Twisted Intercalating Nucleic Acids (TINA) in Guanosine-rich Oligonucleotides

A thesis submitted in the partial fulfilment

of the requirements for the degree of

Doctor of Philosophy in Chemistry

Massey University,

Palmerston North, New Zealand

MASSEY UNIVERSITY

Osman Doluca

2013

Abstract

The main role in the structural diversity of DNA molecules belongs to guanosines due to their array of hydrogen bond donors and acceptors, large aromatic surface and ability to adopt *syn* or *anti* conformations. These properties lead to the formation of various DNA topologies such as triplexes or G-quadruplexes by guanosine-rich oligonucleotides. For a long time these secondary structures were mainly considered to be a fascinating phenomenon with little practical use; it was subsequently realised that these structures are likely to be formed under physiological conditions and therefore might be involved in many important biological processes, including genome recombination, telomere stability and regulation of gene expression. Thus, there is a growing interest in development and control of these non-traditional nucleic acid structures.

Although the secondary structures of nucleic acids can be controlled to a certain extent by the careful design of oligonucleotide sequence this strategy alone is not always sufficient. In this thesis we investigated how to control the assemblies of guanosine-rich oligonucleotides using a novel tool, twisted intercalating nucleic acids (TINAs). The incorporation of pyrene-containing TINA monomers into guanosine-rich oligonucleotides led to the formation of stable triplexes or G-quadruplexes depending on the position of TINA monomers. In the light of our results, we have established a set of rules that helps to create a desired structure of guanosine-rich oligonucleotides using TINA molecules.

In the second half of the thesis we focused on expanding the functionality of TINA conjugated oligonucleotides. In terms of fluorescence, we synthesised several fluorescently-silent triplex-forming oligonucleotides (TFOs) equipped with a dye at different positions in the DNA. Fluorescence properties were strongly dependent on the position of the dye. These fluorescently silent TFOs showed up to an 18-fold increase in fluorescent intensity upon triplex formation.

These findings lay the foundation for the future design of artificial DNA sequences for expanding the repertoire of DNA secondary structures and function.

Acknowledgements

I would like to thank my supervisors Professor Geoff Jameson and Dr Vyacheslav V. Filichev for their support and assistance throughout this project. Especially Vyacheslav for the enthusiasm he showed towards the project and the time taken to show me various techniques of DNA chemistry. I would like to thank Professor Alexandre Boutorine for his support to this project and hospitality.

Thanks to all former and present colleagues in IFS for their assistance over the years, especially Jamie Withers and Emad Al-imarah for their friendship and Pat Edwards for his help with NMR. I am also grateful for assistance of departmental staff over the years I have been at Massey University.

I would like to acknowledge important financial support from the Marsden grant administered by the Royal Society of New Zealand, the Dumont d'Urville grant from NZ/France Science and Technology Support Programme and a doctoral completion bursary from Massey University.

Many thanks to my dear parents, Saliha and Nevzat Doluca, for their infinite support reaching all the way to the other side of the world.

Above all I would like to thank my dear wife, Rose, for her patience and support during completion of this PhD and my kids, Leila and Musa, for making my life cOL*ou***r***f*uL at every step.

Table of Contents

List of Figures

List of Tables

Abbreviations

$(EG)_6$	hexaethylene-glycol
µL	microlitre
µM	micromole/litre
µmol	micromole
A	adenosine
Å	Ångström
ABL	Abelson murine leukemia viral oncogene
Abs	absorbance
ACN	acetonitrile
AcOH	acetic acid
aq	aqueous
ATR	attenuated total reflectance
BCR	breakpoint cluster region
bp	basepair
C	cytosine
CD	circular dichroism
COMBO-FISH	combinatorial oligonucleotide probes in fluorescence *in situ* hybridization
CPG	controlled pore glass
CTAB	cetyl trimethylammonium bromide
CuAAC	copper-assisted azide-alkyne cycloaddition
DCA	dichloroacetic acid
DCI	4,5-dicyanoimidazole
DCM	dichloromethane
ddH_2O	double-distilled water
dH_2O	distilled water
DIEA	di*iso*propylethylamine
DMA	*N*,*N*-dimethylacetamide
DMF	*N*,*N*-dimethylformamide
DMPA	*N*,*N*-dimethylaminopropylamine
DMSO	dimethylsulfoxide
DMT	4,4′-dimethoxytrityl
DNA	deoxyribonucleic acid

Dp	(*N*,*N*-dimethylamino)propylamide
dsDNA	double-stranded DNA
ε	extinction coefficient
EDTA	*N*,*N*-ethylenediaminetetraacetic acid
eq	equivalent
ESI	electrospray ionisation
EtBr	ethidium bromide
EtOAc	ethyl acetate
EtOH	ethyl alcohol
F_c	fluorescence intensity of DNA complex
FISH	fluorescence *in situ* hybridization
FRET	Förster resonance energy transfer
F_{ss}	fluorescence intensity of single-stranded DNA
FT-IR	Fourier transform infrared spectroscopy
G	guanosine
HATU	*O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	human immunodeficiency virus
HPLC	high pressure liquid chromatography
HRAS	Harvey rat sarcoma viral oncogene
Im	*N*-methylimidazole
INA	intercalating nucleic acid
IR	infrared
K_d	dissociation constant
KRAS	Kirsten rat sarcoma viral oncogene
L	litre
LNA	locked nucleic acid
M	mole/litre
MALDI	matrix-assisted laser desorption/ionization
MeOH	methyl alcohol
MGB	minor groove binder
min	minute
mL	millilitre
mM	millimole/litre
mmol	millimole
mRNA	mitochondrial ribonucleic acid

MYC	myelocytomatosis viral oncogene
NaOAc	sodium acetate
N-FISH	non-denaturing fluorescence *in situ* hybridization
nm	nanometre
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
ON	oligonucleotide
P	pyrene
p	terminal phosphate
PAGE	polyacrylamide gel electrophoresis
Ph	phenyl
PhSH	thiophenol
PNA	peptide nucleic acid
ppm	parts per million
Ps	psoralen
Py	*N*-methylpyrrole
qRT-PCR	quantitative real time polymerase chain reaction
RNA	ribonucleic acid
s	second
SNP	single nucleotide polymorphism
SRC	sarcoma viral oncogene
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
T	thymidine
$T_{½}$	mid-transition point
TBTA	tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
TCA	trichloroacetic acid
TDS	thermal difference spectra
TEAA	triethylammonium acetate
TFA	trifluoroacetic acid
TFO	triplex-forming oligonucleotide
THF	tetrahydrofuran
TINA	twisted intercalating nucleic acid
TLC	thin-layer chromatography
T_m	melting temperature
TO	thiazole orange

TOF	time of flight
TSP	trimethylsilyl propionate
U	uridine
UV	ultraviolet
Vis	visible
β	β-alanine
γ	γ-aminobutyric acid
λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
Φ_F	fluorescence quantum yield

Chapter 1. Introduction

1.1 Aim

Even though DNA is commonly found in B-type double-stranded helical topology, it is capable of forming alternative structures, such as quadruplexes, triplexes and junctions. Until the mid-1990s it was thought that the role of DNA in gene expression was no more than encoding the sequence for transcription and that all regulatory roles belonged to RNAs and proteins (Rawal, Kummarasetti et al. 2006). However, in recent years, these non-B-type DNA structures have been discovered in relation with gene regulation (Perez-Martin and de Lorenzo 1997; Pedersen, Jensen et al. 2000; Bacolla, Jaworski et al. 2004; Bacolla and Wells 2004; Palumbo, Memmott et al. 2008). Most of these structures are characterised in the cell and found to be induced by certain nucleobase sequences which are commonly encountered in the sites responsible for gene expression. Today it is clear that these sequences, which form various DNA structures including G-quadruplexes and triplexes, have major roles in gene regulation; however, we have only scratched the surface of understanding how these structures regulate gene expression (Bacolla and Wells 2004). For this reason it is essential to understand and control the formation of these structures. The ability to control their formation may have far-reaching implications in biotechnology, medicine and even nanotechnology.

The topology of DNA is strongly dependent on the sequence (Rawal, Kummarasetti et al. 2006). For example, the triplex formation requires continuous tracts of purines (polypurines) or G-quadruplexes require guanine islands at least 2 nucleotides long. Both of these structures share same sequence characteristic: high guanine content (G-rich). For this reason G-rich sequences are of particular interest for scientists dealing with DNA structures. For this reason our research focuses on G-rich oligonucleotides and the structures formed by them.

Chemical modification is one of the ways that allows controlling DNA-based assemblies. For years scientists have been working on small molecules to be incorporated into DNA structures. These modifications are capable of interfering with the overall topology and promote a defined

structure. Here, we use a recently developed modification called twisted intercalating nucleic acid (TINA). First, we investigate its potential in the making of antiparallel triplex-forming oligonucleotides (TFOs) and define the guidelines about how to incorporate TINA to target different sequences. Next, these triplex-forming probes were improved by incorporation of additional modifications such as DNA-binders or fluorescent molecules. These modifications are expected to improve efficiency of TINA-TFOs for DNA visualisation applications. Finally, we studied TINA's influence on the formation of G-quadruplex structures and aimed to improve our understanding of the effect of TINA incorporation on G-rich sequences.

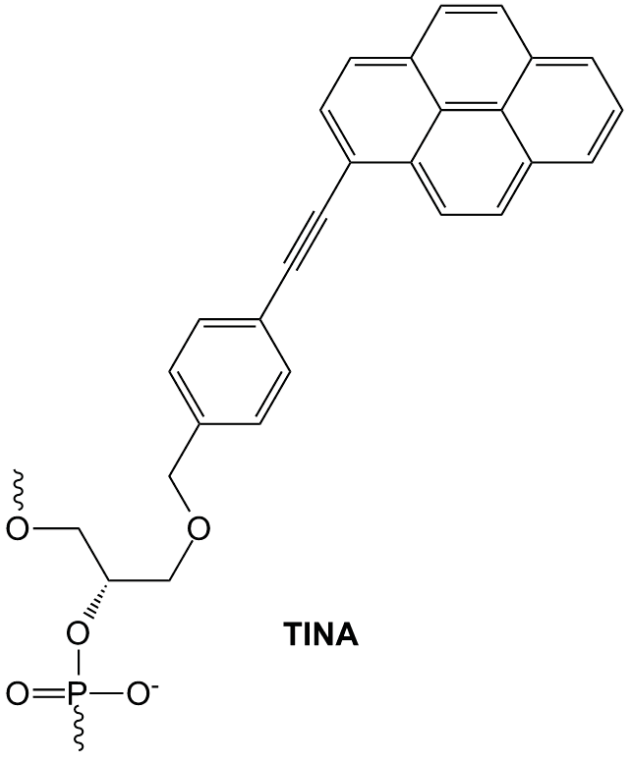


Figure 1.1 Twisted Intercalating Nucleic Acid (TINA).

1.2 Properties of DNA

DNA, or deoxyribonucleic acid, is the hereditary material in humans and almost all other organisms. In eukaryotes DNA is located in the cell nucleus (where it is called nuclear DNA), but a small amount of DNA can also be found in the mitochondria (where it is called mitochondrial DNA or mtDNA).

DNA stores information as a code made up of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T).

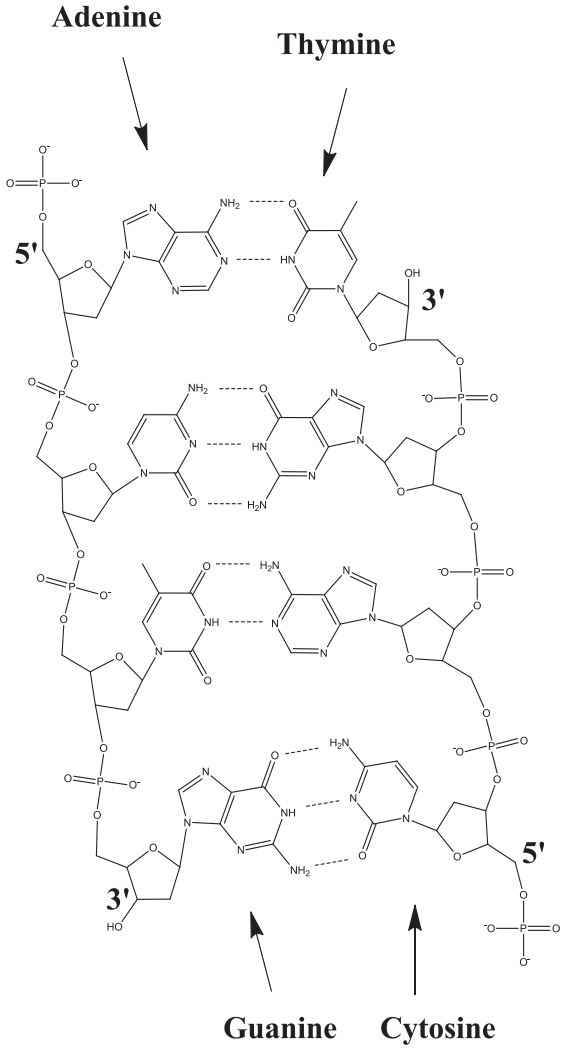


Figure 1.2 A double-stranded DNA (dsDNA) formed by two single-stranded DNA (ssDNA) chains.

Alternating phosphate and sugar residues form the backbone (Figure 1.2). The base is bound to the sugar, 2-deoxyribose, through a β-*N*-glycosidic bond. The sugars are joined together by phosphate groups that form ester bonds between the third and fifth carbon atoms of adjacent sugar rings. These asymmetric bonds give a strand of DNA its polarity. The asymmetric ends of DNA strands are called the 5′ and 3′ ends depending on which carbon of the sugar is closer to the end. Together, a base, a sugar, and a phosphate are called a nucleotide (Clark 2005).

In living organisms, DNA exists as a pair of single-stranded DNA (ssDNA), which is called double-stranded (dsDNA) (Figure 1.2). In dsDNA, two strands are in opposite orientations (antiparallel orientation). The bases of opposite strands are aligned with each other and form the core of the complex while the sugar-phosphate backbone is left outside.

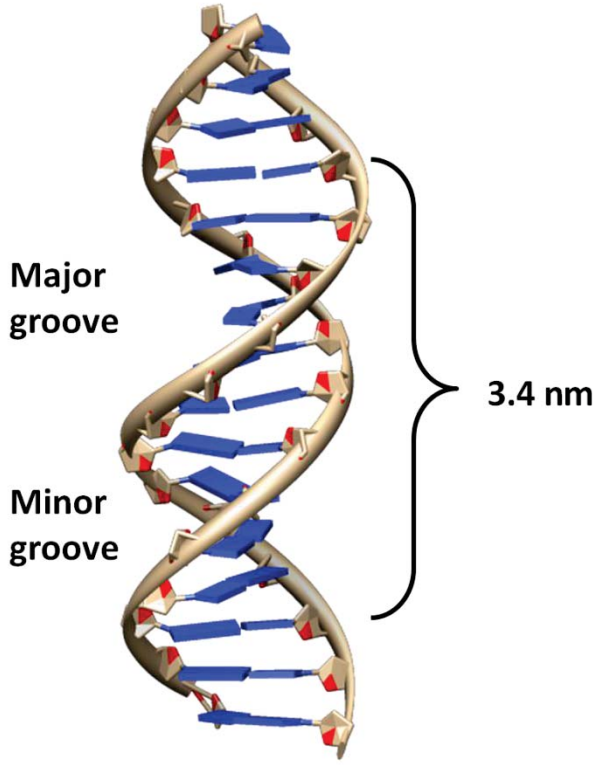
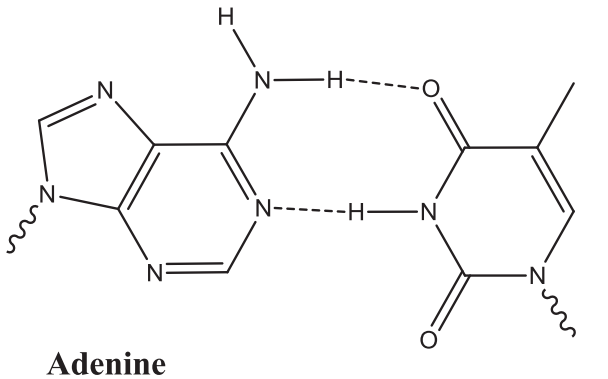


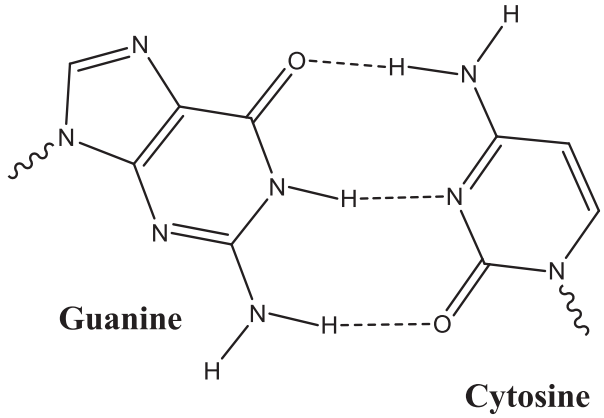
Figure 1.3 B-type DNA model.

In the most common structure, B-type of DNA, the intertwined strands form two grooves of different widths, referred to as the major groove (22 Å) and the minor groove (12 Å) (Figure 1.3). The narrowness of the minor groove means that the edges of the bases are more accessible

in the major groove. As a result, proteins like transcription factors that can bind to specific sequences in double-stranded DNA usually make contacts to the sides of the bases exposed in the major groove. The B-type helix (DNA/DNA) is the net result of a set of forces involved in the formation of DNA complexes. These can be listed as; hydrogen bonds between matching bases (Watson Crick base pairing), the π-π stacking interactions of neighbouring bases, electrostatic repulsion between the negatively charged phosphates on the backbones and also hydration level of the DNA complex (Reddy and Berkowitz 1989; Bruisten, Reiss et al. 1998).



A-T base-pairing



G-C base-pairing

Figure 1.4 Top, an AT base-pair with two hydrogen bonds. Bottom, a GC base-pair with three hydrogen bonds. Hydrogen bonds between the pairs are shown as dashed lines.

The most essential requirement for stable dsDNA formation is that two strands have complementary sequences of bases, in other words, matching base-pairs. According to Watson-Crick base pairing, these hydrogen bonds can only form between a purine base (A or G) and a pyrimidine base (T or C) (Figure 1.4). While between guanine (G) and cytosine (C) three hydrogen bonds form, the hydrogen bond number between adenine (A) and thymine (T) is only two. As a result, double-stranded DNA with high GC-content is more stable than DNA with low GC content (Yakovchuk, Protozanova et al. 2006).

The nucleotides can exist in two distinct orientations according to their *N*-glycosidic bond (Figure 1.5). These conformations are identified as *syn* and *anti*. In the *anti* confirmation H6 of pyrimidines or H8 of purines align above the sugar ring, whereas in *syn* conformation O2 of pyrimidines or N3 of purines switch upwards. In naturally occurring nucleotides, the *anti* conformation predominates, however, the β-*N*-glycosidic bond of purines can rapidly interconvert between the two conformations, although one can assign a preference to each DNA structure. For instance, in B-type DNA the *anti* conformation is adopted and as a consequence the backbone chains run downwards on the right of the minor groove and run upwards on the left of the minor groove.

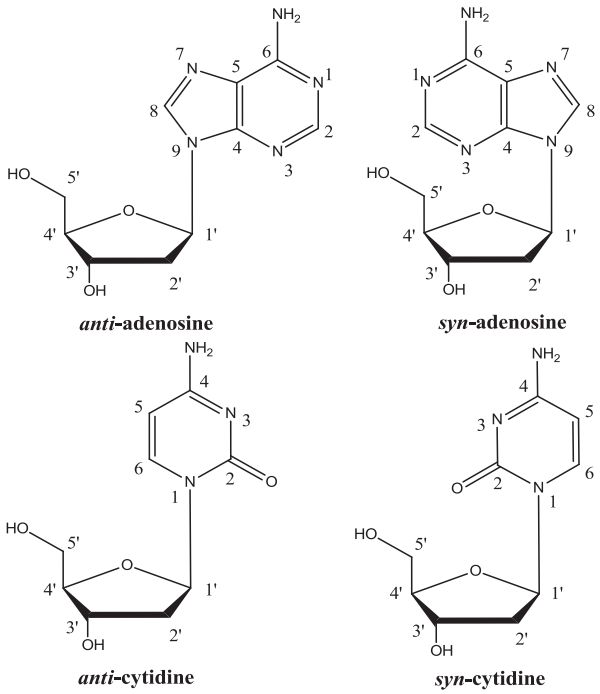


Figure 1.5 *Anti* and *syn* confirmations of adenosine and cytidine.

Of the four common bases of DNA, guanine deserves the most attention. First of all, it has the strongest tendency towards intermolecular stacking interactions due to its two aromatic rings. It has the ability to adopt both *syn* and *anti* conformations and finally, guanine's array of hydrogen bond donors and acceptors located around its rings can form uncommon hyndrogen bonds, such as the Hoogsteen hydrogen bond. As result of these properties, guanines take part in structures as diverse as B-DNA, Z-DNA, hairpins, cruciforms, triplexes and G-quadruplexes; these structures, though varied, are closely related and guanine-rich oligonucleotides can often participate in equilibria between several different assemblies (Son, Guschlbauer et al. 1972; Sundaralingam and Westhof 1981; Henderson, Hardin et al. 1987; Guschlbauer, Chantot et al. 1990). For this reason, in this study we will focus on G-rich oligonucleotides and the two main topologies they participate in; triplexes and G-quadruplexes.

1.2.1 Triple Helical DNA

Formation of a triplex occurs when a third oligonucleotide strand binds to the target duplex in the major groove of dsDNA, which is wide enough to facilitate the wrapping and binding of the third strand around the helix. This third strand is called triplex-forming oligonucleotide (TFO). The stability of the triplex depends on the length of the TFO and base interactions. Because the bases of the duplex are already occupied with Watson-Crick hydrogen bonds, a new set of interactions are established between bases of the TFO and bases of one strand which belongs to the duplex. This strand needs to have a polypurine sequence. It means the strand that TFO binds needs to be formed only by As and/or Gs.

The interactions between the bases of the TFO and the duplex are called Hoogsteen or reverse Hoogsteen base pairing (Figure 1.6). The purine bases involved in the duplex formation have an extra position that has not been occupied by Watson-Crick hydrogen bonds, N7 of the purine bases. This position participates in the formation of Hoogsteen hydrogen bonds in the major groove. For this reason, only a polypurine sequence in the duplex can facilitate the formation of the triplex (Blackburn 2006).

Depending on the orientation of the third strand relative to the polypurine strand of the duplex, two types of triplexes can be formed. A TFO can either bind parallel or antiparallel to the

polypurine strand of the duplex and can form either a parallel or antiparallel triplex, respectively (Figure 1.6).

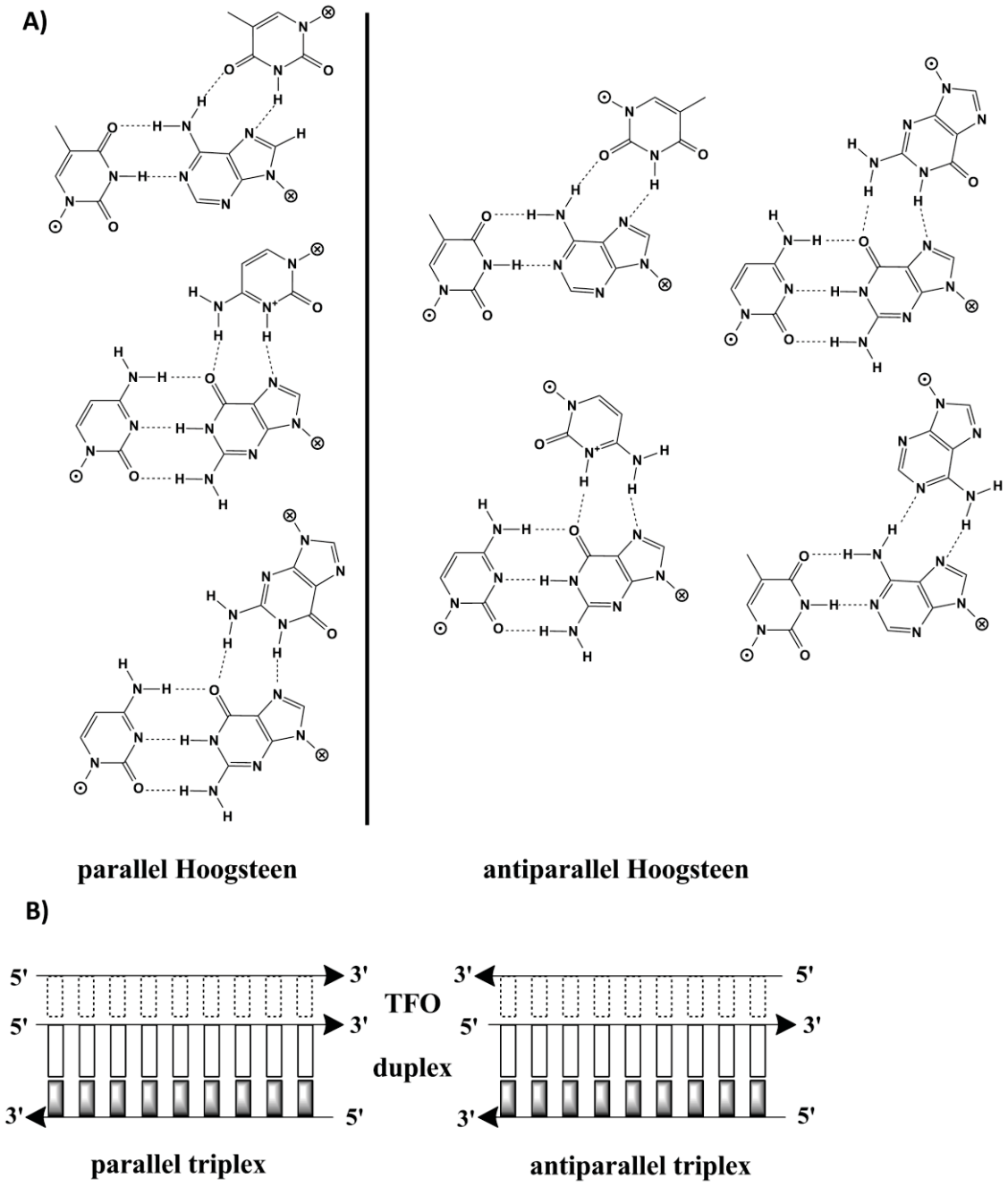


Figure 1.6 A) Models for T•A-T, C^+•G-C, G•G-C, A•A-T base triplets within a triple helix motif. Directionality of the backbones is indicated with ⊗ (3′) and ⊙ (5′). B) Orientations of parallel and antiparallel TFOs relative to the duplex strands. Shaded boxes represent pyrimidines, empty boxes represent purine bases, boxes with dashed borders represent bases of the third strand (Sun, Garestier et al. 1996).

1.2.1.1 Properties of Parallel Triplexes

Three types of Hoogsteen base pairings can be established in a parallel triplex (Figure 1.6, A). According to parallel Hoogsteen base pairing the adenine (A) of an **A-T** dinucleotide can match with a thymidine to form a T•A-T triplet (• : Hoogsteen base pairing, - : Watson-Crick base pairing). For guanine (G) of a **G-C** duplet, there are two possibilities. In theory a guanine can match either a guanine (G•G-C) or a protonated cytosine (C^+•G-C) of the third strand according to Hoogsteen base pairing. However, the ability to form Hoogsteen base pairs is not enough for stable triplex formation. First of all, any polypurine sequence shorter than ten base pairs has been determined to be too short to establish stable triplexes at room temperature (Blackburn 2006). Additionally, stability strongly depends on the sequence. For example, even though guanine is a possible match for a **G-C** dinucleotide, a parallel TFO which has an irregular sequence of guanines and thymidines, cannot form a successful triplex. This is due to the arrangement of the atoms in the plane of base triplets (Figure 1.6, A). Because the C1' atom of guanine is far from thymidine's, the TFO backbone cannot tolerate the distortion. On the other hand, the only combinations that have isomorphous location of their C1' atoms are T•A-T and C^+•G-C triplets (Sun, Garestier et al. 1996).

Another issue of triplex formation is the repulsion between the polyanionic phosphate backbones of the DNA strands. The attraction between the protonated cytosines (C^+) and polyanionic phosphate backbones of dsDNA overcomes this repulsion and enhances the affinity of the TFO to the duplex. As a result, cytosines are preferred in the parallel triplexes. However, cytosines need to be protonated at N3 position to form Hoogsteen hydrogen bonds. The pK_a value of cytosine in isolation was detected to be 4.3 (Blackburn 2006), and this value was calculated to be 4.1 inside the oligonucleotide. This is due to negatively charged phosphates of the oligonucleotide (Pack, Wong et al. 1998). Under physiological conditions, pH is not low enough to facilitate this protonation and the lowering pH in the cytoplasm results in the death of the cell. In conclusion, the use of unmodified parallel TFOs in living cells is limited by requirement of protonation of cytosines.

1.2.1.2 Properties of Antiparallel Triplexes

Reverse Hoogsteen hydrogen bonds are established when the third strand is in antiparallel orientation relative to the polypurine strand of the duplex (Figure 1.6, B). Additional to the Hoogsteen base pairing, an A•A-T triplet is also possible with reverse Hoogsteen base pairing.

Bases on antiparallel TFOs bind to the duplex in the *anti* conformation in the β-*N*-glycosydic bond (Figure 1.6, A). Thus, guanines on the third strand do not produce backbone distortion and are more favorable than protonated cytosines. As a result, antiparallel triplexes can be formed in physiological conditions. However, the antiparallel triplex-forming oligonucleotides are rich in guanines and they show a high tendency to form highly stable G-quadruplexes in physiological conditions instead of participating in triplex formation (Blackburn 2006). This tendency is the major challenge for antiparallel triplex formation. On the other hand, pH independence of antiparallel triplexes attracts interest for triplex studies in live cells.

1.2.2 G-quadruplexes

G-quadruplexes are high-order DNA/RNA structures formed from G-rich sequences, which are built around tetrads of Hoogsteen hydrogen-bonded guanine bases. G-quadruplexes can be formed from up to four separate DNA strands and can display a wide variety of topologies.

They can be defined in general terms as structures formed by a core of at least two stacked G-tetrads, which are held together by loops arising from intervening mixed sequence nucleotides that are not usually involved in the tetrads themselves. The combination of the number of stacked G-tetrads, the polarity of the strands, the location and the length of the loops and *syn* or *anti* conformations of the guanine bases leads to a vast variety of potential G-quadruplex structures (Figure 1.7, B) (Burge, Parkinson et al. 2006). A parallel G-quadruplex has all four strands aligned in the same orientation, possesses guanines in *anti* configuration and it is usually formed by four identical strands. On the other hand, antiparallel G-quadruplexes contain guanines with *anti* as well as guanines with *syn* conformation. These structures are commonly formed by folding of one or two strands.

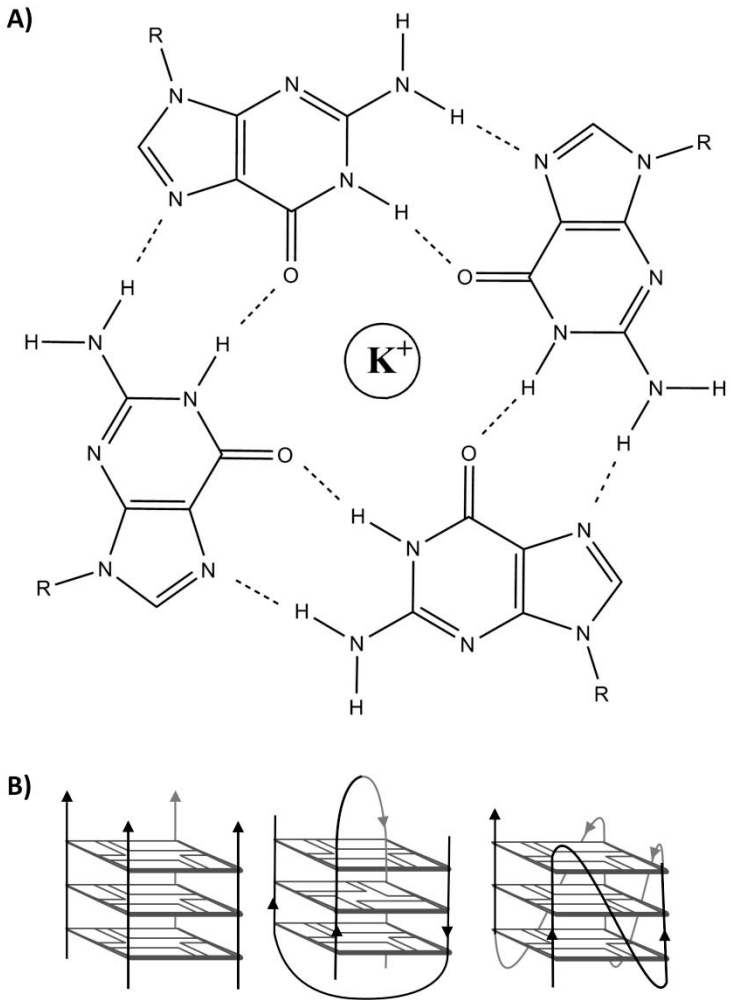


Figure 1.7 A) G-tetrad formation. 'R' refers to backbones of strands that form G-quadruplex. B) Various G-quadruplex structures. Arrows indicate 5′ to 3′ polarity. Squares indicate guanine arrangement inside a G-tetrad.

In a G-tetrad, each guanine is engaged in four hydrogen bonds via Hoogsteen bonding, such that guanines are related by a four-fold rotation axis and are almost coplanar. Each guanine directs its O6 carbonyl oxygen towards the central core of the tetrad (Figure 1.7, A). These carbonyl oxygens repel each other resulting in the destabilization of the complex. In the presence of monovalent ions, such as K^+ or Na^+, the cation is fixed in the core and stabilizes the complex by compensating the repulsion between the carboxyl oxygens. The stabilizing effect of the cation depends on the ionic radius. The order of preference is observed as $K^+ > Na^+ > Cs^+ > Li^+$ (Simonsson 2001). Replacement of K^+ with Li^+ or Na^+ is a common method to reduce stability of G-quadruplex structures in G-quadruplex related studies (Neaves, Huppert et al. 2009).

Stability of G-quadruplexes are also strongly dependent on number of consecutive guanine residues. G-quadruplexes with longer G-tracts can form more G-tetrads and with addition of each G-tetrad stability increases dramatically due to stacking interactions. The stability of G-quadruplexes with more than six tetrads is usually so high that melting temperatures cannot be measured in aqueous buffer (Lane, Chaires et al. 2008; Payet and Huppert 2012). On the other hand, studies focusing on G-quadruplexes with two or three tetrads can show the stabilising effect of the modification introduced into the sequence (Petraccone, Erra et al. 2003; Pagano, Martino et al. 2008).

G-quadruplexes are believed to have significant role in regulation of cell apoptosis and other cell functions. Potential G-quadruplex sequences have been identified in G-rich eukaryotic telomeres, and more recently in non-telomeric DNA (Burge, Parkinson et al. 2006). Repeating G-rich DNA sequences in the telomeres prevent chromosomes from losing base-pair sequences at the ends. However, each time a cell divides, some of the telomere is lost. When the telomere becomes too short, the chromosome reaches a "critical length" and can no longer replicate which leads to aging of the cell. Telomere terminal transferases, also called telomerases, are the enzymes that elongate chromosomes by adding short sequences to the end of existing chromosomes. These guanine-rich sequences are species dependent (ie. dTAAGGG in human). Telomerases, which are expressed in almost all tumour cells, are absent in somatic cells that do not reproduce anymore. Due to the elongation of the telomeres by telomerases, tumour cells do not age. It is thought that stabilization of the telomeres by G-tetrad-interactive compounds may inhibit the activity of this enzyme and stop tumour growth.

Other antiproliferative functions of G-quadruplexes have also been shown (Gomez, Aouali et al. 2003; Cogoi, Paramasivan et al. 2007). Interestingly, in several cases, the biological effects of oligonucleotides designed as antisense agents, which target mRNA molecules to repress expression of a certain protein, were found to be unrelated to the inhibition of the gene translation, but instead were associated with the formation of G-quadruplex structures (Raymond, Soria et al. 2000). In studies of G-rich sequences in the promoter region of various genes such as *MYC* and *KRAS* have been shown to form G-quadruplexes (Simonsson, Pecinka et al. 1998). Because these complexes are recognised by transcription factors, the transcription of these genes can be regulated by G-quadruplex binding molecules; therefore, G-quadruplexes attract interest as targets for drug design (Hurley, Wheelhouse et al. 2000).

1.2.3 Using Synthetic Chemistry for Studying DNA Structures

Synthetic chemistry is a valuable tool to study nucleic acid structures. An automated DNA synthesis enables us to synthesise almost any DNA sequence chemically with a length up to 200 nucleotides. The development of chemical synthesis of DNA has facilitated the base-by-base modification, which has extended the ability to introduce new functional molecules into the DNA.

1.2.3.1 Automated DNA Synthesis

Several methods have been developed to synthesise oligonucleotides chemically. Nucleoside phosphoramidite chemistry is the most common method. A nucleoside phosphoramidite is a derivative of a natural or a synthetic nucleoside with protecting groups added to its reactive exocyclic amine and hydroxyl groups (Figure 1.8, A). The naturally occurring nucleotides (nucleoside-3'- or 5'-phosphates) are insufficiently reactive to afford the synthetic preparation of oligonucleotides. A dramatically more reactive *N*,*N*-di*iso*propyl phosphoramidite group is therefore attached to the 3'-hydroxy group of a nucleoside to form a nucleoside phosphoramidite. To prevent undesired side reactions, all other functional groups of nucleosides have to be protected by attaching protective groups, such as acid-labile DMT (4-4′-dimethoxytrityl) and base-labile 2-cyanoethyl groups (Figure 1.8). Upon the completion of the oligonucleotide chain assembly, all the protecting groups are removed (Reese 2005).

Oligonucleotide synthesis is carried out by a stepwise addition of the phosphoramidites to the 5'-terminus of the growing chain until the desired sequence is assembled. Each addition is referred to as a synthetic cycle and consists of four chemical reactions (Figure 1.9):

Step 1. De-blocking (detritylation): The DMT group is removed with a solution of an acid, such as trichloroacetic acid (TCA) or dichloroacetic acid (DCA), in an inert solvent (dichloromethane or toluene) and washed out, resulting in a free 5' hydroxyl group on the first base bound to the solid support.

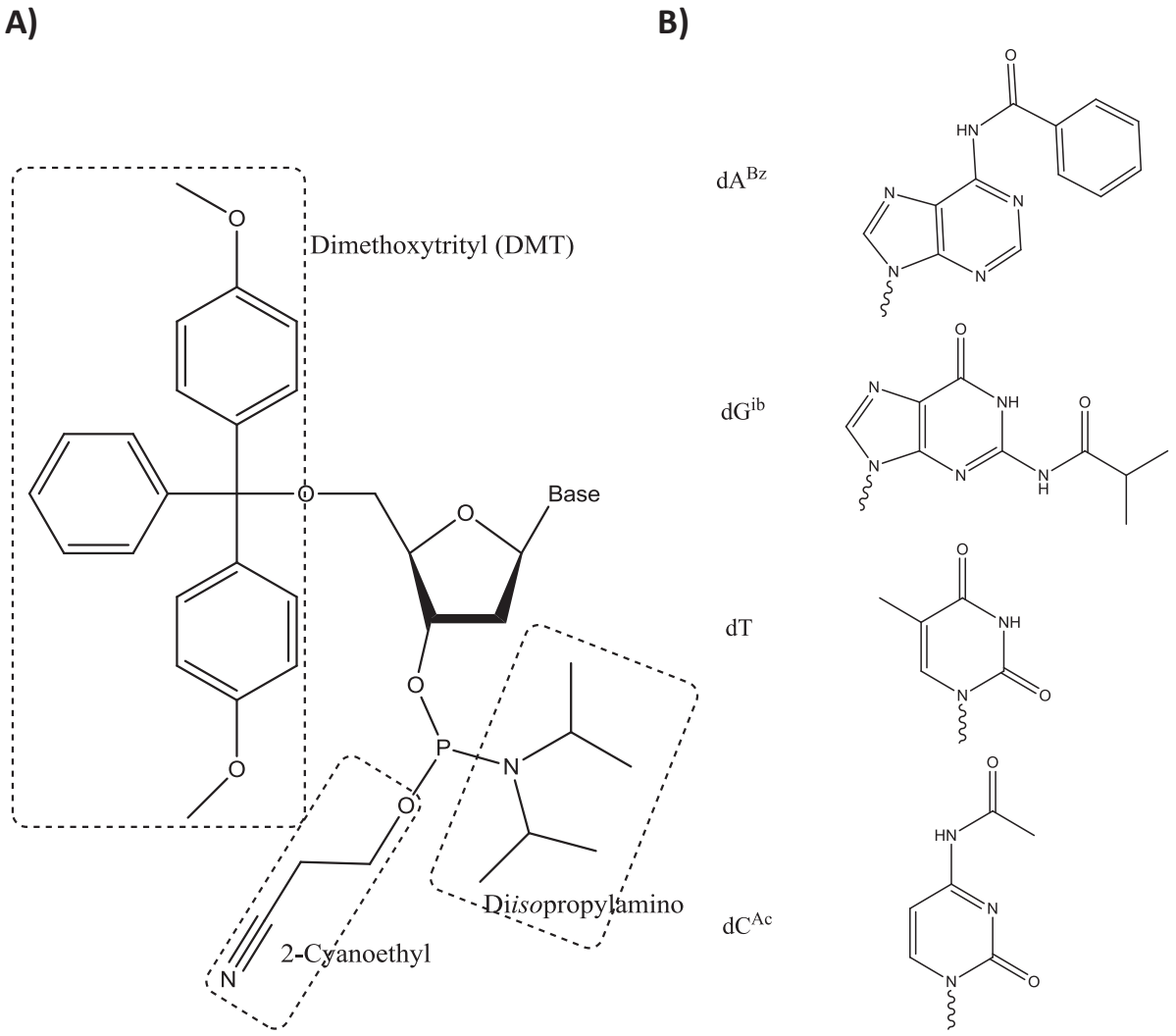


Figure 1.8 Design of nucleoside phosphoramidites (A) and protected nucleic bases (B) for DNA synthesis.

Step 2. Coupling: A nucleoside phosphoramidite (or a mixture of several phosphoramidites) is activated by an acidic azole catalyst, tetrazole, 2-ethylthiotetrazole, 2-benzylthiotetrazole, 4,5-dicyanoimidazole (DCI), or a number of similar compounds. This mixture is brought in contact with the starting solid support (first coupling) or oligonucleotide precursor (following couplings) whose 5'-hydroxy group reacts with the activated phosphoramidite moiety of the incoming nucleoside phosphoramidite to form a phosphite triester linkage. This reaction is very rapid and requires, on small scale, about 20 s for its completion. The phosphoramidite coupling is also highly sensitive to the presence of water and is commonly carried out in anhydrous acetonitrile. Unbound reagents and by-products are removed by washing.

Step 3. Capping: After the completion of the coupling reaction, a small percentage of the solid support-bound 5'-OH groups (0.1 to 1.0 %) remain unreacted and need to be permanently blocked from further chain elongation to prevent the formation of oligonucleotides with an internal base deletion commonly referred to as (n-1) shortmers. This is done by acetylation of the unreacted 5'-hydroxy groups using a mixture of acetic anhydride and 1-methylimidazole as a catalyst. Excess reagents are removed by washing.

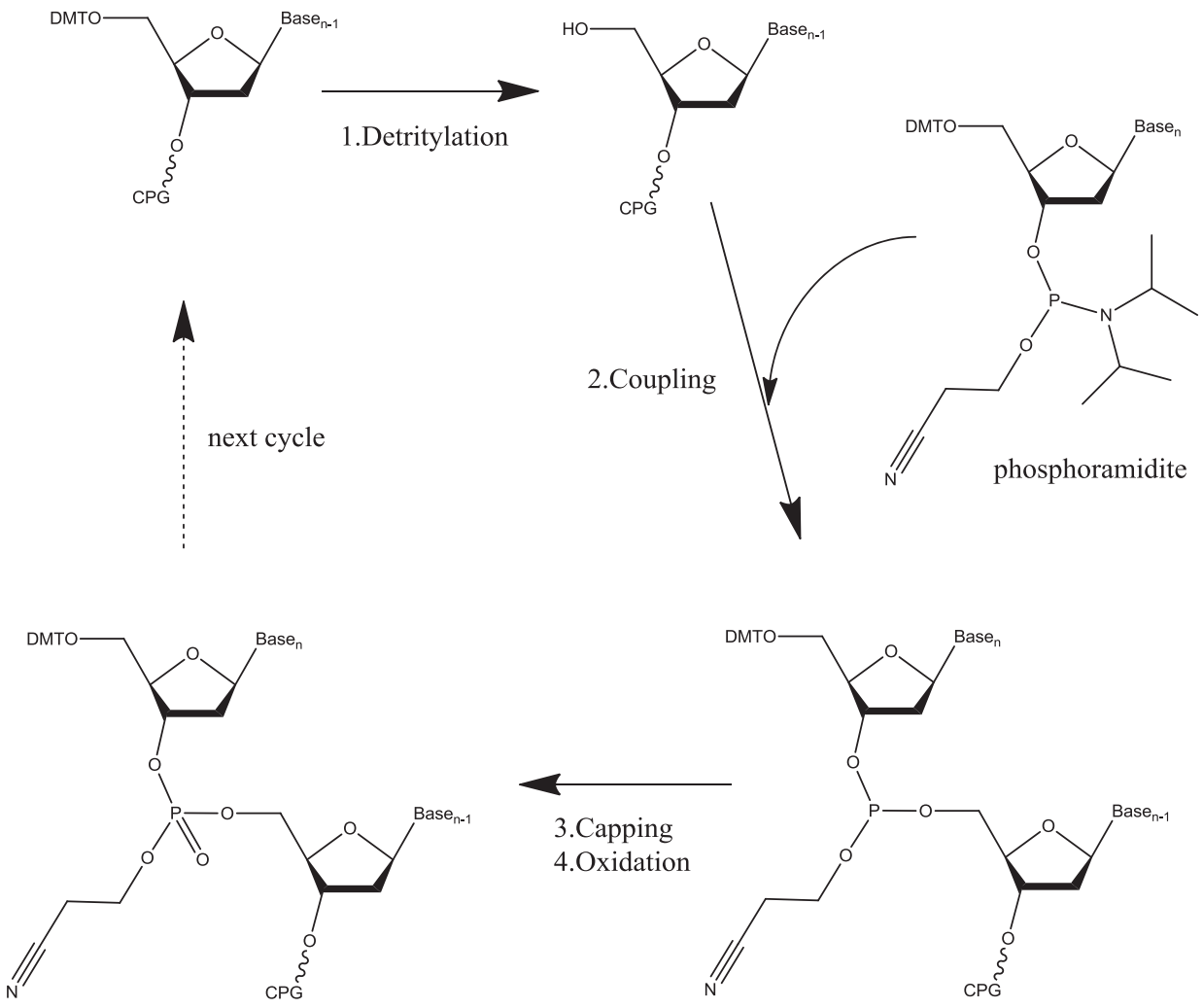


Figure 1.9 Synthetic cycle for preparation of oligonucleotides by phosphoramidite method.

Step 4. Oxidation: The newly formed tricoordinated phosphite triester linkage is not natural and is of limited stability under the conditions of oligonucleotide synthesis. The treatment of the support-bound material with iodine and water in the presence of a weak base (pyridine, lutidine, or collidine) oxidizes the phosphite triester into a tetracoordinated phosphate triester, a protected precursor of the naturally occurring phosphate diester internucleosidic linkage.

The manual addition of phosphoramidites during automated DNA synthesis is referred as hand-coupling. In most cases hand-coupling requires pausing of the automated synthesis right after the detritylation step. The desired phosphoramidite is added directly onto the column that contains the controlled pore glass (CPG) support. Without delay automated synthesis is continued with addition of an activator (*ie.* DCI). A longer coupling period is used for hand-coupling (5-10 min).

Oligonucleotides (ONs) can be modified in several different ways by utilising the active groups of the nucleotide or creating nucleotide analogues. By direct incorporation of modified nucleosides during automated DNA synthesis, modified bases can be incorporated internally or at the 5' end. This site-specific method is constrained by the availability of specialized phosphoramidites. On the other hand, 3' end modifications rely on the existence of modified CPG columns.

1.3 Post-Synthetic Treatments and Purification

After the synthesis is complete, the fully protected, solid support-bound oligonucleotides are subjected to deprotection. Initially, the oligonucleotides are released from the solid-phase and deprotected (base and phosphate) by treatment with aqueous ammonium hydroxide, aqueous methylamine, their mixtures, gaseous ammonia or methylamine or, less commonly, solutions of other primary amines or alkalis at ambient or elevated temperatures. After deprotection and evaporation, the oligonucleotides are purified by reverse-phase HPLC. In this step, the 5'-terminal DMT group serves as a hydrophobic handle for purification. The collected material is then detritylated under aqueous acidic conditions and, finally, desalted by precipitation. Alternatively denaturing gel electrophoresis can be used for purification of DNA instead of HPLC. This method requires cutting of the desired section of the gel and incubating in a buffer

for days to dissolve the gel and release the DNA. Gel purification may be preferred especially when HPLC cannot provide a good separation of the sample; however, it has low yield.

Post-synthetic modifications use functional groups such as amino modifiers, the 3' or 5' hydroxyl groups or the phosphate groups (Goodchild 1990). Although the yields are often low, the reaction of functional groups attached to the bases with dye moieties is widely used to label ONs either on the ends or internally.

1.3.1 Backbone Modifications

One of the main aims with several backbone modifications is the removal of electrostatic repulsion between negatively charged strands of the triplex. *N*,*N*-Dimethylaminopropylamine (DMAP), *N*,*N*-diethylethylenediamine (DEED) and methoxyethylamine (MeOEt) were designed to replace one of the oxygens from the phosphate. These modifications removed the requirement of charge neutralization between phosphates and were capable of forming triplexes even in the absence of $MgCl_2$ (Dagle and Weeks 1996; Bailey, Dagle et al. 1998).

Phosphothioate (PS), on the other hand, was used to enhance nuclease resistance. PS is used to replace one of non-bridging oxygen atoms of the phosphate with a sulfur atom. Due to the larger size of the S atom in comparison to O, enzymes cannot cleave these oligonucleotides. G-rich PS oligonucleotides were found to have negligible effect on antiparallel triplex stability (Paramasivam, Cogoi et al. 2008).

Peptide nucleic acid (PNA) is another backbone modification tested on triplex-forming oligonucleotides. Due to the uncharged backbone, PNA TFOs do not suffer from electrostatic repulsion between strands; however, G-rich PNA TFOs were able to form G-quadruplexes with exceptional stability, thus triplex formation was inhibited (Englund, Xu et al. 2006).

Locked nucleic acids (LNA) refer to a backbone modification formed by an extra bridge connecting the 2′ oxygen with the 4′ carbon of the ribose sugar (Kaur, Arora et al. 2006). LNA modifications show restricted flexibility by forcing the ribose ring to adopt a defined conformation and are one of the most promising modifications. LNA/DNA hybrid GT-TFOs were shown to form stable triplexes in comparison with unmodified or CT-TFOs. The same group has demonstrated an inhibition of firefly luciferase expression with TFO treatment. The

inhibition was enhanced with addition of psoralen (**Ps**, Figure 1.11) on the 5′ terminal of the TFOs (Brunet, Alberti et al. 2005). Another study has been carried out to demonstrate the ability of LNA/DNA TFOs for directed mutagenesis. 14-mer GA LNA/DNA TFO was modified with L-tryptophan and cyanine dye TO1 at 3′ and 5′ termini, respectively, and used to cleave the DNA by production of reactive oxygen species (Biton, Ezra et al. 2010).

