu^I(phen), where B represents any **bp**;^[27] or (B) a consensus catalytic cycle for the hydrolysis of phosphoester bonds from the DNA backbone promoted by metal ions.^[40]

Preceding the design of binary species with Rh(III) complexes and sequence-specific peptide moities (see section 5.1.1), Sigman and co-workers covalently linked Cu(I)-(phen-**IA**) with the tryptophan repressor *trp*, thus obtaining an artificial restriction enzyme of high-specificity (7 **bp** recognition site). Interestingly, DNA cleavage occurs only in the presence of tryptophan, which is a condition for DNA binding of the natural repressor protein.[41] These led to the development of numerous selective artificial nucleases, based on the combination of metal complexes and sequence-specific peptides, capable of cleaving DNA through oxidative,[42-45] but also hydrolytic pathways.[46,47] It is worth mentioning the GCN4**bz**-Fe(II)(EDTA) conjugates (complex is coupled to the N-termini) prepared by Dervan and co-workers, the sequence specific cleavage pattern of which allowed for a better understanding of the protein binding mode, ca that the N-termini of the GCN4**bz** peptide dimer are located in close proximity of the DNA major groove and separated from each other by 9-10 **bp**.[48]

5.1.3 – Aims of the chapter

This chapter describes the fluorescence properties of the polypyridine peptide conjugates and the corresponding Cu(II) complexes and the preparation of new Ru(II) containing polypyridine metal complexes with interesting photophysical properties with respect to sequence selective DNA sensing. The second objective was to test some of the polypyridine peptide conjugates for their ability to perform sequence selective DNA cleavage. Preliminary results are presented.

5.2 – Results and discussion

5.2.1 – Polypyridine-peptide conjugates as sensors

5.2.1.1 – Polypyridine ligand fluorescence applied to DNA sensing

As an extension of the work on polypyridine peptide conjugates complexed with Cu(II) or Zn(II), their fluorescence properties were investigated both in the absence or presence of DNA with specific or non-specific sequences. As a starting point, the emission spectra of

model compounds were evaluated in the absence or presence of CuCl_2 . Upon excitation at 298 nm, the emission spectra of solutions containing 20 μM **bipy-GS₂** or **terpy-GS₂** in 20 mM phosphate pH 7.4 display bands centred at 350 nm and 334 nm (**bipy-GS₂**: $I_{350 \text{ nm}}$ 17.6 a.u.; $I_{334 \text{ nm}}$ 16.1 a.u.), or 351 nm (**terpy-GS₂**: $I_{351 \text{ nm}}$ 317.3 a.u.). Upon addition of CuCl_2 , the intensity of the emission bands decrease and the weak band observed for **bipy-GS₂** is almost not distinguishable after one equivalent of CuCl_2 has been added (**bipy-GS₂**: $I_{350 \text{ nm}}$ 17.6 \rightarrow 6.8 a.u.; $I_{334 \text{ nm}}$ 16.1 \rightarrow 4.8 a.u.). In contrast, it is possible to observe that the more intense emission band of **terpy-GS₂** retains ~23 % of its intensity after one equivalent of CuCl_2 has been added (**terpy-GS₂**: $I_{351 \text{ nm}}$ 317.3 \rightarrow 74.3 a.u.), however, this further decreases to 2 % upon addition of a second equivalent CuCl_2 (**terpy-GS₂**: $I_{351 \text{ nm}}$ 74.3 \rightarrow 7.2 a.u.) (see Figure 5.4).

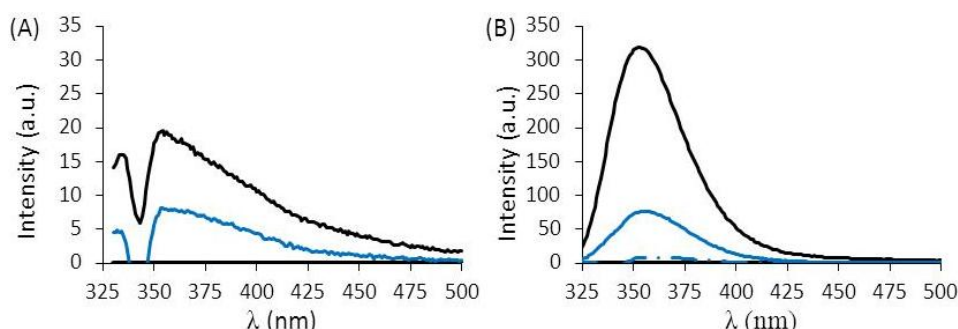


Figure 5.4 – Emission profile ($\lambda_{\text{exc.}} = 298 \text{ nm}$) recorded for 20 μM **bipy-GS₂** (A), or **terpy-GS₂** (B) solutions in 20 mM phosphate buffer pH 7.4, recorded in the absence (—), or presence of CuCl_2 , either 1 (—), or 2 eq. (---).

The emission profile recorded for **bipy-GS₂** at pH 7.4, which is expected to predominantly exist as the *trans*- conformation, displays a very weak emission band, consistent with previous studies of **bipy** fluorescence.[17] Even though the proportion of monoprotonated (*cis-trans*-) over basic form (*trans-trans*-) is expected to be low at this pH,[49] the emission profile of **terpy-GS₂** at pH 7.4 displays much more intense emission bands compared to **bipy-GS₂**, which is consistent with previous reports for **terpy** at neutral pH.[18,23] Interestingly, the decrease in fluorescence of **terpy-GS₂** upon CuCl_2 addition, correlates with the expected amount of unbound **terpy-GS₂** under these conditions (1 eq. Cu :

12-29 %, 2 eq. Cu : 1-9 %), calculated from the two binding constants extracted from the UV titration data (see Chapter 3). Emission spectroscopy therefore represents an alternative method by which to estimate the CuCl_2 affinity of **terpy-GS₂** and related **terpy** ligands.

Histidine was previously hypothesised to bind **terpy**Cu(II) through the vacant metal orbitals, thus explaining the significant increase observed in fluorescence.[23] These vacant orbitals might also be accessible to donor atoms from DNA or peptide, when **terpy**(GCN4bd2)₂Cu(II) is complexed with specific rather than non-specific DNA due to structural constraints, and these could result in contrasted fluorescent properties. The emission profile of 10 μM solutions of **terpy**(GCN4bd2)₂ in 10 mM phosphate pH 7.4, recorded in the absence or presence of 10 μM DNA containing either NON or CRE sites, all display a band centred around 338 nm of differing intensities (alone: $I_{338\text{ nm}}$ 305.9, NON: $I_{342\text{ nm}}$ 109.4, CRE: $I_{342\text{ nm}}$ 140.2). Upon addition of CuCl_2 , the emission bands all decrease in intensity (see Figure 5.5).

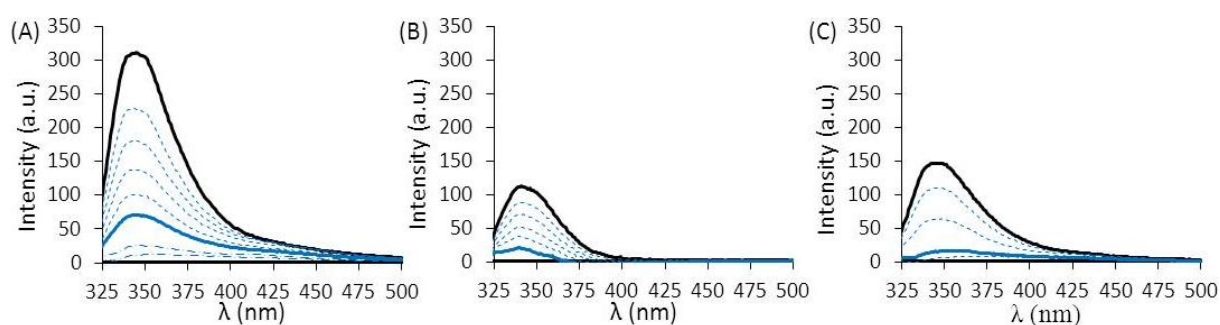


Figure 5.5 – Emission profile ($\lambda_{\text{exc.}} = 298\text{ nm}$) of 10 μM **terpy**(GCN4bd2)₂ solutions in 10 mM phosphate buffer pH 7.4, recorded in the absence (A), or presence of 10 μM duplex DNA, containing either the NON (B), or CRE site (C). CuCl_2 was titrated into each solution: 0 eq. (—), between 0 and 1 eq. (---), 1 eq. (—), more than 1 eq. CuCl_2 added (-.-.).

Plots of emission intensity at 338 nm versus Cu(II) equivalence are largely consistent with formation of the 1:1 complexes previously described, in the absence of DNA or in the presence of DNA containing the CRE target site (see Figure 5.6A and C). However, the number of data points recorded are not sufficient to conclude for spectra recorded in the presence of DNA bearing the non-specific site (see Figure 5.6B).

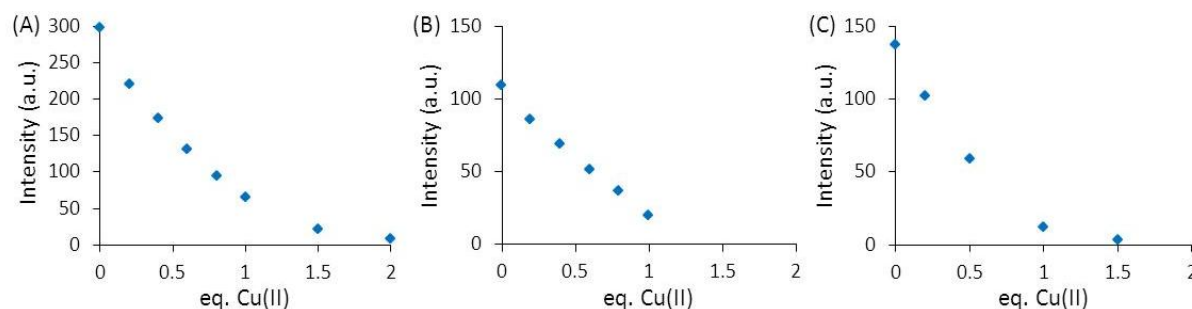


Figure 5.6 – Plot of emission intensity at 338 nm ($\lambda_{\text{exc.}} = 298$ nm) as a function of Cu(II) equivalence for the Cu(II) titration of solutions containing 10 μM **terpy(GCN4bd2)₂** in 10 mM phosphate buffer pH 7.4, recorded in the absence (A), or presence of 10 μM duplex DNA containing either the NON (B), or the CRE target site (C).

Solutions containing 20 μM **terpy-GS₂**, or 10 μM **terpy(GCN4bd2)₂** display similar emission profiles with peaks of similar intensity but shifted by 10 nm. In contrast, the intensity is greatly reduced when **terpy(GCN4bd2)₂** is recorded in the presence of DNA, regardless of the sequence. Considering the different concentrations used, it seems the presence of positively charged peptide moieties results in an increase of the fluorescence intensity, whereas it decreases upon complexation with negatively charged DNA. In all cases, addition of CuCl_2 results in quenching of the fluorescence and do not allow for a satisfactory discrimination between the DNA sequences of interest. However, profiles recorded are consistent with the affinity measured in the absence or presence of DNA by UV spectroscopy, and fluorescence data might therefore be used to independently confirm the affinity calculations.

5.2.1.2 – Attempts at synthesising $[\text{Ru}^{\text{II}}(\text{bipy})_3]^{2+}$ peptide conjugates and investigating their phosphorescent properties

The preparation of some **GCN4bd2** conjugates dimerised through a $[\text{Ru}^{\text{II}}(\text{bipy})_3]^{2+}$ linker was envisioned, the phosphorescence of which might be able to discriminate between specific versus non-specific DNA. In contrast, $\text{Ru}(\text{terpy})_2$ complexes usually display weak

emission signals at room temperature,[50] and therefore preparation of related peptide conjugates was not attempted.

Despite previous reports which describe coupling between $[\text{Ru}^{\text{II}}(\text{bipy})_3]^{2+}$ and peptides (see Figure 5.7, route A),[51-54] a methodology that involved the **bipy**-peptide conjugates previously prepared would involve minimum synthetic steps (see Figure 5.7, routes B and C). In a recent report, McLaughlin and co-workers report the preparation of $[\text{Ru}^{\text{II}}(\text{bipy})_3]^{2+}$ -oligonucleotide conjugates by refluxing **bipy**-oligo conjugates with $[\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2]$ in an ethanol:100 mM phosphate pH 7 (2:3) solvent mixture.[55] This strategy was applied to the preparation of $[\text{Ru}^{\text{II}}(\text{bipy})_3]^{2+}$ -peptide conjugates (see Figure 5.7, route C).

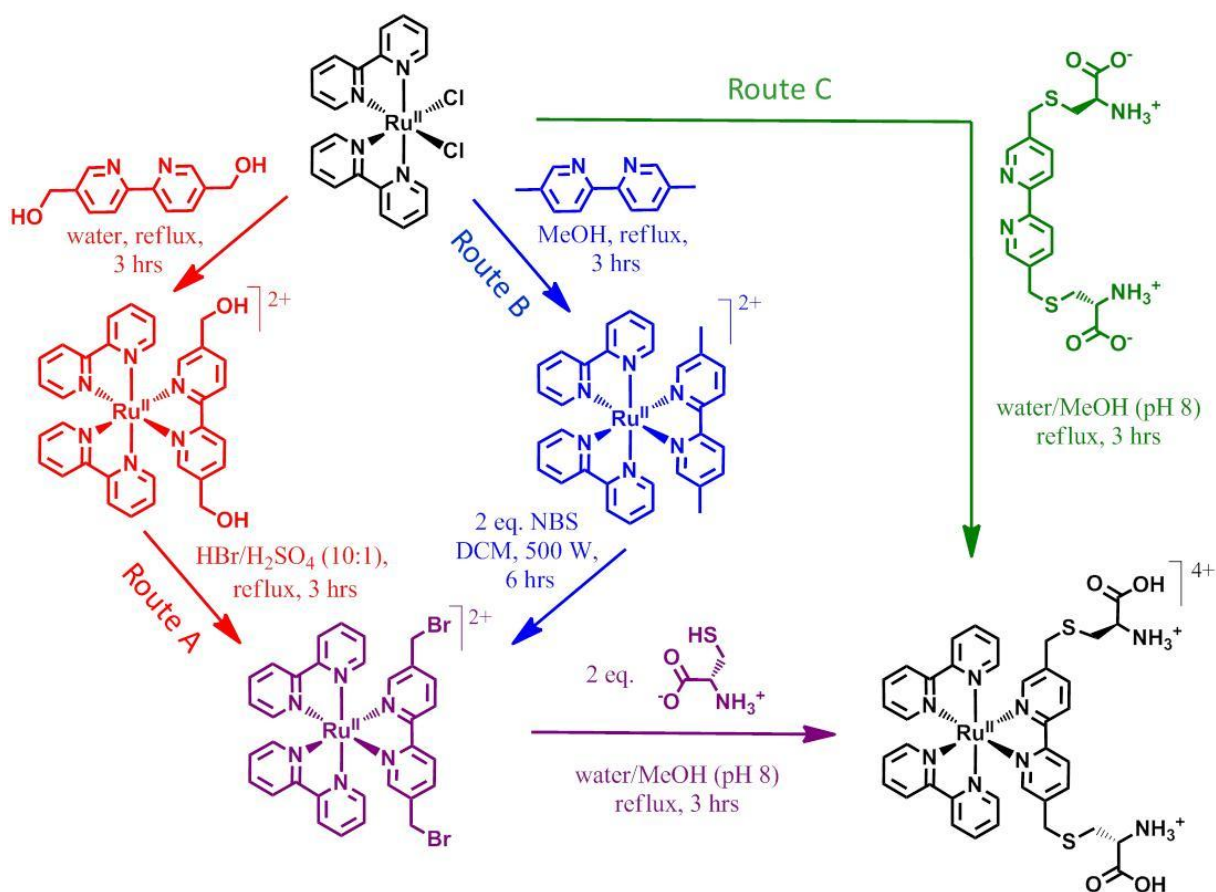


Figure 5.7 – Scheme representing the different potential strategies for the preparation of $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ complex from $[\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2]$. Route A (—) is an adaptation of the Millet *et al.* method,[52] route B (—) remains to be investigated, and route C (—) from the McLaughlin *et al.* method.[55]

The preparation of the conjugate $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ was first attempted, in order to investigate the chemistry. After refluxing **cys₂bipy** and $\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2$ in ethanol:10 mM phosphate pH 7.4 (2:3) for 3.5 hours, and subsequent purification by semi-preparative HPLC (eluted with water (0.05% TFA)/methanol gradients), the conjugate $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ was obtained as a salt, and characterised by ^1H NMR spectroscopy (see Figure 5.8). A new spectrum recorded after 10 days displayed small peaks in the aliphatic region, and integral values for H^7 , H^8 and H^9 decreased in intensity, consistent with degradation of the conjugate complex possibly associated with a loss of the L-cystenyl moieties.

