PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Recognition of DNA by Octahedral Coordination Complexes

Timothy W. Johann and Jacqueline K. Barton

Phil. Trans. R. Soc. Lond. A 1996 **354**, 299-324
doi: 10.1098/rsta.1996.0010

Recognition of DNA by octahedral coordination complexes

By Timothy W. Johann and Jacqueline K. Barton

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Strategies for the site-specific recognition of DNA are described through the design of a family of octahedral metallointercalating complexes. Coordination complexes, which mimic DNA-binding proteins with regard to their affinity and specificity for DNA sites, may be prepared by the specific functionalization of chiral metal complexes so as to achieve the ensemble of non-covalent contacts in the DNA major groove anchored through intercalation.

Abbreviations

HET	2-hydroxyethanethiolate
phen	phenanthroline
TMP	3, 4, 7, 8-tetramethyl-1, 10,-phenanthroline
Me_2trien	(2R,9R)-diamino-4, 7-diazadecane
pyr	pyrimidine
pur	purine
[12]aneS_4	1, 4, 7, 10-tetrathiacyclododecane
en	ethylenediamine
phi	9, 10-phenanthrolinequinonediimine
terpy	2, 2′, 2″-terpyridine
bpy	2, 2′-bipyridyl
TBP	TATA box binding protein
DPB	4, 4′-diphenyl-2, 2′-bipyridyl
DMB	5, 5′-dimethyl-2, 2′-bipyridyl
MGP	4-guanidylmethyl-1, 10-phenanthroline
phen	1, 10-phenanthroline
dppz	dipyridophenazine
EDTA	ethylenediaminetetraacetic acid
DIP	4, 7-diphenylphenanthroline

1. Introduction

DNA is the blueprint for the great diversity of cells in each of us. Exactly 46 molecules of DNA contain all the information necessary to make a human being, and the same 46 molecules, thus the same information, are present in each cell of that human being. The differences between cells that become brain and cells that become muscle do not arise from differences in the information contained in the DNA of these cells,

Phil. Trans. R. Soc. Lond. A (1996) **354**, 299–324
Printed in Great Britain

but from differences in how that information is read or accessed. In undifferentiated cancerous cells, seemingly insignificant changes in the DNA library cascade into large differences in which sequences in the library are read, and uncontrollable cell growth results.

How this DNA information is accessed is determined in large measure by ensembles of non-covalent binding interactions between DNA sites and proteins, which regulate and control gene expression (Watson *et al.* 1987). Repressor proteins bind selectively to 8–12 base pair sequences of DNA and block the transcription of certain genes. Transcriptional amplifiers, in contrast, in binding to their specific targets, can enhance the expression of selected genes by more than a thousand-fold. Each of these events constitutes an example of molecular recognition. How do these DNA-binding proteins selectively identify one site on the DNA duplex rather than another?

A major focus of our laboratory has been in the design of small coordination complexes as mimics of these large DNA-binding proteins (Dupureur & Barton 1995; Sitlani & Barton 1995). What are the chemical principles that govern site-selective recognition of nucleic acids, and can we construct small synthetic metal complexes rationally which selectively discriminate among sites utilizing these principles? From a practical standpoint, the construction of these small synthetic complexes would lay the foundation for the rational design of new chemotherapeutics targeted to nucleic acids. Fundamentally, in constructing and applying generations of complexes to DNA recognition, we hope to elucidate some of the principles which govern these remarkably selective non-covalent interactions with macromolecules.

Here we describe our efforts to prepare and apply a range of octahedral metal complexes designed for non-covalent interaction with DNA. Over the past 20 years, as chemists have learned to synthesize and manipulate DNA molecules, the field of metal–nucleic acid chemistry has burgeoned (Tullius 1990; Pyle & Barton 1990). Rather than providing an exhaustive review of the literature, we intend here to focus on efforts from our own laboratory in illustrating some of these principles of DNA recognition.

(*a*) *DNA structure and modes of recognition*

As shown in figure 1, DNA is a polymer of nucleotide units consisting of a ribose sugar, a phosphate and a heterocyclic aromatic base (Saenger 1984). The information in the DNA polymer is contained in the ordering or sequence of the four different bases: guanine, adenine, cytosine and thymine. Two of these polymers associate into a double helix via hydrogen bonds between the bases. Adenines associate with thymines and cytosines with guanines in an arrangement referred to as a base pair. As can also be seen in figure 1, this pairing presents two different faces, two unique sets of possible hydrogen bonds and van der Waals contacts available for non-covalent interaction with proteins or small molecules which associate with the DNA polymer. When base pairs are viewed in the context of the DNA double helix, as is also illustrated, the faces of the base pairs combine to form the major and minor grooves of DNA. These grooves differ from each other in width and depth as well as in the number and variety of possible hydrogen bonds and van der Waals contacts.

Historically, it was thought that proteins might recognize specific DNA substrates by matching particular amino acid hydrogen bonds and van der Waals functionalities with those of the individual DNA bases (Seeman *et al.* 1976). This type of recognition was termed direct readout and was supported by the fact that most proteins bind primarily in the major groove of DNA, which has a greater number and variety

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

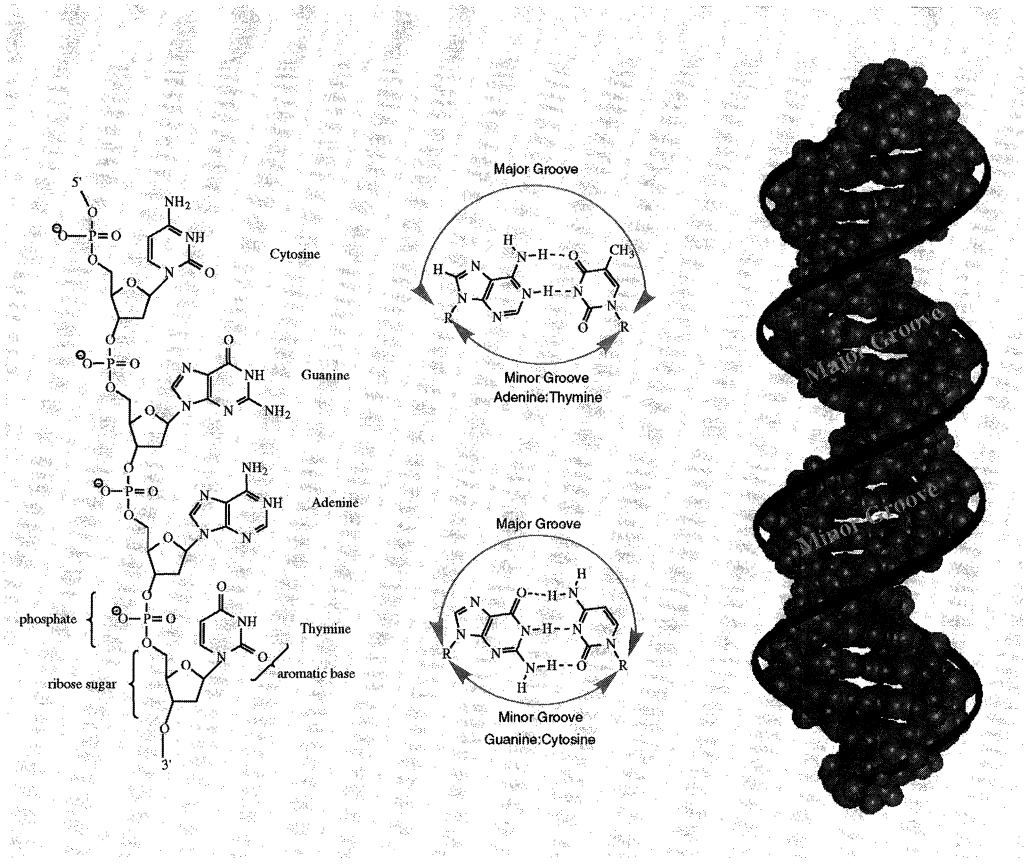


Figure 1. The building blocks of a DNA polymer (left). The major and minor groove sides of adenine:thymine and guanine:cytosine base pairs (middle). Space filling model of the major and minor grooves in a DNA polymer (right). The ribbons trace the phosphate backbone.

of functionalities to specify sequence. Over the past decade, a range of studies have made it increasingly clear that the act of recognition is much more complex (Steitz 1990). Despite a growing body of structural data, no general rules for the recognition of DNA by amino acid side chains have emerged (Pabo & Sauer 1992). Specific amino acids have been shown to recognize more than one base and, inversely, each base has been shown to be recognized by a number of different amino acids. The recognition properties of individual amino acids are largely dependent on their three-dimensional orientation relative to the DNA. These orientations are in turn determined by the local and global structure of the protein. Furthermore, the three-dimensional characteristics of double-stranded DNA are just as complex (Saenger 1984). The sequence of a particular DNA determines its characteristics, such as groove width and depth as well as the positions and orientations of the bases, sugars and phosphates relative to each other in space. The DNA double helix is not simply a uniformly spiraling staircase of linear base pair sequence. Instead, the three-dimensional structure of DNA is polymorphic and in a way which depends upon the sequence (Rich 1993). Recognition which depends upon these local variations in the three-dimensional structure of the double helix has been termed shape selection (Chow & Barton 1992) or indirect

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

readout (Otwinowski *et al.* 1988), and this binding mode plays at least as important a role as direct readout in the recognition of DNA.

(*b*) *The ensemble of non-covalent interactions*

DNA binding proteins tend to have affinities for DNA of at least 10^6 M^{-1}, which corresponds to contacts with the double helix worth 8 kcal more than the contacts they make with the solvent. Hydrogen bonds and methyl–methyl van der Waals contacts are worth approximately 0.25–2 kcal for DNA binding versus water (Krotz *et al.* 1993*b*; Donner *et al.* 1982; Singleton & Dervan 1992). The task of designing a small synthetic molecule which would make enough specific contacts with the complex three-dimensional structure of DNA to achieve high affinity is indeed daunting. Instead, both DNA-binding proteins and their small molecular mimics, in achieving high affinity binding to DNA, take advantage of a range of non-specific interactions, as well as those which specify site contacts. Indeed, as a result of these non-specific contacts, achieving high affinity is not difficult. The challenge lies in obtaining specificity, or in other words a high ratio in affinity constant for specific versus non-specific sites.

An important means of generating non-specific interaction of molecules with DNA is based upon electrostatics. DNA is a polyanion with a −2 charge for each base pair unit. Positively charged molecules will therefore associate with DNA non-specifically on the basis of electrostatic interactions alone. Thus, many DNA-binding proteins tend to be basic, and small molecule mimics are similarly positively charged.

We take advantage of intercalation as a sequence-neutral non-covalent interaction of our coordination complexes with DNA. As first proposed by Lerman (1961), intercalation occurs when a planar aromatic heterocycle inserts itself between the base pairs of the double helix. Figure 2 illustrates the structurally characterized intercalation of Pt(terpy)HET^+ into the dinucleotide $(dCpG)_2$ (Wang *et al.* 1978). Note that the base step is expanded from the normal 3.4 Å distance to 6.8 Å and unwound by 23° to make room for the intercalator. A binding energy of 5–8 kcal is gained by the interaction of delocalized π-orbitals of intercalating molecules with the π-orbitals of the heterocyclic aromatic base stack (Berman & Young 1981). In the 1970s, Lippard and co-workers carried out a series of studies demonstrating that square planar platinum(II) complexes containing aromatic heterocyclic ligands could intercalate into duplex DNA in a fashion reminiscent of classical organic natural product intercalators (Jennette *et al.* 1974; Bond *et al.* 1975). Our studies have been based upon intercalation by octahedral complexes. This expansion of the coordination number converts a two-dimensional interaction to three dimensions.

The intercalation by aromatic heterocycles into dinucleotides has actually been crystallographically characterized in many instances (Berman & Young 1981). Most of these natural products insert into the base stack from the minor groove. An exception is the platinum complex which inserts into the duplex via the major groove. This major groove orientation may very well be a general feature of metallointercalation. Many ^{1}H-NMR (David & Barton 1993; Dupureur & Barton 1994; Rehmann & Barton 1990*a*) and biophysical (Sitlani *et al.* 1992; Krotz *et al.* 1993*a*) studies have consistently demonstrated a similar major groove orientation for the octahedral metallointercalators. In addition to expanding to three dimensions, this major groove intercalation by the octahedral complex has the advantage of delivering functional groups on the coordination complex with stereochemical control to sites in the major groove, where DNA sequence-specifying functional groups are most abundant.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

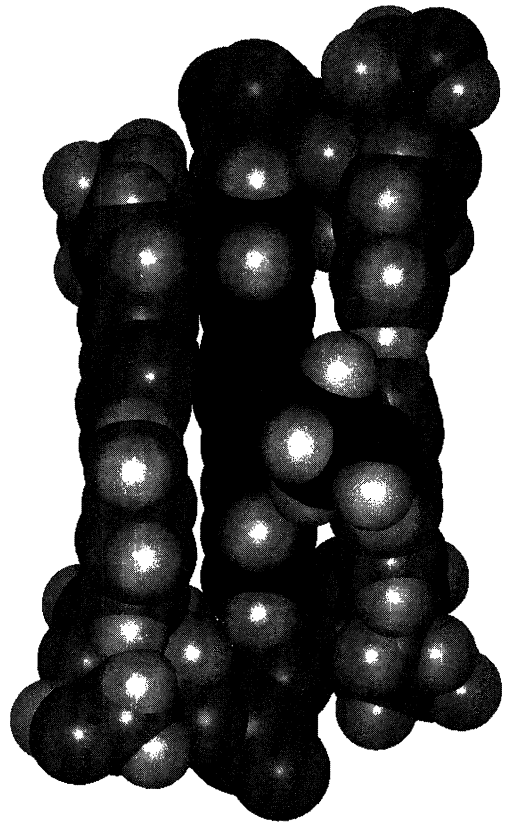


Figure 2. Space filling illustration of Pt(terpy)HET^+ (HET = 2-hydroxyethanethiolate) intercalated into the dinucleotide $(dCpG)_2$. View is perpendicular to the helix axis (adapted from Wang *et al.* 1978).

(*c*) *Coordination complexes in probing recognition*

Transition metal chemistry offers tremendous utility in defining interactions with DNA. This utility and chemical diversity are illustrated by the many natural and more recent unnatural metal complex-DNA interactions explored in our laboratory and elsewhere.

