UNIVERSITY OF SOUTHAMPTON

Studies on the Stability of Multistranded DNA

by **Nicholas Matthew Brown**

A Thesis Submitted for the Degree of
Doctor of Philosophy.

Division of Biochemistry & Molecular Biology,
Faculty of Medicine, Health & Life Sciences.
December 2003.

University of Southampton
Abstract
Faculty of Medicine, Health and Life Sciences
School of Biological Sciences, Division of Biochemistry & Molecular Biology
Doctor of Philosophy

Studies on the Stability of Multistranded DNA

by Nicholas Matthew Brown

G-quadruplexes are unusual four-stranded DNA structures that contain stacks of guanine quartets. Many guanine-rich sequences with the potential to adopt this structure are found in eukaryotic genomes, and have been shown to form quadruplexes in vitro. Such sequences are found in telomeres and in the promoters of several genes. These structures are particularly interesting, as quadruplex formation inhibits the enzyme telomerase in vitro. These structures have also been implicated in the regulation of the c-myc oncogene. Since telomerase is activated in many cancers it is a potential target for developing novel anticancer agents. Ligands which facilitate the formation of quadruplexes are therefore potentially important therapeutic molecules. The work described in this thesis examines the stability of G-quadruplexes in a range of different conditions and studies their interaction with a series of small molecule ligands.

A novel technique, using fluorescently-labelled oligonucleotides, has been developed to measure the thermal stability of quadruplex-forming DNA sequences. This has been used to compare the stability of human $(TTAGGG)_n$ and Oxytricha $(G_4T_4)_n$ telomeric repeats and to examine their kinetic properties. These sequences are stabilised more by potassium than sodium and the Oxytricha sequence is the most stable. Both sequences produce hysteresis between the melting and annealing profiles, especially at higher rates of temperature change, indicative of slow dynamics of folding and unfolding. These dynamics were studied by temperature-jump experiments, which confirmed that the kinetics of these telomeric repeats is very slow. Both sequences show slower dynamics in potassium than sodium, though the kinetic pathway is complicated, requiring two exponentials for its complete description. These kinetics are much slower for the Oxytricha repeat, which displays a $t_{½}$ of about 30 minutes even at the melting temperature.

The interaction of a series of substituted anthraquinones with quadruplex DNA was studied by fluorescence melting experiment and bandshift analysis. Structure-activity relationships for these compounds were derived by comparing closely related ligands. These studies showed that the propionamide sidechains are an important determinant of the activity and showed that (i) 2,6- substitution of the anthraquinone ring is better than 2,7- substitution; (ii) attachment via the nitrogen rather than carbon of the amide of the propionamide sidechains is often superior; (iii) longer sidechains are better especially if they contain positive charges and (iv) bulky groups in the sidechains are generally unfavourable.

Triplexes are formed by the sequence-specific interaction of synthetic oligonucleotides within the major groove of duplex DNA. Oligonucleotide-directed triple helix formation can be exploited for triplex-mediated transcriptional inhibition (antigene approach). Triplex stability is known to be influenced by the precise sequence composition and arrangement. The final part of this thesis describes the (unsuccessful) development of a novel SELEX methodology (REPSA) for identifying strong triplex binding sites from a pool of potential target sites.

List of Contents

List of Figures

Chapter 1

Chapter 3

Chapter 6

List of Tables

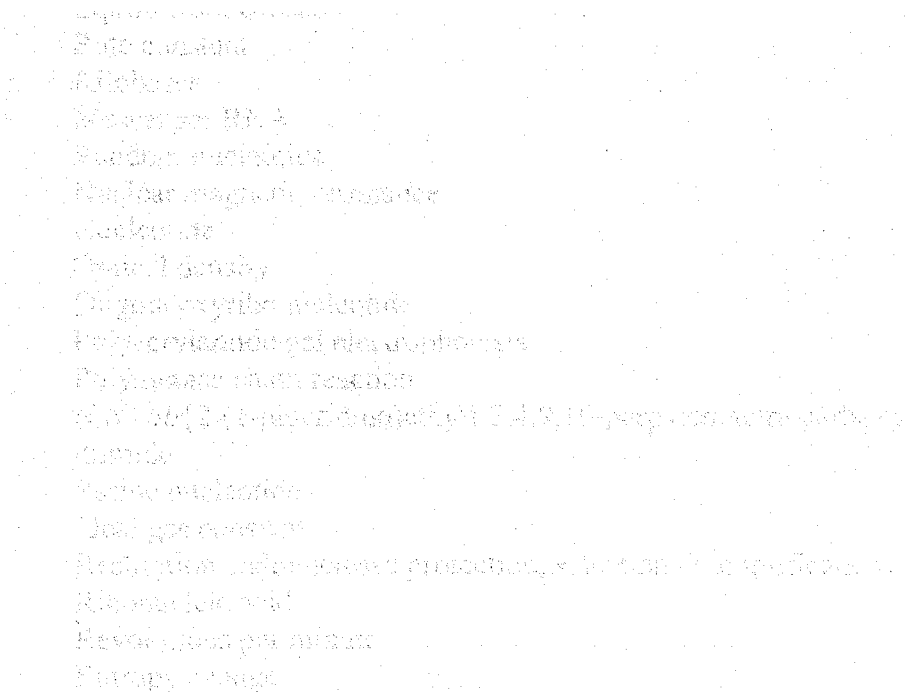
Chapter 5

Acknowledgements

I would like to thank Professor Keith Fox for his help and guidance, and the BBSRC for funding my studentship.

I would also like to thank all my friends and colleagues throughout the Fox group, Biochemistry Dept and the School of Biological Sciences, past and present. I'm thinking particularly of: Dave, Pete and Paul for all the laughs, at lunch, coffee breaks ('coffee at 3, anyone???') and down the pub; and Owain and Matt from Neuroscience for their friendship.

Finally, a very sincere thank you to Steve, Rob, Sarah and Michelle, for your help and many favours, of which: 'I've just lost count…'.



Abbreviations

A	Adenosine
ACI	Acridine
ACO	Acridone
AQ	Anthraquinone
bp	Base pairs
C	Cytidine
CD	Circular dichroism
cps	Counts per second
dATP	Deoxyadenosine triphosphate
dF/dt	Rate of fluorescence change
dT/dt	Rate of temperature change
DNA	Deoxyribonucleic acid
DNaseI	Deoxyribonuclease I
DODC	3,3'-diethyl-oxyadicarbocyanine
ds	Double stranded
DTC	*N,N'*-diethyl-thiocarbocyanine
EDTA	Ethylediaminetetraacetic acid
F	Fluorescence
ΔG	Gibbs free energy change
G	Guanosine
G4	Quadruplex
ΔH	Enthalpy change
IPTG	Isopropyl –D-thiogalactopyranoside
K	Guanosine or thymidine
K	Equilibrium constant
k	Rate constant
kb	Kilobases
mRNA	Messenger RNA
N	Random nucleotide
NMR	Nuclear magnetic resonance
nt	Nucleotide
OD	Optical density
ODN	Oligodeoxyribonucleotide
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PIPER	*N,N'-bis*[2-(1-piperidino)ethyl]-3,4,9,10-perpylenetetra-carboxylic diimide
R	Purine nucleotide
R	Ideal gas constant
REPSA	Restriction endonuclease protection, selection & amplification
RNA	Ribonucleic acid
rpm	Revolutions per minute
ΔS	Entropy change
SAR	Structure activity relationship
ss	Single stranded
$1/\tau$	Relaxation time
T	Thymidine
TBE	Tris-borate-EDTA buffer

TE	Tris-EDTA buffer
Tel01	*N,N'-bis*[3-(4-morpholino)propyl]-3,4,9,10-perpylenetetra-carboxylic diimide
TEMED	*N',N',N',N'*-tetramethylethylenediamine
TFO	Triplex forming oligonucleotide
T_m	Temperature midpoint / melting temperature
TMPyP	Meso-tetra(n-methyl-4-pyridyl) porphine
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
U	Uridine
UV	Ultra-violet (radiation)
X-gal	5-bromo-4-chloro-3-indolyl β-galactosidase
Y	Pyrimidine nucleotide

Note regarding quadruplex nomenclature: in this thesis, $[d(T_4G_4)]_4$ specifies a quadruplex composed of four strands of d(TTTTGGGG). [ ] denotes quadruplexes, whereas the subscript outside these brackets indicates the number of strands.

Note regarding triplets: the third strand oligonucleotide is represented by the first letter, a midline dot (·) refers to the Hoogsteen hydrogen bonds, while the second and third letters represent the base pair of the duplex.

1 Introduction

1.1 DNA structure

Watson and Crick proposed the structure for DNA in 1953, and this has successfully explained many of its biological properties as the carrier of genetic information. The details of the structure have since been refined, but the overall architecture has remained the same. The two strands are arranged antiparallel to each other, with the hydrophobic nucleobases in the centre (with A paired with T, and G with C) and the charged phosphates on the outside. This is not a rigid structure and the precise structural details vary in a sequence-dependent fashion. Three have been described termed A-, B- and Z-DNA, though B-DNA is assumed to be the biologically relevant form.

B-DNA is a right-handed helix, which arises from *anti*-nucleotides with $C^{2'}$-endo pucker. A-DNA is also a right-handed duplex formed by dehydrating B-DNA. It is comprised of nucleotides in the *anti*-configuration and $C^{3'}$-endo sugar pucker. Z-DNA is a left-handed helix consisting of zig-zag dinucleotide repeats, with an alternating *anti/syn* backbone formed by G/C rich sequences at high salt. Models of each duplex form are shown in Figure 1.1 and the dimensions given in Table 1.1. It is also clear that DNA can adopt a variety of three- and four-stranded forms, which may have biological or therapeutic roles. These form the basis of the work described in this thesis.

