

Biophysical Studies of Oligonucleotides Containing Duplex Stabilising Modifications

A thesis submitted to the Board of Mathematical, Physical and Life Sciences Division, Department of Chemistry in partial fulfilment for the degree of

Doctor of Philosophy

Pietro Marafini

Wadham College, University of Oxford, Trinity 2016

Nam veluti pueri trepidant atque omnia caecis
in tenebris metuunt, sic nos in luce timemus
inter dum, nihilo quae sunt metuenda magis quam
quae pueri in tenebris pavitant finguntque futura.
Hunc igitur terrorem animi tenebrasque necessest
non radii solis neque lucida tela diei
discutiant, sed naturae species ratioque.



Titus Lucretius Carus
De Rerum Natura, Liber II

Acknowledgments

The first person I would like to thank is my supervisor Professor Tom Brown for welcoming me in his group and offering me a DPhil position. I am also thankful for his ideas and inspiration which have made the research possible.

Secondly, I am grateful to ATDBio Ltd. and the University of Oxford for funding me.

I would also like to express my thanks to Professor Afaf El-Sagheer for supervising me in the lab and always encouraging me to do more and aim higher.

A particularly special thank you goes to Anna. We started this adventure together and through thick and thin you proved to be a better friend than I could have ever hoped for. For your endless support I am forever thankful.

Thank you Sara for being such a good friend. Talking with you was not only extremely useful when I had problems with my experiments, but you also always managed to make me see things from a different perspective. And, of course, thank you so much for checking my thesis.

Thank you Arun for always being so helpful and for being so keen on answering my questions. I am also very thankful for all the effort you put in correcting my thesis.

I cannot forget all the amazing companions I had during my DPhil. Agnes, Xiaomei, Jieqiong, Marie, I am ever so grateful of how you welcomed me in the group and for the opportunity I had to share most of these four year together as friends and colleagues. I also would like to thank Adam, Joanna, Lapatrada, Nayoun, Jack and all the other group members, for making my DPhil an unforgettable experience.

I am really grateful for all the support I received from my friends and my family and I also wish to thank all the people I am forgetting to mention here that I have been so lucky to cross my path with during the past years.

Last but not least I would like to thank my parents for their continued encouragement and support, even financial. Without you I would have not been able to embark on this life changing journey.

Abstract

Chemically modified DNA and RNA probes have been used in numerous diagnostic (*e.g.* fluorogenic PCR and *in situ* hybridisation) and therapeutic (*e.g.* exon-skipping and antisense) applications to increase duplex stability and facilitate mismatch discrimination. There is a constant need for well-studied modifications, and to this end X-pyrene was evaluated as a new cytosine analogue that combines enhanced π-stacking, hydrogen bonding and electrostatic interactions to greatly increase the stability of bulged DNA duplexes and DNA/RNA hybrid duplexes. X-pyrene is highly selective for guanine, and even more importantly duplex stability is reduced dramatically when X-pyrene or a neighbouring base is mismatched. NMR and fluorescence studies indicate that the pyrene moiety stacks within the DNA double helix. To enable easier access to this modification, a reductive amination labelling strategy was developed to produce pyrene-functionalised nucleobase analogues which allowed an easier synthesis of X-pyrene. Amino C6 dT was successfully functionalised with this methodology as well, obtaining the duplex stabilising mono- and di-labelled products. Additionally, a much simpler pyrene-modified thymidine monomer (U-pyrene) was selected for further studies. Single and multiple additions of U-pyrene were found to significantly increase the melting temperature of dsDNA without perturbing the B-DNA duplex, as shown by UV and CD spectroscopy. Interestingly, thermodynamic studies showed that four additions of U-pyrene increased the binding constant of a 14-mer oligonucleotide to its complementary strand 10000-fold when compared to the unmodified control. Two oligonucleotides containing a single addition of U-pyrene were also crystallised and their structures were studied by X-ray diffraction. To the author's knowledge these represent the first examples of X-ray crystal structures of a chemically modified pyrene oligonucleotide. They also proved to be of great interest as they highlight the capacity of U-pyrene to facilitate the production of interesting nanostructures.

Abbreviations

A	Adenine
A	Absorbance
Å	Ångström (10^{-10} m)
aq	Aqueous
C	Cytosine
C or c	Molar concentration
°C	Degrees centigrade
CD	Circular dichroism
$CDCl_3$	Deuterated chloroform
cm	Centimetre
COSY	^1H – ^1H correlation spectroscopy
d	Deoxy
d	Doublet (NMR)
DCM	Dichloromethane
DEPT	Distortionless Enhancement through Polarisation Transfer
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxy(nucleotide)triphosphate
ds	Double strand
eq	Equivalents

ESI	Electrospray ionisation
EtOAc	Ethyl acetate
EtOH	Ethanol
FRET	Fluorescence/Förster Resonance Energy Transfer
G	Guanine
G	Gibbs energy
g	Gram
H	Enthalpy
HMBC	Heteronuclear Multiple Bond Correlation spectroscopy
HMQC	Heteronuclear Multiple Quantum Correlation spectroscopy
HPLC	High-performance liquid chromatography
Hz	Hertz
J	Coupling constant
K_b	Binding constant
K_d	Dissociation constant
K_{eq}	Equilibrium constant
KF	Klenow fragment
l	Light path length in cm
M	Molar concentration
m	Metre
m	Multiplet (NMR)
MAD	Multiple wavelength anomalous dispersion/diffraction

MeCN	Acetonitrile
MeOH	Methanol
mg	Milligram
MHz	Megahertz
mL	Millilitre
mM	Millimolar concentration
MR	Molecular replacement
MS	Mass Spectrometry
MW	Molecular weight
NHS	*N*-Hydroxysuccinimide
nm	Nanometre
NMR	Nuclear Magnetic Resonance
nM	Nanomolar concentration
ODN	Oligonucleotide
PCR	Polymerase Chain Reaction
pM	Picomolar concentration
pol α	Human DNA polymerase α
ppm	Parts per million
q	Quartet (NMR)
R	Ideal constant of gases
R_0	Förster distance
R_f	Retention factor

RNA	Ribonucleic acid
RP	Reversed-phase
RT	Room temperature
S	Entropy
s	Singlet (NMR)
S_0	Ground state
S_n	nth singlet state
SAD	Single wavelength anomalous dispersion/diffraction
ss	Single strand
T	Thymine
T	Temperature
t	Triplett (NMR)
T_m	Melting temperature
T_n	nth triplet state
T7 RNAP	T7 RNA polymerase
tC	1,3-diaza-2-oxophenothiazine
tC_{nitro}	7-nitro-1,3-diaza-2-oxophenothiazine
tC^O	1,3-diaza-2-oxophenoxazine
TLC	Thin Layer Chromatography
U	Uracil
UV	Ultraviolet radiation of the electromagnetic spectrum
Δ	Variation

δ	Chemical shift in parts per million
ε	Extinction coefficient
λ	Wavelength
µL	Microlitre
µM	Micromolar
Φ	Quantum yield

Table of Contents

Chapter 1 Introduction 1

1.1 Nucleic Acids 3

1.1.1 Structure of Nucleotides 3

1.1.2 Structure of Oligonucleotides and Duplexes 5

1.2 Biophysical Techniques 9

1.2.1 UV Melting 9

1.2.2 Circular Dichroism 10

1.2.3 Fluorescence 12

1.3 Modified Oligonucleotides 14

1.4 Fluorescent Base Analogues: A Brief Introduction 15

1.5 The Tricyclic Cytosine Family 17

1.5.1 Basic Properties of tC and its Derivatives 17

1.5.2 Biocompatibility of tC and its Derivatives 21

1.6 Pyrene as a Nucleotide-Functionalising Moiety 25

1.7 Objectives and Thesis Overview 32

Chapter 2 X-pyrene 35

2.1 Preface 37

2.2 Introduction 37

2.3 Materials and Methods 39

2.3.1 Oligonucleotides Synthesis, Purification and Analysis 39

2.3.1.1 Synthesis 39

2.3.1.2 Purification 39

2.3.1.3 Analysis 40
2.3.2 Oligonucleotides Labelling with NHS Ester Labels 41
2.3.2.1 6-Azidohexanoic Acid NHS Ester Labelling of ODN-2 41
2.3.2.2 1-Pyrenebutyric Acid NHS Ester Labelling of ODN-2 and ODN-5 41
2.3.3 UV Melting 41
2.3.4 Fluorescence Studies 42
2.3.5 Circular Dichroism 43
2.3.6 Error Calculations 43
2.3.7 Molecular Visualisation Based on NMR Data 44
2.4 UV Melting Studies 44
2.5 Structural Studies and Molecular Visualisation 53
2.6 Fluorescence Studies 55
2.7 Conclusions 62
Chapter 3 Pyrene Labelling of Oligonucleotides by Reductive Amination 65
3.1 Introduction 67
3.1.1 Post-Synthetic Labelling of Modified Oligonucleotides 67
3.1.2 Reductive Amination Labelling 68
3.2 Materials and Methods 70
3.2.1 Synthesis of 1-Pyrenebutyraldehyde 70
3.2.2 Oligonucleotides Synthesis, Purification and Analysis 72
3.2.2.1 Synthesis 72
3.2.2.2 Purification 73
3.2.2.3 Analysis 73

3.2.3 Reductive Amination Labelling with 1-Pyrenebutyraldehyde 74
3.2.3.1 Labelling of G-Clamp 74
3.2.3.2 Labelling of Amino C6 dT 75
3.2.3.3 Labelling of 5′-Amino-Modifier C6 and 3′-Amino-Modifier C7 76
3.2.4 UV Melting 76
3.2.5 Error Calculations 77
3.3 Reductive Amination 77
3.4 UV Melting Studies 84
3.5 Conclusions 91
Chapter 4 U-pyrene 93
4.1 Preface 95
4.2 Introduction 95
4.2.1 UV Melting and Thermodynamics 96
4.2.2 Crystallography of Oligonucleotides 99
4.3 Materials and Methods 101
4.3.1 Oligonucleotides Synthesis, Purification and Analysis 101
4.3.2 UV Melting and Thermodynamics 102
4.3.2.1 General UV Melting Procedure 102
4.3.2.2 UV melting for thermodynamics studies 103
4.3.3 Error Calculations 104
4.3.4 CD Spectroscopy 104
4.3.5 X-Ray Crystallography 105
4.3.5.1 Crystallisation of Oligonucleotides 105
4.3.5.2 Diffraction Data Collection 105
4.3.5.3 Data Processing, Phase Determination, Model Building and Refinement 106

4.4 UV Melting and Thermodynamics 107
4.5 Circular Dichroism 114
4.6 X-Ray Crystallography 117
4.6.1 Preface 117
4.6.2 Crystallisation 117
4.6.3 Data Collection and Crystallographic Statistics 118
4.6.4 Experimental Phasing, Model Building and Refinement 120
4.6.5 CRY-1 122
4.6.6 CRY-2Br 126
4.7 Conclusions 130

Chapter 5 Conclusions and Future Work 133

References 139

Chapter 1

Introduction

1.1 Nucleic Acids

The discovery of the substance that would later be known as deoxyribonucleic acid (DNA) was made by Johann Friedrich Miescher in 1869.[1] He extracted the nuclei of leucocytes, taken from the pus on fresh surgical bandages, and he obtained a precipitate called "nuclein". When the acidic properties of "nuclein" were discovered, the name was changed to "nucleic acid" (Richard Altmann, 1889). Most importantly, in 1944 Oswald, MacLeod and McCarty suggested that DNA carries genetic information,[2] and their finding was confirmed in 1952 by Alfred Hershey and Martha Chase.[3] A year later James D. Watson and Francis H. C. Crick proposed the double helix structure for DNA.[4]

1.1.1 Structure of Nucleotides

DNA and RNA are biological polymers built up from monomers called nucleotides that are linked together by phosphodiester linkages. Nucleotides have three main components: a nitrogen-containing heterocyclic base, a pentose sugar (2′-deoxy-D-ribose in DNA and D-ribose in RNA) and a phosphate (Figure 1.1). There are five different fundamental nitrogenous bases and only four are present in DNA. These are divided in two groups (Figure 1.2): the purines (adenine and guanine) and the pyrimidines (cytosine, thymine and uracil).

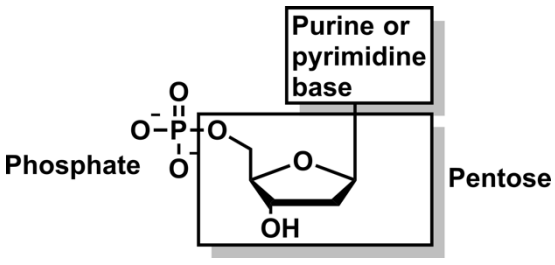


Figure 1.1 Structure of Nucleotides. The pentose here depicted is 2′-deoxy-D-ribose.

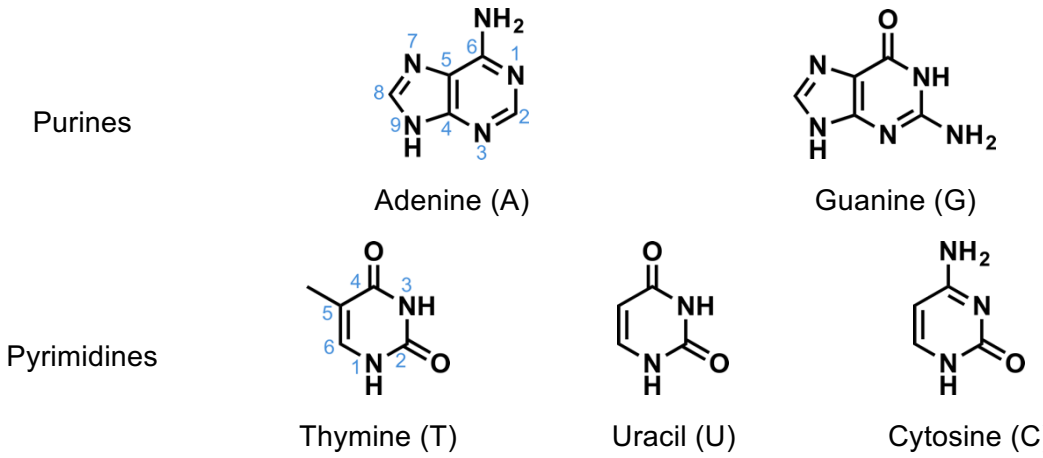


Figure 1.2 The five nitrogenous bases present in DNA and RNA. Thymine is present in DNA only, while uracil is present only in RNA. The numbering of the nucleobases is in light blue.

The nucleobase is linked to the pentose sugar *via* an *N*-glycosidic bond connecting the C-1′ of the sugar with N-9 of purines or N-1 of pyrimidines. The nitrogenous base can be located on the same side of the pentose ring as its 5′-hydroxyl group (β-anomer) or on the opposite side (α-anomer), with naturally occurring oligonucleotides presenting nucleosides in the β-configuration (Figure 1.3).

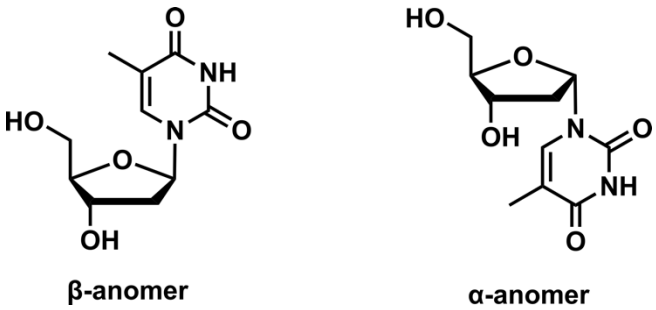


Figure 1.3 α- and β-anomer of a DNA nucleoside.

The nucleobase can rotate around the *N*-glycosidic bond, having two extreme conformations (Figure 1.4). The *syn* conformer has the bigger N-3 (purine) or O-2 (pyrimidine) above the pentose ring, and it is therefore more sterically hindered than the *anti* conformer that has the smaller H-8 (purine) or H-6 (pyrimidine) above the sugar. For this reason, naturally occurring nucleic acids prefer the *anti* conformation.

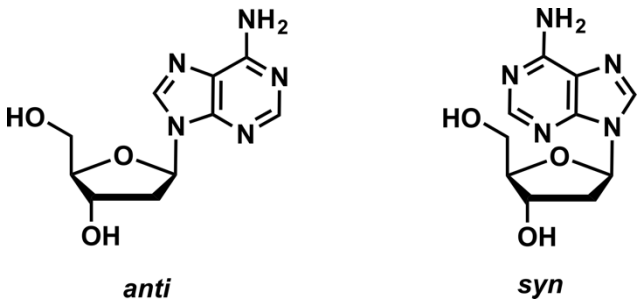


Figure 1.4 *Anti* and *syn* conformations of deoxyadenosine.

In order to reduce electronic and steric interactions between the substituents of furanose, the ring is twisted out of plane generating the so called sugar puckering. There are two form of sugar pucker, identifiable by the position of the 2′ and 3′ carbons of the pentose ring (Figure 1.5). The South (S) configuration presents the C-2′ on the same side of the ring (*endo*) as the C-5′, while the North (N) configuration presents the C-3′ in the *endo* position. In DNA 2′-deoxy-D-ribose usually adopts the C-2′ *endo* (S) conformation, while in RNA D-ribose usually prefers the C-3′ *endo* (N) conformation.

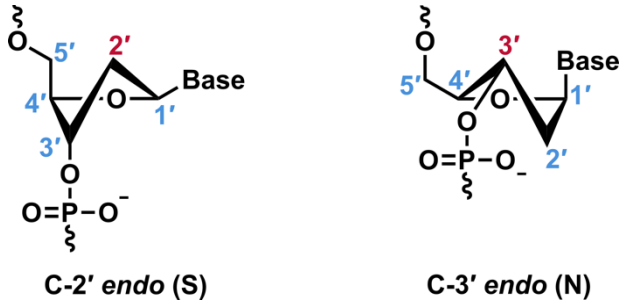


Figure 1.5 C-2′ *endo* and C-3′ *endo* sugar conformations.

1.1.2 Structure of Oligonucleotides and Duplexes

One of the fundamental steps in understanding DNA structure was made by Erwin Chargaff who noted that the ratio of purines to pyrimidines in cells is equal (*i.e.* 1:1).[5-7] This was followed by X-ray diffraction studies by Rosalind Franklin and Maurice Wilkins, that showed that the secondary structure of DNA is helical and humidity dependent.[8-10] The key discovery in consolidating this information was published in

1953 by James Watson and Francis Crick,[4] who proposed the structure of B-DNA as a double helix with hydrogen bonds joining the DNA nucleobases of the two different strands in what are now known as Watson-Crick base pairs (Figure 1.6).

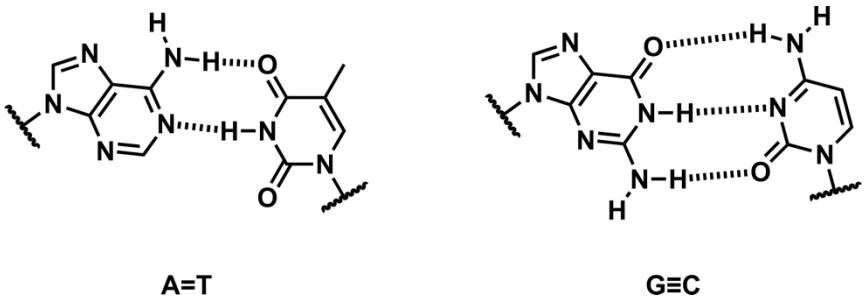


Figure 1.6 Watson-Crick base pairing. The G≡C base pair stabilises the DNA more than the A=T base pair because of the additional hydrogen bond.

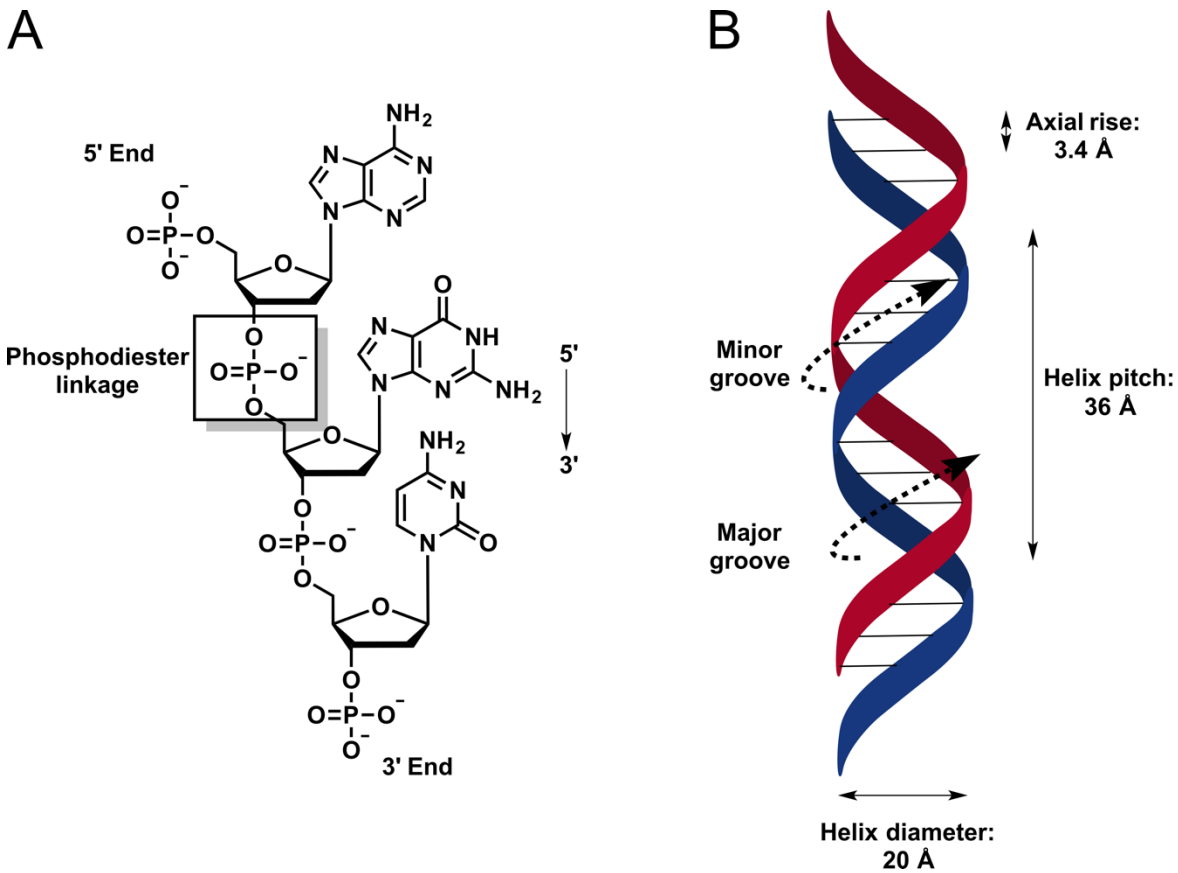


Figure 1.7 DNA structure. **A)** Single-stranded DNA (ssDNA) with a phosphodiester linkage highlighted. **B)** Schematic representation of the B-double helix, there are 10.5 base pairs per turn of the helix (3.6 nm).

In an oligonucleotide, as for example a single strand of DNA, nucleotides are covalently linked through the phosphate groups forming a phosphodiester linkage. As shown in Figure 1.7 A, the 5′-phosphate group of one nucleotide is joined to the 3′-hydoxyl group of the next nucleotide. The B-DNA double helix (Figure 1.7 B) is made by two antiparallel single strands that interact through hydrogen bonding between nitrogenous bases, as proposed by Watson and Crick (Figure 1.6). The base pairs that are preferentially formed are always the same: cytosine binds specifically to guanine (G≡C) and adenine to thymine (A=T).

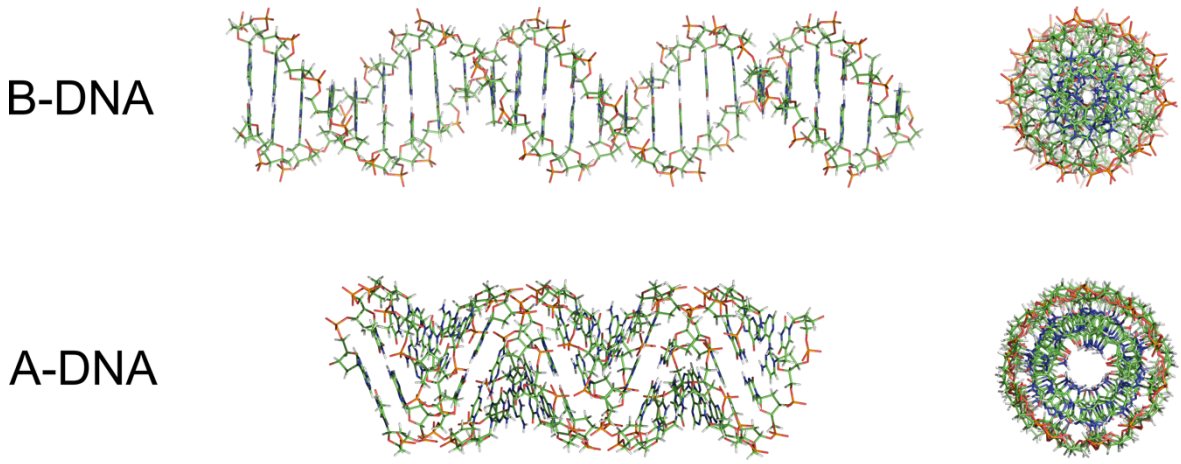


Figure 1.8 Side and top view of the structure of B-DNA and A-DNA. The two structures have the same sequence, the models were produced using Hyperchem 7.

Two major conformations of DNA are known: A- and B-DNA (Figure 1.8).[11] Both B-DNA and A-DNA are right-handed helices, but while the B-form of DNA, which is the most biologically important and common conformation, is formed in environments with a high humidity level, A-DNA is present when the humidity is low. The A-form is also the major conformation adopted by RNA-RNA and RNA-DNA duplexes.

B-form duplexes (Figure 1.7 B and Figure 1.8) present two grooves of similar depth, but one (major groove, about 12 Å) is wider than the other (minor groove, about 6 Å). The nucleobases are stacked upon each other, with an interbase distance of 3.4 Å, and

they are almost perpendicular to the helix axis. There are 10.5 bases for each 36 Å helix pitch and the helix diameter is 20 Å. The furanose sugars are in the C-2′ *endo* conformation with *anti* nucleobases.[11-13]

A-form duplexes (Figure 1.8) are slightly wider (23 Å of helix diameter) and more compact, with 11 bases for each 28 Å helix pitch. Therefore, the nucleobases are closer to each other, with an interbase distance of only 2.6 Å. Because of that, in order to maintain the normal van der Waals separation, the nucleobases are significantly tilted compared to the B-form. Another major difference with B-form nucleic acids is the fact that base pairs are almost centred over the helical axis in B-form, while in A-form they are displaced away from the central axis creating a 3 Å wide hollow core. The furanose sugars are in the C-3′ *endo* conformation with *anti* nucleobases. As a result of these features, the minor groove is wide and shallow and the major groove is narrow and deep.[11-13]

The overall stability of double-stranded DNA (dsDNA) depends on the sequence, which determines the base pair composition. The higher the G≡C/A=T ratio the more stable the resulting dsDNA, since the G≡C base pair has an additional hydrogen bond compared to the A=T base pair. Another factor that stabilises the dsDNA is the stacking of the nitrogenous bases when the double helix is formed, the overlapping π-orbitals of neighbouring base pairs interact stabilising the molecule.[11,14]

1.2 Biophysical Techniques

1.2.1 UV Melting

The capacity of oligonucleotides to absorb light in the ultraviolet (UV) region arises mainly from the pyrimidine and purine nucleobases and their n-π* and π-π* transitions.[15] Each nucleotide has a different extinction coefficient (ε) and a different absorption maximum (λ_{max}) and therefore oligonucleotides have an average λ_{max} around 260 nm.[11] The intensity and the maximum of absorption not only depend on the sequence of the oligonucleotide under study, but are also influenced by, for example, the salt concentration, the pH of the solution and importantly base-base stacking.[11] In fact, interactions between nucleobases stacking on each other cause a decrease in the extinction coefficient. This hypochromic effect is very useful for following the transitions from duplexes to single strands of oligonucleotides, as the absorption intensity of the former is lower than the latter. By monitoring the UV absorption while increasing the temperature it possible to record a melting curve and derive from it the melting temperature (T_m), as a measurement of the thermal stability of the duplex under study. If there is only one melting transition, the T_m will correspond to the mid-point of the sigmoidal curve (Figure 1.9).

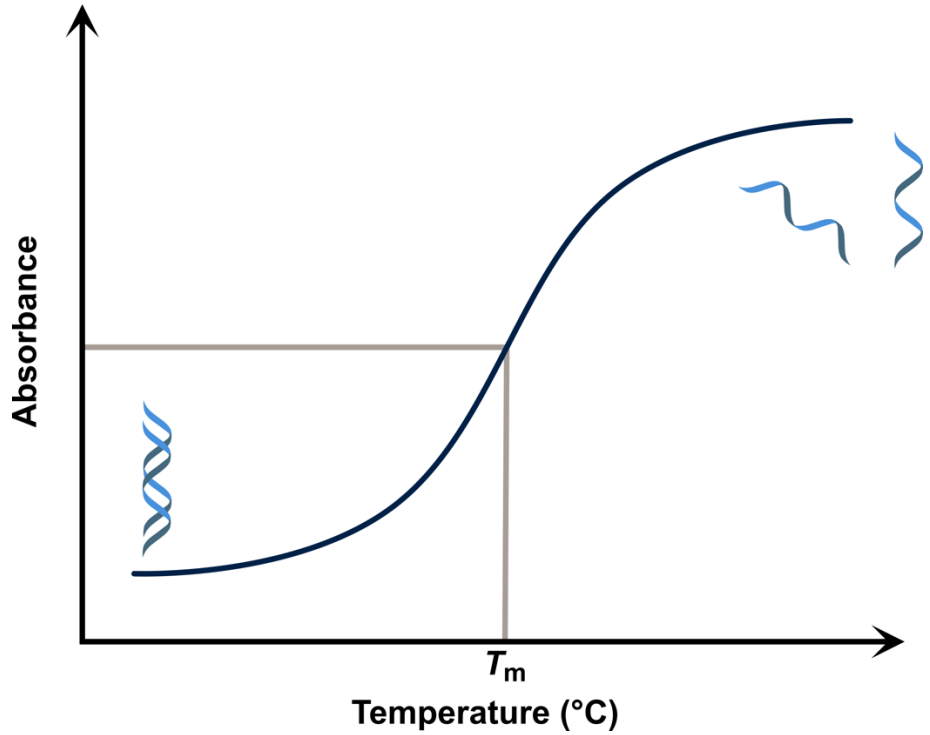


Figure 1.9 UV melting curve of a transition from dsDNA to ssDNA when increasing the temperature. The T_m correspond to the mid-point of the sigmoidal curve.

1.2.2 Circular Dichroism

Circular dichroism (CD) spectroscopy is a method commonly used to determine the ensemble conformation of nucleic acids and proteins. A CD spectrum is generated by the differential absorption of the left (L) and right (R) circularly polarised light by chiral molecules,[11,15,16] that can be expressed as

$$\Delta A = A_L - A_R.$$

Considering the Lambert-Beer law, it is possible to relate the differential absorbance (ΔA) to the differential extinction coefficient ($\Delta\varepsilon$):

$$\Delta A = c \cdot l \cdot \Delta\varepsilon,$$

where

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R. \qquad \text{Equation 1.1}$$

A CD spectrum is often presented as a plot of $\Delta\varepsilon$ *versus* the wavelength of light or the molar ellipticity $[\theta]$ can be used instead of $\Delta\varepsilon$. The two quantities can be converted using the following relationship:

$$[\theta] = 3298 \cdot \Delta\varepsilon. \qquad \text{Equation 1.2}$$

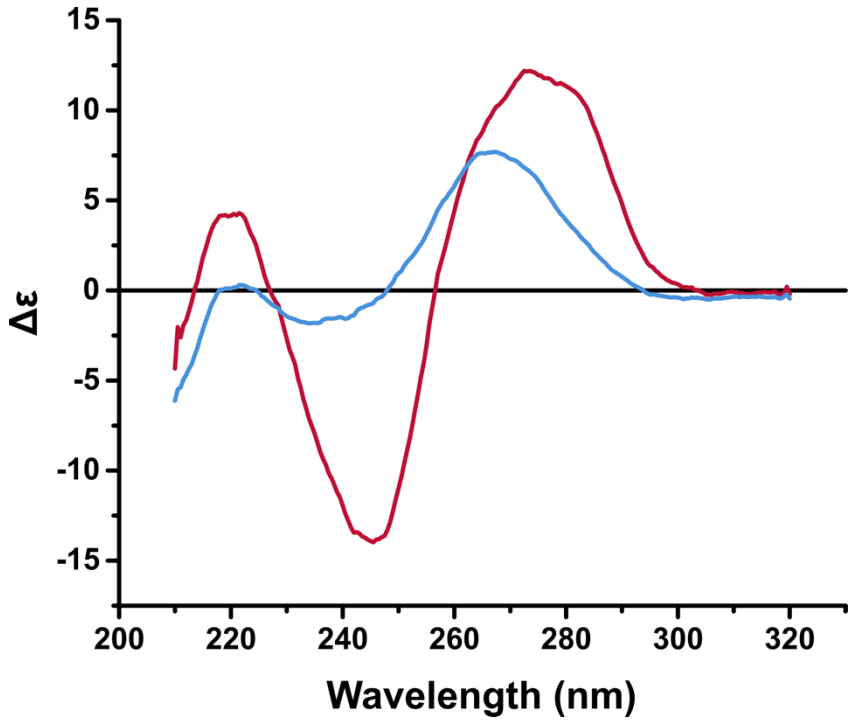


Figure 1.10 CD spectra of a B-form duplex (red) and an A-form duplex (light blue).

In the field of nucleic acids, CD spectroscopy is useful to quickly determine the form of oligonucleotide duplexes, as different conformations present different spectra. Figure 1.10 reports the typical spectra for B-form and A-form duplexes. Nucleobases are not CD-active on their own, however they present an induced activity when in a chiral environment as in a DNA duplex.[11,15] For the same reason CD spectroscopy is useful to study nucleic acid-ligand interactions, with the ligand acquiring an induced signal when interacting with a duplex. As the intensity of the peaks in a CD spectrum follows the Lambert-Beer law, the intensity of the signal depends on the amount of the nucleic

acid-ligand complex formed.[11] CD spectroscopy can also help to study and differentiate between different types of nucleic acid-ligand interactions (*e.g.* intercalation, groove binding) and it can even provide details about the relative orientation with respect to the environment of the ligand.[17-19]

1.2.3 Fluorescence

The process during which a fluorophore absorbs a photon and then emits another photon of lower energy is defined as fluorescence. To schematically describe the process a Jablonski diagram (Figure 1.11) is often used.[20-22] When a molecule absorbs a photon (absorption, 10^{-15} s), it is excited from the ground state (S_0) to a higher energy electronic state (*e.g.* S_1 or S_2). Fast non radiative processes (internal conversion, 10^{-14}–10^{-11} s, and vibrational relaxation, 10^{-14}–10^{-11} s) take the high energy molecule to the lowest vibrational level of the S_1 electronic state. From here a photon is emitted (fluorescence, 10^{-9}–10^{-7} s) and so the molecule returns to its ground state. The wavelength of the emitted photon is higher and therefore the energy is lower. The difference in wavelength between the absorbed and the emitted photon is called the Stokes shift, after the physicist George Stokes.

An alternative way by which the molecule can return from the high energy state to the ground state is intersystem crossing (ISC, 10^{-8}–10^{-3} s). ISC changes the spin multiplicity, which takes the molecule from the lowest vibrational level of the S_1 electronic state to the triplet state T_1. From here, after vibrational relaxation to the lowest vibrational level, a photon is emitted in a process called phosphorescence. This transition from S_1 to T_1 is "forbidden" by classical quantum mechanics, in fact the rate of phosphorescence is very slow (10^{-4}–10^{-1} s).

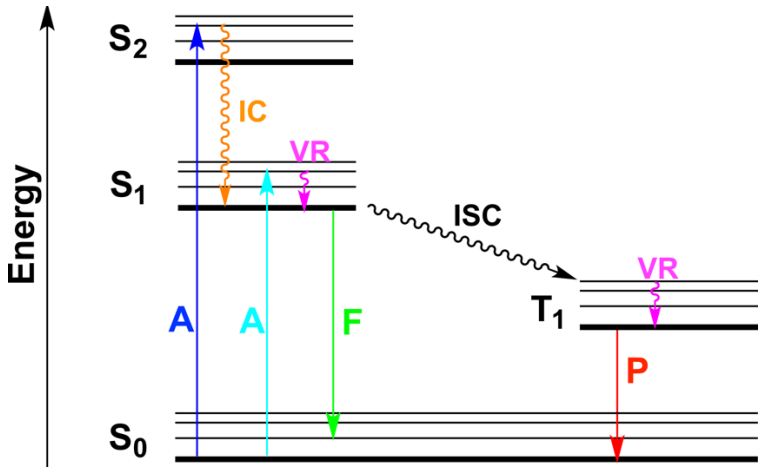


Figure 1.11 Jablonski diagram (energy diagram) depicting the transition between different electronic states of a molecule. A (absorption of a photon), F (fluorescence), P (phosphorescence), IC (internal conversion), VR (vibrational relaxation), ISC (intersystem crossing), S_0 (ground state), S_n (singlet states), T_n (triplet states).

The fundamental properties required to determine the brightness of a fluorophore are the extinction coefficient and its quantum yield.[20-22]

The extinction coefficient, which is derivable from the Lambert-Beer law (Equation 1.3), is a measure of the efficiency with which the light is absorbed by the fluorophore.

$$\varepsilon = \frac{A}{c \cdot l} \qquad \text{Equation 1.3}$$

Where ε is the extinction coefficient at a given wavelength (λ) in $M^{-1}cm^{-1}$, A is the absorbance at a given λ, c is the concentration of the fluorophore in M and l is the path length in cm.

The quantum yield (Equation 1.4) is the ratio between the photons emitted and the photons absorbed by the fluorophore.

$$\Phi = \frac{\text{number of photons emitted}}{\text{number of photons absorbed}} \qquad \text{Equation 1.4}$$

A bright fluorophore is capable of absorbing and emitting a high number of photons and therefore, it has a high extinction coefficient and quantum yield.

1.3 Modified Oligonucleotides

With the discovery of the solid-phase DNA/RNA synthesis, the introduction of unnatural modifications into oligonucleotides has become increasingly common. These modifications have been designed to expand the possible applications of nucleic acids by enhancing their natural properties and introducing new ones as well. One of the areas where nucleic acids are frequently modified is for antisense applications. Antisense agents are short oligonucleotides capable of binding their complementary sites on the messenger RNA (mRNA), preventing the expression of that particular genetic information.[23-26] Antisense oligonucleotides have been chemically modified in order to improve their cellular uptake and to increase their biostability by increased resistance to nucleases.[24-27] Additionally, strong hybridisation of the antisense oligonucleotides to their targets is fundamentally important for their function, and modifications increasing the thermal stability of the duplexes formed are therefore extremely useful. In fact, the correlation between antisense activity with hybridisation affinity and melting temperature has been observed.[28-31]

Another area where the modification of oligonucleotides is highly beneficial is the design and production of probes for specific DNA sequences, which could be used for diagnostic or forensic applications. Probes need to be long enough to be specific for their target, capable of binding it tightly, and to be able to differentiate a mutated sequence from the original one.[32,33] It is also necessary to have methodologies for detecting and differentiating among different probes and their responses. For example, the modification of oligonucleotides with moieties that are fluorescent is very useful for real-time PCR (polymerase chain reaction) probes like HyBeacons,[34-38] while the use of dyes that are Raman active is essential to follow the electrochemical

denaturation on a nanostructured gold surface *via* surface enhanced Raman spectroscopy (SERS).[39-44]

As a final example of the use of modified oligonucleotides is it worth mentioning the nanostructure field. Modified oligonucleotides have been use to stabilise, study, control, conjugate and enhance the properties of different simple nanoconstructs.[45-50]

1.4 Fluorescent Base Analogues: A Brief Introduction

Fluorescent base analogues are interesting because the fluorescent moiety is not attached externally to the DNA, but is directly part of modified nucleotides.[51] An advantage of having intrinsically fluorescent base analogues is that perturbations to the DNA structure and behaviour are minimal and it is therefore possible to study nucleic acid systems in conditions that are more similar to the natural ones, preserving so the "native behaviour". Moreover, having fluorescent probes near to the site of interest allows higher level of details. On the other hand, fluorescent base analogues are on average far less fluorescent and photostable than commonly used "external" dyes.[51]

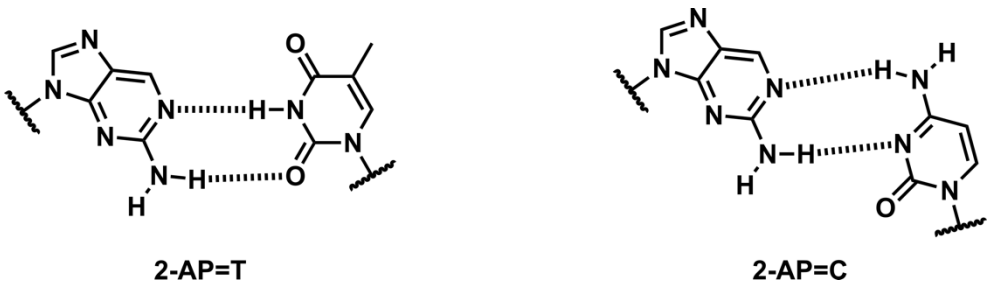


Figure 1.12 2-aminopurine base paired with thymine (left) and with cytosine (right).

One of the most used fluorescent base analogues is 2-aminopurine (2-AP, Figure 1.12).[51] 2-AP is an adenine analogue and so it forms stable base pairs with thymine and uracil, but unfortunately it also forms quite a stable base pair with cytosine.[52] The

lowest absorption band of 2-AP is at 305 nm (extinction coefficient ε = 6000 $M^{-1}cm^{-1}$), therefore it is possible to excite 2-AP selectively with a limited involvement of natural nucleotides or amino acids. The quantum yield of the free monomer in solution is high (0.68) but decreases around 100-fold when 2-AP is incorporated in single- or double-stranded nucleic acids.[53]

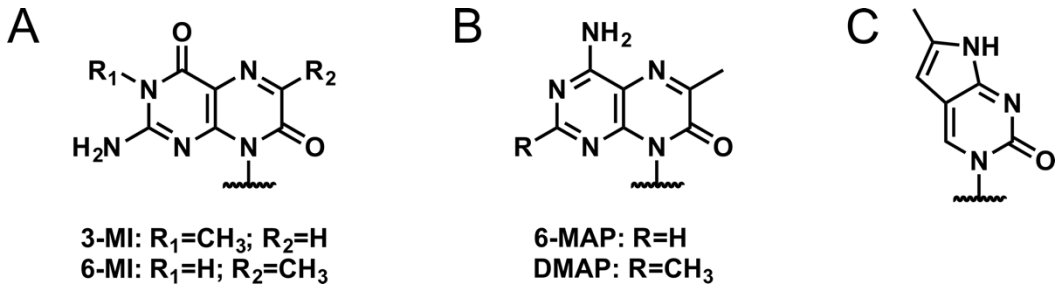


Figure 1.13 Fluorescent base analogues. **A)** 3-MI and 6-MI. **B)** 6-MAP and DMAP. **C)** pyrrolo-dC.

Another family of fluorescent base analogues are the pteridines (Figure 1.13 A and B).[54,55] Among them there are 3-methylisoxanthopterin (3-MI), 6-methylisoxanthopterin (6-MI), 4-amino-8-(2-deoxy-β-D-ribofuranosyl)5′-O-dimethoxy-trityl-6-methyl-7(8H)-pteridone (6-MAP) and 4-amino-8-(2-deoxy-β-D-ribofuranosyl)5′-O-dimethoxy-trityl,-2,6-dimethyl-7(8H)-pteridone (DMAP). 3-MI and 6-MI are guanine analogues, while 6-MAP and DMAP are adenine analogues. They are highly quenched when incorporated into DNA and the quantum yields vary from <0.01 to 0.3 depending on the local base pairs.[55,56] For all four pteridines, the lowest absorption band is between 310 nm and 350 nm and so a selective excitation is possible. A small problem with these analogues (except 6-MI) is that the melting temperature of dsDNA is lowered, which means that the DNA stability is reduced.

