

Interactions between metal ions and DNA

Book or Report Section

Accepted Version

Cardin, C. J. ORCID: <https://orcid.org/0000-0002-2556-9995>
(2019) Interactions between metal ions and DNA. In: Structure
and Bonding. Springer, pp. 203-237. doi:
10.1007/430_2019_42 Available at
<https://centaur.reading.ac.uk/88302/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: http://dx.doi.org/10.1007/430_2019_42

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Interactions between metal ions and DNA

Christine J Cardin, Department of Chemistry, University of Reading

Abstract

84 years elapsed between the announcements of the Periodic Table and that of the DNA double helix in 1953, and the two have been combined in many ways since then. In this chapter an outline of the fundamentals of DNA structure leads into a range of examples showing how the natural magnesium and potassium ions found in nature can be substituted in a diversity of applications. The dynamic structures found in nature have been studied in the more controlled, but artificial environment of the DNA crystal using examples from sodium to platinum, and also in a range of DNA-binding metal complexes. While nmr is an essential technique for studying nucleic acid structure and conformation, most of our knowledge of metal ion binding has come from X-ray crystallography. These days the structures studied, and therefore also the diversity of metal binding, go beyond the double helix to triplexes, hairpin loops, junctions and quadruplexes, and the chapter describes briefly how these pieces fit into the DNA jigsaw. In a final section, the roles of metal cations in the crystallisation of new DNA structures are discussed, along with an introduction to the versatility of the Periodic Table of absorption edges for nucleic acid structure determination.

Abbreviations

DNA bases are A-adenine, C-cytosine. G-guanine, T- thymine.

DNA – deoxyribonucleic acid

NDB – Nucleic Acid Database

PDB - Protein Data Bank

DDD – Drew-Dickerson dodecamer

MPD – 2,4-dimethylpentanediol

RNA – ribonucleic acid

FRET – fluorescence resonance energy transfer

phen – 1,10-phenanthroline

TAP – 1,4,5,8-tetraazaphenanthrene

dppz – dipyridophenazine

MLCT – metal to ligand charge transfer

Bpy – 2,2' -bipyridyl

TEM – transmission electron microscopy

Salphen - N,N'-phenylenebis(salicylideneimine)

1.Introduction

1.1 DNA is the biggest ligand

The Periodic Table is a 19th century scientific ‘icon’ and the DNA Double Helix is a 20th century ‘icon’. The main aim of this Chapter is to consider the consequences that result when they are considered simultaneously. At this stage we cannot look into the future and predict those structures which will be seen as having changed our thinking in the 21st Century. Nevertheless, it represents an interesting challenge to consider the consequences of considering the interactions of the DNA molecule with those metal ions which form the major part of the Periodic Table. At its simplest the DNA molecule

can be viewed as a very large and complex ligand which is capable of co-ordinating to a very large number of metal ions. The intellectual structure of the Periodic Table enables chemists and biochemists to explore the way in which metal ions interact with DNA molecules in a systematic way. In this short chapter it is not possible to give more than a few examples of the diversity of possibilities which can arise from combining the unique individual elements with nucleic acids, but nonetheless it will illustrate the depth of the relationship between these two icons. A linking technique is that of X-ray crystallography. A key property of X-rays is that the intensity of scattering is determined by the electron cloud, and hence in most cases leads directly to a determination of atom type. It has therefore since its inception been a key technique for inorganic chemists, and has gone from being an inherently complex and intellectually demanding technique to being an extremely efficient, cost-effective and routine technique since the double helix was first proposed.¹

Initially DNA may be viewed as an 'organic' molecule, and its base pairing abilities lead to the double helix structure first proposed by Crick and Watson using the X-ray diffraction data of Franklin and Wilkins. It also has a phosphate backbone which plays a very important role in the determining its chemical and physiological properties. For each base in a single DNA strand there is a phosphate group which has a negative charge and forms part of the backbone. The stability of the DNA molecule arises principally from the stacking interaction between the bases, not any inherent property of the phosphate, but the hydrophobic bases in the interior of the DNA duplex are solubilised by the deoxyribose and the phosphate components, making the molecule overall a water-soluble species. The size of the molecule is mind-boggling for a chemist. We have 23 pairs of chromosomes, which for evolutionary reasons are arranged by size, so that chromosome 1 is the largest. This chromosome consists of about 247 million base pairs, or roughly 15 billion atoms, remembering that the double helix actually consists of two complementary molecules (Figure 1). So how long is it? Fortunately the unrolled length is determined by the stacking interaction, otherwise this would be difficult to estimate. Assuming that a single stacking repeat is 3.34 Å, the overall length would be about 8.5 cm. Within each molecule there are therefore 247 million phosphate anions, which are neutralised by cations. Within the coiled and supercoiled chromatin of the cell nucleus, which is about 6 µm in diameter, much of the charge is neutralised by the positively charged histone proteins, around which the double helix coils, but there is also an important role for metal cations, both those naturally occurring and those which can be useful as tools to investigate structure and those which can be of therapeutic benefit.

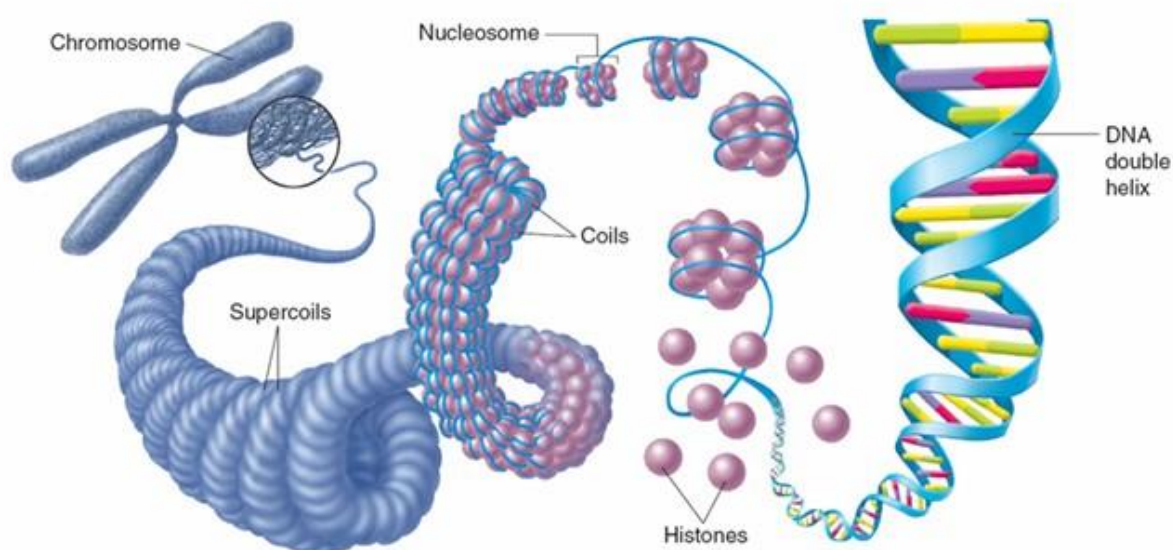


Figure 1. From the chromosome to the double helix.

<https://socratic.org/questions/56116c7e11ef6b17032ddc8b>

1.2 Binding sites on DNA

The presence of the negative charge on each phosphate group implies the presence of a cation which leads to electroneutrality. The cations can be organic, as in the tetra-positive spermine molecule, or in the protein histones. In native DNA the metal counterion is Mg^{2+} , which raises structural questions concerning whether this ion is directly co-ordinated to the phosphate group or is a hydrated species $[\text{Mg}(\text{OH}_2)_6]^{2+}$ and whether the ions are linked to the phosphates by hydrogen bonding. This knowledge immediately opens up the possibility of modifying the DNA conformation or topology by changing the metal ions using principles derived from the Periodic Table. Specifically, vertical relationships in the Periodic Table suggest the replacement of Mg^{2+} by Ca^{2+} , Sr^{2+} or Ba^{2+} , and these substitutions have been explored by X-ray crystallographers wishing to elucidate the structures of these complex biological molecules.² These replacements may also affect the structural and biochemical properties of DNA. If the ions are present as hydrated species then the range of heavy atom substitutions increases, because one may access the hydrated ions known for a large number of main group, transition metals and rare earth metal ions. A collaborating inorganic chemist may also suggest substitutions based on isoelectronic principles and suggest exploring the structural and biochemical consequences of replacing Mg^{2+} by Na^+ or Al^{3+} or related metal ions from these alternative groups of the periodic table. There have been unexpected results from such exercises.^{3,4} Inorganic chemists tend to classify metal ions as having hard or soft characteristics and the 'hard'- 'soft' principle may be applied to produce variants. For the DNA molecule, this leads to the conclusion that broadly speaking there are two sorts of binding – the 'hard' cations would be expected to be associated with the phosphate backbone oxygen atoms, whereas the 'softer' cations would bind to the nucleobases. The essential ions of the first row transition series (Mn^{2+} , Fe^{2+} , Co^{2+} , Cu^{2+} , Zn^{2+}) are often described as having intermediate hardness because they can display the most versatile coordination chemistry, which means that with nucleic acids they might bind at several sites. For an

interesting example of the switchover, see the discussion of the d(GCATGCT) structures in section 4. In living systems the organisation of the cell dictates what ions can be present with any biomolecule, and in the case of the DNA of humans, Mg^{2+} is the essential ion in the chromatin of the cell nucleus. Other metal ions which interact with the DNA of a healthy cell, such as the zinc of the so-called zinc finger proteins, are tightly coordinated to their protein ligand and cannot interact directly. Any other free ion can be classified as toxic, non-toxic or possibly of medicinal benefit, which often correlates with the 'hard'/'soft' classification.

Figure 2 highlights the structural effect of these differences, showing that the essential role of magnesium, in Figure 2a, is mainly that of charge neutralisation. It is labile, creates water ordering with its large hydrated radius, and often not directly bound. The most important example of nucleobase binding, well studied since 1963 for its medicinal importance, is the formation of the $[\text{Pt}(\text{NH}_3)_2]^{2+}$ adduct at two adjacent guanine N7 positions on the same DNA strand (intra-strand binding), which lead to the introduction of a $\sim 50^\circ$ kink in the DNA double helix and opens up the minor groove to the binding of HMG and other proteins, as shown in Figure 2b. ⁵ This is an unnatural intervention, and should be classified as toxic, but nevertheless provides the basis of an important cancer treatment. The initial discovery of the therapeutic benefit of platinum was accidental, but the subsequent exploration of other possibly useful elements was very much inspired by the Periodic Table. This topic is covered in a later section.

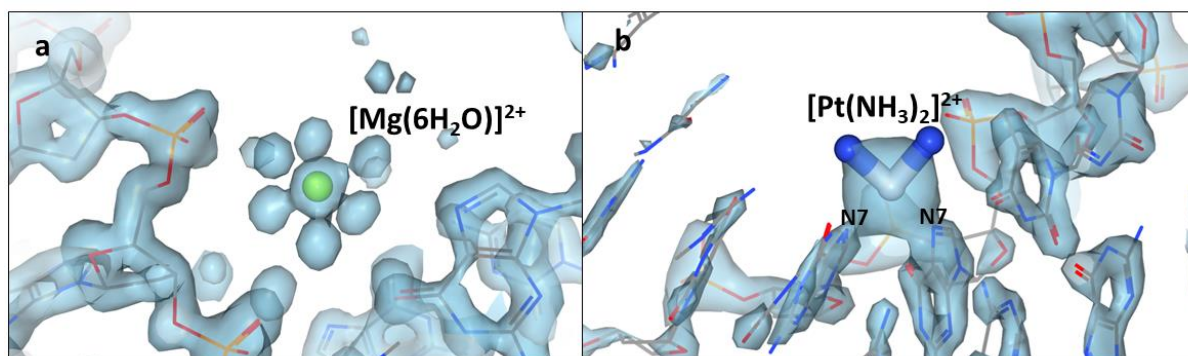


Figure 2 Electron density maps for the visualisation of metal ion binding to DNA. a) A fully hydrated magnesium cation (from PDB code 4KWX) b) A directly bound platinum(II)bis(amine) cation (from PDB 3LPV))

In addition to the classic Lewis acid-Lewis base type interactions between metal ions and alternative sites on the DNA strands, cationic planar aromatic species may intercalate between the base pairs. The strength of this interaction, as with the stability of the duplex, is determined largely by the donor-acceptor nature of the stacking interaction. Typically guanine is an electron-rich donor and the cation is the electron-poor acceptor. Metal cations have a role to play here, as part of coordination compounds in which the electronic properties can be tuned through the properties of the metal. Some of the classic intercalating drugs such as acridine-based chromophores, daunomycin and the camptothecins are known to be topoisomerase poisons (they interfere with the DNA unwinding enzymes). ⁶ This has important consequences for the replication process of DNA. Metal complexes, particularly those with planar arrangements of ligands are also able to intercalate. ⁷

2. DNA structures – the building blocks and the diversity

2.1 Natural and unnatural DNA

The double helix of DNA was iconic partly because of its simplicity.⁸ The idea of linear information storage and a code appealed to the computer age, and the first computers appeared around the same time (the 1950s) as the world learnt that the 'ribbon' of the twisted helix contained, in a single topological dimension, the code of life. The famous last sentence of Watson and Crick's paper (It has not escaped our notice...) implied that the code just had to be read for the secret of life to be revealed. It has turned out to be not as simple as that. Fifty years later (the 2000's), and the human genome is now known to contain about 19,000- 20,000 genes (codes for proteins) accounting for maybe 2% of our DNA, with the other 98% of so-called 'junk' DNA has no precise description of its role, but is probably just something we don't yet understand. The regulation of gene expression is a whole subject in itself, as is the mechanism by which the 'code' is assembled by splicing together the appropriate fragments. And in addition to the double helix we now have a repertoire of other structures – hairpins, cruciforms, triple helices, DNA junctions, the G-quadruplex, the i-motif – none of which have that 'iconic' simplicity but which increasingly appear to have defined biological roles, and which extend also our ideas of what is 'natural'. The natural ions of DNA have therefore to be extended at least as far as K⁺ in the G-quadruplex, only shown to be present in cells since 2013, and of which more later.⁹

The unnatural DNA is that from the test tube. The original 'test tube baby' is now a healthy middle-aged woman. While it is not realistic to synthesise a whole eukaryotic genome from scratch, DNA synthesis is nevertheless a 'cottage industry' with a range of start-up companies supplying fragments which can be used for genetic modification. For the chemist, in 1980 the availability of short synthetic DNA fragments – oligonucleotides – allowed the possibility of working with DNA as a chemical entity, more or less like any other, and hence to explore the whole Periodic Table of possible interactions. It also allowed the organic chemist's equivalent of exploiting the trends in Group 16 – the replacement of oxygen by sulphur and selenium.¹⁰ The aim of the work described in this reference is to increase the stability of a DNA triplex, but that is one of many rationales for such work.

2.2 The building blocks of DNA and their assembly

There are three components to the DNA molecule – phosphate, sugar and base. Together they make up one residue in the DNA sequence.

