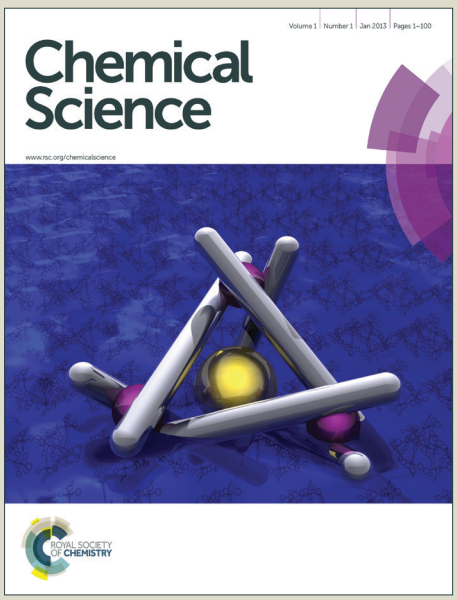
Chemical Science

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemicalscience

Journal Name

RSCPublishing

ARTICLE

Enhanced H-bonding and π-stacking in DNA: A potent duplex-stabilizing and mismatch sensing nucleobase analogue

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Chenguang Lou,[a] Andre Dallmann,[c] Pietro Marafini,[a,b] Rachel Gao[a] and Tom Brown[a,b]*

X-pyrene is a new nucleic acid duplex stabilizing cytosine analogue that combines enhanced π-stacking, hydrogen bonding and electrostatic interactions to greatly increase the stability of bulged DNA duplexes and DNA/RNA hybrids. X-pyrene is highly selective for guanine as a partner and duplex stability is reduced dramatically when X-pyrene or a neighboring base is mismatched. An NMR study indicates that the pyrene moiety stacks within the helix, and large changes in fluorescence emission on duplex formation are consistent with this. These favorable properties make X-pyrene a promising cytosine analogue for use in a variety of biological applications.

Introduction

Hydrogen-bonding and base stacking interactions between Watson-Crick G:C and A:T base pairs control replication and transcription and are fundamental to applications such as DNA sequencing and genetic analysis.[1,2-5] Various chemically modified DNA and RNA probes have been used to increase duplex stability and discriminate against mismatches;[6-8] and this has been achieved *via* artificial backbones (PNA, LNA and UNA),[8-11] modified nucleobases[8, 12-15] and increased π-stacking interactions.[16-18] Nucleobase analogues can be incorporated during solid-phase oligonucleotide synthesis to provide a site-specific method of increasing DNA duplex stability.[19] The best example is G-clamp, a cytosine analogue which specifically recognizes guanine and stabilizes duplexes *via* a network of four-hydrogen bonds.[14, 20] Oligonucleotides containing G-clamp and analogues have increased nuclease resistance, improved cellular penetration, and enhanced antisense activity.[21-23] Stabilization by G-Clamp has been found to vary widely depending on the neighboring sequence.[24]

Nucleic acid duplex stability can also be enhanced by intercalation of large aromatic/heterocyclic molecules, and in this context pyrene[25-26] has been shown to provide extra stability by strong π-stacking. [17,27] It can be used to promote site-specific targeting of enzymes such as uracil and cytosine DNA glycosylases and DNA methyltransferases. [28-31] In the present study we have enhanced the duplex stabilizing and mismatch discrimination properties of G-clamp and pyrene by combining them into a single cytosine analogue (X-pyrene, Figure 1A) with resultant highly favorable H-bonding, electrostatic and π-stacking interactions (Figure 1B). We demonstrate the ability of X-pyrene to form very stable duplexes with DNA and RNA strands, particularly when an unpaired nucleobase is located nearby in the complementary strand. Such bulged duplexes occur in biology[32] and can be engineered into any duplex at will. The mismatch sensing and fluorescent properties of X-pyrene oligonucleotides are also investigated.

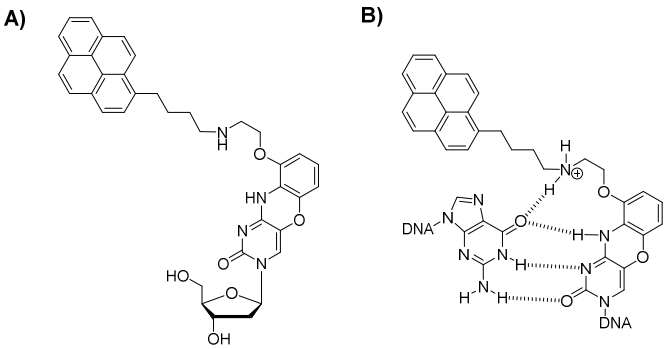


Figure 1. A) The X-pyrene nucleoside. **B)** Base pairing with guanine in DNA. Hydrogen bonds are shown as dashed lines.

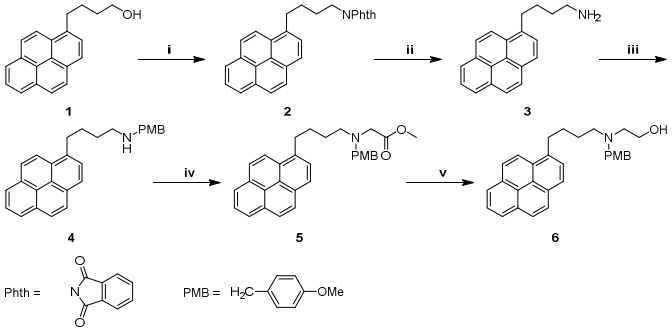
Results and discussion

Synthesis of X-pyrene phosphoramidite monomer and oligonucleotides

In principle it would be possible to conjugate commercially available pyrene butyric acid directly to the tethered primary amine of the G-clamp nucleoside to produce a pyrene G-clamp

monomer. However, the amide linkage might restrict the conformational flexibility and prevent pyrene from finding an optimum site for intercalation, and an undesirable intramolecular H-bond could form between the amide oxygen and the NH proton on the attached nucleobase. Either of these effects could reduce duplex stabilization,[14, 33-35] so instead we elected to join pyrene to G-clamp *via* a secondary amino group.

