Crystal structures of Λ-[Ru(phen)2dppz]2+ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes

Article

Accepted Version

Crystal structures of Λ-$[Ru(phen)_2dppz]^{2+}$ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes

Hakan Niyazi[a,d], James P. Hall[a], Kyra O'Sullivan[c], Graeme Winter[b], Thomas Sorensen[b], John M. Kelly[c] and Christine J. Cardin[a]

[a] Department of Chemistry, University of Reading, Whiteknights, Reading, RG6 6AD, United Kingdom

[b]Diamond Light Source Ltd, Diamond House, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, United Kingdom

[c] School of Chemistry, Trinity College, Dublin 2, Ireland.

[d] Institut Laue-Langevin, 6 Rue Jules Horowitz, 38042 Grenoble, France; European Synthrotron Radiation Facility, 6 Rue Jules Horowitz, 38043 Grenoble, France

Corresponding author

Correspondence to Christine J Cardin (c.j.cardin@rdg.ac.uk)

The ruthenium complex $[Ru(phen)_2(dppz)]^{2+}$ (where phen is a phenanthroline and dppz a dipyridyl–phenazine ligand) is known as a 'light-switch' complex because its luminescence in solution is significantly enhanced in the presence of DNA. This property is poised to serve diagnostic and therapeutic applications, but its binding mode with DNA needs to be further elucidated. Here we describe the crystal structures of the Λ enantiomer bound to two oligonucleotide duplexes. The dppz ligand intercalates symmetrically and perpendicularly from the minor groove of the $d(CCGGTACCGG)_2$ duplex at the central TA/TA step — but not at the central AT/AT step of $d(CCGGATCCGG)_2$. In both structures, however, a second ruthenium complex links the duplexes by a combination of a shallower angled intercalation into the C_1C_2/G_9G_{10} step at the end of the duplex, and semi-intercalation into the G_3G_4 step of an adjacent duplex. The TA/TA specificity of the perpendicular intercalation arises from the packing of phenanthroline ligands against the adenosine residue.

The binding of small molecules to DNA is key in understanding the mode of action of many drugs, and increasingly useful too in nanotechnological applications of nucleic acids. Intercalation as a binding mode was described by Lerman in his pioneering studies, based on observations with acridine (10-aza-anthracene) derivatives, and using fluorescence and dichroic methods[1].

Crystallographic studies[2-6] later showed that the acridine chromophore was indeed intercalated between base pairs of the B-DNA double helix, in parallel fashion, that is with the long axis of the acridine molecule parallel to those of the base-pair. This is accompanied by significant local unwinding of B-DNA; the helix is typically lengthened by about 3.4 Å. Further structural studies have shown that a variety of intercalation geometries is possible with different chromophores, for example parallel, perpendicular, or — as we report here — angled.

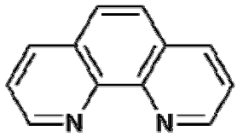
Several of these reports deal with molecules which are clinically important anticancer agents, such as the anthracycline antitumour antibiotic daunomycin[7], used in the treatment of leukaemias. The polypeptide antibiotic actinomycin D[8] inhibits both DNA and RNA synthesis by blocking chain elongation. These species, however, often have bulky functionalised substituents on the central heteroaromatic portion of the molecule, which undergo specific interactions in the grooves of the DNA, and might therefore have a determining role on the orientation of the molecule. For many other molecules, studies of intercalation are based on biophysical and spectroscopic studies in solution[9], which demonstrate, as expected, an orientation of the intercalated chromophore, helix lengthening, overall unwinding (in most cases, and as distinct from local twist angle), as well as the anticipated shifts in the positions of spectroscopic features of the intercalated molecule. Although nuclear Overhauser effect nuclear magnetic resonance (NOE NMR) studies have

been possible in a few cases[10], structural models of intercalation from NMR data alone are rare[11].

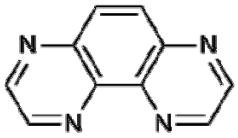
Because of their potential as biosensors and therapeutic agents, there has been very considerable interest in metal complexes where one of the ligands can intercalate between the base pairs of DNA, and the area has been reviewed recently[12-14]. Octahedral complexes with a dipyridophenazine (dppz) ligand have been particularly intensively studied, because of the so-called'light-switch'effect. These complex cations show no photoluminescence in water at ambient temperatures, but display intense photoluminescence when added to solutions of calf thymus and synthetic DNAs. The luminescence is also seen in non-aqueous solutions, leading to the assumption that the key property is the exclusion of water from the chromophore in DNA. These unusual photophysical properties[15,16] make them suitable as DNA light-switching sensors[17,18], as agents to target photo-oxidation in DNA[19,20] and as probes for DNA conductivity[21]. it is known from spectroscopic and biophysical methods that the dppz ligand is intercalated between the base-pairs of the polynucleotide[22-24] but so far no definitive evidence for any binding site for **1** had been available.

We have now carried out the crystallographic study of a complex between Λ-$[Ru(phen)_2(dppz)]^{2+}$ (**1**) and an oligonucleotide. Here we describe the specific perpendicular, symmetrical, classical intercalation mode we observed for the dppz ligand of (**1**) at the TA/TA central step of a $d(CCGG\mathbf{TA}CCGG)_2$ duplex. We contrast this with the lack of such binding at the AT/AT central step of the $d(CCGG\mathbf{AT}CCGG)_2$ duplex, which differs only in the switch from TA to AT at the central step. With both sequences, we also observe a second metal complex intercalated with the oligonucleotide for which, in contrast to previous studies, the intercalation of dppz at the terminal GG/CC step that occurs is angled (neither parallel nor perpendicular with the base pairs). The angle also enables another interaction, not observed previously, to take place: the semi-intercalation of phen (1,10 phenanthroline) at the

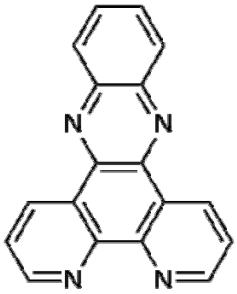
G_3G_4/C_7C_8 step — similar to that observed for Λ-$[Ru(TAP)_2(dppz)]^{2+}$ (**2**) with the d(TCGGCGCCGA)$_2$ duplex[25]. There is a minor end-effect of a *syn*-guanine at G10, but otherwise, this work provides striking confirmation in the solid state of the evidence already gathered in solution for more than one type of binding site[26], and may explain the presence of two fluorescence decay lifetimes found for **1** bound to DNA in solution, even for single enantiomers and non-alternating double-stranded homopolymers.