2.2.1 Phosphate The phosphate component contributes the overall negative charge on each residue and, by forming only single bonds to the ribose oxygen atoms, contributes to the flexibility of the backbone. The directionality of a nucleic acid strand is defined by the two of the phosphate oxygen atoms linked to the sugar, and by convention the strand runs from the 5' end to the 3' end, with the two strands of the duplex running in opposite directions (antiparallel). The formal description of a backbone conformation starts at the phosphorus atom and defines six backbone torsion angles between it and the next phosphorus atom, a distance of about 7 Å around the outside of the helix in a typical B-DNA structure. In the cartoon representation of a nucleic acid structure, used in some of the figures in this chapter, the backbone is defined by the phosphorus atom positions (e.g. Figure 7 showing the d(CGCGAATTCGCG)₂ dodecamer structure.¹¹

2.2.2 Sugar The chirality of each nucleotide (residue) of the DNA molecule arises from the deoxyribose sugar component, as does the great flexibility of the molecule, with the sugar having a low energy barrier to interconversion between several forms of the five-membered ring. In normal B-form DNA, the conformation of the ring is different from that in RNA (which contains an additional -OH at the 2' position). In RNA the 2'-OH is in the axial orientation, but in B-DNA the 2'-H at the same position is equatorial. This change affects the whole structure, changing it from B-form to A-form. Originally, the A-form was induced in DNA fibres by dehydration, and gave a characteristically different fibre diffraction pattern as seen (but not interpreted) by Rosalind Franklin. Today the A-form is principally associated with helical RNA, whenever it is found. It has been found that a DNA crystal can be reversibly dehydrated using controlled humidity and showing a partial B-A transition in some of the sugar rings.¹²

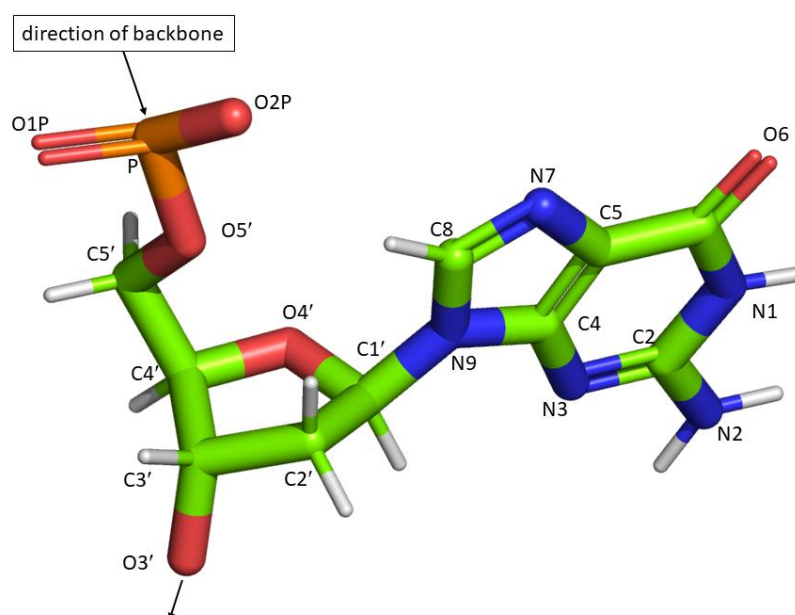


Figure 3 A single residue of DNA. The guanine base is in the *anti* orientation with respect to the sugar ring. The purine base atoms are numbered 1-9, while the sugar component numbering always includes a prime (').

The sugar is connected by a single bond to the DNA base, and there is some rotational freedom about this bond, with a pattern of preferred conformations. In the normal B-DNA structure, the sugar ring and the base are approximately at right angles, the *anti* conformation, shown in Figure 3, with the Watson-Crick base-pairing atoms (O6, 1-NH and 2-NH₂) remote from the sugar ring. The principal alternative, the *syn* conformation, places the base much closer to the sugar and is sterically less favourable. It is of importance because it is partly responsible for the diversity of topologies which are possible for the G-quadruplex, a topic of great current interest. Historically, it was important because it was found in the first ever single crystal structure of a DNA oligonucleotide, the hexamer duplex d(CGCGCG)₂. In that structure, the so-called Z-DNA, the zig-zag appearance arises from the alternating *anti* and *syn* conformations, with the *syn* conformation adopted by all the guanine residues. The formation of this structure can be induced by cations such as [Co(NH₃)₆]³⁺, which form specific groove interactions and stabilise the crystal packing. The Z-DNA conformation was primarily seen as an (undesirable) academic curiosity for many years, but it is now known to form transiently as part of a

specific enzyme recognition process, so has taken its place among the DNA structures known to have a biological role.¹³ It can also be recognised specifically by some metal complexes.¹⁴

2.2.3 Bases and pairing schemes The basepairing specificity and recognition properties of a DNA sequence arises from the combination of donors and acceptors on the bases. In contrast to the sugar component, the bases are rigid planar heterocyclic ring systems, with each of the four bases having a distinct distribution of polarity and potential hydrogen bonding patterns. The original Watson-Crick base-pairing scheme explained the specificity of CG and AT base-pair formation and hence defined our understanding of the basis for the fidelity of the DNA copying process. A much less well known name than Watson and Crick is that of Karst Hoogsteen.^{15–17} A graduate of Groningen, he left for the US in 1956 to work with Linus Pauling. Later he worked as an industrial crystallographer for 29 years, living until 2015. Nevertheless, he was the first to demonstrate that there were completely different hydrogen bonding possibilities associated with what is sometimes referred to as the ‘Hoogsteen face’ of the DNA bases, which he originally saw when he determined the X-ray structure of 9-methyladenine and 1-methylthymine, but which had been proposed to account for the formation of triplexes.

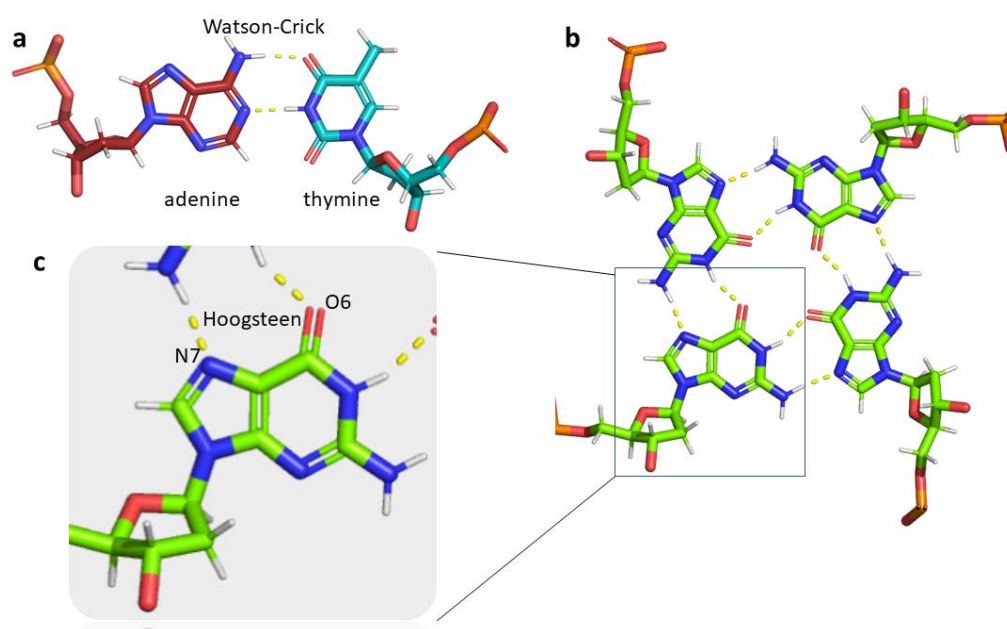


Figure 4 Watson-Crick and Hoogsteen hydrogen bonding. a) Watson-Crick base-pair between an adenine (red) and Thymine (cyan) residue. Note that only two hydrogen bonds can be formed. b) Hydrogen bonding in the G-quartet. Note that each guanine base forms four hydrogen bonds, using two different faces of the base, sometimes referred to as the ‘Watson-Crick face’ and the ‘Hoogsteen face’. c) zoom on one of the four bases to show the ‘Hoogsteen face’. The remaining lone pair on O6 forms the coordinate bond to K^+ , required for the stability of the overall assembly. The base colour scheme follows that of the Nucleic Acid Databank, and is guanine -green; adenine- red; thymine – blue; cytosine – yellow.

Hoogsteen-type bonding has been observed in a range of contexts, including the formation of the guanine quadruplex known as the G-quartet, which is now known to bind K^+ specifically, and therefore an interesting structure for its cation recognition abilities¹⁸ (Figure 4). The authors of this paper worked with Tl^+ as a monovalent ion substitute for K^+ , and compare their work with other structures

containing Na^+ , NH_4^+ and Rb^+ as well as Tl^+ and K^+ , highlighting the importance of size and coordination number of these monovalent ions. Other examples of Hoogsteen-type bonding involve AT-rich DNA sequences, loop regions in DNA, in structures of DNA bound to antibiotics and proteins, damaged and chemically modified DNA, and in polymerases that replicate DNA via Hoogsteen pairing. These structures are often dynamic and transient in the cell and so their existence may not be suspected, but can be seen *in vitro* by nmr.

2.2.4 Cytosine protonation The Watson-Crick scheme gives two distinct pairing patterns which rely on the approximately neutral pH of the system for correct functioning, with guanine protonated on N1 and thymine on N3. Specifically the scheme breaks down if cytosine is protonated (pK_a 4.45 of cytosine N3), whereas adenine (pK_a N1 = 4.1) will be less affected. In slightly acidic conditions, therefore, there are other possibilities. One which has significant therapeutic potential is the use of DNA triplexes.¹⁹ A third strand can be added in the major groove which has the pairing schemes T-AT and $\text{C}^+\text{-GC}$, which is a favoured design option because two triplexes thus formed are structurally isomorphic, as are the Watson-Crick pairs themselves, allowing the formation of a smooth triple helical structure. (Figure 5) This concept is important in the design of gene-targeting agents, and lends itself to strategies for the design of oligonucleotides modified in base, sugar or phosphate components to increase the binding strength of the resultant triplex assembly. The synthetic pyrimidine analogue pseudoisocytosine (ref?) is an example of a modified base which can be used to create stable triplexes at pH 7. Can't find a PNA triplex structure, do you have a ref?

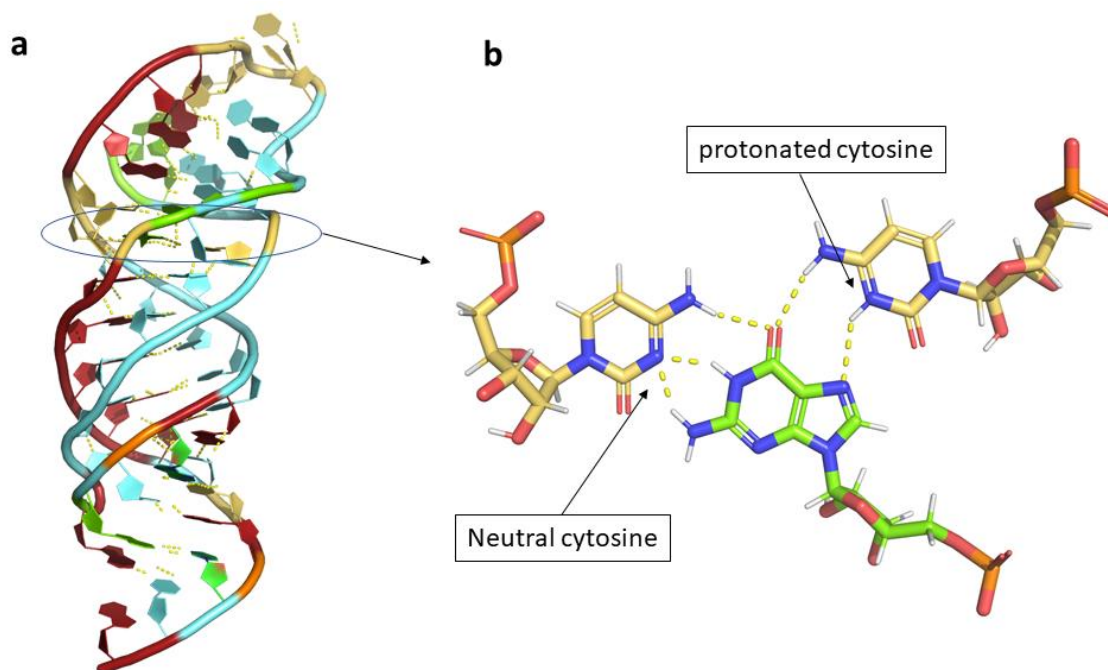


Figure 5 Triplex formation. There is very limited structural information available on triplex formation. The example shown here is from the nmr structure of a so-called RNA pseudoknot, PDB code 2MK8. a) the complete fold, with residues coloured adenine -red, cytosine- yellow, guanine – green and uracil – pale cyan. Thymine is 5'-methyluracil, and is the corresponding DNA base. b) a triplex from this structure showing the presence of both neutral and protonated cytosine.

A second way in which cytosine protonation can play a role is in the formation in cytosine-rich DNA sequences of the so-called i-motif structure.²⁰ If a DNA sequence is potentially G-quadruplex forming, as will be described later, the complementary strand will be C-rich and therefore potentially i-motif forming. This motif is only formed by pairing between two cytosine bases, one of which is protonated (Figure 6). The structure does not require a cation in the ways that the G-quadruplex does, but the role of the linear two-coordinate central H^+ can be played by a monovalent cation such as Cu^+ or Au^+ in the context of making use of the Periodic Table for nanotechnological and design applications.²¹ In this example, the redox properties of copper and its co-ordination preferences can be exploited to make a redox-switchable system. Such a naturally occurring structure will obviously have a pH optimum for formation, which is not necessarily 4.45, as the pK_a of the cytosine base is very sensitive to environment. Following on from the 2013 demonstration that G-quadruplexes were formed in cells, the same antibody technology was used in 2018 to show that the i-motif structure forms transiently in living cells.²² Rather convincingly, the number of i-motifs detected by this fluorescence method is clearly reduced even in slightly alkaline cellular pH, as would be expected from *in vitro* experiments.

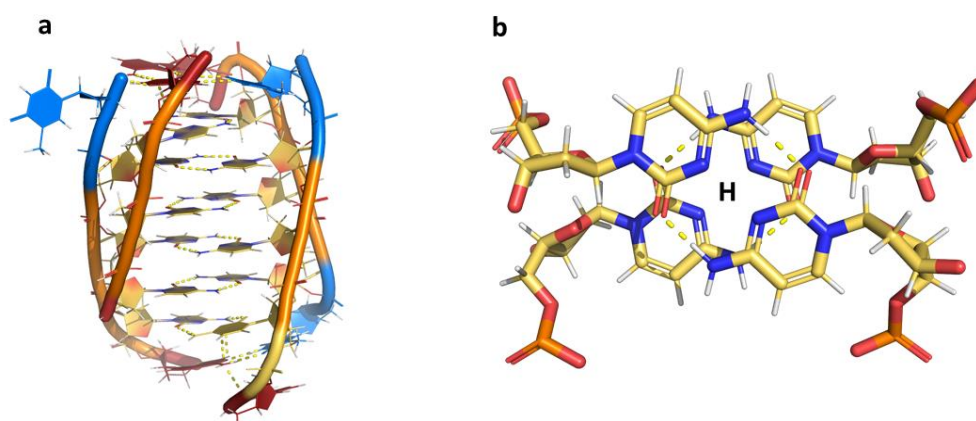


Figure 6 The i-motif a) The X-ray structure assembled from four strands of the DNA sequence d(ACCCT). (PDB 1BQJ). b) two successive steps of the 'ladder' showing the characteristic alternation of steps. Each cytosine pair must be linked by a single proton, indicated by H in the figure but not located by the experiment. The H^+ are thus linear two coordinate. This structure is remarkable because it lacks stacking interactions between successive CH^+-C pairs.

2.2.5 Base mispairing. Mistakes in the copying of DNA before cell division occur even in a healthy human being, and may be corrected by several DNA repair enzymes to maintain the integrity of the genome, with mismatches if uncorrected capable of giving rise to mutations which can lead to cancer. DNA damage can also cause potentially cancerous mismatches, of which a well known example is the oxidation of guanine. Guanine is the most readily oxidised of the four DNA bases, giving a guanine radical cation, and a major product in the presence of singlet oxygen is the species 8-oxoguanine. The presence of 8-oxoguanine can be recognised by a repair enzyme, but if 8-oxoguanine passes through the copying process, the mispairing which can occur is to thymine, which in the next round of copying will be matched with adenine, the overall process hence leading to the replacement of guanine by adenine.²³ Metal complexes are useful in both the recognition of mismatches and in the study of guanine oxidation.^{24–27}

2.3 A complete assembly - Mg^{2+} ions and the Drew-Dickerson dodecamer

The natural ion found associated with double helical (duplex) DNA is Mg^{2+} , which has a radius of 0.67 Å in MgO but an estimated hydrated radius of 3.0 Å.²⁸ A variety of methods have been used to estimate that the magnesium concentration in cells is 17-20 mM.²⁹ The structure shown in Figure 7 is the classic 'Drew-Dickerson dodecamer', often referred to just as DDD. It was the first such self-complementary B-DNA structure to be studied, using a synthetic oligonucleotide of sequence d(CGCGAATTCGCG)₂. It has become the prototype DNA molecule since it was first determined in 1980 to a resolution of 2.3 Å.^{11,30} Since then it has proved a workhorse for an enormous variety of experiments, which aimed to understand the many chemical properties of the DNA molecule, because it can be thought of simply as a molecule of manageable size. It contains eight CG base-pairs four at each end of the sequence, with a central core of four AT base pairs. The thinking was that a CG base-pair is held together by three hydrogen bonds whereas an AT base-pair only has two, therefore the duplex would 'fray' at the ends without several CG base-pairs. In fact, when the structure was determined, it turned out that the real stability of the crystal came from the intermolecular interactions between the terminal two CG base-pairs which introduced additional hydrogen bonds. The central region of the structure was not affected and hence available for a range of studies including cation replacement. The so-called minor groove binding drugs could also be studied with this model system. These are organic cations which can curve to follow the groove contours and form additional hydrogen bonds within the groove.