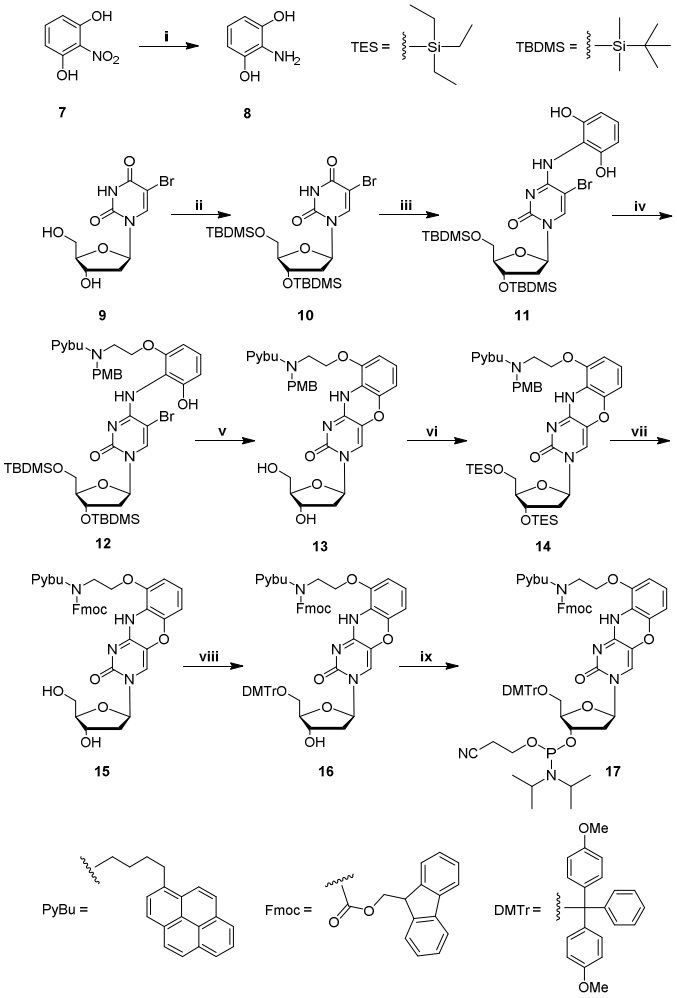
Synthesis of the X-pyrene phosphoramidite monomer was accomplished starting from 4-(1-pyrenyl)-1-butanol and 5-bromo-2′-deoxyuridine (Scheme **1** and **2**). The hydroxyl group of 4-(1-pyrenyl)-1-butanol (**1**) was quantitatively displaced by phthalimide under Mitsunobu conditions to yield the 1-phthalimidobutyl derivative **2**. This was treated with hydrazine monohydrate to produce the primary amine **3** which was then mono-functionalized with *p*-methoxybenzyl chloride (PMBCl) to yield protected pyrene-butylamine **4**. The low yield for this step was due to uncontrolled benzylation, producing a large amount of dibenzylated product. This secondary amine was alkylated using methyl bromoacetate to give compound **5**, which was reduced with $LiBH_4$ to generate the desired compound **6** with a free hydroxyl group, ready for conjugation to the nucleoside moiety (Scheme **1**).



Scheme 1. *Reagents and conditions:* (i) phthalimide, PPh_3, DEAD, THF, RT, 1 h; (ii) $H_2NNH_2{\cdot}H_2O$, EtOH/DCM (1:1), RT, 72 h, 87 % for two steps from **1** to **3**; (iii) PMBCl, NaH, DMF, 0 ˚C, 24 h, 43 %; (iv) methyl bromoacetate, NaH, DMF, 0 ˚C, 1 h; (v) $LiBH_4$, THF, RT, 72 h, 49 % for two steps from **4** to **6**.

Synthesis of nucleoside **13** (Scheme 2) followed a similar route to that of Matteucci.[13-14] The exocyclic nitro group of 2-nitroresorcinol (**7**) was reduced to give the aniline derivative (**8**) by Pd/C-catalyzed hydrogenation, and 5-bromo-2′-deoxyuridine (**9**) was protected by reaction with *tert*-butyldimethylsilyl chloride (TBDMSCl) to produce compound **10**. The uracil base of **10** was activated using PPh_3/CCl_4[36-37] and reacted with 2-aminoresorcinol (**8**) in the presence of DBU to yield compound **11**. One of the two phenolic hydroxyl groups of **11** was coupled to the hydroxyl group of compound **6** under Mitsunobu conditions to produce the desired conjugate **12**, which was desilylated and cyclized using potassium fluoride and ethanol under reflux to give compound **13**.[13, 38] Hydrogenolysis was unsuccessful in removing the PMB group from nucleoside **13** and generated a complex mixture. Fortunately, after the 3′- and 5′- positions of deoxyribose were protected as triethylsilyl ethers, PMB could be removed from compound **14** by selective debenzylation using α-chloroethyl chloroformate in DCM followed by refluxing in methanol.[39] This generated the desired fully deprotected nucleoside which was treated *in situ* with Fmoc-OSu in pyridine to give **15**.[40] The deoxyribose 5′-hydroxyl group was then selectively protected by reaction with 4,4′-dimethoxytrityl chloride in pyridine to give **16**, which was finally phosphitylated to produce X-pyrene phosphoramidite monomer **17**. Although the synthesis of **17** is lengthy and low-yielding we were able to produce sufficient quantities to make a series of X-pyrene oligonucleotides and thoroughly investigate their properties.

The synthesis of oligonucleotides containing X-pyrene was carried out using standard solid-phase phosphoramidite chemistry. Fmoc protection of the side-chain amino group of X-pyrene is compatible with the standard oligonucleotide synthesis cycle, including capping with acetic anhydride, and it permits oligonucleotide deprotection under mild conditions.[40-41] Full oligonucleotide synthesis and purification protocols are provided in the supporting information (S4 Oligonucleotide synthesis, purification and analysis).



Scheme 2. *Reagents and conditions:* (i) Pd/C, H_2, MeOH, 50 ˚C, 12 h, 93 %; (ii) TBDMSCl, imidazole, DMF, RT,1.5 h, 94 %; (iii) PPh_3, CCl_4, DCM, reflux, 16 h; **8**, DBU, DCM, RT,18 h, 47 % for two steps; (iv) **6**, PPh_3, DEAD, DCM, RT, 2 h, 81 %; (v) KF, ethanol, 82 ˚C, 24 h, 54 %; (vi) TESCl, imidazole, DMF, RT, 24 h, 86 %; (vii) α-chloroethyl chloroformate, DCM, RT, 1 h; methanol, reflux, 2 h; Fmoc-OSu, pyridine, RT, 0.5 h, 48 % for three steps; (viii) DMTrCl, pyridine, RT, 6 h, 74 %; (ix) 2-cyanoethyl-*N*,*N*-diisopropyl chlorophosphine, DIPEA, DCM, 0.5 h, RT, 79 %.

Biophysical studies on X-pyrene oligonucleotides

Two types of duplex were constructed from X-pyrene oligonucleotides and their complementary strands (Table 1, Figure 2A); normal fully-paired duplexes (ND) and bulged duplexes (BD) which have an unpaired nucleobase in the

complementary strand opposite to X-pyrene (Figure 2A). The inclusion of this additional unpaired nucleobase primes the duplex for pyrene intercalation by creating a pocket of non-optimal base stacking. Initial studies indicated that this is effective for the stabilization of 3′-bulged duplexes, in which the unpaired base is in the opposite strand to X-pyrene on the immediate 3′-side. In contrast, 5′-bulged duplexes containing X-pyrene were found to be less stable in both DNA/DNA and DNA/RNA contexts (Figures S3, S4, S5, S6). Steric constraints imposed by the linker between pyrene and the attached nucleobase may contribute to the lower stability of 5′-bulged duplexes relative to 3′-bulged duplexes (Figure S3 and Figure S4). If so, a longer tether between X-pyrene and G-clamp might allow the pyrene moiety to stabilize 5′-bulged duplexes to a greater degree. In subsequent studies we focused on the very stable 3′-bulged duplexes, and BD in Table 1 refers to these constructs.