Finally, a pyrrolo-dC family is also known.[51] Pyrrolo-dC (Figure 1.13 C) base pairs selectively with G (G>>T>A>C) and does not negatively influence dsDNA stability.[57] Its

lowest absorption band is at 350 nm (ε = 5900 $M^{-1}cm^{-1}$) and therefore it is possible to selectively excite it. The quantum yield of the free monomer is 0.20 but it is highly reduced (down to 0.03) when incorporated into oligonucleotides.[51,57]

1.5 The Tricyclic Cytosine Family

1.5.1 Basic Properties of tC and its Derivatives

While all the fluorescent base analogues presented before have a quantum yield that decreases significantly upon incorporation into oligonucleotides, the tricyclic cytosine analogues maintain their high quantum yield. This is essential for many applications such as Fluorescence/Förster Resonance Energy Transfer (FRET) and anisotropy measurements. There are three tricyclic cytosine analogues commercially available (Figure 1.14): tC (1,3-diaza-2-oxophenothiazine), tC^O (1,3-diaza-2-oxophenoxazine) and tC_{nitro} (7-nitro-1,3-diaza-2-oxophenothiazine). tC and tC^O are highly fluorescent, whereas the quantum yield of tC_{nitro} is essentially zero and so it acts as a quencher.[51]

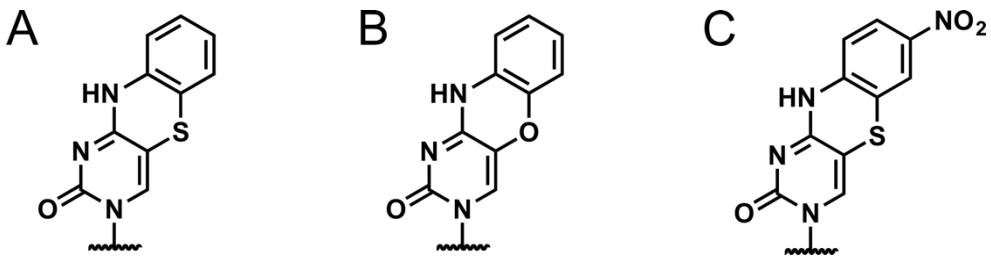


Figure 1.14 Commercially available tricyclic cytosine analogues. **A)** tC. **B)** tC^O. **C)** tC_{nitro}.

Matteucci *et al.* developed tC and tC^O for antisense purposes, without studying their fluorescent properties.[58] In fact tC, tC^O and also tC_{nitro} increase the thermal stability of DNA duplexes by on average 2.7 °C per C substituted.[58-61] This is probably due to the extra aromatic ring being in the major grove, where it can interact with neighbouring bases, which in turn enhances stabilising π-stacking interactions. Moreover, Matteucci

et al. showed that tC^O favours cellular permeation (and so the antisense activity) of short oligonucleotides.[27] Only afterwords, Wilhelmsson *et al.* studied the fluorescent properties of first tC, then tC^O and finally tC_{nitro}.[61-65]

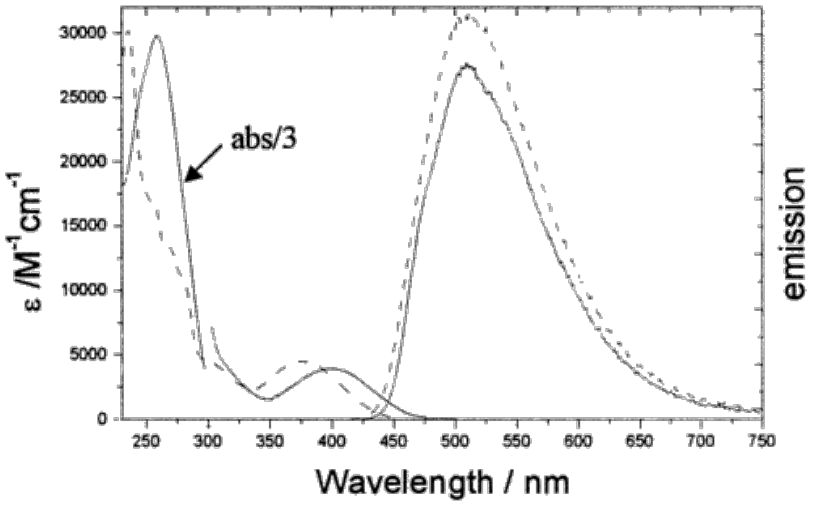


Figure 1.15 Tricyclic Isotropic absorption spectra (left) and emission spectra (right) of free tC (dashed curves) and tC-PNA (solid curves). Adapted with permission from [62]. Copyright 2001 American Chemical Society.

The free tC base presents the lowest energy absorption peak at 375 nm (ε = 4500 M^{-1} cm^{-1}), but when incorporated into oligonucleotides the band is red shifted to 395 nm with a lower intensity (ε = 4000 $M^{-1}cm^{-1}$, Figure 1.15).[51,62-64] The fact that the excitation band is far from 260 nm and 280 nm means that tC can be excited selectively in the presence of natural nucleotides or amino acids. The emission maximum of tC is at 505 nm. Another interesting property of tC is its stable quantum yield. The free monomer has a high quantum yield (0.20), which is not significantly modified when tC is incorporated in oligonucleotides. In fact, irrespective of neighbouring bases, the quantum yield is still about the same in ssDNA (0.17–0.24) and in dsDNA (0.16–0.21).[62,64]

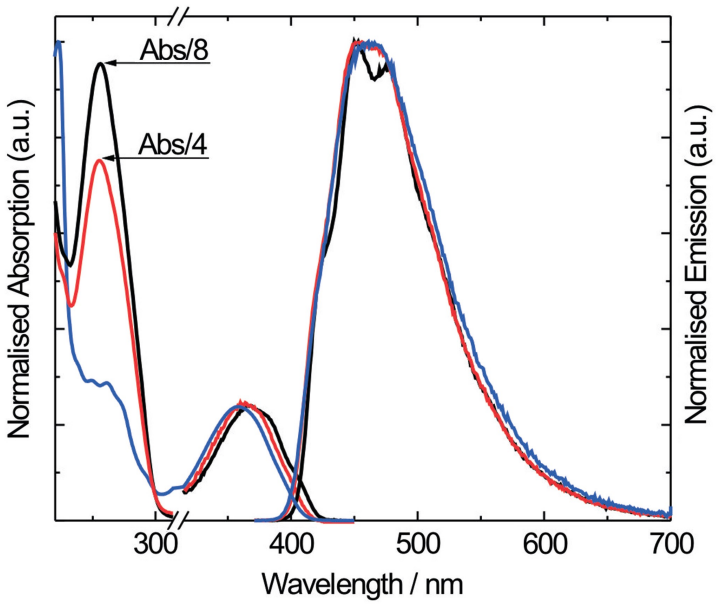


Figure 1.16 Isotropic absorption and emission spectra of the tC^O nucleoside (blue), tC^O-containing single strand (red) and tC^O-containing double strand (black). Adapted with permission from [61]. Copyright 2008 Oxford University Press.

The lowest energy absorption peak of free tC^O is at 360 nm but, similarly to tC when incorporated into single- or double-stranded oligonucleotides the band is slightly red shifted to 395 nm (Figure 1.16). The extinction coefficient is reasonably high (ε = 9000 $M^{-1}cm^{-1}$) and the emission maximum is 465 nm.[61] The free monomer has a high quantum yield (0.30), which is modified to some extent when tC^O is incorporated in oligonucleotides. While tC maintains the same quantum yield irrespective of neighbouring bases, this is not the same for tC^O. On average the quantum yield for ssDNA is still around 0.3 but practically, since tC^O is sensitive to the surrounding bases, the quantum yield ranges between 0.14 and 0.41. For dsDNA the quantum yield is lower on average (0.22) and there is a small variation depending on the neighbouring bases (0.17–0.3), if compared to the ssDNA case.[51,61] With an intelligent design of the bases surrounding tC, it is possible to have little or significant differences in the quantum yield between single- and double-stranded DNA. The high, stable quantum yield and the high extinction coefficient make tC^O one of the brightest fluorescent base

analogues commercially available at the present. tC^O has been used in various applications, for example it was used by Börjesseon *et al.* as a probe to detect individual melting processes in a complex nucleic acid structure containing various secondary structure motives.[66] Similarly Sandin *et al.* used tC^O to detect the denaturing of each side of a small DNA hexagon.[45] tC^O was also used by Stengel *et al.* to label DNA in PCR reactions with Deep Vent DNA polymerase.[67] The resulting PCR product was fluorescent enough to be directly visible in a gel under a longwave UV light, without staining with ethidium bromide.

The lowest energy absorption peak of tC_{nitro} is at 440 nm (ε = 5400 $M^{-1}cm^{-1}$, Figure 1.17) and its quantum yield is essentially zero (it increases only below 225 K).[51,59,65] Due to its lack of emission, tC_{nitro} is a quencher and importantly its absorption peak overlaps with the emission peaks of both tC and tC^O. This allows the creation of a FRET pair of nucleic acid base analogues. Börjesseon *et al.* were able to use tC^O and tC_{nitro} to carry out FRET measurements, thus reporting the first working nucleic acid base analogue FRET pair in the literature.[59] The advantages of using two nucleic acid base analogues for FRET experiments are evident. Since the modified nucleotides are firmly fixed inside the oligonucleotide structure, it is possible to know their position and orientation with great precision, thereby increasing FRET signal accuracy.

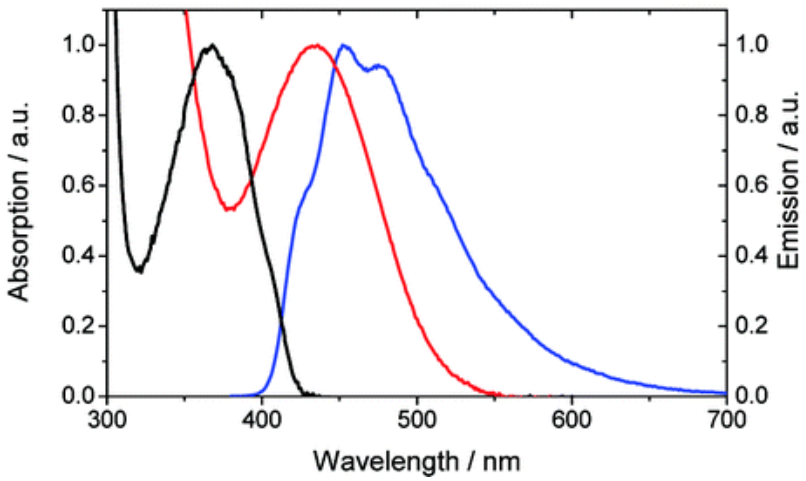


Figure 1.17 Representative normalized absorption (black) and emission (blue) spectra of FRET donor tC^{O} and absorption spectrum (red) of virtually non-fluorescent acceptor tC_{nitro} within dsDNA showing the donor/acceptor spectral overlap. Adapted with permission from [59]. Copyright 2009 American Chemical Society.

1.5.2 Biocompatibility of tC and its Derivatives

The tricyclic cytosine analogues do not negatively influence the structure or stability of dsDNA. The first evidence for this is the fact that they all increase the thermal stability of dsDNA.[51] A useful particularity of tC^{O} is that, with an intelligent design of the bases surrounding tC^{O}, it is possible to have little or no increase in the melting temperature.[61] Therefore tC^{O} can be used without influencing positively the duplex stability. Secondly, the tC derivatives maintains the natural B-form conformation;[59,61,64] the general appearance of the CD (circular dichroism) spectra is that of B-DNA irrespective of the derivative. This suggests they could be bio- or more specifically polymerase-compatible; an important criterion for assessing their ability to mimic the C base. From Figure 1.18 A, it is clear that tC and its derivatives can and should pair with G as they maintain the C moiety with its hydrogen bond pattern.

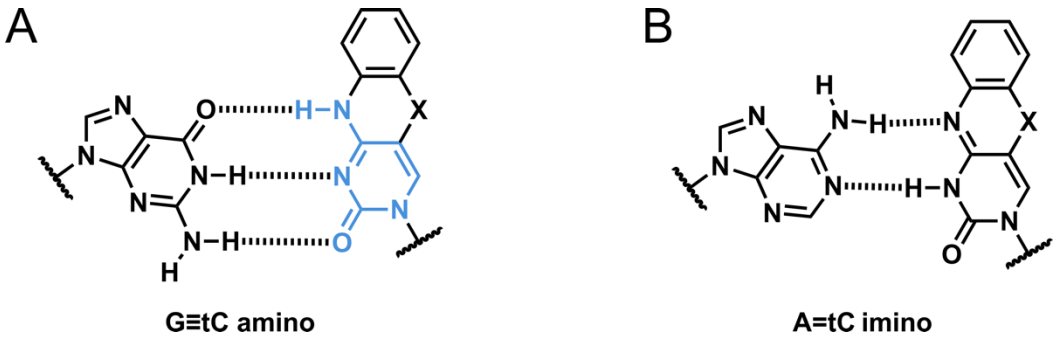
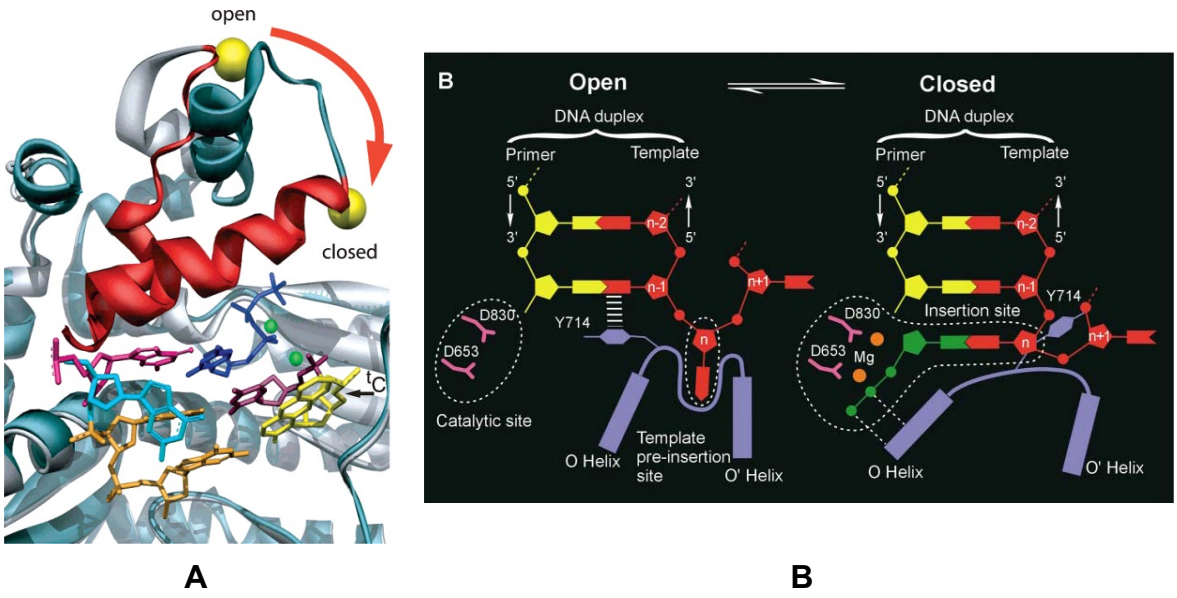


Figure 1.18 tC (and derivatives) base pairs. **A)** tC amino tautomer/G. The C moiety is in light blue. **B)** tC imino tautomer/A.

A study by Walsh *et al.* showed that the *Escherichia Coli* Y-family DNA polymerase DinB (Pol IV) is able to incorporate dtCTP opposite G selectively but cannot extend beyond the new G≡tC base pair.[68] Another report by Sandin *et al.* showed that the Klenow fragment (KF) incorporates dtCTP and dtC^{o}TP opposite G efficiently and that the analogues do not act as chain terminators.[69] Unfortunately, there is a loss of selectivity as the analogues are mis-incorporated opposite A.

A more comprehensive study by Stengel *et al.*, discussing the effects of tC and tC^{o} both in the template and primer strands with KF or human DNA polymerase α (pol α), gives more details regarding the problem of mismatches.[70] Inserting dtCTP or dtC^{o}TP opposite A is only ~4–11-fold less efficient than inserting dTTP opposite A for both KF and pol α. When tC or tC^{o} act as the template the situation is not much better. Inserting dGTP opposite the analogues is <10 fold less efficient than inserting it opposite C. Moreover, dATP addition opposite to the analogues is not significantly more inefficient than dATP addition opposite T. The presence of the tC=A mismatches could be explained by the formation of the imino tautomer of tC (Figure 1.18 B), which base pairs with A. The normal cytosine imino tautomer is essentially negligible compared to the amino tautomer, however Anand *et al.* showed that for the methoxy derivative of C mo^{4}C (N^{4}-methoxycytosine) the imino tautomer is very important.[71] The methoxy

electron withdrawing group makes the amino protons more acidic and therefore the tautomerisation is easier. It is likely that a similar effect facilitates the formation of the tC imino tautomer, given the presence of the electron delocalising aromatic ring. The central ring of tC and its derivatives is aromatic (the one that contains N-4). Therefore N-4 is sp^2 hybridised and bound to two electron withdrawing sp^2 carbons, thereby facilitating N-4 deprotonation and imino tautomer formation.[70]



A **B**

Figure 1.19 Pol I DNA polymerase active sites. **A)** KF active site. Adapted with permission from [72]. Copyright 2007 American Chemical Association. **B)** Schematic depiction of the functioning of a general pol I DNA polymerase. Adapted with permission from [73]. Copyright 2004 Elsevier.

Stengel *et al.* showed that T7 RNA polymerase (T7 RNAP) incorporates tC very efficiently opposite G and is also very selective, as it does not present the same decrease of specificity observed with KF and pol α.[74] They demonstrated that the ribonucleotide form of tC could replace natural C in an ~800 bases long RNA. A possible rational for this could be due to how these different polymerases function. Stengel *et al.* used tC and Alexa-555 as, respectively, donor and acceptor of a FRET pair to study the "open" and the "closed" KF conformations.[72] The Alexa dye was attached at the end of the O-helix of the finger subdomain of KF, while tC was at the 3′-end of

the primer strand. The active closed conformation was triggered only by the presence of the correct dNTP and it was detected by augmented FRET efficiency (Figure 1.19 A). The functioning of the active site in both "open" and "closed" conformations is well explained in the article by Temiakov *et al.* (Figure 1.19 B and Figure 1.20).[73] In the open conformation, the next templating nucleotide is hidden in a pocket between the O-helix and the O′-helix and then, with the arrival of the dNTP, the closed form is triggered. In this way the templating nucleotide is released and it is possible to form the base pair with the dNTP (Figure 1.19 B). So the selection of the dNTP happens mainly in the closed form when templating and incoming nucleotide are facing and can form hydrogen bonds.

The mechanism of the T7 RNAP is different because the templating nucleotide is not hidden in a pocket when the polymerase is in the open form (Figure 1.20). The selection of the right dNTP begins in the open conformation because the incoming nucleotide can immediately face the templating one. With the transition to the closed conformation, the Y-helix pushes the dNTP to the reactive position of the active site. With T7 RNAP the incoming dNTP faces the templating nucleotide from the very beginning, so the hydrogen bonds are formed immediately giving no time for tC to tautomerise. Conversely, with KF and pol α the incoming dNTP first enters in the active site and then faces the templating nucleotide to form hydrogen bonds. This gives enough time to form the imino tautomer of tC and derivatives.[74]

T7 RNAP is not the only selective enzyme, in fact Urban *et al.* showed that human DNA primase can readily add tCTP opposite G but not opposite A.[75]

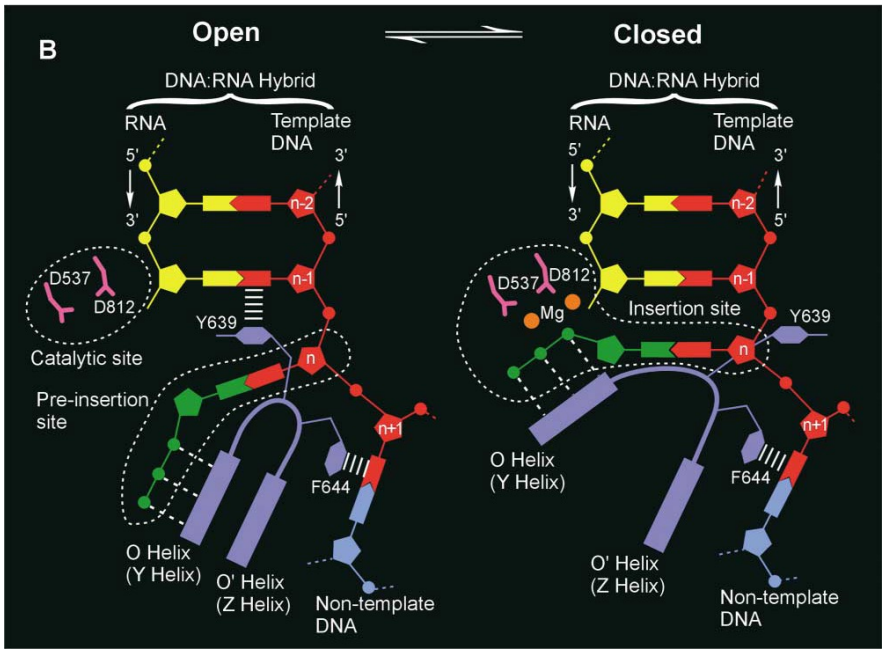


Figure 1.20 Schematic depiction of the functioning of T7 RNAP. Adapted with permission from [73]. Copyright 2004 Elsevier.

1.6 Pyrene as a Nucleotide-Functionalising Moiety

Pyrene is a widely used moiety to modify nucleotides,[76] this section gives a brief introduction to its properties and presents some of the modifications that have been developed, focusing on their effect on the stability of nucleic acid duplexes.

One of the most interesting features of pyrene is that its surface area is comparable to that of an A=T base pair (~184 Å^2 for pyrene and ~223 Å^2 for the A=T base pair).[76,77] Moreover, in the crystal structure of pyrene the interplanar distance is 3.54 Å,[78] interestingly close to the 3.4 Å between two adjacent bases in a B-form dsDNA.[11,12] These properties explain the tendency of pyrene to π-stack with the aromatic nucleobases and its capacity to intercalate inside B-type duplexes. The duplex has to locally unwind about 3.4 Å to make space for pyrene, increasing its thermal stability thanks to the extended π-stacking system.[79] Intercalation inside an A-type duplex is

less likely to occur, as the distance between the nucleobases is significantly reduced and the resulting duplex is compressed.[11,12,76]

When measured in deoxygenated dichloromethane, pyrene has absorption maxima at 275, 320 and 337 nm. Pyrene is also fluorescent with emission maxima at 372, 378, 383, 393 and 413 nm (although only the first and third peak are usually visible due to broadening of the signal) and a quantum yield of 0.41.[76,80] When pyrene is incorporated into nucleic acids its fluorescence can be quenched by neighbouring nucleobases, especially when intercalating, *via* photo-induced electron transfer. Usually adenine is the weakest quencher, thymine or even better cytosine are moderately strong quenchers (*via* reductive electron transfer to the nucleobase), and the strongest quenching comes from guanine *via* oxidative electron transfer from guanine itself.[76,81-83] Therefore, it is possible to follow the melting or the annealing of duplexes containing pyrene by monitoring the variations in the intensity of pyrene fluorescence emission.[84,85]

Pyrene can form excimers when an electronically excited pyrene unit forms a dimer by π-stacking with a ground state pyrene unit that is within 3–4 Å distance.[84,86-88] The resulting excimer emission peak is featureless, broad and centred around 490 nm.[88-90] It too can be used to follow the melting of duplexes containing pyrene.[84]

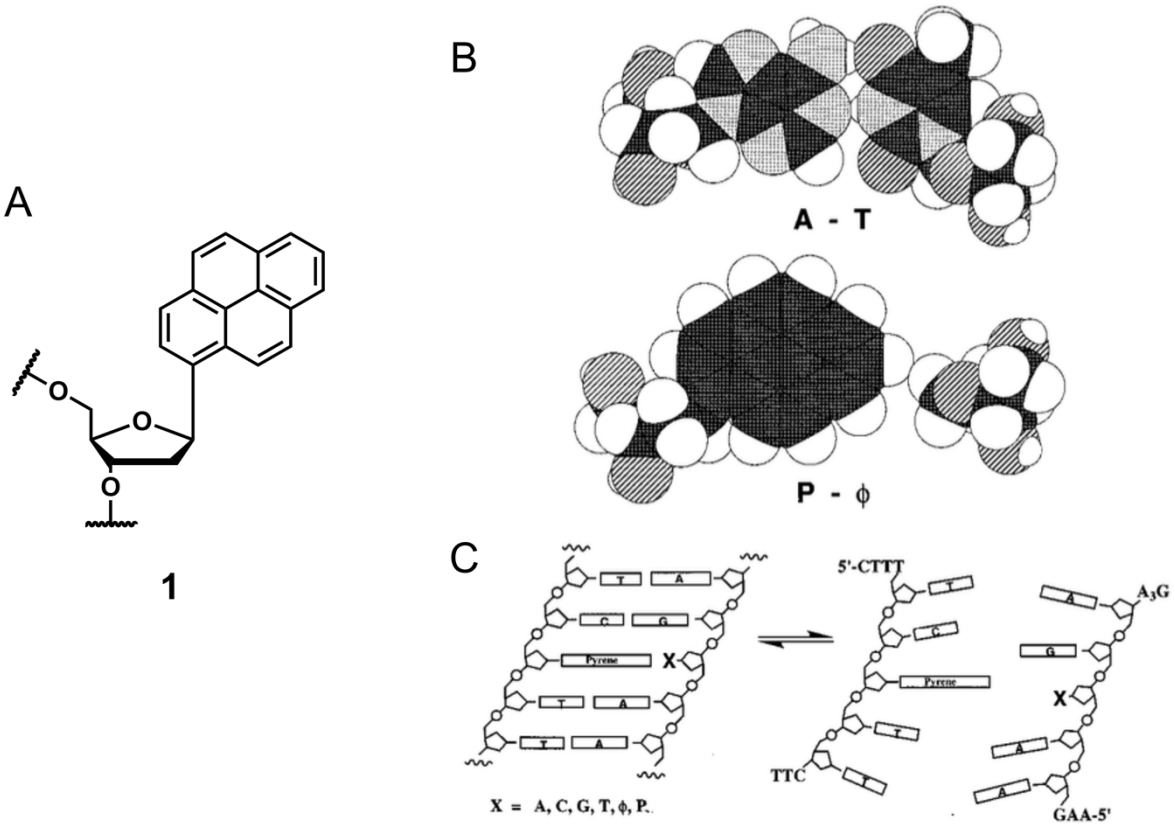


Figure 1.21 A) Structure of the nucleoside having pyrene as a base synthesised by Kool and collaborators.[91] **B)** Space filling models of the base pairs A-T and P-ϕ (pyrene-abasic site) in a B-DNA context. **C)** Schematic model of the melting of a duplex where the pyrene base is stacking inside the dsDNA). Adapted with permission from [92]. Copyright 1998 American Chemical Society.

Kool and co-workers synthesised nucleotide analogue **1** that has a pyrene in place of the normal nucleobase (Figure 1.21 A).[91] When adenine was paired with an abasic site instead of thymine there was a strong destabilisation of the resulting duplex (5.3 kcal/mol or −21.2 °C),[92] probably due to disruption of the continuous π-stacking system.[93] However, when instead monomer **1** was paired with the same abasic site, the dsDNA was only slightly less stable that the control duplex having an A=T base pair (by 0.6–0.7 kcal/mol or −1.6–2.2 °C).[92] The lower destabilisation with the pyrene nucleobase analogue probably derives from the fact that the surface area available for π-stacking with this modification was very close to the one covered by an A=T base pair (Figure 1.21 B).[92] When monomer **1** was paired with one of the natural

nucleotides (A, T, G, C, Figure 1.21 C), the resulting dsDNA was significantly destabilised compared to the duplex having **1** pairing with the abasic site (by 0.5–1.8 kcal/mol or −2.3–4.6 °C).[92] This showed the selectivity of this pyrene nucleobase analogue towards an abasic site compared to a natural nucleobase. Furthermore, NMR studies confirmed that the pyrene moiety intercalated opposite the abasic site of dsDNA and stacked with the neighbouring bases without influencing the overall B-DNA conformation (Figure 1.22 A).[94] Monomer **1** was also used as a "dangling" residue to study stacking interactions at the end of a duplex (Figure 1.22 B). Pyrene, not surprisingly, stacks on top of terminal base pairs better than the natural nucleobases do and, in fact, it increased the melting temperature by 23.1 °C compared to the duplex without dangling residues and by 12.5 °C compared to the duplex having A as a dangling residue.[95]

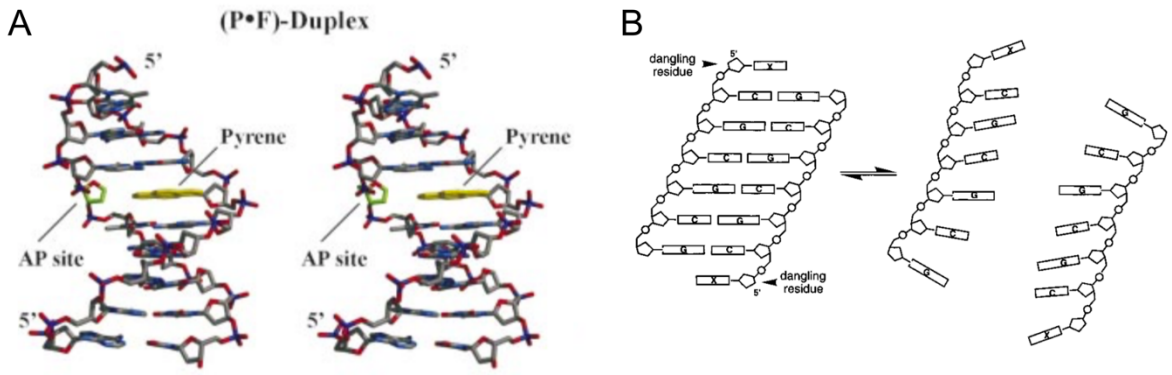


Figure 1.22 A) Cross-eye stereo view of the NMR structure of a dsDNA containing a pyrene-abasic site (P-F or P-AP) base pair. Adapted with permission from [94]. Copyright 2002 Oxford University Press. **B)** Schematic model of the melting of a duplex where the pyrene base is located at the end of the sequence as a dangling residue. Adapted with permission from [95]. Copyright 1996 American Chemical Society.

Monomers presenting a bulky modification like pyrene attached to C-5 of a pyrimidine *via* a short rigid linker tend to be destabilising. In fact, both nucleotide **2** and **3** (Figure 1.23) destabilised the DNA duplexes in which they were incorporated. A single addition of monomer **3** decreased the melting temperature between 6.5 and 9.0 °C,

while monomer **2**, with its slightly longer linker, destabilised the resulting duplexes of only 1.0–5.5 °C.[96]

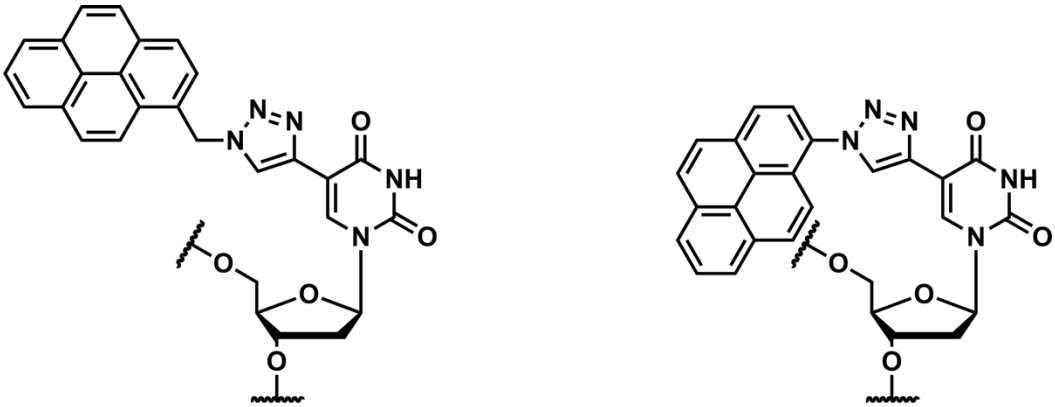


Figure 1.23 Pyrene-functionalised triazole-linked uracil analogues.

Also monomer **4** (Figure 1.24 A), which has another short rigid linker, destabilised dsDNA by 2.2–5.2 °C.[97] However, when it was used as a dangling end residue stacking at the bottom of a duplex it increased the melting temperature (+4 °C), as monomer **5** also did (+6 °C).[98] Both **4** and **5** were used to follow the B-to-Z-DNA transition by monitoring changes in their fluorescence emission (Figure 1.24 B). When the DNA was in the B-form, the pyrenes were stacking at the end of the duplex and were quenched by the neighbouring nucleobases, while when the DNA was in the Z-form pyrenes were unable to π-stack and their fluorescence was restored. The lack of stacking was also confirmed by the thermal stability. The duplex containing monomer **4** had a T_m of 70 °C as B-DNA and 60 °C as Z-DNA, while for monomer **5** the B-DNA T_m was 68 °C and the Z-DNA T_m was 61 °C. The unmodified control duplex had a B-DNA T_m of 64 °C and a Z-DNA T_m of 59 °C, highlighting how the B-DNA was significantly stabilised by the stacking of the pyrene moiety, while the Z-DNA was not affected to the same degree.[98]

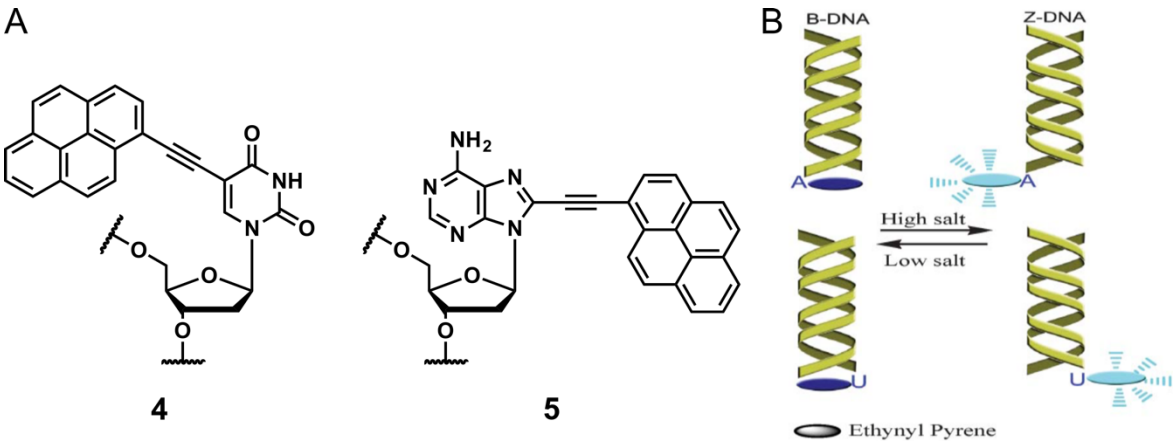


Figure 1.24 A) Nucleobases with an ethynyl pyrene unit. **B)** Schematic depiction of the probing system for the B-to-Z-DNA transition using monomers **4** and **5**. Adapted with permission from [98]. Copyright 2006 The Royal Society of Chemistry.

Modifications with longer and more flexible linkers connecting pyrene to the nucleobase tend to be more stabilising and, for example, modifications **6–9** (Figure 1.25) always stabilised the DNA duplexes into which they were incorporated.[85] Monomer **6** increased the melting temperature by 5.5–10.0 °C and monomer **7**, even with two bulky pyrenes, stabilised the duplex by 3.5–5.5 °C. Monomers **8** and **9** were less effective at increasing the melting temperature than their pyrimidine counterparts, with a ΔT_m = +5 °C and +1 °C respectively .

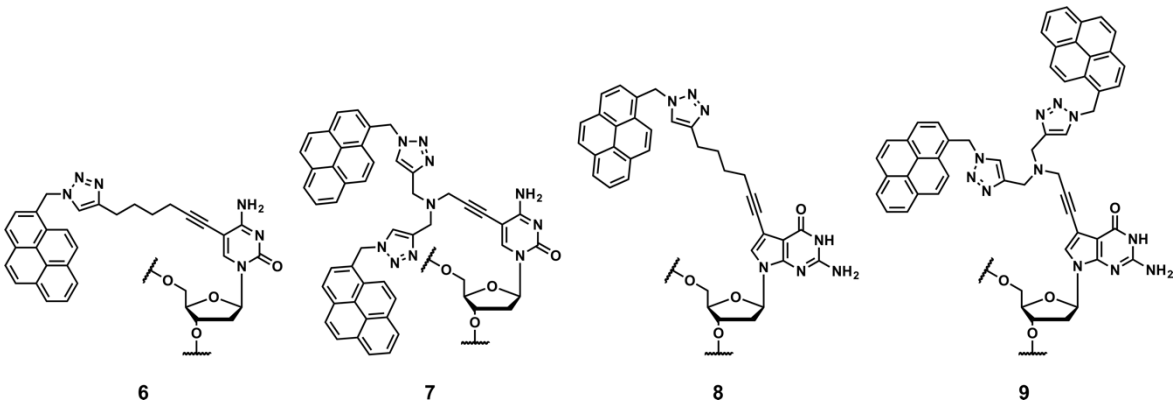
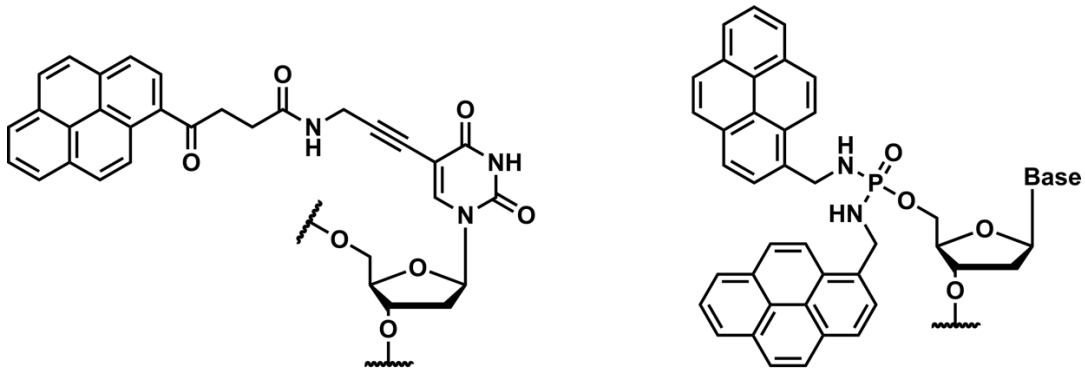


Figure 1.25 Nucleotides containing pyrene attached with a long linker synthesised by Seela and co-workers.[85]

Monomer **10** (Figure 1.26) also presents a pyrene with a long linker and when it was incorporated into dsDNA it significantly increased the melting temperature by 8.0–

11.1 °C. However, when two units of this modification were next to each other their stabilising effect was diminished, as the T_m increased by only 3.0 °C.[99]

The presence of two pyrenes on the 5′-phosphate of an oligonucleotide (**11**, Figure 1.26), stacking at the end of the resulting dsDNA, increased the thermal stability by 2–3 °C. Even more interestingly, the intramolecular structures of the single strands containing this modification were greatly influenced by the presence of the two pyrenes, with the ssDNA T_m increased by 23–24 °C.[100]



10 **11**

Figure 1.26 On the left a nucleotide containing a pyrene connected to the nucleobase *via* a long linker and on the right a double pyrene modification on the 5′-phosphate of an oligonucleotide.

Monomer **12** (Figure 1.27 A) is a nucleotide modified with pyrene on the 2′-position of the ribose sugar, which shows higher affinity for DNA than RNA.[101] When **12** was incorporated in a DNA strand the resulting DNA:DNA duplex was stabilised by 7–13 °C if a purine was flanking the modification on its 3′-side or by about 2 °C if there was a pyrimidine instead. When a DNA:RNA duplex was formed the modification had a destabilising effect as the melting temperature was reduced by 1–5 °C if a purine was flanking the modification on its 3′-side or by about 10 °C with a pyrimidine in the same position.[101] Incorporating monomer **12** in an RNA strand confirmed the preference for a DNA target over an RNA one. In fact while the RNA:DNA duplex was slightly stabilised

(from −0.2 to +2.9 °C), the RNA:RNA duplex was always destabilised (−3.2–4.1 °C).[102] NMR and molecular dynamics studies helped with understanding this different behaviour. In the case of dsRNA (Figure 1.27 B), while the overall A-form was maintained, the pyrene moiety was free from π-stacking interactions with the nucleobases, pointing outside of the duplex. In the case of dsDNA (Figure 1.27 C) the duplex maintained the B-form, with the pyrene moiety intercalating inside the double helix and π-stacking with the neighbouring nucleobases.[103] Naturally the latter imparts greater duplex stability.

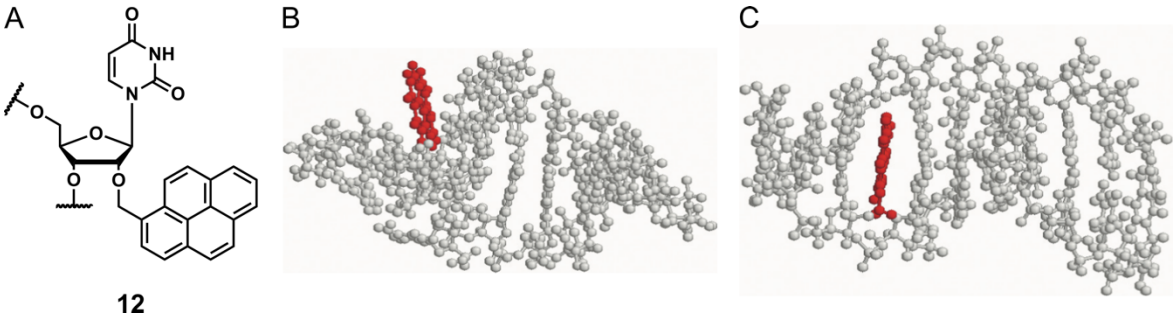


Figure 1.27 A) Structure of a 2′-pyrene-modified nucleotide. **B)** NMR and molecular dynamics structure of a dsRNA containing monomer **12** (in red). **C)** NMR and molecular dynamics structure of a dsDNA containing monomer **12** (in red). Adapted with permission from [103]. Copyright 2005 Oxford University Press.

1.7 Objectives and Thesis Overview

Thorough biophysical studies are necessary to gain a better understanding of the properties of modified oligonucleotides, with structural data being particularly important when trying to obtain mechanistic information on the functioning of these modifications. This work sets out to explore modifications containing pyrene using different techniques to obtain a wider picture of their properties, also by employing some structural studies. Investigating the effect these pyrene modifications have on oligonucleotides is fundamental to be able to better evaluate their potential

applications and improve the design of further more focused experiments. The first main modification, X-pyrene, was designed to study the cooperative effect of different elements that positively contribute to the thermal stability of duplexes. UV melting and fluorescence experiments allow to quickly evaluate the system under study, while NMR can contribute explaining the properties at a molecular level. A reductive amination labelling is explored as a quicker way to obtain X-pyrene and other simple pyrene modifications, with their capacity of stabilising dsDNA assessed by UV melting experiments. U-pyrene was instead chosen as a simple monomer to investigate the consequences that multiple additions of highly hydrophobic modifications like pyrene have on short oligonucleotides. Thermodynamic studies are essential to have a more complete view of the binding interactions of DNA strands containing this modification and X-ray crystallography allows to obtain the amount of molecular details necessary to understand the mechanism with which U-pyrene works.

What follows is an overview of the work presented in this thesis.

Chapter 2 presents the properties of X-pyrene, a novel nucleobase analogue designed to contain multiple duplex stabilising factors including electrostatic interactions, extra hydrogen bonds and extended π-stacking with the intercalating pyrene moiety. Fluorescence and NMR studies were used to confirm the intercalation of the aforementioned pyrene moiety.

Chapter 3 is focused on the use of a reductive amination labelling reaction to produce pyrene-functionalised nucleobase analogues and on their impact on the thermal stability of DNA duplexes. Both G-clamp and amino C6 dT were successfully functionalised with this methodology, obtaining both mono- and di-labelled products.

Finally, chapter 4 reports thermal stability and thermodynamic studies on the effect of multiple additions of U-pyrene, showing its great influence on increasing the binding affinity towards a complementary DNA target. X-ray crystallography was also employed to study the structural effects of this modification. The structures reported are to the author's knowledge the first X-ray crystal structures of dsDNA containing a pyrene-functionalised nucleotide analogue and they highlight U-pyrene capacity of facilitating the production of interesting nanostructures.

Chapter 2

X-pyrene

2.1 Preface

The majority of the work presented in this chapter features in a paper published in Chemical Science.[104] Dr Chenguang Lou initiated this project by synthesising the X-pyrene monomer and performing the initial biophysical studies with oligonucleotides that contain the modification. These oligonucleotides were synthesised and purified by ATDBio Ltd (with special thanks to Rachel Gao). Dr Andre Dallmann performed the NMR studies.

