PCCP

PAPER

A DFT study of 2-aminopurine-containing dinucleotides: prediction of stacked conformations with B-DNA structure

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Darren A. Smith,†[a] Leo F. Holroyd,†[b] Tanja van Mourik[b] and Anita C. Jones*[a]

The fluorescence properties of dinucleotides incorporating 2-aminopurine (2AP) suggest that the simplest oligonucleotides adopt conformations similar to those found in duplex DNA. However, there is a lack of structural data for these systems. We report a density functional theory (DFT) study of the structures of 2AP-containing dinucleotides (deoxydinucleoside monophosphates), including full geometry optimisation of the sugar-phosphate backbone. Our DFT calculations employ the M06-2X functional for reliable treatment of dispersion interactions and include implicit aqueous solvation. Dinucleotides with 2AP in the 5'-position and each of the natural bases in the 3'-position are examined, together with the analogous 5'-adenine-containing systems. Computed structures are compared in detail with typical B-DNA base-step parameters, backbone torsional angles and sugar pucker, derived from crystallographic data. We find that 2AP-containing dinucleotides adopt structures that closely conform to B-DNA in all characteristic parameters. The structures of 2AP-containing dinucleotides closely resemble those of their adenine-containing counterparts, demonstrating the fidelity of 2AP as a mimic of the natural base. As a first step towards exploring the conformational heterogeneity of dinucleotides, we also characterise an imperfectly stacked conformation and one in which the bases are completely unstacked.

Introduction

2-Aminopurine (2AP) is the archetypal fluorescent base analogue; its close structural similarity to adenine (A) and its extraordinary photophysical sensitivity to inter-base interactions have led to its widespread use as a fluorescent probe of nucleic acid conformation.[1] Time-resolved fluorescence measurements of 2AP in DNA reveal the structural heterogeneity of the duplex; the excited 2AP population is partitioned between several different local conformational environments that provide distinctly different quenching efficiencies, resulting in a number of different fluorescence lifetimes. 2AP-containing DNA duplexes generally show fluorescence decays that can be described by four exponential components with typical lifetimes of <100 ps, ~0.5 ns, ~2 ns, and ~10 ns.[2-7] The very short lifetime component is attributed to a highly stacked conformation, in which excited 2AP is rapidly quenched by inter-base interaction. This is the dominant conformation, typically accounting for more than 70% of the emitting population. The long, ~10 ns, lifetime is comparable with that of free 2AP-riboside[8] and is attributed to an unstacked conformation in which 2AP is extrahelical, free from quenching interactions. This is a minor conformation, typically accounting for <5% of the emitting population. The intermediate lifetimes correspond to imperfectly or partially stacked structures, in which 2AP is intrahelical, but is not subject to rapid quenching.

The conformational variability and flexibility of the DNA duplex (even within the restricted environment of a crystal lattice) is evident from the range of values of base-step parameters, sugar-phosphate backbone torsional parameters and pseudorotation phase angles (sugar pucker modes) that are found in the numerous X-ray structures that are available.[9-11] The idealised view of a single canonical structure is far from reality. A detailed conformational analysis of a large number of X-ray crystal structures of naked and complexed DNA has been reported by Svozil *et al.*[9] Each dinucleotide step (two consecutive bases contained within a longer sequence) was analysed to determine sugar-phosphate and glycosidic torsional angles as well as sugar puckering modes. A clear outcome from this study was that a DNA duplex could not simply be characterised by any one conformational family: individual structures generally showed considerable variation, sometimes even between sequential dinucleotide steps. Although the structures were dominated by several major conformational families (namely AI, AII, BI, and BII, which are subsets of the established Watson-Crick A- and B-form structures), there was a large number of minor conformers identified by the analysis. This observation was interpreted to mean that there are many energetically low-lying states within the

[a] EaStCHEM School of Chemistry, The University of Edinburgh, David Brewster Road, Edinburgh, EH9 3FJ, UK. E-mail: a.c.jones@ed.ac.uk
[b] EaStCHEM School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST, UK.

† The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors

Electronic Supplementary Information (ESI) available: base-step parameters of twisted dinucleotides and dimers; sugar-pucker parameters of twisted dinucleotides; base-step parameters, backbone torsional angles and sugar pucker parameters of unstacked dinucleotides; comparison of structures of twisted dinucleotides and free dimers; Cartesian coordinates of all optimised structures. See DOI: 10.1039/x0xx00000x

conformational space of DNA that play a significant role in its dynamic behaviour. Indeed, it has been shown that, although sometimes obscured, polymorphism is prevalent within DNA structure.[12, 13]

Dinucleotides (deoxydinucleoside monophosphates) of 2AP with the natural bases display remarkably similar fluorescence decay parameters to 2AP in duplex DNA; four decay components are observed with lifetime values analogous to those summarised above.[14] This suggests that the simplest of oligonucleotide systems adopts similar conformational states (ranging from highly stacked to completely unstacked) to those found in the duplex. Moreover, the decay parameters imply that well-stacked states are highly populated in dinucleotides. While these results point to broadly similar conformational properties of dinucleotides and duplexes, they provide no insight into the structures of the conformers. For example, the extent to which a 'highly stacked' conformation of a dinucleotide might resemble the structure of a base-step in DNA remains to be determined. There is also the important underlying question as to whether the replacement of adenine by 2AP might cause structural perturbations that could undermine the relevance of such studies to the conformational properties of the natural system.

NMR spectroscopy of deoxydinucleoside monophosphates has given some insight into conformational structures through the measurement of scalar coupling constants (J-couplings), which yield information on the sugar-phosphate backbone.[15, 16] These measurements indicate the presence in solution of an equilibrium between conformations with South (C2'-endo) or North (C3'-endo) sugar pucker, which are loosely attributed to stacked and unstacked states, respectively. Such studies suggest preferential population of stacked states. However, the acquisition of detailed structural information is inhibited by the inability of NMR to distinguish between conformations that interconvert on timescales faster than milliseconds, and the limited ability to determine proton-proton distances by nuclear Overhauser effect (NOE) spectroscopy because of the scarcity of relevant protons.

Gas-phase structures of dinucleotides have been studied by ion-mobility mass spectrometry, which can distinguish between conformers which have significantly different collisional cross-sections and do not interconvert on the experimental timescale of 500 μs.[17] By comparing measured cross-sections with those of structures generated by molecular dynamics simulations, three conformational families were identified at low temperature (80 K): a low-energy stacked form; an extended, open (completely unstacked) form; and a structure with an intra-nucleotide hydrogen bond, in which the two bases are approximately co-planar. Although the structural information that can be extracted from these results is limited, the direct experimental observation of conformational multiplicity in dinucleotides is significant.

