THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

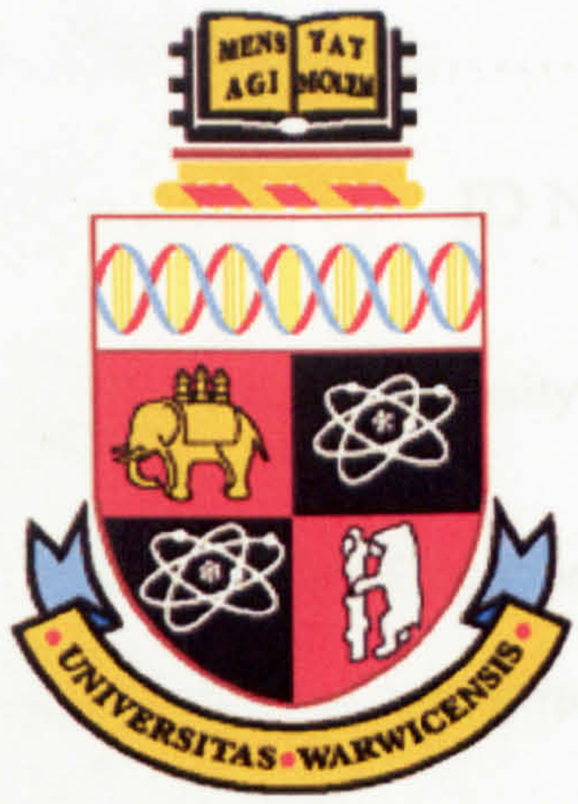
**A Thesis Submitted for the Degree of PhD at the University of Warwick**

http://go.warwick.ac.uk/wrap/4069

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.




# Biomacromolecule–ligand Interactions

**By**

**Jemma C. Peberdy**

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Chemistry

University of Warwick, Department of Chemistry
December 2004

*For my Dad and Mum*

*I couldn't have done this without you.*

*All my love now, always and forever*

*Jemma*

## Contents


Contents i

Acknowledgments v

Declaration vi

Abbreviations vii

Abstract xi

**Chapter 1 – Introduction** 1

1.1 Introduction 3

1.2 Introduction to DNA structure and function 3

1.2.1 Structure and geometry of A-, B- and Z-DNA 5

1.2.2 The history of nucleic acids and DNA 6

1.3 DNA-ligand binding modes 7

1.3.1 Intercalation between DNA base pairs 7

1.3.1.1 Metallo-intercalators 8

1.3.2 External association 9

1.3.3 Major and minor groove binding 10

1.3.4 Covalent binding of metal complexes 11

1.3.5 Non-covalent binding of metal complexes 13

1.4 Zinc finger proteins 14

1.4.1 DNA sequence recognition 15

1.5 Supramolecular chemistry 16

1.5.1 Supramolecular helicates 17

1.6 Techniques 19

1.6.1 UV/visible absorbance spectroscopy 19

1.6.2 Circular dichroism 20

1.6.3 Linear dichroism 20

1.6.4 Fluorescence spectrophotometry 22

1.6.5 Molecular dynamics simulations 22

1.7 Objectives 24

1.8 References 24

**Chapter 2 – Metal complex synthesis, characterisation and enantiomeric separation** 31

2.1 Introduction 33

2.2 Bimetallo iron cylinders 33

2.3 Chromatography 35

2.3.1 Chiral resolution of transition metal complexes 35

2.3.2 Cellulose chromatography 37

2.4 Bimetallo cylinder spectroscopy 38

2.5 Materials and methods 40

2.5.1 Synthesis of methylated metal complexes 41

2.5.1.1 Characterisation data 42

2.5.2 Determination of ε 43

2.5.3 Stability studies 43

2.5.3.1 Photostability studies 43

2.5.3.2 Thermal stability of enantiomers 44

2.5.4 Paper chromatography 44

2.5.5 DNA-paper chromatography 44

2.5.6 Cellulose column chromatography 45

2.5.7 High-pressure liquid chromatography 46

2.5.8 High temperature docking calculations 47

2.6 Results 47

2.6.1 Synthesis of methylated metal complexes 47

2.6.2 Determination of ε 49

2.6.3 Photostability experiments 50

2.6.4 Thermal stability experiments 54

2.6.5 Paper chromatography 56

2.6.6 DNA-paper chromatography 58

2.6.7 Cellulose column chromatography 58

2.6.8 High-pressure liquid chromatography 66

2.6.9 High temperature docking calculations 66

2.7 Discussion 68

2.8 References 69

**Chapter 3 – DNA binding studies** 73
3.1 Introduction 75
3.1.1 Data Analysis 78
3.2 Materials 79
3.2.1 UV/Visible absorbance spectroscopy of DNA and the metal complexes 80
3.3 Methods 81
3.3.1 Determination of ethidium bromide binding constant at 50 mM NaCl 81
3.3.2 Determination of metal complex binding constants 81
3.3.3 DNA binding studies 82
3.3.3.1 *LD* binding studies 82
3.4 Results 83
3.4.1 Determination of ethidium bromide binding constant at 50 mM NaCl 83
3.4.2 Determination of metal complex binding constants 85
3.4.3 DNA binding studies 88
3.4.3.1 DNA binding of metal complex $L_3$ 88
3.4.3.2 DNA binding of metal complex $L_5$ 98
3.5 Conclusions 107
3.6 References 110

**Chapter 4 – Probing the metal complex DNA-binding geometry** 113
4.1 Introduction 115
4.1.1 Hoechst 33258 115
4.1.2 Cobalt (III) hexammine 116
4.1.3 High temperature docking calculation 117
4.2 Materials and methods 118
4.2.1 Hoechst 33258 competitive binding assay 118
4.2.2 Cobalt (III) hexammine competitive binding assay 118
4.2.3 High temperature docking calculations 119
4.3 Discussion 119
4.3.1 Hoechst 33258 competitive binding assay 119
4.3.2 Cobalt (III) hexammine competitive binding assay 122
4.3.3 High temperature docking calculations 128
4.4 Discussion 136
4.5 References 137

**Chapter 5 – Oligonucleotide-protein interactions** **139**

5.1 Introduction 141

5.2 Guanosine rich oligonucleotides 141

5.2.1 Anti-proliferative activity of GRO's 145

5.2.2 Summary of GRO structure 146

5.3 Protein-DNA interactions 151

5.3.1 Protein secondary structure 152

5.3.2 Nucleolin 153

5.4 Materials and methods 154

5.4.1 Structural characterisation of oligonucleotides 155

5.4.1.1 Investigating the effect of freezing and annealing 155

5.4.1.2 Cation and concentration effect on quadruplex formation 156

5.4.2 Structural characterisation of nucleolin-RBD 156

5.4.3 Oligonucleotide-nucleolin-RBD interactions 158

5.4.3.1 Kinetic and thermodynamic stability of oligonucleotides and nucleolin-RBD 158

5.5 Results 159

5.5.1 Structural characterisation and biological activity of oligonucleotides 159

5.5.1.1 Investigating the effect of freezing and annealing 161

5.5.1.2 Cation and concentration effect on quadruplex formation 164

5.5.2 Structural characterisation of nucleolin-RBD 168

5.5.3 Oligonucleotide-nucleolin-RBD interactions 169

5.5.3.1 Kinetic and thermodynamic stability of oligonucleotides and nucleolin-RBD 171

5.6 Discussion 175

5.7 References 176

**Chapter 6 – Conclusions and future work** **181**

Appendix I 189

Appendix II 203

Appendix III 227

**Acknowledgements**

First and foremost I would like to thank my supervisor, Prof. Alison Rodger for all of her help, guidance and support over the last three years. I would also like to take this opportunity to thank Prof. P. Mark Rodger and Alison for letting me live with them during the last few difficult weeks of my research. I will be forever grateful.

Many people have helped and supported me during my time at Warwick and I would like to thank all of them. I would like to thank Dr. Syma Khalid for all of her help with the molecular simulations and writing of this thesis. I'd also like to thank Dr. Mike Hannon and the Hannon and Rodger research groups, in particular, Dr. Rachel Marrington, for proof reading some of my chapters and Ursula Mc Donnell, Jascindra Rajendra, Dr. Jarek Malina and Raul Pacheco-Gomez for their friendship and support.

I would like to thank Dr. Ralph Noble, Prof. Paula Bates and Prof. John O. Trent for their helpful comments during the writing of this thesis and I would also like to thank Prof. P. Mark Rodger again and Mark Broome for their help with the molecular simulations and all of the people in the molecular modelling laboratory for their friendship, help and support.

I'd like to thank my partner and best friend Steven Harrison for all of his love and support during my studies and to thank him for being my 'shoulder to cry on' and 'personal punching bag' during the more frustrating times.

Finally I want to thank my family, especially my Dad, Brian, for helping me to realise that dreams can become a reality if you work hard enough. I know that I've been difficult to live with over the last three years, but I'd like to thank you for just being you. Thank you for giving me the motivation to continue, when times were really hard. I couldn't have done this without you. I'd like to thank my mum, Linda, and sisters, Candice and Olivia, for all of their love and support over the last three years. I couldn't have done this without the love and support given to me by my family and friends. Thank you to you all.

## Declaration

The observations and recommendations in this thesis are those of the author except where acknowledgement has been made to results and ideas previously published. The research was undertaken at the Department of Chemistry, University of Warwick between October 2001 and December 2004 and has not been previously submitted for a degree at any other institution.

## Abbreviations

| | |
|---|---|
| μ | micro, $10^{-6}$ |
| λ | wavelength |
| a.a. | amino acid residues |
| A | adenine |
| A/bA | acrylamide/bis-acrylamide |
| approx. | approximately |
| APS | ammonium persulphate |
| AT- | adenine-thymine synthetic DNA homopolymers, poly[d(A-T)$_2$] |
| bipy | bipyridine |
| bp | base pairs |
| B.P. | boiling point |
| C | cytosine |
| *CD* | circular dichroism |
| CHN | carbon nitrogen hydrogen analysis |
| CRO | cytosine-rich oligonucleotide |
| CRO1 | structurally characterised CRO |
| ct- | calf thymus |
| DNA | deoxyribonucleic acid |
| ds-DNA | double stranded DNA |
| ε | extinction coefficient, molar absorptivity |
| EB | ethidium bromide |
| ENT | enantiomer |
| ESI-MS | electrospray ionisation mass spectrometry |
| EtOH | ethanol |
| fs | femtosecond |
| FTIR | Fourier transform infra red |
| G | guanine |
| GC- | guanine-cytosine synthetic DNA homopolymers, poly[d(G-C)$_2$] |
| GRO | guanosine-rich oligonucleotide |
| GRO1 | structurally characterised GRO |
| -H | Hoechst 33258 |
| HIV | human immunodeficiency virus |

| | |
|---|---|
| *ICD* | induced circular dichroism |
| ISC | ideal separating conditions |
| kDa | kilo Dalton |
| *LD* | linear dichroism |
| m | medium intensity IR signal |
| MC | metal complex |
| mins | minutes |
| mm | millimetres |
| m/z | mass to charge ratio |
| (M)- | left handed cylinder |
| M | molar concentration mol $dm^{-3}$ |
| mdeg | millidegrees |
| MD | molecular dynamics |
| MeCN | acetonitrile |
| MeOH | methanol |
| MLCT | metal to ligand charge transfer |
| MWT | molecular weight |
| n-BuLi | n-butyl lithium |
| ns | nanosecond |
| NMR | nuclear magnetic resonance |
| Nu | nucleolin-RBD |
| NVT | constant number of molecules, volume and temperature |
| (P)- | right handed cylinder |
| phen | phenanthroline |
| pI | isoelectric point of protein |
| ps | picosecond |
| *rac* | racemic |
| RBD | RNA binding domain |
| $R_f$ | resolution factor |
| RNA | ribonucleic acid |
| s | strong IR signal |
| s.s. | stock solution |
| SDS-PAGE | Sodium dodecyl sulphonate polyacrylamide gel electrophoresis |
| T | thymine |

| | |
|---|---|
| THF | tetrahydrofuran |
| UV | ultraviolet |
| vs | very strong IR signal |
| v/v | volume per volume |
| w | weak IR signal |

**Abstract**

The interactions and binding of various ligands to biomacromolecules e.g. DNA and proteins finds widespread application in the design and development of novel pharmaceuticals. DNA has been identified as the target molecule for a number of drugs and carcinogens and the supramolecular synthetic approach has led to the discovery of a range of bimetallo iron cylinders that bind to DNA inducing remarkable structural effects. The cylinders are chiral and the enantiomers were separated on cellulose packed in paper or in a column. The optimum mobile phase for efficient separation was found to be 90% acetonitrile:10% 0.02 M NaCl. The (M)-enantiomers of the parent cylinder have been found to bind to DNA in the major groove. Hydrophobic methyl groups were added at various positions on the ligand backbone. UV/visible absorbance, circular and linear dichroism were used to investigate any interactions of the metal complex with DNA with the aim of investigating any sequence preference or selectivity upon binding. Competitive binding studies and molecular dynamics simulations were used to probe the binding geometries of the enantiomers of the parent cylinder and two methylated cylinders to DNA as the exact site of interaction of the (P)-enantiomers of the parent cylinder was unclear. It was concluded that the methylated bimetallo iron cylinders bind to DNA and provide major groove recognition and may show some sequence preference.

Circular dichroism was used to structurally characterise a range of guanosine-rich oligonucleotides (GRO's) and to investigate their interactions with a nucleolar protein — nucleolin. Biological/anti-proliferative activity has been related to the ability of the oligonucleotide to bind to this protein. It was found that nucleolin does bind to a biologically active GRO in the presence of $K^+$ and induces a structural change in it.

# CHAPTER 1 – INTRODUCTION

## 1.1 Introduction

The work presented in this thesis is based on probing the interaction of different ligands with biomacromolecules including DNA and proteins. The first part of this thesis is focussed on the synthesis, characterisation, enantiomeric resolution and DNA binding of bimetallo supramolecular cylinders with different surface topologies. The cylinders are chiral and have been shown to be able to be separated on cellulose packed in paper or in a column using 0.02 M NaCl and other organic:aqueous eluents (see chapter 2.0). The chirality of the complex has been found to have a significant effect on the DNA binding of the enantiomers. UV/visible absorbance, circular and linear dichroism (*CD/LD*) binding experiments have been completed to determine any binding specificity or DNA sequence preference of two methylated cylinders to ct-, AT- and GC-DNA (see chapter 3.0). A series of competitive binding experiments using fluorescent dyes (Hoechst 33258 and ethidium bromide) and a major groove binding cation (cobalt (III) hexammine) were then designed to investigate the site of interaction of the enantiomers of the parent compound and two methylated cylinders. High-temperature molecular dynamics (MD) simulations were used to investigate the binding geometries of the enantiomers of the methylated cylinders to a DNA dodecamer (see chapter 4.0).

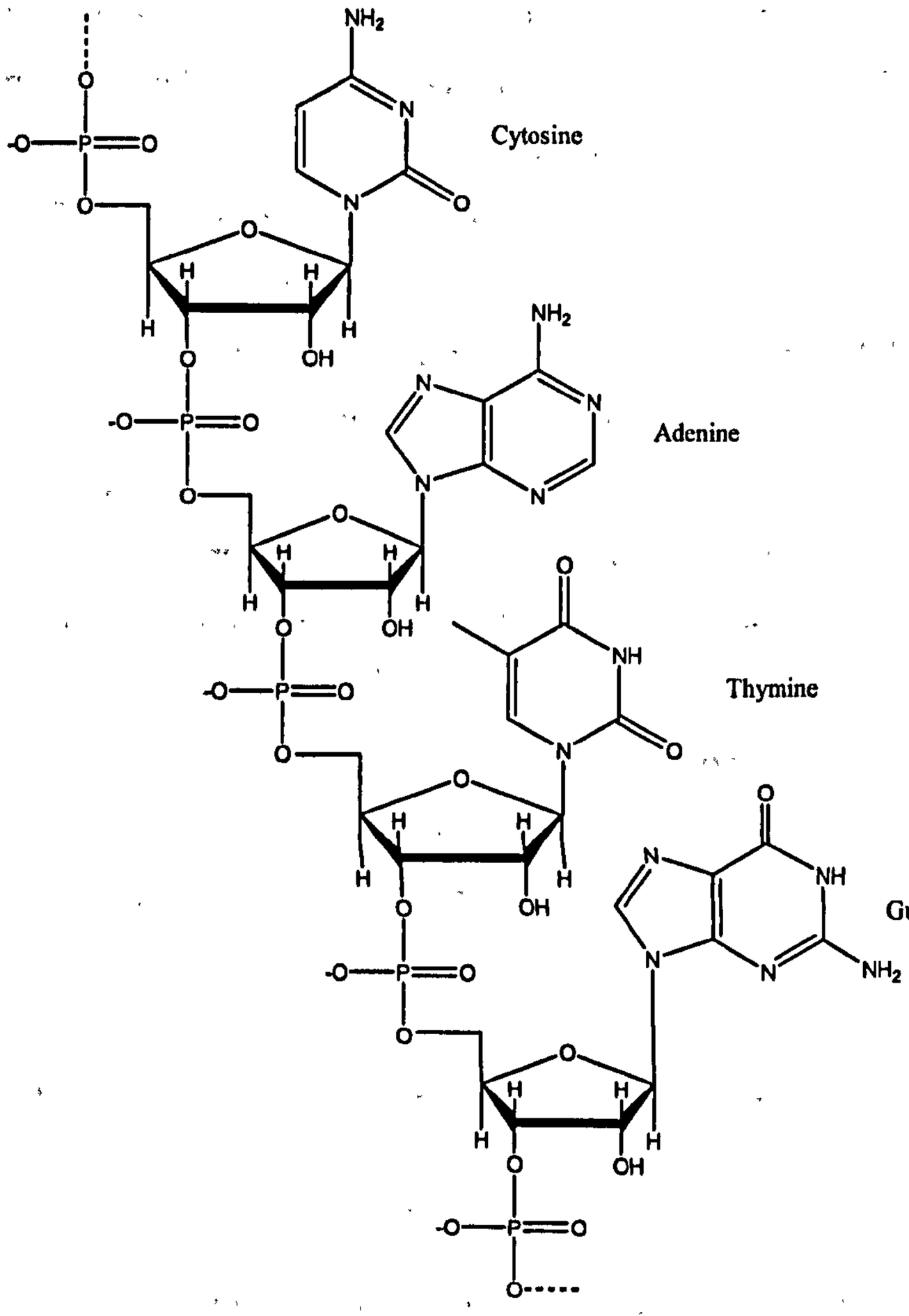
The DNA binding of the bimetallo iron cylinders has been complemented by the research presented in chapter 5. The focus of this research is on the investigation of biomacromolecule–ligand interactions and so UV/visible absorbance and *CD* spectroscopy were used to structurally characterise a range of guanosine-rich and cytosine-rich oligonucleotides (GRO/CRO) and their interactions and binding to a RNA-binding domain fragment of human nucleolin (denoted nucleolin-RBD, a ~25 kDa nucleolar protein).

## 1.2 Introduction to DNA structure and function

Deoxyribonucleic acid (DNA) is the basic hereditary material in the nucleus of all cells and contains all of the information necessary to make proteins. DNA is a linear polymer that consists of many nucleotide units.[1]

Each nucleotide has three main structural features (see figure 1.1): 1. a five membered sugar ring or furanose, which is deoxyribose in DNA; 2. a heterocyclic base, adenine

(A), guanine (G), cytosine (C) or thymine (T), attached to each sugar ring via a β-glycosyl linkage; and 3. a phosphodiester linkage joining each of the nucleoside units.[2]




A nucleoside is a combination of the heterocyclic base and the five membered sugar ring. With the addition of the phosphate group, a nucleotide is generated. Each nucleotide is connected via the O3' and O5' atoms forming the phosphodiester linkages. These two linkages mean that DNA has a 'sense' and by convention we read the DNA sequence from the 5' end of the first base.

**Figure 1.1 Single stranded DNA oligomer showing backbone and its link to each base.**

Most DNA's exist as hydrogen bonded duplexes with A and T joined by 2 H–bonds and G and C joined by 3 H–bonds. There are three common forms of duplex DNA (see figure 1.2) that are known to exist in biological systems. B-DNA is the most common physiological structure and is a right-handed helix. A-DNA is produced by the dehydration of B-DNA and is also a right-handed helix. Z-DNA is a left-handed helix with essentially no major groove.[3]

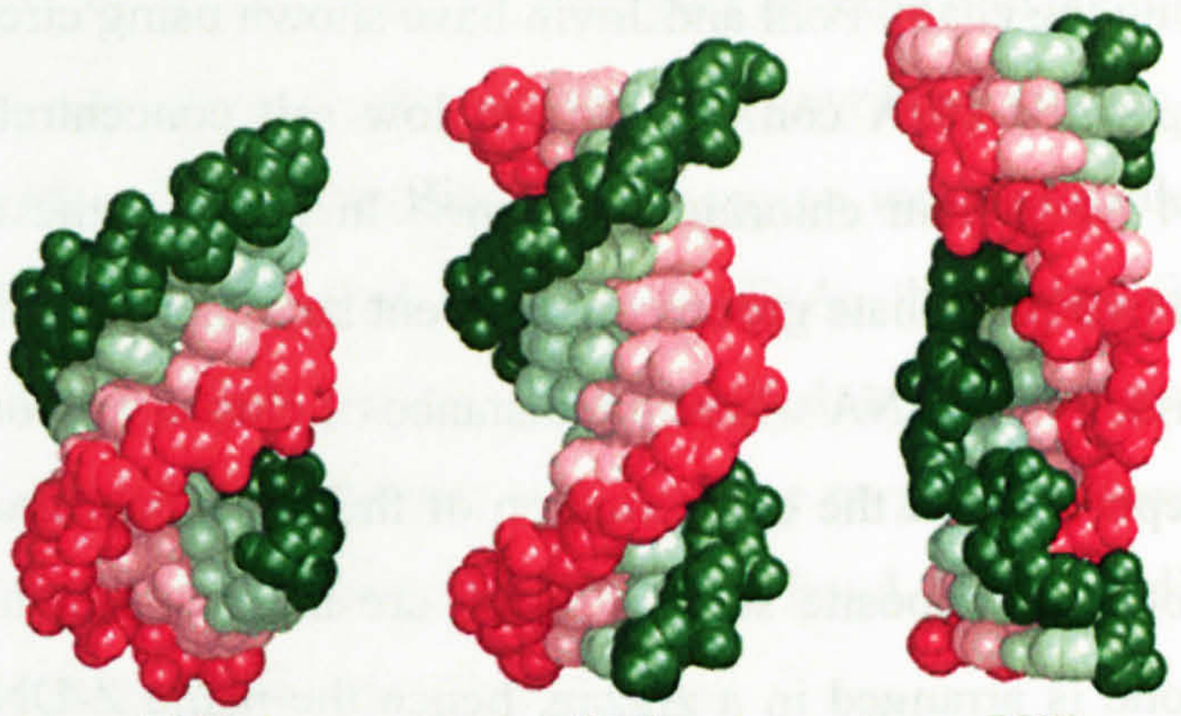



**Figure 1.2 Schematic diagram showing structures of A-DNA, a right handed helix, B-DNA, the most common physiological structure and Z-DNA, a left handed helix.[1]**

### 1.2.1 Structure and geometry of A- B- and Z-DNA

A-DNA (see figure 1.2) is formed by the dehydration of B-DNA e.g. in the presence of aqueous EtOH or excess NaCl. It is a shorter, wider helix than B-DNA and has a deep, narrow major groove which is not easily accessible to proteins. There are 11 base pair residues per turn of the helix. Its structure consists of a very shallow minor groove and the base pairs are tilted to the helical axis (see table 1.1 for geometry).[4]

B-DNA (see figure 1.3) has approximately ten base pairs per turn of the helix. It has a helical diameter of 2 nm and its pitch is 3.4 nm. Since there are ten base pair residues per turn of the helix, the distance between each pair is 0.34 nm. The base pairs are at a right angle to the helical axis and the pairing of the non-symmetric bases and their stacking at a 36° torsion angle, as well as the asymmetric link to the sugar units, gives rise to major and minor grooves on the DNA[5] (see table 1.1 for geometry).

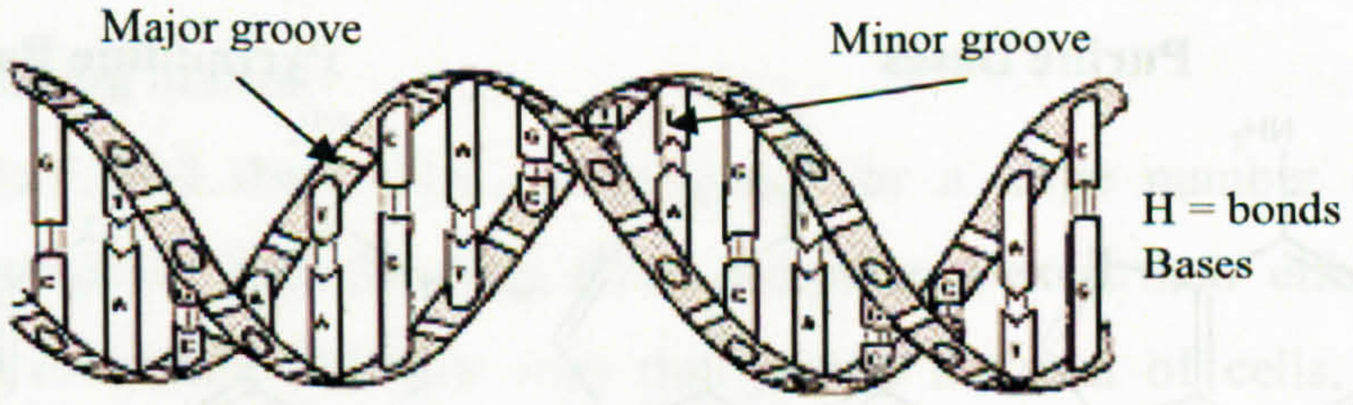



**Figure 1.3 The double helix of B-DNA as depicted by Watson and Crick showing the major and minor grooves.[6]**

The structure of Z-DNA (see figure 1.2) was first elucidated by Rich *et al.* in 1979 using high resolution X-ray crystallography.[7] Z-DNA is a left-handed helix which has 12 base pair residues per turn of the helix and essentially no major groove; every other base is rotated around the glycosyl bonds so that the bases alternate in anti- and syn-

conformations along the chain. Pohl and Jovin have shown using circular dichroism that poly(dG-dC) adopts a B-DNA conformation at low salt concentration and a Z-DNA conformation in 4 M sodium chloride solution.[8] In B-DNA the distance of closest approach between two phosphate groups on different strands is 11.5 Å across the minor groove, in comparison to Z-DNA where the distance of closest approach varies between 7.7 and 13.5 Å depending on the conformation of the phosphate group.[9] The Z-DNA has phosphate groups in opposite strands which are much closer together than in B-DNA. The backbone is arranged in a zigzag, hence the name Z-DNA.[10] Its structure shows no major groove but a shallow groove involving the same side of the bases as in the minor groove in B-DNA. Table 1.1 gives geometry parameters for the helices.

| | A | B | Z |
|---|---|---|---|
| Helix sense | Right-handed | Right-handed | Left-handed |
| Repeating unit | 1 bp | 1 bp | 2 bp |
| Rotation/bp | 33.6 ° | 35.9 ° | 60 ° per 2 bp |
| Mean bp/turn | 10.7 | 10.0 | 12.0 |
| Pitch/turn of helix | 24.6 Å | 33.2 Å | 45.6 Å |
| Sugar Pucker | C3'-endo | C2'-endo | C: C2'-endo, G: C2'-endo |
| Diameter | 26 Å | 20 Å | 18 Å |

**Table 1.1 Summary of A- B- and Z-DNA geometries to describe the differences in structure.[4]**

### 1.2.2 The history of nucleic acids and DNA

In 1868 Meischer isolated cell nuclei from pus cells and observed the presence of a phosphorous containing compound.[3] This led to the discovery of the nucleic acid. Upon hydrolysis of the DNA the purine bases: adenine and guanine and the pyrimidine bases: cytosine and thymine were observed (figure 1.4).

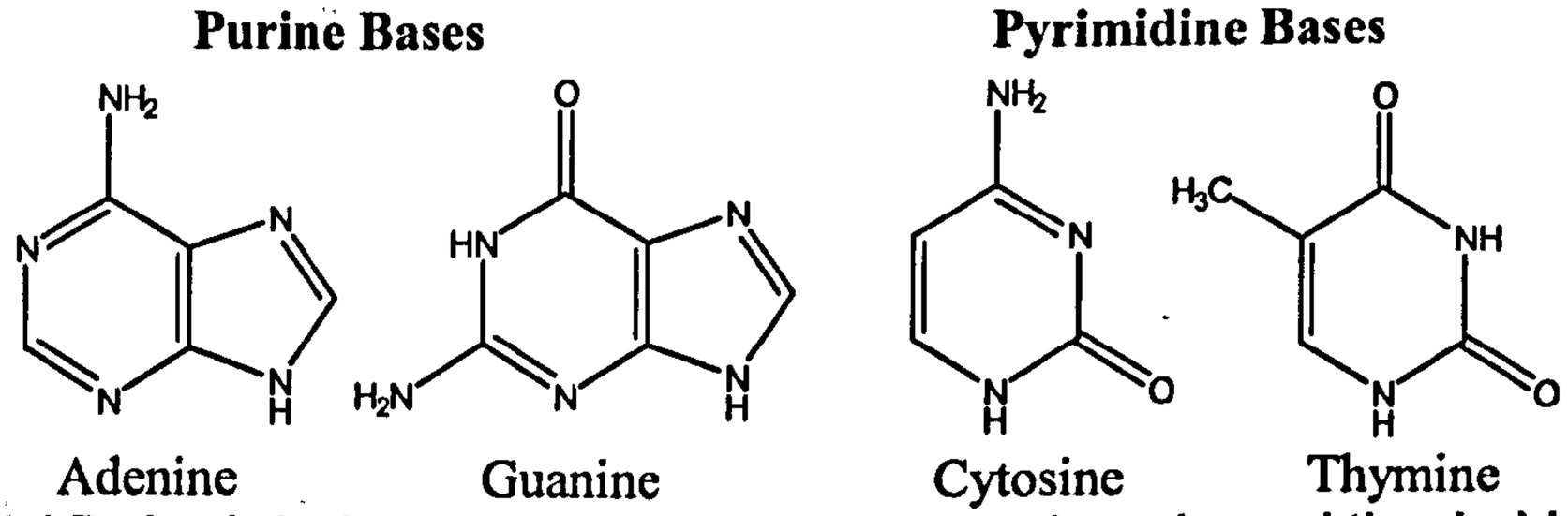



**Figure 1.4 Purine (adenine and guanine) and pyrimidine (cytosine and thymine) bases of DNA.**

The structure of DNA was first elucidated in 1953, when Watson and Crick proposed that the DNA structure consisted of two helical chains of phosphate diester groups joining deoxyribofuranose residues, largely held together by hydrogen-bonding between purine and pyrimidine bases.[11]

It had been observed earlier, in 1950, by Chargaff that the molar ratios of total purines to total pyrimidines was close to one and that in deoxyribose nucleic acid, the amounts of adenine and thymine, and of cytosine and guanine were equivalent.[12] This led to Watson and Crick's proposal that in normal DNA, the bases formed pairs in a complementary manner where purine bases bonded to pyrimidine bases and more specifically: adenine bonds to thymine and cytosine to guanine (see figure 1.5). This obeyed Chargaff's rule in that purine bases must bond to pyrimidine bases so their numbers would be approximately equal.[12, 13]

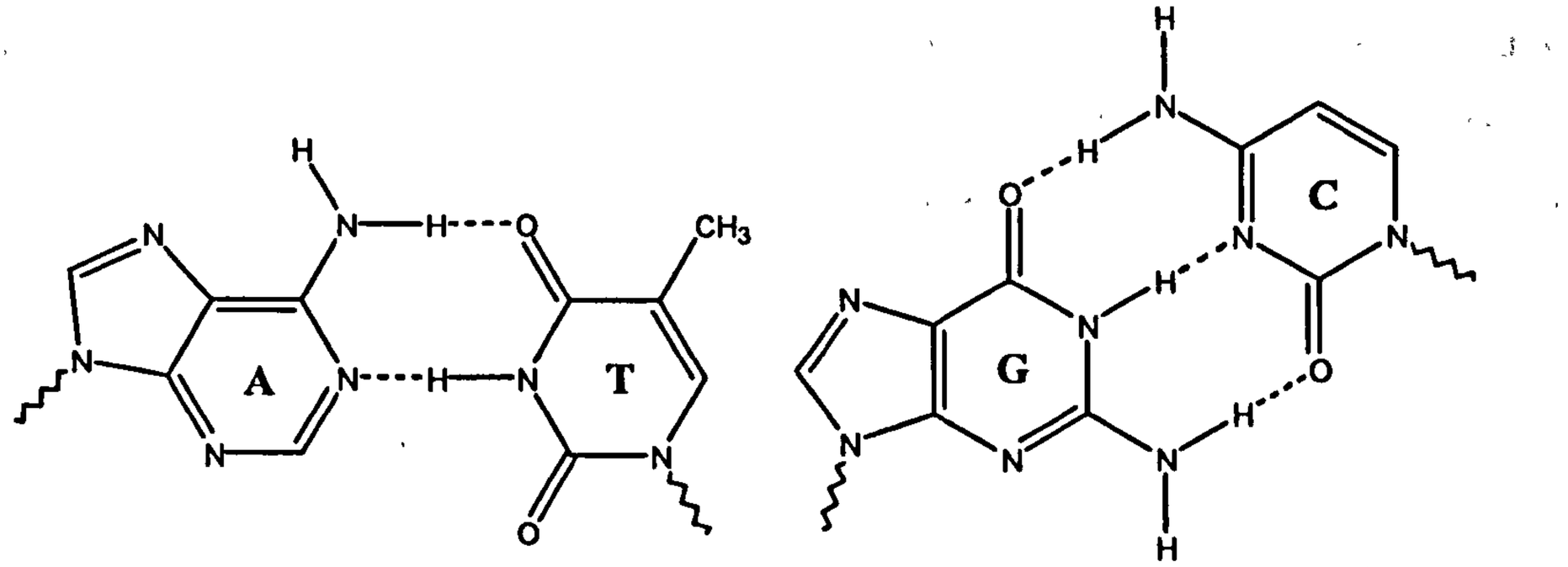



**Figure 1.5 Schematic diagram showing AT (2 hydrogen bonds) and GC (3 hydrogen bonds) base pairs.**

The guanine/cytosine base pair is held together by three hydrogen bonds which impart stability and a degree of flexibility to the base pair. In contrast, only two hydrogen bonds hold the adenine/thymine base pair together resulting in a less stable structure with increased flexibility. In addition to this hydrogen bonding, the DNA double helix configuration is stabilised by base stacking. The DNA bases are planar aromatic ring compounds and they interact via $\pi$-$\pi$ stacking forces of attraction.

## 1.3 DNA-ligand binding modes

It is now well established that DNA is the target for a large number of drugs, carcinogens and mutagens. Upon binding, these compounds exert their effect on the DNA primarily by interacting directly with the genetic material of cells, i.e. with nuclear DNA.[14 – 17] Three different modes of binding for DNA-binding ligands have been identified: intercalation, external association and (major and minor) groove binding.

### 1.3.1 Intercalation between DNA base pairs

Mainwaring defined intercalation as the sandwiching of a molecule between two adjacent pairs of bases in the DNA double helix (see figure 1.6).[18, 19] This association

involves the insertion of a planar aromatic ring system between the DNA base pairs, leading to significant π-electron overlap. This mode of binding is stabilised by stacking interactions and is thus less sensitive to ionic strength compared with the other two binding modes. Intercalation is usually favoured by the presence of an extended electron-deficient planar aromatic ligand.[20]

The structure of the DNA double helix is significantly distorted upon interaction with an intercalating compound. There is a decrease in the twist angle between the base pairs about the intercalation site, the length of the DNA molecule is increased and there is a decrease in mass per unit length.[21] All of these effects are reversible upon removal of the intercalator at room temperature.[22]

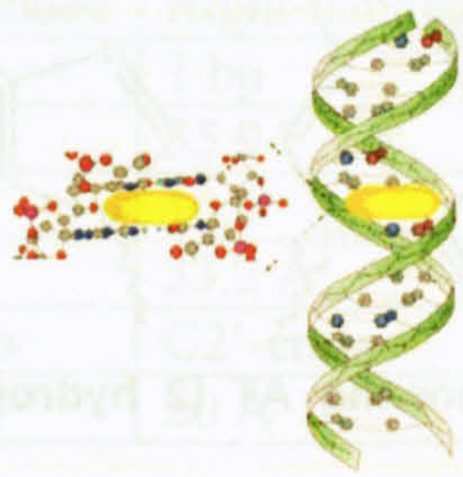


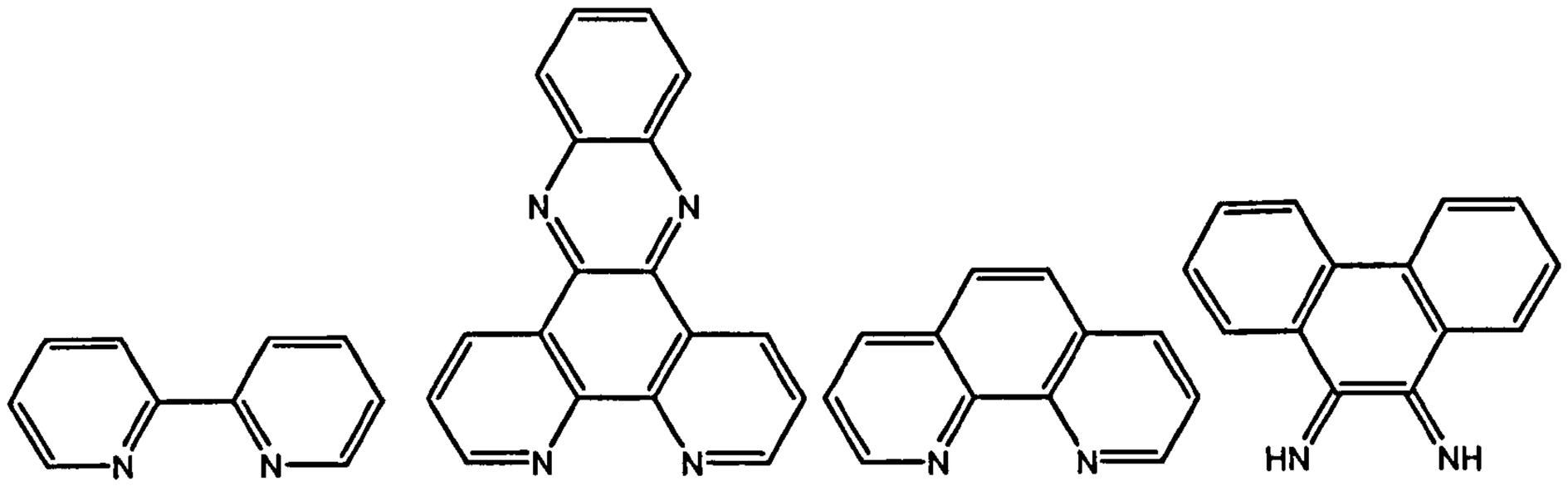
**Figure 1.6 Schematic diagram of molecule with planar aromatic ring system (yellow) intercalating between base pairs in DNA.[19]**

### 1.3.1.1 Metallo-intercalators

Over the past thirty years the use of transition metal complexes that reversibly bind to DNA has attracted a great deal of attention.[23 – 25] Such architectures find application in the development of novel DNA probes for site and shape selective interactions[26, 27] and for spectroscopic studies of DNA-mediated electron transfer, as the intercalator represents another base pair simply inserted within the stack.[28] Octahedral $d^6$ and square planar $d^8$ metal ions provide an ideal scaffold for the construction of metal-based DNA binding agents due to their well defined coordination geometries and distinctive electrochemical or photophysical properties.[25]

In 1978, Lippard and co-workers showed that positively charged platinum complexes with planar aromatic ligands (see figure 1.7) could intercalate between adjacent base pairs of the DNA.[29] Electrochemical and photophysical measurements of $[Ru(bipy)_2(dppz)]^{2+}$ carried out by Friedman *et al.* in 1990 showed that upon addition of DNA to aqueous solutions of $[Ru(bipy)_2(dppz)]^{2+}$ intense luminescence occurred, indicating that the ring nitrogens of the dppz ligands were shielded from the solvent via

intercalation into the DNA base pair stack. These properties lead to what has been termed the 'light switch' effect.[30, 31]



(a) (b) (c) (d)

**Figure 1.7 Structure of (a) bipy (b) dppz (c) phen and (d) phi ligands.**

Spectroscopic studies on a rhodium complex of phi have shown that the phi ligand intercalates with the long axis parallel to the long axis of the base pair increasing the binding affinity, $K_b$ to $10^7$ $M^{-1}$.[32]

### 1.3.2 External association

A molecule is said to be externally associated with DNA if it binds to the negatively charged phosphate backbone or interacts with the electron donor groups of the bases (see figure 1.8).[19, 33] This external association can be due to electrostatic or covalent interactions. The factors that influence this mode of binding are predominantly electrostatic interactions and include the charge of the molecule, the ligand hydrophobicity and the size of the complex ion.[20] Manning defined the two ligand binding modes to a charged species as 1. site specific binding and 2. territorial binding where territorial binding involves the metal complex binding to the whole polyanion by coulombic interactions without associating with any specific sites.[34, 35] In 1986, Nordén *et al.* showed the presence of two different DNA-binding modes for the metal complexes $[Fe(bipy)_3]^{2+}$ and $[Ru(bipy)_3]^{2+}$ using circular and linear dichroism (see figure 1.7 for ligand structure). The strongest DNA-binding mode at low salt concentration (<0.2 M NaCl) was dominated by electrostatic association with the DNA backbone.[36] Such a binding for the bimetallo cylinders is easy to visualise due to the cationic/electrostatic interaction of the metal complex with the negatively charged phosphate backbone of the DNA helix.

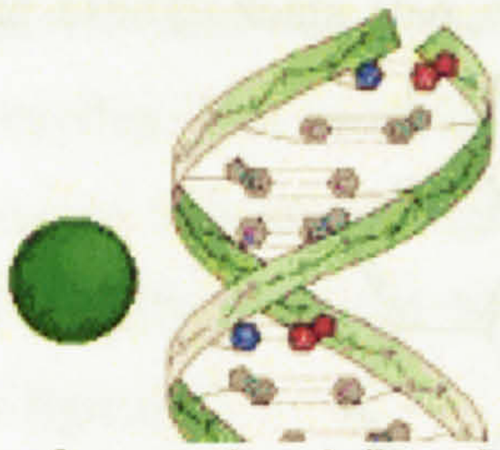


**Figure 1.8 Schematic diagram of a molecule/ion binding to DNA in an external associative manner.[19]**

### 1.3.3 Major and minor groove binding

Small groove binding molecules typically bind to B-DNA in or via the minor groove of the double helix, while proteins or gene-targeted oligonucleotides often interact with the major groove (see figure 1.9).[14, 19] These larger molecules bind in the major groove because it presents a better receptor in terms of its size, flexibility, electrostatic potential, water binding properties and H–bonding patterns.

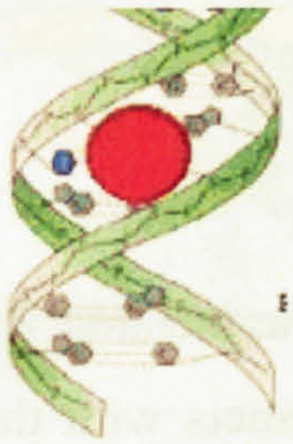


**Figure 1.9 Diagram of a molecule binding in the major groove of DNA.[19]**

The major groove represents a target for sequence-specific interactions and is rich in chemical information. The edge of an AT base pair in the major groove (see figure 1.10) displays two H-bond acceptors (the N7 of adenine and C4 of thymine) a H-bond donor (C6 of adenine), and a bulky hydrophobic surface (the methyl group on C5 of thymine). Similarly the edge of a GC base pair in the major groove (see figure 1.10) displays two H-bond acceptors (N7 and C6 of guanine) a H-bond donor (C4 of cytosine) and a small non-polar hydrogen (the H at C5 of cytosine).[4]

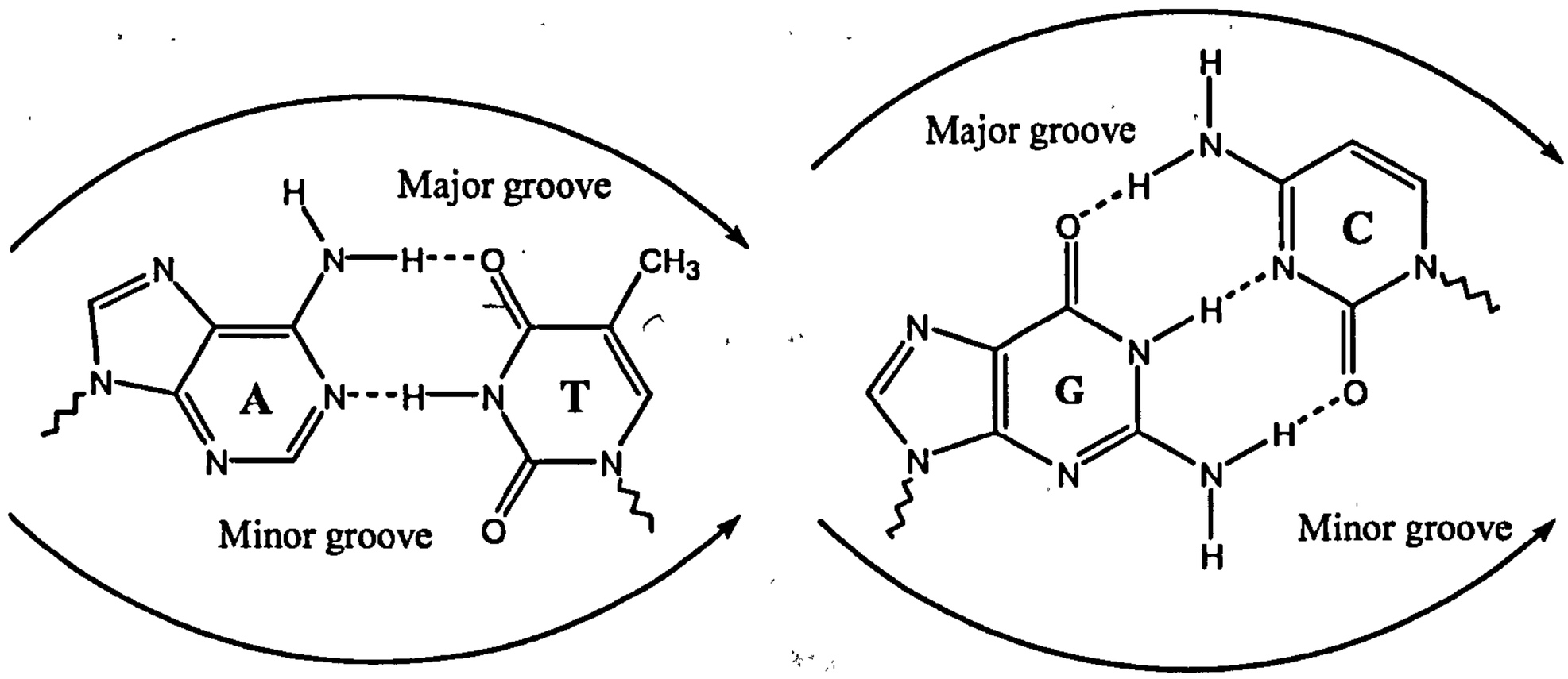


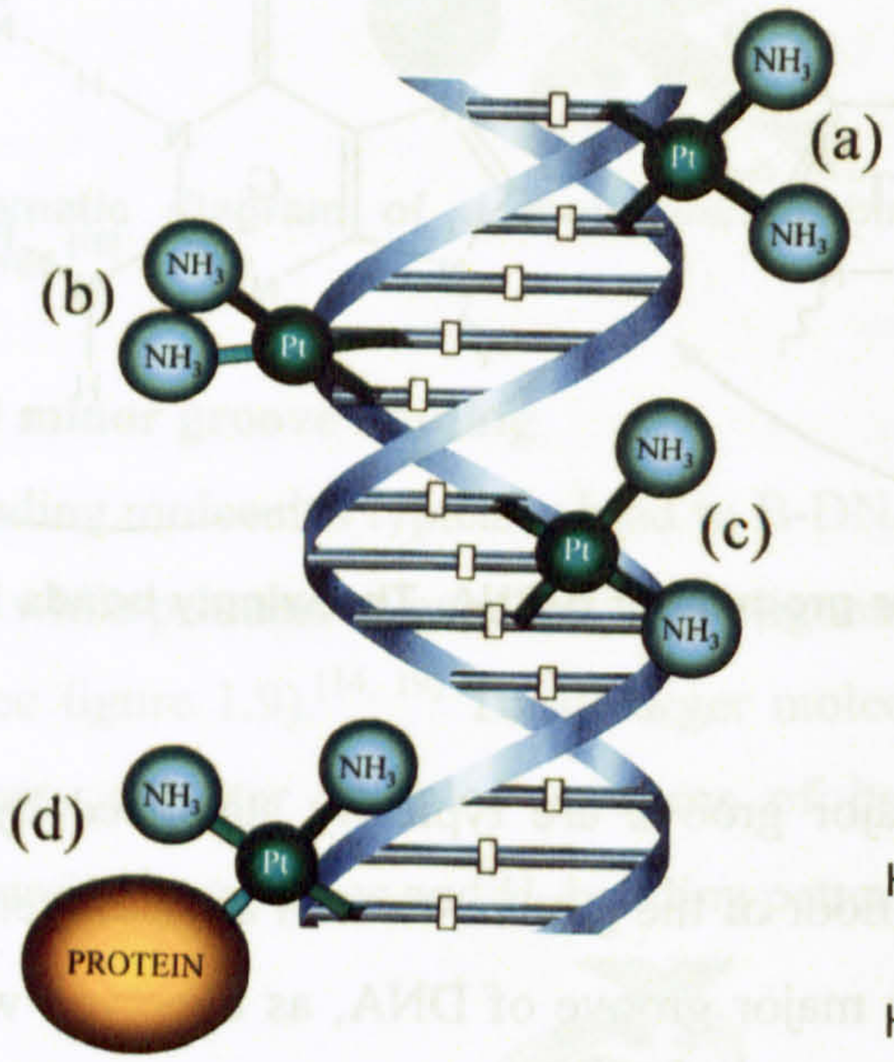

**Figure 1.10 The major and minor grooves of B-DNA. The wiggly bonds indicate the links to the sugar backbone.[4]**

Molecules that bind in the major groove are typically stabilised by van der Waals interactions with the walls and floor of the groove, as well as hydrogen bonds.[37] Some small molecules do bind in the major groove of DNA, as observed with the platinum anti-cancer drug cisplatin where a covalent bond to the N7 of G or A is found. Groove binding interactions involve the convex curvature of the inner surface of the molecule complementing the concave surface of the floor of the DNA groove. The surface matching concept was termed isohelicity and has been found to be a useful tool in the design of novel minor groove-binder agents.[38, 39] Small molecules typically bind in the minor groove of DNA where sequence recognition is limited compared to the major groove, in large part due to less varied H–bonding patterns and more limited shape variations with sequence.

### 1.3.4 Covalent binding of metal complexes

Cisplatin (cis-diamminedichloroplatinum (II)), a major anti-cancer agent, was first synthesized in 1844 by Peyrone,[40] its structure (see figure 1.11) was later elucidated by Werner in 1893.[41] In the early 1960's, Rosenberg at the Michigan State University identified cisplatin as a highly effective agent in eliminating tumours.[42] Cisplatin is believed to kill cancer cells by binding to DNA and interfering with its repair mechanism, eventually leading to cell death. In 1978, Mansy *et al.* established that cisplatin binds with high selectivity to purine bases and furthermore that it showed a preferential binding with guanine at the N7 position inducing extensive unstacking of the bases.[43] Cisplatin has biochemical properties similar to that of bifunctional alkylating agents, producing interstrand, intrastrand and monofunctional adduct cross-linking in DNA, (see figure 1.11).[44, 45] The most prevalent form of binding is the 1,2-

intrastrand crosslink. In this adduct, the platinum is covalently bound to the N7 position of adjacent purine bases. In the 1,2-intrastrand crosslink, the DNA is unwound and bent towards the major groove.[44]



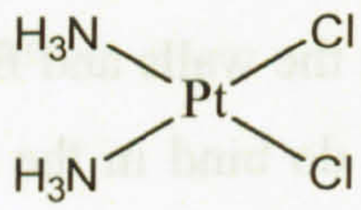




**Figure 1.11 Structure of cisplatin and cisplatin DNA-binding modes (a) interstrand, (b) 1,2-intrastrand, (c) 1,3-intrastrand and (d) DNA-protein cross link.[45]**

While cisplatin has been used successfully in the treatment of cancer patients since its introduction in 1979, it does produce toxic side effects and certain tumours can develop resistance to the drug. Second generation platinum drugs such as carboplatin and oxaliplatin have been developed (see figure 1.12). Other developments involve polynuclear platinum drugs and transition metals other than platinum such as ruthenium.[46 – 49] The major challenge still remains in developing drugs that bind to DNA in a sequence-specific manner, are efficient at low doses, have few side effects and are able to be excreted from the body after being used. Any molecules that fulfil these requirements and bind with high affinity to any predetermined DNA sequence in the human genome would be useful tools for molecular biology and potentially for human medicine.[50]

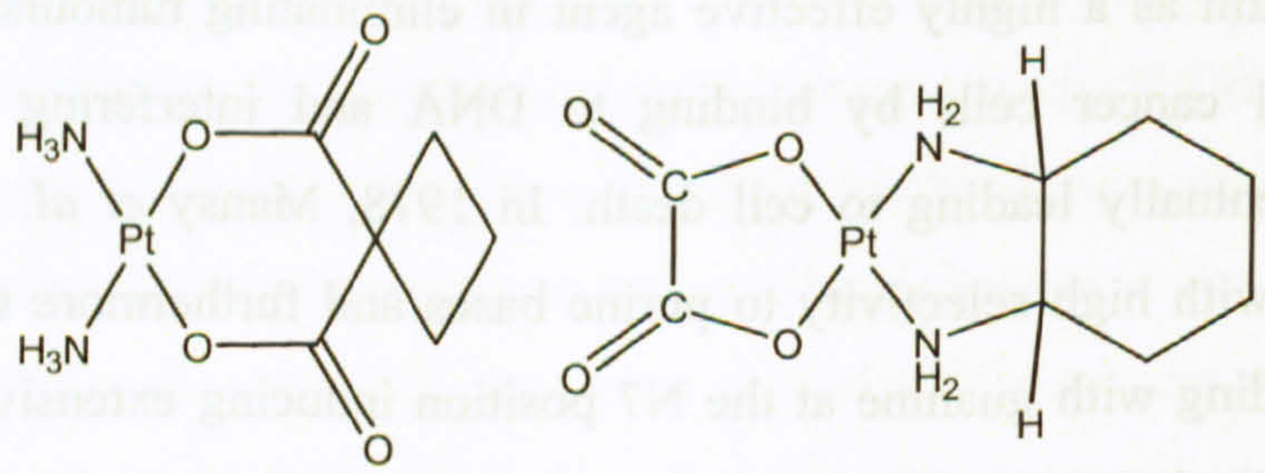



**Figure 1.12 Structure of carboplatin and oxaliplatin.[51]**

### 1.3.5 Non-covalent binding of metal complexes

In many ways a non-covalent binder, where a greater range of sequence selectivity can be programmed in, is very attractive. The interactions of metal complexes with DNA and direct binding of metals with DNA base pairs has great importance in the development of novel pharmaceuticals.[52 - 54] Small molecules are stabilised when bound to DNA through a series of weak interactions such as the π-stacking interactions associated with intercalation and hydrogen bonding and van der Waals interactions of functionalities bound along the groove of the DNA helix. The metal centres of metal complexes can impart a cationic charge to the ligand which provides a significant energetic contribution to the non-covalent binding to DNA. Electrostatic interactions arise from the ability of the charged group to attract to an oppositely charged group. π-stacking interactions occur between aromatic rings and exist as two main types: face-to-face where the two aromatic rings are 3.5 Å apart and face-to-edge where the centroids are ~ 5 Å apart.[55] Hydrogen bonding occurs when a δ positive hydrogen atom attached to an electronegative atom (or electron-withdrawing group) is attracted by a hydrogen acceptor (or electronegative) atom. Hydrogen bonds play an important role in biological systems, being responsible, for example, for the double stranded structure of DNA. Van der Waals bonds are considerably weaker non-specific, non-directional forces of attraction. They result from the polarisation of an electron cloud by the proximity of an adjacent nucleus and are referred to as induced dipole-induced dipole interactions. In general terms, electrostatic attraction provides non-specific attraction, however, often the binding Gibbs free energy is dominated by an entropy increase between cationic metal complexes and DNA due to displacement of water upon binding of the metal complex.[56]

Ligands that have been extensively studied include phen (1,10-phenanthroline) and phi (9,10-phenanthrenequinone) (see figure 1.7 for ligand structure). The metal complex $[Ru(phen)_3]^{2+}$ exists as two isomers: Λ-$[Ru(phen)_3]^{2+}$ (the left handed isomer) which binds into the major groove of B-DNA with a single phen ligand approximately parallel to the base pair planes and Δ-$[Ru(phen)_3]^{2+}$ (the right handed isomer) which binds with two phen ligands inserted into the minor groove at low binding ratios, and via insertion of a single phen ligand into both the minor and major grooves at higher complex loading (see figure 1.13).[57] This work also showed that methyl substituents on the ligands did affect binding strength and sequence selectivities. An extreme example is

substitution at the 5,6-position on the phenanthroline which switches off groove binding.[57]

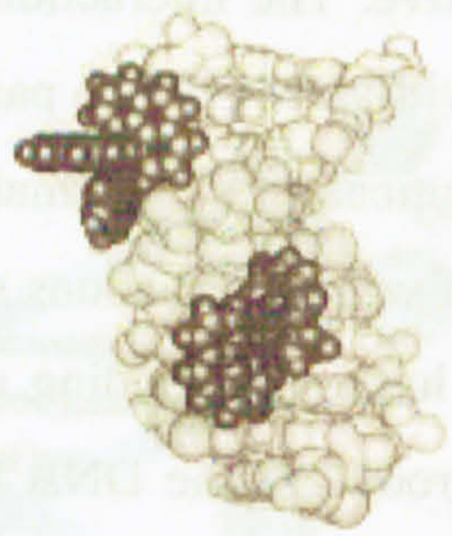


**Figure 1.13 Non-covalent binding of octahedral $[Ru(phen)_3]^{2+}$ with DNA, Δ-$[Ru(phen)_3]^{2+}$ intercalated into the major groove and Λ-$[Ru(phen)_3]^{2+}$ surface bound against the minor groove.[54]**

## 1.4 Zinc finger proteins

The class of protein motifs, zinc fingers, bind specifically to a very wide range of DNA sequences and have many roles. Some of them also bind non-specifically. As such they should provide templates for sequence selectivity. In 1985, Aaron Klug and his colleagues at the MRC Laboratory of Molecular Biology in Cambridge, UK described the first classic zinc finger protein motif.[58] It was observed from the analysis of an amino acid sequence of the transcription factor TFIIIA that the motif consisted of nine repeated sequences of approximately thirty residues each, with two invariant pairs of cysteines and histidines, which bind to a zinc (II) ion.[59] A tetrahedrally co-ordinated finger structure is formed by the independent folding of the repeated sequences around the central zinc atom (figure 1.14).[60] The finger region of this protein motif is so named due to the linker regions' (which is the region between the last cysteine and the first histidine residue) similarity to a finger. Zinc finger motifs are small independently folded protein domains that play a vital role in regulating a remarkable array of biological functions. They can be described as 'mini globular proteins' which have a hydrophobic core and polar side chains on the surface.[60]

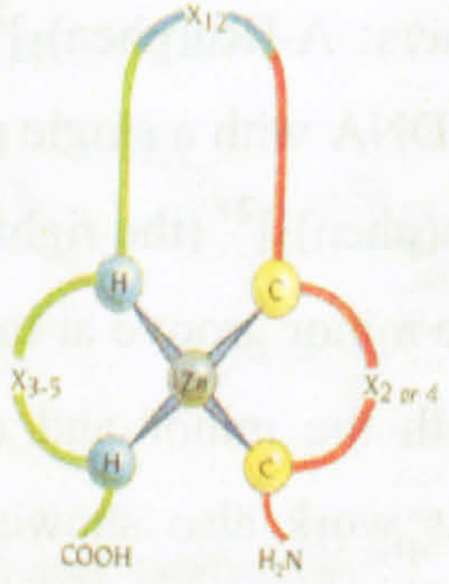



**Figure 1.14 Schematic diagram of the classic zinc finger motif.[59]**

Zinc fingers are DNA-binding motifs of proteins that provide a three-dimensional scaffold that matches the contours of DNA. This allows sequence specific and non-specific interactions to occur.[59] Zinc finger proteins are among the most abundant proteins in the eukaryotic genome.[60] Due to their size and structure they have a diverse range of functions including: RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, lipid binding and DNA recognition.

### 1.4.1 DNA sequence recognition

DNA recognition is important because this is the process by which proteins that regulate transcription of DNA recognise specific DNA sequences through discrete DNA-binding domains within the polypeptide chains. Zinc finger proteins recognize a diverse set of DNA sequences in a sequence specific manner and therefore offer an attractive framework for the design of novel DNA binding ligands.[61] They bind into the major groove of DNA, as they are usually too big to bind in the minor groove. They thus behave as recognition elements for the major groove and typical interactions involve hydrogen bonding between amino acid side-chains and the edges of the base pairs. One finger typically recognises three contiguous base pairs of a DNA sequence and is mediated through base contacts with the side chains of specific amino acids located on the recognition helix.[62] Protein recognition of DNA sequence motifs is thought to be a combination of the protein having affinity for DNA with a certain topology (local conformation and configuration) and its ability to form specific H-bonds with the bases in the sequence.

In 1993, Paveltich and Pabo obtained the first crystal structure of a zinc finger/DNA complex, showing each of the fingers' α-helix region bound into the major groove of DNA, with each of their $NH_2$-terminal groups closest to the bases.[63] However, it has been observed that only four amino acid positions on the α-helix provide critical contacts; three primary positions (–1, +3 and +6) and the auxiliary (+2) position, see figure 1.15.[64]

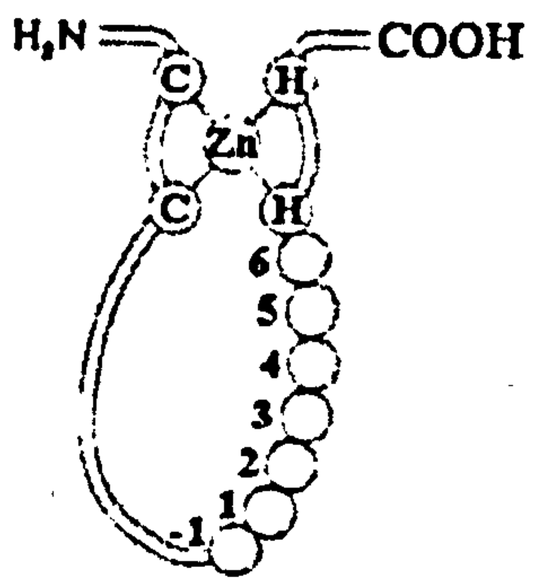



**Figure 1.15 Sequence specific binding positions on a zinc finger motif.[64]**

Choo and Klug found two factors that governed the zinc finger/DNA sequence selectivity: (i) the electronic properties (electron donor or acceptor) and the geometrical complementarities between the amino acid side chains of the α-helix and the accessible groups of the base pairs; and (ii) the base sequence, which changes the DNA conformation.[62, 65 – 67] From the perspective of the research outlined in this thesis the design and synthesis of a supramolecular cylinder that binds in the major groove of DNA appears to be the best option to achieve binding specificity. Binding specificity may be achieved by adding electron donor and acceptor groups along the backbone of the ligand.

## 1.5 Supramolecular chemistry

The field of supramolecular chemistry is involved with the self-assembly and organisation of small units which are programmed to fit together to form large complicated molecules. It is the chemistry of the intermolecular bond and concerns the structure and functions of the entities formed by the association of two or more chemical species.[68, 69] Transition metal ions such as iron (II), cobalt (II) and nickel (II) are the ideal foundation upon which to build functional supramolecular architectures due to their predictable geometry and lability.

The supramolecular self-assembly approach has distinct advantages to the synthetic chemist providing a powerful tool involving the non-covalent synthesis of a supramolecular architecture which positions the components, often followed by post-assembly modification through covalent bond formation.[69] The supramolecular synthetic approach has led to the discovery and assembly of an exciting range of species with complex architectures including, knots,[70] grids,[71] boxes[72 – 74] and cylinders.[75 – 78]

The fact that metal-ligand interactions occur in certain biological systems has led to the field of bioinorganic supramolecular chemistry and this area has created considerable interest in recent years. One of the best examples of a self-assembling biological supramolecular complex is that of haemoglobin[79] (figure 1.16). This protein-based species is composed of four separate components: two α and two β subunits, each are capable of reversibly binding oxygen. There is a multitude of precisely positioned non-covalent interactions that exist between the individual subunits in haemoglobin which control structure, function and oxygen-binding.[80]

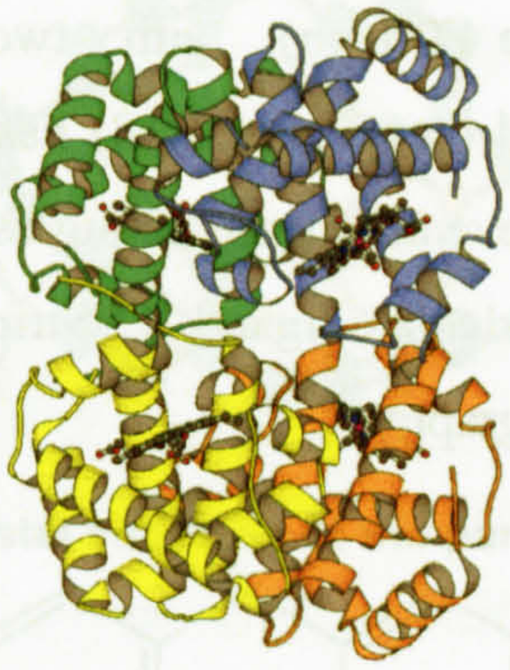


**Figure 1.16 Structure of haemoglobin.[79]**

## 1.5.1 Supramolecular helicates

The use of metal ions to control the self-assembly of topologically novel molecular structures has significant appeal.[81] The binding of metal ions with specific co-ordination geometry to organic ligands of poly(2,2'-bipyridyl) (bipy) nature sometimes results in the spontaneous organization of an inorganic double helix (see figure 1.17). Lehn and co-workers found the first of these and introduced the term *helicate* in 1987 for the description of a double-stranded metal containing helix.[82, 83] The word *helicate* is derived from the combination of *helix*, which denotes 'winding' in Greek and the suffix *–ate* which characterises host-guest complexes between (pre)organised receptors and metal ions.

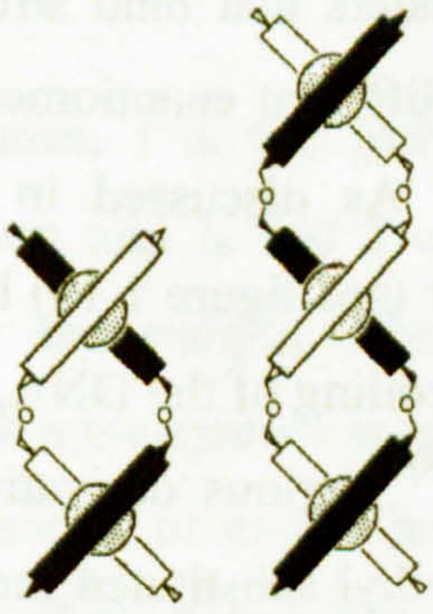


**Figure 1.17 Schematic representation of $[Cu_2(bipy_2)_2]^{2+}$ and $[Cu_3(bipy_3)_2]^{3+}$.[82]**

Helicates in more general terms are defined as species formed by the spontaneous assembly of one or more organic ligands about a metal centre. A discrete linear dinuclear oligomer is formed when two or three ligand strands bridge between a pair of metal centres, to give the corresponding double or triple stranded helicate respectively. In 1958, Busch *et al.* synthesised the first triple-stranded helicate by adding pyridinaldazine to an aqueous solution of metal (iron (II) or nickel (II)) salt.[84] The triple-stranded architecture was not structurally characterised until 30 years later when Williams and co-workers mixed three equivalents of bis[1-methyl-2-(6'-methyl-2'-pridyl)benzimidazl-5-yl]methane (Mepb) with two equivalents of cobalt (II) perchlorate.[85] The Mepb ligand (see figure 1.18) has four co-ordination sites but the rigidity of the aromatic rings prevents it from acting as a tetradentate ligand and makes it ideally suited to act as a bis-bidentate ligand. The triple-stranded helical structure was confirmed using X-ray crystallography.

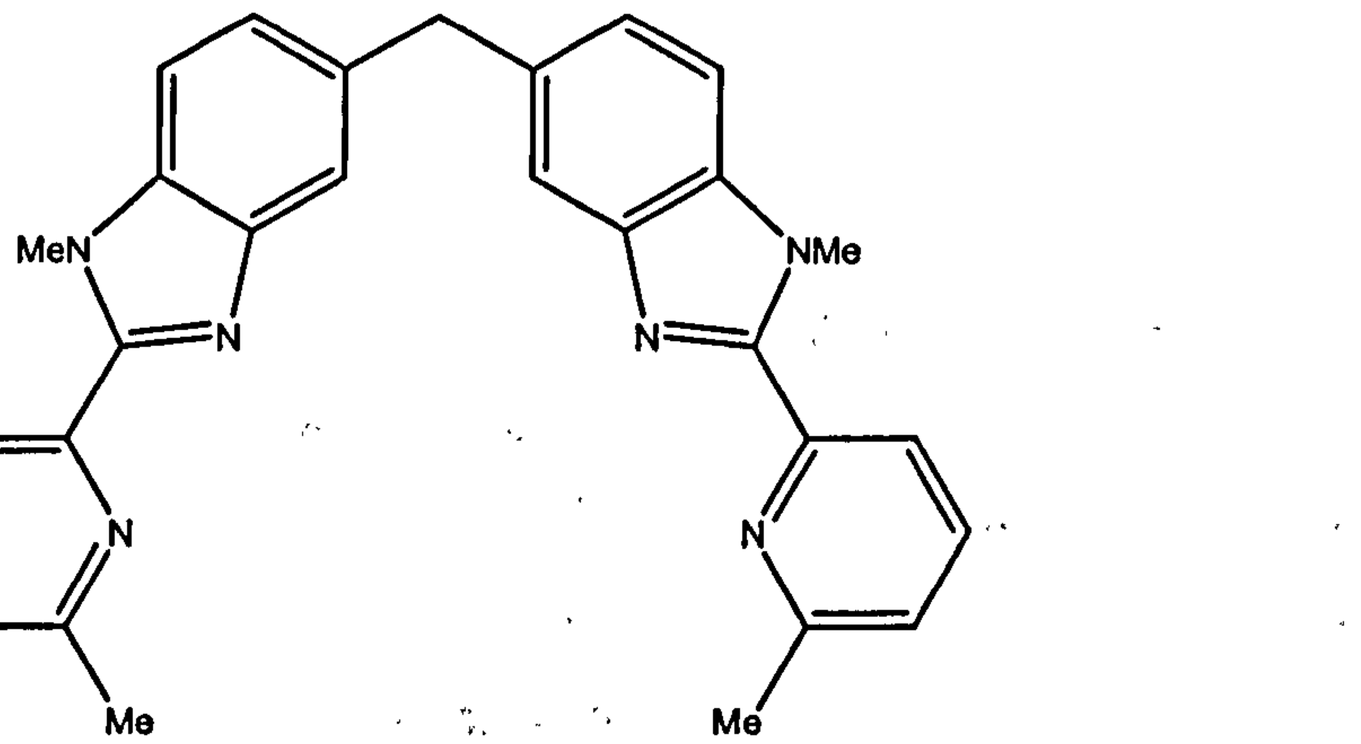



**Figure 1.18 Structure of Mepb ligand.[85]**

The supramolecular synthetic approach led to the discovery at the University of Warwick, of a range of triple helicate bimetallo iron compounds (see figure 1.19). These compounds are tetracationic cylinders that bind strongly to DNA. The cylinders are helical and therefore chiral. The different enantiomers can be separated on a cellulose column using 0.02 M NaCl.[86] As discussed in more detail in chapter 3.0 the enantiomers of the parent cylinder (see figure 1.19) bind to B-DNA independently and differently inducing bending and coiling of the DNA, such effects are unprecedented in synthetic DNA binding agents.[78] Various derivatives show different DNA-binding properties as illustrated by the methyl substituted molecules that form the basis of the work reported in chapters 3 and 4. The left handed, ΛΛ enantiomer will be referred to

as the (M)-enantiomer and the right handed, ΔΔ enantiomer will be referred to as the (P)- enantiomer throughout this thesis, due to the sign of the *CD* signal at long wavelength and ease of use.

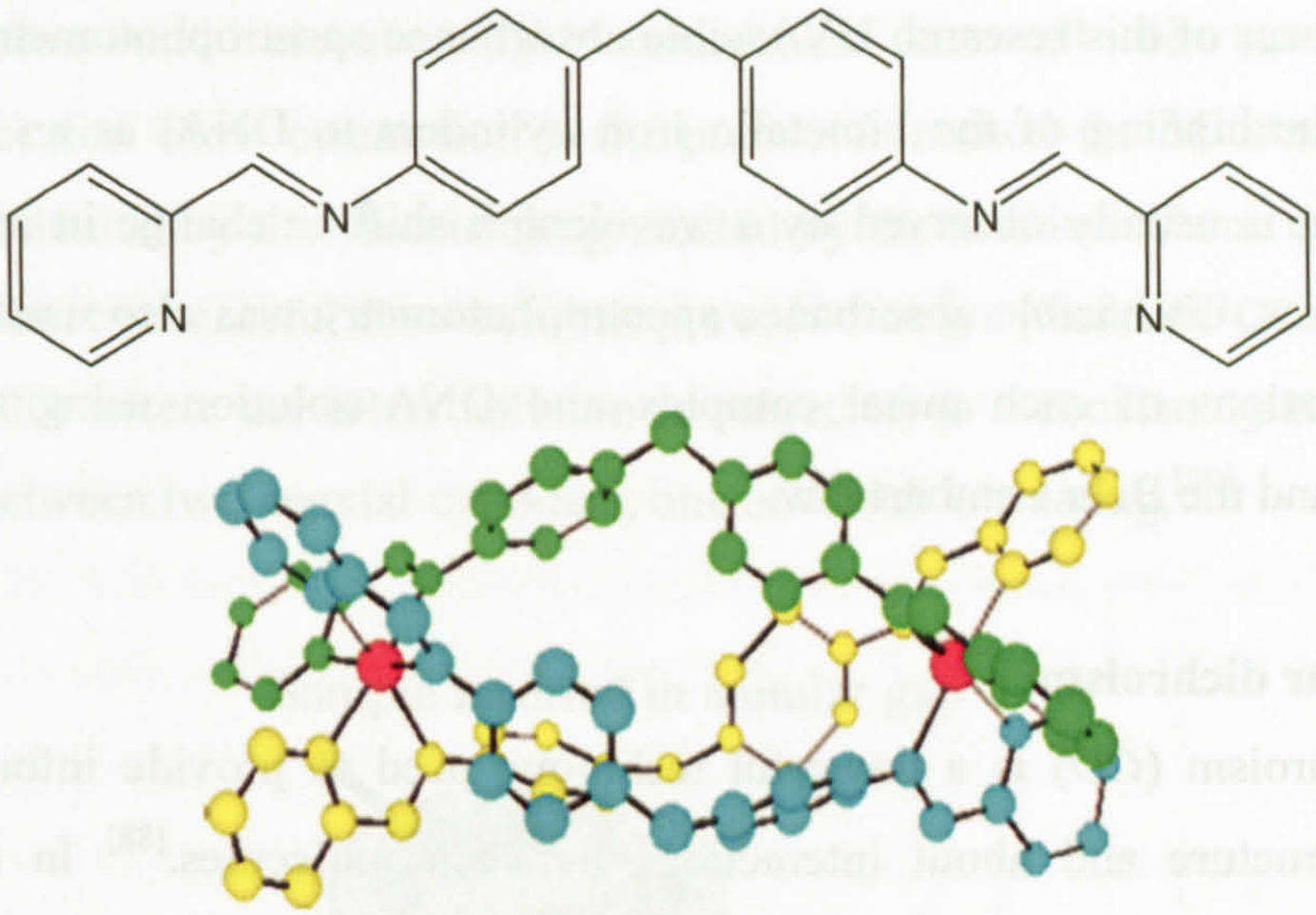


**Figure 1.19 Ligand and X-ray crystal structure of the parent cylinder, $L_1$.[78]**

## 1.6 Techniques

A range of analytical techniques including UV/visible absorbance spectrophotometry, circular and linear dichroism, fluorescence and molecular modelling, have been employed to investigate various biomacromolecule-ligand interactions. A brief description of each is outlined below.

### 1.6.1 UV/visible absorbance spectroscopy

Absorption measurements based upon ultraviolet or visible radiation find widespread application for the qualitative and quantitative analysis of molecular species.[87] The data are often analysed according to the Beer Lambert Law:

$$A = \log I_o/I = \varepsilon lc$$

Where A is the optical absorbance, *l* is the path length of light, ε is the molar absorptivity, c is the concentration and $I_0$ and I are respectively the light intensity incident on and transmitted by the sample. The two principles that govern the absorption of light passing through a transparent solution are that the absorption of light is exponentially related to the number of molecules of the absorbent in solution and exponentially related to the length of the light pathway through the sample.[88] DNA has a characteristic absorption spectrum with a maximum observed at ~260 nm and a

minimum at ~230 nm. The concentration of the DNA can then be determined and calculated at 260 nm using the Beer Lambert Law as outlined above.

For the purposes of this research UV/visible absorbance spectrophotometry was used to investigate the binding of the bimetallo iron cylinders to DNA, as an interaction or binding event is usually observed by a wavelength shift or change in intensity of the binding species. UV/visible absorbance spectrophotometry was also used to determine the concentrations of each metal complex and DNA solution using the extinction coefficients and the Beer Lambert law.

### 1.6.2 Circular dichroism

Circular dichroism (*CD*) is a powerful technique used to provide information about molecular structure and about interactions between molecules.[88] In this research, circular dichroism has been used to determine whether the pure enantiomers of the triple helical-metal complexes have been resolved and to investigate the binding of two methylated cylinders to DNA. Circular dichroism is defined as the difference in absorption, A, of left and right circularly polarised light:[88]

$$CD\,(\lambda) = A_l\,(\lambda) - A_r\,(\lambda)$$

*CD* of solutions is sensitive to chiral molecules and because chiral molecules do not have a reflection plane, any rearrangement of the electrons will not have one either, therefore the interaction between the chiral molecule and the left- and right-handed photons will be different, giving rise to a *CD* signal. When a ligand binds to DNA (even if it is achiral or racemic) it acquires an induced *CD* (*ICD*) that is characteristic of the interaction.[89] If an *ICD* signal is observed upon addition of the metal complex to a DNA solution it can be concluded that the metal complex does bind to DNA.

### 1.6.3 Linear Dichroism

Linear dichroism (*LD*) is a technique used with systems that are either intrinsically oriented or are oriented during the experiment. It is used to probe electronic transitions within a molecule when it absorbs a photon of radiation of the correct energy to excite it from its ground state. *LD* is the difference in absorption of light linearly polarised parallel and perpendicular to an orientation axis at a given wavelength:[90]

$$LD(\lambda) = A_{//} - A_{\perp}$$

Long polymers such as DNA can be orientated using various methods; stretched film, flow orientation and electric field orientation can all be used to orient biomacromolecules. Flow orientation is the preferred method for *LD* studies of DNA. The DNA is oriented by the viscous drag created by the rotation of one cylinder inside another. The most successful flow cell has proved to be a cylindrical Couette flow cell (see figure 1.20) where the DNA solution is subjected to a constant gradient over the annular gap between two coaxial cylinders, one of which is rotating.[90]

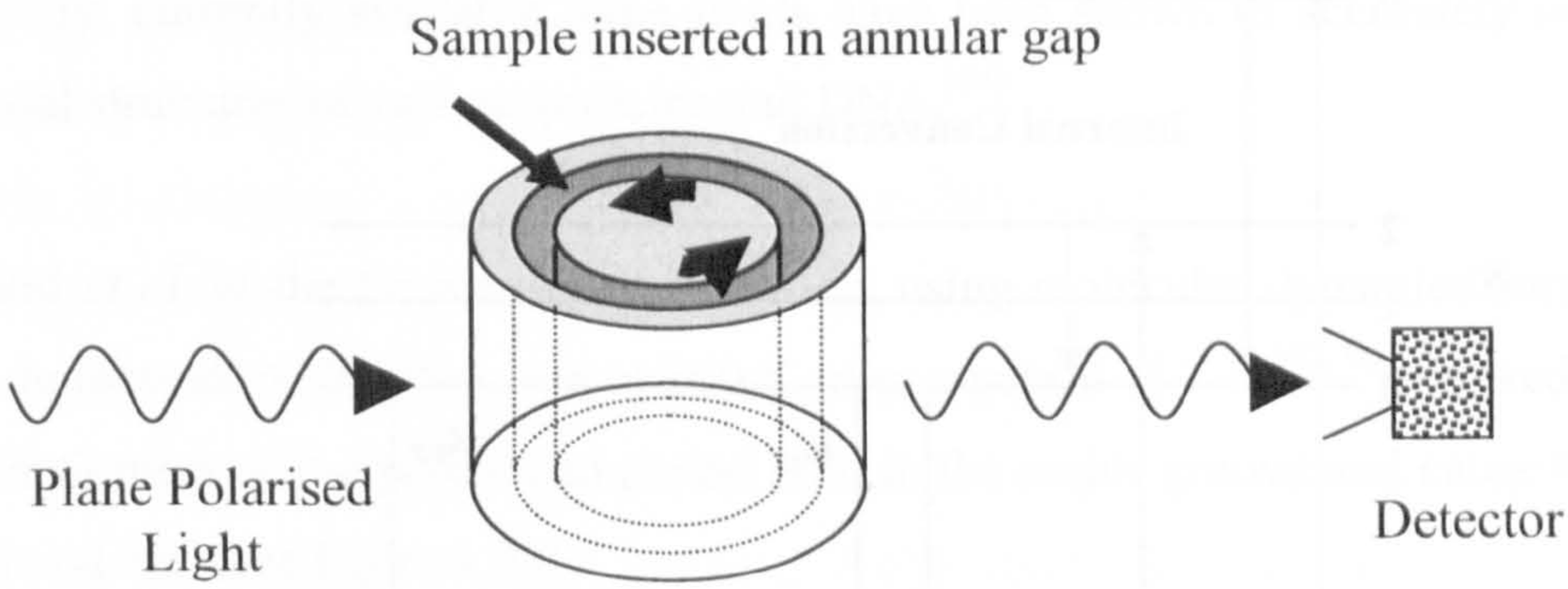



**Figure 1.20 Schematic diagram of couette flow cell.**[91]

In this research, *LD* was used (see chapter 3) to probe the binding of two methylated bimetallo iron cylinders to DNA and to determine the binding constant of ethidium bromide to ct-DNA at 50 mM NaCl. DNA has a characteristic *LD* spectrum, with no transitions between 750 – 300 nm and a large absorbance band at approximately 260 nm due to the $\pi$-$\pi^*$ transitions and stacking of the base pairs. The metal complexes show no intrinsic *LD*, but upon binding to the DNA, show an *LD* signal in the metal to ligand charge transfer (MLCT) region of the spectrum and in the in-ligand region of the spectrum just above 300 nm. It was shown previously that for both low and high loadings of the parent cylinder on DNA, an *LD* signal was observed in the MLCT and in-ligand regions of the spectrum, confirming its DNA-binding is not random. Addition of the parent cylinder to ct-DNA, even at very low loadings caused a significant reduction in the *LD* signal intensity at 260 nm which was consistent with shortening of the DNA through bending.[76]

### 1.6.4 Fluorescence spectrophotometry

The emission of photons from electronically excited states in any substance is called luminescence. Fluorescence is the emission which results from the return of the molecule to its ground state.[92] The photophysical processes that occur from absorption to emission are often summarised in a Jabloński diagram, (see figure 1.21). Three electronic levels are depicted along with three vibrational energy levels, $h\upsilon_A$ and $h\upsilon_F$ denote absorption and fluorescence respectively, $k_r$ is the rate constant for fluorescence and $k_{nr}$ is the rate constant for the competing non-radiative route.

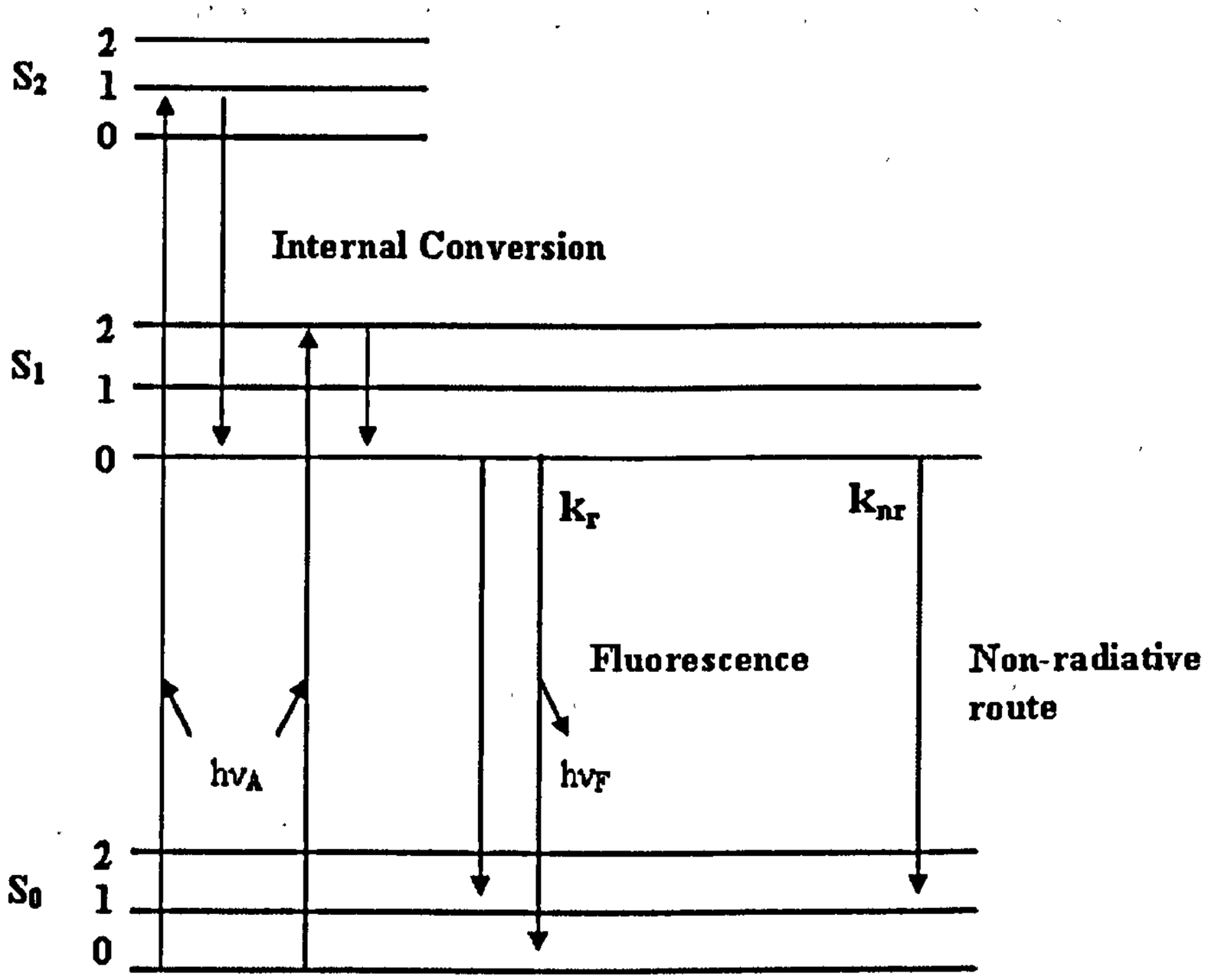



**Figure 1.21 Jabloński diagram.**[92]

Fluorescence spectrophotometry was used in this work to determine binding constants for the non-fluorescent parent cylinder and two methylated bimetallo iron cylinders in a competitive binding assay with the fluorescent dye ethidium bromide. It was also used to investigate the competition between the bimetallo iron cylinders and the DNA minor groove binder Hoechst 33258.

### 1.6.5 Molecular Dynamics simulations

Molecular simulations can broadly be defined as the process by which the nature of a chemical process is simulated. In this thesis molecular dynamics simulations were employed to investigate glucose–cylinder and cylinder–DNA interactions. Molecular dynamics simulations proceed by applying Newton's equation of motion to generate

successive configurations of a system, this is called a trajectory.[93] The chemical nature of the different atom types is described by the force field. Controllable parameters in a simulation include the temperature, pressure, volume, and of course length of simulation. The advantages of simulation techniques over experimental methods are that simulations can be performed under extreme conditions such as high temperature or pressure, these conditions are often difficult to achieve experimentally. One drawback of simulations is that the length of the simulation performed is dependent upon the available computational resources. Currently simulations are performed on a nanosecond timescale which can be too short to observe certain biological phenomena. One other drawback is that the accuracy of the simulation depends on the force field. Fortunately, currently available force fields have been shown to accurately reproduce the crystal structures of various proteins and DNA.[94]

S. Khalid *et al.* at the University of Warwick, using molecular dynamics simulations, with a dodecamer of B-DNA, run at 300 K over a period of 2 ns[95, 96] showed that the (M)-enantiomers of the parent compound bind in the major groove and cause the DNA to bend and coil (see figure 1.22).

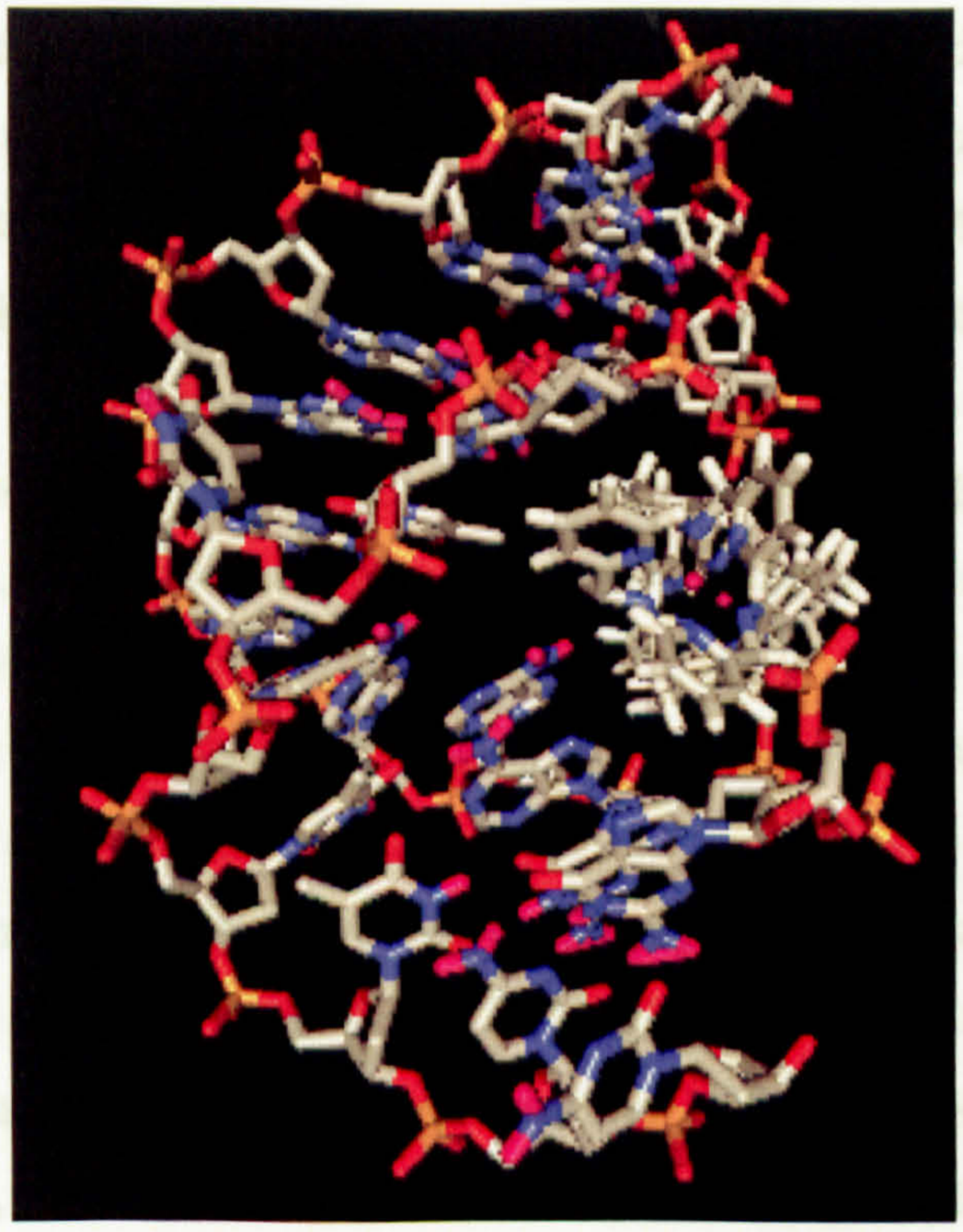


**Figure 1.22 (M)-$[Fe_2(L^1)_3]^{4+}$ binding in the major groove of B-DNA.**[96]

The interactions of the enantiomers of the parent compound with a glucose monomer were investigated using high temperature docking calculations. It was hoped that a

difference in configurational energy with respect to binding of each of the enantiomers to the glucose monomer would be observed. Experimentally the (M)-enantiomers of the bimetallo cylinders elute first from the column suggesting a lower energy of interaction and it was hoped that the same trend would be observed in the molecular simulations. Similar high temperature docking calculations were also used to investigate the binding of the enantiomers of two methylated bimetallo iron cylinders to a B-DNA dodecamer.

## 1.7 Objectives

The primary objectives of the work reported in this thesis were to investigate the binding of chiral bimetallo iron complexes to DNA, with the aim of determining whether substituents on the ligand backbone of the parent cylinder gave sequence selectivity to the DNA interactions (chapter 3). To achieve this it was necessary to separate the enantiomers of the various cylinders studied. A method was developed to efficiently purify and enantiomerically separate the parent cylinder and a large number of derivatives with different surface topologies (chapter 2). The DNA binding studies focussed on differences between the resolved (M)- and (P)-enantiomers. A big challenge is to determine the binding site of the (M)- and (P)-enantiomers of cylinders on DNA when no crystals can be grown. This is the focus of the work reported in chapter 4 where experimental and molecular modelling results are combined. The structures of the RNA-binding domain fragment of nucleolin-RBD and a range of guanosine-rich oligonucleotides form the background to a study of the interaction between a phase I clinical trial guanosine rich oligonucleotide and nucleolin using circular dichroism (chapter 5).

## 1.8 References

1. Stryer, L., *'Biochemistry'* Fourth Ed. New York: Freeman, **1995** and references therein

2. Sarma, R. H., '*Nucleic acid geometry and dynamics*' Pergamon Press, **1980** and references therein

3. Adams, R. L. P., Knowler, J. T. and Leadrer, D. P., '*The Biochemistry of the nucleic acid*'. Chapman and Hall, **1992** and references therein

4. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V. and Kopka, M. L., *Science*, **1982**: 216, 475 – 485

5. Guschlbauer, W., '*Nucleic Acid Structure*' Springer-Verlag, New York Inc. **1976** and references therein

6. Smith, E. L., Hill, R. L., Lehman, I. R., Handler, R. J. and White, A., '*Principles of Biochemistry – General Aspects*' Seventh Edition, McGraw-Hill Inc. **1983** and references therein

7. Wang, A. H., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom J. H., van der Marel, G. and Rich, A., *Nature*, **1979**; 13, 282, 680 – 686

8. Pohl, F. M. and Jovin, T. M., *J. Mol. Biol.* **1972**; 67, 375 – 396

9. Crawford, J. L., Kolpak, F. J., Wang, A. H., Quigley, G. J., van Boom, J. H., van der Marel, G. and Rich, A., *Proceedings of the National Academy of Sciences, USA* **1980**; 77, 7, 4016 – 4020

10. Rich, A. and Zhang, S., *Nature Reviews, Genetics*, **2003**; 4, 566 – 572

11. Watson, J. D. and Crick, F. H. C., *Nature*, **1953**; 171, 737 – 738

12. Chargaff, E., *Experientia* **1950**; 6, 201 – 209

13. Neidle, S., *Progress in Medicinal Chemistry*, **1979**; 16, 152 – 221

14. Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H. and Waring, M. J. '*The Molecular Basis of Antibiotic Action*' John Wiley, London, **1972** and references therein

15. Corcoran, J. W. and Hahn, F. E. (Eds.) '*Antibiotics III Mechanism of action of Anti-microbial and Anti-tumour agents*' Springer Verlag, Berlin, **1974** and references therein

16. Drake, J. W. and Baltz, R. H. *Annu. Rev. Biochem.*, **1976**; 45, 11 – 37

17. Heidelberger, C., *Annu. Rev. Biochem.*, **1975**; 44, 79 – 121

18. Mainwaring, W. I. P., Parish, J. H., Pickering, J. D. and Mann, N. H., '*Nucleic Acid Biochemistry and Molecular Biology*' Blackwell Scientific Publications, **1982** and references therein

19. http://www.photobiology.com/photoiupac2000/pierard/Interactionmain.html

20. Meistermann, I., '*DNA major groove recognition by supramolecular helicates*', Ph.D. thesis, University of Warwick, **2001**; 31 – 34

21. Lerman, L. S., *Proc. Natl. Acad. Sci. USA*, **1963**; 49, 94 – 102

22. Waring, M., *J. Mol. Biol.* **1970**; 54, 247 – 279

23. Johnson, D. S. and Boger, D. L., *'Binding Agents in Comprehensive Suprmolecular Chemistry'* (Atwood, J. L., Davies, J. E. D., MacNicol, D. D., and Vogtle, F., Eds.) **1996**; Vol. 4, Pergamon, Oxford.