2.2 Introduction

The tricyclic cytosine analogues are a family of fluorescent base analogues with interesting properties. Unlike other widely used fluorescent base analogues such as 2-aminopurine (2-AP), their quantum yield remains almost unchanged upon incorporation into oligonucleotides.[51,61,62] This is a highly desirable property for many applications, such as Fluorescence/Förster Resonance Energy Transfer (FRET) measurements.[59] Furthermore, they do not negatively influence the structure and stability of the DNA duplex, and are readily incorporated by DNA and RNA polymerases. However, they are not perfect analogues of cytosine.[60,61,69,70,74] 1,3-diaza-2-oxophenothiazine (tC) and 1,3-diaza-2-oxophenoxazine (tC^O) possess a tautomeric form (Figure 1.18) that base pairs with A instead of G.[70] Since one of the potential applications of tC^O could be its use for *in cellula* fluorescence imaging, the transition mutations formed this way are potentially problematic because, if enough mutations are formed, they could trigger cell apoptosis.[70,105] A solution to the problem could be the use of a tC^O analogue named G-clamp (Figure 2.1 A). G-clamp (9-(2-aminoethoxy)phenoxazine) maintains the tC^O moiety and is therefore still

fluorescent.[106] The fundamental difference from the parent molecule is that G-clamp forms an additional hydrogen bond with the Hoogsteen face of the complementary guanine.[107] This confers the G-clamp a higher specificity towards G than tC^O and, in principle, this should reduce the ocurrence of mismatches during replication. Other interesting properties of oligonucleotides containing tC^O and G-clamp include their enhanced antisense activity and their increased nuclease resistance.[108,109]

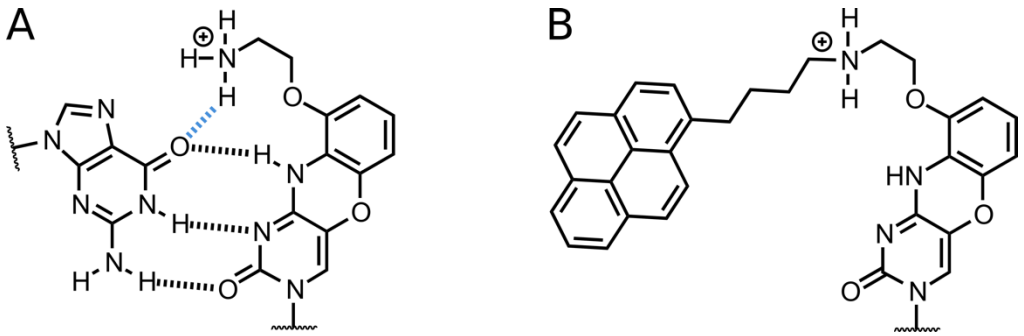


Figure 2.1 A) The tricyclic cytosine analogue G-clamp and its base pair with guanine. Hydrogen-bonds are shown as dashed lines, the Hoogsteen H-bond is in light blue. **B)** Structure of X-pyrene.

Large polycyclic aromatic molecules like pyrene have been used to increase dsDNA stability by intercalating and improving the π-stacking system.[76,110,111] Pyrene has also been used to promote site-specific targeting of enzymes like uracil DNA glycosylases and DNA methyltransferases.[112,113]

In this chapter, the biophysical properties of X-pyrene, a novel nucleobase analogue, are reported. X-pyrene (Figure 2.1 B) was designed as a duplex stabilising nucleoside combining the properties of G-clamp and pyrene. From the former it inherits the extended aromatic nucleobase, which should increase duplex stability, and the additional hydrogen bond to the Hoogsteen face of the complementary guanine, which should increase both duplex stability and target discrimination. From the latter it gains a large hydrophobic moiety, which should intercalate between base pairs. Therefore,

in principle, X-pyrene has the potential to form very stable duplexes with DNA and RNA, while all of these properties could also render X-pyrene sensitive to mismatches.

2.3 Materials and Methods

2.3.1 Oligonucleotides Synthesis, Purification and Analysis

2.3.1.1 Synthesis

Oligonucleotide synthesis was carried out on an Applied Biosystems 394 automated DNA/RNA synthesiser using a standard 0.2 µmol or 1.0 µmol phosphoramidite cycle of acid-catalysed detritylation, coupling, capping and iodine oxidation. All X-pyrene and G-clamp phosphoramidite monomers were dissolved in anhydrous DCM or acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 35 s and this was extended to 6 min for all modified monomers. Stepwise coupling efficiencies and overall yields were determined by automated trityl cation conductivity monitoring and in all cases were >98.0 %. The oligonucleotides attached to the synthesis columns were treated with 20 % diethylamine in acetonitrile for 20 min then washed with acetonitrile (5 × 1 mL). This procedure removes cyanoethyl groups from the phosphotriesters and scavenges the resultant acrylonitrile, preventing cyanoethyl adducts being formed at the secondary amine of X-pyrene. After this, oligonucleotides were deprotected and cleaved from the solid support using concentrated aqueous ammonia for 5 h at 55 ˚C.

2.3.1.2 Purification

Purification of oligonucleotides was carried out by reversed-phase HPLC on a Gilson system using a Brownlee Aquapore column (C8, 8 mm x 250 mm, 300 Å pore) with a

gradient of CH_3CN in NH_4OAc increasing from 0 % (buffer A) to 50 % (buffer B) over 20 min with a flow rate of 4 mL/min (buffer A: 0.1 M NH_4OAc, pH 7.0; buffer B: 0.1 M NH_4OAc with 50 % CH_3CN, pH 7.0). After HPLC purification, oligonucleotides were desalted using NAP-10 Sephadex G-25 columns (GE Healthcare) according to the manufacturer's instructions.

2.3.1.3 Analysis

Mass spectra of oligonucleotides were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES− mode. All the oligonucleotides were found with an experimental mass corresponding to the expected mass (±2 Da, eventual exceptions are reported in the table below). Table 2.1 presents the MS results for some representative sequences.

Table 2.1 Representative oligonucleotide sequences and MS analysis

Code	Oligonucleotide sequence (5′–3′)	MS (Da)	
		Calc.	Found
ODN-1	GCACYCTACG	3378	3378
ODN-4	CGACAYCACAGC	3990	3990
ODN-7	GCACYYATCG	3784	3783
ODN-9	GCYATACYAG	3808	3808
ODN-11	GCAAYATACG	3426	3426
ODN-12	GCACZCTACG	3517	3517
ODN-13	GCACWCTACG	3392	3390
ODN-14	CGACAWCACAGC	4004	4001

Note: Mass spectra of modified oligonucleotides were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES− mode. Y = X-pyrene, Z = 6-azidohexanoic acid NHS ester labelled G-clamp, W = 1-pyrenebutyric acid NHS ester labelled G-clamp.

2.3.2 Oligonucleotides Labelling with NHS Ester Labels

2.3.2.1 6-Azidohexanoic Acid NHS Ester Labelling of ODN-2

6-Azidohexanoic acid NHS ester (1 mg, 4 µmol) dissolved in DMF (80 µL) was added to a solution of the G-clamp modified oligonucleotide ODN-2 (0.1 µmol) in a 0.5 M Na_2CO_3/$NaHCO_3$ pH 8.75 buffer (80 µL). The reaction mixture was incubated at 55 ˚C for 4 hours.[114] The fully labelled oligonucleotides were desalted on NAP-25 Sephadex columns (GE Healthcare) and purified by reverse-phase HPLC (ODN-12, MS: calculated 3517 Da, found 3517 Da).

2.3.2.2 1-Pyrenebutyric Acid NHS Ester Labelling of ODN-2 and ODN-5

The G-clamp modified oligonucleotide (0.065 µmol) was freeze-dried and then dissolved in 40 µL of buffer (0.5 M Na_2CO_3/$NaHCO_3$ pH = 8.75). 1-Pyrenebutyric acid NHS ester (0.5 mg, 1.3 µmol, 20 eq.) was dissolved in DMSO (40 µL). After mixing the label and the oligonucleotide solution the reaction mixture was incubated at 55 °C for 4.5 hours. The excess pyrene was removed by NAP-25 gel-filtration (GE Healthcare). The labelled oligonucleotides were purified by reversed-phase HPLC (ODN-13, ODN-14).

2.3.3 UV Melting

UV melting samples were prepared in 5 mM sodium phosphate buffer with 140 mM KCl, 1 mM $MgCl_2$ at pH 7.2. Prior to use, the buffer was filtered with Kinesis regenerated cellulose 13 mm, 0.45 µm syringe filters and degassed by sonication for 10 min. Concentration of both target and probe (labelled) oligonucleotides was 2.5 µM. Melting analysis was performed on a Cary 4000 UV-Vis spectrometer fitted with a Cary temperature controller. Samples were analysed in 1.2 mL cuvette (Hellma Precision

cells made of synthetic quartz glass, QG, 1 cm path length). Before melting curves were recorded, the samples were melted and then slowly annealed. UV melting was carried out from 18 °C to 84 °C with ramping of 1 °C/min. UV absorbance at 260 nm was recorded every 0.1 °C. The melting temperatures (T_m) were calculated using Cary Win UV thermal application software, the values reported are an average of the three or four melting curves. The following table summarises the procedure used:

Stage	Data interval (°C)	Rate (°C/min)	End temperature (°C)	Hold time (min)
Melting 0	1.00	10	84	2
Annealing 1	0.10	1	18	2
Melting 1	0.10	1	84	2
Annealing 2	0.10	1	18	2
Melting 2	0.10	1	84	2
Annealing 3	0.10	1	18	2
Melting 3	0.10	1	84	2
Annealing 4	0.10	1	18	2
Melting 4	0.10	1	84	2
Annealing 5	1.00	10	20	0

Note: Annealing 4 and melting 4 were carried out only for some of the experiments. Therefore, for some oligonucleotides three melting temperatures were measured and for others four.

2.3.4 Fluorescence Studies

The experiments were carried out using a BMG Labtech CLARIOstar plate reader with a Greiner 96 wells fluorescence microplate. In order to maximise the emission intensities, the excitation wavelength used were 344 nm for X-pyrene and 364 nm for G-clamp, while the detection was at 372–612 nm for X-pyrene and 390–612 nm for G-clamp. All samples were prepared with a concentration of 0.15 μM of probe strand (containing either X-pyrene or G-clamp) and 0.2 μM of complementary strand in a phosphate buffer (5 mM sodium phosphate, 1 mM $MgCl_2$, pH 7.2 containing 140 mM KCl). 240 μL of solution per well were used and each sample was divided in 4 wells. Before the analysis, the samples were denatured by heating until 95 °C and then they

were left to slowly cool down to room temperature overnight, in order to allow a slow annealing. All the spectra reported are an average of four spectra from which the average blank spectrum was subtracted and they are normalised to the ssDNA spectrum present in each graph.

Titrations reported in Figure 2.12 B were done by adding increasing quantities (from 1 μL to 4 μL in 0.5 μL steps) of the complementary oligonucleotide to ODN-1 containing a bulged G to a solution of the duplex BD1-3′G (with the following initial concentrations: 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). For the blank and the solution of ODN-1 alone, buffer was added each time to obtain the same volume as the BD1-3′G solution.

2.3.5 Circular Dichroism

CD spectra were measured on a Jasco J-815 spectropolarimeter at 2.5 μM duplex concentration in phosphate buffer (5 mM sodium phosphate, 1 mM $MgCl_2$, pH 7.2 containing 140 mM KCl).[115] Spectra were recorded at 50 nm min^{-1} (accumulation = 5) with a response time of 1 s, a bandwidth of 1 nm and a data pitch of 0.2 nm. A buffer baseline was subtracted from each spectrum.

2.3.6 Error Calculations

Melting temperature errors were calculated using the confidence intervals derivable with a Student's t-test with 95 % confidence and the appropriate degrees of freedom. For ΔT_m the errors were calculated using the normal error propagation formula without the covariance factor. In graphs errors are represented by error bars.

2.3.7 Molecular Visualisation Based on NMR Data

The models, based on a qualitative interpretation of the NMR data, were constructed for suggestive visualisation using Hyperchem 7 (Hypercube, Inc.) and Avogadro (an open-source molecular builder and visualisation tool, version 1.1.1 http://avogadro.openmolecules.net/). Geometrical optimisation *in vacuo* was performed in Hyperchem using the MM+ force field and the Polak-Ribiere algorithm, RMS gradient of 0.01 Kcal/(Å·mol). All the models were rendered with PyMOL (Molecular Graphics System, Version 1.5.0.3 Schrödinger, LLC).

2.4 UV Melting Studies

The biophysical studies were carried out using two types of duplexes (Table 2.2, Figure 2.3 A): normal fully-paired duplexes (ND) and bulged duplexes (BD), where an unpaired nucleobase opposite X-pyrene (or G-clamp or C) is present in the unmodified complementary strand (Figure 2.2 and Figure 2.3 A). This unpaired base allows the formation of a pocket where pyrene can intercalate. Initial studies were carried out placing the unpaired base both on the 3′- and 5′-side of X-pyrene. 5′-bulged duplexes containing X-pyrene were less stable than the corresponding 3′-bulged duplexes both in DNA/DNA and DNA/RNA contexts (Figure 2.2). For this reason, all the subsequent studies were focused on the 3′-bulged duplexes and the terminology "BD" refers to 3′-bulged duplexes except where stated otherwise.

In order to establish the importance of the positive charge of the exocyclic amine of X-pyrene, separate from the role of the pyrene moiety, a set of experiments involving labelling G-clamp oligonucleotides with pyrene and non-intercalative NHS esters was designed. When the secondary amine on the G-clamp exocyclic amine (corresponding

to the amine on X-pyrene linker chain) was functionalised with 6-azidohexanoic acid NHS ester after oligonucleotide synthesis, a large decrease in T_m compared to X-pyrene (−16.4 ˚C) was observed in the stability of the BD1 bulged duplex (Figure 2.4). This is probably because, aside from the major effect due to the loss of the intercalating pyrene unit, the conversion of the secondary amine to an amide removed the positive charge which might have made it more difficult for X-pyrene to form the extra Hoogsten hydrogen bond in the major groove with guanine (similarly to G-clamp, Figure 2.1 A), given the reduced freedom of movement deriving from a more rigid amide bond. However, it cannot be ruled out that this decrease in stability may also be partly caused by the steric hindrance from the 6-azidohexanoyl group. Nonetheless, it is interesting to note how this labelled BD1 duplex was also 2 ˚C less stable than its G-clamp counterpart (49.8 ˚C *vs* 51.8 ˚C, Figure 2.3 and Figure 2.4). To further understand the role of the exocyclic amine of X-pyrene in duplex stabilisation, the G-clamp oligonucleotides were labelled with 1-pyrenebutyric acid NHS ester. This allowed direct comparison between the amine (X-pyrene) and the amide (NHS ester labelling) versions of the linker, as the rest of the structure was unchanged. It is clear from the results presented in Figure 2.4 that, independently of the sequence (1 or 2) or the duplex context (BD or ND), the amide linker always destabilised the resultant duplex. This is consistent with what hypothesised previously about the absence of the positive charge and the possible negative effect of the more rigid amide bond, which could also render proper intercalation more difficult. Moreover, these results substantiate the design choices for X-pyrene.

		5′-intercalation		3′-intercalation	
DNA/DNA	CYC form	3'-CGTG A GGATGC-5' 5'-GCAC pyrene XCTACG-3'	T_m = 55.7 °C (0.4)	3'-CGTGG A GATGC-5' 5'-GCACX pyrene CTACG-3'	T_m = 61.7 °C (0.7)
		3'-CGTG C GGATGC-5' 5'-GCAC pyrene XCTACG-3'	T_m = 58.3 °C (2.5)	3'-CGTGG C GATGC-5' 5'-GCACX pyrene CTACG-3'	T_m = 66.2 °C (1.3)
	AYA form	3'-CGTT A GTATGC-5' 5'-GCAA pyrene XATACG-3'	T_m = 37.2 °C (1.4)	3'-CGTTG A TATGC-5' 5'-GCAAX pyrene ATACG-3'	T_m = 49.0 °C (0.8)
DNA/RNA	CYC form	3'-CGUG C GGAUGC-5' 5'-GCAC pyrene XCTACG-3'	T_m = 57.6 °C (0.4)	3'-CGUGG C GAUGC-5' 5'-GCACX pyrene CTACG-3'	T_m = 60.1 °C (0.5)

Figure 2.2 Top: Melting temperatures (T_m) of 5′- and 3′- X-pyrene bulged DNA/DNA duplexes in two different neighbouring environments (CYC and AYA, where Y = X-pyrene). ODN-1 (CYC form) and ODN-11 (AYA form) were used to construct X-pyrene bulged duplexes. Different flanking neighbourhoods are underlined. **Bottom:** Melting temperatures (T_m) of 5′- and 3′- X-pyrene bulged DNA/RNA hybrid duplexes. The errors are reported in brackets.

In the case of bulged duplexes, a single insertion of X-pyrene confers a far greater duplex stability than G-clamp or C (Figure 2.3 B). In the case of the BD1 duplex a significant increase in melting temperature was observed relative to C and G-clamp (ΔT_m = +32.2 ˚C and +14.4 ˚C respectively). The same trend was also observed for the longer BD2 duplex (ΔT_m = +15.3 ˚C and +8.9 ˚C respectively). It is interesting to note that previous studies on the sequence dependence of G-clamp duplex stabilisation revealed that CYC and AYA are the two extreme neighbouring environments, giving considerable variations in melting temperature (T_m(CYC) – T_m(AYA) = 19.0 ˚C).[116] When X-pyrene was incorporated into bulged duplexes, a lower degree of sequence dependence was observed, and the difference between CYC and AYA was reduced to 12.7 ˚C (Figure 2.5). This decrease in sequence dependence is an advantage when planning the design and use of X-pyrene in genetic probes.

Table 2.2 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′-GCACYCTACG	ND	5′-CGTAGGGTGC
	ODN-2	G-clamp			5′-CGUAGGGUGC (RNA)
				BD	5′-CGTAGCGGTGC
	ODN-3	C			5′-CGUAGCGGUGC (RNA)
2	ODN-4	X-pyrene	5′-CGACAYCACAGC	ND	5′-GCTGTGGTGTCG
	ODN-5	G-clamp		BD	5′-GCTGTGTGTGTCG
	ODN-6	C			
3	ODN-7	X-pyrene	5′-GCACYYATCG	BD	5′-CGATTGTGGTGC
	ODN-8	C		ND	5′-CGATGGGTGC
4	ODN-9	X-pyrene	5′-GCYATACYAG	BD	5′-CTTGGTATTGGC
	ODN-10	C		ND	5′-CTGGTATGGC

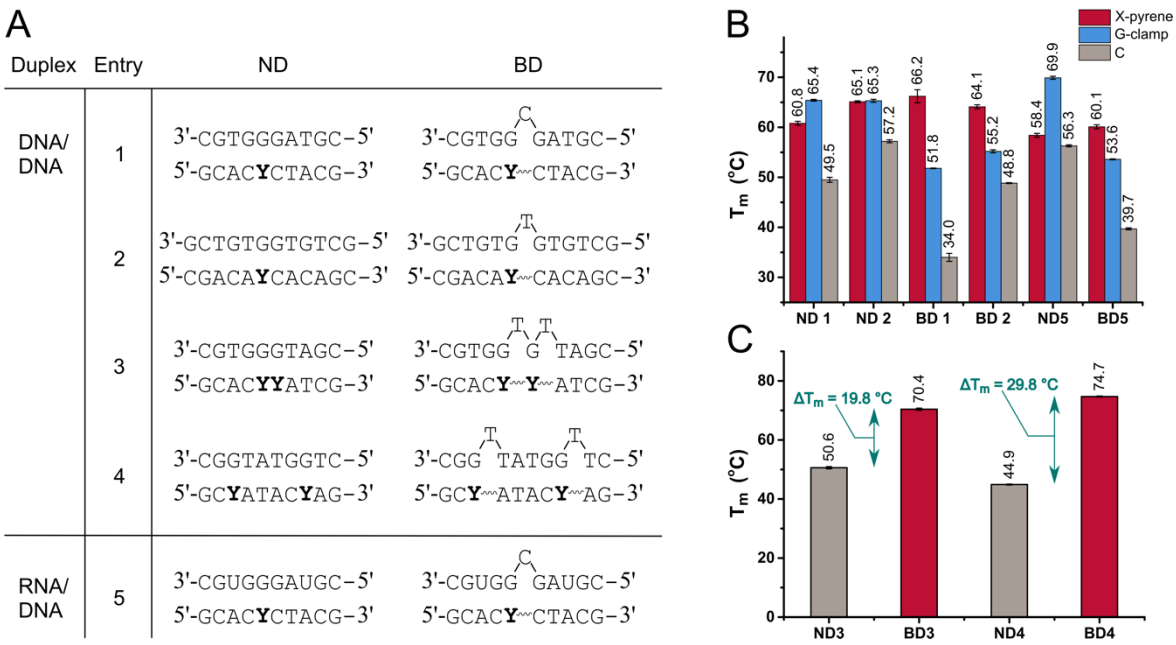


Figure 2.3 A) Duplexes used in the initial UV melting experiments. Each ND duplex has a corresponding BD duplex in the same row. Complementary RNA strands were used in Entry 5. ND = normal fully paired duplexes, BD = bulged duplexes. Y = X-pyrene, G-clamp or C. **B)** Melting temperatures of duplexes with a single X-pyrene added and of the corresponding duplexes with G-clamp or C. **C)** Melting temperatures of duplexes with a double X-pyrene insertion and of the corresponding duplexes with C. 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 sodium phosphate buffer (pH 7.2, containing 140 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2.5 μM).

	Y = X-pyrene	Y = G-clamp + label	$\Delta T_m = T_m$ (X-pyrene) – T_m (G-clamp + label)
BD1 3'-CGTGG A GATGC-5' 5'-GCAC**Y**~CTACG-3'	T_m = 66.2 °C (1.3)	Y = 6-azidohexamido G-clamp T_m = 49.8 °C (0.2)	ΔT_m = 16.4 °C (1.3)
		Y = 1-pyrenebutamido G-clamp T_m = 58.6 °C (0.6)	ΔT_m = 7.6 °C (1.4)
ND1 3'-CGTGGGATGC-5' 5'-GCAC**Y**CTACG-3'	T_m = 60.8 °C (0.4)	Y = 1-pyrenebutamido G-clamp T_m = 57.1 °C (0.4)	ΔT_m = 3.7 °C (0.6)
BD2 3'-GCTGTG T GTGTCG-5' 5'-CGACA**Y**~CACAGC-3'	T_m = 64.1 °C (0.4)	Y = 1-pyrenebutamido G-clamp T_m = 61.5 °C (0.2)	ΔT_m = 2.6 °C (0.5)
ND2 3'-GCTGTGGTGTCG-5' 5'-CGACA**Y**CACAGC-3'	T_m = 65.1 °C (0.2)	Y = 1-pyrenebutamido G-clamp T_m = 61.5 °C (0.2)	ΔT_m = 3.6 °C (0.2)

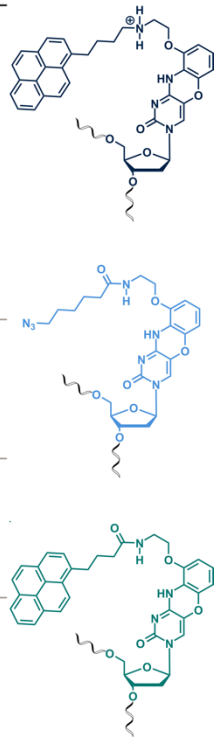


Figure 2.4 Effects on the thermal stability (T_m) of BD and ND duplexes after G-clamp labelling with 1-pyrenebutyric acid NHS ester (green) and 6-azidohexanoic acid NHS ester (light blue). The first column reports the T_m of the duplexes containing X-pyrene, the middle column presents the T_m of the duplexes containing the labelled G-clamp and the third column highlights the difference in thermal stability. The errors are reported in brackets.

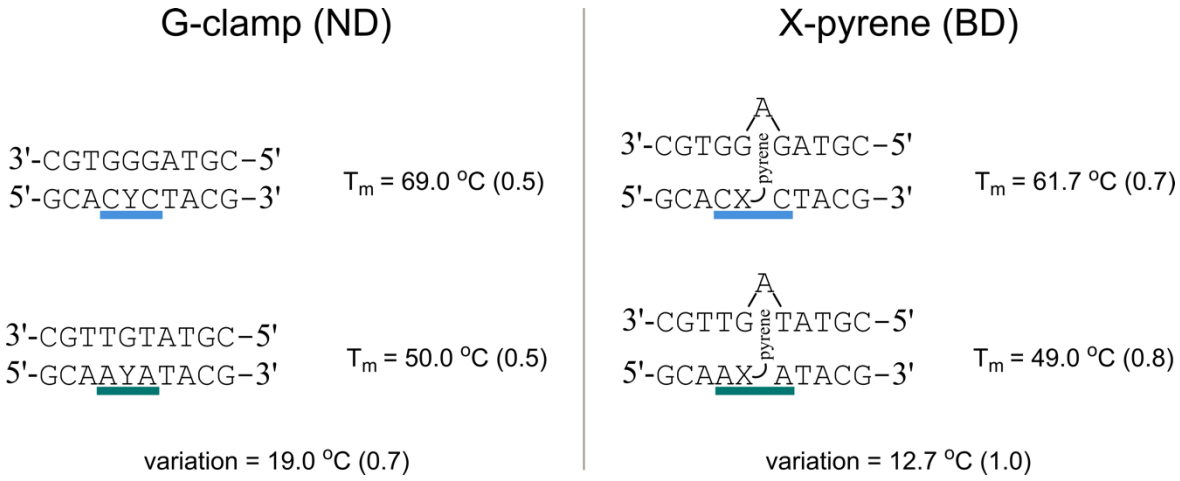


Figure 2.5 Comparison of the variation of thermal stability (T_m) of 3′- X-pyrene bulged duplexes (BD) to G-clamp normal standard duplexes (ND) in two extreme neighbouring environments (CYC and AYA, Y = X-pyrene or G-clamp). The T_m values of duplexes containing G-clamp are quoted from Ortega's work.[116] ODN-1 (CYC form) and ODN-11 (AYA form) were used to obtain X-pyrene bulged duplexes, the melting studies were performed under the same conditions as Ortega's (5 mM NaH_2PO_4/Na_2HPO_4 buffer at pH 7.2 with 1 mM $MgCl_2$ and 140 mM KCl, 2.5 μM of both probe and target strand.). The errors are reported in brackets.

The lack of the unpaired base in the fully complementary strand of normal duplexes (ND) prevents the formation of the pocket where the pyrene moiety of X-pyrene intercalates. Therefore, the stabilising effect of pyrene is lost and a negative effect (compared to the G-clamp) is even seen in the case of ND1, probably because the pyrene moiety has less space to stack in the duplex and it is so more exposed to the aqueous environment (Figure 2.3 B, ND1 and ND2).

Bulged DNA/RNA duplexes show the same trends of X-pyrene stabilisation as DNA/DNA duplexes. X-pyrene in BD5 stabilises the DNA more than G-clamp and its natural counterpart (ΔT_m = +20.4 ˚C and +6.5 ˚C respectively, Figure 2.3 B). Similarly, in the fully complementary normal duplex ND5, X-pyrene is destabilising compared to G-clamp. Moreover, 3′-intercalation is more stabilising than 5′-intercalation (Figure 2.2), as was also observed for DNA/DNA duplexes.

Next, the consequences of multiple insertions of X-pyrene were studied and the stabilising effect of X-pyrene was found to be cumulative; a double modification stabilised the duplex more than a single one. Unsurprisingly, when the modifications were dispersed (BD4, 4 bp away from each other) the stabilisation was stronger than when they were clustered (BD3, next to each other) (Figure 2.3 A and C). However, even the clustered X-pyrene units can strongly stabilise dsDNA.

Studies were also carried out to see if the nature of the unpaired base had any influence on the duplex stability (Figure 2.6). Unpaired pyrimidines (red) resulted in more stable duplexes than purines (light blue), which were more destabilising than an abasic site (green). However, the difference between the most (C) and the least (A) stable nucleobase was only 4.5 °C. Interestingly, a possible explanation for this different

behaviour of pyrimidines and purines was suggested by subsequent structural studies, which will be discussed in Section 2.5.

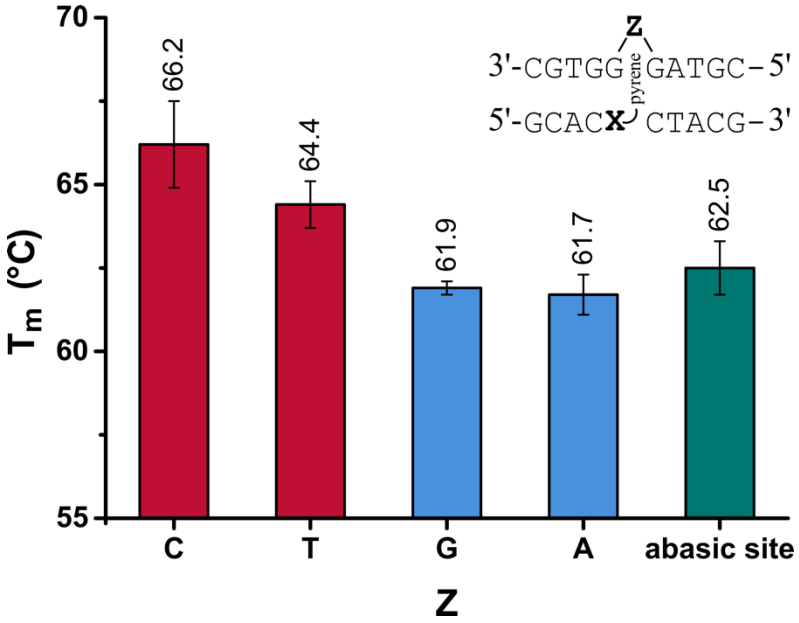


Figure 2.6 X-pyrene with different nucleobases in the unpaired position in bulged duplex (BD1). Z indicates A, G, C, T or an abasic site (tetrahydrofuran). Pyrimidines are in red, purines in light blue and the abasic site is in green. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 140 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2.5 μM).

The mismatch discrimination ability of X-pyrene was initially studied by varying the base opposite to it in BD (Figure 2.7). For G-clamp and C this is not the most favourable form for mismatch discrimination, as these two monomers prefer the ND context. Therefore, for a fairer comparison, matched and mismatched duplexes containing X-pyrene were in the BD1 form, while the duplexes containing G-clamp or C were in the ND1 form. Both X-pyrene and G-clamp were able to distinguish between the matched and mismatched base pairs with greater differences in T_m than C. Interestingly X-pyrene showed an almost uniform decrease in the melting temperature ($\Delta T_m \approx -35$ °C) independently of the nature of the mismatched nucleobase.

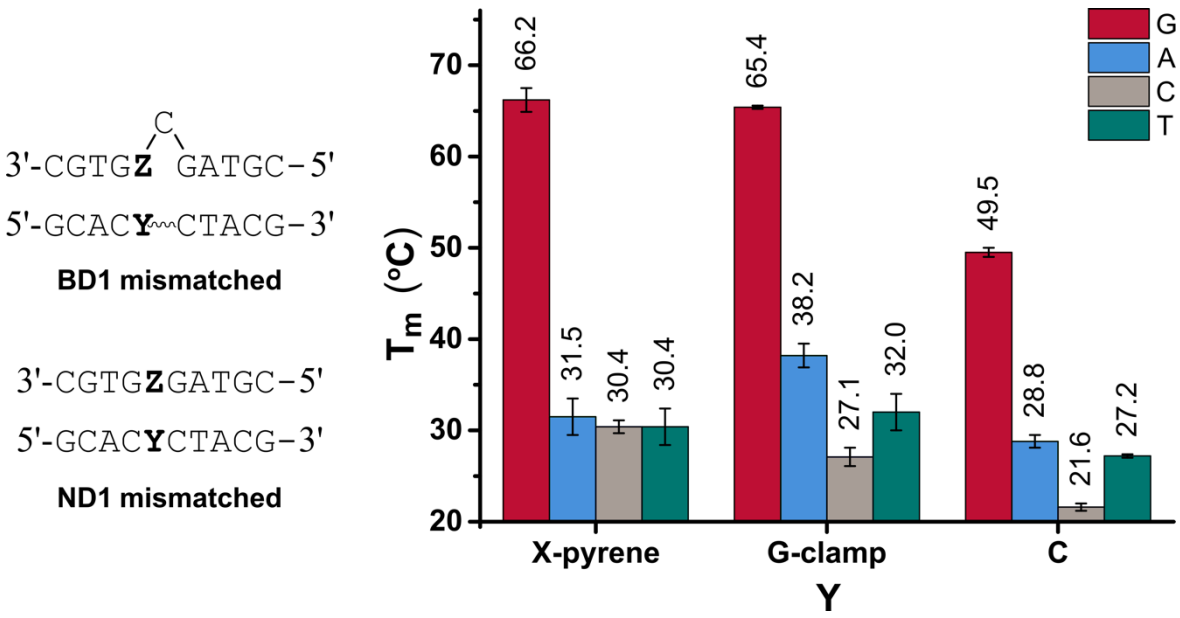


Figure 2.7 Comparison of matched and mismatched duplex stability. In ND1, Y = G-clamp or C, in BD1, Y = X-pyrene. The varied nucleobase Z and its partner are marked in bold. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 140 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2.5 μM).

As briefly mentioned in the introduction, the uniqueness of X-pyrene lies in the synergistic stabilising effect of three key structural features:

1. The protonated exocyclic secondary amine belonging to the G-clamp moiety, which forms an additional major groove hydrogen bond with O-6 of guanine on its Hoogsteen side and, since it is positively charged, it can also electrostatically interact with the negatively charged backbone;
2. The substituted phenoxazine ring (tC^O moiety), which provides an extended surface to enhance base stacking;
3. The pyrene moiety, which acts as a DNA intercalator providing so strong π-stacking with adjacent base pairs.

Due to these factors, bulged duplexes containing X-pyrene should be particularly sensitive to mismatches near this modification, even when it is not directly base pairing with the mutated base. In order to support this hypothesis, further studies

were performed on duplexes presenting mismatched base pairs around the X-pyrene:G base pair. The remote mismatch sensing ability of X-pyrene was studied using the BD1 duplex, while for G-clamp and C the normal duplex ND1 was used instead (Figure 2.8).

Neither X-pyrene (BD1) nor G-clamp (ND1) were very sensitive to mismatches very far from the modification (positions 1, 5 and 6 in Figure 2.8). However, both monomers showed enhanced sensing capacities for the base pairs closer to the modification. Particularly X-pyrene excelled at this task, showing a substantial decrease in melting temperature for both a TC mismatch next to the modification at position 3 (ΔT_m = −36.4 ˚C) and for an AA mismatch placed one base pair further away on the 5′-side of X-pyrene at position 2 (ΔT_m = −38.5 ˚C). The capacity to detect an AA mismatch at position 2, especially with such a considerable $\Delta T_{m,}$ was unique to X-pyrene. The G-clamp probe proved to be the only one able to improve the detection of a TC mismatch on the 3′-side of the modification at position 4 (ΔT_m = −26.3 ˚C), while detecting also the TC mismatch on its 5′-side at position 3 (ΔT_m = −29.2 ˚C), albeit not as well as X-pyrene. These results might indicate that the two Watson-Crick base pairs on the 5′-side of X-pyrene play a major role in the overall stability of this bulged duplex. The fact that even a mismatch two base pairs away on the 5′-side of X-pyrene influences the thermal stability of BD1 so strongly might be linked with a complex and dynamic structure. This tallies with the NMR studies, where a secondary conformation was detected.[104] 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.

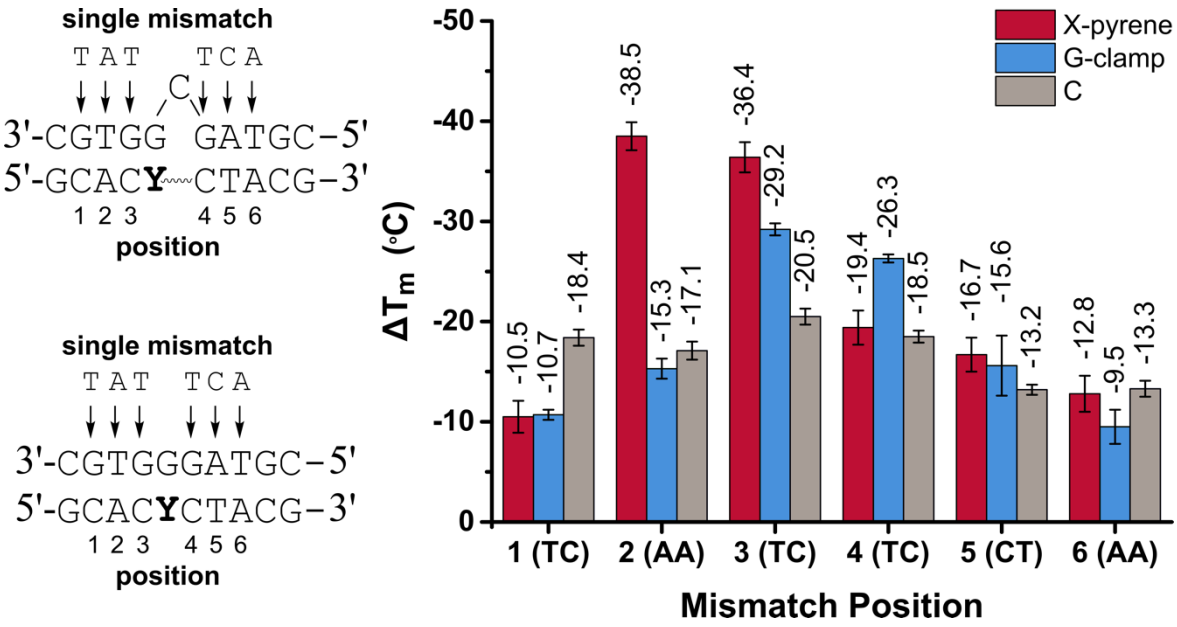


Figure 2.8 Efficient remote mismatch sensing with X-pyrene and comparison with G-clamp and C. A single mismatch was introduced at positions 1–6 of BD1 (X-pyrene) and ND1 (G-clamp or C). Y = G-clamp or C or X-pyrene. The bar chart shows ΔT_m *vs* mismatch position where ΔT_m = T_m(mismatch) – T_m(match).

2.5 Structural Studies and Molecular Visualisation

In order to understand the effects of X-pyrene on dsDNA conformation, circular dichroism spectra were obtained for two bulged duplexes (Y = X-pyrene or G-clamp) and three normal ones (Y = X-pyrene, G-clamp or C – Figure 2.9). Incorporating X-pyrene did not significantly affect the DNA conformation in either ND1 or BD1, however there is an evident hypsochromic shift (about 15 nm) of the main negative peak. Nevertheless, in both cases the overall shape is consistent with a B-conformation, similar to that of the standard duplex (Y = C, ND1). The presence of G-clamp, in both a bulged and a normal duplex, resulted in a pseudo-positive Cotton effect in the 233–252 nm region (absent when Y = X-pyrene or C). This is very similar to what was previously observed on tC^O-containing oligonucleotides.[117] Overall, based on the CD spectra, the

introduction of a G-clamp does not appear to perturb the B-conformation in both bulged and normal duplexes.

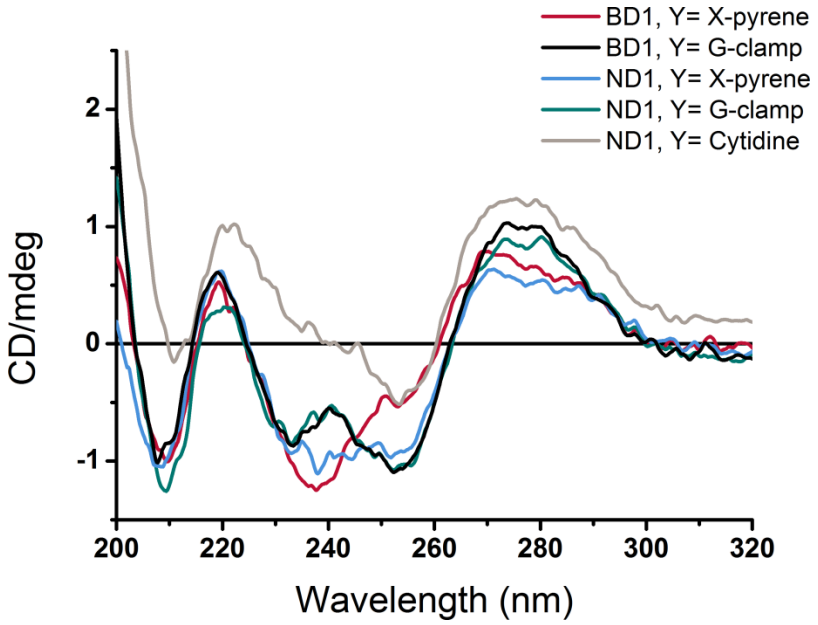


Figure 2.9 Figure adapted from Lou *et al.*[104] CD spectra of bulged and normal duplexes (Y = X-pyrene, G-clamp or C) recorded from 200 nm to 320 nm. The signals were smoothed using the Savitzky-Golay method (20 pts, polynomial order = 2) in Origin 8.6.

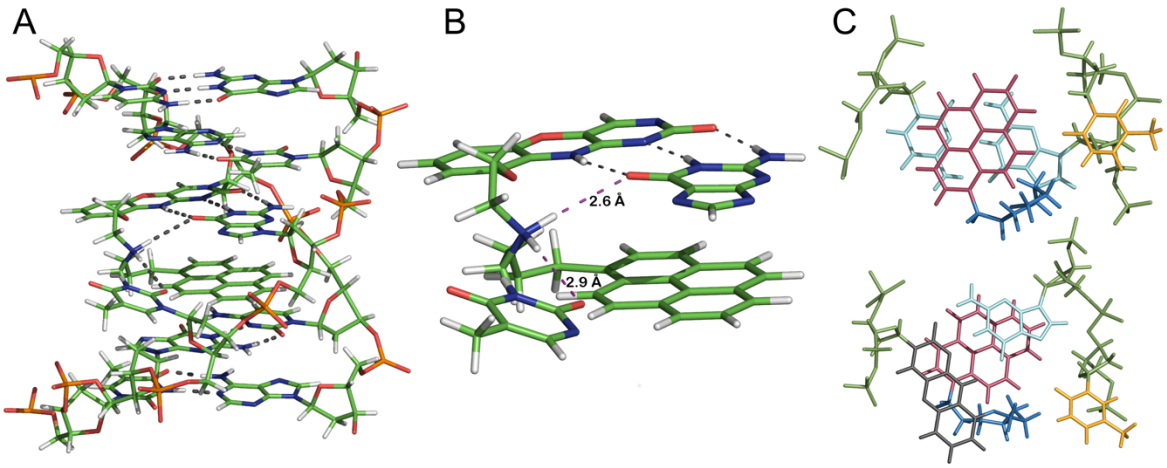


Figure 2.10 Molecular visualisation of the central part of BD2. **A)** Full duplex with hydrogen bonds highlighted; **B)** Close-up of X-pyrene region showing hydrogen bond pattern of the linker. Shown in purple are the H-bond from the linker to the Hoogsteen face of the G base paired with X-pyrene and the H-bond from the linker to O-2 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 tC^{O} moiety is in dark grey, the backbones are in green and the natural bases are in light blue.

Naturally, this global structural information is insufficient to understand why X-pyrene excelled at stabilising bulged duplexes. Therefore, qualitative NMR studies were carried out by Dr Andre Dallmann on the BD2 duplex and then molecular visualisation was carried out based on the information derived from NMR. In general, both NMR and CD agreed in stating that X-pyrene does not modify the B-form of DNA apart from some small local deviations. Intercalation and stacking of the pyrene were observed and, surprisingly, the unpaired T residue is not completely looped out, but is only pushed out towards the phosphate backbone (Figure 2.10 C). Although not stabilised by base stacking interactions, the unpaired base is stacked on the adjacent deoxyribose sugar and is capable of forming a hydrogen bond with the amino proton of the linker. While this H-bond is implied by the model, the flexibility of the linker unfortunately makes the H-bond too transient to be directly observed by NMR spectroscopy. This hypothetical hydrogen bond might also have a role in explaining the different behaviour of unpaired pyrimidines and purines seen in Figure 2.6. Such H-bond, highlighted in Figure 2.10 B, would prevent the unpaired nucleobase from being completely looped out of the duplex and, as mentioned before, it is possible to see the base stacking on the backbone (Figure 2.10 C). When the unpaired pyrimidine is replaced with a purine, the possibility of forming this extra hydrogen bond is lost, therefore explaining the reduced stability of a duplex containing a bulged purine compared to one containing a bulged pyrimidine.