Molecular dynamics simulations provide important insight into the conformational heterogeneity and dynamics of DNA,[18, 19] but they do not deliver the precise structures and energies of specific conformations that can be obtained, in principle, from quantum chemical calculations. Traditionally, *ab initio* quantum mechanical studies of base-stacking interactions in nucleic acid systems have been performed using wave-function-based methods that reliably treat long-range dispersion interactions.[20, 21] However, the high computational demands restrict this approach to small structural fragments. Density functional theory (DFT) calculations, which scale more favourably with system size, are potentially advantageous for the study of larger nucleic acid structures, but some commonly used density functionals, such as B3LYP, give very poor results for systems in which dispersion interactions are important.[22, 23] However, over the past decade or so, density functionals that do describe dispersion, either through explicit correction or *via* parameterisation, have become available and have been applied successfully to nucleobase stacking.[22, 24-28] In particular, the M06-2X functional,[29-31] which is employed in the present work, has been shown to give excellent performance for the prediction of base-stacked structures and stacking energies.[22, 25, 27]

There have been numerous quantum chemical calculations on stacking interactions in free nucleobase dimers (i.e. in the absence of the sugar-phosphate backbone); see, for example, the study of Morgado *et al.*[32] and references therein. In general, minimum-energy structures obtained for stacked nucleobase dimers do not resemble canonical DNA base-step structures (indeed would not be attainable in an oligonucleotide) as exemplified in a recent DFT (M06-2X) study of adenine and 2AP dimers.[33] For A|A, A|2AP, and 2AP|2AP (where the vertical line denotes that the bases are stacked but not covalently bonded) minimum-energy structures with twist angles of about 60° were found, on the verge of what can be achieved in DNA, but these were not the lowest energy minima; the most stable structures had much higher twist angles. This demonstrates the influence of the backbone on the conformational energy landscape of oligonucleotides and the need for calculations that include optimisation of the backbone structure. There have been a few such calculations on natural dinucleotides, as reviewed below, but none, to our knowledge on 2AP-containing dinucleotides.

Several studies have used quantum chemical calculations to investigate excited-state properties of free dimers consisting of 2AP and a natural base (2AP|X), where a canonical B-DNA structure has been imposed on the dimer to simulate inter-base interaction in DNA.[34-39] Recent studies by Matsika and coworkers[37-39] examined relaxation of 2AP|X dimers, along the excited-state surface, from an initially excited B-form conformation and found different quenching pathways depending on whether 2AP was in the 5' or 3' position. Although the prediction of conformational influence on the non-radiative decay mechanism was significant, the relevance of this computational scenario to experimental observations is questionable, since the constraints of the sugar-phosphate backbone on

conformational relaxation and the effects of solvation were neglected.

Churchill and Wetmore[22] investigated the ability of three different density functionals (B3LYP, MPWB1K, and M06-2X) to accurately reproduce the structural features of a dinucleoside monophosphate unit within DNA. Three different phosphate models, anionic, neutral (protonated) and (sodium) counter-ion, were also assessed. Geometry-optimised structures of three guanine (G) dinucleotides, 5'-d(GpX)-3', where X was thymine, uracil or 5-bromouracil, were compared with typical B-DNA structures. To judge the ability of each method to emulate base-stacking, the relative orientations of the bases were classified as distorted, repelled, tilted or stacked. The structures were further scrutinised by considering the torsional angles of the sugar-phosphate backbone. The results heavily favoured the use of M06-2X, which reliably predicted structures resembling B-DNA. The other two functionals generally failed to optimise to a base-stacked form. Anionic and counter-ion phosphate models were found to exhibit better performance than the neutral model, which could not consistently predict stacked structures.

Barone *et al.*[26] studied all 16 possible permutations of deoxydinucleoside monophosphates, using dispersion-corrected DFT, in the presence of sodium counter-ions and with aqueous solvation modelled by the conductor-like screening model. As this study was concerned mainly with energetics, analysis of the optimised structures was limited to the backbone torsional angles. On this basis, the optimised geometries were deemed to have B-DNA conformation. To gain further insight into the conformations reported, we have undertaken further analysis of the optimised structures containing 5'G or 5'A, which correlate with the systems (5'2AP or 5'A) investigated in the present work. We applied the 3DNA program (see Experimental) to the structural Cartesian coordinates supplied by the authors in the Supplementary Information to derive base-step parameters. We found that, despite exhibiting backbone torsional angles that could be considered similar to typical B-DNA form (especially after averaging), the base-step parameters showed considerable distortion from the canonical conformation, as illustrated in Table S1. Many of the optimised structures had large tilt and/or roll angles and only two had twist angles close to 36° (the ideal twist angle in B-DNA). All four purine-purine dinucleotides had large twist angles, between 50° and 60°. This highlights the risk of relying on backbone torsional angles alone to characterise the dinucleotide structure, as is done in many studies.

Poltev *et al.*[40-44] have published a series of papers concerning the full geometry optimisation of dinucleoside monophosphates (with sodium counter-ion) using DFT, neglecting dispersion. These studies focused on analysis of the sugar-phosphate backbone. The base-step parameters were not evaluated beyond a rough measure of the twist angle and mutual planarity. It was found that the geometry-optimised backbone structures of dinucleotides reproduced the torsional angles and sugar pucker found in DNA crystal structures. Although the authors concede that base-stacking interactions must have some importance in determining the precise conformational state in nucleic acids, their main conclusion is that the sugar-phosphate backbone is the dominant structure-determining element and it is inferred that duplexes are predisposed to a particular conformational family (AI, AII, BI, BII) by the backbone structure of the single strands. Since stacking interactions are not properly modelled in these studies, the influence of the backbone structure is probably overstated. Moreover, the conformational analysis considers only the broad classification into the four canonical families and overlooks the diversity of duplex structures that exists within these classes.

In this paper, we report the first DFT study of the deoxydinucleoside monophosphates of 2-aminopurine with each of the natural bases, including reliable treatment of dispersion interactions (employing the M06-2X functional), full geometry optimisation of the sugar-phosphate backbone and the presence of implicit aqueous solvation. The analogous adenine-containing dinucleotides have also been studied. We present a comprehensive analysis of the computed conformational structures in terms of the conventional structural parameters of DNA: the base-step parameters (shift, slide, rise, tilt, roll, and twist), the torsional angles of the sugar-phosphate backbone and the glycosidic bonds, and the pseudorotation phase angles (sugar pucker modes). These structural parameters are compared in detail with those characteristic of B-DNA, as derived from the analysis of large numbers of X-ray crystal structures by Svozil *et al.*[9] and Olson *et al.*[10] We investigate whether 2AP-containing dinucleotides adopt stacked structures that closely resemble B-DNA and whether 2AP can faithfully mimic adenine in reporting the conformational properties of DNA. We begin to explore the conformational diversity of these systems by characterising three different minimum-energy structures that display different degrees of base stacking.