24. Phillips, T., Haq, I., Meijer, A. J. H. M., Adams, H., Soutar, I. Swanson, L., Sykes, M. J. and Thomas, J. A., *Biochemistry*, **2004**; 43, 13657 – 13665

25. Metcalfe, C. and Thomas, J. A., *Chem. Soc. Rev.* **2003**; 32, 215 – 224

26. Metcalfe, C., Webb, M. Thomas, J. A., *Chem. Commun.* **2002**; 18, 2026 – 2027

27. Metcalfe, C., Adams, H., Haq, I. and Thomas, J. A., *Chem. Commun.*, **2003**; 10, 1152 – 1153

28. Núñez. M. E. and Braton, J. K., *Current Opinion in Chemical Biology*, **2000**; 4, 199 – 206

29. Lippard, S. J., *Accounts of Chemical Research*, **1978**; 11, 211 – 217

30. Friedman, A. E., Chambron, J.-C., Sauvage, J.-P., Turro, N. J. and Barton J. K., *J. Am. Chem. Soc.*, **1990**; 112, 4960 – 4962

31. Hiort, C., Lincoln, P. and Nordén B. *J. Am. Chem. Soc.*, **1993**; 115, 3448 – 3454

32. Sitlani, A., Long, E. C., Pyle, A. M. and Barton, J. K., *J. Am. Chem. Soc.* **1992**; 114, 2303 – 2312

33. Eichorn, G. L. and Shin, Y. A., *J. Am. Chem. Soc. USA*, **1988**; 85, 99 – 102

34. Manning, G. S., *J. Am. Chem. Soc.*, **1979**; 12, 443 – 449

35. Friedman, R. A. G. and Manning, G. S., *Biopolymers,* **1984**; 23, 2671 – 2714

36. Hard, T. and Nordén, B., *Biopolymers*, **1986**; 25, 1209 – 1228

37. Neidle, S., '*DNA Structure and Recognition*' Oxford University Press, **1994** and references therein

38. Singh, M. P., Plouvier, B., Hill, G. C., Gueck, J., Pon, R. T. and Lown, J. W., *J. Am. Chem. Soc.* **1994**; 116, 7006 – 2020

39. Goodsell, D. and Dickerson, R. E., *J. Med. Chem.* **1986**; 29, 727 – 733

40. Peyrone, M., *Liebigs Ann. Chem.***1844**; 51, 1

41. Werner, A., *Z. Anorg. Chem.*, **1893**; 3, 267

42. Rosenberg, B., Van Camp, L., Grimley, E. B. and Thomson, A. J., *J. Biol. Chem.* **1967**; 242, 6, 1347 – 1352

43. Mansy, S., Chu, G. Y. H., Duncan, R. E. and Tobias, S., *J. Am. Chem. Soc.* **1978**; 100, 607 – 616

44. Eastman, A., *Cisplatin, Chemistry and Biochemistry of a leading Anti-Cancer drug,* Wiley-VCH, Weinheim, Germany, **1999**; 113 – 134

45. Gonzalez, V. M., Feurtes, M. A., Alonso,C. and Perz J. M., *Mol. Pharmacol.* **2001**; 59, 4, 657 -663

46. Legendre, F and Chottard J.-C., *Cisplatin, Chemistry and Biochemistry of a leading Anti-Cancer drug,* Wiley-VCH, Weinheim, Germany, **1999**; 223 – 245

47. Manzotti, C., Pratesi, G., Menta, E., Di Domenico, R., Cavallettit, E., Fiebig, H. H., Kelland, L. R., Farrell, N., Polizzi, D., Supino, R., Pezzoni, G. and Zunino, F., *Clin. Cancer Res.* **2000**; 6, 2626 – 2634

48. Clarke, M. J., Zhu, F. and Frasca, D. R., *Chem. Rev.*, **1999**; 99, 2511 – 2533

49. Sava, G., Cappozzi, I., Clerici, K. Alessio, E., Mestroni, G., *Clin. Exp. Metastasis*, **1998**; 16, 371 – 379

50. Trauger, J. W., Baird, E. E., and Dervan P. B., *Nature*, **1996**; 382, 559 – 561

51. Rixe, O., Ortuzar, W. O., Alvarez, M., Parker, R., Reed, E., Paull, K. and Fojo, T., *Biochem. Pharmacol.* **1996**; 52, 1855 – 1865

52. Berman, H. M., Young, P. R., *Annu. Rev. Biophys. Bioeng.* **1981**; 10, 87 – 114

53. Wilson, W. D., Wang, Y.-H., Kusuma, S., Chandrasekaran, S., Yang, N. C. and Boykin, D. W., *J. Am. Chem. Soc.*, **1985**; 107, 4989 – 4995

54. Pyle, A. M., Rehmann, J. P. Meshoyrer, R., Kumar, C. V., Turro, N. J. and Barton, J. K., *J. Am. Chem. Soc.* **1989**; 111, 3051 – 3058

55. Steed, J. W. and Atwood, J. L., *'Supramolecular Chemistry'* John Wiley & Sons, Ltd. **2000** and references therein

56. Haq, I., Lincoln, P., Suh, D. C., Nordén, B., Chowdrhy, B. Z. and Chaires, J. B., *J. Am. Chem. Soc.*, **1995**; 117, 4788 – 4796

57. Coggan, D. Z. M., Haworth, I. S., Bates, P. J., Robinson, A. and Rodger, A., *Inorg. Chem.* **1999**; 38, 4486 – 4497

58. Miller, J., McLachlan, A. D. and Klug, A., *The EMBO Journal*, **1985**; 4, 1609 – 1614

59. Branden, C. and Tooze, J., *'Introduction to Protein Structure'* Second Ed. Garland Publishing **1999**; 176 – 184

60. Laity, J. H., Lee, B. M., and Wright, P. E., *Current Opinion in Structural Biology*, **2001**; 11, 39 – 49

61. Sugiura, Y., *RIKEN Review*, **2001**; 35, 102 – 104

62. Choo, Y. and Klug, A., *Proc. Natl. Acad. Sci. USA.*, **1994**; 91, 11163 – 11167

63. Paveltich, N. P. and Pabo, C. O., *Science* **1993**; 261, 1701 – 1707

64. Desjarleis, J. R. and Berg, J. M., *Proc. Natl. Acad. Sci. USA*, **1992**; 89, 7345 – 7349

65. Choo, Y. and Klug, A., *Curr. Opin. Struct. Biol.* **1997**; 7, 117 – 135

66. Choo, Y. and Isalan, M. *Curr. Opin. Struct. Biol.* **2000**; 10, 411 – 416

67. Klug, A., *J. Mol. Biol.* **1999**; 293, 215 – 218

68. Lehn, J.-M., *Angew. Chem. Int. Ed. Engl.* **1990**; 29, 1304 – 1319

69. Lehn, J.-M., *Proc. Natl. Acad. Sci.* **2002**; 99, 8, 4763 – 4768

70. Chambron, J. C., Dietrich-Buckecker C. O., Heitz, V., Nierengarten, J. J., Sauvage, J. P., Pascard, C. and Guilheim, J., *Pure. Appl. Chem.* **1995**; 67, 233 – 240

71. Lehn, J.-M., *'Supramolecular Chemistry, Concepts and Perspectives'*, VCH Weinheim, **1995** and references therein

72. Hannon, M. J., Painting, C. L. and Errington, W., *Chem. Commun.* **1997**; 307 – 308

73. Fujita, M. and Ogura, K., *Bull. Chem. Soc. Jpn.* **1995**; 69, 1471 – 1482

74. Stang, P. J. and Chen, K., *J. Am. Chem. Soc.* **1995**; 117, 1667 – 1668

75. Hannon, M. J., Painting, C. L. Jackson, A., Hamblin, J. and Errington, W., *Chem. Commun.* **1997**; 1807 – 1808

76. Hannon, M. J., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Meistermann, I., Isaac, C. J. Sanders, K. J. and Rodger, A. *Angew. Chem. Int. Ed.* **2001**; 40, 5, 880 – 884

77. Hannon, M. J., Painting, C. L. and Alcock, N. W. *Chem. Commun.* **1999**; 2023 – 2024

78. Meistermann, I., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Khalid, S., Rodger, P. M., Peberdy, J. C., Isaac, C. J., Rodger, A. and Hannon, M. J. *Proc. Natl. Acad. Sci. USA.* **2002**; 99, 8, 5069 – 5074

79. Lawrence, D. S., Jiang, T. and Levett, M., *Chem. Rev.* **1995**; 95, 2229 – 2260

80. Dickerson, R. E., Geis, I., '*Haemoglobin: Structure, Function, Evolution and Pathology*'.Benjamin Cummings Pub. Co. USA **1983** and references therein

81. Baxter, P. N. W., Lehn, J-M., Fischer, J. and Youinou, M-T., *Angew. Chem. Int. Ed. Engl.* **1994**; 33, 22, 2284 – 2287

82. Lehn, J.-M., Rigualt, A., Siegel, J., Harrowfield, J., Chevrier, B. and Moras, D., *Proc. Natl. Acad. Sci. USA,* **1987**; 84, 2565 – 2569

83. Saenger, W. '*Principles of Nucleic Acid Structure*', Springer New York, **1984**

84. Stratton, W. J. and Busch, D. H., *J. Am. Chem. Soc.* **1958**; 80, 3191 – 3195

85. Williams, A. F., Piquet, C., Bernardinelli, G., *Angew. Chem. Int. Ed. Engl.* **1991**; 30, 1490 – 1492

86. Hannon, M. J., Meistermann, I. Isaac, C., Blomme, C., Aldrich-Wright, J. and Rodger, A., *Chem. Commun.*, **2001**; 1078 – 1079

87. Skoog, D. A. and Leary, J. J., '*Principles of Instrumental Analysis*' Fourth Ed. Saunders College Publishing, **1992**; 579

88. Rodger, A. and Nordén, B. '*Circular Dichroism and Linear Dichroism*' Oxford University Press, **1997** and references therein

89. Craig, D. P., Power, E. A. and Thirunamachandran, T., *Chem. Phys. Letts.* **1974**; 2, 149 – 153

90. Wada, A. and Kosawa, S., *J. Polymer Sci.* **1964**; 2, 853 – 864

91. Dafforn, T., '*personal communication*' **2004**

92. Lakowicz, J. R., *'Principles of Fluorescence Spectroscopy'* Plenum Press, New York and London **1983**; 1 – 18

93. Allen, M. P., Tildersley, D. J., *Computer Simulation of Liquids*, Oxford University Press, Oxford, UK **1987** and references therein

94. Mac Kerell, A. D. Jnr., Wiórkiewicz-Kuczera, J. and Karplus, M., *J. Am. Chem. Soc.* **1995**; 117, 11946 – 11975

95. Khalid, S., *'Molecular simulation studies of the interaction between DNA and a novel macromolecular ligand'* PhD thesis, University of Warwick, **2004**

96. Khalid, S., Rodger A. and Rodegr, P. M., *in press*, **2005**

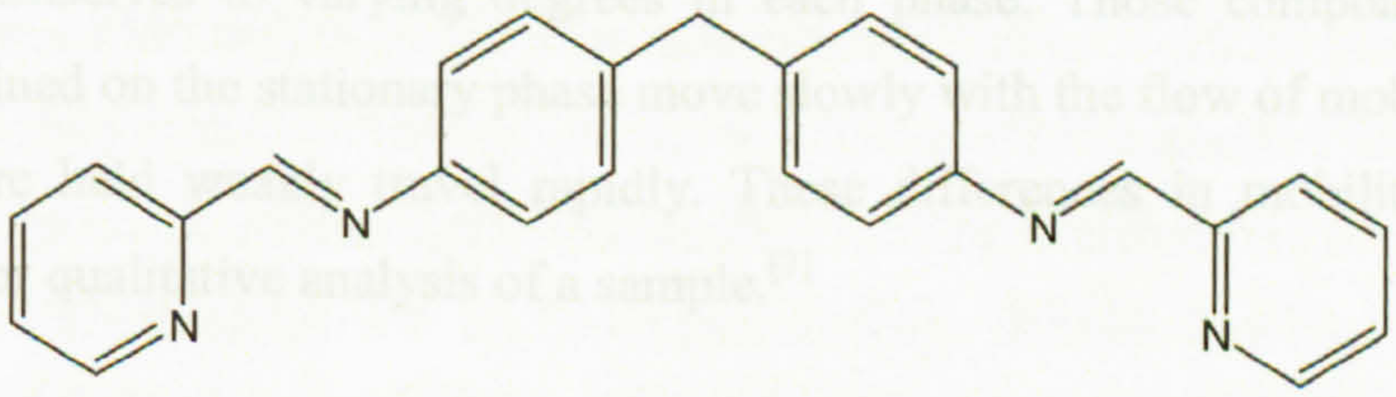
# CHAPTER 2 – METAL COMPLEX SYNTHESIS, CHARACTERISATION AND ENANTIOMERIC RESOLUTION

## 2.1 Introduction

The results presented in this chapter are focussed on the synthesis, characterisation and enantiomeric separation of a range of bimetallo iron cylinders (see figure 2.2) on cellulose in the form of paper or packed in a column using various eluents. The ligand structure of the parent compound is denoted $L^1$, and the metal complex is denoted $L_1$. This notation also applies to all of the derivatives shown in figure 2.2. Initial investigations involved the use of paper and 0.02 M NaCl as eluent. Cellulose chromatography has proved to be an ideal method for the purification and enantiomeric separation of the parent compound (see figure 2.1) and a range of its derivatives and the optimum mobile phase was found to be 90% acetonitrile:10% 0.02 M NaCl. This chapter also contains a description of the determination of an accurate extinction coefficient for the parent cylinder and results from the kinetic and photostability studies of the enantiomers of three cylinders.

## 2.2 Bimetallo iron cylinders

A group of metallo-supramolecular cylinders have been prepared by Hannon *et al.* at the University of Warwick, using a simple route for assembly from inexpensive commercially available starting materials.[1 – 3] The compounds have been designed to have a cylindrical shape (see figure 2.1) and to be similar in size and architecture to DNA major groove binding units of proteins. It has indeed been shown that the (M)-enantiomer of the parent compound does bind to the major groove of B-DNA spanning five base pairs.[2] The length of the parent cylinder is 1.8 nm and the radius is 0.53 nm.[4]



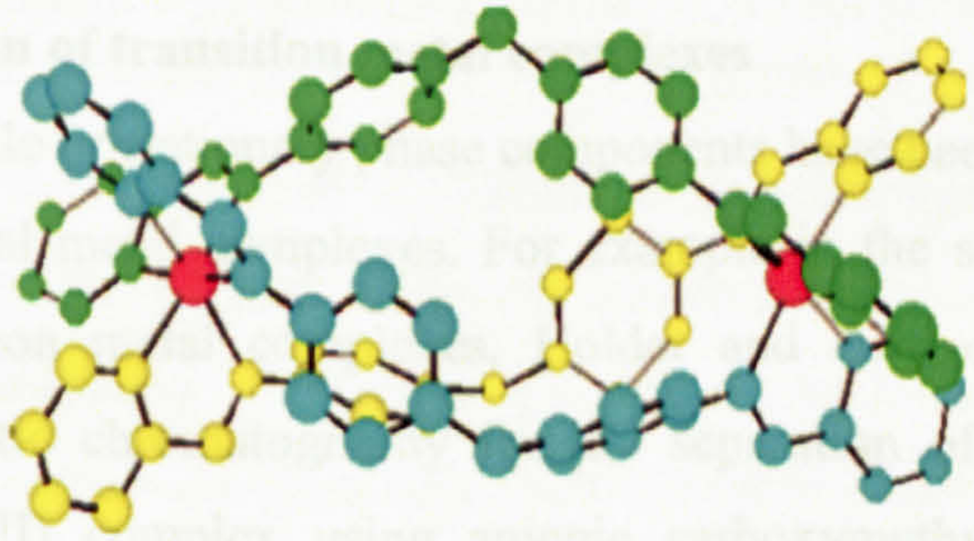



**Figure 2.1 Ligand and X-ray crystal structure of the parent compound $L_1$.[1]**

A high yield of the parent ligand is obtained simply by mixing ethanolic solutions of pyridine-2-carboxaldehyde and 4,4'-methylenedianiline. The simplicity of a single step reaction and high yield of product makes this an extremely attractive ligand system.[1] However, purification of the compounds has proved to be a challenge. The cylinder structure of Hannon *et al.* which is the primary focus of this research, represents the best developed and most investigated supramolecular architecture to date.[1 – 4] The cylinders contain two $Fe^{2+}$ metal ions that have a high positive charge and this should enhance their binding ability to the negatively charged DNA. The ligand structure (figure 2.1) incorporates two pyridylimine metal binding sites separated by a phenylene spacer. The phenylene groups sterically hinder and prevent the two metal binding sites from co-ordinating to a single metal centre.[1] Flexibility is introduced into the backbone of the ligand by the central methylene unit and this permits the ligand to support triple helical ligand arrays. Octahedral metal ions such as iron (II), nickel (II) and cobalt (II) require three bidentate pyridylimine units to satisfy their co-ordination requirements and so the formation of the triple-helical architecture was expected and indeed found.[1]

In this chapter the results presented are focussed on the synthesis, characterisation and resolution of a range of cylinders with allyl substituents at different positions on the cylinder backbone (see figure 2.2) where a particular aim of the work has been to reduce the amount of NaCl co-eluted with the cylinder enantiomers. It should be noted that the metal complexes outlined in figure 2.2, other than $L_1$, $L_3$ and $L_5$ were synthesised by Dr. C. Isaac and Dr. J. Kerckchoffs at the University of Warwick.

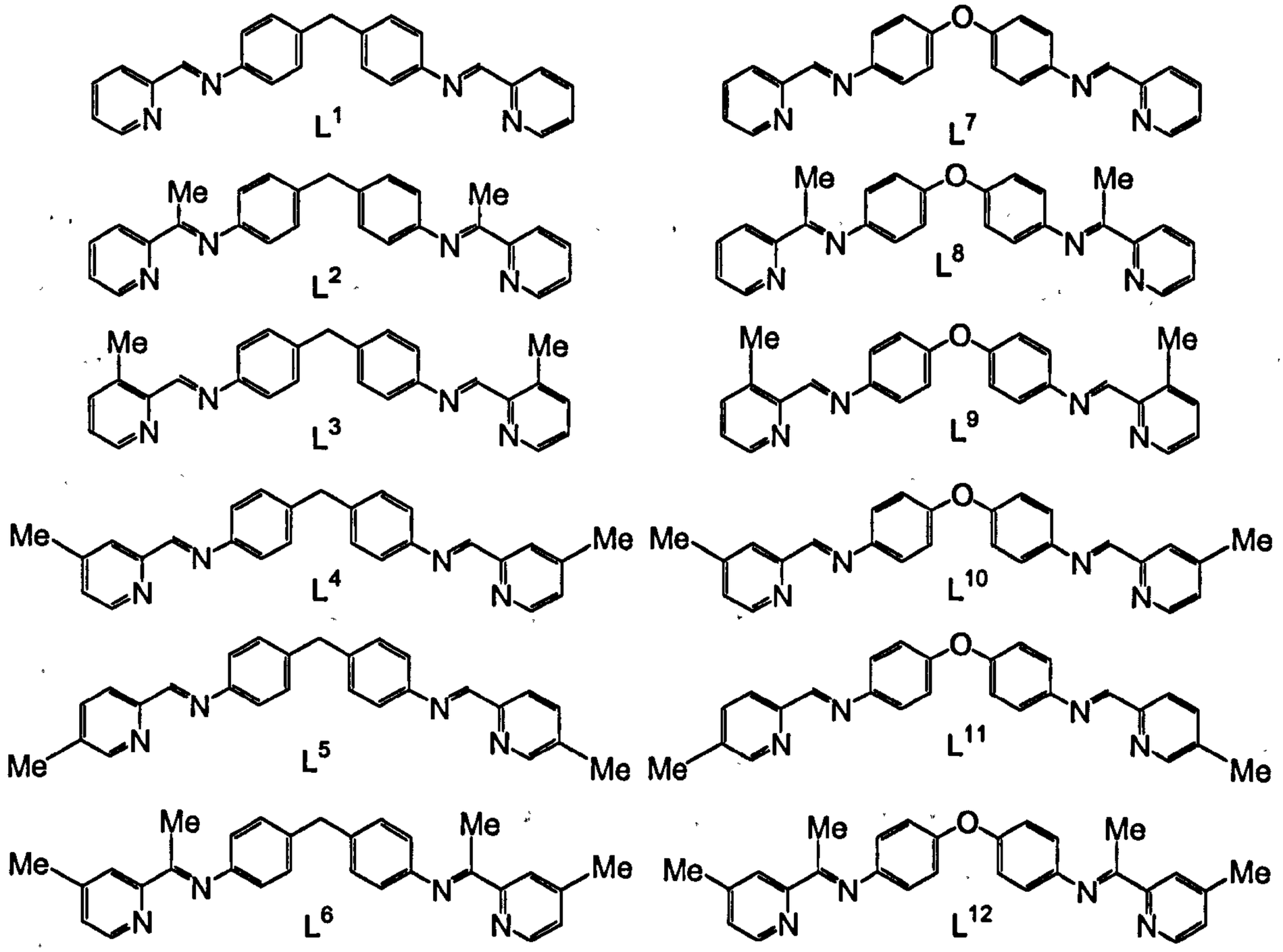



**Figure 2.2 Schematic diagram of the ligand structures.**

## 2.3 Chromatography

Chromatography encompasses an important and diverse range of methods that allow the separation of closely related components of complex mixtures.[5, 6] In all chromatographic separations the sample is dissolved in a suitable solvent and then forced through an immiscible stationary phase, which is fixed in place in a column, or on a solid surface. The two phases are chosen so that the components of the sample distribute themselves to varying degrees in each phase. Those components that are strongly retained on the stationary phase move slowly with the flow of mobile phase but those that are held weakly travel rapidly. These differences in mobility permit the quantitative or qualitative analysis of a sample.[7]

### 2.3.1 Chiral resolution of transition metal complexes

Cyclodextrins as mobile or stationary phase components have been successfully used by others to resolve chiral metal complexes. For example in the search for a method to resolve chiral transition metal complexes, Holder and co-workers[8, 9] successfully employed electrokinetic chromatography for the separation of a modified tris(2,2'-bipyridine)ruthenium(II) complex using anionic carboxymethyl-β-cyclodextrin as a chiral mobile phase additive. Alternatively, immobilised cyclodextrin has been used as a

stationary phase for the separation and resolution of chiral metal complexes by Armstrong *et al.*[10] Green *et al.*[11] and Yamanari and Nakamichi.[12]

Aldrich-Wright and co-workers have developed stationary phases which exploit DNA as the chiral discriminator. A covalently bound DNA stationary phase for HPLC applications was used to optically resolve $[Ru(dipyrido[6,7\text{-}d:2',3'\text{-}f]\text{-}quinoxaline)_3]^{2+}$ and $[Ru(1,10\text{-}phenanthroline)_3]^{2+}$ complex ions.[13] Due to the chiral nature of DNA and the different sites available for ligand binding, DNA has the ability to discriminate the enantiomers of these complexes and thus resolve them. Baker *et al.* have separated the enantiomers of *rac*-$[Ru(phen)_3]^{2+}$ using an adsorption-based DNA-hydroxylapatite chromatographic column.[14] The use of adsorption-based columns has many drawbacks, but the use of covalently bound DNA chromatographic columns as described by Aldrich-Wright and co-workers overcomes some of these problems.[13] The development of DNA-cellulose paper has provided a quantitative measure of the enantiomeric separation of several chiral metal complexes and this chromatographic technique was investigated for the separation of the bimetallo iron cylinders (see § 2.6.6).

Sephadex, usually with chiral mobile phases, has also been successfully used by a number of groups.[15, 16] In our endeavour to improve the efficiency and enantiomeric separation of the parent compound, sodium antimonyl-L- and D- tartrate were considered as potential eluent components. The use of these chiral mobile phase additives, however, did not improve the enantiomeric separation. We found these and more traditional methods of co-crystallisation with chiral anions only achieved limited success for our cylinders. For reasons that remain unclear, cellulose packed in aqueous saline seems to be the most effective stationary phase for these compounds. It was hoped that a universal method could be employed for the enantiomeric separation of a library of cylinders analogous to the parent compound (see figure 2.2).

For the purposes of this research, cellulose has been found to be the only stationary phase capable of efficiently separating the enantiomers of the range of bimetallo iron cylinders of figure 2.2. Our preliminary work was undertaken after an exhaustive attempt to find literature methods that separated the enantiomers of the parent compound failed to lead to a method for their separation. Initial investigations involved the use of paper and 0.02 M NaCl as eluent and this provided an ideal technique for the

separation of the enantiomers of the parent compound,[2] however, when applied to the enantiomeric separation of a range of derivatives, this method gave less satisfactory resolution and so other methods were investigated.

### 2.3.2 Cellulose chromatography

There is more cellulose on Earth than any other organic substance: it is the most abundant of all biomolecules.[17] Cellulose (see figure 2.3) is a natural polymer consisting of a long chain of β-D-glucose sugar molecules, linked by 1,4-glucosidic bonds that form a highly ordered helical structure. It is found as linear crystalline units which, singly or in groups or associated with other polymers, forms microfibrils.[18] It is also the main constituent of plant cell walls. When cellulose is used as a stationary phase, intermolecular forces of attraction between the sample and the cellulose, involving polar/ionic interactions, hydrophobic effects and hydrogen bonding, achieve chiral recognition and separation.

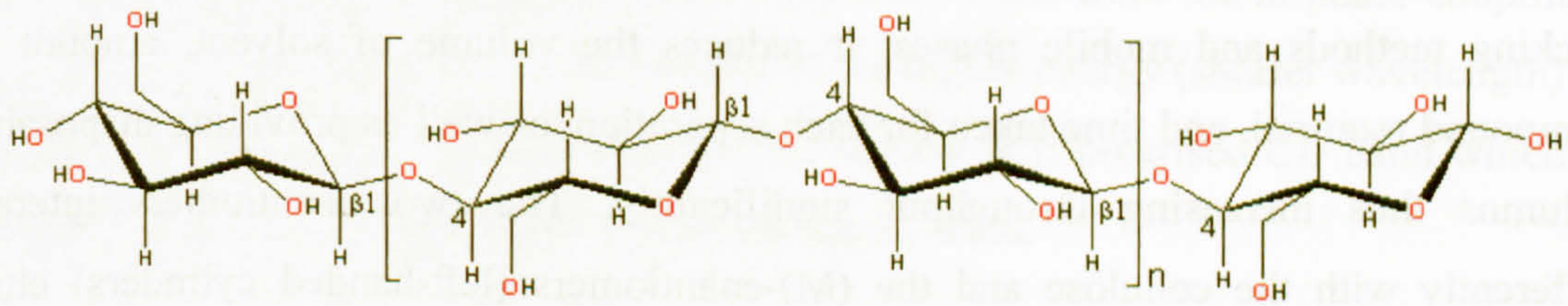



**Figure 2.3 Structure of cellulose.[18]**

Cellulose exists in two allomorphic forms: cellulose $I_\alpha$ and cellulose $I_\beta$. The crystal structure of each has been elucidated using Synchrotron X-ray and neutron fibre diffraction.[19, 20] In cellulose $I_\alpha$, alternating glucose units in each chain differ slightly in conformation but all chains are the same. Cellulose $I_\beta$, has chains of two kinds, termed 'origin' and 'centre' arranged in alternating sheets. Within each chain, all glucose units are identical, although they face in alternate directions (see figure 2.4). Each cellulose chain approximates to a flat ribbon, with alternate glucose units facing in opposite directions.[21] Hydrogen bonding between a hydroxyl group of one glucose unit and the ring oxygen of the next and weak C—H---O hydrogen bonds between sheets plays a prominent role in the cohesion of cellulose.

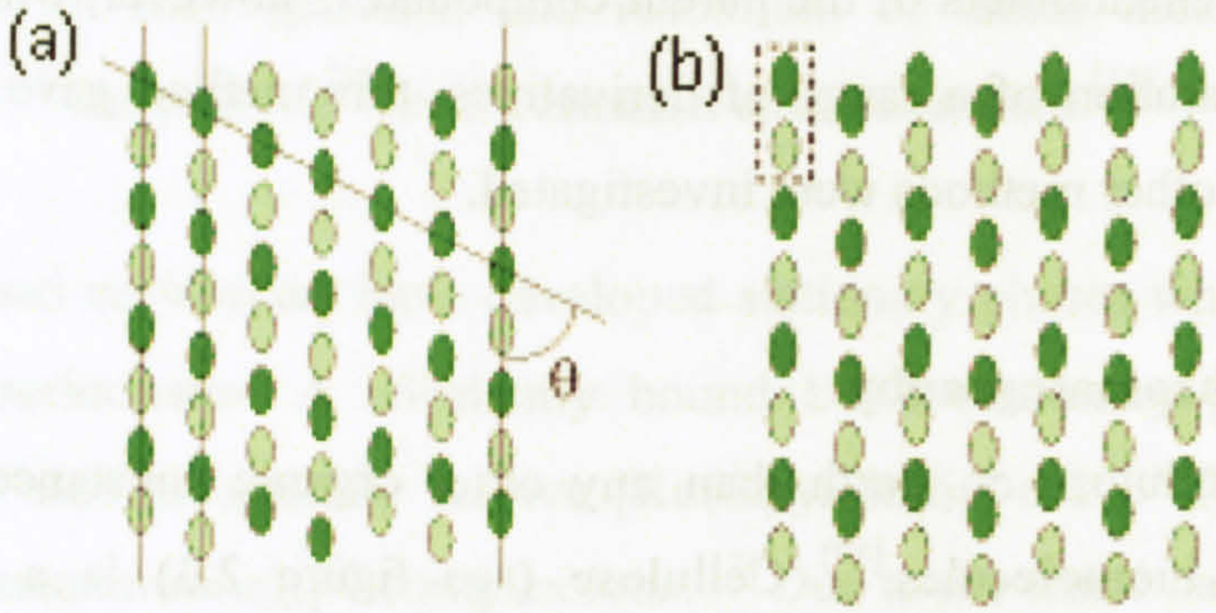



**Figure 2.4 Schematic diagram of (a) Cellulose $I_\alpha$, and (b) cellulose $I_\beta$.[22]**

We have used cellulose in the form of sheet paper, or packed in a column to separate the enantiomers of a range of bimetallo iron cylinders. A range of mobile phases were investigated to see if the separation and speed of elution could be increased and the amount of NaCl decreased. Methanol, ethanol and acetonitrile were considered, together with aqueous NaCl/organic mixtures. During this method development Pasteur pipettes were adopted as 'mini' columns. This provided an ideal method for investigating packing methods and mobile phases: it reduces the volume of solvent, amount of compound required, and time taken for each separation, as well as providing disposable columns thus increasing throughput significantly. The two enantiomers interact differently with the cellulose and the (M)-enantiomers (left-handed cylinders) elute from the column first.

## 2.4 Bimetallo cylinder spectroscopy

In order to be able to use *CD* spectroscopy as a fast means of confirming the identity of fractions eluted from the cellulose columns it was important to understand the origin of the sign of the bands and to be able to relate this to the handedness of the bimetallo helicate. The following analysis follows that of mono-metallo tris chelate complexes.[23 – 25]

*In-ligand transitions:* The key feature of a chromophore, as of a functional group, is that it can be considered as a spectroscopically well-defined sub-unit of a molecule that is only slightly perturbed by the rest of the system. For the in-ligand transitions of metal complexes, the individual ligands are natural chromophores. The *CD* of in-ligand transition of *tris* chelate mono-metallo complexes have therefore been very successfully modelled as arising from the coupling of electric dipole allowed (eda) transition moments within the ligands.[24, 26] Since the bridge between the two halves of the

bimetallo triple helical ligands is saturated, the conjugation along the ligand backbone is broken, so the bimetallo triple helical complexes can be viewed as two tris-chelate metal complexes stacked along the three fold axis. Thus in-ligand *tris* chelate transition metal complex *CD* equations can be used after multiplying by a factor of two. We may then write the *CD* intensity, *R*, for the $A_2$ and E polarised in-ligand bands of the P bi-metallo cylinder as,[23 – 25]

$$R\left(z/x/y\right) = R\left(A_2/E\right) = \pm\frac{\varepsilon\mu^2\rho}{\sqrt{2}\hbar} \qquad (1)$$

with transition energies

$$\varepsilon\left(z/x/y\right) = \varepsilon\left(A_2/E\right) = \varepsilon \pm \frac{\mu^2}{12\sqrt{3}\rho^3} \qquad (2)$$

where $\varepsilon$ is the energy of the unperturbed transition, $\mu$ is the magntiude of its transition dipole moment, and $\rho$ is the distance from the metal to the centre of a chelating part of the ligand. *R(z)* is the *z* (or $A_2$) polarized *CD* that results from the in-phase coupling of the three dipoles. Thus the $A_2$ band occurs at the higher energy (shorter wavelength) and is positive. The out-of-phase couplings give the *x/y* or E polarised *CD* band which has the same magnitude but opposite sign from the $A_2$ band.

The observed *CD* spectra (Figures 2.13 – 2.16) are a complicated overlay of different MLCT and in-ligand bands. There is, however, a clear sharp exciton signal centred at ~280 nm indicating this is due to the exciton coupling of a single in-ligand transition. This lets us identify the handedness of enantiomers eluting from the column: if the higher energy (shorter wavelength) band is positive then the complex is the P enantiomer. Thus the second eluting enantiomer is P.

*Metal to ligand charge transfer transitions:* The definition of the chromophore for the metal-ligand charge transfer (MLCT) transitions is less clear than for the in-ligand transitions. Any model must account for the fact that the available data on transition polarisations of long wavelength charge transfer transitions for low spin iron(II) and ruthenium(II) (including our own stretched film *LD* assignments of transition polarisations in $Fe_2L_3$) show significantly more (>90%) E than $A_2$ polarisation of the long wavelength end of the absorbance spectrum. By way of contrast the *CD* spectra show similar magnitude postive and negative bands (which correlates with different polarisations having similar magnitudes).

For systems, such as low spin $d^6$ systems, where all the *d* electrons are in the between-bond $t_{2g}$ *d* orbitals with one lobe pointing to the middle of a chelate, the two simple alternative ways of modelling MLCT transitions for unsaturated *tris* chelate systems are to assume either (i) that the electrons transfer between the metal and the ligating atoms (thus defining an approximately $D_{3d}$ $ML_6$ chromophore for *tris* chelate metal complexes) as in reference 26 or (ii) that the MLCT transition is from a *d* orbital to the middle of a ligand (into a ligand $\pi^*$ orbital).[27] With case (i) the $A_2$ and E polarised transitions will have equal and opposite *CD* intensity[26] and the normal absorbance intensity of the E polarisation will be twice that of the $A_2$.[28] Thus case (i) is not a good description of the spectroscopy of these compounds. With case (ii), an MLCT transition will involve motion of electron density from the metal into ligand orbitals directed either along the ligand short axis or long axis. The short axis $\pi^*$ states will give rise only to planar E polarised MLCT bands (from the out-of-phase coupling of the individual *d* to ligand orbitals as the in-phase coupling has no net electric dipole transition moment). The long axis $\pi^*$ states will give rise to both E and $A_2$ MLCT bands and will be the only ones giving rise to a *CD* signal. The net effect will be that the absorbance spectrum may be dominated by E polarised transitions in regions where the MLCT bands include short axis polarised transitions, whereas the *CD* spectrum will have similar contributions from the two polarisations — in accord with experimental observation. So we conclude that case (ii) is relevant here and the in-ligand and MLCT bands of the same polarisation will have the same sign. Since the enantiomeric assignment has been made using the 280 nm in-ligand band, the *CD* signs can be used to assign the $A_2$ and E components of the MLCT bands and the longer wavelength (320 nm) in-ligand band that overlaps with MLCT bands.

## 2.5 Materials and methods

All materials were purchased from Sigma-Aldrich unless otherwise stated. Ultra pure water (18.2 ΩM) was used in all experiments. 2-Pyridine carboxaldehyde was used without further purification. 3- and 5-methyl pyridine carboxaldehyde were prepared from their respective commercially available bromo-methyl pyridines according to literature methods.[29] Anhydrous tetrahydrofuran (THF, 99.9%), n-butyl lithium (n-BuLi, 1.6 M solution in hexanes), acetic anhydride (99+%), N-N-dimethylformamide (DMF, 99.8% anhydrous), 4,4-methylenedianiline and iron (II) chloride tetrahydrate (99%) were purchased from Sigma-Aldrich and used without further purification. Toluene, dichloromethane, molecular sieves (3 Å, 1.6 – 2.5 mm) and silica (60 Å, 35 –

70 μM) were purchased from Fisher. Chromatography cellulose paper (Whatman, 3MM CHR, Fisher) and cellulose (powder ~ 20 micron, Sigma-Aldrich) were used to separate the enantiomers. Methanol, ethanol and acetonitrile were purchased from Fisher Scientific. Sodium antimonyl-L-tartrate (93+%) was purchased from Avocado Research chemicals. Sodium antimonyl-D-tartrate was synthesised according to literature methods.[30]

UV/visible absorbance spectra were recorded on a Jasco V-550 spectrometer and *CD* and *LD* spectra were recorded on a Jasco J-715 spectropolarimeter. The methods used to synthesise and enantiomerically separate a range of cylinders are outlined below.

### 2.5.1 Synthesis of methylated metal complexes

Synthesis of the 3' and 5' methylated cylinders ($L_3$ and $L_5$) involves the reaction of 4,4-methylenedianiline with the corresponding *x*-methylpyridine carboxaldehyde, which is prepared from the commercially available bromo-methylpyridine according to literature methods:[29] 2-bromo-*x*-methylpyridine (4.370 g, 23.6 mmols) in anhydrous THF (40 mL) was syringed into a thoroughly dried round bottom flask and cooled in a dry ice/acetone bath. n-BuLi in hexane (1 mol $dm^3$, 18.875 mL, 1.3 M eq.) was added slowly over a period of 10 minutes. The solution was left to stir for 15 minutes at –78°C. After this time, anhydrous DMF (1 $cm^3$) was added slowly and the solution allowed to stir for a further 15 minutes at room temperature (see figure 2.5).

The reaction was then quenched with water (50 mL), followed by extraction with diethyl ether (3 × 50 mL). The ether extracts were reduced *in vacuo* to yield an oil. This product was then purified on silica using dichloromethane as the mobile phase to yield a colourless oil. The presence of the final product was confirmed using thin layer chromatography and NMR.

The ligand was prepared by mixing 2 equivalents of the *x*-methyl pyridine carboxaldehyde (1.90 mL, 0.020 mol) with 1 equivalent of 4,4-methylenedianiline (1.99 g, 0.010 mol) in toluene (50 mL) containing vacuum dried 3 Å molecular sieves (5 g) and allowed to stir overnight. The resultant yellow precipitate (3.40 g, 0.009 mol) was filtered through celite and recrystallised from hot ethanol. Yield: 90%.

The metal complexes (see figure 2.6) were synthesised by adding iron (II) chloride tetrahydrate (0.48 g, 0.002 mol) to the ligand (1.36 g, 0.003 mol) in methanol (180 mL). The purple solution characteristic of iron (II) tris-pyridylimine compounds was refluxed for three hours. After cooling, ammonium hexafluorphosphate dissolved in methanol was added in excess. The $PF_6$ salt of the cylinder precipitates out of solution and is isolated by subsequent filtration. Following filtration, the solid was washed with ether and allowed to dry in a vacuum desiccator. The $PF_6$ salt is chemically purer than the corresponding $Cl_4$ salt, but is not soluble in aqueous solution. The chloride salt was obtained by anion metathesis in acetonitrile using tetrabutylammonium chloride. The same complexes could be prepared directly by treating mixtures of the diamine and aldehyde/ketone in methanol solution in the appropriate ratios.

The purple solution was taken to dryness *in vacuo* and freeze-dried. The methylated metal complexes are termed $L_3$ and $L_5$ respectively to indicate the position of the methyl group on the outer pyridyl ring.

### 2.5.1.1 Characterisation data

NMR spectra were recorded on a Bruker DPX 300/400 MHz instrument using standard Bruker software. Electrospray ionisation (ESI) analyses were performed by the EPSRC National Mass Spectrometry Service Centre, Swansea, on a Micromass Quatro (II) instrument in positive ionisation mode. Microanalyses were conducted on a Leeman Labs CE44 CHN analyser by the University of Warwick Analytical Service. Infrared spectra were recorded with a Perkin Elmer Paragon 1000 FTIR spectrometer. NMR, CHN, ESI-MS and IR data confirm the identity of the complexes.

**$L_3$.** Anal. Calc. for $[Fe_2(C_{27}H_{24}N_4)_3][PF_6]_4$: C, 51.1; H, 3.8; N, 8.8%. Found: C, 50.8; H, 4.1; N, 8.8%. Positive-ion ESI (MeCN): m/z ($[Fe_2(L^2)_3(PF_6)_3]^+$), 1760.1 ($[Fe_2(L^2)_3(PF_6)_2]^{2+}$), 807.2, ($[Fe_2(L^2)_3(PF_6)]^{3+}$), 489.9 ($[Fe_2(L^2)_3]^{4+}$), 331.2. $^1$H NMR ($(CD_3CN)$ 300 MHz, 300 K): δ2.85 (3H, s, Me), 4.02 (1H, s, Me), 5.28 (1H, bs, $H_{Ph}$), 5.74 (1H, bs, $H_{Ph}$), 6.63 (1H, bs, $H_{Ph}$), 7.06 (1H, d, $H_5$ J = 4.5 Hz), 7.31 (1H, bs, $H_{Ph}$), 7.61 (1H, bt, $H_4$ J = 6.0 Hz), 8.15 (1H, d, $H_3$ J = 7.0 Hz), 8.96 (1H, s, $H_{im}$); Selected IR data ($cm^{-1}$): 3386w, 1626w, 1588w, 1558w, 1503s, 1474m, 1441m, 1380m, 1335m, 1308w, 1256w, 1166w, 1110w, 1060w, 1019w, 828vs, 771vs, 750vs, 691m, 674m.

$L_5$. Anal. Calc. for $[Fe_2(C_{27}H_{24}N_4)_3][PF_6]_4$: C, 51.1; H, 3.8; N, 8.8%. Found: C, 50.8; H, 4.0; N, 8.6%. Positive-ion ESI (MeCN): m/z ($[Fe_2(L^3)_3(PF_6)_3]^+$), 1761.1 ($[Fe_2(L^3)_3(PF_6)_2]^{2+}$), 807.4 ($[Fe_2(L^3)_3(PF_6)]^{3+}$), 490.2 ($[Fe_2(L^3)_3]^{4+}$), 331.4. $^1$H NMR ((MeOD) 400 MHz, 300 K): δ2.48 (3H, s, Me), 4.08 (1H, s, Me), 5.59 (2H, vbs, $H_{Ph}$), 7.04 (2H, vbs, $H_{Ph}$), 7.31 (1H, s, $H_5$), 8.33 (1H, d, $H_{3/4}$, J = 7.5 Hz), 8.62 (1H, d, $H_{3/4}$, J = 8.0 Hz), 9.16 (1H, s, $H_{im}$); Selected IR data($cm^{-1}$): 3137w, 2350w, 1625m, 1597w, 1561w, , 1499s, 1354w, 1221m, 1197s, 1106w, 1040m, 1016w, 912s, 836s, 770m, 748w, 658w.

### 2.5.2 Determination of ε

An accurate value of ε was required so that the concentration of the cylinders could be determined spectrophotometrically using the Beer Lambert Law. An accurate value for ε was determined for the parent compound using UV/visible absorbance spectroscopy. Three independent masses of each metal complex were weighed accurately using a seven-figure balance and dissolved in ultra pure water (50 mL). Each solution was diluted twice and then analysed by UV/visible absorbance spectroscopy. Several attempts were made to determine the extinction co-efficients of (R)-$L_3$ and (R)-$L_5$, but due to the low purity (even after several recrystallisations from hot ethanol, and diffusion recrystallisation using benzene and acetonitrile) and solubility of the compounds, it was concluded that it was not possible to obtain an accurate ε value for the methylated cylinders and so ε of the parent compound was used to determine the concentration.

UV/visible absorbance parameters: scan range: 750 – 200 nm; scan speed: 200 nm/min; accumulation: 1.

### 2.5.3 Stability studies

#### 2.5.3.1 Photostability experiments

Solutions of (R)-$L_1$, (R)-$L_3$ and (R)-$L_5$ were prepared. A concentrated (500 μM) and a dilute (50 μM) solution of each of the metal complexes were placed in the dark and in the light and UV/visible absorbance spectra were recorded weekly over a period of 22 days. To obtain accurate UV/visible absorbance spectra for the concentrated aqueous metal complex solutions, a 1 in 32 dilution was made and the UV/visible absorbance spectra collected. Refer to § 2.5.3.2 for instrumental parameters.

### 2.5.3.2 Thermal stability of enantiomers

The enantiomers of (R)-$L_1$, (R)-$L_3$ and (R)-$L_5$ were separated on a cellulose column using NaCl (0.02 M) and the solutions freeze-dried. 50 μM solutions of each enantiomer were prepared and placed in an oven at 37°C, on the bench at 20°C, in a refrigerator at 4°C and in a freezer at –20°C. UV/visible absorbance and *CD* spectra were recorded weekly over a period of 22 days.

UV/visible absorbance parameters: scan range: 750 – 200 nm; scan speed: 200 nm/min; accumulation: 1. *CD* parameters: scan range: 750 – 200 nm; scan speed: 200 nm/min; step resolution: 0.5 nm; response: 1.0 s; band width: 2.0 nm; sensitivity: 20 mdeg; accumulation: 4 scans.

### 2.5.4 Paper chromatography

Chromatography cellulose paper was cut into 5 cm × 20 cm pieces. The paper was spotted with a concentrated aqueous solution (approx. 5 mg/mL) of the metal complex and the enantiomers eluted by capillary action with a 0.02 M sodium chloride solution, in which the paper was suspended. Once the solvent front reached nearly to the top of the paper, the paper was removed from the eluent and allowed to dry. A range of mobile phases were investigated including: 75% ethanol:25% 0.02 M NaCl, 80% methanol:20% 0.02 M NaCl, and 90% acetonitrile:10% 0.02 M NaCl. The resolution factor gives a quantitative measure of the separation of each of the metal complexes (see table 2.2) and can be calculated using the following equation:

$$R_f = \frac{\text{distance moved by component}}{\text{distance moved by solvent front}}$$

### 2.5.5 DNA-paper chromatography

The enantiomeric separation of twelve cylinders was investigated using DNA-paper chromatography as a comparison to the results obtained using only cellulose paper chromatography. ct-DNA (0.3 g) and sodium periodate (3.0 g) were dissolved in water (160 mL) and the solution stirred and warmed for 30 minutes. The solution was then transferred to a shallow tray. A solution of sodium borohydride (2.0 g) in water (160 mL) was placed in a second tray. Chromatography cellulose paper was cut into 5 cm × 20 cm pieces and each sheet allowed to soak in the DNA solution prior to its being transferred to the borohydride tray, where the reaction is quenched. The DNA-

immobilised paper was then removed and allowed to air dry for several hours on a flat surface. The DNA paper was spotted with a concentrated aqueous solution of the metal complex and the enantiomers eluted by capillary action with a 0.02 M sodium chloride solution, in which the paper was suspended. Sodium cacodylate buffer (pH 6.8), 90% MeCN: 10% 0.02 M NaCl, 80% MeOH: 20% 0.02 M NaCl, 75% EtOH: 25% 0.02 M NaCl and 0.02 M NaCl were used as eluents to investigate the enantiomeric separation of a range of cylinders. Once the solvent front reached nearly to the top of the paper, the paper was removed from the eluent and allowed to dry.[13]

### 2.5.6 Cellulose column chromatography

Cellulose columns were packed into 2 cm × 30 cm unsintered columns using cellulose particles (~20 micron) as the stationary phase and aqueous 0.02 M sodium chloride as the solvent in which the cellulose was suspended for packing. To 6 g of cellulose, 40 mL of 0.02 M sodium chloride was added and the solution stirred to a smooth consistency. The column was packed by pouring the aqueous saline suspension of cellulose onto a glass wool pad located just above the stopcock and excess solvent was eluted. The sample, as the chloride salt (the equivalent $PF_6$ salt is not soluble in aqueous solution), was then loaded onto the column as a saturated aqueous solution (approx. 5 mg/mL) and the column eluted with the mobile phase of choice. To increase through-put Pasteur pipettes were used as 'mini' columns for the separation of the metal complex derivatives. A range of eluents was prepared by mixing organic solvents with 0.02 M NaCl with organic content ranging from 95% to 50%. Enantiomeric separation was also investigated using non-aqueous organic solvents, less concentrated sodium chloride solutions and with sucrose in the eluent. Sodium antimonyl-L- and D- tartrate were also investigated as potential eluent components for the separation of the enantiomers. The fraction collection was guided by visual inspection of the profile determined by whether a visible separation of two bands was present or not. If no visible separation of two bands was observed a number of small fractions were collected across the purple band. If separation was apparent, two fractions were collected, one from each purple band. The UV/visible absorbance and *CD* of each fraction were recorded.

The bimetallo iron cylinder is a chiral molecule, so the *CD* was recorded to confirm its enantiomeric composition: a flat line signal indicates that the solution is racemic, i.e. the enantiomers are present in solution in equal proportion. The *CD* spectra of the

enantiomers of all the compounds discussed in this work were converted to Δε assuming all compounds had the same metal to ligand charge transfer extinction coefficient as the parent compound. Solutions were assumed to be enantiomerically pure if the *CD* spectra, expressed as Δε, of the two enantiomers were equal and opposite. It should be noted that if these spectra, which have been normalised to the parent ε, are equal and opposite then the cylinders are enantiomerically pure, even if the value for Δε is in error by the factor ε(cpd)/ε(parent cpd).

UV/visible absorbance parameters: scan range: 750 – 200 nm; scan speed: 200 nm/min; accumulation: 1. *CD* parameters: scan range: 750 – 200 nm; scan speed: 200 nm/min; step resolution: 0.5 nm; response: 0.5 s; band width: 2.0 nm; sensitivity: 20 mdeg; accumulation: 4 scans.

### 2.5.7 High-pressure liquid chromatography

Initial investigations involved the use of a standard $C_{18}$ reverse phase column and 90% MeCN: 10% $H_2O$ as eluent. A dilute solution of the racemic parent cylinder was prepared in water (approx. 0.5 mg/mL) and injected directly onto the column. Chromatograms were collected using a Jasco HPLC system (Jasco PU-1580 pump, DG-980-50 degasser and AS-950 auto-sampler). The column was purchased from Phenomenex (Hypersil 5 μ, ODS $C_{18}$, 120 Å, 250 × 2.0 mm). The mobile phase flow rate was set to 0.5 mL/min and chromatograms collected. The analytes were detected using UV detection set at 254 nm and 580 nm for the detection of the in-ligand and metal to ligand charge transfer absorbance bands respectively (Jasco UV-975). See table 2.1 for summary of chromatographic conditions.

| | |
|---|---|
| Column: | Hypersil 5 μ, ODS $C_{18}$, 120 Å, 250 × 2.0 mm |
| Mobile phase: | 90% MeCN: 10% $H_2O$<br>Degassed by sonicating under reduced pressure |
| Flow rate: | 0.5 mL/min |
| Detection method: | UV @ λ = 254 nm or 580 nm |
| Injection volume: | 10 μL (full loop injection) |
| Run time: | 25 mins |

**Table 2.1 Summary of chromatographic conditions used for the enantiomeric separation of the parent cylinder.**

### 2.5.8 High temperature docking calculations

Compounds were modelled using the CHARMM22 all-atom force-field and simulations performed using DL_POLY.[31, 32] The conversion from CHARMM to DL_POLY force-field formats was achieved using a purpose built program that interprets the CHARMM prm and crd files. A Coulomb sum was used to evaluate electrostatic interactions, with a cut off of 15 Å. Simulations were performed in the *NVT* ensemble (constant number of molecules, constant volume and constant temperature) using the Hoover thermostat with a time constant of 0.1 ps and a time step of 2 fs. Orthorhombic periodic boundary conditions were used to mimic an infinite system (box size: 30 × 30 × 30 Å). A dielectric constant of 80 $C^2 J^{-1} m^{-1}$ was used to mimic the screening effects of water.

Two docking calculations were performed.[33] The cylinder was immobilised while the sugar was mobile (although restrained such that the chair conformation was retained throughout). *NVT* MD simulations were performed for 1 ns at 900 K. The conformation with the lowest configurational energy was extracted from each trajectory and then two further systems were set up in which the cylinders were replaced with the opposite enantiomer, with both of the enantiomers in each of the two different sites. These configurations were then energy minimised. Single point calculations of the sugars after minimisation gave the configurational energies. These were subtracted from the total configurational energies to give the interaction energies. It is important to note that the cylinder does not contribute to the total configurational energy as it was frozen.

## 2.6 Results

### 2.6.1 Synthesis of methylated metal complexes

The $L_3$ and $L_5$ cylinders were synthesised according to the schematic outlined in figure 2.5. The purity of the metal complex is dependent on the initial purity of the ligand so the ligand was recrystallised from hot ethanol repeatedly to improve the purity. This was continued until CHN analysis confirmed chemical purity. The $PF_6$ salt of the metal complex is of higher purity than the corresponding chloride salt and as described in § 2.4.1 the chloride salt was obtained by anion metathesis.

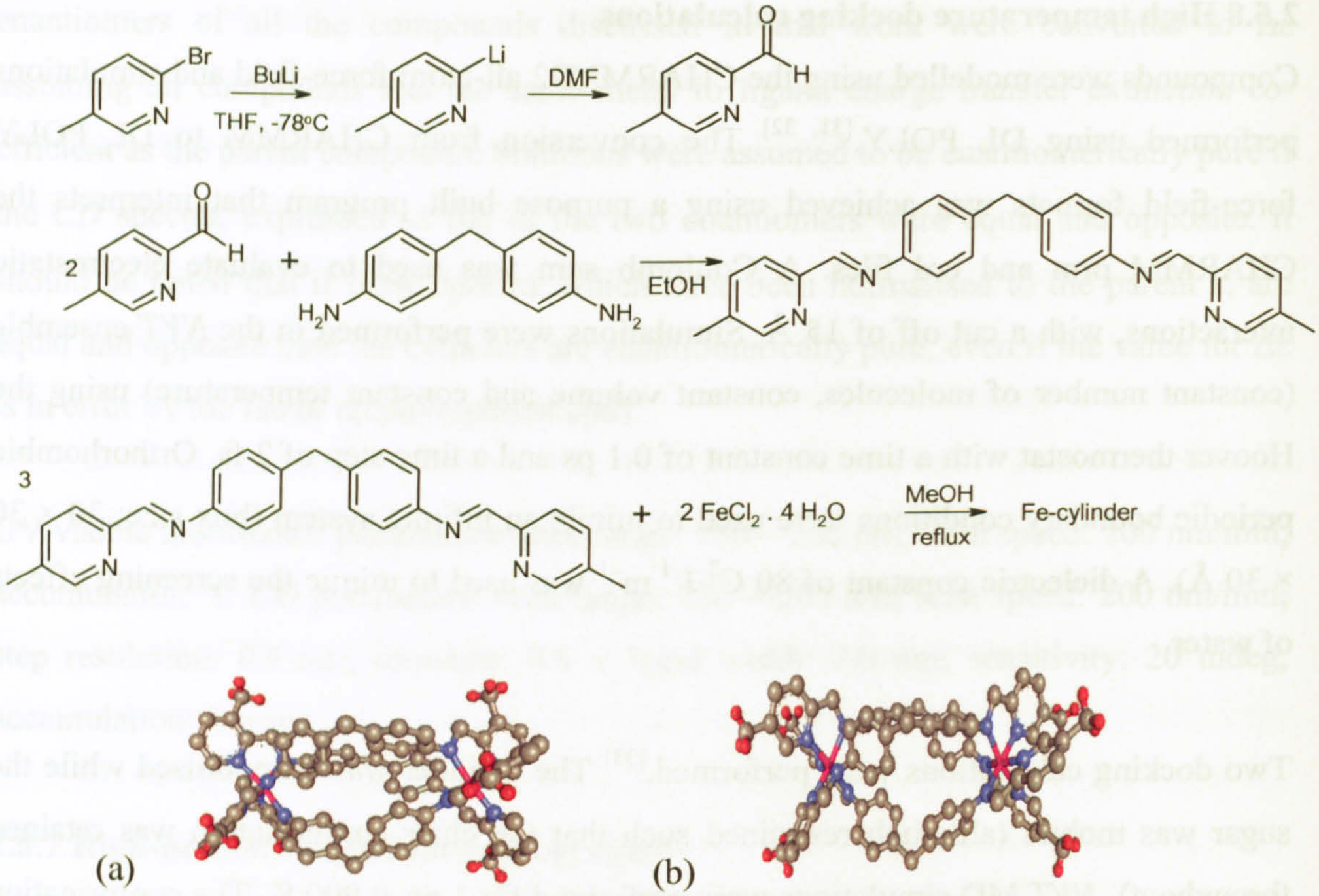



**Figure 2.5 Schematic diagram showing synthesis of 5-methylpyridine carboxaldehyde and structures of (a) $L_3$ and (b) $L_5$.**

The UV/visible absorbance spectra obtained for the $L_3$ and $L_5$ metal complexes (50 μM, 1 cm path length cuvette) had very similar spectral shapes (see figure 2.6) with a slight variation in the metal to ligand charge transfer and in-ligand wavelengths for some of the complexes.

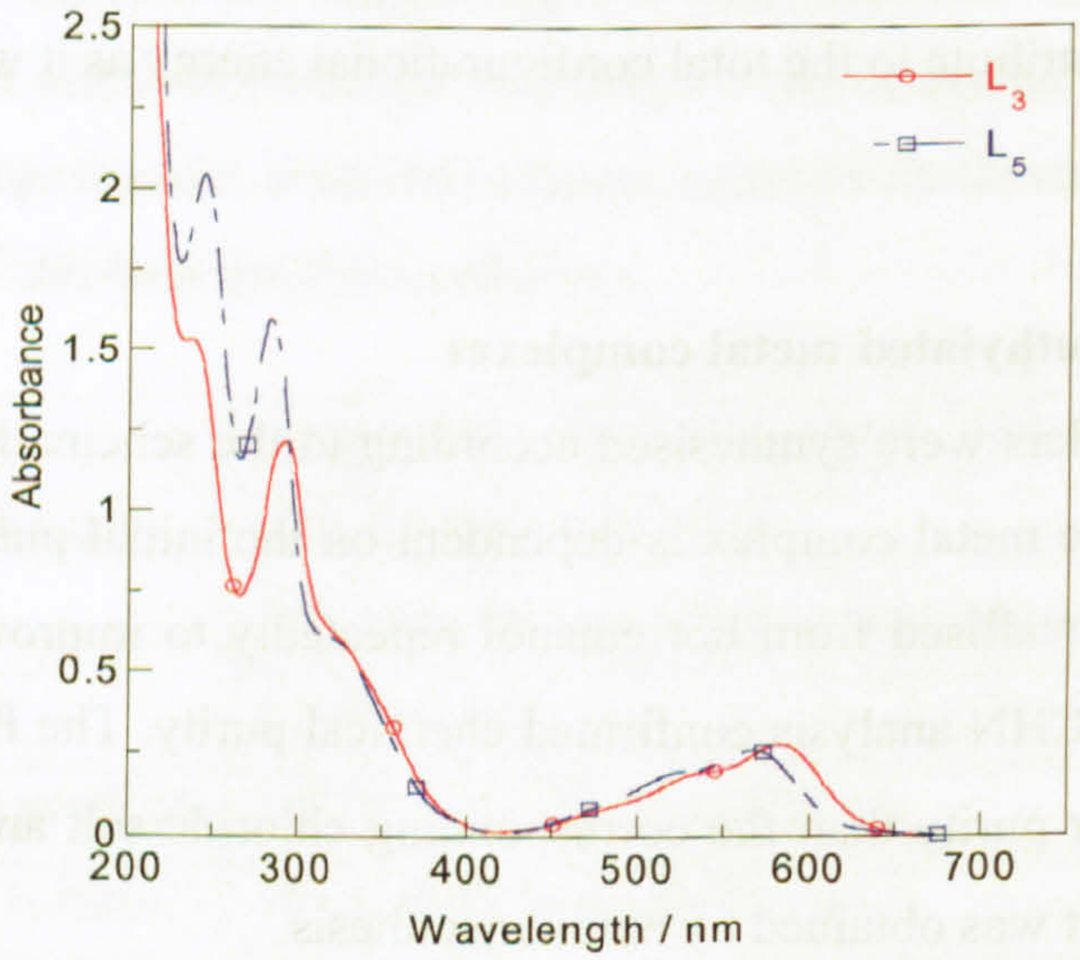



**Figure 2.6 UV/visible absorbance spectrum of (R)-$L_3$ and (R)-$L_5$ metal complexes showing differences in MLCT and in-ligand regions of the spectrum. Absorbance data above 2.5 has been omitted.**

### 2.6.2 Determination of ε

An accurate value of ε was determined for the parent compound, $L_1$ (see figure 2.7). Unfortunately, due to problems with purity and solubility, an accurate value for the extinction co-efficient of the $L_3$ and $L_5$ cylinders was not obtained and so the ε value of the parent compound was used to determine the concentration of these metal complexes.

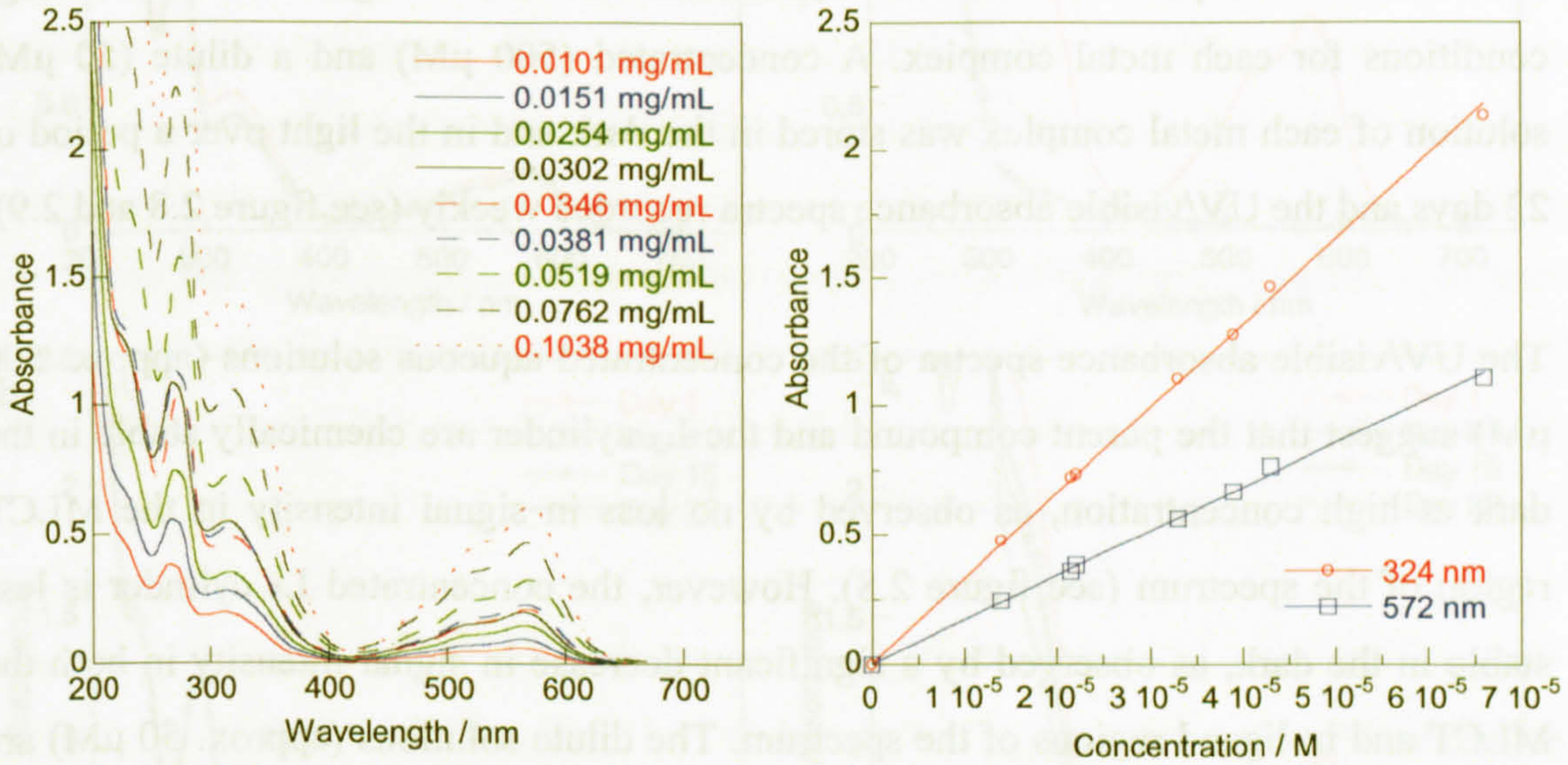



**Figure 2.7 UV/visible absorbance data for the determination of ε for the parent compound $L_1$. Gradient of each line gives the corresponding ε value in the MLCT and in-ligand regions of the spectrum. Absorbance values above 2.5 were omitted.**

The gradient of the slope in figure 2.7 gives the extinction co-efficient at the corresponding wavelength and was used to determine an accurate concentration for the metal complexes using UV/visible absorbance spectrophotometry and the Beer Lambert Law (see table 2.2).

| Metal complex | In-ligand ε / $cm^{-1}$ $mol^{-1}$ $dm^3$ | MLCT ε / $cm^{-1}$ $mol^{-1}$ $dm^3$ |
|---|---|---|
| $L_1$ | $\varepsilon_{324} = 32{,}900$ | $\varepsilon_{574} = 16{,}900$ |

**Table 2.2 ε values for the parent compound, $L_1$ in the MLCT (574 nm) and in-ligand (324 nm) regions of the spectrum.**

In an attempt to determine the accurate extinction co-efficients for the methylated bimetallo iron cylinders $L_3$ and $L_5$ it was observed that although reproducible values for ε were obtained from different batches of metal complex and repeat experiments, the ratio of in-ligand ε to MLCT ε was significantly larger than 2 (the value for the parent compound) suggesting the presence of unreacted starting ligand. The metal complexes were recrystallised from hot ethanol and diffusion recrystallisation using benzene and

acetonitrile. The ratio of in-ligand to MLCT absorbance reduced but an accurate value for ε was still not obtained.

### 2.6.3 Photostability experiments

It was important to investigate the stability of the metal complexes due to the length of time taken to complete each *CD/LD* experiment and to investigate the ideal storage conditions for each metal complex. A concentrated (500 μM) and a dilute (50 μM) solution of each metal complex was stored in the dark and in the light over a period of 22 days and the UV/visible absorbance spectra recorded weekly (see figure 2.8 and 2.9).

The UV/visible absorbance spectra of the concentrated aqueous solutions (approx. 500 μM) suggest that the parent compound and the $L_3$ cylinder are chemically stable in the dark at high concentration, as observed by no loss in signal intensity in the MLCT region of the spectrum (see figure 2.8). However, the concentrated $L_5$ cylinder is less stable in the dark, as observed by a significant decrease in signal intensity in both the MLCT and in-ligand regions of the spectrum. The dilute solutions (approx. 50 μM) are considerably less stable in the dark. The parent compound and the $L_5$ cylinder are chemically unstable at the lower concentration; and the dilute $L_3$ cylinder is more stable in the dark, but does show some loss in signal intensity in the MLCT region of the spectrum.

In summary, this experiment has confirmed that concentrated solutions of the metal complexes are more stable than the dilute solutions and that the order of stability of the metal complexes is (R)-$L_1$ > (R)-$L_3$ > (R)-$L_5$.

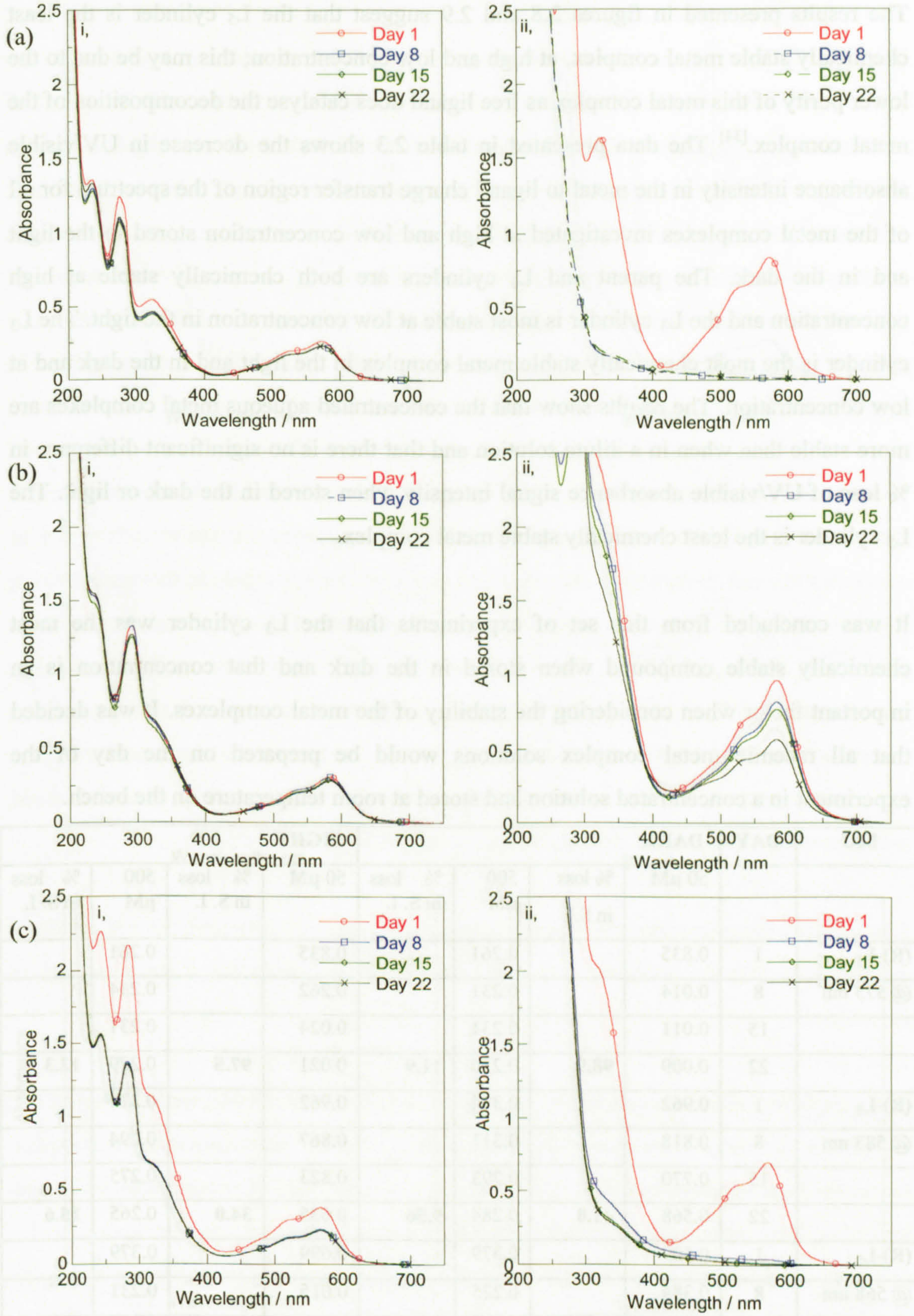



**Figure 2.8 UV/visible absorbance spectra of (a) (R)-$L_1$ (b) (R)-$L_3$ and (c) (R)-$L_5$ at i, high (500 μM) and ii, low (50 μM) concentration, stored in the dark. Absorbance data above 2.5 have been omitted. A 1 in 32 dilution was made of the 500 μM solutions and the spectra collected in a 1 cm path length cuvette.**

The results presented in figures 2.8 and 2.9 suggest that the $L_5$ cylinder is the least chemically stable metal complex, at high and low concentration; this may be due to the lower purity of this metal complex as free ligand does catalyse the decomposition of the metal complex.[34] The data presented in table 2.3 shows the decrease in UV/visible absorbance intensity in the metal to ligand charge transfer region of the spectrum for all of the metal complexes investigated at high and low concentration stored in the light and in the dark. The parent and $L_3$ cylinders are both chemically stable at high concentration and the $L_3$ cylinder is most stable at low concentration in the light. The $L_3$ cylinder is the most chemically stable metal complex in the light and in the dark and at low concentration. The results show that the concentrated aqueous metal complexes are more stable than when in a dilute solution and that there is no siginificant difference in % loss of UV/visible absorbance signal intensity when stored in the dark or light. The $L_5$ cylinder is the least chemically stable metal complex.

It was concluded from this set of experiments that the $L_3$ cylinder was the most chemically stable compound when stored in the dark and that concentration is an important factor when considering the stability of the metal complexes. It was decided that all racemic metal complex solutions would be prepared on the day of the experiment in a concentrated solution and stored at room temperature on the bench.

| MC | DAY | DARK | | | | LIGHT | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 50 μM | % loss in S.I. | 500 μM | % loss in S. I. | 50 μM | % loss in S. I. | 500 μM | % loss in S. I. |
| (R)-$L_1$ | 1 | 0.835 | | 0.261 | | 0.835 | | 0.261 | |
| @ 573 nm | 8 | 0.014 | | 0.231 | | 0.262 | | 0.234 | |
| | 15 | 0.011 | | 0.231 | | 0.024 | | 0.231 | |
| | 22 | 0.009 | **98.9** | 0.230 | **11.9** | 0.021 | **97.5** | 0.229 | **12.3** |
| (R)-$L_3$ | 1 | 0.962 | | 0.314 | | 0.962 | | 0.314 | |
| @ 583 nm | 8 | 0.818 | | 0.311 | | 0.867 | | 0.294 | |
| | 15 | 0.770 | | 0.293 | | 0.823 | | 0.275 | |
| | 22 | 0.568 | **41.0** | 0.284 | **9.56** | 0.635 | **34.0** | 0.265 | **15.6** |
| (R)-$L_5$ | 1 | 0.699 | | 0.379 | | 0.699 | | 0.379 | |
| @ 568 nm | 8 | 0.384 | | 0.225 | | 0.015 | | 0.231 | |
| | 15 | 0.012 | | 0.222 | | 0.005 | | 0.226 | |
| | 22 | 0.007 | **99.0** | 0.219 | **42.2** | 0.004 | **99.4** | 0.225 | **40.6** |

**Table 2.3 $\lambda_{max}$ values in the MLCT region of the UV/visible absorbance spectrum for (R)-$L_1$, (R)-$L_3$ and (R)-$L_5$ at high and low concentration stored in the light and in the dark. Also shown is the % loss in S. I. signal intensity at the wavelength maximum.**

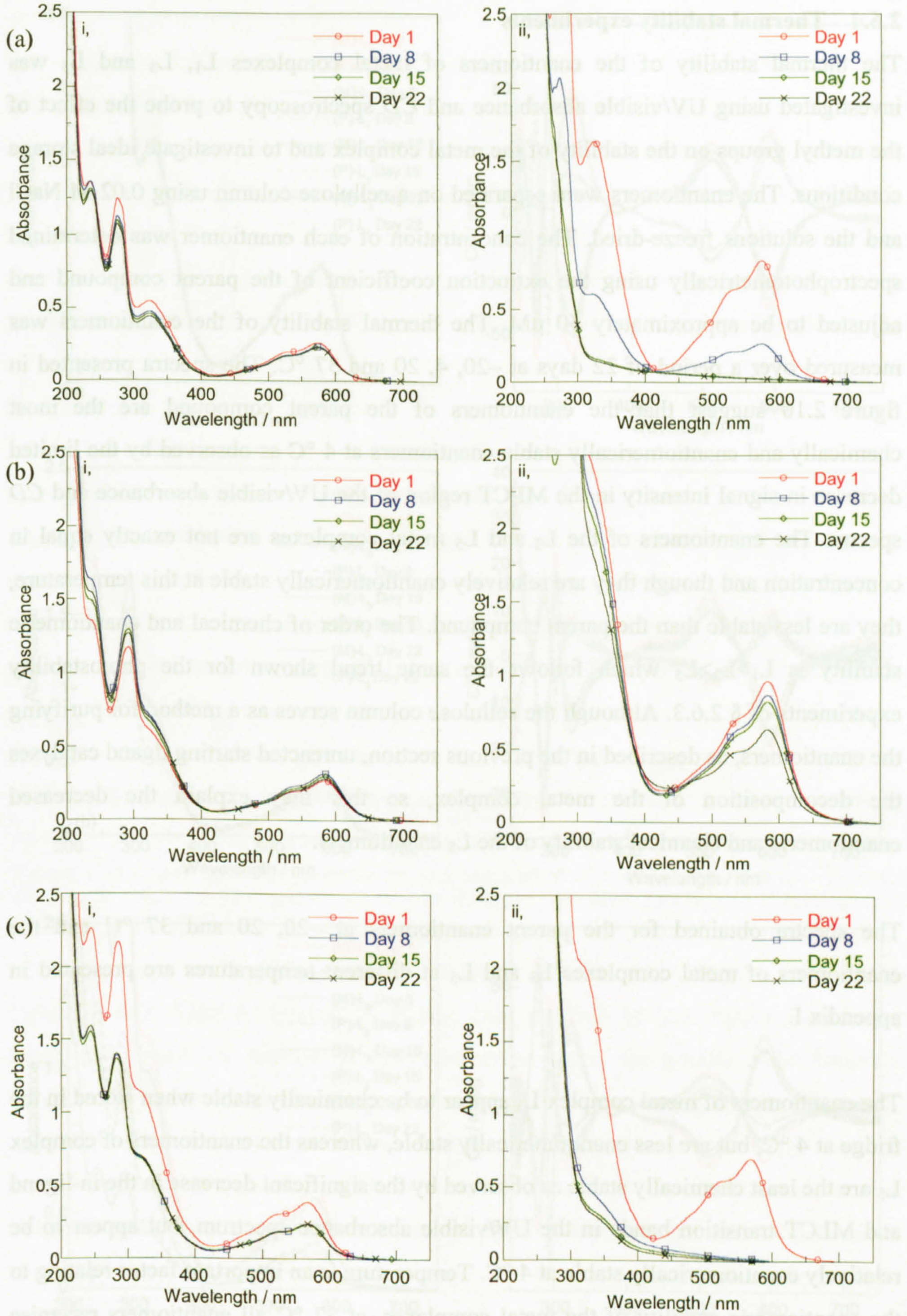



**Figure 2.9 UV/visible absorbance spectra of (a) (R)-$L_1$, (b) (R)-$L_3$ and (c) (R)-$L_5$ at i, high (500 μM) and ii, low (50 μM) concentration, stored in the light. Absorbance data above 2.5 have been omitted. A 1 in 32 dilution was made of the 500 μM solutions and the spectra collected in a 1 cm path length cuvette.**

### 2.6.4 Thermal stability experiments

The thermal stability of the enantiomers of metal complexes $L_1$, $L_3$ and $L_5$ was investigated using UV/visible absorbance and *CD* spectroscopy to probe the effect of the methyl groups on the stability of the metal complex and to investigate ideal storage conditions. The enantiomers were separated on a cellulose column using 0.02 M NaCl and the solutions freeze-dried. The concentration of each enantiomer was determined spectrophotometrically using the extinction coefficient of the parent compound and adjusted to be approximately 50 μM. The thermal stability of the enantiomers was measured over a period of 22 days at –20, 4, 20 and 37 °C. The spectra presented in figure 2.10 suggest that the enantiomers of the parent compound are the most chemically and enantiomerically stable enantiomers at 4 °C as observed by the limited decrease in signal intensity in the MLCT region of the UV/visible absorbance and *CD* spectra. The enantiomers of the $L_3$ and $L_5$ metal complexes are not exactly equal in concentration and though they are relatively enantiomerically stable at this temperature, they are less stable than the parent compound. The order of chemical and enantiomeric stability is $L_1 > L_3 > L_5$ which follows the same trend shown for the photostability experiments of § 2.6.3. Although the cellulose column serves as a method for purifying the enantiomers, as described in the previous section, unreacted starting ligand catalyses the decomposition of the metal complex, so this may explain the decreased enantiomeric and chemical stability of the $L_5$ enantiomers.

The spectra obtained for the parent enantiomers at –20, 20 and 37 °C and the enantiomers of metal complexes $L_3$ and $L_5$ at different temperatures are presented in appendix I.

The enantiomers of metal complex $L_3$ appear to be chemically stable when stored in the fridge at 4 °C, but are less enantiomerically stable, whereas the enantiomers of complex $L_5$ are the least chemically stable as observed by the significant decrease in the in-ligand and MLCT transition bands in the UV/visible absorbance spectrum, but appear to be relatively enantiomerically stable at 4 °C. Temperature is an important factor relating to the enantiomeric stability of the metal complexes, at 37 °C all enantiomers racemise within 7 days. The data suggests that the enantiomers are most stable when stored in a freezer at –20 °C. It was concluded that all enantiomeric solutions would be prepared on the day of the experiment and stored in the fridge at 4 °C during the day.

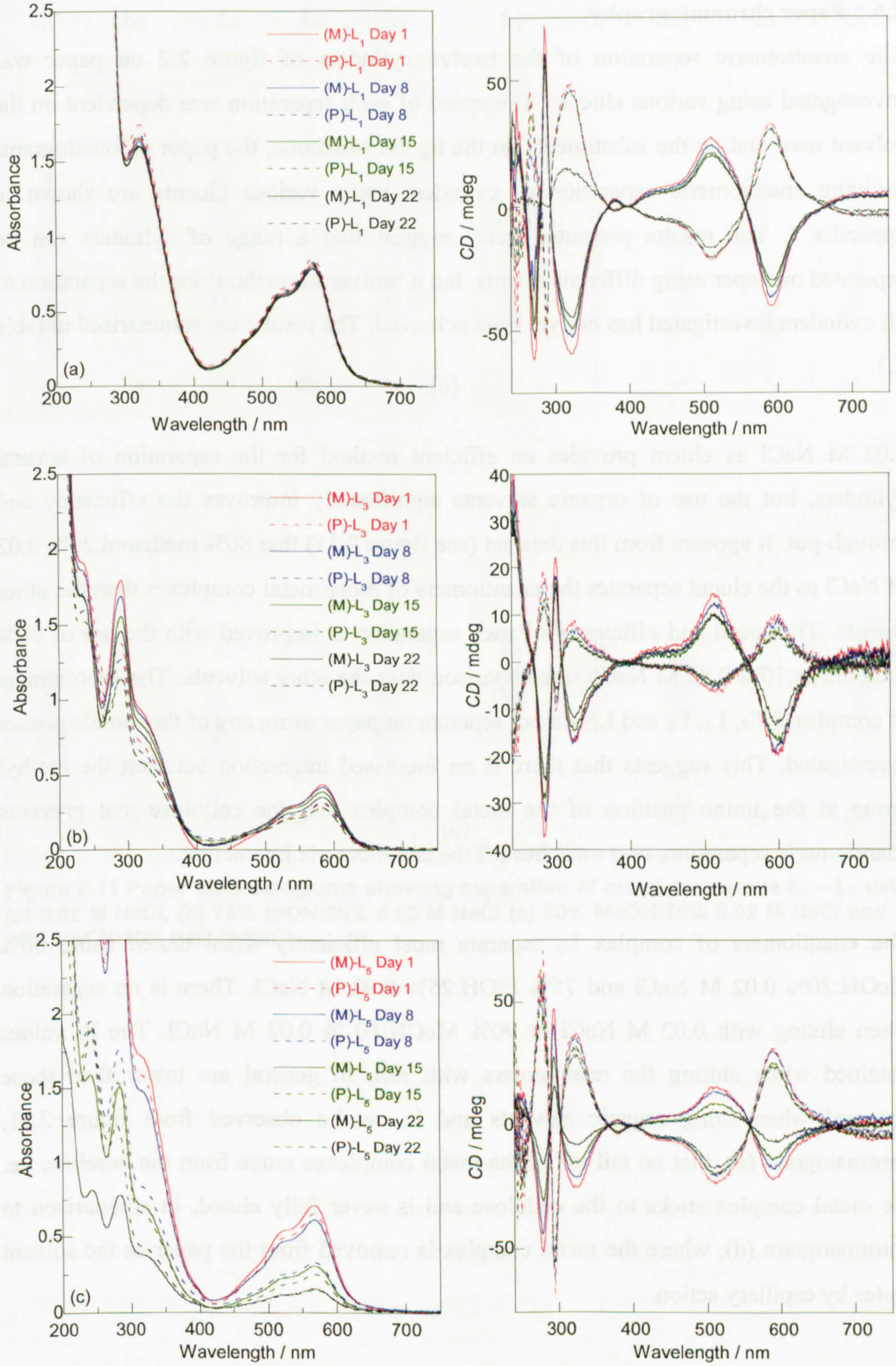



**Figure 2.10 UV/visible absorbance and *CD* spectra of (a) $L_1$ (b) $L_3$ and (c) $L_5$ stored at 4 °C over a period of 22 days. Concentration = 50 µM. Spectra collected in a 1 cm path length cuvette, water was subtracted from each sample spectrum and the data zeroed at 400 nm. Absorbance data above 2.5 have been omitted.**

### 2.6.5 Paper chromatography

The enantiomeric separation of the twelve cylinders of figure 2.2 on paper was investigated using various eluents. The speed of each separation was dependent on the solvent used and on the substituents on the ligand backbone, the paper chromatograms showing enantiomeric separation of cylinders using various eluents are shown in appendix I. The results presented here, suggest that a range of cylinders can be separated on paper using different eluents, but a 'universal method' for the separation of all cylinders investigated has not yet been achieved. The results are summarised in table 2.3.

0.02 M NaCl as eluent provides an efficient method for the separation of several cylinders, but the use of organic solvents significantly improves the efficiency and through-put. It appears from this data set (see figure 2.11) that 80% methanol:20% 0.02 M NaCl as the eluent separates the enantiomers of more metal complexes than the other eluents. The speed and efficiency of each separation is improved with the use of 90% acetonitrile:10% 0.02 M NaCl, in comparison with the other solvents. The enantiomers of complexes $L_2$, $L_6$, $L_8$ and $L_{12}$ do not separate on paper using any of the mobile phases investigated. This suggests that there is an increased interaction between the methyl group at the imine position of the metal complex and the cellulose that prevents enantiomeric separation, or it switches off the enantiomeric interaction.

The enantiomers of complex $L_7$ separate most efficiently when eluted using 80% MeOH:20% 0.02 M NaCl and 75% EtOH:25% 0.02 M NaCl. There is no separation when eluting with 0.02 M NaCl or 90% MeCN:10 % 0.02 M NaCl. The $R_f$ values obtained when eluting the enantiomers with salt, in general are lower than those obtained when using organic solvents and it can be observed from figure 2.11, chromatogram (a), that no tail off of the metal complexes move from the baseline, i.e. the metal complex sticks to the cellulose and is never fully eluted, in comparison to chromatogram (d), where the metal complex is removed from the paper as the solvent elutes by capillary action.

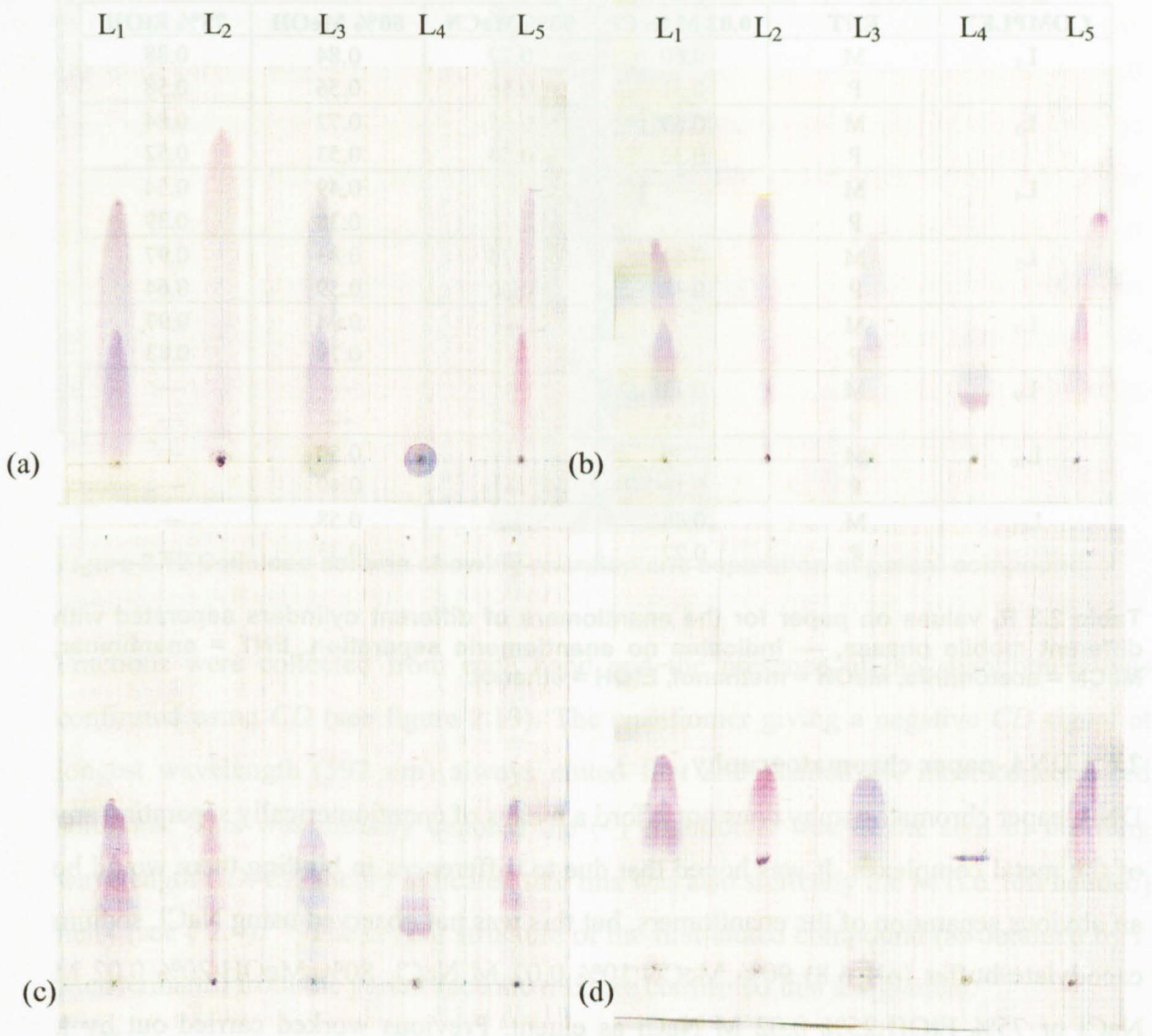



**Figure 2.11 Paper chromatograms showing separation of metal complexes $L_1$ – $L_5$ using (a) 0.02 M NaCl, (b) 75% EtOH:25% 0.02 M NaCl (c) 80% MeOH:20% 0.02 M NaCl and (d) 90% MeCN:10% 0.02 M NaCl.**

| COMPLEX | ENT | 0.02 M NaCl | 90% MeCN | 80% MeOH | 75% EtOH |
|---|---|---|---|---|---|
| $L_1$ | M | 0.60 | 0.72 | 0.84 | 0.88 |
| | P | 0.31 | 0.56 | 0.56 | 0.58 |
| $L_3$ | M | 0.63 | 0.65 | 0.72 | 0.84 |
| | P | 0.32 | 0.54 | 0.53 | 0.52 |
| $L_4$ | M | — | — | 0.49 | 0.54 |
| | P | — | — | 0.38 | 0.39 |
| $L_5$ | M | 0.64 | 0.74 | 0.83 | 0.97 |
| | P | 0.31 | 0.60 | 0.59 | 0.64 |
| $L_7$ | M | — | — | 0.94 | 0.97 |
| | P | — | — | 0.79 | 0.83 |
| $L_9$ | M | 0.73 | — | — | — |
| | P | 0.45 | — | — | — |
| $L_{10}$ | M | 0.30 | — | 0.50 | — |
| | P | 0.18 | — | 0.40 | — |
| $L_{11}$ | M | 0.48 | — | 0.58 | — |
| | P | 0.27 | — | 0.33 | — |

**Table 2.3 $R_f$ values on paper for the enantiomers of different cylinders separated with different mobile phases, — indicates no enantiomeric separation. ENT = enantiomer, MeCN = acetonitrile, MeOH = methanol, EtOH = ethanol.**

### 2.6.6 DNA-paper chromatography

DNA-paper chromatography does not afford a means of enantiomerically separating any of the metal complexes. It was hoped that due to differences in binding there would be an obvious separation of the enantiomers, but this was not observed using NaCl, sodium cacodylate buffer (pH 6.8) 90% MeCN:10% 0.02 M NaCl, 80% MeOH:20% 0.02 M NaCl or 75% EtOH:25% 0.02 M NaCl as eluent. Previous worked carried out by I. Meistermann,[35] had showed an element of separation but it may have been due to free cellulose sites and not enantiomeric resolution by the DNA.

### 2.6.7 Cellulose column chromatography

The paper chromatography technique was scaled up to incorporate the use of cellulose packed in a column and the results are presented below. The cellulose column (see figure 2.12) simultaneously chemically purifies (orangey brown band at the top of the column) and enantiomerically separates (two purple bands) a range of cylinders and although a variety of eluents were investigated a 'universal method' for the separation of all cylinders has not yet been achieved as was the case with the paper chromatography.

**Figure 2.12 Cellulose column showing enantiomeric separation of parent compound.**

Fractions were collected from each band and the presence of the enantiomers was confirmed using *CD* (see figure 2.13). The enantiomer giving a negative *CD* signal at longest wavelength (592 nm) always eluted first and yielded the most concentrated solutions. This was initially denoted the (–) enantiomer due to the sign of the long wavelength *CD*. *CD* theory indicated that this was also sterically the M (i.e. left handed) helix (see § 2.4).[36] The crystal structure of the first eluted compound (as obtained by I. Meistermann) from the parent racemic mixture confirmed this assignment.[35]

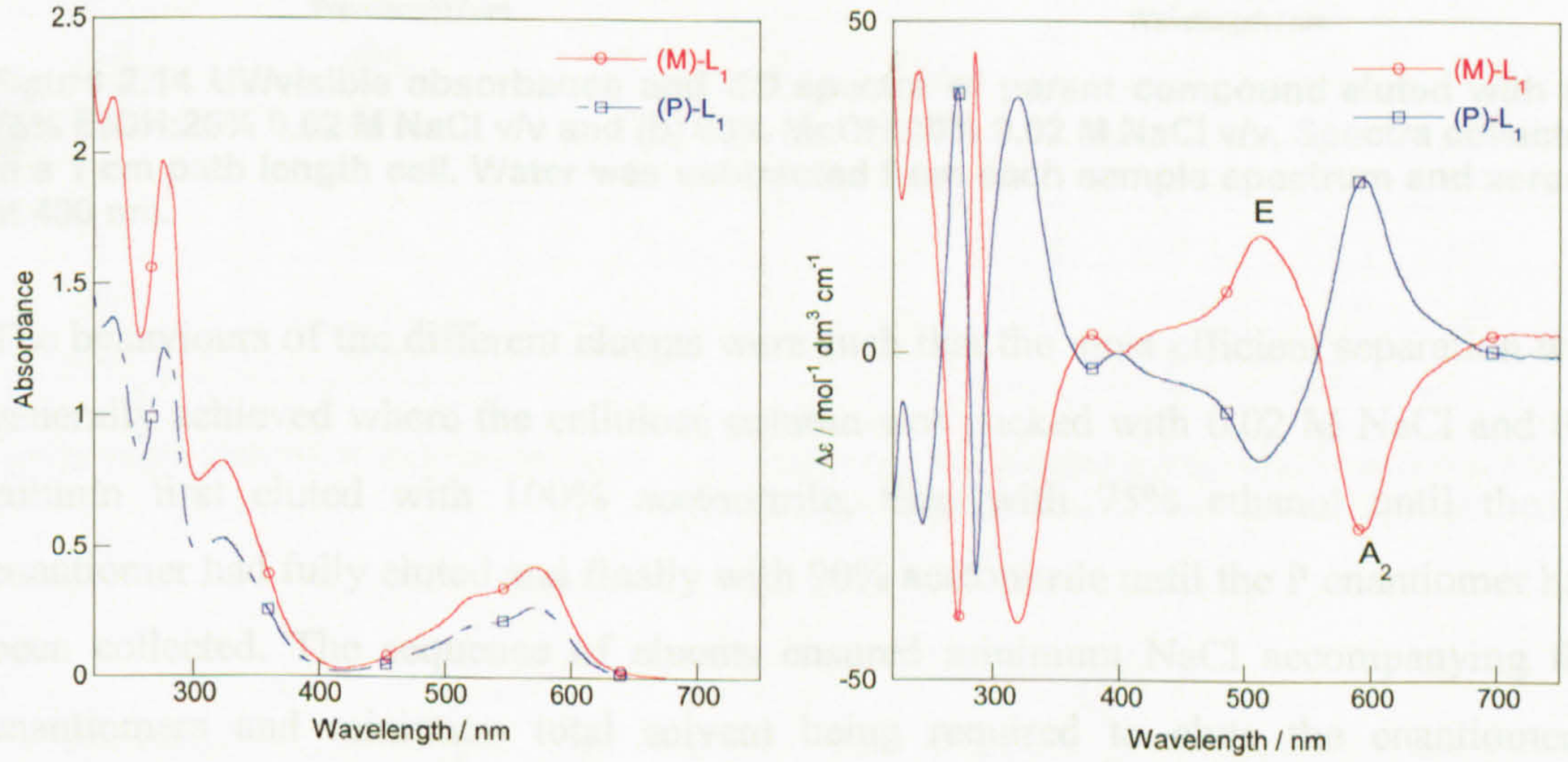



**Figure 2.13 UV/visible absorbance and *CD* spectra of the parent cylinder eluted with 90% MeCN:10% 0.02 M NaCl v/v. Spectra collected in a 1 cm path length cell. Water was subtracted from each sample spectrum and zeroed at 400 nm.**

Aqueous sodium chloride eluent (0.02 M) provided a suitable and efficient separation method for the parent compound, but when applied to the separation of the other metal complexes it gave less satisfactory resolution, with rather dilute solutions (and hence

high NaCl content) due to slow elution. A range of different eluents were investigated. 0.02 M NaCl proved to be the minimum viable salt concentration for effective separation of the parent compound. Addition of an organic solvent to the aqueous NaCl eluted the components from the column more quickly than the fully aqueous eluent, thus reducing the time taken and producing more concentrated solutions of the enantiomers with less NaCl present. The optimum mobile phase for the parent compound was found to be 90% MeCN:10% 0.02 M NaCl v/v (figure 2.13). The separation was fast and clearly visible on the column, the enantiomers were eluted in highly concentrated aqueous solution and due to the high organic content the amount of NaCl in the solution was drastically reduced. Ethanol/aqueous 0.02 M NaCl 75:25 v/v and methanol/aqueous NaCl 80:20 v/v were also effective eluents (figure 2.14). However, different variations were required for the enantiomeric separation of the other compounds (see appendix I).

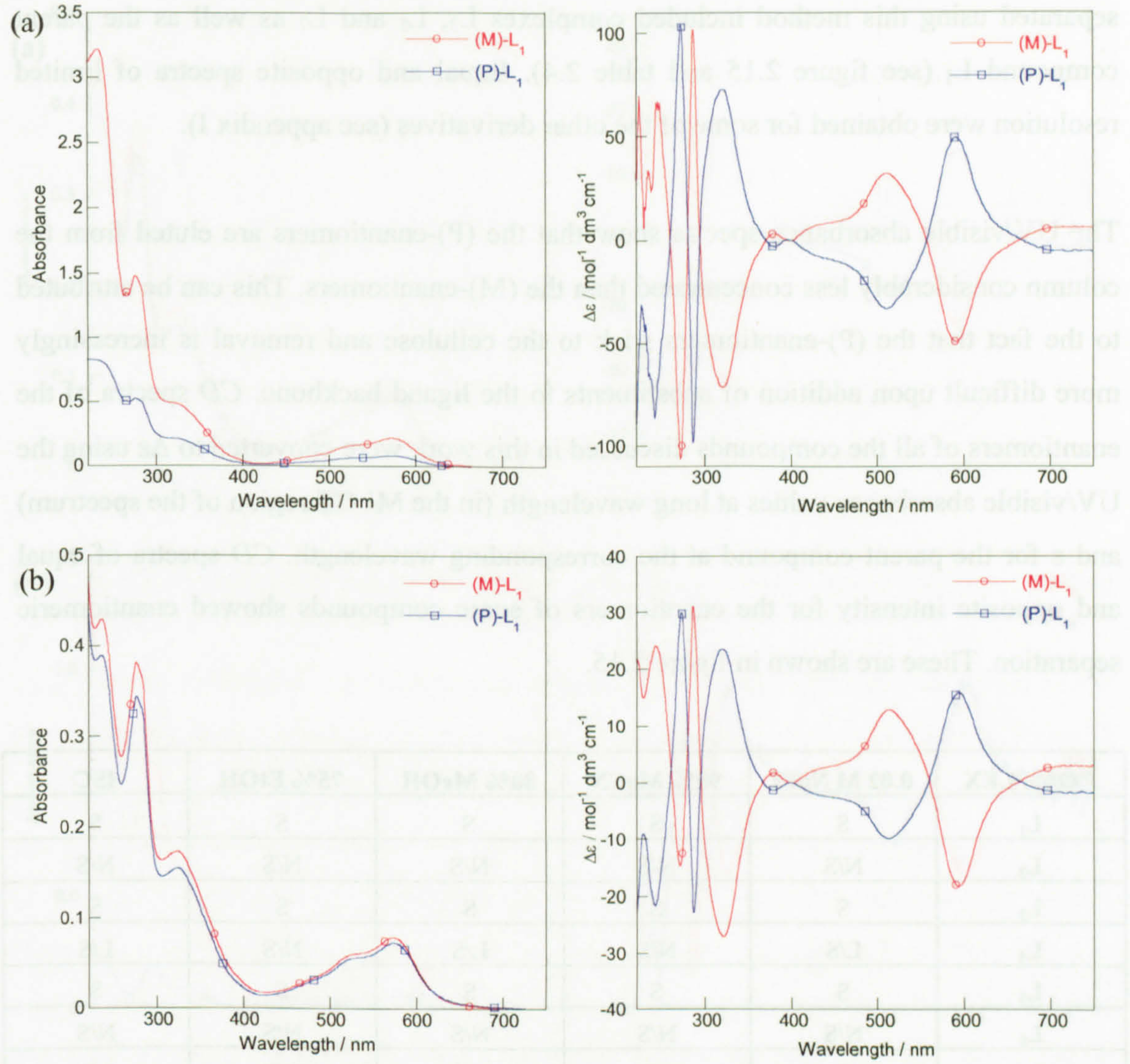



**Figure 2.14 UV/visible absorbance and *CD* spectra of parent compound eluted with (a) 75% EtOH:25% 0.02 M NaCl v/v and (b) 80% MeOH:20% 0.02 M NaCl v/v. Spectra collected in a 1 cm path length cell. Water was subtracted from each sample spectrum and zeroed at 400 nm.**

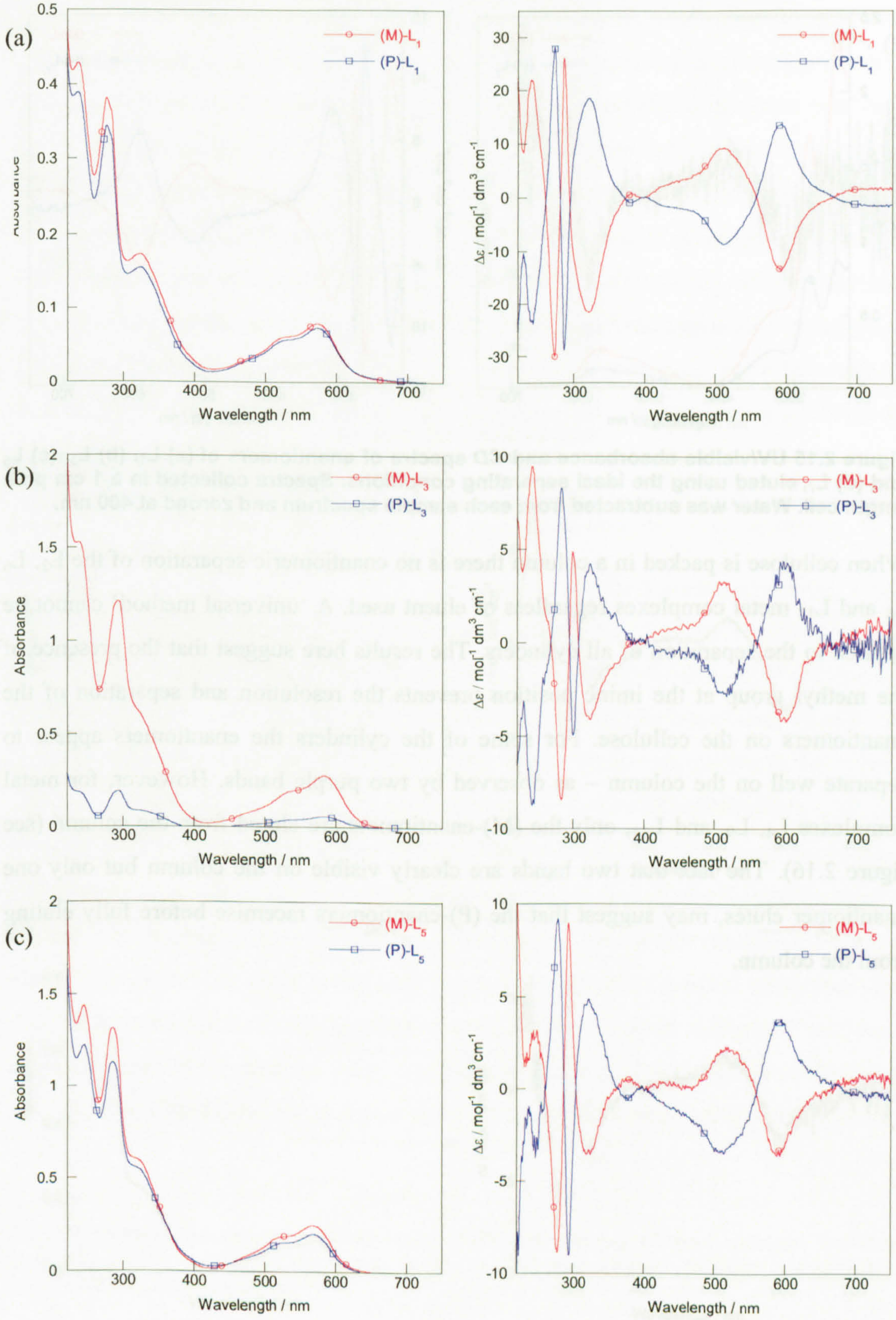
The behaviours of the different eluents were such that the most efficient separation was generally achieved where the cellulose column was packed with 0.02 M NaCl and the column first eluted with 100% acetonitrile, then with 75% ethanol until the M enantiomer had fully eluted and finally with 90% acetonitrile until the P enantiomer had been collected. The sequence of eluents ensured minimum NaCl accompanying the enantiomers and minimum total solvent being required to elute the enantiomers, especially the second one, from the column. Acetonitrile provided fastest movement of the purple bands down the column and reduced the time taken for the enantiomers to be separated compared with the other solvents. This suggests that it decreased the interaction between the enantiomers of the metal complex and the cellulose. The ethanol then facilitates the enantiomeric resolution and the acetonitrile:NaCl mobile phase efficiently removes the P enantiomers from the column. The metal complexes that were

separated using this method included complexes $L_3$, $L_5$ and $L_7$ as well as the parent compound $L_1$ (see figure 2.15 and table 2.4). Equal and opposite spectra of limited resolution were obtained for some of the other derivatives (see appendix I).