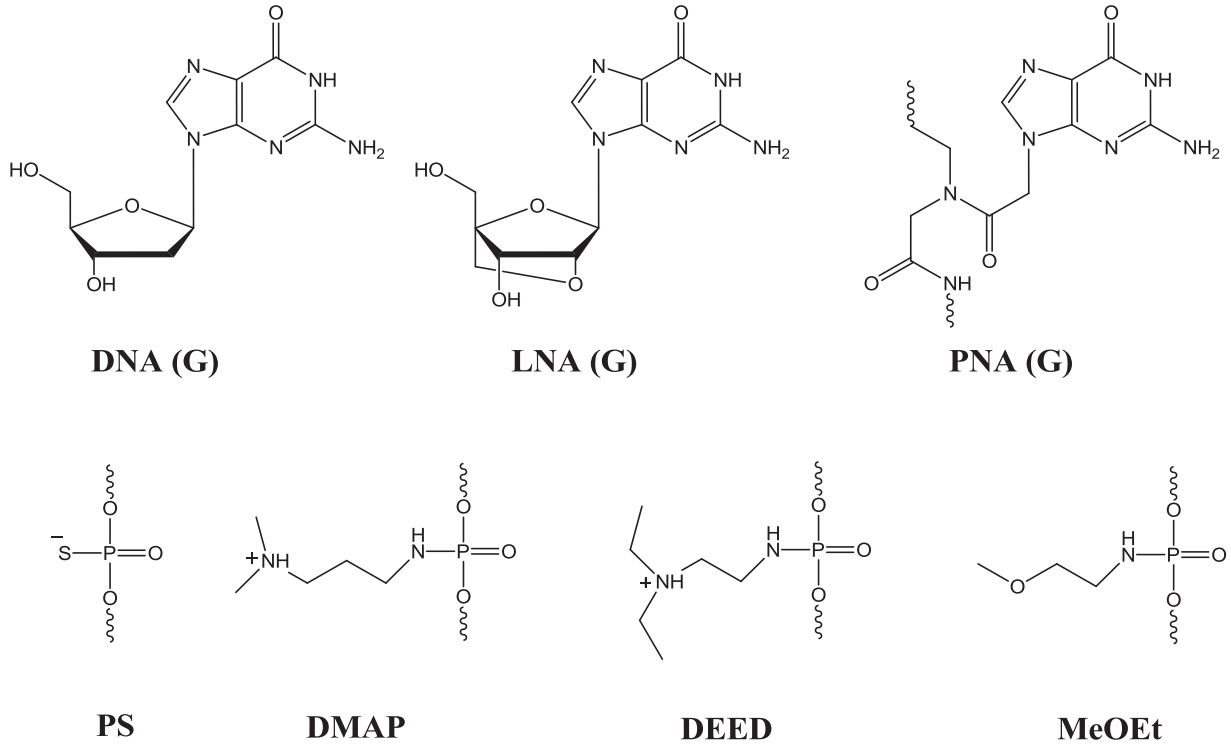


Figure 1.10 Unmodified deoxyribonucleic acid (DNA) and backbone modifications: locked nucleic acid (LNA), peptide nucleic acid (PNA), phosphothiate (PS), N,N-Dimethylaminopropylamine (DMAP), N,N-diethylethylenediamine (DEED) and methoxyethylamine (MeOEt).

1.3.2 Base Modifications

Several guanine modifications were designed to control their ability to form hydrogen bonds. The structures mediated by guanine residues can be controlled by changing the hydrogen bond patterns. Most studies that focused on changing hydrogen bond capabilities of guanines aimed to destabilise G-quadruplex formation. Guanine substitutions, 6-thioguanine (**6SG**), 7-deazaguanine (**^{7da}G**) and 6-thio-7-deazaguanine (**^{7da}6SG**) were designed by replacing essential atoms participating in H bond formation in G-quartets (Figure 1.11). **^{7da}G** and **^{7da}6SG** have decreased

stability of G-quadruplexes, as well as antiparallel triplexes (Milligan, Krawczyk et al. 1993; Durland, Rao et al. 1995; Olivas and Maher 1995). 7-Chloro-7-deazaguanine (^{7Clda}**G**) was found to resist K^+ mediated inhibition and to form stable triplexes when used in TFOs (Aubert, Perrouault et al. 2001). Alternatively, 7-deazaxanthines (^{7d}**X**) has also shown to stabilize triplexes formed when used to substitute thymidines of GT-TFOs in comparison to unmodified TFO (Milligan, Krawczyk et al. 1993).

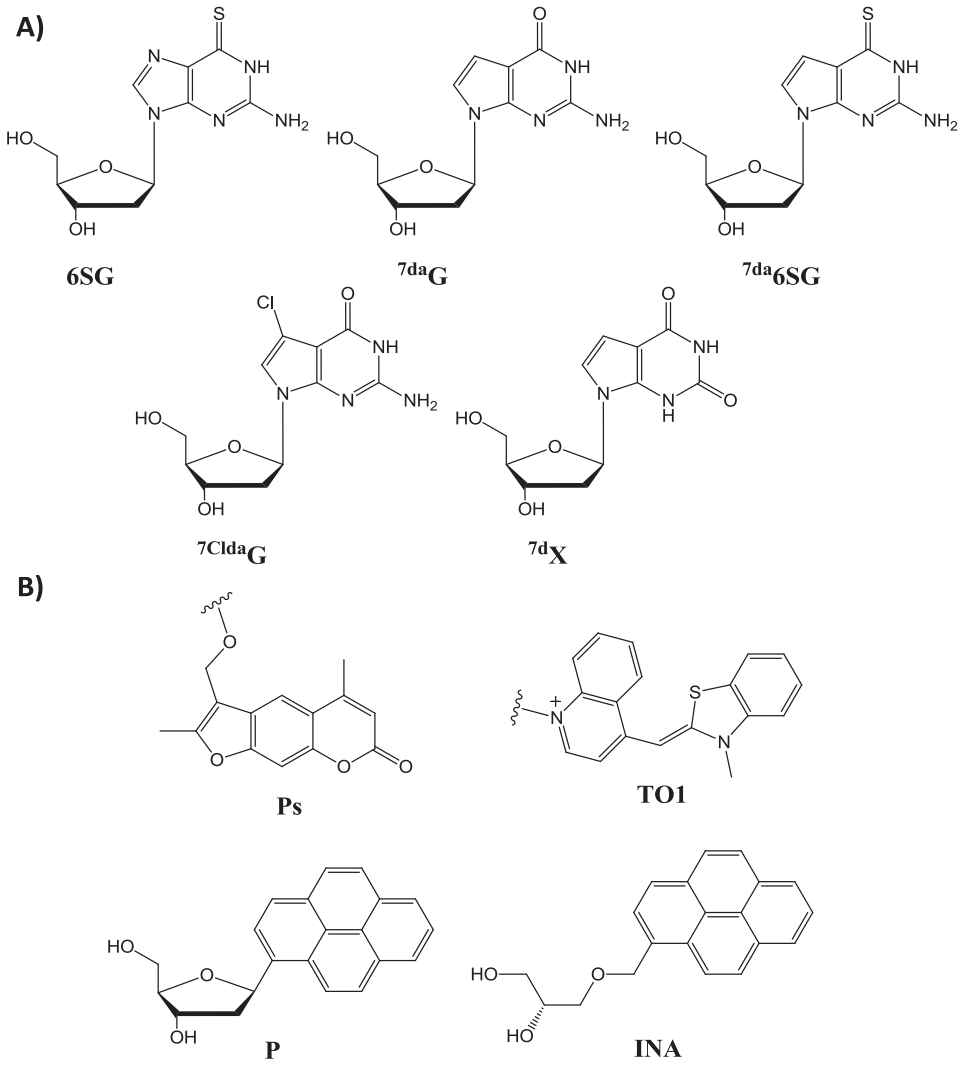


Figure 1.11 Base modifications (A) and intercalators (B) for triplex-forming oligonucleotides: 6-thioguanine (**6SG**), 7-deazaguanine (^{7da}**G**) and 6-thio-7-deazaguanine (^{7da}**6SG**), 7-Chloro-7-deazaguanine (^{7Clda}**G**), 7-deazaxanthines (^{7d}**X**), psoralen (**Ps**), thiazole orange (**TO1**), pyrene (**P**) and intercalating nucleic acid (INA).

1.3.3 Intercalators

Another group of modifications features intercalators. Intercalators are polycyclic aromatic hydrocarbons which can interact with nucleobases. The contribution of π–π stacking between nucleobases and intercalators increases the stability of DNA structures. Intercalation of a large chromophore unwinds the helix of a DNA complex by pushing nucleobases apart (Figure 1.12, path a). Alternatively such moieties may interact by stacking at the ends of duplexes, triplexes, or quadruplexes, where the hydrophobic surface of nucleic bases is exposed. They can participate in additional π–π interactions and also protect the H bonds from interaction with aqueous solution (Figure 1.12, path b). Traditionally, intercalators are considered as free ligands; however, covalent attachment of intercalators to nucleic acids is also used to enhance triplex stability.

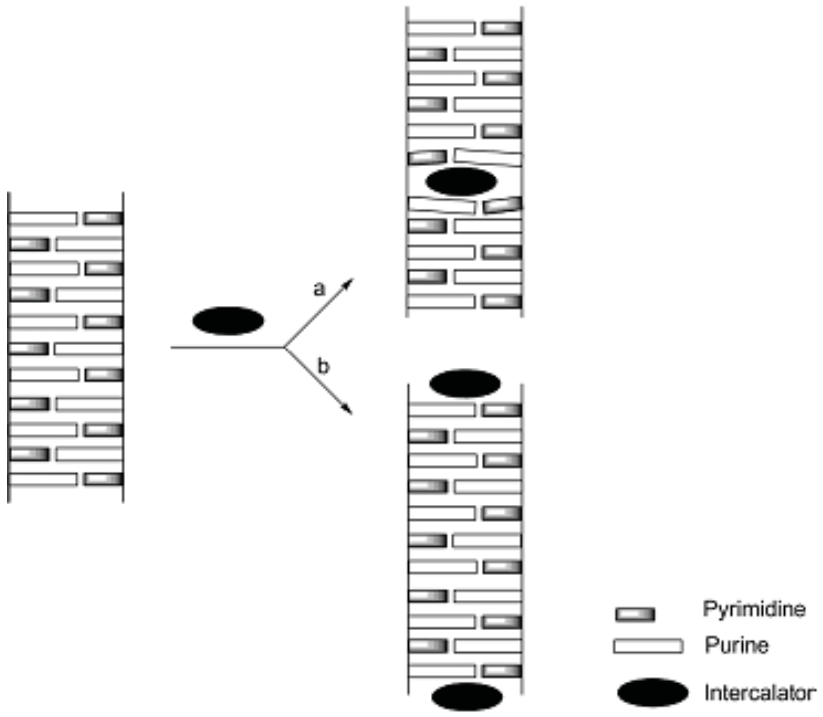


Figure 1.12 Different ways of intercalation: a) Intercalator is placed between adjacent base pairs, and b) intercalator is stacking at the ends of the duplex.

Unfortunately, the studies on TFO conjugates of polycyclic hydrocarbons focusing on antiparallel triplexes have been considerably less intense than for parallel triplexes. Psoralen (**Ps**) is a popular 5′ modification due to its photoreactivity (Figure 1.11). **Ps**-conjugated antiparallel TFOs were used to deliver DNA damage to a target gene *via* cross-linking (Vasquez 2010).

Successful regulatory effects of **Ps**-conjugated TFOs were detected by several other groups (Macaulay, Bates et al. 1995; Miller, Bi et al. 1996; Song, Intody et al. 2004). Alternatively, a pyrene (**P**) modification was used in TFOs for DNA damage (Benfield, Macleod et al. 2008).

1.3.3.1 Twisted Intercalating Nucleic Acid (TINA)

Within the last decade, a novel modification was derived from an intercalating nucleic acid (INA, Figure 1.11, B). INA was used as a bulge insertion in order to stabilise duplex DNA. INA also decreased stability in triplexes by 5 °C (Christensen, Wamberg et al. 2004). Inspired by INA's ability to stabilise duplexes a new intercalator was designed in order to adapt INA to triplexes. Pyrene containing (*R*)-1-*O*-(4-(1-pyrenylethynyl)phenylmethyl)glycerol residue, also known as twisted intercalating nucleic acid (TINA), was developed (Figure 1.1) (Filichev and Pedersen 2005). Early studies have shown that while oligonucleotides bearing this modification as a bulge have decreased duplex stability, each TINA modification has increased the thermal stability of parallel triplexes by ~19 °C. With this property TINA is the only known intercalating nucleic acid with discrimination of Hoogsteen-type triplexes over Watson-Crick-type duplexes (Filichev, Gaber et al. 2006; Geci, Filichev et al. 2006)

It was assumed that upon triplex formation the pyrene moiety was positioned in the dsDNA part of the triplex while the phenyl group coaxially stacked with nucleobases of the TFO (Figure 1.13). Another interesting property of TINA is the twisting arm carrying the pyrene. The triple bond allows the pyrene groups to fit properly in the duplex. The 1-pyrenyl derivative of TINA was found to be the best one among the tested modifications (acridine, naphthalene, *m*-phenylethynyl and 4-biphenyl) for binding to the Hoogsteen-type triplexes. Contrary to TINA, pyrenemethyl glycerol (INA) (Figure 1.11) led to destabilization of parallel triplexes when used as a bulge insertion (Christensen, Wamberg et al. 2004).

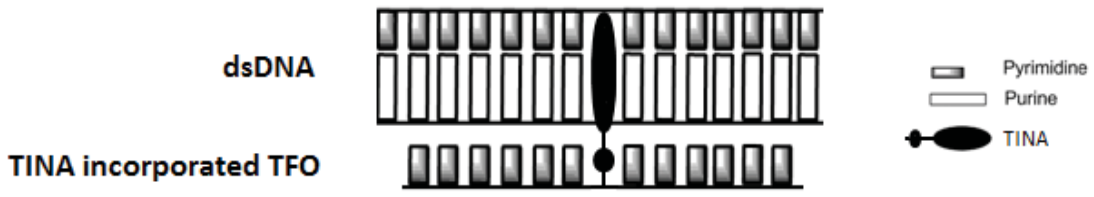


Figure 1.13 Parallel triple helical DNA complex formed by TINA incorporated TFO.

Even though TINA insertions have stabilized parallel triplexes, certain drawbacks of these structures persist, such as pH sensitivity. Recently, it was shown that TINA incorporated homopurine (GA) sequences can be used to form antiparallel triplexes in the presence of complementary dsDNA from the human and murine *KRAS* promoter (Figure 1.14) (Paramasivam, Cogoi et al. 2008). Twenty bases long GA-TFOs containing two or three insertions of TINA were shown to recognise their target successfully in the presence of 50 mM K^+ and pH 7.2 with dissociation constants as low as 130 nM. At the same time TINA insertions between continuous guanine tracts showed a high tendency to negate TFO's self-association in the form of G-quadruplex formation. Antigene activity of these probes was also confirmed when *KRAS* promoter driven chloramphenicol acetyltransferase expression was inhibited by 50 % when treated with complementary TINA-TFOs in comparison to mismatching TINA-TFOs.

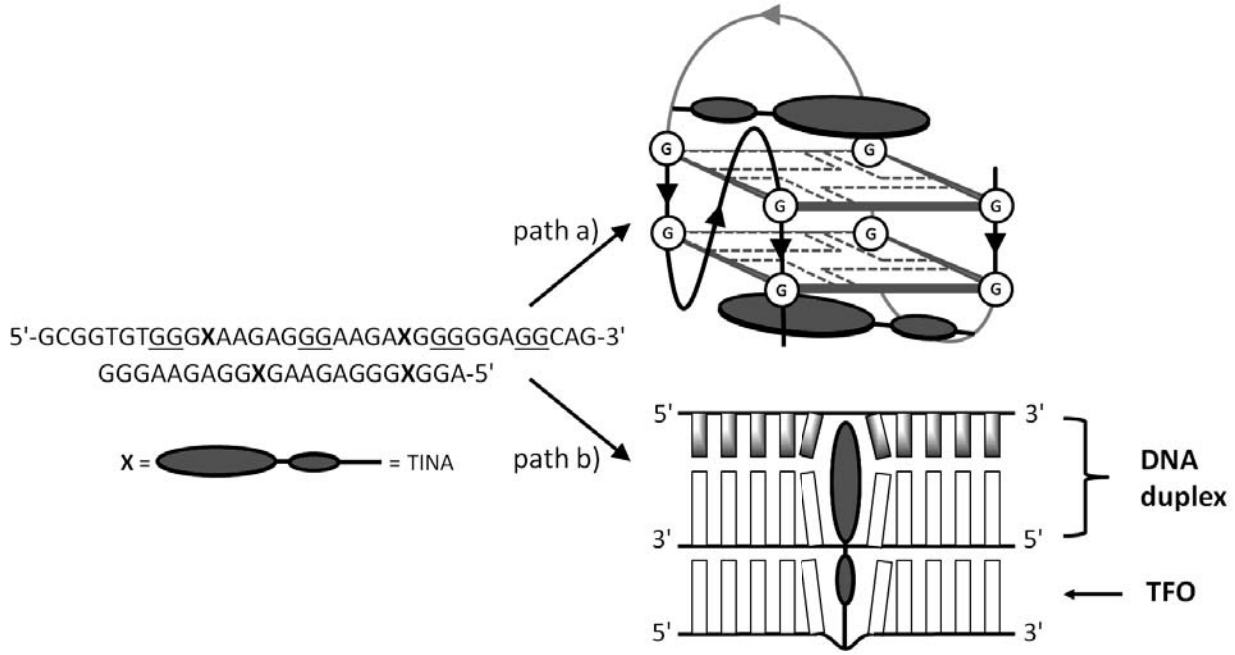


Figure 1.14 Representation of the TINA molecule in the design of G-quadruplexes (path a) and antiparallel DNA triplexes (path b).

On the other hand, when TINA insertions were located outside G-tracts, oligonucleotides formed G-quadruplexes with higher thermal stability in comparison to unmodified oligonucleotides (Figure 1.14). Other studies have shown significant increases in the stability of these structures when TINA modifications were located at the top or bottom of the G-tetrad stacks. It was

demonstrated that both intermolecular and intramolecular G-quadruplexes can be stabilised using TINAs.

Complementary to the antiparallel TINA-TFO study, the G-quadruplex forming motif from human *KRAS* promoter was studied. TINAs positioned at the top and the bottom of the G-tetrad stacks increased the thermal stability of the G-quadruplex by 32 °C (Cogoi, Paramasivam et al. 2009). At the same time, 5′-dTGGGAG and 5′-dGTGGTGGGTGGGTGGGT (T30177) sequences, known to inhibit HIV-1 replication on cell cultures, increased inhibition efficiency at 1.0 μM by up to 87 % from 10 % compared to unmodified T30177 (Pedersen, Nielsen et al. 2011). These studies indicate that the effect of TINA is strongly dependent on its position in the sequence (Figure 1.14). Unfortunately, studies focusing on this promising subject are limited and systematic work is necessary to provide further understanding of the structural effects of TINA incorporation into G-rich oligonucleotides.

1.4 Challenges

The developlement of nucleic acid based molecules is a popular research area of chemistry. Especially the development of DNA-binding probes has an ever-growing demand of molecular biology applications. DNA-binding probes must have three main features; ability to bind to their DNA targets strongly, ability to bind with high sequence specificity, ability to be detected using commercially available methods *e.q.* fluorescence. The improvement of these properties is possible using chemical modifications.

One of the common techniques in cellular biology that requires the use of DNA-binding probes is fluorescence *in situ* hybridisation (FISH). This technique, which uses single-stranded ONs labelled with non-selective fluorescent molecules, requires fixation of the cells, removal of the proteins from the chromatin, denaturation of dsDNA and several washing steps. In an attempt to improve FISH, triplex-forming oligonucleotides have been used as DNA-binding probes to visualise chromosomes under non-denaturing conditions. For instance, triplex formation has been observed using TFOs and triplex-specific antibodies (Ohno, Fukagawa et al. 2002) and this method was described as N-FISH (non-denaturing FISH). The advantages of visualising chromatin without denaturation have been explored in COMBO-FISH which relies on the combinatorial library of oligonucteotides (COMBO) designed to form triplexes on a gene of

interest (Wu, Gaddis et al. 2007). Otherwise disruptive in standard FISH, harsh treatments of the chromatin were eliminated in COMBO-FISH and the Abelson murine leukaemia (ABL) gene and the breakpoint cluster region (BCR) were successfully visualised in fixed cells (Schwarz-Finsterle, Stein et al. 2007). Several obstacles prevent gene visualisation in live cells using DNA triplex technology. Both N-FISH and COMBO-FISH were performed in acidic pH to insure formation of parallel triplexes. Washing steps were required to remove unbound probes and reduce a background noise that comes from a constantly emitting dye. Getting rid of the washing step will eliminate the neccessity of cell fixation. Non-discriminative binding of commercial FISH probes to DNA and RNA requires the elimination of RNA molecules which is usually done by pre-treatment of cells with RNases. This means that chemistry has a lot to offer in the development of gene visualisation methods *in vivo*. We believe that this PhD thesis made an important progress in DNA triplex technology for chromosome painting in live cells.

1.5 Thesis Outline

This thesis discusses the development of artificial DNA modifications in order to provide a better control of G-quadruplex and triplex formation.

Chapter 3 describes the synthesis of twisted intercalating nucleic acids (TINAs) and their incorporation into antiparallel triplex-forming oligonucleotides (TFOs). The ability of TINA monomers to disrupt self-aggregation of a variety of G-rich TFOs and to form antiparallel triplexes with their targets, which become more stable than parallel triplexes will be demonstrated. The proposed approach and established rules for the design of TINA-TFOs for the formation of antiparallel triplexes can be widely applied and should advance the applicability of DNA triplex technology in molecular biology, biochemistry and biotechnology. This work was published in *ChemBioChem* (Doluca, O.; Boutorine, A. S.; Filichev, V. V., *ChemBioChem*, 2011, 12 (15), 2365-2374).

The following chapter will focus on investigating the effect of TINA insertion into G-quadruplex forming sequence dTG_4T. This is the first systematic study about TINA's influence on tetramolecular G-quadruplex structure that improves our ability to control G-quadruplex topology. A manuscript discussing this work is under preparation for publication.

In Chapter 5, the conjugation of the TINA oligonucleotides with minor groove binders (MGBs), amino acid-based DNA-binding agents, is discussed in order to improve the affinity of TFOs and expand the length of target sequences.

The Chapter 6 will focus on the improvement of the fluorescent capabilities of TINA-TFOs for visualisation purposes. Pyrene-containing TINA monomers are fluorescently active when excited at 373 nm. However, most of the commercial lasers operate at wavelengths higher than 460 nm. For this reason the fluorescent properties of TINA cannot be exploited effectively in live cells. The most obvious solution for this obstacle is using TINA monomers in combination with another modification with the desired fluorescent properties. This led to the development of a new class of nucleic acids called, assembly-dependent fluorescence enhancing nucleic acids (AFENA). In this chapter AFENA's fluorescent properties and ability to be used together with the TINA monomer in G-rich TFOs for antiparallel triplex formation will be investigated. A patent on AFENA and a manuscript discussing this work is under preparation.

Future directions and conclusions are discussed in Chapter 7 and experimental procedures are included in Chapter 8.

Chapter 2. Methods for Studying DNA Structures

2.1 Introduction

There is a large number of biophysical techniques available to study biomolecules. In this section we will briefly mention the most common techniques with special emphasis on their role in the study of DNA structures.

2.2 Gel Electrophoresis

Gel electrophoresis is a technique that relies on the migration of biomolecules in an electric field. Under normal conditions the phosphate groups of DNA carry a negative charge. This results in migration of DNA molecules in a gel matrix on application of an electric field. The speed of the migration is dependent on the overall charge, the size of DNA molecules, the topology of the strands and the size of pores of the gel. As a net result of these factors, bulkier DNA molecules migrate slower in comparison to smaller ones. This results in formation of bands on the gel that can be detected. Visualisation of these gels is performed either by staining of the DNA molecules using dyes such as ethidium bromide (EtBr) and Stains-All® or using UV-shadowing which relies on UV absorbance of DNA molecules. Also, visualisation of radioactive or fluorescently labelled DNA molecules can be achieved using screen exposure or laser, respectively.

The pore size can be controlled using different polymers in various concentrations. Agarose gel is a common solution for DNA molecules with sizes bigger than 150 base-pairs. On the other hand, polyacrylamide is preferred when the length of the DNA is relatively shorter. Polyacrylamide gel electrophoresis (PAGE) can be performed under denaturing or non-denaturing (native) conditions.

2.2.1 Denaturing Gel Electrophoresis

Denaturing gel electrophoresis is a method that helps to identify the size or purity of the DNA molecules in a sample. Generally a preliminary treatment is required. DNA samples are heated up to 90 °C and incubated in concentrated urea or formamide solution in order to denature any complex and convert it to ssDNA. Denaturing conditions are maintained throughout the electrophoresis by addition of denaturing agents into the gel. This method is also used to determine the purity of a sample because each biomolecule with different molecular size appear on the gel as a separate band. However, biomolecules with similar molecular sizes may be indistinguishable using denaturing gel electrophoresis.

2.2.2 Non-denaturing Gel Electrophoresis

Non-denaturing gel electrophoresis is used to collect information about the topology of the DNA complex. For this reason, the samples are needed to be preserved from denaturing conditions such as overheating that may occur during electrophoresis. Usually, the electrophoresis buffer conditions are chosen to be similar to native conditions, such as ion concentrations.

The level of migration in non-denaturing conditions depends on the bulkiness of the complex or molecularity of DNA molecules. The molecularity refers to the order of association or number of biomolecules assembled to form the complex.

2.2.3 Determination of Dissociation Constant (K_d)

Non-denaturing gel electrophoresis is a very useful tool that permits the measurement of binding affinity between two molecules. The affinity, expressed as the dissociation constant (K_d), is calculated from the relative amounts at equilibrium of bound versus unbound ligands at different concentrations.

$$AB \rightleftarrows A + B \qquad K_d = \frac{[A]\times[B]}{[AB]} \qquad \text{(Formula 1)}$$

In case of DNA complexes, a set of samples needs to be prepared including a target DNA at a specific concentration and a ligand DNA at various concentrations. After equilibrium is reached

between bound and unbound ligand DNA, a non-denaturing gel electrophoresis is performed. Since folded structures would have a higher size and lower mobility in comparison to single-stranded DNA, the ratio between folded and unfolded structures can be easily calculated after visualisation of the gel. In most cases the target DNA is labelled by fluorescent molecules or radioactive elements for easy visualisation. The data representing the percentage of the conversion, relative to the total fluorescence or radioactivity in various ligand concentrations are interpreted to calculate the dissociation constant (K_d) using the following formula, where a value of the ligand concentration equals to K_d when half of the target compound is involved in the complex formation:

$$\log(S_0 - D_0 x) = \log\left(\frac{x}{(1-x)}\right) + \log(K_d) \qquad \text{(Formula 2)}$$

When D_0 and S_0 are initial concentrations of the target and the probe, respectively, and x is the proportion of conversion, $\log(K_d)$ equals to $\log(S_0 - D_0 x)$ where $\log[x/(1-x)]$ equals to zero. Computer software such as ImageQuant, Microsoft Excel or KaleidaGraph can be used to facilitate the calculations (Boutorine and Escude 2007). The dissociation constant (K_d) obtained from this formula is a representation of the affinity of the ligand DNA towards the target DNA. It is important to note that dissociation constants are inversely proportional to the affinity – a lower dissociation constant indicates higher affinity.

2.3 UV-Vis Spectroscopy

DNA molecules are known to absorb UV light with peak absorption around 260 nm. This feature is exploited in UV-Vis spectroscopy. This technique provides information about the concentration of DNA, purity of DNA versus RNA and thermal stability of the secondary structure.

2.3.1 Determination of Concentration

The absorbance of DNA is dependent on its extinction coefficient (ε). This value indicates the strength of the absorption at a given wavelength. 2′-Deoxyoligonucleotides have maximum absorption at 260 nm. The extinction coefficients of DNA molecules at this wavelength (ε_{260}) can

be determined by a number of mathematical methods and depend on the number of nucleotides and their neighbouring bases. Concentration is determined using following formula of Beer-Lambert law where Abs refers to absorbance, *c* refers to the concentration and *l* refers to the pathlength of cell.

$$c = \frac{\mathrm{Abs}_{260}}{\varepsilon_{260} \times l} \qquad \text{(Formula 3)}$$

2.3.2 Determination of Melting Temperature

Comparison of the change in UV absorbance versus temperature (melting curve) is a common method to study the stability of a DNA complex. The middle point of the transition is refered as melting temperature (T_m, Figure 2.1). At this point half of the oligonucleotides exist as ssDNA and the other half as part of the secondary structure. In this report we determined the T_m using the first derivative of the melting curve. The point where the rate of change in absorbance is maximum is chosen as the point of midtransition. In some cases more than one transition may occur, such as DNA triplexes. T_m is detected separately for each transition. An alternative method is used of drawing baselines at the top and bottom of the melting curve. A median line is placed between these lines. The point where this line intersects the curve is chosen as T_m. This method can be preferred for curves that do not have flat baselines; however, it is not suitable for melting curves with multiple transitions.

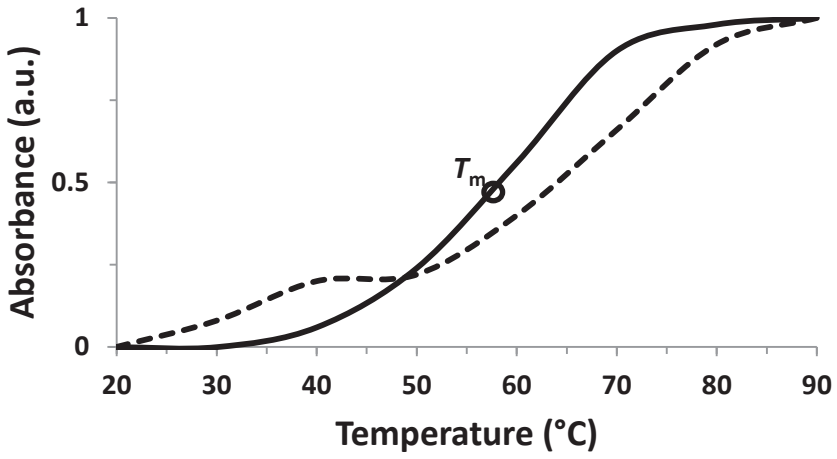


Figure 2.1 DNA melting curve examples. Solid line and dashed line represent melting curves with a single and double transition, respectively. Circle represents the point of midtransition.

Hysteresis is a common issue during melting/annealing studies. Briefly, hysteresis refers to the discrepancy in a value under the same conditions but with different previous conditions (*e.q.* temperature). This may cause melting and annealing experiments to give different values of T_m for oligonucleotides which form multimeric DNA complexes. In such cases, the midpoint of the transition does not accurately indicate the melting temperature (T_m). For this reason in multimeric DNA complexes that show hysteresis, such as G-quadruplexes, this point is refered as midtransition point ($T_{½}$). The hysteresis depends on several factors, such as the rate of heating/cooling, sample concentration, and the rate of association or dissociation. Higher sample concentrations or slower heating/cooling rates are preferred to overcome this issue.

2.3.3 Thermal Difference Spectra (TDS)

The difference between UV-Vis absorption spectrum of a DNA complex and that of its ssDNA form is defined as thermal difference spectrum (TDS). The TDS is usually obtained simultaneously with melting profiles. The TDS is obtained by subtraction of the UV-Vis absorption spectrum of folded DNA a complex at low temperatures from UV-Vis absorption spectra following complete melting at an elevated temperature. The TDS is shown to be specific to secondary structure and can be used for determination of the topology of a DNA complex. The B-form of dsDNA is shown to have positive TDS band between 259 and 276 nm depending on GC content (Mergny, Li et al. 2005). The TDS of parallel triplexes show a positive peak between 245 and 270 nm and may also show a weaker negative 295 nm band (Table 2.1).

Table 2.1 Characteristic wavelengths of TDS and CD spectra for different nucleic acid structures.

	Characteristic TDS signature[a] (nm)	Characteristic CD signature[a] (nm)
B-type duplex	260-280(+)	245(-), 260-290(+)
Parallel/antiparallel triplex	240-270(+), 295(-)[b]	210-215(-), 245(-), 260-290(+)
Parallel G-quadruplex	243(+), 273(+)	240(-), 260(+)
Antiparallel G-quadruplex	243(+), 273(+)	245(+), 265(-) 295(+)

[a] (+) and (-) refer to positive and negative peaks, respectively. [b] in some cases.

The TDS of antiparallel triplexes share a common peak at around 270 nm (Kolganova, Shchyolkina et al. 2012). In the case of G-quadruplexes, two positive bands at 243 nm and 273 nm may be detected. The strength of 273 nm peak and the depth of the valley between 243 and 273 nm depends on the sequence. The relative strength of these peaks also depends on the cation present in the solution. An additional negative band at 295 nm may also detected; however, this band is not specific to G-quadruplex formation (Mergny, De Cian et al. 2005).

2.4 Fluorescence Spectroscopy

Fluorescence spectroscopy is a common tool to analyse samples excited by irradiation at a certain wavelength. This technique requires the presence of a fluorescent moiety attached to the DNA sequence because unmodified oligonucleotides show negligible fluorescence. After being excited, the fluorescent molecule relaxes back to the ground state from an excited state by emitting radiation at a longer wavelength. The radiation is then detected over a range of wavelengths. The excitation spectrum is measured by recording the light emission at a particular wavelength resulting from excitation over a range of wavelengths. Two types of fluorescence spectra are usually recorded for fluorescent systems. The emission spectrum is obtained by recording the light emission at a range of wavelengths resulting from excitation from a single wavelength.

Visualisation of biological samples and Förster resonance energy transfer (FRET) are common applications of fluorescence. Förster resonance energy transfer refers to a non-radiative energy transfer between two fluorophores in close proximity. FRET occurs as the excited-state “donor” fluorophore transfers most of its energy to the ground-state “acceptor” fluorophore in close proximity (<10 nm); however, this process requires that emission spectra of the “donor” fluorophore overlap the excitation spectra of the “acceptor” molecule. FRET is a common method for detection of interactions between molecules such as DNA and proteins.

Fluorescence quantum yield (Φ_F) is a parameter used to describe the efficiency of a fluorescent system and is defined as the ratio of the number of photons emitted to the number of photons absorbed (Lakowicz 1999). One of the techniques for measuring the fluorescence quantum yield uses sophisticated fluorescence spectrophotometers with an integrated sphere accessory (also known as the Ulbricht sphere). The integrated sphere is an optical component consisting of a

hollow spherical cavity with its interior covered with a diffuse white reflective coating and it is possible to measure the total number of photons absorbed and emitted by the sample which is positioned in the centre of the sphere using this accessory (Porres, Holland et al. 2006). It must also be noted that the error for fluorescence quantum yield measurements are usually around 10 % of the value; however, in reality much more uncertainty should be anticipated due to experimental errors (Fery-Forgues and Lavabre 1999). For better quantum yield calculation, the testing of a set of samples with a range of concentrations is advised.

2.5 Circular Dichroism (CD) Spectroscopy

Circular dichroism (CD) spectroscopy is a common tool for studying DNA structures. It has been shown that the difference between absorption of right and left circularly polarised light by DNA is dependent on secondary structure (Baase and Johnson 1979). This is the result of the interactions between molecules with chiral centers and electromagnetic radiation of circularly polarised light. The absorption difference of left-handed and right-handed circularly polarised light is collected for a range of wavelengths and transformed into a CD spectrum, where positive and negative peaks are formed. The spectral signature is specific to the secondary structure of the DNA complex. A B-form dsDNA shows a positive band between 260 and 280 nm and a negative band around 245 nm (Table 2.1) (Kypr, Kejnovska et al. 2009). Triplex formation can be detected through a negative band between 210 and 215 nm (Gray, Hung et al. 1995). Also a negative peak at around 240 nm and a positive broad band between 260 and 280 nm is shown by triplexes (Kolganova, Shchyolkina et al. 2012). Determination of G-quadruplex structures using CD spectroscopy is complicated due to the variety of potential topologies. The CD spectra of G-quadruplexes have shown that there is a correlation between the CD spectrum and the alignment of the G-tetrads that form the G-quadruplex. When all four strands are parallel to each other, the CD spectrum is dominated by a positive band around 260 nm and a negative band at 240 nm. When two of these strands are in opposite orientation, then the CD spectrum shows a positive band at ~295 nm and a negative band ~ 265 nm (Masiero, Trotta et al. 2010). It should be remembered that even though CD spectra are commonly used to distinguish the strand orientation of G-quadruplex structures, a given spectrum is mainly controlled by the alignment of guanines and not by the orientation of the strands.

2.6 Workflow for Studying Nucleic Acid Structures

The methods described in previous section are required to be used in combination in order to obtain adequate data about DNA structures. The following flowchart is designed to give a general idea on the workflow in the field of structural analysis of nucleic acids.

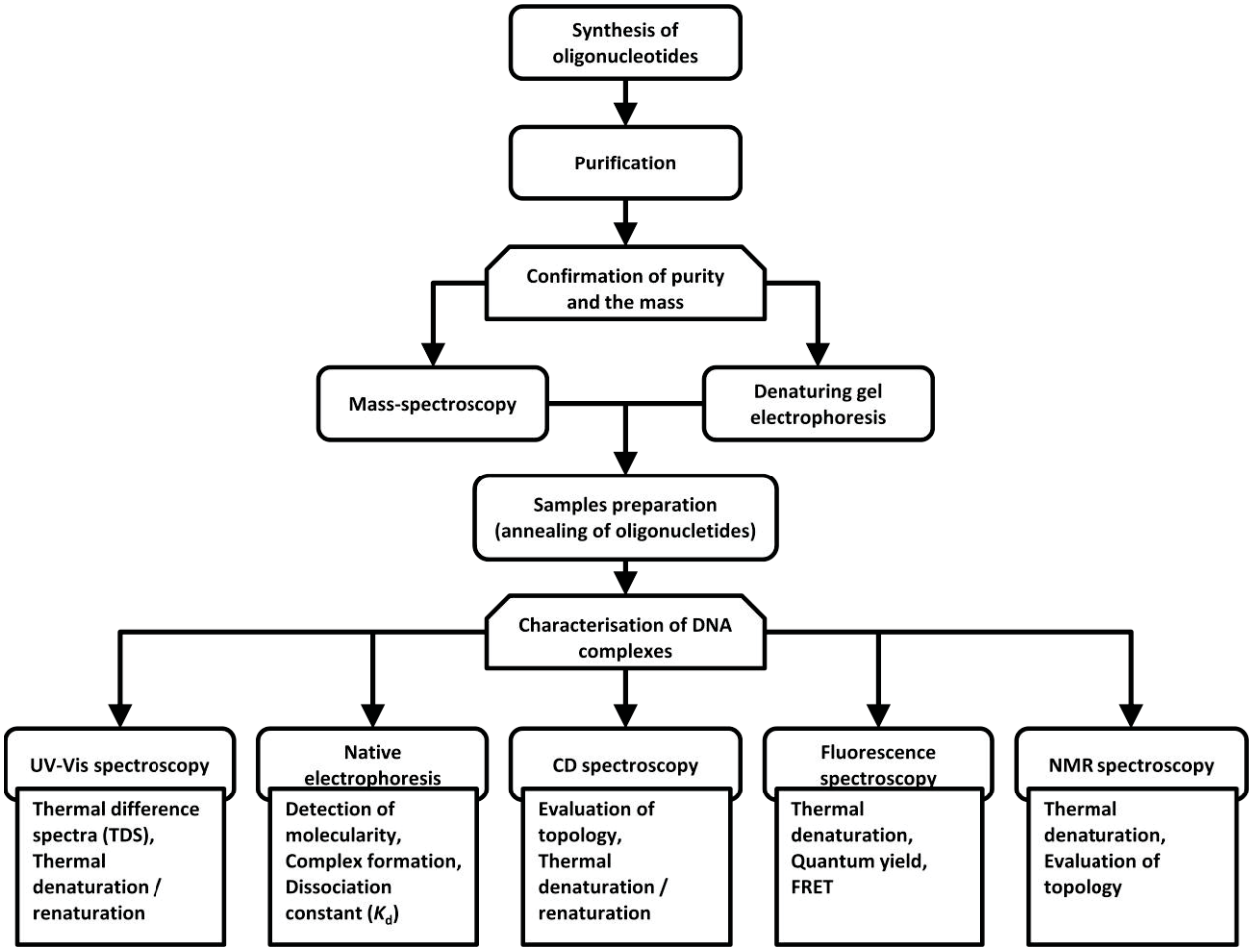


Figure 2.2 Workflow for studies of nucleic acid structures.

Chapter 3. Synthesis and Biophysical Studies of TINA-TFOs

3.1 Introduction

For over 20 years scientists have been working on agents that would recognize DNA in its native form and environment because the ability to target specific DNA sequences inside the cell has far-reaching implications in biotechnology and medicine, such as gene regulation, visualisation of chromosomes and, most of all, cancer treatment (Barton 1988; Rosenthal 1995; Lavery 2005). Targeting DNA is more advantageous in some ways than targeting RNAs or proteins. DNA has fewer copies in the cell and it is easier to repress a gene by inhibiting its transcription at the DNA level, rather than inhibiting translation of many RNAs or inhibiting corresponding proteins after translation (Tonelli, Purgato et al. 2005). Today a number of DNA-binding agents are available, thanks to efforts of scientists from various disciplines; however, several obstacles still prevent their practical applications.

Triplex-forming oligonucleotides (TFOs) have been attracting a lot of attention due to their potential as dsDNA-binding agents. These molecules consist of oligonucleotides with the ability to form triple helical DNA (or triplex) with a target dsDNA. With the exception that TFOs can recognize only homopurine (GA) sequences, TFOs are not limited by recognition length. Moreover, recognition sequences can be easily modified without complex modelling, mutagenesis or selection methods.

To date, over 30 human and rodent genes, such as *BCR/ABL*, *HRAS*, *MYC*, *SRC*, have been identified to have high-affinity TFO-binding sequences. Moreover, it has been found with computational methods that 97.8 % of human genes have at least one potential sequence in the promoter or transcribed region of genes that can be targeted by TFOs with high affinity with a minimum length of 15 base pairs (Wu, Gaddis et al. 2007).

One of the promising TFO targets is a 16 bp homopurine sequence from HIV-1 proviral DNA, 5′-dAAAAGAAAAGGGGGGA. Previously psoralen-conjugated G-rich TFOs were used to target the sequence in a parallel fashion and successfully demonstrated gene regulation on infected cell lines (Giovannangeli, Diviacco et al. 1997). In similar studies, gene regulation on eukaryotic cells using triplex-forming oligonucleotides has also been demonstrated; however, no direct evidence was provided that this inhibition was due to triplex formation and not due to alternative pathways such as G-quadruplex formation (Orson, Thomas et al. 1991; Thomas, Faaland et al. 1995; Porumb, Gousset et al. 1996). Also it was established by several studies that G-rich TFOs designed for HIV-1 proviral DNA prefer to self-aggregate instead of forming a triplex structure (Noonberg, Francois et al. 1995; Noonberg, Francois et al. 1995; Arimondo, Barcelo et al. 1998). For instance, 5′-dAGGGGGGGTTTTGTTTT was one of these TFOs designed to bind HIV-1 proviral DNA in antiparallel orientation. This formed instead thermally stable G-quadruplex structures which could be destroyed only partially at 90 °C (Figure 3.1).

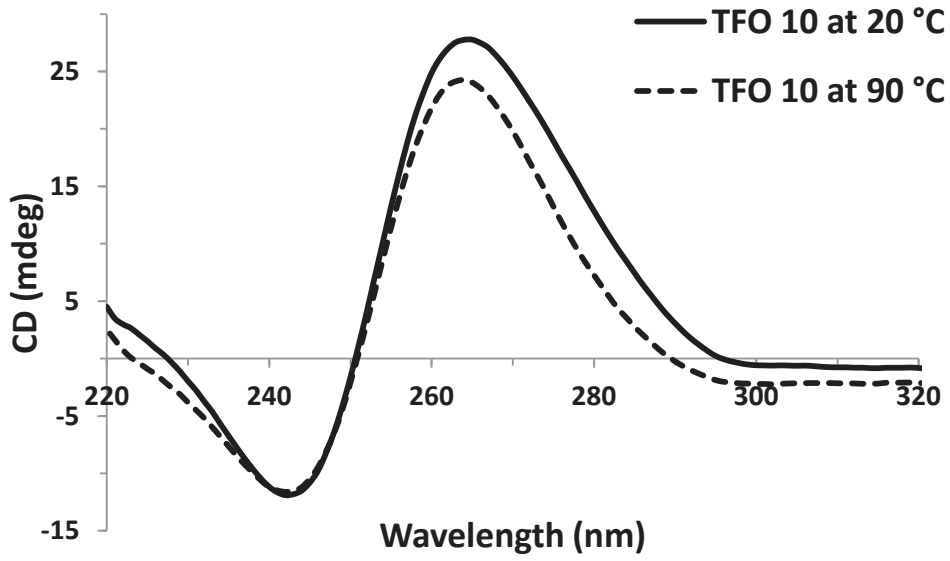


Figure 3.1 CD spectra of G-quadruplexes formed by 5′-dAGGGGGGGTTTTGTTTT (TFO 10, Table 3.1) sequence (10 μM) at 20 °C (solid line) and 90 °C (dashed line) in the presence of 10 mM Na cacodylate buffer, 100 mM NaCl and 10 mM $MgCl_2$ at pH 7.2.

A recent study has shown that TINA monomers can be used to discriminate formation of triplexes over G-quadruplexes (Paramasivam, Cogoi et al. 2008). G-rich sequences from the *KRAS* promoter, which are known to form G-quadruplexes, have been modified by TINA

monomers between continuous guanine tracts. As a result these sequences preferred to form triplexes instead of undergoing self-aggregation in the presence of 140 mM K^+. Inspired by this study, we have decided to investigate TINA-mediated triplex formation in HIV-1 proviral DNA, 5′-dAAAAGAAAAGGGGGGA, aiming to negate TFO self-aggregation by incorporating TINA in the middle of the 6 nucleotide long G-tract.

3.2 Chapter Outline

First, we describe the synthesis of TINA monomers and synthesis of TINA-TFOs by incorporation of TINA into oligonucleotides using automated DNA synthesis. Next, we investigate the properties of antiparallel triplexes formed between TINA-TFOs and HIV-1 proviral DNA. The tendency of TINA-TFOs to form self-aggregates was observed. A set of rules on how to optimise TINA-TFOs was compiled using these results. In the second part of this chapter we demonstrate the applicability of these rules to design TFOs targetting to polypurine sequences from *ABL* gene.

3.3 Synthesis of TINA Monomer

The TINA monomer was synthesised starting from [(*S*)-2,2-dimethyl-1,3-dioxolan-4-yl]methanol (**1**), and 4-iodobenzyl bromide (**2**) in the presence of KOH in dry toluene under Dean-Stark conditions. Compound **3** was deprotected by CF_3COOH in H_2O at 65 °C for three days. Because of the low yield and impurities detected by TLC, the deprotection step was repeated using Amberlite (H^+) resin instead of CF_3COOH. After deprotection, compound **4** was washed with MeOH from the resin and concentrated in vacuo (Figure 3.2).

A Sonogashira reaction was performed with 1-ethynyl pyrene (**5**) in the presence of $Pd(PPh_3)_4$, CuI and Et_3N in dry DMF. Solution was kept under argon for four days.

4,4′-Dimethoxytrityl (DMT) was introduced to compound **6** under argon in dry DCM overnight, using Et_3N (1 % v/v) as a base. After silica gel column purification (0-100 % EtOAc in hexane), pure compound **7** was obtained. 2-Cyanoethyl-*N*,*N*,*N*′,*N*′-tetra*iso*propyl-phosphordiamidite was reacted with compound **7** in the presence of di*iso*propylammonium tetrazolide under argon overnight in dry DCM. The final product, compound **8**, was purified using silica gel column

chromatography with an overall yield of 24.6 %. The product was confirmed by ^{1}H, ^{13}C and ^{31}P NMR before using in the DNA synthesis.

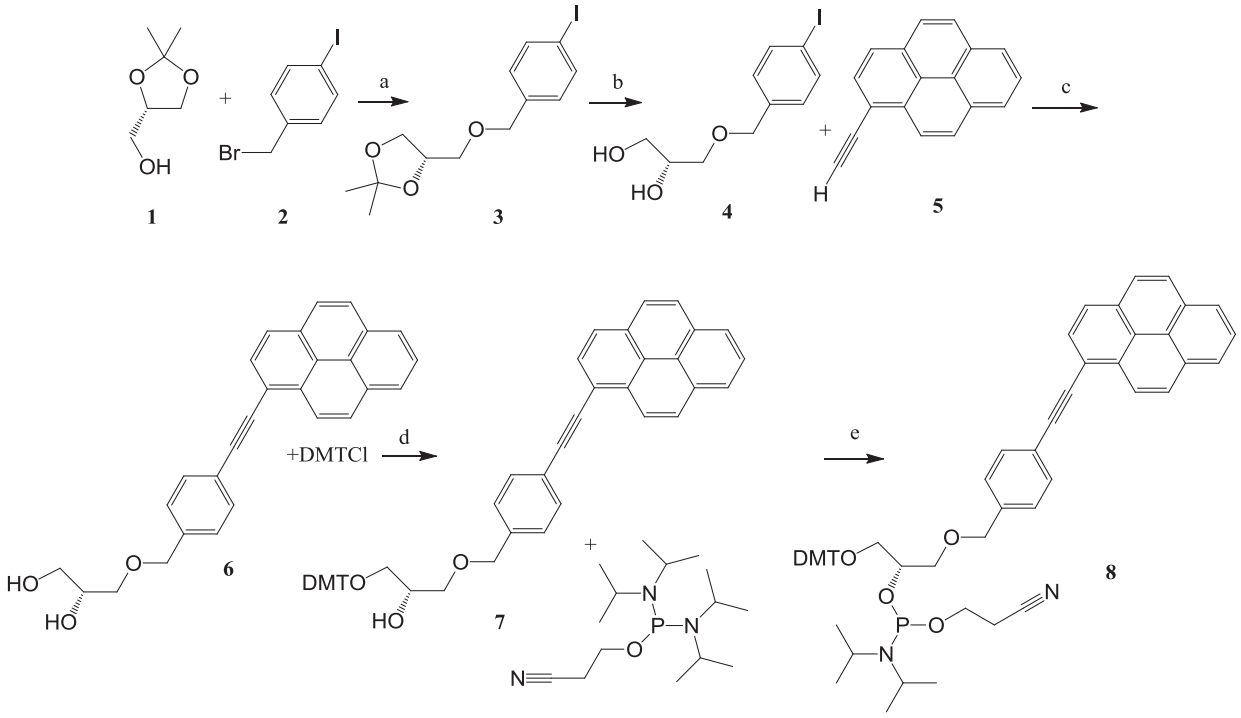


Figure 3.2 Synthesis of TINA monomer. Reagents and conditions: a) KOH, toluene, reflux. b) Amberlite (H^+) resin, H_2O, 65 °C. c) $Pd(PPh_3)_4$, CuI, DMF/Et_3N, argon at 20 °C. d) DCM/Et_3N, argon at 20 °C. e) DCM, di*iso*propylammonium tetrazolide, argon at 20 °C.

3.4 Design and Synthesis of Triplex-forming Oligonucleotides (TFOs)

TINA-incorporated TFOs were synthesised on a Mer-Made-4 automated DNA synthesiser under standard conditions with the exception of the TINA coupling. The TINA monomer was hand-coupled using ca. 10 mg of the phosphoramidite (compound **8**) in 750 µL of an activator (0.25 M 4,5-dicyanoimidazole (DCI) in dry acetonitrile) per coupling per column for 5-10 min. TFOs were synthesised in DMT-ON mode and were purified on reverse-phase C18 HPLC before detritylation. After evaporation, the DMT group was removed using 80 % aqueous AcOH at 20 °C for 20 min. After quenching using 3 M NaOAc, the TFOs were precipitated from EtOH.

Purity of TFOs was confirmed by denaturing 20 % PAGE and composition was confirmed using ESI mass-spectrometry.

Table 3.1 Sequences of TFOs synthesised or purchased.

Oligonucleotide	Sequence 5′-3′ [a]
TFO **1**	dAGG**X**GG**X**GGTTT**X**TGTTTT
TFO **2**	dAGG**X**GGG**X**GTTT**X**TGTTTT
TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT
TFO **4**	dAG**X**GG**X**GG**X**GTTT**X**TGTTTT
TFO **5**	dAG**X**GGGG**X**GTTT**X**TGTTTT
TFO **6**	dAGG**X**GGGG**X**TTT**X**TGTTTT
TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT
TFO **8**	dAG**X**GGG**X**GGAAA**X**AGAAAA
TFO **9**	dAG**X**GGG**X**GGTTT**X**TGT**X**TTT
TFO **10**	dAGGGGGGTTTTGTTTT
TFO **11**	dAGGGTGGGTTTTGTTTT
TFO **12**	dTTTTCT**X**TTTCCC**X**CCCT
TFO **13**	dTTTTCT**X**TTTGGG**X**GGGT
TFO **14**	dTTTTCTTTTCCCCCCT
TFO **15**	dTTTTGTTTTGGGGGGT
TFO **16**	dAAAACAAAACCCCCCT
TFO **17**	dAAAACAAAACCC
TFO **18**	dCAAAACCCCCCT
TFO **19**	dAGGGGGGAAAAGAAAA
TFO **20**	dTGGGGGGTTTTGTTTT
TFO **21**	dGGGTGGGTTTTGTTTT
TFO **22**	dGGGGGGTTTTGTTTT

[a] **X** represents TINA monomer.

