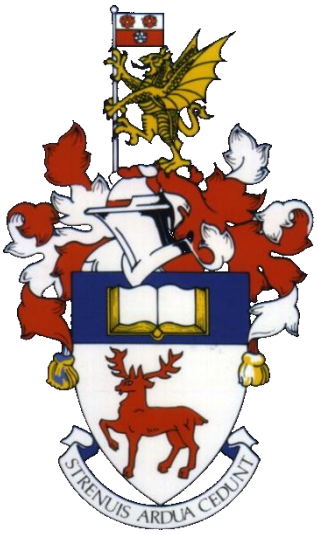
UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

School of Chemistry



PORPHYRIN SUBSTITUTED DNA: BUILDING BLOCKS FOR NOVEL NANOSTRUCTURES

by

Daniel George Singleton

Thesis for the degree of Doctor of Philosophy

October 2012

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMETAL SCIENCES
SCHOOL OF CHEMISTRY

Doctor of Philosophy

PORPHYRIN SUBSTITUTED DNA: BUILDING BLOCKS FOR NOVEL NANOSTRUCTURES

By Daniel George Singleton

Porphyrin-modified DNA duplexes have been studied in detail for the last 10 years. It is well known that G-rich DNA sequences of the type $(GGGX)_4$ can form stable intramolecular quadruplex structures containing stacks of G-quartets in the presence of monovalent cations such as potassium. The synthesis of novel G-quadruplex forming DNA sequences based on the human telomeric repeat $(GGGTTA)_x$ where porphyrins are covalently attached to thymidine is presented. It is shown that the porphyrin modification offers a large degree of stabilisation to the G-quadruplex. Structural studies have revealed that the porphyrin modification does not bias the formation of a single G-quadruplex. However, one exception was observed for a G-quadruplex with two porphyrin modifications. Further insight into the interactions between the porphyrin and the G-quadruplex were achieved by molecular dynamic simulations.

The synthesis of novel porphyrin-modified locked nucleic acids (LNA) is presented. It was hoped that the LNA modification would provide stability to the porphyrin-modified DNA duplex as porphyrin modifications have previously been found to destabilise duplex DNA. The structure and stability of the DNA was assessed in detail using circular dichroism spectroscopy.

Finally, a new method of synthesising porphyrin-modified DNA using polymerases to incorporate porphyrin triphosphates into oligonucleotides is presented. The synthesis of novel porphyrin triphosphates is described. A variety of polymerases were assessed on whether they could incorporate these triphosphates into oligonucleotides. Multiple incorporations of these porphyrin-modified triphosphates showed that they were a good substrate for certain polymerases and with further development this method could be a viable route to synthesis of porphyrin-modified DNA.

Contents

List of figures

List of schemes

List of tables

DECLARATION OF AUTHORSHIP

I, Daniel George Singleton declare that the thesis entitled:

Porphyrin substituted DNA: Building blocks for novel nanostructures

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 or mainly 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;
- parts of this work have been published as:

  Singleton, D. G.; Burns, J. R.; Fox, K. R.; Siligardi, G.; Hussain, R.; Várnai, P.; Stulz, E.; Modulation of G-quadruplex structure and stability through selective covalent modification with porphyrins. *Manuscript in preparation*.

Signed: ……

Date:……

Acknowledgements

Firstly, I would like to thank Dr. Eugen Stulz for providing me with the opportunity to carry out this research and also for the countless amount of advice given to me throughout the course of my PhD. I would also like to thank the other members of the Stulz research group and associates of lab 5003; Ashley Brewer, Tom Bandy, Jon Burns, Michele Carbone, ThaoNguyen Nguyen, Gabi Marth, Simon Gerrard, James Wood, Warren Duffy, Andy Naylor and Stuart Perkins. Not only has their advice has been invaluable over the years, but they created a work environment that was great to be a part of. Special thanks in particular must go to Warren and James for taking the time to proof read my thesis.

I would like to thank Giuliano Siligardi and Rohanah Hussain at the Diamond Light Source for their help and advice on circular dichroism and I am especially grateful to Peter Várnai for carrying out some molecular modelling for this project. I would also like to thank David Rusling and Annie Cardew for their help with radiolabelling DNA and taking the time to help me with gel electrophoresis experiments. I would also like to thank Stuart Findlow for carrying out NMR experiments on G-quadruplex DNA.

I would like to thank Michal Hocek for providing me with opportunity to go to Prague and spend some time working in his lab. I would also like to thank Jan Riedl and Radek Pohl whose help during this time was invaluable.

Finally, I would like to thank my family and friends, in particular my parents for their continuous support throughout these last four years.

Definitions and abbreviations

A	Adenine
ATP	Adenosine triphosphate
aq.	Aqueous
br. s	Broad singlet
C	Cytosine
°C	Degrees centigrade
CD	Circular dichroism
$CDCl_3$	Deuterated chloroform
CD_3OD	Deuterated methanol
CEP-Cl	2-*O*-Cyanoethyl-*N*-*N*-di-*iso*-propylchlorophosphoramidite
conc.	Concentrated
CPG	Controlled porous glass
d	Doublet
dATP	2'-Deoxyadenosine triphosphate
DBStv	Dynabeads M-270 Streptavidin
DCM	Dichloromethane
dCTP	2'-Deoxycytidine triphosphate
DDQ	2,3-Dichloro-5,6-dicyanobenzoquinone
dGTP	2'-Deoxyguanosine triphosphate
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DMT	4,4'-Dimethoxytrityl
DNA	2'-Deoxyribonucleic acid
dNTP	2'-Deoxynuceloside triphosphate
dTTP	2'-Deoxythymidine triphosphate
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray Ionisation
G	Guanine
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HFIP	1,1,1,3,3,3-Hexafluoro-2-propanol
HOBt	Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
hr	Hours
Hz	Hertz

J	Coupling constant
LNA	Locked Nucleic Acid
m	Multiplet
m	Meta
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionisation Time of Flight
MeCN	Acetonitrile
MeOH	Methanol
min	Minutes
MsCl	Methanesulfonyl chloride
m/z	Mass to charge ratio
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
o	Ortho
p	Para
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PEX	Primer Extension
ppm	Parts per million
PyBOP	(Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
PyBrOP	Bromotripyrrolidinophosphonium hexafluorophsophate
q	Quartet
R_f	Retention factor
RT	Room temperature
s	Singlet
T	Thymine
t	Triplet
TCA	Trichloroacetic acid
TEA	Triethylamine
TEAA	Triethylamine acetate
TEAB	Triethylamine bicarbonate
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
Tm	Melting temperature
UV-Vis	Ultra violet and visible spectroscopy

1 Introduction

1.1 Structure of DNA

1.1.1 Primary structure

The elucidation of the structure of double-stranded DNA by Watson, Crick, Wilkins and Franklin is arguably one of the most important discoveries of the twentieth century.[1-3] The variation in the sequence of bases in DNA is responsible for the huge diversity of life on earth. The role of 2'-deoxyribose nucleic acid (DNA) is to store genetic information. The single strand form of DNA comprises of three parts; the nucleobase, the sugar ring and the phosphate backbone. There are four heterocyclic nucleobases present in DNA, adenosine (A), guanine (G), cytosine (C) and thymine (T) (Figure 1.1). Adenine and guanine are both purines whereas cytosine and thymine are pyrimidines.

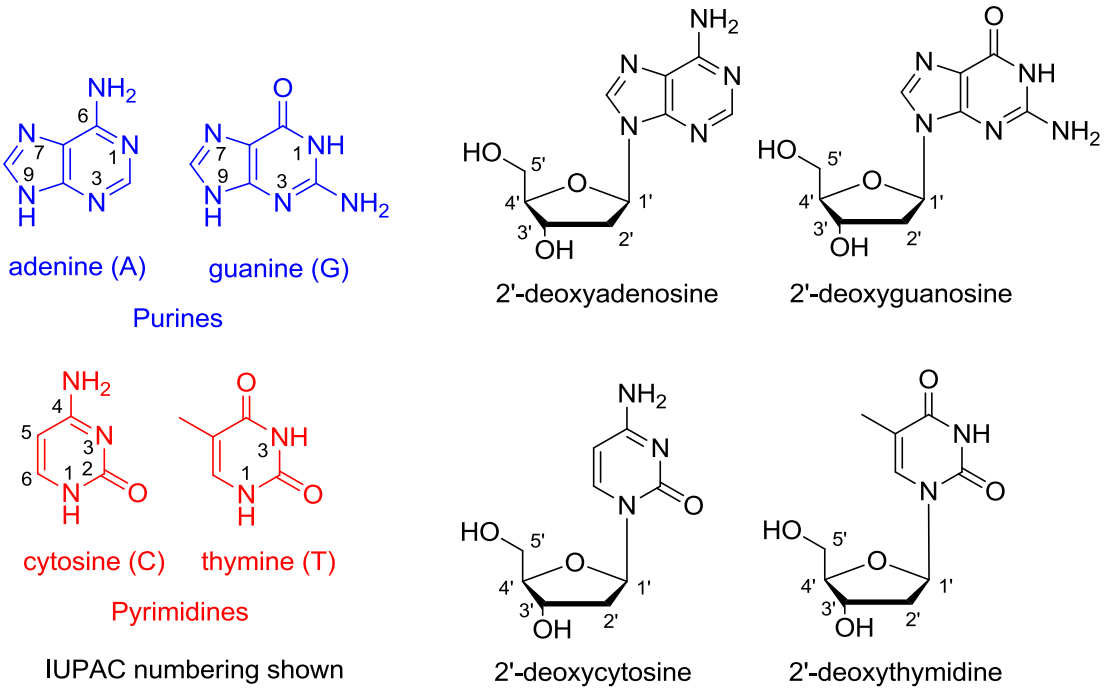


Figure 1.1 The four heterocyclic bases and their 2'-deoxyribonucleoside equivalents found in DNA

Ribonucleic acid (RNA) is similar to DNA in that its role is to encode genetic information; however, there are two structural differences. The first is the sugar is ribofuranose, which has a hydroxyl group at the 2'-position. The second is that thymine is replaced with uracil, which lacks a methyl group at the 5-position

(Figure 1.2). The roles of DNA and RNA are described by the central dogma of molecular biology. The dogma gives an understanding into the transfer of genetic information, starting with DNA, and then onto RNA. The information within the RNA sequence is used to direct the synthesis of proteins.

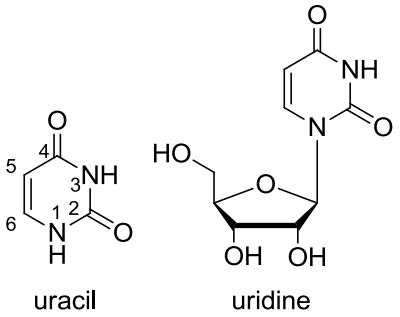


Figure 1.2 Uracil and its corresponding ribonucleoside, uridine

When the nucleobase is joined with a 2'-deoxy-D-ribose ring a nucleoside is formed. The heterocyclic base is covalently bonded to the 1'-position of the sugar via an *N*-glycosidic bond at the N^1-atom for pyrimidine bases and at the N^9-atom for purine bases. Rotation about the *N*-glycosidic bond can occur which gives rise to the *syn* and *anti* conformations of DNA (Figure 1.3). The *anti* configuration is generally favoured due to sterics but both *syn* and *anti* configurations can be observed in some DNA secondary structures such as G-quadruplexes.

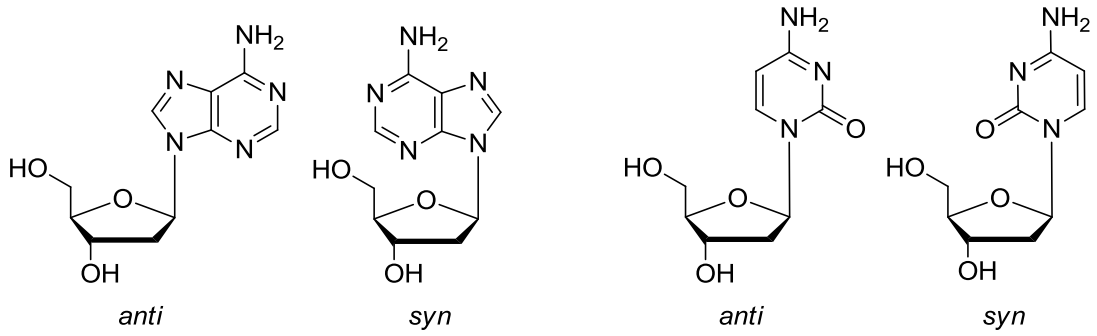


Figure 1.3 *Syn* and *anti* conformations of nucleoside

There are two possible conformations for the furanose ring of the nucleoside. These are the C2'-*endo* conformation which is mostly seen in DNA duplexes and C3'-*endo* conformation which is more often seen in RNA duplexes (Figure 1.4). There are other instances where it is possible that one conformation will be prevalent over the other and this will often have an effect of the overall structure of the oligonucleotide. These conformations are generally referred to as "sugar puckering".

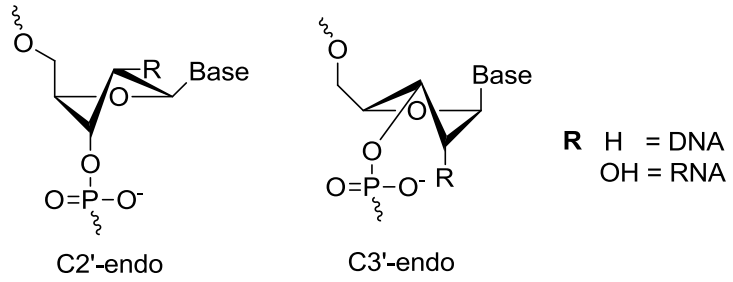


Figure 1.4 C2'-*endo* and C3'-*endo* furanose ring conformations

Nucleosides are connected together as an oligonucleotide via a phosphodiester backbone (Figure 1.5). The sugar phosphate backbone is chiral in nature and has directionality. Nucleosides are joined to the phosphate backbone by the 5' and 3' positions of the ribose ring. Because of this directionality the sequence 5' – ACGT – 3' is not the same as 5' – TGCA – 3'.

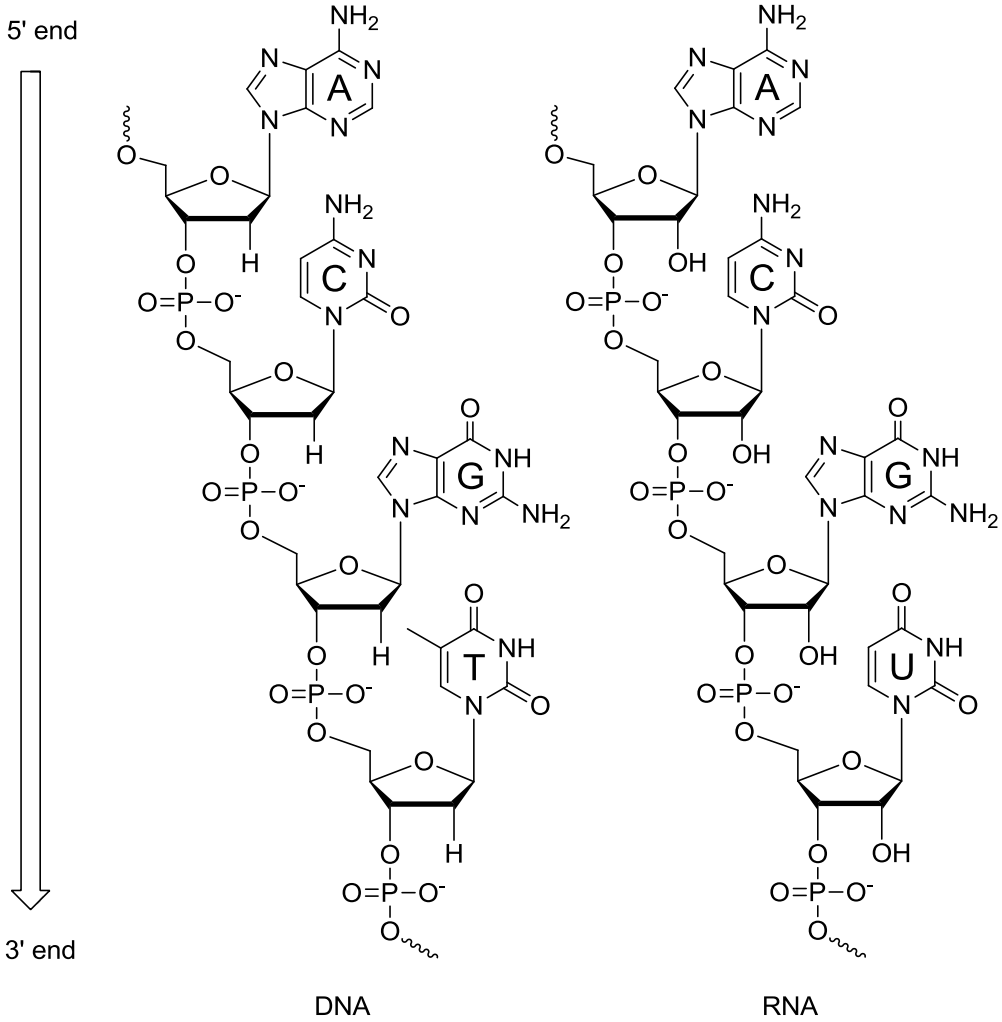


Figure 1.5 Nucleic acid primary structure

1.1.2 Secondary structure of DNA – double helix

The most well-known form of DNA is the double helix. This structure comprises of two complementary strands of DNA that hybridise together via Watson-Crick base pairing. When the strands are paired they coil around each other give a right-handed helical structure with the hydrophobic bases in the centre and the hydrophilic phosphate backbone exposed on the outside (Figure 1.7). The term "right-handed" with respect to the double helix indicates that the backbone at the front of the molecule facing the observer slopes down from top right to bottom left. The planar heterocyclic bases stack upon on another with a separation between successive base pairs along the helical axis of around 0.34 nm. One helical turn (a full 360° turn of the double helix) is repeated every 10 to 11 base pairs. There are two grooves that run through the entire length of the double helix. These grooves are termed the major groove and the minor groove, named due to their respective size to each other. The major groove is typically 22 Å wide and the minor groove 12 Å wide for B-DNA.[4] The two strands that make up this structure run in antiparallel directions.

Watson-Crick base pairing is responsible for the overall structure and also the specific nature of the complementary strand. With this type of base pairing, nucleobases interact with each other by forming intermolecular hydrogen bonds (Figure 1.6). It was Chargaff who found that the number of adenosine bases present in the DNA was equal to the number of thymidine bases and the same for cytosine and guanine bases.[5] This led to the conclusion that adenine will pair with thymine and that cytosine will pair with guanine which was found to be a consequence of the complementary hydrogen-bonding pattern. For an adenosine – thymidine pairing two hydrogen bonds are formed with a bonding strength of 15-25 kJ mol^{-1} per base pair, whereas for a cytosine – guanine pairing three hydrogen bonds are formed with a bonding strength of 25-40 kJ mol^{-1} per base pair.[6, 7]

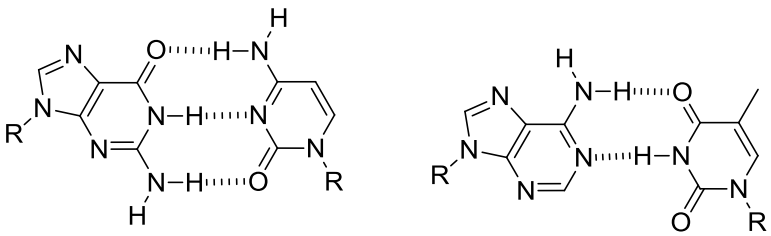


Figure 1.6 Watson-Crick basing pairing of natural nucleobases

So far we have just looked at the most observed form of duplex DNA – B-DNA. There are two other forms of DNA duplexes that have been observed, A-DNA and Z-

DNA (Figure 1.7). A-DNA structure is generally more associated with RNA duplexes and is more of a compressed right-handed helix with a deep major groove and a very narrow minor groove.[8] Z-DNA is the left-handed helical structure observed by Wang, Rich and co-workers when the hexamer d(CGCGCG) was crystallised.[9, 10] This structure is very different to that of the A- and B- forms of DNA, One helical turn contains twelve base pairs and there is little difference in size between the major and minor grooves. This structure is only formed in GC rich sequences in high salt conditions. Proteins have been found to bind to Z-DNA but as of yet it is not known whether this form of DNA has any biological significance.

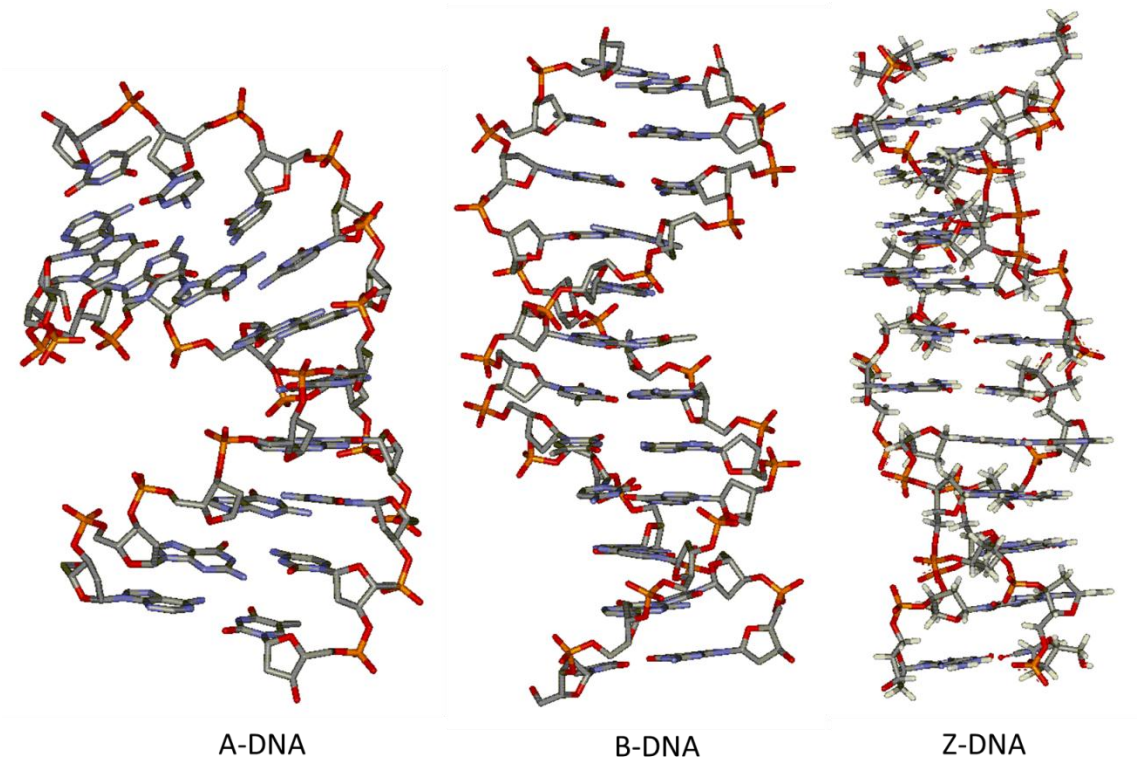


Figure 1.7 DNA duplex forms from crystal structures[9, 11, 12]

1.2 Other secondary DNA structures

1.2.1 Triplexes, G-quadruplexes and i-motifs

Secondary structures of DNA are not just limited to DNA duplexes, other arrangements such as DNA-triplexes, G-quadruplexes and i-motifs have been observed.[13] Triple stranded DNA (or triplex) was discovered by Felsenfeld and Rich in 1957[14] when a 2:1 complex of polyuridylic acid (polyU) and polyadenylic acid (polyA) was formed. The major groove of a DNA duplex is large enough to accommodate and bind a third strand of DNA (or RNA) and form the triplex structure[15] (Figure 1.8). The

third strand may be parallel or anti-parallel relative to the purine strand. The parallel triplexes are more stable than their anti-parallel equivalents. Binding of the third DNA strand to the major groove is possible as Hoogsteen hydrogen bonds can be formed to the N-7 of the purine bases of the Watson-Crick base pairs. In a parallel triplex a GC base pair is recognised by a protonated cytosine to give C^+.GC and an AT base pair is bound to thymine to give T.AT. Triplexes have since been used in gene therapy and as a tool in molecular biology and biotechnology applications.[16-18]

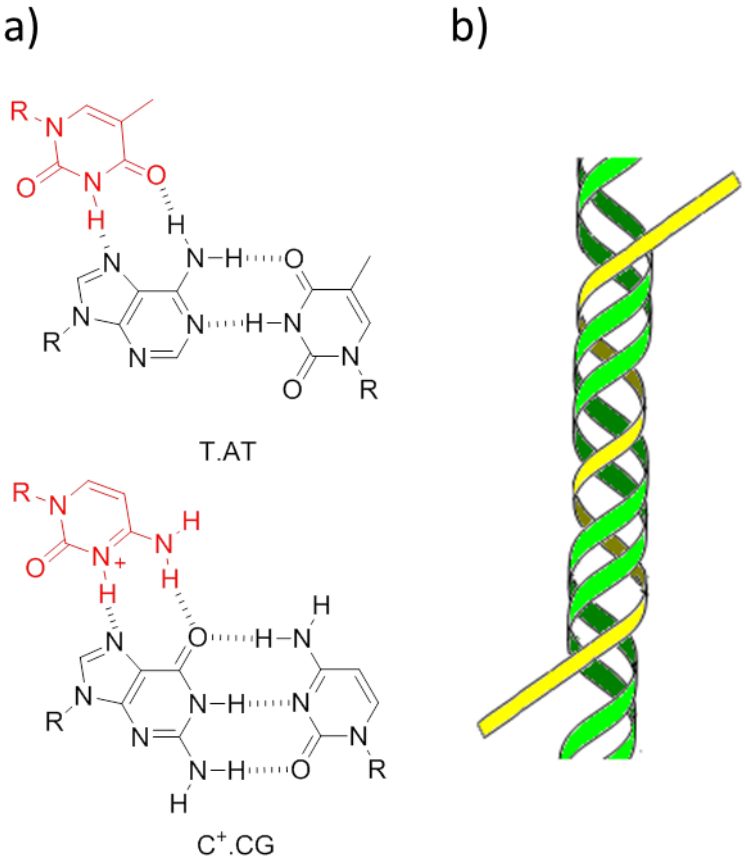


Figure 1.8 a) Hydrogen bonding between nucleobases present in DNA triplex formation b) Schematic of a DNA triplex. The duplex strands are shown in green and the third (triplex-forming) strand is shown in yellow

G-quadruplexes are formed in G-rich regions of DNA where four guanines associate together using their Hoogsteen and Watson-Crick faces (N_1-O_6 and N_2-N_7) to form a G-quartet with a cation present in the centre. The G-quartets can stack on top of each other to form the G-quadruplex structure (Figure 1.9). G-quadruplexes can be

formed from one, two or four strands of DNA. Theory on G-quadruplex formation, structure and study will be covered in more detail in Chapter 4.

A similar quadruplex-forming sequence of DNA is the i-motif which can be formed from a cytosine rich strand with the protonation of the intercalated $C.C^+$ base pairs (Figure 1.9).[19, 20] As one of the cytosines needs to be protonated, i-motifs are formed at an acidic or neutral pH and like G-quadruplex DNA, the i-motif can be unimolecular, bimolecular or tetramolecular.[21]

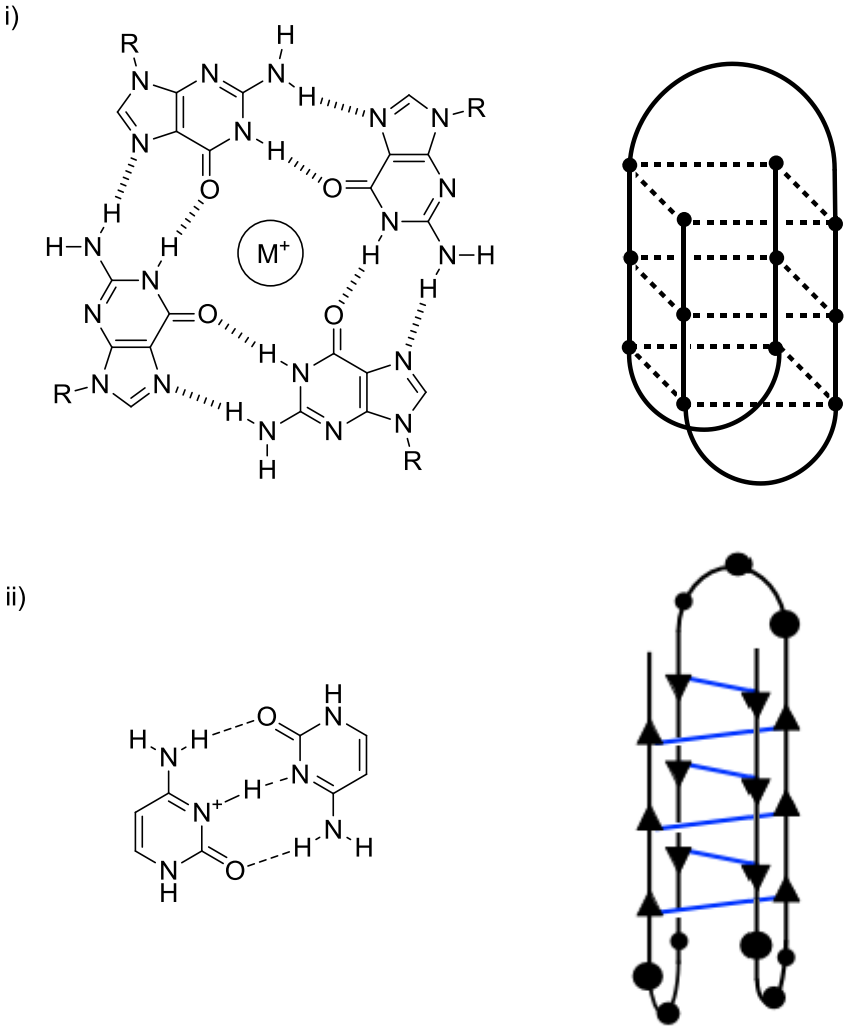


Figure 1.9 (i) G-quadruplex and (ii) i-motif secondary DNA structures

1.2.2 Higher complexity DNA structure and nanotechnology

Over the past 30 years, the field of DNA nanotechnology has expanded due to the ability of DNA to selectively self-assemble with complementary strands. The work of Nadrian Seeman was crucial for the development of nanoscale assemblies[22-24] and was further explored by Rothemund who introduced the concept of DNA origami. The principle of DNA origami is that a long viral single strand of DNA is folded into a new shape by addition of hundreds of single strand oligonucleotides called "staple strands". Each of the staple strands is designed to bind to the viral DNA in specific positions to facilitate the formation of a new nanostructure. This was famously demonstrated with the formation of larger 2D shapes such as stars and smiley faces (Figure 1.10).[25]

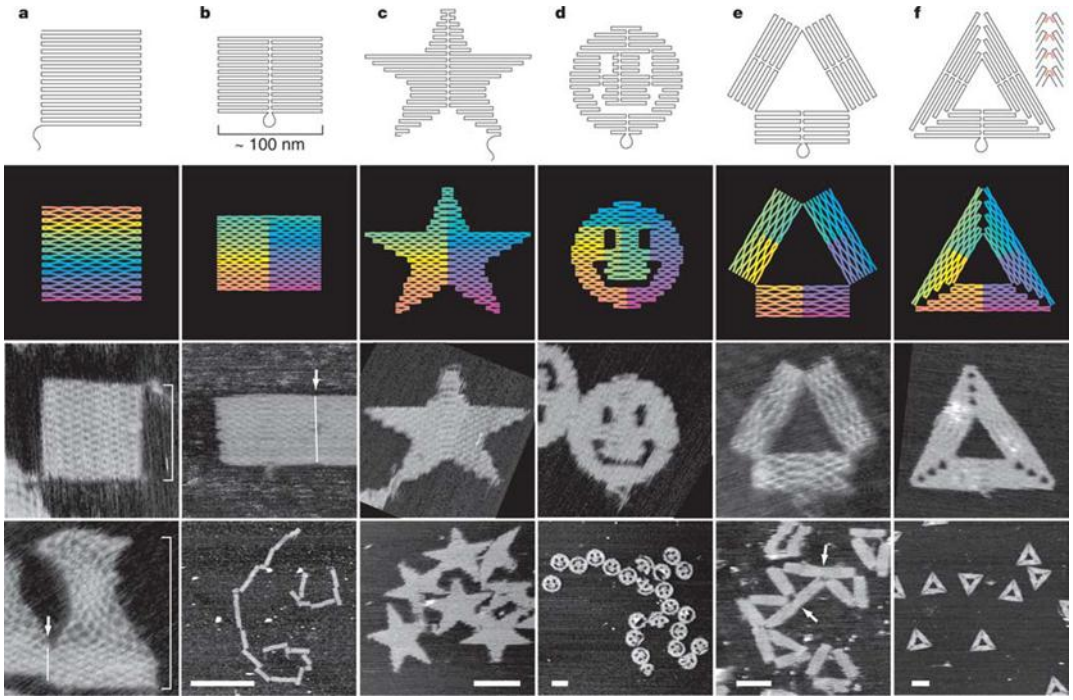


Figure 1.10 First examples of DNA origami. The top row illustrates the designs. The second row from top show diagrams of the folding indicating the bend of the helices at crossovers (where helices touch) and away from crossovers (where helices bend apart). The bottom two rows contain the resulting DNA structures as imaged by atomic force microscopy[25]

Since then the concept of DNA origami has been explored further to create nanostructures of DNA that have potentially useful applications. One example is the construction of a DNA box that can be opened with specific DNA keys (Figure 1.11).[26] This 3D example of DNA origami has the potential to transport smaller molecules and perform a controlled release of them when required.

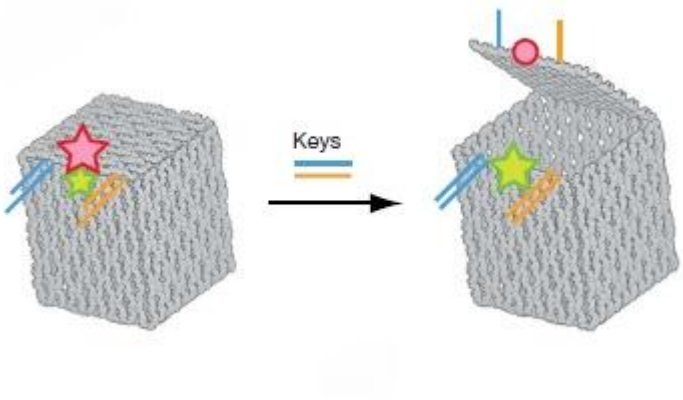


Figure 1.11 Schematic of a DNA box. The controlled opening of the box is carried out with the addition of DNA keys[26]

DNA origami is not just limited to using Watson-Crick base pairing to create DNA nanostructures. A dynamic system reported by Sugiyama[27] takes advantage of the formation and dissociation of a G-quadruplex by the addition and removal of a buffer solution containing potassium chloride (Figure 1.12). A DNA origami frame was designed with two parallel helices that bridge the frame. When potassium is introduced into the system a G-quadruplex is formed between the two helices to create an X-shaped state, which was observed by atomic force microscopy.

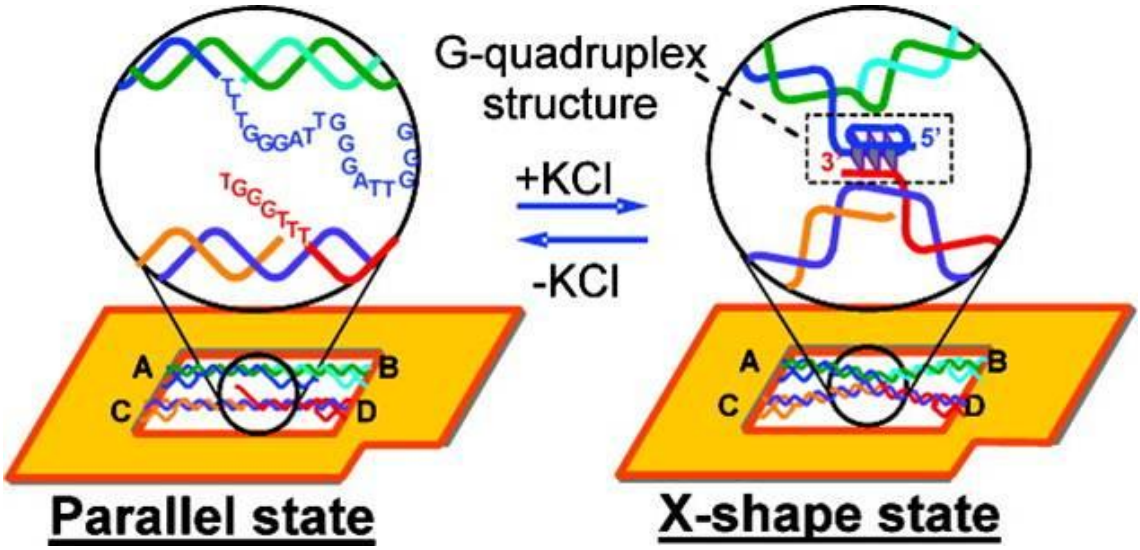


Figure 1.12 Dynamic G-quadruplex formation within a DNA origami frame. The top schematic shows the formation of a G-quadruplex when potassium is introduced[27]

1.3 DNA modifications

Although DNA origami using unmodified oligonucleotides is an exciting area of DNA technology, the modification of DNA unlocks a huge number of possibilities in taking advantage of the unique properties of DNA. The introduction of modifications into DNA can be achieved by several methods and at various positions for example 5'- and 3'-terminal positions, 2'- and 4'- positions on the ribose ring, the phosphate backbone and natural base modifications are all options. It is also possible to substitute a nucleobase with artificial nucleobases or molecules that mimic the function and structure of the DNA. Developments in automated solid phase DNA synthesis has opened the door to a wide range of possible DNA modifications provided they are compatible with the chemistry employed in this process.

1.3.1 5'-Terminus modifications

5'-Terminus modifications have been used for a variety of reasons; these include increased DNA stability from enzyme degradation, increased thermal stability through end-capping, and labelling of the DNA. Modifications at the 5'-position can be achieved in one of two ways. The first is to attach the modifier via a phosphate group and extend the DNA backbone. This can be carried out if the modifier contains a hydroxyl group that can be phosphitylated to the phosphoramidite for use in automated solid phase DNA synthesis. The second method that can be used is to functionalise the 5'-position of a monomer prior to addition using DNA synthesis. Popular 5-terminal modifications include the attachment of an amino or thiol modifier (Figure 1.13).[28, 29] A variety of molecules, such as fluorescent dyes or biotin, can then be attached via this modifier to the DNA post solid phase synthesis.

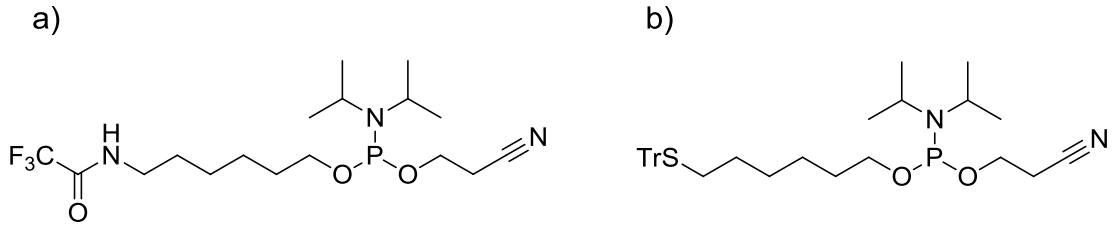


Figure 1.13 Examples of 5'-terminus modifiers a) 5'-amino modifier b) 5'-thiol modifier

1.3.2 3'-Terminus modifications

Modification of the 3'-terminal of oligonucleotides is important in the design of novel diagnostic probes and antisense oligonucleotides. By modifying the 3'-terminus, oligonucleotides have been found to be more resistant to digestion by 3'-exonucleases.[30] An important consideration when modifying the 3'-terminus is that automated DNA synthesis is carried out in the 3' to 5' direction with the 3'-position directly attached to the solid support. It is possible to carry out DNA synthesis in the opposite direction, as it is done in nature, but commercial 5'-phosphoramidites are expensive and the secondary 3'-hydroxyl is less reactive than the primary 5'-hydroxyl. For these reasons solid phase DNA synthesis in the 5' to 3' direction is generally not a popular choice. However, it is possible to modify the 3'-position and then attach the modifier directly to the solid support. This will allow the possibility of synthesis in the 3' to 5' direction and the solid support can be removed as standard after the synthesis is complete. Popular examples of this are 3'-amino-modifers[31] and 3'-thiol modifiers [32] (Figure 1.14).

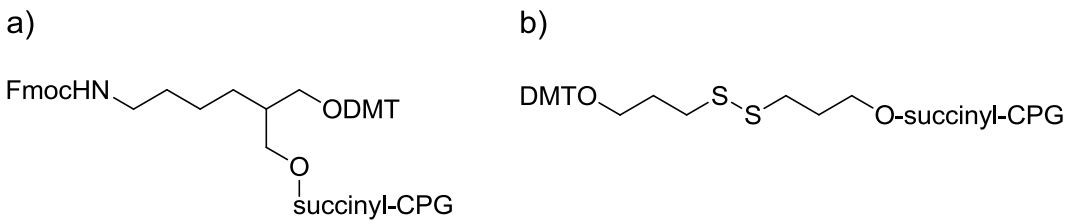


Figure 1.14 Examples of 3'-terminus modifiers on solid supports a) 3'-amino modifier b) 3'-thiol modifier

1.3.3 Ribose ring modifications

Modifications at the 2'-position can be made by taking a ribonucleoside and using its 2'-hydroxyl group to attach new moieties (Figure 1.15a).[33] Alternatively, thiols[34] and amines[35] can be added to this position so subsequent modifications can be made (Figure 1.15b and c respectively). Two other examples of ribose ring modifications are locked nucleic acids (LNA) and unlocked nucleic acids (UNA) (Figure 1.15d and e respectively) .[36] LNA has a methylene bridge between the 2'-oxygen and 4'-carbon of the ribose ring and locks the ribose ring in an RNA-like, C3'-*endo* conformation. UNA is a highly flexible structure due to a cleavage in the ribose ring between the 2'- and 3'-carbons. When modifying the ribose ring it is important to consider the nature of the modification and its relation to the structure of the DNA

duplex. Large modifications are generally not possible as the modification will be positioned within the helix itself and perturb the secondary structure.

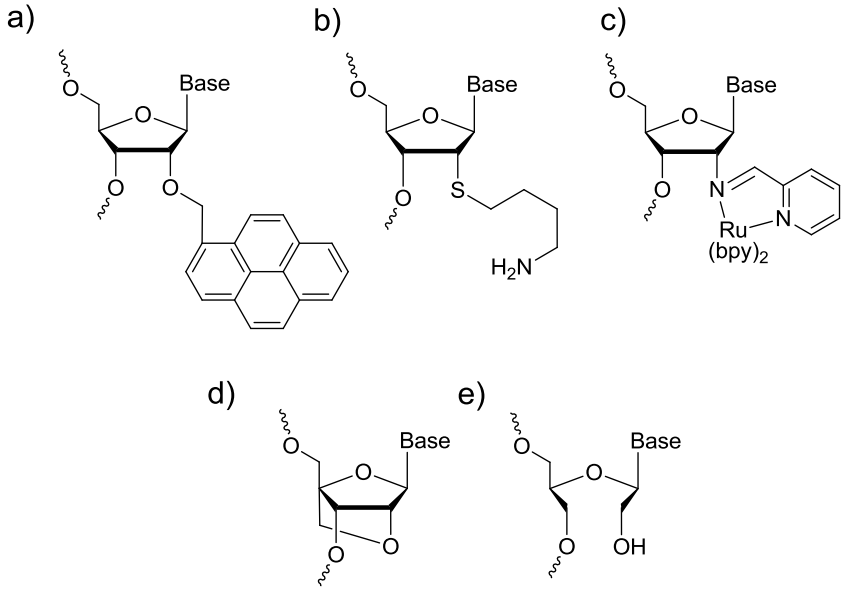


Figure 1.15 Examples of ribose ring modification[33-36]

1.3.4 Artificial nucleobases for internal DNA modification

The replacement of a natural nucleobase with an artificial one opens the door to a wide range of possible DNA modifications (Figure 1.16). Although replacing a natural nucleobase removes the Watson-Crick base paring that gives DNA its structure and stability, a duplex will still often be formed around the modification meaning that the overall structure is retained. Many replacements are flat aromatic compounds such as phenyl, bipyridiyl, pyrene and porphyrin. Although these molecules cannot offer any stability through hydrogen bonding, they can π-stack with neighbouring bases in the DNA duplex. In some cases these type of base replacements can increase the stability of the duplex, as observed by Leumann when employing a biphenyl replacement.[37] Base replacements using flat aromatic compounds are usually incorporated to act as a spectroscopic probe that report on the structure and function of enzymes and nucleic acids.[38, 39] There is also a growing interest in the extension of the genetic code by nucleobase replacement.[40-43]

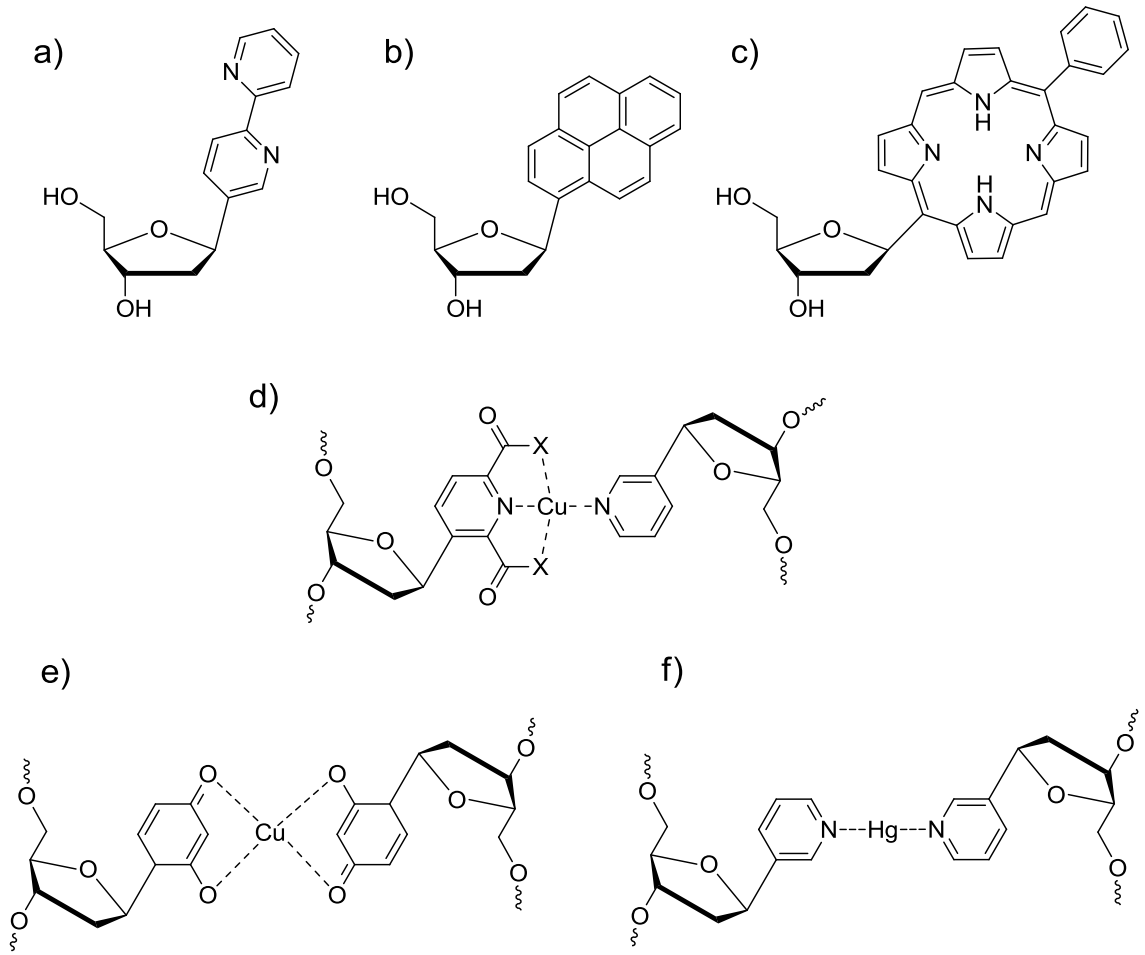


Figure 1.16 Examples of artificial base replacements

Other examples of artificial base replacements attempt to incorporate electrostatic interactions to help stabilise the DNA duplex structure. One strategy for this is to design the nucleobases so that two can chelate a metal ion when the DNA duplex is formed.[44, 45] Using these modifications one can build up a chain of metals within the double helix which could be used in future applications for energy transfer and nanowire formations.

1.3.5 Backbone modifications

Some of the larger modifications to natural DNA come in the shape of backbone replacement. In this example, part of the phosphate backbone is substituted with small organic molecules at either the end[46] or within the sequence of the oligonucleotide.[47] Häner *et al.* substituted the phosphate backbone in the mid-section of a DNA duplex with pyrene molecules and found, despite the achiral nature of pyrene, chirality was

induced to this backbone modification as a helical structure was formed. A backbone modification can also be used to extend the length of an oligonucleotide.[47] Sheppard *et al*.[48] placed an aldehyde modification at the 5'-terminal end of an oligonucleotide and the 3'-terminal end of a second oligonucleotide and connected them together using a diamine to form an imine. Preorganisation of the oligonucleotides was achieved by using a complementary oligonucleotide to act as a template.

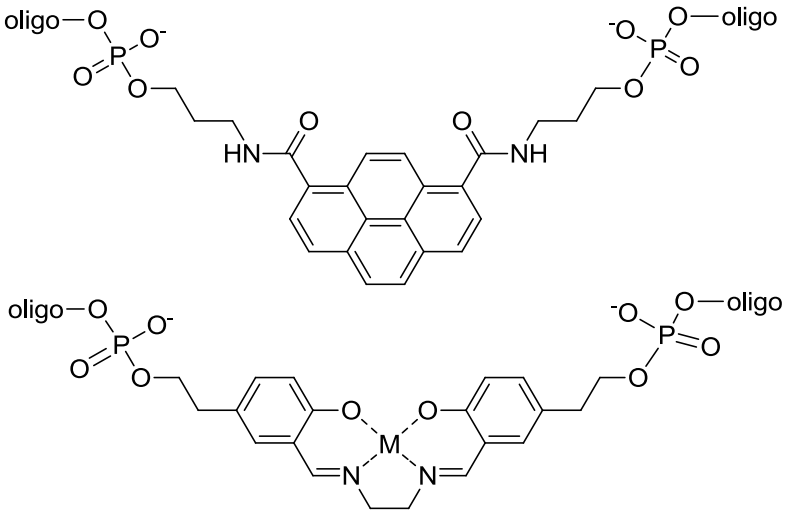


Figure 1.17 Examples of DNA backbone replacements[46]

1.3.6 Nucleobase modifications

Modifications on the nucleobases offer an attractive route for using DNA as a scaffold for supramolecular applications. Functional molecules such as porphyrins,[49] pyrenes,[50, 51] metallated bipyridines,[52, 53] terpyridines[54, 55] and other various alkyl and aryl substituents[56] can be attached to the nucleobases using palladium cross-coupling chemistry. Again, the helical nature of the DNA needs to be considered when functionalising the nucleobase. The ideal positions for attaching moieties are the C7 position for deazapurines and the C5 position for pyrimidines (Figure 1.18). In these positions, the modification will situate itself within the major groove of the double helix which is more accommodating for large modifications. The iodo nucleobases are all commercially available but the most often used one is 5-iodo-2'-deoxyuridine. This is because the nucleobase does not require any protection of additional functional groups and it is also the most cost effective starting material. The modifier is attached to the nucleobase using Sonogashira cross-coupling[57] or by Stille cross-coupling reactions.[58] These reactions are generally found to be facile and offer a wide range of possibilities in nucleobase functionalisation.

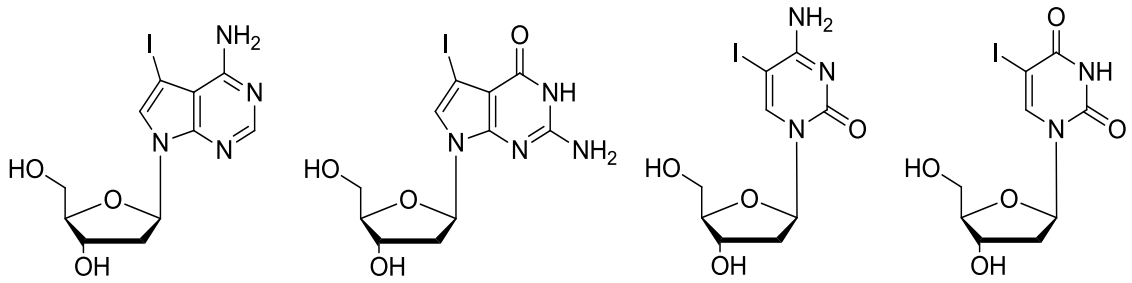


Figure 1.18 Iodinated nucleobases used to add substituents using Sonogashira palladium cross-coupling chemistry

1.4 Porphyrins

1.4.1 Structure of porphyrins

Research within this thesis will concern the modification of oligonucleotides with porphyrins. Before a discussion about the developments in this field of research, it is important to understand why the porphyrin modification is of great interest. Porphyrins are aromatic macrocycles composed of four pyrrole rings that are connected at the α-positions with four methine linkers. The aromaticity of the porphyrin comes from the conjugation of 18 π electrons (Figure 1.19) which satisfies Hückel's rule for aromaticity (where n = 4 for 4n + 2π electrons). There are a further 4 π electrons present from two double bonds, but these are not part of the aromatic system.[59]

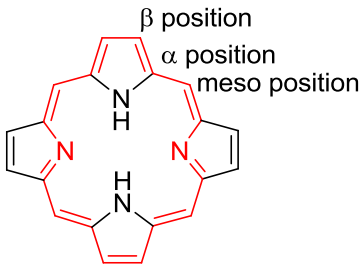


Figure 1.19 Structure of porphine with naming of positions. Aromatic region shown in red

The internal cavity of a porphyrin is approximately 2 Å across and is readily metallated to form the *21,23*-metallo porphyrin; example metals include zinc, copper,

lead, chromium, molybdenum and iron.[60-62] Smaller metal ions, such as Co^{2+} and Fe^{2+}, are able to sit within the plane of the porphyrin and not distort its structure whereas larger metal ions such as Hf^{4+} and Re^{+} are too big so they sit above the plane of the ring and distort the planarity of the ring.[63]

The insertion of metal ions combined with the substitution pattern of the porphyrin allows fine tuning of the system's electronic characteristics.[64] Substituents can be added to the porphyrin at the β-pyrrole positions, *meso* positions and the pyrrolic nitrogens. The addition of substituents at the β-pyrrole positions is usually carried out prior to formation of the porphyrin.[65] It is possible to add substituents after porphyrin synthesis but control over the stereospecificity of the addition is difficult. Substitutions at the *meso* positions are easily carried out prior to the porphyrin synthesis by using substituted aldehydes in the porphyrin synthesis.[66] Using benzaldehyde in the synthesis gives the tetraphenylporphyrin. To carry out multiple substitutions of the *meso* positions a variety of aldehydes must be used in specific ratios to obtain the desired product. Alternatively substituted dipyrromethanes can be used to gain more control over the *meso* substituted porphyrins that are formed during synthesis. Substitutions at the *meso* position are also possible after porphyrin synthesis, but again control over the position of the substitution can be difficult.

The electronic properties of porphyrins are responsible for their interesting absorption spectra. Free base porphyrins have many bands in the visible light region of the electromagnetic spectrum, the largest of these are the Soret bands or B-bands which come at 400-430 nm. They absorb light with an extremely high molar extinction coefficient that is around 1 $x10^5$ M^{-1} cm^{-1}. [67] These two bands come very close to each other and are usually observed as just a single peak, known as the Soret band. The free base porphyrin has a further four absorbance peaks called the Q-bands that are present at a higher wavelength (480-600 nm) but, have a lower molar extinction coefficient in the order of 1 $x10^4$ M^{-1} cm^{-1} (Figure 1.20). [67]

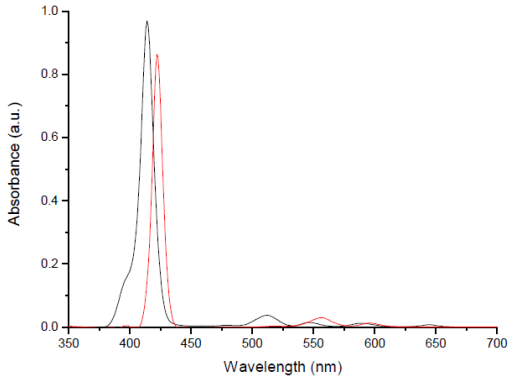


Figure 1.20 UV-Vis absorbance spectra of freebase (black) and zinc metallated (red) porphyrin

Gouterman *et al*. explained the origins of the absorbance spectra using his four orbital model.[68-70] For simplicity, a symmetrical porphyrin, such as porphine or tetraphenylporphyrin, will be considered in this discussion. The free base form of this type of porphyrin has D_{2h} symmetry due to the presence of the two hydrogen atoms on the pyrrole nitrogens. This gives the porphyrin two orthogonal dipoles, B_x and B_y that are shown in Figure 1.21. The two ground states have quasi allowed $\pi \rightarrow \pi^*$ transitions to both the first and second excited states, producing four Q-bands and two B bands respectively. When the porphyrin is metallated, the symmetry of the porphyrin is changed from D_{2h} to D_{4h} which causes B_x ad B_y transitions to become degenerate. The outcome of this is that there is now only one B band absorbance and two Q band absorbances (Figure 1.20). The addition of other substituents will have large effects on the electronic properties of the porphyrin, but further discussion is beyond the scope of this introduction.

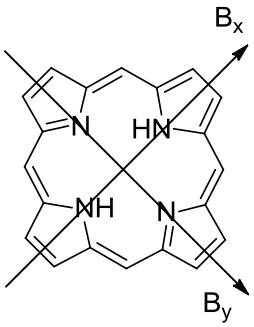


Figure 1.21 B_x and B_y transitions of porphine

1.4.2 Porphyrins in nature

Porphyrins are often found in biological systems with haem[71] and chlorophyll[72] being two well known examples (Figure 1.22). Haem is a vital component of the protein haemoglobin which is used to transport oxygen molecules to all cells of the body. Haem contains iron (II) in the central cavity of the porphyrin (ferrous haem) that can bind an oxygen molecule when the concentration of oxygen in the blood is high and carbon dioxide concentration is low. The bound oxygen is then transported through the bloodstream via an iron (III) metallated porphyrin (ferric haem) before being released where the carbon dioxide concentration in the blood is high and the oxygen concentration is low. This reduces the iron back to iron (II) where it can bind another oxygen molecule and repeat the process. There are four molecules of haem present in each haemoglobin protein. When oxygen is bound to one of the haem molecules there is a change in the configuration of the protein which makes the subsequent binding of oxygen to the other haem molecules more facile. Cytochrome c is another protein that contains haem. This protein is a component of the respiratory

chain in mitochondria where it transports electrons between complexes III and IV. This is achieved by the reduction and oxidation of the central iron ion.

Chlorophyll is a magnesium metallated reduced porphyrin where one of the pyrrole rings has been reduced to give a macrocycle known as a chlorin. Chlorophyll acts as a photoreceptor that is present in photosystems I and II of plants.[73, 74] They strongly absorb visible light from the sun. The molar extinction coefficient is higher than 1 $x10^5$ M^{-1} cm^{-1}, which is extremely high for an organic compound. When chlorophyll absorbs light, it is excited from its ground energy level state. If an electron acceptor is present and in close proximity, an electron is transferred from chlorophyll in a process called photoinduced charge separation. Several electron transfer steps take place which lead to the reduction of NADP (nicotinamide adenine dinucleotide phosphate), which can then be used to reduce atmospheric carbon dioxide to create sugars.[75]

Vitamin B12 (Figure 1.22c) is another biologically important molecular that contains a porphyrin derivative, in this case a corrin ring. A corrin is similar to the porphyrin in that it has four pyrrole-like subunits that are connected at the α-carbon positions to form a ring. However, one of the linkers lacks a carbon atom that the others possess and the ring is not aromatic like a porphyrin. The central cavity of vitamin B12 coordinates a cobalt cation. Vitamin B12 plays an important role in the metabolism of fatty acids.

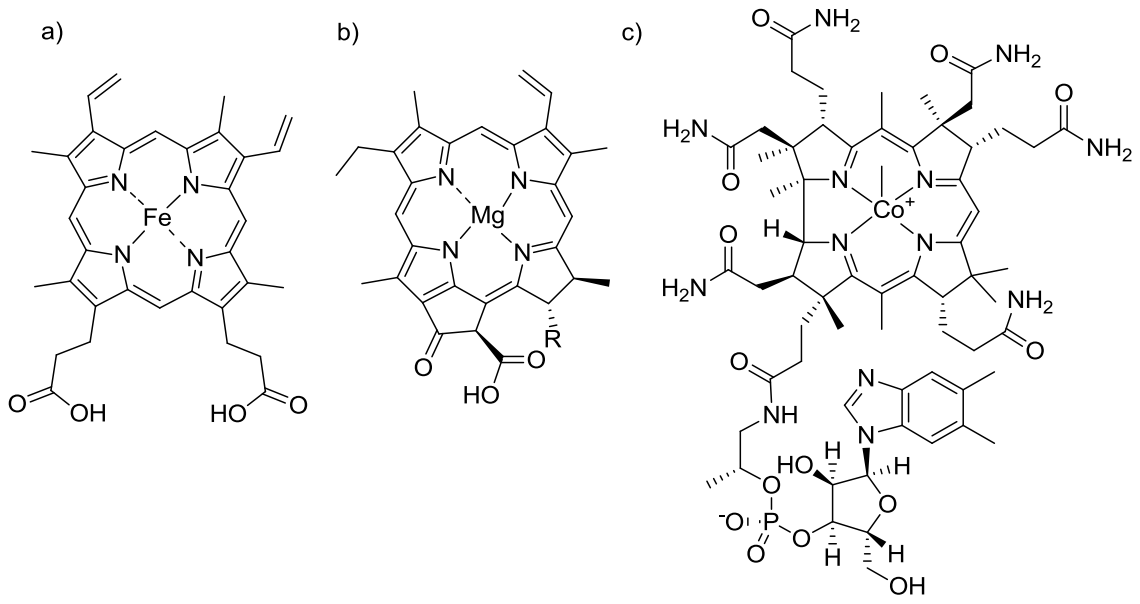


Figure 1.22 Examples of biological porphyrins: a) Haem B b) Chlorophyll A c) Vitamin B12

1.5 Supramolecular porphyrin assemblies

Porphyrins have long been incorporated in supramolecular assemblies due to their interesting physical properties. These assemblies are often made to mimic photosynthetic systems,[76, 77] harvest light,[78] transfer energy or electrons,[79] act as sensors[80] or be used in photodynamic therapy.[81] The porphyrins can either be covalently bound to a scaffold or can interact non-covalently with a preformed structure (Figure 1.23). Porphyrins have been covalently attached successfully to many different supports such as oligonucleotides,[82] peptides,[83, 84] isocyanide polymers,[85] nanoparticles[79] and cyclodextrins.[81] Non-covalent porphyrin assemblies can also be created in a huge variety of ways with examples using carbon nanotubes,[78] graphene,[86] backbone substituted oligonucleotides,[87] peptide nanotubes[88] or templated on viruses.[89]

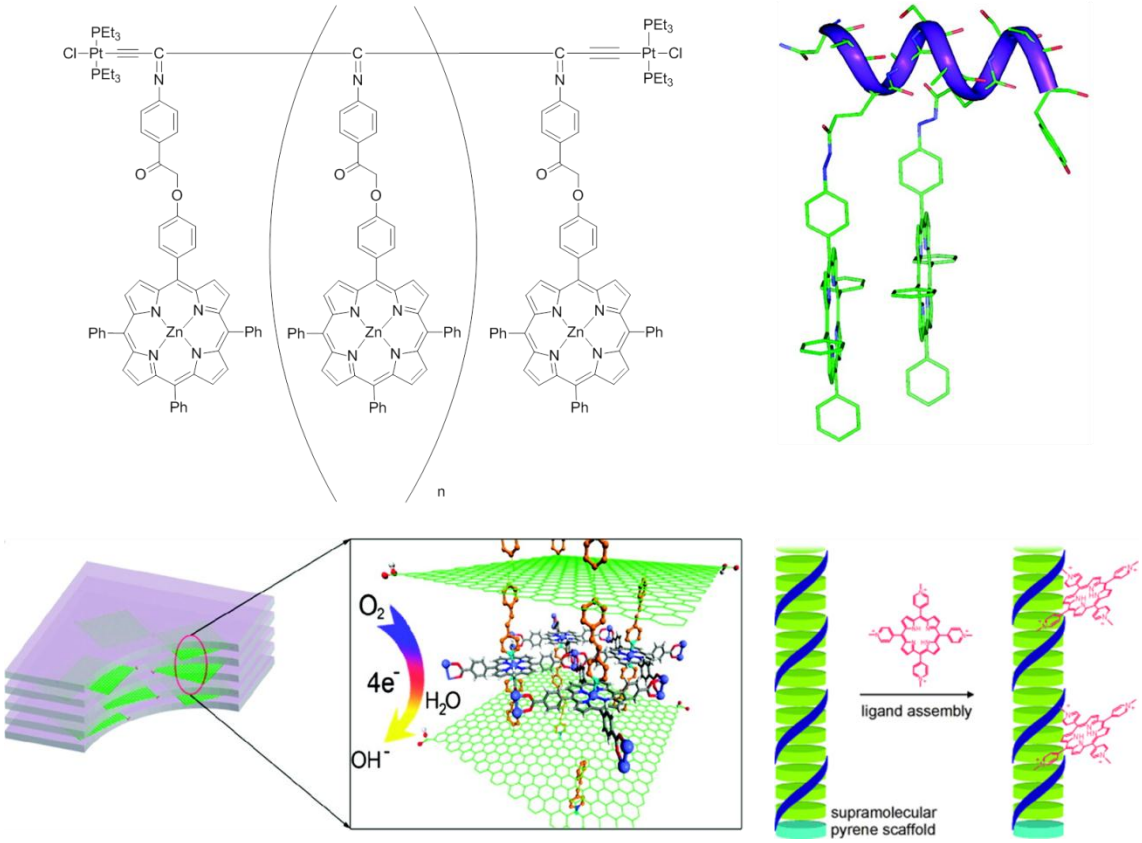


Figure 1.23 Examples of supramolecular porphyrin assemblies[83, 85-87]

1.6 Porphyrin-modified nucleosides, nucleotides and oligomers

Due to the ability of DNA to self-assemble into more complex structures and the interesting electronic properties of porphyrins there have been many examples of where porphyrins have been used to modify nucleosides that can then be further incorporated into oligonucleotides. Early examples started by covalently attaching a porphyrin to a nucleoside. Examples of this include the porphyrin-modified 2'deoxyuridine[90] and a porphyrin-modified guanosine that could hydrogen bond to a quinone modified cytosine in solution to form a redox couple.[91]

The first examples of incorporating porphyrins into oligonucleotides came by attaching the porphyrin to the 5'-terminal of the oligonucleotide after synthesis. This could be achieved by an amide coupling reaction between the porphyrin and oligonucleotide.[92, 93] Although this method was successful in synthesising the porphyrin oligonucleotide conjugates, it was limited in the fact only one porphyrin modification could be added in one specific position. To navigate this problem porphyrin-modified nucleosides that could be then phosphitylated and incorporated into standard solid phase oligonucleotide synthesis were developed. This was first achieved by using a Sonogashira cross-coupling reaction to attach a porphyrin to the C5 position of thymidine.[90, 94] An alternative to this approach was to replace the nucleobase completely with a porphyrin, as demonstrated by Kool *et al*.[95]

A third way of making porphyrin oligonucleotide conjugates is to attach the porphyrin to the nucleoside at the 5'- or 3'-hydroxyl positions. This method has been successfully used by Berova *et al*.[96, 97] When porphyrins are attached to DNA they can be used as a chiral marker for monitoring the structure of DNA and any changes that occur to it. When the porphyrin is bound to the DNA scaffold, the chiral information of the double helix is transferred to the porphyrin. As the porphyrin strongly absorbs light at 420 nm it has been found by the Berova research group that by using circular dichroism spectroscopy, structural information about the DNA can be obtained by studying the porphyrin Soret band region. One example where this has been shown to be successful is monitoring a change in structure from B-DNA to Z-DNA with an increasing salt concentration.[98, 99] The Berova group have also shown that by functionalising the porphyrin modification, the spectroscopic properties can be changed to their advantage.[100, 101]

The Stulz research group were the first to synthesise porphyrin oligonucleotide conjugates with the porphyrin attached to the nucleobase.[102] The advantage of using this method to incorporate the porphyrins into DNA is that you are not limited to having the porphyrin at the terminal end of the DNA and that multiple porphyrins can

be incorporated into a single oligonucleotide. Using this method, tetraphenyl porphyrins[49] and diphenyl porphyrins[103] with different linkers to the nucleobase[82] have been successfully incorporated into oligonucleotides. When a single porphyrin is incorporated in an oligonucleotide, its spectroscopic properties are generally unaffected with a single broad absorbance at 420 nm observable. However, once more porphyrins are incorporated in close proximity a large broadening of the Soret band is observed. This effect is indicative of exciton coupling between the porphyrin moieties.[104] Exciton coupling occurs when the dipoles of two molecules within a short distance of each other interact. The exciton coupling observed is dependent on the distance and the angle of the dipoles are at with respect to each other.[105]

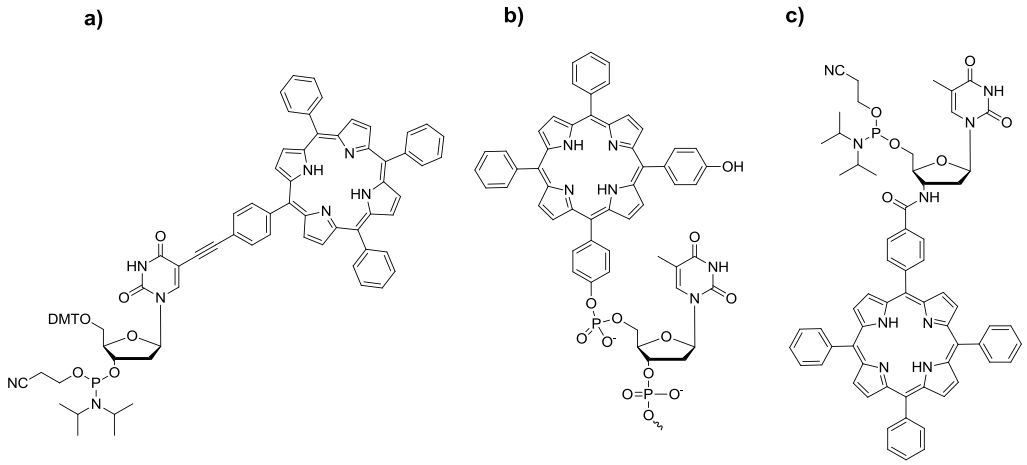


Figure 1.24 Examples of porphyrin-modified nucleosides[82, 97, 100]

Much of the research on porphyrin-modified oligonucleotides by the Stulz group has concerned the effect the porphyrin modification has on the structure and stability of the DNA. It has been observed that the porphyrin modification has a tendency to destabilise duplex DNA. This is especially the case for the diphenyl porphyrin used in the work by Nguyen,[103] where it is thought the large pendant hexyl chains that were attached to the porphyrin at the β positions to aid the solubility of the monomer are causing the destabilising effect. Large numbers of porphyrin modifications have been introduced into oligonucleotides with 11 consecutive porphyrins on a single strand being the largest reported to date (Figure 1.25).[49] Further investigations have shown that the positioning of the porphyrins in double-stranded DNA has large consequences on the overall stability of the duplex. By placing the porphyrin modifications on one of the strands it was observed there is a decrease in duplex stability. This was confirmed by a – 3 °C decrease per porphyrin modification in the melting temperature of the DNA duplex.[49] However, if the porphyrins are placed in alternate positions on both strands

of DNA in a zipper fashion (Figure 1.25) the stability of the duplex is found to increase.[103] Other research groups have since followed the lead of the Stulz group to further develop porphyrin-modified oligonucleotides.[106, 107]

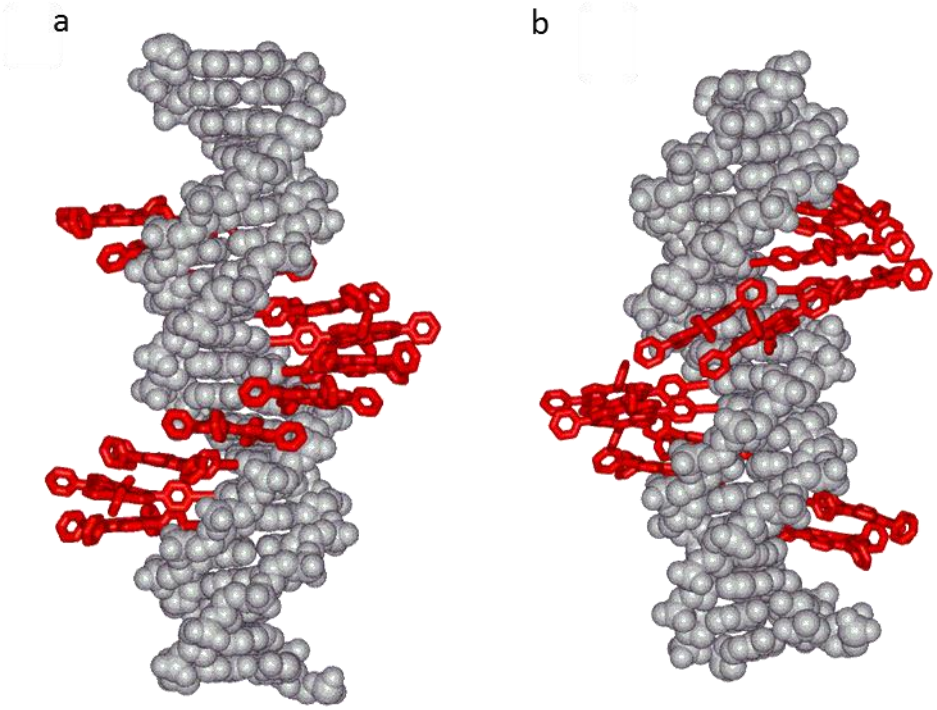


Figure 1.25 Molecular modelling of multiple porphyrin-modified duplexes a) 11 consecutive porphyrins on one strand[49] b) porphyrins in a "zipper" type arrangement[103]

2 Objectives

There were a number of objectives set out at the beginning of this project which have been developed as the project progressed. To begin with, it was proposed to synthesise the first G-quadruplex forming DNA sequences where porphyrins are covalently attached to a thymidine, either within the loop or at the periphery. The G-quadruplex forming sequence chosen for this project was based on the human telomeric sequence, $(TTAGGG)X_n$. The structural influence and G-quadruplex stability of these sequences were to be studied using spectroscopic methods and compared with their unmodified analogues.

As this project progressed it was clear that the synthesis of porphyrin-modified DNA needed reviewing. The synthesis of these modified oligonucleotides is a challenging process so it would be highly beneficial for future work if the methods of synthesis and purification could be improved. If this is successful, the analysis of porphyrin DNA would be possible using techniques such as NMR spectroscopy and X-ray crystallisation.

A number of related side projects arose during this time. The first of these was to synthesise a new type of porphyrin DNA modification that was based on LNA. The influence the LNA modification has on the DNA structure and stability was to be assessed. A second project concerned a new method of synthesising porphyrin-modified DNA. This was to be achieved by using polymerases to incorporate porphyrin-modified triphosphates into DNA. As the number of modified 2'-deoxynuceloside triphosphates (dNTPs) that have been synthesised and incorporated into DNA by primer extension and polymerase chain reaction increases, it would seem viable to attempt to synthesise porphyrin DNA using this route.

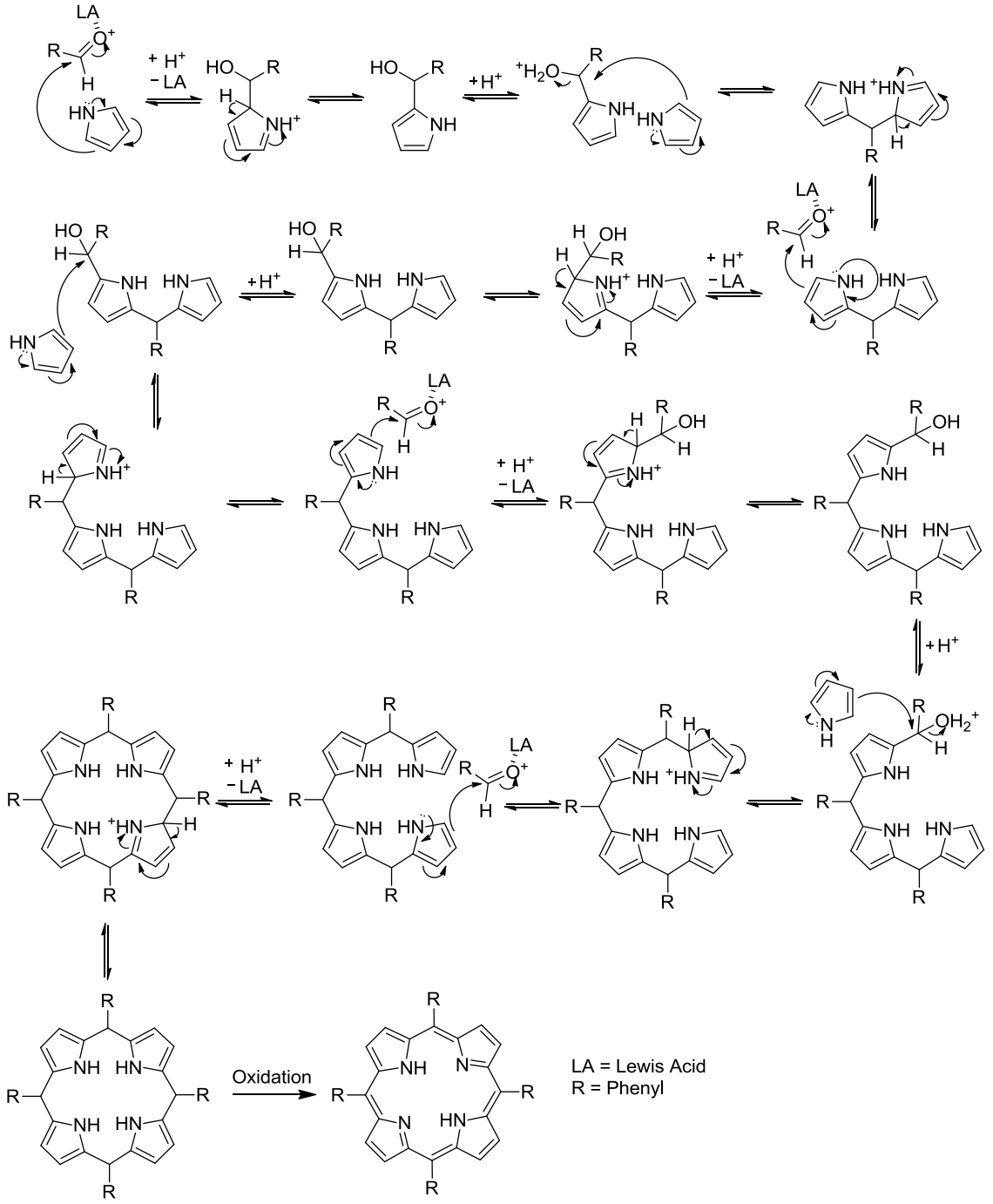
3 Porphyrin DNA synthesis

3.1 An introduction to porphyrin synthesis

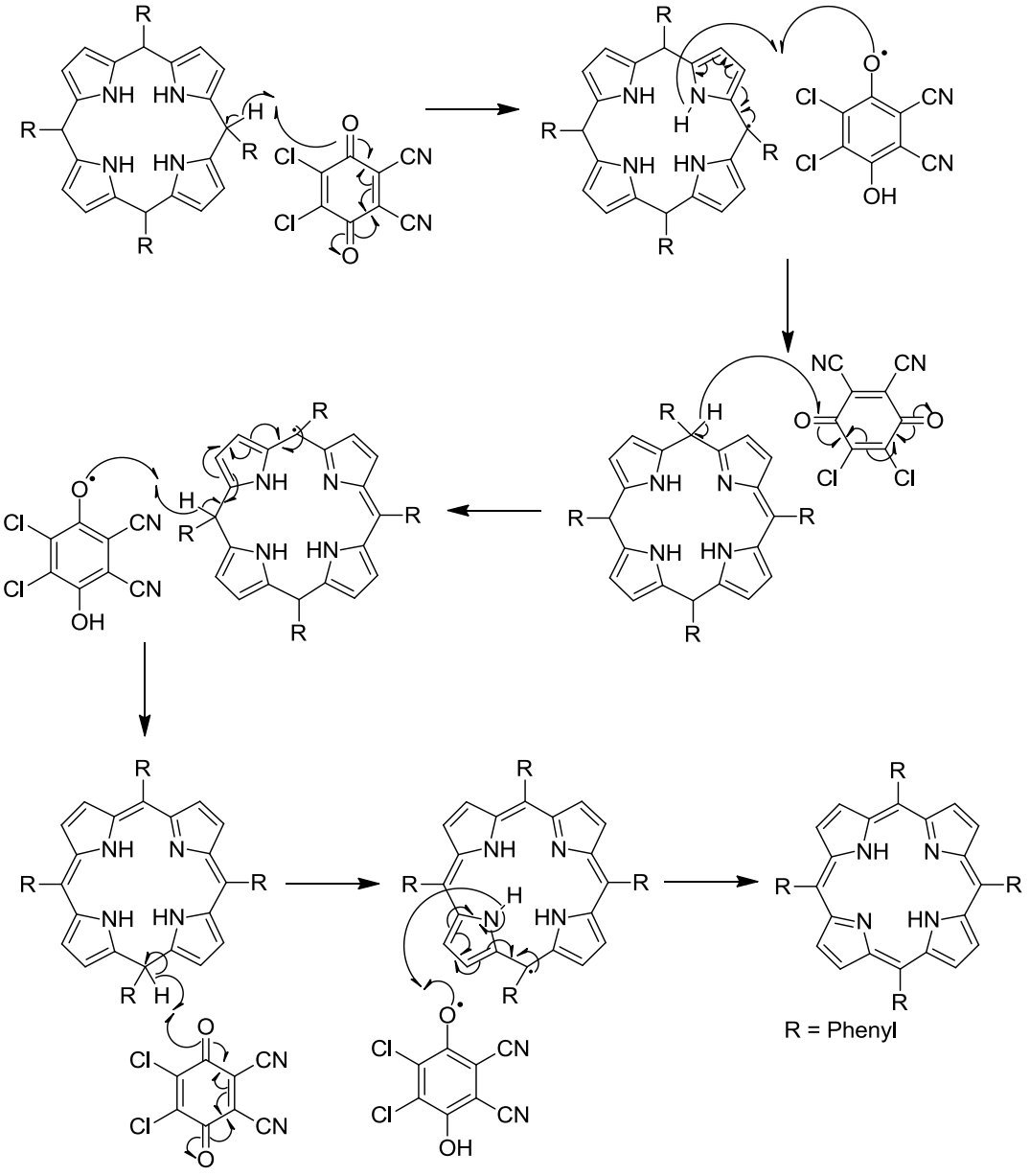
The synthesis of porphyrins has been widely studied since Fischer and Zeile first reported their synthesis of haem in 1929.[108] Rothemund further developed the synthesis in the 1930s by introducing a one pot synthesis of tetraphenylporphyrin from pyrrole and aldehydes.[109-111] The pyrrole and aldehyde were dissolved in pyridine before being heated to 150 °C for 24 h in a sealed bomb. Low yields and severe conditions meant that only a few substituted benzaldehydes could be converted to the corresponding porphyrin without degradation. The synthesis was further improved by Addler and Longo, who developed a simplified synthesis of *meso*-tetraphenylporphyrin.[112] This method refluxes benzaldehyde and pyrrole in propionic acid (141 °C) whilst open to air. Although the conditions for this method could be considered milder than the Rothemund synthesis, they are still too harsh for reactions with benzaldehydes bearing sensitive functional groups. Another disadvantage to this method is that oxidation of the porphyrinogen by molecular oxygen is not the most efficient or reliable method available.

