Molecular aspects of biomolecule structure and function

Alison Rodger

Doctor of Science
University of Sydney

2002
1. Introduction

This thesis is submitted in accordance with the By-laws of The University of Sydney relating to the award of Doctorates and additional requirements imposed by Resolutions of the Faculty of Science in relation to the Degree of Doctor of Science.

The By-laws state that the degree of Doctor of Science shall be awarded on the recommendation of the Faculty of Science "for published work which, in the opinion of the examiners, has been generally recognized by scholars in the field concerned, as a distinguished contribution to knowledge." The By-laws require that the principal publications shall have been published at least one year before submission, and shall be a record of original research undertaken by the candidate, who shall state the sources from which this information is derived, the extent to which the work of others has been made use of, and the portion of the work claimed as original. If the publications submitted, whether published in the candidate's sole name or under cojoint authorship, record work carried out cojointly, the candidate shall state the extent to which the candidate was responsible for the initiation, conduct or direction of such cojoint research, however published. Additional Faculty Resolutions require the candidate to describe, by way of introduction, the theme of the published work submitted, and to state how the publications are related to one another and to the theme.

This thesis summarizes the original research work published by Alison Rodger and international collaborators in Australia, North and South America, and Europe over a twenty year period. Two theses previously submitted for the degrees of BSc (Sydney) and PhD (Sydney) are given as publications [1] and [2] respectively. Work arising from the BSc thesis was published in [17]. Publications [19, 20, 21, 22] were the direct result of the PhD thesis work. [24, 25, 26] are based on work initiated during the PhD research and continued independently in Cambridge, UK. The remaining 85 peer reviewed publications are a record of the original research undertaken by Alison Rodger either independently or in collaboration with international scientists. They include 3 books, 2 translations of one of the books, 8 refereed book or encyclopedia chapters and 72 peer reviewed journal articles.

In multi author work in an interdisciplinary area such a biophysical chemistry it is often difficult to attribute the origin of all original ideas. In fact, part of the excitement of this research area is that the ‘whole’ is greater than the ‘sum of the parts’. However, an estimate of the contribution of Alison Rodger to the originality of the published work is given on the right hand side of the publication list. The author or authors who are responsible for the research programme represented by each publication are indicated in bold.

There is no doubt that the research summarized in this thesis would be of much less value without the collaborations that have been a key feature of the research reported. Particularly important people include PE Schipper, who supervised both my Honours project and PhD; BFG Johnson (Cambridge, UK) with whom I developed much of my understanding.
1. Introduction

of transition metal complex and cluster geometries; PM Rodger (Warwick, UK) whose rigorous theoretical approach and molecular modeling expertise have contributed to many projects over the years; B Nordén (Chalmers University of Technology, Gothenburg, Sweden) whose challenges to develop spectroscopic theory have invariably led to new insights; IS Haworth to whom I owe my introduction to biomolecule molecular modeling; and MJ Hannon (Warwick, UK) whose bioinorganic synthetic expertise is the ideal complement to the biophysical research environment I have established at Warwick.

The research is summarized thematically in §3—§7. Viewed chronologically, the first ten years can be seen as establishing a unified theoretical approach to ultra violet-visible spectroscopic and molecular structure techniques for biophysical chemistry. Over the second ten years this foundation has been implemented experimentally with a range of case studies and more general method and instrumentation development. The net effect of these phases of activity is internationally recognized expertise in spectroscopy of biomacromolecules, particularly circular dichroism and linear dichroism of nucleic acids and proteins and their interactions with substrates.
2. Themes of the research

All biological processes are fundamentally inter-molecular interactions. In order to understand, and hence control, biomolecule structure and function, methods are required that probe biological systems at the molecular level, ideally with those molecules being in their native environment. The research summarized herein has at its core the development and application of ultra violet (UV)-visible spectrophotometric techniques for this purpose, in particular circular dichroism (CD) and linear dichroism (LD) but also absorbance, fluorescence and resonance light scattering. The spectroscopy is complemented by fundamental theoretical work on molecular structure and reactivity that forms the basis for designing molecules to bind to biomolecules for a particular structural or functional effect.

