Structural analysis of the binding of the diquaternary pyridophenazine derivative dqdppn to B-DNA oligonucleotides†

Philip Waywell,[a,b] James A. Thomas*[a] and Mike P. Williamson*[b]

Received 3rd September 2009, Accepted 31st October 2009
First published as an Advance Article on the web 7th December 2009
DOI: 10.1039/b918252g

The interaction of the ethylene-bipyridyldiylium-naphthaphenazine dication, dqdppn, with several hexa- and octanucleotide duplexes has been studied using CD and NMR. Taken together, these studies reveal that with the hexanucleotide, dqdppn intercalates into the terminal base pair, and causes a large twisting of the terminal base pair. In contrast, with all three octanucleotides, dqdppn intercalates more centrally within the sequence. The NMR-derived structures of two of the binding complexes demonstrate that dqdppn intercalates from the major groove in an unusual 'side-on' geometry, rather than threading through the helix. An analysis of these results indicates that the preferred binding site is not sequence-specific, but primarily at the most conformationally flexible DNA step.

Introduction

DNA metallointercalator complexes based on dipyrido[3,2-*a*:2′,3′-*c*]phenazine (dppz), such as $[Ru(bpy)_2(dppz)]^{2+}$ and $[Ru(phen)_2(dppz)]^{2+}$ (Fig. 1), have interesting photophysical properties: in aqueous solution they display virtually no luminescence, but when bound to DNA their emission is enhanced by several orders of magnitude.[1,2] The exact mechanism of this 'light switch' effect is still under investigation but it clearly involves subtle interplay between several excited states.[3] Variable temperature studies in aprotic solvents by Brennaman *et al.* led to the hypothesis that excitation can populate an entropically-favored "light" state which is analogous to the well characterized ^{3}MLCT found for the $[Ru(bpy)_3]^{2+}$ cation, or a more charge separated, enthalpically favored "dark" state centered on the phenazine fragment of dppz.[4] A subsequent theoretical study reported by Pourtois *et al.* indicated that the system possesses a ligand-centered triplet state, ^{3}LC that may be non-emissive (but, not it was suggested, the dark state in water) and that emission occurs when a close-lying ^{3}MLCT is thermally occupied.[5] Further experimental and theoretical studies by Önfelt and Lincoln also revealed anomalous temperature effects in hydroxylic solvents such as glycerol and ethylene glycol. These data were interpreted as being a consequence of an equilibrium involving three excited states that differ in the extent of solvent hydrogen-bonding, with the double hydrogen-bonding state being the dark state.[6]

This switching effect has led to a wide range of applications for these complexes, including components of DNA sensors,[7] and tools to probe the electron[8] and energy[9] transfer properties of the DNA double helix. It has been shown using a range of spectroscopic (including NMR) and calorimetric methods that these and closely related complexes intercalate into DNA.[2,10-20] However, the binding geometries of these systems are still somewhat unclear, with studies indicating different intercalation modes for different optical isomers[13,16,17] and intercalation from either the major or minor groove. It also remains a matter of debate exactly what the binding geometry is, and it is likely that the complexes bind in multiple geometries in equilibrium.[14,21,22] Given these uncertainties, and the difficulties in derivatizing a chiral kinetically inert complex, the design and characterization of more specific systems based on such architectures is often extremely demanding.[23]

Recently, we have reported on the DNA binding properties of purely organic ligands based on the dppz unit and its analogs.[24,25]

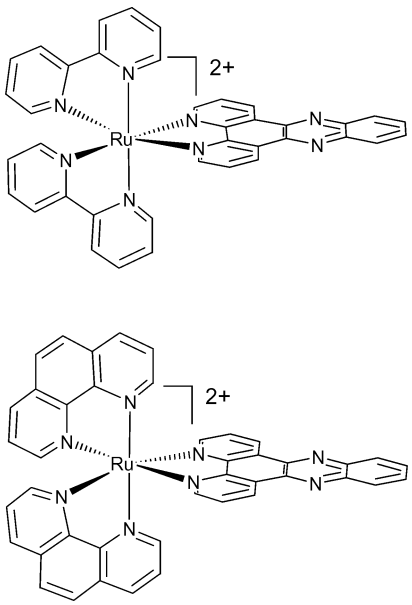

Fig. 1 The bisbipyridyl and bisphenanthroline ruthenium(II) complexes of dipyrido[3,2-*a*:2′,3′-*c*]phenazine, $[Ru(bpy)_2(dppz)]^{2+}$ (top) and $[Ru(phen)_2(dppz)]^{2+}$ (bottom) respectively.

[a]*Dept. of Chemistry, University of Sheffield, Brook Hill, Sheffield, S3 7HF, UK. E-mail: james.thomas@sheffield.ac.uk; Fax: +44 114 222 9346*
[b]*Dept. of Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK. E-mail: m.williamson@sheffield.ac.uk; Fax: +44 114 222 2800*

† Electronic supplementary information (ESI) available: ***S1***. The ^{31}P–^{1}H COSY spectrum of free $d(AGAGCTCT)_2$. ***S2***. ^{31}P NMR experiments showing the step-wise addition of dqdppn to $d(AGAGCTCT)_2$. ***S3***. An overlay of the 2D ^{1}H/^{31}P correlation spectra of free $d(AGAGCTCT)_2$. ***S4***. Intermolecular NOE contacts in the 2 : 1 complex formed by dqdppn and $d(GAGCTC)_2$. See DOI: 10.1039/b918252g

These compounds have the ruthenium(II) center replaced by an ethylene bridge, to give ligands such as dqdppz and dqdppn (Fig. 2). Like the parent ruthenium complexes, these ligands bind *via* intercalation; although unlike the parent[19] they bind with an apparent slight preference for GC-rich sequences in the case of dqdppz,[24] and a strong preference for poly(dG).poly(dC) over poly(dA).poly(dT) for dqdppn.[25] Unlike the ruthenium complexes, the organic cations are luminescent when free in aqueous solution. Experimental and theoretical data have revealed that this emission is due to an intramolecular charge transfer. The luminescence of dqdppz is quenched on addition of DNA, probably due to electron transfer processes involving photo-oxidation of the guanine and adenine bases.[24] The quenching of dqdppn depends on the composition of the DNA. Poly(dG).poly(dC) gives a significant reduction of luminescence, while poly(dA).poly(dT) shows a fourfold increase.[25] These observations are consistent with the lower oxidizing potential of the photo-excited state of dqdppn compared to that of dqdppz. Like the parent ruthenium complexes, no details are known of the structure of the intercalated cation–DNA complex, although it has been suggested to bind from the minor groove.