Lindsey *et al.* [66] developed the one pot synthesis further by using a Lewis acid, such as boron trifluoroetherate, or trifluoroacetic acid and carrying out the reaction in an apolar solvent at room temperature (Scheme 3.1). After an hour of reacting, the porphyrinogen would be irreversibly oxidised to the porphyrin by the addition of 2,3,5,6-tetrachlorobenzoquinone (*p*-chloranil) or 2,3-dichloro-5,6-dicyanobenzoquinone (DDQ) (Scheme 3.2). Yields typically around 50 % were observed for this method and the mild conditions used meant that more sensitive benzaldehydes could be used.

The Lindsey method is ideal for synthesising symmetric porphyrins, but for the synthesis of asymmetric porphyrins a statistical mixture of products will be formed. However, by changing the ratios of the pyrrole and benzaldehydes added to the reaction mixture you can bias the formation of the desired porphyrin. In the case of the synthesis of a monosubstituted porphyrin, a 6:6:1 ratio of pyrrole : benzaldehyde-A : benzaldehyde-B will give two major porphyrin products. The first will be a tetrasubstituted benzaldehyde-A porphyrin and the second will be the desired mono benzaldehyde-B substituted porphyrin. Work by Stulz *et. al.*[49] have shown this method to be successful. However, the synthesis of higher substituted porphyrins using this method is not very successful and a different strategy must be employed.

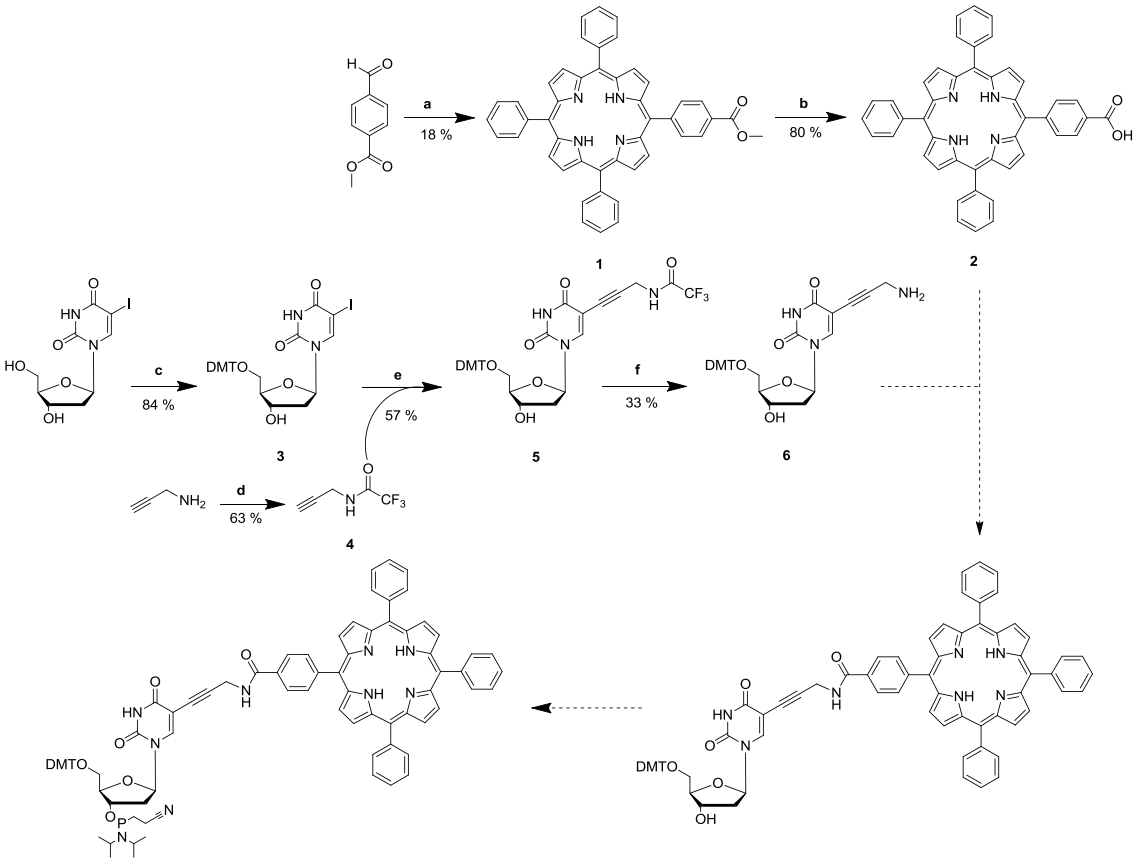


Scheme 3.1 One pot porphyrin synthesis mechanism



Scheme 3.2 Mechanism showing the oxidation of the porphyrinogen using DDQ

3.2 Synthesis of the amide linked porphyrin monomer



a - i) pyrrole, benzaldehyde, $BF_3.Et_2O$, $CHCl_3$ ii) DDQ, 18 % **b** - KOH, pyridine, H_2O, 40 °C, 80 %
c - 4-4'-dimethoxytrityl chloride, pyridine, 84 % **d** - ethyltrifluoroacetate, TEA, MeOH, 63 %
e - $Pd(PPh_3)_4$, CuI, TEA, DMF, 57 % **f** - $MeNH_2$, MeOH, H_2O, 33%

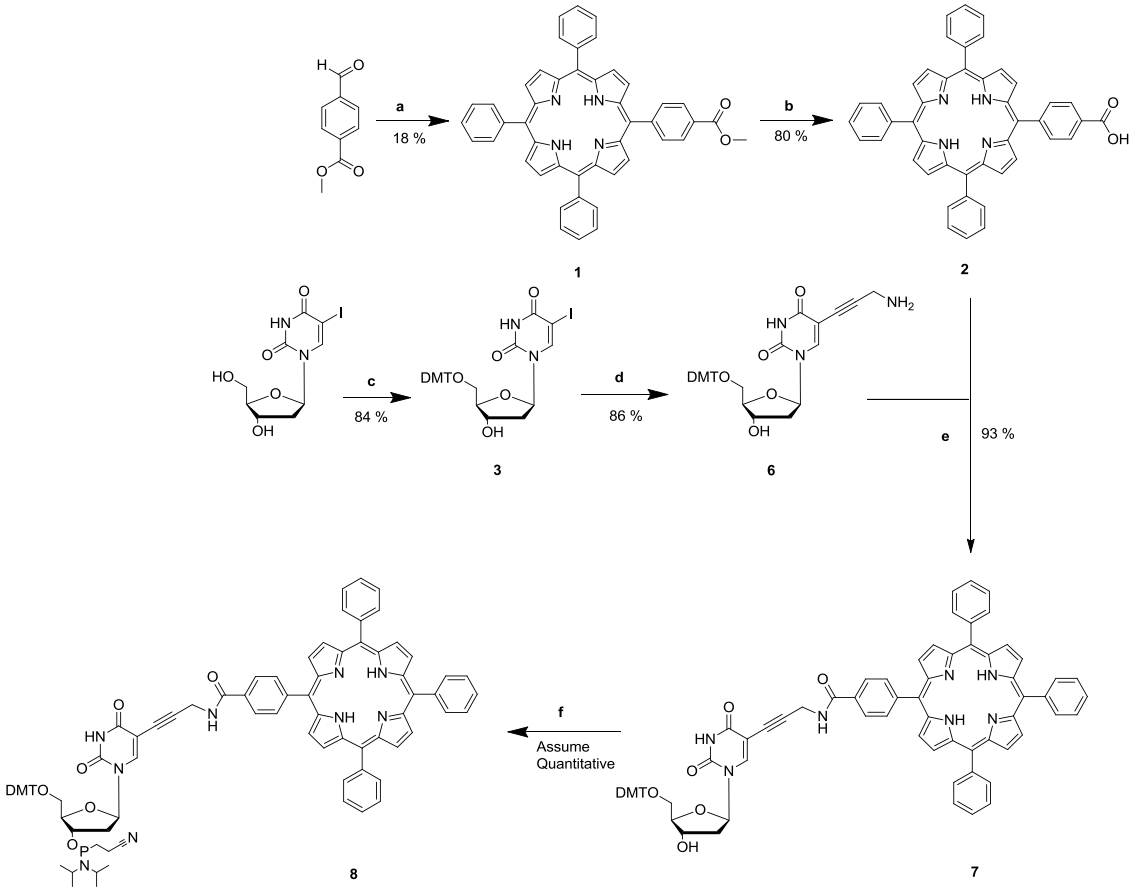
Scheme 3.3 Original proposed synthetic route of the porphyrin-modified thymine

The synthetic route begins with the synthesis of a tetraphenyl mono methyl ester porphyrin **1**. Benzaldehyde, pyrrole and methyl-*p*-formylbenzoate in the optimum ratio of 6:6:1 were reacted together to give the monosubstituted porphyrin.[82] Purification of this porphyrin was relatively straightforward once the polypyrrole side products had been removed by two rounds of column chromatography. Yields typically around 18 % were observed from this reaction. The methyl ester of the porphyrin was then hydrolysed to the carboxylic acid **2** by treating with 50 equivalents of potassium hydroxide. The reaction went to completion when left to stir at 40 ˚C for 24 hours. It was observed that if a lower amount of potassium hydroxide was added to the mixture, the reaction would not go to completion. Careful neutralisation of the reaction mixture with dilute hydrochloric acid is required once the reaction is complete as the deprotonated porphyrin is water soluble. Problems can occur if too much acid is added

and the protonated porphyrin species is formed. This is immediately noticeable as the protonated porphyrin is bright green in colour. This phenomenon is only observed for tetraphenyl porphyrins, where the protonation causes a change in conformation of the porphyrin macrocycle.[59] This new conformation allows the four phenyl rings at the *meso* positions to become conjugated with the porphyrin macrocycle. The extended conjugation is responsible for the observed colour change. When this occurs the porphyrin will be partially soluble in both aqueous and organic phases, which is problematic when extracting the porphyrin into an organic phase and performing aqueous washings. Once neutralised and the porphyrin completely dissolved in DCM, the product was thoroughly washed with water to remove the large excess of salt impurities and then purified by column chromatography.

5'-DMT-5-iodo-dU **3** was prepared by a standard S_N1 reaction between the primary 5'-hydroxyl group of 2'-deoxy-5-iodouridine and 4,4'dimethoxytrityl chloride (DMT-Cl) in dry pyridine. The DMT-Cl was added portionwise as there is the possibility it could also react with the less reactive 3'-hydroxyl group. By keeping the concentration of the DMT-Cl low, one can bias the addition of DMT towards the less hindered 5'-hydroxyl position. Careful addition prevents the formation of any bis DMT substituted product. However, if this occurs separation can easily be achieved by column chromatography on silica. It is important to neutralise the silica with triethylamine before addition of the crude product, as the slightly acidic nature of the silica can cleave the DMT protecting group. The neutralisation of the silica is carried out prior to any column carried out on DMT protected compounds.

In the initial synthetic route (Scheme 3.1Scheme 3.3), the plan was to protect propargylamine with trifluoroacetic acid **4** and then couple this compound with DMT-iodouridine using a Sonogashira cross-coupling reaction to yield **5**. Once this was complete, the TFA protecting group would be removed under basic conditions. The Sonogashira cross-coupling reaction was found to work well[57] but removal of the TFA protecting group after in the next step was not so simple. Ammonia, 30 % aqueous ammonia, and methylamine were tried to remove the TFA protecting group, but with little success. This is contrary to what has been reported before.[113] It was found that carrying out the Sonogashira cross-coupling reaction between the unprotected propargylamine and DMT-iodouridine was a more viable route for **6** with a yield of 86 % observed. This reaction was successful and removes two steps from the total synthetic route (Scheme 3.4).



a - i) pyrrole, benzaldehyde, $BF_3.Et_2O$, $CHCl_3$ ii) DDQ, 18 % **b** - KOH, pyridine, H_2O, 40 °C, 80 %
c - 4-4'-dimethoxytrityl chloride, pyridine, 84 % **d** - $Pd(PPh_3)_4$, CuI, TEA, DMF, 86 %
e - EDC, HOBt, DMAP, DCM, 93 % **f** - CEP-Cl, DIPEA, DCM, Assume quantitative

Scheme 3.4 Modified synthetic route to the porphyrin-modified linker

The next step was to find an efficient way of coupling porphyrin acid **2** and DMT-iodouridine propargylamine **6** to give **7**. There are many possible methods in achieving this so I worked alongside Ashley Brewer of the Stulz research group in trying to find the ideal conditions for this coupling. Ashley explored the possibility of forming an acid chloride *in situ* and then reacting with the amine. This was first achieved by using oxalyl chloride. The acid chloride solution was then added dropwise to a solution of 5'-DMT-5-propargylamino-dU **6** to give **7** in 30 % yield. The acid chloride was then synthesised *in situ* using cyanuric chloride but it was found that various cyanuryl esters were formed and not converted to the acid chloride. Despite these issues the reaction of the acid chloride of **2** with 5'-DMT-propargylamino-dU **6** proceeded well, giving an improved overall yield of 51 %.

The phosphonium salts, PyBOP and PyBrOP were investigated next as alternative peptide coupling reagents for this reaction. No product formation was observed when PyBOP was used so we moved onto the more reactive PyBrOP where the reaction was found to proceed well at room temperature. However, many porphyrin side products

were formed. One major product that was observed had a mass of +53 of the starting porphyrin acid **2**. It was found that this side product was the porphyrin pyrrolidide (Figure 3.1a), which is a common side product when using PyBrOP if the coupling reaction is slow.[114] Despite this, the desired product **7** was separated by column chromatography, but the reaction was found to give a low yield of 37 %.

Next, after finding a report showing a similar peptide coupling reaction, [100] it was decided to use EDC and DMAP as peptide coupling reagents. This reaction formed the desired product, but again many porphyrin side products were formed. One major side product was the tautomer of the EDC active ester (Figure 3.1b), a common, stable side product when using carbodiimides in peptide coupling reactions. The product was obtained by column chromatography but with a poor yield of 30 %. The next logical step was to repeat the coupling with EDC and add something that will form an active ester that is less likely to tautomerise to the urea side product. HOBt is often used to solve this problem. A similar coupling with a porphyrin carboxylic acid was found where HOBt was used with EDC.[115] Using these conditions the preparation of **7** was found to be extremely successful with a yield of 93 % observed after a facile purification process. Once the porphyrin amide monomer **7** has been synthesised the next step was to phosphorylate the 3'-hydroxyl group of the nucleoside.

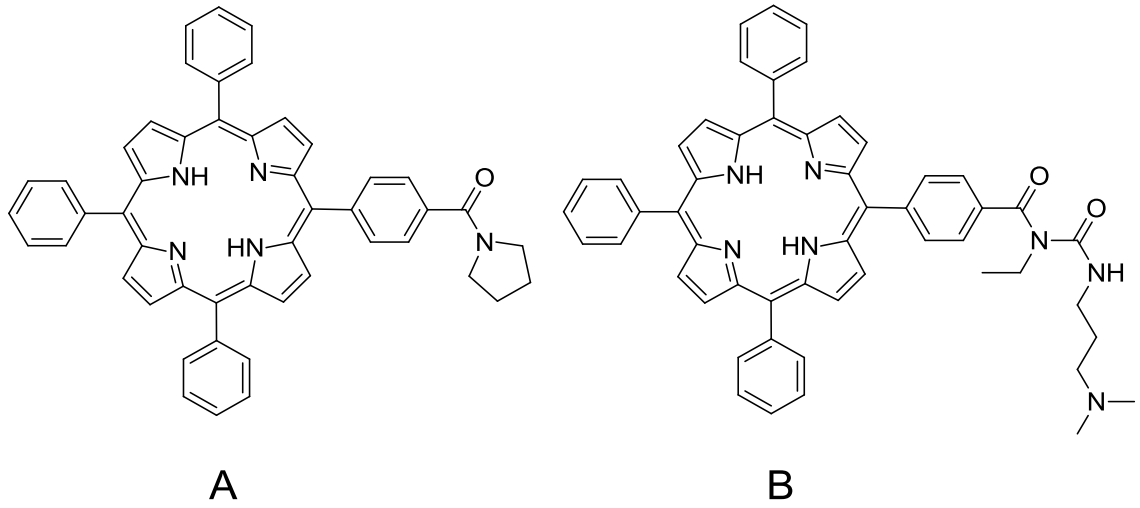


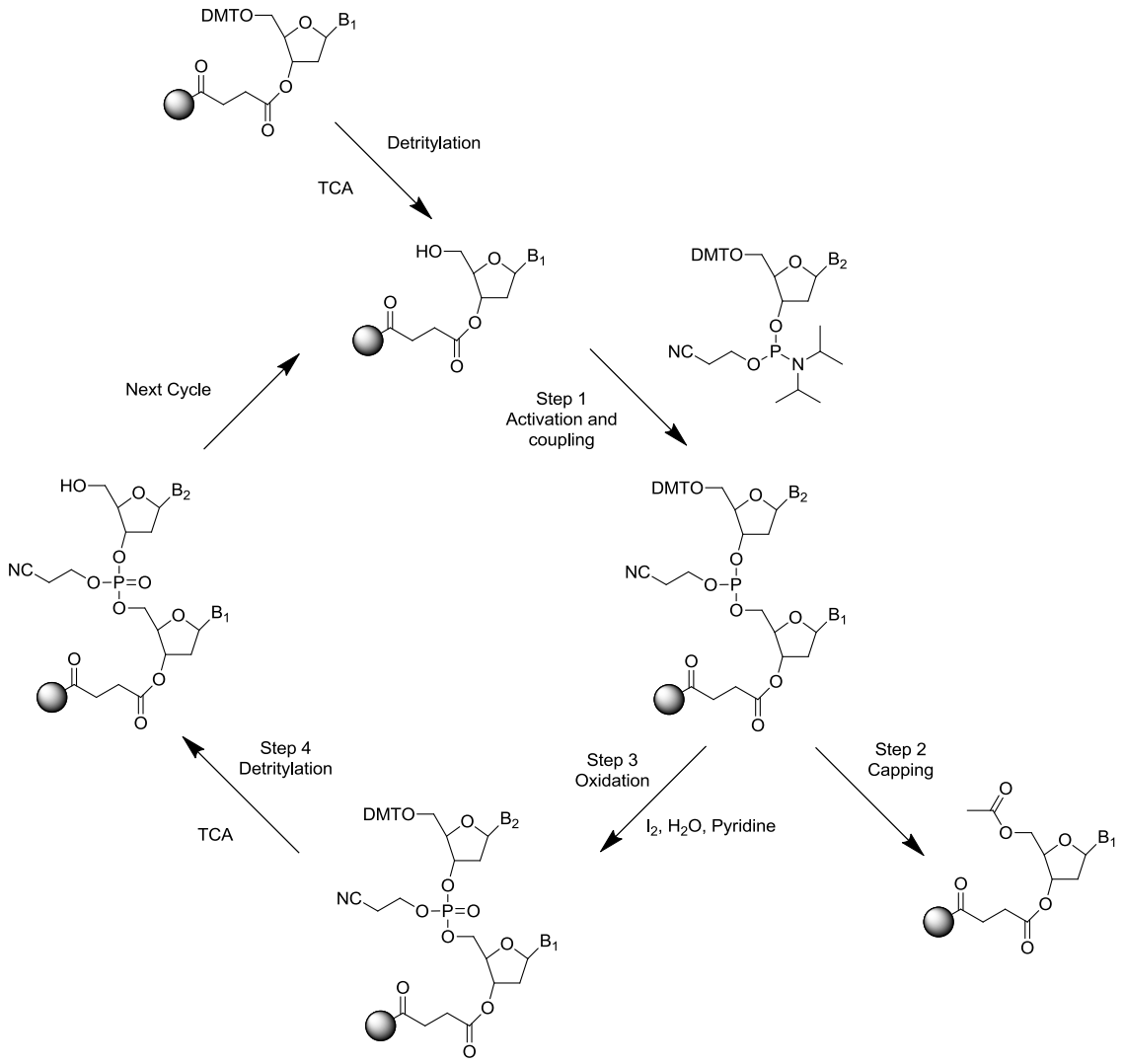
Figure 3.1 Major side products from **A** amide coupling using PyBrOP and **B** amide coupling using EDC without HOBt

The phosphorylation of **7** is a reasonably facile synthesis, carried out with CEP-Cl and DIPEA in dry DCM under an inert atmosphere. The reaction is usually complete within 2 hours. Purification of the porphyrin phosphoramidite **8** can be achieved using column chromatography under an inert gas with a degassed aprotic eluent. The other alternative which we employ is to precipitate the porphyrin phosphoramidite with

hexane and wash the excess CEP-Cl and DIPEA away with more hexane. Once dried, **8** must be immediately used for DNA synthesis. This is due to the instability of this compound. The phosphorus (III) will readily be oxidised to the more thermodynamically stable phosphorus (V) species. Oxidation of the phosphorus can occur in atmospheric conditions with molecular oxygen. The attached porphyrin will increase the rate of oxidation due to its ability to convert molecular triplet oxygen to the more reactive singlet oxygen.[116] For this reason the porphyrin phosphoramidite is only stable for a couple of hours. This makes the synthesis of porphyrin DNA and characterisation of the porphyrin phosphoramidite difficult.

3.3 Solid phase DNA synthesis

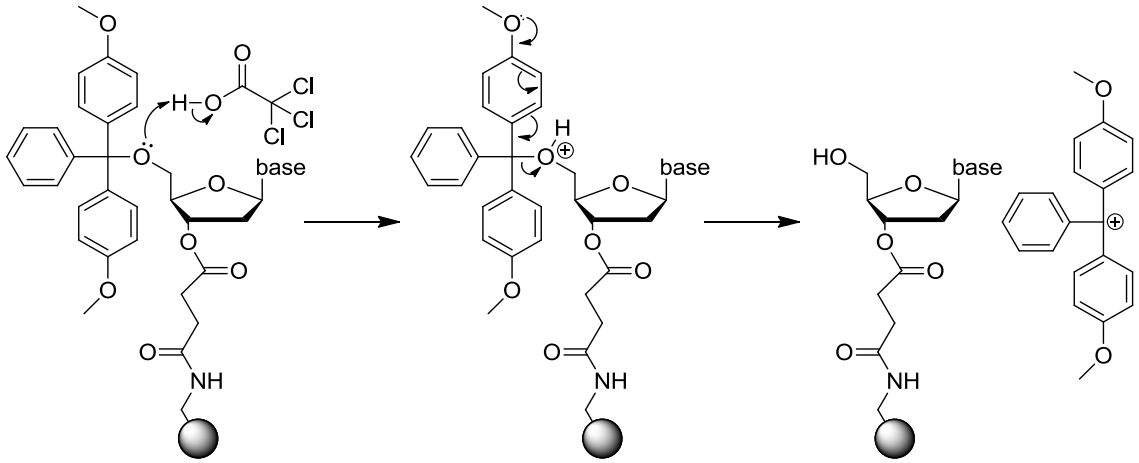
Solid phase synthesis was originally developed in the 1960s by Bruce Merrifield, which led to him winning a Nobel Prize in 1984. This technology was adapted to provide an efficient way of synthesising large quantities of oligonucleotides in a relatively short time frame. The solid phase is typically a glass bead with pores of a known size (controlled pore glass or CPG). The size of the pore is important when deciding the length of DNA to be synthesised, usually a 500Å pore size is suitable for small oligomers (<30 bases). The bead is loaded with the first base of the oligomer. The linker that attaches the first base to the solid support must be stable during solid phase synthesis but be labile in basic conditions so that the finished oligomer can be cleaved from the bead. As opposed to enzymes such as DNA polymerase which synthesise DNA in a 5' →3' direction, solid phase synthesis goes in a 3' →5' direction. Solid phase DNA synthesis is carried out using an automated DNA synthesiser and all reagents are stored in an inert atmosphere as the phosphoramidite used are especially vulnerable to oxidation. One nucleobase is added to the growing oligonucleotide during each cycle (Scheme 3.5). There are four steps that occur during each cycle: activation and coupling, capping, oxidation and detritylation. The cycle is repeated until the desired number of nucleobases is added to the oligonucleotide.



Scheme 3.5 Solid phase DNA synthesis cycle

3.3.1 Detritylation of the support bound 3'-nuceloside

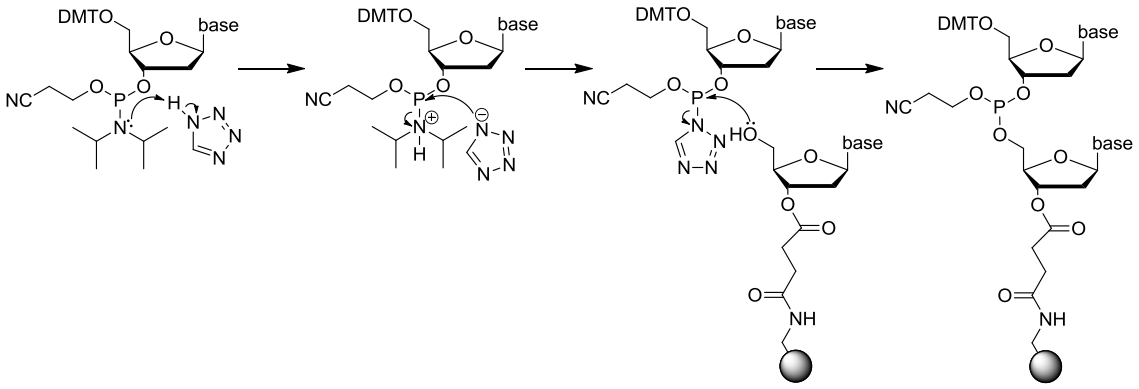
Before DNA synthesis can proceed, the 5'-DMT protecting group on the support bound nucleoside must be cleaved (Scheme 3.6). This is done by introducing 3 % trichloroacetic acid (TCA) in DCM to the nucleoside. The DMT protecting group is labile under acidic conditions and cleavage needs to be carried out before any nucleosides can be added to the 5'-hydroxyl position of the starting nucleotide.



Scheme 3.6 Detritylation of the support bound 3'-nucleoside

3.3.2 Activation and coupling – Step 1

The first step of the oligonucleotide synthesis cycle is to activate the di-*iso*-propylamino group of the 3'-phosphoramidite (Scheme 3.7). The tertiary amine is first protonated and then substituted with the tetrazolium anion. The leaving group of this intermediate is displaced through a nucleophilic attack on the neighbouring phosphorus atom from the free 5'-hydroxyl of the growing oligonucleotide, and a new phosphorus-oxygen bond is formed.

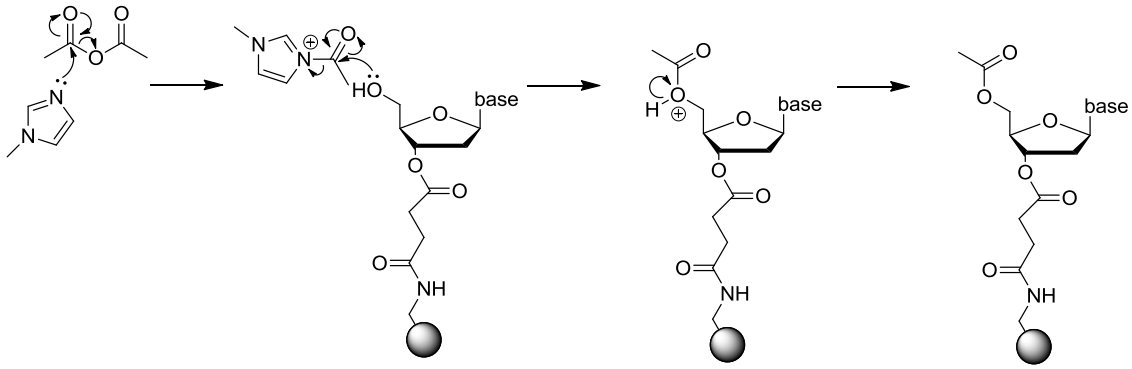


Scheme 3.7 Mechanism of the activation and coupling step

3.3.3 Capping – Step 2

Yields greater than 99 % are usually observed for the coupling step, but despite this a capping step is needed to prevent the small quantity of uncoupled oligonucleotide from taking part in further couplings. If they were left uncoupled,

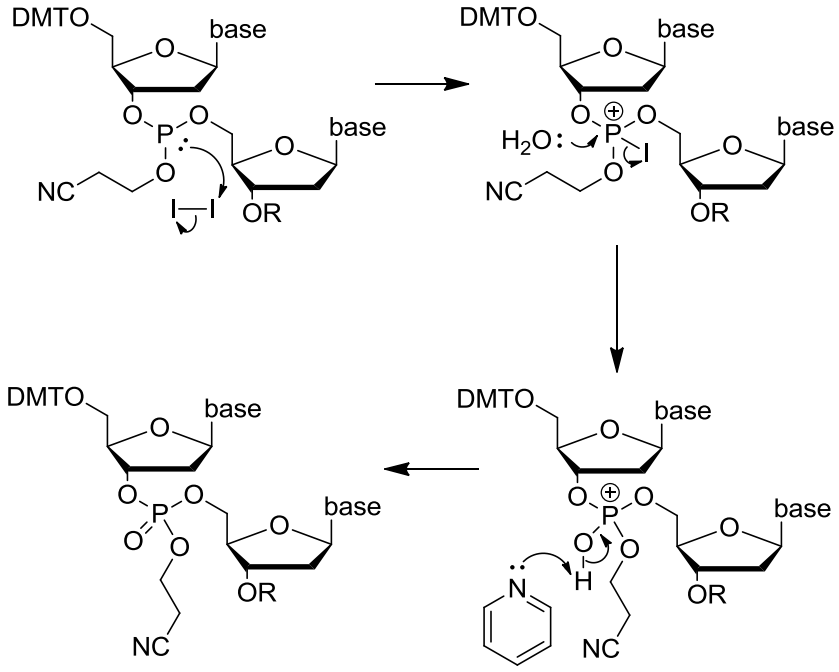
truncated oligonucleotide side products can accumulate after each successive cycle and the final product would be a complex mixture of oligonucleotides that would be difficult to purify. Capping is achieved by blocking the unreacted 5'-hydroxyl groups of any unreacted oligonucleotides (Scheme 3.8). The synthesiser introduces two separate solutions of acetic anhydride and *N*-methylimidazole (NMI) that are mixed together prior to delivery to the synthesis column. The NMI forms the activated *N*-acetyl-*N'*-methylimidazonium species with the acetic anhydride, efficient acetylation with the 5'-hydroxyl then occurs.



Scheme 3.8 Mechanism of the capping step

3.3.4 Oxidation – Step 3

The phosphite-triester (P(III)) formed in the coupling step is unstable to acid and must be converted to a stable (P(V)) species prior to the next TCA detritylation step (Scheme 3.9). The oxidation is carried out with a mixture of iodine, water and pyridine. A phosphorus iodine bond is first formed before the iodine is displaced by water. Deprotonation of the positively charged intermediate by pyridine leads to the generation of a new phosphorus oxygen double bond and the phosphorus (V) species.



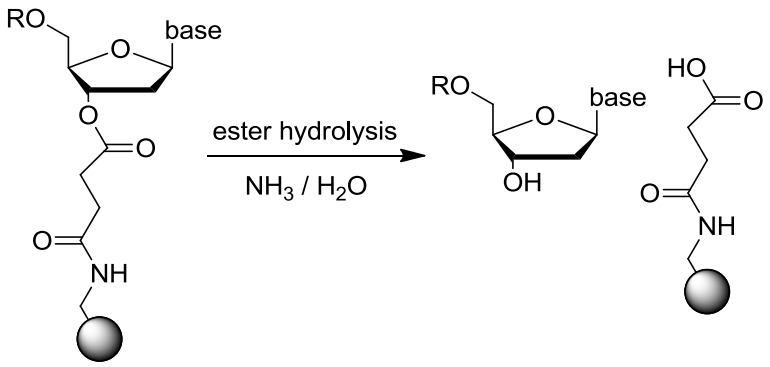
Scheme 3.9 Mechanism of the oxidation step

3.3.5 Detritylation – Step 4

The final step of the oligonucleotide synthesis cycle is to remove the DMT group that protects the 5'-hydroxyl of the recently coupled nucleoside so it is ready to participate in the next cycle. The method used for this is the same as the detritylation step for the first base attached to the solid support. During the detritylation step a bright orange colour is observed from the cleaved DMT carbocation, which absorbs in the visible region of the electromagnetic spectrum at 495 nm. The intensity of the absorbance can be used as an indicator for the efficiency of the coupling. The DNA synthesiser is able to measure this absorbance to record the trityl yield and provide real time analysis of each coupling. After this step, the cycle is repeated for the next base to be added to the growing oligonucleotide.

3.3.6 Cleavage from the solid support

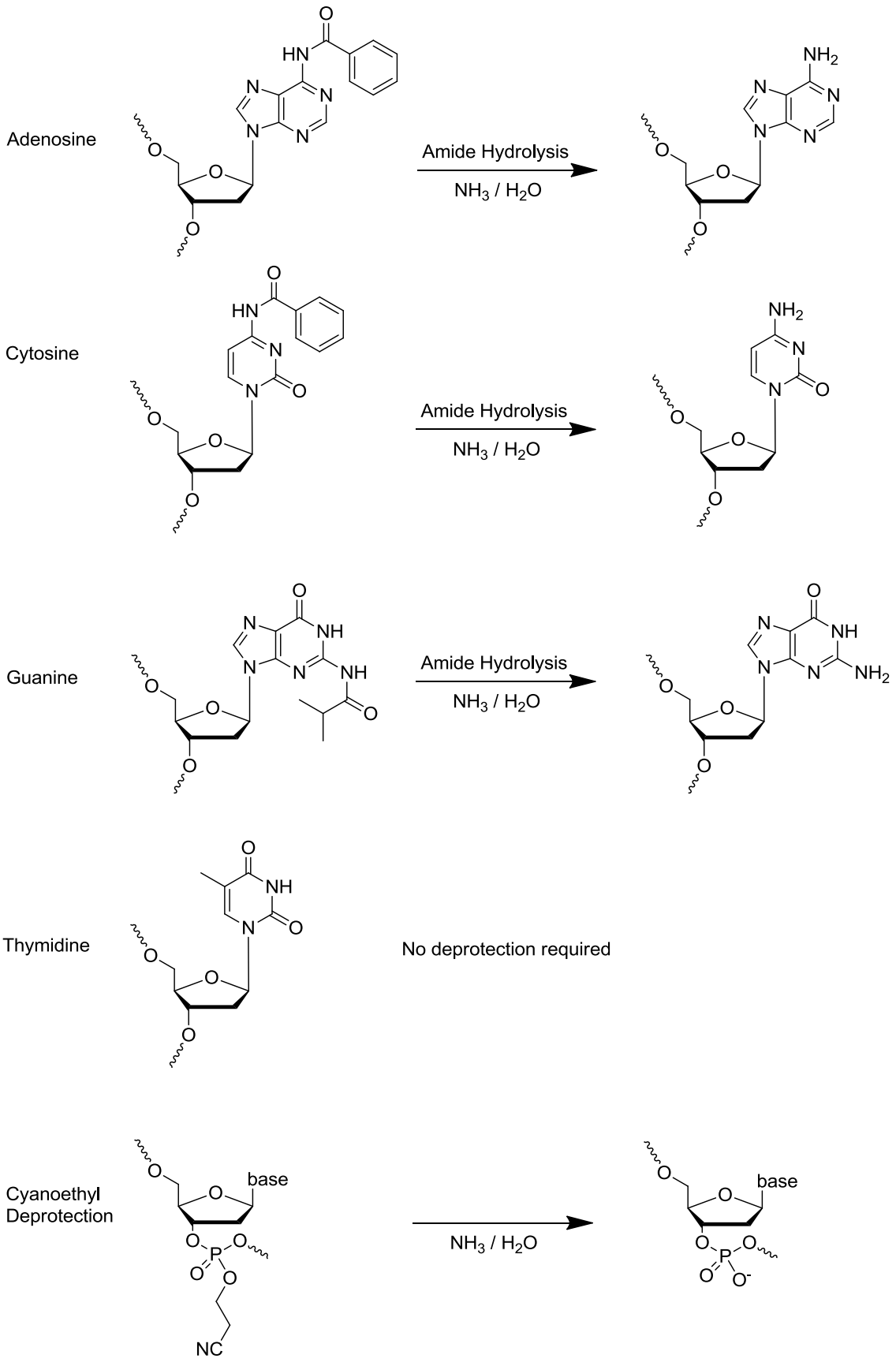
Once the desired oligonucleotide sequence is complete, cleavage from the solid support needs to take place. The oligonucleotide is attached to the solid support at the 3' end via a succinyl linker. This can be cleaved by adding concentrated ammonium hydroxide at room temperature for one hour (Scheme 3.10). This is done by taking the column off the synthesiser and attaching a syringe at either end of the column. Concentrated ammonium hydroxide solution is then passed between the two syringes via the oligonucleotide on the solid support.



Scheme 3.10 Cleavage of an oligonucleotide from its solid support

3.3.7 Oligonucleotide deprotection

When the oligonucleotide is dissolved in concentrated ammonia, deprotection of the various protecting groups on the nucleobases and the cyanoethyl protecting group on the phosphate backbone (Scheme 3.11) takes place. To ensure complete deprotection of the oligonucleotides the reaction is carried out at 55 °C for 4 hours.


Adenosine
Amide Hydrolysis
NH3 / H2O
Cytosine
Amide Hydrolysis
NH3 / H2O
Guanine
Amide Hydrolysis
NH3 / H2O
Thymidine
No deprotection required
Cyanoethyl
Deprotection
base
NH3 / H2O
base

Scheme 3.11 Deprotection of nucleobases used in solid phase DNA synthesis

3.4 Purification of DNA

DNA can be purified using many different methods. One of the most effective and routinely used methods is to use high pressure liquid chromatography (HPLC) on C_{18} reverse-phase modified silica. Different oligonucleotides will be separated on the column due to how they are retained on the column. There are many factors that can be altered to aid the separation of the oligonucleotides. These include the solvents used and the gradient employed when changing from one solvent to the other. HPLC was carried out using Waters XBridge OST C_{18} analytical and semi-prep columns. A minor drawback to this method is that resolution decreases with an increasing oligonucleotide length, making the separation of long oligonucleotides challenging. Modifications to the DNA can also introduce difficulties in purification. This is especially evident for porphyrin-modified DNA, as the porphyrin has a tendency to encourage the DNA to "streak" across the column decreasing any resolution of products. It must also be considered that some DNA sequences, such as G-rich DNA, can form secondary structures on the column, which will have different retention times. This makes their purification more challenging, although heating the column can help to reduce the amount of secondary structures formed.

Polyacrylamide denaturing gel electrophoresis can also be used to purify oligonucleotides. This technique separates DNA based on their electrical charge and their hydrodynamic properties, which are a function of their chain length. This method is more commonly used for larger oligonucleotides of around 100 bases or longer, as other methods are not as effective. Oligonucleotides of high purity are obtained using this method at the expense of the overall yield.

Another effective method of purifying oligonucleotides is to use affinity columns such as Poly-Pak or Glen-Pak. These purification techniques take advantage of the hydrophobic nature of the DMT protecting group. The stationary phase of these columns has a high affinity for the large aromatic DMT group. When the crude oligonucleotide is loaded onto the column the desired DMT protected product is retained on the column, whereas any failure products without the DMT group are washed away. The desired oligonucleotide can be retrieved from the column by cleaving its DMT group with the addition of 2 % TFA. The column is then neutralised and the product is eluted from it. This method of purification is favoured as it is relatively cheap and quick to do. A downside to this method is that it will fail to remove any n-1 failure sequences from the synthesis that also contain the final DMT protecting group.

Fluoro-Pak type columns take advantage of fluorous affinity to purify oligonucleotides.[117] The stationary phase of these columns are fluorinated so they have

a high affinity to other highly fluorinated compounds. Oligonucleotides are not highly fluorinated species (unless they have been modified to be) so during the last step of the DNA synthesis a nucleobase modified with a fluorous tag is added. The fluorous tag used is 4-(1*H*,1*H*,2*H*,2*H*-perfluorodecane)-4',4''-dimethoxytrityl (FDMT). This tag is similar to the traditionally used 5'-hydroxyl DMT protecting group with the addition of a "fluorous ponytail" (Figure 3.2). When the crude oligonucleotide bearing the FDMT group is loaded on the Fluoro-Pak column, it is retained on the column and the waste unfluorinated products are washed away. As with the Poly-Pak and Glen-Pak columns, the product can be cleaved from the stationary phase with the addition of 3 % TFA. This cleaves the FDMT group from the oligonucleotide allowing it to be neutralised and then washed off the column. Fluorous affinity purification is favoured over DMT-on purification as it is more effective at purifying larger oligonucleotides. For DMT-on purification, as the oligonucleotide length increases the relative contribution of the lipophilic DMT is diminished, thus reducing the affinity of the oligonucleotide to the adsorbent. This results in a lower overall yield and a reduced selectivity for the desired DMT-on oligonucleotide over the non-DMT failure sequences. However, there are similar issues to the DMT-on purification techniques in that n-1 failure sequences are not removed because they will also have the FDMT protecting group attached to them.

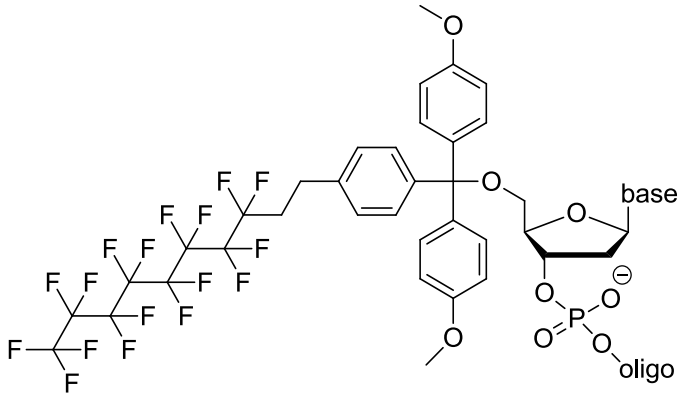


Figure 3.2 Fluorous dimethoxytrityl (FDMT) protected oligonucleotide

3.5 Synthesis of porphyrin-modified DNA

The porphyrin amide phosphoramidite **8** was used for oligonucleotide synthesis. A feature of most DNA synthesisers is that there are auxiliary lines present in which modified phosphoramidites can be added onto the synthesiser. Phosphoramidites are usually dissolved in acetonitrile when placed on the DNA synthesiser, however due to solubility issues with the porphyrin phosphoramidites a 1:1 mixture of acetonitrile : dichloromethane was utilised. Solubility is achieved, but the change in solvent used effects how much reagent is taken up by the synthesiser on each coupling. DCM could not be used on its own since the measurement of solvent volume by the DNA synthesiser is unreliable with high vapour pressure solvent. Even by using a 1:1 mixture of acetonitrile and DCM can have an effect on the measured volume of reagent delivered to the column. This effect needs to be taken into consideration when setting up coupling parameters so that the desired amount of phosphoramidite is delivered to the column.

Between 2.5 and 4 equivalents of the porphyrin phosphoramidites were required per coupling; less than 2 equivalents led to very poor coupling efficiencies, while greater than 4 equivalents showed no great improvement in coupling efficiency and as such is simply a waste of the phosphoramidites. A coupling time of 5 minutes was used for the porphyrin phosphoramidites. Previously it has been observed in the Stulz group that when varying the coupling times between 2 and 12 minutes, a coupling time of 5 minutes gave the most efficient yields. Addition of the porphyrin phosphoramidite to the oligonucleotide can be observed by the purple colour the column takes once coupling has occurred and also by monitoring of the trityl yield of the following detritylation step.

Once the synthesis is complete, the porphyrin-modified oligonucleotide is cleaved form the solid support and all protecting groups removed using the method as described previously. The striking purple colour of the ammonium hydroxide solution should be observed when the porphyrin DNA is cleaved from the solid support.

Both Glen-Pak and Fluoro-Pak columns were used in the purification of porphyrin-modified oligonucleotides. For purification using Glen-Pak columns, the DMT group must be removed prior to purification. The porphyrin modification is hydrophobic enough to be retained on the column. Any non-porphyrin bearing failure oligonucleotides can be washed away and the porphyrin-modified DNA collected afterwards. This preparation is recommended by Glen Research when purifying oligonucleotides modified with cyanine dyes. A similar method was employed by the Berova group to porphyrin-modified DNA followed by purification by RP-HPLC.[100] Analysis and any further purification afterwards is then carried out using RP-HPLC.

3.6 Porphyrin-modified Dickerson dodecamer DNA

A lot of work has been carried out within the Stulz group in synthesising and characterising duplex DNA modified with porphyrins.[49, 82, 118, 119] One method of characterisation that has not been achieved yet is X-ray diffraction of crystallised porphyrin DNA. Single crystal X-ray diffraction is the gold standard of characterisation and any results we obtain will be excellent to compare with the computation models of porphyrin DNA previously made within the group.[119] Crystallisation of DNA can be difficult and the addition of a large hydrophobic molecule certainly changes how the DNA may behave during crystallisation. It was decided that the porphyrin DNA synthesised would be based on a famous sequence of DNA known as the Dickerson Dodecamer. This self-complementary oligonucleotide with the sequence 5' – CGC GAA TTC GCG – 3' was the first example of B-DNA that had been crystallised and its structure determined by X-ray crystallography.[4, 12]

A large amount of pure porphyrin-modified oligonucleotide will be needed for the crystallisation experiments. Synthesising porphyrin DNA of high purity on this scale has not been required previously. So this becomes a good opportunity to assess the synthesis and purification techniques used to make porphyrin DNA.

The first thymidine from the 5'-position of the Dickerson dodecamer will be replaced with the flexible amide porphyrin monomer **dU^{2HTPP}** (Figure 3.3) to give:

5' – CGC GAA **P**TC GCG – 3' – **DD1**
3' – GCG CT**P** AAG CGC – 5' – **DD1**

Where **P** = **dU^{2HTPP}**

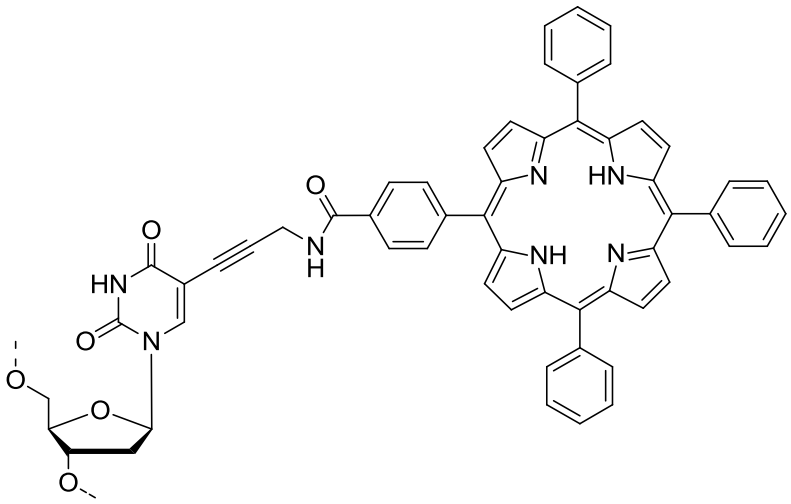


Figure 3.3 The flexible amide linked porphyrin (**dU^{2HTPP}**) used in the synthesis of the porphyrin-modified Dickerson dodecamer

This modification is very convenient because the self-complementary nature of the Dickerson dodecamer allows the two porphyrin modifications to be in close proximity to each other when in the duplex form. It has been observed that the porphyrins can interact with each other when in these positions on two complementary strands. However, the tetraphenyl modification of the two porphyrins may provide some steric hindrance, which would limit the amount of possible π-stacking interactions. However, to start with a single porphyrin modification was placed at the 5'-position of the Dickerson dodecamer:

5' – **P** CGC GAA TTC GCG – 3' – **DD2**
3' – GCG CTT AAG CGC **P** – 5' – **DD2**

It was decided to put a porphyrin modification in this position as this is the final base that is added to the oligonucleotide. This means that there is no concern about the coupling efficiency of any subsequent bases after the porphyrin modification and purification should be much easier. This sequence was synthesised on a 1 μmol scale using standard coupling conditions for the standard bases and an extended 5 minute coupling time for the porphyrin phosphoramidite. The oligonucleotide was synthesised DMT-off so the porphyrin at the 5'-position would act as the hydrophobic group needed for Glen-Pak purification. Deprotection and cleavage of the porphyrin-modified oligonucleotide was achieved by adding 35 % aqueous ammonia solution and heating to 55 °C for 4 hours. The porphyrin-modified oligonucleotide was then purified using Glen-Pak purification. The oligonucleotide would next need to be purified using HPLC, as the method used for Glen-Pak purification will not remove any failure sequences containing the porphyrin modification.

200 nmol of DNA was obtained after the Glen-Pak purification. Taking this into account, HPLC purification was carried out in two batches using a semi-prep reverse-phase column. A two solvent system was used comprising of 100 mM hexafluoroisopropanol and 8.6 mM triethylamine in water as the buffer and methanol as the second solvent (Table 3.1). Three problems were encountered during this purification. The first was due to the hydrophobic nature of the porphyrin, a large amount of streaking of products was observed on the column. There is not a lot that can be done about this problem, but by reducing the rate of change in solvent gradient the degree of streaking could be reduced.

The second problem after the purification was complete was that some of porphyrin-modified DNA was retained on the column. Not only does this mean product is being lost, but all material must be removed from the column prior to each purification attempt. Unfortunately, removal of the porphyrin DNA from the column proved to be more difficult than anticipated. The column was washed at least 5 times

before the majority of the porphyrin DNA had been removed. This meant that the purification using this method was wasteful in both terms of material and time. There were two ways of countering this problem. The first was to load a smaller amount of DNA at a time and the second was to use a smaller column so that less porphyrin DNA would be retained during the purification. The downside to this is that many runs must be carried out to purify all of the porphyrin DNA, which is very time consuming.

Time (min)	Buffer A (%)	Methanol (%)
0	100	0
25	60	40
60	20	80
70	0	100
75	0	100
85	100	0
90	100	0

Table 3.1 Solvent gradient used for HPLC purification of free base **DD2** Buffer A = 100 mM hexafluoroisopropanol and 8.6 mM triethylamine in water. Column temperature = 60 °C

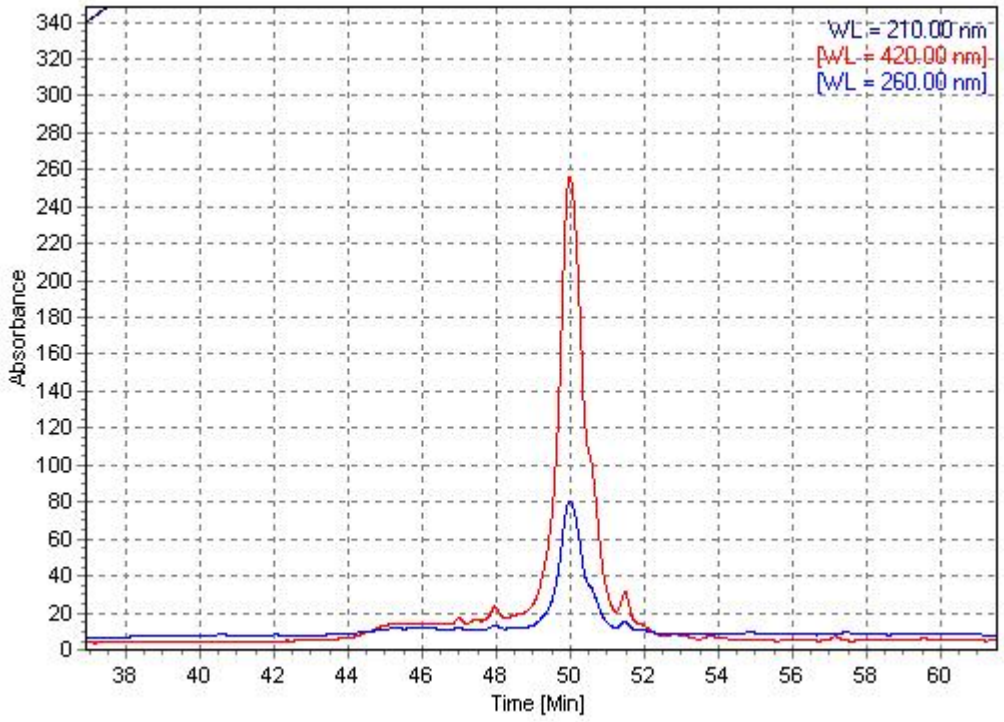


Figure 3.4 HPLC trace of **DD2Zn** – reinjection

The third issue observed was despite the fact separation of the product from failure sequences was not too difficult, there were a lot more porphyrin-modified DNA products present than expected. Upon close inspection of the UV data obtained from the HPLC trace, it was observed that about half of the porphyrins had become metallated. This affected the retention time (the metallated porphyrin products eluted before the free base porphyrin products) and doubled the amount of porphyrin-modified DNA products present. The source of the metallation is most likely coming from the HPLC solvents and is taking place during the run where the temperature is high enough to allow metallation of the porphyrin to occur. This issue has been observed previously by the Berova group, which they attributed to spontaneous metallation of the porphyrin during the cleavage/deprotection step with aqueous ammonia.[115] It is known that concentrated ammonia contains numerous trace metals and that the long exposure of the porphyrin at room temperature to them could be responsible for the metallation. However, these observations are inconsistent with the standard way of metallating porphyrins with an excess of the metal acetate under reflux. One would think the conditions during the cleavage/deprotection step are not ideal for efficient metallation to occur. However, using different batches of ammonia, it was observed that either zinc or copper could be introduced into the porphyrin. It is very unlikely that this is happening in our study as UV-Vis spectroscopic analysis of the porphyrin-modified DNA prior to HPLC purification showed the porphyrin present to be

completely in the free base form (Figure 3.5). Metallation during HPLC purification was also observed by the Berova group and the ratio of metallated to free base porphyrin (60:40 respectively) products was similar to what we had observed. The retention times for the metallated porphyrin DNA compared to the free base porphyrin DNA were found to be longer in this paper. This is inconsistent with what we have observed with the opposite being the case despite the same solvent system being used with a similar gradient.

Even though a pure product was obtained from this purification (Figure 3.4), we needed to find out what metal had been introduced into the porphyrin. UV-Vis spectroscopic analysis of the product (Figure 3.6) showed a completely metallated porphyrin was present with no sign of any free base porphyrin. The introduction of a metal into the porphyrin has an effect on both the Soret band and the Q band absorbances. There was a red shift in the Soret band of the porphyrin from 420 nm to 425 nm. In the free base porphyrin absorbance spectrum there are four Q bands at 519 nm, 556 nm, 591 nm and 646 nm, but when it is metallated with zinc only two Q bands at 558 nm and 599 nm are observed. This spectrum is consistent with what has previously been observed for zinc metallated porphyrins.[120] Zinc metallation of the porphyrin has no effect on the absorbance of the oligonucleotide at 260 nm. Further analysis was carried out by recording the fluorescence spectrum of the zinc metallated porphyrin at 425 nm. The emission spectrum showed two bands at 608 nm and 658 nm compared to 651 nm and 714 nm as seen with the free base porphyrin.

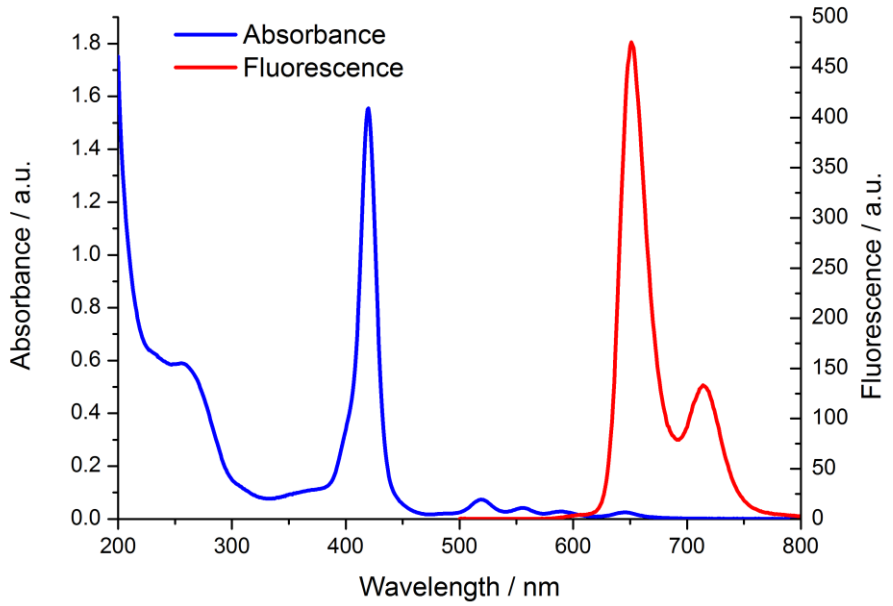


Figure 3.5 UV-Vis and fluorescence spectra of **DD2** before HPLC purification

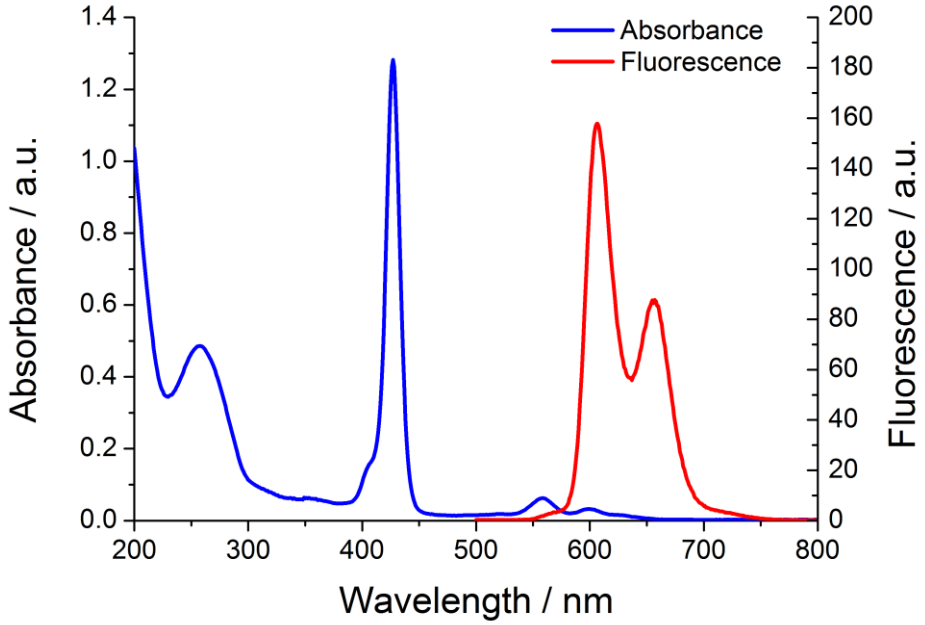


Figure 3.6 UV-Vis and fluorescence spectra of **DD2Zn** after HPLC purification

HPLC-Mass spectrometry was used to further characterise **DD2Zn** (Figure 3.7). A mass of 4694.8 was expected, and 4693.8 was observed in the ESI positive mode. The mass spectrum result showed that the correct product had been synthesised and

purified, and that it had to be metallated. The accuracy of this result is not good enough to confirm whether copper or zinc had been incorporated, but on the strength of the absorbance and emission spectra we are confident that the porphyrin is zinc metallated.

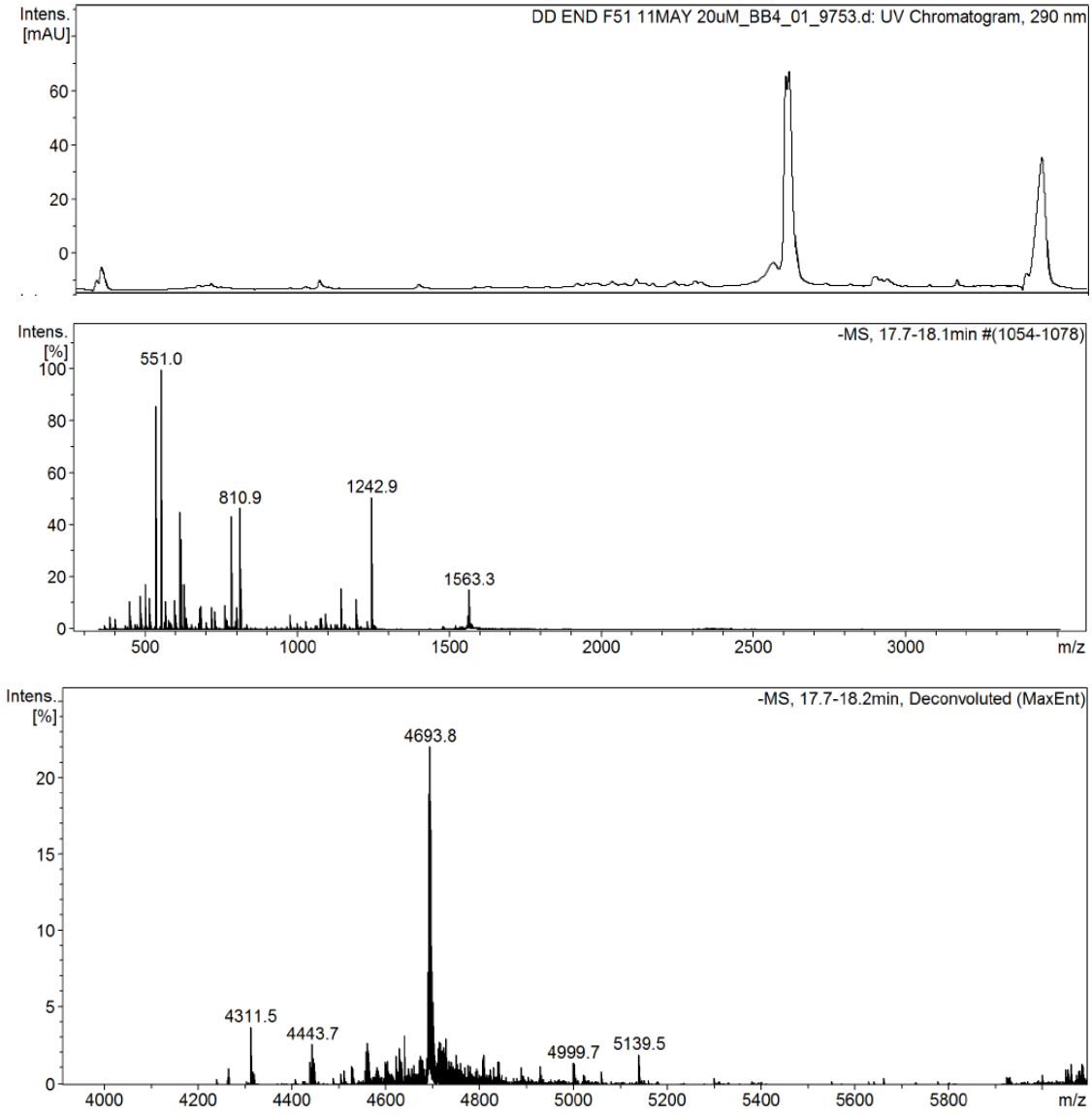


Figure 3.7 HPLC-MS results for **DD2Zn**. The top box shows the HPLC trace, the middle box shows the mass spectrum for the product peak and the bottom box shows the deconvoluted mass spectrum

Many previous attempts at analysing porphyrin DNA have been unsuccessful with samples being analysed in house and also at the EPSRC National Mass Spectrometry Centre at Swansea University. The exceptions to this were porphyrin nucleotide dimers two bases in length[102] and short oligonucleotides (6 to 12 bases in length) with a single porphyrin modification as observed by other research groups.[115, 121] In the case of the

latter paper, the published mass spectrometry data was for a test sequence as the 14-mer with one modification and the 39-mer with two modifications had no mass spectrometry data attributed to them. This will be worth considering when synthesising future porphyrin-modified oligonucleotides.

After purification, only 15 nmols of **DD2Zn** were obtained. This is a very small amount and would not be enough for crystallography. Taking this into consideration, the synthesis of **DD1** was carried out on a larger scale (4 x 1 μmol). Deprotection and cleavage were carried out using the same method as before and the four separate synthesises were initially purified using Glen-Pak purification. The next step was to metallate the porphyrins prior to HPLC purification, as this would prevent any problems made by metallation during HPLC. A method for zinc metallation of porphyrin-modified DNA had been developed previously within the Stulz research group.[82, 119] The free base porphyrin DNA is heated to 85 °C with zinc acetate (200 equivalents per porphyrin modification) for five minutes. During this heating the porphyrin-modified DNA precipitates due to the co-ordination of zinc to the phosphate backbone. After five minutes EDTA solution is added to redissolve the DNA and sequester the excess metal ions. Sodium chloride solution was added to the metallated porphyrin DNA sample before purification using a Glen-Pak column. The hydrophobic nature of the porphyrin means the porphyrin DNA sticks to the Glen-Pack column whilst excess salts can be washed away and sodium ion exchange can be performed on the phosphate backbone. The metallated porphyrin DNA can then be eluted from the column and analysed by UV-Vis and fluorescence spectroscopy.

Incomplete metallation of the porphyrin DNA was observed when carrying out this method with **DD1**. This was confirmed by the mixture of porphyrin species present in the UV-Vis and fluorescence spectra. To achieve complete zinc metallation of **DD1** the amount of zinc acetate was increased to a 1000 fold excess. This increase was successful and multiple attempts to achieve complete metallation using the previous method were not needed.

Once zinc metallated, **DD1Zn** purification was attempted using reverse-phase HPLC. This time, separation of the product from the failure sequences was slightly more difficult as the porphyrin modification was added midway through the porphyrin DNA synthesis. It was found that there were more porphyrin DNA peaks in the HPLC trace than observed for **DD2**. The porphyrin modification may play a part in hindering the coupling efficiency of any subsequent nucleobases once added to the oligonucleotide due to its large size, hence the increase in failure products. Drawing from the previous experience of HPLC, it was decided to load less porphyrin DNA onto the column at a time. It was also decided to use a smaller analytical column so to reduce the chance of porphyrin DNA being retained on the column. Previously, the column temperature during the HPLC purification was set to 60 °C, but now the

porphyrins had been zinc metallated the column was kept at room temperature to avoid any demetallation. However, by doing this there is an increased chance secondary structures can be formed on the column and make the separation more difficult. To address this potential problem a shallower solvent gradient was used to help separate products (Table 3.2).

Time (min)	Buffer A (%)	Methanol (%)
0	100	40
25	60	40
100	20	80
105	0	100
107	0	100
117	100	0
120	100	0

Table 3.2 Solvent gradient used for HPLC purification of zinc metallated **DD1Zn**. Buffer A = 100 mM hexafluoroisopropanol and 8.6 mM triethylamine in water. Column temperature = 20 ˚C

To begin with 20 nmol of **DD1Zn** was injected onto the column. This was initially thought to be a sensible amount to inject, but large amounts of streaking, and poor separation was observed. 5 nmol injections were found to be much more suitable even though some streaking was still observed. The downside to this was that many injections of **DD1Zn** were required to get the desired amount of pure material for crystallography. Using this method 110 nmol was obtained that was pure enough to be used for crystallisation (Figure 3.8).

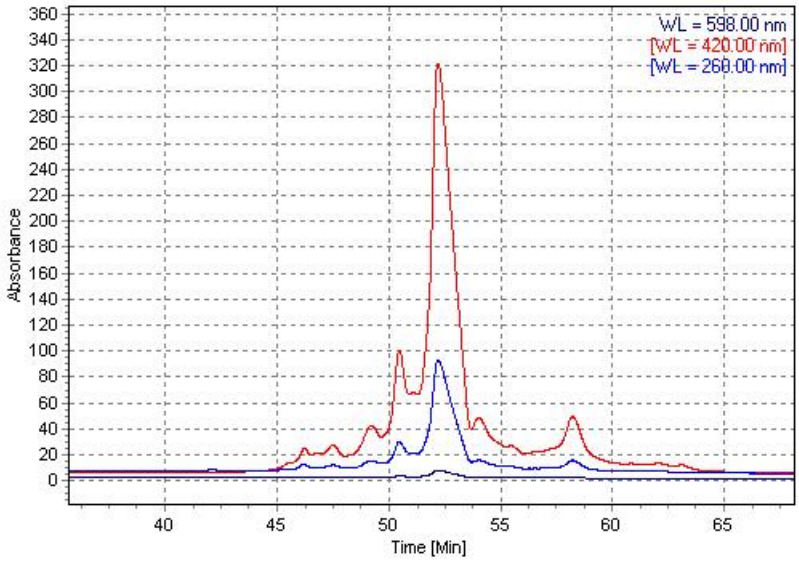


Figure 3.8 Analytical reinjection of **DD1Zn**

3.7 Crystallisation of the Dickerson dodecamer

Attempts to grow crystals of **DD1Zn** were carried out with the guidance of Dr. Ivo Tews at the Institute for Life Sciences, University of Southampton. **DD1Zn** was dissolved in 10 mM HEPES buffer to give a final concentration of 1 mM. The sample was supplemented with 10 mM magnesium chloride. A second set of samples also contained 2 mM spermine, which has been previously found to be beneficial for crystallising the unmodified Dickerson dodecamer. However, when spermine tetrachloride was added the porphyrin-modified DNA precipitated out of solution. This was a result we were not expecting and made that sample unsuitable for the subsequent screening. The addition of spermine is meant to stabilise the helical structure of the DNA duplex. Here, it seems it is having a similar effect to when the porphyrin DNA was in a zinc acetate solution during metallation. Spermine is a polycationic molecule (4+) that can interact with the negatively charged phosphate backbone of DNA. When this occurs less of the phosphate backbone is able to interact with the water molecules that keep the DNA in solution and so the DNA precipitates.

Screening took place using 96 different pre-made conditions (Natrix HT deep well block, Hampton Research). A robot was used to pipette the 96 different conditions onto a 96 well plate and then pipetted the porphyrin DNA solution into each of these wells. Each well was checked under a microscope immediately after setting up the screen to ensure the pipetting of all the solutions had been successful. The screen was then stored at 20 °C before being examined a week later. This time it was found that the porphyrin DNA had precipitated in every sample. Two more screenings were set up using PEGRx HT (Hampton Research), which is designed for macromolecules and PACT

premier HT (Molecular Dimensions), which is designed for proteins. As the large hydrophobic porphyrin modification appears to have a large effect on the physical properties of the DNA, it was thought by using screening conditions more suited towards proteins it could be possible to have more success with crystallisation. Again, a robot was used to prepare the screens, which were then stored at 20 °C. Examinations of the screens took place weekly. It was observed again that the majority of the porphyrin DNA precipitated out of the solution. Unfortunately, no material was suitable for X-ray crystallography.

3.8 Conclusions and outlooks

We have looked into the synthesis and purification processes for porphyrin-modified oligonucleotides. The purification of these oligonucleotides has always been difficult because of the addition of the hydrophobic porphyrin. A single porphyrin modification can have a surprisingly large effect on the physical properties of the oligonucleotide. This makes traditional purification techniques difficult to employ. Reverse-phase HPLC was found to be effective for purifying porphyrin-modified oligonucleotides, but only small amounts can be purified at a time due to the issues with streaking and porphyrin DNA sticking to the column. This can be an issue if a large amount of material is needed for analysis by NMR spectroscopy or X-ray crystallography. Care must be taken with HPLC purification as zinc metallation of the porphyrin is possible with the solvents used in this work. Metallation of the porphyrin prior to HPLC purification eliminates this problem, but is not so useful if the study of the free base porphyrin is required. Considering these issues and the overall difficulty of making porphyrin-modified DNA, it would be worth investigating other methods of synthesising it. This could be achieved by addition of the porphyrin to the oligonucleotide after solid phase synthesis. Another alternative, which is covered in Chapter 6, is the incorporation of a porphyrin-modified nucleotide into the DNA using DNA polymerases.

We briefly explored the crystallisation of porphyrin DNA during this work. Obtaining a crystal structure of the porphyrin-modified DNA would be highly beneficial for determining the effect the porphyrin has on the DNA structure and for comparisons with past work carried out in the Stulz research group. Unfortunately all attempts so far have been unsuccessful. This is hardly surprising considering the unpredictable nature of crystallisation. More screenings should be set up in the future but more porphyrin-modified DNA needs to be synthesised first. This is a major disadvantage to this study as crystallography requires a large amount of highly purified sample to work. We have established that to do this for porphyrin-modified DNA is a far from simple task. If new methods of purification can be developed to improve this process,

then the task of finding the ideal conditions for crystallography could be achieved much more quickly. However, it also must be considered that it might not be possible to crystallise this porphyrin-modified oligonucleotide.

One surprising observation from the screening attempts was that the porphyrin DNA precipitated out of the HEPES buffer upon addition of spermine. Spermine is often added to DNA samples to help stabilise the duplex structure. The crystallisation of the original unmodified Dickerson dodecamer DNA was found to need the addition of spermine so it was disappointing that the porphyrin DNA did precipitate upon the addition of spermine. At this time we cannot give an exact explanation for why the porphyrin DNA does precipitate but we can hypothesise what is occurring. The porphyrin modification on the DNA is highly hydrophobic so it could be that when the spermine, which itself has three hydrophobic linking carbon chains, is added associates with the DNA the extra hydrophobicity causes the DNA to precipitate from the HEPES buffer.

4 Porphyrin-modified G-quadruplex DNA

4.1 Aim

It is known that G-rich DNA sequences of the type $(GGG)X_n$ can form stable intramolecular quadruplex structures containing stacks of G-tetrads.[122] These secondary structures are stabilised by the presence of K^+ ions and play a crucial role in telomere structure and the control of gene expression. Large aromatic molecules, such as porphyrins, can further stabilise the G-quadruplex structure by end stacking on top of the loops or in a side-on mode in the groove of the G-quadruplex.[123] So far such stabilising agents have not been covalently connected to the DNA, but the development of G-quadruplexes with covalently bound porphyrins could lead to a new diagnostic tool to select and possibly inhibit G-quadruplex binding proteins.

4.2 An introduction to G-quadruplexes

G-quadruplex is the general name given to the four-stranded structure that is made from a guanine-rich sequence of DNA. Four guanine bases interact by forming eight hydrogen bonds using their Hoogsteen and Watson-Crick faces (N_1-O_6 and N_2-N_7) to form a G-tetrad (Figure 4.1), whose structure was first identified in 1962.[124] The G-tetrad is a square planar structure and is stabilised by the presence of monovalent cations such as potassium or sodium.

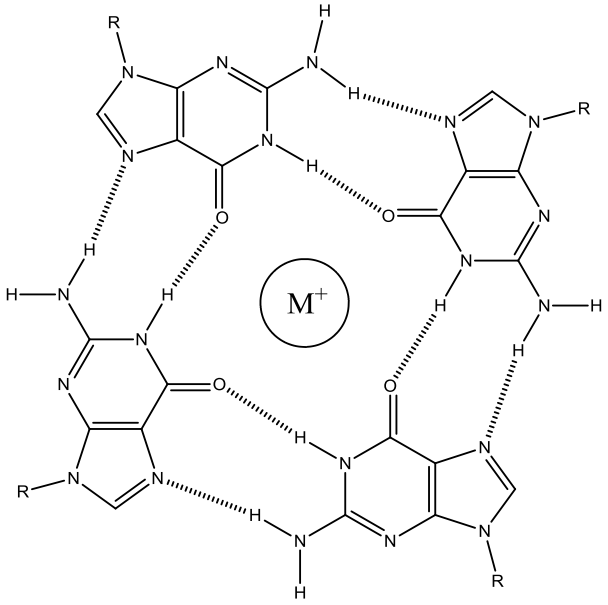


Figure 4.1 Hydrogen bonding pattern in a G-tetrad

4.2.1 Stabilising the G-quadruplex

To maximise these interactions bases in both duplex DNA and G-quadruplex DNA stack on top of each other. Due to their large π-surfaces, G-tetrads prefer to stack on top of each other maximising favourable π-π stacking. The stacked G-tetrads align themselves in a similar manner to that of duplex DNA, characterised by a regular rise and twist between the G-tetrad planes to give a right-handed helical structure.

Cations are important to the stabilisation of the G-quadruplex. The O6 of guanine can co-ordinate monovalent and divalent cations.[13, 125] The degree of stability offered by the cation depends on its valency and size. It has been shown that potassium ions provide the largest amount of stabilisation and sodium ions are also effective at stabilising G-quadruplexes. The potassium ions are too large to sit directly in the plane of the G-tetrad so sit slightly above the plane. In this position the potassium ion can coordinate to eight guanine O6 atoms from two G-tetrads (Figure 4.2). Smaller ions such as sodium will sit more in the plane of the G-tetrad and therefore do no co-ordinate to the eight guanine O6 atoms as well.[126] A general trend in amount of stability offered by monovalent cations ions from the most stable to the least is as follows: $K^+ > Na^+ > Rb^+ > NH_4^+ > Cs^+ > Li^+$.[127] The concentration of the cation also has an effect on the overall stability of the quadruplex with higher melting temperatures observed at higher cation concentrations. Without a bound cation present it is impossible for a G-quadruplex to form, as the cyclic arrangement of the four oxygen atoms clustered in the centre of the G-quartet would become electronically unfavourable.[128]

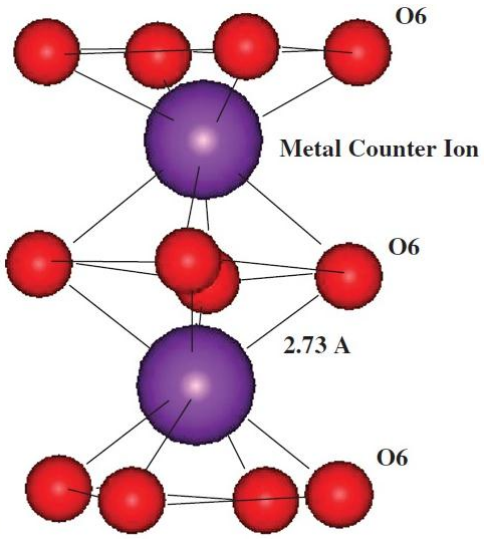


Figure 4.2 Two potassium ions each coordinated between eight carbonyl oxygens of two G-tetrads[129]

4.2.2 Folding and topology of G-quadruplexes

G-quadruplexes can be formed from one, two or four separate strands of DNA (or RNA) and often display a wide variety of topologies, which is mainly due to the various possible combinations of strand direction, loop size and sequence. The simplest form of a G-quadruplex is made from four DNA strands containing short guanine runs that can self-associate with each other in the presence of a suitable stabilising cation (Figure 4.3a), for example d(TGGGGT). Complexity can arise in tetramolecular G-quadruplexes when the directionality of each strand is considered. If the backbone of all the strands run in the same direction and all the bases are in the *anti* glycosidic orientation, the quadruplex is said to be parallel.[126] If at least one of the DNA strands run in the opposite direction and the bases are in a mixture of *anti* and *syn* glycosidic orientations, the G-quadruplex is termed antiparallel. However, only parallel arrangements of tetramolecular G-quadruplexes have been observed experimentally.[126, 130, 131] The rate of formation and dissociation of tetramolecular G-quadruplexes has been found to be very slow in comparison to intramolecular quadruplexes but they are exceptionally stable in the presence of potassium ions.[132]

The next level of complexity can be found in the bimolecular G-quadruplex. Formation of these G-quadruplexes requires two DNA strands containing two runs of guanine residues separated by a loop sequence. Eg.

$$X_a G_b X_c G_b X_a$$

Where X_a is any non-guanine nucleotide of length *a*, G_b is any number of guanines involved in tetrad formation of length *b*, and X_c is any nucleotide of length *c* involved in loop formation (Figure 4.3b).