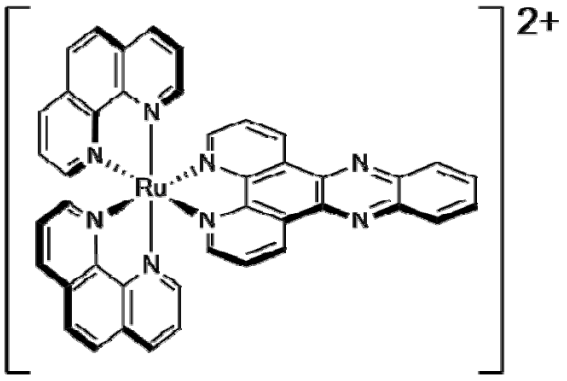
phen



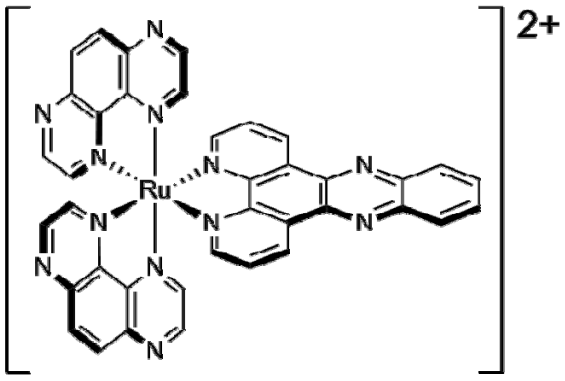
TAP



dppz



1



2

Results

Crystallisation, measurement and structure solution. Racemic $[Ru(phen)_2(dppz)]Cl_2$ and a decamer — either d(CCGGTACCGG)[20-21] or d(CCGGATCCGG)[22] — in a buffer containing spermine, 2-methyl-2,4-pentanediol, sodium cacodylate pH 7.0, NaCl and $BaCl_2$ were allowed to stand in the dark, at 20°C for d(CCGGTACCGG) and 4°C for d(CCGGATCCGG). Crystals formed after a few days, as orange rhombs. The crystals were flash-frozen, and their structure subsequently characterized by X-ray crystallography (data measured at 100 K). Full details of the data collection and refinement statistics are given in the Methods section and Supplementary Information (Text T1 and T2, Tables S1 and S2).

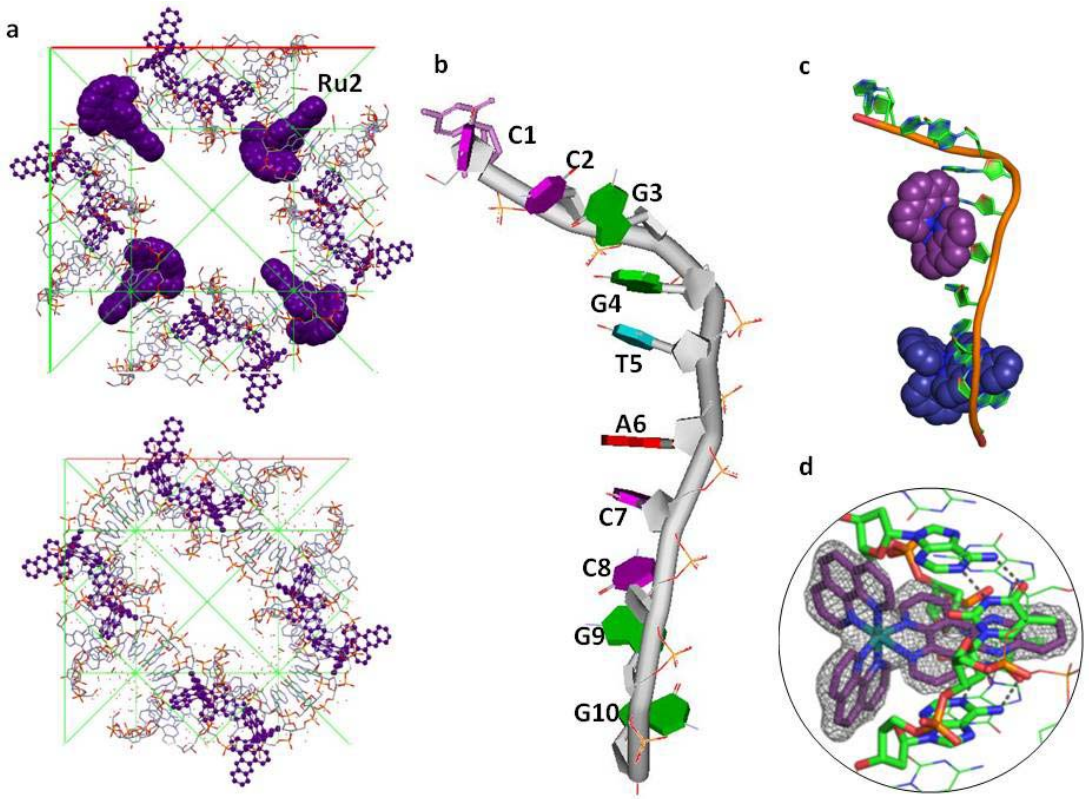


Figure 1 | Crystallisation of the ruthenium cation Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ and oligonucleotides. ***(a)*** *Packing diagrams of the two structures projected down the c axial direction, to highlight the symmetrical intercalation of Ru2 at the T_5A_6/T_5A_6 step in the d(CCGGTACCGG)$_2$ duplex (above) compared with the d(CCGGATCCGG)$_2$ duplex (below). Ru2 shown space-filling in purple, Ru1 shown ball-and-stick in purple.* ***(b)*** *Conformation of a single strand of the decamer d(CCGGTACCGG), with the base numbering. The disordered C1 is shown in the two main conformations (each 0.3 occupancy). Note the syn conformation of guanine residue G10. Colour code for bases is cytosine: magenta, guanine: green, thymine: cyan, adenine : red. The DNA backbone, bases and sugars are shown in cartoon mode in grey.* ***(c)*** *Location of the ruthenium cations in the asymmetric unit of the single strand. Ru 1 in dark blue, Ru2 in purple, both in spacefilling mode, DNA backbone (connecting P positions) orange, carbon: green, oxygen: red, nitrogen: blue, all in cartoon mode. The central Ru2 cation is treated as a whole complex of 0.5 occupancy per oligonucleotide strand, and the view direction is down the crystallographic twofold axis into the minor groove* ***(d)*** *The final electron density map at the T_5A_6/T_5A_6 step of d(CCGGTACCGG)$_2$, showing the intercalated ruthenium complex Ru2, drawn in stick mode. Map drawn, in grey, at the 1.0 sigma contour level. Nucleic acid strands in stick mode, phosphorus: orange.*