A brief summary of the contributions of the listed publications to our understanding of ‘Molecular aspects of biomolecule structure and function’ is given below under five headings:

Circular dichroism theory
Molecular geometry and reactivity
Small molecule-macromolecule interactions: spectroscopic probes of inter-molecular geometries
Molecular design for nucleic acid structure and control
Spectroscopic probes of biomolecule structure: instrumentation and application

In general terms these correspond to successive phases of the research programme, however, all areas have been present since the first publications in 1983 and can be traced weaving through all subsequent activity.
3. Circular dichroism theory

Prior to the circular dichroism (CD) linear dichroism (LD) theory developed in references [17, 19, 20, 21, 23, 24, 25, 30, 31, 32, 38, 41, 42, 44, 47, 48, 52, 53, 57, 69] and unified in [6] ‘Circular dichroism and linear dichroism’ published by Oxford University Press in 1997, the CD literature was confused by conflicting notations and often erroneous work that was accepted as correct. One does after all have a 50:50 chance of getting the right sign (if not magnitude) for a prediction CD since the definition of circular dichroism is

$$CD = A_\ell - A_r = (\varepsilon_\ell - \varepsilon_r) c \ell$$

where $A_\ell$ is the absorbance of left circularly polarized light (and correspondingly $r$ stands for right circularly polarized light), $\varepsilon_\ell$ is the extinction coefficient for left circularly polarized light, $c$ is the concentration of the sample and $\ell$ is the path length of sample through which the light passes. [6] has become the standard text book in many research laboratories round the world for CD and LD theory and applications.

The first CD theory paper ‘Symmetry rules for the determination of the intercalation geometry of host/guest systems using circular dichroism: A symmetry adapted coupled-oscillator model’ [17] was based on an honours project and the subsequent work undertaken by AR under the supervision of PE Schipper at Sydney University. The focus of this work was on CD arising from the coupling of electric dipole transition moments in different chromophores where some of the chromophores were related by rotational symmetry. However, it also provides a general formalism for all electric dipole coupled oscillator CD systems. The applications of this paper included molecules binding intercalatively to DNA (i.e. sandwiched between two base pairs by untwisting the DNA helix and creating a pocket) and also metal to ligand charge transfer (MLCT) and in-ligand transitions of tris chelate metal complexes. This research was the beginning of what became a general independent systems perturbation theory approach to CD theory.

These two themes of transition metal complex spectroscopy and DNA binding ligands probed by CD are present in nearly every publication listed in §8 and in combination form the most exciting aspects of the research that has been accomplished. The theory of tris chelate transition metal complex CD upon DNA binding was extended in collaboration with B Nordén [38]. Molecular modeling studies with IS Haworth and WG Richards [45, 48] on the same system was the first molecular modeling research undertaken. An extensive CD and LD experimental and theoretical project on derivatives of [Ru(1,10-phenanthroline)_3]^{2+} binding to DNA was subsequently published in [70]. Among the most recent publications are studies of the DNA binding and structure control by a new class of dimetallo self-assembling triple helicate metal complexes [72, 75, 82, 88] synthesized by MJ Hannon. These molecules can be viewed as two transition metal tris chelates complexes joined together (Figure 1). These
molecules are currently being designed to have particular DNA structure control effects in a sequence specific manner. A European Research Training Network involving five other European Laboratories has just been funded for this project, following the initial collaborative research reported in [75, 88]. Fundamental to this project is the symmetry adapted approach to spectroscopy theory initiated in [17].

![Image of a molecule]

*Figure 1* Di-iron metallo helicate [Fe$_2$(LL)$_3$]$^{4+}$, synthesis and structure.

The CD of carbonyl n→π* transitions (Figure 2), and in particular the empirical octant rule, was well established when the formal CD theory work summarized herein was initiated, though its wider context and limitations had not been appreciated until [31] was published. In reference [24] a general formalism for the CD of all magnetic dipole allowed transitions located in an essentially achiral chromophore was applied to a range of systems of different symmetry. This, among other things, established the limitations of the octant rule for carbonyl n→π* transitions and generated a firm theoretical basis (and hence understanding of its limitations) for the dominant E-band rule for magnetic dipole allowed d-d transitions in tris chelate metal complexes. The symmetry adaptation methodology underlying this paper, as well as the derivations of coupled-oscillator CD and dispersion [19, 20, 30] and associated induced CD [23] mechanisms, is given in [21, 25]. Planar zigzag carbonyls were analyzed in collaboration with MG Moloney in reference [44].

Applications of the independent systems perturbation theory for electric dipole allowed transitions to oxiranes [32, 42] and for magnetic dipole allowed transitions of carbonyls requiring two substituents for chirality [41] were published in collaboration with A. Gedanken.