The best example of the application of metal–nucleic acid chemistry to chemotherapy to date rests in the simple coordination complex $(NH_3)_2PtCl_2$, *cis*-platin, which is a remarkably effective antitumor drug (Sundquist & Lippard 1990). *Cis*-platin is used worldwide as an anticancer agent and is considered a cure for testicular cancer. The mode of action of *cis*-platin represents an area of intensive research. Studies have established that, once *cis*-platin is activated by chloride loss and hydrolysis, the diammineplatinum(II) moiety may form a bidentate crosslink to adjacent guanines on the DNA strand. The crosslink involves direct coordination of the platinum centre to the N7 nitrogen atoms of guanines (Sherman *et al.* 1985). This crosslink leads to a severe kink in the DNA helix, which is recognized by structure-specific DNA-binding proteins (Chow *et al.* 1995). How these various interactions with DNA-binding proteins, promoted by platinum coordination, yields the preferential toxicity

to cancerous cells is still not understood. For the purpose of this discussion, what is important to keep in mind is that the *cis*-platin lesion involves direct coordination to nucleophilic sites on DNA rather than an ensemble of sequence-specific non-covalent contacts. Although effective as an anticancer agent, *cis*-platin shows cytotoxicity in non-cancerous cells as well. New generations of transition metal complexes must, therefore, be designed to achieve cell-specific cytotoxicity, presumably based upon DNA sequence-selective interaction.

Perhaps the closest analogue to our studies of non-covalent binding by metallointercalators is represented by the family of DNA-binding proteins containing the zinc finger domain (Berg 1993). It has been discovered that this zinc finger structural motif is ubiquitous to eukaryotic transcriptional activators. In this structural motif, a zinc ion coordinates two cysteines and two histidines within a ~ 30 residue polypeptide so as to fold the peptide into a structurally well defined and stable domain. Multiple domains then interact sequence-specifically in the major groove of DNA (or RNA). Thus the zinc centre defines a structural unit with the surrounding peptide ligands. This structural unit non-covalently associates with DNA, poising functional groups for an ensemble of specific contacts.

This zinc finger motif therefore illustrates well the utility of the metal centre in defining the geometry and stereochemistry for the ligand structure which encapsulates it. Using octahedral coordination complexes in synthetic structures targeted to DNA, one may similarly control the structure and stereochemistry of surrounding ligands. Furthermore, ligands may be varied systematically and with synthetic facility to generate a family of structures targeted to DNA. Finally, the potential for rigidity in these coordination complexes offers an additional element of definition and control in the design of DNA-binding complexes.

Transition metal complexes furthermore serve as powerful handles to probe DNA recognition. The spectroscopic properties of these complexes afford useful tools for the determination of affinities, modes and orientations of binding. For example, the octahedral immine complexes of ruthenium(II) have strong visible absorbance and luminescence properties which change upon binding to DNA (Murphy & Barton 1993). Paramagnetic metal complexes allow the use of NMR techniques to determine positions of binding via shifted and broadened DNA peaks (Rehmann & Barton 1990*b*). The presence of heavier metals aids in the elucidation of complex crystal structures of large nucleic acids by X-ray crystallography (Ohalloran *et al.* 1987; Jack *et al.* 1977).

Importantly, metal complexes also provide reactive probes of DNA. The first chemical DNA footprinting reagent was developed by appending FeEDTA onto methidium, a common sequence-neutral intercalator (Hertzberg & Dervan 1984). In the presence of a reducing agent and hydrogen peroxide, hydroxyl radicals could be generated near the DNA using Fenton chemistry associated with the DNA-bound ferrous centre. High concentrations of hydroxyl radicals promote DNA strand scission, which may be monitored using high resolution gel electrophoresis techniques. Tullius and co-workers (Price & Tullius 1992) then developed FeEDTA itself as a means of generating hydroxyl radicals in solution to probe DNA. Dervan and coworkers also applied this oxidative metal chemistry in developing the more general technique of affinity cleaving by appending FeEDTA onto specific, as well as nonspecific, DNA-binding molecules (Dervan 1986). The reactive metal, in promoting DNA strand scission through redox chemistry, could be applied to mark sites of binding by different appended moieties.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

In addition, oxidative DNA cleavage as a probe of DNA structure and reactions has been applied in studies using copper phenanthroline complexes (Sigman *et al.* 1993), nickel macrocycles (Burrows & Rokita 1994) and various metalloporphyrins (Magda *et al.* 1995; Mastruzzo *et al.* 1994). Perhaps the natural analogue for metal-promoted oxidative chemistry of DNA lies in metallobleomycin, a natural product which chelates a metal and is also used in chemotherapy (Hecht 1979). Bleomycin coordinates ferrous ion to form a complex which binds DNA in the minor groove with intercalation of the pendant bithiazole moiety (Wu *et al.* 1994; Manderville *et al.* 1994). A series of elegant mechanistic studies have shown that reaction with dioxygen leads to a ferryl intermediate which abstracts the C4′-H atom of the deoxyribose, promoting DNA strand scission (Absalon *et al.* 1995). This metal-promoted DNA cleavage is thought to be a critical element of the efficacy of the antitumor drug.

In our own studies of octahedral metallointercalators, we have primarily exploited photooxidation by rhodium(III) complexes to produce photoinduced DNA cleavage. In these reactions, photolysis of DNA bound complexes of rhodium(III) most likely activates the intercalated ligand, which is positioned in the major groove of duplex DNA. This activation takes the form of a ligand to metal charge transfer which leads to the direct abstraction of the C3′ hydrogen of deoxyribose and results in strand scission. Product analyses have been consistent with this mechanism (Sitlani *et al.* 1992). This photoinduced DNA cleavage chemistry may therefore similarly be applied in high resolution gel electrophoresis experiments to mark the sites of binding by the metallointercalator.

2. Recognition based on shape selection

Given metal complexes with well defined rigid structures, we examined first recognition based upon shape selection (Pyle & Barton 1990). Complexes have been prepared with shapes and symmetries which match those of sites on DNA (Chow & Barton 1992). Indeed, these same complexes may be applied in probing the local variations in structure along the DNA helix. In general, what we learned from these initial studies, time and again, was the power of shape selection in achieving high site-selectivity. Complexes lacking hydrogen bond donors or acceptors may be targeted to sites with specificities rivaling DNA-binding proteins.

(*a*) *Coordination complex chirality and the right-handed double helix*

One of the simplest examples of shape selection lies in the chiral interactions of tris(phenanthroline) metal complexes with the right-handed B-DNA helix (Barton 1986). Photophysical (Barton *et al.* 1986) and ^{1}H-NMR (Rehmann & Barton 1990*a*, *b*) studies indicate that each enantiomer binds to DNA via two modes: one in which the complex partially intercalates between the bases in the major groove, and one in which the complex binds against the surface of the minor groove. These studies also indicated a preference for the Δ enantiomer upon intercalation into the right-handed helix and a preference for the complementary symmetry of the Λ-isomer in surface-binding to right-handed DNA.

These results can be best explained if one considers the geometry of the ancillary ligands when the complex binds via intercalation. As can be seen from figure 3*a*, the Δ-isomer of the complex fits nicely into the major groove of right-handed DNA with the ancillary ligands extending along the groove, thus minimizing steric clashes with the right-handed phosphate backbone. On the other hand, intercalation of the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Λ-isomer positions the ancillary ligands against the groove and steric clashes with the phosphate backbone occur so as to destabilize the binding by this mode. A similar effect is seen with isomers of $Rh(phen)_2phi^{3+}$ (David & Barton 1993) and $Ru(phen)_2dppz^{2+}$ (Dupureur & Barton 1994), shown in figure 3*b*, where the presence of the phi and dppz ligands with enhanced surface areas for stacking leads to high affinity intercalation for both enantiomeric forms. In these cases, the enantioselective preference for the Δ-isomer is maintained (Pyle *et al.* 1990; Hiort *et al.* 1993). More support for this model is offered by the observation that the size of the ancillary ligands determines the extent of the enantioselective effect. As shown in figure 3*c*, for the family of metal complexes $Rh(bpy)_2phi^{3+}$, $Rh(DMB)_2phi^{3+}$ and $Rh(DPB)_2phi^{3+}$, the enantioselectivity increases from left to right with increasing ancillary ligand size (Sitlani *et al.* 1993; Sitlani & Barton 1994). Indeed, Λ-$Rh(DPB)_2phi^{3+}$ shows no detectable binding to B-DNA at 1000 times the concentration of that where Δ-$Rh(DPB)_2phi^{3+}$ binds specifically.

(*b*) *Size and shape of the grooves*

The major and minor grooves of a nucleic acid duplex offer different surfaces and shapes for recognition. Groove shapes become more varied still as a function of nucleic acid sequence and local conformation. A simple example of the heterogeneity of groove shape can be seen in the comparison of double stranded DNA, which is primarily in the B-conformation, with double stranded RNA, which adopts the A-conformation. Metal complex probes can be constructed which are sensitive to this variability in groove width and depth.

One such complex is $Ru(TMP)_3^{2+}$ (Mei & Barton 1988). The complex binds nucleic acids via van der Waals contacts and electrostatics in the minor groove. For B-form nucleic acids, the minor groove is deep and narrow with the base stack protected from solution, as shown in figure 4. The bulky TMP ligands on the complex are too large to get past the phosphate backbone and make van der Waals contacts with the base pairs. The minor groove of A-form nucleic acids adopts a different shape. In this case, the groove is shallow and wide, presenting the base stack to solution where the complex can easily interact with it.

A good probe of the shape of the major groove of nucleic acids is the intercalating metal complex $Rh(phen)_2phi^{3+}$. While the complex binds avidly to B-form DNA, it shows low affinity for A-form nucleic acids like duplex RNA. This observation can be explained by the general groove shape of A- and B-form nucleic acids. As illustrated in figure 4, A-form nucleic acids possess a very narrow and deep major groove that does not allow enough space for the complex to reach and intercalate into the base stack. B-form nucleic acids, however, have a much shallower and wider major groove so that the complex can reach the base stack to gain the stabilization afforded by intercalative stacking (Campisi *et al.* 1994).

The fact that $Rh(phen)_2phi^{3+}$ does not bind to A-form nucleic acids makes it an excellent probe of RNA tertiary structure (Chow *et al.* 1992). On yeast $tRNA^{phe}$, for which extensive crystallographic data exist, photocleavage studies have shown the complex to bind sites of triple-base interactions and structured loops, while the complex shows no detectable binding to the A-form double-stranded regions. Having characterized the binding properties of the complex to tRNA, binding studies were then carried out on tDNA, which was proposed to have a tertiary structure similar to RNA, although crystallographic data was lacking. As shown in figure 5, the complex targeted the same sites on tDNA as on tRNA (Lim & Barton 1993). In

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

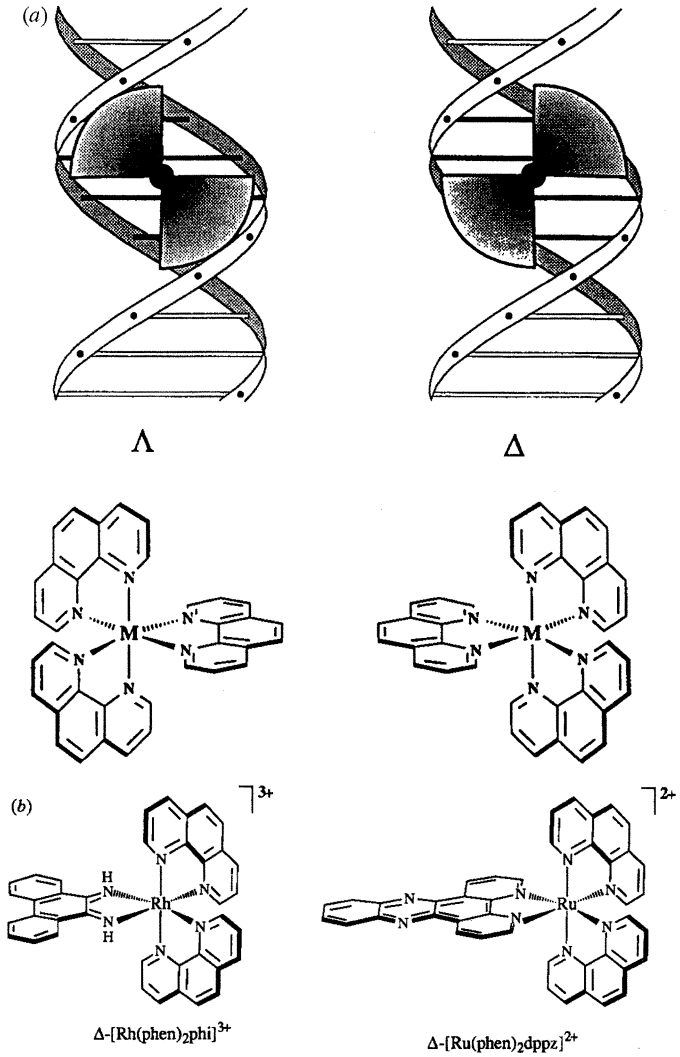


Figure 3. (*a*) An illustration of both enantiomers of the tris(phenanthroline) metal complexes intercalated into DNA. Note the steric clashes of the phenanthroline ligands of the *Λ*-enantiomer with the backbone of the right-handed duplex. (*b*) The metal complexes $Rh(phen)_2phi^{3+}$ and $Ru(phen)_2dppz^{2+}$.

addition, however, targeting of 5′-pyrimidine-pyrimidine-purine-3′ (5′-pyr-pyr-pur-3′) sites were observed in putative double helical regions. This specificity indicated that the tDNA was forming globally the same structure as the tRNA, but that its double-stranded regions adopted the B-form instead of A-form.

(*c*) *Local variations in DNA propeller twisting*

A more subtle example of shape selection is illustrated by the sequence-selective binding of $Rh(phen)_2phi^{3+}$ to B-form DNA itself. The rhodium complex contains no hydrogen bonding functionalities, yet DNA photocleavage studies have demonstrated a selective binding to 5′-pyr-pyr-pur-3′ sequences with intercalation between the pyr-pur base step (Pyle *et al.* 1989). This result can be explained by steric clashes

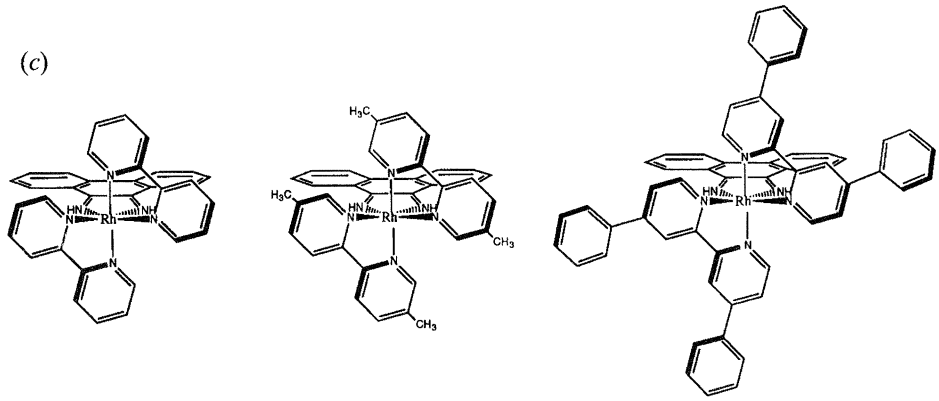


Figure 3. (*c*) Effect of ancillary ligands on enantioselectivity. Enantioselectivity increases with increasing ligand size from left to right.

between the ancillary ligands and the DNA. Modelling studies have shown that the 2,9 hydrogens of the phenanthroline ligands may clash with the DNA bases upon intercalation of the phi ligand. Propeller twisting of the DNA bases can open up the intercalation site to allow more room for these hydrogens in a conformation referred to as an open major groove. As can be seen from figure 6, the open major groove creates a chiral site for the complex in addition to that present due to the handedness of the helix. This affords even greater enantioselectivity in binding. This notion is supported by the quantitative correlation of enantioselective cleavage by $Rh(phen)_2phi^{3+}$ with propeller twisting at such sites measured through X-ray crystallographic studies of different B-form oligonucleotides (Campisi *et al.* 1994). Thus, shape-selective targeting by $Rh(phen)_2phi^{3+}$ may be applied as a chemical probe of the 'openness' of the major groove.