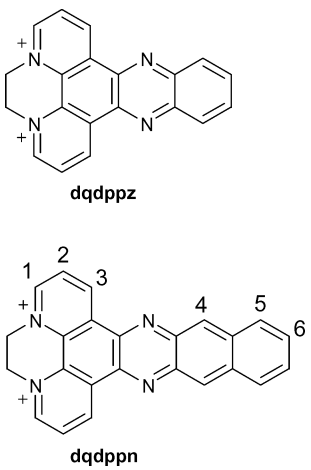

Fig. 2 Diquaternised dppz derivatives dqdppz and dqdppn, showing the numbering of aromatic protons in dqdppn.

Herein we report on the structure and mode of binding of dqdppn to several oligonucleotides, and show that it intercalates from the major groove in a non-standard 'side-on' geometry, parallel to the base pair long axis, a binding geometry that is related to previous suggestions for Ru(dppz) systems. We suggest that the organic cation recognizes, not a specific sequence, but a flexibility or deformability in the double helix. This also appears to be the requirement for a number of ligands that bind at operator and transcription factor sites on DNA, suggesting that dqdppn derivatives may have novel therapeutic applications.

Results and discussion

Studies were carried out on the self-complementary hexanucleotide $d(GAGCTC)_2$, and on three octanucleotides: $d(AGAGCTCT)_2$ and $d(CGAGCTCG)_2$, which contain the same hexanucleotide sequence flanked by either an AT base pair or a GC base pair; and $d(GCTATAGC)_2$.

Preliminary stability studies

To confirm that the oligonucleotides were duplex as opposed to single stranded under the experimental conditions a range of studies were carried out. The presence of dqdppn in general stabilised the duplex form, as – to some extent – did addition of salts, but by far the biggest effect on duplex stability was the concentration of oligonucleotide. Therefore in all the studies reported here, in order to maintain the DNA as a duplex, the concentration of nucleotide was at least 1 mM and usually approximately 2 mM.

Circular dichroism studies

As expected from previous studies, the duplexes have intense CD signals around 260 nm: indeed, at the high concentrations used in the binding experiments, the absorbance in this region saturated the detector and therefore no useful CD signals were obtained. Although B-DNA has no CD signal above 310 nm, when dqdppn binds to DNA the chiral DNA induces a CD signal in dqdppn. The induced CD (ICD) signal can be either positive or negative, depending on the orientation of the ligand transition relative to the DNA base pairs within the binding site. Theoretical calculations[26,27] have shown that a positive ICD signal is expected when the ligand transition is perpendicular to the long axis of the DNA base pairs, whereas a negative ICD signal is expected when the ligand transition is parallel to the long axis of the DNA base pairs. This behavior has been confirmed experimentally through the analysis of the binding of intercalating species such as 9-aminoacridine and 2,7-diazapyrene.[28,29]

Addition of dqdppn to the hexanucleotide $d(GAGCTC)_2$ results in a negative ICD signal at 353 nm, the intensity of which initially increases linearly with concentration, after which it is affected by self-association of the dqdppn and saturation of the detector (Fig. 3). The negative signal indicates that dqdppn is intercalating with its long axis parallel to the base pair long axis, 'side-on' into the DNA.

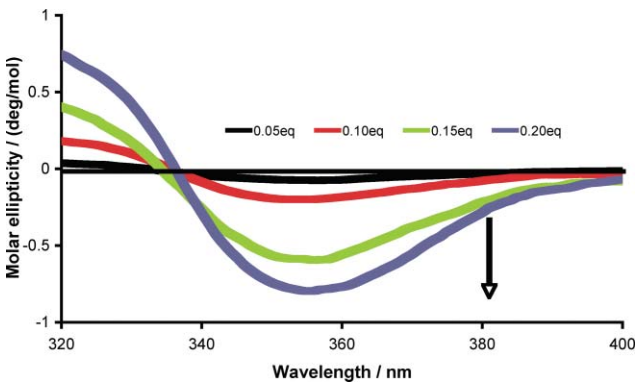

Fig. 3 Circular dichroism (CD) spectra produced during the step-wise addition of dqdppn to $d(GAGCTC)_2$.

Addition of dqdppn to the three octanucleotides $d(AGAGCTCT)_2$, $d(CGAGCTCG)_2$ and $d(GCTATAGC)_2$ gave similar results: a negative ICD signal at approximately 350 nm, which is linearly dependent on dqdppn concentration up to approximately 0.5 mM. This indicates that in all cases the

Table 1 Intermolecular NOE contacts between dqdppn and d(GAGCTC)$_2$ in the 1 : 1 complex

	d(GAGCTC)$_2$					
dqddpn	G1	A2	G3	C4	T5	C6
CH_2						
H1	4′, 5′, 5″, 2′/2″				3′	
H2		4′/5′/5″, 1′				4′/5′/5″
H3				1′		
H4	1′					
H5					H6, Me, 2′, 2″	H5, 1′, 2′, 2″, 3′
H6				1′	H6, Me, 2′, 2″	H5, 1′, 2′, 2″, 3′

cation binds with the long axis of dqdppn parallel to the base pair long axis.