In the work reported in [19, 20], sugars were used to induce chirality and hence CD into the transitions of achiral molecules, and the signs and magnitudes of these signals were used to assign transition moment polarizations as a prelude to using a wide range of spectroscopic techniques in other work such as DNA binding studies. The underlying principles developed in that work have led to recent work with chromatographic enantiomeric separations using...
different sugars. In particular, rather than using sugars to induce CD into transition metal complexes, cellulose has been used to resolve the bimetallo tris helicates illustrated in Figure 1. All other literature methods that were investigated had proved to be unsuccessful for resolving the compounds. Needless to say such a simple chromatographic method has attracted considerable interest. A molecular modeling project has also recently been undertaken in collaboration with PM Rodger and AJ Clark following preliminary indications that the induced CD of sugars and monomers might correlate with favourable chiral interactions on sugar-based high performance liquid chromatography stationary phases [86]. A previous collaboration with PM Rodger involved probing solvent contributions due to chiral packing about the analyte [52, 57]. Such issues are crucial when one begins to use spectroscopic techniques to probe biomolecules that are dissolved in solvents and interacting with all species present in solution.

![Diagram](image)

**Figure 2** (a) Axis system used in the octant rule. According to the octant rule, the carbonyl CD induced by substituents in any given octant has the sign opposite from the xyz product of its coordinates with an r−4 distance and polarisability dependence. (b) β-equatorially substituted adamantaneone indicating the net perturber in bold, the induced CD from such a perturber is positive in sign.

When using spectroscopic techniques in general and CD in particular to study biomolecules one can often use the data empirically to deduce that something has occurred or to determine a binding constant (as discussed below), however, ideally to relate biomolecule structure to spectroscopic observables one would measure the spectrum and determine the molecular geometry by self consistent calculations of the CD spectrum. In collaboration with B Nordén [47, 53] calculations of DNA CD spectra were used to assess the validity of literature spectroscopic parameters and to probe the binding of a range of intercalators and groove binders to DNA. 9-hydroxyellipticine experimental and calculated CD signals were used to guide molecular modeling studies of its DNA binding [67]. More recently the same Matrix Method approach has been used in collaboration with JD Hirst to determine polarizations of transition moments of proteins and protein structural motifs in order to use LD to study orientations of membrane proteins in model bilayer membrane systems [90] as discussed below (Figure 3).

Magnetic CD (MCD), whereby a CD signal is induced into achiral molecules in solution by application of a magnetic field, is a somewhat specialized technique. This is in part because
3. Circular dichroism theory

of the instrumentation required and in part because the formalism that is commonly used is unique to MCD. During writing ‘Circular dichroism and linear dichroism’ [6] it became apparent that MCD should be amenable to the same formalism as used for all other CD theory in the book. This was published in [69] and formally brings MCD into the wider CD community.

![Diagram showing transition moments for different structures](image)

Figure 3 Transition moment for (a) α-helix, (b) β-sheet, (c) polyproline II helix, (d) tryptophan, (e) tyrosine and (f) adenine. Such data are essential for theoretical analysis of both CD and LD spectra in biological systems. They can be determined by a mixture of calculation and experiments, such as stretched film LD.

The applications of CD theory not outlined above are in the context of small molecules binding to biomacromolecules and are discussed below.
4. Molecular geometry and reactivity

The constraints imposed on spectroscopic transitions by the symmetry of a molecule are well known and were the starting point for much of the CD theory outlined above. It was shown formally that one could work just with operators rather than the actual transition moments to determine symmetry adapted equations for CD for different systems using the independent-systems perturbation approach discussed above. This successful application of symmetry, however, raised the whole question of why do molecules so often adopt high symmetry structures? Since the molecular symmetry is determined by (or determines) the symmetry of the electron density of a molecule, the same principles controlling electronic spectroscopy must also underlie and control molecular geometry and reactivity. In the foundation work undertaken at Sydney it was shown [22, 26] that high symmetry arrangements of atoms in molecules are either high or low energy arrangements. Hence the highest possible symmetry is often the ground state arrangement of atoms, and the symmetry changes along a reaction pathway are related to the energy changes. This work can be seen to follow that of Jahn and Teller and also Bader and Pearson.

The potential of the symmetry selection rules for reaction mechanisms with application not only to organic molecules but also to transition metal complexes and transition metal cluster compounds became clear in the late 1980’s [9, 10, 27, 28, 29, 33, 34, 35, 36, 37, 39, 40, 49, 51]. Much of this work benefited from a period of close collaboration with, among others, BFG Johnson at Cambridge. As with the CD theory, a unified nomenclature and approach enabled application to all classes of molecules with symmetry elements. Transition metals afford the widest range of high symmetry systems and so this area was particularly fruitful. For example it enabled the relative merits of the Bailar and Ray Dutt twists as mechanisms for the isomerisation of tris chelate transition metal complexes to be assessed and quantified [34] (Figure 4). This shape flexibility and a unified symmetry based approach for understanding how molecular geometry changes during a reaction of any kind, including binding to a biomacromolecule, has been used subsequently as the underlying structural basis for much of the biophysical work of the last ten years. This phase of molecular geometry work culminated in the publication of ‘Molecular Geometry’ by Butterworth-Heinemann [4] in 1995 co-authored by PM Rodger.