1.2 G-quadruplexes

A G-quadruplex is a four-stranded structure comprised of stacked G-quartets (or tetrads). In a G-quartet, four guanines associate via Hoogsteen hydrogen-bonding to form a planar ring (as shown in Figure 1.2A). This model was first proposed by Gellert *et al.* (1962). It explained the unusual property seen during UV melting experiments on synthetic poly(G) oligonucleotides, which implied the existence of strong self-associated structures.

These structures have since generated renewed interest. While definitive evidence for the existence *in vivo* of G-quadruplexes is still outstanding (remaining a major obstacle), much indirect evidence supports their likely existence. Runs of guanines exist in the eukaryotic genome, which should be capable of forming G-quadruplexes similar to the behaviour of identical synthetic sequences under physiological pH and cation concentrations.

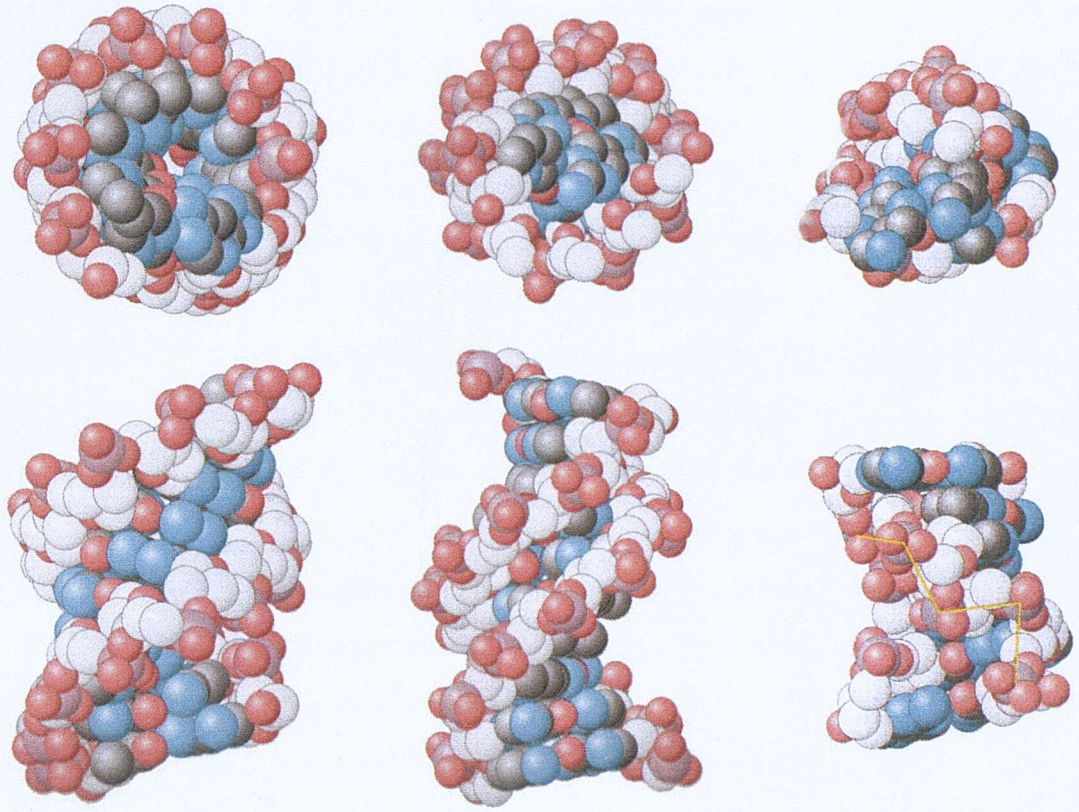


Figure 1.1. Space-filling models of idealised forms of A-, B-, and Z-DNA. Reproduced from Berg J.M., Tymoczko J.L., Stryer L. Biochemistry 5th edition.

Table 1.1. Comparison of A-, B-, and Z-DNA.

	Helix type		
	A	B	Z
Shape	Broadest	Intermediate	Narrowest
Rise per base pair	2.9 Å	3.3 Å	3.7 Å
Helix diameter	25.5 Å	23.7 Å	18.4 Å
Screw sense	Right-handed	Right-handed	Left-Handed
Glycosidic bond	*anti*	*anti*	*anti* for C,T *syn* for G
Base pairs per turn	11	10.4	12
Pitch per turn of helix	25.3 Å	35.4 Å	45.6 Å
Tilt of base pairs from normal to helix axis	19°	1°	9°
Major groove	2.7 Å wide 13.5 Å deep	11.7 Å wide 8.8 Å deep	8.8 Å wide 3.7 Å deep
Minor groove	11 Å wide 2.8 Å deep	5.7 Å wide 7.5 Å deep	2.0 Å wide 13.8 Å deep

Such sequences exist at the end of chromosomes in the telomeric region (as repeating units), and were first demonstrated to form such structures by Henderson *et al.* (1987). This was soon extended to include other G-rich sequences within transcription regulatory regions, such as within the Ig switch (Sen & Gilbert, 1988), the *c-myc* oncogene (Simonsson *et al.*, 1998), the HIV genome (Sundquist & Heaphy, 1993), the insulin gene promoter (Hammond-Kosack *et al.*, 1992), and the retinoblastoma susceptibility gene (Murchie & Lilley, 1992). Such abundance implies several possible biological roles for quadruplex, the most significant being in transcriptional regulation. Proteins have also been identified *in vivo* which interfere with quadruplex formation and could act as molecular chaperones to participate in the biological events that involve these structures (see section 1.7). Most compelling are two recent reports, which support the presence *in vivo* of quadruplexes in *Stylonychia lemnae* nuclei (Schaffitzel *et al.*, 2001), and demonstrate that a quadruplex in a promoter region can repress c-myc transcription *in vivo*, and that a G4 ligand can influence this repression (Siddiqui-Jain *et al.*, 2002).

However, several obstacles complicate the idea that quadruplexes exist *in vivo*. Firstly, G-rich regions form stable Watson-Crick base pairs throughout the cell cycle and quadruplex formation will usually have to compete with duplex DNA. Secondly, specific and non-specific proteins interact with genomic dsDNA which may either hinder or facilitate quadruplex formation. Thirdly, intermolecular quadruplexes require relatively high DNA concentrations *in vitro* and these may not occur *in vivo*. Fourthly, although quadruplexes can have high stability, they have exceptionally slow kinetics of formation and dissociation.

In 1991, it was demonstrated that G-quadruplex structures in *Oxytricha* telomeres were able to inhibit telomerase (Zahler *et al.*, 1991). This key observation suggested that these structures might represent novel molecular targets for interactive agents, should they be biologically significant, and initiated renewed interest in their properties. One objective therefore, is to devise small molecules that can recognise and stabilise these structures, and this is one of the focuses of this thesis.

1.3 Quadruplex structural polymorphism

Quadruplex structures have been examined using NMR, X-ray crystallography, circular dichroism (CD), chemical and nuclease protection, and electrophoresis. These studies have shown that the sequences including human $d(G_3T_2A)_n$, *Tetrahymena* $d(G_4T_2)_n$ and *Oxytricha* $d(G_4T_4)_n$ telomeric repeats form highly diverse structures based on the same hydrogen bonding scheme, with different strand stoichiometry, strand orientation, loop orientation and glycosyl conformation (see Figure 1.2). The structural variants are described below.

1.3.1 Tetrameric quadruplex

Sequences with single G-tracts must form intermolecular tetrameric quadruplexes consisting of four strands each contributing at least a single G-repeat (Sen & Gilbert, 1988; Wang & Patel, 1993a; Laughlan *et al.*, 1994; Phillips *et al.*, 1997). This topology consists of parallel strands, and all the guanosines are the *anti* glycosyl conformers as in the B-form of dsDNA. This produces a complex with four equally sized grooves, each of which is roughly equal to the minor groove of B-DNA. See Figure 1.2B.

1.3.2 Dimeric quadruplex

DNA possessing two G-rich repeats can form intermolecular dimeric quadruplexes. An antiparallel organisation results from two G-repeats running opposite to the other two. The glycosidic conformation alternates *syn-anti* along the strands, similar to Z-DNA (Kang *et al.*, 1992; Smith & Feigon, 1992; Schultze *et al.*, 1999; Haider *et al.*, 2002). Loops may be lateral to produce so-called *chair-* or *edge loop* quadruplexes; or diagonal in structures termed *basket-* or *diagonal loop-* quadrupelxes. Lateral loops (see Figure 1.2C and D) may be head-to-head or head-to-tail, and have the G(*syn*)·G(*anti*)·G(*syn*)·G(*anti*) quartet arrangement, which produces two narrow and two wide grooves. Diagonal loops (see Figure 1.2E) result in G(*syn*)·G(*syn*)·G(*anti*)·G(*anti*) quartets, resulting in one narrow, two medium and one wide groove. The association of two G·G hairpins may form diagonal-looped quadruplexes, whereas within diagonal looped structures each guanine is hydrogen bonded only to those of the other strand and so these structures clearly cannot form in this way.