TFOs were designed to target the 16 bp homopurine sequence 5′-dAAAAGAAAAGGGGGGA since it is an extensively studied sequence for triplex formation. TFOs **1** - **11** were designed as

antiparallel TFOs and TFOs **12 - 15** were designed for parallel triplex formation (Table 3.1). TFO **10** was known to form a highly stable G-quadruplex which could not be negated even at 90 °C (Figure 3.1).

The distance between the backbone of the TFO and the backbones of the duplex is significant in the design of TFOs. Ideally, the arrangement of nucleobases in neighbouring triplets should be isomorphous. By isomorphous we mean that the distances between the points where nucleobases are attached to their backbones should be identical; however, this distance changes depending on the nucleobase type of the TFO, so the backbone of the TFO fluctuates (Figure 1.6, A). Such fluctuations in the TFO backbone conformation might result in the inhibition of triplex formation. A single adenine preceding a G-tract is more preferable than thymidine alternative, because this gives a shorter distance between the third strand backbone and the duplex. On the other hand, it was found that a long T-tract is more preferable than an A-tract (Keppler, Read et al. 1999). TFO **8** was designed to show this preference in comparison with TFO **3**.

Table 3.2 Sequences of target duplexes **I**, **II** and **III**. Target **II** contains an additional A-T base-pair, was designed as target for unmodified TFO **11**. Target **III** was designed as a mismatching target.

Target duplex	Sequence
Target **I**	5′-dGCCACTTTTTAAAAGAAAAGGGGGGACTGG-3′
	fluorescein-3′-dCGGTGAAAAATTTTCTTTTCCCCCCTGACC-5′
Target **II** [a]	5′-dGCCACTTTTTAAAAGAAAAGGG**A**GGGACTGG-3′
	fluorescein-3′-dCGGTGAAAAATTTTCTTTTCCC**T**CCCTGACC-5′
Target **III** [a]	5′-dGCCACTTTTTAAAAG**C**AAAGG**T**GGGACTGG-3′
	fluorescein-3′-dCGGTGAAAAATTTTC**G**TTTCC**A**CCCTGACC-5′

[a] Bold nucleotides represent inserted or substituted bases in comparison with the Target **I**.

Three target duplexes, which include wild type (Target **I**) or mutated editions (Target **II** and **III**) of this sequence, were purchased from Integrated DNA Technology (USA). These targets were purchased labelled with fluorescein at 3′-end of the polypyrimidine strand (Table 3.2). Target **I** contains a wildtype HIV-1 sequence and is extended at both ends to promote faster and stable duplex formation. Target **II** was designed for TFO **11** in which an extra thymidine insertion in the G tract matches a T-A duplet of the target. Since specificity is another required property of DNA probes, Target **III** was designed to possess two mismatches for triplexes.

3.5 Properties of TINA-Incorporated Antiparallel Triplexes

3.5.1 Determination of Dissociation Constants of Triplexes (K_d)

Determination of the dissociation constant, K_d, was performed as described previously on the native PAGE in HEPES buffer with variations in K^+ and Na^+ concentrations (Boutorine and Escude 2007). An assortment of TFO concentrations was mixed with the target duplex labelled with fluorescein (0.06 μM). After an overnight preliminary incubation at 4 °C, the samples were loaded into 20 % non-denaturing PAGE, made with the same buffer. After electrophoresis at 37 °C, the movement of the fluorescein-attached structures was detected using an FLA5000 fluorescent image reader. Due to the high molecular weight, duplex-TFO complexes had a lower mobility when compared to the duplexes. In case of successful association duplex-TFO complex should indicate the triplex formation. With increasing TFO concentrations higher conversions are observed (Figure 3.3).

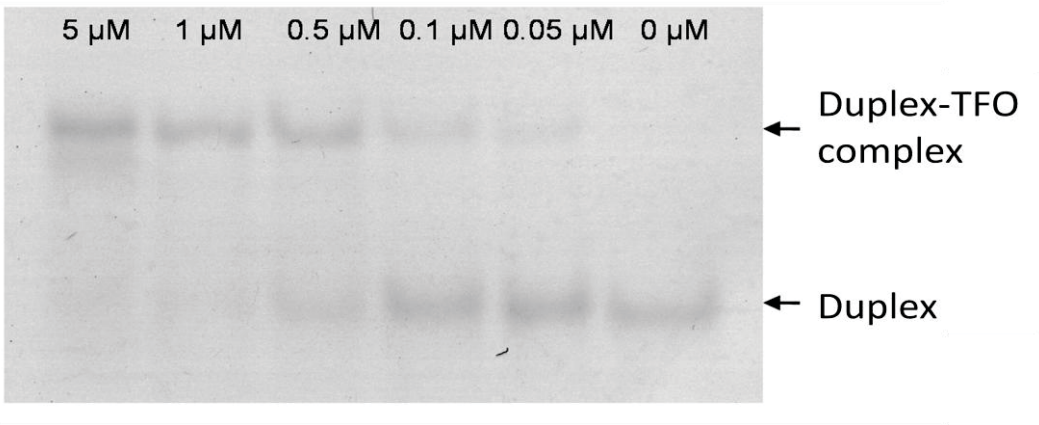


Figure 3.3 Non-denaturing gel analysis of TFO **7** with Target **I**. Duplex concentration: 60 nM, probe concentrations: 5, 1, 0.5, 0.1, 0.05, 0 μM (lanes 1 to 6, respectively) in 0.05 M HEPES buffer, 50 mM NaCl, 5 mM $MgCl_2$ at pH 7.2, 37 °C.

Proportions of the conversion in each lane were detected using ImageQuant software and dissociation constants were calculated from the conversion extents using Microsoft Excel (See Experimental Section).

We have reached several conclusions about the incorporation of TINA monomers into antiparallel TFOs. These conclusions are summarised below.

3.5.2 TINA Insertion as a Bulge in the TFO is Essential for Triplex Formation and Stability, Especially at Physiological pH

As a result of TINA incorporation, TFOs have successfully formed stable triplexes with target **I** after preliminary incubation in near-physiological conditions (50 mM HEPES buffer, 50 mM NaCl, 5 mM $MgCl_2$, pH 7.2). In comparison, under same conditions none of the unmodified TFOs formed triplexes with target **I**. Under these conditions the antiparallel triplex formed by TFO **7** had the lowest dissociation constant (K_d = 0.16 ± 0.02 μM), which was the highest affinity towards target **I**. It is important to note that TFO **7** is a GT sequence containing two TINA monomers and has the lowest TINA density among antiparallel TINA-TFOs. Thermal denaturation experiments have shown that thermal dissociation temperatures in the majority of antiparallel TFOs were very close and almost identical to the melting point of the duplex (T_m = 68 °C).

3.5.3 TINA-Conjugated Triplexes Preserve Sequence Selectivity

Sequence selectivity is very important for applications of TFOs. This property needs to be preserved during any chemical modification for TFOs to be used as probes. We have observed that TINA-modified TFOs strongly bind to target **I**. A scrambled version of target **I** was designed and synthesised with two mismatches (target **III)**. The mismatches were designed as **G-C** to **T-A** and **A-T** to **C-G** as shown in bold (Figure 3.2). The dissociation constant of the triplex formed between TFO **7** and target **III** increased by 100-fold (K_d =17.8 ± 2.5 μM) in comparison to TFO **7** and target **I**. This indicates that a significant portion of the sequence selectivity was preserved.

3.5.4 Antiparallel Triplexes are Significantly More Stable than Parallel Triplexes

TFO **12** and **13** were designed for parallel oriented triplex formation. TINA monomers were positioned similarly to TFO **7**, which yielded the highest affinity to target **I** in 50 mM Na^+. GT-TFO **13** showed the lower affinity, while CT-TFO **12** had relatively better affinity despite

pH-sensitivity of TFO **12**. TFO **12** showed comparable affinity to most of the antiparallel TFOs; however, its affinity did not exceed that of TFO **7** for target **I**.

Table 3.3 Modified and unmodified TFOs targeting HIV-1 proviral DNA and their dissociation constants [μM] from triplexes with corresponding duplex, target **I** (60 nM), in HEPES buffer (50 mM) containing $MgCl_2$ (5 mM), at pH 7.2, 37 °C, supplemented with 50 mM NaCl, 150 mM NaCl or 150 mM KCl.

Oligonucleotide	Sequence 5′-3′	50 mM Na^+	150 mM Na^+	150 mM K^+
TFO **1**	dAGG**X**GG**X**GGTTT**X**TGTTTT	1.32 ± 0.03		
TFO **2**	dAGG**X**GGG**X**GTTT**X**TGTTTT	0.64 ± 0.06		
TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT	0.25 ± 0.03	8.58 ± 0.26	3.9 ± 0.86
TFO **4**	dAG**X**GG**X**GG**X**GTTT**X**TGTTTT	2.70 ± 0.06		
TFO **5**	dAG**X**GGGG**X**GTTT**X**TGTTTT	0.48 ± 0.06		
TFO **6**	dAGG**X**GGGG**X**TTT**X**TGTTTT	1.49 ± 0.03		
TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT	0.16 ± 0.01	1.90 ± 0.03	8.50 ± 1.96
TFO **8**	dAG**X**GGG**X**GGAAA**X**AGAAAA	0.70 ± 0.03		
TFO **9**	dAG**X**GGG**X**GGTTT**X**TGT**X**TTT	1.63 ± 0.21		
TFO **10**	dAGGGGGGTTTTGTTTT	n/d [a]		
TFO **11**	dAGGGTGGGTTTTGTTTT	n/d [a]		
TFO **12**	dTTTTCT**X**TTTCCC**X**CCCT	0.32 ± 0.02	n/d [a]	n/d [a]
TFO **13**	dTTTTCT**X**TTTGGG**X**GGGT	10.67 ± 0.64		
TFO **14**	dTTTTCTTTTCCCCCCT	n/d [a]		
TFO **15**	dTTTTGTTTTGGGGGGT	n/d [a]		

[a] n/d - No triplex formation was detected below 50 μM TFO and 60 nM duplex concentrations. **X** represents TINA monomer.
[b] '-' indicates that the experiment was not performed.

3.5.5 Triplexes Formed by GT Sequences are More Stable than GA sequences

According to reverse Hoogsteen base pairing, the nucleotide backbone of a **G•G-C** triplet has a similar alignment to that of a **T•A-A** triplet. This means a triplex formed by a GA sequence is more isomorphous in comparison to a GT sequence. However, when thymidines of TFO **3** were converted to adenosines, the K_d of resulting TFO **8** increased threefold (K_d = 0.70 ± 0.03 μM). Comparison of TFO **3** and **8** suggests that GT-TFOs are more stable than GA-TFOs for this target **I**. This conclusion was also supported in a previous study, where long T-tracts were shown to result in more stable antiparallel triplexes compared to A-tacts (Keppler, Neidle et al. 2001). Still, we have chosen to use dA as the 5′ cap of our TFOs in order to maintain backbone isomorphism. This cap was seen to be necessary in order to prevent dimerization to form G-quadruplexes (Wang and Patel 1992) rather than assisting in triplex stability. The resulting complexes in the absence of the 5′-dA capping had a slower migration on non-denaturing gel (Figure 3.4; lanes 21, 22). For this reason we advise that the choice of the sequence should depend on the number of the homonucleotide stretches in the target.

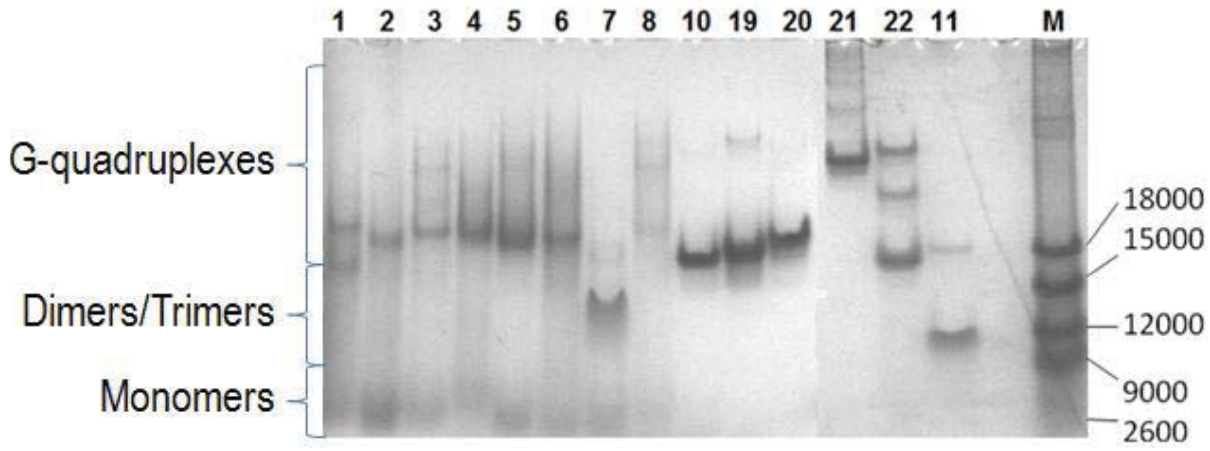


Figure 3.4 Non-denaturing 20 % PAGE analysis of TFOs in HEPES buffer (50 mM), KCl (150 mM), $MgCl_2$ (5 mM), at pH 7.2, 37 °C. Lanes 1 - 11, 19 - 22 refer to TFOs **1** - **11**, **19** - **22**, respectively. M: Marker consisting oligothymidines with shown molecular mass in g/mol. Oligonucleotide concentration was set to 100 μM. Gel was stained by Stains-All® and destained in water.

3.5.6 TINA Incorporation at the Junction Points of Different Nucleotide Tracts Should be Avoided

The positioning of TINA monomers is essential to reach the maximum affinity. We have observed that insertion of TINA monomers between two tracts of different nucleotide types has resulted in decreased affinity in comparison to that of TINA monomers inside tracts of same type nucleotides. When the triplex formed by TFO **6** is compared to those formed by TFOs **2**, **3** or **5**, it can be seen that the positioning of TINA between the G-tract and the T-tract in TFO **6** results in significant decrease in affinity in binding to target **I**. A similar conclusion was also obtained for CT-TINA-TFO that formed a parallel triplex (Filichev and Pedersen 2005).

3.5.7 Insertions of TINA should be Located at Least 3 Bases Apart

It is expected that each TINA incorporation distorts the backbone isomorphism further. If it is incorporated too densely the affinity tends to drop. When designed antiparallel TFOs were compared in terms of TINA density; TFO **7** has the highest affinity with only two TINA incorporations whereas TFO **4** has the lowest affinity for target **I** with four TINA incorporations.

3.5.8 TINA Insertion into the G-tract Helps to Disrupt the G-quadruplex Formation

Preferably, TFOs should not participate in the formation of any complexes except triplexes; however, G-rich sequences show a strong tendency to form highly stable G-quadruplexes. Insertion of TINA monomers into the G-tracts can be used to negate possible G-quadruplex formation and push the equilibrium towards formation of antiparallel triplexes. We have observed the formation of multimeric complexes, such as G-quadruplexes, by non-denaturing gel retardation experiments. After preliminary incubation at 100 μM strand concentration in 50 mM HEPES buffer, 50 mM NaCl, 5 mM $MgCl_2$, pH 7.2 at 4 °C, followed by gel electrophoresis, TFOs were visualised by UV radiation at 362 nm (Figure 3.5, A), where TINA monomers were excited at and by staining with “Stains-All®” (Figure 3.5, B). These experiments indicate that,

depending on TINA position and number, TFOs adopted various complexes with different molecular sizes.

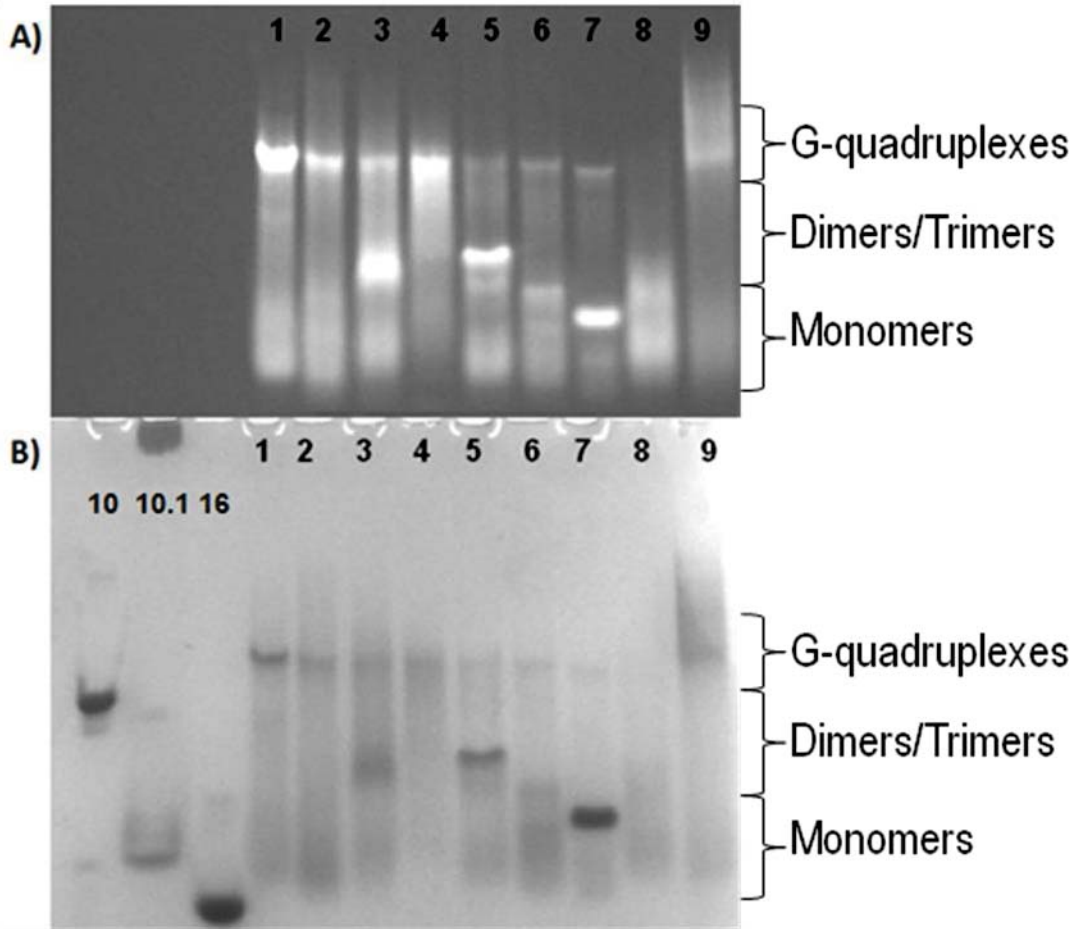


Figure 3.5 Native PAGE (20 %) analysis of TFOs **1**-**10** and **16** in HEPES buffer (50 mM), NaCl (50 mM), $MgCl_2$ (5 mM), at pH 7.2, 37 °C. Lanes 1 - 10 refers to TFOs **1** - **10**, respectively. Lane 10.1 refers to 1M urea and LiOH treated TFO **10** in order to destabilise G quadruplexes, while lane 16 refers to TFO **16**. Oligonucleotide concentration was set to 100 μM. Gel was initially observed under UV light (A) and then stained by Stains-All® and destained in water (B).

With six contiguous guanines, unmodified TFO **10** is expected to form a tetramolecular G-quadruplex (Arimondo, Riou et al. 2000). A thermal denaturation study using CD spectroscopy of TFO **10** showed that this G-quadruplex did not melt until 90 °C in the presence of 50 mM HEPES buffer, 50 mM NaCl and 5 mM $MgCl_2$ at pH 7.2. TFO **10** was treated by 7 M urea, LiOH (0.1 M), heated up to 90 °C for 5 min and finally quenched with $HClO_4$ (0.1 M) in

order to denature the G-quadruplex and yield ssDNA (Figure 3.5, 10.1). No triplex formation with target **I** twas detected with either denatured or non-denatured TFO **10**.

Untreated and denatured TFO **10** represent G-quadruplex and single-stranded, respectively (Figure 3.5, 10 and 10.1). In contrast, modified TFOs showed various mobilities. TFO **7** was shown to exist mostly in a single-stranded form. These results were also confirmed by CD spectroscopy. It is expected that CD profile of a single-stranded disordered oligonucleotide would be featureless, as was observed for TFO **7**. With a negative band at 241 nm and a positive band at 264 nm, TFO **10** showed a CD spectrum typical for a parallel G-quadruplex (Figure 3.6).

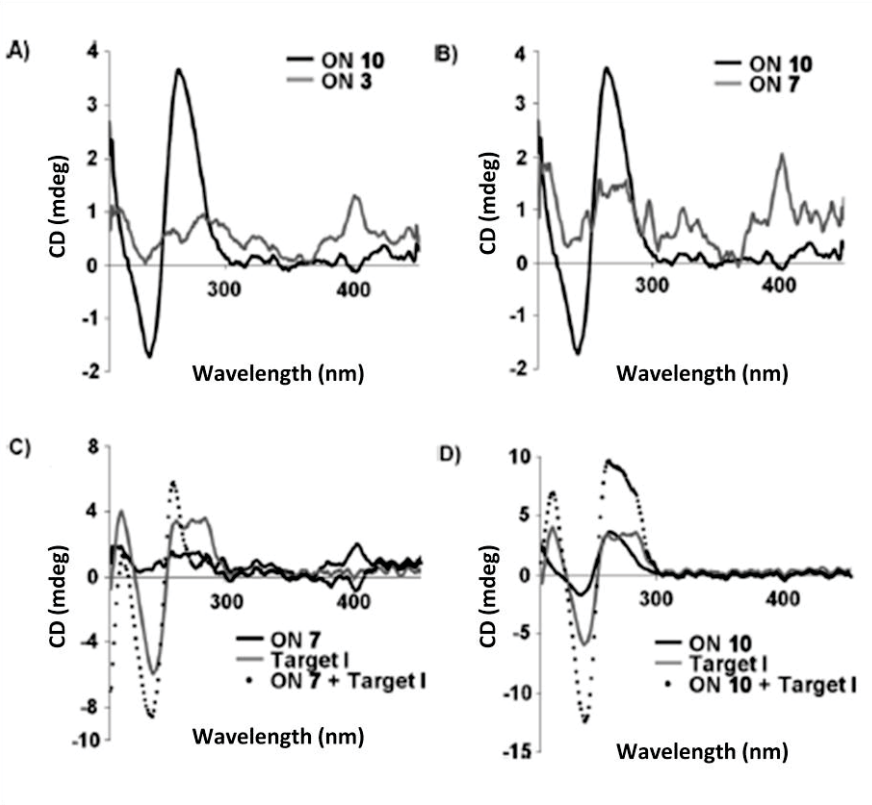


Figure 3.6 CD spectra of A) TFO **3** and B) TFO **7** in comparison with TFO **10**. CD profiles of target duplex, target I, in the absence and presence of C) TFO **7** and D) TFO **10** in Na cacodylate buffer (20 mM) with NaCl (50 mM) and $MgCl_2$ (5mM) at pH 7.2, 20 °C. Concentrations of oligonucleotides are 1.0 μM.

Insertion of a TINA into the G tract pushed the single-stranded versus G quadruplex equilibrium towards lower order structures for most of the TFO designs and generated irregular CD profiles. A distinct band at 395 nm characteristical belongs to the phenylethylpyren-1-yl moiety. However, this equilibrium is still dependent on environmental factors. In 150 mM KCl or NaCl

the G-quadruplex formation was more favourable in comparison to 50 mM NaCl. TFOs were analysed in 150 mM KCl using native PAGE, in which TFO **7** migrated as a dimer and TFO **3** formed G-quadruplex-like structures instead of being single-stranded (Figure 3.4). This may explain the difference in K_d values for dissociation of the triplex complex with target **I** under different salt concentrations and clearly shows the relationship between the stability of triplexes and the ability of G-rich sequences to aggregate.

3.5.9 TINA Oligonucleotides Show Tendency to Self-aggregate

Although TINA insertion is an efficient way to disrupt G-quadruplex structures, extensive TINA insertions may lead to self-association of TINA-TFOs due to hydrophobic interactions between pyrene moieties. Previously, similar effects were observed when four pyrene residues had been inserted into 14-mer CT-sequence with three nucleotides between each insertion (Geci, Filichev et al. 2007). While no triplex formation was observed, an excimer band of high intensity in the fluorescence spectrum indicated strong interactions between pyrene moieties. Self-association of TFOs obviously reduces triplex formation efficiency. In our case, both TINA-TFOs **4** and **9** were mostly found as multimeric complexes on the native gel and almost no monomer form was detected. These TFOs also yielded highest dissociation constants among antiparallel TFOs (K_d = 2.7 and 1.63 μM, respectively). On the other hand, TFOs **3** and **5** showed a slower migration in the gel possibly corresponding to a dimer or a trimer and had a lower affinity towards the duplex than TFO **7**, which exists mostly in single-stranded form in 50 mM HEPES buffer, 50 mM NaCl, 5 mM $MgCl_2$, pH 7.2.

However, these aggregates can easily be disrupted by urea treatment prior to electrophoresis. Disruption of G-quadruplexes requires urea, heat and an alkaline treatment (90 °C, 5-10 min). Fluorescence spectroscopy studies support the existence of pyrene-pyrene interactions with an excimer band at 490 nm. This excimer band is strongly expressed by TFOs **4** and **9** (Figure 3.8). In contrast, the intensity of the excimer band is significantly lower for TFOs **3** and **8**, which form fewer higher-order structures. Interestingly, in the presence of target **I**, the excimer intensity decreases even further for these TFOs, showing the displacement of the equilibrium towards triplex formation. We can conclude that in order to maximise the affinity potential, TINA content should not exceed 14-15 % (Geci, Filichev et al. 2007).

3.5.10 Antiparallel Triplex Formation by TINA-TFOs is Slower than Formation of Parallel Triplexes

The kinetics of TINA-triplex formation were also studied by comparing the level of triplex conversion at different times. Starting from the initiation of the incubation at 20 °C, samples were loaded from the solutions onto an already-running non-denaturing gel at 10, 20, 40, 60, 120, 180, 240 and 300 min.

Previously, three hours was shown to be sufficient for parallel triplex formation (Boutorine and Escude 2007). However, in the case of antiparallel triplex formation, the conversion was not complete after five hours of incubation of 5 μM TFOs **3** or **7** in the presence of target **I**, 50 mM NaCl, 5mM $MgCl_2$, pH 7.2, 20 °C as detected using native PAGE (Figure 3.7).

The full conversion was finally reached after 16 hours at 20 °C. The presence of an extra TINA monomer at TFO 3 in comparison to TFO 7 does not seem to have a strong effect on the triplex conversion rate. In order to observe full conversion overnight incubation should be preferred.

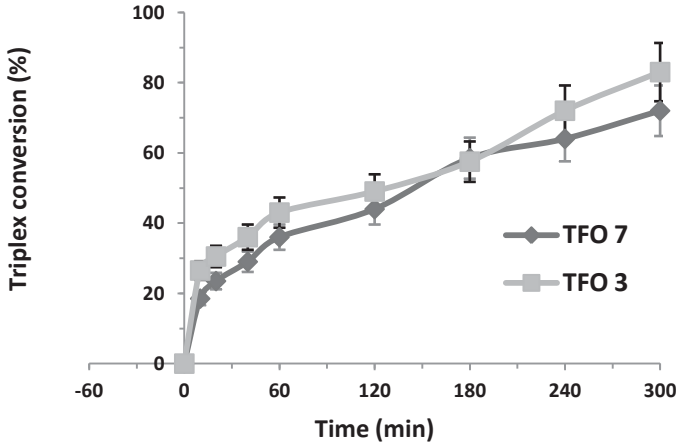


Figure 3.7 Triplex conversion versus time (min) by TFO **3** and **7** at 100 μM strand concentration in the presence of 0.06 μM Target **I** in HEPES buffer (50 mM), NaCl (50 mM), $MgCl_2$ (5 mM), pH 7.2 at 20 °C.

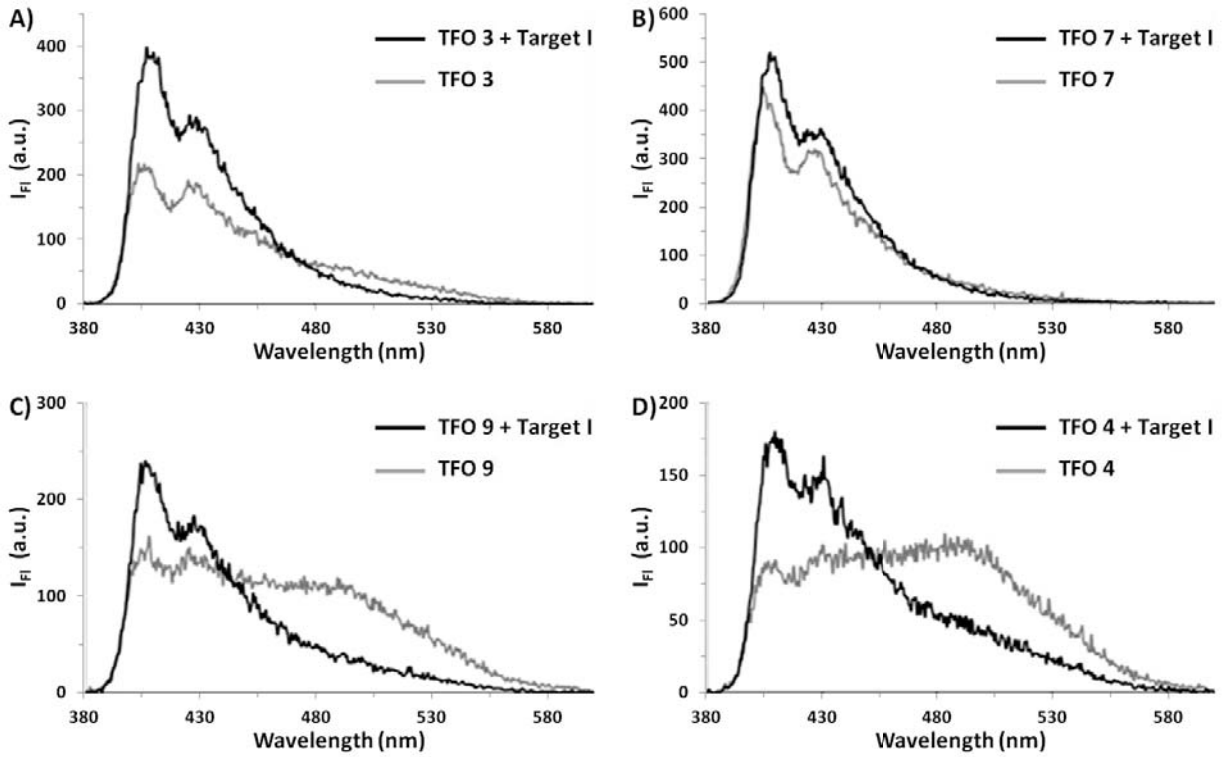


Figure 3.8 Fluorescence spectra of TFO **3** (A), **4** (D), **7** (B) and **9** (C) in the presence and in the absence of the target duplex, Target **I**, λ_{ex}:373 nm, λ_{em}:380 nm – 600 nm, in Na cacodylate buffer (20 mM), NaCl (50 mM), $MgCl_2$ (5 mM), at pH 7.2, 20 °C. Concentration of each strand is set to 1.0 μM.

3.5.11 Thymidine Insertion Instead of TINA can be Used to Simulate TINA's Effect of G-quadruplex Disruption

It can be expensive to synthesise several TINA-TFOs to find the most effective TFO design for one particular target. Fortunately, thymidines can be used instead of TINA monomers to simulate the effect of TINA on G-quadruplexes. Just as TINA, thymidines are sufficient to disrupt guanine continuity and G-tetrad stacking. Thymidines negate association of G-quadruplexes at low oligonucleotide concentrations. TFO **11** which contains an extra thymidine residue in the middle of the G-tetrad, was shown to migrate on non-denaturing gel electrophoresis as single-stranded structure even in the presence of KCl (Figure 3.4). Although aggregates were negated, TFO **11** did not form any triplex with target **I**. This was not a surprise since thymidine insertion into the G-tract would create a bulge with the target. In order to observe the triplex-forming potential of this sequence, a new target with corresponding **T-A** base-pair was

synthesised (target **II**) so the duplex target and TFO **11** would be fully matched. However, TFO **11** did not form any triplexes with target **II** even at 10 μM TFO concentration. This sequence is an important control because incubation of TFO **11** with target **II** shows that TINA does not only disrupt G-quadruplex but also greatly increases the stability of the triplexes through intercalation. At the same time, these results show that more accessible and cheap dT can be used to simulate G-quadruplex disruption. This can simplify the screening of correct TINA positioning in the G rich TFOs using native gel electrophoresis.

3.6 Demonstration of Applicability of the Rules for TINA-TFO Design

We have chosen to design antiparallel TINA-TFOs to target human Abelson murine leukimia viral proto-oncogene in order to demonstrate the applicability of the rules defined for TINA incorporation. We listed TFO targets in this gene that were previously used for labelling and visualisation of chromosomes via parallel triplex formation at pH 5.5 (Table 3.4, ABL **1 - 6**) (Schwarz-Finsterle, Stein et al. 2007).

In this part of the study sequences ABL **1 - 6** with high guanosine content were screened in non-denaturing PAGE. As expected, most of them existed in equilibrium between multimeric aggregates and monomers under near-physiological conditions (50 mM NaCl, 5 mM $MgCl_2$, pH 7.2, Figure 3.9). ABL **1** and **6**, were abundantly found as aggregates. We designed sequences based on ABL 1 and 6 in which G-tracts were interrupted by thymidine(s) (ABL **1-2T1**, ABL **1-2T2**, ABL **6-1T**, ABL **6-2T**). Four sequences were designed based on ABL **1** and **6** (Table 3.5); all of them migrated individually as single-stranded DNA in the native gel (Figure 3.9).

ABL **1-2T2** and ABL **6-2T** were selected as suitable designs and re-synthesised with TINA monomers instead of dT. ABL **1-2X** and ABL **6-2X** showed considerably faster migration in comparison to ABL **1** and **6**, respectively. However, ABL **1-2X** formed an aggregate, probably a dimer, due to hydrophobic TINA interactions. We also observed that both TINA-TFOs have improved binding affinity towards to corresponding targets, **IV** and **V** (Table 3.4). These results demonstrate the feasibility of the protocol described.

Table 3.4 Modified and unmodified TFOs targeting the *ABL* region and their dissociation constants [μM] from triplexes with corresponding duplexes (Table 3.5, 60 nM) in HEPES (50 mM), NaCl (50 mM) and $MgCl_2$ (5 mM) at pH 7.2, 37 °C.

Oligonucleotide	Sequence (5'-3') [a]	K_d [μM]
ABL 1	dGGGGAAGAGAGGGGGAG	1.48 ± 0.02
ABL 1-2T1	dGG**T**GGAAGAGAGG**T**GGGAG	- [b]
ABL 1-2T2	dGG**T**GGAAGAGAGGG**T**GGAG	n/d [c]
ABL 1-2X	dGG**X**GGAAGAGAGGG**X**GGAG	0.82 ± 0.10
ABL 2	dAGAAAAGGAGGGGAGAGG	- [b]
ABL 3	dGAAGGAGGAAGGGAGGGAGGGGG	- [b]
ABL 4	dGAAGAAAGAAGGAAGGGGAGGA	- [b]
ABL 5	dAGAAGGGAAAAAAGGGGA	- [b]
ABL 6	dGAAGGGGGAGAGGGAAG	1.24 ± 0.03
ABL 6-1T1	dGAAGG**T**GGGAGAGGGAAG	- [b]
ABL 6-1T2	dGAAGGG**T**GGAGAGGGAAG	- [b]
ABL 6-2T	dGAAGGG**T**GGAGAGG**T**GAAG	n/d [c]
ABL 6-2X	dGAAGGG**X**GGAGAGG**X**GAAG	0.65 ± 0.05

[a] **X** represents TINA monomer. Inserted oligothymidines are presented in bold. [b] '-' indicates that the experiment was not performed. [c] 'n/d' indicates that no triplex formation was detected below 50 μM TFO at 60 nM duplex concentrations.

Table 3.5 Sequences of target duplexes **IV**, **V**. Target **IV** and **V** are designed for probes based on ABL **1** and ABL **6**, respectively.

Target duplex	Sequence
Target **IV** [a]	5′-dACTGAGGGGGAGAGAAGGGGTCT-3′
	fluorescein-3′-dTGACTCCCCCTCTCTTCCCCAGA-5′
Target **V** [b]	5′-dTACGAAGGGAGAGGGGGAAGTAC-3′
	fluorescein-3′-dATGCTTCCCTCTCCCCCTTCATG-5′

[a] The target **IV** was designed as a complementary duplex for ABL **1**. [b] The target **V** was designed as a complementary duplex for ABL **6**.

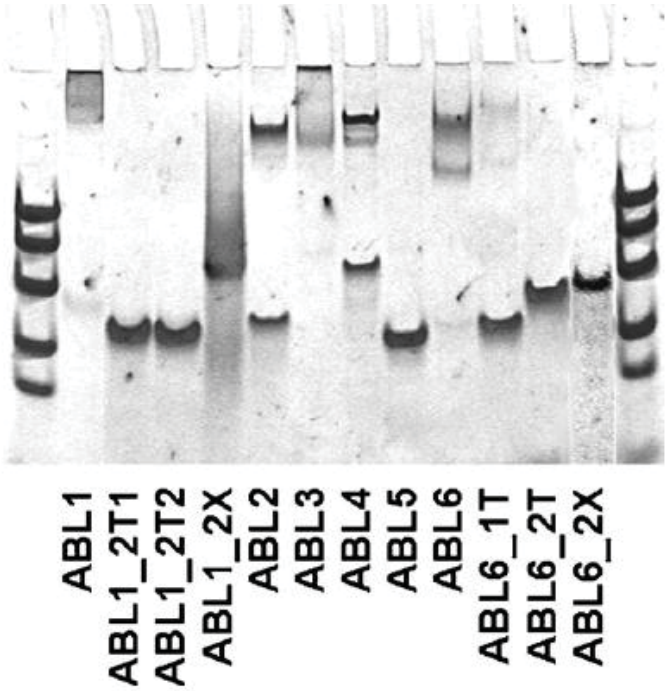


Figure 3.9 Non-denaturing PAGE (20 %) analysis of ABL sequences listed in Table 3.4 in HEPES buffer (50 mM), NaCl (50 mM), $MgCl_2$ (5 mM) at pH 7.2, 37 °C with DNA ladders consisting of 50-, 40-, 30-, 20- and 15-mer oligothymidylates. Oligonucleotide concentration is 100 μM. Gel was stained with Stains-All® and destained in water.

3.7 Discussion

DNA triplex technology is an attractive approach to label, modify or alter gene expression both *in vitro* and *in vivo* (Faria and Giovannangeli 2001; Vasquez and Glazer 2002; Duca, Vekhoff et al. 2008; Simon, Cannata et al. 2008). However, there are several problems that have to be solved to make TFOs efficient sequence-selective DNA targeting agents. TFOs must overcome unfavourable negative charge-charge repulsion between phosphates in the complex, be able to penetrate the cell membrane, have a resistance to nucleases for *in vivo* applications and be able to form stable triplexes in physiological conditions so the resulting complex will interfere with bioprocesses on DNA. Chemical modifications of TFOs have been the obvious choice to overcome these limitations (Vasquez and Glazer 2002; Buchini and Leumann 2003; Duca, Vekhoff et al. 2008).

Most studied native CT-TFOs form parallel triplexes at acidic conditions and are not suitable for these tasks. Parallel GT and antiparallel GT/AT triplexes, despite their attractiveness to be formed at physiological pH, have been generally less explored. The main reason behind this is the self-aggregation of G-rich TFOs mediated by guanines. Extra modifications were explored to inhibit these structures (Milligan, Krawczyk et al. 1993; Durland, Rao et al. 1995; Olivas and Maher 1995; Aubert, Perrouault et al. 2001). There is also always the question of if it is really a triplex, and not some self-aggregated G-quadruplex, that causes an effect when oligonucleotides were used on cell lines, as in the case of the discovery of AGRO100, which was designed as TFO but currently in clinical studies as a G-quadruplex forming oligonucleotide under the name AS1411 (Bates, Laber et al. 2009). It should also be noted that the use of backbone-modified nucleic acids, such as peptide nucleic acid (PNA) and 2′-*O*-modified RNAs (including locked nucleic acid (LNA)), in the formation of antiparallel triplexes is very limited. This is because PNA and LNA modified sequences have even higher tendency for self-aggregation than unmodified structures (Wittung, Kim et al. 1994; Wittung, Nielsen et al. 1994; Brunet, Alberti et al. 2005). Therefore, in the case of G-rich ONs, LNA and PNA-based sequences may be excluded from the triplex formation in any orientation. In that regard, the development of molecules that are able to tune properties of G-rich oligonucleotides and form antiparallel triplexes is of high importance.

Organic intercalators have been introduced in DNA triplex technology to increase triplex stability as separate molecules. It was demonstrated that acridine, benzo(e)pyridoindole or other similar intercalators are able to stabilise parallel triplexes (Escude, Mohammadi et al. 1996; Escude, Nguyen et al. 1998). Moreover, benzo(e)pyridoindole derivatives were also shown to induce formation of triplexes by antiparallel GT-TFO, which was not formed in the absence of the intercalator (Escude, Sun et al. 1996). The next step was to incorporate the aromatic molecules into the TFO in the middle, at the 5′ end or instead of a nucleobase. Later acridine (Sun, Francois et al. 1989), perylene (Aubert and Asseline 2004; Filichev and Pedersen 2005), psoralen (Bates, Macaulay et al. 1995; Miller, Kipp et al. 1999) and pyrene (TINA)(Filichev and Pedersen 2005) analogues were inserted as a bulge in the middle of TFOs and led to an increased thermal stability of parallel triplexes. Among these, psoralen and TINA were also demonstrated to form antiparallel triplexes (Miller, Bi et al. 1996; Keppler, Neidle et al. 2001; Paramasivam, Cogoi et al. 2008). It was found that TINA was not only able to form antiparallel triplex but was

also capable of disrupting self-aggregation of GA-TFOs and enable them to form antiparallel triplex within the promoter region of *KRAS* proto-oncogene at physiological pH and ion concentrations. The designed TINA-TFOs were found to inhibit the formation of a DNA-protein complex and to down-regulate the transcription of the chloramphenicol acetyltransferase driven by the murine *KRAS* promoter. It is important to mention that parallel TINA-TFOs designed to bind to the same target did not form a triplex spiecies at physiological pH. Therefore, organic chromophores covalently attached to TFOs in the middle of the sequence represent an important class of compounds. These compounds may be used to overcome self-aggregation by G-rich oligonucleotides and be used as efficient triplex-forming agents.

In this chapter we have investigated the effect of TINA position and density on antiparallel TFO efficiency. We selected the best TINA-TFOs targeting an HIV-1 polypurine sequence, and established the rules for their design. The selected GT-TFO consist of a single run of six guanines and forms a tetramolecular G-quadruplex of high thermal stability, whereas in a previous study TINA was inserted into GA-TFOs containing several 2-5 nt long runs of guanines. With this study we advanced our knowledge of the design of TINA-TFOs for different types of TFO and genomic sequences. Bulged TINA monomers were incorporated into G_6-tract of the G-rich TFO and a series of TINA-TFOs that formed stable antiparallel triplexes was obtained. The triplex affinity was greatly dependent on the design of the TINA insertions, as discussed earlier.

The disruption of G-quadruplex-like structures was achieved by placing TINA in the middle of the G-tracts. A direct relationship between G-quadruplex disruption and increased affinity towards the target duplex was observed. The sequence specificity typical of the TFO was maintained, despite the introduction of intercalating moieties into the sequence. It was not possible to compare the loss of specificity due to introduction of the TINA monomer because the unmodified TFOs would not form any triplex even without mismatch. The total number of TINA insertions must not exceed 28-30 % of the nucleotide content as the lipophilic nature of the phenylethynylpyrene can contribute to the formation of aromatic aggregates that might affect the TFO equilibrium.

While insertion of the TINA monomer in the middle of a G-tract disrupts the formation of G-quadruplexes, a previous study has shown that TINA incorporation outside the G-tract

improved the thermal stability of G-quadruplexes (Cogoi, Paramasivam et al. 2009). This effect is thought to be a result of capping of G-tetrad stacks at the bottom and the top by TINA monomers. These opposite effects on G-quadruplex formations by TINA monomers need further investigation. For this reason, the next chapter will focus on the relationship of TINA incorporation with G-quadruplex structure and stability. We will scan thermodynamic properties of TINA-G-quadruplexes by incorporating TINA monomers in different positions of a dTG_4T sequence.

Chapter 4. TINA Incorporated G-quadruplexes

4.1 Introduction

Guanine-rich sequences are known to form highly stable structures called G-quadruplexes. The significance of these structures has been rising as more and more studies reveal their functions in the cell. With growing experimental evidence of the G-quadruplex binding capabilities of many proteins, understanding and controlling the G-quadruplex topology can assist in applications for molecular biology, bio/nanotechnology and medicine.

Previous studies have shown that pyrene-containing TINA molecules may the affect assembly of G-rich oligonucleotides. When a TINA molecule was positioned at the bottom or top of a G-tetrad stack, the resulting G-quadruplexes, either parallel or antiparallel, were more thermally stable and biologically active in comparison to unmodified complexes (Cogoi, Paramasivan et al. 2007; Cogoi, Paramasivam et al. 2009; Membrino, Cogoi et al. 2011; Pedersen, Nielsen et al. 2011; Rohrbach, Fatthalla et al. 2012). On the other hand, incorporation of TINA between contiguous guanines that participate G-tetrad formation results in significant destabilisation of the overall structure (Paramasivam, Cogoi et al. 2008). The findings of the previous chapter also support this hypothesis, as incorporation of TINA into the G-tract has resulted in a decrease in G-quadruplex formation in all cases.

To get further insights into the ability of TINA molecule to stabilise or destabilise G-quadruplex structures, the well-studied sequence 5′-dTGGGGT from the *Tetrahymena* telomere and its TINA conjugates were investigated. The unmodified sequence forms a tetramolecular parallel G-quadruplex with guanines arranged in an *anti* glycosidic conformation in the presence of Na^+ and K^+ ions. It is characterised by slow association kinetics with 2.6 hours half-association time at 4 °C, in 100 mM NaCl, 10 mM Na cacodylate pH 7.0 buffer (Mergny, De Cian et al. 2005).

4.2 Design of G-quadruplexes

For the DNA sequence design we chose to insert the TINA moiety at the 5'-end of the d(TG_4T) sequence, in the middle of the G-tract and between T and dG at the 5'-end. To evaluate the potential behind multiple insertions of chromophores in the tetramolecular sequences we synthesised DNAs possessing both TINA and fluorescein (Table 4.1).

Table 4.1 Sequences used in this chapter and their abbreviations.

Abbreviation	Sequence 5'-3'
TG4T	dTGGGGT
XTG	d**X**TGGGGT [a]
TXG	dT**X**GGGGT [a]
GTG	dTGGTGGT
GXG	dTGG**X**GGT [a]
GXX	dTGG**XX**GGT [a]
G3X	dTGGG**X**GGGT [a]
GXGF	dTGG**X**GGT**F** [a, b]
TXGF	dT**X**GGGGT**F** [a, b]
TINA duplex	dAGCTTG**X**CTTGAG / dCTCAAG**X**CAAGCT [a]

[a] **X** refers to TINA monomer. [b] **F** refers to fluoroscein.

Oligonucleotides were obtained using standard phosphoramidite chemistry and were purified, where necessary, using reverse-phase HPLC. The formation of G-quadruplexes were performed by incubation of the oligonucleotides at 10 μM concentration in 10 mM Li cacodylate buffer, pH 7.2 supplemented with 110 mM NaCl or KCl at 90 °C for 30 min. After cooling to 20 °C samples were left incubating between 2-7 days at 4 °C. After annealing modified G-quadruplexes were studied by means of gel retardation, UV-Vis, circular dichroism (CD), fluorescence and NMR spectroscopy and were compared with the parent **TG4T** sequence. All PAGE experiments were performed at 37 °C.

4.3 Properties of TINA Incorporated G-quadruplexes

4.3.1 TINA Located Outside the G-tract

The presence of TINA in the structure of G-quadruplex forming sequences changed the profiles of thermal difference spectra (TDS), which are often used to characterize the topology of G-quadruplexes (Mergny, Li et al. 2005). TDS are obtained through subtraction of the UV-Vis spectra of folded and unfolded oligonucleotides (see Chapter 2). TDS profiles of complexes formed by TINA-modified **XTG** and **TXG** sequences in the presence of KCl did not resemble any of the G-quadruplex TDS described in the literature (Figure 4.2, Figure 4.3). Two positive peaks at 243 and 273 nm and a negative band at 295 nm are seen in the **TG4T** TDS profile (Figure 4.1) as described in literature for other G-quadruplexes (Mergny, Li et al. 2005). In the case of TINA-containing G-quadruplexes **XTG** and **TXG**, when in the presence of KCl a single positive band at 277 nm is observed with a shoulder at 295 nm. On the other hand, in the presence of NaCl an additional positive band at 241 nm is seen (Figure 4.2, Figure 4.3). The presence of peaks in the range of 300 – 420 nm indicates that TINA participates in the formation of folded structures.