Bulged duplexes containing unpaired pyrimidine bases and single (BD1, BD2), or double X-pyrene units (BD3, BD4), were synthesized to investigate the duplex-stabilizing properties of X-pyrene. The equivalent control 2′-deoxycytidine (C) and G-clamp bulged duplexes were used for comparison. In the doubly modified oligonucleotides, X-pyrene was incorporated in two different arrangements; either clustered in the centre (BD3) or dispersed (BD4). To evaluate the duplex-stabilizing property of X-pyrene in different sequence contexts, several different nucleobase combinations were chosen to flank X-pyrene; CYC, AYC and AYA as single modifications and CYA as a double modification (Figure 2A, Figure S3).

Table 1. Oligonucleotides used in biophysical studies.
ND = normal fully-paired duplex. BD = bulged duplex. Y = X-pyrene, C or G-clamp.

	Code	Y	Sequence	Duplex	Complementary Strands
1	ODN-1	X-pyrene	5′-GCAC**Y**CTACG	ND	5′-CGTAGGGTGC
	ODN-2	G-clamp			5′-CGUAGGGUGC (RNA)
				BD	5′-CGTAG**C**GGTGC
	ODN-3	C			5′-CGUAG**C**GGUGC (RNA)
2	ODN-4	X-pyrene	5′-CGACA**Y**CACAGC	ND	5′-GCTGTGGTGTCG
	ODN-5	G-clamp		BD	5′-GCTGTG**T**GTGTCG
	ODN-6	C			
3	ODN-7	X-pyrene	5′-GCAC**YY**ATCG	BD	5′-CGAT**T**G**T**GGTGC
	OND-8	C		ND	5′-CGATGGGTGC
4	ODN-9	X-pyrene	5′-GC**Y**ATAC**Y**AG	BD	5′-CT**T**GGTAT**T**GGC
	ODN-10	C		ND	5′-CTGGTATGGC

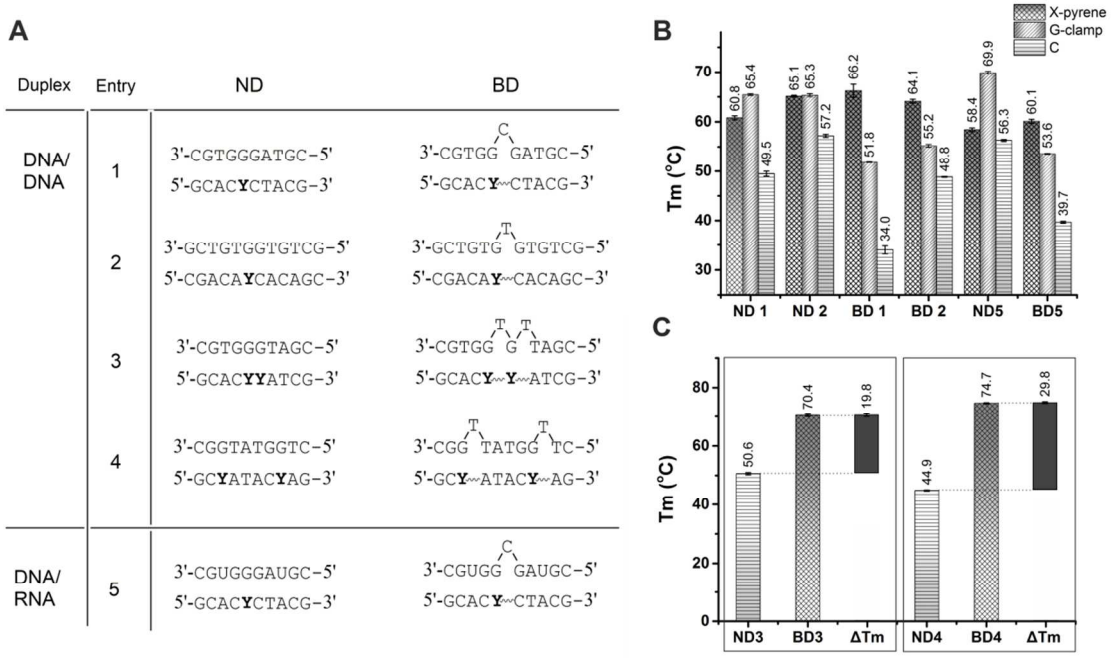


Figure 2. X-pyrene produces very stable bulged duplexes. A: Each ND duplex in column 1 has a corresponding BD duplex in the same row in column 2. Complementary RNA strands were used in Entry 5. ND = normal fully paired duplexes, BD = bulged duplexes. Y = X-pyrene, C or G-clamp. **B, C:** Duplex stability. In **2C**, Y = C (ND3, ND4) or X-pyrene (BD3, BD4). The ultraviolet melting curves were recorded at 260 nm. T_m values are an average of three measured melting temperatures (°C). The experiments were performed in 5 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 7.2, containing 140 mM KCl, 1 mM $MgCl_2$, oligonucleotide strand concentration = 2.5μM).

A single X-pyrene unit produced far greater duplex stability in the BD1 and BD2 sequence contexts than G-clamp or C (Figure 2B), presumably due to intercalation of the pyrene moiety. In the case of the BD1 duplex (Figure 2B, Table S2) a huge increase in melting temperature was observed relative to C and G-clamp (ΔT_m = +32.2 ˚C and +14.4 ˚C respectively) and the same trend was recorded for the longer BD2 duplex (ΔT_m = +15.3 ˚C and +8.9 ˚C respectively, Figure 2B, Table S2). Previous studies on the sequence dependence of G-clamp duplex stabilization revealed that CYC and AYA are the two extreme neighboring environments, giving huge variations in melting temperature (T_m(CYC) – T_m(AYA) = 19.0 ˚C).[24] Interestingly, when X-pyrene is embedded into bulged duplexes, a lower degree of sequence dependence is observed, and the difference between CYC and AYA is reduced to 12.7 ˚C (Figure S7). This relaxation of sequence-dependence is an advantage when contemplating the design and use of X-pyrene in genetic probes.

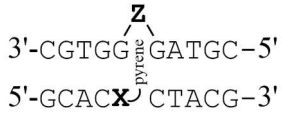
The lack of an unpaired nucleobase in the fully complementary strand of normal duplexes appears to disfavor intercalation of X-pyrene, as without a hydrophobic pocket the pyrene may become more exposed to the aqueous environment, and its stabilizing effect will be lost. Consequently ND1 with X-pyrene is less stable than with G-clamp, and for ND2 the X-pyrene and G-clamp modifications confer equal stability.

The trends in X-pyrene stabilization of bulged DNA/RNA hybrid duplexes are similar to those observed for bulged DNA duplexes. When X-pyrene oligonucleotide (ODN-1) was hybridized to an RNA strand yielding BD5 (Figure 2B), it produced substantially higher duplex stability than its natural and G-clamp counterparts, with ΔT_m values of + 20.4 ˚C and + 6.5 ˚C respectively. In addition, as noted for bulged DNA duplexes, 3′-intercalation is more stabilizing than 5′-intercalation. (Figure S5 and Figure S6). In the non-bulged duplex (ND5 in Figure 2) X-pyrene conferred greater stability than C (+ 2.1 ˚C) but was inferior to G-clamp (− 11.5 ˚C).