The UV/visible absorbance spectra show that the (P)-enantiomers are eluted from the column considerably less concentrated than the (M)-enantiomers. This can be attributed to the fact that the (P)-enantiomers stick to the cellulose and removal is increasingly more difficult upon addition of substituents to the ligand backbone. *CD* spectra of the enantiomers of all the compounds discussed in this work were converted to $\Delta\varepsilon$ using the UV/visible absorbance values at long wavelength (in the MLCT region of the spectrum) and $\varepsilon$ for the parent compound at the corresponding wavelength. *CD* spectra of equal and opposite intensity for the enantiomers of some compounds showed enantiomeric separation. These are shown in figure 2.15.

| COMPLEX | 0.02 M NaCl | 90% MeCN | 80% MeOH | 75% EtOH | ISC |
|---|---|---|---|---|---|
| $L_1$ | S | S | S | S | S |
| $L_2$ | N/S | N/S | N/S | N/S | N/S |
| $L_3$ | S | S | S | S | S |
| $L_4$ | L/S | N/S | L/S | N/S | L/S |
| $L_5$ | S | S | S | S | S |
| $L_6$ | N/S | N/S | N/S | N/S | N/S |
| $L_7$ | S | L/S | N/S | N/S | L/S |
| $L_8$ | N/S | N/S | N/S | N/S | N/S |
| $L_9$ | L/S | L/S | L/S | L/S | L/S |
| $L_{10}$ | L/S | L/S | L/S | L/S | L/S |
| $L_{11}$ | N/S | L/S | N/S | N/S | L/S |
| $L_{12}$ | N/S | N/S | N/S | N/S | N/S |

**Table 2.4 Enantiomeric separation of cylinders, S = separation, L/S = limited separation, N/S = no separation. ISC = ideal separating conditions.**


(a)
(M)-L1
(P)-L1
Absorbance
Δε / mol-1 dm3 cm-1
Wavelength / nm
(b)
(M)-L3
(P)-L3
Absorbance
Δε / mol-1 dm3 cm-1
Wavelength / nm
(c)
(M)-L5
(P)-L5
Absorbance
Δε / mol-1 dm3 cm-1
Wavelength / nm

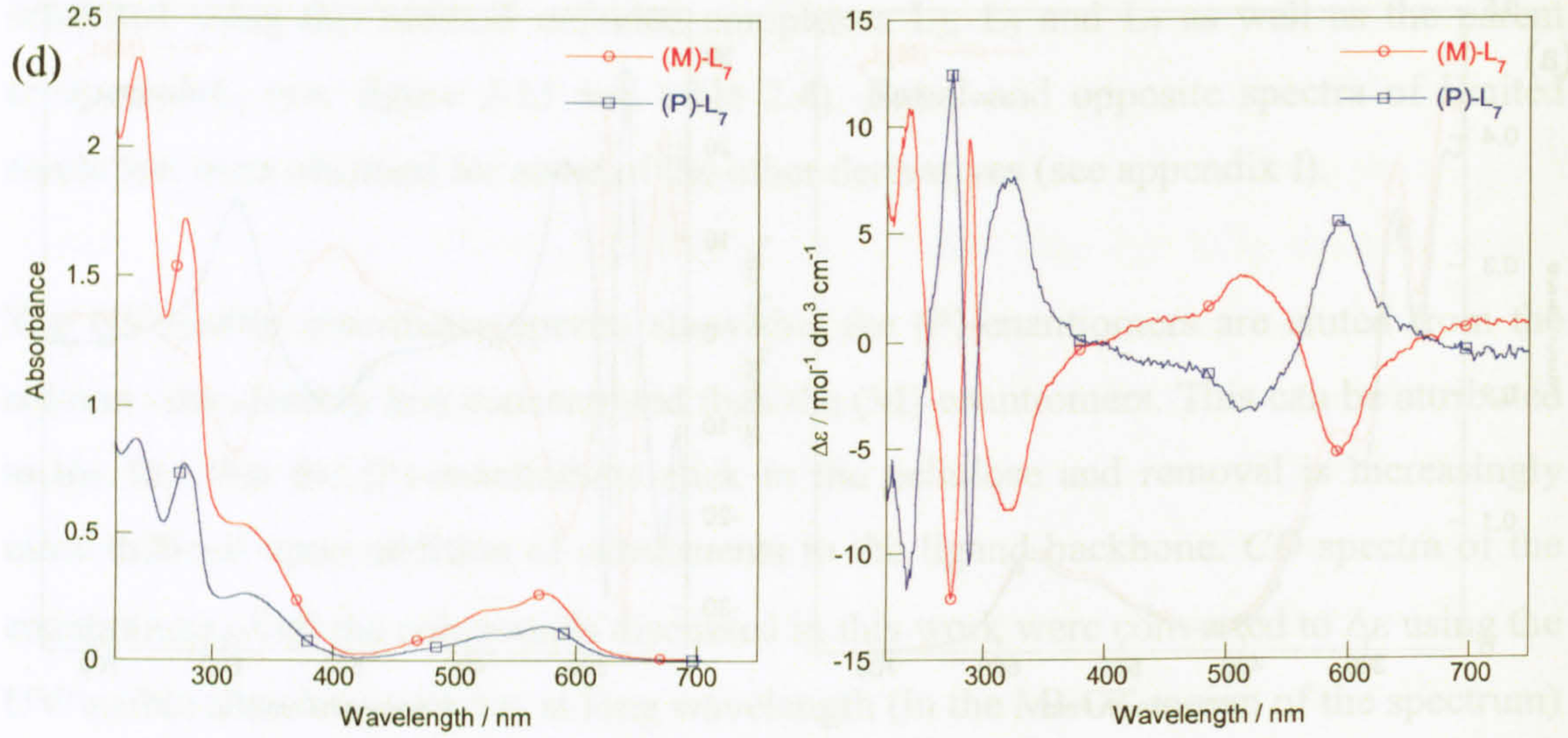



**Figure 2.15 UV/visible absorbance and *CD* spectra of enantiomers of (a) $L_1$, (b) $L_3$, (c) $L_5$ and (d) $L_7$, eluted using the ideal separating conditions. Spectra collected in a 1 cm path length cell. Water was subtracted from each sample spectrum and zeroed at 400 nm.**

When cellulose is packed in a column there is no enantiomeric separation of the $L_2$, $L_6$ $L_8$ and $L_{12}$ metal complexes regardless of eluent used. A 'universal method' cannot be applied to the separation of all cylinders. The results here suggest that the presence of the methyl group at the imine position prevents the resolution and separation of the enantiomers on the cellulose. For some of the cylinders the enantiomers appear to separate well on the column – as observed by two purple bands. However, for metal complexes $L_4$, $L_9$, and $L_{11}$, only the (M)-enantiomers are eluted from the column (see figure 2.16). The fact that two bands are clearly visible on the column but only one enantiomer elutes, may suggest that the (P)-enantiomers racemise before fully eluting from the column.

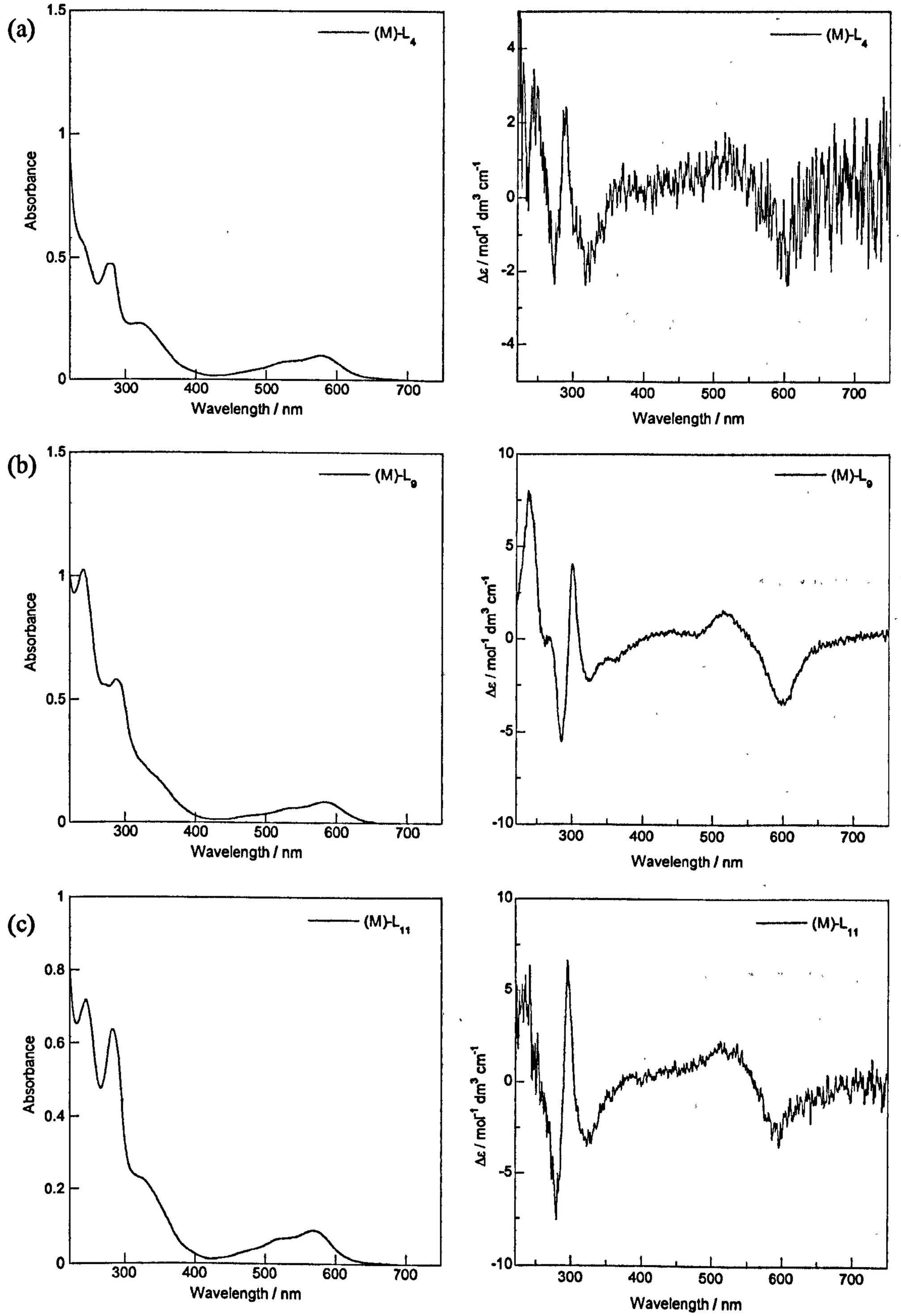



**Figure 2.16 UV/visible absorbance and *CD* spectra of (M)-enantiomers of (a) $L_4$ (b) $L_9$ and (c) $L_{11}$ eluted using the ideal separating conditions. Spectra collected in a 1 cm path length cell. Water was subtracted from each sample spectrum and zeroed at 400 nm.**

### 2.6.8 High-pressure liquid chromatography

After several attempts to develop a method for the enantiomeric separation of the parent cylinder it was concluded that reverse phase HPLC could not be used for the separation of the enantiomers. The results obtained suggested that the metal complex, when injected directly onto the relatively acidic silica $C_{18}$ column, was decomposing into its substituent components. The solution eluting from the column was collected and it was observed that the purple colouring characteristic of the bimetallo iron complexes had disappeared. This suggested that either (i) the metal complex was sticking to the column (as observed when separating the enantiomers using the cellulose column) or (ii) that the metal complex was decomposing on the column. It was concluded, by injecting the ligand onto the column and recording the retention time that the latter was more likely to be the reason for why no enantiomeric separation occurred.

### 2.6.9 High temperature docking calculations

High temperature docking calculations were used to investigate the interactions of the enantiomers of the parent cylinder, $L_1$, with a monomer of glucose (see figure 2.17).[33] MD simulations were performed with the cylinder frozen, while the glucose was mobile. As described previously (see § 2.3.2) cellulose is a natural polymer consisting of a long chain of glucose sugar molecules. It has been shown (see § 2.6.7) that the enantiomers of a range of bimetallo iron cylinders can be separated on cellulose using various eluents. It was hoped that trends would be observed in the high temperature docking calculations that would explain and re-enforce the experimental data in that the (M)-enantiomer is always the first to be eluted from the column and the (P)-enantiomer elutes last. Molecular modelling of the glucose-cylinder system revealed three binding sites with more favourable interaction energies. Two of these were with the sugar bound near one of the metal centres and the other was near the bridging $CH_2$ group.

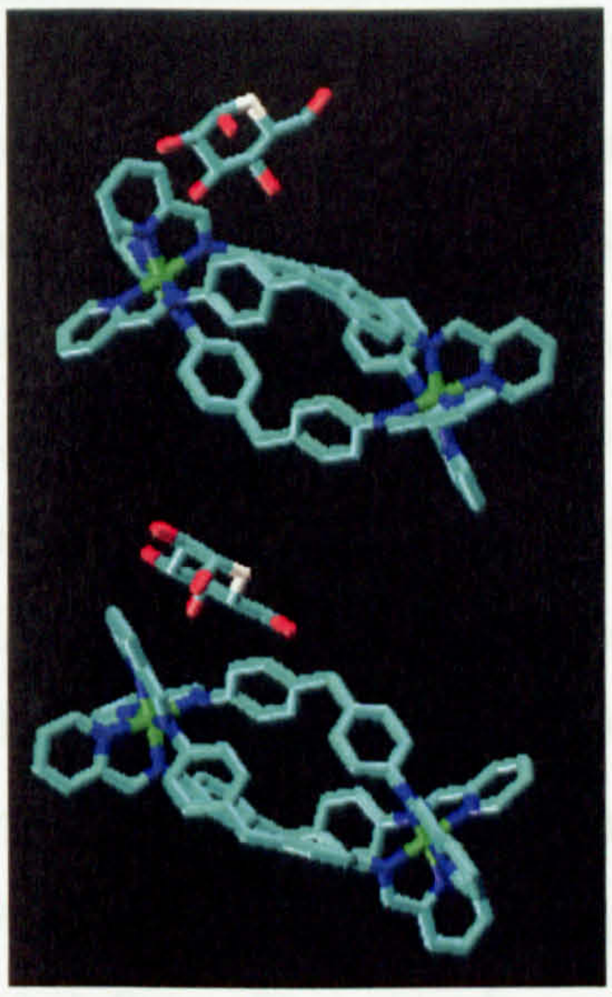


**Figure 2.17 Diagram showing interaction of the (M) and (P)-enantiomers of the parent cylinder with a monomer of glucose in favourable site 1. Top = (M)-enantiomer and bottom = (P)-enantiomer. MD simulations were performed for 1 ns at 900 K. These simulations were carried out by Dr. S. Khalid.[33]**

The data presented in table 2.5 does indeed provide evidence for why the (M)-enantiomers always elute first from the column. The data from the docking calculations shows that the interaction energy between the (M)-enantiomer and the glucose molecule is lower than that of the (P)-enantiomer and hence resolution occurs on the column because of these energy differences. Although a rather simplistic calculation, this level of sophistication is appropriate to compare to similar interaction sites.

| **SITE 1** | **Interaction Energy kcal/mol** |
|---|---|
| M | – 7.9 |
| P | – 9.0 |
| **SITE 2** | |
| M | – 7.1 |
| P | – 8.4 |
| **SITE 3** | |
| M | – 6.2 |
| P | – 11.0 |

**Table 2.5 interaction energies of the (M) and (P)-enantiomers of the parent compound $L_1$ interacting with a glucose monomer during a high temperature docking calculation in each of the favourable sites.[33]**

Molecular modelling of the glucose monomers and both (M)- and (P)-enantiomers has shown that the (P)-enantiomers interact more favourably with the glucose monomer, so would be retained on the cellulose column.

## 2.7 Discussion

The metal complexes are easily synthesised from commercially available starting materials, however purifying the complexes has proved to be a big challenge. Purification of the ligand using recrystallisation from hot ethanol is essential for the synthesis of higher purity metal complexes. The purity of the metal complex is an important factor affecting the kinetic and thermodynamic stability of the enantiomers and the determination of the extinction co-efficients for the methylated cylinders. Unreacted starting ligand was assumed to be present in the metal complex solutions that jeopardised the accurate determination of ε and even after several recrystallisations from hot ethanol and diffusion recrystallisation using benzene and acetonitrile an ε value for the methylated cylinders was not obtained. The cellulose column provides a technique for the purification and enantiomeric separation of the metal complexes, but due to the significant amount of NaCl co-eluted with the enantiomers, the ε value for each methylated cylinder could still not be obtained.

An alternative explanation for the reduced chemical stability of the methylated cylinders may be due to steric effects. The addition of the methyl group particularly in the 5'-position on the outer pyridyl ring may adversely affect the stability of the metal complex due to steric hindrance.

It has been shown that the (M)-enantiomers always elute from the column first. Looking at the ligand structures of each of the metal complexes (see figure 2.19) that were not separated on the cellulose column using any of the eluents investigated, shows that they all have the methyl group at the imine bond and the only difference between them is the bridging spacer, i.e. $CH_2$ for $L_2$ and $L_6$ and O for $L_8$ and $L_{12}$. The separation of metal complexes with the oxygen bridging moiety is generally much slower and the fractions eluting from the column are not so well resolved.

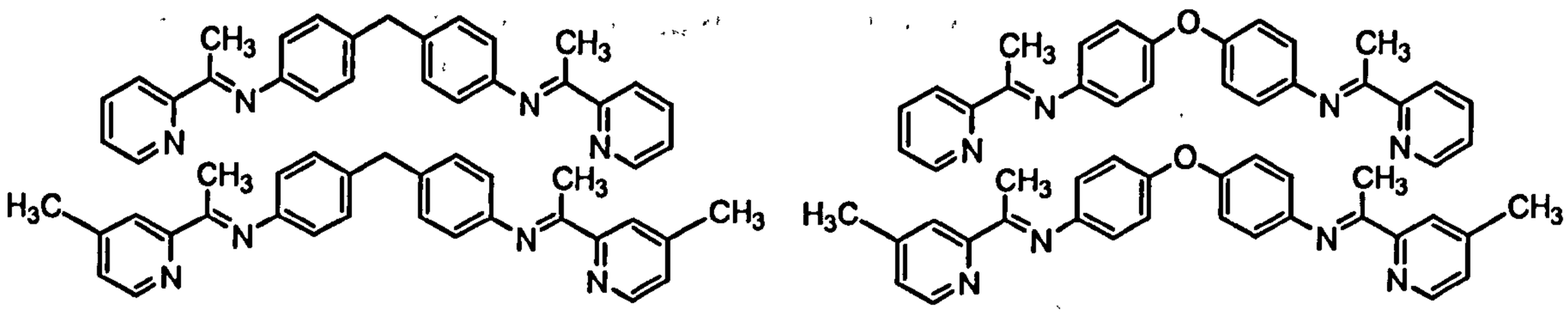


**Figure 2.18 Ligand structures of the metal complexes that cannot be separated on cellulose using any of the eluents investigated.**

In an attempt to find a universal method for the separation of all 12 cylinders, HPLC was employed. A racemic mixture of the parent compound was first injected directly onto a $C_{18}$ reverse phase column and eluted with 90% MeCN:10% $H_2O$. Several peaks were observed in the chromatogram, suggesting decomposition of the metal complex on the column. It was also observed that no 'purple' solution (characteristic of the metal complexes) eluted from the column, so this method was not further developed. It should be noted that a sample of the parent cylinder was sent to Hichrom for chromatographic analysis and it was concluded after a thorough investigation of different columns and mobile phases that the metal complex could not be separated into the two enantiomers using reverse phase high-pressure liquid chromatography. Future investigations could involve the use of chiral HPLC columns for the enantiomeric separation of the metal complexes.

The cylinder-glucose system used as a model for the investigation of interaction energy between the enantiomers of the parent cylinder and cellulose may have some fundamental drawbacks. As described in § 2.3.1 cellulose is a natural polymer consisting of a long chain of β-D-glucose sugar molecules, linked by 1,4-glucosidic bonds that form a highly ordered helical structure. In the simulations we used one glucose molecule and so the interactions between the cylinder and the glucose molecule may not be representative of the interaction between the cylinder and cellulose due to variations in hydrogen bonding and/or other electrostatic interactions. For the purposes of this research, this simplistic calculation was shown to agree with the experimental results in that the (M)-enantiomer always elutes from the column first and that the (P)-enantiomer elutes last i.e. there is a lower interaction energy between the cellulose and the (M)-enantiomers than the cellulose and the (P)-enantiomers.

## 2.8 References

1. Hannon, M. J., Painting, C. L. Jackson, A., Hamblin, J. and Errington, W., *Chem. Commun.* **1997**; 1807 – 1808

2. Hannon, M. J., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Meistermann, I., Isaac, C. J. Sanders, K. J. and Rodger, A. *Angew. Chem. Int. Ed.* **2001**; 40, 5, 880 – 884

3. Hannon, M. J., Painting, C. L. and Alcock, N. W. *Chem. Commun.* **1999**; 2023 – 2024

4. Meistermann, I., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Khalid, S., Rodger, P. M., Peberdy, J. C., Isaac, C. J., Rodger, A. and Hannon, M. J. *Proc. Natl. Acad. Sci. USA.* **2002**; 99, 8, 5069 – 5074

5. Paisner, S. N., and Bergman, R. G., *J. Organometallic Chem.* **2001**; 621, 1 – 2, 242 - 245

6. Ray, R. K., and Kauffman, G. B., *J. Chrom. A.* **1988**; 442, 381 – 385

7. Skoog, D. A. and Leary, J. J., '*Principles of Instrumental Analysis*' Fourth Ed. Saunders College Publishing, **1992**; 579

8. Holder, E., Schoetz, G., Schurig, V. and Lindner, E. *Tetrahedron Asymmetry* **2001**; 12, 2289 – 2293

9. Holder, E., Trapp, G., Grimm, J. C., Schurig, V. and Lindner, E., *Tetrahedron Asymmetry*, **2002**; 13, 2673 – 2678

10. Armstrong, D. W., DeMond, W. and Czech, B. P., *Anal. Chem.* **1985**; 57, 481 – 484

11. Green, J. M., Jones, R., Harrison, R. D., Edwards D. S. and Glajch, J. L., *J. Chromatography,* **1993**; 635, 203 – 209

12. Yamanari, K. and Nakamichi, M., *J. Chem. Soc. Commun.* **1989**; 1723 – 1724

13. Aldrich-Wright, J. R., Greguric, I., Vagg, R. S., Vickery, K. and Williams, P. A.,*J. Chromatography A,* **1995**; 718, 436 – 443

14. Baker, D. Morgan, R. J. and Strekas, T. C., *J. Am. Chem. Soc.* **1991**; 4, 113, 1411 – 1412

15. Pellegrini, P. P. and Aldrich-Wright, J. R., *J. Chem. Soc., Dalton Trans.*, **2003**; 176 – 183

16. Fletcher, N. C., Junk, P. C., Reitsma, D. A. and Keene, F. R., *J. Chem. Soc. Dalton Trans.* **1998**; 133 – 138

17. Hannon, M. J., Meistermann, I. Isaac, C., Blomme, C., Aldrich-Wright, J. and Rodger, A., *Chem. Commun.*, **2001**; 1078 – 1079

18. Vietor, R. J., Newman, R.H., Ha, M.-A., Apperley, D. C. and Jarvis, M. C., *The Plant Journal,* **2002**; 30. 6 721 – 731

19. Nishiyama, Y., Sugiyama, J., Chanzy, H. and Langan, P., *J. Am. Chem. Soc.*, **2003**; 125, 14300 – 14306

20. Nishiyama, Y., Langan, P. and Chanzy, H., *J. Am. Chem. Soc.*, **2002**; 124, 9074 – 9082

21. Jarvis, M., *Nature*, **2003**; 426, 611 – 612

22. http://spm.phy.bris.ac.uk/research/cellulose

23. Rodger, A. and Nordén, B., *Circular Dichroism and Linear Dichroism,* Oxford University Press, Oxford **1997**

24. Schipper, P. E., Rodger A. *Journal of the American Chemical Society*, **1983**; 105, 4541 – 4550

25. Peberdy, J. C., Reudegger, V., Kerchoffs, J., Meistermann, I., Khalid, S., Rodger, P. M., Alcock, N. W., Hannon, M. J. and Rodger A. *'in press'* **2005**

26. Bosnich, B.. *Acc. Chem. Res,* **1969**; 2, 266 – 273

27. Rodger, A., Nordén, B., Ardhammar, M. and Lincoln, P., *Chem. Phys. Lett.* **2002**; 354, 44 – 50

28. Rodger, A., Nordén, B., Rodger, P. M. and Bates, P. J., *Eur. J. Inorg. Chem.* **2002**; 49 – 53

29. Windscheif, P. M. and Vogtle, F. *Synthesis,* **1994**; 87 – 92

30. Coggan, Z. Haworth, I. S., Bates, P. J., Robinson, A. and Rodger, A., *Inorg. Chem.*, **1999**; 38, 4486 – 4497

31. Allen, M. P., Tildersley, D. J., *Computer Simulation of Liquids*, **1987**; Oxford University Press, Oxford, UK

32. Mac Kerell, A. D. Jnr., Wiórkiewicz-Kuczera, J. and Karplus, M., *J. Am. Chem. Soc.* **1995**; 117, 11946 – 11975

33. Khalid, S. Hannon, M. J., Rodger A. and Rodger, P. M., *in press*, **2005**

34. Kerckhoffs, J. M. C. A., Benito-Lopez, F., Reinhoudt, D. N. and Hannon, M. J., *submitted* **2004**

35. Meistermann, I. '*DNA major groove recognition by supramolecular helicates*' Ph.D. thesis, University of Warwick **2001**

36. Rodger, A. '*personal communication*' **2004**

# CHAPTER 3 – DNA BINDING STUDIES

## 3.1 Introduction

The bimetallo iron cylinders (see § 2.2) which are the primary focus of this research, are chiral compounds. This is an important property for their interaction with DNA as the helicity of the complex has been found to have a significant effect on the DNA binding properties of the enantiomers. Circular and linear dichroism, atomic force microscopy (AFM) (see figure 3.1) and NMR studies had been previously used to show that the left handed or (M)-enantiomer of the parent cylinder does indeed provide major groove recognition and induce remarkable structural effects, wrapping up double stranded DNA in an intramolecular manner.[1 – 4] The binding mode of the (P)-enantiomers remains unclear. To probe the DNA-cylinder interactions responsible for the groove binding mode of the cylinder, in this work hydrophobic methyl groups were placed at different positions on the ligand backbone. The way in which these affect the DNA binding will provide further insights into the binding modes of the two enantiomers of the parent cylinder. This in turn will enable the control and fine tuning of the DNA binding with the ultimate aim of introducing sequence selective substituents onto the backbone of the ligand.

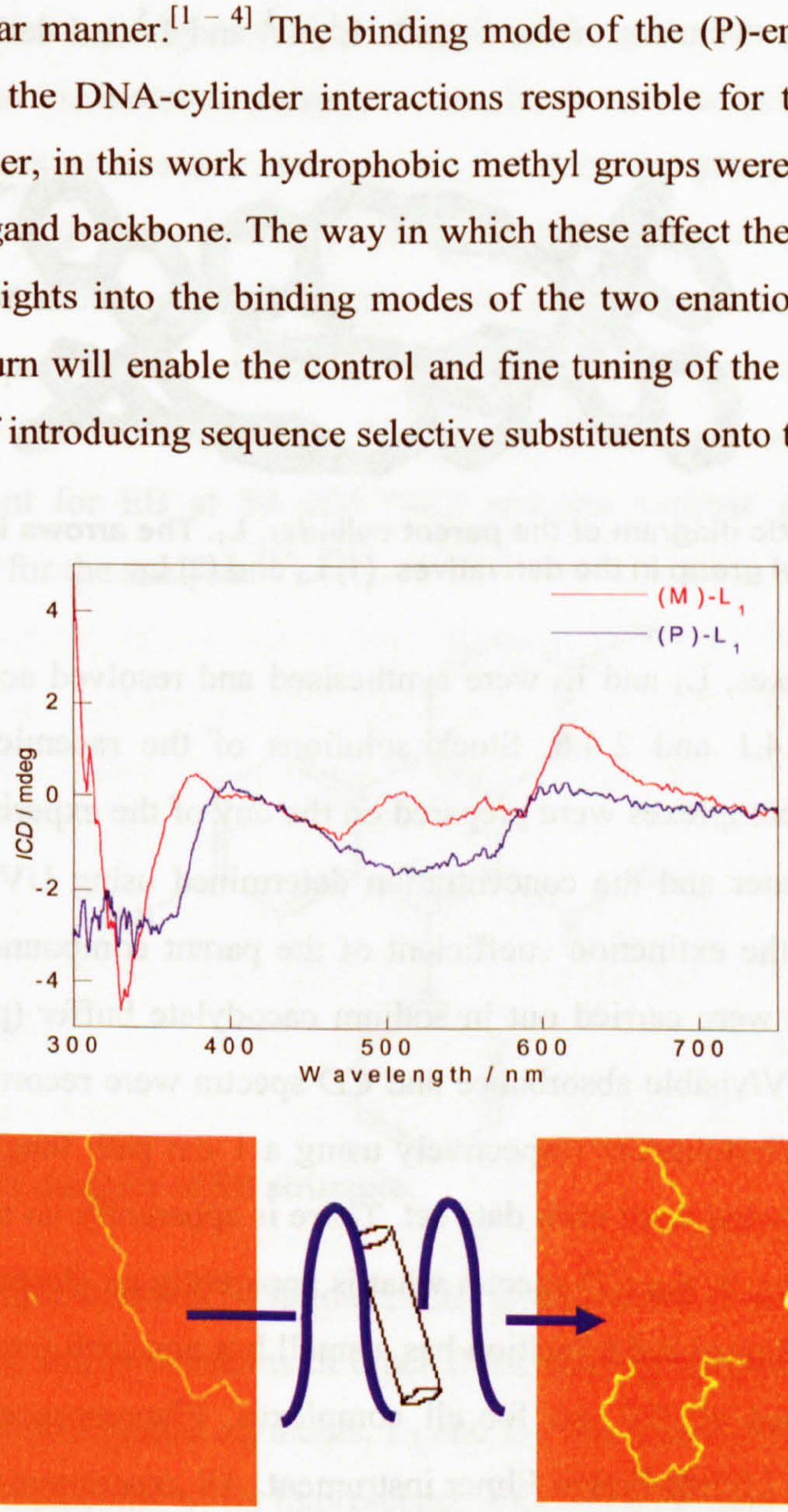



**Figure 3.1 *ICD* spectra of the (M)- and (P)-enantiomers of the parent cylinder $L_1$, binding to ct-DNA at a ratio of 30 DNA:1 metal complex. DNA concentration = 600 μM and metal complex concentration = 20 μM. AFM images of (M)-$L_1$ wrapping up double stranded DNA.[4]**

The aims of the research presented in this chapter were to (i) determine the ct-DNA binding constants for the enantiomers of the parent compound $[Fe_2(L^1)_3]^{4+}$, $L_1$ and for the enantiomers of cylinders $L_3$ and $L_5$ (see figure 3.2) using an ethidium bromide competitive binding assay and (ii) probe the effect of hydrophobic substituents on the DNA binding of the methylated cylinders. UV/visible absorbance, circular and linear dichroism were used to determine whether there are differences between the DNA binding of the parent and methyl substituted cylinders and whether DNA sequence affects binding. It was hoped that sequence selectivity upon binding to DNA might be observed. The $Fe_2$ cylinders of the ligands $L^1$, $L^3$ and $L^5$ are denoted $L_1$, $L_3$ and $L_5$ throughout.

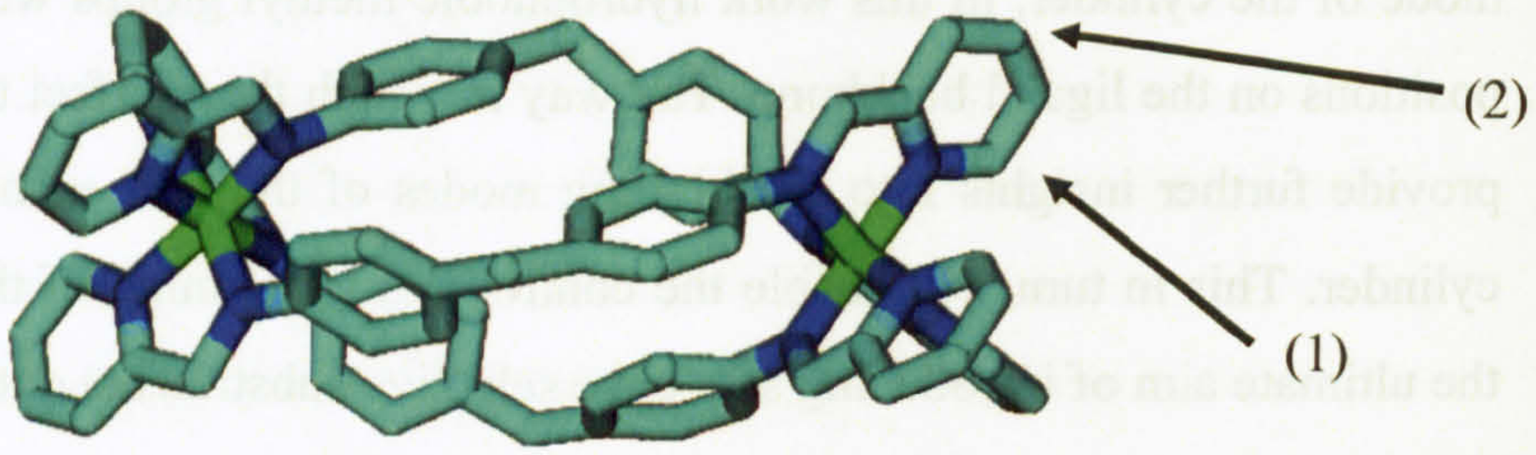



**Figure 3.2 Schematic diagram of the parent cylinder, $L_1$. The arrows indicate the position of the added methyl group in the derivatives. (1) $L_3$ and (2) $L_5$.**

The metal complexes, $L_3$ and $L_5$ were synthesised and resolved according to methods described in § 2.4.1 and 2.4.6. Stock solutions of the racemic and enantiomeric methylated metal complexes were prepared on the day of the experiment from the solid chloride salt in water and the concentration determined using UV/visible absorbance spectroscopy and the extinction coefficient of the parent compound (see § 2.4.2). All DNA experiments were carried out in sodium cacodylate buffer (pH 6.8, 1 mM) and NaCl (50 mM). UV/visible absorbance and *CD* spectra were recorded on Jasco V-550 and J-715 spectrophotometers respectively using a 1 cm path length cuvette. A water baseline was subtracted from each data set. There is apparently no absorbance intensity at 750 nm, however, in the *CD* spectra what is apparently an electric dipole forbidden, magnetic dipole allowed d-d transition has a small but non-zero intensity. Thus the *CD* spectra were zeroed at 420 nm for all complexes. Fluorescence experiments were carried out using a LS50B Perkin Elmer instrument. All experiments were performed in duplicate to ensure reproducibility.

To determine the binding constants of the metal complexes, it was first necessary to determine the binding constant for ethidium bromide at 50 mM NaCl concentration. The

binding of EB to DNA has been extensively studied[8 – 12] and the binding constant, $K_e$ of EB has been determined as $9.5 \times 10^6$ $M^{-1}$ at 0.01 M ionic strength using fluorescence spectrophotometry at low EB to nucleotide ratios[13] and $2.5 \times 10^5$ $M^{-1}$ at 0.2 M ionic strength using linear dichroism.[14] A wide range of values with and without experimental justification can also be found.[15 – 17] Ethidium bromide is a positively charged polycyclic aromatic compound, which binds to DNA by inserting itself between the base pairs. In addition to a large increase in fluorescence of the EB molecule upon binding to DNA, the process of intercalation causes lengthening and untwisting of the DNA helix by around 26°.[8] The binding strength is in part due to $\pi$–$\pi$ interactions between the aromatic moieties of the bases and the intercalator. However, for most intercalators electrostatic interactions between the phosphates of the DNA backbone and cationic groups of the intercalator dominate the binding energy.[9] For this work, we found it difficult to obtain entirely reproducible data using the fluorescence method. The *LD* approach of Nordén *et al.*[14] was found to be more reliable and broadly consistent with the fluorescence results. Thus linear dichroism was used to determine the binding constant for EB at 50 mM NaCl and the intrinsic and Scatchard plot methods were used for the analysis.[18 – 20]

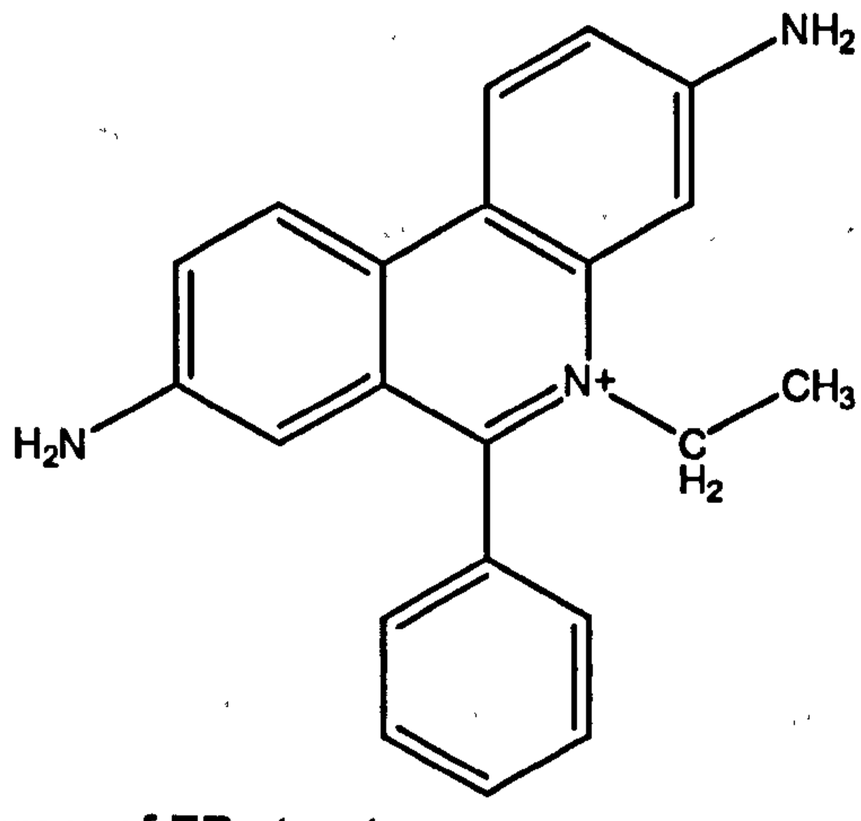



**Figure 3.3 Schematic diagram of EB structure.**

Once the EB binding constant is known, EB displacement assays can be used to estimate the binding ability/constants of other DNA binding species such as the parent cylinder, $L_1$ and two methylated cylinders, $L_3$ and $L_5$. The displacement of the ethidium bromide leads to a quenching of its fluorescence which can be used as a direct measure of the binding of the non-fluorescent iron cylinder. However, it should be noted that EB-EB interactions on the DNA can affect the fluorescence intensity as might EB-ligand interactions or ligand-induced DNA conformational changes. The DNA-binding of the parent compound had been previously investigated by I. Meistermann at a salt

concentration of 20 mM NaCl and the binding constant of the parent cylinder was estimated to be in excess of $10^7$ $M^{-1}$.[21] The experiments reported here were undertaken at higher ionic strength as it was observed that for the ligands with hydrophobic methyl groups on the ligand backbone, addition of the metal complex to DNA caused precipitation of the DNA at quite low ligand loading. The salt concentration was therefore increased to 50 mM NaCl — this was sufficient to keep the DNA in solution in the presence of the bimetallo iron cylinders.

A series of spectroscopic titration experiments keeping either the DNA or metal complex concentration constant were undertaken by adding the DNA to a solution of water, metal complex, salt and buffer (in that order). These experiments were designed to investigate the effect of constant metal complex concentration on a decreasing concentration of DNA.

### 3.1.1 Data analysis

Scatchard analysis is a method used to linearise data from a saturation binding experiment in order to determine binding constants. The Scatchard equation describes the binding of a molecule to a polymer containing N sites, each of which binds the molecule independently and with the same binding constant K when the binding interaction is non-cooperative. The fluorescence data were analysed by assuming that:

$$E_b = \frac{E_{tot}(F - F_{free})}{(F_b - F_{free})}$$

where $E_{tot}$ is the total EB concentration in solution, $E_b$ is the concentration of bound EB, $F$ is the fluorescence of a solution of EB and DNA at known concentration, $F_{free}$ is the fluorescence of the same concentration of free EB in the same buffer and salt, $F_b$ is the fluorescence of the same concentration of bound EB under the same conditions (i.e. excess DNA). A Scatchard plot from the data in the curved portion of the binding curve (i.e. where neither nearly all EB is bound nor all sites are occupied) was then used to give $K_e$ and $n_e$, the binding association constant for EB and the number of bases occupied by each bound EB respectively:

$$\frac{r}{E_f} = K_e - rK \quad \frac{}{n_e}$$

where $r$ is the binding ratio (bound EB concentration divided by DNA concentration in bases) and $E_f$ is the concentration of free EB.

The *LD* data were analysed by using the intrinsic method as described in ref 17. In this case the starting point is that

$$E_b = \alpha\rho$$

where $\rho$ is the *LD* signal of the EB induced upon binding to DNA. One then plots,

$$y = \frac{E^k_{tot} - E^j_{tot}}{\rho^k - \rho^j}$$

versus

$$x = \frac{\frac{E^k_{tot}}{\rho^k} - \frac{E^j_{tot}}{\rho^j}}{\rho^k - \rho^j}$$

where the superscripts denote two different EB concentrations at fixed DNA concentration. The slope of this plot equals [DNA]/($n_e\alpha$) and the y-intercept is α. This then yields $E_b$ and hence $E_f$ which can be used in a Scatchard plot to determine $K_e$ (see figure 3.6 (c)).

### 3.2 Materials

Calf thymus DNA (ct-DNA) was purchased from Sigma Aldrich. Synthetic double stranded DNA alternating homopolymers, poly[d(A-T)$_2$] (AT) and poly[d(G-C)$_2$] (GC) were obtained from Pharmacia Biochemicals. Ultrapure water (18.2 ΩM) was used in all experiments. All polynucleotides were dissolved in water without any further purification and kept frozen until the day of the experiment. Polynucleotide concentrations were determined spectrophotometrically using the molar extinction coefficients:[5, 6] ct-DNA $\varepsilon_{260}$ = 6,600 cm$^{-1}$ mol$^{-1}$ dm$^3$, AT-DNA $\varepsilon_{262}$ = 6,600 cm$^{-1}$ mol$^{-1}$ dm$^3$ and GC-DNA $\varepsilon_{254}$ = 8,400 cm$^{-1}$ mol$^{-1}$ dm$^3$. A 100 mM stock solution of sodium cacodylate buffer was prepared by mixing 50 mL of 0.2 M sodium cacodylate (4.28 g of $Na(CH_2)_2AsO_2.3H_2O$ in 100 mL) with 9.3 mL of 0.2 M hydrochloric acid, and diluting to a total volume of 200 mL. Ethidium bromide was purchased from Sigma-Aldrich and dissolved in water and used without any further purification. The concentration was

determined spectrophotometrically using the molar extinction coefficient $\varepsilon_{480}$ = 5,600 $cm^{-1}\ mol^{-1}\ dm^3$.[7]

### 3.2.1 UV/Visible absorbance spectroscopy of DNA and the metal complexes

DNA has a characteristic absorbance band at 260 nm, from which its concentration can be calculated using the Beer Lambert law. The absorbance spectra of the bimetallo iron cylinders are somewhat more complex than that of DNA (see figure 3.4). They possess characteristic bands due to different types of transitions. The band at approximately 580 nm is characteristic of the iron (II) metal → ligand charge transfer (MLCT) transitions. The band observed around 320 nm is due to the imine bond in-ligand transitions. Absorbance bands in the DNA region of the spectrum are also in-ligand transitions.

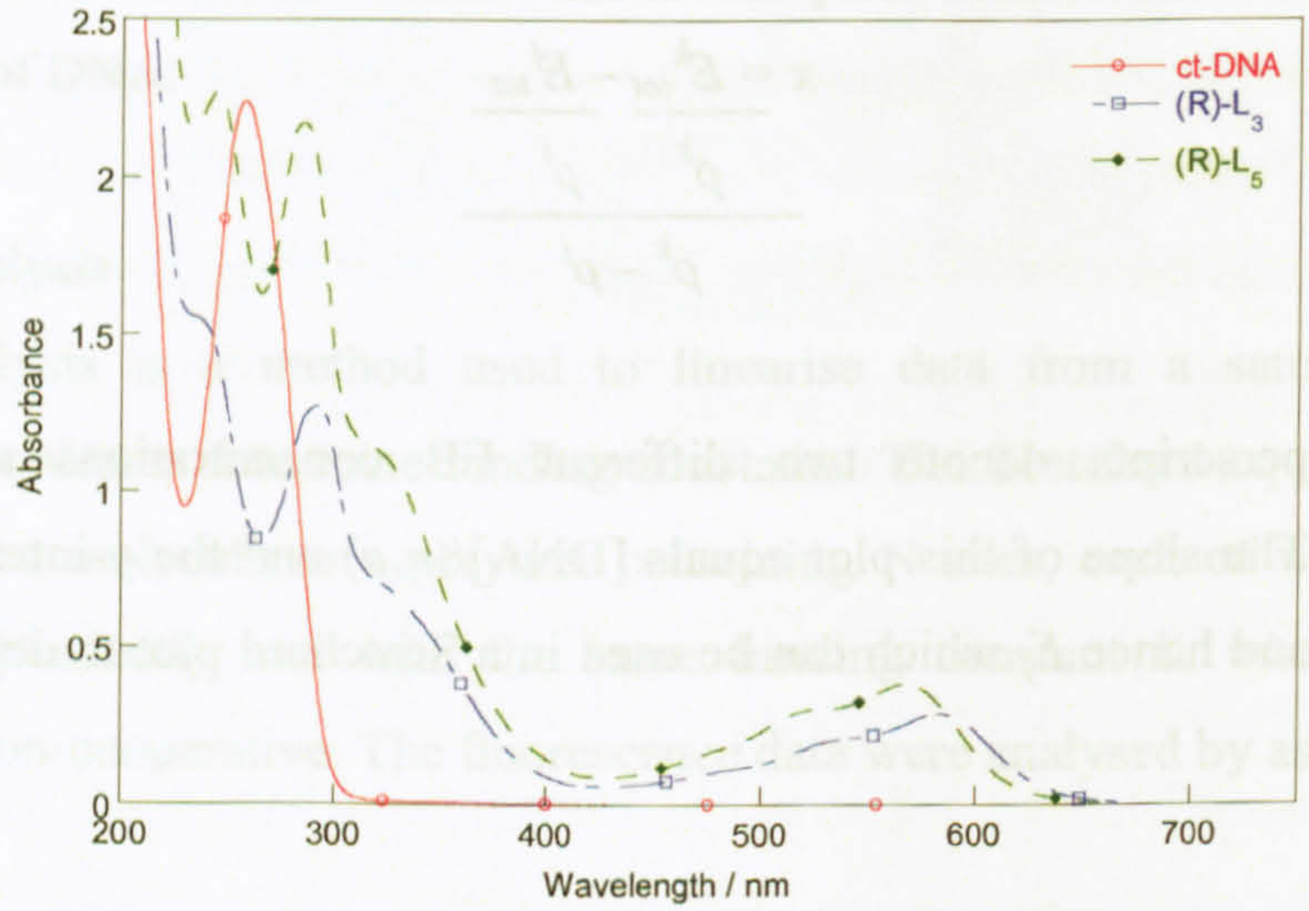



**Figure 3.4 UV/visible absorbance spectra of ct-DNA (200 μM), (R)-$L_3$ (50 μM) and (R)-$L_5$ (50 μM). Solid chloride salt of metal complexes dissolved in water and concentration determined from accurately weighed masses and molecular weight of each metal complex. Spectra collected in a 1 cm path length cuvette.**

UV/visible absorbance spectra of the cylinders with and without DNA were recorded according to the method described in § 3.3.3. UV/visible absorbance and *CD* measurements were recorded keeping the metal complex concentration constant and *LD* measurements were recorded keeping the DNA concentration constant. For the racemic cylinders, on addition of DNA, any *CD* observed above 300 nm, may be taken to show an interaction of the metal complex with DNA. The DNA has a characteristic absorbance pattern from 220 nm to 300 nm and no signal from 300 nm to 750 nm. The enantiomers of the metal complexes have a characteristic absorbance pattern below 750 nm (see § 2.3) and so to observe the induced circular dichroism (*ICD*) which is typically a few millidegrees in intensity, the *CD* of the enantiomer is subtracted from the *CD* of the enantiomer bound to DNA, to give the *ICD*.

## 3.3 Methods

### 3.3.1 Determination of EB binding constant at 50 mM NaCl

A fluorescence experiment was designed to determine the binding constant of EB at constant concentration to ct-DNA. Independent solutions (3.0 mL) of ct-DNA (10, 20, 40, 80, 200 μM), EB (3 μM), NaCl (50 mM) and buffer (1 mM) were prepared. An *LD* titration experiment was designed to investigate the binding of ethidium bromide to ct-DNA at constant concentration. A 2.0 mL solution of ct-DNA (200 μM), NaCl (50 mM) and buffer (1 mM) was prepared. Ethidium bromide (200 μM) was added to give a resulting concentration in solution of between 5 – 50 μM. A stock solution (1 mL) containing DNA, (400 μM), NaCl, (100 mM) and buffer (2 mM) was used to maintain constant DNA concentration upon addition of the EB. Several experiments were completed using different batches of EB at different starting concentrations to ensure reproducibility. Experiments were performed in duplicate.

LS50B parameters: emission: 600 nm; excitation: 480 nm; scan range: 750 – 500 nm; excitation slit: 10.0 nm; emission slit: 15.0 nm; scan speed 500 nm/min; accumulation: 3 scans. *LD* parameters: scan range: 600 – 200 nm; step resolution: 0.5 nm; speed: 500 nm/min; response: 1.0 s; band width: 1.0 nm; sensitivity: 20 mdeg; accumulation: 4 scans.

### 3.3.2 Determination of metal complex binding constants

A 3.0 mL solution of DNA (12 μM), ethidium bromide (EB) (15 μM) NaCl (50 mM) and buffer (1 mM) was prepared. The UV/visible absorbance and emission spectra were recorded, and the metal complex ((R)- (M)- and (P)-$L_1$, $L_3$ and $L_5$) concentration was slowly increased from DNA:metal complex ratios of 75:1 to 1.5:1, keeping the DNA and EB concentrations constant. After each addition, the fluorescence and UV/visible absorbance spectra were recorded.

UV/visible absorbance parameters: scan range: 750 – 200 nm; scan speed: 200 nm/min; response: fast; bandwidth: 2.0 nm. LS50B parameters: emission: 600 nm; excitation: 480 nm; scan range: 750 – 500 nm; excitation slit: 10.0 nm; emission slit: 15.0 nm; scan speed 500 nm/min; accumulation: 3 scans.

### 3.3.3 DNA binding studies

Stock solutions of NaCl (500 mM) and sodium cacodylate buffer (100 mM) were prepared. A 1.6 mL solution of DNA (600 μM), metal complex (20 μM), NaCl (50 mM) and buffer (1 mM) was prepared at a starting ratio of DNA: metal complex, 30:1. A metal complex stock solution (25 mL) containing metal complex (20 μM), NaCl (50 mM) and buffer (1 mM) was added to provide a solution of specified DNA concentration. Table 3.1 shows the volume of stock solution required to dilute the DNA to a given concentration. Solutions of each enantiomer were stored in the fridge at 4 °C to prevent racemisation and the laboratory air conditioning was set to 18 °C. All experiments were performed in duplicate. Titrations were carried out using ct-, AT- and GC-DNA with (R)- (M)- and (P)-$L_3$ and $L_5$. UV/visible absorbance and *CD* measurements were recorded at each DNA concentration and experiments performed in duplicate.

| DNA:MC | [DNA] / μM | $V_{MC}$ / μL |
|---|---|---|
| 30:1 | 600 | 0 |
| 25:1 | 500 | 320 |
| 20:1 | 400 | 480 |
| 18:1 | 360 | 267 |
| 16:1 | 320 | 333 |
| 14:1 | 280 | 429 |
| 12:1 | 240 | 571 |
| 10:1 | 200 | 800 |
| 8:1 | 160 | 1200 |
| 6:1 | 120 | 2000 |
| 4:1 | 80 | 4000 |
| 2:1 | 40 | 12000 |

**Table 3.1 $V_{MC}$, the volume of metal complex stock solution (20 μM) required to provide a solution of specified DNA concentration (MC = metal complex).**

UV/visible absorbance parameters were: scan range: 750 – 200 nm; scan speed: 200 nm/min; response: fast; band width: 2.0 nm; data pitch: 0.5 nm; accumulation: 1. *CD* parameters were: scan range: 750 – 300 nm; scan speed: 200 nm/min; step resolution: 0.5 nm; response: 1.0 s; band width: 2.0 nm; sensitivity: 20 mdeg; accumulation: 16 scans.

#### 3.3.3.1 *LD* binding studies

The *LD* experiment was designed to investigate the effect of increasing metal complex concentration on constant DNA concentration solutions. Stock solutions of NaCl (500 mM) and sodium cacodylate buffer (100 mM) were prepared. A 2.0 mL solution of DNA (500 μM), NaCl (50 mM) and buffer (1 mM) was used to show the orientation of

the DNA with no metal complex present. A metal complex stock solution (approx. 300 μM, 1 mL, determined using the extinction co-efficient of the parent compound and the Beer Lambert law) was prepared. A stock solution (1 mL) containing DNA, (1 mM), NaCl, (100 mM) and buffer (2 mM) was used to maintain constant DNA concentration upon addition of the metal complex. Table 3.2 shows the volume of metal complex and DNA stock solution required for increasing metal complex concentration. Titrations were carried out to investigate the binding of (R)-, (M)- and (P)-$L_3$ and $L_5$ to ct-, AT- and GC-DNA. The highest metal complex loading investigated was 50 μM as it was observed that concentrations of metal complex higher than this value caused precipitation of the DNA.

| DNA:MC | $V_{init}$ | [MC] / μM | [DNA] / μM | $V_{MC}$ | $V_{tot}$ | $V_{MC\ s.s}$ | $V_{DNA\ s.s}$ |
|---|---|---|---|---|---|---|---|
| | 2000 | 0 | 500 | 0 | 2000 | | |
| 500:1 | | 1.00 | 500 | 6.71 | 2013.42 | 6.71 | 6.71 |
| 400:1 | | 1.25 | 500 | 8.40 | 2016.81 | 1.69 | 1.69 |
| 300:1 | | 1.67 | 500 | 11.24 | 2022.47 | 2.83 | 2.83 |
| 200:1 | | 2.50 | 500 | 16.95 | 2033.90 | 5.71 | 5.71 |
| 100:1 | | 5.00 | 500 | 34.48 | 2068.97 | 17.53 | 17.53 |
| 80:1 | | 6.25 | 500 | 43.48 | 2086.96 | 9.00 | 9.00 |
| 60:1 | | 8.33 | 500 | 58.82 | 2117.65 | 15.35 | 15.35 |
| 40:1 | | 12.50 | 500 | 90.91 | 2181.82 | 32.09 | 32.09 |
| 20:1 | | 25.00 | 500 | 200.00 | 2400.00 | 109.09 | 109.09 |
| 10:1 | | 50.00 | 500 | 500.00 | 3000.00 | 300.00 | 300.00 |

**Table 3.2 Volumes of metal complex (300 μM), $V_{MC\ s.s}$ and DNA stock solution (1 mM), $V_{DNA\ s.s}$ required in *LD* experiments. $V_{MC\ s.s}$ and $V_{DNA\ s.s}$ are the volumes required to be added at each titration step. (MC = metal complex, V = volume required to achieve desired concentration, s.s = stock solution, $V_{Init}$ = initial volume, $V_{tot}$ = total volume in the *LD* cell).**

*LD* parameters were: scan range: 750 – 200 nm; step resolution: 0.5 nm; speed: 500 nm/min; response: 1.0 s; band width: 1.0 nm; sensitivity: 20 mdeg; accumulation: 4 scans. Experiments were performed in duplicate.

## 3.4 Results

### 3.4.1 Determination of EB binding constant at 50 mM NaCl

The DNA-binding constant of ethidium bromide at 50 mM NaCl concentration was determined using both fluorescence (see figure 3.5) and linear dichroism (see figure 3.6). The value of this number is a very important parameter for determining the binding strengths of the bimetallo iron cylinders.

The EB fluorescence was measured at different concentrations of ct-DNA and the data at 590 nm used for the analysis (see figure 3.5). The fluorescence analysis (see figure 3.5) gives $K_e \sim 6.4\ (\mu M)^{-1}$ with a site size of approximately 6 bases. This value is based

on three data points and thus is not very accurate. The experiment was repeated several times and the results obtained were reproducible but an important factor to note was that free ethidium bromide which has a low intrinsic fluorescence had a higher fluorescence intensity at 590 nm than when all of the EB had been displaced from the DNA and the metal complex had bound, so free EB fluorescence > DNA-MC + free EB fluorescence.

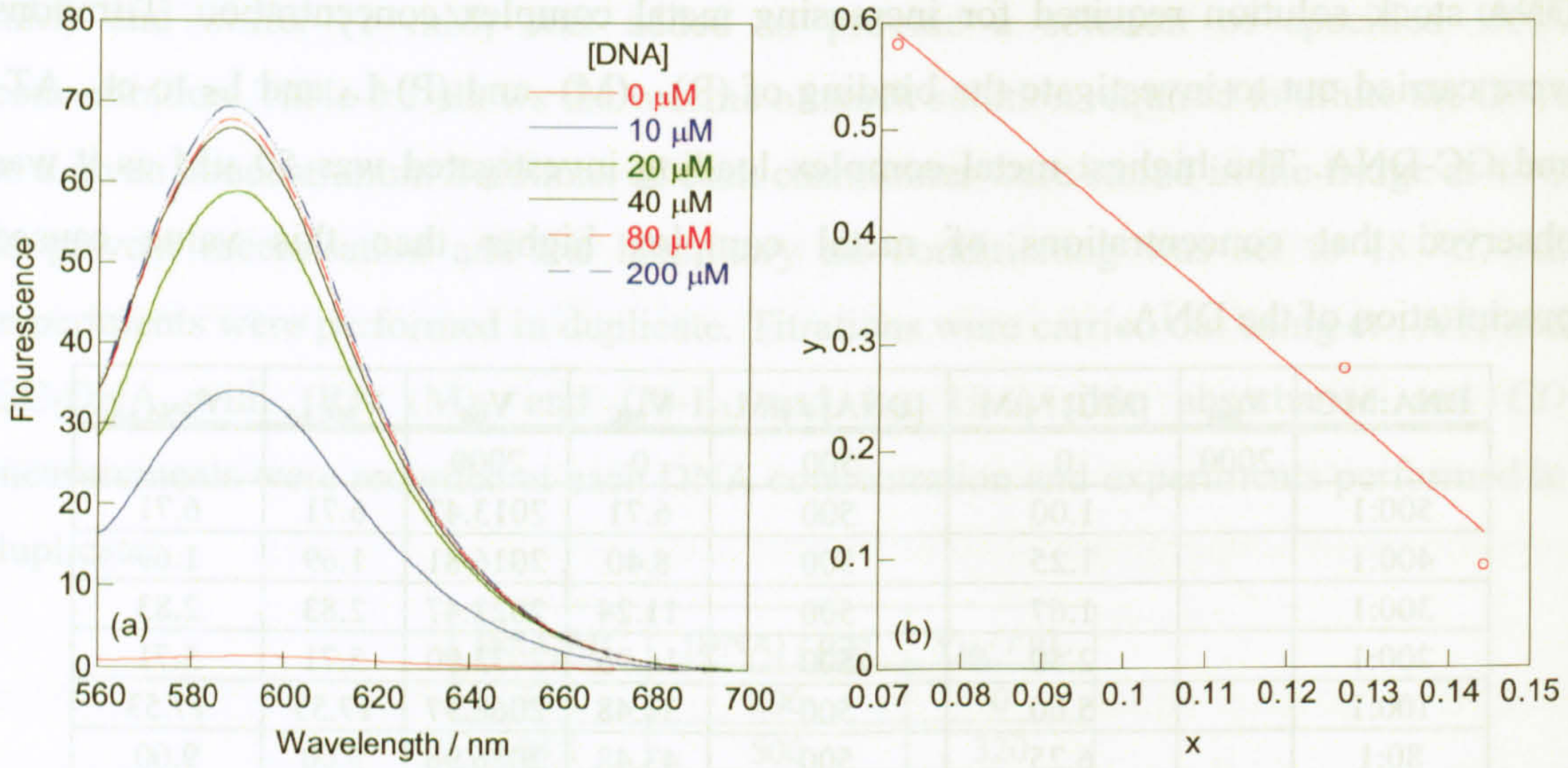



**Figure 3.5 (a) Fluorescence spectra of EB (3 μM) in the presence of ct-DNA at concentrations indicated in the figure. NaCl = 50 mM, buffer = 1 mM. The $\lambda_{max}$ value was averaged over 20 data points (580 – 600 nm). The EB fluorescence spectrum was subtracted from all DNA-EB spectra and the data zeroed at 700 nm. (b) Determination of EB binding constant using the Scatchard plot method, x = $E_b$/[DNA], y = [($E_b$/[DNA])/$E_f$], $E_b$ and $E_f$ = ligand (EB) bound and ligand free respectively, [DNA] = DNA concentration. The EB binding constant determination is based on three data points.**

The *LD* spectrum at constant DNA concentration (see figure 3.6 (a)) shows that upon addition of EB, the DNA signal at 260 nm increases, suggesting stiffening of the DNA. A signal is also observed that has a negative maximum at the 520 nm absorbance of EB which is characteristic of intercalating compounds. The EB *LD* experiment was repeated several times (see figure 3.6) and it was concluded that the binding constant of EB to ct-DNA at 50 mM NaCl concentration is 5.4 (± 1.0) × $10^5$ $M^{-1}$ and the binding site size, n, is 4.5 (± 0.5) bases.

Previous studies investigating the binding of EB to ct-DNA have shown that the mode of binding is via intercalation and the binding is non-cooperative.[8 – 12, 14] The results presented here suggest a co-operative binding mode, but since the binding of EB to ct-DNA has been extensively studied and confirmed as binding via non-cooperative intercalation, the Intrinsic and Scatchard methods[18 – 20] will be used to determine the binding constants of EB and the metal complexes and not a method that incorporates co-operativity such as the Mc Ghee–von Hippel method.[22, 23]

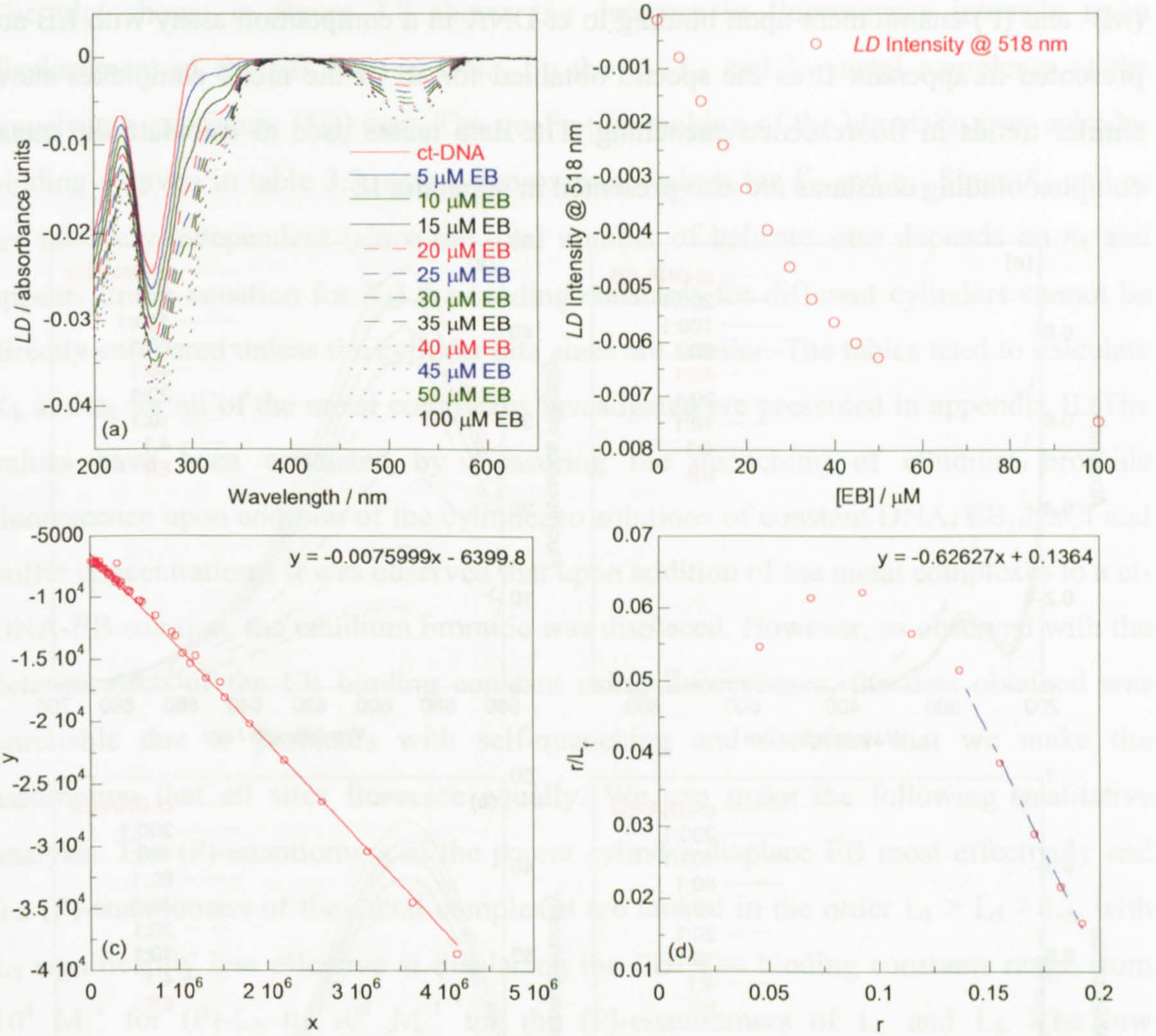



**Figure 3.6 (a) *LD* spectra of ct-DNA (200 μM) upon addition of EB at concentration indicated in figure. (b) plot of EB concentration versus *LD* intensity at 518 nm (c) Determination of EB binding constant using the Intrinsic plot method, $x = [(E^k_{tot}/\rho^k - E^j_{tot}/\rho^j)/(\rho^k - \rho^j)]$, $y = [(E^k_{tot} - E^j_{tot})/(\rho^k - \rho^j)]$. $[DNA]/(n_e\alpha) = -0.008$ and $\alpha = -6400$, $n_e = 4.1$. $E^k$ and $E^j$ refer to the concentration of EB, $\rho^k$ and $\rho^j$ refer to the *LD* signal at the corresponding EB concentration, $S_{tot}$ is the total DNA concentration which is constant at 200 μM, α is a function of wavelength which is constant over the range of binding ratios being investigated and $n_e$ is the binding site size. (d) Determination of EB binding constant using the Scatchard plot method and α from (c). The DNA concentration is constant at 200 μM. The first five data points were omitted from the binding constant calculation.**

### 3.4.2 Determination of metal complex binding constants

EB has little intrinsic fluorescence, but upon binding to ct-DNA the fluorescence is significantly enhanced. If the EB fluorescence decreases upon addition of the metal complex then the cylinder is (i) competing with the EB for binding sites or (ii) upon binding to the DNA is displacing the EB and hence quenching the fluorescence. The fluorescence spectra (see figure 3.7) show that all of the cylinders at low concentration displace EB, so their binding constants are all higher than that of EB at 50 mM NaCl concentration. The order of binding strength for the racemic metal complexes is (R)-$L_1$ ≥ (R)-$L_5$ > (R)-$L_3$. The UV/visible absorbance and fluorescence spectra obtained for the

(M)- and (P)-enantiomers upon binding to ct-DNA in a competition assay with EB are presented in appendix II as the spectra obtained for all of the metal complexes show similar trends in fluorescence quenching. The data tables used to calculate the metal complex binding constants are also presented in appendix II.

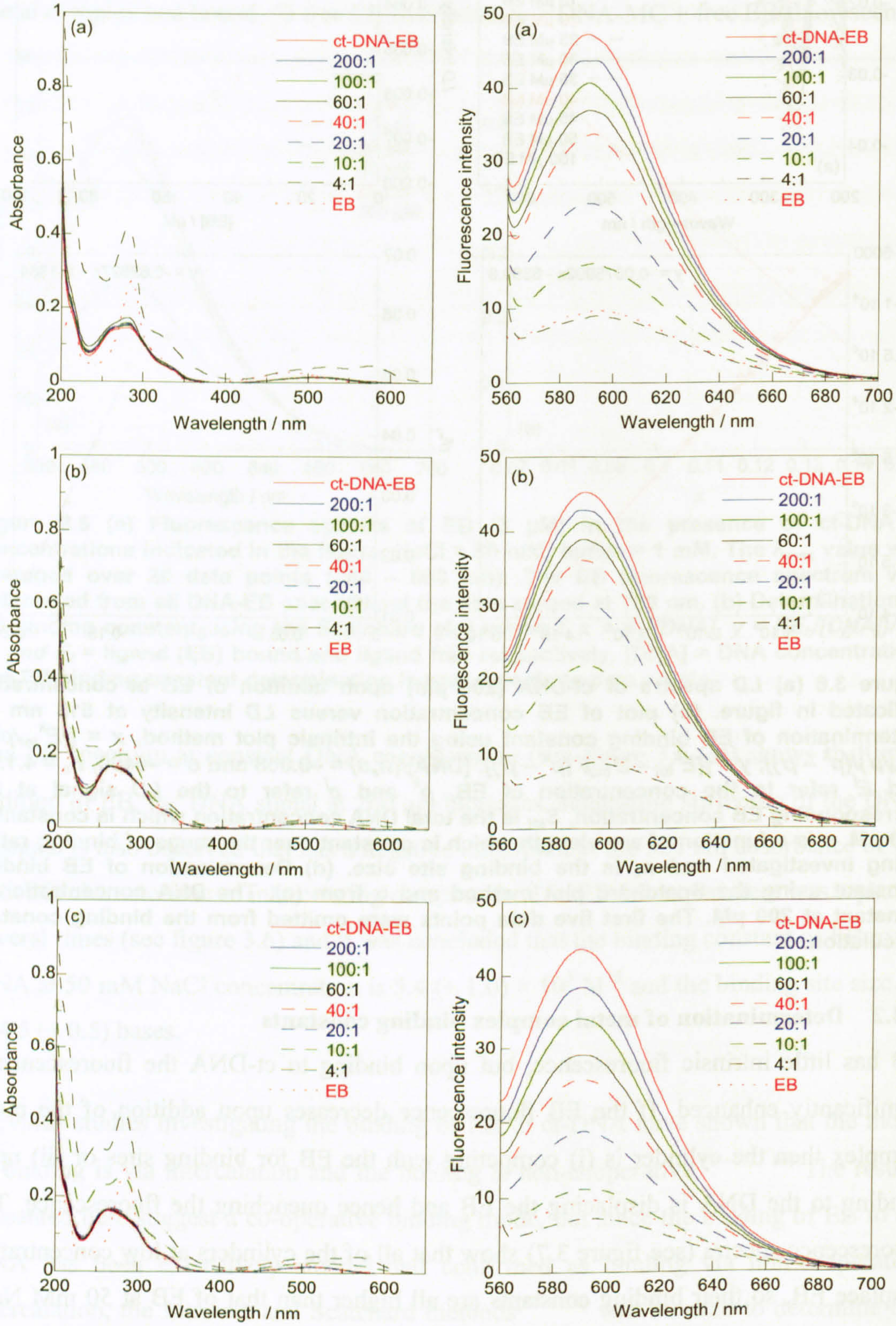



**Figure 3.7 UV/visible absorbance and fluorescence spectra of (a) (R)-$L_1$ (b) (R)-$L_3$ and (c) (R)-$L_5$ displacing EB (15 μM) from ct-DNA (12 μM). Ratio shows DNA concentration to metal complex concentration. Data collected in a 1 cm path length cuvette.**

The plot shown in figure 3.8 shows the decrease in fluorescence intensity upon displacement of the EB from ct-DNA by the $L_1$, $L_3$ and $L_5$ metal complexes at the wavelength maximum (590 nm). The qualitative ranking of the bimetallo iron cylinder binding is given in table 3.3, as are approximate values for $K_h$ and $n_h$. Since $K_h$ and $n_h$ are not truly independent (since the total number of helicate sites depends on $n_h$ and appears in the equation for $K_h$) the binding constants for different cylinders cannot be directly compared unless the cylinder site sizes are similar. The tables used to calculate $K_h$ and $n_h$ for all of the metal complexes investigated are presented in appendix II. The values have been estimated by measuring the quenching of ethidium bromide fluorescence upon addition of the cylinder to solutions of constant DNA, EB, NaCl and buffer concentrations. It was observed that upon addition of the metal complexes to a ct-DNA-EB solution, the ethidium bromide was displaced. However, as observed with the determination of the EB binding constant using fluorescence, the data obtained was unreliable due to problems with self-quenching and the fact that we make the assumption that all sites fluoresce equally. We can make the following qualitative analysis. The (P)-enantiomers of the parent cylinder displace EB most effectively and the (P)-enantiomers of the metal complexes are ranked in the order $L_1 > L_5 > L_3$, with $L_3$ significantly less effective at displacing the EB. The binding constants range from $10^6$ $M^{-1}$ for (P)-$L_3$ to $10^8$ $M^{-1}$ for the (P)-enantiomers of $L_1$ and $L_5$. The low loading/high site size for the (P)-enantiomers is intriguing as no cylinder can block 50 or more base pairs. It must reflect the significant conformational change induced by these cylinders making the DNA unconducive to intercalative binding.

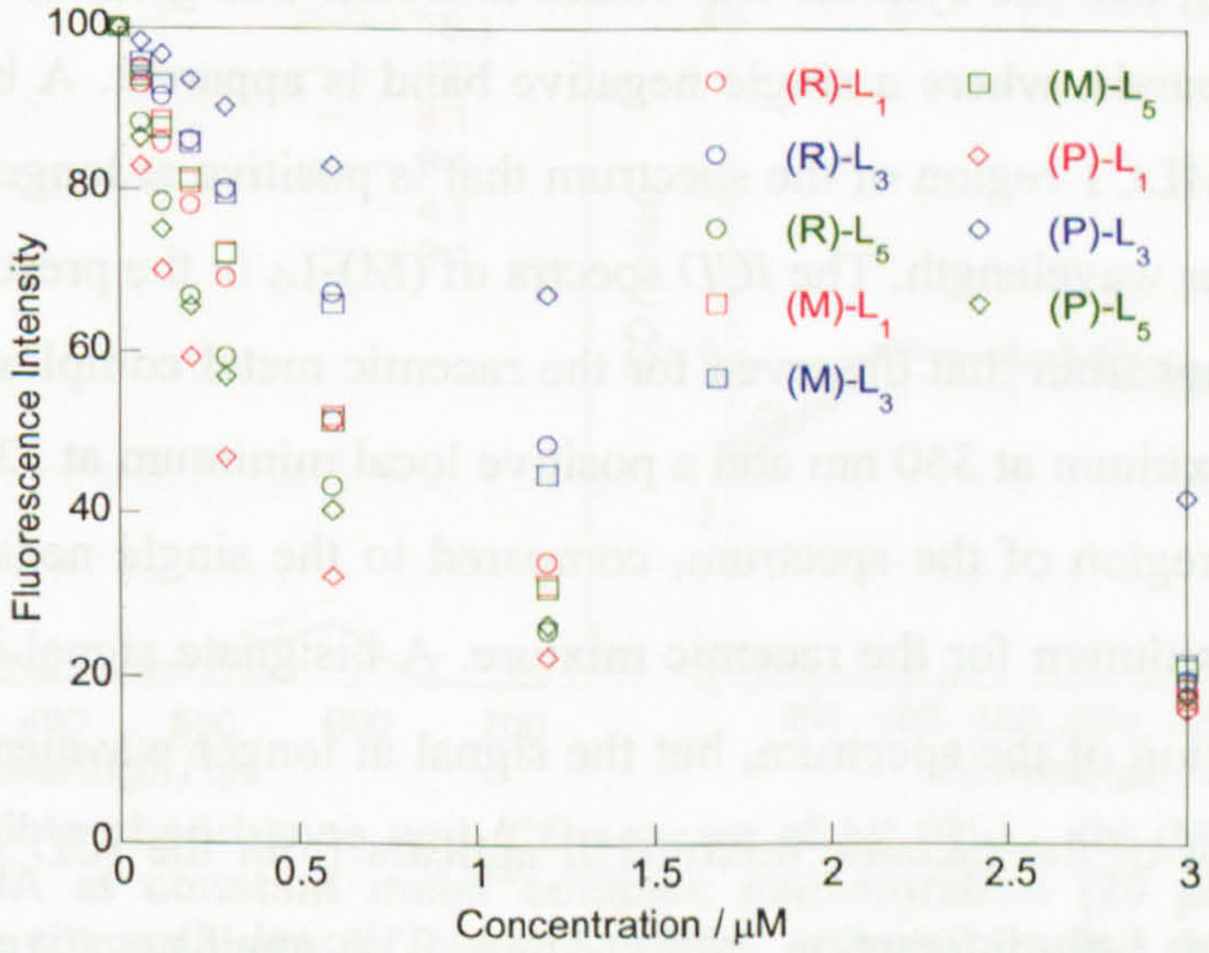



**Figure 3.8 Fluorescence intensity at different concentrations of (R)-, (M)- and (P)-$L_1$, $L_3$ and $L_5$ with constant DNA (12 μM) and EB (15 μM) concentration (data from figure 3.7). Data averaged over 20 data points (580 nm to 600 nm) and normalised to a starting intensity of 100. (Fluorescence spectra presented in appendix II).**

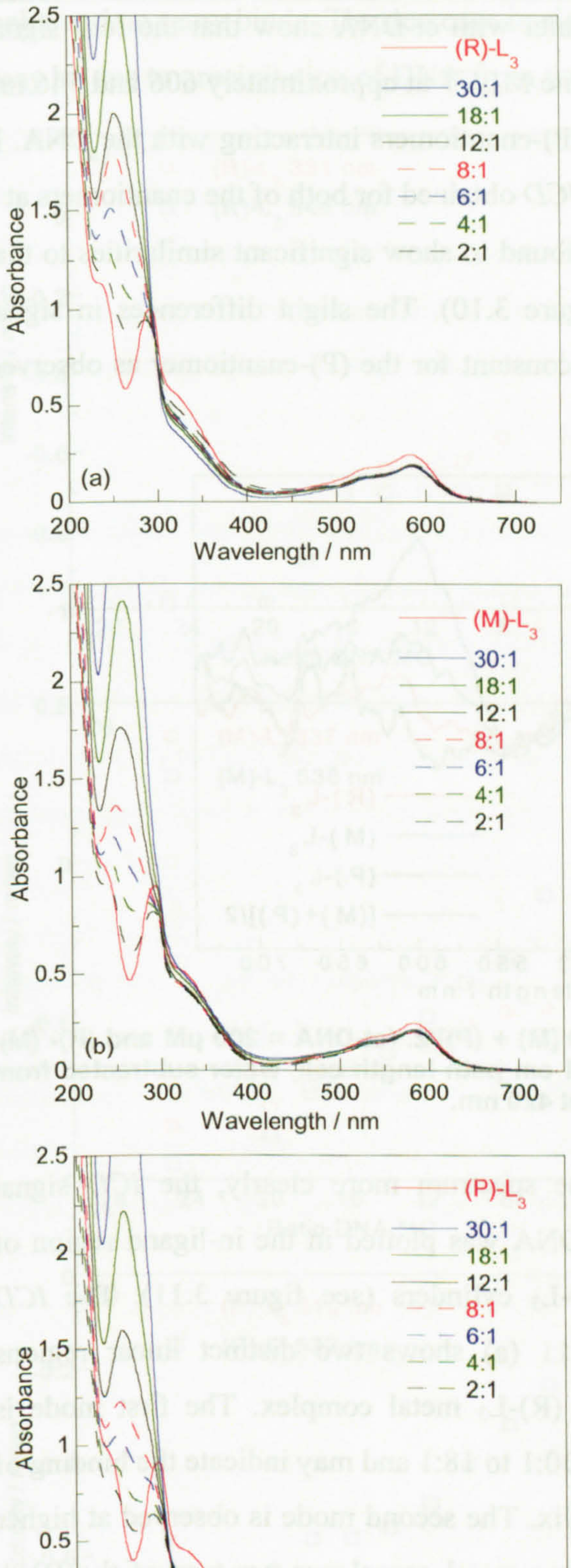
| EB Displacement Ranking | Least | | | | | | | | Most |
|---|---|---|---|---|---|---|---|---|---|
| | (P)-$L_3$ | (R)-$L_3$ | (M)-$L_3$ | (M)-$L_1$ | (R)-$L_1$ | (M)-$L_5$ | (R)-$L_5$ | (P)-$L_5$ | (P)-$L_1$ |
| $K_h$ / ($10^6$ $M^{-1}$) | 1 | 8 | 19 | 29 | 31 | 36 | 53 | 240 | 280 |
| $n_h$ / bases | 21 | 9 | 20 | 26 | 71 | 62 | 30 | 29 | 105 |

**Table 3.3 Order of effectiveness of EB displacement by the bimetallo iron cylinders. Also shown is the approximate binding constant $K_h$ and the metal complex site size $n_h$ calculated using the Intrinsic and Scatchard plot methods. Errors (as determined by experiment to experiment variation) are of the order 20 – 50%.**

## 3.4.3 DNA binding studies

### 3.4.3.1 DNA binding of metal complex $L_3$

The UV/visible absorbance spectrum (see figure 3.9) shows that there is a small wavelength shift in the metal to ligand charge transfer band for (R)-$L_3$ ((R)-$L_3$ maximum at 583 nm, (R)-$L_3$ in the presence of DNA maximum at 585 nm). There is also a small change in intensity in the MLCT band for the enantiomers, suggesting that the complexes are interacting or binding to DNA. Another important feature of the UV/visible absorbance spectra is that upon addition of the metal complex to the DNA solution, there is a wavelength shift in the DNA spectroscopic region ((R)-$L_3$: wavelength shift from 260 nm to 254 nm; (M)-$L_3$: from 260 nm to 261 nm and (P)-$L_3$: from 260 nm to 257 nm) suggesting a definite binding interaction. The induced circular dichroism (*ICD*) of (R)-, (M)- and (P)-$L_3$ at all wavelengths decreases as the DNA concentration decreases with constant cylinder concentration. This is almost certainly due to the decrease in available sites on the DNA to which the metal complex may bind. *CD* spectra of (R)-$L_3$ in solution with ct-DNA show induced *CD* signals at 330 nm and 546 nm indicating that the cylinder has bound to DNA. The greatest effect is observed in the in-ligand bands, where a single negative band is apparent. A bisignate signal is observed in the MLCT region of the spectrum that is positive at longer wavelength and negative at shorter wavelength. The *ICD* spectra of (M)-$L_3$ in the presence of ct-DNA is somewhat different from that observed for the racemic metal complex. A signal with a positive local maximum at 350 nm and a positive local minimum at 335 nm is observed in the in-ligand region of the spectrum, compared to the single negative band with a negative local maximum for the racemic mixture. A bisignate signal is again observed in the MLCT region of the spectrum, but the signal at longer wavelength has a greater intensity than that of the racemic mixture. It appears from the *ICD* spectrum that the (P)-enantiomer has little interaction with ct-DNA. The small negative minimum in the in-ligand region of the *ICD* spectrum and small changes in the UV/visible absorbance spectrum were assumed to be indicative of a binding interaction of the (P)-enantiomer with ct-DNA.



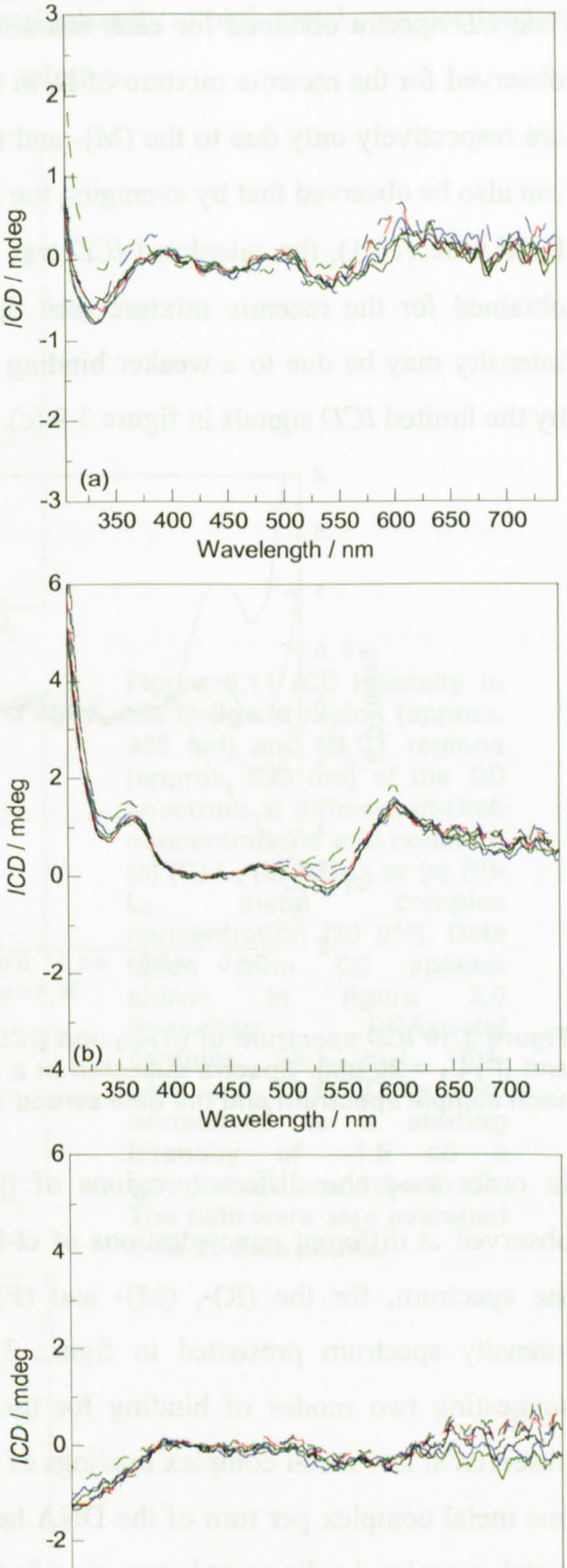




**Figure 3.9 UV/visible absorbance and *ICD* spectra of (a) (R)-$L_3$, (b) (M)-$L_3$ and (c) (P)-$L_3$ binding to ct-DNA at constant metal complex concentration (20 μM). Spectra were recorded in a 1 cm path length cuvette. Water was subtracted from each sample spectrum and the data zeroed at 420 nm for the racemic metal complex and enantiomers. *CD* spectra were collected at 30:1, 25:1 20:1, 18:1, 16:1, 14:1, 12:1, 10:1, 8:1, 6:1, 4:1 and 2:1 DNA:metal complex, ratios shown are indicated in figure, some were omitted for clarity. UV/visible absorbance data above 2.5 absorbance units not shown.**

The *CD* spectra obtained for each enantiomer with ct-DNA show that the *ICD* signal observed for the racemic mixture of $L_3$ in the MLCT at approximately 606 and 546 nm are respectively only due to the (M)- and (P)-enantiomers interacting with the DNA. It can also be observed that by averaging the *ICD* obtained for both of the enantiomers at a fixed ratio (10:1), the calculated *ICD* was found to show significant similarities to that obtained for the racemic mixture (see figure 3.10). The slight differences in signal intensity may be due to a weaker binding constant for the (P)-enantiomer as observed by the limited *ICD* signals in figure 3.9 (c).

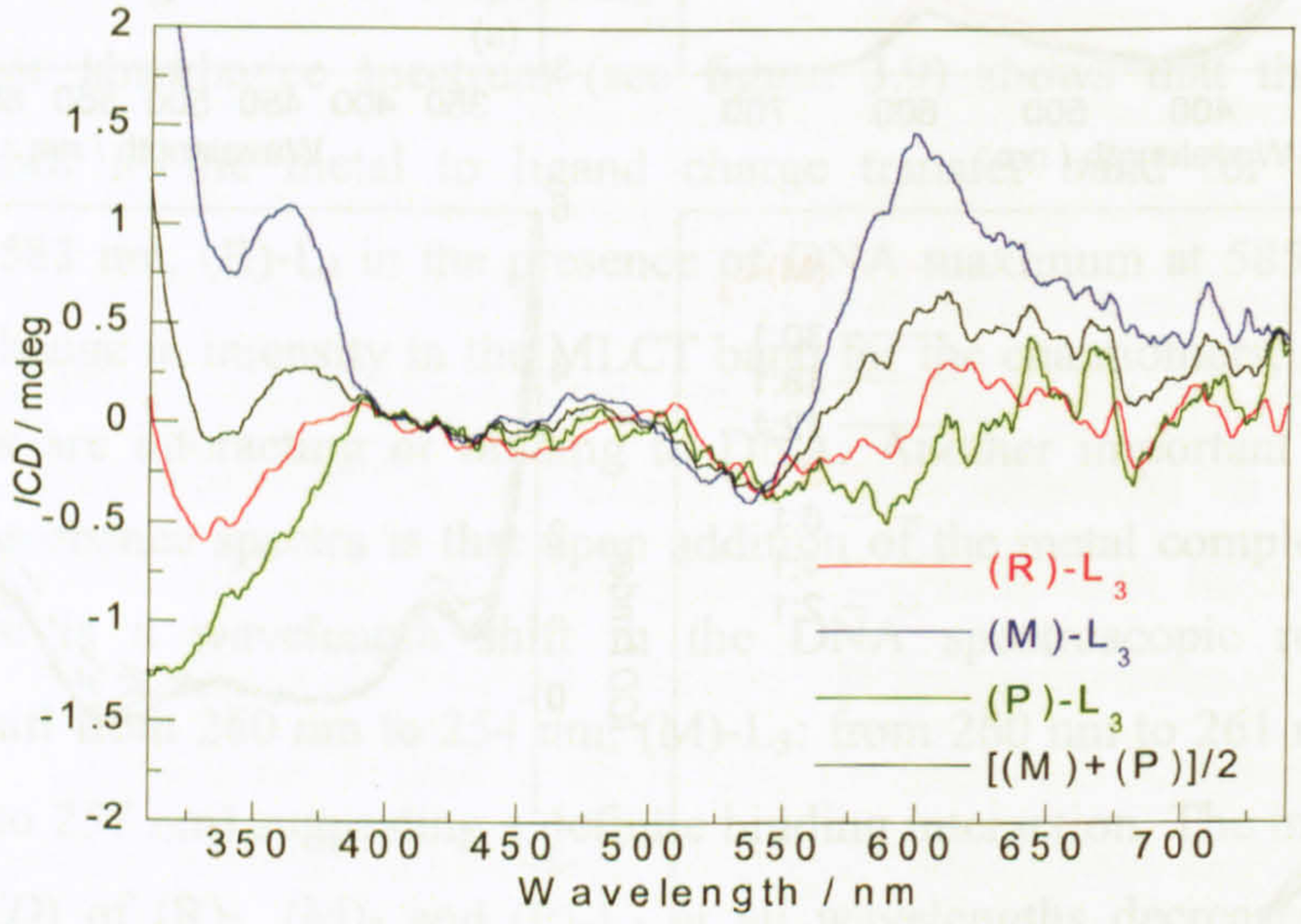



**Figure 3.10 *ICD* spectrum of (R)-$L_3$ and [*ICD* (M) + (P)]/2. (ct-DNA = 200 µM and (R)- (M)- and (P)-$L_3$ = 20 µM). Spectra collected in a 1 cm path length cell, water subtracted from each sample spectrum and the data zeroed at 420 nm.**

In order to probe different regions of the spectrum more clearly, the *ICD* signal observed at different concentrations of ct-DNA was plotted in the in-ligand region of the spectrum, for the (R)-, (M)- and (P)-$L_3$ cylinders (see figure 3.11). The *ICD* intensity spectrum presented in figure 3.11 (a) shows two distinct linear regions suggesting two modes of binding for the (R)-$L_3$ metal complex. The first mode is observed at low metal complex loadings of 30:1 to 18:1 and may indicate the binding of one metal complex per turn of the DNA helix. The second mode is observed at higher metal complex loadings and may signify two metal complexes per turn of the DNA helix. It appears from this intensity plot that the (M)-$L_3$ enantiomers have two modes of binding: the first is observed at low metal complex loadings of 30:1 to 18:1 and the final mode is observed from 16:1 to 2:1. Assuming that the small signal in the in-ligand region of the *ICD* spectrum shows a binding interaction between the (P)-enantiomers and ct-DNA, the data suggest that the (P)-enantiomers bind in one single mode over a large concentration range, for example one metal complex binding per turn of the DNA

helix and no more binds. The decrease in signal intensity at high metal complex loading may be due to precipitation of DNA from the solution.

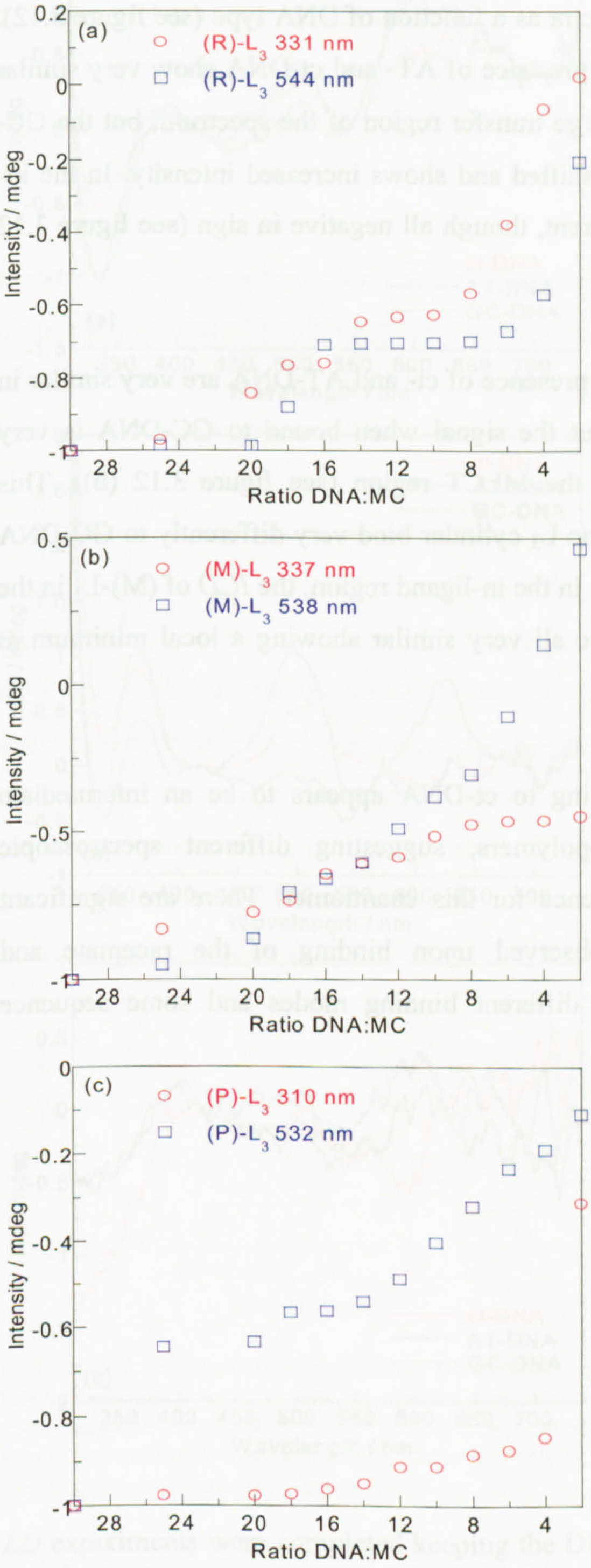



**Figure 3.11 *ICD* intensity in the in-ligand region (approx. 300 nm) and MLCT regions (approx. 530 nm) of the *CD* spectrum at different ct-DNA concentrations and constant (a) (R)-$L_3$ (b) (M)-$L_3$ or (c) (P)-$L_3$ metal complex concentration (20 μM). Data taken from *CD* spectra shown in figure 3.9 (including DNA:metal complex, 25:1 20:1, 16:1, 14:1 and 10:1). Data were normalised to a starting intensity of –1.0 so a comparison could be made. The data were also averaged over 20 data points.**

UV/visible absorbance and *CD* experiments were completed using AT- and GC-DNA, to determine if the metal complex had any preference for the different synthetic DNA

homopolymers. UV/visible absorbance, circular and linear dichroism spectra of (R)-$L_3$, (M)-$L_3$ and (P)-$L_3$ binding to AT- and GC-DNA are presented in appendix II. From this it was possible to compare the *CD* spectra as a function of DNA type (see figure 3.12). The spectra obtained for (R)-$L_3$ in the presence of AT- and ct-DNA show very similar *ICD* spectra in the metal to ligand charge transfer region of the spectrum, but the GC-DNA signal is somewhat wavelength shifted and shows increased intensity. In the in-ligand region they are all slightly different, though all negative in sign (see figure 3.12 (a)).

The spectra obtained for (M)-$L_3$ in the presence of ct- and AT-DNA are very similar in the MLCT region of the spectrum, but the signal when bound to GC-DNA is very different, in fact opposite in sign in the MLCT region (see figure 3.12 (b)). This suggests that the (M)-enantiomers of the $L_3$ cylinder bind very differently to GC-DNA compared to the (R)- and (P)-cylinders. In the in-ligand region, the *ICD* of (M)-$L_3$ in the presence of ct-, AT- and GC-DNA are all very similar showing a local minimum at approximately 320 nm.

The *ICD* signal of (P)-$L_3$ upon binding to ct-DNA appears to be an intermediate between the two alternating homopolymers, suggesting different spectroscopic interactions but little sequence preference for this enantiomer. There are significant differences in the spectral shape observed upon binding of the racemate and enantiomers to the DNA suggesting different binding modes and some sequence preference.

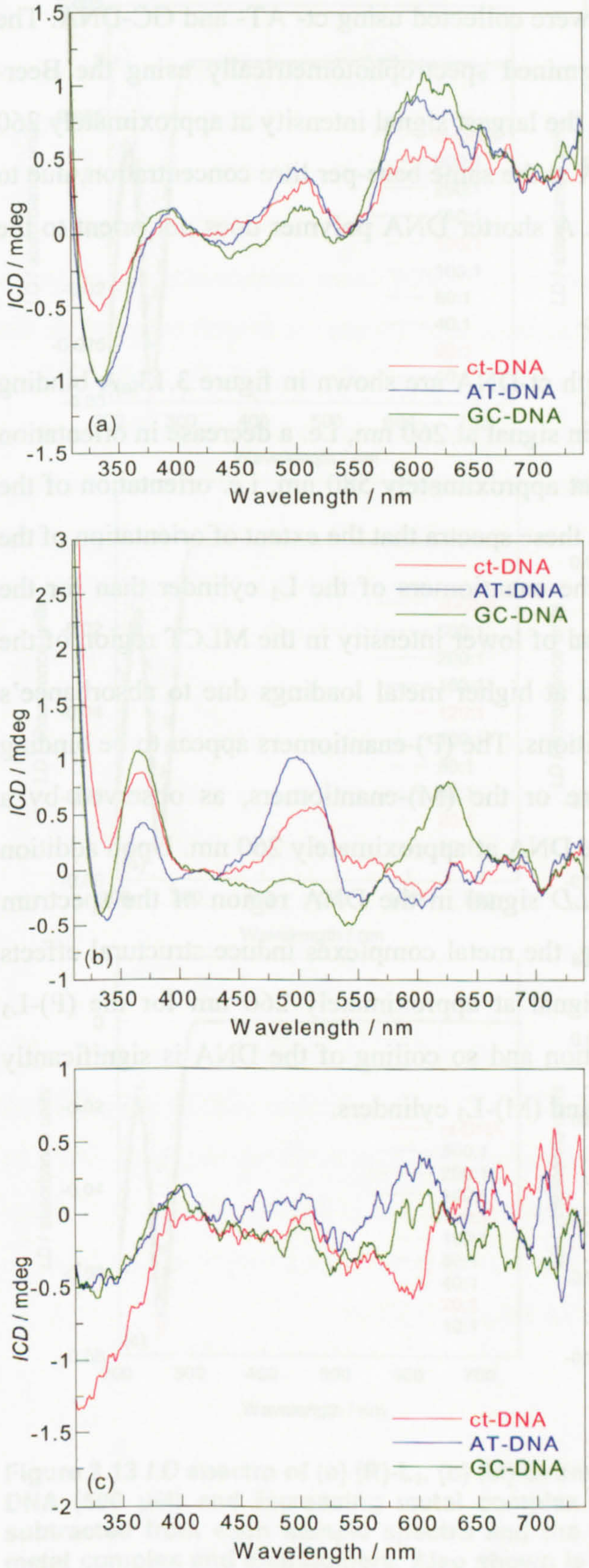



**Figure 3.12 *ICD* spectra of (a) (R)-$L_3$, (b) (M)-$L_3$ and (c) (P)-$L_3$ binding to ct-, AT- and GC-DNA at constant DNA (400 μM) and constant metal complex (20 μM) concentration. Spectra collected in a 1 cm path length cuvette and water subtracted from each sample spectrum. The spectra were zeroed at 420 nm for the racemic metal complex and enantiomers and averaged over 20 data points.**

*LD* experiments were completed keeping the DNA concentration constant as described in § 3.3.3.1. DNA has a characteristic negative *LD* signal from 300 nm to 220 nm due to the base pairs lying perpendicular to the DNA axis. Signals in the MLCT and in-ligand regions of the spectrum, as well as a decrease in the DNA signal intensity indicates a

metal complex-DNA interaction. Data were collected using ct- AT- and GC-DNA. The concentration of each DNA was determined spectrophotometrically using the Beer-Lambert Law. Ct-DNA generally gives the largest signal intensity at approximately 260 nm in comparison to AT- and GC-DNA at the same base-per litre concentration, due to the difference in polymer chain length. A shorter DNA polymer does not orient to the same extent as a longer strand of DNA.

