# UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

School of Chemistry


# Rigid Fluorescent Monomers For Incorporation Into Synthetic DNA

**by**


**Lucy Marie Hall**


Thesis for the degree of Doctor of Philosophy

**September 2011**

**UNIVERSITY OF SOUTHAMPTON**

# ABSTRACT

**FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES**

## DOCTOR OF PHILOSOPHY

**RIGID FLUORESCENT MONOMERS FOR INCORPORATION INTO SYNTHETIC DNA**

**By Lucy Marie Hall**

Fluorescence is used extensively to probe DNA structure and dynamics. Natural nucleic acids possess no detectable fluorescence; therefore chemically modified fluorescent oligonucleotides are used. It is possible to study nucleic acid dynamics and interactions by inserting a donor and acceptor chromophore into the DNA. Fluorescence resonance energy transfer (FRET) can then be used to measure distances between the two fluorophores.

In this thesis is presented the synthesis and properties of a range of rigid fluorescent monomers in synthetic DNA. These monomers were designed to maximise fluorescence, whilst maintaining a rigid linker to the DNA to fix the dye precisely in space to enable accurate FRET distance measurements to be made. The rigidity is also important to minimise dye-DNA interactions and to reduce fluorescence quenching. A fluorescent cytosine base analogue, CPP, was initially prepared followed by the cyanine dye analogues, Cy3dT and Cy5dT, which have higher extinction coefficients (112,000 $M^{-1}cm^{-1}$ and 208,000 $M^{-1}cm^{-1}$ respectively) than CPP (14,000 $M^{-1}cm^{-1}$).

The properties of these monomers were investigated and they were shown to be viable thymidine analogues in terms of base pair formation. Up to ten Cy3dT additions were successfully incorporated into one duplex, demonstrating the power of this methodology. FRET studies were carried out by inserting combinations of the above dyes into a DNA helix and measuring energy transfer between them. Distance information relating to the position of the dyes in the double-helix was obtained.

Finally a range of highly fluorescent Cy3B monomers are reported (quantum yield ~0.8 compared to 0.05 for Cy3dT) and their use was demonstrated in biochemical/biological applications. The Cy3B ‘molecular toolkit’ was used to create a set of probes for use in real-time PCR amplification. Taqman probes, Molecular Beacon probes, HyBeacon probes and Scorpion primers were synthesised and evaluated.

# CONTENTS

# DECLARATION OF AUTHORSHIP

I, Lucy Marie Hall, declare that the thesis entitled 'Rigid Fluorescent Monomers For Incorporation Into Synthetic DNA' and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission,

**Signed:** ……..

**Date:**…….

## Acknowledgements

I would firstly like to thank my supervisor, Professor Tom Brown, for giving me the opportunity to work on such a varied and interesting research project. Not many people get the chance to work with such colourful compounds so I'm glad I was one of them! I would like to thank him for his support and encouragement throughout my PhD, which never failed, even in times of my own low self-confidence.

Secondly, I would like to thank Dr. Dorcas Brown and everyone at ATDBio Ltd. for their help during my PhD with oligonucleotide synthesis and purification. Thank you to Dr. John Langley, Julie Herniman and Dr. Neil Wells for their help with my mass spectrometry and NMR questions. Thank you also to Neville Wright for his help with the CD-spectrometer.

A big thank you to Brown group members past and present, including Dr. Radha Tailor, Dr. Harvey Lou, Montse Shelbourne and Eli Chen; I have enjoyed working with you all. I would like to say a special thanks to Dr. Afaf El-Sagheer, Dr. Nittaya Gale, Dr. Petr Kočalka, Dr. Simon Gerrard and Dr. James Richardson for the extensive knowledge and expertise they have shared with me. A massive thank you to Marta Gerowska; working on a linked project was invaluable for ideas and support throughout the PhD (it also made it easier not to be the only person struggling with those columns and walking around with fluorescent dye on my face!).

Thank you to all my friends and family who have always been there for a chat (or a moan) whenever I have needed it. You have helped me to put the stressful moments in to perspective.

Finally, I would like to say the biggest thank you of all to my Mum, Dad and Sam. You have supported me unwaveringly for the last four years; through all the highs and lows. I have appreciated your endless patience and I couldn't have got through it without you all.

## Abbreviations

| | |
|---|---|
| 2-AP | 2-Aminopurine |
| A | Adenine |
| abs | Absorbance |
| $A_{max}$ | Absorbance maximum |
| APS | Ammonium persulfate |
| aq. | Aqueous |
| BHQ | Black hole quencher |
| br | Broad (IR) |
| C | Cytosine |
| CD | Circular dichroism |
| CFTR | Cystic fibrosis trans conductance regulatory gene |
| COSY | Correlation spectroscopy |
| CPG | Controlled pore glass |
| CPP | *p*-Chlorophenylpyrrolopyrimidine-2'-deoxyriboside |
| $C_T$ | Cycle threshold (number of cycles required to generate fluorescent signal) |
| d | Doublet (NMR) |
| dA | 2'-Deoxyadenosine |
| DABCYL | 4-((4-(Dimethylamino)phenyl)azo)benzoic acid |
| dC | 2'-Deoxycytidine |
| DCM | Dichloromethane |
| DEPT | Distortionless enhancement through polarisation transfer |
| dG | 2'-Deoxyguanosine |
| DIPEA | *N*,*N*-Diisopropylethylamine |
| DMAP | *N*,*N*-Dimethylaminopyridine |
| DMF | *N*,*N*-Dimethylformamide |
| DMSO | Dimethylsulfoxide |
| DMTCl | 4,4'-Dimethoxytrityl chloride |
| DNA | 2'-Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide triphosphate |
| ds | Double-stranded |
| dT | 2'-Deoxythymidine |
| dU | 2'-Deoxyuridine |
| EDC | 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide |
| EDTA | Ethylenediamine tetraacetic acid |

| | |
|---|---|
| em | Emission |
| $Em_{max}$ | Emission maximum |
| eq | Equivalents |
| ESI | Electrospray ionisation |
| ex | Excitation |
| FAM | Fluorescein addition monomer |
| FRET | Fluorescence resonance energy transfer |
| G | Guanine |
| h | Hours |
| HEG | Hexaethylene glycol |
| HMBC | Heteronuclear multiple bond correlation |
| HMQC | Heteronuclear multiple quantum correlation |
| HRM | High resolution melting |
| HRMS | High resolution mass spectrometry |
| IR | Infrared spectroscopy |
| *J* | Coupling constant (Hz) |
| LRMS | Low resolution mass spectrometry |
| m | Multiplet (NMR), medium (IR) |
| *m/z* | Mass to charge ratio |
| MALDI-TOF | Matrix assisted laser desorption ionisation- time-of-flight |
| min | Minutes |
| Mp | Melting point |
| mt | Mutant-type (DNA sequence) |
| NAP | Nucleic acid purification |
| NHS | *N*-Hydroxysuccinimide |
| NMR | Nuclear magnetic resonance |
| OD | Optical density |
| ODN | Oligodeoxynucleotide |
| Oligo | Oligodeoxynucleotide |
| PAGE | Polyacrylamide gel electrophoresis |
| PCR | Polymerase chain reaction |
| PNA | Peptide nucleic acid |
| q | Quartet (NMR) |
| QF-MB | Quencher-free Molecular Beacon |
| $R_f$ | Retention factor |
| RNA | Ribonucleic acid |

| | |
|---|---|
| RP-HPLC | Reverse-phase high performance liquid chromatography |
| rt | Room temperature |
| rT | Ribothymidine |
| s | Singlet (NMR), strong (IR) |
| ss | Single-stranded |
| SNP | Single nucleotide polymorphism |
| STR | Short tandem repeat |
| T | Thymine |
| t | Triplet (NMR) |
| TBAF | Tetrabutylammonium fluoride |
| TBE | Tris/Borate/EDTA buffer |
| TCA | Trichloroacetic acid |
| THF | Tetrahydrofuran |
| TLC | Thin layer chromatography |
| $T_m$ | DNA melting temperature |
| TEMED | *N*,*N*,*N'*,*N'*-tetramethylethylene diamine |
| TMS | Trimethylsilyl |
| UV | Ultra-violet |
| w | Weak (IR) |
| wt | Wild-type (DNA sequence) |
| α | Alpha |
| β | Beta |
| δ | Chemical shift (in parts per million) |
| ε | Extinction coefficient |
| Φ | Quantum yield |

# Chapter 1

# Introduction

# Chapter 1 – Introduction

## 1.1 DNA structure

2'-Deoxyribonucleic acid (DNA) is a large polymeric molecule containing the genetic information essential for the development and function of nearly all living organisms. It is composed of two polymers which run in opposite, or anti-parallel, direction to one another. Each of the two strands of DNA has a backbone made from alternating phosphate and pentose sugar groups (2'-deoxy-D-ribofuranose). On each sugar, at the 1' position, is one of four heterocyclic bases; the bicyclic purines, adenine (A) and guanine (G), and the monocyclic pyrimidines, cytosine (C) and thymine (T). The repeating unit containing a phosphate, sugar and base is called a nucleotide.

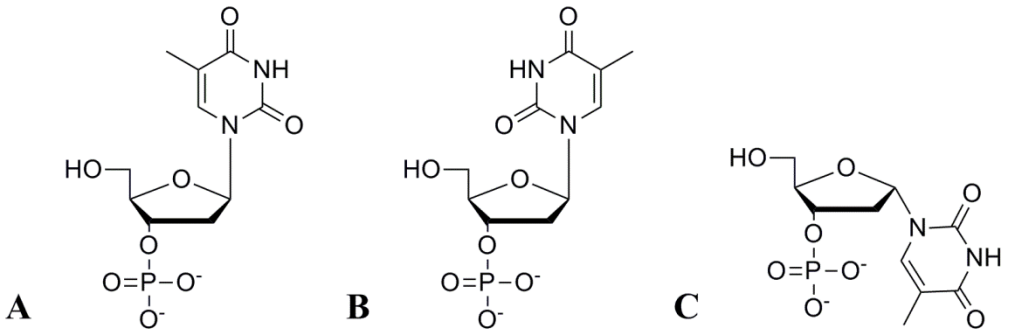



**Figure 1.1.** Thymidine-3'-monophosphate (nucleotide unit) showing the phosphate, sugar and heterocyclic base; (A) β-anomer in *anti*-conformation; (B) β-anomer in *syn*-conformation; (C) α-anomer.

Attachment of the base at the 1' sugar position may be in one of two configurations; either the β-anomer (figure 1.1A) with the substituent on the same face of the sugar as the 5' hydroxyl; or the α-anomer with the substituent on the opposite face (figure 1.1C). The β-anomer occurs in DNA. Restricted rotation about the glycosidic bond (that which joins the base to the 1' position of the sugar) gives rise to either a *syn* (figure 1.1B) or an *anti* (figure 1.1A) conformation. The *anti*-conformation is favoured in DNA both on steric grounds and to enable base-pairing.

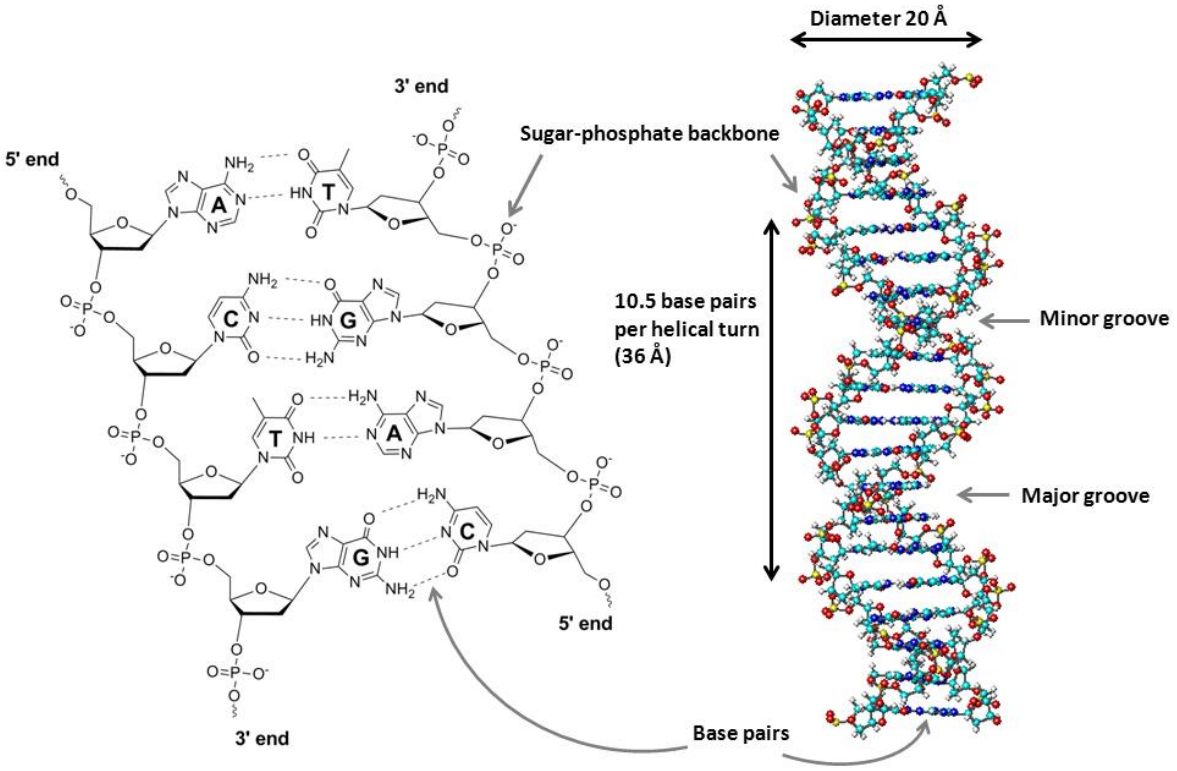



**Figure 1.2.** Base pairing in a section of the DNA double helix and the helical structure of DNA. The base-pairs are at the core of the double-helix in a hydrophobic environment. The phosphodiester groups, which run down the backbone of the helix, are anions at neutral pH and so the DNA helix can exist in a stable form in the aqueous cell environment. The DNA helix has a diameter of 20 Å and a helical turn of 36 Å.

In 1953 significant advances were made in our understanding of the structure of DNA. Based on both x-ray diffraction data acquired by Franklin[1] and Wilkins,[2] and Chargaff's rules, James Watson and Francis Crick elucidated the double-helix structure of DNA.[3] Chargaff's rules specified certain ratios between the DNA bases; A and T are always present in equimolar amounts, as are G and C. Watson and Crick showed that the two strands of DNA are held together by hydrogen bonds between bases on opposite strands; A to T and G to C (figure 1.2). The two strands are complementary and run antiparallel with the 5' end of one next to the 3'end of the other. The strands coil together to form a right-handed double helix with the negatively charged phosphate groups forming the external backbone and the base-pairs forming a hydrophobic core. The phosphodiester groups exist as anions at physiological pH and so the DNA helix can exist in a stable form in the aqueous cell environment.

The principle form of the DNA double helix is B-form (figure 1.2) which has a helical turn of 10-11 base pairs. The B-form has a major groove and smaller minor groove running around the helix for its entire length in which a variety of proteins, drug molecules and dyes can bind. Planar aromatic molecules such as dyes, can intercalate into the DNA duplex between the base-pairs.

## 1.2 Solid-phase oligonucleotide synthesis

The chemical synthesis of small pieces of DNA (referred to as oligonucleotides) is essential for the study of nucleic acids and their function in biological systems. Oligonucleotides are also used in the fields of diagnostic and forensic science where they are in high demand. They are commonly prepared by the quick and convenient solid-phase synthesis method, invented by Bruce Merrifield[4] in which the synthesis is carried out on a solid support. Controlled pore glass (CPG) resin is commonly used. The resin is held between two filters so that during the successive steps of the synthesis cycle, the reagents pass through the solid-support and are then washed off again. In this way reactions can proceed quickly as reagents are used in large excess, in addition, the process can be automated by computer-controlled synthesisers making the technique very convenient.

The synthesis itself takes place in pores on the resin. Resins with a pore size of 1000 Å are often used for the synthesis of oligonucleotides up to 100 bases in length, as smaller pore sizes can become blocked by the oligonucleotide as it increases in length. Resins are normally prepared with a nucleoside loading of 20-30 μmol/g. Higher loadings result in steric hindrance between the growing oligonucleotide chains and so give inefficient synthesis.

Solid-phase oligonucleotide synthesis proceeds in the 3'-5' direction and involves the successive addition of nucleotide phosphoramidites (one per synthesis cycle).[5-8] Each cycle involves a series of steps which are carried out under argon gas; detritylation, activation and coupling, capping and oxidation, which are interposed with acetonitrile washing steps (figure 1.3). Cleavage of the dimethoxytrityl group during the detritylation step liberates the 4,4'-dimethoxytrityl cation. This is orange and absorbs light at 495 nm and so may be utilised to monitor the coupling efficiency of the phosphoramidites.

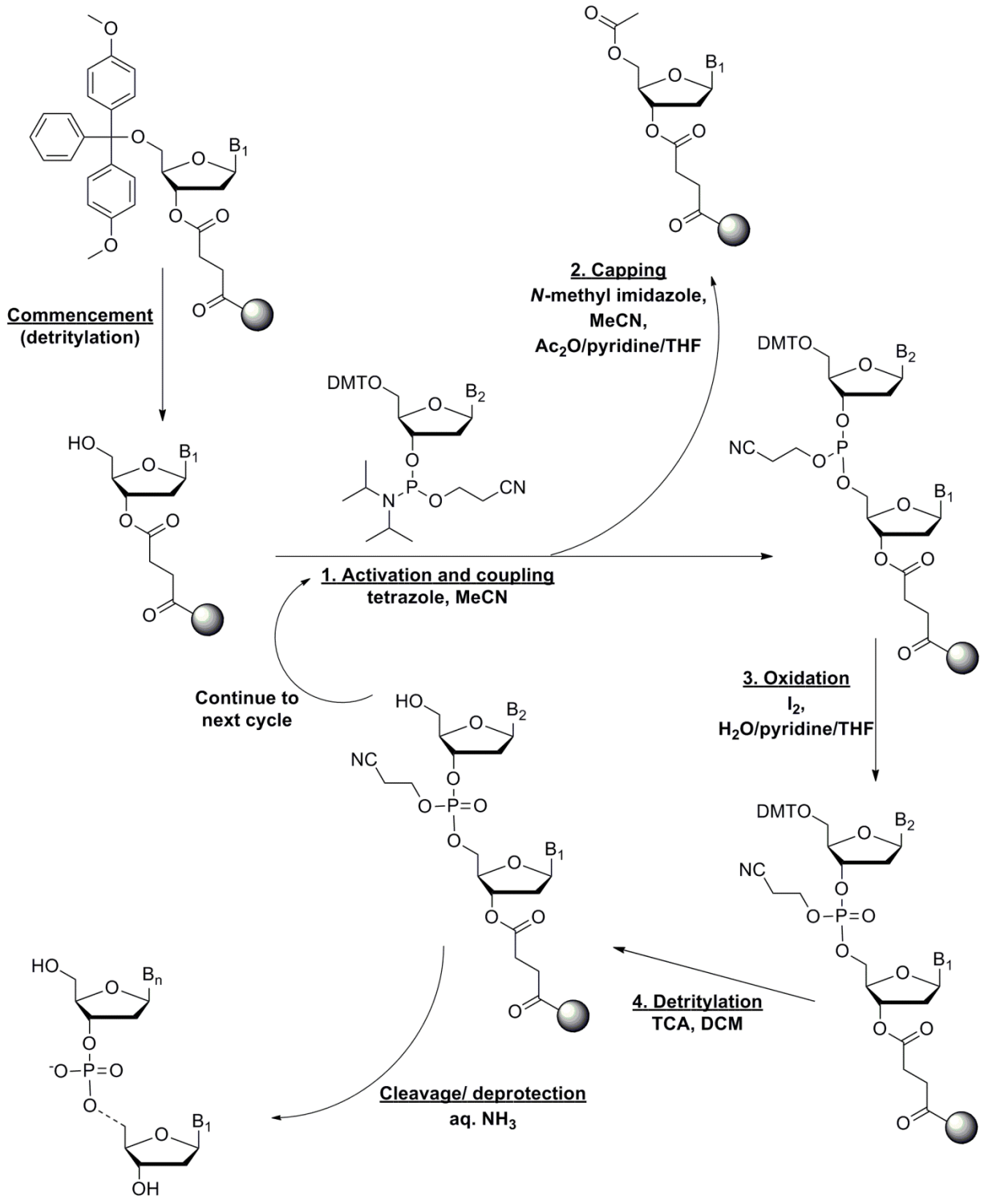



**Figure 1.3.** The phosphoramidite oligonucleotide solid-phase synthesis cycle. (DMT= 4,4'-dimethoxytrityl, $B_1$,$B_2$,$B_n$=any of the DNA bases).

## 1.3 Studying the properties of nucleic acids by UV-Melting[9,10]

DNA bases absorb UV light across the ultraviolet spectrum, with a peak at around 260 nm although the precise details vary between the nucleobases. In single-stranded DNA the absorption of UV light depends on the composition of bases and their stacking interactions. Double stranded DNA absorbs less UV light as π–stacking between the bases increases upon duplex formation; this change is called *hypochromicity.* As a DNA duplex is heated the hydrogen bonds between the bases are broken, the stacking of the bases is disrupted and the UV absorption of the system therefore increases. Hence UV melting can be used for the convenient measurement of DNA duplex stability. To do this a DNA duplex sample (equimolar amounts of each strand) is prepared in aqueous buffer and heated (20-80 ºC). The concerted melting of the duplex into two strands gives a smooth transition in UV absorbance, the mid-point of which is known as the melting temperature ($T_m$) (figure 1.4). Thermodynamic parameters (for example the change in Gibbs energy for the duplex formation) can be determined from UV melting experiments.[11-13]

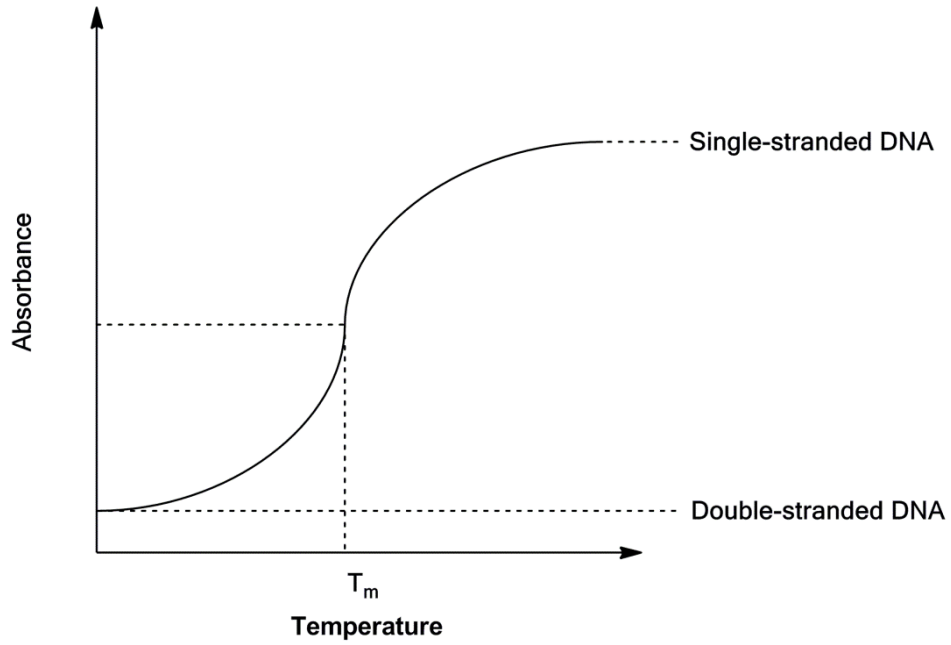



**Figure 1.4.** UV melting curve of a DNA duplex showing the transition from double-stranded DNA to single-stranded DNA as the temperature increases. The mid-point of the transition is the melting temperature ($T_m$).

## 1.4 Fluorescence

### 1.4.1 Principles of fluorescence

Fluorescence is the emission of energy, in the form of light, by a molecule which has absorbed energy of a different wavelength. When light absorption occurs the fluorophore is excited from its ground state, $S_0$, to a higher electronic state, $S_1$ or $S_2$ (figure 1.5). Internal conversion and vibrational relaxation, non-radiative processes, rapidly take the molecule down to the lowest vibrational level of the $S_1$ state. From here the molecule emits the remaining energy in the form of a photon returning it to the original ground state energy. The energy of the emission is generally less than that of the absorbance; therefore fluorescence typically occurs at lower energies or longer wavelengths than the absorbance. This difference in wavelength is known as the Stokes shift, and the extent varies between fluorophores. It is demonstrated in many fluorescent dyes which absorb in the ultra-violet region but fluoresce in the visible region. From the lowest vibrational level of the lowest excited state the molecule can also undergo intersystem-crossing, where the electron changes spin multiplicity, taking it into an excited triplet state, $T_1$. Here the molecule again undergoes vibrational relaxation before releasing energy to drop back to the ground state, a process known as phosphorescence. This transition is 'forbidden' by classic quantum rules and as such the rate is much slower ($10^{-4}$-$10^{-1}$ s compared to fluorescence $10^{-9}$-$10^{-7}$ s, or radiationless decay $10^{-14}$-$10^{-11}$ s). Quenching is an example of radiationless decay which can occur by various processes including collisional quenching, contact quenching and FRET quenching. In all cases the process results in decreased fluorescence exhibited by the fluorophore.

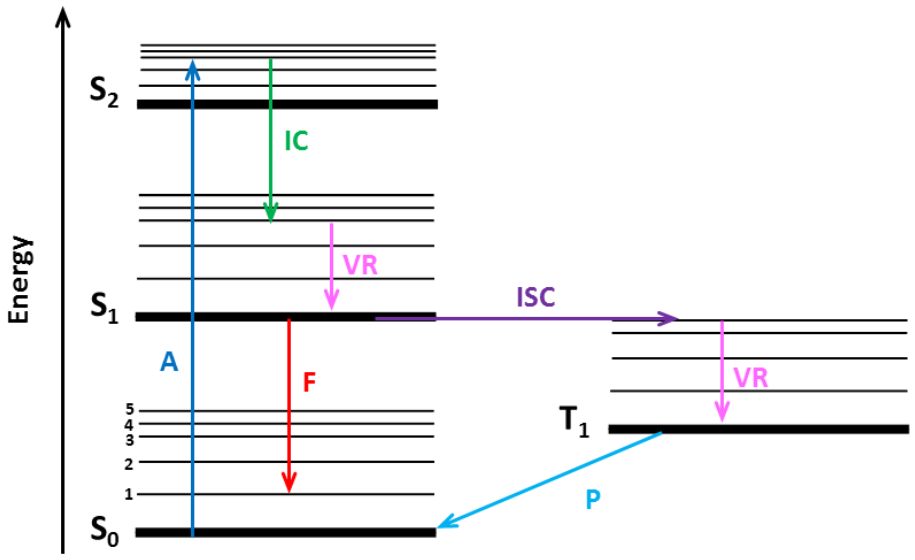



**Figure 1.5.** Energy diagram illustrating transitions between electronic states of a molecule. Absorbance (A, blue, timescale $10^{-15}$ s); internal-conversion (IC, green, timescale $10^{-14}$-$10^{-11}$ s); vibrational-relaxation (VR, pink, timescale $10^{-14}$-$10^{-11}$ s); fluorescence (F, red, timescale $10^{-9}$-$10^{-7}$ s); intersystem-crossing (ISC, purple, timescale $10^{-8}$-$10^{-3}$ s); phosphorescence (P, cyan, timescale $10^{-4}$-$10^{-1}$ s).

### 1.4.2 Extinction coefficient and quantum yield

Two important properties of fluorophores are the quantum yield and extinction coefficient which together determine the effective brightness of the fluorophore. The extinction coefficient describes how effectively the fluorophore absorbs light at a given wavelength relative to its molar concentration and can be calculated by using the Beer-Lambert law, $A=\varepsilon cl$ (where A is UV absorbance at $A_{max}$, c is concentration in M, l is pathlength in cm and ε is extinction coefficient $M^{-1}cm^{-1}$).

The quantum yield refers to the number of photons emitted relative to the number absorbed. In situations where the rate of radiative decay is far higher than the rate of radiationless decay, the quantum yield can be close to 1.0, and in these cases the fluorophore will exhibit very bright emission. Quantum yield can be calculated in relation to other known fluorophores by using the following equation (where QY=quantum yield, I=fluorescence emission area, OD=optical density, η=refractive index of solvent, R=reference and S=sample).

$$QY_S = QY_R \times (I_S/I_R) \times (OD_R/OD_S) \times (\eta_S^2/\eta_R^2) \qquad \text{equation 1}$$

### 1.4.3 Fluorescence Resonance Energy Transfer (FRET)

FRET is the energy transfer between two molecules which are coupled by an interaction between their respective dipoles. For FRET to occur the emission spectrum of a fluorophore (the donor) must overlap with the absorbance spectrum of an acceptor molecule, which may be another fluorophore. A good example of a FRET pair is Cy3 and Cy5 which have a good spectral overlap. The efficiency of the energy transfer process depends strongly upon the distance between the donor and acceptor (r) and the Förster distance ($R_0$) of the donor-acceptor pair (equation 2). The Förster distance is that between the donor and acceptor at which the FRET efficiency is 50 %. The Förster distance, described by equation 3, is unique for every donor-acceptor pair and depends on the relative orientations of the dipoles (κ) (figure 1.6), the refractive index of the medium ($\eta$), the quantum yield of the donor ($Q_D$) and the overlap integral of the donor emission and acceptor absorbance spectra ($J(\lambda)$).

$$E = R_0^6 / (R_0^6 + r^6) \qquad \text{equation 2}$$

$$R_0 = 0.211\ (\kappa^2 \eta^{-4} Q_D J(\lambda))^{1/6} \text{ in Å} \qquad \text{equation 3}$$

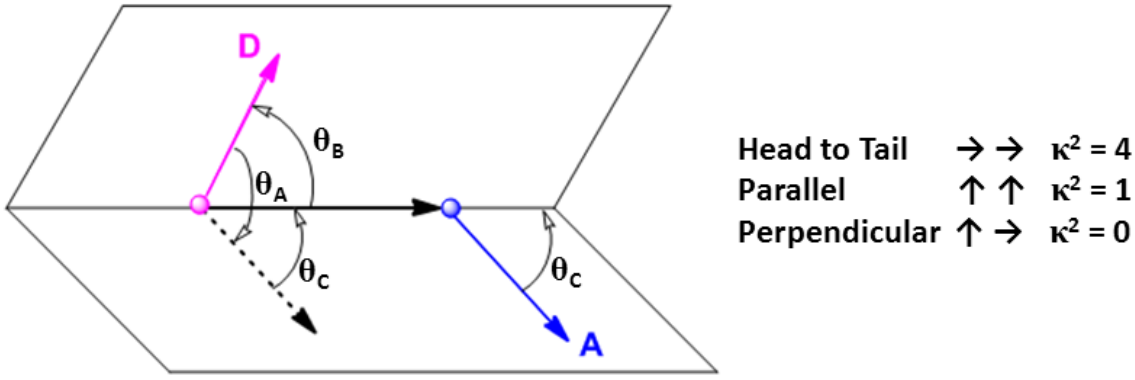



**Figure 1.6.** Dipole orientation angles in FRET between the donor chromophore (pink) and the acceptor chromophore (blue). Using equation 4; a perpendicular orientation between the two dipoles results in a $\kappa^2$ value of 0 leading to FRET efficiency of 0 in these cases.

$$\kappa^2 = (\cos\theta_A - 3\cos\theta_B \cos\theta_C)^2 \qquad \text{equation 4}$$

#### *1.4.3.1 Applications of FRET in the study of DNA*

FRET is an important technique for studying DNA[14] and RNA[15] bending, nucleic acid dynamics and interactions with proteins.[16-19] The distance information which may be elucidated from a FRET study can impart important information about conformation within the biological system studied. For example, in the study by Millar *et al* the conformational dynamics of the complex between Klenow-Fragment DNA polymerase and DNA have been studied. This was done by incorporating a donor fluorophore (tC monomer) into the DNA template and an acceptor (Alexa-555) into the polymerase on a engineered cysteine residue within the KF polymerase 'finger' region. By monitoring the FRET efficiency they found that when a correct nucleotide substrate was bound, the finger closed over the active site; indicated by an increase in FRET efficiency.[17] In recent years FRET has also been increasingly used for single molecule studies.[20-23] Cy3 and Cy5 have been extensively used as a FRET dye pair due to their good spectral overlap and large extinction coefficients.[18,22,24] However in all these cases any uncertainty in the dye position, for example as the result of a flexible dye-DNA linker, could impose large errors onto any conclusions drawn from experimental results.

### 1.4.4 Fluorescent dyes in DNA

Techniques involving fluorescence analysis of a biological system often require the incorporation of one or more fluorescent dyes into either a protein or synthetic DNA or RNA oligonucleotide. Currently there is an extensive market for the commercial synthesis of dyes and dye-labelled oligonucleotides, and there is a large choice of fluorophores and quenchers available to the researcher. Dyes can be obtained with a wide range of spectral and physical properties to fit the requirements of the user. However with so many options available it is important that the fluorophore is carefully selected to gain optimum experimental results.

### *1.4.4.1 Fluorescent Base Analogues*

A key family of fluorophores for incorporation into synthetic DNA are the fluorescent base analogues.[25] These are characterised by an internal fluorophore modification covalently attached to, or replacing, the nucleobase inside the DNA helix. The advantage of base analogues is that they may be incorporated into a DNA system at the site of interest with negligible disruption to the double helix.

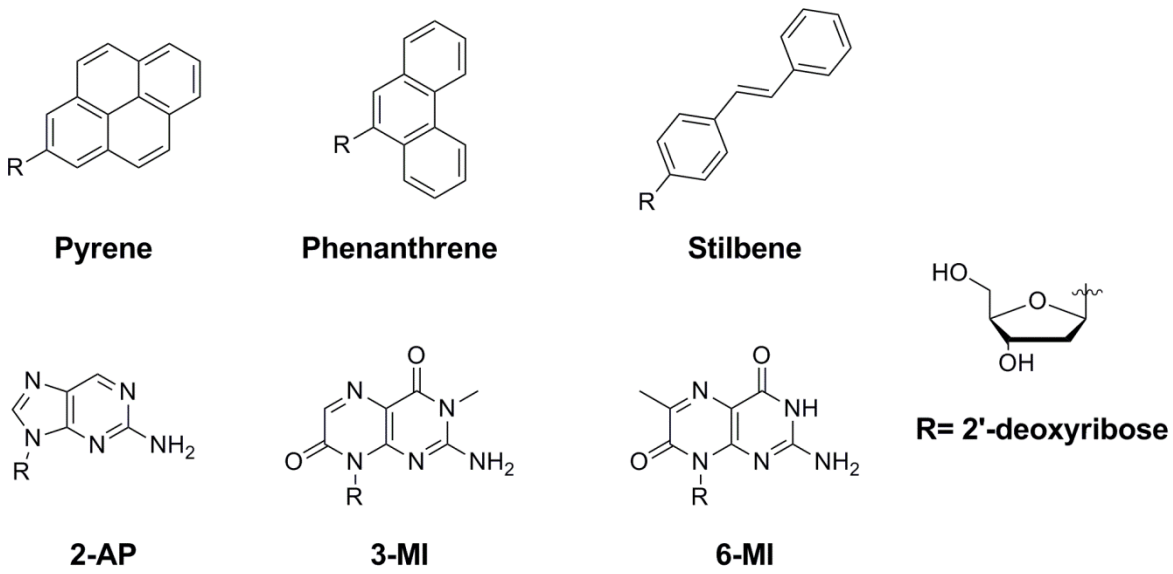



**Figure 1.7.** Chemical structure of; internal polycyclic hydrocarbons for incorporation into DNA (pyrene, phenanthrene and stilbene); fluorescent base analogue 2-aminopurine (2-AP); and the pteridines (3-MI and 6-MI).

Kool *et al* replaced the base moiety with pyrene, phenanthrene and stilbene (figure 1.7) to generate fluorescent base replacements which can stack inside the DNA helix with only minor disruption to the structure.[26-29] However, in most cases it is important for the monomer to be able to form hydrogen bonds to its complementary strand; therefore the fluorescent base analogues were developed to resemble the shape of a natural nucleobase. One of the earliest studied fluorescent base analogues is 2-aminopurine (2-AP); an adenine analogue with capacity to hydrogen bond to thymine to form a stable base pair.[30] More recently the pteridines have been studied. 3-Methylisoxanthopterin (3-MI) and 6-methylisoxanthopterin (6-MI) are two pteridine analogues which replicate the hydrogen

bonding capacity of guanine. The 3-position methyl group of 3-MI however inhibits the hydrogen bonding to cytosine causing a reduction in thermal stability of the duplex approximately equal to that of a single base mismatch.[31,32]

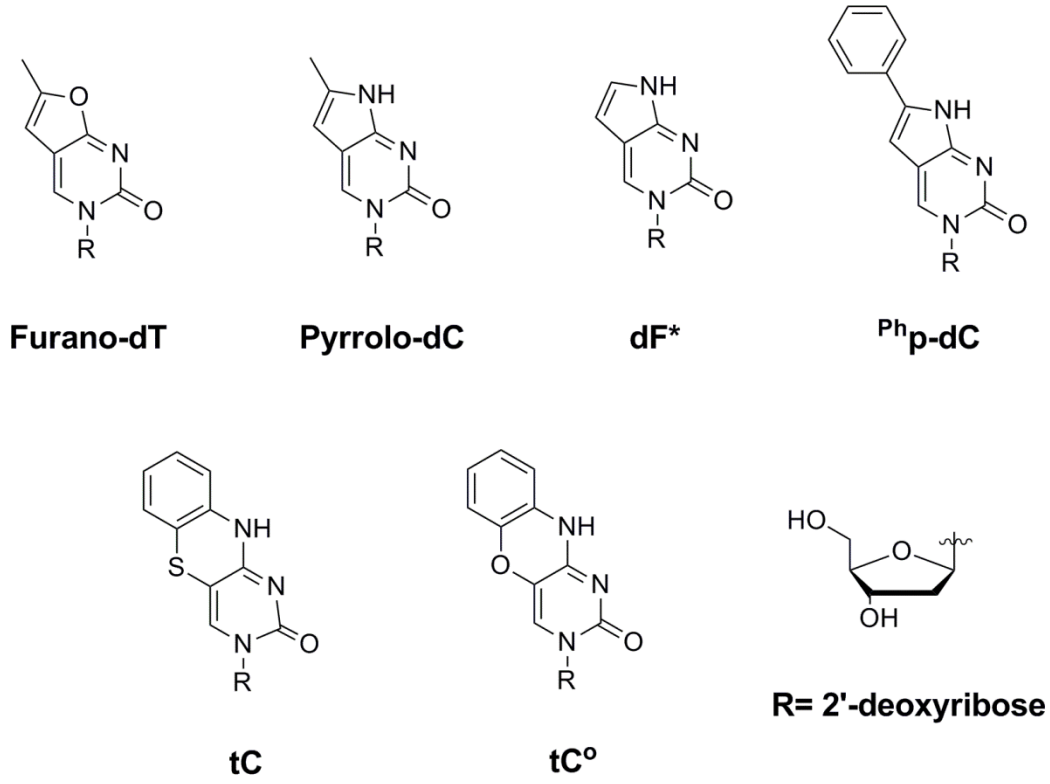



**Figure 1.8.** Chemical structure of fluorescent base analogues from the pyrrolo-dC family (furano-dT, pyrrolo-dC, dF* and $^{Ph}$p-dC) and the tricyclic cytosine analogues tC and tC$^{O}$.

Another widely studied series are the furano and pyrrolo derivatives of thymine and cytosine (figure 1.8). Woo *et al* reported the structure of furano-dT however it was found that in the final stages of DNA synthesis, during treatment with ammonia, it was converted to pyrrolo-dC.[33] In addition conjugation of the oxygen atom into the furano ring means that the nitrogen atom (equivalent to N3 of thymine) does not carry a hydrogen atom, so it is not available for hydrogen bonding to adenine (figure 1.9). Pyrrolo-dC was originally described by Berry *et al*[34] and has subsequently been the subject of much derivatisation. For example Inoue *et al* have investigated the analogue dF*,[35,36] whereas Hudson has focussed intensively on replacing the 6-methyl group, for example with a phenyl group.[37-42] It was found that phenyl-pyrrolo-dC ($^{Ph}$p-dC) increases duplex stability and has approximately 50 % higher quantum yield than pyrrolo-dC (0.32 compared to 0.2 for pyrrolo-dC).[43]

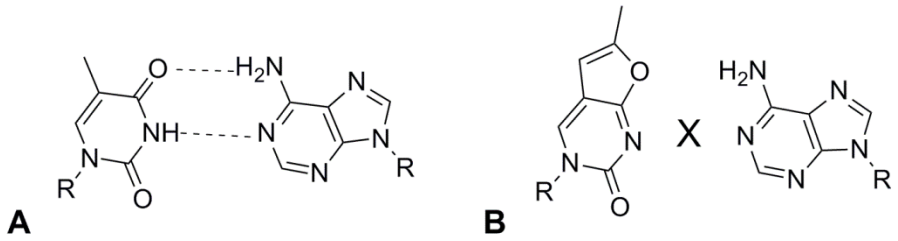



**Figure 1.9.** The TA base pair conformation (A). In comparison, Furano-dT lacks the hydrogen necessary to pair with adenine (B).

The fluorescent base analogues are generally quite sensitive to their immediate environment, showing variation in quantum yield upon incorporation into DNA and on duplex formation. An environment-insensitive base analogue, 1,3-diaza-2-oxophenothiazine (tC) was developed by Lin *et al*[44] and was found to have a quantum yield which is virtually unaffected by its surrounding environment; ranging from 0.16-0.24 in single and double stranded forms.[45-47] In addition, DNA containing the tC monomer was found to adopt the B-DNA structure by CD analysis and, on average, stabilised a 10-mer duplex by ~3 °C.[48] An oxo-homologue, 1,3-diaza-2-oxophenoxazine ($tC^O$),[44] was also developed which has a quantum yield of 0.22 and extinction coefficient of 9000 $M^{-1}cm^{-1}$. It is consequently the most highly fluorescent base analogue available for use in duplex DNA.[49] It was found to stabilise a duplex by ~3 °C on average, and did not perturb the DNA B-form structure. The quantum yield however was found to be far more sensitive to surrounding bases, ranging from 0.14 to 0.41 in the single strand.

### *1.4.4.2 Fluorophores and Quenchers*

A variety of fluorophores are available from numerous commercial sources covering most of the visible spectrum.[50-52] Most fluorophores are traditionally based on fluorescein, rhodamine and cyanine structures and the current range demonstrates a variety of synthetic modifications to these core dyes (figure 1.10). Subtle variations in structure result in altered fluorescence properties, such as absorbance maxima, emission maxima, extinction coefficient and quantum yield, providing a unique dye in each case (table 1.1). Addition of certain functional groups enables insertion of the fluorophore into a biomolecule. For example a fluorophore containing a phosphoramidite reactive group may be incorporated into DNA during automated synthesis. For post-synthetic labelling of DNA or proteins,

fluorophores with carboxylic acid, *N*-hydroxysuccinimide ester, maleimide and hydrazine reactive groups may be used.

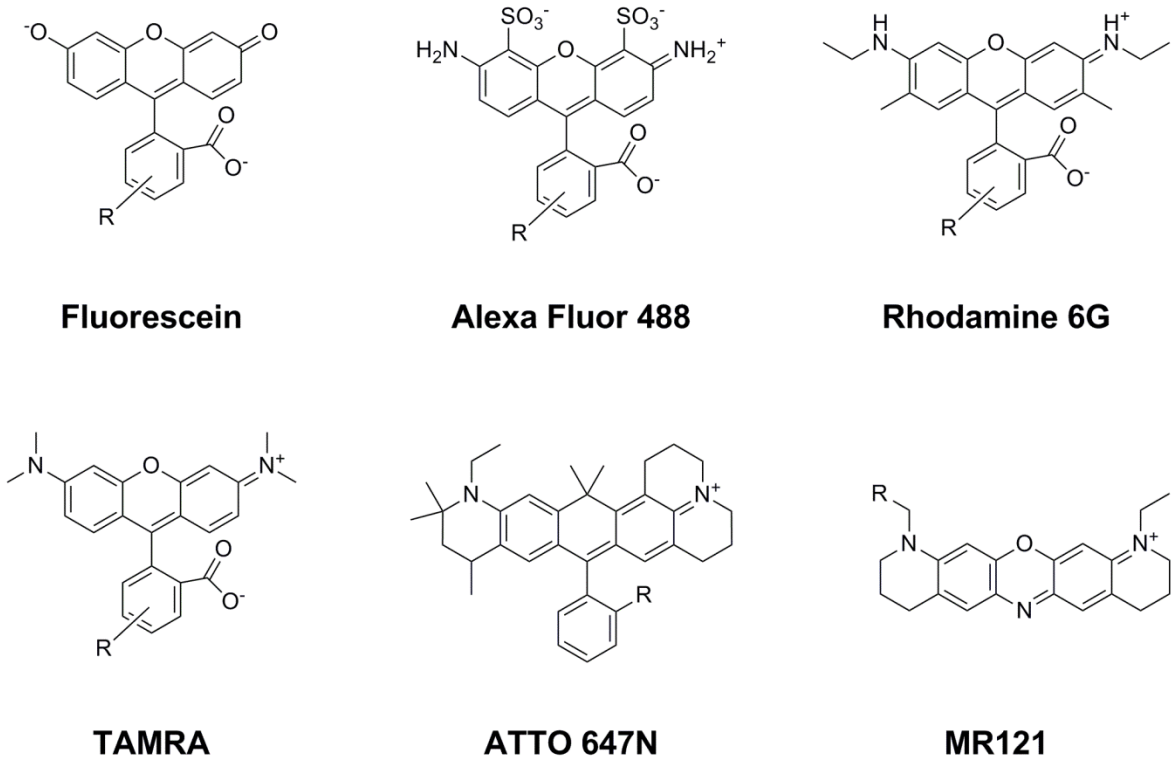



**Figure 1.10.** Structures of some commonly used fluorophores; all have a similar core structure with different synthetic modifications. R= reactive group for labelling of an oligonucleotide (such as phosphoramidite, active ester, carboxylic acid and maleimide groups). An extensive range of fluorophores can be found in reviews of the area.[50-52]

**Table 1.1.** Photophysical properties of some commonly used fluorophores;[a] UV-Vis absorbance maxima ($A_{max}$), fluorescence emission maxima ($Em_{max}$), extinction coefficient (**ε**) and quantum yield (**Φ**).

| Dye | $A_{max}$ (nm) | ε ($M^{-1}cm^{-1}$) | $Em_{max}$ (nm) | Φ |
|---|---|---|---|---|
| Fluorescein* | 490 | 75 000 | 514 | 0.92 |
| Alexa Fluor 488[†] | 495 | 73 000 | 519 | 0.92 |
| Rhodamine 6G | 524[†] | 92 000[†] | 552[†] | 0.95[¥] |
| TAMRA* | 547 | 77 000 | 574 | 0.35 |
| ATTO 647N[‡] | 644 | 150 000 | 669 | 0.65 |

[a] Parameters reported for free dyes.[50] * Data from literature.[53] [†] Data from supplier; Invitrogen. [‡] Data from supplier; ATTO-TEC. [¥] Data from literature.[54]

Quencher molecules are also readily available (figure 1.11). Quenchers should be chosen with a specific corresponding fluorophore in mind as each have unique absorbance properties. For example the subtle variation in structure between BHQ-1 and BHQ-2 results in a shift in absorbance maxima from 534 nm to 580 nm. If it is intended the fluorophore should quenched by FRET, then a quencher with absorbance overlapping the emission of the fluorophore should be selected. If however it is intended the fluorophore should be quenched by contact only then a good contact quencher should be chosen with minimal overlap between the absorbance and emission spectra (figure 1.12).

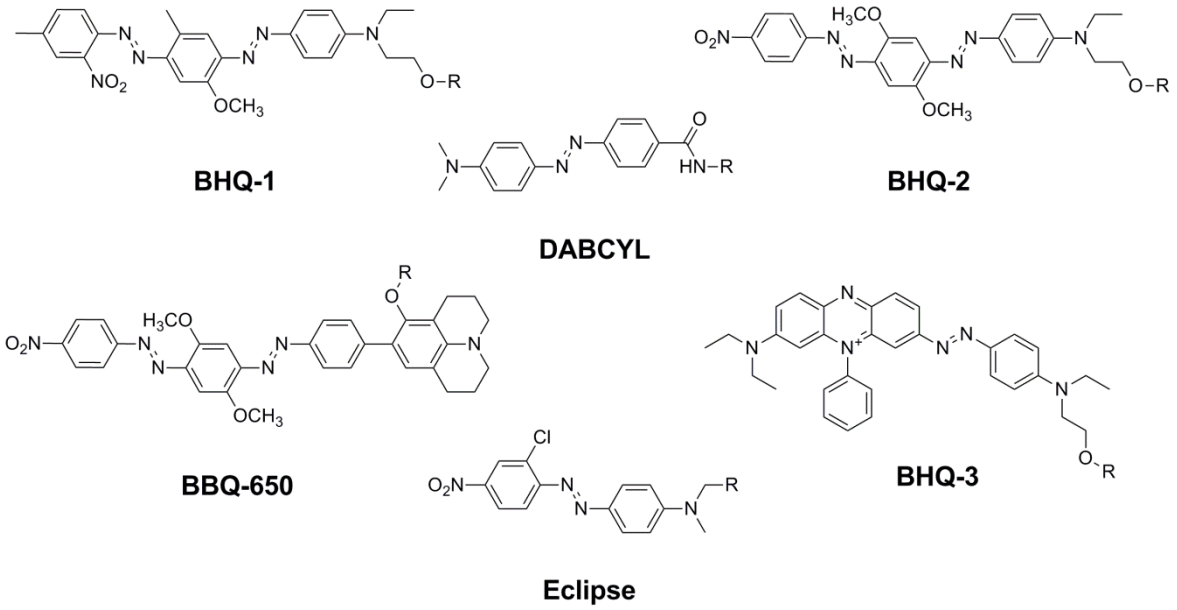



**Figure 1.11.** Structure of some commonly used quenchers in DNA; Black-Hole Quenchers (BHQ), BBQ-650, DABCYL and Eclipse. R= reactive group for labelling of an oligonucleotide (such as phosphoramidite, active ester, carboxylic acid and maleimide groups). A wide range of quenchers can be found from commercial sources such as Berry associates, ATTOtec, Integrated DNA technologies and Molecular Probes.

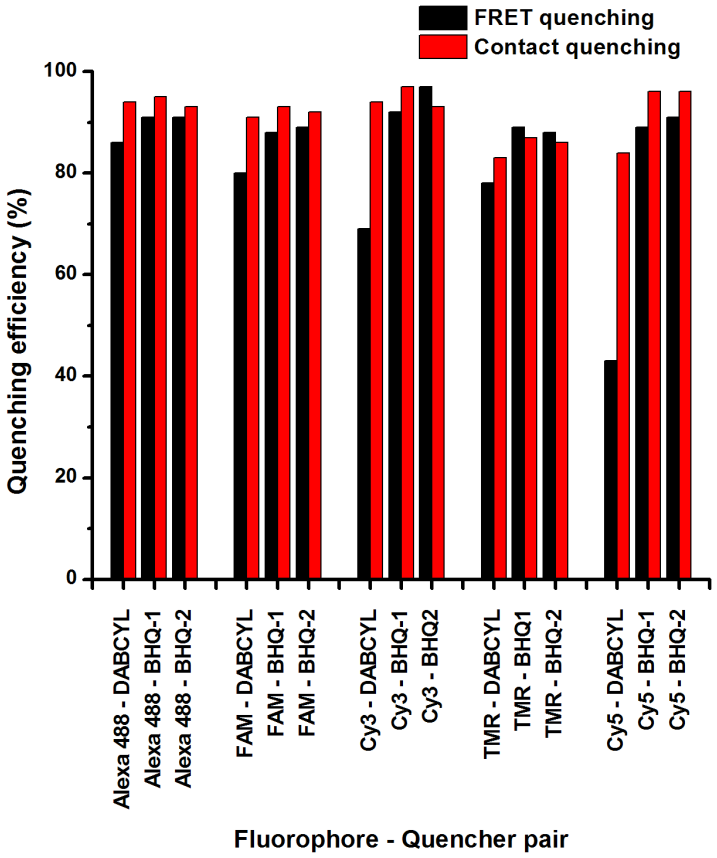



**Figure 1.12.** Quenching efficiency (%) of a selection of fluorophores by quenchers DABCYL ($A_{max}$ 475 nm), BHQ-1 ($A_{max}$ 534 nm) and BHQ-2 ($A_{max}$ 580 nm). Data obtained from literature.[55]

### *1.4.4.3 Cyanine Dyes*

The cyanine dyes are a family of fluorophores which together span most of the visible spectrum making them popular choices for fluorescent labels. The structure of cyanine dyes in general is based on a polymethine chain, containing an odd number of carbons, which links two nitrogen atoms. At each end of the chain is a heterocyclic group, for example an indole (figure 1.13). The most popular cyanine dyes are those with indole groups (the indocarbocyanines) and are generally referred to as the CyDyes; Cy3 (n=1) and Cy5 (n=2).[56,57] Altering the structure of a cyanine dye will alter its photophysical properties; for example an increase in the polymethine linker length results in a red-shift in the absorbance and emission spectra.

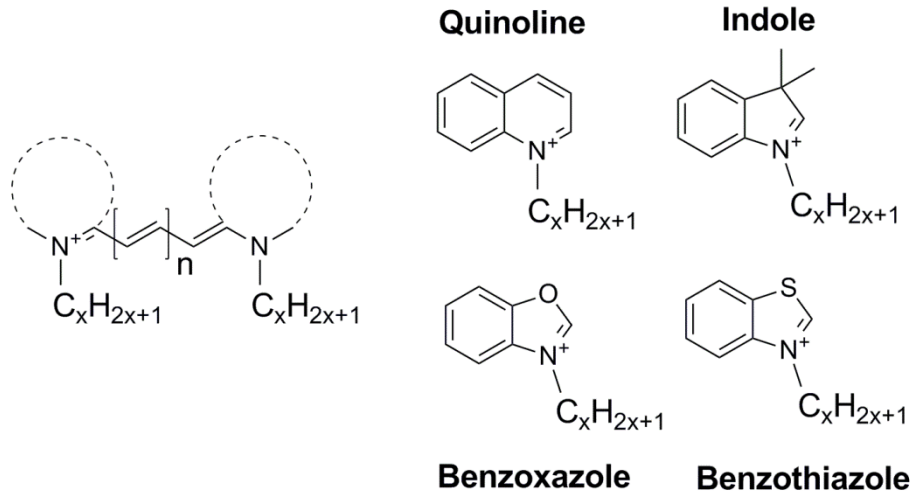



**Figure 1.13.** General structure of the cyanine dyes. Left: core structure of the polymethine cyanine dye where the dotted line represents a heterocyclic structure. Right: Heterocyclic structures commonly found in cyanine dyes.

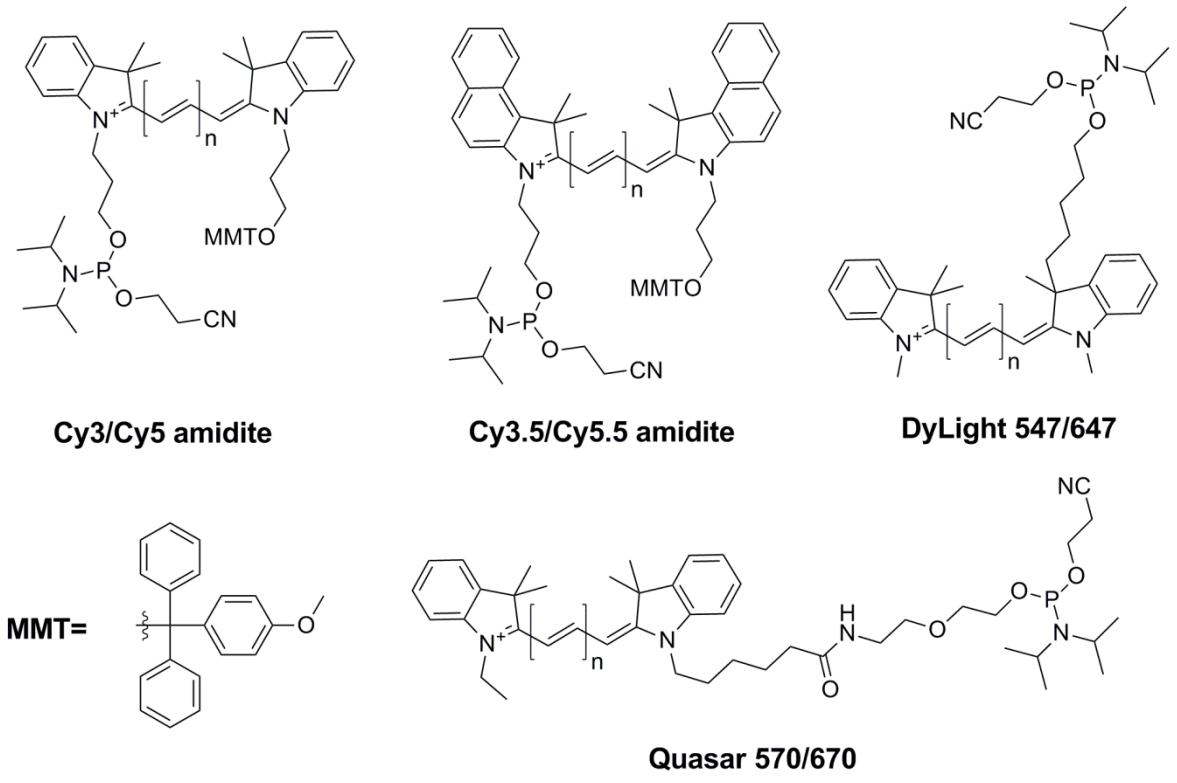



**Figure 1.14.** Phosphoramidite derivates of Cy3 (n=1) and Cy5 (n=2). The commercial alternatives, DyLight 547 (n=1), DyLight 647 (n=2), Quasar 570 (n=1) and Quasar 670 (n=2), contain subtle alterations in the structure.

As described previously, their range in spectral properties makes the CyDyes ideal in FRET applications. In the amidite form the CyDye is usually added to the 5' end of an oligonucleotide during solid-phase synthesis. Subtle variations in the structure of CyDye amidites has given rise to a number of commercially available alternatives which have the same spectral properties as Cy3 and Cy5 (figure 1.14). For DNA addition the use of dye-amidites is preferable due to the ease of incorporation and the resulting high coupling yields obtained. The addition of a sulfonate group onto the dye reduces the tendency of aggregation in aqueous solution which can be an issue at high concentrations.[58,59] The sulfonate group also aids water solubility, which is important for post-synthetic oligonucleotide labelling carried out in aqueous conditions. For amidite incorporation, carried out in acetonitrile, water-solubility is not an issue and so dye-amidites are often synthesised without sulfonate groups.

CyDyes are also popular in various forms for post-synthetic labelling of proteins and nucleic acids (figure 1.15). The NHS-ester and carboxylic acid forms allow labelling of free amino groups (commonly amino-C6-dT residues in DNA). The maleimide reactive group is used for labelling of free thiols, for example those on cysteine residues in proteins. Finally the hydrazine form of the dyes is used for labelling aldehyde and ketone groups. The post-synthetic labelling technique often results in a long flexible linker between the dye and the DNA.

The photophysical properties of the CyDyes can vary dramatically depending on their local environment. In cases where a cyanine dye is added at the 5' end of an oligonucleotide the dye may stack on the end of the DNA holding it in a rigid position.[60] However Levitus showed that in the double-stranded form 18 % of the dyes are un-stacked, but this proportion decreases in the single-stranded form.[61] They found that oligonucleotides with Cy3 on a long linker contained 24 % of Cy3 molecules which were free to rotate around their linkers, whilst the rest were able to fold back and stack into the DNA. It has been regularly demonstrated that the nature of the interaction between cyanine dyes and biomolecules can significantly affect their photophysical properties; stacking of the dye into DNA increases its fluorescence quantum yield considerably but it is very sensitive to DNA sequence,[61-63] so CyDye-DNA conjugates exhibit a large variance in fluorescent output. Based on this, for any quantitative application of cyanine dyes, the knowledge of the structure of the dye-biomolecule linker, and the resulting interactions with DNA are essential.

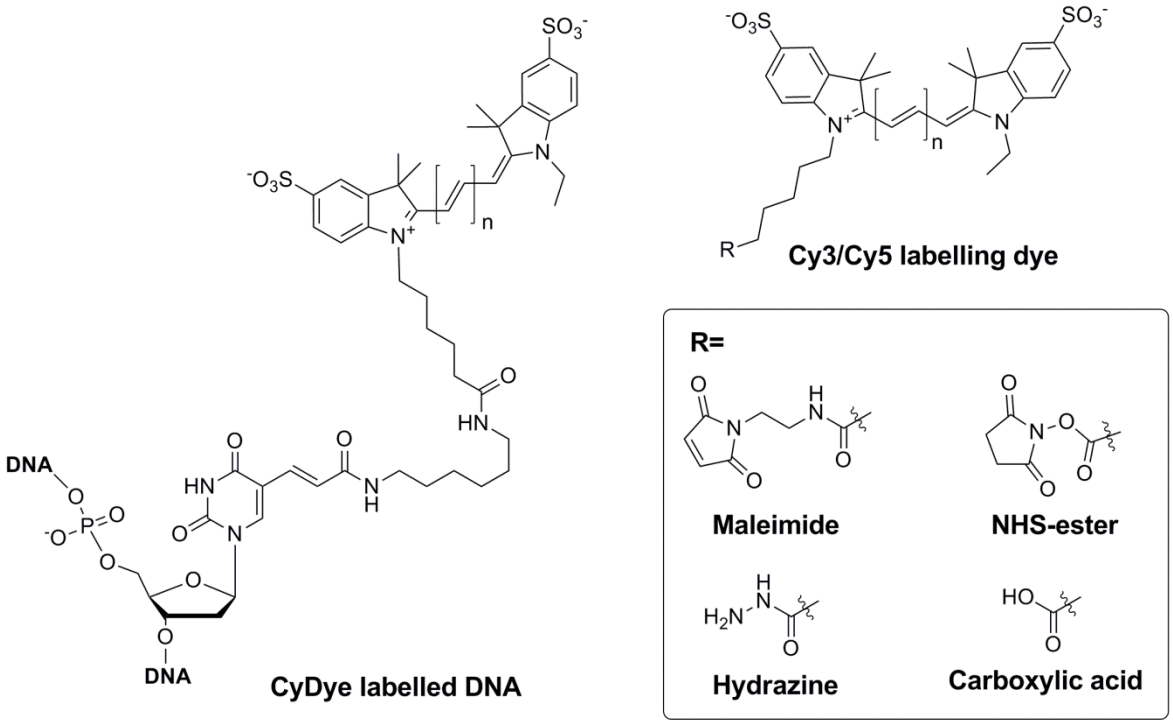



**Figure 1.15.** CyDye labelled amino-C6-dT residue (Glen Research) in DNA and variations of CyDye labelling dyes (Cy3 n=1, Cy5 n=2) where R= maleimide, *N*-hydroxysuccinimide ester (NHS-ester), hydrazine or carboxylic acid functional group.

**Table 1.2.** Photophysical properties of CyDyes;[a] UV-Vis absorbance maxima ($A_{max}$), fluorescence emission maxima ($Em_{max}$), extinction coefficient (**ε**) and quantum yield (**Φ**).

| Dye | $A_{max}$ (nm) | ε ($M^{-1}cm^{-1}$) | $Em_{max}$ (nm) | Φ |
|---|---|---|---|---|
| Cy3-NHS ester | 550 | 150,000 | 570 | >0.15 |
| Cy5-NHS ester | 649 | 250,000 | 670 | >0.28 |
| Cy3B-NHS ester | 559 | 130,000 | 570 | >0.70 |

[a] Parameters reported for free dyes. All data acquired from GE Healthcare, measured in MeOH.

CyDyes suffer from complex non-fluorescent states, largely caused by *cis/trans* isomerisation around the polymethine linker. This gives rise to the changes in fluorescence exhibited between rigid and flexible dye-environments; the mechanism is discussed in detail in chapter 3. This behaviour initiated the development of Cy3B, an analogue of Cy3 which has a rigid backbone along the polymethine linker (figure 1.16).[64] This alteration in structure enhances the fluorescent properties relative to Cy3 (table 1.2) making it one of the most fluorescent dyes currently available. Whilst the syntheses of Cy3 and Cy5 have been well described in the literature, the details regarding Cy3B are less accessible. In

addition Cy3B is not available commercially as a phosphoramidite monomer, but only as a post-synthetic labelling dye which dramatically reduces its practicality.

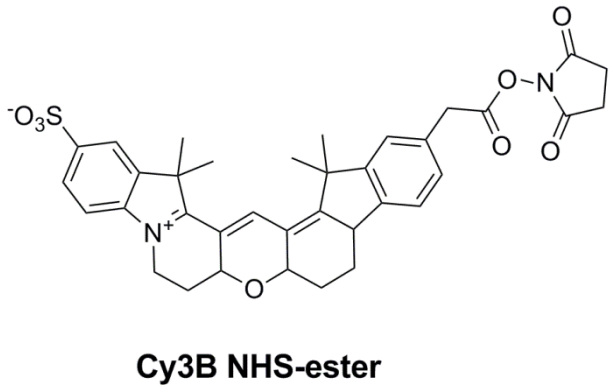



**Figure 1.16.** Chemical structure of Cy3B-NHS ester; a rigid highly fluorescent alternative to Cy3.

## 1.5 PCR

### 1.5.1 Principles of PCR

The polymerase chain reaction (PCR) is a technique used to amplify a specific region of DNA, sometimes from only a few molecules of DNA sample. The technique is therefore key in forensic science, molecular genetics and diagnostics.

The technique requires the design of two PCR primers (short oligonucleotides) which bind at the 3' end of the region of interest on each of two DNA template strands. The template is put through a series of thermal cycles with the addition of the primers, Taq DNA polymerase, deoxyribonucleotide triphosphates (dNTPs) and $Mg^{2+}$. In the first step of the cycle the temperature is increased to around 95 °C to separate the two strands. This is followed by a cooling step to around 55 °C at which point the primers anneal to their template strands. The temperature is then increased to around 72 °C whereupon the Taq DNA polymerase, at its optimum temperature, incorporates the dNTPs to extend the primers along the template DNA to synthesise a complementary sequence; producing two new strands of DNA (figure 1.17).

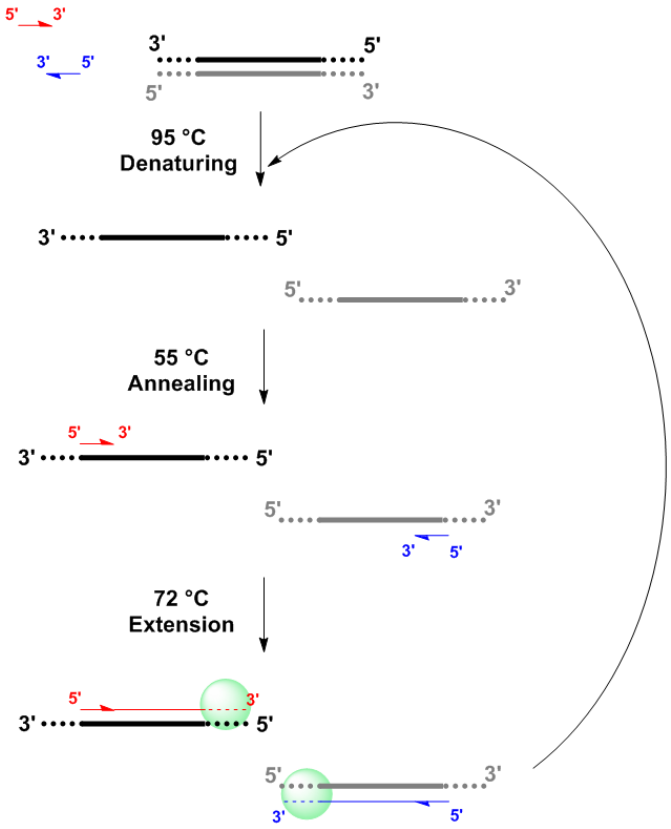



**Figure 1.17.** PCR thermal cycle example; template DNA (thick black and grey) to amplified DNA (thin red and blue) using PCR primers (red and blue arrows) and Taq polymerase (green).

The second cycle follows the same thermal protocol with the newly synthesised amplicon strands now able to act as template strands. The extension step in this instance can only go as far as the locus of the first primer and a sequence of a specific length is generated. All subsequent amplification products are defined in length by the locus of the two primers and, with exponential increase, will far outnumber the original target DNA strand (figure 1.18).

PCR amplification requires only a tiny amount of starting DNA template sample, and hence is a very sensitive technique. Therefore contamination of the PCR reaction must be avoided to prevent misleading results.

Forensic DNA analysis uses PCR to amplify short tandem repeat (STR) regions of human genomic DNA. STRs are regions of DNA containing multiple nucleotide repeats; the lengths of which vary between individuals. The amplified DNA from the sample can then be analysed by gel electrophoresis and compared with an equivalent sample from a suspect. There are many different STR regions in the human genome and therefore the

simultaneous analysis of several STRs can unambiguously link a scene-of-crime sample to a specific person.

In recent years there has been much effort to investigate real-time PCR monitoring techniques in order to improve timescales of forensic and genetic sample analysis.

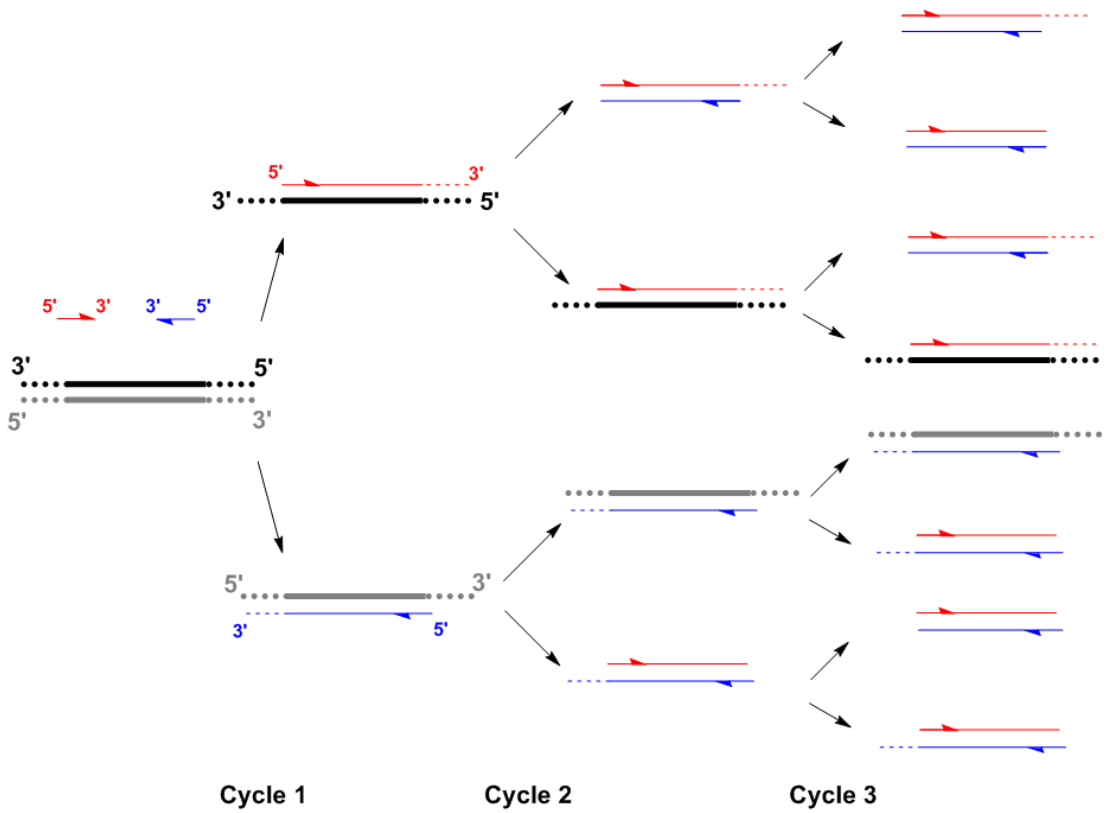



**Figure 1.18.** PCR amplification showing exponential growth from template DNA (thick black and grey) to amplified DNA (thin red and blue) using PCR primers (red and blue arrows). Each cycle involves the steps specified in figure 1.17.