(*d*) *Site-specificity rivaling a DNA binding protein*

An example of the power of shape selection in determining sequence specificity is seen in the metal complex $Rh(DPB)_2phi^{3+}$, shown in figure 7. The diphenylbipyridyl ancillary ligands of this complex contain no potential hydrogen bonding functionalities, yet the specificity of the complex spans eight base pairs (Sitlani *et al.* 1993). The complex must be recognizing its target site based on the shape of this recognition sequence, 5′-CT**CT**AGAG-3′. Molecular modelling of the complex with DNA indicates that it is only large enough to span six base pairs, and photocleavage data indicate that the complex intercalates between the boldfaced C and T. The cleavage site places the complex in a position such that it cannot reach the two 3′ bases of its recognition sequence. This result led us to propose a model where two complexes dimerize across the recognition sequence. Photocleavage and DNA footprinting studies were carried out on DNAs containing the eight base pair full recognition site, as well as a six base pair monomeric subsite 5′-CTCTAG-3′. The binding affinity of the complex to the eight base pair site was found to exceed the binding affinity to the six base pair site. These results were consistent with two complexes binding to the eight base pair site in a cooperative manner with a 2 kcal interaction energy. This non-covalent dimerization in fact reflects a strategy exploited by DNA binding proteins (Steitz 1990) to substantially enhance site-specificity, a strategy which may also be utilized in subsequent generations of metal complexes being designed.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

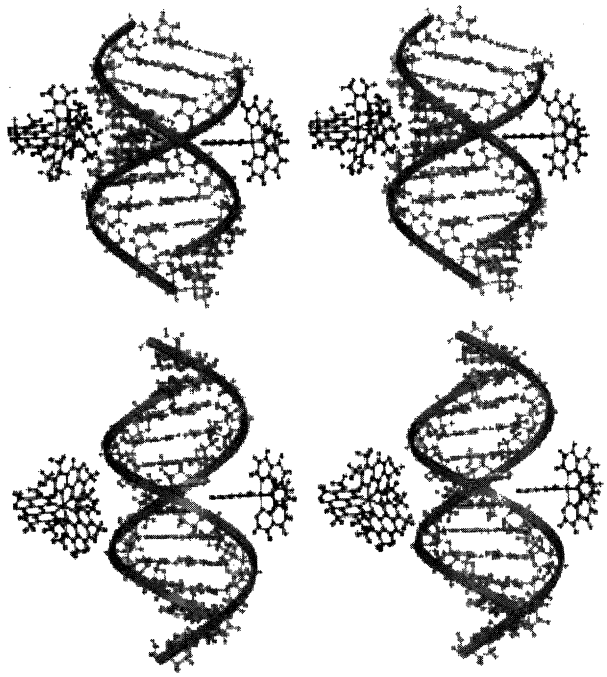


Figure 4. Stereoviews illustrating the size and shape complementarity of Λ-Ru(TMP)$_3^{2+}$ (left) and Δ-Rh(phen)$_2$phi^{3+} (right) with the grooves of A-form (top) and B-form (bottom) nucleic acids. Λ-Ru(TMP)$_3^{3+}$ is illustrated bound against the minor groove and Δ-Rh(phen)$_2$phi^{3+} is poised for intercalation in the major groove. In A-form nucleic acids, the shallow minor groove presents an accessible surface for binding by Λ-Ru(TMP)$_3^{2+}$. Meanwhile the backbone in the major groove has closed down, preventing the intercalation of Δ-Rh(phen)$_2$phi^{3+}. In contrast, the tetramethylphenanthroline ligands of Λ-Ru(TMP)$_3^{2+}$ are too large to contact the base stack in the minor groove of B-form nucleic acids without clashing with the phosphate backbone. In the major groove, Δ-Rh(phen)$_2$phi^{3+} has ample room to intercalate into the base stack.

3. Direct readout

Shape selection can provide a powerful influence on the sequence specificity of a molecule binding to DNA. In fact, Rh(DPB)$_2$phi^{3+} with its eight base pair recognition sequence is the most specific complex we have created to date. However, binding by shape selection is a tremendously complicated process involving many weak interactions and steric factors. While there is an increasing body of structural data becoming available for oligonucleotides, the rules governing the subtle heterogeneities of the double helix shape based on sequence is still poorly understood. This rudimentary understanding severely limits the use of shape selection in the *predictive* design of sequence specific DNA binding molecules.

A potentially more predictive approach to the design of sequence-specific molecules is the use of direct hydrogen bonds and van der Waals contacts to functional groups on the DNA base pairs. As mentioned earlier, coordinatively saturated octahedral metal complexes make an excellent scaffolding for the placement of functional groups

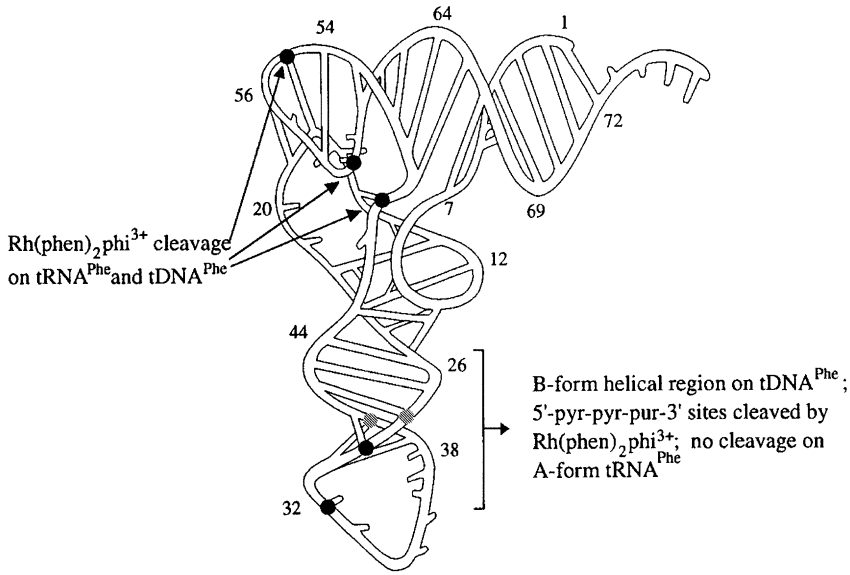


Figure 5. Illustration of $Rh(phen)_2phi^{3+}$ photocleavage on $tRNA^{Phe}$ and $tDNA^{Phe}$. Black circles denote positions of cleavage on both $tRNA^{Phe}$ and $tDNA^{Phe}$, while grey circles denote positions of cleavage on $tDNA^{Phe}$ only. Note that $Rh(phen)_2phi^{3+}$ cleaves double helical regions only on $tDNA^{Phe}$, indicating these regions are B-form only in $tDNA^{Phe}$. On both folded polymers, sites of tertiary interaction are targeted.

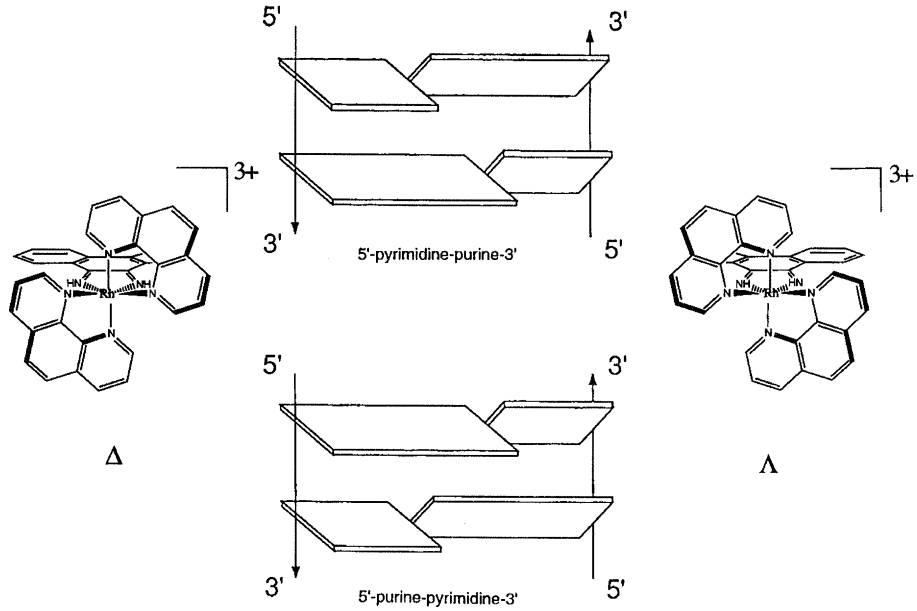


Figure 6. Schematic illustration of the chiral binding site created by propeller twisting in an open major groove. The enantiomers of $Rh(phen)_2phi^{3+}$, also shown, provide probes of this propeller twisting.

for such contacts. In order to better understand the energetics and constraints involved in the design of molecules which would recognize DNA via these contacts, we

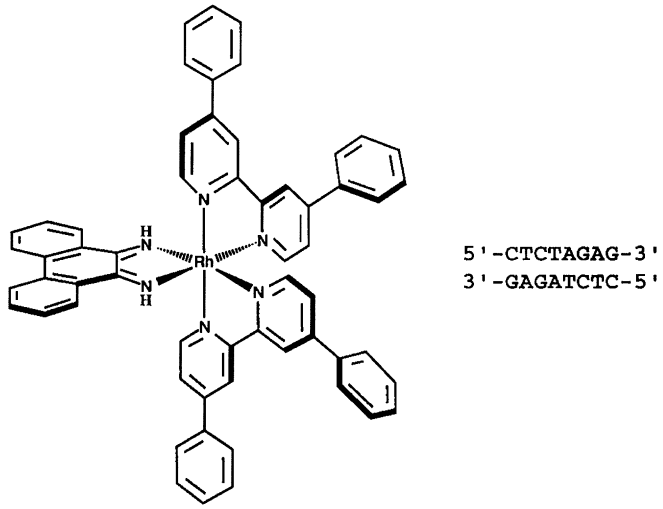


Figure 7. Δ-Rh(DPB)$_2$phi^{3+} and its palindromic recognition site.

can apply coordination chemistry, starting with quite simple complexes, to characterize these interactions.

(*a*) *Hydrogen bonds*

An example of a simple synthetic metal complex which recognizes a sequence based on hydrogen bonding contacts is Rh(NH$_3$)$_4$phi^{3+}. ^{1}H-NMR (Collins *et al.* 1994) and DNA photocleavage (Krotz *et al.* 1993*b*) studies have shown a preference for the complex to intercalate between 5′-GC-3′ steps. ^{1}H-NMR data on Rh(NH$_3$)$_4$phi^{3+} bound to d(TGGCCA)$_2$ indicate that the complex is deeply intercalated between the central 5′-GC-3′ base step, and that the axial ammines of the complex are therefore placed in position to form specific hydrogen bonds with the O6 of the guanines above and below (figure 8). A series of related compounds, constructed to test if the amine hydrogen bonding contacts may be generally responsible for the 5′-GC-3′ specificity, is shown in figure 9. Indeed, all of the Rh(phi)$^{3+}$ complexes containing saturated axial amines also show the preference for 5′-GC-3′ recognition. In contrast, Rh([12]aneS$_4$)phi^{3+}, which lacks amines and thus the ability to donate hydrogen bonds to the DNA bases, lacks this 5′-GC-3′ specificity.

(*b*) *van der Waals interactions*

DNA binding studies with enantiomers of Rh(en)$_2$phi^{3+}, both of which contain axial amines for hydrogen bonding, but which differ in disposition of the en ligands, furthermore showed differing specificities (Shields & Barton 1995). While both enantiomers did indeed display a strong affinity for 5′-GC-3′ base steps, Λ-Rh(en)$_2$phi^{3+} showed also significant binding to 5′-TX-3′ base steps; this site recognition is absent for the Δ-enantiomer. This result cannot be explained based on differences in hydrogen bonding, as both complexes contain axial amines which are similarly disposed. Shape selection due to steric clashes of the ancillary ethylene diamine ligands with the phosphate backbone is also not responsible, as these ligands are too small to account sterically for this interaction. Molecular modelling, however, indicates that the methylene groups of the ancillary en ligands are positioned in the Λ-enantiomer and not in the Δ-enantiomer for specific van der Waals contacts with the methyl

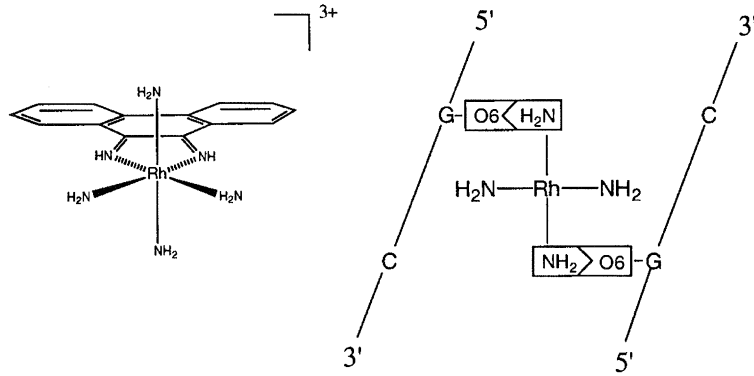


Figure 8. $Rh(NH_3)phi^{3+}$ and a schematic of its hydrogen bonding contacts to a DNA 5′-GC-3′ step upon intercalation in the major groove.