2.6 Fluorescence Studies

In addition to being a useful indicator of structural changes, fluorescence can be used in biochemical assays in the DNA 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,[64,118] which leads to quenching by surrounding base pairs.[119] Both X-pyrene and G-clamp contain a tC^{O} (1,3-diaza-2-oxophenoxazine) moiety, which is fluorescent. tC^{O} has a high quantum yield and extinction coefficient and it is therefore a bright fluorophore.[51,61] For these reasons, fluorescence studies were carried out concentrating on the emission properties of the tC^{O} moiety.

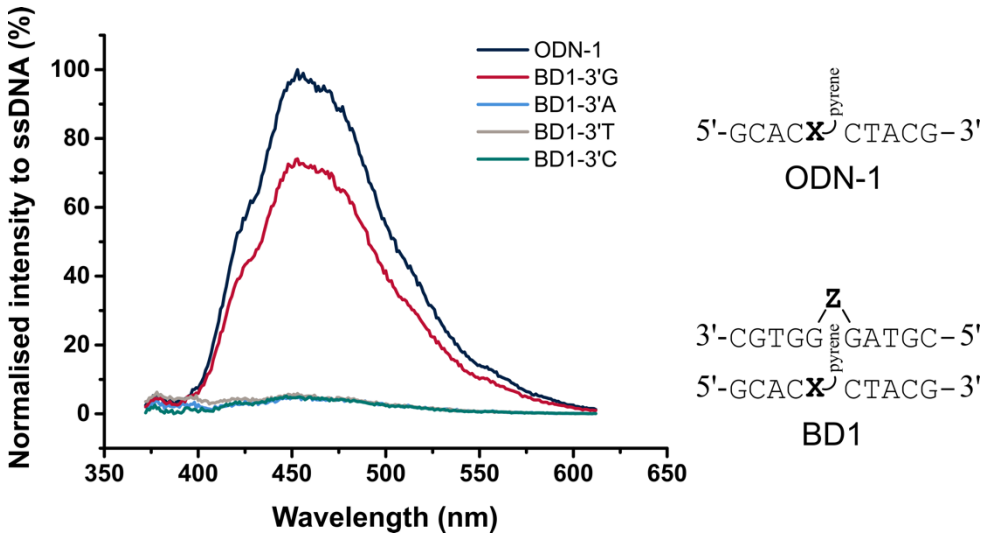


Figure 2.11 Fluorescence measurements on 10-mer X-pyrene modified oligonucleotide (ODN-1, black) and its bulged duplexes with complementary strands which have different nucleobases (Z) in the unpaired position (BD1-G, BD1-A, BD1-C, BD1-T). λ_{ex} = 344 nm. Concentrations: 0.15 μM of probe strand containing 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 emission, bulged duplexes with one strand modified with X-pyrene were used (ODN-1, BD1). A large decrease in fluorescence (about 95 %) was observed for bulged A, C and T, but not for bulged G, where the reduction was only about 25 % (Figure 2.11). The different behaviour of the bulged guanine oligonucleotide is probably caused by the presence of a G-rich core that leads to tetramolecular G-quadruplex formation.[120] Since G-quadruplexes are stable structures,[120,121] it is possible that their formation is

preferred over the formation of a normal duplex. Therefore, most of the X-pyrene oligonucleotide will remain in single stranded form and will not be quenched. In agreement with this hypothesis, the fluorescence diminished when more of the target strand was added, and almost 11 equivalents were necessary to produce an emission intensity reduction of more than 70 % (Figure 2.12).

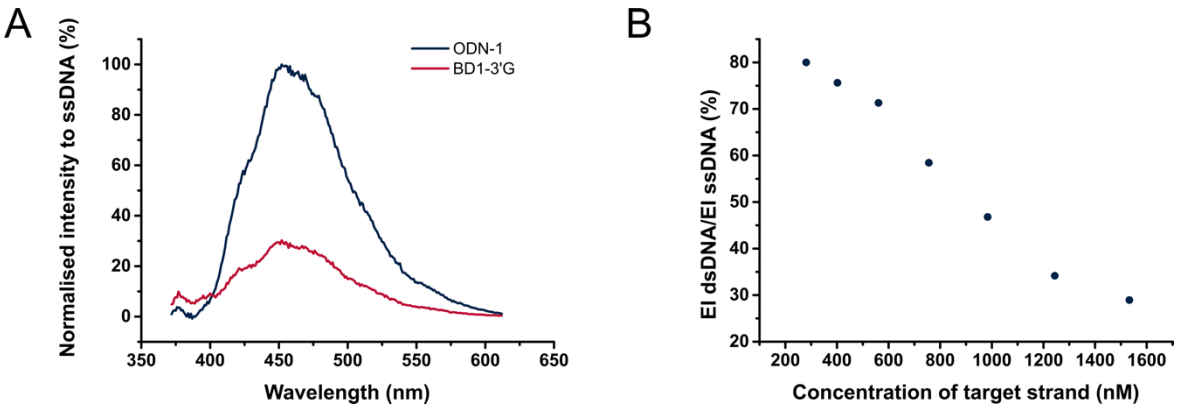


Figure 2.12 Fluorescence of X-pyrene and G-rich complementary strand. **A)** Fluorescence emission spectra of an X-pyrene containing oligonucleotide (ODN-1) and its duplex with the complementary strand presenting a G in the unpaired position (BD1-3′G). The target strand is in a large excess compared to the probe strand (almost 11 equivalents). **B)** Titration curve of ODN-1 with the target strand containing a G in the unpaired position. The graph plots the ratio between the emission intensity of the BD1-3′G duplex and the emission intensity of the ODN-1 alone measured at 454 nm versus the concentration of the target strand.

To confirm the quadruplex hypothesis, complementary oligonucleotides with 2′-deoxyinosine replacing one of the four consecutive deoxyguanosines of the G-rich core were used to prevent G-quadruplex formation (Figure 2.14). Inosine does not have the exocyclic amine present in guanine, which is important for the H-bond pattern that gives rise to G-quadruplexes (Figure 2.13). The addition of deoxyinosine led to normal duplex formation and strong fluorescence quenching with less than 1.5 equivalents of the complementary strand. The fluorescence reduction was almost 60 % when the deoxyinosine was at the 3′-end of the G4 repeat and almost 90 % when the deoxyinosine was in the middle, as in this position deoxyinosine disrupts the

formation of G-quadruplexes more (Figure 2.14). These results further confirm the role of G-quadruplexes in preventing the duplex formation and so the fluorescence quenching of X-pyrene. The above fluorescence studies are in general agreement with the NMR structural analysis, suggesting that the pyrene moiety, the tC^{O} moiety and the adjacent base pair overlap in the core of the duplex.

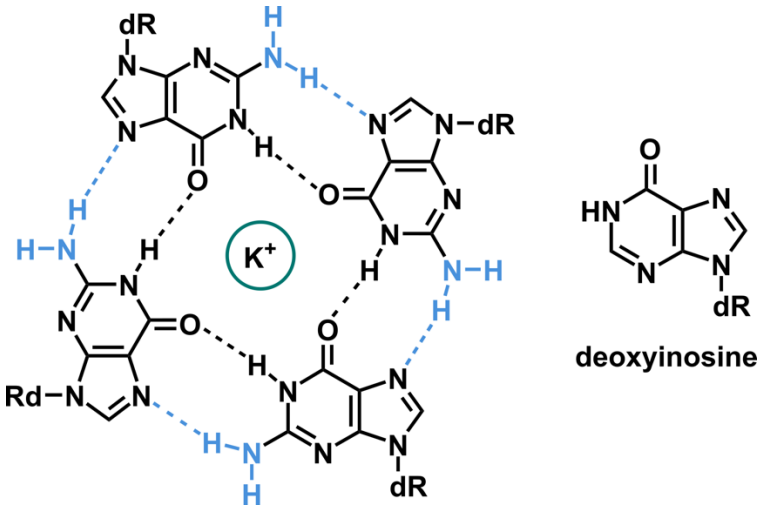


Figure 2.13 Structure of a G-tetrad, the building block of a G-quadruplex. The substitution of a deoxyguanosine with a deoxyinosine, which lacks of the amine in light blue, would prevent the formation of one of the two hydrogen bonds (in light blue) essential for a G-tetrad.

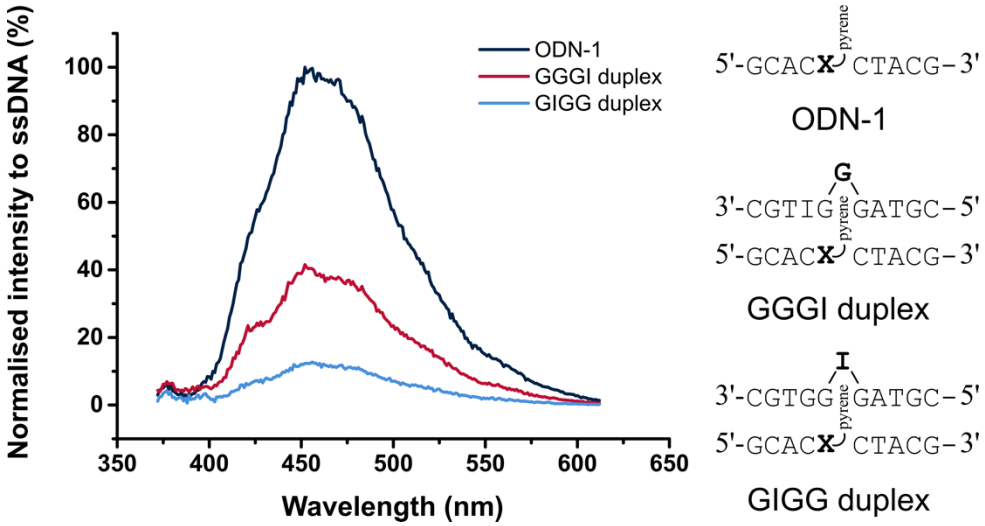


Figure 2.14 Fluorescence of X-pyrene and deoxyinosine containing complementary strands. Fluorescence emission spectra of an X-pyrene containing oligonucleotide (ODN-1) and its duplexes with the complementary strands presenting a deoxyinosine (I) at the 3′-end (GGGI duplex) and in the middle (GIGG duplex) of the G-rich core. The target strand is in the same concentration as the experiments in Figure 2.11. (200 nM).

Further investigations were carried out using ODN-1/ODN-2 and ODN-4/ODN-5 (Table 2.2, Figure 2.15 A, Figure 2.17 A) to determine how the fluorescence emission varies in different duplex contexts. The absolute fluorescence intensity clearly depends on the sequence; single stranded sequence 1 (ODN-1, ODN-2) is less fluorescent than sequence 2 (ODN-4, ODN-5). This could be due to the presence of two Cs as neighbouring bases to the fluorescent nucleotide compared to one C and one A (Figure 2.15, Figure 2.17). Such behaviour would be in agreement with what expected for a photo-induced electron transfer (PET) quenching, since purines have lower oxidation potentials than pyrimidines.[122-124] Furthermore, it is important to point out that a similar trend was also observed for tC^{O}.[61] Interestingly, the relative quenching in the different duplex contexts does not seem to depend on the sequence. When the fully complementary duplex is formed (ND1 and ND2, Figure 2.15), the fluorescence of X-pyrene is quenched by up to 60 %. This is not unexpected, as the quantum yield of tC^{O} is reduced in dsDNA because of electronic interactions with the other bases.[61] In the case of X-pyrene, this reduction in fluorescence is even more significant, probably because of additional electronic interactions between the tC^{O} and the pyrene moieties. The reduction of fluorescence upon formation of the fully complementary duplex when G-clamp is the fluorophore (ND1 and ND2, Figure 2.17) is less than for X-pyrene (the reduction is up to 40 %), and is more comparable to tC^{O}.[61]

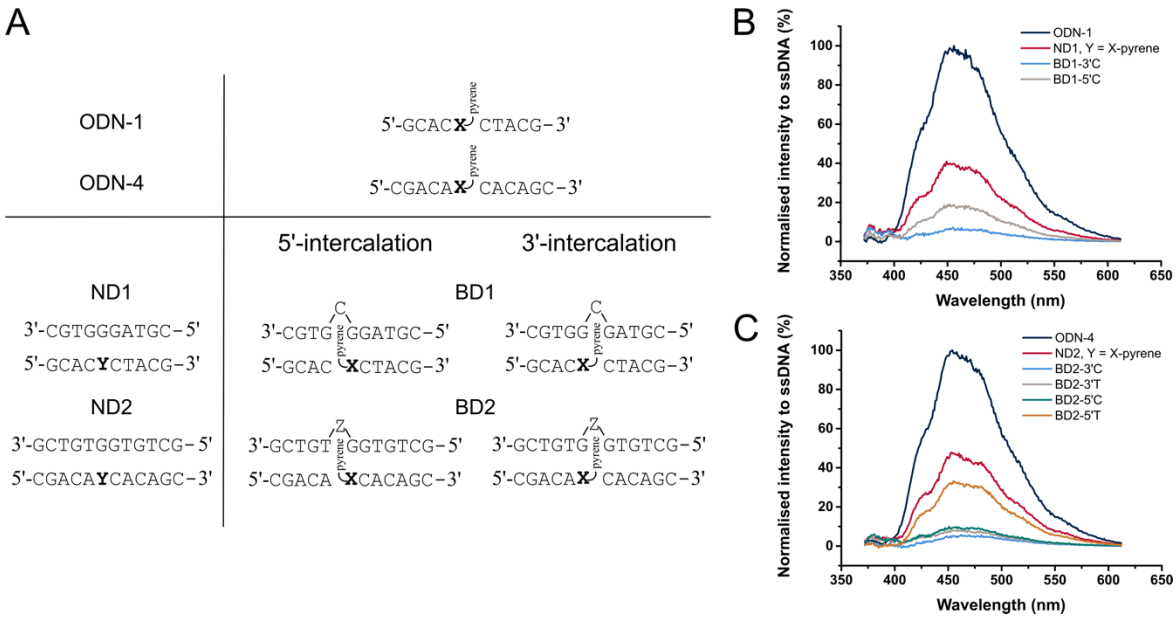


Figure 2.15 Fluorescence of X-pyrene containing oligonucleotides. **A)** Oligonucleotides and corresponding duplexes involved in the fluorescence study; **B)** Fluorescence emission spectra of a X-pyrene containing oligonucleotide (ODN-1) and its three duplexes (ND1, BD1-3′C, BD1-5′C); C) Fluorescence 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).

In almost all X-pyrene bulged duplexes the fluorescence emission was reduced on duplex formation, presumably because of increased surface contact between the pyrene and tC^O moieties. Interestingly, the emission was reduced less upon 5′-intercalation compared to 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 stabilised, the electronic interactions between pyrene and the tC^O moiety are weaker and consequently the fluorescence is less quenched. 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 2.15 B and C). 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 2.16 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, thereby eliminating quenching interactions.

This is not possible with ssDNA which is essentially unstructured. In support of this interpretation, for the same oligonucleotide sequence containing G-clamp instead of X-pyrene (BD6, Figure 2.16 B) the fluorescence is reduced upon duplex formation in a manner similar to all the other duplexes (Figure 2.16 B, Figure 2.17 C).

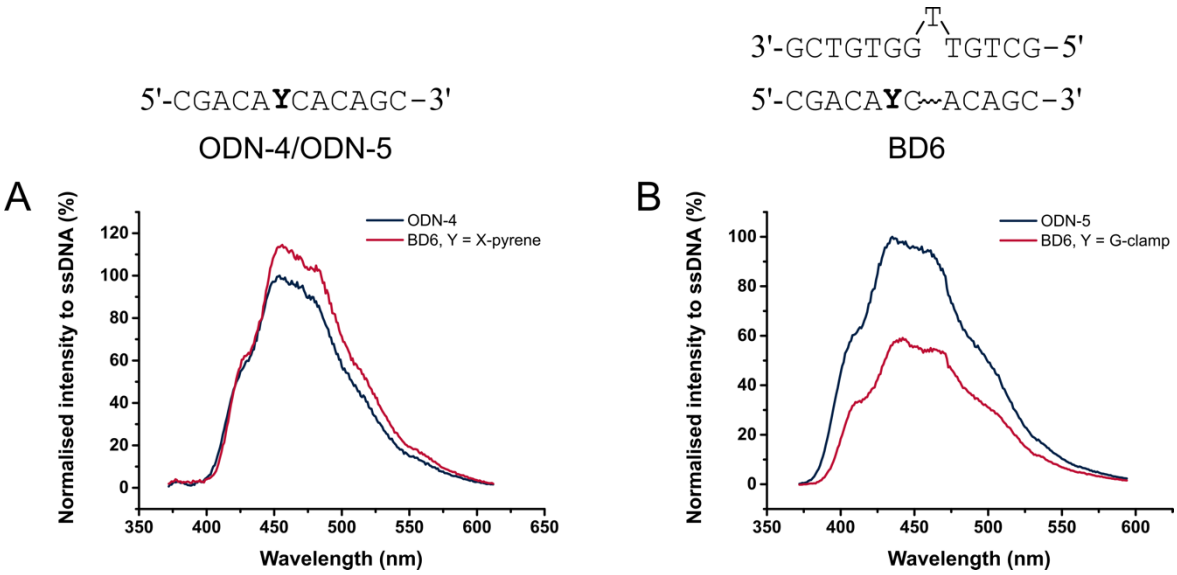


Figure 2.16 Fluorescence of an X-pyrene and a G-clamp oligonucleotide with the unpaired base 1 bp away. **A)** (Y = X-pyrene) Fluorescence emission spectra of a X-pyrene containing oligonucleotide (ODN-4) and its duplex with the bulged base 1 bp away from X-pyrene on the 3′-side (BD6). **B)** (Y = G-clamp) Fluorescence emission spectra of a G-clamp containing oligonucleotide (ODN-5) and its duplex with the unpaired base 1 bp away from G-clamp on the 3′-side (BD6).

Compared to X-pyrene, the fluorescence of G-clamp containing oligonucleotides was affected less by differences in DNA sequence (Figure 2.17). However, it is still possible to recognise a similar pattern for emission intensity as a function of the position of the bulge. When the bulge is in the 5′-position the G-clamp is more fluorescent than the 3′-case. This is probably due to the extra space in the major groove when the bulge is in the 5′-position, which reduces the electronic interactions between the G-clamp and the other bases, so the fluorescence intensity is higher.

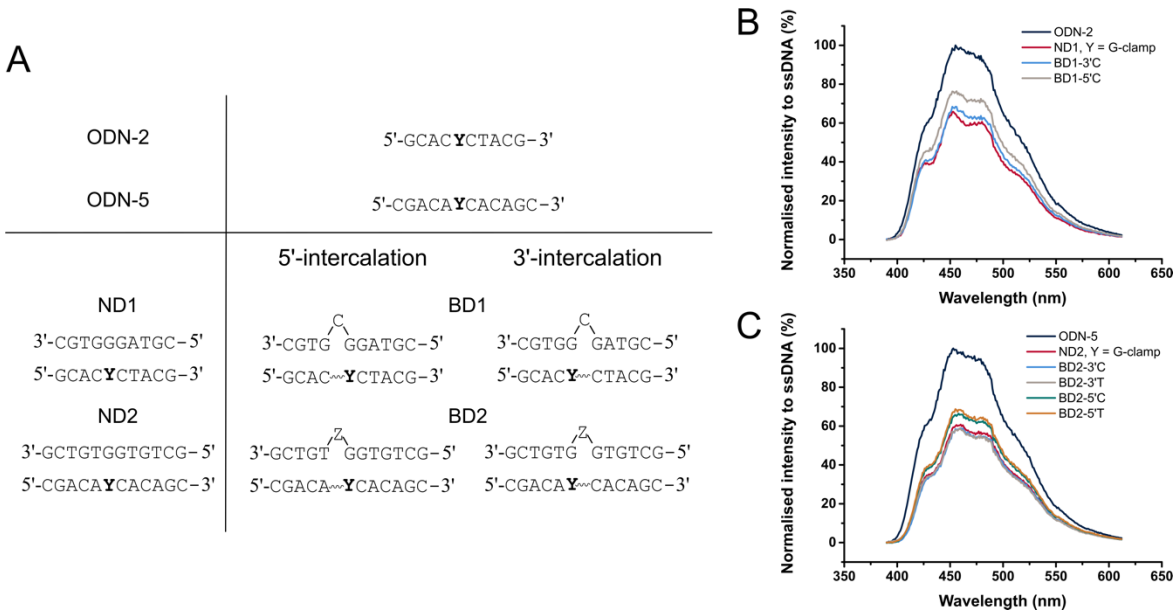


Figure 2.17 Fluorescence of G-clamp containing oligonucleotides. **A)** Oligonucleotides and corresponding duplexes involved in the fluorescence study; **B)** Fluorescence emission spectra of a G-clamp containing oligonucleotide (ODN-2) and its three duplexes (ND1, BD1-3′C, BD1-5′C); **C)** Fluorescence emission spectra of a G-clamp containing oligonucleotide (ODN-5) and its five duplexes (ND2, BD2-3′C, BD2-3′T, BD2-5′C, BD2-5′T).

2.7 Conclusions

X-pyrene combines additional non-nucleobase π-stacking, H-bonding and electrostatic attraction into a single monomer which forms a very stable base pair with guanine and therefore fully matched bulged duplexes have very high stability in DNA/DNA and DNA/RNA contexts. The presence of a single adjacent or remote mismatch greatly disrupts this stability, proving how X-pyrene could be useful for detecting mutations. Upon duplex formation, the pyrene moiety, the tC^{O} moiety and the adjacent bases extensively stack, causing fluorescence emission to decrease dramatically. This is a useful property that could be exploited in sensing applications. Due to all of these characteristics, X-pyrene could be used as a modification to enhance antisense oligonucleotides and it also could be used to produce improved fluorogenic probes for real-time PCR. The principle of combining three cooperative stabilising interactions in

a single monomer could be used as a general approach for the production of new nucleotide analogues with different nucleobases and intercalators.

Chapter 3

Pyrene Labelling of Oligonucleotides by Reductive Amination

3.1 Introduction

3.1.1 Post-Synthetic Labelling of Modified Oligonucleotides

Highly modified oligonucleotides are in constant demand for a wide variety of applications in biology, genetics, forensics or nanotechnology, *e.g.* for therapeutic or diagnostic applications such as antisense oligonucleotides or real-time PCR probes.[24,35-38,125,126] These modifications often involve either changing the structure of the backbone, the sugar or the nucleobases as well as the addition of functional moieties such as fluorophores.[24,38,51,127,128] For example, in the case of PCR probes, functionalisation with different dyes allows their use in more complex multiplex assays.[38] Modifications to DNA or RNA strands are commonly added using phosphoramidite monomers in solid-phase oligonucleotide synthesis. However, the synthesis of such phosphoramidite monomers can be challenging and the modifications need to be stable to the conditions used during the solid-phase synthesis cycle. Alternatively, it is possible to follow a different approach by post-synthetically labelling the oligonucleotides. This involves the addition of simpler phosphoramidite monomers in the DNA or RNA strands which have a chemical handle (*e.g.* a primary amine, an alkyne, *etc.*), which can react with an appropriate label. Three commonly used methodologies are the copper-catalysed azide and alkyne cycloaddition (CuAAC) reaction, the strain-promoted azide and cycloalkyne cycloaddition (SPAAC) reaction and the amine and active ester coupling reaction (Figure 1.1).[38,114,127,129,130] The linkers formed with these three labelling methods are uncharged and confer a certain rigidity to the structure *via* the resultant triazole or amide bond. When the presence of a charge

or a higher degree of flexibility is desired, it is necessary to consider alternative labelling strategies.

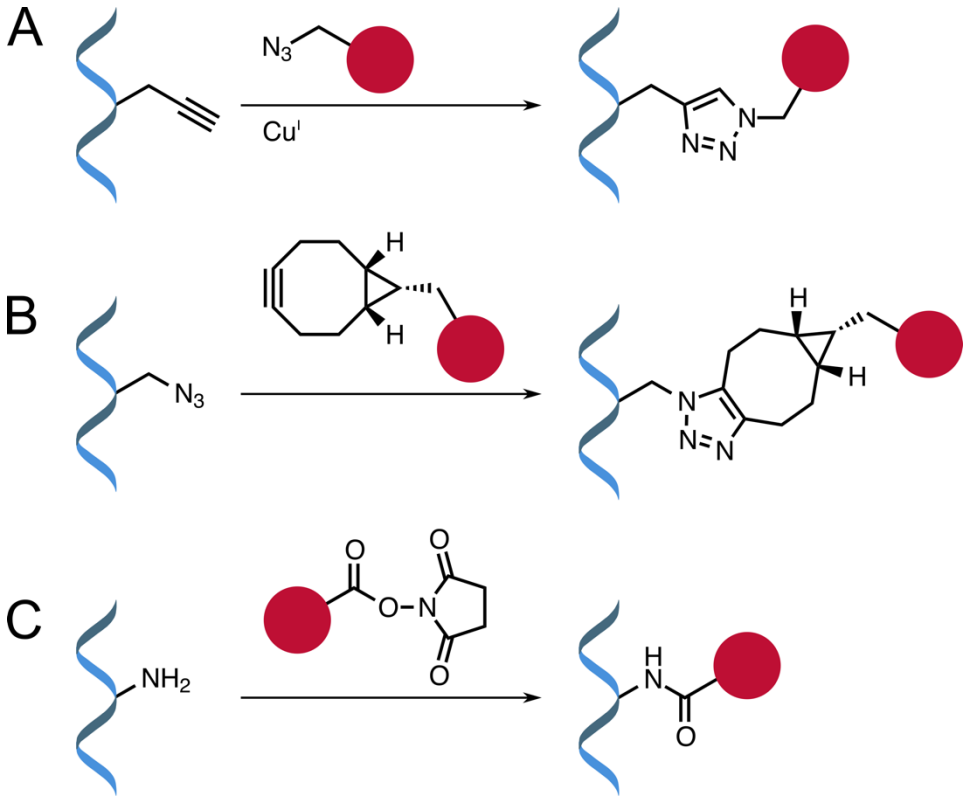


Figure 3.1 Examples of post-synthetic labelling. **A)** Copper-catalysed azide and alkyne cycloaddition reaction. **B)** Strain-promoted azide and cycloalkyne cycloaddition reaction. **C)** Amine and active ester coupling reaction.

3.1.2 Reductive Amination Labelling

The previous chapter presented the biophysical properties of X-pyrene, which proved to be a highly stabilising cytosine analogue. Its capacity to form extremely stable bulged duplexes and its mismatch sensing properties potentially make it a very good modification to probe genetic mutations, for example using real-time PCR. X-pyrene was previously synthesised as a monomer for DNA solid-phase synthesis and while this methodology allows an easy addition of multiple modifications, it also relies on an efficient synthesis of the phosphoramidites. Unfortunately, the synthesis of X-pyrene is lengthy (14 steps) and low yielding (~2 %).[104] As explained in the previous chapter,

the parent molecule of X-pyrene is G-clamp, which can be purchased as a phosphoramidite monomer. It is possible to functionalise G-clamp-modified oligonucleotides by labelling the primary exocyclic amine with 1-pyrenebutyric acid NHS ester to produce a modification that, albeit similar to X-pyrene, is not as stabilising (see previous chapter). Therefore, in the attempt to post-synthetically obtain X-pyrene with its stabilising amine linker, a reductive amination strategy was investigated. In theory, when a G-clamp containing oligonucleotide reacts with 1-pyrenebutyraldehyde in the presence of sodium cyanoborohydride, the corresponding X-pyrene containing oligonucleotide should be obtained (Figure 3.2). Reductive amination has already been used on oligonucleotides to ligate short sequences *via* an amine backbone, to conjugate DNA strands to proteins and to link a DNA duplex to an RNA sequence in order to control its folding.[49,131-133] Reductive amination labelling with formaldehyde has also been used in proteomics to tag proteins before mass spectrometry.[134,135] However, small molecule labelling of oligonucleotides with this methodology is relatively unexplored.

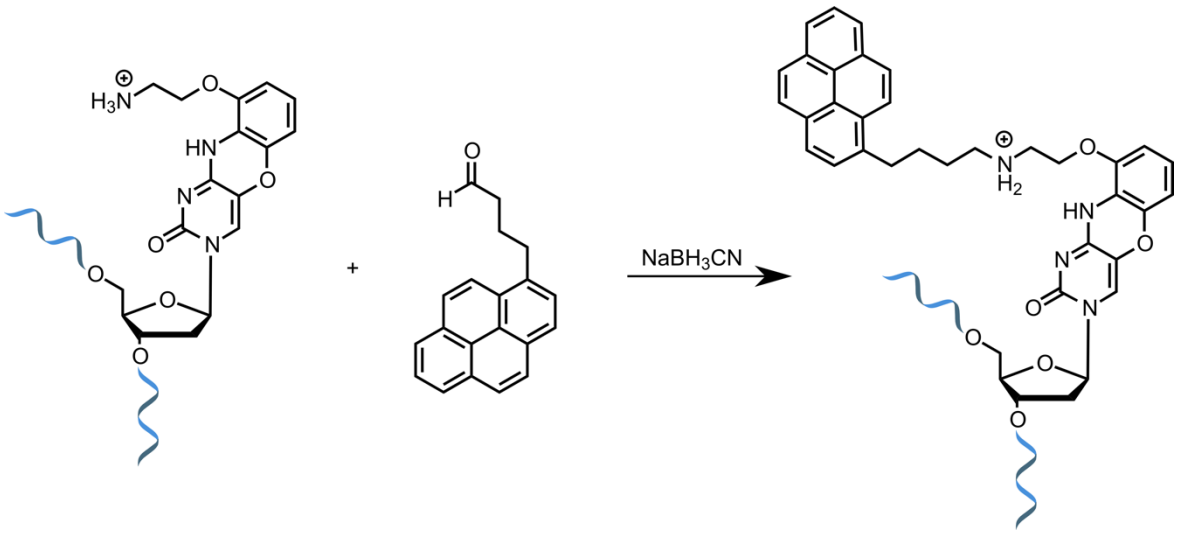


Figure 3.2 Post-solid-phase synthesis of X-pyrene. A G-clamp containing oligonucleotide can react with 1-pyrenebutyraldehyde in the presence of sodium cyanoborohydride to yield the corresponding X-pyrene containing oligonucleotide.

This chapter discusses the synthesis of X-pyrene and its derivative with two pyrenes (X-2*pyrene), directly at an oligonucleotide level, using a reductive amination labelling approach. Other modified oligonucleotides, containing monomers with a primary amine, are also labelled with pyrene using the same methodology and the effect on the thermal stability of the corresponding duplexes is studied.

3.2 Materials and Methods

3.2.1 Synthesis of 1-Pyrenebutyraldehyde

General Procedures

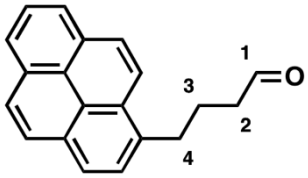
Reagents were from Sigma-Aldrich and were used without purification.

Thin layer chromatography (TLC) was carried out using Merck Kieselgel 60 F24 silica gel plates, 0.22 mm thickness, aluminium backed. The compounds were visualised under UV (254 and 365 nm).

Column chromatography was performed under pressure (air) using Fisher Scientific DAVISIL 60 Å (35–70 μm) silica.

^{1}H NMR spectra were measured at 400 MHz and ^{13}C NMR spectra were measured at 100 MHz on a Bruker DPX400 (AVIIIHD 400) spectrometer. Chemical shifts are given in ppm, relative to tetramethylsilane and *J* coupling values are quoted in Hz. ^{1}H and ^{13}C assignments were supported by DEPT-135, ^{1}H-^{1}H COSY, HMQC, HMBC experiments. Spectra were reprocessed using ACDlab program and internally referenced to the appropriate residual undeuterated solvent signal.[136]

Synthesis of 1-Pyrenebutyraldehyde



The Dess-Martin oxidation methodology was adapted from method b from the work of Meyer and Schreiber.[137]

1-Pyrenebutanol (0.5 g, 1.82 mmol, 1 eq) and Dess-Martin periodinane (1.16 g, 2.73 mmol, 1.5 eq) were dissolved in CH_2Cl_2 (10.8 mL). While stirring the reaction mixture vigorously, a solution of water in CH_2Cl_2 (36.2 µL of water in 36.2 mL of CH_2Cl_2, containing 2.01 mmol/1.1 eq of water) was added dropwise for 30 minutes. After 2 hours at room temperature the reaction mixture was washed with aqueous solutions of saturated sodium chloride, saturated potassium chloride and saturated sodium bicarbonate. The organic layer was separated, dried over sodium sulfate, filtered and then the solvent was removed *in vacuo*. The residue was purified by column chromatography on silica gel (EtOAc/hexane 3:7) after being dissolved in EtOAc, CH_2Cl_2 and hexane and adsorbed onto silica. An off-white solid of 1-pyrenebutyraldehyde was obtained with 69 % yield (0.343 g, 1.26 mmol). MW = 272.35 g/mol ($C_{20}H_{16}O$).

R_f: 0.51 [EtOAc/hexane 4:6]

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.22 (2 H, quin, *J* = 7.4 Hz, H-3), 2.59 (2 H, td, *J* = 7.0, 0.61 Hz, H-2), 3.40 (2 H, t, *J* = 7.7 Hz, H-4), 7.86 (1 H, d, *J* = 7.7 Hz, H-Ar), 8.01 (1 H, t, *J* = 7.6 Hz, H-Ar), 8.04 (2 H, s, H-Ar), 8.12 (1 H, d, *J* = 2.8 Hz, H-Ar), 8.14 (1 H, d, *J* = 4.4 Hz, H-Ar), 8.19 (2 H, dd, *J* = 7.5, 3.24 Hz, H-Ar), 8.30 (1 H, d, *J* = 9.3 Hz, H-Ar), 9.82 (1 H, t, *J* = 1.5 Hz, H-1).

^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 24.0 (C-3), 32.6 (C-4), 43.4 (C-2), 123.2 (C-Ar), 124.8 (C-Ar), 125.0 (C-Ar), 125.1 (C-Ar), 125.9 (C-Ar), 126.8 (C-Ar), 127.3 (C-Ar), 127.46 (C-Ar), 127.48 (C-Ar), 128.7 (C-Ar), 130.0 (C-Ar), 130.9 (C-Ar), 131.4 (C-Ar), 135.5 (C-Ar), 202.1 (C-1).

NMR data consistent with the literature.[138,139]

3.2.2 Oligonucleotides Synthesis, Purification and Analysis

3.2.2.1 Synthesis

Oligonucleotide synthesis was carried out on an Applied Biosystems 394 automated DNA/RNA synthesiser using a standard 0.2 µmol or 1.0 µmol phosphoramidite cycle of acid-catalysed detritylation, coupling, capping and iodine oxidation. All G-clamp, amino-modifier C6 dT and 5′-amino-modifier C6 phosphoramidite monomers were dissolved in anhydrous DCM or acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 35 s and this was extended to 6 min for all modified monomers. Stepwise coupling efficiencies and overall yields were determined by automated trityl cation conductivity monitoring and in all cases were >98.0 %. The oligonucleotides attached to the synthesis columns were treated with 20 % diethylamine in acetonitrile for 20 min then washed with acetonitrile (5 × 1 mL). This procedure removes cyanoethyl groups from the

phosphotriesters and scavenges the resultant acrylonitrile; preventing cyanoethyl adducts being formed at the primary amines of the modified monomers. After this, oligonucleotides were deprotected and cleaved from the solid support using concentrated aqueous ammonia for 5 h at 55 ˚C.

3.2.2.2 Purification

HPLC

Purification of oligonucleotides was carried out by reversed-phase HPLC on a Gilson system using a Luna® 10 µm C8 100 Å pore Phenomenex column (10 × 250 mm) with a gradient of acetonitrile in triethylamonium bicarbonate buffer (TEAB) increasing from 0 % (buffer A) to 100 % (buffer B) over 30 min with a flow rate of 4 mL/min (buffer A: 0.1 M TEAB, pH 7.5; buffer B: 0.1 M TEAB with 50 % CH_3CN, pH 7.5). After HPLC purification, the oligonucleotides were desalted by removing the solvent in a vacuum concentrator.

Ethanol Precipitation

The reductive amination reaction mixture containing the labelled oligonucleotides (1 × volume) was mixed with 3 M sodium acetate pH 5.3 (0.1 × volume) and ethanol (3 × volume). The resulting solution was left at −78 °C for 15 minutes and then at −20 °C for 20 minutes. Samples were centrifuged with a speed of 15000 × g at 4 °C for 30 minutes and the supernatant removed.

3.2.2.3 Analysis

The oligonucleotides were characterised by HPLC-MS using an Acquity UPLC system with a BEH C18 1.7 µm column (Waters) and a Bruker micro-TOF mass spectrometer (ESI− mode). A gradient of methanol in triethylamine (TEA) and

hexafluoroisopropanol (HFIP) was employed, increasing from 0 % to 70 % buffer B over 8 min, with a flow rate of 0.2 mL/min (buffer A: 8.6 mM TEA, 200 mM HFIP in 5 % methanol/water; buffer B: 20 % buffer A in methanol). All the oligonucleotides were found with an experimental mass corresponding to the expected mass (±2 Da, except when stated otherwise). Table 2.1 presents the MS results for some representative sequences.

Table 3.1 Representative oligonucleotide sequences and MS analysis

Code	Oligonucleotide sequence (5′–3′)	MS (Da)	
		Calc.	Found
ODN-1$_l$	GCAC**Y**CTACG	3378	3378
ODN-2	GCAC**X**CTACG	3122	3122
ODN-15$_l$	GCAC**Z**CTACG	3635	3634
ODN-16	GCAC**W**CTACG	3142	3142
ODN-17$_l$	GCAC**J**CTACG	3399	3398
ODN-18$_l$	GCAC**K**CTACG	3655	3655
ODN-19	**1**TATCTATATT	3151	3152
ODN-20$_l$	**2**TATCTATATT	3683	3684
ODN-21	GACGCTTCTG**3**	3181	3182
ODN-22$_l$	GACGCTTCTG**4**	3740	3741

Note: Mass spectra of modified oligonucleotides were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES− mode. Y = X-pyrene, X = G-clamp, Z = X-2*pyrene, W = amino C6 dT, J = amino C6 dT + 1 pyrene, K = amino C6 dT + 2 pyrenes, 1 = 5′-amino-modifier C6, 2 = 5′-amino-modifier C6 + 2 pyrenes, 3 = 3′-amino-modifier C7, 4 = 3′-amino-modifier C7 + 2 pyrenes. The subscript "l" is used for the oligonucleotides obtained after labelling *via* reductive amination.

3.2.3 Reductive Amination Labelling with 1-Pyrenebutyraldehyde

3.2.3.1 Labelling of G-Clamp

ODN-2 (40 nmol) was lyophilised and then dissolved in 70 µL of sodium phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH 7.2). An equal volume of 1-pyrenebutyraldehyde dissolved in dimethylformamide (2000 nmol, 50 eq) was added to the DNA solution, followed by 8 µL of concentrated sodium cyanoborohydride dissolved in water (8000 nmol, 200 eq). The reaction mixture was incubated at 55 °C

for 1 hour to obtain the mono addition product and for 7 hours to obtain the double addition product. The reaction was stopped *via* ethanol precipitation, which simultaneously desalted the labelled oligonucleotides that were then purified by reverse-phase HPLC.

3.2.3.2 Labelling of Amino C6 dT

Mono Addition Product

ODN-16 (50 nmol) was lyophilised and then dissolved in 70 µL of sodium phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH 7.2). An equal volume of 1-pyrenebutyraldehyde dissolved in dimethylformamide (800 nmol, 16 eq) was added to the DNA solution, followed by 8 µL of concentrated sodium cyanoborohydride dissolved in water (4000 nmol, 80 eq). The reaction mixture was incubated at 55 °C for 23 hours to obtain the mono addition product. The reaction was stopped *via* ethanol precipitation, which simultaneously desalted the labelled oligonucleotides that were then purified by reverse-phase HPLC.

Double Addition Product

ODN-16 (50 nmol) was lyophilised and then dissolved in 70 µL of sodium phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH 7.2). An equal volume of 1-pyrenebutyraldehyde dissolved in dimethylformamide (1200 nmol, 24 eq) was added to the DNA solution, followed by 8 µL of concentrated sodium cyanoborohydride dissolved in water (4000 nmol, 80 eq). The reaction mixture was incubated at 55 °C for 23 hours to obtain the double addition product. The reaction was stopped *via* ethanol precipitation, which simultaneously desalted the labelled oligonucleotides that were then purified by reverse-phase HPLC.

3.2.3.3 Labelling of 5′-Amino-Modifier C6 and 3′-Amino-Modifier C7

ODN-19 or ODN-21 (50 nmol) was lyophilised and then dissolved in 70 µL of sodium phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH 7.2). An equal volume of 1-pyrenebutyraldehyde dissolved in dimethylformamide (2000 nmol, 40 eq) was added to the DNA solution, followed by 8 µL of concentrated sodium cyanoborohydride dissolved in water (8000 nmol, 160 eq). The reaction mixture was incubated at 55 °C for 30 hours to obtain the double addition product. The reaction was stopped *via* ethanol precipitation, which simultaneously desalted the labelled oligonucleotides that were then purified by reverse-phase HPLC.

3.2.4 UV Melting

UV melting samples were prepared in 5 mM sodium phosphate buffer with 140 mM KCl, 1 mM $MgCl_2$ at pH 7.2. Prior to use, the buffer was filtered with Kinesis regenerated cellulose 13 mm, 0.45 µm syringe filters and degassed by sonication for 10 min. Concentration of both target and probe (labelled) oligonucleotides was 2.5 µM. Melting analysis was performed on a Cary 4000 UV-Vis spectrometer fitted with a Cary temperature controller. Samples were analysed in 1.2 mL cuvette (Hellma Precision cells made of synthetic quartz glass, QG, 1 cm path length). Before melting curves were recorded, the samples were melted and then slowly annealed. UV melting was carried out from 18 °C to 84 °C with ramping of 1 °C/min. UV absorbance at 260 nm was recorded every 0.1 °C. The melting temperatures (T_m) were calculated using Cary Win UV thermal application software, the values reported are an average of the four melting curves. For the tandem probe system, the data were analysed with Origin 8.6 and for simplicity the curves reported are for a single melting only. The following table summarises the procedure used:

Stage	Data interval (°C)	Rate (°C/min)	End temperature (°C)	Hold time (min)
Melting 0	1.00	10	84	2
Annealing 1	0.10	1	18	2
Melting 1	0.10	1	84	2
Annealing 2	0.10	1	18	2
Melting 2	0.10	1	84	2
Annealing 3	0.10	1	18	2
Melting 3	0.10	1	84	2
Annealing 4	0.10	1	18	2
Melting 4	0.10	1	84	2
Annealing 5	1.00	10	20	0

3.2.5 Error Calculations

Melting temperature errors were calculated using the confidence intervals derivable with a Student's t-test with 95 % confidence and the appropriate degrees of freedom. For ΔT_m the errors were calculated using the normal error propagation formula without the covariance factor. In graphs errors are represented by error bars.

3.3 Reductive Amination

In order to label G-clamp with pyrene *via* reductive amination, it was necessary to first synthesise 1-pyrenebutyraldehyde, which was easily obtained in a one-step procedure *via* a Dess-Martin oxidation following the methodology of Meyer and Schreiber (Figure 3.3).[137]

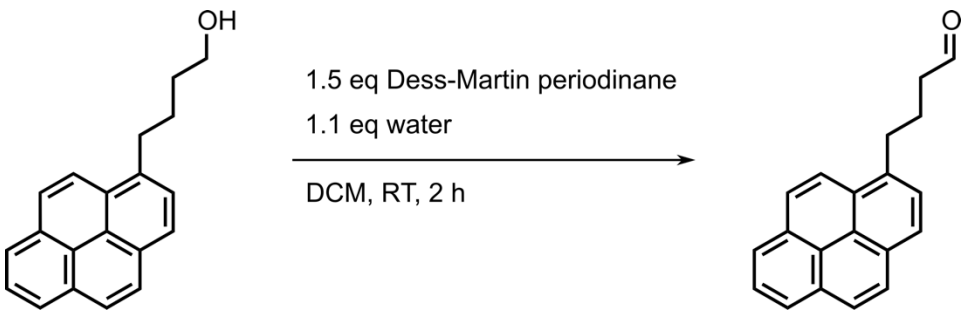


Figure 3.3 Synthesis of 1-pyrenebutyraldehyde *via* a Dess-Martin oxidation.

Initial test reductive aminations evaluated the organic solvent required to dissolve 1-pyrenebutyraldehyde, the pH of the buffer and the temperature of the reaction. DMF was selected to dissolve the label over dimethyl sulfoxide (DMSO) and methanol as its use resulted in slightly higher yields, while the pH of the buffer had little influence on the outcome when between 6 and 8 and therefore, the commonly used phosphate buffer at pH 7.2 was chosen. The reaction temperature was selected to be 55 °C, as this is also often used for active ester labelling and after discovering that at lower temperatures the reductive amination was considerably slower and at higher temperatures the decomposition of the reducing agent was more likely to occur.