### 1.5.2 Real-time PCR monitoring

PCR may be monitored during thermal cycling by the addition of a fluorescent intercalator, such as SYBR green I (figure 1.19) which becomes up to 1000-fold more fluorescent when it is bound to double-stranded DNA.[65,66] It is commonly used as it has minimal binding to single-stranded DNA and exhibits similar spectral properties to fluorescein. Many genetic analysis real-time PCR platforms have excitation and emission wavelengths designed for the monitoring of fluorescein and so this makes SYBR green I an ideal fluorophore to use. There are drawbacks however; use of an intercalator provides no allelic discrimination. In

addition any primer-dimers or amplicons from mis-priming at an incorrect site will result in signal that is non-specific to the locus of interest.

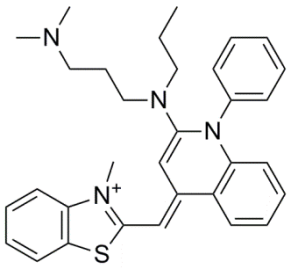


**Figure 1.19.** SYBR green I intercalating dye.

Recently however, fluorescently labelled oligonucleotides designed to 'probe' an amplicon for a region of interest have been widely investigated. Many types of probe have been designed, most of which involved a probe that is quenched (by FRET or collisional quenching) until hybridised to a complementary sequence. Upon hybridisation the fluorophore and quencher are usually separated by distance and so the probe exhibits an increase in fluorescence which may be detected by most commonly used real-time PCR platforms.

In most cases the mechanism of action of the PCR probe allows the DNA sample under investigation to be defined as wild-type or mutant-type, either during PCR or by fluorescence melting analysis immediately following the PCR amplification.

### *1.5.2.1 Taqman probes*

Taqman probes contain a fluorophore and an acceptor dye covalently bound to an oligonucleotide sequence which is designed to be complementary to a locus of interest which is located between two PCR primers. The fluorophore, generally at the 5' end, is quenched by the acceptor, generally at the 3' end, by means of FRET.[67] During PCR cycling the probe hybridises to its complementary amplicon strand and is then enzymatically cleaved during the primer extension step, releasing the fluorophore and creating an increase in fluorescence which is cumulative (figure 1.20). The enzymatic cleavage relies on the 5'-3' exonuclease activity of Taq polymerase which hydrolytically cleaves any DNA fragment obstructing its path of replication.[68] A mutant-type DNA

template will generate a less stable probe-template duplex and so will not be hybridised at the extension temperature, hence producing no fluorescent increase and thus providing allelic discrimination.

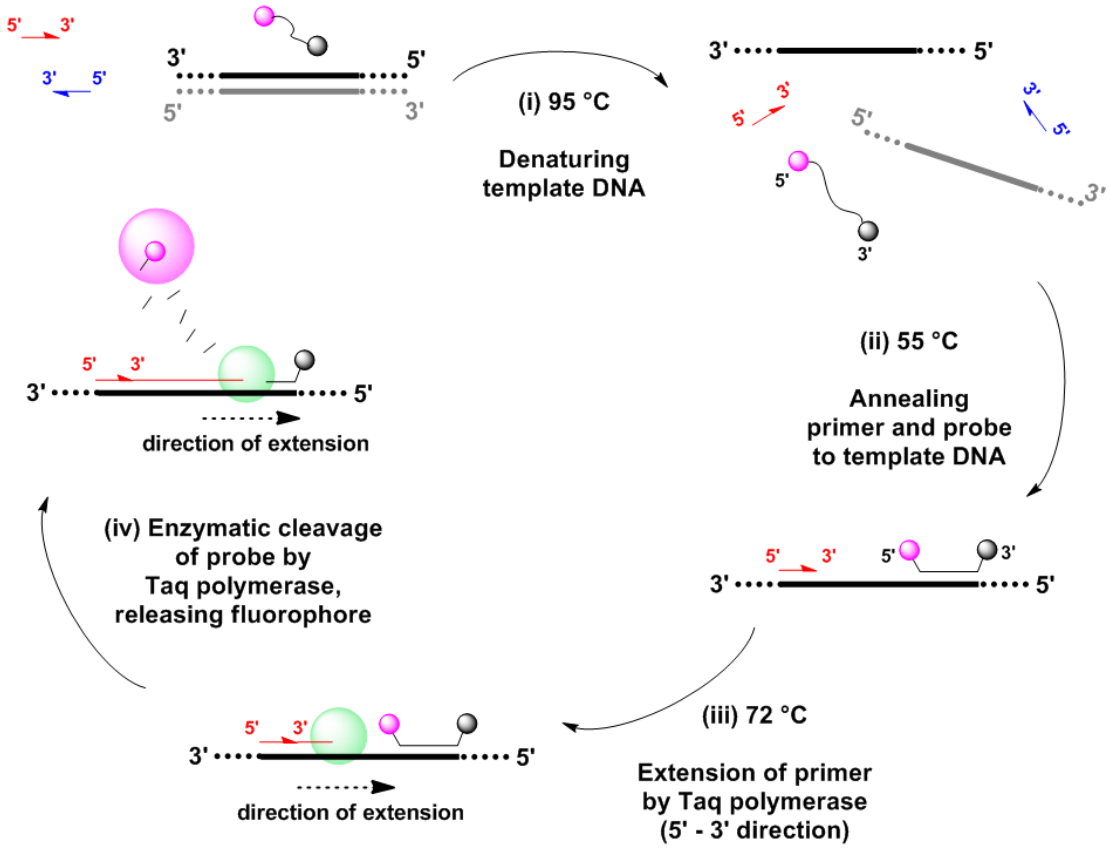



**Figure 1.20.** Mode of action of Taqman probes. Template DNA (thick black and grey) is amplified to amplicon DNA (thin red) using PCR primers (red and blue arrows). The probe, which exhibits quenching between the fluorophore (pink sphere) and the quencher (grey sphere) in its natural form, is enzymatically cleaved by Taq polymerase (green sphere) to release the fluorophore, generating a fluorescent signal.

### *1.5.2.2 Molecular Beacon probes*

Molecular Beacon probes are comprised of a stem and loop structure with a fluorophore and quencher at either end of the oligonucleotide.[69-74] The loop region is designed to be complementary to the region of interest between the primers on the template DNA. When the stem/loop structure is closed the fluorophore and quencher are in close proximity and so contact quenching occurs. When the correct amplicon is present the loop region of the

probe hybridises to it and the fluorophore and quencher are held apart, resulting in an increase in fluorescence (figure 1.21).

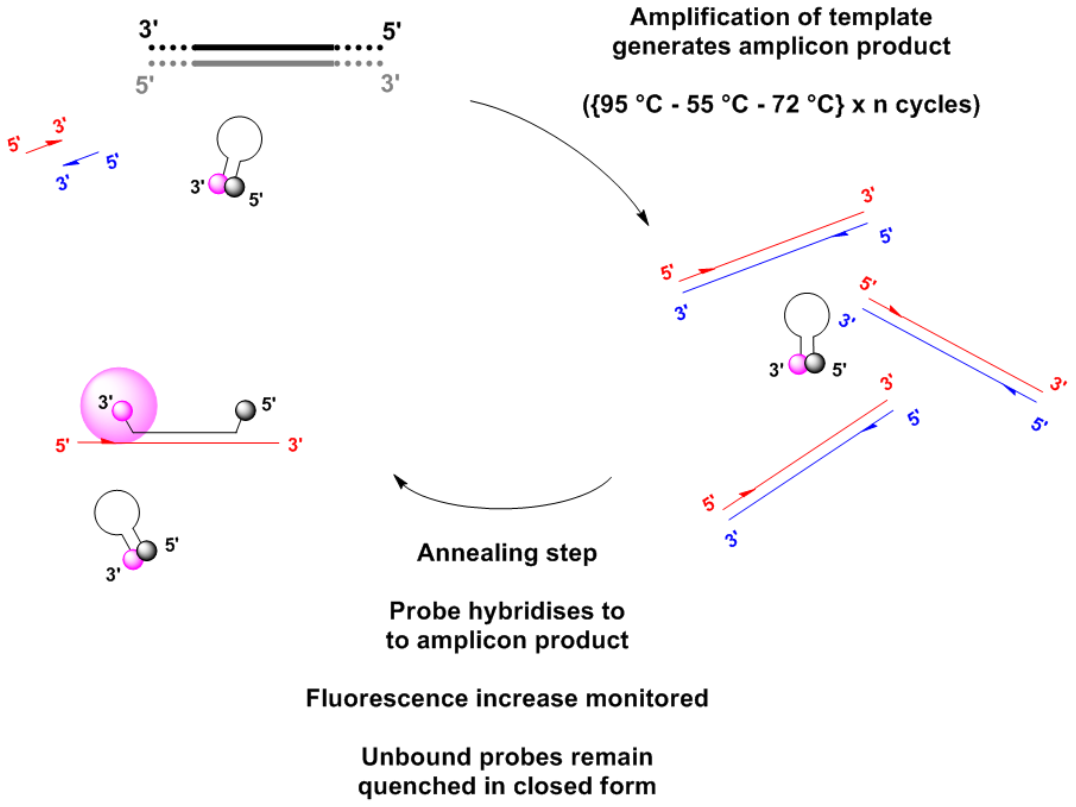



**Figure 1.21.** Mode of action of Molecular Beacon probes. Template DNA (thick black and grey) is amplified to amplicon DNA (thin red and blue) using PCR primers (red and blue arrows). The probe, which exhibits quenching between the fluorophore (pink sphere) and the quencher (grey sphere) in its closed form, generates fluorescent signal when hybridised to complementary amplicon.

The probe is designed with a stem of 5-6 base-pairs and a loop region 25-30 nucleotides in length, so that the probe-target hybridisation is thermodynamically favoured over the closing of the stem structure. The fluorophore and quencher pair must be chosen carefully to ensure there is minimal quenching by FRET in the open form. A mutant amplicon will generate a less stable probe-amplicon duplex and so will generate decreased fluorescence at the monitoring temperature of the real-time PCR (it will not hybridise to the template); in this way Molecular Beacon probes can be used for allelic discrimination.

### *1.5.2.3 Scorpion primers*

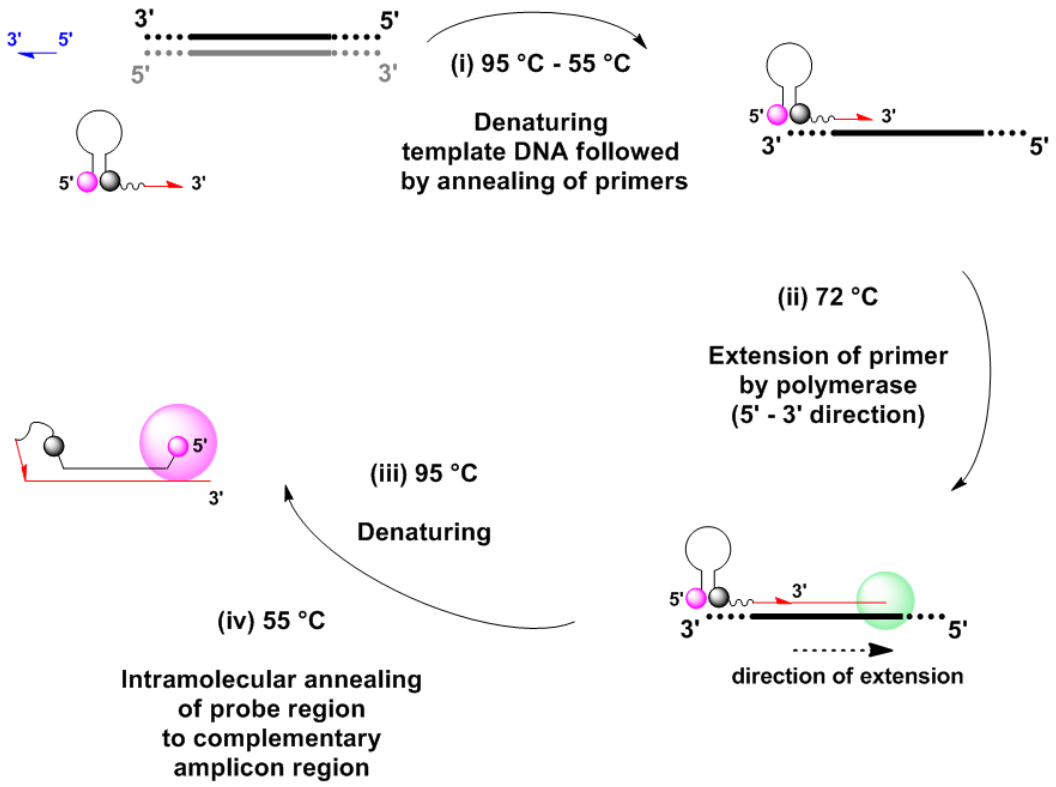



**Figure 1.22.** Mode of action of Scorpion primers. Template DNA (thick black and grey) is amplified (thin red) using PCR primer (blue arrows) and Scorpion primer (red arrow and probe). The Scorpion primer, which exhibits quenching between the fluorophore (pink sphere) and the quencher (grey sphere) in its closed form, is extended by the polymerase (green sphere). The HEG linker (wavy line) between the primer and the probe prevents copying of the probe region by the polymerase. In the next annealing step, after an intramolecular rearrangement, the probe region hybridises to the amplicon region, separating the fluorophore and quencher, resulting in increased fluorescence.

Scorpion primers combine a stem/loop probe with a PCR primer in one oligonucleotide.[75-77] A PCR stopper, commonly hexaethylene glycol (HEG, figure 1.23) is situated between the probe region and primer to prevent copying of the probe region and opening of the hairpin loop during the extension step. The loop region is complementary to a specific region of the amplicon resulting from the extension of the Scorpion primer. The probe region is longer than the stem region and so hybridisation of the probe to the amplicon is thermodynamically favoured over stem reformation. This intramolecular loop-template interaction imposes a conformational change which separates the fluorophore and quencher

to give an increase in fluorescence (figure 1.22). The advantage of this technique over Molecular Beacons is the increased speed and fluorescent signal stability resulting from the unimolecular process. The hybrid duplex between the probe and the wild-type amplicon is more stable than that with a mutant-type amplicon. As a result the fluorescence output at the monitoring temperature is reduced and so allelic discrimination is observed.

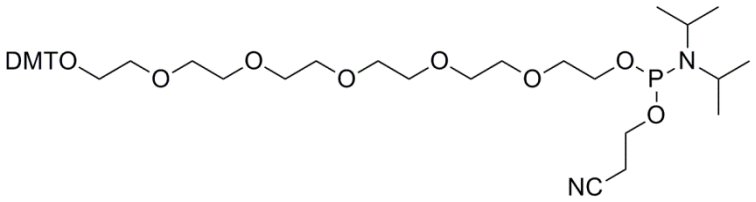



**Figure 1.23.** Hexaethylene glycol (HEG) monomer used as a PCR stopper in the Scorpion primer.

### *1.5.2.4 HyBeacon probes*

HyBeacons are linear probes which function by non-competitive hybridisation. As such they can be shorter, more specific and are easier to design than Molecular Beacons or Scorpions.[78-80] HyBeacon probes commonly contain two fluorophores, covalently attached to the oligonucleotide at either the nucleobase (major groove) or 2' position of the sugar (minor groove). In the single-stranded form the oligonucleotide wraps around itself to minimise repulsive interactions between the nucleobases and water molecules. In this form quenching of the fluorophores occurs by both dye-dye contact quenching and by quenching from the surrounding nucleobases. When a complementary amplicon is present the HyBeacon probe hybridises to it, and the dyes are held apart generating an increase in fluorescence which is cumulative and can be monitored in real-time (figure 1.24).

Post-PCR melting of the sample generates additional information in the form of a fluorescence melting curve. The derivative gives the specific $T_m$ for wild-type or mutant-type samples due to the reduced stability of the probe-target complex with a mismatched sequence. HyBeacon probes are synthesised with a 3' PCR blocker to prevent the probe acting as a PCR primer and giving rise to non-specific signals, propanol is commonly used as the PCR stopper (figure 1.25).

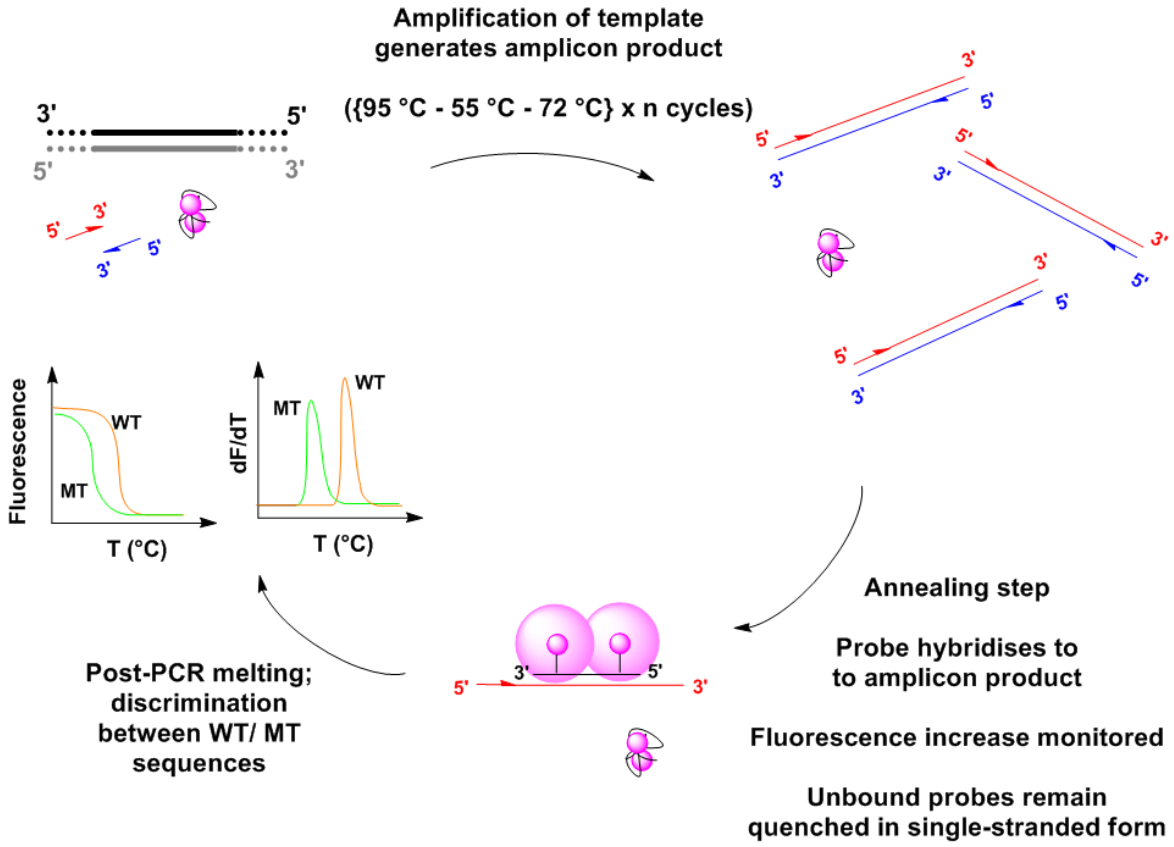



**Figure 1.24.** Mode of action of HyBeacon probes. Template DNA (thick black and grey) is amplified (thin red and blue) using PCR primers (red and blue arrows). The HyBeacon probe, which exhibits quenching of the fluorophores (pink spheres) in the single-stranded state, hybridises to its complementary amplicon product at the annealing steps, separating the fluorophores and resulting in increased fluorescence.

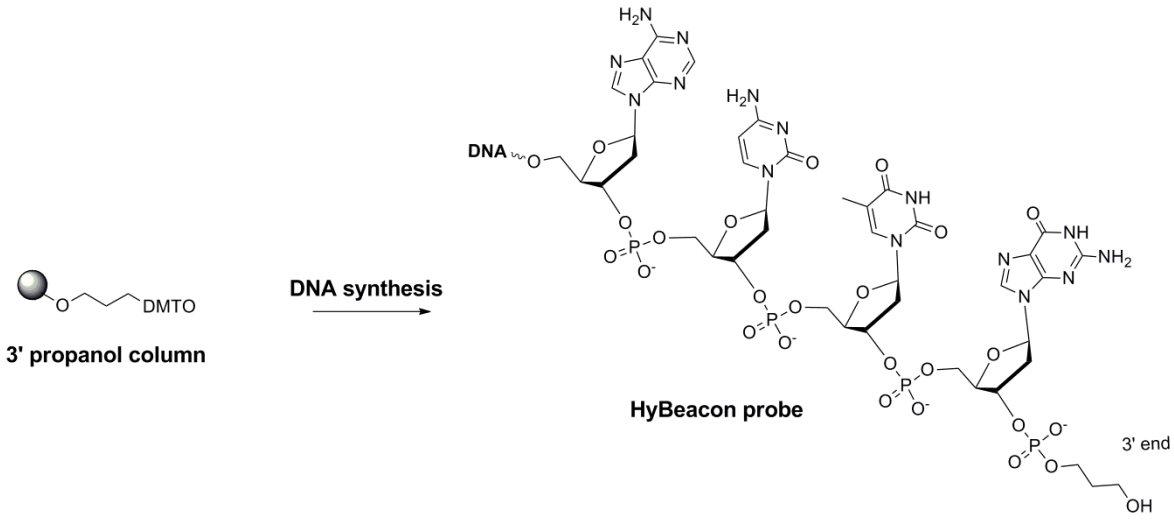



**Figure 1.25.** Incorporation of 3' propanol into a HyBeacon probe at the 3' end prevents the probe acting as a PCR primer and giving rise to non-specific signals.

# Objectives of the research and structure of the thesis

The main aim of the PhD project was to synthesise and investigate the properties of rigid fluorescent monomers in synthetic DNA. These monomers were carefully designed to maximise fluorescence whilst maintaining a rigid linker to the DNA. Rigidity was thought to be important to minimise dye-DNA interactions, to reduce fluorescence quenching and to fix the dyes precisely in space to enable accurate quantitative FRET distance measurements to be made in biological systems.

The synthesis of a fluorescent cytosine base analogue, CPP, is described in chapter 2, the cyanine dye analogues Cy3dT and Cy5dT are presented in chapter 3 and a 'molecular toolkit' containing the highly fluorescent derivative Cy3B is reported in chapter 4.

The photophysical properties of each new monomer have been studied and compared to literature sources. Their suitability as nucleotide analogues was evaluated by carrying out UV-Vis and fluorescence emission studies on oligonucleotides containing the various monomers. The results of these studies are presented in chapters 2-4.

FRET studies were carried out to evaluate the fluorescent dyes in biological and physical applications, and to determine whether the rigid monomers were suitable for energy transfer studies and accurate distance calculations in DNA systems. By inserting combinations of the dyes into a DNA helix and observing the energy transfer between them, information can be obtained about the conformation and dynamics of a system. The results of these studies are presented in chapter 3.

Finally it was important to demonstrate the use of the new fluorescent monomers in biochemical/biological applications. To this end the Cy3B 'molecular toolkit' was used to create a set of PCR probes for use in real-time PCR amplification. Taqman probes, Molecular Beacon probes, HyBeacon probes and Scorpion primers were synthesised and evaluated and the results are presented in chapter 4.

# Chapter 2

# Fluorescent Base Analogue

# Chapter 2 – Fluorescent Base Analogue

## 2.1 Fluorescent base analogues

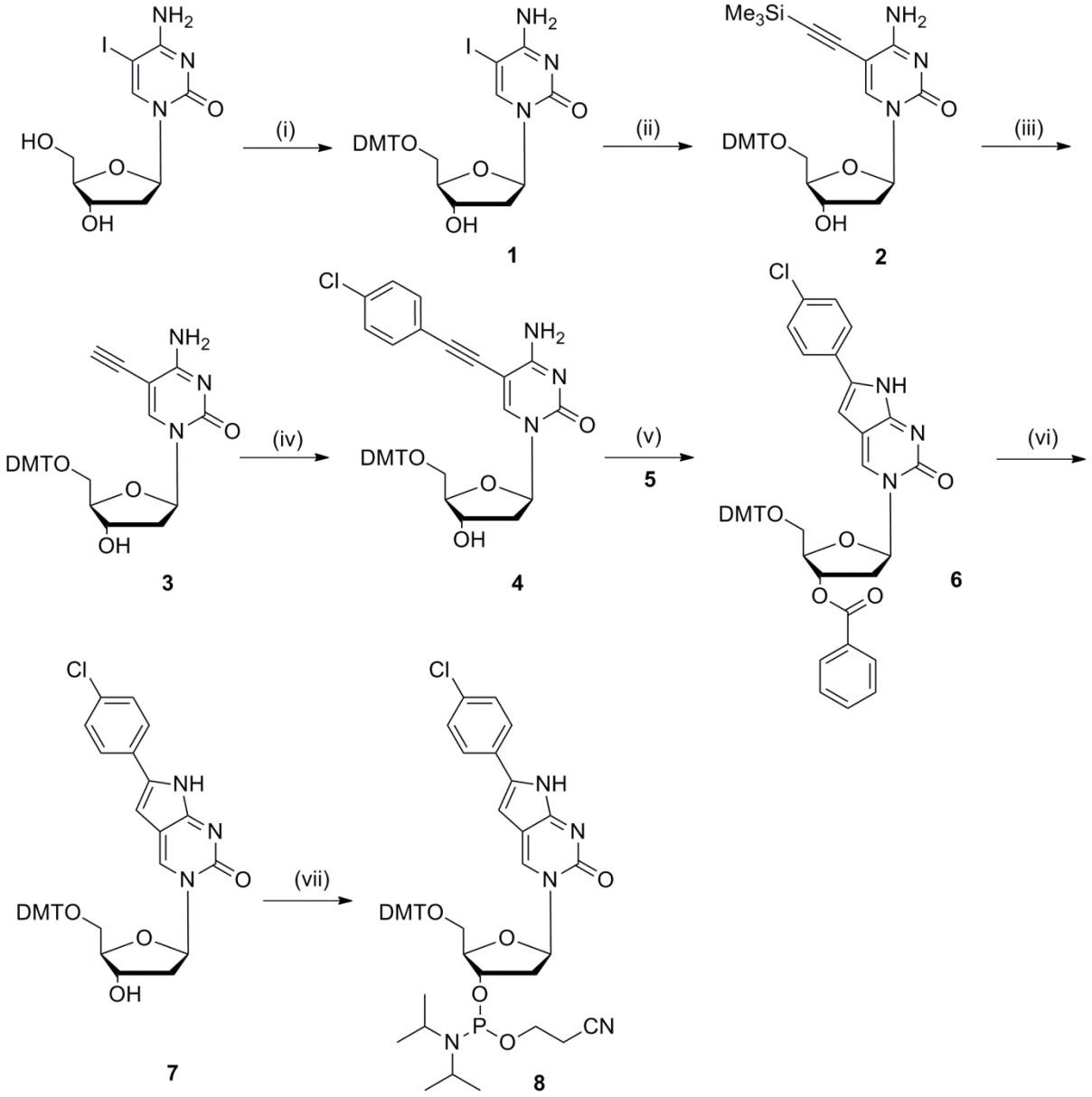
Fluorescent base analogues have been designed to allow incorporation of chromophores into the structure of the nucleobases to create completely rigid fluorescent moieties, some of which have high quantum yields. Various examples have been studied, for example 2-AP,[81-85] pteridines,[31,32,86] pyrrolo-dC[34,87] and tC,[45,88,46] some of which have been discussed in the introduction (section 1.4.4.1).

The synthesis of pyrrolo-pyrimidines, on both DNA and PNA backbones, has been described by Hudson.[37,39,42] He found that pyrrolo-cytosines with phenyl substituents on the 5-membered ring had higher extinction coefficients and quantum yields than those with alkyl substituents.[38] In particular, phenylpyrrolo-cytosines with electron-withdrawing groups in the *para* position have higher extinction coefficients than those with electron donating groups. Based on this evidence, and on our own previous experience with pyrrolo- and furano-base analogues, we decided to investigate the synthesis of a pyrrolo-dC derivative with a *para*-chlorophenyl group.[89-91] It was hoped that the inductively electron-withdrawing substituent on the phenyl ring would help to generate a good fluorophore, which would allow site-specific fluorescent labelling of oligonucleotides without perturbing the DNA structure or duplex stability.

## 2.2 Synthesis of *p*-chlorophenylpyrrolopyrimidine-2'-deoxyriboside (CPP)

The synthesis of CPP (scheme 2.1) was started by tritylation of the commercially available compound 5-iodo-2'-deoxycytidine.[92] This was coupled with trimethylsilylacetylene using Hobbs' variation of the Sonogashira palladium cross-coupling reaction.[93,94] As opposed to standard Sonogashira conditions (palladium catalyst, copper iodide and diethylamine), the variation described by Hobbs, modified specifically for nucleosides, utilises triethylamine and DMF to aid nucleoside solubility. It also specifies that CuI and Pd(0) must be in a molar ratio of 0.2:0.1. These conditions have been found in our research group to afford the optimum results.[95] Following deprotection with TBAF,[96,97] the useful intermediate

5-ethynyl-dC (**3**) was obtained. This compound, recently reported by Qu *et al*,[98] could potentially be reacted, either by the Sonogashira reaction or 'Click chemistry', with any number of chemical moieties to incorporate them in the major groove of a DNA cytidine nucleotide.[99]




**Scheme 2.1.** Synthesis of *p*-chlorophenylpyrrolopyrimidine-2'-deoxyriboside phosphoramidite monomer. Reagents and conditions: (i) DMTCl, pyridine, 5 h, rt, 67 %; (ii) trimethylsilylacetylene, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, 16 h, rt, 86 %; (iii) TBAF, THF, 15 min, rt, 60 %; (iv) 4-chloro-1-iodobenzene, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, 1 h, rt, 94 %; (v) benzoic acid pentafluorophenolester **5**, $Et_3N$, DMF, 48 h, 80 °C, 63 % (plus **7**, 12 %); (vi) 0.04 M $K_2CO_3$/MeOH, 2 h, rt, 91 %; (vii) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 1 h, 26 %.

4-Chloro-1-iodobenzene was coupled to the free ethynyl group on compound **3**, via palladium-catalysed cross-coupling chemistry. The aim was then to utilise the established cyclisation conditions for furano-dT described by McGuigan[100] (triethylamine, methanol and copper-iodide at reflux). Cyclisation of furano-dT has previously been followed by oxygen to NH conversion using ammonia in methanol (at 55 °C) to yield pyrrolo-dC.[41] Unfortunately, these cyclisation conditions did not work for our cytosine compound. To overcome this problem, the approach was altered to include an initial acylation of the exocyclic amino group, as described by Hudson.[38,37] Benzoic acid pentachlorophenol ester was reacted with compound **4**, which resulted in the addition of a benzoyl group to the exocyclic amino group of the cytidine ring. This increased the reactivity of the amino nitrogen, thus initiating the spontaneous cyclisation of the compound to form a pyrrolo ring. It is likely that trace amounts of copper left from the previous reaction would have aided the cyclisation. During this reaction, the benzoyl group migrated to the 3'-OH. Subsequent deprotection of the 3'-OH afforded the product **7**, which was then phosphitylated to afford the phosphoramidite monomer **8**, ready for oligonucleotide synthesis.

## 2.3 The photophysical properties of CPP monomer and oligonucleotides containing CPP phosphoramidite

The CPP monomer **7** was evaluated for its photophysical properties by UV-Vis absorption and fluorescence emission spectroscopy, and was compared to the oligonucleotides containing the CPP monomer (table 2.1). Three CPP-containing oligonucleotides were synthesised, all with deoxyadenosine nucleotides flanking the CPP to negate any discrepancy in neighbouring-nucleobase quenching, often displayed by fluorescent base analogues.

**Table 2.1.** Oligonucleotide sequences (mass data in appendix A.1); **X** = CPP monomer.

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN01 | CGA TCA CAC ACA AGG ACG AGG ATA AGG AGG AGG |
| ODN02 | CGA TCA CAC A**X**A AGG ACG AGG ATA AGG AGG AGG |
| ODN03 | CGA TCA CA**X** ACA AGG ACG AGG ATA AGG AGG AGG |
| ODN04 | CGA TCA **X**AC ACA AGG ACG AGG ATA AGG AGG AGG |
| ODN05 (complement) | CCT CCT CCT TAT CCT CGT CCT TGT GTG TGA TCG |

The extinction coefficient and quantum yield of CPP in the free monomer form are relatively high (14000 $M^{-1}cm^{-1}$ and 0.75 respectively). The extinction coefficient of the oligonucleotides containing CPP ranges between 13500 and 17500 $M^{-1}cm^{-1}$, and hence deviates little in the oligonucleotide environment. The extinction coefficients are in the higher end of the range typical for nucleobase analogues (5900 $M^{-1}cm^{-1}$ for pyrrolo-dC, 6000 $M^{-1}cm^{-1}$ for 2-AP, 13000 $M^{-1}cm^{-1}$ for 3-MI, figure 2.1).[81,82,86] In comparison, tC, a tricyclic phenothiazine analogue of cytosine, has an extinction coefficient of 4500 $M^{-1}cm^{-1}$.[45] The absorbance maxima of the CPP duplexes are red-shifted by 5 nm from the CPP free monomer.

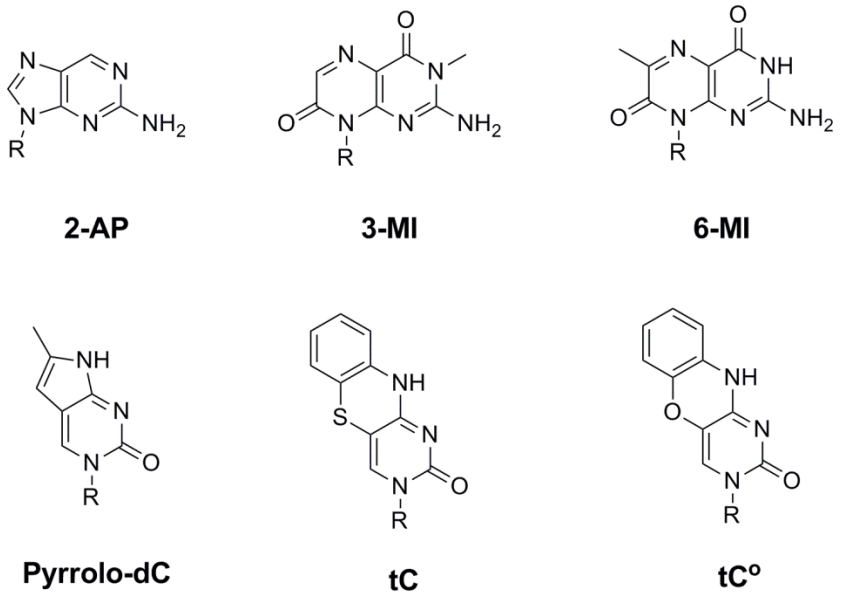



**Figure 2.1.** Chemical structures of some common fluorescent base analogues; 2-aminopurine (2-AP); the pteridines (3-MI and 6-MI); pyrrolo-dC; and the tricyclic cytosine analogues tC and $tC^{O}$. R=2'deoxyribose.

The quantum yield of the CPP decreases upon incorporation into oligonucleotides and decreases further upon duplex formation (figure 2.2). This pattern is the same for pyrrolo-dC; the quantum yield of which is 0.2 in the free monomer form, decreasing to 0.06 in single stranded DNA, then reducing again to 0.03 in double-stranded DNA.[87] Upon incorporation of CPP into DNA, the reduction in quantum yield is accompanied by a red-shift in the emission maxima of 8-12 nm (table 2.2). Upon hybridisation to a complementary strand, the fluorescence emission is again reduced dramatically and the emission maximum is further red-shifted by up to 6 nm. The quantum yield of the double-stranded form is only slightly lower than that seen for the tC monomer. The tC monomer however exhibits very little change between single- and double-stranded forms; 0.17-0.24 and 0.16-0.21 respectively, suggesting it is independent of neighbouring nucleobases.[46]

The phenoxazine derivative of tC, tC$^o$ (figure 2.1), has a relatively high quantum yield (0.3) and extinction coefficient (9000 $M^{-1}cm^{-1}$) but is very sensitive to its surrounding bases in DNA and shows a large range in quantum yield upon DNA incorporation (0.14-0.41).[49] Albinsson and co-workers maintain that this makes it possible to use tC$^o$ as a probe of nucleic acid secondary structure. The pteridine base analogues also exhibit this kind of behaviour. For example, 3-MI, which has a quantum yield of 0.88 in the free monomer, has a reduced quantum yield in DNA of <0.01-0.3 depending on the sequence.[31]

Despite the potential problems of environment-sensitivity, the overall fluorescence properties of the CPP monomer are encouraging.

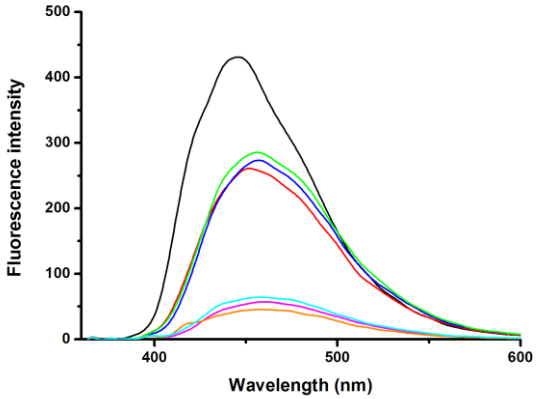



**Figure 2.2.** Fluorescence emission of CPP free monomer compared to single-stranded (ss) and double-stranded (ds) oligonucleotide samples. All samples evaluated at 1 µM; free monomer in MeOH and oligonucleotides in phosphate buffer (25 mM with 100 mM NaCl, pH 7.5). Free monomer (black); ODN02 ss (red), ds (pink); ODN03 ss (blue), ds (orange); ODN04 ss (green), ds (cyan). Fluorescence is highest in the free CPP monomer, decreases upon incorporation into oligonucleotides and is lowest in double-stranded DNA.

**Table 2.2.** Photophysical properties[a]; UV-Vis absorbance maxima ($A_{max}$), fluorescence emission maxima ($Em_{max}$), extinction coefficient (**ε**) and quantum yield (**Φ**).

| **Oligo code** | **$A_{max}$ (nm)** | **ε ($M^{-1}cm^{-1}$)** | **$Em_{max}$ (nm)** | **Φ** |
|---|---|---|---|---|
| Free CPP monomer | 370 | 14,000 | 445 | 0.75 |
| ODN02 | 370 | 15,100 | 457 | 0.48 |
| ODN02 duplex | 375 | 13,500 | 459 | 0.13 |
| ODN03 | 375 | 13,800 | 457 | 0.63 |
| ODN03 duplex | 375 | 14,300 | 459 | 0.12 |
| ODN04 | 370 | 17,500 | 453 | 0.59 |
| ODN04 duplex | 375 | 14,500 | 459 | 0.15 |

[a]All measured in phosphate buffer (25 mM with 100 mM NaCl, pH 7.5) apart from free monomer, measured in MeOH.

**Table 2.3.** $T_m$ data from UV melting for CPP-modified duplexes [a]

| Duplex | Oligo1 | Oligo2 | $T_m$/ °C | Δ $T_m$/ °C |
|---|---|---|---|---|
| Unmodified control | ODN01 | ODN05 | 74.2 | --- |
| CPP modified | ODN02 | ODN05 | 74.6 | +0.4 |
| CPP modified | ODN03 | ODN05 | 73.7 | -0.5 |
| CPP modified | ODN04 | ODN05 | 74.1 | -0.1 |

[a]Average $T_m$ value was calculated from three successive melting curves. Standard deviation for all $T_m$ data is ± 0.1 ºC.

The thermal stability of the oligonucleotides containing CPP was tested to ensure that it did not perturb the DNA duplex. In some cases, the CPP was found to destabilise the DNA, however in the case of ODN02 the duplex was stabilised (table 2.3). In all cases, the change in $T_m$ is very small and therefore it is reasonable to propose that the addition of the CPP monomer into DNA does not significantly affect duplex stability. Hudson found that the DNA duplex was destabilised by 4.7 ºC upon incorporation of pyrrolo-dC, although with 6-methoxymethyl-3-(2-deoxy-*β*-D-ribofuranosyl)-*3H*-pyrrolo[2,3-*d*]-pyrimidin-2-one ($^{MMe}$p-dC, figure 2.3) the duplex was stabilised by 1.3 ºC.[40] In the case of 6-phenylpyrrolocytosine ($^{Ph}$p-dC) he observed a 2 ºC stabilisation of the duplex, which was attributed to increased π-stacking. The derivative $^{boPh}$p-dC, which has bis-*ortho*-aminoethoxy groups, stabilised the duplex by an astounding 10 ºC, which is probably due to the additional hydrogen bond which can form between $^{boPh}$p-dC and guanine (figure 2.3).[43]

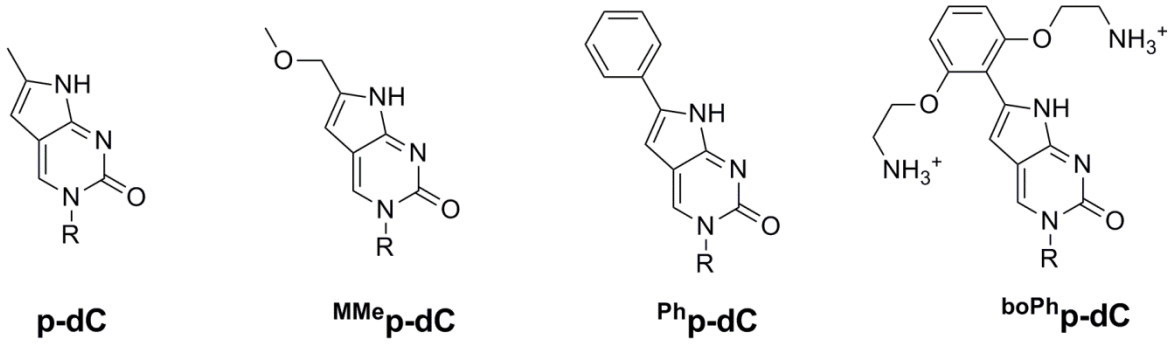



**Figure 2.3.** Chemical structures of fluorescent base analogue pyrrolo-dC (pdC) and its derivatives; methoxymethyl pyrrolo-dC ($^{MMe}$p-dC); phenyl-pyrrolo-dC ($^{Ph}$p-dC); and bis-*ortho*-aminoethoxyphenyl pyrrolo-dC ($^{boPh}$p-dC). R= 2'deoxyribose.

As indicated by the change in fluorescence upon incorporation into oligonucleotides and duplex formation, the CPP monomer appears to be sensitive to its environment. Base-analogues are often found to exhibit these properties and it is largely attributed to quenching by surrounding bases, with more significant quenching caused by purine bases, in particular guanine which has better electron donating properties.[101-103] A study was undertaken in which deoxyguanosine was replaced by 2'deoxyinosine nucleotides (figure 2.4) in order to confirm this theory (hypoxanthine is a less efficient quencher than guanine).

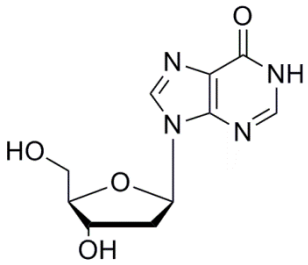



**Figure 2.4.** Chemical structure of 2'deoxyinosine (nucleobase = hypoxanthine) which can base-pair with cytosine but is a less efficient quencher than guanine.

A range of oligonucleotides complementary to the CPP-containing strands were synthesised with either one inosine replacement (opposite the CPP site), five inosine replacements (replacing every guanine except opposite the CPP site) and six inosine replacements (replacing every guanine) (table 2.4).

**Table 2.4**. Sequence of inosine-modified oligonucleotides (complementary to ODN02, ODN03, ODN04); **I**= 2'deoxyinosine. (mass data in appendix A.1)

| **Oligonucleotide code** | **Sequence** | **To test with:** |
|---|---|---|
| ODN06 | CCT CCT CCT TAT CCT CGT CCT T**I**T GTG TGA TCG T | ODN02 |
| ODN07 | CCT CCT CCT TAT CCT C**I**T CCT TGT **I**T**I** T**I**A TC**I** T | ODN02 |
| ODN08 | CCT CCT CCT TAT CCT C**I**T CCT T**I**T **I**T**I** T**I**A TC**I** T | ODN02/ ODN03/ ODN04 |
| ODN09 | CCT CCT CCT TAT CCT CGT CCT TGT **I**TG TGA TCG | ODN03 |
| ODN10 | CCT CCT CCT TAT CCT CGT CCT TGT GT**I** TGA TCG | ODN04 |
| ODN11 | CCT CCT CCT TAT CCT C**I**T CCT T**I**T GT**I** T**I**A TC**I** T | ODN03 |
| ODN12 | CCT CCT CCT TAT CCT C**I**T CCT T**I**T **I**TG T**I**A TC**I** T | ODN04 |

In general, replacement of the guanine bases with hypoxanthine bases resulted in an increase in fluorescence quantum yield, with replacement of all the guanines having the most significant effect (table 2.5). This confirms that the decrease in quantum yield exhibited by the CPP monomer upon oligonucleotide incorporation may be attributed, in part to quenching from neighbouring guanine bases, although it is clear that other factors also operate.

**Table 2.5.** Quantum yield of duplexes; oligonucleotides containing CPP and complementary strands containing 0-6 inosine additions.

| **Duplex** | **# Inosine additions** | **Φ** |
|---|---|---|
| ODN02_ODN05 | None | 0.13 |
| ODN02_ODN06 | 1 | 0.18 |
| ODN02_ODN07 | 5 | 0.19 |
| ODN02_ODN08 | 6 | 0.22 |
| ODN03_ODN05 | None | 0.12 |
| ODN03_ODN09 | 1 | 0.14 |
| ODN03_ODN11 | 5 | 0.15 |
| ODN03_ODN08 | 6 | 0.19 |
| ODN04_ODN05 | None | 0.15 |
| ODN04_ODN10 | 1 | 0.14 |
| ODN04_ODN12 | 5 | 0.16 |
| ODN04_ODN08 | 6 | 0.19 |

A further study was undertaken to determine whether a change in pH would result in a change in the fluorescence quantum yield. Such behaviour is exhibited by the tC monomer, which is pH independent in the range 4-10 but has a drop in quantum yield at extreme pH. Between pH 4-10, the quantum yield of tC is 0.16-0.17. At pH 14, in its deprotonated form, the quantum yield is 0.11 and at pH -0.5, in its protonated form, it is <0.01.[88] An example CPP-oligonucleotide was evaluated for fluorescence quantum yield in the pH range 5.5 to 9.0. The results show that within the range pH 6.0 to 8.0 the quantum yield remains fairly constant but then drops below pH 6.0 (table 2.6). This suggests that small changes of pH around neutral pH should not significantly affect the fluorescence efficiency of CPP-oligonucleotides.

**Table 2.6.** Quantum yield (**Φ)** of CPP-containing oligonucleotides in the single-stranded form, at varying pH.

| pH | Φ |
|---|---|
| 5.5 | 0.33 |
| 6 | 0.58 |
| 7.5 | 0.59 |
| 8 | 0.63 |

## 2.4 Conclusions

A new fluorescent base analogue was successfully synthesised, which did not perturb DNA duplex stability. It was found to have a high fluorescence quantum yield and extinction coefficient in the free monomer form, relative to other fluorescent base analogues (although it is still dramatically less fluorescent than fluorophores such as fluorescein). However, the fluorescence quantum yield was found to be significantly affected by the surrounding DNA environment upon oligonucleotide incorporation, which may be problematic in some applications. For example, for most DNA-related applications, such as PCR and single-molecule DNA studies, it is preferable if the dye becomes more fluorescent in double-stranded DNA or at least retains its fluorescence. In the case of CPP, the fluorescence decreases and so may not be completely suitable in these cases. It was therefore of interest to initiate investigations into alternative fluorescent compounds which could fulfil this requirement.

# Chapter 3

# Cyanine Dye Analogues

# Chapter 3 – Cyanine Dye Analogues

## 3.1 Cyanine dyes in DNA

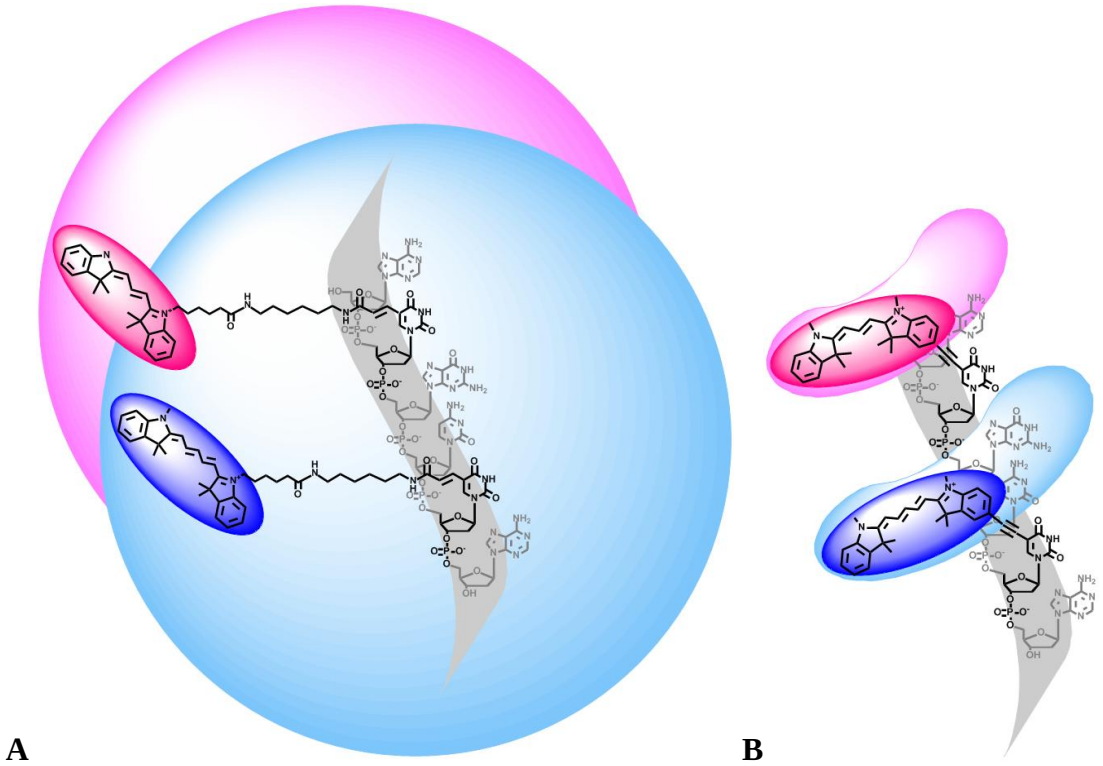
Based upon the lower fluorescence exhibited by base analogues (due mainly to relatively low extinction coefficients), it was of interest to investigate the properties of dyes conjugated directly to nucleobases. Commonly used fluorescent dyes such as rhodamine and cyanine dyes have very high extinction coefficients making them more fluorescent than simple fluorescent base analogues such as 2-aminopurine or pyrrolo-dC. The advantage of using fluorescent base analogues, as opposed to external modification with a covalently attached fluorophore, is that the fluorophore can be rigidly incorporated into the DNA single-strand/duplex. To achieve this objective we decided to incorporate dye structures into DNA via a rigid ethynyl linker at the C5-position of 2'-deoxythymidine to give "CydT" monomers.

CyDyes were selected due to their high fluorescent efficiency and the broad range of spectral properties that are exhibited by the various analogues. As discussed in section 1.4.4.3, CyDyes are available as NHS-esters for post-synthetic oligonucleotide labelling. The NHS-esters can be reacted with amino-functionalised DNA containing amino-C6-dT residues (Glen Research), generating a long flexible linker to the C5-base position, allowing binding of CyDyes into the major groove of DNA.[104] More than one binding-conformation has been reported and it has been suggested the dye may spend some time not bound to the DNA.[105] This means that the position of the dye cannot be accurately predicted which may be a drawback for FRET studies. Although planar heterocycles generally preferably intercalate into DNA, studies have shown that free CyDye monomers have a preference for minor-groove binding.[106-109] This may be due to the flexible polymethine linker between the indole rings, which would allow the dye to fit into the space neatly. In a detailed $^{1}$H NMR study, Armitage and co-workers proved that cyanine dyes preferentially bind in the minor groove as dimers, spanning six base pairs, the efficiency of which is sequence dependent.[110] However in a previous study they found that that the dimethylindole groups of Cy5 suppressed dimerization in comparison to benzoxazole, benzothiazole and quinoline groups.[111] This is thought to be due to clashing of the methyl groups with the walls of the minor groove and/or less optimal stacking of the

π-systems in this analogue. The variable interactions between the dye and the DNA could significantly complicate interpretation of the fluorescence properties of these systems.

In addition to the post-synthetic labelling methodology, the other option for CyDye incorporation is to use the commercially available CyDye-amidites which allow the dye to be incorporated at the 5' end of the DNA during solid phase oligonucleotide synthesis. In this way the CyDye can stack on the end of the duplex, making its position and fluorescence properties more predictable. However it has been shown that a certain percentage of the dyes are not stacked on the end of the duplex, hence introducing some uncertainty into the conclusions that are drawn from studying the fluorescence properties of the system.[60,61] Furthermore, 5'-addition of the dye is severely limiting in the design of potential oligonucleotide sequences, as the dye may only be added to the end of the sequence and not internally.

The planned CydT phosphoramidite monomers should generate a DNA-addition option that is more flexible in terms of sequence design and applications. The short rigid tether was designed to make the dye position more precise and prevent interaction between the dye and the DNA (figure 3.1).



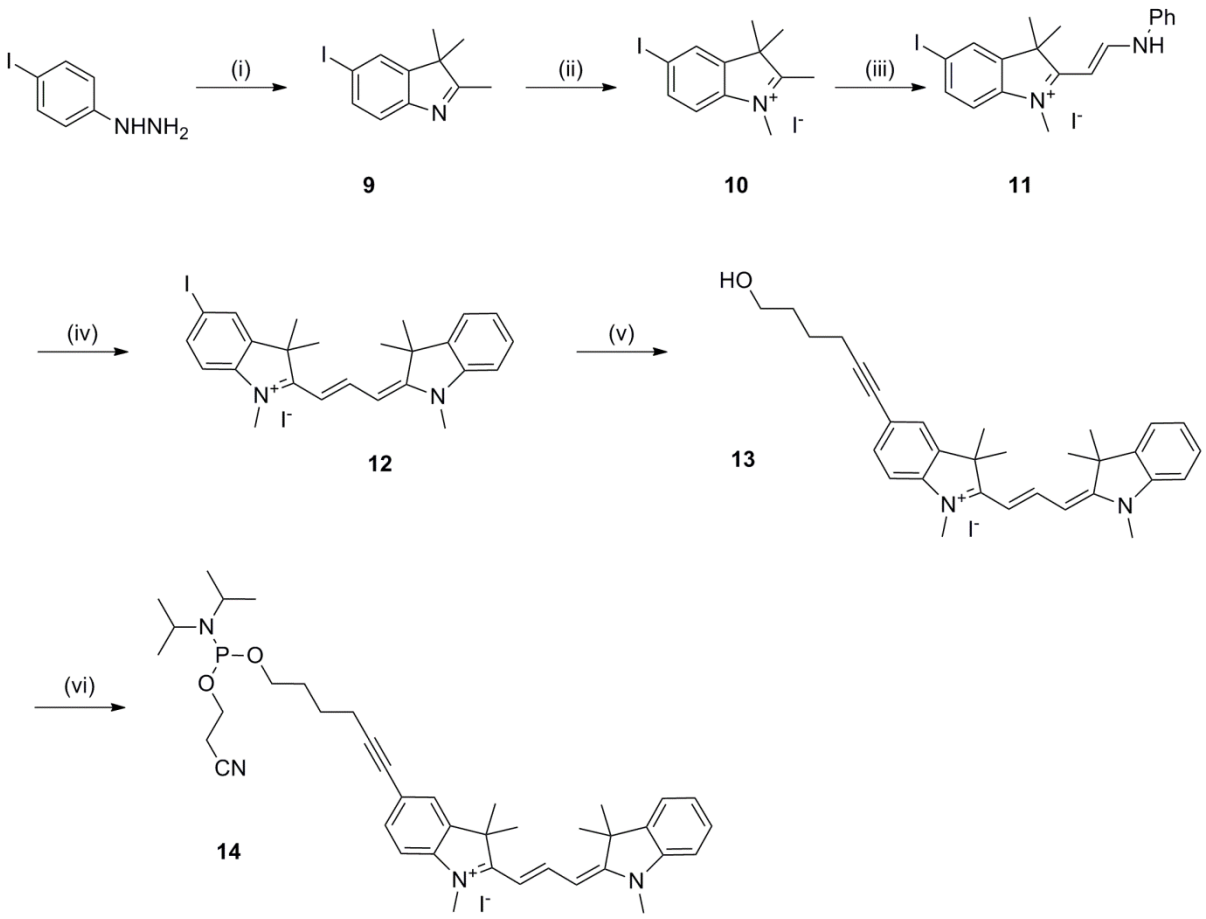
**A** **B**

**Figure 3.1.** (A) Flexible dye linker and (B) short rigid dye linker in the context of a DNA helix cross-section (only labelled strands shown). Blue and pink shaded areas show the volume of space that can be sampled by the dyes. A short rigid linker to DNA greatly increases the accuracy of the dye position.

Similar work was carried out by Balasubramanian and co-workers by the post-synthetic Sonogashira reaction between ethynyl-Cy3 or ethynyl-Cy5 and 5-iodo-dU in a support-bound oligodeoxynucleotide.[112] However, in comparison to the solid-phase phosphoramidite monomer incorporation, this post-synthetic technique is limiting as it adds additional steps to the overall oligonucleotide synthesis. In addition, it does not allow for multiple or mixed dye additions, hence it was deemed important to investigate the potentially more useful CydT phosphoramidite approach.

## 3.2 CyDye synthesis and CyDye monomer properties

The synthesis of CyDyes is well documented. Initially a solid-phase synthesis[113] was evaluated but it was found to be inappropriate for the quantity of product required. Synthesis of the basic cyanine structure was then carried out according to a combination of literature procedures to yield 5-iodo-Cy3.[113-115,112] To test whether addition of an ethynyl group conjugated to one of the indole rings would affect the fluorescence properties of the dye, 5-hexyn-1-ol was added to 5-iodo-Cy3, **12**, by Sonogashira reaction (scheme 3.1). The primary alcohol **13** was then phosphitylated to give a 5'Cy3 phosphoramidite monomer **14**. Purification of the phosphoramidite monomer was problematic as the high polarity of the dye meant that elution of the product by column chromatography was slow, therefore precipitation in hexane from a minimum of DCM was used instead for purification. Without purification by column chromatography (which is the most efficient way to remove excess phosphitylating reagent) additional measures had to be taken to ensure purity of the product; firstly the reaction had to be carefully monitored to ensure no more phosphitylating reagent was added than necessary; secondly the precipitation step was repeated several times in an attempt to remove as much remaining reagent as possible.




**Scheme 3.1.** Synthesis of 5-hexyn-1-ol-6-Cy3 phosphoramidite. Reagents and conditions: (i) 3-methyl-2-butanone, $MeCO_2H$, reflux, 2.5 h, 81 %; (ii) MeI, MeCN, reflux, 16 h, 78 %; (iii) *N,N'*-diphenylformamidine, triethylorthoformate, ethanol, reflux, 2.5 h, 95 %; (iv) 1,2,3,3-tetramethyl-3*H*-indolium iodide, $Ac_2O$, pyridine, 50 °C, 48 h, 61 %; (v) 5-hexyn-1-ol, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, rt, 1.5 h, 80 %; (vi) 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 2.5 h, 96 %.

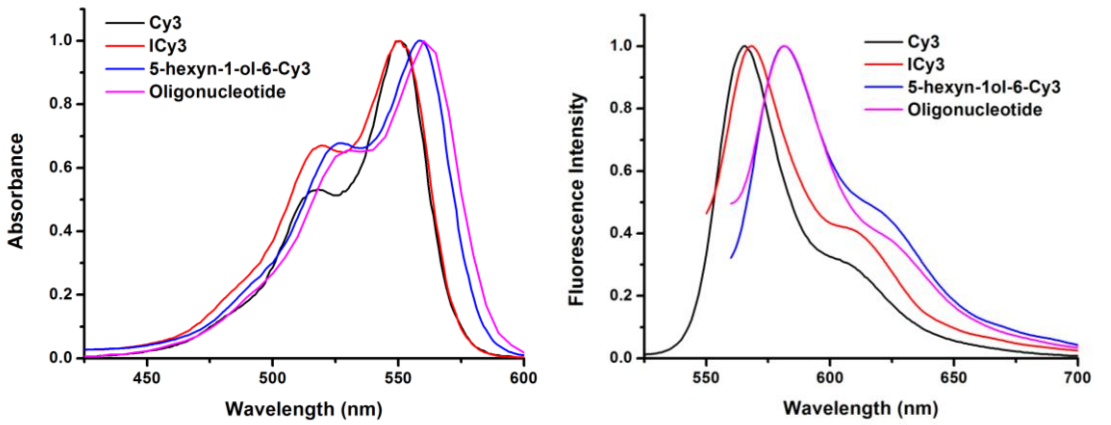
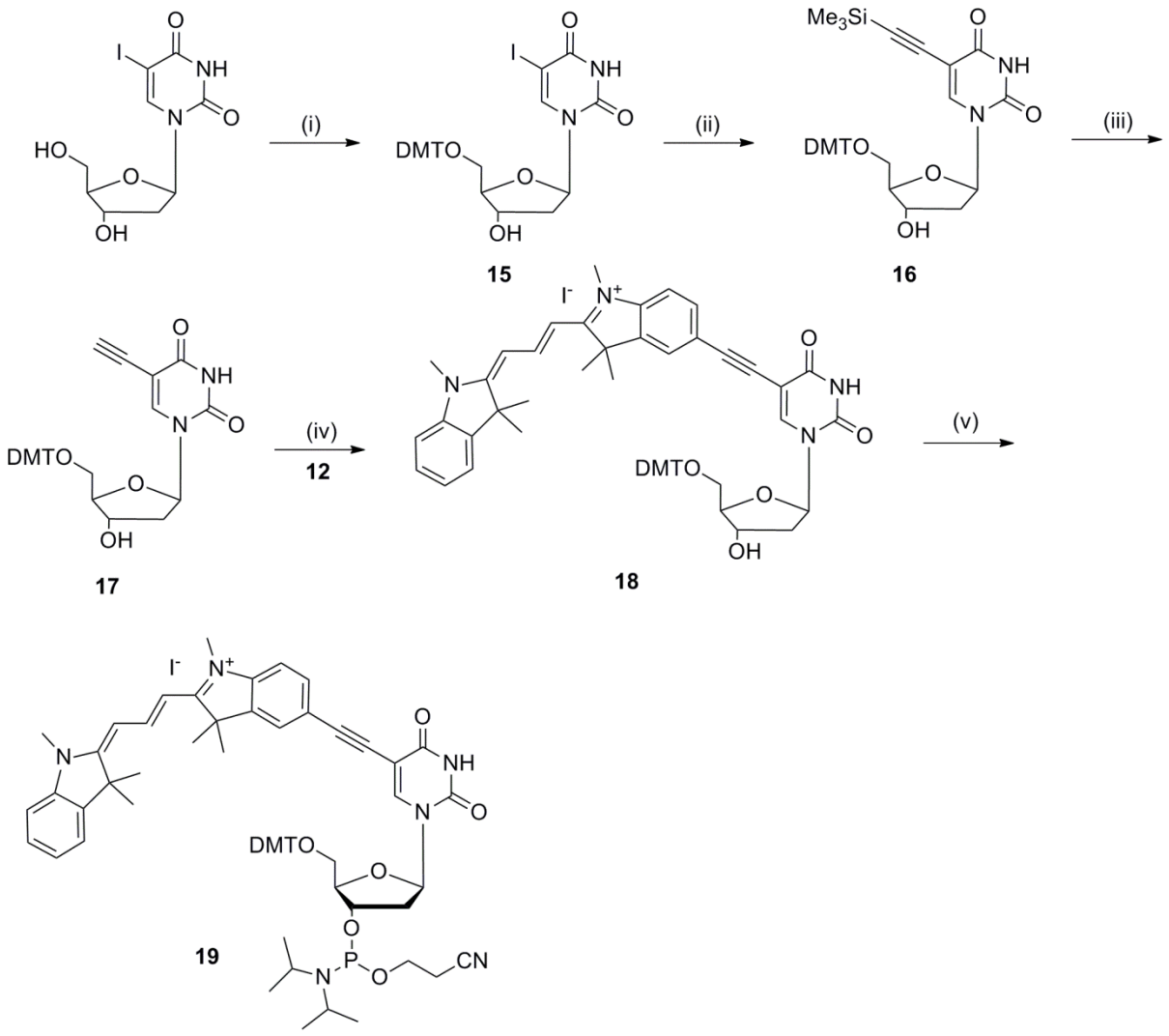



**Figure 3.2.** Normalised absorbance and emission spectra of; Cy3 (NHS-ester); I-Cy3 (compound **12**); 5-hexyn-1-ol-6-Cy3 (compound **13**); 5'-Cy3 oligonucleotide. The properties of I-Cy3 are similar to Cy3 (NHS-ester), addition of a hexynol group and incorporation into an oligonucleotide red-shifts the absorbance and emission maxima.

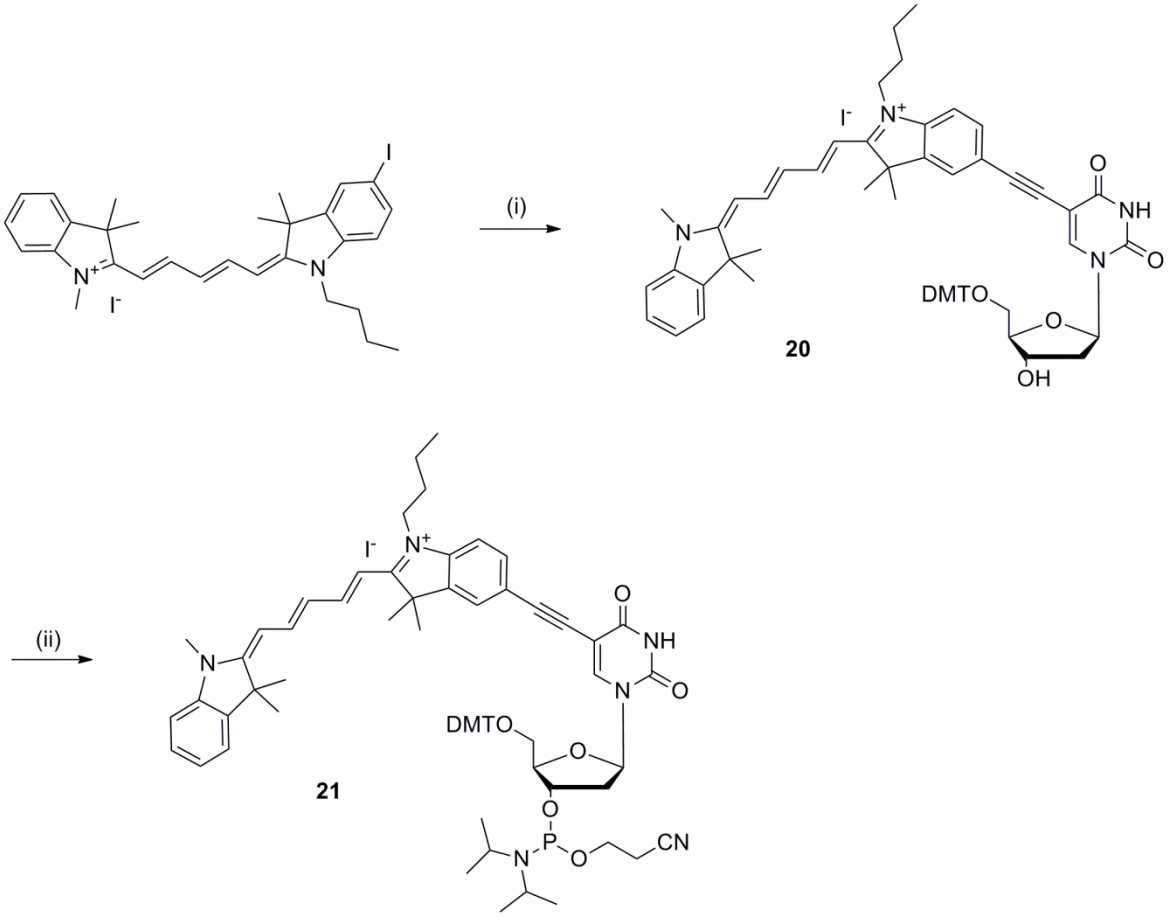
The absorbance and emission spectra of the prepared Cy3 compounds and 5'Cy3-oligonucleotide were compared to standard Cy3 NHS-ester (figure 3.2). The wavelengths of absorbance and emission for I-Cy3 are virtually unchanged from those of Cy3 NHS-ester, however the 5-hexyn-1-ol-6-Cy3 version exhibits a red-shift in the absorbance and emission maxima of ~10 nm (abs 558 nm, em 580 nm compared to abs 550 nm, em 570 nm). With the knowledge that the addition of the ethynyl group did not greatly alter the properties of the dye, the syntheses of the dye-thymine base analogues were undertaken.




**Scheme 3.2.** Synthesis of Cy3dT-phosphoramidite. Reagents and conditions: (i) DMTCl, pyridine, 5.5 h, rt, 88 %; (ii) trimethylsilylacetylene, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, 16 h, rt, 94 %; (iii) TBAF, THF, 20 min, rt, 99 %; (iv) **12**, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, rt, 24 h, 42 %; (v) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 2 h 45 min, 96 %.

The first step in scheme 3.2 was tritylation of the commercially available 5-iodo-dU. The subsequent palladium cross-coupling reaction with trimethylsilylacetylene, followed by

TBAF deprotection of the trimethysilyl group, gave intermediate **17**. This intermediate, much like the 5-ethynyl-dC version (compound **3**, scheme 2.1, chapter 2), is a very versatile compound. The ethynyl group is available for reaction by Sonogashira or click-chemistry to conjugate any number of structures into the major groove position on a DNA thymidine nucleotide. Ethynyl-dT has been studied in depth; in terms of antiviral activity,[116] DNA stabilisation,[117] and for the click-chemistry labelling of DNA with azide-modified fluorophores.[118,119] In the case of scheme 3.2, 5-iodo-Cy3 was added by palladium cross-coupling to compound **17** followed by phosphitylation to yield the Cy3dT phosphoramidite monomer **19**. Problems were experienced with the purification of compound **18**. The polar nature of the compound resulted in a tendency to stick to silica during column chromatography purification; however the need to purify the product after the Sonogashira reaction made the process unavoidable yet probably resulted in the lower than expected yield.




**Scheme 3.3.** Synthesis of Cy5dT-phosphoramidite. Reagents and conditions: (i) 5-ethynyl-DMT-dU, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, rt, 26 h, 91 %; (ii) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 1 h 45 min, 84 %.

