

Recent Advances in Labelling of DNA with Organic Chromophores

Vyacheslav V. Filichev

Institute of Fundamental Sciences - Chemistry, Massey University, Private Bag 11222, Palmerston North
(e-mail: v.filichev@massey.ac.nz)

Introduction

Remarkable changes in the viscosity and sedimentation of complexes of DNA with acridine derivatives suggested to Lerman that intercalation was responsible.¹ Later, intercalative interactions were investigated by using various spectrophotometric methods including X-ray diffraction. Intercalators are typically described as a class of molecules that reversibly bind in the space between two adjacent base pairs of DNA. Traditionally, intercalators are considered as free ligands. If a large aromatic chromophore intercalates into DNA, then unwinding of the DNA double helix occurs, increasing both the flexibility of the DNA and the distance between nucleic acid bases.

Covalent attachment of intercalators to nucleic acids is a useful tool in chemical biology and attachment of fluorophores to either end of a DNA strand is a well-established procedure. Attachment of fluorescent tags via a tether to DNA bases at non-hydrogen bonding sites, e.g. the 5-carbon of C, T or U and the 7-position of 7-deazapurines, allows labelling of many sites in a DNA helix. These techniques found their application in genetic analysis, like DNA sequencing,² real-time PCR (the polymerase chain reaction),³ visualization of chromosomes or DNA/RNA in FISH (fluorescence *in situ* hybridization).⁴ Typical linkers for fluorophore attachment consist of up to a dozen bonds, providing flexibility but also creating an uncertainty in the location and orientation of the chromophore. In contrast, novel methodologies for development of new types of nucleic acids architectures rely on the predictable positioning of labels in the structure.⁵ In that regard, nucleic acids as a helical self-assembled structure provide a clear scaffold, which can be used to develop functional π systems where the correct orientation of each component is vital for the optical properties of the entire system. In the last decade, in particular after sequencing the human genome, the number and variety of intercalating moieties that could be inserted covalently into DNA has increased dramatically, mostly driven by a demand for developing probes that could detect specific nucleic acid sequences. This article presents recent advances in nucleic acid architectures using an example of a logical development of pyrene insertions. Pyrene is one of the most extensively studied chromophores in DNA/RNA structures. Initially, pyrene and other organic chromophores were largely used as *molecular caps*. Later they found use opposite abasic sites, as a counter base, as artificial hydrophobic pairs and, very recently, as helically arranged π - π stacking arrays. Organic chromophores can be introduced into nucleic acid structures via modification of nucleotides, *i.e.* attachments to the base, sugar or phosphate, or via entirely non-nucleosidic linkers. By such means a chromophore can be placed in the interior of nucleic bases or in the grooves of the nucleic acid complexes.

DNA Structures: Duplex, Triplex and G-Quadruplex

Hydrogen bonding and π - π stacking between nucleic bases, electrostatic repulsion between negatively charged phosphates and hydration are responsible for the stability of nucleic acids. Depending on its sequence and media, a nucleic acid can accommodate numerous secondary structures, such as the classical B-type helix, the A-form or left-handed Z (zigzag) form. Alternative (non-Watson-Crick) Hoogsteen and reverse Hoogsteen base-pairing can give rise to high-order DNA and RNA structures that include triplexes and quadruplexes (Fig. 1).

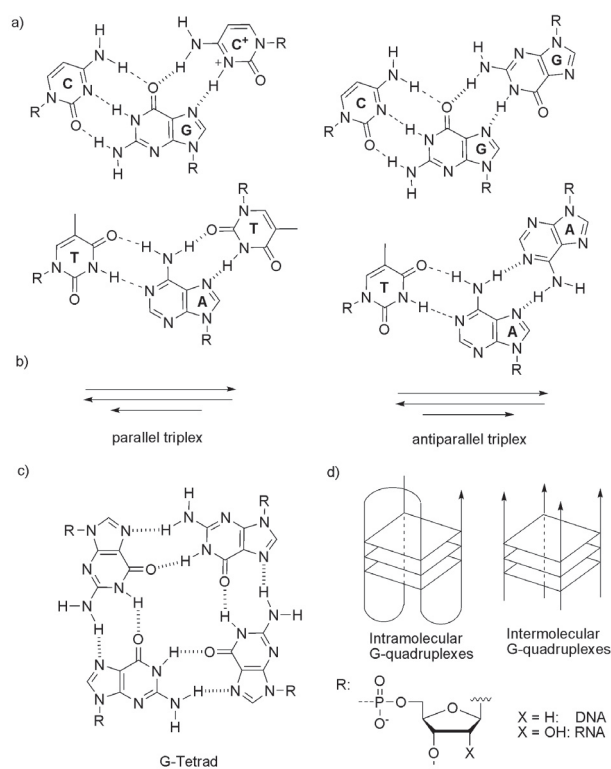


Fig. 1. Structures and secondary structures of DNA/RNA. Parallel and antiparallel triple helices, a): H-bonding formation, Watson-Crick base-pairing are represented by dashed bonds and Hoogsteen or reversed-Hoogsteen are represented by hashed bonds; b): representation of the strand orientation in triplex. G-quadruplexes; c): H-bond formation in G-tetrad; d): intra- and intermolecular G-quadruplexes.

Triple helices are formed when a single-stranded triplex-forming oligonucleotide (TFO) binds in the major groove of double-stranded DNA (dsDNA).⁶ A homopyrimidine TFO can bind in a parallel fashion to the homopurine strand of dsDNA (Fig. 1). A nucleobase T binds to a nucleobase A of the duplex, but cytosine in the TFO must be protonated at the N-3 atom to form Hoogsteen base-pairing with dsDNA. This makes formation of parallel triplexes at neutral pH problematic. A homopurine third

strand can also bind to a homopurine-homopyrimidine duplex using reversed Hoogsteen patterns. In this triplex, a nucleobase A binds to a T-A base pair and a G to a C-G pair. Here, the homopurine TFO is antiparallel to the homopurine strand of the original dsDNA. GT-rich TFOs can also form parallel and antiparallel triplexes depending on the sequence.⁷ G-Rich sequences can fold into stable G-quadruplexes, especially in the presence of K^+ (Fig. 1c, 1d).⁸ The topology and molecularity of G-quadruplexes are determined by several factors including variations in strand polarity, loop geometry, the presence of metal ions, *etc.*⁹ The extraordinary stability of G-quadruplexes has some unfortunate outcomes for G-rich sequences: TFO probes designed to form triplexes but possessing stretches of guanines do not bind to duplexes at all because of the probes' self-aggregation.¹⁰