Preliminary ^{1}H NMR studies of dqdppz binding to hexameric DNA

The binding of dqdppz to a range of hexameric oligonucleotide sequences was analyzed using 1D ^{1}H NMR experiments. In all cases the resultant spectra were severely broadened and proved to be non-assignable at all temperatures. Therefore, we chose to concentrate on the binding of dqdppn to a range of short oligonucleotides as the resultant NMR spectra for these systems proved to be assignable. Details of these studies are presented in the following sections.

NMR studies of the dqdppn : d(GAGCTC)$_2$ complex

NMR spectra of the oligonucleotide were assigned using standard ^{1}H homonuclear 2D methods, with a combination of NOESY, COSY and TOCSY experiments, in which sugar and base spin systems were analyzed using COSY and TOCSY spectra. Intraresidue and interresidue base–sugar connections were made using NOESY.[30] Chemical shift assignments have been submitted to BioMagResBank.[31] Addition of dqdppn to the oligonucleotide caused broadening of the spectrum due to slow to intermediate exchange, as frequently observed with DNA intercalators.[14,32] To some extent this was remedied by increasing experimental temperatures, but at too high a temperature duplex dissociation became a problem. Conversely, lowering temperatures too far produced signals that were too broad to assign. Thus, assignments for free oligonucleotides were carried out at 298 K, whereas measurements on the complexes were carried out at 303 K, the highest temperature at which the hexanucleotides were stable as duplexes.

Titration of dqdppn into the hexanucleotide d(GAGCTC)$_2$ gave gradual chemical shift changes. The spectra of both molecules were re-assigned using homonuclear experiments. At a 1 : 1 ratio of dqdppn to duplex, a single set of resonances was observed for both ligand and the self-complementary duplex, implying that the off-rate of the binding complex(es) is fast enough to lead to averaging of chemical shifts. A number of intermolecular NOEs could be observed (Table 1), which indicate that dqdppn is binding between the terminal base pairs, in an orientation parallel to the base pair axis, in confirmation of the CD results. Somewhat surprisingly, they also clearly imply binding from the *major* groove.

Further addition of dqdppn was conducted up to a 2 : 1 ratio of dqdppn to duplex; a 1 : 1 ratio of ligand to binding site – assuming an asymmetric off-center intercalation position – in the expectation that the spectrum would be sharpened because a fully bound species would be produced. However, the titration only led to further broadening, and to observation of NOEs from dqdppn to protons on all the bases of the hexanucleotide (Supporting Information). Protons from different parts of dqdppn also displayed NOEs to the same base protons, indicating that the dqdppn is bound in more than one orientation. It is therefore apparent that, far from giving a single bound species and improving the spectrum, addition of the second equivalent produces a large number of alternative binding geometries.

Therefore, NOEs from the 1 : 1 complex were used to calculate a structure, using a standard simulated annealing procedure within XPLOR. As expected, the resultant structures show dqdppn intercalated into the terminal base pair step, in an unusual geometry parallel to the base pair axis (Fig. 4). This binding mode results in a significant perturbation of the DNA structure at the terminal base pair, in which the base pair plane is almost orthogonal to its position in the starting structure. Such a binding mode is possibly favored because it maximizes π–π stacking interactions between the dqdppn and the bases of the nucleotides within the binding site, in particular with the penultimate AT base pair.

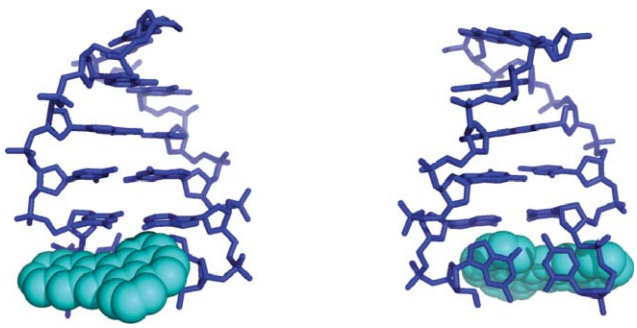

Fig. 4 Energy minimized structure of the 1 : 1 complex formed by dqdppn and d(GAGCTC)$_2$, shown in cyan and blue respectively, as calculated using XPLOR. Hydrogen atoms are omitted for clarity.

NMR studies of the complexes between dqdppn and octanucleotides

Since the dqdppn binds to the hexamer d(GAGCTC)$_2$ at the terminal step, the question arises whether this is due to a preference for a 5′-GA step, or just because the terminal location of the oligonucleotide is more easily distorted to accommodate an intercalating ligand. We therefore carried out studies on two further nucleotides d(AGAGCTCT)$_2$ and d(CGAGCTCG)$_2$, in

which the central hexanucleotide is capped by either a GC base pair or an AT base pair; in this way we aimed to determine the reason for the intercalation preference in the hexamer and to further characterize binding site preferences.

Spectra of free nucleotides were assigned, and dqdppn was titrated in. As seen with the hexamer, the spectra became broader on addition of dqdppn, necessitating an increase in temperature to 318 K. In this case, the 1 : 1 complex produced broad spectra with no useful NOEs, whereas the 2 : 1 complex gave a small number of assignable NOEs (Fig. 5). The nucleotide d(AGAGCTCT)$_2$ gave better spectra and more NOEs than d(CGAGCTCG)$_2$, as it showed less exchange broadening. The intermolecular NOEs indicate that the benzo-phenazine portion of the ligand, containing protons H4, H5 and H6 (Fig. 2), is located close to the 5′-C^5-T^6 base pair step, again approaching from the major groove.