*LD* spectra of (R)-, (M)- and (P)-$L_3$ with ct-DNA are shown in figure 3.13. A binding interaction is observed by the decrease in signal at 260 nm, i.e. a decrease in orientation of the DNA and an increase in signal at approximately 580 nm, i.e. orientation of the metal complex. It can be observed from these spectra that the extent of orientation of the metal complex is less significant for the enantiomers of the $L_3$ cylinder than for the racemic mixture, as observed by a signal of lower intensity in the MLCT region of the spectrum. The DNA signal is distorted at higher metal loadings due to absorbance's from the metal complex in-ligand transitions. The (P)-enantiomers appear to be binding less strongly than the racemic mixture or the (M)-enantiomers, as observed by a decrease in the loss of orientation of the DNA at approximately 260 nm. Upon addition of the metal complexes to DNA, the *LD* signal in the DNA region of the spectrum decreases, suggesting that upon binding, the metal complexes induce structural effects in the DNA conformation. The *LD* signal at approximately 260 nm for the (P)-$L_3$ cylinder shows that the loss of orientation and so coiling of the DNA is significantly lower than that observed for the (R)-$L_3$ and (M)-$L_3$ cylinders.

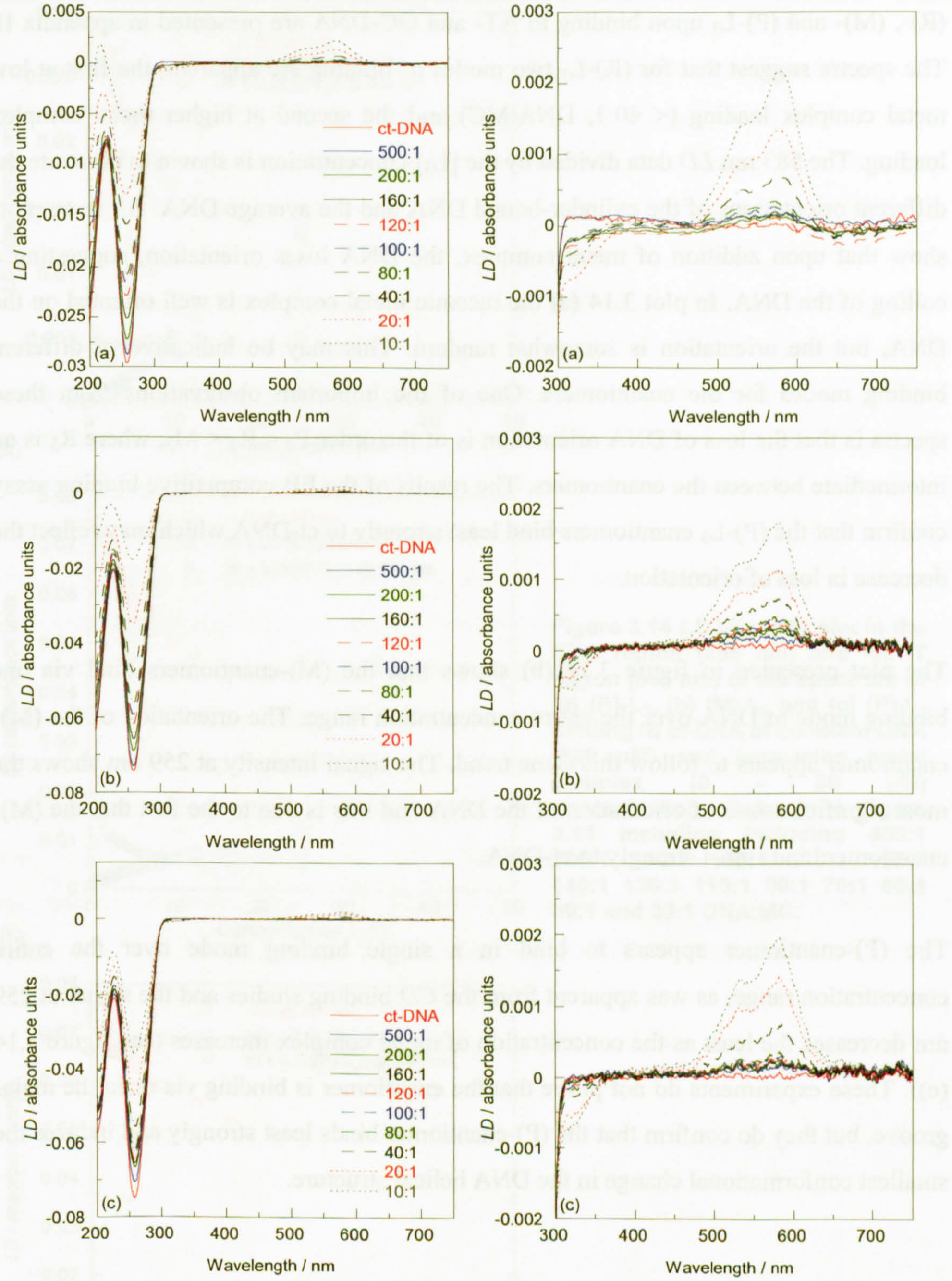



**Figure 3.13 *LD* spectra of (a) (R)-$L_3$, (b) (M)-$L_3$ and (c) (P)-$L_3$ binding to ct-DNA at constant DNA (500 μM) and increasing metal complex (0 – 50 μM) concentration. Water was subtracted from each sample spectra and the data zeroed at 420 nm for the racemic metal complex and enantiomers. Also shown is the signal observed in the MLCT region of the spectrum for (R)- (M)- and (P)-$L_3$.**

The experiment was repeated with AT- and GC-DNA and the data are presented in appendix II. The plots of figure 3.14 show the signal intensity in the DNA and MLCT regions of the spectrum for (R)-, (M)- and (P)-$L_3$ upon binding to ct-DNA. The plots of

(R)-, (M)- and (P)-$L_3$ upon binding to AT- and GC-DNA are presented in appendix II. The spectra suggest that for (R)-$L_3$ two modes of binding are apparent, the first at low metal complex loading (< 40:1, DNA:MC) and the second at higher metal complex loading. The 583 nm *LD* data divided by the [$L_3$] concentration is shown to illustrate the different orientations of the cylinder-bound DNA and the average DNA. All three plots show that upon addition of metal complex, the DNA loses orientation, suggesting a coiling of the DNA. In plot 3.14 (a) the racemic metal complex is well oriented on the DNA, but the orientation is somewhat random. This may be indicative of different binding modes for the enantiomers. One of the important observations from these spectra is that the loss of DNA orientation is of the order $P_3 < R_3 < M_3$, where $R_3$ is an intermediate between the enantiomers. The results of the EB competitive binding assay confirm that the (P)-$L_3$ enantiomers bind least strongly to ct-DNA which may reflect the decrease in loss of orientation.

The plot presented in figure 3.14 (b) shows that the (M)-enantiomers bind via one binding mode to DNA over the entire concentration range. The orientation of the (M)-enantiomer appears to follow this same trend. The signal intensity at 259 nm shows the most significant loss of orientation of the DNA and this is due to the fact that the (M)-enantiomer binds most strongly to ct-DNA.

The (P)-enantiomer appears to bind in a single binding mode over the entire concentration range, as was apparent from the *CD* binding studies and the signal at 259 nm decreases the least as the concentration of metal complex increases (see figure 3.14 (c)). These experiments do not prove that the enantiomer is binding via or in the major groove, but they do confirm that the (P)-enantiomer binds least strongly and induces the smallest conformational change in the DNA helical structure.

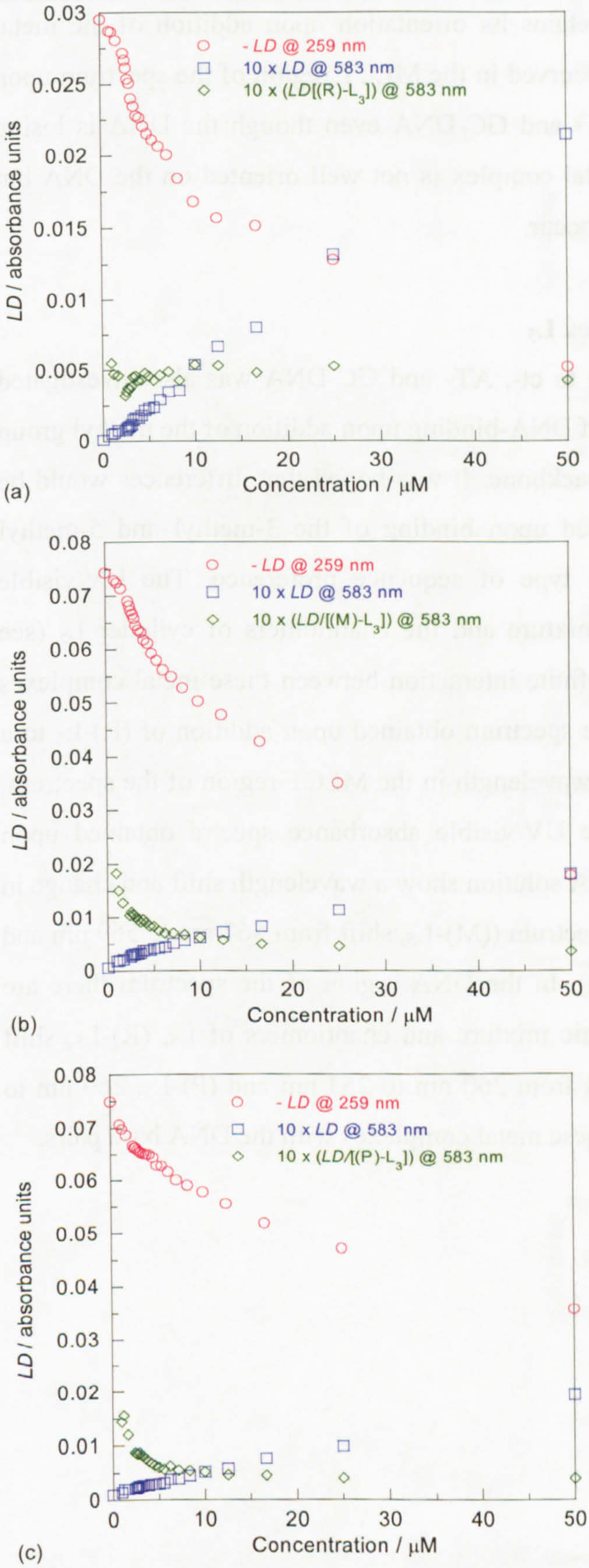



**Figure 3.14 *LD* intensity plot in the DNA region (259 nm) and MLCT region (583 nm) of the spectrum of (a) (R)-$L_3$, (b) (M)-$L_3$ and (c) (P)-$L_3$ binding to ct-DNA at constant DNA (500 μM) and increasing metal complex (0 – 50 μM) concentration. Data from figure 3.13 including including 400:1 300:1 190:1 180:1 170:1 150:1 140:1 130:1 110:1 90:1 70:1 60:1 50:1 and 30:1 DNA:MC.**

The *LD* intensity data obtained for the $L_3$ cylinder in the DNA region of the spectrum suggests that different spectroscopic interactions occur upon addition of (R)-$L_3$ to ct-, AT- and GC-DNA. The rate at which the DNA loses orientation is of the order AT-,

GC- and ct-DNA, i.e. the ct-DNA retains its orientation upon addition of the metal complex. There is very little signal observed in the MLCT region of the spectrum upon addition of the metal complex to AT- and GC-DNA even though the DNA is losing orientation, this suggests that the metal complex is not well oriented on the DNA but some sort of binding interaction does occur.

### 3.4.3.2 DNA binding of metal complex $L_5$

The binding of (R)-, (M)- and (P)-$L_5$ to ct-, AT- and GC-DNA was also investigated using *CD* and *LD* to probe the effect of DNA-binding upon addition of the methyl group at a different position on the ligand backbone. It was hoped that differences would be observed between the spectra obtained upon binding of the 3-methyl and 5-methyl cylinders to DNA to indicate some type of sequence preference. The UV/visible absorbance spectra for the racemic mixture and the enantiomers of cylinder $L_5$ (see figure 3.15) suggest that there is a definite interaction between these metal complexes and DNA. The UV/visible absorbance spectrum obtained upon addition of (R)-$L_5$ to a ct-DNA solution shows no change in wavelength in the MLCT region of the spectrum, but a small change in intensity. The UV/visible absorbance spectra obtained upon addition of the enantiomers to a ct-DNA solution show a wavelength shift and change in intensity in the MLCT region of the spectrum ((M)-$L_5$, shift from 567 nm to 569 nm and (P)-$L_5$, shift from 567 nm to 565 nm). In the DNA region of the spectrum there are again wavelength shifts for the racemic mixture and enantiomers of $L_5$, (R)-$L_5$, shift from 260 nm to 256 nm, (M)-$L_5$, shift from 260 nm to 253 nm and (P)-$L_5$, 260 nm to 252 nm, suggesting an interaction of these metal complexes with the DNA base pairs.

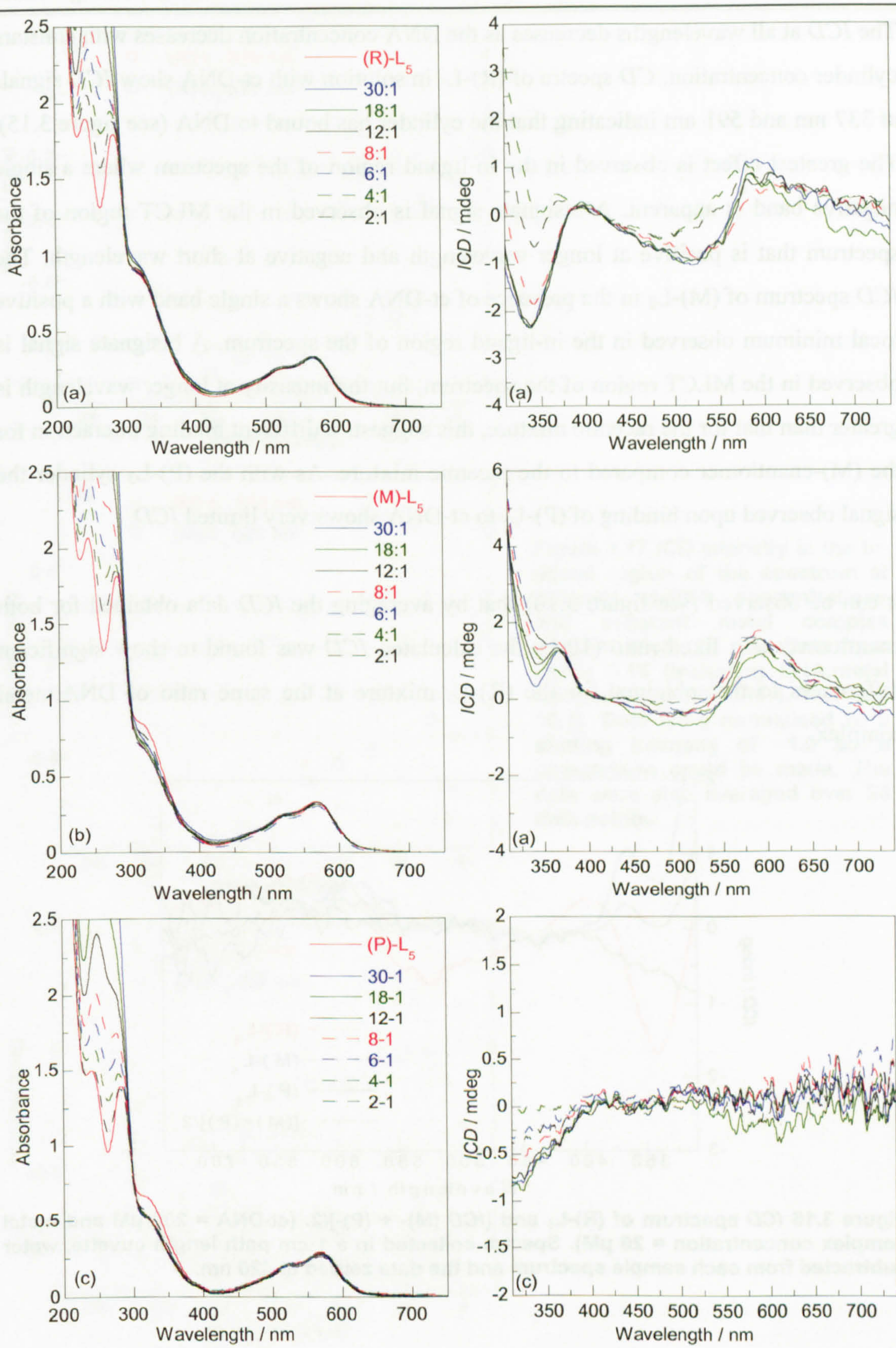



**Figure 3.15 UV/visible absorbance and *ICD* spectra of (a) (R)-$L_5$, (b) (M)-$L_5$ and (c) (P)-$L_5$ binding to ct-DNA at constant metal complex concentration (20 μM). Spectra were recorded in a 1 cm path length cuvette. Water was subtracted from each sample spectrum and the data zeroed at 420 nm for the racemic metal complex and enantiomers. *CD* spectra were collected at 30:1, 25:1 20:1, 18:1, 16:1, 14:1, 12:1, 10:1, 8:1, 6:1, 4:1 and 2:1 DNA:metal complex, ratios shown are indicated in figure, some were omitted for clarity. UV/visible absorbance data above 2.5 absorbance units not shown.**

The *ICD* at all wavelengths decreases as the DNA concentration decreases with constant cylinder concentration. *CD* spectra of (R)-$L_5$ in solution with ct-DNA show *ICD* signals at 337 nm and 591 nm indicating that the cylinder has bound to DNA (see figure 3.15). The greatest effect is observed in the in-ligand region of the spectrum where a single negative band is apparent. A bisignate signal is observed in the MLCT region of the spectrum that is positive at longer wavelength and negative at short wavelength. The *ICD* spectrum of (M)-$L_3$ in the presence of ct-DNA shows a single band with a positive local minimum observed in the in-ligand region of the spectrum. A bisignate signal is observed in the MLCT region of the spectrum, but the intensity at longer wavelength is greater than that for the racemic mixture, this suggests a different binding interaction for the (M)-enantiomer compared to the racemic mixture. As with the (P)-$L_3$ cylinder the signal observed upon binding of (P)-$L_5$ to ct-DNA shows very limited *ICD*.

It can be observed (see figure 3.16) that by averaging the *ICD* data obtained for both enantiomers at a fixed ratio (10:1), the calculated *ICD* was found to show significant differences to that obtained for the (R)-$L_5$ mixture at the same ratio of DNA:metal complex.

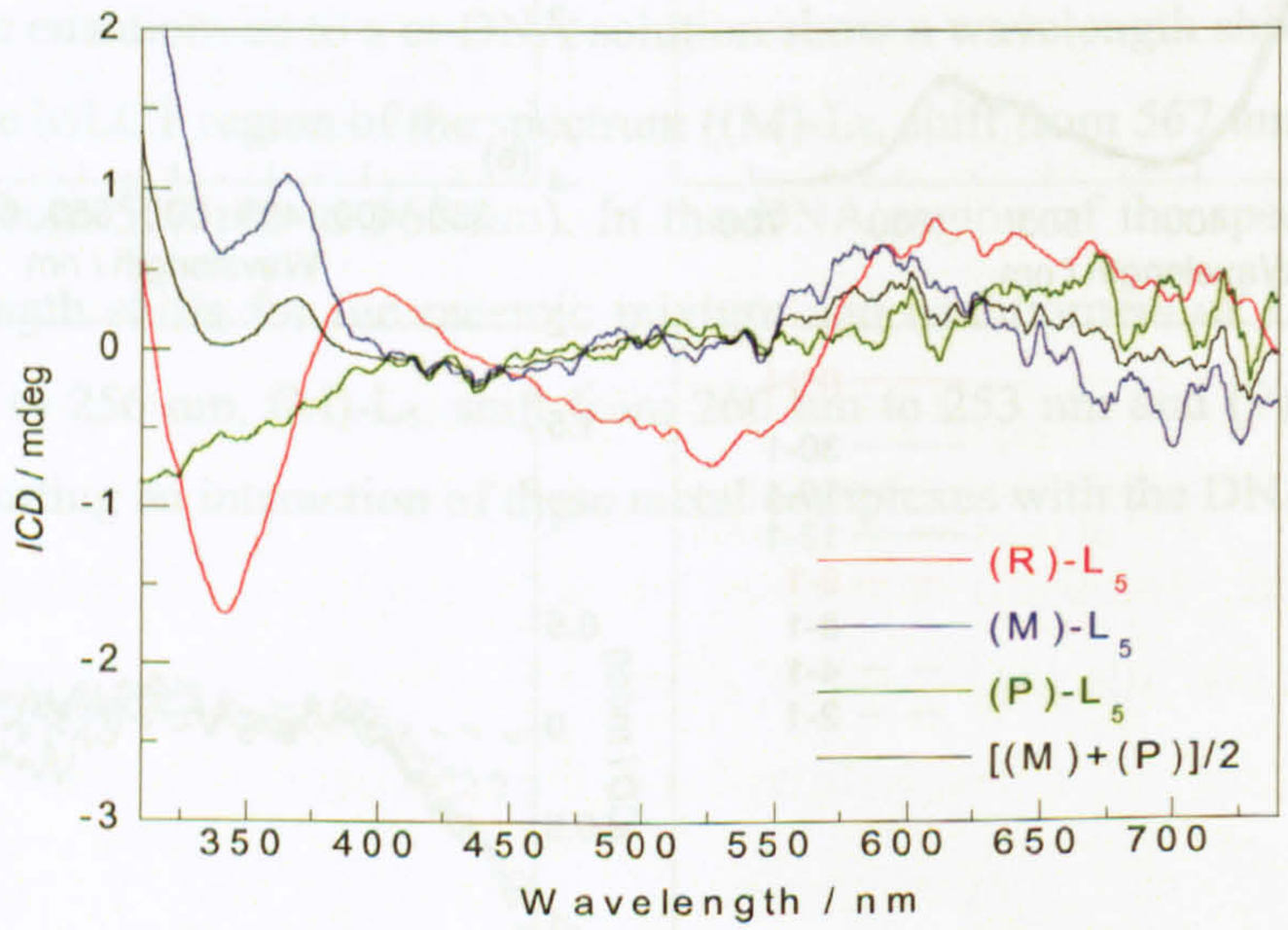



**Figure 3.16 *ICD* spectrum of (R)-$L_5$ and [*ICD* (M)- + (P)-]/2. (ct-DNA = 200 μM and metal complex concentration = 20 μM). Spectra collected in a 1 cm path length cuvette, water subtracted from each sample spectrum and the data zeroed at 420 nm.**

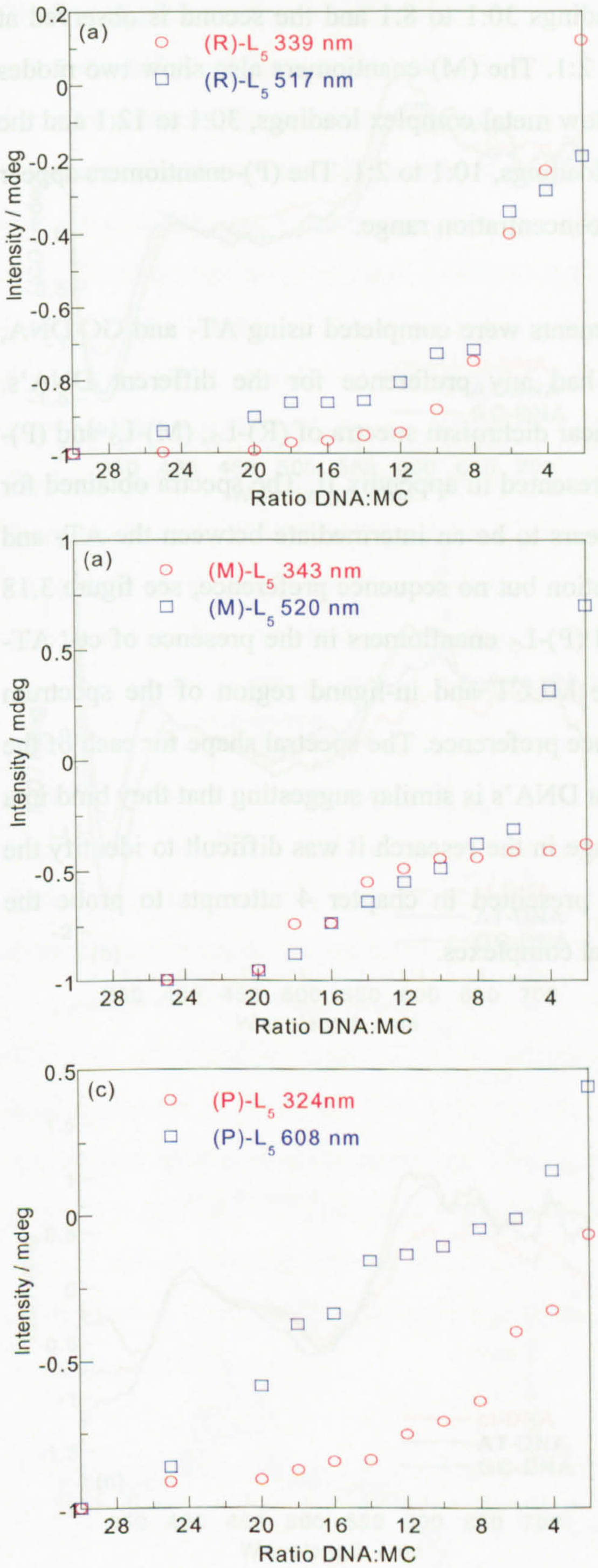



**Figure 3.17 *ICD* intensity in the in-ligand region of the spectrum at different ct-DNA concentrations and constant metal complex concentration (20 μM). Data from figure 3.15 (including DNA:metal complex, 25:1 20:1, 16:1, 14:1 and 10:1). Data were normalised to a starting intensity of –1.0 so a comparison could be made. The data were also averaged over 20 data points.**

In order to probe different regions of the spectrum more clearly, the *ICD* signal observed at different concentrations of ct-DNA was plotted at the in-ligand region of the spectrum, for the (R)-, (M)- and (P)-$L_5$ cylinders (see figure 3.17). It can be observed from this intensity plot that the (R)-$L_5$ metal complex has two modes of binding: the

first is observed at metal complex loadings 30:1 to 8:1 and the second is observed at higher metal complex loadings, 6:1 to 2:1. The (M)-enantiomers also show two modes of binding; the first mode observed at low metal complex loadings, 30:1 to 12:1 and the second mode at higher metal complex loadings, 10:1 to 2:1. The (P)-enantiomers appear to bind in a single mode across a wide concentration range.

UV/visible absorbance and *CD* experiments were completed using AT- and GC-DNA, to determine if the metal complex had any preference for the different DNA's. UV/visible absorbance, circular and linear dichroism spectra of (R)-$L_5$, (M)-$L_5$ and (P)-$L_5$ binding to AT- and GC-DNA are presented in appendix II. The spectra obtained for (R)-$L_5$ in the presence of ct-DNA appears to be an intermediate between the AT- and GC-DNA spectra suggesting an interaction but no sequence preference, see figure 3.18 (a). The spectra obtained for (M)- and (P)-$L_5$ enantiomers in the presence of ct-, AT- and GC-DNA are very similar in the MLCT and in-ligand region of the spectrum suggesting an interaction but no sequence preference. The spectral shape for each of the complexes upon binding to the different DNA's is similar suggesting that they bind in a similar manner to the DNA. At this stage in the research it was difficult to identify the exact modes of binding. The results presented in chapter 4 attempts to probe the different binding geometries of the metal complexes.

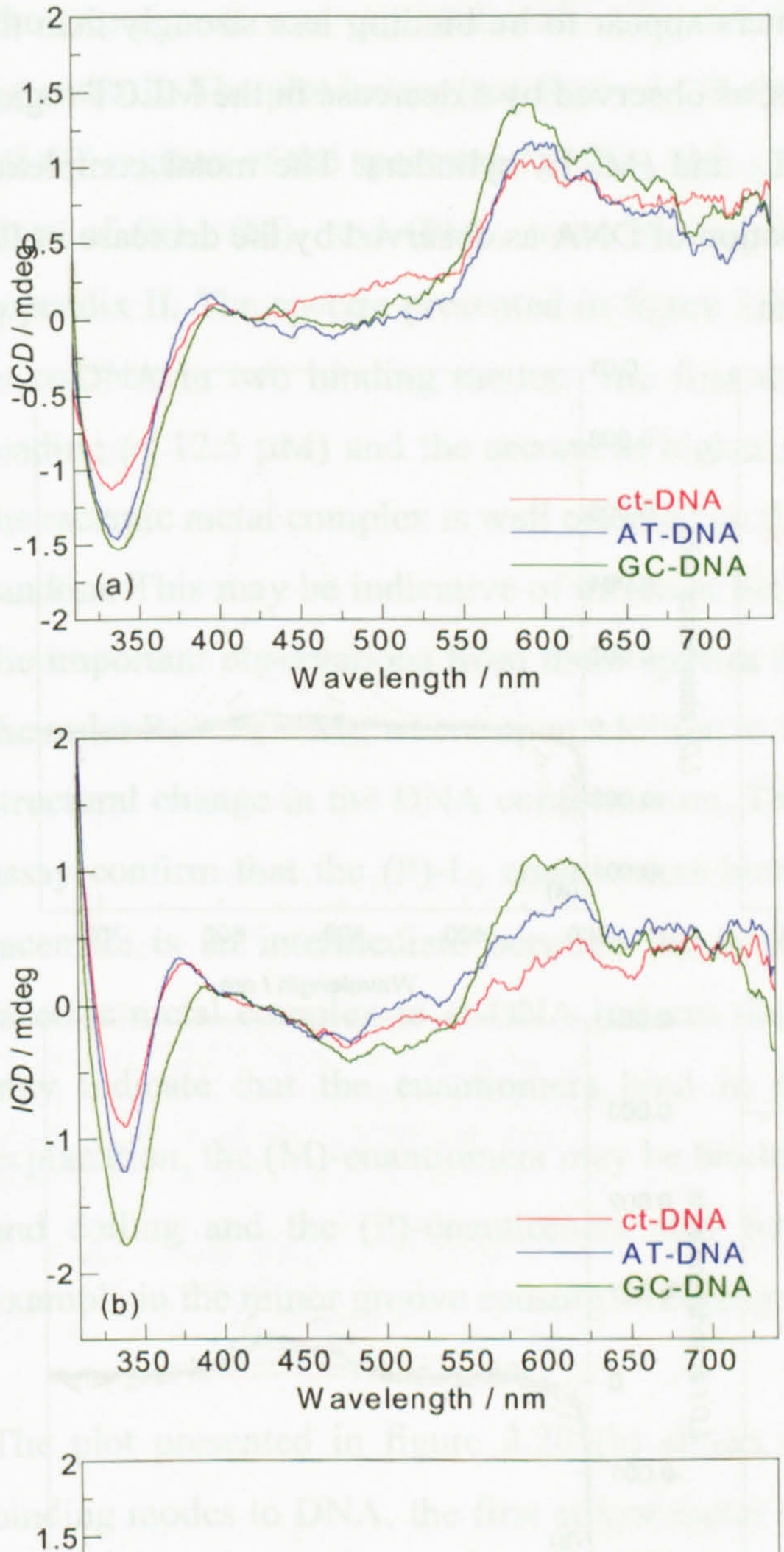



**Figure 3.18 *ICD* spectra of (a) (R)-$L_5$, (b) (M)-$L_5$ and (c) (P)-$L_5$ binding to ct-, AT- and GC-DNA at constant DNA (400 μM) and constant metal complex (20 μM) concentration. Spectra collected in a 1 cm path length cuvette and water subtracted from each sample spectrum. The spectra were zeroed in a region of no absorbance at 420 nm for the racemic metal complex and enantiomers.**

*LD* spectra of (R)-, (M)- and (P)-$L_5$ with ct-DNA are shown in figure 3.19. It can be observed from these spectra that the extent of orientation of the metal complex is less significant for the enantiomers of the $L_5$ cylinder than for the racemic mixture, as observed by a signal of lower intensity in the MLCT region of the spectrum. The DNA signal is distorted at higher metal loadings due to absorbance's from the metal complex

in-ligand transitions. The (P)- enantiomers appear to be binding less strongly than the racemic mixture or the (M)-enantiomers, as observed by a decrease in the MLCT region of the spectrum compared to the (R)-$L_5$ and (M)-$L_5$ cylinders. The metal complexes have a significant effect on the conformation of DNA as observed by the decrease in *LD* signal in the DNA region of the spectrum.

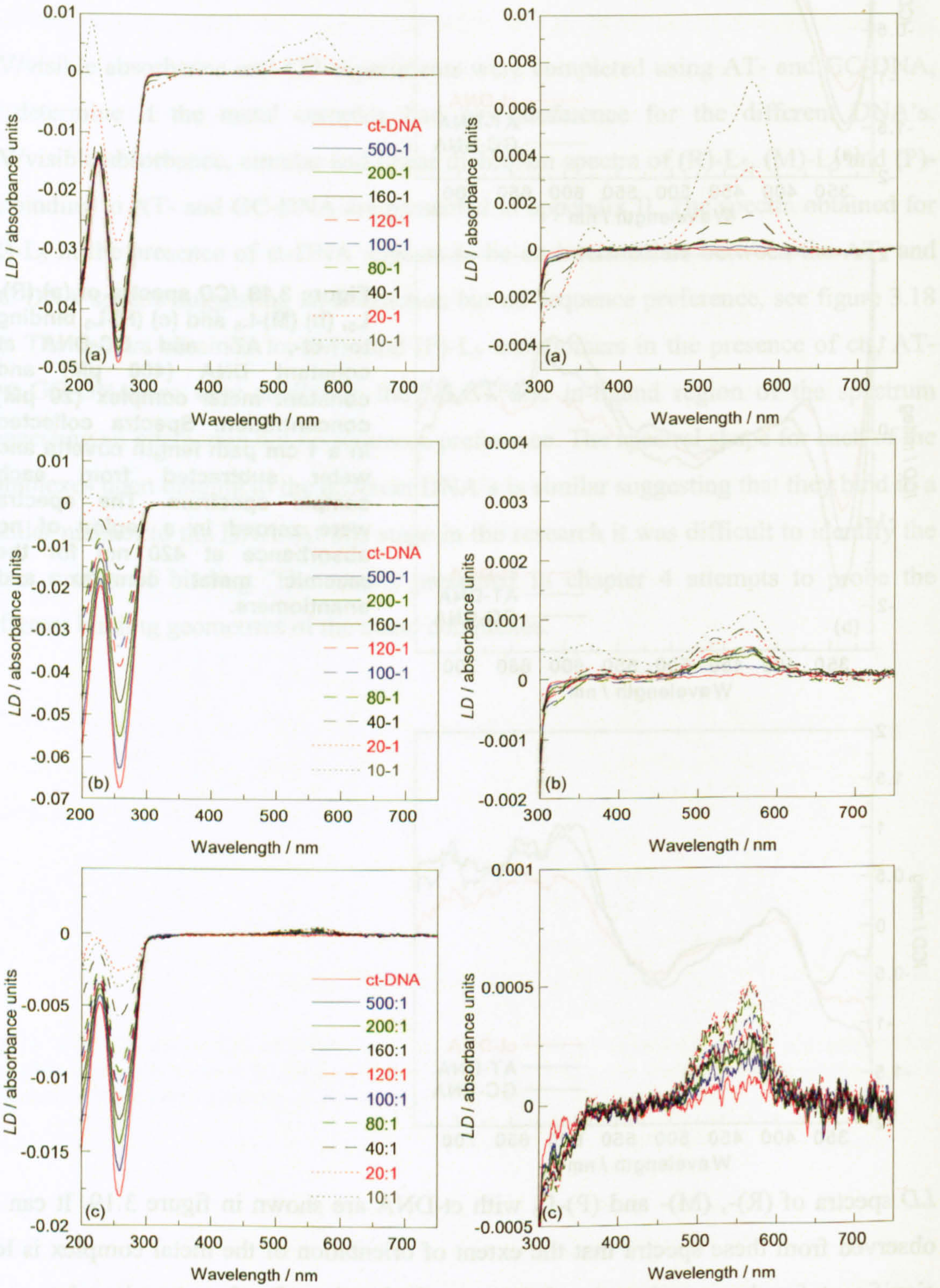


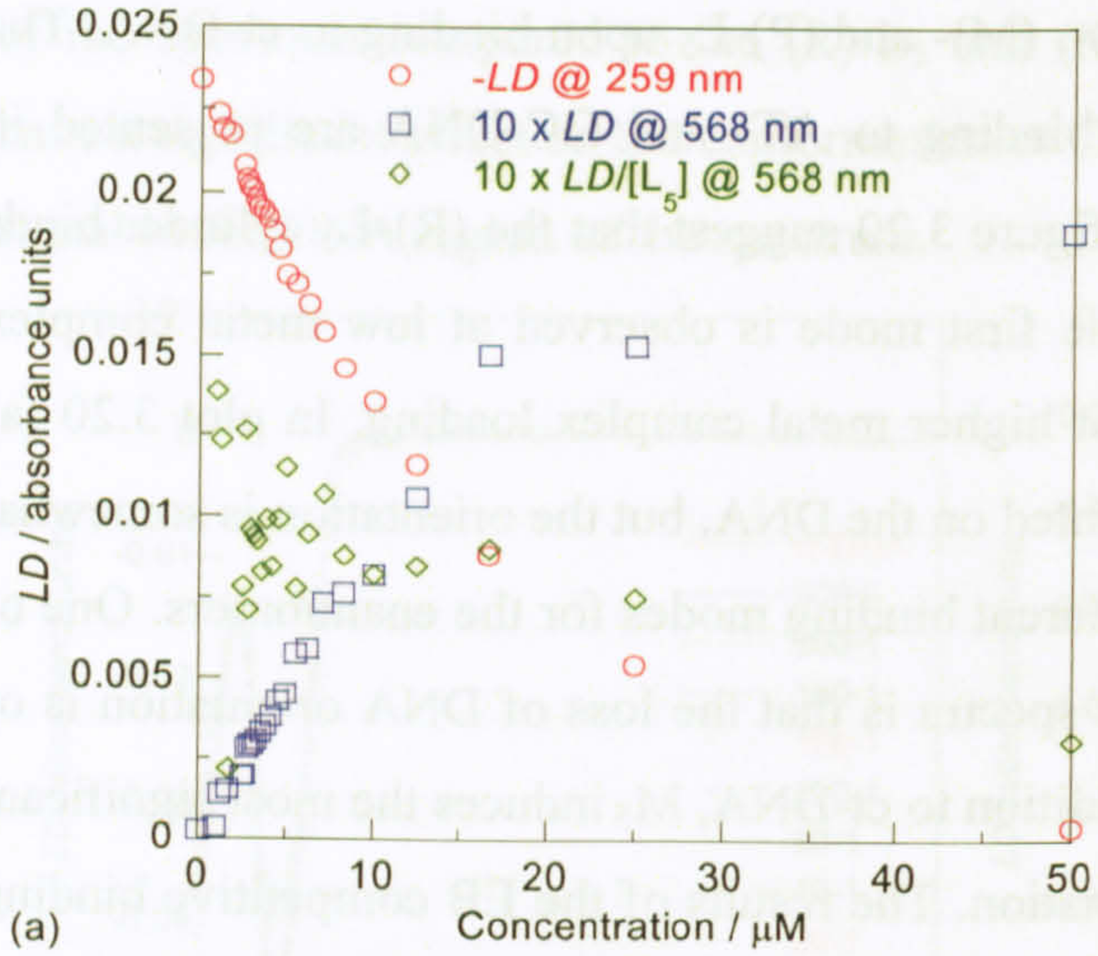

**Figure 3.19 *LD* spectra of (a) (R)-$L_5$, (b) (M)-$L_5$ and (c) (P)-$L_5$ binding to ct-DNA at constant DNA (500 µM) and increasing metal complex (0 – 50 µM) concentration. Also shown is the signal observed in the MLCT region of the spectrum for (R)- (M)- and (P)-$L_5$.**

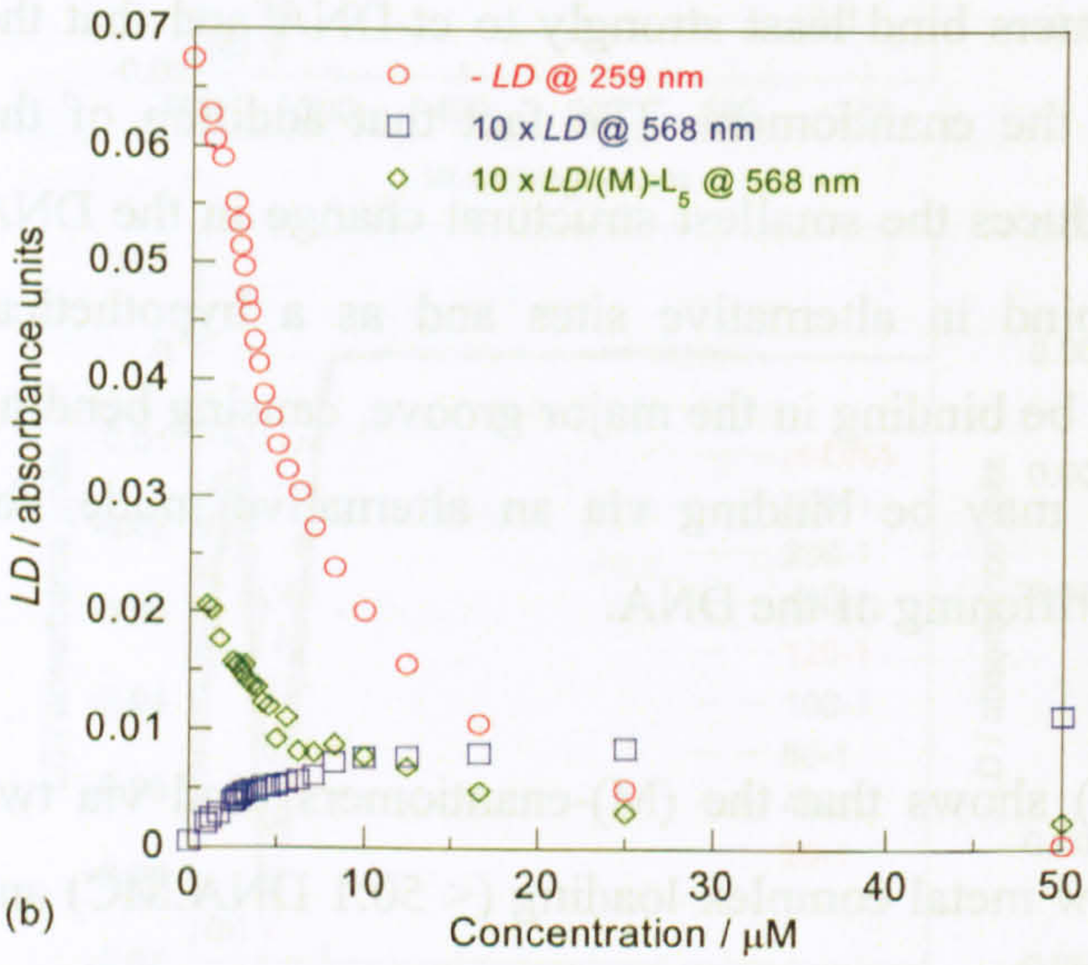
The experiment was repeated with AT- and GC-DNA and the data are presented in appendix II. The plot below (see figure 3.20) shows the signal intensity in the DNA and MLCT regions of the spectrum for (R)-, (M)- and (P)-$L_5$ upon binding to ct-DNA. The plots of (R)-, (M)- and (P)-$L_5$ upon binding to AT- and GC-DNA are presented in appendix II. The spectra presented in figure 3.20 suggest that the (R)-$L_5$ cylinder binds to ct-DNA in two binding modes. The first mode is observed at low metal complex loading (< 12.5 μM) and the second at higher metal complex loading. In plot 3.20 (a) the racemic metal complex is well oriented on the DNA, but the orientation is somewhat random. This may be indicative of different binding modes for the enantiomers. One of the important observations from these spectra is that the loss of DNA orientation is of the order $R_5 < P_5 < M_5$, where upon addition to ct-DNA, $M_5$ induces the most significant structural change in the DNA conformation. The results of the EB competitive binding assay confirm that the (P)-$L_5$ enantiomers bind least strongly to ct-DNA and that the racemate is an intermediate between the enantiomers. The fact that addition of the racemic metal complex to ct-DNA induces the smallest structural change in the DNA may indicate that the enantiomers bind in alternative sites and as a hypothetical explanation, the (M)-enantiomers may be binding in the major groove, causing bending and coiling and the (P)-enantiomers may be binding via an alternative mode, for example in the minor groove causing stiffening of the DNA.

The plot presented in figure 3.20 (b) shows that the (M)-enantiomers bind via two binding modes to DNA, the first at low metal complex loading (< 50:1 DNA:MC) and the second at higher metal complex loading. The orientation of the (M)-enantiomer appears to follow this same trend, the first mode at low metal complex loading and the second at high loading. The signal intensity at 259 nm shows the most significant loss of orientation of the DNA and this is due to the fact that the (M)-enantiomer binds most strongly to ct-DNA.

The (P)-enantiomer appears to bind to ct-DNA via two different modes of binding. The first is observed at low metal complex loading (< 40:1 DNA:MC) and the second at higher metal complex loadings. The (P)-enantiomers binds least strongly to ct-DNA, but the loss of signal at 259 nm is greater than that of the racemic metal complex. This result is somewhat intriguing, indicating that the (P)-enantiomers, irrespective of their binding ability induce a larger structural change in the DNA conformation. These experiments do not prove that the enantiomer is binding via or in the major groove, but

they do confirm that the (P)-enantiomers bind to ct-DNA and induce a significantly greater conformational change in the DNA structure than the racemic metal complex.



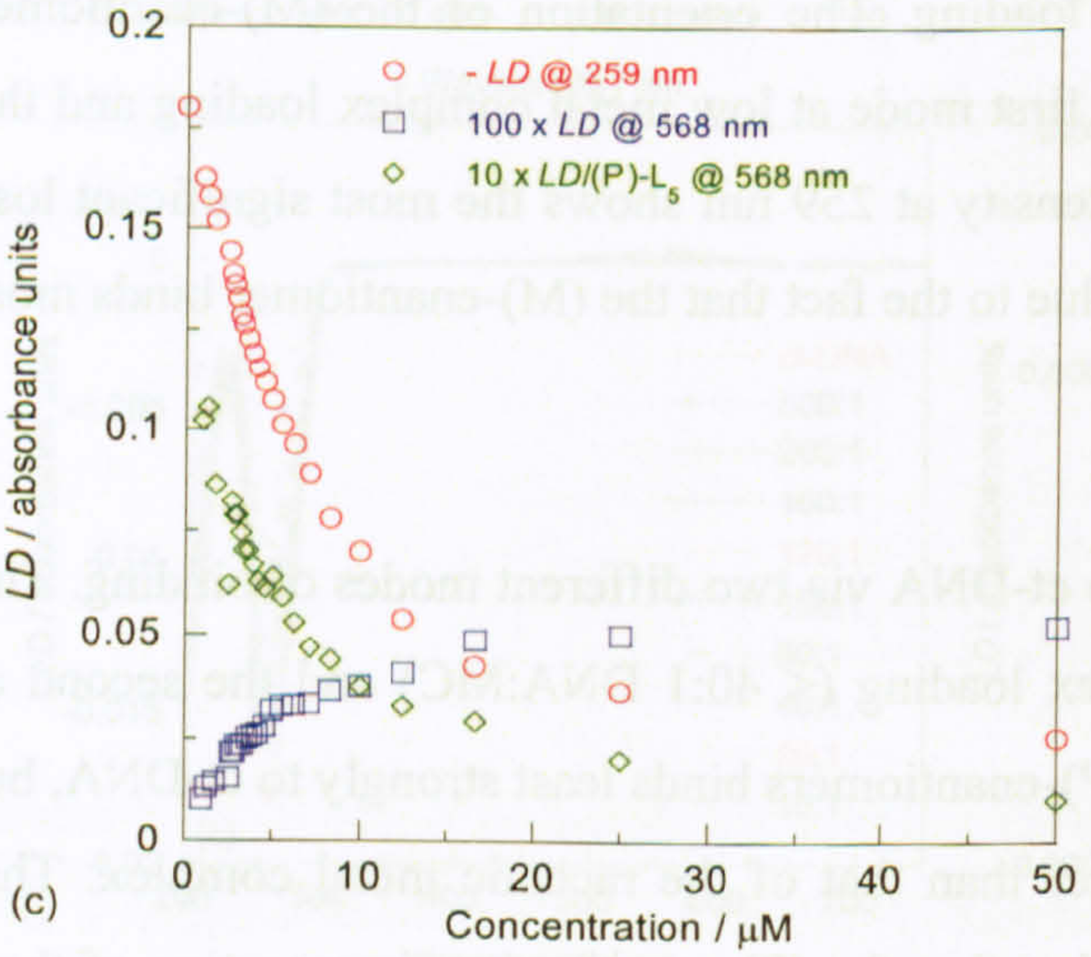




**Figure 3.20 *LD* intensity plot in the DNA region (259 nm) and MLCT region (568 nm) of the spectrum of (a) (R)-$L_5$, (b) (M)-$L_5$ and (c) (P)-$L_5$ binding to ct-DNA at constant DNA (500 μM) and increasing metal complex (0 – 50 μM) concentration. Data from figure 3.19 including 400:1 300:1 190:1 180:1 170:1 150:1 140:1 130:1 110:1 90:1 70:1 60:1 50:1 and 30:1 DNA:MC.**

The *LD* intensity data obtained for the $L_5$ cylinder in the DNA region of the spectrum suggests that different spectroscopic interactions occur upon addition of (R)-$L_5$ to ct-, AT- and GC-DNA. The (M)- and (P)-enantiomers bind to DNA as observed by the increase in the MLCT region of the spectrum. There is very little signal observed in the MLCT region of the spectrum upon addition of the metal complex to AT- and GC-DNA even though the DNA is losing orientation, this suggests that the metal complex is not well oriented on the DNA but some sort of binding interaction does occur.

### 3.5 Discussion

The results presented in this chapter show that:

- The EB binding constant has been determined using linear dichroism to be 5.4 ($\pm$ 1.0) $\times 10^5$ $M^{-1}$ and n, the site size, is 4.5 ($\pm$ 0.5) bases.
- The metal complexes $L_3$ and $L_5$ bind to DNA as observed from the induced circular dichroism spectra.
- The enantiomers of the parent cylinder, $L_1$ and the enantiomers of the methylated bimetallo iron cylinders, $L_3$ and $L_5$ have binding constants higher than 5.4 ($\pm$ 1.0) $\times 10^5$ $M^{-1}$.
- Upon binding the metal complexes induce structural effects on the DNA (as observed from the *LD* binding studies) and as a comparison to the data obtained for the (M)-enantiomer of the parent compound, which has been shown to bind in the major groove, may also show preference for this mode of binding.
- The enantiomers of the $L_3$ cylinder show some sequence preference as observed by the different *ICD* spectra obtained upon binding to the synthetic homopolymers, AT- and GC-DNA.
- The enantiomers of the $L_5$ cylinder appear to show no sequence preference indicated by the similar *ICD* spectra obtained upon binding to ct-, AT- and GC-DNA.

The data obtained from the EB competitive binding assay have confirmed that the (P)-enantiomers of the parent cylinder bind most strongly to ct-DNA and that in general the $L_3$ metal complex binds less strongly than the $L_1$ and $L_5$ cylinders. In order to fully analyse and interpret the results obtained, a comparison of the binding efficiency and extent of coiling (or loss of orientation) upon addition of the metal complexes to ct-DNA is required. The plot, see figure 3.21 shows the loss in *LD* signal intensity at 260 nm, upon addition of the enantiomers of the methylated cylinders to ct-DNA. The data

have been normalised to a starting intensity of –1 mdeg so that a direct comparison of all cylinders can be made. The % loss in *LD* signal at 260 nm has been calculated by assuming that the ct-DNA (with no metal complex) shows 100% orientation and that upon addition of metal complex the signal decreases at different rates for the different enantiomers. The results presented in table 3.4 show the % loss in *LD* signal at 260 nm at a constant metal complex concentration of 10 μM (i.e. 50:1, DNA base pairs: metal complex).

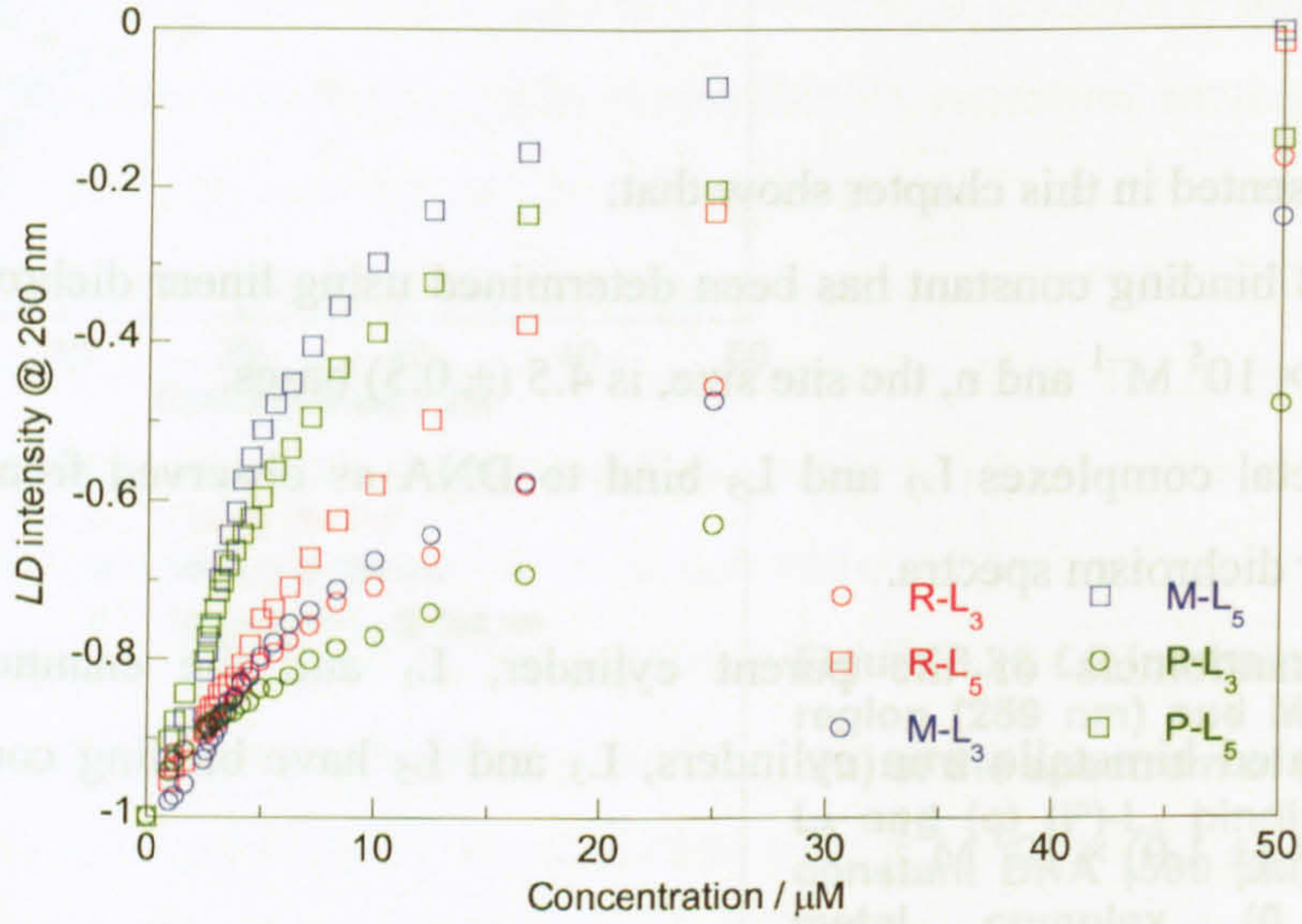



**Figure 3.21 *LD* signal intensity at 260 nm upon addition of metal complex at increasing concentration to a ct-DNA solution, as displayed in legend. Data normalised to starting intensity of –1 mdeg to show 100% orientation of ct-DNA (no metal complex).**

There are several explanations as to why upon addition of the metal complex the *LD* signal at 260 nm decreases: (i) an increase in the flexibility of the DNA or (ii) a shortening of the DNA by kinking, bending, compaction or aggregation. Previous results obtained by I. Meistermann *et al.*[1 – 4] using *CD* and *LD* spectroscopy confirms that the double-stranded B-DNA structure is retained upon addition of the parent cylinder, $L_1$ and that the local structure is not significantly perturbed. This suggests that kinking is not the reason for DNA shortening and that the results from the *LD* binding studies in the MLCT region of the spectrum are consistent only with shortening of the DNA through bending.[2]

As described earlier, the enantiomers of the $L_3$ cylinder bind significantly less strongly than the $L_1$ and $L_5$ metal complexes. This decreased binding efficiency is also reflected in the ability of this metal complex to coil the DNA (see figure 3.21). The results obtained from the *LD* binding experiments confirm that the enantiomers of the $L_3$ cylinder upon binding to ct-DNA show a significant decrease in the % loss of *LD* signal

intensity at 260 nm. It has been shown that the (P)-$L_3$ enantiomer binds least strongly to ct-DNA and it also shows the smallest % loss in *LD* signal intensity at 260 nm (see table 3.4). Further investigations will be required to probe the different binding geometries of this metal complex to DNA. The binding strength and bending are in the same order, with the exception of (R)-L1 and (R)-L5 wherein the bending is less effective than the binding ability suggesting that the (M)- and (P)- enantiomers have opposing bending mechanisms.

The binding efficiency of the $L_5$ metal complex is comparable in strength to that of the parent cylinder, $L_1$. The results from the EB competitive binding assay confirm that the enantiomers of the $L_5$ metal complex displace the EB from ct-DNA in the order (P) < (R) < (M), with the (M)-enantiomer binding most strongly and the racemic metal complex being intermediate in binding strength between the two enantiomers. The results from the *LD* binding experiments suggest that the (M)-$L_5$ enantiomer coils the DNA most significantly while the (R)-$L_5$ cylinder coils the DNA to the least extent (see table 3.4). This may be indicative of different binding modes for the enantiomers and that they bind independently to the DNA when in the racemic mixture with opposing bending effects. At this stage in the research it is difficult to confirm with any confidence the exact binding geometries of the $L_5$ enantiomers.

| | **Weakest** | | | | | | | | **Strongest** |
|---|---|---|---|---|---|---|---|---|---|
| **EB DISPLACEMENT** | P3 | R3 | M3 | P5 | M1 | R5 | M5 | R1 | P1 |
| **% LOSS IN *LD*** | P3 | R3 | M3 | R5 | R1 | P5 | M5 | P1 | M1 |
| **SIGNAL @ 260 nm** | 23 | 29 | 32 | 42 | 60 | 61 | 70 | >95 | >95 |

**Table 3.4 Order of effectiveness of EB displacement by the bimetallo iron cylinders and % loss in *LD* signal at 260 nm. % loss in *LD* signal at 260 nm for the $L_1$ cylinder taken from data presented in reference 2.**

The differences in binding efficiency and % loss in *LD* signal or decrease in orientation of the DNA may be due to the shape and structure of the cylinder upon addition of the methyl group at different positions on the ligand backbone. The aim of the work here was to investigate any sequence preference upon addition of the methyl groups, but the position of the methyl group has also been shown to have a dramatic effect on both the binding and coiling ability of the metal complex. The addition of the methyl group at the 3' position on the outer pyridyl ring, may introduce a certain degree of sequence preference, as the methyl groups project into the groove, but the structure affects the

ability of the metal complex to coil the DNA and the $L_3$ cylinder binds less strongly than the $L_1$ and $L_5$ cylinders. The addition of the methyl group at the 5' position appears to show no sequence selectivity and it can be envisaged that the methyl groups in the 5' position have no additional interactions with the DNA, as they protrude away from the groove. The fact that the methylated metal complexes (i) bind less strongly than the parent cylinder and (ii) induce less dramatic structural changes in the DNA conformation as observed from the *LD* binding studies, suggests that shape and structure are important factors governing the binding geometries of the metal complexes with DNA. The results presented in chapter 4 attempt to investigate further the binding modes and site of interaction of each of the enantiomers.

## 3.6 References

1. Hannon, M. J., Painting, C. L. Jackson, A., Hamblin, J. and Errington, W., *Chem. Commun.* **1997**; 1807 – 1808

2. Hannon, M. J., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Meistermann, I., Isaac, C. J. Sanders, K. J. and Rodger, A. *Angew. Chem. Int. Ed.* **2001**; 40, 5, 880 – 884

3. Hannon, M. J., Painting, C. L. and Alcock, N. W. *Chem. Commun.* **1999**; 20, 2023 – 2024

4. Meistermann, I., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Khalid, S., Rodger, P. M., Peberdy, J. C., Isaac, C. J., Rodger, A. and Hannon, M. J. *Proc. Natl. Acad. Sci. USA.* **2002**; 99, 8, 5069 – 5074

5. Tuite, E and Nordén, B. *J. Am. Chem. Soc.***1994**; 116, 17, 7548 – 7556

6. Waring, M. J., *J. Mol. Biol.* **1965**; 13, 269 – 282

7. Mahler, H. R., Kline, B. and Mehrota, B. D., *J. Mol. Biol.* **1964**; 9, 801 – 811

8. Wang, J. C., *J. Mol. Biol.* **1974**; 89, 783 – 801

9. Mainwaring, W. I. P., Parish, J. H., Pickering, J. D. and Mann, N. H., '*Nucleic acid Biochemistry and Molecular Biology*' Blackwell Scientific Publications, **1982**

10. Lerman, L. S., *J. Mol. Biol.* **1961**; 3, 18 – 30

11. Lerman, L. S., *Proc. Natl. Acad. Sci. USA*, **1963**; 49, 94 – 102

12. Waring, M. J., *J. Mol. Biol.* **1970**; 54, 247 – 279

13. Baguely, B. C. and Falkenhaug, E.-M., *Nucleic Acids Res.* **1978**; 1, 161 – 171

14. Nordén, B. and Tjerneld, F. *Biophys. Chem.*, **1976**; 4, 191 – 198

15. Bugs, M. R. and Cornelio, M. L., *European Biophysics Journal with Biophysics Letters*, **2002**; 31, 3, 232 – 240

16. Vardevanyan, P. O., Antonyan, A. P., Manukyan, G. A., Karapetyan, A. T., Shchyolkina, A. K. and Borisova, O. F., *Mol. Biol.* **2000**; 34, 2, 272 – 276 17. Scaria, P. V. and Shafer, R. H., *J. Biol. Chem.* **1991**; 266, 9, 5417 – 5423

18. Rodger, A. and Nordén, B. '*Circular Dichroism and Linear Dichroism*' Oxford University Press, **1997**

19. Scatchard, G., *Ann. N. Y. Acad. Sci.* **1949**; 51, 660 – 672

20. McCoubrey, A., Latham, H. C., Cook, P. R., Rodger, A., and Lowe, G., *FEBS Letters*, **1996**; 380, 73 – 78

21. Meistermann, I., '*DNA major groove recognition by supramolecular helicates*', Ph.D. thesis, University of Warwick, **2001**

22. Kong, Y., *Biophysical Chemistry*, **2002**; 95, 1 – 6

23. Di Cera, E. and Kong, Y., *Biophysical Chemistry*, **1996**; 61, 107 – 124

# CHAPTER 4 – PROBING THE METAL COMPLEX DNA-BINDING GEOMETRY

## 4.1 Introduction

The results presented in this chapter describe the attempt to investigate and probe the modes of binding of the enantiomers of the parent compound, $L_1$, and the methylated cylinders, $L_3$ and $L_5$. From experiments described above (see § 3.4) it was concluded that the bimetallo iron cylinders do indeed bind to B-DNA as observed by the *ICD* and *LD* spectra upon addition of the metal complex to DNA, but it was not possible to identify the exact binding geometry. A series of competitive binding experiments using Hoechst 33258 (a minor groove binder) and cobalt (III) hexammine (a major groove binding cation) were designed to investigate the binding geometries of the enantiomers of the metal complexes to DNA and are described below. Molecular dynamics, high temperature docking calculations were also used to investigate the binding of the enantiomers of the $L_3$ and $L_5$ metal complexes to a B-DNA dodecamer. Previous research carried out by S. Khalid at the University of Warwick using molecular dynamics simulations, confirmed that the (M)-enantiomers of the parent cylinder $L_1$, bind in the major groove of B-DNA and induce bending and coiling of the DNA.[1] A description of the aims of each experiment is given below followed by the results and discussion.

### 4.1.1 Hoechst 33258

Hoechst 33258 (see figure 4.1) is a non-intercalating compound which binds to the minor groove of DNA with a marked preference for AT sequences.[2 - 4] One dye molecule spans four to five DNA base pairs.[2] The arc shaped, planar unfused aromatic hydrocarbons that have a positive electrostatic potential attracts Hoechst 33258 to the electronegative potential of the minor groove. The size and shape of Hoechst 33258 complements that of the groove, its arc matching the curvature of the DNA double helix.[3] The experiment was designed to investigate the competitive binding of the metal complexes and Hoechst 33258 using fluorescence. It was hoped that differences would be observed in the fluorescence spectra in the presence of the different enantiomers of metal complexes, $L_1$, $L_3$ and $L_5$. The major groove-binding cylinders were not expected to affect the Hoechst 33258-DNA complex, whereas any potential minor–groove binding enantiomers would displace the Hoechst 33258 by virtue of having a stronger binding constant.

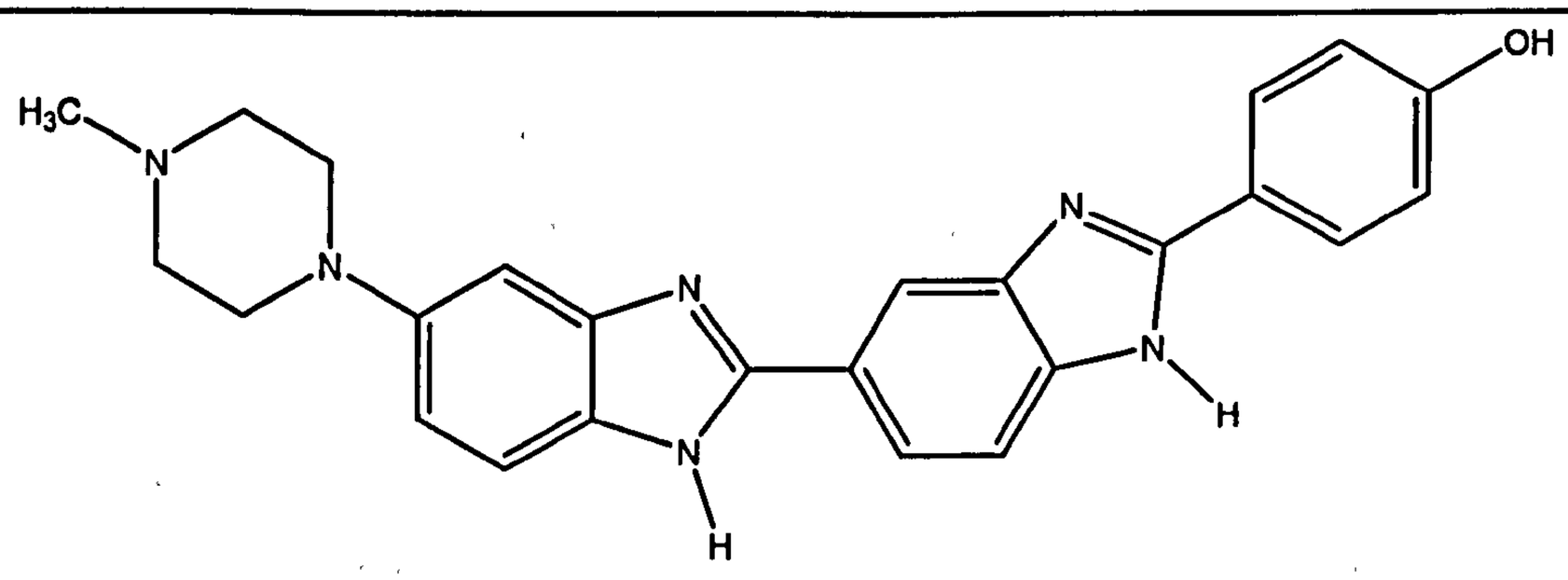



**Figure 4.1 Structure of Hoechst 33258.**

### 4.1.2 Cobalt (III) hexammine

Cobalt (III) hexammine (see figure 4.2) is a DNA-binding cation that converts G-rich B-DNA to Z-DNA by binding in the major groove.[5] It has a ct-DNA binding constant of $1 \times 10^5$ $M^{-1}$ at 11 mM NaCl.[6] G-rich B-DNA is forced into the Z-form using high ionic strength and/or highly charged cations.[7 – 9] It is important to note that Z-DNA, a left-handed helix is not a mirror image of B-DNA, as both owe their chirality to the same ribose sugars in their backbone.[5] An assay was developed by Dr. J. Malina at the University of Warwick,[10] which probed the extent of B → Z-DNA conversion upon addition of cobalt (III) hexammine. B-DNA has a characteristic *CD* spectrum showing a negative *CD* band at approximately 260 nm and a positive *CD* band at approximately 280 nm. In contrast, Z-DNA has a positive *CD* band at approximately 270 nm and a negative absorbance band at 290 nm. The *CD* at 290 nm was monitored upon addition of $[Co(NH_3)_6]^{3+}$ to confirm the conversion from B- to Z-DNA. A solution was prepared which contained the metal complex bound to GC-DNA at 50 mM NaCl concentration. Cobalt (III) hexammine was titrated into this solution. The rationale was that if the bimetallo iron cylinders were binding in the major groove of B-DNA, then addition of cobalt (III) hexammine would not result in conversion from B → Z DNA as the cylinders have a higher binding constant than cobalt (III) hexammine and would not be displaced from the major groove. This would confirm major-groove binding of the cylinders. The competitive binding of cobalt (III) hexammine with (R)-, (M)- and (P)-$L_1$, $L_3$ and $L_5$ was investigated.

**Figure 4.2 Structure of cobalt (III) hexammine.[11]**

### 4.1.3 High temperature docking calculations

A series of high temperature docking calculations were performed to investigate the potential starting configurations of several cylinders to a dodecamer of B-DNA. Molecular dynamics simulations are an ideal complement to experimental techniques as they can reveal insights into the underlying biophysics of systems at the molecular level. It has not been possible to determine the binding mode of all the cylinders with experimental techniques such as NMR and X-ray crystallography. It was hoped that by allowing the cylinders to explore the immobilised DNA molecule and monitoring the energy, potential binding sites could be identified. DNA-cylinder conformations with the most favourable configurational (potential) energy were energy minimised to identify the most favourable configurations from those sampled during the docking procedure. These configurations would then be used to start solvated simulations in which the DNA would be mobile. However, these were beyond the scope of the present study.

The B-DNA (sequence d$(CCCCCAAAAACC)_2$), and cylinders were parameterised using the CHARMM22 all-atom-force-field,[12] and simulations were performed using DL_POLY.[13] The conversion from CHARMM to DL_POLY force-field formats was achieved using a program written by P. M. Rodger that interpreted the CHARMM prm and crd files. All hydrogen atoms were assigned a mass of 2 allowing the use of a 2 fs time step. The Ewald summation[14] was used to evaluate the electrostatic interactions. Simulations were performed in the *NVT* (constant number of molecules, constant volume and constant temperature) ensemble using the Hoover thermostat with a time constant of 0.1 ps and a time step of 2 fs. Orthorhombic periodic boundary conditions were used to mimic an infinite system. The validity of this approach for the cationic cylinders had been verified by S. Khalid.[1]

## 4.2 Materials and methods

Ultrapure water (18.2 ΩM) was used in all experiments. The ct-DNA, GC-DNA, sodium cacodylate buffer, NaCl and metal complexes were prepared as described in § 3.4 and the concentration determined spectrophotometrically using $\varepsilon_{260}$ = 6,600 $cm^{-1}$ $mol^{-1}$ $dm^3$ for ct-DNA[15] $\varepsilon_{254}$ = 8,400 $cm^{-1}$ $mol^{-1}$ $dm^3$ for GC-DNA[16] and $\varepsilon_{572}$ = 16,900 $cm^{-1}$ $mol^{-1}$ $dm^3$ for the metal complexes. Hoechst 33258 and cobalt (III) hexammine were purchased from Sigma-Aldrich, dissolved in water and used without any further purification. Stock solutions of Hoechst 33258 (200 μM, 1.248 mg in 10 mL $H_2O$) and cobalt (III) hexammine (0.01 M, 2.675 mg in 1 mL $H_2O$) were prepared.

### 4.2.1 Hoechst 33258 competitive binding assay

A 3.0 mL solution of ct-DNA (80 μM), Hoechst 33258 (8 μM), NaCl (50 mM) and buffer (1 mM) was prepared. The metal complex concentration was slowly increased for DNA:metal complex ratios, 500:1 to 10:1, keeping the DNA and Hoechst 33258 concentrations constant. After each addition the fluorescence was recorded and the experiments were performed in duplicate to ensure reproducibility.

LS50B parameters: emission: 450 nm; excitation: 350 nm; scan range: 600 – 350 nm; excitation slit: 5.0 nm; emission slit: 5.0 nm; scan speed 500 nm/min; accumulation: 3 scans.

### 4.2.2 Cobalt (III) hexammine competitive binding assay

A 0.5 mL solution of GC-DNA (60 μM), metal complex (6 μM), NaCl (50 mM) and buffer (1 mM) was prepared. The *CD* spectra were recorded, and the cobalt (III) hexammine concentration was slowly increased for DNA:cobalt (III) hexammine concentrations, 5 μM to 50 μM. After each addition the *CD* spectra were recorded. Spectra were collected in a 1 cm masked cuvette. A water spectrum was subtracted from each sample spectrum and the data zeroed in a region of no absorbance at 400 nm. Experiments were performed in duplicate to ensure reproducibility.

*CD* parameters: scan range: 400 – 220 nm; scan speed: 200 nm/min; step resolution: 0.5 nm; response: 0.5 s; band width: 2.0 nm; sensitivity: 20 mdeg; accumulation: 4 scans.

### 4.2.3 High temperature docking calculations

A duplex B-DNA dodecamer and one cylinder molecule were embedded in a neutralizing atmosphere of $Na^+$ ions, with the DNA aligned along the z axis. Further $Na^+/Cl^-$ ion pairs were added to the simulation cell to achieve an ionic strength of 0.8 M, which is similar to that used in other MD studies.[17, 18] The DNA was immobilised and an MD simulation was performed for 2 ns at 900 K. A relative dielectric constant of 80 $C^2 J^{-1} m^{-1}$ is used while the system is not solvated in water, so that the negatively charged DNA is shielded (to prevent aggregation) from the $Na^+$ cations at high temperature. From the trajectory generated during the high temperature docking simulation, conformations of low configurational energy were identified and extracted. These frames were then energy minimised at 210 K for 5 ps using a dielectric constant of 1 $C^2 J^{-1} m^{-1}$ to further optimise the geometry.

## 4.3 Results

### 4.3.1 Hoechst 33258 competitive binding assay

As described above, Hoechst 33258 is a minor groove binding drug. Upon interaction with DNA its fluorescence is significantly increased. The fluorescence spectra presented in figure 4.3 show the displacement of Hoechst 33258 upon addition of the racemic $L_1$, $L_3$ and $L_5$ cylinders. The fluorescence spectra of the (M)- and (P)-enantiomers are presented in appendix III. It can be observed that the metal complexes displace Hoechst 33258 in stoichiometric amounts with increasing cylinder concentration.

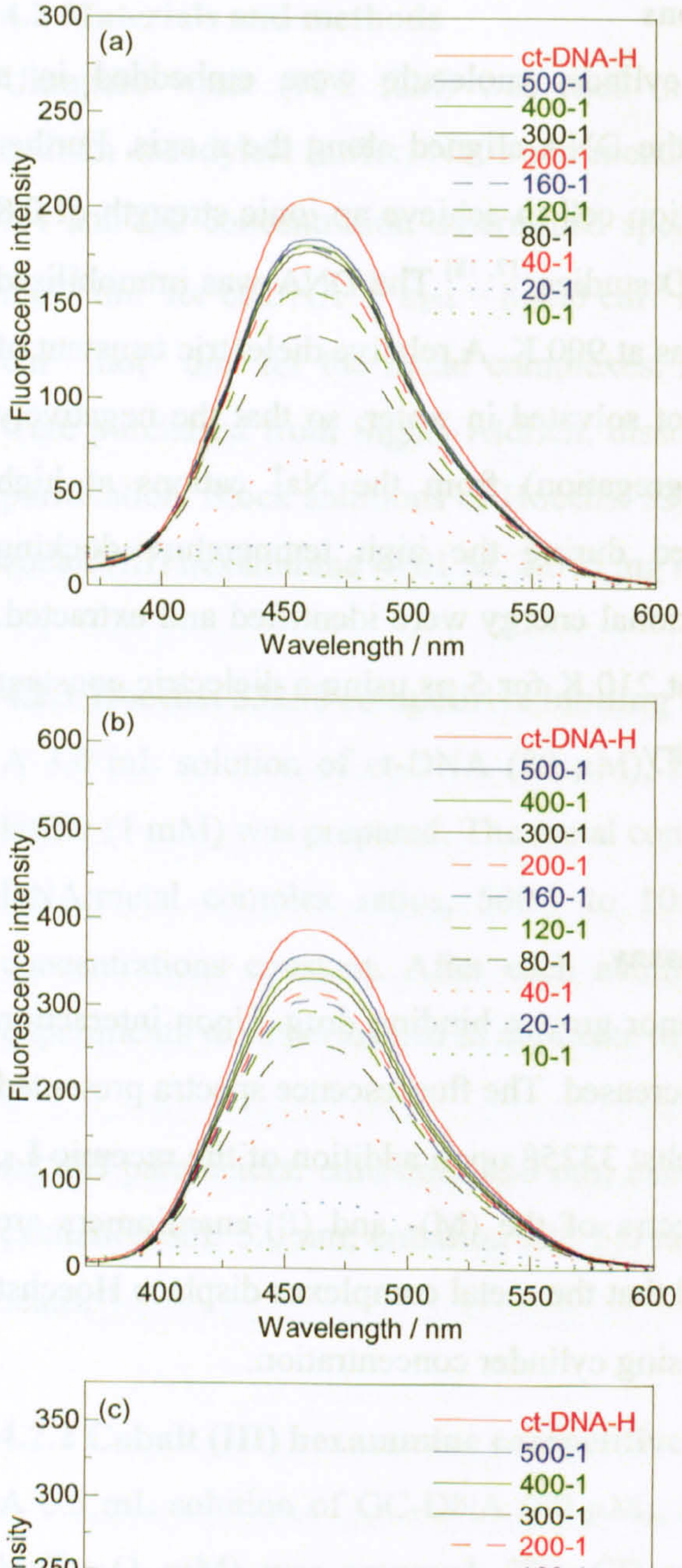



**Figure 4.3 Fluorescence emission spectra of (a) ct-DNA-Hoechst upon addition of (R)-$L_1$, (b) ct-DNA-Hoechst upon addition of (R)-$L_3$ and (c) ct-DNA-Hoechst upon addition of (R)-$L_5$. Ct-DNA-H = ct-DNA-Hoechst 33258 complex (80 μM DNA:8 μM Hoechst 33258. Spectra were collected in a 1 cm path length cuvette and the intrinsic Hoechst 33258 fluorescence spectrum subtracted from each sample spectrum.**

Figure 4.4 shows the fluorescence intensity at the wavelength maximum of the Hoechst 33258 upon binding to ct-DNA, versus cylinder concentration. The first observation that can be made is that the parent cylinder and the methylated derivatives displace Hoechst 33258. This was not expected for the (M)-enantiomer of the parent cylinder which has

been shown to bind in the major groove of B-DNA.[19, 20] The most likely explanation for these results is that when the cylinders bind to DNA they induce remarkable conformational changes. As a consequence of the conformational change the minor groove loses it shape and hence its favourable contacts with Hoechst 33258, forcing the Hoechst 33258 to disassociate from the DNA even with major groove binders. In other words although upon addition of the cylinders, the Hoechst 33258-DNA complex disassociates, this is not necessarily a consequence of competition for the minor groove.

The results of the EB competitive binding assay have confirmed that the (M)-$L_1$ cylinder binds strongly to ct-DNA and NMR studies have shown that the (M)-enantiomers bind in the major groove. The fact that the (M)-enantiomer displaces the Hoechst 33258 least effectively may be a consequence of major groove binding by the (M)-$L_1$ cylinder or that the extent of coiling does not make the minor groove uncondusive to groove binding by the Hoechst 33258. The (P)-$L_1$ cylinders displace the Hoechst most effectively and by comparing the results of the EB competitive binding assay and the % loss of *LD* signal at 259 nm, it can be concluded that the (P)-$L_1$ enantiomers bind strongly to ct-DNA and displace the Hoechst 33258 by binding inducing structural changes in the DNA conformation.

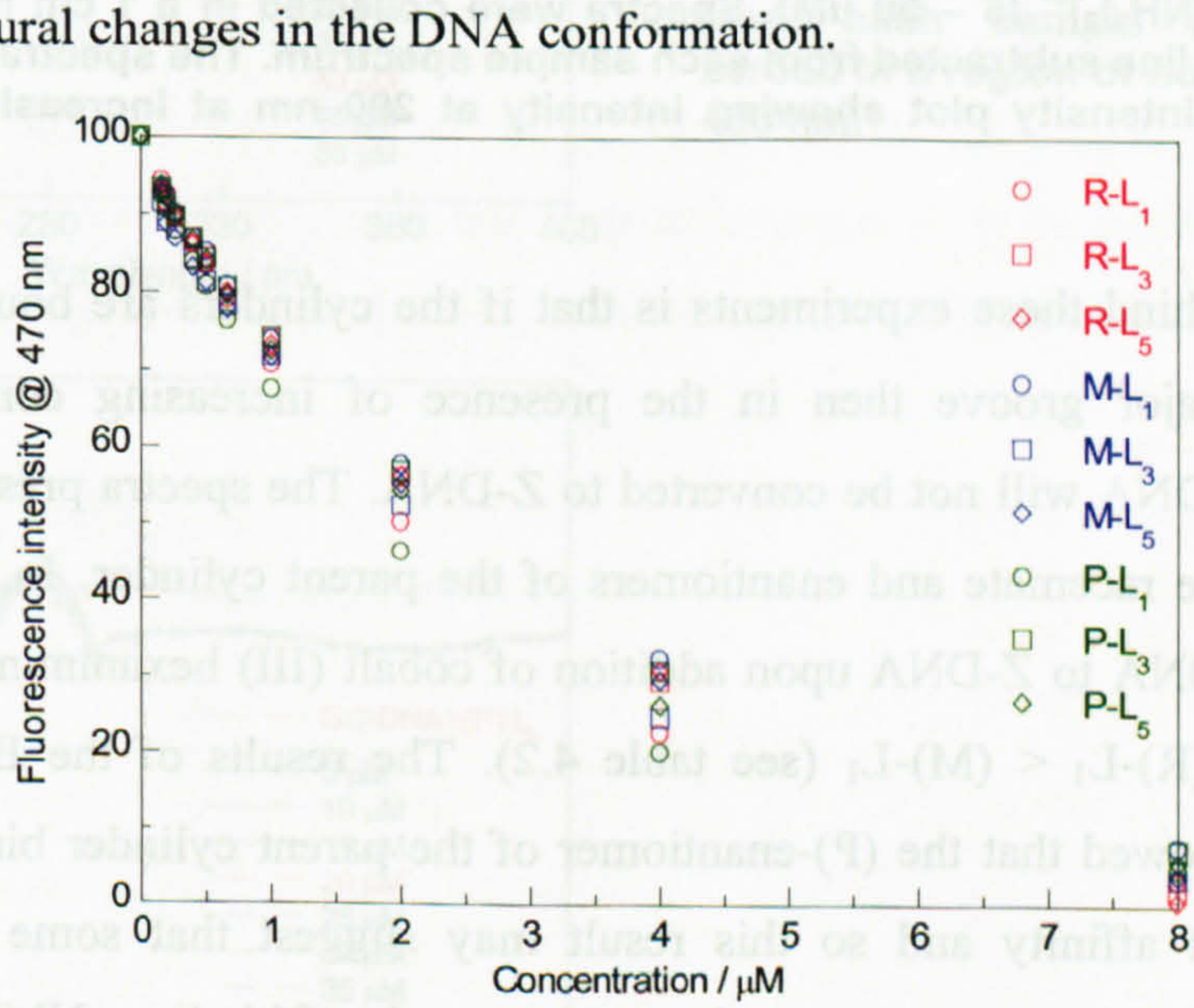



**Figure 4.4 Plot of fluorescence emission as a function of cylinder concentration. The data were normalised to a starting intensity of 100 for comparison.**

| | Weakest | | | | | | | Strongest |
|---|---|---|---|---|---|---|---|---|
| **HOECHST 33258** | (M)-$L_1$ | (M)-$L_5$ | (P)-$L_3$ | (R)-$L_5$ | (R)-$L_3$ | (M)-$L_3$ | (P)-$L_5$ | (R)-$L_1$ | (P)-$L_1$ |

**Table 4.1 Order of effectiveness of Hoechst 33258 displacement by the bimetallo iron cylinders.**

### 4.3.2 Cobalt (III) hexammine competitive binding assay

Circular dichroism was utilised to probe the transition from B- → Z-form DNA upon addition of cobalt (III) hexammine, $[Co(NH_3)_6]^{3+}$ to a GC-DNA solution. The spectra presented in figure 4.5 show the B- → Z-transition of DNA upon addition of increasing concentrations of $[Co(NH_3)_6]^{3+}$ and the intensity of the *CD* signal at 290 nm as the concentration of $[Co(NH_3)_6]^{3+}$ was increased.

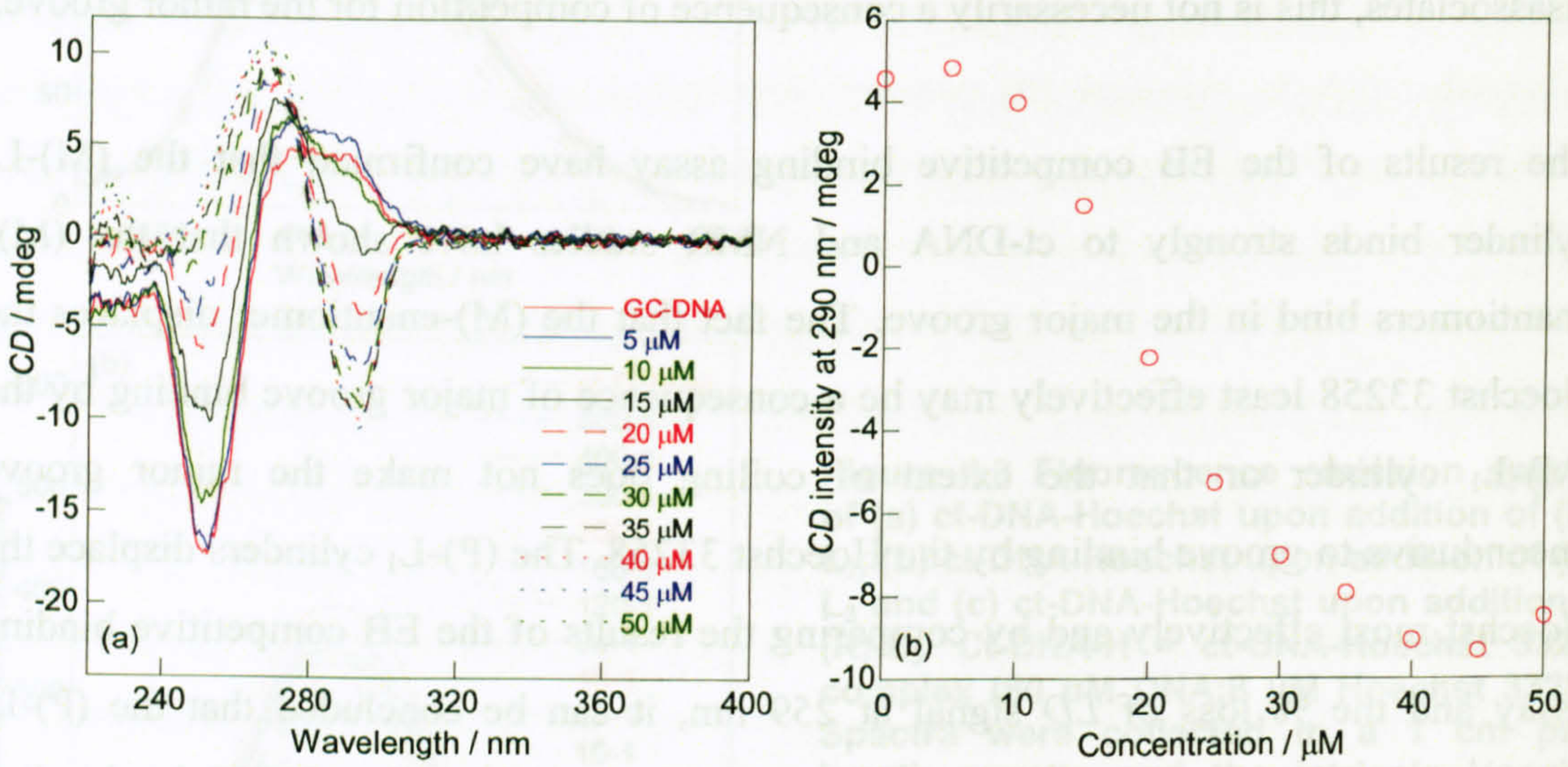



**Figure 4.5 (a) *CD* spectrum of GC-DNA in 50 mM NaCl and 1 mM cacodylate buffer with additions of $[Co(NH_3)_6]^{3+}$ (5 – 50 μM). Spectra were collected in a 1 cm masked cuvette and a water base line subtracted from each sample spectrum. The spectra were zeroed at 400 nm. (b) *CD* intensity plot showing intensity at 290 nm at increasing $[Co(NH_3)_6]^{3+}$ concentration.**

The rationale behind these experiments is that if the cylinders are bound to the GC-DNA in the major groove then in the presence of increasing concentrations of $[Co(NH_3)_6]^{3+}$ B-DNA will not be converted to Z-DNA. The spectra presented in figure 4.6 show that the racemate and enantiomers of the parent cylinder, $L_1$, all inhibit the transition of B-DNA to Z-DNA upon addition of cobalt (III) hexammine and is of the order (P)-$L_1$ < (R)-$L_1$ < (M)-$L_1$ (see table 4.2). The results of the EB competitive binding assay showed that the (P)-enantiomer of the parent cylinder binds to ct-DNA with the highest affinity and so this result may suggest that some of the (P)-$L_1$ enantiomers are binding via a non-major groove mode of binding. NMR studies have confirmed that the (M)-enantiomers bind in the major groove[20, 21] and the results presented here confirm that upon addition of $[Co(NH_3)_6]^{3+}$ to the GC-DNA-(M)-$L_1$ solution the B → Z-DNA transition is less dramatic.

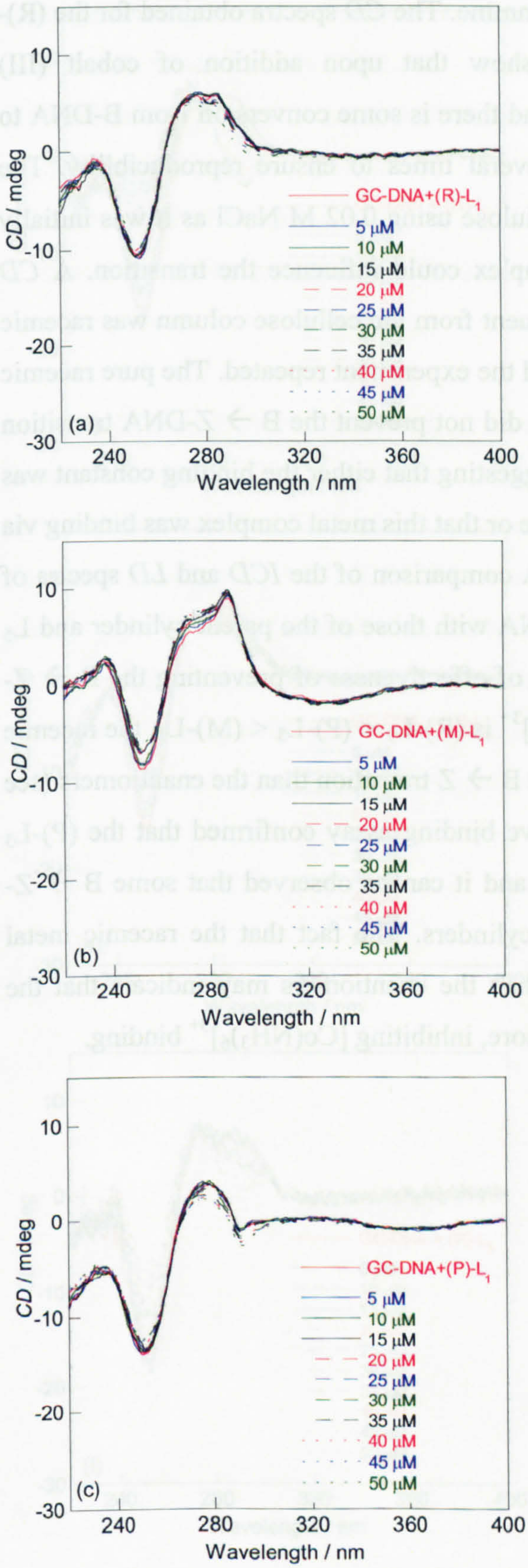



**Figure 4.6 *CD* spectra showing (a) (R)-$L_1$ (b) (M)-$L_1$ and (c) (P)-$L_1$, bound to GC-DNA with increasing $[Co(NH_3)_6]^{3+}$ (5 – 50 μM) concentration. GC-DNA = 60 μM, metal complex = 6 μM, NaCl = 50 mM and pH 6.8 buffer = 1 mM. Data were collected in a 1 cm path length cuvette. Water and the corresponding enantiomer spectra were subtracted from each sample spectrum and zeroed in a region of no absorbance at 400 nm.**

The spectra obtained for the racemate and enantiomers of the $L_3$ cylinder (see figure 4.7) show some interesting results. Preliminary investigations involved the competitive binding of the three racemic cylinders with cobalt (III) hexammine. If the cylinders were major groove binding, it was expected that little or no B → Z transition would be

observed upon addition of cobalt (III) hexammine. The *CD* spectra obtained for the (R)-$L_3$ cylinder when bound to GC-DNA show that upon addition of cobalt (III) hexammine, the iron helicate is displaced and there is some conversion from B-DNA to Z-DNA. This experiment was repeated several times to ensure reproducibility. The racemic metal complex was purified on cellulose using 0.02 M NaCl as it was initially presumed that impurities in the metal complex could influence the transition. A *CD* spectrum was collected to ensure that the eluent from the cellulose column was racemic (flat line signal from 750 nm to 200 nm) and the experiment repeated. The pure racemic $L_3$ metal complex when bound to GC-DNA did not prevent the B → Z-DNA transition upon addition of cobalt (III) hexammine suggesting that either the binding constant was lower than that of the cobalt (III) hexammine or that this metal complex was binding via a non-major groove mode of binding site. A comparison of the *ICD* and *LD* spectra of (R)-$L_3$ upon binding to ct-, AT- and GC-DNA with those of the parent cylinder and $L_5$ cylinder suggest it is the former. The order of effectiveness of preventing the B → Z-DNA transition upon addition of $[Co(NH_3)_6]^{3+}$ is (R)-$L_3$ < (P)-$L_3$ < (M)-$L_3$, the racemic metal complex being less able to prevent the B → Z transition than the enantiomers (see table 4.2). The results of the EB competitive binding assay confirmed that the (P)-$L_3$ enantiomers bind least strongly to ct-DNA and it can be observed that some B → Z-DNA conversion occurs for all of the $L_3$ cylinders. The fact that the racemic metal complex shows more B → Z conversion than the enantiomers may indicate that the enantiomers destabilise the DNA structure more, inhibiting $[Co(NH_3)_6]^{3+}$ binding.

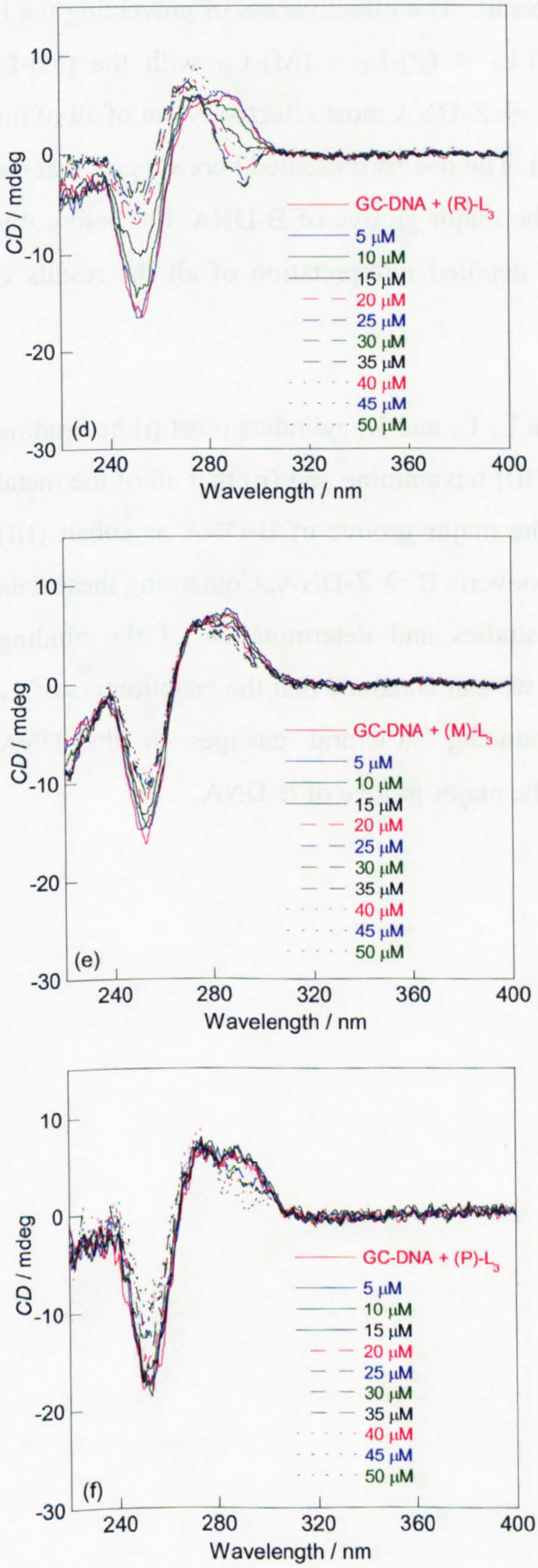



**Figure 4.7 *CD* spectra showing (a) (R)-$L_3$ (b) (M)-$L_3$ and (c) (P)-$L_3$, bound to GC-DNA with increasing $[Co(NH_3)_6]^{3+}$ (5 – 50 μM) concentration. GC-DNA = 60 μM, metal complex = 6 μM, NaCl = 50 mM and pH 6.8 buffer = 1 mM. Data were collected in a 1 cm path length cuvette. Water and the corresponding enantiomer spectra were subtracted from each sample spectrum and zeroed in a region of no absorbance at 400 nm.**

The *CD* spectra presented in figure 4.8 show that as with the $L_3$ cylinder, the racemic $L_5$ metal complex shows more B → Z-DNA conversion than the enantiomers. It was initially assumed that these results were due to impurities present in the racemic metal complex and so it was purified on a cellulose column using 0.02 M NaCl as eluent. *CD*

was used to confirm that the eluent was racemic. The effectiveness of preventing the B → Z-DNA transition is of the order (R)-$L_5$ < (P)-$L_5$ < (M)-$L_5$, with the (M)-$L_5$ enantiomers preventing the conversion of B → Z-DNA most effectively out of all of the metal complexes investigated (see table 4.2). The results presented here suggest that the (M)-$L_5$ enantiomers are binding in or via the major groove of B-DNA, but before this mode of binding can be accepted a more detailed interpretation of all the results is required (see chapter 6).

It was concluded that the enantiomers of the $L_1$ $L_3$ and $L_5$ cylinders must (i) be binding more strongly to GC-DNA than the cobalt (III) hexammine and (ii) that all of the metal complexes investigated were binding via the major groove of B-DNA as cobalt (III) hexammine binds in the major groove and converts B → Z-DNA. Comparing these data to those obtained for the DNA-binding studies and determination of the binding constants for the enantiomers of $L_1$ and $L_5$, we can conclude that the enantiomers of $L_1$ and $L_5$ bind very strongly to DNA inducing structural changes in the DNA conformation and may be binding in or via the major groove of B-DNA.