The synthesis of Cy5dT phosphoramidite **21** followed the above procedure by conjugating 5-iodo-Cy5 (synthesised during a parallel research project by Marta Gerowska) to 5-ethynyl-dU by Sonogashira reaction followed by phosphitylation (scheme 3.3).

Study of the fluorescence properties of the monomers indicated that incorporation of the dyes into the CydT structure leads to a red-shift in the absorbance and emission spectra by around 15-20 nm relative to the NHS-ester (figure 3.3, table 3.1). In addition the extinction coefficients are reduced from 150000 $M^{-1}cm^{-1}$ to 112000 $M^{-1}cm^{-1}$ for Cy3dT and 250000 $M^{-1}cm^{-1}$ to 208000 $M^{-1}cm^{-1}$ for Cy5dT. The quantum yields are also reduced. Incorporation of the monomers into oligonucleotides alters the absorbance and emission maxima slightly (figure 3.4).

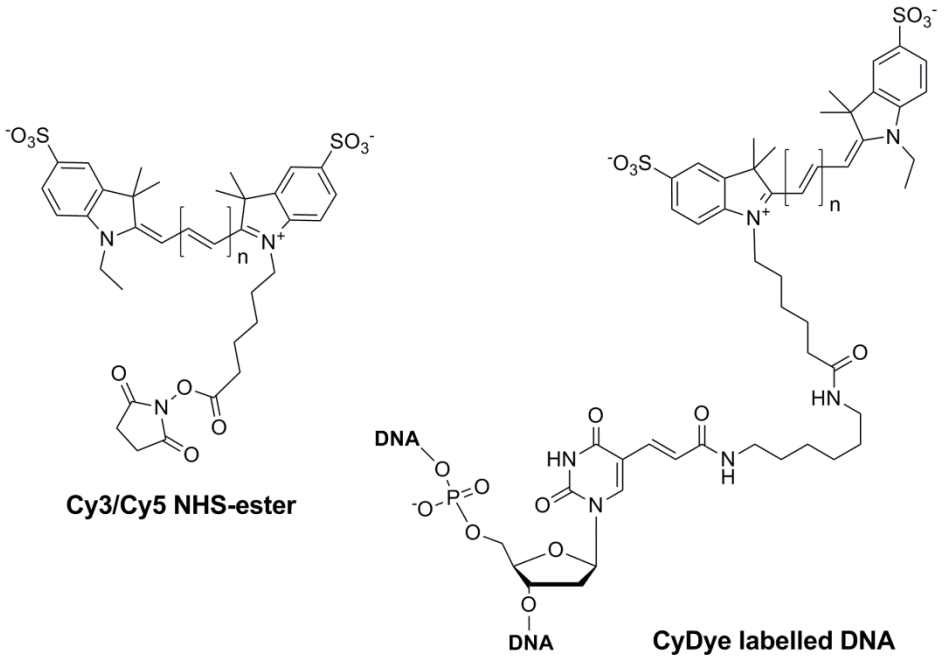



**Figure 3.3.** Chemical structures of; Cy3 (n=1) and Cy5 (n=2) NHS-esters (left); labelling of a DNA oligonucleotide with a CyDye NHS-ester (right) results a very long flexible linker.

**Table 3.1.** Photophysical properties; UV-Vis absorbance maxima ($A_{max}$), fluorescence emission maxima ($Em_{max}$), extinction coefficient (**ε**) and quantum yield (**Φ**) (recorded in MeOH).

| Compound | $A_{max}$ (nm) | ε ($M^{-1}cm^{-1}$) | $Em_{max}$ (nm) | Φ |
|---|---|---|---|---|
| Cy3-NHS ester* | 550 | 150,000 | 570 | >0.15 |
| Cy5-NHS ester* | 649 | 250,000 | 670 | >0.28 |
| ICy3 (**12**) | 551 | 114,000 | 567 | 0.04 |
| Cy3dT (**18**) | 565 | 112,000 | 591 | 0.05 |
| Cy5dT (**20**) | 660 | 208,000 | 685 | 0.06 |

*Data from GE Healthcare (and from literature[58]).

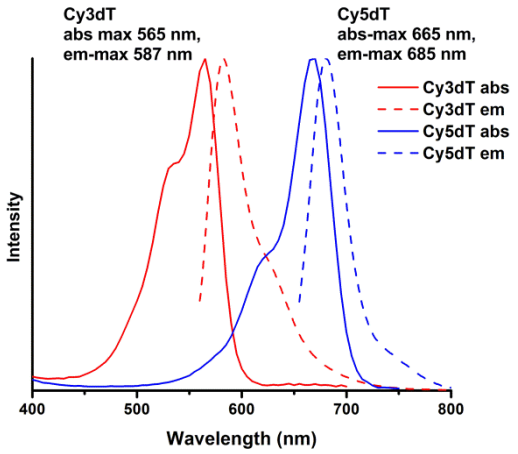



**Figure 3.4.** Normalised UV absorbance and fluorescence emission of Cy3dT (absorbance; red, emission; red dashed) and Cy5dT (absorbance; blue, emission; blue dashed) oligonucleotides. Spectra red-shifted ~15-20 nm from standard Cy3 and Cy5-NHS ester-labelled oligonucleotides ($A_{max}$ 551 nm; $Em_{max}$ 566 nm and $A_{max}$ 648 nm;$Em_{max}$ 668 nm respectively).

To establish whether the CydT monomers were good thymidine analogues, a duplex stability study was undertaken. Duplexes containing the CydT monomers (separately and together, sequences table 3.2) were compared by UV melting to an equivalent unmodified duplex. The results showed the CydT monomers slightly destabilised the DNA duplex, Cy5dT by more than Cy3dT, and the effect was found to be cumulative (table 3.3). A study of the structure of the DNA duplexes by circular dichroism (CD) showed that the duplexes containing the CydT monomers still adopted the B-DNA conformation (figure 3.5). Therefore it was concluded that the CydT analogues were suitable substitutes for thymidine and do not significantly perturb the DNA structure or stability.

**Table 3.2.** Oligonucleotide sequences (mass data in appendix A.2). **5**= Cy5dT, **3**= Cy3dT.

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN17 | GGA **5**TT TCG **3**TT TTA TAA TTG CC |
| ODN18 | GGC AAT TAT AAA AAC GAA AAT CC |
| ODN26 | CGT ATA TTC **3**TT ATT TTT AAA AGC C |
| ODN31 | GGC TTT **5**AA AAA TAA AGA ATA TAC G |
| ODN32 | CGT ATA TTC TTT ATT TTT AAA AGC C |
| ODN33 | GGC TTT TAA AAA TAA AGA ATA TAC G |

**Table 3.3.** $T_m$ data from UV melting for Cy3dT and Cy5dT duplexes [a]

| Duplex | Oligo 1 | Oligo 2 | $T_m$/ °C | Δ $T_m$/ °C |
|---|---|---|---|---|
| Unmodified control | ODN32 | ODN33 | 54.8 | --- |
| Cy3dT control | ODN26 | ODN33 | 52.4 | -2.4 |
| Cy5dT control | ODN32 | ODN31 | 51.7 | -3.1 |
| Cy3dT+Cy5dT opposite-strand | ODN26 | ODN31 | 48.9 | -5.9 |
| Cy3dT+Cy5dT same-strand | ODN17 | ODN18 | 48.6 | -6.2 |

[a]Average $T_m$ value was calculated from three successive melting curves. Standard deviation for all $T_m$ data is ± 0.1 °C.

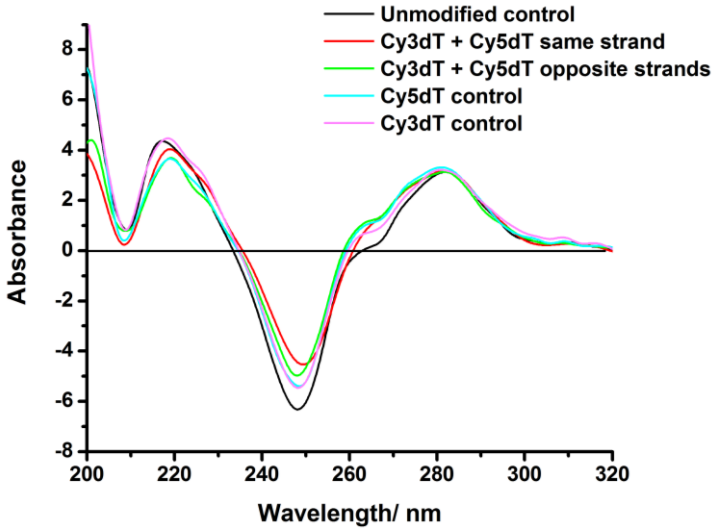



**Figure 3.5.** Circular dichroism spectra of Cy3dT and Cy5dT oligonucleotides (sequences as in Table 3.3) shows that the duplexes containing the CydT monomers still adopt the B-DNA conformation. Ten successive spectra were recorded and averaged.

## 3.3 HyBeacons probes containing Cy3dT and Cy5dT

The CydT monomers were incorporated into sequences based on HyBeacon probes (table 3.4) and tested for fluorescence emission at room temperature followed by fluorescence melting. This was to establish if the probes containing the CydT monomers exhibit any temperature dependent fluorescence properties that might contribute to their potential use as genetic probes. The fluorescence of the single stranded probe at room temperature was found to be higher than in the wild-type (fully-matched) duplex in all cases, with the single addition Cy3dT probe showing the most pronounced effect (figure 3.7). The emission maxima for the duplexes were found to shift upon melting; for the probes with a single-addition of the dye the shift was ~5 nm and the double-addition probes produced a shift of

~8 nm. The emission maxima of the single-stranded probes did not change as a function of temperature (figure 3.8). This suggests that the environment around the dye changes dramatically between single-stranded and double-stranded DNA. The wavelength shift of the probes upon melting is a potentially useful property in the context of SNP analysis.

**Table 3.4**. Oligonucleotide sequences (mass data in appendix A.2). **5**= Cy5dT, **3**= Cy3dT, **P**= propanol (figure 3.6).

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN13 | CAC CAA AGA TGA **3**AT TT**3** CTT TAA TGG **P** |
| ODN14 | CAC CAA AGA TGA TAT TT**3** CTT TAA TGG **P** |
| ODN15 | CCA TTA AAG AAA ATA TCA TCT TTG GTG |
| ODN16 | CAC CAA AGA TGA TAT TT**5** CTT TAA TGG **P** |

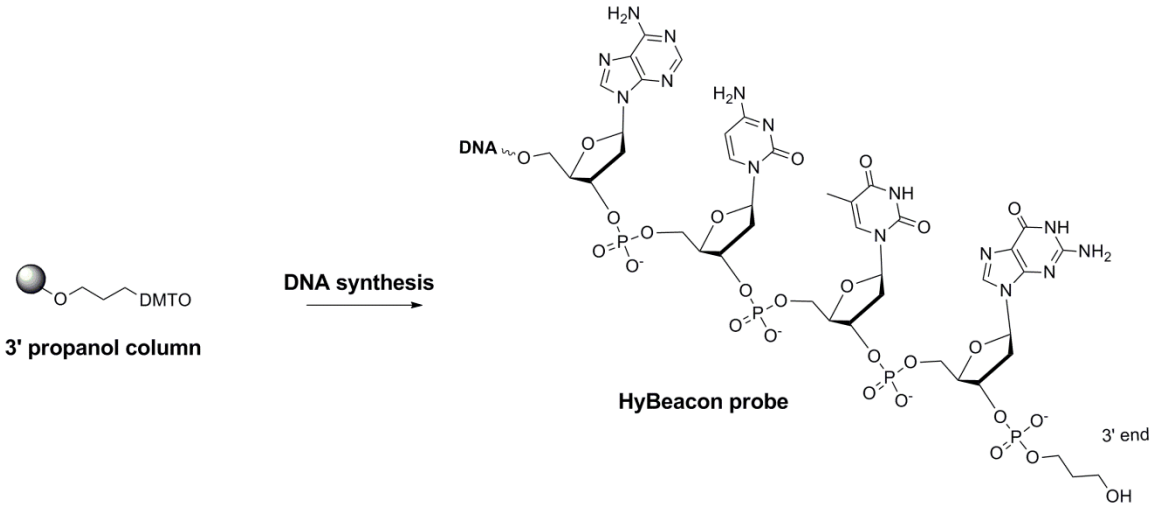



**Figure 3.6.** Chemical structure of 3'propanol stopper; incorporated into HyBeacon probes to prevent the probe acting as a PCR primer and giving rise to non-specific signals.

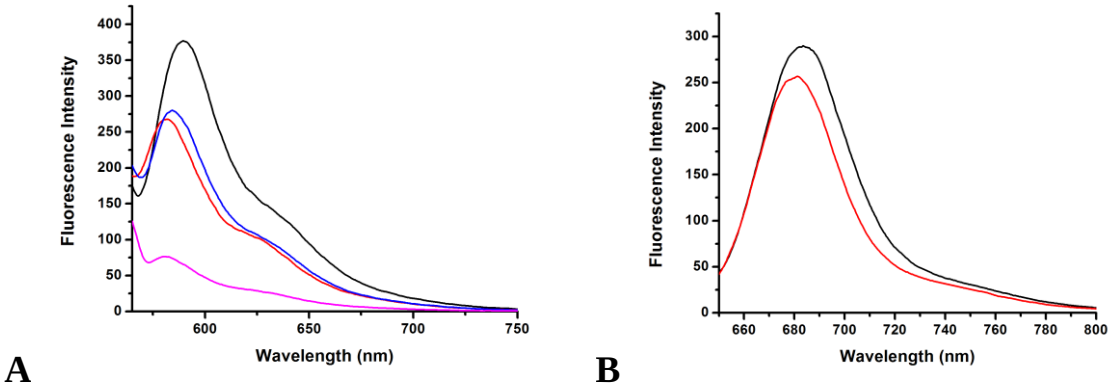



**Figure 3.7.** Room temperature fluorescence emission of; (A) Cy3dT oligonucleotides (ODN13, 2 additions: single stranded (black), double stranded (red). ODN14, 1 addition: single stranded (blue), double stranded (magenta)); (B) Cy5dT oligonucleotide (ODN16: single stranded (black), double stranded (red)). In all cases the fluorescence of the single strand is higher than that of the double-stranded DNA.

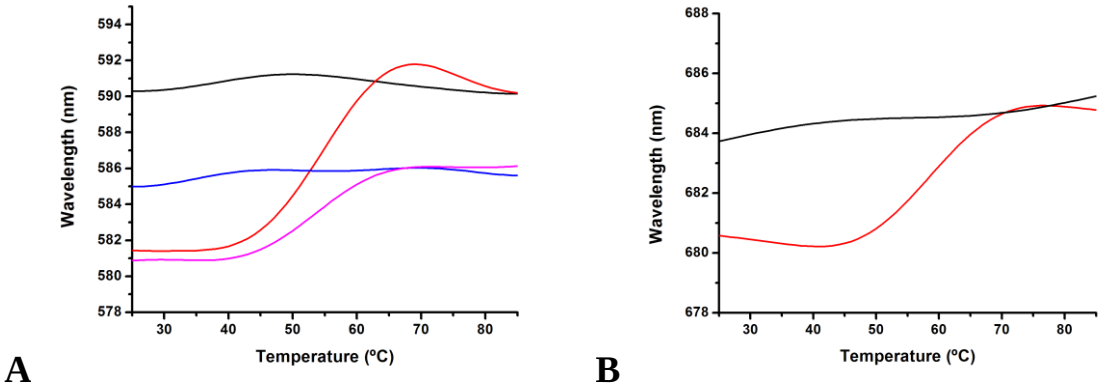



**Figure 3.8.** Monitoring the emission maximum wavelength of the probes containing CydT monomers shows that the maxima shift as the duplex samples melt but the single-stranded samples change by significantly less; (A) Cy3dT oligonucleotides (ODN13, 2 additions: single stranded (black), double stranded (red). ODN14, 1 addition: single stranded (blue), double stranded (magenta)); (B) Cy5dT oligonucleotide (ODN16: single stranded (black), double stranded (red)).

The fluorescence melting curves at constant emission wavelength displayed unusual characteristics (figure 3.9). Firstly the probes have significant general temperature dependence; in the single-strand the fluorescence decreases dramatically as the temperature increases, with the double-addition example showing the trend more clearly. Secondly the shape of the duplex melting for all examples is unusual and not in keeping with the normal melting curve expected for a standard HyBeacon (as shown in figure 1.24, chapter 1). To

explain these phenomena we propose that in the duplex form the rigid attachment of the dye to the C5-position of the base holds the dye into the major groove, here the dye is free to undergo *cis-trans* isomerism which ultimately leads to lower fluorescence (discussed below). As the duplex melts, the strand that contains the CyDye folds around the fluorophore, holding it in a more rigid conformation and therefore increasing the fluorescence. Hence at room temperature the single-stranded form is more fluorescent. As the temperature increases the single strand gradually unfolds and the dye becomes free to isomerise once again.

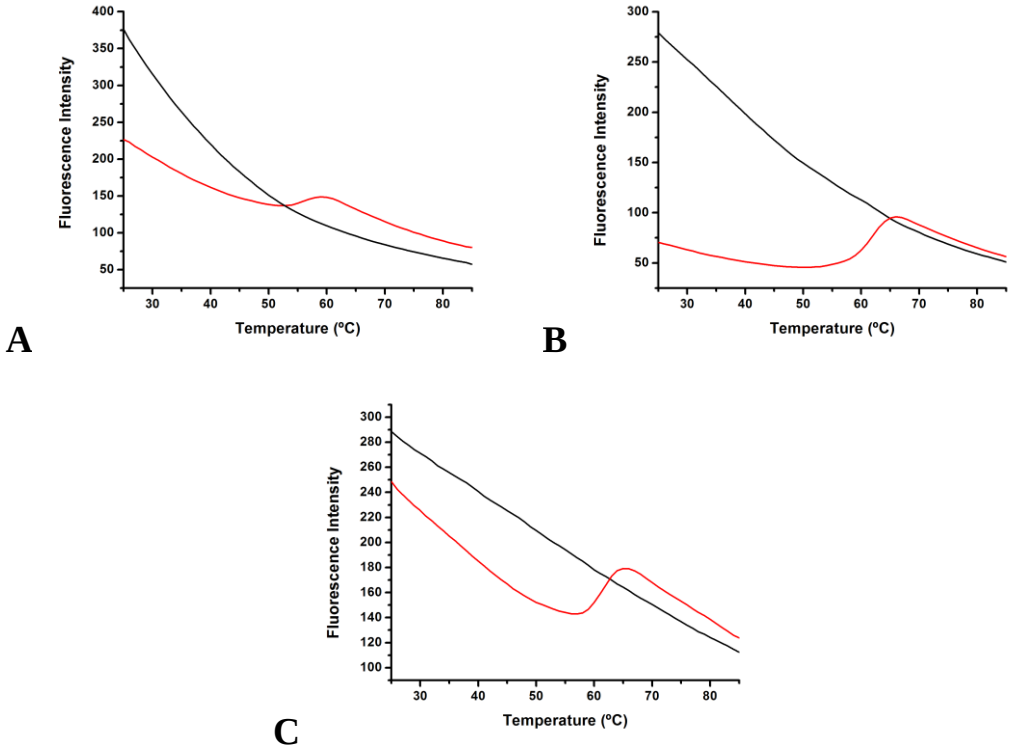



**Figure 3.9.** Fluorescence melting profiles of oligonucleotides containing CydT monomers; (A) Cy3dT oligonucleotide (ODN13, 2 additions: single stranded (black), double stranded (red)); (B) Cy3dT oligonucleotide (ODN14, 1 addition: single stranded (black), double stranded (red)); (C) Cy5dT oligonucleotide (ODN16: single stranded (black), double stranded (red)). All examples show a decrease in fluorescence as the temperature increases.

## 3.4 Isomerisation of the CyDyes

The CyDyes have the ability to isomerise between *trans* and *cis* forms by rotation around the C-C bonds of the polymethine chain, which in turn leads to reduced fluorescence.[120] The photophysical behaviour is described in figure 3.10 and shows the potential processes of isomerism from the first excited singlet state. In the excited *trans*-state, the desired

process of fluorescence emission competes with both internal conversion, reportedly insignificant,[61] and isomerism to the non-fluorescent ground-state photoisomer. The isomerisation process takes the dye through a twisted-intermediate, where the two indoles are at 90° to one another, before dropping to either the photoisomer state, reported to be entirely in the *cis* conformation,[121,122] or returning to the *trans* ground state which is the thermodynamically favoured option.[62,123] The photoisomer, if formed, then undergoes a thermal isomerisation back to the *trans* ground state.

The rate of photoisomerism is dependent on steric hindrance imposed by substituents on the dye, viscosity of the medium and temperature. Increasing the temperature of the system provides the necessary energy to overcome the activation barrier of isomerisation and so leads to more frequent isomerisation of the CyDye. In turn this means that at higher temperature the dye spends more time in the non-fluorescent photoisomer state than at room temperature. This temperature dependent fluorescence can be problematic in fluorescence melting studies of CyDyes, as shown in section 3.3. Increased solvent viscosity has been shown to result in longer fluorescent lifetimes and higher quantum yields of cyanine dyes.[124] Chibisov *et al* were able to link dependence on temperature and solvent viscosity by showing a trend of increased quantum yield with decreasing temperature for different solvents. It was explained that as temperature decreases, solvent viscosity increases and so the barrier for isomerisation cannot be overcome.[122]

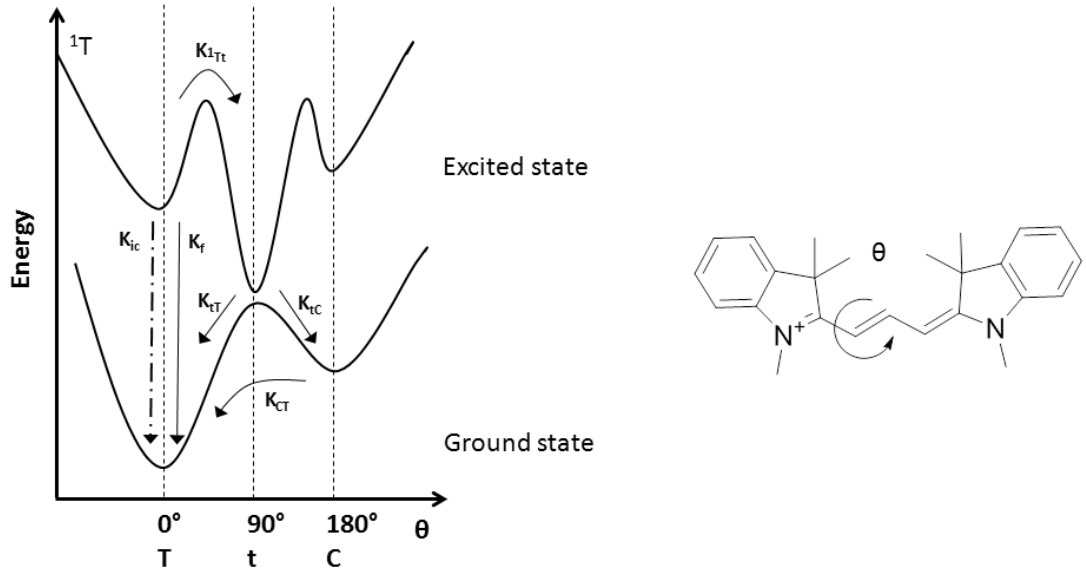



**Figure 3.10.** Potential energy diagram for the cyanine dyes (e.g. Cy3), where θ is the rotation around the polymethine backbone, T is the ground state *trans* form, $^1$T is the excited *trans* form, t is the twisted intermediate form and C is the ground state *cis* photoisomer. $K_{1Tt}$ is the rate of rotation initiating isomerism, $K_{ic}$ is rate of internal conversion and $K_f$ is rate of radiative fluorescence (figure adapted from literature[61]).

## 3.5 Demonstration of FRET in oligonucleotides containing CydT monomers

One of the most useful properties of the CyDyes is the spectral overlap of Cy3 and Cy5 which makes them a useful fluorophore pair in which context they have been utilised in many FRET studies.[60,22,24,125] To test if the new Cy3dT and Cy5dT monomers would function as a useful FRET pair, a duplex containing Cy3dT in one strand and Cy5dT in the opposite strand was excited at 530 nm, the excitation wavelength of the Cy3dT. Fluorescence emission was observed for Cy3dT at ~590 nm but also for Cy5dT at 685 nm, showing that energy was successfully transferred by FRET from Cy3dT to Cy5dT (figure 3.11).

**Table 3.5**. Oligonucleotide sequences (mass data in appendix A.2). **5**= Cy5dT, **3**= Cy3dT.

| **Oligonucleotide code** | **Sequence** | **Nucleotide separation between dyes** |
|---|---|---|
| ODN20 | CGT ATA TTC TTT ATT T**3**T AAA AGC C | 3 |
| ODN21 | CGT ATA TTC TTT ATT **3**TT AAA AGC C | 4 |
| ODN22 | CGT ATA TTC TTT AT**3** TTT AAA AGC C | 5 |
| ODN23 | CGT ATA TTC TTT A**3**T TTT AAA AGC C | 6 |
| ODN24 | CGT ATA TTC TT**3** ATT TTT AAA AGC C | 8 |
| ODN25 | CGT ATA TTC T**3**T ATT TTT AAA AGC C | 9 |
| ODN26 | CGT ATA TTC **3**TT ATT TTT AAA AGC C | 10 |
| ODN27 | CGT ATA T**3**C TTT ATT TTT AAA AGC C | 12 |
| ODN28 | CGT ATA **3**TC TTT ATT TTT AAA AGC C | 13 |
| ODN29 | CGT A**3**A TTC TTT ATT TTT AAA AGC C | 15 |
| ODN30 | CG**3** ATA TTC TTT ATT TTT AAA AGC C | 17 |
| ODN44 | GGC T**5**T TAA AAA TAA AGA ATA TAC G | --- |
| ODN32 | CGT ATA TTC TTT ATT TTT AAA AGC C | --- |
| ODN33 | GGC TTT TAA AAA TAA AGA ATA TAC G | --- |

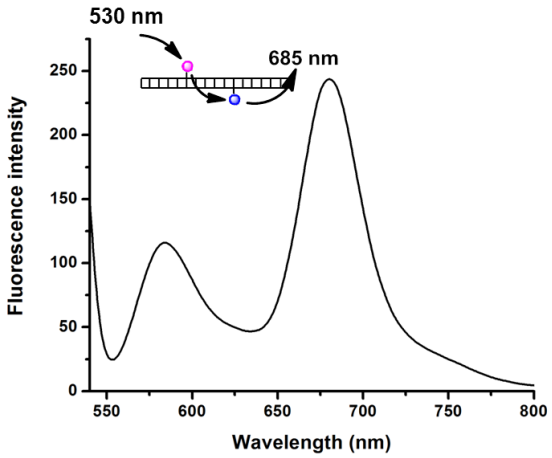



**Figure 3.11.** Fluorescence emission of duplex (ODN23+ODN44) excited at 530nm. FRET efficiency = 0.73 for a separation between dyes of 6 nucleotides.

A FRET study was then undertaken to examine the energy transfer from Cy3dT to Cy5dT more thoroughly. A combination of 11 Cy3dT oligonucleotides and a single Cy5dT oligonucleotide (plus control unmodified oligonucleotides) were synthesised to give base separations between the two dyes ranging from 3-17 base pairs (table 3.5). The fluorescence emission of the Cy5dT in the duplexes was measured after excitation at 530 nm (FRET excitation) and 615 nm (direct excitation). The FRET efficiency of the samples was calculated according to equation 5 and was plotted against both nucleotide separation between the dyes (figure 3.13A) and against the measured distance between the dyes (figure 3.13B). In the second case the distance is taken between the midpoint of the polymethine linker in each CyDye as defined in figure 3.12.

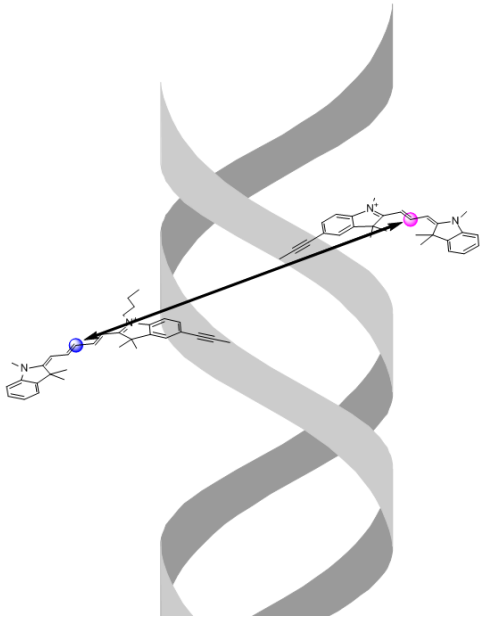


**Figure 3.12.** Schematic DNA helix (grey) demonstrating distance measurement used in FRET efficiency comparison. Distance taken between the middle carbon of the polymethine linker of Cy5dT (blue sphere) and Cy3dT (pink sphere).

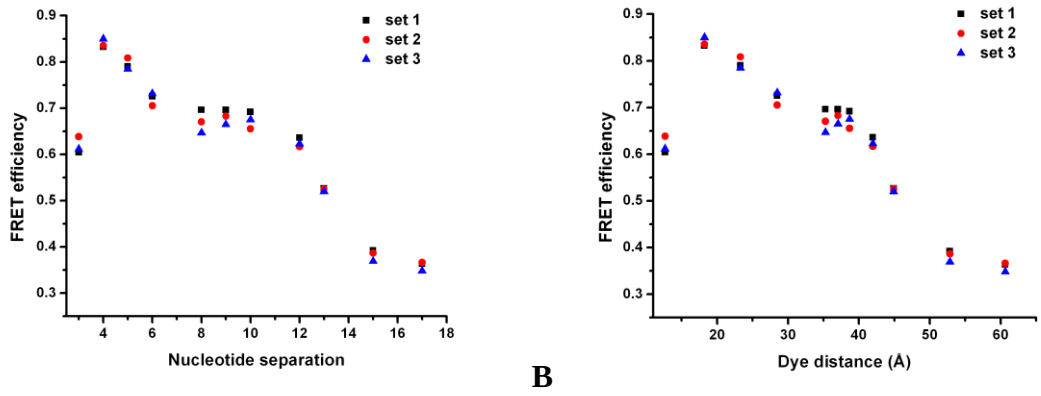



**Figure 3.13**. (A) FRET efficiency vs. number of nucleotides separating dyes (not including CydT nucleobases); (B) FRET efficiency vs. dye-dye distance (as defined in figure 3.12). 3 experimental sets. FRET efficiency calculated according to equation 5.

$$\text{FRET efficiency} = (I_{FRET} \times Abs_{615})/(I_{direx} \times Abs_{530}) \qquad \text{equation 5}$$

Where $I_{FRET}$ is the emission intensity of the Cy5 peak when excited via fret (ex 530 nm), $I_{direx}$ is the emission intensity of Cy5 peak when directly excited (ex 615 nm), $Abs_{615}$ is the absorbance of the sample at 615 nm and $Abs_{530}$ is the absorbance of the sample at 530 nm.

As expected, a decrease in FRET efficiency was exhibited as the dyes were moved further apart. Some modulation of the FRET curve is apparent when FRET efficiency is shown against nucleotide separation (figure 3.13A). However when FRET efficiency is shown against dye-dye distance (in Angstroms, figure 3.13B), the curve is more linear. This is because, in this system where the dye (and hence measurement point) is on the periphery of the helix, an incremental increase in nucleotide separation does not always translate to a regular increase in distance, as demonstrated in figure 3.14. For example as the separation between the dyes is increased from 8 to 10 nucleotides the Cy3 successively moves two nucleotides further along the duplex, increasing the length between the dyes in the z-axis (helical axis), however the helical turn brings the Cy3 closer in angle to the Cy5 when viewed down the z-axis. Therefore the Cy3-Cy5 distance increases only a small amount, from 35.3 Å – 38.7 Å, in this region.

At low nucleotide separation the efficiency dropped, most likely as a result of quenching between the dyes in close proximity. We observed 50 % FRET efficiency in this system at

around 47 Å separation, which compares well with the Förster distance of standard Cy3-Cy5 (50 Å). However, at low base separations, <20 Å, we were expecting FRET efficiency to be higher. The lower than expected FRET might be due to loss of energy through isomerism of the dyes. This is conceivable for CydT, as the dye is free to rotate in the major groove, but would be less prevalent in the standard CyDyes, which tend to stack in the DNA, either on the end of the duplex or in the major or minor groove. Therefore the apparent Förster distance of Cy3dT-Cy5dT may actually be less than that of the standard Cy3-Cy5 pair due to these slightly different properties. In addition, the freedom of mobility of the dyes will result in some movement of the dipole itself which would account for the fact that the maxima and minima of the FRET efficiency curve are not well defined. In comparison a study by Rindermann *et al* a FRET system containing dyes bound by long flexible linkers likewise gave undefined maxima and minima in the FRET efficiency curve.[126]

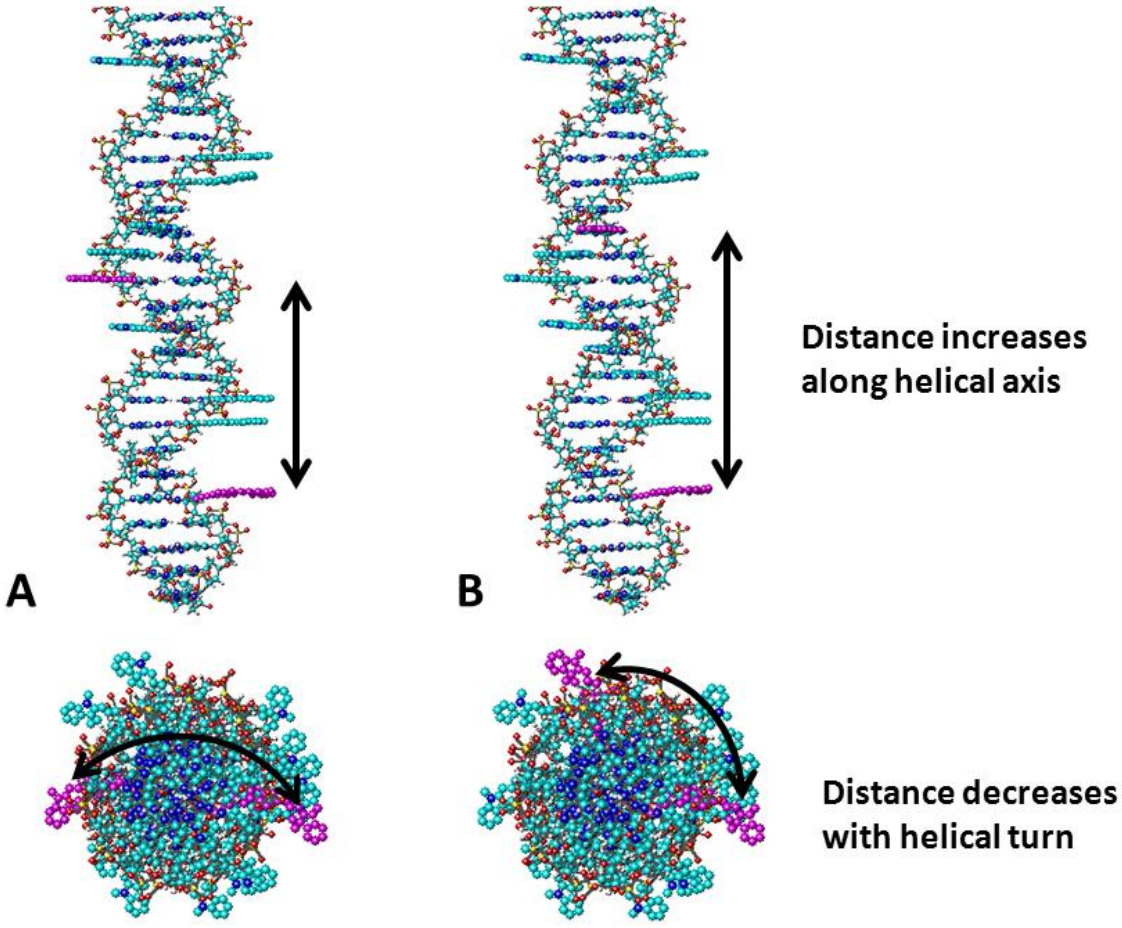



**Figure 3.14.** Demonstration of distance between Cy3dT and Cy5dT (purple) in the DNA helix in terms of helix length (top) and helical turn (bottom); (A) Cy3dT and Cy5dT separated by 8 nucleotides; (B) Cy3dT and Cy5dT separated by 10 nucleotides. An incremental increase of 2 nucleotides between the dyes results in an increase in distance along the helical axis which is counteracted by the helical turn bringing Cy3dT closer to Cy5dT.

The concept of the dye isomerism complicating the FRET study was confirmed by other collaborative efforts involving time resolved fluorescence which indicated that complicated multi-exponential fluorescence decays were present. It is hoped in the future a suitable experiment can be designed in which the *cis-trans* isomerism can be eradicated; most likely by preparing the samples in glycerol and carrying out the measurements at low temperature. Armitage and co-workers have shown that the fluorescence of a cyanine dye increases as a function of solvent viscosity, for example by using glycerol.[127]

Wilhelmsson *et al* have shown that rigid base analogues gave completely predictable dipoles such that FRET efficiency accurately reflected the change in both distance and helical periodicity.[128] Another study with Cy3 and Cy5 stacked on opposite ends of varying length duplexes was intended to achieve the same result.[60] However, the FRET efficiency curve was less defined than predicted. One theory presented is that a certain percentage of the dyes may be un-stacked from the end of the DNA, resulting in free mobility of the dyes around their linkers. This concept is confirmed by the work of Levitus in which dsDNA with 5'-Cy3 exhibited lower fluorescence compared to the equivalent ssDNA.[61] They attribute this behaviour to a proportion of Cy3 molecules becoming un-stacked from the DNA. The efficiency of isomerisation of the dye is increased and therefore the fluorescence decreases. In addition they found that the fluorescence quantum yield of the dsDNA with 5'-Cy3 was more sensitive to temperature, and the fluorescence decay was faster.

## 3.6 Oligonucleotides with multiple CyDyes

Oligonucleotides with multiple dye additions on the same strand were investigated to explore the advantages of the CydT monomers over conventional post-synthetic labelling methods. Oligonucleotides of the following composition were synthesised: two Cy3dT additions ten base pairs apart (ODN36); five Cy3dT additions on alternate bases (ODN37); and six consecutive Cy3dT additions (ODN40) (table 3.6). A complementary oligonucleotide to the five addition example was made also with five alternate Cy3dT additions (ODN38), giving the potential of a duplex containing ten-additions of Cy3dT (figure 3.15). Synthesis and purification of the heavily modified oligonucleotides resulted in clean products which would not be achievable by other established methods (figure 3.16).

**Table 3.6**. Oligonucleotide sequences (mass data in appendix A.2). **3**= Cy3dT.

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN34 | ATG GCT ATT AAA AAC GGA GAA TGT ACG CAA |
| ODN35 | TTG CGT ACA TTC TCC GTT TTT AAT AGC CAT |
| ODN36 | TTG CGT ACA **3**TC TCC GTT T**3**T AAT AGC CAT |
| ODN37 | ATG GCT ATT A**3**A **3**A**3** A**3**A **3**AA TGT ACG CAA |
| ODN38 | TTG CGT ACA T**3**A **3**A**3** A**3**A **3**AT AAT AGC CAT |
| ODN39 | TTG CGT ACA TTA TAT ATA TAT AAT AGC CAT |
| ODN40 | ATG GCT ATT A**33 333 3**AT GTA CGC AA |
| ODN41 | TTG CGT ACA TAA AAA ATA ATA GCC AT |
| ODN42 | ATG GCT ATT ATA TAT ATA TAA TGT ACG CAA |
| ODN43 | ATG GCT ATT ATT TTT TAT GTA CGC AA |

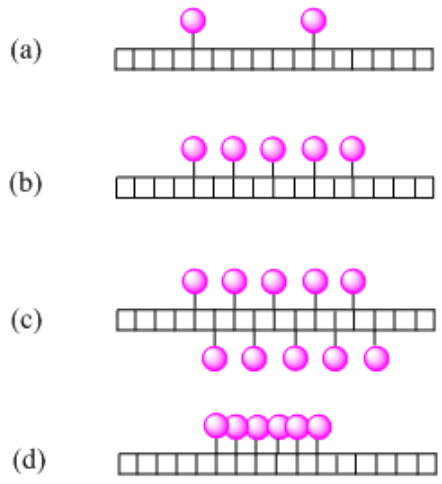



**Figure 3.15.** Oligonucleotide duplexes containing multiple-additions of Cy3dT; (a) HyBeacon format (ODN36+ODN34); (b) five alternating additions in a single strand (ODN37+ODN39); (c) ten alternating additions across the duplex (ODN37+ODN38); (d) six consecutive additions in a single strand (ODN40+ODN41).

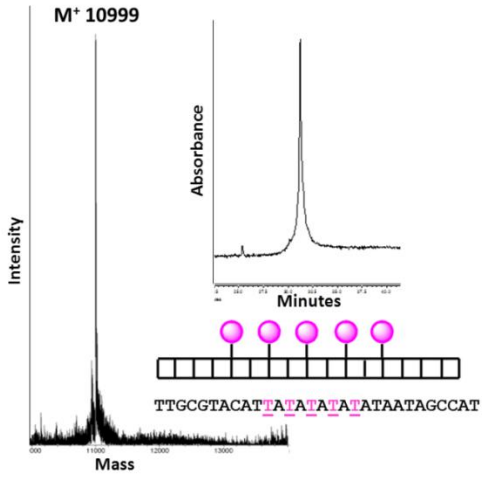



**Figure 3.16.** Mass spectrum and capillary electrophoresis for an oligonucleotide containing five additions of Cy3dT.

Circular dichroism analysis of the CydT-modified oligonucleotide duplexes compared to an equivalent oligonucleotide of the same sequence showed the DNA conformed to B DNA structure; hence the multiple additions did not significantly perturb the structure of the DNA (figure 3.17).

UV melting analysis was used to assess the effect of the additions on the stability of the duplex. As expected, the oligonucleotide duplex with two dye-additions (ODN36+ODN34) shows destabilisation of ~5.3 °C (~2.6 °C per addition, table 3.7). Increasing to five additions (ODN37+ODN39) destabilised the duplex by ~10.7 °C (~2.1 °C per addition). The duplex with ten additions gave a broader melting curve, of which the maximum of the derivative peak is 3.5 °C higher than its unmodified equivalent. This suggests that modifying one strand destabilises the duplex but modifying both strands stabilises it, perhaps by some intrastrand interaction between the dyes, hence accounting for the broad melting seen (figure 3.18). Six consecutive additions gave a melting curve that is too broad to represent a cooperative duplex to single-strand transition; the derivative peak is at ~5.7 °C below its unmodified equivalent but the shape of the curve makes it impossible to assess whether a duplex is actually formed. To summarise, comparing the duplexes containing Cy3dT shows that upon increasing the number of Cy3dT additions from two to ten, then to six consecutive additions, the UV melting curves become progressively broader i.e. the duplex to single-strand transition becomes less cooperative with increasing number of modifications. Nevertheless it is noteworthy that stable DNA duplexes can be formed from oligonucleotide strands containing multiple additions of Cy3dT.

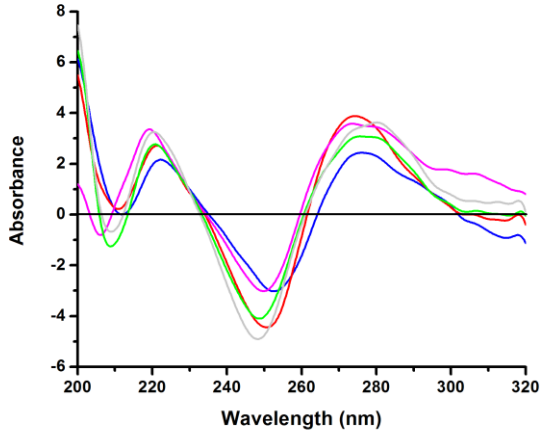



**Figure 3.17.** CD spectra of oligonucleotide duplexes containing multiple-additions of Cy3dT; ODN37+ODN38 (ten additions alternating across duplex, blue); ODN37+ODN39 (five alternating additions in one stand, red); ODN40+ODN41 (six consecutive additions in one strand, magenta); ODN36+ODN34 (two additions in one strand, green); ODN35+ODN34 (unmodified duplex, light grey).

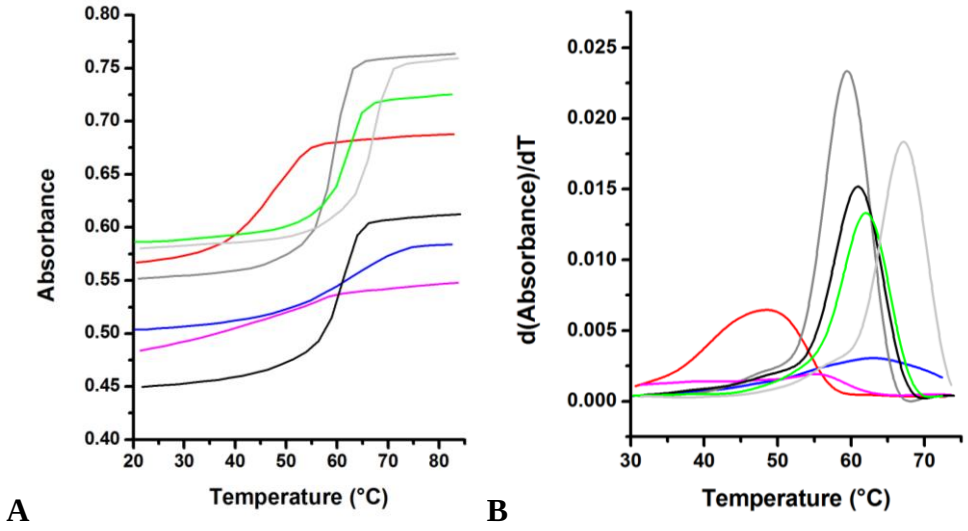



**Figure 3.18.** UV melting of oligonucleotide duplexes containing multiple-additions of Cy3dT; (A) UV melting spectra; (B) UV melting derivatives. ODN37+ODN38 (ten additions alternating across duplex, blue); ODN37+ODN39 (five alternating additions in one strand, red); ODN42+ODN39 (unmodified duplex, dark grey); ODN40+ODN41 (six consecutive additions in one strand, magenta); ODN43+ODN41 (unmodified duplex, black); ODN36+ODN34 (two additions in one strand, green); ODN35+ODN34 (unmodified duplex, light grey).

**Table 3.7.** $T_m$ data from UV melting for multiple addition Cy3dT duplexes [a]

| Duplex | Oligo1 | Oligo2 | $T_m$/ °C | Δ $T_m$/ °C |
|---|---|---|---|---|
| 10 addition- alternating | ODN37 | ODN38 | 63.0 (broad) | +3.5 |
| 5 addition-alternating | ODN37 | ODN39 | 48.9 (broad) | -10.7 |
| Unmodified control | ODN42 | ODN39 | 59.5 | --- |
| 6 addition- consecutive | ODN40 | ODN41 | 55.3 (v. broad) | (-5.7) |
| Unmodified control | ODN43 | ODN41 | 61.0 | --- |
| 2 addition | ODN36 | ODN34 | 61.9 | -5.3 |
| Unmodified control | ODN35 | ODN34 | 67.2 | --- |

[a] Average $T_m$ value was calculated from three successive melting curves. Standard deviation for all $T_m$ data is ± 0.5 ºC.

The duplex with six consecutive additions of Cy3dT (ODN40+ODN41) exhibited B-helix characteristics by CD, but by UV melting appeared to give a very broad transition. It was decided to analyse the duplexes by native PAGE (12 % polyacrylamide gel) to assess whether they were in fact forming duplexes or whether the modifications were causing intra-strand DNA structures. As seen in figure 3.19, duplexes were observed for the two, five and ten addition examples. The duplex with six consecutive additions however did not migrate through the gel but instead remained bound at the top of the gel. The single stranded sample exhibited the same behaviour. It was decided to repeat analysis of the six-addition example with a gel of lower percentage acrylamide (5 % polyacrylamide gel) to give a larger pore size and hence facilitate migration of the oligonucleotide (figure 3.20). Duplex samples were prepared in oligonucleotide strand ratios from 1.0:0.5 to 1.0:1.5 (modified:unmodified) to determine whether any duplex formed. Although the modified oligonucleotide migrated through the gel there was no evidence of duplex formation, even with the modified strand in excess. However it is possible that an unstable duplex will separate into single strands under PAGE conditions and that the individual strands migrate differently on the gel and are therefore kept apart.

To investigate the system further, UV melting and CD analysis were performed on the single strands (figure 3.21). In the CD spectrum the single-stranded unmodified oligonucleotide ODN35 exhibits typical weak B-helix characteristics;[129] this pre-organisation of B-DNA is why the B helical structure is so stable. As the number of Cy3dT additions is increased, the B-helix characteristics are gradually reduced, with ODN40 single strand displaying no defined peaks or troughs. The UV melting spectra of the same

single-stranded sample displays no defined melting curve and is very different from the ODN40+ODN41 duplex sample.

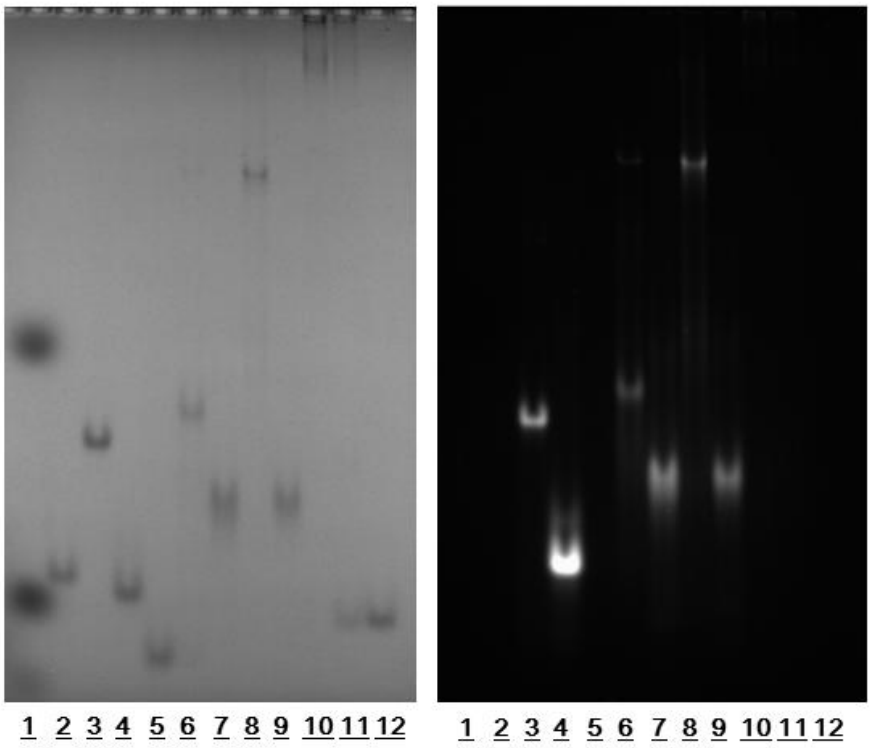



**Figure 3.19.** Native PAGE (12 % acrylamide) of single strands and duplexes containing multiple-additions of Cy3dT; (1) xylene cyanol and bromophenol blue reference dye; (2) ODN34 ss (unmodified); (3) ODN36+ODN34 (two additions in one strand); (4) ODN36 ss (two additions); (5) ODN39 ss (unmodified); (6) ODN37+ODN39 (five alternating additions in one strand); (7) ODN37 ss (five alternating additions); (8) ODN37+ODN38 (ten additions alternating across duplex); (9) ODN38 ss (five alternating additions); (10) ODN40 ss (six consecutive additions); (11) ODN40+ODN41 (six consecutive additions in one strand); (12) ODN41 ss (unmodified). For gel imaging condition see experimental section (6.3.4.1).

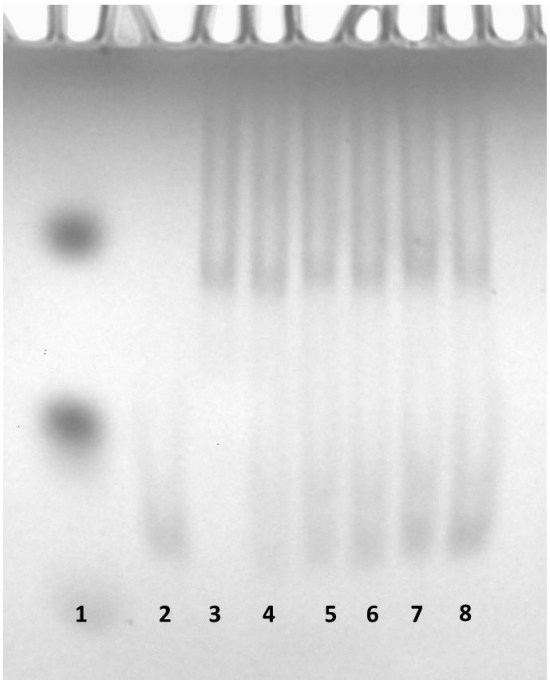



**Figure 3.20.** Native PAGE (5 % acrylamide). (1) xylene cyanol and bromophenol blue reference dye; (2) ODN41 ss (unmodified); (3) ODN40 ss (six consecutive additions); (4) ODN40+ODN41 1:0.5; (5) ODN40+ODN41 1:0.8; (6) ODN40+ODN41 1:1; (7) ODN40+ODN41 1:1.2; (8) ODN40+ODN41 1:1.5. No residual pink oligonucleotide could be seen remaining at the top of the gel in the well cavity. For gel imaging condition see experimental section (6.3.4.1).

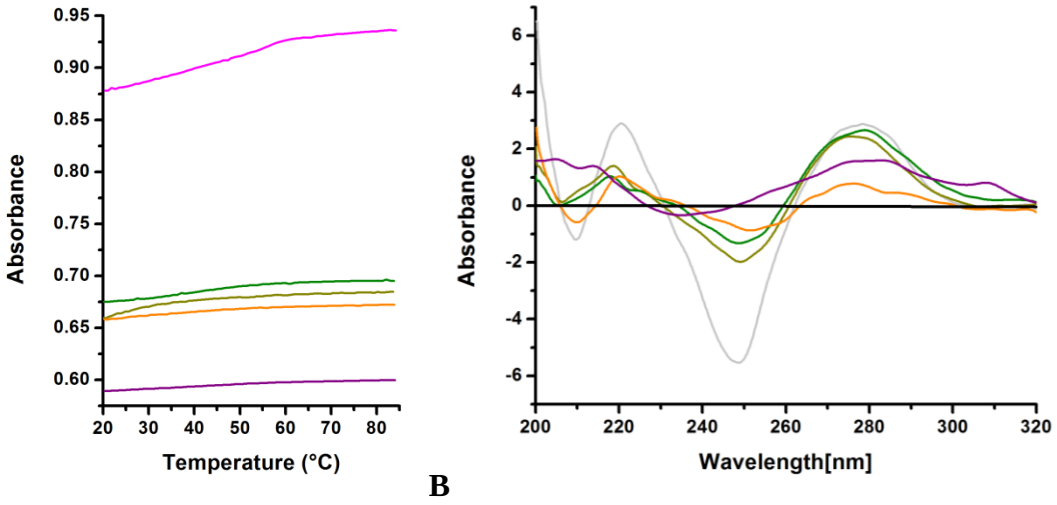



**Figure 3.21.** UV studies of single-stranded oligonucleotides containing multiple-additions of Cy3dT; (A) UV melting curves; (B) CD spectra. ODN35 (unmodified, dark yellow); ODN36 (two additions, dark green); ODN37 (five alternating additions, orange); ODN40 (six consecutive additions, purple); ODN40+ODN41 (six consecutive additions in one strand, pink); ODN35+ODN34 (unmodified, light grey).

To investigate the fluorescence properties of the multiple addition oligonucleotides, the UV-Vis absorbance and fluorescence emission spectra of the duplexes were recorded. The absorbance spectrum exhibited by the duplex with two additions of Cy3dT was the only one that remotely resembled that of a single-addition Cy3dT oligonucleotide spectrum (figure 3.22). The other examples exhibited additional blue-shifted absorbance peaks typical of excimer-like structures. This would suggest that the five, ten and six addition examples have varying degrees of interaction between the dyes. It was found that excitation at the wavelength of the new absorbance peak resulted only in overall decreased fluorescence emission.

The phenomenon of dye-dimers has been studied in depth and, as explained by Hernando *et al*, the blue-shifted peak seen in the absorbance spectrum is characteristic of an H-dimer in which molecules stack with little or no offset.[130] In the multiple addition examples the Cy3dT dyes are able to stack one upon the other to varying degrees, creating strong coupling effects which result in electron delocalisation over the entire chromophore system, causing it to act as an entirely new quantum system. The system possesses a new ground state and the electronically excited state splits into two. Fluorescent emission from the lower of the two excited states is forbidden and so the dimer exhibits a decrease in fluorescence emission.[131] It has been proposed that the dimer has two different absorbing species, the emitting species which displays the same absorbance as a single-addition of the monomer and one which is non-fluorescent and is responsible for the blue-shifted absorption band.[130] This explains why excitation of the new peak does not translate into increased fluorescence emission.

The fluorescence emission of the Cy3dT oligonucleotide duplexes is counterintuitive in terms of fluorescence emission compared to number of dye-additions, with the highest fluorescent output exhibited by the least number of dyes. A significant decrease in fluorescence is seen for the five-addition example, which increases slightly (but not double) for the ten-addition example. The six-consecutive addition example has very low fluorescence. This suggests that increasing the number of Cy3dT dyes that are in close proximity encourages excimer formation resulting in quenching of the system.

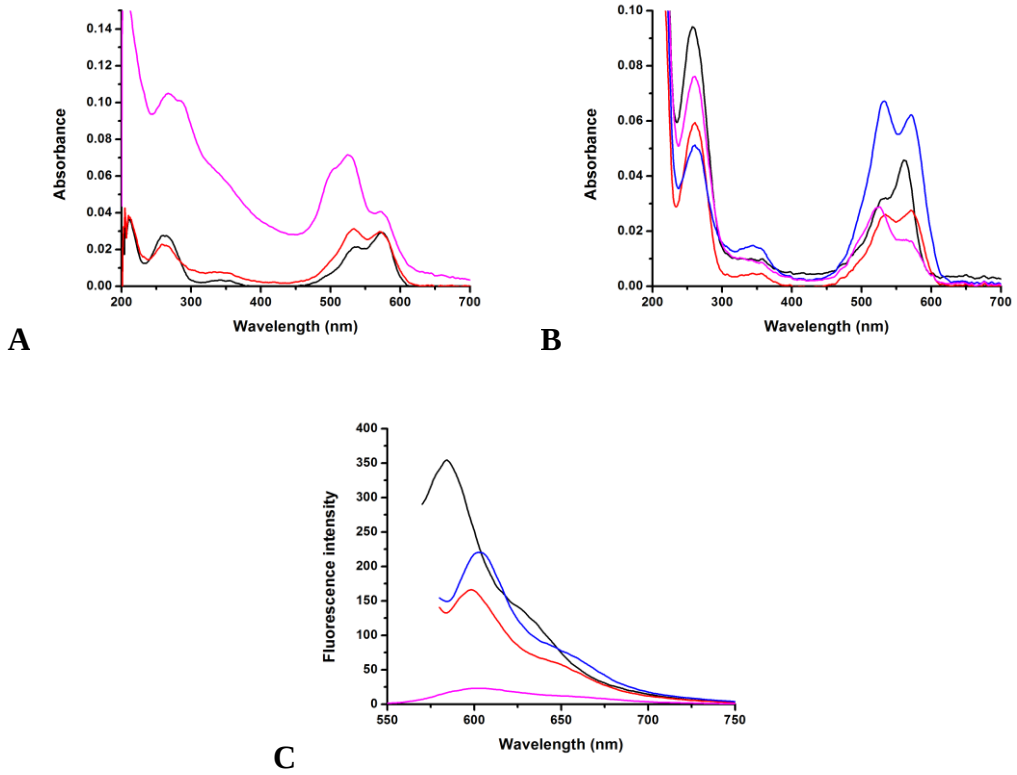



**Figure 3.22.** Absorbance and emission spectra of oligonucleotides containing multiple-additions of Cy3dT; (A) single-stranded UV-Vis absorbance, ODN36 (two additions, black), ODN37 (five alternating additions, red), ODN40 (six consecutive additions, magenta); (B) double-stranded* UV-Vis absorbance, ODN36+ODN34 (two additions in one strand, black), ODN37+ODN39 (five alternating additions in one strand, red), ODN37+ODN38 (ten additions alternating across the duplex, blue), ODN40+ODN41 (six consecutive additions in one strand, magenta); (C) double-stranded* fluorescence emission, ODN36+ODN34 (black), ODN37+ODN39 (red), ODN37+ODN38 (blue), ODN40+ODN41 (magenta).

* In the case of ODN40+ODN41 it is not certain that a double-stranded structure is forming.

Wagenknecht and co-workers have studied excimers in DNA using various fluorescent chromophores such as 5-(pyrene-1-yl)-2'-deoxyuridine, 5-(10-methyl-phenothiazin-3-yl)-2'-deoxyuridine[132] and perylene-3,4:9,10-tetracarboxylic acid bisimides.[133] They reported excimer-like fluorescence in all cases of alternating or consecutive dye-additions. In their study of 1-ethynylpyrene excimers they noted that one or two additions destabilised the duplex but five consecutive additions caused stabilisation.[134] In addition the fluorescence emission intensity increased with increasing numbers of dye-additions, unlike our findings with Cy3dT. However in their study of thiazole orange dimers they observed strong excitonic interactions when the dyes were side-by-side which resulted in fluorescence quenching.[135] When the thiazole orange monomers were in opposite DNA strands they observed stabilisation of the duplex which they attributed to intrastrand interactions, much

like that seen in our ten-addition Cy3dT example. Their recent study reports a Nile-Red analogue which has the same dye-DNA linkage as Cy3dT.[136] They found that one addition of the monomer destabilised the DNA (-3.6 ºC) but for several additions the effect was not so severe (-2.8 ºC per addition). An increased blue-shifted absorbance and huge reduction in emission was observed when increasing from one addition to three or five consecutive additions. This was attributed to the formation of Nile-Red H-aggregates along the DNA backbone, producing strong excitonic interactions, as seen in our multiple Cy3dT examples.

Many more excimer examples have been studied with combinations of bipyridyl- and biphenyl-DNA,[137] pyrenes,[138-140] perylene,[141,142] and porphrins[143]. Armitage and co-workers prevented the formation of quenching dye excimers by allowing the free cyanine dyes to intercalate into DNA.[144] They found that saturation was reached with a dye addition every other base pair, placing them approximately 7 Å apart. The intervening base pairs prevented the dyes from quenching one another so the result was a highly fluorescent nanostructure, in which the brightness increased with increasing numbers of base pairs. In the case of the Cy3dT excimers there is no barrier between the dyes in the major groove to prevent quenching and so increased dye-spacing must be investigated to attain a useful multi-fluorophore system.

Our studies show that oligonucleotides containing multiple additions of Cy3dT can be successfully synthesised, but that incorporating the dyes very close together, in large numbers, can result in both quenching and secondary structure formation. From the combination of UV melting, CD and PAGE analysis it can be seen that increasing numbers of dye additions progressively affects the structure of the single-strand and the formation of the DNA duplex. In the case of the six-consecutive dye-additions (ODN40) no definite conclusion can be drawn, but it is possible that single- or double-stranded structures and higher aggregates caused by dye-dye stacking might be in equilibrium. Further studies on these systems using long-wavelength CD covering the region in which the CyDye absorbs might reveal more information about these structures.

## 3.7 Conclusions

Two new fluorescent phosphoramidite monomers based on CyDyes have been synthesised which permit the addition of multiple and mixed fluorophores into oligonucleotides during solid-phase synthesis. The properties of these monomers have been investigated and they have been shown to be viable thymidine analogues in terms of base pair formation. Multiple additions of the Cy3dT monomer into a single oligonucleotide was successfully achieved. Up to ten dye additions have been successfully incorporated into one duplex, demonstrating the power of this methodology. Increasing the number of dye-additions results in the formation of excimer structures between the dyes, resulting in quenching of the fluorescent system. In the future careful investigations will be undertaken to establish optimum spacing of the dyes to yield maximum fluorescence and duplex stability.

Oligonucleotides containing Cy3dT and Cy5dT were used to constitute a FRET system. By using a series of duplexes with varying spacing between the two dyes, FRET was shown to decrease with increasing inter-nucleotide separation. Oligonucleotides containing Cy3dT and Cy5dT exhibited unusual room-temperature fluorescence emission and fluorescence melting properties. The fluorescence intensity of oligonucleotides containing CydT was found to be lower in double-stranded form and at higher temperatures. It was proposed that, despite the presence of the rigid linker to DNA, in the duplex the location of the dyes in the major groove makes them susceptible to the rotation around polymethine linker, leading to *trans-cis* isomerisation. This leads to reduced fluorescence and inverse temperature dependent fluorescence. The possibility of isomerisation of the CyDye could be a disadvantage in some applications, so we proceeded to investigate the more rigid analogue Cy3B.

# Chapter 4

# Rigid CyDye Analogue

# Chapter 4 – Rigid CyDye Analogue

## 4.1 Cy3B; a rigid CyDye

As discussed in chapter 3, isomerisation of Cy3 and Cy5 around the polymethine linker leads to temperature dependent fluorescence changes in the Cy3dT and Cy5dT monomers. Oligonucleotides containing Cy3dT and Cy5dT also exhibit lower fluorescence in double-stranded DNA than in the single-stranded form. These factors mean that Cy3dT and Cy5dT are not ideal for use in duplex-DNA studies. A more recently developed CyDye, Cy3B,[64] exhibits superior fluorescent properties compared to its Cy3 counterpart due to the rigid backbone along the polymethine linker, figure 4.1. The backbone prevents the *cis/trans* isomerisation and, without the ability to rotate around the polymethine linker, Cy3B is held permanently in the fluorescent *trans* form. As a consequence the fluorescence lifetime and quantum yield of Cy3B are significantly higher than for Cy3 (Cy3B has a lifetime of 2.7-2.8 ns and a quantum yield of 0.67-0.85,[64,61] Cy3 has a lifetime of <0.3 ns and a quantum yield of >0.15 (Cy3 data from GE Healthcare)). The fluorescence of Cy3B is therefore not dependent on temperature, making it a far more versatile dye than Cy3 or Cy5.

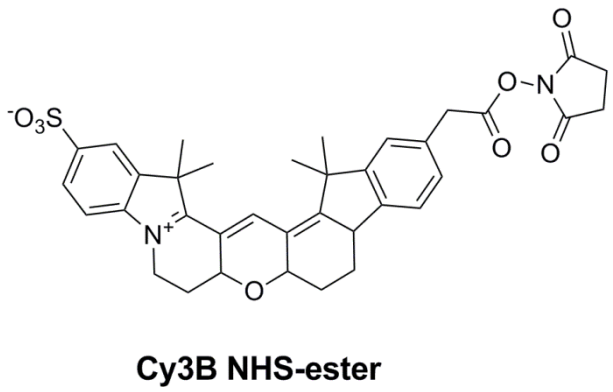



**Figure 4.1.** Chemical structure of Cy3B NHS-ester. The rigid backbone prevents *cis/trans* isomerisation of the polymethine linker.

Cy3B is commercially available as the NHS-ester or maleimide form (for the post-synthetic modification of oligonucleotides or amino- or thiol-modified biomolecules respectively) and not as a phosphoramidite monomer, making it uses limited. It has been utilised in single-molecule studies, cell imaging and fluorescence studies[145-147] but the

photophysical studies on DNA have been surprisingly limited, perhaps due to the high cost of the dye and the limited options for dye-addition.

Therefore it was decided to synthesise an array of Cy3B analogues, concentrating mainly, but not solely, on phosphoramidite addition to oligonucleotides.

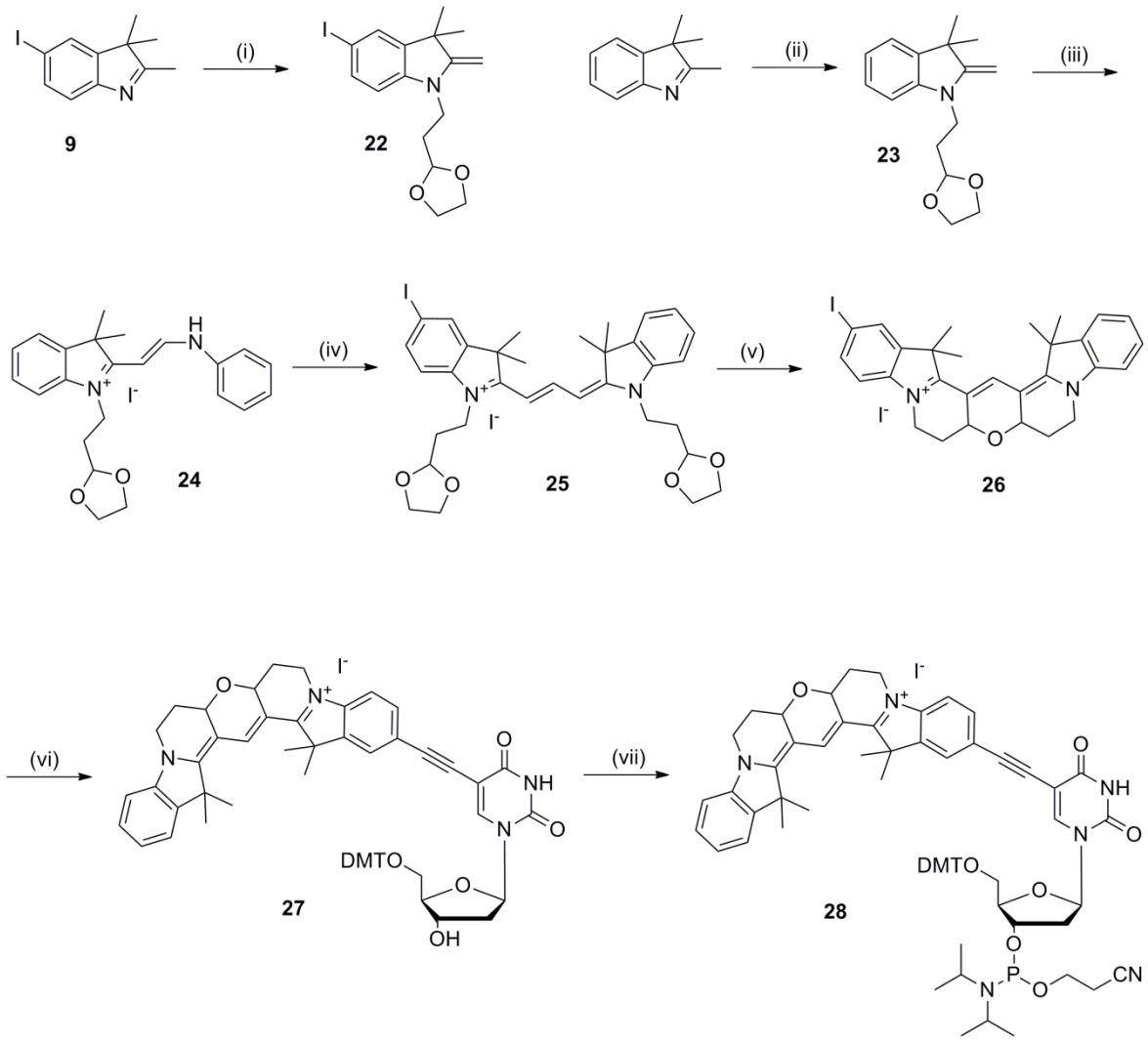
## 4.2 The synthesis of a Cy3B 'molecular toolkit'

### 4.2.1 The synthesis of Cy3BdT phosphoramidite

The synthesis of a Cy3BdT analogue was based on the schemes already designed for Cy3dT and Cy5dT. The synthetic details for the preparation of Cy3B are not as thoroughly documented as for Cy3 and Cy5 and so the commercial patent[148] was examined. Some changes were made to the synthetic procedure from the patent to successfully complete the synthesis. The patent described the addition of a diethylacetal group at the 1N of the indole, however this was found to be too labile for convenient use. It was therefore changed to an ethyl-1,3-dioxolane group (scheme 4.1). The protected indoles were still quite unstable and so were used immediately after their synthesis to prevent degradation. The synthetic steps which followed to give the hemicyanine **24** and cyanine **25** compounds were undertaken using the same conditions as for Cy3 (scheme 3.1). These steps were followed by deprotection of the dioxolane groups which resulted in a rapid tandem cyclisation reaction giving compound **26** in 96 % yield. The palladium cross-coupling of 5-iodo-Cy3B with 5-ethynyl-dU was followed by phosphitylation to give the Cy3BdT phosphoramidite monomer.