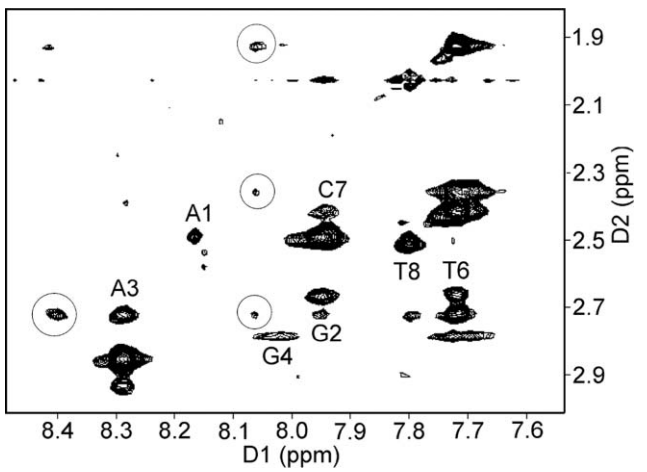

Fig. 5 A selected region from the 2D ^{1}H NOESY spectrum of the 2 : 1 complex formed by dqdppn and d(AGAGCTCT)$_2$. This region corresponds to the contacts between the base (H8 and H6) and sugar (H2′ and H2″) protons of the oligonucleotide. Labels indicate the aromatic protons of the bases, which have both intranucleotide and sequential NOEs. Several intermolecular NOE contacts are also evident in this region and are indicated by circles.

In order to define the binding site more closely, an analysis of the ^{31}P spectrum was carried out. ^{31}P chemical shifts of the backbone phosphates are sensitive to the dihedral angles around the phosphorus, and are therefore good indicators of structural change to the backbone, as expected for intercalative binding.[33–36] Although 1D phosphorus spectra are useful for monitoring a titration, they do not permit assignments. However, 2D ^{31}P–^{1}H COSY spectra are useful for assigning ^{31}P chemical shifts in nucleotides,[37,38] and connect the phosphorus nucleus to its intraresidue H3′ and interresidue H5′/H5″ neighbors (although in our case, correlations to H5′ protons were generally not assignable due to significant overlap between the individual cross peaks). The experiment was carried out in 100% D_2O to reduce the t_1 noise stripe from the residual solvent, which coincides with the H3′ protons. The titration is shown in Supporting Information and indicates that the largest shifts and exchange broadening are seen for the phosphorus nuclei in the A^3-G^4 and C^5-T^6 steps. In this self-complementary duplex, these two phosphates face each other on opposite strands in the C^5-T^6 step, providing strong evidence that the dqdppn intercalates into this site, and confirming the NOESY evidence.

Complexation of dqdppn with the octanucleotide d(GCTATAGC)$_2$ was accompanied by exchange broadening of NMR signals. Assignment of the 1 : 1 complex showed that the normal sequential internucleotide NOEs were of very low intensity in the T^5-A^6 step, suggesting that this is where the dqdppn has bound.[39,40] Only two intermolecular NOEs could be assigned in the 1 : 1 complex, from H4 of dqdppn to two of the H4′, H5′ or H5″ protons of A^4. Due to the position of these protons, these NOE contacts suggest that the likely binding site is the 5′-T^3-A^4 base pair step. This is in agreement with the observed break in the "NOE walk" between T^5 and A^6 step described above, as these nucleotides form the complementary base pairs with A^4 and T^3 respectively. Signal overlap and exchange broadening within the ^{31}P spectrum hindered an analysis of ^{31}P shifts, but they clearly showed that there is no intercalation at the central (A^4-T^5) base pair step.

Thus, for all three octanucleotides, dqdppn binds at the same position, namely the A^3-G^4 step (T^3-A^4 in d(GCTATAGC)$_2$), which by symmetry is equivalent to the C^5-T^6 step. However, only the d(AGAGCTCT)$_2$ complex gave enough observable NOEs to permit calculation of a structure.

Structural model of the d(AGAGCTCT)$_2$–dqdppn complex

Because of exchange broadening effects, there are only 6 measurable intermolecular NOEs in the 1 : 1 complex. These are between the dqdppn protons H4, H5 and H6 and protons in the 5′-C^5-T^6 base pair step: they therefore all occur at the same end of the ligand, and do not define the position of the bispyridyl end of dqdppn within the DNA helix. However, given the CD studies indicating a parallel orientation, the planar nature of the ligand, and the limited size of the 5′-C^5-T^6 step placing a severe limit on any possible binding geometry, a computational structure calculation within these constraints was carried out.

The initial structure (Fig. 6a) contained dqdppn intercalated into the 5′-C^5-T^6 step as expected, but the ligand made little van der Waals contact with the DNA bases, and did not appear to be in an energetically favorable position. On closer inspection, it was clear that this was due to the absence of any intermolecular NOE restraints involving protons H1, H2 and H3 of dqdppn, explained to a large extent by the exchange broadening of these protons (itself evidence that these protons are intimately involved in the complex).

The structure calculation was therefore repeated, with the addition of a small number of non-experimental intermolecular restraints. These additional restraints were equivalent to the intermolecular NOEs observed between H1 and H2 of dqdppn and H2″ of G^1 and H4′ of A^2 in the NOESY spectrum of the 1 : 1 complex formed by dqdppn and d(GAGCTC)$_2$ described above. This resulted in the structure shown in Fig. 6b–d, where the ligand is inserted further into the helix, thus making close van der Waals contacts with the base pairs on either side of the insertion point.