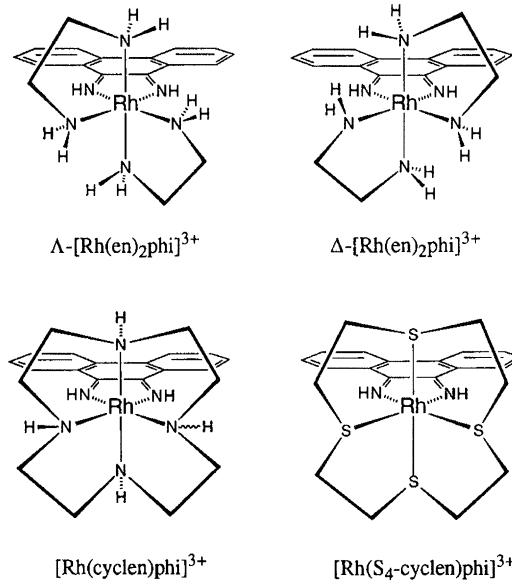


Figure 9. Family of rhodium complexes constructed to test the importance of hydrogen bonding for 5′-GC-3′ specificity. All complexes are found to target 5′-GC-3′ steps except $Rh(S_4\text{-cyclen})phi^{3+}$, which lacks hydrogen bonding functionalities.

group of 5′-thymines in the DNA major groove, as illustrated in figure 10. To test directly the importance of this contact, uracil, which is identical to thymine except it lacks the methyl group, was substituted for thymine in synthetic oligonucleotides used to test recognition. DNA photocleavage studies were conducted on the synthetic oligonucleotides containing both uracil and thymine, and the results were compared. Oligonucleotides containing uracil showed no significant enantioselectivity in cleavage by $Rh(en)_2phi^{3+}$ enantiomers as compared to oligonucleotides containing thymine. This result indicates the essentiality of the methyl group of thymine for the binding by the Λ-enantiomer. This non-covalent contact was found to provide ~ 1 kcal stability. Importantly, the result furthermore supported directly the notion that these metallointercalators access the DNA from the major groove.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

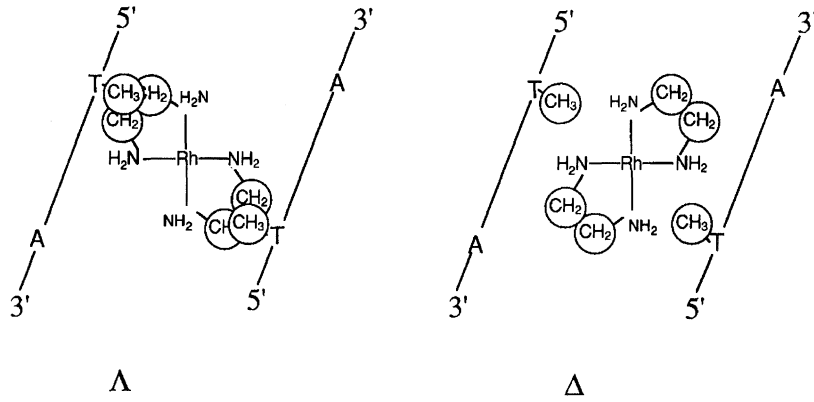


Figure 10. Schematic of methyl-methylene van der Waals interactions between the enantiomers of $Rh(en)_2phi^{3+}$ and 5′-TA-3′.

(*c*) *Assembling non-covalent contacts for predictable specificity*

With the above hydrogen bonding and van der Waals contact data in hand, we next designed the complex Δ-α-Rh[(R,R)-Me_2trien]phi^{3+} (Krotz *et al.* 1993*a*). The complex is a derivative of Δ-$Rh(en)_2phi^{3+}$ which combines hydrogen bonding functionalities and van der Waals contacts to afford predictable four base pair specificity. This complex was designed specifically to target the sequence 5′-TGCA-3′. Modelling studies indicated that the axial amines of the complex are well positioned for hydrogen bonding to the O6 of guanines, while the pendant methyl groups are disposed appropriately for van der Waals contacts with the methyl groups of the thymines two bases to the 5′-side, as shown in figure 11*a*.

As a demonstration of the flexibility of coordination chemistry for new design, we examined differences in recognition by the various structural isomers of $Rh(Me_2trien)phi^{3+}$ (see figure 11*b*). All 12 isomers were isolated, structurally characterized and stereochemistry determined (Krotz & Barton 1994). Each isomer contains the axial amines in position to hydrogen bond with the O6 of guanine. However, these isomers offer eight unique placements of the two methyl groups in three-dimensional space. Modelling indicated that only Δ-α-Rh[(R,R)-Me_2trien]phi^{3+} contains the correct geometry for stereospecific van der Waals contacts between both its methyl groups and the methyl groups of the thymines of the sequence 5′-TGCA-3′.

Importantly, photocleavage studies show that only this Δ-α isomer binds to 5′-TGCA-3′ (Krotz *et al.* 1993*a*). Titration studies of the complex with DNA indicated the preference of binding to sites of 5′-TGCA-3′ > 5′-GCA-3′ > 5′-GC-3′. The affinity of the complex for each of these sites decreases by approximately 1 kcal from left to right consistent with the loss of a methyl–methyl van der Waals contact.

Most recently, a high resolution ^{1}H-NMR structure of the complex bound to the DNA decamer $d(GAGTGCACTC)_2$, illustrated in figure 11*c*, has demonstrated that the complex intercalates between the central 5′-GC-3′ step of the decamer, and that the designed van der Waals interaction is present between the complex methyl groups and the methyl groups of thymine (Hudson *et al.* 1995). This NMR structure represents that of the smallest site-specific molecule bound to the largest DNA yet characterized using high resolution NMR spectroscopy. The site-specificity demonstrates the utility of a minimalist approach. This large duplex oligonucleotide containing a single intercalator bound at one specific site furthermore provides an excellent system for examining the long range effects of intercalation. The ^{1}H-NMR data indicate

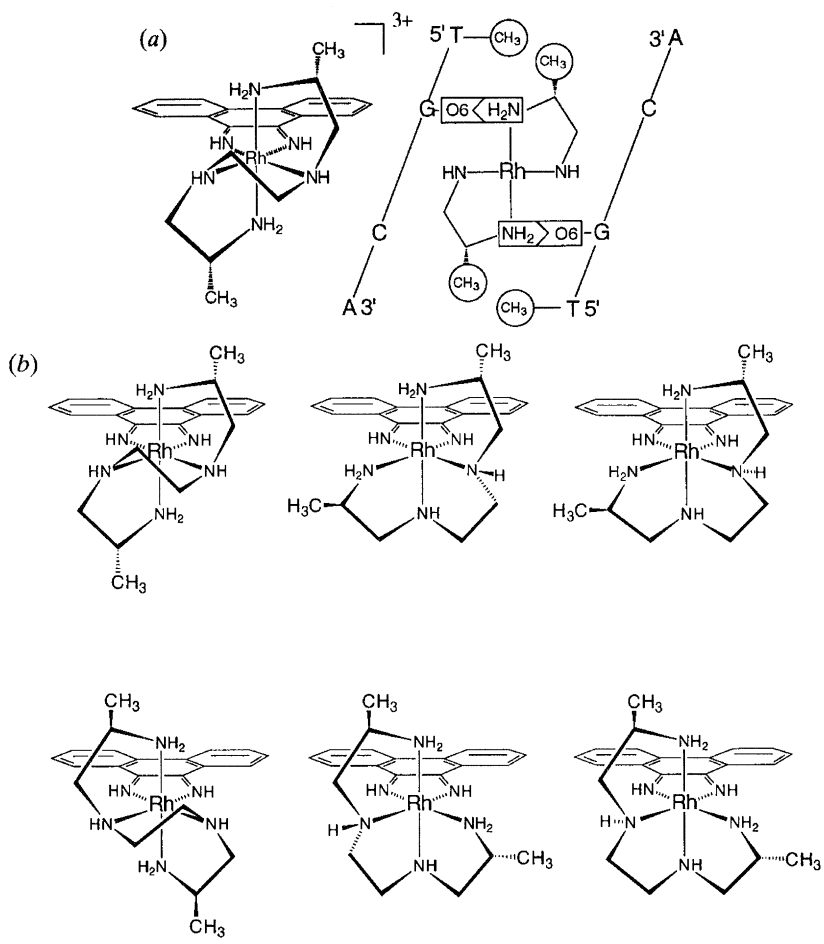


Figure 11. (*a*) Schematic of Δ-α-Rh[(R,R)-Me_2trien]phi^{3+} bound to its recognition sequence. (*b*) Structural isomers of Rh[(R,R)-Me_2trien]phi^{3+} (the enantiomers are not shown).

that little distortion occurs to the DNA outside the site of intercalation; no bending of the duplex is observed. Instead, unwinding is evident at the central 5′-GC-step to accommodate the intercalator with preferential stacking of the phi ligand between guanine bases on opposing strands. This intercalation poises the ancillary Me_2trien ligand in the major groove for direct hydrogen bonding contacts with the guanines and direct methyl–methyl contacts with thymines two bases above and below. This structure represents the first example of a rationally designed synthetic metal complex interacting site-specifically with duplex DNA. Predictable site-specificity may be achieved by building up an ensemble of contacts in the major groove using metallointercalation as a platform.

4. Combining direct readout and shape selection: recognition based on sequence-dependent DNA twistability

Λ-1-Rh$(MGP)_2$phi^{5+}, shown in figure 12, represents an example of a metallointercalator which combines both shape selection and direct readout to effect DNA

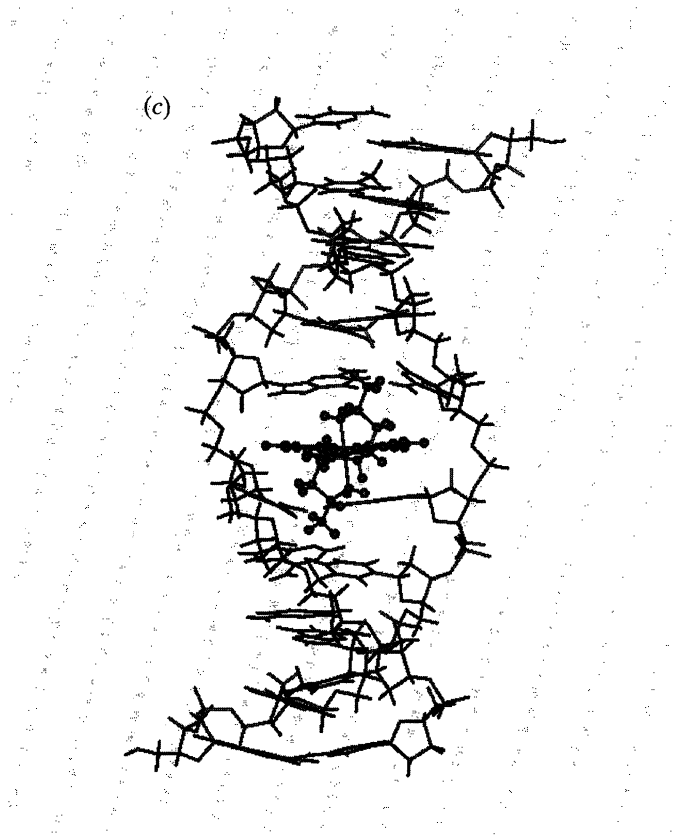


Figure 11. (*c*) NMR structural model of Δ-α-Rh[(R,R)-Me_2trien]phi^{3+} bound to the decamer $d(GAGTGCACTC)_2$.

binding with high affinity and specificity. This complex targets the six base pair site 5′-CATATG-3′ with specificity and at subnanomolar concentrations (Terbrueggen & Barton 1995).

Isomers of $Rh(MGP)_2phi^{5+}$ were first prepared as a second generation analogue of $Rh(phen)_2phi^{3+}$ containing pendant guanidinium groups. While the absence of a general amino acid–DNA recognition code is clear, one frequent peptide recognition element observed in crystal structures, notably of zinc finger proteins, has involved targeting of guanine by the guanidinium side chain of arginine (Pabo & Sauer 1992). Also as illustrated in figure 12, a guanidinium group can donate hydrogen bonds to both the guanine O6 and N7 atoms. The intention was to prepare a derivative of $Rh(phen)_2phi^{3+}$ in which the central element of the recognition site would be a 5′-pyr-pyr-pur-3′ site, governed by the shape-selective interaction of Δ-$Rh(phen)_2phi^{3+}$, and the periphery of the recognition site would be G–C base pairs, targeted by direct readout by the pendant guanidinium groups. The isomer with guanidinium group disposed away from the phi ligand, also shown in figure 12, would offer a useful control for the electrostatic gain presented by introducing two positively charged functionalities onto the metal complex.

The recognition characteristics of the resultant complexes observed were quite remarkable. Although Δ-1-$Rh(MGP)_2phi^{3+}$ showed site-specificities which reflects our design criterion, extremely high specificity and affinity were observed using the Λ-

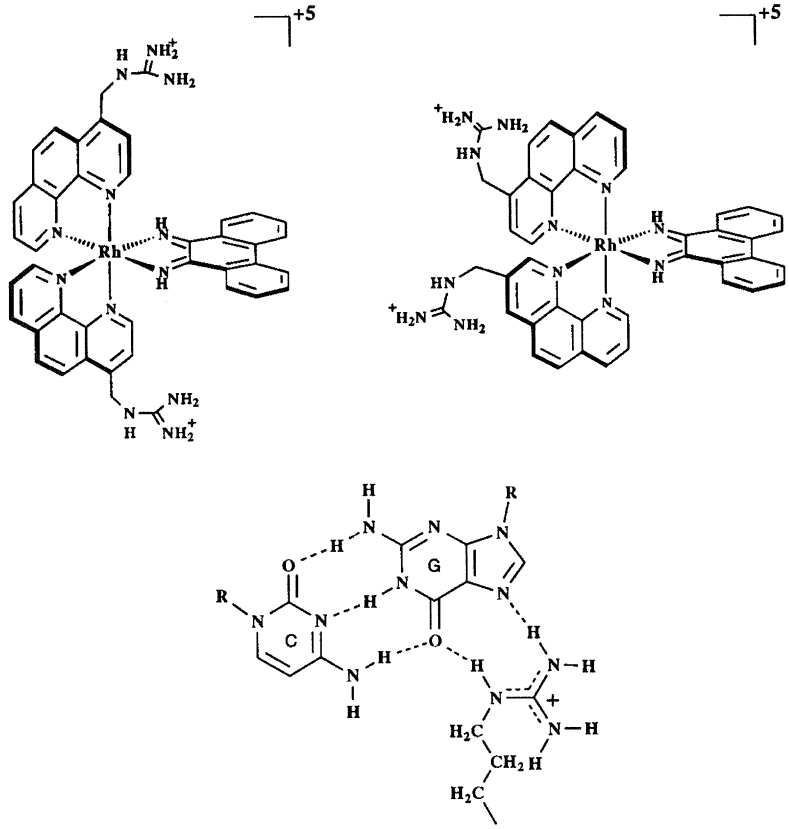


Figure 12. (top) Two of the C2-symmetric isomers of $Rh(MGP)_2phi^{3+}$. (bottom) Hydrogen bonding contacts between the guanidinium group of arginine and a guanine cytosine base pair.

isomer for a site containing a central 5′-ATAT-3′, which is not preferentially targeted by Λ-$Rh(phen)_2phi^{3+}$. Photocleavage studies on a series of sequences indicated the essentiality of the central TA base step for recognition at any concentration, and oligonucleotide substitution studies using 7-deazaguanine confirmed the direct contact of the guanidinium groups of Λ-1-$Rh(MGP)_2phi^{5+}$ with the N7 nitrogen atoms of guanines in the 1 and 6 positions of the six base pair site.