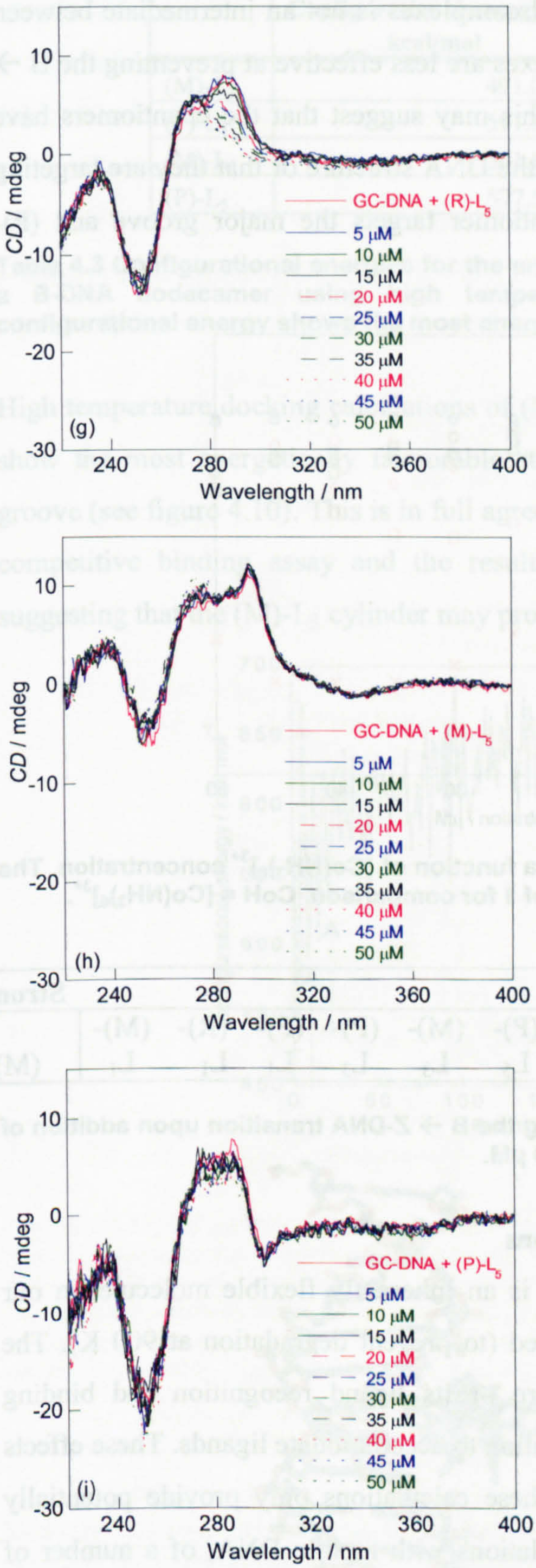



**Figure 4.8 *CD* spectra showing (a) (R)-$L_5$ (b) (M)-$L_5$ and (c) (P)-$L_5$, bound to GC-DNA with increasing $[Co(NH_3)_6]^{3+}$ (5 – 50 μM) concentration. GC-DNA = 60 μM, metal complex = 6 μM, NaCl = 50 mM and pH 6.8 buffer = 1 mM. Data were collected in a 1 cm path length cuvette. Water and the corresponding enantiomer spectra were subtracted from each sample spectrum and zeroed in a region of no absorbance at 400 nm.**

The plot in figure 4.9 shows the *CD* intensity at 290 nm as a function of $[Co(NH_3)_6]^{3+}$ concentration. The most significant B → Z-DNA conversion occurs when the (R)-$L_3$ cylinder is bound to GC-DNA and the least when bound to the (M)-$L_5$ cylinder. The results are quite intriguing, as the order of effectiveness of preventing the B → Z-DNA

transition for the methylated racemic metal complexes is not an intermediate between the enantiomers. The racemic metal complexes are less effective at preventing the B → Z-DNA transition than the enantiomers. This may suggest that the enantiomers have opposing bending effects which destabilise the DNA structure or that they are targeting different sites on the DNA, i.e. (M)-enantiomer targets the major groove and (P)-enantiomer targets the minor groove.

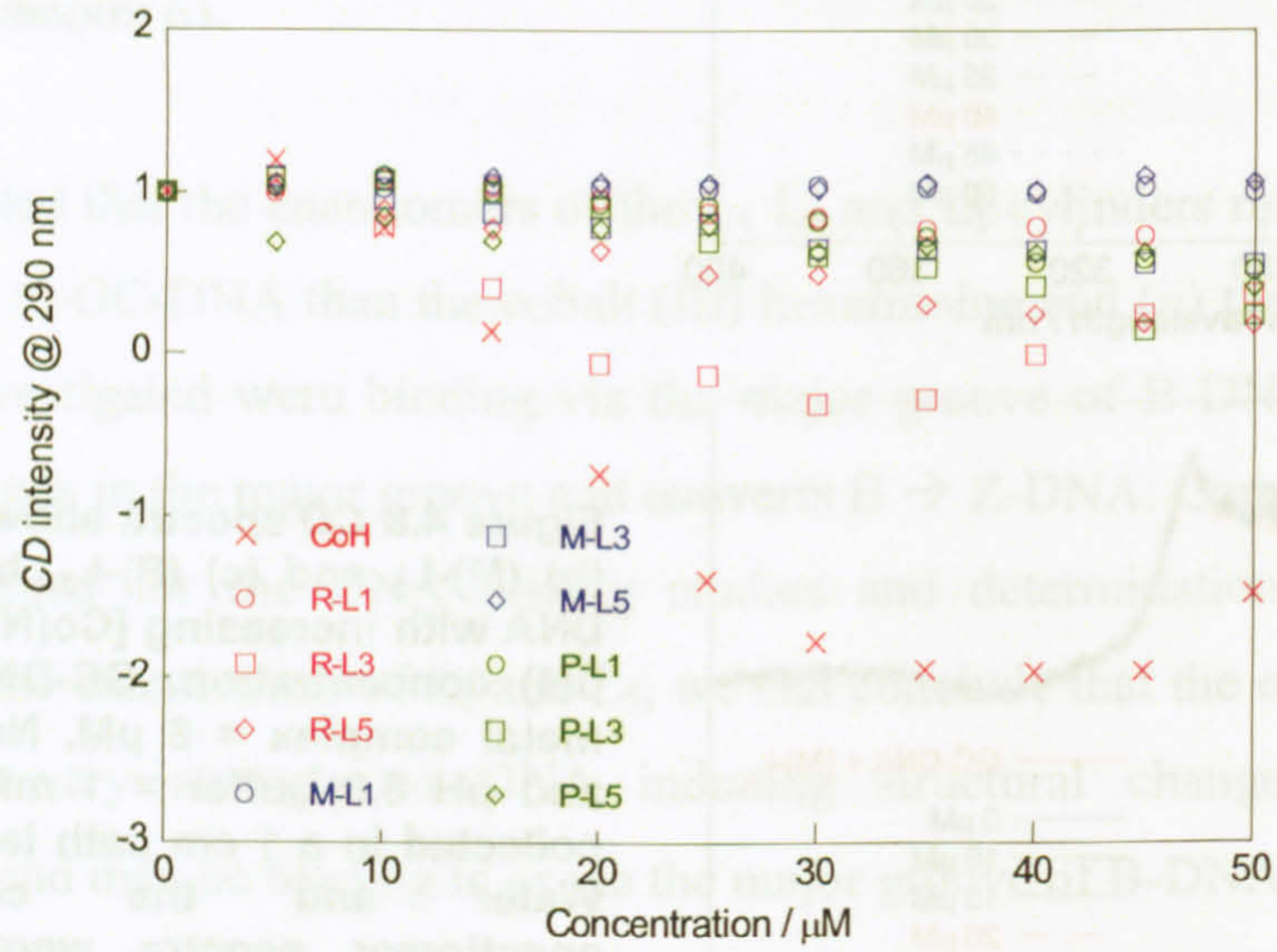



**Figure 4.9 Plot of *CD* intensity at 290 nm as a function of $[Co(NH_3)_6]^{3+}$ concentration. The data were normalised to a starting intensity of 1 for comparison. CoH = $[Co(NH_3)_6]^{3+}$.**

| | Weakest | | | | | | | | Strongest |
|---|---|---|---|---|---|---|---|---|---|
| **COBALT HEXAMMINE** | (R)-$L_3$ | (R)-$L_5$ | (P)-$L_5$ | (M)-$L_3$ | (P)-$L_3$ | (P)-$L_1$ | (R)-$L_1$ | (M)-$L_1$ | (M)-$L_5$ |

**Table 4.2 Order of effectiveness of preventing the B → Z-DNA transition upon addition of $[Co(NH_3)_6]^{3+}$ at a constant concentration of 40 μM.**

### 4.3.3 High temperature docking calculations

It is important to note that whereas DNA is an inherently flexible molecule, in our docking calculations it has been immobilised (to prevent degradation at 900 K). The flexibility of DNA is an important feature of its ligand recognition and binding properties as the DNA can alter its conformation to accommodate ligands. These effects are not considered in our calculations – these calculations only provide potentially suitable binding geometries. Solvated simulations with mobile DNA, of a number of conformations from each docking calculation would provide a more thorough sampling of the binding geometries. These simulations would also reveal insights into the structural effects on the DNA upon cylinder binding. The configurational energy for each starting configuration is given in table 4.3.

| | Configurational energy (a) kcal/mol | Configurational energy (b) kcal/mol |
|---|---|---|
| (M)-$L_3$ | 491.0140 | 529.7920 |
| (P)-$L_3$ | 591.5983 | 571.6043 |
| (M)-$L_5$ | 591.8280 | 552.5046 |
| (P)-$L_5$ | 577.5305 | 579.3221 |

**Table 4.3 Configurational energies for the enantiomers of $L_3$ and $L_5$ upon interaction with a B-DNA dodecamer using high temperature docking calculations. The lowest configurational energy shows the most energetically favourable starting configuration.**

High temperature docking calculations of (M)-$L_3$ interacting with a B-DNA dodecamer show the most energetically favourable starting configurations situated in the major groove (see figure 4.10). This is in full agreement with the data from the Hoechst 33258 competitive binding assay and the results from the cobalt (III) hexammine study suggesting that the (M)-$L_3$ cylinder may provide major groove recognition.

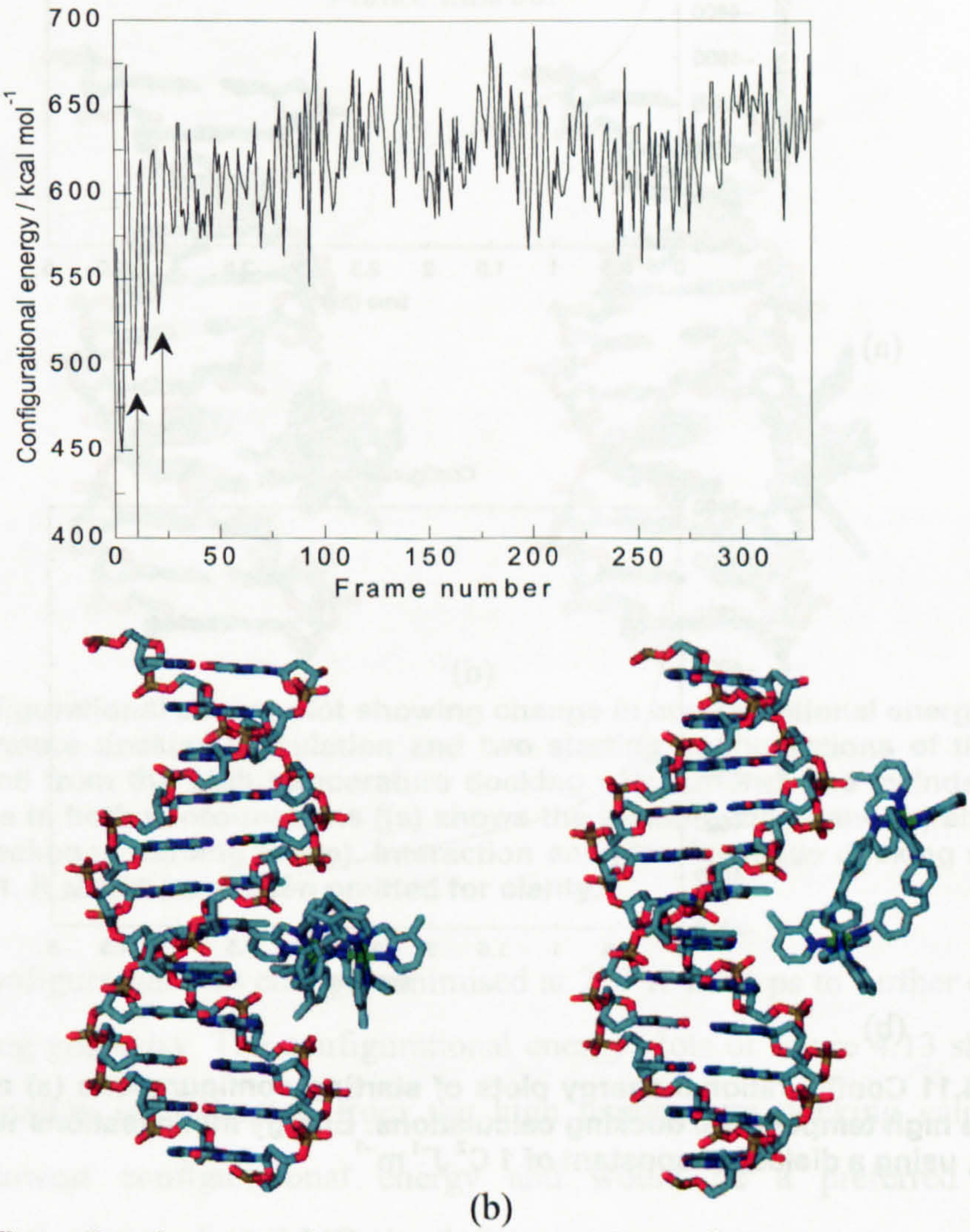
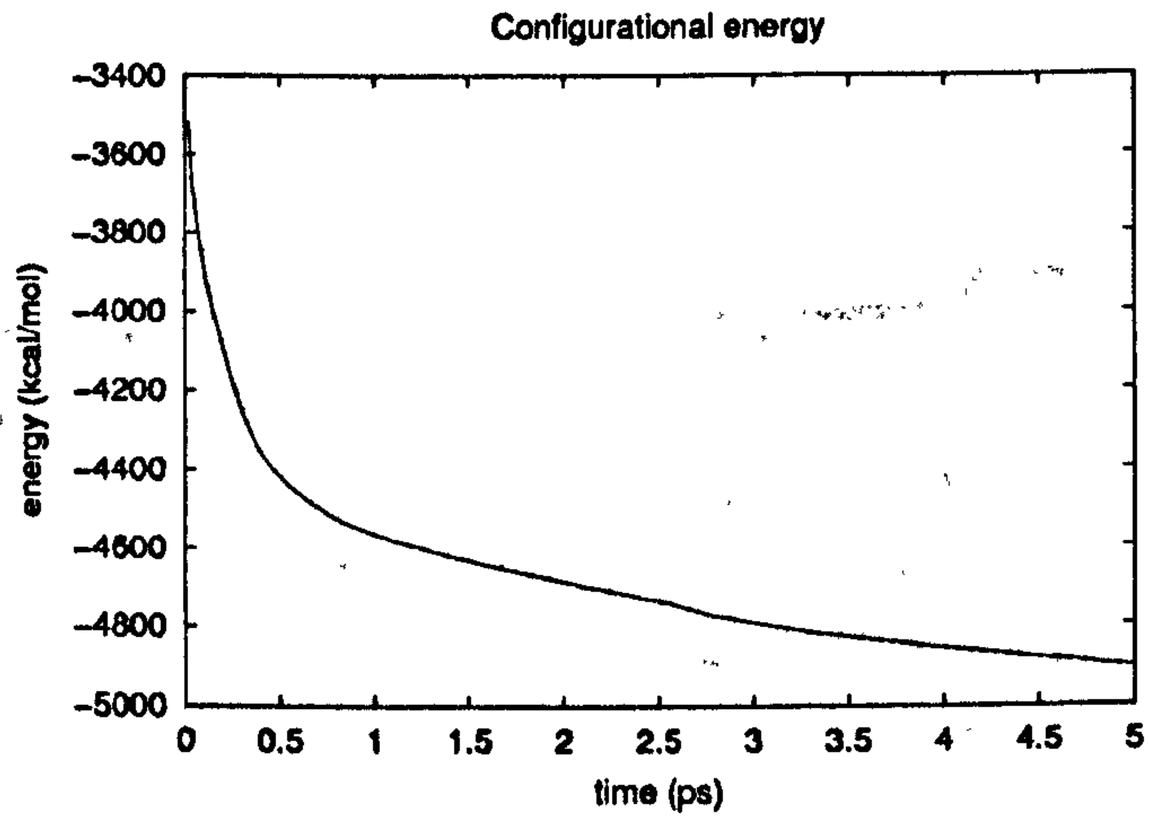



**Figure 4.10 Configurational energy plot showing change in configurational energy during the high temperature docking calculation and two starting configurations of the (M)-$L_3$ cylinder extracted from the high temperature docking calculations. The cylinder lies in the major groove in both conformations ((a) shows the bending mode and (b) shows the major groove backbone binding mode). Interaction energies for these docking sites are given in table 4.1. H atoms have been omitted for clarity.**

Each configuration was further energy minimised at 210 K for 5 ps and the results are presented in figure 4.11. The configurational energy plots showing the change in configurational energy during the energy minimisation process, show that starting configuration (a) from the initial docking calculations is an energetically more favourable starting configuration for the fully solvated MD simulations. This was denoted the major groove bending mode and fully solvated simulations will hopefully show bending and coiling of the DNA as with the (M)-enantiomer of the parent cylinder.




(a)




(b)

**Figure 4.11 Configurational energy plots of starting configurations (a) and (b) extracted from the high temperature docking calculations. Energy minimisations were ran at 210 K for 5 ps, using a dielectric constant of 1 $C^2 J^{-1} m^{-1}$.**

Two starting configurations are shown in figure 4.12 of the (P)-$L_3$ cylinder upon interaction with a B-DNA dodecamer. The high temperature docking calculations have revealed two energetically favourable starting configurations located in the major

groove. Experimental results have suggested that the (P)-$L_3$ enantiomers (i) bind significantly less strongly to DNA and (ii) may bind via an alternative binding mode to that of the other cylinders.

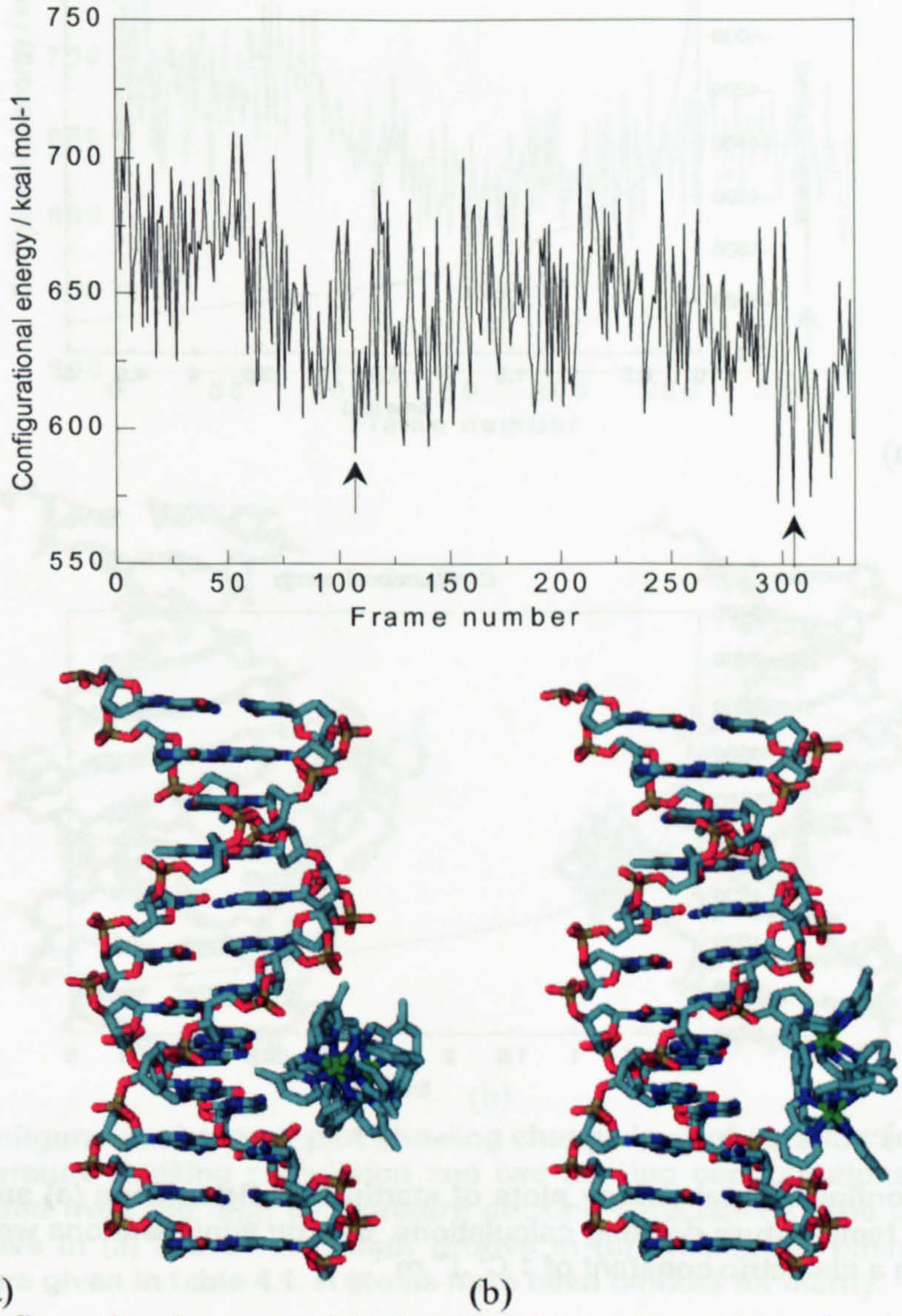
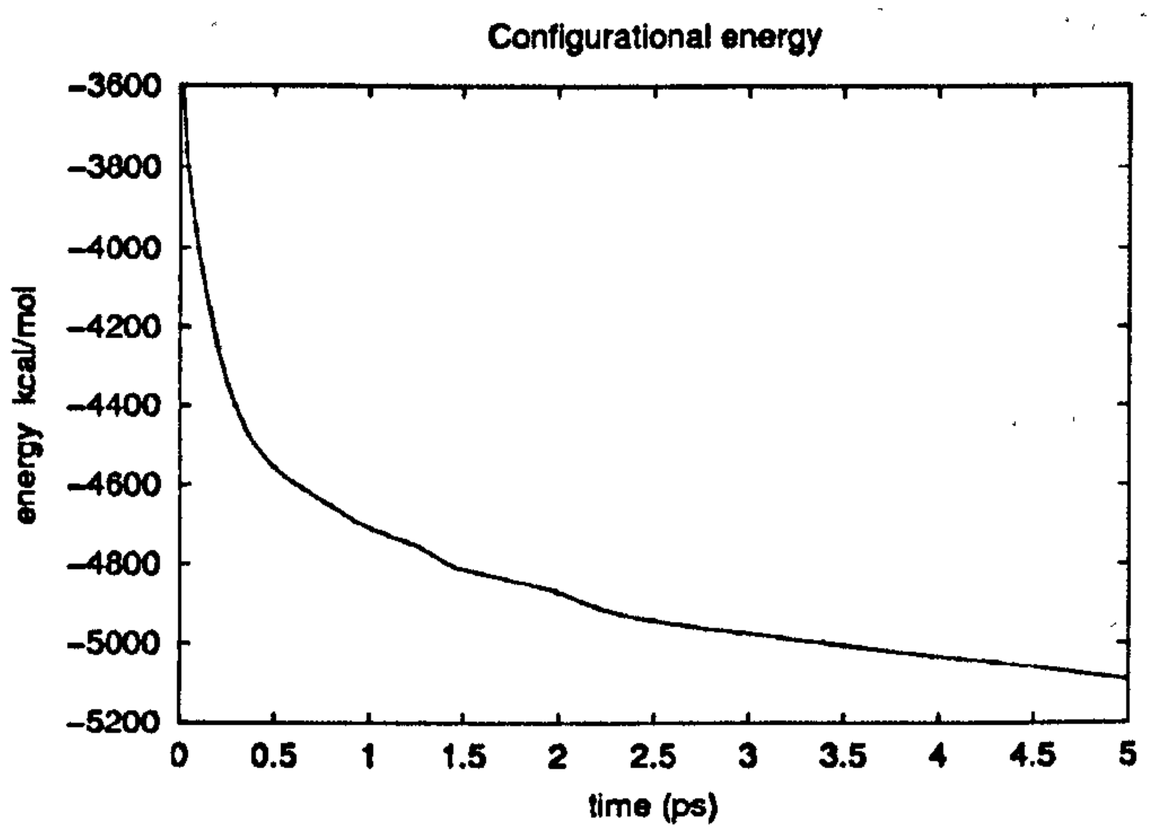
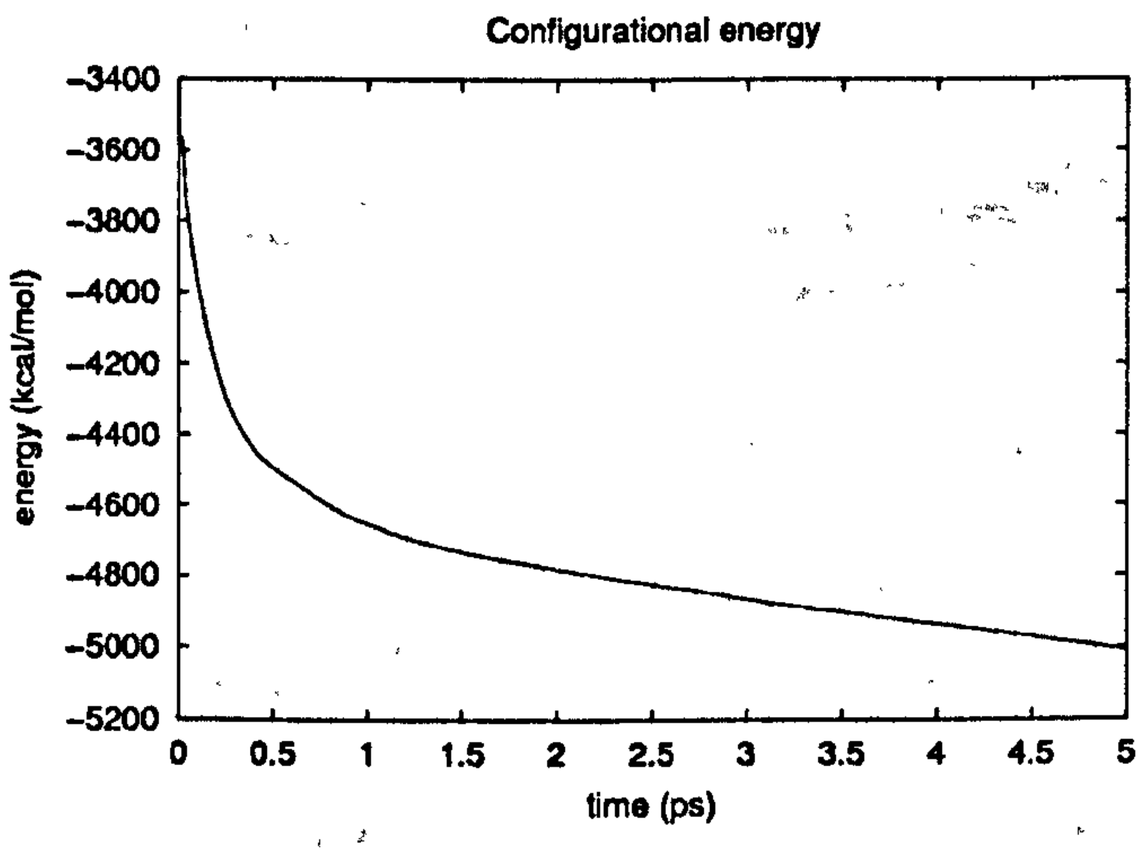



**Figure 4.12 Configurational energy plot showing change in configurational energy during the high temperature docking calculation and two starting configurations of the (P)-$L_3$ cylinder extracted from the high temperature docking calculations. The cylinder lies in the major groove in both conformations ((a) shows the bending mode and (b) shows the major groove backbone binding mode). Interaction energies for these docking sites are given in table 4.1. H atoms have been omitted for clarity.**

Each starting configuration was energy minimised at 210 K for 5 ps to further optimise the initial binding geometry. The configurational energy plots of figure 4.13 show that starting configuration (a) extracted from the high temperature docking calculations provides the lowest configurational energy and would be a preferred starting configuration for the fully solvated MD simulations.




(a)




(b)

**Figure 4.13 Configurational energy plots of starting configurations (a) and (b) extracted from the high temperature docking calculations. Energy minimisations were ran at 210 K for 5 ps, using a dielectric constant of 1 $C^2 J^{-1} m^{-1}$.**

High temperature docking calculations of (M)-$L_5$ interacting with a B-DNA dodecamer show the most energetically favourable starting configurations situated in the major and minor grooves (see figure 4.14). The experimental data suggests that the (M)-$L_5$ cylinder coils the DNA as observed by the decrease in *LD* signal intensity in the DNA region of the spectrum. The two lowest configurational energy configurations show a (a) backbone major groove binding mode and a (b) minor groove bending mode.

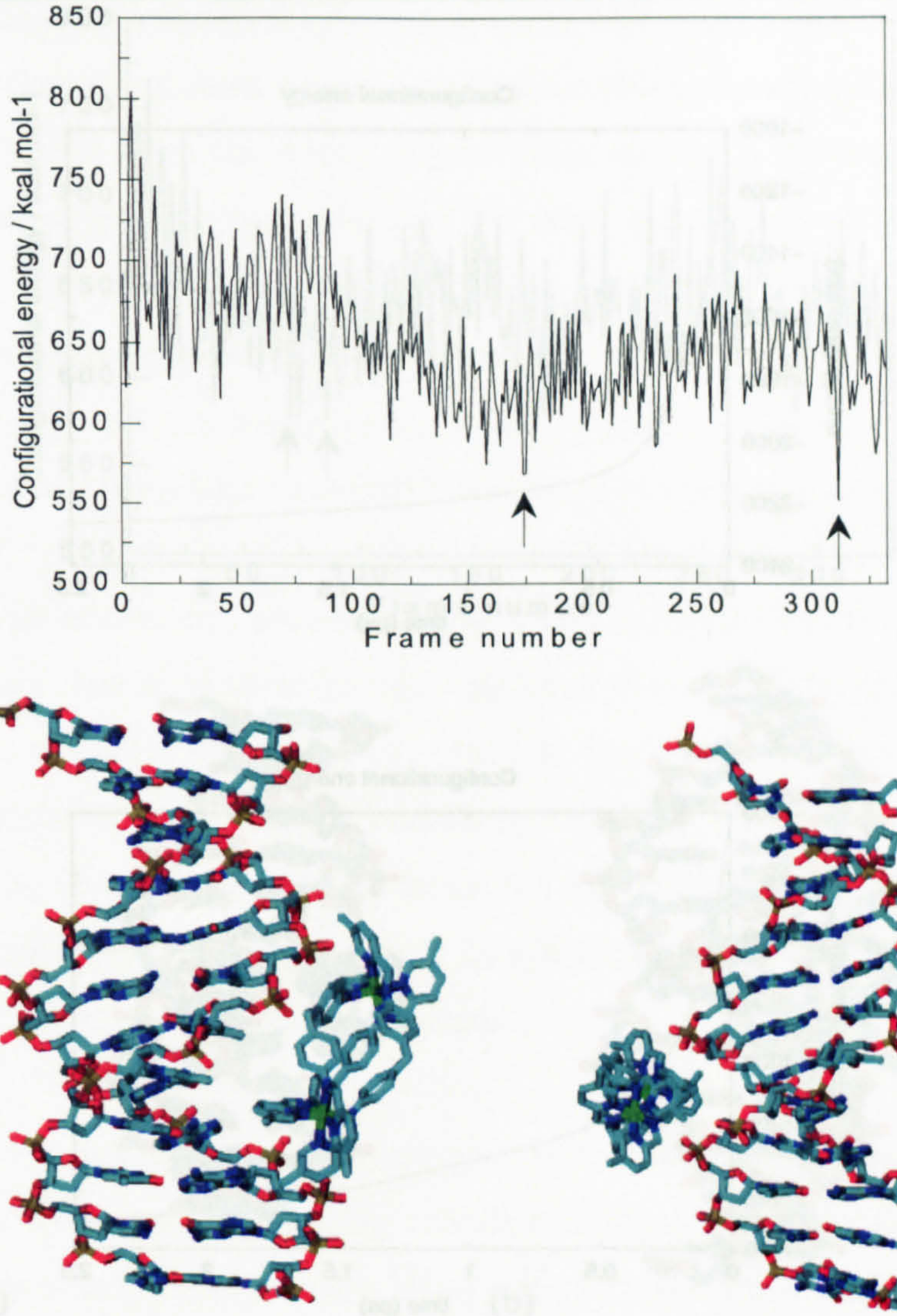
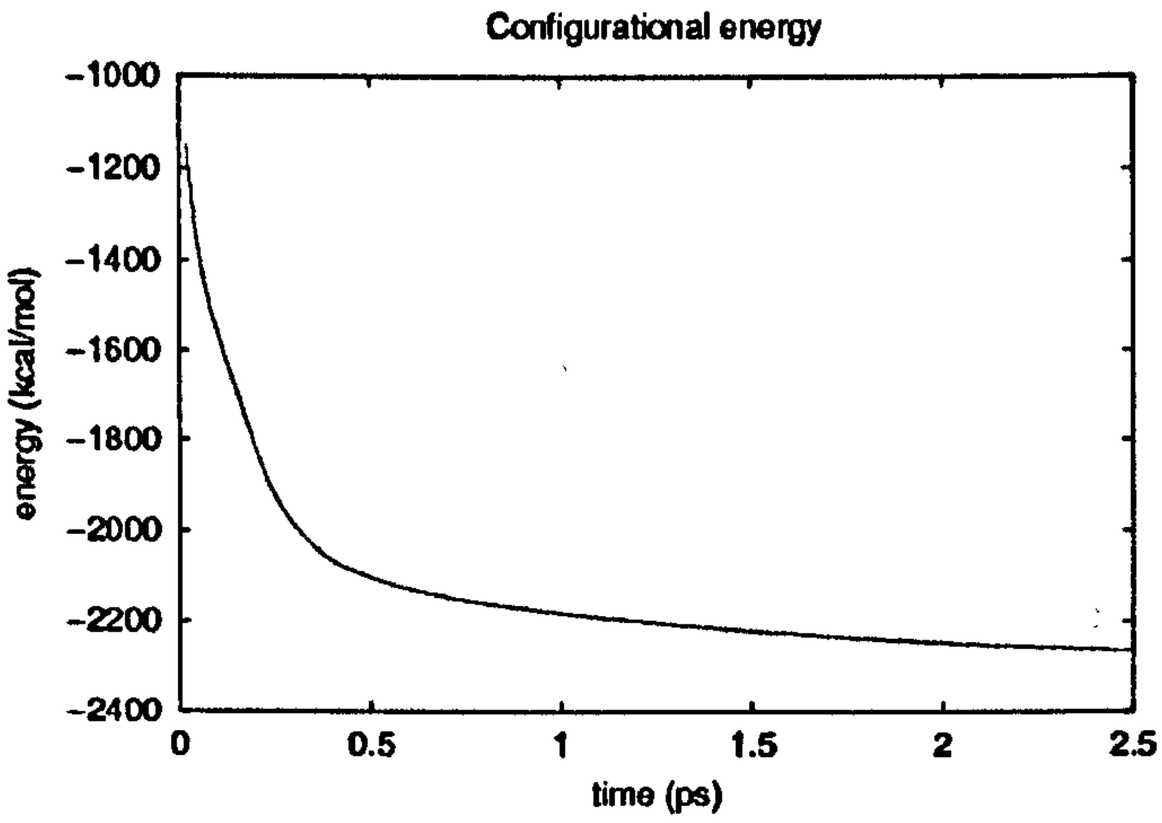
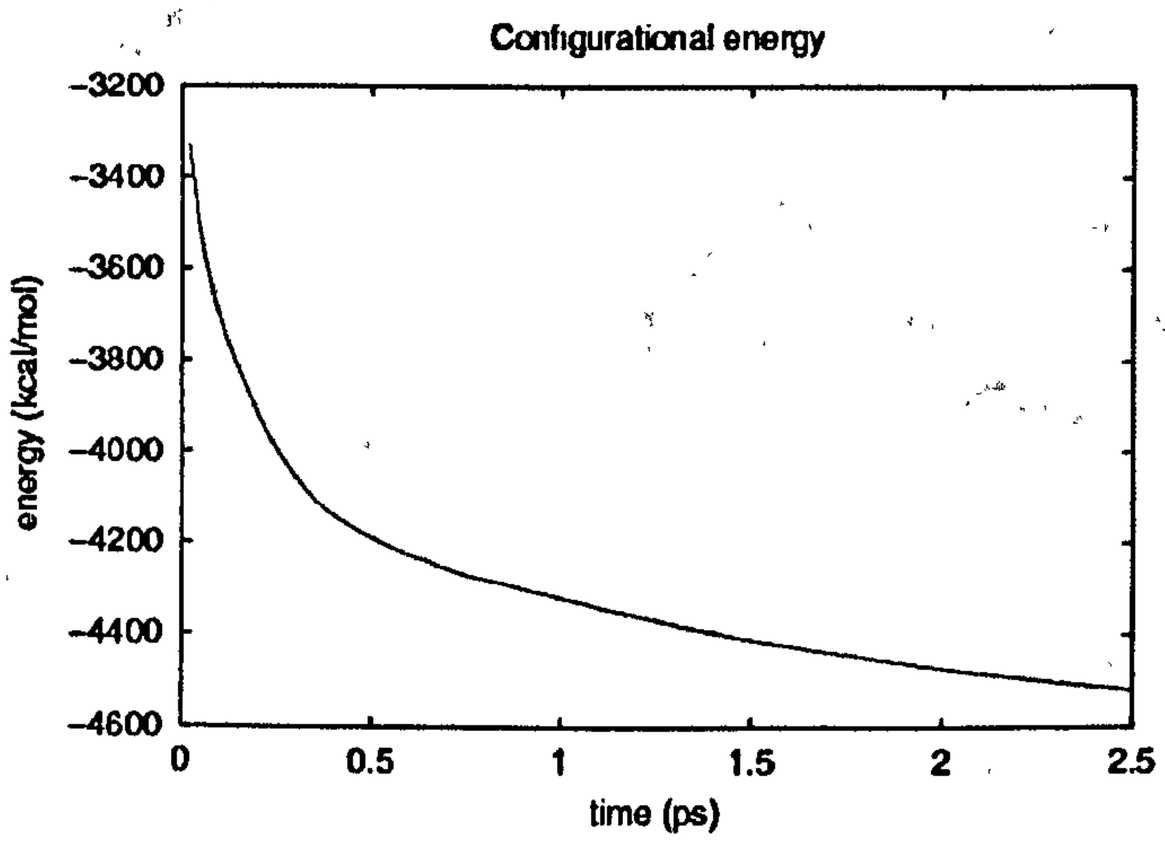



**Figure 4.14 Configurational energy plot showing change in configurational energy during the high temperature docking calculation and two starting configurations of the (M)-$L_5$ cylinder extracted from the high temperature docking calculations. The cylinder lies in the major groove in (a) and in the minor groove in (b). Interaction energies for these docking sites are given in table 4.1. H atoms have been omitted for clarity.**

An example of the energetic differences when the DNA flexibility is taken into account is given below (see figure 4.15). The two binding geometries (one major groove and one minor groove) from the (M)-$L_5$ docking calculations were minimised with mobile DNA and cylinder The two conformations have similar configurational energies initially (see table 4.3) but upon minimisation these energies are substantially different. The minor groove geometry is energetically more favourable than its major groove counterpart.



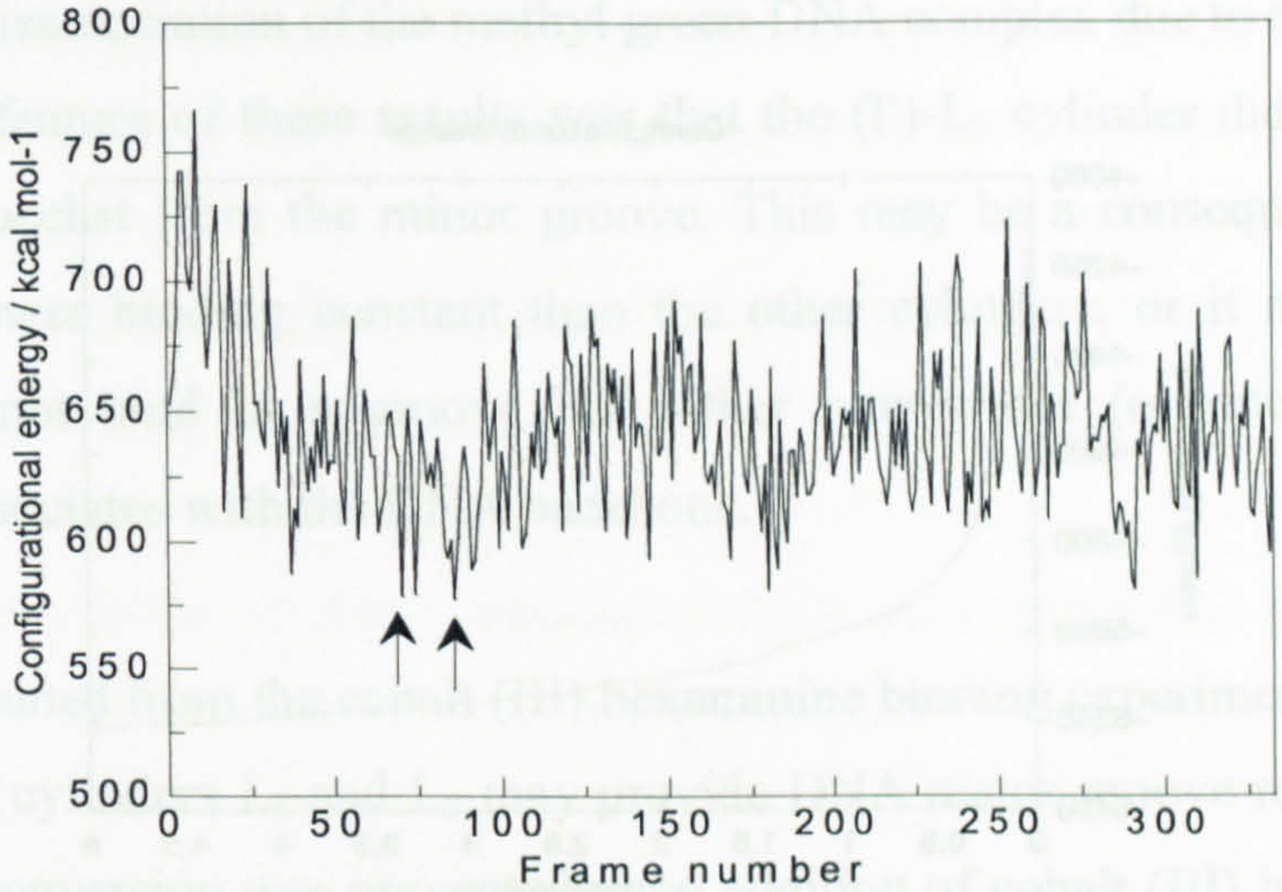

(a)

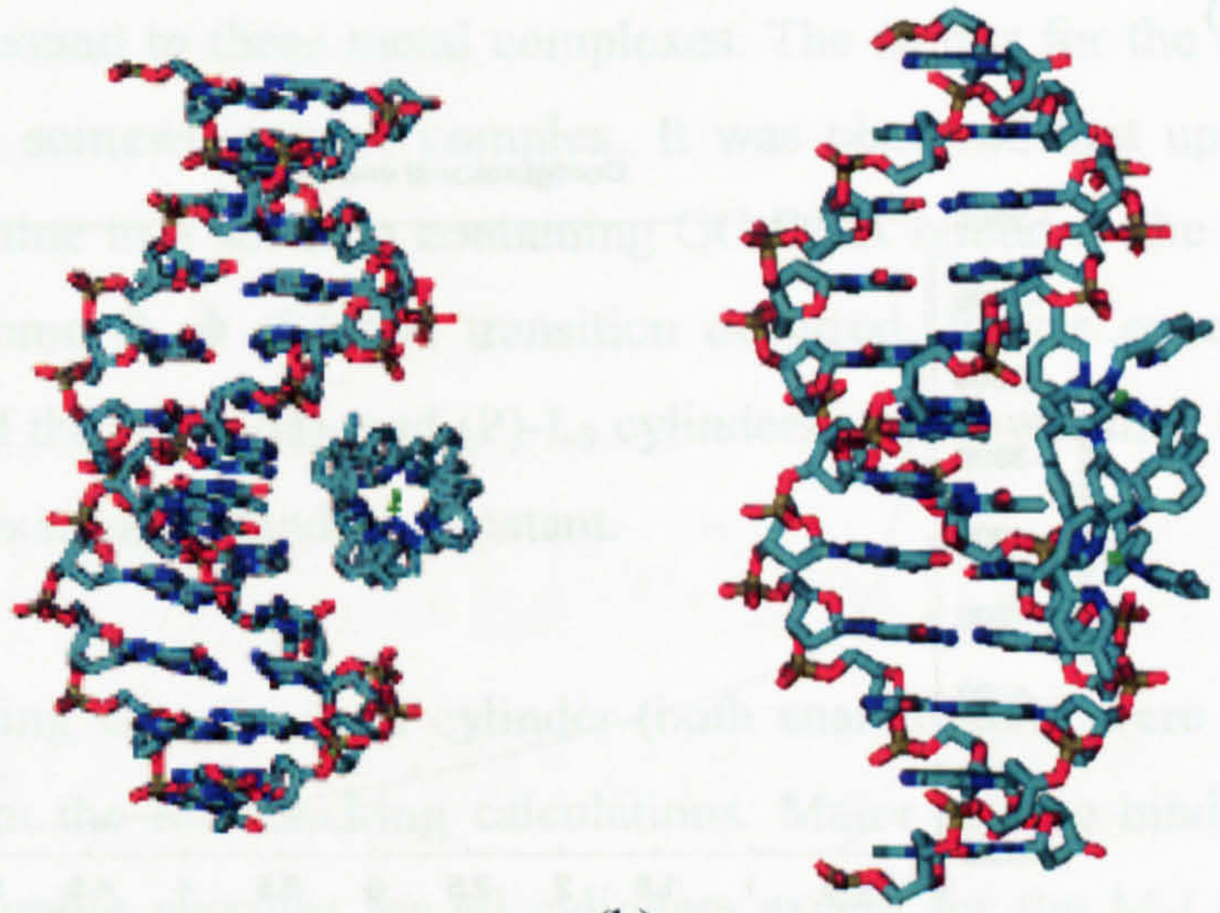



(b)

**Figure 4.15 Configurational energy plots of starting configurations (a) and (b) extracted from the high temperature docking calculations. Energy minimisations were ran at 210 K for 2.5 ps, using a dielectric constant of 1 $C^2 J^{-1} m^{-1}$.**

High temperature docking calculations of (P)-$L_5$ interacting with a B-DNA dodecamer show the most energetically favourable starting configurations situated in the major groove (see figure 4.16). Configuration (a) shows the cylinder situated in the major groove, denoted the bending mode and (b) shows the major groove denoted the backbone binding mode. Experimental data suggest that the (P)-$L_5$ enantiomers bind strongly to DNA inducing conformational changes in the DNA structure, as observed by a decrease in *LD* signal intensity in the DNA region of the spectrum and an increase in the signal intensity in the MLCT region of the spectrum.

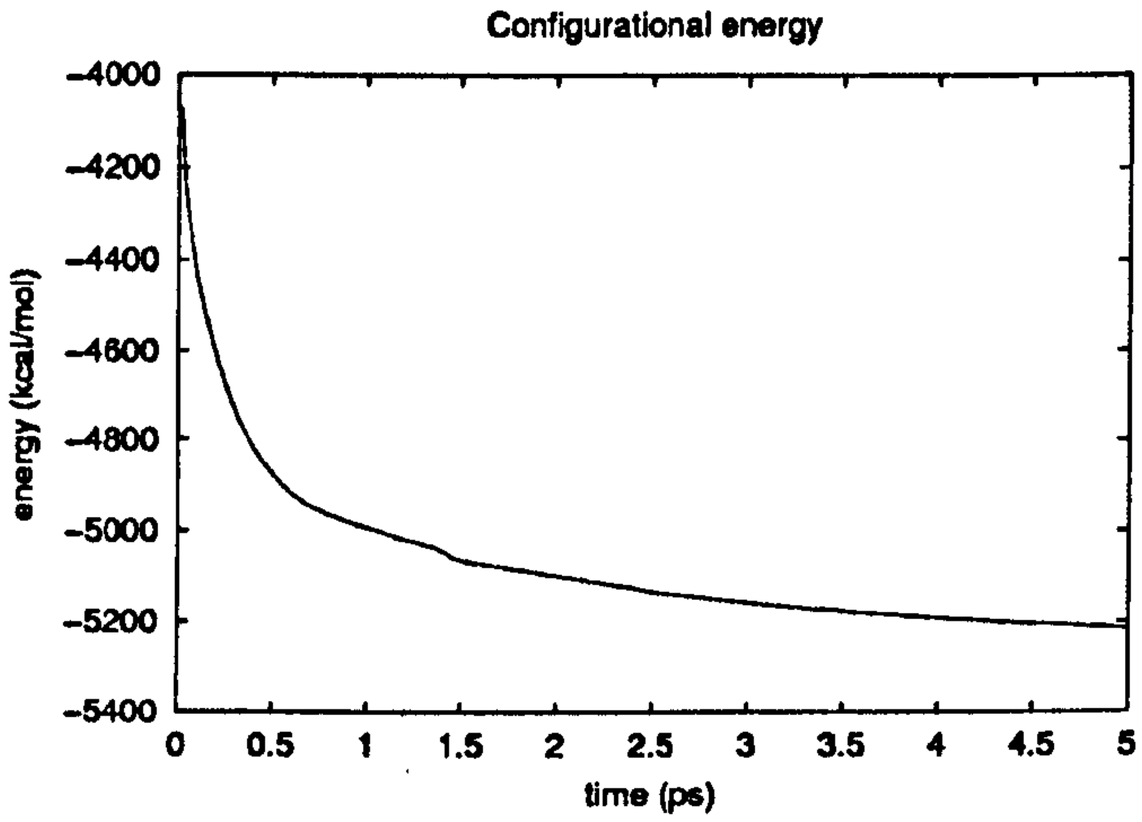


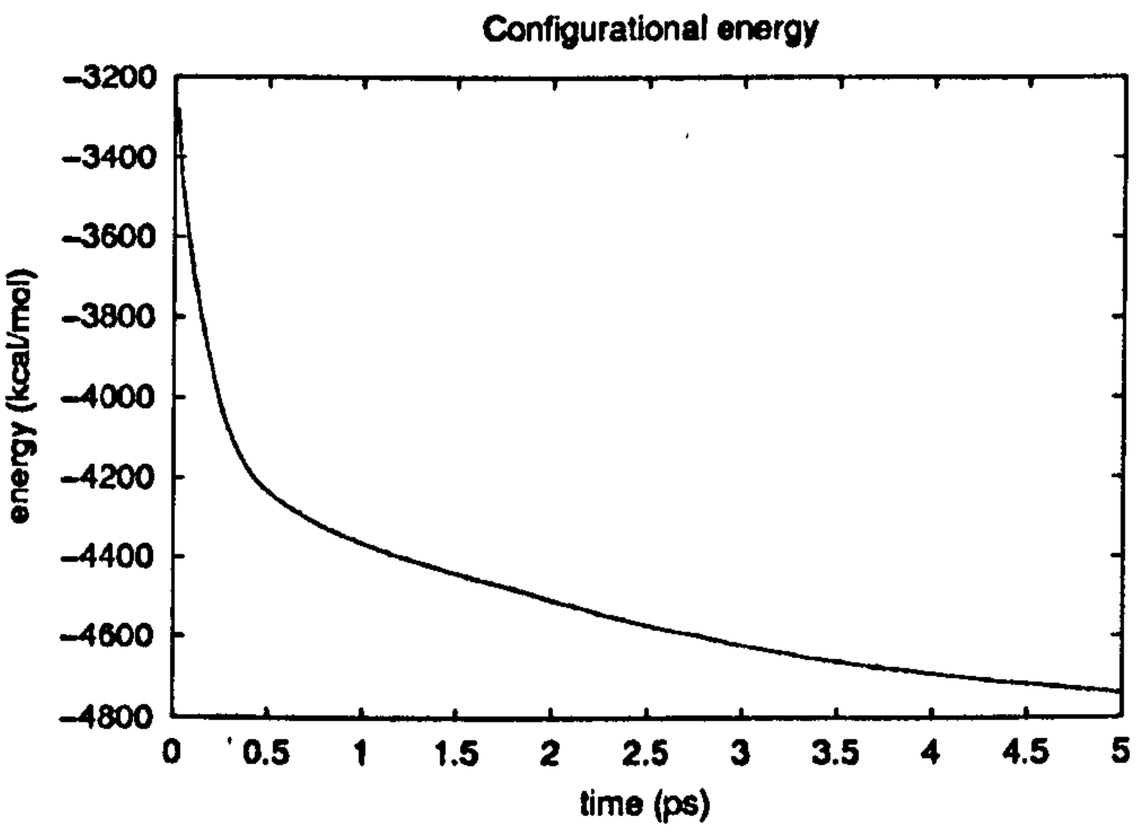
(a) (b)

**Figure 4.16 Configurational energy plot showing change in configurational energy during the high temperature docking calculation and two starting configurations of the (P)-$L_5$ cylinder extracted from the high temperature docking calculations. The cylinder lies in the major groove in both configurations ((a) shows the bending mode and (b) shows the major groove backbone binding mode). Interaction energies for these docking sites are given in table 4.1. H atoms have been omitted for clarity.**

An example of the energetic differences when the DNA flexibility is taken into account is given below. The two binding geometries from the (P)-$L_5$ docking calculations were minimised with mobile DNA and cylinder (see figure 4.17). The two conformations have similar configurational energies initially (see table 4.3) but upon minimisation these energies are substantially different. The major groove bending geometry is energetically more favourable than its major groove backbone binding counterpart.




(a)




(b)

**Figure 4.17 Plots of energy plots of starting configurations (a) and (b) extracted from the high temperature docking calculations. Energy minimisations were ran at 210 K for 5 ps, using a dielectric constant of 1 $C^2 J^{-1} m^{-1}$.**

## 4.4 Discussion

The results of the Hoechst 33258 competitive binding assays unambiguously demonstrated the substantially higher binding constant of the parent cylinder, $L_3$ and $L_5$ compared with Hoechst 33258. Unfortunately a negative aspect of this high binding constant was that all the cylinders were able to induce their characteristic conformational change in the DNA. This change involves a loss of the minor groove shape of canonical B-DNA and hence the Hoechst loses its favourable interactions with this groove and disassociates from the DNA. It was therefore not possible to ascertain if the disassociation of the DNA-Hoechst complex was a result of the cylinder competing for binding in the minor groove or whether this was a result of major groove binding induced conformational changes in the DNA. Competitive binding experiments to investigate the major groove binding with methyl green assays (data not shown) also

suffered from disassociation of the methyl green-DNA complex due to allosteric effects. An interesting feature of these results was that the (P)-$L_3$ cylinder did not completely displace the Hoechst from the minor groove. This may be a consequence of either a substantially lower binding constant than the other cylinders, or it may be that this cylinder does not bind in a groove but rather intercalates (or interacts by partial insertion) or associates with the DNA backbone.

The results obtained from the cobalt (III) hexammine binding experiments show that the enantiomers of cylinders $L_1$ and $L_5$ may provide DNA major groove recognition as the B → Z-DNA conversion was prevented upon addition of cobalt (III) hexammine when the GC-DNA was bound to these metal complexes. The results for the enantiomers of the $L_3$ cylinder are somewhat more complex. It was observed that upon addition of cobalt (III) hexammine to a solution containing GC-DNA bound to the (R)-, (M)- and (P)-$L_3$ cylinders, some B → Z-DNA transition occurred. It was concluded that the binding constants of the (R)-, (M)- and (P)-$L_3$ cylinders were lower than that of the GC-DNA cobalt (III) hexammine binding constant.

Two potential binding sites for each cylinder (both enantiomers) were extracted from each trajectory from the MD docking calculations. Major groove binding geometries had the most favourable energies for all cylinders except for the M-$L_5$, in which the most favourable configurational energy was exhibited by a minor groove binding geometry. At this stage of the research presented in this thesis, molecular modelling gives us a qualitative insight into the binding of the enantiomers of the methylated metal complexes and future studies should include the minimisation of more of the lowest energy configurations extracted from the high temperature docking calculations so that the energetically most favourable starting configuration for each enantiomer can be used as the starting configuration for fully solvated MD simulations.

## 4.5 References

1. Khalid, S. '*Molecular simulation studies of the interaction between DNA and a novel macromolecular ligand*' Ph.D. thesis, University of Warwick, **2004**

2. Waring, M. J., *J. Mol. Biol.* **1965**; 13, 269 – 282

3. Mahler, H. R., Kline, B. and Mehrota, B. D., *J. Mol. Biol.* **1964**; 9, 801 – 811

4. Pjura, P. E., Grzeskowlak, K. and Dickerson, D. E., *J. Mol. Biol.* **1987**; 197, 2, 257 – 271

5. Harshman, K. D. and Dervan, P. B., *Nucleic Acids Res.* **1985**; 13, 4825 – 4835

6. Parkinson, A., Hawken, M., Hall, M., Sanders, K. J. and Rodger, A., *Phys. Chem. Chem. Phys.* **2000**; 2, 5469 – 5478

7. Ali, N. and Ali, R., *Biochem. Mol. Bio. Int.* **1997**; 41, 1227 – 1235

8. Behe, M. J., Felsenfeld, G., Szu, S. C. and Charney, E., *Biopolymers*, **1985**; 24, 289 – 300

9. Loprete, D. M. and Hartman, K. A., *Biochemistry*, **1993**; 32, 4077 – 4082

10. Keller, P. B., Loprete, D. M. and Hartman, K. A., *J. Biomol. Struct. Dyn.*, **1988**; 5, 1221 – 1229

11. Malina, J. '*personal communication*' **2004**

12. Deng, H. and Bloomfield, V. A. *Biophys. J.*, **1999**; 77, 3, 1556 – 1561

13. Mac Kerell, A. D. Jnr., Wiórkiewicz-Kuczera, J. and Karplus, M., *J. Am. Chem. Soc.* **1995**; 117, 11946 – 11975

14. Smith, W., Forester, T. R., *J. Mol. Graphics*, **1996**; 14, 3, 136 – 141

15. Allen, M. P., Tildersley, D. J., '*Computer Simulation of Liquids*', **1987**; Oxford University Press, Oxford, UK

16. Feig, M. and Pettitt, B. M., *Biophys. J.* **1998**; 75, 134 – 149

17. Korolev, N., Lyubartsev, A. P., Laaksonen, A. and Nordenskiöld, L., *Biophys. J.* **2002**; 82, 2860 – 2875

18. Rodger, A., Sanders, K. J., Hannon, M. J., Meistermann, I., Parkinson, A., Vidler, D. S. and Haworth, I. S., *Chirality*, **2000**; 12, 221 – 236

19. Meistermann, I., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Khalid, S., Rodger, P. M., Peberdy, J. C., Isaac, C. J., Rodger, A. and Hannon, M. J. *Proc. Natl. Acad. Sci. USA.* **2002**; 99, 8, 5069 – 5074

20. Hannon, M. J., Moreno, V., Prieto, M. J., Molderheim, E., Sletten, E., Meistermann, I., Isaac, C. J. Sanders, K. J. and Rodger, A. *Angew. Chem. Int. Ed.* **2001**; 40, 5, 880 – 884

# CHAPTER 5 – OLIGONUCLEOTIDE-PROTEIN INTERACTIONS

## 5.1 Introduction

The results presented in this chapter summarise the investigation of a range of different guanosine and cytosine-rich DNA oligonucleotides (GRO/CRO) which adopt different quadruplex conformations in the presence of monovalent cations ($K^+$ or $Na^+$). The single stranded oligonucleotides have a propensity to form quadruplex structures that are stabilised by G-quartets. Experiments were completed to investigate the physical properties and biological activity of a group of 12 oligonucleotides whose structures have been characterised previously, at least to some extent (see § 5.5.1). Thermal denaturation and circular dichroism (*CD*) spectropolarimetry were utilised to investigate the stability, reversibility and ion dependence of G-quartet formation. The results of the initial GRO structural survey have been published.[1] It has been speculated that the biological activity of GRO's is related to their ability to bind to a specific cellular protein and this has been tentatively identified as nucleolin.[1] The second part of this research focuses on the structural characterisation of the RNA binding domain of nucleolin (denoted nucleolin-RBD). Nucleolin-RBD is a ~25 kDa, 237 amino acid residue fragment from human nucleolin. The nucleolin-RBD used in this work was obtained from P. J. Bates and J. O. Trent (University of Louisville, USA). UV/visible absorbance and circular dichroism spectroscopies were used to determine the secondary structure and thermodynamic stability of this protein (see § 5.5.2). The final part of this chapter is focussed on the investigation of the binding of nucleolin-RBD to a selection of the DNA oligonucleotides (screened in § 5.5.1) with the aim of characterising the structures of different G-quadruplexes upon binding to the protein (see § 5.5.3). Cytosine rich oligonucleotides (CRO's) were used as the control in this work as these were not expected to show any quadruplex secondary structure in the presence of buffer, NaCl or KCl. This work has all been undertaken as part of a bigger project taking place in the laboratory of P. J. Bates and J. O. Trent.[1 – 4]

## 5.2 Guanosine rich oligonucleotides

DNA structure is dependent upon intracellular conditions such as ion concentration and the presence of DNA-binding proteins that facilitate the interconversion between different forms or stabilise specific secondary structures.[5] DNA oligonucleotides are polyanionic species (see § 1.2 for discussion of DNA) that are likely to interact with many biomolecules within the cell and also in the extracellular membrane, in addition to sequence-specific interactions with nucleic acids.[3] G-quartets are four-stranded forms of DNA (see e.g. figure 5.1) which are believed to play an important role in living cells,

particularly in telomeres, the protein-DNA assemblies found at the ends of chromosomes[6] and in the inhibition of viral adsorption to the cell surface, giving rise to anti-HIV activity.[7] Oligonucleotide therapies have considerable potential in cancer, viral and cardiovascular disease and it is now clear that the biological effects of oligonucleotides are not solely due to the intended sequence-specific interactions with nucleic acids but may also be due to interactions with numerous cellular proteins owing to their polyanionic character or specific secondary structure.[1]

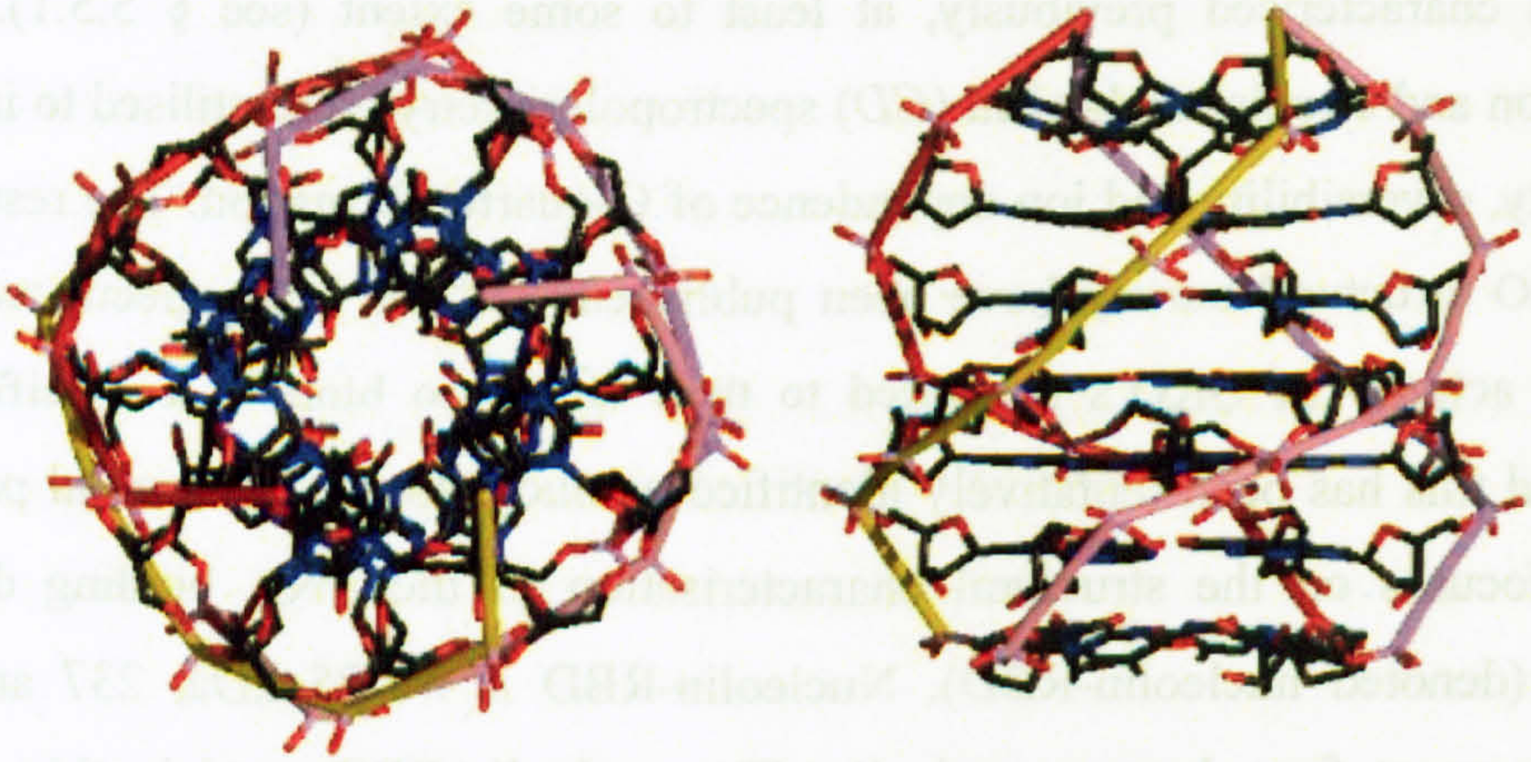


**Figure 5.1 Solution structure of a parallel-stranded G-quadruplex DNA, formed by the *Tetrahymena* telomeric sequence d(TTGGGGT). The structure was obtained by combining a quantitative analysis of the NMR data with molecular dynamics calculations.[8]**

There is currently considerable interest in the therapeutic potential of guanosine rich oligonucleotides and in their ability to fold in the presence of $K^+$ or $Na^+$ to form orderly structures stabilised by guanosine quartets or tetrads. Accumulating evidence has implicated tetraplexes in a number of potential biological roles and it has been shown that biologically important G-rich sequences are capable of forming G-quartets under physiological conditions.[1, 2, 7, 9, 10] It has long been recognised that G-rich nucleic acid sequences can adopt intermolecular and intramolecular quadruplex structures that are stabilised by the presence of G-quartets.[1] The structure of G-quartets (see e.g. figure 5.1) was first elucidated by Gellert *et al.* in 1962[11] using X-ray diffraction data and was proposed to consist of a square co-planar array of four guanine bases in which each base is both the donor and acceptor of two hydrogen bonds with its neighbours. It was proposed that the guanine tetrad arises from the association of four guanines into a cyclic Hoogsten hydrogen bonding arrangement that involves N1, N7, O6 and N2 of each guanine base (see figure 5.2).[11, 12] Monovalent cations such as $K^+$ and $Na^+$ have been shown to stabilise the G-quadruplex structures, presumably by coordinating with eight carbonyl oxygen atoms present between stacked tetrads.[13]

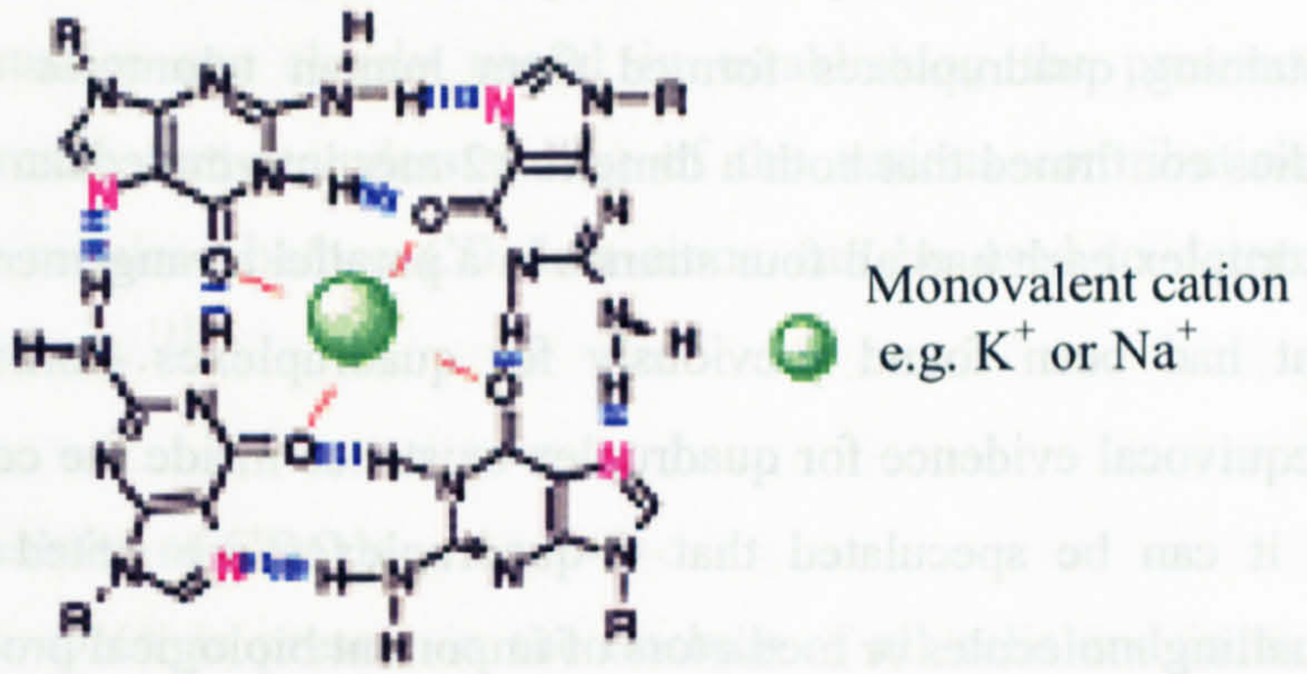



**Figure 5.2 Schematic diagram showing hydrogen bonding in a guanine tetrad motif.[1]**

One of the most intriguing features of G-quadruplexes is their extensive polymorphism which arises from strand stoichiometry, strand polarity and glycosidic torsion angle variation.[5] A variety of structures are known to exist *in vitro* and they can be classified in terms of their molecularity and strand orientation.[14] GRO's can form a variety of possible quadruplex structures involving one, two or four oligonucleotide molecules depending on both thermodynamic and kinetic considerations (see figure 5.3). Monomer and dimer quadruplex structures have been classified according to the position of their loop region, the chair conformer having a lateral loop, and basket having a diagonal loop. The relative strand orientation (5' to 3' polarity) of the four strands of the quadruplex may be parallel, anti-parallel or mixed. The label || or anti-|| is used where there is a choice. Many different polarity configurations have been determined by X-ray crystallography or NMR spectroscopy.[15 – 35] G-quadruplexes are characterised according to the co-ordination of cations, usually $K^+$, as under physiological conditions $K^+$ is best suited to fill the cavity between G-tetrads and different structures are observed under different ionic conditions.[5]

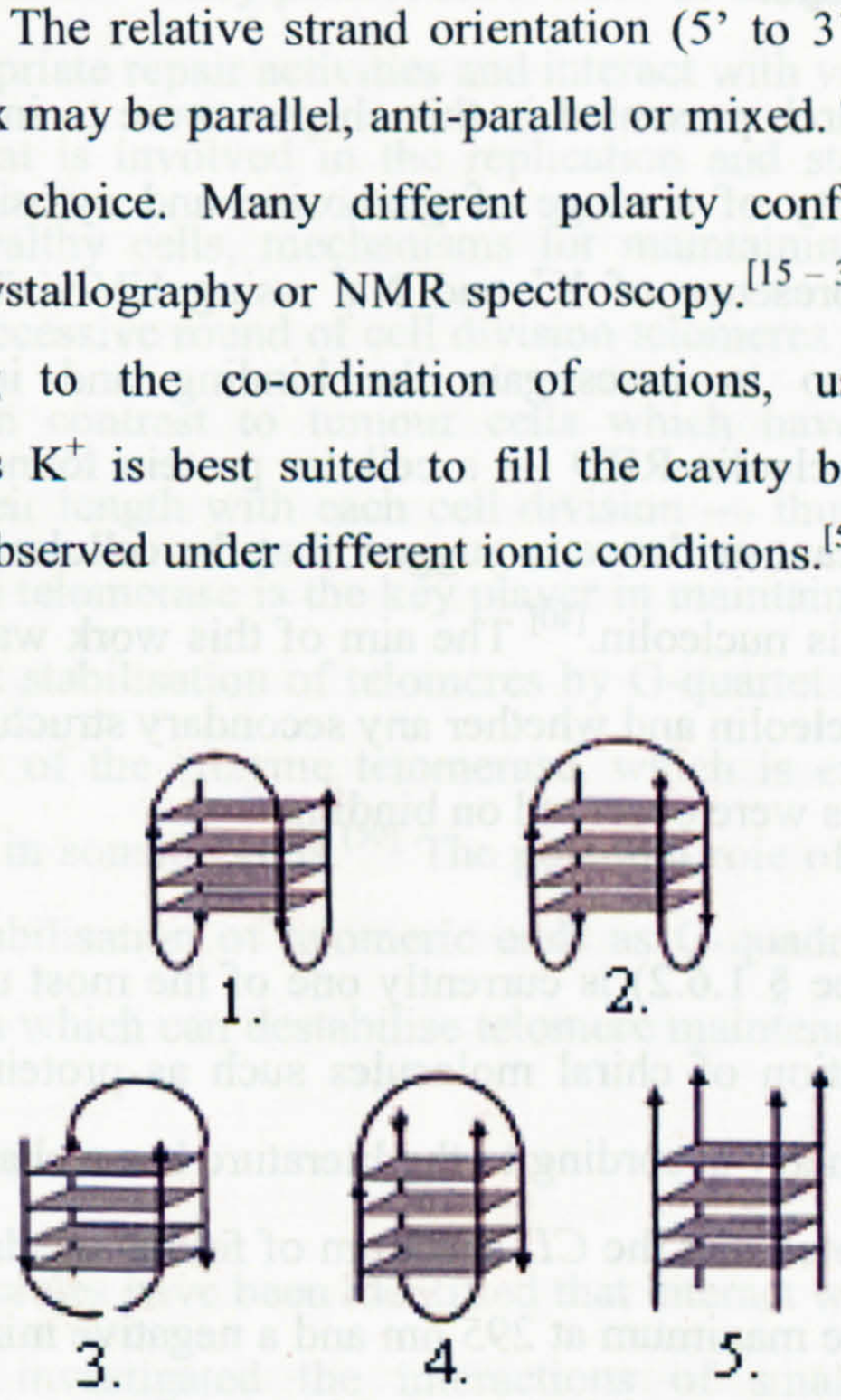



**Figure 5.3 Schematic presentation of molecularity and loop orientation of quadruplexes. 1. monomer chair, 2. monomer basket, 3. dimer chair, 4. dimer basket and 5. tetramer.[1]**

In 2002, Neidle *et al.* reported the crystal structures of both intra- and intermolecular $K^+$ containing quadruplexes formed from human telomeres (see figure 5.4).[36] NMR studies confirmed that both a dimeric 12-mer intermolecular and 22-mer intramolecular quadruplex each had all four strands in a parallel arrangement, which was in contrast to what had been found previously for quadruplexes from *Oxytricha nova*.[32, 37, 38] Unequivocal evidence for quadruplex existence inside the cell has yet to be confirmed, but it can be speculated that G-quadruplexes are suited to the role of biological signalling molecules or mediators of important biological processes.

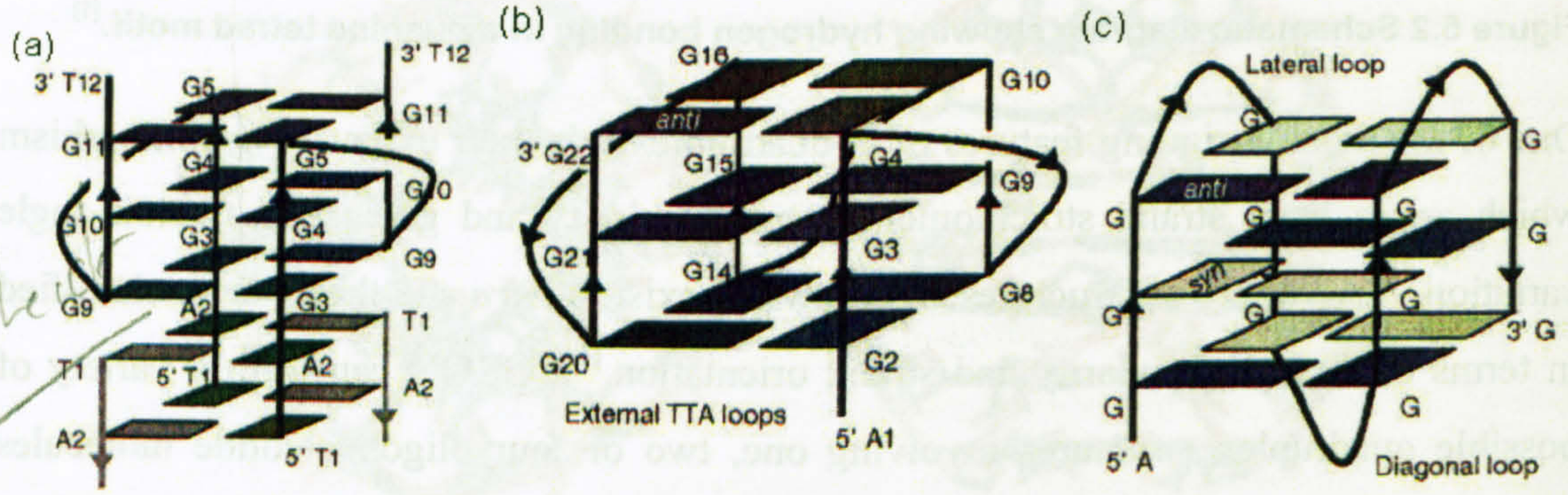



**Figure 5.4 Schematic diagram of human telomeric quadruplex folding topologies. (a) and (b) Intramolecular and dimeric intermolecular G-quadruplexes, all with *anti* glycosidic torsion angles and extended external loops, crystallised with $K^+$ ions, (c) fold determined from the NMR intramolecular G-quadruplex solution structure[39] in $Na^+$ ions only, with lateral and diagonal loops.[36]**

The aims of the research presented in this chapter were to investigate the kinetic and thermodynamic stability of a range of guanosine and cytosine-rich oligonucleotides (GRO/CRO) in the presence of $K^+$ and $Na^+$ using UV/visible absorbance and *CD* spectroscopy and also to investigate the binding and interactions of different oligonucleotides to nucleolin-RBD — a cellular protein found in the nucleolus of all cells. There is significant evidence to suggest that the cellular target of the quadruplex DNA denoted GRO1 is nucleolin.[40] The aim of this work was to investigate whether the GRO1 binds to nucleolin and whether any secondary structural changes of either the protein or nucleic acids were observed on binding.

Circular dichroism (see § 1.6.2) is currently one of the most useful techniques for the structural characterisation of chiral molecules such as proteins and nucleic acids.[41] Each quadruplex geometry according to the literature has a characteristic *CD* spectrum. It is now widely accepted that the *CD* spectrum of folded quadruplexes (monomers and dimers) have a positive maximum at 295 nm and a negative minimum at 265 nm, while parallel quadruplexes (tetramers) have a positive maximum at 264 nm and a negative minimum at 240 nm.[42 – 44] One result from our work reported in § 5.5.1 is that no

characteristic *CD* spectrum can be unequivocally linked to anti-proliferative activity in GRO's, and while *CD* studies are clearly useful in establishing the presence of quadruplex structures, a much better understanding of the various contributions to quadruplex *CD* spectra is required before *CD* data alone can be used to determine quadruplex structure definitively.[1]

### 5.2.1 Anti-proliferative activity of GRO's

As discussed in chapter 1, DNA is present in the nucleus of all cells and is wound around in tight coils to form chromosomes. DNA replication occurs when chromosomes duplicate themselves. The first step of the replication process involves the unwinding of the helix to form two separate strands and, due to the obligatory pairing of A-T and G-C, each parent strand becomes a template for copying a new DNA helix. During the DNA replication process, one strand is incompletely replicated at one end and to prevent loss or mutation of essential genes, the end of each chromosome has a strand of DNA called a telomere. Telomeres are the repetitive DNA sequences (TTAGGG) found at the ends of chromosomes and have a sequence particularly rich in guanine bases. In human somatic cells they are typically 5 – 8 kilobases long, with a single stranded 3' overhang of 100 – 200 bases.[9] They protect the 3' ends of the parent DNA strand from degradation and inappropriate repair activities and interact with various proteins to form a telomeric complex that is involved in the replication and stability of linear DNA molecules.[45 – 48] In healthy cells, mechanisms for maintaining telomere length are absent and with each successive round of cell division telomeres progressively decrease in length.[49] This is in contrast to tumour cells which have markedly shortened telomeres that retain their length with each cell division — thus making tumour cells 'immortal'. The enzyme telomerase is the key player in maintaining telomere length. It has been speculated that stabilisation of telomeres by G-quartet interactive compounds may inhibit the activity of the enzyme telomerase, which is expressed in almost all tumour cells but absent in somatic cells.[50] The potential role of quadruplex structures *in vivo* may involve stabilisation of telomeric ends as G-quadruplex structures using specific small molecules which can destabilise telomere maintenance in tumour cells.[41 – 54]

A number of small molecules have been identified that interact with G-quartets. Several research groups have investigated the interactions of small molecules such as acridines,[54] anthraquinones[55] and cationic porphyrins[56, 57] with G-quartets using

molecular modelling, *CD* and NMR studies. NMR studies have confirmed that 2,6-diamidoanthraquinone BSU1051, interacts with and stabilises the G-quartet, presumably binding via intercalation and inhibits telomerase activity.[58]

The initial telomere anticancer strategy involved designing DNA oligonucleotides complementary to telomeres with the aim of forcing them into duplex not quadruplex structures. However, it appears the biological activity of some such designed oligonucleotides has nothing to do with their antisense activity. Rather the non-antisense anti-proliferative activity of GRO's is associated with their ability to form stable G-quartet containing structures and their subsequent ability to bind to a specific cellular protein or proteins.[1] This then disrupts cellular mechanisms that ensure cellular immortality. Bates *et al.* have recently showed that a 29-mer 3'-modified phosphodiester oligonucleotide (GRO29A) exerts a potent growth inhibitory effect against several cancer cell lines *in vitro* and that the activity of the GRO is related to its binding to a cellular protein. It has been shown that GRO29A specifically arrests the cells in the S-phase of the cell cycle and inhibits DNA synthesis.[3, 4] Nucleolin was identified as the GRO-binding protein and a strong correlation was observed between binding to this protein and anti-proliferative activity.[2]

### 5.2.2 Summary of GRO structure

It has been observed that guanosine-rich oligonucleotides can form a variety of possible quadruplex structures, depending on both thermodynamic and kinetic considerations. The structures formed can be influenced by oligonucleotide base sequence and concentration, as well as the conditions (temperature and buffer) used for annealing. The presence of monovalent cations such as $K^+$ or $Na^+$ has a particularly significant effect. The research presented in this chapter attempts to characterise GRO quadruplex structure using *CD* under different experimental conditions. Initial experiments (§ 5.4.1) involved the structural characterisation of 12 oligonucleotides using *CD* to investigate the structures formed in the presence of different monovalent cations. The structure of each oligonucleotide investigated has previously been characterised and a summary of the structures is presented below (see table 5.1 for summary of oligonucleotide sequence and structural properties). In many of the cases where both NMR and XRC structures are available, there is disagreement between the two techniques regarding the loop structure (basket/chair) or conformation (syn/anti) of the quartet guanosines, although the molecularities of the structures formed are in accord. It is possible that

both structures are correct and that discrepancies reflect the different experimental conditions used for quadruplex annealing (see § 5.4.1.1).[1]

Oligonucleotide KS-A, d(TGGGGT), is derived from the *Oxytricha nova* telomeric sequence and NMR and X-ray crystallographic techniques have confirmed that it forms a parallel-stranded tetrameric quadruplex in the presence of sodium ions (see figure 5.5 (a)). A right-handed helix is formed that is stabilised by H-bonding tetrads of co-planar guanine bases. Sodium ions are found, between, and at defined points, within the tetrad planes and co-ordinated with the guanine O6 groups. [15 - 17] In 1994, Aboul-ela *et al.* using NMR confirmed that the guanine-tetrad containing structures are stabilised by the presence of $K^+$.[16]

KS-B, d(GGTTGGTGTGGTTGG) known as the thrombin-binding aptamer, forms a monomeric chair structure that is strongly stabilised by potassium ions.[18 - 20] KS-B has a well defined solution structure, eight of the conserved guanine residues form a central core composed of two guanine quartets (see figure 5.5 (b)). The guanine residues are 5'-syn-anti-3' along each of the four edges of the guanine core.[18] The elucidation of the structure of KS-B should now allow for the rational design approach, increasing the stability and efficacy of this potential anti-thrombotic drug.[19]

KS-C, d($G_4T_4G_4$) is derived from the *Oxytricha nova* telomeric sequence and forms a dimeric hairpin quadruplex with four G-quartets and thymine loops at the opposite ends of the G-quartets in the presence of $K^+$, $Na^+$ or $NH_4^+$ (see figure 5.5 (c)).[21] X-ray crystallographic data suggests the formation of a chair conformation,[22] in contrast to NMR data which suggests the formation of a basket conformation.[23] Recent studies using X-ray crystallography indicates that a similar basket conformation can also be formed in the crystalline state.[24] NMR solution studies of KS-C in the presence of $Na^+$ and $K^+$ ions shows that both possess the same folding topology and similar stacked G-quartet geometry but display a notable difference in the conformation of the diagonal $T_4$ connecting loops.[21]

The KS-D, d(GCATGCT) sequence is a heptamer that forms a four-stranded structure that does not involve G-quartets (see figure 5.5 (d)). Rather it forms a novel dimer structure that is stabilised by the non-classical base-pairing interactions of two folded molecules.[25] The asymmetric unit comprises a single strand that folds back on itself to

form a loop structure. Two strands associate in an anti parallel manner related by a two fold axis of symmetry. The base pairs in oligonucleotide KS-D are not coplanar with the partners in the quartets — instead they are tilted by ~30°.[25]

KS-E, d(GCGGTTTGCGG) represents the fragile X gene repeat sequence and has been shown to form a dimeric chair structure, in the presence of sodium ions, containing two G-quartets and two G·C·G·C tetrads (see figure 5.5 (e)). This novel NMR-molecular dynamics based solution structure contains internal G·C·G·C tetrads sandwiched between terminal G·G·G·G tetrads. Adjacent strands are anti-parallel to each other around the symmetric G-quadruplex which contains two distinct narrow and two symmetric wide grooves.[26]

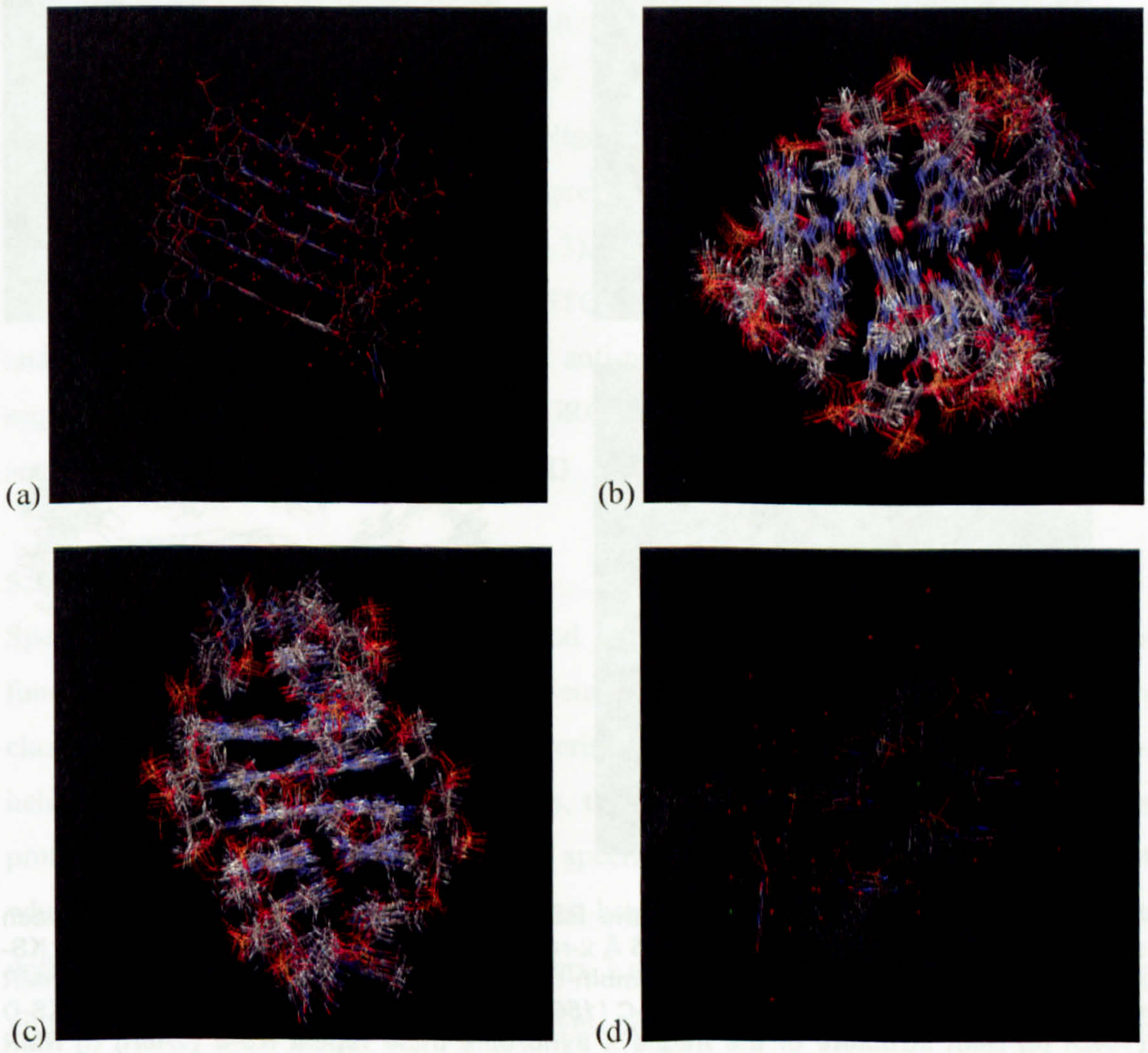
Oligonucleotide KS-F, d(TAGG) is a short analogue of the *Bombyx mori* telomere repeat (se figure 5.5 (f)). Analysis by a combined NMR-molecular modelling approach has suggested the formation of an unusual structure containing two G-quartets, in which each strand has a parallel and an anti-parallel neighbour.[27] A two-fold symmetric four-stranded G-quadruplex containing both syn and anti guanines with both parallel and anti-parallel alignment of strands is formed in the presence of $Na^+$ ions.[27]

The KS-G, d($G_3T_4G_3$) sequence is another *Oxytricha nova* telomere sequence that forms a dimeric quadruplex structure in the presence of $K^+$ or $Na^+$, but is stabilised preferentially by potassium (see figure 5.5 (g)).[28 – 30] NMR studies show that this quadruplex is in the basket form, but that the glycosidic conformations are different from the similar KS-C sequence.

Oligonucleotides KS-H, d($G_4T_4G_4T_4G_4T_4G_4$) and KS-I, d($T_2G_4T_2G_4T_2G_4T_2G_4$) are derived from *Oxytricha nova* and *Tetrahymena thermophila* telomere sequences respectively (see figures 5.5 (h) and (i) respectively). They both form folded monomeric quadruplexes but have considerably different structures.[31 – 33] The KS-H monomer is stabilised by four stacked G-quartets with two lateral T4 loops and a diagonal central loop and one anti-parallel strand (basket monomer), while due to the shorter sequence between the G4 repeats, KS-I can only form three stacked G-quartets. KS-I forms a lateral GTTG loop, a central lateral TTG loop and an unusual TT loop that spans from the top of the quadruplex to the bottom such that the fourth strand polarity is parallel to the adjacent first and third strands.[33]

T30695, d($G_3TG_3TG_3TG_3T$) is a potent anti-HIV oligonucleotide that has been analysed by a combined NMR-molecular modelling approach and shown to form a potassium-stabilised chair monomer containing three stacked G-quartets (see figure 5.5 (j)).[34] After binding three $K^+$ ions T30695 is composed of a central most pair of G-quartets and the loop bases are aligned so that they lie roughly parallel to the G-quartet bases.

The structures of GRO20A and GRO23A have been determined by Marathias and Bolton using NMR analysis.[35] These authors showed that a structure similar to GRO20A (some of the loop thymines were substituted by uracil) formed a well defined monomer basket quadruplex in the presence of sodium ions and did not change significantly upon addition of potassium. In contrast, the spectrum of GRO23A was affected by the addition of potassium and it was concluded that GRO23A formed a complex mixture of chair and basket conformations.[35]



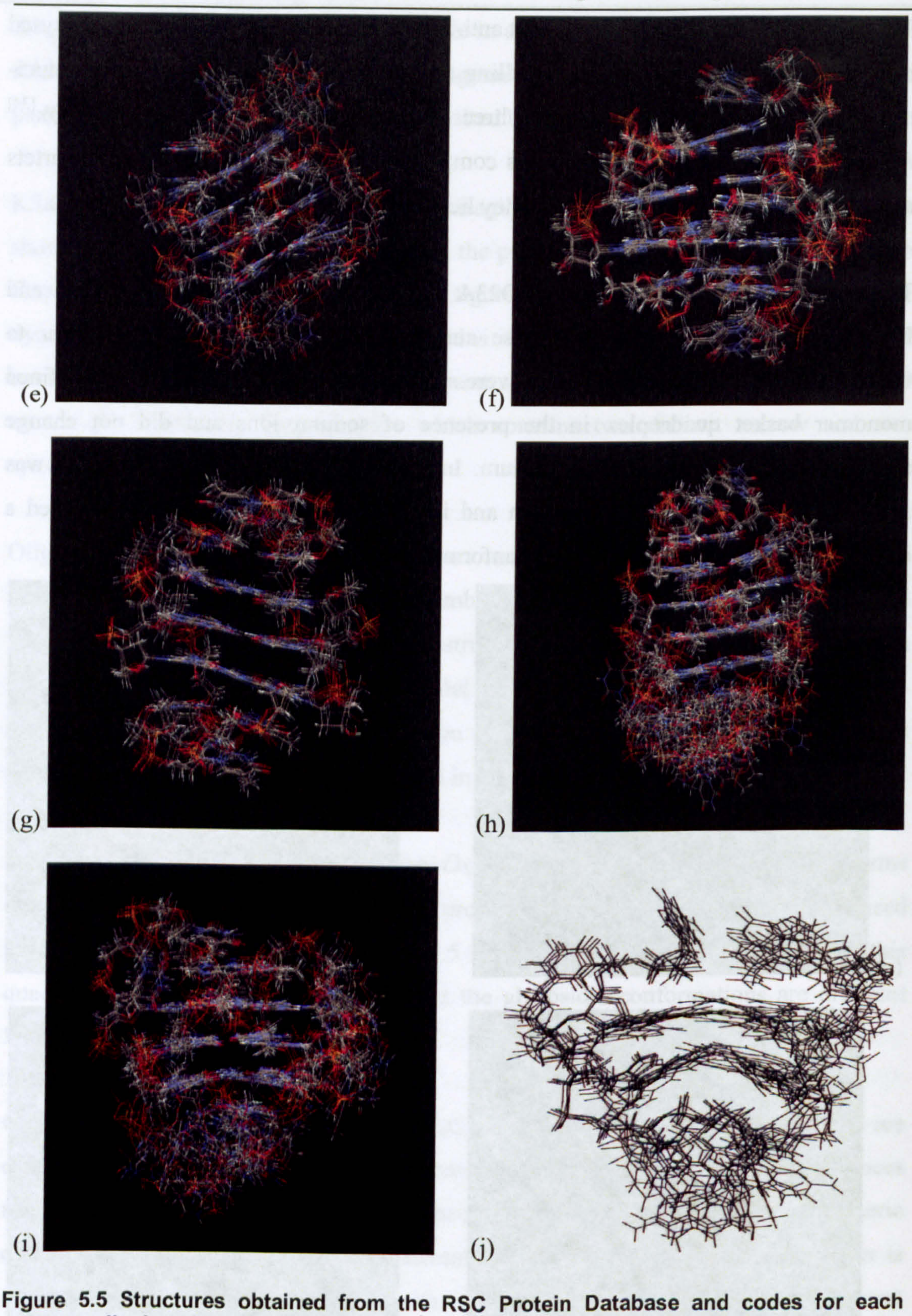



**Figure 5.5 Structures obtained from the RSC Protein Database and codes for each structure displayed in brackets. (a) 0.95 Å x-ray crystal structure of parallel stranded KS-A (*352D*) (b) NMR structure of the thrombin-binding DNA aptamer KS-B (*148D*) (c) NMR structure of the dimeric quadruplex KS-C (*156D*) (d) 1.8 Å x-ray crystal structure of KS-D (*184D*) (e) NMR structure of the fragile X syndrome triple repeat KS-E (*1A6H*) (f) NMR structure of KS-F, a G-quadruplex capped by base triads (*1AFF*) (g) NMR structure of KS-G, a dimeric guanine quadruplex in $Na^+$ (*1FQP*) (h) NMR structure of the telomeric repeat KS-H (*201D*) (i) NMR structure of G-tetraplex telomeric repeat KS-I (*186D*) and (j) NMR structure of the unimolecular tetrad structure of T30695 stabilised by a pair of G-quartets (as in reference 34).**

| Oligo | Sequence | Description | Refs. |
|---|---|---|---|
| KS-A | d(TGGGGT) | Tetramer, telomere | 15 – 17 |
| KS-B | d(GGTTGGTGTGGTTGG) | Monomer, thrombin aptamer | 18 – 20 |
| KS-C | d(GGGGTTTTGGGG) | Dimer, telomere | 21 – 24 |
| KS-D | d(GCATGCT) | Dimer, non-GQ | 25 |
| KS-E | d(GCGGTTTGCGG) | Dimer, fragile-X | 26 |
| KS-F | d(TAGG) | Tetramer, telomere | 27 |
| KS-G | d(GGGTTTTGGG) | Dimer, telomere | 28 – 30 |
| KS-H | d(GGGGTTTTGGGGTTTTGGGGTTTTGGGG) | Monomer, telomere | 31 – 33 |
| KS-I | d(TTGGGGTTGGGGTTGGGGTTGGGG) | Monomer, telomere | 31 – 33 |
| T30695 | d(GGGTGGGTGGGTGGGT) | Monomer, anti-HIV | 34 |
| GRO20A | d(GGTTTTGGTTTTGGTTTTGG) | Monomer basket | 35 |
| GRO23A | d(GGGGTTGGGGTGTGGGGTTGGGG) | Mixed basket/chair | 35 |
| GRO1 | d(GGTGGTGGTGGTTGTGGTGGTGGT) | Experimental oligo | |
| CRO1 | d(CCTCCTCCTCCTTCTCCTCCTCCT) | Control | |

**Table 5.1 Table showing oligonucleotide sequence and structural properties.**

Subsequent to the work reported in reference 1, we have also worked with GRO1, a 24-mer guanosine-rich oligonucleotide which was shown to adopt different quadruplex structures in the presence of $K^+$ or $Na^+$. The second oligonucleotide, CRO1, the corresponding 24-mer cytosine-rich oligonucleotide was used as the control. Differences in binding interactions in the presence of different monovalent cations were expected and indeed observed (see § 5.4.3). A 26-mer guanosine-rich oligonucleotide (5'-GGTGGTGGTGGTTGTGGTGGTGGTGG-3') is currently in Phase I clinical trials and has been shown to exhibit significant anti-proliferative activity, which is similar in sequence to GRO1.[40] It was hoped that GRO1 would show increased anti-proliferative activity upon binding to the nucleolin-RBD.

## 5.3 Protein-DNA interactions

Specific interactions between DNA and DNA-binding proteins (DBP's) are of fundamental importance in biology. Proteins bind to double stranded DNA and are classified based on their structural characteristics into four main groups: the helix–turn–helix proteins, the homeodomain proteins, the leucine zipper proteins and zinc finger proteins.[59] Each protein binds sequence specifically to the major groove of ds-DNA where the size, shape, number of hydrogen bonds, donors and acceptors differs more, with base sequence as discussed in chapters 3 and 4. In addition, many proteins target unusual DNA structures, of specific interest to this work are DNA quadruplexes as described above.

### 5.3.1 Protein secondary structure

The α-helix (see figure 5.6 (a)) is a common protein secondary structure motif. It was described by Pauling and Corey as the simplest arrangement that the polypeptide chain can adopt with its rigid peptide bonds forming a helical structure.[60] The polypeptide backbone is tightly wound around an imaginary axis drawn longitudinally through the middle of the helix and the R groups of the amino acid residues protrude outward from the helical backbone. Every peptide bond participates in hydrogen bonding between its electronegative carbonyl oxygen atom and the hydrogen atom attached to the electronegative nitrogen atom of a peptide linkage n + 4 residues away. The combined hydrogen bonds give the entire helical structure considerable stability.