The structure shown in Figure 7 is a redetermination completed in 1999 to 1.1 Å resolution, with one specific aim that of understanding the role of cations in stabilising the structure. At that time it had been suggested that Na^+ ions could be present in the narrowed minor groove of the central section, and a consequence of the AT base-pair geometry. Unlike a GC base-pair, which has an amino group protruding into the minor groove (labelled N2 in Figure 3), the AT base-pair presents a smooth surface in the minor groove. When this high resolution structure was first reported, it was the highest resolution B-DNA structure determined to date, and for the first time allowed the magnesium ions to be located. There were three ions in the asymmetric unit, one of which is labelled in Figure 7. One is in the major groove, and seems to be partly responsible for the relative ease with which this duplex can be crystallised, whereas the other two are associated with the phosphate oxygen atoms, both via direct coordination and through water bridges. Crucially, these authors found no evidence for the presence of sodium ions in the minor groove. High resolution data are essential to distinguish between sodium ions and water molecules, as both have 10 electrons associated with them and so scatter very similarly. With good data they can be distinguished both by the higher coordination number of a sodium ion and the somewhat shorter distances. A typical Na-O distance in a nucleic acid structure is 2.4 Å, whereas the typical water-water distance in these structures is closer to 2.8 Å. The authors add

'To further investigate the possibility of alkali metal ion coordination in the minor groove we crystallized the DDD duplex in the presence of either Rb^+ or Cs^+ . In the 1.2 Å structure of a Rb^+ -form DDD duplex a Rb^+ ion replaces the inner spine water with H bonds to O2 atoms of residues T8 and T20.'

They also report that replacing the Mg^{2+} with Ca^{2+} in the crystallisation mixture gives rise to a different, rhombohedral crystal form.

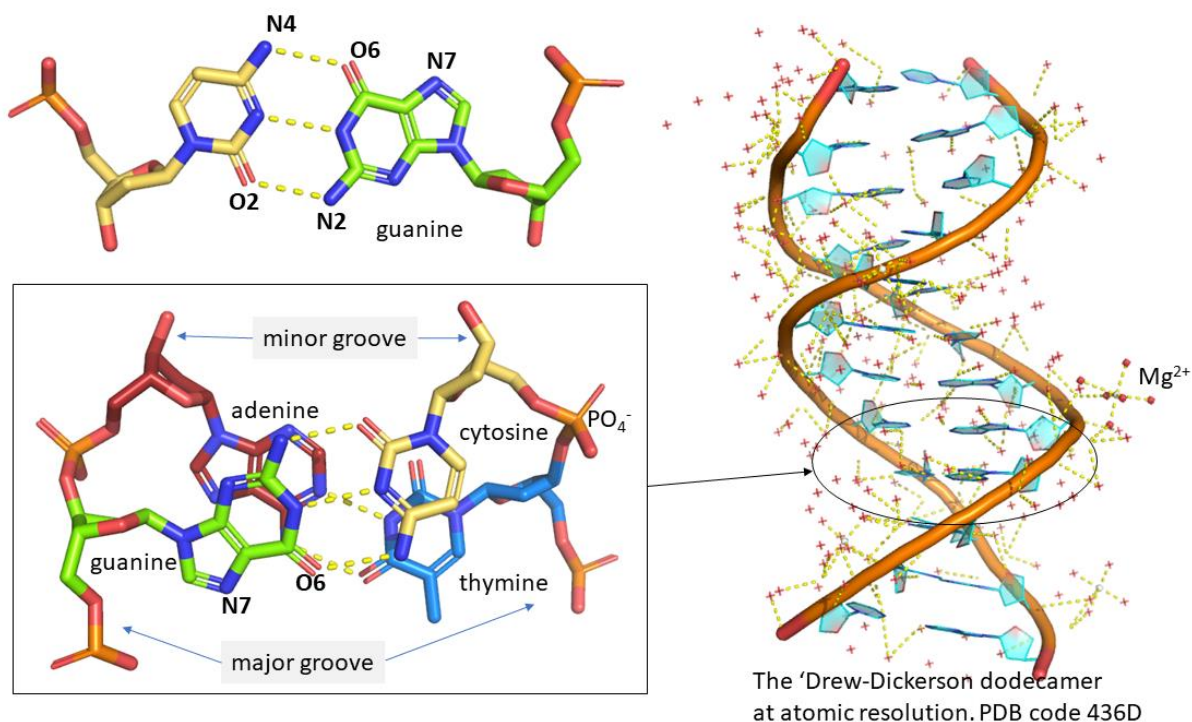


Figure 7 Aspects of the Drew-Dickerson dodecamer (DDD), from PDB code 436D, a 1.1 Å redetermination of the original d(CGCGAATTCGCG)₂ duplex structure. Crystallisation conditions - sitting drop vapor diffusion; a 20-mL droplet (1.2 mM DNA, 20 mM sodium cacodylate, pH 6.9, 25 mM Mg(OAc)₂, 3 mM spermine-4HCl) was equilibrated against a reservoir of 25 mL 40% MPD. Space group $P2_12_12_1$; cell dimensions $a = 24.64$ Å, $b = 39.63$ Å, $c = 65.53$ Å. The final refined model includes 143 water molecules. Details - (top left) a cytosine-guanine basepair from this model (C9-G16) and (below left) the same CG basepair showing the link to the adjacent AT basepair (T8-A17) highlighting the phosphate linkages, the right-handed turn of the helix and the creation of the major and minor grooves. Base colour codes as in Figure 4. The full duplex (right) with all water molecules and hydrogen bonding, with one of the coordinated $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ ions labelled.

3. Metal cations and the stabilisation of different DNA structures

The roles of Group I and Group II ions in the stabilisation of nucleic acids reflect their nature as highly mobile ions, consistent with the highly dynamic system of the living cell. The typical cell contains DNA stored in the nucleus, but there is also a constantly shifting population of RNA species, with the never-ending cycle of replication, transcription and translation. The binding of metal cations to RNA is also an important, but separate, topic.³¹ In replication, the DNA duplex is transiently unwound so that the DNA polymerase and associated enzymes can copy the two DNA strands before cell division, thus this process is correlated with the natural cell cycle. In transcription, the information contained in particular DNA sequences is copied and manipulated by a range of enzymes to generate a single stranded RNA molecule, but still as coded information. A complex set of interactions are required to regulate this process. In a normal healthy cell, the proteins required for health are then synthesised (translation of the code) on the ribosome. In cancerous cells, inappropriate processes occur, often due to DNA damage. Metal ions are involved at every step, but this is beyond the scope of this review.

The lack of useful uv/visible spectroscopic properties and nmr nuclei means that much of what we know about the detailed location of Group I and Group II cation binding comes from the highly static experiments of X-ray crystallography. The data on these interactions is stored in the Nucleic Acid Database (NDB) and is freely downloadable. The Protein Data Bank (PDB) in fact includes all macromolecular structures, including what is also available from the NDB, and both have many useful software resources which mean they are readily accessible to non-experts. The PDB allows the generation of electron density maps, such as the example in Figure 2a of a bound Mg^{2+} ion, (PDB code 4KWV); where the water coordination can be clearly seen. This ion is clearly not directly coordinated to DNA, but is indirectly linked to two phosphate residues through water bridges, and lies in the curve of the helix. Notably, nmr, which is a very powerful tool in nucleic acid structure determination, can define the path of the nucleic acid backbone but reveals nothing about the role played by cations in stabilising the structure. In contrast, in Figure 2b the platinum ion (which in that case dominates the scattering) is directly bound to two guanine bases at the N7 positions, constraining the coordination around platinum to adopt a local square planar geometry.⁵

3.1 Kinking of the DNA helix

The possibility of DNA kinking, and what it would do to the DNA structure, had been foreseen by Francis Crick himself.³² The kinking of the helix by DNA manipulating enzymes³³ is well known, and work of Lippard and his colleagues showed how the introduction of kinks by platination could widen the DNA minor groove. The compound $\text{cis-}[\text{PtCl}_2(\text{NH}_3)_2]$ and its analogues, first introduced into cancer therapeutics in the 1960s, have remained part of some standard treatment regimes ever since. The preparative methods for this important class of compound have recently been reviewed in detail.³⁴ Platinum resistance to chemotherapy, as well the inherent toxicity of platinum, are both major issues in the use of these compounds. Crystallography has shown, not only why DNA kinking is an effective therapy, but also, some of the mechanisms of platinum resistance³⁵, due to the existence of specialised DNA polymerases which can bypass the kinked DNA. Perhaps surprisingly, the obvious strategy of searching the Periodic Table for other suitable and hopefully less toxic elements has not really generated an obviously superior class of compound, although a range of elements such as palladium³⁶ and titanium³⁷ have been explored. The key to the success of the platinum compounds is the chemical inertness and strength of the interaction between the $[\text{Pt}(\text{NH}_3)_2]^{2+}$ moiety and the guanine N7 atoms of DNA. The crystallographic demonstration of this strong bond⁵ showed the power of the platinum square plane – the coordination around platinum was essentially that seen in small molecule structures, whereas the flexible DNA molecule was kinked at a 50° angle. As the N7 positions of the guanine bases are in the major groove, on kinking the direction of bend is such as to compress the major groove and open up the minor groove. So far, it has not been demonstrated that coordination by any other element can have this far-reaching effect on DNA structure. The initial accidental discovery, which at the time suggested that all sorts of other chemistry might be much more effective if rationally designed, has not really been borne out. It has been suggested that such kinking could be the mechanism by the potent anticancer activity induced using photolabile ruthenium compounds,³⁸ and it is an attractive idea, but the active moiety has to be sterically compatible with the restricted size of the curved major groove.

3.2 The Holliday junction

The Holliday junction is formed in DNA during processes which involve the swapping of DNA strands, leading to the exchange of genetic information. Some DNA repair mechanisms operate through such

a strand exchange, with a defective strand exchanged for an intact one. Genetic recombination is the exchange of genetic material between different organisms which leads to production of offspring with combinations of traits that differ from those found in either parent, and also goes through this junction intermediate. David Lilley showed that Mg^{2+} ions determined the folding equilibria of synthetic junctions in solution, and studied the process by FRET and by single molecule methods.³⁹ Junction formation clearly has specific charge neutralisation requirements, and in the closed form, which is Mg^{2+} stabilised, the phosphate negative charges are clearly brought close together. The open form has been crystallised with at least one of its associated enzymes⁴⁰ and the closed form was crystallised serendipitously as part of a study of DNA mismatches.⁴¹

While Mg^{2+} is the naturally occurring ion bound to DNA duplexes, the heavier Group II ions have all been found useful in the stabilisation of particular DNA conformations and in some cases of higher order DNA structures. The role of Ca^{2+} ions in the stabilisation of the Holliday junction has been studied in some detail as a function of DNA sequence.⁴² The DNA decamers of sequence D(CCnnnnNNNGG) are self-complementary, that is, they will form a duplex when the sequence nnn is the complement of NNN. In this work, the sequence ACC, as in the decamer d(CCGGTACCGG) turned out to be the only one which consistently formed the junction structure with Ca^{2+} ions.

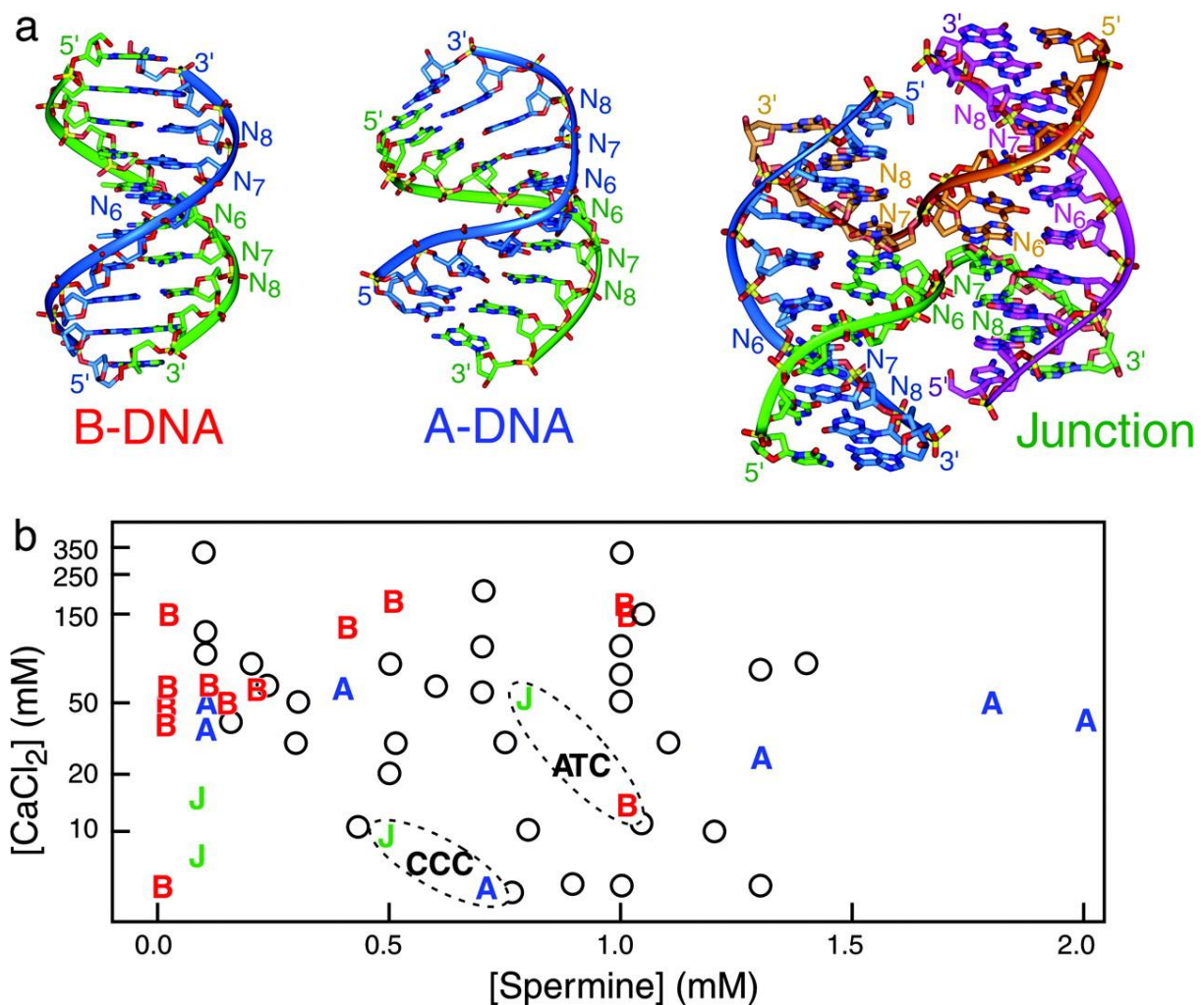


Figure 8 How sequence defines structure: A crystallographic map of DNA structure and conformation.