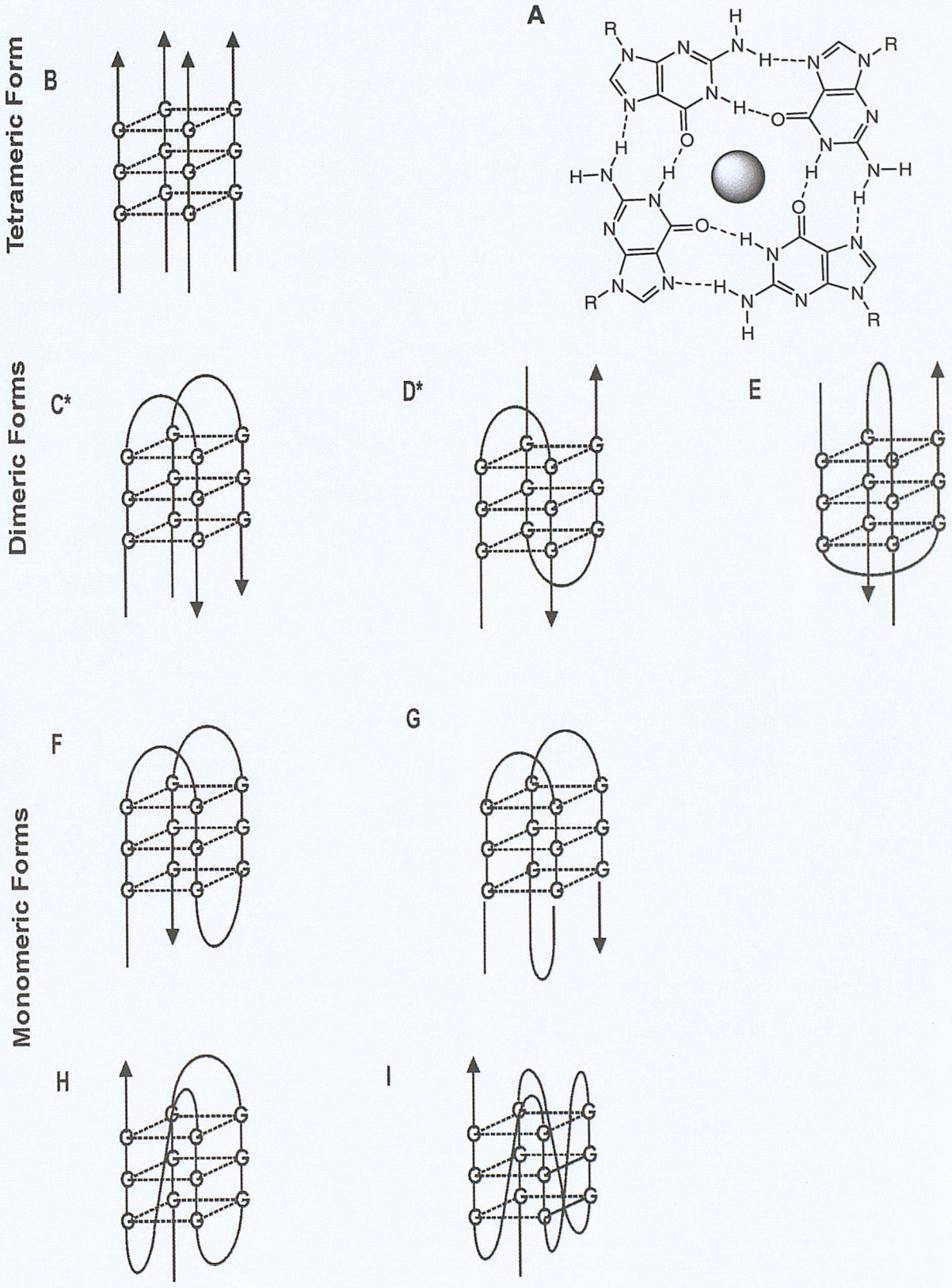


Figure 1.2. (A) Structure of a G-quartet which gives rise to four-stranded nucleic acids, structural variants of which are labelled (B to I). *For simplicity only one of the two possible hairpin orientations is shown.

1.3.3 Monomeric quadruplex

Sequences containing four G-tracts can fold into intramolecular monomeric quadruplexes. These structures have antiparallel strand polarities and the same patterns of glycosidic torsional angles both along the strand and within the G-quartets (depending on the position of the loops) as the dimeric quadruplexes (compare structures in Figure 1.1 C to F and D to G). Structure G has its central loop diagonally across the quartets and two lateral loops giving the same patterns of glycosidic configuration and groove width to the diagonal dimer (Smith *et al.*, 1995; Wang & Patel, 1993b; 1995). However, an anomalous structure was determined which contained three parallel strands to one antiparallel (Wang & Patel, 1994). Also, two high resolution crystal structures were recently described which have a novel architecture (Parkinson *et al.*, 2002), the dimeric and monomeric quadruplexes have loops are 'bridged' rather than flanking giving rise to parallel stranded structures with all *anti*-glycosyl nucleotides (see Figure 1.2 H and I, and Figure 1.3).

1.4 Quadruplex stability & kinetics

Quadruplex formation is enthalpy driven, ΔH per tetrad have been measured as – 21 to –26 kcal/mol (Jin *et al.*, 1992). A variety of factors influence the structural preference and stability of a quadruplex structure including temperature, strand concentration, salt concentration, and flanking and loop sequence. Quadruplexes are kinetically slow to form under physiological conditions. The imino protons in G-quartet structures exchange with solvent protons on a very slow time-scale (Deng & Braulin, 1995; Hardin *et al.*, 1993), illustrating the high kinetic stability exceptional among nucleic acid structures. Association rate constants of $0.0017 s^{-1}$ and $0.023 s^{-1}$ in 50 mM Na^+ and K^+ respectively, were deduced at 37°C from bandshifts for $d(T_4G_4T_4G_4T_4G_4T_4G_4)$ (Raghuraman & Cech, 1990). An association rate constant of 5.18 $M^{-1}s^{-1}$ at room temperature for $d(T_4G_4T_4G_4)$ dimerisation was also measured using bandshift analysis (Guo *et al.*, 1993). A similar value of 10^2 $M^{-1}h^{-1}$ ($\approx 2.78 \times 10^{-2}$ $M^{-1}s^{-1}$) for a 24-mer with two human repeats in 100 mM KCl at room temperature was reported again using bandshift experiments (Han *et al.*, 1999b). Sen and Gilbert (1990) deduced by electrophoresis the association rate constant of 4.5×10^{-4} $M^{-1}s^{-1}$ at 60°C in IM NaCl for both a 39-mer and a 49-mer containing two *Tetrahymena* repeats $d(T_2G_4T_2G_4)$ or two *Oxytricha* repeats $d(T_4G_4T_4G_4)$ at 3'respectively. Tetrameric formations have been reported ranging from second to fourth order. The phosphothioate oligonucleotide $d(T_2G_4T_2)$ was found by size exclusion chromatography to have a fourth