An apparently miscellaneous collection of papers and books resulting from associations with co-workers in particular times and places can be seen to fit in with the overall approach to molecular geometry and structure. The first paper published [16] was with LF Lindoy on kinetics of macrocycles; in the same year a small article on the hydrolysis of the chiral molecule sucrose monitored polarimetrically resulted from frustration with an undergraduate practical [18]; [50] arose from work to develop a means of showing school children and undergraduates how the structure of solid and liquid water significantly affected its behaviour,
4. Molecular geometry and reactivity

in particular in biological systems. ‘Foundations of Physical Chemistry’ [5] and its Japanese [6] and Spanish [7] translations took the approaches to molecular geometry developed in a research context back to the late school/early university level by providing a consistent approach to foundations of physical chemistry.

Figure 4 Con concerted symmetry allowed rearrangement mechanism for tris chelate transition metal complexes [34].
The paper ‘Enantioselective DNA binding of [Ru(1,10-phenanthroline)3]2+ studied with linear dichroism’ [38] published in 1990 in collaboration with B Nordén was a very natural development of the previous work on circular dichroism spectroscopy and molecular geometry. Nordén and Hiort had collected an extremely good set of spectroscopic data on the binding of the two enantiomers of [Ru(1,10-phenanthroline)3]2+ to DNA. Due to the literature debate that had grown on the subject more than simple data interpretation was required. The theoretical linear dichroism (LD) work in this paper was the first LD research I published. Rather than the publication being an end in itself it can be seen to be the beginning of a new phase of bioinorganic research activity involving both new CD and LD theory. The programme has ultimately led to the high throughput/minimum volume CD and LD research currently being undertaken as well as the above mentioned design of bimetallo helicates for sequence selective DNA structure control. LD is the differential absorption of light polarized parallel to an orientation axis and light polarized perpendicular to an orientation axis

$$LD = A_\parallel - A_\perp$$

For the LD work on DNA-[Ru(1,10-phenanthroline)3]2+ systems reported in [38], the orientation was provided by couette flow (Figure 5) where the shear forces align the long DNA molecules perpendicular to a light beam that is incident radially. Any molecule bound to DNA is also oriented, so if its transition moment polarizations are known then its orientation (though not position) on the DNA may be deduced. [38] was a significant contribution to the debate on this particular system and also laid foundations to an approach of spectroscopic analysis of DNA-small molecule systems. One of the issues of debate at the time was the polarization of the longest wavelength transitions. The answer required by the CD spectroscopy [17] proved to be at odds with later film LD studies on this class of molecules by Nordén and colleagues.
5. Small molecule-macromolecule interactions: spectroscopic probes of inter-molecular geometries

We have very recently resolved this debate with the realization [89] that CD and LD are dominated by different transition polarizations as not all transitions even of chiral molecules are CD active.

Many aspects of the research reported in the subsequent publications can be traced to reference [38]. No commercial LD cell was available at that time nor, until the publication of [11] were diagrams for construction of such a piece of equipment available. So, having ensured that the CD spectropolarimeter that arrived in Oxford in 1991 had the required electronics for LD, a couette flow cell was designed and built in Oxford as described in [11]. This cell has been extensively used for over ten years in the work that is outlined below.

Recently LD cell design has become a key feature of research activities. In collaboration with Crystal Precision Optics (Rugby, UK), commercial LD couette flow cells have been made available. The construction of a new LD cell (designed as described in [90] and shown in Figure 6) with a very narrow annular gap of 50 µm, rather than the previous 500 µm gap which is typical of cells used in other laboratories has taken LD experimentation a significant step forward. This cell also has calcium fluoride optics with the goal of exploring the possibilities of infra red flow LD. The very short path length of the calcium fluoride cell means that less absorbing sample is placed in the light beam which extends the accessible wavelength range, especially of light scattering samples. In contrast to normal spectroscopic applications, the usual Beer Lambert attenuation of signal when the path length is decreased is avoided since the increased degree of orientation due to viscous drag compensates for this.