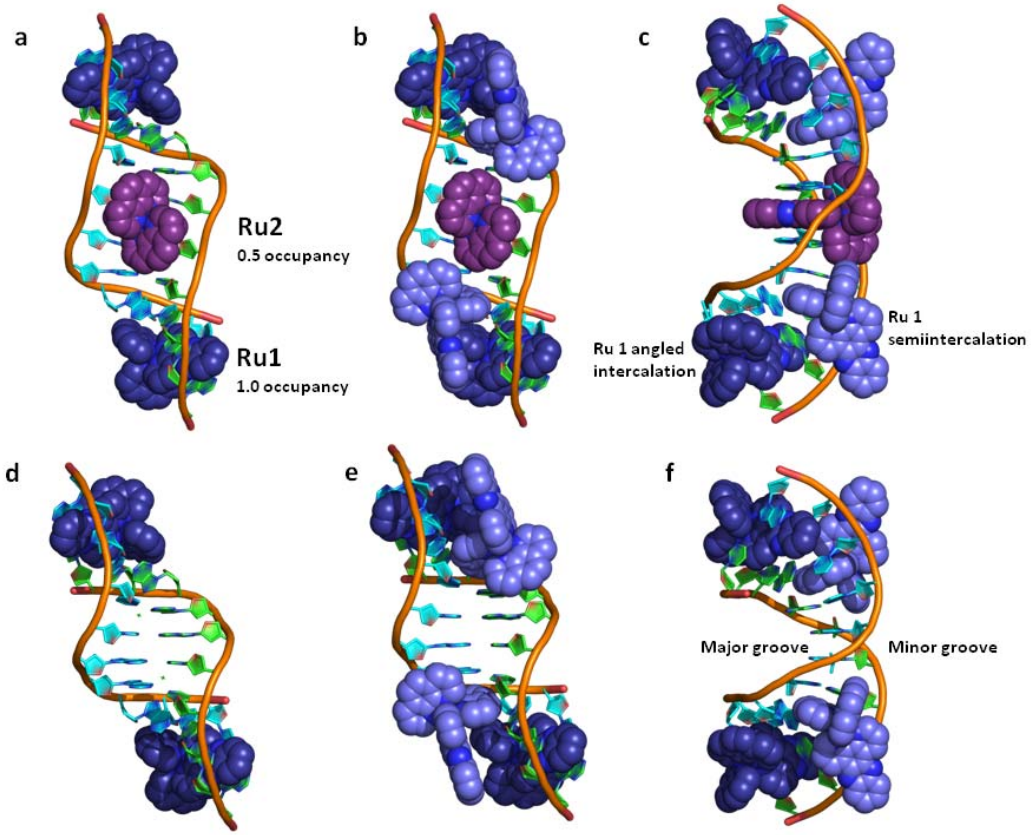


Figure 2 | The assembly of the duplexes. (a) *The twofold rotation axis generates the complete d(CCGG**TA**CCGG)$_2$ duplex with three bound cations; compare Figure 1c. Ru1 in dark blue, Ru2 in purple, both in spacefilling mode, DNA backbone (connecting P positions) orange, carbon: green, oxygen: red, nitrogen: blue, all in cartoon mode. The central Ru2 cation is treated as a whole complex of 0.5 occupancy per oligonucleotide strand, and the view direction is down the crystallographic twofold axis into the minor groove.* ***(b)*** *The same duplex with the addition of the semiintercalated cations (shown in paler blue) from symmetry-related duplexes,* ***(c)*** *the same assembly rotated through 90° about the vertical helix axis. The twofold rotation axis is now in the horizontal direction As throughout (except from Fig 1), the two strands of each duplex are coloured green and cyan for clarity, but all are crystallographically equivalent.*
(d)-(f) *The corresponding duplex (note the wider groove), containing only intercalated Ru1 (dark blue) and semiintercalated Ru1 (light blue)**.** View directions and colour codes are the same.*

Description of the structures

Complex between 1 and d(CCGGTACCGG)

This self-complementary sequence crystallizes to give a symmetrical duplex with a stoichiometry of three cations of **1** per duplex, or 1.5 cations of **1** per decamer strand. All the nucleic acid strands are equivalent in the crystal lattice, with the packing shown in Figure 1a. The conformation of a single resulting oligomer strand is shown in Figure 1b, showing the numbering scheme, and in 1c to show the ligand binding sites. This structure is quite different from that found for d(CCGG**TA**CCGG) itself, which has previously been the study of detailed study[27, 28] because it spontaneously crystallises as the DNA Holliday junction in the presence of Group II metal cations[29]. The quality of the experimental map is shown in Figure 1d with a view of **1** from the minor groove of the duplex.

Two ruthenium cations (Ru1) are symmetrically equivalent, related by the twofold axis running through the centre of the duplex, with a third cation (Ru2) lying on the same axis and being bound at the central step of the duplex (T_5A_6/T_5A_6). The ruthenium atoms lie in the minor groove of the duplex, with a hydrated barium cation in the major groove. There is a large solvent channel, so that the intercalated dppz ligand is embedded within the DNA base stack, while the phen ligands remain solvent-accessible in the minor groove. The asymmetric unit and overall assembly with the bound ruthenium complexes are shown in two orientations in Figure 2. Panels (a)-(c) depict the d(CCGG**TA**CCGG)$_2$ duplex, to show the different locations of angled intercalation (C_1C_2/G_9G_{10}), symmetrical perpendicular intercalation (T_5A_6/T_5A_6), and semiintercalation (G_3G_4/C_6C_7).

The specificity of this symmetrical intercalation for the lambda enantiomer of **1** is striking. We wondered if by reversing the central step sequence, we might see an equally specific interaction of the delta enantiomer. We therefore decided to try crystallizing a sequence with

the central step reversed but, again, using the racemic complex in the crystallization conditions, making no assumptions about the binding of the lambda and delta enantiomers.

Complex between 1 and d(CCGGATCCGG)

The above self-complementary sequence crystallizes to give a symmetrical duplex with only two cations of **1** per duplex, or one per decamer strand, with all strands being equivalent in the crystal lattice. Previous studies have shown that this sequence crystallizes as the duplex[29] in the absence of the ruthenium complex. The key difference between this structure and the TA containing model is that there is no complex bound at the central step (A_5T_6/A_5T_6), and the geometry of the central step (A_5T_6/A_5T_6) is consistent with B-DNA (Figure 2). Panels (d)-(f) show the corresponding d(CCGG**AT**CCGG)$_2$ duplex, showing the absence of the symmetrical intercalation mode here. The groove width at the central step can be seen to be narrowed by intercalation, by comparing the two backbone conformations. The P-P distance across the central step is 18.5 Å in (d), compared to 15.6 Å in (a).

Intercalation site T_5A_6/T_5A_6 and Ru1 geometry in d(CCGGTACCGG)$_2$. This is the first direct observation of a symmetrical perpendicular intercalation by the dppz ligand in an octahedral metal complex (Figure 3). The dppz is deeply inserted so that the ruthenium atom is only about 5 Å from the helix axis. Although the duplex has an overall curvature caused by the semiintercalation at the G_3G_4 step, there is no curvature at this intercalation site. The other distinctive feature of this intercalation geometry is the high local basepair step twist, of nearly 40°. The deep insertion is central to the observed specificity of the intercalation (see Discussion), and a particularly important consequence of the intercalation geometry at this step is that the phen ligands are then unable to semiintercalate into a second duplex. A further consequence of this deep intercalation is that the dppz ligand protrudes from the base stack

into the hydrated major groove. A simple way to compare this symmetrical geometry with the angled geometry, to be described in the next section, is the use of a P-Ru-P triangle (Figure 3b). In this symmetrical geometry, both Ru-P distances are about 8.3 Å, with the P-P separation distinctively short at 15.6 Å, and correlating with the high twist. The buckle at this step is 17°, with an opening of 7.3° and a propeller twist of 13.2° to accommodate the dppz chromophore, and maintaining a B-DNA conformation at these residues. The two hydrogen bonds at this step, giving a T_5(O4)-A_6(N6) distance of 2.97 Å, and a T_5(N3)-A_6(N1) distance of 2.86 Å, are close to normal values.