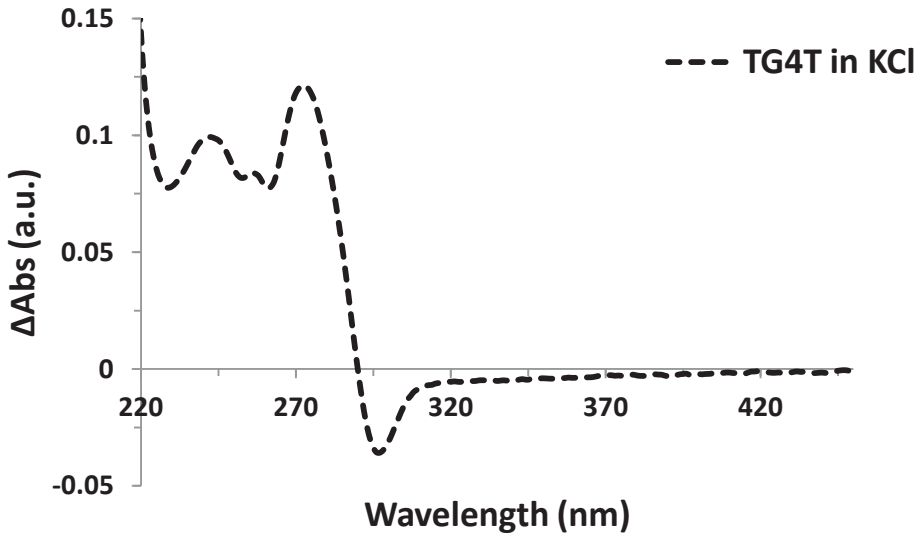


Figure 4.1 Thermal difference spectrum of **TG4T** at 10 μM strand concentration after annealing in the presence of 110 mM KCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

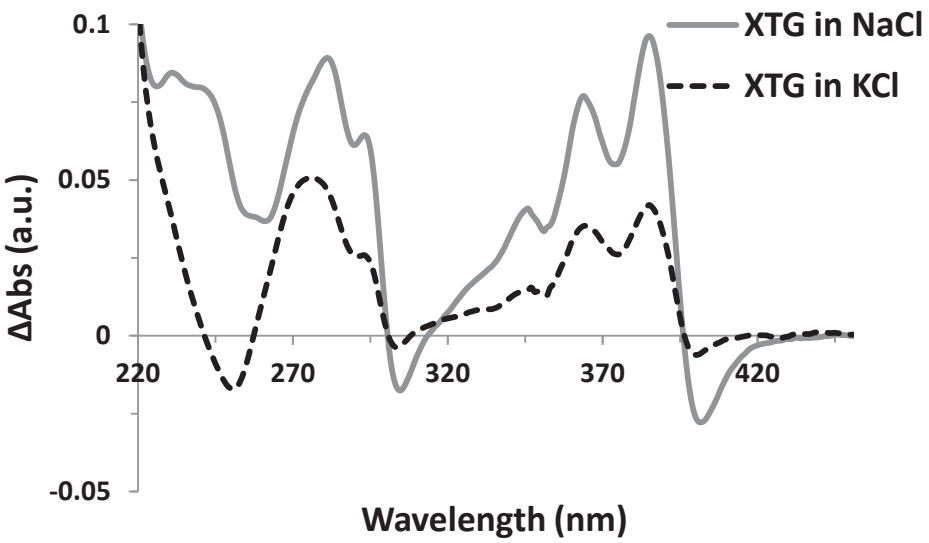


Figure 4.2 Thermal difference spectra of **XTG** at 10 μM strand concentration after annealing in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

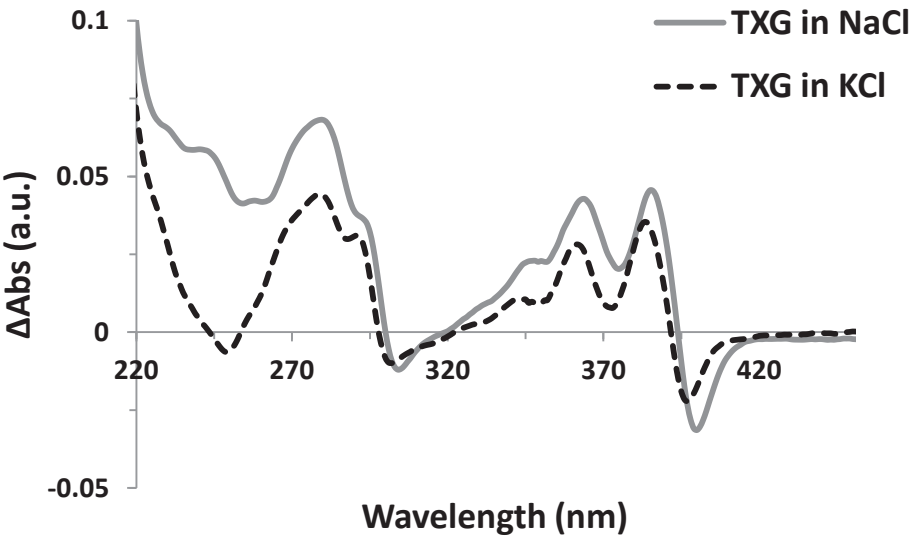


Figure 4.3 Thermal difference spectra of **TXG** at 10 μM strand concentration after annealing in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

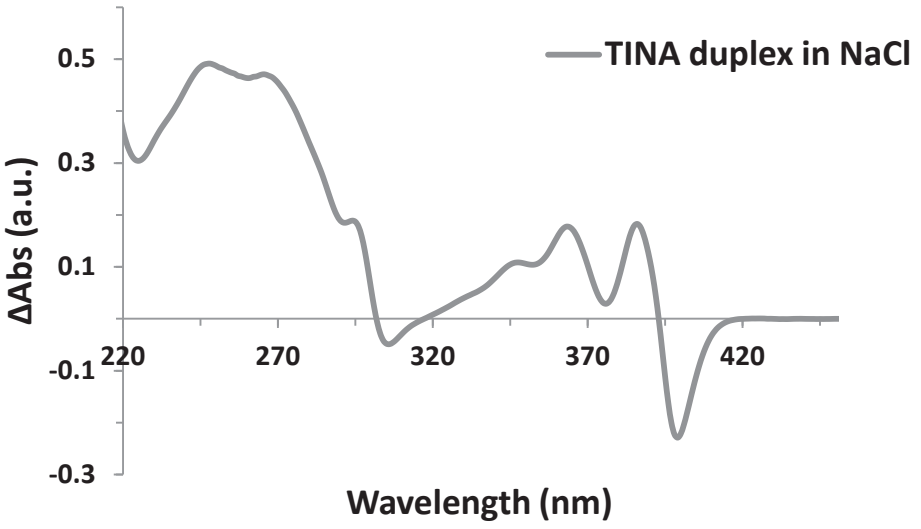


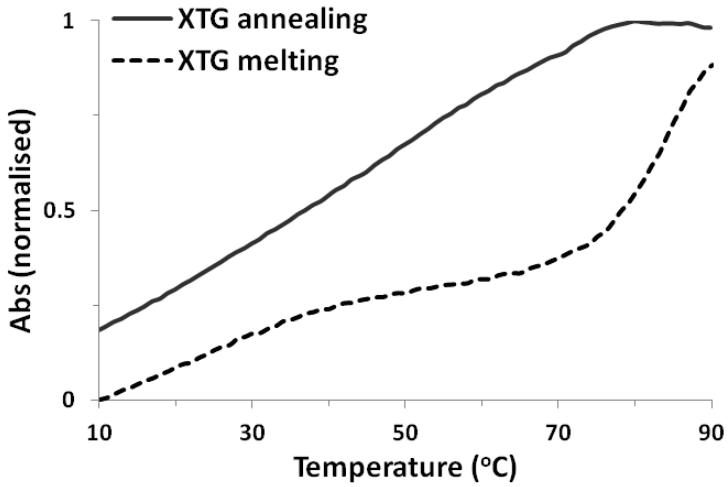
Figure 4.4 Thermal difference spectrum of the **TINA duplex** at 10 μM strand concentration after annealing in the presence of NaCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

We have also observed that the profile in this region for **TXG** is very similar to the TDS profile of a DNA duplex in which two TINA monomers are placed opposite each other (Figure 4.4, **TINA duplex**). The apparent difference in intensity at 240 nm between TDS profiles in the presence of K^+ may indicate incomplete denaturation at 90 °C rather than structural variation. It should also be noted that the overall intensity of peaks in TDS for G-quadruplexes is lower than for the wild-type dTG_4T. This means that monitoring a thermal denaturation of TINA-G-quadruplexes should be performed at several wavelengths: we monitored UV-Vis melting of TINA-containing sequences at 295 and 373 nm. Similar curves were obtained for both wavelengths (see Appendix Figure 1).

According to our UV-Vis thermal stability studies, the formation of G-quadruplexes was not interrupted as long as the TINA molecule was placed outside of the contiguous guanine tract. **XTG** and **TXG** sequences showed enhanced thermal stability with $\Delta T_{½}$ values of 27.5 and 25.6 °C respectively, in comparison with **TG4T** in 110 mM NaCl (Table 4.1). For **XTG** we detected two transitions during melting: the first, at low temperature, is attributed to dissociation of high-order aggregates and the second, at higher temperature, belongs to denaturation of the tetramolecular G-quadruplex (Figure 4.5, A). The annealing profiles indicated faster association rates for TINA containing **XTG** and **TXG** sequences than for the parent G-quadruplex. In the

presence of 110 mM KCl denaturation of these sequences as well as $[d(TG_4T)]_4$ was not complete only after 30 min incubation at 90 °C.

A)



B)

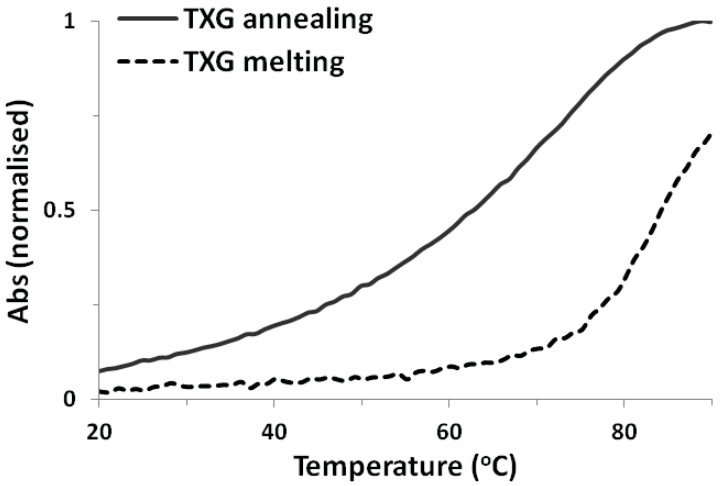


Figure 4.5 Melting (dashed lines) and annealing (solid lines) profiles of **XTG** (A) and **TXG** (B) after incubation in 110 mM NaCl, 10 mM Li cacodylate buffer at pH 7.2. Oligonucleotide concentration was 10 μM. The profiles are based on absorbance data recorded at 373 nm with 0.18 °C/min temperature ramp.

We assessed the molecularity of the complexes formed using native 20 % PAGE in the presence of K^+ and Na^+ ions (Figure 4.6, Table 4.3). Tetramolecular complexes were observed for these sequences except **XTG**, which in 110 mM KCl formed a well defined band above the 50-mer component of the ladder. This suggests dimerisation of the tetramolecular G-quadruplex driven by hydrophobic pyrene residues. This is in accord with observation of two melting transitions

recorded in NaCl. For **TXG** formation of a smear between the 15- and 30-mer ladder components led to the conclusion that this sequence exists as a mixture of re-equilibrating bi- and tetramolecular complexes.

Table 4.2 $T_{½}$ values of G-quadruplexes at 10 μM strand concentration obtained during melting and annealing processes using 0.18 °C/min temperature ramp in the presence of 110 mM NaCl or KCl in 10 mM Li cacodylate at pH 7.2.

	NaCl melting/annealing (°C)	KCl melting/annealing (°C)
TG4T	57.6 / n.d [a]	> 90 / < 20
XTG	85.1 / n.d [a]	> 90 / n.d [a]
TXG	83.2 / 74.8	> 90 / n.d [a]
TXGF	77.1 / < 20	> 90 / n.d [a]
GXG	58.3 / 54.7	56.9 / 55.6
GXGF	54.5 / < 20	62.5 / 35.3
GXX	32.2 / 31.6	35.3 / 38.7
G3X	> 90 / 68.6	>90 / 69.7

[a] not done.

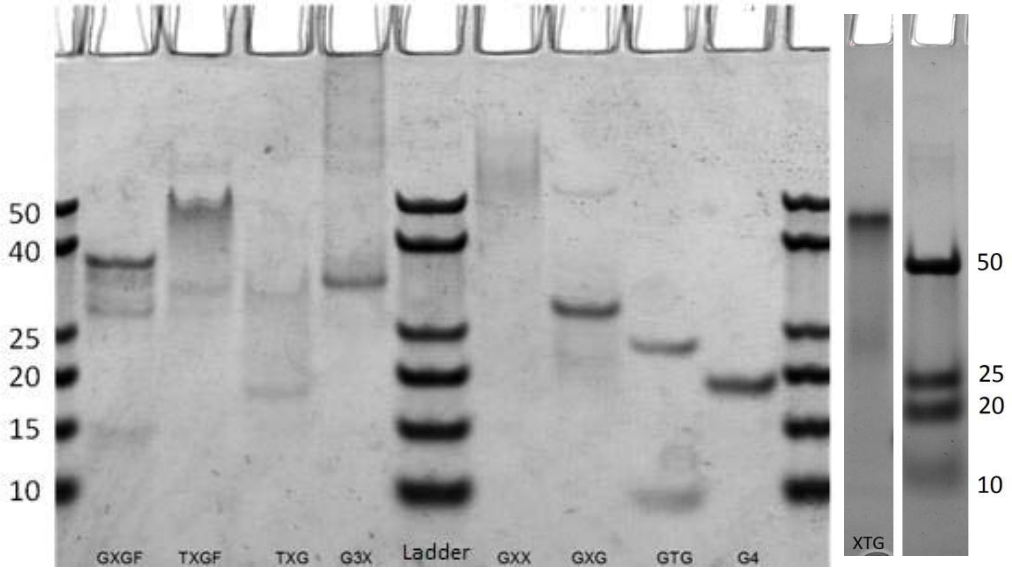


Figure 4.6 Native PAGE (20 %) analysis of oligonucleotides (100 μM) in the presence of 110 mM KCl, 10 mM Li cacodylate at pH 7.2, 4 °C. Ladder contains 10-, 15-, 20-, 25-, 40- and 50-mer oligothymidylates. See Appendix Figure 2 for oligonucleotides in the presence of 110 mM NaCl instead of 110 mM KCl.

Table 4.3 Retardation of the G-quadruplexes formed by TINA-TG_4T sequences in native PAGE (20 %) and molecularity in the presence of 110 mM NaCl or KCl in 10 mM Li cacodylate buffer at pH 7.2, 37 °C.

		NaCl		KCl	
	N [a] (N=nx4)	Mobility [b]	molecularity	Mobility [b]	molecularity
TG4T	24	10-15 mer	4	15-20 mer	4
XTG	28	25-30 mer	4	> 50 mer	aggregation
TXG	28	25-40 mer [c]	4	25-40 mer [c]	4
GTG	28	10-15 mer	ssDNA	~10 mer	ssDNA
				20-25 mer	4
GXG	28	~20 mer	4	25-30 mer	4
GXX	32	25-40 mer [c]	aggregation	> 50 mer [c]	aggregation
G3X	36	~20 mer	4	30-40 mer	4
GXGF	32	20-25 mer	4	~40 mer	4
TXGF	32	~40 mer	aggregation	~50 mer	aggregation

[a] N refers to number of nucleotides/modifications within the tetramolecular G-quadruplex formed by corresponding sequence. [b] Level of mobility in comparison to the oligothymidine ladder. [c] Observed as a smear on gel.

Different interactions between pyrenes in the G-quadruplexes formed by **XTG** and **TXG** can also be seen in the CD spectra (Figure 4.9, Figure 4.10). In the region of nucleobase absorbance, CD profiles of **XTG** and **TXG** in the presence of NaCl and KCl are similar to unmodified **TG4T** sequence (Figure 4.8 – Figure 4.10). All three sequences exhibited a positive ellipticity at 262 nm and a negative ellipticity at 238 nm which is characteristic for parallel G-quadruplexes. Additional signals above 300 nm correspond to the TINA monomer. Two peaks with positive ellipticities are observed for **XTG** at 378 and 395 nm. In contrast, CD spectra of complexes formed by **TXG** showed positive and negative bands centered at 355 and 390 nm, respectively. This indicates a different orientation of the transition dipoles of the pyren-1-yl-ethynylphenyl moieties in comparison to the **XTG** sequence. As a result of pyrene-pyrene interactions in these G-quadruplex structures, a strong excited dimer (excimer) band at 515 nm (λ_{ex} = 373 nm) is seen in the fluorescence emission spectra for both **XTG** and **TXG** complexes, while the monomeric fluorescence bands at 405 and 425 nm due to monomeric species are suppressed in the presence of Na^+ or K^+ ions (Figure 4.7).

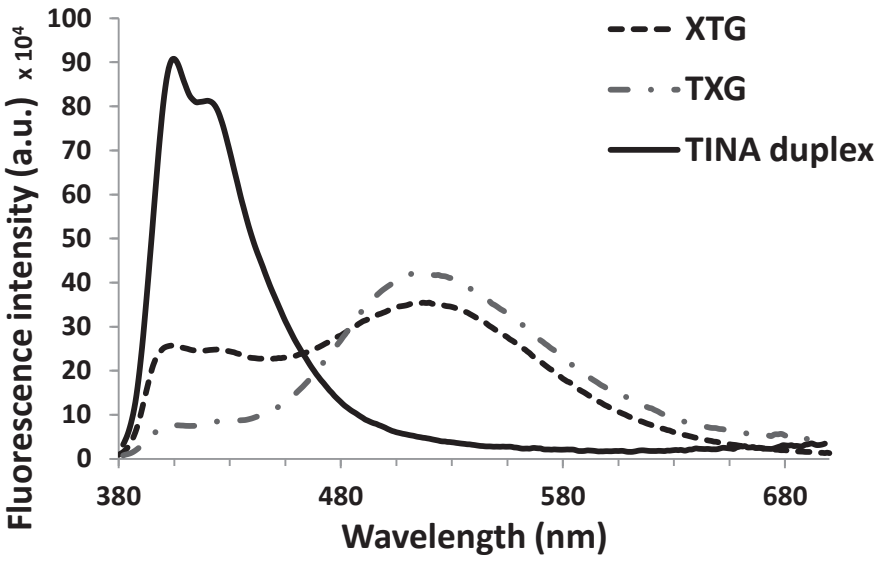


Figure 4.7 Fluorescence emission spectra of **XTG**, **TXG** and **TINA duplex** at 10 μM strand concentration in the presence of 110 mM NaCl in 10 μM Li cacodylate buffer at pH 7.2, 20 °C. λ_{ex} = 373 nm. The fluorescence profiles in the presence of 110 mM KCl instead of 110 mM NaCl are very similar (see Appendix Figure 3).

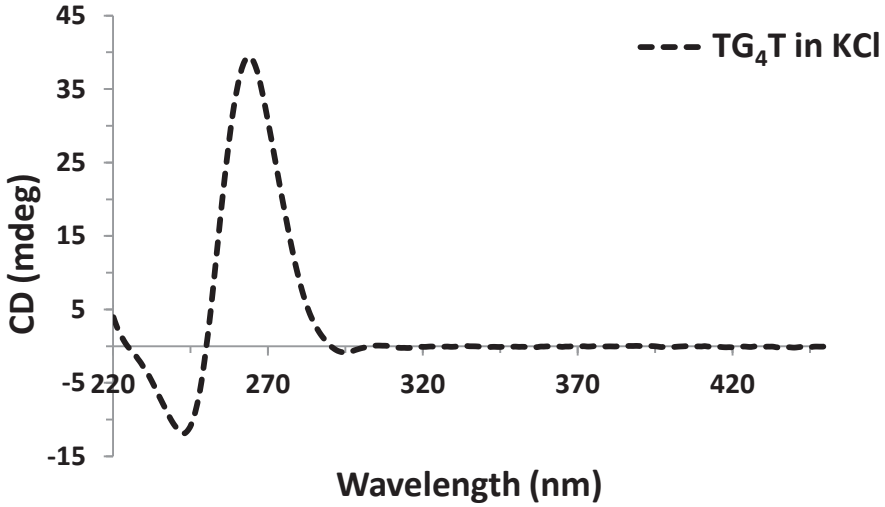


Figure 4.8 CD spectrum of **TG4T** at 10 μM strand concentration in the presence of 110 mM KCl in 10 mM Li cacodylate buffer, pH 7.2, 20 °C.

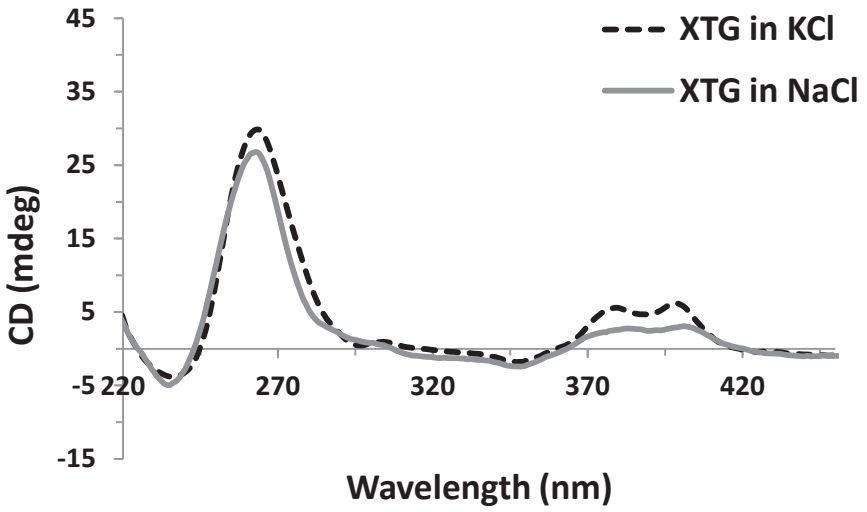


Figure 4.9 CD spectra of **XTG** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer, pH 7.2, 20 °C.

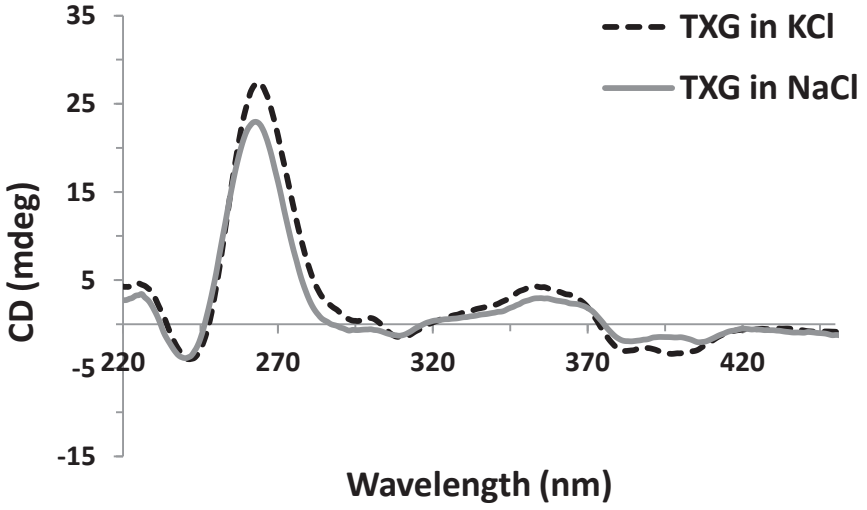


Figure 4.10 CD spectra of **TXG** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer, pH 7.2, 20 °C.

In ^{1}H NMR spectra we observed formation of broad imino-signals in the region of 10.7-11.0 ppm for the **TXG** sequence in 10 mM sodium phosphate buffer (Figure 4.11, A). At elevated temperatures a sharpening of signals was detected with the appearance of three distinct peaks at

55 °C. Addition of K^+ ions resulted in broadening of imino signals indicating formation of aggregates (Figure 4.11, B). The appearance of very few imino signals suggests that the folded structure is symmetrical and resembles that of a wild-type tetramolecular parallel G-quadruplex formed by **TG4T** sequence reported in the literature (Aboul-ela, Murchie et al. 1994; Caceres, Wright et al. 2004).

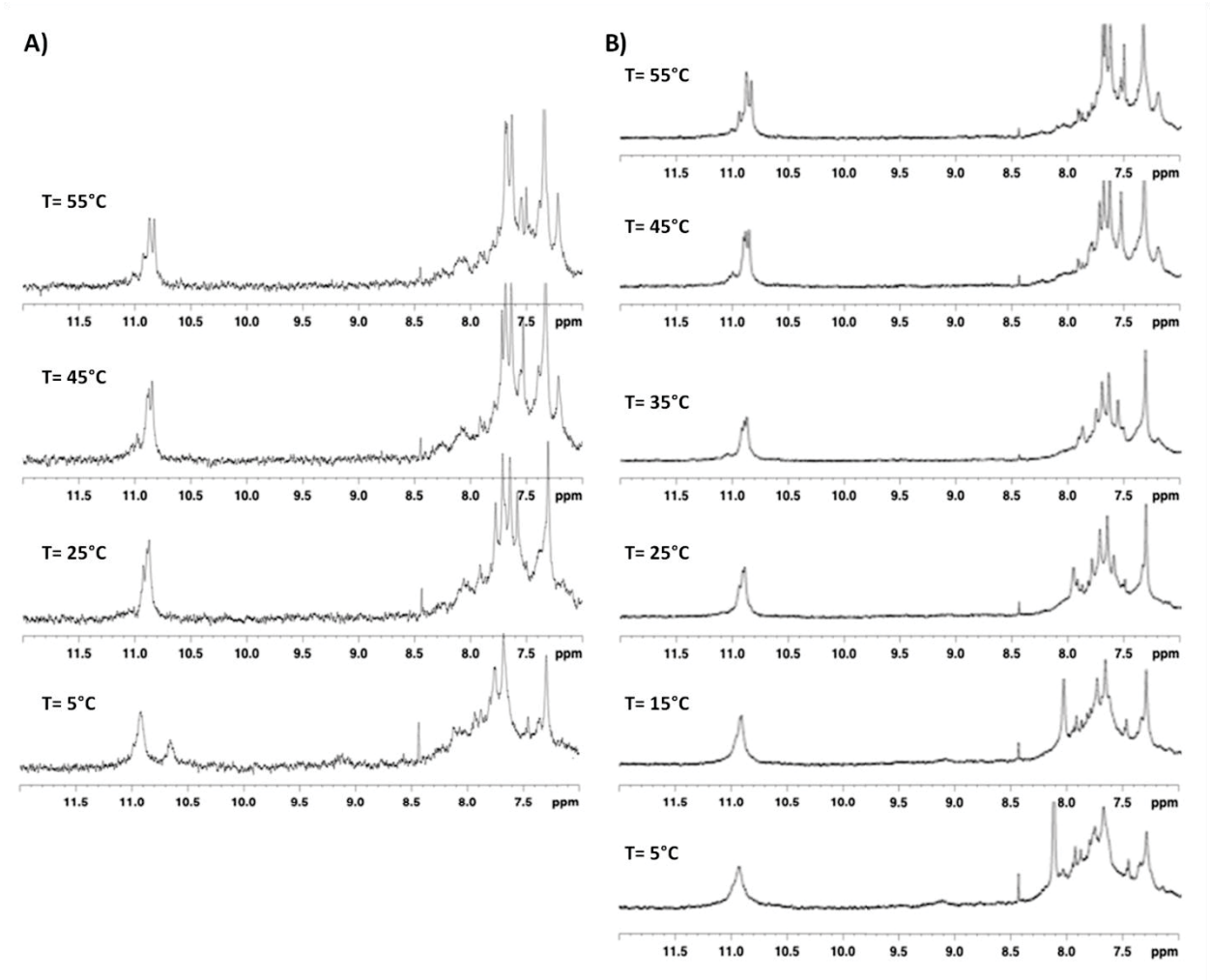


Figure 4.11 A) NMR spectra of **TXG** at 100 µM strand concentration in 10 mM Na^+ phosphate buffer, pH 7.0. B) NMR spectra of **TXG** at 500 µM strand concentration in 10 mM Na^+ phosphate buffer and 10 mM KCl, pH 7.0.

4.3.2 TINA in the Middle of the G-tract

TDS profiles for sequences with TINA in the middle of the G-tract were all similar to each other with positive peaks at 245, 277, 363 and 386 nm and negative peaks at 305 and 400 nm (Figure 4.12- Figure 4.14). It should be noted that a negative peak at 295 nm typical for the TDS profile of **TG4T** G-quadruplex is absent in these TINA-modified structures. Instead, a shoulder in the positive region at 295 nm seems to be typical for these TINA-containing complexes and for the duplex with two zipped TINAs. For some structures, lower intensities of TDS peaks were observed in the presence of KCl as a result of incomplete denaturation of G-quadruplexes at 90 °C.

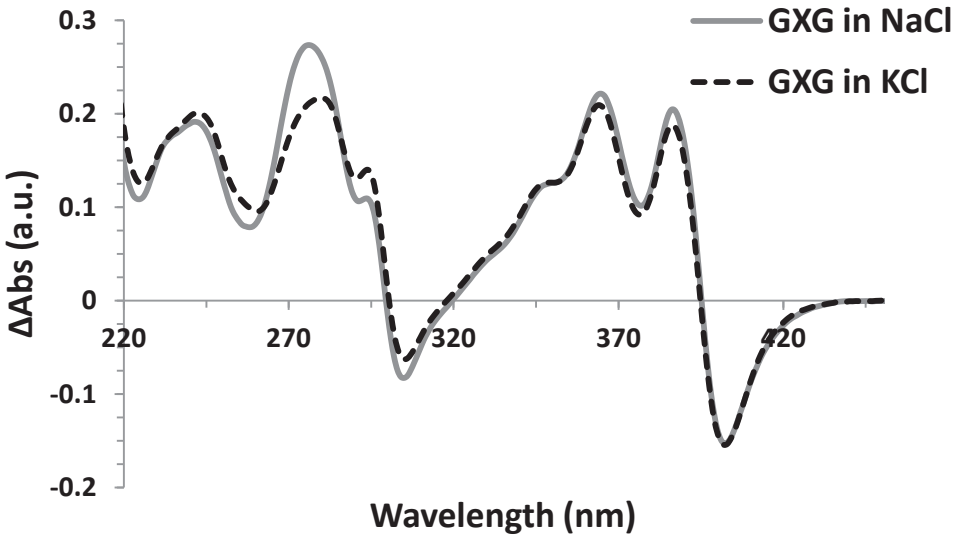


Figure 4.12 Thermal difference spectra of **GXG** at 10 μM strand concentration after annealing in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

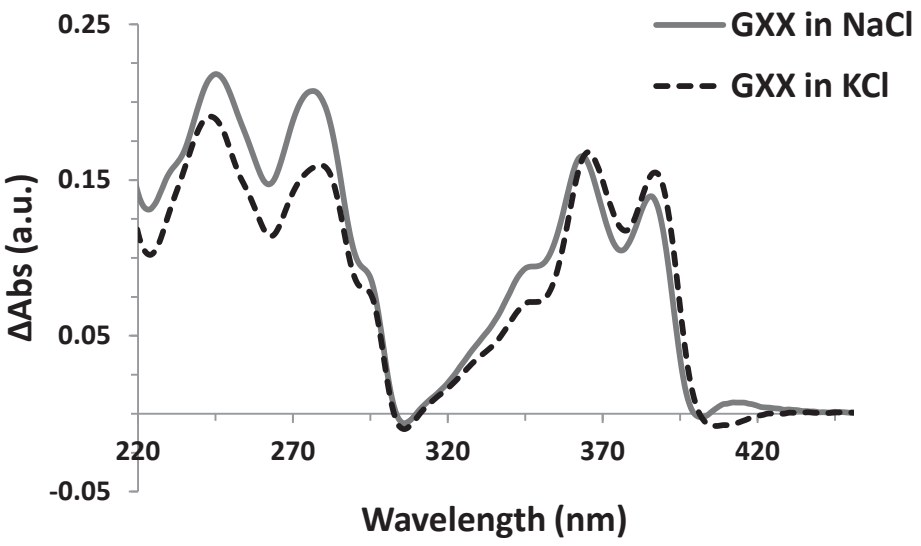


Figure 4.13 Thermal difference spectra of **GXX** at 10 μM strand concentration after annealing in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

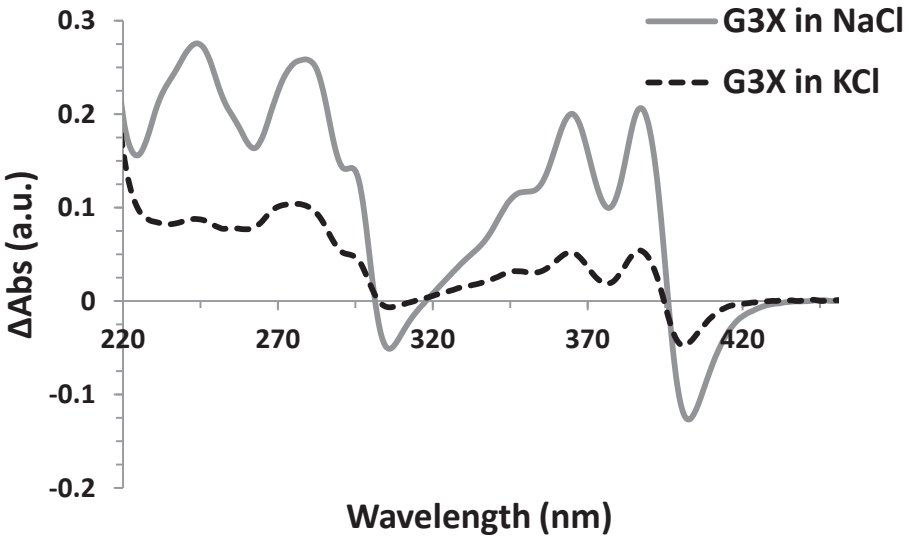


Figure 4.14 Thermal difference spectra of **G3X** at 10 μM strand concentration after annealing in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

Incorporation of thymidine into the middle of the G-tract of dTG_4T destabilizes G-quadruplexes. Only in the presence of 110 mM KCl we observed a mixture of single-stranded DNA and G-quadruplex in the native PAGE for **GTG** (Table 4.3, Figure 4.15). Previously, NMR spectroscopy confirmed formation of a T4-tetrad within the parallel G-quadruplex formed by

5′-dTGGTGGC sequence at high oligonucleotide concentration in the presence of K^+ ions (Aboul-ela, Murchie et al. 1994; Patel and Hosur 1999).

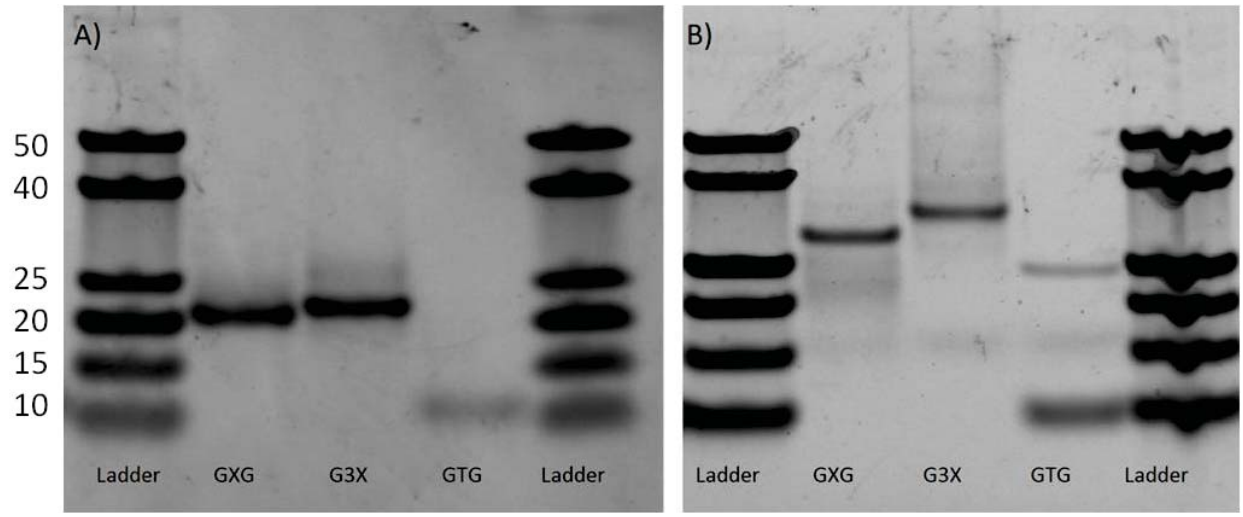


Figure 4.15 Native PAGE (20 %) analysis of oligonucleotides (100 μM) in the presence of 110 mM NaCl (A) or 110 mM KCl (B) in 10 mM Li cacodylate at pH 7.2, 37 °C. Ladder contains 10-, 15-, 20-, 25-, 40- and 50-mer oligothymidines.

UV-Vis thermal melting studies showed a marginal change in $T_{½}$ values for **TG4T** and **GXG** sequences in NaCl (Figure 4.16, Table 4.2). However, insertion of one TINA molecule into the G-tract of d(TG_4T) resulted in significantly lower $T_{½}$ value in KCl: > 90 °C for d(TG_4T) versus 56.9 °C for **GXG**. The incorporation of the second TINA molecule (**GXX**) led to a further decrease in thermal stability in the presence of both Na^+ and K^+ ions. It was interesting to note that the annealing of **GXG** and **GXX** occurred at the rates similar to the temperature ramp used (0.18 °C/min), which resulted in close $T_{½}$ values for melting and annealing (Table 4.2). When faster temperature ramps were used (0.5 and 1.0 °C/min), we observed a hysteresis between melting and annealing profiles with the difference in $T_{½}$ values ranging from 3.4 to 18 °C (Table 4.4). Unfortunately, we were unable to disrupt formation of the G-quadruplex by TINA insertion in the dTG_6T sequence (**G3X**, Table 4.3). Tetramolecular complexes were confirmed for **GXG** and **G3X** sequences using native PAGE, whereas a smear with significantly lower retardation on the gel was observed for **GXX**, which indicated the formation of aggregates with undefined stoichiometry.

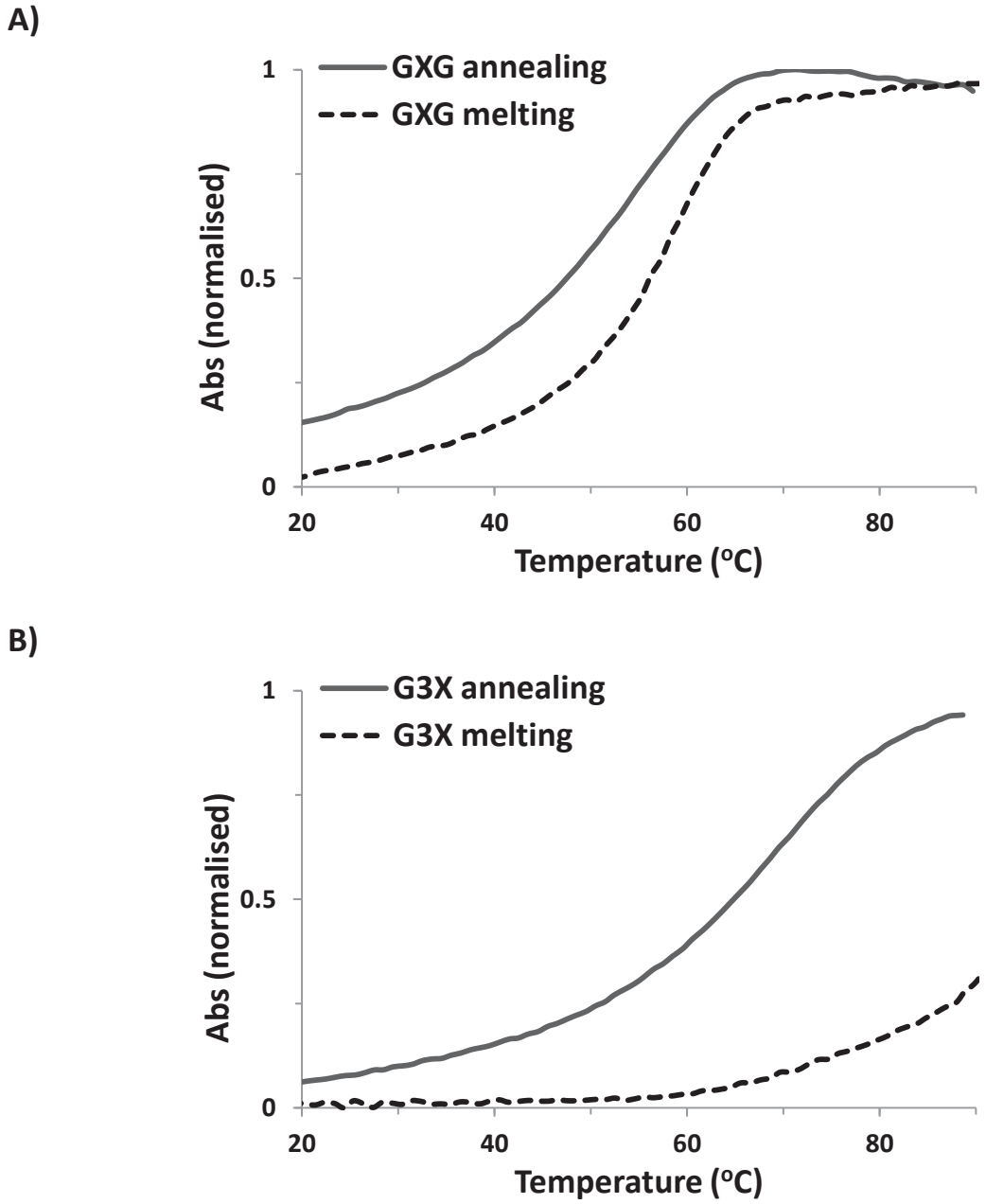


Figure 4.16 Melting and annealing profiles of **GXG** (A) and **G3X** (B) at 10 μM strand concentration after incubation in 110 mM NaCl, 10 mM Li cacodylate buffer at pH 7.2. The profiles are based on absorbance data recorded at 373 nm with a 0.18 °C/min rate.

Table 4.4 $T_{½}$ values of **GXG**, **GXX** and **G3X** obtained during melting and annealing, using 0.5 or 1.0 °C/min temperature ramp in the presence of 110 KCl in 10 mM Li cacodylate buffer at pH 7.2.

	0.5 °C/min melting/annealing (°C)	1.0 °C/min melting/annealing (°C)
GXG	59.4 / 47.8	62.6 / 44.6
GXX	41.2 / 36.8	41.9 / 38.5
G3X	>90 [a] / 41.5	>90 [a] / 41.5

[a] melting process was complete only after 30 min incubation at 90 °C.

In contrast to the CD profiles described in the previous section, CD spectra for **GXG** and **GXX** are different and dependent on the cation present in the solution. In the presence of NaCl, CD profiles in the UV region for both sequences are typical for antiparallel G-quadruplex structures with maxima at 294 and 244 nm, and a minimum at 265 nm (Figure 4.17 - Figure 4.19). However, we cannot rule out the possibility that the appearance of antiparallel signatures in these CD spectra results from a different guanosine alignment in a complex with all strands being in a parallel orientation as a consequence of TINA insertion in the sequence. On the other hand, CD spectra in 110 mM KCl suggest the presence of G-quadruplexes with both parallel and antiparallel topologies because the above mentioned positive ellipticities are seen as shoulders/peaks in the profile of the dominant parallel G-quadruplex. The dissimilarity in the CD profile of **GXG** and **GXX** complexes in the absorption region of TINA indicates different pyrene-pyrene interactions (Figure 4.17, Figure 4.18).

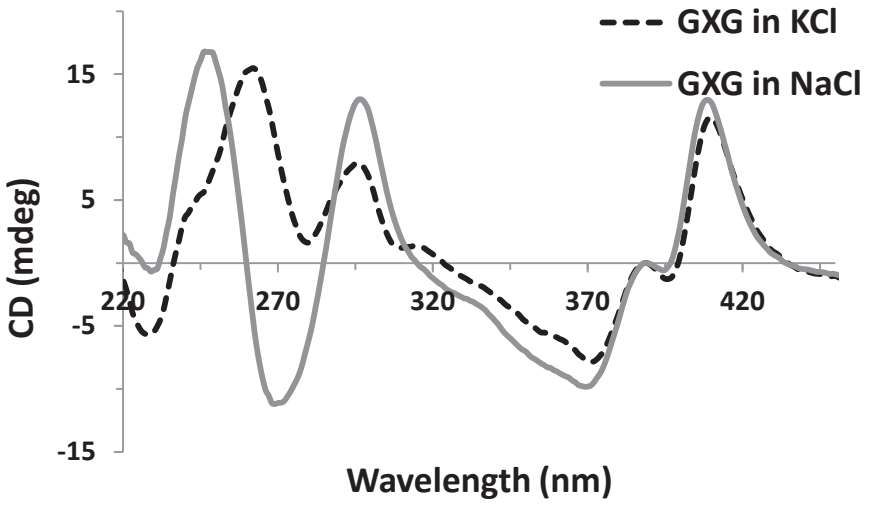


Figure 4.17 CD spectra of **GXG** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

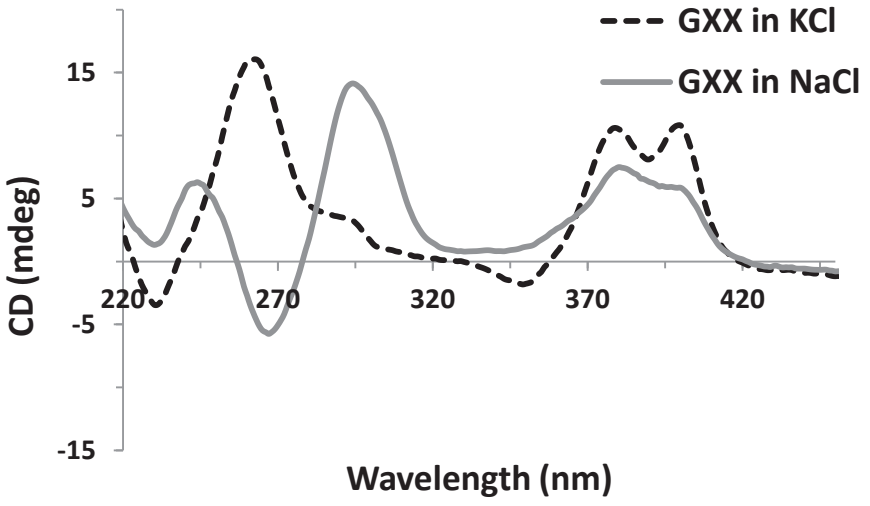


Figure 4.18 CD spectra of **GXX** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

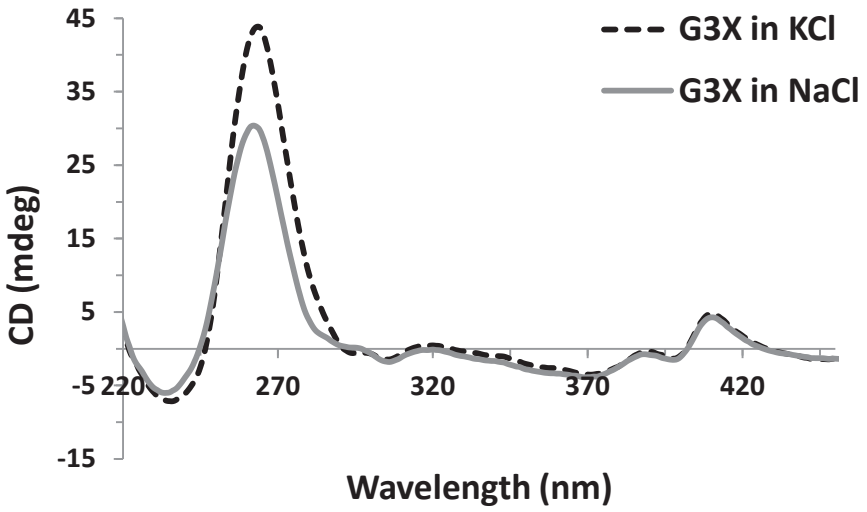


Figure 4.19 CD spectra of **G3X** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

Differences between G-quadruplexes formed by **GXG** and **GXX** can also be seen in the fluorescence spectra as the maximum for the excimer band is shifted from 490 nm in **GXG** to 510 nm in **GXX** irrespective of the cation used (Figure 4.20). It is interesting to compare fluorescence spectra of G-quadruplexes with TINA in the middle of the G-tract (**GXG**, **GXX**, **G3X**) and duplexes having two TINAs in the middle of the sequence arranged in a zipper fashion (**TINA duplex**). The absence of such an excimer band in this duplex is thought to be the result of a poor overlap of pyrenes inside the duplex (Filichev, Astakhova et al. 2008). In contrast to that, all TINA-G-quadruplexes exhibit a strong excimer band with slightly different λ_{em}. This suggests the G-quadruplex can be a scaffold for the arrangement of organic chromophores not only at the 5' or 3' - ends but also in the middle of the sequence.

Elongation of the G-tract in **G3X** resulted in the formation of thermally stable ($T_{½}$ > 90 °C,Table 4.2) parallel G-quadruplexes both in the presence Na^+ or K^+ ions (Figure 4.19). However, the CD and fluorescence spectra in the TINA region of **G3X** are similar to those of **GXG** (Figure 4.19, Figure 4.20).

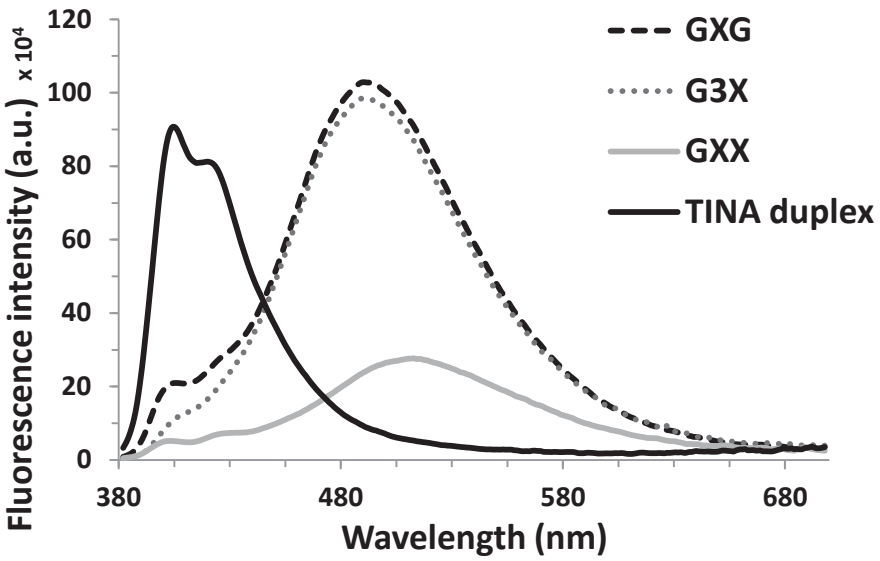


Figure 4.20 Fluorescence emission spectra of **GXG**, **G3X**, **GXX** and **TINA duplex** at 10 μM strand concentration in the presence of 110 mM NaCl in 10 mM Li cacodylate buffer at pH 7.2, λ_{ex} = 373 nm.

The striking difference in CD profiles for **GXG** in Na^+ and K^+ led us to investigate this structure using ^{1}H NMR spectroscopy. To our surprise we observed at least 12 distinct imino signals spread between 9.3 and 11.3 ppm in 100 mM NaCl (Figure 4.21). This indicates the formation of a non-symmetric, probably antiparallel G-quadruplex. In the presence of 100 mM KCl the number of signals in the imino region was reduced and these broadened signals were in line with the observation of the possible existence of mixed topologies in the CD spectrum. Further structural characterisation of **GXG** in NaCl is necessary and will be performed in collaboration with an expert in NMR spectroscopy – Prof. Carlos Gonzalez from Instituto de Quimica Fisica Rocasolano, CSIC, Madrid, Spain.

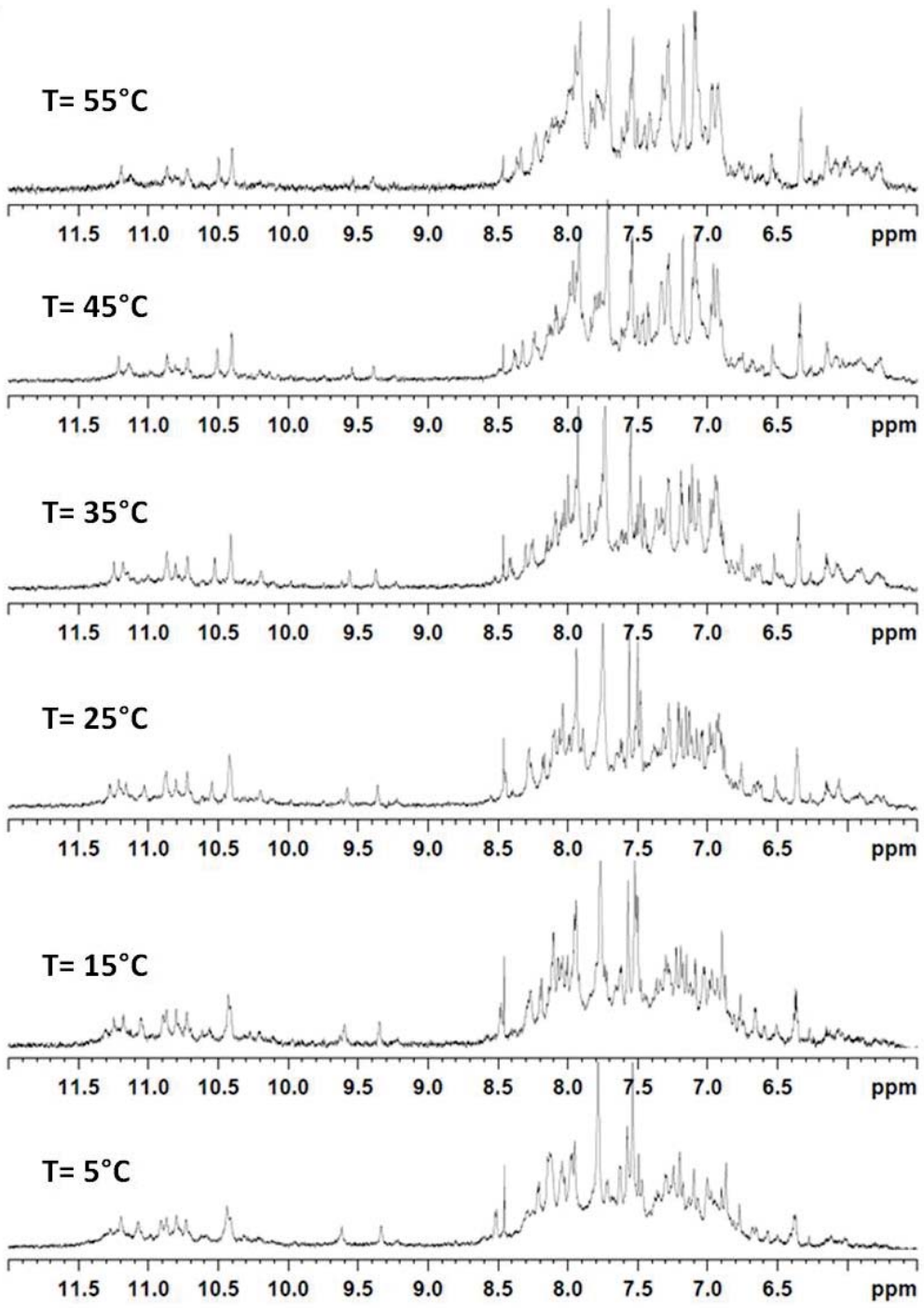


Figure 4.21 NMR spectra of **GXG** at 1.0 mM strand concentration in 10 mM Na^+ phosphate buffer, pH 7.0.