![Figure 6](image.png)

**Figure 6** 50 µm annular gap, CaF<sub>2</sub> window couette flow LD cell [90].
5. Small molecule-macromolecule interactions: spectroscopic probes of inter-molecular geometries

Publications [45, 48] can be traced directly back to [38] and the potential of molecular modeling that became apparent from the collaborative work with WG Richards and IS Haworth on spermine/DNA interaction [43, 54]. In [45, 48] we undertook the first molecular dynamics and general molecular modeling study of a DNA-small molecule system where the calculation was constrained by CD and LD data in particular. In 1999 [70], a combined experimental and molecular modeling study of methyl derivatives of [Ru(1,10-phenanthroline)₃]²⁺ helped to identify that at least some of the literature conflict on the DNA binding of this system was due to significant differences in binding mode as a function of concentration. [70] contains the first ‘high concentration’ molecular modeling studies of metal complexes binding to DNA, showing how close packing can dramatically affect the energetically most favourable binding modes. In [85] the role of DNA in catalyzing the racemisation reactions of the labile analogue of [Ru(1,10-phenanthroline)c₃]²⁺, [Fe(1,10-phenanthroline)c₃]²⁺, was subsequently explored. It is now clear that the DNA catalyses this racemisation reaction. This is probably the first example of DNA catalyzing a reaction.

The value of complementary spectroscopic studies, theoretical interpretation of these data and molecular modeling are best illustrated in [55] where work on an anthracene derivatised spermine molecule is reported. The molecular modeling developed from that of publications [43, 54] mentioned above. By adding the anthracene to spermine, a chromophore was gained as well as a potential intercalator. However, at the concentrations required for both the DNA and the ligand to be probed and binding equilibria to be operative, the 254 nm anthracene band that is under the DNA absorption had to be the probe transition. In addition a 1 mm path length was required for the experiments. This led to the development of a simple titration technique for keeping one concentration constant rather than either making up independent solutions or diluting the analyte by adding a stock solution — both of which methods lead to greater experimental errors. Further, it was the first time LD data had been analyzed in a way that the ligand LD under the DNA band was extracted and evaluated independently from that of the DNA. This programme of work in collaboration with IS Haworth and IS Blagbrough continued through publications [56, 59].

Similarities the DNA structural effects induced by [Co(NH₃)₆]³⁺ and by spermine ([NH₃(CH₂)₃NH₂(CH₂)₃NH₂(CH₂)₃NH₃]₄⁺) or spermidine ([NH₃(CH₂)₃NH₂(CH₂)₄NH₃]₃⁺) had already been established. So a natural development of our spermine/DNA programme was to explore the effects of cobalt ammines on DNA [68, 72, 74, 81]. More recently, the supramolecular tris helicates of Figure 1 that can be viewed as two transition metal tris chelate complexes joined together [72, 75, 88] are being found to have dramatic DNA structure control effects as well as the potential for sequence specific binding (see above and below). The DNA binding modes proposed in [88] for the two enantiomers of the parent tris helicate are shown in Figure 7.
DNA binding molecules come in many shapes and sizes, though most have a positive charge and many, though not all as the spermine work illustrates, have spectroscopic chromophores. In addition to being convenient handles for biophysical spectroscopists, the chromophores are often the key functional groups in the interaction of DNA binding molecules with DNA. This is most obvious with planar aromatic molecules that intercalate due to favourable π-π stacking of the ligand with the DNA bases. Thus spectroscopic techniques such as CD and LD as well as absorbance have proved valuable tools in probing such interactions. Many of the general principles for CD and LD established during the research activity outlined above have been summarized in [6] and in the articles or book chapters [12, 13, 14, 15].

If the DNA binding ligand is fluorescent, as is often the case with aromatic molecules (though not the DNA bases themselves), then fluorescence may be a good tool. Following work by RF Pastermack it was also found that resonance light scattering, whereby excitation and emission monochromaters are synchronously scanned, was an invaluable tool for probing the extent of ligand stacking on DNA. In studies of porphyrins binding to DNA, as well as Hoechst 33258, and 9-hydroxyellipticine this facilitated understanding not only of the binding mode but the binding process [64, 67, 68, 71, 87]. This was particularly relevant for the porphyrin where it transpired that the DNA binding mode was dependent on the solution preparation, not just the final conditions. The same problem was encountered when working
5. Small molecule-macromolecule interactions: spectroscopic probes of inter-molecular geometries

with proflavin — sample preparation methodology proved to be the source of the literature discrepancies with this molecule. In proving this an improved dialysis methodology was developed [76].

The above mentioned spectroscopic and molecular geometry perspectives to DNA binding have also found application to other non-metal complex system namely: intercalative and groove binding ligands on duplex and triplex DNA [61, 65], Hoechst 33258 [62], 9-hydroxyellipticine [64, 67], neutral peptides [66], RNA [63], and quadruplex DNA [91].
6. Molecular design for nucleic acid structure and control

The case studies of the range of different DNA binding molecules mentioned above made it clear not only that a range of different binding modes could be adopted, but also that different effects could be induced into the DNA. It was also apparent that transition metal complexes provided a versatile cationic platform for designing new DNA binders. The unified approach to Molecular Geometry that evolves in the book of that title [4] underlies the molecular design of DNA-binding drugs programme currently being pursued in collaboration with the synthetic supramolecular chemist MJ Hannon.