Numerous examples of the structures of these G-quadruplexes have been reported by both NMR spectroscopy and X-ray crystallography. [133, 134] Not only does the directionality of the guanine tracts need to be considered when studying the structure, but also how the loop affects the overall topology. The two loops can be arranged in different ways, they can be diagonal, lateral (edgewise) or external (propeller) to the quadruplex. The sequence and length of the loop can have a major influence on the type of folding.

The final type of G-quadruplex is the intramolecular quadruplex. This form is made from a single strand of DNA and contains at least four guanine tracts separated by three loop sections (Figure 4.3c). Intramolecular G-quadruplexes have the following general sequence:

$$X_a G_b X_c G_b X_c G_b X_a$$

Where X_a is any non-guanine nucleotide of length *a*, G_b is any number of guanines involved in tetrad formation of length *b*, and X_c is any nucleotide of length *c* involved in loop formation.

Due to the increased possibilities in G-quadruplex folding, the intramolecular G-quadruplexes are the most varied and complex G-quadruplexes to study. The length and sequence of the loop and the number of guanines present in the G-tract are contributing factors in the complexity of these structures. As with bimolecular G-quadruplexes, the same loop types are possible. However, it is not possible for an intramolecular G-quadruplex to begin or end with a diagonal loop. Intramolecular G-quadruplexes are mainly associated with telomeric regions at the ends of chromosomes, but can also be found in G-rich regions in the genomes of some organisms. These G-rich regions are typically double-stranded with the complementary C-rich region. In order for these regions to be able to fold into G-quadruplex structures the DNA duplex must unwind to leave the G-rich region present as a single strand.

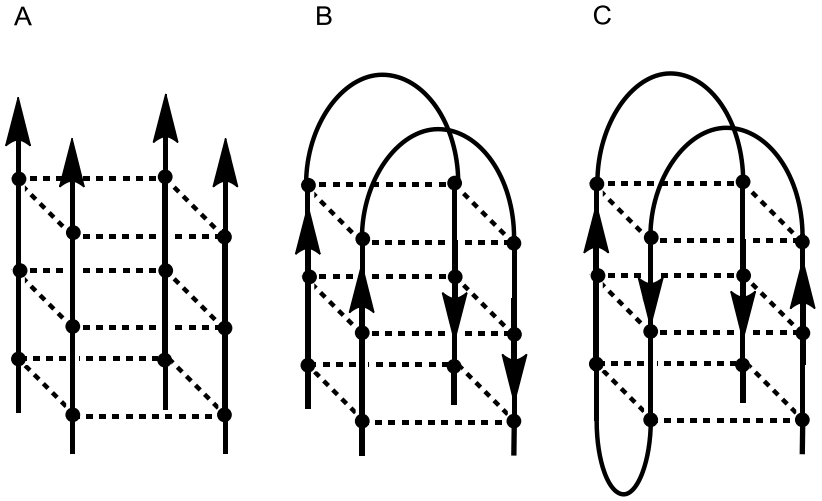


Figure 4.3 Examples of G-quadruplexes A) tetramolecular quadruplex, B) bimolecular quadruplex C) unimolecular quadruplex

The loop length of G-quadruplexes has a large effect on the overall structure and stability of the G-quadruplex.[135, 136] Various loop lengths are feasible, with the shortest loop possible consisting of a single nucleotide in length. The loops can be arranged in several different ways; double chain reversal (propeller) loops link two adjacent parallel strands, while lateral or diagonal loops link two antiparallel strands (Figure 4.4). G-quadruplexes often contain a mixture of loops that can be arranged in different ways. This property means that G-quadruplexes are often highly polymorphic, which can make the study of their structure problematic. However, there are more simple

examples possible, G-quadruplexes containing a single nucleotide in each loop will exclusively form a parallel structure[137]. The number of guanines present in the G-tetrad has an effect on the stability of the G-quadruplex structure,[132] although any changes in the loops will change the degree of stabilisation for intramolecular G-quadruplexes.

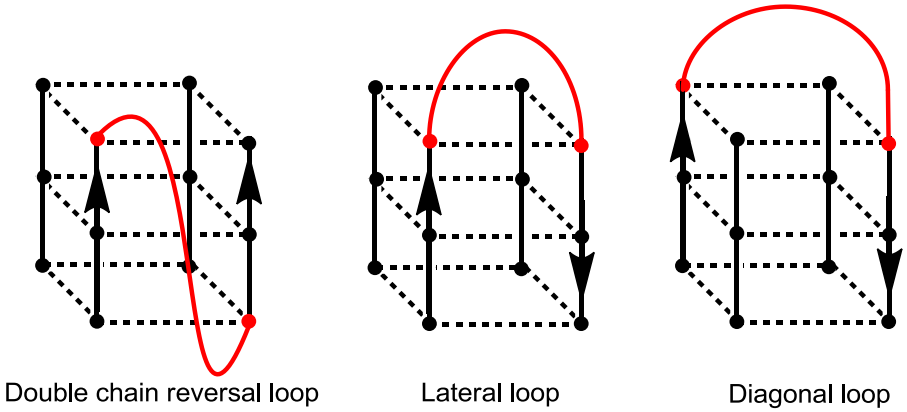


Figure 4.4 G-quadruplex schematics that show possible loop types

The directionality of the strands in G-quadruplexes can be related to the conformational state of the glycosidic bond between the guanine base and the sugar. The glycosidic bond can either be *syn* or *anti* (Figure 4.5). When all four strands are parallel, all the bases are in the *anti* conformation and the grooves between the backbones are of equal size. When two of the strands are joined by a lateral loop (so are antiparallel), one of the paired bases and the glycosidic torsion angle needs to be inverted so that the hydrogen bonding between the guanines is retained. This is usually done in one of two ways, either each strand has a run of alternating *syn*, *anti* glycosidic angles for the guanine bases, or one strand is exclusively in the *syn* conformation and the other exclusively in the *anti* conformation. Depending on the direction of strands and the type of loop between strands, the correct Hoogsteen hydrogen bonding must be retained to form the G-quadruplexes. The mixture of *syn* and *anti* conformations of the glycosidic bonds in antiparallel G-quadruplexes give rise to a variety of different sized grooves in the overall helical structure. This gives a variety of targets to aim for when designing small molecules that can bind to G-quadruplexes.

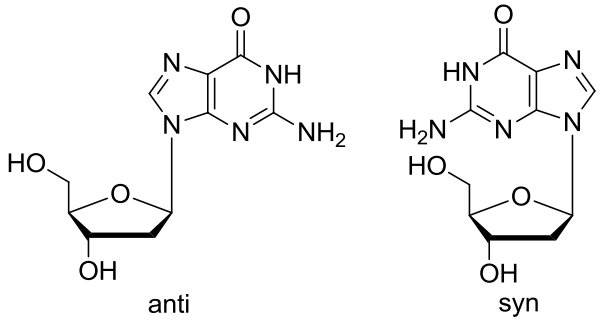


Figure 4.5 *Syn* and *anti* conformations of 2'-deoxyguanosine

4.2.3 Methods of characterisation

There are many experimental techniques used to explore different features of G-quadruplexes, but most of these are mainly descriptive and complete structure determination requires either NMR studies or X-ray crystallography. As stated before, the highly polymorphic nature of G-quadruplexes makes their full characterisation challenging.

4.2.4 X-ray crystallography

X-ray crystallography is extremely important for complete structure characterisation, but the formation of crystals is a relatively slow and uncertain procedure. However, if a crystal can be obtained extensive characterisation can be carried out. To date, over 50 crystal structures of G-quadruplexes are available in the protein or nucleic acid data bank. It should be considered that the crystal structure may not be a fair reflection of the solution phase structure.

4.2.5 Nuclear magnetic resonance (NMR) spectroscopy

NMR spectroscopy is also used for structure determination.[138] ^{1}H NMR spectra can be used to look at the guanine NH_1 imino protons, which have a characteristic shift when hydrogen bonded. For example, if there are 3 G-tetrads in the G-quadruplex, then there should be 12 imino proton signals from the 12 guanine residues assuming there is only one conformation of the G-quadruplex present (Figure 4.6). However, if there are many polymorphs present the number of imino proton signals will be much larger. For more detailed analysis, multidimensional NMR techniques are required, which help to reveal structural features such as backbone conformations and sugar pucker angles.

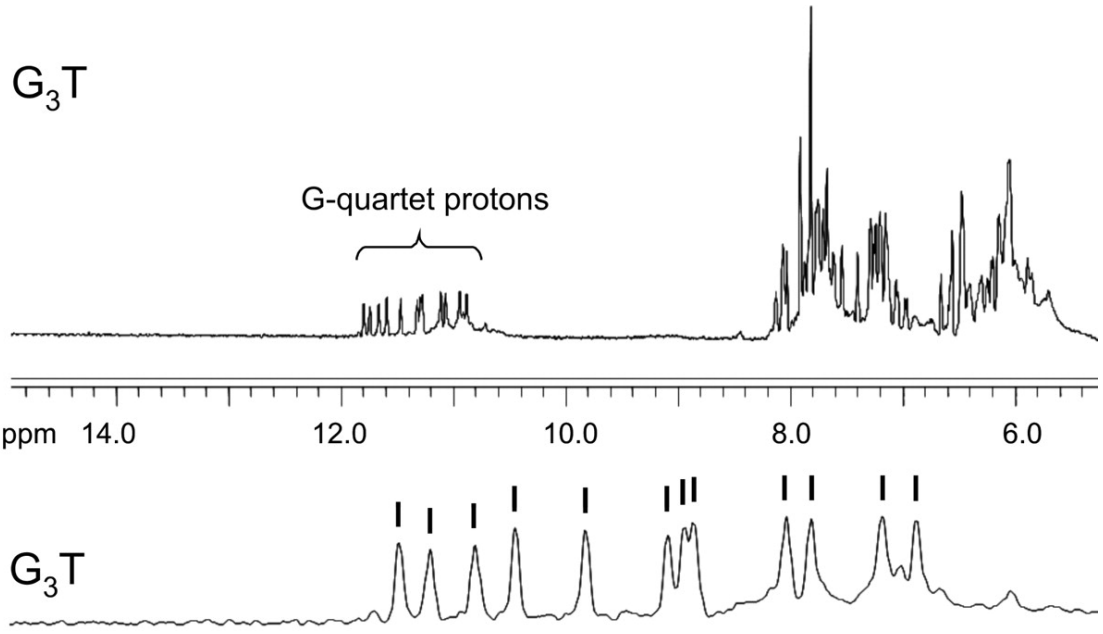


Figure 4.6 ^{1}H NMR spectrum of the parallel quadruplex forming oligonucleotide (G_3T). There are 12 clear imino proton signals indicating no polymorphism present[135]

4.2.6 UV melting analysis

UV melting can be used to study the stability of G-quadruplexes.[139] G-quadruplexes become unstable at higher temperatures and will begin to dissociate. This will cause the wavelength at which the nucleic acids absorb to change. By monitoring the absorbance of UV light at 295 nm, there is a hypochromic shift upon melting.[140] A melting curve measuring absorbance vs. temperature can be plotted to work out the melting temperature (the temperature where half of the G-quadruplexes have denatured). Assuming the melting is a two-step process, a detailed van 't Hoff analysis to extract values for variables such as ΔG, ΔH and ΔS can be used. Many G-quadruplexes are very stable under pseudo physiological conditions (100 mM KCl, pH 7.4) with melting temperatures observed above 60 °C. The loop sequences have large contributions to the melting temperature, for example, a G-quadruplex with a single thymine in all of the loops has a melting temperature in excess of 95 °C.[135]

4.2.7 Circular dichroism spectroscopy

CD spectroscopy (CD) is a ground state absorption technique analogous to UV-Vis spectroscopy. It measures the differential absorption of circularly polarised light. For this to occur, the molecule must be chiral. Therefore, CD is a useful tool in studying different nucleic acid conformations, such as G-quadruplex structures. The technique is limited in how much structural data can be obtained, but discriminating between quadruplex topologies having parallel and anti-parallel strand orientations is possible. In general, a peak in CD at 260 nm wavelength and a trough at 240 nm is indicative of an all-parallel structure, whereas a peak at 295 nm and a trough at 260 nm describes and antiparallel structure (Figure 4.7). Polymorphic forms will contain a superposition of the CD spectra of each individual conformation. Taking this into consideration, care must be taken with drawing conclusions into the structure present as out of the 26 theoretical loop arrangements available to a sequence containing four G-runs, 25 will be anti-parallel and one will be parallel. [141]

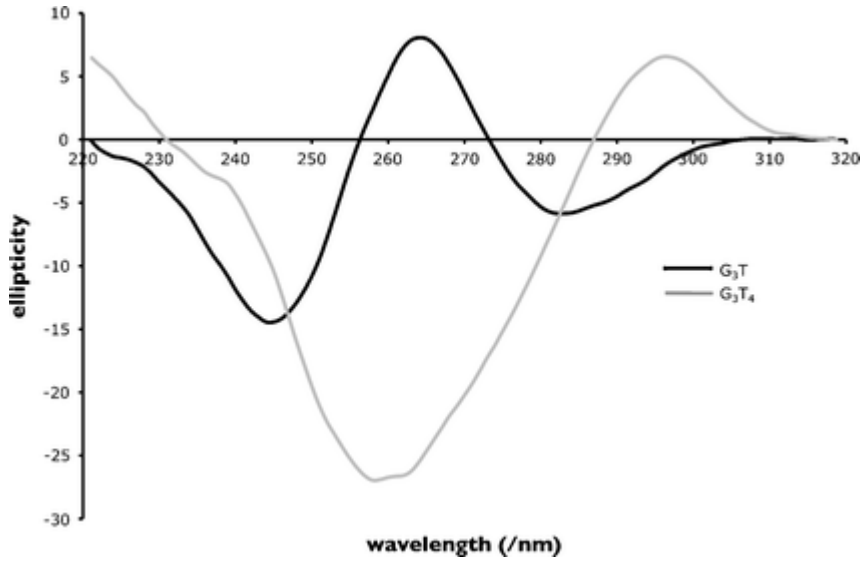


Figure 4.7 Circular dichroism spectra of G-quadruplexes. Sequences shown are G_3T = $d(TG_3TG_3TG_3TG_3T)$ and G_3T_4 = $d(TG_3T_4G_3T_4G_3T_4G_3T)$[137]

4.2.8 Polyacrylamide gel electrophoresis (PAGE)

Native polyacrylamide gel electrophoresis (PAGE) can be used to analyse G-quadruplex formation, as the gel mobility of DNA in this secondary structure is different to that of the unfolded G-quadruplex. It has been observed that the folded structure will migrate more rapidly on the gel compared to the equivalent unfolded structure.

4.3 G-quadruplex functions

4.3.1 Telomeres and telomerase

A telomere is a region of repetitive nucleotide sequences found at the end of a chromosome. In human chromosomes the telomeric DNA typically consists of 5-8 kilobases of a double-stranded tandem repeat of the sequence TTAGGG with a single-stranded 3'-end overhang of 100-200 bases. The purpose of the telomere is to protect the chromosome from degradation as a short amount of DNA is lost during DNA replication in a process known as the end replication problem (Figure 4.8). DNA polymerases are needed to synthesise new DNA strands during replication, but they only work in the 5' to 3' direction. This is an issue since the two strands of DNA in the DNA duplex run in different directions. So polymerase can easily make a complementary DNA strand to the leading strand, which runs in the 3' to 5' direction, but cannot do this for the lagging strand. To solve this problem short sequences of RNA are synthesised to act as primers at regular intervals on the lagging strand. The primers provide a 3'-hydroxyl group for DNA polymerase III to synthesise fragments of DNA between the RNA primers. These DNA fragments are called Okazaki fragments. The RNA is then replaced with DNA by DNA polymerase I and then DNA ligase links the DNA fragments together. However, this process will lead to shortening of the lagging strand. In the step where the RNA primers are converted to DNA there must be some DNA present in front of the RNA primer. This is not the case for the primer found at the terminal end of the lagging strand so it is not converted to DNA. In the end this RNA primer is digested by enzymes and the lagging strand is shortened.

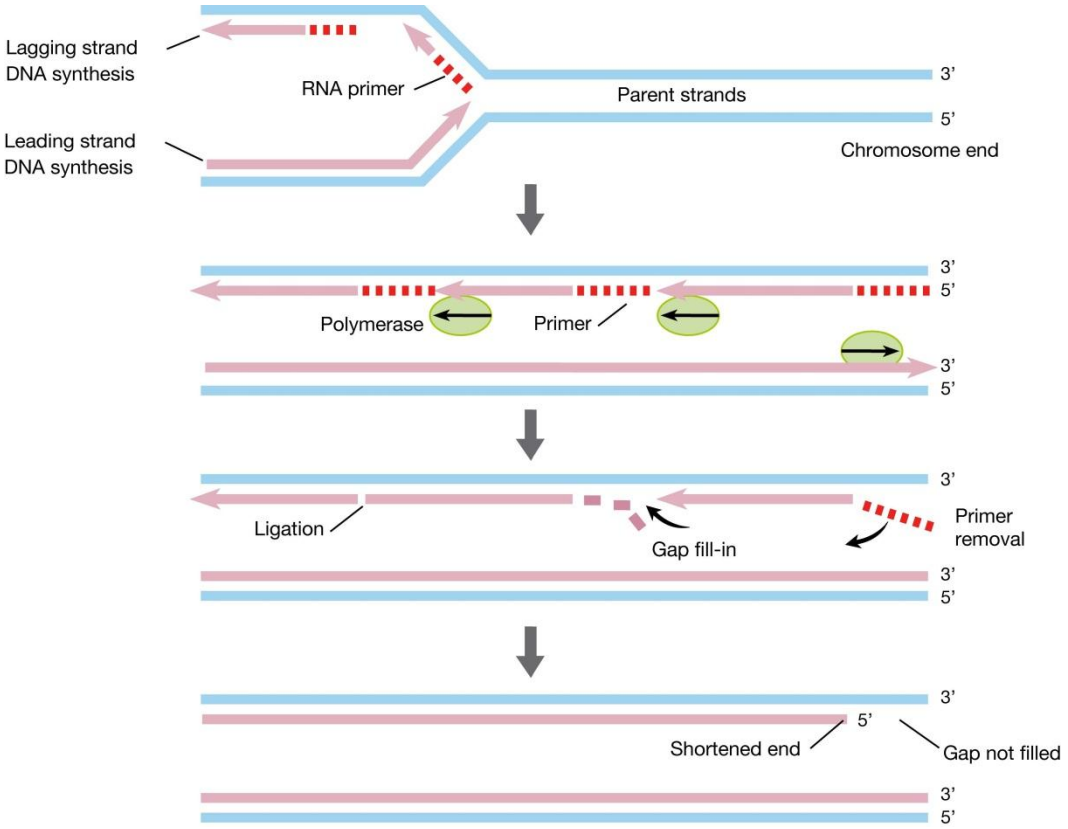


Figure 4.8 The end replication problem. Copying of the lagging strand of a linear DNA molecule by DNA polymerase occurs in a stepwise fashion dependent on an RNA primer that is subsequently degraded. The gap is filled in and ligated to produce an intact DNA molecule. At the very end, DNA polymerases are not able to fill the gap with the result that the end will become shorter upon each round of replication[142]

Telomeres and their importance in protecting chromosomes were discovered by Elizabeth H. Blackburn, Carol W. Greider and Jack Szostak. In 2009 they received the Nobel Prize in Physiology or Medicine for their research on telomeres and the enzyme telomerase. In the late 1970s Szostak had introduced linear plasmid DNA molecules into yeast cells and despite the DNA being highly recombinogenic, long term maintenance of the chromosomes was not possible.[143] Blackburn and Szostak worked together on this problem by introducing *Tetrahymena* chromosome end repeat DNA sequences to linear yeast plasmids in order to stabilise the ends to allow replication. The addition of the repeat DNA sequences at the ends of the linear yeast plasmid provided chromosomal stability.[144] This breakthrough provided evidence that the purpose of the DNA repeats at the end of chromosomes are there to protect the chromosome from the end replication problem.

These experiments also showed evidence of telomere maintenance in yeast, but how this was happening could not be explained.[145] One idea suggested that an enzyme

is responsible for the telomere maintenance. Further experiments carried out by Greider and Blackburn showed that a reverse transcriptase enzyme called telomerase was carrying out the telomere maintenance.[146-148] Telomerase consists of several components that include an endogenous RNA template of 11 nucleotides. This template is required to extend telomeric regions of DNA. A reverse transcriptase subunit then catalyses the addition of telomeric repeats. Telomerase needs the single-stranded telomere at the 3'-end of DNA as a primer in order for telomere extension to occur.[149]

There is currently a lot of interest in telomeres and telomerase as they have significant medical implications in many fields including cancer, ageing, stem cell maintenance and hereditary disease syndromes. It has been observed that 80-90 % of cancers have increased activities of telomerase.[150] This means that through many replication cycles the chromosomes in cancer cells will never shorten. In somatic cells when the telomere length is too short and the chromosome is now unprotected by the telomere, the cell enters a state of senescence where end to end fusion and chromosomal instability ultimately leads to cell death. Targeting telomerase to inhibit its role in telomere lengthening is a viable route in the treatment of cancer. One way of inhibiting telomerase lies in the G-quadruplex forming ability of the human telomeric sequence TTAGGG. Whether the G-quadruplex structure is naturally formed within the human telomere is still debated, but recent studies suggest it is possible.[151] If the G-quadruplex structure can be formed within the telomere, elongation can be negatively regulated. This was shown to be possible by Zahler who demonstrated that G-quadruplex structures could be formed and stabilised in the presence of K^+ ions, which prevented telomerase from elongating the telomere.[152] Since this discovery there has been a drive to find small molecules that can stabilise the telomere G-quadruplex, but a greater understanding of the structure is needed so such molecules can be developed for this purpose.

Due to its sequence a large number of different structures have been observed in the human telomere.[153] However, a small number of structures have been successfully characterised by NMR spectroscopy and X-ray crystallography (Figure 4.9). NMR spectroscopy was first used to study the 22-mer sequence $d[AG_3(T_2AG_3)_3]$ in the presence of sodium and revealed an antiparallel topology with one diagonal and two lateral loops.[154] Later, the same sequence was crystallised from a solution containing potassium ions, but showed an all-parallel topology.[155] This was somewhat surprising but it shows that the type of cation present has an effect on the structure of the G-quadruplex. However, there is a debate as to whether the crystal structure is a fair representation of the telomeric G-quadruplex in solution. There have been further NMR studies of human telomere G-quadruplexes in potassium solutions that show (3 +1) G-quadruplex topologies. The (3+1) nomenclature arises from the distinctive pattern of

3 + 1 *syn* and *anti* glycosidic angles at each G-tetrad. Two forms of this G-quadruplex have been observed, the hybrid-1,[156, 157] which has a single chain-reversal loop followed by two lateral loops and the hybrid-2, which has two lateral loops followed by a single chain-reversal loop.[158, 159] The difference in structure in these cases was due to the flanking sequences of the G-quadruplexes studied. Small changes in the flanking sequences can have large effects on the overall structure of the G-quadruplex and the amount of polymorphism present.

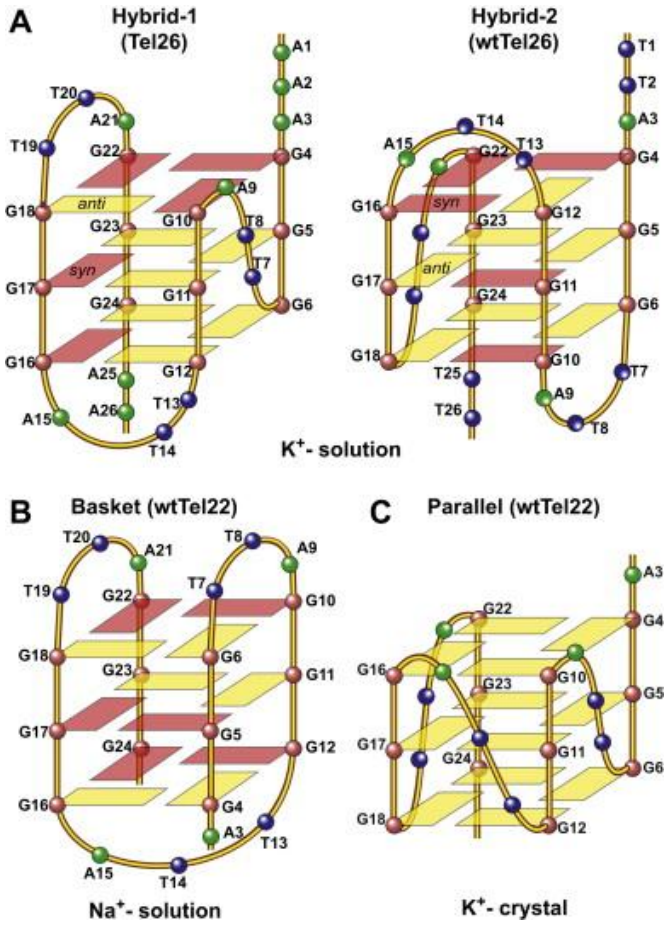


Figure 4.9 (A) Schematic drawing of the folding topologies of the Hybrid-1 (major conformation in Tel26 – $d[A_3G_3(T_2AG_3)A_2]$) [156, 157] and Hybrid-2 (major conformation in wtTel26 - $d[T_2AG3(T_2AG_3)T_2]$) [36, 37] intramolecular telomeric G-quadruplexes in K^+ solution. Yellow box = (anti) guanine, red box = (syn) guanine; red ball = guanine, green ball = adenine, blue ball = thymine. (B) Folding topology of the basket-type intramolecular G-quadruplex formed by wtTel22 ($d[AG_3(T_2AG_3)_3]$) in Na^+ solution as determined by NMR.[154] (C) Folding topology of the propeller-type parallel-stranded intramolecular G-quadruplex formed by wtTel22($d[AG_3(T_2AG_3)_3]$) in the presence of K^+ in crystalline state.[155] The numbering system is based on wtTel26

4.4 Interactions with small molecules

It has been established that stabilising G-quadruplex structures within the human telomere can inhibit the telomere elongation action of telomerase. To achieve this stabilisation many ligands have been designed based on the structural observations of these G-quadruplexes. Most ligands are polyaromatic heterocycles designed for extensive π-stacking interactions with the G-tetrads. This π-stacking can either come from stacking to a terminal G-tetrad or by intercalation between G-tetrads. Figure 4.10 presents a small group of the large family of molecules designed to stabilise the G-quadruplex. Trisubstituted acridine ligands[160, 161] have been found to be effective at selectively binding to G-quadruplexes over DNA duplexes. In particular the 3,6,9-trisubstituted BRACO 19 ligand has been found to bind strongly to the G-quadruplex with effective telomerase inhibitory activity.[162] Another small ligand that has shown good telomerase inhibition is the fluorinated polycyclic quinoacridinium cation RHPS4.[163] A particularly potent telomerase inhibitor is the natural product telomestatin, which is isolated from *Streptomyces anulatus.*[164] Since this discovery other G-quadruplex binding ligands have been synthesised based on this structure.[165] Not only do the ligands need to be selective for G-quadruplex DNA over duplex DNA, but there are many other possible G-quadruplex forming sequences found in the human genome that need to be taken into consideration. Due to the polymorphic nature of many of these G-quadruplexes it is extremely difficult to design a ligand that will be selective based just on the G-quadruplex structure. However, the Balasubramanian group have taken this into account have been designing ligands that show better selectivity for specific G-quadruplexes.[166, 167]

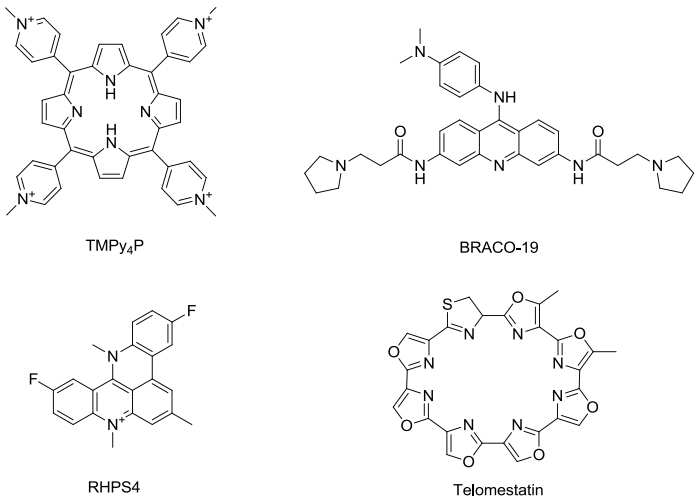


Figure 4.10 A small selection of the many small molecule G-quadruplex binding ligands

G-quadruplex structures have a large π-surface, roughly twice as large as that found in duplex DNA, since there are four coplanar bases instead of two. Porphyrins are a similar size to the G-tetrad, which makes them ideal candidates for binding to G-quadruplexes via π-π stacking interactions. The interactions of 5,10,15,20-tetra(*N*-methyl-4pyridyl)porphyrin ($TMPy_4P$) with G-quadruplex DNA has been widely studied.[168, 169] This cationic porphyrin also takes advantage of the negative backbone of DNA to improve binding affinity. $TMPy_4P$ like many other G-quadruplex binding ligands show no selectivity for G-quadruplex DNA over duplex DNA, which is a problem as duplex DNA far outnumbers G-quadruplex DNA in the human genome.

For $TMPy_4P$, the mode of binding to the G-quadruplex is still debated with many differing reports published. Two main models have been proposed for binding *in vitro*, namely intercalative binding between G-quartets[169, 170] and binding to the terminal exposed G-tetrads of the G-quadruplex via π-π stacking interactions.[168, 171] Neidle *et al.* have reported an X-ray structure of a G-quadruplex-$TMPy_4P$ complex (Figure 4.11), indicating that the $TMPy_4P$ molecules fail to directly interact with G-tetrads in the G-quadruplex[123] (pdb entry 2HRI). The DNA quadruplex in this study was parallel stranded with external double-chain-reversal propeller loops, but bimolecular as opposed to the intramolecular forms used in other studies. $TMPy_4P$ was found to bind to the quadruplex by stacking onto the TTA nucleotides, either as part of the external loop structure or at the 5' region of the stacked quadruplex. However, it must be remembered that many G-quadruplex sequences will fold in different ways to produce different structures, which will have a major effect on the binding of the porphyrin to the G-quadruplex.

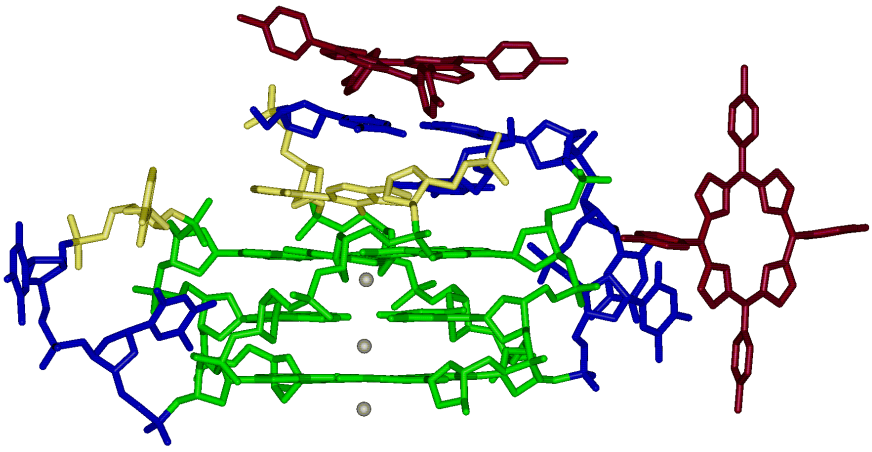


Figure 4.11 X-ray Crystal structure of $TMPyP_4$ (dark red) bound to the parallel bimolecular human telomeric G-quadruplex d(TAGGGTTAGGG) in the presence of K^+ ions - pdb entry 2HRI[123]

4.5 G-quadruplexes with covalently bound porphyrins

Most of the research regarding the interactions of small molecules and G-quadruplexes to date has just been concerned with non-covalent interactions. However, there are a few examples now where modifications to the G-quadruplexes are being covalently attached, such as "unlocked" nucleotides (UNA),[172, 173] LNA and pyrene nucleotides,[174, 175] or thymine to uracil substitution and 2'-OH modification.[176] Some of these G-quadruplex forming oligonucleotides have shown anti-HIV-1 activity, or have improved aptamer based thrombin binding. These modified G-quadruplexes could provide a novel class of bio-active DNA analogues. The covalent attachment of porphyrins is a novel approach to modifying G-quadruplexes. With this modification we intend to observe how the porphyrin modification affects physical properties such as the G-quadruplex structure and stability. It is hoped that the porphyrin-modified G-quadruplexes could then be used in a more applied situation, such as switching between G-quadruplex-duplex structures that could be potentially used in sensing and computing. The G-quadruplex forming oligonucleotide is based on the human telomeric sequence d[TAGGG(TTAGGG)$_3$TT], which is denoted **wtTel25b** for the unmodified oligonucleotide according to the convention used in literature.[177, 178] As interactions between porphyrins and the human telomeric sequence have been comprehensively studied by other research groups in the past, it will be beneficial to compare our observations with this previous research.

4.6 Porphyrin-modified G-quadruplex synthesis

Porphyrin-modified G-quadruplexes were synthesised using solid phase DNA synthesis as described in the previous chapter. The porphyrin is linked to the thymidine using a flexible propargyl-amide linker **dU**^{2HTPP} (Figure 4.12). The extra flexibility offered from this linker has been chosen to aid the porphyrins ability to fold over and π-stack to the exposed terminal G-tetrads.

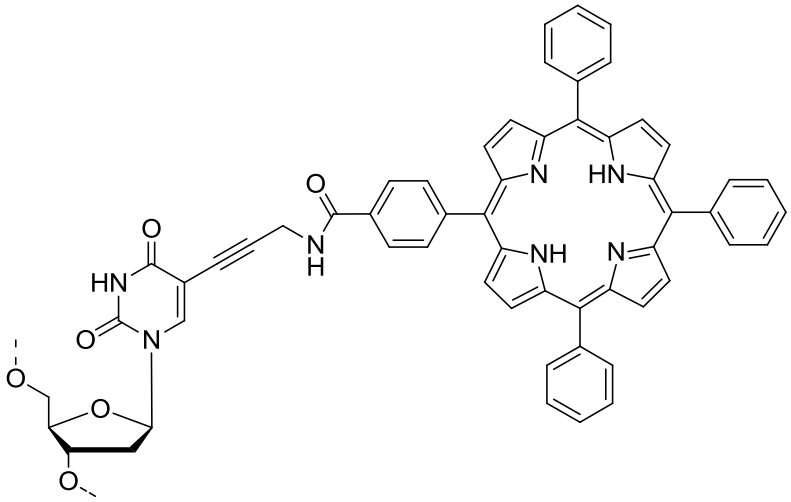


Figure 4.12 Flexible porphyrin monomer **dU**^{2HTPP} used in porphyrin-modified G-quadruplex synthesis

Strand Name	Sequence	Yield (nmoles)
wtTel25b	5' – TA GGG TTA GGG TTA GGG TTA GGG TT – 3'	436
Q1	5' – TA GGG **P**TA GGG TTA GGG TTA GGG TT – 3'	223
Q2	5' – TA GGG T**P**A GGG TTA GGG TTA GGG TT – 3'	291
Q3	5' – TA GGG TTA GGG **P**TA GGG TTA GGG TT – 3'	306
Q4	5' – TA GGG TTA GGG T**P**A GGG TTA GGG TT – 3'	335
Q5	5' – TA GGG TTA GGG TTA GGG **P**TA GGG TT – 3'	427
Q6	5' – TA GGG TTA GGG TTA GGG T**P**A GGG TT – 3'	390
Q7	5' – T**P** GGG TTA GGG TTA GGG TTA GGG **P**T – 3'	219
Q8	5' – TA GGG T**P**A GGG T**P**A GGG TTA GGG TT – 3'	342

Table 4.1 Sequences of the porphyrin-modified G-quadruplexes synthesised where **P** = **dU**^{2HTPP})

All modified strands were synthesised on a 1 μmol scale with the final base containing a fluorous DMT tag for purification following synthesis by fluorous affinity separation (Fluoro-Pak). The unmodified strands were synthesized on the same scale with the final DMT group left on for DMT-on purification using Glen-pack columns. HPLC analysis of all strands was carried out afterwards to assess purity. Concentrations of the DNA strands were calculated by first monitoring the absorption at 260 nm and then using the Beer Lambert law. Molar extinction coefficients were obtained by the molar extinction coefficient calculator provided by Applied Biosystems.[179]

4.7 UV-Vis spectroscopy of G-quadruplexes

The first characterisation of the porphyrin-modified G-quadruplexes was carried out by UV-Vis spectroscopy (Figure 4.13). The DNA was first dissolved in water and its spectrum was recorded. A feature of all the strands was that the peak maxima for the Soret band (420 nm) were always greater than that of the DNA (260 nm). The same DNA sample used for the UV-Vis analysis was then dried and redissolved in 10 mM lithium phosphate buffer solution (pH = 7.4) supplemented with 100 mM potassium chloride solution to induce the formation of a G-quadruplex. The sample was then heated to 95 °C and then left to anneal over a period of four hours. Once the G-quadruplexes had annealed they were analysed by UV-Vis spectroscopy. This time the peak at 260 nm shows a hyperchromic shift and the peak at 420 nm shows a hypochromic shift. This occurred in all samples and the change was large enough in all cases for the Soret band peak (420 nm) to be smaller than the DNA peak (260 nm). It appears that there has been a change in DNA structure from a single strand with no secondary structure to a G-quadruplex DNA structure.

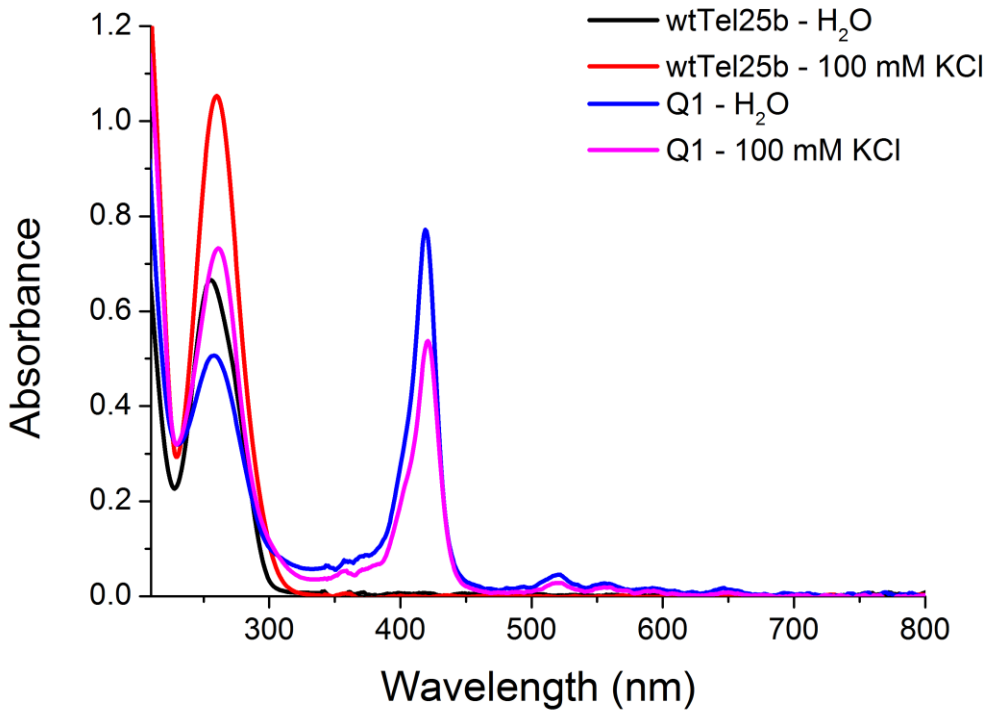


Figure 4.13 UV-Vis spectra of G-quadruplex DNA samples **wtTel25b** and **Q1** in first water (**wtTel25b** – black, **Q1** - blue) and then in 10 mM lithium phosphate buffer (pH = 7.4) supplemented with 100 mM KCl after slow annealing (**wtTel25b** – red, **Q1** – pink)

4.8 UV-Vis melting studies

To assess the stability of the porphyrin-modified G-quadruplexes, UV melting experiments were conducted. The melting temperature (T_m) is the generally accepted measurement of DNA secondary structure stability, the higher the T_m value, the more stable the DNA structure is. The T_m value is found by slowly heating and cooling the G-quadruplex whilst monitoring its absorbance at 295 nm. It is at this wavelength that the largest change in absorbance between the G-quadruplex form and the single strand form can be observed.[180] When performing melting studies on G-quadruplexes it is important that a very slow rate of change in temperature is used (0.2 ˚C / min was used here). Faster rates of change can lead to hysteresis between the melting and annealing curves produced due to the slow kinetics of quadruplex association and dissociation. Using a slow rate of change of temperature helps to reduce the hysteresis between the melting and annealing curves. An example of how the rate of change is related to hysteresis is shown in Figure 4.14, where the fluorescence melting profiles of the sequence F-TG$_4$T$_4$G$_4$TG$_4$T$_4$G$_4$T-Q (F = 6-amidohexylfluorescein and Q = C7 dabcyl) are shown with different rates of temperature change.[139] At a fast rate of melting (6 ˚C / min) there is a 13 ˚C difference in T_m between melting and annealing curves, whereas at a slower rate (0.2 ˚C / min) the melting and annealing curves coincide with a T_m of 70 ˚C.

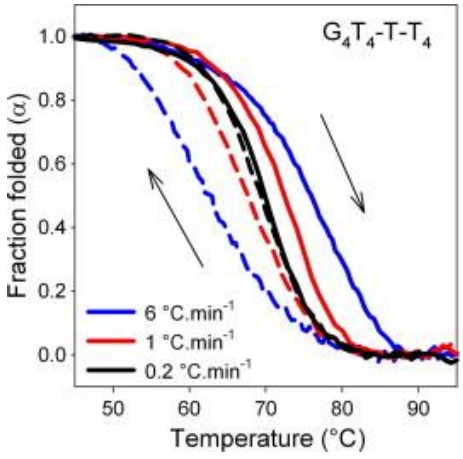


Figure 4.14 Effects of different rates of heating and cooling on the fluorescence melting profiles of F-TG$_4$T$_4$G$_4$TG$_4$T$_4$G$_4$T-Q[139]

When the melting studies were carried out, the absorbance at 420 nm (porphyrin Soret band) was also measured. This is important as we want to know if the porphyrin modification interacts at all with the G-quadruplex structure. Going from the UV-Vis spectrum that was taken before between the single strand form and the G-quadruplex

form of the porphyrin DNA, one would expect to see a hyperchromic shift in the absorbance at 420 nm as the melt progresses.

The oligonucleotides (5 μM) were dissolved in 10 mM lithium phosphate buffer solution (pH 7.4) supplemented with potassium chloride (10 mM). Lithium phosphate was chosen as the buffer for these experiments as lithium does not support the formation of G-quadruplexes. This ensures G-quadruplex formation is only dependent on the addition of sodium chloride or potassium chloride.[139] Samples were heated to 95 °C for 5 minutes and then allowed to cool to room temperature over a period of 4 hours. The absorbance was recorded at 295 nm and at 420 nm.

	Start Temp (°C)	End Temp (°C)	Rate (ΔT/min)
Ramp 1	25	80	10
Ramp 2	80	30	0.2
Ramp 3	30	80	0.2
Ramp 4	80	30	0.2
Ramp 5	30	80	0.2

Table 4.2 Melting temperature experiment parameters

Due to the limitations in how high the temperature could be set to with the cuvettes we were using, it was decided not to go any higher than 80 °C. To ensure that all of the G-quadruplex DNA had been denatured to the single strand form, a low concentration of potassium chloride (10 mM) was added to the lithium phosphate buffer solution. It has been observed in other melting studies of G-quadruplexes that increasing the concentration of potassium ions will increase the G-quadruplex stability, resulting in high T_m values. G-quadruplex formation will still occur at the concentration of potassium used.

The melting profile of a G-quadruplex differs from the melting profile of a DNA duplex. When duplex DNA is denatured there is an increase in absorbance whereas when quadruplex DNA is denatured there is a decrease in absorbance. In all of the oligonucleotides studied by this method, a decrease in absorbance was observed during melting which would imply the G-quadruplex structure is being denatured. In all cases the hypochromic shift at 295 nm was not large, but it was noticeably smaller for the porphyrin-modified oligonucleotides.

The 1st derivative of the melting curve of the unmodified quadruplex (**wtTel25b**) shows clearly that there is one transition present 42.2 °C (Figure 4.15). A

similar transition was observed in G-quadruplexes with one porphyrin modification . However, a second transition occurring at ~57 °C can also be seen in these modified sequences (Figure 4.16). This second transition is clearest in G-quadruplexes where the porphyrin modifications are present in the first loop (**Q1**, **Q2**). It should also be noted that the second transition is clearer in the dissociation than the association of the G-quadruplex in the strands with a single porphyrin modification. The melting curves were also recorded at 420 nm to observe potential changes in the porphyrin environment. As predicted, a large hyperchromic shift was observed in the absorbance at 420 nm. A single melting transition at ~55 °C can be seen in all samples containing a porphyrin. As this transition occurs at about the same temperature as the second transition seen at 295 nm, the second high temperature transition would suggest unstacking of the chromophore from the G-quartet concerted with a structural change in the DNA. The G-quadruplexes containing two porphyrins (**Q7**, **Q8**) show a gradual decrease in absorbance up to about 55 °C with no clear reflection point, and a steep drop similar to the mono-porphyrin sequences with a reflection point at 55.8 °C and 58.0 °C, respectively. This coincides quite well with the transition recorded at 420 nm (T_m 56.5 °C and 57.0 °C). The quadruplex structure is therefore stabilised by the covalently attached porphyrins by about 15 °C, which compares to PNA derived G-quadruplexes.[181]

DNA Strand	T_{m1} [°C]	T_{m2} [°C]	T_{mp} [°C]
wtTel25b	42.2	/	/
Q1	42.3	57.9	55.5
Q2	41.0	57.6	54.5
Q3	43.1	57.7	56.3
Q4	42.1	55.9	53.5
Q5	43.6	57.8	54.2
Q6	42.7	57.0	54.3
Q7	/	55.8	56.5
Q8	/	58.0	57.0

Table 4.3 Mean melting temperatures obtained from porphyrin G-quadruplex denaturing experiments. Thermal denaturation experiments were performed in 10 mM lithium phosphate solution (pH 7.4) supplemented with potassium chloride (10 mM)

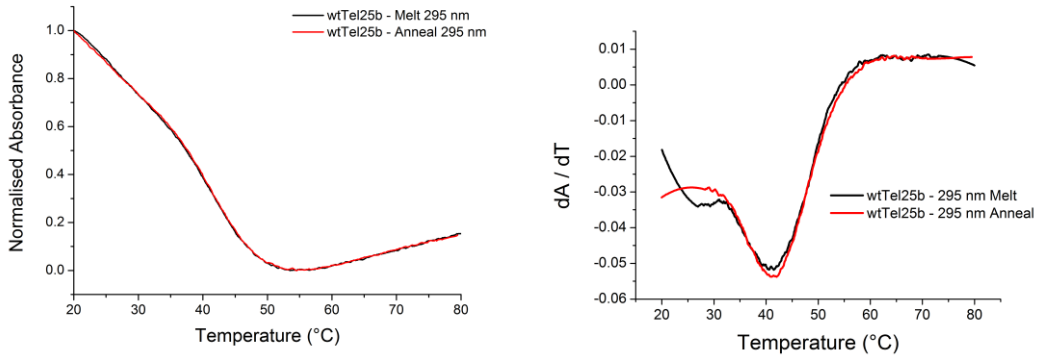


Figure 4.15 Melting curves and first derivative plots for **wtTel25b** in 10 mM lithium phosphate (pH 7.4) supplemented with potassium chloride (10 mM)

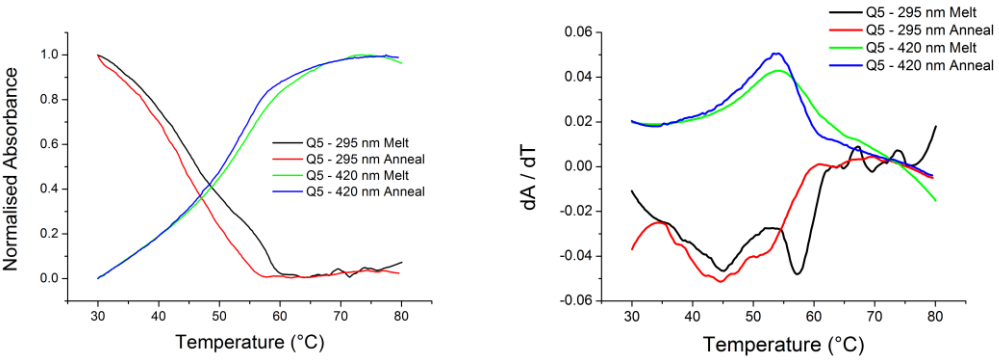


Figure 4.16 Melting curves and 1st derivative plots for **Q5** in 10 mM lithium phosphate (pH 7.4) supplemented with potassium chloride (10 mM)

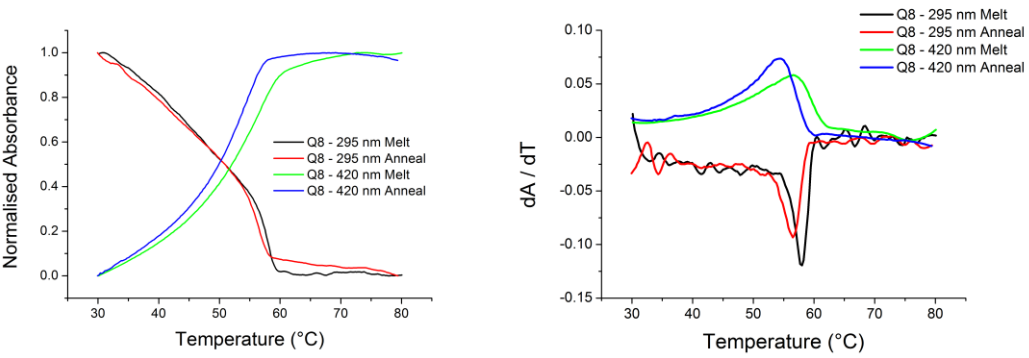


Figure 4.17 Melting curves and first derivative plots for **Q8** in 10 mM lithium phosphate (pH 7.4) supplemented with potassium chloride (10 mM)

4.9 Circular dichroism spectroscopy

We have seen that the covalently bound porphyrin modification can stabilise the G-quadruplex structure irrespective of where it is situated in the DNA strand. However, the amount of stabilisation did seem to be porphyrin position dependent. Any increase in stability could be linked to a change in the structure due to the presence of the porphyrin. Circular dichroism spectroscopy is a technique often used to assess G-quadruplex structure due to the chiral nature of DNA. As mentioned before intramolecular G-quadruplexes can adopt a variety of different topologies and this is certainly the case for the human telomere sequence. Circular dichroism (CD) spectroscopy has long been used to indicate whether the folding of a G-quadruplex is parallel or antiparallel. In general, a peak at 260 nm and a trough at 240 nm is descriptive of an all parallel G-quadruplex, whereas a trough at 260 nm and a peak at 295 nm describe antiparallel topologies;[182, 183] however, assigning structures on the basis of CD peaks at 260 and 295 nm is entirely empirical (and may not be always correct).

For the circular dichroism measurements, the oligonucleotides (5 μM) were dissolved in 10 mM lithium phosphate buffer solution (pH 7.4) supplemented with potassium chloride (100 mM). Samples were heated to 95 °C for 5 minutes and then allowed to cool to room temperature over a period of 4 hours. Data was collected at 20 °C.

CD spectroscopy of the natural and of the porphyrin-modified G-quadruplexes showed that there is generally a mixture of parallel and antiparallel topologies present in solution (Figure 4.18). The sequence **wtTel25b** is reported to adopt a 3 + 1 hybrid-2 structure in K^+ ions;[177, 178] this structure is characterised by a spectrum with a maximum at 295 nm, a shoulder around 270 nm, and a weak minimum at 240 nm. Our CD spectrum reproduces this result but the folding of G-quadruplexes is known to be affected by environmental conditions including sequence, salts, concentration, and rate of annealing to give a large number of accessible polymorphs.[132, 184] As the reported structure for **wtTel25b** was determined by NMR, a much higher concentration of G-quadruplex DNA was used for this measurement. To establish if this result was still applicable at the lower concentration at which our CD measurements were carried out, we recorded the CD spectrum of **wtTel25b** at varying concentrations (Figure 4.19). It was observed that the concentration used had no effect on the CD spectrum obtained and it would be fair to compare our data with the previously reported data. It should be noted that at the highest concentration (400 μM) saturation of the detector was observed at 245-265 nm. This data agrees with previous work in studying the effect of DNA concentration on G-quadruplex formation.[185] It is shown that

the concentration does have an effect on the folding of the G-quadruplex, but only at even higher concentrations than that were used here. At higher concentrations of DNA (200 mM), changes in G-quadruplex formation are attributed to weak interactions between G-quadruplexes rather than an overcrowding effect in the solution. Complexation of G-quadruplexes with TMPyP4 has shown to induce a conformational change towards a more anti-parallel structure,[177] displaying an increase in the peak at 295 nm in the CD spectrum. This is also observed when porphyrins are covalently attached to the telomere sequence (Figure 4.18). Depending on the site of attachment, the CD signature changes to varying degree, and the peak at 295 nm is generally larger than the peak at 260 nm in K^+ ions (apart from **Q8**, *vide infra*). There is a significant influence of the porphyrin on the quadruplex structure. In general, the modified sequences seem to fold into multiple topologies, in the presence of both K^+ or Na^+ (no quadruplex structure could be observed in Li^+). Two sequences stand out in terms of the CD signature: the spectra for **Q5** and **Q8** in K^+ ions show only one positive peak at 290 nm or at 260 nm, respectively. These spectra are characteristic of an antiparallel structure in the case of **Q5**, and of a parallel structure for **Q8**. In these specific sequences, the porphyrin does seem to preferentially stabilise one particular conformation. It was reported recently that also porphyrazines can induce structural changes and stabilise parallel G-quadruplexes.[186]

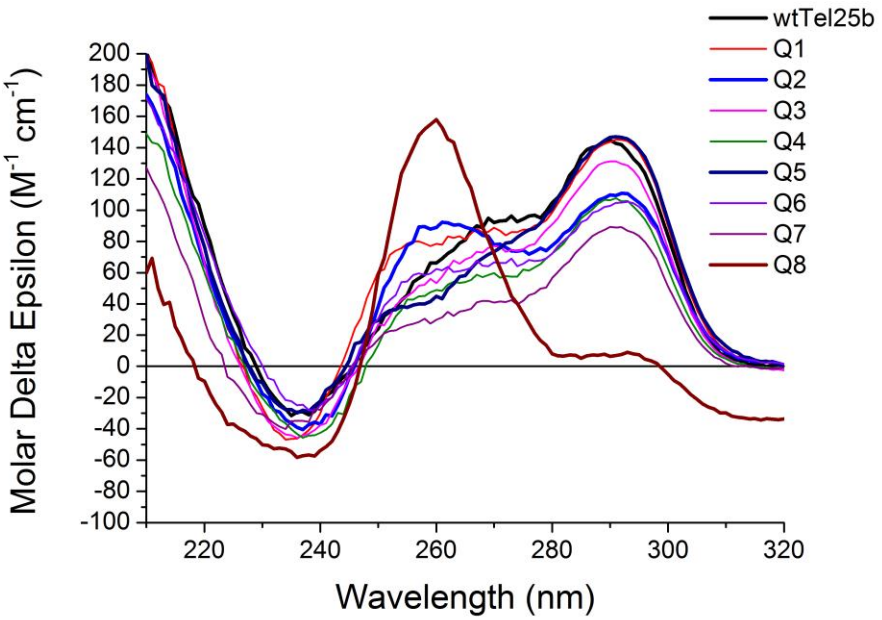


Figure 4.18 Circular dichroism spectra of porphyrin-modified DNA in 10 mM lithium phosphate buffer solution (pH = 7.4) supplemented with 100 mM potassium chloride – DNA region. The unmodified DNA (**wtTel25b**) and strands **Q2**, **Q5**, and **Q8** are highlighted (see text for discussion)

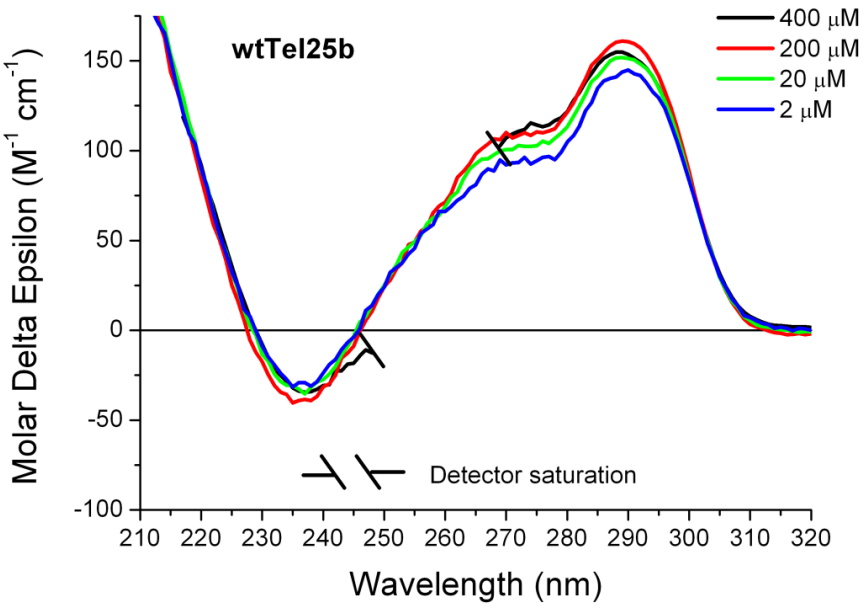


Figure 4.19 Circular dichroism spectra of **wtTel25b** at various concentrations

The CD spectra of the Soret band of the porphyrin at 420 nm show a variety of signatures for the different sequences (Figure 4.20), indicating that the porphyrins are placed into very different environments within the G-quadruplex. Positive and negative Cotton effects are observed in both K^+ and Na^+ with peaks around 443 nm and 418 nm. Most sequences give a positive Cotton effect (− → + with increasing wavelength) except for **Q2**, **Q6** and **Q8**, where a negative Cotton effect (+ → −) is observed. The intensity of these peaks and troughs are generally very similar (Δε around +50 and −50 $M^{-1}cm^{-1}$), except for the sequences **Q2** and **Q8**. For the latter sequences the intensities of the peaks are much larger than those for the other sequences. For **Q8**, the large intensities (Δε = +560 and −424 $M^{-1}cm^{-1}$) could be explained by strong exciton coupling between the chromophores held at close distance in the quadruplex structure (*vide infra*). In the case of the single porphyrin-modified sequence **Q2**, this unusually large Cotton effect (Δε = +375 and −221 $M^{-1}cm^{-1}$) might arise from intermolecular interactions such as stacking of the porphyrins. This has been observed under conditions of high concentration of porphyrin-DNA[82] or at high salt concentration,[101] but these do not apply here. Concentration dependent CD spectroscopy showed that the spectra (porphyrin region) do not change between 1 μM and 50 μM (Figure 4.21), in contrast to an eleven-porphyrin-modified DNA in the same concentration range[82]. Therefore, the unusually large signals must be due to an induced chiral effect and do not arise from intermolecular interactions, but the exact source remains unclear. Interestingly, **Q7** does not show intense CD signals despite having two porphyrins as in **Q8**. The UV region of the CD spectrum is also very different for **Q7** compared to **Q8**, and shows weaker signals and a signature resembling mostly antiparallel structure. It is likely that the porphyrins are in very different orientations/environments in the two quadruplexes, and that in **Q7** there are no strong excitonic couplings. The fact that the porphyrin in **Q2** is attached to the second thymidine in the first loop, where also **Q8** carries a porphyrin, seems to indicate that this position could be special, and that the chirality is transferred onto the porphyrin with an unusually high efficiency.

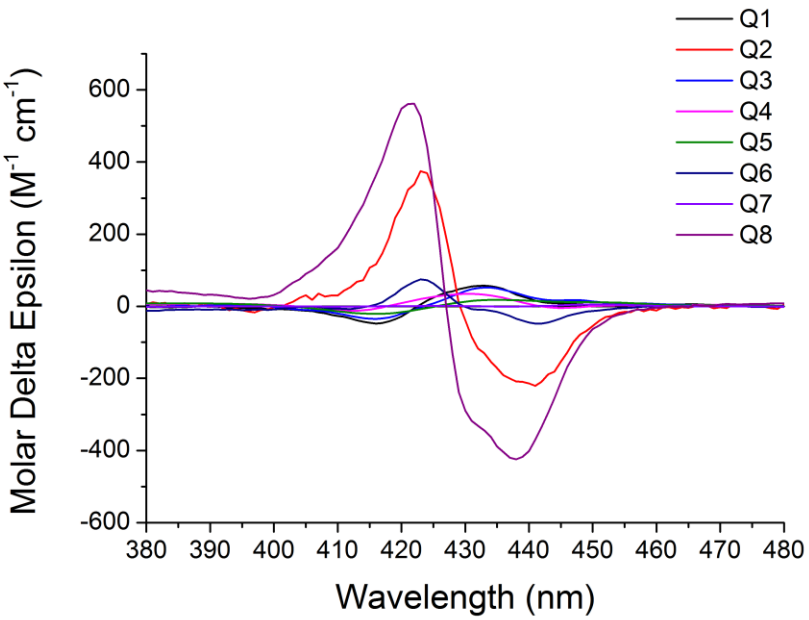


Figure 4.20 Circular dichroism spectra of porphyrin-modified DNA in 10 mM lithium phosphate buffer solution (pH = 7.4) supplemented with 100 mM potassium chloride – Porphyrin Soret band region

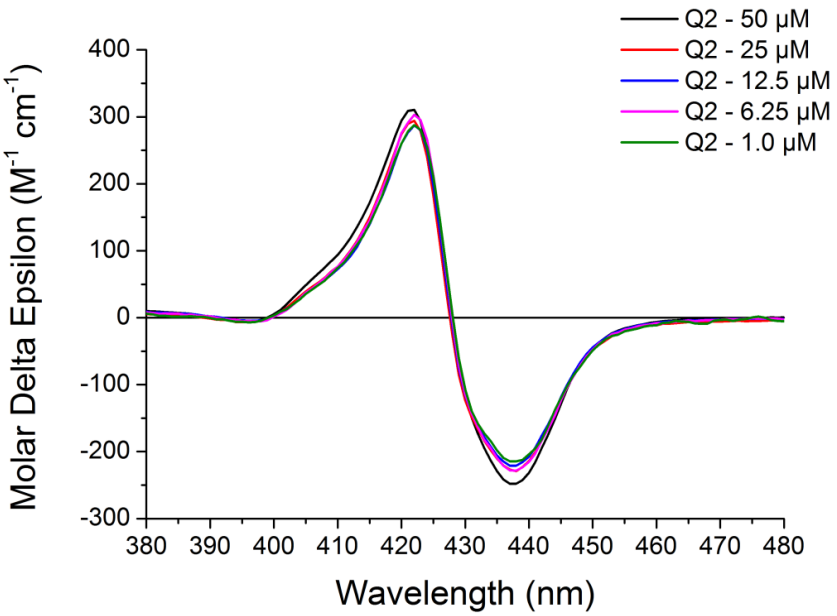


Figure 4.21 Circular dichroism spectra of Q2 at various concentrations

The CD spectra of the porphyrin-modified G-quadruplexes were also measured in the presence of sodium cations. It is known that the type of cation present can have a large effect on the G-quadruplex conformation as described earlier (Figure 4.9). As

expected the CD spectrum for **wtTel25b** in Na^+ shows a different type of G-quadruplex formation is present to that observed in K^+ (Figure 4.22). In this instance, two peak maxima can be observed at 294 nm and 260 nm and a minimum at 236 nm. This type of spectra is observed for the majority of the porphyrin-modified G-quadruplexes with the peak at 260 nm being larger than the one present at 295 nm in most cases. This suggests more parallel character is present in these G-quadruplexes. As circular dichroism spectra show a superposition of all the spectra contributing structures in the solution, one might expect there is a mixture of parallel and hybrid (3+1) G-quadruplexes present. It should be noted that in G-quadruplexes **Q2**, **Q3** and **Q4** an additional shoulder can be observed at 270 nm and the molar delta epsilon values are generally low. This effect could be because G-quadruplex formation in these samples is not particularly strong in Na^+. Based on its CD spectrum, **Q7** appears to be folding into a parallel G-quadruplex with a maximum at 260 nm and a minimum at 240 nm, which is different to the largely anti-parallel folding observed with K^+. However, for both CD spectra in the Soret band region, very little circular dichroism observed. One may have predicted a change in spectrum considering the change in G-quadruplex folding indicated from the DNA region. It should be noted though that the two porphyrin modifications in **Q7** are placed near the terminal ends of the oligonucleotide so the distance for exciton coupling between porphyrins is probably too large hence the small signal.

As observed previously for the G-quadruplexes with K^+, a variety of porphyrin environments were observed in the CD spectra of the Soret band in the presence of Na^+ (Figure 4.23). As before, most sequences exhibit a positive Cotton effect with **Q2**, **Q6** and **Q8** being the exceptions that show a negative Cotton effect. Once again the CD spectra for **Q2** and **Q8** stand out from the rest with larger bisignate signals observed, most likely due to similar reasons as stated for the spectra with potassium. Perhaps the most surprising result was that the Soret band CD spectra did not vary much regardless of whether potassium or sodium was added. This is despite the suggestion in the CD spectra of the DNA region of some of the samples, that there is a change in G-quadruplex conformation. Despite these changes in G-quadruplex structure, we cannot exactly determine how the porphyrin behaves in relation to this. However, we can speculate that the orientation of the porphyrin-modified thymidine within the loop regions of the G-quadruplex is consistently the same despite any change to the overall G-quadruplex structure. This would also imply that the porphyrin is in a similar position with regards to the G-quadruplex. To further study this, one would have to carry out NMR spectroscopy, X-ray crystallography or molecular modelling of these porphyrin-modified G-quadruplexes.

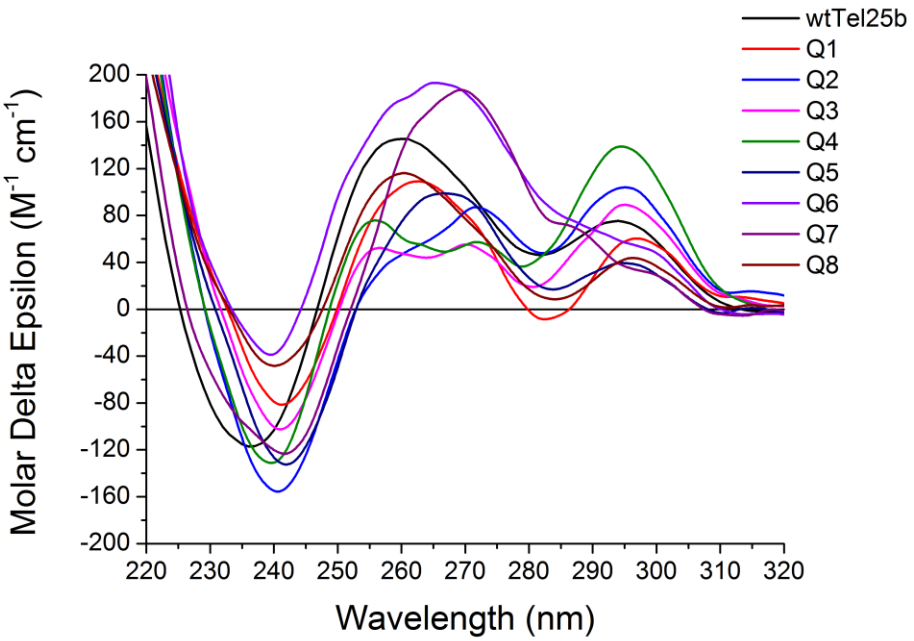


Figure 4.22 Circular dichroism spectra of porphyrin-modified DNA in 10 mM lithium phosphate buffer (pH = 7.4) supplemented with 100 mM sodium chloride – DNA region

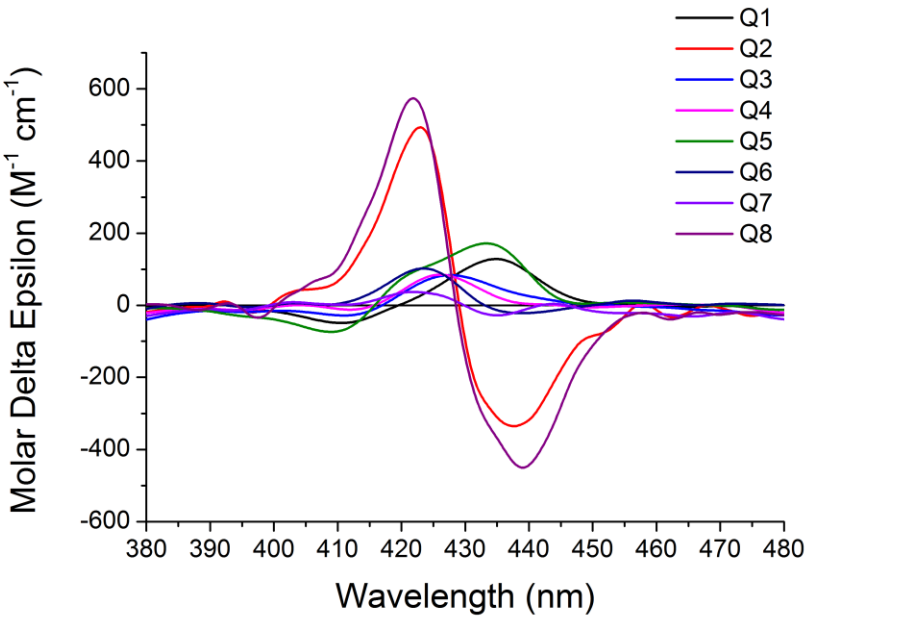


Figure 4.23 Circular dichroism spectra of porphyrin-modified DNA in 10 mM lithium phosphate buffer (pH = 7.4) supplemented with 100 mM sodium chloride – Soret band region

4.10 Circular dichroism melting studies

The CD spectra of **wtTel25b** and porphyrin-modified G-quadruplexes **Q5** and **Q8** were recorded at temperatures between 20˚C and 80 ˚C to gain an insight into the structural changes during DNA melting. For the circular dichroism measurements the oligonucleotides (5 μM) were dissolved in 10 mM lithium phosphate (pH 7.4) and potassium chloride added (100 mM). Samples were heated to 95 ˚C for 5 minutes and then allowed to cool to room temperature over a period of 4 hours. The temperature was allowed to equilibrate for 5 minutes before the spectrum was recorded.

For the unmodified G-quadruplex, **wtTel25b**, an anti-parallel structure is observed. As the temperature increases, the height of the peak at 295 nm decreases (Figure 4.24). This is an effect of the G-quadruplex structure disassociating to form a single strand with no secondary structure. A T_m of 44.2 ˚C was measured by first fitting the data points taken at 292 nm and fitting them with a Boltzmann fitting (Origin 8.6) and then taking the first derivative of this curve (Figure 4.26). T_m values obtained using this method should not be directly compared with the values obtained from the UV-Vis melting studies due to the different temperature ramp and conditions used.

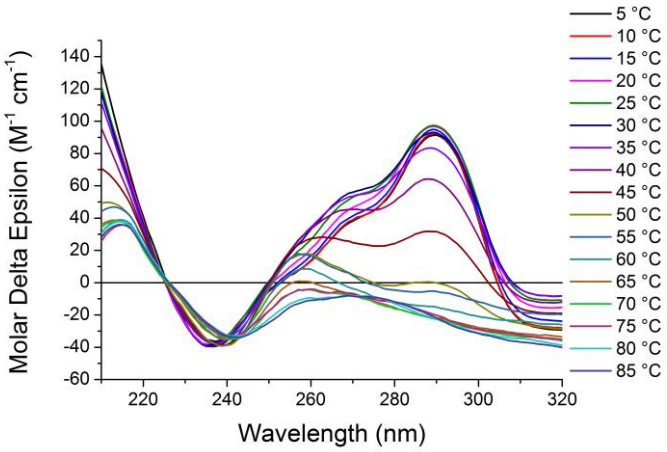


Figure 4.24 Variable temperature circular dichroism spectra of unmodified G-quadruplex **wtTel25b** in 10 mM lithium phosphate (pH 7.4) supplemented with potassium chloride (100 mM)

For **Q5** (Figure 4.25) which is highly antiparallel due to its larger peak at 295 nm, the changes in molar delta epsilon were plotted at 292 nm. A clear melting transition is observed between 40 ˚C and 50 ˚C, although no secondary melting transition is observed. However with the data intervals being 10 ˚C, it is unsurprising that this transition is not observed. At 240 nm the trough tends to become more negative and a

small red shift is observed. The shoulder at 260 nm is also affected with a decrease in molar delta epsilon and a larger red shift to 275 nm observed. These changes are indicative of the G-quadruplex structure being denatured to the random single strand form of DNA. The behaviour of **Q8** (Figure 4.25) was slightly different to that of **Q5** due to the different starting conformation of the G-quadruplex. There was a small decrease in molar delta epsilon for the small peak present at 295 nm, which hints at a melting transition between 40 ˚C to 50 ˚C. As observed before with **Q5**, the small trough at 240 nm becomes more negative and goes through a small red shift, which is indicative of a denaturing of the G-quadruplex to the single stranded form.

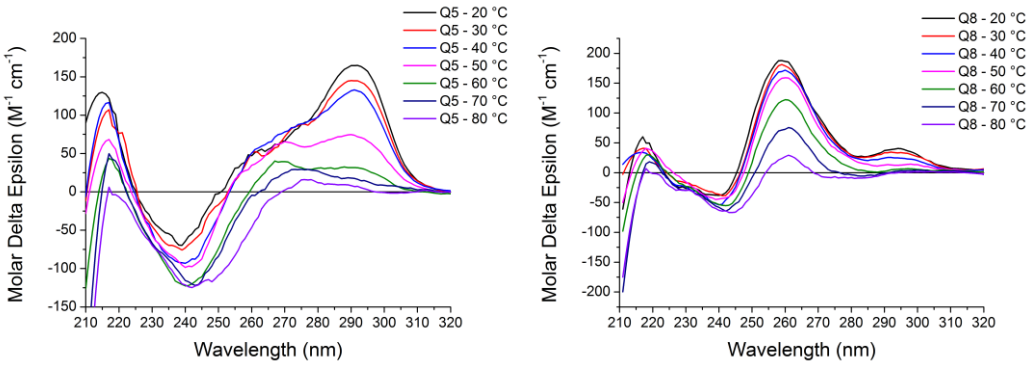


Figure 4.25 Variable temperature circular dichroism spectra of porphyrin-modified G-quadruplexes **Q5** and **Q8** in 10 mM lithium phosphate (pH 7.4) supplemented with potassium chloride (100 mM)

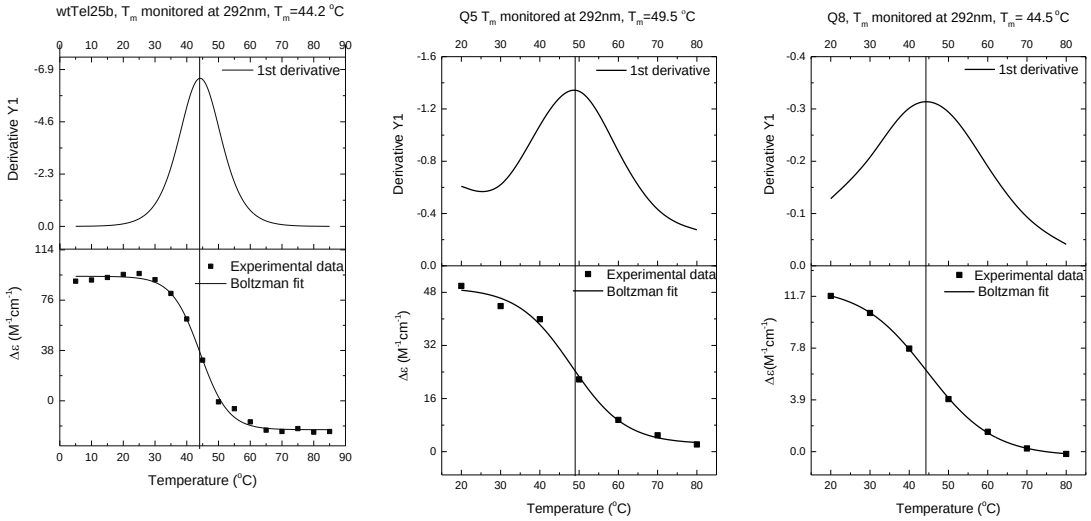


Figure 4.26 CD-Temperature plots with Boltzmann fitting and first derivative (top) to calculate the melting temperature (T_m)

T_m values for **Q5** and **Q8** were taken from data points collected at 292 nm (Figure 4.26). Both values were lower than would be expected but as mentioned before, comparisons with the UV-Vis melting studies are not possible. However, this data can be qualitatively analysed to try and gain an insight into how the G-quadruplex disassociates to a random single stranded form of DNA. Data obtained from thermal CD studies were subjected to principle component analysis (PCA), with the help of Giuliano Siligardi and Rohanah Hussain (Diamond Light Source, Didcot, UK). The eigenvalues of the correlation matrix suggest that more than 98% is made by two components with the remaining components less than 2 % (see table inset in figures). The PCA analysis indicates that for **wtTel25b** and **Q5** qualitatively the two main components behave fairly similarly on increasing the temperature from 20° to 80° C. In particular, the coefficient of component 1 (PC1) remained constant whereas that of component 2 (PC2) changed from negative to positive in a qualitatively similar manner. This was also mirrored by the similar spread of the directions of the eigenvectors (cyan, green, orange and red colours). **Q8** showed differences mainly for the coefficient of PC2 where the change as a function of raised temperature did not follow that of the other two samples. The coefficient of PC2 at 40° C did not decrease in absolute value but increased instead (green colour). At 60° C, the coefficient of PC2 was negligible for **Q8**, increasing for **Q5** and reaching the relative maximum observed for **wtTel25b**, which did not further increase with heating up to 80 °C.