^{1}H NMR spectra for **GXX** was also recorded. As a result of low thermal stability and formation of multimeric species we did not observe any signals in the imino region of **GXX** even in the presence of 100 mM KCl.

4.3.3 TINA Accelerates Association of Tetramolecular G-quadruplexes

It was surprising to see the formation of different structures by **GXG** and **G3X** sequences under identical conditions, e.q. 110 mM NaCl in Li-cacodylate buffer, pH 7.2. To get further insights into the assembly of these sequences we performed kinetic studies as described previously in literature (Mergny, De Cian et al. 2005). During isothermal renaturation experiments we recorded absorbance at different wavelengths as a function of time at constant temperature (Figure 4.22). The association rate constants, k_{on}, were calculated using the following equation;

$$Abs_t = Abs_f + (Abs_i - Abs_f) \times (1 + (n-1) \times k_{on} \times C^{n-1} \times t)^{\frac{1}{1-n}} \qquad \text{(Formula 4)}$$

Abs_i and Abs_f represent the initial and final absorbance values, respectively. C represents the strand concentration, *t* represents the time, *n* represents the molecularity of the folded nucleic acid structure. We found that the association rate constant, k_{on}, was significantly higher for **GXG** in comparison to that earlier reported for **TG4T** (Table 4.5). It is interesting that k_{on} values of **GXG** and dTG_6T (**TG6T**) with a longer G-tract are of the same magnitude (10^{11} M^{-3} s^{-1}). TINA inserted in the middle of the **TG6T** sequence (**G3X**) had a marginal effect on the k_{on} value in comparison to the unmodified sequence. For both **GXG** and **G3X** constructs the order of the reaction was experimentally estimated to be 3.97, which is very close to values reported in the literature for tetramolecular G-quadruplexes (3.4 – 4.1) (Mergny, De Cian et al. 2005).

Table 4.5 Sequences of various ONs and their k_{on} values obtained in NaCl (110 mM).

	Sequence 5′-3′	k_{on} (M^{-3} s^{-1})
TG4T	dTGGGGT [b]	~2.4 x 10^7
GXG	dTGG**X**GGT [a]	(2.73 ± 0.20) x 10^{11}
TG6T	dTGGGGGGT [b]	~1.0 x 10^{10}
G3X	dTGGG**X**GGGT [a]	(2.28 ± 0.15) x 10^{11}

[a] '**X**' refers to TINA. k_{on} values of modified sequences were experimentally obtained in Li cacodylate buffer (10 mM) and NaCl (110 mM) at 20 °C, pH 7.2. [b] k_{on} values of unmodified sequences at 20 °C, pH 7 estimated from data reported previously (Wyatt, Davis et al. 1996; Mergny, De Cian et al. 2005) .

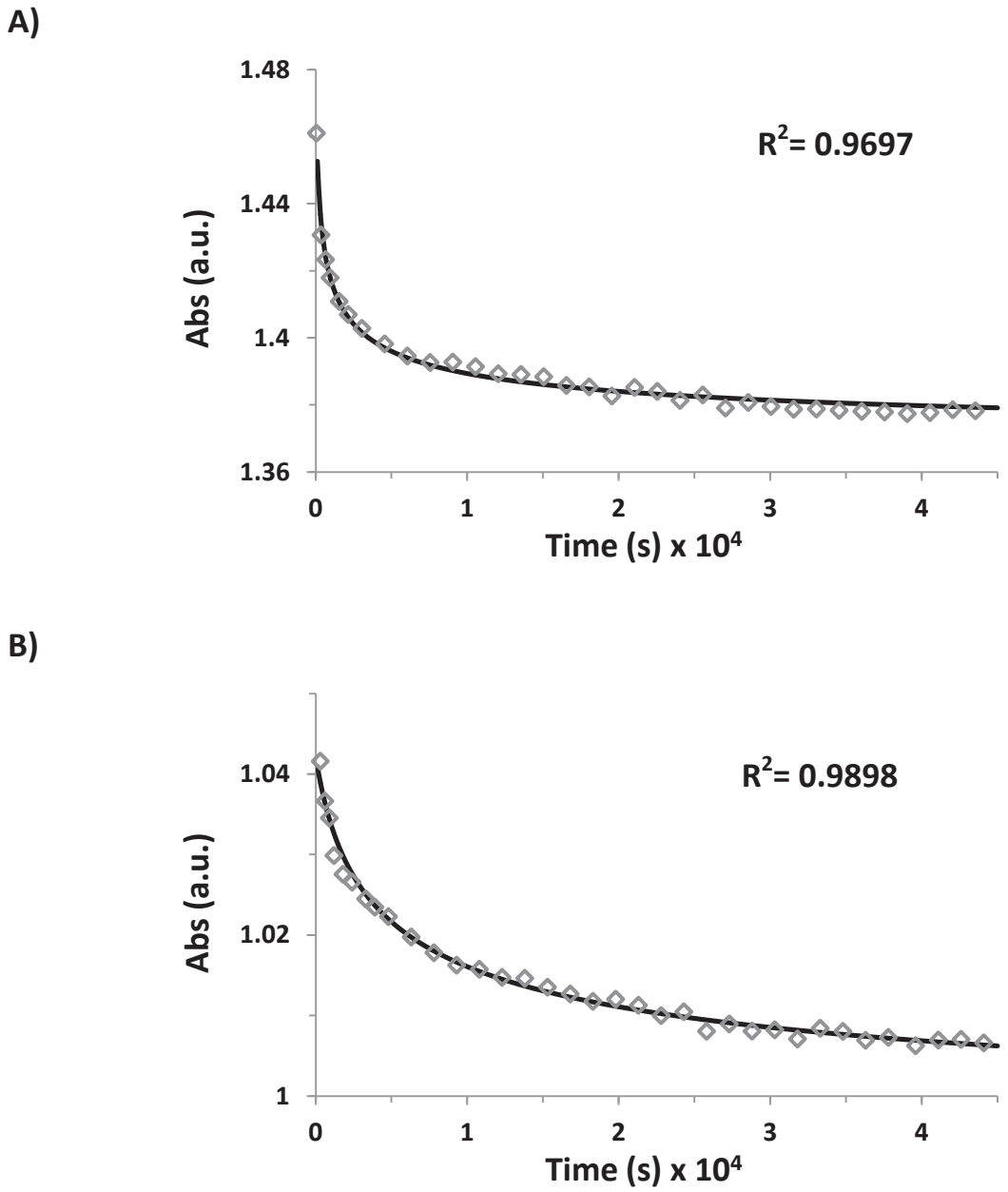


Figure 4.22 Isothermal renaturation processes of **GXG** (20 μM, A) and **G3X** (10 μM, B) recorded at 275 and 280 nm, respectively, in the presence of 110 mM NaCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C. For experiment details see Chapter 8.12.

4.3.4 Effect of 3' Fluorescein Modification

The addition of fluorescein at the 3'-end of TINA-TG_4T sequences destabilised G-quadruplexes in NaCl by 6.1 and 3.8 °C for **TXGF** and **GXGF**, respectively (Table 4.2). An opposite effect

was observed in the presence of KCl for the G-quadruplex derived from **GXGF** as it was more thermally stable than **GXG** ($\Delta T_{½}$ = 5.6 °C). A notably slower annealing was seen for both fluorescein-containing sequences in comparison with TINA-TG_4T, except for **TXGF** in KCl as this complex did not melt. With the addition of the fluorescein comes an extra negatively charged phosphate which can account for the slow G-quadruplex formation. The mobility in the native gel was much slower for **TXGF** in comparison with the sequence with no fluorescein, while the retardation of **GXG** and **GXGF** was very similar (Table 4.3). It is interesting to note that the CD profile in the UV region was marginally changed after fluorescein addition to the **GXG** sequence (Figure 4.17, Figure 4.23). However, the peak intensities for **GXGF** in the TINA absorption region were reduced in KCl solution in comparison to **GXG**. Complexes formed by **TXG** and **TXGF** shared common features in CD spectra in KCl, but the profile in the TINA region was different in NaCl (Figure 4.10, Figure 4.24).

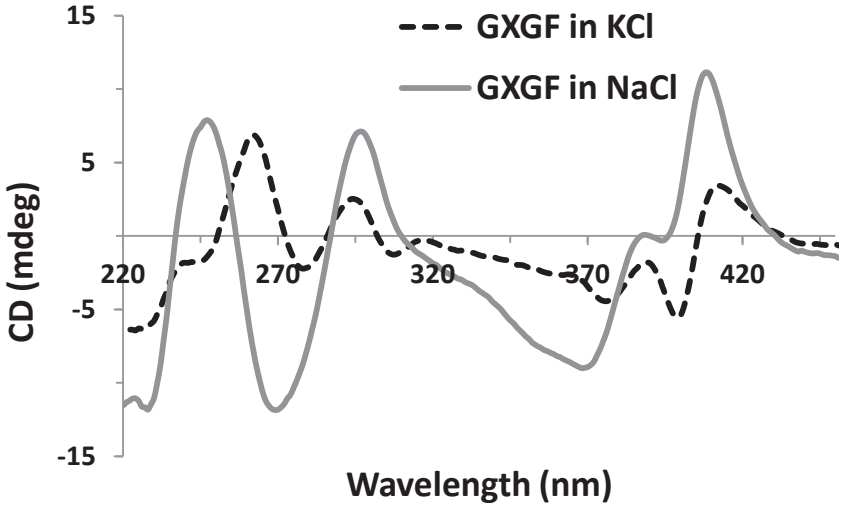


Figure 4.23 CD spectra of **GXGF** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer, pH 7.2.

In the fluorescence emission spectra, the signal originating from the pyrene to fluorescein energy transfer was observed at 520 nm for both **TXGF** and **GXGF** G-quadruplexes, while the pyrene excimer band diminished (λ_{ex} = 373 nm, Figure 4.25). During thermal melting an increase in emission of fluorescein was observed (λ_{ex} = 495 nm, Figure 4.26) indicating the separation of

self-quenching fluorescein molecules, which is in line with previous experiments (Merkina and Fox 2005).

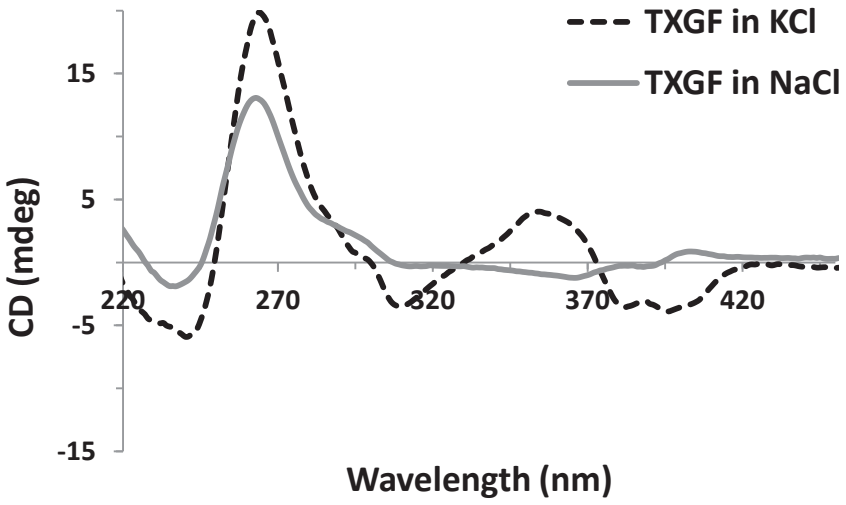


Figure 4.24 CD spectra of **TXGF** at 10 μM strand concentration in the presence of 110 mM KCl (dashed line) or NaCl (grey line) in 10 mM Li cacodylate buffer at pH 7.2, 20 °C.

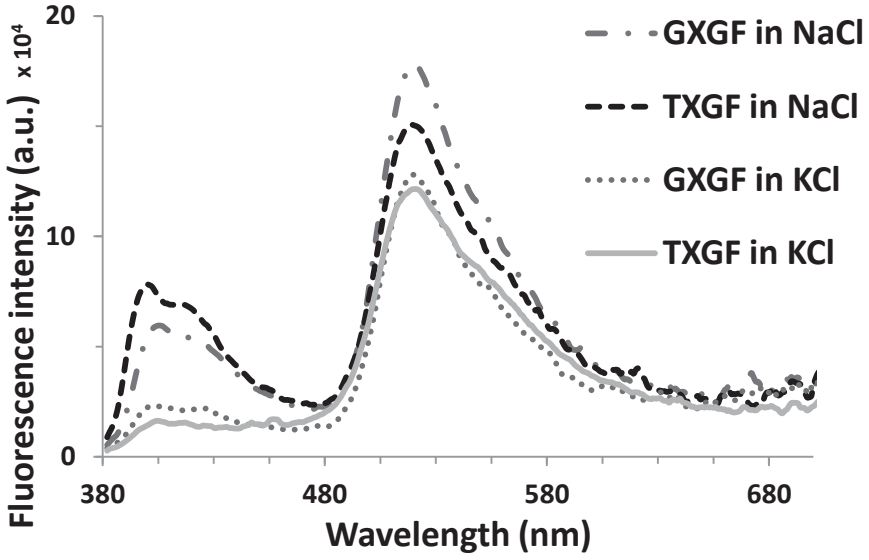


Figure 4.25 Fluorescence emission spectra of **GXGF**, **TXGF** at 10 μM concentration in the presence of 110 mM NaCl or KCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C. λ_{ex} = 373 nm.

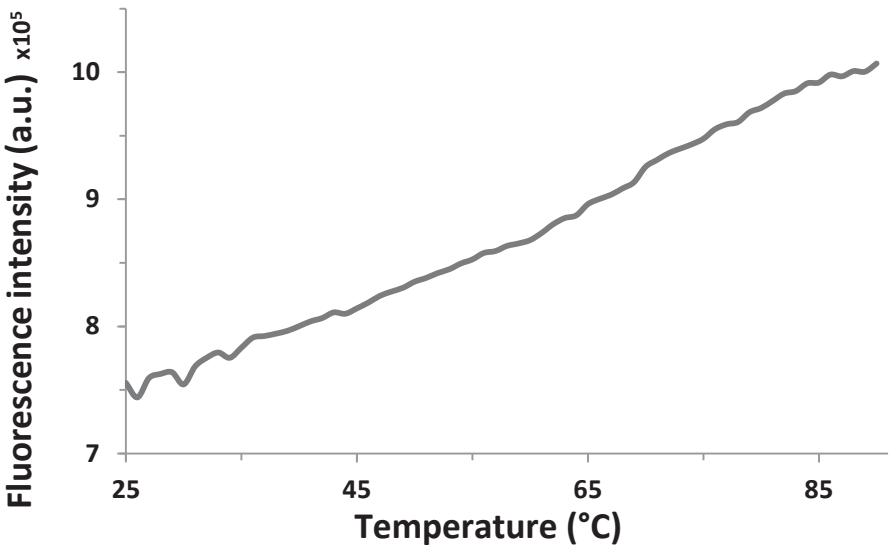


Figure 4.26 The intensity of fluorescence emission at 520 nm during thermal melting of **GXGF** (10 μM) with 0.2 °C/min rate in the presence of 110 mM NaCl in 10 mM Li cacodylate buffer at pH 7.2, 20 °C. λ_{ex} = 465 nm.

4.4 Discussion

The *Tetrahymena* telomere sequence is a well-investigated model for parallel tetramolecular G-quadruplexes. The effect of nucleobase and sugar modifications on the G-quadruplex formation as well as the influence of organic chromophores conjugated at the 3'- and 5'-ends have been studied (Esposito, Virgilio et al. 2005; Merkina and Fox 2005; Gros, Rosu et al. 2007). Here, we focused on the alteration of the tetramolecular G-quadruplex caused by the TINA molecule.

Previous studies on G-quadruplex forming oligonucleotides containing TINA monomers outside of the guanine tract have shown an increase in thermal stability of G-quadruplexes. Similarly, both **XTG** and **TXG** sequences showed a significant increase in $T_{½}$ values in comparison to the wild-type sequence. However; **XTG** which contains TINA monomer (**X**) at the 5'-end of the sequence formed a G-quadruplex dimer according to native gels in K^+, but this was not the case for **TXG**. The higher-order aggregation observed for **XTG** was avoided if thymidine was present at the 5' end as in **TXG**. This led instead to assemblies with stoichiometries between 2 and 4. The situation was reversed for **TXG** after addition of fluorescein at the 3'-end, which resulted in formation of complexes with low mobility in the native gel (above T_{50}). This clearly suggests

that organic dyes present at both ends of a dTG_4T sequence have a propensity to trigger formation of high-order structures especially in the presence of K^+ ions.

Placing a TINA molecule in the middle of the G-tract led to significant G-quadruplex destabilisation especially in 110 mM K^+ ions but not to the structure demolition as seen for the sequence with thymidine in the middle of it, **GTG**. This is surprising as our previous experience with intra- and intermolecular G-rich triplex-forming oligonucleotides (TFOs) showed a strong destabilising effect upon TINA insertion in the middle of the G-tracts. At the same time we observed a dramatic change in the structure of **GXG** depending on the cation present in solution: in the presence of 110 mM NaCl the CD spectrum shows a characteristic antiparallel G-quadruplex signature whereas in the presence of 110 mM KCl the CD spectrum shows features of both parallel and antiparallel structures. Such ion-dependent structural transitions are commonly encountered for *intra*molecular G-quadruplexes, but to the best of our knowledge such a transition has not been observed in *tetra*molecular G quadruplexes.

Similar $T_{½}$ values were detected for **GXG** in NaCl and KCl, and that complexes were formed with a speed close to the temperature ramp of 0.18 °C/min. This behaviour was different to that observed for the wild-type sequence and other TINA-modified sequences. This can be accounted for by the intermolecular interactions of hydrophobic TINA molecules which initiate and guide the assembly of **GXG** and **GXX**. This is clearly seen in NaCl solutions, which results in complexes with topologies different from the parallel tetramolecular G-quadruplex $[d(TG_4T)]_4$. The addition of K^+ ions is known to accelerate the assembly of wild-type G-quadruplexes which competes with TINA-mediated association and, as a result, G-quadruplexes of mixed topologies are formed. This is in agreement with the results from the sequence having an extended guanine tract and TINA in the middle, **G3X**. Previously, it has been shown that elongation of the G-tract by one guanosine leads to a 10-fold larger association rate constant (Mergny, De Cian et al. 2005). The tetramolecular G-quadruplex formed by **G3X** had a parallel topology indicating the preference of guanine-mediated association. Kinetic results support the hypothesis: the association rate constant is comparable for **GXG**, $\mathbf{TG_6T}$ and **G3X** in NaCl solutions (Table 4.5). TINA controls the assembly of the **GXG** which leads to a new G-quadruplex topology but cannot or barely competes with G-mediated association when inserted in the sequence with a longer G-tract or in KCl solutions. These results provide new perspectives in the design of G-

quadruplexes governed by molecules different from guanine – this can potentially expand the G-quadruplex alphabet.

It was interesting to note the difference that the 3′-fluorescein addition makes to the annealing of TINA-TG_4T complexes, which is significantly inhibited. This can be explained by extra negative charges localised on the phosphate and fluorescein, which increase electrostatic repulsion between DNA strands, which is unfavourable for the complex assembly. This effect is in agreement with earlier observations that the addition of extra thymidines, *i.e.* internucleosidic phosphates, at the ends of d(TG_4T) are detrimental to the association rate (Mergny, De Cian et al. 2005). This can also explain our previous results on modification of the triplex-forming oligonucleotide, 5′-d$AG_6T_4GT_4$, designed to target a polypurine region of HIV-1 proviral DNA. A single insertion of TINA in the G-tract, 5′-dAGGG**X**GGGTTT**X**TGTTTT, was enough to destabilise a very stable parallel tetramolecular G-quadruplex ($T_{½}$ > 90 °C in 150 mM NaCl). The modified strand migrated as single-stranded DNA in 150 mM NaCl but existed as a dimer in 150 mM KCl in native gels. In contrast, **G3X** sequence studied here formed a stable G-quadruplex. Analysis of TINA-modified G-rich TFOs described so far in the literature has led to the following conclusions: 1) G-rich ONs capable of forming undesired G-quadruplexes are more vulnerable to TINA insertions in the middle of the G-tracts if the overall G-content decreases and 2) in cases with similar percentage of Gs in the G-rich ONs, it is easier to disrupt G-quadruplex formation if there are several G-tracts and not a single long G-tract.

4.5 Summary

In this chapter we have investigated the introduction of TINA monomers into the dTG_4T sequence which is known to form a highly stable tetrameric G-quadruplex with all strands in a parallel orientation. We have observed that the effect of TINA monomers strongly depends on their position in the sequence. Outside of the G-tract TINA-incorporated G-quadruplexes showed increased thermal stability in comparison to the wild-type structure.

For the first time we studied the structural effects of the organic chromophore, pyrene in a TINA when inserted into the middle of a sequence that forms a highly stable tetramolecular G-quadruplex. We observed that the insertion of TINA monomers into the middle induced transformation of parallel G-quadruplexes into antiparallel G-quadruplexes. In the past, the

control over strand orientation had been shown only by backbone modifications such as inversion sites or abasic sites (Esposito, Virgilio et al. 2005; Esposito, Virgilio et al. 2009; Skolakova, Bednarova et al. 2010). In this respect, for the first time an organic chromophore conjugated to a G-quadruplex is shown to regulate G-quadruplex formation.

We observed that the association of **GXG** was mediated by two different factors. Whereas guanine induced the formation of parallel G-quadruplexes, TINA monomers promoted formation of antiparallel G-quadruplexes. Overcoming one factor by the other is what determines the overall topology. We observed that the topology can be controlled by the guanine/TINA ratio as seen in **G3X** where higher guanine content resulted in the formation of parallel G-quadruplexes. In the next step more structural studies will be performed on **GXG** to provide further insights into the effect of TINA monomers on the topology of this complex. This study presents the potential of TINA modification to induce alternative structures and increases the repertoire of possible structures that can be formed using TINA monomers in G-rich sequences.

Chapter 5. Incorporation of Minor Groove Binders to TINA-Conjugated TFOs

5.1 Introduction

A set of artificial molecules were discovered by Dervan et al. for recognition of DNA (Wade, Mrksich et al. 1992). When polyamides containing *N*-methylimidazole (**Im**) and *N*-methylpyrrole (**Py**, Figure 5.1) amino acids were combined in antiparallel, side-by-side dimeric complexes, they were shown to be capable of recognising dsDNA through hydrogen bond donor/acceptor groups presented by nucleobases in the minor groove. For this reason these polyamides are refered as minor groove binders (MGBs, Figure 5.2). The specificity of sequence recognition can be controlled by modification of the linear sequence of pyrrole and imidazole amino acids. In order to recognise a **G-C** pair, an imidazole/pyrrole (**Im/Py**) alignment, to recognise a **C-G** pair, a pyrrole/imidazole (**Py/Im**) alignment was used. For **A-T** or **T-A** pairs, a a pyrrole/pyrrole (**Py/Py**) alignment was determined to be adequate (Mrksich, Wade et al. 1992; White, Baird et al. 1997). These polyamides were covalently linked to each other by γ-aminobutyric acid (**γ**, Figure 5.1) resulting in a hairpin structure that showed an increase in affinity for dsDNA of ~100 fold. Later, it was also found that an additional tail containing β-alanine (**β**) and (*N*,*N*-dimethylamino)propylamide (**Dp**) would provide further stability and the ability to extend the recognition length (Mrksich, Wade et al. 1992; Wade, Mrksich et al. 1993; Geierstanger, Jacobsen et al. 1994; Geierstanger, Mrksich et al. 1994; Mrksich, Parks et al. 1994).

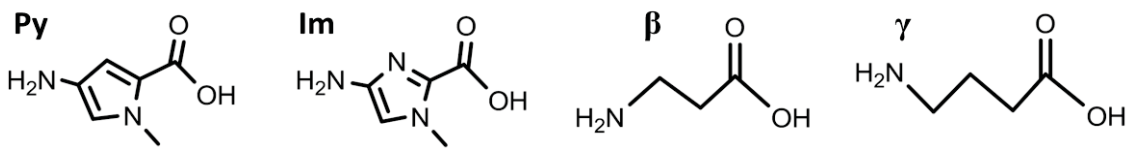


Figure 5.1 Structures of *N*-methylpyrrole (**Py**), *N*-methylimidazole (**Im**), β-alanine (**β**) and γ-aminobutyric acid (**γ**).

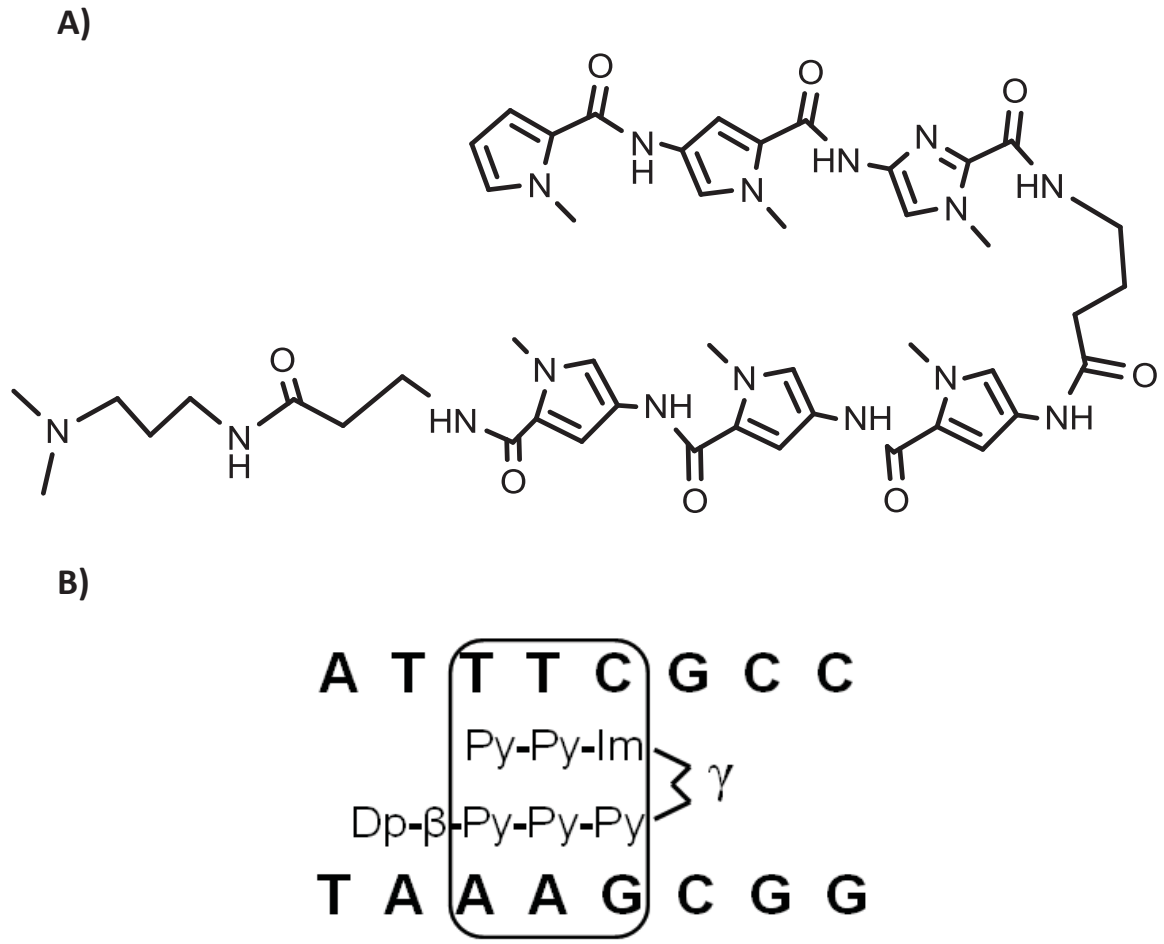


Figure 5.2 Structure of the polyamide hairpin minor groove binder, **Py_2Im-γ-Py_3-β-Dp** (A) and its interaction with double-stranded DNA sequence (B). **Dp** is (*N*,*N*-dimethylamino)propylamide; **β** is β-alanine; **γ** is γ -aminobutyric acid.

Recognition length of a single hairpin MGB is limited to 3 - 5 base-pairs; however, a covalent attachment of two hairpins *via* a flexible linker was shown to provide simultaneous recognition of two MGB binding sites with an increased length of the recognition site (Trauger, Baird et al. 1998; Dervan and Burli 1999; Halby, Ryabinin et al. 2005). These complexes that contain two separate polyamide chains are referred as bis-MGB. It was also found that the sequence specificity decreases as the recognition length increases, while binding affinity is maximised at a continuous residue number (**Py** or **Im**) of five (Kelly, Baird et al. 1996).

5.1.1 MGB-TFO Conjugates

Even though the MGBs show very strong affinity to target DNA, they have limited sequence recognition due to structural restrictions. Many modifications have been tried to expand the length of the binding site. As one of these modifications, conjugates of MGBs and TFOs have been introduced (Szewczyk, Baird et al. 1996). It was shown that short MGB-TFOs can bind to dsDNA and give more stable DNA complexes than DNA triplexes alone. This means that stronger binding of MGB-TFOs allows for more stringent hybridisation conditions to be used for TFO probes.

On the other hand, it was shown that, at basic pH, bis-MGB conjugates of parallel TFOs were attached to the target DNA only via MGBs, leaving the TFOs ineffective. This means that the binding was carried out only by the bis-MGB and TFO did not bind to the target. In the case of a single MGB conjugated to a TFO (mono conjugate), high specificity was achieved, even though affinity was low at temperatures above 37 °C and pH over 6 (Sinyakov, Ryabinin et al. 2001; Boutorine, Ryabinin et al. 2003; Halby, Ryabinin et al. 2007). This showed that low specificity of parallel TFOs, especially at higher pH, was an obstacle to successful MGB-TFO conjugates. On the other hand, no studies have been undertaken focused on conjugation of MGBs with pH insensitive, antiparallel TFOs.

5.1.2 Aim

In this chapter we combine MGBs with antiparallel TINA-TFOs, which were discussed in Chapter 3, in order to extend the length of recognition site and expand it into non-homopurine sequences. TFOs **3** and **7** from Chapter 3 were identified as the TINA-TFOs with highest efficiency in terms of target affinity. Here we propose to conjugate these TFOs with MGBs that can recognise the sequence of a neighbouring TFO recognition site. New conjugates are expected to be capable of recognising a longer sequence beyond the polypurine tract.

5.2 Probe Design

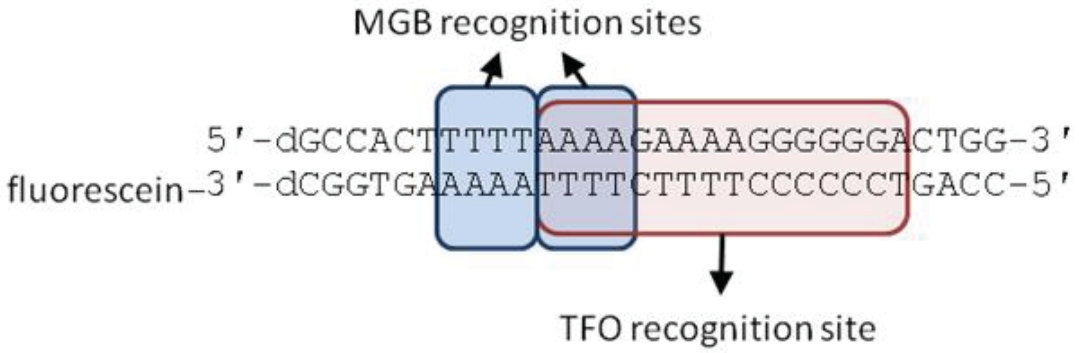


Figure 5.3 The recognition sites of TFO and MGB on the duplex, Target **I**.

A 4 base-pairs long polyT-polyA tract flanking the 5′ of the polypurine TFO target represents a potential target for MGB. At the same time, a polyA-polyT tract can be found within the recognition site of the TFO. Both of these sites are potentially acceptable by MGBs since TFOs bind in the major groove, leaving the minor groove available for MGB binding.

In order to combine the properties of MGB and TFO in one molecule a bridge is necessary to crossover from minor to major groove. For this reason hexaethyloxy-glycol $(EG)_6$ was chosen as a linker for covalent attachment of hexamethylpyrroles (MGB) at the 3′ terminal phosphate. This linker is commercially available and does not interfere with phosphoramidite chemistry. Sequence of the MGB is determined as **Py_3-γ-Py_3-βDp** since this sequence is capable of binding both MGB recognition sites on target **I** (Novopashina, Sinyakov et al. 2003).

Table 5.1 Probes and their sequences.

ON	Sequence 5′ - 3′[a]
MGB-TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT-$(EG)_6$p-**Py_3-γ-Py_3-βDp**
p$(EG)_6$-TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT-$(EG)_6$p
MGB-TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT-$(EG)_6$p-**Py_3-γ-Py_3-βDp**
p$(EG)_6$-TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT-$(EG)_6$p

[a] '**X**' represents TINA monomer; $(EG)_6$ is a hexaethylene-glycol linker; p is a terminal phosphate.

5.2.1 Solid-phase Synthesis of Minor-Groove Binders

Solid-phase peptide Boc-chemistry was applied in the synthesis of these polyamides. For the elongation of the chain, the C-termini of the polyamides were attached to commercially available Boc-β-Pam-resin. The resin provided the ability to cleave the polyamides after the synthesis, in a single step. A polyamide synthesis cycle commenced with deprotection with TFA/phenol/water mixture (92.5:5:2.5) and continued with washing (DCM/DMF), coupling and washing steps (Table 5.2). The synthesis was concluded by aminolysis of the chain from the column by (*N*,*N*-dimethylamino)propylamine treatment (Baird and Dervan 1996; Krutzik and Chamberlin 2002). Synthesis of MGB with **Py_3-γ-Py_3-βDp** sequence was performed by Prof. A. Boutorine in Muséum National d'Histoire Naturelle (MNHN, Paris).

Table 5.2 Standard protocol for manual solid-phase synthesis of pyrrole-imidazole polyamides.

Steps	Reagent	Incubation time/mode
Deprotection	2 × TFA, phenol, water (92.5:5:2.5)	1+2 min mix
Washing	DCM	30 s flow
	DMF	1 min flow
Coupling	Activated monomer	20 min mix
Washing	DMF	30 s flow
	DCM	1 min flow

5.3 Synthesis of Probes

Two TFOs were synthesised as described in Chapter 2 using the automated DNA synthesiser. Additionally hexaethylene glycol and 5′-phosphate were attached to the CPG columns prior to ON synthesis by hand-coupling using commercially available DMA-hexaethyloxy-glycol phosphoramidite and 5′-phosphate reagent. ONs synthesised were referred as $p(EG)_6$-TFOs. Then, the CPG support was removed from the column and the ONs were cleaved from the CPG support and deprotected by incubation in 32 % NH_4OH at 55 °C. After freeze-drying the ONs were purified using reverse-phase HPLC. Fractions were treated with 80 % *aq.* AcOH solution (20 min). After quenching with 3 M NaOAc (100 μL), the fractions were precipitated from

EtOH. The ONs were dissolved in 100 µL H_2O at 55 °C for 30 min and their purity was confirmed on 20 % denaturing PAGE (7 M urea).

5.3.1 Conjugation with TFOs

A conjugation reaction between TFOs and MGBs was carried out by activation of the 3′ terminal phosphate of the oligonucleotides with triphenylphosphine and dipyridylsulfide (Figure 5.5). Because this reaction is water sensitive, ONs were treated with 8 % *aq.* cetyl trimethylammonium bromide (CTAB) until all TFOs were precipitated from water and dried *in vacuo* prior to the reaction. After 15 min activation of the terminal phosphate, the dry TFA salt of the MGB was dissolved in DMSO in the presence of triethylamine and activated TFOs.

After 16 hours of coupling, MBG-TFOs were precipitated with 3 % $LiClO_4$ in acetone. The efficiency of the MGB-TFO coupling reaction was evaluated using denaturing 20 % PAGE. As a consequence of this reaction two MGBs may be attached to a terminal phosphate instead of one. Denaturing gel shows that a small portion of bis-conjugate was also produced after 2 or 16 hours of reaction (Figure 5.4 - Figure 5.6).

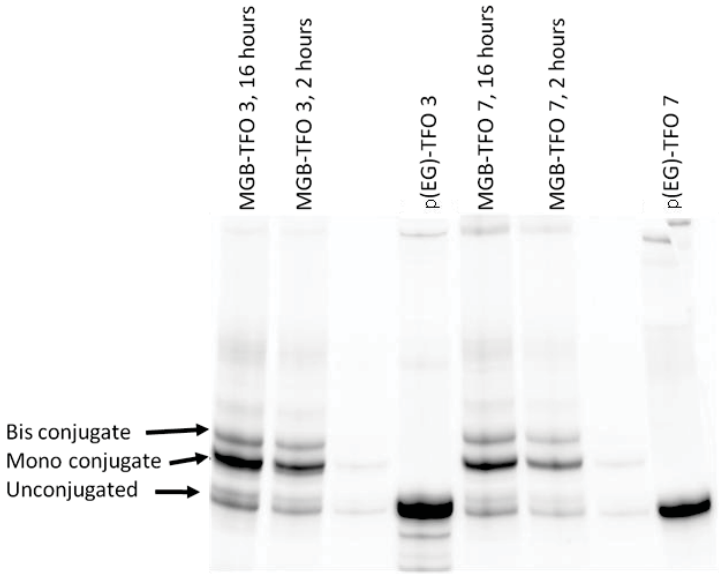


Figure 5.4 Denaturing PAGE (20 %) of a coupling reaction between MGB and $p(EG)_6$-TFO after 2 and 16 hours.

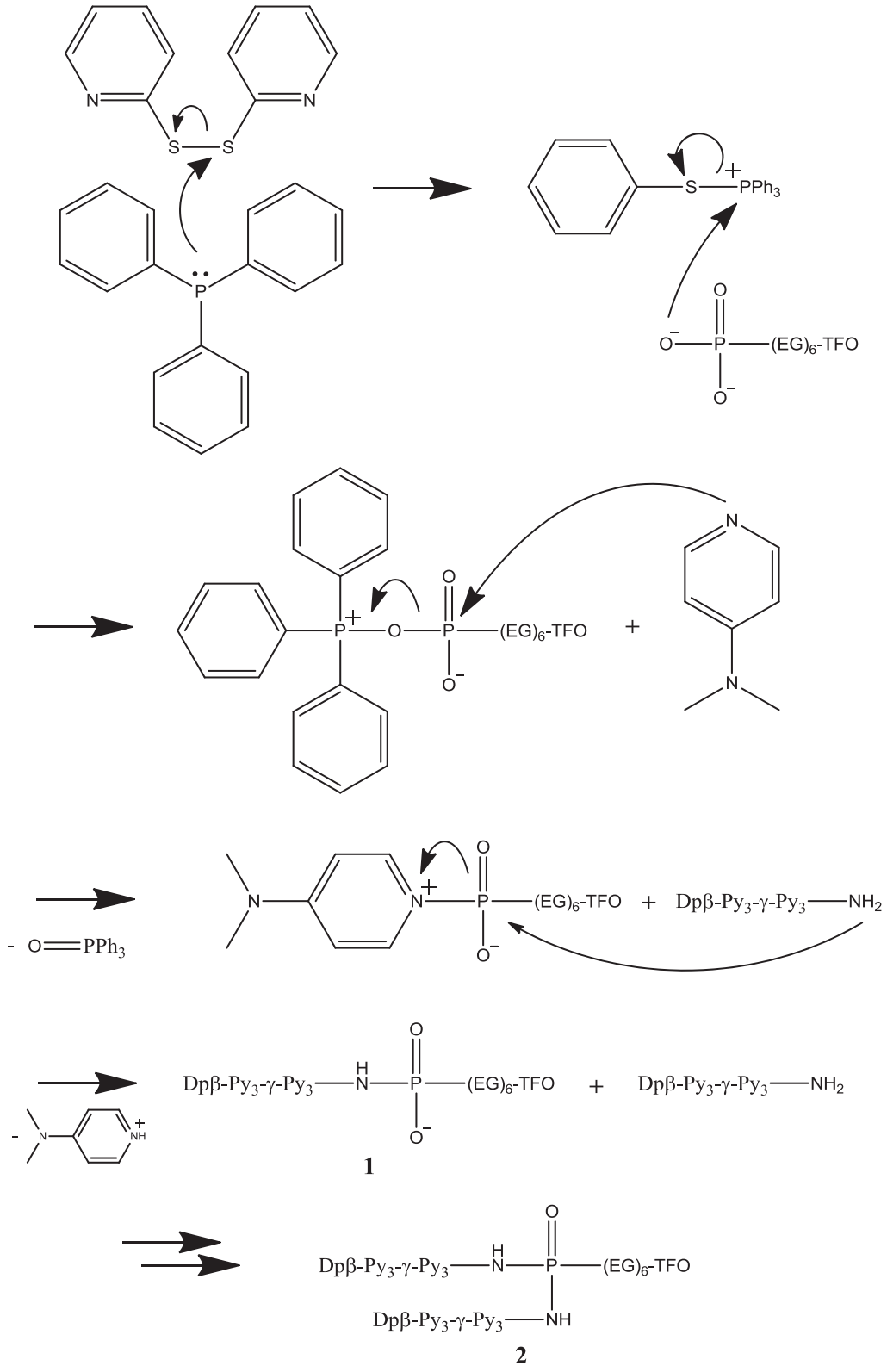


Figure 5.5 Activation and coupling of MGBs with TFO-p$(EG)_6$; compound **1**, mono-conjugated TFO-p$(EG)_6$-MGB; compound **2**, bis-conjugated TFO-p$(EG)_6$-MGB_2.

Initially, in order to isolate mono-conjugated MGB-TFOs from bis-conjugated and unconjugated TFOs, we performed reverse-phase HPLC. The UV-active fractions were freeze-dried and precipitated from 3 % $LiClO_4$ in acetone. Denaturing 20 % PAGE revealed that all fractions that contained mono-conjugated MGB-TFOs also contained bis-conjugated MGB-TFOs. This indicated that the separation by reverse-phase HPLC was not effective to purify mono-conjugated MGB-TFOs.

To improve the purification efficiency we chose to use gel purification method. Impure fractions from previous step was combined and run in 20 % denaturing PAGE (7 M urea). Bands were visualised by UV-shadowing and bands were cut out of the gel. The middle band, which corresponds to mono conjugates, was incubated in aqueous buffer to extract MGB-TFOs. Urea was removed by repetitive precipitation from 3 % $LiClO_4$ in acetone.

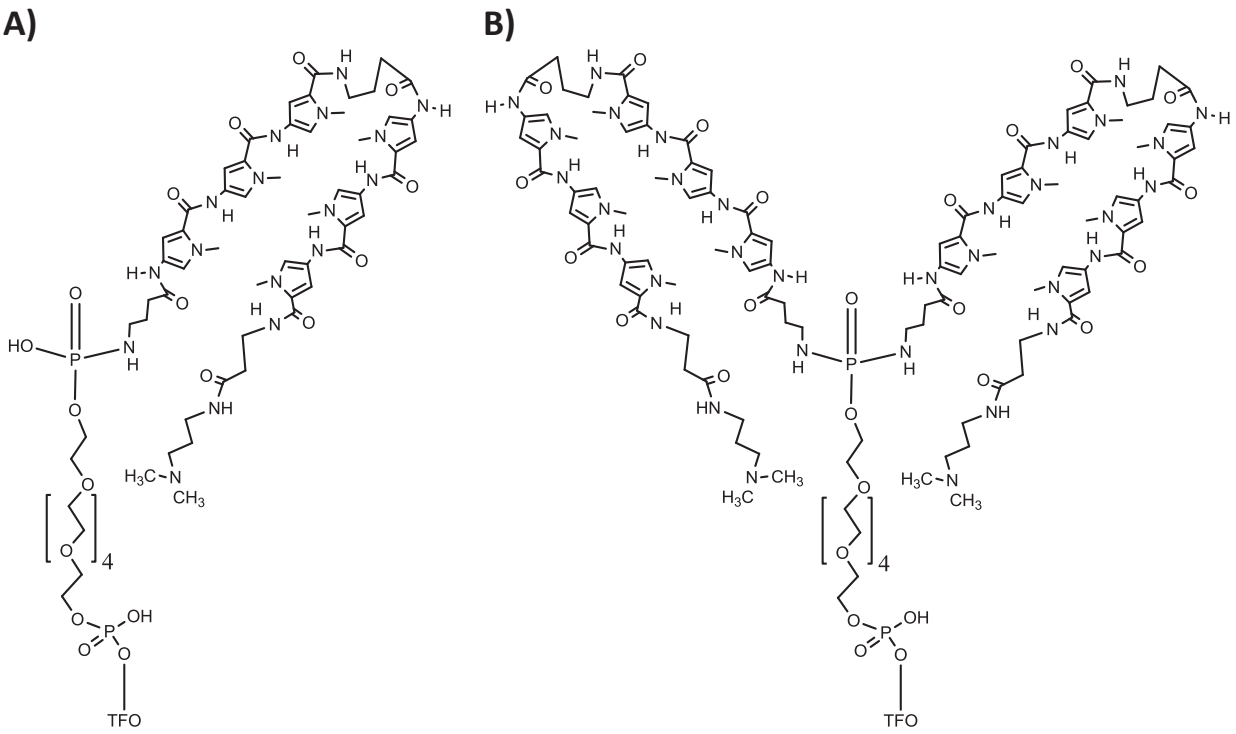


Figure 5.6 A) TFO-p$(EG)_6$-Py_3-γ-Py_3-β-Dp, mono conjugate; B) TFO-p$(EG)_6$-(Py_3-γ-Py_3-β-Dp)$_2$, bis conjugate.

5.4 Dissociation Constants (K_d) of MGB-TFO Conjugates from Target I

Dissociation constants (K_d) were measured for mono conjugate MGB-TFOs and p(EG)-TFOs from target **I** as described in Chapter 2 using 20 % non-denaturing PAGE at 37 °C. The results were compared to non-conjugated TFOs.

Table 5.3 Sequences and dissociation constants (K_d) of triplexes formed with target I in 10 mM HEPES buffer in 50 mM NaCl, 5 mM $MgCl_2$ at pH 7.2, 37 °C.

ON	Sequence 5′ - 3′[a]	K_d (μM)
TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT	0.25 ± 0.06
p$(EG)_6$-TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT-$(EG)_6$p	0.37 ± 0.08
MGB-TFO **3**	dAG**X**GGG**X**GGTTT**X**TGTTTT-$(EG)_6$p-**Py_3-γ-Py_3-βDp**	0.13 ± 0.03
TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT	0.16 ± 0.01
p$(EG)_6$-TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT-$(EG)_6$p	0.45 ± 0.06
MGB-TFO **7**	dAGGG**X**GGGTTT**X**TGTTTT-$(EG)_6$p-**Py_3-γ-Py_3-βDp**	0.30 ± 0.03

[a] '**X**' represents TINA monomer; $(EG)_6$ is a hexaethylene glycol linker; p is a phosphate.

5.5 Discussion

Several studies have been performed on conjugation of MGBs and ONs for improving affinity and stability of DNA complexes (Ryabinin, Denisov et al. 1999; Kutyavin, Afonina et al. 2000; Ryabinin, Boutorine et al. 2004). In the case of triplexes, these studies have focussed on formation of parallel triplexes and were aimed to improve the binding affinity of parallel TFOs for *in vivo* studies where physiological pH prevents the formation of parallel triplexes (Boutorine, Ryabinin et al. 2003; Halby, Ryabinin et al. 2007).

Antiparallel TFOs have never been studied as MGB-TFO conjugates. Even though the physiological pH is not an obstacle for antiparallel triplex formation, addition of MGBs could improve the affinity and may also extend the length of recognition site towards non-homopurine sequences. For this reason, we investigated the conjugation of MGBs with antiparallel TFOs.

Among MGB-TFO conjugates synthesised in this study, only MGB-TFO **3** showed marginally improved affinity towards the target duplex in comparison to TFO **3** (ΔK_d = -0.12 μM). However, MGB-TFO **7** showed less affinity towards the duplex than TFO **7** (ΔK_d = 0.14 μM). It should be remembered that both MGB-TFOs **3** and **7** had higher affinity than their precursors p(EG)-TFOs **3** and **7** ($\Delta K_{d(\text{TFO } \mathbf{3} - \text{p(EG)-TFO } \mathbf{3})}$ = -0.24 μM and $\Delta K_{d(\text{TFO } \mathbf{7} - \text{p(EG)-TFO } \mathbf{7})}$ = -0.15 μM, respectively). This indicates that the incorporation of the MGBs had a limited success.

Synthesis was complicated by purification of MGB-TFO conjugates. Coupling reactions of TFOs and MGBs have yielded both mono- and bis-conjugated MGB-TFOs and isolation of mono-conjugated MGB-TFOs using reverse-phase HPLC was problematic due to the fact that all fractions contained both mono- and bis-conjugates. Instead, denaturing preparative PAGE was used which gave pure mono-conjugated MGB-TFOs.

We could not confirm composition of MGB-TFOs using MALDI-TOF or ESI mass-spectrometry. This can be accounted to the nature of MGB-TFO conjugates and limitations of current protocols used in mass-spectrometry of nucleic acids.

5.6 Summary

We have attached MGBs to antiparallel TFOs in an attempt to increase the affinity and expand the recognition site. We have selected two antiparallel TFOs that showed highest duplex DNA binding affinity from Chapter 3 and covalently linked them to MGBs through a polyethyloxy-glycol chain between phosphate groups as described previously (Kutyavin, Lokhov et al. 2003). The incorporation of the chain itself has resulted in a negative impact on the affinity of antiparallel triplexes. We assume that the linker length can be further optimised which can result in an improved affinity of MGB-TFO for dsDNA. Additionally, different conjugation methods which can provide a single product as a result of the chemical reaction between TFOs and MGBs should be developed.

Chapter 6. Fluorescently Silent TFO probes

6.1 Introduction

As discussed in previous chapters, TINA is an effective mediator of DNA triplex and G-quadruplex formation. The action of TINA can be controlled by its position in a DNA sequence. We have shown that even exceptionally stable G-quadruplexes can be disrupted and DNA triplexes can be formed instead. Further exploitation of TINA-TFOs in diagnostics is limited by the inadequate fluorescence of pyrene. The excitation wavelength of pyrene in TINA is 373 nm, whereas the majority of commercial lasers used in DNA detection techniques operate in the visible and infra-red spectral regions. Moreover, cells autofluoresce upon excitation at 350-370 nm. This warrants the development of new dyes for TINA-TFOs. Two strategies can be envisioned. In the *first strategy* the pyrene in the TINA structure can be substituted by another molecule. However, such structural alterations may change the properties of the TINA molecule significantly and result in poor DNA triplex stability. In the *second strategy* novel dye molecules can be incorporated into the structure of the nucleotides providing the required fluorescent properties, while TINA can still be used as DNA-triplex mediator in the sequence.

In an ideal situation novel dye molecules should have following properties:

- Be excited by commercial lasers at 473, 532, 635, 685 or 785 nm;
- Have a large Stokes shift, so that the fluorescent signal is detected against a dark background;
- Must be stable in biological media and able to be stored for a long time;
- Have low photobleaching;
- Have no or very low fluorescence when the DNA probes are unbound;
- Have large fluorescent enhancement after the hybridization event;

Most of the commercial dyes constantly fluoresce which means that the washing step is required to remove unbound probes. This is a serious limitation especially for *in vivo* imaging techniques. We chose cyanine molecules as they fit the above mentioned criteria including the fact that large fluorescence enhancements are usually observed only after binding of these dyes to DNA structures (Lee, Chen et al. 1986). Moreover, these compounds can be synthesised from commercially available chemicals.

6.1.1 Aim

We aim to synthesise cyanine dyes suitable for incorporation into DNA structure and investigate fluorescent and triplex stability properties of dye-DNA conjugates obtained.