As non-experimental distance restraints were used in the structural model, ^{1}H chemical shifts were calculated for the two complexes and used to validate the structure calculation. ^{1}H chemical shifts of DNA protons are largely determined by aromatic ring current shifts,[41,42] and consequently the location

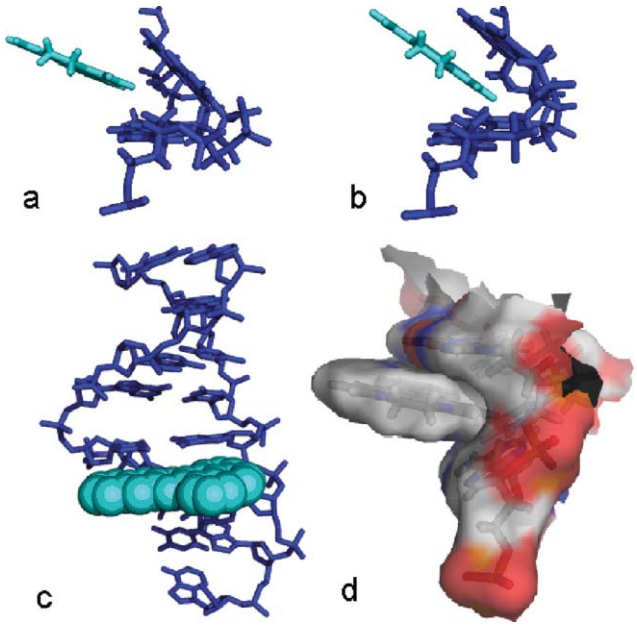

Fig. 6 Structures of dqdppn within the binding site of the 2 : 1 complex formed with d(AGAGCTCT)$_2$. (a) Detail of structure calculated using only the experimental NOE restraints. (b) Detail of structure calculated using a combination of experimental and pseudo-NOE restraints. (c) Complete structure calculated using a combination of experimental and pseudo-NOE restraints. (d) van der Waals surfaces of DNA and dqdppn, showing the close complementarity of the surfaces.

of an aromatic intercalator can be very sensitively probed by comparing chemical shifts in the free and complexed forms. This technique has been previously successfully exploited by us and others.[43–45]

Chemical shift differences between free and bound oligonucleotide were measured and compared to calculations for the two structures. The results are shown in Fig. 7. Most protons undergo only a small shift change on complexation, and their calculated shifts in general are similar for both complexes. Therefore only the protons that show the largest calculated difference between the

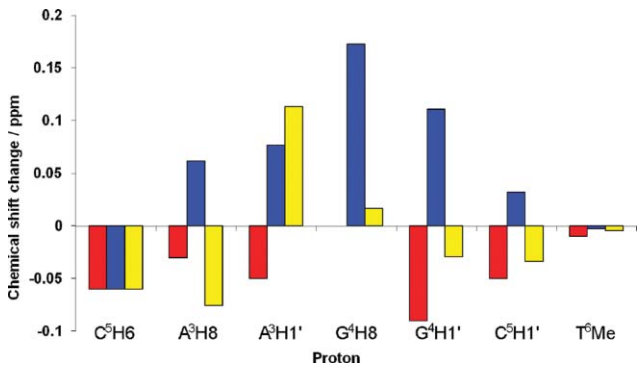

Fig. 7 Experimental and calculated chemical shift changes for the complex formed by dqdppn and d(AGAGCTCT)$_2$. The experimental chemical shift changes are shown in red; the calculated chemical shift changes for the model containing only experimentally derived distance restraints are in blue; and the calculated chemical shift changes for the model containing both experimental and non-experimental distance restraints are in yellow. The C^{5}H6 shift was used as the calibration (see text).

two complexes are shown. Since the fraction of bound ligand was unknown, and because conformational averaging in the complex leads to averaging of calculated shifts (and therefore to a general reduction in shift magnitude), calculated shift magnitudes were normalized to the experimental value for C^{5}H6, whose shift is essentially identical in both complexes.

It is clear from Fig. 7 that the complex calculated using the additional non-experimental restraints provides a significantly better match to the experimental data for all protons except A^{3}H1′, which matches poorly for both calculations. Thus, the shift calculations support the use of the additional restraints, and imply that the conformation shown in Fig. 6c is closer to the true averaged structure in solution

The calculated structures – for hexanucleotide in Fig. 4, and octanucleotide in Fig. 6c – are based on a low number of NOE restraints, and therefore are of low precision. In addition, the structure of the octanucleotide complex relies on a number of additional non-experimental restraints. However, CD and NMR data clearly demonstrate that the complexes have the same unusual 'side-on' geometry with the dqdppn long axis parallel to the base pair long axis, rather than being intercalated through the helix axis; indeed this geometry has been suggested as one possible orientation for the [Ru(phen)$_2$(dppz)]$^{2+}$ complex.[2b,7,11] Once dqdppn is intercalated in this way, there is relatively little conformational freedom in the system. This indicates that the overall features of the structures are correct, although the detailed geometries may be considerably less precise. In particular, the introduction of dqdppn into one side of the helix clearly introduces a bend in the helix axis in both complexes, but the angle of twist is not well defined by the data, although it is clear that the angle of twist is considerably greater for the hexanucleotide in which the intercalation site is at the terminal step. Similarly, the depth of insertion of dqdppn into the octamer is not well defined, although a more deeply intercalated structure fits the chemical shifts better. The structures are therefore best defined as models, consistent with all experimental data, rather than as detailed structures. The complex has considerable dynamics, as evidenced by line broadening and chemical shift changes, and it is likely that it is in equilibrium between several conformers, as observed for example with the [Ru(phen)$_2$dppz]$^{2+}$ complex,[16,17] thus rendering the structure determination of a single precise complex geometry in solution impossible even in principle. The data however suggest that these structures are the dominant form in solution.