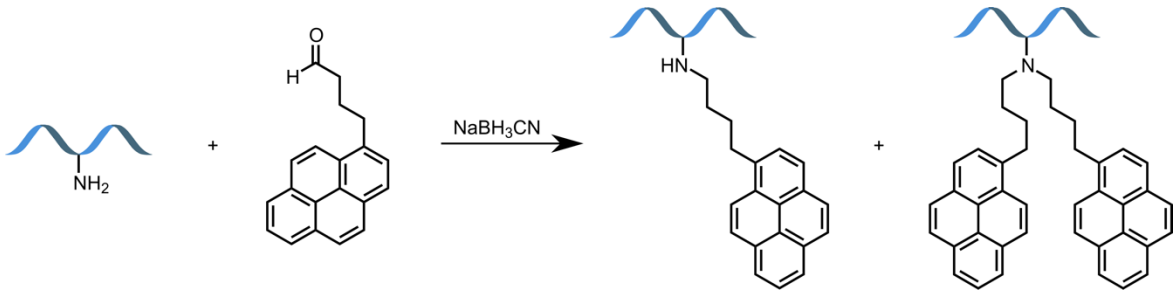


Figure 3.4 Products of reductive amination with 1-pyrenebutyraldehyde. Both the mono addition product and the more favourable double addition product are obtained.

For the general reductive amination labelling procedure adopted, the amine containing oligonucleotides were lyophilised and then dissolved in a sodium phosphate buffer (10 mM NaH_2PO_4, 200 mM NaCl, pH 7.2). An equal volume of 1-pyrenebutyraldehyde dissolved in dimethylformamide (DMF) was added to the DNA solution, followed by sodium cyanoborohydride dissolved in water. The reaction mixtures were incubated at 55 °C. Unfortunately, the reductive amination consistently yielded two products (Figure 3.4), in analogy to what was previously reported for proteins.[134,135] The aldehyde reacts with the primary amine in the DNA strand forming a Schiff base, which is then reduced by sodium cyanoborohydride forming a secondary amine, which is comparatively more reactive than the starting primary amine. Therefore, if there is

enough 1-pyrenebutyraldehyde and reducing agent and if the reaction is allowed to go to completion, the main product obtained will be the double addition product. Unfortunately, it is not possible to obtain only the mono addition product, even if only a slight excess of the aldehyde is used, as the reaction on an oligonucleotide is not efficient: the equilibrium of the formation of the Schiff base is unfavourable in an aqueous environment and therefore a higher amount of aldehyde is necessary to drive the formation of the imine intermediate.[133]

In order to attempt the post-synthetic synthesis of an X-pyrene oligonucleotide, ODN-2 (5′-GCACYCTACG, Y = G-clamp) was selected to perform the experiments. The oligonucleotide was reacted with an excess of 1-pyrenebutyraldehyde (50 eq) and sodium cyanoborohydride (200 eq), in order to overcome the unfavourable imine formation and to ensure a fast reduction even if part of the reducing agent degraded over time. After 1 hour (Figure 3.5 A), although there still was a considerable amount of starting material left, the main product was the mono-labelled oligonucleotide (ODN-1_l, containing a single X-pyrene) with only traces of the double addition product (ODN-15_l, containing a single X-2*pyrene). After 2 hours (Figure 3.5 B), the situation changed considerably as the main product was ODN-15_l (X-2*pyrene), although significant amounts of both ODN-2 (G-clamp) and ODN-1_l (X-pyrene) were still present. After 7 hours (Figure 3.5 C), the reaction was almost completed, with only traces left of both starting material and mono addition product. These results showed that obtaining the double addition oligonucleotide with this methodology was trivial, as it was essentially the only product obtained when many equivalents of the reagents were used and the reaction times were long enough. However, obtaining the mono addition product (X-pyrene) was far more challenging, as it reacts again to produce

X-2*pyrene. While decreasing the equivalents of aldehyde and reducing agent that were used and consequentially increasing the reaction time worked as a strategy to improve the yield of the mono-labelled product, the inherent lack of reproducibility of this reaction, also due to the small amounts used when working with oligonucleotides, makes it suboptimal as a consistent and easy to use mono-labelling method. The efficiency of the reaction depends on too many factors, as every time fresh solutions of the reagents need to be prepared and for this purpose minute amounts need to be weighted. Also the pipetting errors tend to be significant, as the volumes used are small. Moreover, it is difficult to accurately measure the very small amounts of oligonucleotides used. Finally, every time the label or the labelling site is changed it is often necessary to perform new optimisation cycles to improve the yields. All of this is highly time and material consuming and therefore, it was necessary to adopt another more convenient strategy. It is important to note that purification of the unlabelled, mono-labelled and di-labelled products was simple by reversed phase HPLC (RP-HPLC), making separation of product mixtures relatively easy. Considering this, the most productive approach to obtain a mono-labelled oligonucleotide involved stopping the reaction early to recover the mono addition oligonucleotide and the unreacted started material, before reusing the starting material for another reductive amination cycle (*i.e.* a batch process).

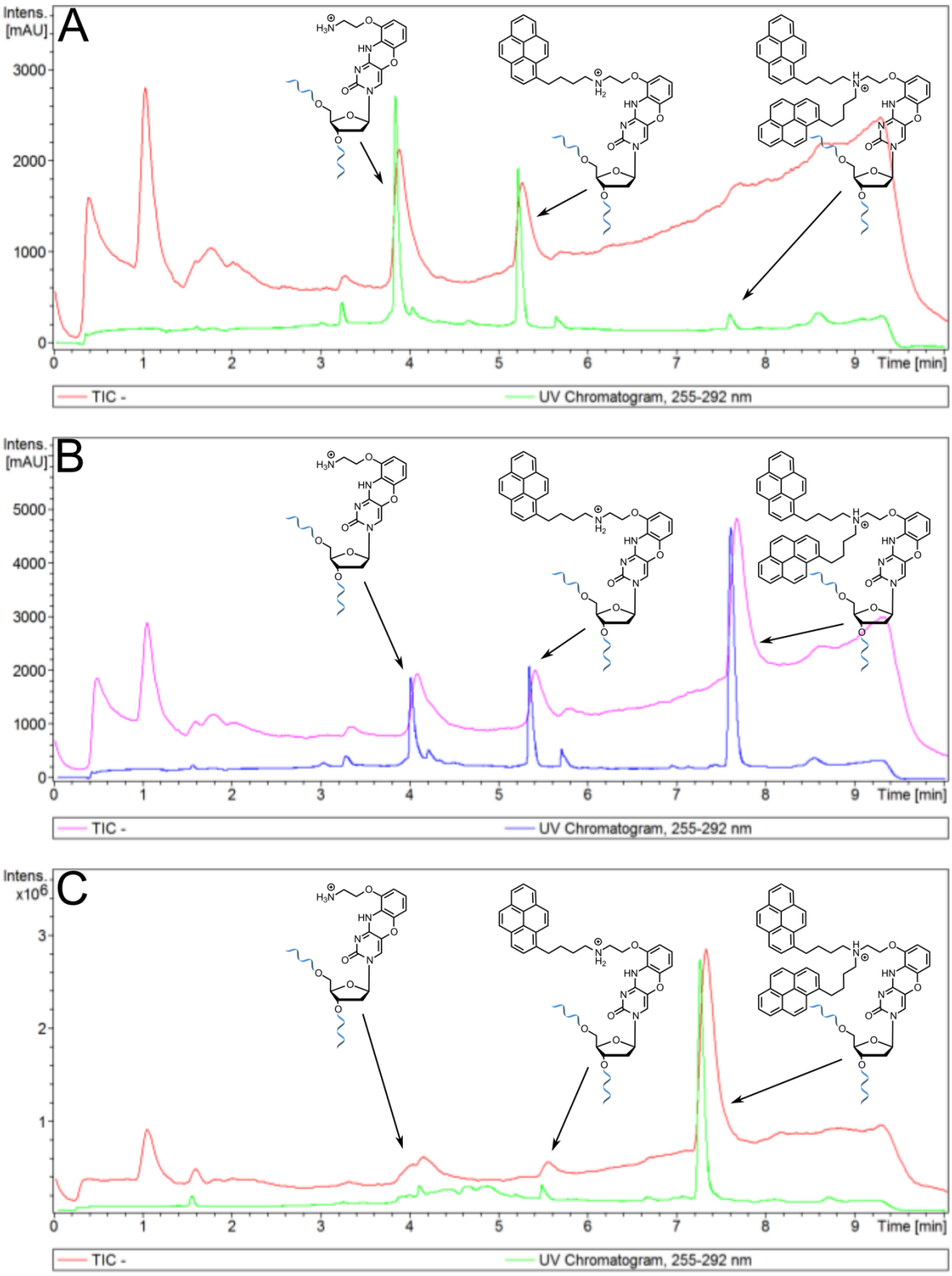


Figure 3.5 Chromatograms of the labelling of the G-clamp containing ODN-2 with 1-pyrenebutyraldehyde *via* reductive amination. X-pyrene was the mono-labelled product and X-2*pyrene was the di-labelled product. The identity of the peaks was confirmed by mass spectrometry. **A)** Reaction after 1 hour. **B)** Reaction after 2 hours. **C)** Reaction after 7 hours. The amounts used were: 40 nmol of ODN-2 (at a concentration of 270 μM), 50 eq of 1-pyrenebutyraldehyde and 200 eq of sodium cyanoborohydride.

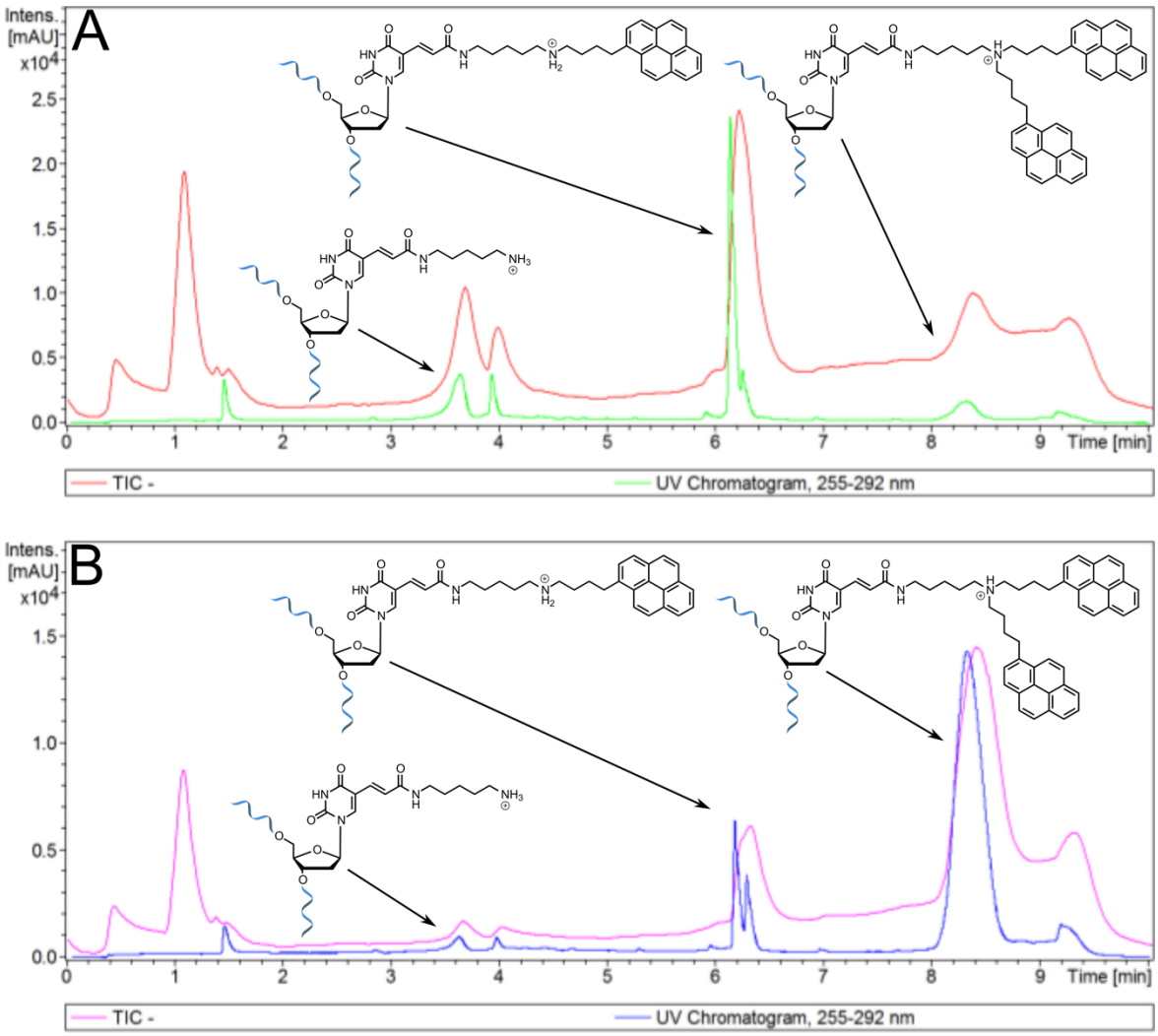


Figure 3.6 Chromatograms of the labelling of the amino C6 dT containing ODN-16 with 1-pyrenebutyraldehyde *via* reductive amination. The identity of the peaks was confirmed by mass spectrometry. **A)** Production of the mono-labelled product ODN-17$_l$: 50 nmol of ODN-16 (at a concentration of 338 μM), 16 eq of 1-pyrenebutyraldehyde and 80 eq of sodium cyanoborohydride. **B)** Production of the di-labelled product ODN-18$_l$: 50 nmol of ODN-16 (at a concentration of 338 μM), 24 eq of 1-pyrenebutyraldehyde and 80 eq of sodium cyanoborohydride.

Another amine modified oligonucleotide used for reductive amination labelling was ODN-16 (5′-GCACYCTACG, Y = amino C6 dT), which contained a single amino C6 dT (see first structure in Figure 3.6). In this case, two sets of reaction conditions were investigated, one to obtain the mono-labelled product and one for the double addition one. In both cases the reactions were left at 55 °C for 23 hours with a large excess of sodium cyanoborohydride (80 eq), while 16 eq of 1-pyrenebutyraldehyde were used

to obtain the mono addition product (ODN-17$_l$, Figure 3.6 A) and 24 eq to obtain the di-labelled one (ODN-18$_l$, Figure 3.6 B). In both cases the main product was the desired one.

Considering how easy it was to obtain a double addition for both the modifications studied using the reductive amination labelling, two oligonucleotides, probing for the cystic fibrosis transmembrane conductance regulator (CFTR) gene of human DNA (ABCC7), were selected to study whether the extra hydrophobic interactions generated by the presence of two pyrenes could be used for a split-oligonucleotide (tandem) probe system (see Figure 3.10 and the end of section 3.4). ODN-19 (5′-1TATCTATATT, 1 = 5′-amino-modifier C6) and ODN-21 (5′-GACGCTTCTG2, 2 = 3′-amino-modifier C7) were reacted at 55 °C for 30 hours with a large excess of 1-pyrenebutyraldehyde (40 eq) and sodium cyanoborohydride (160 eq). Both reactions yielded the desired di-labelled products (Figure 3.7).

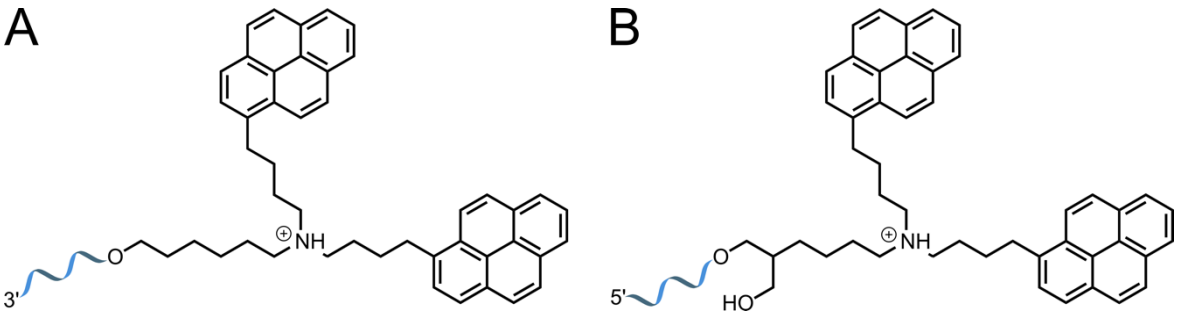


Figure 3.7 Reductive amination labelling of terminal amines. **A)** ODN-20$_l$, product of the double labelling of the 5′-amino-modifier C6 containing ODN-19. **B)** ODN-22$_l$, product of the double labelling of the 3′-amino-modifier C7 containing ODN-21. The amounts used were: 50 nmol of ODN-19 or ODN-21 (at a concentration of 338 μM), 40 eq of 1-pyrenebutyraldehyde and 160 eq of sodium cyanoborohydride.

3.4 UV Melting Studies

When ODN-2 (containing a single G-clamp) was labelled with 1-pyrenebutyraldehyde two products were obtained, the mono- (ODN-1_l, containing a single X-pyrene) and the di-labelled (ODN-15_l, containing a single X-2*pyrene) G-clamps. While X-pyrene possesses significant stabilising properties, the effect of the second pyrene of X-2*pyrene were unknown. Therefore, its thermal stability was studied when this modification was present in a normal duplex (ND), a 3′-bulged duplex (3′-BD) and a 3′- and 5′-bulged duplex (3′, 5′-BD) (Figure 3.8). Predictably, given the extra steric hindrance produced by two pyrenes, when X-2*pyrene was present in a fully complementary duplex (ND1) it significantly destabilised the dsDNA compared to X-pyrene (−13.4 °C) and G-clamp (−8.8 °C). Yet, remarkably, the bulky modification still formed a more stable duplex than a natural C (+2.5 °C). Adding a 3′-bulge (3′-BD1) allowed X-2*pyrene to sit more comfortably in the duplex, making it more stabilising than a G-clamp (+9.4 °C) or an unmodified cytosine base (+27.2 °C). However, there was probably not enough space for two pyrenes to properly intercalate, as X-2*pyrene was 5 °C less stabilising than X-pyrene. In order to verify if this loss of stability was linked to reduced intercalation, another bulge was added on the 5′-side of the modification (3′, 5′-BD1), thereby creating another pocket where pyrene could potentially intercalate. In agreement with this hypothesis, having a 3′- and a 5′-bulge increased the thermal stability by 2.7 °C compared to a duplex with only a 3′-bulge.

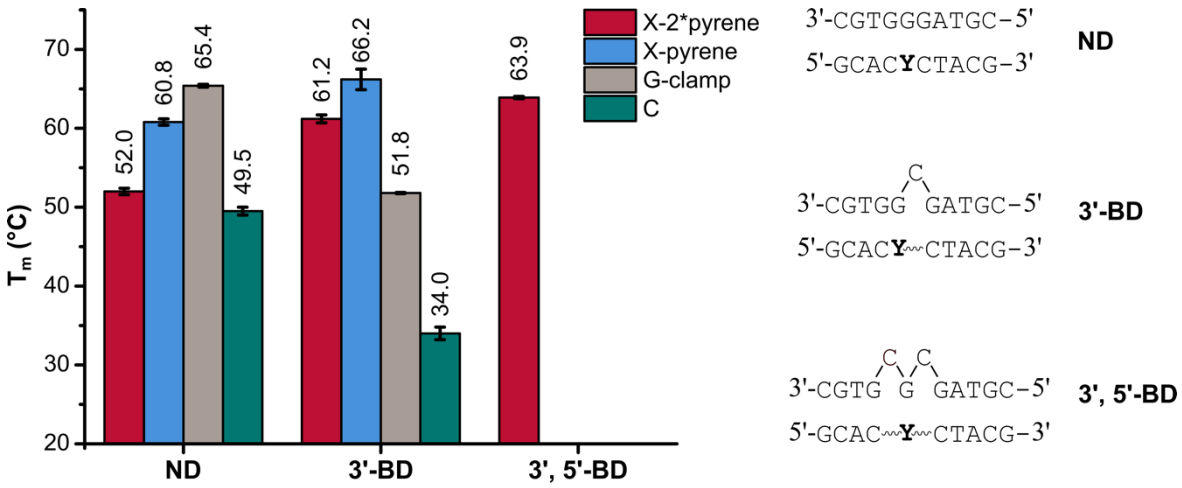


Figure 3.8 Effects of X-2*pyrene on the thermal stability of ND1, 3′-BD1 and 3′, 5′-BD1. Y = X-2*pyrene, X-pyrene, G-clamp or C. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 140 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2.5 μM).

The labelling of ODN-16 (containing a single amino C6 dT, see structure in Figure 3.6) with 1-pyrenebutyraldehyde increased the thermal stability of a normal duplex (ND), a 3′-bulged duplex (3′-BD) and a 5′-bulged duplex (5′-BD) (Figure 3.9). The substitution of T with amino C6 dT slightly increased the melting temperature of all the duplexes, probably due to the positively charged amine. After the amine was labelled with a single pyrene, the resulting normal duplex was 9.6 °C more stable than when T was present instead. This was in contrast to X-pyrene, which destabilised normal duplexes compared to G-clamp. The difference might be linked to the fact that G-clamp has an extended nucleobase, protruding into the major grove and leaving so less space for intercalation. This feature is absent in amino C6 dT, which also has a longer and more flexible linker that might facilitate good intercalation. The stabilising effect of a mono-pyrene-labelled amino C6 dT was more prominent in both 3′- (+18 °C) and 5′- (+11 °C) bulged duplexes as, in an analogous way to X-pyrene, the pyrene moiety probably sits in the pocket produced by the bulged nucleotide, creating a more continuous pi-stacking system. Another similarity to X-pyrene was the preference of a 3′-bulge over a 5′-bulge, since pyrene might again intercalate better on the 3′-side.

Predictably, adding a second pyrene destabilised the normal duplex (−6.5 °C) compared to the case with only one pyrene. Remarkably, this bulky modification was also more thermally stabilising than amino C6 dT (+2.1 °C) and T (+3.1 °C). Moreover, it is interesting to note that, in both the 3′- and 5′-bulge cases, the second pyrene did not seem to have a predominant role, as the melting temperature varied by about 1 °C in both instances.

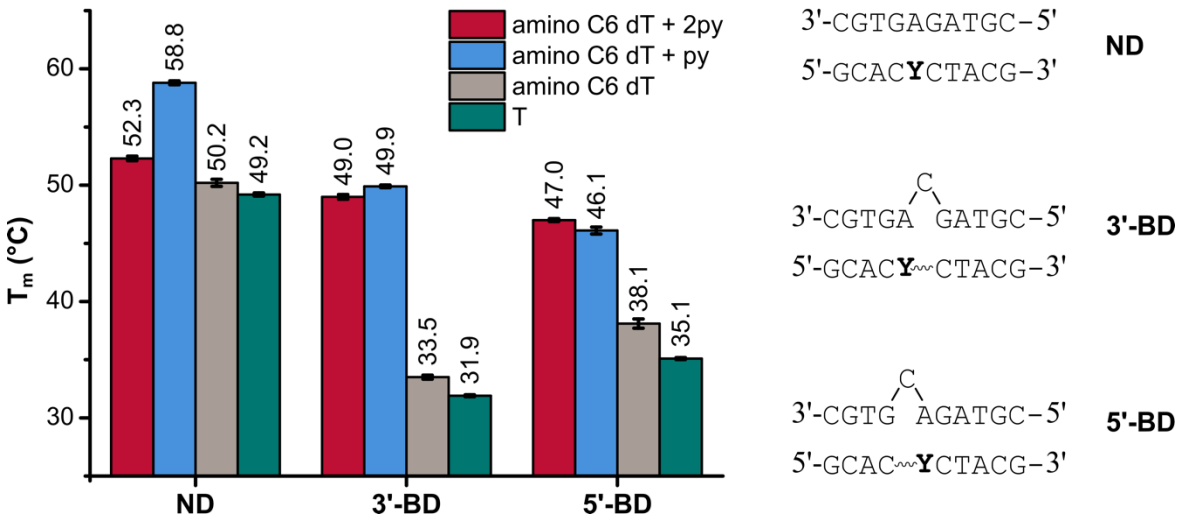


Figure 3.9 Effects of the labelling of amino C6 dT (ODN-16) on the thermal stability. Y = amino C6 dT, amino C6 dT labelled with two pyrenes (+ 2py), amino C6 dT labelled with one pyrene (+ py) or T. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 140 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2.5 μM).

In order to be specific for a single DNA sequence in a human genome, a probe generally needs to be at least 16 nucleotides long.[32] However, once a duplex of this length is formed, the melting temperature of single mismatches is only slightly reduced compared to a fully matched duplex.[32] Short probes (less than 10 nucleotides) are instead highly sensitive to mismatches. In order to obtain the specificity of a long probe and the discrimination capacity of a short probe, a split-oligonucleotide (tandem) probe system can be exploited. This involves the use of two short probes that, after hybridising to the target contiguously without a gap, can interact to form a single

longer probe. Tandem probe systems containing pyrene have been used many times for exciplex or excimer based fluorescence detection of DNA.[84,86,87,89,90,140-143] Pyrenes tend to form excimers when two units are 3–4 Å apart from each other and, in fact, in the crystal structures of pyrene the interplanar distance is 3.54 Å, interestingly close to the 3.4 Å between two adjacent bases in a B-form dsDNA.[11,12,78,84,87,89] Given how easy it was to achieve a doubly-labelled product with reductive amination labelling, a tandem probe system for the human cystic fibrosis transmembrane conductance regulator (CFTR) gene was designed and synthesised by reductive amination labelling with the 3′- or the 5′-end of two short probes with 1-pyrenebutyraldehyde. The aim was to exploit the potential interaction among the four pyrenes, which could promote the formation of a longer probe (Figure 3.10).

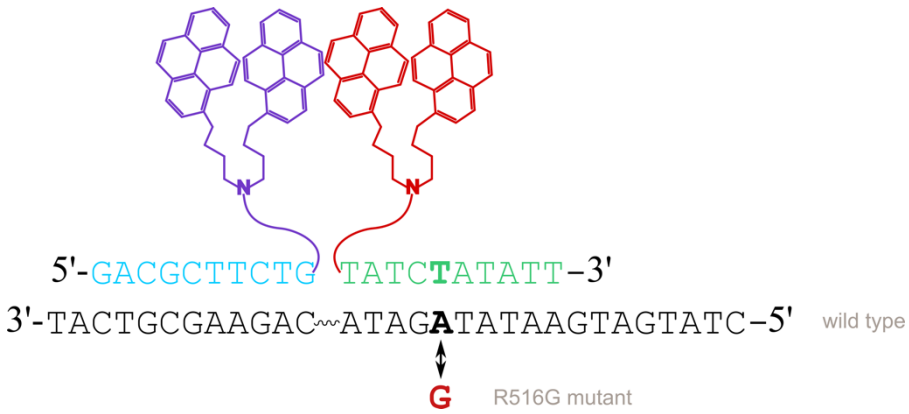


Figure 3.10 Split-oligonucleotide (tandem) probe system for the CFTR gene of human DNA (ABCC7). ODN-20_l is in green/red, while ODN-22_l is in light blue/purple. The wild type gene contains an A (black), while the R516G mutant contains a G (red) opposite to ODN-20_l.

Surprisingly, this tandem probe system gave more than one melting temperature both when hybridised to the wild type target (Figure 3.11 A) and the R516G mutant (Figure 3.11 B), implying that the two short probes are not working cooperatively.

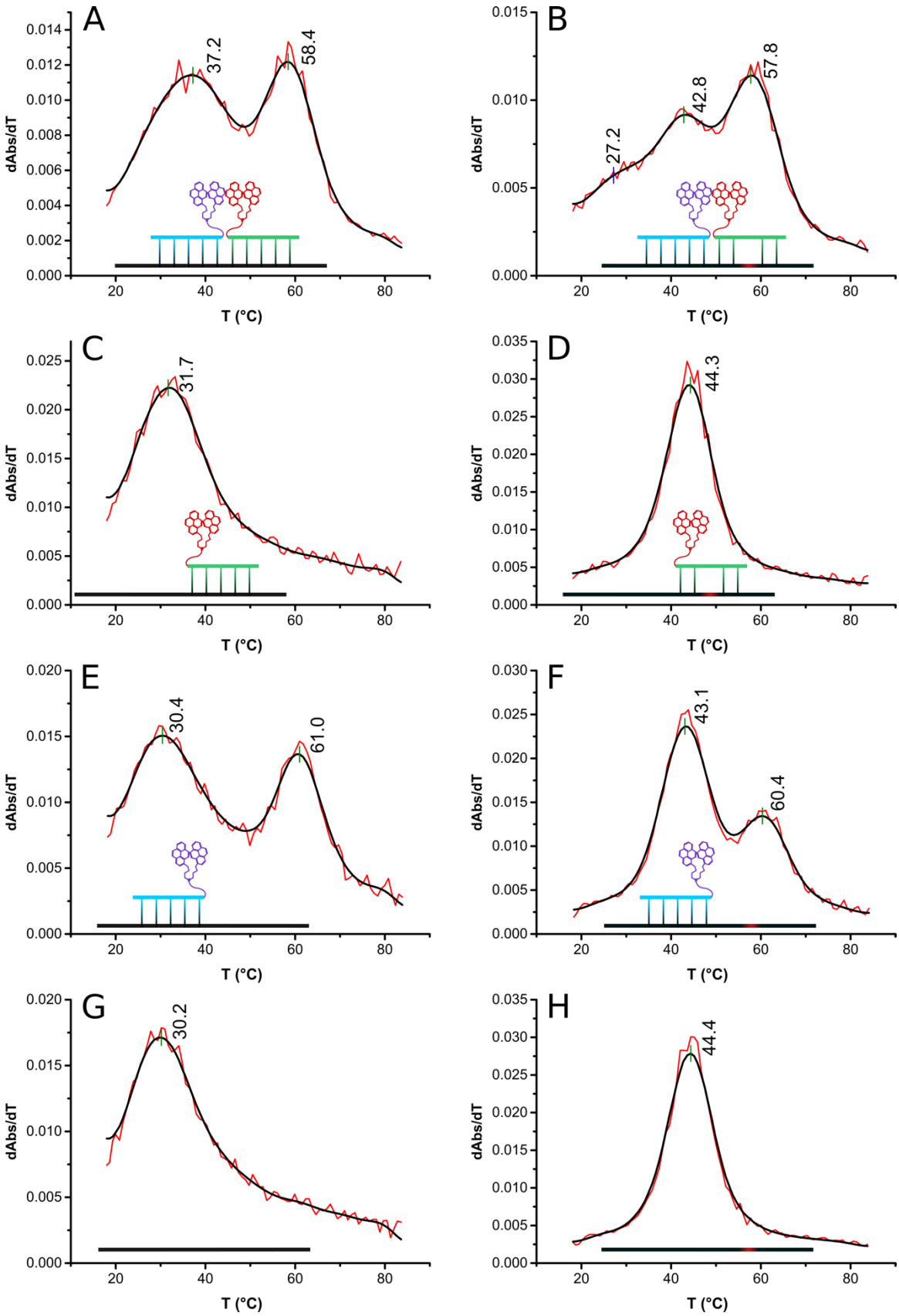


Figure 3.11 UV melting first derivative plots of the split-oligonucleotide (tandem) probe system for the CFTR gene of human DNA. In red is the first derivative of the raw data (collected at 260 nm), while in black is the curve smoothed with Origin 8.6. T_m values are reported next to the peaks.

In order to identify the different peaks, the same UV melting experiments were repeated using only one of the two short probes. When only the probe opposite to the mutation site was present (ODN-20_l, in green), there was a single melting transition, both when it was hybridised to the wild type target (Figure 3.11 C) and to the R516G mutant (Figure 3.11 D). The melting temperatures of these duplexes, albeit slightly different, were relatable to one of the T_m values measured for the corresponding two probes system. When ODN-22_l (in light blue) was the only probe present, two transitions were observed in the first derivatives of the melting curves, both with the wild type target (Figure 3.11 E) and the R516G mutant (Figure 3.11 F). The highest temperature peaks were related to the rightmost peaks of the corresponding two probes system, while the remaining peaks were comparable to the peaks present when only ODN-20_l was used. This result suggested that the melting transitions in Figure 3.11 C and D are not related to the melting of the probes, but they are instead probably linked to the unfolding of a secondary structure of the targets. This was confirmed by the UV melting of the wild type target (Figure 3.11 G) and the R516G mutant target (Figure 3.11 H) with no probes. The formation of these stable secondary structures was further supported by analysis of the targets on the mfold web server, which found that those sequences can form hairpins (Figure 3.12).[144] According to the results obtained from the mfold web server, the region targeted by ODN-20_l is involved in the hairpin formation, thus explaining why ODN-20_l did not hybridise. Instead the region targeted by ODN-22_l is available and therefore ODN-22_l was able to anneal to its target as highlighted by the presence of the high temperature T_m peak in Figure 3.11 E and F. In order to further characterise the system and to confirm the conclusions derived from the mfold results, UV melting experiments performed at different concentrations would be useful, as the melting temperature of intramolecular structures is not

concentration dependent. Additionally, fluorescence experiments would be beneficial, especially if aimed at the detection of the formation of eventual pyrene excimers. In conclusion, the pyrene-pyrene interactions were not strong enough in this case to completely counteract the formation of the secondary structures of the targets during annealing, rendering this particular tandem probe system ineffective.

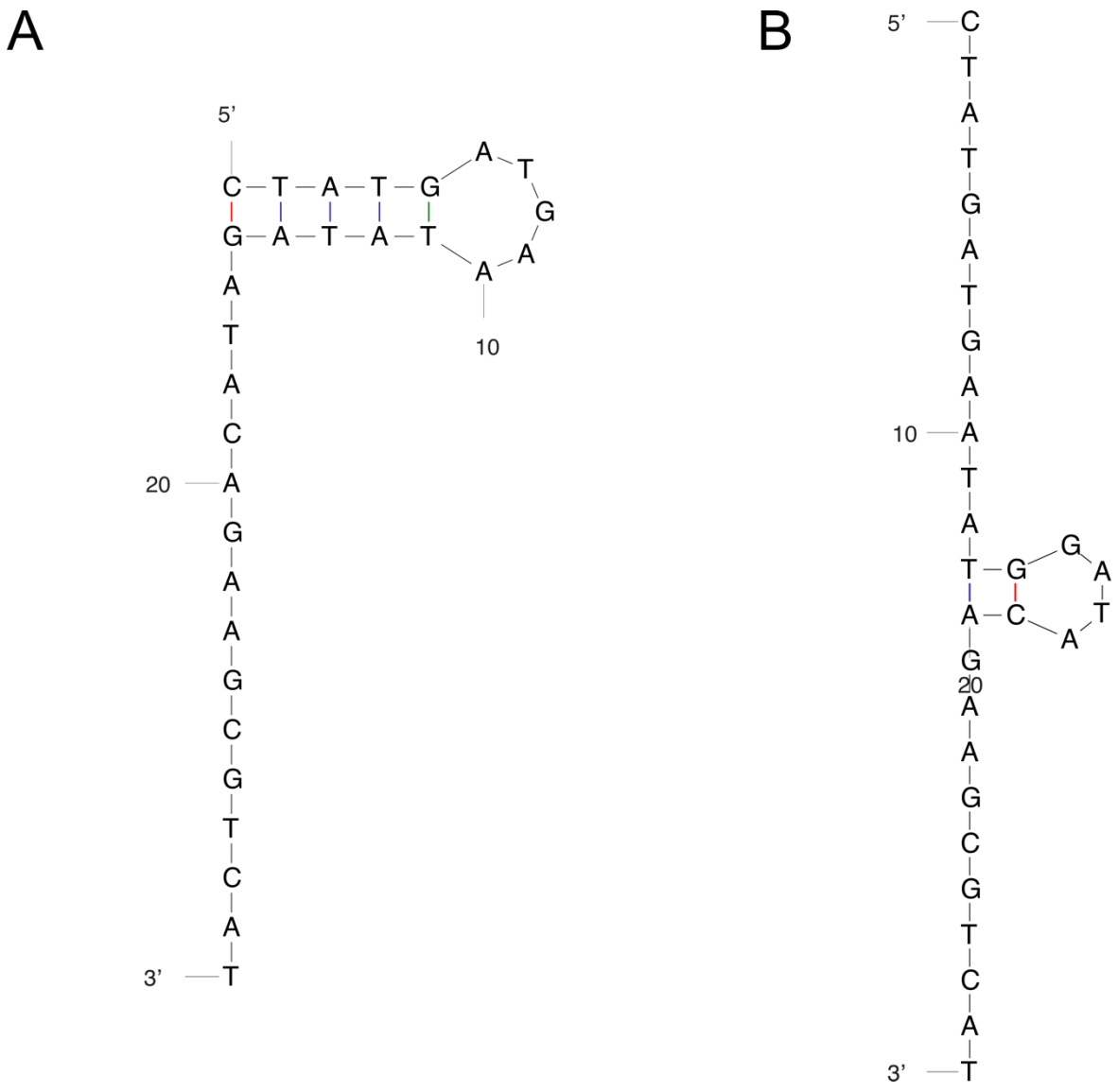


Figure 3.12 Secondary structures predicted by the mfold web server for the wild type target (**A**) and for the R516G mutant target (**B**).

3.5 Conclusions

An X-pyrene containing oligonucleotide was successfully prepared *via* a post-synthetic reductive amination labelling of a G-clamp monomer with 1-pyrenebutyraldehyde. This methodology preferentially yielded the double addition product X-2*pyrene, which was not as stabilising as X-pyrene, although it still formed significantly stable normal and bulged duplexes even with its high steric hindrance. Considering its ease of preparation, it would be interesting to continue studying the biophysical properties of X-2*pyrene and eventually its potential applications as well.

An amino C6 dT containing oligonucleotide was also successfully labelled with 1-pyrenebutyraldehyde exploiting the same methodology. The single addition product significantly stabilised both normal and bulged duplexes and, surprisingly, the favoured double addition product behaved very similarly with bulged duplexes. However, it did not stabilise the normal duplex as much. Further studies are necessary to fully understand the behaviour of these modifications in different sequence environments and in the presence of mismatches.

3′- and 5′-terminal amino modified oligonucleotides were easily labelled with two pyrenes to produce probes for a split-oligonucleotide (tandem) probe system. This was used to analyse a mutation in the CFTR gene of human DNA. In the specific example that was studied this approach proved to be unsuccessful as the pyrene-pyrene interactions were not strong enough to force the hybridisation of the probes over the formation of the secondary structures of the target. It would be interesting to see if an improved probe design could overcome the problems. This could be achieved by changing the probing sequences and/or by shortening the linker to reduce the

flexibility and the freedom of movement of the pyrene moieties to increase the chances of pyrene-pyrene interactions.

It would be interesting to try to invert the approach explained in this chapter for the reductive amination, having an aldehyde-modified oligonucleotide and an amine-modified label. This change could help to obtain a more selective reaction towards single addition products, as it would be feasible to have a large excess of the small molecule amine in order to facilitate the production of the mono-labelled product. It is also important to highlight that aldehyde monomers, usable with solid-phase oligonucleotides synthesis, are already commercially available.

Overall, the reductive amination labelling with 1-pyrenebutyraldehyde, although not particularly reproducible, produced modifications with good duplex stabilising properties. Moreover, since it was relatively easy to obtain both a mono and a double addition, it is possible to envisage its use for fast production of other modifications for screening purposes.

Chapter 4

U-pyrene

4.1 Preface

The author would like to thank Dr Nittaya Gale at ATDBio Ltd. for the synthesis of the U-pyrene monomer and for the synthesis and purification of the oligonucleotides. Many thanks also to Dr Denis Ptchelkine, who collaborated on the X-ray crystallography studies.

4.2 Introduction

As highlighted in the previous chapters, pyrene modified oligonucleotides can help stabilise dsDNA, and both X-pyrene and the mono-pyrenated amino C6 dT proved to be good at this. The only drawback for both modifications lies in their synthesis. X-pyrene was made either with a lengthy and low yielding phosphoramidite monomer synthesis or by post synthetic labelling *via* reductive amination, in the same way as the mono-pyrenated amino C6 dT was obtained. The limit with the reductive amination labelling was the inherent difficulty of avoiding the formation of the double addition product, which would naturally make labelling multiple sites on the same oligonucleotide problematic. Therefore, especially if there is an interest in studying the effects of multiple pyrene modifications in the same sequence, it is necessary to select a synthetically more tractable modification. U-pyrene (5-(pyren-1-ylbutanamidopropargyl)dU, Figure 4.1) was previously studied by Richert and collaborators and it was synthesised in only three steps (66 % overall yield).[145,146] Moreover, multiple additions of this modification on a single sequence were reported and even a single U-pyrene significantly stabilised the duplexes formed (+6.4–7.1 °C).[145,146] For all these reasons U-pyrene was selected for further studies to investigate the thermodynamic effects of multiple additions. Moreover, X-ray

crystallography was employed to study the influence U-pyrene has on dsDNA structure.

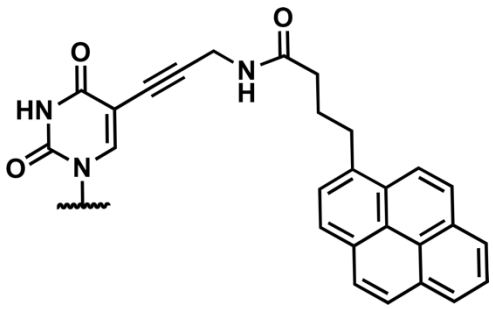


Figure 4.1 Structure of U-pyrene.

4.2.1 UV Melting and Thermodynamics

Numerous publications report methods for deriving thermodynamic parameters from UV melting experiments, a notable example being Marky and Breslauer's report.[147-151] This section explains the theory behind the methodology used in this chapter.

The two thermodynamic relations defining the variation of Gibbs energy of a reaction

$$\Delta_r G = -RT \ln K_{eq} \qquad \text{Equation 4.1}$$

and

$$\Delta_r G = \Delta_r S - T\Delta_r S \qquad \text{Equation 4.2}$$

can be combined to give the following van 't Hoff plot equation

$$\ln K_{eq} = \frac{\Delta_r S}{R} - \frac{\Delta_r H}{RT}$$

where K_{eq} is the equilibrium constant, R is the ideal constant of gases (8.3145 $J \cdot K^{-1} \cdot mol^{-1}$), T is the temperature in kelvins, $\Delta_r S$ is the entropy variation of the reaction

and $\Delta_r H$ is the enthalpy variation of the reaction. Rearranging this equation allows to obtain the following relation

$$\frac{1}{T} = \frac{\Delta_r S}{\Delta_r H} - \frac{R}{\Delta_r H} \cdot \ln K_{eq} \qquad \text{Equation 4.3}$$

that can be plotted as a line when graphed as $1/T$ *vs* $\ln K_{eq}$ (Figure 4.2). This allows the values of the variation of enthalpy and entropy to be obtained after the intercept ($q = \Delta_r S/\Delta_r H$) and the slope ($m = -R/\Delta_r H$) are known:

$$\Delta_r H = -\frac{R}{m} \quad \text{and} \quad \Delta_r S = q \cdot \Delta_r H = q \cdot \left(-\frac{R}{m}\right).$$

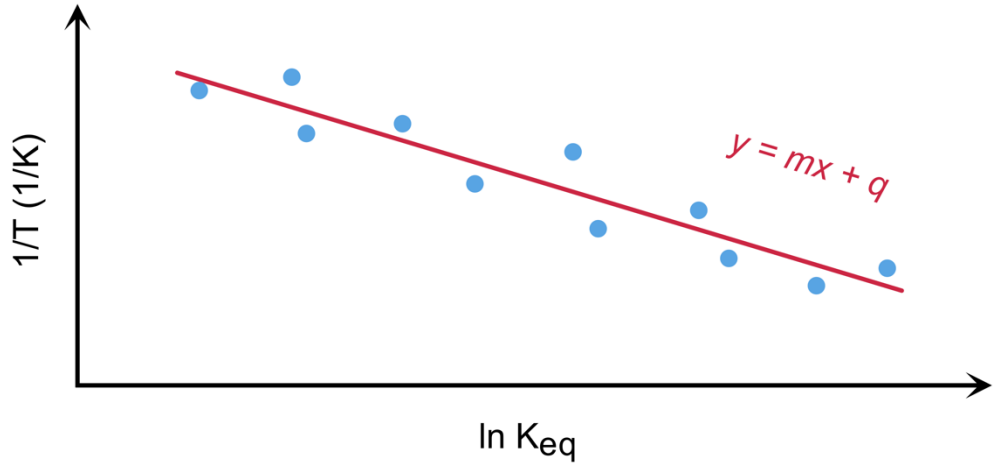


Figure 4.2 $1/T$ *vs* $\ln K_{eq}$ plot to determine thermodynamic parameters. A linear regression (red) is performed on the experimental data (light blue) to obtain intercept (q) and slope (m).