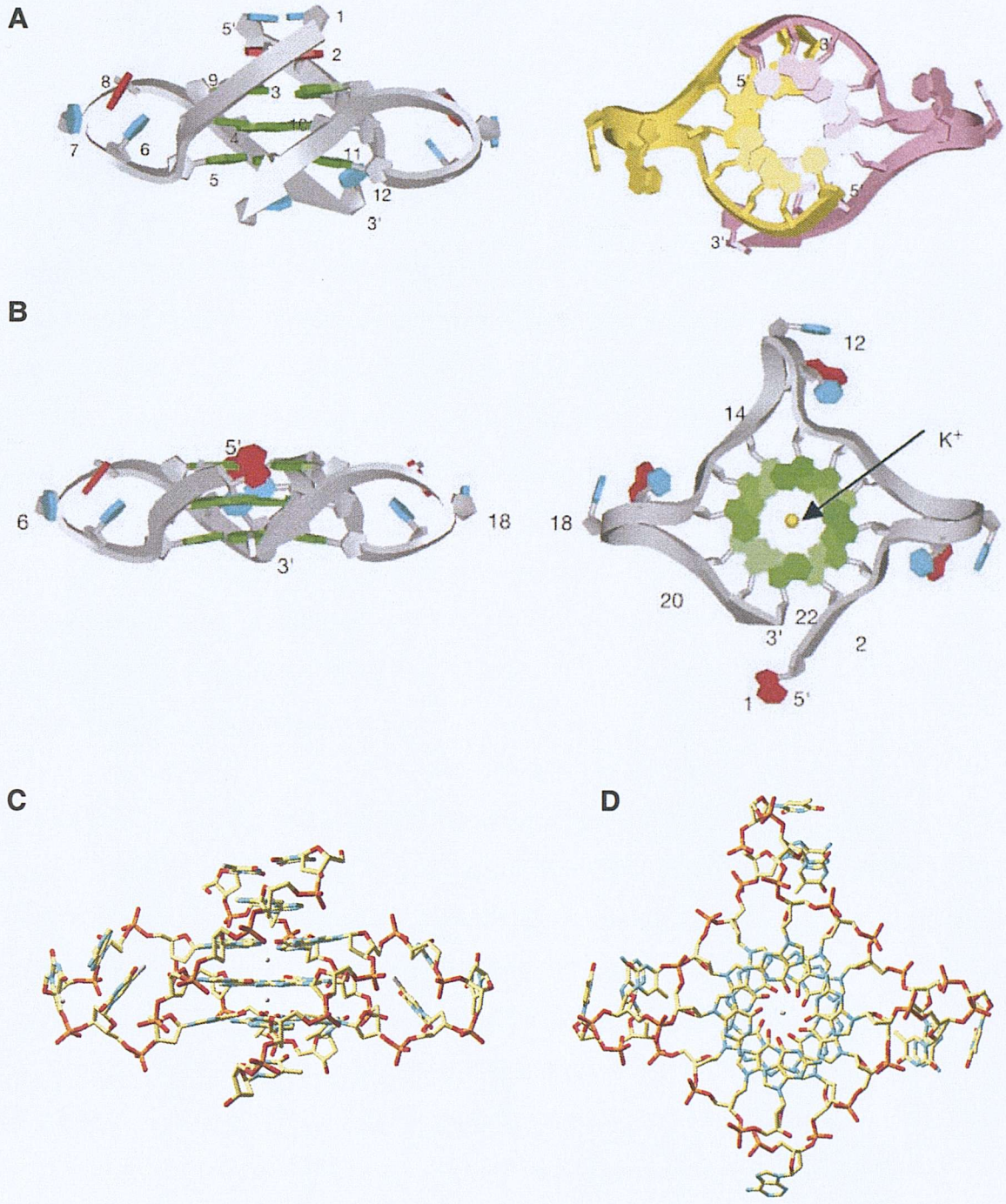


Figure 1.3. The overall folding topology of (A) two 12-mers, and (B) a 22-mer forming inter- and intramolecular quadruplexes, respectively. Side views and end views are shown of the quadruplexes with the phosphate sugar backbone drawn as a ribbon showing 5'-3' directionality. The guanines are in green, thymines are blue, and adenines are red. The two strands of the intermolecular quadruplex are yellow and magenta. Stick representations coloured by atom type show (C) the side view of the intermolecular quadruplex, and (D) the end view of the intramolecular quadruplex and are shown. Figures adapted from Parkinson *et al.* (2002).

order association rate constant of 6.1 x 10^4 $M^{-3}s^{-1}$, in phosphate buffered saline, and an order of magnitude faster than the phosphodiester (Wyatt *et al.*, 1996).

1.5 Cation interactions

1.5.1 M^+ cations

The formation of stacked G-quartets generates an internal pocket of eight guanine carbonyl oxygen atoms (GO6). Internal cation coordination within the GO6 core is essential to drive nucleobase alignment and allow stable quadruplex formation. Crystal structures of K^+-forms of several different quadruplexes all describe K^+ ions as lying between G-tetrads rather than coplanar (Kang *et al.*, 1992; Schultze *et al.*, 1999; Parkinson *et al.*, 2002; Haider *et al.*, 2002). By contrast, solution structures of Na^+-quadruplexes have found the Na^+ ions to lie within each quartet plane (Laughlan *et al.*, 1994; Phillips *et al.*, 1997; Schultze *et al.*, 1994). Presumably, an ionic radii of 1.33 Å for K^+ prevents the ions from chelation within the plane of the G-quartets but allows these cations to more comfortably reside within the GO6 pocket.

The monovalent cations produce differing quadruplex stability (which is unusual among nucleic acid structure). The overall trend for their stabilisation appears to be K^+>Rb^+>Na^+>Cs^+>Li^+. Although the T_m of the quadruplex $[d(G_3T_2AG_3)]_2$ in 70 mM K^+ or Na^+ containing buffer were 42°C compared to 31°C respectively, K^+ induced quadruplex was only marginally more stable by a ΔG of 1.1 kcal/mol than its Na^+-induced structure (Balagurumoorthy & Brahmachari, 1994). This order is thought to reflect the decreasing ionic radius, allowing a more optimal fit within the finite GO6 ion binding pocket. This factor is balanced against overcoming the increasing hydration energies with decreasing ionic radii from 70 to 120 kcal/mol. It was previously assumed that the fit was more significant, however, it has been suggested that the $\Delta\Delta G°$ associated with the conversion from the Na^+ to K^+ form of $[d(G_3T_4G_3)]_2$ of –1.7 kcal/mol reflects the energetic penalty of Na^+ and K^+ dehydration more than the relative $\Delta G°$ of Na^+ and K^+ binding (Hud *et al.*, 1996).

The differing interactions of K^+ and Na^+ result in different structures being induced in similar conditions. K^+ or Na^+ are able to stabilise tetrameric and diagonal loop dimeric quadruplexes, while K^+ alone appeared to stabilise edge-loop type quadruplex (Marathias

& Bolton, 1999). It was suggested that K^+ binding in basket-types was between adjacent quartets and that it interacts within loops.

Sen & Gilbert (1990) suggested a cationic switch mechanism to explain the large influence of different ratios of Na^+ and K^+ on the quadruplex topology formed by a particular sequence (Figure 1.4). Tetrameric quadruplex can form via a G-G duplex, this model proposes that this can also proceed via intermediates (a G-G hairpin and a dimeric intermediate) which are promoted by K^+. At low K^+ / Na^+ concentrations, K^+ increases Na^+ induced tetramer formation. High K^+ concentrations decrease tetramer formation by excessively stabilising and trapping dimeric foldback quadruplexes (dimer 2 and dimer 3).

1.5.2 M^{2+} cations

Divalent cations have binding affinities up to 100-fold that of monovalent cations, due to improved electrostatic interactions, and can induce quadruplex formation at low millimolar levels (Sen & Gilbert, 1991). The relative stabilisation afforded by divalent cations follows the order: $Sr^{2+}>Ba^{2+}>Ca^{2+}>Mg^{2+}$, mirroring those of their ionic radius counterparts (Hardin *et al.*, 1992; Hardin *et al.*, 1993; Venczel & Sen, 1993). Studies have shown quadruplex is increasingly stable when the concentration of M^{2+} is below 10 mM (or 10-40 mM) and the concentration of M^+ is above 50 mM (not below) (Hardin *et al.*, 1992; Hardin *et al.*, 1993). Though low divalent concentrations stabilise quadruplexes, evidence suggests that higher concentrations actually antagonise quadruplex induction by monovalent cations (Lee, 1990; Blume *et al.*, 1997). At concentrations greater than ~10 mM they may induce dissociation by binding to hydrogen-bond acceptors in guanines. The underlining mechanism suggested to explain this, is that divalent cations disrupt the quadruplex strands by entering transiently opened G-quartets and binding to the GN7 and GO6 ligands, thus preventing reformation (Hardin *et al.*, 2001). Elevated monovalent concentrations alleviate this by preventing the transient opening and also compete for ligands. Interestingly, Ca^{2+} has been suggested to induce a transition from antiparallel quadruplex to high molecular-weight parallel quadruplex (Miyoshi *et al.*, 2003).

1.6 Sequence effects upon quadruplex formation

Residues which flank G-quartets can have a major influence on structural stability. Guo *et al.* (1993) studied a series of $d(T_nG_4)$ and $d(T_nG_4T_nG_4)$ sequences, where n was varied from

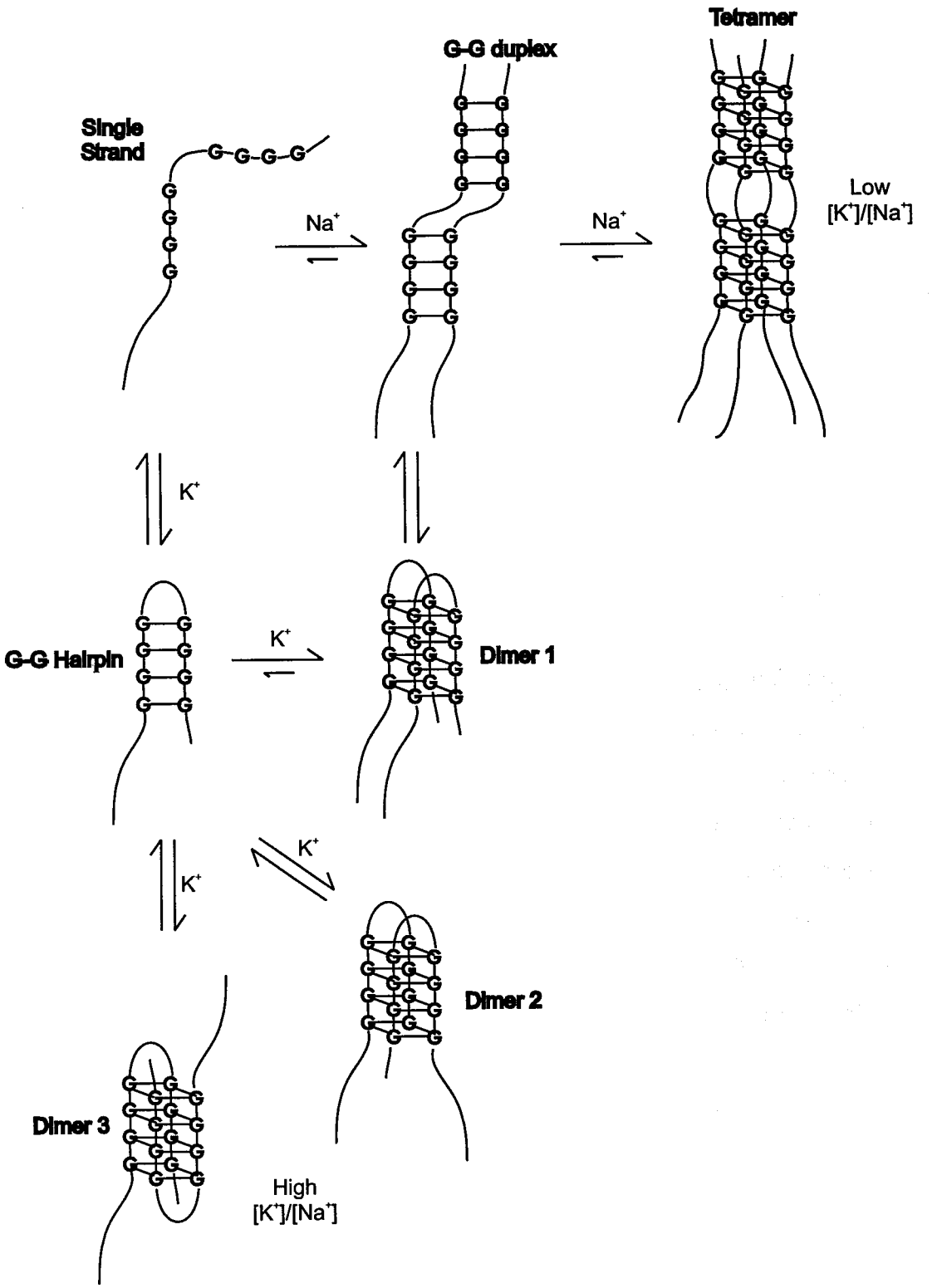


Figure 1.4. The 'cationic switch' scheme for the formation of G4-DNA. Adapted from Sen & Gilbert (1990).