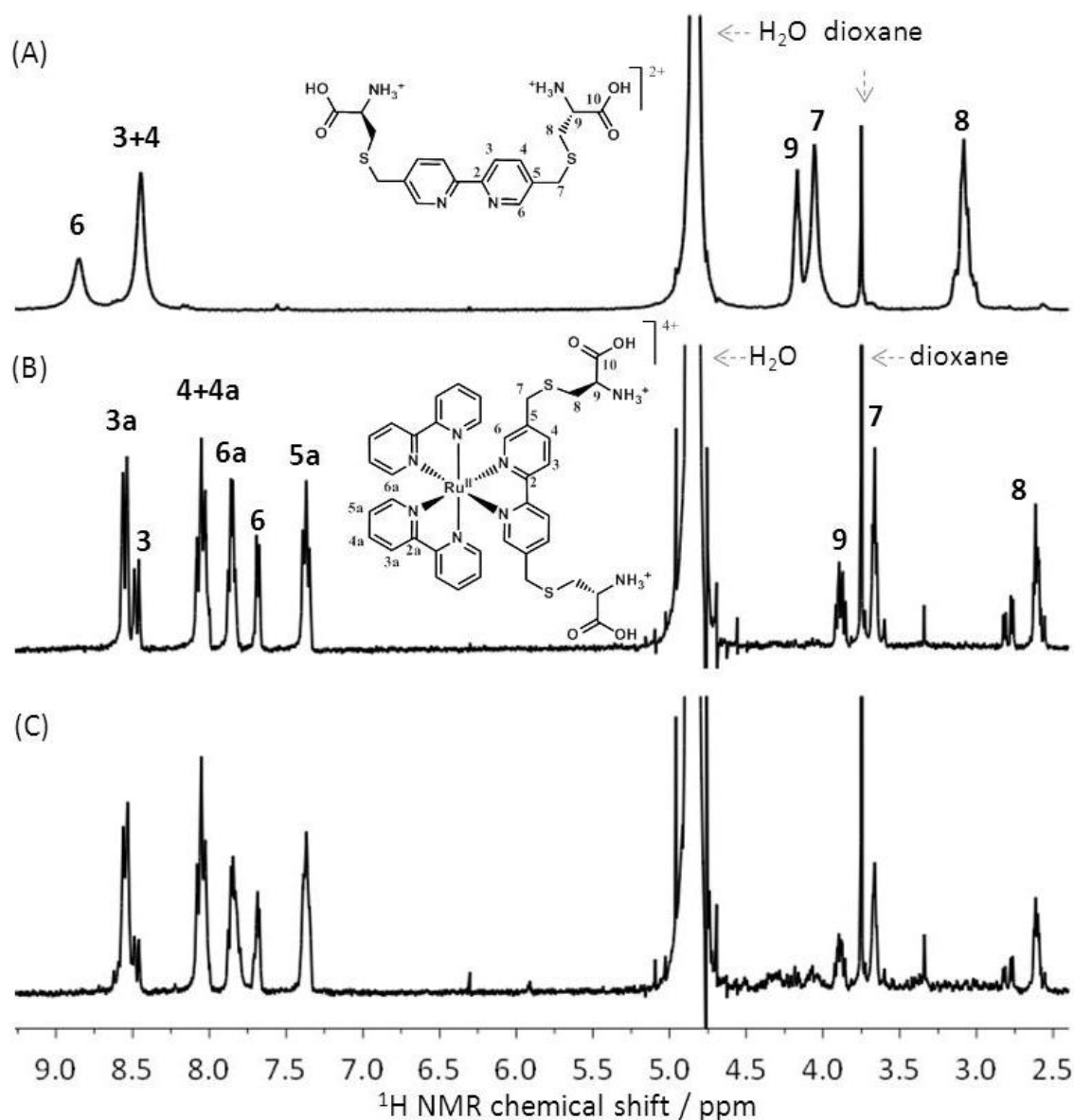


Figure 5.8 – Comparison of ^1H NMR spectra for **cys₂bipy** (A) or $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ complex recorded at 300 MHz in D_2O (pD ~1) directly (B) and 10 days after preparation (C); dioxane was added as internal standard ($\delta = 3.75$ ppm).

The absorption spectra of 6.7 μM $\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2$ recorded in ethanol:10 mM phosphate pH 7.4 (2:3), display peaks at 517 ($\epsilon_{517\text{ nm}} 3,238\text{ M}^{-1}\text{ cm}^{-1}$, $^1\text{MLCT}$), 365 ($\epsilon_{365\text{ nm}} 3,253\text{ M}^{-1}\text{ cm}^{-1}$, $^1\text{MLCT}$), 296 ($\epsilon_{296\text{ nm}} 23,238\text{ M}^{-1}\text{ cm}^{-1}$, $\pi\rightarrow\pi^*$) and 250 nm ($\epsilon_{250\text{ nm}} 7,451\text{ M}^{-1}\text{ cm}^{-1}$, $\pi\rightarrow\pi^*$). In contrast, the absorption spectra of $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ recorded under similar conditions displays peaks at 429 ($\epsilon_{429\text{ nm}} 3,427\text{ M}^{-1}\text{ cm}^{-1}$, $^1\text{MLCT}$), 289 ($\epsilon_{289\text{ nm}} 41,829\text{ M}^{-1}\text{ cm}^{-1}$, $\pi\rightarrow\pi^*$) and 244 nm ($\epsilon_{244\text{ nm}} 24,776\text{ M}^{-1}\text{ cm}^{-1}$, $\pi\rightarrow\pi^*$) (assignments based on references 56-58) (see Figure 5.9A). Upon excitation at 456 nm, the emission spectra of a 50 μM $\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2$ solution displays a weak band centred at 533 nm ($I_{533\text{ nm}} 8.8\text{ a.u.}$). In contrast, the emission spectra of $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ recorded under identical conditions displays two intense peaks at 618 ($I_{618\text{ nm}} 198.6\text{ a.u.}$), and 526 nm ($I_{526\text{ nm}} 97.4\text{ a.u.}$) (see Figure 5.9B).

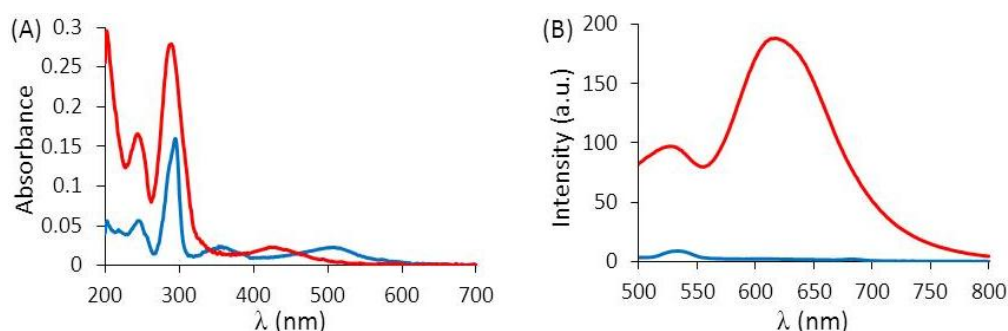


Figure 5.9 – Comparison of absorption (A), and emission spectra (B) of $\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2$ (—), and $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ (—) recorded in an ethanol:10 mM phosphate pH 7.4 (2:3) mixture. Concentration of complexes were either 6.7 μM (A), or 50 μM (B); $\lambda_{\text{exc}} = 456\text{ nm}$.

Absorption and emission ($\lambda_{\text{exc}} = 456\text{ nm}$) spectra recorded for $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ compared to those of $\text{Ru}^{\text{II}}(\text{bipy})_2\text{Cl}_2$ are consistent with coordination of an additional **bipy** unit to the metal center.[59,60] The contrast between the two emission spectra provides an efficient way to monitor the formation of $[\text{Ru}^{\text{II}}(\text{bipy})_2((\text{peptide})_2\text{bipy})]^{2+}$ complexes, specifically using HPLC coupled to a fluorescence detector.

The preparation of the conjugate $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{bipy}(\text{GCN4bd2})_2)]^{2+}$ was attempted following a similar experimental methodology, but no evidence suggested formation of the desired product. The reaction was then attempted starting from **bipy**(GCN4bd2)₂ and $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{acetone})_2]^{2+}$ which is already charged and might be more reactive towards a strained bipy unit, however this was also unsuccessful. At this stage, it remains unclear

whether the complex $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{bipy}(\text{GCN4bd2})_2)]^{2+}$ is easier to prepare by a peptide-**bipy** coupling step followed by complexation (see Figure 5.7, route C), or by synthesis of the intact ruthenium linker and subsequent coupling to the peptides (see Figure 5.7, routes A or B).

5.2.2 – Polypyridine-GCN4bd peptides as sequence-specific nuclease agents

5.2.2.1 – Modelling of GCN4bd peptide conjugates

Metal-based nuclease agents cleave DNA either through hydrolysis or oxidative mechanisms. DNA cleavage through hydrolysis involves coordination of the metal center to oxygen from the DNA backbone, thus facilitating cleavage/hydrolysis of a phosphoester bond (-P-O-).[61] In contrast, DNA cleavage through oxidative mechanisms involve production of ROS that can diffuse and abstract protons from sugar [27] or nucleobases [62] (see Figure 5.10), ultimately leading to DNA cleavage through rearrangement reactions. The efficiency is highly dependent on “target” hydrogens accessibility to the ROS.[27] As an indication, OH^\bullet radicals generated by Cu^{I} -**phen** complexes were estimated to diffuse and cleave DNA over a 15 Å radius,[63] however diffusion is prevented in the presence of radical quenchers such as glycerol.[64]

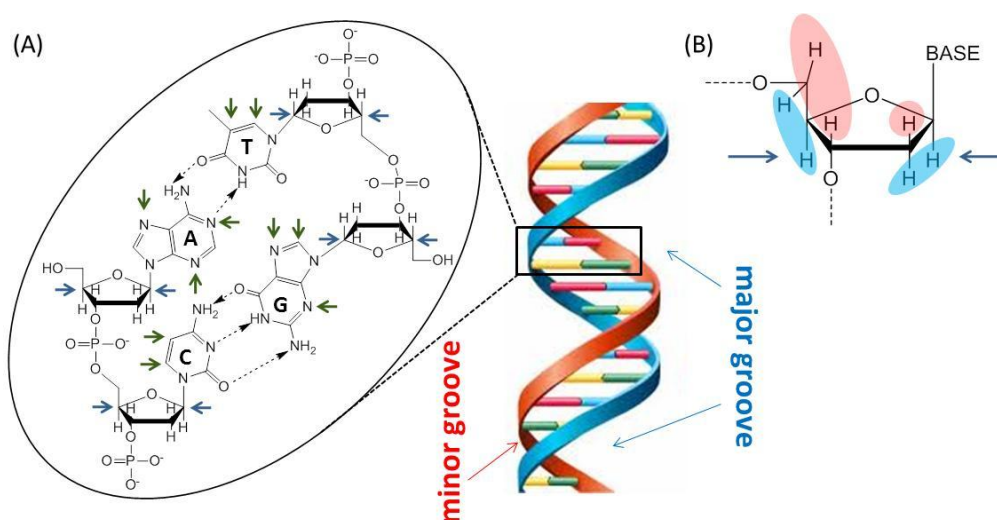


Figure 5.10 – (A) Cartoon representation of the DNA double helix and localisation of the minor and major groove: zoom and chemdraw representation of two **bp**, and potential oxidation sites are marked by arrows (blue = sugar oxidation; green = base oxidation). (B) Chemdraw representation of DNA deoxyribose, showing the accessibility of the proton (potential oxidation sites) from the minor (red) or major groove (blue); adapted from reference 65.

Simple models generated from the DNA bound GCN4 crystal structure (pdb code 1YSA[66]), indicate that metal-ions coordinated to polypyridine linkers of **bipy(GCN4bd1)₂** or **bipy(GCN4bd2)₂** complexes would be located up to 28-30, or 11-12 Å away from the nearest backbone oxygen, respectively (see Figure 5.11). These represent distances which are too long for direct coordination of the metal to DNA oxygens, therefore, both conjugates are not expected to be able to cleave DNA through a hydrolytic process. However, 11-12 Å is a suitable distance for oxidative cleavage and production of ROS by the shorter **GCN4bd2** conjugates, which might therefore be able to cleave DNA, consistent with previous designs of sequence-specific nuclease agents.[42]

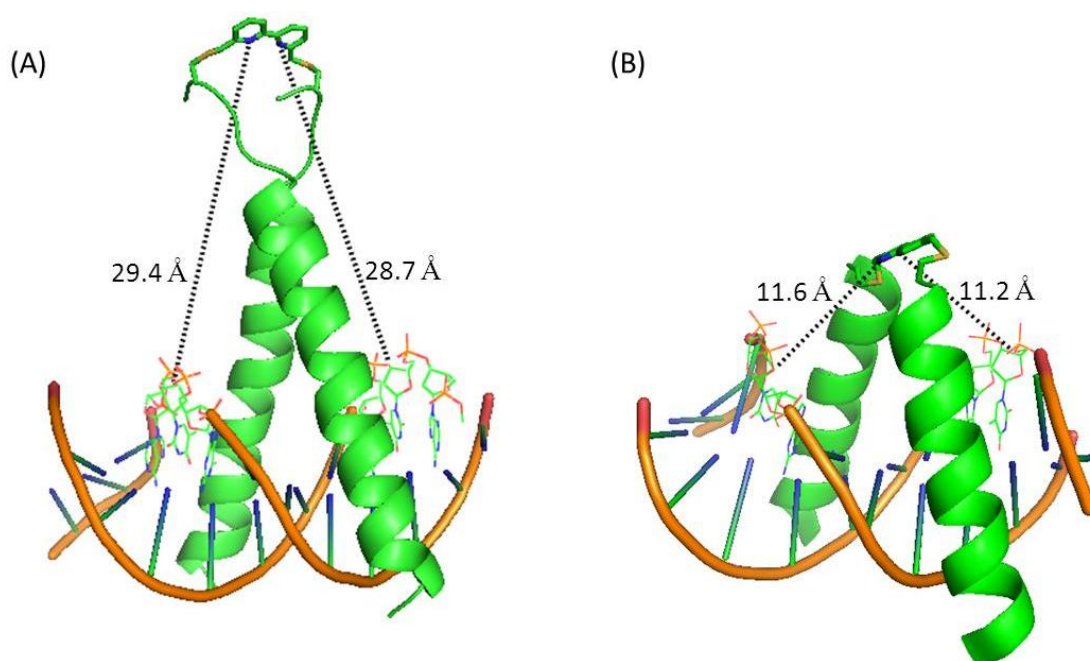


Figure 5.11 – Cartoons showing distances separating polypyridine linker nitrogen from the closest oxygens from the DNA backbone in the **bipy(GCN4bd1)₂:AP1** (A), and **bipy(GCN4bd2)₂:AP1** complexes (B). Cartoons were obtained by qualitative modification of pdb file 1YSA [66] using pymol, without energy minimisation. The glycine-glycine-cysteine domain in (A) is represented as an unfolded domain (consistent with our CD measurements) and therefore fully extended.

5.2.2.2 – Cleavage studies of supercoiled plasmid containing specific site by Cu:**bipy(GCN4bd2)₂**

In order to study sequence specific oxidative cleavage, a procedure similar to that reported by Giovannangeli and co-workers was used.[67] A plasmid vector *pUC57* (> 95 %

supercoiled) consisting of 2,710 **bp**, plus 20 **bp** oligonucleotide CRE duplex inserted at the *EcoRV* site by blunt end ligase was purchased from Genscript (see Figure 5.12). This artificial construct contains only one 5'-TGACGTCA-3' repeat, which is a target site of **bipy(GCN4bd)₂** (see Chapter 5).