Clearly it is important that multiple additions of a duplex-stabilizing monomer produce an enhanced effect. In this context, and unlike G-clamp, the effect of X-pyrene on duplex stability is cumulative (Figure 2C); two X-pyrenes give a greater increase in duplex stability than a single modification. The disposition of the dual X-pyrene units in the DNA strand has an additional influence on duplex stability, dispersed modifications (ΔT_m +29.8 ˚C) being more stabilizing than clustered (ΔT_m +19.8 ˚C) when compared to unbulged duplexes with normal G:C base pairs (Figure 2C). The fact that even clustered X-pyrene units can strongly stabilize DNA duplexes is interesting, considering that the structure has three base pairs interspersed with two unpaired bases.

Effects of different unpaired nucleobases opposite to X-pyrene:

In addition to studies with C or T as the unpaired base in the bulged duplexes (Figure 2) the BD1 duplex was also evaluated with an unpaired purine base and also an abasic analogue (tetrahydrofuran) in this position. The results show that X-pyrene can tolerate all of these, with a ΔT_m of 4.5 ˚C between cytosine, which is the most stable, and adenine which is the least stable (Figure 3). This simplifies the design of probes containing X-pyrene. However, another design aspect should be considered. In certain cases BD duplexes have similar stability to ND duplexes (e.g. BD2 and ND2, Figure 2B) so when designing probes for biological applications this should be taken into account.



	BD1	BD1-T	BD1-G	BD1-A	BD1-0
Z	C	T	G	A	abasic site
T_m(˚C)	66.2	64.4(−1.8)	61.9 (−4.3)	61.7 (−4.5)	62.5 (−3.7)

Figure 3. X-pyrene with different nucleobases in the unpaired position in bulged duplex (BD1). Z indicates A, G, C, T or an abasic site (tetrahydrofuran, 0). The table gives duplex T_m values; values in parentheses are ΔT_m = T_m(Z = T, G, A, 0) – T_m(Z =C). Experiments were carried out under the conditions described in Figure 2.

NMR studies on DNA duplexes containing X-pyrene

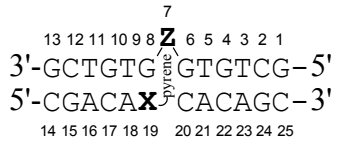


Figure 4. Bulged duplexes studied by NMR. Z = T (BD2) and Z = C (BD2-C). The unpaired base can loop out or intercalate.

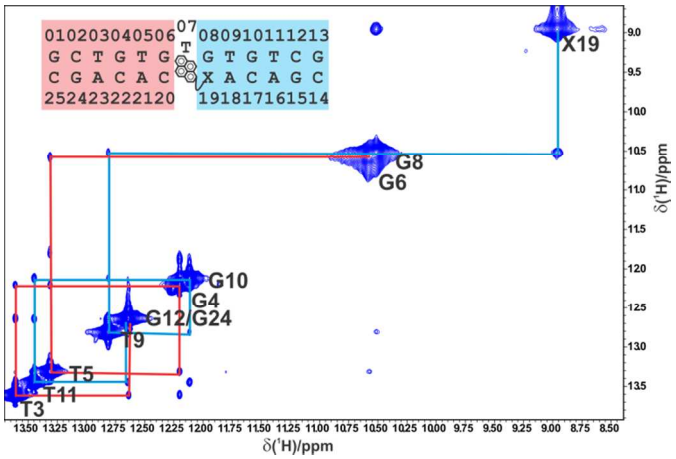


Figure 5. Imino-imino region of the homonuclear NOESY of BD2 at 310 K. Assignments are shown for the duplex 3′ (red) and 5′ (blue) of the pyrene. For the looped out base as well as for the linker region no imino signal is observed due to increased exchange with water. The pronounced shift of the G6 and G8 imino protons is probably due to ring current effects from the pyrene and indicates that the pyrene moiety is stacking between these two residues.

NMR studies were carried out to examine the effects of X-pyrene on DNA structure. Two duplexes were analyzed, BD2 and the equivalent sequence with C instead of T as the widowed (unpaired) base (BD2-C in Figure 4). A combination of ^{1}H-^{1}H-NOESY and natural abundance ^{1}H-^{15}N-SOFAST-HMQC in H_2O/D_2O 9:1 was used to assign all observable exchangeable protons and aromatic H2 protons of adenine (Figure 5). Sequential imino-imino connectivity on both sides of the pyrene intercalation site was observed, indicative of regular B-DNA geometry and stable base pair formation.

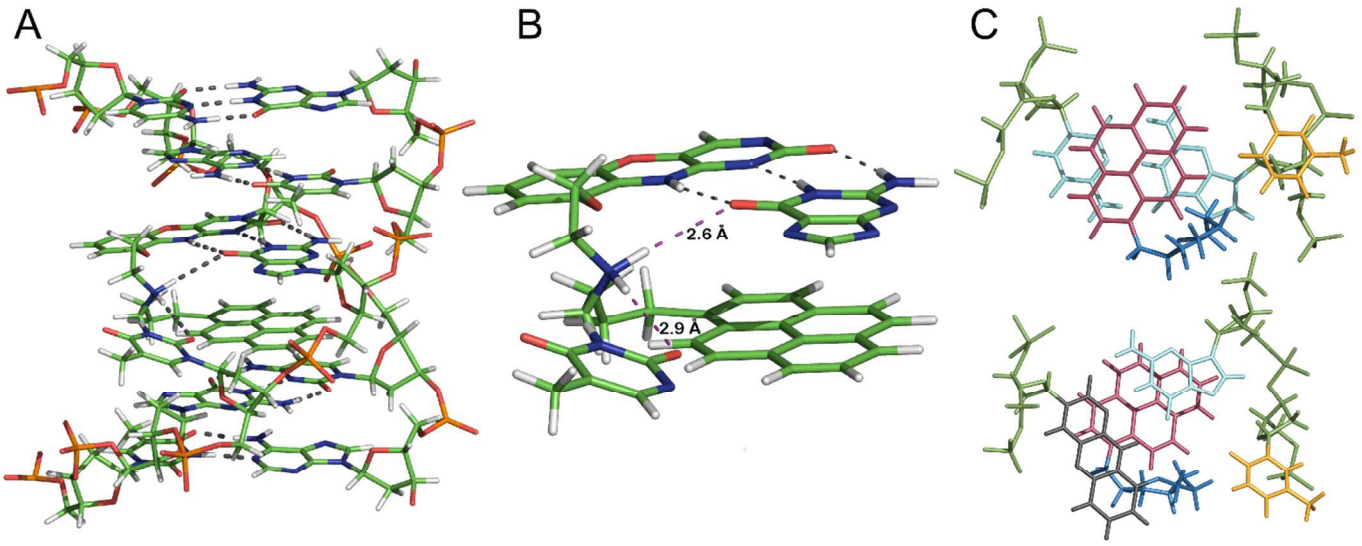


Figure 6. X-pyrene is compatible with B-DNA; molecular modelling of BD2. **A)** Full duplex with hydrogen bonds highlighted; **B)** Close-up of X-pyrene region showing hydrogen bond pattern of the linker. Shown in purple are the H-bond from the linker to the Hoogsteen face of the G base paired with X-pyrene and the H-bond from the linker to O2 of the unpaired T; **C)** Close-up highlighting the stacking of pyrene with the base pair below it (top) and above it (bottom), and stacking of the widowed thymine on the deoxyribose sugar. Pyrene is in wine red, the unpaired base is in yellow, the linker chain is in blue, the modified nucleobase moiety (tC^O) is in dark grey, the backbones are in green and the natural bases are in light blue.