Purification of compounds **26** and **27** (with the dye in the cyclised Cy3B form) was problematic and several methods were tested including silica gel purification, alumina purification and precipitation. Purification by silica gel chromatography gave the best separation but the majority of the product remained stuck to the silica. Conversely, alumina chromatography allowed passage of the compound through the column but without giving any separation from impurities. Likewise, precipitation was not successful in removing all the impurities. The optimum system was found to be purification by silica column using methanolic ammonia in DCM as the eluent. In the case of compound **27**, the product still remained bound to the silica and could not be removed resulting in a lower yield than we would typically expect for a palladium cross-coupling reaction. For the phosphoramidite

**28**, the crude material was purified by precipitation only to avoid the chromatographic purification.




**Scheme 4.1.** Synthesis of Cy3BdT-phosphoramidite. Reagents and conditions: (i) Bromoethyl-1,3-dioxolane, KI, MeCN, reflux, 45 h, 29 %; (ii) bromoethyl-1,3-dioxolane, KI, MeCN, reflux, 45 h, 56 %; (iii) *N,N'*-diphenylformamidine, EtOH, triethylorthoformate, 97 °C, 16 h, 65 %; (iv) **22**, pyridine, $Ac_2O$, 50 °C, 24 h, 79 %; (v) $CHCl_3$, 50 % $H_2SO_4$, rt, 20 min, 96 %; (vi) 5-ethynyl-dU, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, rt, 1.5 h, 42 %; (vii) 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 45 min, 63 %.

### 4.2.2 Biophysical studies of oligonucleotides containing Cy3BdT

The Cy3BdT monomer was incorporated successfully into 3 oligonucleotide sequences (ODN51-53 table 4.1). Due to the hydrophobic nature of the Cy3B they were purified by polyacrylamide gel electrophoresis (PAGE, 12 % polyacrylamide). Subsequent sequences were purified using a RP-HPLC method modified for hydrophobic oligonucleotides (see

experimental section 6.2.1). As can be seen in figure 4.2 the purification was successful and the oligonucleotides were subsequently analysed by capillary electrophoresis and mass spectrometry. The oligonucleotides synthesised were HyBeacon probe format with either one or two Cy3BdT additions.

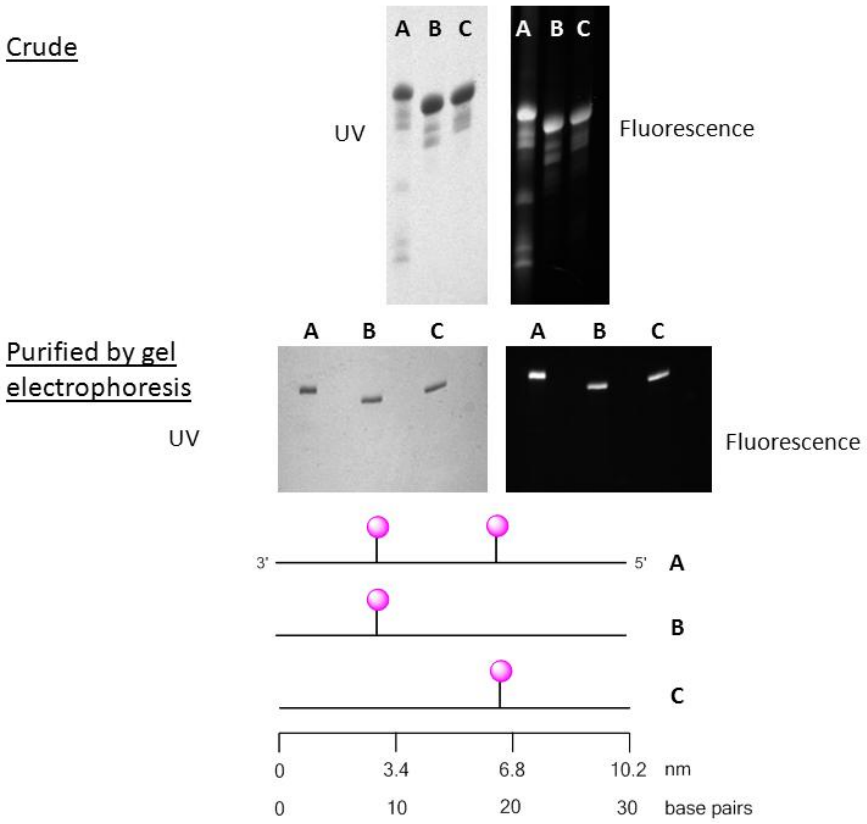



**Figure 4.2.** PAGE purification (12 % polyacrylamide) of oligonucleotides ODN51 (A), ODN52 (B), ODN53 (C). The gels were imaged using G:Box imaging system (Syngene) and Genesnap software (v7.08, Syngene). UV active bands were imaged with Epi shortwave UV darkroom lighting and shortwave bandpass filter. Fluorescent bands were imaged with transluminator darkroom lighting and EtBr/UV filter.

**Table 4.1.** Oligonucleotide sequences (mass data in appendix A.3). **1**= Cy3BdT; **2**= Cy3dT; **3**=5'Cy5; **4**=3'FAM; **F**= FAMdT (chemical structures in figure 4.3)

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN34 | ATG GCT ATT AAA AAC GGA GAA TGT ACG CAA |
| ODN36 | TTG CGT ACA **2**TC TCC GTT T**2**T AAT AGC CAT |
| ODN47 | **3**AT GGC TAT TAA AAA CGG AGA ATG TAC GCA A |
| ODN48 | **3**AT GGC TAT TAA AAA CGG AGA ATG TAC GCA A**4** |
| ODN51 | TTG CGT ACA **1**TC TCC GTT T**1**T AAT AGC CAT |
| ODN52 | TTG CGT ACA TTC TCC GTT T**1**T AAT AGC CAT |
| ODN53 | TTG CGT ACA **1**TC TCC GTT TTT AAT AGC CAT |
| ODN80 | TTT AGC ATT AGA TAC CTT TGC ATA CCT TA**F** T |
| ODN81 | ATA AGG TAT GCA AAG GTA TCT AAT GCT AAA |

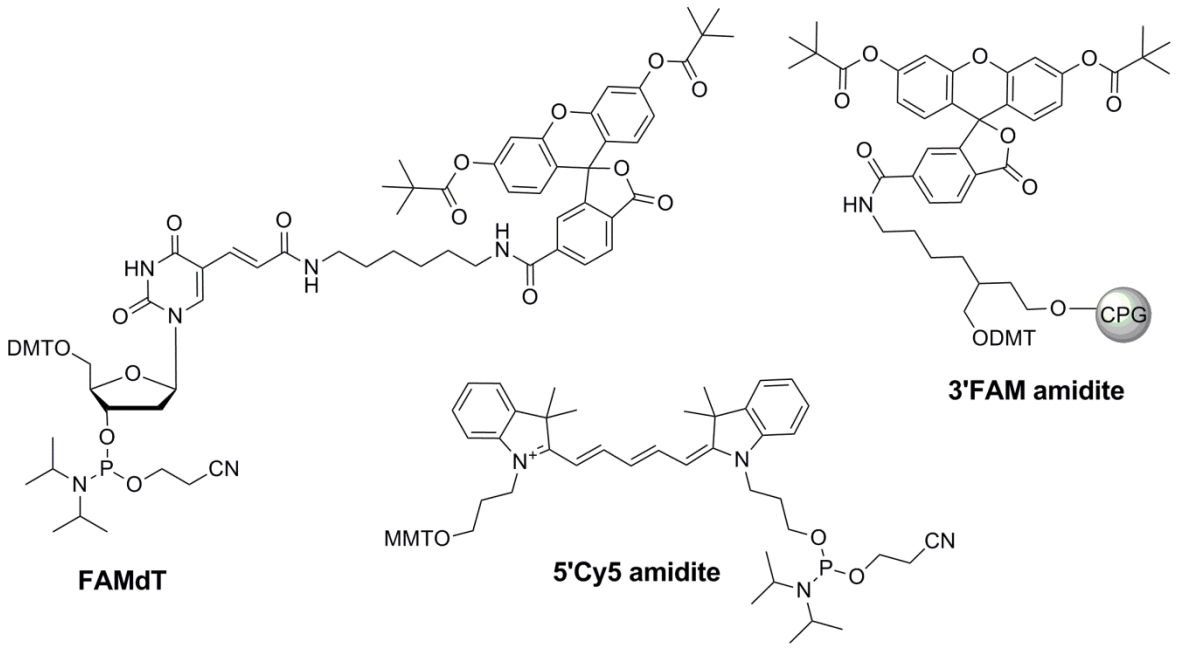



**Figure 4.3.** Chemical structures of FAMdT, 3'FAM amidite and 5'Cy5 amidite used in oligonucleotides ODN47, ODN48 and ODN80 (DMT= dimethoxytrityl protecting group and MMT= methoxytrityl protecting group).

The UV/visible absorbance and fluorescence properties of the oligonucleotides containing Cy3BdT were compared to an equivalent Cy3dT oligonucleotide (ODN36) and a FAMdT oligonucleotide (ODN80). The oligonucleotides containing Cy3BdT exhibited red-shifted absorbance and emission maxima in comparison to those containing Cy3dT (absorbance max red-shifted by 16 nm, emission max red-shifted by 4 nm). The fluorescence emission intensity of the oligonucleotides containing Cy3BdT was significantly higher than the corresponding oligonucleotides containing Cy3dT or FAM (figure 4.4).

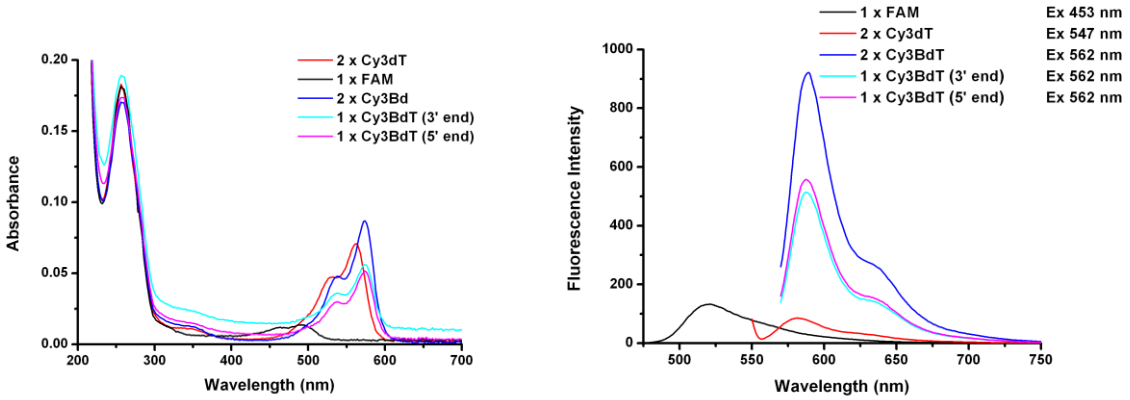



**Figure 4.4.** Absorbance and emission (excitation ¾ absorbance max) spectra of Cy3BdT duplexes in comparison to Cy3dT and FAM duplexes. (1xCy3dT= ODN36, 1xFAM= ODN80, 2x Cy3BdT= ODN51, 1x Cy3BdT= ODN52, 1x Cy3BdT=ODN53). All duplex samples were excited at ¾ absorbance max (453 nm for FAM; 547 nm for Cy3dT; 562 nm for CyBdT) to avoid excitation/emission overlap on the spectrum.

As previously discussed (section 3.3), oligonucleotides containing Cy3dT exhibit a biphasic melting curve. This is probably due to isomerisation of the Cy3 moiety in the major-groove of double-stranded DNA, which does not occur so readily in the single-stranded form as the dye is more constrained (and hence more fluorescent). Oligonucleotides containing Cy3BdT however exhibit melting curves typical of fluorescently labelled HyBeacon probes such as fluorescein HyBeacons (figure 4.5b). This is because the Cy3B cannot isomerise and so it is more fluorescent in double-stranded DNA than Cy3. As it is protruding into the major groove it cannot interact with the DNA nucleobases. In contrast, in the single-stranded form it is quenched by interactions with neighbouring bases. Consequently the room temperature fluorescence of the oligonucleotides containing Cy3BdT increases significantly from single- to double-stranded form, whereas the fluorescence of the Cy3dT analogue decreases (figure 4.5a). For HyBeacon probes to be useful in PCR-based applications, the fluorescence of the single-stranded probe should be far lower than that of the double-stranded form to avoid the background fluorescence drowning out the RT-PCR signal. The preliminary study described above shows that Cy3B is likely to be a suitable fluorophore for use in HyBeacons.

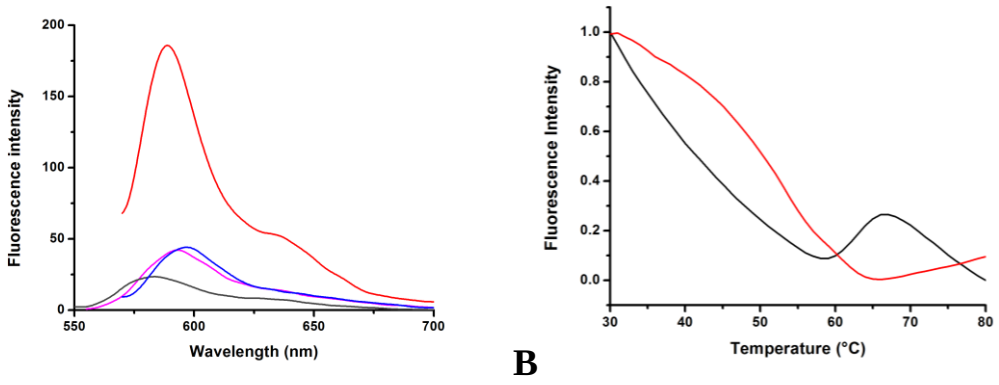



**Figure 4.5.** Comparison of the fluorescence emission of Cy3dT and Cy3BdT HyBeacon oligonucleotides; (A) room temperature emission spectra (Cy3dT: ODN36 ss (pink), ds (black), Cy3BdT: ODN51 ss (blue), ODN51 ds (red)); (B) normalised fluorescence melting profiles of duplexes (Cy3dT: ODN36 ds (black), Cy3BdT: ODN51 ds (red)). Cy3BdT HyBeacon probes show greater fluorescence in double-stranded form than in single-stranded form but the reverse is seen for Cy3dT HyBeacon probes. Cy3BdT HyBeacon probes exhibit melting curves typical of fluorescently labelled HyBeacon probes such as fluorescein HyBeacons.

The temperature dependent fluorescence of single-stranded Cy3BdT oligonucleotides with one or two additions of the dye was shown to be insignificant (figure 4.6a). The small increase in fluorescence as the temperature increases is likely to be the result of the unfolding of the single-strand (which partially quenches the dye at room temperature). ODN51 (with two Cy3BdT additions), exhibited highest double-stranded fluorescence and lowest single-stranded fluorescence and so is the best HyBeacon probe (figure 4.6b). This result is expected based on the mechanism of HyBeacon probes (discussed in section 1.5.2.4); the two dyes in ODN51 quench each other in the folded single-strand but are held apart (and are therefore fluorescent) in the double-stranded form. The two dye addition, having twice as many fluorophores than either of the single-addition examples (ODN52 or ODN53), naturally gives higher double-stranded fluorescence. The mechanism of dye-quenching was confirmed by studying the UV absorbance spectrum (figure 4.6c) which showed, in the single stranded form, ODN51 exhibits an excimer-like spectrum with an increased shoulder peak at ~530 nm which is indicative of CyDye stacking. This phenomenon is discussed in detail in section 3.6.

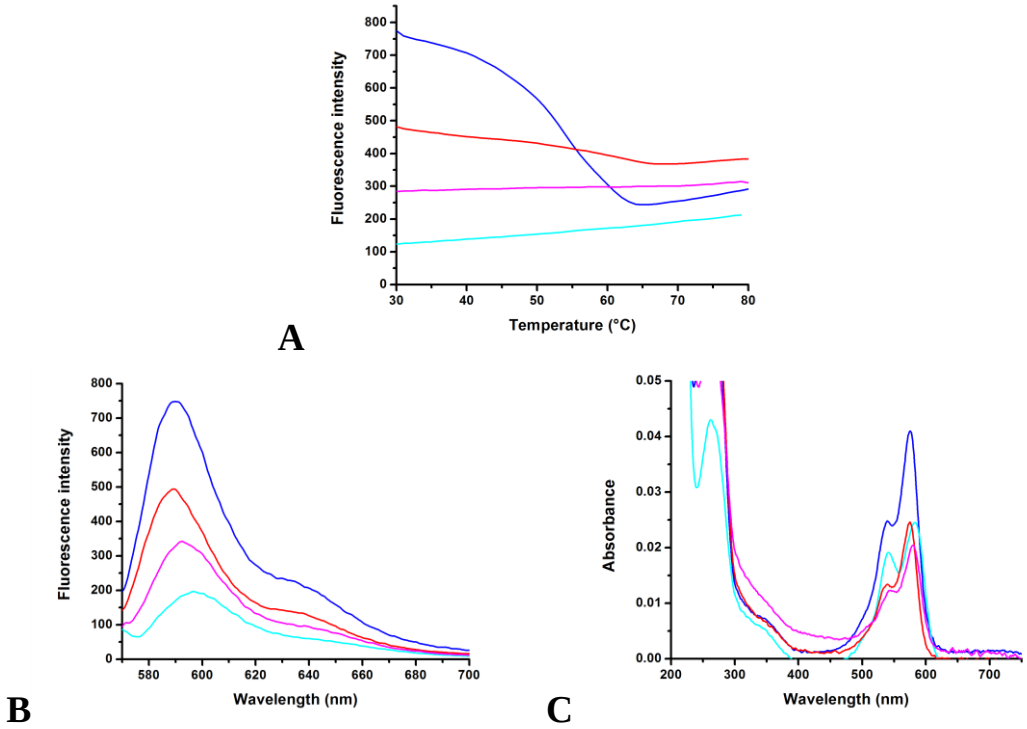



**Figure 4.6.** Cy3BdT oligonucleotides (with 1 or 2 Cy3BdT additions) (A) fluorescence melting profiles, (B) fluorescence emission spectra, (C) UV absorbance spectra; ODN51 ds (2 Cy3BdT, blue), ODN51 ss (cyan), ODN53 ds (1 Cy3BdT, red), ODN53 ss (pink). Fluorescence melting of single-stranded oligonucleotides containing Cy3BdT show very little temperature dependent fluorescence, whereas the duplex samples exhibit typical HyBeacon probe melting profiles.

The use of Cy3BdT in a FRET system was compared with the equivalent system with Cy3dT (figure 4.7). The system under investigation contained FAM and Cy5 in one strand (one dye at each end of the oligonucleotide) and two Cy3BdT dyes in the complementary strand. By exciting the system at 490 nm (the excitation of FAM), fluorescence is transferred successfully through Cy3B to Cy5. In the control system (with no Cy3BdT dyes) the emission spectrum is dominated by the emission of FAM, with very little direct excitation of Cy5, proving that Cy3B is a key component in the energy transfer system. In comparison, the corresponding Cy3dT example shows less efficient energy transfer to the Cy5, probably due to the lower quantum yield of the Cy3dT (figure 4.8). The examples with a single Cy3BdT addition show less efficient energy transfer than the dual addition example, suggesting that there is energy transfer between the Cy3B dyes themselves.

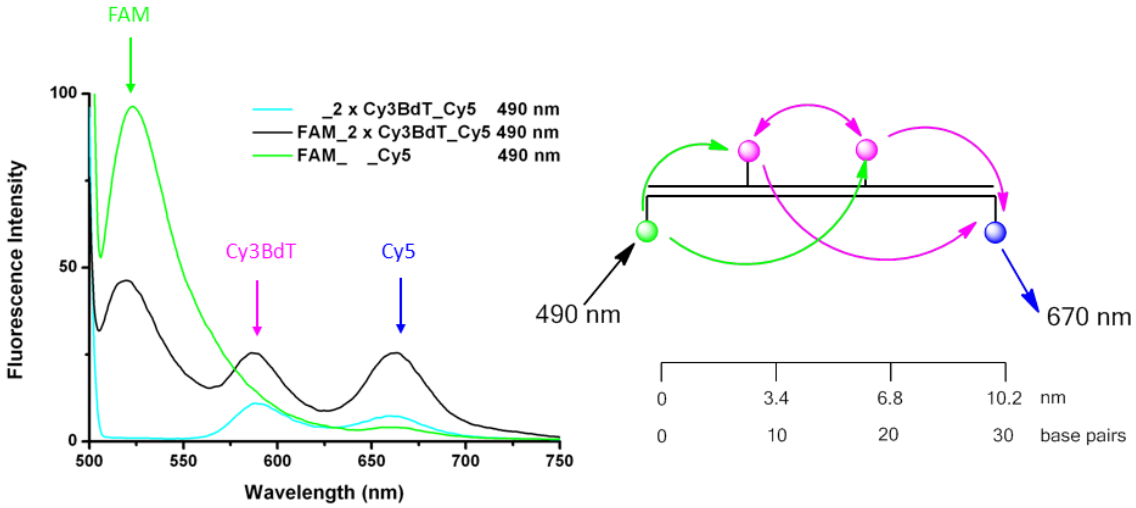



**Figure 4.7.** Fluorescence emission of the FRET system excited at 490 nm. Schematic representation of the FRET system. The fluorescein moiety is excited at 490 nm and fluorescence is transferred successfully through Cy3B to Cy5 (which can be seen to emit at 670 nm).

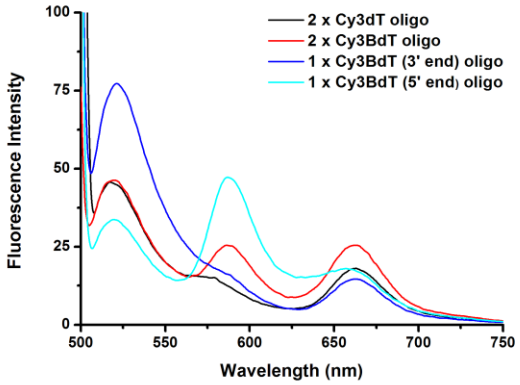
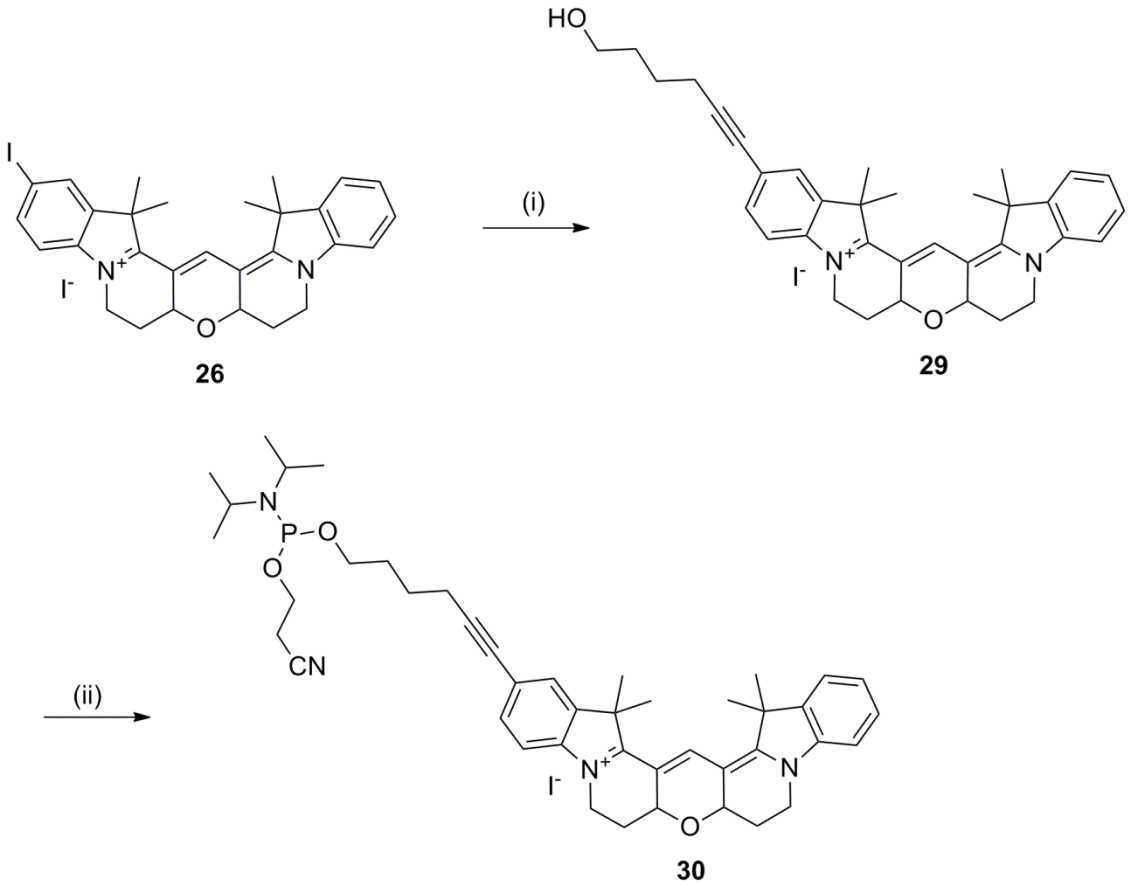



**Figure 4.8.** Comparison of Cy3BdT and Cy3dT in the FRET system; Cy3dT (2 dye example ODN36), Cy3BdT (2 dye example ODN51), Cy3BdT (1 dye examples ODN52 and ODN53). Energy is transferred most successfully in the example with two Cy3BdT dyes. The energy transfer is lower for the example with two Cy3dT dyes probably because of the lower quantum yield of Cy3dT.

### 4.2.3 The synthesis of a 5'Cy3B phosphoramidite and biophysical studies of oligonucleotides containing 5'Cy3B

To complement the Cy3BdT monomer a Cy3B phosphoramidite was designed for addition to the 5'-end of oligonucleotides (5-hexyn-1-ol-6-Cy3B). It was made by a similar synthetic approach as was developed for the 5'Cy3 phosphoramidite. 5-Iodo-Cy3B was conjugated to 5-hexyn-1-ol by a palladium cross-coupling reaction, followed by phosphitylation to give the 5'Cy3B phosphoramidite monomer (scheme 4.2). The flexible linkage between the dye and DNA (in contrast to the rigid CydT monomers) allows stacking of the dye on the end of the duplex making it suitable for applications such as Molecular Beacon probes where it may interact with a quencher in the complementary strand.




**Scheme 4.2.** Synthesis of 5-hexyn-1-ol-6-Cy3B phosphoramidite. Reagents and conditions: (i) 5-hexyn-1-ol, CuI, $Pd(PPh_3)_4$, DMF, $Et_3N$, rt, 5 h, 82 %; (ii) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 45 min, 93 %.

The temperature dependent fluorescence of the oligonucleotides containing 5'Cy3B (table 4.2) was evaluated by fluorescence melting. No significant temperature dependence was exhibited upon melting for either the single-stranded or double-stranded forms proving that Cy3B is invulnerable to the isomerisation typically exhibited by the other CyDyes at high

temperatures (figure 4.9). The oligonucleotide containing 5'Cy3B was annealed to both short and long complementary oligonucleotides to assess whether there was any difference in the fluorescence melting. It was thought in the case of the short complement that the Cy3B would be able to stack on the end of the duplex whereas in the overhanging complement the Cy3B may stack in either the major or minor groove of the duplex. Both however exhibited the same fluorescence melting properties suggesting the fluorescence properties of the Cy3B are not affected by the surrounding DNA structure/environment. The oligonucleotide was used subsequently as a 5'labelled PCR primer, discussed later in section 4.3.5.

**Table 4.2.** Oligonucleotide sequences (mass data in appendix A.3). **8**= 5-hexyn-1-ol-6-Cy3B phosphoramidite

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN45 | TCT CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GTA AGA AAC TAT GTG AAA ACT TTT TGA |
| ODN66 | GCA TAA TCC AGG AAA ACT GA |
| ODN67 | **8**TC AGT TTT CCT GGA TTA TGC |

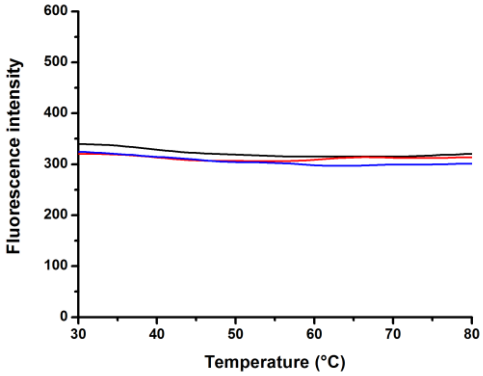
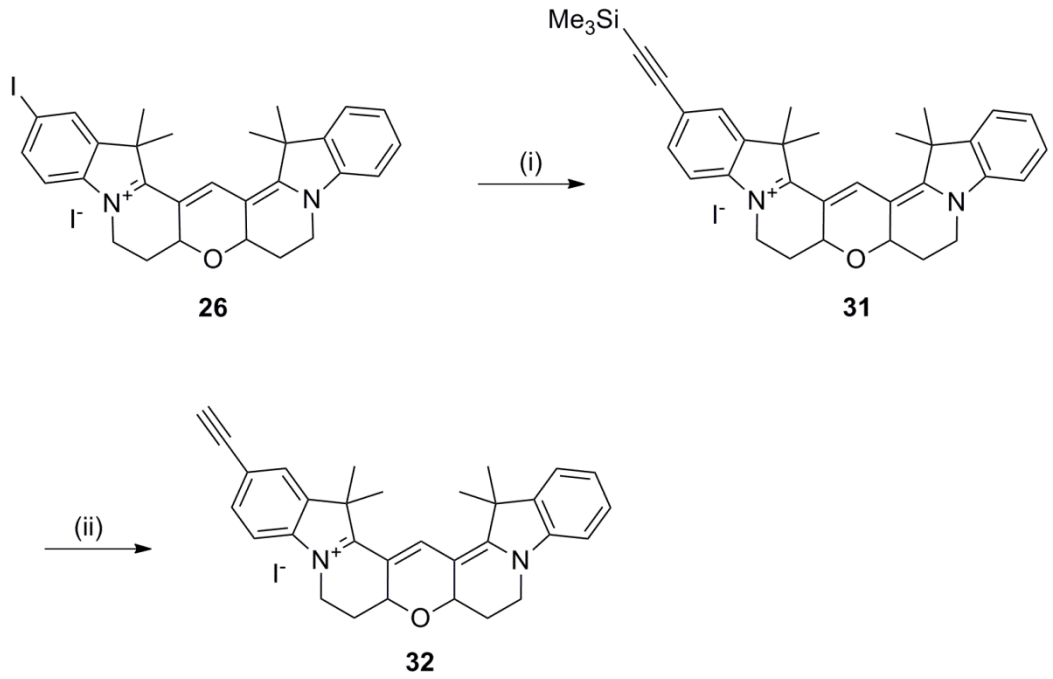



**Figure 4.9.** Fluorescence melting profile of an oligonucleotide containing 5'Cy3B; ODN67 ss (black), duplex with short complementary oligonucleotide (red), duplex with long complementary oligonucleotide (blue). Virtually no temperature dependent fluorescence can be seen proving that Cy3B is invulnerable to the isomerisation typically exhibited by the other CyDyes at high temperatures.

### 4.2.4 Labelling of oligonucleotides with Cy3B using click-chemistry

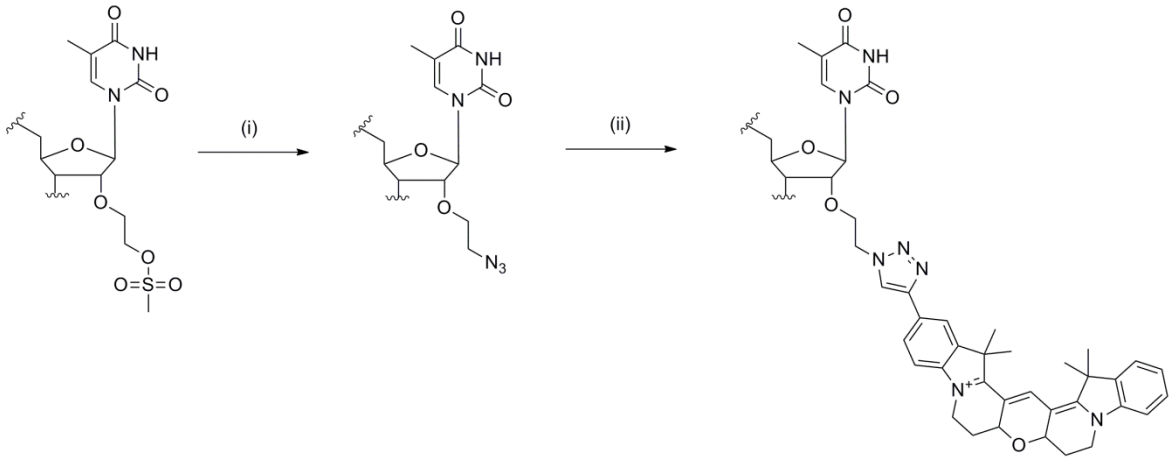
'Click-chemistry' is the term used to describe highly efficient and selective reactions and is commonly applied to the copper-catalysed 1,3-dipolar cycloaddition reaction between a terminal alkyne and an azide (CuAAC).[149,150] It has been used for the interstrand crosslinking of DNA in our research group[99] and also for the postsynthetic labelling of alkyne-modified oligonucleotides with azide-modified fluorophores by Carell[118] and others.[151] Click-chemistry is an important method to attach fluorophores to DNA and, in the work of Marta Gerowska (in our research group), a series of alkyne-modified CyDyes were added to azide-modified HyBeacon probes.[152]

The literature has shown that HyBeacon probes work successfully with the dye moiety in either the major or the minor groove.[79,78,80] With this in mind, and in parallel with the work described above, 5-ethynyl-Cy3B was synthesised for the click-chemistry labelling of oligonucleotides in either the minor groove or major groove (scheme 4.3). The synthesis was achieved using the palladium cross-coupling reaction between 5-iodo-Cy3B and trimethylsilylacetylene, followed by deprotection of the trimethylsilyl group with tetrabutylammonium fluoride. The absorbance and emission maxima of the 5-ethynyl-Cy3B compound were red-shifted relative to those of the commercial Cy3B NHS-ester (the absorbance by 8 nm and the emission by 14 nm).

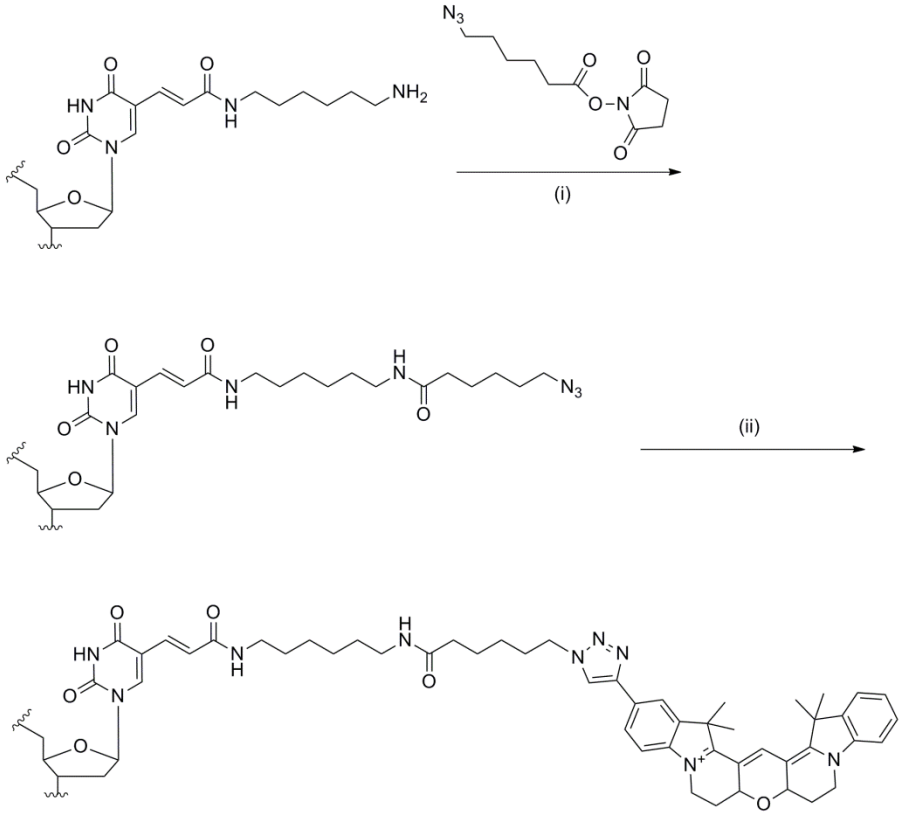



**Scheme 4.3.** Synthesis of 5-ethynyl-Cy3B. Reagents and conditions: (i) trimethylsilylacetylene, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, 16 h, rt, 71 %; (ii) TBAF, THF, 5 min, rt, 75 %.

To demonstrate the labelling of oligonucleotides with Cy3B in the minor groove by click-chemistry, three HyBeacon oligonucleotides were synthesised containing two 2'mesyloxyethyl rT monomers at 3, 5 or 7 nucleotide spacing (ODN54, ODN55, ODN56 respectively, table 4.3). The oligonucleotides must be synthesised containing a mesyl group which is then converted to azide on the solid support because of the incompatibility of the azide functionality with P(III) chemistry (the Staudinger reaction).[153] The 2'mesyloxyethyl rT monomer was synthesised by Montse Shelbourne in our research group. The probe sequence was designed to probe the R516G mutation of the cystic fibrosis trans conductance regulatory (CFTR) gene and was based on previous work in our research group.[154] An on-resin conversion of the mesyl group to azide was undertaken followed by cleavage of the oligonucleotide from the resin (scheme 4.4). The solution-phase click-chemistry reaction between the oligonucleotide and 5-ethynyl-Cy3B afforded the dual-labelled oligonucleotide. By labelling at the 2'-position of the sugar in the thymidine residues, the dye is positioned in the minor groove of the DNA helix when the probe hybridises to its complementary oligonucleotide. To investigate other labelling positions on the nucleoside, corresponding sequences were synthesised with amino-C6-dT phosphoramidite in the same positions in the oligonucleotides as the 2'mesyloxyethyl rT residues (scheme 4.5). In solution phase, the amino group was coupled to an azide active-ester, followed by a click-chemistry reaction with 5-ethynyl-Cy3B. This resulted in a triazole linkage between the oligonucleotide and 5-ethynyl-Cy3B, giving the dual-labelled oligonucleotides (ODN57, ODN58, ODN59, table 4.3). By labelling at the C5-position of the thymine bases, the dye is located in the major groove of the DNA helix when the probe is hybridised to its complementary strand. Previous work in our research group had utilised the same azide functionality in the major groove for interstrand DNA crosslinking.[99]



**Scheme 4.4.** Click-chemistry labelling reaction of Cy3B at the 2'-sugar position of thymidine in an oligonucleotide. Conditions; (i) sodium azide, 18-crown-6, DMF, 65 °C, 20 h, then $NH_3$, rt, 16 h; (ii) 0.2 M NaCl, tris-hydroxypropyl ligand, sodium ascorbate, $CuSO_4.5H_2O$, 5-ethynyl-Cy3B, DMF, rt, 2 h.



**Scheme 4.5.** Click-chemistry labelling reaction of Cy3B at the C5 position of thymidine in an oligonucleotide. Conditions; (i) 0.5 M $Na_2CO_3/NaHCO_3$, azide active ester, DMSO, rt, 2 h; (ii) 0.2 M NaCl, tris-hydroxypropyl ligand, sodium ascorbate, $CuSO_4.5H_2O$, 5-ethynyl-Cy3B, DMF, 55 °C, 2 h.

### 4.2.5 Comparison of HyBeacon probes containing Cy3BdT phosphoramidite with those labelled with Cy3B by click-chemistry

To compare Cy3BdT and the click-chemistry labelled HyBeacons (ODN54-56 and ODN57-59) the same sequences were synthesised containing Cy3BdT (table 4.4; ODN62, ODN63, ODN64). Room temperature fluorescence emission scans (figures 4.10) were compared for all the HyBeacon probes and show that for all three addition-types the 7 base spacing between the dyes gives the best single-strand quenching compared to double-strand (wild-type) fluorescence (summarised in table 4.3). Quenching may occur between the two dyes in the closer-spaced duplex examples (3 and 5 nucleotide separation) as the duplex fluorescence is generally lower in these cases.

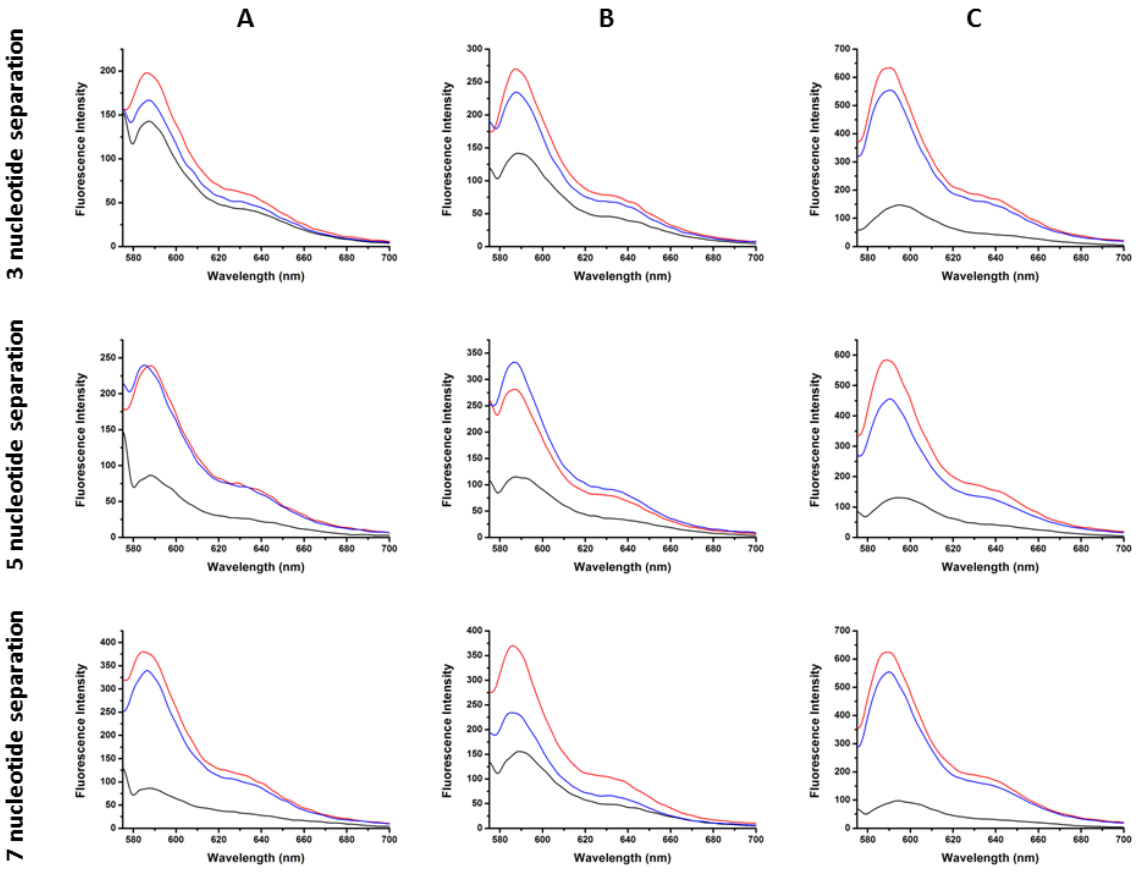



**Figure 4.10.** Comparison of room-temperature fluorescence emission scans for HyBeacon probes containing Cy3BdT and those labelled with Cy3B by click-chemistry. (A) (ODN54-56; triazole-Cy3B at 2'sugar position; 3,5,7 bp spacing); (B) (ODN57-59; triazole-Cy3B at C5 position of thymidine; 3,5,7 bp spacing); (C) (ODN62-64 Cy3BdT addition 3,5,7 spacing); ss (black), wt (red), mt (blue). The HyBeacon probes containing Cy3BdT are most fluorescent, followed by those labelled with Cy3B in the major groove and then those labelled with Cy3B in the minor groove. The examples where the dyes are separated by 7 nucleotides give the best wt-duplex to single-strand fluorescence intensity ratio.

**Table 4.3.** Summary of room-temperature fluorescence emission (at emission maxima) for HyBeacon probes containing Cy3BdT and those labelled with Cy3B by click-chemistry.

| Probe | Fluorescence intensity at emission maxima | | | Quenching efficiency % (wt duplex:ss) |
|---|---|---|---|---|
| | wt | mt | ss | |
| ODN54 | 198 | 167 | 143 | 28 |
| ODN55 | 240 | 240 | 87 | 64 |
| ODN56 | 380 | 340 | 87 | 77 |
| ODN57 | 269 | 235 | 141 | 48 |
| ODN58 | 281 | 334 | 118 | 58 |
| ODN59 | 375 | 235 | 155 | 59 |
| ODN62 | 632 | 553 | 148 | 77 |
| ODN63 | 583 | 455 | 131 | 78 |
| ODN64 | 625 | 555 | 98 | 84 |

**Table 4.4.** Oligonucleotide sequences (mass data in appendix A.3). **1**= Cy3BdT; **5**= minor groove-triazole-Cy3B; **6**= propanol; **7**= major groove-triazole-Cy3B.

| Oligonucleotide code | Sequence |
|---|---|
| ODN49 | CTA TGA TGA ATA TAG ATA CAG AAG CGT CAT |
| ODN50 | CTA TGA TGA ATA TGG ATA CAG AAG CGT CAT |
| ODN54 | CGC TTC **5**GT A**5**C TAT ATT CAT C**6** |
| ODN55 | CGC TTC **5**GT ATC **5**AT ATT CAT C**6** |
| ODN56 | CGC TTC **5**GT ATC TA**5** ATT CAT C**6** |
| ODN57 | CGC TTC **7**GT A**7**C TAT ATT CAT C**6** |
| ODN58 | CGC TTC **7**GT ATC **7**AT ATT CAT C**6** |
| ODN59 | CGC TTC **7**GT ATC TA**7** ATT CAT C**6** |
| ODN62 | CGC TTC **1**GT A**1**C TAT ATT CAT C**6** |
| ODN63 | CGC TTC **1**GT ATC **1**AT ATT CAT C**6** |
| ODN64 | CGC TTC **1**GT ATC TA**1** ATT CAT C**6** |
| ODN70 | CGC TTC TGT ATC TAT ATT CAT C |

By comparing all the HyBeacon probes in duplex-form under the same conditions (figure 4.11) it was found that those containing Cy3BdT have the highest overall fluorescence. This is probably because the dye in Cy3BdT is held rigidly away from the DNA helix, preventing any quenching interactions, whereas the dyes in the probes labelled in the minor-groove by click-chemistry are stacked closely against the helix where they will be partially quenched. In the examples where the dyes are labelled in the major-groove by

click-chemistry, the linker is sufficiently long and flexible to allow the dyes to bend back and stack onto the helix, where they will be quenched by the neighbouring DNA bases. The fluorescence melting of the HyBeacons containing Cy3BdT (wild-type versus mutant-type) was compared (figure 4.12). Although ODN64 gave the best single strand to double strand fluorescence ratio at room temperature, the wild-type/mutant-type discrimination was not quite as distinct, so it was decided to pursue further RT-PCR studies with ODN63 (the Cy3BdT HyBeacon with a 5 nucleotide spacing).

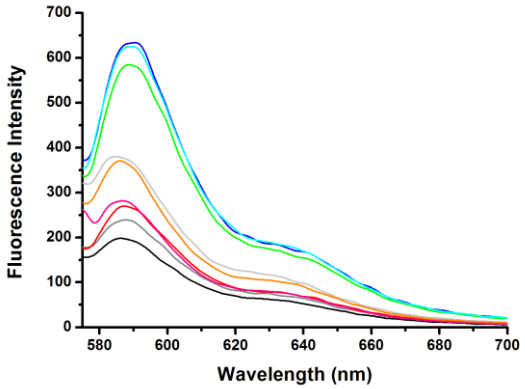



**Figure 4.11.** Room temperature fluorescence emission spectra of wild-type HyBeacon duplex samples; ODN54 (black), ODN55 (dark grey), ODN56 (light grey), ODN57 (red), ODN58 (pink), ODN59 (orange), ODN62 (blue), ODN63 (green), ODN64 (cyan). It can be seen that the probes containing Cy3BdT exhibit the highest fluorescence intensity overall.

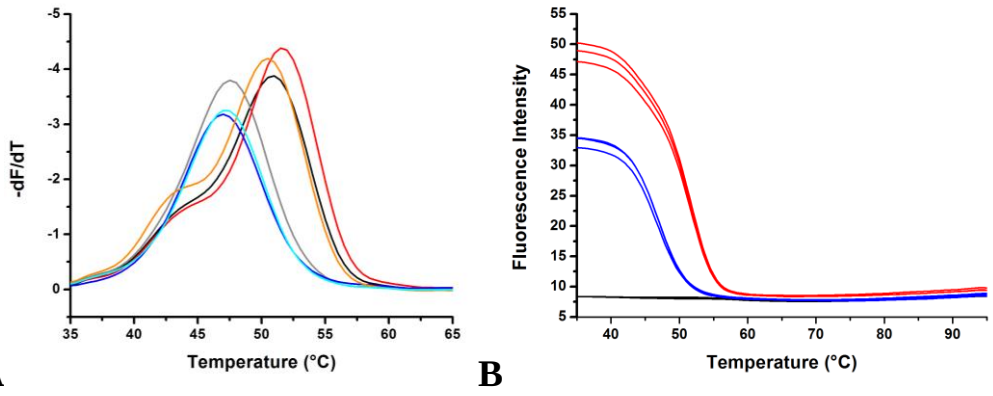



**Figure 4.12.** Comparison of Hybeacons containing Cy3BdT; (A) fluorescence melting derivatives of HyBeacons containing Cy3BdT, ODN62 wt (black), ODN62 mt (grey), ODN63 wt (red), ODN63 mt (blue), ODN64 wt (orange), ODN64 mt (cyan); (B) ODN63 fluorescence melting curves, wt (red), mt (blue), ss (black).

The thermal stability of the HyBeacon probes were compared in a UV melting experiment in which the duplexes containing Cy3B were compared to an unmodified duplex

(table 4.5). For each addition-type the 5-nucleotide separation example was evaluated. ODN56, a 7-nucleotide separation example, was also tested to ensure no significant change resulted from the nucleotide spacing.

The results show that the HyBeacons labelled with two Cy3B dyes by click-chemistry are destabilised by around 4.0 °C, with the example with the dyes added in the minor-groove being destabilised by 2.1 °C per addition. The example where the dyes are added into the major-groove is destabilised less (-1.8 °C per addition) which may be due to the flexible linker allowing the dye to intercalate into the DNA. The Cy3BdT HyBeacon was destabilised by 6.8 °C (-3.4 °C per addition) which is in the same range as that seen previously for the Cy3dT and Cy5dT examples (-2.4 ºC for Cy3dT, -3.1 ºC for Cy5dT). The rigidity of the dyes in the major groove may exert some strain on the DNA structure.

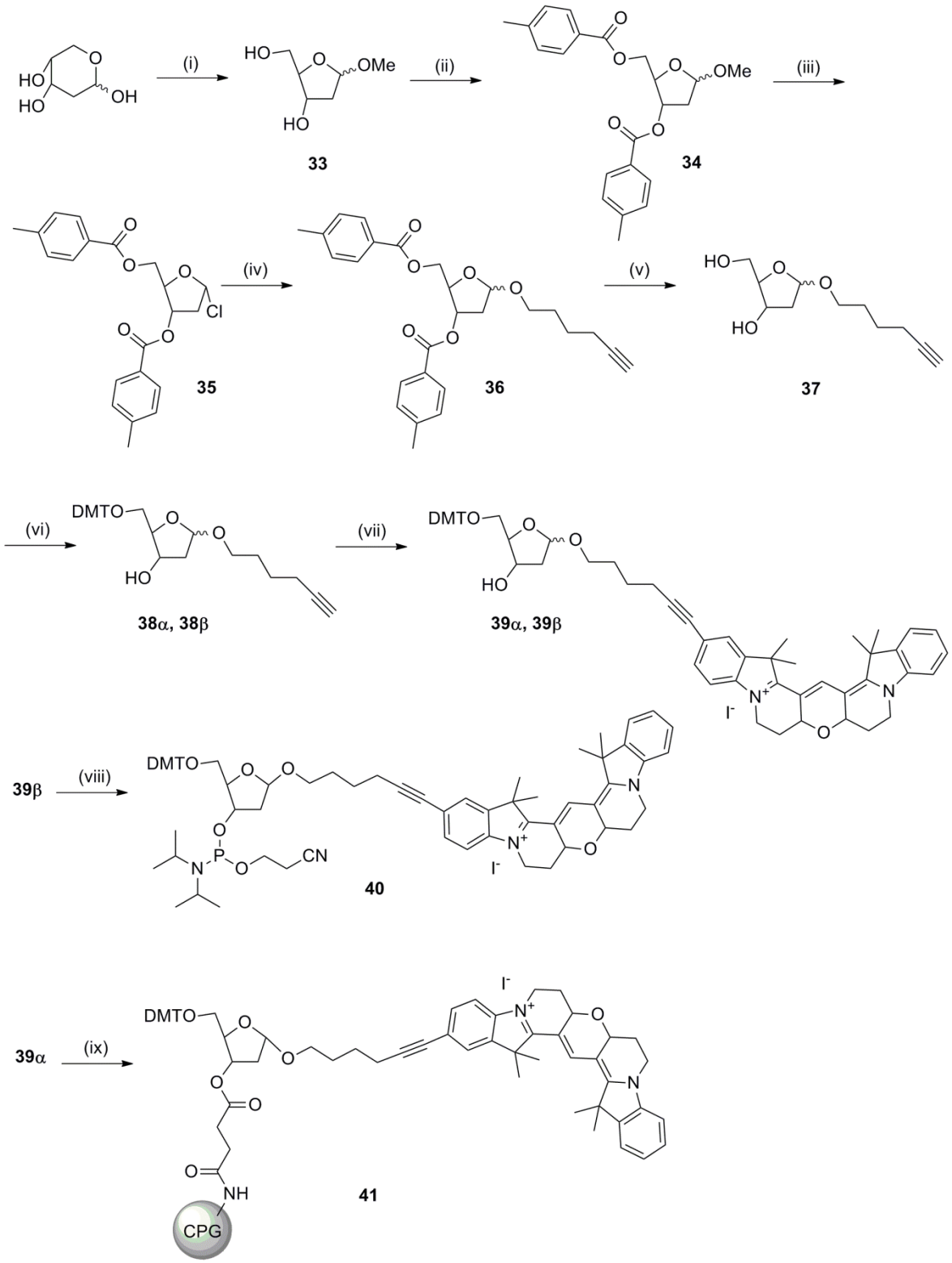
**Table 4.5.** $T_m$ data from UV melting for duplexes modified with Cy3B (all with complementary oligonucleotide ODN49) [a]

| Oligonucleotide code | Dye addition-type | $T_m$/ °C | $\Delta T_m$/ °C | $\Delta T_m$ per addition/ °C |
|---|---|---|---|---|
| ODN70 | Unmodified | 57.2 | --- | --- |
| ODN55 | Cy3B minor groove (5 bp spacing) | 53.1 | -4.1 | -2.1 |
| ODN56 | Cy3B minor groove (7 bp spacing) | 52.9 | -4.3 | -2.2 |
| ODN58 | Cy3B major groove (5 bp spacing) | 53.6 | -3.6 | -1.8 |
| ODN63 | Cy3BdT | 50.4 | -6.8 | -3.4 |

[a]Average $T_m$ value was calculated from three successive melting curves. Standard deviation for all $T_m$ data is ± 0.1 ºC.

### 4.2.6 The synthesis of Cy3B-deoxyribose phosphoramidite and Cy3B-deoxyribose resin

To complete the Cy3B DNA 'molecular toolkit' a synthetic scheme was devised for the synthesis of Cy3BdR (deoxyribose). This was designed to give two anomers; the α-anomer was to be coupled to an oligonucleotide synthesis resin for 3'-addition to oligonucleotides and the β-anomer was made for phosphitylation to provide monomers for either 5' or internal addition (scheme 4.6). The theory was that the Cy3BdR should be less hydrophobic and therefore easier to purify than the 5-hexyn-1-ol-6-Cy3B monomer, giving a convenient alternative for 5' oligonucleotide addition.




**Scheme 4.6.** Synthesis of Cy3BdR-phosphoramidite and resin. Conditions; (i) MeOH, 1 % methanolic HCl, rt, 30 min; (ii) pyridine, *p*-toluoyl-chloride, 0 °C then rt, 16 h; (iii) acetic acid, sat. HCl in acetic acid, acetyl chloride, 0 °C then rt, 64 % over 3 steps; (iv) THF, DCM, DMAP, 5-hexyn-1-ol, rt, 3.5 h, 53 %; (v) methanolic ammonia, 45 °C, 16 h, 82 %; (vi) DMTCl, pyridine, DMAP, 1 h, rt, α 47 %, β 41 % (total 88 %); (vii α) **26**, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, 52 h, rt, 66 %; (vii β) **26**, CuI, $Pd(PPh_3)_4$, $Et_3N$, DMF, 52 h, rt, 73 %; (viii) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, rt, 45 min, 94 %; (ix) amino-link resin, *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride, DMAP, $Et_3N$.

The synthetic route to compound **35** progressed according to the literature.[155] 5-Hexyn-1-ol was added to compound **35** at the 1' position followed by deprotection of the toluoyl groups. Subsequent DMT protection at the 5'OH position yielded a key intermediate for the synthesis of a variety of possible monomers. At this stage the α and β-anomers were separated. 5-Iodo-Cy3B was coupled, by palladium cross-coupling reaction to the terminal alkyne at the 1' position of **38α** and **38β**. The α-anomer was coupled to an amino-link resin whilst the β-anomer was phosphitylated to give the Cy3BdR phosphoramidite monomer. In general the α-anomer is used for resin-modification as it gives less steric hindrance to the incoming monomer during solid-phase synthesis.[156] The β-anomer is phosphitylated as the resulting amidite should be more reactive than the α-amidite because the side chain (and dye in this case) is further away from the phosphoramidite moiety.[157]

Oligonucleotides containing each monomer were synthesised to evaluate coupling efficiency and purity (see appendix section A3). The 3'Cy3BdR resin was used in a sequence designed for use in a surface enhanced Raman spectroscopy (SERS) experiment by a collaborator (table 4.6). The Cy3BdR phosphoramidite monomer was incorporated as an internal addition in a test sequence (ODN79). Both the 3'Cy3BdR and the Cy3BdR phosphoramidite were then used in the synthesis of PCR probe sequences (sections 4.3.3-4.3.4).

**Table 4.6.** Oligonucleotide sequences (mass data in appendix A.3). **R**= Cy3BdR resin; **S**= Cy3BdR phosphoramidite

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN75 | AAT ATC ATC TTT GGT GTT TCC T**R** |
| ODN79 | TTT T**S**T TT |

### 4.2.7 A discussion of the fluorescence properties of the Cy3B monomers

A comparison of the absorbance and fluorescence properties of the synthesised compounds shows that despite its lower extinction coefficient in comparison to Cy5, Cy3BdR has a considerably higher quantum yield than both the Cy3dT and Cy5dT monomers, so it exhibits higher fluorescence overall (table 4.7). This shows that a high fluorescence output relies on a combination of both high extinction coefficient and high quantum yield. Whilst

Cy3 and Cy5 seem to be significantly affected by the conjugation of additional groups (the dT monomers give a lower extinction coefficient and quantum yield than the corresponding commercial NHS-esters), the Cy3B variations show less fluctuation in the extinction coefficient.

Parallel work by Marta Gerowska shows that addition of a 5-ethynyl group to Cy3 or Cy5 causes a significant reduction in extinction coefficient compared to the commercial NHS-ester.[152] This is mirrored in the 5-ethynyl-Cy3B version but to a lesser extent. The quantum yield values obtained for Cy3B agree with the literature values ranging from 0.67-0.85.[61,64]

**Table 4.7.** Photophysical properties; UV-Vis absorbance maxima ($A_{max}$), fluorescence emission maxima ($Em_{max}$), extinction coefficient (**ε**) and quantum yield (**Φ**). (recorded in MeOH).

| Compound | $A_{max}$ (nm) | ε ($M^{-1}cm^{-1}$) | $Em_{max}$ (nm) | Φ |
|---|---|---|---|---|
| Cy3-NHS ester* | 550 | 150,000 | 570 | >0.15 |
| Cy5-NHS ester* | 649 | 250,000 | 670 | >0.28 |
| Cy3B-NHS ester* | 559 | 130,000 | 570 | >0.70 |
| ICy3 | 551 | 114,000 | 567 | 0.04 |
| Cy3dT | 565 | 112,000 | 591 | 0.05 |
| Cy5dT | 660 | 208,000 | 685 | 0.06 |
| ICy3B | 562 | 141,000 | 575 | 0.96 |
| 5-ethynyl-Cy3B | 566 | 118,000 | 579 | 0.88 |
| Cy3BdT | 576 | 122,000 | 596 | 0.50 |
| Cy3BdR | 571 | 130,000 | 585 | 0.80 |

*Data from GE Healthcare.

## 4.3 Demonstration of Cy3B monomers in PCR probes

To demonstrate some applications of the Cy3B monomers, several oligonucleotide sequences were synthesised for use as PCR probes. Examples of the commonly used probe types were made, incorporating the different Cy3B additions into a Taqman probe, a HyBeacon probe, a Molecular Beacon probe, a Scorpion primer and a 5'labelled primer.

Synthetic template DNA (wild-type and mutant-type) and primer sequences for the CFTR R516G mutation site (table 4.8) were previously prepared and in use by Dr James Richardson in our research group.

Initial optimisation of template concentration was undertaken to attain threshold cycle ($C_T$) values that were in accordance with the literature values.[79,78,75,76,69] A PCR assay was set up with template (ODN45) concentrations ranging from 0.07 fg up to 100 pg per 20 µL reaction. Precision-HRM polymerase mastermix (Primer design) was used and the reactions were monitored by the addition of SYBR green intercalating dye. A $C_T$ value of ~20 cycles, in line with the majority of work in the area, was achieved with a template concentration of 1 pg of target DNA per 20 µL reaction.

As the experiments were aimed at demonstrating that Cy3B works well in a variety of probes, all probes were synthesised to be complementary to the wild-type sequence only. Additional probes could, if required, be synthesised to correspond to mutant-type sequences. If the experiment was to be aimed at genotyping, the additional probes could be added to the PCR assays, as demonstrated by Piatek *et al* in their study of Molecular Beacons for detecting drug resistance in Mycobacterium tuberculosis.[69]

**Table 4.8.** PCR templates (ODN45 WT, ODN46 MT mutation site underlined) and primers (ODN60 forward, ODN61 reverse). (mass data in appendix A.3)

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN45 | TCT CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GTA AGA AAC TAT GTG AAA ACT TTT TGA |
| ODN46 | TCT CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT GGA TAC AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GTA AGA AAC TAT GTG AAA ACT TTT TGA |
| ODN60 | CAG TTT TCC TGG ATT ATG CC |
| ODN61 | CAA AAA GTT TTC ACA TAG TTT CTT |

### 4.3.1 Demonstration of a Taqman probe containing 5'Cy3B phosphoramidite

A Taqman probe with 5'Cy3B and internal BHQ2 was synthesised, ODN72 (table 4.9). BHQ2 was chosen as the FRET quencher for Cy3B due to its extensive absorbance at the emission wavelength of Cy3B (BHQ2 abs max 580 nm). The fluorophore is traditionally located at the 5' end of a Taqman probe so that it will be cleaved first by the DNA

polymerase.[67] In a typical Taqman probe the quencher would be at the 3'end of a probe of ~20 bases in length, but due to the locus of the probe in the CFTR sequence, the probe had to be longer (to generate a suitable $T_m$) and so the quencher was added at a central position in the probe. An alternative forward primer was also synthesised to place it in closer proximity to the 5' end of the Taqman probe (ODN74).

Symmetric RT-PCR was undertaken using Precision-HRM mastermix which contains a thermostable polymerase with the 5'-3' exonuclease activity necessary for Taqman probe cleavage. As seen in figure 4.13 successful PCR was achieved which gave discrimination between the wild-type and mutant-type synthetic template. The sample with wild-type template has a $C_T$ value of 16 which shows the probe cleavage during the extension step is very efficient. The sample with a template mutation has a $C_T$ value of 19 because in this case the probe-template duplex is unstable and so the probe is cleaved at a far slower rate. The ratio of final fluorescence intensity (~6.5:1 wt:mt) is also a clear indicator of wt/mt discrimination and proves less probe is cleaved overall in the sample with the mutated template.

**Table 4.9.** Oligonucleotide sequences (mass data in appendix A.3). **6**= propanol; **8**= 5-hexyn-1-ol-6-Cy3B phosphoramidite; **Y**= BHQ2dT.

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN72 | **8**TT CCT ATG A**Y**G AAT ATA GAT ACA GAA GCG **6** |
| ODN74 | GGC ACC ATT AAA GAA AAT ATC A |

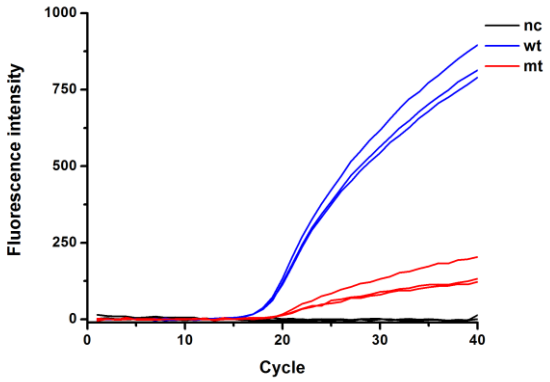



**Figure 4.13.** Fluorescence accumulation of RT-PCR using a Taqman probe containing Cy3B and BHQ2 (ODN72). The sample with wild-type (wt, blue) template has a $C_T$ value of 16 whilst the sample with template mutation (mt, red) has a $C_T$ value of 19. In the sample with the mutated-template the probe-template duplex is unstable and so the probe is cleaved at a far slower rate.

### 4.3.2 Demonstration of a HyBeacon probe containing Cy3BdT phosphoramidite

As discussed in section 4.2.5, the best HyBeacon probe was found to be ODN63 containing two Cy3BdT monomers. Asymmetric RT-PCR was carried out using the Precision-HRM mastermix. Although the mastermix contains 5'-3' exonuclease activity the design of the probe and the assay means that at the extension temperature the probe is not annealed and so probe cleavage is avoided. The design of a HyBeacon probe means that although it can be used to monitor the PCR in real-time its major benefit is from a post-amplification melt, from which the wt/mt discrimination may be unambiguously assigned.[79,78,80]

It can be seen from figure 4.14 that the use of HyBeacon probes in PCR is highly advantageous. As shown in the RT-PCR fluorescence accumulation graph (figure 4.14-left), the samples containing wild-type template (wt) have a $C_T$ value of 22 whilst those containing mutated-template (mt) have a later $C_T$ value of 25. In addition, the ratio of fluorescence intensity is ~4:1 (wt:mt) as a result of the probe-mt duplex being less stable than the probe-wt duplex at the monitoring temperature (50 ºC). By taking the derivative of a post-amplification fluorescence melt (figure 4.14-right) clearly defined $T_m$ values are given for the wild-type (50 ºC) and mutant-type (45 ºC) probe-template duplexes, making it possible to identify them unambiguously.

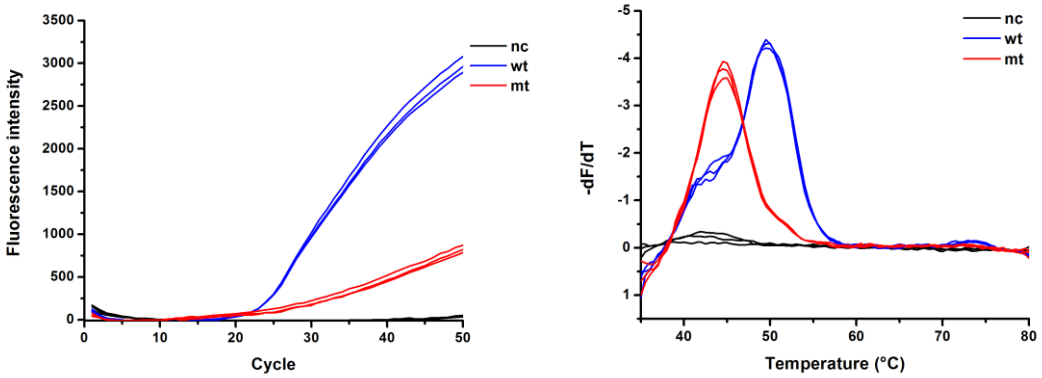



**Figure 4.14.** Fluorescence accumulation of RT-PCR using a HyBeacon probe containing Cy3BdT (ODN63) (left). The samples containing wild-type template (wt, blue) have a $C_T$ value of 22 whilst those containing mutated-template (mt, red) have a $C_T$ value of 25. Taking the derivative of a post-amplification fluorescence melt (right) gives clearly defined $T_m$ values for the wild-type (50 ºC) and mutant-type (45 ºC) probe-template duplexes, making it possible to identify them unambiguously.

### 4.3.3 Demonstration of Molecular Beacon probes containing 5'Cy3B and Cy3BdR phosphoramidites

The initial Molecular Beacon was designed to probe the same target DNA locus as the HyBeacon probe, however the melting temperature of the probe between the wild-type and the mutant target was not adequate to discriminate between them. Mismatch base pairs in DNA are generally different in shape from Watson-Crick base pairs resulting in unstable base stacking interactions within the DNA double-helix. The thermodynamic instability is different for each type of mismatch.[158] It was found that by probing the opposite strand of the DNA locus an improved discrimination in melting temperature between the probe:wt complex and probe:mt complex was achieved, as an AC mismatch is probed instead of a TG mismatch.[159] AC mismatches are particularly unstable at neutral and high pH (figure 4.15).[160]

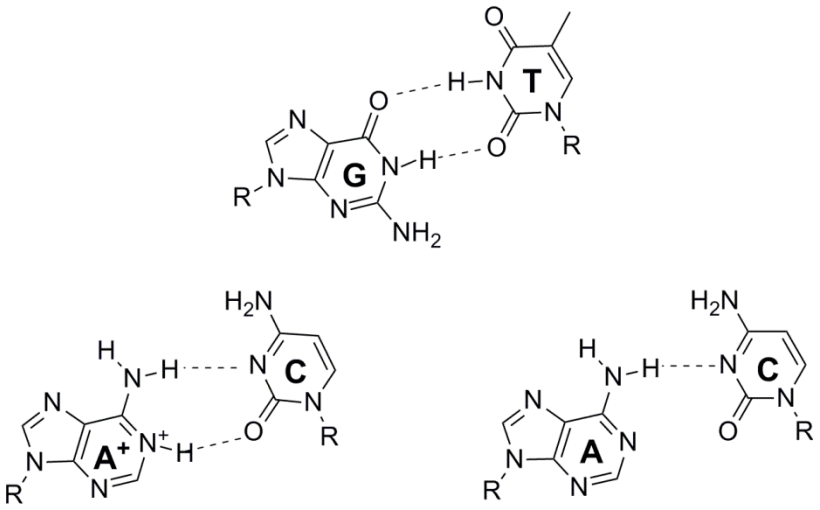


**Figure 4.15.** G.T, $A^+$.C and A.C mismatch base pairs. The $A^+$.C mismatch requires protonation of adenine (pKa ~3-4) so is unlikely to occur at neutral or high pH. The A.C mismatch can only form one hydrogen bond which makes it very unstable.