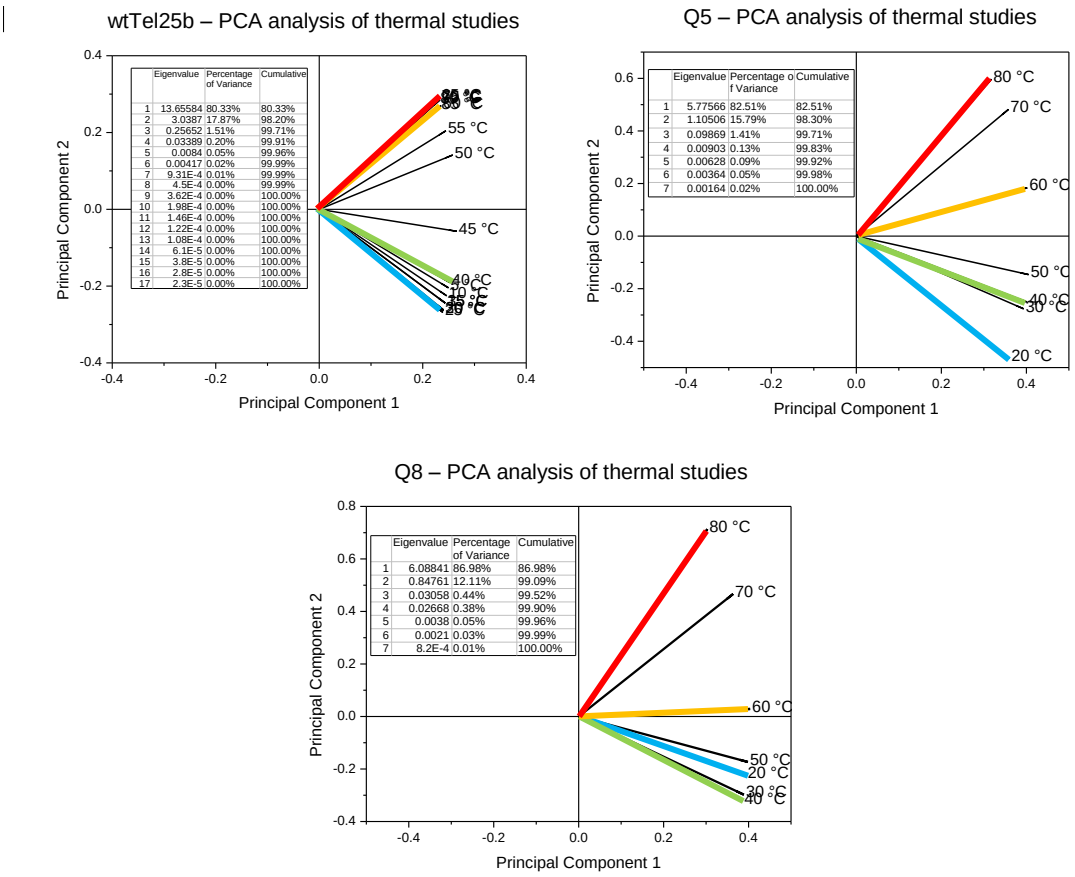


Figure 4.27 Loading plot of PCA presented as vectors of thermal studies of **wtTel25b**, **Q5** and **Q8** (cyan, green, orange and red represent the directions of the eigenvectors at 20˚, 40˚, 60˚ and 80˚ C respectively to illustrate, at glace, the spread of the directions)

4.11 ^{1}H NMR spectroscopy of porphyrin G-quadruplexes

Further structural characterisation of the porphyrin G-quadruplexes was carried out using ^{1}H NMR spectroscopy. It was decided to start at a basic level with this study and first use this technique to further confirm that the porphyrin-modified DNA is folding into a G-quadruplex structure. This can be achieved by observing the guanine NH_1 imino protons, which have a characteristic shift when hydrogen bonded. By knowing how many imino protons are involved in the hydrogen bonding present in G-quadruplexes, one can easily predict how many peaks you would see if one topology is present. For example, if there are three G-quartets in the G-quadruplex, then there should be twelve imino proton signals. If there are a mixture of topologies present the number of observed will greatly increase making any further characterisation by NMR spectroscopy difficult.

All oligonucleotides (100 μM) were dissolved in potassium phosphate buffer (0.1 M, pH = 7.4). Samples were heated to 95 °C and cooled slowly over a period of 4 hours to room temperature to allow the quadruplexes to anneal. The NMR experiment for the unmodified G-quadruplex, **wtTel25b** was run overnight.

The ^{1}H NMR spectrum of **wtTel25b** showed evidence of G-quartet formation (Figure 4.28). In the region between 10 and 12 ppm, the imino protons are easily distinguished. From this spectrum it looks as though there is one dominant structure formed in solution with other structures present in smaller amounts. As there are 12 imino protons in the G-quadruplex DNA, there should be 12 individual signals if only one structure is present. In the ^{1}H NMR spectrum, there are 12 larger peaks between 10 and 12 ppm. There were many smaller peaks present which hint that there are other structures present. This particular G-quadruplex forming sequence has been used previously for analysis by NMR and it was found that many topologies were present but the major conformation present was a hybrid-2 G-quadruplex.[159] The data collected was similar to what had been seen before, but small differences could be attributed to a difference in DNA concentration as no exact value was stated in the previous work (0.1 – 2.5 mM DNA was used). Due to the polymorphic nature of **wtTel25b**, it was decided not to pursue any further characterisation of this particular DNA sequence.

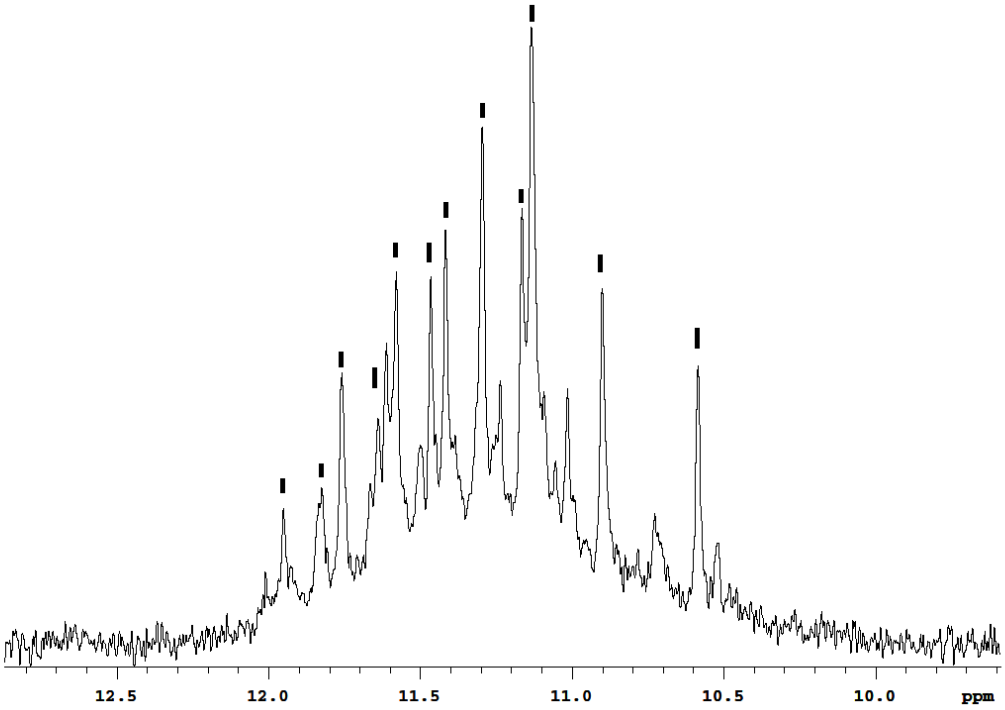


Figure 4.28 1D imino proton NMR spectrum of the unmodified G-quadruplex forming oligonucleotide **wtTel25b**. Major peaks are indicated

^{1}H NMR spectra of all the porphyrin-modified G-quadruplexes proved to be less clear than the unmodified analogue. In the imino proton region there is evidence of G-quartet formation, some fine detail can be made out within a small hump between 10 and 12 ppm but overall the data is poor (Figure 4.29). Surprisingly, it was not possible to detect the porphyrin proton signals from the spectra of all the porphyrin-modified samples. This is surprising as one would have predicted to at least observe the β-pyrrole protons of the porphyrin in the spectra, which are present between 8 and 10 ppm. From the circular dichroism spectroscopy experiments described earlier, we expected to see a mixture of different topologies with the possibility of one major conformation present. Although the absence of any signature peaks from the porphyrin cannot be explained, the poor quality of the NMR results could be linked to the method of purification used to prepare the samples. All porphyrin G-quadruplex DNA strands as mentioned earlier were purified using fluorous affinity chromatography. Whilst this method is reasonably effective at purifying these modified oligonucleotides it is not good enough for future NMR analysis. A better solution would have been to further purify the samples using reversed phase HPLC but considering the amount of sample needed for NMR analysis and the difficulty of purifying this amount of modified oligonucleotide using this method it was decided not to carry on any further with this analysis. Another possibility is that the broad signals may be linked to

an aggregation of the oligonucleotides due to the high concentrations used in these experiments.

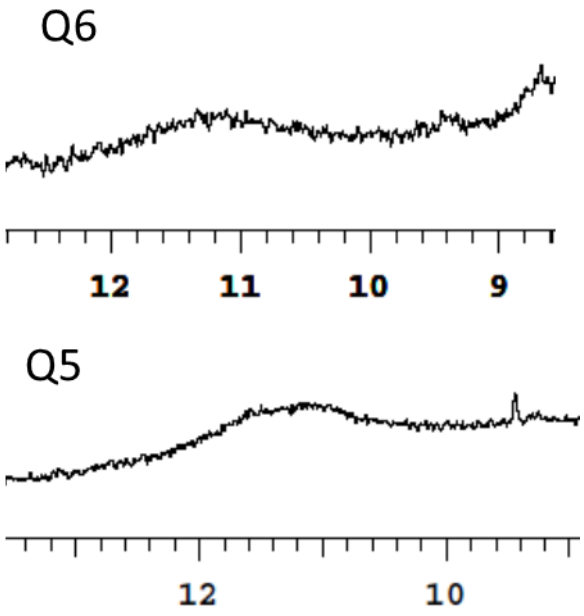


Figure 4.29 1D imino proton NMR spectrum of quadruplex-forming oligonucleotides **Q6** and **Q5**

4.12 Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel electrophoresis was used to gain a greater understanding of the G-quadruplex folding and to find out what role the bound porphyrin plays in interacting with it. Early non-denaturing gel electrophoresis experiments were carried out using 16 % polyacrylamide gels, which were run in TBE buffer that had been supplemented with 20 mM KCl. We found that there were issues with visualising the DNA once the gel was run. We originally thought that the covalently bound porphyrin would fluoresce enough so that we could see it on the gel reader, but no DNA was observed. Staining with ethidium bromide or with SYBR Gold were both unsuccessful in visualising the DNA. This surprising result could possibly be due to the DNA streaking heavily on the gel to an extent that it was not observed once the gel had been run. Streaking tends to occur more in non-denaturing gels and it has been observed before within our research group that porphyrin DNA has a large tendency to streak during electrophoresis.

To try and reduce the problem of streaking, it was decided to load a smaller amount of porphyrin-modified DNA on the gels at a time. This could be achieved by labelling the oligonucleotides with ^{32}P ATP. Radiolabeling of porphyrin DNA was carried out with help from Annie Cardew and Dr David Rusling of the Fox research group, University of Southampton. Samples were labelled with ^{32}P and then purified by denaturing polyacrylamide gel electrophoresis. On a 30 mL polyacrylamide gel, the unmodified G-quadruplex showed only one major band present. Porphyrin-modified G-quadruplexes **Q2** and **Q8** both showed 3 major bands. All major bands were collected and purified. Denaturing and non-denaturing gel electrophoresis was then carried out on the 7 samples. Samples for the denaturing gel were prepared in distilled water. The gel (Figure 4.30) shows that **Q2** and **Q8** both clearly have 3 distinct species present. It is impossible to tell which one of these, if any, is the original porphyrin-DNA quadruplex. The extra species present are most likely degradation products of the porphyrin DNA.

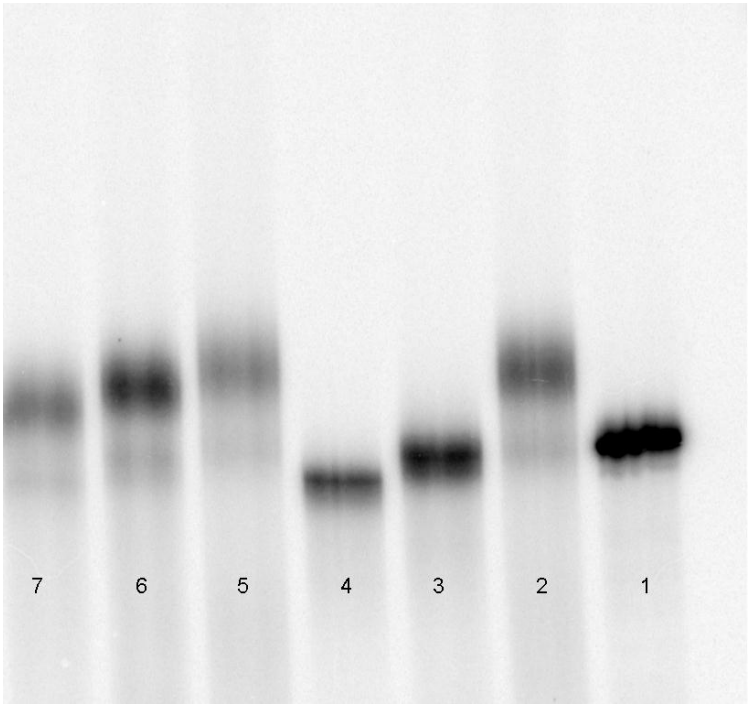


Figure 4.30 Denaturing gel of 7 DNA samples containing one porphyrin modification

Lane 1 = wtTel25b

Lanes 2-4 are the three products from the purification of Q2

Lanes 5-7 are the three products from the purification of Q8

A non-denaturing gel using 30 mL polyacrylamide gel was carried out on all 7 samples. The samples were divided into two, one half was dissolved in 10 mM lithium phosphate with 100 mM potassium chloride and the other half dissolved in distilled water. The samples suspended in the salt solution were heated to 95 °C and allowed to cool slowly to room temperature prior to the gel electrophoresis experiment. The porphyrin DNA samples were found to streak on the non-denaturing polyacrylamide gel, but this was to be expected. However, this made most bands undistinguishable from one another. It was hoped from this experiment to see whether folding the DNA into a quadruplex structure would change the mobility of the DNA. However, due to the excessive streaking and possibly poor quality of samples used we could gain nothing from this analysis. The experiment was repeated with porphyrin DNA samples that had been resynthesised and purified but similar issues were observed after radiolabelling. We decided to stop any further analysis using radiolabelling and run the porphyrin DNA on a shorter gel. By running the electrophoresis experiments with a shorter gel length it was hoped that the streaking would be minimised whilst still being able to separate different products. We loaded a larger amount of porphyrin DNA on the gels than usual as to take advantage of the fluorescent nature of porphyrin for visualisation. Denaturing PAGE was first carried out using this new technique

(Figure 4.31 and Figure 4.32), followed by non-denaturing PAGE to attempt to distinguish between the G-quadruplex structure and the single stranded form with no secondary structure (Figure 4.33 and Figure 4.34).

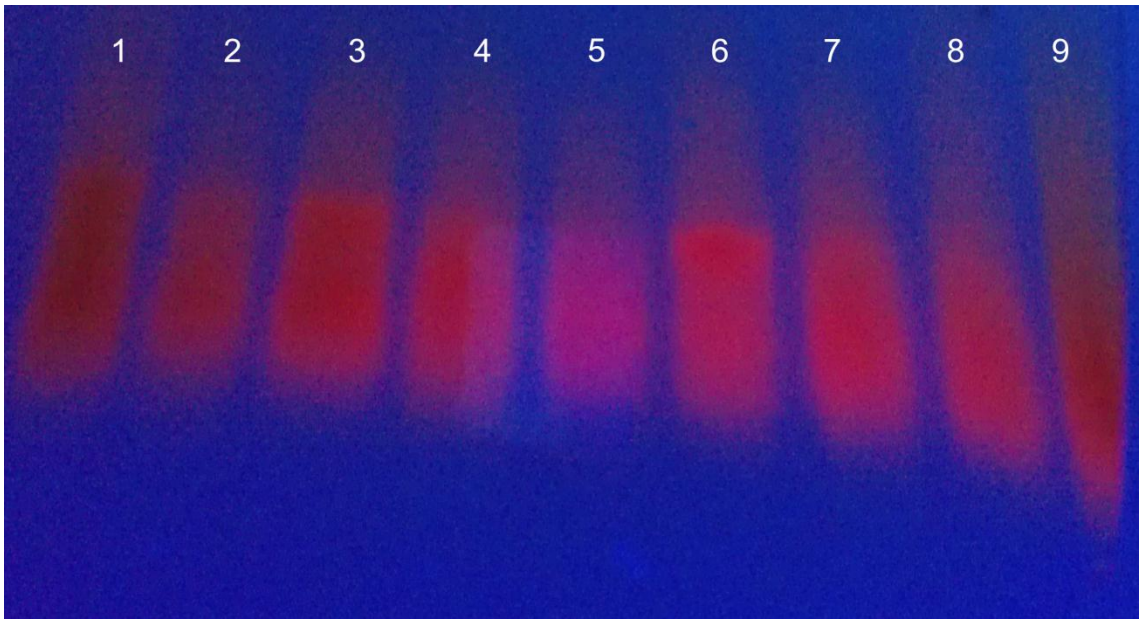


Figure 4.31 Denaturing PAGE visualised at 365 nm.

Lane 1 = **Q6** Lane 2 = **Q5** Lane 3 = **Q4** Lane 4 = **Q3**
Lane 5 = **Q2** Lane 6 = **Q1** Lane 7 = **Q7** Lane 8 = **Q8**

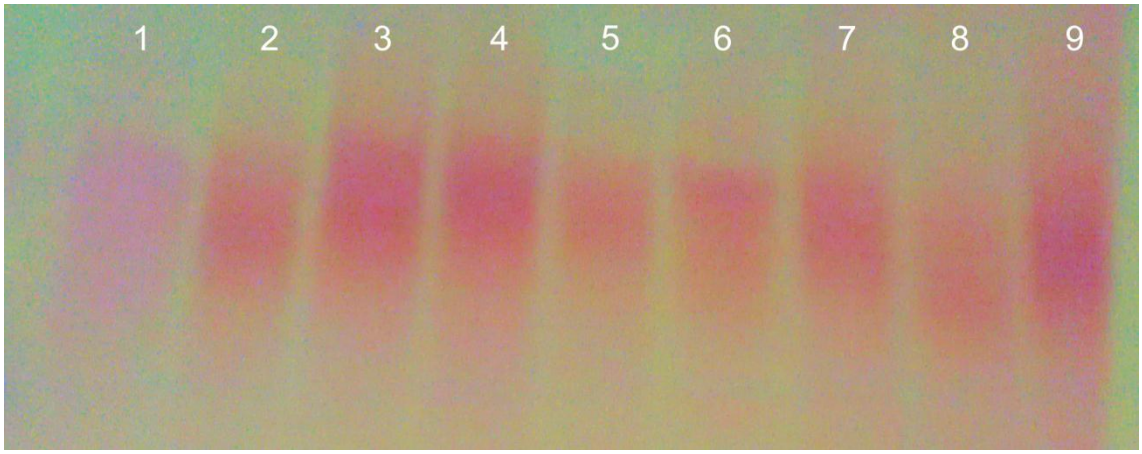


Figure 4.32 Denaturing PAGE visualised at 254 nm.

Lane 1 = **wtTel25b** Lane 2 = **Q6** Lane 3 = **Q5** Lane 4 = **Q4**
Lane 5 = **Q3** Lane 6 = **Q2** Lane 7 = **Q1** Lane 8 = **Q7**
Lane 9 = **Q8**

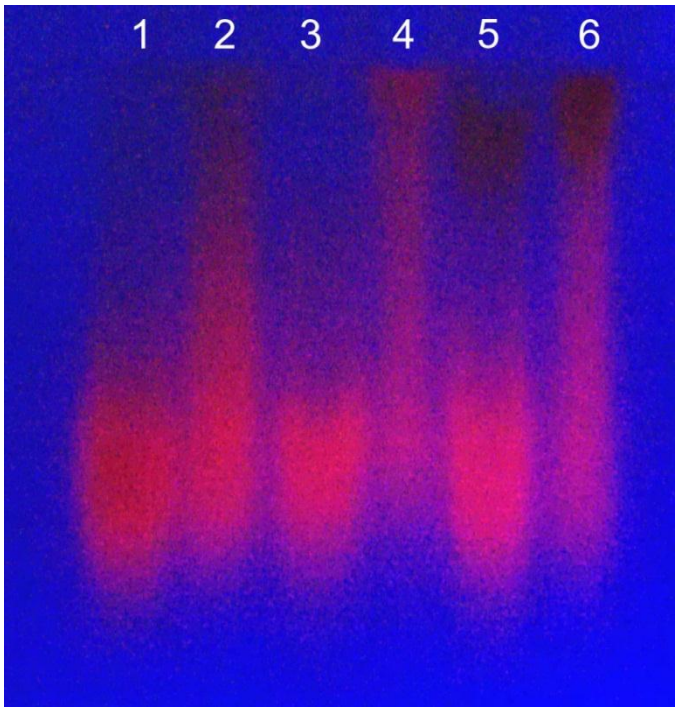


Figure 4.33 Native PAGE visualised at 365 nm

Lane 1 = **Q5** denatured form Lane 2 = **Q5** G-quadruplex form

Lane 3 = **Q3** denatured form Lane 4 = **Q3** G-quadruplex form

Lane 5 = **Q7** denatured form Lane 6 = **Q7** G-quadruplex form

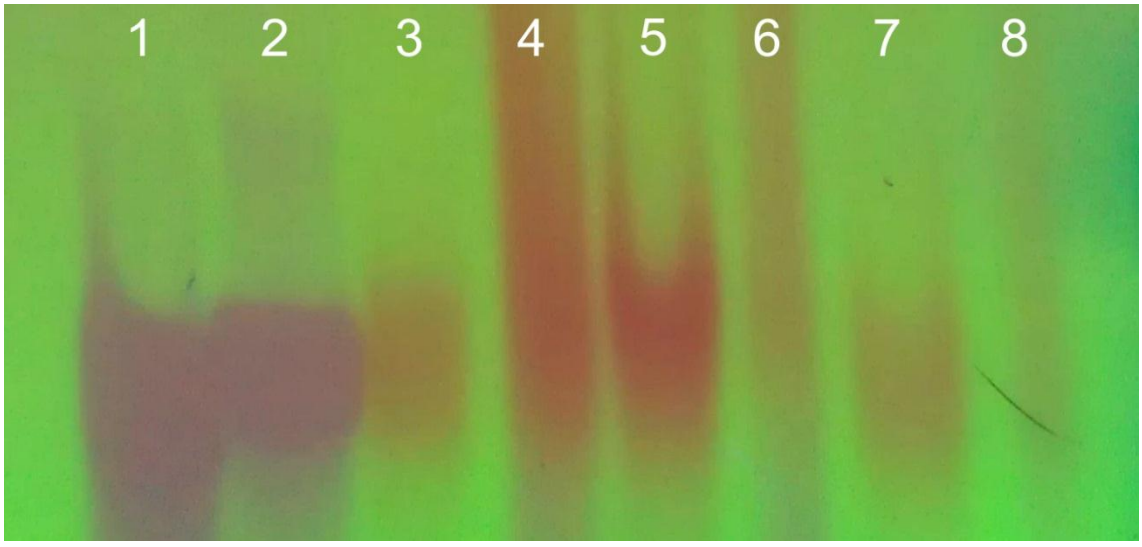


Figure 4.34 Native PAGE visualised at 254 nm

Lane 1 = **wtTel25b** denatured form Lane 2 = **wtTel25b** G-quadruplex form

Lane 3 = **Q3** denatured form Lane 4 = **Q3** G-quadruplex form

Lane 5 = **Q5** denatured form Lane 6 = **Q5** G-quadruplex form

Lane 7 = **Q7** denatured form Lane 8 = **Q7** G-quadruplex form

As observed with previous attempts at electrophoresis with porphyrin DNA, streaking was present in all samples. The non-denaturing PAGE experiments seem to hint that there is a difference in gel mobility between the denatured form of the G-quadruplex and the annealed form. However, due to the streaky nature of bands it cannot be conclusively said that this is the case. The hydrophobic nature of the porphyrin is clearly the cause of the streaking on the gels and this is preventing any further electrophoresis analysis. One could add charged functional groups to the porphyrin in order to attempt to improve the mobility of the porphyrin on the gel so more conclusive results could be obtained. However, for the time being it was decided not to carry on any further with the PAGE experiments.

4.13 Molecular modelling

In order to gain insight into the interactions of the porphyrins with the G-quadruplex structure, we performed molecular dynamics simulations with the help of Peter Várnai of the University of Sussex. Since all complexes show either mixtures of structures or ambiguous topologies when supplemented with potassium, only **Q8** could be assumed to have a definite starting geometry. The CD spectrum suggests that **Q8** folds into a parallel quadruplex, the structure of which is known from X-ray diffraction (XRD) analysis.[155] Based on this structure (pdb entry 1KF1), a molecular dynamics simulation study (using the Amber11 package, for details see 8.5.6)[187] with the two porphyrins attached to the thymidines showed a stable structure with both porphyrins closely associated with the terminal quartets, irrespective of the starting geometry at the sites of attachment (Figure 4.35). Since porphyrins do not have a suitable size to fully stack to a G-quartet through π-π interactions, the role of the *meso*-phenyl substituents appear essential. The hydrophobic nature of the porphyrins may also be responsible for the observed conformation. Interestingly, the porphyrin modifications appear to rigidify the corresponding loop structure. The structure is different to the reported XRD structure of a non-covalent $TMPy_4P$-quadruplex structure, where either groove binding or binding to the loops without direct ligand interactions with the G-tetrad was observed,[123] but it is not too dissimilar to the reported NMR structure of a TMPyP4 complex with a five-guanine-tract G-quadruplex structure from the human MYC promoter.[188] In the latter structure, the porphyrin was found to interact closely with the G-tetrad. However, the covalent attachment of the porphyrins clearly has a very different effect, and particularly sequence **Q8** stands out in terms of stable parallel structure and strong excitonic coupling, as seen by CD spectroscopy.

Further simulations carried out by Peter Várnai showed there are two other possible ways the porphyrin can interact with the G-quadruplex. In Figure 4.36, only

one porphyrin associates with a terminal G-tetrad. The second porphyrin interacts with the first porphyrin through π-π interactions to create a stacked porphyrin arrangement. It was also observed that it is possible for the two porphyrins not to interact at all with the terminal G-tetrads. Instead the two porphyrins closely associate with each other using π-π interactions away from the G-quadruplex (Figure 4.37). It could be suggested that as the porphyrins in these two examples closely associated with each other, it is more likely that they will exhibit excitonic coupling between each other than they would if they were to associate with the two terminal ends of the G-quadruplex. This excitionic coupling was manifested in the porphyrin circular dichroism spectrum for **Q8**, so it can be argued that the structures shown in Figure 4.36 and Figure 4.37 are more likely to be present than in Figure 4.35.

A simulation of the dissociation of the G-quadruplex with a single porphyrin modification (**Q5**) from an increase in temperature was also carried out. In this simulation the porphyrin begins associated with a terminal G-tetrad. As the G-quadruplex begins to dissociate, it is the porphyrin free terminal G-tetrad that dissociates first. The other two G-tetrads remain intact with the porphyrin still associated with the terminal tetrads until they all dissociate at the same time. It can be argued this simulation is similar to what was described from the UV melting studies of these G-quadruplexes.

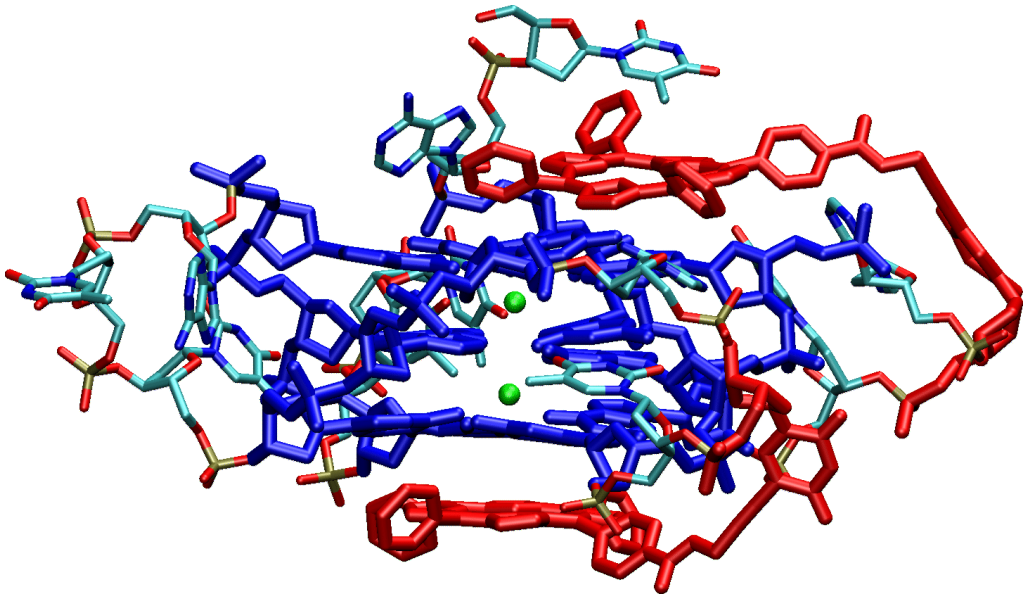


Figure 4.35 Putative structure of **Q8** from molecular dynamics simulations – Porphyrins interacting with the two terminal G-tetrads

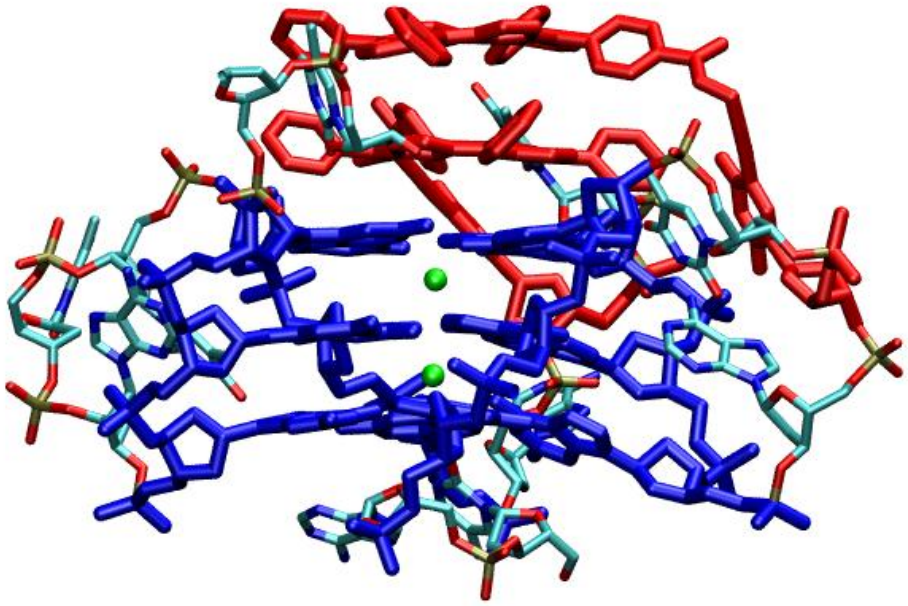


Figure 4.36 Putative structure of **Q8** from molecular dynamics simulations – Two porphyrins found on top of one terminal G-tetrad

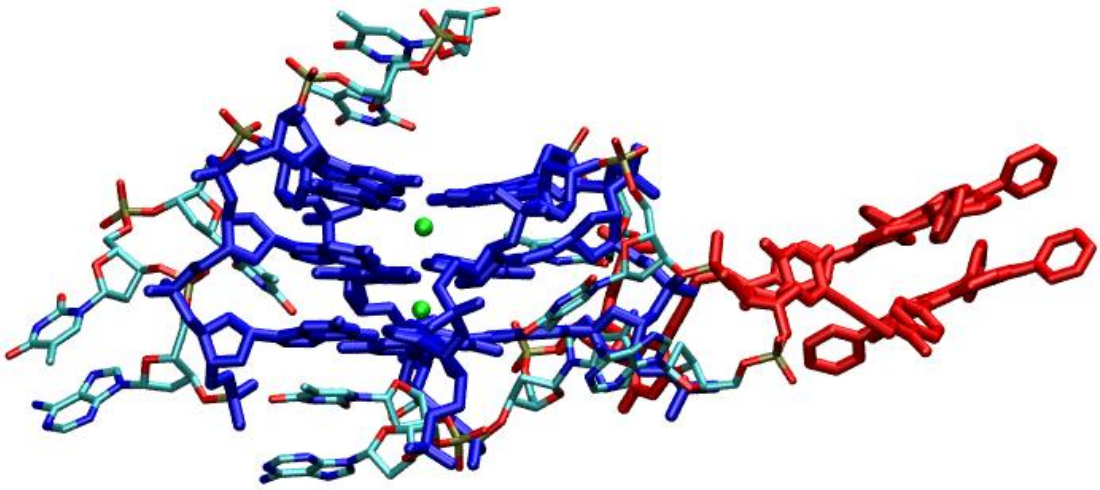


Figure 4.37 Putative structure of **Q8** from molecular dynamics simulations – Two porphyrins interacting with each other, but not with the G-quadruplex

4.14 Competition experiments

Switching between G-quadruplex and duplex structures has been used as the basis for DNA nanomachines.[189-192] One of the first examples of a nanomachine designed on this principle was achieved by Mergny.[193] He presented a cycle that could be considered the equivalent to a two-stroke engine (Figure 4.38).

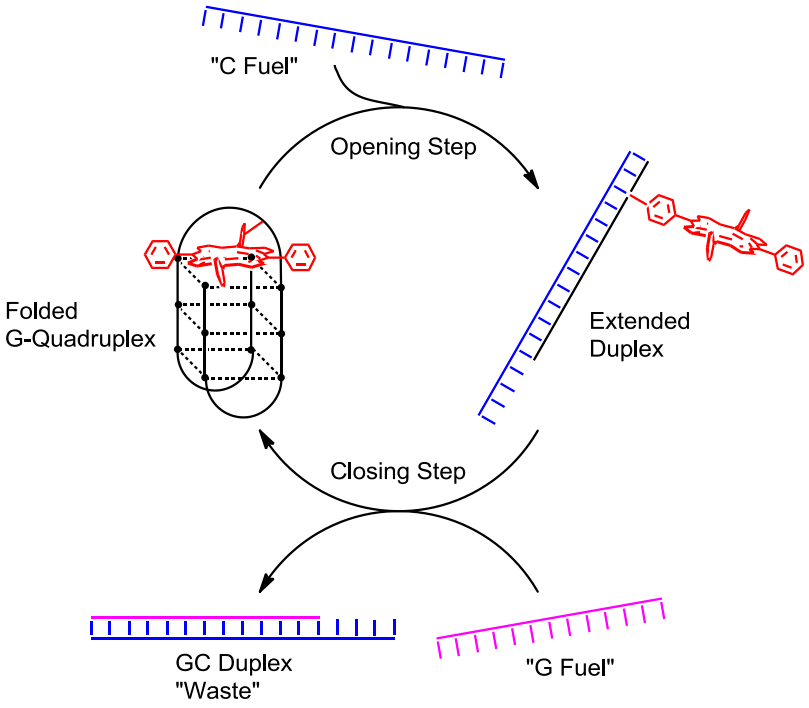


Figure 4.38 A schematic describing the process of switching between G-quadruplex and duplex structures

There are two states present in the cycle, a closed and open state. The machine starts off in the closed state as the G-quadruplex form. To get to the open state the complementary C-rich strand is added to the system. In many cases the DNA duplex will be formed preferentially to the G-quadruplex.[194] Once the duplex is formed the machine is said to be in the open state. To return to the closed state of the G-quadruplex, the complementary of the C-rich strand is added. Obviously, the complementary of the C-rich strand will be the same as the original G-quadruplex. This is undesirable as there is no driving force for the duplex to separate and form a new duplex with the recently added G-rich strand.

To solve this problem and improve the efficiency of the nanomachine, some mismatches were added to the C-rich strand. The amount of mismatches added was designed so that the duplex is still formed over the G-quadruplex. If too many

mismatches are included, the DNA duplex may not be very stable and no switching between the open and closed states will occur. Mismatches were positioned so that the complementary G-fuel needed to close the system could not form a G-quadruplex with itself. Short complementary flanking sequences were also added to the C and G fuels to increase the stability of the GC waste duplex over the duplex formed in the opening step.

F-5'-GGG TTA GGG TTA GGG TTA GGG-3'-T
C Fuel 3' - TGCAAT CC**G** AAT C**G**C AAT C**A**C CAA TCC C-5'
G Fuel 5' - ACGTTA GG**C** TTA G**C**G TTA G**T**G TTA-3'

Figure 4.39 DNA used in the Mergny duplex-G-quadruplex nanomachine.[193] Bases in bold indicate mismatches from the original G-quadruplex forming sequence. Underlined regions indicate the complementary flanking sequences. The starting G-quadruplex was labelled with fluorescein (F) and tetramethylrhodamine (T) so opening and closing of the G-quadruplex could be monitored by FRET

Taking inspiration from this idea, we wanted to see if we could apply the duplex-G-quadruplex switching to the porphyrin G-quadruplexes. However, to start with a more fundamental experiment needed to be carried out. This was to assess the stability of the G-quadruplex when in the presence of the complementary C-rich strand **C9**. It has been observed previously that addition of the C-rich strand to a G-quadruplex will induce switching to the duplex.[194] The easiest way of observing this switching is to add a pair of fluorescent dyes that will interact via fluorescence resonance energy transfer (FRET) to the terminal ends of the G-quadruplex. When in the closed state of the G-quadruplex, the donor dye will be in close proximity to the acceptor dye, so FRET is observed. When in the open state of the DNA duplex the donor and acceptor dyes will be far apart and no FRET will be observed. However, this method is not possible for the system we are studying; therefore CD spectroscopy was used to monitor structural changes of the quadruplex upon addition of the complementary DNA.

Switching was observed in the unmodified G-quadruplex **wtTel25b** as expected. However, the addition of **C9** to strands **Q2**, **Q5**, and **Q8** did not induce spontaneous disruption of the G-quadruplex structure at ambient temperature, neither in 100 mM K^+ nor in 100 mM Na^+, where the stability of the G-quadruplex is lower. This is contrary to what has been observed previously with unmodified G-quadruplexes when in the presence of the complementary C rich strand.[194] When the mixtures are heated to 80 °C and cooled slowly to room temperature over 4 h, duplex formation could be observed

for strands **Q2**, **Q5**, and **Q8** in both K^+ and Na^+ through a change in the porphyrin CD signature (Figure 4.40). This is generally manifested in a decrease in ellipticity, but is not accompanied by a large change of the CD signature. Unfolding is therefore prohibited by the porphyrins at ambient temperature where the activation barrier is too high, but is possible at high temperature upon thermal disruption of the G-quadruplexes. This is consistent with the large stabilisation effect seen by the melting curves. From these observations it was clear that the porphyrin-modified G-quadruplexes would not be suitable for a duplex-G-quadruplex switching machine.

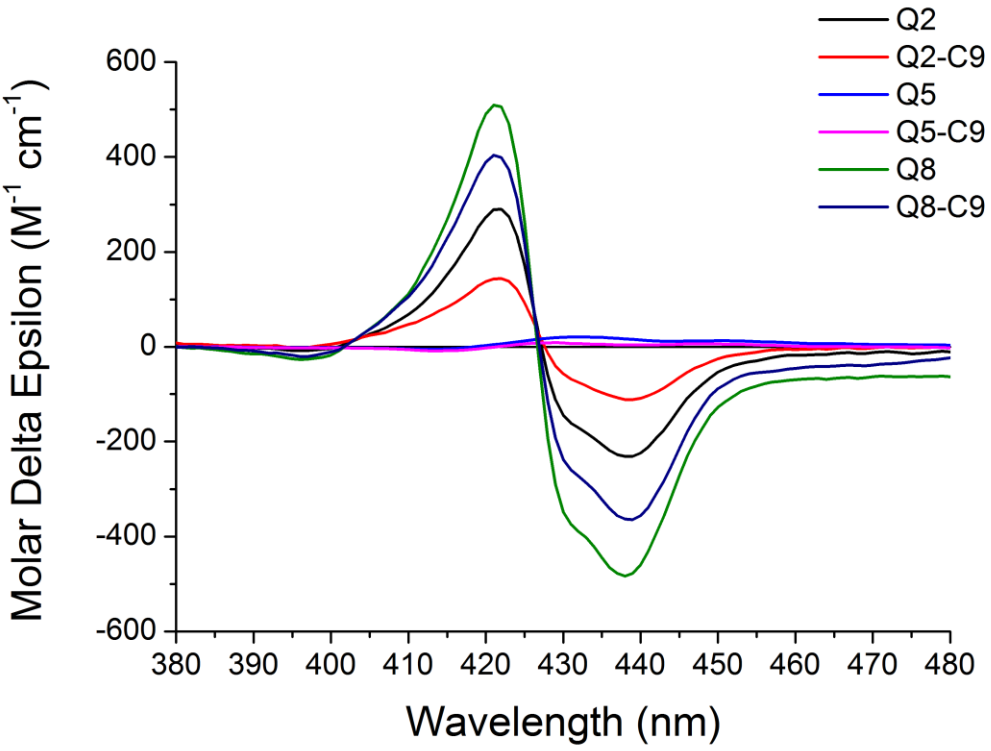


Figure 4.40 CD spectra of the strands **Q2**, **Q5** and **Q8** in quadruplex form, and as duplex **Q2-C9**, **Q5-C9** and **Q8-C9** after annealing in 100 mM NaCl ([DNA] = 2.5 μM)

4.15 Conclusions and remarks

In summary, we have described the first synthesis of G-quadruplexes with covalently bound porphyrins which greatly stabilise the G-quadruplex structure. The spectroscopic measurements reveal that the covalently bound porphyrin does not promote the formation of a single quadruplex structure apart from one specific two-porphyrin sequence (**Q8**), which folds into a parallel quadruplex. Due to the polymorphic nature of the strands studied in this work, detailed structural analysis

proved to be difficult. However, computer simulation has helped to suggest a number of possible ways of how these systems behave. Experimental data is needed to confirm these predictions though. NMR spectroscopy and X-ray diffraction would be ideal techniques that would give us a better understanding of what is actually happening.

From the discussion in the previous chapter, the synthesis and purification of large quantities of porphyrin-modified DNA is difficult due to the highly hydrophobic modification and the unstable nature of the porphyrin-modified phosphoramidite used in the synthesis. To add to this issue the polymorphic nature of G-quadruplex DNA makes it difficult to purify by HPLC due to the formation multiple secondary structures during purification. This issue however can be slightly alleviated by increasing the column temperature to encourage denaturing of any DNA secondary structures. These purification problems make any potential studies using NMR and X-ray diffraction difficult as large quantities of high purity oligonucleotide are needed for successful results. It is also important to bear in mind that if high polymorphism is present, the NMR spectra will be a superposition of all of these topologies and crystallisation will be difficult to achieve. There are G-quadruplex sequences based on the human telomeric sequence that are known to give one topology and have been crystallised. It could be worth considering for any future work that using these sequences with the porphyrins covalently bound may make study by NMR and X-ray crystallography easier. However, any changes in the DNA sequence used will make comparisons with the work carried out so far impossible.

All work so far with the porphyrin DNA G-quadruplexes has just been concerned with their structure and stability. Our next goal is to use these G-quadruplexes in more applied situations. One area that could be considered is the use of porphyrin-modified G-quadruplexes as a tool for sensing proteins that bind to G-quadruplex structures. This would involve synthesising a selection of G-quadruplex forming DNA that represents the many possible sections of DNA present in the human genome that can fold into G-quadruplex structures. It has been observed that some of these G-quadruplex forming sequences are found in the promoter regions of some genes. The theory is that these genes can be switched on and off with the formation of a G-quadruplex within these regions, and that proteins play a role with this process. By taking advantage of the fluorescent properties of the porphyrin we should be able to use porphyrin-modified G-quadruplexes to sense the proteins that bind to these sequences. There is also the possibility that the porphyrin G-quadruplex could be used to inhibit specific proteins and thus effect the transcription of this gene.

5 Porphyrin-modified LNA

5.1 Introduction

Locked nucleic acids (LNA) are a recent development in oligonucleotide chemistry. The idea behind locked nucleic acids is to create a modification that is structurally rigid and also increases the binding affinity to a complementary oligonucleotide. What gives LNA these structural properties is the methylene bridge that links the 2'-oxygen and the 4'-carbon of the ribose sugar (Figure 5.1), which locks it in an RNA-like, C3'-endo conformation.[36] Adding this small modification to the nucleoside has great implications on how binding takes place with a complementary oligonucleotide and the structure and overall stabilisation of the newly formed DNA duplex.

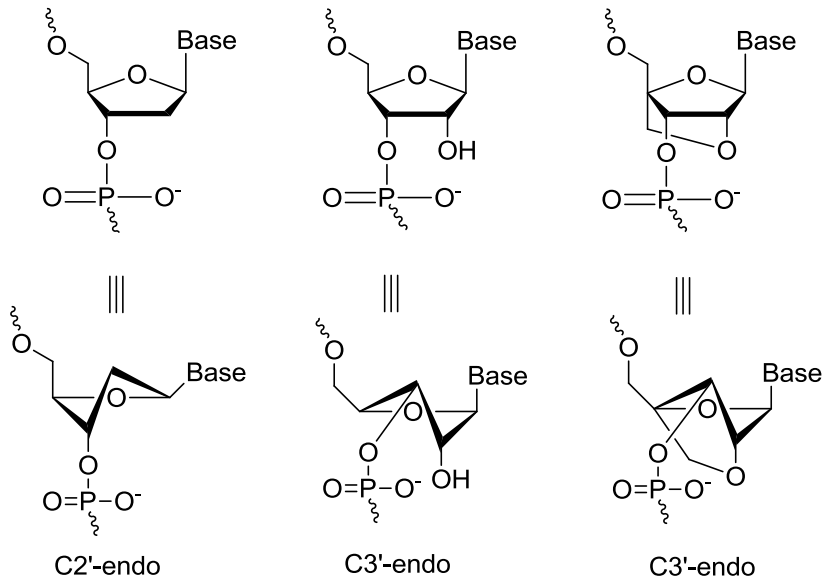
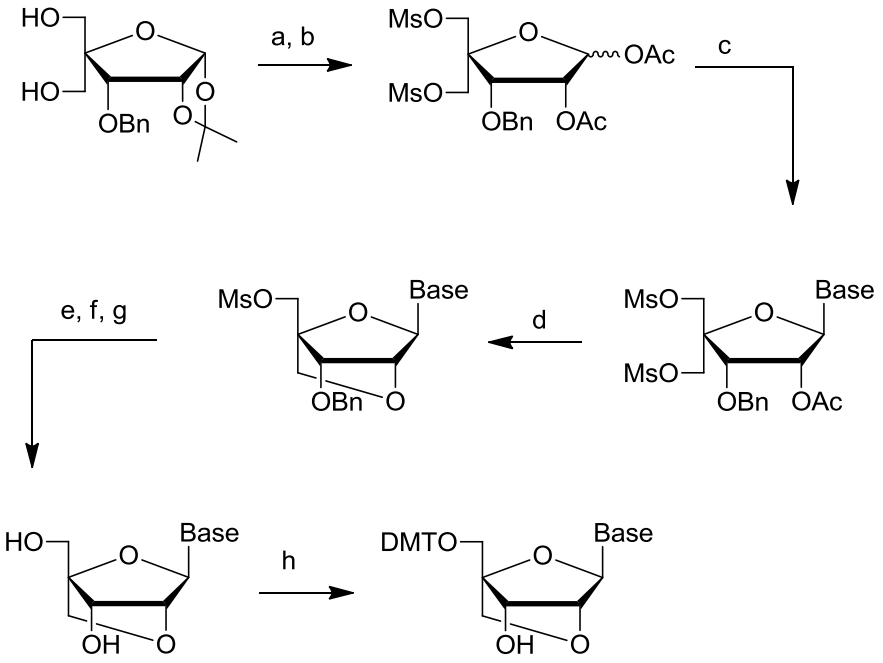


Figure 5.1 Structures and furanose conformation of canonical and LNA nucleotides

5.1.1 LNA synthesis

Synthesis of locked nucleic acids is far from a facile process. Initial developments of this synthesis used either a linear strategy using nucleosides as the starting material[195] or a convergent strategy where an appropriate modified glycosyl donor is synthesised and then coupled to the nucleobase to give the modified nucleoside.[196] This has since been improved to a seven step synthetic route that shows improved yields (Scheme 5.1).[197] O3'-phosphitylation of the 3'-hydroxyl functional group has found to be relatively simple, as with standard DNA nucleosides, despite the presence

of the methylene bridge. LNA phosphoramidites can then be incorporated into oligonucleotides using standard solid phase synthesis, although coupling times that are extended to 6 minutes are recommended for 98-99 % yields.[196]



(a) MsCl, pyridine; (b) (i) 80 % TFA; (ii) Ac_2O, pyridine;
(c) nucleobase, BSA, TMSOTf, MeCN or 1,2-dichloroethane;
(d) NaOH or LiOH, THF or 1,4-dioxane; (e) NaOBz, DMF;
(f) 20 % $Pd(OH)_2/C$, HCO_2NH_4, MeOH;
(g) NH_4OH; (h) DMTCl, pyridine;

Scheme 5.1 Synthetic route for LNA nucleosides

When LNA is incorporated into oligonucleotides it has been observed that there is an increase in stability when hybridised to the complementary DNA or RNA strand.[198] In some cases an increase in melting temperature between 2-10 °C per LNA modification has been achieved. However, it should also be noted that as the number of LNA modifications increases in the oligonucleotide, their overall stabilising effect per base is reduced. Another general observation is that LNA will form a more stable duplex with a complementary RNA strand than a complementary DNA strand.[196, 198] LNA though, forms the most stable interactions when it is paired to a complementary LNA strand. One example of this is the all-LNA duplex that is nine base pairs in length which has a T_m value of greater than 93 °C compared to the same unmodified DNA duplex which has a T_m of 29 °C.[199] This level of DNA duplex stabilisation is quite simply outstanding and it is little wonder as to why the LNA modification is of great interest in oligonucleotide research.

Due to its locked rigid conformation, incorporation of the LNA nucleoside into an oligonucleotide can have significant structural implications when hybridised to the complementary strand. Using NMR spectroscopy as a tool to study the structure of a DNA duplex containing an LNA modification, it has been observed that a C3'-*endo* conformation can be induced onto the neighbouring nucleotides from the LNA modification.[200] NMR analysis also showed that single stranded DNA containing the LNA modification are to some degree preorganised for duplex formation.[200, 201] This preorganisation causes a decrease in entropy, which results in stronger binding to complementary DNA and RNA. Another consequence of the C3'-*endo* conformation found in the LNA:DNA duplex is that there are improvements in the efficiency of base stacking around the LNA modification which leads to increased duplex stability.

5.1.2 LNA uses in biotechnology

There is an interest in applying LNA technology into antisense therapy. The theory behind this is that when the genetic sequence of a particular gene is known to be the cause of a particular disease, it is possible to synthesise an oligonucleotide that will bind to the messenger RNA (mRNA) produced by the gene and silence it. This works as mRNA has to be single stranded for it to be translated. LNA is particularly good for this due to its high binding affinity to specific RNA and resistance to degradation from 3'-exonucleases.[202] The RNA:LNA duplex formed can also silence the desired gene by inducing cleavage of the RNA by RNase H, an enzyme that degrades RNA:DNA hybrids.[203, 204]

The Hrdlicka group, Moscow, USA, have progressed research in this area using C5-functionalised LNA.[205] C5-functionalised pyrimidine DNA building blocks have gathered attention due to their ability to accommodate functional entities in the major groove of nucleic acid duplexes. It was thought that these modifications would be able block any potential enzyme degradation of the oligonucleotide. Coupled with the high affinity for RNA from the LNA modification a new class of effective antisense oligonucleotides could be synthesised.

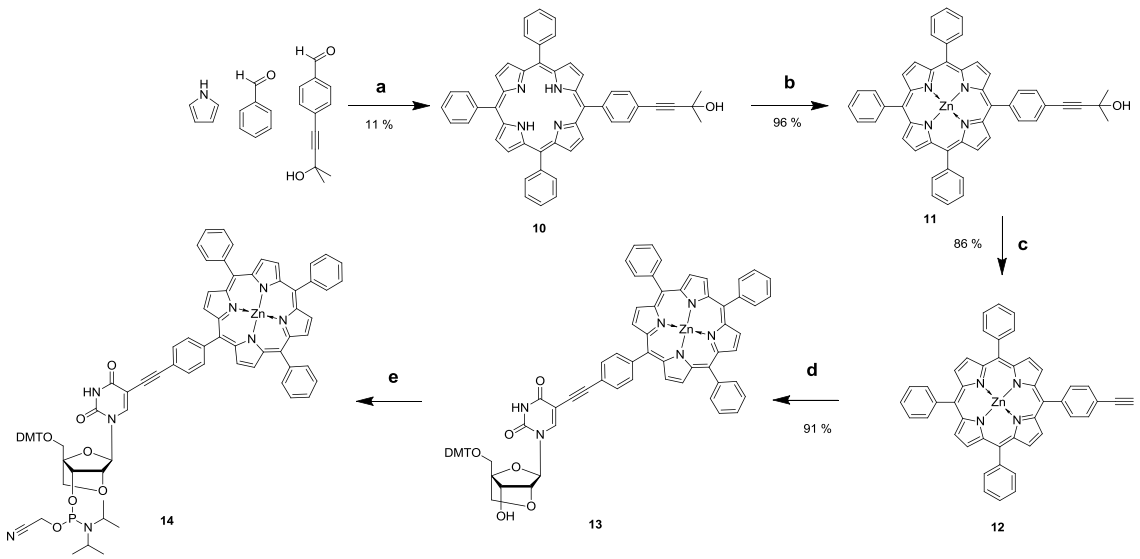
Perhaps more closely related to the general work carried out by the Stulz group, the Hrdlicka group also have an interest in adding hydrophobic fluorescent molecules to DNA, in this case pyrene. The fluorescent nature of the pyrene makes it a good choice for designing nucleic acid detection probes. Reliable detection of specific nucleic acid targets is an integral component of a wide range of applications including pathogen detection, quantification of DNA/RNA in real-time PCR and the detection of single nucleotide polymorphisms (SNPs).[206, 207]

5.2 Aim

The next variation on porphyrin DNA that we have investigated is porphyrin LNA. Locked nucleic acids have the ability to increase the stability of DNA:DNA duplexes and also DNA:RNA duplexes. After observations from previous work by ThaoNguyen Nguyen and Ashley Brewer, both of the Stulz group, it is known that covalently attaching porphyrins to one strand of a DNA duplex can destabilise the duplex structure. We propose that synthesising porphyrin LNA monomers and incorporating them into DNA, it may be possible to offer greater stability to porphyrin duplex DNA. The increased stability of porphyrin DNA could be useful for future applications with porphyrin DNA such as antisense therapy and SNP detection as well as improving on what has already been achieved within the Stulz group.

5.3 Synthesis of porphyrin-modified LNA

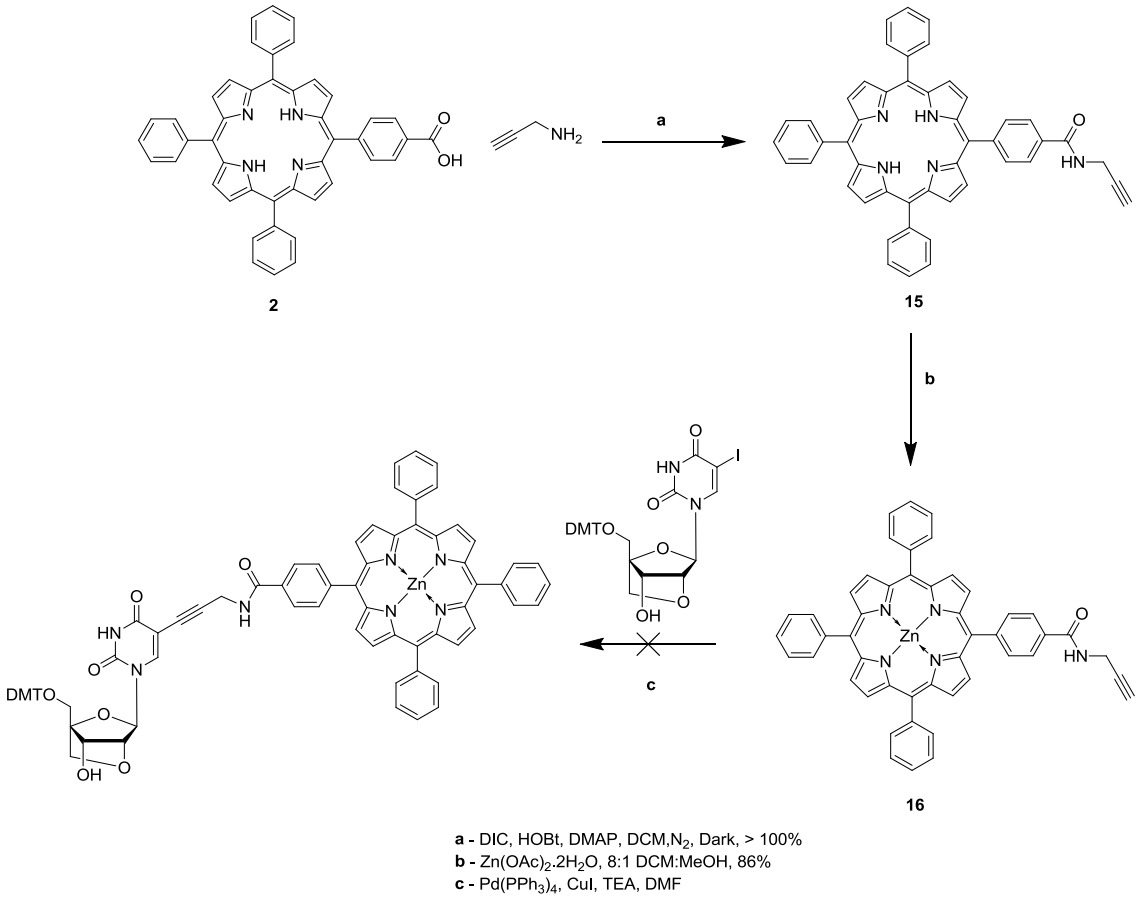
Synthesis of the porphyrin-modified LNA monomers followed a similar procedure to the porphyrin-modified monomers **7** and **21**. The Hrdlicka group kindly supplied 5-iodo-5'-ODMT-LNA-uracil as a starting material for the synthesis. The first monomer that we attempted to synthesise was the rigid linked porphyrin LNA monomer. The proposed synthetic route is shown in Scheme 5.2. Synthesis of the zinc metallated mono alkyne porphyrin, **12**, was carried out as described previously by the Stulz group.[49] Once this porphyrin had been synthesised, a Sonogashira cross-coupling reaction was required to add the iodo-LNA-uracil. The conditions used for this coupling were the same as for the synthesis of **21** but the reaction was found to progress much more quickly, 7 hours as opposed to over 24 hours. A recurring feature of these Sonogashira cross-couplings with iodo uracil is that the final product has a very similar rf value to the iodo uracil. This makes purification potentially quite tricky when it is essential that all of the iodouracil is removed as this could be phosphitylated and incorporated into the DNA synthesis in the following step. To aid separation of the product from any side materials silica H was used during column chromatography. The product was obtained with an excellent yield of 91 %. The next step was to phosphitylate this monomer, **13**, at the 3'-hydroxyl position using standard phosphityalting conditions as described before. As usual it was important to carry out this experiment in an inert atmosphere and in the absence of light. It was also found as with previously synthesised porphyrin-modified phosphoramidites that stability is an issue and DNA synthesis must take place immediately after synthesis.



Scheme 5.2 Synthetic route for acetylene linked porphyrin LNA monomer

Synthesis of the flexible amide linked porphyrin LNA monomer **18** was found to be more difficult than first expected. Synthesis of the monocarboxylic acid tetraphenylporphyrin **2** was carried out as described in chapter 4. Two options were considered to connect **2** to the LNA uridine, as the Sonogashira cross-coupling reaction of propargylamine with iodo uridine had proved to be difficult in the past and we only had a limited amount of the iodo-LNA-uridine. As a Sonogashira cross-coupling reaction between an alkyne porphyrin with iodo-uracil had been successful in the past, it made sense to try this with the flexible linked porphyrin (Scheme 5.3). An amide coupling between propargylamine and carboxylic acid porphyrin **2** was found to be facile and give the amide porphyrin **15** in a good yield. The next step was to metallate the porphyrin prior to the Sonogashira cross-coupling reaction. This is essential as free base porphyrins do not tend to couple very well in a Sonogashira cross-coupling, possibly because they have the ability to bind to the copper in the reaction mixture. The porphyrin was zinc metallated using the standard preparation of dissolving the porphyrin in a mixture of dichloromethane and methanol, adding zinc acetate and then thoroughly heating the reaction mixture with a heat gun for five minutes. This metallation of the porphyrin was found to be facile. A slight colour change is observed when the zinc is bound to the porphyrin. Metallation of the porphyrin was confirmed by NMR spectroscopy and mass spectrometry. A crystal structure was also obtained for this product, **16** (details of which can be found in the experimental section). The next step was to couple this metallated porphyrin to the iodo-LNA-uridine, however, attempts at this were unsuccessful. This was somewhat surprising considering there

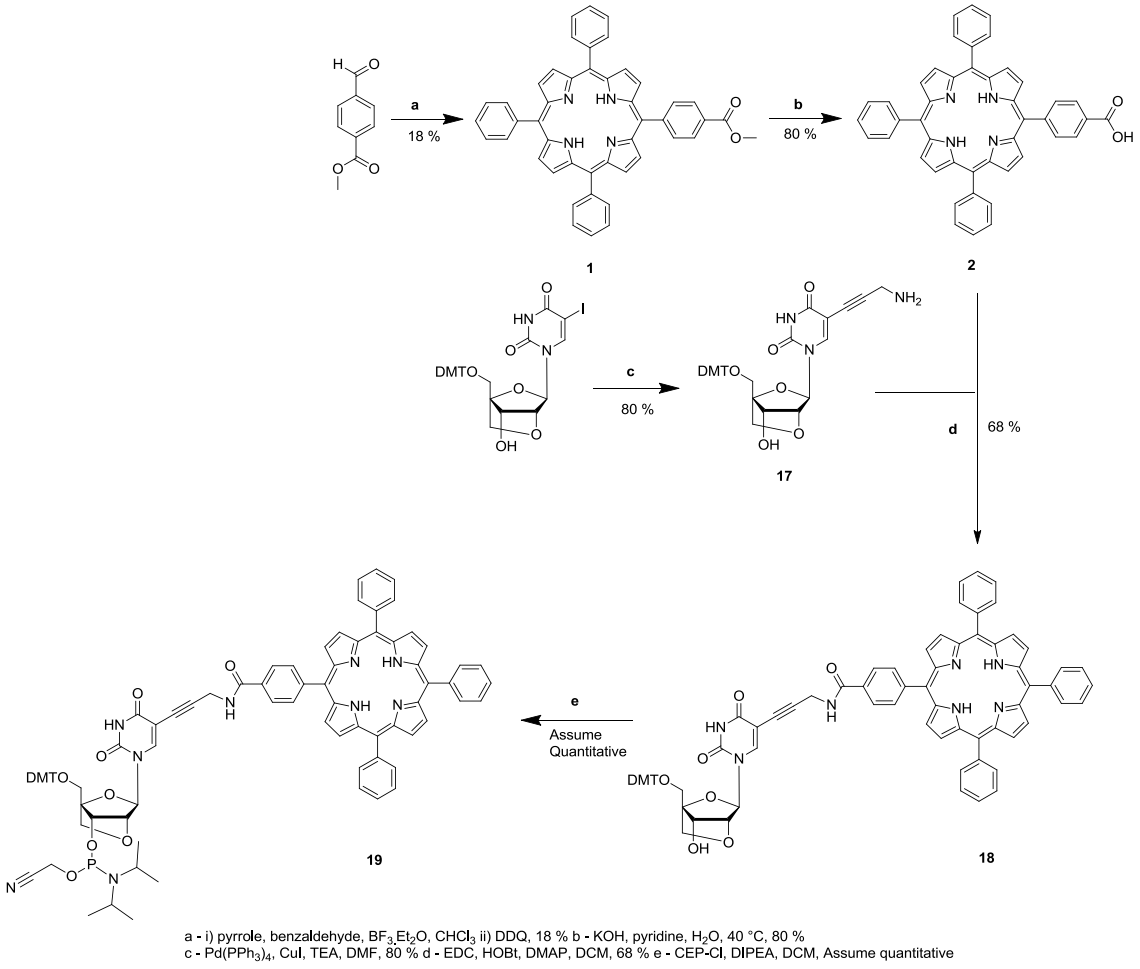
are other examples of synthesis of modified nucleobases with a propargylamine linker that use this approach.[206]



Scheme 5.3 Unsuccessful synthetic route towards the synthesis of the flexible amide linked porphyrin LNA monomer

After this set back, it was decided that the propargylamine would be coupled to the iodo-LNA-uridine first, followed by an amide coupling to join the porphyrin (Scheme 5.4). Previous experiences of the Sonogashira cross-coupling reaction between iodo-uridine and propargylamine have shown the reaction to be slightly unpredictable. There has been occasions were very poor yields of the final product had been obtained but there is no obvious reason for this. However, this time the coupling with iodo-LNA-uridine was found to proceed smoothly and the pure product, **17**, was obtained with a good yield (80 %). The method used to couple the porphyrin carboxylic acid to the propargylamine-LNA-uridine was analogous to that used in the synthesis of **7**. The reaction proceeded in a similar fashion to that of **7**, but the final yield was poorer in comparison. The flexible amide porphyrin monomer, **18**, was then phosphitylated at the 3'-OH position prior to DNA synthesis. Again, the

phosphoramidite was found to be highly unstable so had to be used immediately for DNA synthesis.



Scheme 5.4 Synthetic route for flexible amide linked porphyrin LNA monomer

5.4 Solid phase oligonucleotide synthesis

The LNA porphyrin-modified oligonucleotides were synthesised using standard solid phase synthesis. The sequences used were taken from previous work carried out by the Hrdlicka group.[205] All syntheses were carried out on a 1 µmol scale using standard coupling conditions for unmodified phosphoramidites. The coupling time for the porphyrin-modified phosphoramidites was extended to 6 minutes. As found from previous experiences, the porphyrin-modified phosphoramidites had to be dissolved in a mixture of acetonitrile and dichloromethane (1:1). The unmodified sequences **U1** and **U2** were synthesised with the final DMT group and then purified with Glen-Pak columns using the DMT-on purification protocol. The porphyrin-modified strands were synthesised DMT-on with the final base bearing a fluorous DMT group off, as was carried out for the porphyrin-modified G-quadruplexes. The crude oligonucleotides

were purified using Fluoro-Pak columns (Berry and Associates). Once purified, yields of the oligonucleotide synthesis were assessed by UV-Vis spectroscopy.

Strand Name	Sequence	Yield (nmoles)
U1	5' – GTG ATA ACG – 3'	611
U2	5' – GCA TAT CAC – 3'	576
B1	5' – GTG A**B**A TGC – 3'	562
B2	5' – GCA TA**B** CAC – 3'	347
B3	5' – GCA **B**AT CAC – 3'	644
B4	5' – GCA **B**A**B** CAC – 3'	708
F1	5' – GTG A**F**A TGC – 3'	228
F2	5' – GCA TA**F** CAC – 3'	362
F3	5' – GCA **F**AT CAC – 3'	371
F4	5' – GCA **F**A**F** CAC – 3'	342

Table 5.1 Oligonucleotides synthesised for this study. **B** denotes a modification using the rigid linked porphyrin-modified LNA monomer and **F** denotes a modification using the flexible amide linked porphyrin LNA monomer

5.5 Circular dichroism spectroscopic analysis

Structural analysis was carried out by using circular dichroism spectroscopy. Separate analysis was carried out on the DNA and porphyrin regions. We were particularly interested in the DNA region as the LNA modification has the possibility of having an impact on the overall structure of the DNA. As mentioned in the introduction, the methylene bridge between the 2'-oxygen and the 4'-carbon locks the ribose ring in an RNA like, C3'-*endo* conformation. It is unlikely that this will have a major effect on the overall DNA duplex structure, like changing the duplex from B-DNA to A-DNA, but there could be some subtle changes that may be linked to any effect the porphyrin LNA modification has on the overall stability of the duplex.

5.5.1 Double-stranded DNA containing the rigid linked porphyrin LNA modification

For circular dichroism experiments the oligonucleotide solutions (4.0 μM) were prepared in 100 mM sodium phosphate (pH 7.4) supplemented with 100 mM sodium chloride and 1 mM Na_2EDTA. The samples were heated to 95 °C for 5 minutes and allowed to slowly cool to ensure duplex formation. We first observed the duplex formation of oligonucleotides containing just the complementary strand with the strands that are modified with the rigid porphyrin LNA and compared this to the analogous unmodified DNA duplex. All of the strands show a similar signature with a small peak around 220 nm followed by a trough around 250 nm and finally a larger peak at 275 nm (Figure 5.2). All of the strands exhibit a circular dichroism spectrum that is not too dissimilar to what would be expected for B-DNA. The most striking difference between the unmodified duplex with the modified duplexes is that the trough at 250 nm for the unmodified duplex is much larger ($\Delta\varepsilon$ = –60 and –23 $M^{-1}cm^{-1}$ for **U1.U2** and **U1.B2** respectively). This change is possibly due to the base stacking present within the DNA duplex, with more efficient stacking observed for the unmodified duplex. Less efficient stacking in the porphyrin-modified duplexes could be due to ribose ring conformation of the LNA modification or also due to the presence of the large hydrophobic porphyrin. In all strands there is relatively no deviation in wavelength of the peak at 220 nm although the intensity of the peak varies from strand to strand.

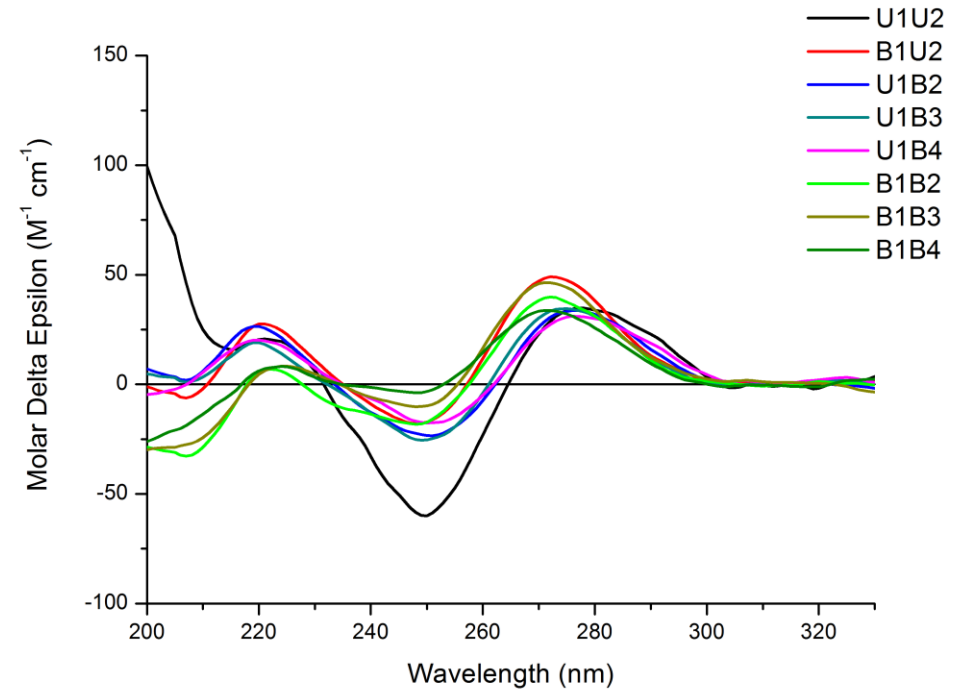


Figure 5.2 Circular dichroism spectra of double-stranded DNA containing just the rigid porphyrin LNA modification – DNA region shown

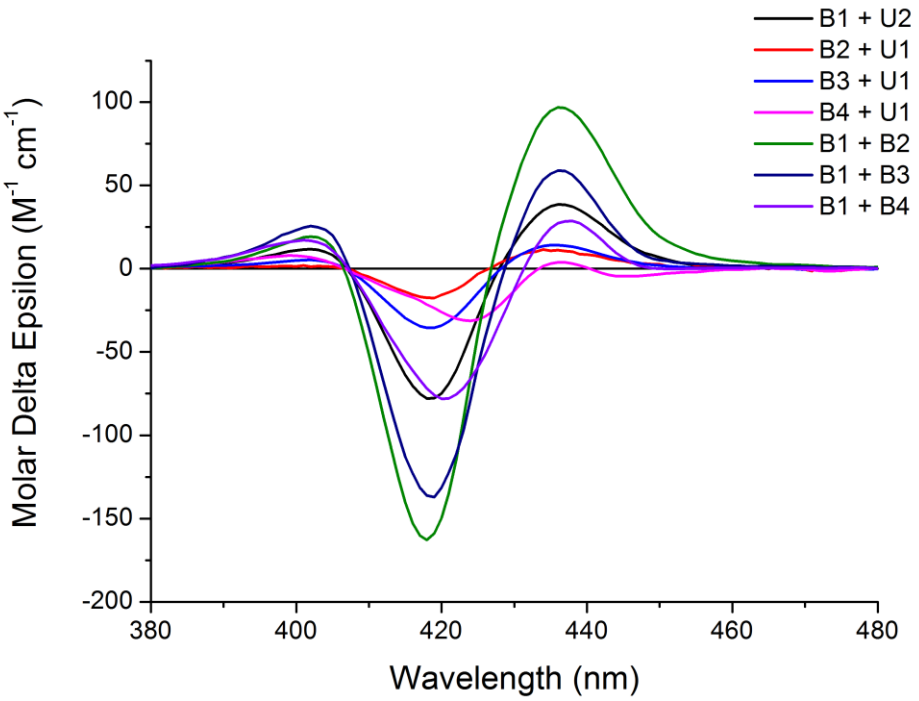


Figure 5.3 Circular dichroism spectra of double-stranded DNA containing just the rigid porphyrin LNA modification – Porphyrin Soret band region shown

The porphyrin region of the duplexes containing just the rigid porphyrin LNA monomer tend to give a positive Cotton effect (– → + with increasing wavelength) with varying degrees of circular dichroism intensity. One feature that can be observed from this set of data (Figure 5.3) is that the trough and peaks at 419 nm and 436 nm respectively are greater for duplexes **B1.B2** and **B1.B3** than any of the others ($\Delta\varepsilon$ = –163 and +97 $M^{-1}cm^{-1}$ for **B1.B2**). It can be hypothesised that because there are two porphyrins within close proximity of each other the larger circular dichroism signal is from an excitonic coupling effect between these porphyrins. However, this effect is not as great for **B4.U1** ($\Delta\varepsilon$ = –78 and +29 $M^{-1}cm^{-1}$), which is another system with two porphyrins present. This is due to the position of the porphyrins within the DNA duplex. The extent of excitonic coupling between porphyrins is dependent on the angle and distance between the porphyrins.[104] The porphyrins in the duplex **B4.U1** are present on the same strand and are separated by one nucleotide. They are a further distance apart from each other than they would be in the case of the duplexes with both strands containing porphyrin modifications. This distance seems to be too far for excitonic coupling to occur so a smaller circular dichroism signal is observed. When another porphyrin is introduced into this duplex (**B1.B4**), there are three porphyrins all within in close proximity to each other in a "zipper type" formation. This arrangement should allow excitonic coupling to take place between porphyrins which leads to the larger circular dichroism signal. However, this prediction does not agree with the observed data as the signal is reduced in comparison to **B1.B2** and **B1.B3**.

5.5.2 Double-stranded DNA containing the flexible amide linked porphyrin LNA modification

Next we observed the effect of using the flexible amide linker as opposed to the rigid linker in duplex DNA. As with the rigid linker, the spectra observed do not deviate greatly form what would be expected from B-DNA (Figure 5.4). Most samples give maxima at +240/−250/+275 nm, which is similar to what was observed before with the rigid linked porphyrin LNA. The exact wavelength of the peak maxima at −250/+275 nm is shifted from sample to sample depending on where the porphyrin LNA modification is present. It was observed that the maximum around +275 nm is shifted by roughly 5 nm to +270 nm when a duplex with a porphyrin LNA modification is present on both strands. This is also the case when the rigid porphyrin LNA is present. In both examples, as the number of porphyrin modifications increase, the amount of deviation from the expected B-DNA spectrum increases. As suggested before, this is most likely an effect of the porphyrin modification perturbing the duplex DNA structure.

When looking at the circular dichroism spectra of the porphyrin Soret band region of the duplex containing the flexible amide porphyrin LNA modification (Figure 5.5) it is clear that the change in the linker has led to an increase in complexity of the spectrum. The amide linker provides a higher degree in linker flexibility than the alkyne linker used in the rigid systems. Because of this it is possible for the porphyrin to freely move in a larger space so that any effect of induced chirality from the DNA duplex is decreased. This can be observed by comparing the peak maxima intensities to the same strands that use the rigid linker. For example duplexes **B1.B2** and **B1.B3**, both have much larger peak maxima that the flexible amide equivalent duplexes **F1.F2** and **F1.F3** ($\Delta\varepsilon$ = −163 and +97 $M^{-1}cm^{-1}$ for **B1B2** and $\Delta\varepsilon$ = −44 and +37 $M^{-1}cm^{-1}$ for **F1.F2**). It can also be observed that the shape of the circular dichroism signals from some of the duplexes are quite different from the usual negative to positive bisignate signal. **F4.U1** is a good example of this with just one peak maxima at +436 nm present.

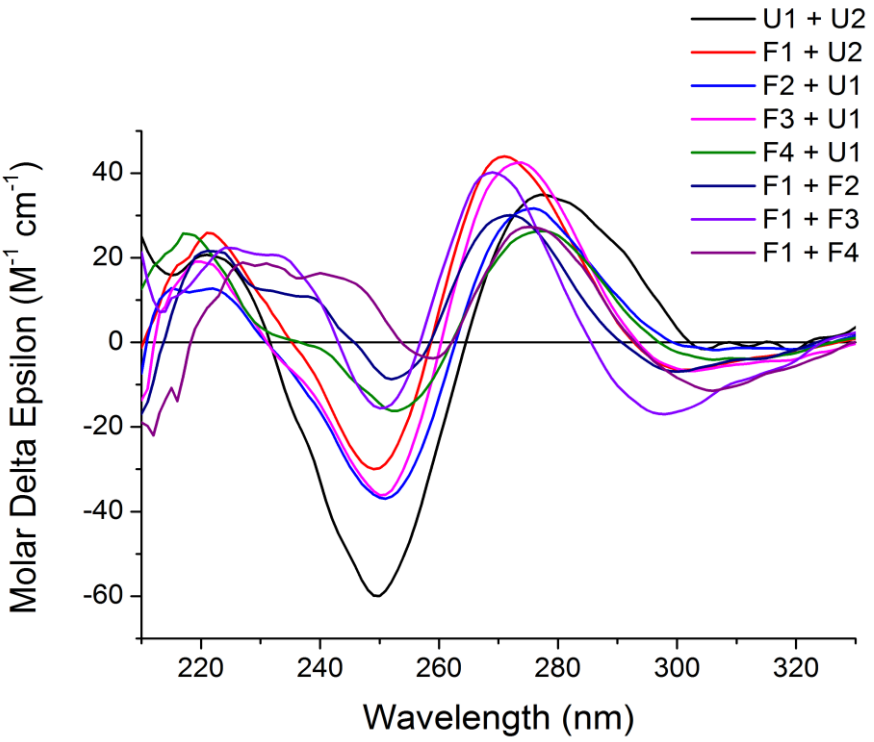


Figure 5.4 Circular dichroism spectra of double-stranded DNA containing just the flexible porphyrin LNA modification – DNA region shown

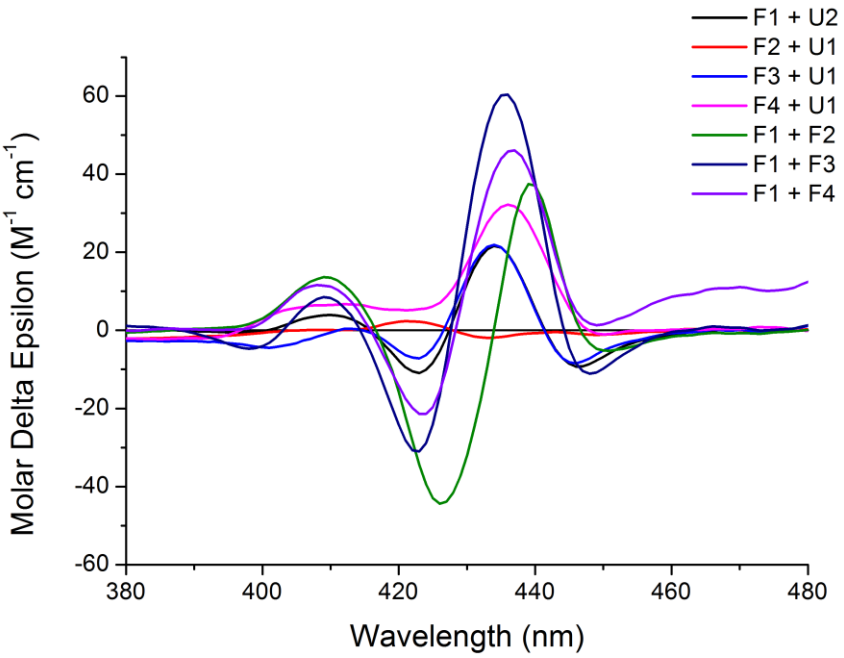


Figure 5.5 Circular dichroism spectra of double-stranded DNA containing just the flexible amide porphyrin LNA modification – Porphyrin Soret band region shown

5.5.3 Double-stranded DNA containing both rigid linked porphyrin and LNA flexible amide linked porphyrin LNA modifications

We next investigated DNA duplexes formed using DNA that has the rigid linked porphyrin LNA modification on one strand and the flexible amide linked porphyrin LNA modification on the complementary strand. The DNA region (Figure 5.6) of the mixed porphyrin duplexes are reasonably similar to what has been observed before, with peak maxima present at +220/−250/+270 nm. The duplexes which contain three porphyrin modifications, **B1.F4** and **F1.B4**, show smaller signals than those with two modifications present ($\Delta\varepsilon$ = +7, −10 and +26 $M^{-1}cm^{-1}$ for **F1.B4**). This is especially clear at 220 and 240nm. Between these two peaks there are other signals present which may indicate that the overall B-DNA structure is increasingly perturbed with less efficient base stacking present as more LNA modifications are present.