Introduction of additional moieties to DNA has an impact on the thermodynamic stability of DNA complexes. When UV/VIS, fluorescence or CD spectrometers are fitted with a Peltier temperature programmer, melting of DNA complexes or thermal stability studies can be undertaken. The melting temperature (T_m , °C) is usually defined as the maximum of the first derivative of the curves obtained by measuring absorbance at 260 nm or CD against increasing temperature. Control experiments should be performed at other wavelengths also, depending on the DNA structure and the absorbance/fluorescence region of the intercalator. Luminescence is a widely explored feature of intercalators. The fluorescent sensitivity of some fluorophores depends highly on changes in the microenvironment. For example, the fluorescence of pyrene is quenched in low-polarity environments and increased in high-polarity environments. By inserting such intercalating units as reporter groups, one can gain a deeper understanding of fundamental features of nucleic acids, for example, folding and unfolding mechanisms, local dynamics within DNA helices, polymorphism in DNA, the type of forces involved in formation of different secondary structures, and so forth.

Excitation (λ_{ex}) and emission wavelengths (λ_{em}) are important descriptive values for each chromophore. Usually, organic chromophores exist in an excited state for a finite time (typically 1–10 ns). The difference in energy or wavelength, represented by $(h\nu_{ex} - h\nu_{em})$, is called the Stokes shift and is fundamental to the sensitivity of fluorescence techniques because it allows emission photons to be detected against an unpolluted background, isolated from excitation photons. The fluorescence quantum yield (Φ_f) is defined as the ratio of the number of photons emitted by fluorescence at a certain wavelength to the number of photons absorbed at this wavelength. In cases where there are multiple insertions of chromophores, particularly if two fluorescent molecules are positioned and overlapped in close proximity to each other (*ca.* 3.4 Å), formation of excimers or exciplexes can be observed. During *excited dimer* (excimer) formation, which can be obtained in saturated solutions of a free ligand, one molecule is in an electronically excited state and the other is in an electronic ground state. In contrast, an *excited duplex* (exciplex) is a bimolecular complex of two different chromophores.

Molecular Caps and DNA Interstrand Stacking Interactions

Pyrene's size (220 Å²) allows it to occupy the area usually covered by two natural purine:pyrimidine base pairs (269 Å²).¹¹ It is an ideal lid for canonical base pairs¹² because its size cannot fully cover purine:purine base pairs, whereas pyrimidine:pyrimidine pairs are too small to accommodate it. However, the choice of linker is crucial. This can be demonstrated by insertion of pyrene at the 5'- or 3'-terminus of oligonucleotides (Fig. 2). Introduction of such molecular caps can lead to an increased thermal stability of the resulting complexes with ssDNA/RNA that can exceed 10 °C per modification.¹³ The thermodynamic stabilization gained is a result of the stacking pyrene with the nearest nucleic bases.¹⁴ Additionally, the dangling-end residue acts as a hydrophobic cap and restricts bulk water access to the terminal base pair. This makes these terminal base pairs energetically comparable to the corresponding internal base pairs.¹⁵ In the case of flexible linkers, such as butyric acid, pyrene can adopt a number of different conformations, whereas in structures such as 2-deoxyribofuranose (**P**) or pyrrolidine (**azaP**) the pyrene moiety remains in a more rigidly defined position on the top or the bottom of the helix. The defined position leads to slightly higher thermal stability and hyperchromicity, as it has been shown for **azaP** in comparison with 1-pyrenylbutyric acid.¹² These properties have already been used in the design of probes that combine enhanced affinity and base pair fidelity. An unmodified DNA probe can barely discriminate the perfectly matched and terminal mismatched DNAs ($\Delta T_m = -0.3 - +1.7$ °C). However, when pyrene pyrrolidine **azaP** ($n=0$) as a 5'-end cap was combined with anthraquinone derivative (**AQ**) as a 3'-end cap,¹⁶ the prepared dodecamer DNA gave a melting point decrease for a double-terminal mismatch of $\Delta T_m = +6.2 - 7.4$ °C. Similarly, a doubly capped matched DNA duplex showed increased thermal stability in comparison with an unmodified one ($\Delta T_m = +11.1$ °C).¹²

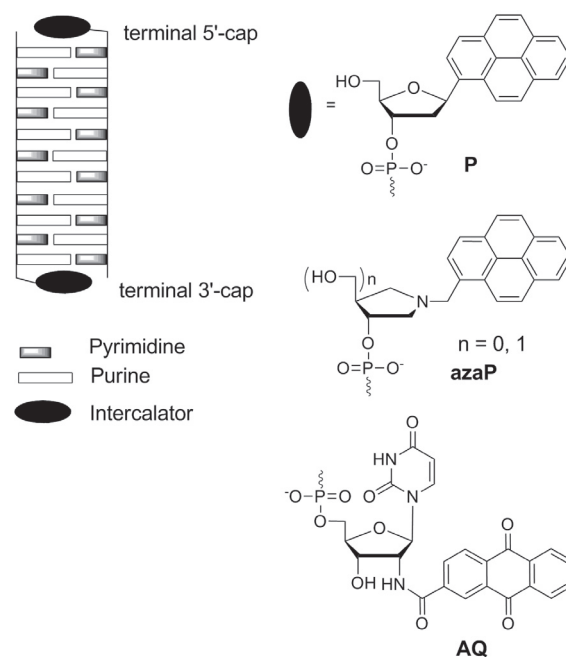


Fig. 2. Molecular caps - chromophores covalently attached to the 3'- or 5'-end of the duplex.

Insertion of an intercalating moiety in the core of the DNA duplex can be performed in three ways: opposite the natural base (Fig. 3), opposite an abasic site (Fig. 4), and as a bulge (Fig. 5). In contrast to free intercalating ligands, the covalently attached chromophores as base substituents do not have to unwind the DNA for intercalation. Hydrophobic base analogues do not form pairs with the natural bases and these combinations are usually strongly destabilized. This is an outcome of the absence of hydrogen bonding formation and imperfect steric fit within the DNA duplex.¹⁷ Also, if the size of intercalator is too large the counter nucleobase may be flipped out of the helix.