Structures from the crystallographic screen of the IR sequence d(CCnnnN6N7N8GG), where all 64 combinations of the N6 N7 N8 trinucleotide are sampled. (a) The conformations observed in the single crystal structures of this sequence include standard B-DNA (AGC structure shown), the altered A-DNA duplex (GGG structure shown), and the four-stranded Holliday junction (ACC structure shown). The positions of the N6 N7 N8 trinucleotide are labelled in each of the structures. (b) The CaCl_2 and spermine concentrations yielding crystals of B-DNA (B), A-DNA (A), and junctions (J) are compared, with the concentration of CaCl_2 plotted on a logarithmic scale. Open circles indicate conditions that yielded crystals, but where the conformation has not been determined. Only one label for each form is denoted in cases where crystallisation conditions overlap. The conditions for crystallisation of the amphimorphic sequences ATC and CCC are encompassed in ovals and labeled by the trinucleotide sequence.⁴² (downloaded from <https://doi.org/10.1073/pnas.0409455102>)

The sequence GCC, as in d(CCGGCGCCGG) could crystallise as B-DNA or junction DNA depending on the Ca^{2+} concentration and most other sequences from this set were simply B-DNA or in some cases A-DNA. Closer examination of these structures showed that in some cases the Ca^{2+} ions could be readily located in the minor groove of the DNA, in contrast to the typical behaviour of Mg^{2+} . There are a few examples of Sr^{2+} binding,⁴³ typically also in the minor grooves, whereas Ba^{2+} binds to the same sequence in a disordered way, with several fractionally occupied sites (site occupation being a parameter which can be readily determined for such a strong X-ray scatterer).⁴⁴ In all these cases, the characteristic coordination number of the ion can be seen in the water coordination, and it is typically exactly as would be expected from inorganic structures.

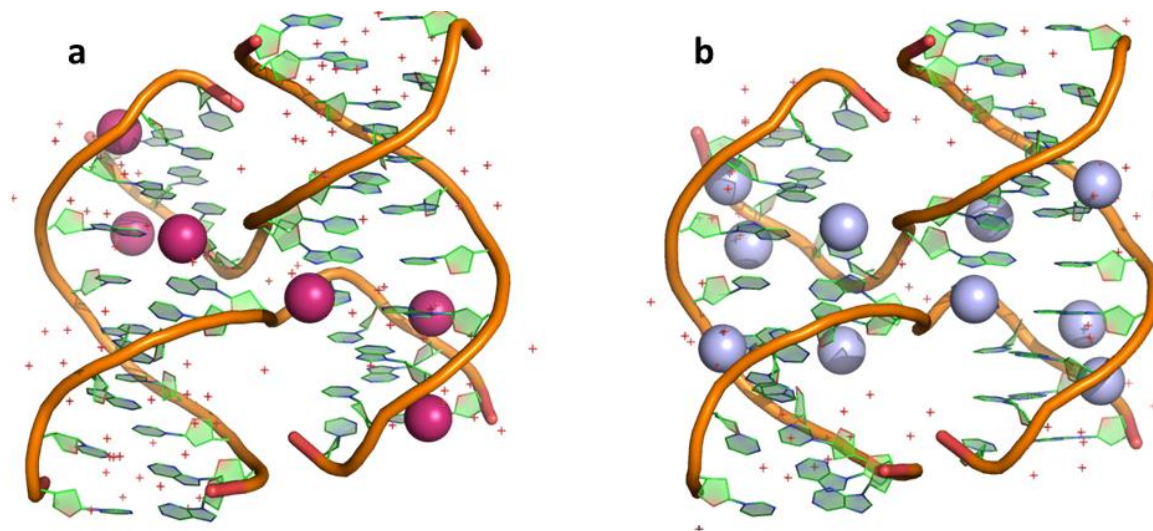


Figure 9 Strontium and barium ions bound to the Holliday junction forming sequence d(TCGGCGCCGA). a) Sr^{2+} crystallised with this sequence (PDB : 3GOO)⁴⁵ b) Ba^{2+} with the same sequence (PDB: 3GOJ).⁴⁴

3.3 The G-quadruplex

The G-quadruplex is a specialised DNA structure first crystallised as a tetramer⁴⁶ and recently in forms with greater biological relevance^{47–49}. At first this structure was thought to be a curiosity but more recently it has turned out to be important not only in telomeres (the single stranded ends of chromosomes) but also in the promoter regions of some important genes implicated in cancer⁵⁰, with perhaps about 300,000 potential G-quadruplex forming sequences in the human genome. The sequence requirements are approximately ---GGG---GGG---GGG---GGG---. Notably, monovalent cations play a key role in stabilising this structure. In human cells, the K^+ concentration is about 100 mM, but it is only 4 mM in blood plasma. The Na^+ concentration in cells is about 10 mM, but 100–200 mM in blood plasma. Therefore, it is striking that the G-quadruplex structure is under normal circumstances specific for K^+ , with an 8-coordinate environment formed by the eight guanine 6-CO groups in the centre of this structure (Figure 10) in which each K-O distance is 2.8 Å. Although nmr is an essential tool for studying G-quadruplex structures in solution, the K^+ ions are invisible to this technique. In crystal structures the K^+ are seen to be aligned down the centre of the assembly, generating almost flat parallel layers of G-quartet bases (Figures 4 and 10a). These structures are assumed to arise in nature from a single DNA strand as it is separated from its complementary strand during, e.g. replication, and formed dynamically in cells in 100 mM K^+ . The creation of G-quadruplex specific antibodies has demonstrated the presence of these assemblies in human cells, making their study of great current interest.^{9,51} The human telomeric sequence --- (GGGTTA)*n* ---, where there are several hundred repeats, has been shown by nmr to give several interconverting structures, but specifically, the Na^+ and the K^+ forms are distinct.

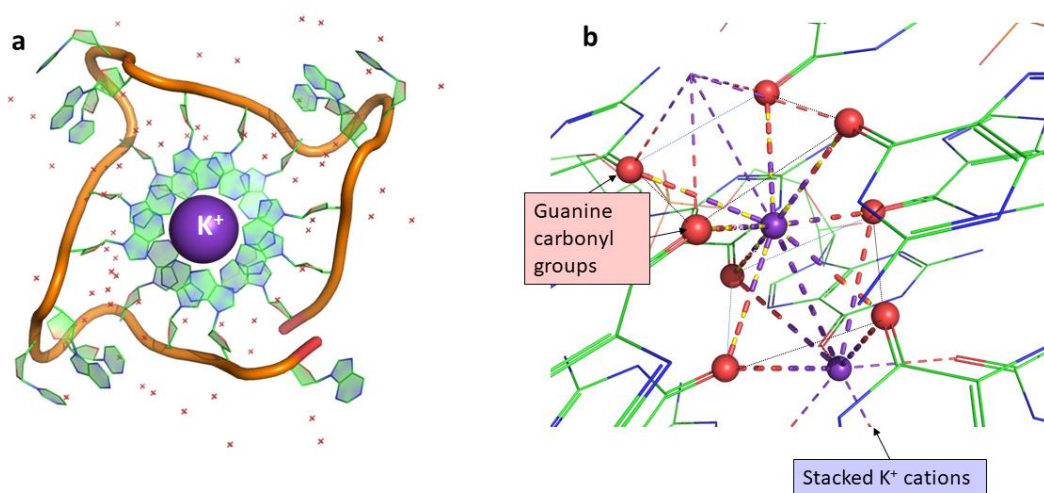


Figure 10 Crystal structure of the human telomeric G-quadruplex with potassium ions. a) the single strand d(AGGGTTAGGGTTAGGGTTAGGG) crystallised with K^+ ions. The structure is a parallel stranded quadruplex in which each of the four runs of ---GGG--- sequences run in the same direction. b) the potassium coordination (from PDB file 1KF1). The K^+ ion shown is 8-coordinate, with successive G-quartet layers offset by right-handed rotation. The faint grey lines show the individual G-quartet square arrangement of carbonyl oxygen atoms. These are the atom type labelled O6 in Figure 4c. The G-quartet is held together by Hoogsteen-type hydrogen bonds.^{15,17}

The original G-quadruplex structure was reported in 1994⁴⁶ and was built from four strands of the short sequence d(TGGGGT). The structure was thought to be an interesting crystallographic curiosity, with its column of K^+ ions running through the structure and linking two such assemblies. Figure 11a

shows enough of the crystal packing to indicate why this model may have appeared as elegant but not very useful. The stacking interaction clearly dominates the packing and the units are packed in long parallel columns in a fourfold symmetric arrangement. Yet, twenty years later, as shown in Figures 11b and 11c, the same central feature is seen in the crystal structure of human native ckit proto-oncogene promoter quadruplex DNA (PDB code 4WO2). This sequence, d(AGGGAGGGCGCTGGGAGGAGGG), shows the typical complexity of these important 'switch' regions in the human genome, and this structure was recently compared to that formed by another important gene promoter region, that of the c-myc oncogene.⁵² Both are parallel-type quadruplexes and both regulate genes which are dysregulated in many cancers. Therefore the G-quadruplex has evolved from being an academic curiosity into an important therapeutic target.⁵³ The structural problem is that these more complex but important sequences can fold in several ways, all held together by a central K⁺ ion assembly but with more than one possible way the G-quartet can stack, and the strength of the stacking interaction is that, as with B-DNA, it is the prime force holding the assembly together. Lower potassium concentrations can be found in cancer cells⁵⁴ also contributing to the dysregulation of genes like c-myc, since the existence of the G-quadruplex structure acts as an inhibitor of the gene expression process. But unlike B-DNA, where the complementarity of the basepairs normally generates a unique double stranded assembly, and a specificity which can be exploited therapeutically, the G-quadruplex assembly is inherently polymorphic, and indeed this shape-shifting property is probably related to its normal biological role. Hence in this area, both X-ray structures showing the exact cation locations, and nmr structures which only define the nucleic acid component have important and complementary roles to play.

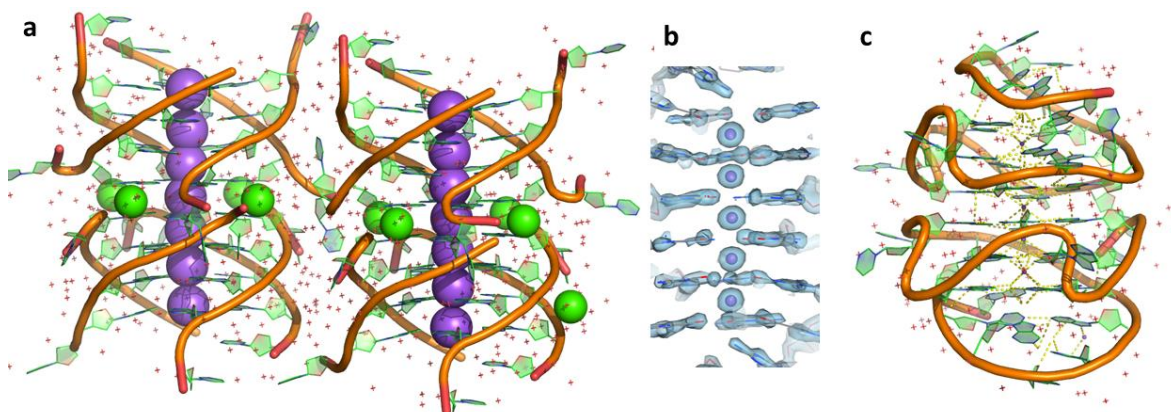


Figure 11 a) The first G-quadruplex structure crystal packing. This assembly is built from the d(TGGGGT) oligonucleotide, and the figure shows 16 such strands, with K⁺ ions in purple and Ca²⁺ ions in green. b) a similar stack of K⁺ cations in structure with PDB code 4WO2, the crystal structure of human native c-kit proto-oncogene promoter quadruplex DNA c) the unimolecular parallel stranded G-quadruplex formed by the sequence used in b).

3.4 DNA and luminescent probes of structure

Flat positively charged aromatic cations have long been known to bind to DNA by intercalation.⁶ The intercalative binding mode to natural duplex DNA is seen when such a cation is inserted into a certain position in the DNA stack, unwinding the twist of the helix and lengthening the helix by the equivalent of an extra basepair. The resulting overall conformation is that of a B-DNA, even if the natural local sequence is A or Z. This binding mode is relatively easy to detect, because the biophysical properties of the DNA molecule are changed; the temperature at which the strands separate (the melting temperature) is typically increased, the viscosity is increased and using circular DNA the unwinding angle can be estimated. Some of the classic cancer therapeutics such as daunomycin bind to DNA in this way.⁵⁵ More recently, metal complexes which can act as structural probes have been designed, making use of both the characteristic properties of the metal and the design pointers from ligands such as daunomycin. There is an extensive literature on this subject, and for many in the field, the aim is to design molecules which will luminesce when bound to specific DNA structures in the living cell.⁵⁶ For this purpose, the metal itself should be nontoxic, and the complexes should have suitable luminescence properties as well as solubility behaviour which allows transport to the cell nucleus. For photodynamic therapy applications, in which cancer is treated by irradiation of the tumour already exposed to a suitable sensitiser, there is a restricted therapeutic wavelength range corresponding to the transparent window for skin (650-950 nm).^{57,58}

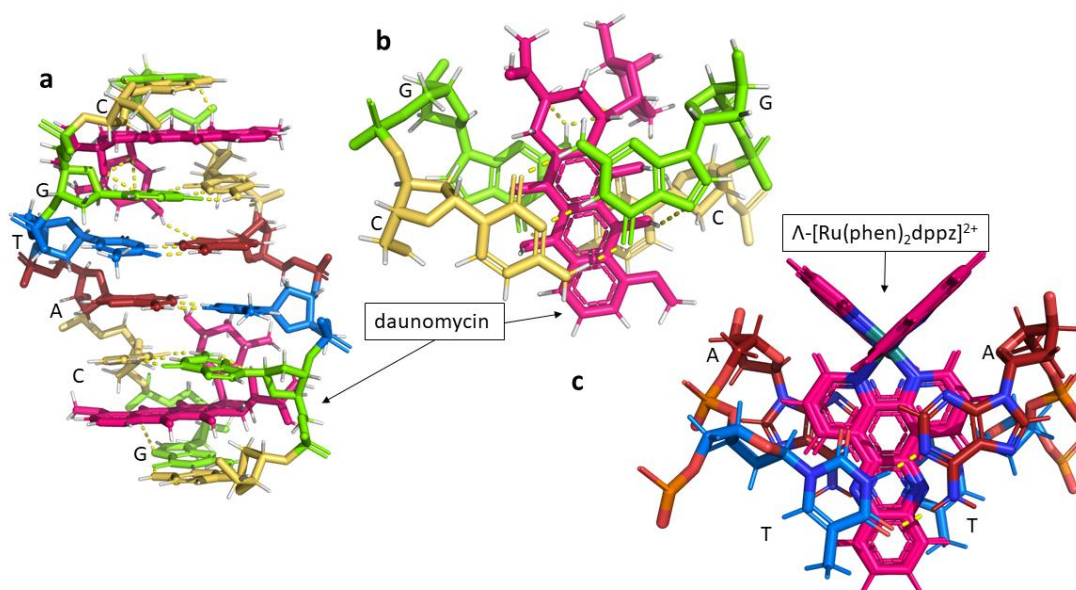


Figure 12 Perpendicular intercalation. a) the intercalation of two daunomycin cations into the DNA duplex $d(\text{GCTACG})_2$. The bases of one strand are labelled. Cytosine C - yellow; guanine G - green; adenine A - red; thymine T - blue. The daunomycin ligand is shown in pink. b) the intercalation cavity of daunomycin showing the perpendicular intercalation from the minor groove side of the double helix. The daunomycin chromophore is approximately perpendicular to the long axes of the CG basepairs. (in contrast, an intercalated acridine would have an approximately parallel orientation). c) perpendicular intercalation of the 'light-switch' complex Λ -[Ru(phen)₂dppz]²⁺ between two TA basepairs (complex shown in pink).

Several lanthanide cations, such as Sm^{3+} , Eu^{3+} and Tb^{3+} have luminescence properties which have been exploited⁵⁹ for DNA recognition, though no structural studies are currently available, and the binding modes are beyond the scope of this review. The ruthenium polypyridyl complexes typified by the 'light-switch' complex $[\text{Ru}(\text{bpy})_2(\text{dppz})]^{2+}$ are better understood (dppz=dipyridophenazine). Ru(II) has useful luminescence properties even in the commercially available $[\text{Ru}(\text{bpy})_3]^{2+}$, due to the $^3\text{MLCT}$ transition on excitation at 400 nm from the Ru(II) $4d^6$ ground state to a Ru(III) excited state in which the excited electron is on the bpy ligands. This is an extensively studied and well understood photophysical system, originally studied for its possible use in solar energy. Luminescence changes on binding to natural DNA (such as the commercially available calf thymus DNA) can be observed even with this complex. Since 1990, when the idea of extending one of the bpy ligands was introduced by Chambrion and coworkers⁶⁰ the dppz and related polypyridyl complexes have been studied for their potential uses as cancer therapeutics and as sensitive luminescent probes for the presence of particular DNA structural features. The modularity of the octahedral metal complexes, their kinetic inertness and their inherent lack of toxicity have led to many variations on the original theme, with increased specificity for important structural targets, improved luminescence properties, specific DNA cleavage properties and the like. Structural studies have highlighted the enantiospecificity of the DNA binding and the importance, for fundamental spectroscopic binding studies, of working with enantiomerically pure material.⁶¹ The success of these studies has naturally resulted in an exploration of the surrounding areas of the Periodic Table, including rhodium⁶², rhenium^{63,64} and iridium⁶⁵ complexes, and has created a very active, multidisciplinary research field which is now resulting in patentable inventions and useful therapeutics.⁶⁶ An osmium analogue⁶⁷ can be used as a high-resolution TEM agent, particularly for the imaging of intranuclear detail.