Experimental

The dinucleotides (deoxydinucleoside monophosphates) studied were of the form 5'-d(2pN)-3' or 5'-d(ApN)-3', where 2 = 2-aminopurine (2AP), A = adenine, and N = adenine, guanine (G), cytosine (C), or thymine (T), as illustrated in Figure 1. Dinucleotides containing a 3'-purine will be abbreviated collectively as d(ApR) or d(2pR) and those containing a 3'-pyrimidine as d(ApY) or d(2pY). Starting structures for geometry optimisation of d(2pN) in stacked conformations were obtained by mutating A to 2AP in appropriate dinucleotides extracted from two alternative DNA crystal structures, one exhibiting typical B-form structure and the other showing a somewhat twisted (twist angle ~50°) base-step structure (PDB codes 4C64 and 3R86, respectively). Starting structures for d(2pN) in unstacked conformations were obtained from the crystal structure of a base-flipped complex of DNA with methyltransferase

M. TaqI, where 2AP is the flipped base (PDB code 2IBS). The starting geometry of each d(ApN) dinucleotide was created from the geometry-optimised structure of the corresponding d(2pN) by mutating 2AP to A. The dinucleotides were in the anionic form, in accordance with the recommendations of previous work.[22]

Geometry optimisation was performed using the Gaussian 09 package,[45] employing density functional theory (DFT) with the M06-2X functional[30] and 6-31+G(d) basis set. The M06-2X/6-31+G(d) combination was previously found to give counterpoise-corrected interaction energies in excellent agreement with high-level CCSD(T) results for stacked uracil dimers (U|U).[25] Gaussian's tight convergence criteria and ultrafine integration grid (containing 99 radial shells and 590 angular points per shell) were used. Harmonic vibrational frequencies were computed at the same level of theory to verify the nature of the stationary points as minima and to compute free energies. Gibbs free energies were calculated at 298.15 K and 1 atmosphere of pressure, from the harmonic vibrational frequencies, by standard thermochemical analysis, using the principal isotope for each element. Aqueous solvation was modelled using the Polarizable Continuum Model (PCM).[46]

Molecules were visualised and manipulated using a combination of Jmol,[47] PyMol (The PyMOL Molecular Graphics System, Version 1.6.0 Schrödinger, LLC), Molden,[48] and MATLAB (R2013b, The MathWorks, Inc., Natick, Massachusetts, United States of America). Dinucleotide geometry was evaluated using 3DNA[49, 50] and w3DNA.[51] These comprehensive analysis programs provided stacking parameters (slide, shift, rise, tilt, roll, and twist), torsional angles, and sugar pucker.

Results and Discussion

B-form structures

Optimisation of starting geometries derived from the B-form crystal structure (PDB code 4C64) yielded the dinucleotide geometries (local energy minima) shown in Figure 2. It is evident visually that the structures of all the dinucleotides resemble the typical base-step structure of B-DNA duplexes. This is confirmed by comparison of the base-step parameters of the computed structures with those derived from B-DNA crystal structures,[10, 49] as shown in Table 1. For both d(2pN) and d(ApN) dinucleotides, the stacking parameters are generally within the range of typical B-form values. It is apparent that the rise values are generally a little less than the idealised value of 3.34 Å. This discrepancy may be due to the computational methodology. There is some indication in the literature that M06-2X may underestimate inter-base distances[22, 27] and we also note that a high-level CCSD(T) structure for U|U displayed an inter-base distance of 3.3 Å, in closer agreement with the idealised value.[52] Nevertheless, despite the slightly shorter distances predicted by M06-2X, the M06-2X/6-31+G(d) combination gave counterpoise-corrected interaction energies in very close agreement with CCSD(T) results for stacked U|U.[25]

The discrepancy between the computed rise values and the idealised value may, however, also genuinely reflect the difference in stacking interactions between different pairs of bases. Indeed, the rise is seen to depend on the identity of the 3'-base, but is independent of whether the 5'-base is 2AP or A. In the dinucleotides only a single stacking interaction is being optimised and we are only considering structures in which 2AP or A is at the 5' position. The crystal-derived parameters are representative of larger constructs in which the rise must reflect a compromise between the stacking interactions of multiple consecutive bases, leading to a more homogeneous structure.

The backbone structures of the dinucleotides also conform to that typical of B-form DNA. As shown in Table 2 the torsional angles of the backbone and the glycosidic bonds are in good agreement with values typical of B-DNA. The sugar-pucker parameters, Table 3, are also consistent with the range seen in B-DNA structures. Although the dinucleotides do not generally show the classical C2'-endo sugar conformation, the pseudorotation angles lie close to the C2'-endo range (144° to 180°)[53] and correspond to C3'-exo conformation (seen most commonly for the 5'-nucleoside) and C1'-exo conformation (most common for the 3'-nucleoside). The B-DNA helix permits a broad range of sugar conformations and crystal structures frequently show C1'-exo and C3'-exo conformations.[11, 54] Indeed, on the basis of crystal data, Dickerson[11] has proposed that C1'-exo sugar pucker, rather than C2'-endo, should be deemed typical of B-DNA. The O4'-endo conformation found for the 3'-nucleoside in d(2pC) and d(ApT) is also included by Dickerson in the range found in B-DNA structures.

Twisted stacked structures

Optimisation from an alternative starting geometry (based on PDB code 3R86) disclosed another set of minimum energy dinucleotide structures in which the bases are substantially stacked, but deviate from the B-form arrangement, as illustrated in Figure 3. Examination of the base-step parameters (Table 4) shows that these conformations have twist and/or slide values that differ significantly from the average values for B-DNA. The dinucleotides containing a 3'-purine show particularly high twist values (>50°), while the pyrimidine-containing dinucleotides show larger slide values. These structures, especially the twist angles, are similar to those of the corresponding dinucleotides reported by Barone *et al.*,[26] as shown in Table S1, although the latter were designated as B-DNA structures by the authors, on the basis of the backbone torsional angles. The prediction of twisted structures by Barone's calculations may be due to their starting geometries, which were built using the TINKER molecular design program package, and/or the level of theory used, BLYP-D with an STO-based TZ2P basis set.

A notable feature of the twisted structures is the presence of intramolecular hydrogen bonds. Two types of hydrogen bond were observed, as defined in Figure 1 and illustrated in

Figure 3. In 2AP-containing dinucleotides, an inter-nucleotide H-bond (designated HB1) is present, between the 2AP amino group (H22) and the neighbouring deoxyribose sugar group (O4'). This H-bond cannot be formed by adenine because of the different position of the amino group. The presence of HB1 correlates with a greater twist angle for d(2pN) than d(ApN). A similar interaction to HB1 is apparent in computed structures of d(GpN) dinucleotides reported by Barone *et al.*,[26] but is not commented on by the authors. In pyrimidine-containing dinucleotides, both d(2pY) and d(ApY), an intra-backbone H-bond (HB2) is present between the terminal H3' atom and the O2P oxygen atom of the phosphate group. (It should be noted that formation of HB2 in a DNA duplex is precluded by the absence of H3', as a result of the continuation of the polymeric backbone structure.)

The effect of these hydrogen bonds on the backbone structure is evident from the torsional angles in Table 5. The dinucleotides that have HB2, d(2pY) and d(ApY), show backbone structures that are quite distinct from those of d(2pR) and d(ApR), but very similar to each other. The effect of HB2 is seen in the values of δ_1, ε_1, ζ_1, β_2, and χ_1, and appears to override any influence of the additional presence of HB1 in d(2pY). The presence of HB1 does, however, affect the backbone structure of the d(2pR) dinucleotides, as compared with d(ApR) which have no H-bonds. The torsional angles of d(2pA) and d(2pG) are similar to each other but differ from those of d(ApR), particularly ε_1, α_2, β_2, χ_1, and χ_2. The backbone torsional angles of d(ApR), the only structures that contain no hydrogen bonds, are much closer to the B-form values than those of the other dinucleotides.