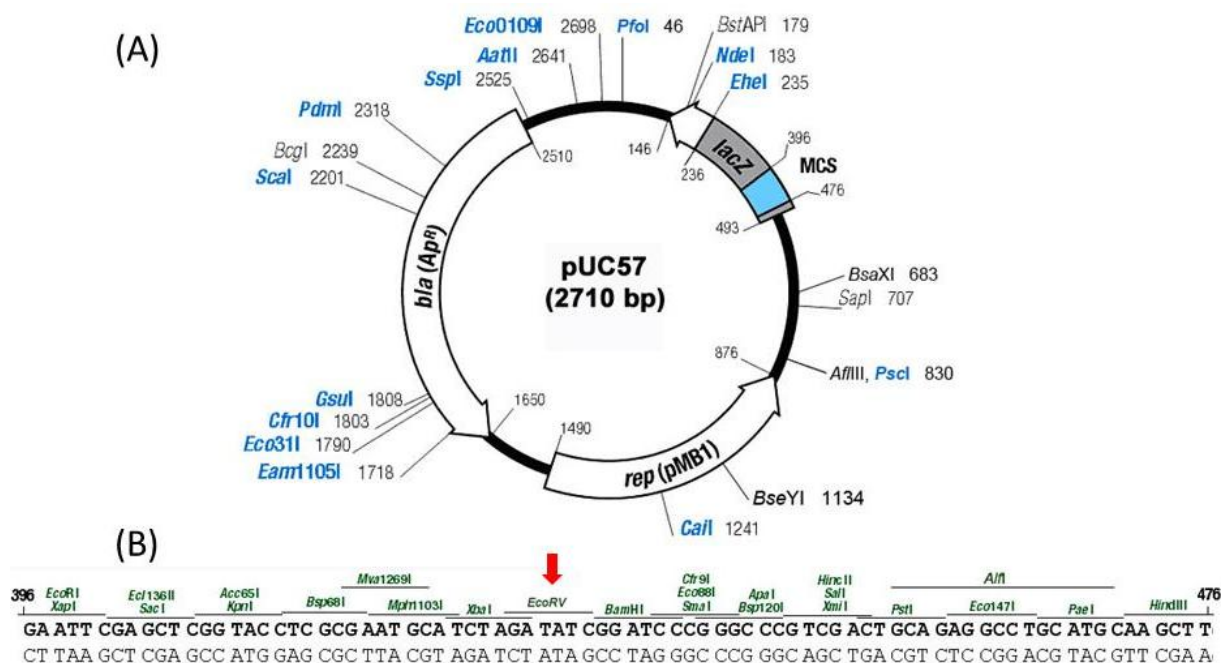


Figure 5.12 – (A) Diagram representing the different domains and restriction sites in the *pUC57* artificial vector, for which **bp** numbering is indicated, reproduced from Genscript.[68] (B) Full sequence display for the Multiple Cloning Site (MCS) domain of *pUC57* (396-476): CRE 20 **bp** duplex was inserted at *EcoRV* site (marked by an arrow).

The supercoiled plasmid *pUC57*-CRE was incubated at 310 K either alone, or in the presence of the restriction enzyme *AatII*, which cleaves 5'-GACGTC-3' sites including *AatII* (see Figure 5.12) and CRE sites, thus affording two linear fragments of 513 and 2,217 **bp**, or in the presence of Cu(II), 3-mercaptopropionic acid (MPA), and increasing amount of **bipy(GCN4bd)₂** this would result in a 2,730 **bp** linear fragment if cleavage occurs at CRE. The different solutions were resolved by horizontal electrophoresis, and visualised by UV transillumination following staining with ethidium bromide.

The gel radiogram obtained indicates that plasmid *pUC57*-CRE alone elutes as two bands: a more intense fast-migrating band attributed to supercoiled content, and a weaker

slow-migrating band consistent with the presence of circular plasmid (see Figure 5.13 lane 1). Following co-incubation with *AatII* at 310 K for 1 hour, *pUC57*-CRE eluted as a single band with retention different from that of DNA alone and consistent with the expected ~2,200 bp linear fragment (see Figure 5.13 lane 2). The short fragment (513 bp) would not be expected to be retained on the gel under the conditions used. After incubation at 310 K for 2 hours in the presence of 25 μ M Cu(II) and 100 μ M MPA, *pUC57*-CRE still elutes as two bands with retentions similar to the plasmid alone, but the intensity of the slow- and fast-migrating band increases and decreases respectively, consistent with single-strand cleavage (see Figure 5.13 lane 3). In contrast, elution indicates the presence of a third band when increasing amounts of **bipy(GCN4bd2)₂** (1-5 μ M) were also present during incubation (see Figure 5.13 lane 4-8). The new band, consistent with a linear 2,730 bp fragment associated with a double-strand cut, has intermediary retention and its intensity increases with **bipy(GCN4bd2)₂** concentration, accompanied by a decrease of the fast-migrating band intensity, whereas the slow-migrating band remains unchanged. Finally, a *pUC57*-CRE solution incubated with 5 μ M **bipy(GCN4bd2)₂** and 100 μ M MPA, but in the absence of Cu(II), displays an elution profile similar to that of *pUC57*-CRE alone (see Figure 5.13 lane 9).

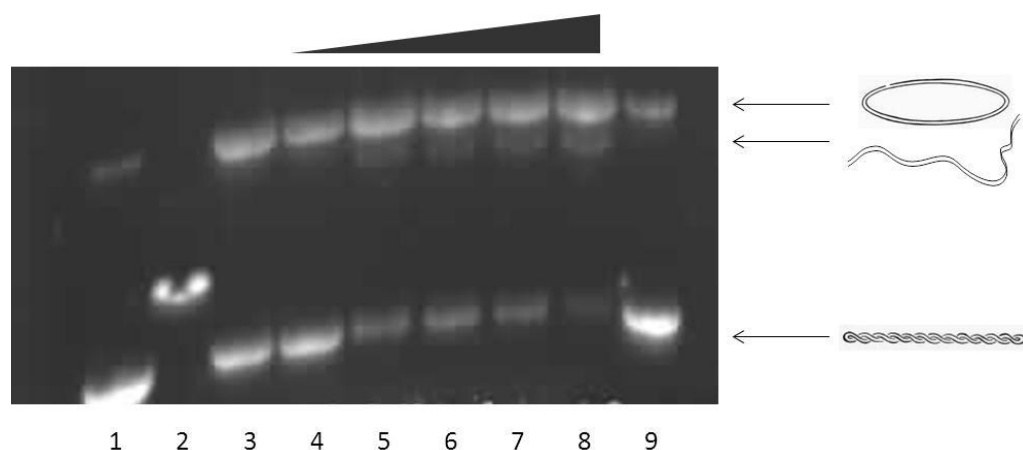


Figure 5.13 – Gel radiogram recorded after electro-elution of solutions containing 150 ng *pUC57*-CRE either alone (lane 1), in the presence of *AatII* enzyme (lane 2), in the presence of 100 μ M MPA, 25 μ M CuCl₂, buffer, and increasing amounts of **bipy(GCN4bd2)₂** (lanes 3 to 8), or in the presence of 100 μ M MPA, buffer, and 5 μ M **bipy(GCN4bd2)₂** (lane 9). Concentration of **bipy(GCN4bd2)₂** was 0, 1, 2, 3, 4 and 5 μ M (lane 3, 4, 5, 6, 7 and 8 respectively), in 20 mM phosphate buffer pH 7.4 and 20 mM NaCl.

The short linear fragment observed upon co-incubation of *pUC57*-CRE with *AatII* is consistent with the insertion of the CRE duplex in the *Eco*-RV site. The gel indicates that Cu(II) promotes DNA cleavage, but that only in the presence of **bipy(GCN4bd2)₂** are double-strand breaks observed.

In order to investigate the specificity of cleavage attributed to the Cu(II):**bipy(GCN4bd2)₂** complex, the linear fragments obtained were post-treated with a restriction enzyme in order to establish the initial cleavage site based on the size of the different fragments,[41,67] using *NdeI* enzyme (see Figure 5.14A). Supercoiled *pUC57*-CRE plasmid was incubated at 310 K with: (1) 5 μ M **bipy(GCN4bd2)₂**, 10 μ M CuCl₂, 100 μ M MPA for 2 hours followed by (2) *NdeI* enzyme for 0.5 hours. A 1 kilobase DNA ladder, and the *pUC57*-CRE solutions prior to and post-treatment with *NdeI* were loaded on to the gel and eluted. The radiogram confirms that incubation with Cu(II):**bipy(GCN4bd2)₂** results in the formation of circular and linearised fragments of ~2,900 bp (see Figure 5.14B, compare lanes 1 and 2). However, only the ~2,900 bp linear fragment is observed upon post-treatment with *NdeI* (see Figure 5.14B, compare lane 1 and 3).

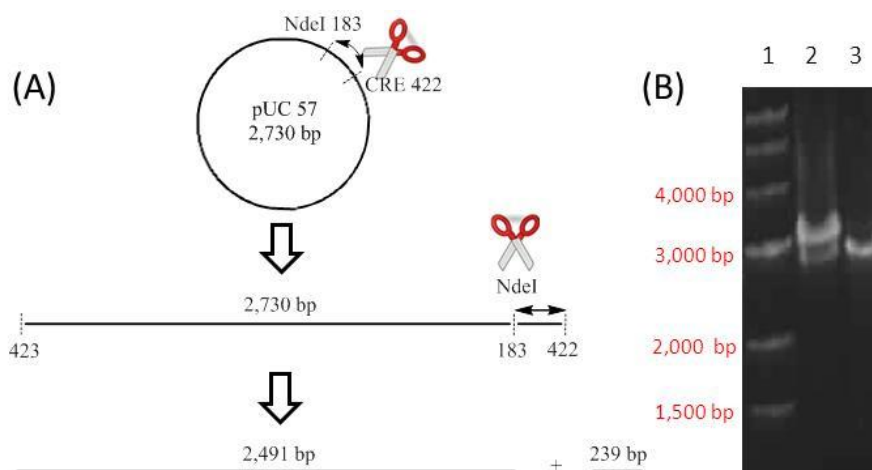


Figure 5.14 – (A) Scheme representing the strategy used to establish *pUC57*-CRE cleavage specificity towards CRE sites using *NdeI* enzyme. Following post-treatment with *NdeI*, the 2,730 bp linear fragment would be converted to 2,491 and 239 bp fragments. (B) Radiogram for 150 ng of *pUC57*-CRE incubated with 100 μ M MPA, 10 μ M CuCl₂, 5 μ M **bipy(GCN4bd2)₂**, 20 mM phosphate pH 7.4 and 20 mM NaCl at 310 K for 2 hours: 1 kb DNA ladder (lane 1), *pUC57*-CRE prior to (lane 2) and post-treatment with *NdeI* (lane 3).

Successive treatment of *pUC57*-CRE with Cu(II):**bipy**(GCN4bd2)₂ and *NdeI* did not result in the appearance of a band corresponding to the ~2,500 bp fragment, suggesting that DNA cleavage by Cu(II):**bipy**(GCN4bd2)₂ occurring under these conditions was not specific towards the introduced CRE site. This constitutes a preliminary negative result, and no further cleavage conditions have been tested for Cu(II):**bipy**(GCN4bd2)₂. As an example, the sequence-specific cleavage reported by Giovannangeli *et al.* involved incubation of plasmid at 310 K for 17 hours in the presence of 3 mM MPA. However, the supercoiled *pUC57*-CRE was quantitatively converted to the linear form when exposed to similar conditions (15 hours at 310 K, 3 mM MPA).

5.3 – Summary

Preliminary fluorescence studies on polypyridine peptide conjugates indicate that unlike **bipy** conjugates, metal free **terpy** conjugates display an intense emission band upon excitation at 298 nm. Addition of Cu(II) to **terpy**-GS₂ or **terpy**(GCN4bd2)₂ results in fluorescence quenching, which is largely independent of the sequence of the DNA present. Any differences were attributed to different metal affinities previously estimated, rather than second-sphere coordination. In an extension of this study, the model conjugate [Ru^{II}(**bipy**)₂(cys₂**bipy**)]²⁺ was prepared, the optical properties of which are shown to be consistent with [Ru^{II}(**bipy**)₃]²⁺ type complexes. However, following the same synthetic approach, the conjugate [Ru^{II}(**bipy**)₂(**bipy**(GCN4bd2)₂)]²⁺ could not be obtained.

The ability of the Cu(II):**bipy**(GCN4bd2)₂ complex to cleave supercoiled DNA at an inserted CRE recognition site was investigated by horizontal electrophoresis. Our results indicate that in the presence of excess Cu(II) and reducing agent the conjugate is able to induce double-strand breaks which are not directed specifically towards the recognition site, as shown by post-treatment and fragment size analysis. However, the nuclease ability of **terpy**(GCN4bd2)₂ was not investigated, and more studies are required in order to conclude

on the potential of polypyridine:Cu(II) GCN4**bd** peptide conjugates to find applications as sequence specific oxidative nuclease agents.

5.4 – Experimental

5.4.1 – Equipment and reagents

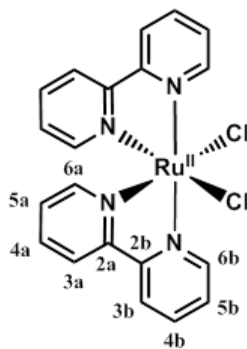
Mono- and dihydrogen potassium phosphate salts, sodium chloride, lithium chloride, ethanol (EtOH), agarose, and ultrapure or DNA grade water were all obtained from Fisher Scientific. Copper chloride (CuCl₂), glycerol, 5,5'-dimethyl-2,2'-bipyridine, 2,2'-bipyridine, silver nitrate, and 3-mercaptopropionic acid (MPA) were obtained from Sigma-Aldrich. Deuterated solvents (DMSO-d₆ and D₂O) were obtained from Cambridge Isotope Laboratory Inc. Dimethylformamide (DMF) was obtained from AGTC Bioproducts and acetone from VWR. RuCl₃ hydrate was obtained from Alfa Aesar and ammonium hexafluorophosphate from Strem Chemicals. *NdeI* and *AatII* enzyme, 10x restriction buffer, and 1 kb DNA ladder were obtained from New England Biolabs. The *pUC57* vector > 95 % supercoiled content (SC grade) with 5'-ACGAGATGACGTCATCTCCA-3' oligonucleotide and its complementary strand inserted by blunt end ligase at *EcoRV* cloning site, was obtained as a 1 µg µL⁻¹ solution in distilled deionised water from Genscript and used directly.

HRES and ES-TOF MS were recorded on a Microwaters LCT TOF spectrometer equipped with a 3000 V capillary voltage, and a cone voltage of 35 V. UV-visible spectra were recorded in a 1 cm pathlength quartz cuvette at 298 K on a CARY50 spectrometer. Wavelengths are given in nm, and extinction coefficients in M⁻¹ cm⁻¹. Emission spectra were recorded on a Shimadzu RF-5301PC spectrophotometer with excitation and emission slit width both set at 5 nm, and using 1 mL quartz cuvette with 1 cm pathlength. Wavelengths are given in nm and intensity in arbitrary unit (a. u.). All ¹H NMR spectra were collected on AVIII300 (300 MHz ¹H, T = 293 K), and ¹³C NMR on AVIII400 (400 MHz ¹H, 100 MHz ¹³C, T = 298 K) spectrometer, both equipped with a 5 mm probe. Solvent residual signal, DMSO quintet (δ = 2.50 ppm), was used as an internal reference. As an exception, dioxane

was used as an internal standard ($\delta = 3.75$ ppm) for all experiments recorded in D_2O . For pD measurement, pH paper was used. Chemical shifts (δ) are given in parts per million (ppm) to higher frequency compare to the methyl signal of either trimethylsilane (organic solvent), or the sodium salt of 3-(trimethylsilyl)propanesulfonicacid (D_2O) at 0 ppm.[69] Data were processed using Bruker Topspin version 2.1 (300 and 400 MHz) or 1.3 (500 MHz) and Mestrenova Lite version 5.2.5. For the electrophoresis experiments, gel were visualised using an Uvitec Cambridge imager.