The highly specific recognition was found to depend upon a sequence-specific twistability of DNA. This sequence-dependent twistability was determined in a DNA unwinding assay we developed, schematically illustrated in figure 13, and based upon earlier elegant assays for DNA bending designed by Crothers and co-workers (Wu & Crothers 1984). Earlier studies had shown that six adenines in a row, an A tract, cause the DNA helix to bend roughly 20°. If a series of these A tracts are placed in proper phase with the helical repeat of DNA, the bend becomes additive, and large curves in the DNA polymer are produced. Two sets of five additive A tracts, producing a 100° bend, were therefore ligated together with a variably sized central oligonucleotide containing the specific recognition sequence for Λ-1-$Rh(MGP)_2phi^{5+}$ (Terbrueggen & Barton 1995). When the length of the linker oligonucleotide was a multiple of 10.5 base pairs, which corresponds to a 360° twist of the helix, the A tracks are in phase with each other and produce a 200° overall curvature in the DNA. This bent oligomer is well positioned for intramolecular cyclization by the enzyme

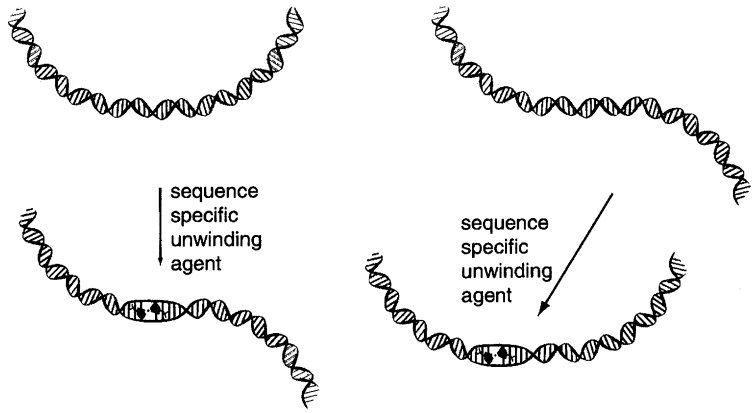


Figure 13. Assay used to determine the extent of DNA unwinding upon binding by 1-Λ-$Rh(MGP)_2phi^{3+}$. As illustrated, the metal complex, in sequence-specifically unwinding the DNA, macroscopically alters the structure of a bent DNA. A bent structure is converted to an S-shaped or an S-shaped structure to a bent one. These structural alterations can be detected in solution through measurements of the rate of intramolecular cyclization of these DNA substrates.

ligase. If the metal complex specifically binds to the central region so as to unwind the DNA, altering the relative phasing of the two bent segments to an 'S'-shaped structure, then intramolecular cyclization by ligase is inhibited. Perhaps the better demonstration of sequence-specific unwinding, however, is apparent in comparing intramolecular cyclization for the oligonucleotide originally phased in the 'S'-shape in the absence and presence of metal. Here, addition of a metal complex, also illustrated in figure 13, would serve to unwind the central segment so as to position the two bent segments for facile cyclization; in this case, ligation by the enzyme is accelerated by the metal complex.

Using this assay, we determined that in the presence of Λ-1-$Rh(MGP)_2phi^{3+}$, the site is unwound by $70 \pm 10°$, a substantial unwinding of the duplex. We have not yet determined if this sequence is particularly twistable in the absence of metal, and binding by the rhodium complex, with its pendant guanidinium groups properly disposed, traps the unwound site, or whether the metal complex promotes the sequence-specific duplex unwinding. It is important to note in this context, however, that an essential transcriptional activator, the TATA-box binding protein (TBP), recognizes the sequence 5′-TATATAAA-3′. This site is similar to that recognized by the rhodium complex and is substantially bent and also unwound by more than 80° in the presence of TBP. Studies with the rhodium complex therefore offer remarkable analogies to critical protein-DNA interactions. Moreover, these studies indicate quite clearly how the coupling of direct readout strategies to some conformational switch, or sequence-dependent twisting, can offer a remarkably high level of specificity in DNA site-recognition.

5. Metal–peptide chimeras

In nature, the recognition domains of proteins can often be broken down into discrete units. Often, the functionalities of a protein which make specific contacts with the DNA are localized to an α-helix which lies in the DNA major groove, while substantial non-specific DNA binding affinity is gained from a variety of electrostatic interactions with the phosphate backbone (Steitz 1990). Frequently, site-specific affini-

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

ties of DNA binding proteins are found to be 10^9–10^{10} M^{-1}, while affinities for non-specific DNA range from 10^4–10^7 M^{-1}. The site-specific recognition helices, however, cannot be separated from their parent proteins and used as sequence-specific DNA binding molecules because, without the remainder of non-specific contacts, they lack sufficient affinity for detectable binding. Furthermore, in the context of their natural proteins, these recognition units are specifically directed to the major groove in the correct orientation and conformation for binding.

Once again, transition metal chemistry is a valuable tool for the construction of sequence-specific DNA-binding molecules where pendent peptides are used to achieve selectivity. To overcome the aforementioned problems, we have attached a series of recognition helices to the sequence-neutral metallointercalator $Rh(phi)_2phen^{3+}$ (Sardesai *et al.* 1994). In this chimera (figure 14*a*), the metal complex provides the non-specific binding energy necessary to bring the peptide into contact with the DNA in the major groove while the peptide provides the sequence specificity.

A family of metal peptide chimeras was prepared to investigate the binding properties of a peptide recognition helix. Attachment of this small recognition peptide to the metal complex resulted in a chimera which was remarkably specific for the sequence 5′-CCA-3′. No cleavage at these sites was found by the metal complex lacking the appended peptide. Hence, site-specific recognition depended upon interactions by the peptide. Systematic substitutions of amino acids in the appended peptide were made to determine which amino acids were responsible for the conformation and the sequence-specificity of the 'recognition helix'. Specific recognition of the chimera was found to depend sensitively upon the glutamic acid at position 10. Any mutation at this position resulted in a loss of specificity for DNA as well as a loss in α-helicity. Even conservative mutations such as replacing the glutamate with an aspartate, glutamate methyl ester, or glutamine, disrupted binding. Importantly, alanine substitution at this position preserved α-helicity but site-selectivity was nonetheless lost. Hence, the glutamate appeared essential, both to form an α-helix and for a specific contact in the DNA major groove. We have proposed the model shown in figure 14*b* for this site-specific interaction, in which the metal complex is intercalated into the DNA with the α-helical peptide directed along the major groove to achieve a direct contact between the essential glutamate and the amine group of cytosine. We propose 5′-CA-3′ recognition to depend on the shape-selective opening of the major groove at the CA step. Importantly, as we saw in recognition by $Rh(MGP)_2phi^{3+}$, here too, a high level of specificity is derived from a sequence (peptide)-dependent conformational switch. For the peptide positioned in the α-helix to achieve contacts with the DNA major groove, a high level of specificity results.

6. Tests of sequence-specific biological function

The ultimate goal in designing metal complexes which target specific DNA sequences is their application as therapeutics in the treatment of molecular diseases. In this regard, the most sequence-specific molecule is useless if it fails to show an effect on biological function. Early studies on $Co(DIP)_3^{3+}$ have shown that despite the high positive charge on this complex, it is avidly concentrated in the nuclei of mammalian cells (Chapnick *et al.* 1988). To date, molecules of high specificity which we have constructed include Δ-$Rh(DPB)_2phi^{3+}$ and 1-Λ-$Rh(MGP)_2phi^{5+}$, which target eight and six base pair sites, respectively. These two compounds differ substantially in solubilities, with the first being readily soluble in chloroform and the second being

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

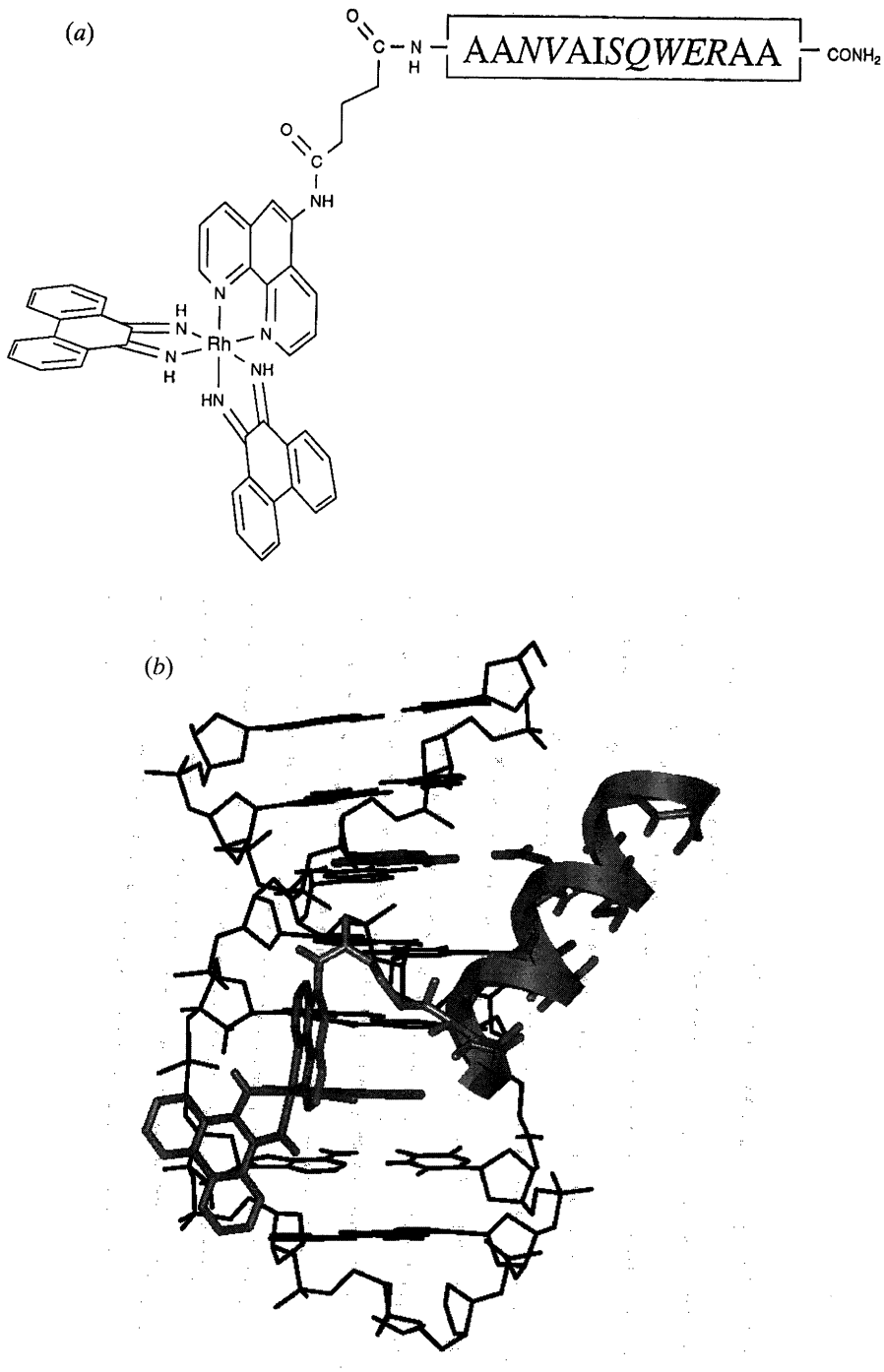


Figure 14. (*a*) A metal–peptide intercalating chimera and (*b*) a model of the complex bound to its recognition sequence. Note the α-helical conformation of the appended peptide with the glutamate side chain disposed for a specific contact with cytosine.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

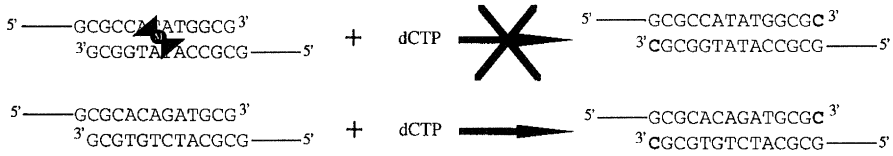


Figure 15. Illustration of template-dependent inhibition of a DNA polymerase by a metal complex.

very soluble in water. It is encouraging that the DNA selectivity of the metal complex is not predominantly influenced by its solubility; this characteristic will certainly require fine-tuning for issues of bioavailability. Both complexes are being tested with respect to biological function, and with encouraging success.

The high specificity and affinity of Δ-Rh(DPB)$_2$phi^{3+} for the eight base pair palindromic site 5′-CTCTAGAG-3′ allowed the investigation of whether or not the complex could successfully compete with a DNA binding protein that would target the same sequence (Sitlani *et al.* 1993). The metal complex shares a binding site with the restriction enzyme Xba I. Actually, the restriction enzyme shows lower specificity than the metal complex, in that Xba I targets only the internal six base pairs. Competition assays have shown that the complex can inhibit Xba I's ability to cleave DNA at its recognition sequence. Two important control experiments were also carried out. First, Xba I cleavage was examined in the presence of Rh(phi)$_2$bpy^{3+}, a sequence-neutral rhodium intercalator. In the presence of this metal complex, no similar inhibition of Xba I cleavage was observed. Next, we examined whether Rh(DPB)$_2$phi^{3+} promoted inhibition of restriction enzymes which targeted alternate sites; again no effect was observed. Hence, Δ-Rh(DPB)$_2$phi^{3+} sequence-specifically inhibits Xba I. The metal complex is, therefore, found to rival a natural DNA-binding protein, not only in affinity and specificity, but also as a direct functional competitor.

Perhaps in a more functionally relevant set of experiments, the metal complex 1-Λ-Rh(MGP)$_2$phi^{5+} has been shown to inhibit the replication of DNA by DNA polymerase in a sequence-selective manner (Johann & Barton 1995, unpublished work). Two DNA polymerase templates were constructed, one containing a binding site for 1-Λ-Rh(MGP)$_2$phi^{5+} and one lacking a specific duplex recognition site. DNA synthesis would then be sequence-specifically inhibited if the metal complex preferentially bound to its target, blocking the polymerase from binding to initiate DNA extension. When both templates were added to the same test tube and polymerization reactions carried out in the presence of the metal complex, extension of the metal specific template was found to be significantly inhibited compared to extension of the template lacking the target site (figure 15). This inhibition was furthermore enantiospecific; no similar template-specific inhibition by Δ-1-Rh(MGP)$_2$phi^{3+} was observed. Additionally, this specific inhibition was not found to depend on the non-specific template chosen for competition. Furthermore, the same effect was observed using three different DNA polymerases. These results clearly demonstrate that the metal complex may be applied to inhibit DNA synthesis in a sequence-specific manner.