Given two non-self-complementary oligonucleotides A and B having the same concentration ($C_T/2$) with a total concentration of ssDNA C_T ($C_T = [A] + [B]$ with $[A] = [B] = C_T/2$), it is possible to write the process of the formation of a duplex AB as

$$A + B \rightleftharpoons AB.$$

At equilibrium, the reaction quotient is equal to the equilibrium constant which is

$$K_{eq} = \frac{[AB]}{[A]\cdot[B]} = \frac{\alpha \cdot {}^{C_T}/_{2}}{\left(\frac{C_T}{2} - \alpha \cdot \frac{C_T}{2}\right)^2} = \frac{2\cdot\alpha}{(1-\alpha)^2 \cdot C_T}$$ Equation 4.4

where α is the extent of association, *i.e.* the fraction of single strands in the duplex state. Since at the melting temperature (T_m) α = 1/2, Equation 4.4 can be rewritten as

$$K_{eq} = K_{T_m} = -\frac{4}{C_T}$$ Equation 4.5

and therefore, if Equation 4.5 is substituted in Equation 4.3 the following is obtained:

$$\frac{1}{T_m} = \frac{R}{\Delta_r H} \cdot \ln\frac{C_T}{4} + \frac{\Delta_r S}{\Delta_r H}.$$ Equation 4.6

Equation 4.6 can be fitted as a straight line when plotted as $1/T_m$ *vs* $\ln(C_T/4)$ (Figure 4.3). Once the intercept ($q = \Delta_r S/\Delta_r H$) and the slope ($m = R/\Delta_r H$) are measured, it is possible to obtain the values of the variation of enthalpy and entropy:

$$\Delta_r H = \frac{R}{m}$$ Equation 4.7

and

$$\Delta_r S = q \cdot \Delta_r H = q \cdot \left(\frac{R}{m}\right).$$ Equation 4.8

It possible to obtain the variation of Gibbs energy of the formation of a duplex ($\Delta_r G$) at a temperature T by substituting the values found with Equation 4.7 and Equation 4.8 in Equation 1.2:

$$\Delta_r G = \Delta_r S - T\Delta_r S.$$ Equation 4.9

The binding constant (K_b) of the process at the same temperature T can also be easily derived by combining the result obtained from Equation 4.9 with Equation 4.1:

$$K_b = e^{-\left(\Delta_r G / RT\right)}.$$ Equation 4.10

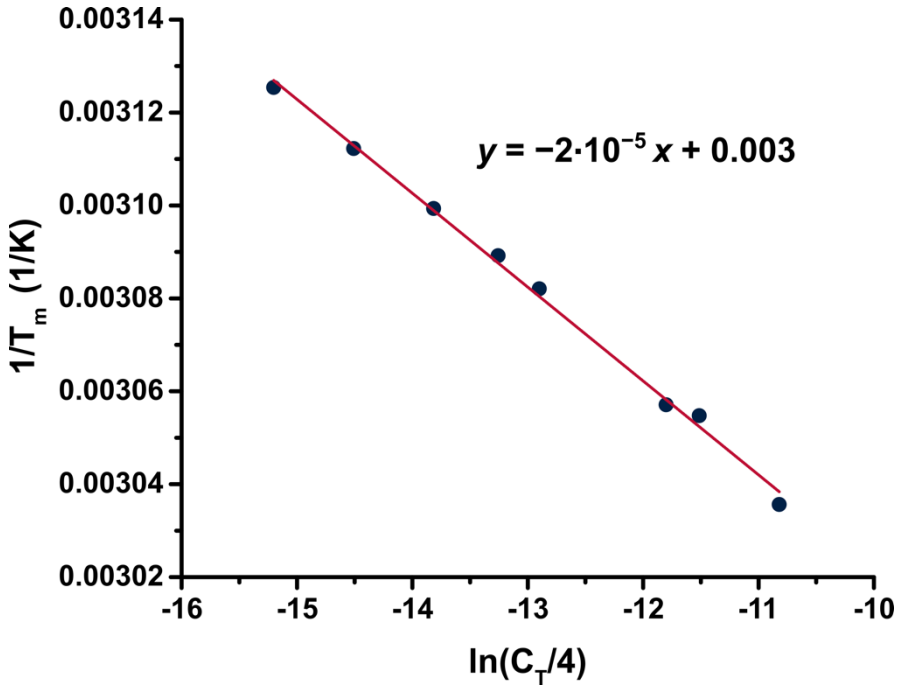


Figure 4.3 Example of a $1/T_m$ *vs* $\ln(C_T/4)$ plot to determine thermodynamic parameters. The linear regression is in red.

4.2.2 Crystallography of Oligonucleotides

Since the pioneering work of Watson, Crick and Franklin with the DNA double helix model proposed in 1953,[4] 3D structures of nucleic acids, very often obtained *via* crystallographic studies, have helped with understanding their biological functions. While Rosalind Franklin used a fibre diffraction method, macromolecular X-ray crystallography evolved preferring single-crystal methods. Nowadays, X-ray crystallography is the preferred technique for characterising macromolecules, also because medium-resolution structures (~2 Å) have an uncertainty in atomic position of only 0.1–0.2 Å.[152] The first protein structure (myoglobin at 6 Å resolution) was published in 1958,[153] while studies of oligonucleotides lagged behind for more than a decade. Although DNA was first crystallised by Giannoni who published his work in

1969,[154] it took a few more years to obtain the first crystal structures. The first oligonucleotide that was characterised this way was a tRNA in 1974,[155,156] while the first DNA structure was reported only in 1979.[157] There were, and in part there still are, many reasons behind this difference between nucleic acids and proteins.[158] DNA is intrinsically plastic and polymorphic,[13,159,160] making it far more difficult to obtain single monomorphic crystals. In fact, it was shown that in a single crystal there can be more DNA forms coexisting together.[161] Moreover, the same oligonucleotide sequence can be used to produce different structures and crystal packing organisations.[162] It is important to highlight that, in contrast with proteins, DNA can undergo significant structural alterations because the packing environment can induce or stabilise them,[163,164] for example the crystallisation conditions could induce the high salt B-Z transition.[165,166]

Despite all these crystallisation problems, over the years X-ray crystallography has been used to improve the understanding of the mechanisms of many biological processes such as DNA damage and repair, ribozyme catalysis, translation, transcription and replication.[167-174] The level of detail obtainable with X-ray crystallography of nucleic acids has also proved to be useful for structure-based drug design.[175,176]

Unfortunately, although often being the rate-limiting step, crystallisation is not the only problem that has to be overcome in X-ray crystallography. In a diffraction experiment, the detector acts as a photon counter and the data collected can be used to obtain information only on the intensity and the pattern of the diffraction. During the experiment no data about the phases of the different diffractions can be preserved, which gives rise to the so called “phase problem”, as both amplitudes and phases are

necessary in order to reconstruct the electron density map from a diffraction experiment.[177] It is possible to recover phase information in several ways and one of the most used methodologies is Molecular Replacement (MR).[177,178] This can be used to solve a crystal structure when there is already an appropriate model for a reasonable proportion of it. Of course, the more similar structures that have been solved previously the higher the probability of finding one that is related to the crystal under study, therefore rendering MR more powerful over the years. When there is no structure close enough available it is necessary to use experimental phasing to obtain phase information from the experimental data. Different methodologies can be used such as single wavelength anomalous dispersion/diffraction (SAD), and multiple wavelength anomalous dispersion/diffraction (MAD).[177,179-186] Both methods depend on the presence of usually heavy atoms in the macromolecule, which can generate an anomalous scattering that can be used to solve the phase problem. Common solutions, especially for nucleic acids, are the use of halogen atoms, such as 5-Br modifications (as 5-BrdU). Alternatively, the use of sulfur or phosphorus present in the backbone can be used for native SAD which, given the weak anomalous scattering of those atoms, requires highly redundant data.[179,181-186]

4.3 Materials and Methods

4.3.1 Oligonucleotides Synthesis, Purification and Analysis

The oligonucleotides were provided by ATDBio Ltd., where they were synthesised and purified. The sequences for crystallography were obtained already mixed as a duplex.

Analysis

The oligonucleotides were characterised by HPLC-MS using an Acquity UPLC system with a BEH C18 1.7 µm column (Waters) and a Bruker micro-TOF mass spectrometer (ESI− mode). All the oligonucleotides were found with an experimental mass corresponding to the expected mass (±2 Da, except when stated otherwise). Table 2.1 presents the MS results for some representative sequences.

Table 4.1 Representative oligonucleotide sequences and MS analysis

Code	Oligonucleotide sequence (5′–3′)	MS (Da)	
		Calc.	Found
ODN-23 (0 pyrenes)	CAGTATATATAGAC	4171	4270
ODN-24 (1 pyrene)	CAGTATA**X**ATAGAC	4580	4580
ODN-25 (2 pyrenes near)	CAGTA**X**A**X**ATAGAC	4890	4889
ODN-26 (2 pyrenes far)	CAG**X**ATATA**X**AGAC	4890	4889
ODN-27 (4 pyrenes)	CAG**X**A**X**A**X**A**X**AGAC	5508	5507

Note: Mass spectra of modified oligonucleotides were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES− mode. X = U-pyrene.

4.3.2 UV Melting and Thermodynamics

4.3.2.1 General UV Melting Procedure

UV melting samples were prepared in 5 mM sodium phosphate buffer with 135 mM KCl, 1 mM $MgCl_2$ at pH 7.2. Prior to use, the buffer was filtered with Kinesis regenerated cellulose 13 mm, 0.45 µm syringe filters and degassed by sonication for 10 min. Concentration of both target and probe (labelled) oligonucleotides was 2 µM. Melting analysis was performed on a Cary 4000 UV-Vis spectrometer fitted with a Cary temperature controller. Samples were analysed in 1.2 mL cuvette (Hellma Precision cells made of synthetic quartz glass, QG, 1 cm path length). Before melting curves were recorded, the samples were melted and then slowly annealed. UV melting was carried out from 18 °C to 84 °C with ramping of 1 °C/min. UV absorbance at 260 nm was

recorded every 0.1 °C. The melting temperatures (T_m) were calculated using Cary Win UV thermal application software, the values reported are an average of the four melting curves. The following table summarises the procedure used:

Stage	Data interval (°C)	Rate (°C/min)	End temperature (°C)	Hold time (min)
Melting 0	1.00	10	84	2
Annealing 1	0.10	1	18	2
Melting 1	0.10	1	84	2
Annealing 2	0.10	1	18	2
Melting 2	0.10	1	84	2
Annealing 3	0.10	1	18	2
Melting 3	0.10	1	84	2
Annealing 4	0.10	1	18	2
Melting 4	0.10	1	84	2
Annealing 5	1.00	10	20	0

4.3.2.2 UV melting for thermodynamics studies

The samples for the thermodynamics UV melting experiments were studied as described in section 4.3.2.1, with a few differences. The concentration of both target and probe (labelled) oligonucleotides was varied between 0.5–40 μM (data were collected at 0.5 μM, 1 μM, 2 μM, 3.5 μM, 5 μM, 15 μM, 20 μM, 40 μM). The procedure to collect data with the Cary 4000 changed slightly as follows:

Stage	Data interval (°C)	Rate (°C/min)	End temperature (°C)	Hold time (min)
Melting 0	1.00	10	84	2
Annealing 1	0.10	0.5	18	2
Melting 1	0.10	0.5	84	2
Annealing 2	0.10	0.5	18	2
Melting 2	0.10	0.5	84	2
Annealing 3	0.10	0.5	18	2
Melting 3	0.10	0.5	84	2
Annealing 4	0.10	0.5	18	2
Melting 4	0.10	0.5	84	2
Annealing 5	1.00	10	20	0

In order to derive $\Delta_r S$, $\Delta_r H$, $\Delta_r G$ and K_b the various melting temperatures were plotted against the concentration of oligonucleotides as explained in section 4.2.1 and shown in Figure 4.3. A linear regression was performed with Equation 4.6, allowing so to obtain $\Delta_r S$ and $\Delta_r H$ with Equation 4.7 and Equation 4.8. $\Delta_r G$ and the binding constant (K_b) were obtained at 37 °C using Equation 4.9 and Equation 4.10 respectively.

4.3.3 Error Calculations

T_m, $\Delta_r S$ and $\Delta_r H$ errors were calculated using the confidence intervals derivable with a Student's t-test with 95 % confidence and the appropriate degrees of freedom. For ΔT_m, $\Delta_r G$, K_b and K_d the errors were calculated using the normal error propagation formula without the covariance factor. In graphs errors are represented by error bars.

4.3.4 CD Spectroscopy

CD spectra were collected on a Chirascan Plus spectropolarimeter (Applied Photophysics) using a quartz cuvette with a 10 mm or a 2 mm path length. The oligonucleotides were dissolved in a 5 mM sodium phosphate buffer with 135 mM KCl, 1 mM $MgCl_2$ at pH 7.2, with a final duplex concentration of 4 µM or 10 µM. Data were collected over the range of 200–420 nm at 25 °C. Spectra were recorded with a step size of 0.5 nm, a response time of 0.8 s and a bandwidth of 1 nm. A buffer baseline was subtracted from each spectrum. The trace plotted are the average of four identical scans with a further smoothing using the Savitzky-Golay method available within the instrument software (smoothing window = 7).

4.3.5 X-Ray Crystallography

4.3.5.1 Crystallisation of Oligonucleotides

All the crystals were obtained using the sitting-drop vapour diffusion method. Initial trials were performed using the Natrix HT (Hampton Research) and Nucleix (Qiagen) screens, dispensed into 96-well sitting-drop crystallisation plates (CrystalQuick, Greiner) using a high-throughput dispenser (Hydra 96, Robbins Scientific). In each drop, 100 nL of the crystallisation solution were mixed with 100 nL of a 1 mM aqueous DNA solution using a crystallisation robot (Cartesian Microsys). The plates were sealed and incubated at 20 °C. The most promising conditions were further manually optimised using 24-well sitting-drop crystallisation plates (Chryschem, Hampton Research), in which 1 µL of the 250 µL well solution (see Table 4.4) was combined with 1 µL of a 0.6 mM DNA solution. The plates were sealed and incubated at 20 °C. Typically, crystals appeared within a week and were collected within 3 weeks.

4.3.5.2 Diffraction Data Collection

Crystals were flash-frozen in liquid nitrogen with no further cryoprotection and all diffraction data were collected at 100 K at the Diamond Light Source synchrotron science facility (Harwell) on beamlines I02 and I04 using Pilatus 6M hybrid pixel array detectors. To minimise radiation damage during data collection, when possible, data were collected using the line scan mode and low transmission (10 % or 15 % attenuated beam). Table 4.2 reports other relevant information about data collection.

Table 4.2 Parameters used for data collection.

	CRY-1	CRY-2Br
Beamline	I04	I02
Type of data	P-SAD	Br-SAD
Detector	Pilatus 6M	Pilatus 6M
Detector distance (mm)	209.16	374.07
Images collected	9600	900
Wavelength (Å)	1.5400	0.9196
Energy (eV)	8051	13482
Oscillation angle per image: $\Delta\Phi_{image}$ (°)	0.15	0.4
Total oscillation: $\Delta\Phi_{total}$ (°)	1440	360
Exposure time (s)	0.05	0.20
Transmission (%)	10 (first 4800 frames) 15 (last 4800 frames)	10
Flux (ph/s)	*	$4.53 \cdot 10^{10}$

*because of a software bug the value was not recorded.

4.3.5.3 Data Processing, Phase Determination, Model Building and Refinement

The data obtained were indexed and scaled using XDS[187] and then Aimless,[188] in order to obtain the space group, the cell dimensions, the resolution and to estimate the quality of the data. Metthews_coeff[188] was used to estimate the solvent content and the number of duplexes per asymmetric unit. The experimental phases were determined from the SAD data sets using the SHELX software suite.[189] The models were manually built inside the electron density maps produced by SHELX using COOT.[190] In the case of CRY-1, before building the macromolecule in COOT, molecular replacement was performed with Phaser,[191] using a computational short B-DNA without U-pyrene as a search model. A number of cycles of refinement were performed with PHENIX[192] to obtain the final solution. All the models were rendered with PyMOL (Molecular Graphics System, Version 1.5.0.3 Schrödinger, LLC).

4.4 UV Melting and Thermodynamics

The effects of multiple additions of U-pyrene on the stability of dsDNA were studied using the sequences in Figure 4.4 A. A single addition of U-pyrene increased the melting temperature of the resulting duplex by 8.2 °C compared to the unmodified one (Figure 4.4 B). A second U-pyrene in the same sequence had a smaller impact and moreover, the stabilisation was dependent on the distance between the two modifications. Considering the size of U-pyrene it was predictable that when the two pyrene monomers were closer to each other the stabilisation effect was smaller (+13.1 °C) than when they were further apart (+14.4 °C). When the same sequence was modified with a total of four U-pyrenes, there was a further increase in the melting temperature (+17.0 °C). However, it is possible to notice from Figure 4.4 B that there was an apparent plateauing effect, where each further addition of U-pyrene had a smaller beneficial effect on the T_m.

A

Entry	X = U-pyrene	
1	3'-GTCATATATATCTG-5' 5'-CAG**T**A**T**A**T**A**T**AGAC-3'	0 pyrenes
2	3'-GTCATATATATCTG-5' 5'-CAGTATA**X**ATAGAC-3'	1 pyrene
3	3'-GTCATATATATCTG-5' 5'-CAGTA**X**A**X**ATAGAC-3'	2 pyrenes near
4	3'-GTCATATATATCTG-5' 5'-CAG**X**ATATA**X**AGAC-3'	2 pyrenes far
5	3'-GTCATATATATCTG-5' 5'-CAG**X**A**X**A**X**A**X**AGAC-3'	4 pyrenes

B

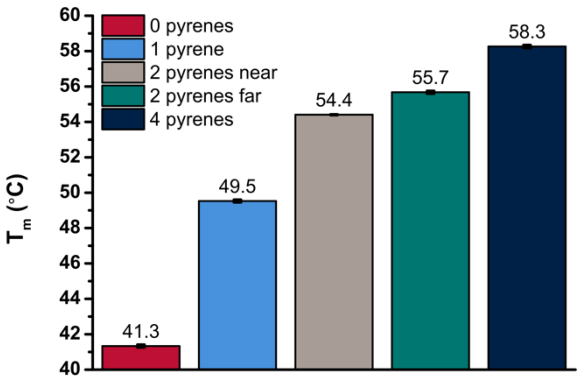


Figure 4.4 A) Sequences used for the study of U-pyrene. X = U-pyrene (in orange). The four T of the control sequence (0 pyrenes) that are substituted by U-pyrene in the other sequences are in bold. **B)** Thermal stability of the various duplexes. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 135 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2 μM).

In order to understand the effects of multiple additions of U-pyrene, a more comprehensive thermodynamic study was performed, according to the principles explained in section 4.2.1. Melting temperatures were measured varying the concentration of the duplexes from 0.5 µM to 40 µM. The results were then plotted as in Figure 4.3 and a linear regression was performed with Equation 4.6, allowing $\Delta_r S$ and $\Delta_r H$ to be obtained. $\Delta_r G$ and the binding constant (K_b) were obtained at 37 °C using Equation 4.9 and Equation 4.10 respectively. The dissociation constant (K_d) was obtained as $1/K_b$. Table 4.3 presents the values of all the thermodynamic parameters.

Table 4.3 Thermodynamic parameters obtained for the duplexes in Figure 4.4 A. $\Delta_r G$, K_b and K_d were calculated at 37 °C. Errors are reported in brackets. The number of significant figures reported is limited by the first significant figure of the error.

Duplex	$\Delta_r H$ (kJ/mol)	$\Delta_r S$ (kJ/K·mol)	$\Delta_r G$ (kJ/mol)	K_b (1/M)	K_d (M)
0 pyrenes	−403 (5)	−1.17 (0.017)	−41.30 (0.05)	$9.0 \cdot 10^{6}$ (0.17)	$1.11 \cdot 10^{-7}$ (0.02)
1 pyrene	−417 (6)	−1.18 (0.018)	−51.9 (0.3)	$5.5 \cdot 10^{8}$ (0.6)	$1.8 \cdot 10^{-9}$ (0.2)
2 pyrenes near	$-4.2 \cdot 10^{2}$ (0.11)	−1.16 (0.03)	−57.9 (0.6)	$6 \cdot 10^{9}$ (1.4)	$1.8 \cdot 10^{-10}$ (0.4)
2 pyrenes far	$-4.2 \cdot 10^{2}$ (0.10)	−1.15 (0.03)	−59.3 (0.6)	$10 \cdot 10^{9}$ (2)	$1.0 \cdot 10^{-10}$ (0.2)
4 pyrenes	−504 (9)	−1.41 (0.03)	−68.0 (0.6)	$2.9 \cdot 10^{11}$ (0.4)	$3.5 \cdot 10^{-12}$ (0.8)

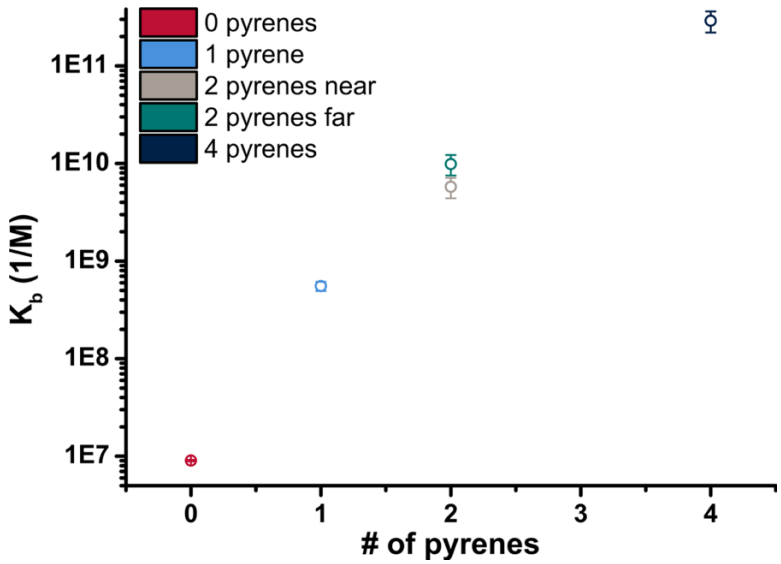


Figure 4.5 Binding constants at 37 °C calculated for the duplexes in Figure 4.4 A, plotted against the number of U-pyrenes in the sequences. Note that the K_b scale is logarithmic.

Unsurprisingly, the process of duplex formation was enthalpy driven as the resulting reduction in entropy was unfavourable. The bigger decrease in enthalpy with a higher number of U-pyrenes may be due to the hydrophobic pyrene moieties being better shielded from the aqueous solvent in a duplex environment. Interestingly, both for $\Delta_r S$ and $\Delta_r H$, the values did not vary drastically when zero to two U-pyrenes were present. However, having four pyrenes changed the values significantly (Table 4.3). The most interesting feature of these results was the calculation of $\Delta_r G$ and the binding constant at the biologically relevant temperature of 37 °C. Figure 4.5 highlights the dramatic effects of the addition of U-pyrenes on the binding affinity: going from zero pyrenes to four pyrenes increased the binding constant of this 14-mer duplex to its complementary strand 10000-fold. Therefore, the use of U-pyrene provides an easy method to obtain a 14-mer probe with a remarkable pM affinity for its target. These results also underline the limited information single UV melting experiments provide and the need for more thorough thermodynamic studies to understand the selected systems.

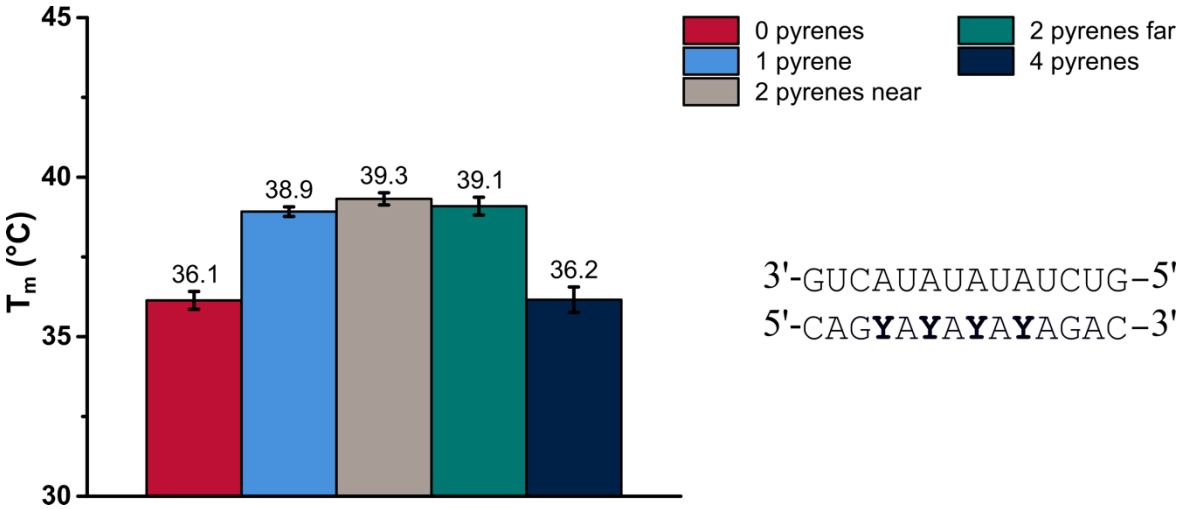


Figure 4.6 Thermal stability of the probes presented in Figure 4.4 A with RNA targets. Y = T or U-pyrene. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 135 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2 μM).

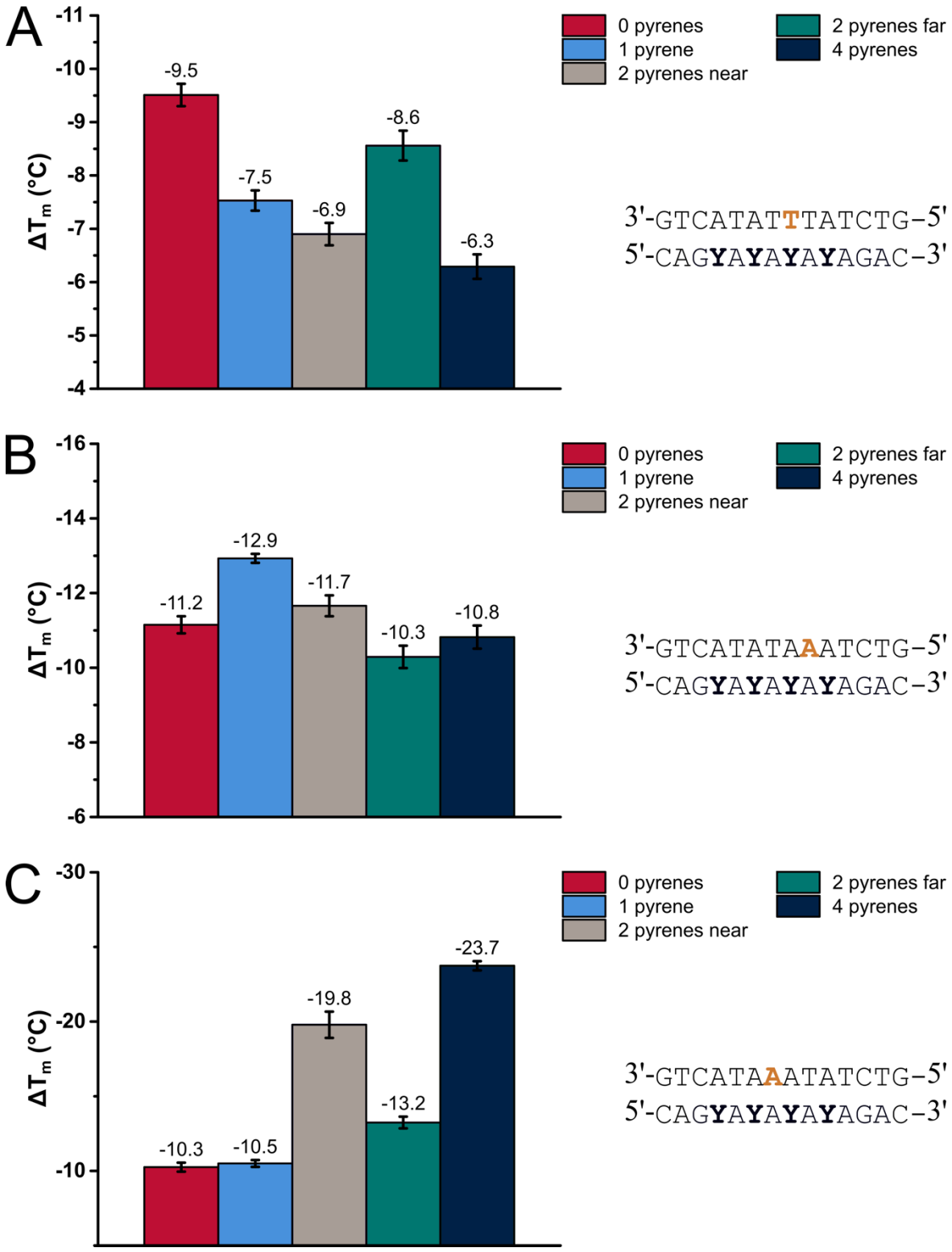


Figure 4.7 Detection of single mismatches with U-pyrene. The plots report the ΔT_m of the mismatched duplexes compared to the fully complementary ones (see Figure 4.4, also for the sequences of the probes). Y = T or U-pyrene. The single point mutations are highlighted in orange. **A)** A->T mutation opposite to the only U-pyrene present in the 1 pyrene sequence (light blue). **B)** T->A mutation 3′ to the U-pyrene present in the 1 pyrene sequence. **C)** T->A mutation 5′ to the U-pyrene present in the 1 pyrene sequence. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 135 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2 μM).

This kind of high affinity probe would be extremely beneficial for antisense oligonucleotides. Therefore, the same probes reported in Figure 4.4 were tested against their RNA target (Figure 4.6). Unfortunately, the additive stabilising effect of U-pyrene was lost, as with one or two pyrenes the melting temperature increased by only about 3 °C, without any significant difference between the three different probes used. Even more surprising was the total lack of any stabilising effect of the probe with four pyrenes, as its T_m was found to be identical to the unmodified control sequence. These negative results with RNA could be explained by the fact the duplex would be in its A-form and not its B-form as with a DNA target. As an A-type duplex is compressed with a significantly reduced distance between the nucleobases,[11,12] this conformation would affect the intercalating capacity of the pyrene moiety, preventing it from stabilising the resulting duplex.

As these results preclude U-pyrene's use for antisense purposes, the high affinity with complementary DNA suggested potential applications as mismatch probes. Therefore, initial studies were performed to determine its ability to detect single base mismatches. Unfortunately, the results were mostly underwhelming (Figure 4.7). For example, when only one U-pyrene was present (light blue in Figure 4.7), the capacity of detecting mismatches was overall not improved compared to the control sequence (in red). The U-pyrene mismatched duplex was 2 °C less destabilised than the control duplex when the mutation was opposite to U-pyrene, it was 1.7 °C more destabilised than the control when the mutation was 3′ to U-pyrene and it was undistinguishable from the control when the same mutation was on the 5′-side of U-pyrene. In most of the cases studied, U-pyrene probes resulted in a reduced destabilisation or at best a slightly improved destabilisation compared to the unmodified control. Considering the

data collected, it seems like the number and the position of these U-pyrene monomers needs to be optimised to reduce unfavourable effects, which impairs an easy probe design. There was only one exception to this trend, which was when a mismatch was between two U-pyrenes (Figure 4.7 C, in grey and dark blue). Only in this particular scenario the destabilisation of the modified probes was doubled compared to the control. This is probably linked to the pyrene intercalating moieties. If the π-stacking in between such pyrenes is disrupted by a mismatch, their intercalation becomes more difficult as well, destabilising the duplex formed.

In order to verify the intercalation mode of U-pyrene and to study its influence on dsDNA structure, two duplexes, either made of non-self-complementary (CRY-1) or self-complementary (CRY-2) oligonucleotides, were chosen for X-ray crystallographic studies. Both CRY-1 and CRY-2 sequences were chosen as their unmodified counterparts (CRY-1C and CRY-2C) have been crystallised and studied before.[193-195]

The thermal stability of CRY-1 and CRY-2 was measured and compared to the corresponding control duplexes (Figure 4.8). Both modified duplexes were more stable than the unmodified ones, however, while CRY-2 had a T_m 11.3 °C higher than the control (comparable to the values observed for the sequences with two U-pyrenes of the thermodynamic studies, +13.1 °C, Figure 4.4 B), CRY-1 was only 2.9 °C more stable than the control. This was a considerably lower stabilisation effect than the one obtained with one U-pyrene for the thermodynamic studies (+8.2 °C, Figure 4.4 B). This is probably due to a U-pyrene sequence dependence, as the oligonucleotides used present different nucleobases neighbouring the modification.

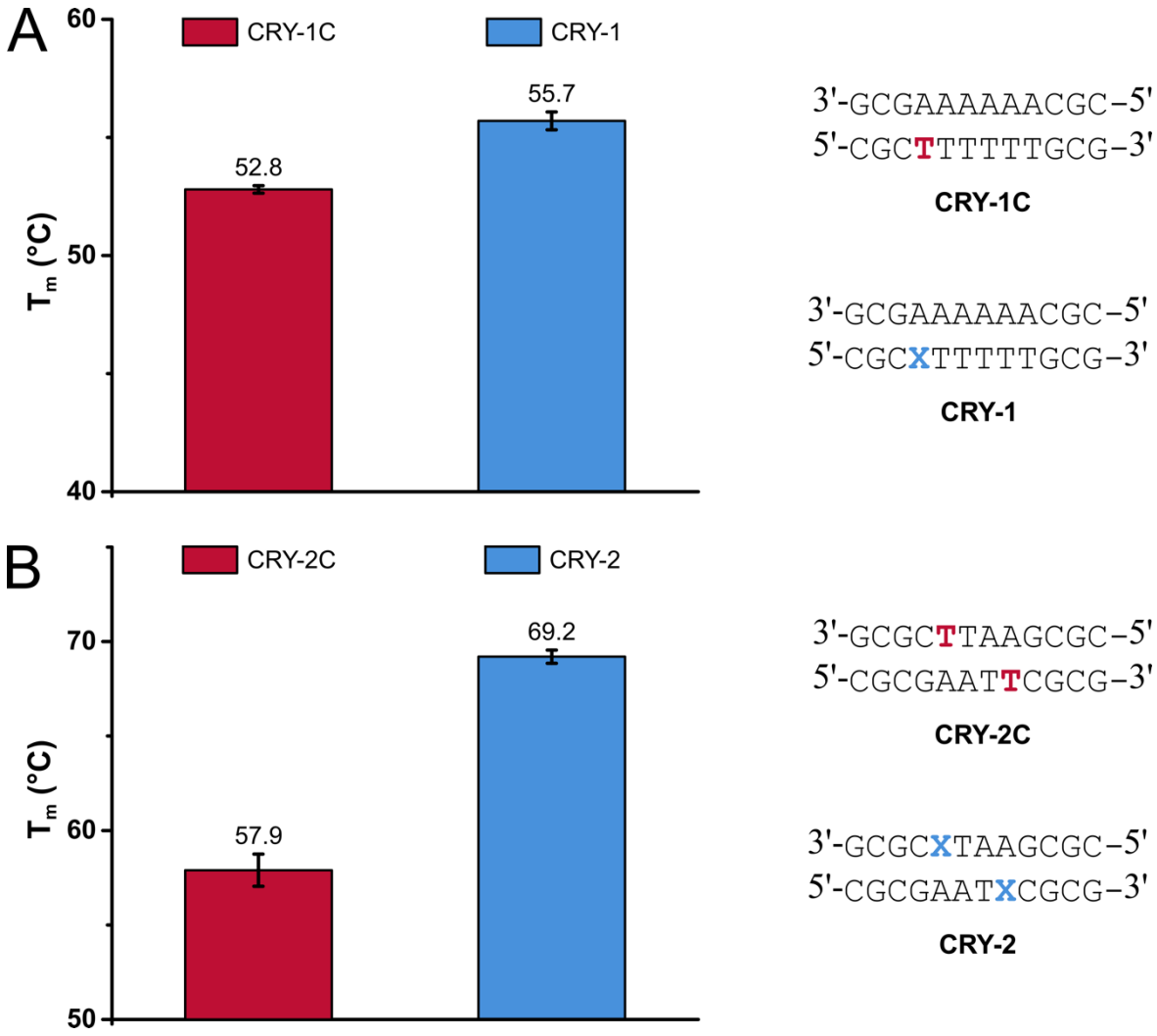


Figure 4.8 Thermal stability of the sequences used for crystallisation studies. X = U-pyrene. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 135 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 2 μM).

4.5 Circular Dichroism

U-pyrene showed a considerable capacity for increasing the binding affinity of oligonucleotide probes to their targets, with a higher number of modifications giving higher duplex stability. Therefore, as a preface to crystallography, circular dichroism (CD) spectroscopy was performed to confirm if the dsDNA conformation was altered by U-pyrene and to provide preliminary evidence of pyrene intercalation. CD demonstrated that all five duplexes under examination, whether with or without U-pyrene incorporated, maintained the overall shape of a typical B-form dsDNA (Figure 4.9).[196,197] This was very promising as, even four U-pyrenes did not significantly distort the unmodified duplex B-conformation. The two peaks present in the 330–360 nm region are pyrene adsorption peaks and they are associated with transitions along the major axis of pyrene.[198,199] Thanks to the work of Lyng *et al.* it was possible not only to confirm the intercalation of pyrene inside dsDNA, but also to speculate the orientation of pyrene inside the duplex.[18] As pyrene peaks are present on the CD spectra, it means that the achiral pyrene is in a chiral environment, as would happen when intercalating inside a duplex. Moreover, since these two peaks are negative and they are connected to electronic transitions along the major axis of pyrene, it means that pyrene is probably oriented inside the duplex with its major axis parallel to the axis connecting the nucleobases of a base pair. The intensity of the pyrene peaks should be connected to the amount of pyrene intercalating. From Figure 4.9 it is possible to notice that the peaks for the sequence with four pyrenes are more intense than for the sequence with one pyrene. However, because of the large variation between the two cases with two pyrenes, it is clear that there is no unambiguous linear

relationship between the number of U-pyrenes and the intensity of those peaks, suggesting that the system under study might be significantly dynamic.

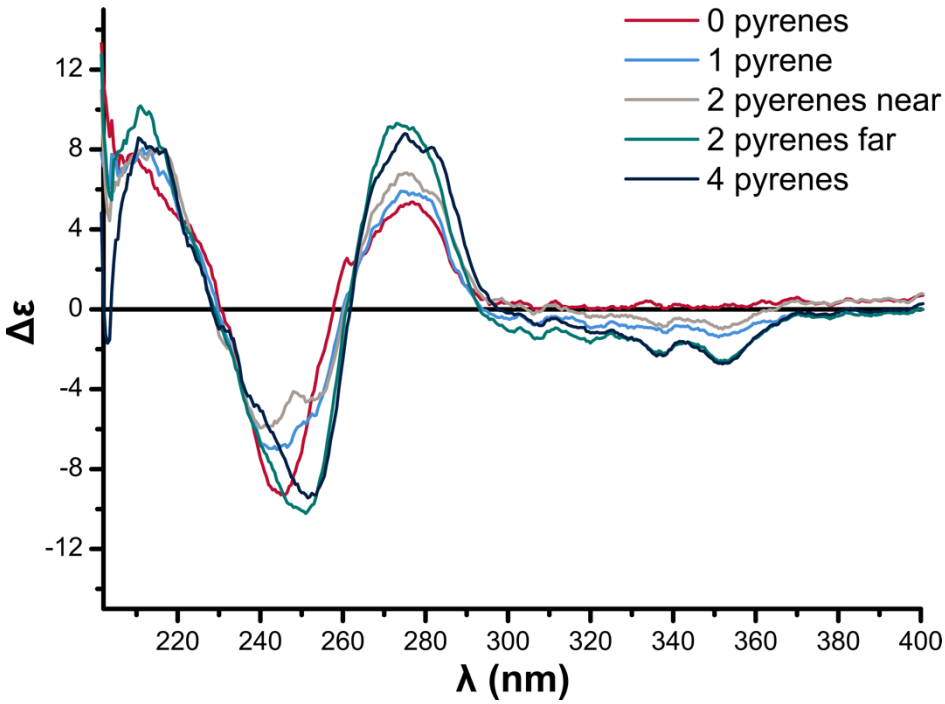


Figure 4.9 CD spectra of the duplexes presented in Figure 4.4 A. The spectra were recorded from 200 nm to 420 nm. The signals were smoothed using the Savitzky-Golay method available within the instrument software. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 135 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 4 μM).

Similar CD experiments were performed for the duplexes used for the crystallographic studies (Figure 4.10). Both CRY-1 and CRY-2 showed spectra typical of a B-form duplex, exactly as their corresponding controls. Also the negative pyrene peaks, indicating its intercalation, were present, albeit being faint in both cases.

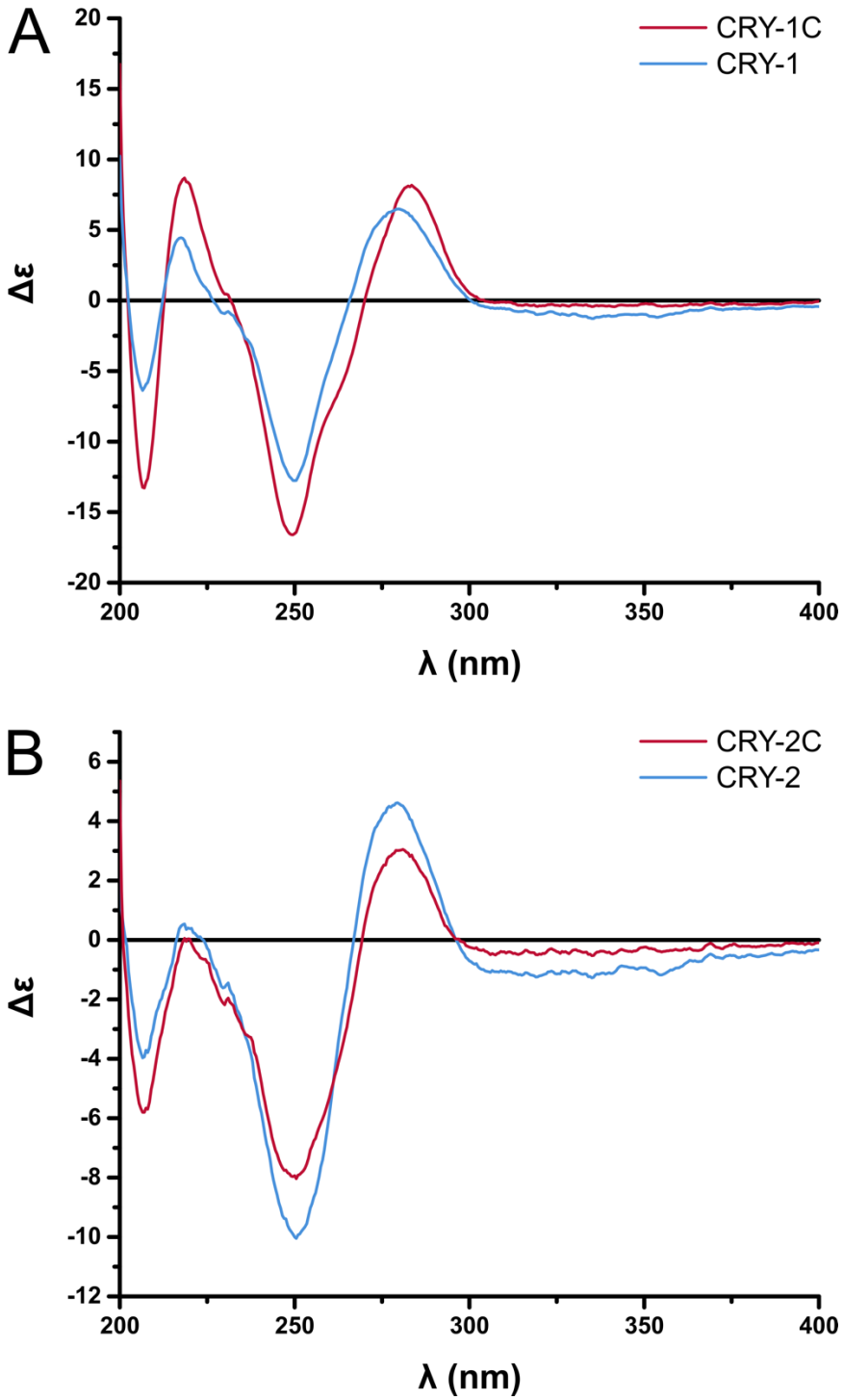


Figure 4.10 CD spectra of the duplexes presented in Figure 4.8. The spectra were recorded from 200 nm to 420 nm. The signals were smoothed using the Savitzky-Golay method available within the instrument software. The experiments were performed in 5 mM sodium phosphate buffer (pH 7.2, containing 135 mM KCl, 1 mM $MgCl_2$, oligonucleotide concentration = 10 μM).