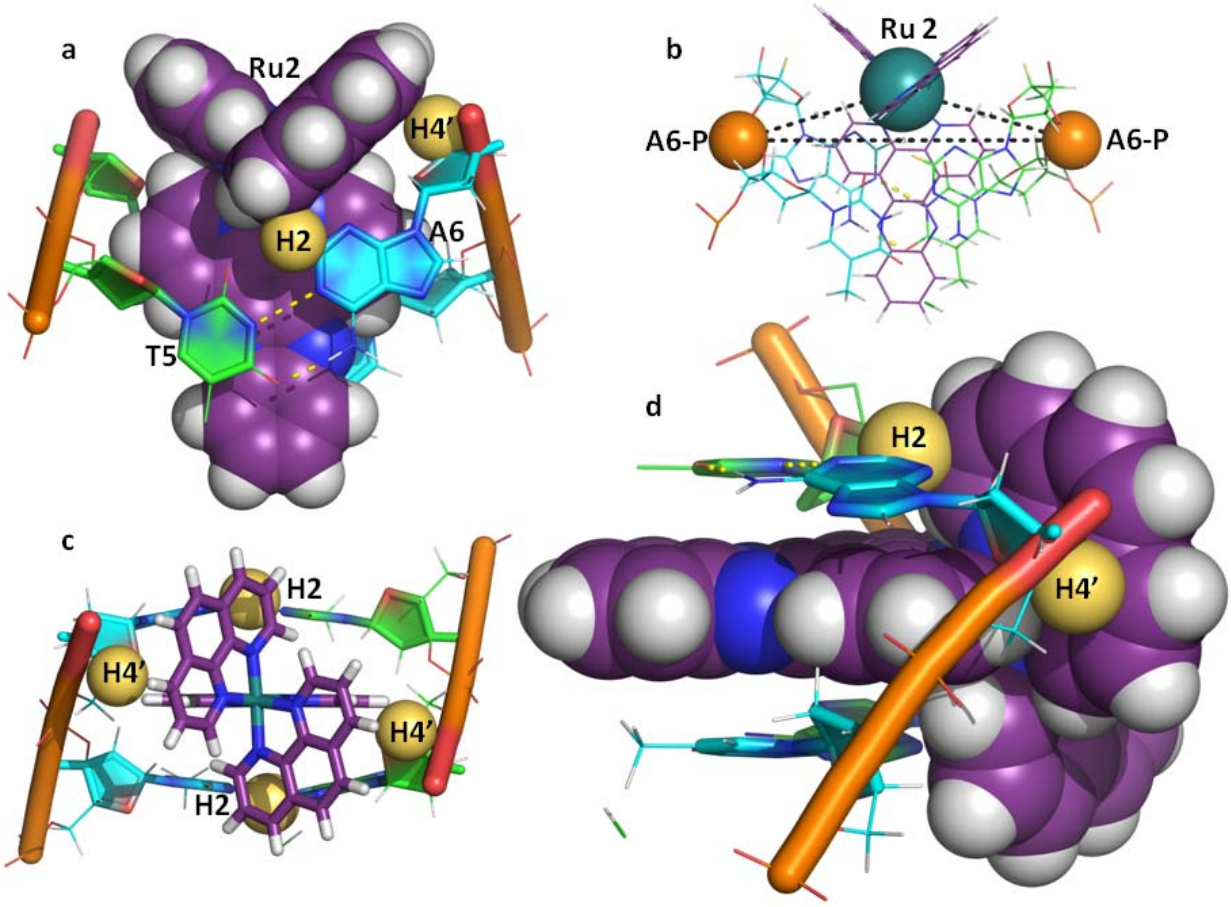


Figure 3. The symmetrical perpendicular intercalation geometry at the central TA/TA step. ***(a)*** *The Ru2 complex, shown in purple, intercalated between the T5-A6 basepairs. The twofold rotation axis runs vertically down the middle, so that the view from the other side is identical, although the two strands are coloured in green and cyan for clarity. The T_5-A_6 hydrogen bonds are shown by yellow dashes. The phenanthroline ring contacts the adenine*

H2 (on C2) and H4' (on C4' of the deoxyribose) as shown by rendering just these two hydrogen atoms in yellow space-filling mode. ***(b)*** *The use of a P-Ru-P triangle to define the intercalation geometry, using the same projection as in a). The P-P distance is short, at 15.6 Å, and the two Ru-P distances are 8.3 Å. Ruthenium atom in teal, phosphorus atoms in orange* ***(c)*** *View down the twofold axis from the minor groove side. In this orientation it is clear that the symmetrical orientation is maintained by the four contacts highlighted in a).* ***(d)*** *A sideways view, to show the extent of exposure of the dppz ligand in this symmetrical orientation. Note the projection of the dppz ligand into the major groove, also evident in (a).*

Terminal intercalation site G_9G_{10}/C_1C_2. Both these structures contain a second ruthenium cation, Ru1, bound by intercalation from the minor groove at the ends of the decamer duplex (Figure 4). This is a GG/CC site, and has the same geometry in both structures. Notably, only the guanine side of the base pair interacts fully with the complex, although both basepairs are complete. The terminal guanosine has adopted a *syn*-conformation, perhaps to maximise the stacking interaction between the purine ring and the dppz ligand, and giving a Hoogsteen-type base pairing. (Figure 4b). Although the C_1 cytosine is only partly ordered, there remains a single hydrogen bond to this *syn* guanine G_{10}. In this cavity, the ruthenium atom is about 6 Å from the helix axis, with the purine six-membered ring stacking onto the pyrazine ring of the dppz chromophore (Figure 4a). It appears that the guanine 2-NH_2 substituent limits the depth to which the dppz can intercalate, making for the shallower intercalation. The angled geometry shown in Figure 4c can be contrasted with the symmetrical geometry of Figure 3b, and here, the closest Ru-P distance is 7.9 Å.