Interestingly, the imino protons exhibited no shift upon changing the bulged residue 7 from T to C, indicating that the nature of this nucleobase does not affect the adjacent base pair geometry (Figure S20). This is consistent with the UV-melting studies. The remarkable upfield chemical shift of the G6 and G8 imino protons (1.5-2 ppm compared to expected range, Table S7) can be explained by the strong ring current effects induced by stacking of the pyrene moiety between the two surrounding base pairs. Stacking of the pyrene also explains the lack of cross-peaks between the G6 and G8 imino protons. In case of Z = T, the imino proton was exchange-broadened beyond detection, confirming as expected that no stable base pair is formed with this base. The 0.5 ppm upfield shift of the T9 imino proton is notable as well, and indicates tighter stacking with the G8:X19 due to steric pressure from pyrene intercalation and ring current effects from the extended aromatic surface of X compared to C.

For assignment of the non-exchangeable protons in BD2, a combination of ^{1}H-^{1}H-COSY-, ^{1}H-^{1}H-TOCSY and ^{1}H-^{1}H-NOESY in 99.98 % D_2O was used. The NOESY spectrum showed the expected connectivity pattern typical of B-DNA,[42] indicating that no major disruption of the helical arrangement takes place. However, weaker than expected peaks to the 3′-side of pyrene at the intercalation site indicate that stacking of the pyrene moiety between the base pairs induces slight deviations from regular geometry.

Stacking of the pyrene moiety between base pairs G8:X19 and G6:C20 is indicated by the observation of several NOE cross-peaks between protons of the adjacent DNA residues and protons from the pyrene moiety. Due to the overlap and high number of similar aromatic protons within the pyrene moiety, unambiguous assignment of the pyrene aromatic protons was not possible, and thus the exact orientation of pyrene within the duplex was not defined. The methylene protons of the linker chain could not be assigned either, due to its flexibility and positioning outside the helix. However, unusually strong downfield shifts of protons of adjacent residues (Table S7) clearly position the pyrene between base pairs G8:X19 and G6:C20.

Based on the NMR data, we developed a model for the NMR structure (Figure 6). Surprisingly, the unpaired T7 residue is not completely looped out, but is merely pushed out towards the phosphate backbone (Figure 6 C). Although not stabilized by *base* stacking interactions in this position, it is stacked on the adjacent deoxyribose sugar, and is capable of forming a hydrogen bond with the linker amino proton. This is postulated from the model, but the inherent flexibility of the linker makes the H-bond too transient to be observed by NMR spectroscopy (water exchange broadens the amino proton beyond detection limits).

Notably, at low temperature (278 K) a second set of imino protons is observed for the X-pyrene and the adjacent base pairs, which is in slow (millisecond timescale) exchange with the major set of peaks (Figure S20). This suggests a second conformation that the pyrene moiety and unpaired base can adopt. Severely broadened, low-intensity peaks in the NOESY in D_2O at 298 K and even at 310 K suggest that this second conformation is still present at higher temperature. However, due to lower populations and exchange broadening, the corresponding imino protons are below the detection limit. More sophisticated temperature dependence experiments were prohibited by the resonance overlap with the major conformation.

Overall the NMR analysis confirms intercalation of the pyrene moiety into the DNA helix, and somewhat surprisingly, just a mild displacement of the widowed base over the phosphate backbone, rather than complete looping out. Consistent with NMR analysis, the CD spectra of bulged and normal duplexes indicate that the presence of X-pyrene does not severely disturb the structure of the B-DNA double helix (Figure S19).

X-pyrene strongly destabilizes mismatched duplexes

The mismatch discrimination ability of X-pyrene was studied by varying its nucleobase partner (Figure 7). To make a fair comparison, matched and mismatched duplexes containing X-pyrene were in the BD1 form, while control duplexes

containing G-clamp or C were in the ND1 form. This allowed us to study each different monomer in its most favorable conditions. Both X-pyrene and G-clamp showed much greater differences than C between their correct match with guanine and mismatches with A, T and C. This is probably because the fourth hydrogen-bond to the Hoogsteen face of the target guanine requires a Watson-Crick base pair, as does efficient π-stacking.[14] Unlike G-clamp and C, X-pyrene gave an almost uniform huge decrease in stability ($\Delta T_m \approx -35$ °C) regardless of the mismatched partner. This confirms the exquisite selectivity of X-pyrene for guanine, which could be valuable in genetic probe and related applications.

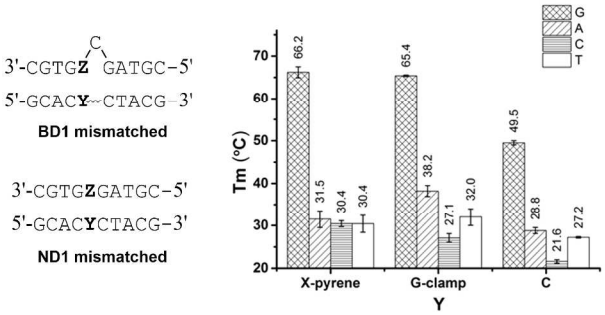


Figure 7. X-pyrene gives rise to efficient mismatch discrimination; comparison of matched and mismatched duplex stability. In ND1, Y = G-clamp or C, in BD1, Y = X-pyrene. The varied nucleobase Z and its partner are marked in bold. UV-melting experiments were performed under the conditions described in Figure 2.

X-pyrene senses remote mismatches

The extreme duplex stabilizing effect of X-pyrene in bulged duplexes may derive from a synergistic effect of three key structural features; i) the protonated exocyclic secondary alkylamine which forms electrostatic interactions and an additional major groove hydrogen bond to O6 from the Hoogsteen side of guanine, ii) the substituted phenoxazine ring which provides an extended surface to enhance base stacking, and iii) the pyrene moiety which acts as a DNA intercalator to provide strong π-stacking with adjacent base pairs. This should cause bulged duplexes with X-pyrene to be hyper-sensitive to any mismatch in the neighborhood, even if they do not directly involve X-pyrene. To test this hypothesis we studied the effect of mismatched base pairs surrounding the X-pyrene:G base pair. This remote mismatch sensing ability of X-pyrene was studied in the BD1 model duplex, and for comparison, G-clamp and C were evaluated in the normal duplex ND1 (Figure 8 and Table S5).