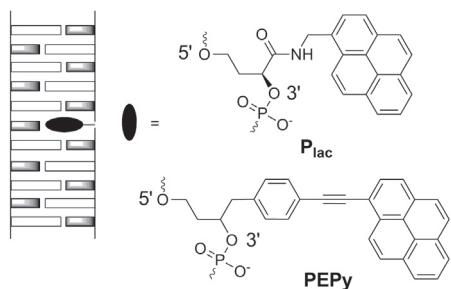


Fig. 3. Chromophores in the middle of a DNA duplex opposite the natural base.

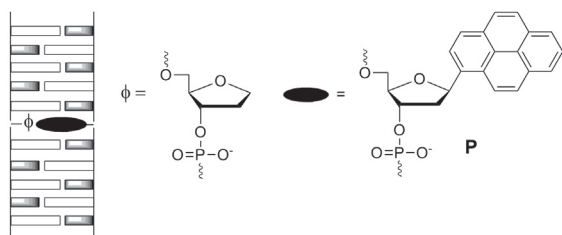


Fig. 4. Chromophores in the middle of a DNA duplex opposite an abasic site.

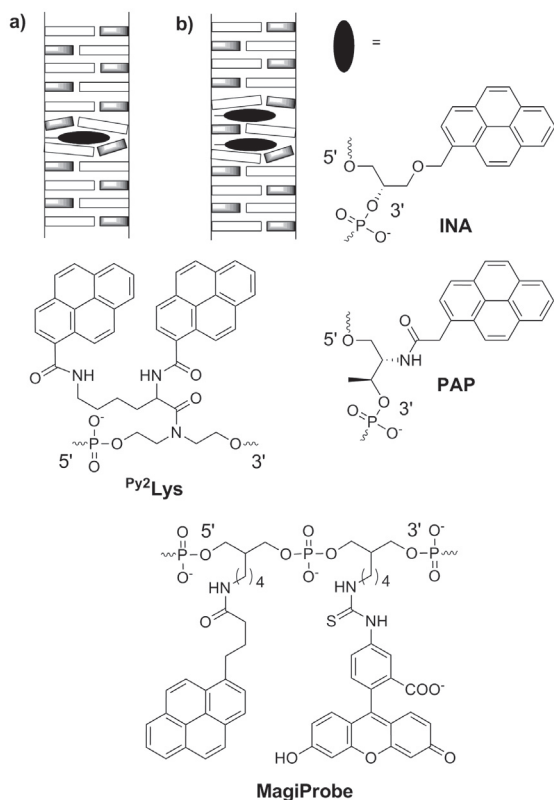


Fig. 5. Fluorescent probes designed using the principle of bulge insertions; a): a bulge insertion; b): chromophore insertions as next-nearest neighbours.

Generally, DNA duplexes can more easily accommodate non-native structures incorporated close to the ends than in the middle of the helix. Thus, incorporation of P_{lac} in the midst of the sequence led to much more pronounced destabilization ($\Delta T_m = -6.3$ °C) than for insertions close to the 3'- or 5'-ends ($\Delta T_m = -2.7$ °C).¹⁸ 1-(Phenylethynyl)pyrene (**PEPy**) has been placed in a DNA duplex using a 1,3-butanediol linker instead of the nucleobase and a slight destabilization of the duplex was observed ($\Delta T_m = -1.4$ °C).¹⁹ Spectroscopic properties of **PEPy** analogues are beneficial compared with those of the initial pyrene. Nowadays, standard genetic analysis platforms are primarily designed to detect fluorescein at 520 nm,³ which unfortunately eliminates many potentially useful labels based on the pyrene moiety ($\lambda_{ex} \sim 330\text{--}340$ nm; $\lambda_{em} \sim 400\text{--}410$ nm). Moreover, biomolecules are also excited upon irradiation at the same wavelength as pyrene.²⁰ For **PEPy** there is a bathochromic shift in the absorption spectra (373 nm vs 343 nm for pyrene) and fluorescence maxima are shifted to longer wavelength (400–410 nm vs 380 nm for the first peak in a monomer fluorescence and 500–510 nm, compared with 480 nm for excimer fluorescence).^{19–21} This is also accompanied by high fluorescence quantum yields in buffer solutions even in the presence of oxygen. A further shift in fluorescence maxima to longer wavelengths was observed upon introduction of the second and the third 1-phenylethynyl residues in the structure of pyrene (positions 1, 6 and 8).^{20,22} Insertion of two **PEPy** units opposite the variable nucleotide sites has been successfully used in the design of fluorescence probes for detection of single polymorphisms in the gene fragment of 23S rRNA *Helicobacter pylori* showing remarkable sensitivity for the presence of a T or of a non-T base opposite to the **PEPy** pair.²³

If the DNA base opposite is missing a so-called *abasic site* (Fig. 4), some aromatic molecules can actually fit within the DNA duplex and restore the π - π DNA duplex base stack. A good example of this is pyrene C-nucleoside **P** (Fig. 2).¹¹ The 5'-triphosphate of **P** was incorporated into DNA using a Klenow fragment opposite an abasic site with a selectivity and efficiency which were greater than those for the natural DNA triphosphates. The termination of the DNA biosynthesis was detected after **P** incorporation,²⁴ which makes this fluorescent pseudo-nucleoside useful for distinguishing abasic mutations. Insertion of **P** opposite an abasic site was more energetically favourable than the use of natural bases. The stabilization of the duplex was detected in a ΔT_m range of 18–23 °C in comparison with native duplexes having an internal abasic site.¹¹

Incorporation of organic chromophores as a bulge (Fig. 5) was used in the development of DNA probes for various applications. Chromophores actually intercalate between two adjacent base pairs in a way similar to free ligands. In general, acyclic linkers that can be more easily accommodated in the helix are used for bulged constructions.