The porphyrin region for these mixed duplexes (Figure 5.7) was the most complicated set of spectral data collected in this study. A variety of spectra were observed from the samples which was surprising as all previously analysed duplexes gave a positive Cotton effect (− → + with increasing wavelength). This was not observed in any of the duplexes this time, however, a negative cotton effect (+ → −) was observed in duplexes **F1.B2** and **F1.B3**. Other samples such as **B1.F2** and **B1.F4** exhibited a single positive peak at 440 nm. A feature all of these duplexes shared was that a very broad positive signal was observed from 400 nm to 420 nm. This effect may be coming from the flexible amide linked porphyrin modification as a peak had been observed in this region from the previous strands containing this modification. It is difficult to say why such a variety of spectra is observed for these strands where the previous spectra for DNA duplexes with porphyrin LNA modifications on both strands were reasonably similar to each other.

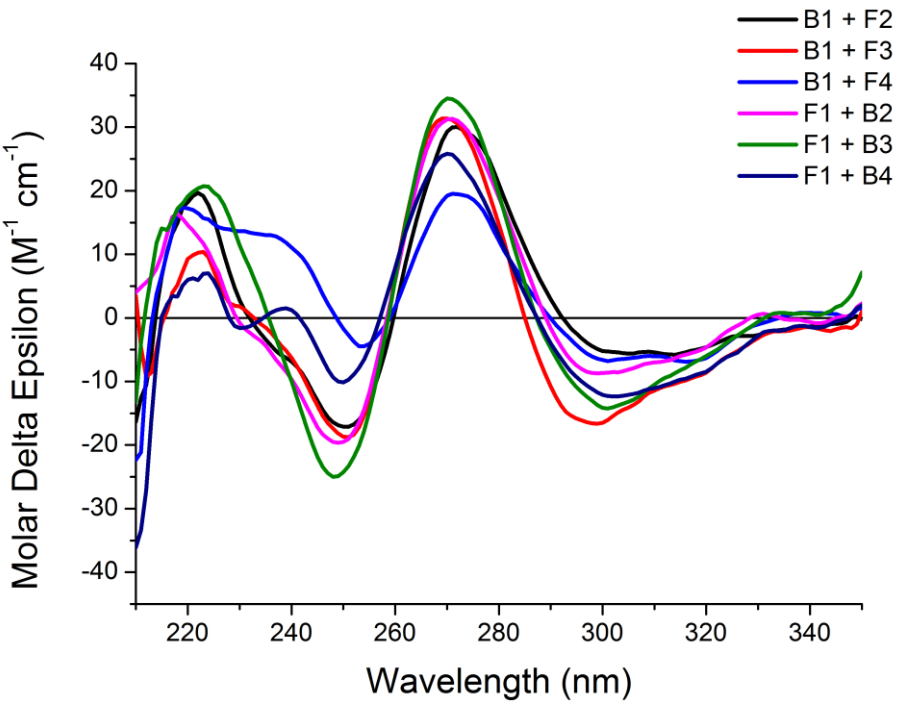


Figure 5.6 Circular dichroism spectra of double-stranded DNA containing both rigid linked porphyrin LNA and flexible amide porphyrin LNA modifications – DNA region shown

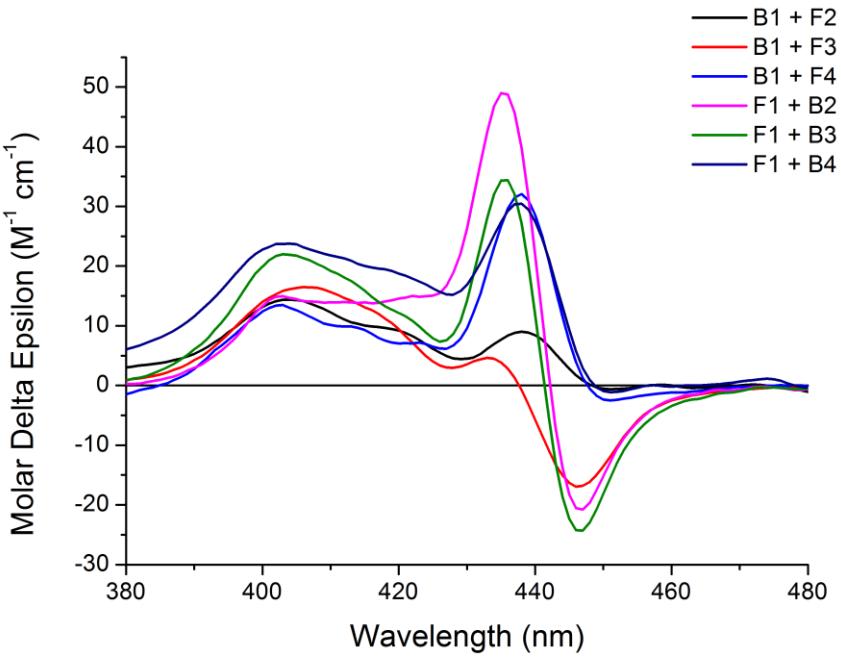


Figure 5.7 Circular dichroism spectra of double-stranded DNA containing both rigid linked porphyrin LNA and flexible amide porphyrin LNA modifications – porphyrin Soret band region shown

5.5.4 Zinc metallation of the porphyrin LNA modification

The final part of this investigation was to metallate one of the porphyrin-modified strands and to see how this affects the porphyrin circular dichroism signal (Figure 5.8). It was decided that only the oligonucleotides that contain two porphyrin modifications (**B4** and **F4**) would be metallated. The metallation of these porphyrins was carried out post DNA synthesis by heating the DNA to 85 °C for five minutes in an aqueous solution containing a large excess of zinc acetate. The porphyrin DNA was then purified using a desalting protocol for a Glen-Pak cartridge. Confirmation of complete metallation was obtained through UV-Visible spectroscopy, the number of Q bands present decreases from four (free base porphyrin) to two (zinc metallated porphyrin). The complementary DNA strand was then added to the porphyrin metallated DNA and the circular dichroism spectra recorded. The decrease in peaks is due to a reduction in the symmetry of the porphyrin ring when the zinc is present in the centre. As there is a change in symmetry of the porphyrin, one may expect that this will have an effect on the circular dichroism spectrum that is produced from these molecules. This was indeed found to be the case, with most strands exhibiting a bisignate signal (– → +) with peak maxima at –420/+436 nm. The exception to this is the trisignate signal (– → + → –) seen for **B4Zn.B1** and **B4Zn.F1** with the final trough coming at 447 nm. The observed changes in circular dichroism signals for the zinc metallated porphyrin samples compared to the free base porphyrin samples agree with a previous example of metallation of porphyrin DNA.[82]

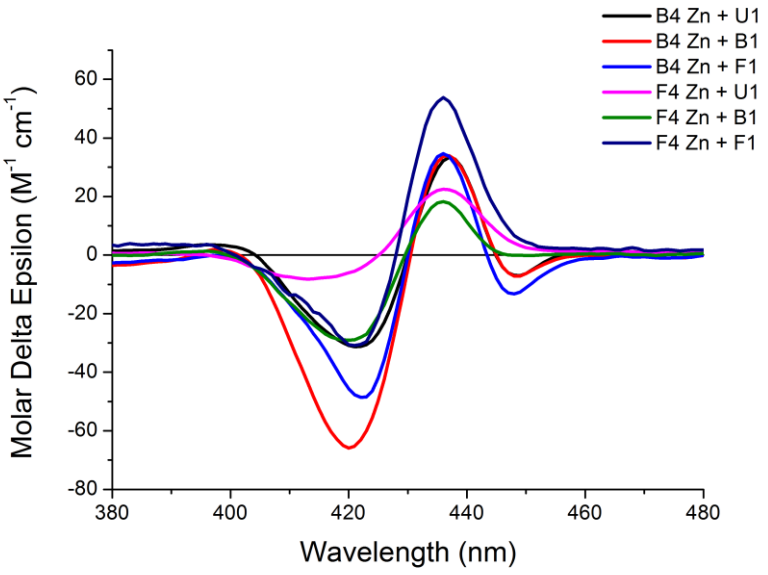


Figure 5.8 Circular dichroism spectra of double-stranded DNA with a zinc metallated porphyrin LNA modification present

5.6 UV-Vis melting studies of porphyrin-modified LNA

UV melting experiments were carried out to assess the stability of the porphyrin LNA duplexes. As described previously, the melting temperature (T_m) is used to measure the DNA secondary structure stability. For assessing DNA duplexes, the absorbance is recorded at 260 nm. As the temperature of the system is gradually increased, a hyperchromic shift will be observed due to the denaturation of the DNA duplex. When the DNA duplex denatures, the hydrogen bonds that hold the duplex together are broken. The hydrogen bonding limits the resonance of the aromatic nucleobases which in turn limits the absorbance of the DNA. So as the hydrogen bonds are broken, the absorbance of the DNA increases. The absorbance of the porphyrin Soret band was also monitored during the experiment. The behaviour of the porphyrin is difficult to predict in these DNA duplex systems so it will be interesting to see what effect the denaturing of the DNA duplex has on the porphyrin absorbance. As we observed in chapter 4, changes in the porphyrin absorbance can provide hints into the porphyrins interaction with the DNA as it is denatured. For the experiment, the oligonucleotides (2.5 μM) in 100 mM sodium phosphate buffer, pH 7.0 supplemented with 100 mM NaCl and 1 mM Na_2EDTA, where first heated up to 95 °C for 5 minutes and then allowed to slowly cool to room temperature to ensure formation of the duplex. Temperature ramps (Table 5.2) were repeated three times and the data averaged and normalised.

Stage	Start Temperature (°C)	End Temperature (°C)	Rate (ΔT/min)
Melt	10	70	0.5
Anneal	70	10	0.5

Table 5.2 Melting temperature experiment parameters

The UV melting of the unmodified DNA duplex **U1.U2** was measured as a control for this experiment. This duplex gave the expected sigmoidal curves for the melting and annealing phases with a T_m of 34.7 °C. For all DNA duplexes containing the porphyrin LNA modification, no sigmoidal curves were observed, with only a small increase in absorbance at 260 nm (Figure 5.9). This result was surprising as from the previous circular dichroism data collected it would appear that these systems are forming DNA duplexes. The small change in absorbance could be attributed to the porphyrin modification perturbing the DNA duplex structure so that the stacking of nucleobases is reduced. This effect has been previously observed by the Hrdlicka

research group with pyrene modified LNA modifications.[206] If this is the case, the effect is more pronounced for our samples possibly because the oligonucleotides used in this study are quite short (9 nucleotides in length). It could also be argued that the DNA duplex is not being formed at all, hence the lack of sigmoidal curve produced. No assessment of DNA duplex stability could be made from the data collected.

The porphyrin Soret region (420 nm) also gave interesting results that are difficult to interpret. It was stated before the experiment that it would be difficult to predict the behaviour of the porphyrin during DNA denaturation, and this was found to be the case. In all samples, a hyperchromic shift was observed as the temperature increased. However, the nature of this shift was different from sample to sample (Figure 5.9), which made the interpretation of the porphyrin behaviour near impossible. In one instance (**U1.B3**) a sigmoidal curve was observed from the porphyrin absorbance that gave a T_m of 37.3 ˚C but as the change in states was very gradual, this value is not so accurate.

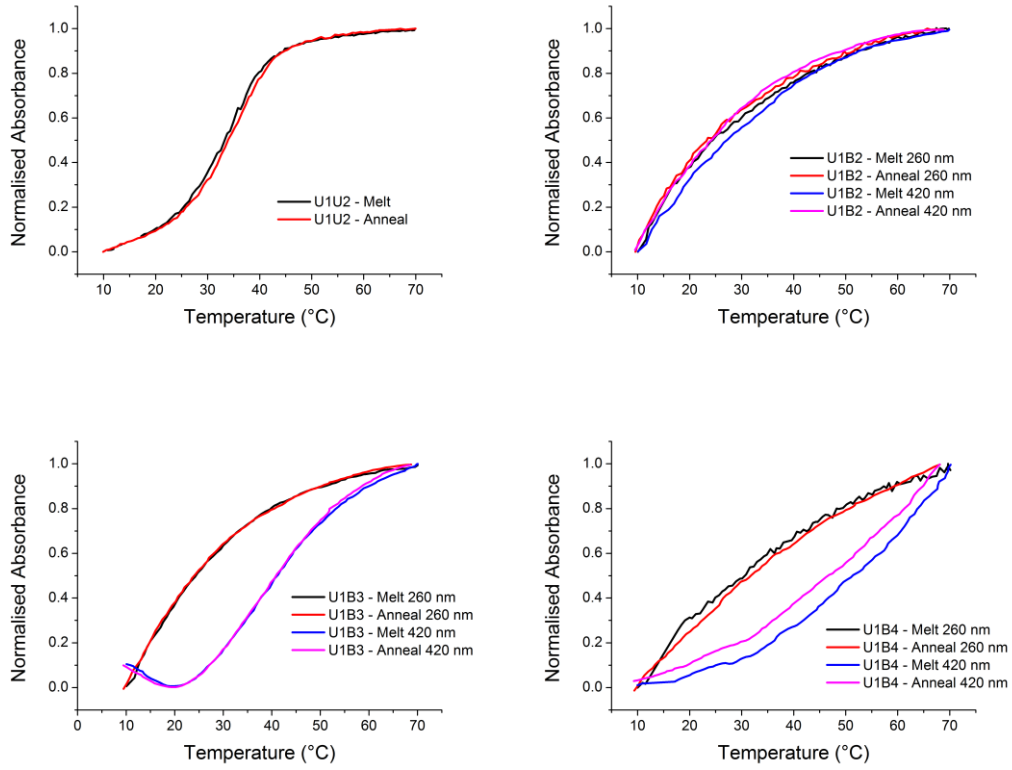


Figure 5.9 UV melting curves for a selection of LNA duplexes. All experiments were performed in 100 mM sodium phosphate buffer, pH 7.0 supplemented with 100 mM NaCl and 1 mM Na_2EDTA

5.7 Circular dichroism melting of porphyrin-modified LNA

Circular dichroism melting was used to assess the stability of the porphyrin-modified LNA. The experiments were carried out in with the same conditions used to measure the CD spectra. Data was collected in 10 °C intervals with an equilibration time of 5 minutes for each temperature. Data was fitted using Boltzmann distribution in Origin. Data points were taken from 250 nm, as this point was observed to give the largest change in molar delta epsilon value as the temperature increased. As with the UV melting experiments described in chapter 4, the point at which half of the DNA states are in the duplex form and half are in the random coil state, T_m, will be used to assess stability (Table 5.3).

DNA Duplex	T_m (°C)	T_m error (°C)	ΔT_m vs. **U1.U2** (°C)
U1.U2	33.1	0.2	N/A
B1.U2	36.2	3.3	+ 3.1
U1.B2	32.0	5.2	– 1.1
U1.B3	34.8	4.6	+ 1.7
U1.B4	30.9	3.9	– 2.2
B1.B2	N/D	N/D	N/D
B1.B3	43.9	1.4	+ 10.8
B1.B4	38.3	0.5	+ 5.2

Table 5.3 Melting temperatures obtained from rigid porphyrin LNA circular dichroism denaturing experiments. Thermal denaturation experiments were performed in 100 mM sodium phosphate (pH 7.4) supplemented with 100 mM sodium chloride and 1 mM Na_2EDTA

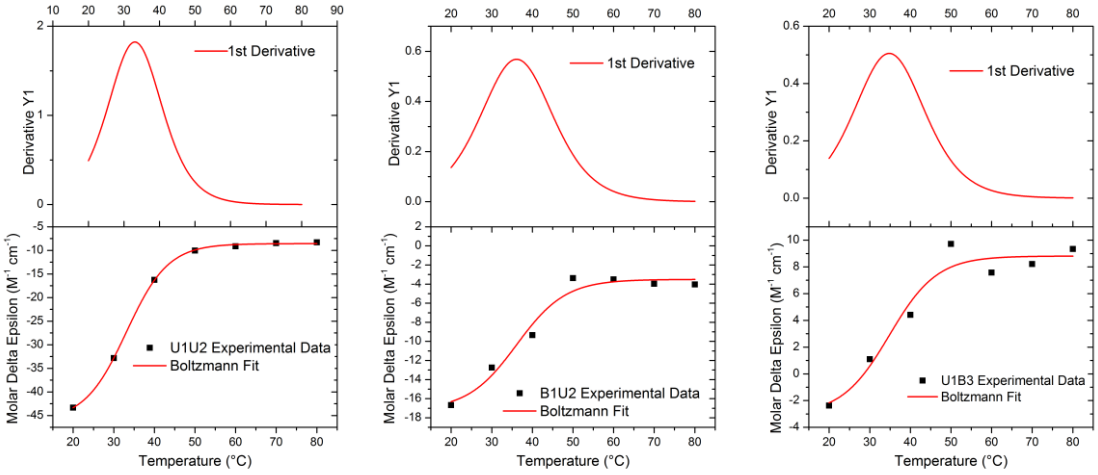


Figure 5.10 Boltzmann fitted melting curves with first derivative for porphyrin-modified LNA

Due to limited nature of the time available on the beamline at the diamond light source, data was only collected for DNA duplexes containing just the rigid LNA modification. As the data was collected in 10 ˚C intervals, the T_m values obtained will not be accurate but will help to gain an understanding on what effect the porphyrin LNA modification has on the DNA duplex stability. All samples apart from **B1.B2** could be fitted using a Boltzmann fit (carried out with Origin 8.6) from which the first derivative could be used to calculate a T_m value (Figure 5.10 and Table 5.3). As the data could be fitted to a sigmoidal curve, it could be assumed that there is a two state system present. This could be the DNA in duplex form at the beginning and then as a random coil and the end of the melting experiment, as you would normally expect to see. The T_m obtained of the unmodified control duplex **U1.U2** is reasonably close to the one obtained from the UV melting studies (33.1˚C for the circular dichroism melting and 34.7 ˚C for the UV melting experiment).

In the samples that contain the porphyrin LNA modification on just one of the strands in the DNA duplex, the degree of stability offered by the porphyrin LNA modification to the DNA duplex differs depending on its position in the DNA duplex but in all four cases, the difference in melting temperature is quite small. Duplexes containing porphyrin LNA modifications on both DNA strands were analysed next. Apart from **B1.B2**, which could not be fitted to a sigmoidal curve, the melting temperature of the DNA duplexes was found to be greater than the unmodified DNA duplex. **B1.B3** gave the largest increase in stabilisation with an increase of 10.8 ˚C observed. The DNA duplex **B1.B4**, which contained three porphyrins in a zipper arrangement also showed an increase in melting temperature (5.2 ˚C). The larger increase in duplex stability is consistent with what has been observed in other

LNA.DNA duplexes, where it has been observed that as the stability of the duplex increases as number of LNA modifications increases. This could be a factor for the stabilisation observed here, but it should be noted that largest increase in stability came from a duplex containing two modifications instead of three, which is not what would be expected.

One major drawback to this data collected is that the temperature was only measured in 10 °C intervals, which has a negative impact on the accuracy of the T_m values obtained. Considering this, we repeated the circular dichroism melting experiments whilst recording more data points over a larger temperature range. Due to time constraints, the data was recorded using an Applied Photophysics Chirascan circular dichroism spectrometer rather than the beamline, which the previous experiments were measured on. The experiments were carried out in with the same conditions used to measure the CD spectra at 5 °C. Data was collected in 5 °C intervals with an equilibration time of 2 minutes at each temperature. Data was fitted using Boltzmann distribution in Origin 8.6 and the T_m values were taken from the first derivative of the fitted curve (Table 5.4). Data points were taken from 250 nm, as this point was observed to give the largest change in molar delta epsilon value as the temperature increased.

DNA Duplex	T_m (°C)	T_m error (°C)	ΔT_m vs. **U1.U2** (°C)
U1.U2	26.9	1.6	N/A
B1.U2	17.2	2.2	– 9.7
U1.B2	16.7	1.0	– 10.2
U1.B3	19.1	1.1	– 7.8
U1.B4	20.0	2.0	– 6.9
U1.F2	17.1	1.1	– 9.8
U1.F3	18.5	0.6	– 8.4
U1.F4	17.6	1.4	– 9.2
B1.B2	19.2	2.0	– 7.7
B1.B3	19.6	1.7	– 7.3
B1.B4	17.1	0.8	– 9.8
B1.F2	19.9	1.9	– 7.0
B1.F3	20.9	3.0	– 5.9
B1.F4	18.2	3.2	– 8.7

Table 5.4 Melting temperatures obtained from rigid porphyrin LNA circular dichroism denaturing experiments. Thermal denaturation experiments were performed in 100 mM sodium phosphate (pH 7.4) supplemented with 100 mM sodium chloride and 1 mM Na_2EDTA

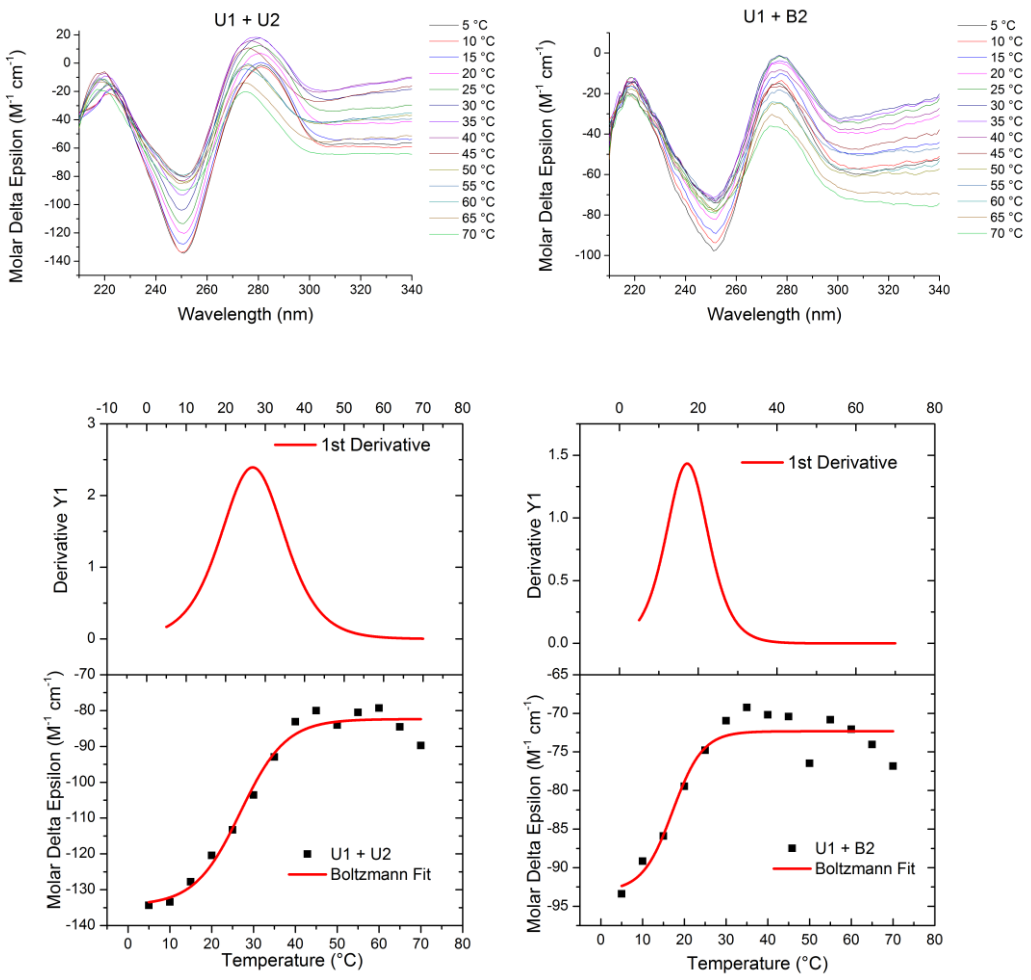


Figure 5.11 Top - examples of circular dichroism melts. Bottom – corresponding Boltzmann fitted melting curves with first derivative for porphyrin-modified LNA from values taken at 250 nm

As the data was collected in 5 ˚C intervals, the T_m values obtained will not be accurate but will help to gain an understanding on what affect the porphyrin LNA modification has on the DNA duplex stability. The data recorded was surprising in a number of ways. For the unmodified DNA duplex **U1.U2**, the T_m obtained was 26.9 ˚C. This value is almost 10 ˚C lower that what we would have expected. The next surprise was that it was observed in all cases that the porphyrin LNA modification destabilises the DNA duplex by around 6 to 10 ˚C. This behaviour has been observed previously with porphyrin DNA duplexes[118] and also with pyrene LNA duplexes[206]. In the case of pyrene C5 modified thymidine, both the standard nucleoside and its LNA equivalent show a destabilisation effect on the DNA duplex and the behaviour appears to be similar with the porphyrin modification. Porphyrin and pyrene are similar in that they

are both hydrophobic molecules. If they are affecting the efficiency of the base stacking within the double helix, as mentioned earlier, it would be expected that the stability of the DNA duplex would be decreased.

Analysis of this data was found to be problematic due to some issues we believe are related to the measurement of the data. A clear issue present in all of the data collected was that the baseline at the higher wavelengths (320-350 nm) was inconsistent as the temperature increased. A subtraction of the buffer was carried out on all of the data but there was still a large amount of variation on the baselines. Attempts to correct by subtracting the baseline had a negative effect on the data. For most of the data, the Boltzmann fitting could only be performed up to 50 °C (see appendix). Around this temperature the Δε values rapidly decreased. This effect seems related to the issue with the baseline. Considering these issues, more work needs to be carried out in assessing the stability of the porphyrin LNA duplexes as I do not believe these are reliable measurements.

5.8 Conclusions and outlooks

Here, we have shown a new class of porphyrin-modified LNA monomers. When they are incorporated in an oligonucleotide, it was observed that this oligonucleotide could form a duplex DNA structure with its complementary sequence. A B-DNA duplex was observed in all examples despite the LNA modifications ability to influence the formation of A-DNA. However, if more LNA modifications were present within the oligonucleotide one may expect the formation of A-DNA. The circular dichroism spectroscopy of the porphyrin regions of these duplexes was also studied. When only one porphyrin modification with a rigid linker was used, a bisignate signature was observed. The complexity of the data increased as flexibility was introduced to the porphyrin linker. This effect could be due to a superposition of signals as the porphyrin has a greater degree of freedom. When more than one porphyrin was present in the duplex more complexities arose. This would be due to the interaction of the porphyrins by excitonic coupling.

The stability of these modified duplexes was also studied. Usually this is carried out by UV melting, but this was unsuccessful with the porphyrin modifications. This is possibly due to the porphyrin modification affecting the base stacking present in the DNA duplex, which reduces hypochromicity during UV melting. Circular dichroism melting was then used to try and get an understanding of the affect the porphyrin modification has on the duplex stability. Results from the circular dichroism melting were inconclusive. The first measurements showed that it is possible the porphyrin LNA could be stabilising the duplex, but due to the small number of data points collected the T_m values are not very accurate. Further studies indicated a decrease in T_m

value, but problems with the data measurement makes these values unreliable. In future circular dichroism melting experiments, a single wavelength (250 nm) should be monitored so more data points can be collected as to determine a more accurate T_m value. The recording of the circular dichroism spectrum from 210 nm to 350 nm takes at least five minutes to do. To do this over a larger number of temperatures than shown in this study would be impractical.

Due to the limited time available to carry out variable temperature circular dichroism, it would be sensible to look for other ways of assessing the stability of these duplexes. One method that would be ideal to prove the formation of the duplex structure would be native polyacrylamide gel electrophoresis. However, as discussed in the previous chapter, the porphyrin modification is responsible for large amounts of streaking on the gels. This makes the analysis of secondary structure formation almost impossible.

Further work in the future will be carried out by the Hrdlicka group using these porphyrin LNA modified oligonucleotides. These experiments will look into forming duplexes with RNA and observing the effect the LNA modification has on the stability compared with the DNA analogue. It is known that LNA will form more stable duplexes with RNA. From this we hope that we will not encounter the similar issues observed with assessing the duplex stability with the complementary DNA. Mismatches will also be introduced into the RNA to assess the porphyrin LNAs potential at detecting single nucleotide polymorphisms (SNPs).

6 Porphyrin-modified triphosphates

6.1 Aim

All porphyrin-modified DNA so far has been synthesised by standard solid phase oligonucleotide chemistry. As the number of modified 2'-deoxynucleoside triphosphates (dNTPs) synthesised and incorporated into DNA by primer extension and polymerase chain reaction increases, it would seem viable to attempt to synthesise porphyrin DNA using this route. Incorporation of the porphyrin modifier into DNA using solid phase oligonucleotide synthesis is a difficult process due to the porphyrin phosphoramidite instability. This project was carried out as a collaboration with Dr. Michal Hocek at the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague.

6.2 Introduction

Modified oligonucleotides are usually synthesised using one of two methods: solid phase oligonucleotide synthesis using functionalised phosphoramidites, or by post-synthetic oligonucleotide modification. Modified oligonucleotide synthesis using functionalised phosphoramidites can be difficult. The substituted nucleoside phosphoramidites must be stable to the conditions used during the oligonucleotide synthetic cycle (acidic detritylation, coupling, capping, oxidation and final acyl protecting groups cleavage by ammonia). It should also be noted that the functionalisation of the phosphoramidite can have an effect on the stability of the compound. Phosphoramidites are susceptible to oxidation as it is thermodynamically favourable for the phosphorus (III) atom to be irreversibly oxidised to phosphorus (V). Once oxidised, the phosphoramidite can no longer be added to the oligonucleotide. For standard phosphoramidites, oxidation is not such a problem if they are handled in the correct manner. However, it has been observed that some modifications, such as the porphyrin modifier, can decrease the phosphoramidite stability and their lifetime. In the case of post-synthetic modifications, they require selective and mild reactions and in most cases are further transformations of an already modified oligonucleotide.

Another way of preparing oligonucleotides is to use a class of enzymes called polymerases. The function of these enzymes is to polymerise new DNA or RNA against an existing DNA or RNA template in the processes of replication and transcription. Polymerases transcribe DNA or RNA in the 5' to 3' direction. Polymerases are used in polymerase chain reaction (PCR) (Figure 6.1), an important technique to amplify specific DNA sequences, developed by Kary Mullis in 1984. PCR has been used to

provide valuable diagnostic information in medicine and forensic science. PCR starts off with a parent double-stranded DNA sample that will act as two templates for the reaction. The DNA duplex is denatured by heating the solution to 95 °C for 15 seconds. Primer extension requires short primers, typically 20 to 30 nucleotides long that are recognised by the polymerase and are a short complement of the template DNA strand. The primers must hybridise to the single stranded DNA template strands, which happens when the solution is cooled to 54 °C. The solution is then heated to 72 °C where the polymerase extends both primers in the 5' to 3' direction by adding the complementary nucleoside triphosphates present in the solution. The temperature 72 °C is used as this is the optimal temperature for *Taq* DNA polymerase to operate at. This heat-stable polymerase comes from *Thermus aquaticus*, a thermophilic bacterium that lives in hot springs. This is one cycle of the polymerase chain reaction. To amplify more DNA the cycle is repeated many times. After *n* cycles, the sequence is amplified 2^n-fold so the amplification is a billionfold after 30 cycles. This can be carried out in less than an hour. A similar technique called primer extension (PEX) is used to synthesise oligonucleotides using polymerases. The primer extension is one cycle of the overall PCR process. Primer extension can be used to investigate the incorporation of modified nucleoside triphosphates with different polymerases.

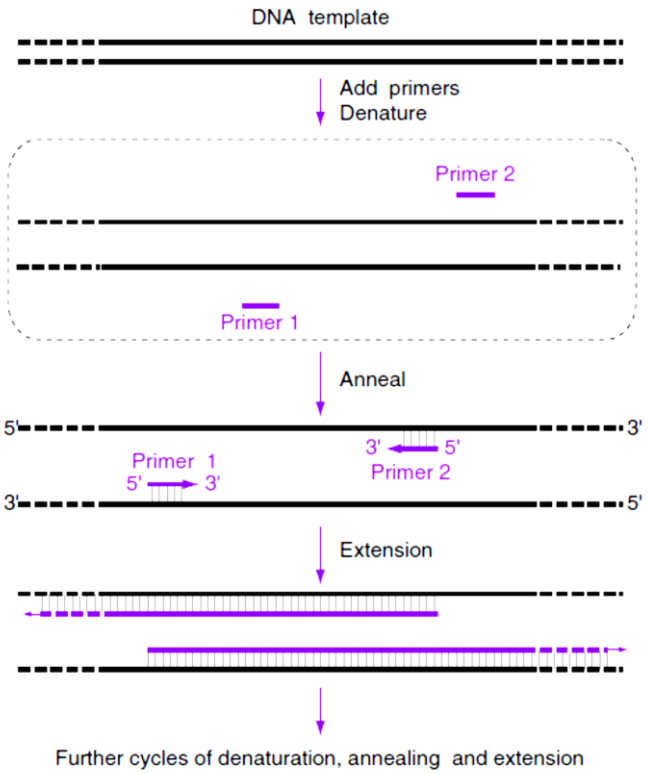


Figure 6.1 Polymerase chain reaction: first cycle[208]

Not only can polymerases incorporate standard nucleoside triphosphates (dNTPs) into an oligonucleotide, it has been shown that it is possible to incorporate a wide range of modified dNTPs. This has led to a viable route in synthesising modified oligonucleotides. The first polymerase incorporation of base-functionalised dNTPs was published in 1981 by Langer *et al.* who prepared dUTP biotinylated at the C5 position and found that it was a substrate for several DNA polymerases.[209] Over time more functionalised dNTPs have been synthesised and shown to be suitable substrates for incorporation by polymerases.[210] Famalouk *et al.* were the first to prepare a complete set of four modified dNTPs and study their incorporation during PCR.[211, 212] They found for the best results pyrimidines should be modified at the C5 position and deazapurines at the C7 position (Figure 6.2). It was observed that purines modified at the C8 position are poor substrates for the polymerase.[213] This is most likely because modifications at the C8 position are sterically unfavourable when the DNA is in its duplex form, as they will point towards the phosphate backbone. Modifications at the C5 position for pyrimidines and at the C7 position for deazapurines are suitable, as at these positions the modification will sit comfortably in the major groove of the DNA duplex.

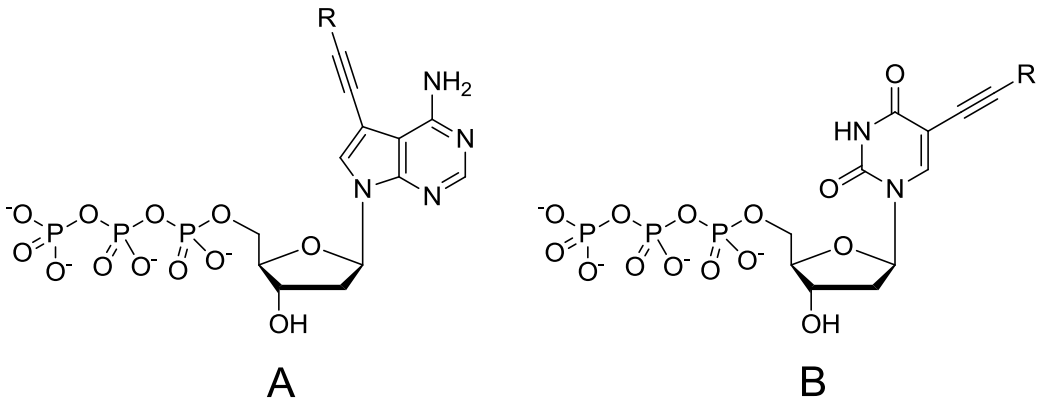
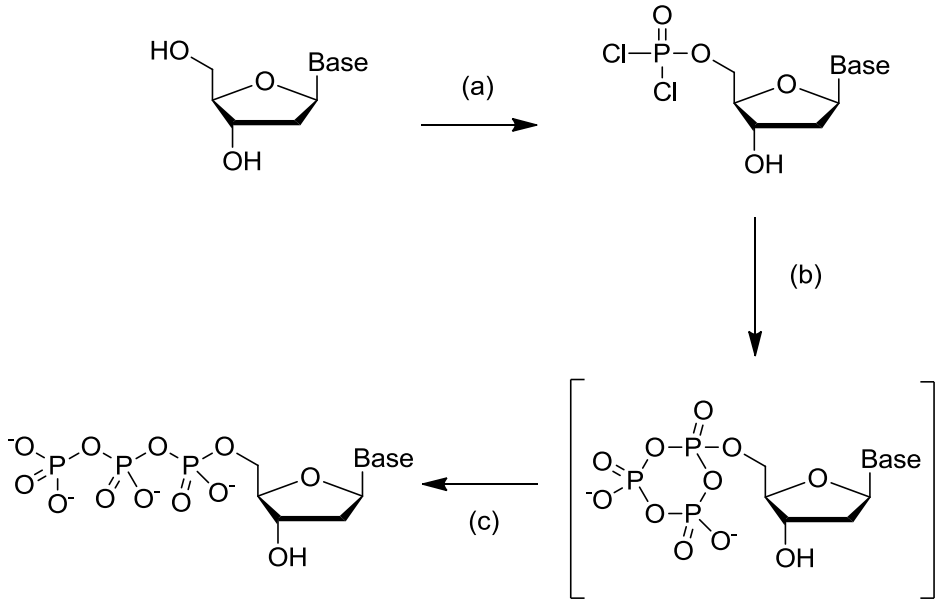


Figure 6.2 Nucleoside triphosphates with an alkyne linked modification. A: deazapurine modified at the C7 position, B: pyrimidine modified at the C5 position

Base modified dNTPs can be prepared by two methods. The first of these is the classical triphosphorylation of modified nucleosides (Figure 6.3).[214] This method can often be difficult and function groups on the modifiers often need to be protected prior to the reaction. Yields for this reaction are typically low, which is highly undesirable if a large multistep route has been used to synthesise the modified nucleoside. The first step in the triphosphorylation is to convert the 5'-hydroxyl group to a monophosphate. This is achieved by dissolving the nucleoside in dry trimethyl phosphate and adding dry phosphoryl chloride ($POCl_3$). Proton sponge (1,8-

bis(dimethylamino)naphthalene) can be added to the reaction mixture to remove any protons generated from the reaction between the 5'-hydroxyl of the nucleoside and the phosphoryl chloride. This reaction is carried out on ice and is the slower of the two steps involved in triphosphorylation. Once this reaction is complete, the monophosphate is converted to the triphosphate by the addition of tributylammonium pyrophosphate. The reaction is quenched by the addition of 1 M triethylammonium bicarbonate (TEAB).



(a) $POCl_3$, $PO(OMe)_3$; (b) $(Bu_3NH)_2H_2P_2O_7$, Bu_3N, DMF; (c) 1 M TEAB

Figure 6.3 Classical nucleoside triphosphorylation

A second, more popular way of synthesising functionalised dNTPs is to add the modification directly to the triphosphate via C-C bond formation using a cross-coupling reaction. There have been developments in palladium cross-coupling reactions that allow the reaction to take place in aqueous conditions,[215] which is ideal when considering the solubility of the triphosphate involved. It was Burgess *et al.* who first used this development to add fluorescent dyes with terminal acetylenes to 5-iodo-dUTP.[216] This method of synthesis is particularly favoured by the Hocek research group with many examples of modified dNTPs being produced.[217-219] In both methods of modified dNTP synthesis, the final products need to be purified by reverse-phase HPLC. Modified dNTPs are then lyophilised and stored at –20 °C where they should be stable for several months.

A variety of factors need to be taken into consideration when using modified dNTPs for oligonucleotide synthesis by primer extension. The type of modification can have an effect on how well the polymerase can incorporate the substrate into the

oligonucleotide. The size of the modification is a possible concern, if it is too large it may hinder the polymerases ability to incorporate it. However, somewhat surprisingly it is possible to incorporate some dNTPs with large modifications attached to them. One example of this is the incorporation of dNTPs modified with a large bis bipy substituted Ruthenium, which could be used as a redox-active DNA labelling tag.[220] How the modification is attached to the dNTP can affect how well the substrate is incorporated by the polymerase. For modifications that have bulky aromatic substituents, it has been found that a longer acetylene linker is incorporated more efficiently than if the modification was directly attached to the dNTP. Another point to take into consideration is the type of polymerase that is used. There is not a universal DNA polymerase that will incorporate all modified dNTPs. However, for the incorporation of C5-substituted pyrimidine dNTPs it has been observed that B-family polymerases such as KOD Dash, Pwo and Vent (*exo*-) are generally more efficient than A-family polymerases.[221] When attempting to incorporate a novel modified dNTP it is advised that a selection of polymerases should be screened in order to find an efficient one for that particular nucleotide.

6.3 Synthesis of porphyrin-modified triphosphates

Synthesis of the porphyrin-modified deoxyuridines was based on previous synthetic schemes developed within the group. This was mainly for convenience as compounds **7** and **21** just had to be deprotected at the 5'-hydroxyl position. However, this reaction proved to be surprisingly difficult for both porphyrin monomers. The reactions were carried out using 3 % TCA in DCM as the acid to cleave the DMT protecting group, as this mixture is used widely in the DMT cleaving step of automated DNA synthesis. In the case of the rigidly linked porphyrin monomer **21**, the starting porphyrin was zinc metallated. However, the zinc is removed from the porphyrin when it is exposed to acidic conditions to leave the free base porphyrin. The reactions progressed quickly as would be expected, but a number of unwanted side products were formed. Separation of these undesired compounds was difficult and was not helped by the general insolubility of the final product in any eluent suitable for column chromatography. It was found that the deprotected porphyrin monomer with the flexible amide linker was more difficult to purify than the porphyrin monomer with the rigid linker. Many methods were attempted but none were completely satisfactory. Deprotection of the flexible amide linked porphyrin was found to go with a 78 % yield, whereas a 74 % yield was observed for the rigidly linked porphyrin. These yields were only achieved on a small scale (~50 mg), attempts on a larger scale resulted in significantly reduced yields. This meant that only a small amount of product could be

synthesised at a time. Due to the solubility issues of these products, it was found that the only suitable solvent for NMR analysis was pyridine-d_5.

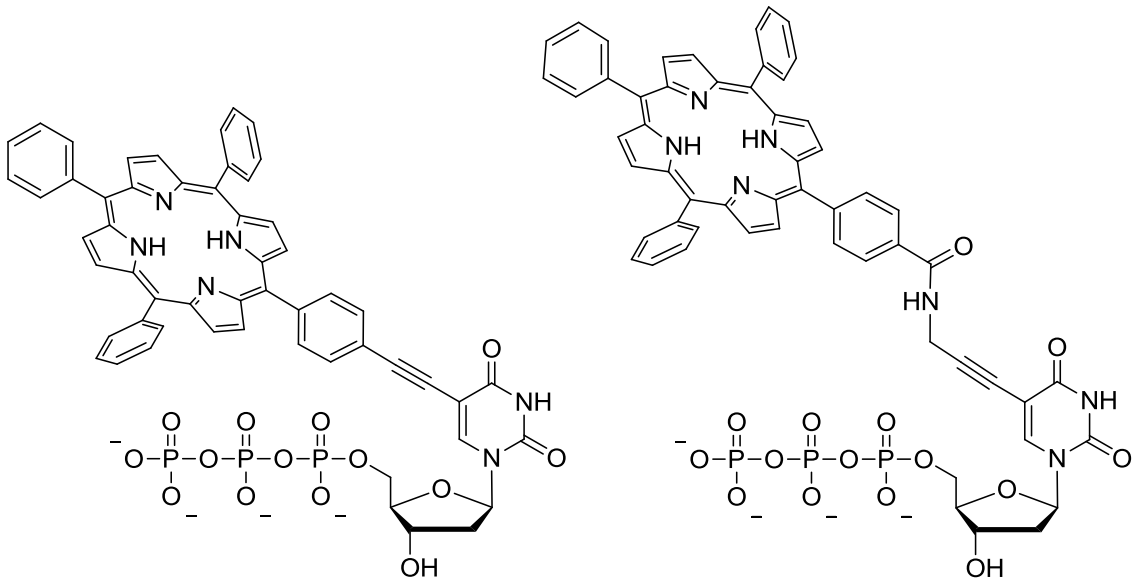


Figure 6.4 Porphyrin-modified nucleoside triphosphates

It would be useful for any future work with this project if the synthetic route could be improved so that larger amounts of porphyrin monomers **20** and **22** could synthesised in a more facile manner. The most obvious way of doing this would be to try the reactions without DMT protecting the 5'-hydroxyl of 5-iodo-2'-deoxyuridine. A Sonogashira cross-coupling reaction between the porphyrin alkyne, **12** and 5-iodo-2'-deoxyuridine, was carried out using a similar procedure to other porphyrin Sonogashira cross-coupling reactions that have been successful in the Stulz group before. The reaction took over 24 hours to go to completion but the following work up and purification was found to be facile compared to the previous synthesis of **22**. A decent yield of 76 % was obtained for this reaction and it can be performed on a larger scale than the previous route. The largest advantage from this synthesis is that it removes two steps from the previous route making it logical to synthesise this monomer in this manner in future. However, one major difference using this route is that the zinc ion is retained within the porphyrin ring. To synthesise the flexible amide porphyrin monomer **20** without protection of the 5'-hydroxyl group is less trivial. It has been shown to be possible to couple propargylamine onto 5-iodo-2'-deoxyuridine.[222] However, the subsequent amide coupling with the porphyrin carboxylic acid **2** has the potential to be difficult as the 5'-hydroxyl group can react with the porphyrin active ester to give an unwanted side product

Once in Prague, triphosphate synthesis was investigated. The first step of the reaction was to convert the 5'-hydroxyl to a monophosphate in a reaction with $POCl_3$. This reaction for both porphyrin-modified monomers was found to go to completion after 24 hours with the addition of twice as much $POCl_3$ than is usually required. This is generally not encouraged as there is a large risk of adding a monophosphate to the 3'-hydroxyl position. In both reactions a colour change from purple to green was observed when $POCl_3$ was added, due to protonation of the porphyrin. This is due to the $POCl_3$ reacting with the hydroxyl group and liberating HCl as a side product, which protonates the porphyrin to give a green colour. Once completed, the monophosphate was converted to the triphosphate in a step which is generally a facile process. The reaction was quenched with 2M TEAB (triethylamine bicarbonate) solution and the solvent removed *in vacuo*.

Purification of the triphosphates was completed by reverse-phase HPLC. HPLC purification was found to be difficult due to the very high number of impure porphyrin products that were present. Also both desired porphyrin triphosphates were eluted in neat methanol, which is not ideal in HPLC. For the rigid alkyne linked porphyrin, the triphosphate form **24** was collected and was found to be pure enough to consider carrying out further experiments with. The flexible amide linked porphyrin triphosphate **25** was obtained pure by HPLC with the exception of the presence of some trimethylphosphate. However, this could be removed by lyophilisation. Characterisation of the triphosphates was achieved using ^{1}H and ^{31}P NMR spectroscopy. We found the best way to the do this was using the triphosphate as a salt with triethylamine. When in this form the triphosphate can be dissolved in deuterated methanol, which gives a much clearer spectrum compared to when dissolved in deuterated water. This is due to the porphyrins tendency to aggregate in water, which leads to very broad signals in the ^{1}H NMR spectrum. Once characterised by NMR, the triphosphates were converted to the more stable sodium salt using Dowex 50WX8 in a Na^+ cycle.

We also attempted to synthesise the flexible amide linked porphyrin triphosphate using a water based Sonogashira cross-coupling reaction. The reaction usually takes place in a solution containing acetonitrile and water at temperatures between 60 ºC and 80 ºC. This is a problem for the coupling of the acetylene porphyrin because it is insoluble in this solvent system, although the addition of DMF can aid solubility. Our attempts at this were unsuccessful and this route was abandoned.

6.4 Primer extension using porphyrin-modified triphosphates

The first attempt at the primer extension experiment was carried out with the rigid linked porphyrin. We initially looked at using polymerases Vent (*exo*-) and KOD XL as the Hocek group have found that these two tend to incorporate C5 modified pyrimidine dNTPs into DNA reasonably well. Two templates were investigated, one which could be used to incorporate a single porphyrin dNTP (Oligo 1T) and one that could incorporate four porphyrin dNTPs (Oligo 4T)

Oligo 1T (1 modification)
5' – CCC ACC CAT GCC GCC GCC CAT G – 3'

Oligo 4T (4 modifications)
5' – CTA GCA TGA GCT CAG TCC CAT GCC GCC CAT G – 3'

Underlined bases indicate where the primer binds.
Red bases indicate where the porphyrin dNTP is incorporated.

The primer was labelled using γ-^{32}P ATP and purified prior to primer extension, to allow radiographic detection of the primer extension products after polyacrylamide gel electrophoresis. Three experiments were set up for each attempt at primer extension; the first a positive control which contains all four standard dNTPs, the second a negative control with the absence of thymidine triphosphate and the third which has the porphyrin-modified thymidine triphosphate plus the other three standard dNTPs. The concentration of all dNTPs remained the same in each experiment. The primer extension reaction was carried out at 60 ºC for 20 minutes for experiments with the Oligo 1T template and 30 minutes for experiments with the Oligo 4T template. The primer extension reaction was then quenched by addition of a stop solution (formamide, EDTA, bromophenol blue and xylene cyanol) and the samples denatured at 95 ºC for 5 minutes before separation by 12.5 % denaturing PAGE.

Analysis by denaturing polyacrylamide gel electrophoresis showed neither polymerase incorporated the rigid linked porphyrin dNTP into the DNA. We then tried a range of polymerases in an attempt to incorporate one porphyrin modification into the DNA (Figure 6.5). Once again no porphyrin incorporation was observed. As well as trying different polymerases, we tried increasing the concentration of the porphyrin and the enzyme but these experiments still showed no incorporation.

It has been observed in previous work that the length of the linker is important in whether the polymerase will incorporate the modified triphosphate. The rigid linked porphyrin has a short linker, and extending the linker could mean that the porphyrin could be incorporated. One issue we had with the rigid linked porphyrin triphosphate was the presence of a small amount of the monophosphate analogue present when the PEX studies were attempted. There is a chance that this impurity could inhibit the polymerase. Another observation from this experiment was the anomalous result using the Therminator polymerase. All experiments using this polymerase, including the negative control, gave products that were larger than the ones observed for the positive controls of the other five experiments. It was decided that there was a problem with this polymerase and it was not used in any further experiments.

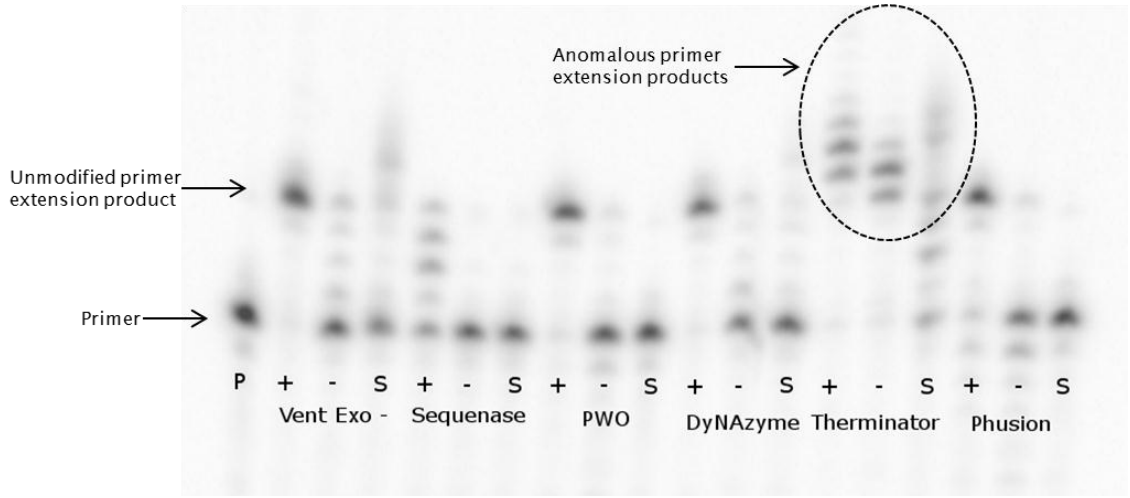


Figure 6.5 Denaturing PAGE results of incorporation of the rigid linked porphyrin triphosphate using the Oligo 1T template (1 modification). P = Primer, + = Positive control, - = Negative control and S = Primer extension with porphyrin triphosphate

PEX experiments with the flexible amide linked porphyrin triphosphate **25** were investigated next. We were able to purify this triphosphate successfully, which meant we would not have any issues with impurities inhibiting the polymerases. The first experiment was to try a single incorporation of **25** with a selection of polymerases (Figure 6.6). Denaturing PAGE showed that the polymerases Vent (*exo*-), KOD XL and DyNAzyme incorporated the flexible porphyrin triphosphate successfully into the DNA.

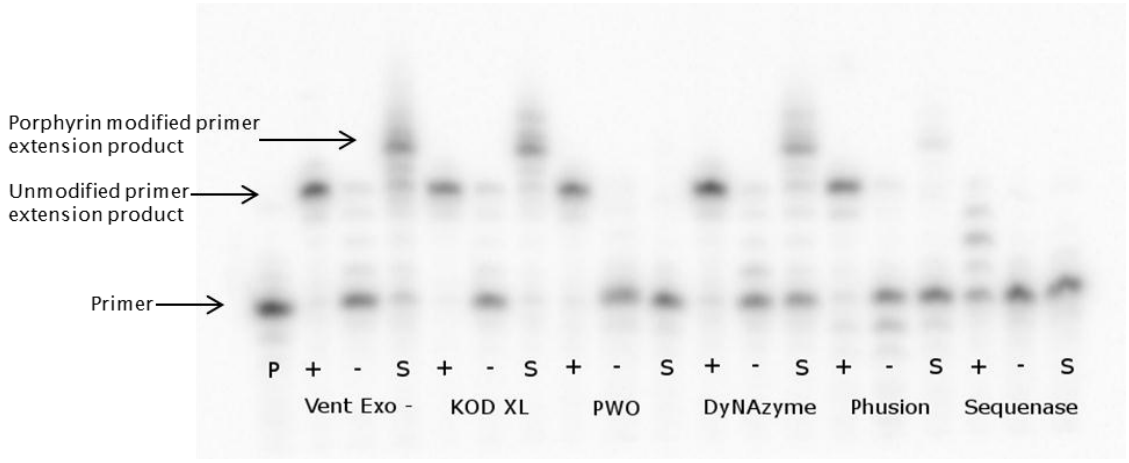


Figure 6.6 Denaturing PAGE results of incorporation of the flexible amide linked porphyrin triphosphate using the Oligo 1T template (1 modification). P = Primer, + = Positive control, - = Negative control and S = Primer extension with porphyrin triphosphate

We followed on from this experiment by attempting to incorporate four flexible porphyrin triphosphates into the DNA (Figure 6.7). Results from this experiment showed that the same three polymerases successfully incorporated the porphyrins into the DNA. By analysing the intensity of the bands on the gel, it can be assumed that KOD XL was the most efficient at incorporating the porphyrin dNTPs whereas DyNAzyme was the least efficient of the successful polymerases.

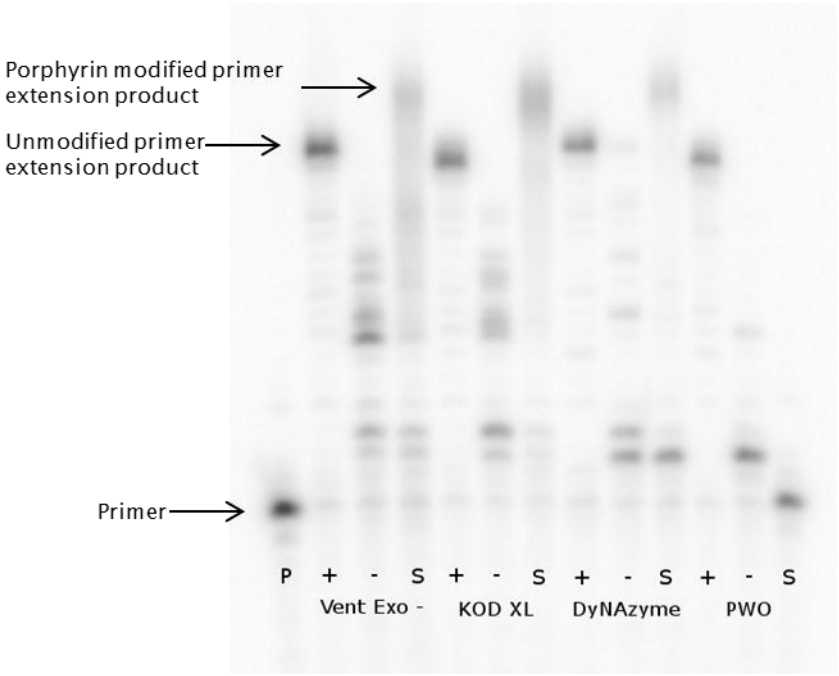


Figure 6.7 Denaturing PAGE results of incorporation of the flexible amide linked porphyrin triphosphate using the Oligo 4T template (4 modifications). P = Primer, + = Positive control, - = Negative control and S = Primer extension with porphyrin triphosphate

Primer extension studies with the rigid porphyrin triphosphate **24** were repeated once we had resynthesised the triphosphate and purified it to a higher standard. The ^{31}P NMR spectrum for this triphosphate was noticeably cleaner than the triphosphate we had previously used for this experiment. The primer extension was repeated using the same conditions as before with the same six polymerases. This time incorporation of the porphyrin was observed by four polymerases, Vent (*exo*-), KOD XL, DyNAzyme and Phusion during analysis by denaturing PAGE (Figure 6.8). This result proves that it is crucial that the triphosphate used for the primer extension is of a high purity to ensure adequate incorporation into the oligonucleotide. It also disproves our original hypothesis that the short length of the rigid linker was responsible for no previous incorporation.

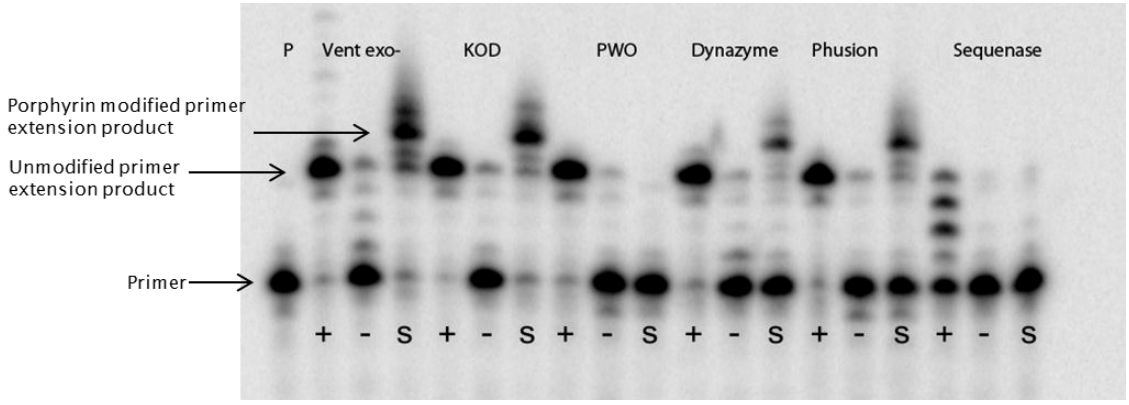


Figure 6.8 Denaturing PAGE results of incorporation of the rigid linked porphyrin triphosphate using the Oligo 1T template (1 modification). P = Primer, + = Positive control, - = Negative control and S = Primer extension with porphyrin triphosphate

As previously carried out with the flexible linker, we wanted to look further into how effective the successful polymerases are at incorporating the rigid porphyrin triphosphate **24** in the oligonucleotide. A primer extension experiment was carried out to incorporate four of the rigid porphyrin modifications into an oligonucleotide using the Oligo 4T template. The same conditions were used as for the analogous experiment with the flexible amide porphyrin triphosphate **25** and the results were analysed by denaturing PAGE (Figure 6.9). It was observed that the porphyrin had been incorporated successfully into the oligonucleotide by polymerases Vent (*exo*-) and KOD XL. The primer extension experiments with polymerases DyNazyme and Phusion were not as successful, although the gel does suggest that some of the porphyrin had been incorporated. This result is similar to what was observed with the flexible amide porphyrin triphosphate **25** in that Vent (*exo*-) and KOD XL were the only polymerases successful at incorporating a larger number of porphyrin modifications into the oligonucleotide. It is only worth using Vent (*exo*-) and KOD XL for any future experiments with these porphyrin triphosphates judging from the results we have obtained so far.

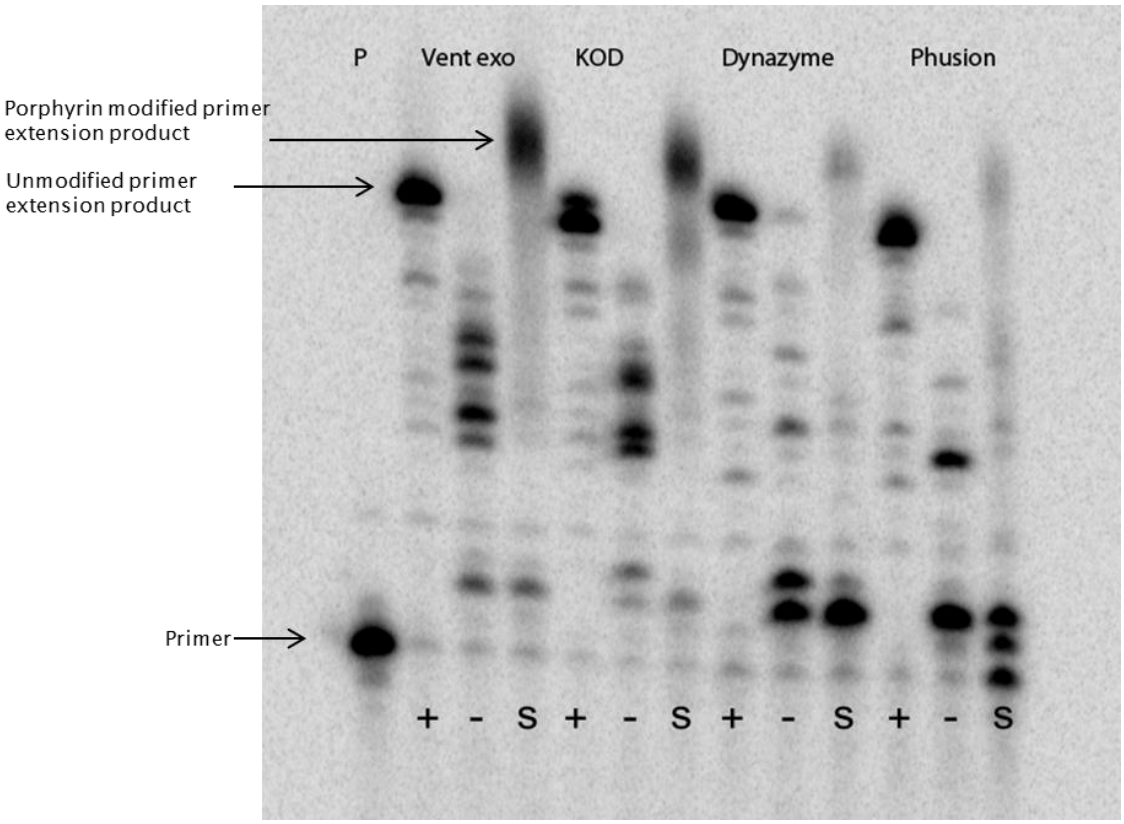


Figure 6.9 Denaturing PAGE results of incorporation of the rigid linked porphyrin triphosphate **24** using the Oligo 4T template (4 modifications). P = Primer, + = Positive control, - = Negative control and S = Primer extension with porphyrin triphosphate

Our final primer extension experiments concerned seeing how many porphyrin triphosphates with the flexible linker we could incorporate into an oligonucleotide (Figure 6.10). To start with we tried the incorporation of four consecutive porphyrins. It was thought that the polymerase may struggle to incorporate the porphyrin triphosphates when they are next to each other as the porphyrin could sterically hinder the polymerase. However, analysis of the gel showed that incorporation of the four porphyrins was successful. It should be noted that the band produced for the product on the gel was very broad, most likely due to the hydrophobic nature of the porphyrin, which has been observed before. The next experiment was to try seven porphyrins in a row. It is difficult to say whether this experiment was successful as the band produced is quite faint and even broader than when four porphyrins were added. More smearing of the product band would be expected due to the increased number of porphyrins present, but it would be fair to say this is not conclusive proof that seven porphyrins have been incorporated. The final experiment was to incorporate seven alternate porphyrins modifications into an oligonucleotide. This experiment was found to be much more successful than with consecutive porphyrin incorporations, with a much

stronger band for the product observed. Again, the product band has streaked across the gel, but this time it appears more likely that the polymerase has incorporated the porphyrins.

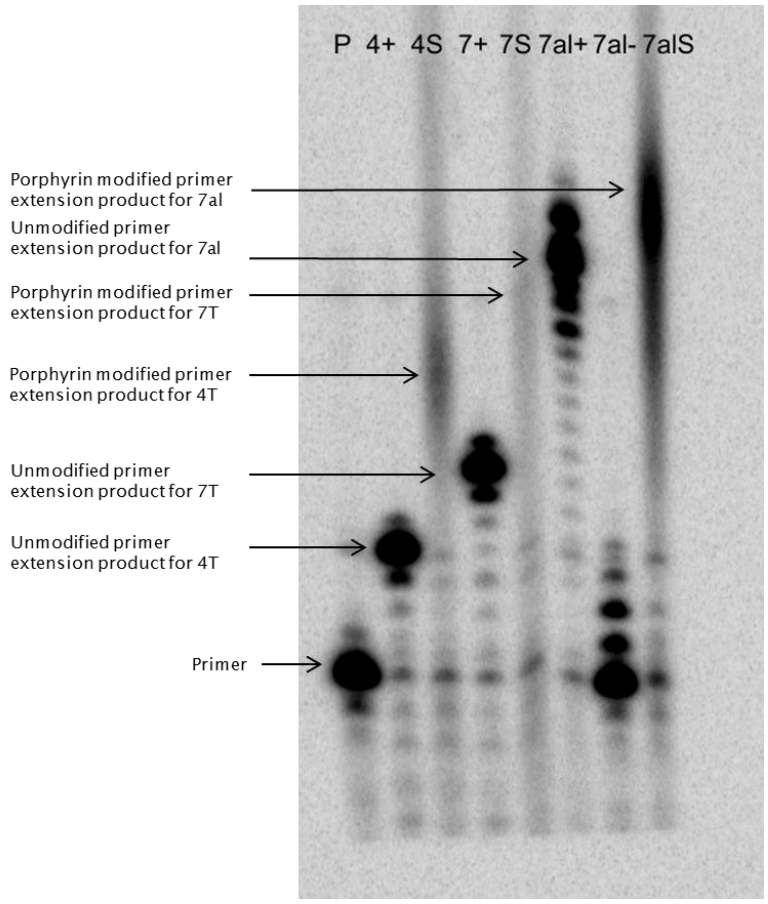


Figure 6.10 Denaturing PAGE results of incorporation of the flexible amide linked porphyrin triphosphate **25** using the 4T template (4 consecutive modifications), 7T template (7 consecutive modifications) and 7al T template (7 alternate modifications). + = Positive control, - = Negative control and S = Primer extension with porphyrin triphosphate

To gain further analysis of the primer extension products, mass spectrometry and UV-Vis spectroscopy were required. However, these methods require more oligonucleotide than was made in the previous experiments so a scale up was required. To start with, it was decided to incorporate one flexible amide porphyrin triphosphate **25** using the 1oligoT template with Vent (*exo*-) as the polymerase. The concentration of the template and primer was increased by a factor of ~ 100 and the triphosphates by a factor of ~ 2.5. The template was modified with biotin which would be used in the purification after the primer exchange.

It was observed during the primer exchange reaction that when the reaction mixture was heated to 60 ˚C the porphyrin triphosphate **25** began to aggregate and some porphyrin precipitated from the reaction solution. This clearly is a major problem as it will affect the efficiency of the porphyrin triphosphate being incorporated into the oligonucleotide. Once the time for the reaction was complete, the reaction was cooled and the stop solution added. The reaction mixture was heated to 95 ˚C for five minutes. The reaction was then cooled and streptavidin coated magnetic beads (Dynabeads) were added to bind to the biotin modified DNA template (Figure 6.11). Once bound, the double-stranded DNA (template and any extended primer) could be removed from the solution with a magnet and thoroughly washed to remove any impurities from the reaction mixture. The magnetic beads were then suspended in water and heated to 75 ˚C so the duplex DNA would separate into single strands. The magnetic beads containing the template could then be removed from the solution leaving just the extended primer in solution.

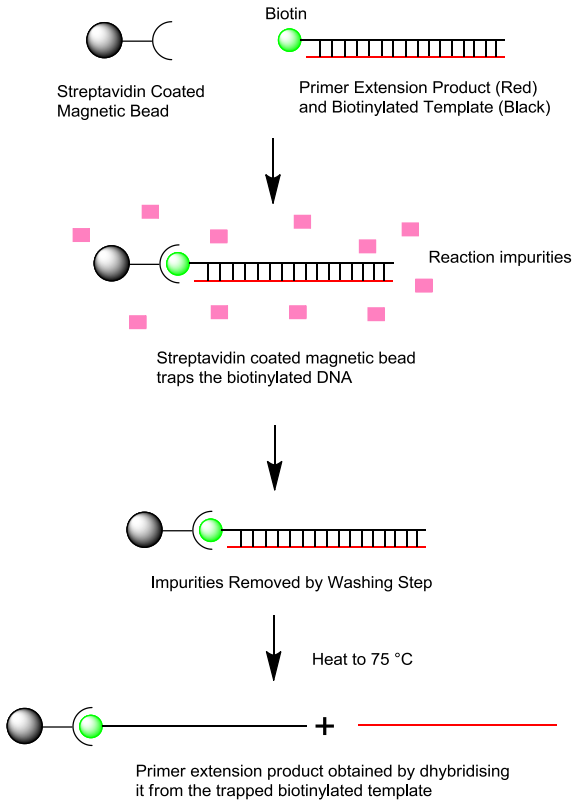


Figure 6.11 Purification of the primer extension product using streptavidin coated magnetic beads and a biotinylated template DNA strand

It was clearly noticeable from the colour of the primer extension product solution that very little porphyrin triphosphate had been incorporated. This was confirmed by UV-Vis spectroscopy where a large DNA peak was observed followed by a very small porphyrin Soret band peak. A drawback to the method of purification used for this experiment is that it will not remove any incomplete primer extension products. Because of this, the UV-Vis spectrum we obtained is mainly showing these impurities. Changing the conditions of the primer extension experiment such as changing the porphyrin triphosphate concentration failed to show any improvements. The final products were submitted for analysis by MALDI-TOF mass spectrometry but unsurprisingly the desired porphyrin DNA product was not observed. To observe the product would have been surprising using this technique, as past experiences with it in the Stulz group have been largely unsuccessful. It is difficult to determine why the porphyrin triphosphate aggregated during this experiment. One possibility is that a lot more polymerase and its required buffer were needed for this reaction as compared to the small scale reactions. This buffer contains a higher concentration of salt, so it is possible that an increase in ionic strength of the solution could cause the hydrophobic porphyrins to aggregate.

6.5 Conclusions and outlooks

In this short project we have demonstrated it is possible to incorporate 2'-deoxynuceloside triphosphates modified with porphyrins into oligonucleotides using various polymerases. It was observed that the length of the linker that attaches the porphyrin to the C5 position of the thymine nucleobase is not important for incorporation. This disproved our initial hypothesis that the shorter the linker, the less likely the polymerase will recognise the modified dNTP for primer extension on the grounds of steric hindrance. We have also shown that multiple porphyrin modifications can be incorporated into a single oligonucleotide, although the distance between each porphyrin modification does appear to have an effect on the efficiency of subsequent incorporations.

These initial advances in this area of porphyrin DNA synthesis are certainly promising, but there is still more work to do with this project. An improvement in the characterisation of the porphyrin DNA products would be an important area to look into. To start, the porphyrin triphosphate aggregation problem we observed upon the primer extension scale up must be solved or it will not be possible to synthesise enough material for thorough characterisation. It would also be interesting to attempt to use the porphyrin-modified dNTPs in a polymerase chain reaction in order to make a larger amount of porphyrin DNA. This study would be particularly interesting as after the first step the porphyrin modification would be incorporated into the template and

it is not known whether the polymerase will have trouble recognising this new base when extending the primer.

It would also be useful to compare the efficiency in the incorporation of the porphyrin-modified dNTPs with the natural ones. This can be performed with a simple kinetic study when incorporating one modified dNTP. The reactions with both the modified and natural dNTPs can be carried out and stopped at specific times. By comparing the resulting primer extension products by denaturing PAGE, it should be possible to qualitatively assess how efficient the polymerase is at incorporating the porphyrin-modified dNTPs.

One idea that relates this work to the study of G-quadruplexes from earlier would be to see if these porphyrin-modified dNTPs can be incorporated by the reverse transcriptase, telomerase, when extending the human telomere sequence. If it is possible, the inclusion of the porphyrins could induce the formation of G-quadruplex structures in the telomere sequence, hindering further telomere elongation. However, it remains to be seen if G-quadruplex formation would be possible when you consider there will be two porphyrins present in each loop. The steric hindrance due to the presence of many porphyrins could prevent the formation of a G-quadruplex. However, with the porphyrins incorporated it may be possible that the telomerase will not be able to interact with this modified telomere. With the porphyrins incorporated into the elongated telomere sequence, it may be possible to use photodynamic therapy to destroy cancer cells where telomerase is expressed. A porphyrin derivative called Photofrin has been utilised for this purpose. A red light source emitting at 630 nm is used to excite the porphyrin. This leads to the production of singlet oxygen, which is used to kill cancer cells.

This approach to synthesising porphyrin DNA has its pros and cons. The ability to synthesise a large oligonucleotide with many porphyrin modifications in a short time is certainly attractive. This is especially the case when comparing to standard solid phase oligonucleotide synthesis, where the time for synthesis is generally a lot longer and the synthesis and handling of the porphyrin monomer phosphoramidites is difficult. Another benefit of using the porphyrin-modified dNTPs is that they are relatively stable and can be stored in a freezer for a reasonable length of time. The opposite of this is true for the porphyrin-modified phosphoramidites, which are only stable for a very short period of time. This means that oligonucleotide synthesis must take place as soon as the phosphoramidite has been made. However, using modified triphosphates for DNA synthesis limits the amount of customisation that is possible for the oligonucleotide that is to be synthesised. This is because a porphyrin-modified dNTP will be incorporated at every position where the template strand directs it. It is not possible to synthesise a DNA strand containing a porphyrin-modified nucleobase and further along in the strand have the analogous unmodified base present, it is a

case of all or nothing. This is not a problem with solid phase oligonucleotide synthesis where pretty much any strand with multiple modifications in any position can easily be synthesised. Bearing this in mind, it is still advantageous to have different options available for when synthesising porphyrin-modified DNA.

7 Concluding remarks

In this project, we have developed the synthesis of porphyrin DNA and created new DNA structures using it. The synthesis of DNA with covalently bound porphyrins is a challenging process when you consider factors such as the stability of the porphyrin phosphoramidite and also the huge changes the porphyrin modification has on the properties of the DNA. Unfortunately, the stability of the porphyrin phosphoramidites will always be an issue with this work, but with careful working techniques this problem can be reduced. However, the purification of these oligonucleotides is a greater problem due to the highly hydrophobic nature of the porphyrin. We have shown HPLC purification is the best method for this, but it is still difficult and time consuming as only a small amount can be purified at once. It would be worth considering making the porphyrin less hydrophobic to make the purification process more facile. This could be achieved by either synthesising a cationic or anionic porphyrin monomer (Figure 7.1). Not only could these charged porphyrin monomers aid purification, but they may also help with further analysis using polyacrylamide gel electrophoresis. It was found that the hydrophobic nature of the porphyrins was causing large amounts of streaking on the gels. If we reduce the hydrophobic nature of the porphyrin we could possibly reduce the amount of streaking.

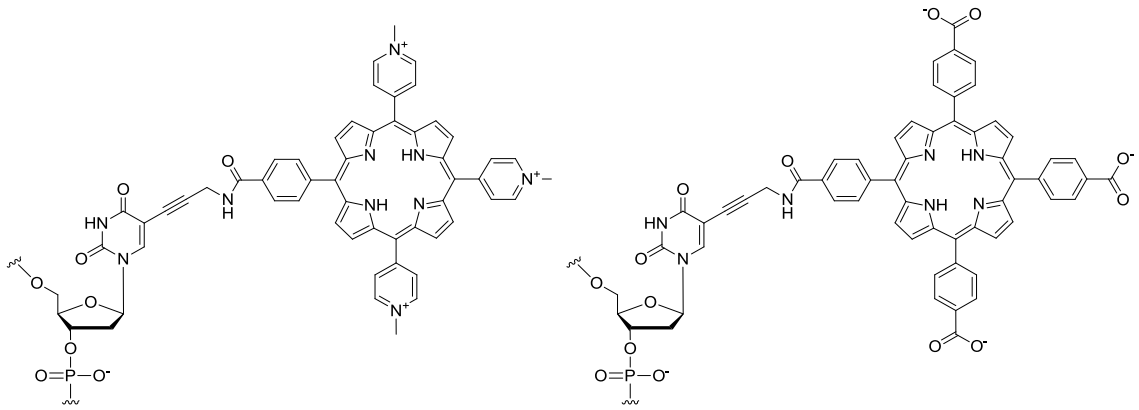


Figure 7.1 Examples of cationic and anionic porphyrin monomers

We have developed a new method of synthesising porphyrin DNA using polymerases to incorporate porphyrin-modified triphosphates. This method has its advantages in that the porphyrin triphosphate is more stable compared to the phosphoramidite and that the time it takes to carry out the synthesis of the porphyrin DNA is relatively short. However, the synthesis and purification of the porphyrin triphosphates was found to be particularly difficult. Another issue with this method was that when the polymerase extension experiment was scaled up, the porphyrin

triphosphate was not incorporated into the DNA very well. Further work will be needed to improve this method of synthesis, as at the moment it only works on a small scale.

As well as developing the techniques used to synthesise porphyrin-modified DNA, new structures were developed such as G-quadruplexes and LNA, that have expanded on the original work carried out with porphyrin DNA duplexes. Porphyrins are well known to stabilise G-quadruplexes through non-covalent interactions. We have demonstrated that this stabilising effect is still present when the porphyrin is covalently attached to the G-quadruplex. One issue with studying the structure of G-quadruplexes arises from their polymorphic nature. We have seen that the porphyrin modification generally does not promote the formation of a single G-quadruplex based on the human telomeric sequence. However, there was an exception with one G-quadruplex with two porphyrins present that forms a parallel G-quadruplex. More structural analysis could be carried out by further NMR studies and also X-ray crystallography. However, these techniques require large amounts of high purity material, which for porphyrin-modified G-quadruplex DNA is very difficult to prepare.

Finally we have developed the first porphyrin-modified LNA systems. It was hoped that the structural rigidity provided by the LNA modification would help improve the stability of duplex DNA containing porphyrin modifications. However, current measurements have been unable to determine what effect the porphyrin has on the stability so further work must be carried out on this.

8 Experimental

8.1 General methods

8.1.1 Suppliers

Chemicals were supplied by Fisher Scientific, Sigma-Aldrich, Apollo Scientific, Link Technologies, Glen Research, and Berry and Associates and used as received. DNA purification columns were supplied by Glen Research and Berry and Associates. Spin filters were supplied by Costar. Ultrafiltration units were supplied by Millipore. Nylon filtration membranes were supplied by Supelco. Desalting columns were supplied by GE Healthcare.

8.1.2 Column chromatography and TLC

Column chromatography was carried out using silica gel (Kieselgel 60), silica gel type H and/or basic alumina (50-200 μm, Brockmann activity I). TLC was carried out on Merck aluminium backed sheets of silica gel 60 F_{254} or aluminium backed sheets of alumina 60 F_{254} and were visualised using UV light (254 nm and 365 nm), anisaldehyde, phosphomolybdic acid (5 % in ethanol), Mary's Reagent (4,4'*Bis*-dimethylamino benzhydrol, 0.4 g in acetone, 100 mL), iodine on silica, and potassium permanganate in water.