6.2 Synthesis of Fluorescently Silent DNA Probes

Two functionalisation methods are usually used for modification of DNA (Weisbrod and Marx 2008). The first method utilises modified DNA building blocks in the form of phosphoramidites which are used in automated DNA synthesis. This requires the synthesis of individual monomers containing the dye molecule. It has been shown that purification of novel cyanine phosphoramidites is problematic and long coupling times are usually required for their incorporation into DNA (up to 1 hour in contrast to 2 min coupling for standard DNA phosphoramidites) (Menacher, Rubner et al. 2008; Berndl and Wagenknecht 2009). In this regard, the second, post-synthetic method for DNA modification is more appealing, where a functional group of the dye reacts with the complementary functional group on the DNA strand. This also opens a possibility to evaluate cyanine dyes in different locations of the DNA structure using different DNA building blocks.

We chose the second method for DNA modification and used Cu(I)-catalysed azide-alkyne cycloaddition reaction (CuAAC) for post-synthetic attachment of the dye to DNA. This reaction occurs between an organic azide and a terminal alkyne providing 1,2,3-triazole bridge; and this reaction does not attack nucleotides (Bouillon, Meyer et al. 2006; Gierlich, Burley et al. 2006; Geci, Filichev et al. 2007).

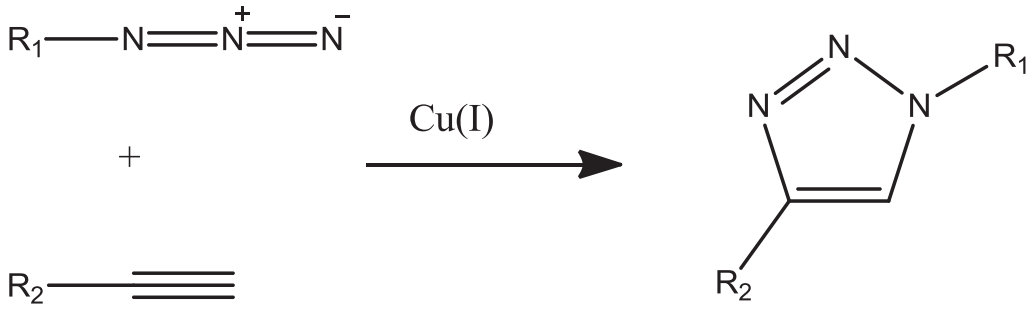


Figure 6.1 Copper-catalysed azide-alkyne cycloaddition reaction (CuAAC).

We have chosen to focus on cyanine dyes for our DNA modification. The cyanine-based dyes are formed by two aromatic moieties linked to each other through a methine chain of an odd number of carbons. In our case a benzothioazole ring was linked to pyridine through a single carbon (Figure 6.2). In order to react with the functional groups on the DNA, an azide group was attached to pyridine through a propane linker. This linker was intended to provide necessary flexibility for intercalation of the dye into the nucleobases of the DNA triplex.

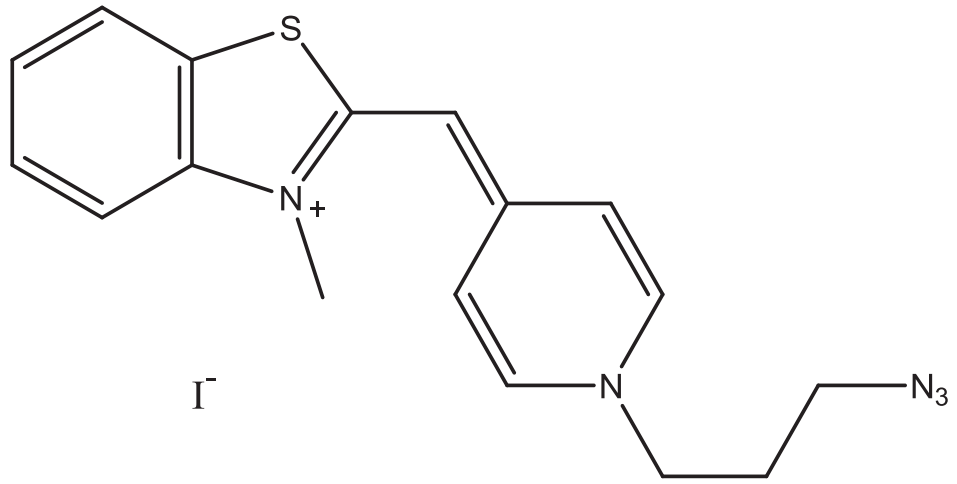


Figure 6.2 Structure of the cyanine-based dye (compound **20**) to be incorporated into DNA.

Four different DNA building blocks were chosen which bear alkynes located in different positions of nucleotides (Figure 6.3). Linker **9** was designed to incorporate a dye molecule in TINA in place of pyrene (*Strategy 1*). Attachment of the dye to 5-ethynyl-2′-deoxyuridine (**10**) and 2′-*O*-propargyl nucleotides (**11** and **12**) allows us to evaluate the *second strategy* for the

incorporation of the dye into TFO. Phosphoramidites of linkers **11** and **12** are commercially available, while linkers **9** and **10** were prepared according to the published procedures (Graham, Parkinson et al. 1998; Geci, Filichev et al. 2007) and were available in our laboratory.

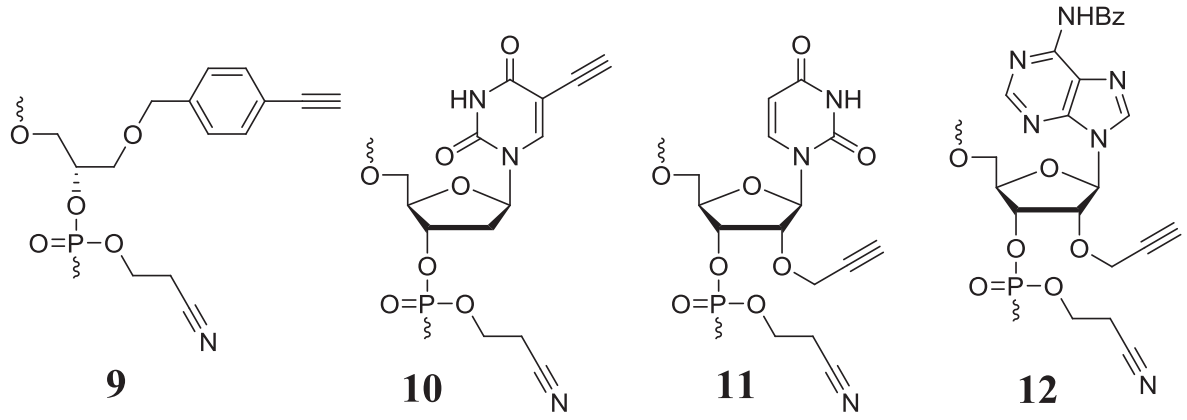


Figure 6.3 The structures of the linkers after DNA synthesis prior to CuAAC reaction.

6.2.1 Synthesis of Fluorescent Moiety

Synthesis of the fluorescent moiety was started from commercially available 1,3-diiodopropane and 4-methylpyridine, which were refluxed in acetonitrile for 20 hours (Figure 6.4). The resulting yellow solid was precipitated after addition of EtOAc and used in the substitution reaction with NaN_3 in acetonitrile (70 % yield over two steps). Compound **19** was synthesised in 84 % yield by fusion of methyl 4-methylbenzenesulfonate and 2-(methylthio)benzothiazole at 130 °C for 1 hour followed by reflux in acetonitrile overnight. Compounds **16** and **19** were mixed in ethanol in the presence of piperidine and refluxed overnight. The resulting compound **20** was purified by silica gel column chromatography, and its structure and composition were confirmed using ^{1}H and ^{13}C NMR spectroscopy and IR spectroscopy and ESI-mass spectrometry.

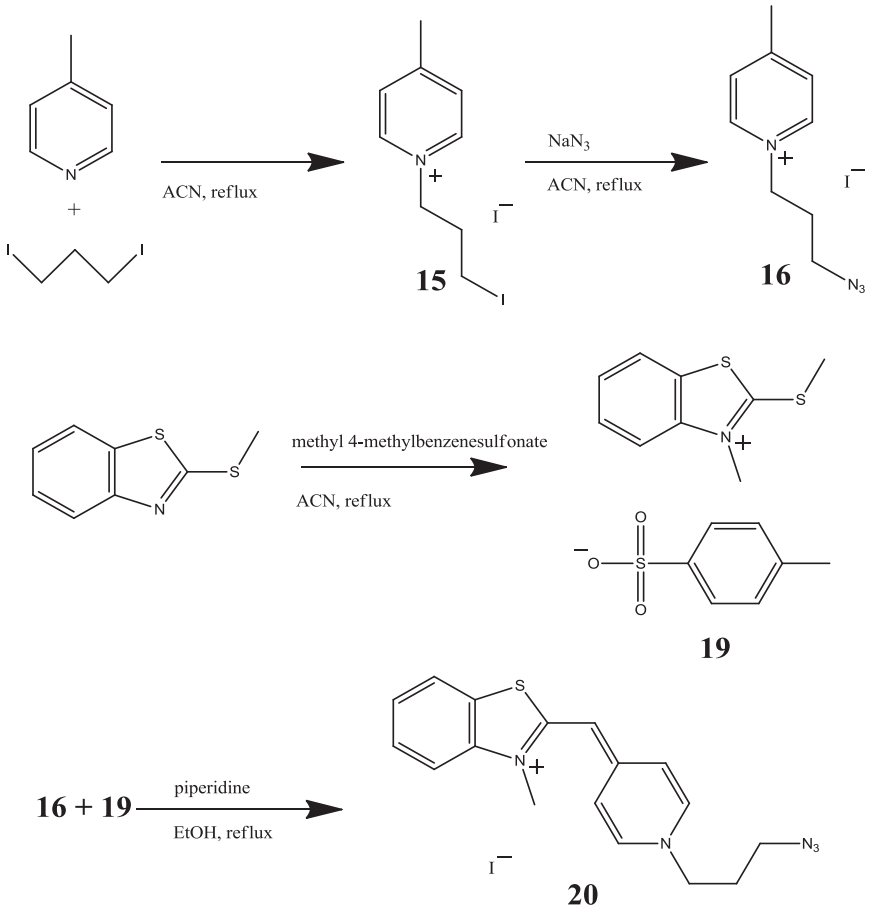


Figure 6.4 Synthesis of the azide-containing fluorescent moiety for incorporation into TFOs.

6.2.2 Synthesis of Oligonucleotides

As our duplex targets, we selected three polypurine sequences which form parallel and antiparallel triplexes (Table 6.1) (Filichev and Pedersen 2005; Doluca, Boutorine et al. 2011). Triplex-forming oligonucleotides (TFOs) complementary to these duplexes were designed with a single alkyne residue (Table 6.2). Linker **9** was used as a bulge insertion whereas others were used as a substitute for a nucleotide (**10** and **11** instead of dT, **12** instead of dA). Sequences **T1-T4** were designed to form parallel triplexes with duplex **D1**. Sequences **T5** and **T6** were designed to form antiparallel triplexes with target **D2**. In order to form a antiparallel triplex with **D3**, sequences **T8** and **T9** were designed using nucleotide **11** based on **T7** which is refered to as **ABL6** in Chapter 3 and was already shown to form an antiparallel triplex. **D2m** and **D3m** were designed to evaluate the sequence specificity of antiparallel TFOs **T5**, **T6**, **T8** and **T9**. Both **D2m**

and **D3m** were based on **D2** and **D3** with two basepairs altered. In order to evaluate the sequence specificity of parallel TFOs, **T2-T4**, we used duplex **D2** as a scrambled target.

Table 6.1 Abbreviations and sequences of target duplexes, complementary DNAs and RNAs containing polypurine binding sites.

ON	Sequence 5′ - 3′ [a]
D1	dGAC **GGG GAA AGA AAA AA** / dTTT TTT CTT TCC CCG TC
D2	dGCC ACT TTT T**AA AAG AAA AGG GGG GA**C TGG / dCCA GTC CCC CCT TTT CTT TTA AAA AGT GGC
D2m	dGCC ACT TTT TAA AAG CAA AGG TGG GAC TGG / dCCA GTC CCA CCT TTG CTT TTA AAA AGT GGC
D3	dTAC **GAA GGG AGA GGG GGA AG**T AC / dGTA CTT CCC CCT CTC CCT TCG TA
D3m	dTAC GAA GGG AGC GGG GTA AGT AC / dGTA CTT ACC CCG CTC CCT TCG TA
C1	dAAA AAA GAA AGG GG
C2	dAAA ACA AAA CCC CCC T
C3	dCTT CCC TCT CCC CCT TC
R1	rAAA AAA GAA AGG GG
R2	rAAA ACA AAA CCC CCC U
R3	rCUU CCC UCU CCC CCU UC

[a] Bold nucleotides in duplexes **D1-D3** represent nucleotides that form Hoogsteen/reverse Hoogsteen bonds with TFOs during triplex formation. Underlined nucleotides represent the mismatched sites in the duplexes **D2m** and **D3m**

Previously we have shown that sequence 5′-dAGGGGGGTTTTGTTTT (**TFO 10**) does not form an atiparallel triplex due to the formation of a highly thermally stable parallel G-quadruplex (see Chapter 2). For this reason antiparallel TFOs for target **D2** were designed with TINA molecules (**X**) to ensure disruption of G-quadruplexes. In the case of the target duplex **D3** we observed that unmodified TFO **T7** formed a G-quadruplex but was still able to form antiparallel triplexes in the presence of the duplex (Chapter 2). To evaluate the ability of modified fluorescent TFOs to form duplexes with complementary single-stranded DNAs and RNAs we included sequences **C1-C3** (ssDNA) and **R1-R3** (ssRNA) in our study.

All modified sequences were synthesised using solid-phase chemistry on an automated DNA synthesiser using phosphoramidite analogues of TINA and linkers **9-12**. For all modified phosphoramidites a hand-coupling step was used during DNA synthesis. Prior to the start of the coupling step, modified phosphoramidites were directly placed on top of the DNA columns followed by an automatic addition of activator solution, which dissolved and activated amidites for an extended coupling for 5 min under argon.

The fluorescent moiety was conjugated into TFOs post-synthetically (Figure 6.5). After solid-phase automated DNA synthesis, the DNAs bound to the controlled pore glass (CPG) were removed from the columns and transferred into microwave vessels along with azide-containing fluorescent molecule (10 equiv.) for CuAAC. Microwave-assisted CuAAC followed by overnight shaking at 20 °C provided higher yields in a shorter amount of time, in comparison to a similar CuAAC reaction over 3 days without microwave assistance (Stephenson, Partridge et al. 2011). After a noticeable change in color, the CPG was washed with DMSO, H_2O and EtOH and the DNA was cleaved from the CPG and deprotected using 32 % NH_4OH.

Table 6.2 Modified and unmodified triplex-forming oligonucleotides (TFO) and their sequences. See Figure 6.5 for definition of **K**, $\mathbf{L^U}$, $\mathbf{M^U}$ and $\mathbf{M^A}$.

TFO	Sequence 5′ - 3′
T1	dCCC CTT TCT TTT TT [b]
T2	dCCC CTT T**K**C TTT TTT [a, b]
T3	dCCC CTT $\mathbf{L^U}$CT TTT TT [a, b]
T4	dCCC CTT $\mathbf{M^U}$CT TTT TT [a, b]
T5	dAGG G**X**G GGT T$\mathbf{L^U}$T GTT TT [a, c]
T6	dAGG G**X**G GGT T$\mathbf{M^U}$T GTT TT [a, c]
T7	dGAA GG GG GAG AGG GA AG [c]
T8	dGAA GG GG GAG $\mathbf{M^A}$GG GA AG [a, c]
T9	dGAA GG G**X**G GAG $\mathbf{M^A}$GG **X**GA AG [a, c]

[a] all modified nucleotides are represented in bold. [b] CT-TFOs were designed to form triplexes in a parallel fashion. [c] GT/GA-TFOs were designed to form triplexes in an antiparallel fashion.

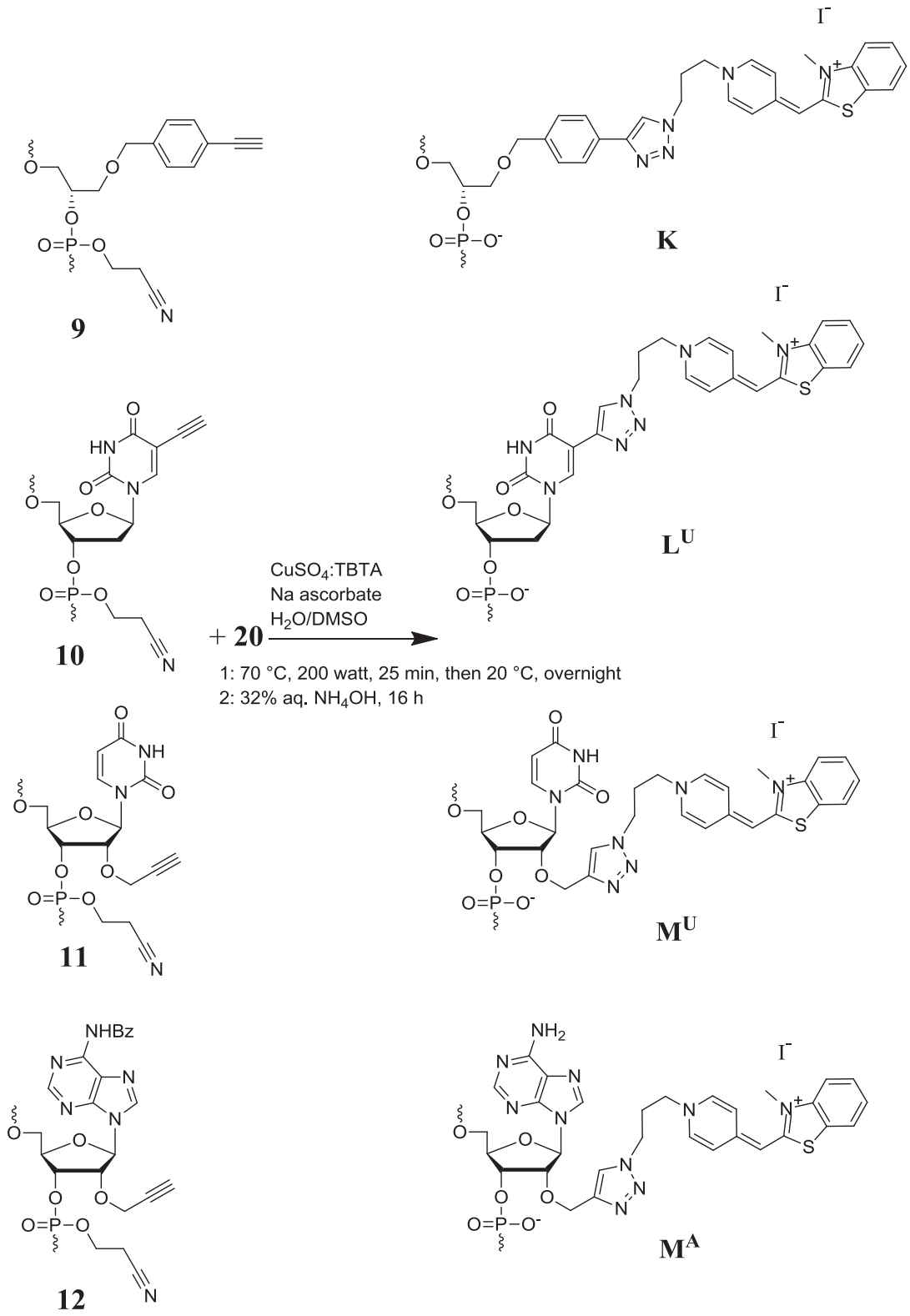


Figure 6.5 Deprotection of ethynyl-containing nucleotides and the CuACC reaction with fluorescent dye (compound **20)**.

6.2.3 Purification of Modified TFOs

A reverse-phase HPLC or denaturing PAGE was used to purify modified TFOs. TFOs **T2 - T4** were purified on reverse-phase HPLC. Buffer A (0.05 M triethylammonium acetate (TEAA) in H_2O at pH 7.0) and Buffer B (75 % acetonitrile in H_2O) were used in chromatography in the following order: 2 min 100 % A, linear gradient to 70 % B in 38 min, linear gradient to 100 % B in 10 min and then 100 % A for 10 min. ONs **T5**, **T6**, **T8** and **T9** were purified using denaturing 20 % PAGE; this is because HPLC purification of these TFOs could not provide pure products. Gels were run in denaturing conditions (7 M urea), sliced to separate bands and cut into pieces. After 2 hours of incubation in formamide, the solution was transferred to new Eppendorf tubes and the TFOs were precipitated from NaOAc in EtOH or $LiClO_4$ in acetone. The pellets were dried at 55 °C for 10 min and dissolved in 100 µL ddH_2O, followed by incubation at 55 °C for 30 min and occasional vortexing, which was required to dissolve TFO completely.

6.3 Properties of Fluorescently Silent Probes

The TFO probes were mixed with the target duplex, complementary DNA or complementary RNA to form triplexes or duplexes in 10 mM Li cacodylate buffer supplemented with 100 mM NaCl and 10 mM $MgCl_2$. An additional set of samples was prepared for each TFO without any target or complementary sequences. For samples containing TFOs **T1 - T4**, the buffer was set to pH 5.0. For all other samples the pH was set to 7.2. Strand concentration was fixed at 1.0 µM for all sequences. After 10 min incubation at 90 °C samples were left to cool down and stored at 4 °C overnight.

Fluorescence emission, UV-Vis, CD and thermal difference spectra (TDS) of these structures were recorded. Quantum yields and melting temperatures (T_m) were calculated. The excitation wavelength (λ_{ex}) was determined according to the maximum fluorescence intensity in the excitation spectrum monitoring emission at 520 nm.

6.3.1 Parallel TFOs

6.3.1.1 Properties of Parallel Triplexes

Parallel TFOs in single-stranded form showed a weak 280 ± 1 nm peak in their CD spectra. Characteristic CD spectra of their duplex match, **D1**, had several positive peaks at 220, 262, and 282 nm in addition to a negative peak at 247 nm. Upon triplex formation by **T1** and **D1**, **T1 + D1**, the positive peaks at 220 and 262 nm disappeared (Figure 6.6). In comparison, a strong negative peak at 210 nm has appeared and the strength of 282 nm peak has increased marginally. This characteristic CD signature was observed previously in CT triplexes (Xodo, Manzini et al. 1990) and has been very similar for all tested parallel triplexes with tiny variations (Figure 6.7 – Figure 6.9).

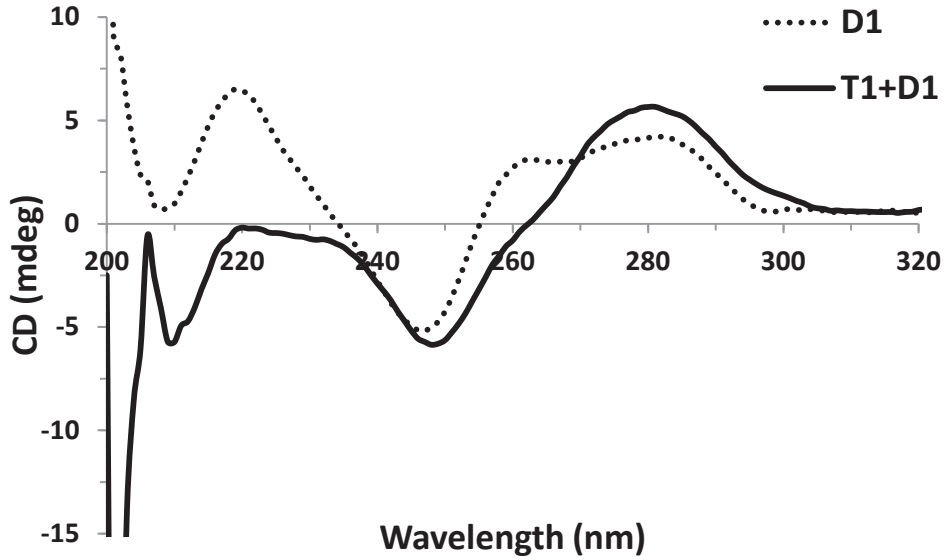


Figure 6.6 CD spectra of duplex **D1** and a parallel triplex **T1+D1** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

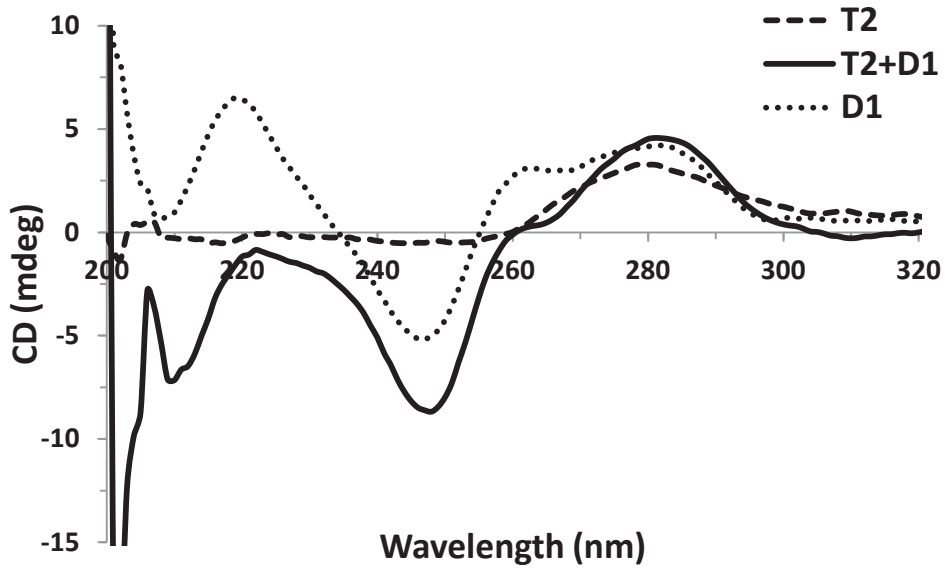


Figure 6.7 CD spectra of TFO **T2**, duplex **D1** and a parallel triplex **T2+D1** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

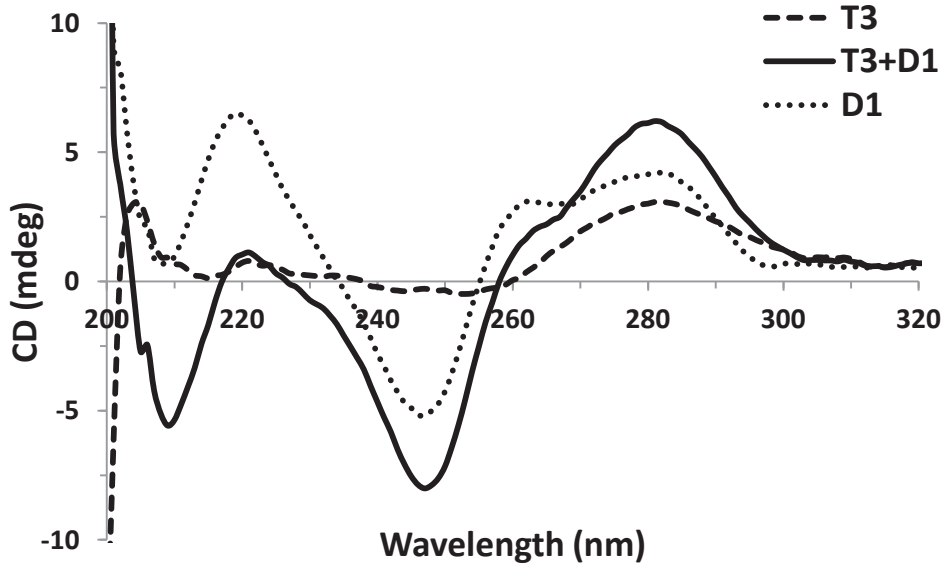


Figure 6.8 CD spectra of TFO **T3**, duplex **D1** and a parallel triplex **T3+D1** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

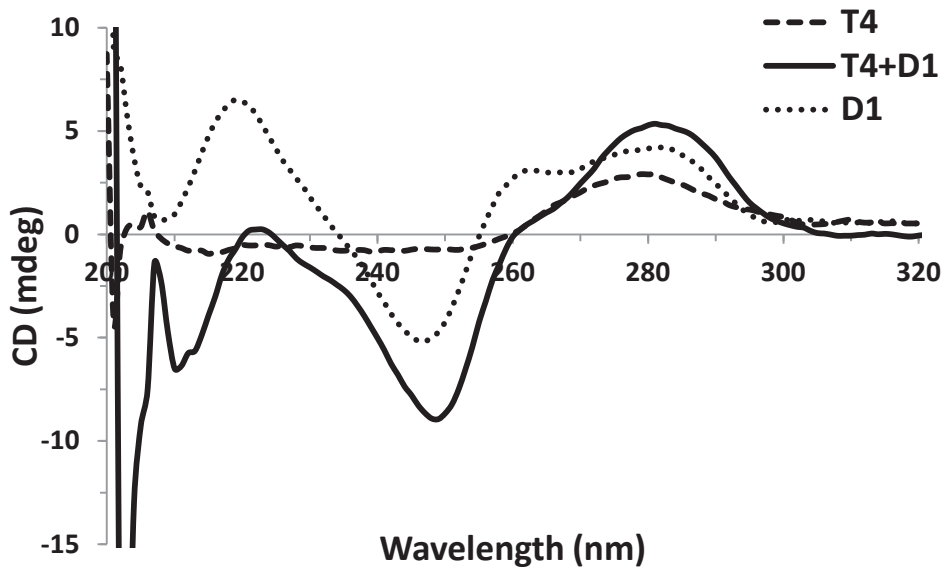


Figure 6.9 CD spectra of TFO **T4**, duplex **D1** and a parallel triplex **T4+D1** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

The fluorescent enhancement (F_c/F_{ss}) is calculated as a ratio of the fluorescence intensities of triplex vs single-stranded TFO. These fluorescent properties were dependent on the alkyne monomer used for the attachment of the cyanine dye in TFOs. Formation of parallel triplexes by incubation of **T2-T4** in the presence of the target duplex **D1** resulted in an increase in the fluorescence intensity between 3 and 16 fold, in comparison to single-stranded TFOs (Table 6.3, Figure 6.10 - Figure 6.12). TFO **T4** with modification $\mathbf{M^U}$ had the lowest quantum yield ($\Phi_F = 0.9$ %) among the single-stranded TFOs **T2-T4**, but the highest quantum yield ($\Phi_F = 25.3$ %) and fluorescence intensity (Figure 6.12) were observed after triplex formation (**T4**+**D1**). These led to the fluorescent enhancement of 15.8 for the triplex vs isolated TFO. In contrast, fluorescent quantum yield and fluorescence intensity changed marginally when TFO **T3** bearing label $\mathbf{L^U}$ was mixed with the duplex **D1** ($F_c/F_{ss} = 2.3$). This is because TFO **T3** exhibits significantly stronger fluorescence in the single-stranded form in comparison with **T2** and **T4** (Φ_F of **T3** = 9.8 %).

Table 6.3 Fluorescence data of parallel TFOs in 10 mM Li cacodylate, 100 mM NaCl and 10 mM $MgCl_2$ at pH 5.0, 20 °C at 1.0 μM strand concentration (See Fluorescence Spectroscopy in Chapter 2).

Sample	λ_{em}, nm	Φ_F [a], %	F_c/F_{ss} [b]	T_m, °C
T1 + D1				56.0
T2	486	2.5 ± 0.03		
T2 + C1	486	8.7 ± 0.08	4.0	42.5
T2 + R1	486	27.4 ± 0.48	8.6	52.6
T2 + D1	486	18.9 ± 0.19	6.7	58.2
T2 + D2	486		1.1	
T3	480	9.8 ± 0.11		
T3 + C1	483	16.7 ± 0.11	3.0	52.5
T3 + R1	486	23.3 ± 0.16	3.7	56.3
T3 + D1	486	10.2 ± 0.06	2.3	57.9
T3 + D2	480		1.1	
T4	480	0.9 ± 0.01		
T4 + C1	480	19.0 ± 0.10	13.7	47.5
T4 + R1	480	8.2 ± 0.06	4.7	53.5
T4 + D1	486	25.3 ± 0.28	15.8	57.7
T4 + D2	480		1.3	

[a] Fluorescence quantum yield, λ_{ex} is set to 464 nm. [b] $F_{complex}/F_{single\text{-}stranded}$, ratios were calculated from the intensities measured at the λ_{em} of the duplexes or triplexes vs single-stranded TFOs.

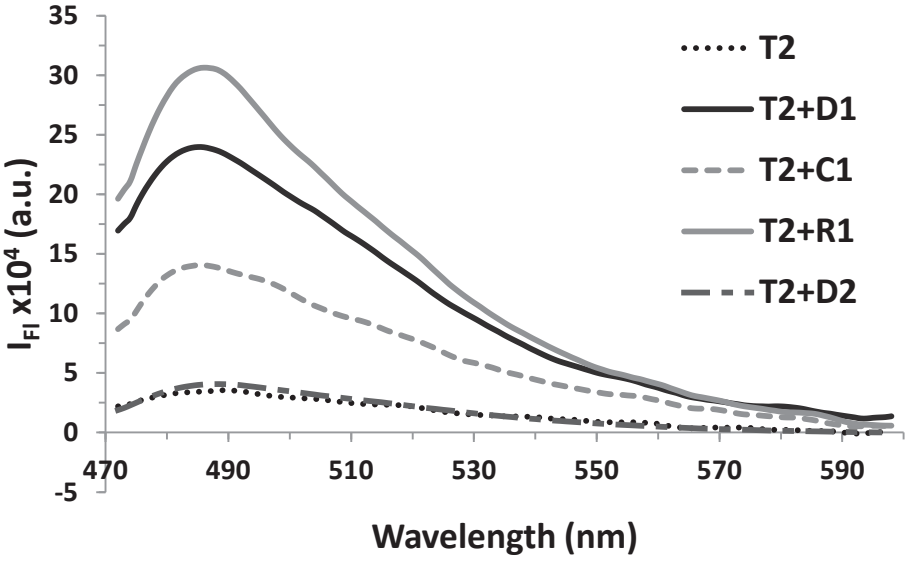


Figure 6.10 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T2**, duplexes **T2+C1** and **T2+R1**, parallel triplex **T2+D1** and mismatched triplex **T2+D2** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

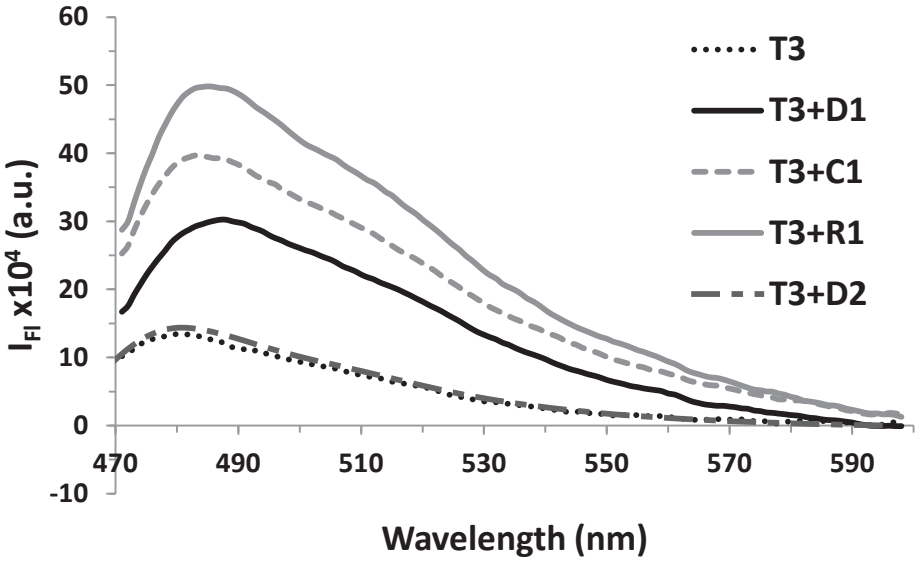


Figure 6.11 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T3**, duplexes **T3+C1** and **T3+R1**, parallel triplex **T3+D1** and mismatched triplex **T3+D2** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

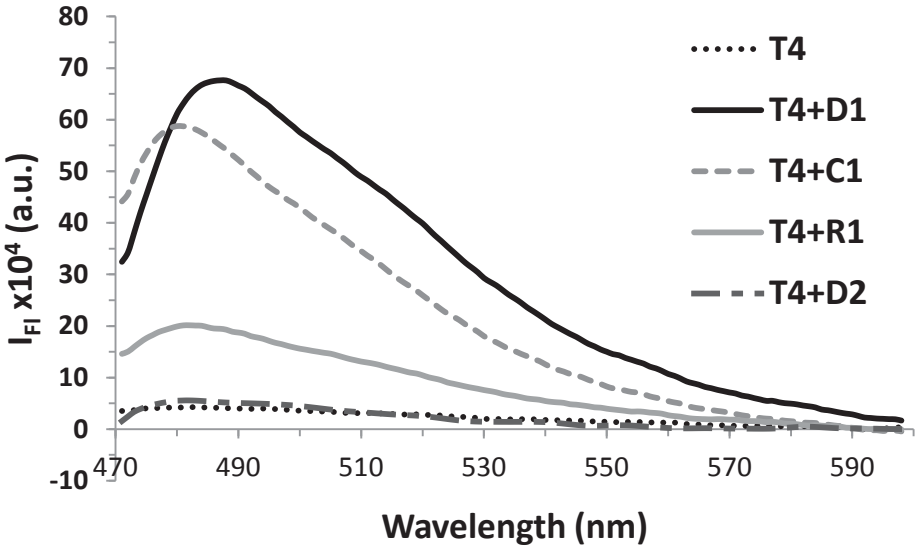


Figure 6.12 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T4**, duplexes **T4+C1** and **T4+R1**, parallel triplex **T4+D1** and mismatched triplex T4+D2 in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0, 20 °C. Strand concentration was set to 1.0 μM.

While the thermal difference spectra of triplexes (**T1-T3 + D1**) and duplex (**D1**) were very similar, **D1** can be distinguished from triplexes by a weak increase in absorbance at 295 nm during melting. In comparison to the unmodified parallel triplex (**T1 + D1**), all modified triplexes (**T2 + D1**, **T3 + D1** and **T4 + D1**) showed increased thermal stability by +2.2, +1.9 and +1.7 °C, respectively. Insignificant changes in the fluorescent intensity of modified probes in the presence of the target duplex at pH 7.2 indicate that triplexes are not formed at neutral pH.

We decided to use duplex **D2** as a negative control for parallel triplex formation. When TFOs **T2-T4** were incubated with duplex **D2**, we observed a slight fluorescent enhancement ($F_c/F_{ss} < 1.3$). This indicates that due to the mismatching with the scrambled sequence there was little or no interaction between TFOs **T2-T4** and duplex **D2**.

6.3.1.2 Properties of DNA/RNA and DNA/DNA Duplexes

The high RNA content in the intracellular environment is one of the biggest obstacles to *in vivo* use of DNA-based probes. RNA may form stable duplexes with these probes and prevent interaction with the intended targets (*i.e.* duplex DNA). For this reason we decided to investigate

the properties of duplexes that are formed by our TFOs and complementary strands (RNA and DNA).

Interestingly, we observed higher quantum yields and fluorescent enhancement for probes **T2** and **T3** in the structure of RNA/DNA duplexes than in parallel triplexes (Φ_F = 27.4 % and 23.3 %, respectively). In comparison, when these probes formed DNA/DNA duplexes quantum yields were lower than for RNA/DNA (**T2+C1** vs. **T2+R1** and **T3+C1** vs. **T3+R1**, Table 6.3). These results indicate that these probes prefer targeting RNA instead of DNA. To our delight, duplexes formed by **T4** were not as bright as the target triplex.

We observed significant differences in thermal stability of DNA/DNA duplexes formed by the parallel TFO probes, **T2 + C1**, **T3 + C1** and **T4 + C1**. The lowest T_m value among these samples was detected for **T2 + C1** (42.5 °C) and the highest for **T3 + C1** (52.5 °C). A difference of 10 °C indicates that the type of the alkyne monomer has a very strong effect on the thermal stability of the duplexes. This was also the case for the RNA/DNA duplexes. Among RNA/DNA duplexes, **T2 + R1** and **T3 + R1** had the lowest (52.6 °C) and the highest (56.3 °C) thermal stabilities, respectively. This means that the linker used in **T3**, monomer **10**, is advantageous over monomers **9** and **11** in terms of duplex thermal stability.

6.3.2 Antiparallel TFOs

6.3.2.1 Properties of Antiparallel Triplexes

The design of sequences **T5** and **T6** was based on TFO **7** from Chapter 3, which targeted a 16-bp polypurine region of HIV-1 proviral DNA (Target I). A TINA monomer was used to disrupt G-quadruplex formation mediated by six contiguous guanosines. In agreement with the parallel triplex studies, TFO **T5** containing modified 5-ethynyl uridine ($\mathbf{L}^{\mathbf{U}}$) showed relatively high fluorescence quantum yield (15.6 %, Table 6.4) in a single-stranded form in comparison with TFO **T6**, which was modified with 2′-*O*-propargyl uridine ($\mathbf{M}^{\mathbf{U}}$, Φ_F = 1.5 %). This limits the use of monomer $\mathbf{L}^{\mathbf{U}}$ in TFOs for fluorescent *in vivo* applications. Moreover, only a small increase in fluorescent quantum yield and fluorescent intensity was detected after triplex formation of **T5** with the target duplex **D2**. On the other hand, triplex **T6 + D2** showed almost 20 times higher

fluorescence quantum yield in comparison with the single-stranded **T6**, which resulted in a strong fluorescent enhancement (F_c/F_{ss} = 6.2) of the triplex against the fluorescently silent TFO. See Appendix Figure 4 for CD spectra of **T5** and **T6**.

Table 6.4 Fluorescence data of antiparallel TFOs in 10 mM Li cacodylate, 100 mM NaCl and 10 mM $MgCl_2$ at pH 7.2, 20 °C at 1.0 μM strand concentration (See Fluorescence Spectroscopy in Chapter 2).

Sample	λ_{em}, nm	Φ_F [a], %	F_c/F_{ss} [b]	T_m, °C
T5	482	15.6 ± 0.37		
T5 + C2	483	23.6 ± 0.20	1.5	58.1
T5 + R2	482	18.5 ± 0.11	1.1	54.6
T5 + D2	482	32.7 ± 0.58	1.8	74.0
T5 + D2m	482		0.4	
T6	484	1.5 ± 0.02		
T6 + C2	484	15.1 ± 0.50	3.5	55.8
T6 + R2	486	11.2 ± 0.30	3.4	53.5
T6 + D2	486	23.5 ± 1.0	6.2	74.5
T6 + D2m	484		1.6	
T7				>90 [c]
T7 + D3				73.0
T8	485	11.2 ± 0.13		>90 [c]
T8 + C3	486	21.7 ± 0.09	3.8	64.2
T8 + R3	486	19.7 ± 0.06	3.6	76.2
T8 + D3	485	29.8 ± 0.13	7.1	73.1
T8 + D3m	485		1.0	
T9	486	1.1 ± 0.03		
T9 + C3	484	17.8 ± 0.88	4.0	49.7
T9 + R3	486	17.6 ± 0.14	3.9	43.7
T9 + D3	486	34.9 ± 1.10	17.8	72.5
T9 + D3m	486		1.3	

[a] λ_{ex} is set to 464 nm. [b] $F_{complex}/F_{single\text{-}stranded}$, ratios were calculated from the intensities measured at the λ_{em} of duplexes or triplexes vs single-stranded TFOs. [c] G-quadruplex formation was detected.

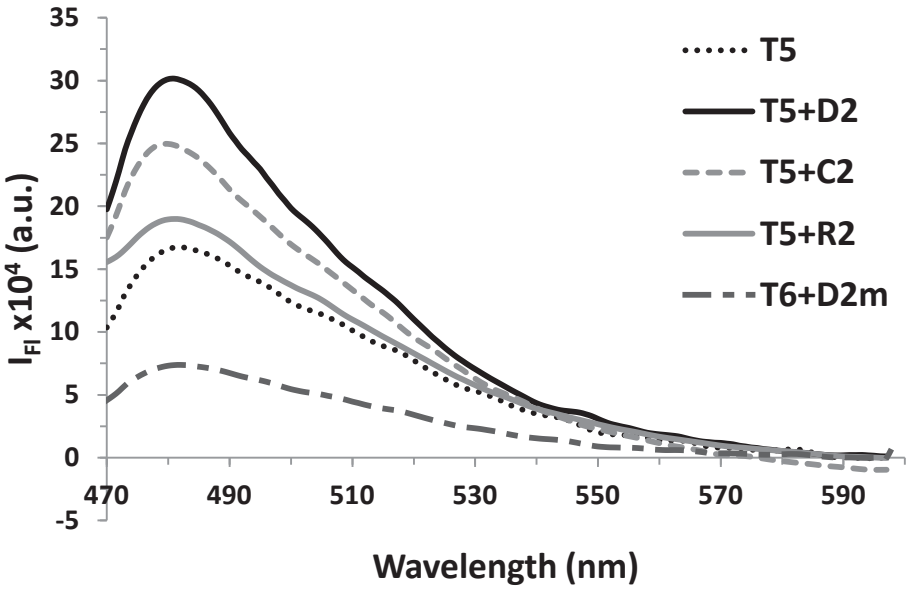


Figure 6.13 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T5**, duplexes **T5+C2** and **T5+R2**, parallel triplex **T5+D2** and mismatched triplex **T5+D2m** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

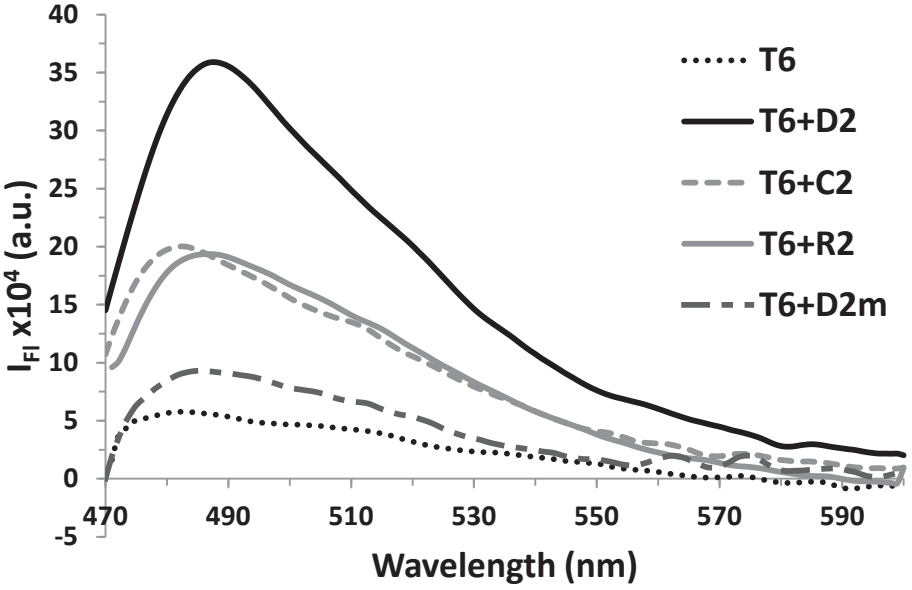


Figure 6.14 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T6**, duplexes **T6+C2** and **T6+R2**, parallel triplex **T6+D2** and mismatched triplex **T6+D2m** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

Duplex **D2m,** which was identical to duplex **D2** except for two mismatches, was used to evaluate sequence specificity of TFOs **T5** and **T6**. To our surprise, when **D2m** was incubated with **T5** the fluorescence intensity was halved (F_c/F_{ss} = 0.4). On the other hand, fluorescence enhancement of triplex **T6 + D2m** was limited to 1.61.

TFO **T7**, which is identical to **ABL6** used in Chapter 3, formed a parallel G-quadruplex according to native PAGE and CD experiments. We designed two fluorescent sequences which contain monomer $\mathbf{M^A}$: TFOs **T8** and **T9**. A TINA molecule was used in the sequence **T9** to disrupt potential formation of G-quadruplexes. It is interesting that TFO **T8** showed a high Φ_F value of 11.2 % in the absence of the target duplex. CD spectroscopy shows that this sequence exists as a G-quadruplex, similar to unmodified TFO **T7** (Figure 6.17 and Figure 6.18). In both cases positive ellipticity at 265 nm and valleys with minima at 245 and 290 nm indicate G-quadruplex formation. We have previously shown that TFO **T7** forms an antiparallel triplex in the presence of the target (see Chapter 3). The triplex **T8 + D3** was also formed according to CD spectra and showed a quantum yield value of 31.9 % which is almost three times higher than Φ_F of the G-quadruplex formed by TFO **T8**. Triplexes **T7 + D3** and **T8 + D3** have almost identical melting temperatures ($\Delta T_{m[(\mathbf{T8}+\mathbf{D3})-(\mathbf{T7}+\mathbf{D3})]}$ = 0.1 °C). This indicates that the incorporation of monomer $\mathbf{M^A}$ did not have any significant effect on the thermal stability of antiparallel triplexes.