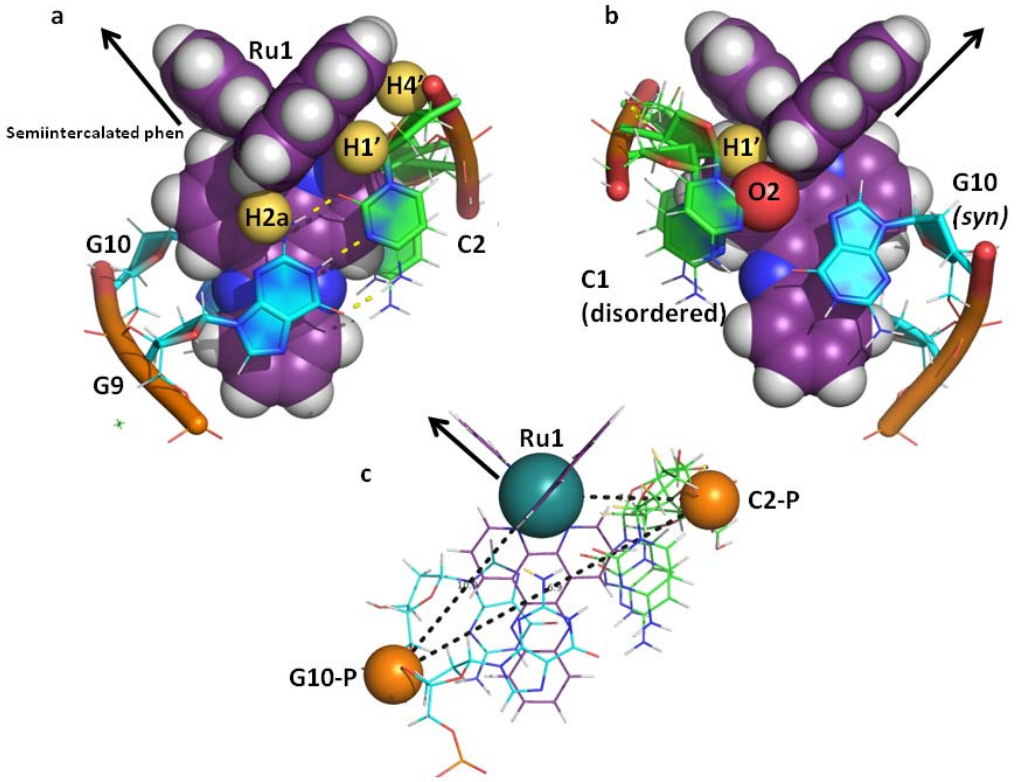


Figure 4 | The CC/GG angled intercalation geometry.(a) *The C_2-G_9 basepair, showing the stacking between the six-membered ring of the purine base and the pyrazine ring of the dppz ligand. The contacts between the phenanthroline ligand and residue C_2 are highlighted as H1' and H4' of the deoxyribose, rendering the hydrogen atoms only as yellow spheres. H2a of the G_9 residue (the 2-amino substituent unique to guanine) is similarly highlighted. These contacts define the phen orientation, and the angled geometry, leaving the other phen ligand free, and in this lattice it semiintercalates into a symmetry-related duplex at the G_3-G_4 step, in the direction of the arrow;* ***(b)*** *the view from the other side, in this case different, and showing the C_1-G_{10} step. The combination of disordered cytosine and syn-guanine mean that this is not a standard basepair (closer to Hoogsteen-type) and probably should not be regarded as having the same influence on ligand orientation as the very well defined C_2-G_9 step. The arrow again shows the direction of the semiintercalated interaction with the symmetry-related duplex;* ***(c)*** *the angled geometry as a P-Ru-P triangle. The shorter Ru-P distance is here 7.9 Å, the longer one 10.9 Å, with a P-P separation of 16.9 Å (compare the very short P-P separation in Figure 3)*

The GG/CC semiintercalation site. The angled intercalation mode, seen with both d(CCGG**TA**CCGG)$_2$ and d(CCGG**AT**CCGG)$_2$, creates asymmetry with respect to the phenanthroline interactions, one forming several close contacts, while the other is free, as shown by the arrow in Figure 4. A feature of the packing seen here is that this free phen is able to semiintercalate into a second duplex at the G_3-G_4 step in both structures (Figure 5a and b). This phen ligand forms a series of close contacts within the semiintercalation cavity, as highlighted by the yellow spheres of Figure 5a, and it is not possible to single out one specific interaction. A comparable finding was the highlight of our earlier study of the d(TCGGCGCCGA)$_2$ duplex bound to **2**[25], with semiintercalation of the TAP chromophore.

In that work, we were unable to obtain any crystals with phen, so it is important to emphasise that semiintercalation, well known from solution studies, but not previously from crystallography, has now been seen for the phen ligand also. The crosslinking of two duplexes which results, and which generates the overall packing of Figure 1, is shown in Figure 5c, which shows all the contacts made by a single angled ruthenium cation in this enantiospecific packing. The geometry at this step is close to that previously observed, with the proviso that there is a closer stacking of the phen chromophore onto the G_4 purine ring, with the angle between the G_3 purine plane and the phen plane opened up to 43°. It is this step which causes the overall curved appearance of the bound duplex, in fact due to the sharp kink at this point. The kink also creates a favourable binding site for the barium ion, which is coordinated to G_3, G_4 and T_5, as well as five water molecules, in a curved but strongly hydrated major groove (Supplementary Figure S2).

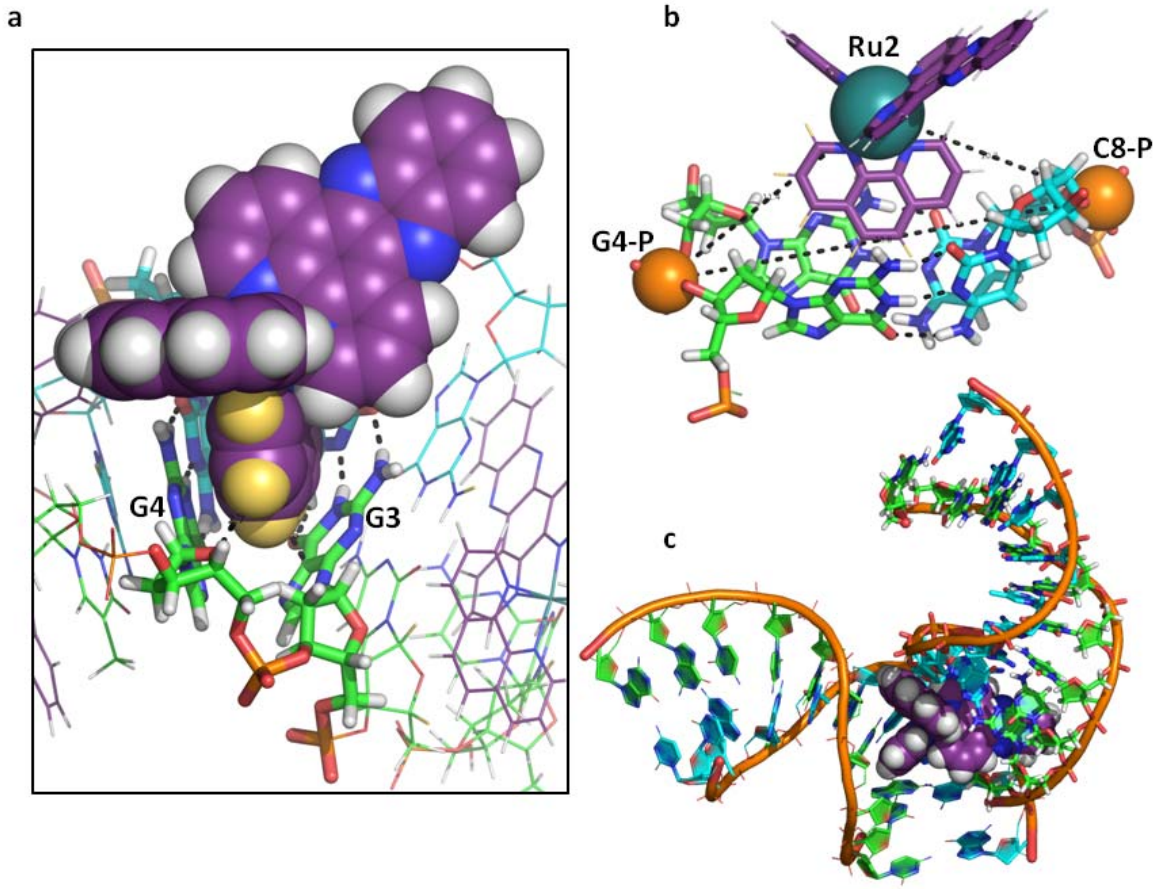


Figure 5 | Semiintercalation of one phenanthroline ligand. ***(a)*** *Semiintercalation into the G_3G_4/C_6C_7 step of the d(CCGG**AT**CCGG)$_2$ duplex. There is no single short contact; rather, the phen surface highlighted by colouring the hydrogens yellow contacts the surfaces of the two purine rings at G_3 and G_4 on either side.* ***(b)*** *the semiintercalation geometry represented as a P-Ru-P triangle for comparison with Figures 3b and 4c. The shorter Ru-P distance is 10.2 Å (to C_8-P) and the longer one (to G_4-P) is 11.4 Å, with the P-P spacing lengthened to 18.8 Å. The dppz ligand is on the right.* ***(c)*** *the overall environment of the ruthenium cation* ***1*** *both intercalated and semiintercalated into two duplexes of d(CCGG**AT**CCGG)$_2$.*

Discussion

This work provides the first structural evidence that the DNA 'light-switch' complex **1** binds to an oligonucleotide by both symmetrical (perpendicular, or 'head-on'), and by angled (canted) intercalation modes, and in each case through the minor groove of B-DNA. Angled intercalation is here combined with semiintercalation of one phen ligand.