Molecular Beacon probes were evaluated with several different quenchers to assess which works best in combination with Cy3B. BHQ1 (ODN71), BHQ2 (ODN65) and DABCYL (ODN76) were incorporated into Molecular Beacon probes along with Cy3B (table 4.10). A dual-dye "quencher-free" Molecular Beacon (ODN78) was also synthesised. To yield an optimum Molecular Beacon, the quencher should provide good contact quenching and low

FRET quenching. This yields low fluorescence in the closed-form and high fluorescence in the hybridised-form.[69-73] From the literature[55] it was found that DABCYL would be the best contact quencher for Cy3B with least FRET quenching, however it was not known how well a Cy3B dual-dye Molecular Beacon would function. The two initial BHQ probes were synthesised with the 5'Cy3B addition but, due to lack of sufficient quantities of 5'Cy3B phosphoramidite and the availability of the Cy3BdR monomer, the DABCYL and dual-dye probe were synthesised with the Cy3BdR phosphoramidite instead. For the purpose of comparison it was assumed both would behave similarly as they both have a hexynol linkage to the DNA. The dual-dye Molecular Beacon was synthesised with a 3'Cy3BdR using the prepared synthesis resin.

Dual-dye Molecular Beacons (quencher-free Molecular Beacons, QF-MBs), are described in a review by Venkatesan *et al*.[161] In the closed form the two dyes stack and quench each other, and in the hybridised-form they give double the fluorescence output of standard dye-quencher Molecular Beacons. Variations on the dual-dye QF-MBs have been studied using several chromophores[162-164] with varied success but it still remains a challenge to obtain good closed form quenching yet retain high enough fluorescence in the hybridised form for reliable RT-PCR or single-molecule studies. It was hoped both these requirements could be fulfilled using Cy3B dual-dye QF-MB's.

The UV/Vis absorbance spectrum indicated that the dual-dye QF-MB exhibited excimer-like properties in the closed-form. As seen in figure 4.16a the closed-form has a new absorbance maxima at ~525nm, blue-shifted from the normal absorbance maximum of Cy3B by around 50 nm. Interestingly this is the same wavelength as the typical shoulder-peak of Cy3B. Since most RT-PCR fluorescence-monitoring equipment has fixed wavelengths, it was important to test whether excitation of the probe at the shoulder-wavelength would affect the probing ability. This was to establish whether a channel exciting at 525 nm would result in a high closed-form fluorescence and lower hybridised-form fluorescence, hence corrupting the results. It was found that exciting the probe at 525 nm in fact resulted in decreased fluorescence emission for both the hybridised and closed-forms (figure 4.16b). It was therefore concluded that excitation at the wavelength of the new absorbance maximum in the excimer-form does not translate to increased (or even equivalent) fluorescence emission. This phenomenon was likewise established in the excimer study of chapter 3.6 where dye-dimer (or excimer) theory is discussed more fully.

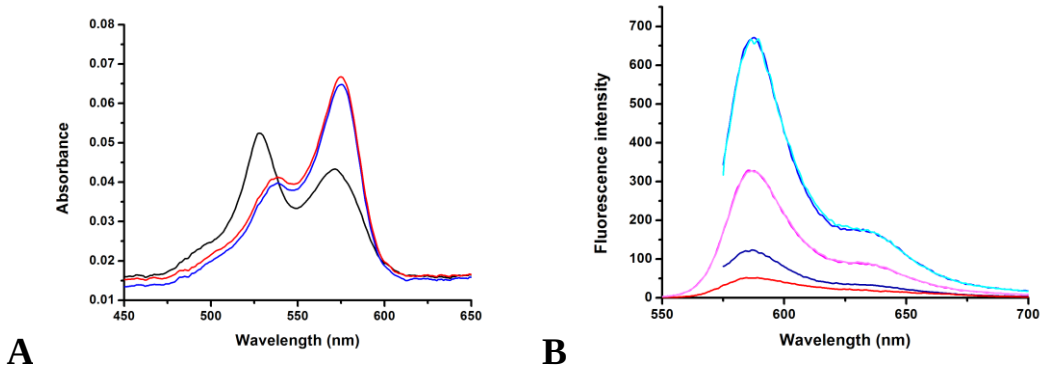



**Figure 4.16.** Photophysical properties of the dual-dye Molecular Beacon probe (ODN78); (a) absorbance spectra, ss (black), wt (blue), mt (red); (b) emission spectra: excitation 570 nm, ss (black), wt (blue), mt (cyan) and excitation (525 nm), ss (red), wt (magenta), mt (light pink). The peak at 525 nm in the absorbance spectra and decreased fluorescence emission suggest the single-stranded sample has excimer properties.

**Table 4.10.** Oligonucleotide sequences (mass data in appendix A.3). **8**= 5-hexyn-1-ol-6-Cy3B phosphoramidite; **9**= BHQ2; **Z**= BHQ1; **R**= Cy3BdR resin; **S**= Cy3BdR phosphoramidite; **D**= DABCYL.

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN65 | **8**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **9** |
| ODN68 (wt) | TGA TGA CGC TTC TGT ATC TAT ATT CAT CAT AGG A |
| ODN69 (mt) | TGA TGA CGC TTC TGT ATC CAT ATT CAT CAT AGG A |
| ODN71 | **8**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **Z** |
| ODN76 | **S**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **D** |
| ODN78 | **S**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **R** |

Upon comparing the fluorescence emission of the different Molecular Beacon variations in both the hybridised and closed form, it was found that ODN78 (dual-dye Molecular Beacon) has the highest fluorescence intensity in the hybridised form, followed by ODN76 (DABCYL Molecular Beacon). ODN71 and ODN65 (BHQ1 and BHQ2 Molecular Beacons) have a similar fluorescence (figure 4.17). The quenching efficiency in the closed-form is best in the DABCYL and BHQ1 probes (96 % quenching each), the BHQ2 example gives 94 % quenching and the dual-dye probe gives 82 % quenching. Since it was not known whether the quenching ability or the hybridised-form fluorescence would be more important in RT-PCR, it was decided to test both ODN76 and ODN78 in the RT-PCR experiment.

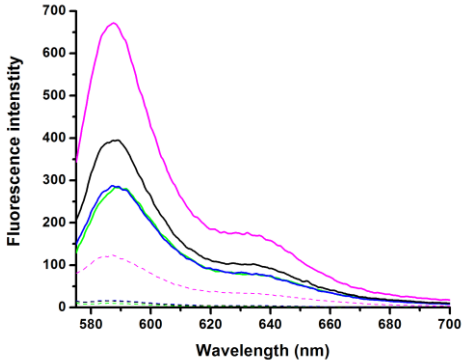



**Figure 4.17.** Comparison of Molecular Beacon probes by room temperature fluorescence emission. ODN71 probe with BHQ1 quencher (green); ODN65 probe with BHQ2 quencher (blue); ODN76 probe with DABCYL quencher (black); ODN78 probe with dual-Cy3B dyes (pink); ss (dashed line), wt (solid line). The dual-Cy3B probe exhibits highest fluorescence in the double-stranded form, although the probes with DABCYL, and BHQ quenchers give better quenching in the closed-form.

Both the dual-dye Molecular Beacon (ODN78) and DABCYL Molecular Beacon (ODN76) were evaluated by fluorescence melting to establish the optimum monitoring temperature for wt/mt discrimination (figure 4.18). A monitoring temperature of 58 ºC was chosen to give highest wild-type fluorescence and lowest mutant-type fluorescence. The asymmetric RT-PCR protocol was set-up as described in the experimental section 6.3.3 and for both ODN76 and ODN78 it gave successful wt/mt discrimination in RT-PCR (figures 4.19). In both probe examples the sample containing wild-type template has a $C_T$ value of 20 whereas the samples containing the mutant-type template show no real-time fluorescence accumulation. This is because the duplexes between the mutated-template and the Molecular Beacon probes are less stable than the equivalent wild-type duplexes and so are not formed at the monitoring temperature of the PCR (58°C).

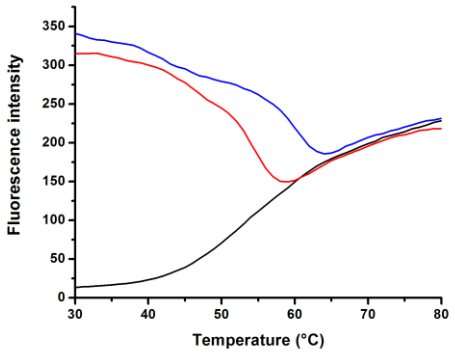



**Figure 4.18.** Fluorescence melting profile for the Molecular Beacon probe ODN76; ss (black), wt (blue), mt (red). The melting curves indicate the optimum monitoring temperature is 58 °C (at which the highest wt fluorescence and lowest mt fluorescence would be achieved).

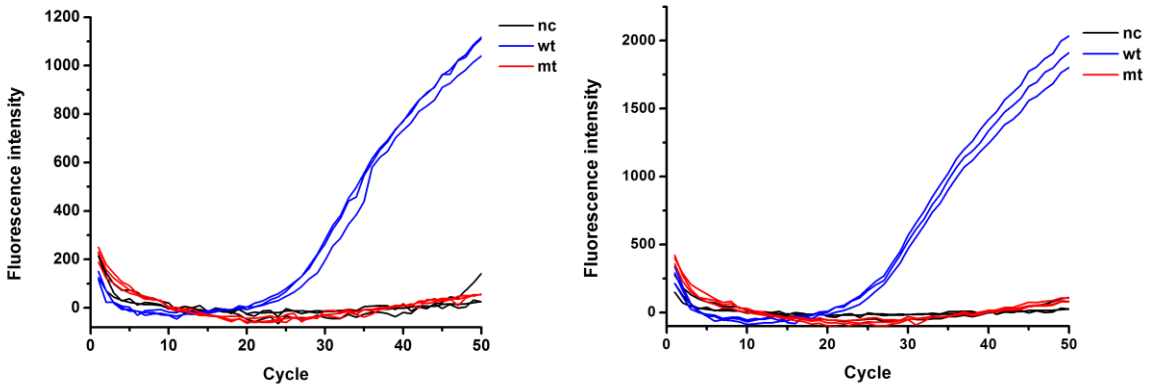



**Figure 4.19.** Fluorescence accumulation of RT-PCR using a Molecular Beacon probe containing Cy3B and DABCYL (ODN76, left) and a dual-dye QF-MB (ODN78, right). In both cases the sample containing wild-type template has a $C_T$ value of 20 whereas the samples containing the mutant-type template show no real-time fluorescence accumulation. This is because at the monitoring temperature (58 °C) the (mt)template:probe duplexes are very unstable.

### 4.3.4 Demonstration of Scorpion primers containing 5'Cy3B and Cy3BdR phosphoramidites

Two Scorpion primer sequences were synthesised containing Cy3B and either a BHQ1 quencher (ODN73) or a DABCYL quencher (ODN77) (table 4.11). From figure 4.20 it can be seen that although the quenching ability is similar for both probes (~93 %), ODN77 gave superior hybridised-form fluorescence. Symmetric RT-PCR was carried out using ODN77 (the conditions are described in the experimental section 6.3.3). The probe exhibited very good wt/mt discrimination and an excellent real-time signal (figure 4.21). The sample containing the wild-type template has a $C_T$ value of 18 and the final fluorescence intensity is high which shows the probe-template duplex formation is very efficient. The advantage of the Scorpion primer in RT-PCR is that the probe-target binding is kinetically favoured over re-annealing of the amplicon duplex, and thermodynamically favoured over the intrastrand hairpin structure, giving a fast and reliable signal.[75] The mechanism of the primer/probe was confirmed as being unimolecular by Thelwell *et al*, in comparison to the slower bimolecular mechanism of equivalent Taqman and Molecular Beacon probes.[76]

**Table 4.11.** Oligonucleotide sequences (mass data in appendix A.3). **8**= 5-hexyn-1-ol-6-Cy3B phosphoramidite; **X**= BHQ1dT; **W**= HEG; **S**= Cy3BdR phosphoramidite; **Q**= DABCYL-dT.

| **Oligonucleotide code** | **Sequence** |
|---|---|
| ODN73 | **8**CC GCG GGA TGA ATA TAG ATA CAG AAG CGC CGC GG**X W**TC TTC TAG TTG GCA TGC T |
| ODN77 | **S**CC GCG GGA TGA ATA TAG ATA CAG AAG CGC CGC GG**Q W**TC TTC TAG TTG GCA TGC T |

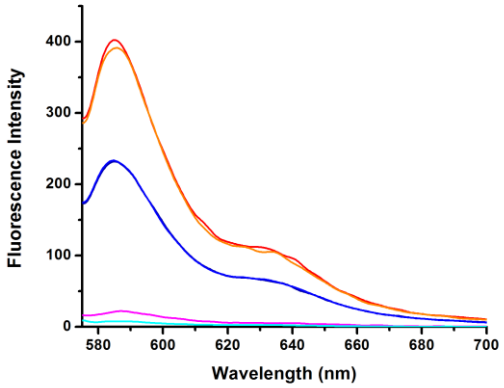



**Figure 4.20.** Room temperature fluorescence emission spectra of Scorpion primers (ss, wt and mt forms). ODN73 with BHQ1 quencher (5'Cy3B monomer); ss (cyan), wt (navy), mt (blue), ODN77 with DABCYL quencher (Cy3BdR monomer); ss (pink), wt (red), mt (orange). ODN77 (Scorpion primer containing DABCYL) gives highest fluorescence intensity in the hybridised form although quenching efficiency is similar (~93 %) for both examples.

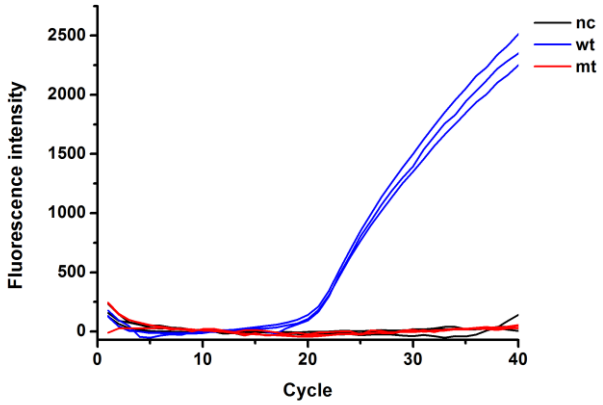



**Figure 4.21.** Fluorescence accumulation of RT-PCR using a Scorpion primer containing Cy3B and DABCYL (ODN77). The sample containing the wild-type template has a $C_T$ value of 18 and the final fluorescence intensity is high which shows the probe-template duplex formation is very efficient.

### 4.3.5 Demonstration of a PCR primer labelled with Cy3B at the 5'-end

To demonstrate further use of the Cy3B 'toolkit' in PCR, a forward primer with a 5'Cy3B was synthesised and used in a PCR assay (table 4.2). The PCR could not be monitored in real-time since the fluorescence of the Cy3B is unaltered between single- and double-stranded forms. Therefore the PCR was monitored by running a control with the addition of SYBR-green in tandem. The PCR reaction was consequently analysed by agarose gel electrophoresis to visualise the product (figure 4.22). It can be seen in the figure that the PCR reaction samples (lanes 4 and 5, duplicate samples) which contained synthetic template DNA resulted in a PCR product (135 bases in length). The duplicate negative control samples in lanes 2 and 3 (those without synthetic template in the reaction) exhibit only the un-extended primer as expected.

5'dye labelled oligonucleotides are used commonly in DNA sequencing[165,166] and so the convenient phosphoramidite incorporation of the highly fluorescent Cy3B, which would have lower detection limits due to its high fluorescence, would be extremely useful.

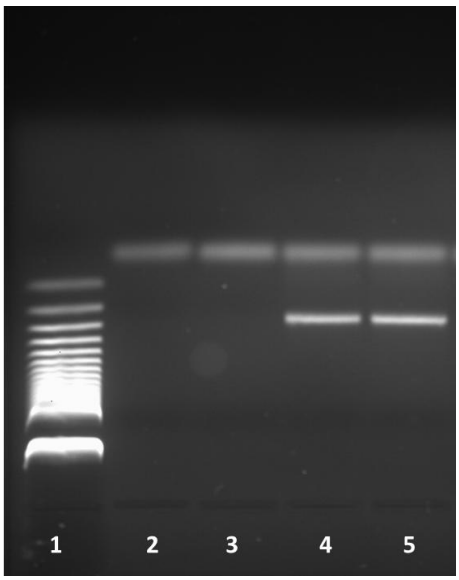



**Figure 4.22.** Product visualisation of a PCR experiment using a 5'Cy3B- labelled primer (ODN67) by agarose gel (2 %) electrophoresis. Lane 1- DNA stepladder; lane 2&3- negative control; lane 4&5- crude reaction product. The PCR reaction samples (lanes 4 and 5, duplicate samples) which contained synthetic template DNA exhibit a PCR product (135 bases in length). The negative control samples (those without synthetic template in the reaction, lanes 2 and 3, duplicate samples) exhibit only the un-extended primer. The gel was imaged using G:Box imaging system (Syngene) and Genesnap software (v7.08, Syngene). Bands were imaged with transluminator darkroom lighting and EtBr/UV filter.

## 4.4 Conclusions

A Cy3B DNA 'molecular toolkit' was synthesised for the addition of the highly fluorescent Cy3B into multiple sites of oligonucleotides. Cy3B addition has been demonstrated both internally and at either and both ends of oligonucleotides. The monomers allow for addition of Cy3B on both a ribose sugar (Cy3BdR phosphoramidite and resin or by click-chemistry at the 2' position) and at the C5 base-position of thymidine; by either a short rigid linker (Cy3BdT) or by a long flexible linker (click-chemistry addition). Upon the formation of the double-helix these correspond to addition in either the minor or major groove of DNA respectively. The highly fluorescent Cy3B dye is currently only commercially available for the modification of oligonucleotides after solid-phase synthesis and it is hoped that the new options presented will greatly extend the use of Cy3B in research and commercial applications.

It has been demonstrated that various Cy3B monomers can be used in a variety of applications. Examples of the commonly used PCR probes were synthesised, incorporating different Cy3B additions into a Taqman probe, a HyBeacon probe, a Molecular Beacon, a Scorpion primer and a 5'labelled primer. In a more detailed study multiple probes of each type could be synthesised corresponding to different mutations or SNPs and used in the same PCR assay to allow genotyping/ point mutation detection.

Several useful intermediates have been described including 5-ethynyl-Cy3B (for click-chemistry or palladium cross-coupling addition) and 1'-hexynol-dR. These compounds, together with the previously described 5-ethynyl-dT and 5-ethynyl-dC,[98,116-119] provide a useful array of compounds from which a wider variety of dye or quencher phosphoramidites may be synthesised.

For future work it would be interesting to use the Cy3B monomers in oligonucleotide synthesis to produce structures with multiple dye additions (introduced in chapter 3). Nanostructures with many additions of Cy3B could be synthesised to generate highly fluorescent DNA nanoparticles.

# Chapter 5

# Conclusions and Future Work

# Chapter 5 – Conclusions and Future Work

It was of interest to us to synthesise phosphoramidite monomers for incorporation into synthetic DNA which would impart maximum fluorescence whilst maintaining a rigid linker to the DNA. The concept was that the rigid linker would fix the dye precisely in space to enable accurate quantitative FRET distance measurements to be made. The rigidity was also thought to be important to minimise dye-DNA interactions and to reduce fluorescence quenching.

Initially a new fluorescent base analogue, CPP, was synthesised which did not perturb DNA duplex stability. It was found to have a high fluorescence quantum yield (0.75) and extinction coefficient (14,000 $M^{-1}cm^{-1}$) in the free monomer form, relative to other fluorescent base analogues such as the tricyclic cytosine analogue tC (QY ~0.2, extinction coefficient 4,500 $M^{-1}cm^{-1}$). However, the fluorescence quantum yield of CPP significantly dropped upon oligonucleotide incorporation and again upon duplex formation. This behaviour could potentially be used for sensing the DNA environment (such as sensing hybridisation in sequence detection experiments) but it may be problematic in other applications. For example, for most DNA-related applications, such as PCR and single-molecule DNA studies, it is preferable if the dye retains its fluorescence in double-stranded DNA or in an ideal case it should become more fluorescent. Due to the low extinction coefficient of CPP (in comparison to other fluorophores) it was clear it did not possess the highly fluorescent properties we had hoped to achieve to enable further studies of FRET and biological-systems.

Commercially available fluorophores tend to have far higher extinction coefficients than fluorescent base analogues (and often have high quantum yields). We therefore decided to develop alternative compounds by conjugating known fluorophores into the nucleotide structure which could provide highly fluorescent DNA monomers

New fluorescent phosphoramidite monomers based on CyDyes were synthesised which permitted the addition of multiple and mixed fluorophores into oligonucleotides during solid-phase synthesis. The properties of the Cy3dT and Cy5dT monomers were investigated and they have been shown to be viable thymidine analogues in terms of base pair formation. Although each addition of Cy3dT thermally destabilised the duplex

by -2.4 ºC, and Cy5dT by slightly more (-3.1 ºC), oligonucleotide duplexes containing the monomers still conform to B-form helix structure. Multiple additions of the Cy3dT monomer into a single oligonucleotide was successfully achieved. Up to ten dye additions have been successfully incorporated into one duplex, which would be unachievable with NHS-ester labelling, demonstrating the power of this methodology. Increasing the number of dye-additions resulted in the formation of excimer structures between the dyes. These excimers were found to result in quenching of the fluorescent system.

Oligonucleotides containing Cy3dT and Cy5dT were used to establish a FRET system. By using a series of duplexes with varying spacing between the two dyes, FRET was shown to decrease with increasing inter-nucleotide separation. Oligonucleotides containing Cy3dT and Cy5dT exhibited unusual room-temperature fluorescence emission and fluorescence melting properties. The fluorescence intensity of oligonucleotides containing CydT was found to be lower in double-stranded form and at higher temperatures. It was proposed that, despite the presence of the rigid linker to DNA, in the duplex the location of the dyes in the major groove makes them susceptible to rotation around the polymethine linker, leading to *trans-cis* isomerisation. This leads to reduced fluorescence and inverse temperature dependent fluorescence. These factors could be a disadvantage in some applications, so we proceeded to investigate the more rigid analogue Cy3B.

A Cy3B DNA ‘molecular toolkit’ was synthesised for the incorporation of the highly fluorescent Cy3B into multiple sites of oligonucleotides. Cy3B addition was demonstrated both internally and at either and both ends of oligonucleotides by using the range of new monomers. 5-Ethynyl-Cy3B was also developed to provide the option of ‘click-chemistry’ addition to DNA and proteins. The highly fluorescent Cy3B dye is currently only commercially available for the modification of oligonucleotides after solid-phase synthesis. Therefore the new phosphoramidite monomer range is a great advantage for research and commercial applications as it provides a way to incorporate Cy3B with ease and in high yield.

We demonstrated by synthesising examples of the commonly used PCR probes that the Cy3B phosphoramidites can be used in a variety of applications. A Taqman probe, a HyBeacon probe, a Molecular Beacon, a Scorpion primer and a 5’labelled primer, which incorporate the different Cy3B addition options, demonstrate the versatility of the new monomers.

In the future we would be very interested in applying the range of new CyDye phosphoramidites to nanotechnology applications. In particular, the highly fluorescent Cy3B could be used for incorporation into nanostructures. By exploiting the ease of incorporation of multiple CydT monomers into DNA, many Cy3B additions could be made to generate highly fluorescent DNA nanoparticles. Investigations will be undertaken to evaluate what spacing between the dyes can be used to prevent the excimer quenching effect.

# Chapter 6

# Experimental

# Chapter 6 - Experimental

## 6.1. Synthesis

### 6.1.1. General:

Chemicals were purchased from Sigma-Aldrich, Fisher Scientific, Alfa Aesar, Apollo Scientific and Link Technologies. The following solvents were purified by distillation (over calcium hydride) before use: pyridine, $Et_3N$, DCM, MeCN and MeOH. THF was purified by distillation over sodium. Anhydrous DMF was purchased from Sigma-Aldrich. When mentioned, deoxygenated solutions/solvents were prepared by bubbling argon through the solution for 15 min. Reactions were carried out using air-sensitive techniques, under a degassed atmosphere of argon in cleaned, dry glassware.

Reactions were monitored by thin layer chromatography which was carried out on Merck Kieselgel 60 $F_{254}$ plates (0.22 mm thickness, aluminium backed). They were visualised by UV irradiation at 254/365 nm. Column chromatography was carried out under air-pressure using Fisher Scientific DAVISIL 60 Å (35-70 micron) silica gel. Silica was pre-equilibrated with $Et_3N$ or pyridine before purification of acid-sensitive compounds.

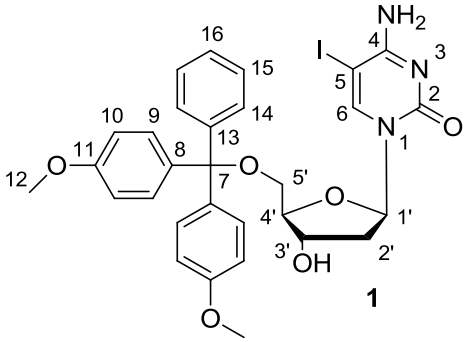
$^1H$ NMR spectra were measured on a Bruker AC300 spectrometer at 300 or Bruker DPX400 spectrometer at 400 MHz. $^{13}C$ spectra were measured at 75 or 100 MHz respectively. $^{31}P$ spectra were measured at 121 MHz. All shifts were referenced to the residual solvent peak and are given in ppm. *J* coupling values are correct to 0.5 Hz. Assignment was aided by DEPT-135, $^1H$-$^1H$ COSY, HMQC and HMBC experiments. Experiments were carried out using deuterated $CDCl_3$, $d_6$-DMSO, MeOD or $CD_3CN$. Low-resolution mass spectra were measured on a Waters ZMD quadrupole mass spectrometer for +/- electrospray ionisation (ESI). These were carried out in HPLC grade MeOH or MeCN.

High-resolution mass spectra were recorded in MeOH or MeCN on a Bruker APEX III FT-ICR mass spectrometer (ESI) or a VG Analytical 70-250-SE mass spectrometer (EI). Melting points were determined on a Gallenkamp Electothermal melting point apparatus.

Common procedures including tritylation,[92] palladium cross-coupling,[94,93] trimethylsilyl deprotection[96,97] and phosphitylation[167] were based on literature procedures.
Quantities of reagents are quoted to 3 significant figures or 2 decimal places as appropriate.

### 6.1.2. Experimental procedure

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(iodo)cytosine (1)[92]**




5-iodo-deoxycytidine (1.49 g, 4.22 mmol) was co-evaporated with distilled pyridine (3 x 6 mL) and dried overnight *in vacuo* before re-suspending in distilled pyridine (25 mL) under an argon atmosphere.
Over 3.5 h a solution of DMTCl (1.71 g, 5.06 mmol, in 12 mL distilled pyridine) was added drop-wise to the reaction which was stirred at rt. After 5 h the reaction was quenched with MeOH (10 mL) and left to stir for 10 min before reducing the solvent volume by half *in vacuo*. The remaining solution was diluted with DCM (200 mL) and washed with $NaHCO_3$ (3 x 200 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*.
The crude material was purified by column chromatography (3:97 MeOH/DCM with 1 % pyridine) affording **1** (1.86 g, 2.84 mmol, 67 %) as a fine white powder.

$R_f$ : 0.48 (MeOH/DCM 1:9).
LRMS: [ESI+, MeOH] m/z (%): 678 ($(M + Na)^+$, 100).
HRMS: [ESI+, MeOH] for $C_{30}H_{30}I_1N_3Na_1O_6$ $(M+Na)^+$: calcd 678.1071, found 678.1058.
$^1$H (300 MHz, $d_6$-DMSO): δ 7.97 (s, 1H, $H^6$), 7.87 (br.s, 1H, $NH_2$), 7.41-7.20 (m, 9H, $H^{Ar}$), 6.91 (d, *J*=8.9, 4H, $H^{Ar}$), 6.63 (br.s, 1H, $NH_2$), 6.11 (t, *J*=6.7, 1H, $H^{1'}$), 5.28 (d, *J*=4.1, 1H, $OH^{3'}$), 4.19 (m, 1H, $H^{3'}$), 3.91 (ddd, *J*=3.2, 3.2, 3.2, 1H, $H^{4'}$),

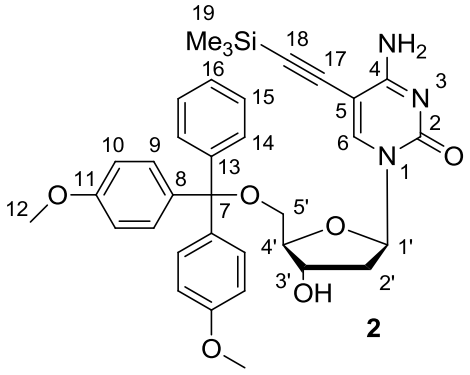
3.74 (s, 6H, $H^{12}$), 3.21-3.14 (m, 2H, $H^{5'}$), 2.22 (ddd, *J*=15.0, 6.0, 3.0, 1H, $H^{2'}$), 2.09 (ddd, *J*=13.7, 6.7, 6.7, 1H, $H^{2'}$) ppm.

$^{13}$C (75 MHz, $d_6$-DMSO): δ 163.8 ($C^2$), 158.1 ($C^4$), 153.8 ($C^{Ar}$), 146.3 ($CH^6$), 144.7 ($C^{Ar}$), 135.5 ($C^{Ar}$), 129.7 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 126.7 ($CH^{Ar}$), 113.3 ($CH^{Ar}$), 85.8 ($C^7$), 85.8 ($CH^{4'}$), 85.4 ($CH^{1'}$), 70.7 ($CH^{3'}$), 63.7 ($CH_2{}^{5'}$), 56.9 ($C^5$), 55.1 ($CH_3{}^{12}$), 40.8 ($CH_2{}^{2'}$) ppm.

Mp: 182 ºC (decomposes).

Characterisation data recorded matches previously reported literature values.[92]

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(trimethylsilyleth-1-ynyl)cytosine (2)**




To a solution of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(iodo)cytosine (**1**) (1.72 g, 2.62 mmol) in anhydrous DMF (5 mL) under an argon atmosphere was added CuI (0.10 g, 0.54 mmol), distilled $Et_3N$ (1.83 mL, 13.0 mmol) and trimethylsilylacetylene (1.11 mL, 7.85 mmol). The reaction was stirred at rt in the dark for 10 min before the addition of $Pd(PPh_3)_4$ (0.30 g, 0.26 mmol). The reaction was stirred at rt in the dark for 16 h. The reaction solvent was removed *in vacuo*. Following purification by column chromatography (2:98 EtOH/$CHCl_3$ with 1 % pyridine) the product **2** (1.42 g, 2.27 mmol, 86 %) was afforded as a yellow foam.

$R_f$ : 0.33 (EtOH/$CHCl_3$ 1:9).

LRMS: [ESI+, MeOH] m/z (%): 1274 ($(2M + Na)^+$, 6), 664 ($(M + K)^+$, 4), 648 ($(M + Na)^+$,100).

HRMS: [ESI+, MeOH] for $C_{35}H_{39}N_3Na_1O_6Si_1$ $(M+Na)^+$: calcd 648.2500, found 648.2498.

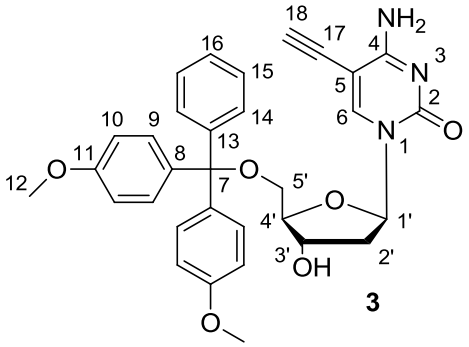
$^{1}$H (400 MHz, $d_6$-DMSO): δ 7.98 (s, 1H, $H^{6}$), 7.83 (br.s, 1H, $NH_2$), 7.42-7.19 (m, 9H, $H^{Ar}$), 6.88 (dd, *J*=8.9, 1.8, 4H, $H^{Ar}$), 6.62 (br.s, 1H, $NH_2$), 6.11 (t, *J*=6.7, 1H, $H^{1'}$), 5.29 (d, *J*=4.3, 1H, $OH^{3'}$), 4.23 (m, 1H, $H^{3'}$), 3.96 (m, 1H, $H^{4'}$), 3.73 (s, 6H, $H^{12}$), 3.16 (m, 2H, $H^{5'}$), 2.25 (ddd, *J*=13.3, 5.8, 2.9, 1H, $H^{2'}$), 2.08 (ddd, *J*=13.7, 6.6, 6.6, 1H, $H^{2'}$), 0.07 (s, 9H, $H^{19}$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): δ 164.0 ($C^{2}$), 158.0 ($C^{Ar}$), 153.1 ($C^{4}$), 144.6 ($C^{Ar}$), 144.2 ($C^{6}$), 135.6 ($C^{Ar}$), 135.4 ($C^{Ar}$), 129.7 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 127.6 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 113.2 ($CH^{Ar}$), 99.7 ($C^{18}$), 96.1 ($C^{17}$), 89.9 ($C^{5}$), 85.9 ($CH^{4'}$), 85.8 ($C^{7}$), 85.7 ($CH^{1'}$), 70.6 ($CH^{3'}$), 63.6 ($CH_2{}^{5'}$), 55.0 ($CH_3{}^{12}$), 40.9 ($CH_2{}^{2'}$), 0.0 ($CH_3{}^{19}$) ppm.

Mp: 84 °C.

IR $\nu_{Max}/cm^{-1}$: 3325 (m, amine N-H), 2951, 2668 (m, alkane C-H), 2150 (m, alkyne C≡C), 1636 (s, carbonyl C=O).

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(eth-1-ynyl)cytosine (3)**




**3**

To a solution of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(trimethylsilyleth-1-ynyl)cytosine (**2**) (1.38 g, 2.21 mmol) in distilled THF (7 mL) under an argon atmosphere, was added 1M TBAF in THF (2.73 mL, 2.73 mmol). The reaction was stirred for 15 min, then diluted with EtOAc (200 mL) and washed with brine (3 x 200 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. Following purification by column chromatography (5:95 MeOH/DCM with 1 % pyridine) the product **3** (0.73 g, 1.32 mmol, 60 %) was afforded as an orange/beige solid.

$R_f$: 0.35 (MeOH/DCM 1:9).

LRMS: [ESI+, MeOH] m/z (%): 1130 ($(2M + Na)^+$, 11), 592 ($(M + K)^+$, 5), 576 ($(M + Na)^+$, 100).

HRMS: [ESI+, MeOH] for $C_{32}H_{31}N_3Na_1O_6$ $(M+Na)^+$: calcd 576.2105, found 576.2100.

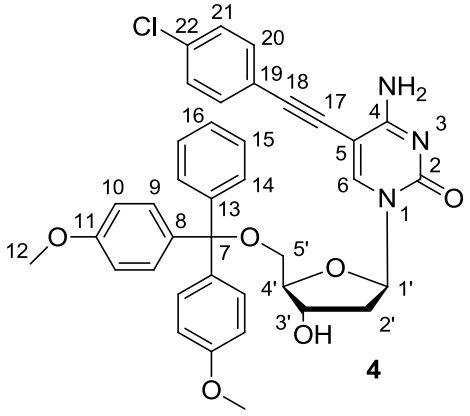
$^1$H (300 MHz, MeOD): δ 8.16 (s, 1H, $H^6$), 7.41-7.13 (m , 9H, $H^{Ar}$), 6.81 (dd, *J*=8.9, 2.1, 4H, $H^{Ar}$), 6.12 (t, *J*=6.4, 1H, $H^{1'}$), 4.41 (ddd, *J*=6.0, 3.1, 1H, $H^{3'}$), 4.05 (dt, *J*=3.1, 3.1, 1H, $H^{4'}$), 3.73 (s, 6H, $H^{12}$), 3.57 (s, 1H, $H^{18}$), 3.28-3.26 (m, 2H, $H^{5'}$), 2.49 (ddd, *J*=13.7, 6.1, 3.3, 1H, $H^{2'}$), 2.22 (ddd, *J*=13.6, 6.6, 6.6, 1H, $H^{2'}$) ppm.

$^{13}$C (75 MHz, MeOD): δ 166.5 ($C^2$), 160.2 ($C^{Ar}$), 156.6 ($C^4$), 146.1 ($C^6$), 137.1 ($C^{Ar}$), 136.9 ($C^{Ar}$), 131.3 ($CH^{Ar}$), 129.2 ($CH^{Ar}$), 128.9 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 114.3 ($CH^{Ar}$), 91.8 ($C^5$), 88.5 ($CH^{4'}$), 88.3 ($CH^{18}$), 88.2 ($CH^{1'}$), 85.5 ($C^7$), 75.3 ($C^{17}$), 72.7 ($CH^{3'}$), 64.7 ($CH_2{}^{5'}$), 55.7 ($CH_3{}^{12}$), 42.9 ($CH_2{}^{2'}$) ppm.

Mp: 140 ºC (decomposes to dark foam).

IR $\nu_{Max}/cm^{-1}$: 3272 (m, alkyne C-H overlap with; m, amine N-H), 3050 (m, aromatic C-H), 2931, 2835 (m, alkane C-H), 2106 (m, alkyne C≡C), 1635 (s, carbonyl C=O).

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(4-chlorophenyl eth-1-ynyl)cytosine (4)**




**4**

To a solution of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(eth-1-ynyl)cytosine (**3**) (2.02 g, 3.65 mmol) in anhydrous DMF (18.3 mL) under an argon atmosphere, was added CuI (0.14 g, 0.73 mmol), 1-chloro-4-iodobenzene (1.74 g, 7.31 mmol) and distilled $Et_3N$ (11.7 mL). After stirring for 10 min at rt, $Pd(PPh_3)_4$ (0.42 g, 0.37 mmol) was added in one portion. The reaction was stirred at rt in the dark for 1 hour

before the solvent was evaporated *in vacuo*. Following purification by column chromatography ($CHCl_3$ with 1 % pyridine→ 1:9 EtOH/$CHCl_3$ with 1 % pyridine) the product **4** (2.28 g, 3.44 mmol, 94 %) was afforded as an off-white solid.

$R_f$ : 0.59 (EtOH/$CHCl_3$ 1:9).

LRMS: [ESI+, MeOH] m/z (%): 1350 ($(2M+Na)^+$, 5), 686 ($(M+Na)^+$, 100).

HRMS: [ESI+, MeOH] for $C_{38}H_{34}Cl_1N_3Na_1O_6$ $(M+Na)^+$: calcd 686.2028, found 686.2044.

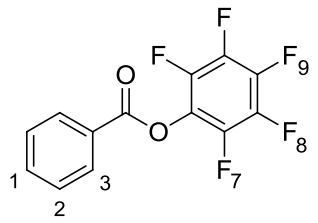
$^1$H (400 MHz, $d_6$-DMSO): δ 8.13 (s, 1H, $H^6$), 7.84 (br.s, 1H, $NH_2$), 7.44-7.15 (m, 13H, $H^{Ar}$) 7.11 (br.s, 1H, $NH_2$),6.85 (dd, *J*=8.8, 3.6, 4H, $H^{Ar}$), 6.15 (t, *J*=6.5, 1H, $H^{1'}$), 5.35 (br s, 1H, $OH^{3'}$), 4.29 (m, 1H, $H^{3'}$), 3.99 (d, *J*=3.1, 1H, $H^{4'}$), 3.66 (s, 6H, $H^{12}$), 3.18-3.17 (m, 2H, $H^{5'}$), 2.31 (m, 1H, $H^{2'}$), 2.14 (ddd, *J*=13.4, 6.6, 6.6, 1H, $H^{2'}$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): δ 163.7 ($C^2$), 158.0 ($C^{Ar}$), 153.2 ($C^4$), 144.7 ($C^{Ar}$), 144.1 ($CH^6$), 135.6 ($C^{Ar}$), 135.4 ($C^{Ar}$), 132.9 ($C^{Ar}$), 132.7 ($CH^{Ar}$), 129.6 ($CH^{Ar}$), 128.2 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 127.6 ($CH^{Ar}$), 126.7 ($C^{Ar}$), 121.2 ($C^{Ar}$), 113.2 ($CH^{Ar}$), 92.6 ($C^{18}$), 89.6 ($C^5$), 86.0 ($CH^{4'}$), 85.9 ($C^9$), 85.8 ($CH^{1'}$), 82.3 ($C^7$), 70.5 ($CH^{3'}$), 63.5 ($CH_2{}^{5'}$), 54.9 ($CH_3{}^{12}$), 41.2 ($CH_2{}^{2'}$) ppm.

Mp: 152 ºC (decomposes).

IR $\nu_{Max}/cm^{-1}$: 3389 (w, amine N-H), 3053 (br. w, O-H), 2910, 2834 (m, alkane C-H), 1629 (s, carbonyl C=O).

**Synthesis of benzoic acid pentafluorophenyl ester (5)[168]**




**5**

To a solution of benzoic acid (1.27 g, 10.0 mmol) and pentafluorophenol (2.00 g, 11.0 mmol) in distilled DCM (30 mL), under an argon atmosphere, was slowly added DCC (2.27 g, 11.0 mmol). After stirring at rt for 1 hour the reaction mixture was filtered to remove the white precipitous side product (dicyclohexylurea, DCU). The filtrate was concentrated *in vacuo* and the residue re-dissolved in DCM (10 mL) before repeating the

filtering procedure. The crude product was dissolved in warm hexane (3 mL) and filtered hot before being left to crystallise at rt then at 4 °C. The resulting crystals were collected by filtration, affording the product **5** (2.61 g, 9.06 mmol, 91 %) as clear/white crystals.

$R_f$ : 0.64 (EtOAc/hexane 5:95)

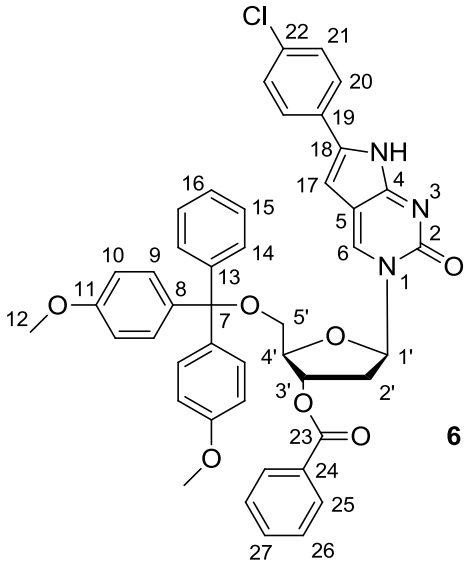
LRMS: [ESI+, MeOH] m/z (%): 288 ($(M)^+$, 3).

$^1$H (300 MHz, $CDCl_3$): δ 8.20 (dd, *J*=8.4, 1.1, 2H, $H^3$), 7.71 (tt, *J*=7.0, 1.5, 1H, $H^1$), 7.56 (td, *J*=8.1, 1.5, 2H, $H^2$) ppm.

$^{19}$F (280 MHz, $CDCl_3$): δ -152.7 (d, *J*=18.3, 2F, $F^7$), -158.2 (t, *J*=22.8, 1F, $F^9$), -162.6 (t, *J*=18.3, 2F, $F^8$) ppm.

Characterisation data recorded matches previously reported literature values.[168]

**Synthesis of 1-[2'-(deoxy)-3'-*O*-(benzoyl)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-6-(4-chloropheynyl)-3 *H*-pyrrolo-[2,3-*d*]pyrimidine-2-one (6)**




To a solution of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(4-chlorophenyl eth-1-ynyl)cytosine (**4**) (0.68 g, 1.02 mmol) in anhydrous DMF (10 mL) under an argon atmosphere was added distilled $Et_3N$ (3 mL) followed by the slow addition of benzoic acid pentafluorophenylester **5** (0.59 g, 2.05 mmol). After stirring for 48 h at 80 ºC the reaction was dried *in vacuo*. The residue was re-dissolved in DCM (20 mL) and was washed with $NaHCO_3$ (3 x 20 mL). The organic phase was dried ($Na_2SO_4$) and the

solvent removed *in vacuo*. Following purification by column chromatography (3:97 EtOH/$CHCl_3$ with 1 % pyridine) the product **6** was afforded as a yellow/brown foam (0.49 g, 0.69 mmol, 63 %).

Some deprotected product **7** (0.08 g, 0.12 mmol, 12 %) was collected during the column chromatography (1:9 EtOH/$CHCl_3$ with 1 % pyridine) and was afforded as a yellow foam.

$R_f$ : 0.58 (EtOH/$CHCl_3$ 1:9)

LRMS: [ESI+, MeOH] m/z (%): 793 ((M($^{37}$Cl)+Na)$^+$, 7), 791 ((M($^{35}$Cl)+Na)$^+$, 20).

HRMS: [ESI+, MeOH] for $C_{45}H_{38}Cl_1N_3Na_1O_7$ $(M+Na)^+$: calcd 790.2290, found 790.2315.

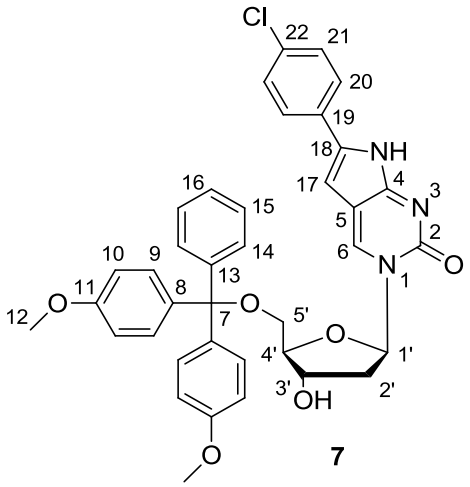
$^1$H (300 MHz, $d_6$-DMSO): δ 11.85 (br.s, 1H, NH), 8.63 (s, 1H, $H^6$), 8.02-7.14 (m, 18H, $H^{Ar}$), 6.87 (dd, *J*= 8.9, 1.2, 4H, $H^{Ar}$), 6.40 (t, *J*= 6.4, 1H, $H^{1'}$), 6.17 (s, 1H, $H^{17}$), 5.62 (ddd, *J*= 6.7, 3.5, 3.5, 1H, $H^{3'}$), 4.42 (dt, *J*= 3.6, 3.6, 1H, $H^{4'}$), 3.66 (br.s, 6H, $H^{12}$), 3.51 (dd, *J*=12.0,4.0, 1H, $H^{5'}$), 3.40 (dd, *J*= 10.7, 3.3, 1H, $H^{5'}$), 2.79 (ddd, *J*= 14.3, 6.2, 3.5, 1H, $H^{2'}$), 2.58-2.52 (m, 1H, $H^{2'}$) ppm.

$^{13}$C (75 MHz, $d_6$-DMSO): δ 166.2 ($C^{23}$), 161.1 ($C^2$), 159.1 ($C^{Ar}$), 154.6 ($C^4$), 145.4 ($C^{Ar}$) 137.4 ($CH^6$), 136.4 ($C^{Ar}$), 136.1 ($C^{Ar}$), 134.7 ($C^{18}$), 133.8 ($C^{Ar}$), 130.8 ($CH^{Ar}$), 130.7 ($CH^{Ar}$), 130.4 ($CH^{Ar}$), 129.9 ($CH^{Ar}$), 129.8 ($CH^{Ar}$), 129.2 ($CH^{Ar}$), 129.0 ($CH^{Ar}$), 128.7 ($CH^{Ar}$), 127.6 ($CH^{Ar}$), 126.3 ($CH^{Ar}$), 114.3 ($CH^{Ar}$), 110.2 ($C^5$), 98.2 ($CH^{17}$), 88.1 ($CH^{1'}$), 87.3 ($C^7$), 84.4 ($CH^{4'}$), 75.4 ($CH^{3'}$), 64.0 ($CH_2{}^{5'}$), 56.0 ($CH_3{}^{12}$), 39.7 ($CH_2{}^{2'}$) ppm.

Mp: 60 ºC (decomposes)

IR $\nu_{Max}$/$cm^{-1}$: 3066 (w, 2º amine N-H), 2931, 2836 (w, alkyl C-H), 1722 (m, amide C=O), 1659 (s, ester C=O).

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-6-(4-chloropheynyl)-3 *H*-pyrrolo-[2,3-*d*]pyrimidine-2-one (7)**




**7**

1-[2'-(Deoxy)-3'-*O*-(benzoyl)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-6-(4-chloropheynyl)-3 *H*-pyrrolo-[2,3-*d*]pyrimidine-2-one (**6**) (0.44 g, 0.58 mmol) was dissolved in 0.04 M $K_2CO_3$ in MeOH (21.8 mL, 0.87 mmol) and stirred under an argon atmosphere at rt for 2 h. The reaction was diluted with DCM (200 mL) and washed with $NaHCO_3$ (3 x 100 mL). The organic layers were dried ($Na_2SO_4$) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (4:96 EtOH/$CHCl_3$ with 1 % pyridine) to afford the product **7** (0.35 g, 0.53 mmol, 91 %) as a dark yellow foam.

$R_f$ : 0.51 (EtOH/$CHCl_3$ 1:9)

LRMS: [ESI+, MeOH] m/z (%): 689 ((M($^{37}$Cl)+Na)$^+$, 3), 687 ((M($^{35}$Cl)+Na)$^+$, 19).

HRMS: [ESI+, MeOH] for $C_{38}H_{34}Cl_1N_3Na_1O_6$ (M+Na)$^+$: calcd 686.2028, found 686.2033.

$^1$H (300 MHz, $CDCl_3$): δ 11.60 (br s, 1H, NH), 8.81 (s, 1H, $H^6$),7.50-7.08 (m, 13H, $H^{Ar}$), 6.74 (dd, *J*=9.0, 3.0, 4H, $H^{Ar}$), 6.56 (t, *J*=5.5, 1H, $H^{1'}$), 5.56 (s, 1H, $H^{17}$), 4.62 (ddd, *J*= 5.3, 5.3, 5.3, 1H, $H^{3'}$), 4.17 (d, *J*= 4.2, 1H, $H^{4'}$), 3.61 (d, *J*=6.0, 6H, $H^{12}$), 3.51-3.39 (m, 2H, $H^{5'}$), 3.04 (m, 1H, $H^{2'}$), 2.34 (m, 1H, $H^{2'}$) ppm.

$^{13}$C (100 MHz, $CDCl_3$): δ 160.3 ($C^2$), 159.1 ($C^{Ar}$), 155.4 ($C^4$), 144.7 ($C^{Ar}$), 139.8 ($C^{18}$), 136.8 ($CH^6$), 136.2 ($C^{Ar}$), 136.0 ($C^{Ar}$), 134.5 ($C^{Ar}$), 130.6 ($CH^{Ar}$), 130.5 ($CH^{Ar}$), 129.3 ($CH^{Ar}$), 128.8 ($C^{Ar}$), 128.7 ($CH^{Ar}$), 128.5 ($CH^{Ar}$), 127.1 ($CH^{Ar}$), 127.0 ($CH^{Ar}$), 113.8 ($CH^{Ar}$), 111.0 ($C^5$), 97.5 ($CH^{17}$), 88.5 ($CH^{1'}$),

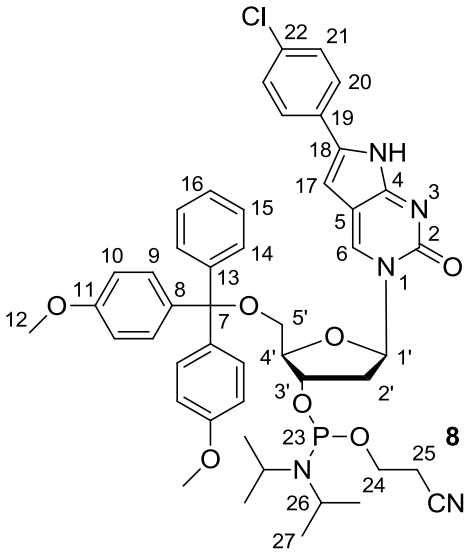
87.5 ($C^{7}$), 86.8 ($CH^{4'}$), 71.1 ($CH^{3'}$), 63.1 ($CH_2^{5'}$), 55.6 ($CH_3^{12}$), 43.0 ($CH_2^{2'}$) ppm.

Mp: 150 ºC (decomposes)

IR $\nu_{Max}/cm^{-1}$: 3065 (w, 2º amine N-H), 2930, 2836 (w, alkyl C-H), 1659 (s, amide C=O).

UV/Vis (MeOH): $A_{max}$ 370 nm, ε 14,000 $M^{-1}cm^{-1}$, $Em_{max}$ 445 nm.

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-6-(4-chloropheynyl)-3 *H*-pyrrolo-[2,3-*d*]pyrimidine-2-one-3'-*O*-(2-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite (8)**
**(*p*-chlorophenylpyrrolopyrimidine-2'-deoxyriboside, CPP)**




To a solution of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-6-(4-chloropheynyl)-3 *H*-pyrrolo-[2,3-*d*]pyrimidine-2-one (**7**) (0.29 g, 0.44 mmol) in distilled DCM (5 mL), with activated molecular sieves (3 Å), under an argon atmosphere, was added distilled DIPEA (0.026 mL, 1.18 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.126 mL, 0.57 mmol) and the reaction was stirred at rt for 1 hour. The reaction mixture was diluted with distilled DCM (20 mL) and was washed with deoxygenated sat. KCl (20 mL). The organic phase was dried ($Na_2SO_4$) and the solvent was evaporated *in vacuo*. Following purification by column chromatography (dry and degassed silica, 1:9 hexane/EtOAc with 1 % pyridine) the product **8** (0.10 g, 0.12 mmol, 26 %) was afforded as a yellow solid.

$R_f$ : 0.47, 0.34 (EtOAc/$Et_3N$ 99:1)

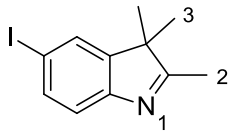
LRMS: [ESI+, MeCN] m/z (%): 1752 ((2M($^{35}$Cl)($^{37}$Cl)+Na)$^+$, 19), 889 ((M($^{37}$Cl)+Na)$^+$, 50), 887 ((M($^{35}$Cl)+Na)$^+$, 100).

HRMS: [ESI+, MeOH] for $C_{47}H_{51}Cl_1N_5Na_1O_7P_1$ $(M+Na)^+$: calcd 886.3107, found 886.3109.

$^{31}$P (300 MHz, $CDCl_3$): δ 149.6 (s, P), 149.4 (s, P) ppm.

$^1$H (300 MHz, $CDCl_3$): δ 10.26 (s, 1H, NH), 8.75 (s, 1H, $H^6$), 7.65-7.26 (m, 13H, $H^{Ar}$), 6.87 (dd, *J*=9.2, 2.9, 4H, $H^{Ar}$), 6.24 (t, *J*=4.4, 1H, $H^{1'}$), 5.75 (s, 1H, $H^{17}$), 4.78 (m, 1H, $H^{3'}$), 4.18 (m, 1H, $H^{4'}$), 3.71 (d, *J*=3.3, 6H, $H^{12}$), 3.70-3.58 (m, 4H, $H^{25,26}$), 3.50-3.47 (m, 2H, $H^{5'}$), 2.64 (m, 1H, $H^{2'}$), 2.54 (t, *J*=5.9, 2H, $H^{24}$), 2.45 (m, 1H, $H^{2'}$), 1.20-1.15 (m, 12H, $H^{27}$) ppm.

**Synthesis of 5-iodo-2,3,3-trimethyl-3*H*-indole (9)[114]**




**9**

To a solution of 4-iodophenylhydrazine (1.51 g, 6.46 mmol) in AcOH (120 mL) was added 3-methyl-2-butanone (0.75 mL, 7.05 mmol) and the reaction was stirred at reflux for 2.5 h. The reaction mixture was allowed to cool to rt, diluted with $H_2O$ (300 mL) and washed with DCM (2 x 200 mL). The organic phase was collected and concentrated *in vacuo*. The oily residue was diluted with DCM (16 mL). The product was extracted by addition of hexane (300 mL) which was decanted from the resultant purple oily precipitate. The yellow hexane layer was dried *in vacuo* to afford the product **9** (1.48 g, 5.19 mmol, 81 %) as an orange oil.

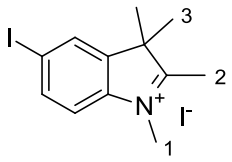
$R_f$ : 0.41 (DCM/$Et_3N$ 99:1)

LRMS [ESI+, MeOH] m/z (%): 286 ((M + H)$^+$, 100).

$^1$H (300 MHz, $CDCl_3$) δ: 7.63-7.59 (m, 2H, $H^{Ar}$), 7.31 (t, *J*=8.0, 1H, $H^{Ar}$), 2.26 (s, 3H, $H^2$), 1.29 (s, 6H, $H^3$) ppm.

Characterisation data recorded matches previously reported literature values.[114]

**Synthesis of 5-iodo-1,2,3,3-tetramethyl-3*H*-indolium iodide (10)[112,115]**




**10**

To a solution of 5-iodo-2,3,3-trimethyl-3*H*-indole (**9)** (1.42 g, 4.99 mmol) in MeCN (10 mL) was added MeI (0.34 mL, 5.44 mmol) and the reaction was stirred under reflux for 16 h. The reaction mixture was allowed to cool to rt and the orange precipitate was collected by filtration, washed with $Et_2O$ and the solvent removed *in vacuo* to afford the product **10** (1.67 g, 3.91 mmol, 78 %) as a light orange powder.

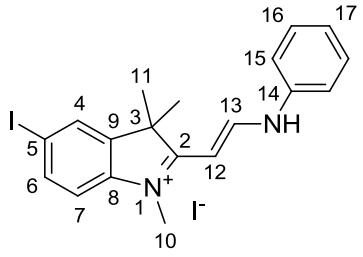
$R_f$ : 0.74 (MeOH/DCM 1:9)

LRMS [ESI+, MeOH] m/z (%): 300 ($M^+$, 100).

$^1$H (300 MHz, $d_6$-DMSO) δ: 8.28 (s, 1H, $H^{Ar}$), 8.00 (d, *J*=8.4, 1H, $H^{Ar}$), 7.72 (d, *J*=8.4, 1H, $H^{Ar}$), 3.93 (s, 3H, $H^1$), 2.73 (s, 3H, $H^2$), 1.52 (s, 6H, $H^3$) ppm.

Characterisation data recorded matches previously reported literature values.[112,115]

**Synthesis of 2-[(E)-2-(phenylamino)vinyl]-5-iodo-1,3,3-trimethyl-3*H*-indolium iodide (11)** (experimental method based on literature[113])




**11**

To a mixture of 5-iodo-1,2,3,3-tetramethyl-3*H*-indolium iodide (**10**) (1.10g, 2.58 mmol) and N,N'-diphenylformamidine (0.51g, 2.58 mmol) in ethanol (10 mL) was added triethylorthoformate (0.43 mL, 2.58 mmol) and the reaction was heated under reflux for 2.5 h.

The mixture was allowed to cool to rt, the solvent removed *in vacuo* and the residue resuspended in ice-cold $CHCl_3$. The resulting orange precipitate was filtered and washed with ice-cold $CHCl_3$. A further 2 isolations of precipitate were made from the filtrate by repeating the procedure, affording the product **11** (1.30 g, 2.45 mmol, 95 %) as an orange powder.

$R_f$ : 0.31 (MeOH/DCM 5:95)

LRMS [ESI+, MeOH] m/z (%): 403 ($M^+$, 100).

HRMS: [ESI+, MeCN] for $C_{19}H_{20}IN_2$ $(M)^+$: calcd 403.0666, found 403.0659.

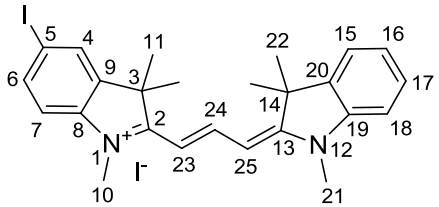
$^1$H (400 MHz, $d_6$-DMSO) δ: 11.99 (s, 1H, NH), 8.67 (d, *J*=12.0, 1H, $H^{13}$), 8.08 (d, *J*=1, 1H, $H^4$), 7.80 (dd, *J*=8.0, 1.0, 1H, $H^6$), 7.54-7.48 (m, 4H, $H^{15/16/17}$), 7.36 (d, *J*=8.0, 1H, $H^7$), 7.30 (t, *J*=7.0, 1H, $H^{15/16/17}$), 6.09 (d, *J*=12.0, 1H, $H^{12}$), 3.63 (s, 3H, $H^{10}$), 1.68 (s, 6H, $H^{11}$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO) δ: 177.5 ($C^2$), 151.9 ($CH^{13}$), 143.2 ($C^9$), 142.2 ($C^8$), 138.2 ($C^{14}$), 136.9 ($CH^6$), 131.1 ($CH^4$), 129.8 ($CH^{15/16/17}$), 126.2 ($CH^{15/16/17}$), 118.3 ($CH^{15/16/17}$), 113.9 ($CH^7$), 91.1 ($CH^{12}$), 90.1 ($C^5$), 49.5 ($C^3$), 32.3 ($CH^{10}$), 27.2 ($CH^{11}$) ppm.

Mp: 190°C (decomposes).

IR $\nu_{Max}/cm^{-1}$: 3129 (w-m) (N-H), 2944, 2875 (m) (C-H, C-$H_3$), 1628 (s) (C=N).

**Synthesis of 5-iodo-1,3,3-trimethyl-2-((1E,3E)-3-(1,3,3-trimethylindolin-2-ylidene)prop-1-enyl)-3*H*-indolium iodide (5-iodo-Cy3) (12)[112]**




**12**

To a mixture of 2-[(E)-2-(phenylamino)vinyl]-5-iodo-1,3,3-trimethyl-3*H*-indolium iodide (**11**) (1.88g, 3.54 mmol) and 1,2,3,3-tetramethyl-3*H*-indolium iodide (1.07 g, 3.54 mmol) in distilled pyridine (20 mL) was added $Ac_2O$ (0.67 mL, 7.09 mmol) and the reaction was stirred at 50 °C for 24 h. Additional portions of 1,2,3,3-tetramethyl-3*H*-indolium iodide (0.426g, 1.42 mmol) and $Ac_2O$ (1.11 mL, 11.9 mmol) were added to the reaction and it

was stirred at 50 °C for a further 24 h. The mixture was allowed to cool to rt and the addition of $Et_2O$ (20 mL) afforded a brown precipitate. The precipitate was isolated by filtration and following purification by column chromatography (DCM → 5 % MeOH/DCM) the product **12** (1.31 g, 2.15 mmol, 61 %) was afforded as a dark purple solid.

$R_f$ : 0.29 (MeOH/DCM 5:95)

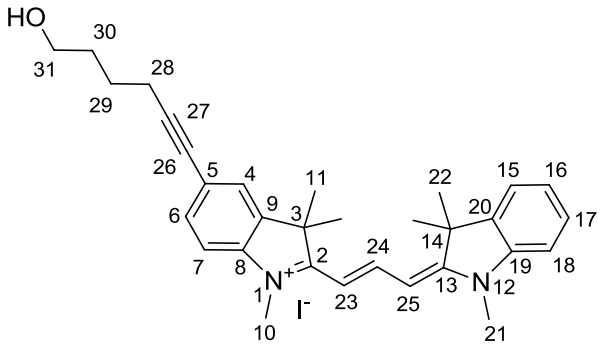
LRMS [ESI+, MeOH] m/z (%): 483 ($M^+$, 100).

$^1$H (400 MHz, $CD_3CN$) δ: 8.43 (t, *J*=13.6, 1H, $H^{24}$), 7.85 (d, *J*=1.9, 1H, $H^{Ar}$), 7.77 (dd, *J*=8.3, 1.5, 1H, $H^{Ar}$), 7.53 (d, *J*=7.5, 1H, $H^{Ar}$), 7.46 (m, 1H, $H^{Ar}$), 7.33 (t, *J*=7.6, 2H, $H^{Ar}$), 7.05 (d, *J*=8.5, 1H, $H^{Ar}$), 6.39 (d, *J*=13.6, 1H, $H^{23}$), 6.29 (d, *J*=13.1, 1H, $H^{25}$), 3.61 (s, 3H, $H^{10}$), 3.52 (s, 3H, $H^{21}$), 1.69 (s, 6H, $H^{11}$), 1.67 (s, 6H, $H^{22}$) ppm.

UV/Vis (MeOH): $A_{max}$ 551 nm, ε 114,000 $M^{-1}cm^{-1}$, $Em_{max}$ 568 nm.

Characterisation data recorded matches previously reported literature values.[112]

**Synthesis of 5-hexyn-1-ol-6-Cy3 (13)**




**13**

To a mixture of 5-iodo-1,3,3-trimethyl-2-((1E,3E)-3-(1,3,3-trimethylindolin-2-ylidene)prop-1-enyl)-3*H*-indolium iodide (**12**) (0.35 g, 0.57 mmol) and CuI (0.02 g, 0.12 mmol) in anhydrous DMF (2.90 mL) under an argon atmosphere was added distilled $Et_3N$ (2 mL) and 5-hexyn-1-ol (0.07 mL, 0.63 mmol) and the mixture was stirred for 10 min at rt. $Pd(PPh_3)_4$ (0.07g, 0.06 mmol) was added in one portion and the reaction was stirred in the dark for 1.5 h at rt. The reaction solvent was removed *in vacuo* and following

two purifications by column chromatography ([1] 100 % DCM → 2 % MeOH/DCM, [2] 7:3 EtOAc/EtOH → 1:9 MeOH/DCM) the product **13** (0.26 g, 0.45 mmol, 80 %) was afforded as a purple/pink iridescent solid.

$R_f$ : 0.46 (MeOH/DCM 1:9)

LRMS [ESI+, MeCN] m/z (%): 453 ($M^+$, 100).

HRMS [ESI+, MeCN] for $C_{31}H_{37}N_2O_1$ $(M)^+$: calcd 453.2900, found 453.2896.

$^1$H (300 MHz, $CDCl_3$): 8.40 (t, *J*=13.4, 1H, $H^{24}$), 7.44-7.23 (m, 7H, $H^{23,25,Ar}$), 7.15 (d, *J*=7.9, 1H, $H^4$), 7.02 (d, *J*=8.3, 1H, $H^7$), 3.83 (s, 3H, $H^{10}$), 3.78 (s, 3H, $H^{21}$), 3.73 (t, *J*=6.0, 2H, $H^{31}$), 2.48 (t, *J*=6.6, 2H, $H^{28}$), 1.79-1.73 (m, 4H, $H^{29,30}$), 1.71 (2s, 12H, $H^{11,22}$) ppm.

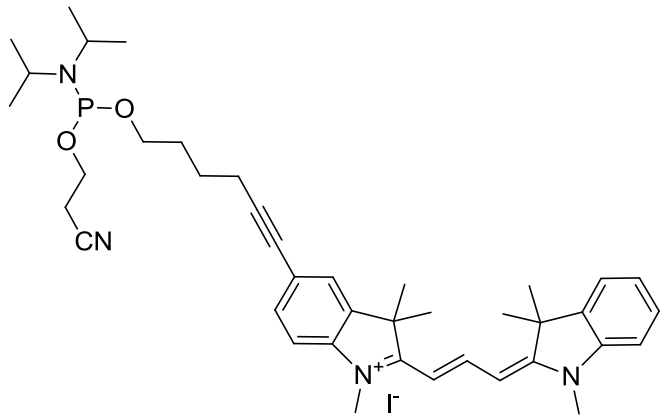
$^{13}$C (75 MHz, $CDCl_3$): 174.9 ($C^2$), 173.9 ($C^{13}$), 150.9 ($CH^{24}$), 142.9 ($C^{8,13}$), 142.3 ($C^{8,13}$), 140.9 ($C^{9,20}$), 140.7 ($C^{9,20}$), 132.8 ($CH^{Ar}$), 129.2 ($CH^{Ar}$), 125.8 ($CH^{Ar}$), 125.4 ($CH^{Ar}$), 122.3 ($CH^{Ar}$), 121.1 ($C^5$), 111.3 ($CH^4$), 110.8 ($CH^7$), 106.1 ($CH^{23,25}$), 105.7 ($C^{23,25}$), 91.5 ($C^{26,27}$), 80.6 ($C^{26,27}$), 62.7 ($CH_2{}^{31}$), 49.2 ($C^{3,14}$), 48.8 ($C^{3,14}$), 33.3 ($CH_3{}^{10,21}$), 32.2 ($CH_2{}^{29,30}$), 28.4 ($CH_3{}^{11,22}$), 28.4 ($CH_3{}^{11,22}$), 25.3 ($CH_3{}^{29,30}$), 19.6 ($CH_2{}^{28}$) ppm.

Mp: 108-110°C.

IR $\nu_{Max}/cm^{-1}$: 3357 (br, m. O-H), 2925, 2859 (m. C-H, C-$H_3$).

UV/Vis (MeOH): $A_{max}$ 558 nm, $Em_{max}$ 580 nm.

**Synthesis of 5-hexyn-1-ol-6-Cy3 phosphoramidite (14)**



**14**

5-Hexyn-1-ol-6-Cy3 (**13**) (0.195 g, 0.34 mmol) was co-evaporated with distilled pyridine (3 x 5 mL) followed by distilled DCM (5 x 5 mL) before suspending in distilled DCM (5 mL), with activated molecular sieves (3 Å), under an argon atmosphere. To this solution was added distilled DIPEA (0.15 mL, 0.84 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.07 mL, 0.34 mmol) and the reaction was stirred at rt for 45 min. An additional 0.2 equivalents 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite was added and the reaction stirred for a further 2.5 h. The reaction was diluted with distilled DCM (10 mL) and washed with deoxygenated sat. KCl (10 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. The crude product was precipitated in deoxygenated hexane (50 mL) from a minimum of distilled DCM (2 mL). The resulting precipitate was washed with deoxygenated hexane (2 x 50 mL). The precipitate was dried *in vacuo* to afford the product **14** (0.26 g, 0.33 mmol, 96 %) as a dark pink solid.

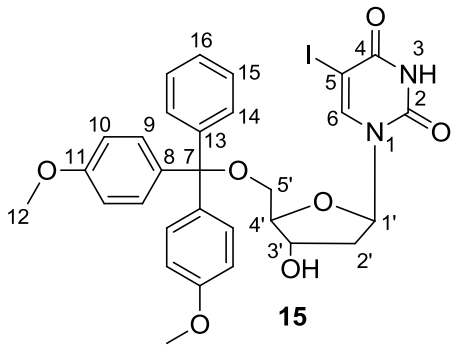
$R_f$ : 0.49 (MeOH/DCM 1:9 with 1 % pyridine)

LRMS [ESI+, MeCN] m/z (%): 653 ($M^+$, 100).

HRMS [ESI+, MeCN] for $C_{40}H_{54}N_4O_2P_1$ $(M)^+$: calcd 653.3979, found 653.3966.

$^{31}P$ (121 MHz, $CD_3CN$): 148.4 (s, P).