8.1.3 NMR spectroscopy

Proton and carbon NMR spectra were recorded at 300/400 MHz and 75/100 MHz, respectively, using either a Bruker AC300 or Bruker DPX400 spectrometer. Fluorine NMR spectra were recorded at 282 MHz using a Bruker AC300 spectrometer. Chemical shifts are given in ppm and spectra are calibrated to the residual solvent peak. Coupling constants (*J*) are given in Hertz (Hz). Assignment was aided by DEPT-135, ^{1}H-^{1}H COSY, HMQC and HMBC experiments and Lorentz-Gauss resolution enhancement data reprocessing (ACD/SpecManager 12.0, ACD Labs).

8.1.4 Mass spectrometry

Low resolution electrospray mass spectrometry was conducted on a Walters ZMD (in house). MALDI-TOF was conducted on a ThermoBioAnalysis Dynamo using a *p*-nitroaniline matrix and referenced against TPP (Mw – 614.25) and 2,8,12,18-tetrahexyl-

3,7,13,17-tetramethyl-5,15-di(*p*-(3-hydroxy-3-methyl but-2-ynyl)phenyl porphyrin (Mw – 1082.88) or using a Micromass TOFSpec2E using external calibrants of terfenadine, bradykinin, angiotensin 1, substance P, renin substrate and ACTH clip for masses under 5000 Da, for larger molecules a range of oligonucleotides ranging from 5000 to 15000 Da are used.

8.1.5 Infrared spectroscopy

Infrared spectra were collected using a Thermo Electron Corporation Nicolet 380 FT-IR, spectra were obtained of samples in the solid state between 500 and 4000 cm^{-1}.

8.1.6 UV-visible spectroscopy

UV-Vis spectroscopy experiments were conducted using a Varian Cary 300 Bio spectrophotometer using quartz cells (supplied by Hellma and Starna) with 1 mm, 2 mm or 1 cm path lengths. The temperature of experiments was controlled using a Varian Cary Temperature Controller and peltier system with a Varian Cary Series II Temperature Probe.

8.1.7 Fluorescence spectroscopy

Fluorescence spectroscopy experiments were conducted using a Varian Cary Eclipse spectrophotometer using quartz cells (supplied by Hellma and Starna). Excitation for fluorescence experiments was carried out at the samples λ_{max}. The temperature of experiments was controlled using a Varian Cary Temperature Controller and peltier system with a Varian Cary Series II Temperature Probe.

8.1.8 DNA synthesis

DNA synthesis was carried out on an Applied Biosystems Expedite machine using 500 Å pore CPG beads. DNA synthesis reagents (SAFC or Link Technologies) were used as received. Deblocking steps used 3 % TCA in DCM solution, activation steps use 0.1 M 'Activator 42' (5-(*bis*-3,5-trifluoromethylphenyl)-1*H*-tetrazole) in MeCN, capping steps used acetic anhydride in THF (Cap A) and pyridine and NMI in THF (Cap B), oxidizing steps used 0.02M iodine, pyridine and water in THF, washing steps used MeCN. Cleavage of the oligonucleotide from the solid support was achieved by passing concentrated ammonium hydroxide through the column using two syringes at either

end for one hour at room temperature. This solution was then heated to 50 °C for 4 hours to deprotect the oligonucleotide.

8.1.9 Drying of DNA samples

DNA samples were dried using an Eppendorf Concentrator 5301 using the appropriate setting for the solvent at room temperature.

8.1.10 Centrifugation

Centrifugation was achieved using a Techno mini or Eppendorf 5415D centrifuge.

8.1.11 Thermomixing

Agitation of samples (with or without heating) was done using either an Eppendorf Thermomixer Compact or Eppendorf Thermomixer Comfort.

8.1.12 Pipetting

Micropietting was carried out using Gilson Pipetman, Socorex Acura 825 or Fisherbrand micropippetes. Microsyringing was conducted using Hamilton or SGE microsyringes.

8.1.13 Sonication

Sonication was conducted using a Fisherbrand 37 kHz, 120 W ultrasonic bath.

8.1.14 pH measurement

pH values were determined using a Metrohm pH Meter 744.

8.2 DNA purification methods

8.2.1 Glen-Pak™ cartridge purification

Glen-Pak™ cartridges can be used to purify oligonucleotide sequences which have been synthesised with the 5'-DMT on. The DMT group will bind to the solid support, which allows the removal of failure sequences. After this, the DMT can be removed allowing for the elution of the purified product.

Following DNA synthesis, the DMT-on oligonucleotide is deprotected in ammonium hydroxide (1.0 mL). Sodium chloride solution (1 mL, 100 mg/mL) is added to the deprotected DMT-on oligonucleotide for a final volume of 2.0 mL. Next, the cartridge needs to be prepared. A female luer port with a 5 ml disposable syringe to act as an eluent reservoir is attached to the cartridge. The flow rate can be controlled via the syringe. The column is conditioned with acetonitrile (0.5 mL) followed by 2M TEAA (1.0 mL). The oligo/salt mixture is applied to the cartridge in 1.0 mL aliquots, DMT-ON oligos should stick to the column packing material and be retained.

The cartridge is then washed with salt wash solution (5 % acetonitrile in 100 mg/mL sodium chloride solution, 2 x 1.0 mL) to rinse away the remainder of failure sequences. The cartridge is next rinsed with 2 % TFA (2 x 1.0 mL) to remove the DMT group from the bound oligonucleotide. The cartridge is then washed with deionised water (2 x 1.0 mL) to remove the TFA and any excess salts. The desired oligonucleotide is washed out with 50 % acetonitrile in water containing 0.5 % ammonium hydroxide (1 x 1.0 mL). A second elution is usually performed to ensure all of the desired oligonucleotide is collected.

8.2.2 Fluorous affinity purification of oligonucleotides using Fluoro-Pak™ II columns

Fluorous affinity purification was used to purify porphyrin-modified oligonucleotides. Glen-Pak™ columns could not be used without further HPLC purification as the porphyrin, like the DMT group, is a large aromatic molecule that could be retained by the column. This would mean truncated porphyrin containing oligonucleotides would be retained on the column and then eluted with the final desired product. Fluorous affinity purification requires the oligonucleotides to have a fluorous dimethoxytrityl (FDMT) at the 5' terminus as explained previously. Fluoro-Pak™ columns contain an adsorbent that has fluorinated organic groups bound to a pH-stable resin. The Fluoro-

Pak™ II columns contain 150 mg of adsorbent and are designed for 1 micromole purifications.

After synthesis, the FDMT-on oligonucleotide is deprotected in ammonia. The deprotected oligonucleotide is diluted with an equal volume of loading buffer (Berry & Associates, contains water, DMF and NaCl) so that the final sample volume is between 2 and 6 mL. A luer syringe adaptor is attached to the column to allow addition of liquids to the column via syringe needle. The syringe can be used to control the flow rate of liquid through the column. A flow rate of 2 seconds per drop should be used except for when loading the sample.

The Fluoro-Pak™ II column is conditioned with acetonitrile (2 mL), then 0.1M aqueous TEAA (2 mL) and finally loading buffer (2 mL). The FDMT-on oligonucleotide is then added onto the column at a flow rate of 5 seconds per drop. The slow flow rate is essential to ensure all oligonucleotide is bound to the column. The failure sequences are eluted from the column with 10 % acetonitrile in 0.1 M aqueous TEAA (2 mL) and then deionised water (2 mL).

Detritylation occurs next, on the column with the addition of 3 % aqueous TFA (3 mL), followed by 0.1 M aqueous TEAA (1 mL) and then deionised water (1 mL). The final detritylated oligonucleotide can then be eluted from the column with 20 % acetonitrile in water (1 mL). 50 % acetonitrile in water (1 mL) and then methanol (1 mL) are used to elute any remained oligonucleotide from the column if required.

8.2.3 Purification, desalting and buffer exchange of oligonucleotides using NAP columns

NAP columns, from GE Healthcare, are disposable columns prepacked with Sephadex™ G-25 DNA grade resin and require only gravity to run. They are principally used for desalting or buffer exchange of oligonucleotides. They are available in a number of different cross-sections for use with different volumes of sample solution. NAP-5 columns were used for purification.

Molecules larger than the pore size in the matrix are excluded from the matrix and have a shorter path length through the gel bed and elute first. Smaller molecules elute last as they can enter the pores of the matrix, which increases the path length they must take through the column.

The NAP column is prepared by first removing the top and bottom caps. The liquid inside is allowed to flow out of the column. The column is then equilibrated with deionised water (10 ml). The oligonucleotide solution is added to the column (maximum volume of 0.5 mL). If less than 0.5 mL of sample is added, make the volume up to 0.5 mL with deionised water. The oligonucleotide is then eluted with an appropriate amount of deionised water.

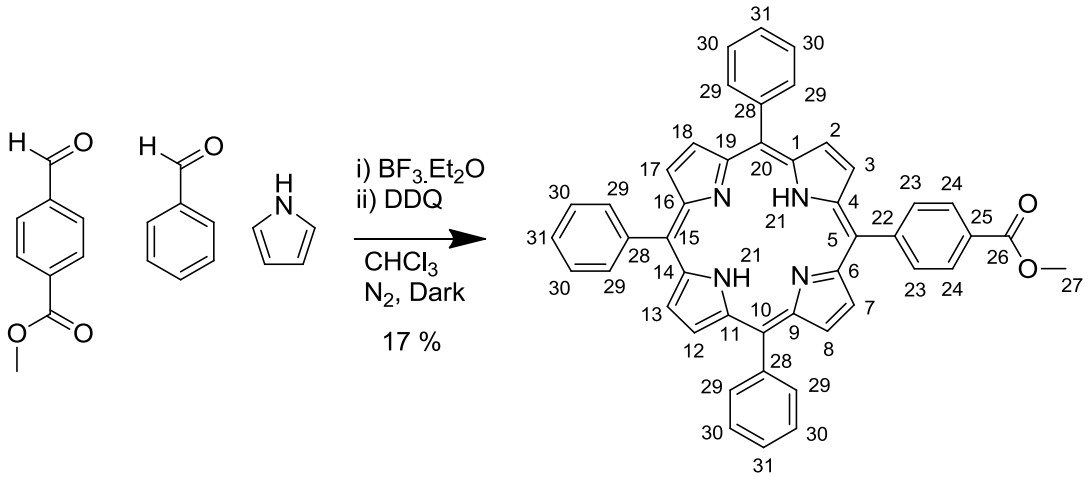
8.2.4 High pressure liquid chromatography (HPLC) of DNA

HPLC was carried out using a Varian Galaxie system using a Waters XBridge OST C18, 2.5 µm 4.6 x 50mm column. Eluents used were 8.6 mM TEA / 100 mM HFIP buffer and MeOH. Eluents were filtered through a Supelco Nylon 66 Membrane filter (0.45 µm pore size) before use. Flow rates were set to 1 mL min^{-1}. Eluent gradients varied and are stated with the relevant data.

8.3 Synthesis

Experiments conducted under an inert atmosphere were carried out using oven and/or flame dried glassware and purged of air using standard vacuum manifold techniques. 4 Å molecular sieves were oven-dried at greater than 120 °C prior to use. Chloroform was neutralised by filtering through basic alumina when required.

8.3.1 Synthesis of 5-(*p*-methylbenzoate)-10,15,20-triphenyl porphyrin (1)[82]



1

Pyrrole (2.50 mL, 36.0 mmol, 6.0 eq.), benzaldehyde (3.60 mL, 36.0 mmol, 6.0 eq.) and methyl-*p*-formylbenzoate (0.96 g, 6.0 mmol, 1.0 eq.) were dissolved in chloroform (500 mL) and purged with N_2 for 1 hr in the dark before boron trifluoride etherate (0.69 mL, 5.4 mmol, 0.9 eq.) was added and the reaction allowed to stir at room temperature. After one hour DDQ (8.14 g, 36.0 mmol, 6.0 eq.) was added and left to stir overnight. The reaction mixture was concentrated *in vacuo* before purification by column chromatography (alumina, eluent – DCM) to isolate the porphyrin containing fractions. The product was isolated by column chromatography (silica, eluent – toluene) and crystallised from chloroform and petroleum ether (0.68 g, 1.01 mmol, 17 %).

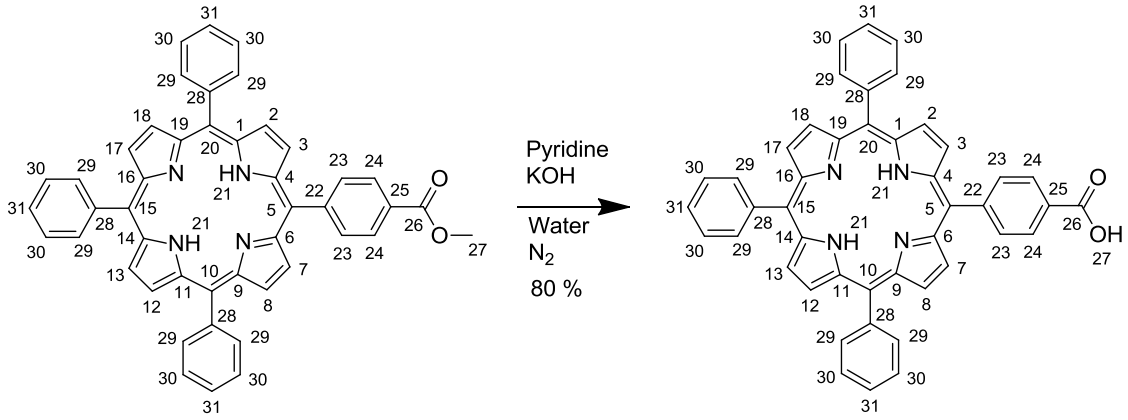
$\mathbf{R}_f$ 0.79 (silica, DCM), 0.38 (silica, toluene)

^{1}**H NMR** (400 MHz, $CDCl_3$): δ = 8.88 - 8.95 (m, 6H, **H**^{2}, **H**^{8}, **H**^{12}, **H**^{13}, **H**^{17}, **H**^{18}), 8.83 - 8.87 (m, 2H, **H**^{3}, **H**^{7}), 8.49 (d, *J* = 8.1 Hz, 2H, **H**^{24}), 8.36 (d, *J* = 8.1 Hz, 2H, **H**^{23}), 8.27 (d, *J* = 5.6 Hz, 6H, **H**^{29}), 7.73 - 7.88 (m, 9H, **H**^{30}, **H**^{31}), 4.15 (s, 3H, **H**^{27}), -2.70 (br. s., 2H, **H**^{21})

^{13}C NMR (100 MHz, $CDCl_3$): δ = 167.8 (C, **C26**), 147.5 (C, **C25**), 142.5 (C, **C5**, **C10**, **C15**), 135.0 (CH, **C29**), 131.8 (C, **C3**, **C7**), 131.3 (C, **C1**,**C9**, **C11**, **C14**, **C16**, **C19**), 130.0 (C, **C4**, **C6**), 129.3 (C, **C5**), 128.4 (CH, **C24 / C23**), 128.2 (C, **C C24 / C23**), 127.2 (CH, **C30, C31**), 121.0 (C, **C25**), 120.9 (C, **C25**), 119.0 (C, **C22**), 52.9 (OCH_3, **C27**)

MALDI-TOF ($C_{46}H_{32}N_4O_2$): Calculated mass 672.77; observed mass 673.65 $[M+H]^+$

8.3.2 Synthesis of 5-(*p*-benzoic acid)-10,15,20-triphenyl porphyrin (2)[82]



1 **2**

5-(*p*-methylbenzoate)-10,15,20-triphenyl porphyrin **1** (630 mg, 0.93 mmol, 1.0 eq.) was dissolved in pyridine (15 mL) under a nitrogen atmosphere. Potassium hydroxide (2.62 g, 46.7 mmol, 50.0 eq.) and distilled water (1.5 mL) were added to the reaction and the reaction stirred in the dark at 40 °C for 24 hr. The reaction mixture was poured into distilled water (300 mL) and DCM (300 mL). Dilute hydrochloric acid (0.1 M) was added until the porphyrin had been completely neutralised and found in the organic phase. The organic phase was then washed with water (2 × 300 mL), dried (Na_2SO_4) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica, eluent – DCM → 9:1 DCM/ MeOH). The purple product was collected and dried *in vacuo* (700 mg, 1.06 mmol, 80 %).

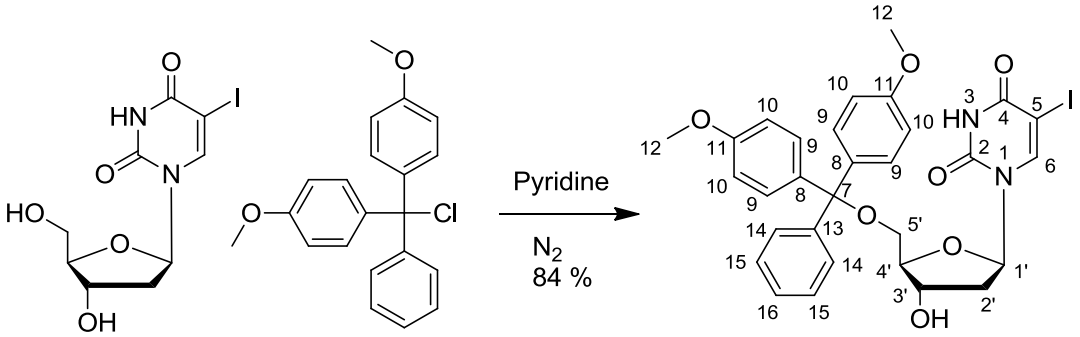
R_f 0.28 (silica, DCM + 10 % MeOH)

^{1}**H NMR** (400 MHz, $CDCl_3$): δ = 8.84 - 8.92 (m, 6H, **H**^{2}, **H**^{8}, **H**^{12}, **H**^{13}, **H**^{17}, **H**^{18}), 8.80 - 8.84 (m, 2H, **H**^{3}, **H**^{7}), 8.51 (d, *J* = 8.7 Hz, 2H, **H**^{24}), 8.37 (d, *J* = 8.2 Hz, 2H, **H**^{23}), 8.22 (d, *J* = 7.4 Hz, 6H, **H**^{29}), 7.71 - 7.84 (m, 9H, **H**^{30}, **H**^{31}), -2.75 (br. s., 2H, **H**^{21})

^{13}**C NMR** (100 MHz, $CDCl_3$): δ = 168.9 (C, **C26**), 146.8 (C, **C25**), 141.8 (C, **C5**, **C10**, **C15, C20**), 134.4 (CH, **C23, C29**), 128.0 (CH, **C31**), 127.7 (CH, **C24**), 126.6 (CH, **C30**), 120.2 (C, **C28**), 118.3 (C, **C22**)

MALDI-TOF ($C_{45}H_{30}N_4O_2$): Calculated mass 658.75, 658.55 observed mass $[M+H]^+$

8.3.3 Synthesis of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine (3)[49]



3

5-iodo-2'-deoxyuridine (5.00 g, 14.1 mmol, 1.0 eq.) was dried by co-evaporation with distilled pyridine (2 × 10 mL) under high vacuum, and then dissolved in distilled pyridine (30 mL). 4,4'-dimethoxytrityl chloride (5.02 g, 14.8 mmol, 1.05 eq.) was added in six portions over 5 hr in an inert atmosphere with constant stirring. An orange solution was observed when the 4,4'-dimethoxytrityl chloride was first added which turned to a yellow colour after around 20 mins. The reaction mixture was stirred for a further 30 mins and then quenched with methanol/water (1:1 v/v, 5 mL), stirring for 10 mins at rt. The reaction mixture was concentrated *in vacuo* and by co-evaporation with toluene to give an orange syrup. The syrup was then dissolved in DCM (150 mL), washed with water (2 × 150 mL), sat. KCl (150 mL), dried (Na_2SO_4) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica pretreated with TEA, eluent – DCM/hexane 8:2 → DCM → DCM + 1 % MeOH). A second column was needed to complete purification (silica pretreated with TEA, eluent – ethyl acetate/hexane 7:3 → ethyl acetate + 1 % MeOH → ethyl acetate + 5 % MeOH). The solvent was removed by rotary evaporation and the product dried under high vacuum to give a white foam (7.82 g, 11.9 mmol, 84 % yield).

$\mathbf{R}_f$ 0.19 (silica, DCM + 5 % MeOH)

^{1}**H NMR** (400 MHz, DMSO-d_6): δ = 11.73 (br. s, 1H, N**H**), 8.01 (s, 1H, **H**^{6}), 7.40 (d, *J* = 7.5 Hz, 2H, **H**^{14}), 7.32 (t, *J* = 7.4 Hz, 2H, **H**^{15}), 7.29 (d, *J* = 8.9 Hz, 4H, **H**^{9}), 7.22 (t, *J* = 7.3 Hz, 1H, **H**^{16}), 6.90 (d, *J* = 8.8 Hz, 4H, **H**^{10}), 6.10 (t, *J* = 6.8 Hz, 1H, **H**$^{1'}$), 5.30 (br. s., 1H, 3'O**H**), 4.23 (br s, 1H, **H**$^{3'}$), 3.87 - 3.93 (m, 1H, **H**$^{4'}$), 3.74 (s, 6H, OC**H**$_3$), 3.19 (dd, *J*=10.5, 4.9 Hz, 1H, **H**$^{5'}$), 3.16 (dd, *J*=10.6, 3.3 Hz, 1H, **H**$^{5'}$), 2.24 (dt, *J*=13.4, 6.9 Hz, 1H, **H**$^{2'}$), 2.19 (ddd, *J*=13.3, 6.0, 3.4 Hz, 1H, **H**$^{2'}$)

^{13}**C NMR** (100 MHz, DMSO-d_6): δ = 161.1 (C, **C4**), 158.6 (C, **C11**), 150.5 (C, **C2**), 145.5 (C, **C13**), 144.6 (CH, **C6**), 135.9 (C, **C8**), 130.2 (CH, **C9**), 128.4 (CH, **C15**), 128.1 (CH, **C14**), 127.2 (CH, **C16**), 113.8 (CH, **C10**), 86.4 (C **C7**), 86.3 (CH, **C4'**), 85.4 (CH, **C1'**), 71.2 (CH, **C3'**), 64.3 (CH_2, **C5'**), 55.6 (CH_3, **C12**), 40.4 (CH_2, **C2'**)

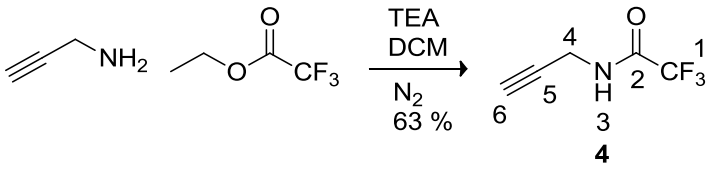
ESI Positive ($C_{30}H_{29}IN_2O_7$): Calculated mass 656.46, observed *m/z* 679 $[M+Na]^+$

ESI negative ($C_{30}H_{29}IN_2O_7$): Calculated mass 656.46, observed *m/z* 655.1 $[M-H]^-$

UV-Vis (MeCN, 38 μM): λ_{max} (log ε) 231 shoulder (4.27), 279 (3.89)

Emission (MeCN, 38 μM): No emission observed

8.3.4 Synthesis of 2,2,2-trifluoro-*N*-(prop-2-ynyl)acetamide (4)



4

Propargylamine (0.34 mL, 5.0 mmol, 1.0 eq.) and triethylamine (0.91 mL, 6.5 mmol, 1.3 eq.) were dissolved in anhydrous methanol (5.0 mL) and degassed with N_2 for 25 mins. The reaction mixture was cooled (0 °C), ethyl trifluoroacetate (0.77 mL, 6.5 mmol, 1.3 eq.) was added dropwise and the reaction mixture allowed to warm to room temperature and left to stir overnight. The reaction mixture was concentrated *in vacuo* before redissolving in DCM (40 mL), washing with 10 % w/v aq. citric acid (2 × 50 mL) sat. KCl (2 × 50 mL), dried (Na_2SO_4) and the solvent removed *in vacuo* to give an orange brown oil. Following column chromatography (silica, eluent – DCM), and drying *in vacuo*, the product was isolated as a clear oil (480 mg, 3.2 mmol, 63 %)

R_f 0.65 (silica, DCM + 10 % MeOH)

^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.83 (1H, br s, **H^3**), 4.08 (2H, d, *J* = 2.5 Hz, **H^4**), 2.27 (1H, t, *J* = 2.5 Hz, **H^6**)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 159.5 (q, *J* = 37.6 Hz, **2**), 117.1 (q, *J* = 287.5 Hz, **1**), 77.4 (C, **5**), 73.5 (CH, **6**), 30.1 (CH_2, **4**)

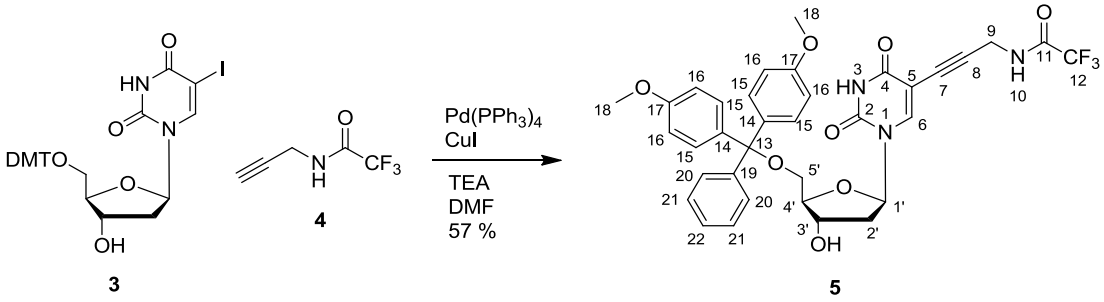
^{19}F NMR (282 MHz, $CDCl_3$): δ = 76.2 (s, **1**)

IR (film): *v* 3301 (br m, CC-H), 3090 (br w, N-H), 1704 (vs, C=O), 1552 (m, N-H), 1152 (vs, C-F), 725 (m, C-F)

ESI Positive ($C_5H_4F_3NO$): Calculated mass 151.09, observed *m/z* 206 $[M+CH_3OH+Na]^+$, 174 $[M+Na]^+$

ESI Negative ($C_5H_4F_3NO$): Calculated mass 151.09, observed *m/z* 150.1 $[M-H]^-$

8.3.5 Synthesis of 1'-*β*-5'-*O*-(4,4'-Dimethoxytrityl)-5-(3-trifluoroacetamidopropynyl)-2'deoxyuridine (5)[223]



To a stirred solution of 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'deoxyuridine **3** (250 mg, 0.38 mmol, 1.0 eq.) in anhydrous DMF (5.0 mL) under a nitrogen atmosphere was added triethylamine (0.37 mL, 2.67 mmol, 7.0 eq.) followed by, in absence of light, copper(I) iodide (15 mg, 0.08 mmol, 0.2 eq.) and 3-(trifluoroacetamido)prop-1-yne **4** (80 mg, 0.53 mmol, 1.4 eq.). After stirring for 20 mins, tetrakis(triphenylphosphine)palladium(0) (44 mg, 0.04 mmol, 0.1 eq.) was added and the reaction stirred at room temperature for 1.75 hr. The reaction mixture was diluted with ethyl acetate (100 mL) and washed with 5 % w/v Na_2EDTA (pH 8, 3 × 50 mL) and sat. NaCl (2 × 100 mL), dried (Na_2SO_4), and the solvent removed *in vacuo* to give a golden brown foam. The crude product was purified by column chromatography (silica, eluent – 1:1 ethyl acetate / toluene + pyridine). The desired product was collected, dried under high vacuum as a very pale orange foam (149 mg, 0.22 mmol, 57 %).

R_f 0.09 (silica, DCM + 10 % acetone + 0.5 % pyridine), 0.50 (silica, 1:1 ethyl acetate/toluene + 0.5 % pyridine)

^{1}**H NMR** (400 MHz, $CDCl_3$): δ = 9.42 (br. s., 1H, **H**^{10}), 8.24 (s, 1H, **H**^{6}), 7.42 (d, *J* = 7.9 Hz, 2H, **H**^{20}), 7.33 (d, *J* = 8.8 Hz, 4H, **H**^{15}), 7.28 - 7.32 (m, 2H, **H**^{21}), 7.23 (t, *J* = 7.4 Hz, 1H, **H**^{22}), 6.85 (d, *J* = 8.8 Hz, 4H, **H**^{16}), 6.36 (t, *J* = 6.6 Hz, 1H, **H**$^{1'}$), 4.57 - 4.64 (m, 1H, **H**$^{3'}$), 4.10 - 4.16 (m, 1H, **H**$^{4'}$), 3.86 - 4.00 (m, 2H, **H**^{9}), 3.80 (s, 6H, **H**^{18}), 2.98 - 3.10 (m, 2H, **H**$^{5'}$), 2.53 - 2.61 (m, 1H, **H**$^{2'}$), 2.30 - 2.39 (m, 1H **H**$^{2'}$)

^{13}**C NMR** (100 MHz, $CDCl_3$): δ = 162.5 (C, **4**), 159.1 (C, **17**), 157.2 (q, *J* = 37.50 Hz, C, **11**) 149.6 (C, **2**), 144.0 (CH, **6**), 135.8 (C, **14**), 130.4 (CH, **15**), 128.5 (C, **20**), 128.3 (C, **20**), 127.5 (C, **22**), 116.1 (C, **12**), 113.8 (CH, **16**), 99.5 (C, **8**), 87.8 (C, **13**), 87.6 (C, **7**), 87.3 (CH, **4'**), 86.4 (CH, **1'**), 72.8 (CH, **3'**), 63.9 (CH_2, **5'**), 55.7 (CH_3, **18**), 42.1 (CH_2, **2'**), 30.8 (CH_2, **9**)

^{19}**F NMR** (282 MHz, $CDCl_3$): δ = 75.8

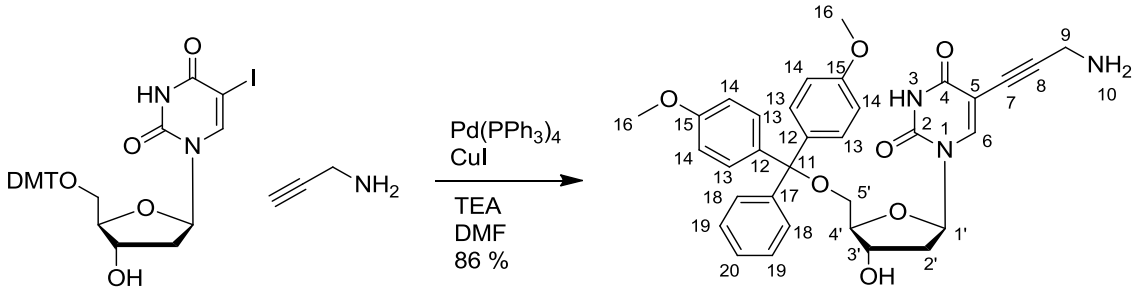
ESI Positive ($C_{35}H_{32}F_3N_3O_8$): Calculated mass 679.64, observed *m/z* 702.3 $[M+Na]^+$

ESI negative ($C_{35}H_{32}F_3N_3O_8$): Calculated mass 679.64, observed *m/z* 678.5 $[M-H]^-$

UV-Vis (MeCN, 25.7 μM): λ_{max} (log *ε*) 226 (6.39), 283 (2.01)

Emission (MeCN, 25.7 μM): No fluorescence observed

8.3.6 Synthesis of 5'-DMT-5-propargylamine-dU (6)[224]



5'-*O*-(4,4'-dimethoxytrityl-5-iodo-2'-deoxyuridine **3** (500 mg, 0.75 mmol, 1.0 eq.) was dissolved in DMF (5 mL) and degassed with nitrogen for 1 hr. In the absence of light, copper(I) iodide (36 mg, 0.19 mmol, 0.25 eq.) and propargylamine (103 µL, 1.50 mmol, 2.0 eq.) were added. After stirring for 20 mins, tetrakis(triphenylphosphine)palladium(0) (87 mg, 0.08 mmol, 0.1 eq.) followed by triethylamine (0.73 mL, 5.20 mmol, 7.0 eq.) were added and the reaction stirred for 1.75 hr. The reaction mixture was diluted with ethyl acetate (100 mL) and washed with 5 % w/v Na_2EDTA (pH = 9) (2 × 100 mL). The aqueous phases were back extracted with ethyl acetate (200 mL). The organic phases were combined, washed with sat. KCl (3 × 150 mL), dried ($MgSO_4$), solvent removed by rotary evaporation and dried under high vacuum to give a yellow / white foam. The crude product was purified by column chromatography (silica pretreated with 1 mL TEA, eluent: DCM + 5 % MeOH) to give the desired product as a white / yellow solid (376 mg, 0.38 mmol, 86 %).

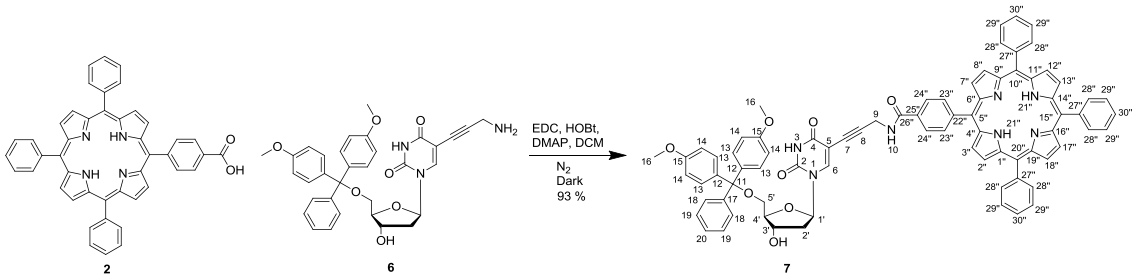
R_f 0.33 (silica, DCM + 10 % MeOH)

^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.15 (s, 1H, **H^6**), 7.44 (d, *J* = 7.6 Hz, 2H, **H^{18}**), 7.34 (d, *J* = 7.7 Hz, 4H, **H^{13}**), 7.20 - 7.15 (m, 2H, **H^{19}**), 7.18 (t, *J* = 7.2 Hz, 1H, **H^{20}**), 6.83 (d, *J* = 8.5 Hz, 4H, **H^{14}**), 6.31 (t, *J* = 6.6 Hz, 1H, **H$^{1'}$**), 4.51 (br s, 1H, **H$^{3'}$**), 4.10 (br s, 1H, **H$^{4'}$**), 3.74 (br s, 6H, **H^{16}**), 3.40 (d, *J* = 9.0 Hz, 1H, **H$^{5'}$**), 3.28 (d, *J* = 7.9 Hz, 1H, **H$^{5'}$**), 3.19 (br s, 2H, **H^9**), 2.60 - 2.43 (m, 1H, **H$^{2'}$**), 2.35 - 2.21 (m, 1H, **H$^{2'}$**)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 161.5 (C, **4**), 157.6 (C, **15**), 148.7 (C, **2**), 143.7 (C, **17**), 141.7 (CH, **6**), 134.7 (C, **12**), 134.6 (C, **12**), 129.1 (CH, **13**), 129.0 (CH, **13**), 127.0 (CH, **19**), 125.9 (CH, **18**), 112.4 (CH, **14**), 99.0 (C, **7**), 93.2 (C, **8**), 85.9 (C, **11**), 85.7 (CH, **4'**), 84.8 (CH, **1'**), 72.5 (C, **5**), 70.7 (CH, **3'**), 62.6 (CH_2, **C5'**), 54.3 (CH_3, **1**), 40.6 (CH, **2'**), 30.6 (CH_2, **9**)

ESI positive ($C_{33}H_{33}N_3O_7$): Calculated mass 583.63, observed *m/z* 584.3 $[M+H]^+$

8.3.7 Synthesis of *N*-(5'DMT-5-propargyl-dU)-5"-(*p*-benzamide)-10",15",20"-triphenyl-21"-*H*-23"-*H*-porphyrin (7)[82]



5-(*p*-methylbenzoate)-10,15,20-triphenyl porphyrin **2** (115 mg, 0.18 mmol, 1.0 eq.) was dissolved in anhydrous DCM (6 mL). The mixture was purged with N_2 whilst stirred in the absence of light. Propargylamine deoxyuridine **6** (131 mg, 0.23 mmol, 1.29 eq.), HOBt (28 mg, 0.18 mmol, 1.04 eq.), DMAP (42 mg, 0.35 mmol, 2.0 eq.) and EDC (48 µL, 0.35 mmol, 2.0 eq.) were added to the purple solution and stirred for 5 hr at room temperature. The crude reaction mixture was loaded directly onto a column (silica pretreated with TEA, eluent – DCM + 0.5 % MeOH → 5 % MeOH). The isolated product was coevaporatated with toluene (3 × 20 mL) and neutralised chloroform (3 × 20 mL) to give a purple solid (198 mg, 0.16 mmol, 93 %).

R_f 0.46 (silica, DCM + 10 % MeOH)

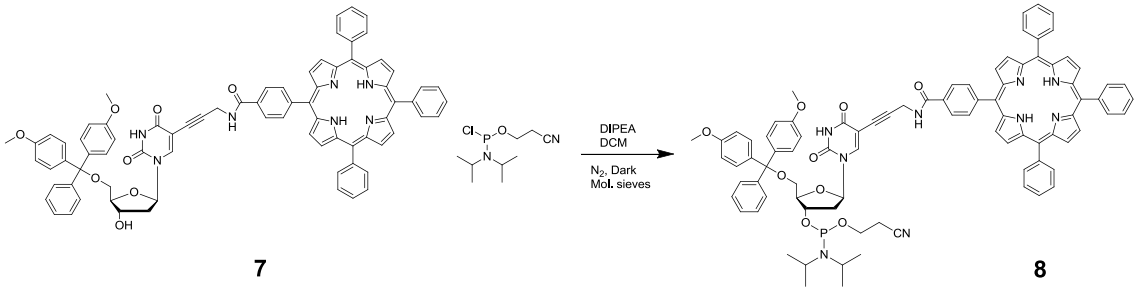
^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.84 - 8.86 (m, 6H, **H$^{2"}$**, **H$^{8"}$**, **H$^{12"}$**, **H$^{13"}$**, **H$^{17"}$**, **H$^{18"}$**), 8.75 - 8.79 (m, 2H, **H$^{3"}$**, **H$^{7"}$**), 8.31 (s, 1H, **H$^{6'}$**), 8.19 - 8.23 (m, 8H, **H$^{24"}$**, **H$^{28"}$**), 7.93 (d, *J* = 8.2 Hz, 2H, **H$^{23"}$**), 7.71 - 7.82 (m, 9H, **H$^{29"}$**, **H$^{30"}$**), 7.48 (d, *J* = 7.3 Hz, 2H, **H^{18}**), 7.38 (d, *J* = 8.9 Hz, 4H, **H^{13}**), 7.31 (t, *J* = 7.7 Hz, 2H, **H^{19}**), 7.23 (t, *J* = 7.0 Hz, 1H, **H^{20}**), 6.86 (d, *J* = 8.9 Hz, 4H, **H^{14}**), 6.42 - 6.48 (m, 1H, **H^{10}**), 6.38 (t, *J* = 6.4 Hz, 1H, **H$^{1'}$**), 4.55 - 4.59 (m, 1H, **H$^{3'}$**), 4.39 (dd, *J* = 17.9, 4.9, 1H, **H^{9}**), 4.36 (dd, *J* = 17.9, 4.9, 1H, **H^{9}**), 4.13 - 4.17 (d, *J* = 2.3, 1H, **H$^{4'}$**), 3.75 (s, 6H, **H^{16}**), 3.33 - 3.44 (m, 2H, **H$^{5'}$**), 2.53 - 2.60 (m, 1H, **H$^{2'}$**), 2.26 - 2.34 (m, 1H, **H$^{2'}$**), -2.77 (br. s., 2H, **H^{21}**)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 167.0 (C, **C26"**), 162.2 (C, **C4**), 158.6 (C, **C15**), 149.3 (C, **C2**), 145.6 (C, **C25"**), 144.5 (C, **C17**), 143.2 (CH, **C6**), 142.0 (C, **C5"**, **C10"**, **C15"**, **C20"**), 135.5 (C, **C12**), 134.5 (CH, **C28"**), 131.7 – 130.6 (CH, **C2"**, **C3"**, **C7"**, **C8"**, **C12"**, **C13"**, **C17"**, **C18"**) 129.9 (CH, **C13**), 128.1 (CH, **C19**), 127.8 (CH, **C20**), 127.7 (CH, **C18**), 127.0 (CH, **C30"**), 126.7 (CH, **C29"**), 125.5 (CH, **C23"**), 120.3 (C, **C27"**), 113.4 (CH, **C14**), 99.4 (C **C7**), 89.6 (C **C8**), 87.1 (CH, **43'**), 86.7 (CH, **C1'**), 86.0 (C, **C11**), 74.6 (C, **C5**), 72.2 (CH, **C3'**), 63.5 (CH_2, **C5'**), 55.2 (CH_3, **C16**), 41.6 (CH_2, **C2'**), 30.8 (CH_2, **C9**)

ESI Positive: ($C_{78}H_{61}N_7O_8$): Calculated mass 1223.46, observed *m/z* 1240.5 $[M+Na]^+$

ESI Negative: ($C_{78}H_{61}N_7O_8$): Calculated mass 1223.46, observed *m/z* 1222.3 $[M-H]^-$

8.3.8 Synthesis of *N*-(5'-DMT-5-propargyl-dU)-5"-(*p*-benzamide)-10",15",20"-triphenyl-21"-*H*-23"-*H*-porphyrin-3'-amidite (8)[82]



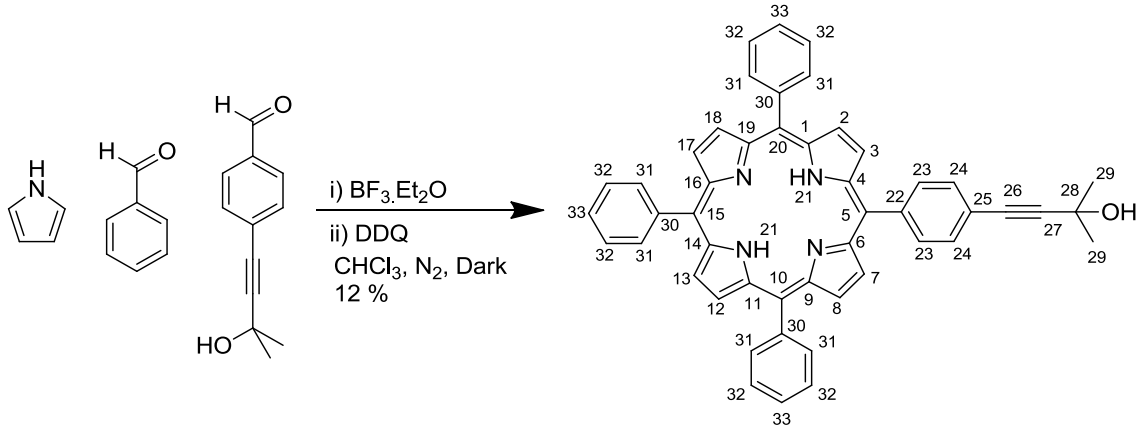
7 **8**

Using oven dried glassware under Schlenk conditions, *N*-(5'-DMT-5-propargyl-dU)-5''-(*p*-benzamide)-10'',15'',20''-triphenyl-21''-*H*-23''-*H*-porphyrin **7** (100 mg, 82 µmol, 1.0 eq.) was dissolved in anhydrous DCM (2 mL) with molecular sieves in the dark. DIPEA (57 µL, 327 µmol, 4.0 eq.) and CEP-Cl (58 µL, 245 µmol, 3.0 eq.) were added and the reaction allowed to stir at room temperature for 1.75 hr. The reaction mixture was filtered and concentrated to a volume of 1 mL. The product was precipitated from degassed hexane (20 mL) and collected by filtration. The purple solid was dried *in vacuo* (131 mg, 92 µmol, >100 %).

Full characterisation of this product was not achieved due to its instability. The product was immediately used for DNA synthesis.

R_f 0.71 (silica, DCM + 10 % MeOH + 0.5 % Triethylamine)
MALDI-TOF ($C_{87}H_{78}N_9O_9P$): Calculated mass 1424.58, observed mass 1428.64 $[M+H]^+$

8.3.9 Synthesis of 5- (*p*-(3-methyl-3-hydroxybutynyl)phenyl-10,15,20-triphenyl-21H,23H-porphyrin (10)[49]



10

Pyrrole (4.20 mL, 60 mmol, 6.0 eq.), benzaldehyde (6.36 g, 60 mmol, 6.0 eq.) and 4-(3-hydroxy-3-methylbut-1-ynyl) benzaldehyde (1.88.g, 10.0 mmol, 1.0 eq.) were dissolved in chloroform (1 L) and purged with N_2 for 30 mins in the dark before boron trifluoride etherate (1.14 mL, 9.0 mmol, 0.9 eq.) was added and the reaction allowed to stir at room temperature. After 1 hr, DDQ (13.6 g, 60 mmol, 6.0 eq.) was added and left to stir overnight. The reaction mixture was concentrated *in vacuo*, DCM (300 mL) was added and the crude product was purified by column chromatography (silica/alumina, eluent – 10 % methanol in DCM) to isolate the porphyrin containing fractions. The product was isolated by column chromatography (silica/alumina, eluent – DCM → DCM + 2 % MeOH) and, crystallised from chloroform and hexane to yield dark purple crystals (0.83 g, 1.2 mmol, 12 %).

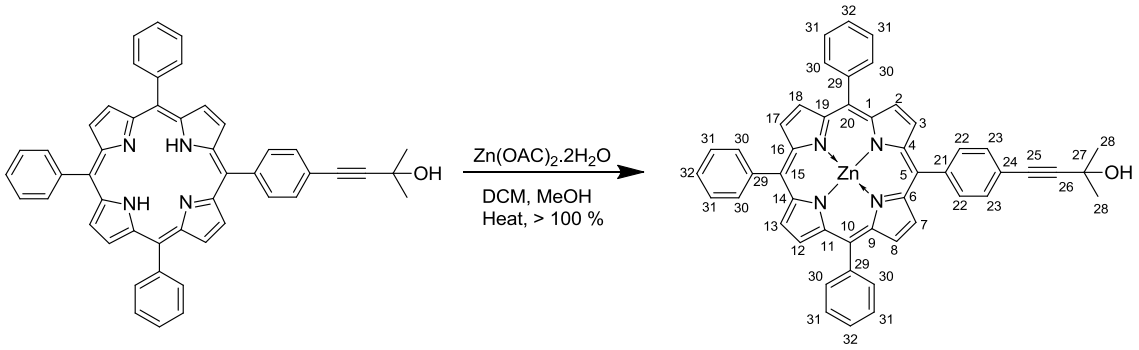
$\mathbf{R}_f$ 0.28 (silica, DCM)

^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.86-8.91 (m, 6H, **H^2**, **H^8**, **H^{12}**, **H^{13}**, **H^{17}**, **H^{18}**), 8.83-8.86 (m, 2H, **H^3**, **H^7**), 8.22-8.27 (m, 6H, **H^{31}**), 8.20 (d, *J* = 8.0 Hz, 2H, **H^7**), 7.84 (d, *J*=8.0 Hz, 2H, **H^7**), 7.73-7.81 (m, 9H, **H^{32}**, **H^{33}**), 1.79 (s, 6H, **H^{29}**), -2.74 (br. s, 2H, **H^{21}**)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 142.29 (C, **C22**), 142.09 (C, **C30**), 134.52 (CH, **C31**), 134.46 (CH, **C23**), 130.32 - 131.98 (m, C, **C2**, **C3**, **C7**, **C8**, **C12**, **C13**, **C17**, **C18**), 130.01 (CH, **C24**), 127.72 (CH, **C33**), 126.67 (CH, **C32**), 122.23 (C, **C14**, **C16**), 120.40 (C, **C1**, **C9**, **C11**, **C19**), 119.13 (C, **C4**, **C6**), 95.02 (C, **C27**), 82.16 (C, **C26**), 65.79 (C, **C28**), 31.63 (CH_3, **C29**)

MALDI-TOF ($C_{49}H_{36}N_4O$): Calculated mass 696.84; observed mass 697.4 $[M+H]^+$

8.3.10 Synthesis of zinc (II) 5-(*p*-(3-methyl-3-hydroxybutynyl)phenyl-10,15,20-triphenyl porphyrin (11)[49]



10 **11**

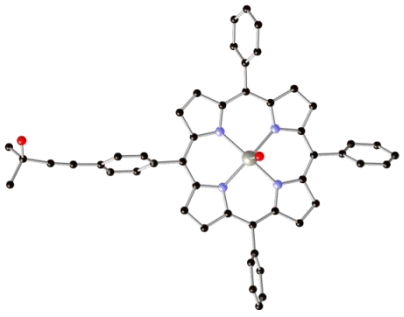
5-(*p*-(3-methyl-3-hydroxybutynyl)phenyl-10,15,20-triphenyl-21H,23H-porphyrin **10** (822 mg, 1.2 mmol, 1.0 eq.) and zinc acetate dihydrate (3.39 g, 15.5 mmol, 13.1 eq.) were dissolved in DCM (100 mL) and methanol (20 mL) before heating gently for 10 mins. The reaction mixture was concentrated *in vacuo*, redissolved in DCM, the excess zinc acetate dihydrate filtered and the filtrate again concentrated *in vacuo*. The product was then washed with brine (2 × 100 mL), dried (Na_2SO_4) and the solvent removed *in vacuo* to give a purple powder (0.913 mg, > 100 %)

R_f 0.14 (silica, DCM)

^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.99-8.95 (m, 8H, **H^2**, **H^3**, **H^7**, **H^8**, **H^{12}**, **H^{13}**, **H^{17}**, **H^{18}**), 8.26-8.22 (m, 6H, **H^{30}**), 8.20 (d, *J* = 8.0 Hz, 2H, **H^{22}**), 7.85 (d, *J* = 8.0 Hz, 2H, **H^{23}**) 7.82-7.72 (m, 9H, **H^{31}**, **H^{32}**), 1.70 (s, 6H, **H^{28}**)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 150.25 (C, **C10**, **C20**) 150.16 (C, **C15**), 149.80 (C, **5**), 143.00 (C, **C21**), 142.75 (C, **C29**), 134.39 (CH, **C30**), 134.31 (CH, **C22**), 132.12 (CH, **C2**, **C3**, **C7**, **C8**), 132.00 (CH, **C12**, **C13**, **C17**, **C18**), 129.82 (CH, **C23**), 127.48 (CH, **C32**), 126.51 (CH, **C31**), 121.85 (C, **C14**, **C16**), 121.28 (C, **C1**, **C9**, **C11**, **C19**), 120.03 (C, **C4**, **C6**), 94.64 (C, **C26**), 82.18 (C, **C25**), 65.68 (C, **C27**), 31.38 (CH_3, **C28**)

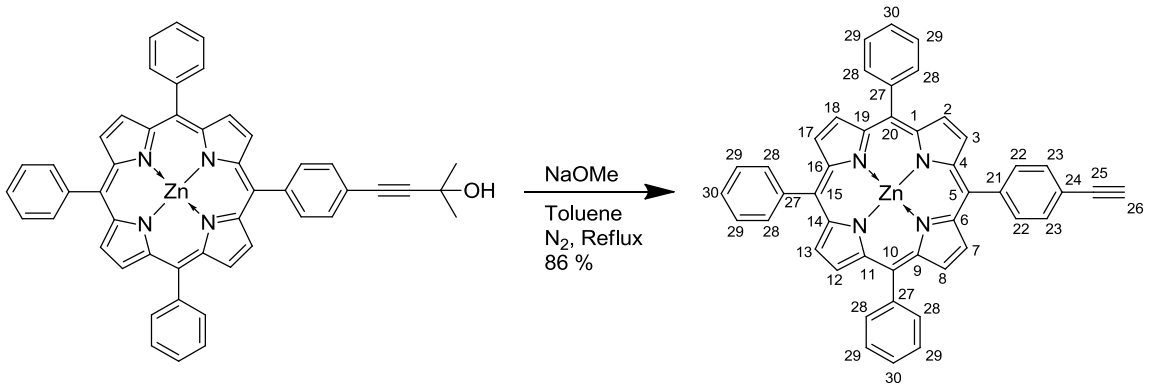
MALDI-TOF ($C_{49}H_{34}N_4OZn$): Calculated mass 758.20, observed mass 759.6 $[M+H]^+$



Crystal data and structure refinement details.

Identification code	**2011SO|T0289** (DS/5896/47)	
	2011SRC0404	
Empirical formula	$C_{49}H_{36}N_4O_2Zn$	
Formula weight	778.19	
Temperature	100(2) K	
Wavelength	0.68890 Å	
Crystal system	Triclinic	
Space group	*P*–1	
Unit cell dimensions	*a* = 10.709(5) Å	α = 77.319(1)°
	b = 13.092(5) Å	β = 82.121(8)°
	c = 14.248(6) Å	γ = 75.541(1)°
Volume	1880.0(14) Å^3	
Z	2	
Density (calculated)	1.375 Mg / m^3	
Absorption coefficient	0.701 mm^{-1}	
F(000)	808	
Crystal	Needle; Red	
Crystal size	0.05 × 0.00 × 0.00 mm^3	
θ range for data collection	2.85 – 24.38°	
Index ranges	$-11 \le h \le 12$, $-12 \le k \le 14$, $-17 \le l \le 17$	
Reflections collected	12624	
Independent reflections	6113 [R_{int} = 0.1205]	
Completeness to θ = 24.38°	90.2 %	
Absorption correction	Semi–empirical from equivalents	
Max. and min. transmission	1.000 and 0.347	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	6113 / 3 / 514	
Goodness-of-fit on F^2	1.012	
Final *R* indices [$F^2 > 2\sigma(F^2)$]	*R1* = 0.0661, *wR2* = 0.1395	
R indices (all data)	*R1* = 0.1166, *wR2* = 0.1778	
Largest diff. peak and hole	0.739 and –0.547 e Å^{-3}	

8.3.11 Synthesis of zinc (II) 5-*p*-ethynylphenyl-10,15,20-triphenyl porphyrin (12)[49]



11 **12**

Zinc (II) 5-(*p*-(3-methyl-3-hydroxybutynyl) phenyl-10,15,20-triphenyl porphyrin **11** (822 mg, 1.2 mmol, 1.0 eq.) and sodium methoxide (1.91 g, 35 mmol, 30 eq.) were dissolved into toluene (100 mL) and purged with N_2 for 1 hr before heating to reflux (125 °C) for 18 hr. When complete, the reaction mixture was cooled, concentrated *in vacuo*, extracted into DCM (100 mL) and washed with sat. KCl (2 × 100 mL) dried (Na_2SO_4), and the solvent removed *in vacuo*. The crude product was purified using column chromatography (silica, eluent – DCM). A purple/red product was obtained (711 mg, 1.0 mmol, 86 %).

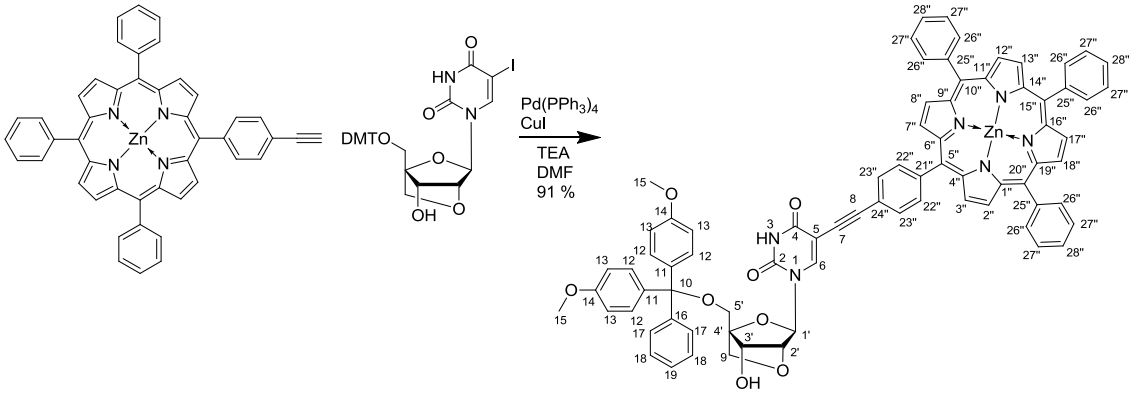
$\mathbf{R}_f$ 0.88 (silica, DCM)

^{1}**H NMR** (400 MHz, $CDCl_3$): δ = 9.03-9.06 (m, 6H, **H**^{2}, **H**^{8}, **H**^{12}, **H**^{13}, **H**^{17}, **H**^{18}), 8.99-902 (m, 2H, **H**^{3}, **H**^{7}) 8.29 (m, 6H, **H**^{28}), 8.26 (d, *J* = 7.9 Hz, 2H, **H**^{22}) 7.93 (d, *J* = 7.9 Hz, 2H, **H**^{23}), 7.77-7.85 (m, 9H, **H**^{29}**, H**^{30}), 3.31 (br. s., 1H, **H**^{26})

^{13}**C NMR** (100 MHz, $CDCl_3$): δ = 150.6 (C, **C10**, **C20**), 150.1 (C, **C15**), 143.8 (C, **C5**), 143.0 (C, **C21**, **C27**), 134.7 (C, **C27**), 134.6 (CH, **C28**, **C22**), 132.6 (CH, **C2**, **C3**, **C7**, **C8**), 132.4 (CH, **C12**, **C13**, **C17**, **C18**), 130.7 (CH, **C23**), 127.9 (CH, **C30**), 125.5 (CH, **C29**), 121.7 (C, **C14**, **C16**), 121.6 (C, **C1**, **C9**, **C11**, **C19**), 120.3 (C, **C4**, **C6**), 84.1 (C, **C25**), 78.4 (CH, **C26**)

MALDI-TOF ($C_{46}H_{28}N_4Zn$): Calculated mass 700.16, observed mass 701.0 $[M+H]^+$

8.3.12 Synthesis of 5'-DMT-5-(5"*p*-ethynylphenyl-10",15",20"-triphenyl-21",23"-zinc (II) porphyrin)-dU-LNA (13)



12 **13**

Using flame dried glassware under an inert atmosphere in the absence of light, zinc (II) 5-*p*-ethynylphenyl-10,15,20-triphenyl porphyrin **12** (102 mg, 0.15 mmol, 1.02 eq.), 5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-LNA-uridine (100 mg, 0.15 mmol, 1.00 eq.), copper(I) iodide (9 mg, 48 μmol, 0.33 eq.) and tetrakis(triphenylphosphine) palladium(0) (28 mg, 24 μmol, 0.17 eq.) were dissolved in DMF (3 mL). The reaction mixture was further purged with N_2 for 20 mins prior to the addition of triethylamine (40 μL, 0.44 mmol, 3.3 eq). The reaction was allowed to stir for 7 hr before extracting into ethyl acetate (100 mL), and washed with 5 % w/v Na_2EDTA (pH = 9) (2 × 100 mL), sat. NaCl (100 mL), dried (Na_2SO_4), and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica containing 20 % silica H neutralised with TEA, eluent - DCM → DCM + 3 % MeOH). The pure product was collected as a purple solid (167 mg, 0.13 mmol, 91 %).

$\mathbf{R}_f$ 0.94 (silica, DCM + 10 % MeOH)

^{1}**H NMR** (300 MHz, $CDCl_3$): δ = 9.03-9.10 (m, 6H, **H**$^{2"}$, **H**$^{8"}$, **H**$^{12"}$, **H**$^{13"}$, **H**$^{17"}$, **H**$^{18"}$),), 8.99-9.03 (m, 2H, **H**$^{3"}$, **H**$^{7"}$), 8.34 (m, 6H, **H**$^{26"}$), 8.19 (s, 1H, **H**^{6}), 8.13 (d, *J* = 7.7 Hz, 2H, **H**$^{23"}$), 7.76-7.89 (m, 9H, **H**$^{27"}$, **H**$^{28"}$), 7.62 (d, *J* = 7.5 Hz, 2H, **H**$^{22"}$), 7.39-7.56 (m, 8H,**H**^{12}, **H**^{17}, **H**^{18}), 7.30 (t, *J* = 7.0 Hz, 1H, **H**^{19}), 6.80-6.90 (m, 4H, **H**^{13}), 5.24 (s, 1H, **H**$^{1'}$), 4.80 (m, 2H, **H**$^{2'}$, **H**$^{3'}$), 4.18 (s, 1H, **H**^{9}), 3.98 (s, 1H, **H**^{9}), 3.63 (s, 6H, **H**^{15}), 3.56 (d, *J* = 9.2 Hz, 2H, **H**$^{5'}$)

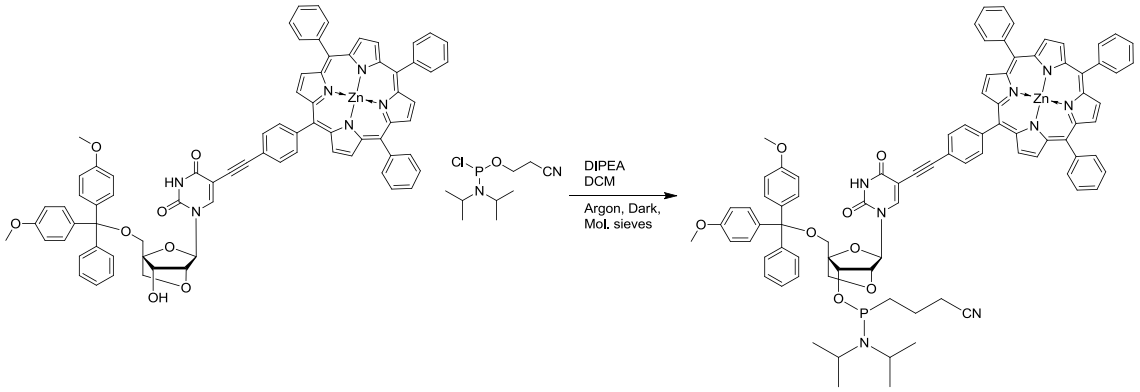
^{13}**C NMR** (100 MHz, $CDCl_3$): δ = 161.3 (CH, **C4**), 158.4 (CH, **C14**), 150.2 (C, **C5"**, **C10"**, **C15"**, **C20"**), 148.1 (C, **C2**), 144.2 (C, **C24"**), 143.0 (C, **C16**), 142.8 (C, **C25"**), 140.9 (CH, **C6**), 135.7 (C, **C11**), 134.5 (CH, **C26"**), 134.3 (CH, **C23"**), 132.0 (CH, **C2"**, **C3"**, **C7"**, **C8"**, **C12"**, **C13"**, **C17"**, **C18"**), 129.8 (CH, **C12**), 129.7 (CH, **C17**, CH, **C18**), 129.0 (CH, **C19**), 128.0 (CH, **C22"**), 126.5 (CH, **C27"**, **C28"**), 113.4 (CH **C13**), 113.1 (CH **C13**), 100.0 (C, **C8**), 94.3 (C, **C7**), (CH, **C1'**), 86.6 (C, **C10**), 80.8 (C, **C5**), 78.6 (CH, **C2'**), 71.4 (CH_2, **C9**), 70.2 (CH, **C3'**), 58.5 (CH_2, **C5'**)

ESI negative ($C_{55}H_{38}N_6O_5Zn$): Calculated mass 1256.35, observed mass 1255.6 $[M-H]^-$

UV-Vis (DCM, 8 μM): λ_{max} (log ε) 420 nm (5.15), 548 nm (3.73), 588 nm (2.89)

Emission (DCM, 8 μM): λ_{ex} 420 nm, λ_{em}(rel int) 598 nm (1), 645 nm (0.91)

8.3.13 Synthesis of 5'-DMT-5-(5"*p*-ethynylphenyl-10",15",20"-triphenyl-21",23"-zinc (II) porphyrin)-dU-LNA-3'-amidite (14)



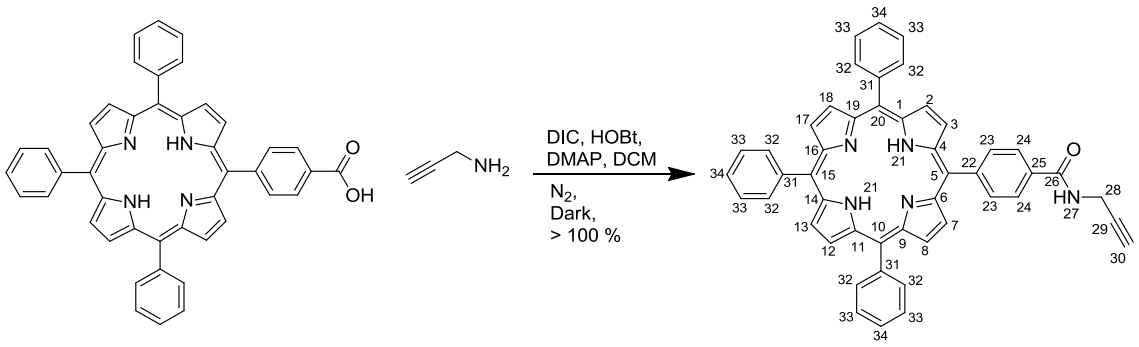
13 14

Using oven dried glassware under Schlenk conditions, 5'-DMT-5-(5''*p*-ethynylphenyl-10'',15'',20''-triphenyl-21'',23''-zinc (II) porphyrin)-dU-LNA **13** (70 mg, 56 µmol, 1.0 eq.) was dissolved in anhydrous DCM (3 mL) with molecular sieves in the dark. DIPEA (39 µL, 0.22 mmol, 4.0 eq.) and CEP-Cl (26 µL, 0.11 mmol, 2.0 eq.) were added and the reaction allowed to stir at room temperature for 2 hr. The reaction hadn't gone to completion in this time so additional DIPEA (39 µL, 0.22 mmol, 4.0 eq.) and CEP-Cl (26 µL, 0.11 mmol, 2.0 eq.) were added. The reaction was found to be complete after 3.5 hr. The crude product was precipitated form hexanes (5 mL) and cooled (-18 °C) for 10 mins, the hexanes were decanted off and the crude product washed with further hexanes (5 mL). The crude product as used immediately for DNA synthesis.

Full characterisation of this product was not achieved due to its instability. The product was immediately used for DNA synthesis.

$\mathbf{R}_f$ 0.71 (silica, DCM + 10 % MeOH + 0.5 % Triethylamine)

8.3.14 Synthesis of 5-(*N*-(2-propargyl)amidophenyl)-10,15,20-triphenyl porphyrin (15)



2 **15**

Porphyrin acid **2** (200 mg, 0.30 mmol, 1.0 eq.) was dissolved in anhydrous DCM (10 mL). Propargylamine (38 µL, 0.61 mmol, 2.0 eq.), DIC (94 µL, 0.61 mmol, 2.0 eq.), HOBt (46 mg, 0.30 mmol, 1.0 eq.) and DMAP (74 mg, 0.61 mmol, 2.0 eq.) were added to the purple reaction mixture and stirred in the dark under an inert atmosphere for 4 hr. The reaction mixture was diluted with DCM (25 mL) and washed with water (2 × 50 mL), sat. KCl (50 mL), dried (Na_2SO_4) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica, eluent: DCM → DCM + 1 % MeOH) to give the product as a purple solid, (250 mg, 0.36 mmol, > 100 %).

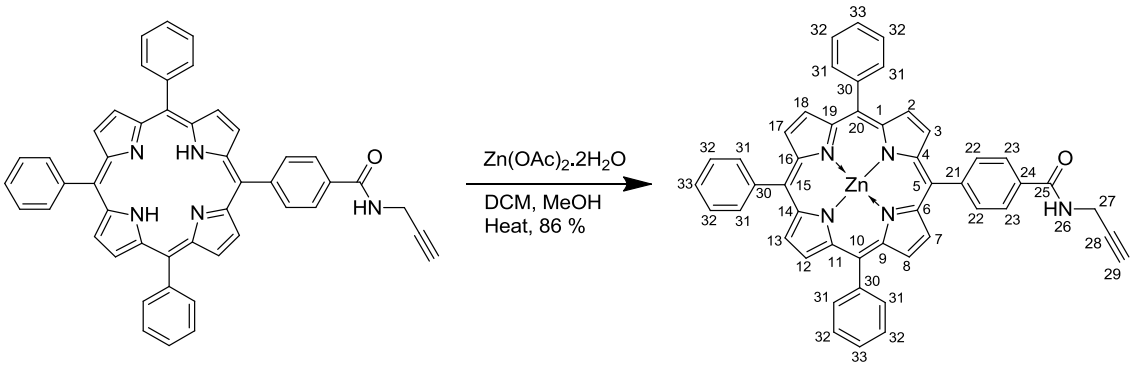
$\mathbf{R}_f$ 0.94 (silica, DCM + 10 % MeOH)

^{1}**H NMR** (400 MHz, $CDCl_3$): δ = 8.85-8.93 (m, 6H, **H**^{2}, **H**^{8}, **H**^{12}, **H**^{13}, **H**^{17}, **H**^{18}), 8.76-8.82 (m, 2H, **H**^{3}, **H**^{7}), 8.29 (d, *J* = 7.9 Hz, 2H, **H**^{23}), 8.21-8.27 (m, 6H, **H**^{32}), 8.15 (d, *J*=7.9 Hz, 2H, **H**^{24}), 7.64-7.88 (m, 9H, **H**^{33}, **H**^{34}), 6.58 (t, *J* = 5.0 Hz, 1H, **H**^{27}), 4.44 (dd, *J* = 4.9, 2.4 Hz, 2H, **H**^{28}), 2.40 (t, *J* = 2.5 Hz, 1H, **H**^{30}), -2.73 (s, 2H, **H**^{21})

^{13}**C NMR** (100 MHz, $CDCl_3$): δ = 167.3 (C, **C26**), 146.1 (C, **C25**), 142.2 (C, **C5**, **C10**, **C15**, **C20**), 134.8 (C, **C1**, **C9**, **C11**, **C14**, **C16**, **C19**), 134.7 (CH, **C32**), 133.2 (C, **C4**, **C6**), 131.5, (CH, **C2**, **C3**, **C7**, **C8**, **C12**, **C13**, **C17**, **C18**), 127.9 (CH, **C34**), 126.8 (CH, **C33**), 125.5 (CH, **C24**), 120.7 (C, **C22**), 120.5 (C, **C31**), 79.6 (C, **C29**), 72.3 (CH, **C30**), 30.2 (CH_2, **C28**)

MALDI-TOF ($C_{46}H_{32}N_4O_2$): Calculated mass 695.27, observed mass 696.83 $[M+H]^+$

8.3.15 Synthesis of zinc (II) 5-(*N*-(2-propargyl)amidophenyl)-10,15,20-triphenyl porphyrin (16)



15 **16**

Propargylamine porphyrin **15** (250 mg, 0.36 mmol, 1.0 eq.) and zinc acetate dihydrate (1.79 g, 8.14 mmol, 23 eq.) were dissolved in DCM (50 mL) and methanol (10 mL) before heating gently for 5 mins. The reaction mixture was concentrated *in vacuo*, redissolved in DCM, the excess zinc acetate dihydrate filtered and the filtrate again concentrated *in vacuo*. The product was then washed with brine (2 × 50 mL), dried (Na_2SO_4) and the solvent removed *in vacuo* to give a purple powder (235 g, 0.31 mmol, 86 %).

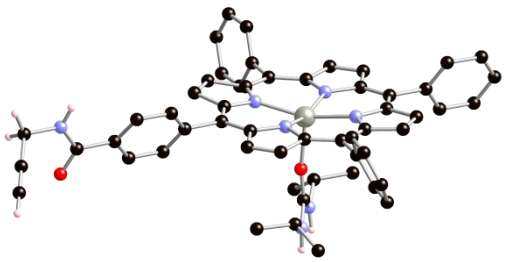
R_f 0.94 (silica, DCM + 10 % MeOH)

^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.90-8.97 (m, 6H, **H^2**, **H^8**, **H^{12}**, **H^{13}**, **H^{17}**, **H^{18}**), 8.82 (d, *J* = 4.7 Hz, 2H, **H^3**, **H^7**), 8.21-8.27 (m, 8H, **H^{22}**, **H^{31}**), 7.90 (d, *J* = 7.8 Hz, 2H, **H^{23}**), 7.66-7.83 (m, 9H, **H^{32}**, **H^{33}**), 6.48 (t, *J* = 5.1 Hz, 1H, **H^{26}**), 4.13 (br. s., 2H, **H^{27}**), 2.32 (t, *J* = 2.5 Hz, 1H, **H^{29}**)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 167.7 (C, **C25**), 150.9 (C, **C5**, **C10**, **C15**, **C20**), 150.2 (C, **C21**), 147.7 (C, **C24**), 143.7 (C, **C30**), 135.2 (CH, **C31**), 132.6 (CH, **C2**, **C3**, **C7**, **C8**, **C12**, **C13**, **C17**, **C18**) 132.0 (CH, C**22**), 128.1 (CH, **C33**), 127.2 (CH, **C32**), 125.7 (CH, **C23**), 121.8 (**C1**, **C4**, **C6 C9**, **C11**, **C14**, **C16**, **C19**), 80.1 (C, **C28**), 72.7 (CH, **C29**), 30.4 (CH_2, **C27**)

ESI Positive ($C_{48}H_{31}N_5OZn$): Calculated mass 757.18, observed *m/z*: 821.2 $[M+Na+CH_3CN]^+$

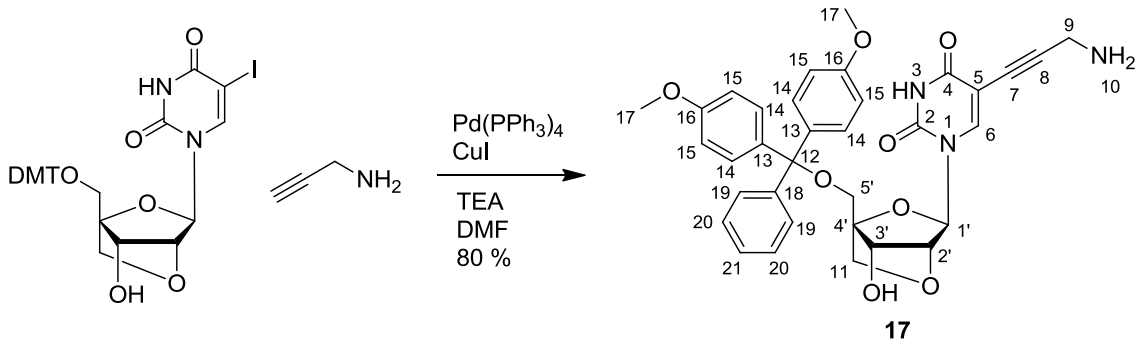
ESI Negative ($C_{48}H_{31}N_5OZn$): Calculated mass 757.18, observed *m/z*: 756.1 $[M-H]^-$



Crystal data and structure refinement details.

Identification code	**2011sot0718** (DS/6101/55)	
Empirical formula	$C_{57}H_{49}Cl_6N_7O_2Zn$	
	$C_{55}H_{47}N_7O_2Zn, 2CHCl_3$	
Formula weight	1142.10	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Triclinic	
Space group	*P*–1	
Unit cell dimensions	*a* = 12.287(3) Å	α = 107.603(13)°
	b = 13.349(3) Å	β = 97.51(2)°
	c = 18.331(5) Å	γ = 102.633(17)°
Volume	2733.3(12) Å^3	
Z	2	
Density (calculated)	1.388 Mg / m^3	
Absorption coefficient	0.791 mm^{-1}	
F(000)	1176	
Crystal	Iridescent blue fragment	
Crystal size	0.10 × 0.06 × 0.02 mm^3	
θ range for data collection	2.93 – 25.02°	
Index ranges	$-14 \le h \le 14$, $-14 \le k \le 15$, $-21 \le l \le 21$	
Reflections collected	21909	
Independent reflections	9615 [R_{int} = 0.0384]	
Completeness to θ = 25.02°	99.6 %	
Absorption correction	Semi–empirical from equivalents	
Max. and min. transmission	0.9843 and 0.9250	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	9615 / 76 / 701	
Goodness-of-fit on F^2	1.121	
Final *R* indices [$F^2 > 2\sigma(F^2)$]	*R1* = 0.0602, *wR2* = 0.1357	
R indices (all data)	*R1* = 0.0701, *wR2* = 0.1416	
Largest diff. peak and hole	0.579 and –0.884 e Å^{-3}	

8.3.16 Synthesis of 5'-DMT-5-propargylamine-dU-LNA (17)



17

5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-LNA-uridine (200 mg, 0.29 mmol, 1.0 eq.) was dissolved in anhydrous DMF (2.5 mL) and degassed with nitrogen for 15 mins. In the absence of light, copper(I) iodide (14 mg, 73 µmol, 0.25 eq.) and propargylamine (37 µL, 0.58 mmol, 2.0 eq.) were added. After stirring for 20 mins, tetrakis(triphenylphosphine)palladium(0) (34 mg, 29 µmol, 0.1 eq.) and triethylamine (0.3 mL, 2.04 mmol, 7.0 eq.) were added and the reaction stirred for 2.5 hr. The reaction mixture was diluted with ethyl acetate (50 mL) and washed with 5 % w/v Na_2EDTA (pH = 9) (2 × 50 mL). The aqueous phases were back extracted with ethyl acetate (50 mL). The organic phases were combined, washed with sat. KCl (100 mL), dried ($MgSO_4$), solvent removed by rotary evaporation and dried under high vacuum to give a yellow/white foam. The crude product was purified by column chromatography (silica, eluent – DCM + 1 % MeOH → 10 % + 0.5 % pyridine) to give the desired product, after drying under high vacuum to give a white/yellow solid (144 mg, 0.24 mmol, 80 %).

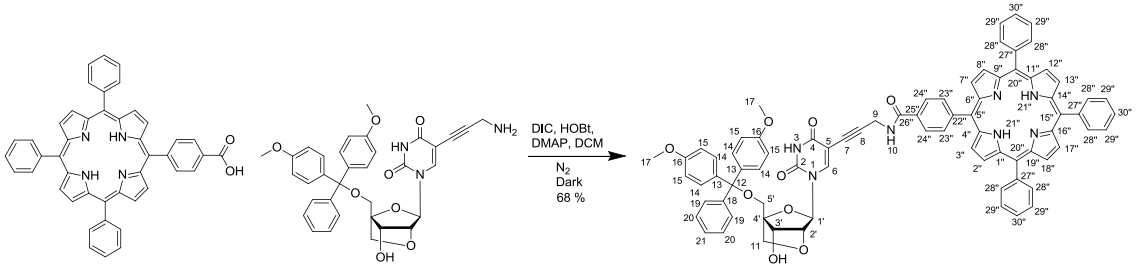
$\mathbf{R}_f$ 0.33 (silica, DCM + 10 % MeOH)

^{1}**H NMR** (400 MHz, $CDCl_3$): δ = 7.95 (s, 1H, **H**^{6}), 7.49 (d, *J* = 7.1 Hz, 2H, **H**^{19}), 7.33-7.40 (m, 4H, **H**^{14}), 7.23-7.32 (m, 2H, **H**^{20}), 7.12-7.21 (m, 1H, **H**^{21}), 6.84 (d, *J* = 7.6 Hz, 4H, **H**^{15}), 5.59 (br. s., 1H, **H**$^{1'}$), 4.66 (br. s., 1H, **H**$^{2'}$), 4.33 (br. s., 1H, **H**$^{3'}$), 3.93 (d, *J* = 6.3 Hz, 1H, **H**^{11}), 3.85 (d, *J* = 6.3 Hz, 1H, **H**^{11}), 3.75 (s, 6H, **H**^{17}), 3.45-3.58 (m, 2H, **H**$^{5'}$), 3.31 (br. s., 2H, **H**^{9})

^{13}**C NMR** (100 MHz, $CDCl_3$): δ = 162.8 (C, **C4**), 158.6 (C, **C16**), 149.1 (C, **C2**), 144.8 (C **C13**), 141.8 (CH, **C6**), 130.2 (CH, **C14**), 130.0 (CH, **C14**), 128.2 (CH, **C20**), 128.0 (CH, **C19**), 125.3 (CH, **C21**), 113.3 (CH, **C15**), 98.9 (C, **C7**), 93.3 (C, **C8**), 87.7 (CH, **C1'**), 86.5 (C, **C12**), 79.2 (CH, **C2'**), 74.2 (C, **C5**), 72.1 (CH_2, **C11**), 70.1 (CH, **C3'**), 59.4 (CH_2, **C5'**), 55.2 (CH_3, **C17**), 31.8 (CH_2, **C9**)

ESI Negative ($C_{34}H_{33}N_3O_8$): Calculated mass 611.23, observed *m/z*: 610.1 ($[M - H]^-$

8.3.17 Synthesis of *N*-(5'DMT-5-propargyl-dU-LNA)-5"-(*p*-benzamide)-10",15",20"-triphenyl-21"-*H*-23"-*H*-porphyrin (18)



2 17 18

Porphyrin acid **2** (113 mg, 0.17 mmol, 1.0 eq.) was dissolved in anhydrous DCM (5 mL). Propargyl deoxyuridine LNA **17** (115 mg, 0.19 mmol, 1.1 eq.), DIC (53 µL, 0.34 mmol, 2.0 eq.), HOBt (26 mg, 0.30, 1.0 eq.) and DMAP (42 mg, 0.34 mmol, 2.0 eq.) were added to the purple reaction mixture and stirred in the dark under an inert atmosphere for 4 hr. The reaction mixture was diluted with DCM (25 mL) and washed with water (2 × 50 mL), sat. KCl (50 mL), dried (Na_2SO_4) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica neutralised with triethylamine, eluent – DCM → DCM + 2 % MeOH) to give the product as a purple solid, (145 mg, 0.12 mmol, 68 %).