4.6 X-Ray Crystallography

4.6.1 Preface

Before starting with the description of the crystallisation and the solution of the crystal structures, it is important to highlight a small sequence change that proved to be necessary. While it was possible to solve the structure of CRY-1 using the anomalous signal from the naturally occurring phosphorus in DNA (phosphorus SAD), the same was not possible with CRY-2, as all the crystals obtained were more sensitive to radiation damage caused by prolonged exposure to X-ray, making the collection of highly redundant data impossible. Therefore, the sequence of CRY-2 was modified by introducing a 5-bromo-2′-deoxyuridine (5-BrdU) instead of a T to enable the use of bromine SAD, obtaining CRY-2Br (5′-CGCGAA**YX**CGCG, Y = 5-BrdU, X = U-pyrene).

4.6.2 Crystallisation

In order to understand which conditions to use to obtain crystals of CRY-1 and CRY-2Br, initial screens were performed using the Natrix HT (Hampton Research) and Nucleix (Qiagen) screens in 96-well sitting-drop crystallisation plates. The most promising conditions were then optimised using 24-well sitting-drop crystallisation plates. The typical initial optimisation screens varied the DNA concentration on one axis and the concentration of the precipitant on the other. Subsequent tests varied the pH and the salts concentration. The fundamental manual screen, which allowed crystallographically suitable crystals to be obtained, had the concentration of spermine varying on one axis, while the concentration of the divalent cation varied on the other. Interestingly, this gave crystals of different forms. Table 4.4 reports the final

conditions used, while Figure 4.11 shows the photographs of the crystals from which it was possible to solve the structure.

Table 4.4 Crystallisation conditions for CRY-1 and CRY-2Br used to obtain the crystals in Figure 4.11.

Duplex	[Duplex] (mM)	Buffer	Salts and Additives	Precipitant
CRY-1	0.6	0.0342 M sodium succinate pH 5.6	20 mM magnesium chloride 0.5 mM spermine tetrahydrochloride	3 M ammonium sulphate
CRY-2Br	0.6	0.04 M sodium cacodylate pH 7	80 mM potassium chloride 0.5 mM barium chloride 12 mM spermine tetrahydrochloride	40 % v/v (±)-2-Methyl-2,4-pentanediol (MPD)

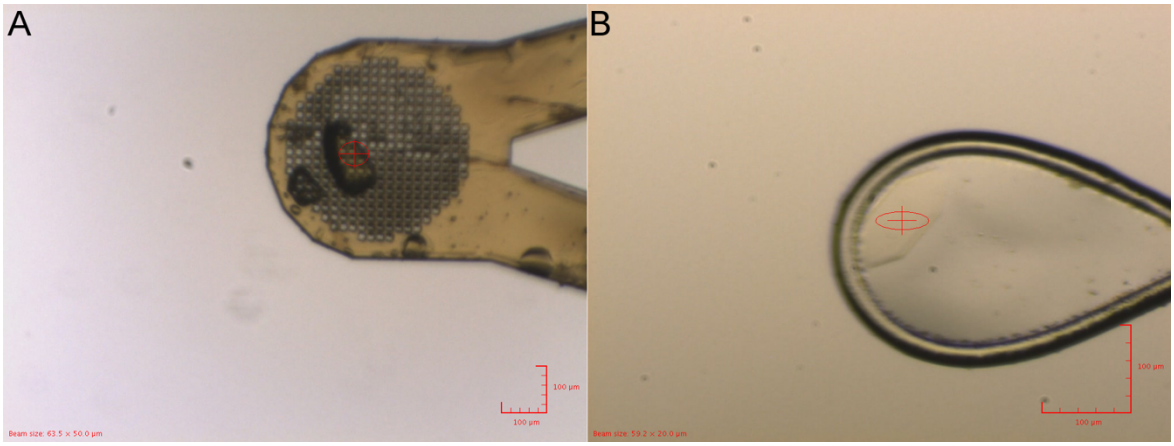


Figure 4.11 Photographs of the crystals used for X-ray diffraction. **A)** CRY-1. **B)** CRY-2Br

4.6.3 Data Collection and Crystallographic Statistics

X-ray diffraction of the two crystals presented in Figure 4.11 was obtained at the Diamond Light Source synchrotron science facility (Harwell) on beamlines I02 and I04. The strategy selected to solve the structure of CRY-1 involved the use of phosphorous SAD (P-SAD) and as the anomalous signal from phosphorus is weak, highly redundant data were necessary for phasing. In order to obtain this, it is necessary to record multiple data sets on the same crystal, if stable enough to radiation damage, or on multiple crystals, if they are perfectly isomorphous.[182] Unfortunately, it was not

possible to obtain different crystals perfectly isomorphous with each other. Therefore, it was fundamental to screen for a single crystal able to withstand a significant amount of radiation without loss of diffraction power to collect data having a high multiplicity. The stronger anomalous signal deriving from bromine was used to solve the structure of CRY-2Br (bromine SAD, Br-SAD) and therefore the same redundancy of data was not required. This difference can be exemplified by the total oscillation angle and the total exposure time: 1440°, 480 s for CRY-1 and only 360°, 180 s for CRY-2Br.

Table 4.5 Characteristics of the crystals under study.

Duplex	Space group	a (°)	b (°)	c (°)	α (Å)	β (Å)	γ (Å)	Complex per AU*	Solvent content (%)	Resolution (Å)
CRY-1	$P4_1$	42.5	42.5	54.9	90	90	90	1	62.4	1.98
CRY-2Br	$C222_1$	44.7	52.7	70.7	90	90	90	1	67.9	1.60

*AU = asymmetric unit

The data obtained were indexed and scaled using XDS[187] and then Aimless,[188] in order to obtain the space group, the cell dimensions and the resolution. Metthews_coeff[188] was used to estimate the solvent content and the number of duplexes per asymmetric unit. The results are reported in Table 4.5. Table 4.6 reports other relevant crystallographic statistics obtained from Aimless.

Table 4.6 Summary of crystallographic data statistics as reported by Aimless.

	CRY-1	CRY-2Br
Low resolution limit (Å)	42.53 (2.02)	35.37 (1.63)
High resolution limit (Å)	1.98 (1.98)	1.60 (1.6)
R_{meas}	0.178 (1.418)	0.087 (0.508)
Number of measured reflections	663577 (30286)	131250 (2826)
Number of unique reflections	6907 (462)	10979 (408)
I/σ	37.4 (4.2)	17.0 (2.7)
Completeness (%)	99.1 (89.0)	96.6 (75.8)
Multiplicity	96.1 (65.6)	12.0 (6.9)
Anomalous completeness (%)	98.8 (84.9)	96.3 (74.5)
Anomalous multiplicity	47.6 (32.5)	6.3 (3.6)

The values for the highest resolution range are given in parenthesis.
R_{meas} is a redundancy independent R-factor as defined by Diedrichs and Karplus.[200]

4.6.4 Experimental Phasing, Model Building and Refinement

For both CRY-1 and CRY-2Br the experimental phases were determined from the SAD data sets using the SHELX software suite.[189] After truncating the data to have a significant anomalous signal (see Figure 4.12), the software gave the position of the heavy atoms (P or Br) and the electron density map, which was extended again to the highest resolution possible by SHELXE.

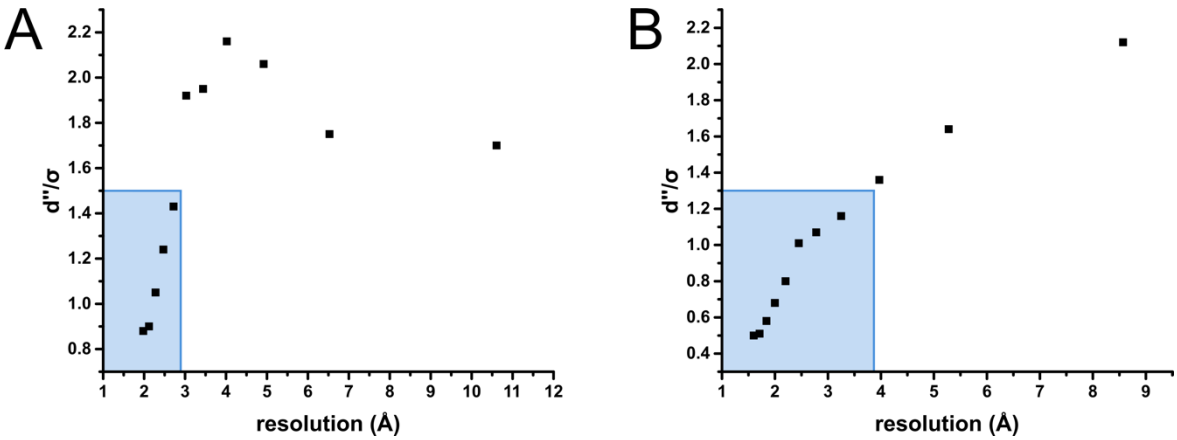


Figure 4.12 Plot of anomalous signal (d''/σ) *vs* resolution as derived from SHELXC. In the light blue box are the data truncated as having $d''/\sigma < 1.3$. **A)** CRY-1. **B)** CRY-2Br.

The models were manually built inside the electron density maps produced by SHELX using COOT.[190] In the case of CRY-1, before building the macromolecule in COOT, molecular replacement was performed with Phaser,[191] using a computational short B-DNA without U-pyrene as a search model. This gave a partial solution with a log likelihood gain function score of 1277 and a Z-score of 5.0 for the rotation function and 20.6 for the translation function. The rest of the structure of CRY-1 was built manually in COOT. For both CRY-1 and CRY-2, a number of cycles of refinement were performed with PHENIX,[192] also employing its simulated annealing function to increase the movement freedom during the rearrangement of the model, allowing the system to escape local minima. U-pyrene was added to the models later in the refinement cycle, when its presence became obvious from the electron density maps. Table 4.7 reports the final refinement statistics for both crystals.

Table 4.7 Refinement statistics for CRY-1 and CRY-2 obtained from PHENIX.

	CRY-1	CRY-2Br
Resolution (Å)	42.53–1.98 (2.49–1.98)	35.37–1.60 (1.76–1.60)
Number of reflections	6780	10927
Total number of non-hydrogen DNA atoms	559	646
Total number of water molecules	50	81
R_{work}	0.2163 (0.2392)	0.2098 (0.2810)
R_{free}	0.2629 (0.3058)	0.2547 (0.3348)
r.m.s deviations for bond lengths (Å)	0.013	0.013
r.m.s deviations for bond angles (°)	1.380	1.680
***B* factor (Å²)**	36.38	18.49
r.m.s deviations for covalently linked atoms (Å²)	5.29	3.33

The values for the highest resolution range are given in parenthesis.
r.m.s = root mean square.

4.6.5 CRY-1

Unfortunately, the structure obtained for the CRY-1 crystal (Figure 4.13 A and Figure 4.14 A) was a crystallographic artefact. As it will be clear from the next few pages, the particular contacts between the different duplexes in the packing of the crystal are very unlikely to occur in solution. For most biological, biochemical and biophysical scenarios, short dsDNAs are usually considerably diluted. For example, all the other experiments were carried out in the low μM range. This implies that on average all the duplexes present should be separated from each other, rendering the crystallographic structure obtained here very unlikely to occur in solution. Considering the high concentrations used during the crystallisation process, the aggregation of oligonucleotides due to the influence of the pyrene moieties could explain the

formation of the structure of this particular crystallographic artefact. Nevertheless, this remains a fascinating structure because it shows how a simple monomer, such as U-pyrene, can lead to the spontaneous formation of a regular nanostructure (Figure 4.13 A). Moreover, to the author's knowledge, this is the first X-ray crystal structure of a chemically modified pyrene oligonucleotide. The reliability of the solution obtained was reinforced by how the electron density map highlighted the presences of U-pyrene before it was added to the model, as shown by the omit map in Figure 4.13 B (R_{work} = 0.2567, R_{free} = 0.2842).

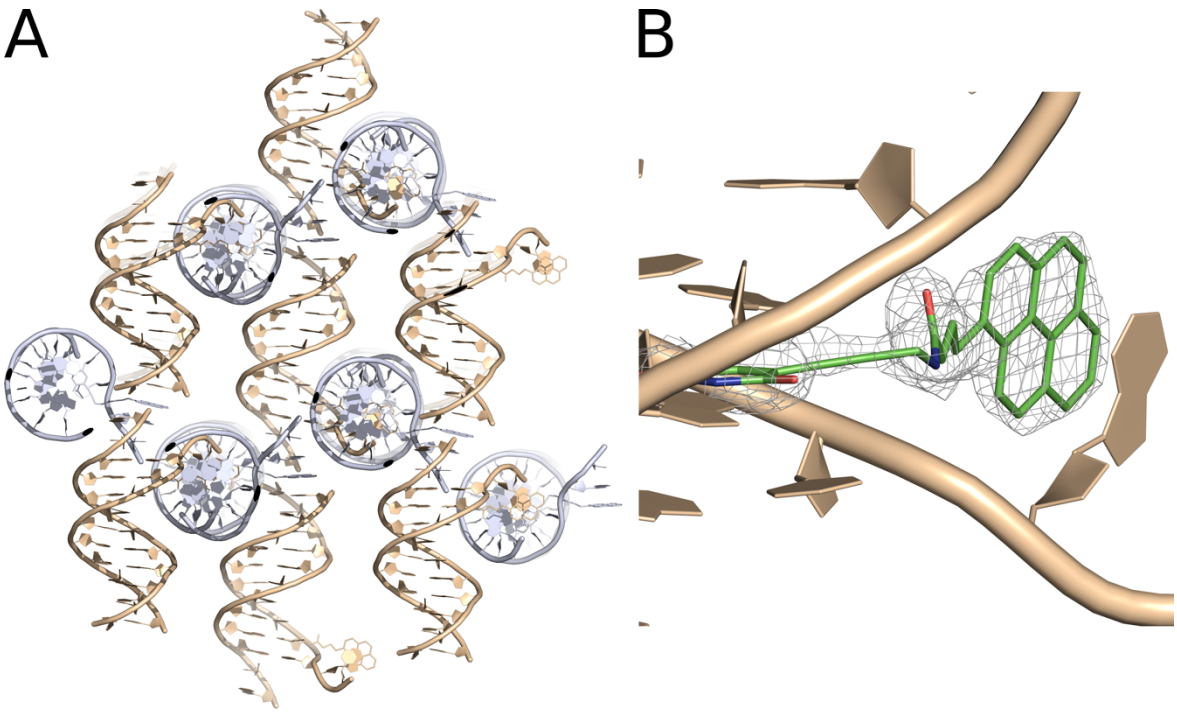


Figure 4.13 A) Fibrous nanostructure formed by CRY-1 in its crystal. Note how the gold fibres are perpendicular to the silver ones. **B)** Omit map showing how the electron density map derived before adding U-pyrene to the structure actually overlaps with U-pyrene. For simplicity and clarity only the electron density around U-pyrene is shown.

A force field simulation published by Kottysch *et al*. showed the pyrene moiety of U-pyrene intercalating in the same duplex, proving the concept that the linker length is sufficient for intercalation.[146] However, this did not occur in the crystal structure obtained (Figure 4.14). Various hydrophobic intramolecular interactions during the

crystallisation process could interfere with same-duplex intercalation, especially if this was leading to stronger crystal contacts and a more stable nanostructure. As it is possible to see from Figure 4.14 B, where CRY-1 is overlaid with the unmodified control (CRY-1C) obtained by Nelson *et al.*,[193] about two thirds of CRY-1 overlapped with the control, demonstrating that the B-conformation was maintained, especially in the centre of the duplex where the two structures were the most similar. Slightly bigger differences at one of the two ends were probably due to the different packing of the two crystals. The staggering difference came at the other end because of the effect of U-pyrene (Figure 4.15). Instead of intercalating in the same duplex, the pyrene moiety of U-pyrene stuck out of the dsDNA it was covalently linked to form intramolecular contacts with neighbouring molecules in the crystal, by intercalating in between the ends of two different symmetry related duplexes (see Figure 4.15 and Figure 4.13 A). The first base pair on the 5′-side of U-pyrene was maintained while the remaining two were disrupted, as the strand to which U-pyrene belonged turned about 90° to form a duplex with the 3′-end of a strand of a symmetry related structure. One of the most interesting elements of this unusual crystal packing was that these inter-duplex base pairs were actually a C:C and G:G mismatch. The overall stability gained from the formation of this nanostructure, formed by perpendicular fibres interlocked with each other and made essentially of B-DNA capped by pyrene (Figure 4.13 A), was able to counterbalance and in fact overcome the negative effect of these two mismatches, rendering this unusual structure stable and easy to obtain.

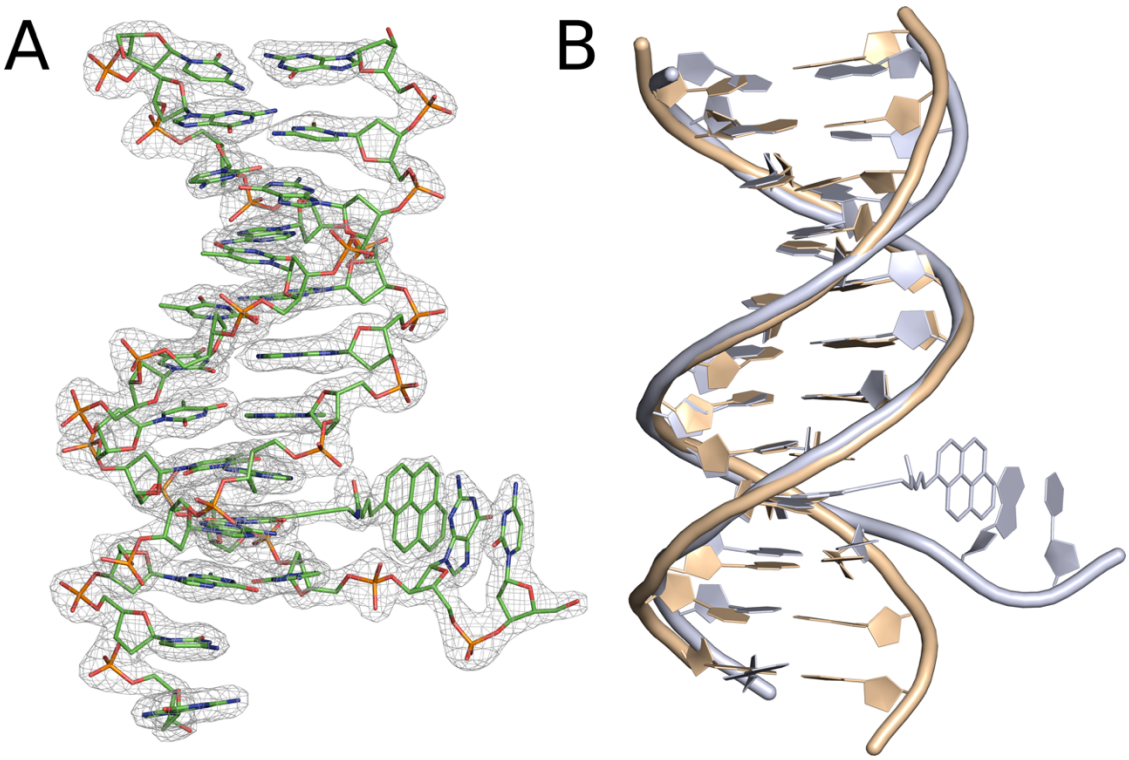


Figure 4.14 Crystallographic structure of CRY-1. **A)** Structure of CRY-1 with its electron density map in grey. **B)** Cartoon of the structure of CRY-1 (in silver) overlapped with the structure of the unmodified control (CRY-1C, in gold) obtained by Nelson *et al.*[193]

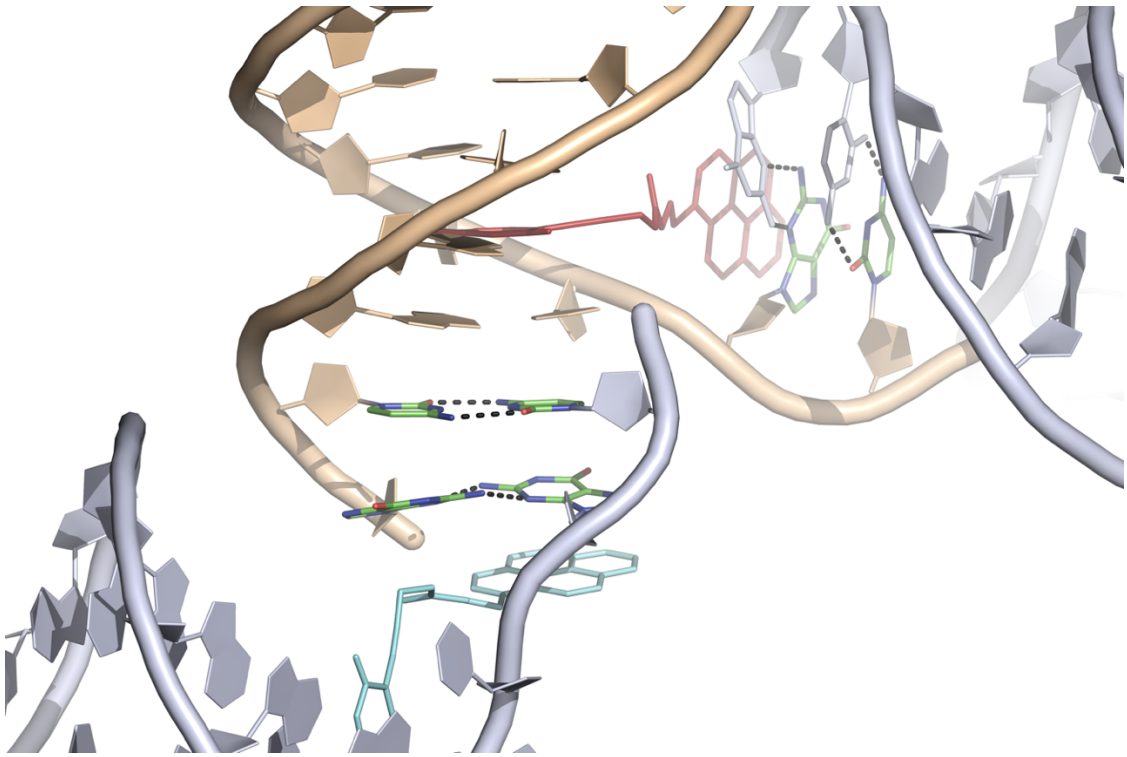


Figure 4.15 Details of the peculiar crystallographic packing of CRY-1. The main structure is in gold, while the symmetry related duplexes are in silver. The pyrene coming from the main duplex is in red, while the one coming from the symmetry related one is in light blue and they are both stacking at the end of one duplex. The bases forming the inter-duplex mismatches are coloured by element. H-bonds are in black.

4.6.6 CRY-2Br

Unfortunately, the structure obtained for the CRY-2Br crystal (Figure 4.16 A and Figure 4.18) was also a crystallographic artefact, for reasons analogous to those of CRY-1. To the author's knowledge, this is the second X-ray crystal structure (after CRY-1) of a chemically modified pyrene oligonucleotide. Once again, the reliability of the solution obtained was reinforced by how the electron density map highlighted the presences of the two U-pyrenes even before they were added to the model, as shown by the omit maps in Figure 4.16 B and C (R_{work} = 0.2668, R_{free} = 0.3271).

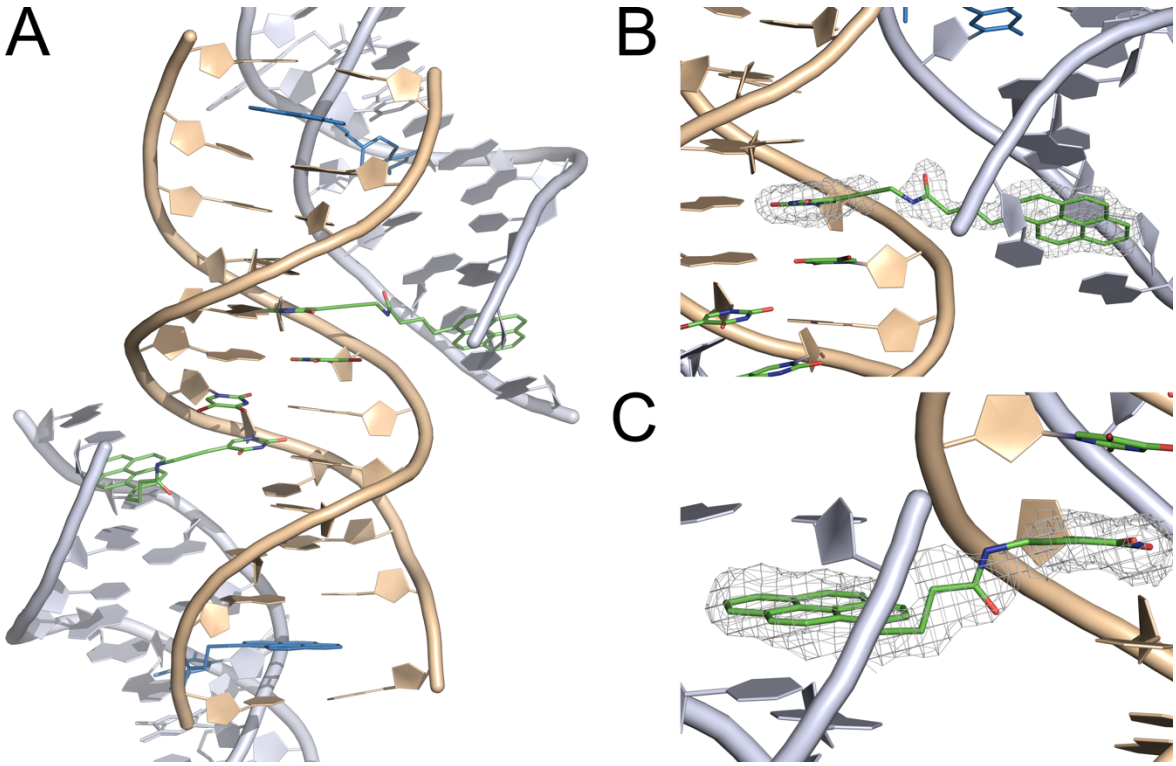


Figure 4.16 A) Overview of the inter-duplex intercalation of U-pyrene forming a "double-double locking" system. The main structure is in gold, while the symmetry related duplexes are in silver. U-pyrene and 5-BrU from the main duplex are coloured by element, while the pyrenes coming from the symmetry related DNA are in blue. **B)** and **C)** Omit maps showing how the electron density map derived before adding the two U-pyrenes to the structure actually overlaps with the two U-pyrenes. For simplicity and clarity only the electron density around U-pyrene is shown.

By using a self-complementary sequence, two U-pyrenes were present per duplex. Hence, the type of interaction obtained changed significantly from what seen for CRY-1. As it is possible to see from Figure 4.16 A and more in details in Figure 4.17, this time the pyrenes intercalated inside a symmetry related duplex between its last two base pairs and they did not cap the duplexes as before. This created four U-pyrene "bridges" connecting the main duplex to two symmetry related ones, forming a "double-double locking" system that helped to increase the stability of the nanostructure of the crystal.

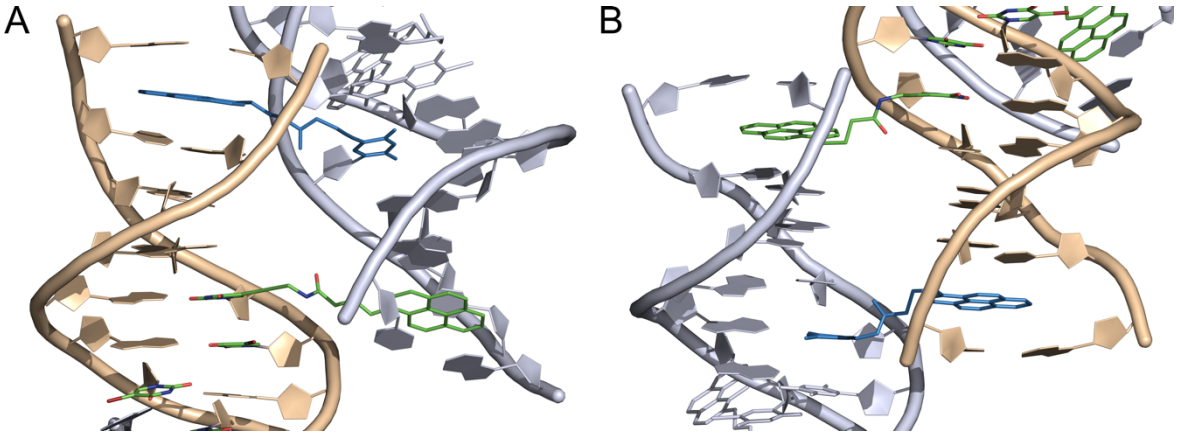


Figure 4.17 Details of the pyrene double locking system presented in Figure 4.16 A. The main structure is in gold, while the symmetry related duplexes are in silver. U-pyrene and 5-BrU from the main duplex are coloured by element, while a pyrene coming from a symmetry related DNA is in blue.

As it is possible to see from Figure 4.18 B, where CRY-2Br is overlaid with the unmodified control (CRY-2C) obtained by Drew *et al.*,[195] the majority of CRY-2Br overlapped with the control, showing that also in this case the B-conformation was maintained, especially in the centre of the duplex where the two structures were the most similar. The only significant difference was at the ends, where the last two base pairs were pushed apart to allow the intercalation of pyrene (Figure 4.18 B and Figure 4.17 for details).

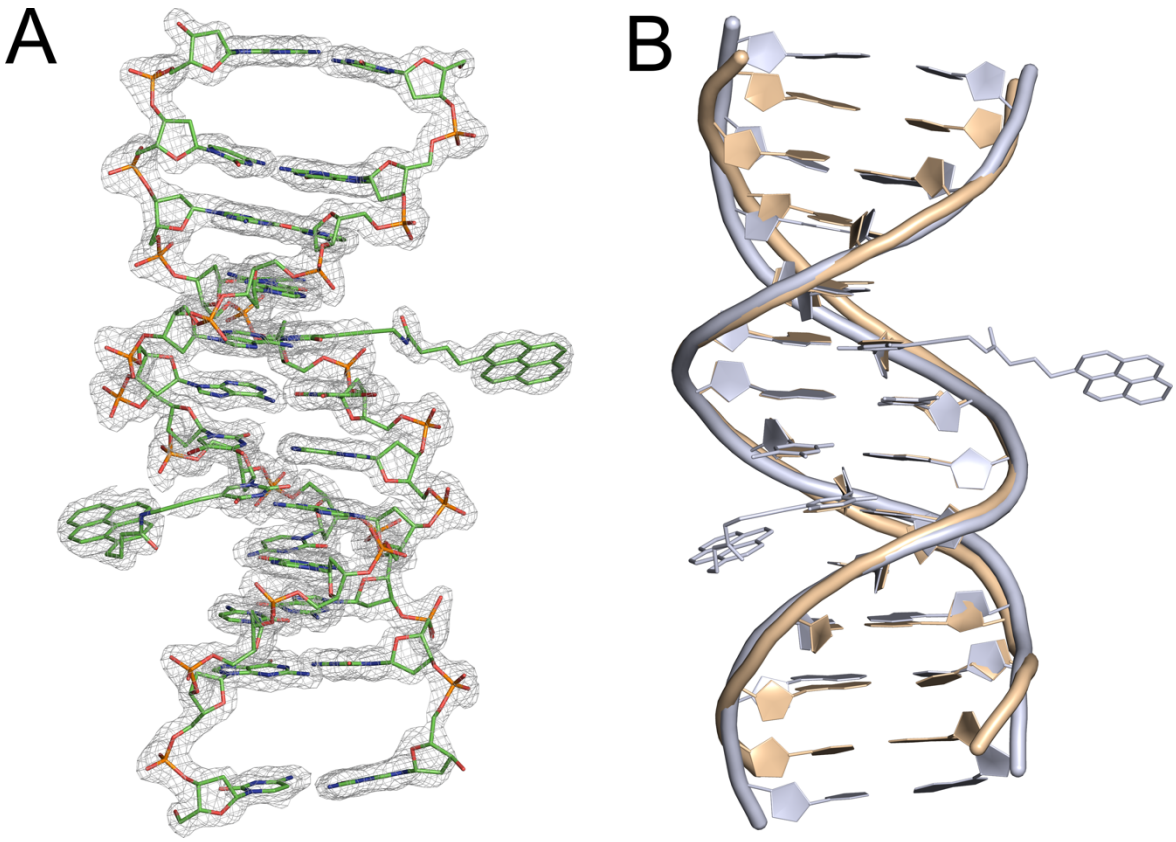


Figure 4.18 Crystallographic structure of CRY-2Br. **A)** Structure of CRY-2Br with its electron density map in grey. **B)** Cartoon of the structure of CRY-2Br (in silver) overlapped with the structure of the unmodified control (CRY-2C, in gold) obtained by Drew *et al.*[195]

Apart from the "double-double locking" system, that contributed to the formation and the stabilisation of the crystallographic contacts, groove-groove interactions also had an important role in the crystallographic packing. Figure 4.19 highlights this by showing the groove-groove interactions between the end of the gold duplex and the start of the symmetry related pink one. It is also interesting to note the location of the pyrenes coming from the silver duplexes, which are intercalating inside the gold and the pink sequences. They were found in a hydrophobic pocket, generated by the two grooves interacting, facing each other almost as in a pyrene-pyrene pseudo base pair.

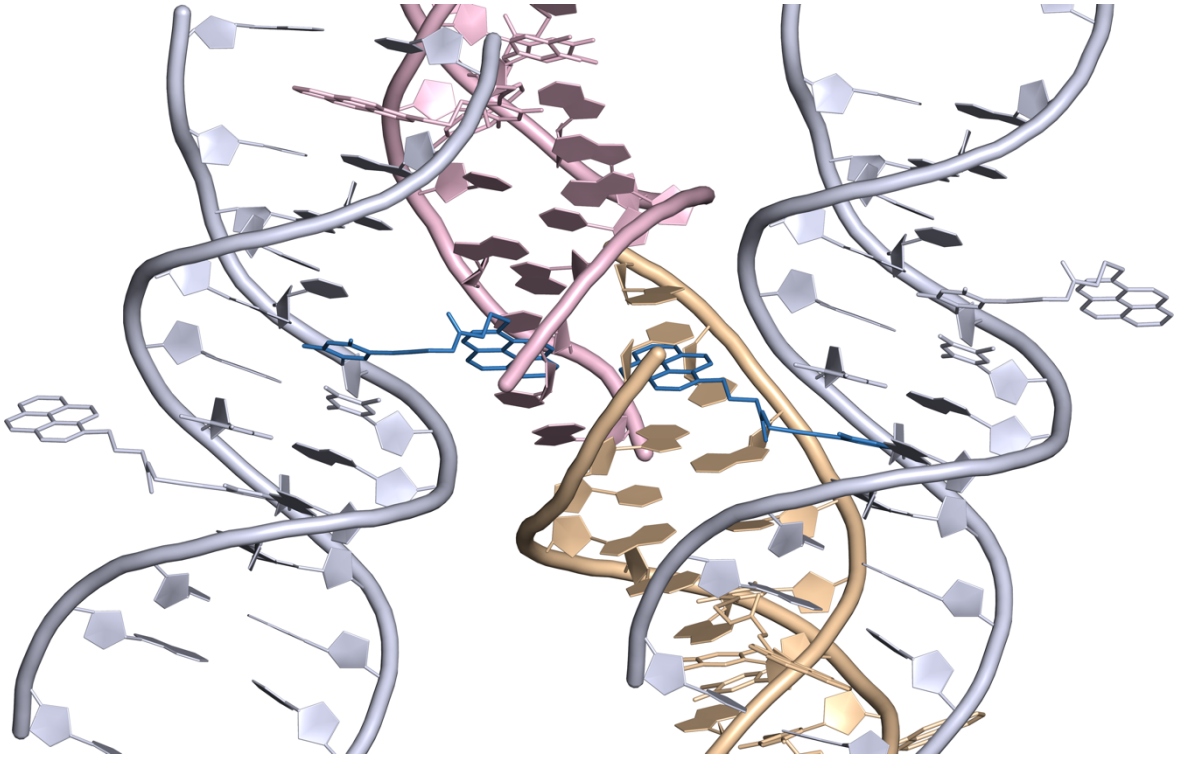


Figure 4.19 Details of the peculiar crystallographic packing of CRY-2Br. The main structure is in gold, while the symmetry related duplexes are in silver or pink depending on their orientation. Note the groove-groove interactions between the gold and the pink duplex. The two pyrenes coming from the silver duplexes stacking in the pink and the gold DNA are in blue.

These groove-groove interactions are important to explain the nanostructure that CRY-2Br formed (Figure 4.20). While CRY-1 had long straight fibres perpendicular to each other, this oligonucleotide formed wavy fibres all parallel to each other. The groove-groove interactions forced the duplexes on the same fibre to be slightly tilted, explaining the wavy aspect. As it is possible to see in Figure 4.20 B also in this case, as in CRY-1, there were two types of fibres. Gold-pink fibres were interlinked with silver-green ones thanks to the U-pyrene "bridges", forcing all the fibres to grow in the same direction and to be essentially parallel to each other (Figure 4.20 A). All these interactions contributed to strengthening the crystal during its formation, rendering this peculiar nanostructure stable and easy to obtain.

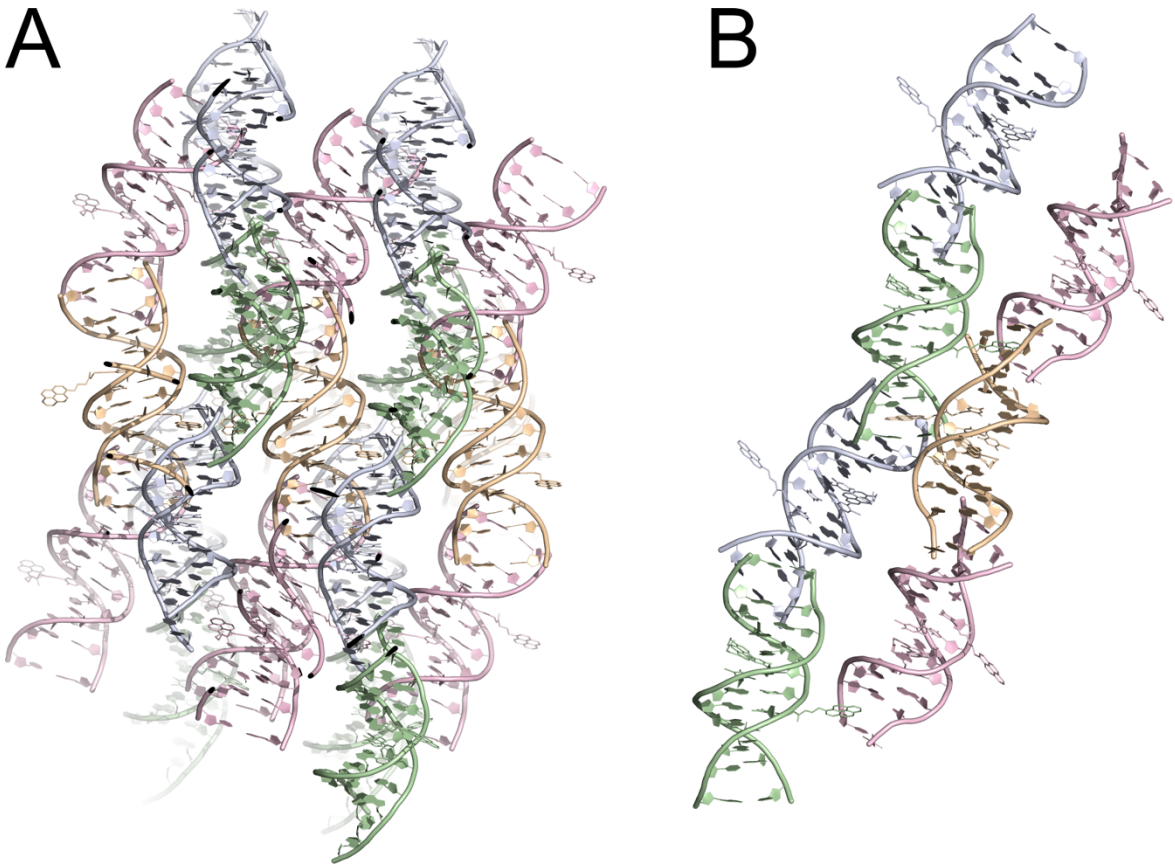


Figure 4.20 A) Fibrous nanostructure formed by CRY-2Br in its crystal. The duplex can have 4 different orientations, each orientation is coloured differently (gold, silver, green pink). **B)** Details of the two fundamental fibres (gold-pink and silver-green) forming the nanostructure.

4.7 Conclusions

The simple U-pyrene monomer proved to have a number of interesting properties. Thermodynamic studies showed that, while even a single U-pyrene was able to significantly stabilise a duplex, four additions of this monomer increased the binding constant (at 37 °C) of a 14-mer oligonucleotide to its complementary strand 10000-fold when compared to the unmodified control. This renders U-pyrene a useful modification when designing high affinity hybridisation probes for various diagnostic applications. Despite this, a single U-pyrene did not improve the discrimination against mismatches. If the mutation occurred between two U-pyrenes the discrimination capacity was essentially doubled. Further studies are needed in order to verify the

capacity of U-pyrene to maintain these properties in different sequence contexts. Unfortunately, the same stabilising properties were not present when the target was RNA, rendering U-pyrene less useful for antisense probes.

An unexpected property of U-pyrene emerged from the crystallographic studies. Although both the structures obtained were crystallographic artefacts that may not explain the behaviour of U-pyrene oligonucleotides in a diluted solution, they were interesting from a nanostructure point of view. Just by varying the sequence and the number of modifications, the simple U-pyrene was able to induce the formation of two very different nanostructures, thanks also to the extra hydrophobic interactions generated by the presence of pyrene. This suggests that U-pyrene could be potentially useful in those applications where unusual conformations are welcomed. Nanotechnology and the generation of nanostructures are the first obvious fields where its use could be envisaged. Another area that could potentially benefit from the use of U-pyrene is aptamer research, as adding a simple modification that can allow the formation of different conformations and interactions could help improving the selection of more chemically varied tight-binding aptamers.

Chapter 5

Conclusions and Future Work

Pyrene-functionalised oligonucleotides have been shown to increase the thermal stability of the DNA duplexes they form and structural studies suggest good pyrene intercalation is important for this effect to occur. The work presented herein examined different nucleobases analogues containing pyrene corroborating these observations.

X-pyrene proved to be an excellent nucleobase analogue capable of greatly stabilising bulged duplexes, while also having impressive mismatch detection properties. X-pyrene was able to achieve this due to its distinctive feature of having different duplex stabilising factors working cooperatively, combining additional non-nucleobase π-stacking, H-bonding and electrostatic attraction into a single monomer. An NMR study also verified that the pyrene moiety intercalated inside the pocket generated by the bulged base, which may help explain the fluorescent properties of X-pyrene.

A potential drawback of the use of X-pyrene is its long and low-yielding synthesis. To overcome this, a method based on reductive amination was successfully developed to obtain X-pyrene by labelling of oligonucleotides containing the commercially available G-clamp monomer. This methodology also produced a doubly pyrene-modified X-pyrene analogue, named here X-2*pyrene. While initial experiments with X-2*pyrene showed that it formed significantly stable normal and bulged duplexes despite its high steric demands, more in depth structural studies will be necessary to fully understand its properties. Molecular dynamics and NMR studies would be particularly beneficial to explore the molecular details of interactions involving this modification.

It would also be interesting to study the effect of the length of the linker between the pyrene moiety and the G-clamp monomer in order to understand the effect on the

stabilisation properties of the X-pyrene family. Molecular dynamics studies would be of great help in trying to predict the behaviour and subsequent NMR experiments would be beneficial to confirm the theoretical predictions.

Reductive amination labelling was also used to obtain the singly and doubly pyrene-modified amino C6 dT. Both modifications proved to be capable of stabilising DNA duplexes, whether fully matched or with a bulged base next to the modification. Further studies are necessary to explore the capacity of detecting mismatches and an eventual sequence dependency of the properties. There is also need for structural studies to understand the molecular details of the stabilisation generated by these two modifications in different contexts.

U-pyrene is another relatively simple dT modification that was found to be able to significantly stabilise the DNA duplexes into which it was incorporated. Impressively, four additions of this monomer increased the binding constant (at 37 °C) of a 14-mer oligonucleotide to its complementary strand 10000-fold when compared to the unmodified control, reaching a remarkably low picomolar range. Further studies are necessary to explore different sequence contexts, in order to analyse the potential uses of U-pyrene.