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyluridine (15)[169]**




**15**

5-iodo-2'-deoxyuridine (10.2 g, 28.9 mmol) was co-evaporated with distilled pyridine (3 x 10 mL) and dissolved in distilled pyridine (80 mL). To this was added drop-wise a solution of DMTCl (9.79 g, 28.9 mmol) in distilled pyridine (50 mL) over a period of 20 min and the reaction was stirred at rt for 3 h. An additional 0.2 eq DMTCl was added and the reaction stirred for a further 2.5 h. The reaction was quenched by the addition of MeOH (80 mL) and was stirred for 20 min. The reaction volume was reduced by half *in vacuo*, diluted with DCM (250 mL) and washed with $H_2O$ (250 mL) and $NaHCO_3$ (2 x 200 mL) then the solvent was removed *in vacuo*. Following purification by column chromatography (DCM with 0.5 % pyridine → 4:96 MeOH/DCM with 0.5 % pyridine) the product **15** (16.6 g, 25.3 mmol, 88 %) was afforded as a beige foam

$R_f$ : 0.44 (MeOH/DCM 1:9)

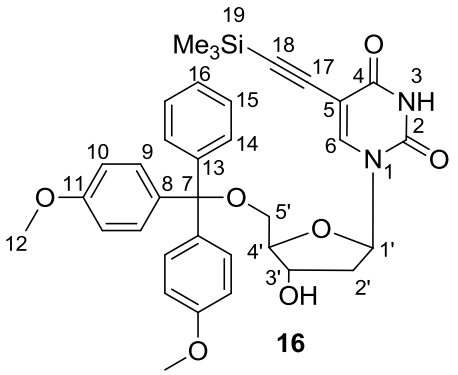
LRMS: [ESI+, MeOH] m/z (%): 679 ($(M+Na)^+$, 100).

$^1$H (300 MHz, $d_6$-DMSO): δ 11.84 (s, 1H, NH), 8.13 (s, 1H, $H^6$), 7.53-7.33 (m, 9H, $H^{Ar}$), 7.02 (d, *J*=12.0, 4H, $H^{10}$), 6.23 (t, *J*=6.7, 1H, $H^{1'}$), 5.42 (d, *J*=4.3, 1H, $OH^{3'}$), 4.35 (m, 1H, $H^{3'}$), 4.03 (m, 1H, $H^{4'}$), 3.86 (s, 6H, $H^{12}$), 3.30-3.27 (m, 2H, $H^{5'}$), 2.37-2.30 (m, 2H, $H^{2'}$) ppm.

$^{13}$C (75 MHz, $d_6$-DMSO): δ 160.2 ($C^2$), 157.8 ($C^4$), 149.8 ($CH^6$), 149.3 ($C^{Ar}$), 144.5 ($CH^{Ar}$), 144.0 ($C^{Ar}$), 135.8 ($CH^{Ar}$), 135.2 ($C^{Ar}$), 135.1 ($C^{Ar}$), 129.4 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 127.4 ($CH^{Ar}$), 126.4 ($CH^{Ar}$), 123.6 ($CH^{Ar}$), 113.0 ($CH^{10}$), 85.6 ($CH^{4'}$), 84.62 ($CH^{1',7}$), 70.2 ($CH^{3'}$), 69.5 ($CH^5$), 63.4 ($CH_2{}^{5'}$), 54.8 ($CH_3{}^{12}$), 39.9 ($CH_2{}^{2'}$) ppm.

Characterisation data recorded matches previously reported literature values.[169]

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(trimethylsilyl eth-1-ynyl)uridine (16)**




**16**

To a mixture of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyluridine (**15**) (1.19 g, 1.81 mmol) and CuI (0.07 g, 0.38 mmol) in anhydrous DMF (4.5 mL) under an argon atmosphere, was added trimethylsilylacetylene (0.77 mL, 5.42 mmol) and $Et_3N$ (0.75 mL, 5.42 mmol) and the reaction was stirred for 10 min at rt. $Pd(PPh_3)_4$ (0.21 g, 0.18 mmol) was added and the reaction stirred at rt for 16 h. The solvent was removed *in vacuo* and purification by column chromatography (4:96 MeOH/DCM with 1 % pyridine) afforded the product **16** (1.06 g, 1.69 mmol, 94 %) as a light brown foam.

$R_f$ : 0.46 (EtOAc/MeOH/aq. $NH_3$ 5:1:1)

LRMS [ESI+, MeOH] m/z (%): 649 ($(M+Na)^+$, 78).

HRMS [ESI+, MeOH] for $C_{35}H_{38}N_2O_7SiNa$ $(M+Na)^+$: calcd 649.2340, found 649.2343.

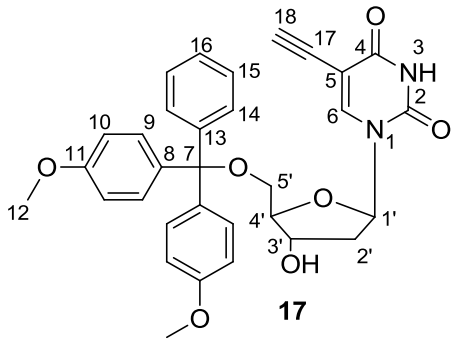
$^1H$ (300 MHz, $CDCl_3$): δ 7.99 (s, 1H, $H^6$), 4.45-7.18 (m, 9H, $H^{Ar}$), 6.83 (d, *J*= 8.8, 4H, $H^{Ar}$), 6.26 (dd, *J*= 7.7, 5.7, 1H, $H^{1'}$), 4.45 (dt, *J*=6.0, 2.0, 1H, $H^{3'}$), 4.08 (dt, *J*=3.0, 2.0, 1H, $H^{4'}$), 3.77 (s, 6H, $H^{12}$), 3.40 (dd, *J*= 11.0, 4.0, 1H, $H^{5'}$), 3.30 (dd, *J*=11.0, 4.0, 1H, $H^{5'}$), 2.48 (ddd, *J*=13.6, 5.6, 2.4, 1H, $H^{2'}$), 2.20 (ddd, *J*=13.7, 7.6, 6.1, 1H, $H^{2'}$), 0.00 (s, 9H, $H^{19}$) ppm.

$^{13}C$ (75 MHz, $CDCl_3$): δ 160.8 ($C^2$), 158.2 ($C^{Ar}$), 148.8 ($C^4$), 144.1 ($C^{Ar}$), 142.3 ($CH^6$), 135.2 ($C^{Ar}$), 129.6 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 127.6 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 113.0 ($CH^{Ar}$), 100.2 ($C^{17}$), 99.4 ($C^{18}$), 94.4 ($C^5$), 86.6 ($C^7$), 86.1 ($CH^{4'}$), 85.4 ($CH^{1'}$), 72.0 ($CH^{3'}$), 63.1 ($CH_2^{5'}$), 54.9 ($CH_3^{12}$), 45.9 ($CH_2^{2'}$), 0.7 ($CH_3^{19}$) ppm.

Mp: 86-88 °C (decomposes)

IR $\nu_{Max}/cm^{-1}$: 3360 (br, O-H alcohol), 3050 (w, C-H aromatic), 2953, 2835 (m, C-H alkyl), 2161 (m, C≡C alkyne), 1682 (s, C=O carbonyl).

**Synthesis of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(eth-1-ynyl)uridine (17)**




**17**

To a solution of 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(trimethylsilyl eth-1-ynyl)uridine (**16**) (0.11 g, 0.17 mmol) in distilled THF (2 mL) under an argon atmosphere was added 1M TBAF in THF (0.26 mL, 0.25 mmol) and the reaction was stirred at rt for 20 min. The reaction was partitioned between EtOAc (100 mL) and brine (100 mL). The organic phase was washed with brine and dried ($Na_2SO_4$). The solvent was removed *in vacuo* and the product **17** (0.093 g, 0.17 mmol, 99 %) was afforded as a light yellow/brown foam.

$R_f$ : 0.38 (MeOH/DCM 7.5:92.5)

LRMS [ESI+, MeOH] m/z (%): 577 ($(M+Na)^+$, 100).

HRMS [ESI+, MeOH] for $C_{32}H_{30}N_2O_7Na$ $(M+Na)^+$: calcd 577.1945, found 577.1951.

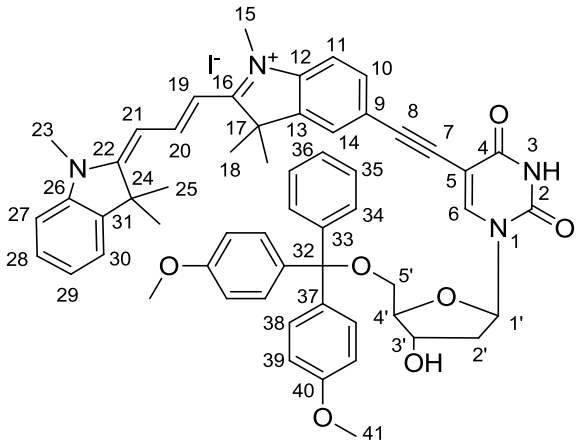
$^1H$ (300 MHz, $CDCl_3$): δ 8.00 (s, 1H, $H^6$), 7.36-7.12 (m, 9H, $H^{Ar}$), 6.77 (d, *J*=8.3, 4H, $H^{Ar}$), 6.20 (t, *J*=6.6, 6.6, 1H, $H^{1'}$), 4.47 (m, 1H, $H^{3'}$), 4.05 (m, 1H, $H^{4'}$), 3.71 (s, 6H, $H^{12}$), 3.35-3.27 (m, 2H, $H^{5'}$), 2.83 (s, 1H, $H^{18}$), 2.45 (ddd, *J*=13.6, 5.6, 2.8, 1H, $H^{2'}$), 2.20 (ddd, *J*=13.6, 6.7, 6.7, 1H, $H^{2'}$) ppm.

$^{13}C$ (75 MHz, $CDCl_3$): δ 161.4 ($C^2$), 158.6 ($C^{Ar}$), 149.1 ($C^4$), 144.4 ($C^{Ar}$), 143.6 ($CH^6$), 135.5 ($C^{Ar}$), 135.3 ($C^{Ar}$), 130.0 ($CH^{Ar}$), 128.0 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 127.0 ($CH^{Ar}$), 113.4 ($CH^{Ar}$), 99.2 ($C^{17}$), 87.1 ($C^5$), 86.5 ($CH^{4'}$), 85.8 ($CH^{1'}$), 82.0 ($CH^{18}$), 74.1 ($C^7$), 72.2 ($CH^{3'}$), 63.4 ($CH_2{}^{5'}$), 55.2 ($CH_3{}^{12}$), 41.5 ($CH_2{}^{2'}$) ppm.

Mp: 95-96 °C (decomposes)

IR $\nu_{Max}/cm^{-1}$: 3425 (br,w, O-H alcohol), 3272 (m, H-C alkyne), 2930, 2836 (m, C-H alkyl), 2162 (w, C≡C alkyne), 1682 (s, C=O carbonyl).

**Synthesis of Cy3dT (18)**



**18**

To a mixture of 5-Iodo-Cy3 (**12**) (1.24 g, 2.03 mmol), 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(eth-1-ynyl)uridine (**17**) (1.24 g, 2.23 mmol) and CuI (0.08 g, 0.41 mmol) in anhydrous DMF (10 mL) under an argon atmosphere was added distilled $Et_3N$ (7 mL) and the mixture stirred for 10 min at rt. $Pd(PPh_3)_4$ (0.24 g, 0.20 mmol) was added and the reaction stirred at rt in the dark for 24 h. The reaction solvent was removed *in vacuo*. Following purification by column chromatography ([1] 0→5 % MeOH/DCM with 1 % pyridine. [2] 10→20 % EtOH/EtOAc with 1 % pyridine. [3] 0→10 % MeOH/DCM with 1 %pyridine) the product **18** (0.88 g, 0.85 mmol, 42 %) was afforded as a purple/gold iridescent solid.

$R_f$ : 0.55 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 910 ($M^+$, 100).

HRMS [ESI+, MeCN] for $C_{57}H_{57}N_4O_7$ $(M)^+$: calcd 909.4222, found 909.4237.

$^1$H (300 MHz, $CD_3CN$) δ: 8.41 (t, *J*=13.6, 1H, $H^{20}$), 8.20 (s, 1H, $H^6$), 7.55-7.09 (m, 16H, $H^{Ar}$), 6.82 (t, *J*=8.5, 4H, $H^{Ar}$), 6.43 (d, *J*=13.6, 1H, $H^{19}$), 6.34 (d, *J*=13.6, 1H, $H^{21}$), 6.19 (t, *J*=6.5, 1H, $H^{1'}$), 4.54 (m, 1H, $H^{3'}$), 4.03 (d, *J*=2.9, 1H, $H^{4'}$), 3.69 (s, 3H, $H^{41}$), 3.68 (s, 3H, $H^{41}$),3.62 (s, 3H, $H^{15}$), 3.53 (s, 3H, $H^{23}$), 3.31 (d, *J*=3.0, 2H, $H^{5'}$), 2.39-2.34 (m, 2H, $H^{2'}$), 1.71 (s, 6H, $H^{18/25}$), 1.61 (s, 3H, $H^{18/25}$), 1.60 (s, 3H, $H^{18/25}$) ppm.

$^{13}$C (75 MHz, $CD_3CN$) δ: 177.3 ($C^{16/22}$), 175.8 ($C^{16/22}$), 162.8 ($C^2$), 160.1 ($C^{40}$), 151.8 ($CH^{20}$), 151.1 ($C^4$), 146.3 ($CH^6$), 144.3 ($C^{Ar}$), 142.5 ($C^{Ar}$), 142.2 ($C^{Ar}$), 137.4 ($C^{Ar}$), 137.2($C^{Ar}$), 133.6 ($CH^{Ar}$), 131.5 ($CH^{38}$), 131.4 ($CH^{38}$), 130.2 ($CH^{Ar}$), 129.4 ($CH^{Ar}$), 127.2 ($CH^{Ar}$), 126.3 ($CH^{Ar}$), 125.2 ($CH^{Ar}$), 123.7

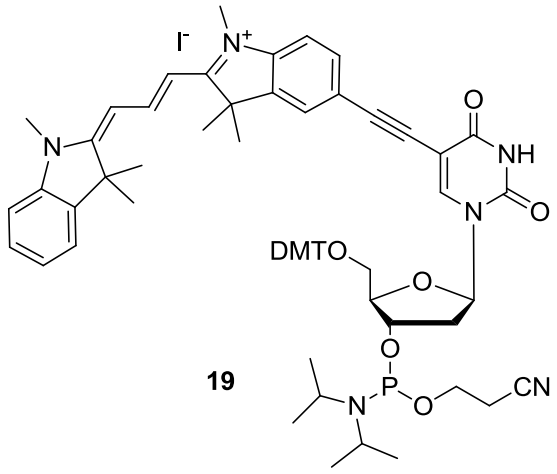
($CH^{Ar}$), 120.4 ($C^{Ar}$), 114.7 ($CH^{39}$), 112.9 ($C^{Ar}$), 112.2 ($C^{Ar}$), 104.9 ($CH^{19}$), 103.9 ($CH^{21}$), 100.4 ($C^{5}$), 93.4 ($C^{8,32}$), 88.1 ($CH^{4'}$), 87.3 ($CH^{1'}$), 83.3 ($C^{7}$), 72.4 ($CH^{3'}$), 64.7 ($CH_2{}^{5'}$), 56.4 ($CH_3{}^{41}$), 50.9 ($C^{17,24}$), 50.2 ($C^{17,24}$), 42.5 ($CH_2{}^{2'}$), 33.0 ($CH_3{}^{15,23}$), 28.5 ($CH_3{}^{18,25}$), 28.4 ($CH_3{}^{18,25}$) ppm.

Mp: >130°C (decomposes).

IR $\nu_{Max}/cm^{-1}$: 3306 (br. w. O-H), 3040 (w. N-H), 2930, 2872 (w. C-H, C-$H_3$), 1704 (s. C=O).

UV/Vis (MeOH): $A_{max}$ 565 nm, ε 112,000 $M^{-1}cm^{-1}$, $Em_{max}$ 592 nm.

**Synthesis of Cy3dT-phosphoramidite (19)**




**19**

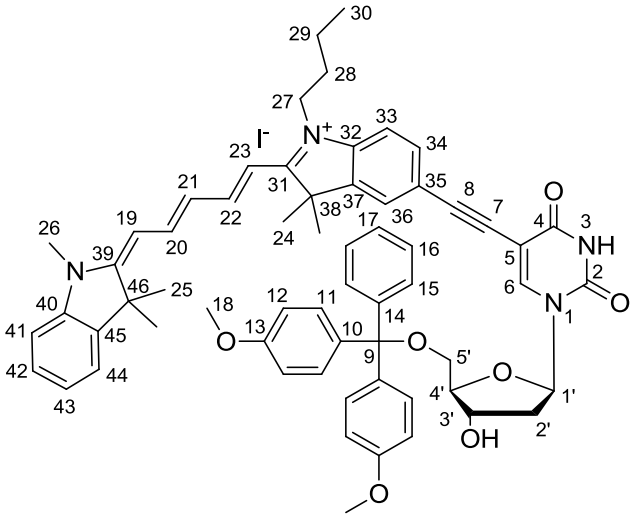
Cy3dT (**18**) (0.63 g, 0.61 mmol) was co-evaporated with distilled pyridine (3 x 5 mL) and distilled DCM (5 x 5 mL) before suspending in distilled DCM (10 mL), with activated molecular sieves (3 Å), under an argon atmosphere. To this was added distilled DIPEA (0.27 mL, 1.53 mmol) followed by the drop-wise addition of 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.14 mL, 0.61 mmol) and the reaction was stirred at rt for 45 min. An additional portion of 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.1 eq) was added and the reaction stirred for a further 2 h. The solution was diluted with distilled DCM (10 mL) and washed with deoxygenated sat. KCl (20 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo* before purification by precipitation in deoxygenated hexane (250 mL) from a minimum of DCM (2 mL). The precipitation process was repeated twice. The precipitate was dried *in vacuo* to afford the product **19** (0.73 g, 0.59 mmol, 96 %) as a dark pink solid.

$R_f$ : 0.39 (1:9 MeOH/DCM + 1 % pyridine)

LRMS [ESI+, MeCN] m/z (%): 1110 ($M^+$, 100).

$^{31}P$ (121 MHz, $CD_3CN$) δ: 149.3 (s, P), 149.2 (s, P) ppm.

**Synthesis of Cy5dT (20)**



**20**

To a mixture of iodo-Cy5 (prepared by Marta Gerowska, 1.00 g, 1.47 mmol), 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(eth-1-ynyl)uridine (**17**) (1.50 g, 2.70 mmol) and CuI (0.06 g, 0.29 mmol) in anhydrous DMF (7.5 mL) under an argon atmosphere was added distilled $Et_3N$ (5.2 mL) and the mixture stirred at rt for 10 min. $Pd(PPh_3)_4$ (0.17 g, 0.15 mmol) was added and the reaction stirred at rt in the dark for 24 h. Additional CuI (0.2 eq) and $Pd(PPh_3)_4$ (0.1 eq) were added and the reaction stirred at rt for a further 2 h. The reaction solvent was removed *in vacuo* and the product purified by column chromatography ([1] 0→20 % MeOH/DCM with 1 % pyridine. [2] 0→30 % EtOH/EtOAc with 1 % pyridine) to afford the product **20** (1.49 g, 1.35 mmol, 91 %) as a blue/red iridescent solid.

$R_f$ : 0.39 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 978 ($M^+$, 100).

HRMS [ESI+, MeCN] for $C_{62}H_{65}N_4O_7$ $(M)^+$: calcd 977.4848, found 977.4859.

$^{1}$H (400 MHz, $CD_3CN$) δ: 8.19 (s, 1H, $H^{6}$), 8.13 (t, *J*=13.8, 1H, $H^{20/22}$), 8.05 (t, *J*=13.1, 1H, $H^{20/22}$), 7.52-7.00 (m, 16H, $H^{Ar}$), 6.81 (dd, *J*=7.04, 7.04, 4H, $H^{12}$), 6.57 (t, *J*=12.5, 1H, $H^{21}$), 6.27 (d, *J*=14.0, 1H, $H^{23}$), 6.20-6.14 (m, 2H, $H^{1',19}$), 4.53 (ddd, *J*=5.5, 3.2, 3.2, 1H, $H^{3'}$), 4.03 (m, 1H, $H^{4'}$), 3.93 (t, *J*=7.5, 2H, $H^{27}$), 3.67 (s, 3H, $H^{18}$), 3.66 (s, 3H, $H^{18}$), 3.59 (s, 3H, $H^{26}$), 3.30 (d, *J*=3.0, 2H, $H^{5'}$), 2.38-2.31 (m, 2H, $H^{2'}$), 1.74-1.61 (m, 8H, $H^{25,28}$), 1.58 (s, 3H, $H^{24}$), 1.57 (s, 3H, $H^{24}$), 1.45-1.40 (m, 2H, $H^{29}$), 0.99-0.94 (m, 3H, $H^{30}$) ppm.

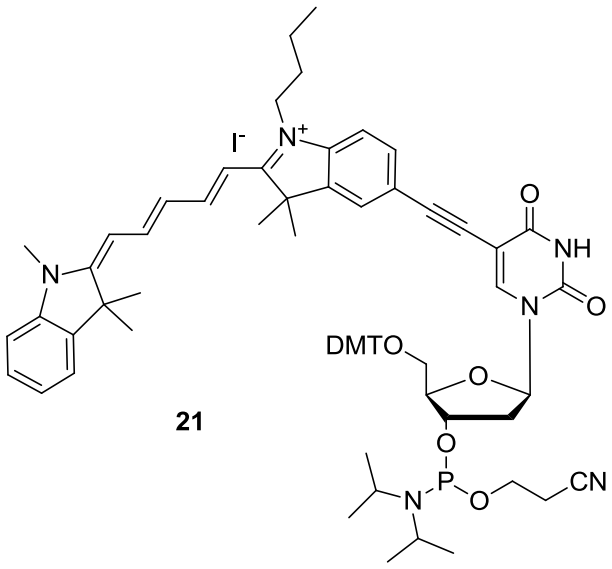
$^{13}$C (100 MHz, $CD_3CN$) δ: 176.2 ($C^{39}$), 172.7 ($C^{31}$), 162.4 ($C^{2/4}$), 159.6 ($C^{13}$), 155.4 ($CH^{20/22}$), 154.0 ($CH^{20/22}$), 150.3 ($C^{2/4}$), 145.9 ($CH^{6}$), 143.9 ($C^{Ar}$), 143.8 ($C^{Ar}$), 143.5 ($C^{Ar}$), 142.6 ($C^{Ar}$), 142.3 ($C^{Ar}$), 137.0 ($C^{10}$), 136.8 ($C^{10}$), 133.0 ($CH^{Ar}$), 131.1 ($CH^{Ar}$), 129.6 ($CH^{Ar}$), 129.1 ($CH^{Ar}$), 129.0 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 125.9 ($CH^{21}$), 123.3 ($CH^{Ar}$), 114.3 ($C^{12}$), 112.4 ($C^{Ar}$), 111.4 ($C^{Ar}$), 105.4 ($CH^{23}$), 103.7 ($CH^{19}$), 100.1 ($C^{5}$), 93.3 ($C^{8,9}$), 87.7 ($CH^{4'}$), 86.9 ($CH^{1'}$), 82.6 ($C^{7}$), 72.0 ($CH^{3'}$), 64.3 ($CH_2{}^{5'}$), 55.9 ($CH_3{}^{18}$), 50.6 ($C^{46}$), 49.6 ($C^{38}$), 44.7 ($CH_2{}^{27}$), 42.1 ($CH_2{}^{2'}$), 32.5 ($CH_3{}^{26}$), 30.0 ($CH_2{}^{28}$), 27.9 ($CH_3{}^{25}$), 27.5 ($CH_3{}^{24}$), 20.8 ($CH_2{}^{29}$), 14.2 ($CH_3{}^{30}$) ppm.

Mp: >130°C (decomposes).

IR $\nu_{Max}/cm^{-1}$: 3325 (br. w. O-H), 2959 (w. N-H), 2929, 2871 (w. C-H, $C-H_3$), 1704 (s. C=O).

UV/Vis (MeOH): $A_{max}$ 660 nm, ε 208,000 $M^{-1}cm^{-1}$,$Em_{max}$ 690 nm.

**Synthesis of Cy5dT-phosphoramidite (21)**




**21**

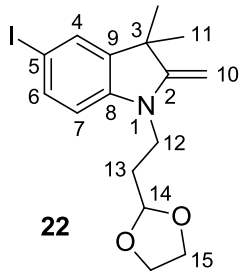
Cy5dT (**20**) (0.24 g, 0.22 mmol) was co-evaporated with distilled pyridine (3 x 5 mL) followed by distilled DCM (5 x 5 mL) before suspending in distilled DCM (5 mL), with activated molecular sieves (3 Å), under an argon atmosphere. To the solution was added distilled DIPEA (0.10 mL, 0.55 mmol). 2-Cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.05 mL, 0.22 mmol) was added drop-wise and the reaction was stirred for 45 min. Additional 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.1 eq) was added before stirring for a further 1 hour. The mixture was diluted with distilled DCM (10 mL) and washed with deoxygenated sat. KCl (10 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. The crude product was purified by precipitation in deoxygenated hexane (50 mL) from a minimum of distilled DCM (2 mL). The precipitation procedure was repeated twice more. The precipitate was dried *in vacuo* to afford the product **21** (0.24 g, 0.18 mmol, 84 %) as a dark blue solid.

$R_f$ : 0.17 (0.5:9.5 MeOH/DCM with 1 % pyridine)

LRMS [ESI+, MeCN] m/z (%): 1178 ($M^+$, 100).

$^{31}P$ (121 MHz, $CD_3CN$) δ: 149.34 (s, P), 149.22 (s, P) ppm.

**Synthesis of 5-iodo-1-ethyldioxolane-3,3-dimethyl-2-methyleneindoline (22)**




**22**

To a suspension of KI (15.4 g, 92.7 mmol) in distilled MeCN (60 mL) under an argon atmosphere was added bromoethyldioxolane (7.25 mL, 61.7 mmol) and the yellow suspension was stirred at 50 °C for 1 hour. 5-Iodo-2,3,3-trimethylindolenine (8.33 g, 29.2 mol) was added in distilled MeCN (20 mL) and the reaction was stirred at reflux for 45 h. The reaction was allowed to cool to rt and the precipitate was removed by filtration (KBr). The solvent was removed *in vacuo* and following purification by column chromatography (100 % DCM with 0.5 % $Et_3N$) the product was afforded as a brown oil (3.28 g, 8.52 mmol, 29 %).

$R_f$ : 0.62 (5:95 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 386 ($(M+H)^+$, 100).

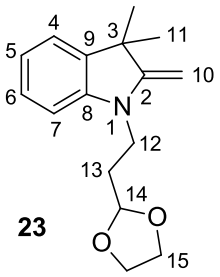
HRMS [ESI+, MeCN] for $C_{16}H_{21}I_1N_1O_2$ $(M+H)^+$: calcd 386.0611, found 386.0607.

$^1$H (400 MHz, $CD_3CN$): δ 7.40 (m, 2H, $H^{Ar}$), 6.46 (d, *J*=9.0, 1H, $H^{Ar}$), 4.85 (t, *J*=4.0, 1H, $H^{14}$), 3.96-3.77 (m, 6H, $H^{10,15}$), 3.60 (t, *J*=7.3, 2H, $H^{12}$), 1.89 (dt, *J*=7.2, 4.8, 2H, $H^{13}$), 1.27 (s, 6H, $H^{11}$) ppm.

$^{13}$C (100 MHz, $CD_3CN$): δ 161.6 ($C^2$), 146.9 ($C^{Ar}$), 141.9 ($C^{Ar}$), 137.5 ($CH^{Ar}$), 131.8 ($CH^{Ar}$), 109.0 ($CH^{Ar}$), 103.5 ($CH^{14}$), 79.9 ($C^5$), 75.5 ($CH_2{}^{10}$), 65.9 ($CH_2{}^{15}$), 45.2 ($C^3$), 38.3 ($CH_2{}^{12}$), 31.1 ($CH_2{}^{13}$), 30.3 ($CH_3{}^{11}$) ppm.

IR $\nu_{Max}/cm^{-1}$: 2960 (m, C-H aromatic), 2928 (m, =C-H), 2871 (m, -C-H).

**Synthesis of 1-ethyldioxolane-3,3-dimethyl-2-methyleneindoline (23)**




**23**

To a suspension of KI (3.98 g, 0.02 mol) in distilled MeCN (20 mL), under an argon atmosphere, was added bromoethyldioxolane (1.55 mL, 13.2 mmol) and the yellow suspension was stirred at 50 °C for 1 h. 2,3,3-Trimethylindolenine (2 mL, 12.0 mmol) was added and the reaction stirred at reflux for 45 h. The reaction was allowed to cool to rt and filtered to remove KBr salt. The solvent was removed *in vacuo* and purified by column chromatography (twice in 100 % DCM with 0.5 % $Et_3N$) to afford the product **23** (1.73 g, 6.68 mmol, 56 %) as an orange oil.

$R_f$ : 0.42 (5:95 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 260 ($(M+H)^+$, 100).

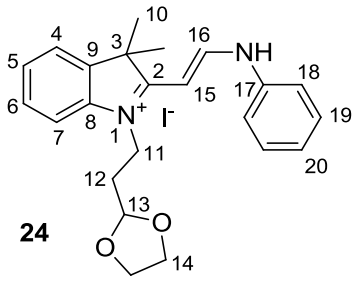
HRMS [ESI+, MeCN] for $C_{16}H_{22}N_1O_2$ $(M+H)^+$: calcd 260.1645, found 260.1645.

$^1$H (400 MHz, $CD_3CN$): δ 7.11 (m, 2H, $H^{Ar}$), 6.74 (t, *J*= 8.0, 1H, $H^{Ar}$), 6.63 (d, *J*=12.0, 1H, $H^{Ar}$), 4.88 (t, *J*=8.0, 1H, $H^{14}$), 3.96-3.78 (m, 6H, $H^{10,15}$), 3.64 (t, *J*=12.0, 2H, $H^{12}$), 1.90 (m, 2H, $H^{13}$), 1.30 (s, 6H, $H^{11}$) ppm.

$^{13}$C (100 MHz, $CD_3CN$): δ 162.2 ($C^2$), 146.7 ($C^{Ar}$), 138.6 ($C^{Ar}$), 128.6 ($CH^{Ar}$), 122.9 ($CH^{Ar}$), 119.5 ($CH^{Ar}$), 106.4 ($CH^{Ar}$), 103.4 ($CH^{14}$), 74.2 ($CH_2{}^{10}$), 65.7 ($CH_2{}^{15}$), 45.0 ($C^3$), 38.1 ($CH_2{}^{12}$), 31.0 ($CH_2{}^{13}$), 30.3 ($CH_3{}^{11}$) ppm.

IR $\nu_{Max}/cm^{-1}$: 2960 (m, C-H aromatic), 2924 (m, =C-H), 2881 (m, -C-H).

**Synthesis of 2-[(E)-2-(phenylamino)vinyl]-1-ethyldioxolane-,3,3-trimethyl-3*H*-indolium iodide (24)**




To a solution of 1-ethyldioxolane-3,3-dimethyl-2-methyleneindoline (**23**) (0.21 g, 0.82 mmol) in EtOH (2 mL), under an argon atmosphere, was added *N,N'*-diphenylformamidine (0.16 g, 0.82 mmol) and triethylorthoformate (0.14 mL, 0.82 mmol) and the reaction refluxed at high temperature (97 °C) in the dark for 16 h. The reaction solvent was removed *in vacuo* and following purification by column chromatography (100 % DCM with 0.5 % $Et_3N$) the product **24** (0.26 g, 0.53 mmol, 65 %) was isolated as an orange foam.

$R_f$ : 0.51 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 363 ($(M)^+$, 100).

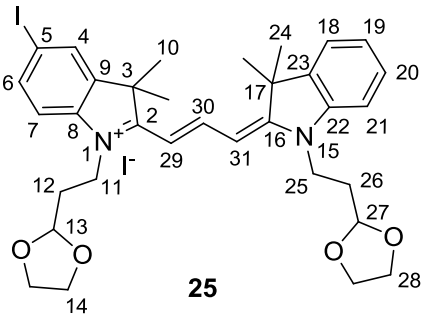
HRMS [ESI+, MeCN] for $C_{23}H_{27}N_2O_2$ $(M)^+$: calcd 363.2067, found 363.2068.

$^1$H (400 MHz, $CD_3CN$): δ 8.60 (d, *J*=12.0, 1H, $H^{16}$), 7.51-7.26 (m, 9H, $H^{Ar}$), 6.59 (d, *J*=12.0, 1H, $H^{15}$), 5.02 (t, *J*=4.0, 1H, $H^{13}$), 4.17 (t, *J*=8.0, 2H, $H^{11}$), 3.92-3.78 (m, 4H, $H^{14}$), 2.19 (dt, *J*=8.0, 4.0, 2H, $H^{12}$), 1.68 (s, 6H, $H^{10}$) ppm.

$^{13}$C (100 MHz, $CD_3CN$): δ 177.9 ($C^2$), 153.3 ($CH^{16}$), 143.1 ($C^{Ar}$), 142.0 ($C^{Ar}$), 141.5 ($C^{Ar}$), 130.8 (CHAr), 129.6 (CHAr), 127.1 (CHAr), 126.4 (CHAr), 123.3 (CHAr), 119.4 ($CH^{Ar}$), 112.3 ($CH^{Ar}$), 102.7 ($CH^{13}$), 92.8 ($CH^{15}$), 65.8 ($CH_2{}^{14}$), 50.5 ($C^3$), 40.7 ($CH_2{}^{11}$), 31.3 ($CH_2{}^{12}$), 28.6 ($CH_3{}^{10}$) ppm.

IR $\nu_{Max}/cm^{-1}$: 3419 (w, br, N-H), 2965 (m, =C-H), 2885 (m, -C-H), 2358 (w, $=N^+$-).

**Synthesis of 5-iodo-1-ethyldioxolane-3,3-trimethyl-2-((1E,3E)-3-(1-ethyldioxolane-3,3-trimethylindolin-2-ylidene)prop-1-enyl)-3*H*-indolium iodide (25)**




**25**

To a solution of 2-[(E)-2-(phenylamino)vinyl]-1-ethyldioxolane-,3,3-trimethyl-3*H*-indolium iodide (**24**) (0.35 g, 0.72 mmol) in distilled pyridine (8 mL) was added 5-iodo-1-ethynyldioxolane-3,3-dimethyl-2-methyleneindoline (**22**) (0.55 g, 1.44 mmol) and $Ac_2O$ (0.68 mL, 7.18 mmol) and the reaction stirred under an argon atmosphere in the dark at 50 °C for 24 h. The reaction solvent was removed *in vacuo* and the crude material was purified by column chromatography (5:95 MeOH/EtOAc with 0.1 % pyridine → 100 % DCM with 0.1 % pyridine) to afford the product **25** (0.45 g, 0.58 mmol, 79 %) as a dark red solid.

$R_f$ : 0.44 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 655 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{33}H_{40}I_1N_2O_4$ $(M)^+$: calcd 655.2027, found 655.2027.

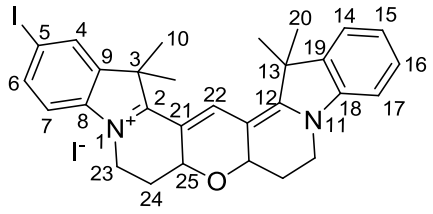
$^1$H (400 MHz, $CD_3CN$): δ 8.44 (t, *J*=12.0, 1H, $H^{30}$), 7.86 (d, *J*=4.0, 1H, $H^{Ar}$), 7.76 (dd, *J*=8.0, 4.0, 1H, $H^{Ar}$), 7.54-7.31 (m, 4H, $H^{Ar}$), 7.07 (d, *J*=8.0, 1H, $H^{Ar}$), 6.45 (d, *J*=16.0, 1H, $H^{29/31}$), 6.36 (d, *J*=12.0, 1H, $H^{29/31}$), 4.98 (m, 2H, $H^{13,27}$), 4.21 (t, *J*=8.0, 2H, $H^{11/25}$), 4.13 (t, *J*=8.0, 2H, $H^{11/25}$), 3.91-3.78 (m, 8H, $H^{14,28}$), 2.18-2.11 (m, 4H, $H^{12/26}$), 1.69 (s, 12H, $H^{10,24}$) ppm.

$^{13}$C (100 MHz, $CD_3CN$): δ 176.8 ($C^{2/16}$), 174.8 ($C^{2/16}$), 152.0 ($CH^{30}$), 144.6 ($C^{Ar}$), 143.7 ($C^{Ar}$), 143.4 ($C^{Ar}$), 142.6 ($C^{Ar}$), 138.9 ($CH^{Ar}$), 132.8 ($CH^{Ar}$), 130.1 ($CH^{Ar}$), 127.2 ($CH^{Ar}$), 123.8 ($CH^{Ar}$), 114.6 ($CH^{Ar}$), 113.1 ($CH^{Ar}$), 104.9 ($CH^{29/31}$), 103.8 ($CH^{29/31}$), 103.0 ($CH^{13,27}$), 89.3 ($C^5$), 66.2 ($CH_2{}^{14,28}$), 51.0 ($C^{3/17}$), 50.5 ($C^{3/17}$), 41.0 ($CH_2{}^{11/25}$), 40.7 ($CH_2{}^{11/25}$), 31.9 ($CH_2{}^{12/26}$), 31.7 ($CH_2{}^{12/26}$), 28.4 ($CH_3{}^{10,24}$) ppm.

Mp: 80°C (decomposes)

IR $\nu_{Max}/cm^{-1}$: 2968 (m, =C-H), 2885 (m, -C-H), 2359 (m, $=N^+$-).

**Synthesis of 5-iodo-Cy3B (26)[148]**



**26**

To a solution of 5-iodo-1-ethyldioxolane-3,3-trimethyl-2-((1E,3E)-3-(1-ethyldioxolane-3,3-trimethylindolin-2-ylidene)prop-1-enyl)-3*H*-indolium iodide (**25**) (0.58 g, 0.74 mmol) in $CHCl_3$ (58 mL) was added 50 % aq. $H_2SO_4$ (12 mL) and the mixture was stirred for 20 min. The reaction mixture was diluted with $CHCl_3$ (60 mL) and washed with $H_2O$ (3 x 60 mL). The aqueous phase was combined, washed with $CHCl_3$ (2 x 60 mL) and the combined organic phase was dried ($Na_2SO_4$) then the solvent was removed *in vacuo* to afford the product **26** (0.48 g, 0.71 mmol, 96 %) as a dark pink solid.

$R_f$ : 0.30 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 549 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{29}H_{30}I_1N_2O_1$ $(M)^+$: calcd 549.1397, found 549.1402.

$^1$H (400 MHz, $CDCl_3$): δ 8.06 (s, 1H, $H^{22}$), 7.71 (dd, *J*=8.3, 1.5, 1H, $H^{Ar}$), 7.64 (d, *J*=1.5, 1H, $H^{Ar}$), 7.45-7.40 (m, 2H, $H^{Ar}$), 7.35-7.22 (m, 2H, $H^{Ar}$), 7.00 (d, *J*=8.3, 1H, $H^{Ar}$), 4.99 (dt, *J*=11.6, 5.7, 2H, $H^{25}$), 4.30-4.19 (m, 4H, $H^{23}$), 2.72-2.68 (m, 2H, $H^{24}$), 2.06-1.93 (m, 2H, $H^{24}$), 1.79 (s, 3H, $H^{10/20}$), 1.77 (s, 3H, $H^{10/20}$), 1.74 (s, 3H, $H^{10/20}$), 1.72 (s, 3H, $H^{10/20}$) ppm.

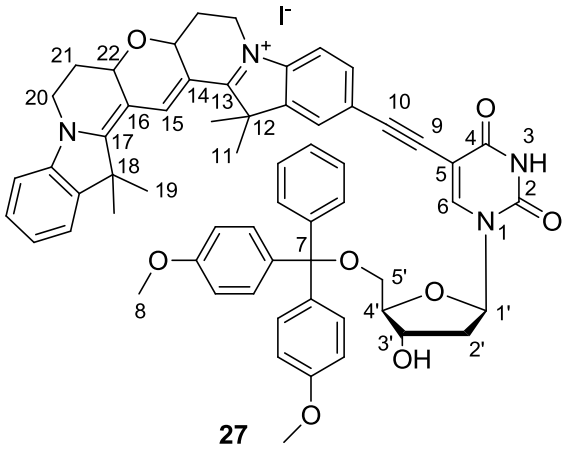
$^{13}$C (100 MHz, $CDCl_3$): δ 168.5 ($C^{2/12}$), 166.3 ($C^{2/12}$), 142.3 ($C^{Ar}$), 141.4 ($C^{Ar}$), 141.3 ($C^{Ar}$), 140.4 ($C^{Ar}$), 138.3 ($CH^{22}$), 137.8 ($CH^{Ar}$), 131.1 ($CH^{Ar}$), 129.1 ($CH^{Ar}$), 126.1 ($CH^{Ar}$), 122.2 ($CH^{Ar}$), 112.6 ($CH^{Ar}$), 111.3 ($C^{21}$), 111.1 ($CH^{Ar}$), 110.3 ($CH^{Ar}$), 88.9 ($C^5$), 69.5 ($CH^{25}$), 48.9 ($C^{3/13}$), 48.4 ($C^{3/13}$), 42.1 ($CH_2{}^{23}$), 42.0 ($CH_2{}^{23}$), 28.4 ($CH_3{}^{10/20}$), 27.8 ($CH_3{}^{10/20}$), 26.5 ($CH_2{}^{24}$), 26.4 ($CH_2{}^{24}$) ppm.

Mp: 110°C (decomposes).

IR $\nu_{Max}/cm^{-1}$: 2960 (m, =C-H), 2923 (m, -C-H), 2359 (m, $=N^+$-).

UV/Vis (MeOH): $A_{max}$ 565 nm, ε 141,000 $M^{-1}cm^{-1}$, $Em_{max}$ 575 nm.

**Synthesis of Cy3B-dT (27)**




**27**

To a mixture of 5-iodo-Cy3B (**26**) (0.48 g, 0.71 mmol), 1-[2'-(deoxy)-5'-*O*-(4,4'-dimethoxytrityl)-*β*-D-erythro-pentafuranosyl]-5-(eth-1-ynyl)uridine (**17**) (0.59 g, 1.06 mmol) and CuI (0.03 g, 0.14 mmol) in anhydrous DMF (4 mL), under an argon atmosphere, was added distilled $Et_3N$ (2.5 mL) and the reaction was stirred for 10 min. $Pd(PPh_3)_4$ (0.08 g, 0.07 mmol) was added and the solution was stirred in the dark for 1.5 h. The reaction mixture was diluted with DCM (100 mL) and washed with 5 % sodium EDTA (100 mL). The organic phase was evaporated to dryness *in vacuo* and purified by column chromatography (0→10 % MeOH/DCM with 0.1 % pyridine) affording the product **27** (0.33 g, 0.30 mmol, 42 %) as a dark pink solid.

$R_f$ : 0.24 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 975 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{61}H_{59}N_4O_8$ $(M)^+$: calcd 975.4327, found 975.4325.

$^1$H (400 MHz, $d_6$-DMSO): δ 11.78 (s, 1H, NH), 8.13 (s, 1H, $H^6$), 7.98 (s, 1H, $H^{15}$), 7.45-7.13 (m, 16H, $H^{Ar}$), 6.86 (dd, 4H, *J*=8.0, 4.0, $H^{Ar}$), 6.16 (t, *J*=4.0, 1H, $H^{1'}$), 5.35 (d, *J*=4.0, 1H, $OH^{3'}$), 4.64 (m, 2H, $H^{22}$), 4.35-4.29 (m, 3H, $H^{3',20}$), 3.97-3.84 (m, 3H, $H^{4',20}$), 3.68 (s, 6H, $H^8$), 3.22-3.19 (m, 2H, $H^{5'}$), 2.51-2.49 (m, 2H, $H^{21}$), 2.32-2.27 (m, 2H, $H^{2'}$), 1.99-1.90 (m, 2H, $H^{21}$), 1.71 (s, 6H, $H^{11/19}$), 1.64 (s, 6H, $H^{11/19}$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): δ 174.0 ($C^{Ar}$), 168.3 ($C^{17/13}$), 166.5 ($C^{17/13}$), 161.4 ($C^{2/4}$), 158.0 ($C^{Ar}$), 149.3 ($C^{2/4}$), 144.7 ($C^{Ar}$), 142.8 ($CH^6$), 141.7 ($C^{Ar}$), 140.7 ($C^{Ar}$), 140.6 ($C^{Ar}$), 137.1 ($CH^{15}$), 135.5 ($C^{Ar}$), 135.3 ($C^{Ar}$), 131.9 ($CH^{Ar}$), 129.7 ($CH^{Ar}$), 128.7 ($CH^{Ar}$), 127.9 ($CH^{Ar}$), 127.57 ($CH^{Ar}$), 126.7 ($CH^{Ar}$), 125.6 ($CH^{Ar}$),

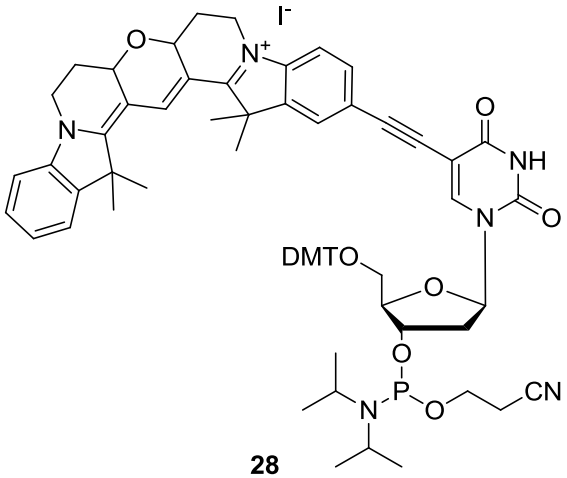
124.8 ($CH^{Ar}$), 122.4 ($CH^{Ar}$), 118.4 ($C^{Ar}$), 113.2 ($CH^{Ar}$), 111.3 ($C^{14/16}$), 110.8 ($C^{14/16}$), 110.4 ($CH^{Ar}$), 109.3 ($CH^{Ar}$), 98.5 ($C^{5}$), 91.9 ($C^{10}$), 86.1 ($C^{9}$), 85.9 ($CH^{4'}$), 85.3 ($CH^{1'}$), 82.3 ($C^{7}$), 70.5 ($CH^{3'}$), 69.3 ($CH^{22}$), 63.5 ($CH_2{}^{5'}$), 55.0 ($CH_3{}^{8}$), 48.5 ($C^{12/18}$), 47.9 ($C^{12/18}$), 41.3 ($CH_2{}^{20}$), 41.0 ($CH_2{}^{2'}$), 27.5 ($CH_3{}^{11/19}$), 26.8 ($CH_3{}^{11/19}$), 26.1 ($CH_2{}^{21}$) ppm.

Mp: >200°C (decomposes).

IR $\nu_{Max}/cm^{-1}$: 3307 (w, br, N-H and –OH), 2960 (m, =C-H), 2929 (m, -C-H), 2359 (m, $=N^{+}$-), 1689 (s, C=O).

UV/Vis (MeOH): $A_{max}$ 581 nm, ε = 122,000 $M^{-1}cm^{-1}$, $Em_{max}$ 596 nm.

**Synthesis of Cy3B-dT-phosphoramidite (28)**




**28**

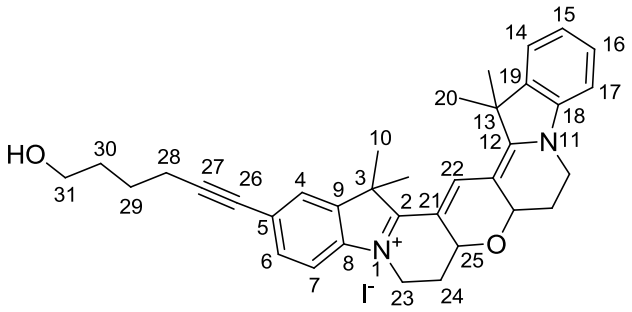
Cy3B-dT (**27**) (0.67 g, 0.61 mmol) was co-evaporated with distilled pyridine (3 x 5 mL) and distilled DCM (3 x 5 mL) then dissolved in distilled DCM (6 mL), with activated molecular sieves (3 Å), under an argon atmosphere. Distilled DIPEA (0.29 mL, 1.53 mmol) was added followed by 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.15 mL, 0.67 mmol) drop-wise and the reaction stirred at rt for 45 min.
The reaction was diluted with distilled DCM (10 mL) and washed with deoxygenated sat. KCl (20 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. The crude product was precipitated 3 times consecutively in deoxygenated hexane (3 x 100 mL) from a minimum of distilled DCM (2 mL). The product was co-evaporated with distilled DCM (5 mL) and trace solvents removed *in vacuo* to afford the product **28** (0.50 g, 0.38 mmol, 63 %) as a dark pink solid.

$R_f$ : 0.44(1:9 MeOH/DCM with 1 % pyridine)

LRMS [ESI+, MeCN] m/z (%): 1176 ($(M)^+$, 100).

$^{31}P$ (121 MHz, $CD_3CN$): 149.3 (s, P), 149.1 (s, P) ppm.

**Synthesis of 5-hexyn-1-ol-6-Cy3B (29)**




**29**

To a solution of 5-iodo-Cy3B **(26)** (0.50 g, 0.70 mmol) and CuI (0.03 g, 0.14 mmol) in anhydrous DMF (3.5 mL), under an argon atmosphere, was added distilled $Et_3N$ (2.5 mL) and 5-hexyn-1-ol (0.08 mL, 0.70 mmol) and the reaction was stirred in the dark at rt for 10 min. $Pd(PPh_3)_4$ (0.08 g, 0.07 mmol) was added and the solution stirred for 3.5 h. An additional portion of 5-hexyn-1-ol (0.07 mL) was added and the reaction stirred for a further 1.5 h. The reaction was diluted with DCM (40 mL) and washed with 5 % sodium EDTA (3 x 50 mL) and 5 % KI solution (2 x 50 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. The crude product was dissolved in DCM (2 mL), filtered through celite and precipitated in hexane (100 mL). The material was filtered through alumina and trace solvents removed *in vacuo* to afford the product **29** (0.37 g, 0.57 mmol, 82 %) as a deep pink solid.

$R_f$ : 0.25 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 519 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{35}H_{39}N_2O_2$ $(M)^+$: calcd 519.3006, found 519.3000.

$^1H$ (400 MHz, $CDCl_3$): δ 8.00 (s, 1H, $H^{22}$), 7.62-7.02 (m, 7H, $H^{Ar}$), 5.00 (2xdd, *J*=11.0, 11.0, 2H, $H^{25}$), 4.20-4.13 (m, 4H, $H^{23}$), 3.65 (t, *J*=6.0, 2H, $H^{31}$), 2.66-2.62 (m, 2H, $H^{24}$), 2.41 (t, *J*=7.0, 2H, $H^{28}$), 1.97-1.92 (m, 2H, $H^{24}$), 1.72-1.65 (m, 16H, $H^{10,20,29,30}$) ppm.

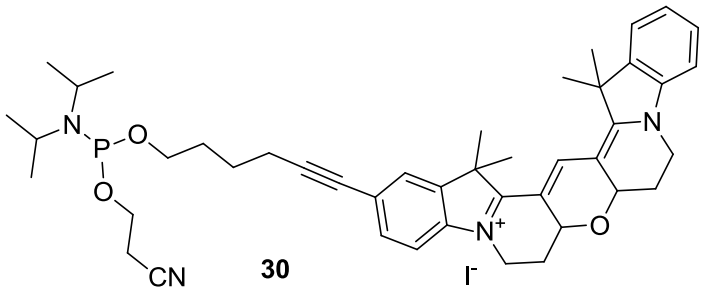
$^{13}C$ (100 MHz, $CDCl_3$): δ 168.5 ($C^{2/12}$), 167.5 ($C^{2/12}$), 141.9 ($C^{Ar}$), 141.2 ($C^{Ar}$), 140.8 ($C^{Ar}$), 140.7 ($C^{Ar}$), 138.6 ($CH^{6}$), 132.4 ($CH^{Ar}$), 128.8 ($CH^{Ar}$), 126.4 ($CH^{Ar}$), 125.7 ($CH^{Ar}$), 122.6 ($CH^{Ar}$), 121.7 ($C^{Ar}$), 111.6 ($C^{21}$), 111.4 ($CH^{Ar}$), 111.0 ($C^{21}$), 110.9 ($CH^{Ar}$), 91.9 ($C^{27}$), 80.8 ($C^{26}$), 70.0 ($CH^{25}$), 62.7 ($CH_2{}^{31}$), 49.2 ($C^{3/13}$), 48.8 ($C^{3/13}$), 42.3 ($CH_2{}^{23}$), 32.3 ($CH_2{}^{29/30}$), 28.8 ($CH_3{}^{10/20}$), 28.3 ($CH_3{}^{10/20}$), 27.0 ($CH_2{}^{29/30}$), 25.4 ($CH_2{}^{24}$), 19.7 ($CH_2{}^{28}$) ppm.

Mp: >130°C (decomposes).

IR $\nu_{Max}/cm^{-1}$: 3363 (br, OH), 2927 (m, =C-H), 2860 (m, -C-H), 2350 (m, C≡C).

UV/Vis (MeOH): $A_{max}$ 573 nm, $Em_{max}$ 585 nm

**Synthesis of 5-hexyn-1-ol-6-Cy3B phosphoramidite (30)**




**30**

5-hexyn-1-ol-6-Cy3B (**29**) (0.35 g, 0.54 mmol) was co-evaporated with distilled pyridine (5 mL) and distilled DCM (3 x 5mL) then dissolved in distilled DCM (5 mL), with activated molecular sieves (3 Å), under an argon atmosphere. To this was added distilled DIPEA (0.24 mL, 1.35 mmol) followed by 2-cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.13 mL, 0.60 mmol) and the reaction was stirred at rt for 45 min. The reaction was diluted with distilled DCM (10 mL) and washed with deoxygenated sat. KCl (10 mL) then dried ($Na_2SO_4$). The solvent was removed *in vacuo* then purified by precipitation in deoxygenated hexane (150 mL) from a minimum of distilled DCM (2 mL). The precipitation was repeated a further two times. The precipitate was co-evaporated with distilled DCM (3 mL) and trace solvents were removed *in vacuo* to afford the product **30** (0.42 g, 0.50 mmol, 93 %) as a purple foam.

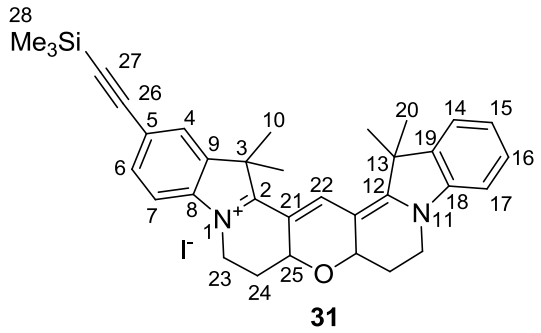
$R_f$ : 0.46 (1:9 MeOH/DCM with 1 % pyridine)

LRMS [ESI+, MeCN] m/z (%): 720 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{44}H_{56}N_4O_3P_1$ $(M)^+$: calcd 719.4085, found 719.4076.

$^{31}P$ (121 MHz, $CD_3CN$): δ 148.2 (s, P) ppm.

**Synthesis of 5-trimethylsilylethynyl-Cy3B (31)**




**31**

To a solution of 5-iodo-Cy3B (**26**) (0.29 g, 0.43 mmol) and CuI (0.02 g, 0.09 mmol) in anhydrous DMF (2.5 mL), under an argon atmosphere was added distilled $Et_3N$ (1.5 mL) and trimethylsilylacetylene (0.09 mL, 0.64 mmol) and the reaction was stirred at rt for 10 min.

$Pd(PPh_3)_4$ (0.05 mg, 0.04 mmol) was added and the reaction was stirred for 16 h.

The reaction was diluted with DCM (100 mL) and washed with 5 % sodium EDTA solution (100 mL) and 5 % KI solution (100 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. Following purification by column chromatography (1→10 % 7M $NH_3$:MeOH/DCM) the product **31** (0.20 g, 0.31 mmol, 71 %) was afforded as a purple iridescent solid.

$R_f$ : 0.35 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 519 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{34}H_{39}N_2O_1Si_1$ $(M)^+$: calcd 519.2826, found 519.2822.

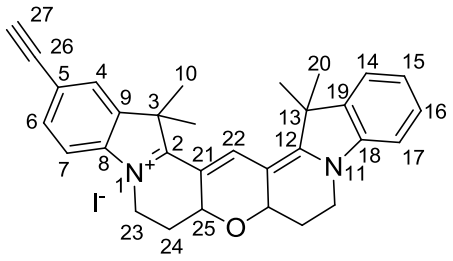
$^1H$ (400 MHz, $CDCl_3$): δ 7.92 (s, 1H, $H^{22}$), 7.36-6.95 (m, 7H, $H^{Ar}$), 4.87 (dd, 1H, *J*=11.0, 5.0, $H^{25}$), 4.82 (dd, 1H, *J*= 11.0, 5.0, $H^{25}$), 4.15-4.01 (m, 4H, $H^{23}$), 2.56-2.53 (m, 2H, $H^{24}$), 1.89-1.83 (m, 2H, $H^{24}$), 1.64 (s, 3H, $H^{10/20}$), 1.61 (s, 3H, $H^{10/20}$), 1.59 (s, 3H, $H^{10/20}$), 1.56 (s, 3H, $H^{10/20}$), 0.11 (s, 9H, $H^{28}$) ppm.

$^{13}C$ (100 MHz, $CDCl_3$): δ 168.6 ($C^2$), 166.7 ($C^{12}$), 141.5 ($C^{Ar}$), 141.4 ($C^{Ar}$), 140.5 ($C^{Ar}$), 140.1 ($C^{Ar}$), 138.2 ($CH^{22}$), 133.2 ($CH^{Ar}$), 129.1 ($CH^{Ar}$), 126.2 ($CH^{Ar}$), 125.5 ($CH^{Ar}$), 122.2 ($CH^{Ar}$), 120.1 ($C^{Ar}$), 111.6 ($C^{21}$), 111.2 ($CH^{Ar}$), 110.5 ($C^{21}$), 110.4 ($CH^{Ar}$), 104.3 ($C^{26}$), 95.6 ($C^{27}$), 69.5 ($CH^{25}$), 48.9 ($C^{3/13}$), 48.2 ($C^{3/13}$), 42.2 ($CH_2{}^{23}$), 41.9 ($CH_2{}^{23}$), 28.4 ($CH_3{}^{10}$), 28.3 ($CH_3{}^{10}$), 27.8 ($CH_3{}^{20}$), 27.8 ($CH_3{}^{20}$), 26.5 ($CH_2{}^{24}$), 26.5 ($CH_2{}^{24}$), 0.1 ($CH_3{}^{28}$) ppm.

Mp: 210-211°C.

IR $\nu_{Max}/cm^{-1}$: 2963 (m, =C-H), 2930 (m, -C-H), 2149 (m, C≡C).

**Synthesis of 5-ethynyl-Cy3B (32)**




**32**

To a solution of 5-trimethylsilylethynyl-Cy3B (**31**) (0.13 g, 0.20 mmol) in distilled THF (3 mL), under an argon atmosphere, was added 1M TBAF in THF (0.31 mL, 0.31 mmol) and the reaction was stirred for 5 min at rt. The reaction was diluted with DCM (100 mL) and washed ($H_2O$ 2 x 500 mL, 5 % KI solution 250 mL). The organic phase was dried ($Na_2SO_4$), the solvent removed *in vacuo* and the crude material purified by column chromatography (1:99 7M $NH_3$:MeOH/DCM) to afford the product **32** (0.09 g, 0.16 mmol, 75 %) as a purple iridescent solid.

$R_f$ : 0.32 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 447 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{31}H_{31}N_2O_1$ $(M)^+$: calcd 447.2431, found 447.2422.

$^1H$ (400 MHz, $CDCl_3$): δ 8.08 (s, 1H, $H^{22}$), 7.53-7.13 (m, 7H, $H^{Ar}$), 5.01 (dd, *J*=11.0, 5.0, 1H, $H^{25}$), 4.96 (dd, *J*=11.0, 5.0, 1H, $H^{25}$), 4.29-4.17 (m, 4H, $H^{23}$), 3.16 (s, 1H, $H^{27}$), 2.73-2.68 (m, 2H, $H^{24}$), 2.03-1.95 (m, 2H, $H^{24}$), 1.79 (s, 3H, $H^{10/20}$), 1.77 (s, 3H, $H^{10/20}$), 1.74 (s, 3H, $H^{10/20}$), 1.72 (s, 3H, $H^{10/20}$) ppm.

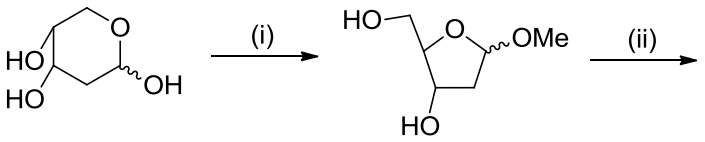
$^{13}C$ (100 MHz, $CDCl_3$): δ 168.7 ($C^{2/12}$), 166.7 ($C^{2/12}$), 141.8 ($C^{Ar}$), 141.4 ($C^{Ar}$), 140.5 ($C^{Ar}$), 140.2 ($C^{Ar}$), 138.4 ($CH^{22}$), 133.3 ($CH^{Ar}$), 129.1 ($CH^{Ar}$), 126.2 ($CH^{Ar}$), 125.7 ($CH^{Ar}$), 122.2 ($CH^{Ar}$), 119.0 ($C^5$), 111.7 ($C^{21}$), 111.2 ($CH^{Ar}$), 110.5 ($CH^{Ar}$), 110.5 ($C^{21}$), 83.0 ($C^{26}$), 78.3 ($CH^{27}$), 69.6 ($CH^{25}$), 49.0 ($C^{3/13}$), 48.2 ($C^{3/13}$), 42.2 ($CH_2{}^{23}$), 41.9 ($CH_2{}^{23}$), 28.4 ($CH_3{}^{10}$), 28.3 ($CH_3{}^{10}$), 27.8 ($CH_3{}^{20}$), 27.7 ($CH_3{}^{20}$), 26.5 ($CH_2{}^{24}$), 26.5 ($CH_2{}^{24}$) ppm.

Mp: >230°C.

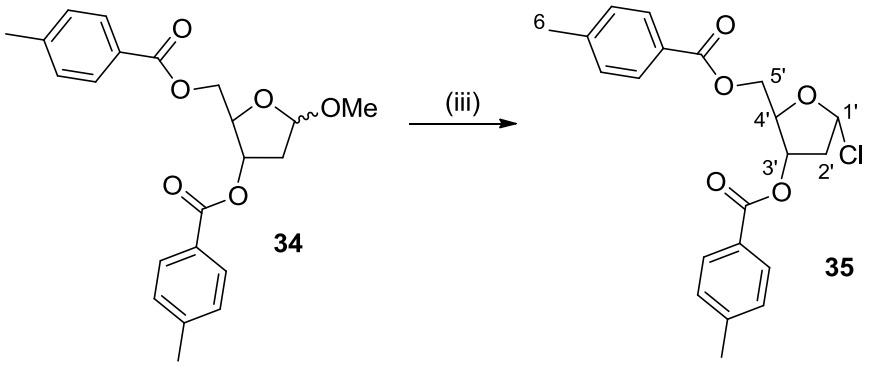
IR $\nu_{Max}/cm^{-1}$: 3167 (w, C≡C-H), 2969 (m, =C-H), 2928 (m, -C-H).

UV/Vis (MeOH): $A_{max}$ 567 nm, ε 118,000 $M^{-1}cm^{-1}$, $Em_{max}$ 584 nm.

**Synthesis of 1'-*α*-chloro-2'-deoxy-3,5-di-*O*-*p*-toluoyl-D-ribose (33-35)[155]**




**33**




**34** **35**

To a solution of 2'-deoxy-D-ribose (10.0 g, 74.5 mmol) in MeOH (120 mL) was added 1 % methanolic HCl (20 mL: made by adding 1.7 mL acetyl chloride to 100 mL MeOH). The resulting yellow solution was stirred for 30 min. The reaction was neutralised by the addition of solid sodium bicarbonate (4.00 g). The solid was removed by filtration, washed with MeOH and the filtrate was evaporated. Residual MeOH was removed by co-evaporation with distilled pyridine (1 x 50 mL, 2 x 25 mL) and the golden coloured syrup was dried on high vacuum for 2.5 h.

To the residue, dissolved in distilled pyridine (60 mL), cooled to 0°C, was added drop-wise *p*-toluoylchloride (22.0 mL, 160 mmol). The cloudy solution was stirred at rt for 16 h. The reaction was diluted with cold water (150 mL) and extracted with DCM (1 x 150 mL, 2 x 100 mL).

The organic phase was washed with $NaHCO_3$ (2 x 150 mL), 2M HCl (150 mL) and $H_2O$ (200 mL), then dried ($Na_2SO_4$) and the solvent removed *in vacuo*.

To the oil (30.8 g) dissolved in AcOH (40 mL) at 0°C, was slowly added sat. HCl in AcOH (63 mL; prepared by adding 16.3 mL acetyl chloride to a mixture of cold AcOH, 81 mL, and water, 4 mL). An additional amount of acetyl chloride (7 mL) was added and the flask was swirled for 10 min resulting in the formation of a thick crystalline precipitate. The precipitate was rapidly filtered and washed thoroughly with cold anhydrous $Et_2O$ (prepared

over molecular sieves, 3 Å). Following drying *in vacuo* the product **35** (18.7 g, 48.1 mmol, 64 %) was afforded as a white solid which was stored under argon.

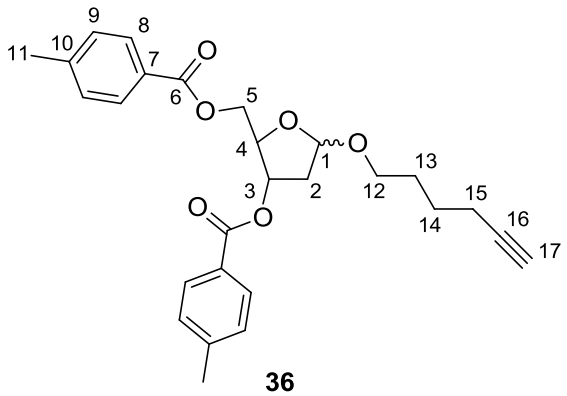
$R_f$ : 0.48 (1:1 EtOAc/ Pet ether)

LRMS [ESI+,MeCN direct probe application] m/z (%): 452 (($M(^{35}Cl)+Na+CH_3CN)^+$, 100), 454 (($M(^{37}Cl)+Na+CH_3CN)^+$, 35).

$^1H$ (300 MHz, $CD_3CN$): δ 8.02 (d, *J*=12.0, 2H, $H^{Ar}$), 7.92 (d, *J*=8.0, 2H, $H^{Ar}$), 7.30-7.24 (m, 4H, $H^{Ar}$), 6.50 (d, *J*=4.0, 1H, $H^{1'}$), 5.59 (dd, *J*=8.0, 4.0, 1H, $H^{3'}$), 4.88 (dt, *J*=8.0, 4.0, 1H, $H^{4'}$), 4.66 (2dd, *J*=16.0, 4.0, 2H, $H^{5'}$), 2.94-2.74 (m, 2H, $H^{2'}$), 2.44 (2s, 6H, $H^6$) ppm.

Characterisation data recorded matches previously reported literature values.[155]

**Synthesis of 1-α/β-*O*-hexyne-2-deoxy-3,5-di-*O*-*p*-toluoyl-D-ribose (36)**




**36**

To a solution of 1'-α-chloro-2'-deoxy-3,5-di-*O*-*p*-toluoyl-D-ribose (**35**) (9.28 g, 23.9 mmol) in distilled THF (50 mL), with molecular sieves (3 Å), under an argon atmosphere, was added DMAP (1.12 g, 9.2 mmol) and distilled DCM (10 mL). Meanwhile 5-hexyn-1-ol (2.03 mL, 18.4 mmol) was stirred in distilled THF (5 mL) over molecular sieves (3 Å). Both solutions were stirred for 30 min before the 5-hexyn-1-ol solution was added to the sugar mixture and the reaction was stirred for 3.5 h. The reaction solvent was removed *in vacuo* then the crude material was purified by column chromatography (100 % DCM) to afford the product **36** (4.43 g, 9.84 mmol, 53 %) as a colourless oil.

$R_f$ : 0.39, 0.47 (100 % DCM)

LRMS [ESI+, MeCN] m/z (%): 473 ($(M+Na)^+$, 100).

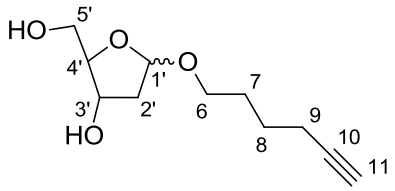
HRMS [ESI+, MeCN] for $C_{27}H_{30}Na_1O_6$ $(M+Na)^+$: calcd 473.1935, found 473.1935.

$^1H$ (400 MHz, $CDCl_3$): δ (both α and β isomers) 8.02-7.95 (m, 8H, $H^8$), 7.32-7.25 (m, 8H, $H^9$), 5.63 (ddd, *J*=7.1, 4.6, 2.5, 1H, $H^3$), 5.48 (ddd, *J*=8.2, 2.5, 2.5, 1H, $H^{3'}$), 5.38 (dd, *J*=5.6, 2.5, 1H, $H^1$), 5.33 (m, 1H, $H^{1'}$), 4.70-4.45 (m, 6H, $H^{4,4',5,5'}$), 3.81 (m, 2H, $H^{12}$), 3.49 (m, 2H, $H^{12'}$), 2.64-2.17 (m, 20H, $H^{11,2,2'15,15'}$), 1.98 (t, *J*=2.5, 2H, $H^{17,17'}$), 1.79-1.56 (m, 8H, $H^{13,13',14,14'}$) ppm.

$^{13}C$ (100 MHz, $CDCl_3$): δ (both α and β isomers) 166.1 ($C^{Ar}$), 165.9 ($C^{Ar}$), 165.8 ($C^{Ar}$), 143.7 ($C^{Ar}$), 143.6 ($C^{Ar}$), 143.5 ($C^{Ar}$), 143.4 ($C^{Ar}$), 129.4, ($CH^{Ar}$), 128.8 ($CH^{Ar}$), 126.8 ($C^{Ar}$), 104.3 ($CH^1$), 103.5 ($CH^{1'}$), 84.0 ($C^{16}$), 81.5 ($CH^4$), 80.8 ($CH^{4'}$), 75.3 ($CH^3$), 74.4 ($CH^{3'}$), 68.0 ($CH^{17}$), 67.2 ($CH_2{}^{12}$), 66.7 ($CH_2{}^{12'}$), 64.9 ($CH_2{}^5$), 64.0 ($CH_2{}^{5'}$), 39.0 ($CH_2{}^{2,2'}$), 28.4 ($CH_2{}^{13}$), 28.3 ($CH_2{}^{13'}$), 25.0 ($CH_2{}^{14}$), 24.8 ($CH_2{}^{14'}$), 21.3 ($CH_3{}^{11}$), 17.8 ($CH_2{}^{15,15'}$) ppm.

IR $\nu_{Max}/cm^{-1}$: 3296 (alkynyl ≡C-H, m), 2946, 2920, 2869 (alkyl C-H, m), 2116 (alkynyl C≡C, w), 1714 (ester C=O, s).

**Synthesis of 1-α/β-*O*-hexyne-2-deoxy-D-ribose (37)**




**37**

To a solution of 1-α/β-*O*-hexyne-2-deoxy-3,5-di-*O*-*p*-toluoyl-D-ribose (**36**) (4.23 g, 9.39 mmol) in $NH_3$:MeOH (7 M, 47 mL) was added conc. aq. $NH_3$ (5 mL) and the reaction was stirred at 45 °C 16 h. The reaction solvent was removed *in vacuo* and the crude material purified by column chromatography (1:1 EtOAc/DCM) to afford the product **37** (1.65 g, 7.71 mmol, 82 %) as a colourless oil.