1 to 8. The sequence $d(TG_4)$ formed the most stable tetrameric quadruplex, while an *n* of 2 was the minimum requirement for formation of an antiparallel quadruplex by $d(T_nG_4T_nG_4)$. Additional thymidines resulted in less stable quadruplex for both types. NMR data has also shown that for the quadruplex $[d(T_4G_4)]_4$ the thymines adjacent to the G-quartets are tightly stacked over the G-tetrads, while additional thymines were more disordered (Gupta *et al.*, 1993). This data further indicated that at least two thymines were needed to form dimeric hairpins.

Alteration of the loop sequence by a single base was shown to be responsible for superior G4-stability of several GT-containing anti-HIV oligonucleotides based on the thrombin aptamer (see section 1.14) measured by thermal denaturation (Jing *et al.*, 1997; Jing & Hogan, 1998). This was due to the role of the loop in cation binding. which increased stability and decreased dissociation. K^+ binding was shown to be a two-step process, one K^+ initially bound to the GO6 core, allowing a second K^+ to bind to the loop making it more ordered and compact.

Systematic substitutions of Ts with Cs in the dimeric hairpin quadruplex $[d(G_3T_4G_3)]_2$ showed that $d(G_3T_2CTG_3)$ formed a higher percentage of four stranded quadruplexes. At pH 4 (where C is protonated) all other substitutions formed linear four stranded quadruplexes, except $d(G_3T_4G_3)$ which tended to form a dimeric hairpin quadruplex (Keniry *et al.*, 1997). This was suggested to result from the formation of hemiprotonated $C{\cdot}C_H^+$ pairs by the cytidines substituted within the loop, which increased the stability of these quadruplexes at low pH relative to their non-C forms. These were shown to contribute half the stabilisation energy at pH 6 and ~10% at pH 7 (Hardin *et al.*, 1993).

NMR has also provided some insight into the role of both the loop and G-tract structure related to K^+ and Na^+ binding (Marathias & Bolton, 1999). Sequences with two to four nucleotides in the loops and two guanines can only form edge-type quadruplexes with K^+ but cannot with Na^+. If all the loops contain four residues, then diagonal type quadruplexes are formed in Na^+ or both Na^+ and K^+, but not in K^+ alone. Runs of three or four guanine with two to four nucleotides in the loops formed either topology with either cation. It was also noted that the presence of purines within the loops sterically hindered K^+ binding, thus preventing formation of edge-type quadruplexes.

1.7 Mechanisms of quadruplex assembly

Three mechanisms for quadruplex assembly have been proposed (Hardin *et al.*, 2001), these are outlined in Figure 1.5. Duplex dimerisation, produces quadruplex simply by the assembly of a dimer, followed by the association of two dimers. Stepwise strand addition involves successive additions of single-strands *i.e.* two single strands form a duplex, then the addition of further single strands form a triplex, followed by a quadruplex. A further putative mechanism is triplex disproportionation. The first two steps occur as with stepwise strand addition to produce triplex. The disproportionation of two triplexes associated via the coordination of a cation (a 'hexaplex' intermediate) produces one quadruplex and one duplex.

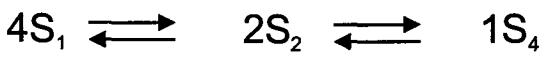
1.8 Quadruplex-protein DNA interactions

A varied class of telomeric and non-telomeric proteins have been discovered which can interact with quadruplexes and in some cases influence their stability and kinetics. Several of these proteins are recombination factors and helicases, which are involved, in the topological control of DNA. The high stability and slow kinetics of quadruplexes has been used to argue against their existence and functioning *in vivo*. However, it is possible that if specific proteins exist *in vivo*, which can influence these properties, this may prove an advantage. Such proteins would allow the precise control of quadruplex formation essential if these structures are involved in regulatory mechanisms.

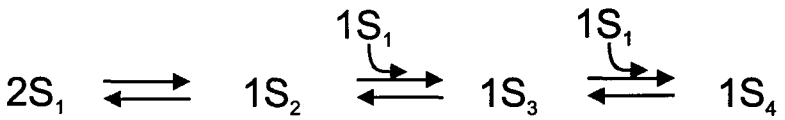
Quadruplex-specific binders

One of the earliest examples of protein recognition of quadruplexes involved MyoD, a transcription factor (Walsh & Gualberto, 1992). A macrophage receptor has been shown to have DNA-binding dependent on quadruplex formation (Pearson *et al.*, 1993). The protein QUAD, was identified from non-histone protein extracts of rabbit hepatocyte chromatin, as binding quadruplex structures (Weisman-Shomer & Fry, 1993). Rat hepatocytes were also found to contain qTBP42 (Sarig *et al.*, 1997), and uqTBP25 (Erlitzki & Fry, 1997), which showed specificity for quadruplexes. Both qTBP42 and uqTBP25 are closely related to human nuclear ribonucleobinding proteins such as hnRBPA1 or A2/B1 and/or their derivative ssDNA binding proteins, UP1 and HDP-1. These proteins are involved in a speculative model for quadruplexes as regulatory control elements (see section 1.13.7).

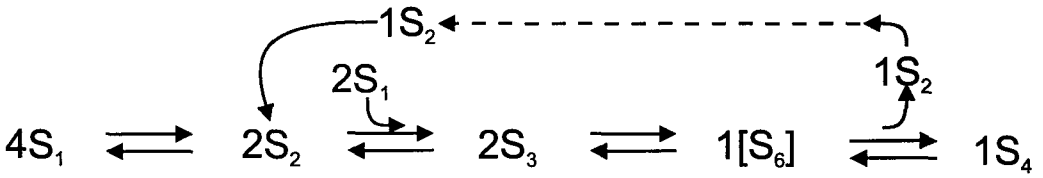
A. Duplex dimerisation



B. Stepwise strand addition



C. Triplex disproportionation



D. Hypothetical mechanism of triplex disproportionation

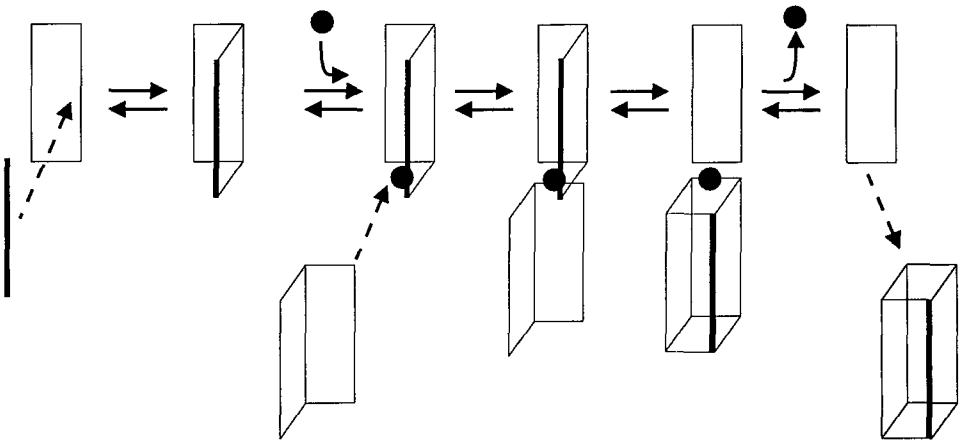


Figure 1.5. Possible mechanisms of quadruplex assembly: (A) Duplex dimerization. (B) Step-wise strand addition, (C) Triplex disproportionation pathway. (D) A hypothetical mechanism for triplex disproportionation adapted from Hardin *et al.* (2001).

G4p1 and G4p2, yeast proteins have been isolated with specificity for binding quadruplexes (Frantz & Gilbert, 1995). LR1, a duplex DNA binding factor that is involved in switch recombination in the ISR (see section 1.13.1), binds more specifically to linear quadruplexes than duplexes (Dempsey *et al.*, 1999). Pur1, a Zn-finger protein that binds several promoters including the insulin promoter (see section 1.13.3), has binding affinity to several DNA sequences which is similar to the ability of the DNA to form quadruplex DNA (Lew *et al.*, 2000).