The symmetrical dppz intercalation mode seen at the central TA/TA **s**tep leaves the helix straight, the ruthenium atom close to 5 Å from the helix axis, and as B-DNA, with only minor buckling of the TA basepairs. The enantiomeric and step specificity of this intercalation geometry is striking. Figure 3 provides a possible starting point to understand the structural features which may be responsible. What we see is that the phen ligands pack orthogonally directly against the adenosine residue, with the closest contact the H4' of the sugar ring, making the two closest contacts this H4' and the adenine 2C-H. These two hydrogens are shown as yellow spheres in Figure 3, which, by virtue of the symmetry is actually four such contacts. Figure 3a shows that they occupy strategic locations at either end of the phen ligand. The overall geometry and high twist angle at this site appear to be dictated by these two contacts, and it is possible that a weak attraction between the phenanthroline system and the ribose hydrogen H4' is responsible for the high twist. These observed contacts, to the adenine 2C-H and to the C4'-H of ribose, contrast the rigidity of the phenanthroline with the flexibility of the TA/TA step. This flexibility is also highlighted by the unusually short P-P separation of 15.6 A, compared to 16.9 Å at the angled intercalation site, 18.8 Å at the semiintercalation site, and 18.5 Å at the corresponding unbound AT/AT step. The result is a deep symmetrical intercalation, specific to the TA/TA step.

In contrast, the angled shallower intercalation seen at the terminal G_9-C_2/G_{10}-C_1 step shows that the helix axis (the centre of the basepair) is displaced, to the rim of the dppz chromophore. This asymmetry can again be contrasted with the symmetrical location of the axis seen with the T_5A_6/T_5A_6 basepair. Additionally, the twist angle between the basepairs is only 26°, compared to the 41° seen at the T_5A_6/T_5A_6 step. These differences are also seen in the P-Ru-P triangle (Figure 4c), which highlights the asymmetry of this angled intercalation, exposing one face of the dppz ligand to solvent but enclosing the other. The shorter Ru-P distance is now only 7.9 Å. This geometry permits the semiintercalation of a phen ligand into the minor groove of a G_3G_4/C_7C_8 step of a symmetry related duplex. This noncovalent crosslinking is only possible because of the shallow intercalation, lower twist, and purine stacking at this step. In turn, this is a selectivity which is only possible because intercalation is from the minor groove side. From the major groove, guanine and adenine are sterically very similar. This form of base selectivity has been studied at length for minor groove binding molecules, but to the best of our knowledge has not been seen for well characterised intercalative binding. An examination of the phen-nucleotide contacts reveals that, as with the symmetrical intercalation, there are contacts between the ribose ring and the phen ligand. These short contacts are shown in Figure 4a and b.

Important comparisons can be made with Actinomycin D and daunomycin, both of which also intercalate from the minor groove. In many literature examples[30], the nucleic acid duplex is formed from a hexanucleotide or smaller, with intercalation at the terminal steps. Actinomycin D is the best example of minor groove intercalation into a non-terminal site, but differs in generating an A-DNA conformation, in showing parallel intercalation of the chromophore at a GC/GC step, and forming a range of specific minor groove interactions. The twist angles here are correspondingly very low. The perpendicularly bound daunomycin

family has also been the subject of many studies, all of which show minor groove binding to a terminal CG step of a symmetrical hexameric duplex, with a twist of about 36° in all the examples. The symmetrical intercalation mode of **1** is at one end of the known range, combines high twist, low P-P separation and deep perpendicular intercalation. In contrast, the intercalation of Δ-$[Rh(bpy)_2chrysi)]^+$ (chrysi= chrysene-5,6-quinone diimine) into a TA/TA step from the **major** groove gave a similar M-helix axis distance as found for **1**, both about 5 Å, but this example has a twist angle of less than 31°, with no interaction between the bipy ligands and the DNA duplex[31].

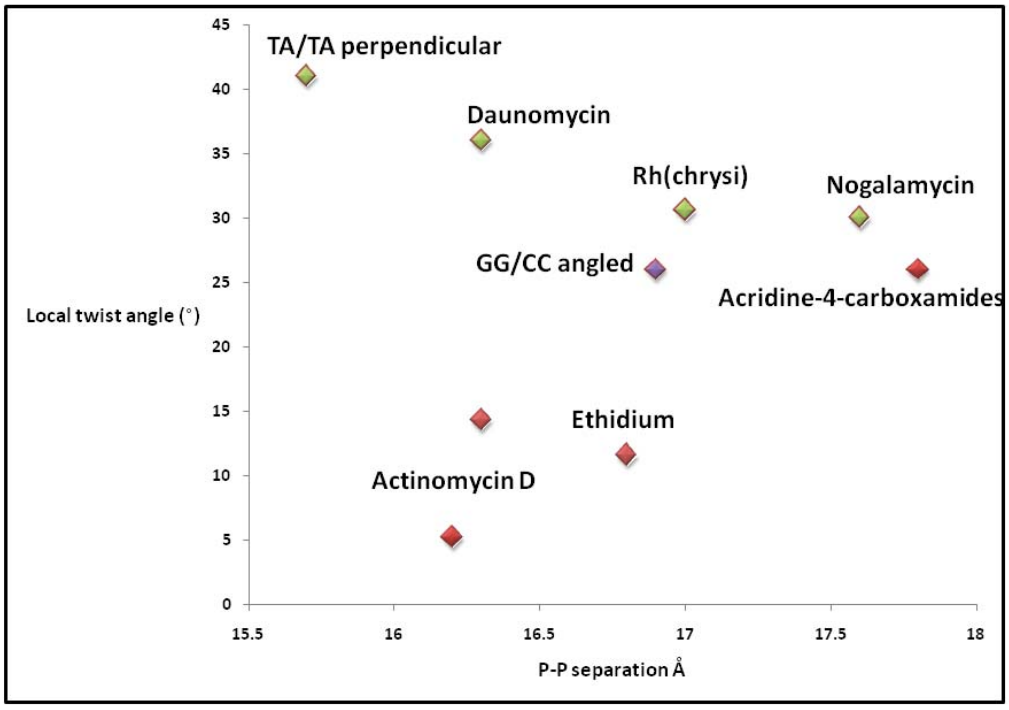


Figure 6 | The relationship between intercalation geometry – parallel, perpendicular or angled chromophore, and the P-P separation across the cavity. *The symmetrical intercalation in this work gives rise to the highest known local twist and the closest known P-P separation. No other data on angled intercalation is available. Green – perpendicular intercalation, red – parallel intercalation, purple – angled intercalation. Data are plotted for daunomycin[7], Actinomycin D [8], $[Rh(bpy)_2(chrysi)]+$[31], nogalamycin[32], an acridine-4-carboxamide derivative [4], and ethidium[2]. There are two points for Actinomycin D (PDB code 1MNV) which occupies the low twist/parallel intercalation extreme, associated with the many specific minor groove interactions. For determination of the twist angles in this work, see the*

Supplementary Information. Our survey of Nucleic Acid Database entries for intercalation, from which these examples are selected, is included as Supplementary Information, Dataset 1.

Although there are clear parallels with daunomycin with respect to chromophore orientation, a striking difference is in the origin of the high binding constant of **1** with natural DNA[24]. The daunomycin molecule can interact with DNA in many more diverse ways than **1**, and in the crystal structure forms two specific short hydrogen bonds to an adjacent guanine base, along with several specific hydrogen bonds to surrounding water molecules. Each of these contributes to the overall energetic, dominated by the favourable enthalpy term. In contrast, the ruthenium complex **1** intercalates without the formation of any specific bonding interaction with either DNA or water other than π-stacking of the dppz ligand. Thus the intercalation releases both bound water and cations from the oligonucleotide, particularly in the minor groove. In this groove the phen ligands displace all the ordered water.