The presence of HB2 also appears to correlate with the significantly smaller twist angle seen for d(2pY) and d(ApY), compared with d(2pR) and d(ApR) (Table 4), suggesting that backbone geometry imposes a constraint on the base-stacking. This was confirmed by removing the backbone from each of the d(2pN) and d(ApN) dinucleotides and re-optimising the structure of the resulting free dimer, designated 2|N and A|N, respectively. As shown in Tables S2 and S3, the base-stacking in 2|R and A|R dimers is essentially unchanged relative to the respective dinucleotide structures. In contrast, the 2|Y and A|Y dimer twist angles are markedly different from those of the corresponding dinucleotides; the twist is much less in 2|Y than in d(2pY) and much greater in in A|Y than in d(ApY). It appears that in d(2pR) and d(ApR) the bases are stacked in an optimum arrangement, whereas in d(2pY) and d(ApY) the backbone structure confers a sub-optimal base-stacking geometry. Comparison of dinucleotide and dimer structures (Figure S1) illustrates that formation of d(2pY) structures that preserved the optimum 2|Y dimer geometry would require very contorted backbone structures.

A common feature of all the twisted dinucleotides is that the 3'-sugar shows a C3'-endo conformation, typical of A-DNA, whereas the 5'-sugars retain B-type pucker (Table S4). Otherwise, the overall impression is that, although derived from a common starting geometry, this family of dinucleotide structures shows much greater conformational diversity than the B-form structures.

Unstacked structures

Optimisation of starting geometries derived from the crystal structure of a base-flipped DNA-enzyme complex (PDB code 2IBS) located a minimum-energy conformation in which the bases are entirely unstacked. (Structures were calculated only for dinucleotides with G or T as the 3'-base, as the overall structure is little affected by the structures of the individual bases.) As illustrated in Figure 4, this conformation has an extended backbone structure; the bases are widely separated (around 13 Å between the base centres) and do not interact with each other or the sugar-phosphate backbone. The existence of a minimum-energy structure of this form is consistent with the observation of a dinucleotide conformation with an open structure (large collision cross-section) in gas-phase ion mobility experiments.[17]

On the basis of visual inspection, the structures of all four dinucleotides appear to be very similar; this is confirmed by the structural parameters in Tables S5-S7 ("Base-step" parameters (Table S5) are presented for completeness, to illustrate the similarity of the four structures, but are not physically meaningful in these unstacked structures). In these conformers, where the backbone is unconstrained by inter-base interactions, the torsional angles of the backbone (Table S6) and the sugar pucker (Table S7) are independent of the identity of the bases; for all four conformers, the backbone structures are virtually identical. As would be expected, the backbone torsional parameters differ considerably from typical B-form values. However, it is interesting to see that, in a structure that differs so greatly from B-DNA, the sugar pucker (both 5' and 3') in all cases is C2'-endo, which is usually considered to be characteristic of the B-form backbone. This throws some doubt on the assumption made in the interpretation of NMR data that South (C2'-endo) pucker is indicative of stacked conformations.

Conformational energies

The relative energies of the B-form, twisted, and unstacked conformations of each dinucleotide are summarised in Table 6. In each case, potential energy (electronic energy) differences, ΔE, and free energy differences, ΔG, relative to the respective lowest-energy conformation are given. Unsurprisingly, in view of the lack of inter-base interaction, the unstacked structures have much higher potential energies than the stacked structures. The potential energy differences between unstacked and B-form structures of d(ApT) and d(ApG), around 10 kcal mol^{-1}, are comparable to the base-base interaction energy differences of 7-10 kcal mol^{-1} between unstacked and stacked forms of these deoxydinucleoside monophosphates determined by Norberg and Nilsson from potential of mean force (PMF) calculations.[55] In the PMF study, stacked and unstacked conformations were defined by a single reaction coordinate,

the distance between glycosidic nitrogen atoms (R_{NN}), with stacked defined as R_{NN} = 4.5 Å, and unstacked as R_{NN} = 9.0 Å. (For our B-form and unstacked conformations the values of R_{NN} are 3.8 Å and 10.6 Å, respectively.)

The free energy differences between unstacked and B-form conformations are considerably smaller than the potential energy differences, indicating substantial entropic contribution to the stability of the unfolded structures. Our ΔG value for d(ApT) is consistent with the PMF profile reported by Norberg and Nilsson,[55] which indicates a free energy difference of 4-6 kcal mol^{-1} between unstacked and stacked states. However, our value for d(ApG) is significantly higher than the 2-3 kcal mol^{-1} predicted in the latter study. Indeed our prediction of similar free energy differences between stacked and unstacked conformations for d(ApG) and d(ApT) appears to be contrary to the consensus from molecular dynamics simulations[19, 55-57] that the stacking propensity of purine-purine dinucleotides is greater than that of purine-pyrimidine dinucleotides. However, the majority of these studies were concerned with RNA, rather than DNA, dinucleotides. It is notable that a recent molecular dynamics study by Brown *et al.*[58] using a revised AMBER force field, confirmed this trend for RNA but predicted that purine-purine and purine-pyrimidine have approximately equal stacking free energies in DNA dinucleotides, supporting our result.

It is important to recognise that the difference in free energy between B-form and unstacked conformations reported here cannot be compared quantitatively with stacking free energies calculated from equilibrium constants (population ratios) derived from molecular dynamics simulations or experimental measurements of dinucleotide melting transitions. The latter calculations are based on the assumption of a two-state equilibrium between stacked and unstacked states, where each state consists of an ensemble of numerous conformations and the transition from stacked to unstacked is defined, somewhat arbitrarily, by the value of a chosen reaction coordinate (in MD simulations) or an experimental measurand. In our case, we are calculating the difference in free energy between two, specific, individual conformational structures.

To our knowledge, the only experimental measurement of stacking free energy that closely approximates to our computational scenario is from the single molecule study of spontaneous flipping of a single DNA base in a mismatched base pair (i.e. in the absence of inter-strand hydrogen bonding) by Yin *et al.*[59] They measured equilibrium constants in the range 10^{-2} to 10^{-4} (at 305 K), for flipping of a base from intrahelical to extrahelical positions, giving free energy differences of 2-5 kcal mol^{-1}.

The stability of the stacked structures relative to the unstacked ones may have been modestly over-estimated by an intramolecular form of the basis-set superposition error (BSSE), a widely encountered issue in computational chemistry. This apparent energy lowering is not physically justified, and since it is conformation-dependent (being greater in compact structures than in extended ones), it is likely to artificially stabilise stacked dinucleotides with close base–base contacts, relative to unstacked ones, where the bases are far apart.[60-62] The error cannot be calculated exactly in an intramolecular case such as this, but the intermolecular BSSEs in stacked A|A and 2AP|2AP dimers[33] and stacked A|T dimers (Holroyd, unpublished results), studied at the same level of theory as herein, have been calculated to be between 1.4 and 1.8 kcal mol^{-1} by the counterpoise procedure.[63] This represents not more than 13-18% of the potential energy differences between stacked and unstacked dinucleotides given in Table 6 and accounting for intramolecular BSSE would, therefore, not affect our conclusions.