Discriminative binding to ssDNA over ssRNA having identical sequences was observed for intercalating nucleic acids (**INA**) having a bulged pyrene moiety connected via a glycerol linkage in the middle of the DNA sequence ($\Delta T_{m\text{ INA/DNA}} = +3.0$ °C, $\Delta T_{m\text{ INA/RNA}} = -4.4$ °C per modifica-

tion).²⁵ Two factors contribute to the discrimination upon binding to DNA and RNA: distortion of the nucleic acid backbone by short ethyleneglycol linker and stabilization of the **INA**/DNA structure by intercalation. Interestingly, using pyrene connected via a flexible linker to form a bulge was energetically more favorable than the use of a nucleobase ($\Delta T_m = -12.8$ °C) and the ethylene glycol linker on its own ($\Delta T_m = -8.2$ °C),²⁵ the latter being a clear indication of intercalation. When two **INA** monomers were placed as next-nearest neighbours (Fig. 5b) a strong excimer fluorescence at 480 nm ($\lambda_{ex} = 340$ nm) was observed for the probe alone.²⁶ The excimer band was completely quenched upon formation of a fully complementary **INA**/DNA duplex. If there was a mismatch between **INA** residues, there was no or little quenching of the excimer band. A deletion polymorphism was effectively detected using two **PAP** monomers as next-nearest neighbours. No excimer fluorescence was detected for the perfectly matched duplex, whereas a deletion of a single-base opposite **PAP** monomers resulted in the formation of the pyrene excimer band.²⁷ A ^{Py}**Lys** backbone with two 1-pyrenyl residues was designed to report insertion mutations.²⁸ The fluorescence intensities for the probe alone and for the perfectly matched duplexes with ^{Py}**Lys** as a bulge were very weak ($\Phi_F = 0.001$ - 0.007). If an extra base was presented in the target opposite the ^{Py}**Lys** insertion, a strong fluorescence peak at 495 nm appeared ($\Phi_F = 0.088$).

Pyrene was used as a quencher of fluorescein in the MagiProbe design.²⁹ When hybridized to the perfectly matched sequence pyrene intercalates into the duplex, this results in the emission of fluorescence from fluorescein. No fluorescence signal was detected for the probe on its own or upon binding to the mismatched sequence close to the pyrene residue.

Bulged intercalators have also been introduced in DNA triplex technology (Fig. 6). We have recently developed an example of a triplex selective intercalator.²¹ Bulge insertions of (*R*)-1-*O*-[4-(1-pyrenylethynyl)phenylmethyl]glycerol (*twisted intercalating nucleic acids*, **TINA**, Fig. 6) into the middle of a homopyrimidine strand led to high triplex thermal stability ($\Delta T_m = +19.0$ °C per modification) and discrimination of Hoogsteen-type triplexes over Watson-Crick type duplexes. Under similar conditions the native TFO was unable to bind to the duplex. The 1-pyrenyl derivative of **TINA** was found to be the most effective among the tested modifications (acridine, naphthalene, *m*-phenylethynyl, 4-biphenyl, 2- or 4-pyrene) for binding to the Hoogsteen-type triplexes.^{21,30} It is interesting to compare the ability of the rather similar structures **INA** (Fig. 5) and **TINA** (Fig. 6) to stabilize parallel triplexes. In contrast to **TINA**, pyrenemethylglycerol (**INA**) destabilized triplexes upon bulged insertions into identical sequences ($\Delta T_m = -5.0$ °C).³¹ These data highlight the importance of the proper intercalator, *e.g.* pyrene positioning in the interior of the dsDNA region of the triple helix. It is important to mention planarity of the intercalating unit is not an absolute requirement because bases are propeller-twisted to a varying degree within the triplets. Rigid linkers, such as single, double or triple C-C bonds, can connect two or three chromophores thus helping the entire intercalator to fit correctly within the base triplets.

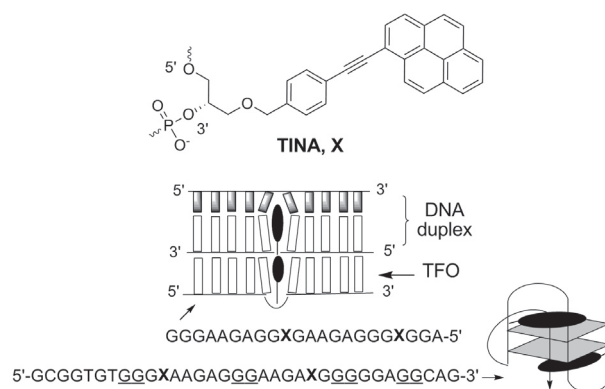


Fig. 6. TINA molecule and its use in the design of triplex-forming oligonucleotides and G-quadruplexes.

We recently investigated the ability of **TINA** to stabilize and destabilize G-quadruplexes. It is known that quadruplex DNA can be stabilized by free ligands such as cationic porphyrins,³² trisubstituted acridines,³³ and others.³⁴ We assumed that the placement of pyrene units on the top or the bottom of the G-quadruplex structures would stabilize them via stacking interactions. On the other hand, insertions *between* adjacent guanines should, in principle, destabilize self-association of the G-rich sequences (Fig. 6). The latter can be useful for antiparallel TFOs. We covalently attached one or two **TINAs** in positions adjacent to the stretches of guanines in the quadruplex motif, mimicking the structure which is located in the promoter region of the human *KRAS* gene. A dramatic increase in the T_m ($\Delta T_m = +22 - 32$ °C) and a strong antiproliferative effect in Panc-1 cells were observed.³⁵ In contrast, when **TINA** monomers were inserted in the middle of G-runs significant destabilization of G-quadruplex-like structures was observed. The final **TINA**-TFOs were able to bind to the DNA duplex in an antiparallel fashion even in the presence of potassium ions.³⁶

Summarizing the effect of using aromatic substituents in the middle and at either end of the DNA sequence, we can conclude that organic chromophores can restore or enhance DNA stacking interactions by virtue of intercalation. In that regard not only the size of the hydrophobic molecule but also the proper coverage of the nucleic base surface and a linker are crucial in the molecules' design. It can be anticipated that future developments will be more focused on non-canonical base pairs and different DNA/RNA secondary structures such as cruciforms, hairpins and three- or four-way junctions.