Quadruplex promoters

Some proteins can actively accelerate quadruplex formation. One example is the β-subunit of *Oxytricha* telomere binding protein, which altered the kinetics of quadruplex formation, resulting in a 10^4- to 10^6-fold rate enhancement (Fang & Cech, 1993). Another early study demonstrated that Rap1p, a yeast telomeric protein, facilitate quadruplex structures (Giraldo *et al.*, 1994). The human topoisomerase I (topoI) has been shown to preferentially interact with quadruplex and can induce a higher rate of quadruplex formation (Arimondo *et al.*, 2000).

Quadruplex-specific nucleases & unwinders

Other proteins are capable of quadruplex dissolution or strand cleavage. Eukaryotic topoisomerase II is one such a protein able to selectively cleave quadruplexes (Chung *et al.*, 1992). A gene product of KEM1 from *S. cerevisiae* has been described as a quadruplex-dependent nuclease (Liu & Gilbert, 1994). SV40 large T-ag helicase, is an enzyme which can unwind duplexes and triplexes, has been shown to unwind several quadruplexes (Baran *et al.*, 1997). Several members of the RecQ DNA helicase family of recombination factors that are significant in several disease states, seem to possess quadruplex unwinding activity. This property may be important to their roles in their respective molecular pathologies (Sun *et al.*, 1998; Fry & Loeb, 1999). If quadruplexes are truly the substrate of these helicases *in vivo*, it is plausible that these structures may play a role in the DNA replication diseases of Blooms and Werner syndrome. Also, *S. cerevisiae* (*SgsI*) helicase (another RecQ helicase) exhibited the most notable preference for unwinding G·G base-paired telomeric substrates over dsDNA (Sun *et al.*, 1999).

1.9 Telomeres & their control

1.9.1 Telomere structure & function

Telomeres are specialised DNA-protein structures found at the ends of chromosomes, which performs two vital functions within the cell. First, the telomere DNA-protein complex acts as a protective cap, allowing them to be distinguished by the cell as 'ends' rather than as dsDNA breaks (DSBs). This prevents their processing as broken DNA which results in chromosome degradation, recombination or fusion. It has been suggested that G-quadruplexes might act as a telomere protective cap (Sundquist & Klug, 1989). Second, telomeres provide a buffer which allows limited shortening of chromosome ends which results from the end-replication problem (Allsopp *et al.*, 1995). This problem arises because DNA polymerisation proceeds from 5' to 3' and requires a 5' RNA primer, a 5' gaps is therefore produced which needs to but cannot be filled, at the very 5' terminus of the lagging strand.

1.9.2 Telomeric DNA

Telomeric DNA in almost all eukaryotes is composed of highly repetitive sequences which extend over several kilobases. Telomeres are double-stranded but for a 3'-terminal single-stranded overhang which may be up to 100 bases long, and are synthesised by telomerase (see below). There is a limited variation in the sequences of the telomeric repeat between even widely divergent species (see Table 1.2), however an universal feature is high G-content. Possible explanations for such similarity, is their common functions which requires them to bind specific proteins or to form G-quadruplex structures (discussed later). The number of repeats varies considerably among eukaryotic species, tissues and cell types from several to over 100 kb.

1.9.3 Telomerase

The reverse transcriptase, telomerase is responsible for synthesising the G-rich strand to maintain the telomeres. It is a ribonucleoprotein, consisting of an essential RNA component and a protein catalytic subunit. The RNA component in humans is a 451nt long RNA termed human Telomere RNA (hTR). It is complementary to, and provides the template for the telomeric repeat, so that synthesis is independent of any exogenous template (Morin, 1989). The catalytic subunit of human telomerase is termed the human Telomere Reverse Transcriptase (hTERT), and has significant sequence identity with other

Table 1.2. Summary of telomeric repeats among eukaryotes.

Group	Organisms	Telomeric Repeat (5'-3')
Mammals	*Homo sapiens.*	TTAGGG
Acellular slime mould	*Physarum polycephalum*	"
	Didymium iridis	"
Filamentous fungus	*Neurospora crassus*	"
Protozoa	*Tetrahymena.*	TTGGGG
	Oxytricha	TTTTGGGG
	Plasmodium	TT(T/C)AGGG
Higher plants	*Arabidopsis*	TTTAGGG
Fission yeast	*Schizosaccharomyces pombe*	$TTAC(A)G_{2-5}$
Budding yeast	*Saccharomyces cerevisiase*	TG_{1-3}
Cellular slime mould	*Dictyostelium discoidum*	AG_{1-8}

Table 1.3. Summary of naturally-occuring G-rich sequences (purine strand only). Gs participating in G-tetrads are in red where known.

G-rich region	Sequence
Telomeres	$(TTAGGG)_n$
Adenovirus major late promoter	AAGGGGGGTATAAAAGGGGGTGGGGGCGC
Ig switch regions	TACAGGGGAGCTGGGG TAGA
Retinoblastoma susceptibilty gene	GGGGGGTTTTGGGCGGCA
Human insulin gene promoter	ACAGGGGTGTGGGG
Fragile X triplet repeats	$(C\underline{GG})_n$
Retroviral DLS	RGGARA
Chicken β-globin gene promoter	G_{16}CGGGTGGTGG
Human c-myc P1 promoter (NHE)	TGGGGAGGGTGGGGAGGTGGGGAAGG

reverse transcriptases (Morin, 1989).

For a primer the telomerase uses the 3'-end of the existing G-rich strand. *In vitro* studies on telomerase have shown a requirement for a G-rich strand of a minimum length of 10 to 12 nts for telomerase activity, similar to the length of the 3'-overhang needed for high affinity recognition as primer *in vivo* (Morin, 1989). Following hTR binding to the telomeric repeat, the telomere is extended by successive rounds of extension then translocates to the new 3' end, to add upwards of several hundred nucleotides. The repeated repositioning of the primer is exceptional in occurring for so many rounds before final separation of the primer and telomerase. It has been suggested that the transient formation of a quadruplex might displace the telomeric DNA from the RNA component, and so drives the translocation step.

1.9.4 Telomeric proteins

Several important telomeric structural proteins have been identified in mammalian, yeast and ciliates, which bind the telomeres in a sequence specific manner both to the double strand and single strand overhang. These are involved not only in protecting the single strand end from nuclease attack but regulate the action of telomerase. Their binding to telomeric DNA to form a telomere complex is likely to be dynamic and highly regulated throughout the cell cycle.

The telomeric repeat–binding factors TRF1 and TRF2 have been shown to interact with telomeric dsDNA in mammalian cells (Smogorzewska *et al.*, 2000). TRF2 has been reported using electron microscopy to promote the formation of novel structures by telomeric DNA *in vitro* termed T-loops (Griffith et al., 1999). These structures were suggested to form by the invasion of the 3′ telomeric overhang into the duplex telomeric repeat region. T-loops may provide a general mechanism for the protection and replication of telomeres.

TRF1 and TRF2 are important in the negative regulation of telomere length, competing with telomerase for access to telomeres. High levels of ADP-ribosylation lead to dissociation and possible degradation of TRF1 which allows telomerase to access and elongate the telomeres. TRF1 binding is regulated by the poly ADP-ribosylase, tankyrase

and may be opposed by poly ADP-ribose glycohydrolases (PARGs) (Smith & deLange, 2000).

1.10 Telomerase activity & cancer

The gradual erosion of telomeres acts as a mechanism for 'counting' divisions, ultimately triggering replicative senescence in many cell types. Telomere length is a function of addition by telomerase and loss by 5' attrition (through successive rounds of cell division). Telomere length varies considerably in eukaryotes, depending on interplay of the many influences both physiological and genetic. Telomere persistence is necessary for long-term proliferation. Actively proliferating (immortal) cells such as germ-line or stem cells possess an active telomerase, arresting telomere erosion and allowing unlimited cell division (Counter *et al.*, 1992). In contrast, this enzyme is not present or is inactive in most somatic cells.

Somatic cells are assumed to possess telomeres of sufficient length to adequately meet the loss accumulated in a typical lifetime of cell divisions. The length acts as a 'biological timer', limiting their proliferative lifespan to a finite number, before cumulative loss from many cell divisions is irreparable and cellular senescence (or apoptosis) is reached. However, telomerase is reactivated in most human tumour cells; a figure of ~85% was given in a review of abnormal telomerase activity in tumour types, whereas 0.5% of normal cells were positive for telomerase activity (Shay & Bacchetti, 1997). This seems to suggest that the ability to maintain telomeres is an essential element of the immortalisation process.

The weight of evidence is sufficient to support the view that telomerase is associated with cellular proliferation and cancer (reviewed in Holt & Shay, 1999). First, transfection of telomerase-negative cell types with telomerase was sufficient to extend their lifespan and sustain immortalisation (Bodnar *et al.*, 1998). Second, the reversal of telomere shortening as a result of telomerase transfection in endothelial and fibroblast cells, provided the first evidence that telomerase activity has a causal relationship to tumourogenesis (Hahn *et al.*, 1999a; Elenbaas *et al.*, 2001). Third, the inhibition of telomerase limits the growth of human cancer cells (Hahn *et al.*, 1999b). Telomerase repression has been associated with tumour or stem cell terminal differentiation, quiescence, and growth arrest (Holt & Shay, 1999). Four, the progressive loss of telomeres due to a lack of telomerase reduces

proliferation and self-renewal potential; telomere erosion due to telomerase deficiency coincides with compromised proliferation and increased apoptosis in germ cells of mice (Lee *et al.*, 1998).