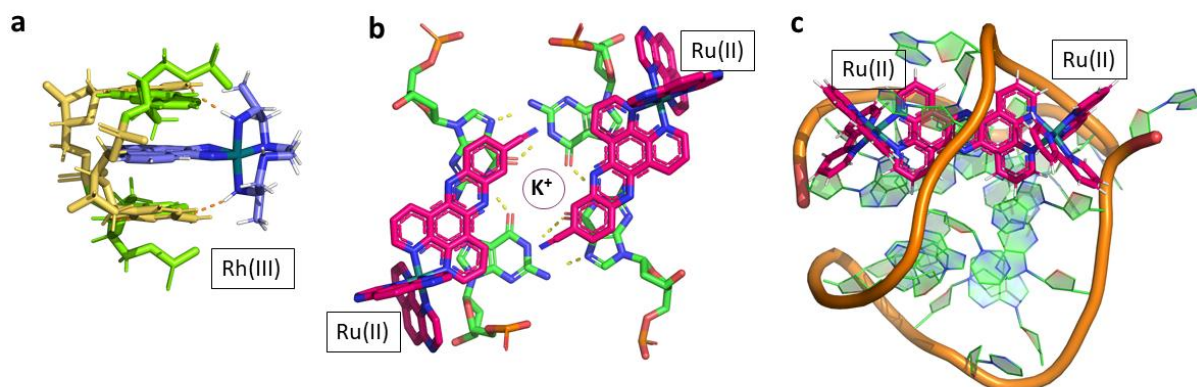


Figure 13 Rhodium and ruthenium complexes. a) a rhodium (III) complex bound to duplex DNA from the major groove at a CG/CG step (PDB code: 454d); b) a ruthenium (II) complex bound to G-quadruplex DNA, giving a mixture of syn and anti base conformations. Compare Figure 4b, which shows all anti conformations and creates four symmetrical grooves; in this example the alternating syn and anti conformations create alternating wide and narrow grooves.(PDB code:5LS8); c) a diruthenium complex threaded through a diagonal loop in the nmr structure of G-quadruplex human telomeric DNA. (PDB code: 2MCO).

The first successful crystallisation of a ruthenium polypyridyl complex with a DNA duplex is described below. It ended a long-running controversy about the binding modes of these complexes by showing that intercalation is from the minor groove of the helix, placing the dppz ligand between the basepairs

in an orientation determined by a secondary interaction between one of the ancillary bpy or phen ligands and one of the sugars. The octahedral geometry and the planar aromatic nature of the ligands creates a rigid three-dimensional metal complex cation which can bind due to the flexibility of the DNA backbone and is inherently enantiospecific. The delta enantiomer, for example, explicitly recognises some mismatched sequences such as the A-A mismatch²⁷, whereas the lambda enantiomer has a high affinity for the weakest step in the DNA double helix, the TA/TA step, and in this case only has a perfectly symmetrical and preferred binding mode.^{68,69} To date there is just one example where both enantiomers are intercalated in a single crystal structure.⁷⁰ This work showed that the two enantiomers of the complex $[\text{Ru}(\text{phen})_2\text{dppz}]^{2+}$ intercalated at different angles and caused small distortions of the phosphate backbone. Later studies using several substituted dppz ligands showed that there was a more general feature of the intercalation, with the cavity always asymmetric. One side of the cavity has a modified backbone conformation, which is significant because the exposure of the dppz ligand to water is the main factor determining the luminescence behaviour. The extended side of the cavity reveals more of the chromophore.⁷¹

The diruthenium complex shown in Figure 13c has been notably successful as a luminescence probe.⁷² It is able to penetrate into the cell nucleus and directly image the DNA without requiring prior membrane permeabilisation. Because there are three possible enantiomers, lambda/lambda, delta/delta and *meso*, specific recognition by the complex must be limited. An nmr study of the separated lambda/lambda and delta/delta enantiomers showed specific binding modes to an antiparallel G-quadruplex structure, with the lambda/lambda enantiomer threading through a diagonal loop at one end of the G-quadruplex stack, but the delta/delta showing end-stacking to the other end of the G-quadruplex stack. Several other metal complexes, the salphens containing Ni(II) and Co(II)⁷³, platinum,⁷⁴ and gold⁷⁵ were recently reviewed as part of a survey of metal complex binding to G-quadruplexes⁷⁶. The ones characterised by X-ray crystallography all show end-stacking to parallel G-quadruplexes by essentially planar chromophores. For more specific G-quadruplex recognition, which would have important therapeutic implications⁵³, the use of an octahedral metal complex, with its enantiomeric possibilities, is a promising strategy. Our recent crystallographic study of the binding of the complex $[\text{Ru}(\text{TAP})_2(11\text{-CN-dppz})]^{2+}$ to the G-quadruplex formed from four strands of the d(TAGGGTTA) sequence showed the specific interaction of the lambda enantiomer with the G-quadruplex stack.⁷⁷ The specificity required the guanine to adopt the *syn* - conformation, which results in an overall antiparallel arrangement of the four DNA strands, alternately up and down around the square of the quadruplex, with two central K^+ ions, as expected. This structure also provided an explanation for the luminescence properties of the 11-Br analogue $[\text{Ru}(\text{bpy})_2(11\text{-Br-dppz})]^{2+}$, which even as the racemic mixture gave selective luminescence with G-quadruplex and i-motif structures in comparison to duplex DNA.⁷⁸ The luminescence must arise principally from the lambda enantiomer, as, using the crystal structure as a model, the substituted dppz ligand would be completely encased by the DNA, lengthening the lifetime of the luminescent excited state. In the crystal structure, which was crystallised from a racemic mixture, the delta enantiomer was also present in the structure, but only acting as the stacking 'glue' between the ligated quadruplex assembly.

4. Working with DNA crystals

4.1 Metal ion diversity in the crystallisation of nucleic acids

The crystallisation of nucleic acids for structure determination is still as much an art as a science. The great importance of protein structure determination has brought automation and high throughput

methodology into the macromolecular structure laboratory, but to date has been less useful in the more specialised area of nucleic acid structure determinations. The ribosome and structures related to it are among the impressive achievements of nucleic acid crystallography, resulting in Nobel prizes for Ramakrishnan, Yonath and Steitz^{79,80}, and the accounts of this work make for awe-inspiring reading. On a somewhat more day-to-day level, the crystallisation of a typical small synthetic oligonucleotide requires the use of a crystallisation screen, several of which are now commercially available. The classic sparse matrix screen⁸¹ has been expanded more recently to a 96 well screen⁸². This screen combines sparse matrix, incomplete factorial approaches and modified orthogonal arrays. These authors have analysed 1450 deposited sets of crystallisation conditions, to reveal the most commonly successful cations. The data show that Mg^{2+} is the most important cation, found in about 500 sets of conditions, with Co^{2+} , Ca^{2+} , Ba^{2+} , Sr^{2+} and Mn^{2+} also important. Of monovalent cations, both Na^+ and K^+ have been widely used, with Li^+ found useful in 10% of cases and with other monovalent cations relatively insignificant. These authors have made a detailed study of crystallisation conditions, including photographs of actual crystals and carried out test crystallisations of a range of sequences, concluding, for example, that successful G-quadruplex crystallisation requires relatively high concentrations of monovalent cations.

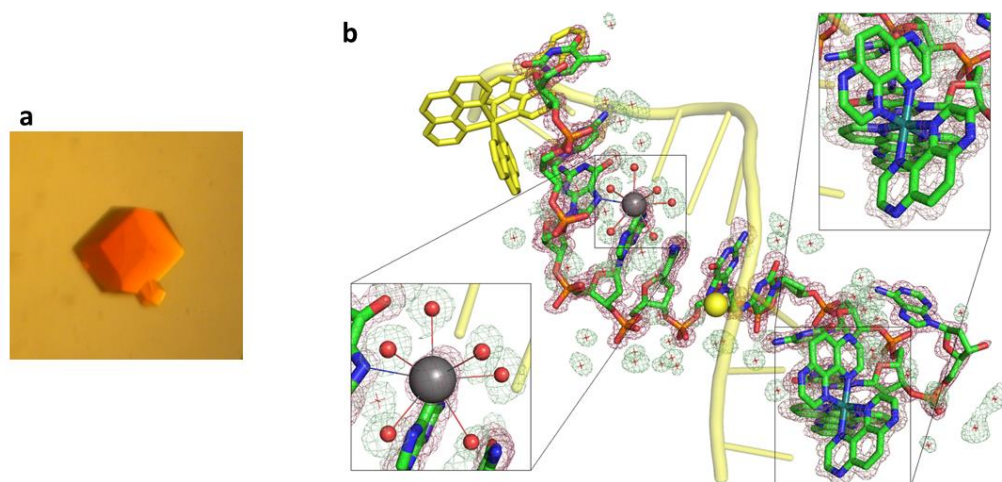


Figure 14 The sequence *d*(TCGGCGCCGA) crystallised with the ruthenium polypyridyl complex Λ -[Ru(TAP)2(dppz)]²⁺ (TAP=tetraazaphenanthrene; dppz=dipyridophenazine). a) A typical crystal; b) The resulting duplex structure. The asymmetric unit is a single strand, with the complex bound at the terminal step of the DNA duplex. The insets show the Ba^{2+} environment (average Ba-O distance 2.8 Å) which shows 8-coordination by 6 water molecules, the N7 of guanine G3 and the O6 of guanine G4.

The Figure shows the first successful crystallisation of a ruthenium polypyridyl complex bound to an oligonucleotide (crystal size about 0.2 mm across) and as suggested by the morphology, the crystal class turned out to be tetragonal (space group $P4_32_12$).⁴⁴ The colour of the crystal is due to the Ru(II) ³MLCT band with λ_{max} around 460 nm. The crystallisation conditions contained 80 mM NaCl and 20mM BaCl_2 . Notably, crystals cannot be obtained at all in the absence of Ba^{2+} ions. Although the

crystallisation mixture contained a mixture of enantiomers, crystallisation was completely specific for the lambda enantiomer of the complex. The DNA sequence here is d(TCGGCGCCGA) and the metal complex has the formula Λ -[Ru(TAP)₂dppz]²⁺, where TAP=tetraazaphenanthrene and dppz=dipyridophenazine. The structure shows why there is a specific requirement for Ba²⁺ ions, as the Ba²⁺ is directly coordinated to two guanine bases in the DNA major groove, the N7 and the O6 positions. The suspicion that the Ba²⁺ was in some way integral to the structure was eventually shown to be false, when the cobalt hexammine cation [Co(NH₃)₆]³⁺ was shown to give crystals in this system but without direct coordination to the DNA bases.²⁶ The [Co(NH₃)₆]³⁺ cation is particularly useful in nucleic acid crystallisations and is included in the standard screens. Famously a kinetically inert cation, it has approximately the same dimensions as the Mg²⁺ cation and therefore can – albeit with a higher charge – bind in what would plausibly be Mg²⁺ sites. The examples shown here are chosen from some of the highest resolution structure determinations in the PDB. Even so, hydrogen atoms are not normally located, though their positions can be determined by neutron diffraction in favourable cases.^{83,84} It is often possible to infer hydrogen locations from water structure, since the phosphate backbone does not protonate and is always the lone pair donor. In the structure shown in Figure 15a, the authors report that the decamer was bent by the presence of Ca²⁺ ions in addition to Mg²⁺ ions.⁸⁵ Co²⁺ seems to turn up unexpectedly in crystal structures, coordinated to the guanine N7 position.⁸⁶ In some crystallisation screens it is present as the hexahydrate, but it may also be that such Co²⁺ ions have been inadvertently liberated from the supposedly completely inert [Co(NH₃)₆]³⁺.⁸⁷ Notably, the binding preferences are different, with [Co(NH₃)₆]³⁺ emulating a Mg(H₂O)₆²⁺ ion ('hard' ligand association with phosphate backbone) but Co²⁺ taking the 'softer' route and targeting the guanine N7 position.

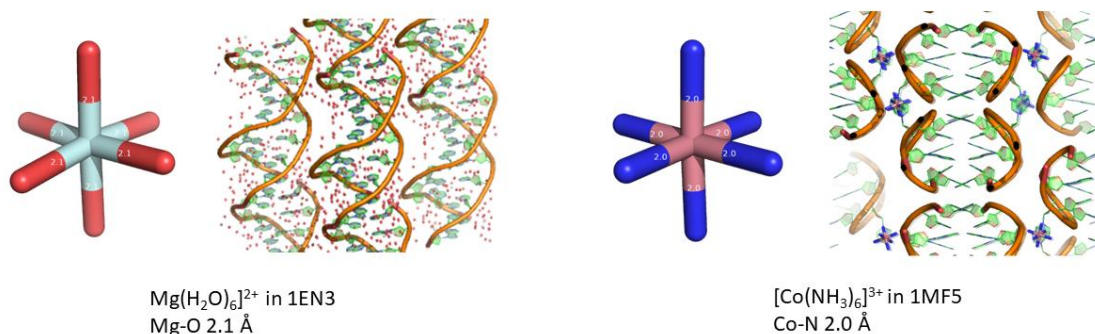


Figure 15 The close isomorphism (for crystallographic purposes) between $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ and $[\text{Co}(\text{NH}_3)_6]^{3+}$. a) the $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ cation as seen in the 0.985 Å resolution structure of the decamer duplex d(CCAACGTTGG)₂ (PDB: 1EN3) b) the $[\text{Co}(\text{NH}_3)_6]^{3+}$ cation seen in the loop structure formed by the sequence d(GCATCGT) (PDB: 1MF5).

The structure containing the cobalt hexammine was an early example of the formation of a quadruplex-like system by the back-to-back hydrogen bonding of two standard G-C base pairs.⁸⁸ The sequence, d(GCATCGT), folds into a loop. The loop is stabilized by association with a cation, by intra-strand hydrogen bonds between guanine and cytosine that are distinct from the normal Watson-Crick hydrogen bonds, and by van der Waals interactions. Two loops associate through the formation of

four GC pairs that exhibit pronounced base-stacking interactions, and the formation of a symmetric AA base pair further stabilizes loop dimerisation. Stacking of the AA pair on a symmetry-related AA pairing assists the formation of a four-stranded assembly. A TT pairing is also observed between symmetry-related loops, giving the packing shown in Figure 15b. There are relatively few water molecules compared with that typically seen surrounding B-DNA structures in the crystal, and later work has shown that this is a characteristic difference associated with these more condensed, higher order assemblies. In the original work the cation used for the crystallisation was Mg^{2+} . The subsequent study carried out in Reading showed that the use of cobalt hexammine did indeed give an almost isomorphous structure, but at 1.1Å resolution (almost atomic) and with the cations in the equivalent positions, thereby providing a validation for the use of cobalt hexammine in crystallisations, discussed in more detail in this paper. This little sequence has also acted as a useful testbed for the behaviour of different divalent ions with DNA. With Ca^{2+} ⁸⁹, Ba^{2+} ⁹⁰ and V^{2+} ⁹¹ the structure was essentially that shown in Fig 16d, but with Co^{2+} ⁹² and Ni^{2+} ⁹³ a completely different conformation is adopted, with direct coordination of the Co^{2+} and Ni^{2+} cations to the guanine N7 positions ('softer' behaviour).

The work of Sugimoto on molecular crowding concluded that important higher order structures such as the G-quadruplexes described above are more stable in the real cellular environment than in the dilute solution studied *in vitro*.⁹⁴ Perhaps this will also turn out to be true for this type of quadruplex, which seemed at the time to be yet another interesting academic curiosity.