7. Future prospects

Our studies to design coordination complexes which bind DNA have provided us with some of the principles of site-specific recognition. We have learned how to design

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

specificity through a combination of shape selection and direct readout. Like DNA-binding proteins, complexes may be prepared using the metal centre as a scaffold from which to append a range of functionalities for sequence-specific contacts in the major groove. Metallointercalation provides a versatile platform for these interactions. We will build on this knowledge by constructing molecules which combine these non-covalent interactions in various ways to provide new sequence specificities and greater affinity. We have also learned how to append peptides to our metallointercalators to achieve a whole new array of sequence-specific DNA-binding complexes. Additionally, from our studies, notions concerning DNA sequence-dependent twistability and conformational switches have emerged. By coupling an ensemble of sequence-specific contacts to a conformational switch, a remarkable level of specificity can be achieved.

However, many challenges and conceptual hurdles remain. As we develop better DNA binding complexes, new questions become important. What parameters besides sequence-specificity determine biological activity? Are exchange rates of complexes on and off the DNA important? Are complex solubilities an issue? Is non-covalent binding affinity sufficient, or is a covalent lesion as with *cis*-platin important, and might such lesions be rapidly repaired? These challenges and others surely await us. Yet it is clear that coordination chemistry provides the rich diversity of structure and reactivity required to meet the challenges ahead.

We thank the National Institutes of Health for their financial support of this work. In addition, we are grateful to our co-workers and collaborators as denoted in the accompanying references.

References

Absalon, M. J., Wu, W., Kozarich, J. W. & Stubbe, J. 1995 Sequence-specific double-strand cleavage of DNA by fe bleomycin.2: mechanism and dynamics. *Biochem.* **34**, 2076–2086.

Barton, J. K. 1986 Metals and DNA: molecular left-handed complements. *Science* **233**, 727–734.

Barton, J. K., Goldberg J. M., Kumar C. V. & Turro, N. J. 1986 Binding modes and base specificity of tris(phenanthroline)ruthenium(II) enantiomers with nucleic-acids: tuning the stereoselectivity. *J. Am. Chem. Soc.* **108**, 2081–2088.

Berg, J. M. 1993 Zinc-finger proteins. *Current Opinion in Structural Biology* **3**, 11–16.

Berman, H. M. & Young, P. R. 1981 The interaction of intercalating drugs with nucleic acids. *Ann. Rev. Biophys. Bioengng* **10**, 87–114.

Bond, P. J., Langridge, R., Jennette, K. W. & Lippard, S. J. 1975 X-ray fiber diffraction evidence for neighbor exclusion binding of a platinum metallointercalation reagent to DNA. *Proc. Natn. Acad. Sci., USA* **72**, 4825–4829.

Burrows, C. J. & Rokita, S. E. 1994 Recognition of guanine structure in nucleic-acids by nickel-complexes. *Acc. Chem. Res.* **27**, 295–301.

Campisi, D., Morii, T. & Barton, J. K. 1994 Correlations of crystal-structures of DNA oligonucleotides with enantioselective recognition by $Rh(phen)_2phi^{3+}$: probes of DNA propeller twisting in solution. *Biochem.* **33**, 4130–4139.

Chapnick, L. B., Chasin, L. A., Raphael, A. L. & Barton, J. K. 1988 In vivo photoreaction of a chiral cobalt complex. DNA cleavage by $Co(DIP)_3^{3+}$ in mammalian cells. *Mutat. Res.* **201**, 17–26.

Chow, C. S. & Barton, J. K. 1992 Transition-metal complexes as probes of nucleic-acids. *Meth. Enzym.* **212**, 219–242.

Chow, C. S., Barnes, C. M. & Lippard, S. J. 1995 A single HMG domain in high-mobility group-1 protein binds to DNAs as small as 20 base-pairs containing the major cisplatin adduct. *Biochem.* **34**, 2956–2964.

Chow, C. S., Behlen, L. S., Uhlenbeck, O. C. & Barton, J. K. 1992 Recognition of tertiary structure in transfer-RNAs by $Rh(phen)_2phi^{3+}$: a new reagent for RNA structure-function mapping. *Biochem.* **31**, 972–982.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Collins, J. G., Shields, T. P. & Barton, J. K. 1994 ^{1}H-NMR of $Rh(NH_3)_4phi^{3+}$ bound to $d(TGGCCA)_2$: classical intercalation by a nonclassical octahedral metallointercalator. *J. Am. Chem. Soc.* **116**, 9840–9846.

David, S. S. & Barton, J. K. 1993 NMR evidence for specific intercalation of Δ-$Rh(phen)_2phi^{3+}$ in $d(GTCGAC)_2$. *J. Am. Chem. Soc.* **115**, 2984–2985.

Dervan, P. B. 1986 Design of sequence-specific DNA-binding molecules. *Science* **232**, 464–471.

Donner, J., Caruthers, M. H. & Gill, S. J. 1982 A calorimetric investigation of the interaction of the lac repressor with inducer. *J. Biol. Chem.* **257**, 4826–4829.

Dupureur, C. M. & Barton, J. K. 1994 Use of selective deuteration and ^{1}H-NMR in demonstrating major groove binding of Δ-$[Ru(phen)_2dppz]^{2+}$ to $d(GTCGAC)_2$. *J. Am. Chem. Soc.* **116**, 10 286–10 287.

Dupureur, C. M. & Barton, J. K. 1995 The assembly of transition metal complexes for site-specific recognition of nucleic acids. In *Comprehensive supramolecular chemistry* (ed. J.-M. Lehn), vol. 5. New York: Pergamon (in the press).

Hecht, S. M. 1979 *Bleomycin: chemical, biochemical, and biological aspects.* Berlin: Springer.

Hertzberg, R. P. & Dervan, P. B. 1984 Cleavage of DNA with methidiumpropyl-EDTA-iron(II): reaction conditions and product analyses. *Biochem.* **23**, 3934–3945.

Hiort, C., Lincoln, P. & Norden, B. 1993 DNA-binding of Δ-$[Ru(phen)_2dppz]^{2+}$ and Λ-$[Ru(phen)_2dppz]^{2+}$. *J. Am. Chem. Soc.* **115**, 3448–3454.

Hudson, B. P., Dupureur, C. M. & Barton, J. K. 1995 ^{1}H NMR structural evidence for the sequence-specific design of an intercalator: Δ-α-$Rh[(R,R)\text{-}Me_2trien]phi^{3+}$ bound to $d(GAGTGCACTC)_2$. *J. Am. Chem. Soc.* **117**, 9379–9380.

Jack, A., Ladner, J. E., Rhodes, D., Brown, R. S. & Klug, A. 1977 A crystallographic study of metal-binding to yeast phenylalanine transfer RNA. *J. Mol. Biol.* **111**, 315–328.

Jennette, K. W., Lippard, S. J., Vassiliades, G. A. & Bauer, W. R. 1974 Metallointercalation reagents, 2-hydroxyethanethiolato(2, 2′, 2″-terpyridine)-platinum(II) monocation binds strongly to DNA by intercalation. *Proc. Natn. Acad. Sci. USA* **71**, 3839–3843.

Krotz, A. H. & Barton, J. K. 1994 Rational construction of metallointercalators: syntheses, structures, and characterizations of isomers of ((2S,9S)-diamino-4, 7,-diazadecane) (9,10-phenanthrenequinonediimine)rhodium(III). *Inorg. Chem.* **33**, 1940–1947.

Krotz, A. H., Hudson, B. P. & Barton, J. K. 1993*a* Assembly of DNA recogntion elements on an octahedral rhodium intercalator: predictive recognition of 5′-TGCA-3′ by Δ-$Rh[(R,R)\text{-}Me_2trien]phi^{3+}$. *J. Am. Chem. Soc.* **115**, 12 577–12 578.

Krotz, A. H., Kuo, L. Y., Shields, T. P. & Barton, J. K. 1993*b* DNA recognition by rhodium(III) polyamine intercalators: considerations of hydrogen-bonding and van der waals interactions. *J. Am. Chem. Soc.* **115**, 3877–3882.

Lerman, L. S. 1961 Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* **3**, 18–30.

Lim, A. C. & Barton, J. K. 1993 Chemical probing of $tDNA^{phe}$ with transition-metal complexes: a structural comparison of RNA and DNA. *Biochem.* **32**, 11 029–11 034.

Magda, D., Wright, M., Miller, R. A., Sessler, J. L. & Sansom, P. I. 1995 Sequence-specific photocleavage of DNA by an expanded porphyrin with irradiation above 700 nm. *J. Am. Chem. Soc.* **117**, 3629–3630.

Manderville, R. A., Ellena, J. F. & Hecht, S. M. 1994 Solution structure of a Zn(II)·Bleomycin A_5-$d(CGCTAGCG)_2$ complex. *J. Am. Chem. Soc.* **116**, 10 851–10 852.

Mastruzzo, L., Woisard, A., Ma, D. D. F., Rizzarelli, E., Favre, A. & Le Doan, T. 1994 Targeted photochemical modification of HIV-derived oligoribonucleotides by antisense oligodeoxynucleotides linked to porphyrins. *Photochem. Photobiol.* **60**, 316–322.

Mei, H. Y. & Barton, J. K. 1988 Tris(tetramethylphenanthroline)ruthenium(II): A chiral probe that cleaves A-DNA conformations. *Proc. Natn. Acad. Sci.* **85**, 1339–1343.

Murphy, C. J. & Barton, J. K. 1993 Ruthenium complexes as luminescent reporters of DNA. *Meth. Enzymol.* **226**, 576–594.

Ohalloran, T. V., Lippard, S. J., Richmond, T. J. & Klug, A. 1987 Multiple heavy-atom reagents for macromolecular X-ray structure determination: Application to the nucleosome core particle. *J. Mol. Biol.* **194**, 705–712.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F. & Sigler, P. B. 1988 Crystal structure of trp repressor/operator complex at atomic resolution. *Nature* **335**, 321–329.

Pabo, C. O. & Sauer, R. T. 1992 Transcription factors: Structural families and principles of DNA recognition. *A. Rev. Biochem.* **61**, 1053–1095.

Price, M. A. & Tullius, T. D. 1992 Using hydroxyl radical to probe DNA structure. *Meth. Enzym.* **212**, 194–219.

Pyle, A. M. & Barton, J. K. 1990 Probing nucleic acids with transition metal complexes. In *Progress in inorganic chemistry* (ed. S. J. Lippard), vol. 38, pp. 413–475. New York: Wiley.

Pyle, A. M., Long, E. C. & Barton, J. K. 1989 Shape-selective targeting of DNA by (phenanthrolinequinonediimine)rhodium(III) photocleavage agents. *J. Am. Chem. Soc.* **114**, 2303–2312.

Pyle, A. M., Morii, T. & Barton, J. K. 1990 Probing microstructures in double-helical DNA with chiral metal-complexes: recognition of changes in base-pair propeller twisting in solution. *J. Am. Chem. Soc.* **112**, 9432–9434.

Rehmann, J. P. & Barton, J. K. 1990*a* ^{1}H-NMR studies of tris(phenanthroline) metal-complexes bound to oligonucleotides: characterization of binding modes. *Biochem.* **29**, 1701–1709.

Rehmann, J. P. & Barton, J. K. 1990*b* ^{1}H-NMR studies of tris(phenanthroline) metal- complexes bound to oligonucleotides: structural characterizations via selective paramagnetic relaxation. *Biochem.* **29**, 1710–1717.

Rich, A. 1993 DNA comes in many forms. *Gene* **135**, 99–109.

Saenger, W. 1984 *Principles of nucleic acid structure.* Berlin: Springer.

Sardesai, N. Y., Zimmermann, K. & Barton, J. K. 1994 DNA recognition by peptide complexes of rhodium(III): example of a glutamate switch. *J. Am. Chem. Soc.* **116**, 7502–7508.

Seeman, N. C., Rosenberg, J. M. & Rich, A. 1976 Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natn. Acad. Sci. USA* **73**, 804–808.

Sherman, S. E., Gibson, D., Wang, A. H. J. & Lippard, S. J. 1985 X-ray structure of the major adduct of the anticancer drug cisplatin with DNA: *cis*-[Pt(NH_3)$_2$(d(pGpG))]. *Science* **230**, 412–417.

Sigman, D. S., Mazumder, A. & Perrin, D. M. 1993 Chemical nucleases. *Chem. Rev.* **93**, 2295–2316.

Singleton, S. F. & Dervan, P. B. 1992 Thermodynamics of oligonucleotide-directed triple helix formation: an analysis using quantitative affinity cleavage titration. *J. Am. Chem. Soc.* **114**, 6957–6965.

Sitlani, A. & Barton, J. K. 1994 Sequence-specific recognition of DNA by phenanthrenequinone diimine complexes of rhodium(III): importance of steric and van der Waals interactions. *Biochem.* **33**, 12 100–12 108.

Sitlani, A. & Barton, J. K. 1995 Recognition of nucleic acids by transition metal complexes. In *Handbook on metal–ligand interactions in biological fluids*, vol 1, pp. 466–487. New York: Marcel Dekker.

Sitlani, A., Dupureur, C. M. & Barton, J. K. 1993 Enantiospecific palindromic recognition of 5′-d(CTCTAGAG)-3′ by a novel rhodium intercalator: analogies to a DNA-binding protein. *J. Am. Chem. Soc.* **115**, 12 589–12 590.

Sitlani, A., Long, E. C., Pyle, A. M. & Barton, J. K. 1992 DNA photocleavage by phenanthrenequinone diimine complexes of rhodium(III): shape-selective recognition and reaction. *J. Am. Chem. Soc.* **114**, 2303–2312.

Shields, T. P. & Barton, J. K. 1995 Sequence-selective DNA recognition and photocleavage: a comparison of enantiomers of Rh(en)$_z$phi^{3+}. *Biochem.* **34**, 15 037–15 048.

Steitz, T. A. 1990 Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding. **23**, 205–280.

Sundquist, W. I. & Lippard, S. J. 1990 The coordination chemistry of platinum anticancer drugs and related-compounds with DNA. *Coord. Chem. Rev.* **100**, 293–322.

Terbrueggen, R. H. & Barton, J. K. 1995 Sequence-specific DNA binding by a rhodium complex: recognition based on sequence-dependent twistability. *Biochem.* **34**, 8227–8234.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Tullius, T. D. 1990 *Metal–DNA chemistry*. American Chemical Society.