1.11 Targeting telomere maintenance by telomerase

Telomerase activity is acknowledged as the primary means of telomere persistence in highly proliferative tumours. Anti-telomerase therapy is therefore a highly selective means of cancer intervention, and many different strategies have been examined (Mergny *et al.*, 2002). First, the catalytic subunit (hTERT) can be targeted by clinically useful reverse transcriptase (RT) inhibitors; or inhibitors of hTERT phosphorylation to interfere with its regulation (Ku *et al.*, 1997). Second, the RNA component (hTR) can be also be targeted by various synthetic antisense RNAs which will not need to compete with the ribosome machinery (Kondo *et al.*, 1998). A similar strategy uses a ribozyme (teloRZ) to specifically cleave hTR (Kanazawa *et al.*, 1996). Expression of mutant hTR in cancer cell lines can be used to produce dysfunctional telomeres which cannot hybridise to the normal telomeric sequence (Kim *et al.*, 2001). The transient heteroduplex formed by hTR and telomeric DNA can be cleaved or specifically stabilised to prevent dissociation (Francis *et al.*, 2001). Third, it is possible to perturb the assembly of active enzyme by molecular chaperones can using drugs (Matsutomi *et al.*, 2000; Stebbins *et al.*, 1997; Wadhwa *et al.*, 2000). Four, the binding of TRFs to the telomeric DNA (discussed in section 1.9.4) can be disrupted. Five, most relevantly to this thesis, the structure of the telomeric DNA can be altered. This can be achieved by ligands that selectively induce intrastrand linkage or cleavage (Burger *et al.*, 1997; Lin *et al.*, 2001; Yoon *et al.*, 1998). More significantly, the folding of a telomerase primer into a quadruplex can be used to inhibit telomerase primer extension (Zahler *et al.*, 1991). Drugs can be used to promote quadruplex formation and exploit this as detailed in section 1.15.

1.12 Limitations to targeting telomeres

The targeting of telomerase is not likely to be the cure for cancer. The upregulation of telomerase may only be an essential prerequisite for some cancers to develop. In others, it is not essential, but may only be associated with poor prognosis, endowing cancer cells with an advantage over normal cells. As such telomerase reactivation is not a definitive cause of cancer, but rather it is a major contributing factor – it has been accepted as one of

the six hallmarks of cancer (Hanahan & Weinberg, 2000).

Fundamentally, several obstacles exist if any anti-telomerase strategy is to clinically useful. Alternative mechanisms, though less frequent, do exist for maintaining the telomeres independently of telomerase in yeast and mammalian cells. A second mechanism exists for overcoming critically shortened telomeres; telomere fusions and recombinational telomere elongation (for a review see McEachern *et al.*, 2000). It is unclear whether these mechanisms can be exploited jointly with telomerase activity.

Another obstacle is the delayed therapeutic effects of anti-telomerase agents. Agents which act to induce progressive telomere shortening, only produce anti-proliferative effects when they reduce the telomeres to a critical length. As such a lag is therefore seen before these agents achieve tumour cell crisis. More active drugs are therefore needed which can deplete the telomeres after shorter exposures. Variation in telomere length and discrepancies in the rates of loss vary considerably, determining in which tumours treatment is viable. It should be noted that an immediate response to telomere interference has been observed, this may arise from interference with the other essential function of telomeres , the 'capping' that prevents telomere fusions.

Telomerase activity is not unique to tumours, it is also found in germ cells and other somatic cells, such as haematopoeitic stem cells, lymphocytes, basal layer cells in skin and hair follicles and small intestine. As such any suppression of telomerase will also be harmful to these cells, leading to unavoidable toxicity. Preliminary data indicate a transitory inhibition in these cells is tolerable as: (*i*) they possess longer initial telomeres than cancers; (*ii*) these normal cells though active still divide less frequently; and (*iii*) once the desired telomere depletion results in cell crisis in the target cells, the drug treatment can be withheld to allow the normal cells to regain their initial telomere lengths (Herbert *et al.*, 1999).

1.13 Other G-rich regions

The abundance of potential quadruplex-forming regions as shown in Table 1.3 implies several biological roles exist for quadruplexes. Quadruplex formation in such regions has attracted far less investigation than have telomeric sequences. As many are within

promoter regions one role may be transcriptional regulation. Many of these G-rich regions are 'hot spots' of recombination, possibly as a result of quadruplex-mediated transcriptional arrest.

1.13.1 Ig switch regions

The idea that a four-stranded nucleic acid structure formed by Hoogsteen bonded G-tetrads might be of biological interest was first suggested by Sen and Gilbert (1988). Serendipitously, oligomers corresponding to G-rich regions in the immunoglobulin-G (IgG) switch region were stored in monovalent-cation containing buffers. This unexpectedly resulted in the formation of these little known structures. The Ig heavy chain genes are repetitive stretches of 1 to 10 kb are found upstream of constant regions, they are composed of tandem repeats of GC rich DNA 20 to 50 bp long (50% guanine content). Repetitive motifs include d(G_4T), d(GAGCT), and d($G_{4/5}$AGCTG_4). These switch regions are known to undergo recombinations. It was therefore proposed that quadruplex structures might act to bring together four homologous chromatids during meiosis. Tight association of individual chromosomes would be allowed due to their unique patterns of G-rich loci. It was also speculated (before any such discoveries) that unknown proteins might melt the duplex and bind the C-strand to free the single-stranded G-strand to associate.

1.13.2 Retinoblastoma susceptibility genes (hRb)

The retinoblastoma susceptibility (Rb) gene encodes a tumour suppressor phospho-protein that may regulate oncogene expression including that of *c-fos*. The *Rb* gene is a further example of a G-rich sequences in the genome. It has an 84% G-C base composition within a 200 bp section that includes the first 100 bp of its coding sequence, and contains several G-stretches. In the presence of Na^+ ions, it was demonstrated that oligomers representative of 5'Rb sequences can form a tetrameric quadruplex (Murchie & Lilley, 1992). Notably, the sequence is part of a gene that itself encodes a DNA binding protein.

1.13.3 Insulin promoter region (ILPR)

The insulin-linked polymorphic region (ILPR) is responsible for genetic susceptibility to insulin-dependent diabetes mellitus (IDDM). It is a polymorphic G-rich mini-satellite, located 363 bp upstream of the human insulin gene within the promoter region. It comprises of tandem repeats related to d(ACAG_4TGT_4), and eleven variants (*a –k*) have

been identified.

ILPR has been shown to form unusual DNA structures, which are supposed to be quadruplex *in vitro*. The ability with which variant ILPR repeats adopt quadruplex structure influences the efficiency with which its transcriptional factor Pur-1 activates the insulin gene. Biochemical studies indicated that the commonest repeat, variant *a*, has the highest propensity to form G-quartets (Hammond-Kosack *et al.*, 1992). The less common repeats had lower transcriptional activity *in vivo,* supporting the idea that the effects of single nucleotide variations within ILPR on transcriptional activity correlate with an altered disposition for quadruplexes. It was shown that both inter- and intra-molecular quadruplex formation can influence this activity. Pur-1 binding (equating to transcriptional activity) to tetraplex was investigated and it was shown that Pur-1 will only bind upwards of 4.0 repeats of variant *a* but not unrelated double- or single-stranded DNA (Lew *et al.*, 2000).

The authors also noted from unpublished data that a library of ILPR interacting proteins had been discovered (from cDNA isolated from insulinoma), among which some exhibited antagonistic binding properties to Pur-1 (*i.e.* higher affinities for repeats with low folding). activity. Whether any others might be quadruplex-catalysing or interactive is unknown. Though diabetes susceptibility is complex (14 distinct loci including *ILPR* are involved), the involvement of quadruplexes may be implicated, as they represent a means for generating single-strands.

It is also noteworthy that the promoter region of the rat preproinsulin II gene was reported to be S_1 nucleases sensitive (Evans *et al.*, 1984). It seems a good candidate sequence for inclusion as a potential quadruplex competent G-rich sequence, though this was not appreciated as this area of research had not been conceived.

1.13.4 Fragile X triplet repeats (FRM1 gene)

Fragile X syndrome is the most common cause of mental retardation. It is associated with a folate-sensitive fragile site that contains an expanded stretch of the d(CGG) trinucleotide repeat in the 5'-untranslated region of the FMR1 gene. A healthy individual will possess typically 6 to 54 repeats, in a carrier this is elevated from 50 to 200, and in a sufferer this reaches between 200 and 2000.

The ability of this G-rich locus to form a quadruplex structure under physiologically relevant conditions has been shown. The oligomer d(^{Me}CGG)$_n$ can form stable structures biochemically characterised as quadruplex of 4, 5 or 7 repeats at pH 8. The 5-methylation of the cytosine residues of the short d(^{Me}CGG)$_5$ tract was essential, except where there are 7 repeats. (Fry & Loeb, 1994). Contrary to the normally slow kinetics of generation, its assembly into quadruplex was rapid at concentrations of 4-15 μM, due to the high G content. Also the structure had a high melting temperature, making it possibly the most stable natural genetic element so far identified. Interestingly, the enhanced quadruplex formation by methylation and the facilitation of methylation by quadruplex formation presents a circular 'cause and effect'.