The β-sheet is another type of repetitive protein secondary structure (see figure 5.6 (b)). As predicted by Pauling and Corey, the structure is built from a combination of several regions of the polypeptide chain in contrast to the α-helix, which is built up from one continuous sequence.[61] The backbone of the polypeptide chain is extended into a zigzag structure. When the zigzag polypeptide chains are arranged side by side a structure resembling a series of pleats is formed. Hydrogen bonding occurs between adjacent segments of the polypeptide chain, which can be either parallel or anti-parallel in conformation (having the same or opposite amino-to-carbonyl orientations respectively). Each of the two forms has a distinctive pattern of hydrogen bonding. All possible main chain hydrogen bonds are formed within the two conformations apart from the two flanking strands of the β-sheet that only have one neighbouring β strand.

β-turns (see figure 5.6 (c)) are also common secondary structures present in proteins. They are strictly the connecting elements that link successive runs of β-conformations.[62] A β-turn usually connects the ends of two adjacent segments of an anti-parallel β-sheet. The structure consists of a 180° turn involving four amino acid residues, with the carbonyl oxygen of the first amino acid residue forming a hydrogen bond with the amino-group hydrogen of the fourth. In practise the label β-turn is used to refer to all turns in proteins linking any structural motifs.

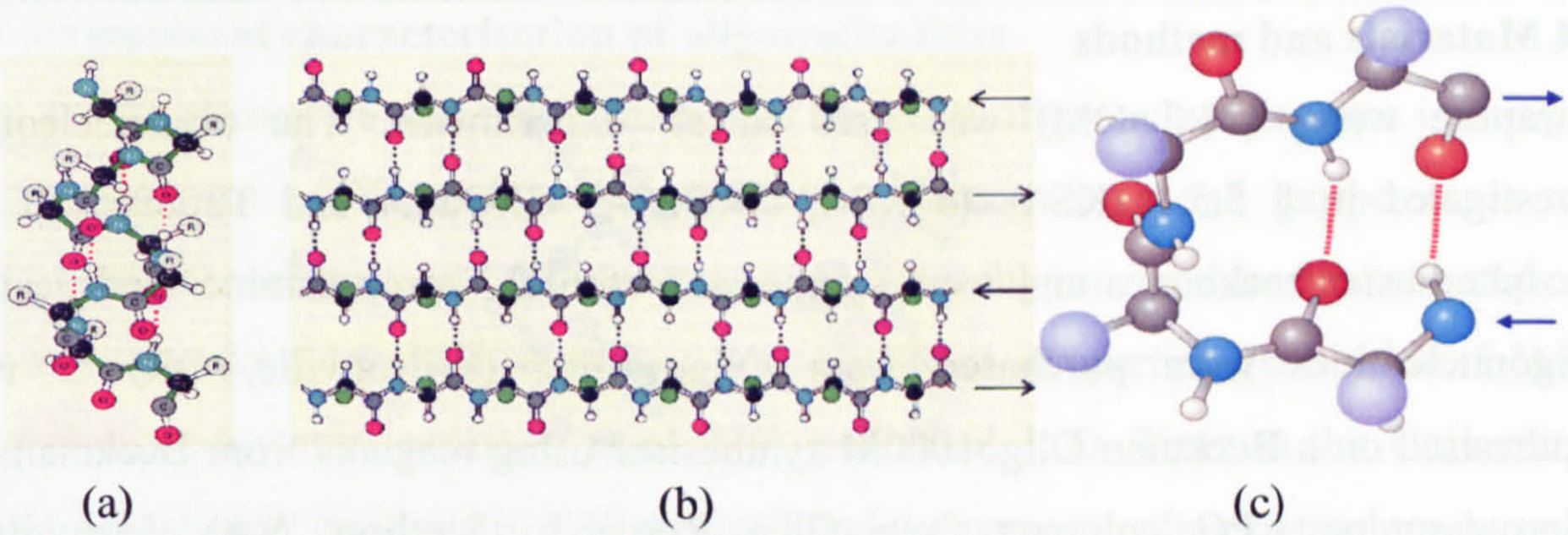



**Figure 5.6 (a) α-helix, (b) anti-|| β-sheet and (c) β-turn protein secondary structures.[48]**

### 5.3.2 Nucleolin

Nucleolin is a nucleolar phosphoprotein of exponentially growing eukaryotic cells and is present in abundance in the nucleolus.[63] Biophysical and biochemical studies have shown that nucleolin has a unique multi-domain structure that is tripartite, consisting of a histone like N–terminus, a central domain containing RNA recognition motifs (see figure 5.7) and a glycine and arginine-rich C–terminus; and is highly conserved during evolution.[64, 65] Each domain performs a specific function by interacting with DNA, RNA or other proteins.[66 – 68] Levels of nucleolin in the body are known to relate to the rate of cellular proliferation,[69, 70] being elevated in rapidly proliferating cells, such as malignant cells, and lower in more slowly dividing cells. For this reason nucleolin has been considered to be an attractive therapeutic target for the treatment of malignant diseases.[3]

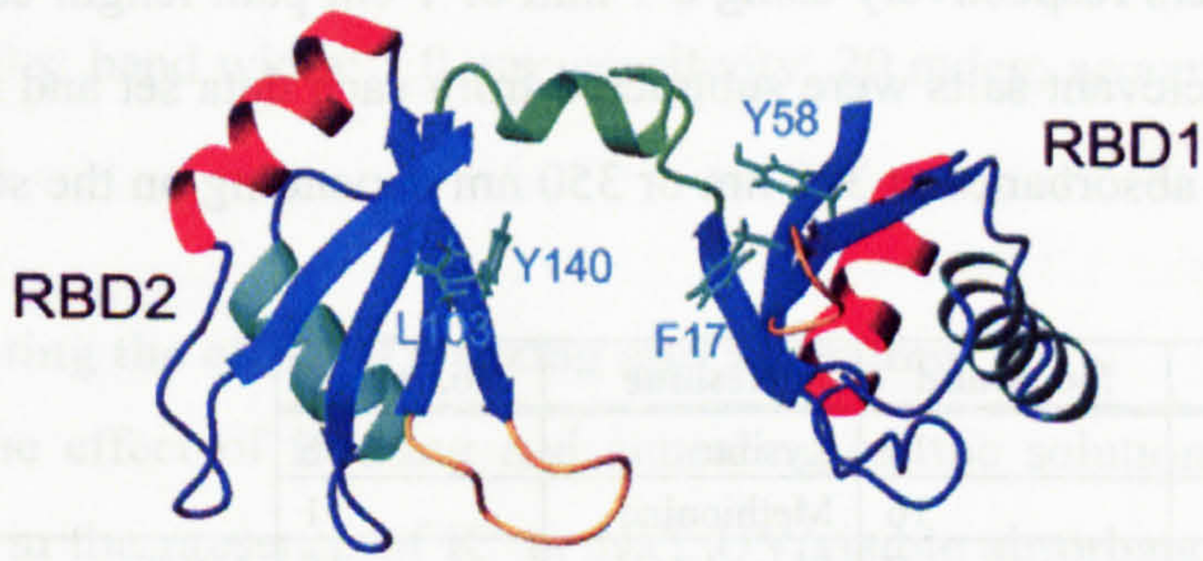



**Figure 5.7 NMR structure of the RNA-binding domains (denoted RBD1 and RBD2) of nucleolin extracted from Mesocricetus auratus.[71]**

The most important feature of nucleolin is its multi-functionality and through direct or in-direct investigation it has been implicated in the following activities: ribosome biogenesis,[71] inducing chromatin decondensation,[72] and as a potential therapeutic agent, interacting with G-quartet forming oligonucleotides, to increase anti-proliferative activity.[1, 3]

## 5.4 Materials and methods

Ultrapure water (18.2 ΩM) was used in all experiments. The oligonucleotides investigated in § 5.5.1 (KS-A to KS-I, GRO20A, GRO23A and T30695) all had phosphodiester backbones and were synthesised with a 3'-propylamine modification. Oligonucleotides were purchased from Oligos Etc. (Wilsonville, OR) or were synthesised on a Beckman Oligo1000M synthesiser using reagents from Beckman and 3'-proplyamine CPG columns from Glen Research (Sterling, VA). Lyophilized oligodeoxynucleotide samples (CRO1 and GRO1) were purchased from Integrated DNA Technologies Inc. Oligonucleotide concentrations were determined spectrophotometrically using the molar extinction coefficients: GRO1 $\varepsilon_{260}$ = 230,000 $cm^{-1}$ $mol^{-1}$ $dm^3$ and CRO1 $\varepsilon_{260}$ = 180,800 $cm^{-1}$ $mol^{-1}$ $dm^3$.[40] Potassium and sodium phosphate buffers (pH 7.0) were prepared. The nucleolin-RBD (two RNA binding domains with a RGG tail) was obtained from P. J. Bates and John O. Trent at the University of Louisville, Kentucky, and used without further purification (see table 5.2 for structural properties). The nucleolin-RBD was dissolved in the corresponding KCl or NaCl phosphate buffer (depending on the oligonucleotide solution) on the day of the experiment and the concentration determined spectrophotometrically using the molar extinction coefficient: $\varepsilon_{277}$ = 6,400 $cm^{-1}$ $mol^{-1}$ $dm^3$.[40] Stock solutions of KCl (0.5 M, 3.728 g in 100 mL) and NaCl (0.5 M, 2.925 g in 100 mL) were prepared. UV/visible absorbance and *CD* spectra were recorded on Jasco V-550 and J-715 spectrophotometers respectively using a 1 mm or 1 cm path length cuvette. Spectra for buffers and the relevant salts were subtracted from each data set and the spectra zeroed in a region of no absorbance at 320 nm or 350 nm depending on the scan range used for each experiment.

| Residue | No. found | Residue | No. found | | |
|---|---|---|---|---|---|
| Alanine | 13 | Lysine | 26 | | |
| Arginine | 16 | Methionine | 1 | | |
| Asparagine | 7 | Phenylalanine | 16 | MWT | 25482.96 |
| Aspartate | 13 | Proline | 6 | pI | 9.38 |
| Glutamine | 3 | Serine | 8 | Charge at pH 7 | 8.6 |
| Glutamate | 22 | Threonine | 13 | a.a. residues | 237 |
| Glycine | 48 | Tyrosine | 5 | | |
| Histidine | 7 | Valine | 10 | | |
| Isoleucine | 8 | Trytophan | 0 | | |
| Leucine | 15 | Cysteine | 0 | | |

**Table 5.2 Structural properties of nucleolin-RBD, MWT = molecular weight of fragment, a.a. = number of amino acids residues in fragment. Data obtained by nucleolin-RBD sequence analysis using ExPASy Protparam tool software.[73]**

### 5.4.1 Structural characterisation of oligonucleotides

The solution structure and conformation of different oligonucleotides was investigated in the presence of $K^+$ and $Na^+$ at different concentrations and temperatures using UV/visible absorbance and *CD* spectroscopy. The oligonucleotides investigated were chosen according to their sequence and structural properties (see table 5.1). All oligonucleotides were obtained from P. J. Bates and J. O. Trent at the University of Louisville already prepared in KCl or NaCl and the corresponding potassium or sodium phosphate buffer. Samples were used, as received. Each oligonucleotide solution in the relevant buffer was boiled for 5 minutes, and immediately cooled on ice before the process of annealing at 60 °C for 56 hours. The solutions were then allowed to cool to room temperature over several hours by suspending eppendorfs containing the oligonucleotide solutions in 1 L of water which was heated and allowed to cool. Prior to analysis using UV/visible absorbance and *CD* spectroscopy the solutions were placed in a freezer at –20 °C. The solutions were kept frozen until the day of the experiment at which point they were removed from the freezer and allowed to thaw to room temperature over several hours. Oligonucleotides at a final concentration of 5 μM were analysed using *CD* in a 1 cm path length masked cuvette. A buffer baseline was collected in the same cuvette and subtracted from the sample spectra. *CD* spectra were collected on a Jasco J-715 spectropolarimeter and experiments performed in duplicate.

*CD* parameters: scan range: 320 – 200 nm; scan speed: 100 nm/min; step resolution: 0.5 nm; response: 1.0 s; band width: 1.0 nm; sensitivity: 20 mdeg; accumulation: 16 scans; path length: 1 cm.

#### 5.4.1.1 Investigating the effect of freezing and annealing

To investigate the effect of freezing and annealing on the solution structures of the oligonucleotides in the presence of $K^+$ or $Na^+$, UV/visible absorbance and *CD* spectra were collected immediately after the frozen solutions had melted and again after the solutions had been re-annealed (see § 5.5.1.1). The oligonucleotides were annealed by heating at 90 °C for 10 minutes and then allowed to cool to room temperature over several hours by suspending eppendorfs containing the oligonucleotide solution in 1 L of water which was heated to 100 °C and allowed to cool.[14] A second method for annealing the oligonucleotides was also investigated. The oligonucleotides were heated at 90 °C for 10 minutes and then allowed to cool to room temperature on the bench (fast cooling). The results obtained from the fast cooling method were comparable to those

obtained using the slow cooling method, so all oligonucleotides after heating at 90 °C were cooled using the fast cooling method. The control for this experiment was analysis of the CRO1 oligonucleotide immediately after freezing and after the solution had been re-annealed. CRO1 was not expected to show any structural differences in the presence of $K^+$ or $Na^+$. It was observed from this experiment that the 'slow cooling' process was not required. *CD* analysis of the oligonucleotide samples after 'fast cooling' (whereby the samples were allowed to cool to room temperature on the bench, after annealing) provided spectra that overlaid exactly with the *CD* spectra obtained after 'slow cooling'.

*CD* parameters: scan range: 350 – 190 nm; scan speed: 100 nm/min; step resolution: 0.5 nm; response: 0.5 s; band width: 2.0 nm; sensitivity: 20 mdeg; accumulation: 16 scans; path length: 1 mm.

#### 5.4.1.2 Cation and concentration effect on quadruplex formation

A series of experiments were then designed to investigate the effect of different cationic salt solutions (KCl and NaCl) on the structure of GRO1 and CRO1 in solution and to investigate the effect of concentration on the formation of quadruplex structure. Stock solutions of KCl (500 mM) and NaCl (500 mM) were prepared. The oligonucleotides were annealed by heating at 90 °C for 10 minutes and then allowed to cool to room temperature over several hours by suspending eppendorfs containing the oligonucleotide solution in 1 L of water which was heated and allowed to cool.[14] 1 mm path length cuvettes were used to record the UV/visible absorbance and *CD* spectra of high concentration oligonucleotide (10 – 20 μM) and 1 cm path length cuvettes were used to record the UV/visible absorbance and *CD* spectra of low concentration oligonucleotide (1 – 5 μM). A solution (0.2 mL – 1 mm path length, or 2.0 mL – 1 cm path length) of oligonucleotide and either KCl or NaCl (100 mM) in the appropriate buffer was prepared. UV/visible absorbance and *CD* spectra were recorded in duplicate.

*CD* parameters: scan range: 350 – 190 nm; scan speed: 100 nm/min; step resolution: 0.5 nm; response: 0.5 s; band width: 2.0 nm; sensitivity: 20 mdeg; accumulation: 16 scans; path length: 1 mm or 1 cm.

### 5.4.2 Structural characterisation of nucleolin-RBD

Structural characterisation of nucleolin has previously been limited to the use of NMR, X-ray crystallography and molecular dynamics simulations. As described above nucleolin may be involved in many cellular processes and a particular aim of this work

has been to probe the quadruplex structure of the GRO's upon binding to a specific cellular protein (which has been tentatively identified as nucleolin), using circular dichroism. Nucleolin-RBD (1.0 mg) was placed in potassium or sodium phosphate buffer (1.0 mL) and the UV/visible absorbance and *CD* spectra recorded. Spectra were collected in a 1 mm path length cuvette. To determine the concentration of nucleolin-RBD in solution, a Bio-rad assay was completed. The Bio-rad protein assay, based on the method of Bradford, is a dye-binding assay in which a differential colour change occurs in response to various concentrations of protein.[74] It involves the addition of an acidic dye to the protein solution and subsequent measurement of the absorbance at 595 nm.[75] A solution was prepared containing water (790 μL), protein (10 μL in buffer) and Bio-rad agent (200 μL). The value of 1.95 is an experimental constant. The experiment was repeated 4 times and the average concentration determined using the following equation:[76]

$$\frac{A_{595}}{0.1} \times 1.95 \times \text{dilution factor} \times \frac{1000}{10} = \text{[nucleolin-RBD] } \mu g/mL$$

Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) was also used to confirm the molecular weight of the nucleolin-RBD. The preparation of the separating and stacking gels is summarised in table 5.3. The running buffer was prepared by mixing tris base (3.0 g), glycine (14.4 g) and SDS (1 g) in water (1 L). The molecular weight marker and nucleolin-RBD were prepared in sample buffer containing water (4 mL), tris-HCl (0.5 M, 1 mL), glycerol (0.8 mL), β-mercaptoethanol (400 μL) and bromophenol blue (0.05% w/v, 200 μL).

| | **Separating Gel** | **Stacking gel** |
|---|---|---|
| $H_2O$ | 4.1 mL | 3.075 mL |
| 1.5 M tris-HCl | 2.5 mL | |
| 0.5 M tris-HCl | | 1.25 mL |
| 10% SDS | 50 μL | 25 μL |
| 30% A/bA | 3.3 mL | 670 μL |
| Temed | 5 μL | 5 μL |
| 10% APS | 50 μL | 25 μL |

**Table 5.3 Preparation of separation and stacking gels for the gel electrophoresis. A/bA = acrylamide/bis-acrylamide and APS = ammonium persulphate. APS is the gel polymerising agent and so was added to the solution last of all.**

The separating gel was poured into the electrophoresis plates and gently overlaid with water. When the gel had polymerised the excess water was removed and the stacking

gel poured on top. The well-comb was inserted and the stacking gel allowed to polymerise. The molecular weight marker and nucleolin-RBD samples were heated at 95 °C for 5 minutes prior to loading and the apparatus was filled with the running buffer. The well-comb was removed and the samples loaded onto the gel. The gel was run at 200 V for one hour. The gel was removed and Coomassie blue stain used to confirm the presence of the protein. These experiments were completed with the help of Dr. S. J. A. Lautru, Dr. T. Dafforn, Dr. D. Roper and R. Pacheco-Gomez.

### 5.4.3 Oligonucleotide-nucleolin-RBD interactions

The main aim of the work presented in this chapter was to see if a proposed target — nucleolin-RBD — bound to GRO's. A nucleolin-RBD stock solution was prepared by dissolving nucleolin-RBD (1.0 mg) in the appropriate $K^+$ or $Na^+$ phosphate buffer (1 mL). A 0.2 mL solution of oligonucleotide (20 μM) and either KCl or NaCl (100 mM) in the corresponding $K^+$ or $Na^+$ phosphate buffer was prepared. The nucleolin-RBD (1 – 5 μM) was titrated into the solution at known concentrations and to eliminate any dilution effects the same volume of double concentration oligonucleotide (40 μM) was added. UV/visible absorbance and *CD* spectra were recorded in duplicate. Identical instrumental parameters were used in all experiments to ensure reproducibility.

UV/visible absorbance parameters were: scan range: 350 – 190 nm; scan speed: 200 nm/min; response: fast; band width: 1.0 nm; data pitch: 0.5 nm; accumulation: 1. *CD* parameters were: scan range: 350 – 190 nm; scan speed: 200 nm/min; step resolution: 0.5 nm; response: 0.5 s; band width: 1.0 nm; sensitivity: 20 mdeg; accumulation: 16 scans; path length: 1 mm.

#### 5.4.3.1 Kinetic and thermodynamic stability of oligonucleotides and nucleolin-RBD

The kinetic and thermodynamic stability of GRO1 and CRO1 was investigated using UV/visible absorbance and *CD* spectroscopy in the presence of $K^+$ and $Na^+$ and upon interaction with nucleolin-RBD. The thermodynamic stability was investigated over a wide temperature range (4 – 90 °C). A 0.2 mL solution of GRO1 or CRO1 (20 μM) and either KCl or NaCl (100 mM) in the corresponding $K^+$ or $Na^+$ phosphate buffer was prepared. The experiment was repeated upon addition of nucleolin-RBD (5 μM) and all experiments performed in duplicate.

## 5.5 Results

The aim of the research presented in this chapter was initially to investigate the kinetic and thermodynamic stability of GRO's and CRO's under different experimental conditions. Previous studies had showed that GRO's can adopt several conformations in the presence of different monovalent cations and at various temperatures,[1, 2, 6, 8, 9] so this phenomenon was further investigated using UV/visible absorbance and *CD* spectroscopy. Cytosine rich oligonucleotides (CRO's) were used as the control as they were not expected to show any quadruplex secondary structure in the presence of buffer, NaCl or KCl.

The second part of this research was focussed on the structural characterisation of nucleolin-RBD. It has been speculated that nucleolin has a role in many sub-cellular processes.[70 – 72] Circular dichroism was used to investigate the secondary structure of the protein and it was observed from the Bio-rad assay that the concentration of protein in solution suggested limited solubility in the buffers used.

The final part of this research was focussed on the investigation of the interactions and binding of GRO1 and CRO1 with nucleolin-RBD. It has also been speculated that the most accurate factor in predicting biological activity of GRO's is the ability of the oligonucleotide to bind to the GRO-binding protein. As with the experiments involving the interaction of metal complexes with DNA, upon binding of the nucleolin-RBD to GRO1, changes in the spectral shape and intensity were expected and indeed observed under certain experimental conditions.

### 5.5.1 Structural characterisation and biological activity of oligonucleotides

The purpose of this part of the investigation was to clarify the relationship between *CD* spectral characteristics and quadruplex properties and also to determine if any features in the *CD* spectra of quadruplexes could be associated with biological activity. The data presented in figure 5.8 shows the *CD* spectra of all oligonucleotides annealed in buffers containing either KCl (0.1 M) or NaCl (0.1 M). Certain trends were observed, but there was no signature *CD* spectrum that was unequivocally linked with anti-proliferative activity.

It can clearly be observed from the *CD* spectra (see figure 5.8) of the different oligonucleotides that different structures are formed in the presence of KCl or NaCl and

that the structures cannot be simply classified as 'parallel' or 'folded'. Many of the oligonucleotides investigated show peaks in the *CD* spectrum at both 265 and 295 nm (see § 5.2) which probably indicates the presence of multiple quadruplex species. It was observed from the *CD* spectra (see figure 5.8) that oligonucleotides with small differences in ellipticity in the presence of $K^+$ or $Na^+$ (KS-A, KS-D, KS-E and KS-F) also had weak anti-proliferative activity and that limited quadruplex formation occurred under these conditions. Oligonucleotides KS-I and GRO23A (see figure 5.8 (i) and (k)) show different *CD* signatures in the presence of KCl and NaCl buffers and clearly exhibited different anti-proliferative activity in the buffers.[1]

The spectra suggest that there is not one structural reason for biological activity, but that several spectral features may be related to activity. These include a large positive *CD* signal at 264 nm or 295 nm, significant ellipticity in the 290 – 310 nm region and a large positive peak at 210 nm.

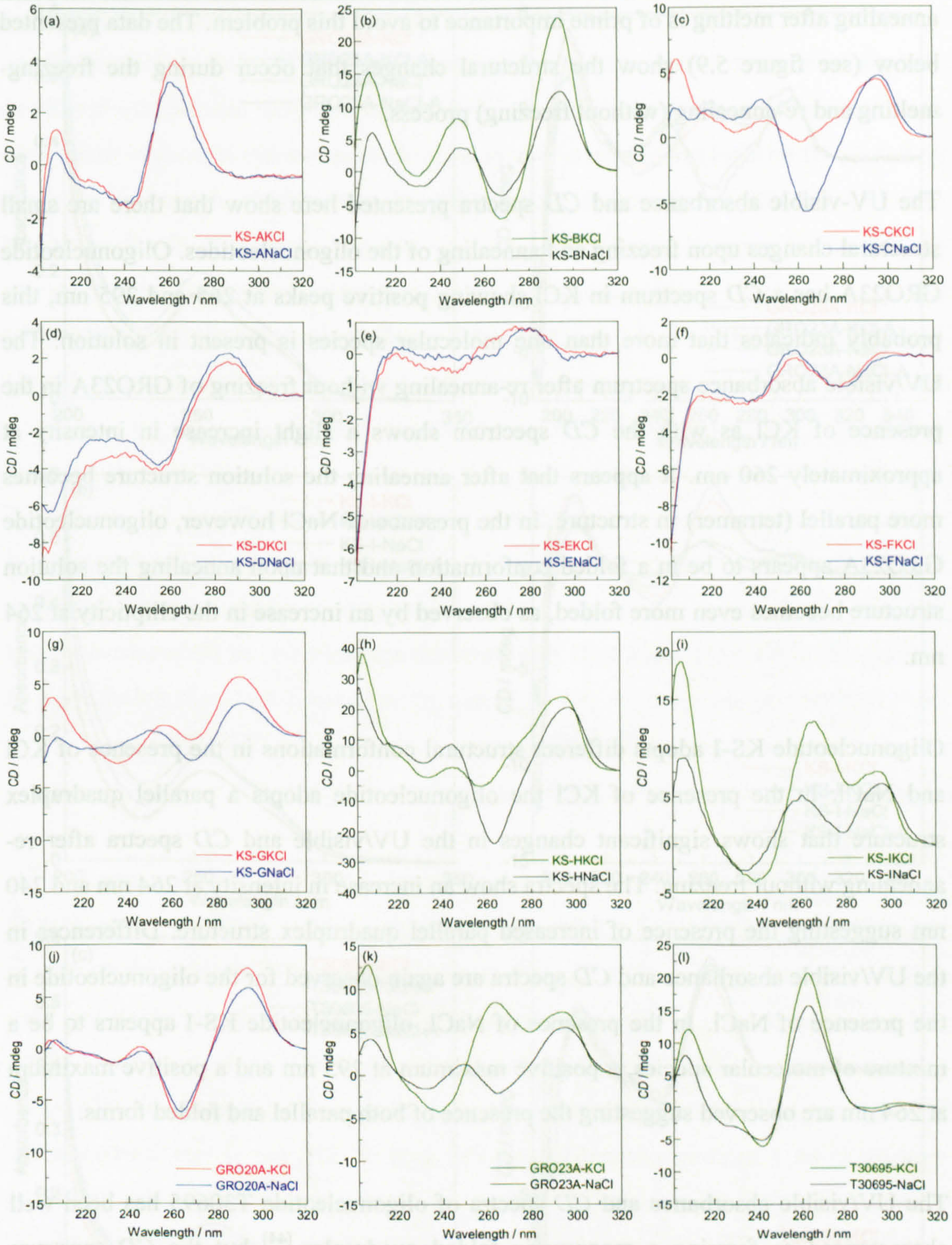



**Figure 5.8 *CD* spectra of oligonucleotides (5 μM final concentration) obtained in the presence of 0.1 M KCl or NaCl and the corresponding phosphate buffer at 25 °C in a 1 cm path length cuvette. (a) KS-A (b) KS-B (c) KS-C (d) KS-D (e) KS-E (f) KS-F (g) KS-G (h) KS-H (i) KS-I (j) GRO20A (k) GRO23A (l) T30695. Key: — and — = no biological activity; — and — = biologically active.[1]**

### 5.5.1.1 Investigating the effect of freezing and annealing

It was observed early in the investigation that the process of freezing and melting changes the structural conformations of both GRO's and CRO's in solution and that

annealing after melting is of prime importance to avoid this problem. The data presented below (see figure 5.9) show the structural changes that occur during the freezing-melting and re-annealing (without freezing) process.

The UV-visible absorbance and *CD* spectra presented here show that there are small structural changes upon freezing and annealing of the oligonucleotides. Oligonucleotide GRO23A has a *CD* spectrum in KCl showing positive peaks at 265 and 295 nm, this probably indicates that more than one molecular species is present in solution. The UV/visible absorbance spectrum after re-annealing without freezing of GRO23A in the presence of KCl as with the *CD* spectrum shows a slight increase in intensity at approximately 260 nm. It appears that after annealing the solution structure becomes more parallel (tetramer) in structure. In the presence of NaCl however, oligonucleotide GRO23A appears to be in a folded conformation and that upon annealing the solution structure becomes even more folded, as observed by an increase in the ellipticity at 264 nm.

Oligonucleotide KS-I adopts different structural conformations in the presence of KCl and NaCl. In the presence of KCl the oligonucleotide adopts a parallel quadruplex structure that shows significant changes in the UV/visible and *CD* spectra after re-annealing without freezing. The spectra show an increase in intensity at 264 nm and 240 nm suggesting the presence of increased parallel quadruplex structure. Differences in the UV/visible absorbance and *CD* spectra are again observed for the oligonucleotide in the presence of NaCl. In the presence of NaCl, oligonucleotide KS-I appears to be a mixture of molecular species, a positive maximum at 295 nm and a positive maximum at 264 nm are observed suggesting the presence of both parallel and folded forms.

The UV/visible absorbance and *CD* spectra of oligonucleotide T30695 has been well characterised as forming a monomeric folded quadruplex,[44] but the *CD* spectrum shows a strong positive ellipticity at 264 nm, suggesting the formation of a parallel quadruplex structure. The presence of KCl or NaCl appears to have little effect on the solution structure of this oligonucleotide. The quadruplex structure in NaCl is not affected by the re-annealing process, but in the presence of KCl, there is a significant increase in the peak intensity at 264 nm in both the UV/visible absorbance and *CD* spectra suggesting that the solution structure becomes even more parallel after re-annealing.

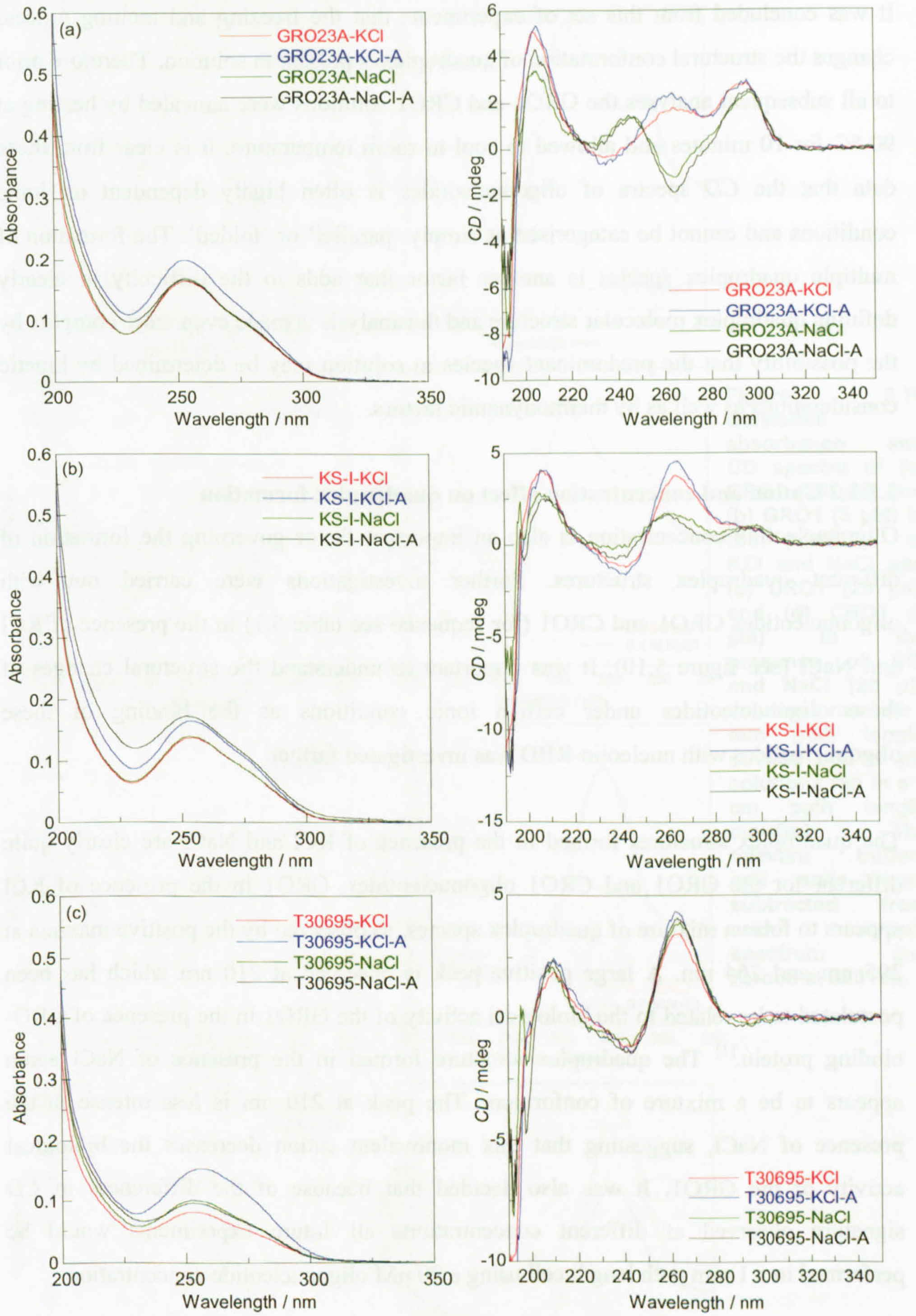


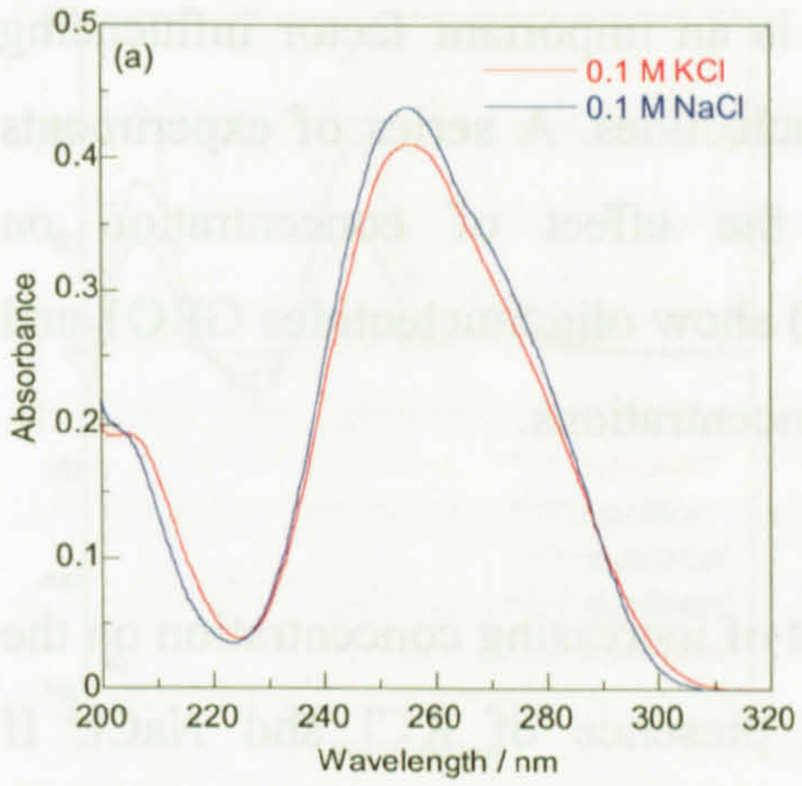

**Figure 5.9 UV/visible absorbance and *CD* spectra of (a) GRO23A, (b) KS-I and (c) T30695 (5 μM final concentration) in KCl and NaCl before and after re-annealing (without freezing) (-A). Spectra collected in a 1 mm path length cell and the relevant buffers and salts subtracted from each sample spectrum and zeroed at 350 nm.**

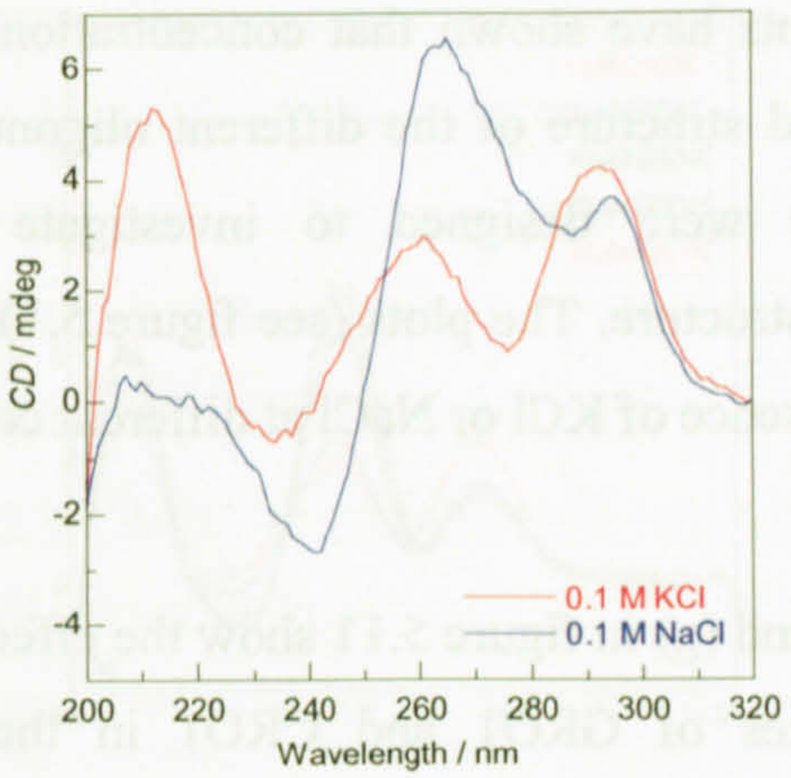
It was concluded from this set of experiments that the freezing and melting process changes the structural conformation of quadruplexes present in solution. Therefore prior to all subsequent analyses the GRO1 and CRO1 solutions were annealed by heating at 90 °C for 10 minutes and allowed to cool to room temperature. It is clear from these data that the *CD* spectra of oligonucleotides is often highly dependent on ionic conditions and cannot be categorised as simply 'parallel' or 'folded'. The formation of multiple quadruplex species is another factor that adds to the difficulty in clearly defining quadruplex molecular structure and the analysis is made even more complex by the possibility that the predominant species in solution may be determined by kinetic considerations as well as by thermodynamic factors.

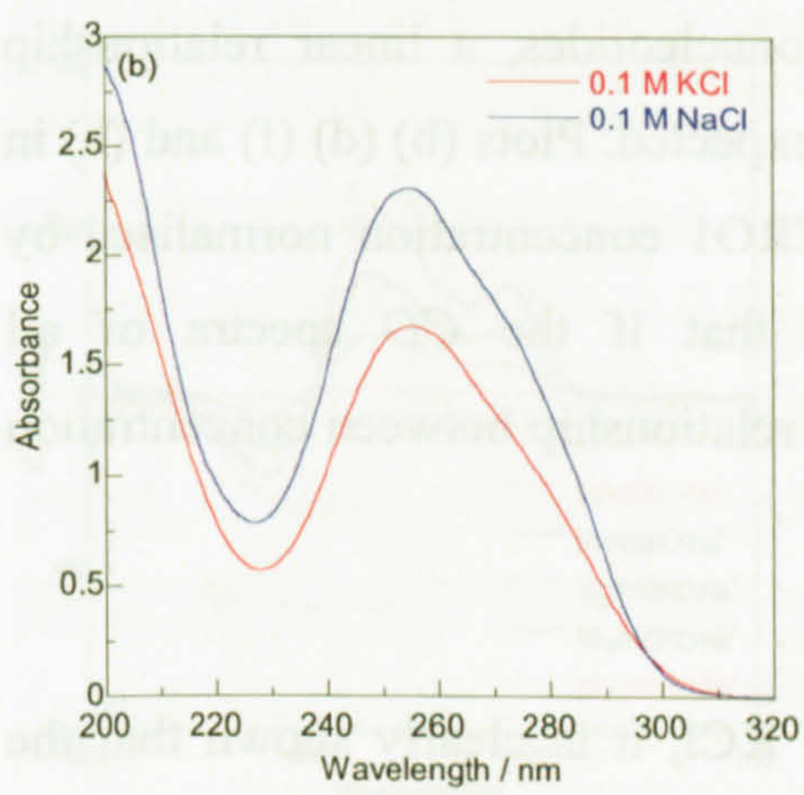
### 5.5.1.2 Cation and concentration effect on quadruplex formation

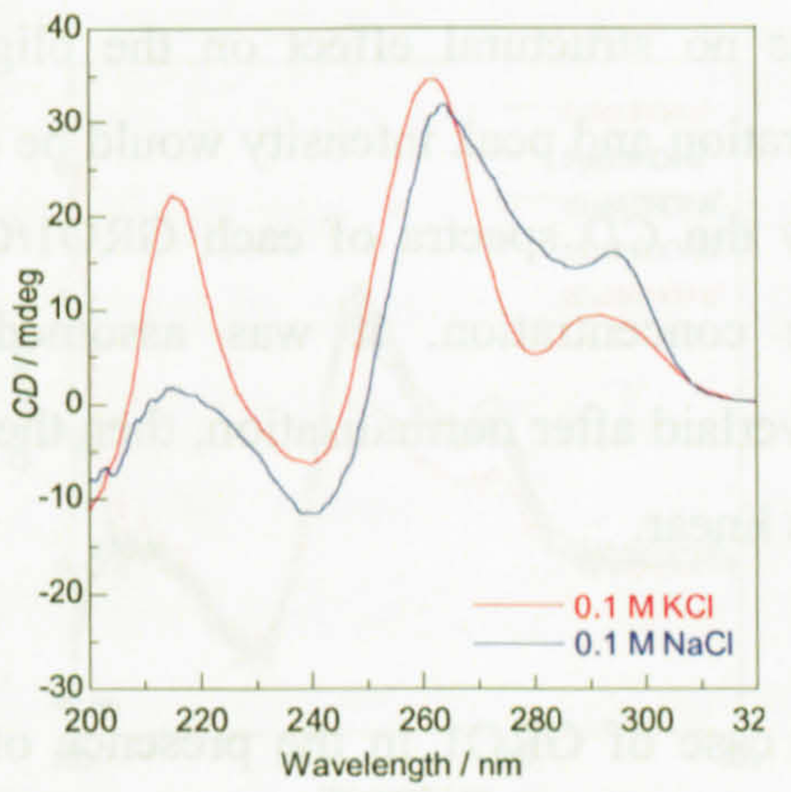
Oligonucleotide concentration is also an important factor governing the formation of different quadruplex structures. Further investigations were carried out with oligonucleotides GRO1 and CRO1 (for sequence see table 5.1) in the presence of KCl and NaCl (see figure 5.10). It was important to understand the structural changes of these oligonucleotides under certain ionic conditions as the binding of these oligonucleotides with nucleolin-RBD was investigated further.

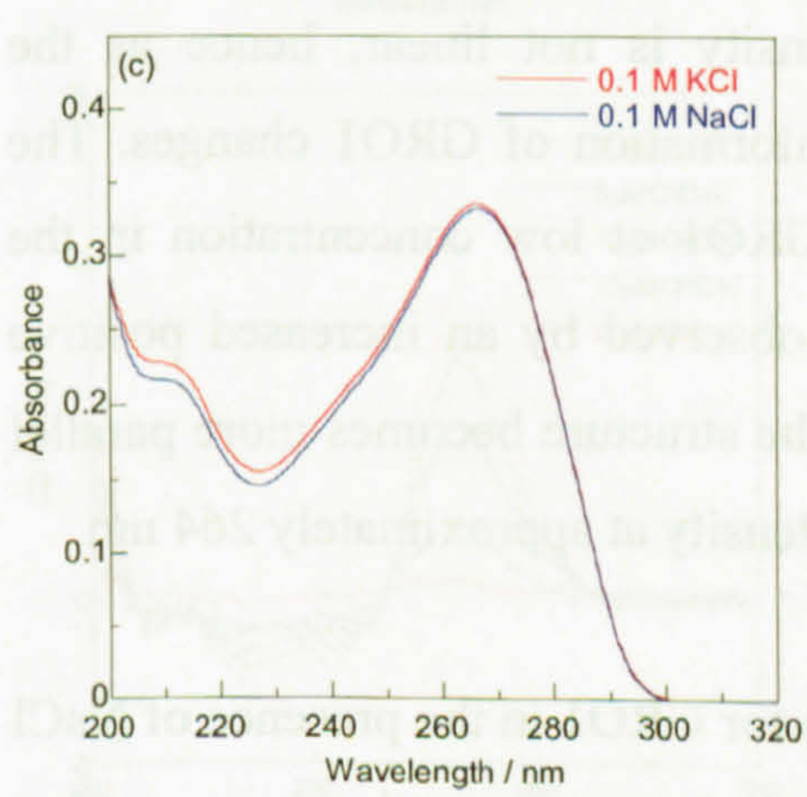
The quadruplex structures formed in the presence of KCl and NaCl are clearly quite different for the GRO1 and CRO1 oligonucleotides. GRO1 in the presence of KCl appears to form a mixture of quadruplex species, as observed by the positive maxima at 295 nm and 264 nm. A large positive peak is observed at 210 nm which has been postulated to be related to the biological activity of the GRO1 in the presence of GRO-binding protein.[1] The quadruplex structure formed in the presence of NaCl again appears to be a mixture of conformers. The peak at 210 nm is less intense in the presence of NaCl, suggesting that this monovalent cation decreases the biological activity of the GRO1. It was also decided that because of the differences in *CD* signature observed at different concentrations all future experiments would be performed in a 1 mm path length cell using a 20 μM oligonucleotide concentration.

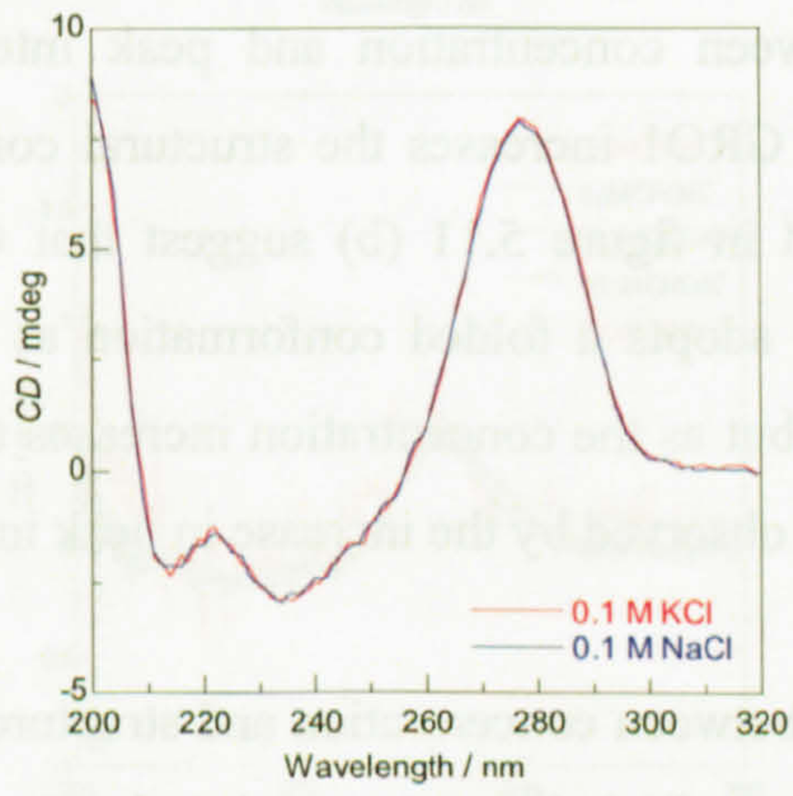


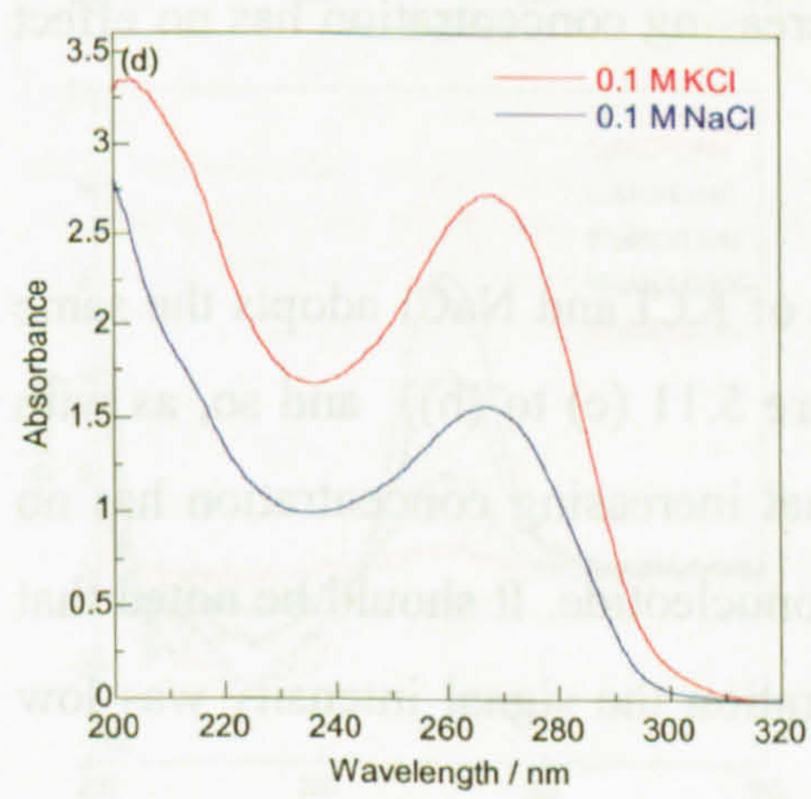



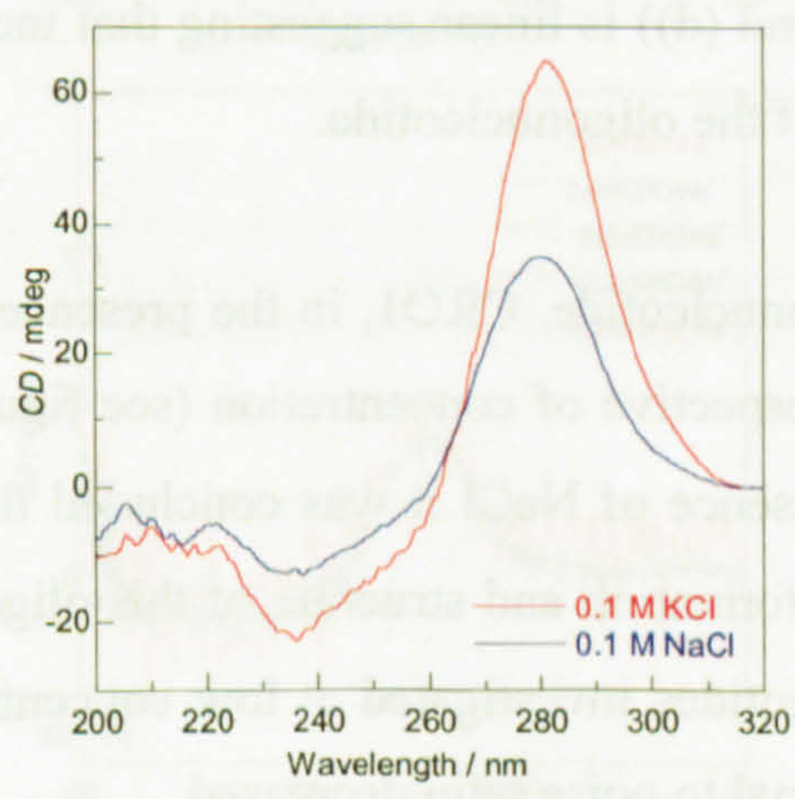




**Figure 5.10 UV/visible absorbance and *CD* spectra of (a) GRO1 (20 μM) and (b) GRO1 (5 μM) in the presence of KCl and NaCl and (c) CRO1 (20 μM) and (d) CRO1 (5 μM) in the presence of KCl and NaCl (20 μM solutions ran in a 1 mm path length cuvette and 5 μM solutions ran in a 1 cm path length cuvette). The relevant buffers and salts were subtracted from each sample spectrum and zeroed at 320 nm.**

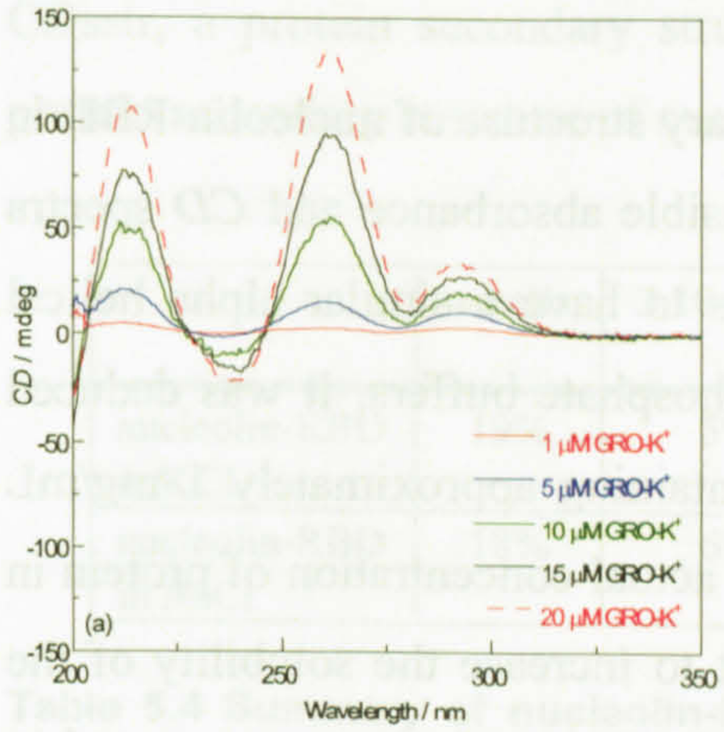
These experiments have shown that concentration is an important factor influencing conformation and structure of the different oligonucleotides. A series of experiments (see § 5.4.1.2) were designed to investigate the effect of concentration on oligonucleotide structure. The plots (see figure 5.11) show oligonucleotides GRO1 and CRO1 in the presence of KCl or NaCl at different concentrations.

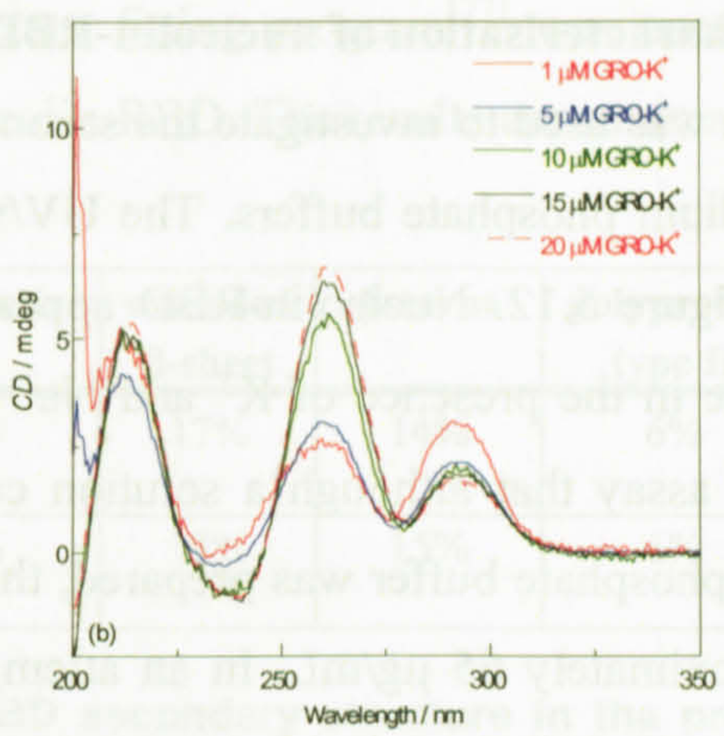
Plots (a) (c) (e) and (g) in figure 5.11 show the effect of increasing concentration on the solution structures of GRO1 and CRO1 in the presence of KCl and NaCl. If concentration has no structural effect on the oligonucleotides, a linear relationship between concentration and peak intensity would be expected. Plots (b) (d) (f) and (h) in figure 5.11 show the *CD* spectra of each GRO1/CRO1 concentration normalised by dividing by the concentration. It was assumed that if the *CD* spectra of all concentrations overlaid after normalisation, then the relationship between concentration and structure was linear.

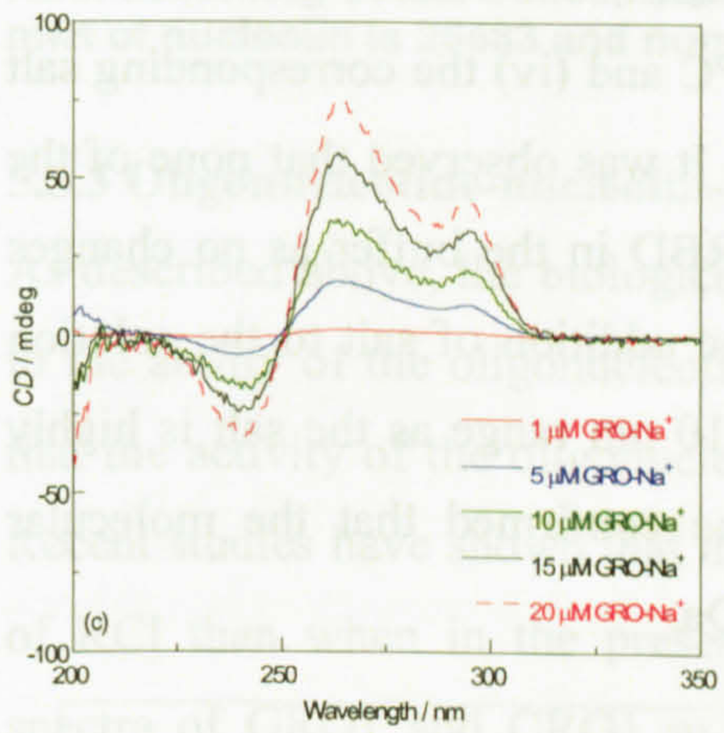
However, in the case of GRO1 in the presence of KCl, it is clearly shown that the relationship between concentration and peak intensity is not linear, hence as the concentration of GRO1 increases the structural conformation of GRO1 changes. The spectra presented in figure 5.11 (b) suggest that GRO1 at low concentration in the presence of KCl adopts a folded conformation as observed by an increased positive peak at 295 nm, but as the concentration increases the structure becomes more parallel and tetrameric as observed by the increase in peak intensity at approximately 264 nm.

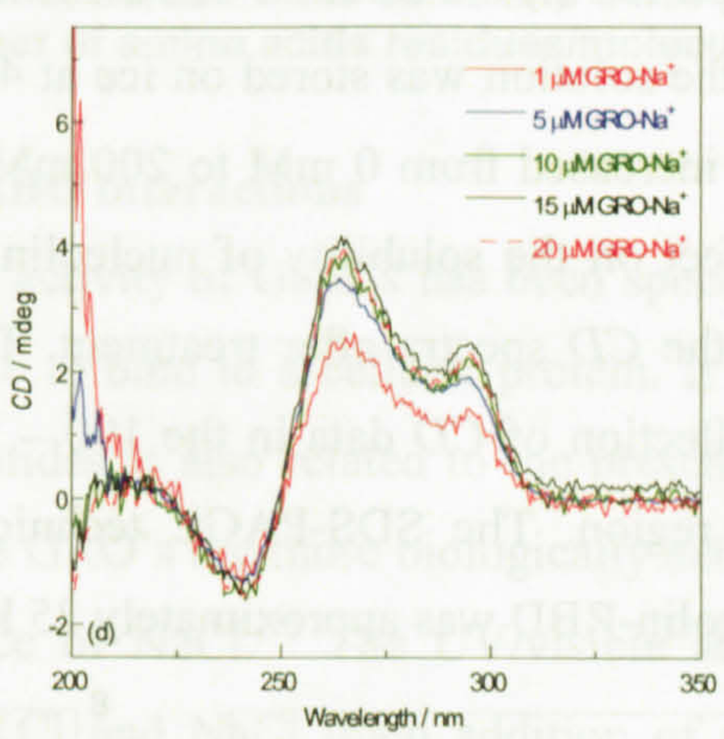
The relationship between concentration and structure for GRO1 in the presence of NaCl (figure 5.11 (c) and (d)) is linear suggesting that increasing concentration has no effect on the structure of the oligonucleotide.

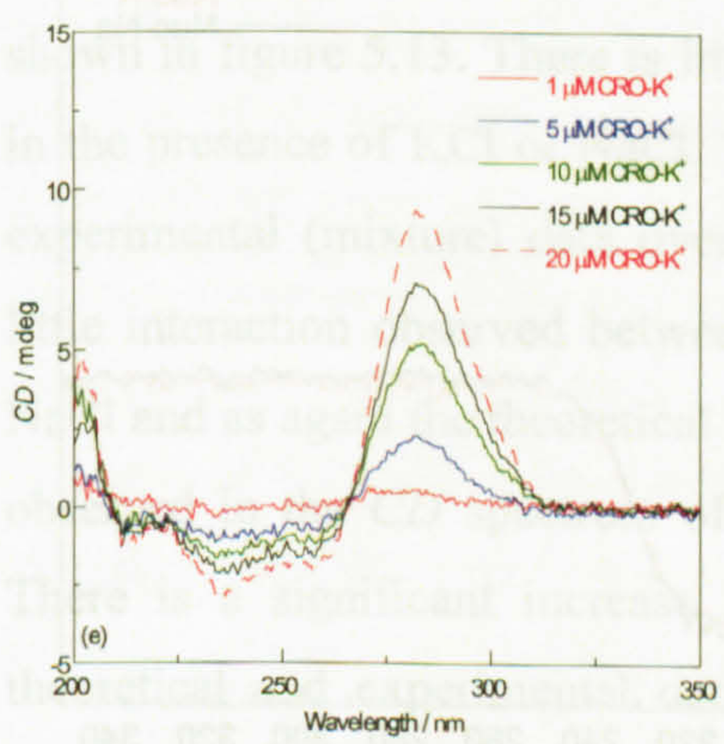
The control oligonucleotide, CRO1, in the presence of KCl and NaCl adopts the same conformation irrespective of concentration (see figure 5.11 (e) to (h)) and so, as with GRO1 in the presence of NaCl it was concluded that increasing concentration has no effect on the conformation and structure of this oligonucleotide. It should be noted that for all oligonucleotides investigated at low concentration the signal intensity was low and hence the signal to noise ratio decreased.

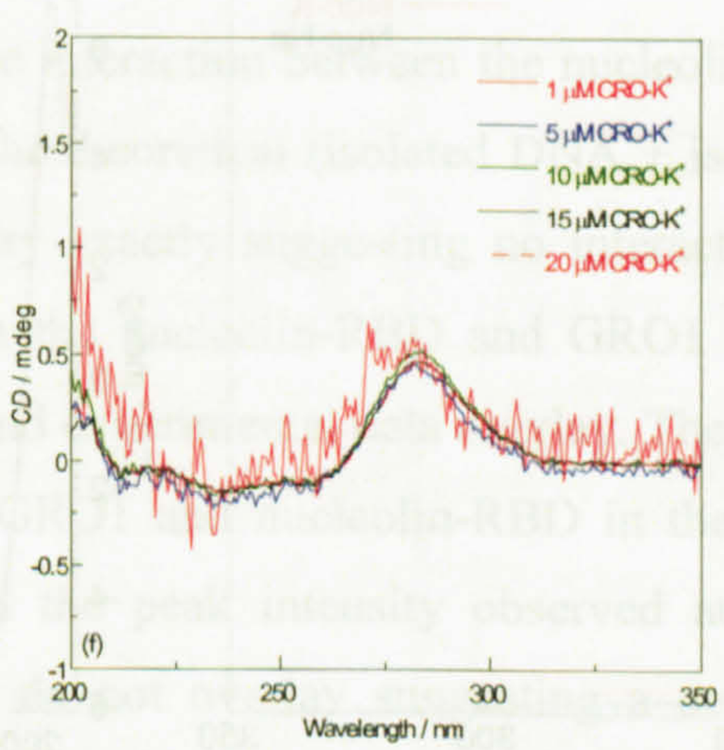


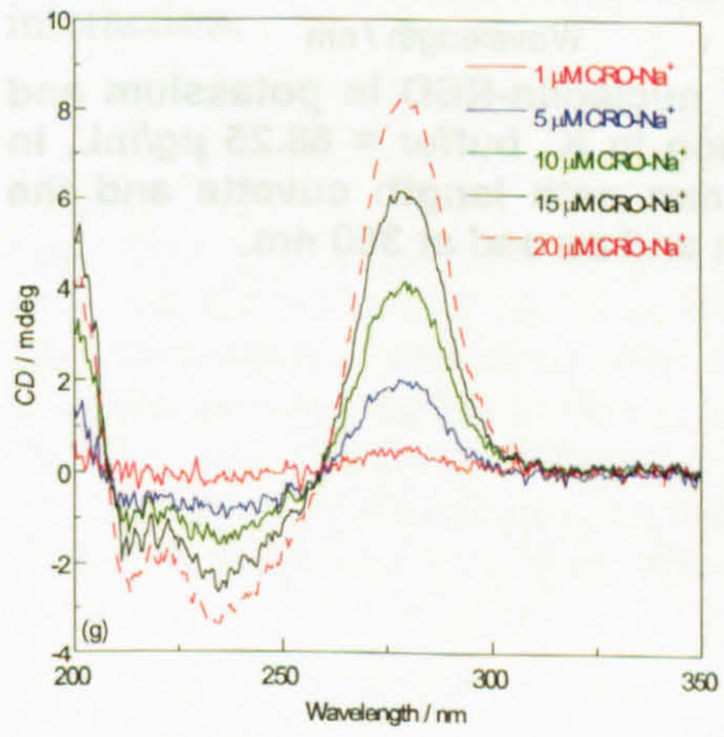



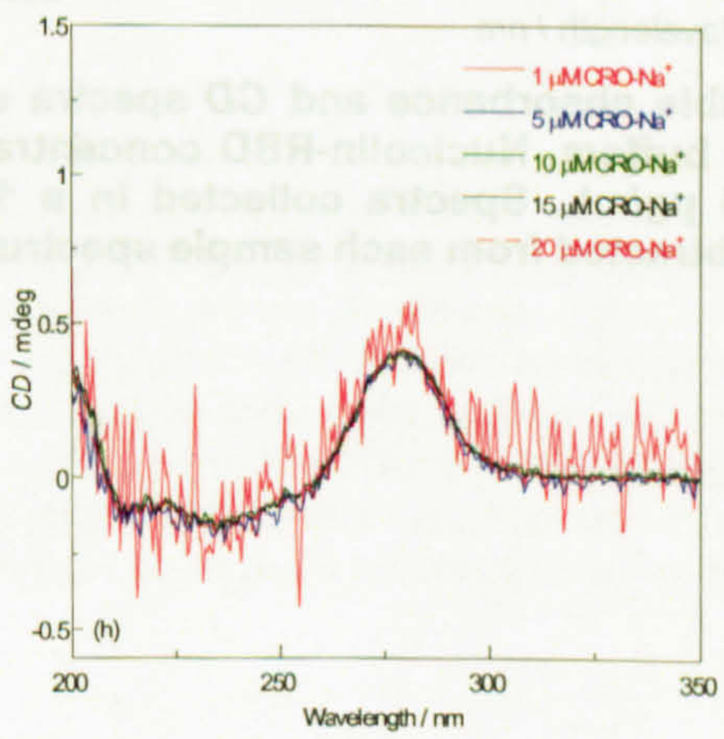




**Figure 5.11 *CD* spectra of (a) GRO1 in the presence of KCl (b) normalised spectra of GRO1 in the presence of KCl (c) GRO1 in the presence of NaCl (d) normalised spectra of GRO1 in the presence of NaCl (e) CRO1 in the presence of KCl (f) normalised spectra of CRO1 in the KCl (g) CRO1 in the presence of NaCl and (h) normalised spectra of CRO1 in the presence of NaCl. The appropriate buffers were subtracted from each sample spectrum and the data zeroed at 350 nm. Spectra were normalised by dividing each sample spectrum by the concentration of each oligonucleotide. High concentration oligonucleotide (10 – 20 μM) data were collected in a 1 mm path length cell and low concentration oligonucleotide (1 – 5 μM) data were collected in a 1 cm path length cell.**

### 5.5.2 Structural characterisation of nucleolin-RBD

Circular dichroism was used to investigate the secondary structure of nucleolin-RBD in potassium and sodium phosphate buffers. The UV/visible absorbance and *CD* spectra are presented in figure 5.12. Nucleolin-RBD appears to have a similar alpha helical secondary structure in the presence of $K^+$ and $Na^+$ phosphate buffers. It was deduced from the Bio-Rad assay that although a solution containing approximately 1 mg/mL nucleolin-RBD in phosphate buffer was prepared, the actual concentration of protein in solution was approximately 65 μg/mL. In an attempt to increase the solubility of the nucleolin-RBD in buffer: (i) the solution was sonicated, (ii) the solution was warmed in a water bath, (iii) the solution was stored on ice at 4 °C and (iv) the corresponding salt concentration was increased from 0 mM to 200 mM. It was observed that none of the above had any effect on the solubility of nucleolin-RBD in the buffer as no changes were observed in the *CD* spectra after treatment. The addition of salt to the solution jeopardised the collection of *CD* data in the 190 – 210 nm range as the salt is highly absorbing in this region. The SDS-PAGE technique confirmed that the molecular weight of the nucleolin-RBD was approximately 25 kDa.

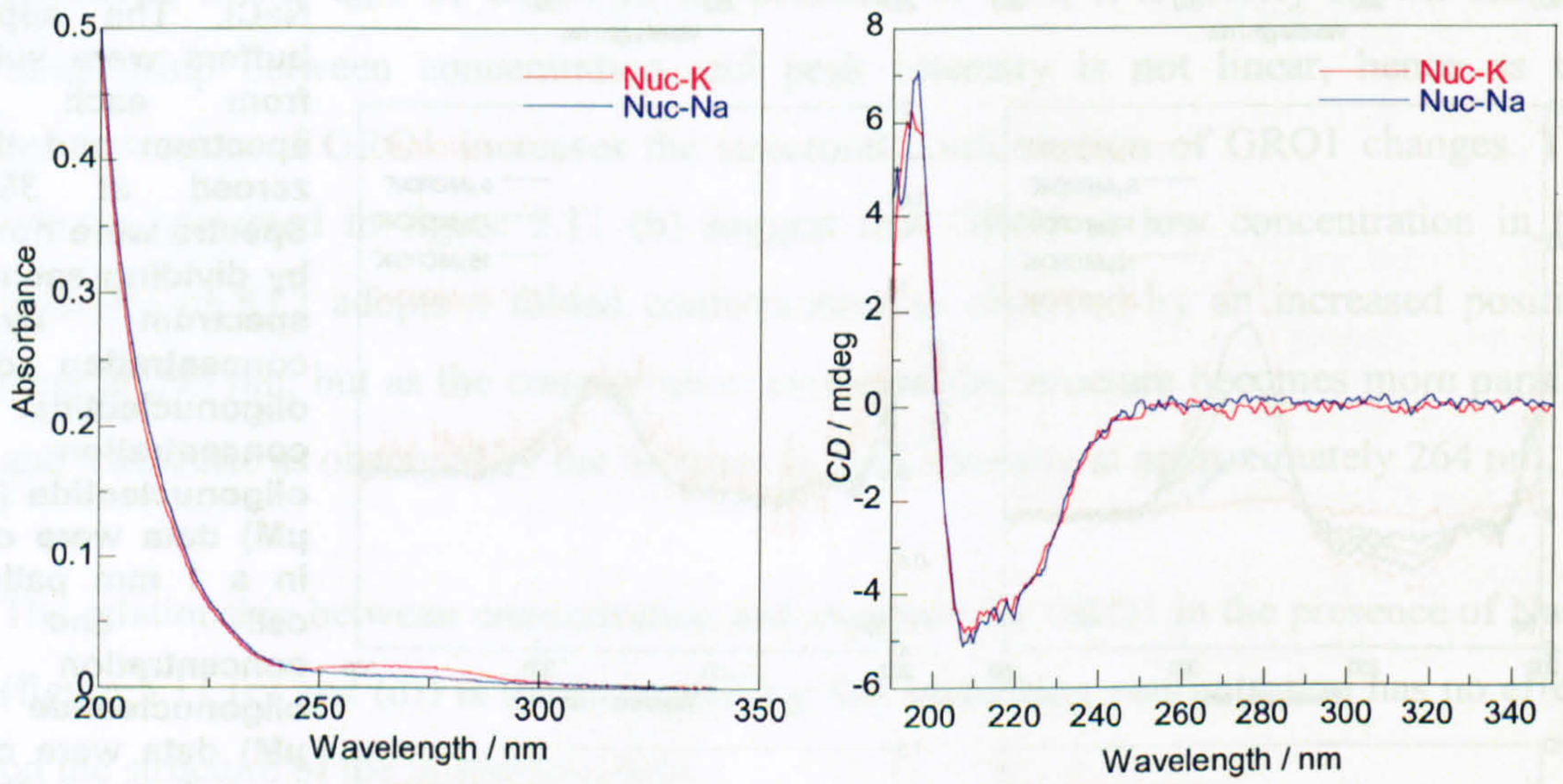



**Figure 5.12 UV/visible absorbance and *CD* spectra of nucleolin-RBD in potassium and sodium phosphate buffers. Nucleolin-RBD concentration in $K^+$ buffer = 68.25 μg/mL, in $Na^+$ buffer = 64.35 μg/mL. Spectra collected in a 1 mm path length cuvette and the relevant buffers subtracted from each sample spectrum and zeroed at 350 nm.**

CDsstr, a protein secondary structure fitting program[77] was used to determine the protein secondary structure of nucleolin-RBD. The results are summarised in table 5.4.

| | α-helix | 3,10-helix | extended β-sheet | β-turns | polyproline type II | other structures |
|---|---|---|---|---|---|---|
| nucleolin-RBD in KCl | 19% | 5% | 17% | 14% | 6% | 37% |
| nucleolin-RBD in NaCl | 18% | 6% | 18% | 15% | 6% | 37% |

**Table 5.4 Summary of nucleolin-RBD secondary structure in the presence of KCl and NaCl assuming actual concentration of solution was 68.25 and 64.35 µg/mL respectively, mwt of nucleolin is 25483 and number of amino acids residues/molecule is 237.**

### 5.5.3 Oligonucleotide-nucleolin-RBD interactions

As described above, the biological activity of GRO's has been speculated to be related to the ability of the oligonucleotide to bind to a cellular protein. It has been observed that the activity of the oligonucleotides is also related to the presence of $K^+$ or $Na^+$.[1] Recent studies have shown that the GRO's are more biologically active in the presence of KCl than when in the presence of NaCl.[1] The UV/visible absorbance and *CD* spectra of GRO1 and CRO1 in KCl and NaCl upon addition of nucleolin-RBD are shown in figure 5.13. There is little interaction between the nucleolin-RBD and CRO1 in the presence of KCl or NaCl. The theoretical (isolated DNA + isolated protein) and experimental (mixture) data overlay exactly suggesting no interaction. There is very little interaction observed between the nucleolin-RBD and GRO1 in the presence of NaCl and as again the theoretical and experimental data overlay. The intriguing result is observed in the *CD* spectrum of GRO1 and nucleolin-RBD in the presence of KCl. There is a significant increase in the peak intensity observed at 240 nm and the theoretical and experimental data do not overlay suggesting a definite well-defined interaction.

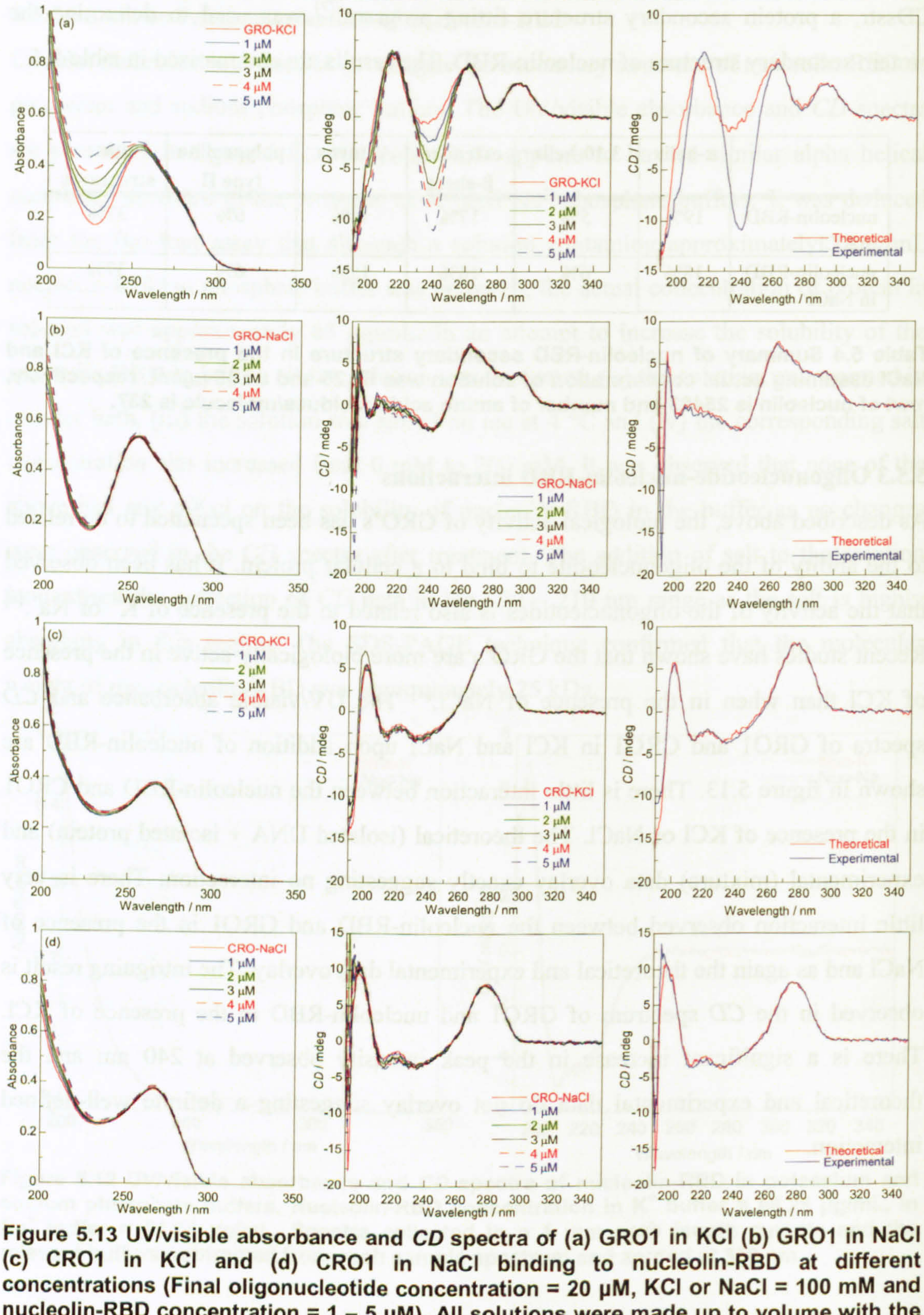



**Figure 5.13 UV/visible absorbance and *CD* spectra of (a) GRO1 in KCl (b) GRO1 in NaCl (c) CRO1 in KCl and (d) CRO1 in NaCl binding to nucleolin-RBD at different concentrations (Final oligonucleotide concentration = 20 μM, KCl or NaCl = 100 mM and nucleolin-RBD concentration = 1 – 5 μM). All solutions were made up to volume with the corresponding $K^+$ or $Na^+$ buffer. Also shown are the theoretical (20 μM oligonucleotide data + 5 μM nucleolin-RBD data) and experimental (20 μM oligonucleotide + 5 μM nucleolin-RBD data) spectra. Data was collected in a 1 mm path length cuvette and the relevant buffers and salts subtracted from each sample spectrum and zeroed at 350 nm.**

### 5.5.3.1 Kinetic and thermodynamic stability of oligonucleotides and nucleolin-RBD

The kinetic and thermodynamic stability of GRO1 and CRO1 were investigated in different buffers and in the presence of $K^+$ or $Na^+$ over a wide temperature range (4 – 90 °C). The experiment was repeated upon addition of nucleolin-RBD to investigate whether the nucleolin-RBD had any effect on the stability of the quadruplex structures formed. The *CD* spectra and melting curves for GRO1 in the presence of KCl or NaCl and with and without nucleolin-RBD (5 μM) are presented in figure 5.14. All of the solutions were prepared at room temperature and then cooled to 4 °C in the Peltier unit on the Jasco J-715 *CD* spectropolarimeter. It can be observed from a comparison of the spectrum obtained for GRO1 in KCl at room temperature (see figure 5.10 (b)) that the structure of the GRO1 is affected by temperature. The *CD* spectrum of GRO1 in KCl at room temperature does not overlay with the *CD* spectrum obtained for GRO1 in KCl at 4 °C. In contrast to the GRO1 in KCl spectrum, upon addition of the nucleolin-RBD, the *CD* spectra at 4 °C and at room temperature (see figure 5.13 (a)) do overlay exactly, suggesting that upon addition of the nucleolin-RBD the GRO1 quadruplex structure is considerably stabilised.

The *CD* spectra presented in figure 5.14 show the melting curves of GRO1 in the presence of $K^+$ and also upon addition of nucleolin-RBD. As described above, it was observed that upon addition of nucleolin-RBD to GRO1 in the presence of KCl the solution structure at 4 °C overlaid exactly with that obtained for the same solution at room temperature. The melting temperature of GRO1 in the presence of KCl and in the presence of KCl and nucleolin-RBD are both approximately 40 °C at 214 nm. The melting curves were plotted at different wavelengths to investigate the effect that the addition of nucleolin-RBD had on the GRO structure. The spectra (see figure 5.14 (iv) and (v)) show that the melting temperature of GRO-1 at 260 nm is approximately the same (45 °C) in the presence or absence of nucleolin-RBD. On the basis that the GRO1 solution contains a mixture of quadruplex conformers (parallel or folded), this may indicate that the parallel conformers are not stabilised by the nucleolin-RBD. However, it can be observed that GRO1 in the presence of nucleolin-RBD and KCl shows an increased melting temperature, suggesting that the folded quadruplex species are stabilised by the nucleolin-RBD.

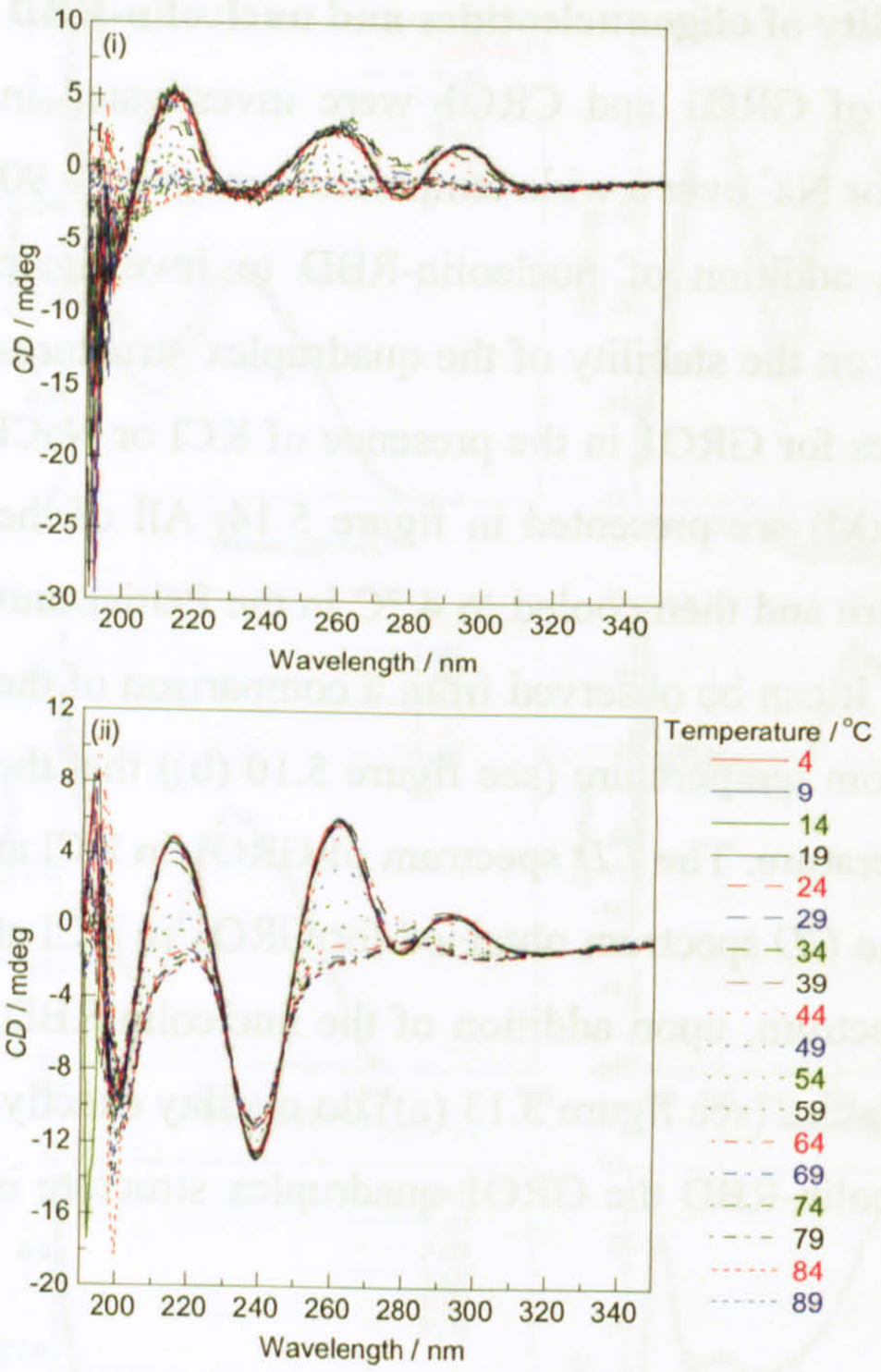



**Figure 5.14 *CD* spectra as a function of T of (i) GRO1 in the presence of KCl, (ii) GRO1 + nucleolin-RBD in the presence of KCl (iii) melting curves of GRO1 and GRO1 + nucleolin-RBD in the presence of KCl at 214 nm (iv) at 260 nm and (v) at 295 nm. Final oligonucleotide concentration = 20 μM and final nucleolin-RBD concentration = 5 μM. All spectra were collected in a 1 mm path length cell and temperature range investigated = 4 – 90 °C.**

It was observed that the stability of the GRO1 in the presence of NaCl does not change upon addition of nucleolin-RBD (see figure 5.15 (iii) and (iv)). This result was as expected based on the previous result of the GRO1-nucleolin-RBD binding experiment in the presence of NaCl where no binding interaction was observed.

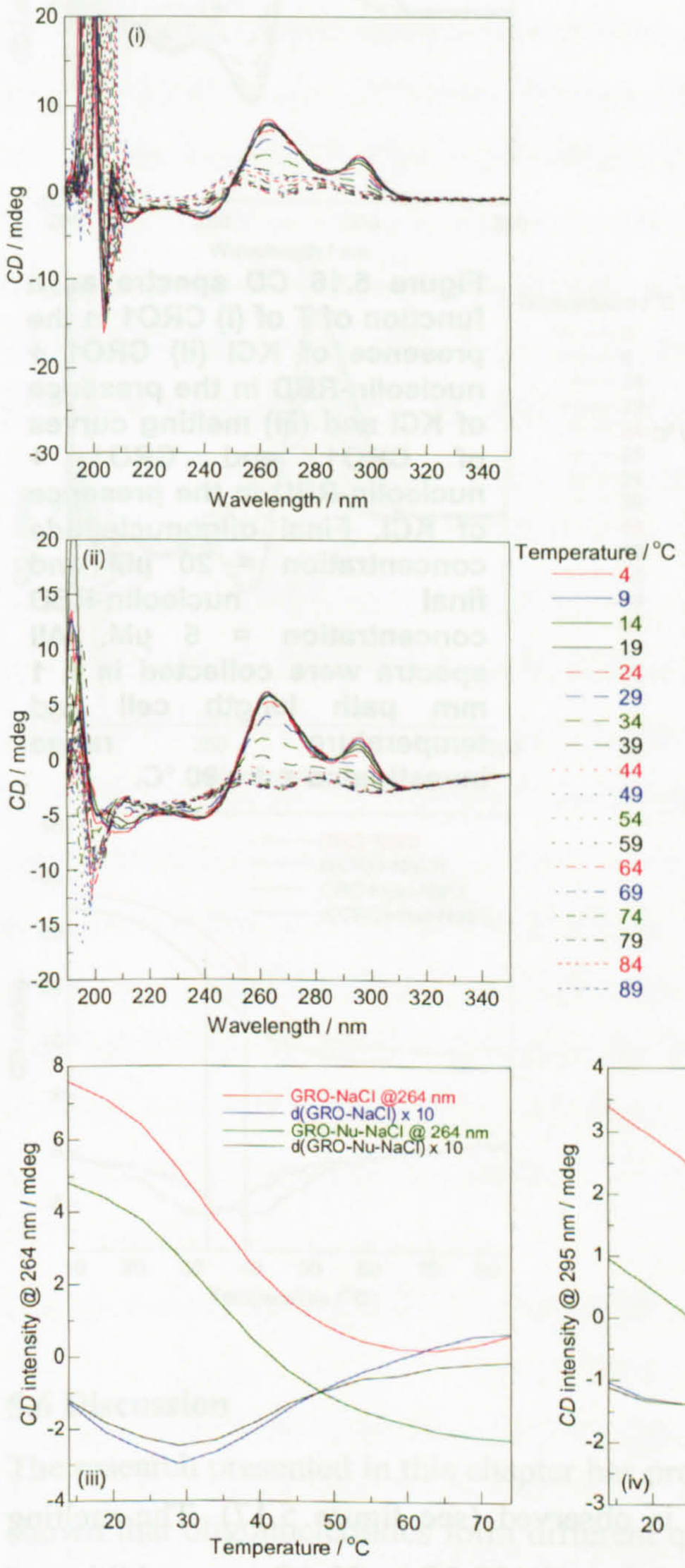



**Figure 5.15 *CD* spectra as a function of T of (i) GRO1 in the presence of NaCl, (ii) GRO1 + nucleolin-RBD in the presence of NaCl and iii, melting curves of GRO1 and GRO1 + nucleolin-RBD in the presence of NaCl at 264 nm and (iv) 295 nm. Final oligonucleotide concentration = 20 μM and final nucleolin-RBD concentration = 5 μM. All spectra were collected in a 1 mm path length cell and temperature range investigated = 4 – 90 °C.**

The *CD* spectra presented in figure 5.16 show the melting curves of CRO1 in the presence of $K^+$ and $Na^+$ and also upon addition of nucleolin-RBD. It was observed from the temperature gradient experiment (4 °C – 90 °C) that upon addition of nucleolin-RBD to the CRO1 the oligonucleotide is considerably destabilised. The melting

temperature decreases from approximately 48 °C to 35 °C upon addition of the nucleolin-RBD in the presence of KCl.

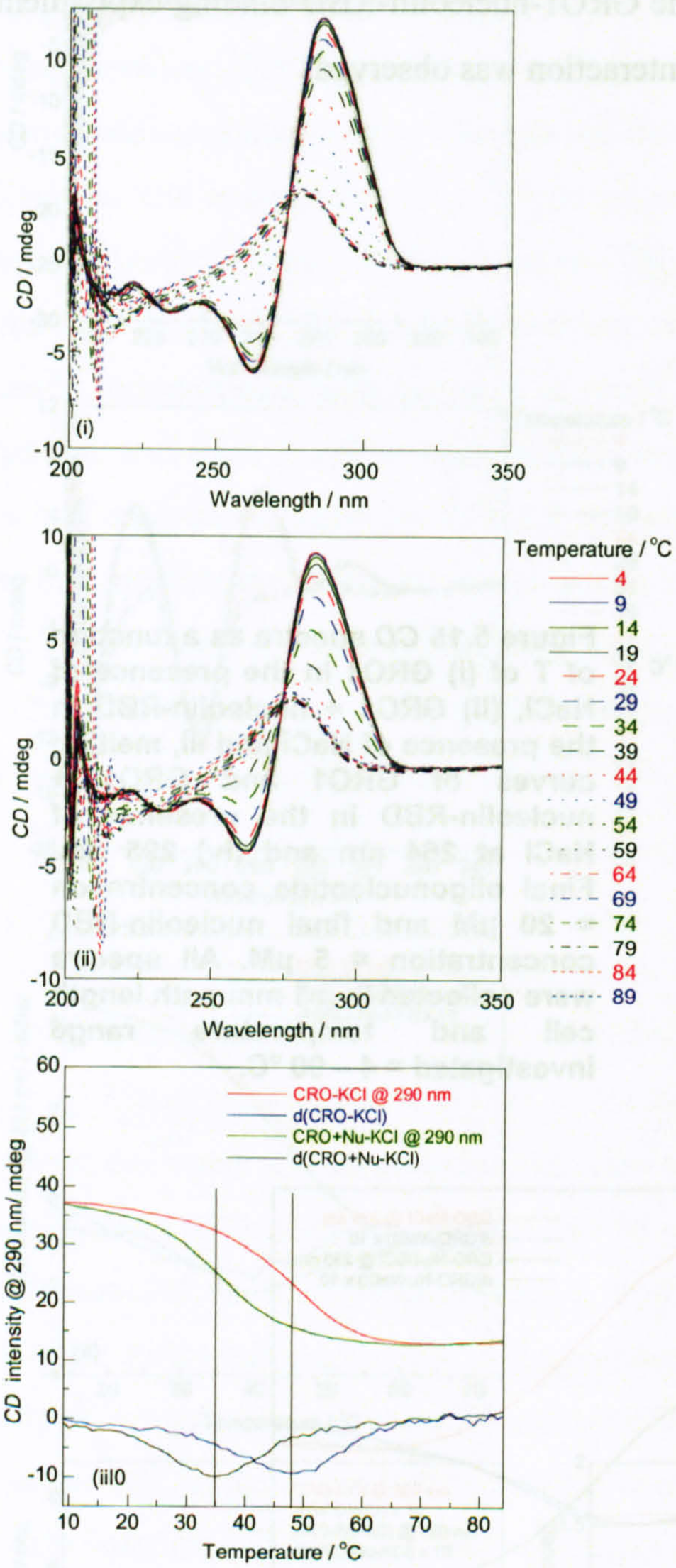



**Figure 5.16 *CD* spectra as a function of T of (i) CRO1 in the presence of KCl (ii) CRO1 + nucleolin-RBD in the presence of KCl and (iii) melting curves of CRO1 and CRO1 + nucleolin-RBD in the presence of KCl. Final oligonucleotide concentration = 20 μM and final nucleolin-RBD concentration = 5 μM. All spectra were collected in a 1 mm path length cell and temperature range investigated = 4 – 90 °C.**

In the presence of NaCl the same trend is observed (see figure 5.17). The melting temperature of CRO1 decreases from approximately 39 °C to 32 °C upon addition of the nucleolin-RBD. This result is intriguing as the results presented in § 5.5.3 suggests no interaction of CRO with nucleolin-RBD.

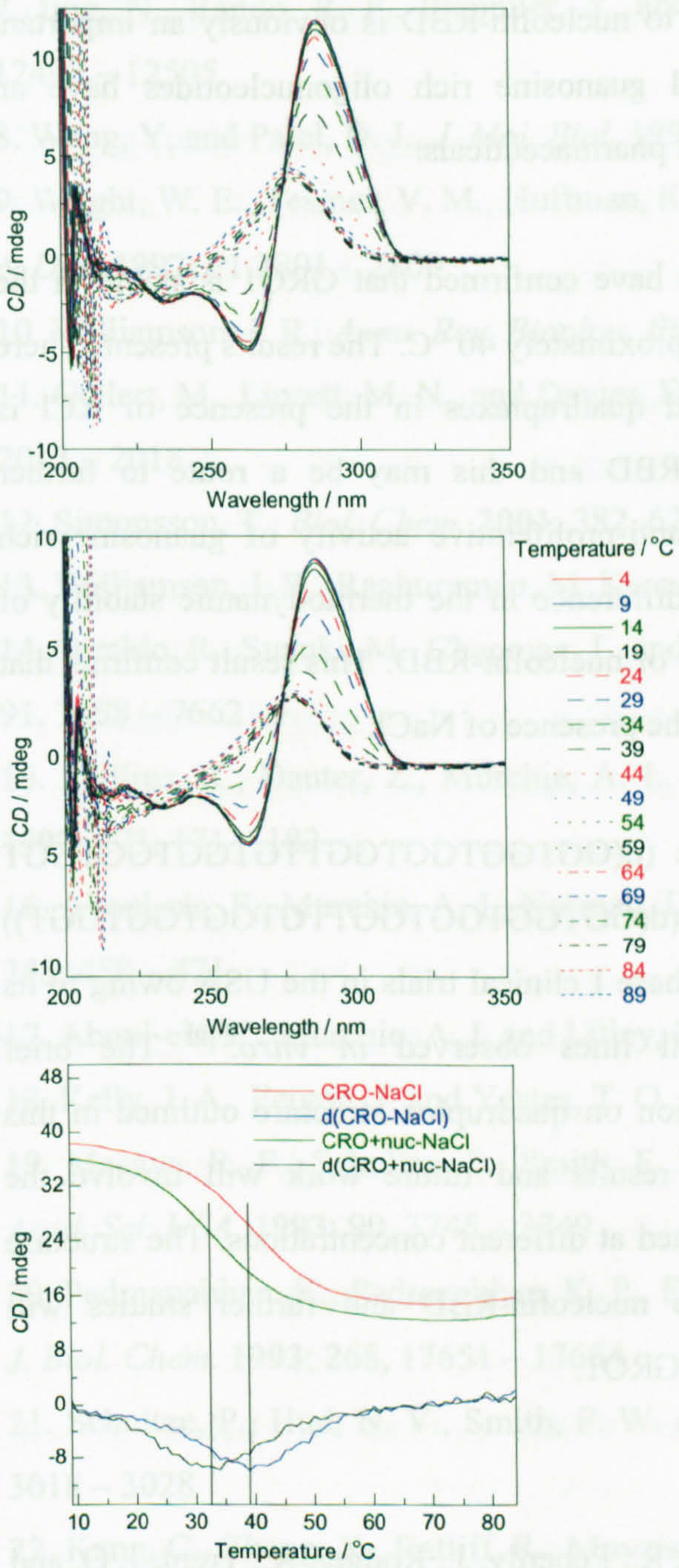



**Figure 5.17 *CD* spectra as a function of T (i) CRO1 in the presence of NaCl (ii) CRO1 + nucleolin-RBD in the presence of NaCl and (iii) melting curves of CRO1 and CRO1 + nucleolin-RBD in the presence of NaCl. Final oligonucleotide concentration = 20 μM and final nucleolin-RBD concentration = 5 μM. All spectra were collected in a 1 mm path length cell and temperature range investigated = 4 – 90 °C.**

## 5.6 Discussion

The research presented in this chapter has produced some interesting results. It has been shown that oligonucleotides form different quadruplex structures in the presence of $K^+$ or $Na^+$ and that biological activity cannot be directly related to any specific structural characteristics as observed in the *CD* spectrum. The oligonucleotide-nucleolin-RBD binding experiments have outlined the importance of the presence of $K^+$ in the ability of the nucleolin-RBD to interact with the guanosine-rich oligonucleotide and it was observed that no interactions occur in the presence of $Na^+$. This result has validated previous findings, in that there is increased biological/anti-proliferative activity in the

presence of KCl. The binding of GRO's to nucleolin-RBD is obviously an important factor governing biological activity and guanosine rich oligonucleotides have an important role in the development of novel pharmaceuticals.

The thermodynamic stability experiments have confirmed that GRO1 is stable in the presence of KCl and NaCl and melts at approximately 40 °C. The results presented here may indicate that the stability of folded quadruplexes in the presence of KCl is increased in the presence of nucleolin-RBD and this may be a route to further investigate any increased biological or anti-proliferative activity of guanosine rich oligonucleotides. There is no significant difference in the thermodynamic stability of GRO1 in NaCl in the presence or absence of nucleolin-RBD. This result confirms that GRO1 does not bind to nucleolin-RBD in the presence of NaCl.

A 26-mer guanosine-rich oligonucleotide (d(GGTGGTGGTGGTTGTGGTGGTGGT GG)) similar in sequence to the GRO1 (d(GGTGGTGGTGGTTGTGGTGGTGGT)) investigated in this work, is currently in Phase I clinical trials in the USA owing to its anti-proliferative activity on cancer cell lines observed *in vitro*.[40] The brief investigation into the effect of concentration on quadruplex structure outlined in this research has produced some interesting results and future work will involve the investigation of quadruplex structures formed at different concentrations. The structure of GRO1 changes upon binding to the nucleolin-RBD and further studies will investigate the anti-proliferative activity of GRO1.