The origin of the 'light-switch' phenomenon has been the subject of many investigations[17,18,24,25], and detailed photophysical measurements[33-35] have shown that this is due to the formation of a dark state in aqueous solution, and emission from the bright state when bound to DNA. Another feature arising from time-resolved emission studies of the complexes when bound to DNA is the presence of two quite distinct emission lifetimes. The use of enantiopure complexes and homogeneous polynucleotides ruled out sequence heterogeneity as the cause, although it was found that the proportion of the two lifetimes was very sensitive to the 'loading' of the complex on the polynucleotide[25]. Two main proposals which have been made are (i) different orientations within the major groove[22] or (ii) isolated or contiguous binding in the minor groove[18]. Neither of these proposals is consistent with our observations. Recent work in the Lincoln laboratory points to an alternative explanation – namely cooperative and anti-cooperative interactions[36] in the minor groove as being the origin of the two lifetimes, with the data only being fittable on the assumption that there are

two orientations in the intercalation pocket, with the shorter lifetime arising from the symmetric perpendicular intercalation mode we see for the first time here. The rationale for this interpretation would be that only one of the un-coordinated dppz nitrogen atoms can form a hydrogen bond to a solvent water molecule in an angled intercalation, but that both can do so in a symmetrical intercalation. As we accumulate crystallographic data on these complexes, we believe that we will be able to carry out a detailed analysis of bound water, which is likely to depend on the ligands employed in the study.

In summary, this study reports two distinct modes of intercalation of the DNA-light-switching complex $[Ru(phen)_2(dppz)]^{2+}$. It also demonstrates that the intercalation proceeds with changes in the DNA conformation, which are similar to those found in the crystal structure of certain classical DNA intercalators (e.g. daunomycin) but different from others (e.g. actinomycin D), and illustrates that there is still much to learn about the detailed geometrical requirements of DNA intercalation. The prevailing view of intercalation in the wider literature is often coloured by the predominance of parallel intercalation geometries in the early work. We have observed two intercalation modes, neither of which fit the parallel pattern: symmetric (perpendicular) and angled (canted), clearly distinguished by both sequence preference and by well-defined geometric criteria, and which correlate with the best solution data. The unambiguous observation of both these interactions significantly contributes to our understanding of this important class of synthetic intercalators.

Methods

Preparation and Analysis of Crystals Containing CCGGTACCGG

Crystals containing d(CCGGTACCGG)2 were grown by vapour diffusion, from sitting drops, at 276 K. A crystal of approximate dimensions 30x30x30 μm was mounted on a MicroMesh (MiTeGen) and data were collected on I04 at Diamond Light Source Ltd. Data were processed using xia2[37], with XDS (X-ray Detector Software)[38] and SCALA (continuous scaling program)[39], giving 2908 unique reflections with a resolution of 2.10 Å. Initial model building was performed using Coot (Crystallographic Object-Oriented Toolkit)[40] and the model was refined using REFMAC5.6 (maximum likelihood refinement) [41] from the CCP4 (Collaborative Computational Program 4) [42] suite. Five percent of reflections were used for the R_{free} test. The model gave a final R_{cryst} of 0.17 and R_{free} of 0.23. The structure, along with experimental data, are deposited in the Protein Data Bank with ID 3U38. For data collection and refinement statistics and full experimental details, including initial phasing with the maximum likelihood phasing programs SHELXC/D/E[43]; see SI Text T1.

Preparation and Analysis of Crystals Containing CCGGATCCGG

Crystals containing $d(CCGGATCCGG)_2$ were grown by vapour diffusion, from sitting drops, at 276 K. A crystal of approximate dimensions 200x200x100 μm was mounted on a MicroMesh (MiTeGen) and data were collected on I02 at Diamond Light Source Ltd. Data were processed using xia2[37], with XDS[38] and SCALA[39], giving 4747 unique reflections with a resolution of 1.70 Å. Initial model building was performed using Coot[40] and the model was refined using REFMAC5.6[41] from the CCP4[42] suite. Five percent of reflections were used for the R_{free} test. The model gave a final R_{cryst} of 0.15 and R_{free} of 0.18. The structure, along with experimental data, are deposited in the Protein Data Bank with ID 4E95. For data collection

and refinement statistics and full experimental details, including initial phasing with SHELXC/D/E[43]; see SI Text T2.

***Acknowledgements** We thank Dr Ger Doorley (formerly of Trinity College Dublin) for preparing the ruthenium complex, and Professor Per Lincoln (Chalmers University, Gothenburg, Sweden) for reading a draft of the manuscript, very helpful discussions, and making available his unpublished observations. HN is funded by a joint ILL/ESRF studentship and the Chemistry Department, University of Reading. JPH is funded by Diamond Light Source and a University of Reading University studentship. We acknowledge additional financial support from The Royal Society, The Royal Irish Academy and Science Foundation Ireland (Grant 06/RF/CHP035).

Author contributions: C.J.C., J.P.H. and J.M.K. conceived and designed the experiments; H.N., J.P.H., K. O'S. and C.J.C. performed the experiments, G.W. and T.S. contributed analytical tools, J.P.H., H.N. and C.J.C. analysed data; C.J.C. wrote the paper, with contributions from J.M.K. , J.P.H. and H.N. All authors discussed the results and commented on the manuscript. H.N. and J.P.H. contributed equally to this work.

References

1. Lerman, L.S., Structural considerations in the interaction of deoxyribonucleic acid and acridines. *J. Mol. Biol.* **3***,* 18-30 (1961).

2. Shieh, H.-S., Berman, H.M., Dabrow, M. & Neidle, S., The structure of a drug-deoxydinucleoside phosphate complex; generalized conformational behaviour of intercalation complexes with RNA and DNA fragments. *Nucl. Acids Res.,* **8***,* 85-98 (1980).

3. Neidle, S. The molecular basis for the action of some DNA-binding drugs, *Prog. Med. Chem*., **16,** 151-221 (1979).

4. Todd, A.K., Adams, A., Thorpe, J.H., Denny, W.A, Wakelin, L.P. & 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-A resolution. *J. Med. Chem.,* **42,** 536-540 (1999).

5. Hopcroft, N. H., Brogden, A. L., Searcey, M., & Cardin, C.J., X-ray crystallographic study of DNA duplex cross-linking: simultaneous binding to two $d(CGTACG)_2$ molecules by a bis(9-aminoacridine-4-carboxamide) derivative, *Nucl. Acids Res.* **34**, 6663-6672 (2006).

6. Brogden, A.l., Hopcroft, N.H., Searcey, M., & Cardin, C.J., Ligand Bridging of the DNA Holliday Junction: Molecular Recognition of a Stacked-X Four-Way Junction by a Small Molecule. *Angew. Chem. Intern. Edn.,* **119**, 3924-3928 (2007).

7. Frederick, C. A. *et al*. Structural comparison of anticancer drug-DNA complexes: adriamycin and daunomycin. *Biochemistry*. **29,** 2538-2549 (1990).

8. Hou, M.-H., Robinson, H., Gao, Y.-G., Wang, A.H.-J. Crystal structure of actinomycin D bound to the CTG triplet repeat sequence linked to neurological diseases, *Nucl. Acids Res.*, **30,** 4910-4917.