5.4.2 – Synthetic procedure and characterisation

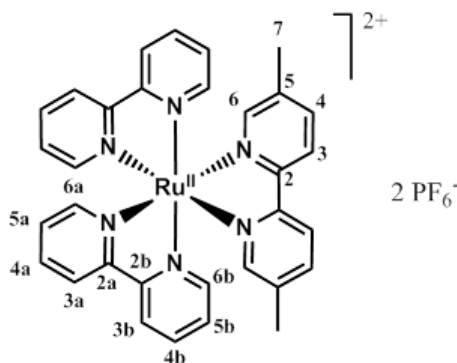
Synthesis of $Ru^{II}(bipy)_2Cl_2$: *cis-bis-(2,2'-bipyridine)dichlororuthenium(II) dihydrate* [70,71]



In a round-bottom flask, trichlororuthenium(III) hydrate (79.3 mg, 298 μ mol), lithium chloride (84.8 mg, 2 mmol), 2,2'-bipyridine (93.6 mg, 0.6 mmol) and DMF (25 mL) were refluxed for 12 hours under a N_2 atmosphere. After the suspension cooled to room temperature, acetone (200 mL) was added and it was stored at 253 K overnight, filtered *in vacuo* using a glass sintered funnel, washed with water (3x5 mL) and diethylether (3x5 mL) and dried *in vacuo*, affording a black solid (72 mg, 46 %). 1H NMR (300 MHz, $DMSO-d_6$): 9.97 (d, $^3J_{H6b-H5b} = 7.5$ Hz, 2H, H^{6b}), 8.64 (d, $^3J_{H3b-H4b} = 7.5$ Hz, 2H, H^{3b}), 8.49 (d, $^3J_{H3a-H4a} = 7.5$ Hz, 2H, H^{3a}), 8.08 (td, $^3J_{H4b-H5b} = ^3J_{H4b-H3b} = 7.5$ Hz, $^4J_{H4b-H6b} = 1.4$ Hz, 2H, H^{4b}), 7.78 (td, $^3J = 6.6$ Hz, $^4J_{H5b-H3b} = 1.1$ Hz, 2H, H^{5b}), 7.69 (td, $^3J = 7.8$ Hz, $^4J_{H4a-H6a} = 1.2$ Hz, 2H, H^{4a}), 7.51 (d, $^3J_{H6a-H5a} = 5.2$ Hz, 2H, H^{6a}), 7.11 (td, $^3J_{H5a-H6a} = ^3J_{H5a-H4a} = 6.6$ Hz, $^4J_{H5a-H3a} = 1.1$ Hz,

^2H , $\text{H}^{5\text{a}}$); ^{13}C NMR ^1H decoupled (100 MHz, DMSO-d_6): 160.3 ($\text{C}^{2\text{a}}$), 158.1 ($\text{C}^{2\text{b}}$), 153.1 ($\text{C}^{6\text{b}}$), 152.1 ($\text{C}^{6\text{a}}$), 134.6 ($\text{C}^{4\text{b}}$), 133.5 ($\text{C}^{4\text{a}}$), 125.3 ($\text{C}^{5\text{b}}$), 125.2 ($\text{C}^{5\text{a}}$), 122.7 ($\text{C}^{3\text{a}}$), 122.5 ($\text{C}^{3\text{b}}$); UV-vis (EtOH/100 mM phosphate pH 7.4 (2/3)): 517 (3,238), 365 (3,253), 296 (23,238), 250 (7,451); Emission (EtOH/100 mM phosphate pH 7.4 (2/3), $\lambda_{\text{exc}} = 456$ nm): 533 (8.8).

Synthesis of $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{Me}_2\text{bipy})](\text{PF}_6)_2$: cis-bis-(2,2'-bipyridine)(5,5'-dimethyl-2,2'-bipyridine)ruthenium(II) hexafluorophosphate

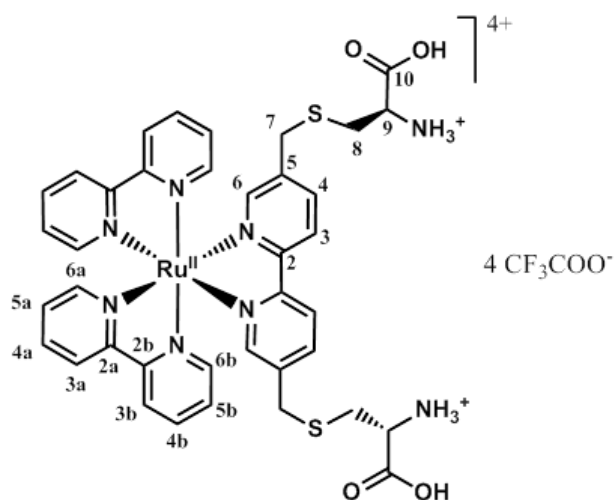


In a round-bottom flask, *cis*-bis-(2,2'-bipyridine)dichlororuthenium(II) dihydrate (15 mg, 28.9 μmol) and 5,5'-dimethyl-2,2'-bipyridine (5.3 mg, 28.9 μmol) were dissolved in EtOH (5 mL) and refluxed for 4 hours, after which the solution was dried *in vacuo*. Upon addition of water (2 mL) and a saturated solution of NH_4PF_6 (5 mL), an orange solid precipitated and was filtered out, washed with diethylether (2x5 mL), collected and dried *in vacuo*, yielding the compound together with excess NH_4PF_6 salts (50.3 mg). ^1H NMR (300 MHz, DMSO-d_6): 8.83 (d, $^3J_{\text{H}^{3\text{a/b}}-\text{H}^{4\text{a/b}}} = 8.2$ Hz, 2H, $\text{H}^{3\text{a/b}}$), 8.82 (d, $^3J_{\text{H}^{3\text{a/b}}-\text{H}^{4\text{a/b}}} = 8.1$ Hz, 2H, $\text{H}^{3\text{a/b}}$), 8.69 (d, $^3J_{\text{H}^3-\text{H}^4} = 8.4$ Hz, 2H, H^3), 8.20-8.12 (m, 4H, $\text{H}^{4\text{a+b}}$), 7.99 (d, $^3J_{\text{H}^4-\text{H}^3} = 8.4$ Hz, 2H, H^4), 7.71 (d, $^3J_{\text{H}^{6\text{a/b}}-\text{H}^{5\text{a/b}}} = 6.1$ Hz, 2H, $\text{H}^{6\text{a/b}}$), 7.68 (d, $^3J_{\text{H}^{6\text{a/b}}-\text{H}^{5\text{a/b}}} = 6.2$ Hz, 2H, $\text{H}^{6\text{a/b}}$), 7.55 (t, $^3J_{\text{H}^{5\text{a/b}}-\text{H}^{6\text{a/b}}} = ^3J_{\text{H}^{5\text{a/b}}-\text{H}^{4\text{a/b}}} = 6.9$ Hz, 2H, $\text{H}^{5\text{a/b}}$), 7.50 (d, $^3J_{\text{H}^{5\text{a/b}}-\text{H}^{6\text{a/b}}} = ^3J_{\text{H}^{5\text{a/b}}-\text{H}^{4\text{a/b}}} = 6.4$ Hz, 2H, $\text{H}^{5\text{a/b}}$), 7.47 (s, 2H, H^6); ^{13}C NMR ^1H decoupled (100 MHz, DMSO-d_6): 156.6 ($\text{C}^{2\text{a/b}}$), 156.6 ($\text{C}^{2\text{a/b}}$), 154.1 (C^2), 151.2 ($\text{C}^{6\text{a/b}}$), 151.1 ($\text{C}^{6\text{a/b}}$), 150.6 (C^6), 138.6 (C^4), 137.8 ($\text{C}^{4\text{a+b}}$), 127.8 ($\text{C}^{5\text{a+b}}$), 124.5 ($\text{C}^{3\text{a/b}}$), 124.4 ($\text{C}^{3\text{a/b}}$), 123.5 (C^3), 18.1 (C^7), C^5 was not observed; HRES-

TOF (CH_2Cl_2) calculated exact mass for $\text{C}_{32}\text{H}_{28}\text{N}_6\text{F}_6\text{P}^{102}\text{Ru}$: 743.1061; measured: 743.1064.

ES-TOF: $m/z = 743.5$ ($[\text{M}+\text{PF}_6]^+$, 33 %), 375.1 (47 %), 299.2 ($[\text{M}]^{2+}$, 100 %), 219.1 (46 %).

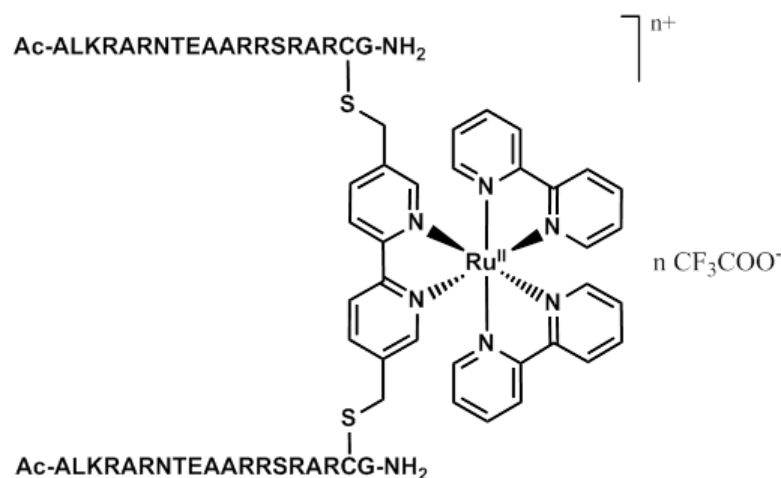
Synthesis of $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})](\text{CF}_3\text{COO})_4$: bis-(2,2'-bipyridine)[5,5'-bis(methyl-S-cysteiny)-2,2'-bipyridine]ruthenium(II) tetra-(trifluoroacetate)



In a round-bottom flask, were added *cis*-bis-(2,2'-bipyridine)dichlororuthenium(II) dihydrate (3.0 mg, 5.8 μmol), a 5,5'-bis(methyl-S-cysteiny)-2,2'-bipyridine aqueous solution (18.9 mM, 305 μL , 5.8 μmol), EtOH (2 mL), 100 mM phosphate buffer pH 7.4 (0.3 mL), and water (2.4 mL). The suspension was refluxed for 3.5 hours, and the solvent was removed *in vacuo* to yield a brown solid which was purified by semi-preparative RP-HPLC (C18 Phenomenex, monitoring absorbance at 210 and 450 nm) using a 0 to 22.5 % gradient methanol in water (containing 0.05 % TFA) over 36 min. The solvent was evaporated *in vacuo* to yield pure bis-(2,2'-bipyridine)[5,5'-bis(methyl-S-cysteiny)-2,2'-bipyridine]ruthenium(II), as a yellow solid (2.5 mg, 52 %). ^1H NMR (300 MHz, D_2O , pD~1): 8.55 (d, $^3J_{\text{H}3\text{a/b}-\text{H}4\text{a/b}} = 8.1$ Hz, 4H, $\text{H}^{3\text{a+b}}$), 8.48 (d, $^3J_{\text{H}3-\text{H}4} = 8.4$ Hz, 2H, H^3), 8.05 (m, 6H, $\text{H}^{4+4\text{a+b}}$), 7.85 (m, 4H, $\text{H}^{6\text{a+b}}$), 7.69 (d, $^5J_{\text{H}6-\text{H}4} = 1.6$ Hz, 1H, H^6), 7.68 (d, $^5J_{\text{H}6-\text{H}4} = 1.6$ Hz, 1H, H^6), 7.37 (m, 4H, $\text{H}^{5\text{a+b}}$), 3.88 (m, 2H, H^9), 3.66 (m, 4H, H^7), 2.69 (m, 4H, H^8); ES-TOF (MeOH): calculated mass for $\text{C}_{38}\text{H}_{38}\text{N}_8\text{O}_4^{102}\text{RuS}_2$: 836.15; measured: $m/z = 871.5$ ($[\text{M}+\text{Cl}]^+$, 28 %), 835.5 ($[\text{M}+\text{H}]^+$, 40 %), 481.2 ($[\text{Ru}(\text{bipy})_2\text{Cl}(\text{MeOH})]^+$, 42 %), 449.2 ($[\text{Ru}(\text{bipy})_2\text{Cl}]^+$,

30 %), 423.3 ([cys₂bipy+H]⁺, 22 %), 418.2 ([M]²⁺, 100 %); UV-vis (EtOH/100 mM phosphate pH 7.4 (2/3)): 429 (3,427), 289 (41,829), 244 (24,776); Emission (EtOH/100 mM phosphate pH 7.4 (2/3), $\lambda_{\text{exc}} = 456$ nm): 618 (198.6), 526 (97.4).

Attempts to synthesise [Ru^{II}(bipy)₂(bipy(GCN4bd2)₂)](CF₃COO)_n: bis-(2,2'-bipyridine)[5,5'-bis(methyl-S-GCN4bd2)-2,2'-bipyridine]ruthenium(II) trifluoroacetate: 2 different approaches were followed (A) and (B).



(A) A solution of *cis*-bis-(2,2'-bipyridine)dichlororuthenium(II) dihydrate (8.66 mM, 20 μ L, 0.176 μ mol) in ethanol was added to a round-bottom flask containing an aqueous **bipy(GCN4bd2)₂** solution (130 μ M, 1.35 mL, 0.176 μ mol), ethanol (980 μ L) and 100 mM phosphate buffer pH 7.4 (150 μ L). The resulting suspension was degassed with N_{2(g)} for 10 minutes and then refluxed for 4 days. In order to monitor the reaction, a small fraction was analysed by analytical C18 RP-HPLC (water +0.05 % TFA/ methanol gradients) after 5 hours, 2 days, and 3 days refluxing and the different fractions were collected, the mass spectrometry analysis of which did not show peaks at expected values.

(B) In a round-bottom flask, *cis*-bis-(2,2'-bipyridine)dichlororuthenium(II) dihydrate (1.4 mg, 2.7 μ mol), AgNO₃ (0.9 mg, 5.4 μ mol) and acetone (1 mL) were stirred at room temperature in the dark for 13 hours, after which the solution turned green and a dark precipitate formed. The suspension was transferred to an eppendorf and centrifuged, and the

green supernatant was collected. UV-vis (water): 645 (1,420), 435 (1,573), 285 (76,800); (adapted from reference 72). A small fraction of the supernatant (22 μL) was added to a round-bottom flask containing an aqueous **bipy(GCN4bd2)₂** solution (130 μM , 0.5 mL, 65 nmol) and acetone (478 μL). The suspension refluxed for 3 days and, again, analytical C18 RP-HPLC and mass spectrometry did not indicate the formation of the desired product.

5.4.3 – Analytical procedures

5.4.3.1 – Polypyridine ligand fluorescence (related to section 5.2.1)

Prior to recording the fluorescence profile of model compounds, emission spectra of 20 mM phosphate buffer pH 7.4 solutions (1 mL) were acquired ($\lambda_{\text{exc.}} = 298 \text{ nm}$; slits width excitation and emission = 5 nm). Aliquots of model compound stock solutions (either **bipy-GS₂** (6.2 mM, 3.2 μL), or **terpy-GS₂** (7.1 mM, 2.8 μL)) were then added to the cuvette such that their final concentrations were equal to 20 μM , and new spectra were recorded. Aliquots of a 1 mM CuCl_2 stock solution were then added, and again new spectra recorded. Spectra recorded for the buffer were subtracted from spectra of model compounds recorded in the absence or presence of CuCl_2 . Fluorescence spectra of **terpy-(GCN4bd2)₂** recorded in the absence or presence of DNA were recorded following a similar methodology, however, the blank solution now contains 10 μM duplex DNA (either CRE or NON) if present (50 μM , 200 μL) and 10 mM phosphate buffer pH 7.4, furthermore, only 10 μM **terpy-(GCN4bd2)₂** (600 μM , 16.7 μL) were added for each measurement.

5.4.3.2 – Nuclease activity of **bipy-(GCN4bd2)₂** (related to section 5.2.2)

For the DNA cleavage assay, binding solution ($V_{\text{total}} = 5 \text{ } \mu\text{L}$) were prepared by successive addition of *pUC57*-CRE (100 ng μL^{-1} , 150 ng, 1.5 μL), and 5 \times buffer or buffer- Cu(II) mixture (1 μL) containing 100 μM phosphate pH 7.4, 100 μM NaCl, and 125 μM CuCl_2 (if present). Gradient of **bipy-(GCN4bd2)₂** were prepared from 10 and 25 μM stock

solutions and the volume adjusted up to 4.5 μL with ultrapure water. Cleavage were initiated by addition of MPA (1 mM, 0.5 nmol, 0.5 μL), and solutions were vortexed, centrifuged and incubated at 310 K for 2 hours. The DNA blank solution contained only *pUC57*-CRE (100 ng μL^{-1} , 150 ng, 1.5 μL), and water (3.5 μL) and was incubated at 277 K. *AatII* control contained *pUC57*-CRE (100 ng μL^{-1} , 150 ng, 1.5 μL), 10x restriction buffer (0.5 μL), *AatII* restriction enzyme (10 unit μL^{-1} , 1 unit, 0.1 μL), water (2.9 μL), and was incubated at 310 K for 1 hour. All eppendorf, tips, and solutions used (except for MPA and **bipy(GCN4bd2)₂** stock and restriction solution) were autoclaved prior to use.

Following incubation, a dye solution (1 μL) prepared in 5x TBE loading buffer and containing 25% glycerol, was added to each eppendorf, and the solutions were stored at 277 K for 10 minutes. 5 μL of each solution was subsequently loaded on a 0.8 % agarose gel soaked in 1x TBE, and electrophoresed horizontally at 7 V cm^{-1} for 3 hours. The gel was then stained in 0.5 $\mu\text{g mL}^{-1}$ aqueous ethidium bromide solution for 45 minutes, destained in deionised water for 15 min, and visualised on a UV transilluminator.

The gel used to confirm the cleavage specificity was run with a similar method. However, the volumes used were three times higher ($V_{\text{total}} = 15 \mu\text{L}$), and CuCl_2 was not present in the buffer but added separately (150 μM , 0.15 nmol, 1 μL). Following incubation at 310 K for 2 hours, the solution was divided equally over 3 eppendorfs. To one of these, 10x restriction buffer (0.5 μL) and *NdeI* restriction enzyme (10 unit μL^{-1} , 1 unit, 0.1 μL) were added, and the solution incubated at 310 K for a further 30 minutes, affording the post-treated solution. Remaining steps were carried out as previously described, except for volume of 1 kb ladder loaded (1 μL), and electro-elution performed at 90 V for 2.3 hours.