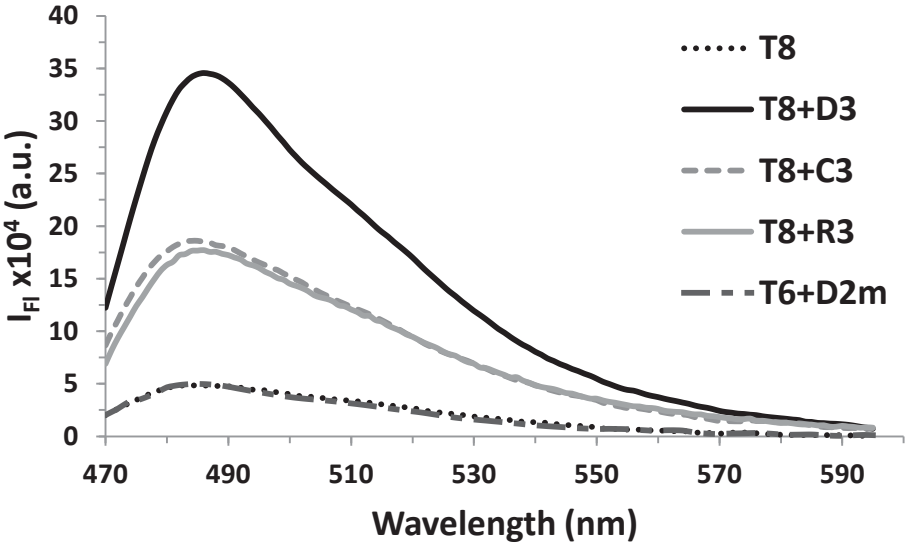


Figure 6.15 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T8**, duplexes **T8+C3** and **T8+R3**, parallel triplex **T8+D3** and mismatched triplex **T8+D3m** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

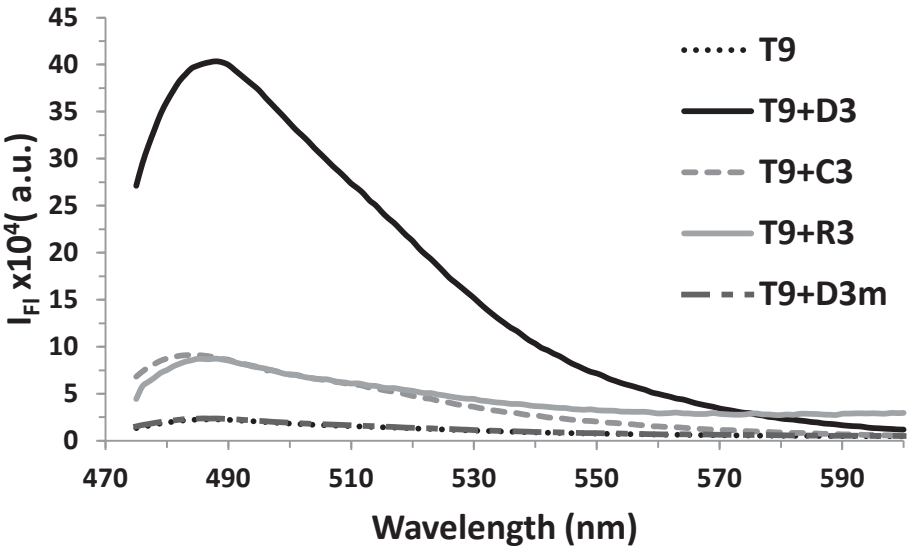


Figure 6.16 Fluorescence emission spectra (λ_{ex} = 464 nm) of TFO **T9**, duplexes **T9+C3** and **T9+R3**, parallel triplex **T9+D3** and mismatched triplex **T9+D3m** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

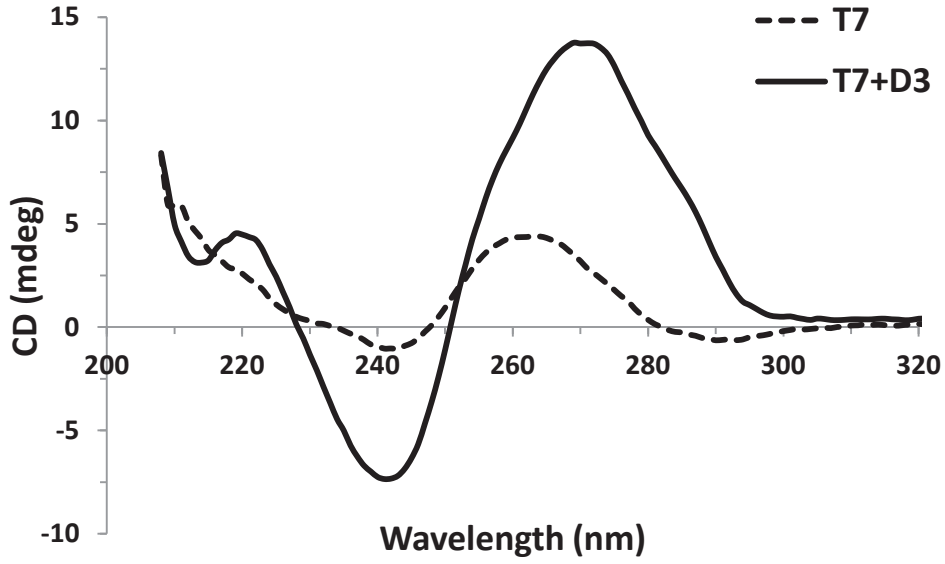


Figure 6.17 CD spectra of TFO **T7** and triplex **T7+D3** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

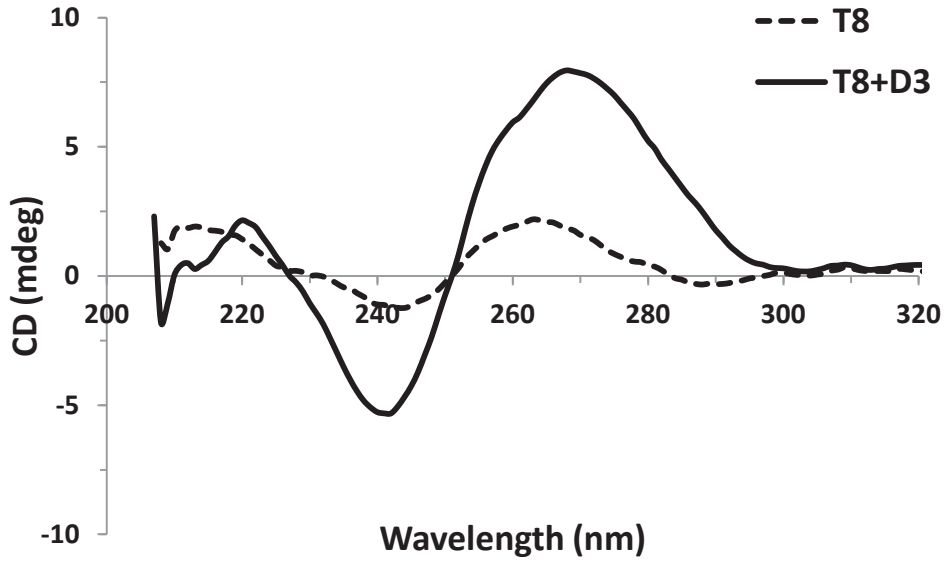


Figure 6.18 CD spectra of TFO **T8** and triplex **T8+D3** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

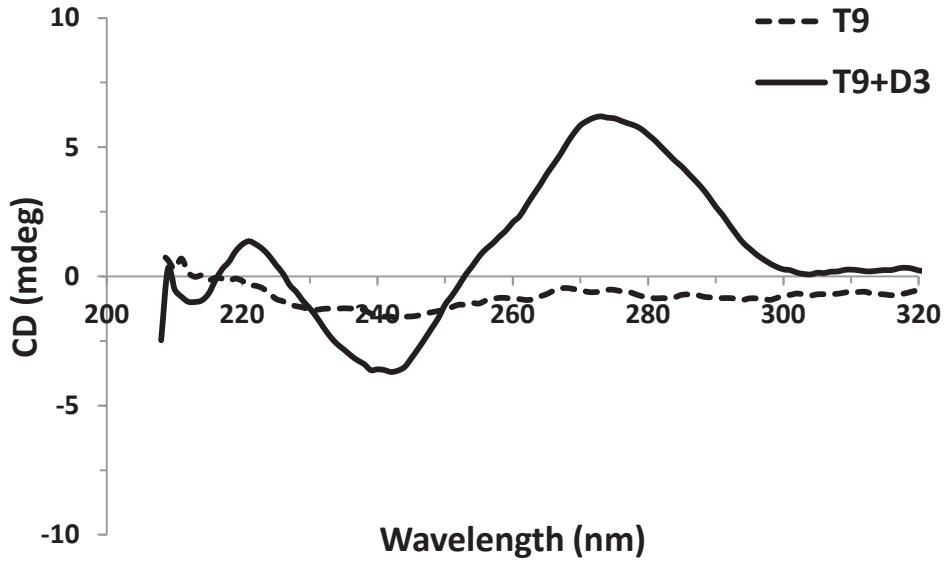


Figure 6.19 CD spectra of TFO **T9** and triplex **T9+D3** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

In order to eliminate high fluorescence of **T8** caused by G-quadruplex formation, we decided to use two TINA monomers (**X**) in the sequence **T9**. This led to a significant decrease in fluorescence intensity of **T9** (Φ_F = 1.1 %) and no G-quadruplex formation was detected using CD spectroscopy (Figure 6.19). To our delight, a highly fluorescent triplex was formed by **T9**

and duplex **D3** (Φ_F = 34.9 %) while thermal stability was not affected ($\Delta T_{m[(\mathbf{T9}+\mathbf{D3})-(\mathbf{T7}+\mathbf{D3})]}$ = -0.5 °C).

Both TFOs, **T8** and **T9**, showed excellent sequence specificity as the mismatched triplexes **T8 + D3m** and **T9 + D3m** showed only a marginal increase in fluorescence intensity in comparison to single-stranded TFOs (F_c/F_{ss} = 1.0 and 1.3, respectively).

6.3.2.2 Properties of DNA/RNA and DNA/DNA Duplexes

TFO **T5** is the least fluorescently discriminating probe in the series of antiparallel TFOs. It has similar fluorescent intensities in single-stranded form, in duplexes and in triplexes. While TFO **T3** with monomer $\mathbf{L}^{\mathbf{U}}$ was ideal for RNA visualisation due to high fluorescence in RNA\DNA duplex, $\mathbf{L}^{\mathbf{U}}$ incorporated into TFO **T5** showed lower fluorescence in RNA\DNA duplex than either the DNA\DNA duplex or antiparallel triplex. This can be explained as a result of the TINA monomer in the sequence of TFO **T5** which promotes triplex formation.

When the attachment of the dye in TFO **T5** was changed from 5-ethynyl uridine to 2′-*O*-propargyl uridine we observed a significantly better discrimination pattern, which is in agreement with the pattern observed for parallel TFOs. Thus, fluorescent intensity of the triplex formed by **T6** was twice as high as the intensity detected for DNA/RNA and DNA/DNA duplexes, while the single-stranded TFO was fluorescently silent (Φ_F = 1.5 %).

Despite the fact that TFO **T8** formed a G-quadruplex in the presence of NaCl, its fluorescence changed upon formation of duplexes with **C2** and **R2** (Table 6.4). It is interesting that upon incorporation of TINA in the fluorescent ABL TFO (TFO **T9**) we obtained a highly fluorescent triplex, whereas quantum yields of the corresponding duplexes with DNA and RNA were lower than those for TFO **T8**. This is also correlated with lower thermal stabilities of duplexes formed by **T9** in comparison with **T8**, i.e. $\Delta T_{m[(\mathbf{T9}+\mathbf{C3})-(\mathbf{T8}+\mathbf{C3})]}$ = -14.5 °C and $\Delta T_{m[(\mathbf{T9}+\mathbf{R3})-(\mathbf{T8}+\mathbf{R3})]}$ = -32.5 °C, which is in agreement with the previously reported preferential binding of TINA-TFOs using Hoogsteen rather than Watson-Crick base-pairing (Filichev and Pedersen 2005). As such, the best discrimination between a DNA triplex versus duplexes and single-stranded probes was seen for TFO **T9** in which TINA molecule destabilised the G-quadruplex and duplex structures, and stabilised DNA triplexes, while fluorescent properties were provided by monomer $\mathbf{M}^{\mathbf{A}}$.

6.4 Discussion

Fluorescent oligonucleotides have been widely used as agents for recognition of nucleic acids for diagnostic purposes. Hybridisation-dependent fluorescence emission is one of the desired properties for DNA probes and in recent years several studies have been focused on 'fluorescently-silent' or 'light-up' oligonucleotide probes. However, only a few studies have focused on the synthesis of fluorescently-silent TFO probes targeting double-stranded DNA (Figure 6.20). A parallel TFO probe labelled with perylene has shown a 3-fold fluorescent enhancement in the presence of the target duplex (Aubert and Asseline 2004). Another example was a ruthenium complex resulting in a 10-fold increase in fluorescence intensity (Grimm, Boutorine et al. 2002). However, among fluorescently-silent TFO probes cyanine dyes have shown the highest fluorescence enhancements. Thus, parallel TFO probes labelled by monomethine cyanine dyes through an 8 carbon long linker at the 5′-end showed between 3 and 37-fold increase in fluorescence intensity upon triplex formation (Renard, Lartia et al. 2008). However, no attention has been given to the synthesis of fluorescently-silent *antiparallel* TFOs.

Here, in this study we synthesised fluorescently-silent antiparallel TFOs and investigated the effect of the position of the dye in the oligonucleotide using four different monomers (**K**, $\mathbf{L^U}$, $\mathbf{M^U}$ and $\mathbf{M^A}$). Monomer **K** was distinguished from the latter three by being used as a bulge insertion and the dye was attached directly to the backbone (*1st Strategy*, page 94). The dye was attached directly to the nucleobase in monomer $\mathbf{L^U}$ and to the sugar in monomers $\mathbf{M^U}$ and $\mathbf{M^A}$. In this respect this is the first study that investigates the effect of different attachment positions of the dye in the TFO. Different attachment positions led to different environments of the dye and hence to different fluorescence properties. On the other hand, our data clearly indicate that monomers $\mathbf{M^U}$ and $\mathbf{M^A}$ have the highest fluorescent enhancement (up to 18-fold) upon triplex formation in comparison to single-stranded TFOs. Moreover, the fluorescent enhancement was specific to triplexes as the fluorescence enhancement for DNA/DNA or DNA/RNA duplexes was limited. We chose to name this type of discriminating fluorescent nucleic acids as assembly-dependent fluorescence enhancing nucleic acids (AFENA). DNA probes based on $\mathbf{M^U}$ and $\mathbf{M^A}$ monomers can be called AFENA-DNA because DNA triplexes can be visualised efficiently in the presence of RNA. DNA probes based on monomer **K** can be called AFENA-RNA because fluorescent signal was higher for the complex with RNA than that with

ssDNA or dsDNA; however fluorescence enhancement of these monomers may be strongly sequence dependent and further investigation is necessary to generalise these assertions. TINA monomers were also investigated in combination with the cyanine dye in the construction of antiparallel TFOs. Presence of TINA monomers not only prevented the formation of G-quadruplexes and thus eliminated the fluorescence caused by G-quadruplex formation but also decreased thermal stability and the fluorescence intensity of alternative duplexes (DNA/DNA or DNA/RNA) that can be formed by the fluorescent TFOs. These properties support the viability of the *2nd Strategy* proposed in this Chapter and lay a foundation for potential *in vivo* applications.

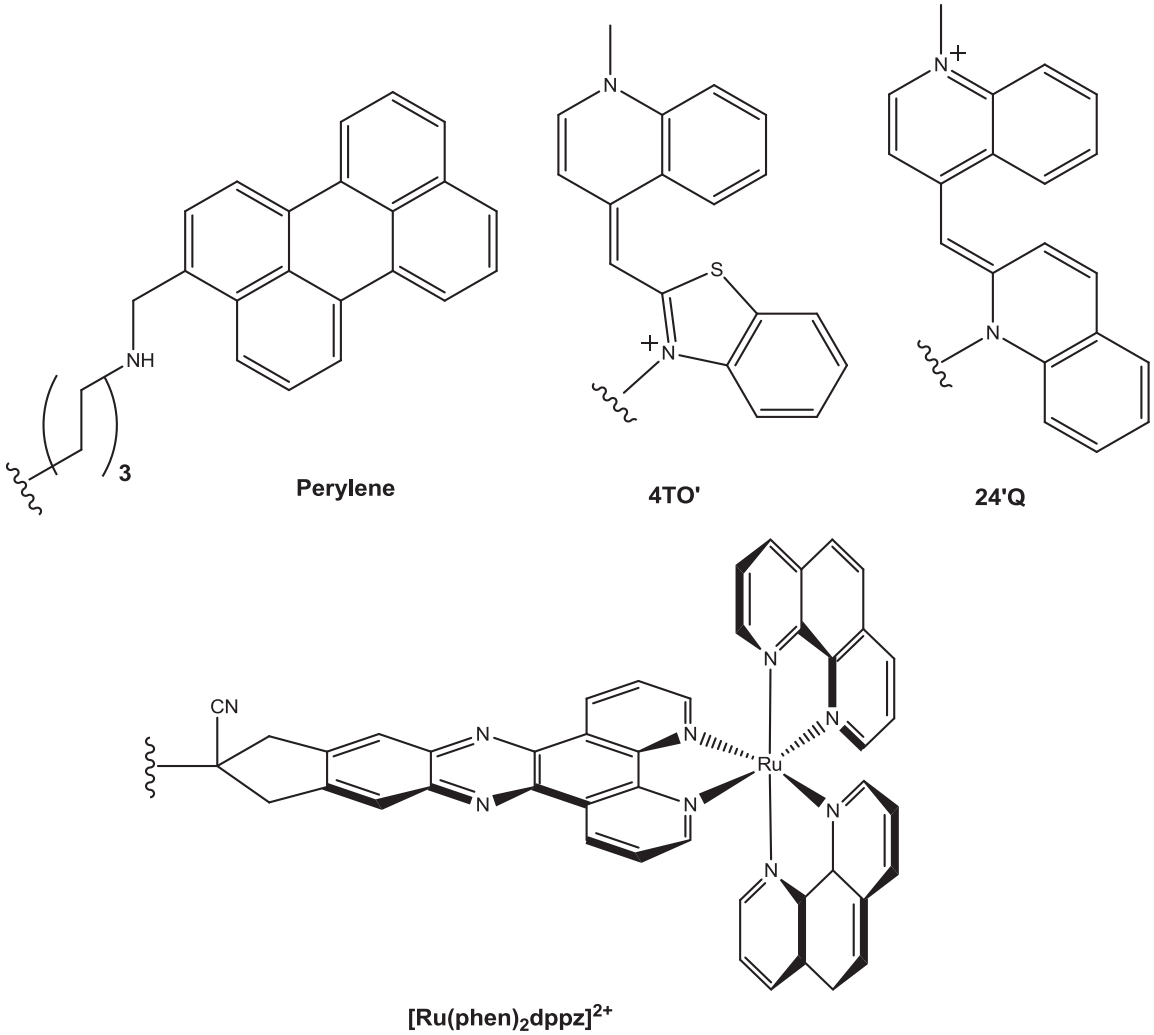


Figure 6.20 Dyes used in the synthesis of fluorescently silent parallel TFOs.

6.5 Summary

In this chapter we investigated strategies aiming to improve the fluorescent properties of TFOs. For this purpose we synthesised a novel cyanine dye, attached it to different positions in a DNA strand and studied the fluorescence properties of these conjugates. This led to the development of novel fluorescently discriminating DNA-conjugates, called AFENA.We demonstrated that up to a 18-fold increase in fluorescence intensity can be obtained upon formation of parallel and antiparallel triplexes. Attachment of the dye at the 2′-*O*-position of the nucleotide led to TFO probes with excellent fluorescent discrimination between triplexes, duplexes and single-stranded DNAs. We also showed that a combination of the cyanine dye with TINA monomers resulted in G-rich TFOs that form antiparallel triplexes while their ability to self-aggregate was diminished.

Our study opens a road for the development of fluorescently silent TFOs based on antiparallel G-rich sequences. This strategy can be applied to other cyanine dyes with different fluorescent properties such as excitation and emission wavelengths, which can be applied for the gene visualisation in specific colours.

Chapter 7. Thesis Conclusion and Future Directions

7.1 Thesis Conclusion

Throughout this thesis, one DNA sequence showed remarkable changes in its structure and functionality. The unmodified G-rich sequence 5′-dAGGGGGGGTTTTGTTTT (TFO **10**) was originally designed as a control TFO for targeting HIV-1 proviral DNA. Due to its 6 nt long guanine tract, this sequence formed a highly stable parallel G-quadruplex and thus was excluded from the triplex formation. In Chapter 3, we investigated how TINA monomers can change the properties of this sequence and lead to antiparallel triplex formation. Insertion of two TINA monomers in this sequence – one of them in the middle of the guanine tract – disrupted the G-quadruplex formation and the new sequence (TFO **7**) was capable of forming a stable antiparallel triplex with the target duplex. In the light of our results, we established a set of rules for the design of TINA-conjugated G-rich TFOs and demonstrated their applicability on other sequences. The dramatic influence of TINA insertion on G-quadruplex stability has inspired us to study this effect further and we expanded our studies on the dTG_4T sequence, which forms a tetramolecular parallel G-quadruplex (Chapter 4). We established that the TINA monomer enhances the speed of association and, to our surprise; this led to an unusual topology for the tetramolecular G-quadruplex (**GXG** sequence). This validates a strategy in which ligands covalently attached to DNA are presumed to overcome guanine-mediated association of G-rich strands and thus control G-quadruplex topology. In the next step, we explored the conjugation of minor groove binders (MGBs) to TINA-TFOs in order to improve the affinity of the TFOs and expand the recognition abilities of MGBs. Two conjugates were synthesised based on TFO targeting an HIV sequence. However, marginal changes in the affinity of the conjugates were observed in comparison to TINA-TFOs. In the final step, we synthesised a fluorescent dye, investigated fluorescent properties of TFO-dye conjugates which led to invention of a new class

of nucleic acids called AFENA that exhibit a distinct visual signal only after binding to double-stranded DNA. It is remarkable that TFO **10** was converted from a highly thermally stable G-quadruplex into an efficient antiparallel fluorescently silent TFO that turns on only after triplex formation (**T6**). This particular sequence serves as an excellent example demonstrating the numerous capabilities behind chemical functionalisation of DNA molecules.

7.2 Future Directions

In this study we have taken G-rich oligonucleotides and converted them into efficient TFO probes for DNA visualisation. Further research will follow two research paths: chemistry and cell biology.

In terms of chemistry, there is a wide range of potential improvements available. The initial step should be modifying the cyanine dye in order to expand fluorophore options with different fluorescence properties. By synthesising new cyanine dyes, TFO probes can be excited by a wider range of commercial lasers.

For *in vivo* applications, TFO probes need further improvement because transfection of TFOs into the cytoplasm and the nucleus is still a significant obstacle. Covalent attachment of lipophilic groups to the TFO backbone may provide the necessary membrane permeability so that the probes can reach the nucleus where they will interact with the DNA in the chromosomes.

In terms of cell biology, TINA-TFOs are ready for *in situ* applications for gene visualisation such as COMBO-FISH (See Introduction, Chapter 3). A study should be undertaken to find the optimal conditions and protocols for TFO applications. Given that further chemical modifications will provide nuclease resistance and membrane permeability, TINA-TFOs can be tested in live cells as the next step.

Gene visualisation is not the only application for which TINA-TFOs can be used. Fluorescently silent TFOs can also be a practical tool in qRT-PCR applications by replacing molecular beacons. Due to their strict sequence selectivity possible for TINA-TFOs, they can be used in the detection of single nucleotide polymorphisms (SNPs) on double-stranded DNA.

Here, we presented a work that leads to new possibilities in the development of functional nucleic acids. We believe this PhD thesis is a milestone in DNA triplex technology and the work

presented here will guide the future studies in developping methods for *in vivo* gene visualisation.

Chapter 8. Experimental Methods

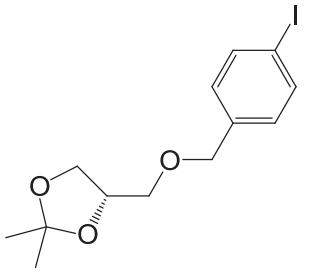
8.1 General Remarks

Thin layer chromatography (TLC) analysis was carried out using 60 F_{254} TLC plates purchased from Merck. All unmodified ONs were purchased from IDT (USA). IR spectra were recorded for azide containing compounds on a Nicolet 5700 FT-IR from Thermo Electron Corporation using an ATR attachment. For NMR spectroscopy the solutions of the compounds were prepared at approximately 20 mM in $CDCl_3$. All ^{1}H NMR, ^{13}C NMR and ^{31}P NMR spectra were obtained on 500 MHz Bruker instruments using Topspin software. $CDCl_3$ was used as an internal standard in ^{1}H NMR (δ = 7.29 ppm) and ^{13}C NMR (δ = 77.0 ppm). High resolution mass spectrometry was performed to identify compounds using a ZMD ESI mass spectrometer from Waters Micromass in MeOH or acetone.

8.2 Synthesis of TINA phosphoramidite

TINA phosphoramidite was synthesised for solid-phase DNA synthesis as described previously (Filichev, Gaber et al. 2006).

(*S*)-4-[(4-Iodobenzyloxy)methyl]-2,2-dimethyl-1,3-dioxolane (3):

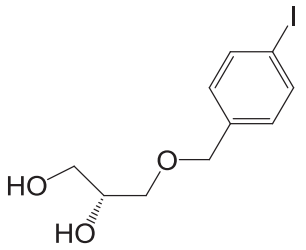


Chemical Formula: $C_{13}H_{17}IO_3$
Exact Mass: 348.02
Molecular Weight: 348.18

[(*S*)-2,2-Dimethyl-1,3-dioxolan-4-yl]methanol (**1**) (1.15 g, 8.7 mmol), 4-iodobenzyl bromide (**2**) (2.6 g, 8.7 mmol) and pulverised KOH (4.4 g, 78 mmol) were mixed and refluxed in dry toluene under Dean-Stark conditions overnight at 125 °C. After cooling, H_2O (30 mL) was added. Phases were separated and the water layer was washed with toluene (2 × 25 mL). Organic layers were

combined, washed with H_2O, dried over $MgSO_4$ and concentrated *in vacuo*. A viscous oil (**3**) was obtained. Yield: 2.9 g (95 %). ^{1}H NMR was in agreement with (Filichev, Gaber et al. 2006).

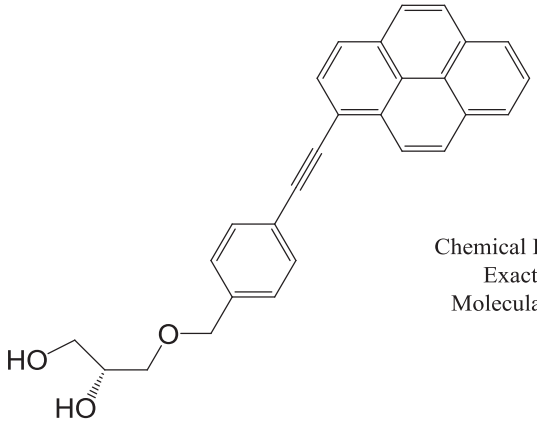
(*R*)-3-(4-Iodobenzyloxy)propane-1,2-diol (4):



Chemical Formula: $C_{10}H_{13}IO_3$
Exact Mass: 307.99
Molecular Weight: 308.11

Amberlite (H^+) resin (9 g) was washed with H_2O (3 × 30 mL) and added to the compound **3** (2.9 g) in H_2O (30 mL). After the mixture was stirred at 65 °C for 3 days, the solution was filtered and Amberlite (H^+) resin was removed. The water layer was removed by freeze-drying. The removed resin was washed with MeOH several times until a significant amount of the product was removed from the resin. The MeOH solution was evaporated. The combined water and MeOH residues were purified by silica gel column chromatography using MeOH (5 %) in DCM as an eluent to afford compound **4**. Yield: 2.5 g (97 %). ^{1}H NMR was in agreement with (Filichev and Pedersen 2005)

(*R*)-3-[4-(Pyrenylethynyl)benzyloxy]propane-1,2-diol (6):

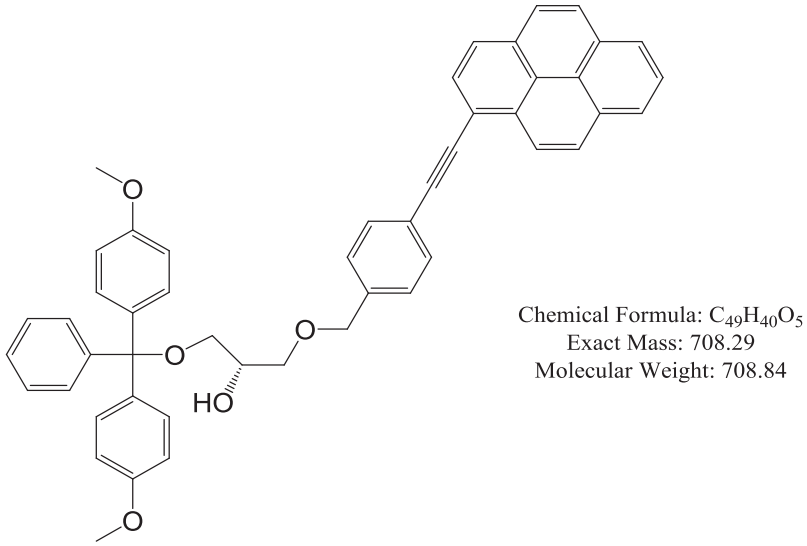


Chemical Formula: $C_{28}H_{22}O_3$
Exact Mass: 406.16
Molecular Weight: 406.47

Compound **4** (2.5 g, 8.1 mmol) was dissolved in Et_3N (10 mL) in dry dimethylformamide (DMF) (66 mL). The solution was bubbled with argon for 30 min. 1-Ethynylpyrene (**5**, 1.65 g, 7.73 mmol), CuI (0.087 g, 0.45 mmol), $Pd(PPh_3)_4$ (0.21 g, 0.18 mmol) were added under argon. The reaction mixture was stirred at 20 °C under argon for 3 days. DCM (150 mL) was added and

washed with disodium ethylenediaminetetraacetic acid (Na_2EDTA) (0.3 M, 150 mL). The organic layer was washed with H_2O (3 × 75 mL), dried with $MgSO_4$ and filtered. Solvents were evaporated *in vacuo* and the residue was co-evaporated with toluene/EtOH (30 mL, 1:1, v/v), affording (*R*)-1-O-[4-(pyren-1-ylethynyl)phenylmethyl]glycerol (**6**). Yield: 3 g (90 %). ^{1}H NMR was in agreement with (Filichev, Gaber et al. 2006).

(*S*)-1-[Bis(4-methoxyphenyl)(phenyl)methoxy]-3-[4-(pyrenylethynyl)benzyloxy]propanol (7):

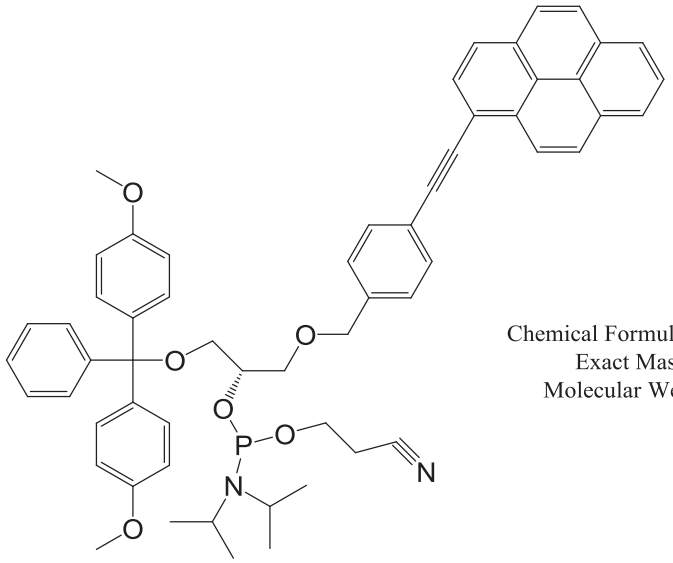


Chemical Formula: $C_{49}H_{40}O_5$
Exact Mass: 708.29
Molecular Weight: 708.84

Compound **6** (2.7 g, 6.63 mmol) was added to Et_3N (5 mL) in dry DCM (50 mL) and bubbled with argon for 30 min. Under argon 4,4′-dimethyloxytrityl chloride (DMTCl) (2.47 g, 7.3 mmol, 1.1 equiv.) was added and stirred overnight at 20 °C. After additional DMTCl (500 mg, 1.4 mmol) was added, the solution was incubated at 20 °C under argon overnight. After starting material disappeared on TLC, the reaction mixture was quenched with MeOH (2 mL), diluted with EtOAc (50 mL) and extracted with std. *aq.* $NaHCO_3$ (2 × 50 mL). The water phase was extracted with dry DCM (2 × 50 mL). The combined organic layers were dried with $MgSO_4$, filtered and concentrated *in vacuo*. The residue was purified by silica gel column chromatography with EtOAc (0-100 %) in hexane to afford compound **7** as yellow foam. Yield: 2.25 g (48 %). ^{1}H NMR ($CDCl_3$): δ= 2.48 (d, 1 H, *J* = 5.0 Hz, OH), 3.27 [m, 2 H, CH(OH)*CH*$_2$OCH$_2$], 3.65 (m, 2 H, CH_2ODMT), 3.80 (s, 6 H, 2 × OCH_3), 4.14 (m, 1 H, *CH*OH), 4.61 (s, 2 H, CH_2Ar), 6.85 (d, 4 H, *J* = 8.5 Hz, DMT), 7.27-7.47 (m, 11 H, DMT), 7.72 (d, 2 H, *J* = 8.0 Hz, phenyl), 8.04-8.69 (m, 9 H, pyren-1-yl) ppm. ^{13}C NMR ($CDCl_3$): δ= 55.2 (OCH_3),

64.3 (CH_2ODMT), 70.0 [CH(OH)CH_2OCH_2], 71.7 (CHOH), 72.9 (CH_2-phenyl), 86.1 [C(Ar)$_3$], 88.7, 94.9 (C≡C), 117.7, 127.7, 138.5, 139.4 (phenyl), 113.1, 124.5-131.8, 136.0, 144.8, 158.5 (DMT, pyren-1-yl) ppm.

(*S*)-1-[Bis(4-methoxyphenyl)(phenyl)methoxy]-3-[4-(pyrenylethynyl)benzyloxy]propanyl-2-cyanoethyl diisopropylphosphoramidite (8):

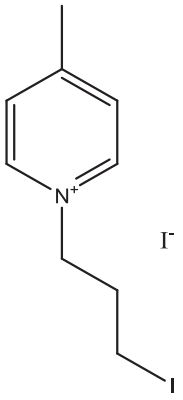


Chemical Formula: $C_{58}H_{57}N_2O_6P$
Exact Mass: 908.40
Molecular Weight: 909.06

Di*iso*propylammonium tetrazolide (1 g, 5.8 mmol) was washed with dry DCM and left under vacuum for 3 hours. Compound **7** (380 mg, 0.53 mmol) was dissolved in dry DCM (25 mL) and stirred under argon for 30 min. Dried di*iso*propylammonium tetrazolide and 2-cyanoethyl tetra*iso*propylphosphordiamidite (210 mg, 0.7 mmol) were added and mixture was stirred under argon at 20 °C overnight. Around 50 % conversion was detected on TLC. An additional 2-cyanoethyl tetra*iso*propylphosphordiamidite (105 mg, 0.35 mmol) was added to the solution and stirred under argon at 20 °C overnight. After overnight incubation, the reaction mixture was washed with H_2O (2 × 30 mL) and the layers were separated. The water layer was washed with dry DCM (25 mL). The combined organic layers were dried with $MgSO_4$ and filtered. Solvent was evaporated under vacuum. The residue was adsorbed to silica gel and purified by silica gel column chromatography with Et_3N (1 %), EtOAc (0-25 %) in hexane. UV-active fractions were combined and evaporated *in vacuo* affording the final compound **8** as yellow foam. Yield: 270 mg (56 %). ^{31}P NMR ($CDCl_3$): δ= 149.3, 149.5 ppm in a 3:2 ratio.

8.3 Synthesis of Cyanine Dye

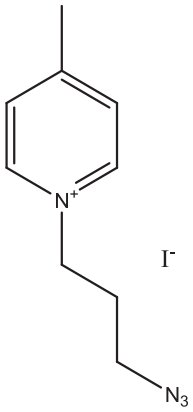
1-(3-Iodopropyl)-4-methylpyridinium iodide (15):



Chemical Formula: $C_9H_{13}I_2N$
Exact Mass: 388.91
Molecular Weight: 389.02

Compound **15** was synthesised as described previously (Kele, Mezo et al. 2009). 1,3-Diiodopropane (20 mmol, 5.92 g) and 4-methylpyridine (5 mmol, 0.47 g) were mixed in acetonitrile (10 mL) and refluxed for 20 hours. The resulting mixture was cooled to 20 °C and concentrated *in vacuo*. EtOAc (10 mL) was added and upon stirring a yellow precipitate was formed. The precipitate was filtered and washed with EtOAc (10 mL). The product was used in the next step without further purification. Yield: 1.69 g (86 %).

1-(3-Azidopropyl)-4-methylpyridinium iodide (16):

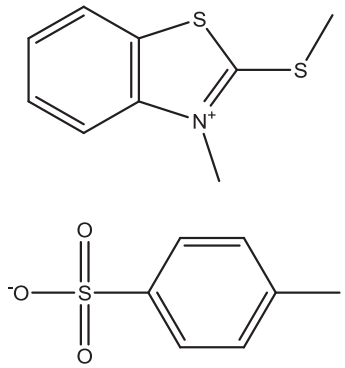


Chemical Formula: $C_9H_{13}IN_4$
Exact Mass: 304.02
Molecular Weight: 304.13

Compound **15** (4.34 mmol, 1.69 g) was mixed with NaN_3 (11.3 mmol, 0.74 g) and refluxed in acetonitrile (20 mL) overnight. The resulting mixture was filtered to remove excess NaN_3 and concentrated *in vacuo*. The residue was thoroughly washed with DCM (20 mL) twice and filtered. Combined fractions were concentrated *in vacuo* to give 1.08 g of oil (Yield: 82 %). ^1H

NMR showed that the mixture contained 1,1′-(propane-1,3-diyl)-bis-(4-methylpyridinium) iodide as impurity (10 %). ^{1}H NMR ($CDCl_3$): δ = 2.40 (m, 2 H, $CH_2CH_2CH_2$), 2.72 (s, 3 H, CH_3), 3.63 (t, J = 6.0, 2 H, $CH_2CH_2N_3$), 5.07 (t, J = 7.5, 2 H, CH_2CH_2Ar), 7.91 (d, J = 6.5, 2 H, Ar), 9.32 (d, J = 6.5, 2 H, Ar) ppm; IR-ATR (cm^{-1}): 2098.6 (azide).

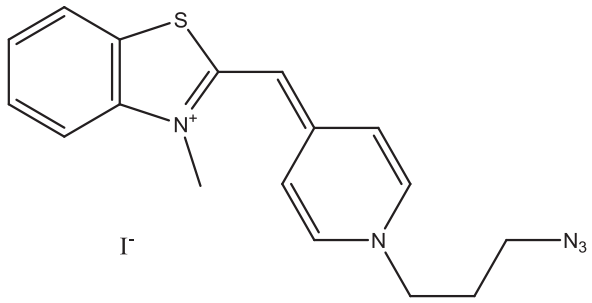
3-Methyl-2-(methylthio)benzothiazolium 4-methylbenzenesulfonate (19):



Chemical Formula: $C_{16}H_{17}NO_3S_3$
Exact Mass: 367.04
Molecular Weight: 367.51

Compound **19** was synthesised as described previously in the literature (Bethge, Jarikote et al. 2008). Methyl 4-methylbenzenesulfonate (35 mmol, 6.6 g) and 2-(methylthio)benzothiazole (33 mmol, 6 g) were fused at 130 °C for 1 hour. The solution was cooled down to 70 °C, acetonitrile (30 mL) was added and refluxed for 30 min. The mixture was left to cool down to 20 °C. Pale yellow solid was precipitated, filtered and washed with acetone (2 × 30 mL). The precipitate was dried in *vacuo* line overnight giving 10.4 g of solid (Yield: 84 %). NMR data are in agreement with the published literature (Bethge, Jarikote et al. 2008). ^{1}H NMR ($CDCl_3$): δ = 2.38 (s, 3 H, CH_3Ar), 3.16 (s, 3 H, CH_3S), 4.18 (s, 3 H, CH_3N), 7.24 (d, J = 8, 2 H, Ar), 7.71 (d, J = 8, 2 H, Ar), 7.76 (d, J = 7, 2 H, Ar), 7.88 (d, J = 7, 2 H, Ar), 8.11 (d, J = 8, 2 H, Ar), 8,23 (d, J = 8, 2 H, Ar) ppm.

(Z)-1-(3-Azidopropyl)-4-[(3-methylbenzothiazolylidene)methyl]pyridinium iodide (20):



Chemical Formula: $C_{17}H_{18}IN_5S$
Exact Mass: 451.03
Molecular Weight: 451.33

Compound **19** (3.55 mmol, 1.5 g) and compound **16** (3 mmol, 0.9 g) were mixed in 20 mL EtOH and 1 mL piperidine (10 mmol) and refluxed overnight. The solvent was evaporated *in vacuo* and the residue was purified by silica gel column chromatography using MeOH/DCM (1:10, v/v), affording 0.1 g of the final product **20**. (Yield: 8 %) ^{1}H NMR ($CDCl_3$): δ = 2.28 (m, 2 H, $CH_2CH_2CH_2$), 3.56 (t, J = 6.5, 2 H, $CH_2CH_2N_3$), 3.79 (s, 3 H, CH_3S), 4.56 (t, J = 7.5, 2 H, CH_2CH_2Ar), 6.15 (s, 1 H, CH), 7.30 (m, 2 H, Ar), 7.45 (d, J = 7, 2 H, Ar), 7.50 (t, J = 8, 1 H, Ar), 7.68 (d, J = 8, 1 H, Ar), 8.42 (d, J = 7.5, 2 H, Ar) ppm. ^{13}C NMR ($CDCl_3$): δ = 30.2 ($CH_2CH_2CH_2$), 33.6 (CH_3), 46.9 ($CH_2CH_2N_3$), 55.8 (NCH_2CH_2), 90.5, 122.5, 124, 124.2, 128 (Ar) ppm. HR-MALDI-MS: *m/z* calcd. for $C_{17}H_{18}N_5S^+$ 324.13; found 324.52; IR-ATR (cm^{-1}): 2096.6.0 (azide).

8.4 Synthesis of Minor Groove Binders (MGBs)

All MGBs used in this study were provided by Prof. A. Boutorine from National Museum of Natural History, Paris. The synthesis of MGBs was performed in cycles for each monomer. Every cycle consisted of four steps, taking approximately 30 min to complete (Table 8.1).

Table 8.1 Manual synthesis cycle of MGB.

Steps	Reagent	Incubation time/mode
Deprotection	2×TFA, phenol, water (92.5:5:2.5)	1+2 min mix
Washing	DCM	30 s flow
	DMF	1 min flow
Coupling	Activated monomer	20 min mix
Washing	DMF	30 s flow
	DCM	1 min flow

At the beginning of the synthesis, Boc-β-alanine-Pam-resin (400 mg for 0.1 mmol scale synthesis) was placed into peptide synthesis vessels, rinsed twice in DCM and mixed for 15 min. Also monomers to be used at each coupling were separated in glass vials (ca. 4 equivalents of necessary monomer, 0.4 mmol of monomer for 0.1 mmol scale synthesis). 3.4 equivalents of *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) were

added to Boc-Pyrrole vials and 7.2 equivalents of HATU to Boc-β-alanine (Boc-β-Ala) and Boc-γ-aminobutyric acid (Boc-γ-Abu) after separation of monomers. Monomers were activated prior to coupling as described below.

Activation of monomers: 0.75 mL 1:1 DMF:NMP (*N*-Methyl-2-pyrrolidone) mixture per 100 mg of resin was mixed with the monomers and HATU. Prior to coupling (ca. 3 min before), 12 equivalents of di*iso*propylethylamine (DIEA) was added for activation of monomers.

Boc Deprotection: For Boc deprotection, ca. 1.6 mL of 92.5 % TFA, 5.0 % phenol, 2.5 % water mixture was added per 400 mg resin and vessels were mixed for 1 min. After draining the vessels, the same amount of mixture was added and mixed for 2 minutes.

Washing of the resin: Resin was washed with DCM (ca. 15-20 mL) for 30 s in vacuo. Resin was then washed with the same amount of DMF for 1 min.

Coupling of the monomers: Proper activated monomer solution was added to the peptide synthesis vessels and mixed for 20 min. When coupling was completed, the mixture was drained from the vessels.

The cycle was repeated until all monomers were coupled. Upon the completion of the synthesis, resin was washed with DMF, DCM, MeOH, and ether, thoroughly. Resin was then dried *in vacuo* before aminolysis.

100 mg of the resin was dried and treated with 0.75 mL dimethylaminopropylamine (DMPA) under argon. After 16 hours of incubation, the resin was cooled down to 20 °C, washed with 0.75 mL of DCM several times and filtered. 5-8 volumes of diethyl ether were added to the amine solution, affording a thick white precipitate. After 30 min incubation at -20 ˚C, the solution was centrifuged for 5 min at 13000 rpm and the supernatant was removed. The crude product was then dissolved in 3 mL of 15 % acetonitrile in water with 0.1 % v/v TFA, filtered and purified on reverse-phase HPLC. Buffer A (0.1 % v/v TFA in water) and Buffer B (acetonitrile). Gradients: 85 % buffer A and 15 % buffer B for 5 min, linear gradient to 25 % buffer B over 60 min. Products were eluted at 21-23 % buffer B.

8.5 Other Modifications

(*S*)-1-[Bis(4-methoxyphenyl)(phenyl)methoxy]-3-[(4-ethynylbenzyl)oxy]propan-2-yl-(2-cyanoethyl)di*iso*propylphosphoramidite (**9**) and 5-ethynyl-2′-deoxyuridine phosphoramidite (**10**) were prepared according to the published procedures (Graham, Parkinson et al. 1998; Geci, Filichev et al. 2007) and available in our laboratory. Modified phosphoramidites, 2′-*O*-propargyl-uridine (**11**) and 2′-*O*-propargyl-adenosine phosphoramidites (**12**) were purchased from GlenReseach (USA).

8.6 Oligonucleotide Synthesis

Oligonucleotides were synthesised with an Mer-Maid 4 automated DNA synthesiser from BioAutomation Corporation using 4,5-dicyanoimidazole (DCI) as an activator. Oxidation and deprotection times were set to 40 s and the activation time at 60 s for 1.0 μmol synthesis scale. For coupling of modified nucleotides, automated DNA synthesis was paused after the deblocking step. TINA phosphoramidites or other modified phosphoramidites (10 mg per coupling) were added into each 1 μmol column and 750 μL of an activator (0.25 M DCI in dry acetonitrile) was directly injected onto the column under argon by the automated DNA synthesiser. Coupling time was extended for hand coupling (up to 5 min). For fluorescein-containing ONs 5′-fluorescein-containing CPG supports from GlenResearch (USA) were used. All oligonucleotides discussed in Chapter 3 and Chapter 5 were synthesised in DMT-on mode.

8.7 Post-synthetic CuAAC Reaction for Fluorescent TFOs

For the synthesis of fluorescent TFOs discussed in Chapter 6, CuAAC reactions were carried out using CPG-bound DNAs possessing ethynyl modifications. Oligonucleotides were synthesised in DMT-off mode. 12 mg of the CPG-bound DNA (0.33 μmol) was weighed and added to a 15 mL microwave reaction vial along with the appropriate azide (Compound **8**, 3.3 μmol, 10 eq, 200 μL of a 16.5 μM stock solution in degassed DMSO). Freshly prepared $CuSO_4 \cdot 5H_2O$ (0.4 μmol, 1.2 eq, 10 μL of a 40 mM stock solution in degassed H_2O), tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA, 0.4 μmol, 1.2 eq, 10 μL of a 40 mM stock solution in degassed

DMSO), sodium ascorbate (1.5 μmol, 4.5 eq, 30 μL of a 50 mM stock solution in degassed H_2O) and triethylammonium acetate (TEAA, 15 μL of a 2 M solution in degassed H_2O at pH 7.0) were added. The reaction mixture was irradiated in a microwave synthesiser (Discover, CEM Corporation, 70 °C, 200 watts, 25 min) and left to shake at 20 °C overnight.

The content of the reactions were transferred into 2.0 mL microcentrifuge tubes. DMSO (1.5 mL) was added and centrifuged (14500 rpm for 1 min). Solvents were decanted and the washing step was repeated until the supernatant no longer showed any colour. The coloured CPG was then washed with H_2O (1.5 mL) and EtOH (1.5 mL) to remove any remaining inorganic salts and to remove DMSO. Residual solvent was evaporated at 55 °C for 30 min.

8.8 Purification of Oligonucleotides

To purifiy the oligonucleotides synthesised in DMT-off mode, DNA-bound CPG supports were removed from columns and treated with 32 % NH_4OH (1 mL). Samples were sealed and after preliminary incubation at 20 °C for an hour, they were kept at 55 °C overnight, except for the fluorescently silent TFOs (**T2-T6**, **T8** and **T9** from Chapter 6), which were incubated at 20 °C overnight. The supernatant was transferred into 16 mL centrifuge tubes with filtering to remove the CPG support. H_2O (3 mL) was added and the mixture was frozen using liquid nitrogen. Samples were freeze-dried overnight *in vacuo*.

TINA-conjugated ONs from Chapter 4 and Chapter 6 (except **T5**, **T6**, **T8** and **T9**) were purified using reverse-phase HPLC instrument (Waters™ 600) using a C-18 column from Alltech with 250 mm length and 10 mm internal diameter. The residue was dissolved in H_2O (1 mL). Buffer A [0.05 M TEAA in H_2O (pH = 7.0)] and buffer B (75 % acetonitrile, 25 % H_2O). Flow 2.5 mL min^{-1}. Gradients: 2 min 100 % buffer A, linear gradient to 100 % buffer B in 48 min, linear gradient to 100 % buffer A in 2 min, 100 % buffer A for 10 min. The corresponding UV-active fractions were freeze-dried.

In cases where separation by HPLC was not efficient (MGB-TFO **3** and MGB-TFO **7** from Chapter 5, TFOs **T5, T6, T8, T9** from Chapter 6), gel purification was performed using 20 % polyacrylamide gels (PAGE) under denaturing conditions (7 M urea). After electrophoresis, bands were visualised using UV shadowing, cut and placed into 2 mL microcentrifuge tubes. Gel

fragments were then soaked with formamide (~150 µL) and shaken at 20 °C for two hours. Formamide solution was transferred into new tubes and DNA was precipitated from NaOAc (3 M, 50 µL) and EtOH (1 mL) mixture (1:10) or using 3 % $LiClO_4$ in acetone. The pellets were dried and dissolved in 100 µL ddH_2O, followed by incubation at 55 °C for 30 min and occasional vortexing, which was required to dissolve TFO completely.

To purify 5′-*O*-DMT-protected ONs (DMT-ONs) prepared in Chapter 3 (all TINA-TFOs) and Chapter 5 ($p(EG)_6$-TFO **3** and $p(EG)_6$-TFO **7**), DNA-bound CPG supports were removed from the columns and treated with 32 % NH_4OH (1 mL). Samples were sealed and after preliminary incubation at 20 °C for an hour, they were kept at 55 °C overnight. Supernatant were transferred into 16 mL centrifuge tubes with filtering to remove the CPG support. H_2O (3 mL) was added and the mixture was frozen using liquid nitrogen. Samples were freeze-dried overnight *in vacuo*. DMT-ONs were purified using reverse-phase HPLC instrument (Waters™ 600) using a C-18 column from Alltech with 250 mm length and 10 mm internal diameter. The residue was dissolved in H_2O (1 mL). Buffer A [0.05 M TEAA in H_2O (pH = 7.0)] and buffer B (75 % acetonitrile, 25 % H_2O). Flow 2.5 mL min^{-1}. Gradients: 2 min 100 % buffer A, linear gradient to 100 % buffer B in 48 min, linear gradient to 100 % buffer A in 2 min, 100 % buffer A for 10 min. The corresponding UV-active fractions were freeze-dried.

After purification, DMT-ONs were treated with 80 % aqueous AcOH (100 µL) for 20 min at 20 °C to remove 5′-*O*-DMT group. NaOAc (3 M, 50 µL) followed by EtOH (1 mL) were added and vortexed. Solutions were cooled to -18 °C for two hours in order to precipitate the ONs. Samples were centrifuged for 20 min at 13000 rpm and supernatant was removed. Pellets were washed with EtOH (2 × 500 µL), centrifuged and the supernatant was removed. Then, the rest of the solvent was evaporated at 50 °C for 15 min. 100 µL H_2O was added to dissolve the ONs.

8.9 Post-synthetic Conjugation of Minor Groove Binders

After purification of p(EG)-TFOs, a portion of the ONs (200 µL) was used for MGB coupling. ONs were treated with 2 µL 8 % *aq.* cetyl trimethylammonium bromide (CTAB) and centrifuged at 13000 rpm for 5 min. CTAB treatment and centrifugation was repeated until all of the DNA was precipitated from water. Solvent was decanted and the DNA was dried *in vacuo* for 30 min. Dry CTAB salt of DNA (200 µmol) was mixed with dimethylaminopyridine (5 mg, 0.04 mmol),

dipyridyldisulfide (6.6 mg, 0.03 mmol) and triphenylphosphine (7.9 mg, 0.03 mmol) in 100 µL DMSO and incubated at 20 °C for 15 min. Dry TFA salt of MGB (1 mg per 600 µg of ON) was dissolved in DMSO in the presence of Et_3N (5 µL) and added to the solution containing the ON. After 16 hours of coupling, MGB-TFOs were precipitated with 3 % $LiClO_4$ in acetone. MGB-bound TFOs were purified using denaturing gel electrophoresis as described in 8.8.

8.10 Polyacrylamide Gel Electrophoresis

Purities of ONs were confirmed by denaturing gel electrophoresis using 20 % polyacrylamide gel (0.75 mm thickness, 19:1 acrylamide/bisacrylamide ratio). Gels were prepared in 1× TBE buffer (100 mM Tris, 90 mM boric acid, and 10 mM EDTA) under denaturing conditions (7 M urea). ONs were loaded onto gels after preincubation at 90 °C for 10 min. For non-denaturing 20 % PAGE, gels were prepared in HEPES buffer (10 mM) in corresponding salt concentrations (50-150 mM NaCl or KCl) with 0.75 mm thickness, 19:1 acrylamide/bisacrylamide ratio. ONs were prepared at 100 µM strand concentration, incubated in 10 mM HEPES in same salt concentrations and heated up to 90 °C for between 10 and 30 min before cooling down and incubating at 4 °C overnight. All gel electrophoresis were performed at 37 °C. After the electrophoresis (denaturing or non-denaturing), gels were first observed using the intrinsic fluorescence of the pyrene moiety in TINA by irradiation of the gel at 362 nm before staining with 5 % Stains-All® in 50 % water/formamide for 5–10 min and then destained in H_2O until complete washing of the dye from the gel background occurred.

8.10.1 Dissociation Constant (K_d) Measurements

In order to determine the dissociation constants (K_d), a set of solutions were prepared for each TFO at various concentrations, in the presence of the target duplexes (target **I-V** from Chapter 3), with a labeled oligopyrimidine strand possessing fluorescein at the 5′-end (60 nm) in HEPES buffer (50 mm), pH 7.2, containing $MgCl_2$ (5 mm). This buffer is most suitable for electrophoresis experiments at the indicated pH and it was also used in order to compare our results with those previously obtained. The salt concentrations used were 50 mm NaCl, 150 mm NaCl, or 150 mm KCl. Glycerol containing bromophenol blue and xylene cyanol (up to 10 % of

the glycerol concentration) was added to the samples (20 mL final volume), which were then heated at 90 °C for 3 min, slowly cooled down to 20 °C and then incubated at 4°C overnight. Samples (10 mL) were charged on a non-denaturing 20 % polyacrylamide gel (0.75 mm thickness, 19:1 acrylamide/bisacrylamide ratio) in HEPES buffer (50 mm) at 20 °C, then PAGE was performed at 37 °C in the same running buffer, with salt concentrations as indicated.