9. Garbett, N.C. & Chaires, J.B., Biophysical tools for Biologists: Vol. 1 In vitro techniques, *Methods in Cell Biology,* **84**, 3-23 (2008).

10. Wheate, N.J., Brodie, C.R., Collins, J.G., Kemp, S., Aldrich-Wright, J.R., DNA intercalators in cancer therapy: organic and inorganic drugs and their spectroscopic tools of analysis, *Mini Reviews Med. Chem.*, **7,** 627-648 (2007)

11. Dai, J., Punchihewa, C., Mistry, P., Ooi, A.T., Yang, D., Novel DNA bis-intercalation by MLN944, a potent clinical bisphenazine anticancer drug., *J. Biol. Chem.,* 279, 46096-46103 (2004).

12. Zeglis, B. M., Pierre, V. C. & Barton, J. K. Metallo-intercalators and metallo-insertors. *Chem. Commun.* **44,** 4549-4696 (2007).

13. Liu, H.-K. & Sadler, P.J., Metal complexes as DNA intercalators. *Acc. Chem. Res.,* **44**, 349-359 (2011).

14. Boer, D. R., Canals, A. & Coll, M., DNA-binding drugs caught in action: the latest 3D pictures of drug-DNA complexes. *Dalton Trans.,* **3**, 399-414 (2009

15. Chambron, J.C., Sauvage, J.P, Amouyal, E. & Koffi, P. $Ru(bipy)_2(dypyridophenazine)^{2+}$: a complex with a long range directed charge transfer excited state. *Nouv. J. Chim.***9**, 527-9 (1985).

16. McKinley, A. W., Lincoln, P. & Tuite, E. M. Environmental effects on the photophysics of transition metal complexes with dipyrido[2,3-1:3',2'-c]phenazine (dppz) and related ligands. *Coord. Chem. Rev.* **255,** 2676-2692 (2011).

17. Friedman, A. E. *et al*. A molecular light switch for DNA: $Ru(bpy)_2(dppz)^{2+}$. *J. Am. Chem. Soc.* **112,** 4960-4962 (1990).

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

19. Smith, J. A., George, M. W. & Kelly, J. M. Transient spectroscopy of dipyridophenazine metal complexes which undergo photo-induced electron transfer with DNA. *Coord. Chem. Rev.* **255,** 2666-2675 (2011).

20. Elias, B., *et al.* Photooxidation of guanine by a ruthenium dipyridophenazine complex intercalated in a double-stranded polynucleotide monitored directly by picosecond visible and infrared transient absorption spectroscopy, *Chem. Eur J.***14**, 369-375.

21. Barton, J. K., Olmon, E. D. & Sontz, P. A. Metal complexes for DNA-mediated charge transport. *Coord. Chem. Rev.* **255,** 619-634 (2011).

22. Hartshorn, R. M. & Barton, J. K. Novel dipyridophenazine complexes of ruthenium(II): Exploring luminescent reporters of DNA. *J. Am. Chem. Soc.* **114,** 5919-5925 (1992).

23. Lincoln, P, Broo, A., and Nordén, B. Diastereomeric DNA-binding geometries of intercalating ruthenium (II) trischelates probed by linear dichroism: $[Ru(phen)_2dppz]^{2+}$ and $[Ru(phen)_2bdppz]^{2+}$. *J. Amer. Chem. Soc.* **118,** 2644-2653.

24. Haq, I. *et al*. Interaction of Δ- and Λ-$Ru(phen)_2dppz]^{2+}$ with DNA: A calorimetric and equilibrium binding study. *J. Am. Chem. Soc.* **117,** 4788-4796 (1995).

25. Hall, J. P. *et al.* Structure determination of an intercalation ruthenium dipyridophenazine complex which kinks DNA by semiintercalation of a tetraazaphenanthrene ligand. *Proc. Natl. Acad. Sci. USA.* **108,** 17610-17614 (2011).

26. Tuite, E., Lincoln, P. & Nordén, B. Photophysical evidence that Δ- and Λ-$[Ru(phen)_2(dppz)]^{2+}$ intercalate DNA from the minor groove. *J. Am. Chem. Soc.* **119,** 239-240 (1997).

27. 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. Cryst. D.* **58,** 567-569 (2002).

28. Eichman, B. F., Vargason, J. M., Mooers, B. H. M. & Shing Ho, P. The Holliday junction in an inverted repeated DNA sequence: Sequence effects on the structure of four-way junctions. *P. Natl. Acad. Sci. USA*. **97,** 3971-3976 (2000).

29. Hays, F. A., Teegarden, A., Jones, Z. J. R., Harms, M., Raup, D., Watson, J., Cavaliere, E. & Shing Ho, P. How sequence defines structure: A crystallographic map of DNA structure and conformation. *P. Natl. Acad. Sci. USA*. **102,** 7157-7162 (2005).

30. Neidle, S. Principles of Nucleic Acid Structure, Academic Press, 2007, pp144-158.

31. Pierre, V. C., Kaiser, J. T. & Barton, J. K. Insights into finding a mismatch through the structure of a mispaired DNA bound by a rhodium intercalator. *P. Natl. Acad. Sci. USA*. **104,** 429-434 (2007).

32. Smith, C. K., Davies, G.J., Dodson. E. J. & Moore, M.H. DNA-Nogalamycin interactions: the crystal structure of d(TGATCA) complexed with nogalamycin. *Biochemistry,* **34,** 415-425 (1995).

33. Brennaman, M.K., Meyer, T.J., Papanikolas, J.M., $[Ru(bpy)_2dppz]^{2+}$ light-switch mechanism in protic solvents as studies through temperature-dependent lifetime measurements. *J. Phys. Chem. (A),* **108**, 9938-9944, (2004).

34. Önfeld, T., Olofsson, J., Lincoln, P., & Nordén, B., Picosescond and steady state emission of $[Ru(phen)_2dppz]^{2+}$ in glycerol: anomalous temperature dependence. *J. Phys. Chem (A)*, **107**, 1000-1009, (2003).

35. Olson, E.J.C., *et al.* First observation of the key intermediate in the 'light-switch' mechanism of $[Ru(phen)_2dppz]^{2+}$. *J.Amer. Chem., Soc.,* **119**, 11458-11467, (1997).

36. Lincoln, P., A generalised McGhee-von Hippel method for the cooperative binding of different competing ligands to an infinite one-dimensional lattice. *Chem. Phys. Lett.,* **288,** 647-656 (1998).

37. Winter, G. Xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43,** 186-190 (2010).

38. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26,** 795-800 (1993).

39. Evans, P. Scaling and assessment of data quality. *Acta. Cryst. D.* **62,** 72-82 (2006).

40. Emsley, P., Lohkamp, B., Scott, W. & Cowtan, K. Features and development of Coot. *Acta. Cryst. D.* **66,** 486-501 (2010).

41. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Cryst. D.* **53,** 240-255 (1997).

42. Collaborative computational project, number 4. The CCP4 suite: programs for protein crystallography. *Acta Cryst. D.* **50,** 760-763 (1994).

43. Sheldrick, G. M. A short history of SHELX. *Acta Cryst. D.* **64,** 113-122 (2008).

TOC legend (suggested)

DNA sensing and photodynamic therapy require specific recognition. Three binding modes of a ruthenium ‘light-switch’ complex with DNA are here seen by X-ray crystallography. The deep symmetric intercalation from the minor groove at the TA/TA step is now thought to correspond to the shorter of two photochemical lifetimes.