5.5 – References

[1] Lerman L. S., *J. Mol. Biol.*, **1961**, 3, 18-30.

- [2] Jenette K. W., Lippard S. J., Vassiliades G. A., Bauer W. R., *Proc. Natl. Acad. Sci. USA*, **1974**, 71, 3839-3843.
- [3] Zeglis B. M., Pierre V. C., Barton J. K., *Chem. Commun.*, **2007**, 4565-4579.
- [4] Sitlani A., Dupureur C. M., Barton J. K., *J. Am. Chem. Soc.*, **1993**, 115, 12589-12590.
- [5] Erkkila K. E., Odom D. T., Barton J. K., *Chem. Rev.*, **1999**, 99, 2777-2795.
- [6] Terbrueggen R. H., Barton J. K., *Biochemistry*, **1995**, 34, 8227-8234.
- [7] Sardesai N. Y., Barton J. K., *J. Biol. Inorg. Chem.*, **1997**, 2, 762-771.
- [8] Sardesai N. Y., Zimmermann K., Barton J. K., *J. Am. Chem. Soc.*, **1994**, 116, 7502-7508.
- [9] Lasey R. C., Banerji S. S., Ogawa M. Y., *Inorg. Chim. Acta*, **2000**, 300-302, 822-828.
- [10] Elias B., Kirsch-DeMesmaeker A., *Coord. Chem. Rev.*, **2006**, 250, 1627-1641.
- [11] Szacilowski K., Macyk W., Drzewiecka-Matuszek A., Brindell M., Stochel G., *Chem. Rev.*, **2005**, 105, 2647-2694.
- [12] Ji L., Zou X., Liu J., *Coord. Chem. Rev.*, **2005**, 105, 2647-2694.
- [13] Pazos E., Vázquez O., Mascareñas J. L., Vázquez M. E., *Chem. Soc. Rev.*, **2009**, 38, 3348-3359.
- [14] Penas C., Pazos E., Mascareñas J. L., Vázquez M. E., *J. Am. Chem. Soc.*, **2013**, 135, 3812-3814.
- [15] Brunner J., Barton J. K., *Biochemisry*, **2006**, 45, 12295-12302.
- [16] Gamba I., Salvadó I., Rama G., Bertazzon M., Sánchez M. I., Sánchez-Pedregal V. M., Brissos R. F., Gamez P., Mascareñas J. L., López M. V., Vázquez M. E., *Chem. Eur. J.*, **2013**, 19, 13369-13375.
- [17] Castellucci E., Angeloni L., Marconi G., Venuti E., Baraldi I., *J. Phys. Chem.*, **1990**, 94, 1740-1745.
- [18] Sarkar A., Chakravorti S., *J. Lumin.*, **1995**, 63, 143-148.
- [19] Henry M. S., Hoffman M. Z., *J. Phys. Chem.*, **1979**, 83, 618-625.
- [20] Kotlicka J., Grabowski Z. R., *J. Photochem.*, **1979**, 11, 413-418.

- [21] Albano G., Balzani V., Constable E. C., Maestri M., Smith D. R., *Inorg. Chim. Acta*, **1998**, 277, 225-231.
- [22] Huang Z., Du J., Zhang J., Yu X., Pu L., *Anal Methods*, **2012**, 4, 1909-1912.
- [23] Huang Z., Du J., Zhang J., Yu X., Pu L., *Chem. Commun.*, **2012**, 48, 3412-3414.
- [24] Umezawa H., Maeda K., Takeuchi T., Okami Y., *J. Antibiot. (Tokyo) Ser. A*, **1966**, 19, 200-209.
- [25] Burger R. M., *Chem. Rev.*, **1998**, 98, 1153-1169.
- [26] Sigman D. S., Graham D. R., D'Aurora V., Stern A. M., *J. Biol. Chem.*, **1979**, 254, 12269-12272.
- [27] Pogożelski W. K., Tullius T. D., *Chem. Rev.*, **1998**, 98, 1089-1107.
- [28] Jiang Q., Xiao N., Shi P., Zhu Y., Guo Z., *Coord. Chem Rev.*, **2007**, 251, 1951-1972.
- [29] Lamour E., Routier S., Bernier J. L., Catteau J.-P., Bailly C., Vezin H., *J. Am. Chem. Soc.*, **1999**, 121, 1862-1869.
- [30] Maheswari P. U., Roy S., den Dulk H., Barends S., van Wezel G., Kozlevčar B., Gamez P., Reedijk J., *J. Am. Chem. Soc.*, **2006**, 128, 710-711.
- [31] Arber W., *Annual Rev. Microbiol.*, **1965**, 19, 365-378.
- [32] Chin J., *Acc. Chem. Res.*, **1991**, 24, 145-152.
- [33] Williams N. H., Takasaki B., Wall M., Chin J., *Acc. Chem. Res.*, **1999**, 32, 485-493.
- [34] Suh J., *Acc. Chem. Res.*, **2003**, 36, 562-570.
- [35] Wang J.-T., Xia Q., Zheng X.-H., Chen H.-Y., Chao H., Mao Z.-W., Ji L.-N., *Dalton Trans.*, **2010**, 39, 2128-2136.
- [36] Simon P., Cannata F., Perrouault L., Halby L., Concordet J.-P., Boutorine A., Ryabinin V., Sinyakov A., Giovannangeli C., *Nucleic Acids Res.*, **2008**, 36, 3531-3538.
- [37] van der Steen S., de Hoog P., van der Schilden K., Gamez P., Pitié M., Kiss R., Reedijk J., *Chem. Commun.*, **2010**, 3568-3570.

- [38] An Y., Lin Y. Y., Wang H., Sun H. Z., Tong M. L., Ji L. N., Mao Z. W., *Dalton Trans.*, **2007**, 1250-1254.
- [39] Bazzicalupi C., Bencini A., Bianchi A., Danesi A., Giorgi C., Lodeiro C., Pina F., Santarelli S., Valtancoli B., *Chem. Comm.*, **2005**, 2630-2632.
- [40] Mancin F., Tecilla P., *New J. Chem.*, **2007**, 31, 800-817.
- [41] Chen C. B., Sigman D. S., *Science*, **1987**, 237, 1197-1201.
- [42] Mack D. P., Iverson B. L., Dervan P. B., *J. Am. Chem. Soc.*, **1988**, 110, 7572-7574.
- [43] Ebright R. H., Ebright Y. W., Pendergrast P. S., Gunasekera A., *Proc. Natl. Acad. Sci. USA*, **1990**, 87, 2882-2886.
- [44] Hu D., Crist M., Duan X., Quijcho F., Gimble F. S., *J. Biol. Chem.*, **2000**, 275, 2705-2712.
- [45] Azbun T. R., Stahura F. L., Mossing M. C., *Biochemistry*, **1997**, 36, 3680-3686.
- [46] Kavacic R., Welch J. T., Franklin S. J., *J. Am. Chem. Soc.*, **2003**, 125, 6656-6662.
- [47] Nomura A., Sugiura Y., *J. Am. Chem. Soc.*, **2004**, 126, 15374-15375.
- [48] Oakley M. G., Dervan P. B., *Science*, **1990**, 248, 847-850.
- [49] Kolarik Z., *Chem. Rev.*, **2008**, 108, 4208-4252.
- [50] Kirchhoff J. R., McMillin D. R., Marnot P. A., Sauvage J., *J. Am. Chem. Soc.*, **1985**, 107, 1138-1141.
- [51] Rau H. K., DeJonge N., Haehnel W., *Proc. Natl. Acad. Sci. USA*, **1998**, 95, 11526-11531.
- [52] Geren L., Hahm S., Durham B., Millet F., *Biochemistry*, **1991**, 30, 9450-9457.
- [53] Yang X., Janiak C., Heinze J., Drepper F., Mayer P., Piotrowski H., Klüfers P., *Inorg. Chim. Acta*, **2001**, 318, 103-116.
- [54] Kise Jr K. J., Bowler B. E., *Inorg. Chem.*, **2002**, 41, 379-386.
- [55] Wiederholt K., McLaughlin L. W., *Nucleic Acids Res.*, **1999**, 27, 2487-2493.
- [56] Brateman P. S., Song J. L., Peacock R. D., *Inorg. Chem.*, **1992**, 31, 555-559.

- [57] Juris A., Balzani V., *Coord. Chem. Rev.*, **1988**, 84, 85-277.
- [58] Fodor L., Lendvay G., Horváth A., *J. Phys. Chem. A*, **2007**, 111, 12891-12900.
- [59] Gleria M., Minto F., Beggiato G., Bortolus P., *JCS Chem. Comm.*, **1978**, 285a-285a.
- [60] Yamagishi A., Naing K., Goto Y., Taniguchi M., Takahashi M., *J. Chem. Soc. Dalton Trans.*, **1994**, 2085-2089.
- [61] Mancin F., Tecilla P., Metal Complex-DNA Interactions, Chapter C-13: Artificial Restriction Agents: Hydrolytic Agents for DNA cleavage, John Wiley & Sons, Chichester, **2009**, ISBN 978-1-4051-7629-3.
- [62] Burrows C. J., Muller J. G., *Chem. Rev.*, **1998**, 98, 1109-1151.
- [63] Hermann T., Heumann H., *RNA*, **1995**, 1, 1009-1017.
- [64] Reed C. J., Douglas K. T., *Biochem. J.*, **1991**, 275, 601-608.
- [65] www.astrochem.org/sci/Nucleobases.php, accessed on 29/08/2013.
- [66] Ellenberger T. E., Brandl C. J., Struhl K., Harrison S. C., *Cell*, **1992**, 71, 1223-1237.
- [67] Simon P., Cannata F., Perrouault L., Halby L., Concordet J., Boutorine A., Ryabinin V., Sinyakov A., Giovannangeli C., *Nucleic Acids Res.*, **2008**, 36, 3531-3538.
- [68] www.genscript.com/document/filecenter/document/1400_20060331011034.jpg, accessed on 05/09/2013.
- [69] Gottlieb H. E., Kotlyar V., Nudelman A., *J. Org. Chem.*, **1997**, 62, 7512-7515.
- [70] Sullivan B. P., Salmon D. J., Meyer T. J., *Inorg. Chem.*, **1978**, 17, 3334-3341.
- [71] McCusker C. E., McCusker J. K., *Inorg. Chem.*, **2011**, 50, 1656-1669.
- [72] Marin V. N., Novel Functional Materials based on Ruthenium(II) and Iridium(II) Polypyridyl Complexes, PhD dissertation, Technische Universiteit Eindhoven, **2006**, p101.

CHAPTER VI: CONCLUSIONS AND FUTURE WORK

6.1 - Summary

This thesis reports the design and study of polypyridine units as peptide dimerisation domains suitable for artificial regulation of biomolecular activity based on conformational transitions. The regulation potential of an allosteric ineffective **bipy** linker was compared with that of an allosteric-effective **terpy** (based on the respective positions of their substituents).

Chapter 2 describes the synthesis and study of low molecular weight artificial regulation sites based on polypyridine conjugates, for which conformational and coordination studies were attempted. The preparation of small dimeric conjugates based on L-glutathione moieties were more straightforward than their smaller counterparts based on L-cysteine moieties consistent with the reaction being driven by the solubility of the resulting product in water. The coordination studies strongly support the hypothesis that both **bipy** and **terpy** linkers bind Cu(II) and Zn(II) mainly in a 1:1 ratio at low concentrations (millimolar range) regardless of the peptide substituents, however CD studies suggests the successive formation of different complexes upon addition of Cu(II) to model conjugates (observed at 100-350 μM). The conformation of metal bound and unbound states could not be studied in detail due to the complexity of the NMR spectra, due to quadrupolar relaxation, second-order effects and dynamic effects. However, Cu(II) and Zn(II) addition monitored by UV spectroscopy is consistent with the expected *trans-to-cis* conformational transition proposed, and hence justify their interest as regulation sites. This represents one of the most complete studies of the conformation of **bipy**, and more importantly **terpy** conjugates, in aqueous solution.

Chapter 3 describes the preparation of larger peptides based on the GCN4 transcription factor, their dimerisation via the **bipy** and **terpy** linkers, and studies of Cu(II) and Zn(II) coordination to the resulting conjugates. Two peptides with N- and C- termini capped incorporating either 17 or 25 residues from GCN4 (both including the entire **bd**) and with different flexibility in the C-terminal regions were designed, prepared and dimerised with the

polypyridine linkers. The conjugates of the longer peptides were not obtained in satisfactory purity, and metal titrations monitored by UV spectroscopy display behaviour different from the model conjugates (only a moderate shift of the bands). In contrast, the dimer conjugates made of the shorter peptides could be obtained in high purity. The addition of Cu(II) and Zn(II) results in coordination similar to the models, however, metal addition monitored by CD did not result in induced signals in contrast to the models. Differences could be due to capping of the N- and C- termini which might otherwise contribute to the coordination sphere.

Chapter 4 describes the DNA binding studies of the designed GCN4**bd** polypyridine conjugates. The dimer conjugates based on the long peptides display both increased helicity and bandshifts in the presence of DNA bearing a target site of native GCN4 (CRE), moreover, the estimated peptide-DNA dissociation constant (~10 nM) is consistent with previous reports. However, the addition of metal ions did not significantly affect DNA binding probably due to a high degree of flexibility in the dimerisation region. In contrast, the dimer conjugates based on the short peptides did not display any bandshift, however the folding into helices in the presence of CRE DNA (associated with DNA binding) was generally improved upon metal ion addition. Interestingly, the addition of either Cu(II) or Zn(II) increased the helicity of **bipy**(GCN4**bd2**)₂ and **terpy**(GCN4**bd2**)₂ conjugates to opposite extents only in the presence of DNA with the CRE site. However, only the helicity of the **terpy**(GCN4**bd2**)₂ conjugate was affected by Cu(II) or Zn(II) addition in the presence of DNA bearing the AP1 site, which is also targeted by native GCN4 but lacks a central **bp** compared to CRE. This result is consistent with both our design and previous reports, which indicate that the sequence specificity of GCN4**bd** dimers is affected by the bulkiness of their linkers.

Chapter 5 presents initial work directed towards studying the potential application of GCN4**bd** polypyridine conjugates as DNA sensors or nuclease agents. Our initial measurements which are consistent with previous photochemical reports on polypyridines, indicate that **terpy** displays a more intense fluorescence than **bipy** at room temperature. In all

cases, the polypyridine fluorescence decreases upon addition of Cu(II) and binding curves are consistent with an increase of Cu(II):**terpy**(GCN4**bd2**)₂ affinity in the presence of DNA bearing the CRE binding site. The preparation of conjugates in which GCN4**bd** peptides are dimerised with a Ru^{II}(**bipy**)₃²⁺ linker have not been successful so far, however, the model conjugate [Ru^{II}(**bipy**)₂(**cys**₂**bipy**)]²⁺ was prepared and studied, displaying intense phosphorescence. Cu(II) and Zn(II) complexes, including polypyridine complexes, have previously been shown to promote DNA strand cleavage, though different mechanisms and intermediates can be involved. Based on qualitative distance estimation using models of the peptide conjugates complexed with their DNA target sites, the DNA cleavage is more likely to occur via the production of diffusible reactive species, hence the choice of Cu which possesses several oxidation states. However, our initial results indicate that cleavage of a 2.7 kilobase supercoiled plasmid, attributed to the conjugate **bipy**(GCN4**bd2**)₂:Cu(I/II) (present in large excess), is not specifically directed towards the CRE target site.

Together, these results indicate that peptide dimerisation with disubstituted polypyridines allows for regulation of biomolecular activity based on the conformational transition resulting from metal coordination. Potential applications could include species using these to impact gene regulation either by occupying or denaturing (cleavage) the recognition site *in vivo*.