When the mismatch is located at position 2 or 3, X-pyrene confers greater mismatch discrimination than C or G-clamp. For G-clamp in ND1, the hot spots for mismatch destabilization are the two adjacent base pairs (positions 3 and 4). In contrast, in the X-pyrene bulged duplex, the AA mismatch at position 2 is extremely destabilizing ($\Delta T_m = -38.5$ °C), whereas in the normal duplexes, ΔT_m is only −15.4 °C for G-clamp and −17.1 °C for C. When the mismatch is placed at any of the other four positions in X-pyrene BD1 (1, 4, 5 and 6), the ΔT_m values are more similar to G-clamp and C, except that G-clamp has a greater influence at position 4, directly to its 3′-side. Overall these results indicate that the geometry of the two Watson-Crick base pairs on the 5′-side of X-pyrene play a major role in maintaining the stable core structure. The observation that bulged duplex stabilization by X-pyrene requires a Watson-Crick pairing two base pairs away on its 5′-side suggests that the structure might be complex and dynamic. This is not surprising, indeed this was indicated in the NMR analysis where multiple conformations were observed. Whatever the structural basis, it is clear that in addition to conferring very high duplex stability, X-pyrene shows great potential in short probes as an immediate and remote mismatch sensor.

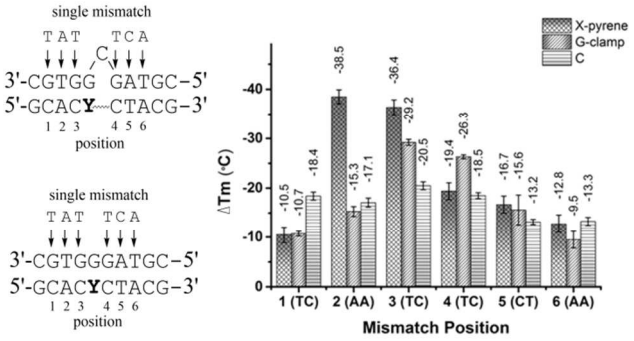


Figure 8. Efficient remote mismatch sensing with X-pyrene (BD1); comparison with G-clamp and C (ND1). A single mismatch was introduced at positions 1-6 of BD1 and ND1. Y = G-clamp or C or X-pyrene. The bar chart shows ΔT_m vs mismatch position where $\Delta T_m = T_m$(mismatch) – T_m(match).

Fluorescence studies

In addition to being a useful indicator of structural changes, fluorescence is a very useful property in biochemical assays in the nucleic acid field. Decreased emission of light from a fluorescent oligonucleotide upon binding to its complementary DNA strand is an indication of strong stacking, for example through intercalation into the double helix, [43-45] which leads to quenching by surrounding base pairs.[43, 46] Both X-pyrene and G-clamp contain a tC^O (1,3-diaza-2-oxophenoxazine) moiety, which is fluorescent. It has a high quantum yield and extinction coefficient and is a bright fluorophore.[47-48] For these reasons, fluorescence studies were carried out to investigate the emission properties of the tC^O moiety. It is important to note that, by exciting at 344 nm, pyrene is also excited. However, its fluorescence is masked by tC^O emission.

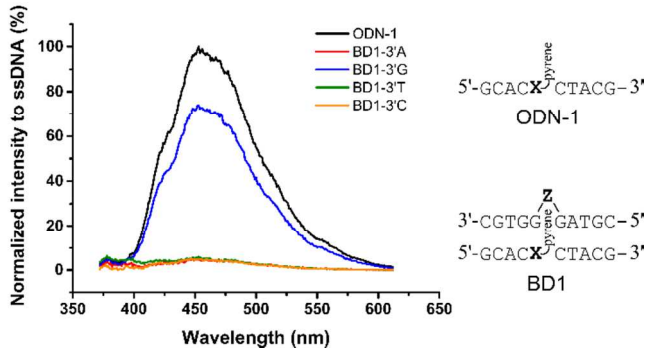


Figure 9. Fluorescence is strongly quenched in X-pyrene duplexes Fluorescence measurements were performed on 10-mer X-pyrene modified oligonucleotide (ODN-1, black) and its bulged duplexes with complementary strands which have different nucleobases (Z) in the unpaired position (BD1-G, BD1-A, BD1-C, BD1-T). 0.15 μM of probe strand containing X-pyrene and 0.2 μM of complementary strand in a phosphate buffer with 5 mM sodium phosphate, 140 mM KCl, 1 mM $MgCl_2$, pH 7.2.

In order to determine the effects of the unpaired base on the fluorescence, bulged duplexes with one strand modified with X-pyrene were used (ODN-1, BD1). A large decrease in

fluorescence (~ 95 %) was observed for bulged A, C and T, but not for bulged G, where the reduction was only about 25 % (Figure 9). The different behavior of bulged guanine is probably caused by the presence of a G-rich core that leads to stable G-quadruplex formation.[49-50] Consequently, under the experimental conditions used, most of the X-pyrene oligonucleotide will remain in single strand form and will not be quenched. In agreement with this hypothesis, the fluorescence was diminished when more of the target strand was added; a large excess produced a reduction in emission intensity of 70 % (Figure S15). To confirm the quadruplex hypothesis, complementary oligonucleotides with deoxyinosine replacing one of the four consecutive deoxyguanosines of the G-rich core were used to prevent G-quadruplex formation (Figure S16). This led to normal duplex formation and strong fluorescence quenching using less than 1.5 equivalents of the complementary strand. Almost 90 % quenching was observed with the deoxyinosine in the middle of the G4 repeat (Figure S16). These fluorescence studies are in agreement with the NMR structural analysis, suggesting that the pyrene moiety, tC^O moiety and the adjacent base pair overlap in the core of the duplex. They also indicate that the fluorescence properties of X-pyrene probes could be utilized effectively in diagnostic and biochemical assays.

Further investigations were carried out comparing X-pyrene to G-clamp (ODN-1/ODN-2 and ODN-4/ODN-5) to determine how fluorescence varies in different duplex contexts (Table 1, Figure 10 A, Figure S17). Fluorescence intensity clearly depends on the sequence; single stranded sequence type-1 (ODN-1, ODN-2) is less fluorescent than sequence type-2 (ODN-4, ODN-5). This could be due to the presence of two cytosines neighboring the fluorescent nucleotide compared to one C and one A (Figure 10 and Figure S17). In fact, similar behavior was observed for tC^O.[47] Nevertheless, the variations of fluorescence intensity in the different duplex contexts do not depend on the sequence. When the fully complementary duplex is formed (ND1 and ND2, Figure 10) the fluorescence of X-pyrene is reduced by up to 60 %. This was expected, as tC^O quantum yield is reduced in dsDNA due to electronic interactions with the other bases.[47] In the case of X-pyrene, the reduction in fluorescence is even more significant, probably because of additional electronic interactions between the tC^O and pyrene moieties. The reduction of fluorescence upon formation of the fully complementary duplex with G-clamp (ND1 and ND2, Figure S17) is not as great (40 %), and is more similar to tC^O.[47]