The ring system of dqdppn is structurally related to that of actinomycin D,[12,14,24] an antibiotic that intercalates from the minor groove.[46] However, the geometry of intercalation is quite different, and in particular the longer and wider aromatic system of dqdppn means that it cannot easily access DNA from the minor groove in the side-on orientation. It is likely that this explains the observation of binding from the major groove. There has been debate over whether Ru^{II}(dppz) complexes bind from the major or minor groove. The results presented here are only relevant if such ruthenium complexes bind in the same side-on orientation as dqdppn, however as discussed in a recent paper,[47] we have discovered that such complexes can bind in a quite different orientation approaching from the minor groove.

Addition of a second equivalent of dqdppn to the hexanucleotide duplex gave rise to further broadening and non-specific NOEs. It thus appears that after the addition of the first nucleotide,

which occurs specifically at the terminal step, a second dqdppn can intercalate almost randomly elsewhere. Although addition of dqdppn stabilized the duplex against dissociation, it is likely to have destabilized the base stacking within the hexanucleotide and therefore increased the probability of nonspecific binding. Inspection of the structure (Fig. 4) shows that the adjacent base pair, 5′-A^2-G^3, has a perturbed geometry. This does not occur to the same extent for the octanucleotide, where it seems likely that the greater stability of the duplex prevents non-specific binding.

Conclusions

Our results demonstrate that in the hexanucleotide $d(G^1AGCTC^6)_2$, dqdppn intercalates at the G^1-A^2 step. However, when the hexanucleotide is capped by an extra base pair at each end to make the octanucleotides $d(AGAGCTCT)_2$ and $d(CGAGCTCG)_2$, dqdppn intercalates not at the equivalent position, namely G^2-A^3, but at the A^3-G^4 (or equivalently the C^5-T^6) step in both cases. Furthermore, dqdppn intercalates at the T^3-A^4 step in $d(GCTATAGC)_2$. It is therefore clear that the site of intercalation is not determined by the local sequence or base type. For the hexanucleotide, it seems likely that the intercalation site is preferred at the G^1-A^2 step because this is the terminal base pair. The 'side-on' intercalation mode seen here causes very significant twisting of the DNA axis, which is most easily accommodated in a terminal step, which is the most flexible and unrestrained by base stacking.

The relative stability and flexibility of individual base pair steps and longer tetranucleotide sequences has been analyzed by Hunter *et al.*[48,49] The flexibility of DNA sequences in general requires analysis at the tetranucleotide level, since there is conformational linkage between adjacent base pairs. This analysis shows that within the capped octanucleotides $d(AGAGCTCT)_2$ and $d(CGAGCTCG)_2$, the terminal AGAG and CGAG sequences are the least flexible. The other two tetranucleotides, GAGC and AGCT, are of comparable flexibility, but GAGC is less stable than AGCT by approximately 2.3 kcal mol^{-1}, implying that the GAGC sequence is more easily deformed. Furthermore, in $d(GCTATAGC)_2$ the tetranucleotide step TATA is by some distance the most flexible, with the TA steps being considerably more flexible than the AT step. Therefore we propose that for all three octanucleotides, the site of intercalation is determined by the ease of deformation of the local sequence, this being particularly important because of the helix deformation caused by cation binding. This result is in agreement with earlier studies on the binding of $[Ru(phen)_2(dppz)]^{2+}$ to bent and straight DNA.[50]

The third octanucleotide $d(GCTATAGC)_2$ was chosen because the analysis of Hunter, *et al.*[49] highlighted that the d(TATA) sequence is especially flexible. The TATA box sequence is of particular interest due to its role as the core promoter sequence within eukaryotic cells. Its flexibility is an important feature in its recognition by the TATA box binding protein which induces large scale bending at the recognition site.[51] The analysis presented herein suggests that dqdppn binds to the same sequences, implying a possible application for this molecule in the regulation of DNA transcription and translation processes, for example in cancer therapeutics.[52]

Experimental section

Materials

The synthesis of the dqdppn nitrate salt was carried out as reported.[25] Oligodeoxyribonucleotides were bought from Sigma-Genosys in a desalted and deprotected form, and were pure by hplc and mass spectroscopy. By NMR it was evident that although the protecting groups had been cleaved off, they had not been completely removed. However this does not appear to have interfered with the subsequent analysis. Duplex formation was analyzed using hyperchromic effects in UV and shift changes in NMR.

CD

Circular dichroism measurements were carried out on a Jasco Spectropolarimeter at 25 °C. DNA duplex formation is highly sensitive to the concentration of oligonucleotide, and therefore 1.8–2.0 mM oligonucleotide was used, to ensure complete duplex formation.

NMR

NMR measurements were carried out using Bruker 500 and 800 MHz spectrometers. Temperatures were calibrated using methanol and ethylene glycol. Standard pulse programs were used, except for the heteronuclear ^{31}P–^{1}H COSY which was adapted from a previous study.[37] The NOESY spectra used mixing times of 100 ms. Two-dimensional spectra were processed and analyzed in Felix (Accelrys Inc., San Diego, CA). For analysis of the exchange-broadened ^{31}P–^{1}H COSY spectrum it was found to be useful to phase the spectrum so that the cross peaks were dispersive in both dimensions (*i.e.*, 90° different from the conventional display) in order to reduce peak overlap.[53] Chemical shift calculations were carried out using the program *total*[54] using the XPLOR-refined coordinates of free octanucleotide and the complexes. Ring current shift intensity factors for the aromatic rings in dqdppn were calculated using the Haigh–Mallion method[55] using ring current intensity factors of 1.1. Ring current intensity factors for DNA bases were taken from published data.[41]

Molecular dynamics

Oligonucleotide structure files were prepared using the program nucgen within AMBER 9.[56] Dqdppn structure files were created using xplo2d[57] starting from the crystal structure.[24] Energy minimization and restrained molecular dynamics were carried out using XPLOR.[58] Throughout the calculations, dqdppn planarity and base pair hydrogen bonding and planarity were maintained, the latter based on the brestraints.inp file in the XPLOR tutorial. Experimental NOEs were classified as strong (1.8–3.8 Å) or weak (3.0–5.0 Å) by comparison to known B-DNA distances.[59] Calculations of complexes started with well-separated B-DNA duplex and ligand, and used energy minimization, followed by MD at 500 K, followed by a gradual exponential cooling over 4000 steps to 100 K with slight expansion of van der Waals radius throughout the cooling from 75% to 100%, followed by a final energy minimization.[60] Point restraints were used on atoms within the DNA bases distant from the ligand binding site, to prevent

DNA unwinding. Structures were analyzed using the programs Rasmol and pymol.[61]

Acknowledgements

We thank EPSRC for a DTA studentship (PW), Dr Tim Phillips for the preparation of dqdppn and the Wellcome Trust and Biotechnology and Biochemical Research Council for equipment grants.