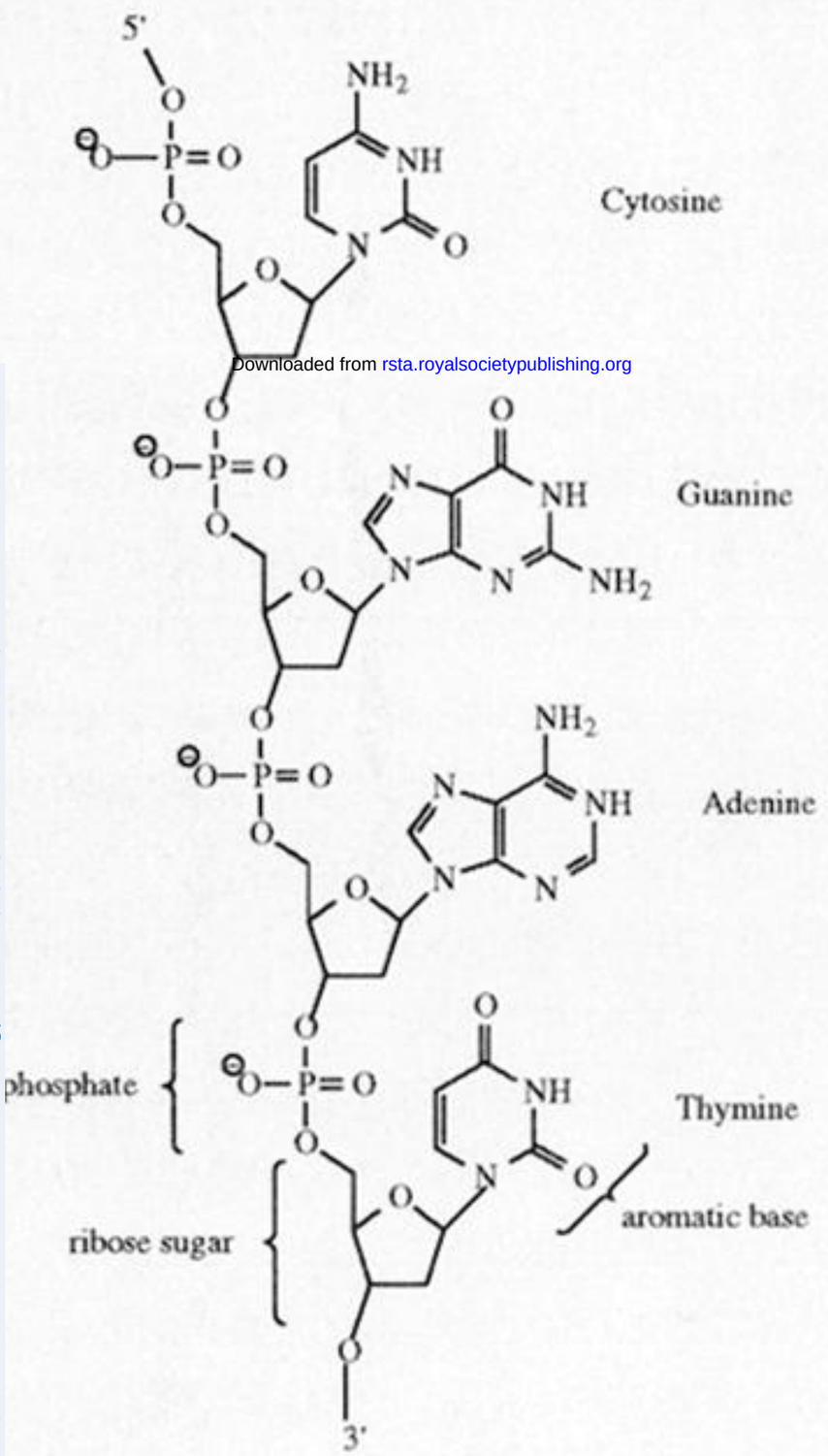
Wang, A. H.-J., Nathans, J., van der Marel, G., van Boom, J. H. & Rich, A. 1978 Molecular structure of a double helical DNA fragment intercalator complex between deoxy CpG and terpyridine platinum compound. *Nature* **276**, 471–474.

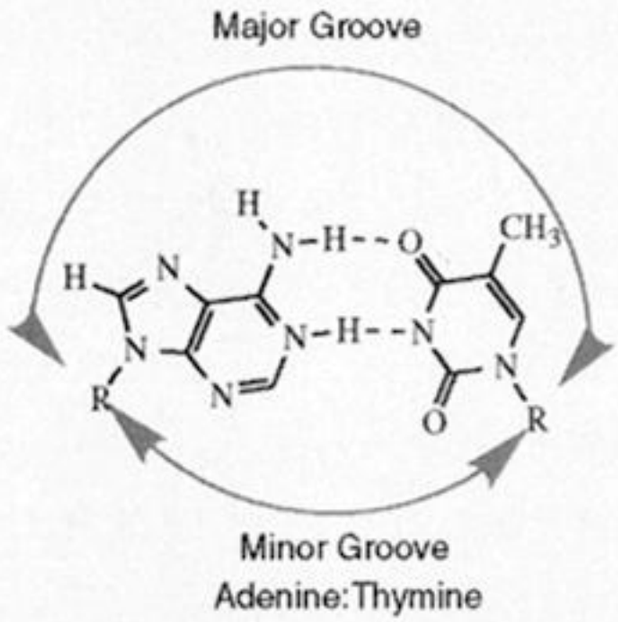
Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A. & Weiner, A. M. 1987 *Molecular biology of the gene*, 4th edn. New York: Benjamin.

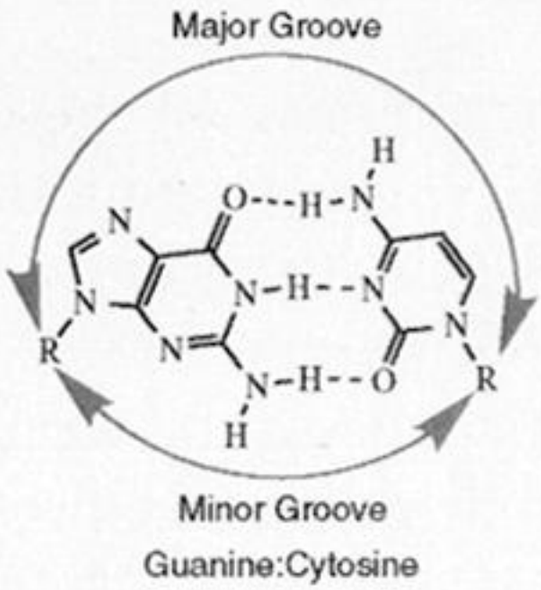
Wu, A., Vanderwall, D. E., Stubbe, J., Kozarich, J. W. & Turner, C. J. 1994 Interaction of Co·Bleomycin A2 (Green) with d(CCAGGCCTGG)$_2$: evidence for intercalation using 2D NMR. *J. Am. Chem. Soc.* **116**, 10 843–10 844.

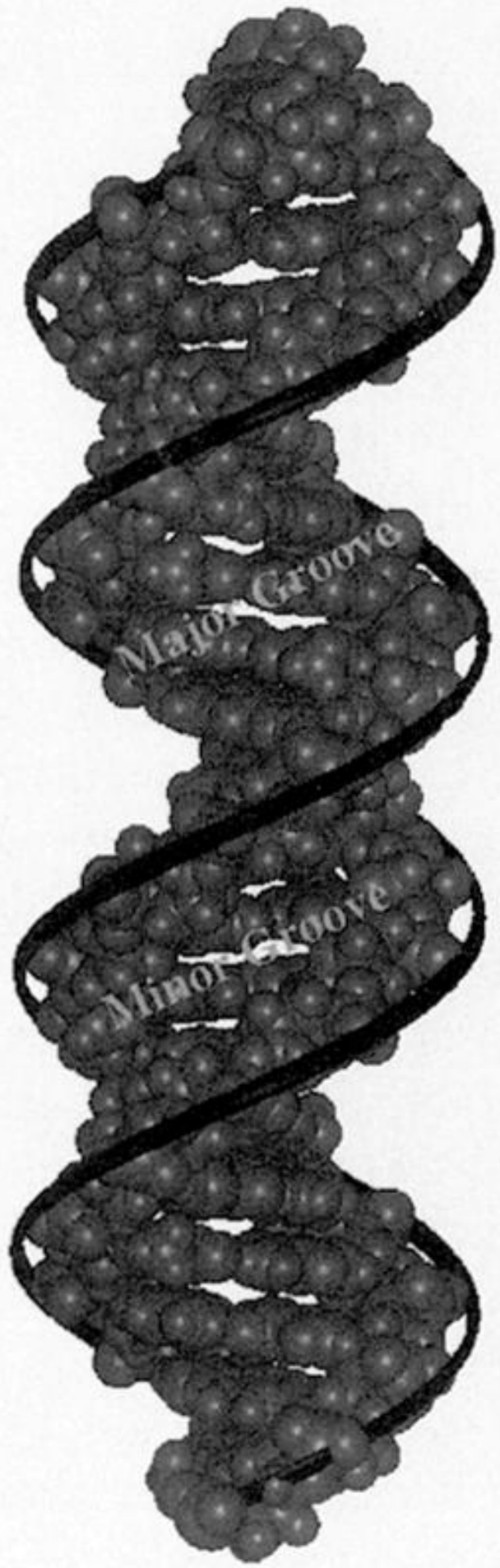
Wu, H. M. & Crothers, D. M. 1984 The locus of sequence-directed and protein-induced DNA bending. *Nature* **308**, 509–513.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES









igure 1. The building blocks of a DNA polymer (left). The major and minor groove sides of denine:thymine and guanine:cytosine base pairs (middle). Space filling model of the major and inor grooves in a DNA polymer (right). The ribbons trace the phosphate backbone.

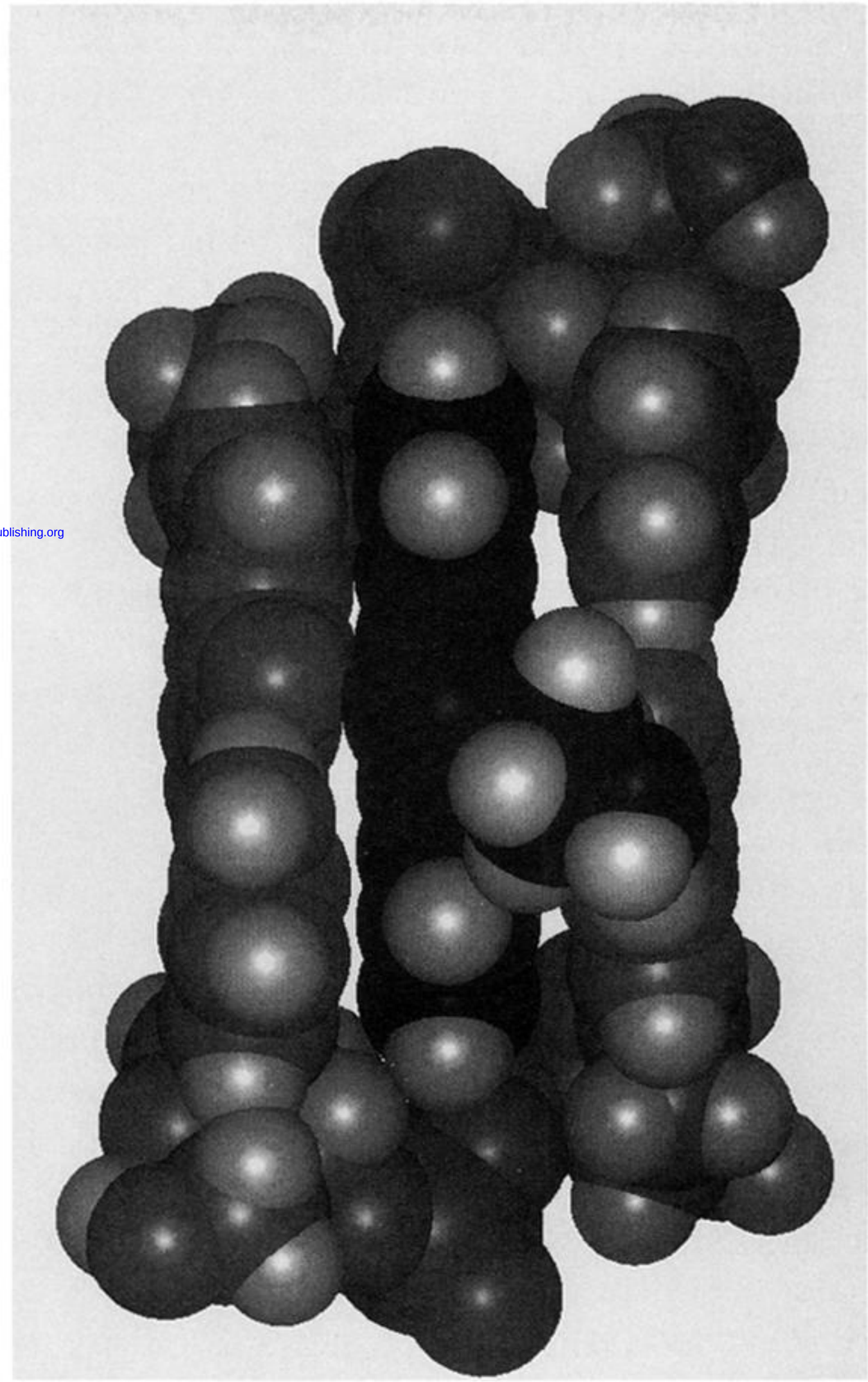


Figure 2. Space filling illustration of Pt(terpy)HET$^+$ (HET = 2-hydroxyethanethiolate) intercalated into the dinucleotide $(dCpG)_2$. View is perpendicular to the helix axis (adapted from Wang *et al.* 1978).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