In almost all X-pyrene bulged duplexes the fluorescence emission was greatly reduced upon duplex formation, presumably because of increased surface contact between the pyrene and tC^O moieties. Interestingly, the emission was reduced less for 5′-intercalation than 3′-intercalation. This tallies with the order of stability of X-pyrene bulged duplexes in the UV-melting studies. 5′-Pyrene intercalation is less efficient, the duplex is less stabilized, the electronic interactions between pyrene and the tC^O moiety are weaker and consequently the fluorescent emission intensity is not reduced as much. The nature of the unpaired base also affects fluorescence; when it is changed from C to T the emission is increased for both 3′- and 5′-intercalation (Figure 10 C), although the effect is very small with 3′-bulged duplexes. In contrast to all previous cases, when the bulge is moved away from X-pyrene by one base pair on the 3′ side (BD6, Figure S18 A), the emission intensity actually becomes higher than that in ssDNA. This is probably because the pyrene is kept away from the tC^O moiety by the intervening base pair and rigid duplex, thereby eliminating mutual quenching interactions. This special case is not possible with ssDNA which is essentially unstructured and will always give rise to some collisional quenching. In support of this interpretation, for the same oligonucleotide sequence containing G-clamp instead of X-pyrene (BD6, Figure S18 B) the fluorescence is not increased but is in fact reduced upon duplex formation in a manner similar to all the other duplexes (Figure S17 C, Figure S18 B).

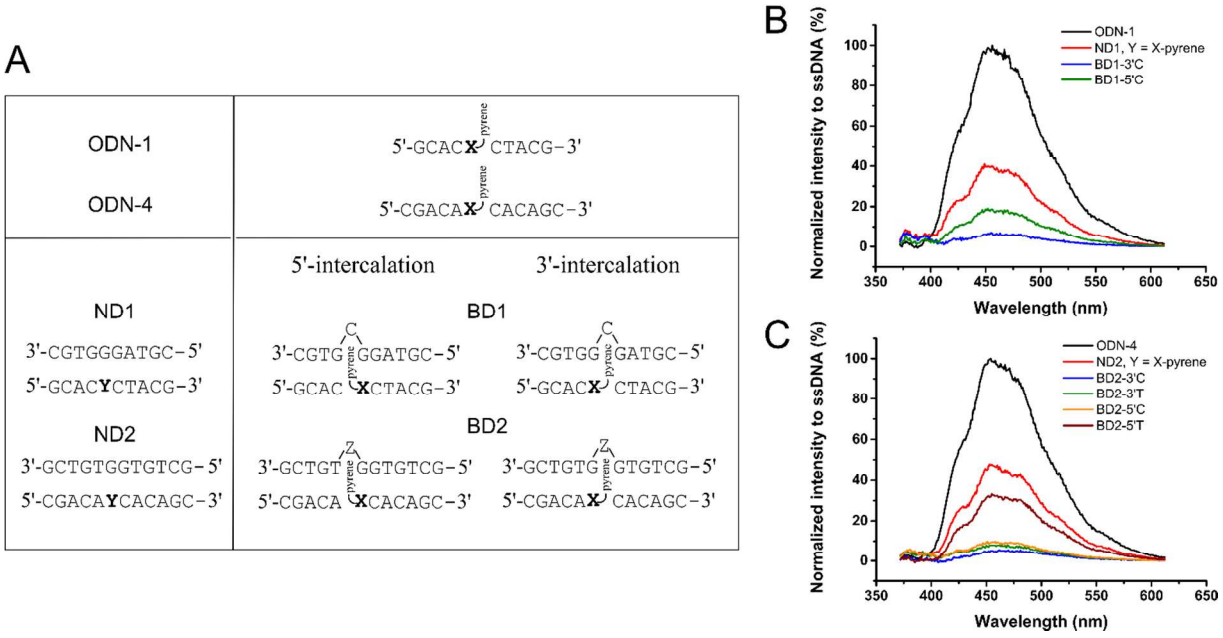


Figure 10. Fluorescence is strongly quenched in X-pyrene bulged duplexes. A) Oligonucleotides and corresponding duplexes involved in the fluorescence study; **B)** Fluorescence emission spectra of a X-pyrene containing oligonucleotide (ODN-1) and its three duplexes (ND1, BD1-3′C, BD1-5′C); **C)** Fluorescence emission spectra of a X-pyrene containing oligonucleotide (ODN-4) and its five duplexes (ND2, BD2-3′C, BD2-3′T, BD2-5′C, BD2-5′T).

Conclusions

X-pyrene combines the three major DNA duplex-stabilizing forces of π-stacking, H-bonding and electrostatic attraction into a single monomer which forms a very stable base pair with guanine. Fully-matched bulged duplexes containing X-pyrene have very high stability in both DNA/DNA and DNA/RNA contexts, whereas the presence of a single immediate or remote mismatch is sufficient to disrupt this stability. Upon duplex formation the pyrene, the tC^{O} moiety and adjacent bases extensively overlap, causing fluorescence emission to decrease dramatically. This is a useful property that could be exploited in sensing applications. The favorable thermodynamic and fluorescence properties of X-pyrene suggest that oligonucleotides containing this new cytosine analogue have great potential in various diagnostic and biomedical applications. The very high observed affinity, sequence selectivity and mismatch discrimination could give rise to improved fluorogenic probes for real-time PCR, targeted DNA or RNA cleavage agents (possibly by enzymatic excision of a widowed DNA nucleobase), microRNA and siRNA silencing, and a new class of antisense oligonucleotides. The principle of combining three cooperative stabilizing interactions in a single monomer is a general concept that could be applied to other intercalators and nucleobase analogues.

Acknowledgements

This work was supported by a research grant to TB from the UK Knowledge Transfer Partnership (KTS). PM is funded by a PhD studentship from ATDBio and the University of Oxford. ATDBio provided assistance with oligonucleotide synthesis and purification.