6.2 – Future work

6.2.1 – Preparation of new linkers and related models

The preparation of new polypyridine linkers suitable for incorporation within large peptide dimer scaffolds might be envisioned, in which case conjugates based on L-glutathione moieties could again be prepared as suitable model ligands. However, capping of the N- and C- termini might be necessary to prevent them contributing to metal coordination. Altering

the positioning of linkages on the pyridinyl rings might allow for the increase in the linker affinity for metal ions. **Terpy** linkers can also be prepared which would deactivate DNA binding upon metal coordination (see Figure 6.1A).[1] This design may allow for the preparation of conjugates for which DNA binding would trigger the release of a selected metal-ion which could be accompanied by an increase in fluorescence. Alternatively, the preparation of polypyridines containing additional substituents able to hinder the pyridine rotation, might stabilise particular conformations and facilitate their assignments using NMR spectroscopy (see Figure 6.1B and C).[2,3,4]

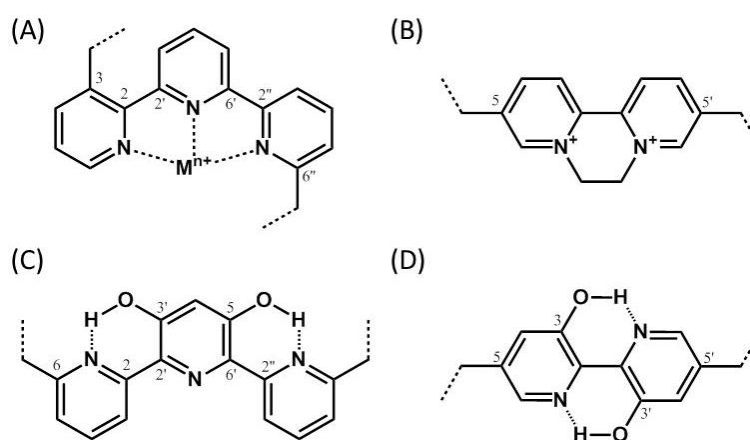


Figure 6.1 – Scheme representing (A) the 3,6''-dimethyl-2,2',6',2''-terpyridine linker the substituent realignment of which would promote the mixed *cis-trans* conformation and the release of a metal-ion (M^{n+}), (B) the annulated **bipy**: ethyl bridge 2,2'-bipyridium salt, or **terpy** (C) and **bipy** (D) bearing hydroxyl groups at positions 3 and/or 5, thus stabilising the *trans-trans*- and *trans*- conformations, respectively.

A logical extension of the coordination studies reported here, would involve investigating the interactions of model compounds with different metal ions, notably Fe(II) which preferentially forms 3**bipy**:1Fe(II) and 2**terpy**:1Fe(II) complexes, thus contrasting with the results obtained with Cu(II) and Zn(II). However, initial work involving the addition of Fe(II) salts to the **bipy-GS₂** conjugates resulted in formation of an unidentified species which is poorly soluble in water at physiological pH.

The first attempts to couple $Ru^{II}(\text{bipy})_2Cl_2$ or $[Ru^{II}(\text{bipy})_2(\text{acetone})_2]^{2+}$ complexes with the **bipy(GCN4bd2)₂** conjugate were unsuccessful. However, there exist alternative synthetic

strategies such as the complexation of a $\text{Ru}^{\text{II}}(\text{DMSO})_4\text{Cl}_2$ salt with **bipy**(GCN4**bd**)₂ followed by coordination with two additional bipy ligands.[5] Considering the low stability of the precursor complex $[\text{Ru}^{\text{II}}(\text{bipy})_2(\text{cys}_2\text{bipy})]^{2+}$ in D_2O , one may consider attaching **bipy** to the GCN4**bd** peptide using more robust bonds, such as amide linkages.[6,7] Aside from the synthetic aspects, it is possible to enhance the changes in photophysical properties experienced by the peptide conjugate upon binding to DNA. For example, an electron acceptor incorporated towards the N-terminus of a GCN4**bd** peptide conjugate could quench the luminescence of a C-terminal luminophore in the absence of a particular structural organisation. In contrast, the luminescence of the peptide conjugate might be restored upon binding to DNA and stabilisation of the α -helix, by similarity with a previous study involving a target antibody (see Figure 6.2).[8]

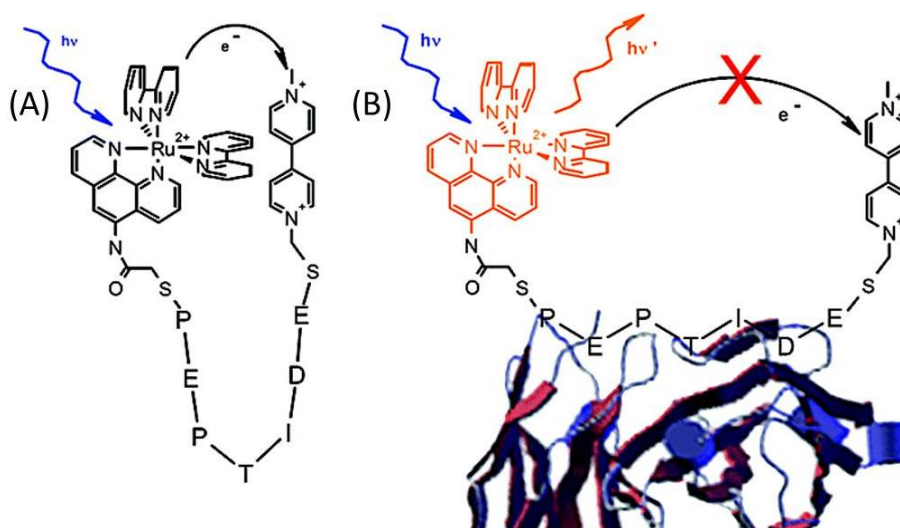


Figure 6.2 – Scheme illustrating a peptide beacon bearing a $\text{Ru}(\text{bipy})_2(\text{phen-IA})$ luminescent moiety and a 4,4-bipyridinium quencher either (A) unbound or (B) bound to the target antibody; adapted with permission from reference 8. Copyright 2007 American Chemical Society.

By extension, one can imagine incorporating artificial regulation sites based on molecular machines within peptide.[9,10] However, highly hydrophobic molecules can be challenging to incorporate within peptides.

6.2.2 – Design of new peptides and conjugates

From the DNA binding results, one can envision the design of a new peptide similar to our short GCN4^{bd} peptide, bearing more residues conserved from GCN4 in order to achieve tighter DNA binding and formation of a more stable peptide dimer conjugate-DNA complex. This would allow us to establish the important relationship between compromising strong DNA binding and the ability to perturb this on switching. Based on literature reports, this can be achieved upon retention of three additional C-terminal residues from GCN4 (approximately one α -helical turn).[11,12] Moreover, the majority of reports on artificial GCN4^{bd} dimers include a C-terminal dimerisation domain, whereas this could alternatively be inserted at the middle of the sequence and followed by a tail made of residues retained from GCN4 or with a high propensity to form helices and/or interact specifically.

As for model conjugates, dimerisation with conformationally restrained linkers and chelation with different metal ions would give more information on the role of the linker conformation in the interaction of the conjugate with the DNA target site. Moreover, DNA binding studies conducted on GCN4-polypyridine monomeric conjugates might allow to account for the peptide assembly with coordination bonds and metal promoting DNA binding, by formation of dimers assembled with coordination bonds. A possible extension of this work would involve testing the activity of peptide conjugates *in vivo* as metal-dependent inhibitors of AP1 transcription factors.[13]

The linkers developed in this thesis could be applied to other DNA binding peptides such as MyoD,[14] or different well-studied biomolecular recognition events such as the hydrolase activity of HN1 artificial enzyme.[15] The insertion of polypyridine in the peptide backbone could alternatively allow for the regulation of pore formation by cyclic peptides (such as gramicidin), or regulate the formation of α -hairpins or single-chain coiled coils important for biomolecular recognition.

6.3 - References

- [1] Zahn S., Reckien W., Kirchner B., Staats H., Matthey J., Lützen A., *Chem. Eur. J.*, **2009**, 15, 2572-2580.
- [2] Campá C., Camps J., Font J., de March P., *J. Org. Chem.*, **1987**, 52, 521-525.
- [3] Rurack K., Radeaglia R., *Eur. J. Inorg. Chem.*, **2000**, 2271-2282.
- [4] Costero A. M., Gil S., Parra M., Huguet N., Allouni Z., Lakhmiri R., Atlamsani A., *Eur. J. Org. Chem.*, **2008**, 1079-1084.
- [5] Bishop B. M., McCafferty D. G., Erickson B. W., *Tetrahedron*, **2000**, 4629-4638.
- [6] Kise K. J., Bowler B. E., *Inorg. Chem.*, **2002**, 41, 379-386.
- [7] Joshi T., Barbante G. J., Francis P. S., Hogan C. F., Bond A. M., Gasser G., Spiccia L., *Inorg. Chem.*, **2012**, 51, 3302-3315.
- [8] Oh K. J., Cash K. J., Hugenberg V., Plaxco K. W., *Bioconjugate Chem.*, **2007**, 18, 607-609.
- [9] Balzani V., Credi A., Raymo F. M., Stoddart J. F., *Angew. Chem. Int. Ed.*, **2000**, 39, 3348-3391.
- [10] Kay E. A., Leigh D. A., Zerbetto F., *Angew. Chem. Int. Ed.*, **2007**, 46, 72-191.
- [11] Ueno M., Murakami A., Makimo K., Morii T., *J. Am. Chem. Soc.*, **1993**, 115, 12575-12576.
- [12] Cammaño A. M., Vásquez M. E., Martínez-Costas J., Castedo L., Mascareñas J. L., *Angew. Chem. Int. Ed.*, **2000**, 39, 3104-3107.
- [13] Zhang F., Timm K., Arndt K., Woolley G. A., *Angew. Chem. Int. Ed.*, **2010**, 49, 3943-3946.
- [14] Morii T., Shimomura M., Morimoto M., Saito I., *J. Am. Chem. Soc.*, **1993**, 115, 1150-1151.
- [15] Lindgren N. J. V., Varedian M., Gogoll A., *Chem. Eur. J.*, **2009**, 15, 501-505.

APPENDICES: EXPERIMENTAL TECHNIQUES

A.1 – Spectroscopic techniques

A.1.1 – Electronic absorption or UV-visible spectroscopy

Spectroscopy is the study of the interactions between light and matter, and electronic absorption spectroscopy measures the fraction of light absorbed by a molecule, termed absorbance (A) (see Equation A.1). Molecules absorb photons of specific energies which correspond to the energetic difference between two energetic states of the electrons (from a stable to a less stable state). The absorption of light by molecules often corresponds to electromagnetic radiation from the far-ultraviolet, ultraviolet or visible energetic domains the energy of which is related to the wavelength (λ) (see Equation A.2). By correlation, it is possible to obtain information on the nature of the chemical bonds (type or electronic density) based on the absorbance wavelength. The most frequent transitions, which correspond to the movement of electrons from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO) (see Figure A.1A), result in the most intense peaks in the absorption spectrum. The absorbance depends on the transition probability, which follow orbital (l), spin (S) and magnetic (M_L) selection rules ($\Delta l = \pm 1$, $\Delta S = 0$, $\Delta M_L = 0, \pm 1$) and the concentration of molecules responsible for this transition (based on Equation A.3, defined as the Beer-Lambert law). [1,2]

$$A = \log \left(\frac{I}{I_0} \right) \quad (\text{A.1})$$

$$E = h \nu = \frac{h c}{\lambda} \quad (\text{A.2})$$

$$A = \epsilon c l \quad (\text{A.3})$$

A corresponds to the absorbance, I_0 and I the intensity of the radiation before and after passing through the sample respectively, E the energy, ν the frequency and λ the wavelength associated with the radiation, h is the Planck constant ($6.63 \times 10^{-34} \text{ m}^2 \text{ kg s}^{-1}$), c the speed of light *in vacuo* ($3.00 \times 10^8 \text{ m s}^{-1}$), ϵ is the molar absorption coefficient for a given compound, c its concentration, and l the pathlength of the material which is crossed by light.

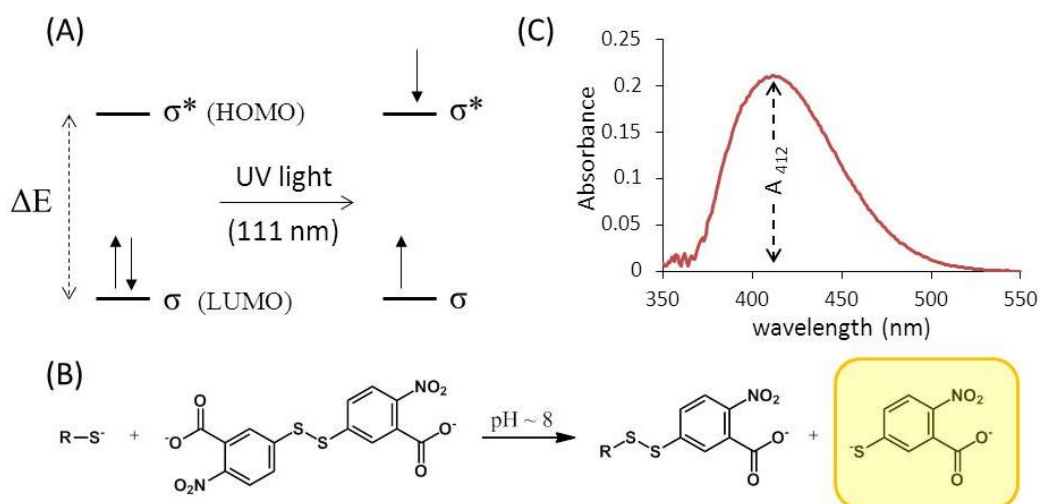


Figure A.1 – (A) Scheme illustrating the absorption of a photon with a wavelength of 111 nm resulting in an electronic transition from the HOMO to the LUMO orbital; adapted from reference 3. (B) Scheme illustrating the reaction between Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid)) and a thiol, resulting in the formation of the 2-nitro-5-thiobenzoate dye; (C) Example of the UV-visible spectra obtained for a L-glutathione solution following Ellman's test procedure. The thiol concentration can be estimated based on the absorbance at 412 nm and the known pathlength of the sample and absorption coefficient for the dye formed.

Proteins display an intense band in the range 185-200 nm attributed to amide $\pi \rightarrow \pi^*$ transitions.[4] Moreover, proteins containing dithiol linkages (cystine), or aromatic residues (tyrosine, phenylalanine and most notably tryptophan) display intense bands in the UV range (200-400 nm). For this reason, the absorbance at 280 nm is sometimes used to calculate the concentration of a protein for which the sequence is known and includes aromatic residues (based on a calculated absorption coefficient value [5]). This can alternatively be achieved using the Ellman's assay if the protein contains a known number of available cysteine residues (the only thiol containing proteinogenic amino-acid). In this procedure, free thiols react with 5,5'-dithiobis-(2-nitrobenzoic acid), generating chromophoric 2-nitro-5-thiobenzoate radicals (see Figure A.1B).[6] The absorbance properties of this radical species, are known ($\epsilon_{412 \text{ nm}} = 14,150 \text{ M}^{-1} \text{ cm}^{-1}$ [7]), and so can be readily quantified by UV-visible spectroscopy (see Figure A.1C), giving an approximation of the available thiol concentration.

A.1.2 – Circular dichroism spectroscopy

Dichroism is the property of certain materials to absorb light to different extents depending on its polarisation. Circular dichroism (CD) is the difference of absorptivity between left-handed and right-handed circularly polarised light, expressed as molar ellipticity (see equations A.4 and A.5), as a function of the wavelength. CD is a variant of electronic absorption spectroscopy which accounts for the chirality of molecular orbitals which can either be intrinsic or induced by bound molecules.[1,8] The CD spectra of chromophoric chiral molecules will display characteristic signals at wavelengths near an absorbance maximum (λ_{\max}) called cotton effects. A positive cotton effect produces a positive CD signal at a wavelength higher than the λ_{\max} , which crosses the wavelength axis at λ_{\max} , and which produces a negative signal at a wavelength lower than λ_{\max} . In contrast, the reverse is observed for a negative cotton effect (see Figure A.2A).

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \quad (\text{A.4})$$

$$[\Phi] = 3298.2 \times \Delta\varepsilon \quad (\text{A.5})$$

ε_L and ε_R correspond to the absorptivity of left and right-handed circularly polarized light, respectively, and $[\Phi]$ the molar ellipticity. The conversion value (3298.2) is obtained upon removal of the concentration dependence from the expression of ellipticity.

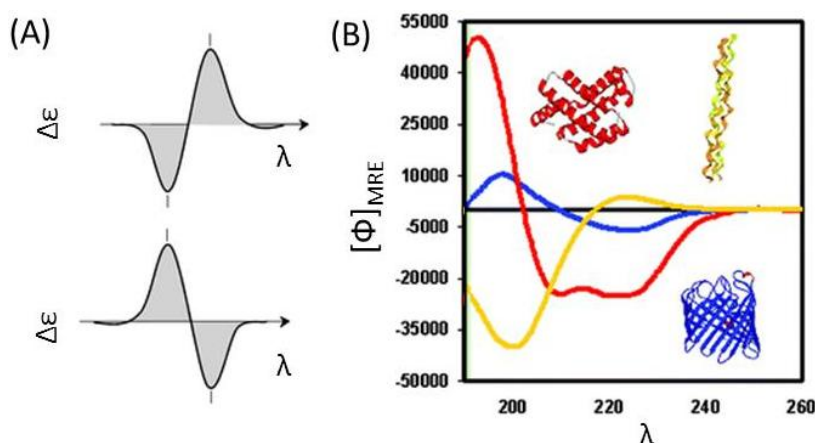


Figure A.2 – (A) CD signal defined as either a positive (top), or negative cotton effect (bottom). (B) CD spectra of proteins with either a high α -helix (red), β -sheet (blue), or polyproline content (yellow). ε represents the molar absorption coefficient, $\Delta\varepsilon$ the difference of absorptivity between left and right circularly polarised light, and $[\Phi]_{\text{MRE}}$ the molar ellipticity per residue in $\text{deg dmol}^{-1} \text{ cm}^2 \text{ res}^{-1}$ and λ the wavelength in nm; adapted with permission from reference 9 and 10. Copyright 2010 and 2006 Royal Society of Chemistry.