One of the best characterized DNA binding modes is intercalation in which an aromatic ligand inserts between DNA base pairs causing helix unwinding and stiffening. This can also be achieved by planar metal complexes such as the four coordinate platinum terpyridine metal complex [60] synthesized by GW Lowe. In this paper a new method for using fluorescence to determine binding constants was developed as well as applications of CD, LD and gel electrophoresis to probe the intercalative binding. By coupling moieties onto the non-metal side of a terpyridine ligand it is possible to design new DNA binding motifs. In [87] an octahedral ruthenium metal architecture was extended by addition of a range of planar aromatic groups which enabled intercalation or groove stacking, depending on the size and shape of the tail.

The attraction of metal complexes as the basis for DNA ligand design is that the cationic metal can be used to assemble different units to accomplish multi-functional tasks. One goal is to develop systems that would allow generic cellular delivery of metal-containing entities (particularly platinum drugs and radionuclides) to specific tissues. Achieving this requires harnessing and applying of molecular-level recognition events prevalent in (or specific to) the desired tissue type. For example, tissues rich in estrogen receptors (ERs), which include a significant percentage of breast cancers, accumulate molecules that have high binding affinities for these receptors. Therefore, molecules that (i) bind to the ER, (ii) have favourable cellular transport properties, and (iii) contain a second functionality (such as a centre that may be used for diagnostic imaging or medical therapy) are exciting synthetic targets in the field of drug delivery. To this end we have prepared a range of hetero-biomolecular arrays involving a metal and a chelate coupled to 17α-ethynylestradiol or other steroid [80] (Figure 8). The estrogen provides a delivery vector to ER positive cells and the metallo unit a covalent DNA binding motif (by analogy with cis-platin) to inhibit cancer cell replication. The binding to the ER both as isolated receptor, and in whole cell assays has been investigated for a number of the compounds. All the compounds prepared and tested exhibit effective binding to the estrogen receptor and are delivered across the cell membrane into ER positive MCF-7 cells. In the whole cell assays, despite their monocationic nature, the palladium and platinum
complexes prepared exhibit similar (and even enhanced) receptor binding affinities compared to their corresponding neutral free ligands. It is unprecedented for a higher ER binding affinity to be observed for a cationic complex than for its metal-free ligand.

Figure 8 Schematic illustration of a steroidal metal complex (SMC) designed to link a protein to a DNA molecule in a heteromolecular array.

While DNA encodes the essential blueprint for life, within biological systems its structure and function is regulated by proteins. Bio-macromolecules frequently achieve DNA recognition via binding in or around its major groove since the size and shape of the major groove of B-DNA varies more than the minor groove or the backbone with base sequence. Such protein recognition of DNA contrasts with synthetic small molecule recognition agents. Being smaller in scale than proteins, they tend to target the minor groove or act via intercalation and often cause little or no bending of the DNA as is the case with most of the molecules discussed above. Exceptions to the minor groove/intercalation preference include the \([\text{Ru}(1,10\text{-phenanthroline})_3]^{2+}\) enantiomers and derivatives which we showed had a significant, though not exclusive, major groove binding component [70]; the cobalt ammine complexes mentioned above; and the \textit{cis}\textendash\textit{platin} type molecules that bind covalently to the major groove side of guanine.

Our aim was to build on the \([\text{Ru}(1,10\text{-phenanthroline})_3]^{2+}\) results and to develop synthetic agents that target the major groove of DNA with recognition through non-covalent surface motifs. The parent compound of the series is the tetracationic supramolecular cylinder illustrated in Figure 1 that is just the right size to fit into the major groove of DNA and is too big to fit \textit{into} the minor groove (Figure 1, Figure 7). As described in detail in [72, 75, 88] the cylinder does indeed recognize the major groove of DNA and excitingly also induces a dramatic intramolecular bending effect resulting in intra-molecular coils of DNA. Such an effect is unprecedented with a synthetic DNA binder. Furthermore the two enantiomers have been shown to have different and independent effects on the DNA sequence [88].
In order to provide samples of the two enantiomers for biophysical characterization a new form of chiral chromatography was required as all the literature methods were ineffective at resolving the compounds. In this case cellulose (paper in the first instance) and aqueous 20 mM sodium chloride eluent were effective at providing 100% resolution [82]. The main techniques used for the biophysical characterization have been CD and LD, but the complementary information provided by NMR and AFM (Figure 9) and the required collaborations with other laboratories have enhanced the work. The current goal is to programme DNA sequence specificity into the backbone of the cylinders. The potential of a sequence specific DNA binder that locally wraps DNA into a little ball is very significant for regulation of cellular processes of all kinds.