After running, gels were scanned on a FLA-5000 imager (FujiFilm) using a laser excitation wavelength of 473 nm and a 510 LP filter to detect the fluorescein. The image was analysed using Image-Quant software to determine the percentage of triplex conversion. K_d values were calculated according to the protocol (Boutorine and Escude 2007) using the formula below where x represents the proportion of triplex conversion, S_0 and D_0 represent initial TFO and duplex concentrations, respectively.

$$\log(S_0 - D_0 x) = \log\left(\frac{x}{(1-x)}\right) + \log(K_d) \qquad \text{(Formula 2)}$$

The term $\log(S_0 - D_0 x)$ was ploted against $\log\frac{x}{1-x}$, and the linear line was fit. The point where x equaled 0.5, $\log(S_0 - D_0 x)$ was equal to $\log(K_d)$. These calculations were facilitated by the graph ploted using convertsion rates obtained from the Image-Quant software.

8.10.2 Kinetics experiments using PAGE

Kinetic experiments for triplex formation by TFO **3** and **7** from Chapter 3 were performed using so-called “stop-in-gel” retardation techniques at 20 °C. The whole mixture (100 µL) contained the fluorescently-labeled DNA duplex and the third strand in HEPES buffer supplemented with glycerol (5 %) and electrophoretic dyes. This mixture was prepared as described above for K_d measurements except that the TFO was added at the last moment without overnight incubation. Aliquots (10 µL) were taken at time intervals of 0, 15, 30, etc. min after mixing and charged consequently into wells of the same non-denaturing gel, and then immediately submitted to electrophoresis. After running, gels were scanned on a FLA-5000 imager and the labeled target DNA in the triplex was calculated using Image-Quant software.

8.11 UV Spectroscopy

UV-Vis spectroscopy was performed using Cary 100Bio UV-Vis spectrometer using quartz cuvettes with 1 cm pathlength and a 2 × 6 multicell block with a Peltier temperature controller.

8.11.1 Determination of Concentration

Concentrations of ONs were calculated using following formula where Abs_{260} refers to absorbance at 260 nm, c is concentration, l is the pathlength and ε is the extinction coefficient.

$$c = \frac{\mathrm{Abs}_{260}}{\varepsilon_{260} \times l} \qquad \text{(Formula 3)}$$

Extinction coefficients of ON were calculated using the extinction coefficients of each nucleoside at 260 nm. Extinction coefficients (ε) of unmodified nucleotides [L/(mol.cm)]: dA (15400), dG (11700), dT (8800), dC (7300). Extinction coefficients (ε) of modified nucleotides [L/(mol.cm)]: **X** (22000, TINA), **FAM** (13000), **K** (19300), $\mathbf{L^U}$ (19300), $\mathbf{M^U}$ (19300), $\mathbf{M^A}$ (25900).

8.11.2 Thermal Difference Spectra

Thermal difference spectra were determined for the DNA complexes (with 10 μM strand concentration) discussed in Chapter 3 by subtraction of UV-Vis spectra obtained in Li cacodylate buffer (10 mM) and NaCl or KCl (110 mM) at 20 °C after 2-7 days on incubation at 4°C from UV-Vis spectra obtained after 30 min incubation at 90 °C under same conditions. Thermal difference spectra were determined for DNA complexes (with 1.0 μM strand concentration) discussed in Chapter 6 by subtraction of UV-Vis spectra obtained in Li cacodylate buffer (10 mM), NaCl (110 mM) and $MgCl_2$ (10 mM) at pH 5.0 or 7.2 at 20 °C after 2-7 days on incubation at 4°C from UV-Vis spectra obtained after 30 min incubation at 90 °C under same conditions.

8.11.3 Determination of Melting and Annealing Temperatures

The melting temperatures (T_m [°C]) were determined for DNA complexes [prepared at 10 μM strand concentration in Li cacodylate buffer (10 mM) and NaCl or KCl (110 mM)] discussed in Chapter 4 as the maxima of the first derivative plots of the melting curves obtained by measuring absorbance at 373 nm against increasing temperatures (0.18, 0.5 or 1 °C/min). The change in absorbance was also recorded at 295 nm. See Appendix Figure 1.

The melting temperatures (T_m [°C]) were determined for DNA complexes [prepared at 1.0 μM strand concentration in Li cacodylate buffer (10 mM), NaCl (110 mM) and $MgCl_2$ (10 mM)] discussed in Chapter 6 as the maxima of the first derivative plots of the melting curves obtained by measuring absorbance at 275 nm against increasing temperatures (0.18 °C/min).

8.12 Association Rate Constant (k_{on}) Measurements

Association rate constants (k_{on}) for G-quadruplex formations were measured using UV spectroscopy as previously described (Mergny, De Cian et al. 2005). ONs were preincubated at 90 °C for 30 min and transferred into Li cacodylate buffer (10 mM) and NaCl (110 mM) solution at 20 °C. As soon as the ONs were diluted in a buffer, changes in absorbance at 275 and 280 nm were then recorded every 5 min for 12.5 h. The data were used to plot the change in absorbance versus time (s) in Kleida Graph software. A curve was fit to the data using the following equation (Mergny, De Cian et al. 2005):

$$Abs_t = Abs_f + (Abs_i - Abs_f) \times (1 + (n-1) \times k_{on} \times C^{n-1} \times t)^{\frac{1}{1-n}} \qquad \text{(Formula 4)}$$

Abs_i refers to absorbance of the sample when all strands exist in unfolded, single-stranded form. Abs_f represent the absorbance of the sample when equilibrium is reached between folded and unfolded forms of the ONs. C represents the strand concentration, *n* represents the order of the reaction, which is 4 for tetramolecular G-quadruplexes, and *t* represents the time elapsed in seconds.

8.13 Circular Dichroism (CD) Spectroscopy

CD spectra were recorded using a Chirascan CD spectrophotometer (150 W Xe arc) from Applied Photophysics with a Quantum Northwest TC125 temperature controller. CD spectra were recorded between 200 and 500 nm with 1 nm intervals, 120 nm/min scan rate and 1 cm pathlength. CD spectra of DNA complexes discussed in Chapter 3 were recorded at 1.0 μM strand concentration in Na cacodylate buffer (20 mM), NaCl (50 mM) and $MgCl_2$ (5 mM) at pH 7.2, 20 °C. CD spectra of DNA complexes discussed in Chapter 4 were recorded at 10 μM strand concentration in Li cacodylate (10 mM) supplemented with NaCl or KCl (110 mM) at pH 7.2, 20 °C. CD spectra of DNA complexes discussed in Chapter 6 were recorded at 1.0 μM strand concentration in Li cacodylate (10 mM) supplemented with NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 5.0 or 7.2, 20 °C.

8.14 Fluorescence Spectroscopy

Fluorescence spectroscopy studies were performed using a FluoroMax-4 Spectrofluorometer from HORIBA Scientific. Fluorescence emission spectra were recorded for all TINA-containing ONs between 380 and 600 nm when excited at 373 nm. Samples were prepared as described in the previous section (Chapter 8.13). Fluorescence emission spectra were recorded for all fluorescein-containing ONs between 380 and 700 nm when excited at 373 nm.

Fluorescence spectroscopy has also been used for studying the melting of fluorescein-containing **GXGF** from Chapter 4. The fluorescence emission at 520 nm was recorded during heating with 0.2 °C/min temperature ramp when excited at 465 nm in 10 mM Li cacodylate buffer and 110 mM NaCl at pH 7.2, 20 °C and 10 μM strand concentration.

8.14.1 Quantum Yield (Φ_F) Measurements

Quantum yield measurements were performed in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) using FluoroMax-4 spectrofluorometer from HORIBA Scientific with an integrating sphere attachment. Quantum yields were measured by division of the number of photons emitted by the number of the photons absorbed. Number of photons absorbed was

measured by subtracting the photons absorbed by the sample from the photons absorbed by the buffer between 450 nm and 470 nm at 20 °C using excitation at 460 nm. Number of photons emitted was measured by subtracting the number of the photons emitted by the sample from the number of photons emitted by the buffer between 470 nm and 600 nm at 20 °C using excitation at 460 nm. The resulting ratio equals the quantum yield (Φ_F) of the particular sample. Calculations were carried out by FluorEssence software from HORIBA Scientific.

8.15 NMR Spectroscopy of Oligonucleotides

NMR spectroscopy of G-quadruplexes formed by **GXG** and **TXG** from Chapter 4 were performed on 700 MHz Brüker instrument. ONs (100, 500 or 1000 μM) were incubated overnight in 10 % D_2O, 10 mM Na^+ phosphate buffer supplemented with KCl (0 or 10 mM), pH 7.0 at 20 °C prior to NMR spectroscopy. Trimethylsilyl propionate (TSP, 25 μM) was used as internal standard in ^1H NMR (δ = 0.003 ppm).

8.16 Mass Spectrometry of Oligonucleotides

Mass spectrometry analysis of modified ONs was performed using following methods.

Method 1: Mass spectrometry analysis was performed using an electrospray Q-TOF (ES Q-TOF MS) on a Q-Star instrument (from Applied Biosystems) in the negative-ion mode in water.

Method 2: Mass spectrometry analysis was performed using obtained using an Autoflex MALDI-TOF instrument (from Brüker Daltonics) in the negative-ion mode using 3-hydroxypicolinic acid as a matrix and dibasic ammonium citrate as a co-matrix. ONs were desalted using C_{18} ZipTips (Millipore) prior to loading in the MALDI plate.

Method 3: Mass spectrometry analysis was performed using a M@LDI micromass instrument (from Waters) in the positive-ion mode using 2-aminobenzoic acid or 6-aza-2-thiothymine as matrices and dibasic ammonium citrate or imidazole as co-matrices. ONs were desalted using C_{18} ZipTips (Millipore) prior to loading in the MALDI plate.

Method 4: Mass spectrometry analysis was performed using a Voyager-DE PRO instrument (from Applied Biosystems) in the negative-ion mode using 3-hydroxypicolinic acid as a matrix

and imidazole as a co-matrix. ONs were desalted using C_{18} ZipTips (Millipore) prior to loading on the MALDI plate.

Table 8.2 Results of mass spectroscopy analysis of ONs synthesised.

Oligonucleotide	Calculated m/z, Da	Observed m/z, Da
TFO **1**	6394.5	6390.3 [a]
TFO **2**	6394.5	6393.3 [a]
TFO **3**	6394.5	6393.2 [a]
TFO **4**	6863.0	6863.2 [a]
TFO **5**	6394.5	6393.6 [a]
TFO **6**	6394.5	6393.7 [a]
TFO **7**	5926.1	5925.5 [a]
TFO **8**	6466.7	6469.9 [b]
TFO **9**	6863.0	6867.0 [b]
TFO **12**	5637.0	5633.3 [b]
TFO **13**	5877.0	5875.6 [b]
ABL1-2X	6391.3	6389.5 [b]
ABL6-2X	6375.3	6374.6 [b]
XTG	2331.7	2330.6 [c]
TXG	2331.7	2330.6 [c]
GXG	2331.7	2329.6 [c]
GXX	2800.1	2801.3 [c]
G3X	2990.1	2989.6 [c]
GXGF	2901.3	2901.2 [c]
TXGF	2901.3	2901.9 [c]
$p(EG)_6$-TFO **3**	6815.3	6815.1 [c]
MGB-TFO **3**	7874.5	- [d]
$p(EG)_6$-TFO **7**	6347.2	6347.0 [c]
MGB-TFO **7**	7406.4	- [d]
T2	4700.2	4703.6 [c]
T3	4459.0	4461.1 [c]
T4	4473.0	4474.9 [c]
T5	5791.0	5791.6 [e]
T6	5820.1	5821.0 [e]
T8	5786.9	5791.6 [e]
T9	6738.9	6736.7 [e]

[a] measured using method 1, [b] measured using method 2, [c] measured using method 3, [d] could not be detected, [e] measured using method 4.

Appendix

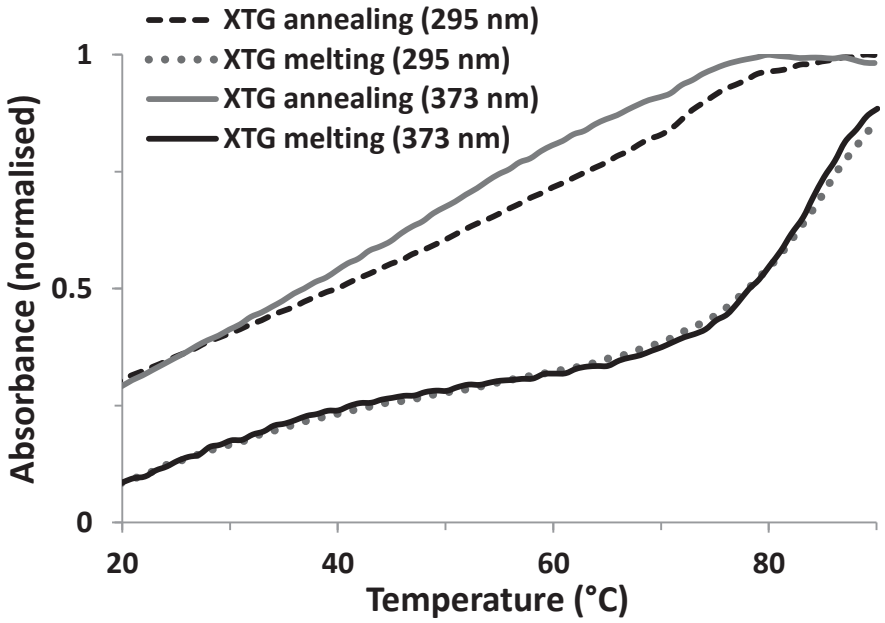


Figure 1 Melting and annealing profiles of **XTG** after incubation in 110 mM NaCl, 10 mM Li cacodylate buffer at pH 7.2. Oligonucleotide concentration was 10 μM. The profiles are based on absorbance data recorded at 373 nm (solid lines) and 295 nm (dashed and dotted lines) with 0.18 °C/min temperature ramp.

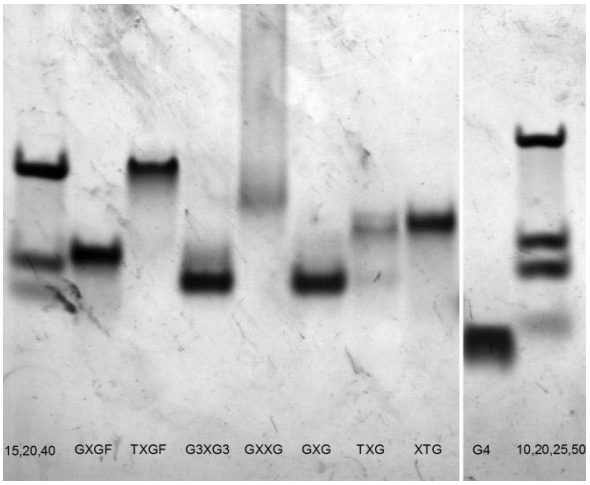


Figure 2 Native PAGE (20 %) at 37 °C, stained with Stains-All® of oligonucleotides (100 μM) in the presence of 110 mM NaCl, 10 mM Li cacodylate at pH 7.2, 4 °C. Ladder contains 10-, 15-, 20-, 25-, 40- and 50-mer oligothymidylates.

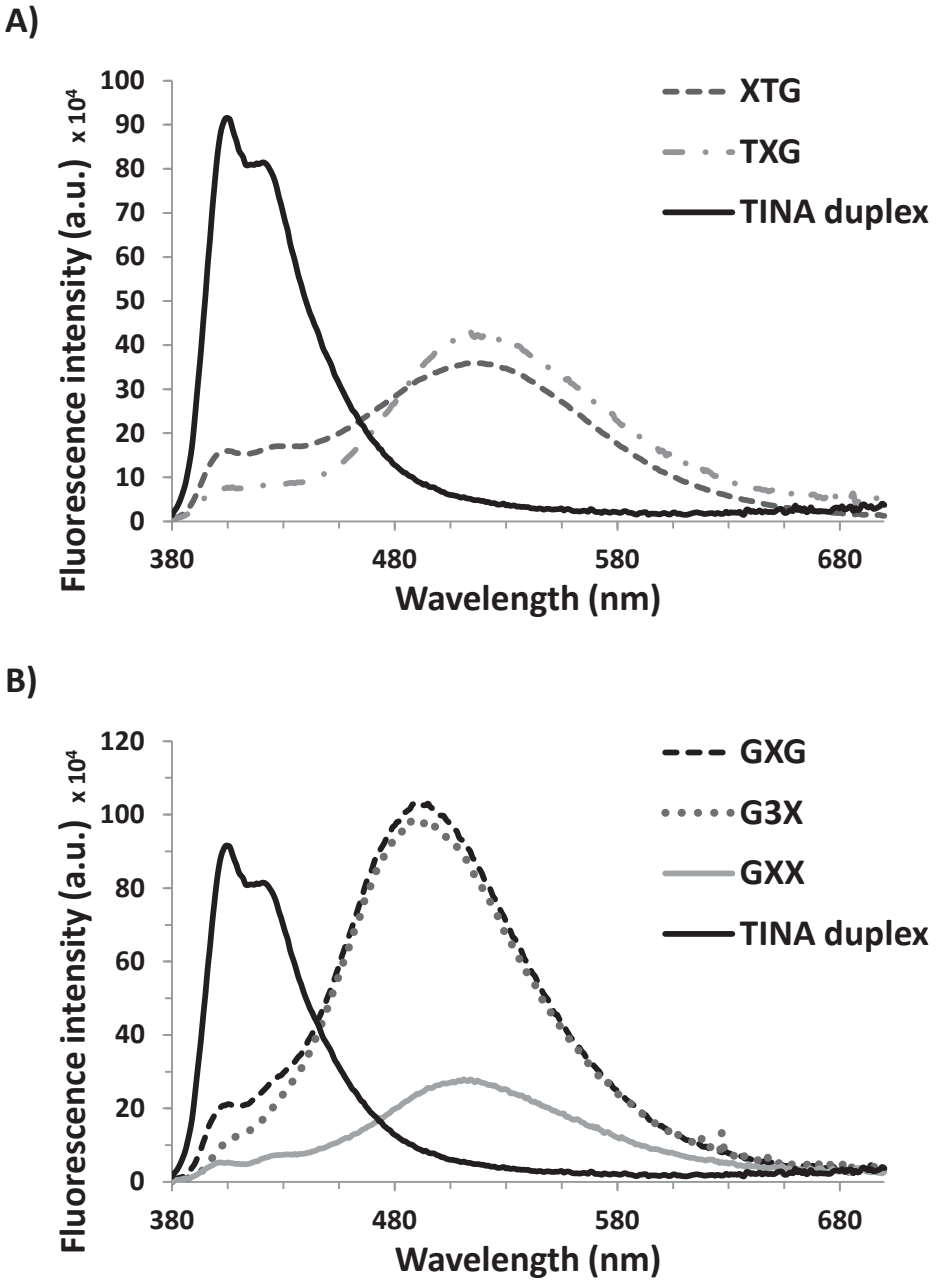


Figure 3 A) Fluorescence emission spectra of **XTG**, **TXG** and **TINA duplex** at 10 μM strand concentration in the presence of 110 mM KCl in 10 μM Li cacodylate buffer at pH 7.2, 20 °C. λ_{ex} = 373 nm. B) Fluorescence emission spectra of **GXG**, **G3X**, **GXX** and **TINA duplex** at 10 μM strand concentration in the presence of 110 mM KCl in 10 μM Li cacodylate buffer at pH 7.2, 20 °C. λ_{ex} = 373 nm.

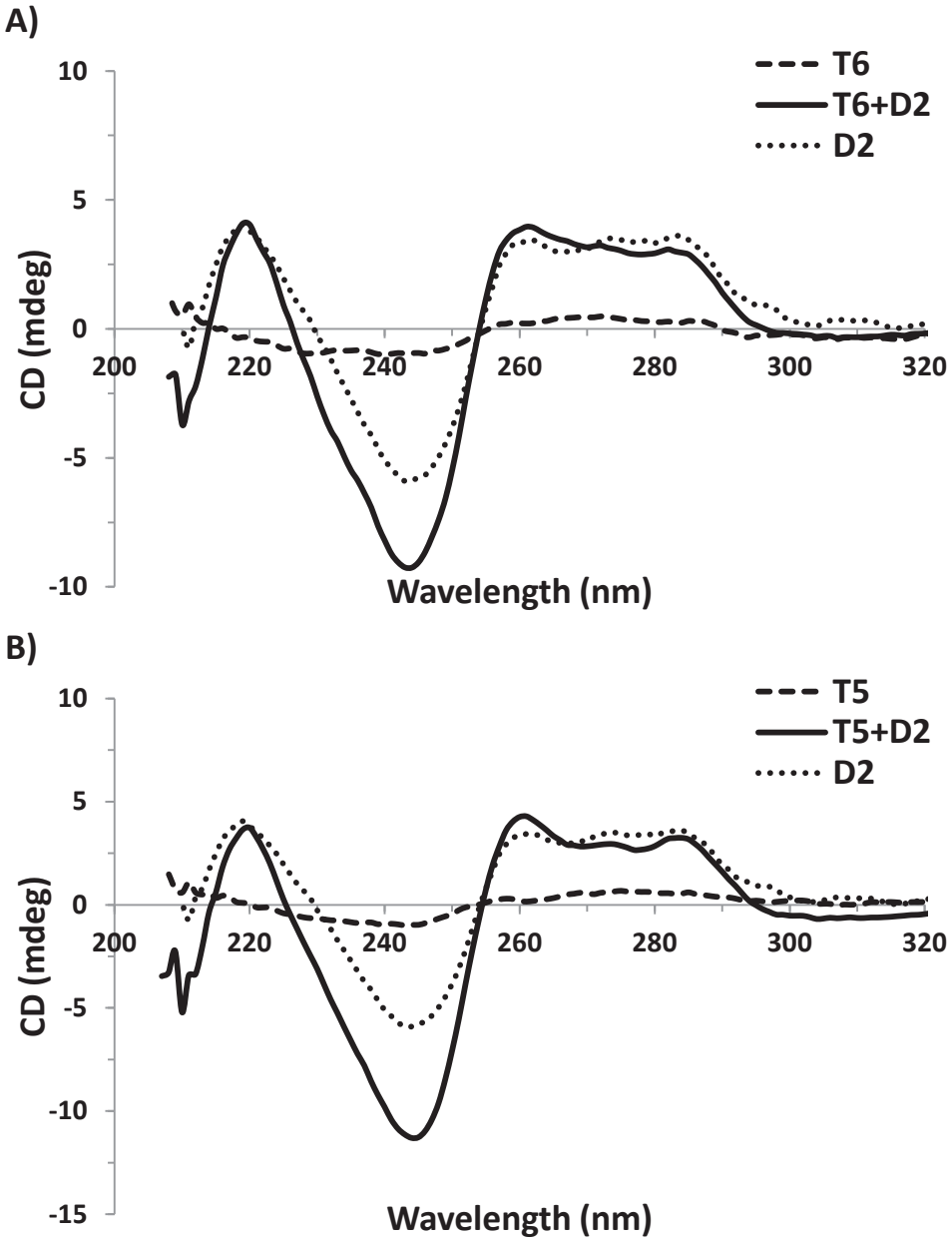


Figure 4 A) CD spectra of TFO **T5**, duplex **D2** and antiparallel triplex **T5+D2** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM. B) CD spectra of TFO **T6**, duplex **D2** and antiparallel triplex **T6+D2** in Li cacodylate buffer (10 mM), NaCl (100 mM) and $MgCl_2$ (10 mM) at pH 7.2, 20 °C. Strand concentration was set to 1.0 μM.

References

Aboul-ela, F., A. I. Murchie, et al. (1994). "Solution structure of a parallel-stranded tetraplex formed by d(TG4T) in the presence of sodium ions by nuclear magnetic resonance spectroscopy." Journal of Molecular Biology **243**(3): 458-471.

Arimondo, P. B., F. Barcelo, et al. (1998). "Triple helix formation by (G,A)-containing oligonucleotides: Asymmetric sequence effect." Biochemistry **37**(47): 16627-16635.

Arimondo, P. B., J. F. Riou, et al. (2000). "Interaction of human DNA topoisomerase I with G-quartet structures." Nucleic Acids Research **28**(24): 4832-4838.

Aubert, Y. and U. Asseline (2004). "Synthesis and hybridization properties of oligonucleotide-perylene conjugates: influence of the conjugation parameters on triplex and duplex stabilities." Organic & Biomolecular Chemistry **2**(23): 3496-3503.

Aubert, Y., L. Perrouault, et al. (2001). "Synthesis and properties of triple helix-forming oligodeoxyribonucleotides containing 7-chloro-7-deaza-2'-deoxyguanosine." Bioorganic & Medicinal Chemistry **9**(6): 1617-1624.

Baase, W. A. and W. C. Johnson, Jr. (1979). "Circular dichroism and DNA secondary structure." Nucleic Acids Research **6**(2): 797-814.

Bacolla, A., A. Jaworski, et al. (2004). "Breakpoints of gross deletions coincide with non-B DNA conformations." Proceedings of the National Academy of Sciences of the United States of America **101**(39): 14162-14167.

Bacolla, A. and R. D. Wells (2004). "Non-B DNA conformations, genomic rearrangements, and human disease." Journal of Biological Chemistry **279**(46): 47411-47414.

Bailey, C. P., J. M. Dagle, et al. (1998). "Cationic oligonucleotides can mediate specific inhibition of gene expression in Xenopus oocytes." Nucleic Acids Research **26**(21): 4860-4867.

Baird, E. E. and P. B. Dervan (1996). "Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids." Journal of the American Chemical Society **118**(26): 6141-6146.

Barton, J. K. (1988). "Recognizing DNA Structures." Chemical & Engineering News **66**(39): 30-42.

Bates, P. J., D. A. Laber, et al. (2009). "Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer." Experimental and Molecular Pathology **86**(3): 151-164.

Bates, P. J., V. M. Macaulay, et al. (1995). "Characteristics of triplex-directed photoadduct formation by psoralen-linked oligodeoxynucleotides." Nucleic Acids Research **23**(21): 4283-4289.

Benfield, A. P., M. C. Macleod, et al. (2008). "Targeted generation of DNA strand breaks using pyrene-conjugated triplex-forming oligonucleotides." Biochemistry **47**(23): 6279-6288.

Berndl, S. and H. A. Wagenknecht (2009). "Fluorescent Color Readout of DNA Hybridization with Thiazole Orange as an Artificial DNA Base." Angewandte Chemie-International Edition **48**(13): 2418-2421.

Bethge, L., D. V. Jarikote, et al. (2008). "New cyanine dyes as base surrogates in PNA: forced intercalation probes (FIT-probes) for homogeneous SNP detection." Bioorganic & Medicinal Chemistry **16**(1): 114-125.

Biton, A., A. Ezra, et al. (2010). "DNA photocleavage by DNA and DNA-LNA amino acid-dye conjugates." Bioconjugate Chemistry **21**(4): 616-621.
Blackburn, G. M. (2006). Nucleic acids in chemistry and biology. Cambridge, UK, RSC Pub.
Bouillon, C., A. Meyer, et al. (2006). "Microwave assisted "click" chemistry for the synthesis of multiple labeled-carbohydrate oligonucleotides on solid support." Journal of Organic Chemistry **71**(12): 4700-4702.
Boutorine, A. S. and C. Escude (2007). "Biophysical analysis of triple-helix formation." Current Protocols in Nucleic Acid Chemistry **Chapter 7**: Unit 7 12.
Boutorine, A. S., V. A. Ryabinin, et al. (2003). "Stabilization of DNA double and triple helices by conjugation of minor groove binders to oligonucleotides." Nucleosides Nucleotides & Nucleic Acids **22**(5-8): 1267-1272.
Bruisten, S. M., P. Reiss, et al. (1998). "Cellular proviral HIV type 1 DNA load persists after long-term RT-inhibitor therapy in HIV type 1 infected persons." AIDS research and human retroviruses **14**(12): 1053-1058.
Brunet, E., P. Alberti, et al. (2005). "Exploring cellular activity of locked nucleic acid-modified triplex-forming oligonucleotides and defining its molecular basis." Journal of Biological Chemistry **280**(20): 20076-20085.
Buchini, S. and C. J. Leumann (2003). "Recent improvements in antigene technology." Current Opinion in Chemical Biology **7**(6): 717-726.
Burge, S., G. N. Parkinson, et al. (2006). "Quadruplex DNA: sequence, topology and structure." Nucleic Acids Research **34**(19): 5402-5415.
Caceres, C., G. Wright, et al. (2004). "A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl+/Na+ ions." Nucleic Acids Research **32**(3): 1097-1102.
Christensen, U. B., M. Wamberg, et al. (2004). "Intercalating nucleic acids: The influence of linker length and intercalator type on their duplex stabilities." Nucleosides Nucleotides & Nucleic Acids **23**(1-2): 207-225.
Clark, D. P. (2005). Molecular biology. Amsterdam ; Boston, Elsevier Academic Press.
Cogoi, S., M. Paramasivam, et al. (2009). "Identification of a new G-quadruplex motif in the KRAS promoter and design of pyrene-modified G4-decoys with antiproliferative activity in pancreatic cancer cells." Journal of Medicinal Chemistry **52**(2): 564-568.
Cogoi, S., M. Paramasivan, et al. (2007). "The effect of INA [(R)-1-O-(1-pyrenylmethyl)glycerol] insertions on the structure and biological activity of a G-quadruplex from a critical KRAS G-rich sequence." Nucleosides Nucleotides & Nucleic Acids **26**(10-12): 1641-1643.
Dagle, J. M. and D. L. Weeks (1996). "Positively charged oligonucleotides overcome potassium-mediated inhibition of triplex DNA formation." Nucleic Acids Research **24**(11): 2143-2149.
Dervan, P. B. and R. W. Burli (1999). "Sequence-specific DNA recognition by polyamides." Current Opinion in Chemical Biology **3**(6): 688-693.
Doluca, O., A. S. Boutorine, et al. (2011). "Triplex-forming twisted intercalating nucleic acids (TINAs): design rules, stabilization of antiparallel DNA triplexes and inhibition of G-quartet-dependent self-association." Chembiochem **12**(15): 2365-2374.
Duca, M., P. Vekhoff, et al. (2008). "The triple helix: 50 years later, the outcome." Nucleic Acids Research **36**(16): 5123-5138.

Durland, R. H., T. S. Rao, et al. (1995). "Azole substituted oligonucleotides promote antiparallel triplex formation at non-homopurine duplex targets." Nucleic Acids Research **23**(4): 647-653.

Englund, E. A., Q. Xu, et al. (2006). "PNA-DNA duplexes, triplexes, and quadruplexes are stabilized with trans-cyclopentane units." Journal of the American Chemical Society **128**(51): 16456-16457.

Escude, C., S. Mohammadi, et al. (1996). "Ligand-induced formation of hoogsteen-paired parallel DNA." Chemistry & Biology **3**(1): 57-65.

Escude, C., C. H. Nguyen, et al. (1998). "Rational design of a triple helix-specific intercalating ligand." Proceedings of the National Academy of Sciences of the United States of America **95**(7): 3591-3596.

Escude, C., J. S. Sun, et al. (1996). "Ligand-induced formation of triple helices with antiparallel third strands containing G and T." Biochemistry **35**(18): 5735-5740.

Esposito, V., A. Virgilio, et al. (2009). "Effects of the introduction of inversion of polarity sites in the quadruplex forming oligonucleotide TGGGT." Bioorganic & Medicinal Chemistry **17**(5): 1997-2001.

Esposito, V., A. Virgilio, et al. (2005). "A new class of DNA quadruplexes formed by oligodeoxyribonucleotides containing a 3 '-3 ' or 5 '-5 ' inversion of polarity site." Chemical Communications(31): 3953-3955.

Faria, M. and C. Giovannangeli (2001). "Triplex-forming molecules: from concepts to applications." Journal of Gene Medicine **3**(4): 299-310.

Fery-Forgues, S. and D. Lavabre (1999). "Are fluorescence quantum yields so tricky to measure? A demonstration using familiar stationery products." Journal of Chemistry Education **76**(9): 1260−1264.

Filichev, V. V., I. V. Astakhova, et al. (2008). "1-, 2-, and 4-Ethynylpyrenes in the structure of twisted Intercalating nucleic acids: structure, thermal stability, and fluorescence relationship." Chemistry **14**(32): 9968-9980.

Filichev, V. V., H. Gaber, et al. (2006). "Twisted intercalating nucleic acids - Intercalator influence on parallel triplex Stabilities." European Journal of Organic Chemistry(17): 3960-3968.

Filichev, V. V. and E. B. Pedersen (2005). "Stable and selective formation of Hoogsteen-type triplexes and duplexes using twisted intercalating nucleic acids (TINA) prepared via postsynthetic sonogashira solid-phase coupling reactions." Journal of the American Chemical Society **127**(42): 14849-14858.

Geci, I., V. V. Filichev, et al. (2006). "Synthesis of twisted intercalating nucleic acids possessing acridine derivatives. Thermal stability studies." Bioconjugate Chemistry **17**(4): 950-957.

Geci, I., V. V. Filichev, et al. (2007). "Stabilization of parallel triplexes by twisted intercalating nucleic acids (TINAs) incorporating 1,2,3-triazole units and prepared by microwave-accelerated click chemistry." Chemistry **13**(22): 6379-6386.

Geierstanger, B. H., J. P. Jacobsen, et al. (1994). "Structural and dynamic characterization of the heterodimeric and homodimeric complexes of distamycin and 1-methylimidazole-2-carboxamide-netropsin bound to the minor groove of DNA." Biochemistry **33**(10): 3055-3062.

Geierstanger, B. H., M. Mrksich, et al. (1994). "Design of a G.C-specific DNA minor groove-binding peptide." Science **266**(5185): 646-650.

Gierlich, J., G. A. Burley, et al. (2006). "Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA." Organic Letters **8**(17): 3639-3642.

Giovannangeli, C., S. Diviacco, et al. (1997). "Accessibility of nuclear DNA to triplex-forming oligonucleotides: The integrated HIV-1 provirus as a target." Proceedings of the National Academy of Sciences of the United States of America **94**(1): 79-84.

Gomez, D., N. Aouali, et al. (2003). "Resistance to the short term antiproliferative activity of the G-quadruplex ligand 12459 is associated with telomerase overexpression and telomere capping alteration." Journal of Biological Chemistry **278**(50): 50554-50562.

Goodchild, J. (1990). "Conjugates of oligonucleotides and modified oligonucleotides: a review of their synthesis and properties." Bioconjugate Chemistry **1**(3): 165-187.

Graham, D., J. A. Parkinson, et al. (1998). "DNA duplexes stabilized by modified monomer residues: Synthesis and stability." Journal of the Chemical Society-Perkin Transactions 1(6): 1131-1138.

Gray, D. M., S. H. Hung, et al. (1995). "Absorption and Circular-Dichroism Spectroscopy of Nucleic-Acid Duplexes and Triplexes." Biochemical Spectroscopy **246**: 19-34.

Grimm, G. N., A. S. Boutorine, et al. (2002). "Formation of DNA triple helices by an oligonucleotide conjugated to a fluorescent ruthenium complex." Chembiochem **3**(4): 324-331.

Gros, J., F. Rosu, et al. (2007). "Guanines are a quartet's best friend: Impact of base substitutions on the kinetics and stability of tetramolecular quadruplexes." Nucleic Acids Research **35**(9): 3064-3075.

Guschlbauer, W., J. F. Chantot, et al. (1990). "Four-stranded nucleic acid structures 25 years later: from guanosine gels to telomer DNA." Journal of Biomolecular Structure & Dynamics **8**(3): 491-511.

Halby, L., V. A. Ryabinin, et al. (2005). "Functionalized head-to-head hairpin polyamides: Synthesis, double-stranded DNA-binding activity and affinity." Bioorganic & Medicinal Chemistry Letters **15**(16): 3720-3724.

Halby, L., V. A. Ryabinin, et al. (2007). "Head-to-head bis-hairpin polyamide minor groove binders and their conjugates with triplex-forming oligonucleotides: Studies of interaction with target double-stranded DNA." Journal of Biomolecular Structure & Dynamics **25**(1): 61-76.

Henderson, E., C. C. Hardin, et al. (1987). "Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs." Cell **51**(6): 899-908.

Hurley, L. H., R. T. Wheelhouse, et al. (2000). "G-quadruplexes as targets for drug design." Pharmacology & Therapeutics **85**(3): 141-158.

Kaur, H., A. Arora, et al. (2006). "Thermodynamic, counterion, and hydration effects for the incorporation of locked nucleic acid nucleotides into DNA duplexes." Biochemistry **45**(23): 7347-7355.

Kele, P., G. Mezo, et al. (2009). "Dual labeling of biomolecules by using click chemistry: a sequential approach." Angewandte Chemie-International Edition **48**(2): 344-347.

Kelly, J. J., E. E. Baird, et al. (1996). "Binding site size limit of the 2:1 pyrrole-imidazole polyamide-DNA motif." Proceedings of the National Academy of Sciences of the United States of America **93**(14): 6981-6985.

Keppler, M. D., S. Neidle, et al. (2001). "Stabilisation of TG- and AG-containing antiparallel DNA triplexes by triplex-binding ligands." Nucleic Acids Research **29**(9): 1935-1942.

Keppler, M. D., M. A. Read, et al. (1999). "Stabilization of DNA triple helices by a series of mono- and disubstituted amidoanthraquinones." European Journal of Biochemistry **263**(3): 817-825.

Kolganova, N. A., A. K. Shchyolkina, et al. (2012). "Targeting duplex DNA with chimeric alpha,beta-triplex-forming oligonucleotides." Nucleic Acids Research **40**(16): 8175-8185.

Krutzik, P. O. and A. R. Chamberlin (2002). "Rapid solid-phase synthesis of DNA-binding pyrrole-imidazole polyamides." Bioorganic & Medicinal Chemistry Letters **12**(16): 2129-2132.

Kutyavin, I., S. Lokhov, et al. (2003). "Chemistry of minor groove binder-oligonucleotide conjugates." Current Protocols in Nucleic Acid Chemistry **Chapter 8**: Unit 8 4.

Kutyavin, I. V., I. A. Afonina, et al. (2000). "3'-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures." Nucleic Acids Research **28**(2): 655-661.

Kypr, J., I. Kejnovska, et al. (2009). "Circular dichroism and conformational polymorphism of DNA." Nucleic Acids Research **37**(6): 1713-1725.

Lakowicz, J. R. (1999). Principles of fluorescence spectroscopy. New York, Kluwer Academic/Plenum.

Lane, A. N., J. B. Chaires, et al. (2008). "Stability and kinetics of G-quadruplex structures." Nucleic Acids Research **36**(17): 5482-5515.

Lavery, R. (2005). "Recognizing DNA." Quarterly Reviews of Biophysics **38**(4): 339-344.

Lee, L. G., C. H. Chen, et al. (1986). "Thiazole orange: a new dye for reticulocyte analysis." Cytometry **7**(6): 508-517.

Macaulay, V. M., P. J. Bates, et al. (1995). "Inhibition of aromatase expression by a psoralen-linked triplex-forming oligonucleotide targeted to a coding sequence." FEBS Letters **372**(2-3): 222-228.

Masiero, S., R. Trotta, et al. (2010). "A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures." Org Biomol Chem **8**(12): 2683-2692.

Membrino, A., S. Cogoi, et al. (2011). "G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy." PLoS One **6**(9).

Menacher, F., M. Rubner, et al. (2008). "Thiazole orange and Cy3: Improvement of fluorescent DNA probes with use of short range electron transfer." Journal of Organic Chemistry **73**(11): 4263-4266.

Mergny, J. L., A. De Cian, et al. (2005). "Kinetics of tetramolecular quadruplexes." Nucleic Acids Research **33**(1): 81-94.

Mergny, J. L., J. Li, et al. (2005). "Thermal difference spectra: a specific signature for nucleic acid structures." Nucleic Acids Research **33**(16).

Merkina, E. E. and K. R. Fox (2005). "Kinetic stability of intermolecular DNA quadruplexes." Biophysical Journal **89**(1): 365-373.

Miller, P. S., G. Bi, et al. (1996). "Triplex formation by a psoralen-conjugated oligodeoxyribonucleotide containing the base analog 8-oxo-adenine." Nucleic Acids Research **24**(4): 730-736.

Miller, P. S., S. A. Kipp, et al. (1999). "A psoralen-conjugated triplex-forming oligodeoxyribonucleotide containing alternating methylphosphonate-phosphodiester linkages: synthesis and interactions with DNA." Bioconjugate Chemistry **10**(4): 572-577.

Milligan, J. F., S. H. Krawczyk, et al. (1993). "An anti-parallel triple helix motif with oligodeoxynucleotides containing 2'-deoxyguanosine and 7-deaza-2'-deoxyxanthosine." Nucleic Acids Research **21**(2): 327-333.

Mrksich, M., M. E. Parks, et al. (1994). "Hairpin peptide potif - A new class of oligopeptides for sequence-specific recognition in the minor-groove of double-helical DNA." Journal of the American Chemical Society **116**(18): 7983-7988.

Mrksich, M., W. S. Wade, et al. (1992). "Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin." Proceedings of the National Academy of Sciences of the United States of America **89**(16): 7586-7590.

Neaves, K. J., J. L. Huppert, et al. (2009). "Direct visualization of G-quadruplexes in DNA using atomic force microscopy." Nucleic Acids Research **37**(18): 6269-6275.

Noonberg, S. B., J. C. Francois, et al. (1995). "Effect of competing self-structure on triplex formation with purine-rich oligodeoxynucleotides containing GA repeats." Nucleic Acids Research **23**(11): 1956-1963.

Noonberg, S. B., J. C. Francois, et al. (1995). "Triplex formation with alpha anomers of purine-rich and pyrimidine-rich oligodeoxynucleotides." Nucleic Acids Research **23**(20): 4042-4049.

Novopashina, D., A. Sinyakov, et al. (2003). "Conjugates of oligo(2'-O-methylribonucleotides) with minor groove binders as new sequence-specific agents recognizing both grooves of double-stranded DNA." Nucleosides, nucleotides & nucleic acids **22**(5-8): 1179-1182.

Ohno, M., T. Fukagawa, et al. (2002). "Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies." Chromosoma **111**(3): 201-213.

Olivas, W. M. and L. J. Maher, 3rd (1995). "Overcoming potassium-mediated triplex inhibition." Nucleic Acids Research **23**(11): 1936-1941.

Orson, F. M., D. W. Thomas, et al. (1991). "Oligonucleotide inhibition of Il2r-alpha messenger-RNA transcription by promoter region collinear triplex formation in lymphocytes." Nucleic Acids Research **19**(12): 3435-3441.

Pack, G. R., L. Wong, et al. (1998). "pK(a) of cytosine on the third strand of triplex DNA: Preliminary Poisson-Boltzmann calculations." International Journal of Quantum Chemistry **70**(6): 1177-1184.

Pagano, B., L. Martino, et al. (2008). "Stability and binding properties of a modified thrombin binding aptamer." Biophysical Journal **94**(2): 562-569.

Palumbo, S. L., R. M. Memmott, et al. (2008). "A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity." Nucleic Acids Research **36**(6): 1755-1769.

Paramasivam, M., S. Cogoi, et al. (2008). "Purine twisted-intercalating nucleic acids: A new class of anti-gene molecules resistant to potassium-induced aggregation." Nucleic Acids Research **36**(10): 3494-3507.

Patel, P. K. and R. V. Hosur (1999). "NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by *Saccharomyces cerevisiae* telomere repeats." Nucleic Acids Research **27**(12): 2457-2464.

Payet, L. and J. L. Huppert (2012). "Stability and structure of long intramolecular G-quadruplexes." Biochemistry **51**(15): 3154-3161.

Pedersen, A. G., L. J. Jensen, et al. (2000). "A DNA structural atlas for Escherichia coli." Journal of Molecular Biology **299**(4): 907-930.

Pedersen, E. B., J. T. Nielsen, et al. (2011). "Enhanced anti-HIV-1 activity of G-quadruplexes comprising locked nucleic acids and intercalating nucleic acids." Nucleic Acids Research **39**(6): 2470-2481.

Perez-Martin, J. and V. de Lorenzo (1997). "Clues and consequences of DNA bending in transcription." Annual Reviews of Microbiology **51**: 593-628.

Petraccone, L., E. Erra, et al. (2003). "Effect of a modified thymine on the structure and stability of [d(TGGGT)]4 quadruplex." International Journal of Biological Macromolecules **31**(4-5): 131-137.

Porres, L., A. Holland, et al. (2006). "Absolute measurements of photoluminescence quantum yields of solutions using an integrating sphere." Journal of Fluorescence **16**(2): 267-272.

Porumb, H., H. Gousset, et al. (1996). "Temporary ex vivo inhibition of the expression of the human oncogene HER2(NEU) by a triple helix-forming oligonucleotide." Cancer Research **56**(3): 515-522.

Rawal, P., V. B. Kummarasetti, et al. (2006). "Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation." Genome Research **16**(5): 644-655.

Raymond, E., J. C. Soria, et al. (2000). "DNA G-quadruplexes, telomere-specific proteins and telomere-associated enzymes as potential targets for new anticancer drugs." Investigational New Drugs **18**(2): 123-137.

Reddy, M. R. and M. Berkowitz (1989). "Hydration forces between parallel DNA double helices: computer simulations." Proceedings of the National Academy of Sciences of the United States of America **86**(9): 3165-3168.

Reese, C. B. (2005). "Oligo- and poly-nucleotides: 50 years of chemical synthesis." Organic & Biomolecular Chemistry **3**(21): 3851-3868.

Renard, B. L., R. Lartia, et al. (2008). "Targeting DNA with "light-up" pyrimidine triple-helical forming oligonucleotides conjugated to stabilizing fluorophores (LU-TFOs)." Organic & Biomolecular Chemistry **6**(23): 4413-4425.

Rohrbach, F., M. I. Fatthalla, et al. (2012). "Chemical maturation of a bivalent aptamer by single domain variation." Chembiochem **13**(5): 631-634.

Rosenthal, N. (1995). "Molecular medicine - Recognizing DNA." New England Journal of Medicine **333**(14): 925-927.

Ryabinin, V. A., A. S. Boutorine, et al. (2004). "Oligonucleotide-minor groove binder conjugates and their complexes with complementary DNA: Effect of conjugate structural factors on the thermal stability of duplexes." Nucleosides Nucleotides & Nucleic Acids **23**(5): 789-803.

Ryabinin, V. A., A. Y. Denisov, et al. (1999). "Stabilization of DNA-duplexes by minor groove binders with the formation of the 1 : 2 type complexes and parallel orientation of ligands in minor groove." Doklady Akademii Nauk **368**(6): 836-838.

Schwarz-Finsterle, J., S. Stein, et al. (2007). "Comparison of triple helical COMBO-FISH and standard FISH by means of quantitative microscopic image analysis of abl/bcr positions in cell nuclei." Journal of Biochemical and Biophysical Methods **70**(3): 397-406.

Simon, P., F. Cannata, et al. (2008). "Targeting DNA with triplex-forming oligonucleotides to modify gene sequence." Biochimie **90**(8): 1109-1116.

Simonsson, T. (2001). "G-quadruplex DNA structures - Variations on a theme." Biological Chemistry **382**(4): 621-628.
Simonsson, T., P. Pecinka, et al. (1998). "DNA tetraplex formation in the control region of c-myc." Nucleic Acids Research **26**(5): 1167-1172.
Sinyakov, A. N., V. A. Ryabinin, et al. (2001). "Stabilization of DNA triple helices using conjugates of oligonucleotides and synthetic ligands." Molecular Biology **35**(2): 251-260.
Skolakova, P., K. Bednarova, et al. (2010). "Quadruplexes of human telomere dG(3)(TTAG(3))(3) sequences containing guanine abasic sites." Biochemical and Biophysical Research Communications **399**(2): 203-208.
Son, T. D., W. Guschlbauer, et al. (1972). "Flexibility and conformations of guanosine monophosphates by the Overhauser effect." Journal of the American Chemical Society **94**(22): 7903-7911.
Song, J., Z. Intody, et al. (2004). "Activation of gene expression by triplex-directed psoralen crosslinks." Gene **324**: 183-190.
Stephenson, A. W. I., A. C. Partridge, et al. (2011). "Synthesis of β-pyrrolic modified porphyrins and their incorporation into DNA." Chemistry-a European Journal **17**(22): 6227-6238.
Sun, J. S., J. C. Francois, et al. (1989). "Sequence-specific intercalating agents: intercalation at specific sequences on duplex DNA via major groove recognition by oligonucleotide-intercalator conjugates." Proc Natl Acad Sci U S A **86**(23): 9198-9202.
Sun, J. S., T. Garestier, et al. (1996). "Oligonucleotide directed triple helix formation." Current Opinion in Structural Biology **6**(3): 327-333.
Sundaralingam, M. and E. Westhof (1981). "Structural motifs of the nucleotidyl unit and the handedness of polynucleotide helices." International Journal of Quantum Chemistry: 287-306.
Szewczyk, J. W., E. E. Baird, et al. (1996). "Sequence-specific recognition of DNA by a major and minor groove binding ligand." Angewandte Chemie-International Edition in English **35**(13-14): 1487-1489.
Thomas, T. J., C. A. Faaland, et al. (1995). "Suppression of c-myc oncogene expression by a polyamine-complexed triplex forming oligonucleotide in mcf-7 breast-cancer cells." Nucleic Acids Research **23**(17): 3594-3599.
Tonelli, R., S. Purgato, et al. (2005). "Antigene peptide nucleic acid specifically inhibits MYCN expression in human neuroblastoma cells leading to cell growth inhibition and apoptosis." Molecular Cancer Therapeutics **4**(5): 779-786.
Trauger, J. W., E. E. Baird, et al. (1998). "Recognition of 16 base pairs in the minor groove of DNA by a pyrrole-imidazole polyamide dimer." Journal of the American Chemical Society **120**(14): 3534-3535.
Vasquez, K. M. (2010). "Targeting and processing of site-specific DNA interstrand crosslinks." Environ Mol Mutagen **51**(6): 527-539.
Vasquez, K. M. and P. M. Glazer (2002). "Triplex-forming oligonucleotides: Principles and applications." Quarterly Reviews of Biophysics **35**(1): 89-107.
Wade, W. S., M. Mrksich, et al. (1992). "Design of peptides that bind in the minor groove of DNA at 5'-(a,T)G(a,T)C(a,T)-3' sequences by a dimeric side-by-side motif." Journal of the American Chemical Society **114**(23): 8783-8794.
Wade, W. S., M. Mrksich, et al. (1993). "Binding affinities of synthetic peptides, pyridine-2-carboxamidonetropsin and 1-methylimidazole-2-carboxamidonetropsin, that form 2:1

complexes in the minor groove of double-helical DNA." Biochemistry **32**(42): 11385-11389.
Wang, Y. and D. J. Patel (1992). "Guanine residues in d(T2AG3) and d(T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution." Biochemistry **31**(35): 8112-8119.
Weisbrod, S. H. and A. Marx (2008). "Novel strategies for the site-specific covalent labelling of nucleic acids." Chemical Communications **44**(44): 5675-5685.
White, S., E. E. Baird, et al. (1997). "On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides." Chemistry & Biology **4**(8): 569-578.
Wittung, P., S. K. Kim, et al. (1994). "Interactions of DNA binding ligands with PNA-DNA hybrids." Nucleic Acids Research **22**(24): 5371-5377.
Wittung, P., P. E. Nielsen, et al. (1994). "DNA-like double helix formed by peptide nucleic acid." Nature **368**(6471): 561-563.
Wu, Q., S. S. Gaddis, et al. (2007). "High-affinity triplex-forming oligonucleotide target sequences in mammalian genomes." Molecular Carcinogenesis **46**(1): 15-23.
Wyatt, J. R., P. W. Davis, et al. (1996). "Kinetics of G-quartet-mediated tetramer formation." Biochemistry **35**(24): 8002-8008.
Xodo, L. E., G. Manzini, et al. (1990). "Spectroscopic and calorimetric investigation on the DNA triplex formed by d(CTCTTCTTTCTTTTCTTTCTTCTC) and d(GAGAAGAAAGA) at acidic pH." Nucleic Acids Research **18**(12): 3557-3564.
Yakovchuk, P., E. Protozanova, et al. (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix (vol 34, pg 564, 2006)." Nucleic Acids Research **34**(3): 1082-+.