For each 2AP-containing dinucleotide, the twisted conformation has somewhat lower potential energy than the B-form structure, but the B-form is more stable in terms of free energy. There is evidently significant entropic contribution to the relative stability of the B-form structures; this can be attributed, at least in part, to the presence of intramolecular hydrogen bonds in the twisted structures. We note also that the PCM continuum solvation model may overestimate the stability of structures with intramolecular hydrogen bonds and, thus, accounting for explicit solvation would further favour the B-form structures. Therefore, in a hypothetical equilibrium (at 298 K) between B-form and twisted conformations, the vast majority of the d(2pN) population would be expected to exist in the B-form structure.

While the adenine-pyrimidine dinucleotides also show free energies that favour the B-form structures, d(ApA) and d(ApG) stand out as having little difference in free energy between B-form and twisted structures (indeed twisted d(ApG) is slightly more stable than its B-form). As noted above, the backbones of twisted d(ApR) are closer to B-form than those of any of the other twisted dinucleotides and this translates into comparable free energies for their twisted and B-form structures.

Conclusions

We have identified and characterised three minimum-energy structures that exemplify the conformational heterogeneity that is manifested experimentally as the multi-exponential fluorescence decay of 2AP-containing dinucleotides. These structures are by no means exclusive, and do not necessarily include the global minimum, but are indicative of the range of conformations, from highly stacked to completely unstacked, that can be populated. The existence of a dinucleotide conformation that complies closely with the B-form structure of DNA, with respect to all structural parameters, demonstrates the importance of base-stacking interactions, modulated by the constraints of backbone geometry, in determining B-DNA structure.

The B-form conformation can plausibly be associated with the very short fluorescence decay component that is observed for d(2pN), while the unstacked conformation is consistent with the observation of a long fluorescence

lifetime that resembles that of free 2AP. The twisted structures demonstrate the existence of well-defined, minimum-energy conformations between the two extremes, which give rise to intermediate fluorescence lifetime components. The diversity of the twisted structures hints at the complexity of the conformational space and the likely existence of a multiplicity of local minima. (We are currently exploring the conformational landscape in more detail and this will be the subject of a future publication.) The computed relative free energies predict that, within the limited context of the three conformations considered here, the B-form should account for the majority of the d(2pN) population. This is in broad agreement with the experimental finding that the short-lived decay component generally has the greatest amplitude in the fluorescence decay of d(2pN).[14] (A quantitative comparison between computationally predicted and experimentally inferred populations is not appropriate until a more extensive exploration of the conformational landscape has been completed.)

It is clear from the structural similarity between the B-form conformation of each 2AP-containing dinucleotide and its adenine-containing counterpart that 2AP faithfully reproduces the stacking interactions of the natural base. However, as seen in the twisted conformations, differences in hydrogen-bonding interactions (the formation of HB1 by 2AP but not by adenine) can result in differences in the respective conformational structures. Such hydrogen-bonding effects may contribute to the experimentally observed impact of 2AP inclusion on DNA melting temperature and base-pair opening times. For example, 2D-NMR measurements[64] showed that, within a duplex structure, the 2AP-T base pair has a shorter lifetime than A-T, and the lifetimes of the neighbouring base-pairs are also reduced when A is replaced by 2AP. The formation by 2AP of a hydrogen bond, such as HB1, with the backbone, would be anticipated to have such an effect, by competing with 2AP-T base-pairing and also perturbing the stacking interaction with neighbouring bases as a result of the increased twist angle.

The overall geometry of the dinucleotide conformations is determined by the interplay of base-stacking interactions and the constraints of backbone structure. In the B-form conformations there is little variation in structure amongst the different dinucleotides; the B-DNA structure can be considered to represent a universally favourable minimum-energy structure, in which a common compromise is found between base-stacking and backbone geometry, for all dinucleotides. In these conformations, we see some influence of inter-base interactions on the backbone structure: there are small differences in the torsional angles of the 3'-section of the backbone (β_2, γ_2, δ_2, χ_2) and the 3'-pseudorotation angle (sugar pucker), depending on whether the 3'-base is a purine or pyrimidine (Table 1). In the twisted conformations, the presence of different hydrogen-bonding motifs, within the backbone and between 2AP and the backbone, results in a variety of structures, amongst which we see clear examples of the optimisation of backbone structure over-riding base-stacking interactions.

2AP has long been accepted by the nucleic acids community to be a base analogue that can be substituted for adenine with minimal perturbation of oligonucleotide structure. This presumption, although partly intuitive (given the close structural similarity between 2AP and adenine), is supported by indirect experimental evidence, such as the small effect of 2AP inclusion on DNA melting temperature and the ability of enzymes to recognise 2AP-containing sequences, and limited direct evidence from a handful of crystal structures of 2AP-containing oligonucleotides (in complex with enzymes). The results presented here add further substance to this premise.

In spite of the success of 2AP as a fluorescent mimic of a natural base, shortcomings in its photophysical properties are stimulating the development of new isomorphic base analogues with higher fluorescence quantum yields and longer emission wavelengths.[65, 66] Isomorphic analogues are designed to closely resemble the corresponding natural bases with respect to their overall dimensions, hydrogen-bonding patterns, and ability to form isostructural Watson-Crick base pairs. The use of DFT calculations in the manner demonstrated here should be very valuable in predicting the ability of such base analogues to simulate the crucial inter-base stacking interactions of natural bases, thereby guiding the design of optimised isomorphic structures.

Acknowledgements

We gratefully acknowledge access to the EaStCHEM Research Computing Facility and thank Dr David Rogers for his assistance and advice. This work was supported by studentship funding from the Engineering and Physical Sciences Research Council Doctoral Training Account to DAS and LFH (EP/K503162/1); EaStCHEM, University of Edinburgh and University of St Andrews; University of Melbourne. The research data supporting this publication can be accessed at http://dx.doi.org/10.17630/45637538-1906-4c3c-83da-0ba6bb20d58b

References

1. A. C. Jones and R. K. Neely, *Q. Rev. of Biophys*, 2015, **48**, 1-36.
2. R. K. Neely, D. Daujotyte, S. Grazulis, S. W. Magennis, D. T. F. Dryden, S. Klimasauskas and A. C. Jones, *Nucleic Acids Res.*, 2005, **33**, 6953-6960.
3. S. V. Avilov, E. Piemont, V. Shvadchak, H. de Rocquigny and Y. Mély, *Nucleic Acids Res.*, 2008, **36**, 885-896.
4. R. K. Neely and A. C. Jones, *J. Am. Chem. Soc.*, 2006, **128**, 15952-15953.
5. T. Ramreddy, B. J. Rao and G. Krishnamoorthy, *J. Phys. Chem. B*, 2007, **111**, 5757-5766.
6. T. Ramreddy, M. Kombrabail, G. Krishnamoorthy and B. J. Rao, *J. Phys. Chem. B*, 2009, **113**, 6840-6846.
7. T. Sabir, A. Toulmin, L. Ma, A. C. Jones, P. McGlynn, G. F. Schröder and S. W. Magennis, *J. Am. Chem. Soc.*, 2012, **134**, 6280-6285.