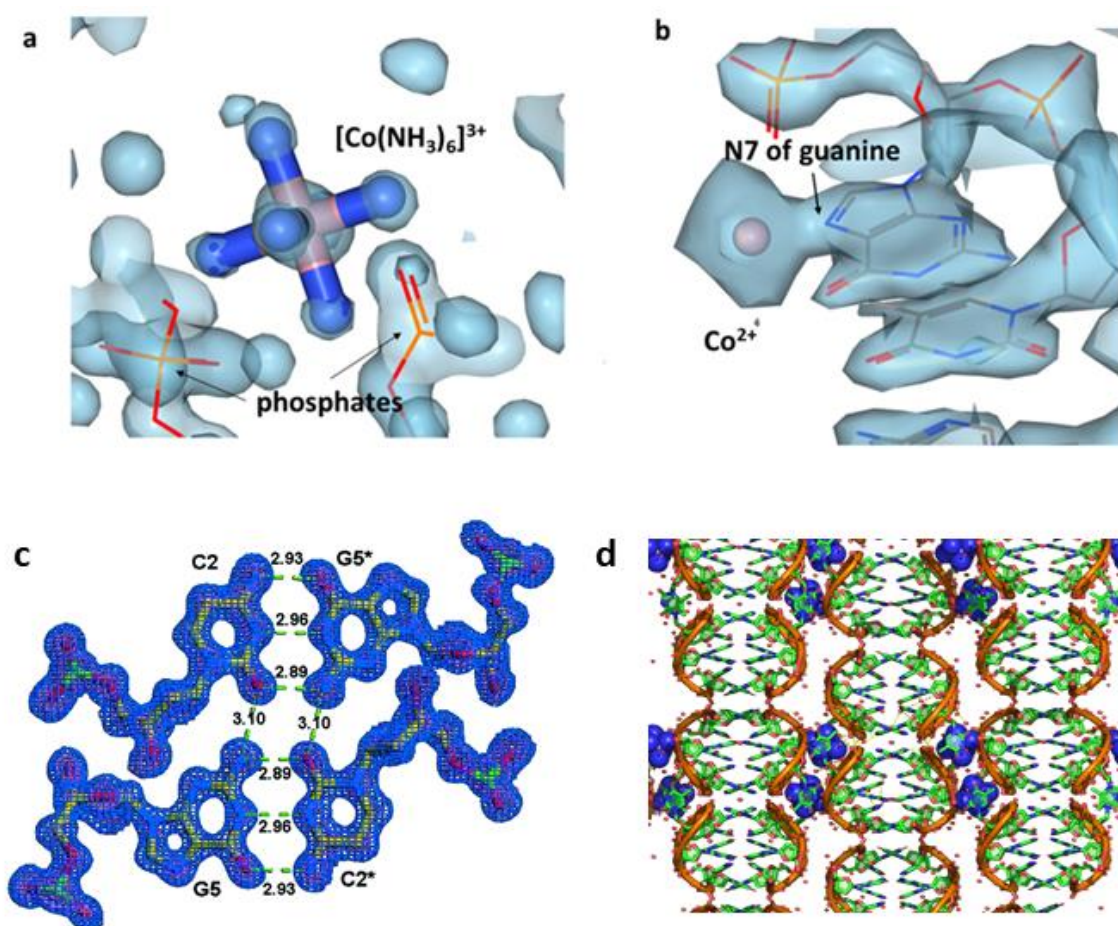


Figure 16 Cobalt and DNA. a) Association of $\text{Co}(\text{NH}_3)_6^{3+}$ with the phosphate backbone b) attachment of Co^{2+} ions directly to the guanine N7 position c) Quadruplex structure in PDB 1MFS (downloaded from our original work in NAR, 2003, 31, 844-8, Figure 3) and d) $\text{Co}(\text{NH}_3)_6^{3+}$ in the crystal packing of this structure.

The more typical crystallisation scenario is that the crystal does not diffract to better than 2 Å, which is the acceptable standard used by industrial protein crystallographers, and at that resolution it is most unlikely that light atom cations will be detectable. Publishable structural data can still be obtained at 3 Å resolution, but the map quality will normally only allow the fitting of a relatively straightforward model, and certainly no cation detection. So our knowledge is derived from these relatively few structural studies refined with almost small molecule precision.

4.2 The Periodic Table of X-ray Absorption edges in nucleic acid structure determination

The determination of X-ray structures requires a method for assigning the phases to the measured diffraction intensity data. Laboratory X-ray sources for macromolecular crystallography use $\text{Cu-K}\alpha$ radiation and are excellent for crystal screening and characterisation, as well as routine measurements of diffraction intensities. What such sources cannot do is provide phase information directly. This may not matter if the dataset being measured is related to known structures, because in that case the structure can be phased (=solved) using the molecular replacement method. This could be for example if a structure is available and the difference is only that a different metal ion has been used, that is to say, the structure is isomorphous with a known one. But of course, the interesting structures are those for which the crystal packing is new, as well as the chemical composition. In that case, synchrotron data collection opens new doors. This is because at a synchrotron the wavelength is tunable, with a standard wavelength range of 0.5-2.0 Å as a minimum. The typical synchrotron is optimised for wavelengths around 1 Å, and these days the technology is such that the changing of the wavelength is an automated procedure. A better known advantage of synchrotron data is the enormous improvement in diffraction intensities, enabling the use of very small crystals (down to 10 µm in some cases) so increasing the range of crystals which are measurable on a routine basis.

The tunability means that structure determination can make use of an inherent property of an element, its X-ray absorption edges. The Periodic Table of absorption edges (Figure 17) shows that many elements of interest to nucleic acid enthusiasts have edges within the synchrotron accessible window.

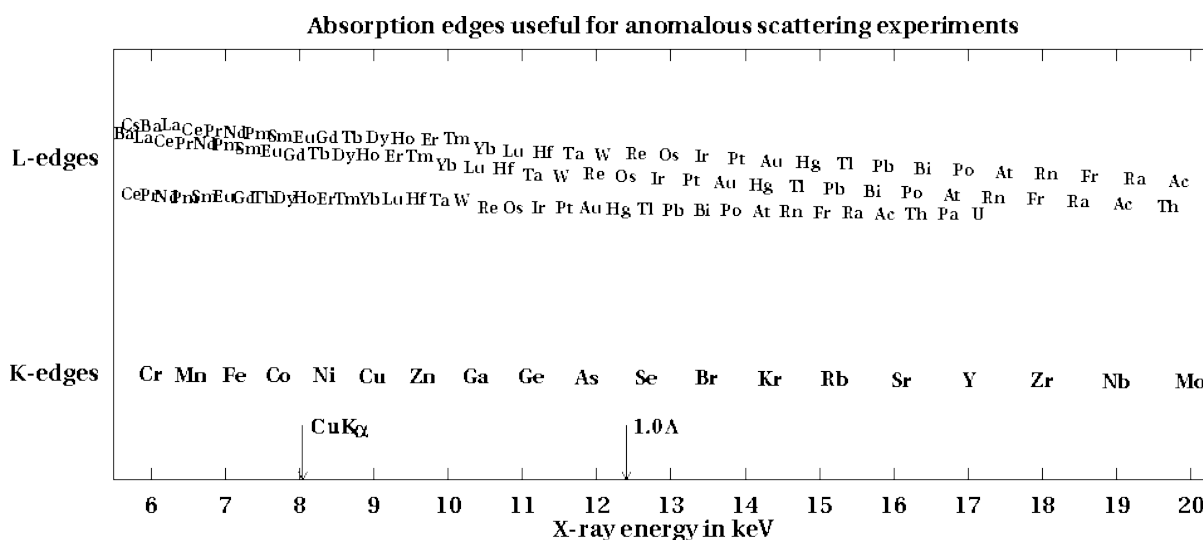


Figure 17 X-ray absorption edges of elements which may be useful in macromolecular crystallography. A tunable beamline for macromolecular crystallography typically allows access to the 0.5-2.0 Å wavelength range, with the maximum intensity tuned to be around the 1 Å mark. The Se and Br K edges have near-optimal intensities; the Ru edge at 22.4 KeV is too high in energy for this standard chart, and the intensity is in fact much lower. (from http://skuld.bmsc.washington.edu/scatter/AS_chart.html)

The phasing experiment near an X-ray absorption edge relies on the phenomenon of anomalous differences. In the normal diffraction experiment, the diffraction pattern is centrosymmetric (it is said to obey Friedel's Law) even though nucleic crystals will not contain a centre of symmetry and are inherently chiral. Just above an absorption edge, this law is not obeyed, so that intensities are different by maybe 2% or so. The small differences are element-specific and so lead directly to the location of the element in question. This also means that, if there can be a question as to which ion has actually bound, it can be determined from the presence or absence of an anomalous diffraction signal at that wavelength and, conversely, unexpected ions can be identified in this way.

As with other aspects of nucleic acid structural work, the enormous importance of protein structure has led to the availability of many sophisticated tools.⁹⁵ The availability of tunable beamlines for macromolecular crystallographers is one such spinoff. Protein crystallographers developed the use of the selenium absorption edge at 0.92 Å, based on the idea that sulphur in proteins could be isomorphously replaced by selenium,⁹⁶ which has this extremely useful K absorption edge. Conveniently, bromine can be introduced into nucleic acids and used similarly (Figure 18) shows the position of the K edge at 13.4737 KeV, or 0.9292 Å, using the same structure solution methods which first derive the position of the anomalous scatterer, and use this as a basis for the phase calculation to derive an electron density map.⁹⁷⁻⁹⁹

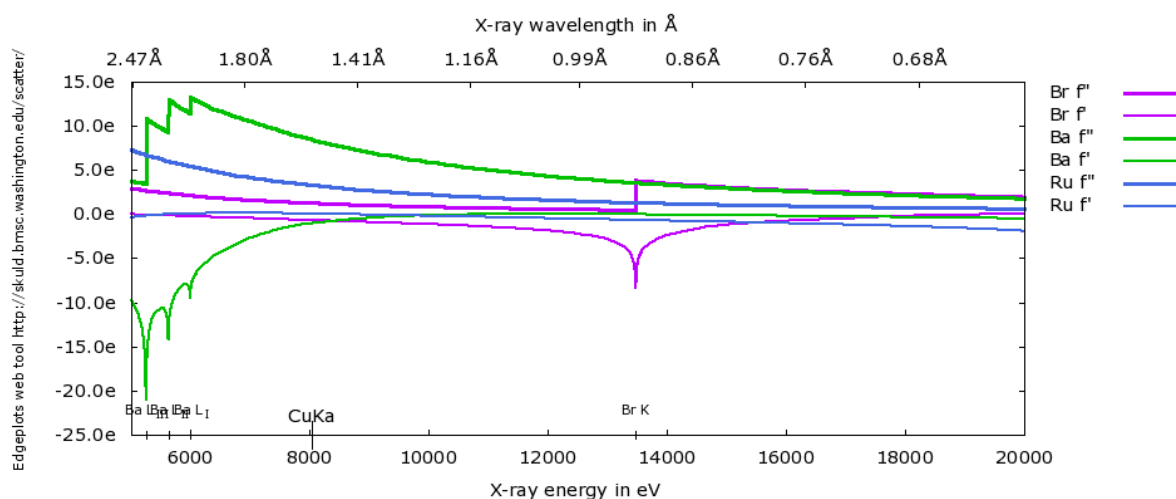


Figure 18 Typical plot of anomalous scattering coefficients for selected elements. From http://skuld.bmsc.washington.edu/scatter/AS_form.html

In many cases, therefore, it may not be necessary to introduce a modified (e.g. brominated) nucleic acid in order to solve a new X-ray structure. As Figure 18 shows, there can be significant anomalous scattering effects even at the Cu $K\alpha$ wavelength of 1.54 Å used by in-house equipment, such as that shown by Ba. The anomalous scattering of Ba was enough to give an in-house solution using sealed-tube Cu- $K\alpha$ radiation to generate the structure shown in Figure 14.⁴⁴ Later, a synchrotron data collection at 1.8 Å gave very high quality phase information and revealed the presence of both Ba and Ru sites from a single anomalous diffraction (SAD) map, phased with the SHELXC,D,E pipeline of crystallographic software.¹⁰⁰ At this long wavelength, the anomalous scattering arises from the ruthenium L-I edge at 3.2240 keV or 3.8457 Å. More recent advances in beamline technology mean that ruthenium-containing oligonucleotide structures can be routinely phased using the same methodology, but using the high energy Ru K edge at 22.1172 KeV or 0.5606 Å.^{71,77} The use of the high energy radiation is to be preferred because it minimises crystal damage. At the longer wavelengths, absorption of X-rays by the heavier elements becomes a problem and leads to sample degradation by the very intense synchrotron radiation. The powerful data collection and processing software now routinely available can diagnose the problem, but it is obviously better to avoid damage by careful choice of wavelength.¹⁰¹

Summary

At the end of his life, around 1881, Charles Darwin wrote, in an appendix to his autobiography, ‘My mind seems to have become a kind of machine for grinding general laws out of large collections of facts’. He was regretting the loss of his ability to appreciate the arts, caused by the massive mental effort of creating order from a vast mass of possibly unrelated pieces of information. One can only sympathise. Since the Periodic Table had brought an immense sense of order to chemists just twelve years earlier, it had created just the ordering in relation to the elements that Darwin could not yet envisage in biology. He might well have been astonished at the underlying comparative simplicity of life at the molecular level, when compared with the biodiversity he witnessed. We are now able to classify, catalogue, even to some extent interpret how different elements behave in those natural

environments, and can at least suggest, for example, why evolution has selected magnesium and potassium for the roles they play in living systems. We can look at the Periodic Table as a toolkit which can help us select the right element for the task in hand. The study of DNA structure has certainly benefited from this approach, as shown by the work of the many scientists quoted in this short account. Crystallographers have sometimes had a feeling of being privileged at being the first to see the actual atoms in some complex structure such as the ribosome, and it is the ability we now have to see each element in structures such as that of DNA, both natural and unnatural, that gives a feeling of order which Darwin, it seems, never quite knew.

This personal view would not be possible without the many students and co-workers I have been privileged to work with; no scientist is an island these days. The development of routine nucleic acid synthesis, the systematisation of crystallisation protocols, and the availability of tuneable synchrotron beamlines has turned the study of metal binding to DNA from an obscure corner of science to a wide-ranging area with applications from nanotechnology to therapeutics. The exploration of the Periodic Table in this context can sincerely be called inspirational.

Acknowledgements

Many thanks to Kane McQuaid, Jim Thorpe, Susy Teixeira and David Cardin for kindly reading and commenting on the manuscript. Thanks also to the many students and coworkers for so much inspiration over the years; I have aimed to cite as many of you as possible, as there are too many to name here. I very much appreciate the financial support given to our work by BBSRC, EPSRC, the Royal Society, Diamond Light Source, and the University of Reading.