This technique is particularly useful for studying the structures of polymeric biomolecules (e.g. DNA and proteins). The region from 260 to 190 nm of a CD spectrum is the most relevant to the study of peptide structures as it contains both the $\pi \rightarrow \pi^*$ and the $n \rightarrow \pi^*$ transitions (215-230 nm), with the latter being electronically forbidden but magnetically allowed.[4] CD spectra displaying two negative minima around 208 and 222 nm and a positive maximum around 190 nm, are associated with a high degree of α -helical content, whereas typical β -sheet spectra display only one negative minima and one positive maximum, which are less intense and centred at a slightly higher wavelength compared to those for the α -helix. In contrast, unstructured (random coil) and polyproline containing peptide strands, both display intense negative minima around 200 nm, and small positive maxima around 210 nm (see Figure A.2B). A first approximation involves estimating the α -helix content of a peptide based on the $[\Phi]$ value at 222 nm, a region where the α -helical domains contribute most to the ellipticity.[11] On a side note, CD data tends to be recorded as ellipticities (Φ), which can readily be converted to the more conventional molar ellipticity ($[\Phi]$) and residual molar ellipticity ($[\Phi]_{MRE}$) following Equation A.6:

$$[\Phi]_{MRE} = \frac{[\Phi]}{n} = \frac{100 \Phi}{n c l} \quad (\text{A.6})$$

n corresponds to the number of residues in the protein, c its concentration, and l to the cell pathlength.

A.1.3 – Emission spectroscopy

Excited states are usually unstable and are therefore characterised by short lifetimes. Following electronic excitation by light, there exist several pathways (both intra- and intermolecular) by which the electron can return to its original ground-state, each associated with different probabilities or quantum yields (Φ). Electronic decay transitions are classified as either fluorescent, phosphorescent, internal conversion or intersystem, depending on

whether energy loss is converted into luminescent radiation or not, and if the transition occurs between two states of the same or different spin multiplicity ($\Delta S \neq 0$, spin-forbidden transition) (see Table A.1). Each pathway involves a number of such electronic transitions coupled with vibrational decay, thus fluorescence radiation occurs at a lower energy compared to the excitation wavelength (see Figure A.3). Spin-forbidden transitions are usually associated with long-lived or metastable excited states and slow rates of decay, which were the original definition for phosphorescence. Quenching commonly refers to intermolecular processes which result in a decrease in the fluorescence quantum yield (energy transfer, reactions, complex formation or collisions).[12]

Table A.1– Summary of the different types of electronic transition decays, and their associated quantum yields (Q.Y.)

Type	Definition	Q.Y.
phosphorescence	Radiative decay between two states with different spin multiplicities	Φ_P
fluorescence	Radiative decay between two states with the same spin multiplicity	Φ_F
intersystem crossing	Non-radiative decay between two states with different spin multiplicity	Φ_{ISC}
internal conversion	Non-radiative decay between two states with the same spin multiplicity	Φ_{IC}

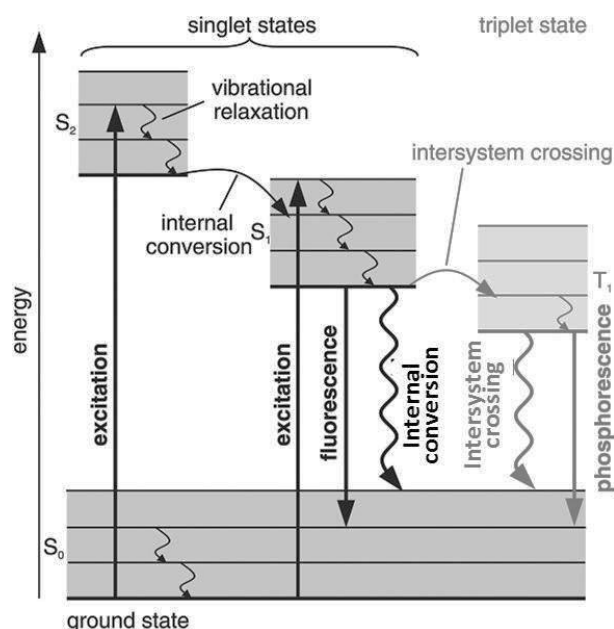


Figure A.3 – Jablonski diagram representing the electronic excitation, the different types of electronic transition decay, and the contribution from vibrational relaxation (not to scale). S corresponds to singlet, T triplet, and S_0 the ground state. The quantum yield for each pathway of electronic decay are defined as $\Phi_F + \Phi_P + \Phi_{IC} + \Phi_{ISC} = 1$; adapted with permission from reference 13. Copyright 2005 Springer.

A.1.4 – Nuclear Magnetic Resonance spectroscopy

Nuclear magnetic resonance (NMR) is a spectroscopic technique which involves measuring transitions between different spin levels of nuclei (I) in the presence of a magnetic field. Any nuclei possessing an even number of both protons and neutrons (such as $^{12}_6\text{C}$) have a nuclear spin $I = 0$, associated with a single energy state ($2I+1 = 1$), and are NMR-silent. In contrast, other nuclei possess more than one spin nuclei state, separated by energies relative to the magnetic field applied and the intrinsic magnetic properties of the nucleus (see Equation A.7). For example, a nucleus with $I = \frac{1}{2}$ (such as ^1H or ^{13}C) has two spin states which correspond to the parallel and antiparallel alignment of the spin with respect to the magnetic field orientation (see Figure A.4A). Since the electron cloud surrounding a nucleus shields it partially from the applied magnetic field B_0 , while being itself affected by the nuclei in its close environment, all non-equivalent nuclei will resonate at slightly different frequencies. The extent of shielding is denoted as the chemical shift (δ), which is independent of the strength of the applied field (see Equation A.8), and is representative of the chemical environment. The indirect transmission of orientation from a spin nucleus to a neighbouring nucleus through chemical bonds (relayed by the electronic spin), is called scalar coupling and results in splitting of the peaks (see Figure A.4B). The resulting number of peaks observed is called the spin multiplicity and is representative of the number of non-equivalent nuclei in the neighbouring environment. The shift between the resulting peaks corresponds to the coupling constant (J), the frequency of which is independent of the field applied, and is representative of the number or types of bonds separating the two nuclei. In ^1H NMR spectroscopy, coupling is usually observed if nuclei are separated by two (2J) to four (4J) chemical bonds. However, there are cases where the chemical shifts and coupling constants may not be extracted from the multiplet centers and line splitting, respectively (second-order effects). This occurs if the coupling between two nuclei is strong compared to the difference in their chemical shift ($\Delta\delta/J < 10$) and can be, in some cases, circumvented by recording the spectra using a higher

magnetic field. Since the Fourier transform (FT) converts time-dependent functions into a frequency spectra (see Figure A.4C), FT-NMR spectrometers usually irradiate samples with the entire energetic range of interest at the same time (pulsed) instead of scanning through (continuous mode), thus allowing for considerable time savings. FT-NMR spectrometers are usually referred to by the frequency at which protons would resonate if placed in the instrument.

$$\Delta E = \frac{h \gamma B_0}{2 \pi} \quad (\text{A.7})$$

$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}} \quad (\text{A.8})$$

ΔE is the energy difference between two spin nuclei states, h the Planck constant, γ the gyromagnetic ratio of the nuclei, and B_0 the magnetic field applied. The frequency, ν , at which a given nucleus will change spin state, and ν_{ref} represents the frequency at which the nucleus of interest will resonate (based on equations A.2 and A.7).

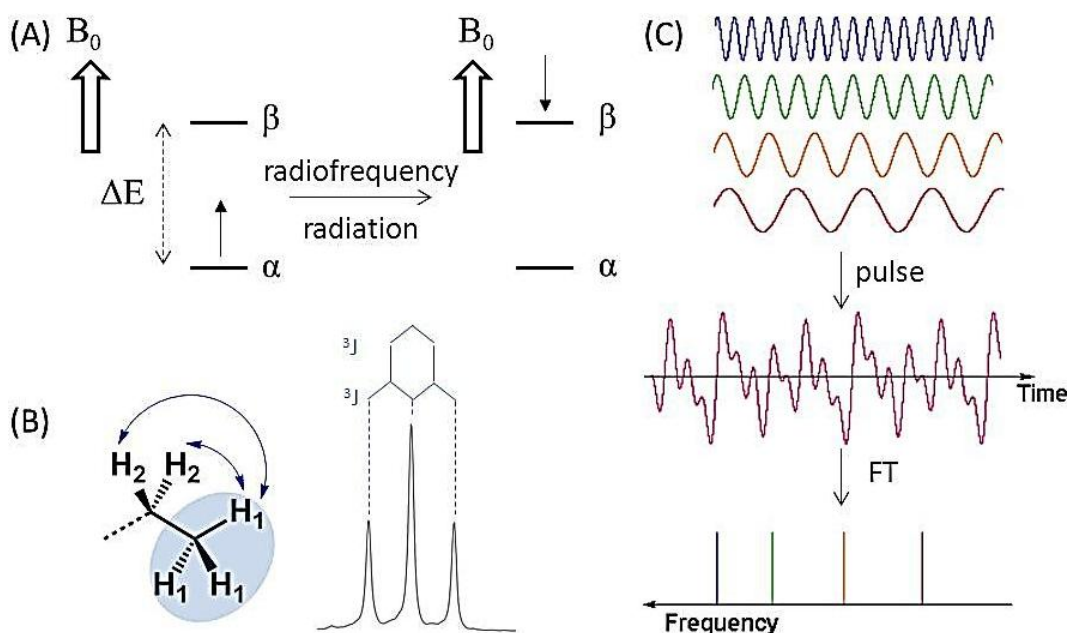


Figure A.4 – Scheme illustrating (A) the absorption of radiation by a nucleus $I = 1/2$ resulting in a spin state transition associated with a change in the spin orientation with respect to the magnetic field applied (B_0), (B) the scalar coupling between non-equivalent protons (H_1 , H_2) and the triplet signal in ^1H NMR spectroscopy associated with protons H_1 . (C) The NMR spectra measured upon irradiation of a sample at different frequencies either separately or together (pulse) and the conversion to a frequency spectrum by fourier transform; adapted from reference 14.

The direct coupling between nuclei, called cross-relaxation or Nuclear Overhauser Effect (NOE), is transmitted through space rather than chemical bonds, and can provide important information on distances separating nuclei. However, NOE are averaged to zero by molecular rotation (tumbling) and therefore do not result in peak splitting upon classic NMR acquisition. NOE spectroscopy (NOESY) uses phasing/dephasing pulse sequences associated with defined mixing times, thus allowing for detection of the NOE coupling between resonances.[2]

Two protons from a CH₂ group, the local environment of which includes a chiral center, are not symmetry related and have different resonances. For examples, the two β hydrogen atoms of a cysteine can be differentiated based on their orientation with respect to the α-proton, and adopt different resonances (β₁ and β₂) which couple with each other (²J) and the α-proton (³J), resulting in an ABX pattern (see Figure A.5A). In some cases, the observed pattern might be complicated by second-order effects.

Another phenomenon encountered in NMR and relevant to this work is the dynamic effect. Molecules can adopt different conformations or form complexes with other molecules, hence different species, which are linked by chemical equilibria, can be present in solution. A proton resonance changes depending on its environment, hence different resonances can be observed for the different species. However, the different species related to a given molecule have, in most cases, chemical exchange processes which occur faster than the measurement, and their resonances are averaged based on their percentage contribution, and a single signal is observed. However, for processes which occur slower than the NMR timescale (e.g. hindered rotation or complexation), the long lived conformations or species will give distinct signals (see Figure A.5B).

If one neglect the magnetic inhomogeneity of the sample, the peak width in NMR is inversely proportional to the transverse relaxation time (T₂). The spin of nuclei with $I > \frac{1}{2}$ (such as ²H or ¹⁴N) which possess more than two spin states, have shorter T₂ due to a

phenomenon called quadrupolar relaxation. Upon complexation of transition metal ions, the relaxation time T_2 of the ligand nuclei might be shortened due to quadrupolar relaxation (due to the organisation of orbitals relative to the different spin levels), resulting in broadening of the spectra. If the metal ion is paramagnetic, the relaxation is very fast and the peaks are significantly broadened, often to the extent that they cannot be observed.

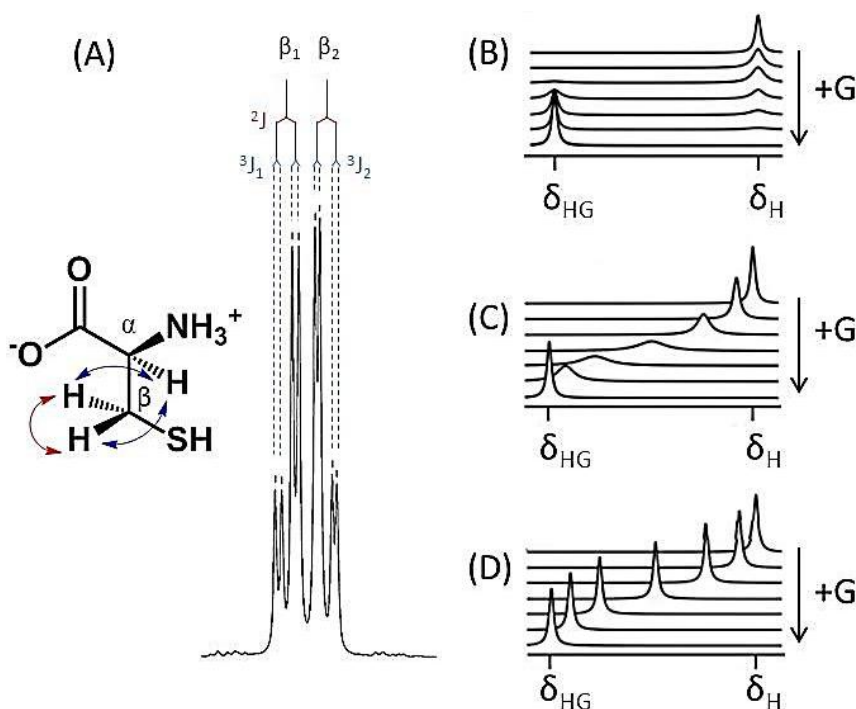


Figure A.5 – (A) The ^1H NMR signals of the two β hydrogens ($\text{H}\beta$) from L-cysteine in D_2O at pH 7.4, 298 K and 500 MHz and a scheme illustrating the scalar coupling between the $\text{H}\beta$ (blue, 2J) and $\text{H}\alpha$ (red, 3J); adapted from reference 15. Scheme monitoring the formation of a complex, HG, upon addition of a guest molecule, G, to a solution containing a host molecule, H, consistent with either (B) a slow, (C) an intermediary, or (D) a fast process on the NMR timescale; adapted with permission from reference 16. Copyright 2013 Elsevier.