From Single Chromophore Incorporation to the Extended Stacking Interactions within DNA Duplexes

Stacking interactions of organic chromophores in the middle of nucleic acid helices represent a promising approach for the design of DNA-based tools for nanobiotechnology. As could be expected, interstrand stacking interactions give rise to interesting and useful UV and fluorescence properties. Impressive changes in optical properties of organic chromophores were observed during formation of oligomers encompassing fluorescent *C*-nucleosides, the so-called fluorosides (Fig. 7).³⁷ Enhanced molar extinc-

tion coefficients and enhanced Stokes shifts up to 145 nm were observed for the oligomer with three pyrenyl derivatives (**3P**, Fig. 7). Using a set of four fluorosides, a combinatorial library was created with a broad array of emission colours from violet to orange. This diversity was achieved with a narrow range of UV excitation (340-380 nm). In a larger combinatorial library composed of eleven fluorosides in oligomers four units long (**4Ch**, Fig. 7), certain constructs exhibited large hypsochromic shifts in their emission spectra after exposure to UV light.³⁸

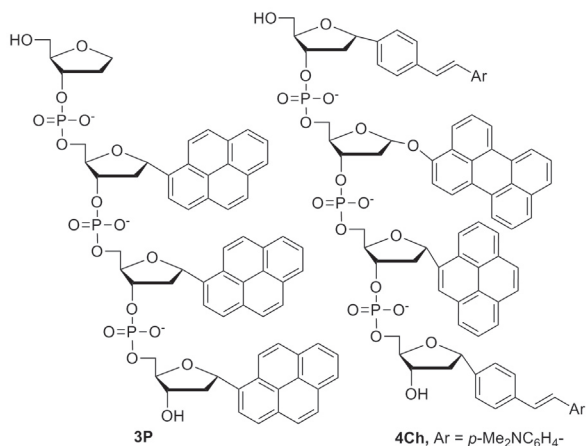


Fig. 7. Examples of polyfluorophores.

Insertion of pyrenyl fragments [INA (Fig. 5), **P**_{1,8}, **P**triazole_{1,8} (Fig. 8)] opposite each other led to thermal destabilization in the range of ΔT_m from -0.2 to -3.0 °C per pair of intercalators when compared to wild-type DNA.^{39,40} Overlapping of pyrene residues in the interior of a DNA duplex, which has been shown by NMR for **INA**,³⁹ led to formation of a strong excimer band. Depending on the structure the λ_{em} values for excimer varied from 480 nm for **INA**⁴¹ to 493 nm for **P**_{1,8},⁴⁰ and to 520 nm for **P**triazole_{1,8}⁴² and **P**triple_{1,8}.⁴³ [λ_{ex} = 340-350 nm (**INA**, **P**_{1,8}, **P**triazole_{1,8}) and 385 nm for **P**triple_{1,8}]. In recent articles homopairs of 1,8-pyrene (**P**_{1,8}) were placed one, two, three and even seven times in the DNA helix.^{44,45} A blue shift in the excimer band from 515 nm to 504 nm was observed with increasing numbers of pyrene pairs, which was also accompanied by a decrease in thermal stability from 70.5 °C for wild-type duplex to 56.5 °C for seven pairs of **P**_{1,8}.⁴⁵

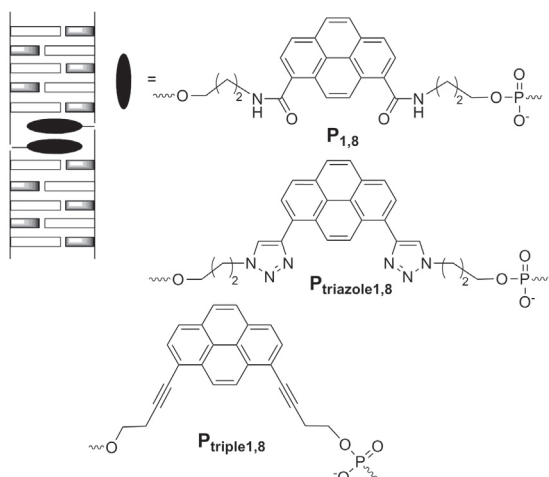


Fig. 8. Representation of a DNA duplex with hydrophobic aromatic moieties positioned opposite each other to form a pair.

Extensive research has been performed on the derivatization of nucleosides at the O2'-position. Depending on the linker length and its flexibility, it may be anticipated that a large chromophore spends a fraction of its time in the major/minor grooves and also partially moves between nucleobases (Fig. 9).

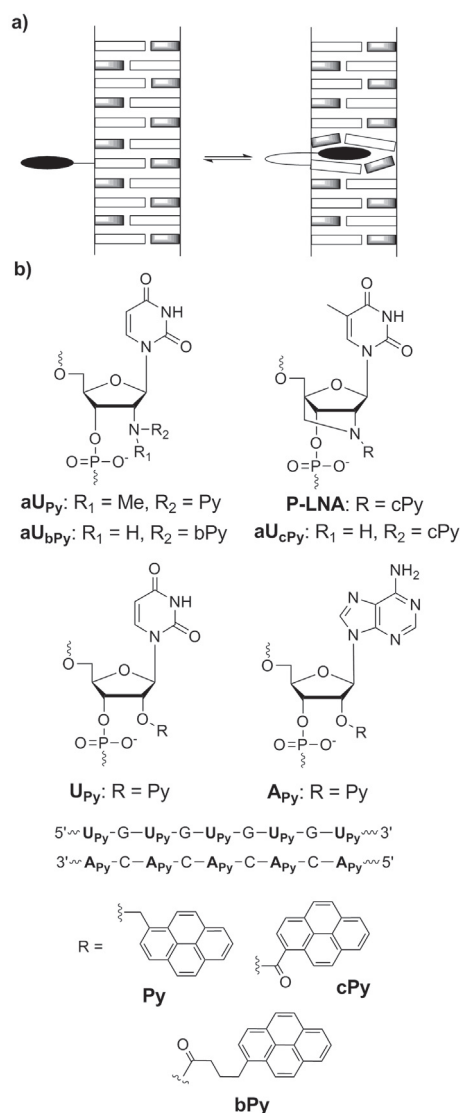


Fig. 9. a): Representation of equilibrium between the states when aromatic moiety positioned in DNA grooves (left) and when it intercalates (right); b): chromophores linked to the 2'-position of nucleotides.