**5.7 References**

1. Đapic, V., Abdomerović, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O. and Bates, P. J., *Nucleic Acids Res.*, **2003**; 31, 8, 2097 – 2107

2. Đapic, V., Bates, P. J., Trent, J. O., Rodger, A., Thomas, S. D. and Miller, D. M., *Biochemistry*, **2002**; 41, 3676 – 3685

3. Bates, P. J., Kahlon, J. B., Thomas, S. D., Trent, J. O. and Miller D. M., *J. Biol. Chem.*, **1999**; 274, 26369 – 26377

4. Xu, X., Hamhouyia, F., Thomas, S. D., Burke, T. J., Girvan, A. C., McGregor, W. G., Trent, J. O., Miller D. M. and Bates, P. J., *J. Biol. Chem.* **2001**; 276, 43221 – 43230

5. Hurley, L. H., *Biochem. Soc. Trans.* **2001**; 29, 6, 692 – 696

6. Petraccone, L., Erra, E., Esposito, V., Randazzo, A., Mayol, L. Nasti, L., Barone, G. and Giancola, C., *Biochemistry*, **2004**; 43, 4877 – 4884

7. Jing, N., Rando, R. F., Pommier, Y. and Hogan, M. E., *Biochemistry,* **1997**; 36, 12498 – 12505

8. Wang, Y. and Patel, D. J., *J. Mol. Biol.* **1993**; 234, 1171 – 1183

9. Wright, W. E., Tesmer, V. M., Huffman, K. E., Levene, S. D. and Shay, S. W., *Genes & Dev.* **1997**; 11 2801 – 2809

10. Williamson, J. R., *Annu. Rev. Biophys. Biomol. Struct.* **1994**; 23, 703 – 730

11. Gellert, M., Lipsett, M. N., and Davies, D. R., *Proc. Natl. Acad. Sci. USA.* **1962**; 48, 2013 – 2018

12. Simonsson, T., *Biol. Chem.* **2001**; 382, 621 – 628

13. Williamson, J. R., Raghuraman, M. K. and Cech, T. R., *Cell*, **1989**; 59, 871 – 880

14. Giraldo, R., Suzuki, M., Chapman, L. and Rhodes, D., *Proc. Natl. Acad. Sci.* **1994**; 91, 7658 – 7662

15. Phillips, K., Dauter, Z., Murchie, A. I., Lilley, D. M. and Luisi, B., *J. Mol. Biol.* **1997**; 273, 171 – 182

16. Aboul-ela, F., Murchie, A. I., Norman, D. G. and Lilley, D. M., *J. Mol. Biol.* **1994**; 243, 458 – 471

17. Aboul-ela, F., Murchie, A. I. and Lilley, D. M., *Nature*, **1992**; 260, 280 – 282

18. Kelly, J. A., Feigon, J. and Yeates, T. O., *J. Mol. Biol.* **1996**; 256, 417 – 422

19. Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. and Feigon, J., *Proc. Natl. Acad. Sci. USA*, **1993**; 90, 3745 – 3749

20. Padmanabhan, K., Padmnabhan, K. P., Ferrara, J. D., Sadler, J. E. and Tulinsky, A., *J. Biol. Chem.* **1993**; 268, 17651 – 17654

21. Schultze, P., Hud, N. V., Smith, F. W. and Feigon J. *Nucleic Acids Res.* **1999**; 27, 3018 – 3028

22. Kang, C., Zhang, X., Ratliff, R., Moyzis, R. and Rich, A., *Nature*, **1992**; 356, 126 – 131

23. Smith, F. W. and Feigon, J., *Biochemistry*, **1993**; 32, 8682 – 8692

24. Haider, S., Parkinson, G. N. and Neidle, S., *J. Mol. Biol.*, **2002**; 320, 189 – 200

25. Leonard, G., Zhang, S., Peterson, M. R., Harrop, S. J., Helliwell, J. R., Cruse, W. B., d'Estaintot, B. L., Kennard, O., Brown, T. and Hunter, W. N., *Structure*, **1995**; 3, 335 – 340

26. Kettani, A., Kumar, R. A., and Patel, D. J., *J. Mol. Biol.* **1995**; 254, 638 – 656

27. Kettani, A., Bouaziz, S., Wang, W., Jones, R. A. and Patel, D. J., *Nature Struct. Biol.* **1997**; 4, 382 – 389

28. Scaria, P. V., Shire, S. J. and Shafer, R. H., *Proc. Natl. Acad. Sci.* **1992**; 89, 10336 – 10340

29. Keniry, M. A., Strahan, G. D., Owen, E. A. and Shafer, R. H., *Eur. J. Biochem.*, **1995**; 233, 631 – 643

30. Hud, N. V., Smith, F. W., Anet, F. A. and Feigon, J., *Biochemistry*, **1996**; 35, 15383 – 15390

31. Wang, Y and Patel, D. J., *J. Mol. Biol.* **1995**; 251, 76 – 94

32. Smith, F. W. and Feigon, J. *Nature*, **1992**; 356, 164 – 168

33. Wang, Y. and Patel, D. J., *Structure*, **1994**; 2, 1141 – 1156

34. Jing, N. and Hogan, M. E., *J. Biol. Chem.* **1998**; 273, 34992 – 34999

35. Marathias, V. M. and Bolton, P. H., *Biochemistry*, **1999**; 38, 4355 – 4364

36. Parkinson, G. N., Lee, M. P. H. and Neidle S. *Nature*, **2002**; 417, 876 – 880

37. Horvath, M. P. and Schultz, S. C., *J. Mol. Biol.* **2001**; 310, 367 – 377

38. Smith, F. W., Schultze, P. and Feigon J. *Structure*, **1995**; 3, 997 – 1008

39. Wang, Y. and Patel, D. J., *Structure*, **1997**; 1, 263 – 282

40. Bates, P. J. and Trent, J. O., *'personal communication'* **2004**

41. Rodger, A. and Nordén, B. *'Circular Dichroism and Linear Dichroism'*, Oxford Chemistry Press, Oxford, **1997**;

42. Balagurumoorthy, P. and Brahmachari, S. K., *J. Biol. Chem.* **1994**; 269, 21858 – 21869

43. Hardin, C. C., Henderson, E., Watson, T. and Prosser, J. K. *Biochemistry,* **1991**; 30, 4460 – 4472

44. Balagurumoorthy, P. and Brahmachari, S. K., Mohanty, D., Bansal, M. and Sasisckharan, V., *Nucleic Acids Res.*, **1992**; 20, 4061 – 4067

45. Rhodes, D., Fairall, L. Simonsson, T., Court, R. and Chapman, L., *EMBO Rep.* **2002**: 3, 1139 – 1145

46. Rezler, E. M., Bearss. D. J., Hurley, L. H., *Annu. Rev. Pharmacol. Toxicol.* **2003**; 43, 359 – 379

47. Kanoh, J., Ishikawa, F., *Cell. Mol. Life Sci.* **2003**; 60, 2295 – 2302

48. Nelson, S. M., Ferguson, L. R., Denny, W. A., *Cell & Chromosome*, **2004**; 3, 1 – 26

49. Simonson, T. Pecinka, P. and Kubista, M. *Nucleic Acids Res.* **1998**; 26, 5 1167 – 1172

50. Newbold, R. F., *Anticancer Drug. Des.* **1999**; 14, 349 – 354

51. Saretzki, G., *Cancer Lett.* **2003**; 194, 209 – 219

52. Rezler, E. M., Bearss, D. J., Hurley, L. H., *Curr. Opin. Pharmacol.* **2002**; 2, 415 – 423

53. Shammas, M. A., Shmookler Reis, R. J., Aklyama, M. Koley, H., Chauhan, D., Hideshima, T., Goyal, R. K., Hurley, L. H., Anderson K. C. and Munshi, N. C., *Mol. Cancer Ther.* **2003**; 2, 825 – 833

54. Gowan, S. M., Harrison, J. R., Patterson, L., Valenti, M., Read, M. A., Neidle, S. and Kelland, L. R., *Mol. Pharmacl.* **2002**; 61, 1154 – 1162

55. Tanious, F. A., Jenkins, T. C., Neidle, S. and Wilson, W. D., *Biochemistry*, **1992**; 31, 11632 – 11640

56. Chiang, S. Y., Welch, J., Rauscher, F. J., 3rd, Beerman, T. A., *Biochemistry,* **1994**; 33, 7033 – 7040

57. Han, H. and Hurley, L. H., *Trends Pharmacol. Sci.* **2000**; 21, 136 – 142

58. Sun, D., Thompson, B., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., Jenkins, T. C., Neidle, S. and Hurley, L. H., *J. Med. Chem.* **1997**; 40, 2113 – 2116

59. Nelson, D. C. and Cox, M. M. '*Lehninger Principles of Biochemistry*' Third Ed. Worth Publishers, **2000**; 163 – 168

60. Pauling, L. Corey, R. B. and Branson, H. R., *Proc. Natl. Acad. Sci. USA*, **1951**; 37, 205 – 211

61. Pauling, L., and Corey, R B., *Proc. Natl. Acad. Sci. USA*, **1951**; 37, 729 – 740

62. Branden, C. and Tooze, J., '*Introduction to Protein Structure*' Second Ed. Garland Publishing **1999**; 176 – 184

63. Tuteja, R. and Tuteja, N. *Crit. Rev. Biochem. Mol. Biol.* **1998**; 33, 407 – 436

64. Lapeyre, B., Bourbon, H. and Amalric, F., *Proc. Natl. Acad. Sci. USA,* **1987**; 84, 1472 – 1476

65. Bugler, B., Bourbon, H. M., Lapeyre, B., Wallace, M. O., Chang, J. H., Amalric, F. and Olson, M. O. J., *J. Biol. Chem.* **1987**; 262, 10922 – 10925

66. Olson, M. O. J., Rivers, Z. M., Thompson, B. A., Kao, W.-Y. and Case, S. T., *Biochemistry*, **1983**; 22, 3345 – 3351

67. Martin, M., Garcia-Fernandez, L. F., Moreno-Diazdela-Espina, S., Noaillac-Depeyre, J., Gas, N. and Medina, F. J., *Exp. Cell. Res.*, **1992**; 199, 74 – 84

68. Schwab, G., Gossweiler, U. and Dreyer, C., *Exp. Cell. Res.*, **1998**; 239, 226 – 234

69. Derenzini, M., Sirri, V., Trere, D. and Ochs, R. L., *Lab. Invest.* **1995**; 73, 497 – 502

70 Roussel, P. and Hernandez-Verdun, D. *Exp. Cell Res.* **1994**; 214, 465 – 472

71. Allain, F. H.-T., Gilbert, D. E., Bouvet, P. and Feigon, J., *J. Mol. Biol.* **2000**; 303, 227 – 241

72. Erard, M., Belenguer, P., Caizergues-Ferrer, M., Pantaloni, A. and Amalric, F., *Eur. J. Biochem.* **1988**; 175, 525 – 530

73. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. and Bairoch, A., *Nucleic Acids Res.* **2003**; 31, 3784 – 3788

74. Bradford, M. *Anal. Biochem.* **1976**; 72, 248

75. Sedmack, J. J. and Grossberg, S. E., *Anal. Biochem.* **1977**; 79, 544

76. Manvalan, P. and Johnson, W. C. Jnr. *Anal. Biochem.* **1987**; 167, 76 – 85

77. Johnson, W. C., *Proteins: Struct. Funct. Genet.* **1999**; 35, 307 – 312

# CHAPTER 6 – CONCLUSIONS AND FUTURE WORK

Supramolecular chemistry has provided a scaffold for the design and synthesis of an exciting range of bimetallo iron cylinders. The main aims of this research were to develop a method for the enantiomeric separation of a range of bimetallo iron cylinders and introduce sequence selective substituents onto the backbone of the parent cylinder to achieve DNA-binding specificity. Cellulose chromatography was used to efficiently purify and enantiomerically separate a range of these cylinders and the optimum mobile phase was found to be 90% MeCN:10% 0.02 M NaCl. A 'universal' method for the separation of all cylinders has not yet been achieved, but future work may involve the use of chiral HPLC.

The DNA binding modes of the enantiomers of two methylated cylinders ($L_3$ and $L_5$) were studied and all were found to bind to DNA inducing structural changes in the DNA conformation. Circular and linear dichroism were used to probe the binding of the methylated cylinders to DNA and signals in the MLCT region of the spectrum and a decrease in the *LD* signal intensity in the DNA region of the spectrum confirmed a binding interaction. A competitive binding assay using the fluorescent dye ethidium bromide confirmed that the enantiomers of the parent cylinder and the methylated cylinders bind strongly to DNA, displacing the ethidium bromide and so have binding constants in excess of $5.4 \times 10^5$ $M^{-1}$. The ultimate aim of the research presented in this thesis was to introduce sequence selective substituents onto the backbone of the parent cylinder enabling the fine tuning of DNA-binding. The results from the binding studies are summarised in figure 6.1 and table 6.1 for comparison and show that some sequence preference has been achieved. The different spectral shape of the plots of figure 6.1 show that the enantiomers of $L_3$ and $L_5$ bind differently to the DNA. The spectral shape is different for the enantiomers of the $L_3$ metal complex confirming that these enantiomers bind via different modes to the DNA. The (R)-$L_3$ and (P)-$L_3$ cylinders shows no sequence preference as the spectral shape upon binding to ct-, AT- and GC-DNA are the same but the (M)-$L_3$ cylinder appears to have a preference for the regular alternating homopolymers and may target AT-rich sequences. The plots of figure 6.1 (b), (d) and (f) show very similar spectral shape, suggesting that the enantiomers of the $L_5$ metal complex show no sequence preference and that the enantiomers bind in a manner analogous to that of the racemic metal complex, i.e. that they all target the same site on the DNA.

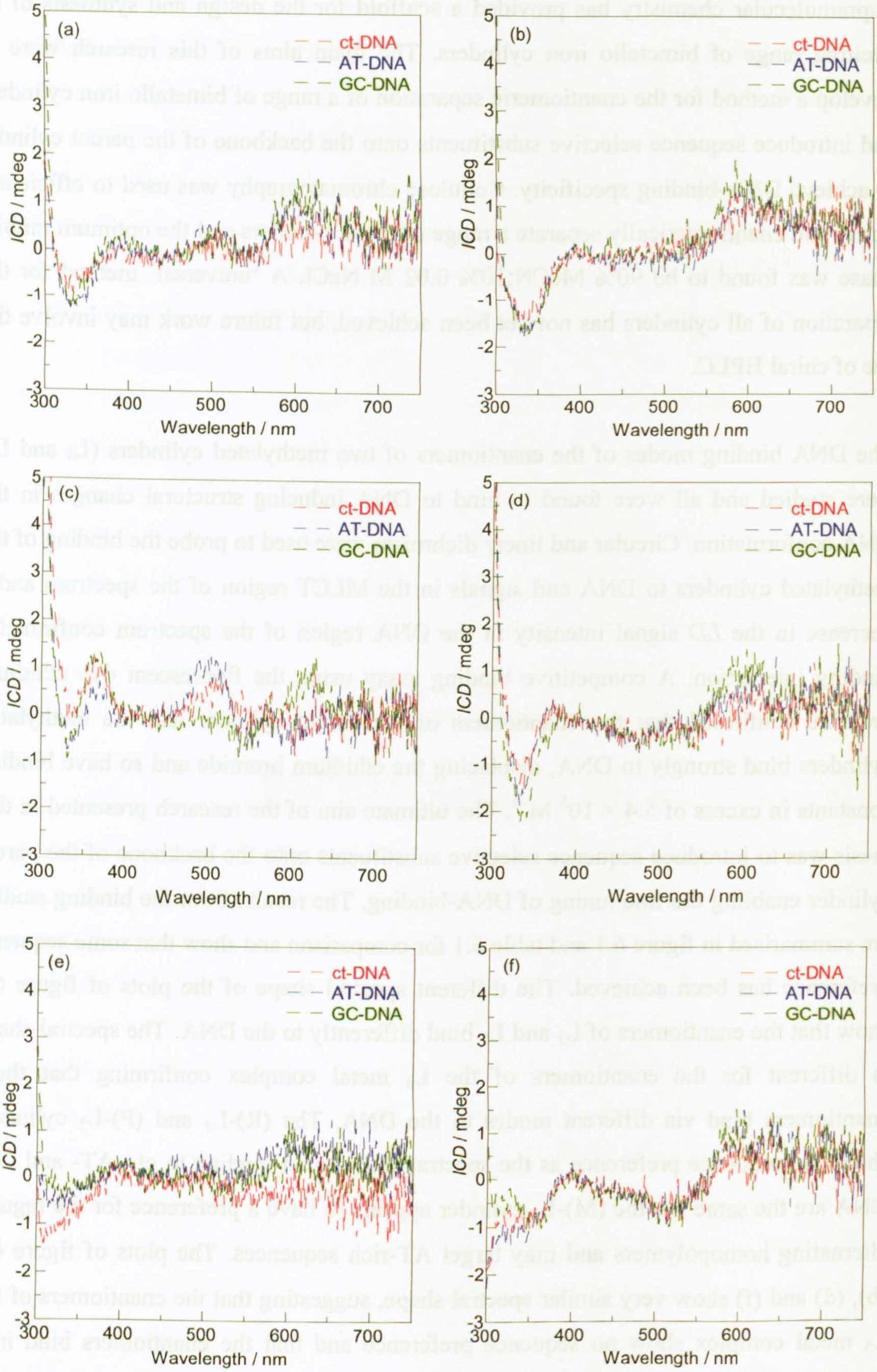



Figure 6.1 *ICD* spectra of (a) (R)-$L_3$ (b) (R)-$L_5$ (c) (M)-$L_3$ (d) (M)-$L_5$ (e) (P)-$L_3$ and (f) (P)-$L_5$ upon binding to ct-, AT- and GC-DNA at constant DNA (400 μM) and constant metal complex (20 μM) concentration. Spectra collected in a 1 cm path length cuvette and water subtracted from each sample spectrum. The spectra were zeroed in a region of no absorbance at 420 nm for the racemic metal complex and enantiomers.

AFM and NMR studies as well as fully solvated molecular dynamics simulations would provide valuable insights into the different binding modes of the enantiomers. Further investigation of the functional groups on zinc finger protein motifs responsible for sequence specific interactions with DNA, may be the route to follow in achieving sequence specific functionalities on the backbone of the parent cylinder

| [MC] | ct-DNA | AT-DNA | GC-DNA | Conclusions |
|---|---|---|---|---|
| (R)-$L_3$ | ✓ | ✓ | ✓ | No sequence preference |
| (M)-$L_3$ | ✓ | × | ✓ | Some sequence preference |
| (P)-$L_3$ | ✓ | ✓ | ✓ | No sequence preference |
| (R)-$L_5$ | ✓ | ✓ | ✓ | No sequence preference |
| (M)-$L_5$ | ✓ | ✓ | ✓ | No sequence preference |
| (P)-$L_5$ | ✓ | ✓ | ✓ | No sequence preference |

**Table 6.1 Summary of results obtained from the *ICD* binding experiments. ✓ indicates significant similarity between *ICD* spectra upon binding to different DNA's. × indicates no similarity between *ICD* spectra upon binding to different DNA's and so some sequence preference.**

A series of competitive binding experiments were designed to investigate the binding geometries of the metal complexes to DNA with the hope of identifying the site of interaction of each of the enantiomers. The results are summarised in table 6.2. At this stage in the research we cannot unequivocally relate the results from the competitive binding assays with the binding site of each of the enantiomers because, as a consequence of the high binding ability of the $L_1$, $L_3$ and $L_5$ cylinders, EB, Hoechst 33258 and cobalt (III) hexaammine all show dissociation from their binding site upon DNA-binding of the metal complexes.

**Summary of $L_1$ binding geometry**

NMR studies have confirmed that the (M)-enantiomers of the parent cylinder bind in the major groove of B-DNA. The results presented here (see table 6.2) confirm that the (M)-$L_1$ enantiomers bind strongly to ct-DNA and induce dramatic structural effects, coiling the DNA. An intriguing result is observed from the Hoechst 33258 competitive binding assay wherein the (M)-$L_1$ enantiomers are the least effective at displacing the Hoechst. By comparing the results of the cobalt hexamine competitive binding assay to the Hoechst 33258 assay, a possible explanation for this result is that the (M)-$L_1$ enantiomers may in fact be binding strongly via the major groove and even upon addition of concentrated $[Co(NH_3)_6]^{3+}$ very little B → Z-DNA transition occurs as a consequence of the higher binding affinity of the (M)-enantiomer. The results presented here suggest that the (P)-$L_1$ enantiomers are binding via the major groove. The (P)-

enantiomers bind very strongly to ct-DNA inducing dramatic structural effects. The (P)-enantiomers are the most effective at displacing the EB and the Hoechst 33258 and induce a significant loss in *LD* signal at 259 nm (see table 6.2).

**Summary of $L_3$ binding geometry**

At this stage in the research circular and linear dichroism and various competitive binding assays have been used to investigate the binding geometries of the $L_3$ enantiomers to DNA. It has been shown that the enantiomers of this metal complex bind least strongly to DNA and induce the least significant % loss in *LD* signal at 259 nm and hence coiling of the DNA (see table 6.2). The results of the Hoechst 33258 and cobalt hexamine competitive binding assays show that the (M)-enantiomers are quite effective at displacing the Hoechst 33258, but do not prevent the B → Z-DNA transition. A possible explanation for this result is that the (M)-enantiomers are binding in or via the major groove and that due to the lower binding constant of the $L_3$ cylinder addition of the cobalt hexamine to the GC-DNA-(M)-$L_3$ solution converts B-DNA to Z-DNA. The (P)-enantiomers bind least strongly to ct-DNA and induce the smallest structural effect in the DNA conformation (see table 6.2). The (P)-$L_3$ enantiomers are not very effective at displacing the Hoechst 33258 and do not prevent the B → Z-DNA transition upon addition of cobalt hexamine. This may indicate that the (P)-enantiomers are binding in the major or minor groove, but due to the fact that they have a lower binding constant than the other metal complexes investigated, other techniques such as NMR will need to be adopted to investigate the site of binding.

**Summary of $L_5$ binding geometry**

The enantiomers of the $L_5$ metal complex have binding constants similar to that of the parent compound and induce structural effects in the DNA conformation comparable to that of the enantiomers of the parent cylinder. The results presented here (see table 6.2) confirm that the (M)-$L_5$ enantiomers bind strongly to ct-DNA and induce dramatic structural effects, coiling the DNA. An intriguing result is observed from the Hoechst 33258 competitive binding assay wherein the (M)-$L_5$ enantiomers are not very effective at displacing the Hoechst. By comparing the results of the cobalt hexamine competitive binding assay to the Hoechst 33258 assay, a possible explanation for this result is that the (M)-$L_5$ enantiomers may in fact be binding strongly via the major groove and even upon addition of concentrated $[Co(NH_3)_6]^{3+}$ very little B → Z-DNA transition occurs as

a consequence of the higher binding affinity of the (M)-enantiomer. The binding constant of the (P)-enantiomers is relatively low, but the (P)-enantiomers are very effective at displacing the Hoechst 33258 (see table 6.2). This may indicate that the (P)-enantiomers are binding in or via the major groove.

| | **Weakest** | | | | | | | | **Strongest** |
|---|---|---|---|---|---|---|---|---|---|
| **EB DISPLACEMENT** | P3 | R3 | M3 | P5 | M1 | R5 | M5 | R1 | P1 |
| **% LOSS** | P3 | R3 | M3 | R5 | R1 | P5 | M5 | M1 | P1 |
| **HOECHST 33258** | M1 | M5 | P3 | R5 | R3 | M3 | P5 | R1 | P1 |
| **COBALT HEXAMMINE** | R3 | R5 | P5 | M3 | P3 | P1 | R1 | M1 | M5 |

**Table 6.2 Summary of the order of effectiveness of EB displacement, % loss at 259 nm (from the *LD* binding studies), Hoechst 33258 displacement and order of effectiveness of preventing the conversion from B → Z-DNA upon addition of $[Co(NH_3)_6]^{3+}$ for the different bimetallo iron cylinders.**

The overall conclusions from this research is that the (M)- and (P)-enantiomers of all of the metal complexes investigated bind strongly and target different sites on the DNA. It is quite probable that the enantiomers have a preferred binding site on the DNA e.g. the major groove, but once the site is filled alternative binding sites are adopted. As the enantiomers are eluted from the column, the sodium chloride mobile phase dilutes the metal complex concentration, so until a method can be developed to efficiently remove the sodium chloride or a technique developed to grow DNA-metal complex crystals, *CD* and *LD* give valuable insights into the modes of binding but cannot be used to unequivocally confirm the binding geometries of the enantiomers of the metal complexes.

High temperature docking calculations were used to give an insight into the preferred starting configuration for each of the enantiomers to a B-DNA dodecamer. The results in general followed that of the experimental findings, in that the lowest configurational energy was observed when the cylinder was located in the major groove. An intriguing result was observed for the (M)-$L_5$ cylinder wherein the most favourable starting configuration with the lowest configurational energy was observed in the minor groove. The results at this stage in the research provide a qualitative analysis of the preferred starting configurations for the fully solvated systems, wherein water is added to the system and the DNA is allowed to relax. In future experiments it will be necessary to further energy minimise several of the lowest energy configurations from the docking calculations.

It can be concluded from the results presented in chapter 5 that oligonucleotides form different quadruplex structures in the presence of $K^+$ or $Na^+$ and that biological activity cannot be directly related to any specific structural characteristics observed in the *CD* spectrum. It was found that nucleolin-RBD (RNA binding domain) binds to GRO1, a 24-mer guanosine-rich oligonucleotide (d(GGTGGTGGTGGTTGTGGTGGTGGT)) in the presence of KCl but not in the presence of NaCl and that no interactions occur between CRO1, the corresponding 24-mer cytosine-rich oligonucleotide, and nucleolin-RBD in the presence of KCl or NaCl. A 26-mer GRO similar in sequence to GRO1 is currently in Phase I clinical trials in the USA owing to its anti-proliferative activity on cancer cell lines observed *in vitro* and future studies will investigate the anti-proliferative activity of GRO1 and the effect of concentration of quadruplex structure.

A range of analytical techniques have been used to investigate different biomacromolecule-ligand interactions where the major aim of the work has been to probe the binding of different ligands to DNA. The results presented in this thesis show that *CD* and *LD* can be used to probe the interactions of the metal complexes to DNA and the binding of proteins to oligonucleotides.

# APPENDIX I

## Appendix I

Characterisation data

**$L_1$.** Anal. Calc. for $[Fe_2(C_{25}H_{20}N_4)_3][Cl_4].(H_2O)_{2.5}$: C, 63.1; H, 4.6; N, 11.8%. Found: C, 63.0; H, 4.5 ; N, 11.6%. Positive-ion ESI (MeOH): m/z ($[Fe_2(L^1)_3(Cl)_3]^+$), 1346.4 ($[Fe_2(L^1)_3(Cl)_2]^{2+}$), 655.7, ($[Fe_2(L^1)_3(Cl)]^{3+}$), 425.3 ($[Fe_2(L^1)_3]^{4+}$), 310.0. NMR ((MeOD) 400 MHz, 300 K): δ4.08 (2H, s, $CH_2$), 5.62 (4H, s, $H_{Ph}$), 7.08 (4H, s, $H_{Ph}$), 7.49 (2H, d, $H_6$ J = 4.2 Hz), 7.90 (2H, t, $H_5$ J = 6.1 Hz), 8.51 (2H, t, $H_4$ J = 6.5 Hz), 8.75 (2H, d, $H_3$ J = 7.0 Hz), 9.21 (2H, s, $H_{im}$). Selected IR data ($cm^{-1}$): 3737w, 3378w, 2360s, 2342s, 2181m, 2148w, 2028w, 1996w, 1977w, 1654m, 1560m, 1474m, 1438m, 1365m, 1267m, 1175w, 1040s, 954w, 900w, 796w, 740m, 674m.

**$L_2$.** Anal. Calc. for $[Fe_2(C_{27}H_{24}N_4)_3][Cl_4].(FeCl_2)_2$: C, 56.6; H, 4.2; N, 9.8%. Found: C, 56.6; H, 4.3 ; N, 9.7 %. Positive-ion ESI (MeOH): ($[Fe_2(L^2)_3]^{4+}$), 331.0; $^1$H NMR ((MeOD) 400 MHz, 300 K): δ4.10 (3H, s, Me), 4.64 (3H, s, Me), 4.79 (2H, dd, $H_{Ph}$ J = 8.3 Hz, J = 1.5 Hz), 5.57 (2H, dd, $H_{Ph}$ J = 8.3 Hz, J = 1.5 Hz), 6.89 (2H, dd, $H_{Ph}$ J = 6.3 Hz), 7.26 (2H, d, $H_6$ J = 6.3 Hz), 7.42 (2H, dd, $H_{Ph}$ J = 6.3 Hz, J = 1.1 Hz), 7.81 (2H, t, $H_5$ J = 6.3 Hz), 8.50 (2H, t, $H_4$ J = 7.8 Hz), 8.78 (2H, d, $H_3$ J = 7.8 Hz); Selected IR data ($cm^{-1}$): 3386w, 1626w, 1588w, 1559w, 1503s, 1474m, 1441m, 1380m, 1334m, 1308w, 1256m, 1166w, 1110w, 1060w, 1019w, 828vs, 771vs, 750vs, 691m, 674m.

**$L_4$.** Anal. Calc. for $[Fe_2(C_{27}H_{24}N_4)_3][PF_6]_4$: C, 51.1; H, 3.8; N, 8.8%. Found: C, 51.3; H, 3.7; N, 8.9%. Positive-ion ESI (MeCN): m/z ($[Fe_2(L^4)_3(PF_6)]^{3+}$), 489.4 ($[Fe_2(L^4)_3]^{4+}$), 330.9. $^1$H NMR (($CD_3CN$) 300 MHz, 300 K): δ2.79 (3H, s, Me), 4.01 (1H, s, Me), 4.66 (1H, bs, $H_{Ph}$), 5.48 (1H, bs, $H_{Ph}$), 6.92 (1H, bs, $H_{Ph}$), 7.02 (1H, d, $H_5$ J = 4.8 Hz), 7.31 (1H, bs, $H_{Ph}$), 7.75 (1H, bt, $H_4$ J = 5.7 Hz), 8.21 (1H, d, $H_3$ J = 6.8 Hz), 8.98 (1H, s, $H_{im}$); Selected IR data ($cm^{-1}$): 3650w, 2931w, 2350w, 2276w, 2046w, 2026w, 1976w, 1622m, 1585m, 1503s, 1478m, 1436m, 1414w, 1302m, 1259m, 1211w, 1172w, 1018m, 954m, 827vs, 740w, 674m.

**$L_6$.** Anal. Calc. for $[Fe_2(C_{29}H_{28}N_4)_3][PF_6]_4$: C, 52.5; H, 4.3; N, 8.5%. Found: C, 52.0; H, 4.2; N, 8.3%. Positive-ion ESI (MeOH): ($[Fe_2(L^6)_3]^{4+}$), 352.4. $^1$H NMR (($CD_3CN$) 300 MHz, 300 K): δ2.35 (3H, s, Me), 2.71 (3H, s, Me), 4.67 (1H, dd, $H_{Ph}$, J = 6.4 Hz, J = 1.9 Hz), 5.47 (1H, dd, $H_{Ph}$, J = 6.0 Hz, J = 2.0 Hz), 6.76 (1H, dd, $H_{Ph}$, J = 8.1 Hz, J = 1.8

Hz), 6.93 (1H, d, $H_{5/6}$, J = 5.6 Hz), 7.31 (1H, dd, $H_{Ph}$, J = 7.6 Hz, J = 1.8 Hz), 7.51 (1H, d $H_{5/6}$, J = 5.3 Hz), 8.45 (1H, s $H_3$); Selected IR data ($cm^{-1}$): 3386w, 1613w, 1590w, 1503s, 1475m, 1441m, 1379m, 1335m, 1310w, 1221w, 1166w, 1110w, 1042w, 1018w, 828vs, 773vs, 750vs, 691m, 674m.

**$L_7$**. Anal. Calc. for $[Fe_2(C_{24}H_{18}N_4O)_3][Cl]_4(H2O)_{12}$: C, 54.1; H, 4.5; N, 10.5%. Found: C, 54.5 ; H, 4.1 ; N, 10.0%. Positive-ion ESI (MeOH): ($[Fe_2(L^7)_3]^{4+}$), 311.0; $^1$H NMR (($D_2O$) 400 MHz, 300 K): δ5.32 (2H, broad d, $H_{Ph}$ J = 6.3 Hz), 5.86 (2H, broad d, $H_{Ph}$ J = 6.0 Hz), 6.38 (2H, bd, $H_{Ph}$ J = 5.1 Hz), 6.98 (2H, bd, $H_{Ph}$ J = 6.0 Hz), 7.25 (2H, d, $H_6$ J = 3.9 Hz), 7.57 (2H, t, $H_5$ J = 6.0 Hz), 8.23 (2H, t, $H_4$ J = 6.9 Hz), 8.43 (2H, d, $H_3$ J = 6.6 Hz), 9.01 (2H, s, $H_{im}$); Selected IR data ($cm^{-1}$): 1626w, 1591w, 1488vs, 1441w, 1357w, 1310w, 1227s, 1195s, 1158m, 1105w, 1043w, 1011w, 834vs, 774vs, 691w, 674w.

**$L_8$**. Anal. Calc. for $[Fe_2(C_{26}H_{22}N_4O)_3][PF_6]_4.1½FeCl_2$ : C, 44.6 H, 3.2; N, 8.0%. Found: C, 44.4; H, 3.4; N, 7.7%. Positive-ion ESI (MeCN): m/z ($[Fe_2(L^8)_3(PF_6)_3]^+$), 1765.5 ($[Fe_2(L^8)_3(PF_6)_2]^{2+}$), 810.2 ($[Fe_2(L^8)_3(PF_6)]^{3+}$), 491.9 ($[Fe_2(L^8)_3]^{4+}$), 332.8. $^1$H NMR (($CD_3CN$) 300 MHz, 300 K): δ2.31 (3H, s, Me), 2.44 (3H, s, Me) 4.88 (1H, d, $H_{Ph}$ J = 8.9 Hz), 5.62 (1H, d, $H_{Ph}$ J = 7.5 Hz), 6.57 (1H, d, $H_{Ph}$ J = 7.1 Hz), 7.10 (1H, d, $H_6$ J = 5.1 Hz), 7.24 (1H, d, $H_{Ph}$ J = 8.1 Hz), 7.69 (1H, t, $H_5$ J = 6.0 Hz), 8.38 (1H, t, $H_4$ J = 7.4 Hz), 8.62 (1H, d, $H_3$ J = 7.7 Hz); Selected IR data ($cm^{-1}$): 3381w, 1626w, 1588w, 1558w, 1503s, 1474m, 1441m, 1380m, 1335m, 1308w, 1256m, 1166w, 1110w, 1060w, 1019w, 827vs, 771vs, 750vs, 691m, 674m.

**$L_9$**. Anal. Calc. for $[Fe_2(C_{26}H_{22}N_4O)_3][Cl]_4 2FeCl_2.3H_2O$: C, 52.6; H, 4.1; N, 9.4%. Found: C, 52.9; H, 4.7; N, 8.5%. Positive-ion ESI (MeOH): m/z ($[Fe_2(L^9)_3(PF_6)_3]^+$), 1765.3 ($[Fe_2(L^9)_3(PF_6)_2]^{2+}$), 810.2 ($[Fe_2(L^9)_3(PF_6)]^{3+}$), 491.9 ($[Fe_2(L^9)_3]^{4+}$), 332.6. $^1$H NMR (($CD_3CN$) 400 MHz, 300 K): δ2.87 (3H, s, Me), 5.49 (1H, s, $H_{Ph}$), 5.94 (1H, s, $H_{Ph}$), 6.44 (1H, s, $H_{Ph}$), 7.04 (1H, d, $H_5$ J = 6.1 Hz), 7.22 (1H, s, $H_{Ph}$), 7.60 (1H, t, $H_4$ J = 7.2 Hz), 8.15 (1H, d, $H_3$ J = 7.0 Hz), 9.04 (1H, s, $H_{im}$). Selected IR data ($cm^{-1}$): 3356w, 1614w, 1590w, 1490vs, 1446w, 1378w, 1311w, 1231s, 1164m, 1108w, 1038w, 1010w, 833vs, 792s, 691w, 674m.

**$L_{10}$**. Anal. Calc. for $[Fe_2(C_{26}H_{22}N_4O)_3][PF_6]_4 \cdot 3½EtOH$: C, 49.0; H, 3.5; N, 8.8%. Found: C, 49.5; H, 4.1; N, 8.3%. Positive-ion ESI (MeCN): m/z ($[Fe_2(L^{10})_3(PF_6)_3]^+$),

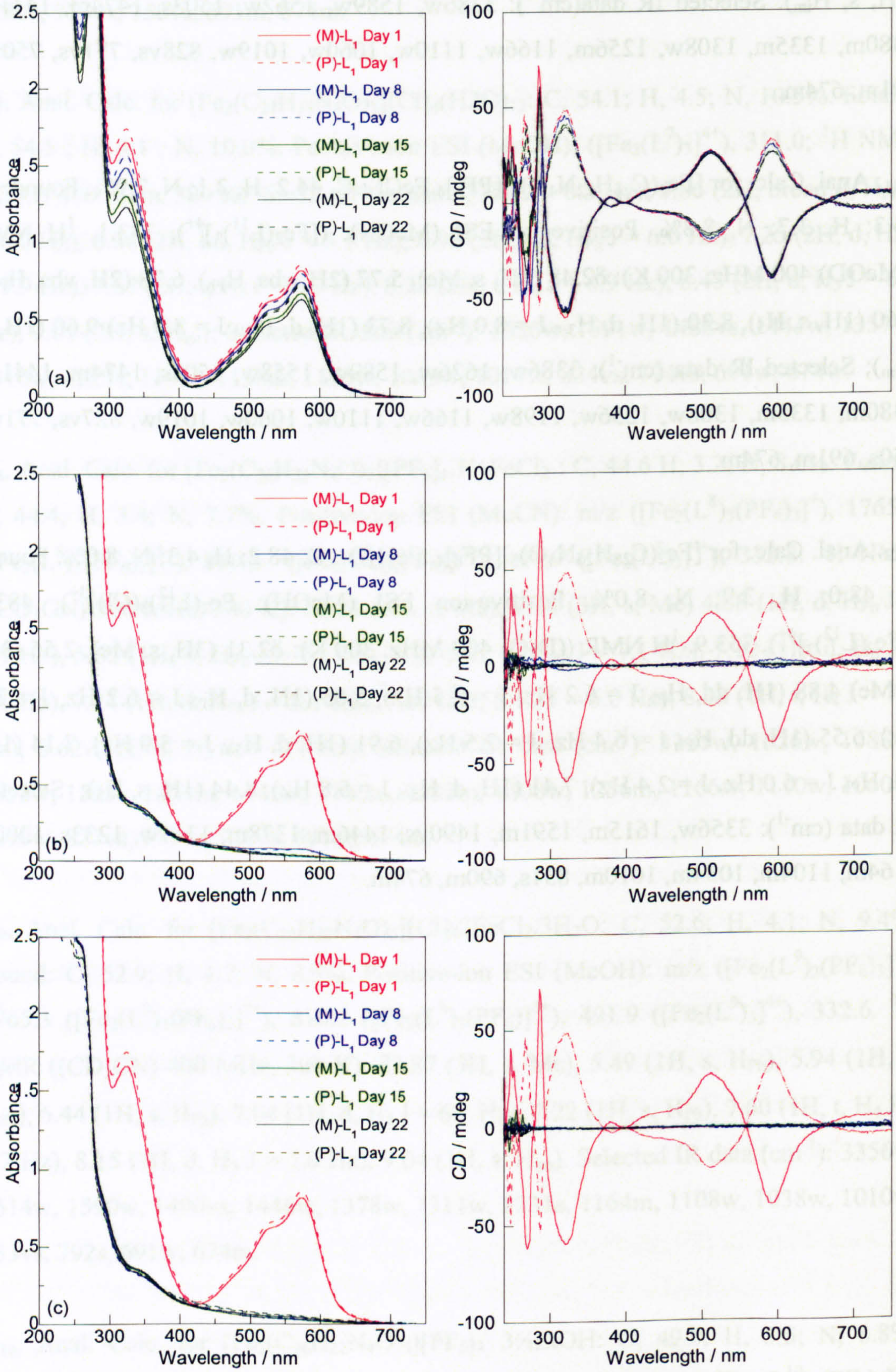
1768.5 ($[Fe_2(L^{10})_3(PF_6)_2]^{2+}$), 810.6 ($[Fe_2(L^{10})_3(PF_6)]^{3+}$), 492.2 ($[Fe_2(L^{10})_3]^{4+}$), 332.9. $^1$H NMR ((CD$_3$CN) 400 MHz, 300 K): δ2.70 (3H, s, Me), 5.70 (2H, vbs, $H_{Ph}$), 6.70 (2H, vbs, $H_{Ph}$), 7.44 (1H, s, $H_4$), 7.72 (1H, d, $H_6$ J = 5.9 Hz), 8.51 (1H, d, $H_3$ J = 7.8 Hz) 9.30 (1H, s, $H_{im}$). Selected IR data(cm$^{-1}$): 3386w, 1589w, 1562w, 1503s, 1474m, 1441m, 1380m, 1335m, 1308w, 1256m, 1166w, 1110w, 1060w, 1019w, 828vs, 771vs, 750vs, 691m, 674m.

**$L_{11}$.** Anal. Calc. for $[Fe_2(C_{26}H_{22}N_4O)_3][PF_6]_4.FeCl_2$: C, 44.2; H, 3.1; N, 7.9%. Found: C, 44.3; H, 3.7; N, 8.8%. Positive-ion ESI (MeCN): ($[Fe_2(L^{11})_3]^{4+}$), 333.1. $^1$H NMR ((MeOD) 400 MHz, 300 K): δ2.48 (3H, s, Me), 5.77 (2H, vbs, $H_{Ph}$), 6.78 (2H, vbs, $H_{Ph}$), 7.60 (1H, s, $H_5$), 8.30 (1H, d, $H_{3/4}$ J = 8.0 Hz), 8.73 (1H, d, $H_{3/4}$ J = 8.0 Hz) 9.60 (1H, s, $H_{im}$); Selected IR data (cm$^{-1}$): 3386w, 1626w, 1589w, 1558w, 1503s, 1474m, 1441m, 1380m, 1335m, 1308w, 1256w, 1198w, 1166w, 1110w, 1060w, 1019w, 827vs, 771vs, 750s, 691m, 674m.

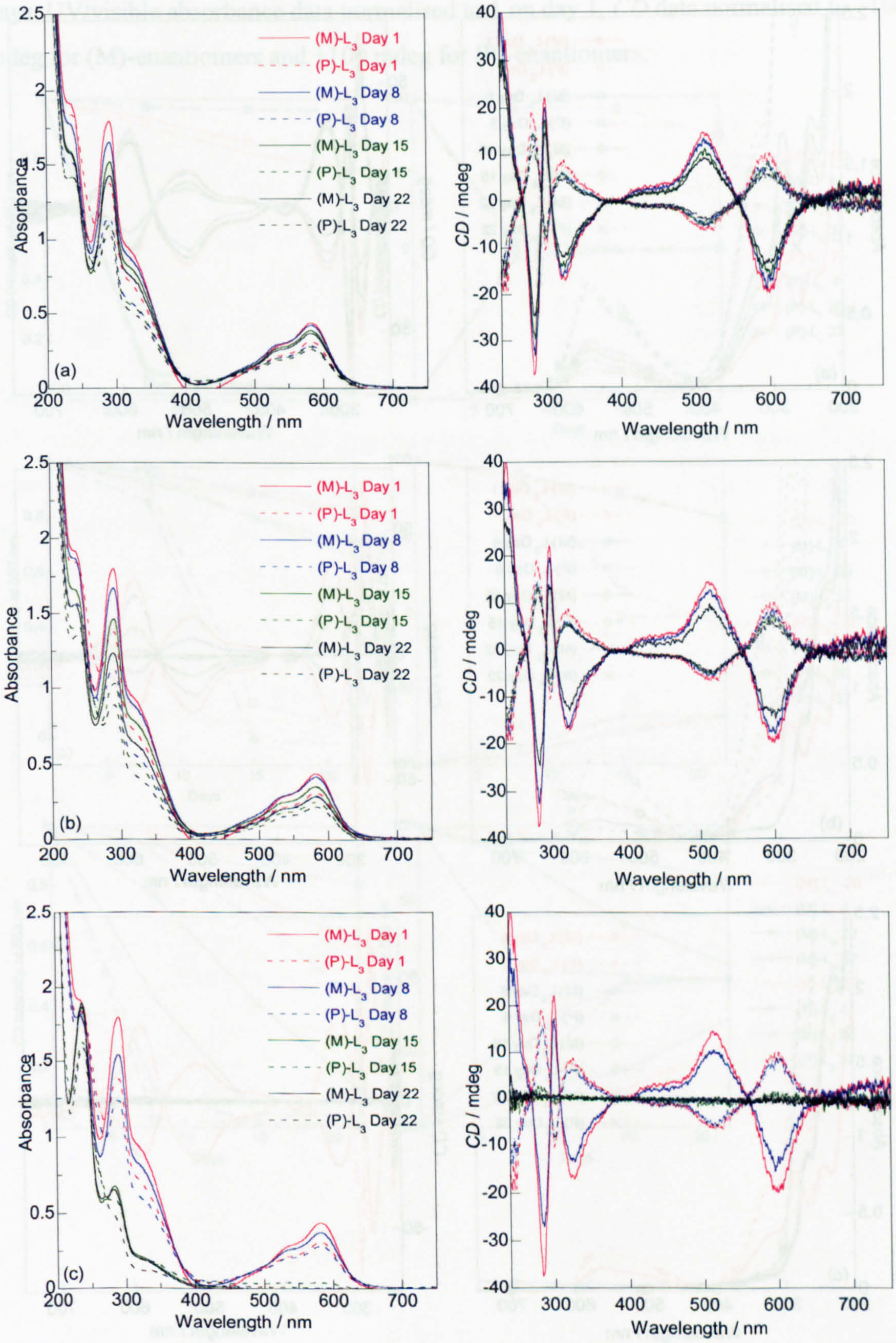
**$L_{12}$.** Anal. Calc. for $[Fe_2(C_{28}H_{26}N_4O)_3][PF_6]_4.5½H_2O$ : C, 48.2; H, 4.3; N, 8.0%. Found: C, 48.0; H, 3.9; N, 8.0%. Positive-ion ESI (MeOH): $Fe_2(L^{12})_3(Cl)]^{3+}$), 483.5 ($[Fe_2(L^{12})_3]^{4+}$), 353.9. $^1$H NMR ((D$_2$O) 400 MHz, 300 K): δ2.31 (3H, s, Me), 2.56 (3H, s, Me) 4.88 (1H, dd, $H_{Ph}$ J = 6.2 Hz, J = 2.5 Hz), 5.65 (1H, d, $H_{Ph}$ J = 6.2 Hz, J = 2.5 Hz), 6.55 (1H, dd, $H_{Ph}$ J = 6.4 Hz. J = 2.5 Hz), 6.91 (1H, d, $H_{5/6}$ J = 5.9 Hz), 7.14 (1H, dd, $H_{Ph}$ J = 6.0 Hz, J = 2.4 Hz), 7.41 (1H, d, $H_{5/6}$ J = 5.8 Hz), 8.44 (1H, s, $H_3$); Selected IR data (cm$^{-1}$): 3356w, 1615m, 1591m, 1490vs, 1446m, 1378m, 1311w, 1233s, 1202s, 1164m, 1104m, 1034m, 1010m, 831s, 690m, 674m.

Thermal stability experiments

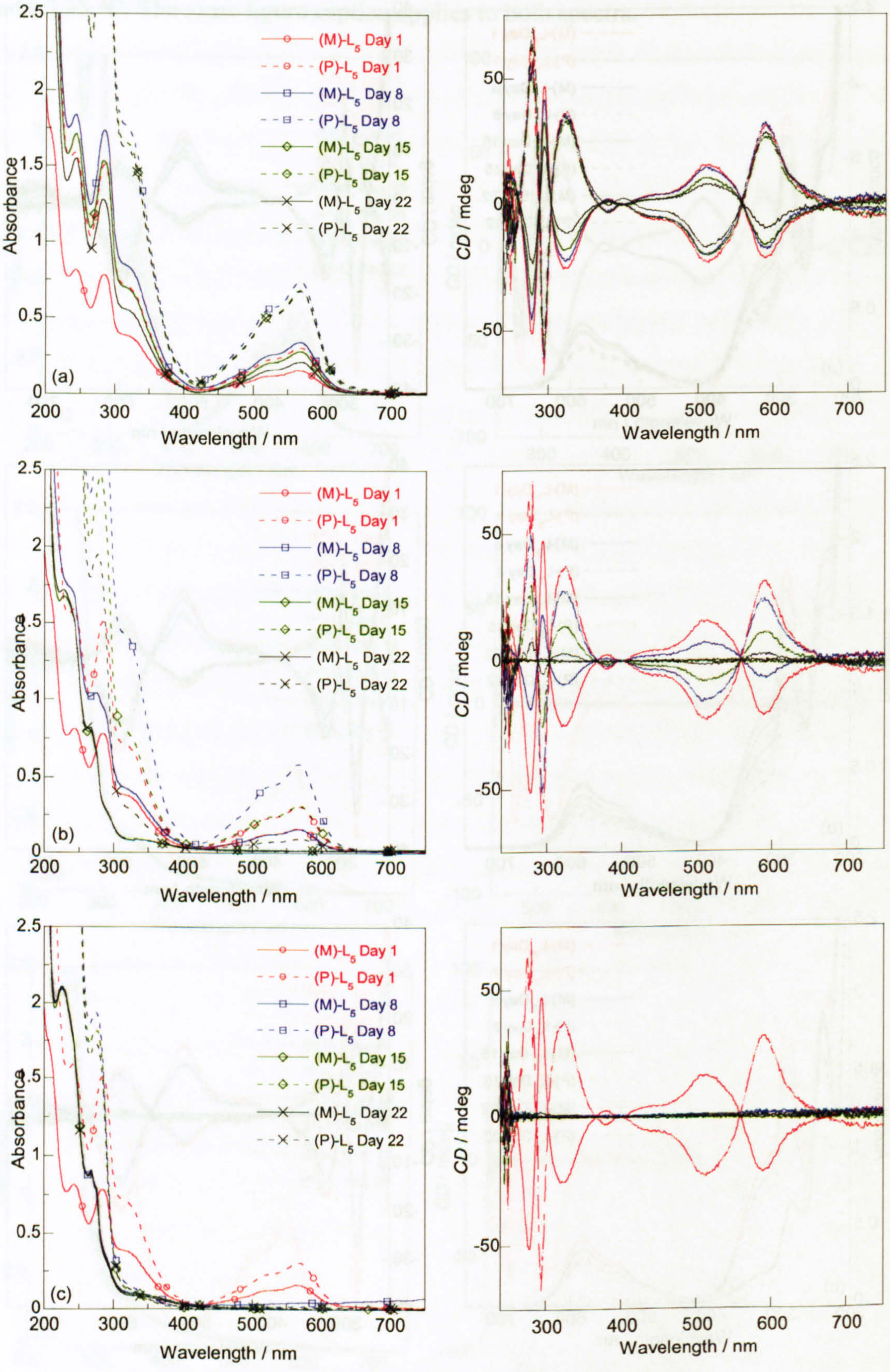
UV/visible absorbance and *CD* spectra of the enantiomers of $L_1$, at (a) –20 °C, (b) 20 °C and (c) 38 °C. The same figure caption applies to both spectra.



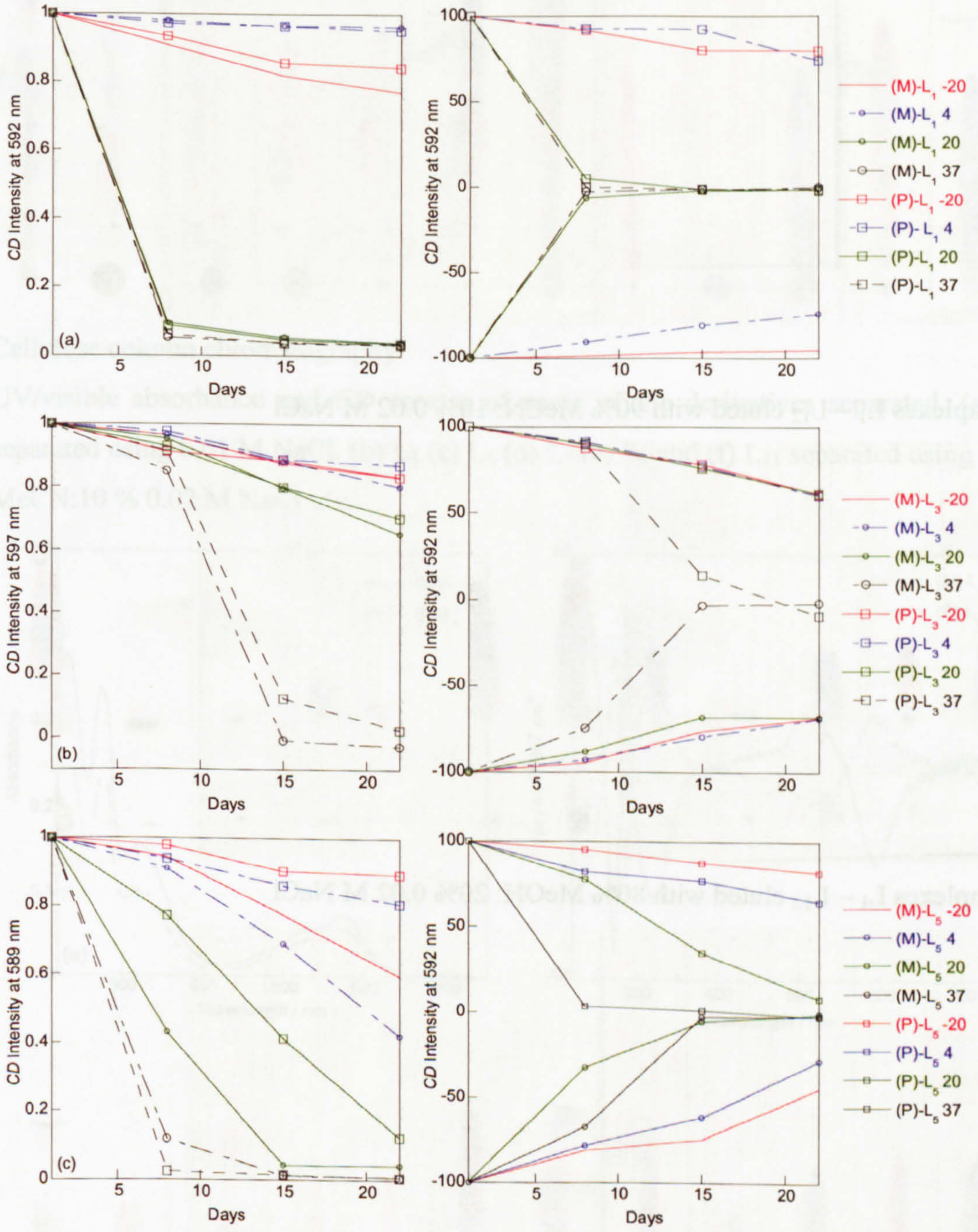
UV/visible absorbance and *CD* spectra of the enantiomers of $L_3$, at (a) –20 °C, (b) 20 °C and (c) 38 °C.



UV/visible absorbance and *CD* spectra of the enantiomers of $L_5$, at (a) –20 °C (b) 20 °C and (c) 38 °C.

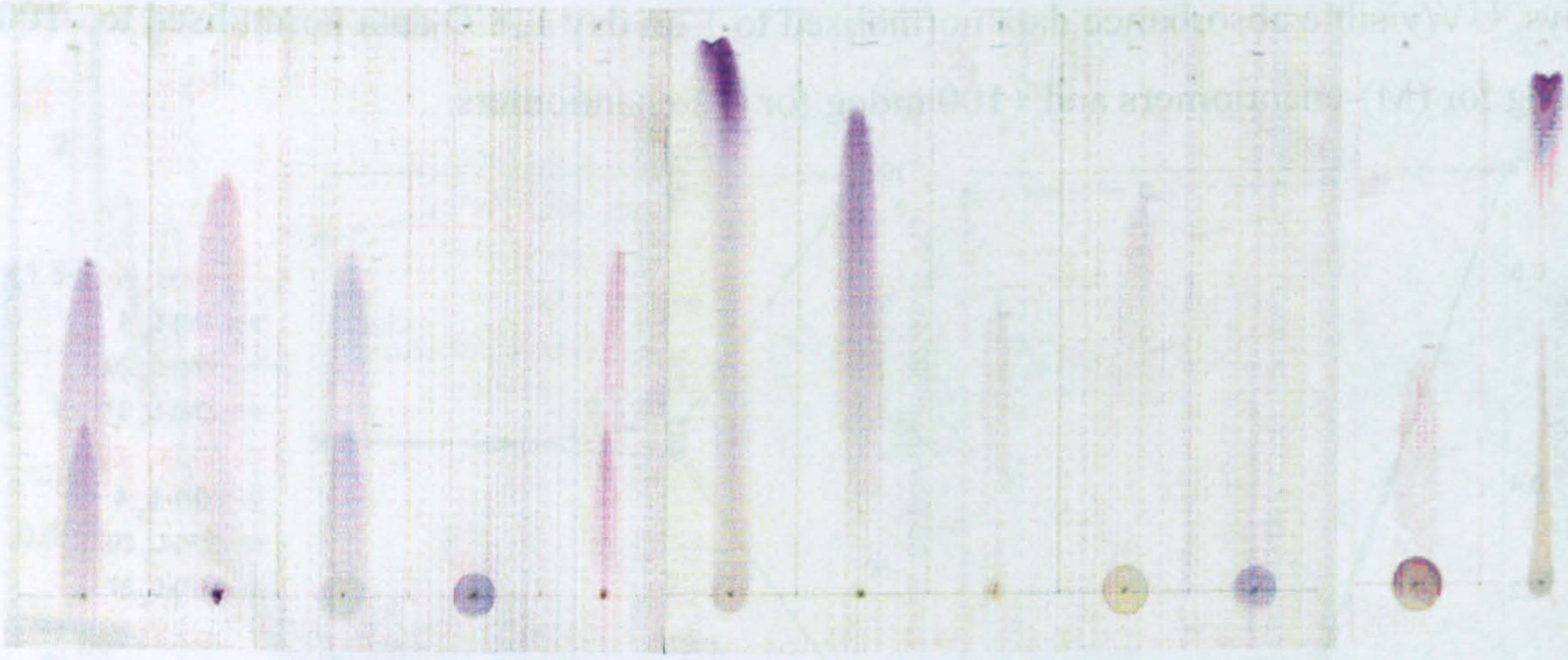


Intensity plots showing decrease in UV/visible absorbance and *CD* in MLCT signal for enantiomers of (a) $L_1$, (b) $L_3$ and (c) $L_5$ at different temperatures over a period of 22 days. UV/visible absorbance data normalised to 1 on day 1, *CD* data normalised to –100 mdeg for (M)-enantiomers and +100 mdeg for (P)-enantiomers.

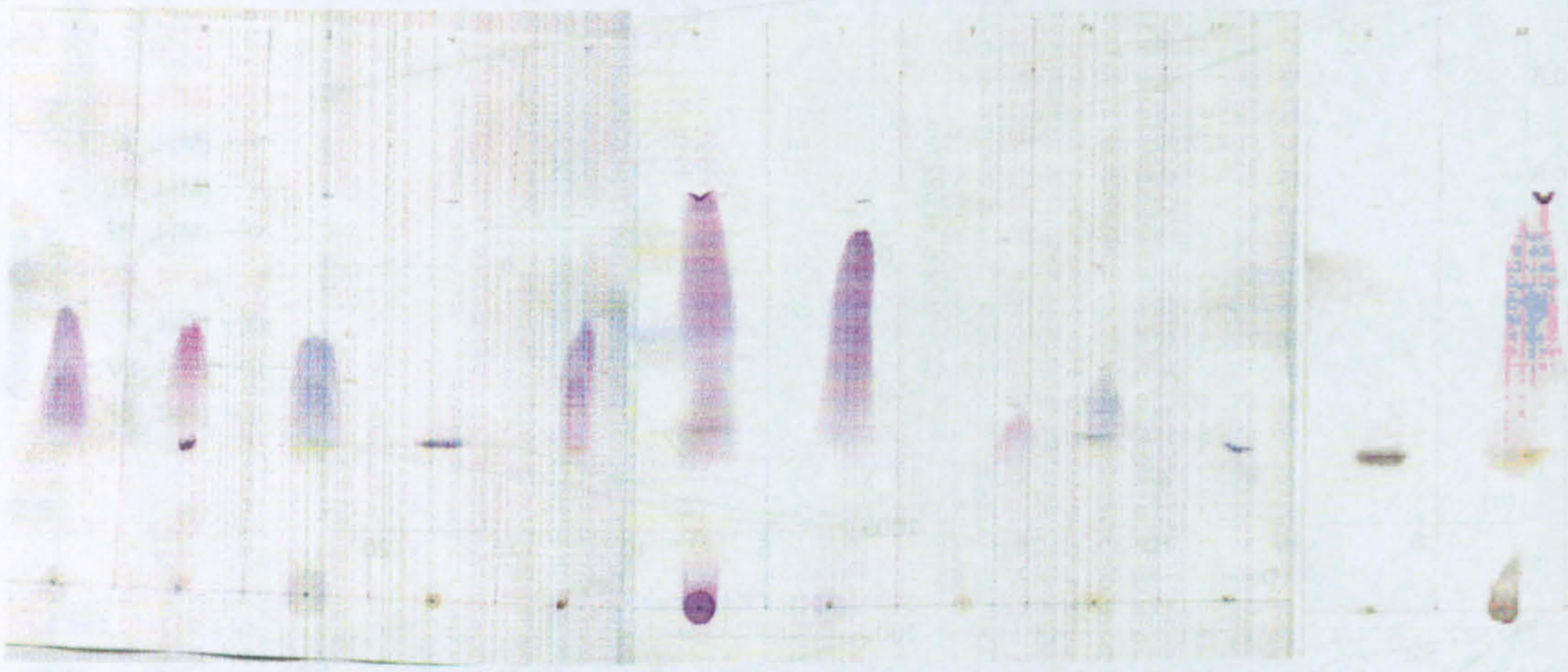


Paper chromatograms

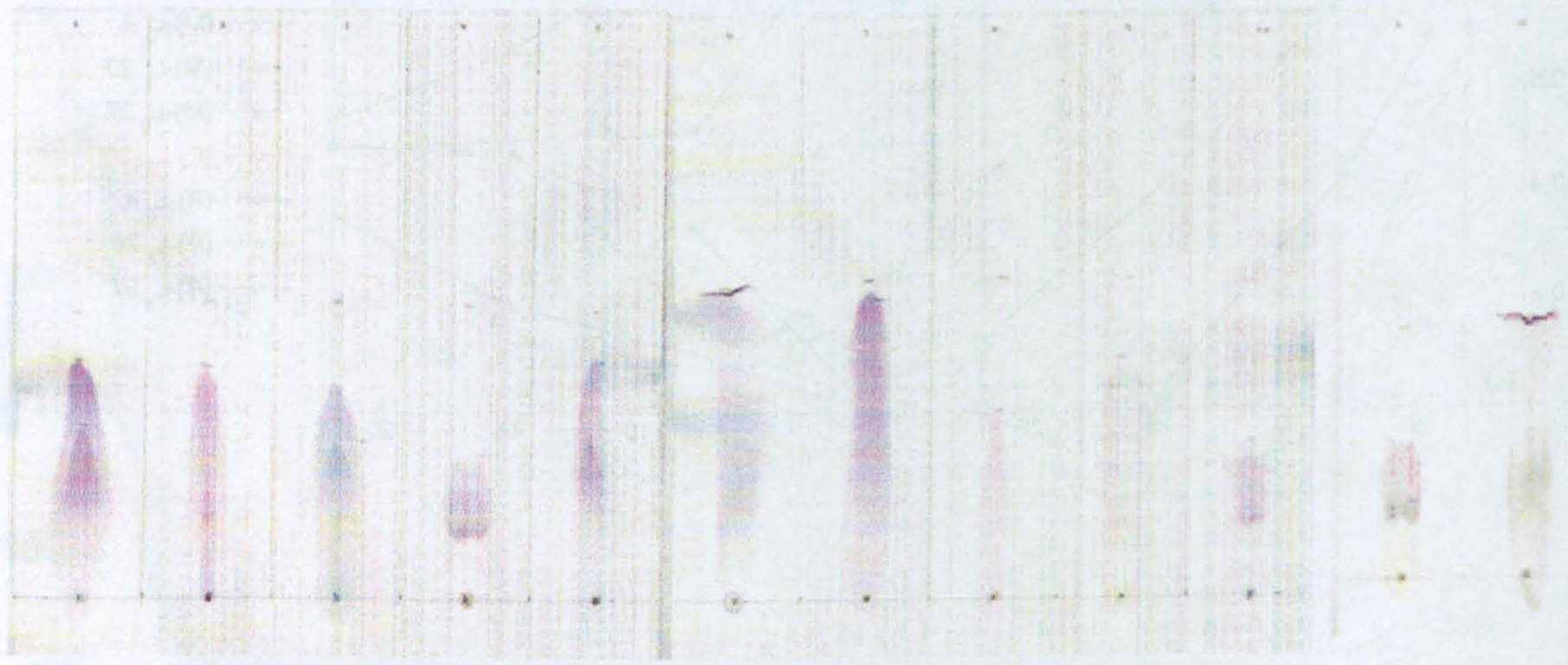
Complexes $L_1 - L_{12}$ eluted with 0.02 M NaCl



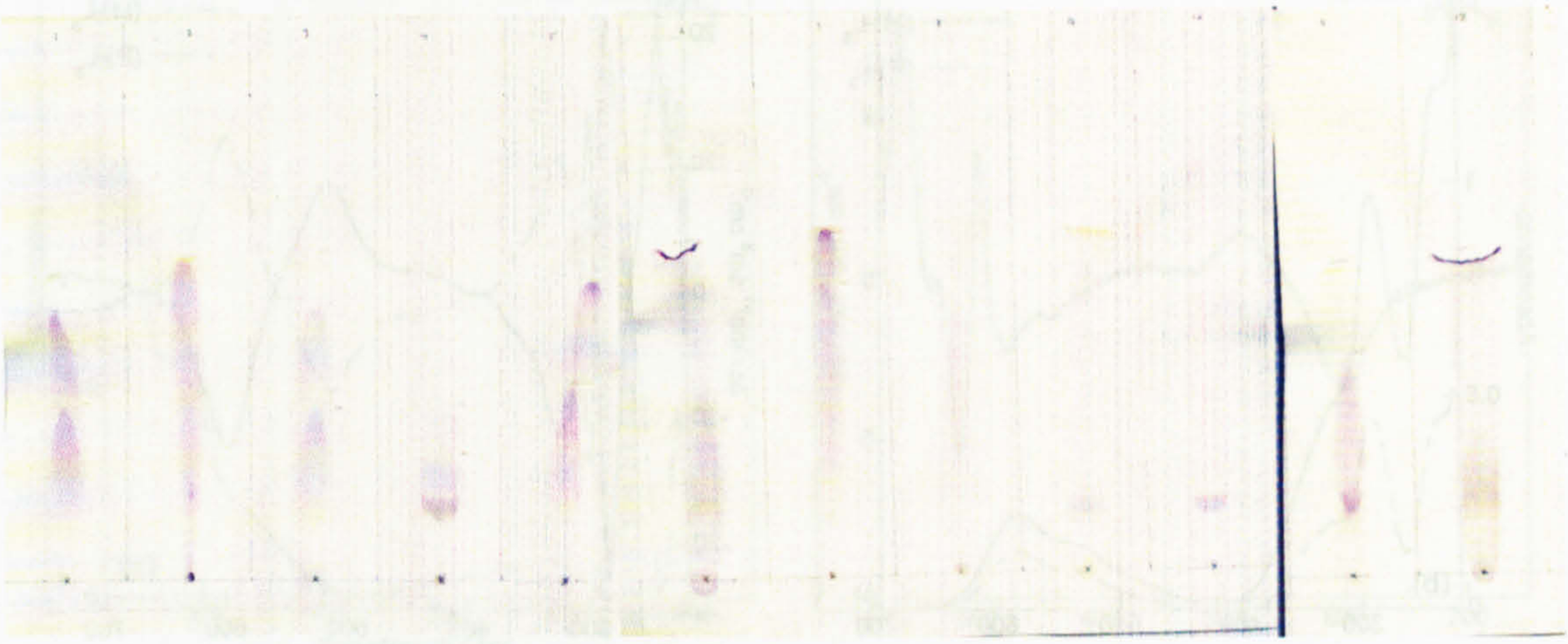
Complexes $L_1 - L_{12}$ eluted with 90% MeCN: 10% 0.02 M NaCl



Complexes $L_1 - L_{12}$ eluted with 80% MeOH: 20% 0.02 M NaCl

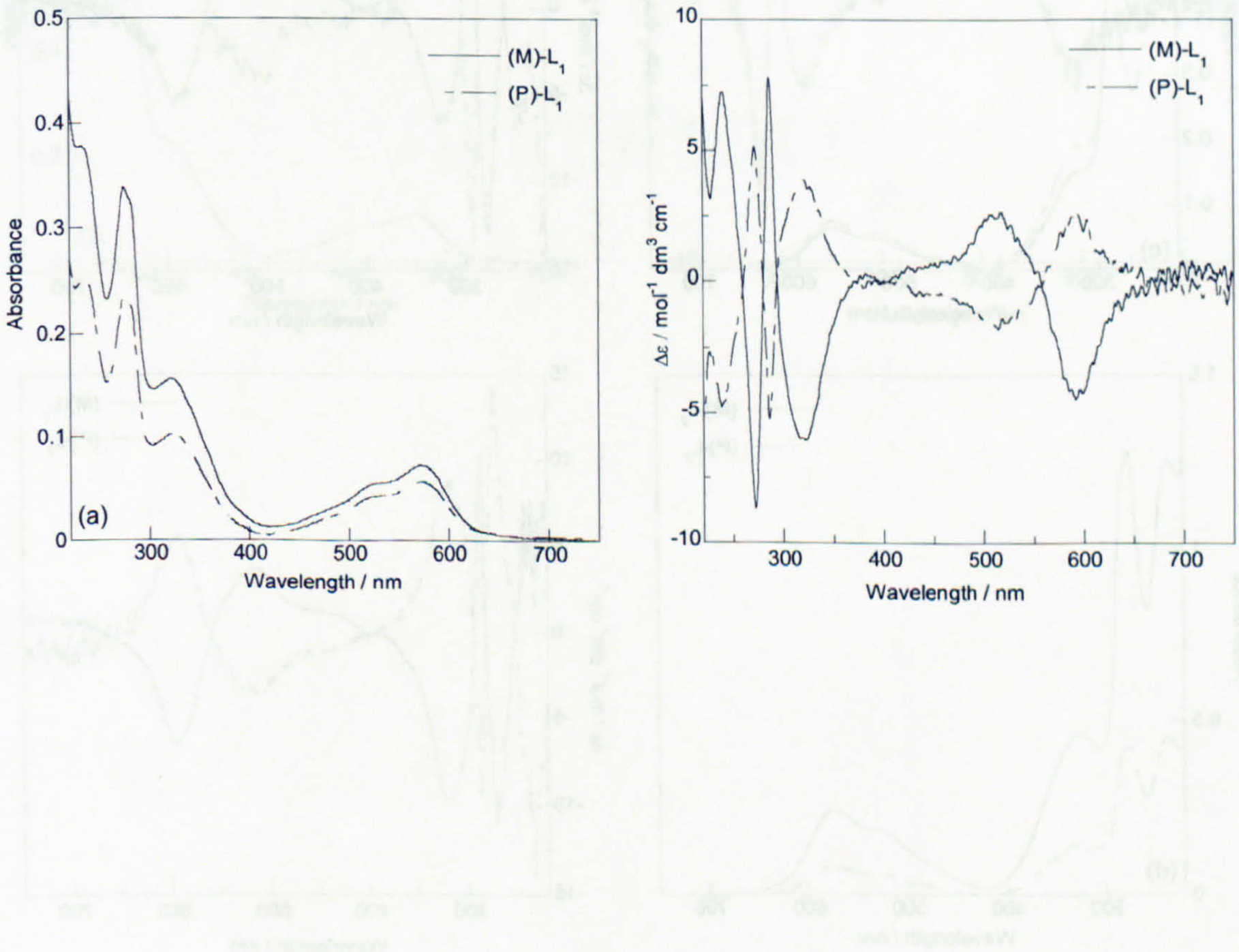


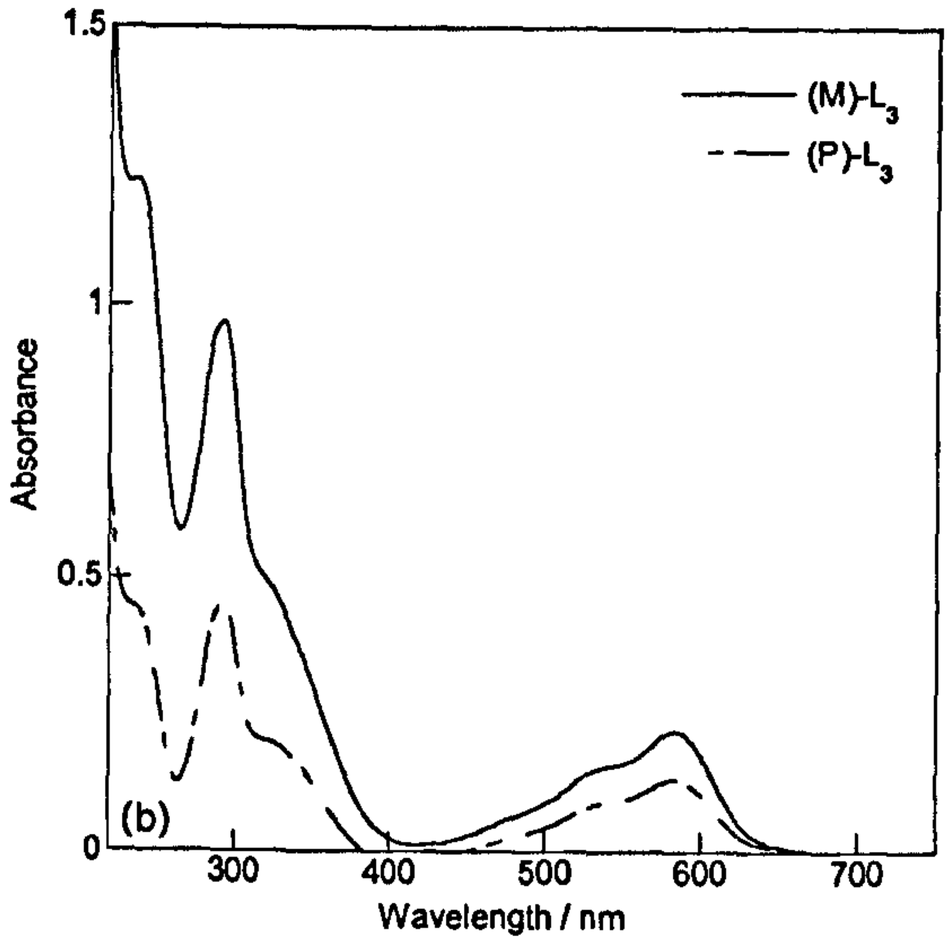
Complexes $L_1 - L_{12}$ eluted with 75% EtOH: 25% 0.02 M NaCl

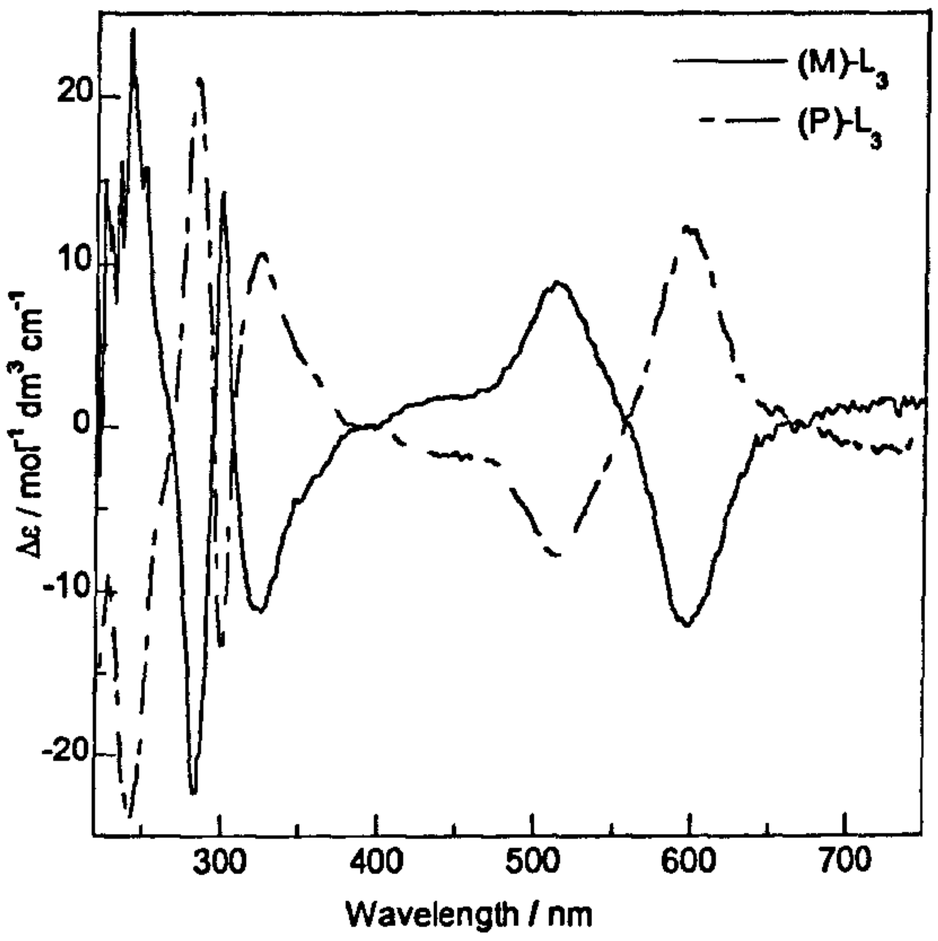


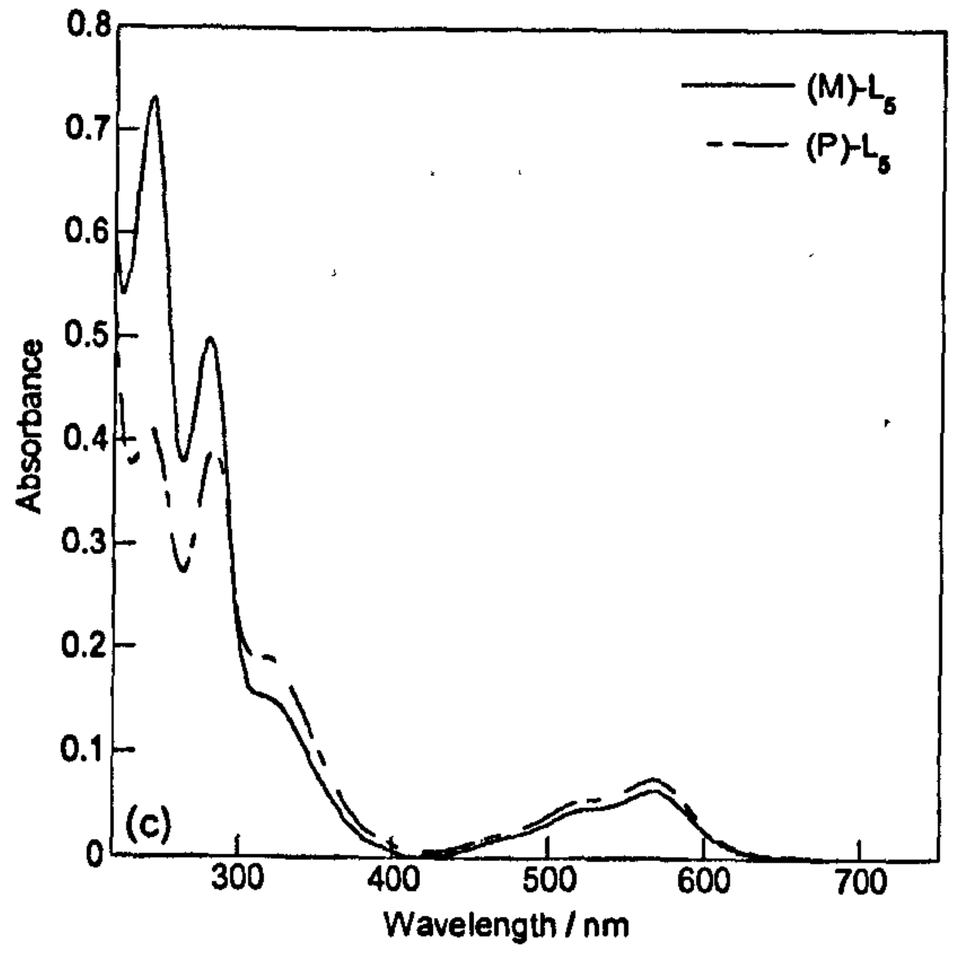
Cellulose column chromatography

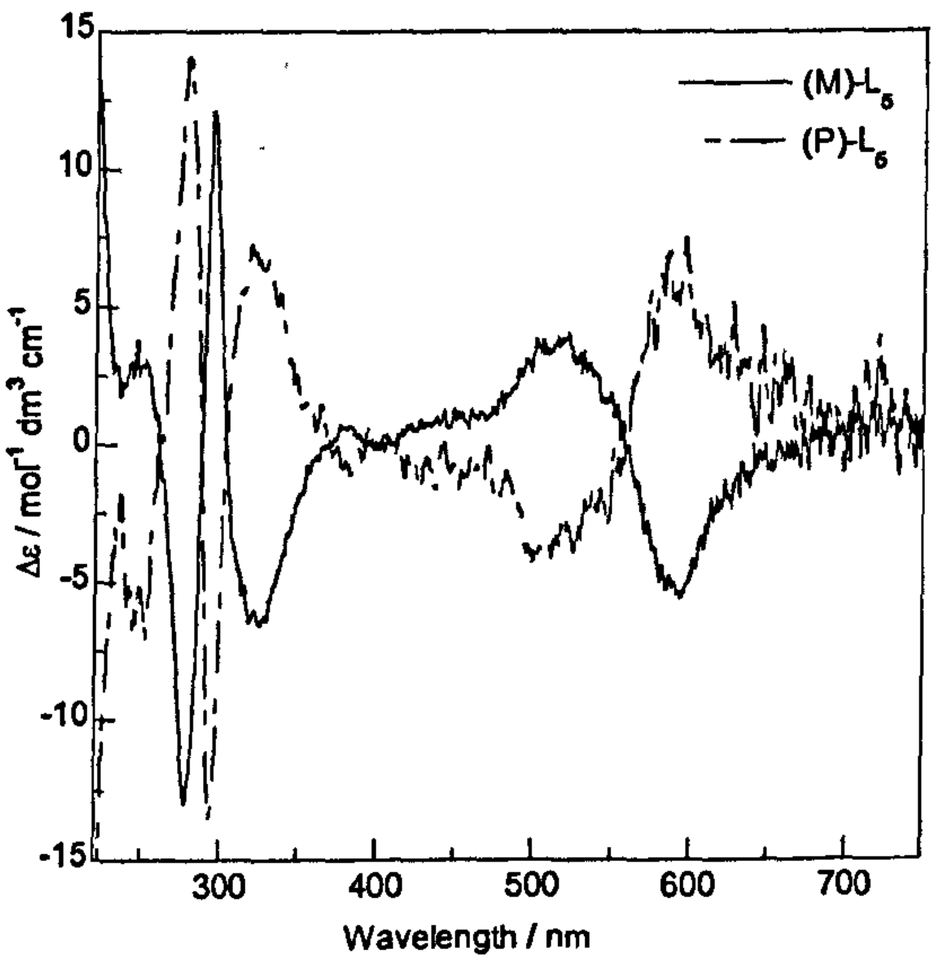
UV/visible absorbance and *CD* spectra of some of the derivatives separated. (a) $L_1$ separated using 0.01 M NaCl. (b) $L_3$ (c) $L_5$ (d) $L_7$ (e) $L_9$ and (f) $L_{11}$ separated using 90% MeCN:10 % 0.02 M NaCl v/v.

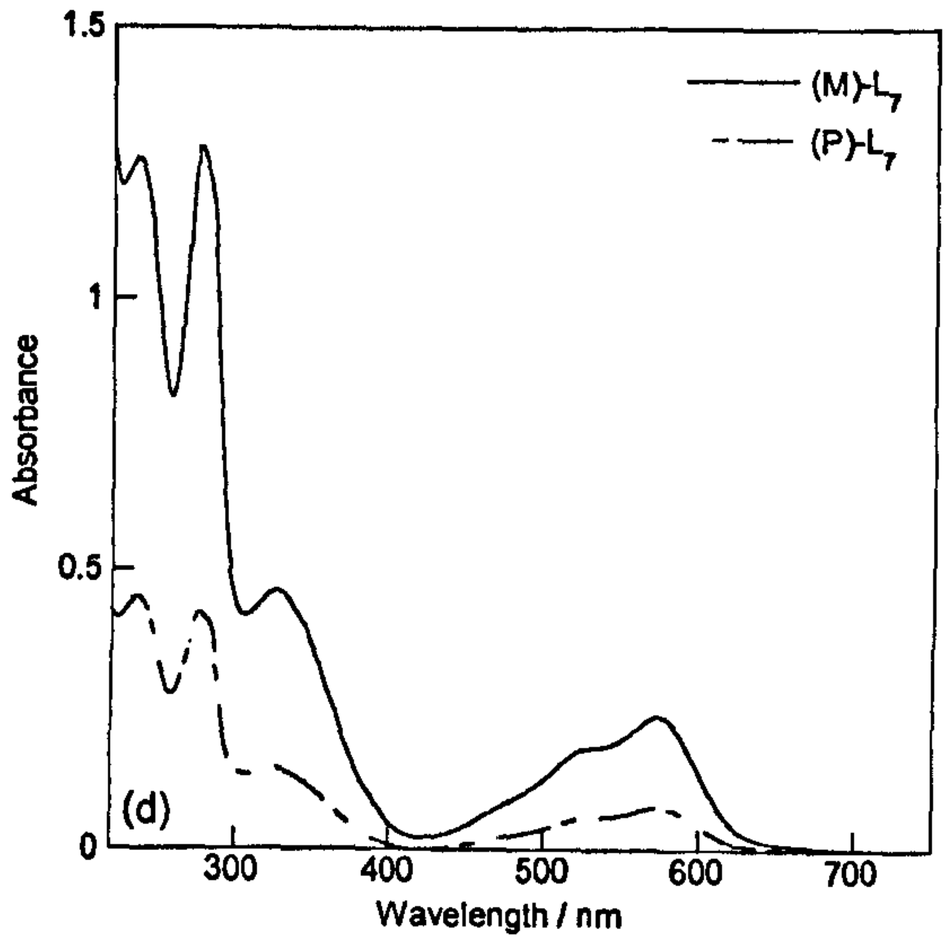


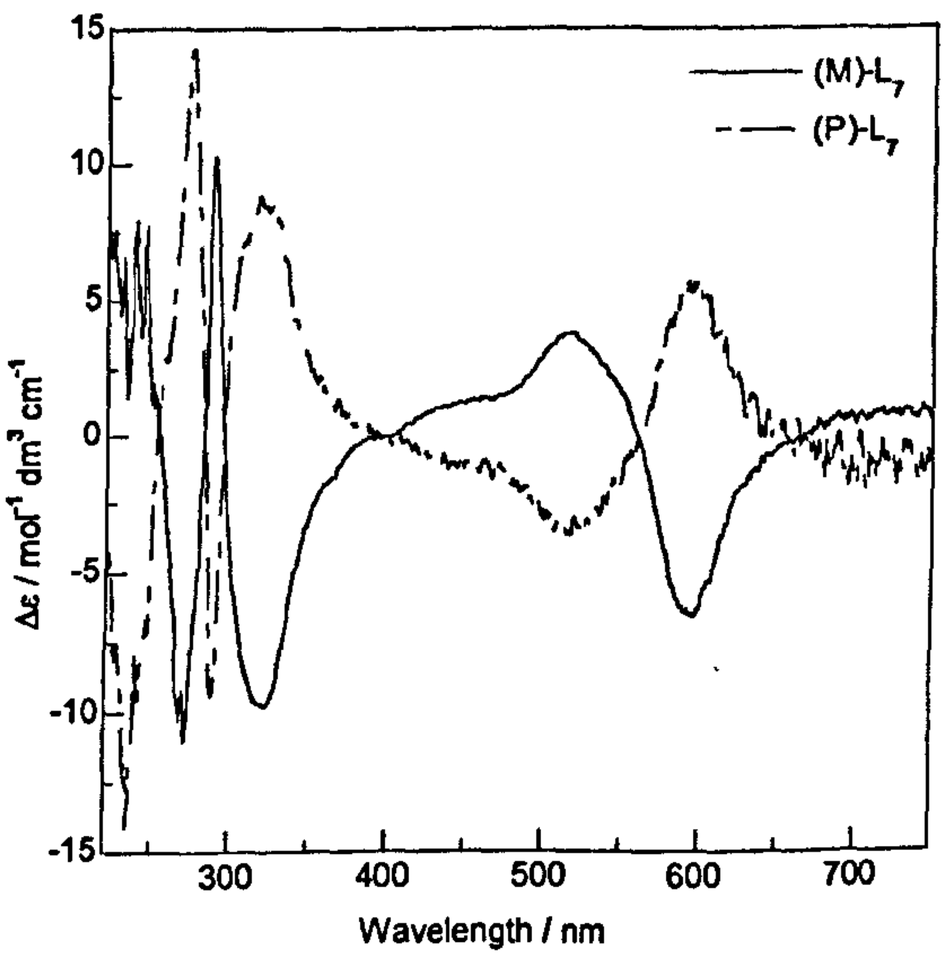

(M)-L3
(P)-L3
Absorbance
(b)
Wavelength / nm

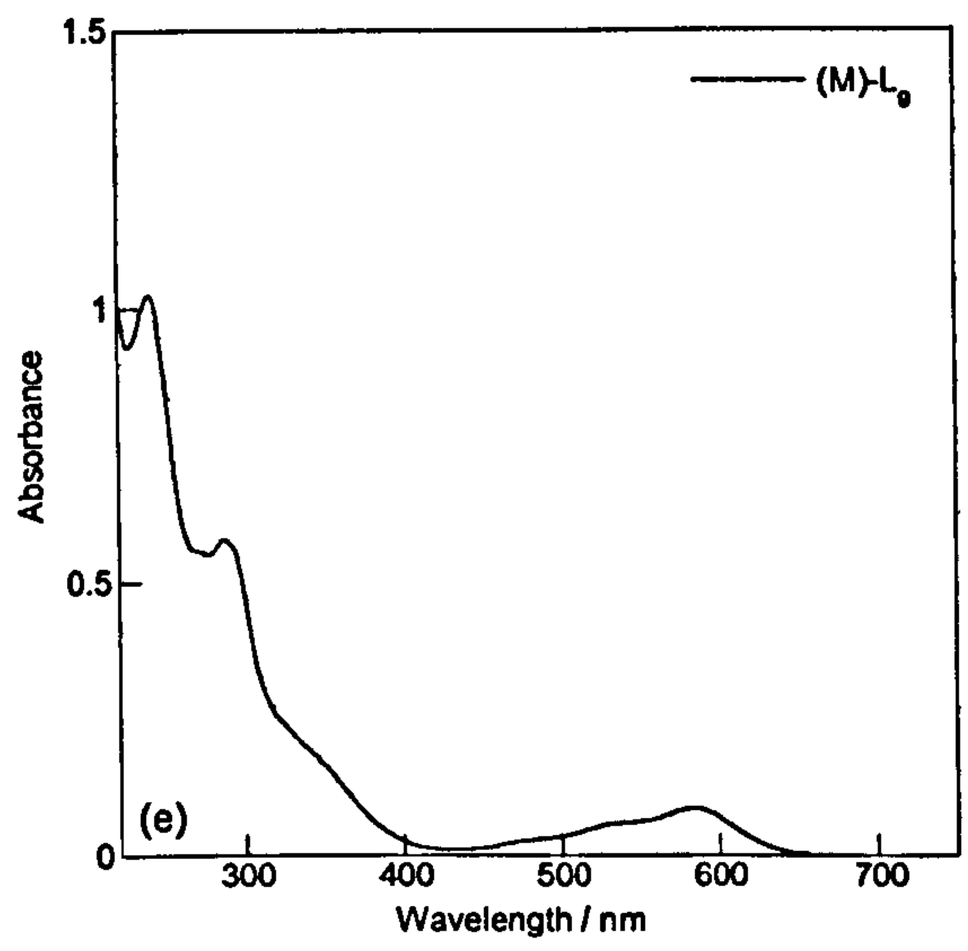

(M)-L3
(P)-L3
Δε / mol-1 dm3 cm-1
Wavelength / nm

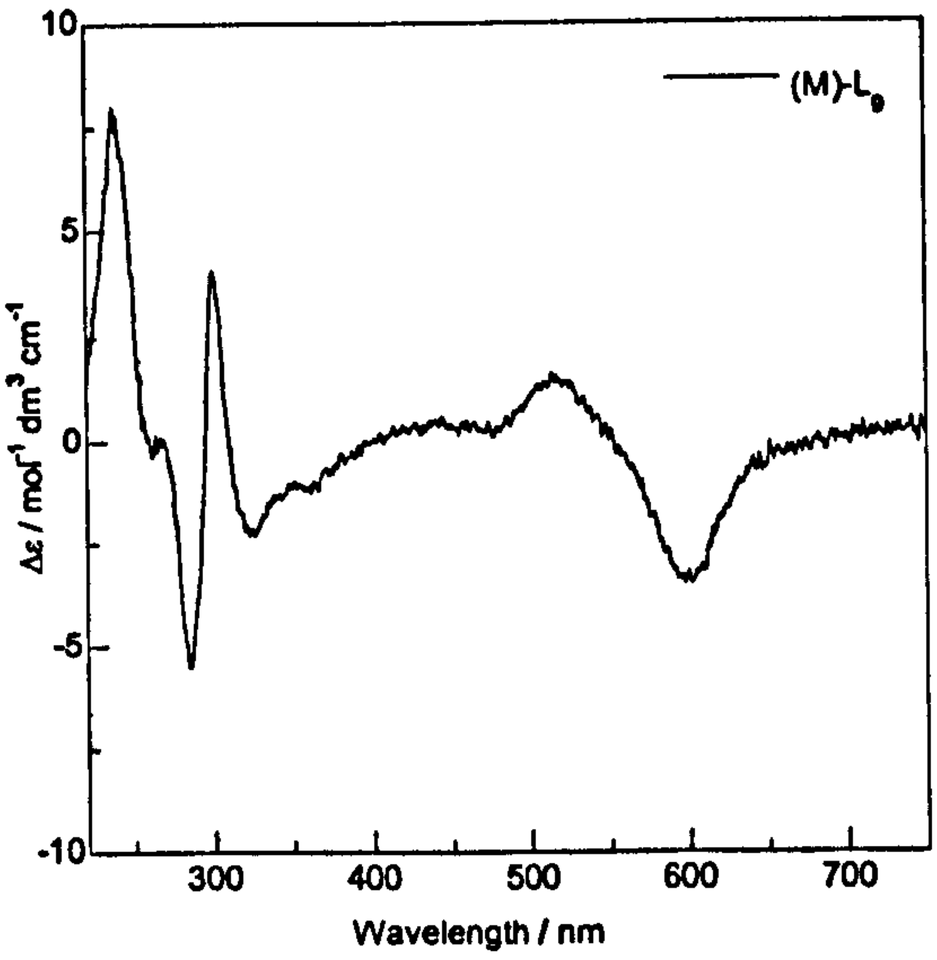

(M)-L5
(P)-L5
Absorbance
(c)
Wavelength / nm

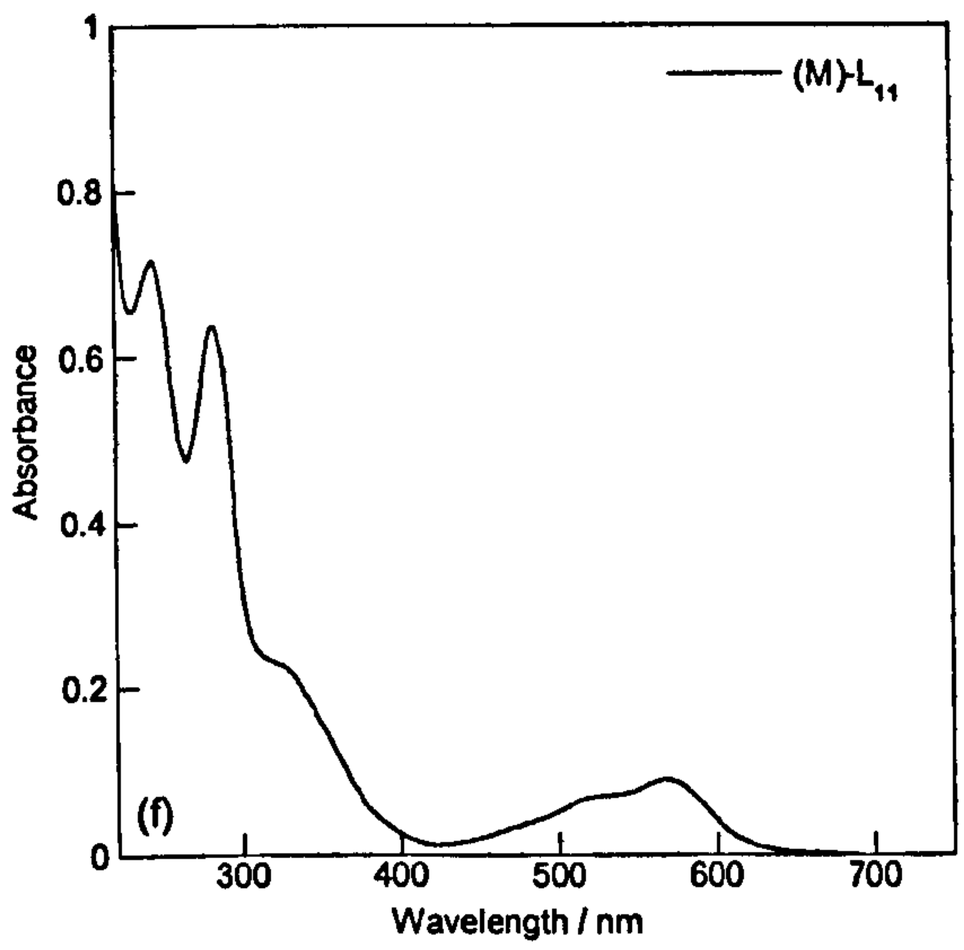

(M)-L5
(P)-L6
Δε / mol-1 dm3 cm-1
Wavelength / nm

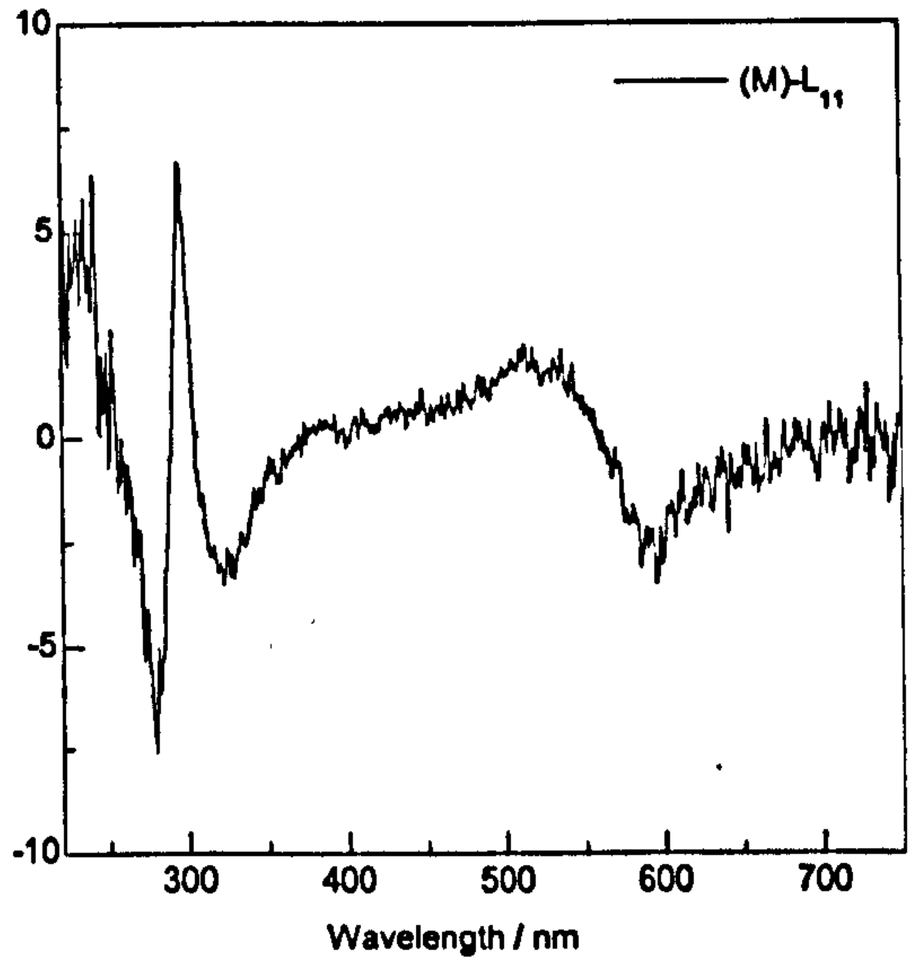

(M)-L7
(P)-L7
Absorbance
(d)
Wavelength / nm


(M)-L7
(P)-L7
Δε / mol-1 dm3 cm-1
Wavelength / nm


(M)-L9
Absorbance
(e)
Wavelength / nm


(M)-L9
Δε / mol-1 dm3 cm-1
Wavelength / nm


(M)-L11
Absorbance
(f)
Wavelength / nm


(M)-L11
Δε / mol-1 dm3 cm-1
Wavelength / nm

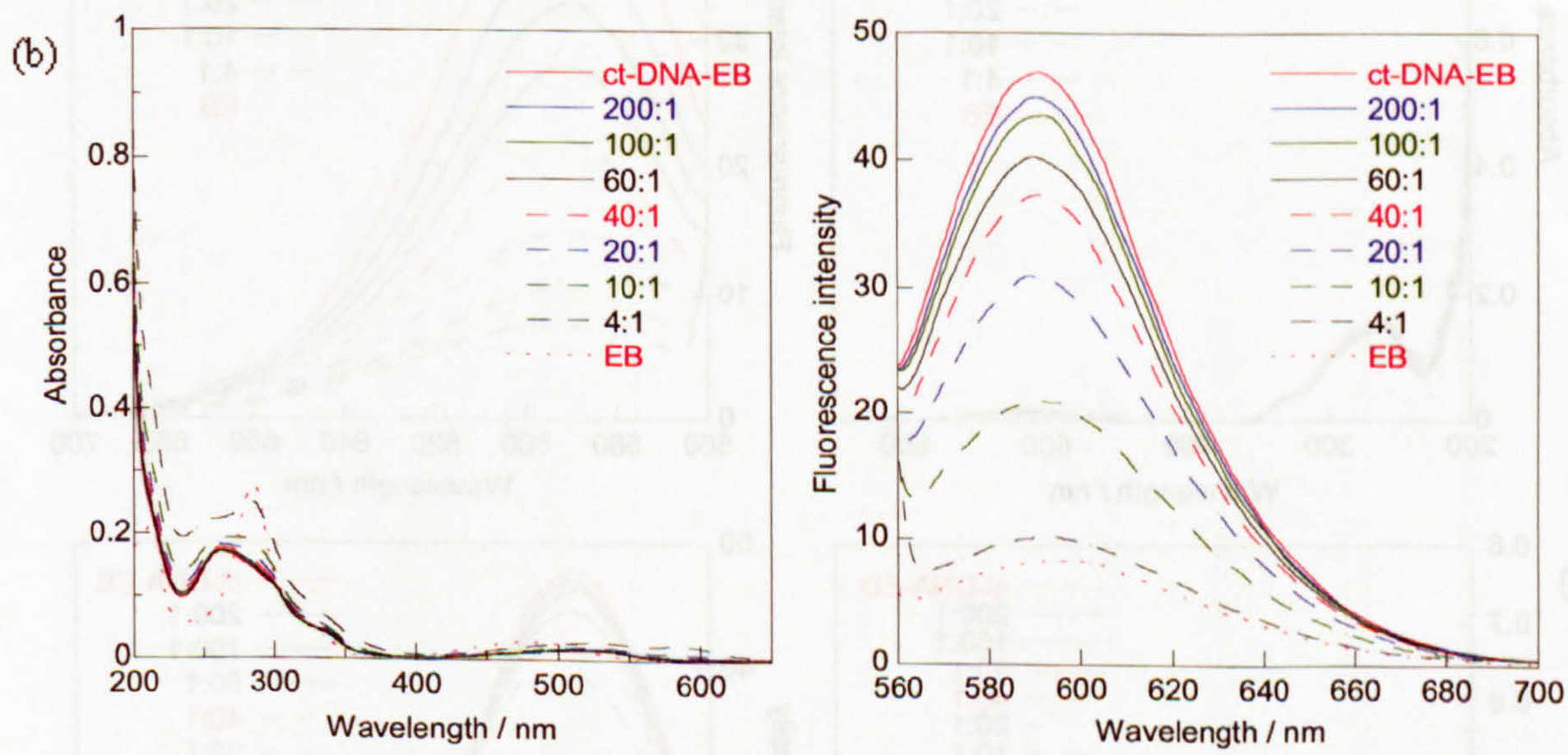
# APPENDIX II

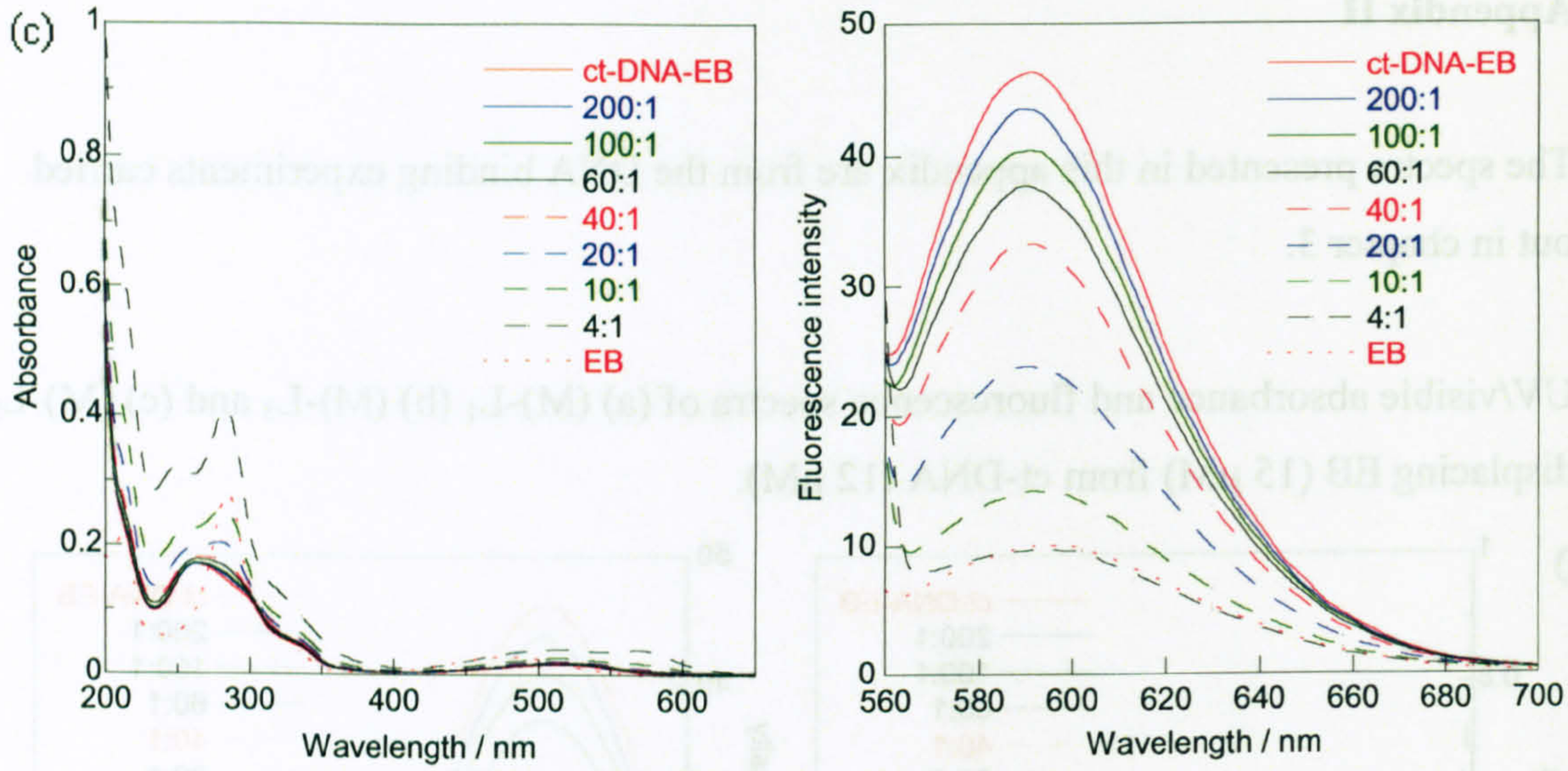
**Appendix II**

The spectra presented in this appendix are from the DNA binding experiments carried out in chapter 3.