8. R. K. Neely, S. W. Magennis, D. T. F. Dryden and A. C. Jones, *J. Phys. Chem. B*, 2004, **108**, 17606-17610.
9. D. Svozil, J. Kalina, M. Omelka and B. Schneider, *Nucleic Acids Res.*, 2008, **36**, 3690-3706.
10. W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X.-J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C.-S. Tung, E. Westhof, C. Wolberger and H. M. Berman, *J. Mol. Biol.*, 2001, **313**, 229-237.
11. R. E. Dickerson, in *International Tables for Crystallography*, International Union of Crystallography, Chester, England, 2006, vol. F, pp. 588-622.
12. T. Maehigashi, C. Hsiao, K. Kruger Woods, T. Moulaei, N. V. Hud and L. D. Williams, *Nucleic Acids Res.*, 2012, **40**, 3714-3722.
13. P. D. Dans, A. Pérez, I. Faustino, R. Lavery and M. Orozco, *Nucleic Acids Res.*, 2012, **40**, 10668-10678.
14. O. J. G. Somsen, A. van Hoek and H. van Amerongen, *Chem. Phys. Lett.*, 2005, **402**, 61-65.
15. C. S. M. Olsthoorn, L. J. Bostelaar, J. H. van Boom and C. Altona, *Eur. J. Biochem. / FEBS J.*, 1980, **112**, 95-110.
16. C. S. M. Olsthoorn, J. Doornbos, H. P. M. de Leeuw and C. Altona, *Eur. J. Biochem. / FEBS J.*, 1982, **125**, 367-382.
17. J. Gidden and M. T. Bowers, *Eur. Phys. J. D*, 2002, **20**, 409-419.
18. R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. a. Case, T. Cheatham, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova and J. Sponer, *Nucleic Acids Res.*, 2010, **38**, 299-313.
19. S. Jafilan, L. Klein, C. Hyun and J. Florián, *J. Phys. Chem. B*, 2012, **116**, 3613-3618.
20. P. Hobza and J. Šponer, *J. Am. Chem. Soc.*, 2002, **124**, 11802-11808.
21. J. Sponer, P. Jurecka, I. Marchan, F. J. Luque, M. Orozco and P. Hobza, *Chem. Eur. J.*, 2006, **12**, 2854-2865.
22. C. D. M. Churchill and S. D. Wetmore, *Phys. Chem. Chem. Phys.*, 2011, **13**, 16373-16383.
23. T. van Mourik and R. J. Gdanitz, *J. Chem. Phys.*, 2002, **116**, 9620.
24. V. R. Cooper, T. Thonhauser, A. Puzder, E. Schro, B. I. Lundqvist and D. C. Langreth, *J. Am. Chem. Soc.*, 2008, 1304-1308.
25. R. S. Hunter and T. van Mourik, *J. Comput. Chem.*, 2012, **33**, 2161-2172.
26. G. Barone, C. Fonseca Guerra and F. M. Bickelhaupt, *ChemistryOpen*, 2013, **2**, 186-193.
27. T. A. Zubatiuk, O. V. Shishkin, L. Gorb, D. M. Hovorun and J. Leszczynski, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18155-18166.
28. L. F. Holroyd and T. van Mourik, *Theor. Chem. Acc.*, 2014, **133**, 1431.
29. Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157-167.
30. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215-241.
31. Y. Zhao and D. G. Truhlar, *Chem. Phys. Lett.*, 2011, **502**, 1-13.
32. C. A. Morgado, P. Jurecka, D. Svozil, P. Hobza and J. Sponer, *Phys. Chem. Chem. Phys.*, 2010, **12**, 3522-3534.
33. T. van Mourik and H. S. W. L., *Struct. Chem.*, 2016, **27**, 145-158.
34. S. J. Hardman and K. C. Thompson, *Biochemistry*, 2006, **45**, 9145-9155.
35. J. M. Jean and K. B. Hall, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 37-41.
36. J. M. Jean and K. B. Hall, *Biochemistry*, 2002, **41**, 13152-13161.
37. J. Liang and S. Matsika, *J. Am. Chem. Soc.*, 2011, **133**, 6799-6808.
38. J. Liang and S. Matsika, *J. Am. Chem. Soc.*, 2012, **134**, 10713-10714.
39. J. Liang, Q. L. Nguyen and S. Matsika, *Photochem. Photobiol. Sci.*, 2013, **12**, 1387-1400.
40. V. I. Poltev, V. M. Anisimov, V. I. Danilov, A. Deriabina, E. Gonzalez, A. Jurkiewicz, A. Leś and N. Polteva, *J. Biomol. Struct. Dynam.,* 2008, **25**, 563-571.
41. V. I. Poltev, V. M. Anisimov, V. I. Danilov, A. Deriabina, E. Gonzalez, D. Garcia, F. Rivas, A. Jurkiewicz, A. Leś and N. Polteva, *J. Mol. Struct. THEOCHEM*, 2009, **912**, 53-59.
42. V. I. Poltev, V. M. Anisimov, V. I. Danilov, T. van Mourik, A. Deriabina, E. Gonzá Lez, M. Padua, D. Garcia, F. Rivas and N. Polteva, *Int. J. Quantum Chem*, 2010, **110**, 2548-2559.
43. V. I. Poltev, V. M. Anisimov, V. I. Danilov, D. Garcia, A. Deriabina, E. González, R. Salazar, F. Rivas and N. Polteva, *Comp. Theor. Chem.*, 2011, **975**, 69-75.
44. V. Poltev, V. M. Anisimov, V. I. Danilov, D. Garcia, C. Sanchez, A. Deriabina, E. Gonzalez, F. Rivas and N. Polteva, *Biopolymers*, 2013, **101**, 640-650.
45. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. B. Ogliaro, M., J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT*, 2009.
46. S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117-129.
47. A. Herráez, *Biochem. Mol. Biol. Educ.*, 2006, **34**, 255-261.
48. G. Schaftenaar and J. H. Noordik, *J. Comput. Aided Mol. Des.*, 2000, **14**, 123-134.
49. X.-J. Lu and W. K. Olson, *Nucleic Acids Res.*, 2003, **31**, 5108-5121.
50. X.-J. Lu and W. K. Olson, *Nat. Protoc.*, 2008, **3**, 1213-1227.
51. G. Zheng, X.-J. Lu and W. K. Olson, *Nucleic Acids Res.,* 2009, **37**, W240-W246.
52. J. Řezáč, K. E. Riley and P. Hobza, *J. Chem. Theory. Comput.*, 2011, **10**, 1359-1360.
53. C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205-8212.
54. R. E. Dickerson and H.-L. Ng, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 6986-6988.
55. J. Norberg and L. Nilsson, *Biophys. J.*, 1995, **69**, 2277-2285.
56. J. Norberg and L. Nilsson, *J. Am. Chem. Soc.*, 1995, **117**, 10832-10840.
57. Z. Vokáčová, M. Buděšínský, I. Rosenberg, B. Schneider, J. Šponer and V. Sychrovský, *J. Phys. Chem. B*, 2009, **113**, 1182-1191.
58. R. F. Brown, C. T. Andrews and A. H. Elcock, *J. Chem. Theory. Comput.*, 2015, **11**, 2315-2328.
59. Y. Yin, L. Yang, G. Zheng, C. Gu, C. Yi, C. He, Y. Q. Gao and X. S. Zhao, *Proc. Natl. Acad. Sci. USA*, 2014, **111**, 8043-8048.
60. L. F. Holroyd and T. van Mourik, *Chem. Phys. Lett.*, 2007, **442**, 42-46.
61. A. E. Shields and T. van Mourik, *J. Phys. Chem. A*, 2007, **111**, 13272-13277.
62. T. van Mourik, P. G. Karamertzanis and S. L. Price, *J. Phys. Chem. A*, 2006, **110**, 8-12.
63. S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553-566.