References

- (1) Kellett, A.; Molphy, Z.; Slator, C.; McKee, V.; Farrell, N. P. Molecular Methods for Assessment of Non-Covalent Metallo drug-DNA Interactions. *Chem. Soc. Rev.* **2019**, *48* (4), 971–988. <https://doi.org/10.1039/c8cs00157j>.
- (2) Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Cardin, C. J. Structural Characterization of a New Crystal Form of the Four-Way Holliday Junction Formed by the DNA Sequence d(CCGGTACCGG)₂: Sequence versus Lattice? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58* (3), 567–569. <https://doi.org/10.1107/S09074444902001555>.
- (3) Kondo, J.; Yamada, T.; Hirose, C.; Okamoto, I.; Tanaka, Y.; Ono, A. Crystal Structure of Metallo DNA Duplex Containing Consecutive Watson-Crick-like T-HgII-T Base Pairs. *Angew. Chemie - Int. Ed.* **2014**, *53* (9), 2385–2388. <https://doi.org/10.1002/anie.201309066>.
- (4) Ennifar, E.; Walter, P.; Dumas, P. A Crystallographic Study of the Binding of 13 Metal Ions of Two Related RNA Duplexes. *Nucleic Acids Res.* **2003**, *31* (10), 2671–2682. <https://doi.org/10.1093/nar/gkg350>.
- (5) Todd, R. C.; Lippard, S. J. Structure of Duplex DNA Containing the Cisplatin 1,2-{Pt(NH₃)₂}₂+·d(GpG) Cross-Link at 1.77 Å Resolution. *J. Inorg. Biochem.* **2010**, *104* (9), 902–908. <https://doi.org/10.1016/j.jinorgbio.2010.04.005>.
- (6) Todd, A. K.; Adams, A.; Thorpe, J. H.; Denny, W. A.; Wakelin, L. P. G.; Cardin, C. J. Major Groove Binding and “DNA-Induced” Fit in the Intercalation of a Derivative of the Mixed Topoisomerase I/II Poison N-(2- (Dimethylamino)Ethyl)Acridine-4-Carboxamide (DACA) into DNA: X-Ray

- Structure Complexed to d(CG(5-Bru)ACG)₂ at 1.3-Å Resolution [2]. *J. Med. Chem.* **1999**, 42 (4). <https://doi.org/10.1021/jm980479u>.
- (7) Cardin, C. J.; Hall, J. P. Structural Studies of DNA-Binding Metal Complexes of Therapeutic Importance. In *DNA-targeting Molecules as Therapeutic Agents*; Royal Society of Chemistry, 2018; pp 198–227. <https://doi.org/http://dx.doi.org/10.1039/9781788012928>.
 - (8) Dennis, C.; Campbell, P. The Eternal Molecule. *Nature* **2003**, 421 (6921), 396. <https://doi.org/10.1038/nature01396>.
 - (9) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. *Nat. Chem.* **2013**, 5 (3), 182–186. <https://doi.org/10.1038/nchem.1548>.
 - (10) Evans, K.; Bhamra, I.; Wheelhouse, R. T.; Arnold, J. R. P.; Cosstick, R.; Fisher, J. Stabilization of a Bimolecular Triplex by 3'-S-Phosphorothiolate Modifications: An NMR and UV Thermal Melting Investigation. *Chem. - A Eur. J.* **2015**, 21 (19), 7278–7284. <https://doi.org/10.1002/chem.201500369>.
 - (11) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R.E. Crystal Structure Analysis of a Complete Turn of B-DNA. *Nature* **1980**, 287 (5784), 755–758.
 - (12) Hall, J. P.; Sanchez-Weatherby, J.; Alberti, C.; Quimper, C. H.; O'Sullivan, K.; Brazier, J. A.; Winter, G.; Sorensen, T.; Kelly, J. M.; Cardin, D. J.; et al. Controlled Dehydration of a Ruthenium Complex-DNA Crystal Induces Reversible DNA Kinking. *J. Am. Chem. Soc.* **2014**, 136 (50), 17505–17512. <https://doi.org/10.1021/ja508745x>.
 - (13) Kim, D.; Hur, J.; Park, K.; Bae, S.; Shin, D.; Ha, S. C.; Hwang, H. Y.; Hohng, S.; Lee, J. H.; Lee, S.; et al. Distinct Z-DNA Binding Mode of a PKR-like Protein Kinase Containing a Z-DNA Binding Domain (PKZ). *Nucleic Acids Res.* **2014**, 42 (9), 5937–5948. <https://doi.org/10.1093/nar/gku189>.
 - (14) Wu, Z.; Tian, T.; Yu, J.; Weng, X.; Liu, Y.; Zhou, X. Formation of Sequence-Independent Z-DNA Induced by a Ruthenium Complex at Low Salt Concentrations. *Angew. Chemie - Int. Ed.* **2011**, 50 (50), 11962–11967. <https://doi.org/10.1002/anie.201104422>.
 - (15) Hoogsteen, K. The Structure of Crystals Containing a Hydrogen-Bonded Complex of 1-Methylthymine and 9-Methyladenine. *Acta Crystallogr.* **1959**, 12 (10), 822–823. <https://doi.org/10.1107/S0365110X59002389>.
 - (16) Hoogsteen, K. The Crystal and Molecular Structure of a Hydrogen-Bonded Complex between 1-Methylthymine and 9-Methyladenine. *Acta Crystallogr.* **1963**, 16 (9), 907–916. <https://doi.org/10.1107/S0365110X63002437>.
 - (17) Nikolova, E. N.; Zhou, H.; Gottardo, F. L.; Alvey, H. S.; Kimsey, I. J.; Al-Hashimi, H. M. A Historical Account of Hoogsteen Base-Pairs in Duplex DNA. *Biopolymers* **2013**, 99 (12), 955–968. <https://doi.org/10.1002/bip.22334>.
 - (18) Gill, M. L.; Strobel, S. A.; Loria, J. P. Crystallization and Characterization of the Thallium Form of the Oxytricha Nova G-Quadruplex. *Nucleic Acids Res.* **2006**, 34 (16), 4506–4514. <https://doi.org/10.1093/nar/gkl616>.
 - (19) Fox, K. R.; Brown, T. Formation of Stable DNA Triplexes. *Biochem. Soc. Trans.* **2011**, 39 (2), 629–634. <https://doi.org/10.1042/bst0390629>.

- (20) Gurung, S. P.; Schwarz, C.; Hall, J. P.; Cardin, C. J.; Brazier, J. A. The Importance of Loop Length on the Stability of I-Motif Structures. *Chem. Commun.* **2015**, 51 (26), 5630–5632. <https://doi.org/10.1039/C4CC07279K>.
- (21) Abdelhamid, M. A. S.; Fábíán, L.; Macdonald, C. J.; Cheesman, M. R.; Gates, A. J.; Waller, Z. A. E. Redox-Dependent Control of i-Motif DNA Structure Using Copper Cations. *Nucleic Acids Res.* **2018**, 46 (12), 5886–5893. <https://doi.org/10.1093/nar/gky390>.
- (22) Zeraati, M.; Langley, D. B.; Schofield, P.; Moye, A. L.; Rouet, R. I-Motif DNA Structures Are Formed in the Nuclei. *Nat. Chem.* **2018**, 10 (June 2018), 631–637.
- (23) Bruner, S. D.; Norman, D. P. G.; Verdine, G. L. Structural Basis for Recognition of the Endogenous Mutagen 8-Hydroxyguanine. *Nature* **2000**, 403, 859–866.
- (24) Keane, P. M.; Poynton, F. E.; Hall, J. P.; Sazanovich, I. V.; Towrie, M.; Gunnlaugsson, T.; Quinn, S. J.; Cardin, C. J.; Kelly, J. M. Reversal of a Single Base-Pair Step Controls Guanine Photo-Oxidation by an Intercalating Ruthenium(II) Dipyridophenazine Complex. *Angew. Chemie* **2015**, 127 (29), 8484–8488. <https://doi.org/10.1002/ange.201502608>.
- (25) Hall, J. P.; Poynton, F. E.; Keane, P. M.; Gurung, S. P.; Brazier, J. A.; Cardin, D. J.; Winter, G.; Gunnlaugsson, T.; Sazanovich, I. V.; Towrie, M.; et al. Monitoring One-Electron Photo-Oxidation of Guanine in DNA Crystals Using Ultrafast Infrared Spectroscopy. *Nat. Chem.* **2015**, 7 (12), 961–967. <https://doi.org/10.1038/nchem.2369>.
- (26) Hall, J. P.; Keane, P. M.; Beer, H.; Buchner, K.; Winter, G.; Sorensen, T. L.; Cardin, D. J.; Brazier, J. A.; Cardin, C. J. Delta Chirality Ruthenium ‘Light-Switch’ Complexes Can Bind in the Minor Groove of DNA with Five Different Binding Modes. *Nucleic Acids Res.* **2016**, 44 (19), 9472–9482. <https://doi.org/10.1093/nar/gkw753>.
- (27) Song, H.; Kaiser, J. T.; Barton, J. K. Crystal Structure of Δ -[Ru(Bpy)₂dppz]²⁺ Bound to Mismatched DNA Reveals Side-by-Side Metalloinsertion and Intercalation. *Nat. Chem.* **2012**, 4 (8), 615–620. <https://doi.org/10.1038/nchem.1375>.
- (28) Thorpe, J. H.; Gale, B. C.; Teixeira, S. C. M.; Cardin, C. J. Conformational and Hydration Effects of Site-Selective Sodium, Calcium and Strontium Ion Binding to the DNA Holliday Junction Structure d(TCGGTACCGA)₄. *J. Mol. Biol.* **2003**, 327 (1), 97–109. [https://doi.org/10.1016/S0022-2836\(03\)00088-3](https://doi.org/10.1016/S0022-2836(03)00088-3).
- (29) Romani, A. Cellular Magnesium Homeostasis. *Arch Biochem Biophys* **2011**, 512, 1–23. <https://doi.org/10.1016/j.abb.2011.05.010>.
- (30) Drew, H. R. R.; Wing, R. M. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. E. Structure of a B-DNA Dodecamer: Conformation and Dynamics. *Proc. Natl. Acad. Sci.* **1981**, 78 (4), 2179–2183. <https://doi.org/10.1073/pnas.78.4.2179>.
- (31) Hiller, D. A.; Strobel, S. A. The Chemical Versatility of RNA. *Philos. Trans. R. Soc. B Biol. Sci.* **2011**, 366 (1580), 2929–2935. <https://doi.org/10.1098/rstb.2011.0143>.
- (32) Crick, F. H. C.; Klug, A. Kinky Helix. *Nature* **1975**, 255, 530–533.
- (33) Juo, Z. S.; Kassavetis, G. A.; Wang, J.; Geiduschek, E. P.; Sigler, P. B. Crystal Structure of a Transcription Factor IIIB Core Interface Ternary Complex. *Nature* **2003**, 422, 534–539.
- (34) Wilson, J. J.; Lippard, S. J. Synthetic Methods for the Preparation of Platinum Anticancer Complexes. *Chem. Rev.* **2014**, 114 (8), 4470–4495. <https://doi.org/10.1021/cr4004314>.

- (35) Reißner, T.; Schneider, S.; Schorr, S.; Carell, T. Crystal Structure of a Cisplatin-(1,3-GTG) Cross-Link within DNA Polymerase η . *Angew. Chemie - Int. Ed.* **2010**, *49* (17), 3077–3080. <https://doi.org/10.1002/anie.201000414>.
- (36) Alam, M. N.; Huq, F. Comprehensive Review on Tumour Active Palladium Compounds and Structure–activity Relationships. *Coord. Chem. Rev.* **2016**, *316*, 36–67.
- (37) Cini, M.; Bradshaw, T. D.; Woodward, S. Using Titanium Complexes to Defeat Cancer: The View from the Shoulders of Titans. *Chem. Soc. Rev.* **2017**, *46*, 1040–1051.
- (38) Howerton, B. S.; Heidary, D. K.; Glazer, E. C. Strained Ruthenium Complexes Are Potent Light-Activated Anticancer Agents. *J. Am. Chem. Soc.* **2012**, *134* (20), 8324–8327. <https://doi.org/10.1021/ja3009677>.
- (39) Lilley, D. M. J. All Change at Holliday Junction. *Proc. Natl. Acad. Sci.* **2002**, *94* (18), 9513–9515. <https://doi.org/10.1073/pnas.94.18.9513>.
- (40) Hadden, J. M.; Déclais, A.-C.; Carr, S. B.; Lilley, D. M. J.; Phillips, S. E. V. The Structural Basis of Holliday Junction Resolution by T7 Endonuclease I. *Nature* **2007**, *449* (7162), 621–624. <https://doi.org/10.1038/nature06158>.
- (41) Coll, M.; Ortiz-Lombardía, M.; González, A.; Eritja, R.; Aymamí, J.; Azorín, F. Crystal Structure of a DNA Holliday Junction. *Nat. Struct. Biol.* **1999**, *6* (10), 913–917. <https://doi.org/10.1038/13277>.
- (42) Hays, F. A.; Teegarden, A.; Jones, Z. J. R.; Harms, M.; Raup, D.; Watson, J.; Cavaliere, E.; Ho, P. S. How Sequence Defines Structure: A Crystallographic Map of DNA Structure and Conformation. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (20), 7157–7162. <https://doi.org/10.1073/pnas.0409455102>.
- (43) Cardin, C. J.; Thorpe, J. H.; Gale, B. C.; Teixeira, S. C. M. Strontium, a MAD target for the DNA Holliday junction <http://www.rcsb.org/structure/1NVY> (accessed May 1, 2019). <https://doi.org/10.2210/PDB1NVY/PDB>.
- (44) Hall, J. P.; O’Sullivan, K.; Naseer, A.; Smith, J. A.; Kelly, J. M.; Cardin, C. J. Structure Determination of an Intercalating Ruthenium Dipyrrophenazine Complex Which Kinks DNA by Semiintercalation of a Tetraazaphenanthrene Ligand. *Proc. Natl. Acad. Sci.* **2011**, *108* (43), 17610–17614. <https://doi.org/10.1073/pnas.1108685108>.
- (45) Naseer, A.; Cardin, C. J. Strontium bound to the Holliday junction sequence d(TCGGCGCCGA)4 <https://www.rcsb.org/structure/3GOO> (accessed May 1, 2019). <https://doi.org/10.2210/PDB3GOO/PDB>.
- (46) Laughlan, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M. J.; Luisi, B. The High-Resolution Crystal Structure of a Parallel-Stranded Guanine Tetraplex. *Science* (80-.). **1994**, *265*, 520–524.
- (47) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. Crystal Structure of Parallel Quadruplexes from Human Telomeric DNA. *Nature* **2002**, *417* (6891), 876–880. <https://doi.org/10.1038/nature755>.
- (48) Haider, S. M.; Neidle, S.; Parkinson, G. N. A Structural Analysis of G-Quadruplex/Ligand Interactions. *Biochimie* **2011**, *93* (8), 1239–1251. <https://doi.org/10.1016/j.biochi.2011.05.012>.
- (49) Neidle, S. Quadruplex Nucleic Acids as Targets for Anticancer Therapeutics. *Nature Reviews*

Chemistry. Nature Publishing Group January 11, 2017. <https://doi.org/10.1038/s41570-017-0041>.

- (50) Lin, C.; Wu, G.; Wang, K.; Onel, B.; Sakai, S.; Shao, Y.; Yang, D. Molecular Recognition of the Hybrid-2 Human Telomeric G-Quadruplex by Epiberberine: Insights into Conversion of Telomeric G-Quadruplex Structures. *Angew. Chemie - Int. Ed.* **2018**, *57* (34), 10888–10893. <https://doi.org/10.1002/anie.201804667>.
- (51) Tan, J.-H.; Liu, H.-Y.; He, J.-H.; Huang, Z.-S.; Zhao, Q.; Ou, T.-M.; Ge, Y.-L.; Lv, P.; Gu, L.-Q.; Li, D.; et al. Conformation Selective Antibody Enables Genome Profiling and Leads to Discovery of Parallel G-Quadruplex in Human Telomeres. *Cell Chem. Biol.* **2016**, *23* (10), 1261–1270. <https://doi.org/10.1016/j.chembiol.2016.08.013>.
- (52) Stump, S.; Mou, T. C.; Sprang, S. R.; Natale, N. R.; Beall, H. D. Crystal Structure of the Major Quadruplex Formed in the Promoter Region of the Human C-MYC Oncogene. *PLoS One* **2018**, *13* (10), 1–15. <https://doi.org/10.1371/journal.pone.0205584>.
- (53) Asamitsu, S.; Obata, S.; Yu, Z.; Bando, T.; Sugiyama, H. Recent Progress of Targeted G-Quadruplex-Preferred Ligands Toward Cancer Therapy. *Molecules* **2019**, *24* (3), 429. <https://doi.org/10.3390/molecules24030429>.
- (54) Tateishi-Karimata, H.; Kawauchi, K.; Sugimoto, N. Destabilization of DNA G-Quadruplexes by Chemical Environment Changes during Tumor Progression Facilitates Transcription. *J. Am. Chem. Soc.* **2018**, *140* (2), 642–651. <https://doi.org/10.1021/jacs.7b09449>.
- (55) Quigley, G. J.; Wang, A. H.; Ughetto, G.; Marel, G. van de; Boom, J. H. van; Rich, A. Molecular Structure of an Anticancer Drug-DNA Complex: Daunomycin plus d(CpGpTpApCpG). *Biochemistry* **1980**, *77* (12), 7204–7208.
- (56) Thomas, J. A. Optical Imaging Probes for Biomolecules: An Introductory Perspective. *Chem. Soc. Rev.* **2015**, *44* (14), 4494–4500. <https://doi.org/10.1039/C5CS00070J>.
- (57) Heinemann, F.; Karges, J.; Gasser, G. Critical Overview of the Use of Ru(II) Polypyridyl Complexes as Photosensitizers in One-Photon and Two-Photon Photodynamic Therapy. *Acc. Chem. Res.* **2017**, *50* (11), 2727–2736. <https://doi.org/10.1021/acs.accounts.7b00180>.
- (58) Mari, C.; Pierroz, V.; Ferrari, S.; Gasser, G. Combination of Ru(II) Complexes and Light: New Frontiers in Cancer Therapy. *Chem. Sci.* **2015**, *6* (5), 2660–2686. <https://doi.org/10.1039/c4sc03759f>.
- (59) He, Y.; Lopez, A.; Zhang, Z.; Chen, D.; Yang, R.; Liu, J. Nucleotides and DNA Coordinated Lanthanides: From Fundamentals to Applications. *Coord. Chem. Rev.* **2019**, *387*, 235–248. <https://doi.org/doi.org/10.1016/j.ccr.2019.02.020>.
- (60) Friedman, A. E.; Chambron, J. C.; Sauvage, J. P.; Turro, N. J.; Barton, J. K. A Molecular Light Switch for DNA: Ru(Bpy)₂(Dppz)²⁺. *J. Am. Chem. Soc.* **1990**, *112* (12), 4960–4962. <https://doi.org/10.1021/ja00168a052>.
- (61) Cardin, C. J.; Kelly, J. M.; Quinn, S. J. Photochemically Active DNA-Intercalating Ruthenium and Related Complexes – Insights by Combining Crystallography and Transient Spectroscopy. *Chem. Sci.* **2017**, *8* (7), 4705–4723. <https://doi.org/10.1039/C7SC01070B>.
- (62) Zeglis, B. M.; Pierre, V. C.; Kaiser, J. T.; Barton, J. K. A Bulky Rhodium Complex Bound to an Adenosine-Adenosine DNA Mismatch: General Architecture of the Metalloinsertion Binding Mode. *Biochemistry* **2009**, *48* (20), 4247–4253. <https://doi.org/10.1021/bi900194e>.