R_f 0.68 (silica, DCM + 10 % MeOH)

^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.79-8.85 (m, 6H, **H$^{2"}$**, **H$^{8"}$**, **H$^{12"}$**, **H$^{13"}$**, **H$^{17"}$**, **H$^{18"}$**), 8.72 (m, 2H, **H$^{3"}$**, **H$^{4"}$**), 8.19 (d, *J* = 7.7 Hz, 2H, **H$^{24"}$**), 8.15 (m, 7H, **H^{6}**, **H$^{28"}$**), 7.96 (d, *J* = 7.7 Hz, 2H, **H$^{23"}$**), 7.62-7.82 (m, 9H, **H$^{29"}$**, **H$^{30"}$**), 7.49 (d, *J* = 7.7 Hz, 2H, **H^{19}**), 7.38 (dd, *J* = 8.7, 2.8 Hz, 4H, **H^{14}**), 7.25-7.31 (m, 2H, **H^{20}**), 7.16 (t, *J* = 7.1 Hz, 1H, **H^{21}**), 6.81 (dd, *J* = 8.7, 3.2 Hz, 4H, **H^{15}**), 5.63 (s, 1H, **H$^{1'}$**), 4.63 (s, 1H, **H$^{2'}$**), 4.29-4.49 (m, 3H, **H$^{3'}$**, **H^{9}**), 3.93 (d, *J* = 8.1 Hz, 1H, **H^{11}**), 3.83 (d, *J* = 8.2 Hz, 1H, **H^{11}**), 3.64-3.67 (m, 1H, **H$^{5'}$**), 3.63 (s, 3H, **H^{17}**), 3.61 (s, 3H, **H^{17}**), 3.56 (d, *J* = 8.2 Hz, 1H, **H$^{5'}$**), -2.78 (br. s., 2H, **H$^{21"}$**)

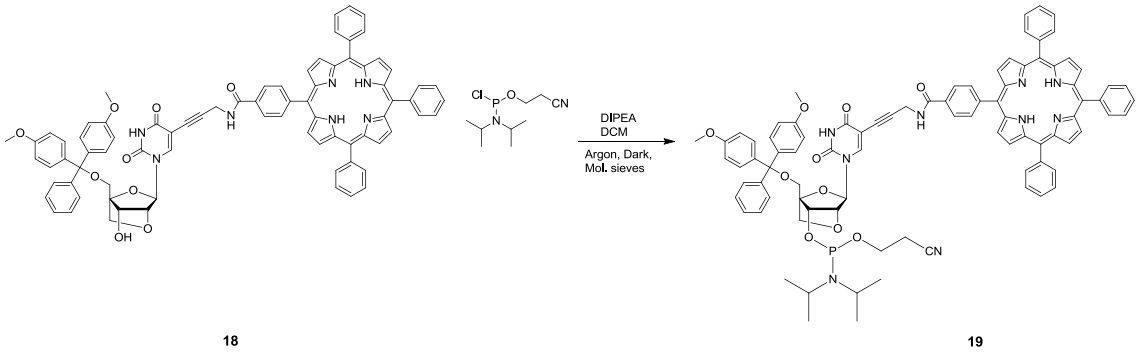
^{13}C NMR (100 MHz, $CDCl_3$): δ = 167.2 (C, **C26"**), 162.3 (C, **C4**), 158.6 (C, **C16**), 148.8 (C, **C2**), 145.6 (C, **C25"**), 144.5 (C, **C18**), 142.0 (CH, **C6**) 134.6 (C, **C13**), 134.5 (CH, **C24"**, **C28"**), 132.9 (C, **C27"**), 131.7 (CH, **C2"**, **C3"**, **C7"**, **C8"**, **C12"**, **C13"**, **C17"**, **C18"**), 130.1 (CH, **C14**), 128.0 (CH, **C20**), 127.7 (CH, **C19**), 126.7 (CH, **C29"**, **C30"**), 127.0 (CH, **C21**), 125.5 (CH, **C23"**), 113.5 (CH, **C15**), 87.5 (C, **C12**), 86.7 (CH, **C1'**), 79.23 (CH, **C2'**) 74.4 (CH_2 **C11**), 70.4 (CH, **C3'**), 58.6 (CH_2, **C5'**), 55.2 (CH_3, **C17**), 30.9 (CH_2, **C9**)

ESI Negative ($C_{79}H_{61}N_7O_9$): Calculated mass 1251.45, observed *m/z* 1250.7 $[M - H]^-$

UV-Vis (DCM, 2.4 µM): λ_{max} (log ε) 417 nm (5.68), 515 nm (4.33), 548 nm (3.99), 589 nm (3.83), 646 nm (3.64)

Emission (DCM, 2.4 µM): λ_{ex} 417 nm, λ_{em} (rel int) 650 nm (1), 717 nm (0.33)

8.3.18 Synthesis of *N*-(5'-DMT-5-propargyl-dU-LNA)-5"-(*p*-benzamide)-10",15",20"-triphenyl-21"-*H*-23"-*H*-porphyrin-3'-amidite (19)



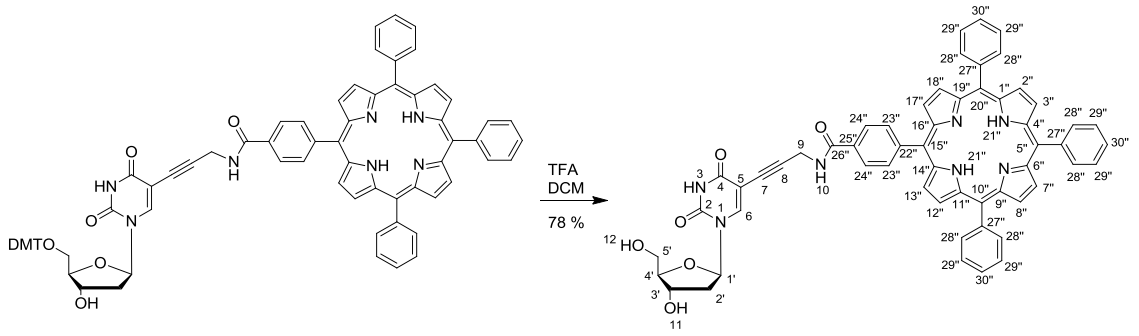
18 19

Using oven dried glassware under Schlenk conditions, *N*-(5'-DMT-5-propargyl-dU-LNA)-5''-(*p*-benzamide)-10'',15'',20''-triphenyl-21''-*H*-23''-*H*-porphyrin **18** (103 mg, 82 µmol, 1.0 eq.) was dissolved in anhydrous DCM (3 mL) with molecular sieves in the dark. DIPEA (57 µL, 0.33 mmol, 4.0 eq.) and CEP-Cl (58 µL, 0.25 mmol, 3.0 eq.) were added and the reaction allowed to stir at room temperature for 1.75 hr. The crude product was precipitated form hexanes (5 mL) and cooled (-18 °C) for 10 mins, the hexanes were decanted off and the crude product washed with further hexanes (5 mL). The crude product was used immediately for DNA synthesis.

Full characterisation of this product was not achieved due to its instability. The product was immediately used for DNA synthesis.

$\mathbf{R}_f$ 0.71 (silica, DCM + 10 % MeOH + 0.5 % Triethylamine)

8.3.19 Synthesis of 5'-hydroxyl amide linked porphyrin monomer (20)



7 **20**

Porphyrin **7** (34 mg, 28 μmol, 1.0 eq.) was dissolved DCM (2mL). TFA (10 μL, 138 μmol, 5.0 eq.) was added to the purple solution and stirred for 2 mins. A change in colour to green was observed upon TFA addition. The reaction mixture was diluted with DCM (100 mL) and washed with 1M HCl (100 mL), sat. sodium bicarbonate solution (100 mL), dried (Na_2SO_4) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica, eluent – 5 % MeOH in DCM). The purple product was collected and dried *in vacuo* (20 mg, 22 μmol, 78 % yield).

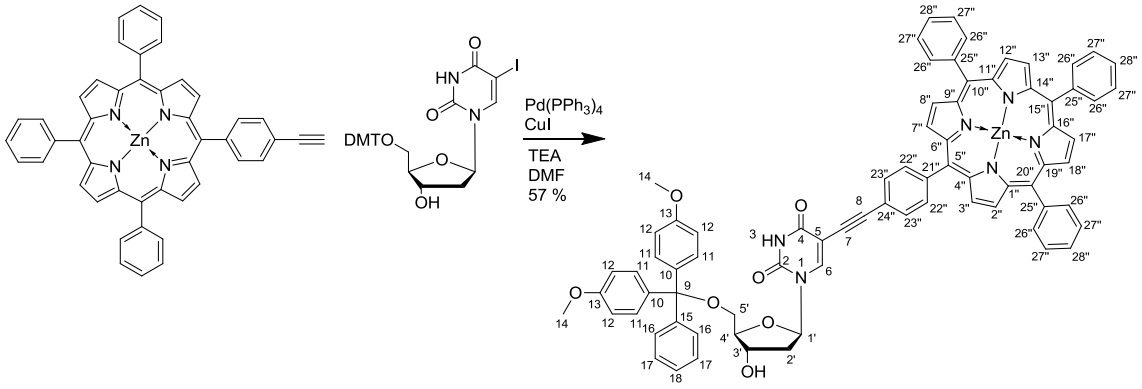
R_f 0.22 (silica, DCM + 10 % MeOH)

^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.66 (br. s, 1H, **H^3**), 9.39 (t, *J* = 5.5 Hz, 1H, **H^{10}**), 8.96 - 8.74 (m, 8H, **H$^{2"}$**, **H$^{3"}$**, **H$^{7"}$**, **H$^{8"}$**, **H$^{12"}$**, **H$^{13"}$**, **H$^{17"}$**, **H$^{18"}$**), 8.39 - 8.30 (m, 4H, **H$^{23"}$**, **H$^{24"}$**), 8.25 (s, 1H, **H^6**), 8.24 - 8.21 (m, 6H, **H$^{28"}$**), 7.91 - 7.78 (m, 9H, **H$^{29"}$**, **H$^{30"}$**), 6.15 (t, *J* = 6.7 Hz, 1H, **H$^{1'}$**), 5.25 (d, *J* = 4.1 Hz, 1H, **H^{11}**), 5.14 (t, *J* = 5.0 Hz, 1H, **H^{12}**), 4.47 (d, *J* = 5.4 Hz, 2H, **H^9**), 4.31 - 4.22 (m, 1H, **H$^{3'}$**), 3.83 (q, *J* = 3.3 Hz, 1H, **H$^{4'}$**), 3.66 (ddd, *J* = 3.6, 5.0, 12.0 Hz, 1H, **H$^{5'}$**), 3.60 (ddd, *J* = 3.8, 4.8, 11.9 Hz, 1H, **H$^{5'}$**), 2.23 - 2.10 (m, 2H, **H$^{2'}$**), -2.89 - -2.94 (m, 2H, **H^{21}**)

^{13}C NMR – Insufficient material for conclusive ^{13}C NMR analysis

MALDI-TOF ($C_{57}H_{43}N_7O_6$): Calculated mass 921.3, observed mass 922.6 $[M+H]^+$

8.3.20 Synthesis of 5'-DMT-5-(5"*p*-ethynylphenyl-10",15",20"-triphenyl-21",23"-zinc (II) porphyrin)-dU (21)[49]



12 **3** **21**

Zinc (II) 5-*p*-ethynylphenyl-10,15,20-triphenyl porphyrin **12** (95 mg, 0.13 mmol, 1.0 eq.), 5'-dimethoxytrityloxy-2'-deoxyribose-5-iodouracil **3** (238 mg, 0.36 mmol, 2.7 eq.), copper(I) iodide (8.0 mg, 0.04 mmol, 0.33 eq.) and triethylamine (36 µL, 0.44 mmol, 3.3 eq.) were dissolved in DMF (3 mL) in N_2 purged, flame dried glassware, in the absence of light. The reaction mixture was further purged with N_2 for 20 mins prior to the addition of tetrakis(triphenylphosphine)palladium(0) (23.2 mg, 0.02 mmol, 0.15 eq.). The reaction was allowed to stir for 36 hr before extracting into ethyl acetate (50 mL), washing with sat. KCl (3 × 50 mL), dried ($MgSO_4$), and concentrated *in vacuo*. Column chromatography was performed twice on TEA neutralised silica containing 20 % silica H (column eluent – 50:1:1 DCM:MeOH:EA). The isolated product was coevaporatated with toluene (3 x 20 mL) and neutralised chloroform (3 x 20 mL) to give a purple solid (95 mg, 0.07 mmol, 57 %).

$\mathbf{R}_f$ 0.48 (silica, DCM + 10 % MeOH)

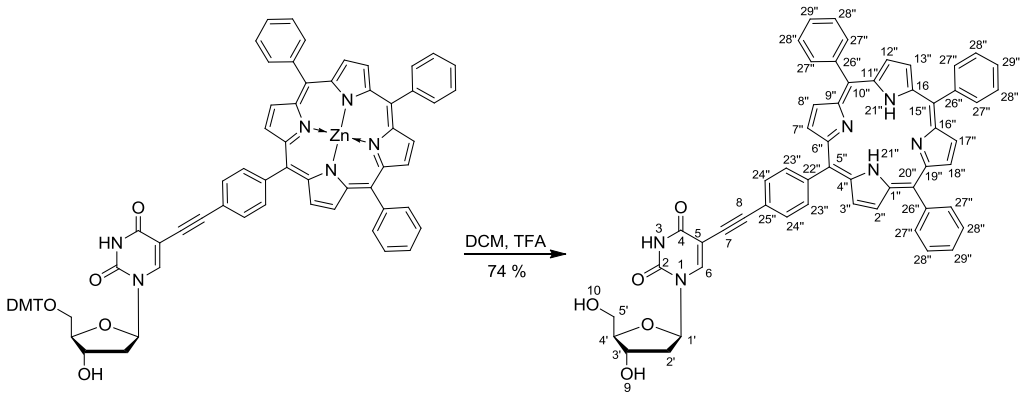
^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.02 (d, *J* = 4.6 Hz, 2H, **H**$^{2"}$, **H**$^{8"}$), 8.99 (s, 4H, **H**$^{12"}$, **H**$^{13"}$, **H**$^{17"}$, **H**$^{18"}$), 8.95 (d, *J* = 4.6 Hz, 2H, **H**$^{3"}$, **H**$^{7"}$), 8.24-8.32 (m, 6H, **H**$^{26"}$), 8.08 (d, *J* = 7.9 Hz, 2H, **H**$^{22"}$), 7.72-7.85 (m, 9H, **H**$^{27"}$, **H**$^{28"}$), 7.54 (s, 1H, **H**^{6}), 7.34 (m, 4H, **H**^{11}), 7.30 – 7.15 (m, 7H, **H**^{16}, **H**^{17}, **H**^{18}, **H**$^{23"}$), 6.76 (d, *J* = 8.7 Hz, 4H, **H**^{12}), 4.97-5.07 (m, 1H, **H**$^{1'}$), 3.87 (br. s., 1H, **H**$^{3'}$), 3.63 (s, 6H, **H**^{14}), 2.83 (br. s., 1H, **H**$^{4'}$), 2.72-2.78 (m, 2H, **H**$^{2'}$), 1.84 (m, 1H, **H**$^{5'}$), 1.69 (m, 1H, **H**$^{5'}$)

^{13}C NMR (100 MHz, $CDCl_3$): δ = 160.6 (C, **C4**), 158.1 (C, **C13**), 149.8 (C, **C15"**), 149.7 (C, **C10"**, **C20"**), 149.4 (C, **C5"**), 148.0 (C, **C2**), 143.9 (C, **C24"**), 142.7 (C, **C15**), 142.5 (C, **C25"**), 140.9 (CH, **C6**), 137.5 (C, **C21"**), 135.1 (C, **C10**), 135.0 (CH, **C26"**), 134.2 (CH, **C22"**) 131.8 (CH, **C2"**, **C8"**), 131.7 (CH, **C12"**, **C13"**, **C17"**, **C18"**), 131.3 (CH, **C3"**, **C7"**), 130.4 (CH, **C17**), 130.3 (CH, **C11**), 128.6 (C, **C12**), 127.8 (CH, **C28"**), 127.5 (CH, **C16**), 127.1 (CH, **18**), 126.7 (CH, **C29"**), 126.2, 121.1 (C, **C1"**, **C9"**), 120.9 (C, **C11"**, **C19"**), 120.8 (C, **C14"**, **C16"**), 119.8 (C, **C4"**, **C6"**), 113.0 (CH, **C12**), 99.8 (C, **C8**), 93.7 (C, **C7**), 86.7 (C, **C9**) , 85.7 (CH_2, **C4'**), 84.7 (CH, **C1'**), 80.4 (C, **C5**), 71.2 (CH, **C3'**), 62.5 (CH_2, **C5'**), 54.8 (CH_3, **C14**), 40.5 (CH_2, **C2'**)

MALDI-TOF ($C_{76}H_{56}N_6O_7Zn$): Calculated mass 1228.35, observed mass 1233.9 $[M+H]^+$

UV-Vis (DCM, 2.04 μM): λ_{max} (log ε) 400 nm (5.29), 417 (6.01), 549 (4.28), 589 nm (4.12)

8.3.21 Synthesis of 5'-hydroxyl acetylene linked porphyrin monomer (22)



Porphyrin **21** (75 mg, 61 μmol, 1.0 eq.) was dissolved DCM (2mL). TFA (23 μL, 305 μmol, 5.0 eq.) was added to the purple solution and stirred for 2 mins. A change in colour to green was observed upon TFA addition. The reaction mixture was diluted DCM (100 mL) and washed with 1M HCl (100 mL), sat. sodium bicarbonate solution (100 mL), dried (Na_2SO_4) and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica, eluent – 5 % MeOH in DCM). The purple product was collected and dried *in vacuo* (39 mg, 45 μmol, 74 %).

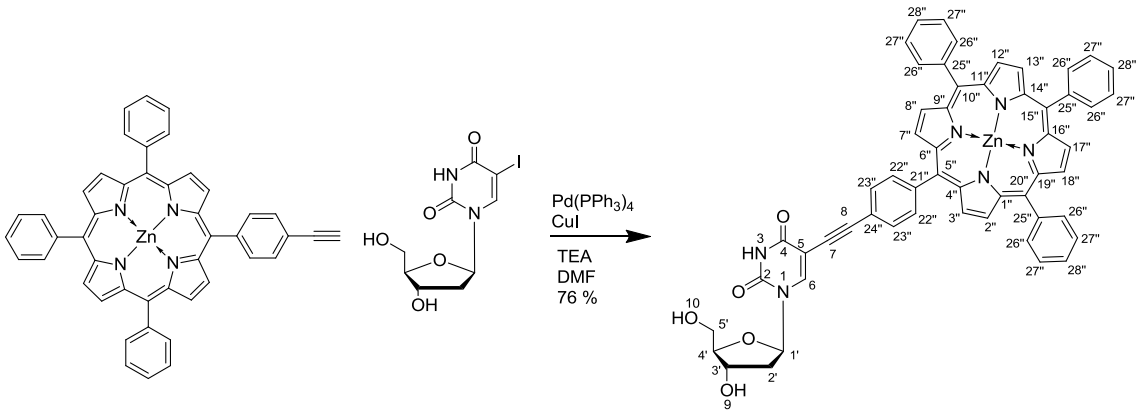
$\mathbf{R}_f$ 0.28 (silica, DCM + 10 % MeOH)

^{1}H NMR (400 MHz, DMSO-d_6): δ = 11.78 (br. s, 1H, **H^{3}**), 8.89 - 8.76 (m, 8H, **H$^{2"}$**, **H$^{3"}$**, **H$^{7"}$**, **H$^{8"}$**, **H$^{12"}$**, **H$^{13"}$**, **H$^{17"}$**, **H$^{18"}$**), 8.56 (s, 1H, **H^{6}**), 8.25 (d, *J* = 8.2 Hz, 2H, **H$^{23"}$**), 8.23 - 8.16 (m, 6H, **H$^{27"}$**), 7.91 (d, *J* = 8.2 Hz, 2H, **H$^{24"}$**), 7.87 - 7.78 (m, 9H, **H$^{28"}$**, **H^{29}**), 6.21 (t, *J* = 6.4 Hz, 1H, **H$^{1'}$**), 5.31 (d, *J* = 4.3 Hz, 1H, **H^{9}**), 5.27 (t, *J* = 4.8 Hz, 1H, **H^{10}**), 4.37 - 4.31 (m, 1H, **H$^{3'}$**), 3.88 (q, *J* = 3.2 Hz, 1H, **H$^{4'}$**), 3.74 (ddd, *J* = 3.3, 4.8, 11.8 Hz, 1H, **H$^{5'}$**), 3.66 (ddd, *J* = 3.4, 4.8, 11.8 Hz, 1H, **H$^{5'}$**), 2.31 - 2.18 (m, 2H. **H$^{2'}$**), -2.91 (br. s, 2H, **H$^{21"}$**)

^{13}C NMR – Insufficient material for conclusive ^{13}C NMR analysis

MALDI-TOF ($C_{55}H_{40}N_6O_5$): Calculated mass 864.3, observed mass 865.4 $[M+H]^+$

8.3.22 Synthesis of 5'-hydroxyl acetylene linked porphyrin monomer (23)



12 **23**

Porphyrin **12** (50 mg, 0.14 mmol, 1.0 eq.), 5-iodo-2'-deoxyuridine and copper(I) iodide (9 mg, 46 µmol, 0.33 eq.) were dissolved in anhydrous DMF (4 mL) and stirred under an inert atmosphere with the absence of light for 10 mins. tetrakis(triphenylphosphine)palladium(0) (27 mg, 23 µmol, 0.17 eq.) and triethylamine (40 µL, 0.29 mmol, 2.0 eq.) were added and the reaction left to stir for 24 hr. Upon completion, the reaction mixture was diluted with ethyl acetate (50 mL) and washed with 5 % w/v Na_2EDTA (pH = 9) (2 x 50 mL), sat. KCl (50 mL), dried (Na_2SO_4), and the solvent removed *in vacuo*. The crude product was purified by column chromatography (silica, eluent – DCM + 2 % MeOH → DCM + 5 % MeOH → EtOAc + 5 % MeOH). The purified product was obtained and dried *in vacuo* to give a purple solid (92 mg, 0.11 mmol, 76 %).

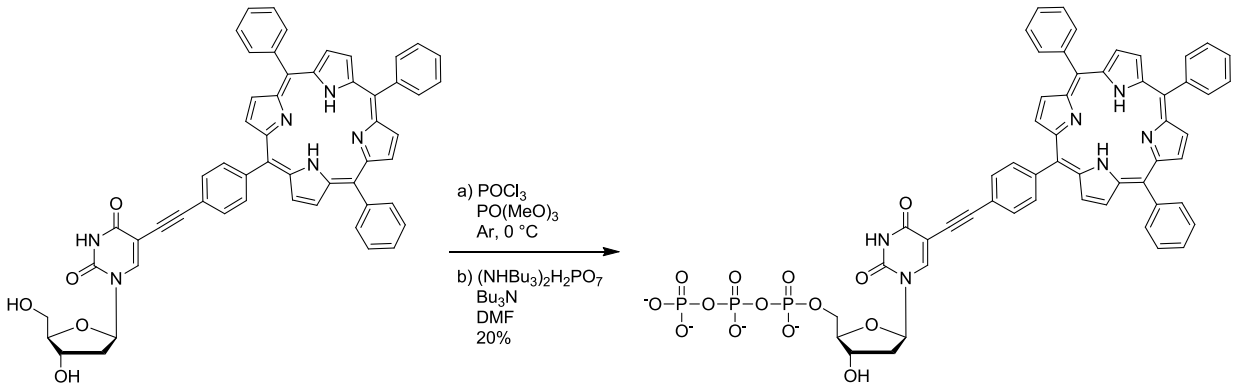
$\mathbf{R}_f$ 0.33 (silica, DCM + 10 % MeOH)

^{1}**H NMR** (400 MHz, DMSO-d_6): δ = 8.74-8.91 (m, 8H, **H**$^{2"}$, **H**$^{3"}$, **H**$^{7"}$, **H**$^{8"}$, **H**$^{12"}$, **H**$^{13"}$, **H**$^{17"}$, **H**$^{18"}$), 8.57 (s, 1H, **H**^{6}), 8.23 (d, *J* = 8.1 Hz, 2H, **H**$^{23"}$), 8.17-8.21 (m, 6H, **H**$^{26"}$), 7.90 (d, *J* = 8.1 Hz, 2H, **H**$^{22"}$), 7.76-7.84 (m, 9H, **H**$^{27"}$, **H**$^{28"}$), 6.23 (t, *J* = 6.6 Hz, 1H, **H**$^{1'}$), 5.33 (d, *J* = 4.5 Hz, 1H, **H**^{9}), 5.29 (t, *J* = 5.1 Hz, 1H, **H**^{10}), 4.31-4.39 (m, 1H, **H**$^{3'}$), 3.86-3.93 (m, 1H, **H**$^{4'}$), 3.64-3.79 (m, 2H, **H**$^{5'}$), 2.18-2.34 (m, 2H, **H**$^{2'}$)

^{13}**C NMR** (100 MHz, DMSO-d_6): δ = 161.6 (C, **C**^{4}), 149.5 (C, **C**^{2}), 149.4 - 149.0 (C, **C**^{5}, **C**^{10}, **C**^{15}, **C**^{20}), 144.1 (CH, **C**^{6}), 134.5 (CH, **C**$^{22"}$), 134.1 (CH, **C**$^{26"}$) 131.8 - 131.4 (CH, **C2"**, **C3"**, **C7"**, **C8"**, **C12"**, **C13"**, **C17"**, **C18"**), 129.4 (CH, **C**$^{23"}$), 127.5 (CH, **C**$^{28"}$), 126.6 (CH, **C**$^{27"}$), 121.6 (C, **C**$^{21"}$), 120.5 (C, **C**$^{25"}$), 120.4, 119.3 (C, **C**$^{24"}$), 98.3 (C, **C**^{7}), 91.9 (C, **C**^{8}), 87.7 (CH_2, **C**$^{4'}$), 85.0 (CH, **C**$^{1'}$), 83.5 (C, **C**^{5}), 70.0 (CH, **C**$^{3'}$), 59.7 (CH_2, **C**$^{5'}$) 40.3 (CH_2, **C**$^{2'}$)

ESI negative ($C_{55}H_{38}N_6O_5Zn$): Calculated mass 926.22, observed mass 925.3 $[M-H]^-$

8.3.23 Synthesis of acetylene linked porphyrin triphosphate (24)

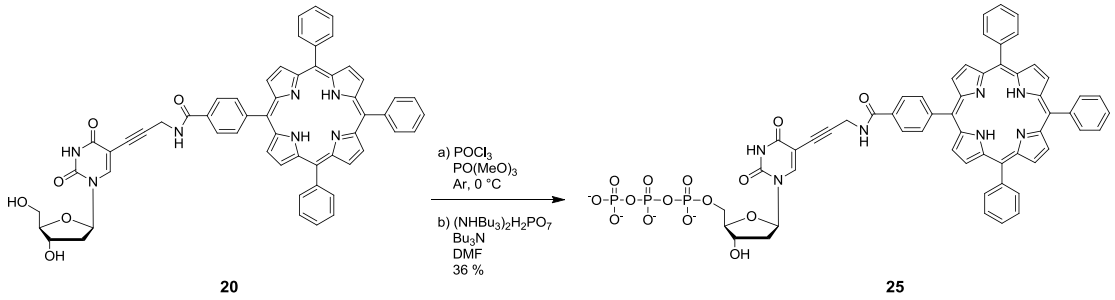


22 **24**

Porphyrin **22** (16 mg, 18.5 μmol, 1.0 eq.) was dried *in vacuo* for 2 hr prior to use. Trimethyl phosphate (0.2 mL) and phosphoryl chloride (2 μL, 24.1 μmol, 1.3 eq.) were added and the reaction stirred at 0 ˚C under argon. A change in colour from purple to green was observed. After 3 hr it was found that the reaction hadn't gone to completion so additional phosphoryl chloride (2 μL, 24.1 μmol, 1.3 eq.) was added and the reaction was stirred for a further 15 hr. Tributylammonium pyrophosphate (51 mg, 92.5 μmol, 5.0 eq.), DMF (1.5 mL) and tributylamine (150 μL) were added and the reaction stirred for 2 hr. Upon addition of these reagents the reaction mixture returned to a purple colour. The reaction was quenched with the addition of 2M triethylammonium bicarbonate (2 mL). The crude product was purified by C18 reverse-phase column chromatography using water/methanol (5 - 100 %) as eluent. The product was isolated as a purple solid (4.1 mg, 3.7 μmol, 20 %).

^{31}**P NMR** (162 MHz, CD_3OD) δ = -8.89 (m, P_γ), -10.24 (m, P_α), -22.02 (m, P_β)

8.3.24 Synthesis of Flexible linked porphyrin triphosphate (25)

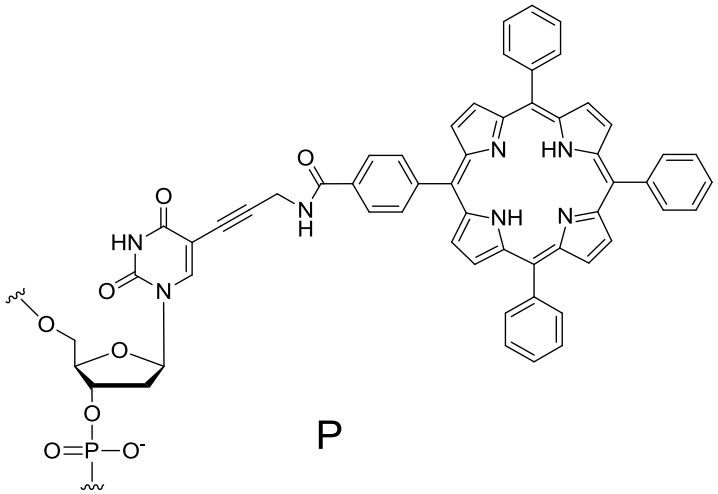


20 25

Porphyrin **20** (22 mg, 24.6 μmol, 1.0 eq.) was dried *in vacuo* for 2 hr prior to use. Trimethyl phosphate (0.2 mL) and phosphoryl chloride (3 μL, 32.0 μmol, 1.3 eq.) were added and the reaction stirred at 0 ˚C under argon. A change in colour from purple to green was observed. The reaction was stirred for 24 hr. Tributylammonium pyrophosphate (134 mg, 246.0 μmol, 10.0 eq.), DMF (1.5 mL) and tributylamine (150 μL) were added and the reaction stirred for 3 hr at 0 ˚C. Upon addition of these reagents the reaction mixture returned to a purple colour. The reaction was quenched with the addition of 2M triethylammonium bicarbonate (2 mL). The crude product was purified by C18 reverse-phase column chromatography using water/methanol (5 - 100 %) as eluent. The product was isolated as a purple solid (11 mg, 9.0 μmol, 36 %).

^{31}**P NMR** (162 MHz, CD_3OD): δ = -10.40 (d, J = 20.8 Hz, P_γ), -11.55 (d, J = 21.8 Hz, P_α), -23.82 (t, J = 20.8 Hz, P_β)

8.4 Synthesis of porphyrin-modified Dickerson dodecamer DNA



Flexible amide linked porphyrin phosphoramidite **8** (108 mg, 76 μmol, 1.0 eq.) was dissolved in DCM:MeCN (1:1, 2.0 mL, 3.8 mM) and coupled by passing the amidite solution (192 μL, 7.3 eq.) through the CGP support over a period of 5 minutes. DNA strands were cleaved from the solid support by passing concentrated ammonia solution (35 % in H_2O, S.G. = 0.88, 1 mL) through the solid support using two syringes at either end for 1 hour. The purple DNA solution was then removed from the solid support and deprotected by heating at 50 °C for 4 hours. Strands were purified by Glen-Pak Columns before concentrating *in vacuo*. Further purification was carried out by reverse-phase HPLC.

Strand	**Sequence**	**Calculated ε_{260} (mol^{-1} dm^3 cm^{-1})**	**Yield (nmoles)**
DD1	5' – CGC GAA **P**TC GCG– 3'	110700	615
DD2	5' – **P** CGC GAA TTC GCG – 3'	119500	426

Concentrations of oligonucleotides were calculated using the Beer Lambert law with the absorption at 260 nm. Molar extinction coefficients were obtained by the molar extinction calculator provided by Applied Biosystems, and by replacing the appropriate thymidines with the value of the porphyrin-dU (ε_{260} = 13460 mol^{-1} dm^3 cm^{-1}) as determined previously.[94]

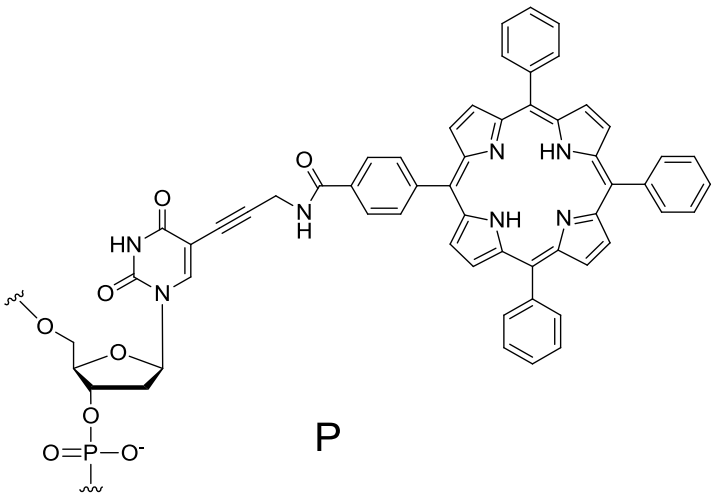
8.4.1 Zinc metallation of porphyrin DNA

Single stranded porphyrin DNA (1.0 eq.) was heated with zinc acetate dihydrate (1.0 M, 1000 eq. per porphyrin) to 85 °C for 5 minutes, the DNA precipitated. EDTA solution (0.5 M, pH 8.1, 100 eq. per Zn ion) was added to redissolve the DNA and sequester excess metal ions. An equal volume of NaCl solution (100 mg mL^{-1} NaCl in H_2O) was added to the metallated porphyrin DNA sample prior to loading onto a conditioned Glen-Pak column. Excess EDTA and zinc salts were eluted with (5 % acetonitrile in 100 mg/mL sodium chloride solution, 2 x 1.0 mL) and then the metallated DNA was eluted with MeCN:H_2O (1:1, 4 mL) and concentrated *in vacuo*. The efficiency of the zinc metallation was checked by UV-Vis and fluorescence spectroscopy.

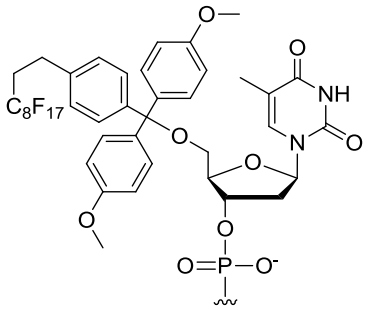
8.4.2 Crystallography details

DD1Zn was dissolved in 10 mM HEPES buffer to give a final concentration of 1 mM. The sample was supplemented with 10 mM magnesium chloride. Three screens were used for crystallisation attempts: Natrix HT deep well block (Hampton Research), PEGRx HT (Hampton Research) and PACT premier HT (Molecular Dimensions). The screens were prepared using a Rigaku Desktop Alchemist liquid handling system and an Art Robbins Gryphon small modular dispenser. Screens were checked weekly for signs of crystal growth.

8.5 Synthesis of porphyrin-modified G-quadruplexes



P



T*

Flexible amide linked porphyrin phosphoramidite **8** (137 mg, 96.4 μmol, 1.0 eq.) was dissolved in DCM:MeCN (1:1, 3.0 mL, 3.1 mM) and coupled by passing the amidite solution (192 μL, 6.0 eq.) through the CGP support over a period of 5 minutes. DNA strands were cleaved from the solid support by passing concentrated ammonia solution (35 % in H_2O, S.G. = 0.88, 1 mL) through the solid support using two syringes at either end for 1 hour. The purple DNA solution was then removed from the solid support and deprotected by heating at 50 °C for 4 hours. Strands were purified by Fluoro-Pak II Columns before concentrating *in vacuo*, desalting with NAP-10 columns and concentrated *in vacuo*.

Strand	Sequence	Calculated ε_{260} (mol^{-1} dm^3 cm^{-1})	Yield (nmoles)
wtTel25b	5' – **T***A GGG TTA GGG TTA GGG TTA GGG TT – 3'	253100	436
Q1	5' – **T***A GGG **P**TA GGG TTA GGG TTA GGG TT – 3'	253100	223
Q2	5' – **T***A GGG T**P**A GGG TTA GGG TTA GGG TT – 3'	253100	291
Q3	5' – **T***A GGG TTA GGG **P**TA GGG TTA GGG TT – 3'	253100	306
Q4	5' – **T***A GGG TTA GGG T**P**A GGG TTA GGG TT – 3'	253100	335
Q5	5' – **T***A GGG TTA GGG TTA GGG **P**TA GGG TT – 3'	253100	427
Q6	5' – **T***A GGG TTA GGG TTA GGG T**P**A GGG TT – 3'	253100	390
Q7	5' – **T*P** GGG TTA GGG TTA GGG TTA GGG **P**T – 3	253100	219
Q8	5' – **T***A GGG T**P**A GGG T**P**A GGG TTA GGG TT – 3'	253100	342

8.5.1 UV-Vis absorption spectroscopy

The thermal melting temperatures of the quadruplexes were determined using UV melting. Oligonucleotide solutions (5 μM) were prepared in 10 mM lithium phosphate buffer (pH = 7.4), which was supplemented with 10 mM potassium chloride to give a total volume of 1 mL. Samples were heated to 95 °C and held at this temperature for 5 minutes and allowed to cool to room temperature over a period of 4 hours.
UV melting experiments:

25 °C →80 °C at 0.2 °C / min - hold for 1 minute at 80 °C
80 °C →30 °C at 0.2 °C / min - hold for 1 minute at 30 °C

Melting curves were recorded as an average of two experiments. Melting transitions were calculated from the first derivatives of the melting profile using Origin software.

8.5.2 Circular dichroism spectroscopy of G-quadruplex DNA

Oligonucleotide solutions (2.5 μM) were prepared in 10 mM lithium phosphate buffer (pH 7.4) supplemented with either 100 mM potassium chloride or 100 mM sodium chloride. The samples were heated to 95 °C for 5 minutes and allowed to cool to room temperature over a period of 4 hours. Spectra were recorded between 200 and 600 nm in 1 cm path length cuvettes. 1 scan was taken with a 1 nm scan step, 2 seconds per point and a bandwidth of 2 nm on an Applied Photophysics Chirascan spectrometer. A buffer baseline was subtracted from each spectrum. The data was recorded in mdeg and corrected to delta epsilon using the formula $\Delta\varepsilon = \theta$ / (10 × conc. x path length x 3298); conc. is in mol/litre, path length is in cm.

Competition experiments were carried out by adding 1.2 equivalents of the complementary DNA strand to one equivalent of the porphyrin-modified DNA. These experiments were carried out in 10 mM lithium phosphate (pH 7.4) supplemented with either 100 mM potassium chloride or 100 mM sodium chloride. No change in spectrum was observed when the complementary strand was added at room temperature except for **wtTel25b**. The samples were then heated to 95 °C for 5 minutes and then allowed to anneal over a period of 4 hours. The spectra were recorded once the samples had finished annealing.

8.5.3 ^{1}H Nuclear magnetic resonance spectroscopy of G-quadruplex DNA

One dimensional ^{1}H NMR spectroscopy experiments were performed on a Varian Inova 600 MHz spectrometer. The concentration of all samples was 100 μM with a volume of 500 μL. Samples were supplemented with 5 % D_2O. The quadruplexes were dissolved in 0.1 M potassium phosphate buffer, pH = 7.4. The samples were heated to 95 ° C for 5 minutes and left to cool to room temperature over 4 hours to ensure formation of the quadruplex structure. Spectra were recorded at 25 °C with WATERGATE water suppression. Data were processed using VNMR software (Varian Inc.) with zero filling and resolution enhancement.

8.5.4 Polyacrylamide gel electrophoresis of G-quadruplexes

Denaturing PAGE was performed using 20 % PAGE (27:1 acrylamide:bis-acrylamide, Fluka) with 7 M urea, TBE running buffer; native PAGE was run in the absence of urea (37.5:1 acrylamide:bis-acrylamide). Samples were loaded either as annealed quadruplexes, or denatured by treatment with 80 % formamide (1 mM EDTA, pH 8) at 90 °C for 5 min. The DNA strands were visualised by UV illumination.

8.5.5 Polyacrylamide gel electrophoresis with ^{32}P labelled G-quadruplexes

Oligonucleotide labelling was carried out by Dr Annie Cardew and Dr David Rusling of the Fox research group, University of Southampton. Oligonucleotide (2 μL), ^{32}P labelled ATP (2 μL), T4 polynucleotide kinase buffer (2 μL) and T4 polynucleotide kinase (1 μL) were mixed in distilled water (13 μL) at 37 °C for 1 hr. DNase 1 stop (10 μL) was added and the mixture heated to 95 °C for 3 minutes. Labelled oligonucleotides were purified by denaturing gel electrophoresis.

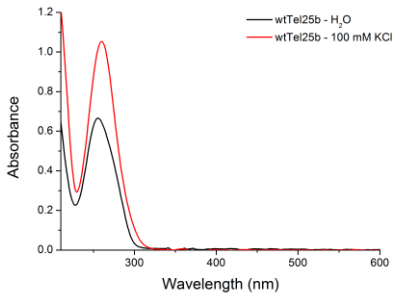
Denaturing gel electrophoresis was performed using polyacrylamide gels. The gel was made up using sequgel (30 mL), 10 × TBE buffer made up in diluent (5 mL), diluent (15 mL), 20 % Ammonium persulfate (200 μL) and TMED (40 μL). Radiolabelled oligonucleotides were supplemented with KCl (100 μM). After electrophoresis the gels were fixed, dried and subjected to autoradiography using a phosphorimager screen (Kodak), which was scanned with a Storm phosphorimager.

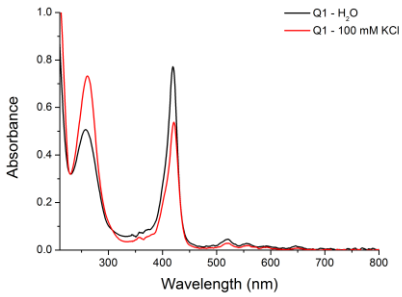
Non-denaturing gel electrophoresis was performed using polyacrylamide gels. The gel was made up using accugel (30 mL), 10 × TBE buffer (5 mL), distilled water (15 mL), 20 % Ammonium persulfate (200 μL) and TMED (40 μL). Radiolabelled oligonucleotides were supplemented with KCl (100 μM). After electrophoresis the gels were fixed, dried and subjected to autoradiography using a phosphorimager screen (Kodak), which was scanned with a Storm phosphorimager.

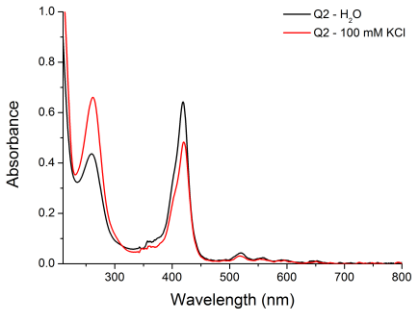
8.5.6 Molecular dynamics simulations

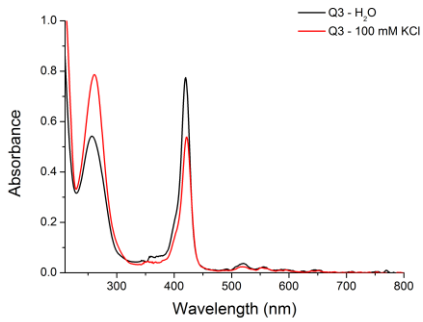
Molecular modelling was carried out by Peter Várnai of the University of Sussex. The crystal structure of wtTel22 (pdb id 1kf1) was used to prepare the parallel quadruplex structure for the modified sequence of wtTel25b (**Q8**) using Nucleic Acid Builder.[225] Parameters for the porphyrin and the propargyl-amide linker were obtained from the General Amber Force Field[226] and AM1-BCC charges were derived using Antechamber,[227] as implemented in AmberTools15. Parameters for DNA were those of the parmbsc0 modifications[228] to the ff99 Amber force field.[229] The starting structures were solvated by TIP3P water molecules to extend to at least 12 Å in a truncated octahedral box and neutralised by K^+ counterions.[230] The system contained ca. 10,000-15,000 water molecules depending on the starting conformation. The system was energy minimised, heated to 300 K at constant volume under restraints on the solute and allowed to equilibrate at constant pressure (1 atm) using periodic boundary conditions and the particle mesh Ewald method to calculate the long-range electrostatic interactions. The unrestrained production run involved 100 ns simulation time for each starting geometry. The trajectory was visualised and the corresponding movie generated by Visual Molecular Dynamics.[231]

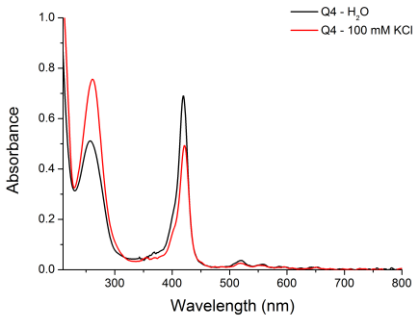
8.5.7 UV-Vis absorption spectra of G-quadruplexes

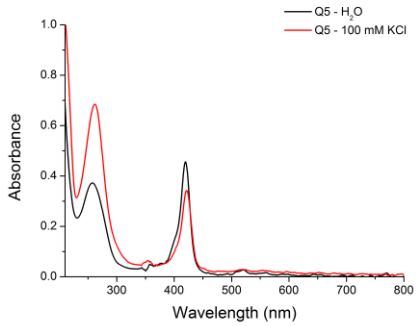


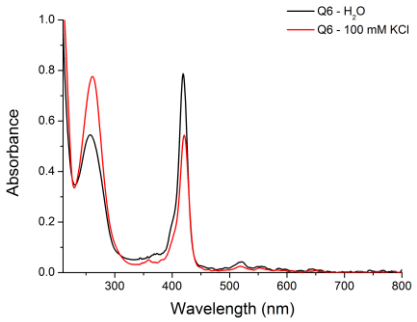


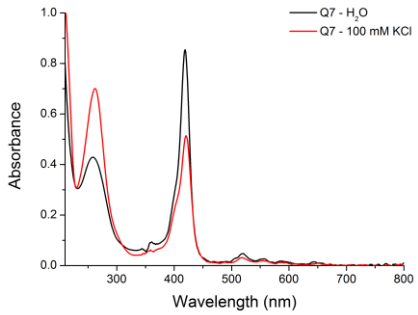


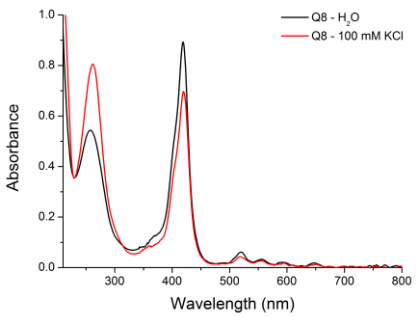




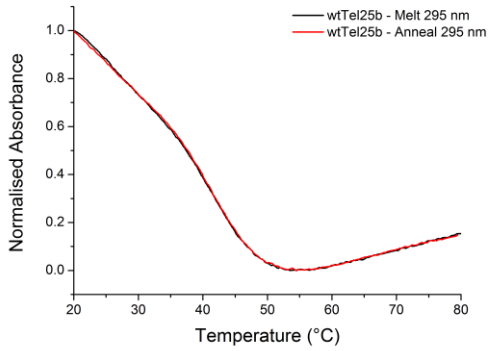


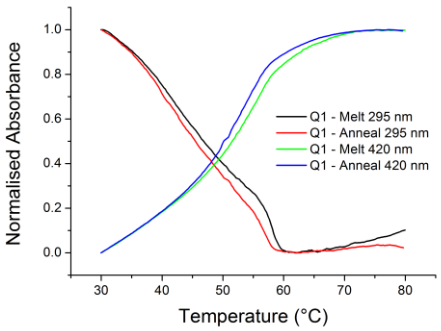


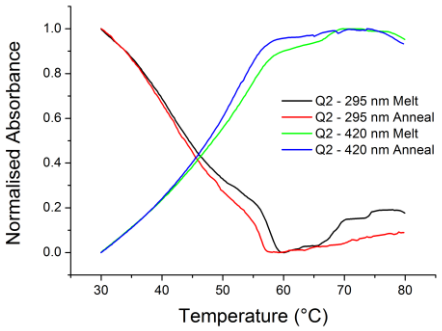


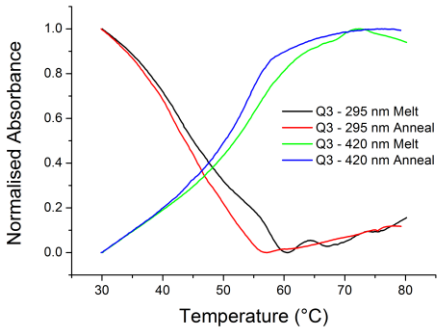


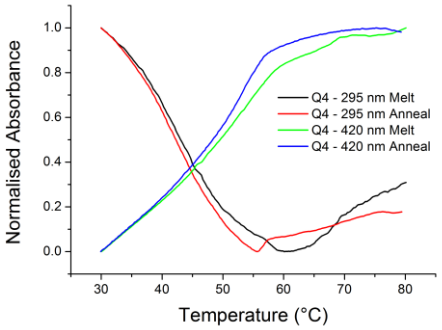
8.5.8 UV-Vis melting of G-quadruplexes

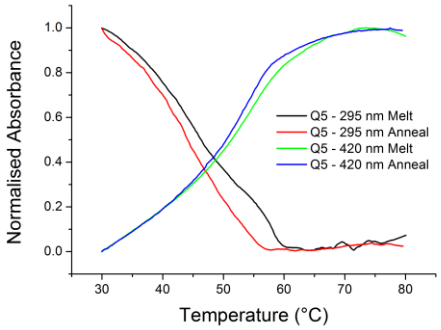


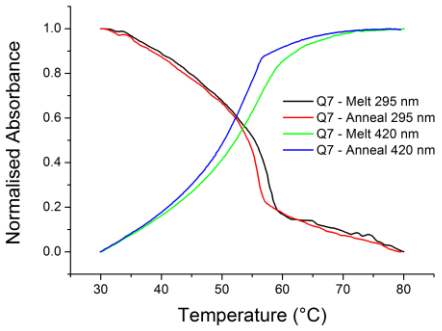


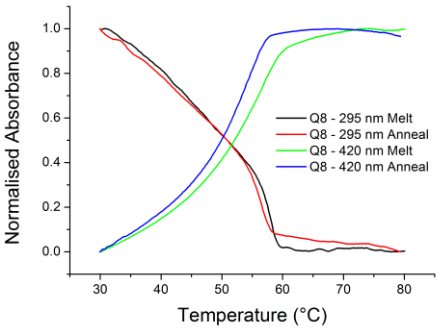












8.5.9 Denaturing PAGE

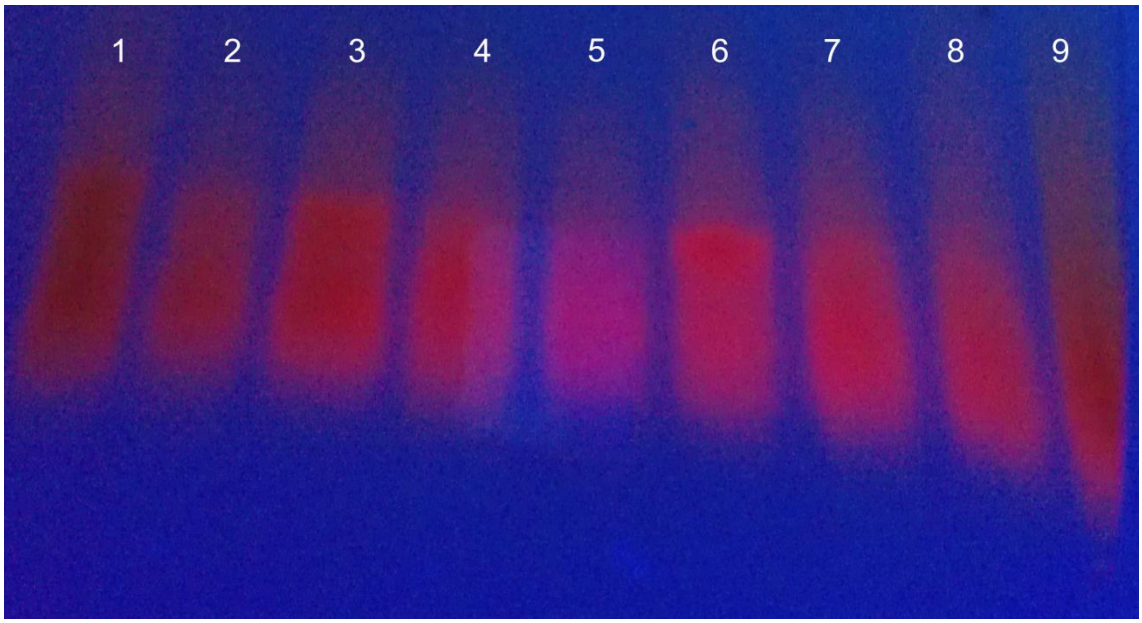


Figure 4.26 Denaturing PAGE visualised at 365 nm.

Lane 1 = **Q6** Lane 2 = **Q5** Lane 3 = **Q4** Lane 4 = **Q3**

Lane 5 = **Q2** Lane 6 = **Q1** Lane 7 = **Q7** Lane 8 = **Q8**

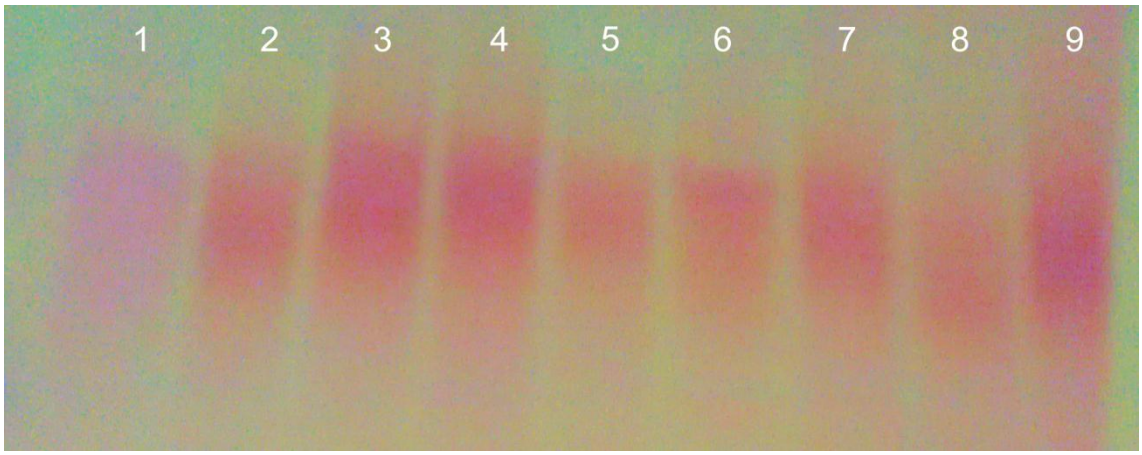


Figure 4.27 Denaturing PAGE visualised at 254 nm.

Lane 1 = **wtTel25b** Lane 2 = **Q6** Lane 3 = **Q5** Lane 4 = **Q4**

Lane 5 = **Q3** Lane 6 = **Q2** Lane 7 = **Q1** Lane 8 = **Q7**

Lane 9 = **Q8**

8.5.10 Native PAGE

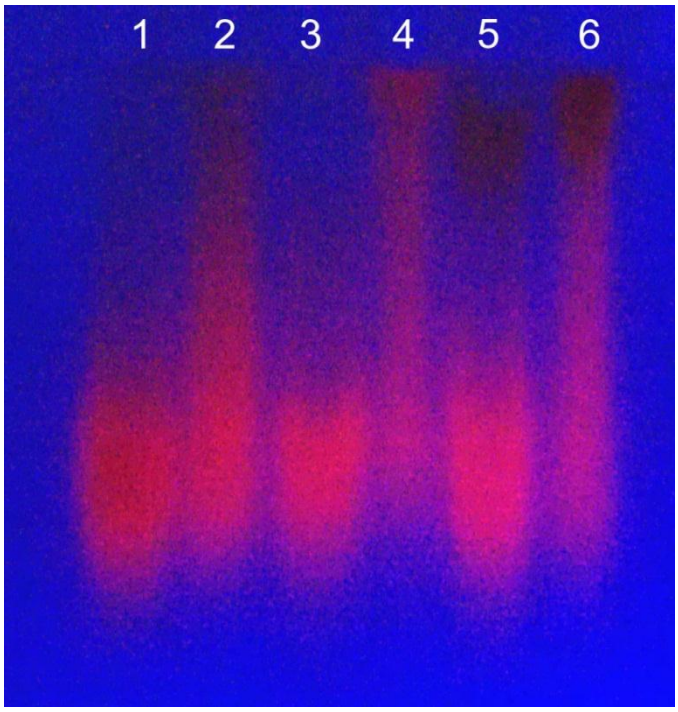


Figure 4.28 Native PAGE visualised at 365 nm.

Lane 1 = **Q5** denatured form Lane 2 = **Q5** G-quadruplex form

Lane 3 = **Q3** denatured form Lane 4 = **Q3** G-quadruplex form

Lane 5 = **Q7** denatured form Lane 6 = **Q7** G-quadruplex form

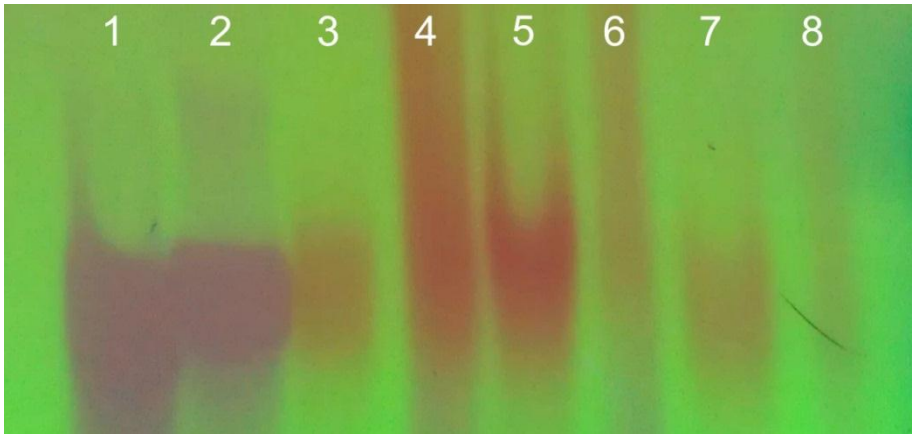


Figure 4.29 Native PAGE visualised at 254 nm.

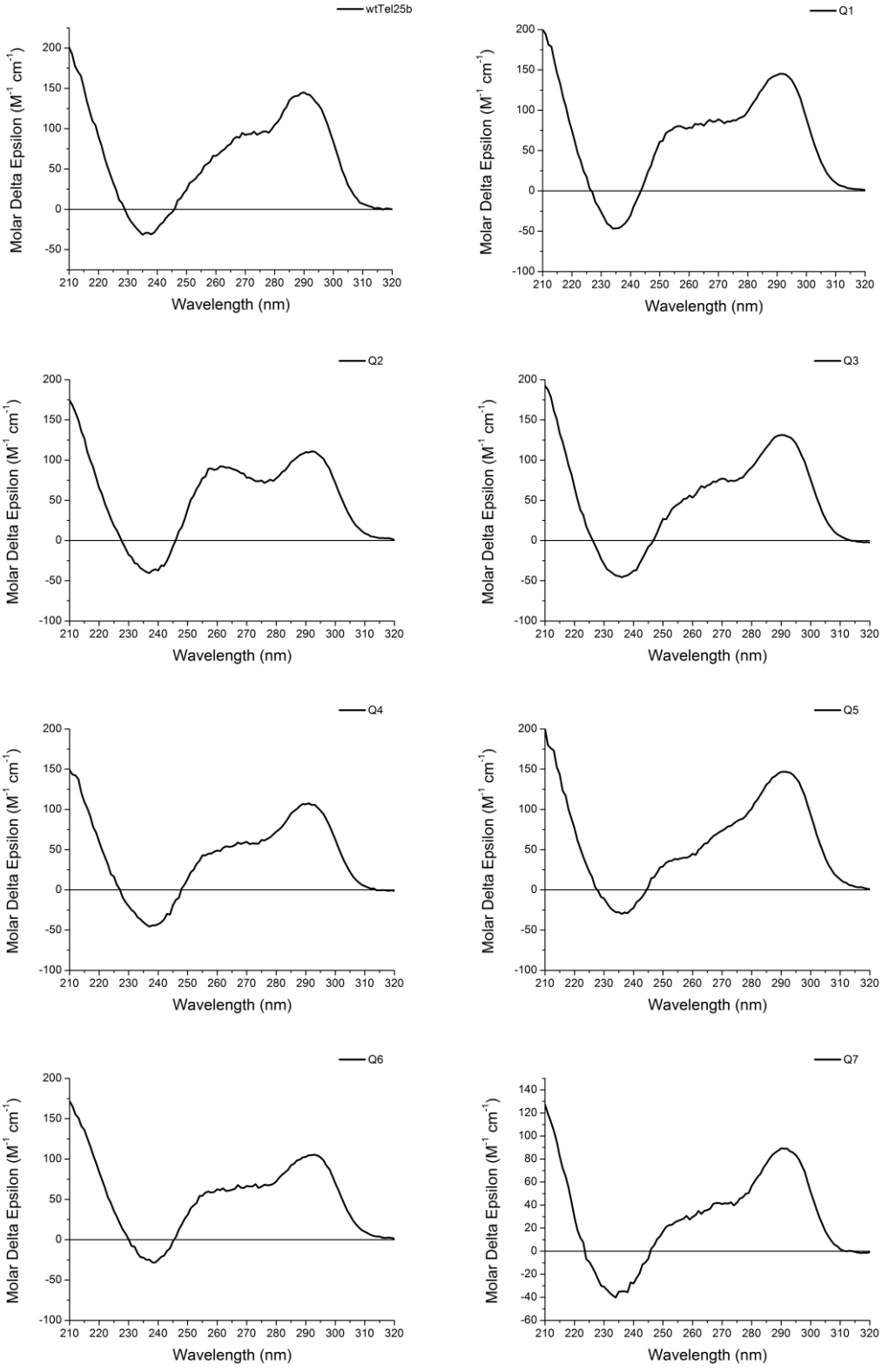
Lane 1 = **wtTel25b** denatured form Lane 2 = **wtTel25b** G-quadruplex form

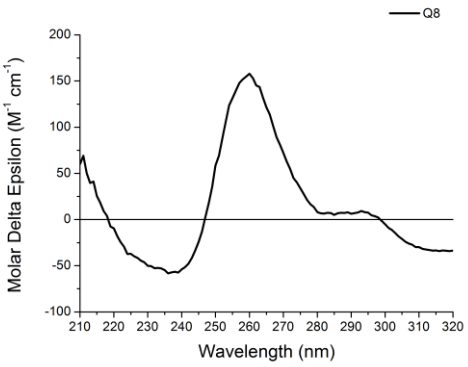
Lane 3 = **Q3** denatured form Lane 4 = **Q3** G-quadruplex form

Lane 5 = **Q5** denatured form Lane 6 = **Q5** G-quadruplex form

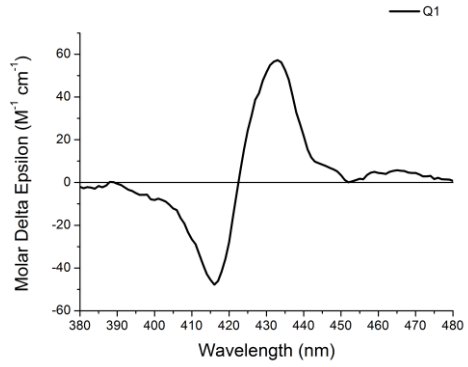
Lane 7 = **Q7** denatured form Lane 8 = **Q7** G-quadruplex form

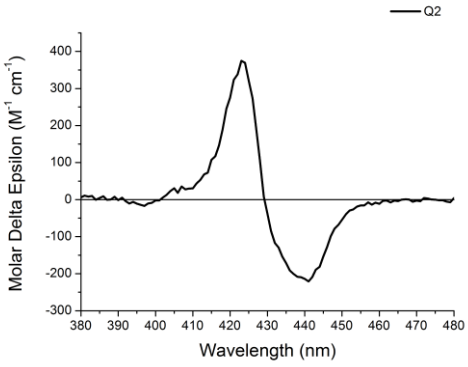
8.5.11 Circular dichroism spectra – 100 mM KCl DNA region

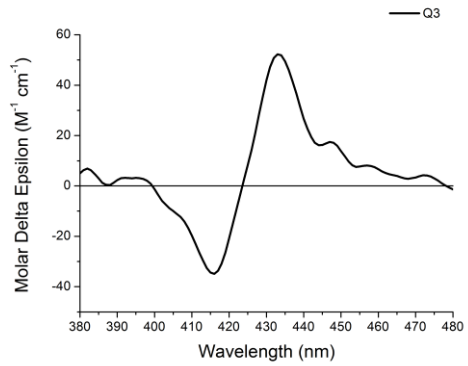


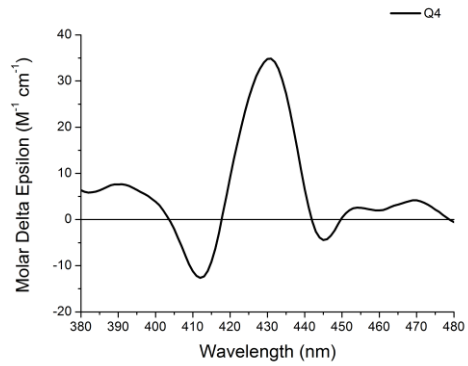


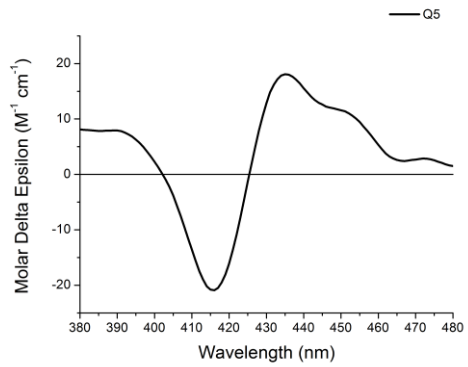
8.5.12 Circular dichroism spectra – 100 mM KCl Porphyrin region

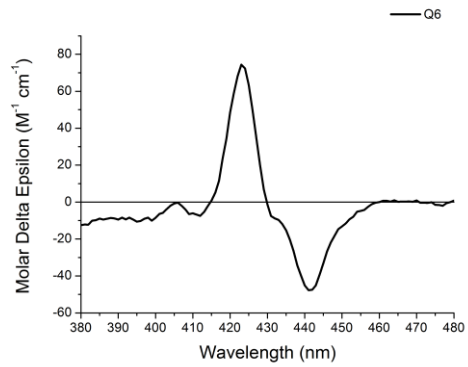


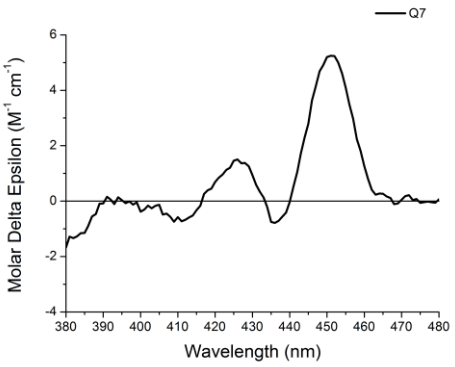


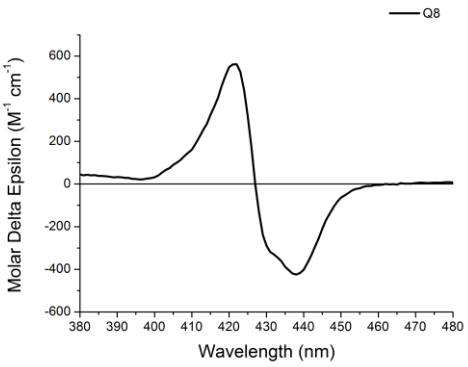




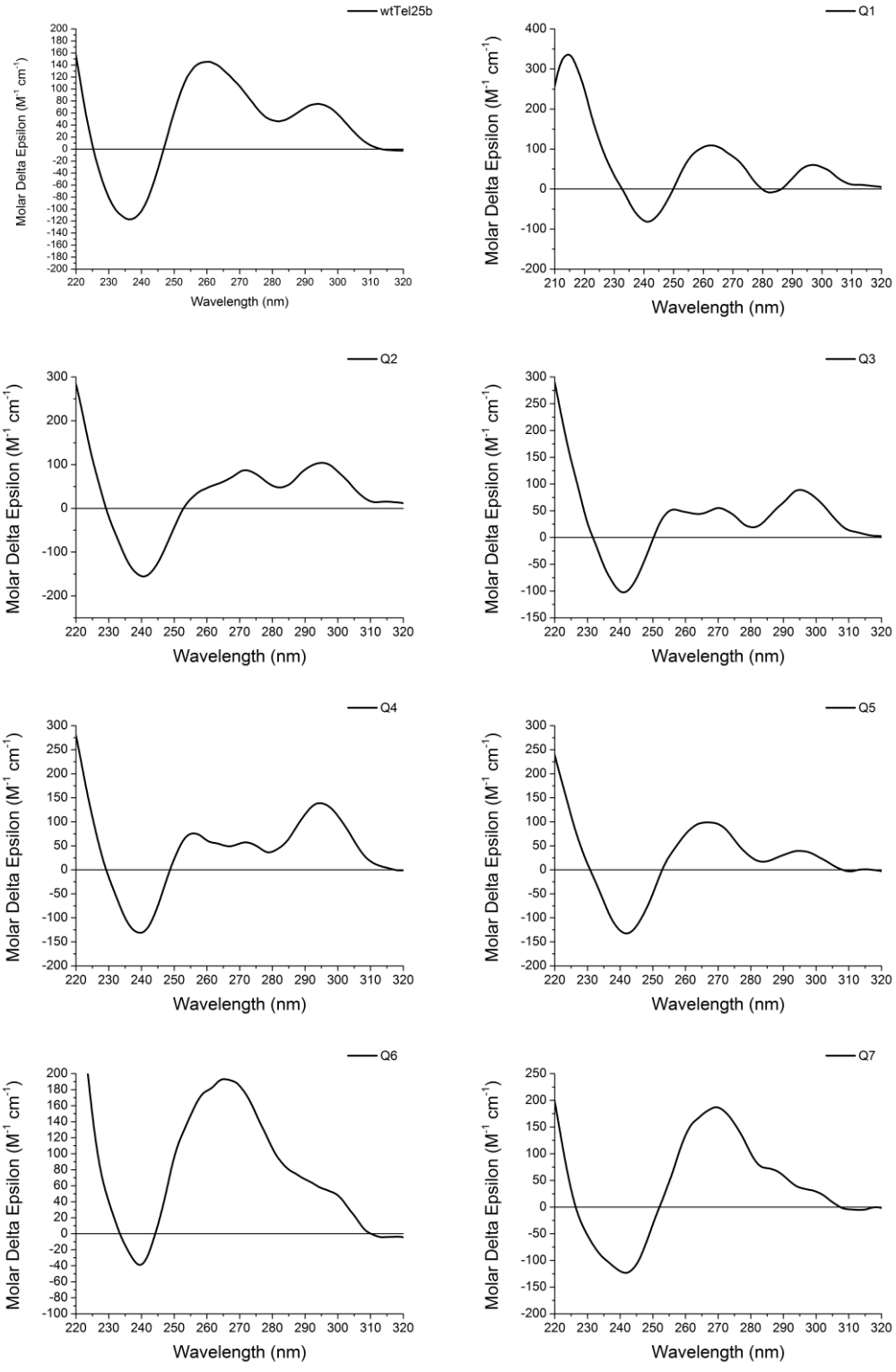


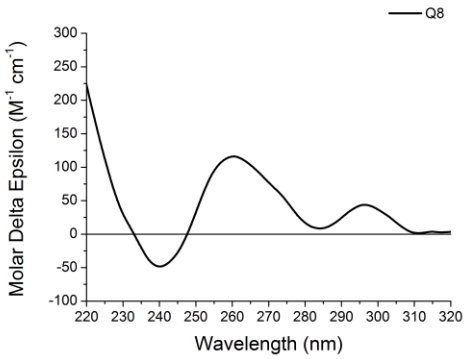



Q7
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)

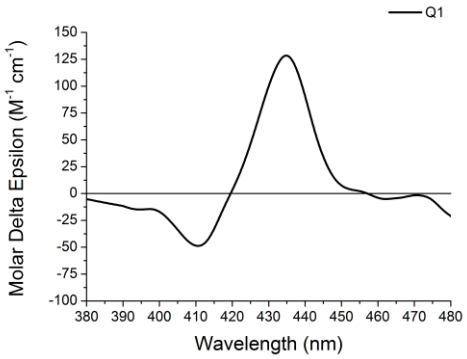

Q8
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)

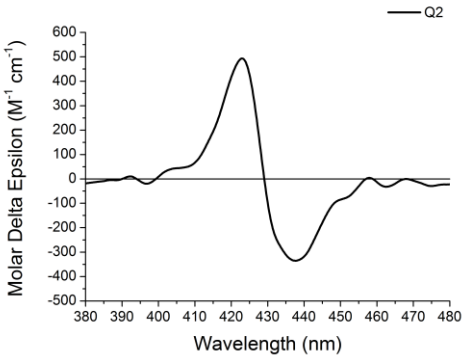
8.5.13 Circular dichroism spectra – 100 mM NaCl DNA region

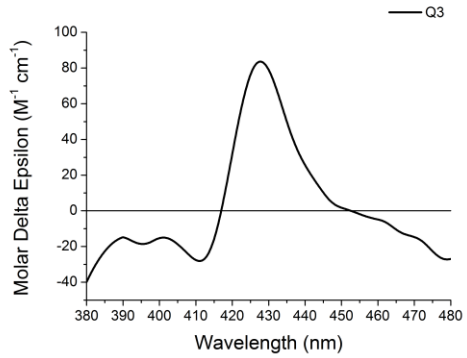


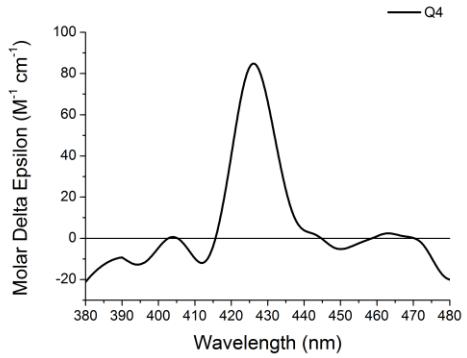


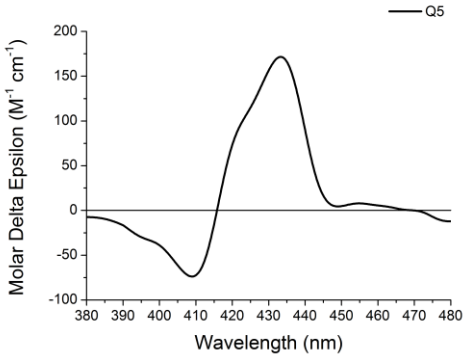
8.5.14 Circular dichroism spectra – 100 mM NaCl Porphyrin region

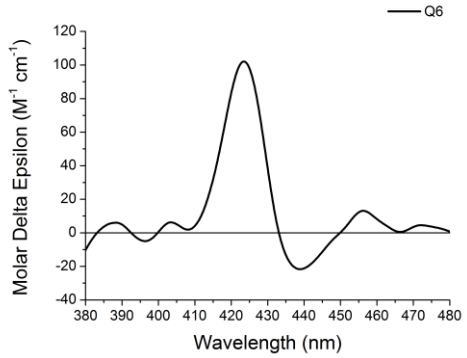


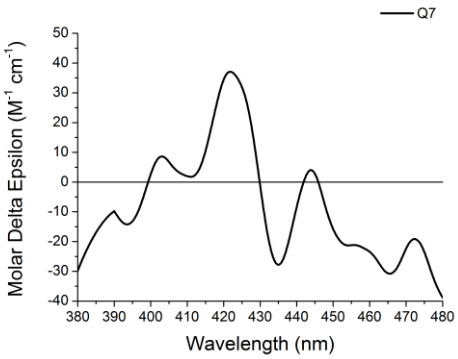


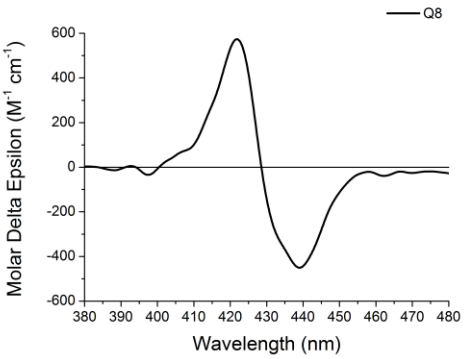




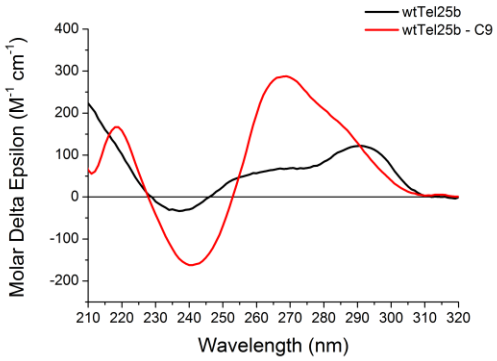


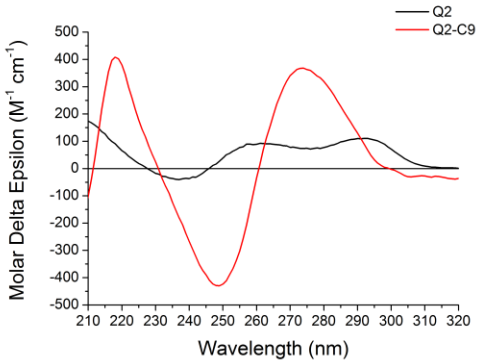


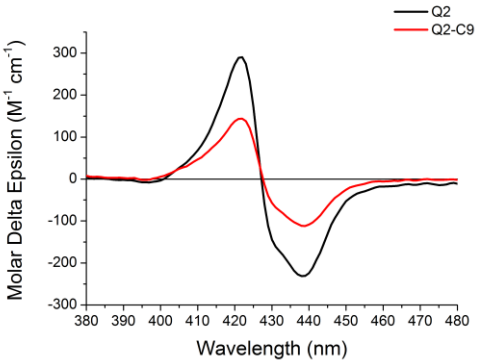

Q7
Molar Delta Epsilon (M-1 cm-1)
50
40
30
20
10
0
-10
-20
-30
-40
380 390 400 410 420 430 440 450 460 470 480
Wavelength (nm)

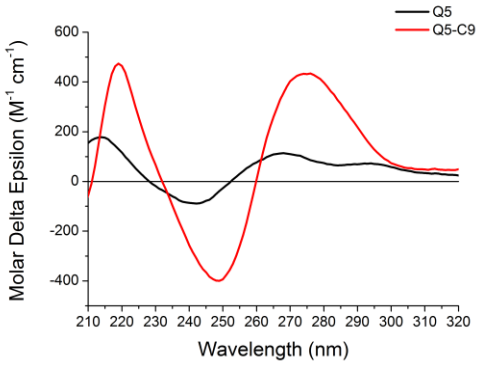

Q8
Molar Delta Epsilon (M-1 cm-1)
600
400
200
0
-200
-400
-600
380 390 400 410 420 430 440 450 460 470 480
Wavelength (nm)

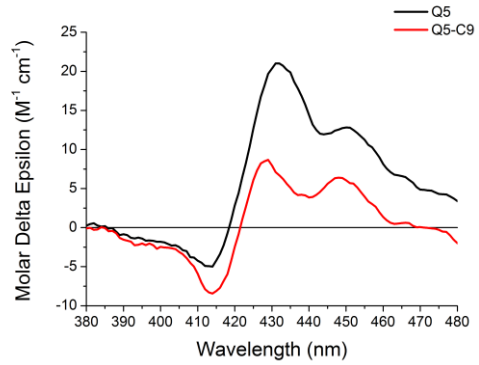
8.5.15 Circular dichroism – G-quadruplex + complementary strand – 100 mM KCl