$R_f$ : 0.34, 0.40 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 237 ($(M+Na)^+$, 100).

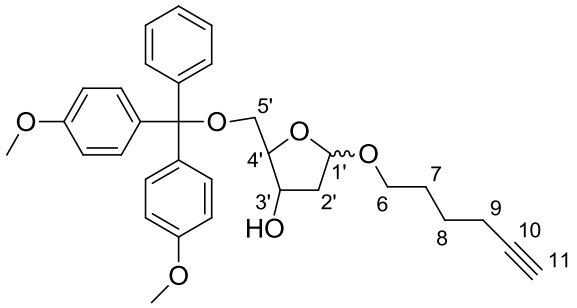
HRMS [ESI+, MeCN] for $C_{11}H_{18}Na_1O_4$ $(M+Na)^+$: calcd 237.1097, found 237.1098.

$^{1}$H (400 MHz, $d_6$-DMSO): δ (both α(a) and β(b) isomers) 5.09 (dd, *J*=5.6, 2.5, 1H, $H^{1'b}$), 5.01 (dd, *J*=5.6, 2.5, 1H, $H^{1'a}$), 4.96 (d, *J*=4.5, 1H, $^{3'b}$OH), 4.82 (d, *J*=5.1, 1H, $^{3'a}$OH), 4.60 (2t, *J*=5.6, 2H, $^{5'b}$OH, $^{5'a}$OH), 4.11 (m, 1H, $H^{3'b}$), 3.91 (dddd, *J*=8.3, 5.3x3, 1H, $H^{3'a}$), 3.72-3.66 (m, 2H, $H^{4'b,4'a}$), 3.63-3.56 (m, 2H, $H^{6}$), 3.48-3.29 (m, 6H, $H^{5'b,5'a,6'}$), 2.75 (m, 2H, $H^{11}$), 2.28 (ddd, *J*=13.6, 8.1, 5.6, 1H, $H^{2'a}$), 2.16 (dddd, *J*=6.8x3, 2.8, 4H, $H^{9,9'}$) 2.01-1.86 (m, 2H, $H^{2'b}$), 1.63-1.44 (m, 9H, $H^{2'a,7,7',8,8'}$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): δ (both α and β isomers) 103.4 ($CH^{1'b}$), 102.8 ($CH^{1'a}$), 86.9 ($CH^{4'a}$), 84.7 ($CH^{4'b}$), 84.4 ($C^{10}$), 71.2 ($CH^{11}$), 70.9 ($CH^{3'a}$), 70.0 ($CH^{3'b}$), 66.2 ($CH_2{}^{6}$), 63.2 ($CH_2{}^{5'a}$), 61.3 ($CH_2{}^{5'b}$), 41.2 ($CH_2{}^{2'b}$), 41.1 ($CH_2{}^{2'a}$), 28.3 ($CH_2{}^{7}$), 24.8 ($CH_2{}^{8}$), 17.4 ($CH_2{}^{9}$) ppm.

IR $\nu_{Max}$/cm$^{-1}$: 3377 (alcohol O-H, br, s), 3291 (alkynyl ≡C-H, s), 2936, 2870 (alkyl C-H, m), 2114 (alkynyl C≡C, w).

**Synthesis of 1'-α/β-*O*-hexyne-5'-*O*-(4-methoxyphenyl)-2'-deoxy-D-ribose (38α and 38β)**




**38**α**, 38**β

1-α/β-*O*-hexyne-2-deoxy-D-ribose (**37**) (1.51 g, 7.06 mmol) was co-evaporated with distilled pyridine (3 x 10 mL) and dissolved in distilled pyridine (20 mL) under an argon atmosphere. To the solution was added DMAP (0.17 g, 1.40 mmol). A solution of DMTCl (3.59 g, 10.6 mmol, in 15 mL distilled pyridine) was added drop-wise over 20 min and the reaction was stirred at rt for 1 hour. MeOH (20 mL) was added to quench the reaction and was stirred for 30 min. The reaction volume was reduced by half *in vacuo*, diluted with DCM (100 mL) and washed with $H_2O$ (100 mL) and $NaHCO_3$ (2 x 100 mL). The organic

phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. Following purification by column chromatography (1:4 EtOAc/ Pet ether with 0.1 % pyridine) the separated anomeric products, **38α** (1.68 g, 3.25 mmol, 47 %) and **38β** (1.51 g, 2.92 mmol, 41 %), were afforded as colourless oils in a total of 88 % yield.

α-anomer:

$R_f$ : 0.18 (1:3 EtOAc/ Pet ether)

LRMS [ESI+, MeOH] m/z (%): 1056 ($(2M+Na)^+$, 25), 539 ($(M+Na)^+$, 100).

HRMS [ESI+, MeCN] for $C_{32}H_{36}Na_1O_6$ $(M+Na)^+$: calcd 539.2404, found 539.2412.

$^1$H (400 MHz, $d_6$-DMSO): 7.42-6.88 (m, 13H, $H^{Ar}$), 5.09 (dd, *J*= 5.6, 2.5, 1H, $H^{1'}$), 4.92 (d, *J*= 5.1, 1H, $^{3'}$OH), 3.90-3.88 (m, 2H, $H^{3',4'}$), 3.75 (s, 6H, $OCH_3$), 3.68 (m, 1H, $H^6$), 3.44 (m, 1H, $H^6$), 3.10 (dd, *J*= 9.9, 2.3, 1H, $H^5$), 2.98 (dd, *J*= 9.6, 5.6, 1H, $H^5$), 2.76 (m, 1H, $H^{11}$), 2.34 (m, 1H, $H^{2'}$), 2.22 (dt, *J*=7.0, 2.8, 2H, $H^9$), 1.68-1.62 (m, 3H, $H^{2',7}$), 1.57-1.53 (m, 2H, $H^8$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): 158.0 ($C^{Ar}$), 145.0 ($C^{Ar}$), 135.7 ($C^{Ar}$), 129.7 ($CH^{Ar}$), 127.8 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 113.1 ($CH^{Ar}$), 103.0 ($CH^{1'}$), 85.2 ($C^{Ar3}$), 84.4 ($C^{10}$), 82.6 ($CH^{4'}$), 71.2 ($CH^{11}$), 70.6 ($CH^{3'}$), 66.2 ($CH_2{}^6$), 63.9 ($CH_2{}^5$), 55.0 ($OCH_3$), 41.1 ($CH_2{}^{2'}$), 28.3 ($CH_2{}^7$), 24.9 ($CH_2{}^8$), 17.4 ($CH_2{}^9$) ppm.

IR $\nu_{Max}/cm^{-1}$: 3525 (alcohol O-H, br m), 3290 (alkynyl ≡C-H, m), 2932, 2868, 2836 (alkyl C-H, m), 2114 (alkynyl C≡C, w).

β-anomer:

$R_f$ : 0.1 (1:3 EtOAc/Pet ether)

LRMS [ESI+, MeOH] m/z (%): 1056 ($(2M+Na)^+$, 7), 539 ($(M+Na)^+$, 100).

HRMS [ESI+, MeCN] for $C_{32}H_{36}Na_1O_6$ $(M+Na)^+$: calcd 539.2404, found 539.2410.

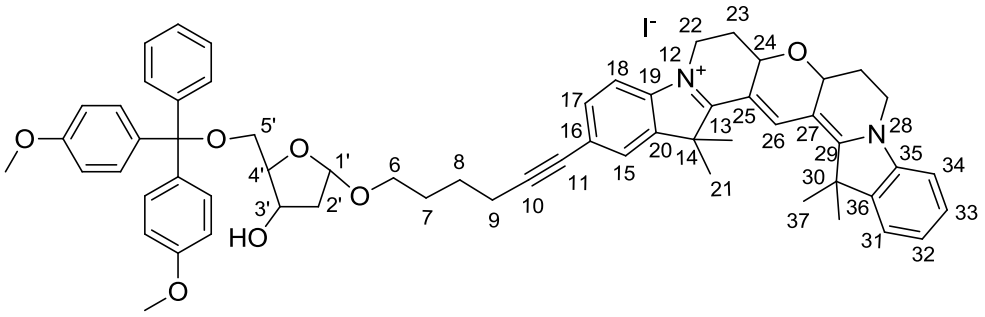
$^1$H (400 MHz, $d_6$-DMSO): 7.44-7.20 (m, 9H, $H^{Ar}$), 6.89 (d, *J*=9.1, 4H, $H^{Ar}$), 5.09 (m, 1H, $H^{1'}$), 5.04 (d, *J*= 5.1, 1H, $^{3'}$OH), 4.10 (dd, *J*= 5.1, 5.1, 1H, $H^{3'}$), 3.87 (ddd, *J*= 4.7x3, 1H, $H^{4'}$), 3.74 (s, 6H, $OCH_3$), 3.54 (m, 1H, $H^6$), 3.29 (m, 1H, $H^6$), 3.04 (dd, *J*=12.0, 4.0, 1H, $H^{5'}$), 2.96 (dd, *J*=12.0, 4.0, 1H, $H^{5'}$), 2.71 (t, *J*=2.3, 1H, $H^{11}$), 2.09-2.04 (m, 2H, $H^9$), 1.90 (ddd, *J*=12.0, 8.0, 4.0, 1H, $H^{2'}$), 1.88 (m, 1H, $H^{2'}$), 1.48-1.43 (m, 2H, $H^7$), 1.36-1.31 (m, 2H, $H^8$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): 158.0 ($C^{Ar}$), 149.6 ($CH^{Ar}$), 145.0 ($C^{Ar}$), 135.8 ($C^{Ar}$), 129.7 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 113.1 ($CH^{Ar}$), 103.3 ($CH^{1'}$), 85.2 ($C^{10}$), 84.6 ($CH^{4'}$), 84.3 ($C^{Ar3}$), 71.1 ($CH^{11}$), 70.8 ($CH^{3'}$), 66.2 ($CH_2{}^6$), 65.1 ($CH_2{}^5$), 55.0 ($OCH_3$), 41.1 ($CH_2{}^{2'}$), 28.2 ($CH_2{}^7$), 24.8 ($CH_2{}^8$), 17.4 ($CH_2{}^9$) ppm.

IR $\nu_{Max}/cm^{-1}$: 3426 (alcohol O-H, br m), 3291 (alkynyl ≡C-H, m), 2934, 2868, 2836 (alkyl C-H, m), 2114 (alkynyl C≡C, w).

The assignment of the anomers was accomplished by NMR experiments (including NOE experiment) and by agreement with literature data of α-nucleosides (J.L. Imbach[170]) and with above reported β-nucleoside NMR data.

**Synthesis of Deoxyribose-1'-α-hexyn-Cy3B (Cy3BdR) (39α)**




**39α**

To 1'-α-*O*-hexyne-5'-*O*-(4-methoxyphenyl)-2'-deoxy-D-ribose (**38α**) (1g, 1.93 mmol) was added 5-iodo-Cy3B (**26**) (1.3 g, 1.93 mmol) and CuI (74 mg, 0.386 mmol) in anhydrous DMF (10 mL) under an argon atmosphere. To this distilled $Et_3N$ (7 mL) was added and the reactions were stirred for 10 min in the dark at rt. $Pd(PPh_3)_4$ (0.22 g, 0.19 mmol) was added and the reactions were stirred for 52 h. Each reaction was diluted with DCM (200 mL) and washed with 5 % sodium EDTA (2 x 200mL) and 5 % KI solution (200 mL). The organic phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. Following purification by column chromatography (twice in 1:9 7M $NH_3$:MeOH/DCM) the products, **39α** (1.37 g, 1.29 mmol, 66 %) was afforded as a purple foam.

$R_f$ : 0.45 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 938 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{61}H_{65}N_2O_7$ $(M)^+$: calcd 937.4786, found 937.4781.

$^1$H (400 MHz, $d_6$-DMSO): 7.98 (s, 1H, $H^{26}$), 7.66-7.19 (m, 16H, $H^{Ar}$), 6.88 (d, *J*=12.0, 4H, $H^{Ar}$), 5.11 (dd, *J*=5.8, 2.5, 1H, $H^{1'}$), 4.95 (d, *J*=4.9, 1H, $^{3'}$OH), 4.67-4.62 (m, 2H, $H^{24}$), 4.37-4.28 (m, 2H, $H^{22}$), 3.93-3.85 (m, 4H, $H^{3',4',22}$), 3.73-3.70 (m, 7H, $OCH_3$,$H^6$), 3.48 (m, 1H, $H^6$), 3.10 (m, 1H, $H^{5'}$), 2.98 (dd, *J*=9.8, 5.5,

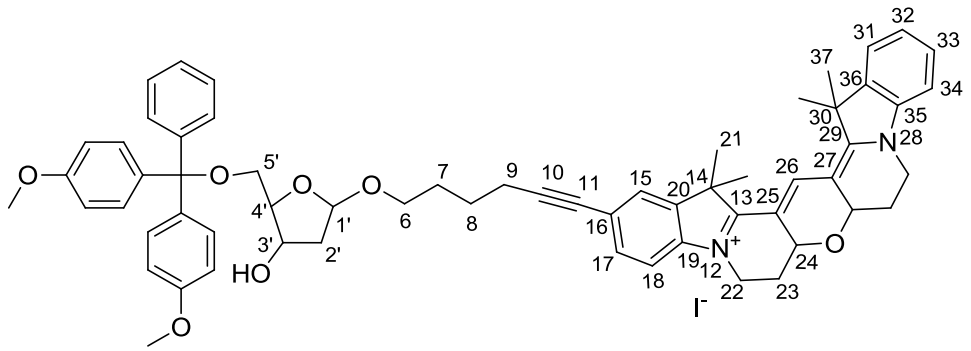
1H, $H^{5'}$). 2.53-2.50 (m, 4H, $H^{23,9}$), 2.35 (m, 1H, $H^{2'}$), 1.96 (dt, *J*=11.9, 4.6, 2H, $H^{23}$), 1.70-1.67 (m, 17H, $H^{21,37,2',7,8}$) ppm.

$^{13}$C (100 MHz, $d_6$-DMSO): 168.0 ($C^{13/29}$), 166.8 ($C^{13/29}$), 158.0 ($C^{Ar}$), 145.0 ($C^{Ar}$), 141.6 ($C^{Ar}$), 141.2 ($C^{Ar}$), 140.8 ($C^{Ar}$), 140.7 ($C^{Ar}$), 137.0 ($C^{Ar}$), 135.7 ($CH^{26}$), 131.9 ($CH^{Ar}$), 129.7 ($CH^{Ar}$), 128.6 ($CH^{Ar}$), 127.8 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 125.5 ($CH^{Ar}$), 125.3 ($CH^{Ar}$), 122.4 ($CH^{Ar}$), 119.7 ($C^{16}$), 113.1 ($CH^{Ar}$), 111.2 ($CH^{Ar}$), 110.9 ($CH^{Ar}$), 110.1 ($C^{25/27}$), 109.3 ($C^{25/27}$), 103.0 ($CH^{1'}$), 91.1 ($C^{10/11}$), 85.2 ($C^{Ar_3}$), 82.6 ($CH^{4'}$), 80.6 ($C^{10/11}$), 70.6 ($CH^{3'}$), 69.3 ($CH^{24}$), 66.3 ($CH_2^6$), 64.0 ($CH_2^{5'}$), 55.0 ($OCH_3$), 48.5 ($C^{14/30}$), 48.0 ($C^{14/30}$), 41.2 ($CH_2^{2'}$), 41.1 ($CH_2^{22}$), 41.0 ($CH_2^{22}$), 28.6 ($CH_2^7$), 27.5 ($CH_3^{21}$), 26.9 ($CH_3^{37}$), 26.1 ($CH_2^8$), 25.1 ($CH_2^{23}$), 18.5 ($CH_2^9$) ppm.

Mp: >150 °C (decomposes)

IR $\nu_{Max}/cm^{-1}$: 3336 (alcohol O-H, br m), 2929, 2866, 2835 (alkyl C-H, m).

**Synthesis of Deoxyribose-1'-β-hexyn-Cy3B (Cy3BdR) (39β)**




**39**β

This was prepared using the procedure as detailed above for **39*α*** from 1'-β-*O*-hexyne-5'-*O*-(4-methoxyphenyl)-2'-deoxy-D-ribose (38β) (1g, 1.93 mmol) to give the product (**39β**) (1.51 g, 1.42 mmol, 73 %) as a purple foam.

$R_f$: 0.34 (1:9 MeOH/DCM)

LRMS [ESI+, MeCN] m/z (%): 938 ($(M)^+$, 100).

HRMS [ESI+, MeCN] for $C_{61}H_{65}N_2O_7$ $(M)^+$: calcd 937.4786, found 937.4788

$^1$H (400 MHz, $d_6$-DMSO): 7.99 (s, 1H, $H^{26}$), 7.65-7.18 (m, 16H, $H^{Ar}$), 6.88 (d, *J*=8.1, 4H, $H^{Ar}$), 5.12 (d, *J*=3.7, 1H, $H^{1'}$), 5.04 (d, *J*= 5.1, 1H, $^{3'}$OH), 4.65 (2t, *J*=5.7,

2H, $H^{24}$), 4.36-4.28 (m, 2H, $H^{22}$), 4.10 (m, 1H, $H^{3'}$), 3.97-3.85 (m, 3H, $H^{4',22}$), 3.73 (d, *J*=0.6, 6H, $OCH_3$), 3.61 (m, 1H, $H^6$), 3.36 (m, 1H, $H^6$), 3.06 (dd, *J*=8.0, 4.0, 1H, $H^{5'}$), 2.97 (dd, *J*=12.0, 8.0, 1H, $H^{5'}$), 2.50-2.49 (m, 2H, $H^{23}$), 2.36 (t, *J*=6.6, 2H, $H^9$), 1.97-1.88 (m, 4H, $H^{2',23}$), 1.70 (s, 6H, $H^{21}$), 1.68 (s, 6H, $H^{37}$), 1.55-1.45 (m, 4H, $H^{7,8}$) ppm.

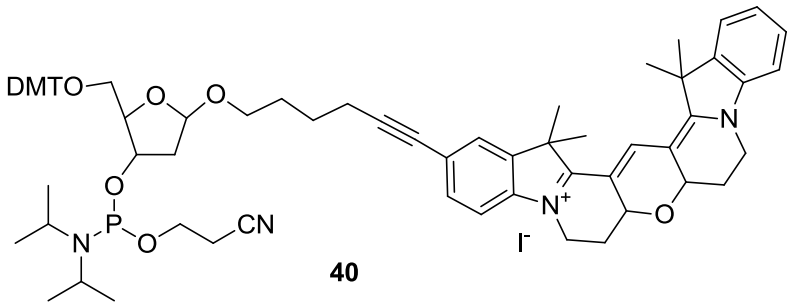
$^{13}C$ (100 MHz, $d_6$-DMSO): 168.0 ($C^{13/29}$), 166.8 ($C^{13/29}$), 158.0 ($C^{Ar}$), 145.0 ($C^{Ar}$), 141.6 ($C^{Ar}$), 141.2 ($C^{Ar}$), 140.7 ($C^{Ar}$), 140.7 ($C^{Ar}$), 137.0 ($CH^{26}$), 135.7 ($C^{Ar}$), 131.8 ($CH^{Ar}$), 129.7 ($CH^{Ar}$), 128.6 ($CH^{Ar}$), 127.7 ($CH^{Ar}$), 126.6 ($CH^{Ar}$), 125.5 ($CH^{Ar}$), 125.3 ($CH^{Ar}$), 122.4 ($CH^{Ar}$), 119.7 ($C^{16}$), 113.1 ($CH^{Ar}$), 111.3 ($CH^{Ar}$), 110.9 ($CH^{Ar}$), 110.1 ($C^{25/27}$), 109.3 ($C^{25/27}$), 103.4 ($CH^{1'}$), 91.0 ($C^{10/11}$), 85.2 ($C^{Ar_3}$), 84.6 ($CH^{4'}$), 80.5 ($C^{10/11}$), 70.8 ($CH^{3'}$), 69.3 ($CH^{24}$), 66.3 ($CH_2{}^6$), 65.1 ($CH_2{}^{5'}$), 55.0 ($OCH_3$), 48.5 ($C^{14/30}$), 48.0 ($C^{14/30}$), 41.2 ($CH_2{}^{2'}$), 41.1 ($CH_2{}^{22}$), 41.1 ($CH_2{}^{22}$), 28.5 ($CH_2{}^7$), 27.5 ($CH_3{}^{21}$), 26.9 ($CH_3{}^{37}$), 26.1 ($CH_2{}^8$), 25.1 ($CH_2{}^{23}$), 18.5 ($CH_2{}^9$) ppm.

Mp: >150 °C (decomposes)

IR $\nu_{Max}/cm^{-1}$: 3336 (alcohol O-H, br m), 2929, 2865, 2835 (alkyl C-H, m).

UV/Vis (MeOH): $A_{max}$ 571 nm, ε 130,000 $M^{-1}cm^{-1}$, $Em_{max}$ 585 nm.

**Synthesis of Cy3BdR phosphoramidite (40)**



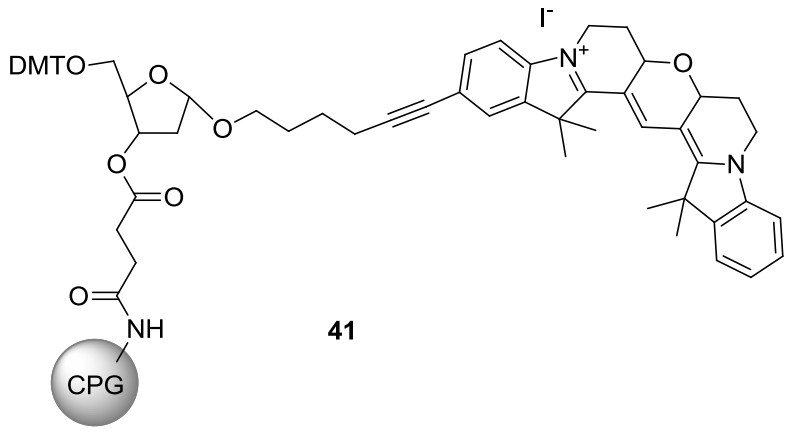

**40**

Deoxyribose-1'-β-hexyn-Cy3B (Cy3BdR) (**39β**) (0.50 g, 0.47 mmol) was co-evaporated with distilled pyridine (3 x 5mL) and distilled DCM (3 x 5mL) then dissolved in distilled DCM (2 mL), with activated molecular sieves (3 Å), under an argon atmosphere. To this was added distilled DIPEA (0.20 mL, 1.18 mmol) and the reaction was stirred for 10 min. 2-Cyanoethyl *N*,*N*-diisopropylchloro phosphoramidite (0.13 mL, 0.56 mmol) was added drop-wise and the reaction was stirred for 45 min. The reaction mixture was diluted with distilled DCM (10 mL) and washed with deoxygenated sat. KCl (10 mL). The organic

phase was dried ($Na_2SO_4$) and the solvent removed *in vacuo*. The crude material was precipitated in deoxygenated hexane (100 mL) from a minimum of distilled DCM (1 mL). The precipitation procedure was repeated twice and the final precipitate was co-evaporated with distilled DCM (2 mL) and trace solvents were removed *in vacuo* affording the product **40** (0.56 g, 0.44 mmol, 94 %) as a dark pink iridescent solid.

$R_f$ : 0.51 (1:9 MeOH/DCM with 1 % pyridine)
LRMS [ESI+, MeCN] m/z (%): 1138 ($(M)^+$, 100).
$^{31}P$ (300 MHz, $CD_3CN$): 148.8 (s, $^{31}P$).

**Synthesis of Cy3BdR resin (41)** (experimental procedure based on literature[171,172])




**41**

The resin used for coupling was AminoSynBase 1000 Å (loading 59 µmol/g) purchased from Link Technologies. All steps were carried out under argon using dry solvents.

Resin (250 mg, 0.015 mmol) was activated by washing with 3 % TCA in DCM (3 mL) and suspending in 3 % TCA in DCM (3 mL) for 4 h at rt. The TCA was removed by filtration and the resin washed with $Et_3N$:DIPEA (9:1) (3 x 3 mL), DCM (3 x 3 mL) and $Et_2O$ (3 x 3 mL). The resin was dried on high vacuum for 1 h then soaked in pyridine (2 mL) for 10 min. To the suspension was added succinic anhydride (112 mg, 1.10 mmol) and 4-DMAP (23.0 mg, 0.189 mmol) in pyridine (2 mL) under an argon atmosphere and the vessel was rotated at rt for 20 h. The reaction solvent was removed and the resin washed with pyridine (3 x 3 mL), DCM (3 x 3 mL) and $Et_2O$ (3 x 3 mL). The resin was dried over high vacuum for 1 h before soaking in pyridine (3 mL) for 10 min. This was removed and a solution of EDC (46.0 mg, 0.30 mmol), 4-DMAP (2.0 mg, 0.015 mmol), $Et_3N$ (0.01 mL)

and deoxyribose-1'-α-hexyn-Cy3B (Cy3BdR) (**39α**) (60.0 mg, 0.056 mmol) in pyridine (3 mL) was added. The vessel was rotated under argon for 20 h at rt.

Pentachlorophenol (20.0 mg, 0.074 mmol) was added to the existing mixture and rotated for a further 3 h. The solvent was removed and the resin washed with pyridine (3 x 3 mL), DCM (3 x 3 mL) and $Et_2O$ (3 x 3 mL). To the resin was added piperidine (10 % in DMF, 3 mL) and the vessel rotated for 1.5 min. The solvent was rapidly removed and the resin washed with DCM (3 x 3 mL), $Et_2O$ (3 x 3 mL) and THF (3 x 3 mL). Capping reagent ($Ac_2O$ /pyridine/THF:*N*-methylimidazole 1:1, 2 mL) was added and the vessel rotated for 1 h. The solution was removed and the resin washed with THF (3 x 3 mL), pyridine (3 x 3 mL), DCM (3 x 3 mL) and $Et_2O$ (3 x 3 mL) before drying *in vacuo* to afford the resin **41** with a loading of 13 μmol/g.

Loading was calculated by the trityl yield method[171].

## 6.2. Oligonucleotide synthesis and purification

### 6.2.1 General

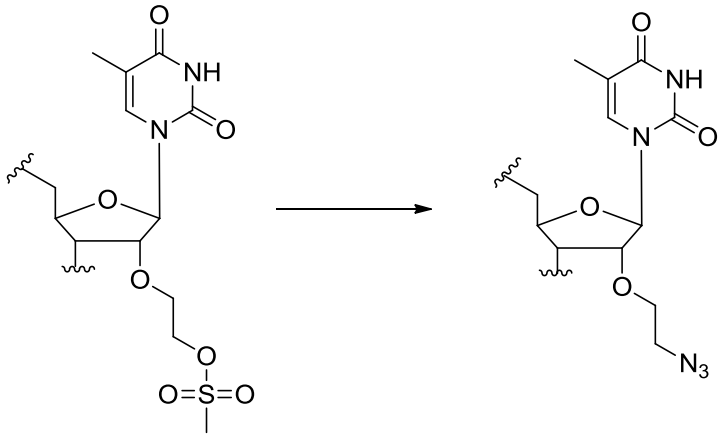
DNA reagents including standard DNA phosphoramidites and solid supports were purchased from Link Technologies Ltd. Oligonucleotides were synthesised on an Applied Biosystems 394 automated DNA/RNA synthesizer. Phosphoramidite cycles, including acid-catalyzed detritylation, coupling, capping and iodine oxidation steps, were undertaken in 0.2 or 1.0 µmole scale. Standard DNA phoshoramidites were used for the majority of oligonucleotide sequences, however fast-deprotecting monomers, dmf-dG and Ac-dC, were used for those sequences containing Cy5dT. Coupling efficiencies and overall oligonucleotide yields were determined by the automated trityl cation conductivity monitoring facility of the synthesiser and were >98.0 % for all cases. Phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for A, G, C and T monomers was set to 25 s however the coupling time for the modified monomers was extended to 360 s. The cleavage of oligonucleotides from the solid support and subsequent deprotection was undertaken by suspending the resin in concentrated aqueous ammonia for 60 min at rt followed by heating in a sealed tube for 5 h at 55 °C. For oligonucleotides containing Cy5dT the deprotection time was reduced to 1h at 55 °C.

A Gilson HPLC system with ABI Aquapore C8 column (8 mm x 250 mm, pore size 300 Å) was used to purify the oligonucleotides by reverse phase. The following protocol was used for the majority of oligonucleotides: run time 20 min, flow rate 4 mL per min, gradient: time in min (% buffer B);0 (0); 3 (0); 3.5 (25); 15 (70); 16 (100); 17 (100); 17.5 (0); 20 (0). For the hydrophobic Cy3B oligonucleotides a modified protocol was used: run time 25 min, flow rate 4mL per min, gradient: time in min (% buffer B); 0 (0); 3 (0); 3.5 (15); 15 (60); 16 (100); 22 (100); 22.5 (0); 25 (0). Elution buffer A: 0.1 M ammonium acetate, pH 7.0, buffer B: 0.1 M ammonium acetate with 50 % acetonitrile pH 7.0. Elution of the oligonucleotides was monitored by ultraviolet absorption at 295 nm and the main peak was collected then desalted using NAP-10 gel filtration columns (GE Healthcare), aliquoted into eppendorf tubes and stored at –20 ºC. All oligonucleotides were characterised either by MALDI-TOF mass spectrometry or MicroTOF electrospray mass spectrometry and capillary electrophoresis.

## 6.2.2. Oligonucleotide click labelling procedures

### *6.2.2.1 Minor groove oligonucleotide labelling procedure (on-resin)*

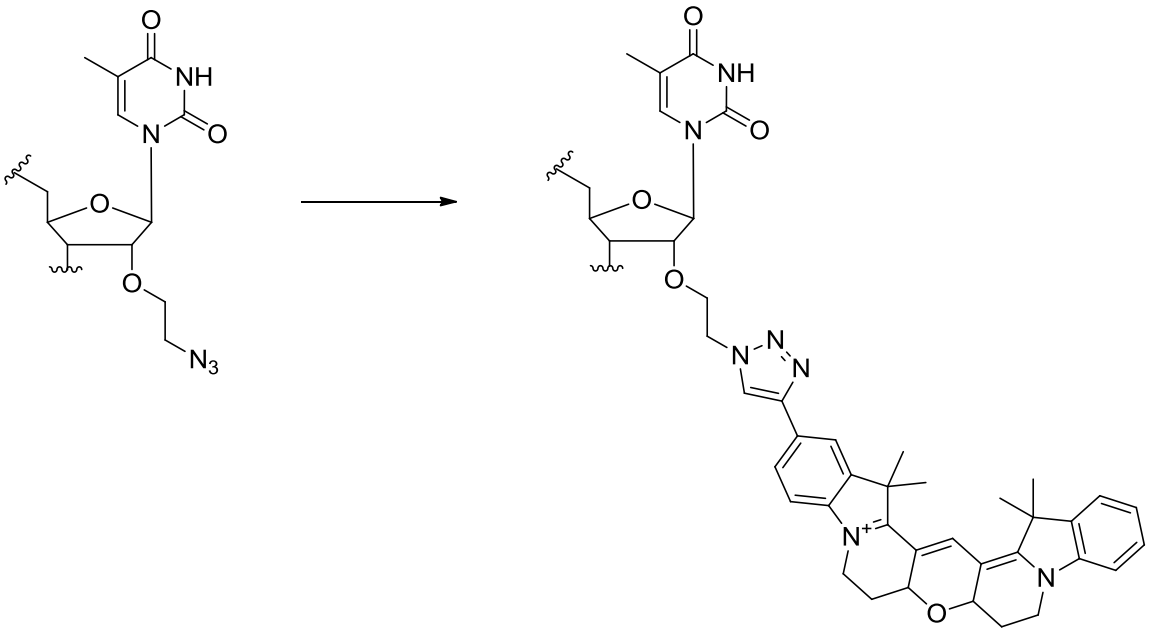
**Mesyl to azide conversion** (experimental procedure based on literature[99])




Oligonucleotides containing 2'mesyloxyethyl rT phosphoramidite monomer were synthesised on a 1μmol scale by standard solid-phase synthesis. These were left on-resin and the contents were transferred into glass vials (0.7-0.8 μmol).

To each was added sodium azide (0.17 g, 2.62 mmol). 18-crown-6 (0.05 g, 0.19 mmol) was added in anhydrous DMF (1 mL) and the suspension was left at 65 °C for 20 h. The mixture was allowed to cool to rt and each oligonucleotide resin was washed with anhydrous DMF (3 x 1 mL), 1:1 EtOH:$H_2O$ (4 x 1mL), dry MeCN (3 x 1 mL) and $Et_2O$ (3 x 1 mL). The resin was dried under a warm gentle air flow and was deprotected in $NH_3$ (0.5 mL) at rt for 16 h. The solution was removed, the resin washed as previously and the solution dried *in vacuo*. The resulting oligonucleotide was desalted by NAP-25 gel-filtration.

**Click chemistry dye labelling** (experimental procedure based on literature[99])



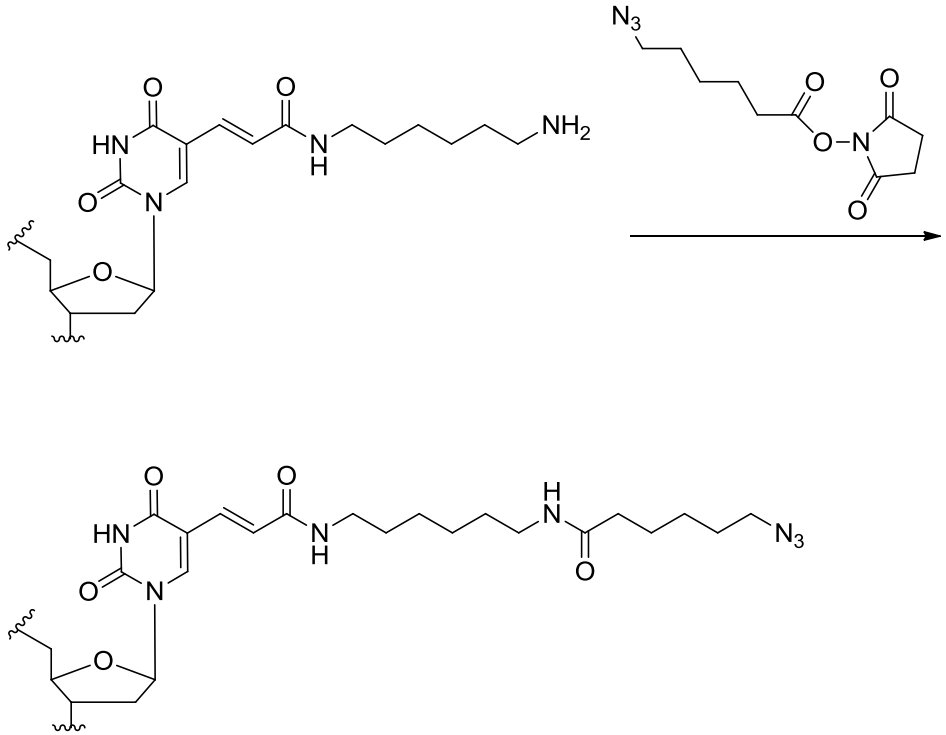
The oligonucleotide in solution (0.04 μmol) was freeze-dried and re-dissolved in NaCl (0.2 M, 44 μL) under an argon atmosphere.

A stock solution was prepared containing tris-hydroxypropyl ligand, sodium ascorbate and $CuSO_4.5H_2O$ (35:50:5 mol:mol). This was added to the oligonucleotide followed by 5-ethynyl-Cy3B (**32**) (in DMF) to give a final reaction containing 35 eq. tris hydroxypropyl ligand, 50 eq. sodium ascorbate, 5 eq. $CuSO_4.5H_2O$ and 5 eq. **32** with a total reaction volume of 60μL. The reaction was purged with argon and left at rt in the dark for 2 h.

The reaction volume was increased to 1 mL ($H_2O$) and purified from free dye by NAP-10 gel filtration (2x). Oligonucleotides were purified by RP-HPLC (in some cases followed by polyacrylamide gel electrophoresis (20 %) if required).

### *6.2.2.2 Major groove oligonucleotide labelling procedure (solution phase)*

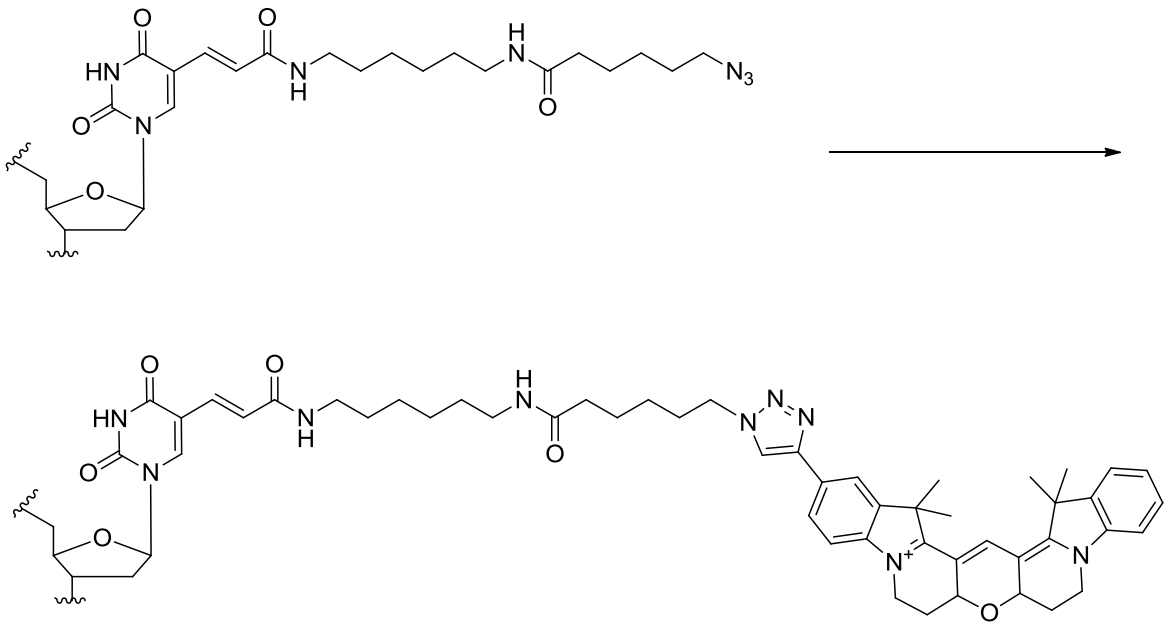
**Mesyl to azide conversion** (experimental procedure based on literature[99])



Oligonucleotides containing amino C6 dT phosphoramidite monomer were synthesised by standard solid-phase techniques and were purified by ion-exchange HPLC.

To each oligonucleotide (0.06-0.17 µmoles) dissolved in 0.5 M $Na_2CO_3$/$NaHCO_3$ (pH 8.75, 80 µL) was added azide active ester (20 eq. from azide active ester stock; 1 mg in 53 µL DMSO) and DMSO to give a total reaction volume of 120 µL. The reactions were left at rt in the dark for 2 h. Reaction volumes were increased to 1 mL ($H_2O$) and were purified by NAP-10 gel filtration before RP-HPLC purification.

**Click chemistry dye labelling** (experimental procedure based on literature[99])



The following stock solutions were prepared immediately prior to reaction commencement; 5-alkyne-Cy3B (**32**) (5 mg in 140 μL DMF), $CuSO_4.5H_2O$ (12.5 mg in 1 mL 0.2 M NaCl), sodium ascorbate (100 mg in 1 mL 0.2 M NaCl) and tris-hyrdoxypropyl ligand (3 mg in 30 μL 0.2 M NaCl). A separate stock solution was prepared under an argon atmosphere containing tris-hydroxypropyl ligand: sodium ascorbate: $CuSO_4.5H_2O$ (35:50:5 mol:mol: mol-oligonucleotide).

To oligonucleotides dissolved in 0.2 M NaCl under an argon atmosphere was added tris-hydroxypropyl ligand: sodium ascorbate: $CuSO_4.5H_2O$ stock (35 eq, 50 eq, 5 eq respectively) and 5-ethynyl-Cy3B (**32**) (50 eq in DMF) to give a total reaction volume of 100 μL. The reaction was left at 55 °C for 2 h.

The reactions were allowed to cool to rt and the reaction volume increased to 1 mL ($H_2O$) before purifying by NAP-10 gel-filtration. The oligonucleotides were purified by RP-HPLC.

## 6.3. Biophysical studies

### 6.3.1 UV analysis

#### *6.3.1.1 UV scan analysis*

Ultraviolet/visible absorption spectra were determined using a Cary 50 Bio UV-Vis spectrophotometer with a scan range of 200-800 nm. Samples were analysed in a 1 mL cuvette (Hellma synthetic quartz 'precision cell QG'; 1 mL volume, 10 mm pathlength). Samples were prepared according to sample conditions 1 or 2 as specified in experimental discussion; (1) 1 μM oligonucleotide concentration in phosphate buffer (25 mM) with a total of 100 mM NaCl at pH 7; (2) 0.15 μM oligonucleotide concentration in GoTaq colourless PCR buffer with a total of 3 mM $MgCl_2$.

#### *6.3.1.2 Extinction coefficient calculation*

Measurements were made on a Cary 4000 UV-Vis spectrometer with Cary temperature controller using Cary Win UV Scan software. Extinction coefficients of compounds (in MeOH or EtOH) were calculated from four readings of UV absorbance at the UV absorbance maxima using the Beer-Lambert law, $A=\varepsilon cl$ (where A is absorbance at UV $A_{max}$, c is concentration in M, l is pathlength 1cm and ε is extinction coefficient $M^{-1}cm^{-1}$). Samples were analysed in a 1 mL cuvette (Hellma synthetic quartz 'precision cell QG'; 1 mL volume, 10 mm pathlength).

In the case of the CyDyes, samples were dissolved in EtOH and passed through Dowex-1-chloride (Dowex 1x2-200 ion exchange resin) to convert all counter ions to $Cl^-$. Samples were then dried in a heating pistol over $P_2O_5$ at 50 ºC for 48 h before a stock solution was prepared of each (1 mg/ 10 mL MeOH). Compounds were diluted from the stock solution to give four readings below an absorbance value of 1. Absorbance values were plotted against concentration and a straight line was drawn through the points intercepting the Y-axis at 0. The extinction coefficient was taken as the gradient of the line.

### *6.3.1.3 UV Melting Analysis*

Measurements were made on a Cary 4000 UV-Vis spectrometer with Cary temperature controller. Cary Win UV Thermal software was used with an absorption wavelength of 260 nm. Samples were analysed in 1 mL cuvettes (Hellma synthetic quartz 'precision cell QG'; 1 mL volume, 10 mm pathlength) and were made to 1 μM oligonucleotide concentration in phosphate buffer (10 mM) with NaCl (100 mM) at pH 7. Three successive melting curves were measured, and average $T_m$ values were calculated with Cary Win UV Thermal application software.

The thermal protocol used was as follows:

Start T 20 °C, end T 20 °C. Thermal protocol as follows:

| **Stage** | **Data interval (°C)** | **Rate (°C/min)** | **End Temp (°C)** | **Hold time (min)** |
|---|---|---|---|---|
| 1 | 1.00 | 10.00 | 84.00 | 2.00 |
| 2 | 0.10 | 1.00 | 20.00 | 2.00 |
| 3 | 0.10 | 1.00 | 84.00 | 2.00 |
| 4 | 0.10 | 1.00 | 20.00 | 2.00 |
| 5 | 0.10 | 1.00 | 84.00 | 2.00 |
| 6 | 0.10 | 1.00 | 20.00 | 2.00 |
| 7 | 0.10 | 1.00 | 84.00 | 2.00 |
| 8 | 0.10 | 10.00 | 20.00 | 0.00 |

### *6.3.1.4 Circular Dichroism Analysis*

Circular dichroism spectra were recorded on a Jasco J-720 spectropolarimeter. Samples were made to 1 μM oligonucleotide concentration in phosphate buffer (10 mM) with NaCl (100 mM) at pH 7. Spectra were recorded between 200-320 nm at a rate of 100 nm/min, with step resolution of 0.2 nm, bandwidth of 1.0 nm and sensitivity of 50 mdeg.

10 successive spectra were recorded and an average was taken. A blank (buffer) baseline was subtracted from each spectrum.

### 6.3.2 Fluorescence analysis

#### *6.3.2.1 Fluorescence scan analysis*

Measurements were made on a Perkin Elmer LS50B fluorimeter equipped with a Perkin Elmer PTP-1 Peltier system. FLWinlab TempScan software was used with settings of 300 nm/s scan speed. Optimum excitation wavelength, scan range wavelengths and excitation/emission slit width were used for each individual sample. In cases where multiple samples were compared the slit width and excitation wavelength settings were kept constant. Where multiple dyes were compared excitation wavelength was taken as the absorption maxima or ¾ absorbance maxima as specified in the experimental discussion.

Samples were analysed in a 200 µL cuvette (Hellma quartz 'SUPRASIL QS'; 200 µL volume, 10 mm pathlength) with a collection angle of 90°. A total sample volume of 200 µL was used, with labelled oligonucleotide concentration of 0.15 µM (0.5 µM for target strand, or 0.15 µM in cases where target strand was also dye-containing). Samples were prepared according to either sample conditions (1) or (2); (1) phosphate buffer (25 mM) with a total of 100 mM NaCl at pH 7; (2) GoTaq colourless PCR buffer (Promega) with a total of 3 mM $MgCl_2$ at pH 8.5. The oligonucleotide samples were heated to 70°C to denature for three min before being allowed to slowly cool to rt to anneal.

#### *6.3.2.2 Fluorescence melting analysis*

For single wavelength analysis, measurements were made on a Perkin Elmer LS50B fluorimeter equipped with a Perkin Elmer PTP-1 Peltier system. FLWinlab TempScan software was used with optimum excitation/emission slit widths, excitation wavelength and emission wavelength settings for each individual sample. In cases where multiple samples were compared the slit width and excitation wavelength settings were kept constant. Where multiple dyes were compared excitation wavelength was taken as the absorption maxima or ¾ absorbance maxima as specified in experimental discussion.

Samples were analysed in a 200 µL cuvette (Hellma quartz 'SUPRASIL QS'; 200 µL volume, 10 mm pathlength) with a collection angle of 90°. A total sample volume of

200 µL was used, with labelled oligonucleotide concentration of 0.15 µM (0.5 µM for target strand, or 0.15 µM in cases where target strand was also dye-containing). Samples were prepared in GoTaq colourless PCR buffer (Promega) with a total of 3 mM $MgCl_2$ at pH 8.5. The oligonucleotide samples were heated to 70°C to denature for three min before being allowed to slowly cool to rt to anneal.
Samples were heated from 30°C to 80°C with a step size of 1°C and 30 s equilibration, the emission intensity was recorded at each step.

For fixed wavelength analysis, fluorescence melting was carried out using a RotorGene-3000 with settings of excitation 530 nm and detection 585 nm with a gain of 5. Samples were analysed in RotorGene thin-walled PCR tubes (0.1 mL. Qiagen). A total sample volume of 20 µL was used, with labelled oligonucleotide concentration of 0.15 µM (0.5 µM for target strand). Samples were prepared in GoTaq colourless PCR buffer with a total of 3 mM $MgCl_2$ at pH 8.5.

Samples were heated to 95°C for 2 min then cooled to 30°C. After holding at 30°C for 1 min the samples were heated from 30°C-95°C with a 0.5°C step (15 sec for first step, 5 sec per step thereafter). Fluorescence emission intensity was recorded at each step.

#### *6.3.2.3 FRET experiments*

Duplexes between Cy5dT oligonucleotide (ODN44) and Cy3dT oligonucleotides (ODN20, ODN21, ODN22, ODN23, ODN24, ODN25, ODN26, ODN27, ODN28, ODN29, ODN30) were prepared in a 1:1 ratio at 0.2 µM concentration of each oligonucleotide. Control samples were also prepared to show the oligonucleotides against an unmodified complement (ODN32 with ODN44, ODN33 with ODN20-30).
A total sample volume of 200 µL was used with sample buffer of 25 mM sodium phosphate with 100 mM NaCl at pH 7.5.

Samples were scanned for UV absorbance from 200-800 nm (a background of buffer only was subtracted automatically from each spectrum). FRET samples were scanned for fluorescence emission from 535-800 nm (excitation 530 nm and 615 nm) with scan speed of 300 nm/min, slit widths of 15 nm excitation slit and 20 nm emission slit. Control samples were scanned with the same conditions.

FRET was calculated by:

$$\text{FRET efficiency} = (I_{FRET} \times Abs_{615})/(I_{direx} \times Abs_{530})$$

$I_{FRET}$ is the emission intensity of the Cy5 peak when excited via fret (ex 530 nm), $I_{direx}$ is the emission intensity of Cy5 peak when directly excited (ex 615 nm), $Abs_{615}$ is the absorbance of the sample at 615 nm and $Abs_{530}$ is the absorbance of the sample at 530 nm.

#### *6.3.2.4 Quantum yield calculation*

Quantum yields were measured in a 1 mL cuvette (Hellma quartz 'SUPRASIL QS'; 1 mL volume, 10 mm pathlength) and were measured in reference to Rh101 (QY = 1.0 at 25°C in EtOH in range $\lambda_{ex}$ 450-565).[173] Samples were initially scanned for OD at the single wavelength of 545 nm using a Cary 50 Bio UV-Vis spectrophotometer. All samples were prepared by dissolving the fluorescent compound in MeOH (analytical grade) and then sequentially diluting with MeOH until a concentration of 0.01 ± 0.008 OD (545 nm) was reached and recorded. In the lone case of the CPP monomer the reference used was quinine sulphate (QY= 0.58 in 0.01M $H_2SO_4$, excitation 350 nm).[174]

Fluorescence emission measurements were made on a Perkin Elmer LS50B fluorimeter equipped with a Perkin Elmer PTP-1 Peltier system. FL Winlab Scan software was used with an emission scan range of 545-800 nm and excitation of 545 nm. Slit widths of excitation slit 3.5 nm and emission slit 4 nm were used together with a scan speed of 300 nm/min. Emission readings were taken at 25°C and the area of each sample was calculated relative to its own baseline. Sample and reference measurements were repeated three times, with an average taken for each compound. The following equation was used to calculate the final quantum yield values. (QY=quantum yield, I=fluorescence emission area, OD=optical density, η=refractive index of solvent, R=reference and S=sample).

$$QY_S = QY_R \times (I_S/I_R) \times (OD_R/OD_S) \times (\eta_S^2/\eta_R^2)$$

### 6.3.3 Polymerase Chain Reaction (PCR) protocols

All PCR reaction samples were prepared under sterilised conditions in a LabCaire PCR workstation. A negative control sample was prepared for each sample (no DNA template) and samples were prepared in triplicate. Reactions were undertaken using a BioRad CFX96 Real-Time PCR Detection System, with CFX Manager software (BioRad) monitoring on channel 3 (excitation range 560-590 nm, detector range 610-650 nm). Reactions were run in 0.2 mL low-profile white 8-tube strips with optically clear lids (BioRad). Sample preparation and thermal protocols differed for each PCR probe as detailed below.

**Asymmetric PCR- Molecular Beacon**

Sample conditions

Template: 1 pg per sample (ODN45 WT, ODN46 MT)

Forward primer: 0.05 μM (ODN60)

Reverse primer: 0.5 μM (ODN61)

Probe: 0.15 μM (ODN71/ODN76/ODN78)

Buffer: PrecisionHRM mastermix (2x) (in reaction 0.25U Taq DNA polymerase, 2.5 mM $Mg^{2+}$, 100 μM each dNTP)

Total volume: 20 μL

Thermal protocol

| Phase | Cycles | T/ °C | Hold/ sec | Measure |
|---|---|---|---|---|
| Activation | 1 | 95 | 480 | No |
| Cycling | 50 | 95 | 15 | No |
| Cycling | 50 | 58 | 15 | Yes |
| Cycling | 50 | 72 | 15 | No |

Followed by melt from 35-80°C.

**Symmetric PCR- Scorpion**

Sample conditions

Template: 1 pg per sample (ODN45 WT, ODN46 MT)

Forward primer: 0.5 μM (ODN60)

Reverse primer/Scorpion: 0.5 μM (ODN73/ODN77)

Buffer: PrecisionHRM mastermix (2x) (in reaction 0.25U Taq DNA polymerase, 2.5 mM $Mg^{2+}$, 100 μM each dNTP)

Total volume: 20 μL

Thermal protocol

| **Phase** | **Cycles** | **T/ °C** | **Hold/ sec** | **Measure** |
|---|---|---|---|---|
| Activation | 1 | 95 | 480 | No |
| Cycling | 40 | 95 | 15 | No |
| Cycling | 40 | 55 | 15 | No |
| Cycling | 40 | 61 | 30 | Yes |

**Symmetric PCR- Taqman**

Sample conditions

Template: 1 pg per sample (ODN45 WT, ODN46 MT)

Forward primer: 0.5 μM (ODN74)

Reverse primer: 0.5 μM (ODN61)

Probe: 0.15 μM (ODN72)

Buffer: PrecisionHRM mastermix (2x) (in reaction 0.25U Taq DNA polymerase, 2.5 mM $Mg^{2+}$, 100 μM each dNTP)

Total volume: 20 μL

Thermal protocol

| **Phase** | **Cycles** | **T/ °C** | **Hold/ sec** | **Measure** |
|---|---|---|---|---|
| Activation | 1 | 95 | 480 | No |
| Cycling | 40 | 95 | 15 | No |
| Cycling | 40 | 55 | 15 | Yes |
| Cycling | 40 | 68 | 30 | No |

**Asymmetric PCR- HyBeacon**

Sample conditions

Template: 1 pg per sample (ODN45 WT, ODN46 MT)

Forward primer: 0.5 μM (ODN60)

Reverse primer: 0.05 μM (ODN61)

Probe: 0.15 μM (ODN63)

Buffer: PrecisionHRM mastermix (2x) (in reaction 0.25U Taq DNA polymerase, 2.5 mM $Mg^{2+}$, 100 μM each dNTP)

Total volume: 20 μL

Thermal protocol

| **Phase** | **Cycles** | **T/ °C** | **Hold/ sec** | **Measure** |
|---|---|---|---|---|
| Activation | 1 | 95 | 480 | No |
| Cycling | 50 | 95 | 15 | No |
| Cycling | 50 | 50 | 15 | Yes |
| Cycling | 50 | 72 | 15 | No |

Followed by melt from 35-80°C on rotorgene3000.

**Asymmetric PCR- Labelled Primer**

Sample conditions

Template: 1 pg per sample (ODN45 WT)

Labelled forward primer: 0.5 μM (ODN67)

Reverse primer: 0.05 μM (ODN61)

SYBRgreen: 1x (from 10,000x stock)

Buffer: PrecisionHRM mastermix (2x) (in reaction 0.25U Taq DNA polymerase, 2.5 mM $Mg^{2+}$, 100 μM each dNTP)

Total volume: 20 μL

Thermal protocol

| **Phase** | **Cycles** | **T/ °C** | **Hold/ sec** | **Measure** |
|---|---|---|---|---|
| Activation | 1 | 95 | 480 | No |
| Cycling | 50 | 95 | 15 | No |
| Cycling | 50 | 50 | 15 | Yes* |
| Cycling | 50 | 72 | 15 | No |

*Monitoring on SYBRgreen excitation and detection channel.

Followed by agarose gel electrophoresis analysis of product.

### 6.3.4. Gel electrophoresis

For gel electrophoresis methods, 10 x TBE buffer was prepared by adding tris base (108 g), boric acid (55 g) and EDTA acid (9.3 g) to $H_2O$ (1 L total volume) at pH 8.3. The buffer was stored at 4 °C in the dark.

### *6.3.4.1 Polyacrylamide gel electrophoresis protocol*

For denaturing polyacrylamide gel electrophoresis (PAGE) oligonucleotide analysis and purification a 20 % polyacrylamide gel was prepared. To acrylamide (35 mL of 40 % acrylamide in $H_2O$, Fisher) and 10 x TBE buffer (7 mL) was added urea (29.4 g) and $H_2O$ to a total volume of 70 mL. Immediately before loading the gel was cross-linked with *N,N,N'N'*-tetramethylethylene diamine (TEMED) (49 µL) and 10 % ammonium persulfate (APS) (490 µL). The gel was loaded into glass plates in a vertical electrophoresis tank with 10 or 15 well comb and allowed to set. The comb was then removed and the gel pre-run in TBE buffer for 1.5 h at an applied wattage of 20 W. A reference dye mixture (xylene cyanol and bromophenol blue) was loaded into the first well. Oligonucleotide samples were prepared at a concentration of 0.5 OD per sample in 30 µL (1:1 $H_2O$:formamide) and heated to 80 °C for 5 min before cooling on ice then loading onto the gel (1 sample per well). The gel was run for 3 h at a constant applied wattage of 20 W. The gel was then removed and imaged using a G:Box imaging system (Syngene) and GeneSnap (v7.08, Syngene) software. Fluorescent bands were imaged with transluminator darkroom lighting and EtBr/UV filter. UV active bands were imaged with Epi shortwave UV darkroom lighting and shortwave bandpass filter.

In cases of oligonucleotide purification, up to 7 OD of oligonucleotide was divided between 10 wells. Following separation required bands were removed from the gel with a surgical blade, crushed and suspended in $H_2O$ (2 mL) then incubated with shaking at 37 °C for 16 h. The supernatant was removed, the volume increased to 2.5 mL ($H_2O$) and the oligonucleotide desalted with a NAP-25 gel filtration column (GE Healthcare).

For native PAGE analysis 12 % polyacrylamide gel was prepared by adding acrylamide (40 % in $H_2O$, 21 mL) to 10 x TBE buffer (7 mL) and $H_2O$ (42 mL). The gel was cross-linked by the addition of TMED (40 µL) and 10 % APS (400 µL). The gel was loaded between glass plates in a vertical electrophoresis tank and left to set then pre-run in TBE buffer at an applied voltage of 120 V for 30 min. The gel was allowed to cool for 30 min before the samples were loaded (1 sample per well), with reference dye mixture (xylene cyanol and bromophenol blue) in the first well. Samples were prepared at ~0.3 OD per sample in a tris-HCl buffer with KCl, a total of 3 mM $MgCl_2$ and a total volume of 40 µL. They were heated to 75 °C in a water bath and allowed to cool to rt. Immediately

prior to loading glycerol (20 µL) was added to each sample. The gel was run with an applied voltage of 120 V for 13 h before removal and imaging as above.

In an alternative case a 5 % native gel was prepared by decreasing the volume of acrylamide and increasing the volume of $H_2O$.

### *6.3.4.2 Agarose gel electrophoresis protocol*

Agarose gel (2 %) electrophoresis analysis was performed to analyse PCR product. To TBE buffer (80 mL) was added agarose (1.6 g) and the mixture heated to boiling in a microwave. EtBr in $H_2O$ (50 µL of 1 mg/mL) was added and the gel was mixed then allowed to cool to <50 °C. The gel was added to a horizontal gel tank, with comb, and allowed to set. After removal of the comb TBE buffer was added to the tank. A 50 bp DNA step ladder (6 µL of 340 µg/µL, Promega) with added loading dye (2 µL of 6x blue/orange loading dye, Promega) was added in the first well. To each subsequent well was added the DNA sample (10 µL of total 20 µL PCR sample) with added loading dye (2 µL). The gel was run with an applied voltage of 120 V until the loading dye reached 2/3 distance along the gel. After gel removal it was imaged using a G:Box imaging system (Syngene) and GeneSnap (v7.08, Syngene) software. Fluorescent bands were imaged with transluminator darkroom lighting and EtBr/UV filter. UV active bands were imaged with Epi shortwave UV darkroom lighting and shortwave bandpass filter.

# Appendix

# Appendix

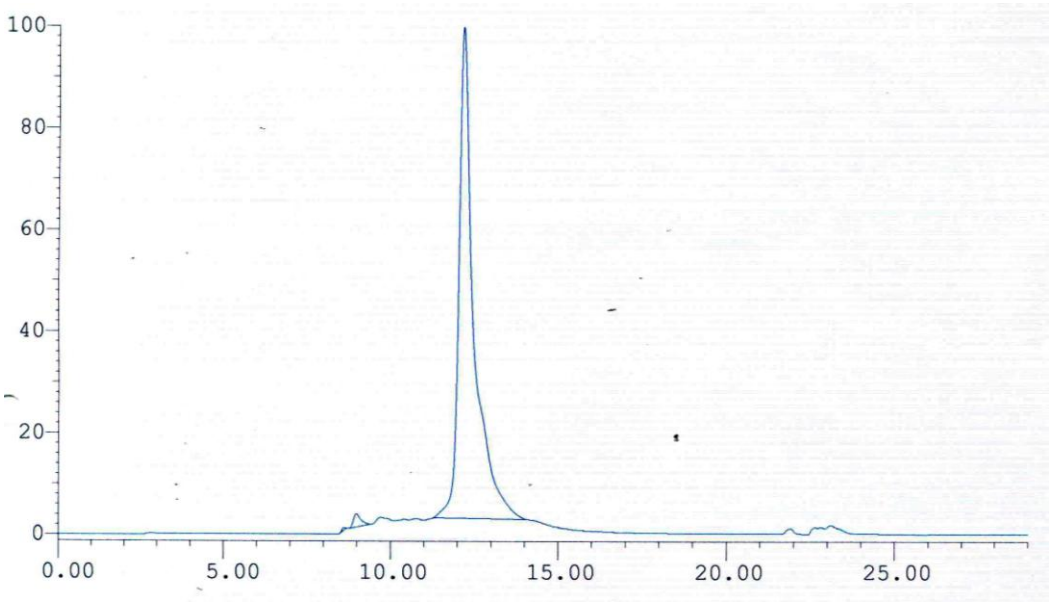
## A.1. Oligonucleotide data; sequences discussed in chapter 2.

**Table A.1.1.** Data (for modified oligonucleotides) and sequences discussed in chapter 2, in ascending numerical order. **X** = CPP monomer, **I**= inosine.