*Figure 9* AFM images of linear plasmid DNA (pBR322, 4361 base pairs, diameter of 475 nm, cleaved with Pst I and Sal I enzymes to give two linear fragments of 1401 base pairs and 2962 base pairs) with the following DNA base: [Fe$_2$(LL)$_3$]$^{2+}$ complex concentration: 0:1 (free DNA), 10:3, and 10:5.
Figure 10  Schematic illustration of the proposed thylakoid membrane insertion mechanism for the precursor protein pre-PsbW indicating a fairly randomly oriented surface bound protein and α-helices in different orientations on or in a lipid bilayer. The mature PsbW (bottom left) has a single trans membrane helix.

CD is the ideal tool for probing the secondary structure of membrane proteins in different environments [77, 83]. However, a key question is often the orientation of the proteins or the membrane spanning motifs. To date this issue has proved an almost insurmountable challenge to biophysicists. LD is generally the ideal tool for answering such questions, but membrane proteins cannot be flow oriented in isolation and in any case need to be measured with reference to their lipid environment. Reference [90] reports the LD of a number of intrinsic or extrinsic membrane proteins oriented by flow orienting liposomes to which the proteins are attached. The experiments are comparatively simple to undertake and in principle can be analyzed to give quantitative measures of membrane protein orientation in bilayers. Even at the level of qualitative analysis, more orientational structural information is immediately available than from other techniques.

In all these applications knowing the protein concentration is an important requirement, especially if protein structure fitting programs are to be applied [77, 83]. If an extinction coefficient is known then UV absorbance is usually the technique of choice. An alternative simple method, if the protein is of known molecular mass and is known to be pure except for solvents of crystallization is CHN analysis as described in reference [79]. Another apparently routine consideration is whether modifications such as di-sulfide bond cleavage have indeed been accomplished. In [78] complementary CD and high resolution mass spectrometry measurements were used to determine conditions for this. Intriguing hints as to the structural roles of disulfide bonds became apparent. In the wheat germ agglutinin studied, these covalent links were there to prevent the protein relaxing into a predominantly α-helical structure.
The biological significance of unusual DNA structures such as quartets, as well as the better known ones such as triplexes and Z-DNA, is just beginning to be established. We have recently shown that the simple assumptions about solution phase quartet structure and observed CD spectra are not correct [91], though the resolution to this dilemma and any relationship between structure and activity remains to be elucidated. A similar level of mystery surrounds the spectroscopy of long naturally occurring RNAs. Literature reports of these are relatively rare due mainly to the difficulty of producing sufficient RNA and keeping it polymeric throughout experiments. A study in collaboration with SF Newbury showed the advantages of CD and UV absorbance melting curves in probing the solution phase structure of a long mRNA of known sequence [63]. Unusual CD spectra (Figure 11) for unusual CNG sequences, compared with isobasic randomized controls, together with gel electrophoresis have recently led to the proposal of a molecular model for RNAs (Figure 12) and perhaps DNAs to account for the trinucleotide repeat expansion diseases such as Huntington’s Chorea [92].

![Figure 11](image-url)

*Figure 11* (A) CD of (CTG)$_5$ DNA (4 µM) at 20°C in 5 mm path length cells in 5 mM sodium phosphate (----), plus 20 mM NaCl (------), and plus 20 mM MgCl$_2$ (- - - - ). (B) CD spectra of chemically synthesised (CUG)$_5$ RNA (8 µM) (-----) and a isobasic randomised control (- - - -) in 5 mM sodium phosphate at 20°C in 5 mm path length cells. (C) Effect of temperature on (CUG)$_5$ RNA: 5°C (-----), 25°C (- - - -), 45°C (--- ---), 65°C (- - - -), 80°C (------).
Flow LD has been applied in only a limited way as so few laboratories have the sample orientation equipment as discussed above. Recent research has been to develop new applications of LD in addition to the membrane protein application outlined above. An example reported in [73] is to determining relative orientations of protein structural motifs or chromophores in peptide fibres. This work was undertaken in collaboration with DN Woolfson following a lecture on potential applications of LD spectroscopy given at the Annual CD conference at Daresbury Laboratory, UK. Another collaboration originating from the same forum led to the application of LD to determine the number of single nucleotide polymorphism mutations in polymerase chain reaction (PCR) DNA samples from Addenbrookes Hospital Clinical Biochemistry Laboratory [84]. This application has also been published as Patent 1 and forms the basis of a wide range of new potential applications of LD in research laboratory and clinical diagnosis settings that will require a range of technical developments for thermal cycling, high throughput etc. Work is in progress to develop the required equipment and to make it available to the wider biophysical and diagnostic community.