Notes and references

[a] School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK..
[b] Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK.
[c] Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany and Center for Integrated Protein Science Munich.
[*] To whom correspondence should be sent: Tel: +44(0)1865 275413, E-mail: tom.brown@chem.ox.ac.uk

Electronic Supplementary Information (ESI) available: [Synthesis of X-pyrene monomer, oligonucleotide synthesis purification and analysis (HPLC and CE chromatograms), ultraviolet duplex melting studies, fluorescence studies, circular dichroism measurements, error calculations, NMR spectroscopy, representative UV-melting curves and derivatives, mass spectrometry data for all oligonucleotides]. See DOI: 10.1039/b000000x/

1. J. D. Watson and F. H. Crick, *Nature*, 1953, **171**, 737-738.
2. P. Tiollais, C. Pourcel and A. Dejean, *Nature*, 1985, **317**, 489-495.
3. B. Hartmann and R. Lavery, *Q. Rev. Biophys.*, 1996, **29**, 309-368.
4. A. Sancar, L. A. Lindsey-Boltz, K. Unsal-Kacmaz and S. Linn, *Annu. Rev. Biochem.*, 2004, **73**, 39-85.
5. E. Birney, I. Adzhubei, J. F. Abril, S. F. Aldred, G. R. Abecasis and etc., *Nature*, 2007, **447**, 799-816.
6. A. Okamoto, *Chem. Soc. Rev.*, 2011, **40**, 5815-5828.
7. D. de Semir and J. M. Aran, *Curr. Gene Ther.*, 2006, **6**, 481-504.
8. S. M. Freier and K. H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429-4443.
9. F. Wojciechowski and R. H. E. Hudson, *Curr. Top. Med. Chem.*, 2007, **7**, 667-679.
10. A. Pasternak and J. Wengel, *Org. Biomol. Chem.*, 2011, **9**, 3591-3597.
11. M. A. Campbell and J. Wengel, *Chem. Soc. Rev.*, 2011, **40**, 5680-5689.
12. A. J. Gutierrez, T. J. Terhorst, M. D. Matteucci and B. C. Froehler, *J. Am. Chem. Soc.*, 1994, **116**, 5540-5544.
13. K. Y. Lin, R. J. Jones and M. Matteucci, *J. Am. Chem. Soc.*, 1995, **117**, 3873-3874.
14. K. Y. Lin and M. D. Matteucci, *J. Am. Chem. Soc.*, 1998, **120**, 8531-8532.
15. F. Wojciechowski and R. H. E. Hudson, *J. Am. Chem. Soc.*, 2008, **130**, 12574-12575.
16. T. Nguyen, A. Brewer and E. Stulz, *Angew. Chem. Int. Edit.*, 2009, **48**, 1974-1977.
17. U. B. Christensen and E. B. Pedersen, *Nucleic Acids Res.*, 2002, **30**, 4918-4925.
18. C. D. Kanakis, S. Nafisi, M. Rajabi, A. Shadaloi, P. A. Tarantilis, M. G. Polissiou, J. Bariyanga and H. A. Tajmir-Riahi, *Spectrosc.-Int. J.*, 2009, **23**, 29-43.
19. T. J. Bandy, A. Brewer, J. R. Burns, G. Marth, T. Nguyen and E. Stulz, *Chem. Soc. Rev.*, 2011, **40**, 138-148.
20. C. J. Wilds, M. A. Maier, V. Tereshko, M. Manoharan and M. Egli, *Angew. Chem. Int. Edit.*, 2002, **41**, 115-117.
21. M. A. Maier, J. M. Leeds, G. Balow, R. H. Springer, R. Bharadwaj and M. Manoharan, *Biochemistry*, 2002, **41**, 1323-1327.
22. W. M. Flanagan, R. W. Wagner, D. Grant, K. Y. Lin and M. D. Matteucci, *Nature Biotechol.*, 1999, **17**, 48-52.
23. W. M. Flanagan, J. J. Wolf, P. Olson, D. Grant, K. Y. Lin, R. W. Wagner and M. D. Matteucci, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 3513-3518.
24. J. A. Ortega, J. R. Blas, M. Orozco, A. Grandas, E. Pedroso and J. Robles, *Org. Lett.*, 2007, **9**, 4503-4506.
25. R. X. F. Ren, N. C. Chaudhuri, P. L. Paris, S. Rumney and E. T. Kool, *J. Am. Chem. Soc.*, 1996, **118**, 7671-7678.
26. M. E. Ostergaard and P. J. Hrdlicka, *Chem. Soc. Rev.*, 2011, **40**, 5771-5788.
27. S. P. Sau and P. J. Hrdlicka, *J. Org. Chem.*, 2012, **77**, 5-16.
28. Y. L. Jiang, K. Kwon and J. T. Stivers, *J. Biol. Chem.*, 2001, **276**, 42347-42354.
29. K. Kwon, Y. L. Jiang and J. T. Stivers, *Chem. Biol.*, 2003, **10**, 351-359.
30. C. Beuck, I. Singh, A. Bhattacharya, W. Heckler, V. S. Parmar, O. Seitz and E. Weinhold, *Angew. Chem. Int. Edit.*, 2003, **42**, 3958-3960.

31. Y. L. Jiang, J. T. Stivers and F. H. Song, *Biochemistry*, 2002, **41**, 11248-11254.
32. J. Zhu and R. M. Wartell, *Biochemistry*, 1999, **38**, 15986-15993.
33. J. D. Dunitz, *Chem. Biol.*, 1995, **2**, 709-712.
34. J. Emsley, *Chem. Soc. Rev.*, 1980, **9**, 91-124.
35. T. Nasr, Z. Li, O. Nakagawa, Y. Taniguchi, S. Ono and S. Sasaki, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 727-730.
36. L. De Napoli, A. Messere, D. Montesarchio, G. Piccialli and C. Santacroce, *Nucleos. Nucleot. Nucl.*, 1991, **10**, 1719-1728.
37. L. Denapoli, A. Messere, D. Montesarchio, G. Piccialli and C. Santacroce, *Bioorg. Med. Chem. Lett.*, 1992, **2**, 315-318.
38. S. C. Holmes, A. A. Arzumanov and M. J. Gait, *Nucleic Acids Res.*, 2003, **31**, 2759-2768.
39. B. V. Yang, D. Orourke and J. C. Li, *Synlett.*, 1993, 195-196.
40. C. G. Lou, M. Shelbourne, K. R. Fox and T. Brown, *Chem.-Eur. J.*, 2011, **17**, 14851-14856.
41. C. Lou, Q. Xiao, R. R. Tailor, N. Ben Gaied, N. Gale, M. E. Light, K. R. Fox and T. Brown, *MedChemComm*, 2011, **2**, 550-558.
42. S. H. Chou, D. E. Wemmer, D. R. Hare and B. R. Reid, *Biochemistry*, 1984, **23**, 2257-2262.
43. M. J. Rist and J. P. Marino, *Curr. Org. Chem.*, 2002, **6**, 775-793.
44. P. Sandin, L. M. Wilhelmsson, P. Lincoln, V. E. C. Powers, T. Brown and B. Albinsson, *Nucleic Acids Res.*, 2005, **33**, 5019-5025.
45. J. Telser, K. A. Cruickshank, L. E. Morrison and T. L. Netzel, *J. Am. Chem. Soc.*, 1989, **111**, 6966-6976.
46. J. N. Wilson, Y. J. Cho, S. Tan, A. Cuppoletti and E. T. Kool, *ChemBioChem*, 2008, **9**, 279-285.
47. P. Sandin, K. Borjesson, H. Li, J. Martensson, T. Brown, L. M. Wilhelmsson and B. Albinsson, *Nucleic Acids Res.*, 2008, **36**, 157-167.
48. L. M. Wilhelmsson, *Q. Rev. Biophys.*, 2010, **43**, 159-183.
49. J. L. Huppert, *Febs J.*, 2010, **277**, 3452-3458.
50. M. L. Bochman, K. Paeschke and V. A. Zakian, *Nat. Rev. Genet.*, 2012, **13**, 770-780.