It is believed that intercalators have a preference in binding to DNA rather than to RNA.⁴⁶ A substantial increase in the fluorescence signal was observed when probes possessing pyren-1-ylmethyl (**U**_{Py}, Fig. 9) were bound to the complementary ssRNA, while marginal changes were observed for the complementary ssDNA.^{47,48} Duplexes with complementary RNA were strongly destabilized relative to unmodified DNA/RNA duplexes. Such discrimination in fluorescence and thermal stability upon binding to complementary ssDNA and ssRNA is a result of different positioning of the chromophore in the interior of ssONs and duplexes. Pyrene fluorescence is quenched in ssONs and dsDNA because of interactions with neighbouring nucleobases, including intercalation.⁴⁹ However, pyrene attached to a duplex with RNAs is located outside of the duplex primarily in the minor groove, resulting in en-

hanced fluorescence. Sequential incorporation of 2-4 U_{Py} groups in the middle of RNA strands resulted in formation of an excimer band at 480 nm for the unhybridized probe ($\lambda_{ex} = 350$ nm).⁵⁰ A two- to ten-fold increase in the excimer fluorescence was observed upon formation of the RNA duplexes. The pyrene absorption band in the double-stranded form appeared at shorter wavelengths than that of the corresponding mono-pyrene labelled RNA duplex. A similar behaviour in fluorescence was observed for a pyrene zipper array along the RNA duplex by hybridization of two strands in which five U_{Py} groups formed a pair with five A_{Py} groups (Fig. 9).⁵¹ CD spectra showed a considerable difference (a positive Cotton effect) for zipper arrays in the region of the pyrene absorption in comparison with the interrupted pyrene stack. These spectroscopic behaviours suggest that the pyrenes in the duplex form are located in the minor groove and are associated with each other as an H-aggregate that induces a hypsochromic shift of the absorption band. Disruption of the pyrene stack in the minor groove resulted in disappearance of the excimer band for duplexes.⁵⁰

A different situation was observed for pyrene-modified N2'-uridines (aU_{Py} and aU_{bPy}). Upon insertion of these analogues into DNA sequences a significant increase in T_m for the corresponding DNA duplexes, +15 and +3 °C, respectively, was observed, whereas duplexes with RNA were slightly destabilized.^{52,53} Such a strong increase in T_m for pyrene analogues toward DNA is no doubt an intercalating effect. However, depending on the site of aU_{Py} insertion, probes showed different fluorescence properties. In some cases the monomer fluorescence of pyrene was quenched, in others only marginal changes were observed upon formation of complementary sequences. One possible explanation for these striking differences is that the pyrene moiety is only partially intercalating and is partially located in the minor groove. This situation was changed when a more rigid amide linkage was used. Thus, an increase in fluorescence was observed for DNAs having a single insertion of aU_{cPy} upon binding to complementary DNA.⁵² Another valuable example in this series is 2'-amino-LNA molecules (**P-LNA**).⁴⁸ LNA or locked nucleic acids are known to induce large increases in thermal stability upon binding to RNA and DNA sequences, which is a result of the constrained 3'-endo configuration of the furanose ring.⁵⁴ During hybridization of DNA sequences comprising from one to three **P-LNA** monomers to either DNA or RNA complements, a large increase in thermal stability and fluorescence (up to 69-fold depending on the probe design) was detected. Molecular modeling of the resulting duplexes with **P-LNA** and aU_{cPy} suggested that the position of the pyrene is fixed in the minor groove.⁴⁸ Significant quenching of pyrene fluorescence in these single-stranded probes can be assigned to the interactions with nucleic bases.

Attachment of chromophores at non-hydrogen bonding positions of nucleic bases is a popular approach in the design of nucleic acids for analysis of singly matched/mismatched base pairs. Linker length and rigidity is critical in the design of these probes.

The fluorescence spectra of single stranded probe and mismatched duplexes possessing pyren-1-yl-modified 7-deazaadenine ($a^{Py}A$, Fig. 10) showed strong fluorescence emission at 390 nm. The flexible propyl linker led to the pyrene intercalation into the duplex that resulted in the decreased fluorescence emission for the fully matched T: $a^{Py}A$ duplex.⁵⁵ However, quenching of the fluorescence signal is not a desired property in the probe design, because a positive rather than a negative signal is required in major assays. In contrast to the ^{Py}A molecule, use of the stiffer propargyl linker in the structures ^{Py}U , ^{Py}C and C-C single and triple bonds (^{P}U , ^{1EP}U , Fig. 10) led to strong fluorescence emission ($\lambda_{em} = 397$ nm for ^{Py}U and ^{Py}C ; $\lambda_{em} = 455$ nm for ^{1EP}U) with Φ_F values of 0.100-0.151 upon hybridization to the *matched* sequences, while a weak fluorescence was detected for the mismatched base pairs.⁵⁶ This strongly suggests that the position of the dye is fixed in the solvated major groove. Continuous stacking in the major groove of at least three pyrenes attached to uridine *via* single or triple C-C bonds (Fig. 10) led to drastic changes in UV/VIS and fluorescence spectra of fully matched duplexes relative to probes alone.^{57,58} Fluorescent enhancement (*ca.* 22-fold) and a blue-shift in a maximum emission from 475 nm for ssDNA to 445 nm for the duplex were observed for five non-interrupted insertions of ^{P}Us (Fig. 10) in the DNA strand. In contrast to that, a red-shift in emission spectrum was detected for five insertions of 1-ethynylpyrene uridine (^{1EP}U , $\lambda_{em} = 485$ -494 nm) in the duplex relative to the single insertion ($\lambda_{em} = 445$ nm).⁵⁷ For duplexes comprised of more than three ^{P}U or ^{1EP}U chromophores, distinct signals related to pyrene were observed in CD spectra that confirm a helical arrangement of pyrene in a major groove of DNA duplexes. This means that at least three chromophores in the intact DNA helix environment are required to obtain an ordered π -stacked array. Right-handed helical arrangement of ^{P}U and ^{1EP}U chromophores could be destroyed by either thermal denaturation or by insertion of mismatches opposite to the pyrenes.

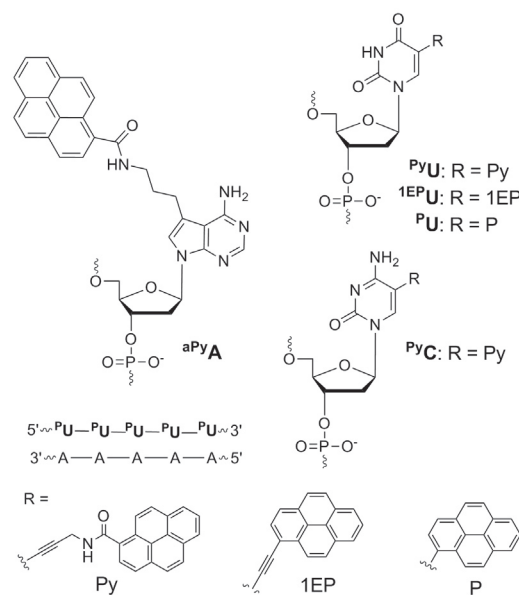


Fig. 10. Chromophores attached to nucleic bases at non-hydrogen bond forming sites.