Notes and references

1 (*a*) K. E. Erkkila, D. T. Odom and J. K. Barton, *Chem. Rev.*, 1999, **99**, 2777–2795; (*b*) C. Metcalfe and J. A. Thomas, *Chem. Soc. Rev.*, 2003, **32**, 215–224; (*c*) B. M. Zeglis, V. C. Pierre and J. K. Barton, *Chem. Commun.*, 2007, 4565–4579.
2 (*a*) A. E. Friedman, J. C. Chambron, J. P. Sauvage, N. J. Turro and J. K. Barton, *J. Am. Chem. Soc.*, 1990, **112**, 4960–4962; (*b*) R. M. Hartshorn and J. K. Barton, *J. Am. Chem. Soc.*, 1992, **114**, 5919–5925.
3 (*a*) E. J. C. Olson, D. Hu, A. Hörmann, A. M. Jonkman, M. R. Arkin, E. D. A. Stemp, J. K. Barton and P. F. Barbara, *J. Am. Chem. Soc.*, 1997, **119**, 11458–11467; (*b*) C. G. Coates, J. Olofsson, M. Coletti, J. J. McGarvey, B. Önfelt, P. Lincoln, B. Nordén, E. Tuite, P. Matousek and A. W. Parker, *J. Phys. Chem. B*, 2001, **105**, 12653–12664.
4 (*a*) M. K. Brennaman, J. H. Alstrum-Acevedo, C. N. Fleming, P. Jang, T. J. Meyer and J. M. Papanikolas, *J. Am. Chem. Soc.*, 2002, **124**, 15094–15098; (*b*) M. K. Brennaman, T. J. Meyer and J. M. Papanikolas, *J. Phys. Chem. A*, 2004, **108**, 9938–9944.
5 G. Pourtois, D. Beljonne, C. Moucheron, S. Schumm, A. Kirsch-De Mesmaeker, R. Lazzaroni and J.-L. Brédas, *J. Am. Chem. Soc.*, 2004, **126**, 683–692.
6 J. Olofsson, B. Önfelt, P. Lincoln and B. Nordén, *J. Phys. Chem. A*, 2003, **107**, 1000–1009; J. Olofsson, B. Önfelt and P. Lincoln, *J. Phys. Chem. A*, 2004, **108**, 4391–4398.
7 Y. Jenkins and J. K. Barton, *J. Am. Chem. Soc.*, 1992, **114**, 8736–8738.
8 (*a*) C. J. Murphy, M. R. Arkin, Y. Jenkins, N. D. Ghatlia, S. Bossmann, N. J. Turro and J. K. Barton, *Science*, 1993, **262**, 1025–1029; (*b*) E. D. A. Stemp, M. R. Arkin and J. K. Barton, *J. Am. Chem. Soc.*, 1997, **119**, 2921–2925.
9 R. E. Homlin, R. T. Tong and J. K. Barton, *J. Am. Chem. Soc.*, 1998, **120**, 9724–9725.
10 A. Greguric, I. D. Greguric, T. W. Hambley, J. R. Aldrich-Wright and J. G. Collins, *J. Chem. Soc., Dalton Trans.*, 2002, 849–855.
11 C. M. Dupureur and J. K. Barton, *J. Am. Chem. Soc.*, 1994, **116**, 10286–10287.
12 P. Lincoln, A. Broo and B. Nordén, *J. Am. Chem. Soc.*, 1996, **118**, 2644–2653.
13 I. Haq, P. Lincoln, D. C. Suh, B. Nordén, B. Z. Chowdhry and J. B. Chaires, *J. Am. Chem. Soc.*, 1995, **117**, 4788–4796.
14 E. Tuite, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 1997, **119**, 239–240.
15 H. Deng, J. Li, K. C. Zheng, Y. Yang, H. Chao and L. N. Ji, *Inorg. Chim. Acta*, 2005, **358**, 3430–3440.
16 C. Hiort, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 1993, **115**, 3448–3454.
17 C. M. Dupureur and J. K. Barton, *Inorg. Chem.*, 1997, **36**, 33–43.
18 Y. Jenkins, A. E. Friedman, N. J. Turro and J. K. Barton, *Biochemistry*, 1992, **31**, 10809–10816.
19 R. E. Holmlin, E. D. A. Stemp and J. K. Barton, *Inorg. Chem.*, 1998, **37**, 29–34.
20 I. Greguric, J. R. Aldrich-Wright and J. G. Collins, *J. Am. Chem. Soc.*, 1997, **119**, 3621–3622.
21 T. Biver, C. Cavazza, F. Secco and M. Venturini, *J. Inorg. Biochem.*, 2007, **101**, 461–469.
22 P. U. Maheswari, V. Rajendiran, M. Palaniandavar, R. Parthasarathi and V. Subramanian, *J. Inorg. Biochem.*, 2006, **100**, 3–17.
23 (*a*) A. H. Krotz and J. K. Barton, *Inorg. Chem.*, 1994, **33**, 1940; (*b*) R. H. Terbrueggen, T. W. Johann and J. K. Barton, *Inorg. Chem.*, 1998, **37**, 6874.
24 T. Phillips, I. Haq, A. J. H. M. Meijer, H. Adams, I. Soutar, L. Swanson, M. J. Sykes and J. A. Thomas, *Biochemistry*, 2004, **43**, 13657–13665.
25 T. Phillips, C. Rajput, L. Twyman, I. Haq and J. A. Thomas, *Chem. Commun.*, 2005, 4327–4329.
26 P. E. Schipper, *J. Am. Chem. Soc.*, 1979, **101**, 6826–6829.