![Figure 12](image1.png)

Figure 12  (A) Schematic picture of the ‘toblerone’ structure of (CNG)$_7$; (B) Alternate views of a ‘toblerone’ energy minimised structure for 5’-CUGCUG resulting from a conformational search where NOE constraints were initially imposed then subsequently removed. Colour coding of atoms is: C, medium grey; O, black; N, dark grey; H, light grey [92].
8. Publications

% of original work belonging to AR is indicated on the right of each publication. All publications whose content has previously been submitted for a degree are indicated with the degree. Senior authors, with whom the research programme might be deemed to have originated, are indicated in bold.

Theses


Books


Book chapters


8. Publications


**Papers**


(18) **Rodger, A.** "The Kinetics of the Acid Catalysed Hydrolysis of Sucrose" *Chemistry in Australia* 1983, 52, 71–71


(29) Rodger, A., **Johnson, B. F. G.** "Polyhedral Rearrangements of Metal Clusters" 50%
8. Publications

*Polyhedron* 1988, 7, 1107–1120

(30) **Schipper, P. E., Rodger, A.** "A Dispersion Induced Circular Dichroism and Normal Absorption Study of the n-\(\pi^*\) Carbonyl Transition for the Isolelectronic Substituent Series -CH\(_3\), -NH\(_2\), -OH" *Spectrochimica Acta* 1988, 44A, 575–580 90%


(34) **Rodger, A., Johnson, B. F. G.** "Which is More Likely: The Ray Dutt Twist or the Bailar Twist" *Inorganic Chemistry* 1988, 27, 3061–3062 75%

(35) **Johnson, B. F. G., Rodger, A.** "Energetics of Isomerization Reactions in Transition Metal Complexes" *Inorganic Chemistry* 1989, 28, 1003–1006 75%

(36) **Rodger, A., Johnson, B. F. G.** "Structure and Reactivity of Transition Metal Complexes: A Stereochemical Approach" *Polyhedron* 1989, 8, 1742–1744 75%

(37) **Rodger, A., Colwell, S. M., Johnson, B. F. G.** "Bonding in Boranes" *Polyhedron* 1990, 9, 1035–1044 75%

(38) **Hiort, C., Nordén, B., Rodger, A.** "Enantioselective DNA Binding of [Ru(1,10-phenanthroline)\(_3\)]\(^{2+}\) Studied with Linear Dichroism" *Journal of the American Chemical Society* 1990, 112, 1971–1982 25%


(40) **Johnson, B. F. G., Bott, A. W., Hugh-Jones, D., Rodger, A.** "Geometrical and Orientational Isomers in Clusters" *Polyhedron* 1990, 9, 1769–1774 20%


(44) **Rodger, A., Moloney, M.** "n-\(\pi^*\) Circular Dichroism of Planar Zig-Zag Carbonyl Compounds" *Journal of the Chemical Society Perkin II* 1991, 919–925 80%

(45) **Haworth, I. S., Elcock, A. H., Freeman, J., Rodger, A., Richards, W. G.** "Sequence Selective Binding to the DNA Major Groove: Tris,10-phenanthroline) Metal Complexes Binding to Poly(dG-dC) and Poly(dA-dT)" *Journal of Biomolecular Structure and Dynamics* 1991, 9, 23–44 25%


8. Publications


(55) Rodger, A., Blagbrough, I. S., Adlam, G., Carpenter, M. L. "DNA binding of a spermine derivative: spectroscopic study of anthracene-9-carbonyl-N^1-spermine with poly(dG-dC) and poly(dA-dT); Biopolymers 1994, 34, 1583–1593


(62) Kim, S. K., Moon, J.-H., Sehlstedt, U., Rodger, A., Nordén, B. "DNA Structural Features Responsible for Sequence Dependent Binding Geometries"
8. Publications

of Hoechst 33258" Biopolymers, 1996, 38, 593–606


8. Publications


Publications in press for less than 12 months

(85) Rodger, A., Nordén, B., Rodger, P.M., Bates, P.J. “DNA as a catalyst and catalytic template for the racemisation of metal tris-phenanthroline complexes” European Journal of Inorganic Chemistry 2002, 1, 49–53


8. Publications

environments” *PCCP 2002*


Patents

(93) Dafforn, T., Halsall, D., Rodger, A. “Method for determining nucleotide mismatches” 0108329.4 *2001*