64. A. Dallmann, L. Dehmel, T. Peters, C. Mügge, C. Griesinger, J. Tuma and N. P. Ernsting, *Angew. Chem. Int. Ed.*, 2010, **49**, 5989-5992.
65. R. W. Sinkeldam, N. J. Greco and Y. Tor, *Chem. Rev.*, 2010, **110**, 2579-2619.
66. M. Sholokh, R. Sharma, D. Shin, R. Das, O. A. Zaporozhets, Y. Tor and Y. Mély, *J. Am. Chem. Soc.*, 2015, **137**, 3185-3188.

Tables

Table 1 Base-step parameters for computed B-form structures of d(2pN) and d(ApN) dinucleotides in comparison with values for idealised B-DNA and mean values from B-DNA crystal structures.

	Shift /Å	Slide /Å	Rise /Å	Tilt /°	Roll /°	Twist /°
B-DNA Ideal[a]	0.00	0.00	3.34	0.00	0.00	36.00
B-DNA Mean[b]	−0.02	0.23	3.32	−0.10	0.60	36.00
(Std. Dev.)	(0.45)	(0.81)	(0.19)	(2.50)	(5.20)	(6.80)
d(2pA)	0.84	−0.27	3.17	0.74	2.06	33.24
d(ApA)	1.32	−0.30	3.13	2.03	1.30	35.92
d(2pG)	0.72	−0.20	3.05	4.28	−0.73	30.12
d(ApG)	1.28	−0.31	3.06	3.35	1.29	34.62
d(2pC)	1.15	−0.39	3.06	2.53	0.00	34.38
d(ApC)	1.03	−0.44	2.96	3.89	4.08	34.20
d(2pT)	0.91	−0.30	2.91	5.59	−0.44	31.66
d(ApT)	1.37	−0.64	2.88	5.08	3.17	33.00

[a] From Lu and Olson.[49] [b] From Olson *et al.*[10]

Table 2 Backbone torsional angles (degrees) for computed B-form structures of d(2pN) and d(ApN) dinucleotides in comparison with mean values from B-DNA crystal structures. Standard deviations in the latter are small (<1.2) and are not shown. Torsional angles are defined in Figure 1.

	γ_1	δ_1	ε_1	ζ_1	α_2	β_2	γ_2	δ_2	χ_1	χ_2
B-DNA Mean[a]	48.4	132.8	−178.3	−96.8	−61.0	179.3	48.4	132.8	−109.7	−109.7
d(2pA)	50.2	146.0	−176.1	−86.2	−67.0	174.0	50.4	132.3	−109.4	−114.5
d(ApA)	50.7	145.8	−175.1	−84.7	−65.5	170.4	52.5	126.3	−110.8	−120.1
d(2pG)	49.6	144.8	−174.0	−85.8	−66.4	172.2	50.1	132.7	−110.0	−110.8
d(ApG)	50.4	145.8	−174.2	−84.3	−65.8	170.2	51.8	127.8	−111.5	−117.8
d(2pC)	50.3	145.6	−173.7	−84.4	−65.3	166.7	55.6	103.6	−109.7	−132.7
d(ApC)	50.2	145.2	−174.4	−84.8	−64.7	169.1	55.1	114.9	−110.7	−126.4
d(2pT)	49.8	144.7	−174.9	−85.3	−64.1	168.7	54.9	116.9	−110.0	−122.1
d(ApT)	51.0	146.7	−174.0	−82.6	−65.1	167.6	54.8	105.6	−110.6	−131.8

[a] From Svozil *et al.*[9]

Table 3 Sugar pucker parameters for computed B-form structures of d(2pN) and d(ApN) dinucleotides in comparison with those of ideal B-DNA. P is the phase angle of pseudorotation.[a]

	5'-Base		3'-Base	
	P /°	Pucker	P /°	Pucker
B-DNA Ideal[b]	144-180	C2'-endo	144-180	C2'-endo
d(2pA)	181.3	C3'-exo	134.8	C1'-exo
d(ApA)	180.1	C3'-exo	127.7	C1'-exo
d(2pG)	181.5	C3'-exo	135.0	C1'-exo
d(ApG)	181.1	C3'-exo	148.5	C2'-endo
d(2pC)	179.6	C2'-endo	101.6	O4'-endo
d(ApC)	180.6	C3'-exo	114.9	C1'-exo
d(2pT)	180.4	C3'-exo	117.1	C1'-exo
d(ApT)	178.7	C2'-endo	105.0	O4'-endo

[a] The range of P values corresponding to each conformational form is as follows: O4'-endo, 72-108°; C1'-exo, 108-144°; C2'-endo, 144-180°; C3'-exo, 180-216°. [b] From Altona and Sundaralingam.[53]

Table 4 Base-step parameters for computed twisted structures of d(2pN) and d(ApN) dinucleotides in comparison with values for idealised B-DNA and mean values from B-DNA crystal structures.

	Shift /Å	Slide /Å	Rise /Å	Tilt /°	Roll /°	Twist /°
B-DNA Ideal[a]	0.00	0.00	3.34	0.00	0.00	36.00
B-DNA Mean[b]	−0.02	0.23	3.32	−0.10	0.60	36.00
(Std. Dev.)	(0.45)	(0.81)	(0.19)	(2.50)	(5.20)	(6.80)
d(2pA)	1.53	−0.13	3.14	0.46	−0.32	60.60
d(ApA)	1.21	0.05	3.24	−2.49	0.43	51.53
d(2pG)	1.46	−0.18	3.07	4.74	−4.05	59.12
d(ApG)	1.45	−0.11	3.19	0.53	−3.59	50.13
d(2pC)	0.49	−1.54	3.68	−8.75	−1.19	45.61
d(ApC)	−0.13	−1.17	3.67	−9.10	6.62	40.42
d(2pT)	−0.14	−1.55	3.43	−2.13	−5.92	38.81
d(ApT)	−0.70	−1.20	3.22	0.79	−3.17	33.81

[a] From Lu and Olson.[49] [b] From Olson *et al.*[10]

Table 5 Backbone torsional angles (degrees) for computed twisted structures of d(2pN) and d(ApN) dinucleotides in comparison with mean values from B-DNA crystal structures. Standard deviations in the latter are small (<1.2) and are not shown. Torsional angles are defined in Figure 1.