27 R. Job and P. E. Schipper, *J. Am. Chem. Soc.*, 1981, **103**, 48–51.
28 M. Wirth, O. Buchardt, T. Koch, P. E. Nielsen and B. Nordén, *J. Am. Chem. Soc.*, 1988, **110**, 932–939.
29 H. C. Becker and B. Nordén, *J. Am. Chem. Soc.*, 1997, **119**, 5798–5803.
30 F. J. M. van de Ven and C. W. Hilbers, *Eur. J. Biochem.*, 1988, **178**, 1–38.
31 E. L. Ulrich, H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. K. Wenger, H. Y. Yao and J. L. Markley, *Nucleic Acids Res.*, 2008, **36**, D402–D408.
32 M. S. Searle, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1993, **25**, 403–480.
33 V. A. Roontga, C. R. Jones and D. G. Gorenstein, *Biochemistry*, 1990, **29**, 5245–5258.
34 D. G. Gorenstein, J. B. Findlay, R. K. Momii, B. A. Luxon and D. Kar, *Biochemistry*, 1976, **15**, 3796–3803.
35 D. G. Gorenstein, S. A. Schroeder, J. M. Fu, J. T. Metz, V. Roongta and C. R. Jones, *Biochemistry*, 1988, **27**, 7223–7237.
36 J. Ott and F. Eckstein, *Biochemistry*, 1985, **24**, 2530–2535.
37 V. Sklenář, H. Miyashiro, G. Zon, H. T. Miles and A. Bax, *FEBS Lett.*, 1986, **208**, 94–98.
38 A. Pardi, R. Walker, H. Rapoport, G. Wider and K. Wüthrich, *J. Am. Chem. Soc.*, 1983, **105**, 1652–1653.
39 B. P. Hudson and J. K. Barton, *J. Am. Chem. Soc.*, 1998, **120**, 6877–6888.
40 M. S. Searle, A. J. Maynard and H. E. L. Williams, *Org. Biomol. Chem.*, 2003, **1**, 60–66.
41 D. A. Case, *J. Biomol. NMR*, 1995, **6**, 341–346.
42 S. S. Wijmenga, M. Kruithof and C. W. Hilbers, *J. Biomol. NMR*, 1997, **10**, 337–350.
43 W. Bocian, R. Kawecki, E. Bednarek, J. Sitkowski, A. Pietrzyk, M. P. Williamson, P. E. Hansen and L. Kozerski, *Chem.–Eur. J.*, 2004, **10**, 5776–5787.
44 C. Giessner-Prettre and B. Pullman, *Q. Rev. Biophys.*, 1987, **20**, 113–172.
45 D. B. Davies, M. P. Evstigneev, D. A. Veselkov and A. N. Veselkov, *Biophys. Chem.*, 2005, **117**, 111–118.
46 S. Kamitori and F. Takusagawa, *J. Am. Chem. Soc.*, 1994, **116**, 4154–4165.
47 P. Waywell, V. Gonzalez, M. R. Gill, H. Adams, A. J. H. M. Meijer, M. P. Williamson, J. A. Thomas, *Chem.–Eur. J. in press*.
48 M. J. Packer, M. P. Dauncey and C. A. Hunter, *J. Mol. Biol.*, 2000, **295**, 71–83.
49 M. J. Packer, M. P. Dauncey and C. A. Hunter, *J. Mol. Biol.*, 2000, **295**, 85–103.
50 R. B. Nair and C. J. Murphy, *J. Inorg. Biochem.*, 1998, **69**, 129–133.
51 (*a*) Y. Kim, J. H. Geiger, S. Hahn and P. B. Sigler, *Nature*, 1993, **365**, 512–520; (*b*) J. L. Kim, D. B. Nikolov and S. K. Burley, *Nature*, 1993, **365**, 520–527.
52 R. H. Terbrueggen and J. K. Barton, *Biochemistry*, 1995, **34**, 8227–8234.
53 D. L. Turner, *J. Magn. Reson., Ser. A*, 1993, **104**, 197–202.
54 M. P. Williamson and T. Asakura, *J. Magn. Reson., Ser. B*, 1993, **101**, 63–71.
55 C. W. Haigh and R. B. Mallion, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1979, **13**, 303–344.
56 M. Bansal, D. Bhattacharya and B. Ravi, *Bioinformatics*, 1995, **11**, 281–287.
57 G. J. Kleywegt, *Newsletter Protein Crystallogr.*, 1995, **31**, 45–50.
58 A. T. Brunger, *X-PLOR, version 3.1: A system for X-ray crystallography and NMR*, Yale University Press: Cambridge, MA, 1992.
59 D. Neuhaus and M. P. Williamson, *The nuclear Overhauser effect in structural and conformational analysis*. 2nd edn.; Wiley-VCH: New York, 2000.
60 M. Nilges and J. Kuszewski, A. T. Brunger, Sampling properties of simulated annealing and distance geometry. In *Computational aspects of the study of biological macromolecules by NMR*, J. C. Hoch, F. M. Poulsen and C. Redfield, ed.; Plenum Press: New York, 1991; pp 451–455.
61 W. L. DeLano, *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA, USA: 2002.