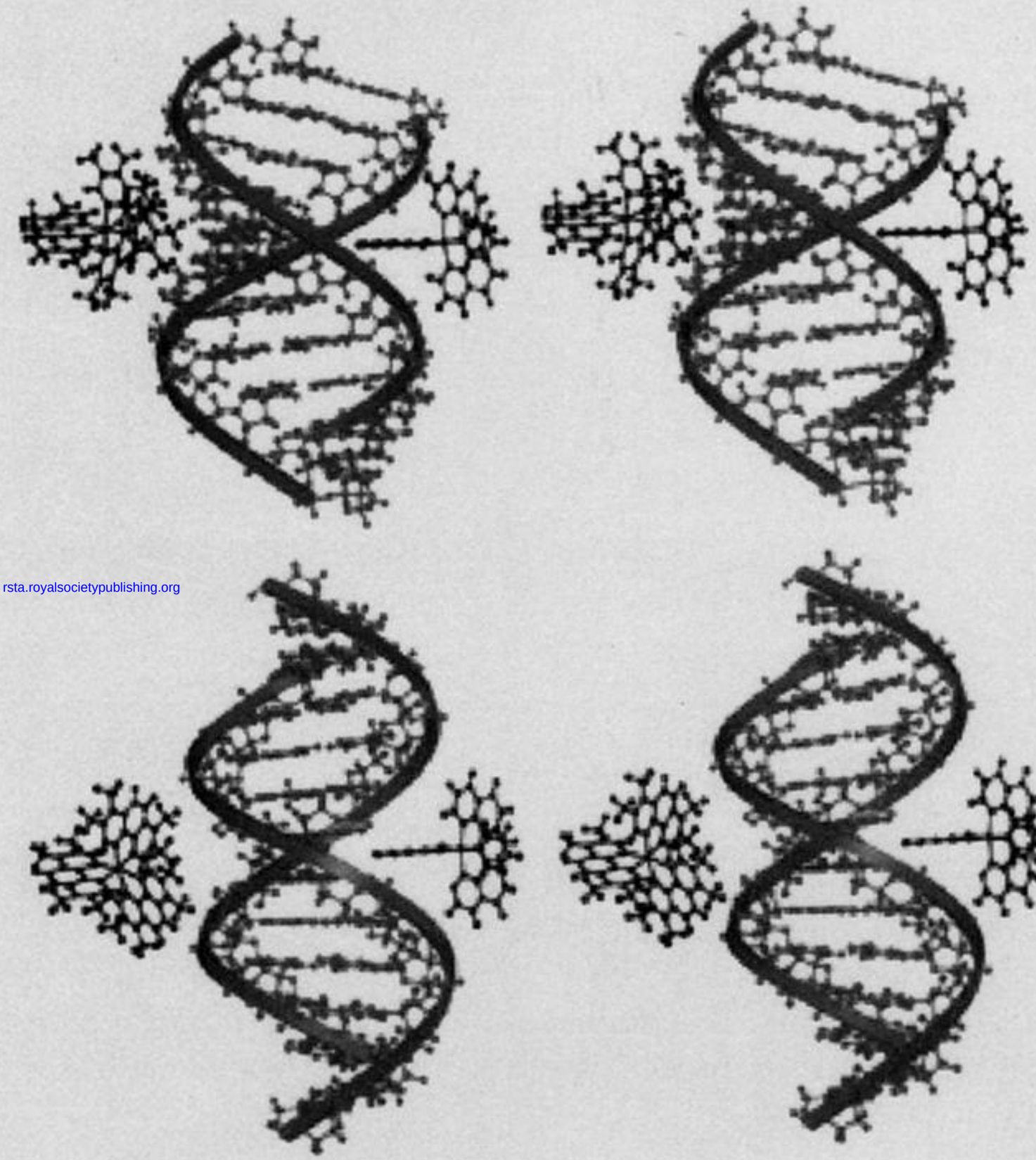


Figure 4. Stereoviews illustrating the size and shape complementarity of Λ-Ru(TMP)$_3^{2+}$ (left) and Δ-Rh(phen)$_2$phi^{3+} (right) with the grooves of A-form (top) and B-form (bottom) nucleic acids. Λ-Ru(TMP)$_3^{3+}$ is illustrated bound against the minor groove and Δ-Rh(phen)$_2$phi^{3+} is poised for intercalation in the major groove. In A-form nucleic acids, the shallow minor groove presents an accessible surface for binding by Λ-Ru(TMP)$_3^{2+}$. Meanwhile the backbone in the major groove has closed down, preventing the intercalation of Δ-Rh(phen)$_2$phi^{3+}. In contrast, the tetramethylphenanthroline ligands of Λ-Ru(TMP)$_3^{2+}$ are too large to contact the base stack in the minor groove of B-form nucleic acids without clashing with the phosphate backbone. In the major groove, Δ-Rh(phen)$_2$phi^{3+} has ample room to intercalate into the base stack.

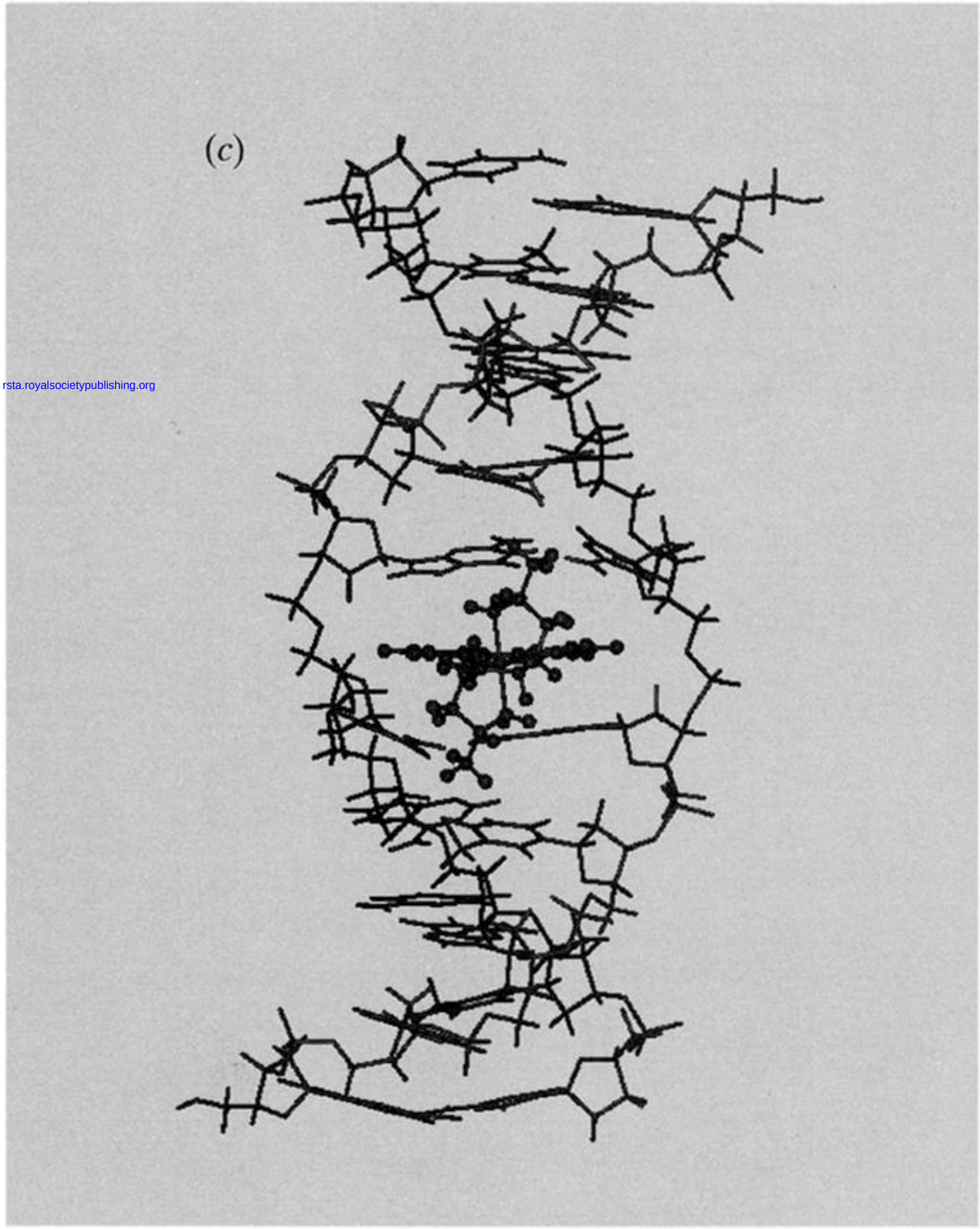
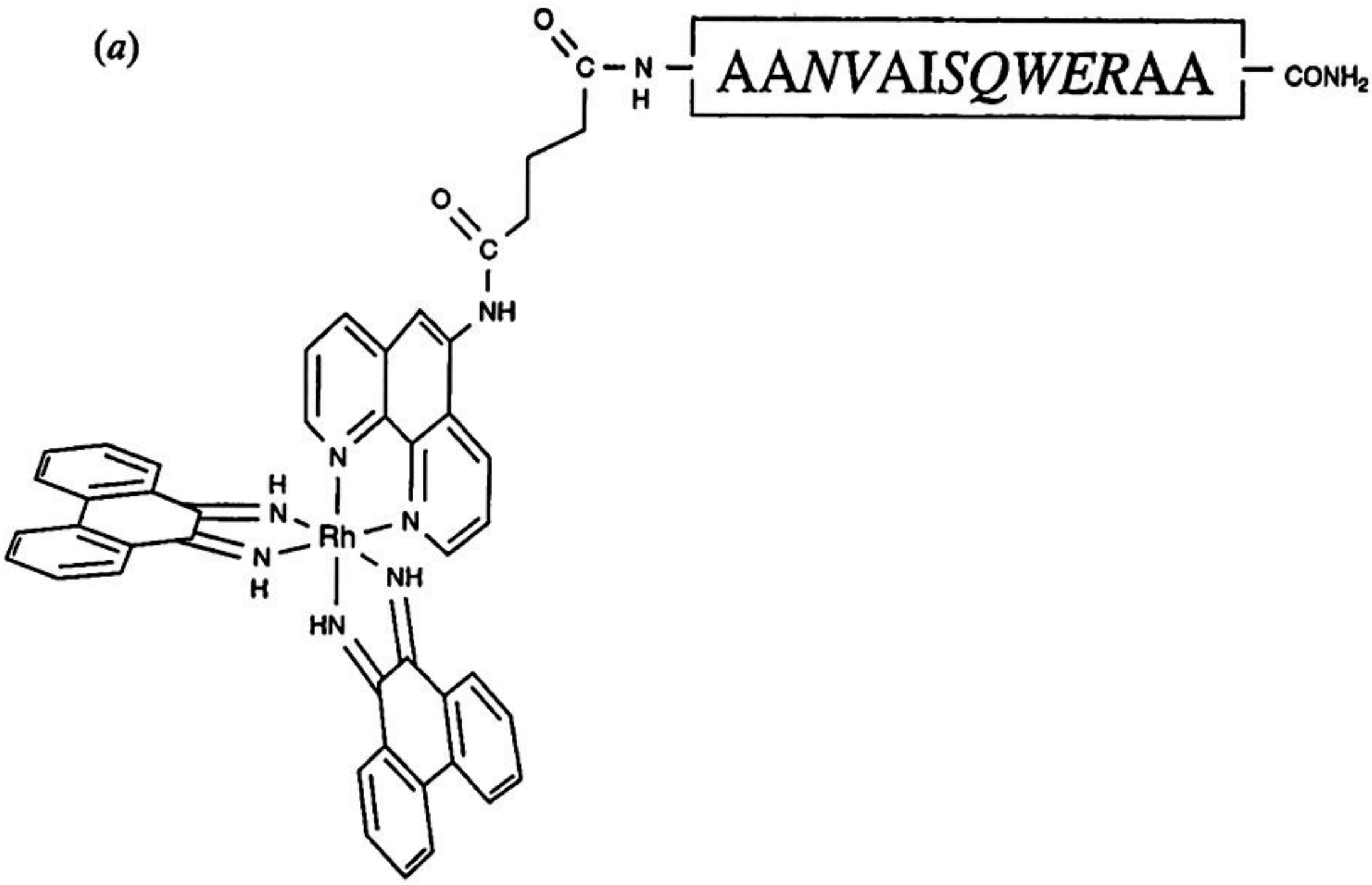


Figure 11. (*c*) NMR structural model of Δ-α-Rh[(R,R)-Me_2trien]phi^{3+} bound to the decamer d(GAGTGCACTC)$_2$.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(a)



(b)

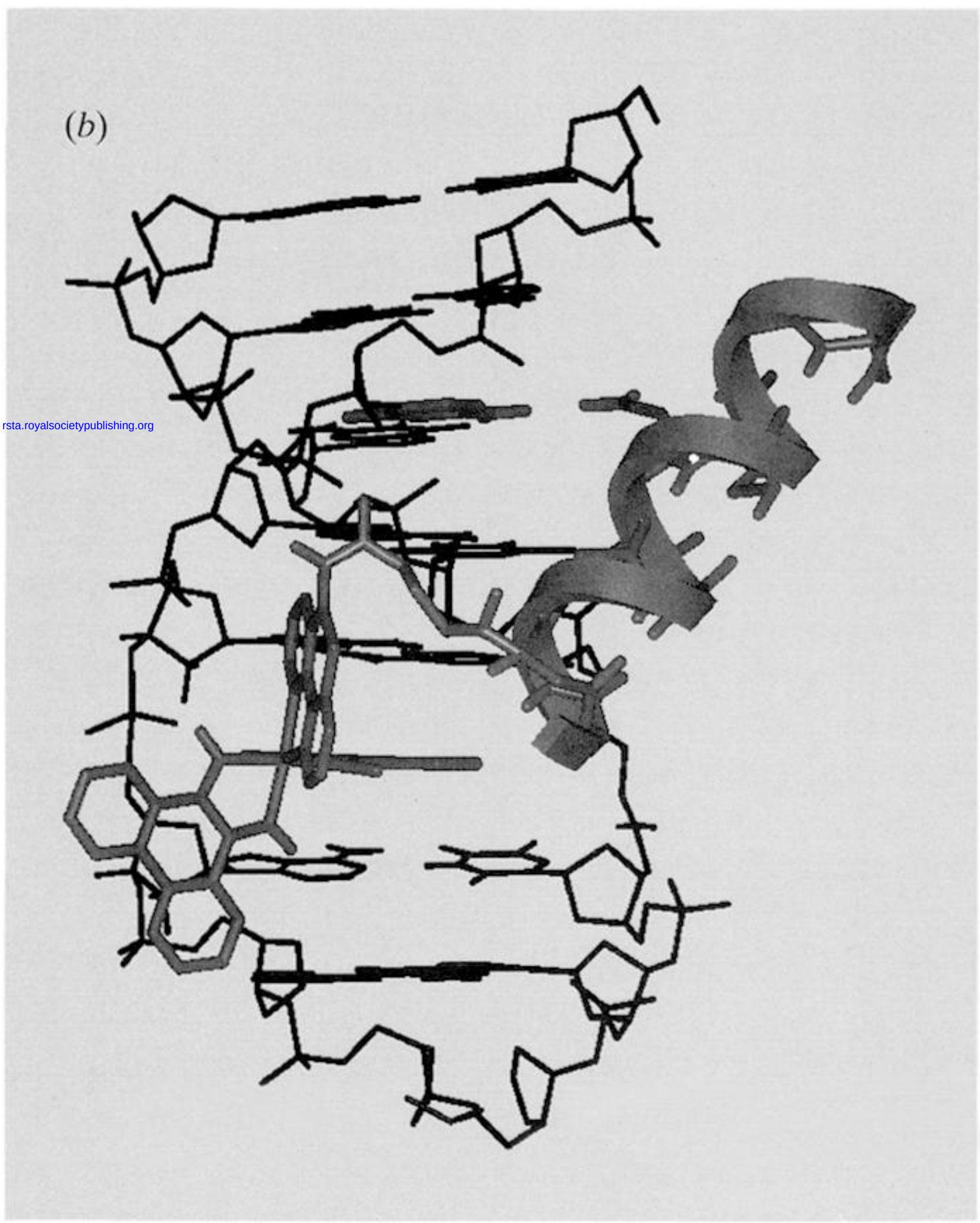


Figure 14. (*a*) A metal–peptide intercalating chimera and (*b*) a model of the complex bound to its recognition sequence. Note the α-helical conformation of the appended peptide with the glutamate side chain disposed for a specific contact with cytosine.