	γ_1	δ_1	ε_1	ζ_1	α_2	β_2	γ_2	δ_2	χ_1	χ_2
B-DNA Mean[a]	48.4	132.8	−178.3	−96.8	−61.0	179.3	48.4	132.8	−109.7	−109.7
d(2pA)	55.2	149.0	156.0	−87.4	−120.3	−113.4	48.9	93.1	−129.5	−114.1
d(ApA)	52.9	149.3	174.7	−97.9	−81.5	−174.7	48.5	85.6	−106.4	−140.7
d(2pG)	55.2	147.9	154.6	−86.2	−120.9	−108.6	48.4	92.5	−133.9	−110.3
d(ApG)	52.8	148.7	172.1	−95.8	−76.7	−178.2	49.7	84.5	−104.1	−141.1
d(2pC)	56.9	125.4	82.1	−55.6	−118.3	−88.1	52.9	75.8	−86.3	−132.5
d(ApC)	56.6	125.6	79.9	−55.3	−116.6	−90.1	52.6	78.0	−77.7	−127.6
d(2pT)	51.0	128.8	79.7	−54.5	−118.7	−87.1	52.5	76.5	−84.0	−122.5
d(ApT)	51.4	129.4	77.3	−53.3	−116.0	−89.9	51.7	78.8	−78.0	−115.2

[a] From Svozil *et al.*[9]

Table 6 Relative potential energies (ΔE) and Gibbs free energies (ΔG) of the B-form (B), twisted (T) and unstacked (U) conformations of each dinucleotide. In each case, the energy is given relative to the lowest energy conformation.

	ΔE /kcal mol^{-1}			ΔG /kcal mol^{-1}		
	B	T	U	B	T	U
d(2pA)	1.4	0.0	-	0.0	0.8	-
d(ApA)	0.9	0.0	-	0.0	0.0	-
d(2pG)	0.4	0.0	10.9	0.0	2.0	5.4
d(ApG)	0.4	0.0	10.5	0.3	0.0	4.2
d(2pC)	1.0	0.0	-	0.0	1.0	-
d(ApC)	0.0	1.0	-	0.0	2.6	-
d(2pT)	0.7	0.0	10.3	0.0	2.9	5.0
d(ApT)	0.0	1.2	9.9	0.0	3.3	3.9

Figures

Figure 1 Schematic structure of the d(2pN) dinucleotides. Torsional angles for the sugar-phosphate backbone (from γ_1 to δ_2 along the backbone, 5′ to 3′) and the two glycosidic bonds (χ_1 and χ_2) are defined. Two H-bonding motifs (HB1 and HB2) found in twisted dinucleotide structures are shown.

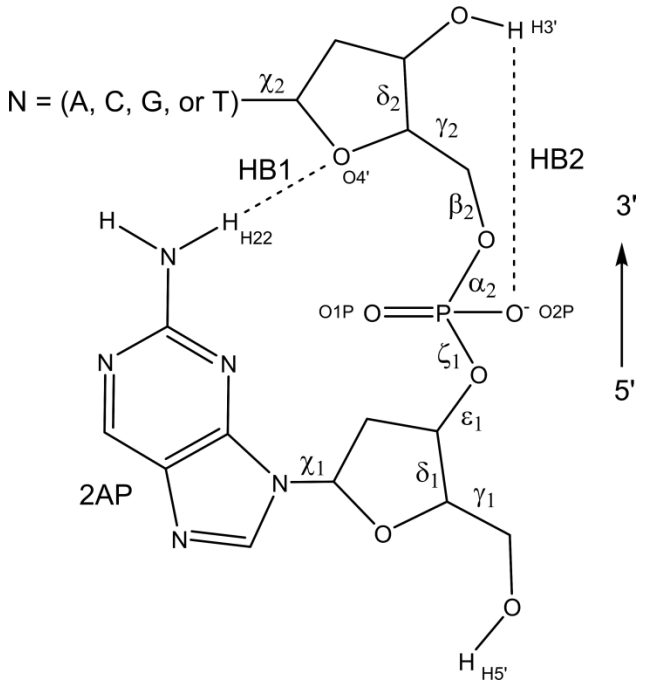

Figure 2 Optimised B-form structures of (a) d(2pN) and (b) d(ApN) dinucleotides. In each case two alternative views are shown: perpendicular to the backbone (top row) and looking along the backbone from the 5′ end (bottom row).

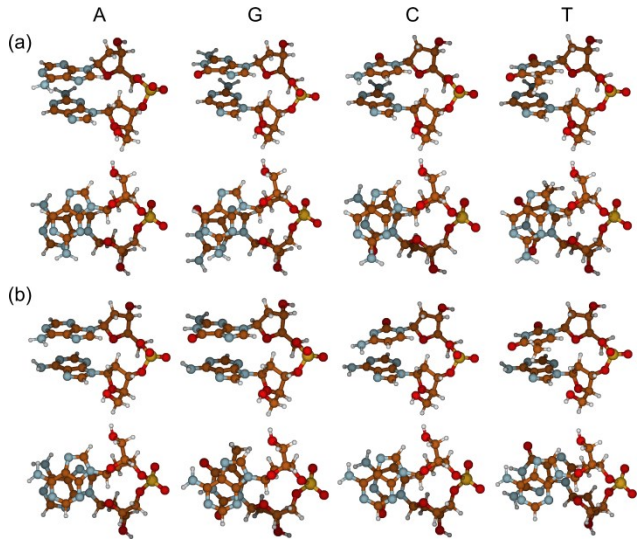

Figure 3 Optimised twisted structures of (a) d(2pN) and (b) d(ApN) dinucleotides. In each case two alternative views are shown: perpendicular to the backbone (top row) and looking along the backbone from the 5′ end (bottom row).

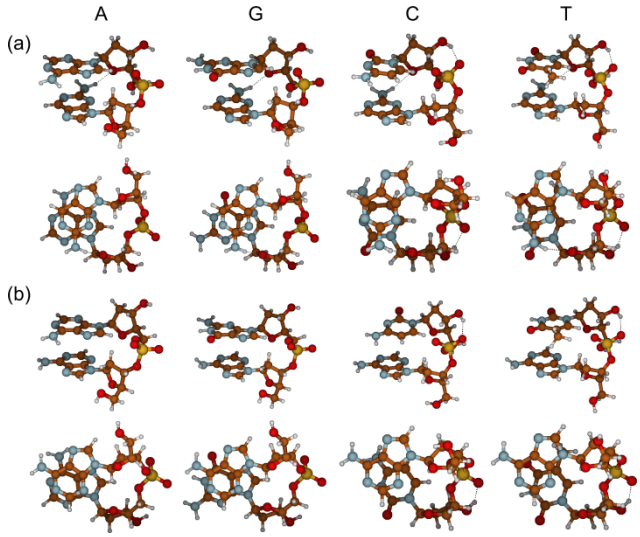

Figure 4 Optimised unstacked structures of (a) d(2pN) and (b) d(ApN) dinucleotides. Two views are shown in each case.