- (63) Foxon, S. P.; Phillips, T.; Gill, M. R.; Towrie, M.; Parker, A. W.; Webb, M.; Thomas, J. A. A Multifunctional Light Switch: DNA Binding and Cleavage Properties of a Heterobimetallic Ruthenium-Rhenium Dipyridophenazine Complex. *Angew. Chemie - Int. Ed.* **2007**, *46* (20), 3686–3688. <https://doi.org/10.1002/anie.200604837>.
- (64) Brueckmann, N. E.; Kuegel, S.; Hamacher, A.; Kassack, M. U.; Kunz, P. C. Fluorescent Polylactides with Rhenium(Bisimine) Cores for Tumour Diagnostics. *Eur. J. Inorg. Chem.* **2010**, No. 32, 5063–5068. <https://doi.org/10.1002/ejic.201000721>.
- (65) Wragg, A.; Gill, M. R.; Turton, D.; Adams, H.; Roseveare, T. M.; Smythe, C.; Su, X.; Thomas, J. A. Tuning the Cellular Uptake Properties of Luminescent Heterobimetallic Iridium(III)-Ruthenium(II) DNA Imaging Probes. *Chem. - A Eur. J.* **2014**, *20* (43), 14004–14011. <https://doi.org/10.1002/chem.201403693>.
- (66) Poynton, F. E.; Bright, S. A.; Blasco, S.; Williams, D. C.; Kelly, J. M.; Gunnlaugsson, T. The Development of Ruthenium(II) Polypyridyl Complexes and Conjugates for: In Vitro Cellular and in Vivo Applications. *Chem. Soc. Rev.* **2017**, *46* (24), 7706–7756. <https://doi.org/10.1039/c7cs00680b>.
- (67) Wragg, A.; Gill, M. R.; Hill, C. J.; Su, X.; Meijer, A. J. H. M.; Smythe, C.; Thomas, J. A. Dinuclear Osmium(II) Probes for High-Resolution Visualisation of Cellular DNA Structure Using Electron Microscopy. *Chem. Commun.* **2014**, *50* (93), 14494–14497. <https://doi.org/10.1039/c4cc05547k>.
- (68) Niyazi, H.; Hall, J. P.; O'Sullivan, K.; Winter, G.; Sorensen, T.; Kelly, J. M.; Cardin, C. J. Crystal Structures of Λ -[Ru(Phen)2dppz]2+ with Oligonucleotides Containing TA/TA and AT/AT Steps Show Two Intercalation Modes. *Nat. Chem.* **2012**, *4* (8), 621–628. <https://doi.org/10.1038/nchem.1397>.
- (69) Keane, P. M.; Poynton, F. E.; Hall, J. P.; Sazanovich, I. V.; Towrie, M.; Gunnlaugsson, T.; Quinn, S. J.; Cardin, C. J.; Kelly, J. M. Reversal of a Single Base-Pair Step Controls Guanine Photo-Oxidation by an Intercalating Ruthenium(II) Dipyridophenazine Complex. *Angew. Chemie - Int. Ed.* **2015**, *54* (29), 8364–8368. <https://doi.org/10.1002/anie.201502608>.
- (70) Hall, J. P.; Cook, D.; Morte, S. R.; McIntyre, P.; Buchner, K.; Beer, H.; Cardin, D. J.; Brazier, J. A.; Winter, G.; Kelly, J. M.; et al. X-Ray Crystal Structure of Rac- [Ru(Phen) 2 Dppz] 2+ with d(ATGCAT) 2 Shows Enantiomer Orientations and Water Ordering. *J. Am. Chem. Soc.* **2013**, *135* (34), 12652–12659. <https://doi.org/10.1021/ja403590e>.
- (71) McQuaid, K.; Hall, J. P.; Brazier, J. A.; Cardin, D. J.; Cardin, C. J. X-Ray Crystal Structures Show DNA Stacking Advantage of Terminal Nitrile Substitution in Ru-Dppz Complexes. *Chem. - A Eur. J.* **2018**, *24* (59), 15859–15867. <https://doi.org/10.1002/chem.201803021>.
- (72) Gill, M. R.; Garcia-Lara, J.; Foster, S. J.; Smythe, C.; Battaglia, G.; Thomas, J. A. A Ruthenium(II) Polypyridyl Complex for Direct Imaging of DNA Structure in Living Cells. *Nat. Chem.* **2009**, *1* (8), 662–667. <https://doi.org/10.1038/nchem.406>.
- (73) Łęczkowska, A.; Gonzalez-Garcia, J.; Perez-Arnaiz, C.; Garcia, B.; White, A. J. P.; Vilar, R. Binding Studies of Metal–Salphen and Metal–Bipyridine Complexes towards G-Quadruplex DNA. *Chem. - A Eur. J.* **2018**, *24* (45), 11785–11794. <https://doi.org/10.1002/chem.201802248>.
- (74) Ang, D. L.; Harper, B. W. J.; Cubo, L.; Mendoza, O.; Vilar, R.; Aldrich-Wright, J. Quadruplex DNA-Stabilising Dinuclear Platinum(II) Terpyridine Complexes with Flexible Linkers. *Chem. - A Eur. J.* **2016**, *22* (7), 2317–2325. <https://doi.org/10.1002/chem.201503663>.

- (75) Bazzicalupi, C.; Ferraroni, M.; Papi, F.; Massai, L.; Bertrand, B.; Messori, L.; Gratteri, P.; Casini, A. Determinants for Tight and Selective Binding of a Medicinal Dicarbene Gold(I) Complex to a Telomeric DNA G-Quadruplex: A Joint ESI MS and XRD Investigation. *Angew. Chemie - Int. Ed.* **2016**, 55 (13), 4256–4259. <https://doi.org/10.1002/anie.201511999>.
- (76) Vilar, R. Nucleic Acid Quadruplexes and Metallo-Drugs. In *Metallo-Drugs: Development and Action of Anticancer Agents*; De Gruyter: Berlin, Boston, Boston, 2018; pp 325–350.
- (77) McQuaid, K.; Abell, H.; Gurung, S. P.; Allan, D.; Winter, G.; Sorensen, T.; Cardin, D. J.; Brazier, J. A.; Cardin, C. J.; Hall, J. P. Structural Studies Reveal the Enantiospecific Recognition of a DNA G-Quadruplex by a Ruthenium Polypyridyl Complex. *Angew. Chemie Int. Ed.* **2019**. <https://doi.org/10.1002/anie.201814502>.
- (78) Wachter, E.; Moyá, D.; Parkin, S.; Glazer, E. C. Ruthenium Complex “Light Switches” That Are Selective for Different G-Quadruplex Structures. *Chem. - A Eur. J.* **2016**, 22 (2), 550–559. <https://doi.org/10.1002/chem.201503203>.
- (79) Brodersen, D. E.; Clemons, W. M.; Carter, A. P.; Wimberly, B. T.; Ramakrishnan, V. Phasing the 30S Ribosomal Subunit Structure. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **2003**, 59 (11), 2044–2050. <https://doi.org/10.1107/S0907444903017669>.
- (80) Ramakrishnan, V. Ribosome Structure and the Mechanism of Translation. *Cell* **2002**, 108 (4), 557–572. [https://doi.org/10.1016/S0092-8674\(02\)00619-0](https://doi.org/10.1016/S0092-8674(02)00619-0).
- (81) Berger, I.; Kang, C.; Sinha, N. D.; Wolters, M.; Rich, A. A Highly Efficient 24-Condition Matrix for the Crystallization of Nucleic Acid Fragments. *Acta Crystallogr D Biol Crystallogr* **1996**, 52, 465–468.
- (82) Viladoms, J.; Parkinson, G. N. HELIX: A New Modular Nucleic Acid Crystallization Screen. *J. Appl. Crystallogr.* **2014**, 47 (3), 948–955. <https://doi.org/10.1107/s1600576714007407>.
- (83) Leal, R. M. F.; Callow, S.; Callow, P.; Blakeley, M. P.; Cardin, C. J.; Denny, W. A.; Teixeira, S. C. M.; Mitchell, E. P.; Forsyth, V. T. Combined Neutron and X-Ray Diffraction Studies of DNA in Crystals and Solutions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66 (11), 1244–1248. <https://doi.org/10.1107/S0907444910017713>.
- (84) Keane, P.; Hall, J.; Poynton, F.; Poulsen, B.; Gurung, S.; Clark, I.; Sazanovich, I.; Towrie, M.; Gunnlaugsson, T.; Quinn, S.; et al. Inosine Can Increase DNA’s Susceptibility to Photo-Oxidation by a Ru(II) Complex Due to Structural Change in the Minor Groove. *Chem. - A Eur. J.* **2017**, No. li. <https://doi.org/10.1002/chem.201701447>.
- (85) Chiu, T. K.; Dickerson, R. E. 1 Å Crystal Structures of B-DNA Reveal Sequence-Specific Binding and Groove-Specific Bending of DNA by Magnesium and Calcium. *J. Mol. Biol.* **2000**, 301 (4), 915–945. <https://doi.org/10.1006/jmbi.2000.4012>.
- (86) Leal, R. M. F.; Teixeira, S. C. M.; Blakeley, M. P.; Mitchell, E. P.; Forsyth, V. T. A Preliminary Neutron Crystallographic Study of an A-DNA Crystal. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* **2009**, 65 (Pt 3), 232–235. <https://doi.org/10.1107/S1744309109002668>.
- (87) Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Cardin, C. J. Crystal Structure of the Complementary Quadruplex Formed by d(GCATGCT) at Atomic Resolution. *Nucleic Acids Res.* **2003**, 31 (3). <https://doi.org/10.1093/nar/gkg168>.
- (88) Leonard, G. A.; Zhang, S.; Peterson, M. R.; Harrop, S. J.; Helliwell, J. R.; Cruse, W. B.; Langlois d’Estaintot, B.; Kennard, O.; Brown, T.; Hunter, W. N. Self-Association of a DNA Loop Creates a

- Quadruplex: Crystal Structure of d(GCATGCT) at 1.8 Å Resolution. *Structure* **1995**, 3 (4), 335–340. [https://doi.org/10.1016/S0969-2126\(01\)00165-4](https://doi.org/10.1016/S0969-2126(01)00165-4).
- (89) Gan, Y.; Thorpe, J. H.; Teixeira, S. C. M.; Naseer, A.; Cardin, C. J. Calcium and the DNA Quadruplex Structure formed by d(GCATGCT) <https://www.rcsb.org/structure/3T86> (accessed May 1, 2019). <https://doi.org/10.2210/PDB3T86/PDB>.
 - (90) Cardin, C. J.; Gan, Y.; Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Moraes, M. I. A. Barium and the DNA Quadruplex Structure Formed by d(GCATGCT) <http://www.rcsb.org/structure/1QYK> (accessed May 1, 2019). <https://doi.org/10.2210/PDB1QYK/PDB>.
 - (91) Cardin, C. J.; Gan, Y.; Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Moraes, M. I. A. Vanadium and the DNA Quadruplex Structure Formed by d(GCATGCT) <http://www.rcsb.org/structure/1QYL> (accessed May 1, 2019). <https://doi.org/10.2210/PDB1QYL/PDB>.
 - (92) Cardin, C. J.; Gan, Y.; Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Moraes, M. I. A. Cobalt and the DNA Quadruplex Structure Formed by d(GCATGCT). *PDB code 1QZL*. <https://doi.org/10.2210/PDB1QZL/PDB>.
 - (93) Cardin, C. J.; Gan, Y.; Thorpe, J. H.; Teixeira, S. C. M.; Gale, B. C.; Moraes, M. I. A. Nickel and the DNA Quadruplex Structure Formed by d(GCATGCT) <http://www.rcsb.org/structure/1R2O> (accessed May 1, 2019). <https://doi.org/10.2210/PDB1R2O/PDB>.
 - (94) Nakano, M.; Tateishi-Karimata, H.; Tanaka, S.; Tama, F.; Miyashita, O.; Nakano, S. I.; Sugimoto, N. Thermodynamic Properties of Water Molecules in the Presence of Cosolute Depend on DNA Structure: A Study Using Grid Inhomogeneous Solvation Theory. *Nucleic Acids Res.* **2015**, 43 (21), 10114–10125. <https://doi.org/10.1093/nar/gkv1133>.
 - (95) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; et al. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, 67 (4), 235–242. <https://doi.org/10.1107/S0907444910045749>.
 - (96) Taylor, G. L. Introduction to Phasing. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66 (4), 325–338. <https://doi.org/10.1107/S0907444910006694>.
 - (97) Todd, A. K.; Adams, A.; Powell, H. R.; Wilcock, D. J.; Thorpe, J. H.; Lausi, A.; Zanini, F.; Wakelin, L. P. G.; Cardin, C. J. Determination by MAD-DM of the Structure of the DNA Duplex d[ACGTACG(5-BrU)]₂ at 1.46 Å and 100 K. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, 55 (4). <https://doi.org/10.1107/S090744499801261X>.
 - (98) Todd, A. K.; Adams, A.; Thorpe, J. H.; Denny, W. A.; Wakelin, L. P. G.; Cardin, C. J. Major Groove Binding and ‘DNA-Induced’ Fit in the Intercalation of a Derivative of the Mixed Topoisomerase I/II Poison N-(2(Dimethylamino)Ethyl)Acridine-4- Carboxamide (DACA) into DNA: X-Ray Structure Complexed to d(CG(5-BrU)ACG)₂ at 1.3-Å Resolution.’ *J. Med. Chem.* **1999**, 42 (4), 536–540. <https://doi.org/10.1021/jm980479u>.
 - (99) Hall, J. P.; Gurung, S. P.; Henle, J.; Poidl, P.; Andersson, J.; Lincoln, P.; Winter, G.; Sorensen, T.; Cardin, D. J.; Brazier, J. A.; et al. Guanine Can Direct Binding Specificity of Ru–dipyridophenazine (Dppz) Complexes to DNA through Steric Effects. *Chem. - A Eur. J.* **2017**, 23 (21), 4981–4985. <https://doi.org/10.1002/chem.201605508>.
 - (100) Sheldrick, G. M. Experimental Phasing with SHELXC/D/E: Combining Chain Tracing with Density Modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66 (4), 479–485. <https://doi.org/10.1107/S0907444909038360>.

- (101) Parkhurst, J. M.; Winter, G.; Waterman, D. G.; Fuentes-Montero, L.; Gildea, R. J.; Murshudov, G. N.; Evans, G. Robust Background Modelling in DIALS. *J. Appl. Crystallogr.* **2016**, *49* (6), 1912–1921. <https://doi.org/10.1107/S1600576716013595>.