In order to better understand the mechanism of stabilisation of U-pyrene, X-ray crystallographic studies were performed. To the author's knowledge, this work produced the first two X-ray crystal structures of a chemically modified pyrene oligonucleotide. Although the structures obtained were crystallographic artefacts, with little connection to the behaviour in a dilute solution, they unexpectedly highlighted that U-pyrene can promote the formation of two different stable nanostructures in the solid phase. It would be interesting to use X-ray crystallography

to continue exploring the production of different nanostructures by varying the number of U-pyrenes and the sequence or also by changing the linker length and type or the pyrene monomer used. At the same time, it would be useful to use high field NMR studies to understand the behaviour of the same duplexes in more dilute solutions. Finally, molecular dynamics could be used to bridge the gap, trying to understand the transition from the solution structure to the crystal one. The use of these techniques could help to rationalise the formation of the nanostructures, eventually allowing to predict the outcome and therefore improving the capacity of designing them.

The work in this thesis focused on the use of pyrene, as it is a reasonably well-studied modification which it is relatively easy to work with, and it showed how a careful study of even comparatively simple modifications can lead to the discovery of interesting and sometimes unexpected properties. Nevertheless, it would be of great importance to explore other moieties in place of pyrene, with anthraquinones and thiazole orange being obvious choices given their known capacity for intercalation. What it is clear is that a combination of stability, thermodynamic and structural studies is highly beneficial in obtaining a thorough understanding of the properties of such modifications.

References

(1) Dahm, R. Friedrich Miescher and the discovery of DNA. *Dev. Biol.* **2005**, *278*, 274-288.

(2) Avery, O. T.; MacLeod, C. M.; McCarty, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal Types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus Type III. *J. Exp. Med.* **1944**, *79*, 137-158.

(3) Hershey, A. D.; Chase, M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **1952**, *36*, 39-56.

(4) Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737-738.

(5) Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **1950**, *6*, 201-209.

(6) Chargaff, E. Structure and function of nucleic acids as cell constituents. *Fed. Proc.* **1951**, *10*, 654-659.

(7) Chargaff, E.; Vischer, E.; Doniger, R.; Green, C.; Misani, F. The composition of the desoxypentose nucleic acids of thymus and spleen. *J. Biol. Chem.* **1949**, *177*, 405-416.

(8) Franklin, R. E.; Gosling, R. G. Molecular configuration in sodium thymonucleate. *Nature* **1953**, *171*, 740-741.

(9) Wilkins, M. H. F.; Seeds, W. E.; Stokes, A. R.; Wilson, H. R. Helical structure of crystalline deoxypentose nucleic acid. *Nature* **1953**, *172*, 759-762.

(10) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. Molecular structure of nucleic acids: molecular structure of deoxypentose nucleic acids. *Nature* **1953**, *171*, 738-740.

(11) *Nucleic Acids in Chemistry and Biology*; 3rd ed ed.; RSC Publishing: Cambridge, 2006.

(12) Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L. Polymorphism of DNA double helices. *J. Mol. Biol.* **1980**, *143*, 49-72.

(13) Dickerson, R.; Drew, H.; Conner, B.; Wing, R.; Fratini, A.; Kopka, M. The anatomy of A-, B-, and Z-DNA. *Science* **1982**, *216*, 475-485.

(14) Lehninger, A. L.; Nelson, D. L.; Cox, M. M.: *Lehninger principles of biochemistry*; 5th ed.; W.H. Freeman: New York, 2008.

(15) Campbell, I. D.: *Biophysical techniques*; Oxford University Press: Oxford, 2012.

(16) Levine, I. N.: *Physical chemistry*; 6th ed.; McGraw-Hill: Boston, 2009.

(17) Eriksson, M.; Nordén, B.: Linear and circular dichroism of drug-nucleic acid complexes. In *Methods Enzymol.*; Academic Press, 2001; Vol. Volume 340; pp 68-98.

(18) Lyng, R.; Härd, T.; Norden, B. Induced CD of DNA intercalators: electric dipole allowed transitions. *Biopolymers* **1987**, *26*, 1327-1345.

(19) Nordén, B.; Kurucsev, T. Analysing DNA complexes by circular and linear dichroism. *J. Mol. Recognit.* **1994**, *7*, 141-155.

(20) Atkins, P. W.; De Paula, J.: *Atkins' Physical chemistry*; 9th ed.; Oxford University Press: Oxford, 2010.

(21) Lakowicz, J. R.: *Principles of fluorescence spectroscopy*; 3rd ed.; Springer: New York, 2006.

(22) Skoog, D. A.; Holler, F. J.; Crouch, S. R.: *Principles of instrumental analysis*; 6th ed.; Thomson Brooks/Cole: Belmont, CA, 2007.

(23) Freier, S. M.; Altmann, K.-H. The ups and downs of nucleic acid duplex stability: Structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.* **1997**, *25*, 4429-4443.

(24) Kurreck, J. Antisense technologies. *Eur. J. Biochem.* **2003**, *270*, 1628-1644.

(25) Lamond, A. I.; Sproat, B. S. Antisense oligonucleotides made of 2'-O-alkylRNA: their properties and applications in RNA biochemistry. *FEBS Lett.* **1993**, *325*, 123-127.

(26) Manoharan, M. Oligonucleotide conjugates as potential antisense drugs with improved uptake, biodistribution, targeted delivery, and mechanism of action. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 103-128.

(27) Flanagan, W. M.; Wagner, R. W.; Grant, D.; Lin, K.-Y.; Matteucci, M. D. Cellular penetration and antisense activity by a phenoxazine-substituted heptanucleotide. *Nat. Biotechnol.* **1999**, *17*, 48-52.

(28) Monia, B. P.; Sasmor, H.; Johnston, J. F.; Freier, S. M.; Lesnik, E. A.; Muller, M.; Geiger, T.; Altmann, K.-H.; Moser, H.; Fabbro, D. Sequence-specific antitumor activity of a phosphorothioate oligodeoxyribonucleotide targeted to human C-raf kinase supports an antisense mechanism of action in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15481-15484.

(29) Wagner, R.; Matteucci, M.; Lewis, J.; Gutierrez, A.; Moulds, C.; Froehler, B. Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. *Science* **1993**, *260*, 1510-1513.

(30) Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *J. Biol. Chem.* **1993**, *268*, 14514-14522.

(31) Monia, B. P.; Johnston, J. F.; Ecker, D. J.; Zounes, M. A.; Lima, W. F.; Freier, S. M. Selective inhibition of mutant Ha-ras mRNA expression by antisense oligonucleotides. *J. Biol. Chem.* **1992**, *267*, 19954-19962.

(32) Leal, N. A.; Sukeda, M.; Benner, S. A. Dynamic assembly of primers on nucleic acid templates. *Nucleic Acids Res.* **2006**, *34*, 4702-4710.

(33) Demidov, V. V.; Frank-Kamenetskii, M. D. Two sides of the coin: affinity and specificity of nucleic acid interactions. *Trends Biochem. Sci.*, *29*, 62-71.

(34) Dobson, N.; McDowell, D. G.; French, D. J.; Brown, L. J.; Mellor, J. M.; Brown, T. Synthesis of HyBeacons and dual-labelled probes containing 2[prime or minute]-fluorescent groups for use in genetic analysis. *Chem. Commun.* **2003**, 1234-1235.

(35) French, D. J.; Archard, C. L.; Andersen, M. T.; McDowell, D. G. Ultra-rapid DNA analysis using HyBeacon™ probes and direct PCR amplification from saliva. *Mol. Cell. Probes* **2002**, *16*, 319-326.

(36) French, D. J.; Archard, C. L.; Brown, T.; McDowell, D. G. HyBeacon™ probes: a new tool for DNA sequence detection and allele discrimination. *Mol. Cell. Probes* **2001**, *15*, 363-374.

(37) Gale, N.; French, D. J.; Howard, R. L.; McDowell, D. G.; Debenham, P. G.; Brown, T. Rapid typing of STRs in the human genome by HyBeacon® melting. *Org. Biomol. Chem.* **2008**, *6*, 4553-4559.

(38) Richardson, J. A.; Gerowska, M.; Shelbourne, M.; French, D.; Brown, T. Six-colour HyBeacon probes for multiplex genetic analysis. *ChemBioChem* **2010**, *11*, 2530-2533.

(39) Johnson, R. P.; Richardson, J. A.; Brown, T.; Bartlett, P. N. A label-free, electrochemical SERS-based assay for detection of DNA hybridization and discrimination of mutations. *J. Am. Chem. Soc.* **2012**, *134*, 14099-14107.

(40) Johnson, R. P.; Richardson, J. A.; Brown, T.; Bartlett, P. N. Real-time surface-enhanced Raman spectroscopy monitoring of surface pH during electrochemical melting of double-stranded DNA. *Langmuir* **2012**, *28*, 5464-5470.

(41) Mahajan, S.; Richardson, J.; Brown, T.; Bartlett, P. N. SERS-melting: a new method for discriminating mutations in DNA sequences. *J. Am. Chem. Soc.* **2008**, *130*, 15589-15601.

(42) Papadopoulou, E.; Gale, N.; Goodchild, S. A.; Cleary, D. W.; Weller, S. A.; Brown, T.; Bartlett, P. N. Strain discrimination of Yersinia pestis using a SERS-based electrochemically driven melting curve analysis of variable number tandem repeat sequences. *Chem. Sci.* **2015**, *6*, 1846-1852.

(43) Papadopoulou, E.; Goodchild, S. A.; Cleary, D. W.; Weller, S. A.; Gale, N.; Stubberfield, M. R.; Brown, T.; Bartlett, P. N. Using surface-enhanced raman spectroscopy and electrochemically driven melting to discriminate yersinia pestis from y. pseudotuberculosis based on single nucleotide polymorphisms within unpurified polymerase chain reaction amplicons. *Anal. Chem.* **2015**, *87*, 1605-1612.

(44) Papadopoulou, E.; Meneghello, M.; Marafini, P.; Johnson, R. P.; Brown, T.; Bartlett, P. N. The effect of temperature on electrochemically driven

denaturation monitored by SERS. *Bioelectrochemistry* **2015**, *106, Part B*, 353-358.

(45) Sandin, P.; Tumpane, J.; Börjesson, K.; Wilhelmsson, L. M.; Brown, T.; Nordén, B.; Albinsson, B.; Lincoln, P. Thermodynamic aspects of DNA nanoconstruct stability and design. *J. Phys. Chem. C* **2009**, *113*, 5941-5946.

(46) Salunkhe, M.; Wu, T.; Letsinger, R. L. Control of folding and binding of oligonucleotides by use of a nonnucleotide linker. *J. Am. Chem. Soc.* **1992**, *114*, 8768-8772.

(47) Tanaka, K.; Clever, G. H.; Takezawa, Y.; Yamada, Y.; Kaul, C.; Shionoya, M.; Carell, T. Programmable self-assembly of metal ions inside artificial DNA duplexes. *Nat. Nanotechnol.* **2006**, *1*, 190-194.

(48) Gothelf, K. V.; LaBean, T. H. DNA-programmed assembly of nanostructures. *Org. Biomol. Chem.* **2005**, *3*, 4023-4037.

(49) Miduturu, C. V.; Silverman, S. K. DNA constraints allow rational control of macromolecular conformation. *J. Am. Chem. Soc.* **2005**, *127*, 10144-10145.

(50) Astakhova, I. K.; Santhosh Kumar, T.; Campbell, M. A.; Ustinov, A. V.; Korshun, V. A.; Wengel, J. Branched DNA nanostructures efficiently stabilised and monitored by novel pyrene-perylene 2′-α-l-amino-LNA FRET pairs. *Chem. Commun.* **2013**, *49*, 511-513.

(51) Wilhelmsson, L. M. Fluorescent nucleic acid base analogues. *Q. Rev. Biophys.* **2010**, *43*, 159-183.

(52) Sowers, L. C.; Boulard, Y.; Fazakerley, G. V. Multiple structures for the 2-aminopurine−cytosine mispair. *Biochemistry* **2000**, *39*, 7613-7620.

(53) Ward, D. C.; Reich, E.; Stryer, L. Fluorescence studies of nucleotides and polynucleotides: I. formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* **1969**, *244*, 1228-1237.

(54) Hawkins, M. Fluorescent pteridine nucleoside analogs. *Cell Biochem. Biophys.* **2001**, *34*, 257-281.

(55) Hawkins, M. E.; Pfleiderer, W.; Jungmann, O.; Balis, F. M. Synthesis and fluorescence characterization of pteridine adenosine nucleoside analogs for DNA incorporation. *Anal. Biochem.* **2001**, *298*, 231-240.

(56) Hawkins, M. E.; Pfleiderer, W.; Balis, F. M.; Porter, D.; Knutson, J. R. Fluorescence properties of pteridine nucleoside analogs as monomers and incorporated into oligonucleotides. *Anal. Biochem.* **1997**, *244*, 86-95.

(57) Berry, D. A.; Jung, K.-Y.; Wise, D. S.; Sercel, A. D.; Pearson, W. H.; Mackie, H.; Randolph, J. B.; Somers, R. L. Pyrrolo-dC and pyrrolo-C: fluorescent analogs of cytidine and 2′-deoxycytidine for the study of oligonucleotides. *Tetrahedron Lett.* **2004**, *45*, 2457-2461.

(58) Lin, K.-Y.; Jones, R. J.; Matteucci, M. Tricyclic 2'-deoxycytidine analogs: syntheses and incorporation into oligodeoxynucleotides which have enhanced binding to complementary RNA. *J. Am. Chem. Soc.* **1995**, *117*, 3873-3874.

(59) Börjesson, K.; Preus, S.; El-Sagheer, A. H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. M. Nucleic acid base analog FRET-pair facilitating detailed structural measurements in nucleic acid containing systems. *J. Am. Chem. Soc.* **2009**, *131*, 4288-4293.

(60) Engman, K. C.; Sandin, P.; Osborne, S.; Brown, T.; Billeter, M.; Lincoln, P.; Norden, B.; Albinsson, B.; Wilhelmsson, L. M. DNA adopts normal B-form upon incorporation of highly fluorescent DNA base analogue tC: NMR structure and UV-Vis spectroscopy characterization. *Nucleic Acids Res.* **2004**, *32*, 5087-5095.

(61) Sandin, P.; Börjesson, K.; Li, H.; Martensson, J.; Brown, T.; Wilhelmsson, L. M.; Albinsson, B. Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Res.* **2008**, *36*, 157-167.

(62) Wilhelmsson, L. M.; Holmén, A.; Lincoln, P.; Nielsen, P. E.; Nordén, B. A highly fluorescent DNA base analogue that forms Watson–Crick base pairs with guanine. *J. Am. Chem. Soc.* **2001**, *123*, 2434-2435.

(63) Wilhelmsson, L. M.; Sandin, P.; Holmén, A.; Albinsson, B.; Lincoln, P.; Nordén, B. Photophysical characterization of fluorescent DNA base analogue, tC. *J. Phys. Chem. B* **2003**, *107*, 9094-9101.

(64) Sandin, P.; Wilhelmsson, L. M.; Lincoln, P.; Powers, V. E. C.; Brown, T.; Albinsson, B. Fluorescent properties of DNA base analogue tC upon incorporation into DNA — negligible influence of neighbouring bases on fluorescence quantum yield. *Nucleic Acids Res.* **2005**, *33*, 5019-5025.

(65) Preus, S.; Börjesson, K.; Kilså, K.; Albinsson, B.; Wilhelmsson, L. M. Characterization of nucleobase analogue FRET acceptor tCnitro. *J. Phys. Chem. B* **2009**, *114*, 1050-1056.

(66) Börjesson, K.; Sandin, P.; Wilhelmsson, L. M. Nucleic acid structure and sequence probing using fluorescent base analogue tCO. *Biophys. Chem.* **2009**, *139*, 24-28.

(67) Stengel, G.; Urban, M.; Purse, B. W.; Kuchta, R. D. High density labeling of polymerase chain reaction products with the fluorescent base analogue tCo. *Anal. Chem.* **2009**, *81*, 9079-9085.

(68) Walsh, J. M.; Bouamaied, I.; Brown, T.; Wilhelmsson, L. M.; Beuning, P. J. Discrimination against the cytosine analog tC by Escherichia coli DNA polymerase IV DinB. *J. Mol. Biol.* **2011**, *409*, 89-100.

(69) Sandin, P.; Stengel, G.; Ljungdahl, T.; Börjesson, K.; Macao, B.; Wilhelmsson, L. M. Highly efficient incorporation of the fluorescent nucleotide analogs tC and tCO by Klenow fragment. *Nucleic Acids Res.* **2009**, *37*, 3924-3933.

(70) Stengel, G.; Purse, B. W.; Wilhelmsson, L. M.; Urban, M.; Kuchta, R. D. Ambivalent incorporation of the fluorescent cytosine analogues tC and tCo by human DNA polymerase α and Klenow fragment. *Biochemistry* **2009**, *48*, 7547-7555.

(71) Anand, N. N.; Brown, D. M.; Salisbury, S. A. The stability of oligodeoxyribonucleotide duplexes containing degenerate bases. *Nucleic Acids Res.* **1987**, *15*, 8167-8176.

(72) Stengel, G.; Gill, J. P.; Sandin, P.; Wilhelmsson, L. M.; Albinsson, B.; Nordén, B.; Millar, D. Conformational dynamics of DNA polymerase probed with a novel fluorescent DNA base analogue. *Biochemistry* **2007**, *46*, 12289-12297.

(73) Temiakov, D.; Patlan, V.; Anikin, M.; McAllister, W. T.; Yokoyama, S.; Vassylyev, D. G. Structural basis for substrate selection by T7 RNA polymerase. *Cell* **2004**, *116*, 381-391.

(74) Stengel, G.; Urban, M.; Purse, B. W.; Kuchta, R. D. Incorporation of the fluorescent ribonucleotide analogue tCTP by T7 RNA polymerase. *Anal. Chem.* **2010**, *82*, 1082-1089.

(75) Urban, M.; Joubert, N.; Purse, B. W.; Hocek, M.; Kuchta, R. D. Mechanisms by which human DNA primase chooses to polymerize a nucleoside triphosphate. *Biochemistry* **2009**, *49*, 727-735.

(76) Østergaard, M. E.; Hrdlicka, P. J. Pyrene-functionalized oligonucleotides and locked nucleic acids (LNAs): Tools for fundamental research, diagnostics, and nanotechnology. *Chem. Soc. Rev.* **2011**, *40*, 5771-5788.

(77) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T. Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J. Am. Chem. Soc.* **2000**, *122*, 2213-2222.

(78) Robertson, J. M.; White, J. G. 72. The crystal structure of pyrene. A quantitative X-ray investigation. *J. Chem. Soc.* **1947**, 358-368.

(79) Dougherty, G.; Pilbrow, J. R. Physico-chemical probes of intercalation. *Int. J. Biochem.* **1984**, *16*, 1179-1192.

(80) Yao, C.; Kraatz, H.-B.; Steer, R. P. Photophysics of pyrene-labelled compounds of biophysical interest. *Photochem. Photobiol. Sci.* **2005**, *4*, 191-199.

(81) Manoharan, M.; Tivel, K. L.; Zhao, M.; Nafisi, K.; Netzel, T. L. Base-sequence dependence of emission lifetimes for DNA oligomers and duplexes covalently labeled with pyrene: relative electron-transfer quenching efficiencies of A, G, C, and T nucleosides toward pyrene. *J. Phys. Chem.* **1995**, *99*, 17461-17472.

(82) Østergaard, M. E.; Kumar, P.; Baral, B.; Guenther, D. C.; Anderson, B. A.; Ytreberg, F. M.; Deobald, L.; Paszczynski, A. J.; Sharma, P. K.; Hrdlicka, P. J. C5-Functionalized DNA, LNA, and α-L-LNA: positional control of polarity-sensitive fluorophores leads to improved SNP-typing. *Chem. Eur. J.* **2011**, *17*, 3157-3165.

(83) Seo, Y. J.; Ryu, J. H.; Kim, B. H. Quencher-free, end-stacking oligonucleotides for probing single-base mismatches in DNA. *Org. Lett.* **2005**, *7*, 4931-4933.

(84) Bichenkova, E. V.; Gbaj, A.; Walsh, L.; Savage, H. E.; Rogert, C.; Sardarian, A. R.; Etchells, L. L.; Douglas, K. T. Detection of nucleic acids in situ: novel oligonucleotide analogues for target-assembled DNA-mounted exciplexes. *Org. Biomol. Chem.* **2007**, *5*, 1039-1051.

(85) Mei, H.; Ingale, S. A.; Seela, F. Pyrene and bis-pyrene DNA nucleobase conjugates: excimer and monomer fluorescence of linear and dendronized cytosine and 7-deazaguanine click adducts. *Tetrahedron* **2013**, *69*, 4731-4742.

(86) Dioubankova, N. N.; Malakhov, A. D.; Stetsenko, D. A.; Korshun, V. A. Detection of point mutations using pyrene-labeled DNA probes. *Russ. Chem. Bull.* **2004**, *53*, 463-470.

(87) Masuko, M.; Ohtani, H.; Ebata, K.; Shimadzu, A. Optimization of excimer-forming two-probe nucleic acid hybridization method with pyrene as a fluorophore. *Nucleic Acids Res.* **1998**, *26*, 5409-5416.

(88) Winnik, F. M. Photophysics of preassociated pyrenes in aqueous polymer solutions and in other organized media. *Chem. Rev.* **1993**, *93*, 587-614.

(89) Ebata, K.; Masuko, M.; Ohtani, H.; Kashiwasake-Jibu, M. Nucleic acid hybridization accompanied with excimer formation from two pyrene labelled probes. *Photochem. Photobiol.* **1995**, *62*, 836-839.

(90) Paris, P. L.; Langenhan, J. M.; Kool, E. T. Probing DNA sequences in solution with a monomer-excimer fluorescence color change. *Nucleic Acids Res.* **1998**, *26*, 3789-3793.

(91) Ren, R. X. F.; Chaudhuri, N. C.; Paris, P. L.; Rumney; Kool, E. T. Naphthalene, Phenanthrene, and Pyrene as DNA Base Analogues: Synthesis, Structure, and Fluorescence in DNA. *J. Am. Chem. Soc.* **1996**, *118*, 7671-7678.

(92) Matray, T. J.; Kool, E. T. Selective and stable DNA base pairing without hydrogen bonds. *J. Am. Chem. Soc.* **1998**, *120*, 6191-6192.

(93) Vesnaver, G.; Chang, C. N.; Eisenberg, M.; Grollman, A. P.; Breslauer, K. J. Influence of abasic and anucleosidic sites on the stability, conformation, and

melting behavior of a DNA duplex: correlations of thermodynamic and structural data. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 3614-3618.

(94) Smirnov, S.; Matray, T. J.; Kool, E. T.; de los Santos, C. Integrity of duplex structures without hydrogen bonding: DNA with pyrene paired at abasic sites. *Nucleic Acids Res.* **2002**, *30*, 5561-5569.

(95) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool, E. T. Experimental measurement of aromatic stacking affinities in the context of duplex DNA. *J. Am. Chem. Soc.* **1996**, *118*, 8182-8183.

(96) Ostergaard, M. E.; Guenther, D. C.; Kumar, P.; Baral, B.; Deobald, L.; Paszczynski, A. J.; Sharma, P. K.; Hrdlicka, P. J. Pyrene-functionalized triazole-linked 2′-deoxyuridines-probes for discrimination of single nucleotide polymorphisms (SNPs). *Chem. Commun.* **2010**, *46*, 4929-4931.

(97) Skorobogatyi, M. V.; Malakhov, A. D.; Pchelintseva, A. A.; Turban, A. A.; Bondarev, S. L.; Korshun, V. A. Fluorescent 5-alkynyl-2′-deoxyuridines: high emission efficiency of a conjugated perylene nucleoside in a DNA duplex. *ChemBioChem* **2006**, *7*, 810-816.

(98) Seo, Y. J.; Kim, B. H. Probing the B-to-Z-DNA duplex transition using terminally stacking ethynyl pyrene-modified adenosine and uridine bases. *Chem. Commun.* **2006**, 150-152.

(99) Bag, S. S.; Kundu, R.; Matsumoto, K.; Saito, Y.; Saito, I. Singly and doubly labeled base-discriminating fluorescent oligonucleotide probes containing oxo-pyrene chromophore. *Biorg. Med. Chem. Lett.* **2010**, *20*, 3227-3230.

(100) Kostenko, E.; Dobrikov, M.; Komarova, N.; Pyshniy, D.; Vlassov, V.; Zenkova, M. 5′-bis-pyrenylated oligonucleotides display enhanced excimer fluorescence upon hybridization with DNA and RNA. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 1859-1870.

(101) Yamana, K.; Iwase, R.; Furutani, S.; Tsuchida, H.; Zako, H.; Yamaoka, T.; Murakami, A. 2′-Pyrene modified oligonucleotide provides a highly sensitive fluorescent probe of RNA. *Nucleic Acids Res.* **1999**, *27*, 2387-2392.

(102) Yamana, K.; Zako, H.; Asazuma, K.; Iwase, R.; Nakano, H.; Murakami, A. Fluorescence detection of specific RNA sequences using 2′-pyrene-modified oligoribonucleotides. *Angew. Chem. Int. Ed.* **2001**, *40*, 1104-1106.

(103) Nakamura, M.; Fukunaga, Y.; Sasa, K.; Ohtoshi, Y.; Kanaori, K.; Hayashi, H.; Nakano, H.; Yamana, K. Pyrene is highly emissive when attached to the RNA duplex but not to the DNA duplex: the structural basis of this difference. *Nucleic Acids Res.* **2005**, *33*, 5887-5895.

(104) Lou, C.; Dallmann, A.; Marafini, P.; Gao, R.; Brown, T. Enhanced H-bonding and π-stacking in DNA: a potent duplex-stabilizing and mismatch sensing nucleobase analogue. *Chem. Sci.* **2014**, *5*, 3836-3844.

(105) Sampath, D.; Rao, V. A.; Plunkett, W. Mechanisms of apoptosis induction by nucleoside analogs. *Oncogene* **2003**, *22*, 9063-9074.

(106) Nakagawa, O.; Ono, S.; Li, Z.; Tsujimoto, A.; Sasaki, S. Specific fluorescent probe for 8-oxoguanosine. *Angew. Chem. Int. Ed.* **2007**, *46*, 4500-4503.

(107) Lin, K.-Y.; Matteucci, M. D. A cytosine analogue capable of clamp-like binding to a guanine in helical nucleic acids. *J. Am. Chem. Soc.* **1998**, *120*, 8531-8532.

(108) Flanagan, W. M.; Wolf, J. J.; Olson, P.; Grant, D.; Lin, K.-Y.; Wagner, R. W.; Matteucci, M. D. A cytosine analog that confers enhanced potency to antisense oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3513-3518.

(109) Maier, M. A.; Leeds, J. M.; Balow, G.; Springer, R. H.; Bharadwaj, R.; Manoharan, M. Nuclease resistance of oligonucleotides containing the tricyclic cytosine analogues phenoxazine and 9-(2-aminoethoxy)-phenoxazine ("G-clamp") and origins of their nuclease resistance properties. *Biochemistry* **2002**, *41*, 1323-1327.

(110) Christensen, U. B.; Pedersen, E. B. Intercalating nucleic acids containing insertions of 1-*O*-(1-pyrenylmethyl)glycerol: stabilisation of dsDNA and discrimination of DNA over RNA. *Nucleic Acids Res.* **2002**, *30*, 4918-4925.

(111) Sau, S. P.; Hrdlicka, P. J. C2′-pyrene-functionalized triazole-linked DNA: universal DNA/RNA hybridization probes. *J. Org. Chem.* **2011**, *77*, 5-16.

(112) Beuck, C.; Singh, I.; Bhattacharya, A.; Hecker, W.; Parmar, V. S.; Seitz, O.; Weinhold, E. Polycyclic aromatic DNA-base surrogates: high-affinity binding to an adenine-specific base-flipping DNA methyltransferase. *Angew. Chem. Int. Ed.* **2003**, *42*, 3958-3960.

(113) Jiang, Y. L.; Kwon, K.; Stivers, J. T. Turning on uracil-DNA glycosylase using a pyrene nucleotide switch. *J. Biol. Chem.* **2001**, *276*, 42347-42354.

(114) Shelbourne, M.; Chen, X.; Brown, T.; El-Sagheer, A. H. Fast copper-free click DNA ligation by the ring-strain promoted alkyne-azide cycloaddition reaction. *Chem. Commun.* **2011**, *47*, 6257-6259.

(115) Kočalka, P.; El-Sagheer, A. H.; Brown, T. Rapid and efficient DNA strand cross-linking by click chemistry. *ChemBioChem* **2008**, *9*, 1280-1285.

(116) Ortega, J.-A.; Blas, J. R.; Orozco, M.; Grandas, A.; Pedroso, E.; Robles, J. Binding affinities of oligonucleotides and PNAs containing phenoxazine and G-Clamp cytosine analogues are unusually sequence-dependent. *Org. Lett.* **2007**, *9*, 4503-4506.

(117) Sandin, P.; Borjesson, K.; Li, H.; Martensson, J.; Brown, T.; Wilhelmsson, L. M.; Albinsson, B. Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Res.* **2008**, *36*, 157-167.

(118) Telser, J.; Cruickshank, K. A.; Morrison, L. E.; Netzel, T. L. Synthesis and characterization of DNA oligomers and duplexes containing covalently attached molecular labels: comparison of biotin, fluorescein, and pyrene labels by thermodynamic and optical spectroscopic measurements. *J. Am. Chem. Soc.* **1989**, *111*, 6966-6976.

(119) Wilson, J. N.; Cho, Y.; Tan, S.; Cuppoletti, A.; Kool, E. T. Quenching of fluorescent nucleobases by neighboring DNA: the “insulator” concept. *ChemBioChem* **2008**, *9*, 279-285.

(120) Huppert, J. L. Structure, location and interactions of G-quadruplexes. *FEBS J.* **2010**, *277*, 3452-3458.

(121) Bochman, M. L.; Paeschke, K.; Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **2012**, *13*, 770-780.

(122) Jovanovic, S. V.; Simic, M. G. One-electron redox potentials of purines and pyrimidines. *J. Phys. Chem.* **1986**, *90*, 974-978.

(123) Narayanan, M.; Kodali, G.; Xing, Y.; Hawkins, M. E.; Stanley, R. J. Differential Fluorescence Quenching of Fluorescent Nucleic Acid Base Analogues by Native Nucleic Acid Monophosphates. *J. Phys. Chem. B* **2010**, *114*, 5953-5963.

(124) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. Nucleobase-specific quenching of fluorescent dyes. 1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *J. Phys. Chem.* **1996**, *100*, 5541-5553.

(125) Okamoto, A.; Ichiba, T.; Saito, I. Pyrene-labeled oligodeoxynucleotide probe for detecting base insertion by excimer fluorescence emission. *J. Am. Chem. Soc.* **2004**, *126*, 8364-8365.

(126) Okamoto, A.; Kanatani, K.; Saito, I. Pyrene-labeled base-discriminating fluorescent DNA probes for homogeneous SNP typing. *J. Am. Chem. Soc.* **2004**, *126*, 4820-4827.

(127) El-Sagheer, A. H.; Brown, T. Click chemistry with DNA. *Chem. Soc. Rev.* **2010**, *39*, 1388-1405.

(128) El-Sagheer, A. H.; Brown, T. Combined nucleobase and backbone modifications enhance DNA duplex stability and preserve biocompatibility. *Chem. Sci.* **2014**, *5*, 253-259.

(129) Paredes, E.; Evans, M.; Das, S. R. RNA labeling, conjugation and ligation. *Methods* **2011**, *54*, 251-259.

(130) Shelbourne, M.; Brown, T.; El-Sagheer, A. H.; Brown, T. Fast and efficient DNA crosslinking and multiple orthogonal labelling by copper-free click chemistry. *Chem. Commun.* **2012**, *48*, 11184-11186.

(131) Goodwin, J. T.; Lynn, D. G. Template-directed synthesis: use of a reversible reaction. *J. Am. Chem. Soc.* **1992**, *114*, 9197-9198.

(132) Haralambidis, J.; Lagniton, L.; Tregear, G. W. The preparation of enzyme-labelled oligonucleotides by reductive amination. *Biorg. Med. Chem. Lett.* **1994**, *4*, 1005-1010.

(133) Raindlová, V.; Pohl, R.; Hocek, M. Synthesis of aldehyde-linked nucleotides and DNA and their bioconjugations with lysine and peptides through reductive amination. *Chem. Eur. J.* **2012**, *18*, 4080-4087.

(134) Hsu, J.-L.; Huang, S.-Y.; Chow, N.-H.; Chen, S.-H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **2003**, *75*, 6843-6852.

(135) Hsu, J.-L.; Huang, S.-Y.; Shiea, J.-T.; Huang, W.-Y.; Chen, S.-H. Beyond Quantitative Proteomics: Signal Enhancement of the a1 Ion as a Mass Tag for Peptide Sequencing Using Dimethyl Labeling. *J. Proteome Res.* **2005**, *4*, 101-108.

(136) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.* **1997**, *62*, 7512-7515.

(137) Meyer, S. D.; Schreiber, S. L. Acceleration of the Dess-Martin oxidation by water. *J. Org. Chem.* **1994**, *59*, 7549-7552.

(138) Maurin, A.; Robert, M. Noncovalent Immobilization of a Molecular Iron-Based Electrocatalyst on Carbon Electrodes for Selective, Efficient CO2-to-CO Conversion in Water. *J. Am. Chem. Soc.* **2016**, *138*, 2492-2495.

(139) Safronov, A. V.; Shlyakhtina, N. I.; Everett, T. A.; VanGordon, M. R.; Sevryugina, Y. V.; Jalisatgi, S. S.; Hawthorne, M. F. Direct Observation of Bis(dicarbollyl)nickel Conformers in Solution by Fluorescence Spectroscopy: An Approach to Redox-Controlled Metallacarborane Molecular Motors. *lnorg. Chem.* **2014**, *53*, 10045-10053.

(140) A, G.; EV, B.; L, W.; HE, S.; AR, S.; LL, E.; A, G.; S, H.; KT, D. New concepts of fluorescent probes for specific detection of DNA sequences: bis-modified oligonucleotides in excimer and exciplex detection. *Libyan J. Med.* **2009**, *4*, 8.

(141) Bichenkova, E. V.; Savage, H. E.; Sardarian, A. R.; Douglas, K. T. Target-assembled tandem oligonucleotide systems based on exciplexes for detecting DNA mismatches and single nucleotide polymorphisms. *Biochem. Biophys. Res. Commun.* **2005**, *332*, 956-964.

(142) Kholodar, S. A.; Novopashina, D. S.; Meschaninova, M. I.; Venyaminova, A. G. Multipyrene tandem probes for point mutations detection in DNA. *J. Nucleic Acids* **2013**, *2013*, 12.

(143) Martí, A. A.; Li, X.; Jockusch, S.; Li, Z.; Raveendra, B.; Kalachikov, S.; Russo, J. J.; Morozova, I.; Puthanveettil, S. V.; Ju, J.; Turro, N. J. Pyrene binary probes for unambiguous detection of mRNA using time-resolved fluorescence spectroscopy. *Nucleic Acids Res.* **2006**, *34*, 3161-3168.

(144) Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, *31*, 3406-3415.

(145) Ahlborn, C.; Siegmund, K.; Richert, C. Isostable DNA. *J. Am. Chem. Soc.* **2007**, *129*, 15218-15232.

(146) Kottysch, T.; Ahlborn, C.; Brotzel, F.; Richert, C. Stabilizing or destabilizing oligodeoxynucleotide duplexes containing single 2′-deoxyuridine residues with 5-alkynyl substituents. *Chem. Eur. J.* **2004**, *10*, 4017-4028.

(147) Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L. A. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 3746-3750.

(148) Mergny, J.-L.; Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **2003**, *13*, 515-537.

(149) Owczarzy, R. Melting temperatures of nucleic acids: discrepancies in analysis. *Biophys. Chem.* **2005**, *117*, 207-215.

(150) Petersheim, M.; Turner, D. H. Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry* **1983**, *22*, 256-263.

(151) Puglisi, J. D.; Tinoco Jr, I.: [22] absorbance melting curves of RNA. In *Methods Enzymol.*; James E. Dahlberg, J. N. A., Ed.; Academic Press, 1989; Vol. Volume 180; pp 304-325.

(152) Mooers, B. H. M. Crystallographic studies of DNA and RNA. *Methods* **2009**, *47*, 168-176.

(153) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* **1958**, *181*, 662-666.

(154) Giannoni, G.; Padden, F. J.; Keith, H. D. Crystallization of DNA from dilute solution *Proc. Natl. Acad. Sci. U. S. A.* **1969**, *62*, 964-971.

(155) Kim, S. H.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Sussman, J. L.; Wang, A. H. J.; Seeman, N. C.; Rich, A. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* **1974**, *185*, 435-440.

(156) Robertus, J. D.; Ladner, J. E.; Finch, J. T.; Rhodes, D.; Brown, R. S.; Clark, B. F. C.; Klug, A. Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* **1974**, *250*, 546-551.

(157) Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **1979**, *282*, 680-686.

(158) Timsit, Y.; Moras, D.: [20] Crystallization of DNA. In *Methods Enzymol.*; Academic Press, 1992; Vol. Volume 211; pp 409-429.

(159) Kennard, O.; Hunter, W. N. Oligonucleotide structure: a decade of results from single crystal X-ray diffraction studies. *Q. Rev. Biophys.* **1989**, *22*, 327-379.

(160) Shakked, Z.; Rabinovich, D. The effect of the base sequence on the fine structure of the DNA double helix. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 159-195.

(161) Doucet, J.; Benoit, J. P.; Cruse, W. B. T.; Prange, T.; Kennard, O. Coexistence of A- and B-form DNA in a single crystal lattice. *Nature* **1989**, *337*, 190-192.

(162) Shakked, Z.; Guerstein-Guzikevich, G.; Eisenstein, M.; Frolow, F.; Rabinovich, D. The conformation of the DNA double helix in the crystal is dependent on its environment. *Nature* **1989**, *342*, 456-460.

(163) Dickerson, R. E.; Goodsell, D. S.; Kopka, M. L.; Pjura, P. E. The effect of crystal packing on oligonucleotide double helix structure. *J. Biomol. Struct. Dyn.* **1987**, *5*, 557-579.

(164) Timsit, Y.; Westhof, E.; Fuchs, R. P. P.; Moras, D. Unusual helical packing in crystals of DNA bearing a mutation hot spot. *Nature* **1989**, *341*, 459-462.

(165) Pohl, F. M.; Jovin, T. M. Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.* **1972**, *67*, 375-396.

(166) Thamann, T. J.; Lord, R. C.; Wang, A. H. J.; Rich, A. The high salt form of poly(dG-dC)·poly(dG-dC) is left-handed Z-DNA: Raman spectra of crystals and solutions. *Nucleic Acids Res.* **1981**, *9*, 5443-5458.

(167) Doudna, J. A.; Lorsch, J. R. Ribozyme catalysis: not different, just worse. *Nat. Struct. Mol. Biol.* **2005**, *12*, 395-402.

(168) Hendrix, D. K.; Brenner, S. E.; Holbrook, S. R. RNA structural motifs: building blocks of a modular biomolecule. *Q. Rev. Biophys.* **2005**, *38*, 221-243.

(169) Jacobs, S. A.; Podell, E. R.; Cech, T. R. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nat. Struct. Mol. Biol.* **2006**, *13*, 218-225.

(170) MacRae, I. J.; Zhou, K.; Doudna, J. A. Structural determinants of RNA recognition and cleavage by Dicer. *Nat. Struct. Mol. Biol.* **2007**, *14*, 934-940.

(171) MacRae, I. J.; Zhou, K.; Li, F.; Repic, A.; Brooks, A. N.; Cande, W. Z.; Adams, P. D.; Doudna, J. A. Structural basis for double-stranded RNA processing by Dicer. *Science* **2006**, *311*, 195-198.

(172) Noller, H. F. RNA structure: reading the ribosome. *Science* **2005**, *309*, 1508-1514.

(173) Stahley, M. R.; Adams, P. L.; Wang, J.; Strobel, S. A. Structural metals in the group I intron: a ribozyme with a multiple metal ion core. *J. Mol. Biol.* **2007**, *372*, 89-102.

(174) Taylor, D. W.; Ma, E.; Shigematsu, H.; Cianfrocco, M. A.; Noland, C. L.; Nagayama, K.; Nogales, E.; Doudna, J. A.; Wang, H.-W. Substrate-specific structural rearrangements of human Dicer. *Nat. Struct. Mol. Biol.* **2013**, *20*, 662-670.

(175) Dailey, M. M.; Hait, C.; Holt, P. A.; Maguire, J. M.; Meier, J. B.; Miller, M. C.; Petraccone, L.; Trent, J. O. Structure-based drug design: from nucleic acid to membrane protein targets. *Exp. Mol. Pathol.* **2009**, *86*, 141-150.

(176) Sheng, J.; Gan, J.; Huang, Z. Structure-based DNA-targeting strategies with small molecule ligands for drug discovery. *Med. Res. Rev.* **2013**, *33*, 1119-1173.

(177) Taylor, G. The phase problem. *Acta Crystallogr. Sect. D* **2003**, *59*, 1881-1890.

(178) Rossmann, M. Molecular replacement - historical background. *Acta Crystallogr. Sect. D* **2001**, *57*, 1360-1366.

(179) Beck, T.; Gruene, T.; Sheldrick, G. M. The magic triangle goes MAD: experimental phasing with a bromine derivative. *Acta Crystallogr. Sect. D* **2010**, *66*, 374-380.

(180) Boggon, T. J.; Shapiro, L. Screening for phasing atoms in protein crystallography. *Structure* **2000**, *8*, R143-R149.

(181) Dauter, Z.; Adamiak, D. A. Anomalous signal of phosphorus used for phasing DNA oligomer: importance of data redundancy. *Acta Crystallogr. Sect. D* **2001**, *57*, 990-995.

(182) Olieric, V.; Weinert, T.; Finke, A. D.; Anders, C.; Li, D.; Olieric, N.; Borca, C. N.; Steinmetz, M. O.; Caffrey, M.; Jinek, M.; Wang, M. Data-collection strategy for challenging native SAD phasing. *Acta Crystallogr. Sect. D* **2016**, *72*, 421-429.

(183) Peterson, M. R.; Harrop, S. J.; McSweeney, S. M.; Leonard, G. A.; Thompson, A. W.; Hunter, W. N.; Helliwell, J. R. MAD phasing strategies explored with a brominated oligonucleotide crystal at 1.65Å resolution. *J. Synchrotron Radiat.* **1996**, *3*, 24-34.

(184) Rose, J. P.; Wang, B.-C.; Weiss, M. S. Native SAD is maturing. *IUCrJ* **2015**, *2*, 431-440.

(185) Todd, A. K.; Adams, A.; Powell, H. R.; Wilcock, D. J.; Thorpe, J. H.; Lausi, A.; Zanini, F.; Wakelin, L. P. G.; Cardin, C. J. Determination by MAD-DM of the structure of the DNA duplex d[ACGTACG(5-BrU)]2 at 1.46 A and 100 K. *Acta Crystallogr. Sect. D* **1999**, *55*, 729-735.

(186) Xiuqi, S.; Mary, E. P.; Leigh Ann, L.; Qi, G.; Craig, O.; Bernard, P. R.; Christiane, G.-J.; Angus, P. W.; Loren Dean, W. Effects of cationic charge on three-dimensional structures of intercalative complexes structure of a bis-intercalated DNA complex solved by MAD phasing. *Curr. Med. Chem.* **2000**, *7*, 59-71.

(187) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D* **2010**, *66*, 133-144.

(188) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D* **2011**, *67*, 235-242.

(189) Sheldrick, G. A short history of SHELX. *Acta Crystallogr. Sect. A* **2008**, *64*, 112-122.

(190) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D* **2010**, *66*, 486-501.

(191) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658-674.

(192) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D* **2012**, *68*, 352-367.

(193) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. The structure of an oligo(dA)·oligo(dT) tract and its biological implications. *Nature* **1987**, *330*, 221-226.

(194) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Crystal structure analysis of a complete turn of B-DNA. *Nature* **1980**, *287*, 755-758.

(195) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78*, 2179-2183.

(196) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37*, 1713-1725.

(197) Ranjbar, B.; Gill, P. Circular dichroism techniques: biomolecular and nanostructural analyses- a review. *Chem. Biol. Drug Des.* **2009**, *74*, 101-120.

(198) Freidzon, A. Y.; Valiev, R. R.; Berezhnoy, A. A. Ab initio simulation of pyrene spectra in water matrices. *RSC Adv.* **2014**, *4*, 42054-42065.

(199) Yoshinaga, T.; Hiratsuka, H.; Tanizaki, Y. Electronic absorption spectra of pyrene and hydropyrenes. *Bull. Chem. Soc. Jpn.* **1977**, *50*, 3096-3102.

(200) Diederichs, K.; Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Mol. Biol.* **1997**, *4*, 269-275.