A.1.5 – Mass Spectrometry

Mass spectrometry (MS) is a gas-phase spectroscopic technique which involves ionising molecules *in vacuo*, separating them based on their mass over charge ratio (m/z), and their detection, thus allowing for deduction of the exact mass. Different techniques exist for ionisation and analysis/detection, and only those methods relevant to this work will be described. Electron impact (EI) is the oldest ionisation technique and involves bombarding the target molecule with a high energy electron beam resulting in loss of electrons and/or fragmentation of the molecule (species type M^+ or $\text{M}^{\bullet+}$, where M is the compound of interest,

are usually obtained). In contrast, electrospray ionisation (ESI) involves passing a solution through a thin charged needle, resulting in the formation of ionised droplets which transit to the gas phase. This soft ionisation technique is particularly useful for measuring the mass of large macromolecules, and tends to result in several adducts with different charges (the most common are of the form $[M+xH]^{x+}$, where M is the compound of interest, H hydrogen atoms, and x the number of hydrogen or charge) are usually obtained. Matrix-assisted laser desorption/ionisation (MALDI) involves co-crystallising the sample in the presence of a matrix material with strong UV absorption properties. The resulting analyte deposited on a plate is irradiated with a UV laser beam resulting in simultaneous desorption, vaporisation and ionisation of the analyte. Due to the protective role of the matrix, MALDI can be classified as a soft ionisation technique, which mainly produces single-charged adducts (of the type $[M+H]^+$). The MS spectrometers used in this study were all equipped with a time-of-flight (TOF) detector which allows for the measurement of the m/z of ionised molecules based on the time they require to cross a chamber and reach the detector.[2]

A.2 – Electrophoretic techniques

A.2.1 – Study of protein-DNA interactions

Electrophoresis refers to the movement of particles in a solid (gel) across which an electric field is applied, and allows for the separation of macromolecules based on their size and/or charge, for both preparative and analytical purposes. The electrophoretic mobility shift assay (EMSA) is a technique suitable for the study of specific or non-specific interactions between proteins and DNA, if the complex formed is stable under the electrophoretic conditions. This is based on the different retention of DNA-protein complex from the DNA strands alone under the gel conditions (see Figure A.6A), the intensity of the resulting bands can be quantified and used to estimate the protein-DNA dissociation constant. If one considers the simple case of a protein binding a particular DNA site as a monomer, the dissociation constant is calculated following equations A.9-A.11:

$$K_D = \frac{[P]_F [D]_F}{[PD]} \quad (\text{A.9})$$

$$[P]_T = [P]_F + [PD] \quad (\text{A.10})$$

$$[D]_T = [D]_F + [PD] \quad (\text{A.11})$$

K_D represents the protein-DNA dissociation constant, $[P]_F$ and $[P]_T$ the DNA unbound and total concentration of the protein, $[D]_F$ and $[D]_T$ the protein unbound and total concentration of the DNA respectively, and $[PD]$ the concentration of the protein-DNA complex. When $[D]_F \ll K_D$, then $[P]_F \gg [PD]$ and one can write $[P]_T = [P]_F$ thus obtaining Equation A.12:

$$K_D = \frac{[P]_T [D]_F}{[PD]} \quad (\text{A.12})$$

Therefore, the protein-DNA dissociation constant can be calculated experimentally if the proportion of free DNA over protein-DNA complex, and the total protein concentration, are known. When the concentration of free DNA is equal to the concentration of protein-DNA complex, the total protein concentration is equal to the protein-DNA dissociation constant.[17]

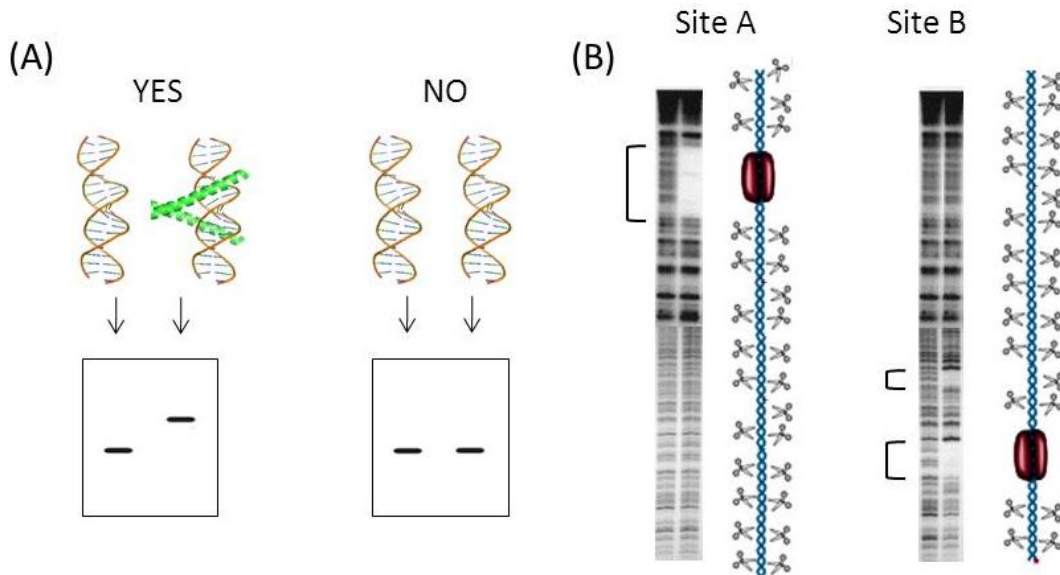


Figure A.6 – Scheme illustrating (A) the use of EMSA to evaluate whether a protein binds to a specific DNA duplex, or (B) the use of DNase footprinting to locate the binding site of a protein within a DNA strand based on the enzymatic cleavage pattern; adapted from references 18 (with permission) and 19. Copyright 2001 Elsevier.

There exist alternative/complementary electrophoretic methods for the study of protein-DNA interaction, including DNase footprinting, which is mentioned in this report. It involves comparing the electrophoretic profiles of DNA fragments resulting from enzymatic cleavage either in the absence or presence of the DNA binding protein, which can protect its binding site from cleavage (see Figure A.6B). Importantly, it allows for confirmation of the protein-DNA binding site, even if the protein-DNA complex is not stable under the electrophoretic conditions.

A.2.2 – Cleavage of supercoiled plasmid DNA

As for proteins, DNA may adopt a large variety of structures. Plasmids are a type of non-chromosomal double-stranded DNA, consisting of several thousand **bp**, which are usually circular (the 5'- end of each strand is directly connected to its 3'- end). In their supercoiled form, the circular plasmids are twisted over themselves forming highly compact structures (see Figure A.7A). A cut to one of the DNA strands will promote the transition to the circular relaxed form, whereas a double stranded cut (or cuts to both strands within 12 **bp**) will promote formation of the linear form (see Figure A.7A). Supercoiled plasmid DNA is largely used to test the ability of a given molecule to cleave DNA, because DNA cleavage is thermodynamically favoured for the large compact form, moreover, the different forms of plasmid can easily be separated and quantified by electrophoresis. For a particular plasmid (sequence and size), the more compact is its form, the faster it migrates on a gel, hence supercoiled plasmid (form I) migrates faster than linear plasmid (form III), itself migrating faster than the relaxed form (form II) (see Figure A.7B). However if several double-stranded cuts occur, the shorter linear DNA strands might migrate quicker than the supercoiled form due to their lower molecular weight. In a typical DNA cleavage assay, the conversion of form I directly into form III indicates that DNA is cut through a double-strand, rather than a single-strand cleavage mechanism.[20]

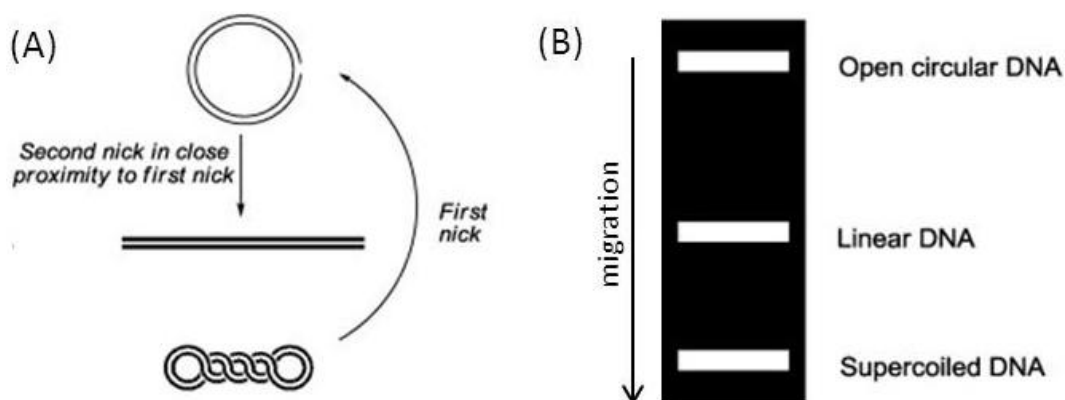


Figure A.7 – Scheme illustrating (A) the shape of the three forms of plasmid DNA and the transition between them promoted upon DNA cleavage, and (B) their theoretical retention in horizontal electrophoresis; reproduced with permission from reference 21. Copyright 2013 Elsevier.

A.3 – Miscellaneous methods

A.3.1 – Solid-phase peptide synthesis (SPPS)

Solid-phase synthesis is a process which involves assembling stepwise the different units of a polymeric molecule on a functionalised solid-support. For each unit, all but one chemical functions are protected, ensuring the coupling of a single unit per step. The excess of reagent, which remains in solution, can be separated from the support by filtration and washing. Then, a function is selectively deprotected to allow for the coupling of a new unit (Figure A.8). Following the last coupling, the polymer product is cleaved from the support.

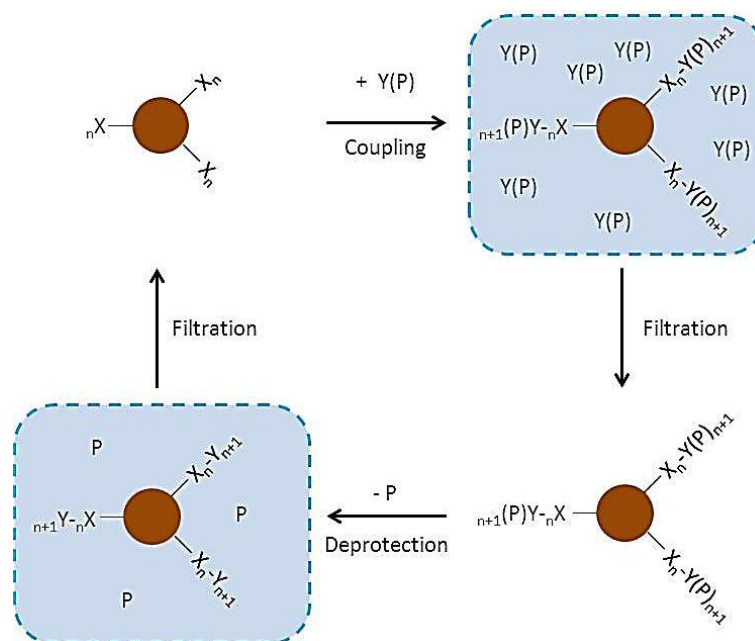


Figure A.8 – Scheme representing the multistep assembly of a polymer on a solid-support, where X and Y are residues, P a protecting group, n and n+1 the positions in the sequence.

Nowadays, SPPS is commonly used for the preparation of short peptides (~ 2-60 residues). Proteinogenic amino-acids have two or three chemical functions, the protection of which has been achieved with a large variety of groups. The fluorenylmethyloxycarbonyl- (Fmoc) group is a chromophore commonly used in SPPS to protect amino-acids N-terminus (amine). Fmoc SPPS involves attaching the C-terminus (carboxylic acid) of such residues to the solid-support, then the Fmoc group is removed using a base affording the free amine to which a new Fmoc-protected residue can be coupled. The side-chains, which usually bear miscellaneous acid-sensitive protecting groups, remain protected during the chain elongation process. In this work, peptides are prepared by Fmoc SPPS on rink amide 4-methylbenzhydrylamine (MBHA) resin. The peptide chain is elongated on an automated peptide synthesiser using (O-(benzotriazol-1-yl)-NN,N',N'-tetramethyluronium) hexafluorophosphate (HBTU) and diisopropylethylamine (DIEA) as acid and amine activators, respectively, in order to increase the rates of coupling. Then, its N-termini is capped with an acetyl (Ac) group to avoid further coupling, and the peptide is simultaneously cleaved from the support and its side-chains deprotected with an acid (cleavage from Rink amide MBHA resin results in C-termini capping with $-NH_2$).[22]

A.3.2 – Reversed-phase high performance liquid chromatography

Liquid chromatography involves eluting a mixture through an adsorbent material (static phase) for which the different components of the mixture display different affinities, using a liquid eluent (mobile phase) as a vector in order to separate the components of the mixture. Several types of static phases can be suitable depending on the properties of the components to be separated. In high-performance liquid chromatography (HPLC), migration through a static phase with small particle size and tightly packed in columns, is made possible by applying a high pressure to the system thus allowing for a better separation (or higher resolution) of mixtures. The content of the eluent leaving the column is usually monitored by

a non-destructive and in-flow spectroscopic technique (such as electronic absorption or emission), allowing for the analysis of the efficiency of the separation, and for recovery of the components. The process can be applied either on an analytical or preparative scale depending on the size of the column and the eluent flow, however resolution is generally greater on thinner columns. Reversed-phase HPLC (RP-HPLC) apparatus involves a nonpolar static phase and a moderately polar eluent and is well suited for the separation of peptides. Using this technique, less polar and low molecular weight components are retained less on the static phase and migrate faster than their nonpolar high molecular weight counterparts. The eluent often consists of a solvent mixture (usually two to four different constituents), and elution processes involving an eluent with constant composition are called isocratic, whereas the eluent composition is gradually varied in gradient processes.[23]

A.3.3 – C H N analysis

C H N analysis is a destructive analytical technique which involves combusting a known amount of a pure sample (ideally a single type of molecule) using an enriched oxygen atmosphere, collecting the different combustion products (typically carbon dioxide, water and nitric oxide), and quantifying them in order to determine the precise elemental carbon, hydrogen and nitrogen composition of the molecule.[24]

A.4 - References

-
- [1] Kobayashi N., Muranaka A., Circular dichroism and magnetic circular dichroism spectroscopy for organic chemists, Chapter 1: Theory of optical spectroscopy, Royal Society of Chemistry, Cambridge, **2011**, ISBN 978-1-84755-869-5.
 - [2] Harwood L. M., Claridge T. D. W., Introduction to organic spectroscopy, Oxford University Press, New York, **1997**, ISBN 978-0-19-855755-5.

- [3] http://chemwiki.ucdavis.edu/Organic_Chemistry/Organic_Chemistry_With_a_Biological_Emphasis/Chapter_4%3A_Structure_Determination_I/Section_4.3%3A_Ultraviolet_and_visible_spectroscopy, accessed on 27/08/2013.
- [4] Woody R. W., Koslowski A., *Biophys. Chem.*, **2002**, 101-102, 535-551.
- [5] Pace C. N., Schmid F. X., Protein Structure, Chapter 10: How to determine the molar absorbance coefficient of a protein, Oxford University Press, New York, **1997**, ISBN 0-19-963618-4.
- [6] Ellman G. L., *Arch. Biochem. Biophys.*, **1959**, 82, 70-77.
- [7] Collier H. B., *Anal. Biochem.*, **1973**, 56, 310-311.
- [8] Wolf C., Bentley K. W., *Chem. Soc. Rev.*, **2013**, 42, 5408-5424.
- [9] Masiero S., Trotta R., Pieraccini S., DeTito S., Perone R., Randazzo A., Spada G. P., *Org. Biomol. Chem.*, **2010**, 8, 2683-2692.
- [10] Miles A. J., Wallace B. A., *Chem. Soc. Rev.*, **2006**, 35, 39-51.
- [11] Wallace B. A., Janes R. W., Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy, IOS press, Amsterdam, **2009**, ISBN 1-60750-000-1.
- [12] Wayne C. E., Wayne R. P., Photochemistry, Oxford University Press, New York, **1996**, ISBN 0-19-855886-4.
- [13] Liu T., Sullivan J. P., Pressure and Temperature Sensitive Paints, Springer, Berlin, **2005**, ISBN 978-3-540-22241-5.
- Lorenz M., Horbach T., Schulz A., Bauer H., *J. Turbomach.*, **2013**, 135, 031003.
- [14] <http://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/spectrpy/nmr/nmr2.htm>, accessed on 27/08/2013.
- [15] http://www.bmrb.wisc.edu/metabolomics/mol_summary/spectrum_display.php?title=1D%201H&mol_dir=L_cysteine&exp=1H&mol=L-cysteine&bmr bid=bmse000034&pH=7.4, accessed on 27/08/2013.

- [16] Kleckner I. R., Foster M. P., *BBA Proteins Proteom.*, **2011**, 1814, 942-968.
- [17] Fairall L., Buttinelli M., Panetta G., DNA-Protein Interactions, Chapter 5: Bandshift, gel retardation or electrophoretic mobility shift assays, Oxford University Press, New York, **2000**, ISBN 978-0-19-963691-4.
- [18] Rodgers J. S., Hocker J. R., Hanas R. J., Nwosu E. C., Hanas J. S., *Biochem. Pharmacol.*, **2001**, 61, 1543-1550.
- [19] <https://www.nationaldiagnostics.com/electrophoresis/article/dnase-i-footprinting>, accessed on 28/08/2013.
- [20] Mancin F., Tecilla P., Metal-Complex DNA interactions, Chapter 13: Artificial restriction agents for DNA cleavage, John Wiley & Sons, Chichester, **2009**, ISBN 978-1-4051-7629-3.
- [21] Desbouis D., Troitsky I. P., Belousoff M. J., Spiccia L., Graham B., *Coord. Chem. Rev.*, **2012**, 256, 897-937.
- [22] Chan, W. C., White, P. D., Fmoc Solid Phase Peptide Synthesis: A Practical Approach, Oxford University Press, New York, **2000**, ISBN 978-0199637249.
- [23] Ahuja S., Rasmussen H., HPLC method development for pharmaceuticals, Elsevier Academic Press, Amsterdam, **2007**, ISBN 978-0-12-370540-2.
- [24] Dean J. A., Analytical chemistry handbook, section 17, McGraw-Hill, New York, **1995**, ISBN 0-07-016197-6.