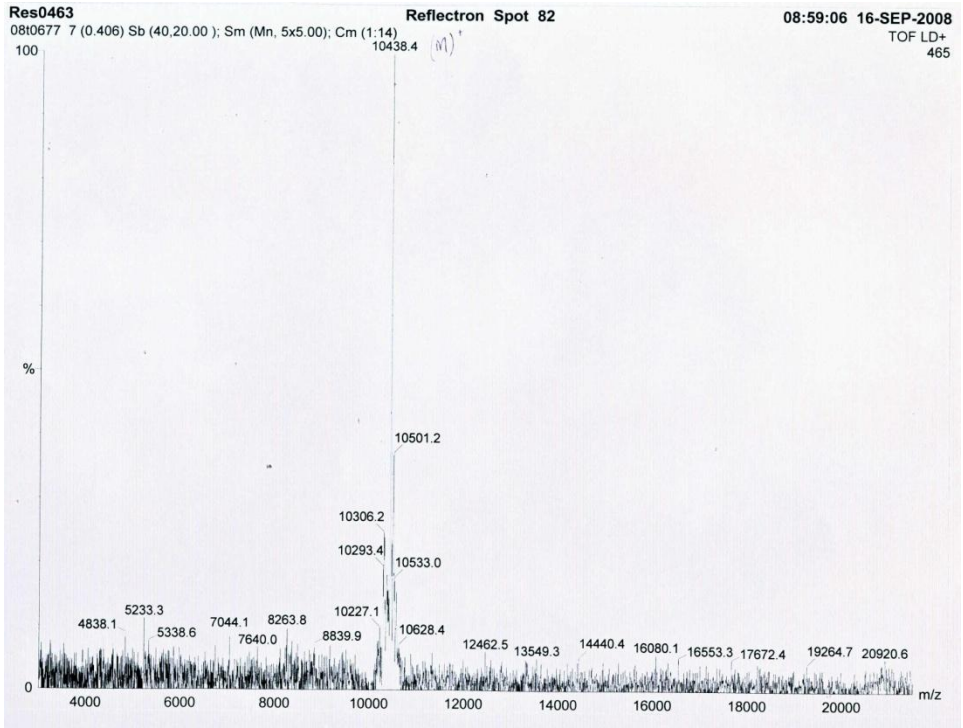
| Oligonucleotide code | Sequence | Mass expected | Mass found |
|---|---|---|---|
| ODN01 | CGA TCA CAC ACA AGG ACG AGG ATA AGG AGG AGG | NA | NA |
| ODN02 | CGA TCA CAC A**X**A AGG ACG AGG ATA AGG AGG AGG | 10437 | 10438.4 |
| ODN03 | CGA TCA CA**X** ACA AGG ACG AGG ATA AGG AGG AGG | 10437 | 10439.1 |
| ODN04 | CGA TCA **X**AC ACA AGG ACG AGG ATA AGG AGG AGG | 10437 | 10459.0 $(M+Na)^+$ |
| ODN05 | CCT CCT CCT TAT CCT CGT CCT TGT GTG TGA TCG | NA | NA |
| ODN06 | CCT CCT CCT TAT CCT CGT CCT T**I**T GTG TGA TCG T | 10253 | 10257 |
| ODN07 | CCT CCT CCT TAT CCT C**I**T CCT TGT **I**T**I** T**I**A TC**I** T | 10193 | 10195.0 |
| ODN08 | CCT CCT CCT TAT CCT C**I**T CCT T**I**T **I**T**I** T**I**A TC**I** T | 10178 | 10179.0 |
| ODN09 | CCT CCT CCT TAT CCT CGT CCT TGT **I**TG TGA TCG | 9949 | 9950.0 |
| ODN10 | CCT CCT CCT TAT CCT CGT CCT TGT GT**I** TGA TCG | 9949 | 9950.0 |
| ODN11 | CCT CCT CCT TAT CCT C**I**T CCT T**I**T GT**I** T**I**A TC**I** T | 10193 | 10190.0 |
| ODN12 | CCT CCT CCT TAT CCT C**I**T CCT T**I**T **I**TG T**I**A TC**I** T | 10193 | 10190.0 |

## Example oligonucleotide analysis

### CPP monomer (ODN02)




***Crude HPLC spectrum***




***MALDI mass spectrum***

## A.2. Oligonucleotide data; sequences discussed in Chapter 3

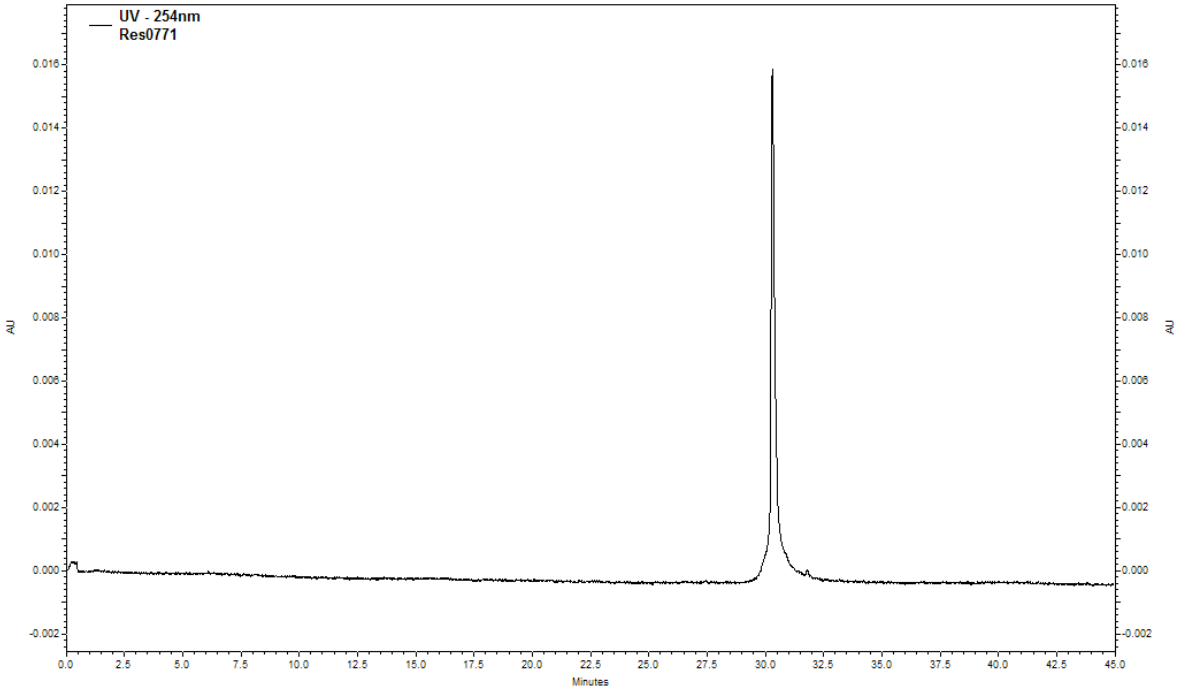
**Table A.2.1** Data (for modified oligonucleotides) and sequences discussed in chapter 3, in ascending numerical order. **3**=Cy3dT, **5**= Cy5dT, **P**= propanol.

| Oligonucleotide code | Sequence | Mass expected | Mass found |
|---|---|---|---|
| ODN13 | CAC CAA AGA TGA **3**AT TT**3** CTT TAA TGG **P** | 9141 | 9150.0 |
| ODN14 | CAC CAA AGA TGA TAT TT**3** CTT TAA TGG **P** | 8775 | 8780.0 |
| ODN15 | CCA TTA AAG AAA ATA TCA TCT TTG GTG | NA | NA |
| ODN16 | CAC CAA AGA TGA TATTT**5** CTT TAA TGG **P** | 8845 | 8849.0 |
| ODN17 | GGA **5**TT TCG **3**TT TTA TAA TTG CC | 7823 | 7824.3 |
| ODN18 | GGC AAT TAT AAA AAC GAA AAT CC | NA | NA |
| ODN20 | CGT ATA TTC TTT ATT T**3**T AAA AGC C | 7960 | 7964.0 |
| ODN21 | CGT ATA TTC TTT ATT **3**TT AAA AGC C | 7960 | 7962.1 |
| ODN22 | CGT ATA TTC TTT AT**3** TTT AAA AGC C | 7960 | 7960.3 |
| ODN23 | CGT ATA TTC TTT A**3**T TTT AAA AGC C | 7960 | 7962.4 |
| ODN24 | CGT ATA TTC TT**3** ATT TTT AAA AGC C | 7960 | 7961.4 |
| ODN25 | CGT ATA TTC T**3**T ATT TTT AAA AGC C | 7960 | 7960.2 |
| ODN26 | CGT ATA TTC **3**TT ATT TTT AAA AGC C | 7960 | 7960.5 |
| ODN27 | CGT ATA T**3**C TTT ATT TTT AAA AGC C | 7960 | 7960.9 |
| ODN28 | CGT ATA **3**TC TTT ATT TTT AAA AGC C | 7960 | 7961.4 |
| ODN29 | CGT A**3**A TTC TTT ATT TTT AAA AGC C | 7960 | 7961.1 |
| ODN30 | CG**3** ATA TTC TTT ATT TTT AAA AGC C | 7960 | 7960.5 |
| ODN31 | GGC TTT **5**AA AAA TAA AGA ATA TAC G | 8153 | 8156.5 |
| ODN32 | CGT ATA TTC TTT ATT TTT AAA AGC C | NA | NA |
| ODN33 | GGC TTT TAA AAA TAA AGA ATA TAC G | NA | NA |
| ODN34 | ATG GCT ATT AAA AAC GGA GAA TGT ACG CAA | NA | NA |
| ODN35 | TTG CGT ACA TTC TCC GTT TTT AAT AGC CAT | NA | NA |
| ODN36 | TTG CGT ACA **3**TC TCC GTT T**3**T AAT AGC CAT | 9840 | 9842.0 |
| ODN37 | ATG GCT ATT A**3**A **3**A**3** A**3**A **3**AA TGT ACG CAA | 11046 | 11048.7 |
| ODN38 | TTG CGT ACA T**3**A **3**A**3** A**3**A **3**AT AAT AGC CAT | 10997 | 10999.3 |
| ODN39 | TTG CGT ACA TTA TAT ATA TAT AAT AGC CAT | NA | NA |
| ODN40 | ATG GCT ATT A**33 333 3**AT GTA CGC AA | 10148 | 10155.1 |
| ODN41 | TTG CGT ACA TAA AAA ATA ATA GCC AT | NA | NA |
| ODN42 | ATG GCT ATT ATA TAT ATA TAA TGT ACG CAA | NA | NA |

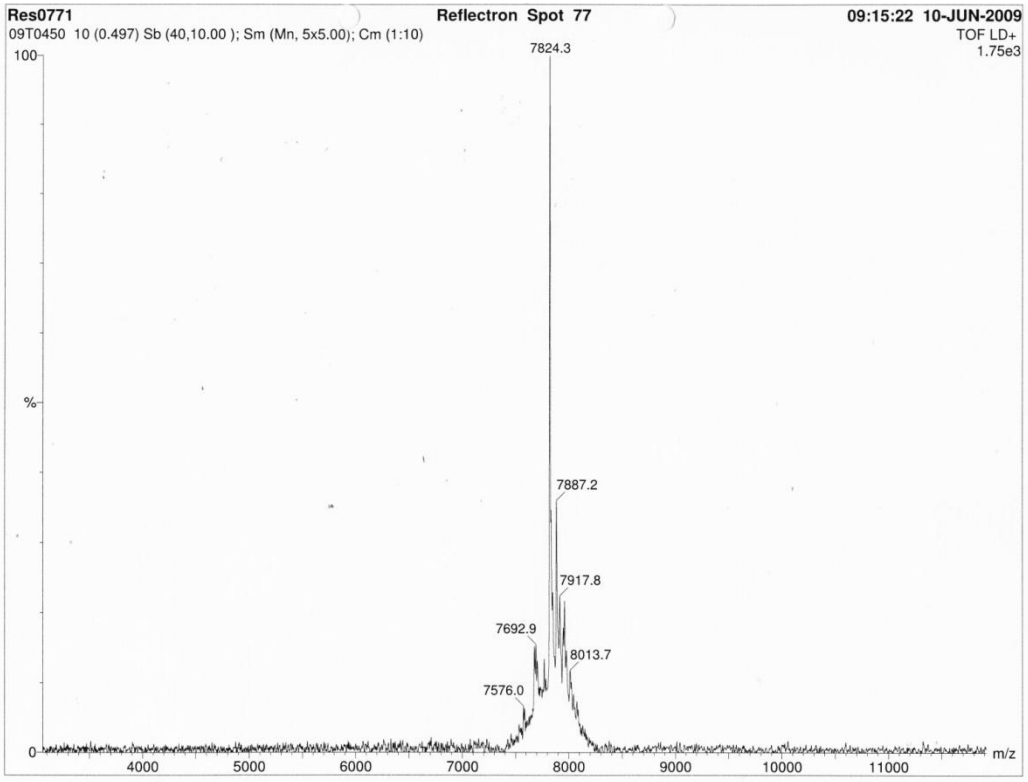
| | | | |
|---|---|---|---|
| ODN43 | ATG GCT ATT ATT TTT TAT GTA CGC AA | NA | NA |
| ODN44 | GGC T**5**T TAA AAA TAA AGA ATA TAC G | 8153 | 8152.3 |

## Example oligonucleotide analysis

### Cy3dT monomer and Cy5dT monomer (ODN17)

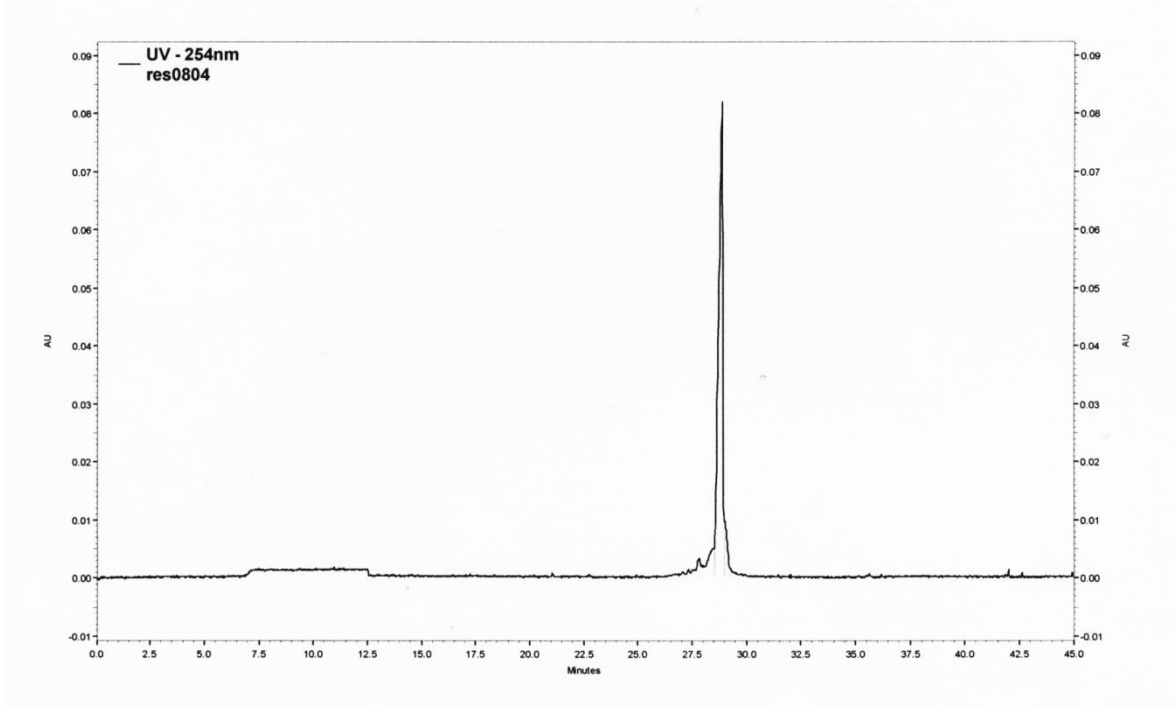



***Capillary electrophoresis***

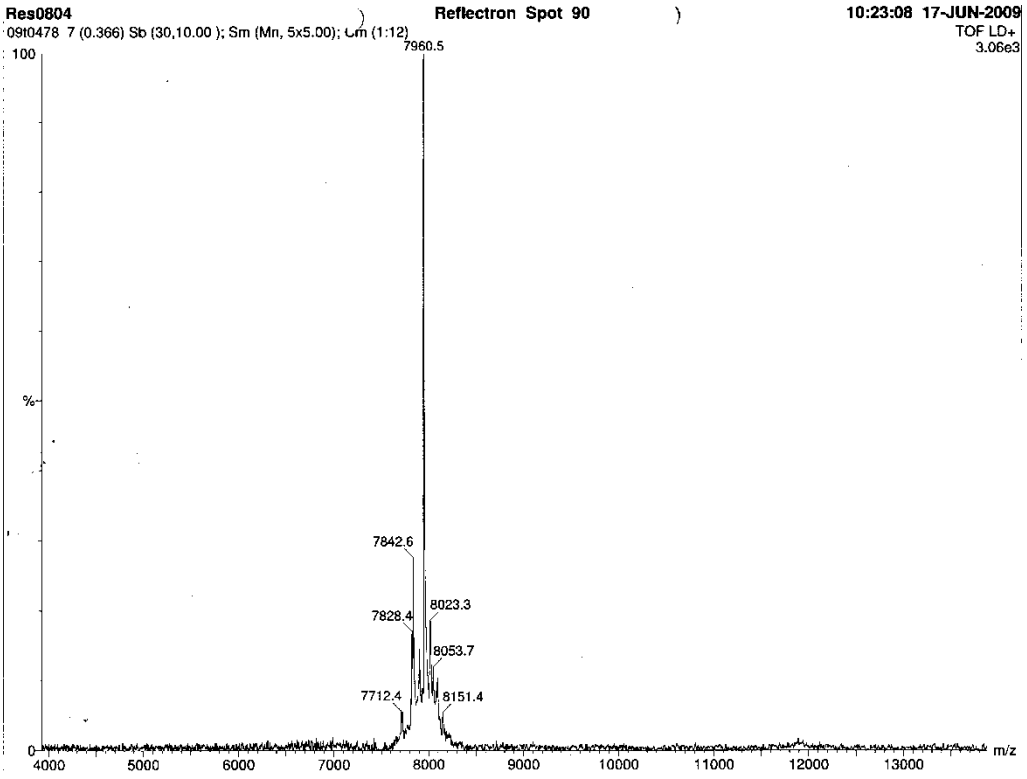



***MALDI mass spectrum***

## Cy3dT monomer (ODN26)

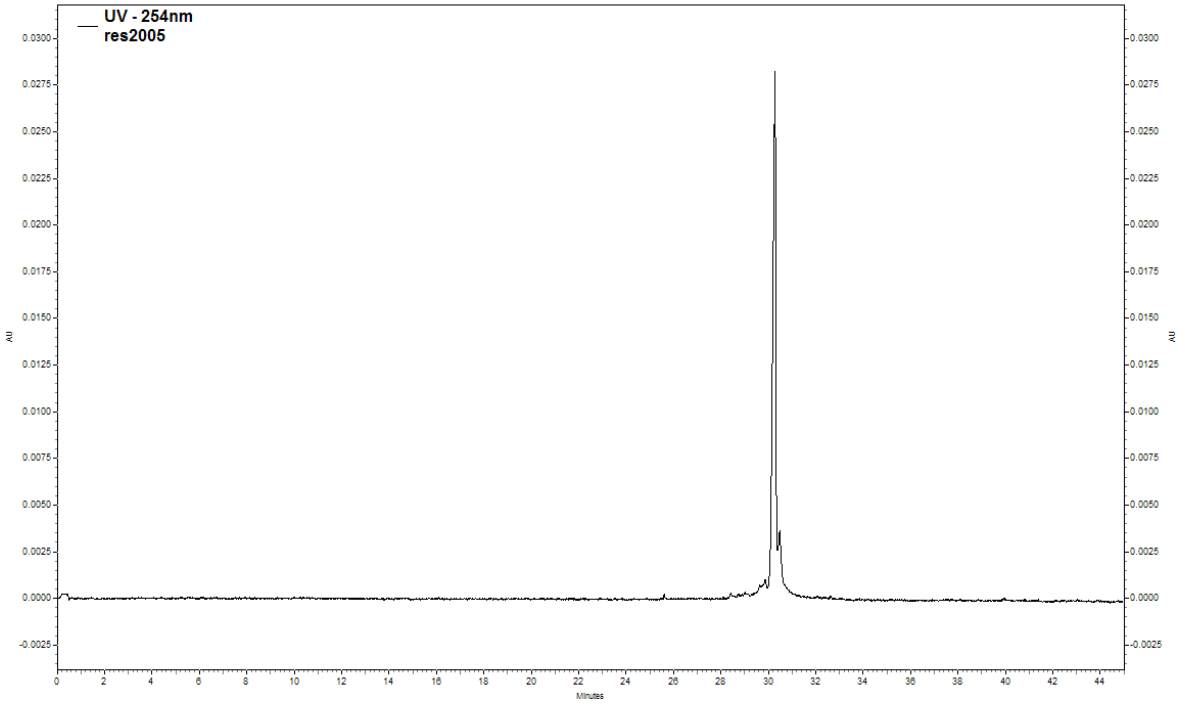



### *Capillary electrophoresis*

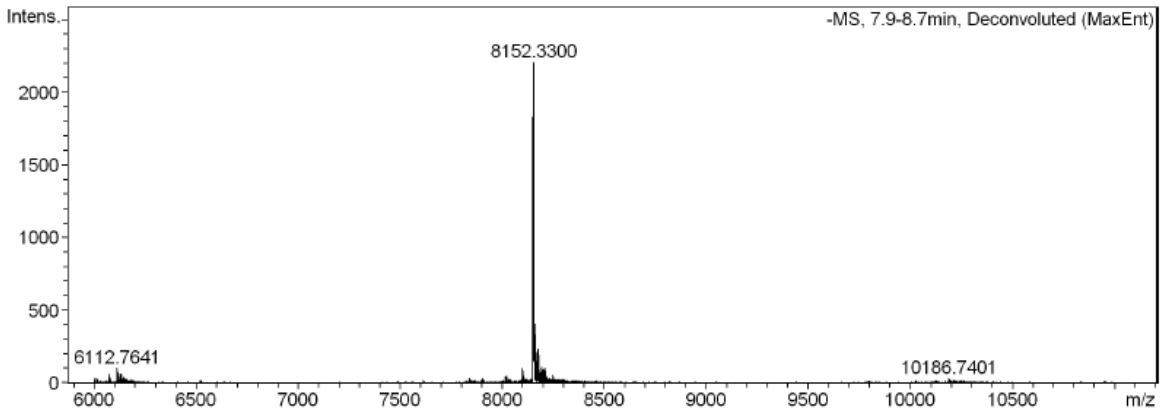



### *MALDI mass spectrum*

## Cy5dT monomer (ODN44)




***Capillary electrophoresis***




***Electrospray mass spectrum***

## A.3. Oligonucleotide data; for sequences discussed in chapter 4.

**Table A.3.1.** Data (for modified oligonucleotides) and sequences discussed in chapter 4, in ascending numerical order. **1**= Cy3BdT; **2**= Cy3dT; **3**=Cy5 phosphoramidite; **4**=FAM C7; **5**= 2'mesyloxyethyl rT converted to azide, labelled with Cy3B; **6**= propanol; **7**= aminoC6dT labelled with azidohexanoic acid and Cy3B; **8**= 5-(hexyn-1-ol)-Cy3B phosphoramidite; **9**= BHQ2; **Z**= BHQ1; **Y**= BHQ2dT; **X**= BHQ1dT; **W**= HEG; **R**= Cy3BdR resin; **S**= Cy3BdR phosphoramidite; **D**= Dabcyl; **Q**= Dabcyl-dT.

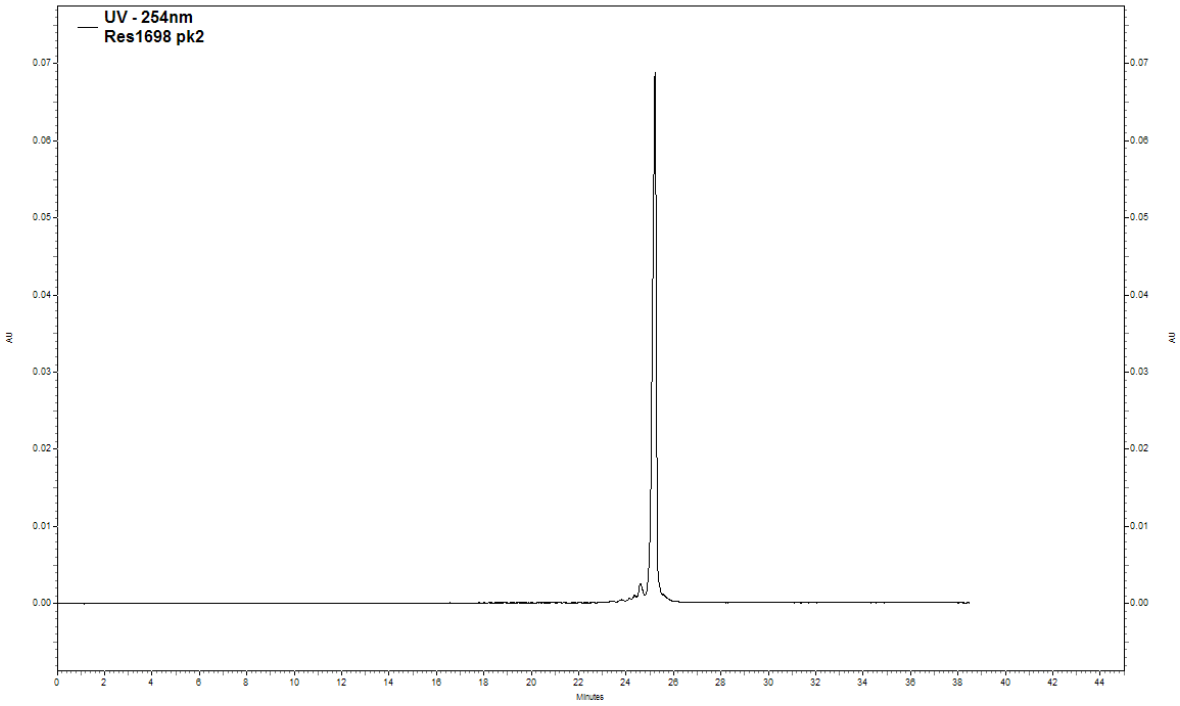
| **Oligonucleotide code** | **Sequence** | **Mass expected** | **Mass found** |
|---|---|---|---|
| ODN45 | TCT CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GTA AGA AAC TAT GTG AAA ACT TTT TGA | NA | NA |
| ODN46 | TCT CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT GGA TAC AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GTA AGA AAC TAT GTG AAA ACT TTT TGA | NA | NA |
| ODN34 | ATG GCT ATT AAA AAC GGA GAA TGT ACG CAA | NA | NA |
| ODN36 | TTG CGT ACA **2**TC TCC GTT T**2**T AAT AGC CAT | 9840 | 9842.0 |
| ODN47 | **3**AT GGC TAT TAA AAA CGG AGA ATG TAC GCA A | 9829 | 9830.8 |
| ODN48 | **3**AT GGC TAT TAA AAA CGG AGA ATG TAC GCA A**4** | 10395 | 10399.7 |
| ODN49 | CTA TGA TGA ATA TAG ATA CAG AAG CGT CAT | NA | NA |
| ODN50 | CTA TGA TGA ATA TGG ATA CAG AAG CGT CAT | NA | NA |
| ODN51 | TTG CGT ACA **1**TC TCC GTT T**1**T AAT AGC CAT | 9974 | 9973.9 |
| ODN52 | TTG CGT ACA TTC TCC GTT T**1**T AAT AGC CAT | 9544 | 9545.4 |
| ODN53 | TTG CGT ACA **1**TC TCC GTT TTT AAT AGC CAT | 9544 | 9545.4 |
| ODN54 | CGC TTC **5**GT A**5**C TAT ATT CAT C**6** | 7829 | 7828.1 |
| ODN55 | CGC TTC **5**GT ATC **5**AT ATT CAT C**6** | 7829 | 7828.1 |
| ODN56 | CGC TTC **5**GT ATC TA**5** ATT CAT C**6** | 7829 | 7829.6 |
| ODN57 | CGC TTC **7**GT A**7**C TAT ATT CAT C**6** | 8244 | 8244.0 |
| ODN58 | CGC TTC **7**GT ATC **7**AT ATT CAT C**6** | 8244 | 8243.9 |
| ODN59 | CGC TTC **7**GT ATC TA**7** ATT CAT C**6** | 8244 | 8243.9 |
| ODN60 | CAG TTT TCC TGG ATT ATG CC | NA | NA |
| ODN61 | CAA AAA GTT TTC ACA TAG TTT CTT | NA | NA |

| | | | |
|---|---|---|---|
| ODN62 | CGC TTC **1**GT A**1**C TAT ATT CAT C**6** | 7625 | 7624.3 |
| ODN63 | CGC TTC **1**GT ATC **1**AT ATT CAT C**6** | 7625 | 7624.3 |
| ODN64 | CGC TTC **1**GT ATC TA**1** ATT CAT C**6** | 7625 | 7624.3 |
| ODN65 | **8**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **9** | 12897 | 12896.9 |
| ODN66 | GCA TAA TCC AGG AAA ACT GA | NA | NA |
| ODN67 | **8**TC AGT TTT CCT GGA TTA TGC | 6671 | 6669.3 |
| ODN68 | TGA TGA CGC TTC TGT ATC TAT ATT CAT CAT AGG A | NA | NA |
| ODN69 | TGA TGA CGC TTC TGT ATC CAT ATT CAT CAT AGG A | NA | NA |
| ODN70 | CGC TTC TGT ATC TAT ATT CAT C | NA | NA |
| ODN71 | **8**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **Z** | 12895 | 12893.9 |
| ODN72 | **8**TT CCT ATG A**Y**G AAT ATA GAT ACA GAA GCG **6** | 10326 | 10325.8 |
| ODN73 | **8**CC GCG GGA TGA ATA TAG ATA CAG AAG CGC CGC GG**X W**TC TTC TAG TTG GCA TGC T | 17971 | 17969.8 |
| ODN74 | GGC ACC ATT AAA GAA AAT ATC A | NA | NA |
| ODN75 | AAT ATC ATC TTT GGT GTT TCC T**R** | 7378 | 7377.3 |
| ODN76 | **S**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **D** | 12918 | 12917.7 |
| ODN77 | **S**CC GCG GGA TGA ATA TAG ATA CAG AAG CGC CGC GG**Q W**TC TTC TAG TTG GCA TGC T | 17835 | 17834.4 |
| ODN78 | **S**CC TAG CAT GAT GAA TAT AGA TAC AGA AGC GTC GCT AGG **R** | 13153 | 13151.7 |
| ODN79 | TTT T**S**T TT | 2764 | 2763.6 |

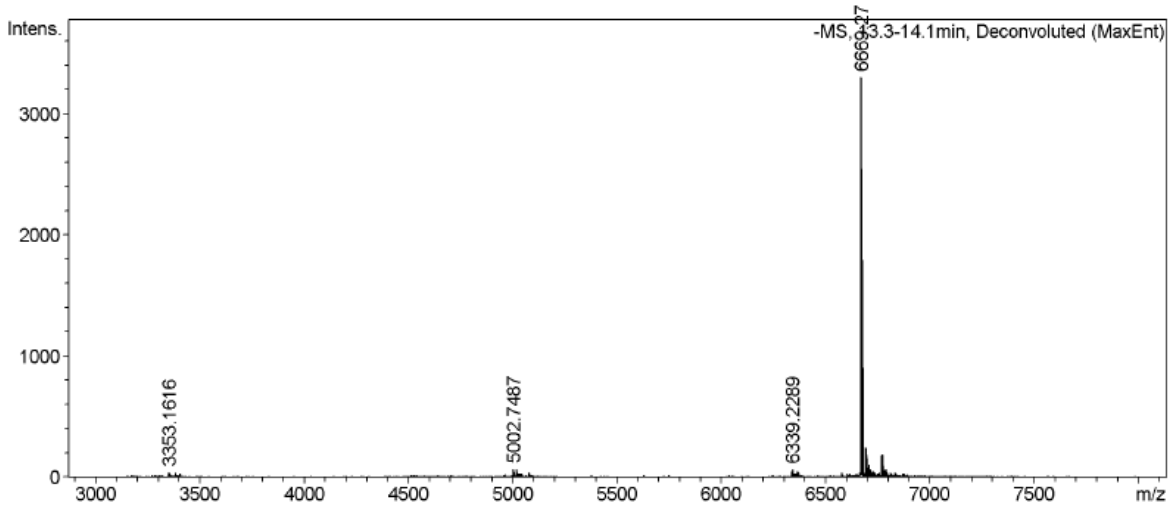
ODN80 and ODN81 were previously prepared and analysed by Dr. James Richardson.

## Example oligonucleotide analysis

### 5’ Cy3B monomer (ODN67)

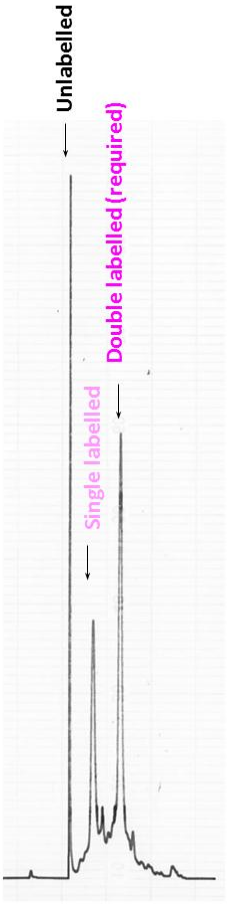


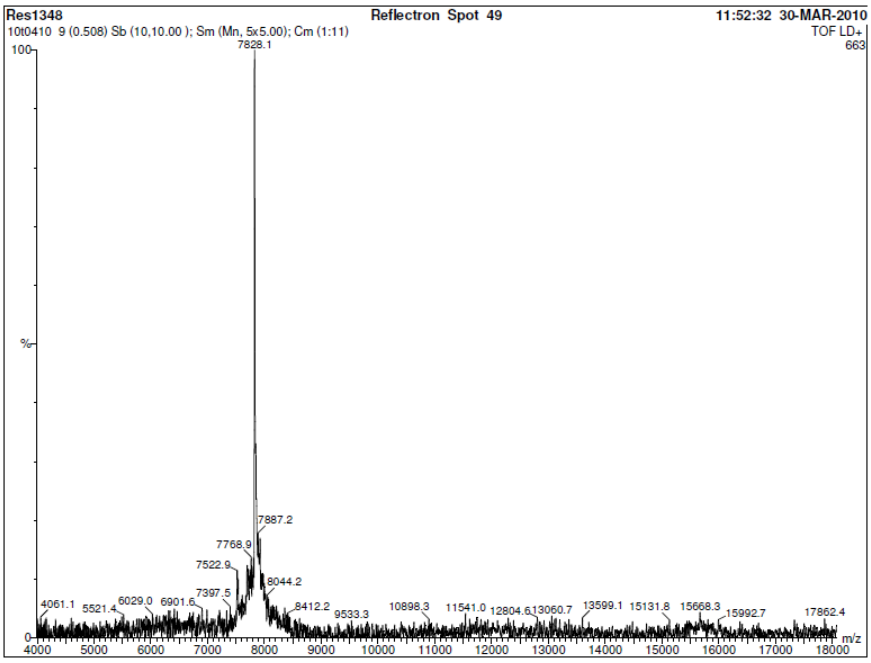

***Capillary electrophoresis***



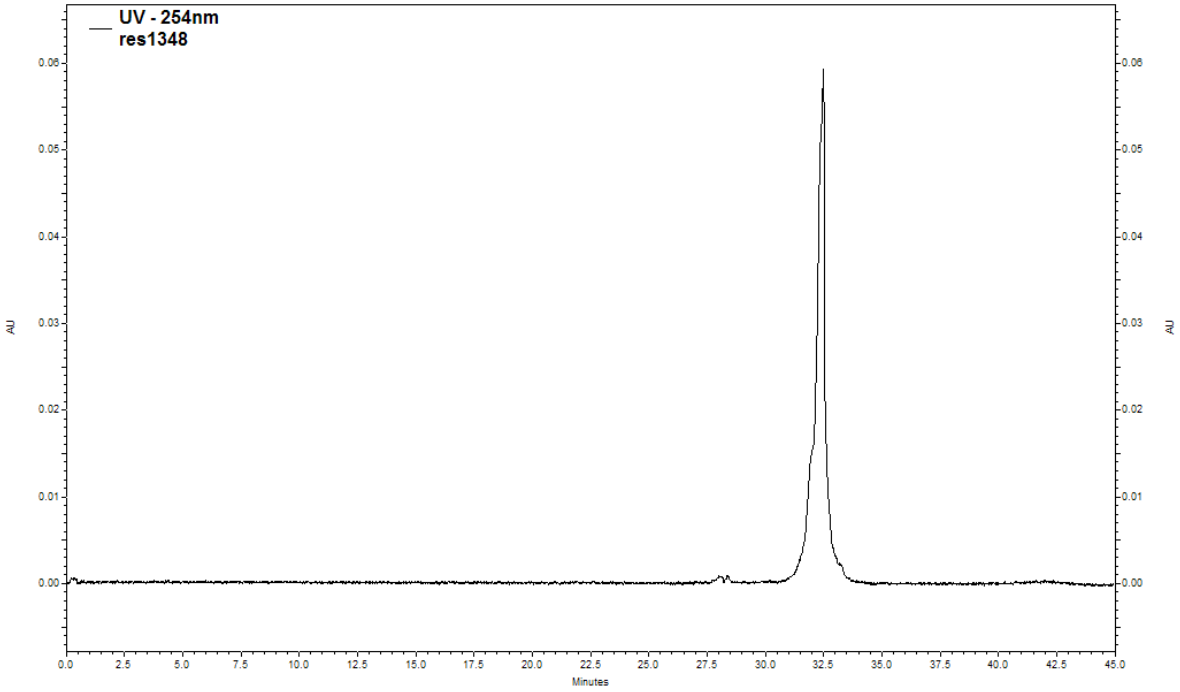

***Electrospray mass spectrum***

## Click chemistry labelling of minor groove with ethynyl-Cy3B (ODN55)



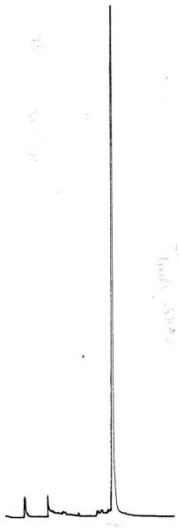




***Crude HPLC spectrum*** ***MALDI mass spectrum***

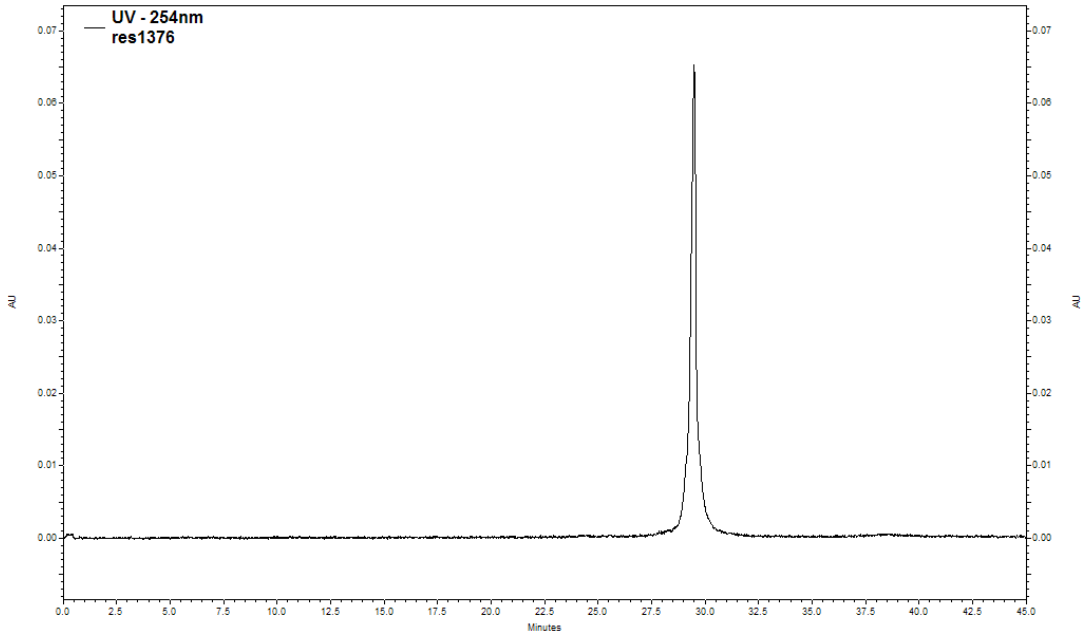



***Capillary electrophoresis***

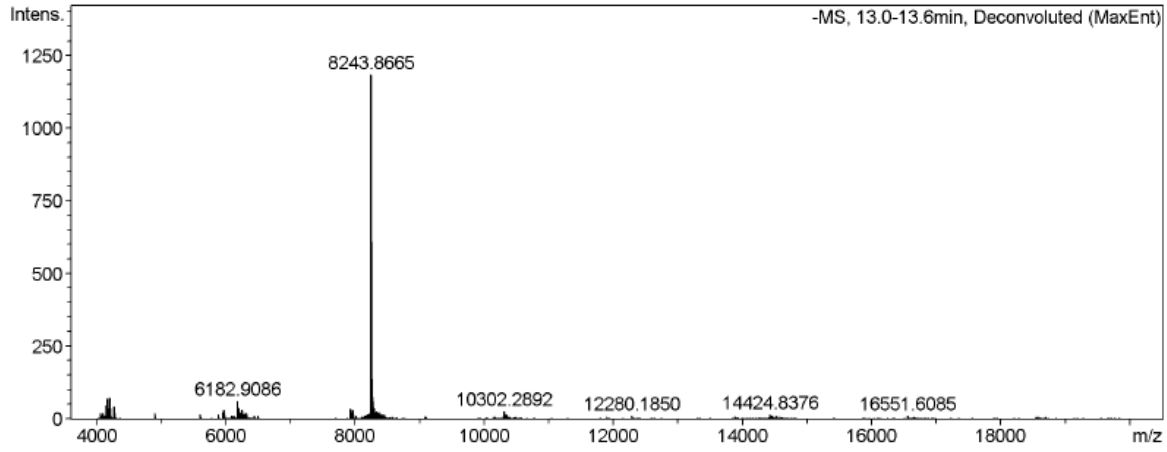
## Click chemistry labelling of major groove with ethynyl-Cy3B (ODN58)



***Crude HPLC spectrum***

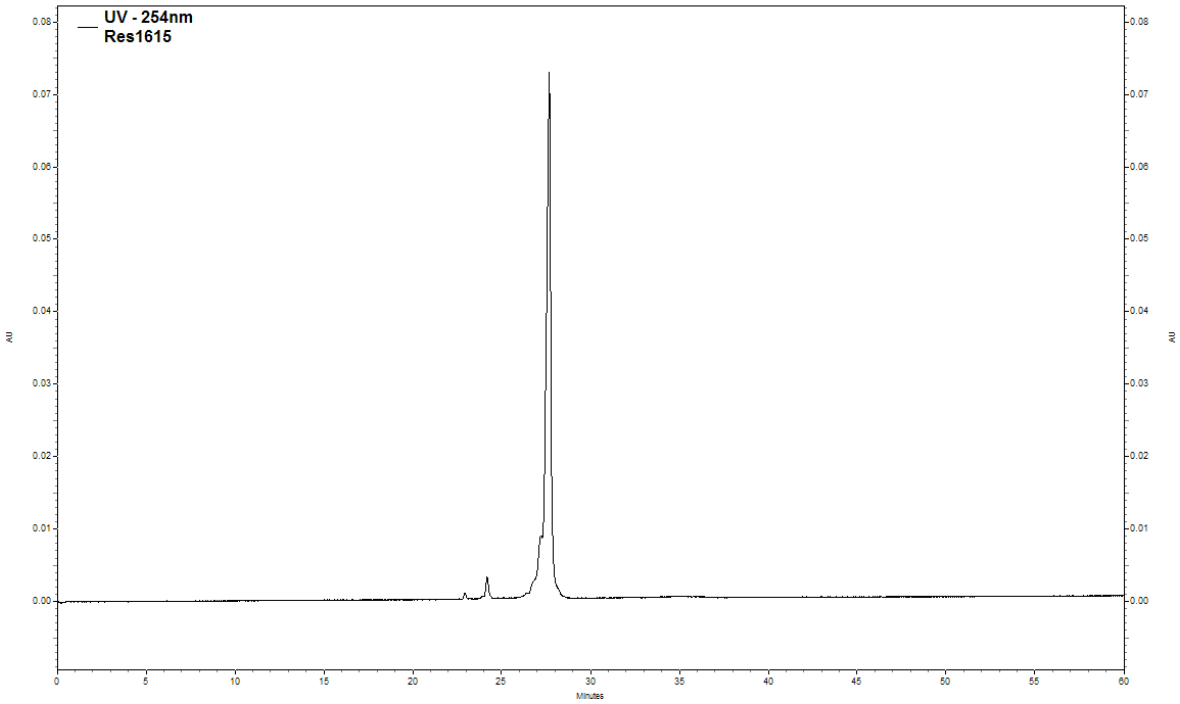



***Capillary electrophoresis***

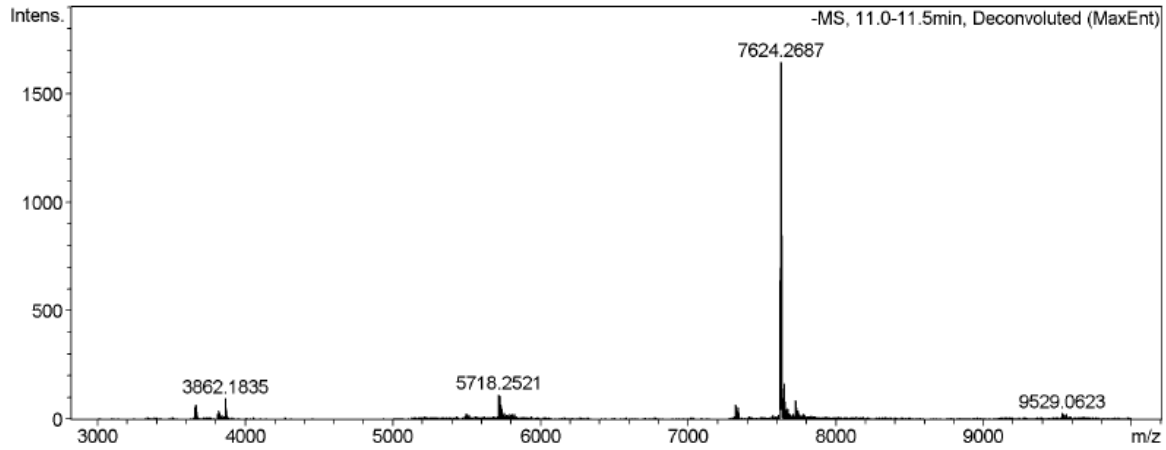



***Electrospray mass spectrum***

## Cy3BdT monomer (ODN63)

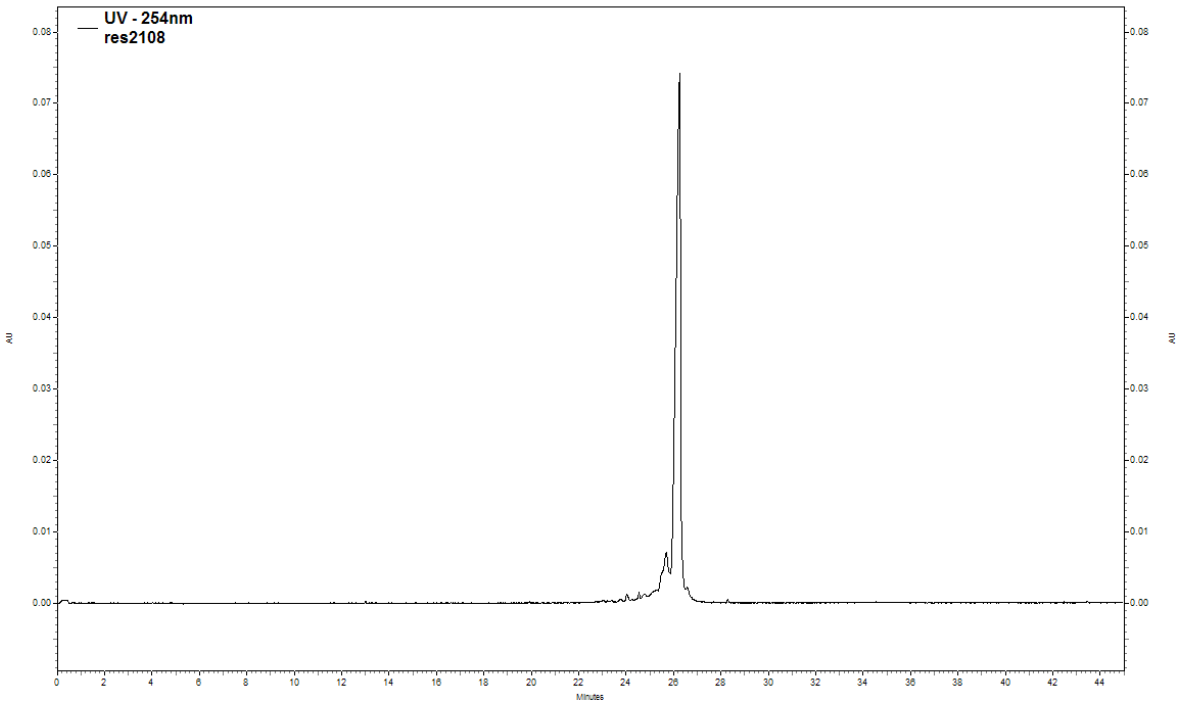



### *Capillary electrophoresis*

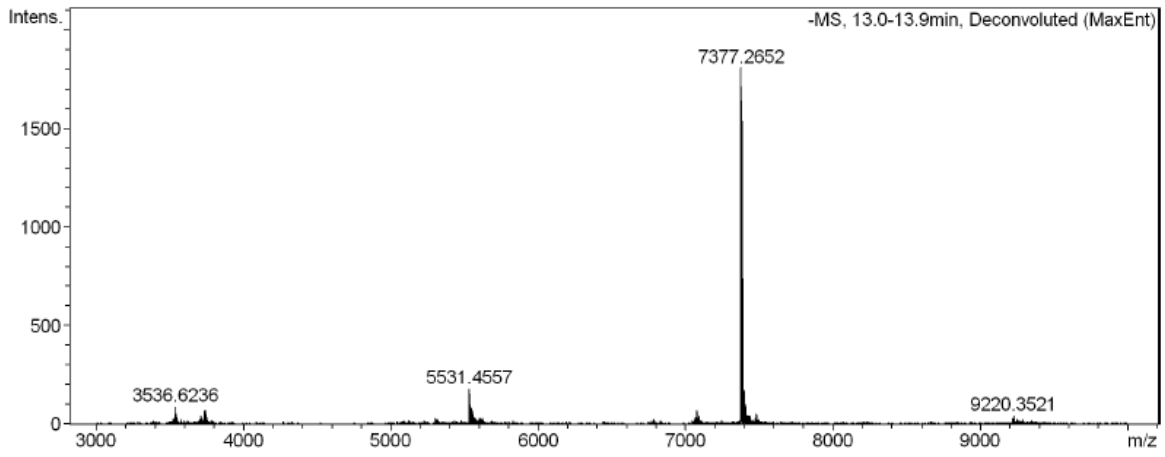



### *Electrospray mass spectrum*

## 3'Cy3BdR monomer (ODN75)

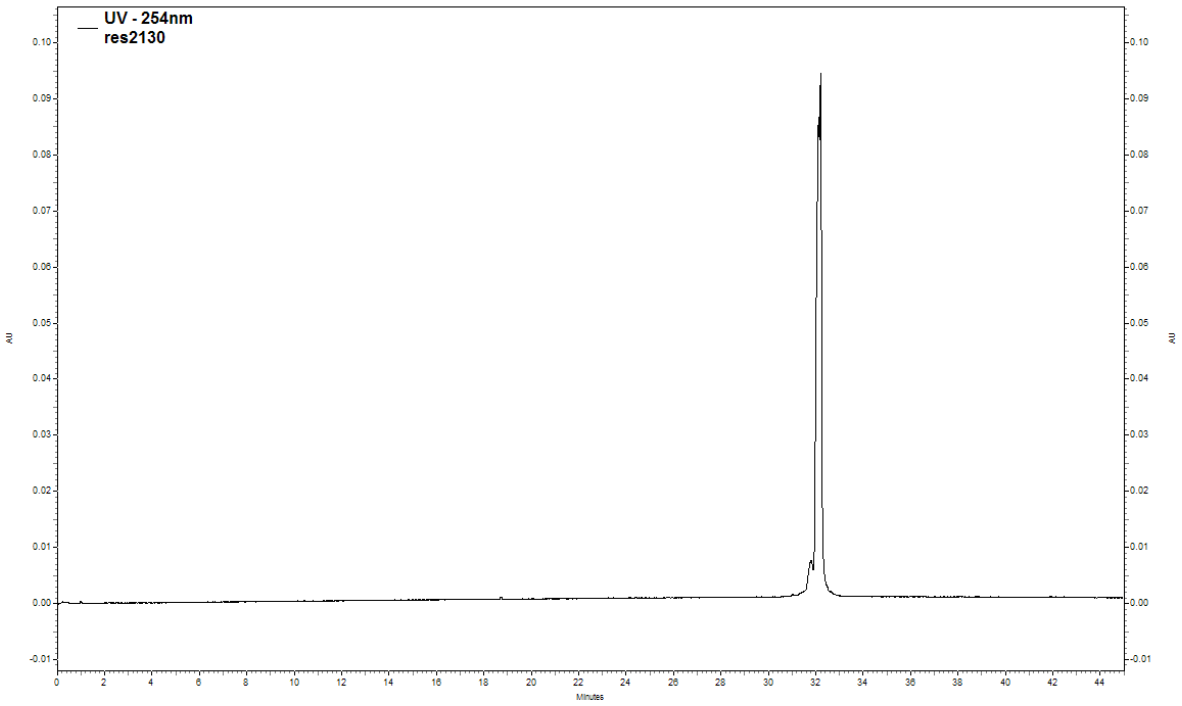



***Capillary electrophoresis***

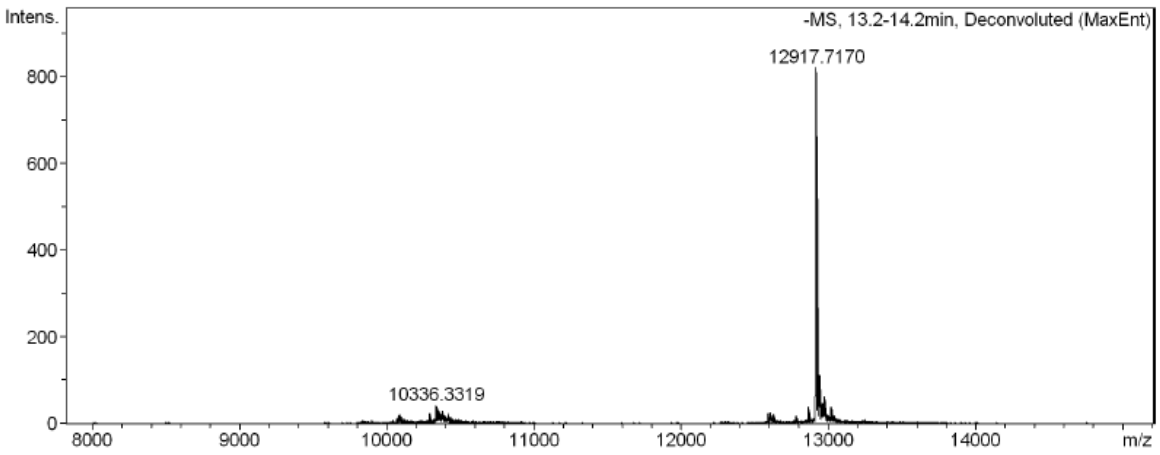



***Electrospray mass spectrum***

## Cy3BdR monomer (ODN76)




### *Capillary electrophoresis*




### *Electrospray mass spectrum*

# References

# References


1. Franklin, R.; Gosling, R. G. *Nature* **1953,** *171*, 740-741.
2. Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953,** *171*, 738-740.
3. Watson, J. D.; Crick, F. H. C. *Nature* **1953,** *171*, 737-738.
4. Merrifield, R. B. *J. Am. Chem. Soc.* **1963,** *85*, 2149-2154.
5. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981,** *22*, 1859-1862.
6. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993,** *49*, 6123-6194.
7. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1993,** *49*, 1925-1963.
8. McBride, L. J.; Caruthers, M. H. *Tetrahedron Lett.* **1983,** *24*, 245-248.
9. Gore, M. G. *Spectrophotometry and Spectrofluorimetry - A practical approach*; Oxford University Press, 2000.
10. Johnson, M. L.; Holt, J. M.; Ackers, G. K. In *Methods in Enzymology; Biothermodynamics, Part A*; Academic Press, 2009; pp. 379-382.
11. Aboul-ela, F.; Koh, D.; Tinoco, I.; Martin, F. H. *Nucleic Acids Res.* **1985,** *13*, 4811-4824.
12. Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U. S. A.* **1986,** *83*, 3746-3750.
13. Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987,** *26*, 1601-1620.
14. Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90*, 2994-2998.
15. Lilley, D. M. J. In *Methods in Enzymology, Vol 469: Biophysical, Chemical, and Functional Probes of Rna Structure, Interactions and Folding, Pt B;* Herschlag, D. Ed.; Elsevier Academic Press Inc: San Diego, 2009; pp. 159-187.
16. Förster, T. *Naturwissenschaften* **1946,** *33*, 166-175.
17. Stengel, G.; Gill, J. P.; Sandin, P.; Wilhelmsson, L. M.; Albinsson, B.; Nordén, B.; Millar, D. *Biochemistry* **2007,** *46*, 12289-12297.
18. Selvin, P. R. *Nat. Struct. Biol.* **2000,** *7*, 730-734.
19. Lilley, D. M. J.; Wilson, T. J. *Curr. Opin. Chem. Biol.* **2000,** *4*, 507-517.
20. Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93*, 6264-6268.
21. Roy, R.; Hohng, S.; Ha, T. *Nat. Methods* **2008,** *5*, 507-516.
22. Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. *Biophys. J.* **2004,** *86*, 2530-2537.

23. Uphoff, S.; Holden, S. J.; Le Reste, L.; Periz, J.; van de Linde, S.; Heilemann, M.; Kapanidis, A. N. *Nat. Methods* **2010,** *7*, 831-836.
24. Ha, T. *Methods* **2001,** *25*, 78-86.
25. Wilhelmsson, L. M. *Q. Rev. Biophys.* **2010,** *43*, 159-183.
26. Gao, J. M.; Strassler, C.; Tahmassebi, D.; Kool, E. T. *J. Am. Chem. Soc.* **2002,** *124*, 11590-11591.
27. Ren, R. X. F.; Chaudhuri, N. C.; Paris, P. L.; Rumney, S.; Kool, E. T. *J. Am. Chem. Soc.* **1996,** *118*, 7671-7678.
28. Strassler, C.; Davis, N. E.; Kool, E. T. *Helv. Chim. Acta* **1999,** *82*, 2160-2171.
29. Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006,** *4*, 4265-4274.
30. Ward, D. C.; Reich, E.; Stryer, L. *J. Biol. Chem.* **1969,** *244*, 1228-1237.
31. Hawkins, M. E. *Cell Biochem. Biophys.* **2001,** *34*, 257-281.
32. Hawkins, M. E.; Pfleiderer, W.; Balis, F. M.; Porter, D.; Knutson, J. R. *Anal. Biochem.* **1997,** *244*, 86-95.
33. Woo, J. S.; Meyer, R. B.; Gamper, H. B. *Nucleic Acids Res.* **1996,** *24*, 2470-2475.
34. Berry, D. A.; Jung, K. Y.; Wise, D. S.; Sercel, A. D.; Pearson, W. H.; Mackie, H.; Randolph, J. B.; Somers, R. L. *Tetrahedron Lett.* **2004,** *45*, 2457-2461.
35. Inoue, H.; Imura, A.; Ohtsuka, E. *Nippon Kagaku Kaishi* **1987**, 1214-1220.
36. Bergstrom, D. E.; Inoue, H.; Reddy, P. A. *J. Org. Chem.* **1982,** *47*, 2174-2178.
37. Hudson, R. H. E.; Dambenieks, A. K. *Heterocycles* **2006,** *68*, 1325-1328.
38. Hudson, R. H. E.; Dambenieks, A. K.; Moszynski, J. M. *Proc. of SPIE* **2005,** *5969*, 59590J:1-10.
39. Hudson, R. H. E.; Dambenieks, A. K.; Viirre, R. D. *Nucleos. Nucleot. Nucl.* **2005,** *24*, 581-584.
40. Hudson, R. H. E.; Ghorbani-Choghamarani, A. *Nucleos. Nucleot. Nucl.* **2007,** *26*, 533-537.
41. Hudson, R. H. E.; Ghorbani-Choghamarani, A. *Synlett* **2007**, 870-873.
42. Hudson, R. H. E.; Moszynski, J. M. *Synlett* **2006**, 2997-3000.
43. Wojciechowski, F.; Hudson, R. H. E. *J. Am. Chem. Soc.* **2008,** *130*, 12574-12575.
44. Lin, K. Y.; Jones, R. J.; Matteucci, M. *J. Am. Chem. Soc.* **1995,** *117*, 3873-3874.
45. Wilhelmsson, L. M.; Holmen, A.; Lincoln, P.; Nielson, P. E.; Nordén, B. *J. Am. Chem. Soc.* **2001,** *123*, 2434-2435.
46. Sandin, P.; Wilhelmsson, L. M.; Lincoln, P.; Powers, V. E. C.; Brown, T.; Albinsson, B. *Nucleic Acids Res.* **2005,** *33*, 5019-5025.

47. Sandin, P.; Lincoln, P.; Brown, T.; Wilhelmsson, L. M. *Nat. Protoc.* **2007,** *2*, 615-623.
48. Engman, K. C.; Sandin, P.; Osborne, S.; Brown, T.; Billeter, M.; Lincoln, P.; Nordén, B.; Albinsson, B.; Wilhelmsson, L. M. *Nucleic Acids Res.* **2004,** *32*, 5087-5095.
49. Sandin, P.; Borjesson, K.; Li, H.; Martensson, J.; Brown, T.; Wilhelmsson, L. M.; Albinsson, B. *Nucleic Acids Res.* **2008,** *36*, 157-167.
50. Ranasinghe, R. T.; Brown, T. *Chem. Commun.* **2011,** *47*, 3717-3735.
51. Fernandez-Suarez, M.; Ting, A. Y. *Nat. Rev. Mol. Cell Biol.* **2008,** *9*, 929-943.
52. Tsien, R. Y.; Ernst, L. A.; Waggoner, A. S. In *Handbook of Biological Confocal Microscopy;* Pawley, J. B. Ed.; SpringerScience+Business Media: New York, 2006; pp. 338-352.
53. Whitaker, J. E.; Haugland, R. P.; Ryan, D.; Hewitt, P. C.; Prendergast, F. G. *Anal. Biochem.* **1992,** *207*, 267-279.
54. Kubin, R. F.; Fletcher, A. N. *J. Lumines.* **1982,** *27*, 455-462.
55. Marras, S. A. E.; Kramer, F. R.; Tyagi, S. *Nucleic Acids Res.* **2002,** *30*, e122;1-8.
56. Ernst, L. A.; Gupta, R. K.; Mujumdar, R. B.; Waggoner, A. S. *Cytometry* **1989,** *10*, 3-10.
57. Southwick, P. L.; Ernst, L. A.; Tauriello, E. W.; Parker, S. R.; Mujumdar, R. B.; Mujumdar, S. R.; Clever, H. A.; Waggoner, A. S. *Cytometry* **1990,** *11*, 418-430.
58. Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S. *Bioconjugate Chem.* **1993,** *4*, 105-111.
59. Mujumdar, S. R.; Mujumdar, R. B.; Grant, C. M.; Waggoner, A. S. *Bioconjugate Chem.* **1996,** *7*, 356-362.
60. Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105*, 11176-11181.
61. Sanborn, M. E.; Connolly, B. K.; Gurunathan, K.; Levitus, M. *J. Phys. Chem. B* **2007,** *111*, 11064-11074.
62. Harvey, B. J.; Levitus, M. *J. Fluoresc.* **2009,** *19*, 443-448.
63. Harvey, B. J.; Perez, C.; Levitus, M. *Photochem. Photobiol. Sci.* **2009,** *8*, 1105-1110.
64. Cooper, M.; Ebner, A.; Briggs, M.; Burrows, M.; Gardner, N.; Richardson, R.; West, R. *J. Fluoresc.* **2004,** *14*, 145-150.
65. Wittwer, C. T.; Herrmann, M. G.; Moss, A. A.; Rasmussen, R. P. *Biotechniques* **1997,** *22*, 130-138.

66. Skeidsvoll, J.; Ueland, P. M. *Anal. Biochem.* **1995,** *231*, 359-365.
67. Livak, K. J.; Flood, S. J. A.; Marmaro, J.; Giusti, W.; Deetz, K. *PCR-Methods Appl.* **1995,** *4*, 357-362.
68. Holland, P. M.; Abramson, R. D.; Watson, R.; Gelfand, D. H. *Proc. Natl. Acad. Sci. U. S. A.* **1991,** *88*, 7276-7280.
69. Piatek, A. S.; Tyagi, S.; Pol, A. C.; Telenti, A.; Miller, L. P.; Kramer, F. R.; Alland, D. *Nat. Biotechnol.* **1998,** *16*, 359-363.
70. Tyagi, S.; Bratu, D. P.; Kramer, F. R. *Nat. Biotechnol.* **1998,** *16*, 49-53.
71. Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996,** *14*, 303-308.
72. Marras, S. A. E.; Kramer, F. R.; Tyagi, S. *Genet. Anal. Biomol. Eng.* **1999,** *14*, 151-156.
73. Vet, J. A. M.; Majithia, A. R.; Marras, S. A. E.; Tyagi, S.; Dube, S.; Poiesz, B. J.; Kramer, F. R. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96*, 6394-6399.
74. Marras, S. A. E.; Tyagi, S.; Kramer, F. R. *Clin. Chim. Acta* **2006,** *363*, 48-60.
75. Whitcombe, D.; Theaker, J.; Guy, S. P.; Brown, T.; Little, S. *Nat. Biotechnol.* **1999,** *17*, 804-807.
76. Thelwell, N.; Millington, S.; Solinas, A.; Booth, J.; Brown, T. *Nucleic Acids Res.* **2000,** *28*, 3752-3761.
77. Solinas, A.; Brown, L. J.; McKeen, C.; Mellor, J. M.; Nicol, J. T. G.; Thelwell, N.; Brown, T. *Nucleic Acids Res.* **2001,** *29*, e96;1-9.
78. French, D. J.; Archard, C. L.; Andersen, M. T.; McDowell, D. G. *Mol. Cell. Probes* **2002,** *16*, 319-326.
79. French, D. J.; Archard, C. L.; Brown, T.; McDowell, D. G. *Mol. Cell. Probes* **2001,** *15*, 363-374.
80. Dobson, N.; McDowell, D. G.; French, D. J.; Brown, L. J.; Mellor, J. M.; Brown, T. *Chem. Commun.* **2003**, 1234-1235.
81. Albert, A.; Taguchi, H. *J. Chem. Soc., Perkin Trans. 2* **1973**, 1101-1103.
82. Holmen, A.; Nordén, B.; Albinsson, B. *J. Am. Chem. Soc.* **1997,** *119*, 3114-3121.
83. Guest, C. R.; Hochstrasser, R. A.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1991,** *30*, 3271-3279.
84. Rachofsky, E. L.; Osman, R.; Ross, J. B. A. *Biochemistry* **2001,** *40*, 946-956.
85. Stivers, J. T. *Nucleic Acids Res.* **1998,** *26*, 3837-3844.
86. Sanabia, J. E.; Goldner, L. S.; Lacaze, P. A.; Hawkins, M. E. *J. Phys. Chem. B* **2004,** *108*, 15293-15300.
87. Liu, C. H.; Martin, C. T. *J. Mol. Biol.* **2001,** *308*, 465-475.

88. Wilhelmsson, L. M.; Sandin, P.; Holmen, A.; Albinsson, B.; Lincoln, P.; Nordén, B. *J. Phys. Chem. B* **2003,** *107,* 9094-9101.
89. Ranasinghe, R. T.; Rusling, D. A.; Powers, V. E. C.; Fox, K. R.; Brown, T. *Chem. Commun.* **2005**, 2555-2557.
90. Gerrard, S. R.; Srinivasan, N.; Fox, K. R.; Brown, T. *Nucleos. Nucleot. Nucl.* **2007,** *26,* 1363-1367.
91. Gerrard, S. R.; Edrees, M. M.; Bouamaied, I.; Fox, K. R.; Brown, T. *Org. Biomol. Chem.* **2010,** *8*, 5087-5096.
92. Booth, J.; Brown, T.; Vadhia, S. J.; Lack, O.; Cummins, W. J.; Trent, J. O.; Lane, A. N. *Biochemistry* **2005,** *44*, 4710-4719.
93. Hobbs, F. W. *J. Org. Chem.* **1989,** *54*, 3420-3422.
94. Sonogashira, K.; Tohda, Y.; Hagihara, N. *Tetrahedron Lett.* **1975**, 4467-4470.
95. Xiao, Q.; Ranasinghe, R. T.; Tang, A. M. P.; Brown, T. *Tetrahedron* **2007,** *63*, 3483-3490.
96. Crisp, G. T.; Flynn, B. L. *J. Org. Chem.* **1993,** *58*, 6614-6619.
97. Cai, C.; Vasella, A. *Helv. Chim. Acta* **1995,** *78*, 732-757.
98. Qu, D. Z.; Wang, G. X.; Wang, Z.; Zhou, L.; Chi, W. L.; Cong, S. J.; Ren, X. S.; Liang, P. Z.; Zhang, B. L. *Anal. Biochem.* **2011,** *417,* 112-121.
99. Kocalka, P.; El-Sagheer, A. H.; Brown, T. *ChemBioChem* **2008,** *9*, 1280-1285.
100. McGuigan, C.; Barucki, H.; Blewett, S.; Carangio, A.; Erichsen, J. T.; Andrei, G.; Snoeck, R.; De Clercq, E.; Balzarini, J. *J. Med. Chem.* **2000,** *43,* 4993-4997.
101. Kelley, S. O.; Barton, J. K. *Science* **1999,** *283,* 375-381.
102. Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. *J. Phys. Chem.* **1996,** *100,* 5541-5553.
103. Steenken, S.; Jovanovic, S. V. *J. Am. Chem. Soc.* **1997,** *119*, 617-618.
104. Dolghih, E.; Roitberg, A. E.; Krause, J. L. *J. Photochem. Photobiol., A* **2007,** *190*, 321-327.
105. Sabanayagam, C. R.; Eid, J. S.; Meller, A. *J. Chem. Phys.* **2005,** *122*, 061103;1-5.
106. Chowdhury, A.; Yu, L.; Raheem, I.; Peteanu, L.; Liu, L. A.; Yaron, D. J. *J. Phys. Chem. A* **2003,** *107*, 3351-3362.
107. Hannah, K. C.; Armitage, B. A. *Accounts Chem. Res.* **2004,** *37*, 845-853.
108. Tomlinson, A.; Frezza, B.; Kofke, M.; Wang, M.; Armitage, B. A.; Yaron, D. *Chem. Phys.* **2006,** *325,* 36-47.
109. Seifert, J. L.; Connor, R. E.; Kushon, S. A.; Wang, M.; Armitage, B. A. *J. Am. Chem. Soc.* **1999,** *121*, 2987-2995.

110. Hannah, K. C.; Gil, R. R.; Armitage, B. A. *Biochemistry* **2005,** *44*, 15924-15929.
111. Garoff, R. A.; Litzinger, E. A.; Connor, R. E.; Fishman, I.; Armitage, B. A. *Langmuir* **2002,** *18*, 6330-6337.
112. Fegan, A.; Shirude, P. S.; Balasubramanian, S. *Chem. Commun.* **2008**, 2004-2006.
113. Mason, S. J.; Hake, J. L.; Nairne, J.; Cummins, W. J.; Balasubramanian, S. *J. Org. Chem.* **2005,** *70*, 2939-2949.
114. Tomasulo, M.; Sortino, S.; Raymo, F. M. *J. Org. Chem.* **2008,** *73*, 118-126.
115. Zimmermann, T.; Hennig, L. *J. Heterocycl. Chem.* **2002,** *39*, 263-269.
116. Lambert, R. W.; Martin, J. A.; Thomas, G. J.; Duncan, I. B.; Hall, M. J.; Heimer, E. P. *J. Med. Chem.* **1989,** *32*, 367-374.
117. Graham, D.; Parkinson, J. A.; Brown, T. *J. Chem. Soc., Perkin Trans. 1* **1998**, 1131-1138.
118. Gierlich, J.; Burley, G. A.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T. *Org. Lett.* **2006,** *8*, 3639-3642.
119. Gierlich, J.; Gutsmiedl, K.; Gramlich, P. M. E.; Schmidt, A.; Burley, G. A.; Carell, T. *Chem.-Eur. J.* **2007,** *13*, 9486-9494.
120. Levitus, M.; Ranjit, S. *Q. Rev. Biophys.* **2011,** *44*, 123-151.
121. Aramendia, P. F.; Negri, R. M.; Sanroman, E. *J. Phys. Chem.* **1994,** *98*, 3165-3173.
122. Chibisov, A. K.; Zakharova, G. V.; Gorner, H. *J. Chem. Soc., Faraday Trans.* **1996,** *92*, 4917-4925.
123. Chibisov, A. K.; Zakharova, G. V.; Goerner, H.; Sogulyaev, Y. A.; Mushkalo, I. L.; Tolmachev, A. I. *J. Phys. Chem.* **1995,** *99*, 886-893.
124. Murphy, S.; Sauerwein, B.; Drickamer, H. G.; Schuster, G. B. *J. Phys. Chem.* **1994,** *98*, 13476-13480.
125. Sako, Y.; Minoguchi, S.; Yanagida, T. *Nature Cell Biol.* **2000,** *2*, 168-172.
126. Rindermann, J. J.; Akhtman, Y.; Richardson, J.; Brown, T.; Lagoudakis, P. G. *J. Am. Chem. Soc.* **2011,** *133*, 279-285.
127. Silva, G. L.; Ediz, V.; Yaron, D.; Armitage, B. A. *J. Am. Chem. Soc.* **2007,** *129*, 5710-5718.
128. Borjesson, K.; Preus, S.; El-Sagheer, A. H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. M. *J. Am. Chem. Soc.* **2009,** *131*, 4288-4293.
129. Johnson, N. P.; Baase, W. A.; von Hippel, P. H. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 7169-7173.
130. Hernando, J.; van der Schaaf, M.; van Dijk, E. M. H. P.; Sauer, M.; Garcia-Parajo, M. F.; van Hulst, N. F. *J. Phys. Chem. A* **2003,** *107*, 43-52.

131. West, W.; Pearce, S. *J. Phys. Chem.* **1965,** *69*, 1894-1903.
132. Mayer-Enthart, E.; Wagner, C.; Barbaric, J.; Wagenknecht, H. A. *Tetrahedron* **2007,** *63*, 3434-3439.
133. Baumstark, D.; Wagenknecht, H. A. *Chem.-Eur. J.* **2008,** *14*, 6640-6645.
134. Barbaric, J.; Wagenknecht, H. A. *Org. Biomol. Chem.* **2006,** *4*, 2088-2090.
135. Berndl, S.; Wagenknecht, H. A. *Angew. Chem.-Int. Ed.* **2009,** *48*, 2418-2421.
136. Varghese, R.; Wagenknecht, H. A. *Chem.-Eur. J.* **2010,** *16*, 9040-9046.
137. Brotschi, C.; Mathis, G.; Leumann, C. J. *Chem.-Eur. J.* **2005,** *11*, 1911-1923.
138. Paris, P. L.; Langenhan, J. M.; Kool, E. T. *Nucleic Acids Res.* **1998,** *26*, 3789-3793.
139. Langenegger, S. M.; Haner, R. *Chem. Commun.* **2004**, 2792-2793.
140. Bouquin, N.; Malinovskii, V. L.; Guegano, X.; Liu, S. X.; Decurtins, S.; Haner, R. *Chem.-Eur. J.* **2008,** *14*, 5732-5736.
141. Wilson, J. N.; Gao, J.; Kool, E. T. *Tetrahedron* **2007,** *63*, 3427-3433.
142. Bouquin, N.; Malinovskii, V. L.; Haner, R. *Chem. Commun.* **2008**, 1974-1976.
143. Bouamaied, I.; Nguyen, T.; Ruhl, T.; Stulz, E. *Org. Biomol. Chem.* **2008,** *6*, 3888-3891.
144. Benvin, A. L.; Creeger, Y.; Fisher, G. W.; Ballou, B.; Waggoner, A. S.; Armitage, B. A. *J. Am. Chem. Soc.* **2007,** *129*, 2025-2034.
145. Vogelsang, J.; Cordes, T.; Tinnefeld, P. *Photochem. Photobiol. Sci.* **2009,** *8*, 486-496.
146. Hern, J. A.; Baig, A. H.; Mashanov, G. I.; Birdsall, B.; Corrie, J. E. T.; Lazareno, S.; Molloy, J. E.; Birdsall, N. J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 2693-2698.
147. Mayilo, S.; Ehlers, B.; Wunderlich, M.; Klar, T. A.; Josel, H. P.; Heindl, D.; Nichtl, A.; Kurzinger, K.; Feldmann, J. *Anal. Chim. Acta* **2009,** *646*, 119-122.
148. Waggoner, A. S.; Mujumdar, R. *International patent WO 99/31181* **1999**.
149. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.-Int. Ed.* **2002,** *41*, 2596-2599.
150. Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002,** *67*, 3057-3064.
151. El-Sagheer, A. H.; Brown, T. *Chem. Soc. Rev.* **2010,** *39*, 1388-1405.
152. Gerowska, M.; Hall, L. M.; Richardson, J. A.; Shelbourne, M.; Brown, T. *Tetrahedron* **2011**, DOI 10.1016/j.tet.2011.11.041.
153. Staudinger, H.; Meyer, J. *Helv. Chim. Acta.* **1919,** *2*, 635-646.
154. Richardson, J. A.; Gerowska, M.; Shelbourne, M.; French, D.; Brown, T. *ChemBioChem* **2010,** *11*, 2530-2533.

155. Rolland, V.; Kotera, M.; Lhomme, J. *Synth. Commun.* **1997,** *27*, 3505-3511.
156. May, J. P.; Brown, L. J.; van Delft, I.; Thelwell, N.; Harley, K.; Brown, T. *Org. Biomol. Chem.* **2005,** *3*, 2534-2542.
157. McKeen, C. M.; Brown, L. J.; Nicol, J. T. G.; Mellor, J. M.; Brown, T. *Org. Biomol. Chem.* **2003,** *1*, 2267-2275.
158. Ikuta, S.; Takagi, K.; Wallace, R. B.; Itakura, K. *Nucleic Acids Res.* **1987,** *15*, 797-811.
159. Ke, S. H.; Wartell, R. M. *Nucleic Acids Res.* **1993,** *21*, 5137-5143.
160. Allawi, H. T.; SantaLucia, J. *Biochemistry* **1998,** *37*, 9435-9444.
161. Venkatesan, N.; Seo, Y. J.; Kim, B. H. *Chem. Soc. Rev.* **2008,** *37*, 648-663.
162. Bernacchi, S.; Mely, Y. *Nucleic Acids Res.* **2001,** *29*, e62.
163. Conley, N. R.; Pomerantz, A. K.; Wang, H.; Twieg, R. J.; Moerner, W. E. *J. Phys. Chem. B* **2007,** *111*, 7929-7931.
164. Nesterova, I. V.; Erdem, S. S.; Pakhomov, S.; Hammer, R. P.; Soper, S. A. *J. Am. Chem. Soc.* **2009,** *131*, 2432-2433.
165. Smith, L. M.; Fung, S.; Hunkapiller, M. W.; Hunkapiller, T. J.; Hood, L. E. *Nucleic Acids Res.* **1985,** *13*, 2399-2412.
166. Kaiser, R. J.; Mackellar, S. L.; Vinayak, R. S.; Sanders, J. Z.; Saavedra, R. A.; Hood, L. E. *Nucleic Acids Res.* **1989,** *17*, 6087-6102.
167. El-Sagheer, A. H.; Brown, T. *Curr. Protoc. Nucleic Acid Chem.* **2008**, 4.33.1-4.33.21.
168. Babadzhanova, L. A.; Kirij, N. V.; Yagupolskii, Y. L.; Tyrra, W.; Naumann, D. *Tetrahedron* **2005,** *61*, 1813-1819.
169. Deglane, G.; Morvan, F.; Debart, F.; Vasseur, J. J. *Bioorg. Med. Chem. Lett.* **2007,** *17*, 951-954.
170. Morvan, F.; Rayner, B.; Leonetti, J. P.; Imbach, J. L. *Nucleic Acids Res.* **1988,** *16*, 833-848.
171. Damha, M. J.; Giannaris, P. A.; Zabarylo, S. V. *Nucleic Acids Res.* **1990,** *18*, 3813-3821.
172. El-Sagheer, A. H.; Brown, T. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 15329-15334.
173. Karstens, T.; Kobs, K. *J. Phys. Chem.* **1980,** *84*, 1871-1872.
174. Eastman, J. W. *Photochem. Photobiol.* **1967,** *6*, 55-72.