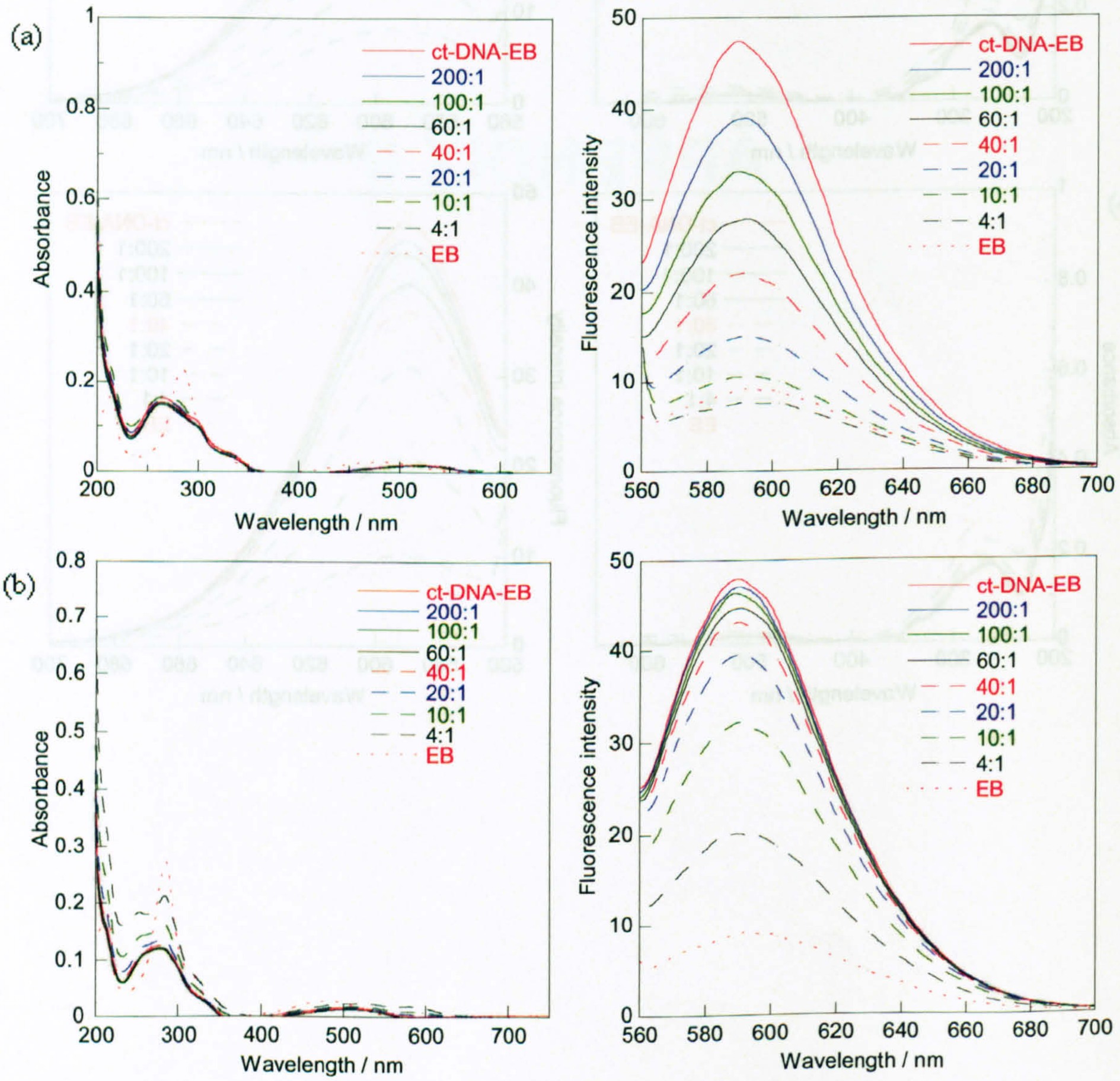
UV/visible absorbance and fluorescence spectra of (a) (M)-$L_1$ (b) (M)-$L_3$ and (c) (M)-$L_5$ displacing EB (15 μM) from ct-DNA (12 μM).

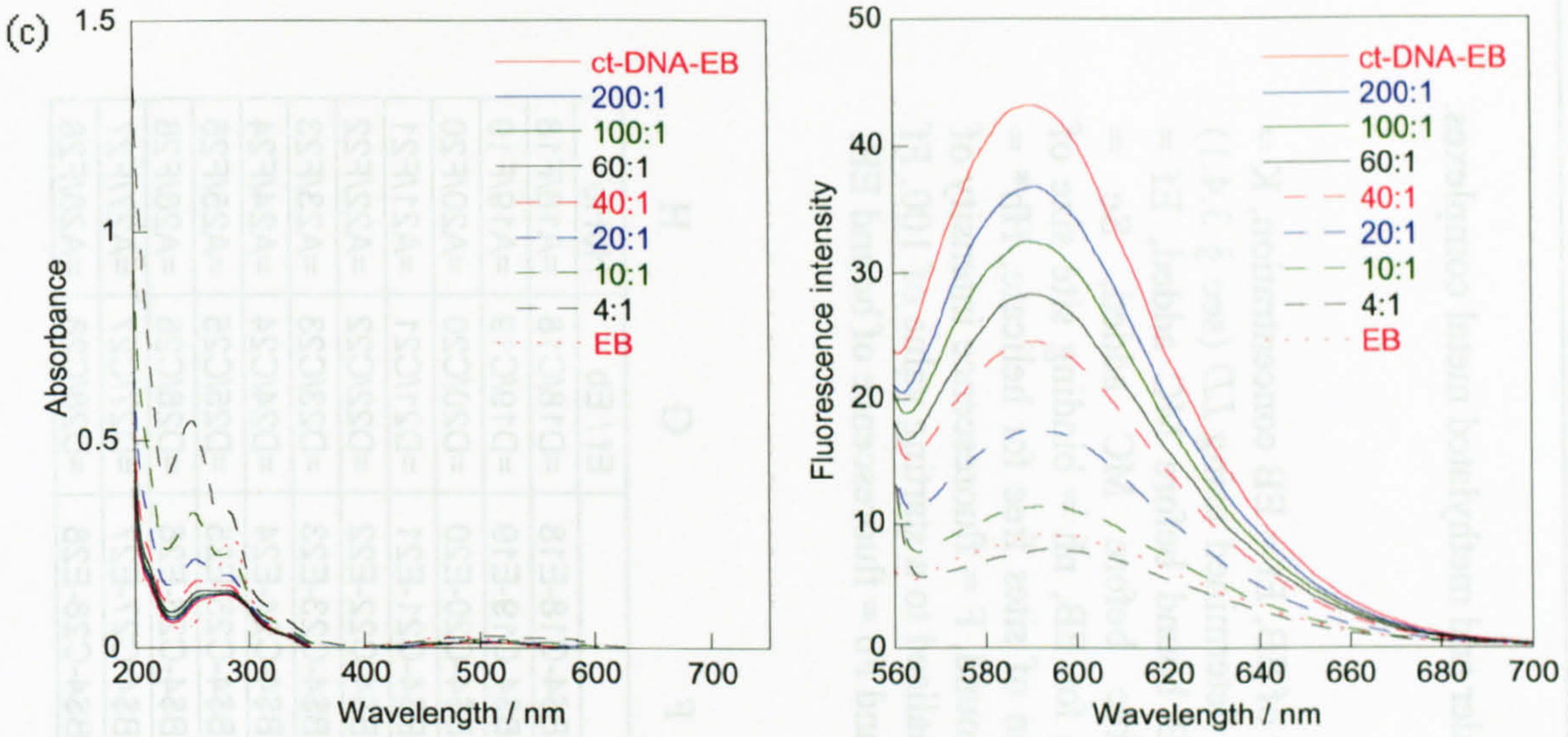





UV/visible absorbance and fluorescence spectra of (a) (P)-$L_1$ (b) (P)-$L_3$ and (c) (P)-$L_5$ displacing EB (15 μM) from ct-DNA (12 μM).




(c)
Absorbance
Wavelength / nm
ct-DNA-EB
200:1
100:1
60:1
40:1
20:1
10:1
4:1
EB
Fluorescence intensity
Wavelength / nm

Data table showing formulae used to calculate the binding constants of the enantiomers of the parent cylinder and methylated metal complexes.

| | |
|---|---|
| DNA | = 12 |
| ne | = 4.5 |
| Et | = 15 |
| K | = 0.54 |
| Eb | = ((B3/B4 + B5 + 1/B6) - SQRT((B3/B4 + B5 + 1/B6)^2 - 4*B5*B3/B4)) /2 |
| Ef | = B5 - B7 |
| Se | = B3/B4-B7 |
| nh | = 5 |
| Sf | = B7 * B4 / (B8 *B6 * B10) |
| Hb* | = B3/B4-B7-B9 |
| F | = 100 |
| Ff | = B28 |
| Fb | = (B5*B13-B8*B14)/B7 |

Key: ne = binding site size of EB, Et = EB concentration, K = binding constant of EB as determined using *LD* (see § 3.4.1) Eb = concentration of EB bound before MC added, Ef = concentration of EB free before MC added, Se = concentration of sites free for EB, nh = binding site size of helicate, Sf = concentration of sites free for helicate, Hb* = concentration of helicate bound, F = fluorescence intensity of ct-DNA-EB complex normalised to a starting value of 100. Ff = fluorescence of free EB and Fb = fluorescence of bound EB.

| A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|
| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
| 0 | 100 | =$B$5*(B18 - $B$14)/($B$15-$B$14) | =$B$5-C18 | =C18/(D18*$B$6) | =$B$3/$B$4-C18-E18 | =D18/C18 | =A18/F18 |
| 0.16 | 82.864 | =$B$5*(B19 - $B$14)/($B$15-$B$14) | =$B$5-C19 | =C19/(D19*$B$6) | =$B$3/$B$4-C19-E19 | =D19/C19 | =A19/F19 |
| 0.2 | 80.097 | =$B$5*(B20 - $B$14)/($B$15-$B$14) | =$B$5-C20 | =C20/(D20*$B$6) | =$B$3/$B$4-C20-E20 | =D20/C20 | =A20/F20 |
| 0.267 | 74.866 | =$B$5*(B21 - $B$14)/($B$15-$B$14) | =$B$5-C21 | =C21/(D21*$B$6) | =$B$3/$B$4-C21-E21 | =D21/C21 | =A21/F21 |
| 0.4 | 68.356 | =$B$5*(B22 - $B$14)/($B$15-$B$14) | =$B$5-C22 | =C22/(D22*$B$6) | =$B$3/$B$4-C22-E22 | =D22/C22 | =A22/F22 |
| 0.5 | 64.687 | =$B$5*(B23 - $B$14)/($B$15-$B$14) | =$B$5-C23 | =C23/(D23*$B$6) | =$B$3/$B$4-C23-E23 | =D23/C23 | =A23/F23 |
| 0.667 | 58.714 | =$B$5*(B24 - $B$14)/($B$15-$B$14) | =$B$5-C24 | =C24/(D24*$B$6) | =$B$3/$B$4-C24-E24 | =D24/C24 | =A24/F24 |
| 1 | 54.483 | =$B$5*(B25 - $B$14)/($B$15-$B$14) | =$B$5-C25 | =C25/(D25*$B$6) | =$B$3/$B$4-C25-E25 | =D25/C25 | =A25/F25 |
| 2 | 47.751 | =$B$5*(B26 - $B$14)/($B$15-$B$14) | =$B$5-C26 | =C26/(D26*$B$6) | =$B$3/$B$4-C26-E26 | =D26/C26 | =A26/F26 |
| 4 | 44.767 | =$B$5*(B27 - $B$14)/($B$15-$B$14) | =$B$5-C27 | =C27/(D27*$B$6) | =$B$3/$B$4-C27-E27 | =D27/C27 | =A27/F27 |
| 8 | 39.446 | =$B$5*(B28 - $B$14)/($B$15-$B$14) | =$B$5-C28 | =C28/(D28*$B$6) | =$B$3/$B$4-C28-E28 | =D28/C28 | =A28/F28 |

(R)-$L_1$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | 0.00E+00 | 5.446937 | #DIV/0! |
| 0.16 | 82.864 | 1.668264 | 1.33E+01 | 2.32E-01 | 7.67E-01 | 7.991382 | 0.208694 |
| 0.2 | 80.097 | 1.561947 | 1.34E+01 | 2.15E-01 | 8.89E-01 | 8.6034 | 0.224852 |
| 0.267 | 74.866 | 1.360954 | 1.36E+01 | 1.85E-01 | 1.12E+00 | 10.02168 | 0.238196 |
| 0.4 | 68.356 | 1.110818 | 1.39E+01 | 1.48E-01 | 1.41E+00 | 12.50356 | 0.284143 |
| 0.5 | 64.687 | 0.969843 | 1.40E+01 | 1.28E-01 | 1.57E+00 | 14.46642 | 0.318712 |
| 0.667 | 58.714 | 0.740341 | 1.43E+01 | 9.61E-02 | 1.83E+00 | 19.26094 | 0.364445 |
| 1 | 54.483 | 0.577772 | 1.44E+01 | 7.42E-02 | 2.01E+00 | 24.96182 | 0.49635 |
| 2 | 47.751 | 0.319106 | 1.47E+01 | 4.03E-02 | 2.31E+00 | 46.00636 | 0.866811 |
| 4 | 44.767 | 0.204451 | 1.48E+01 | 2.56E-02 | 2.44E+00 | 72.36738 | 1.641614 |
| 8 | 39.446 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(R)-$L_3$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 91.608 | 1.995194 | 1.30E+01 | 2.84E-01 | 3.87E-01 | 6.518066 | 0.41305 |
| 0.2 | 88.467 | 1.871122 | 1.31E+01 | 2.64E-01 | 5.32E-01 | 7.016582 | 0.376209 |
| 0.267 | 85.819 | 1.766523 | 1.32E+01 | 2.47E-01 | 6.53E-01 | 7.491257 | 0.408918 |
| 0.4 | 82.513 | 1.635933 | 1.34E+01 | 2.27E-01 | 8.04E-01 | 8.169081 | 0.497485 |
| 0.5 | 81.425 | 1.592956 | 1.34E+01 | 2.20E-01 | 8.54E-01 | 8.416458 | 0.585697 |
| 0.667 | 75.975 | 1.377675 | 1.36E+01 | 1.87E-01 | 1.10E+00 | 9.887906 | 0.605424 |
| 1 | 72.65 | 1.246335 | 1.38E+01 | 1.68E-01 | 1.25E+00 | 11.03529 | 0.79839 |
| 2 | 61.375 | 0.800961 | 1.42E+01 | 1.04E-01 | 1.76E+00 | 17.7275 | 1.135561 |
| 4 | 51.829 | 0.423885 | 1.46E+01 | 5.39E-02 | 2.19E+00 | 34.38696 | 1.827378 |
| 8 | 41.098 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(R)-$L_5$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 85.014 | 1.731298 | 1.33E+01 | 2.42E-01 | 6.94E-01 | 7.664019 | 0.230634 |
| 0.2 | 84.15 | 1.696972 | 1.33E+01 | 2.36E-01 | 7.33E-01 | 7.839276 | 0.272677 |
| 0.267 | 79.82 | 1.524942 | 1.35E+01 | 2.10E-01 | 9.32E-01 | 8.836437 | 0.286433 |
| 0.4 | 74.051 | 1.295742 | 1.37E+01 | 1.75E-01 | 1.20E+00 | 10.57638 | 0.334495 |
| 0.5 | 69.9 | 1.130824 | 1.39E+01 | 1.51E-01 | 1.38E+00 | 12.26466 | 0.36105 |
| 0.667 | 65.352 | 0.950134 | 1.40E+01 | 1.25E-01 | 1.59E+00 | 14.78725 | 0.419154 |
| 1 | 55.573 | 0.561618 | 1.44E+01 | 7.20E-02 | 2.03E+00 | 25.70854 | 0.49188 |
| 2 | 50 | 0.340205 | 1.47E+01 | 4.30E-02 | 2.28E+00 | 43.09109 | 0.875854 |
| 4 | 46.852 | 0.215136 | 1.48E+01 | 2.69E-02 | 2.42E+00 | 68.72336 | 1.649767 |
| 8 | 41.437 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(M)-$L_1$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -1.05E-15 | 5.446937 | 0 |
| 0.16 | 87.654 | 1.852735 | 1.31E+01 | 2.61E-01 | 5.53E-01 | 7.096141 | 0.289349 |
| 0.2 | 86.483 | 1.807781 | 1.32E+01 | 2.54E-01 | 6.05E-01 | 7.297466 | 0.330513 |
| 0.267 | 79.663 | 1.545967 | 1.35E+01 | 2.13E-01 | 9.08E-01 | 8.702664 | 0.294083 |
| 0.4 | 75.346 | 1.380242 | 1.36E+01 | 1.88E-01 | 1.10E+00 | 9.867663 | 0.364048 |
| 0.5 | 71.681 | 1.239545 | 1.38E+01 | 1.67E-01 | 1.26E+00 | 11.10121 | 0.396729 |
| 0.667 | 65.726 | 1.010938 | 1.40E+01 | 1.34E-01 | 1.52E+00 | 13.8377 | 0.438267 |
| 1 | 58.147 | 0.719987 | 1.43E+01 | 9.34E-02 | 1.85E+00 | 19.8337 | 0.539575 |
| 2 | 51.417 | 0.461629 | 1.45E+01 | 5.88E-02 | 2.15E+00 | 31.49364 | 0.931864 |
| 4 | 47.644 | 0.316787 | 1.47E+01 | 4.00E-02 | 2.31E+00 | 46.35046 | 1.731657 |
| 8 | 39.392 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(M)-$L_3$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 93.922 | 2.097902 | 1.29E+01 | 3.01E-01 | 2.68E-01 | 6.15 | 0.597795 |
| 0.2 | 90.436 | 1.966685 | 1.30E+01 | 2.79E-01 | 4.21E-01 | 6.627049 | 0.475575 |
| 0.267 | 86.36 | 1.813258 | 1.32E+01 | 2.55E-01 | 5.99E-01 | 7.2724 | 0.445916 |
| 0.4 | 80.353 | 1.587147 | 1.34E+01 | 2.19E-01 | 8.60E-01 | 8.450921 | 0.464906 |
| 0.5 | 76.334 | 1.435866 | 1.36E+01 | 1.96E-01 | 1.03E+00 | 9.446654 | 0.4832 |
| 0.667 | 71.321 | 1.24717 | 1.38E+01 | 1.68E-01 | 1.25E+00 | 11.02723 | 0.532934 |
| 1 | 64.417 | 0.987295 | 1.40E+01 | 1.30E-01 | 1.55E+00 | 14.19303 | 0.645621 |
| 2 | 53.284 | 0.568234 | 1.44E+01 | 7.29E-02 | 2.03E+00 | 25.39759 | 0.987401 |
| 4 | 45.554 | 0.277266 | 1.47E+01 | 3.49E-02 | 2.35E+00 | 53.09966 | 1.698856 |
| 8 | 38.188 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(M)-$L_5$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 89.003 | 1.879509 | 1.31E+01 | 2.65E-01 | 5.22E-01 | 6.980806 | 0.306584 |
| 0.2 | 85.239 | 1.726452 | 1.33E+01 | 2.41E-01 | 6.99E-01 | 7.68834 | 0.28598 |
| 0.267 | 78.591 | 1.456121 | 1.35E+01 | 1.99E-01 | 1.01E+00 | 9.301345 | 0.263977 |
| 0.4 | 72.212 | 1.196728 | 1.38E+01 | 1.61E-01 | 1.31E+00 | 11.53418 | 0.305487 |
| 0.5 | 68.593 | 1.049567 | 1.40E+01 | 1.39E-01 | 1.48E+00 | 13.29161 | 0.338346 |
| 0.667 | 62.214 | 0.790174 | 1.42E+01 | 1.03E-01 | 1.77E+00 | 17.98316 | 0.376089 |
| 1 | 54.959 | 0.49516 | 1.45E+01 | 6.32E-02 | 2.11E+00 | 29.29325 | 0.474318 |
| 2 | 50.169 | 0.300382 | 1.47E+01 | 3.78E-02 | 2.33E+00 | 48.93649 | 0.858943 |
| 4 | 48.096 | 0.216086 | 1.48E+01 | 2.71E-02 | 2.42E+00 | 68.41679 | 1.650496 |
| 8 | 42.782 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(P)-$L_1$ data table

| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 82.745 | 1.671632 | 1.33E+01 | 2.32E-01 | 7.63E-01 | 7.973267 | 0.20976 |
| 0.2 | 79.75 | 1.557932 | 1.34E+01 | 2.15E-01 | 8.94E-01 | 8.628147 | 0.223687 |
| 0.267 | 74.647 | 1.364206 | 1.36E+01 | 1.85E-01 | 1.12E+00 | 9.995405 | 0.238992 |
| 0.4 | 67.117 | 1.078344 | 1.39E+01 | 1.43E-01 | 1.44E+00 | 12.91022 | 0.276839 |
| 0.5 | 63.948 | 0.958038 | 1.40E+01 | 1.26E-01 | 1.58E+00 | 14.65699 | 0.315999 |
| 0.667 | 59.544 | 0.790849 | 1.42E+01 | 1.03E-01 | 1.77E+00 | 17.96697 | 0.376252 |
| 1 | 55.27 | 0.628594 | 1.44E+01 | 8.10E-02 | 1.96E+00 | 22.86278 | 0.510967 |
| 2 | 50.4 | 0.443713 | 1.46E+01 | 5.64E-02 | 2.17E+00 | 32.8056 | 0.923146 |
| 4 | 46.213 | 0.284762 | 1.47E+01 | 3.58E-02 | 2.35E+00 | 51.67563 | 1.70498 |
| 8 | 38.712 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

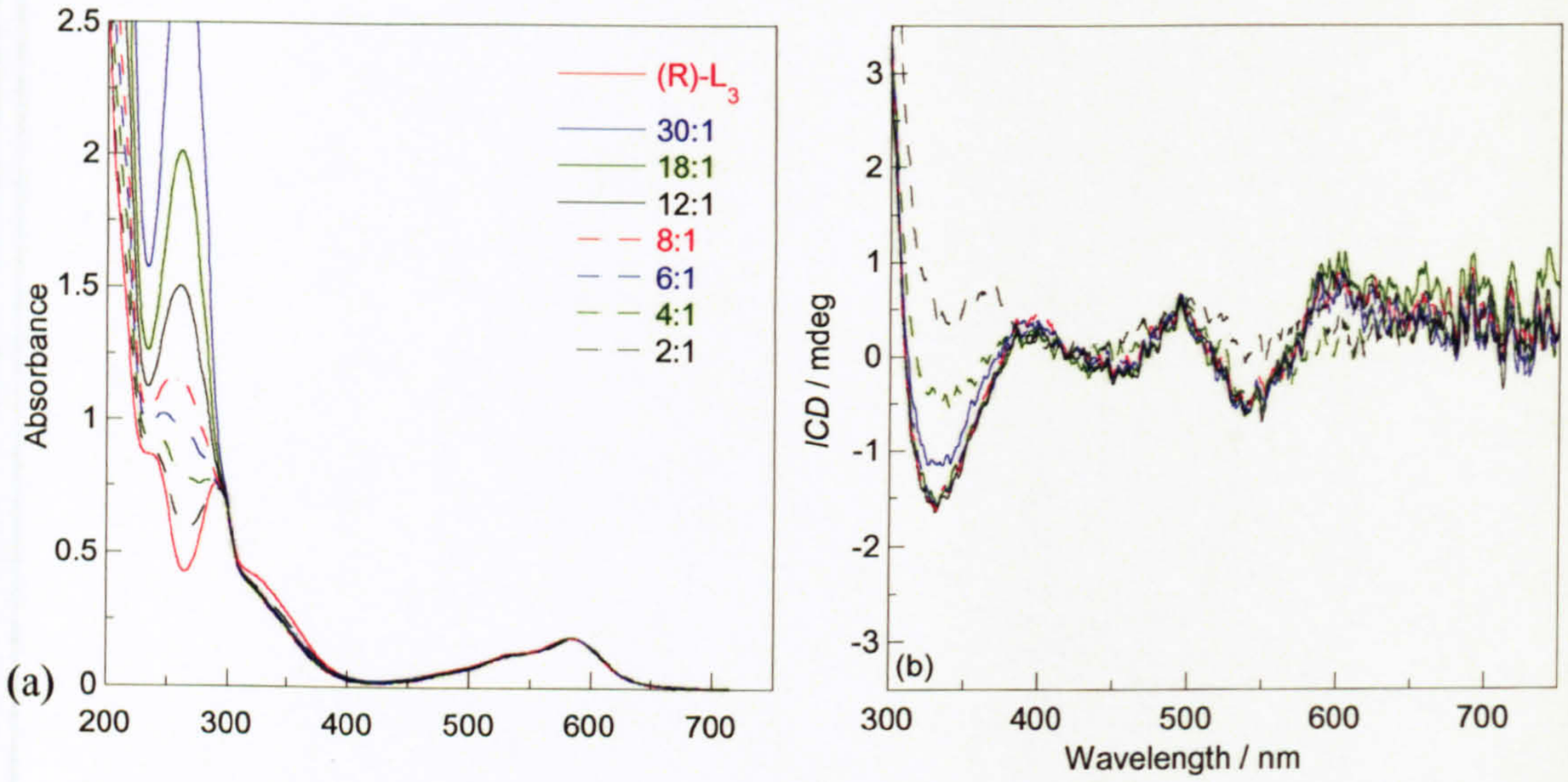
(P)-$L_3$ data table

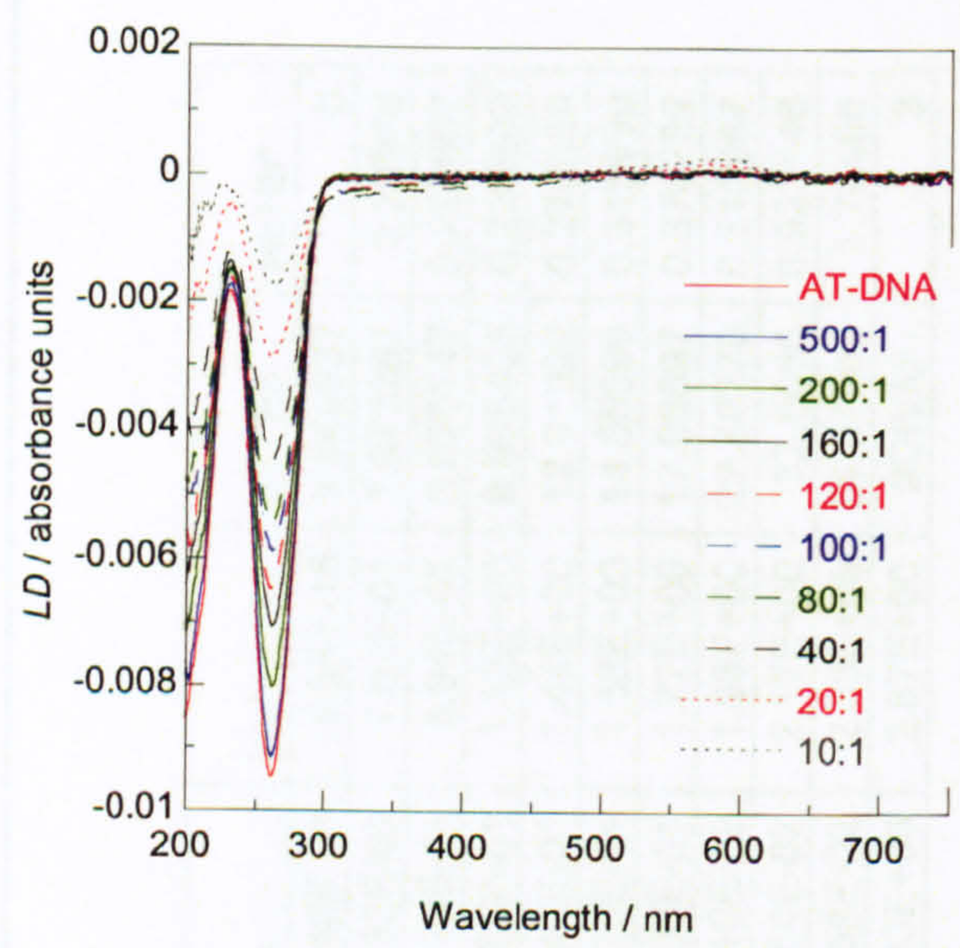
| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 90.747 | 1.931074 | 1.31E+01 | 2.74E-01 | 4.62E-01 | 6.767698 | 0.346349 |
| 0.2 | 87.357 | 1.786134 | 1.32E+01 | 2.50E-01 | 6.30E-01 | 7.398024 | 0.317352 |
| 0.267 | 83.61 | 1.625931 | 1.34E+01 | 2.25E-01 | 8.16E-01 | 8.225482 | 0.327367 |
| 0.4 | 78.747 | 1.418014 | 1.36E+01 | 1.93E-01 | 1.06E+00 | 9.578179 | 0.379035 |
| 0.5 | 74.957 | 1.255972 | 1.37E+01 | 1.69E-01 | 1.24E+00 | 10.94294 | 0.402749 |
| 0.667 | 70.523 | 1.066396 | 1.39E+01 | 1.42E-01 | 1.46E+00 | 13.06607 | 0.457306 |
| 1 | 64.602 | 0.813244 | 1.42E+01 | 1.06E-01 | 1.75E+00 | 17.44466 | 0.572322 |
| 2 | 55.592 | 0.428021 | 1.46E+01 | 5.44E-02 | 2.18E+00 | 34.04504 | 0.915645 |
| 4 | 49.714 | 0.176707 | 1.48E+01 | 2.21E-02 | 2.47E+00 | 83.88649 | 1.620821 |
| 8 | 45.581 | 0 | 1.50E+01 | 0.00E+00 | 2.67E+00 | #DIV/0! | 3 |

(P)-$L_5$ data table

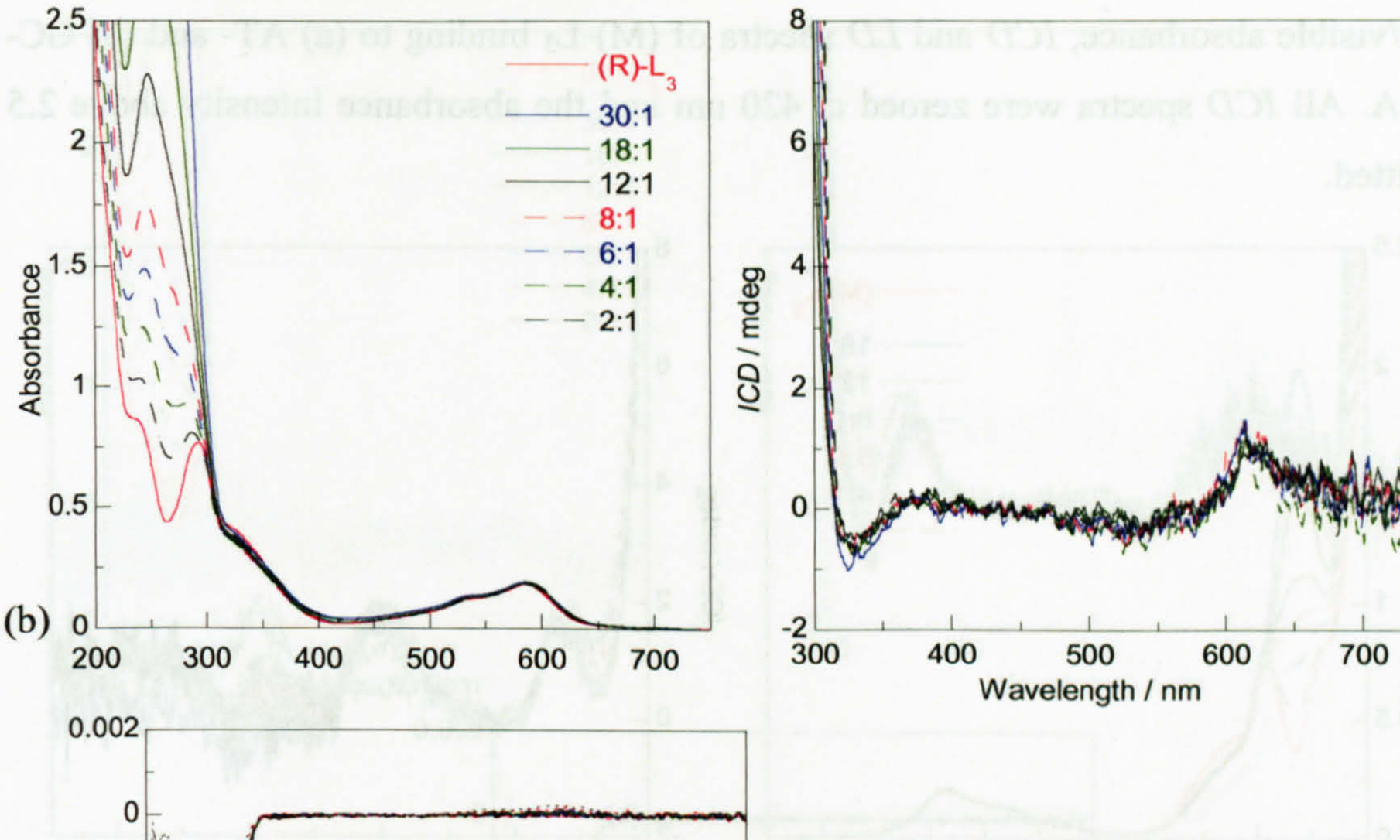
| Ht | F measured | Eb | Ef | Se | Hb* | Ef / Eb | Ht/Hb* |
|---|---|---|---|---|---|---|---|
| 0 | 100 | 2.326686 | 1.27E+01 | 3.40E-01 | -5.55E-16 | 5.446937 | 0 |
| 0.16 | 93.605 | 2.08597 | 1.29E+01 | 2.99E-01 | 2.82E-01 | 6.190899 | 0.568239 |
| 0.2 | 90.814 | 1.980913 | 1.30E+01 | 2.82E-01 | 4.04E-01 | 6.572266 | 0.495067 |
| 0.267 | 87.241 | 1.84642 | 1.32E+01 | 2.60E-01 | 5.60E-01 | 7.123827 | 0.476535 |
| 0.4 | 82.533 | 1.669205 | 1.33E+01 | 2.32E-01 | 7.66E-01 | 7.986314 | 0.522477 |
| 0.5 | 79.574 | 1.557824 | 1.34E+01 | 2.15E-01 | 8.94E-01 | 8.628814 | 0.55914 |
| 0.667 | 74.268 | 1.358099 | 1.36E+01 | 1.84E-01 | 1.12E+00 | 10.04485 | 0.593306 |
| 1 | 67.965 | 1.120846 | 1.39E+01 | 1.50E-01 | 1.40E+00 | 12.38275 | 0.716194 |
| 2 | 56.954 | 0.706377 | 1.43E+01 | 9.15E-02 | 1.87E+00 | 20.23511 | 1.070221 |
| 4 | 49.256 | 0.416614 | 1.46E+01 | 5.29E-02 | 2.20E+00 | 35.00452 | 1.820541 |
| 8 | 41.42 | 0 | 1.49E+01 | 1.51E-02 | 2.53E+00 | #DIV/0! | 3 |

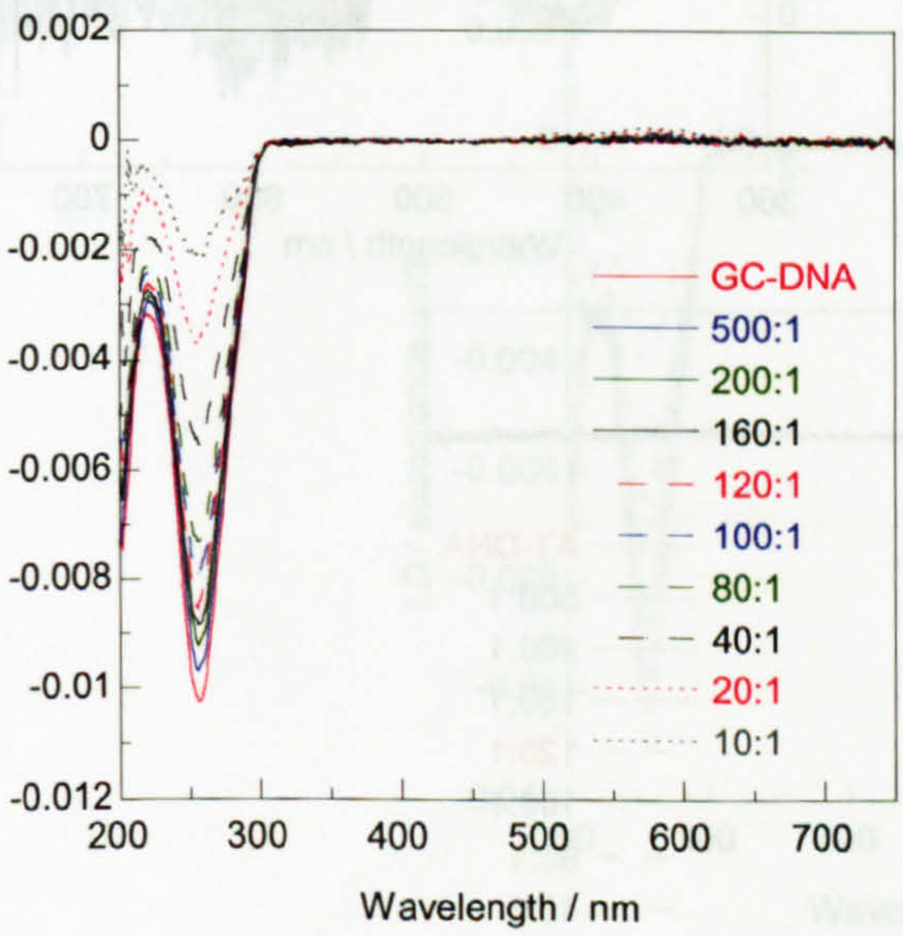
UV/visible absorbance, *ICD* and *LD* spectra of (R)-$L_3$ binding to (a) AT- and (b) GC-DNA. All *ICD* spectra were zeroed at 420 nm and the absorbance intensity above 2.5 omitted.



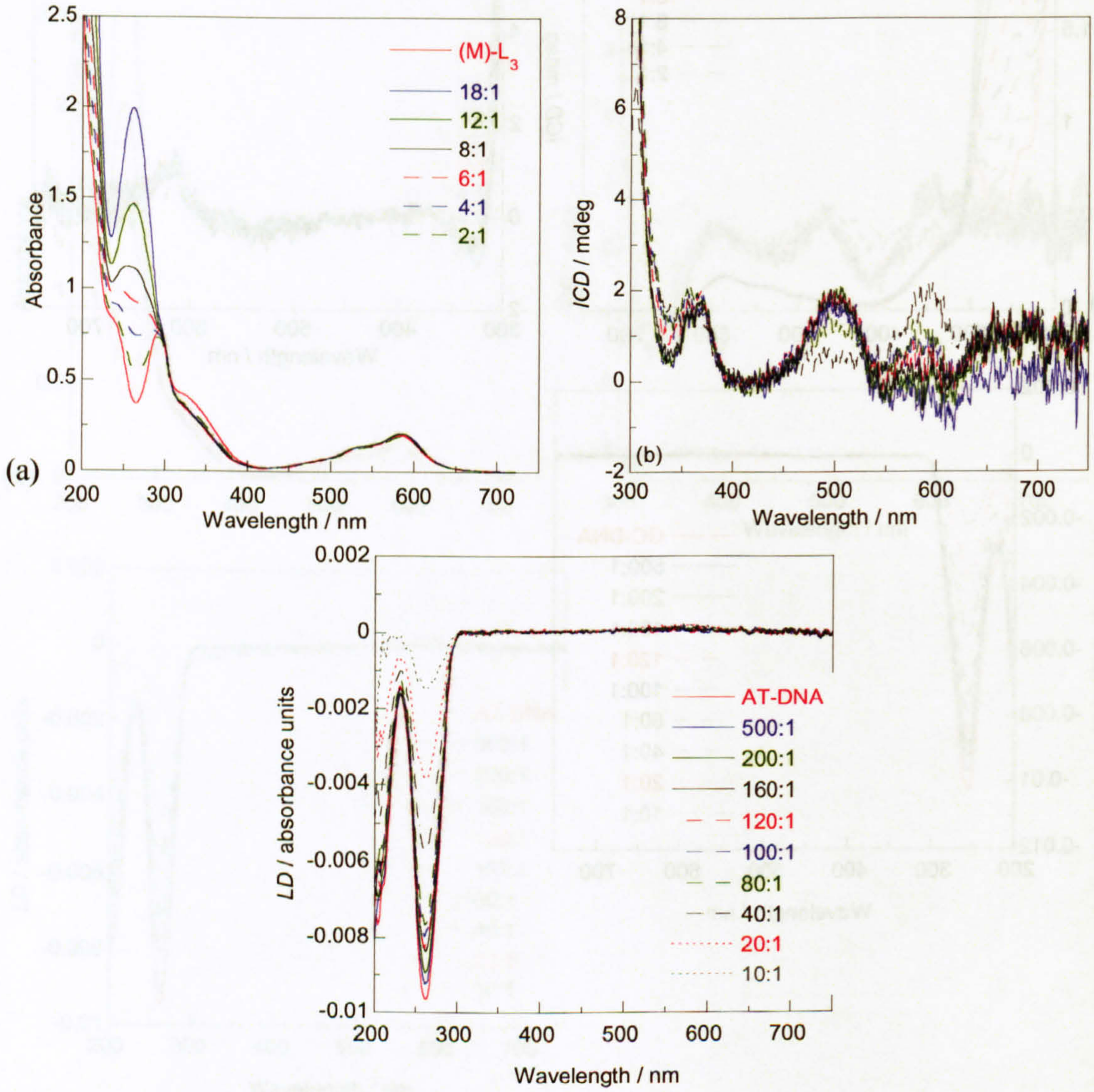



(b)

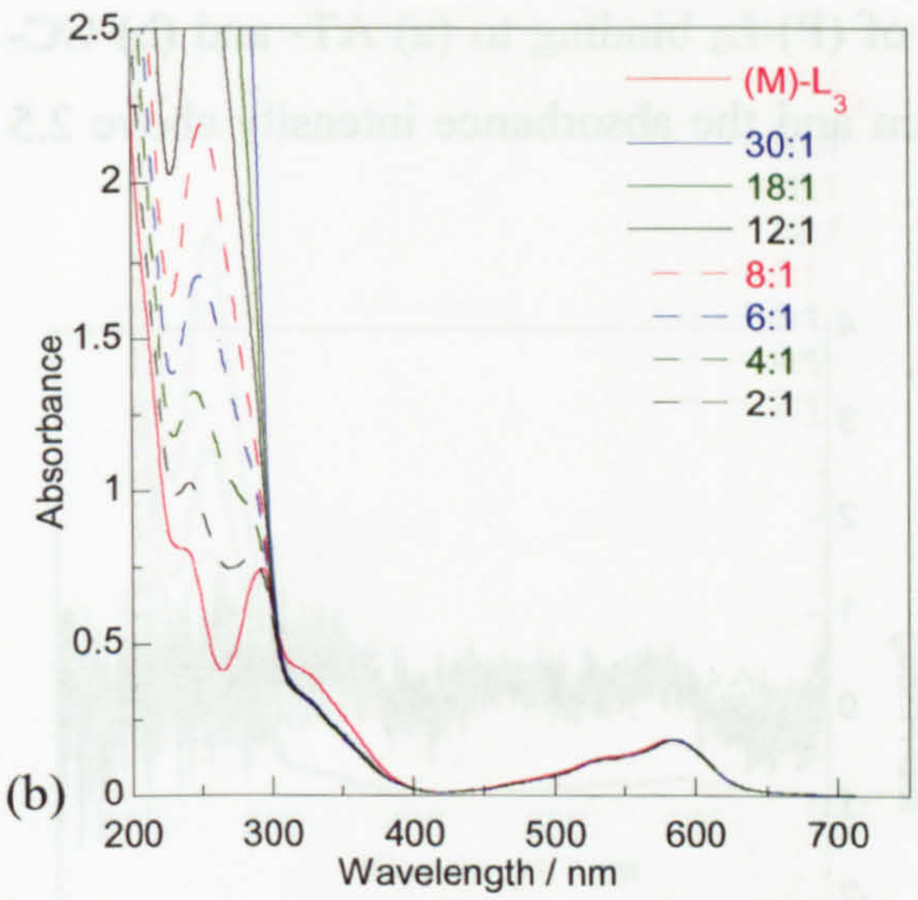

Absorbance
2.5
2
1.5
1
0.5
0
(R)-L3
30:1
18:1
12:1
8:1
6:1
4:1
2:1
ICD / mdeg
8
6
4
2
0
-2
200
300
400
500
600
700
Wavelength / nm

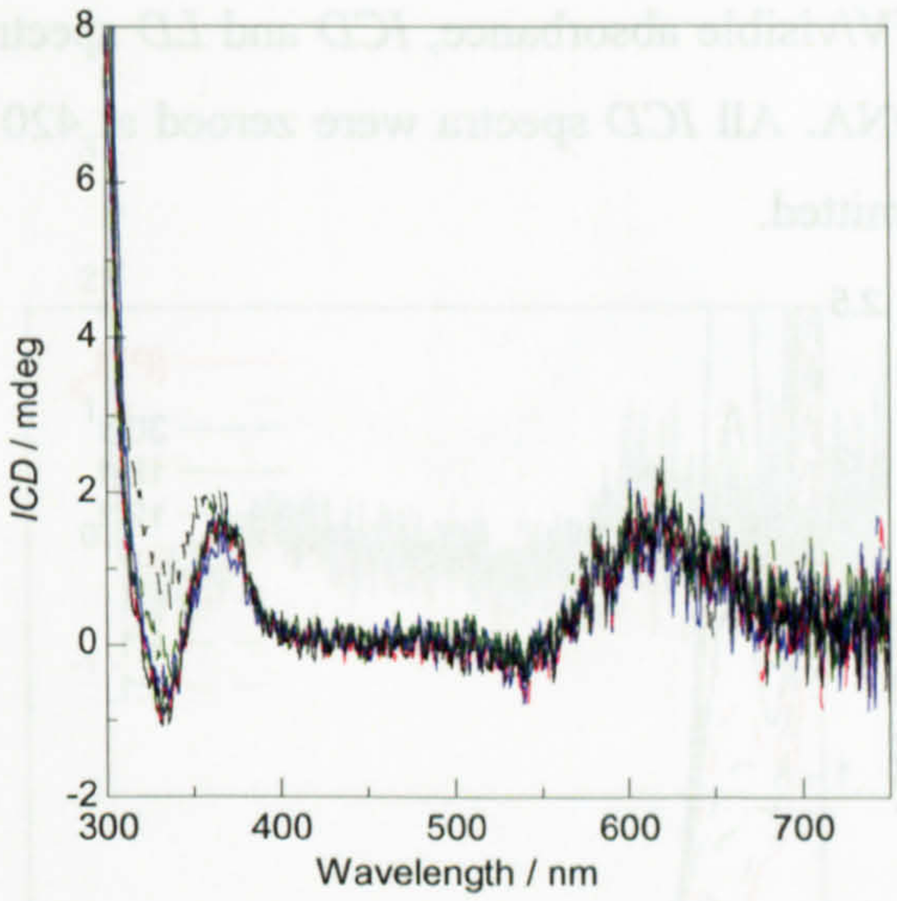

LD / absorbance units
0.002
0
-0.002
-0.004
-0.006
-0.008
-0.01
-0.012
200
300
400
500
600
700
Wavelength / nm
GC-DNA
500:1
200:1
160:1
120:1
100:1
80:1
40:1
20:1
10:1

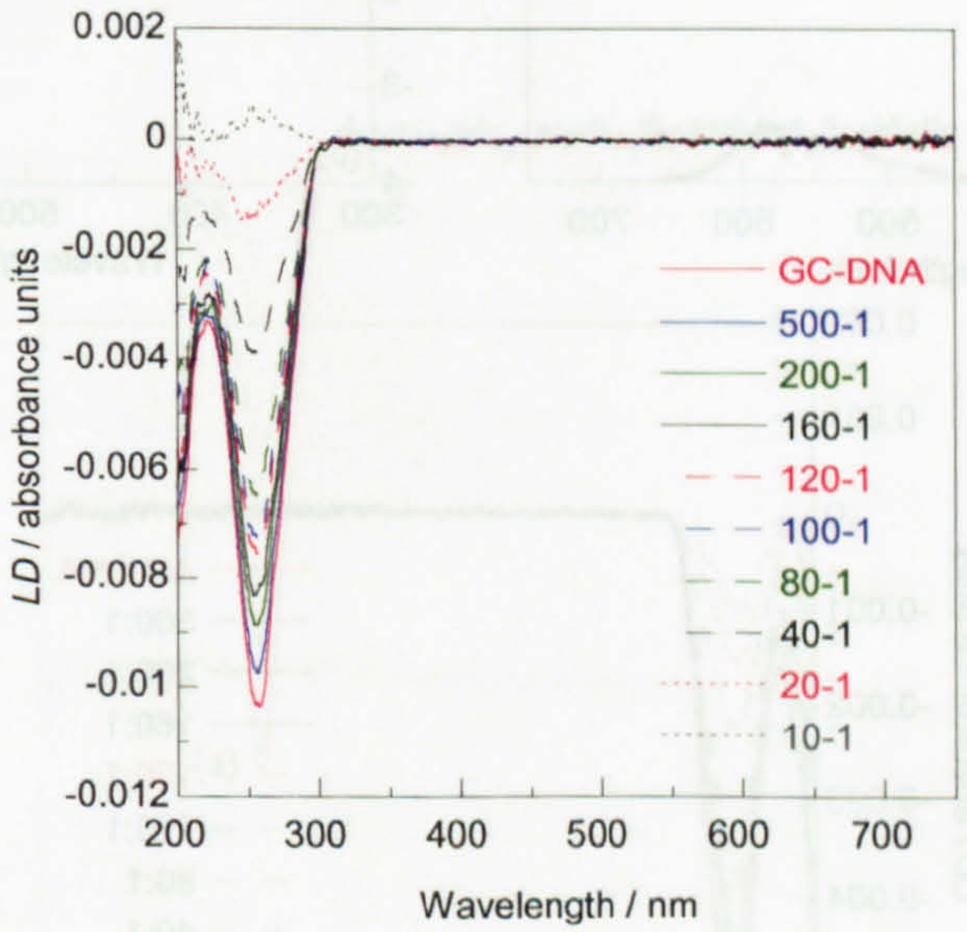
UV/visible absorbance, *ICD* and *LD* spectra of (M)-$L_3$ binding to (a) AT- and (b) GC-DNA. All *ICD* spectra were zeroed at 420 nm and the absorbance intensity above 2.5 omitted.



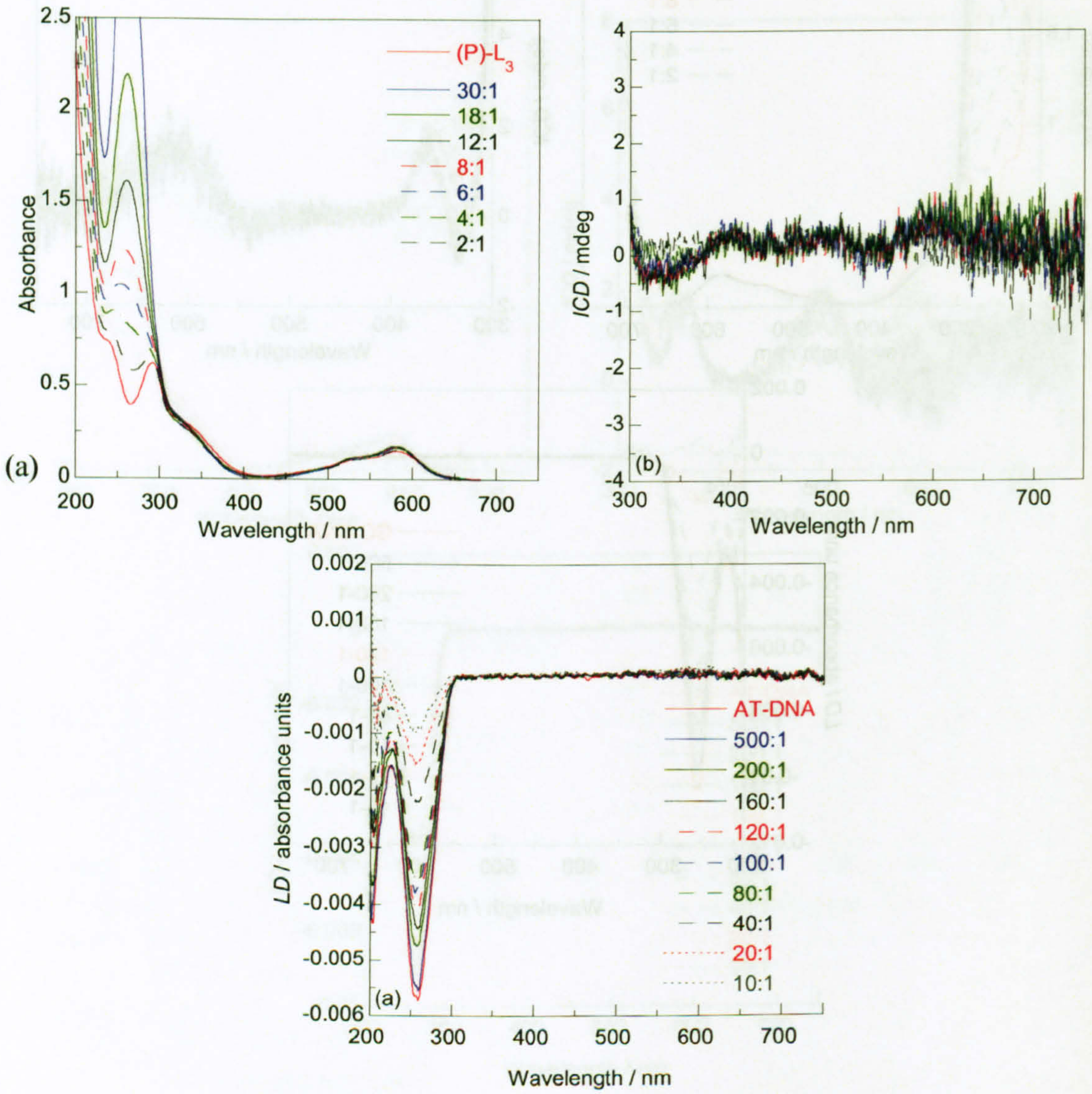
(b)

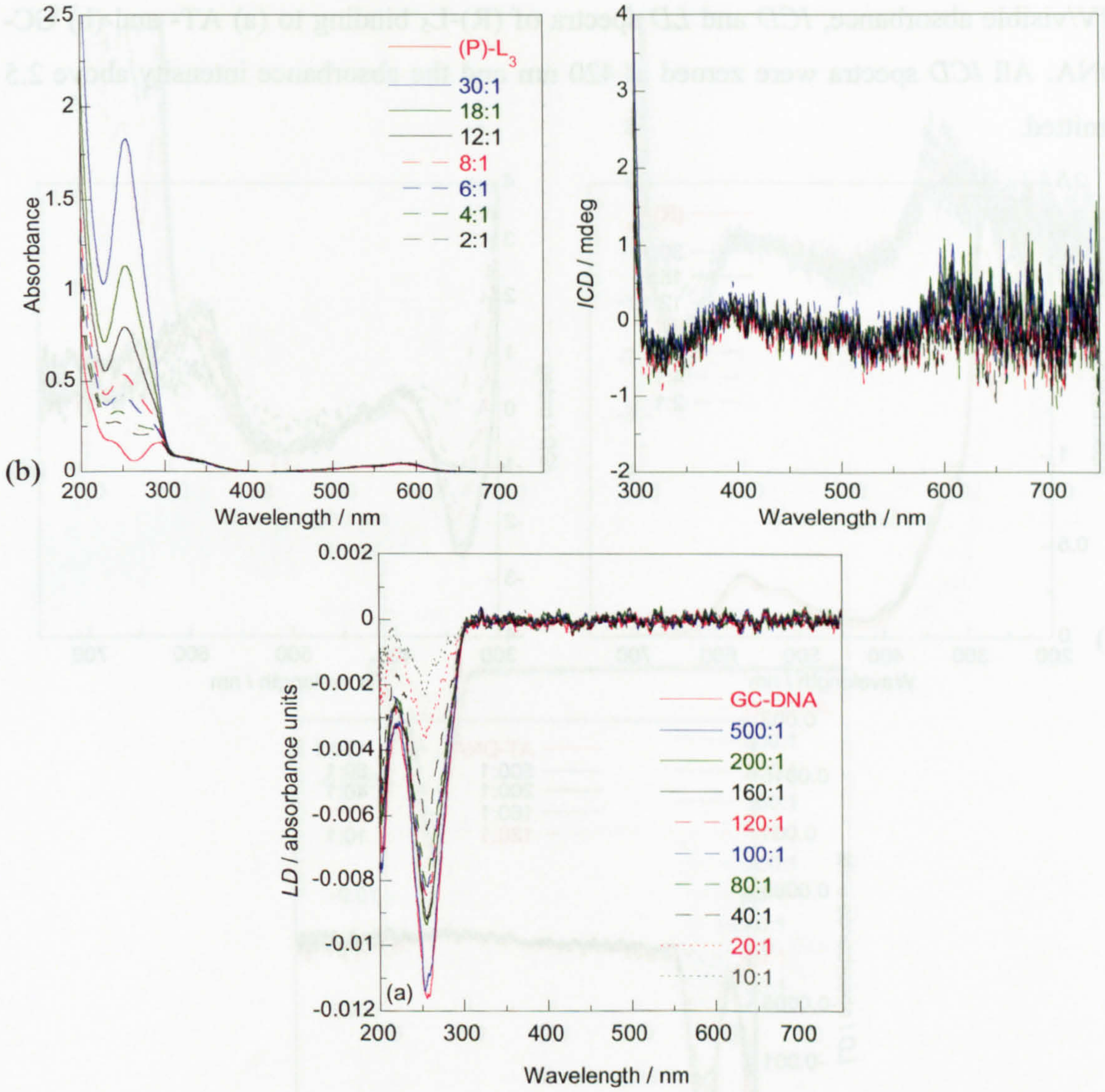

Absorbance
Wavelength / nm
(M)-L3
30:1
18:1
12:1
8:1
6:1
4:1
2:1


ICD / mdeg
Wavelength / nm

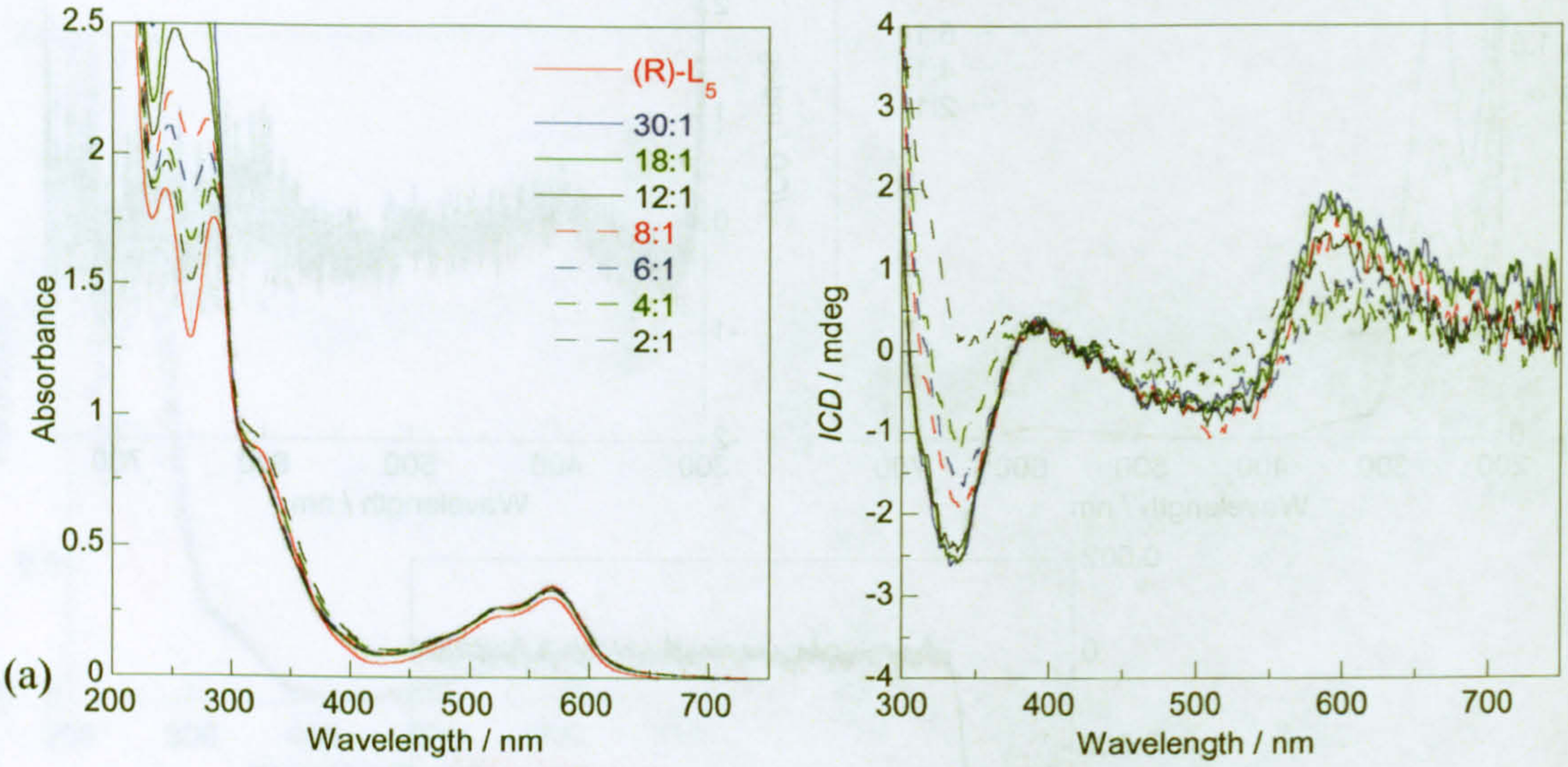

LD / absorbance units
Wavelength / nm
GC-DNA
500-1
200-1
160-1
120-1
100-1
80-1
40-1
20-1
10-1

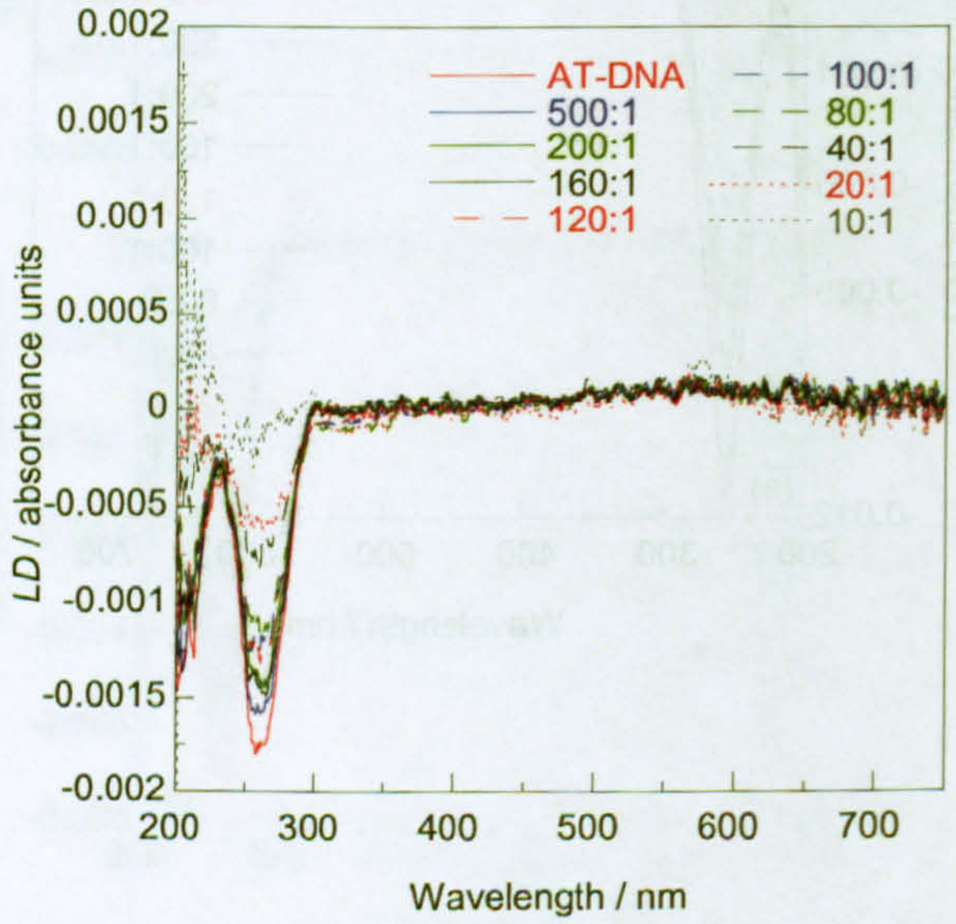
UV/visible absorbance, *ICD* and *LD* spectra of (P)-$L_3$ binding to (a) AT- and (b) GC-DNA. All *ICD* spectra were zeroed at 420 nm and the absorbance intensity above 2.5 omitted.



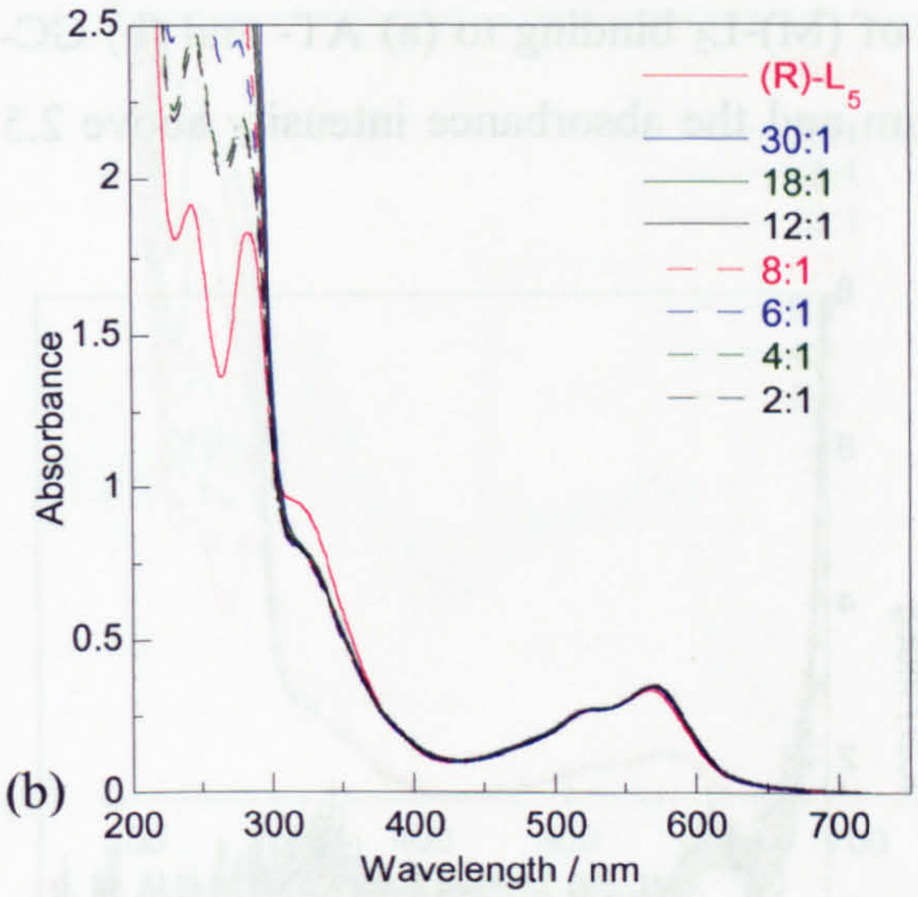

(b)
Absorbance
Wavelength / nm
(P)-L3
30:1
18:1
12:1
8:1
6:1
4:1
2:1
ICD / mdeg
Wavelength / nm
(a)
LD / absorbance units
Wavelength / nm
GC-DNA
500:1
200:1
160:1
120:1
100:1
80:1
40:1
20:1
10:1

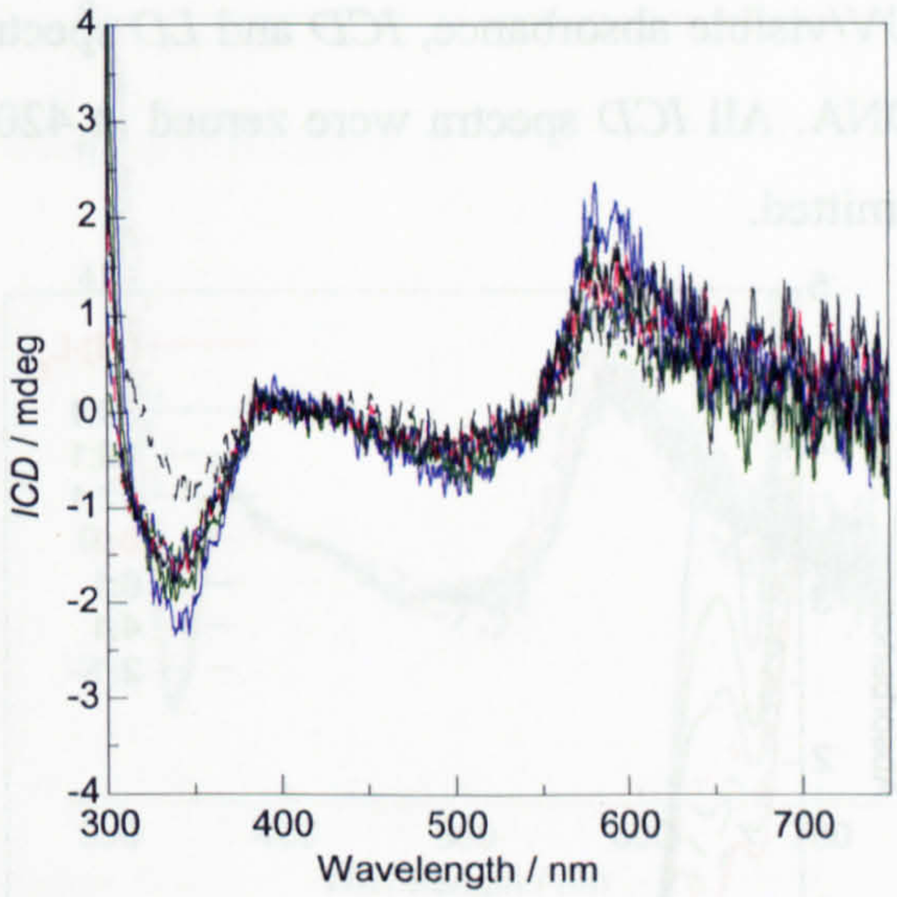
UV/visible absorbance, *ICD* and *LD* spectra of (R)-$L_5$ binding to (a) AT- and (b) GC-DNA. All *ICD* spectra were zeroed at 420 nm and the absorbance intensity above 2.5 omitted.

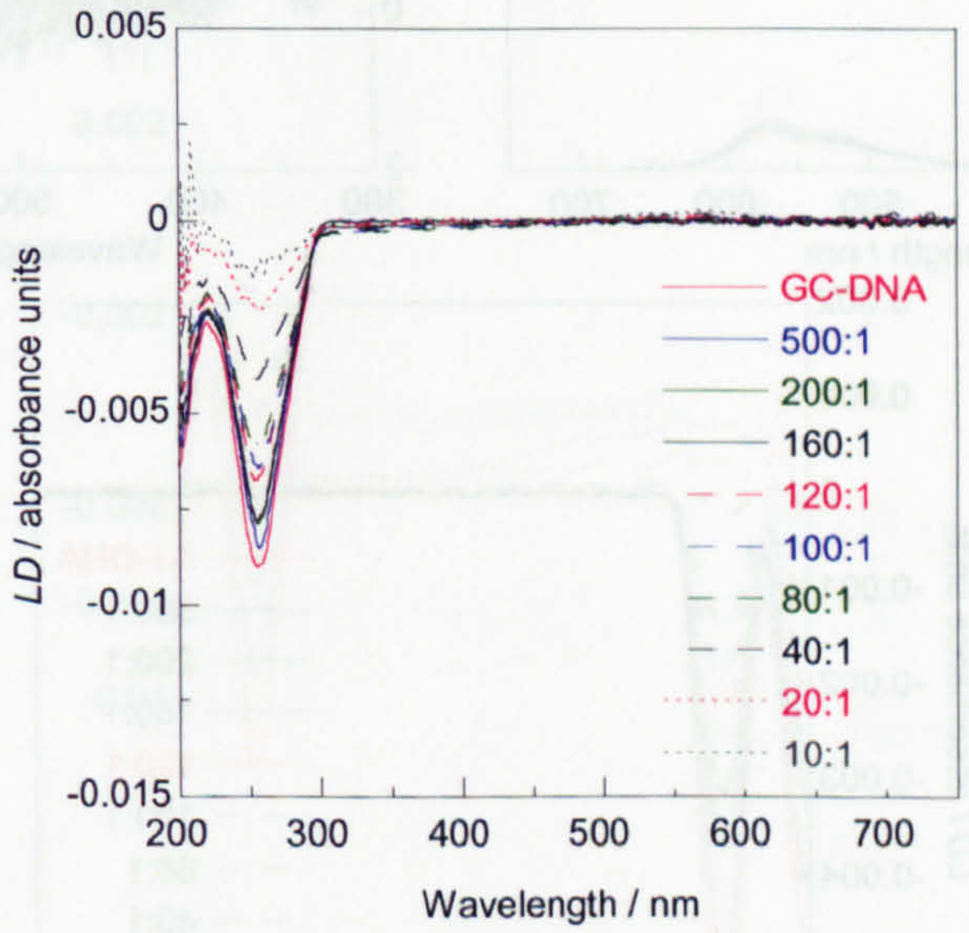


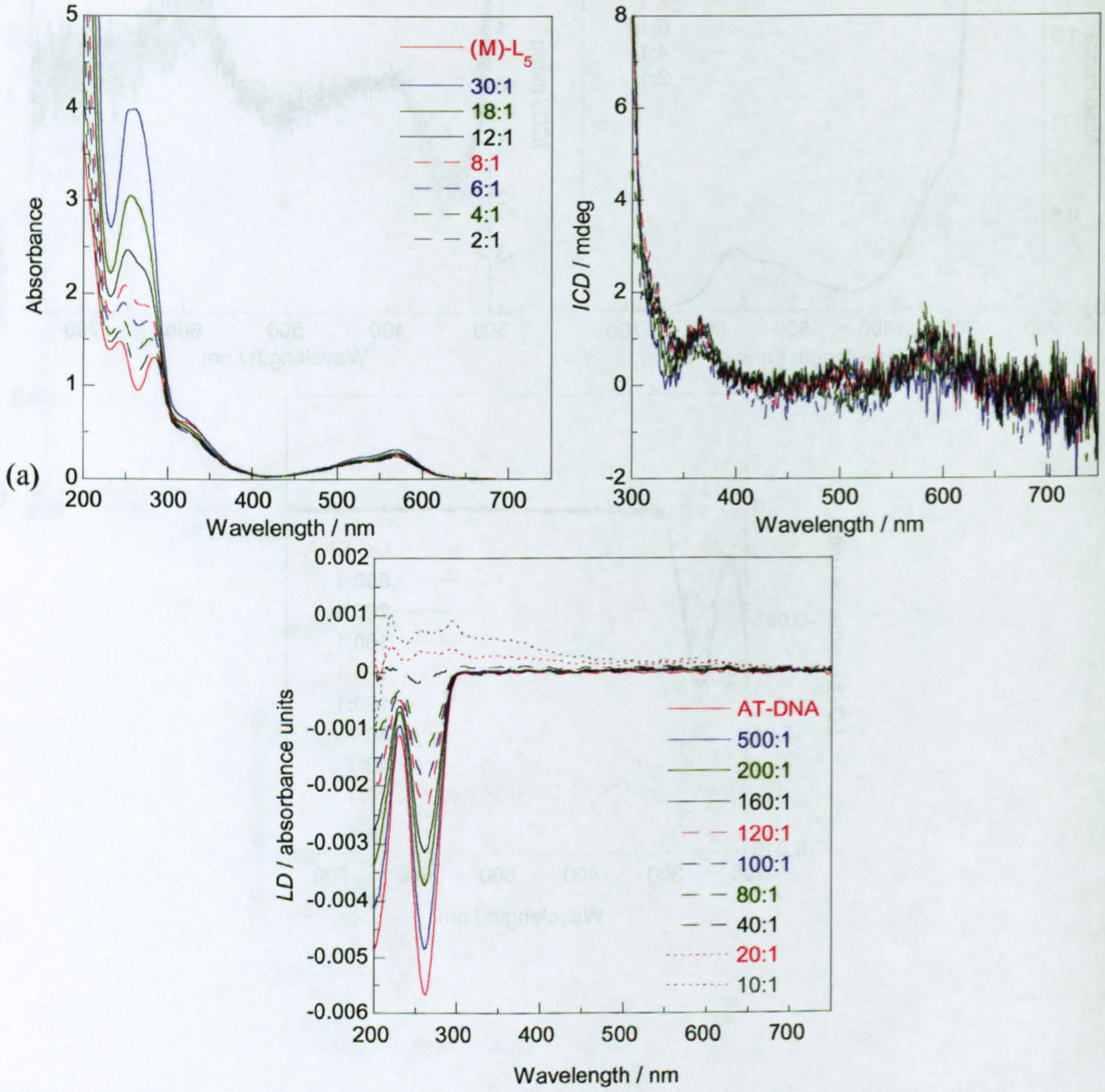
(b)


Absorbance
Wavelength / nm
(R)-L5
30:1
18:1
12:1
8:1
6:1
4:1
2:1

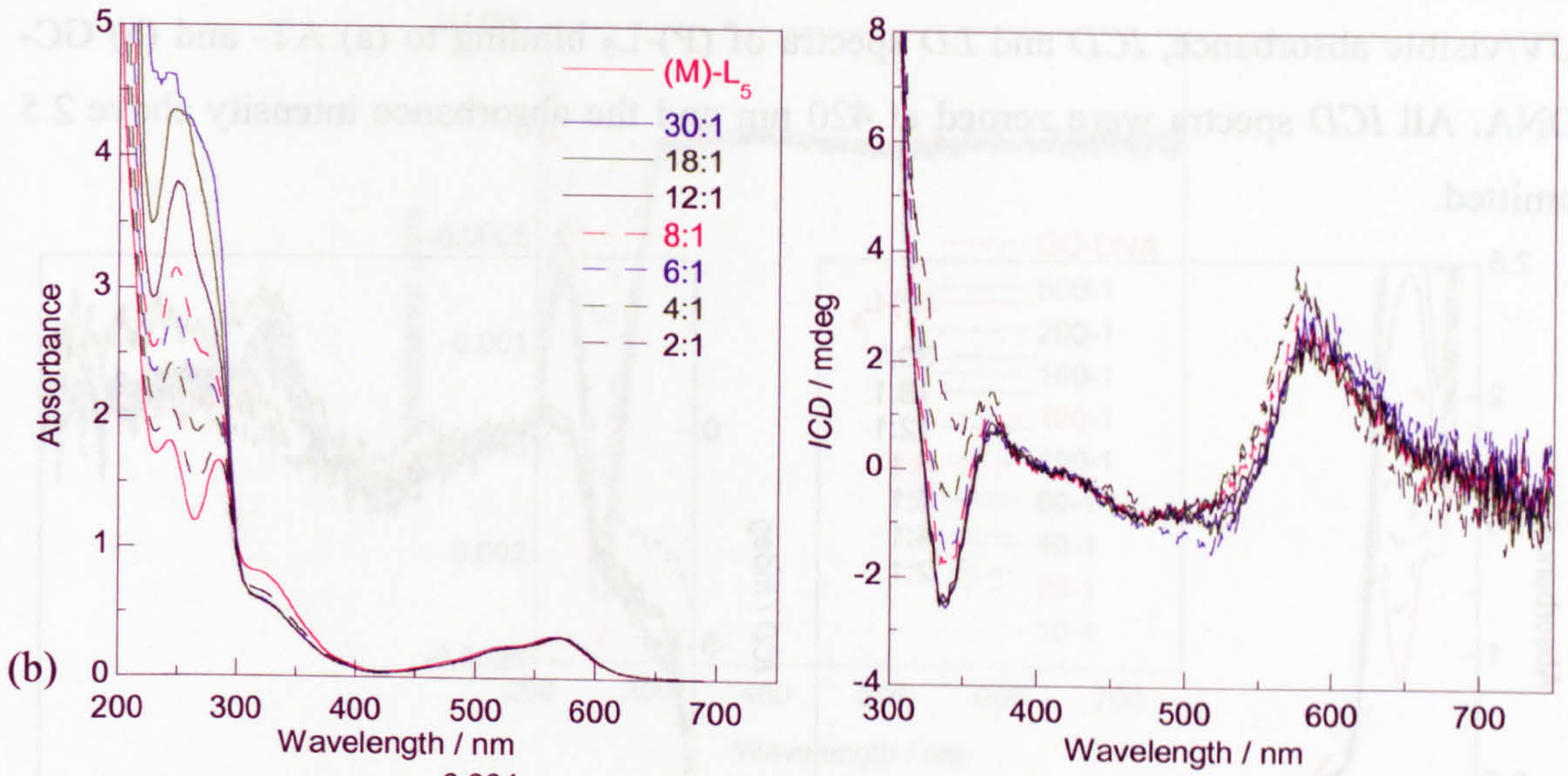

ICD / mdeg
Wavelength / nm

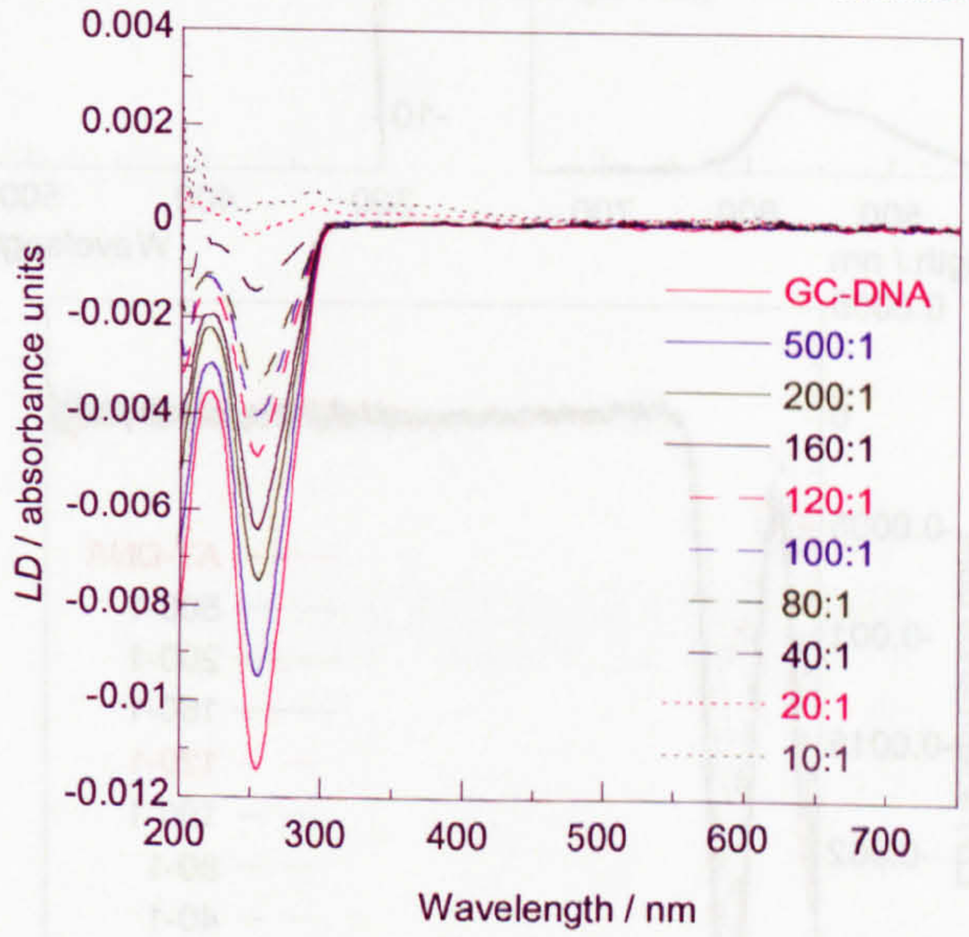

LD / absorbance units
Wavelength / nm
GC-DNA
500:1
200:1
160:1
120:1
100:1
80:1
40:1
20:1
10:1

UV/visible absorbance, *ICD* and *LD* spectra of (M)-$L_5$ binding to (a) AT- and (b) GC-DNA. All *ICD* spectra were zeroed at 420 nm and the absorbance intensity above 2.5 omitted.

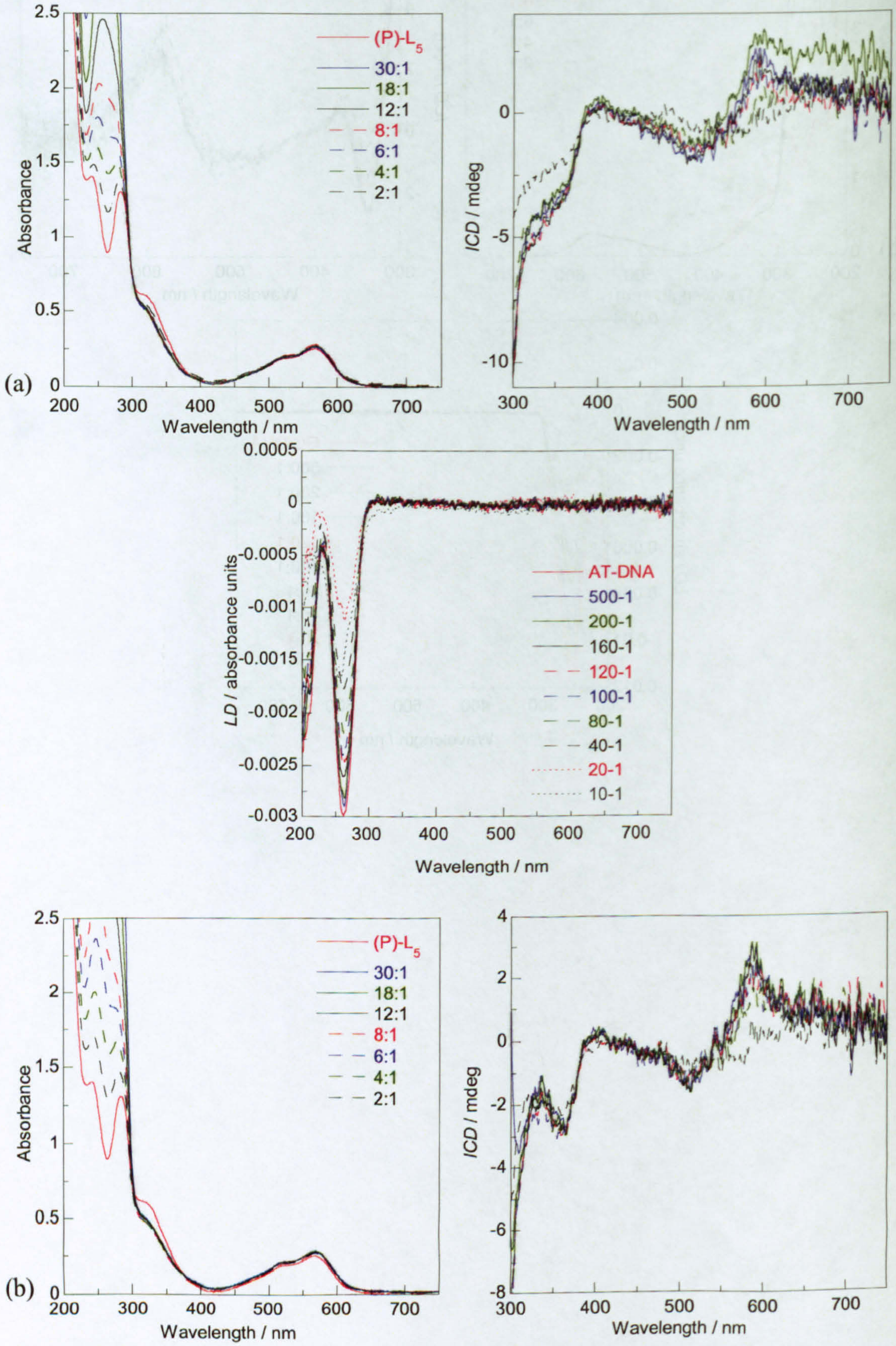


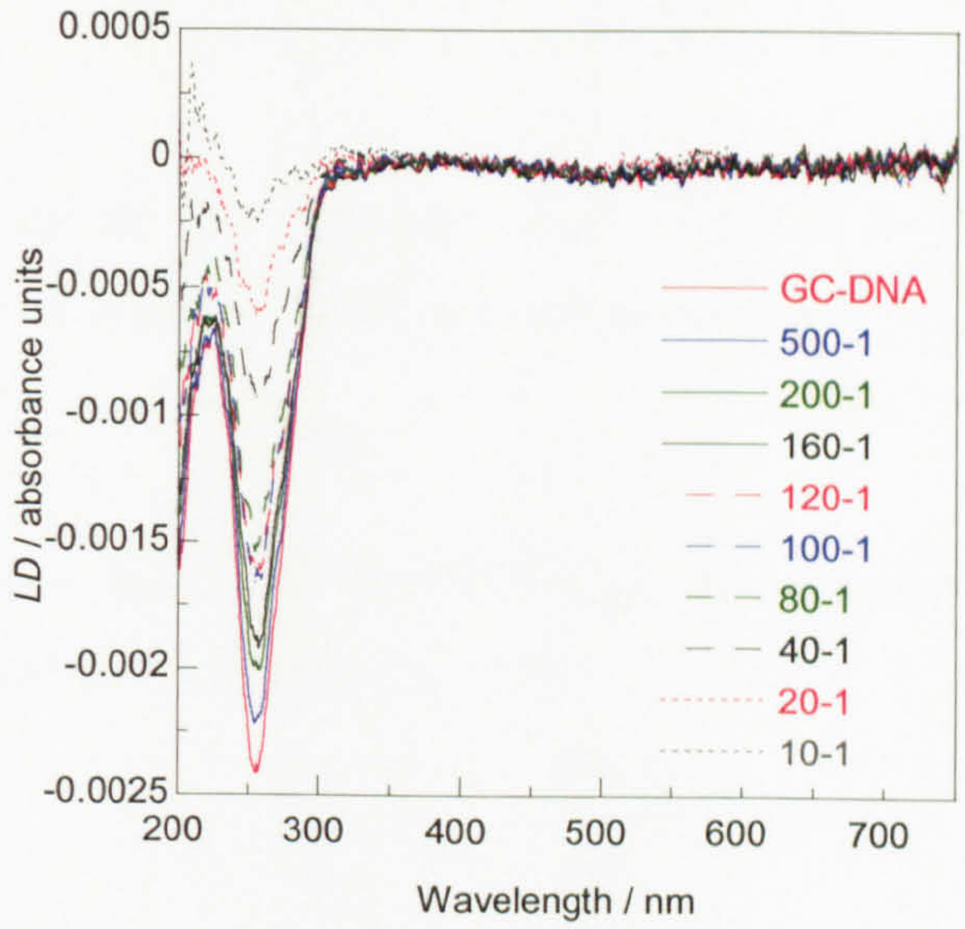
(b)


Absorbance
Wavelength / nm
(M)-L5
30:1
18:1
12:1
8:1
6:1
4:1
2:1
ICD / mdeg
Wavelength / nm


LD / absorbance units
Wavelength / nm
GC-DNA
500:1
200:1
160:1
120:1
100:1
80:1
40:1
20:1
10:1

UV/visible absorbance, *ICD* and *LD* spectra of (P)-$L_5$ binding to (a) AT- and (b) GC-DNA. All *ICD* spectra were zeroed at 420 nm and the absorbance intensity above 2.5 omitted.



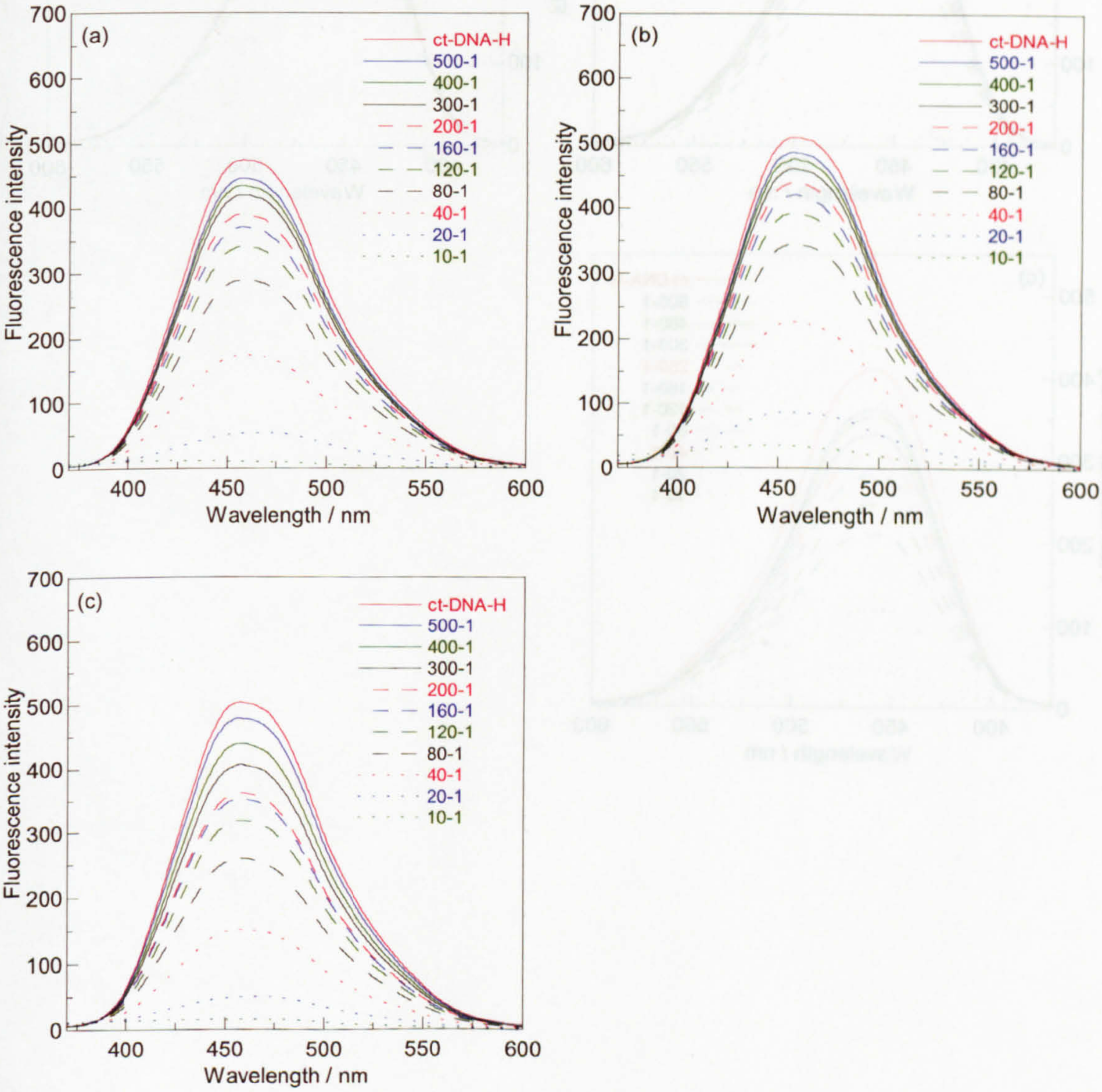

0.0005
0
-0.0005
-0.001
-0.0015
-0.002
-0.0025
LD / absorbance units
GC-DNA
500-1
200-1
160-1
120-1
100-1
80-1
40-1
20-1
10-1
200
300
400
500
600
700
Wavelength / nm

# APPENDIX III

**Appendix III**

The results presented in this appendix were obtained from the investigation of competitive binding of Hoechst 33258 and the enantiomers of metal complexes $L_1$, $L_3$ and $L_5$.

Fluorescence emission spectra of (a) (M)-$L_1$ (b) (M)-$L_3$ and (c) (M)-$L_5$ displacing Hoechst 33258 (8 μM) from ct-DNA (80 μM).



Fluorescence emission spectra of (a) (P)-$L_1$ (b) (P)-$L_3$ and (c) (P)-$L_5$ displacing Hoechst 33258 (8 μM) from ct-DNA (80 μM).