The Future

Advances in the automated synthesis of DNA and post-synthetic oligonucleotide modifications led to the vigorous development of novel types of probes possessing organic chromophores. In the last decade there has been a considerable change in the design of probes. The predictable positioning of labels in the nucleic acid structure is vital to obtain the desired optical and binding properties. This goal is achieved by using short and rigid linkers that bring certainty in the chromophores' arrangements within the DNA/RNA helix. Those efforts described in this article lay a foundation for the generation of novel nucleic acid architectures which will have an enormous impact on the design of probes in molecular biology, biotechnology, and nanotechnology. Despite such an impressive development, a disadvantage is the use of aromatic molecules that are available on the market or easily modified synthetically. Their large size, photobleaching, quenching of fluorescence by surrounding nucleobases, formation of aggregates upon multiple insertions into DNA, overlapping of absorption or emission wavelengths with biomolecules are the main drawbacks of many organic chromophores. From that point of view, pyrene derivatives can be considered as extremely ineffective molecules. However, the efforts presented have already expanded our knowledge on stacking interactions within DNA and we have learned much about the properties of organic chromophores and their interactions with nucleic acid bases and themselves. It can be anticipated that this knowledge will be used in the construction of novel DNA architectures that will bring probe design to a conceptually new level. There is a considerable demand for novel chromophores with tunable wavelengths, high emission intensities and brightness, sensitivity to microenvironmental changes and insensitivity to surrounding nucleobases. For nanometre-sized materials availability of additional functional groups on organic chromophores would be of a high value for vigorous development of DNA entities as catalysts, nanomachines, nanorobotics, nanoelectronics, etc. In that regard, chemically modified pyrene, perylene, acridine and coumarine entities that have improved optical properties would be extremely valuable. Dramatic development of DNA-based approaches involving the use of nanoparticles and quantum dots in combination with fluorescent molecules can be expected.⁵⁹ This will expand areas of application and advance labelling techniques by taking the advantages of different classes of compounds. Post-synthetic labelling procedures, some of them recently highlighted,⁶⁰ will play a significant role in the synthesis, and more importantly, in the screening of DNA-conjugated organic chromophores. It will allow a rapid discovery of probes showing discriminative recognition of certain nucleic acid secondary structures with tunable luminescence properties. Moreover, it seems to pave the way for novel multichromophore architectures that will be of use in the development of photoactive materials and photonic nanodevices.

Acknowledgments

The work is funded by the Marsden Fund Council from Government funding, administrated by the Royal Society of New Zealand.

References

- Lerman, L. S. *J. Mol. Biol.* **1961**, *3*, 18-30.
- Glazer, A. N.; Mathies, R. A. *Curr. Opin. Biotechnol.* **1997**, *8*, 94-102; Smith, L. M.; Sanders, J. Z.; Kaiser, R. J.; Hughes, P., *et al. Nature* **1986**, *321*, 674-679.
- Ranasinghe, R. T.; Brown, T. *Chem. Comm.* **2005**, 5487-5502.
- Pinkel, D.; Gray, J. W.; Trask, B.; Vandenengh, G., *et al. Cold Spring Harb. Symp. Quant. Biol.* **1986**, *51*, 151-157; Luo, D.; Saltzman, W. M. *Nat. Biotechnol.* **2000**, *18*, 33-37.
- Wengel, J. *Org. Biomol. Chem.* **2004**, *2*, 277-280; Wagenknecht, H. A. *Angew. Chem. Int. Ed.* **2009**, *48*, 2838-2841.
- Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024; Frankkamenetskii, M. D.; Mirkin, S. M. *Ann. Rev. Biochem.* **1995**, *64*, 65-95; Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.
- Thuong, N. T.; Helene, C. *Angew. Chem. Int. Ed.* **1993**, *32*, 666-690.
- Davis, J. T. *Angew. Chem. Int. Ed.* **2004**, *43*, 668-698.
- Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K., *et al. Nucleic Acids Res.* **2006**, *34*, 5402-5415.
- Noonberg, S. B.; Francois, J. C.; Garestier, T.; Helene, C. *Nucleic Acids Res.* **1995**, *23*, 1956-1963; Cheng, A. J.; Wang, J. C.; Van Dyke, M. W. *Antisense Nucleic Acid Drug Dev.* **1998**, *8*, 215-225.
- Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191-6192.
- Narayanan, S.; Gall, J.; Richert, C. *Nucleic Acids Res.* **2004**, *32*, 2901-2911.
- Asseline, U.; Delarue, M.; Lancelot, G.; Toulme, F., *et al. Proc. Natl. Acad. Sci. USA. Biol. Sciences* **1984**, *81*, 3297-3301; Asseline, U.; Toulme, F.; Thuong, N. T.; Delarue, M., *et al. EMBO J.* **1984**, *3*, 795-800; Filichev, V. V.; Pedersen, E. B. *Org. Biomol. Chem.* **2003**, *1*, 100-103.
- Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J., *et al. J. Am. Chem. Soc.* **2000**, *122*, 2213-2222; Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J., *et al. J. Am. Chem. Soc.* **1996**, *118*, 8182-8183.
- Isaksson, J.; Chattopadhyaya, J. *Biochemistry* **2005**, *44*, 5390-5401.
- Connors, W. H.; Narayanan, S.; Kryatova, O. P.; Richert, C. *Org. Lett.* **2003**, *5*, 247-250.
- Schweitzer, B. A.; Kool, E. T. *J. Am. Chem. Soc.* **1995**, *117*, 1863-1872; Loakes, D.; Brown, D. M. *Nucleic Acids Res.* **1994**, *22*, 4039-4043.
- Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Korshun, V. A., *et al. Org. Lett.* **2002**, *4*, 4607-4610.
- Malakhov, A. D.; Skorobogatyi, M. V.; Prokhorenko, I. A.; Gontarev, S. V., *et al. Eur. J. Org. Chem.* **2004**, 1298-1307.
- Maeda, H.; Maeda, T.; Mizuno, K.; Fujimoto, K., *et al. Chem. Eur. J.* **2006**, *12*, 824-831.
- Filichev, V. V.; Pedersen, E. B. *J. Am. Chem. Soc.* **2005**, *127*, 14849-14858.
- Astakhova, I. V.; Korshun, V. A.; Wengel, J. *Chem. Eur. J.* **2008**, *14*, 11010-11026.
- Prokhorenko, I. A.; Malakhov, A. D.; Kozlova, A. A.; Momynaliev, K., *et al. Mutat. Res.* **2006**, *599*, 144-151.
- Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704-708.
- Christensen, U. B.; Pedersen, E. B. *Nucleic Acids Res.* **2002**, *30*, 4918-4925.
- Christensen, U. B.; Pedersen, E. B. *Helv. Chim. Acta* **2003**, *86*, 2090-2097.
- Kashida, H.; Asanuma, H.; Komiyama, M. *Chem. Comm.* **2006**, 2768-2770.