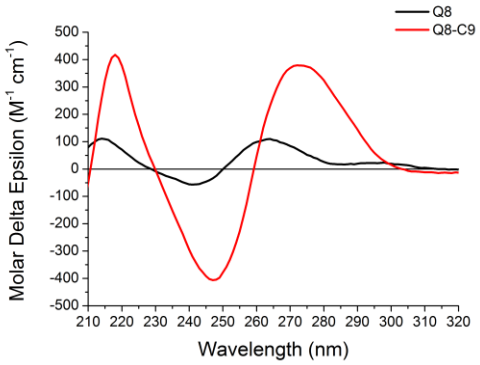


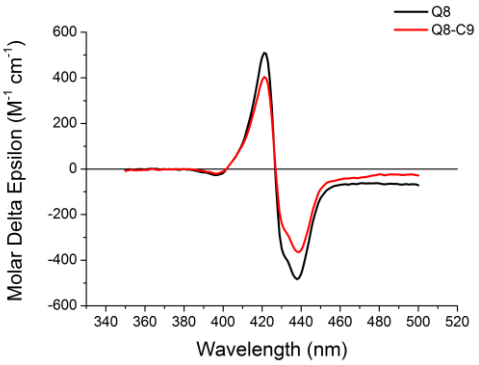




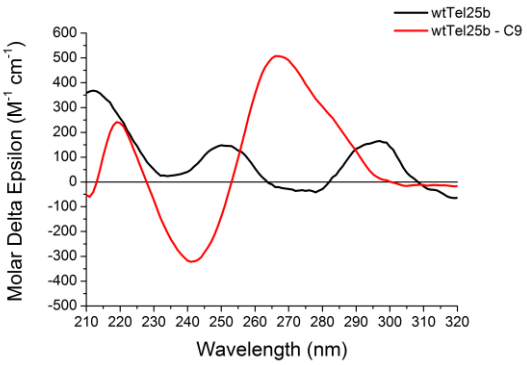


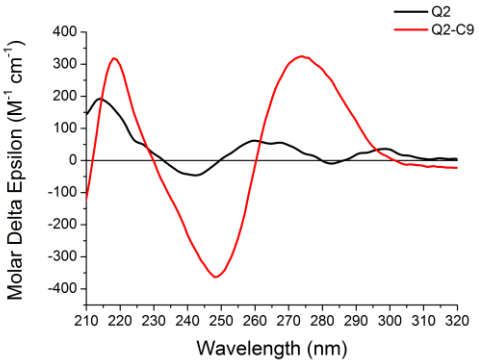


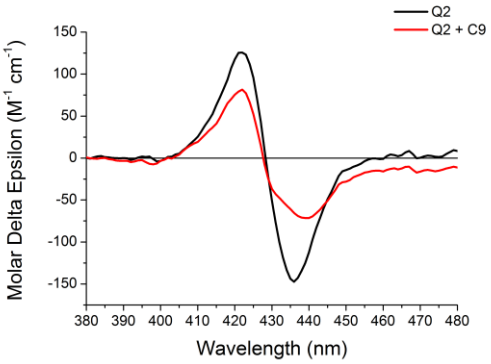


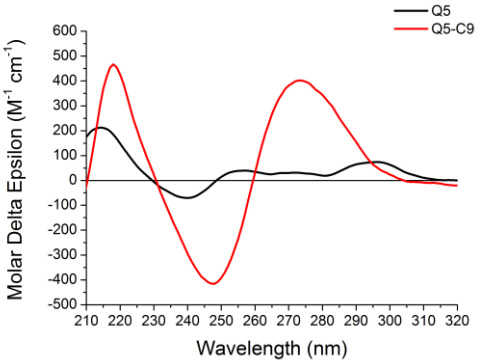


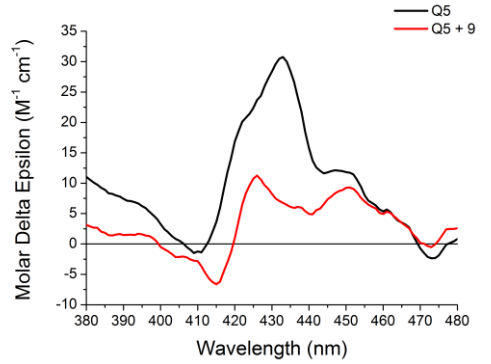
8.5.16 Circular dichroism – G-Quadruplex + Complementary Strand – 100 mM NaCl

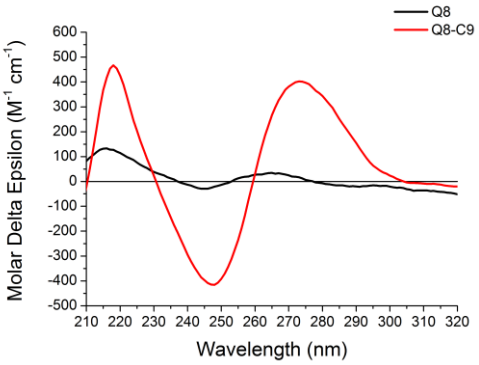


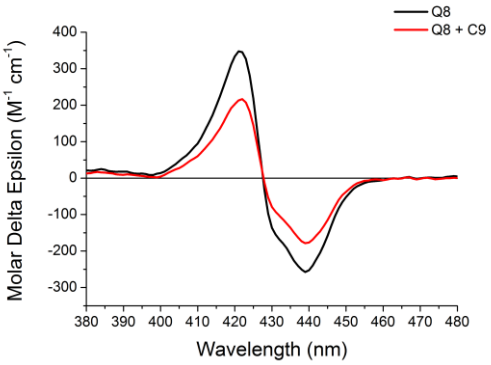






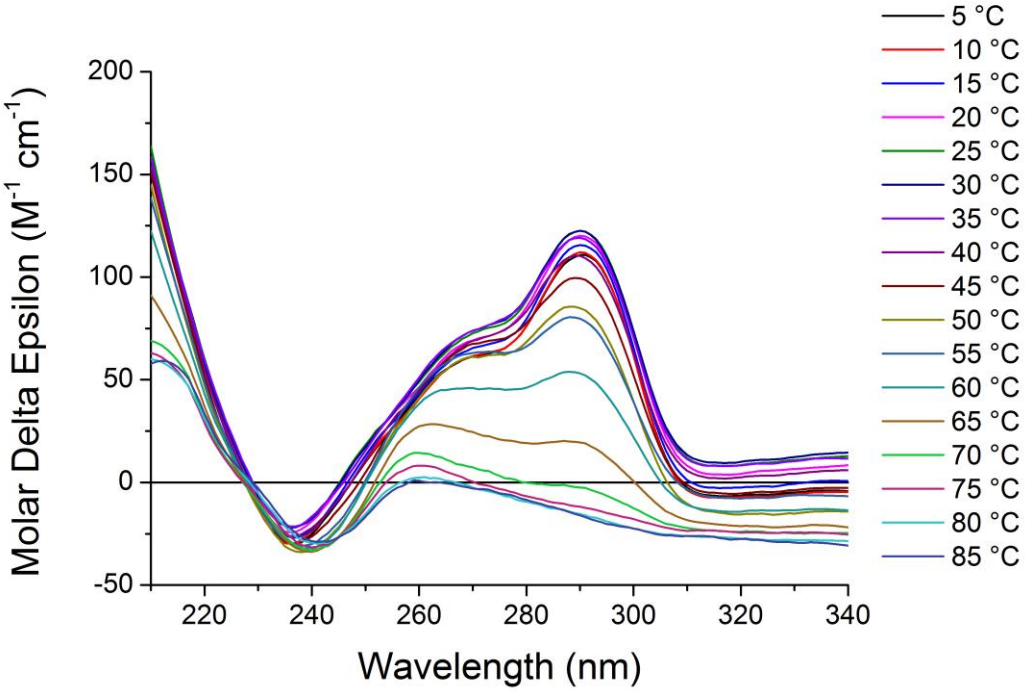




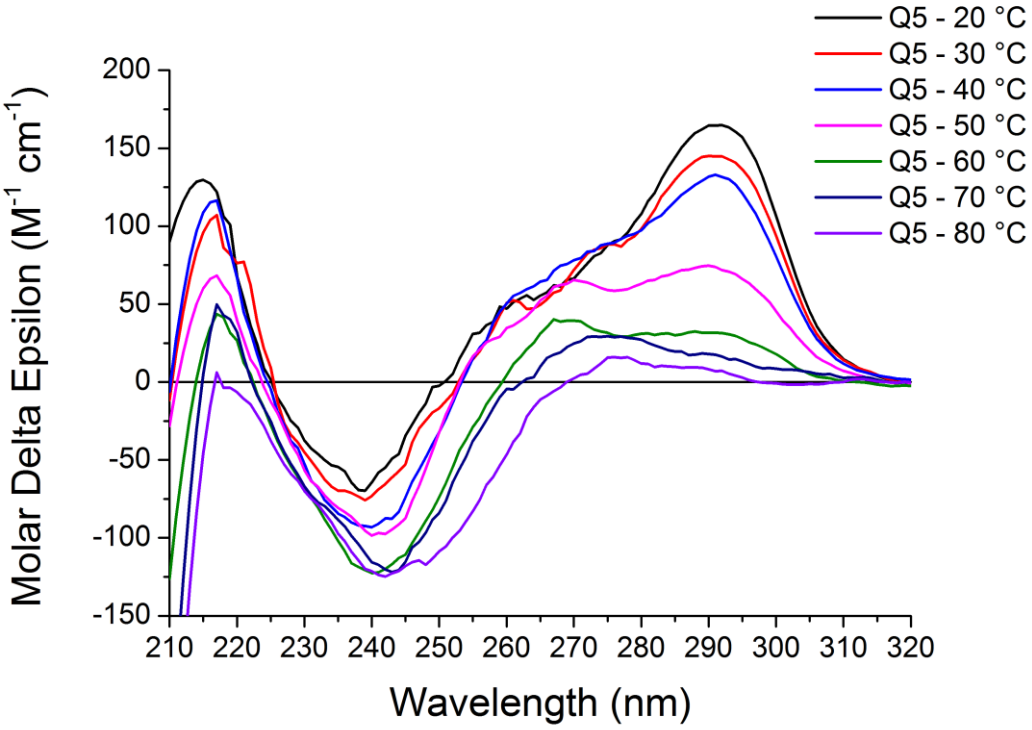


8.5.17 Circular dichroism melting in 100 mM potassium chloride

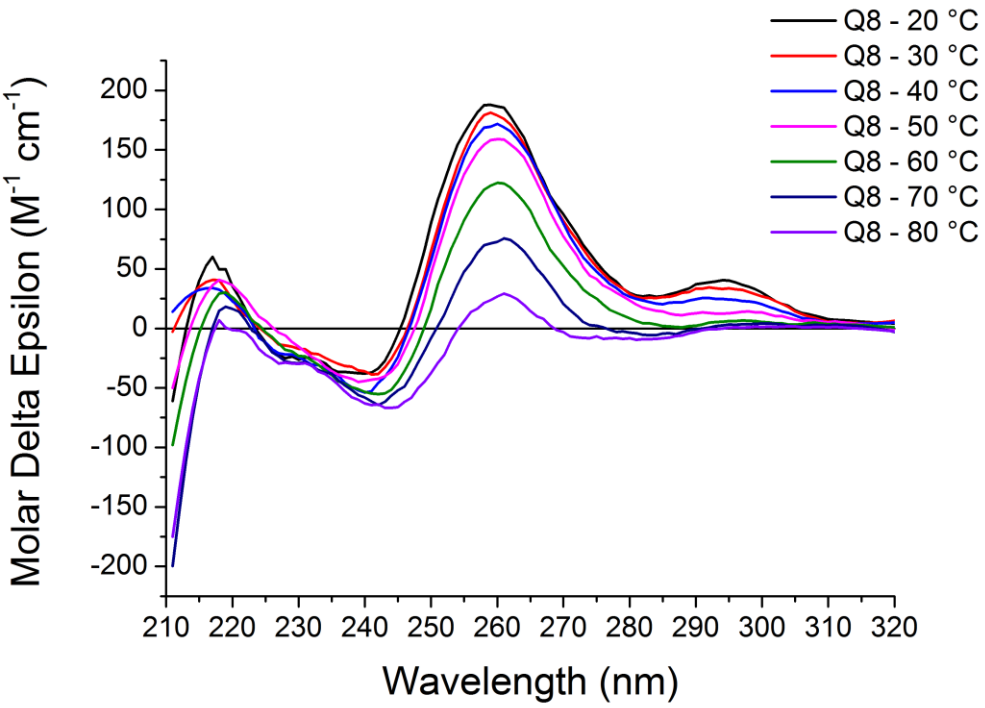
wtTel25b



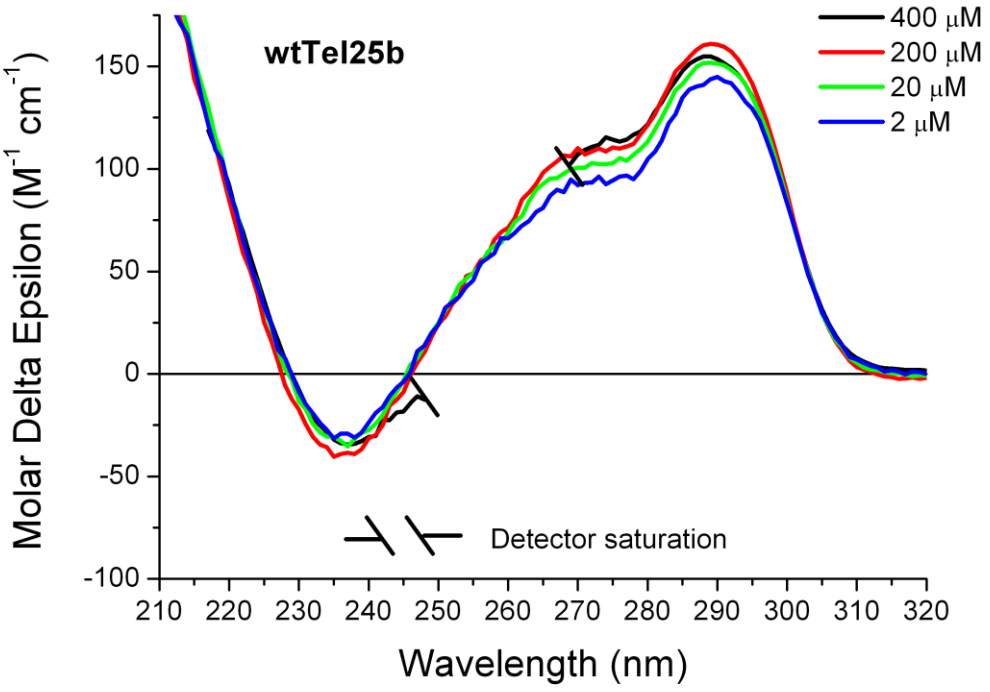
Q5



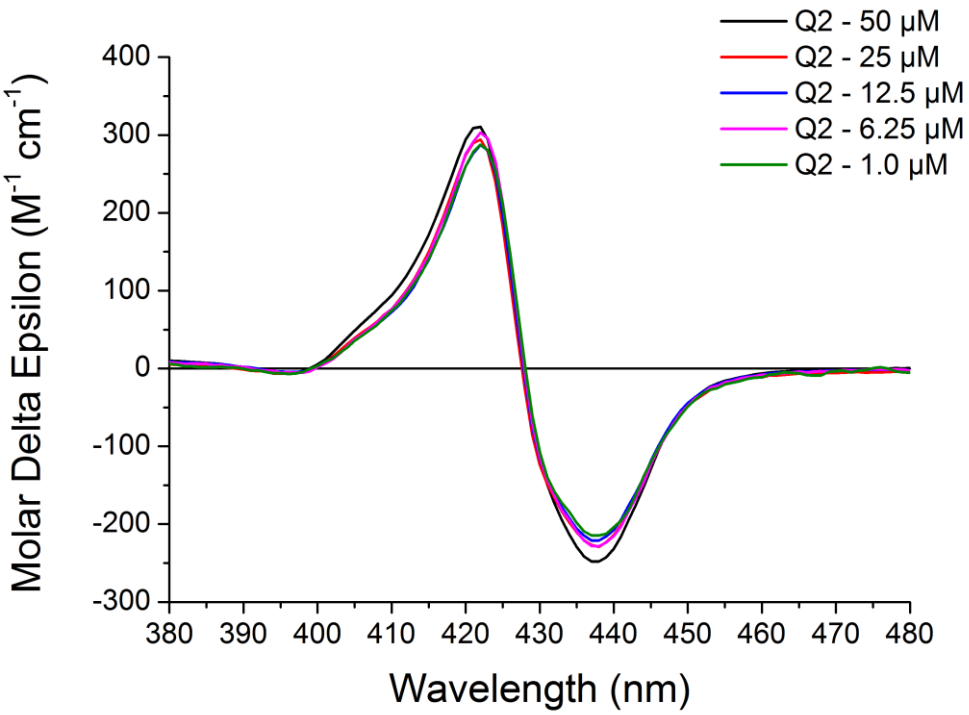
Q8



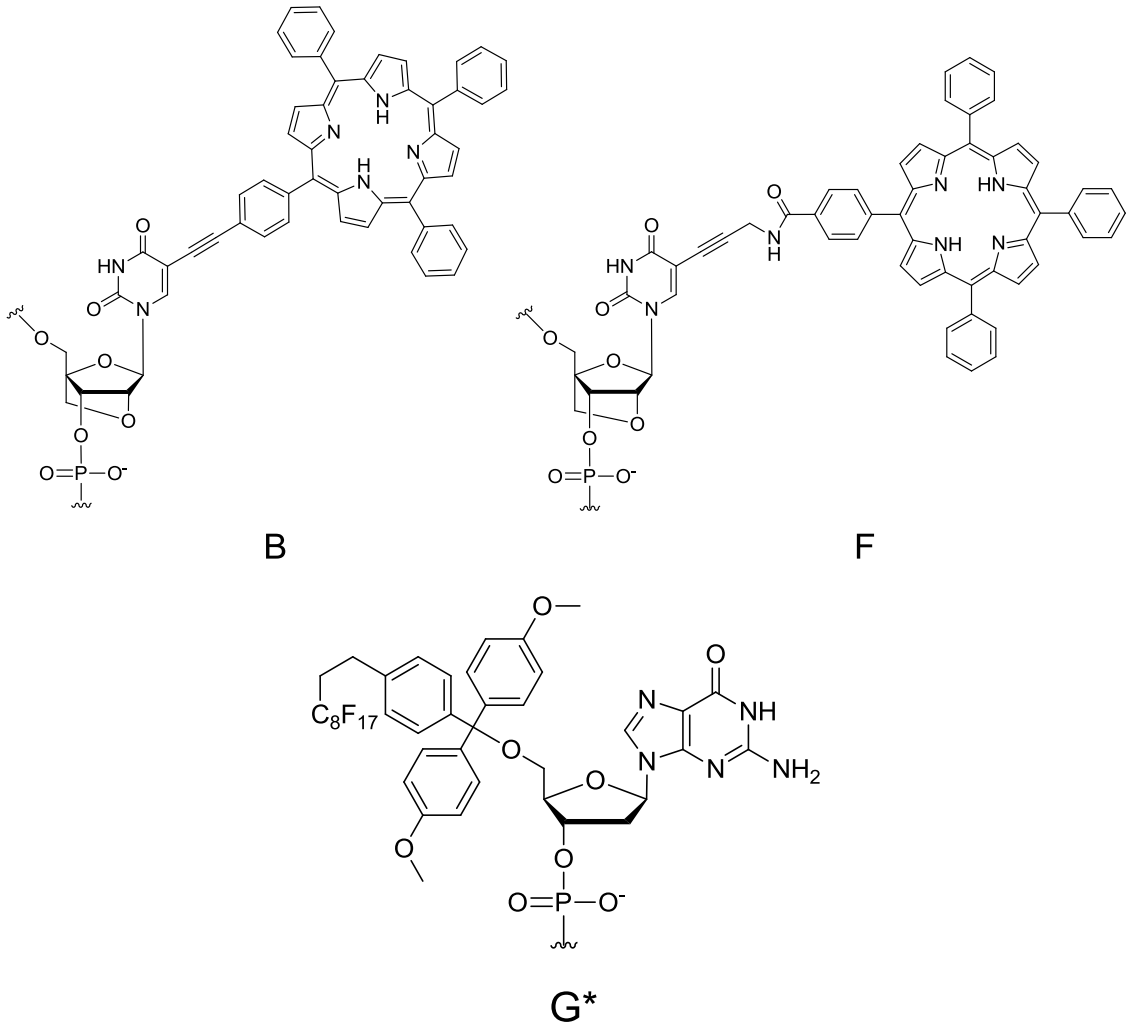
8.5.18 Concentration dependent circular dichroism spectra of G-quadruplex wtTel25b



8.5.19 Concentration dependent circular dichroism spectra of Q2



8.6 Synthesis of porphyrin-modified LNA



8.6.1 Rigid alkyne linked porphyrin LNA synthesis

Rigid alkyne linked porphyrin LNA phosphoramidite **14** (81 mg, 55.6 μmol, 1.0 eq.) was dissolved in DCM:MeCN (1:1, 1.5 mL, 3.7 mM) and coupled by passing the amidite solution (192 μL, 7.1 eq.) through the CGP support over a period of 5 minutes. DNA strands were cleaved from the solid support by passing concentrated ammonia solution (35 % in H_2O, S.G. = 0.88, 1 mL) through the solid support using two syringes at either end for 1 hour. The purple DNA solution was then removed from the solid support and deprotected by heating at 50 °C for 4 hours. Strands were purified by Fluoro-Pak II Columns before concentrating *in vacuo*.

8.6.2 Flexible amide linked porphyrin LNA synthesis

Flexible amide linked porphyrin LNA phosphoramidite **19** (119 mg, 82 μmol, 1.0 eq.) was dissolved in DCM:MeCN (1:1, 3.0 mL, 2.7 mM) and coupled by passing the amidite solution (192 μL, 5.3 eq.) through the CGP support over a period of 5 minutes. DNA strands were cleaved from the solid support by passing concentrated ammonia solution (35 % in H_2O, S.G. = 0.88, 1 mL) through the solid support using two syringes at either end for 1 hour. The purple DNA solution was then removed from the solid support and deprotected by heating at 50 °C for 4 hours. Strands were purified by Fluoro-Pak II Columns before concentrating *in vacuo*.

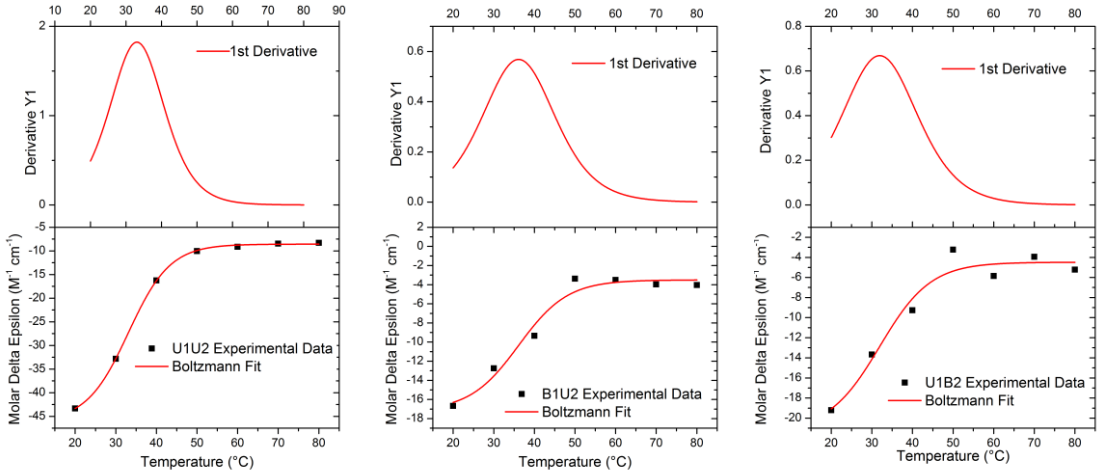
Strand	Sequence	Calculated ε_{260} (mol^{-1} dm^3 cm^{-1})	Yield (nmoles)
U1	5' – GTG ATA ACG – 3'	89900	611
U2	5' – GCA TAT CAC – 3'	88000	576
B1	5' – **G***TG A**B**A TGC – 3'	89900	562
B2	5' – **G***CA TA**B** CAC – 3'	88000	347
B3	5' – **G***CA **B**AT CAC – 3'	88000	644
B4	5' – **G***CA **B**A**B** CAC – 3'	88000	708
F1	5' – **G***TG A**F**A TGC – 3'	89900	228
F2	5' – **G***CA TA**F** CAC – 3'	88000	362
F3	5' – **G***CA **F**AT CAC – 3'	88000	371
F4	5' – **G***CA **F**A**F** CAC – 3'	88000	342

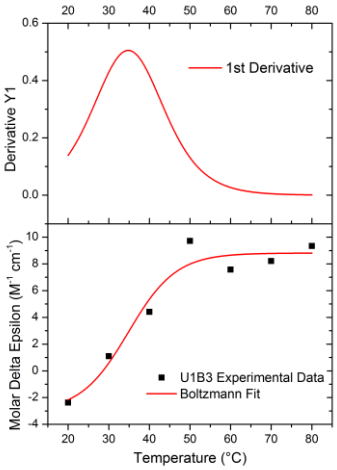
8.6.3 LNA circular dichroism spectroscopy

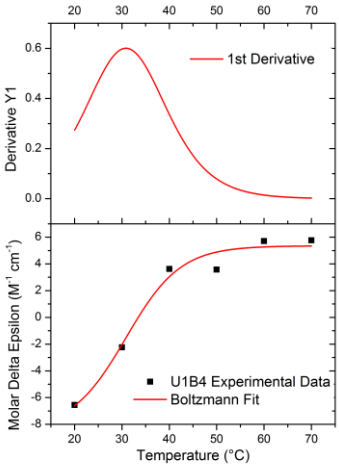
Oligonucleotide solutions (4.0 μM) were prepared in 100 mM sodium phosphate buffer (pH 7.4) supplemented with 100 mM sodium chloride and 1 mM Na_2EDTA. The samples were heated to 95 ˚C for 5 minutes and allowed to slowly cool. Spectra were recorded between 200 and 600 nm in 1 cm path length cuvettes. 1 scan was taken with a 1 nm scan step, 2 seconds per point and a bandwidth of 2 nm on a Chirascan spectrometer, or on B23 Station B, Diamond. A buffer baseline was subtracted from each spectrum. The data was recorded in mdeg and corrected to delta epsilon using the formula $\Delta\varepsilon = \theta$ / (10 x conc. x path length x 3298); conc. is in mol/litre, path length is in cm.

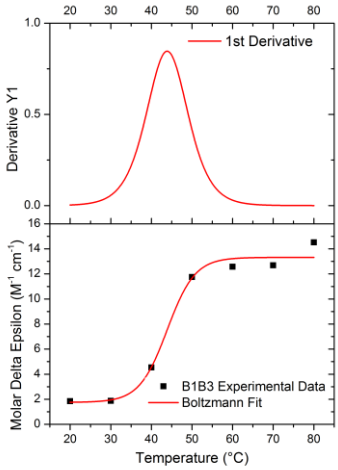
8.6.4 LNA circular dichroism melting – data collected at B23 Station B, Diamond

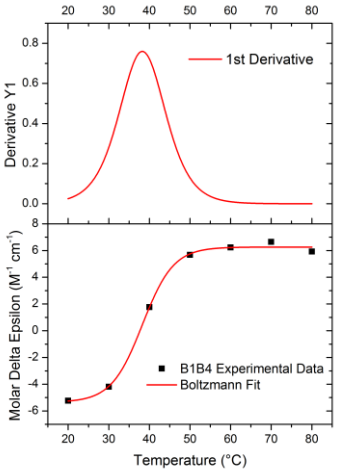
Oligonucleotide solutions (4.0 μM) were prepared in 100 mM sodium phosphate buffer (pH 7.4) supplemented with 100 mM sodium chloride and 1 mM Na_2EDTA. The samples were heated to 95 ˚C for 5 minutes and allowed to slowly cool. Spectra were recorded between 200 and 350 nm in 0.3 cm path length cuvettes. 1 scan was taken with a 1 nm scan step, 2 seconds per point and a bandwidth of 2 nm on using the beamline at B23 Station B, Diamond. Spectra were measured in 10 ˚C intervals from 20 ˚C to 80 ˚C with a 300 second equilibration between scans.





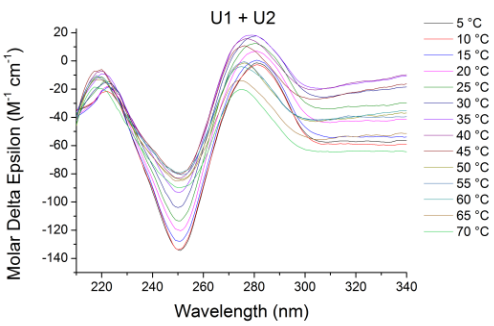


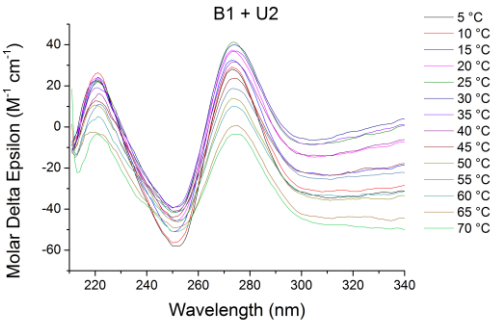


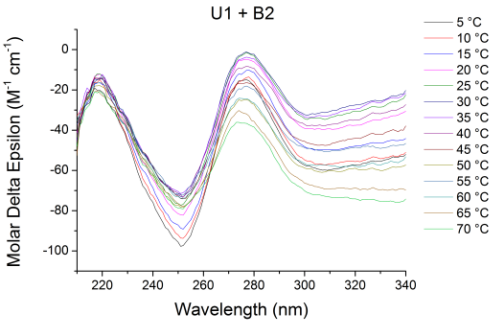


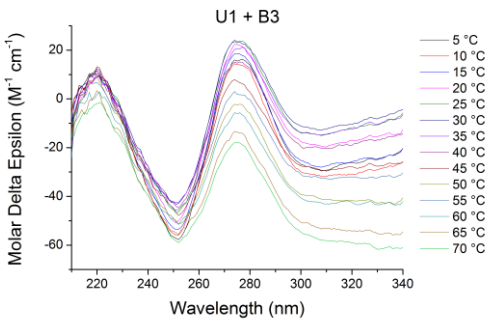
8.6.5 LNA circular dichroism melting – data collected with the Chirascan

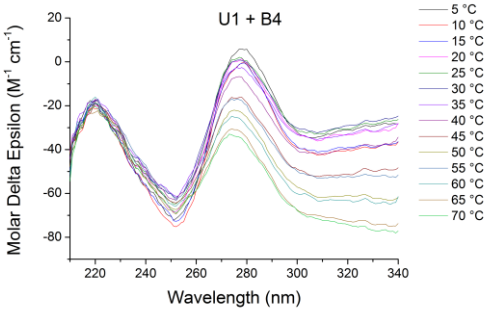
Oligonucleotide solutions (4.0 μM) were prepared in 100 mM sodium phosphate buffer (pH 7.4) supplemented with 100 mM sodium chloride and 1 mM Na_2EDTA. The samples were heated to 95 ˚C for 5 minutes and allowed to slowly cool. Spectra were recorded between 200 and 340 nm in 0.3 cm path length cuvettes. 1 scan was taken with a 1 nm scan step, 2 seconds per point and a bandwidth of 2 nm on an Applied Photophysics Chirascan spectrometer. Spectra were measured in 5 ˚C intervals from 5 ˚C to 70 ˚C with a 120 second equilibration between scans.

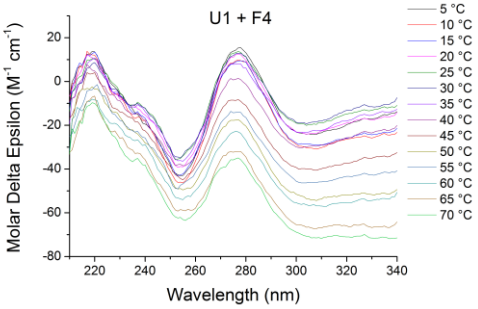

U1 + U2
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

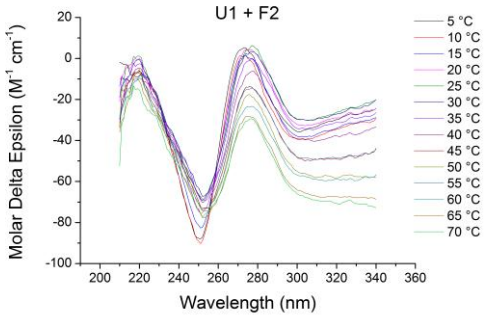

B1 + U2
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

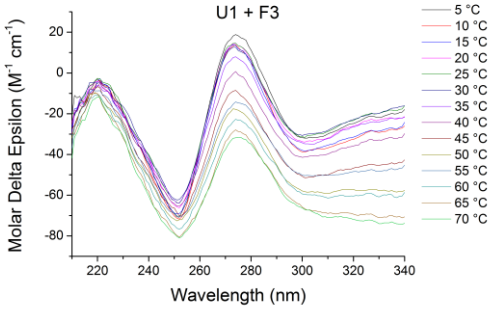

U1 + B2
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

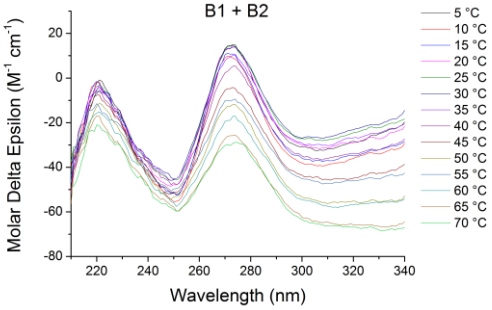

U1 + B3
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

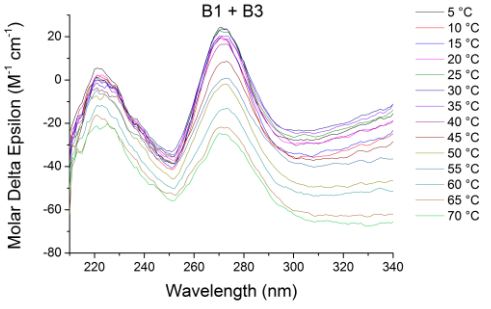

U1 + B4
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

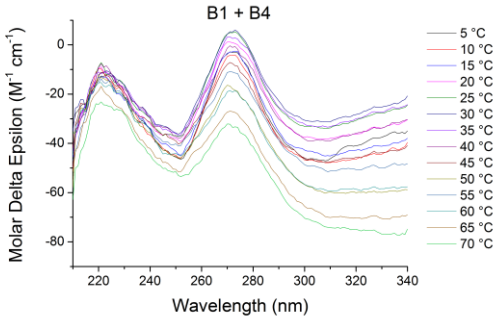

U1 + F4
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

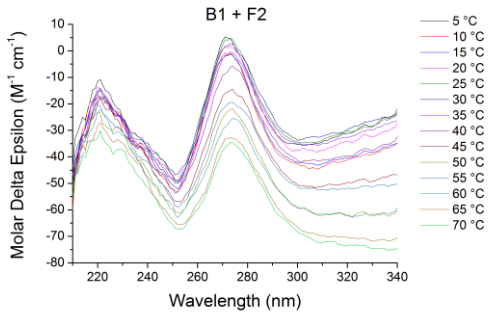

U1 + F2
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

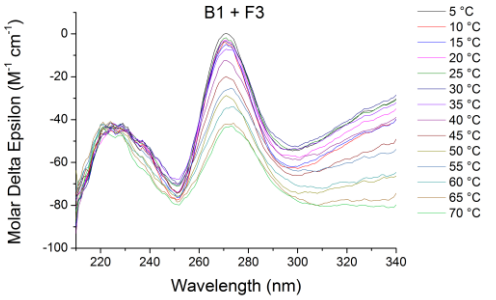

U1 + F3
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

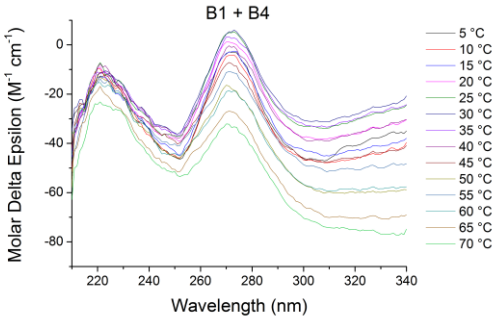

B1 + B2
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

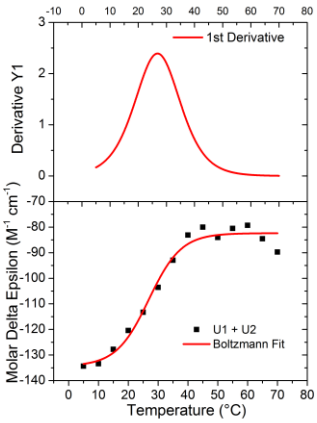

B1 + B3
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

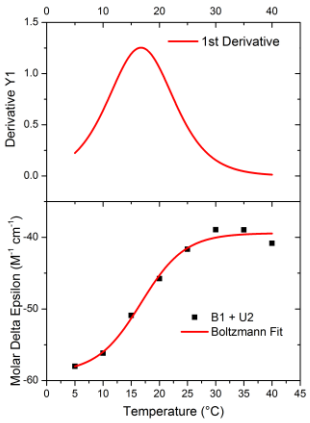

B1 + B4
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

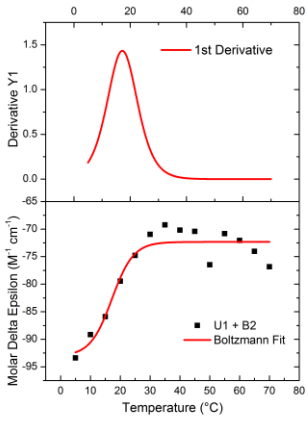

B1 + F2
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

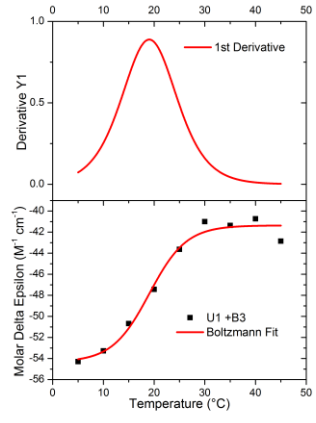

B1 + F3
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

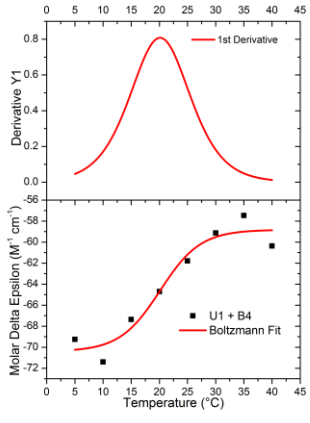

B1 + B4
Molar Delta Epsilon (M-1 cm-1)
Wavelength (nm)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C
65 °C
70 °C

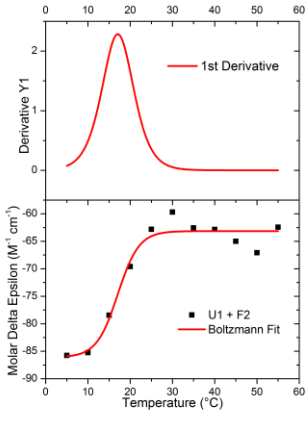

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 + U2
Boltzmann Fit
Temperature (°C)

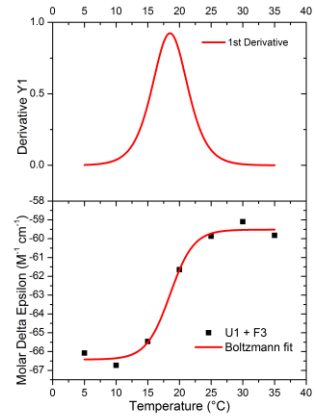

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
B1 + U2
Boltzmann Fit
Temperature (°C)

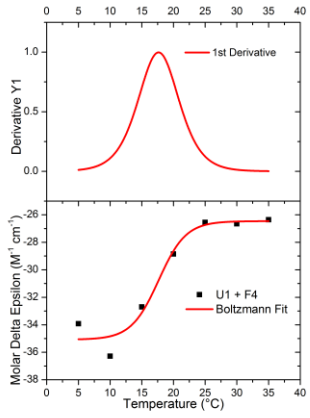

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 + B2
Boltzmann Fit
Temperature (°C)

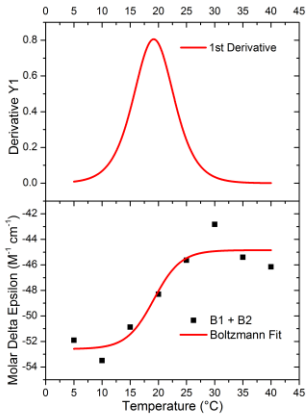

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 +B3
Boltzmann Fit
Temperature (°C)

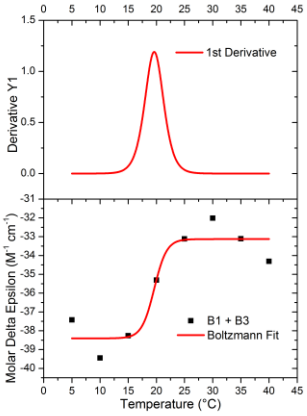

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 + B4
Boltzmann Fit
Temperature (°C)

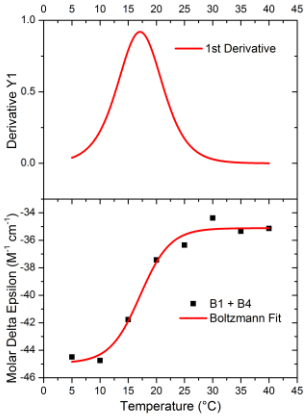

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 + F2
Boltzmann Fit
Temperature (°C)

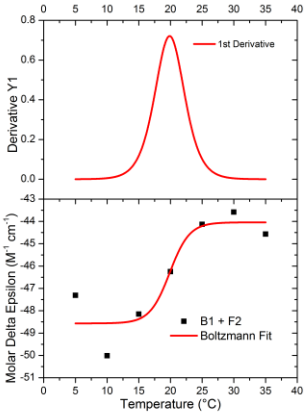

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 + F3
Boltzmann fit
Temperature (°C)

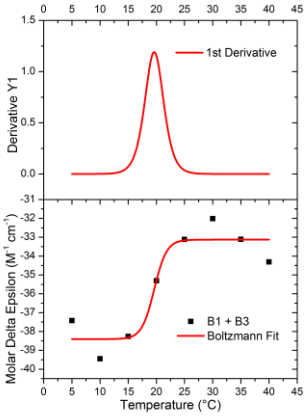

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
U1 + F4
Boltzmann Fit
Temperature (°C)

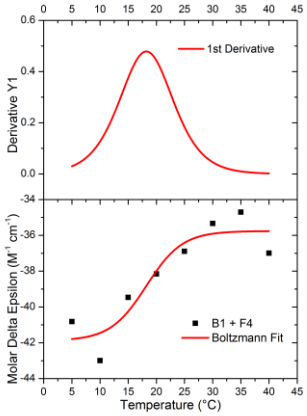

1st Derivative
Derivative Y1
Molar Delta Epsilon (M^{-1} cm^{-1})
B1 + B2
Boltzmann Fit
Temperature (°C)


1st Derivative
Derivative Y1
Molar Delta Epsilon (M-1 cm-1)
B1 + B3
Boltzmann Fit
Temperature (°C)


1st Derivative
Derivative Y1
Molar Delta Epsilon (M-1 cm-1)
B1 + B4
Boltzmann Fit
Temperature (°C)


1st Derivative
Derivative Y1
Molar Delta Epsilon (M-1 cm-1)
B1 + F2
Boltzmann Fit
Temperature (°C)


1st Derivative
Derivative Y1
Molar Delta Epsilon (M-1 cm-1)
B1 + B3
Boltzmann Fit
Temperature (°C)


1st Derivative
Derivative Y1
Molar Delta Epsilon (M-1 cm-1)
B1 + F4
Boltzmann Fit
Temperature (°C)

8.7 Primer extension methods

8.7.1 Materials for biochemistry

Synthetic oligonucleotides were purchased from VBC Genomics (Austria). Dynabeads® M-270 Streptavidin (DBStv) were obtained from Dynal A.S. (Norway), Vent (*exo*-), Pwo, Dynazyme, Phusion and Therminator were purchased from New England Biolabs (Great Britain), KOD XL from Novagen, unmodified triphosphates (dATP, dTTP, dCTP, dGTP) from Sigma Aldrich, and γ^{32}P-ATP from MP Empowered Discovery (USA).

8.7.2 Primer extension experiments – single porphyrin incorporation

The reaction mixture (20 μL) contained DNA polymerase (2.5 U μL^{-1}, 0.05 μL), polymerase buffer (2 μL), dGTP (4 mM, 0.5 μL), dTTP (2 mM, 1.0 μL, positive control), porphyrin-modified dTTP (4 mM, 1.0 μL, porphyrin incorporation), (negative control contains neither dTTP or porphyrin-modified dTTP), primer (3 μM, 1 μL) and template (3 μM, 1.5 μL) in water. The primer was labelled by use of [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated at 60 °C for 15 mins and were stopped by the addition of stop solution (40 μL, 80 % [v/v] formamide, 10 mM EDTA, 0.025 % [w/v] bromophenol blue, 0.025 % [w/v] xylene cyanol) and heated for 5 mins at 95 °C. Reaction mixtures were separated by use of 12.5 % denaturing PAGE. Visualisation was performed by phosphoimaging

8.7.3 Primer extension experiments – multiple porphyrin incorporations

The reaction mixture (20 μL) contained DNA polymerase (2.5 U μL^{-1}, 0.05 μL), polymerase buffer (2 μL), natural triphosphates dATP, dCTP and dGTP (4 mM, 0.7 μL), dTTP (2 mM, 1.4 μL, positive control), porphyrin-modified dTTP (4 mM, 1.0 μL, porphyrin incorporation), (negative control contains neither dTTP or porphyrin-modified dTTP), primer (3 μM, 1 μL) and template (3 μM, 1.5 μL). The primer was labelled by use of [γ^{32}P]-ATP according to standard techniques. Reaction mixtures were incubated at 60 °C for 15 mins and were stopped by the addition of stop solution (40 μL, 80 % [v/v] formamide, 10 mM EDTA, 0.025 % [w/v] bromophenol blue, 0.025 % [w/v] xylene cyanol) and heated for 5 mins at 95 °C. Reaction mixtures were separated by use of 12.5 % denaturing PAGE. Visualisation was performed by phosphoimaging.

8.7.4 Primer extension experiment – single porphyrin incorporation on a larger scale

The reaction mixture (325 μL) contained KOD XL (2.5 U μL^{-1}, 4.0 μL), polymerase buffer (30 μL), dGTP (4 mM, 21 μL), porphyrin-modified dTTP (4 mM, 21 μL) and primer (100 μM, 6 μL) and template (100 μM, 6 μL) in water. Reaction mixtures were incubated at 60 °C for 45 mins where precipitation of the porphyrin-modified dTTP was observed. Isolation of the porphyrin-modified oligonucleotide was carried out by the DBStv magnetoseparation procedure.

8.7.5 Isolation of single strand oligonucleotides by the DBStv magnetoseparation procedure

Reaction mixture (325 μL) was supplemented with 2.5 M NaCl (26 μL) and was added to DBStv (45 μL of the stock solution washed three times with 150 μL of buffer (0.3 M NaCl, 10 mM TRIS, pH = 7.4)). The suspension was shaken at room temperature for 30 °C to allow the oligonucleotides to bind to the DBStv beads. The DBStv beads were washed three times with 400 μL of PBS solution (0.14 M NaCl, 3 mM KCl, 4 mM sodium phosphate, pH = 7.4) with 0.01 % Tween 20 and then three times with 400 μL of buffer (0.3 M NaCl, 10 mM TRIS, pH = 7.4) and finally three times with 400 μL PCR grade water. Porphyrin-modified oligonucleotides were released by shaking and heating the sample in PCR grade water (150 μL) at 75 °C for 3 mins. Each medium exchange was performed using a magnetoseparator (Dynal, Norway).

9 Bibliography

1. J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737-738.
2. R. E. Franklin and R. G. Gosling, *Nature*, 1953, **171**, 740-741.
3. M. H. F. Wilkins, A. R. Stokes and H. R. Wilson, *Nature*, 1953, **171**, 738-740.
4. R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Nature*, 1980, **287**, 755-758.
5. E. Chargaff, *Experientia*, 1950, **6**, 201-209.
6. I. R. Gould and P. A. Kollman, *J. Am. Chem. Soc.*, 1994, **116**, 2493-2499.
7. P. Jurečka and P. Hobza, *J. Am. Chem. Soc.*, 2003, **125**, 15608-15613.
8. M. H. Wilkins, W. E. Seeds, A. R. Stokes and H. R. Wilson, *Nature*, 1953, **172**, 759-762.
9. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich, *Nature*, 1979, **282**, 680-686.
10. A. Rich and S. Zhang, *Nat. Rev. Genet.*, 2003, **4**, 566-572.
11. Y.-G. Gao, H. Robinson and A. H. J. Wang, *Eur. J. Biochem.*, 1999, **261**, 413-420.
12. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 2179-2183.
13. D. E. Gilbert and J. Feigon, *Curr. Opin. Struct. Biol.*, 1999, **9**, 305-314.
14. G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc.*, 1957, **79**, 2023-2024.
15. K. R. Fox and T. Brown, *Biochem. Soc. Trans.*, 2011, **39**, 629-634.
16. Z. Ma and J.-S. Taylor, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11159-11163.
17. C. Zhao, K. Qu, C. Xu, J. Ren and X. Qu, *Nucleic Acids Res.*, 2011, **39**, 3939-3948.
18. S. Buchini and C. J. Leumann, *Curr. Opin. Chem. Biol.*, 2003, **7**, 717-726.
19. K. Gehring, J.-L. Leroy and M. Gueron, *Nature*, 1993, **363**, 561-565.
20. C. H. Kang, I. Berger, C. Lockshin, R. Ratliff, R. Moyzis and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 11636-11640.
21. M. Guéron and J.-L. Leroy, *Curr. Opin. Struct. Biol.*, 2000, **10**, 326-331.
22. N. R. Kallenbach, R. I. Ma and N. C. Seeman, *Nature*, 1983, **305**, 829-831.
23. J. H. Chen and N. C. Seeman, *Nature*, 1991, **350**, 631-633.
24. E. Winfree, F. R. Liu, L. A. Wenzler and N. C. Seeman, *Nature*, 1998, **394**, 539-544.
25. P. W. K. Rothemund, *Nature*, 2006, **440**, 297-302.
26. E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf and J. Kjems, *Nature*, 2009, **459**, 73-76.
27. Y. Sannohe, M. Endo, Y. Katsuda, K. Hidaka and H. Sugiyama, *J. Am. Chem. Soc.*, 2010, **132**, 16311-16313.
28. B. A. Connolly, *Nucleic Acids Res.*, 1987, **15**, 3131-3139.
29. B. A. Connolly and P. Rider, *Nucleic Acids Res.*, 1985, **13**, 4485-4502.
30. J. G. Zendegui, K. M. Vasquez, J. H. Tinsley, D. J. Kessler and M. E. Hogan, *Nucleic Acids Res.*, 1992, **20**, 307-314.
31. C. R. Petrie, M. W. Reed, A. D. Adams and R. B. Meyer, *Bioconjugate Chem.*, 1992, **3**, 85-87.
32. M. Antopolsky and A. Azhayev, *Nucleos. Nucleot. Nucl.*, 2001, **20**, 539-550.
33. M. Nakamura, Y. Shimomura, Y. Ohtoshi, K. Sasa, H. Hayashi, H. Nakano and K. Yamana, *Org. Biomol. Chem.*, 2007, **5**, 1945-1951.
34. L. Zhu, P. S. Lukeman, J. W. Canary and N. C. Seeman, *J. Am. Chem. Soc.*, 2003, **125**, 10178-10179.
35. E. S. Krider, J. J. Rack, N. L. Frank and T. J. Meade, *Inorg. Chem.*, 2001, **40**, 4002-4009.
36. M. A. Campbell and J. Wengel, *Chem. Soc. Rev.*, 2011, **40**, 5680-5689.
37. C. Brotschi, A. Haberli and C. J. Leumann, *Angew. Chem. Int. Ed.*, 2001, **40**, 3012-3014.

38. M. M. Somoza, D. Andreatta, C. J. Murphy, R. S. Coleman and M. A. Berg, *Nucleic Acids Res.*, 2004, **32**, 2494-2507.
39. S. Hainke and O. Seitz, *Angew. Chem. Int. Ed.*, 2009, **48**, 8250-8253.
40. J. Gao, H. Liu and E. T. Kool, *Angew. Chem. Int. Ed.*, 2005, **44**, 3118-3122.
41. D. A. Malyshev, D. A. Pfaff, S. I. Ippoliti, G. T. Hwang, T. J. Dwyer and F. E. Romesberg, *Chem. Eur. J.*, 2010, **16**, 12650-12659.
42. J. Chelliserrykattil, H. Lu, A. H. F. Lee and E. T. Kool, *ChemBioChem*, 2008, **9**, 2976-2980.
43. Y. J. Seo and F. E. Romesberg, *ChemBioChem*, 2009, **10**, 2394-2400.
44. N. Zimmermann, E. Meggers and P. G. Schultz, *Bioorg. Chem.*, 2004, **32**, 13-25.
45. K. Tanaka, G. H. Clever, Y. Takezawa, Y. Yamada, C. Kaul, M. Shionoya and T. Carell, *Nat. Nanotechnol.*, 2006, **1**, 190-U195.
46. H. Yang, A. Z. Rys, C. K. McLaughlin and H. F. Sleiman, *Angew. Chem. Int. Ed.*, 2009, **48**, 9919-9923.
47. V. L. Malinovskii, F. Samain and R. Haener, *Angew. Chem. Int. Ed.*, 2007, **46**, 4464-4467.
48. J. L. Czlapinski and T. L. Sheppard, *J. Am. Chem. Soc.*, 2001, **123**, 8618-8619.
49. L. A. Fendt, I. Bouamaied, S. Thoni, N. Amiot and E. Stulz, *J. Am. Chem. Soc.*, 2007, **129**, 15319-15329.
50. J. Barbaric and H. A. Wagenknecht, *Org. Biomol. Chem.*, 2006, **4**, 2088-2090.
51. E. Mayer-Enthart and H. A. Wagenknecht, *Angew. Chem. Int. Ed.*, 2006, **45**, 3372-3375.
52. M. Vrabel, P. Horakova, H. Pivonkova, L. Kalachova, H. Cernocka, H. Cahova, R. Pohl, P. Sebest, L. Havran, M. Hocek and M. Fojta, *Chem. Eur. J.*, 2009, **15**, 1144-1154.
53. S. I. Khan, A. E. Beilstein and M. W. Grinstaff, *Inorg. Chem.*, 1999, **38**, 418-+.
54. T. Rühl and E. Stulz, *Supramol. Chem.*, 2010, **22**, 103-108.
55. J. R. Burns, J. Zekonyte, G. Siligardi, R. Hussain and E. Stulz, *Molecules*, 2011, **16**, 4912-4922.
56. T. Kottysch, C. Ahlborn, F. Brotzel and C. Richert, *Chem. Eur. J.*, 2004, **10**, 4017-4028.
57. F. W. Hobbs, *J. Org. Chem.*, 1989, **54**, 3420-3422.
58. M. Hocek, *Eur. J. Org. Chem.*, 2003, 245-254.
59. E. B. Fleischer, *Acc. Chem. Res.*, 1970, **3**, 105-112.
60. D. Ostfeld and M. Tsutsui, *Acc. Chem. Res.*, 1974, **7**, 52-58.
61. A. Prodi, M. T. Indelli, C. J. Kleverlaan, F. Scandola, E. Alessio, T. Gianferrara and L. G. Marzilli, *Chem. Eur. J.*, 1999, **5**, 2668-2679.
62. D. Mansuy, J. P. Battioni, D. Dupre, E. Sartori and G. Chottard, *J. Am. Chem. Soc.*, 1982, **104**, 6159-6161.
63. S. Ryu, J. Kim, H. Yeo and K. Kim, *Inorg. Chim. Acta.*, 1995, **228**, 233-236.
64. K. Kalyanasundaram, *Photochemistry of polypyridine and porphyrin complexes*, Academic Press, 1992.
65. T. Balasubramanian and J. S. Lindsey, 1999, **55**, 6771-6784.
66. J. S. Lindsey, I. C. Schreiman, H. C. Hsu, P. C. Kearney and A. M. Marguerettaz, *J. Org. Chem.*, 1987, **52**, 827-836.
67. X. Huang, K. Nakanishi and N. Berova, *Chirality*, 2000, **12**, 237-255.
68. M. Gouterman, *J. Chem. Phys.*, 1959, **30**, 1139-1161.
69. M. Gouterman, G. H. Wagnière and L. C. Snyder, *J. Mol. Spectrosc.*, 1963, **11**, 108-127.
70. P. J. Spellane, M. Gouterman, A. Antipas, S. Kim and Y. C. Liu, *Inorg. Chem.*, 1980, **19**, 386-391.
71. S. W. Ryter and R. M. Tyrrell, *Free Radical Biol. Med.*, 2000, **28**, 289-309.
72. E. Rabinowitch, *Rev. Mod. Phys.*, 1944, **16**, 226-235.
73. P. Fromme and P. Mathis, *Photosynth. Res.*, 2004, **80**, 109-124.
74. K.-H. Rhee, *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 307-328.
75. B. B. Buchanan, *Annu. Rev. Plant Physiol.*, 1980, **31**, 341-374.
76. I. Cohen-Ofri, M. van Gastel, J. Grzyb, A. Brandis, I. Pinkas, W. Lubitz and D. Noy, *J. Am. Chem. Soc.*, 2011, **133**, 9526-9535.

77. V. L. Gunderson, A. L. Smeigh, C. H. Kim, D. T. Co and M. R. Wasielewski, *J. Am. Chem. Soc.*, 2012, **134**, 4363-4372.
78. J. K. Sprafke, S. D. Stranks, J. H. Warner, R. J. Nicholas and H. L. Anderson, *Angew. Chem. Int. Ed.*, 2011, **50**, 2313-2316.
79. D. Conklin, S. Nanayakkara, T.-H. Park, M. F. Lagadec, J. T. Stecher, M. J. Therien and D. A. Bonnell, *Nano Lett.*, 2012, **12**, 2414-2419.
80. M. B. Winter, E. J. McLaurin, S. Y. Reece, C. Olea, D. G. Nocera and M. A. Marletta, *J. Am. Chem. Soc.*, 2010, **132**, 5582-5583.
81. J. Králová, Z. k. Kejík, T. s. Bříza, P. Poučková, A. Král, P. Martásek and V. r. Král, *J. Med. Chem.*, 2009, **53**, 128-138.
82. A. Brewer, G. Siligardi, C. Neylon and E. Stulz, *Org. Biomol. Chem.*, 2011, **9**, 777-782
83. J. R. Dunetz, C. Sandstrom, E. R. Young, P. Baker, S. A. Van Name, T. Cathopolous, R. Fairman, J. C. de Paula and K. S. Åkerfeldt, *Org. Lett.*, 2005, **7**, 2559-2561.
84. N. Solladié, N. Aubert, J. P. Gisselbrecht, M. Gross, C. Sooambar and V. Troiani, *Chirality*, 2003, **15**, S50-S56.
85. M. Fujitsuka, A. Okada, S. Tojo, F. Takei, K. Onitsuka, S. Takahashi and T. Majima, *J. Phys. Chem. B*, 2004, **108**, 11935-11941.
86. M. Jahan, Q. Bao and K. P. Loh, *J. Am. Chem. Soc.*, 2012, **134**, 6707-6713.
87. V. L. Malinovskii, A. L. Nussbaumer and R. Haener, *Angew. Chem. Int. Ed.*, 2012, **51**, 4905-4908.
88. J. H. Kim, M. Lee, J. S. Lee and C. B. Park, *Angew. Chem. Int. Ed.*, 2012, **51**, 517-520.
89. Y. S. Nam, T. Shin, H. Park, A. P. Magyar, K. Choi, G. Fantner, K. A. Nelson and A. M. Belcher, *J. Am. Chem. Soc.*, 2010, **132**, 1462-1463.
90. N. Solladie´ and M. Gross, *Tetrahedron Lett.*, 1999, **40**, 3359-3362.
91. A. Berman, E. S. Izraeli, H. Levanon, B. Wang and J. L. Sessler, *J. Am. Chem. Soc.*, 1995, **117**, 8252-8257.
92. A. S. Boutorine, D. Trung Le, J. P. Battioni, D. Mansuy, D. Dupre and C. Helene, *Bioconjugate Chem.*, 1990, **1**, 350-356.
93. I. Dubey, G. Pratviel and B. Meunier, *CR Acad Sci II C*, 1998, **1**, 259-267.
94. I. Bouamaied and E. Stulz, *Synlett*, 2004, 1579-1583.
95. H. Morales-Rojas and E. T. Kool, *Org. Lett.*, 2002, **4**, 4377-4380.
96. M. Balaz, A. E. Holmes, M. Benedetti, P. C. Rodriguez, N. Berova, K. Nakanishi and G. Proni, *J. Am. Chem. Soc.*, 2005, **127**, 4172-4173.
97. M. Balaz, J. D. Steinkruger, G. A. Ellestad and N. Berova, *Org. Lett.*, 2005, **7**, 5613-5616.
98. M. Balaz, B. C. Li, J. D. Steinkruger, G. A. Ellestad, K. Nakanishi and N. Berova, *Org. Biomol. Chem.*, 2006, **4**, 1865-1867.
99. A. D'Urso, A. Mammana, M. Balaz, A. E. Holmes, N. Berova, R. Lauceri and R. Purrello, *J. Am. Chem. Soc.*, 2009, **131**, 2046-2047.
100. M. Balaz, A. E. Holmes, M. Benedetti, G. Proni and N. Berova, *Bioorg. Med. Chem.*, 2005, **13**, 2413-2421.
101. A. Mammmana, G. Pescitelli, T. Asakawa, S. Jockusch, A. G. Petrovic, R. R. Monaco, R. Purrello, N. J. Turro, K. Nakanishi, G. A. Ellestad, M. Balaz and N. Berova, *Chem.-Eur. J.*, 2009, **15**, 11853-11866.
102. I. Bouamaied and E. Stulz, *Chimia*, 2005, **59**, 101-104.
103. T. Nguyen, A. Brewer and E. Stulz, *Angew. Chem. Int. Ed.*, 2009, **48**, 1974-1977.
104. H. L. Anderson, *Inorg. Chem.*, 1994, **33**, 972-981.
105. C. A. Hunter, J. K. M. Sanders and A. J. Stone, *Chem. Phys.*, 1989, **133**, 395-404.
106. Y. Ohya, N. Hashimoto, S. Jo, T. Nohori, T. Yoshikuni, T. Ouchi and H. Tamiaki, *Supramol. Chem.*, 2009, **21**, 301-309.
107. K. Borjesson, J. Tumpane, T. Ljungdahl, L. M. Wilhelmsson, B. Norden, T. Brown, J. Martensson and B. Albinsson, *J. Am. Chem. Soc.*, 2009, **131**, 2831-2839.
108. H. Fischer and K. Zeile, *Justus Liebigs Ann. Chem.*, 1929, **468**, 98-116.
109. P. Rothemund, *J. Am. Chem. Soc.*, 1935, **57**, 2010-2011.
110. P. Rothemund, *J. Am. Chem. Soc.*, 1936, **58**, 625-627.

111. P. Rothemund, *J. Am. Chem. Soc.*, 1939, **61**, 2912-2915.
112. A. D. Adler, F. R. Longo, J. D. Finarelli, J. Goldmacher, J. Assour and L. Korsakoff, *J. Org. Chem.*, 1967, **32**, 476-476.
113. L. Clima and W. Bannwarth, *Helv. Chim. Acta.*, 2008, **91**, 165-U165.
114. J. Alsina, G. Barany, F. Albericio and S. A. Kates, *Lett. Pept. Sci.*, 1999, **6**, 243-245.
115. A. Mammana, T. Asakawa, K. Bitsch-Jensen, A. Wolfe, S. Chaturantabut, Y. Otani, X. X. Li, Z. M. Li, K. Nakanishi, M. Balaz, G. A. Ellestad and N. Berova, *Bioorg. Med. Chem.*, 2008, **16**, 6544-6551.
116. A. Karotki, M. Kruk, M. Drobizhev, A. Rebane, E. Nickel and C. W. Spangler, *IEEE J. Sel. Top. Quantum Electron.*, 2001, **7**, 971-975.
117. W. H. Pearson, D. A. Berry, P. Stoy, K. Y. Jung and A. D. Sercel, *J. Org. Chem.*, 2005, **70**, 7114-7122.
118. I. Bouamaied, T. Nguyen, T. Rühl and E. Stulz, *Org. Biomol. Chem.*, 2008, **6**, 3888-3891.
119. T. Nguyen, A. Brewer and E. Stulz, *Angew. Chem. Int. Ed.*, 2009, **48**, 1974-1977.
120. G. D. Dorough, J. R. Miller and F. M. Huennekens, *J. Am. Chem. Soc.*, 1951, **73**, 4315-4320.
121. K. Börjesson, J. Wiberg, A. H. El-Sagheer, T. Ljungdahl, J. Mårtensson, T. Brown, B. Nordén and B. Albinsson, *ACS Nano*, 2010, **4**, 5037-5046.
122. E. E. Merkina and K. R. Fox, *Biophys. J.*, 2005, **89**, 365-373.
123. G. N. Parkinson, R. Ghosh and S. Neidle, *Biochemistry*, 2007, **46**, 2390-2397.
124. M. Gellert, M. N. Lipsett and D. R. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 1962, **48**, 2013-2018.
125. S. Arnott, Chandras.R and C. M. Marttila, *Biochem. J.*, 1974, **141**, 537-&.
126. K. Phillips, Z. Dauter, A. I. H. Murchie, D. M. J. Lilley and B. Luisi, *J. Mol. Biol.*, 1997, **273**, 171-182.
127. A. Wong and G. Wu, *J. Am. Chem. Soc.*, 2003, **125**, 13895-13905.
128. N. a. Špačková, I. Berger and J. Šponer, *J. Am. Chem. Soc.*, 1999, **121**, 5519-5534.
129. G. N. Parkison, in *Quadruplex Nucleic Acids*, eds. S. Neidle and S. Balasubramanian, RSC Publishing, 2006, pp. 1-30.
130. F. Aboul-ela, A. I. H. Murchie and D. M. J. Lilley, *Nature*, 1992, **360**, 280-282.
131. Y. Wang and D. J. Patel, *J. Mol. Biol.*, 1993, **234**, 1171-1183.
132. P. A. Rachwal, T. Brown and K. R. Fox, *Biochemistry*, 2007, **46**, 3036-3044.
133. G. D. Balkwill, T. P. Garner, H. E. L. Williams and M. S. Searle, *J. Mol. Biol.*, 2009, **385**, 1600-1615.
134. P. Hazel, G. N. Parkinson and S. Neidle, *J. Am. Chem. Soc.*, 2006, **128**, 5480-5487.
135. P. A. Rachwal, I. S. Findlow, J. M. Werner, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2007, **35**, 4214-4222.
136. A. Guédin, J. Gros, P. Alberti and J.-L. Mergny, *Nucleic Acids Res.*, 2010, **38**, 7858-7868.
137. P. Hazel, J. Huppert, S. Balasubramanian and S. Neidle, *J. Am. Chem. Soc.*, 2004, **126**, 16405-16415.
138. M. Webba da Silva, *Methods*, 2007, **43**, 264-277.
139. P. A. Rachwal and K. R. Fox, *Methods*, 2007, **43**, 291-301.
140. J. L. Mergny, A. T. Phan and L. Lacroix, *FEBS Lett.*, 1998, **435**, 74-78.
141. M. Webba da Silva, *Chem. Eur. J.*, 2007, **13**, 9738-9745.
142. *The 2009 Nobel Prize in Physiology or Medicine - Advanced information*, http://www.nobelprize.org/nobel_prizes/medicine/laureates/2009/advanced.html, Accessed 11 Sept 2012.
143. T. L. Orr-Weaver, J. W. Szostak and R. J. Rothstein, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 6354-6358.
144. J. W. Szostak and E. H. Blackburn, *Cell*, 1982, **29**, 245-255.
145. J. Shampay, J. W. Szostak and E. H. Blackburn, *Nature*, 1984, **310**, 154-157.
146. C. W. Greider and E. H. Blackburn, *Cell*, 1985, **43**, 405-413.
147. C. W. Greider and E. H. Blackburn, *Cell*, 1987, **51**, 887-898.

148. C. W. Greider and E. H. Blackburn, *Nature*, 1989, **337**, 331-337.
149. J. Lingner, T. R. Hughes, A. Shevchenko, M. Mann, V. Lundblad and T. R. Cech, *Science*, 1997, **276**, 561-567.
150. N. Kim, M. Piatyszek, K. Prowse, C. Harley, M. West, P. Ho, G. Coviello, W. Wright, S. Weinrich and J. Shay, *Science*, 1994, **266**, 2011-2015.
151. N. Maizels, *Nat. Struct. Mol. Biol.*, 2006, **13**, 1055-1059.
152. A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718-720.
153. S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402-5415.
154. Y. Wang and D. J. Patel, *Structure*, 1993, **1**, 263-282.
155. G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, 2002, **417**, 876-880.
156. K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix and D. J. Patel, *J. Am. Chem. Soc.*, 2006, **128**, 9963-9970.
157. A. Ambrus, D. Chen, J. X. Dai, T. Bialis, R. A. Jones and D. Z. Yang, *Nucleic Acids Res.*, 2006, **34**, 2723-2735.
158. A. T. Phan, K. N. Luu and D. J. Patel, *Nucleic Acids Res.*, 2006, **34**, 5715-5719.
159. J. X. Dai, M. Carver, C. Punchihewa, R. A. Jones and D. Z. Yang, *Nucleic Acids Res.*, 2007, **35**, 4927-4940.
160. M. Read, R. J. Harrison, B. Romagnoli, F. A. Tanious, S. H. Gowan, A. P. Reszka, W. D. Wilson, L. R. Kelland and S. Neidle, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 4844-4849.
161. J. E. Redman, J. M. Granadino-Roldan, J. A. Schouten, S. Ladame, A. P. Reszka, S. Neidle and S. Balasubramanian, *Org. Biomol. Chem.*, 2009, **7**, 76-84.
162. N. H. Campbell, G. N. Parkinson, A. P. Reszka and S. Neidle, *J. Am. Chem. Soc.*, 2008, **130**, 6722-6724.
163. S. M. Gowan, R. Heald, M. F. G. Stevens and L. R. Kelland, *Mol. Pharmacol.*, 2001, **60**, 981-988.
164. K. Shin-ya, K. Wierzba, K.-i. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa and H. Seto, *J. Am. Chem. Soc.*, 2001, **123**, 1262-1263.
165. C. M. Barbieri, A. R. Srinivasan, S. G. Rzuczek, J. E. Rice, E. J. LaVoie and D. S. Pilch, *Nucleic Acids Res.*, 2007, **35**, 3272-3286.
166. Z. A. E. Waller, P. S. Shirude, R. Rodriguez and S. Balasubramanian, *Chem. Commun.*, 2008, 1467-1469.
167. S. Muller, G. D. Pantos, R. Rodriguez and S. Balasubramanian, *Chem. Commun.*, 2009, 80-82.
168. R. T. Wheelhouse, D. Sun, H. Han, F. X. G. Han and L. H. Hurley, *J. Am. Chem. Soc.*, 1998, **120**, 3261-3262.
169. N. V. Anantha, M. Azam and R. D. Sheardy, *Biochemistry*, 1998, **37**, 2709-2714.
170. I. Haq, J. O. Trent, B. Z. Chowdhry and T. C. Jenkins, *J. Am. Chem. Soc.*, 1999, **121**, 1768-1779.
171. F. X. G. Han, R. T. Wheelhouse and L. H. Hurley, *J. Am. Chem. Soc.*, 1999, **121**, 3561-3570.
172. A. Pasternak, F. J. Hernandez, L. M. Rasmussen, B. Vester and J. Wengel, *Nucleic Acids Res.*, 2011, **39**, 1155-1164.
173. A. Pasternak and J. Wengel, *Org. Biomol. Chem.*, 2011, **9**, 3591-3597.
174. S. Cogoi, M. Paramasivam, V. Filichev, I. Gel□ci, E. B. Pedersen and L. E. Xodo, *J. Med. Chem.*, 2008, **52**, 564-568.
175. E. B. Pedersen, J. T. Nielsen, C. Nielsen and V. V. Filichev, 2011, **39**, 2470-2481.
176. B. Sacca, L. Lacroix and J. L. Mergny, *Nucleic Acids Res.*, 2005, **33**, 1182-1192.
177. L. Martino, B. Pagano, I. Fotticchia, S. Neidle and C. Giancola, *J. Phys. Chem. B*, 2009, **113**, 14779-14786.
178. J. Dai, M. Carver and D. Yang, *Biochimie*, 2008, **90**, 1172-1183.
179. *Applied Biosystems Oligo Calculator*, http://www.ambion.com/techlib/misc/oligo_calculator.html, Accessed 2009.
180. J.-L. Mergny, A.-T. Phan and L. Lacroix, *FEBS Lett.*, 1998, **435**, 74-78.
181. N. K. Sharma and K. N. Ganesh, *Org. Biomol. Chem.*, 2011, **9**, 725-729.
182. S. Paramasivan, I. Rujan and P. H. Bolton, *Methods*, 2007, **43**, 324-331.

183. A. I. Karsisiotis, N. M. a. Hessari, E. Novellino, G. P. Spada, A. Randazzo and M. Webba da Silva, *Angew. Chem. Int. Ed.*, 2011, **50**, 10645-10648.
184. M. C. Miller, H. T. Le, W. L. Dean, P. A. Holt, J. B. Chaires and J. O. Trent, *Org. Biomol. Chem.*, 2011, **9**, 7633-7637.
185. D. Renčiuk, I. Kejnovská, P. Školáková, K. Bednářová, J. Motlová and M. Vorlíčková, *Nucleic Acids Res.*, 2009, **37**, 6625-6634.
186. I. Manet, F. Manoli, M. P. Donzello, E. Viola, G. Andreano, A. Masi, L. Cellai and S. Monti, *Org. Biomol. Chem.*, 2011, **9**, 684-688.
187. D. A. Case, T. A. Darden, I. T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Z. K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. P. J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, **2010**; AMBER 11, University of California, San Francisco.
188. A. T. Phan, V. Kuryavyi, H. Y. Gaw and D. J. Patel, *Nat. Chem. Biol.*, 2005, **1**, 167-173.
189. P. Alberti, A. Bourdoncle, B. Sacca, L. Lacroix and J. L. Mergny, *Org. Biomol. Chem.*, 2006, **4**, 3383-3391.
190. W. U. Dittmer, A. Reuter and F. C. Simmel, *Angew. Chem. Int. Ed.*, 2004, **43**, 3550-3553.
191. F. Pu, C. Y. Wang, D. Hu, Z. Z. Huang, J. S. Ren, S. Wang and X. G. Qu, *Mol. Biosyst.*, 2010, **6**, 1928-1932.
192. Y. Xu, Y. Hirao, Y. Nishimura and H. Sugiyama, *Bioorg. Med. Chem.*, 2007, **15**, 1275-1279.
193. P. Alberti and J. L. Mergny, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1569-1573.
194. A. Risitano and K. R. Fox, *Biochemistry*, 2003, **42**, 6507-6513.
195. S. Obika, D. Nanbu, Y. Hari, K.-i. Morio, Y. In, T. Ishida and T. Imanishi, *Tetrahedron Lett.*, 1997, **38**, 8735-8738.
196. A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen and J. Wengel, *Tetrahedron*, 1998, **54**, 3607-3630.
197. A. A. Koshkin, J. Fensholdt, H. M. Pfundheller and C. Lomholt, *J. Org. Chem.*, 2001, **66**, 8504-8512.
198. S. Obika, D. Nanbu, Y. Hari, J. Andoh, K. Morio, T. Doi and T. Imanishi, *Tetrahedron Lett.*, 1998, **39**, 5401-5404.
199. A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh and J. Wengel, *J. Am. Chem. Soc.*, 1998, **120**, 13252-13253.
200. M. Petersen, C. B. Nielsen, K. E. Nielsen, G. A. Jensen, K. Bondensgaard, S. K. Singh, V. K. Rajwanshi, A. A. Koshkin, B. M. Dahl, J. Wengel and J. P. Jacobsen, *J. Mol. Recognit.*, 2000, **13**, 44-53.
201. S. O. Konorov, H. Georg Schulze, C. J. Addison, C. A. Haynes, M. W. Blades and R. F. B. Turner, *J. Raman Spectrosc.*, 2009, **40**, 1162-1171.
202. H. Kaur, B. R. Babu and S. Maiti, *Chem. Rev.*, 2007, **107**, 4672-4697.
203. C. Wahlestedt, P. Salmi, L. Good, J. Kela, T. Johnsson, T. Hökfelt, C. Broberger, F. Porreca, J. Lai, K. Ren, M. Ossipov, A. Koshkin, N. Jakobsen, J. Skouv, H. Oerum, M. H. Jacobsen and J. Wengel, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 5633-5638.
204. J. Kurreck, E. Wyszko, C. Gillen and V. A. Erdmann, *Nucleic Acids Res.*, 2002, **30**, 1911-1918.
205. M. E. Østergaard, P. Kumar, B. Baral, D. J. Raible, T. Santhosh Kumar, B. A. Anderson, D. C. Guenther, L. Deobald, A. J. Paszczynski, P. K. Sharma and P. J. Hrdlicka, *ChemBioChem*, 2009, **10**, 2740-2743.
206. M. E. Østergaard, P. Kumar, B. Baral, D. C. Guenther, B. A. Anderson, F. M. Ytreberg, L. Deobald, A. J. Paszczynski, P. K. Sharma and P. J. Hrdlicka, *Chem. Eur. J.*, 2011, **17**, 3157-3165.
207. M. E. Ostergaard and P. J. Hrdlicka, *Chem. Soc. Rev.*, 2011, **40**, 5771-5788.

208. J. W. Dale and M. v. Schantz, *From Genes to Genomes: Concepts and Applications of DNA Technology*., John Wiley & Sons Ltd., 2002.
209. P. R. Langer, A. A. Waldrop and D. C. Ward, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 6633-6637.
210. K. Sakthivel and C. F. Barbas Iii, *Angew. Chem. Int. Ed.*, 1998, **37**, 2872-2875.
211. O. Thum, S. Jager and M. Famulok, *Angew. Chem. Int. Ed.*, 2001, **40**, 3990-+.
212. S. Jager and M. Famulok, *Angew. Chem. Int. Ed.*, 2004, **43**, 3337-3340.
213. S. Jager, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum and M. Famulok, *J. Am. Chem. Soc.*, 2005, **127**, 15071-15082.
214. T. Gourlain, A. Sidorov, N. Mignet, S. J. Thorpe, S. E. Lee, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1898-1905.
215. J. P. Genet and M. Savignac, *J. Organomet. Chem.*, 1999, **576**, 305-317.
216. L. H. Thoresen, G. S. Jiao, W. C. Haaland, M. L. Metzker and K. Burgess, *Chem.-Eur. J.*, 2003, **9**, 4603-4610.
217. J. Balintová, R. Pohl, P. Horáková, P. Vidláková, L. Havran, M. Fojta and M. Hocek, *Chem. Eur. J.*, 2011, **17**, 14063-14073.
218. J. Riedl, R. Pohl, L. Rulíšek and M. Hocek, *J. Org. Chem.*, 2011, **77**, 1026-1044.
219. L. Kalachova, R. Pohl and M. Hocek, *Org. Biomol. Chem.*, 2012, **10**, 49-55.
220. H. Weizman and Y. Tor, *J. Am. Chem. Soc.*, 2002, **124**, 1568-1569.
221. H. Sawai, J. Nagashima, M. Kuwahara, R. Kitagata, T. Tamura and I. Matsui, *Chem. Biodivers.*, 2007, **4**, 1979-1995.
222. M. Nuzzolo, A. Grabulosa, A. M. Z. Slawin, N. J. Meeuwenoord, G. A. van der Marel and P. C. J. Kamer, 2010, **2010**, 3229-3236.
223. L. Clima and W. Bannwarth, *Helv. Chim. Acta.*, 2008, **91**, 165-175.
224. L. J. Brown, J. P. May and T. Brown, *Tetrahedron Lett.*, 2001, **42**, 2587-2591.
225. T. Macke and D. A. Case, in *Molecular Modelling of Nucleic Acids*, eds. N. B. Leontes and J. SantaLucia Jr., American Chemical Society, Washington DC, 1998, pp. 379-393.
226. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
227. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247-260.
228. A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham III, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817-3829.
229. J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049-1074.
230. I. S. Joung and T. E. Cheatham, *J. Phys. Chem. B*, 2008, **112**, 9020-9041.
231. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33-38.