28. Okamoto, A.; Ichiba, T.; Saito, I. *J. Am. Chem. Soc.* **2004**, *126*, 8364-8365.
29. Yamane, A. *Nucleic Acids Res.* **2002**, *30*, e97.
30. Filichev, V. V.; Gaber, H.; Olsen, T. R.; Jorgensen, P. T., et al. *Eur. J. Org. Chem.* **2006**, 3960-3968; Geci, I.; Filichev, V. V.; Pedersen, E. B. *Bioconjugate Chem.* **2006**, *17*, 950-957; Filichev, V. V.; Astakhova, I. V.; Malakhov, A. D.; Korshun, V. A., et al. *Chem. Eur. J.* **2008**, *14*, 9968-9980.
31. Christensen, U. B.; Wamberg, M.; El-Essawy, F. A. G.; Ismail, A. E. H., et al. *Nucleosides Nucleotides Nucleic Acids* **2004**, *23*, 207-225.
32. Han, H. Y.; Langley, D. R.; Rangan, A.; Hurley, L. H. *J. Am. Chem. Soc.* **2001**, *123*, 8902-8913.
33. Read, M.; Harrison, R. J.; Romagnoli, B.; Tanious, F. A., et al. *Proc. Natl. Acad. Sci. USA.* **2001**, *98*, 4844-4849.
34. De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N., et al. *Biochimie* **2008**, *90*, 131-155.
35. Cogoi, S.; Paramasivam, M.; Filichev, V.; Geci, I. et al. *J. Med. Chem.* **2009**, *52*, 564-568.
36. Paramasivam, M.; Cogoi, S.; Filichev, V. V.; Bomholt, N., et al. *Nucleic Acids Res.* **2008**, *36*, 3494-3507.
37. Gao, J. M.; Strassler, C.; Tahmassebi, D.; Kool, E. T. *J. Am. Chem. Soc.* **2002**, *124*, 11590-11591.
38. Gao, J. M.; Watanabe, S.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 12748-12749.
39. Nielsen, C. B.; Petersen, M.; Pedersen, E. B.; Hansen, P. E., et al. *Bioconjugate Chem.* **2004**, *15*, 260-269.
40. Langenegger, S. M.; Haner, R. *Chem. Comm.* **2004**, 2792-2793.
41. Filichev, V. V.; Vester, B.; Hansen, L. H.; Pedersen, E. B. *Nucleic Acids Res.* **2005**, *33*, 7129-7137.
42. Werder, S.; Malinovskii, V. L.; Haner, R. *Org. Lett.* **2008**, *10*, 2011-2014.
43. Bittermann, H.; Siegemund, D.; Malinovskii, V. L.; Haner, R. *J. Am. Chem. Soc.* **2008**, *130*, 15285-15287.
44. Haner, R.; Samain, F.; Malinovskii, V. L. *Chem. Eur. J.* **2009**, *15*, 5701-5708.
45. Malinovskii, V. L.; Samain, F.; Haner, R. *Angew. Chem. Int. Ed.* **2007**, *46*, 4464-4467.
46. Wilson, W. D.; Ratmeyer, L.; Zhao, M.; Streckowski, L., et al. *Biochemistry* **1993**, *32*, 4098-4104.
47. Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y., et al. *Nucleic Acids Res.* **2005**, *33*, 5887-5895; Yamana, K.; Zako, H.; Asazuma, K.; Iwase, R. et al. *Angew. Chem. Int. Ed.* **2001**, *40*, 1104-1106; Yamana, K.; Iwase, R.; Furutani, S.; Tsuchida, H., et al. *Nucleic Acids Res.* **1999**, *27*, 2387-2392.
48. Hrdlicka, P. J.; Babu, B. R.; Sorensen, M. D.; Harrit, N., et al. *J. Am. Chem. Soc.* **2005**, *127*, 13293-13299.
49. Manoharan, M.; Tivel, K. L.; Zhao, M.; Nafisi, K., et al. *J. Phys. Chem.* **1995**, *99*, 17461-17472.
50. Nakamura, M.; Shimomura, Y.; Ohtoshi, Y.; Sasa, K., et al. *Org. Biomol. Chem.* **2007**, *5*, 1945-1951.
51. Nakamura, M.; Murakami, Y.; Sasa, K.; Hayashi, H., et al. *J. Am. Chem. Soc.* **2008**, *130*, 6904-6905.
52. Kalra, N.; Parlato, M. C.; Parmar, V. S.; Wengel, J. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3166-3169.
53. Kalra, N.; Babu, B. R.; Parmar, V. S.; Wengel, J. *Org. Biomol. Chem.* **2004**, *2*, 2885-2887.
54. Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J., et al. *Tetrahedron Lett.* **1998**, *39*, 5401-5404; Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K., et al. *Tetrahedron* **1998**, *54*, 3607-3630; Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J. *JCS Chem. Commun.* **1998**, 455-456.
55. Saito, Y.; Miyauchi, Y.; Okamoto, A.; Saito, I. *Chem. Comm.* **2004**, 1704-1705.
56. Skorobogatyi, M. V.; Malakhov, A. D.; Pchelintseva, A. A.; Turban, A. A., et al. *ChemBioChem.* **2006**, *7*, 810-816; Hwang, G. T.; Seo, Y. J.; Kim, S. J.; Kim, B. H. *Tetrahedron Lett.* **2004**, *45*, 3543-3546.
57. Barbaric, J.; Wagenknecht, H. A. *Org. Biomol. Chem.* **2006**, *4*, 2088-2090.
58. Mayer-Enthart, E.; Wagenknecht, H. A. *Angew. Chem. Int. Ed.* **2006**, *45*, 3372-3375.
59. Asseline, U. *Curr. Org. Chem.* **2006**, *10*, 491-518.
60. Weisbrod, S. H.; Marx, A. *Chem. Comm.* **2008**, 5675-5685.