Fragile X is characterised by transcriptional inactivation of the FRM1 gene, delayed replication ≥150kb 5' and ≥34kb 3' of the expanded region, methylation of CpG islands, and adjacent d(CGG) repeats. It is not understood how d(CGG) expansion and hypermethylation causes suppression of the FRM1 and delayed replication. However, it is hypothesised that its folding into hairpin structures and/or assembly into quadruplex structures may be the cause. Their formation *in vivo* may block the replication fork and cause expansion by replication slippage. It is possible that fewer than 50 repeats may be manageable, but a greater number might overwhelm any repair system. The availability of specific binding proteins, (normally sufficient) may be insufficient to prevent this in longer, expanded stretches, thus leaving it exposed and susceptible to such events.

1.13.5 Retroviral dimer linkage site (DLS)

Two identical RNA molecules that are non-covalently associated constitute the genome of a retrovirus. How this genomic RNA assembles is unclear, though it is know that the most stable contact occurs at a site near the 5'end termed the dimer linkage site (DLS). Although the DLS sequences are disparate in retroviruses, a consensus appears to be RGGARA. This stable association is considered important in the retroviral infectious cycle, and DLS alignment is thought to facilitate subsequent recombination mediated by the nucleocapsid (gag) protein.

Though unclear, association is believed to involve a non Watson-Crick base-paired RNA-RNA interaction, which aligns the strands in a parallel orientation. The conserved structural motif mediating association is identical rather than complementary. No protein association is thought to be necessary and RNA genomes such as that of HIV-1 can dimerise *in vitro* in the absence of co-factors, and can be stably isolated without proteins.

Two different purine-rich RNA fragments of the HIV-1 DLS have been independently shown to dimerise at high ionic strength (without any protein co-factors) (Sundquist & Heaphy, 1993; Awang & Sen, 1993). Deletion analyses of a 127nt fragment determined that a 94nt region at its 3'- end was needed for association, and the loss of G_5AGA_2 was non-permissible (Sundquist & Heaphy, 1993). The antisense oligomers used to define the essential region of a 111 nt fragment of the DLS were found to be even more purine-rich than the DLS and five contiguous pyrimidines were not needed for dimer formation (Awang & Sen, 1993).

The kinetics, purine-rich nature and ionic dependencies are all suggestive of a G-tetrad structure within the DLS, though AG-quartets cannot be ruled out. This strongly supports the idea that quadruplex formation mediates the dimerisation *in vivo*. Although, caution is needed before this is concluded, as the rate of formation under optimal conditions is too low for viral infection, and the stability once formed is too great. A highly basic HIV-1 protein, such as the nucleocapsid protein, may overcome these unfavourable kinetics.

1.13.6 β-Globin

A K^+-dependent DNA synthesis arrest site was identified in the G-rich strand of the promoter region of the chicken β-globin gene: $G_{16}CG(GGT)_2GG$. (Woodford *et al.* 1994). The G-rich sequence, the absolute requirement for K^+, the template concentration-independence, the stability at elevated temperatures and non-Watson-Crick base pairing are all consistent with blockage by adoption of an intramolecular quadruplex.

Unusually, the arrest site lacked the typical four G-repeats, and a novel quadruplex structure was proposed termed 'cinched'. This tetrahelical structure lacked four perfect G-tetrads, rather incorporated non-canonical bases into the stem, and 5'- and 3'-flanking regions hydrogen-bonded to form a 'cinch' which holds one end together. This was

supported by experiments which compared a known quadruplex [$d(T_2G_5)_4$] to an identical one where one tetrad was disrupted by substitution of a single guanine for a cytidine: $d(T_2G_5T_2G_2C^*G_2(T_2G_5)_2)$ (Howell *et al.*, 1996). Both the arrest of DNA synthesis and the DMS protection of the stem were also still observed (though decreased). This supported for the idea that non-guanines can be accommodated into a quadruplex. It was speculated that the quadruplex might have a role in transcriptional regulation, as the sequence might function as a K^+ switch modulating expression of the gene. Interestingly, these results extends the range of potential sequences which may form quadruplexes, particularly those where flanking sequences may interact similarly to stabilise and reduce perturbation to the framework of G-tetrads of unconventional forms.

1.13.7 *c-myc* promoter (NHE)

The human *c-myc* oncogene belongs to a family of oncogenes that encode phospho-proteins, which activate genes, vigourously stimulating cell growth. Illegitimate activation of the human *c-myc* oncogene is detectable in and is key to the progression of a variety of malignancies. The mechanism responsible for *c-myc* over-expression is obscured by the use of four promoters, however, it is known that a major control element is the nuclease hypersensitive element III (NHE), found 115 bp upstream of the P1 promoter. The NHE is a purine·pyrimidine rich 27 bp region. Its exact role is unclear, though it is believed to control upwards of 95% of *c-myc* transcription and its presence is associated with the activation of this oncogene (Siebenlist *et al.*, 1984).

This region interestingly has a tendency to form atypical DNA structures under superhelical stress. A tandem H-DNA structure comprised of two intramolecular pyrimidine·purine·pyrimidine triplexes was suggested for this region which activated *c-myc* (Kinniburgh *et al.*, 1994). However, the need for 5-20 mM Mg^{2+} for antiparallel triplex formation and low pH for parallel triplexes, is incompatible with the intracellular environment. In contrast, K^+-dependent tetraplex appears to be adopted *in vitro* by the guanine-rich strand (Simonsson *et al.*, 1998). Fluorescence studies (Simonsson & Sjöback, 1999) indicated that optimal FRET was obtained with the central 22-mer motif of fluorescein-5'-$d(G_3TG_4AG_3TG_4)$-3'-tetramethyl rhodamine. A 16-mer or a 33-mer capable of forming G-tetraplex showed reduced fluorescence transfer. This suggested that the adoption of a preferred structure composed of three G-tetrads planes; two two-base

intervening loops (foremost G_4pA_5 and endmost $T_{18}pG_{19}$) possibly containing two intrastrand base pairs A_5-T_{18} and T_9-A_{14}; and a central six base loop (TG_4A). The cytosine-rich strand has also been reported to form several distinct i-motif structures (Simonsson *et al.*, 2000). These observations have led to the suggestion that non B-DNA structures act as *cis*-elements for the transcription initiation of *c-myc*: preventing nucleosome packaging and chromatin unwinding that cause the nuclease sensitivity.

The interaction of several proteins with the NHE has been implicated in c-myc activation. The proteins NDPK-B and hnRNP K are *c-myc* transcriptional factors NDPK-B has helicase activity, binds to the pyrimidine single strand and augments NHE expression (Hildenbrandt *et al.*, 1995). hnRNP K seems oblivious to duplex NHE but binds the pyrimidine strand (Michelotti *et al.*, 1996), and has been shown to interact *in vivo* with the RNA polymerase II transcription machinery (Tomonaga & Levens, 1996). hnRBP A1 has been also demonstrated to bind the NHE, and being homologous to other G4 binding protein (previously discussed in section 1.7) might act likewise to facilitate tetraplex. A zinc finger protein, CNBP has purine strand binding activity and activate transcription of *c-myc* (Michelotti *et al.*, 1995).

Incorporating this research into a working model, Simonsson and co-workers conceived a scheme whereby tetraplex formation might act in concert with these proteins. (i) Normally *c-myc* is silent, and the chromatin must first unfold to trigger activation by an unknown mechanism (possibly by alteration of cellular K^+). (ii) Once the nucleosome is dislodged and no longer constrains the chromatin, NDPK-B unwinds duplex NHE and binds the pyrimidine single strand. (iii) hnRBP A1 and/or CNBP bind to the tetraplex formed by the purine strand, while hnRBP K can then bind the exposed pyrimidine strand and attract the RNA polymerase machinery and activate *c-myc*. Indeed, recently a report has directly implicated quadruplex formation in *c-myc* regulation *in vivo* (Siddiqui-Jain *et al.*, 2002). Mutations which destabilised an intramolecular quadruplex in the NHE were reported to result in a 3-fold increase in *c-myc* transcription. Further, a known G4 ligand, TMPyP4 was shown to suppress *c-myc* transcriptional activation, but only in cell lines where the NHE element was not deleted.

1.14 Biological effects of quadruplex-forming-oligomers

Quadruplexes are formed by a number of G-rich aptamers that show potential therapeutic effects. Using selective enrichment of combinatorial DNA libraries, the thrombin aptamer was identified (Bock *et al.*, 1992). It is a single-stranded DNA with a highly conserved sequence d($G_2NTG_2N_{2\text{-}5}G_2NTG_2$) essential for high affinity binding to thrombin. Several further GT- containing anti-HIV oligonucleotides have been identified which fold into quadruplex structures based on d($GTG_2T\ G_3TG_3TG_3T$) (Rando *et al.*, 1995). Their activity is believed to be derived from their quadruplex forms, which selectively bind to HIV integrase and so interfere with the insertion of retroviral dsDNA transcripts into the host genome. Their quadruplex stability (based on their loop structure) is highly related to their activity (Mazumder *et al.*, 1996; Jing *et al.*, 1997; Jing & Hogan, 1998).

Oligonucleotides have also been developed that **suppress *c-myc***. Nanomolar concentrations of a G-rich 22-mer reduced the growth of Burkitt's lymphoma cells which constitutively express the *c-myc* gene, while the effectiveness of the oligonucleotide was enhanced by preforming quadruplex (Simonsson & Henriksson, 2002). It was suggested that the unstructured oligomer bound the pyrimidine-rich strand of NHE within the *c-myc* promoter, preventing its recognition by hnRNP K necessary for activation of transcription. While the oligonucleotide folded into a quadruplex competed with the purine-rich strand for binding of hnRBP and CNBP which otherwise promote its dissociation from the pyrimidine strand into tetraplex (see section 1.13.7). This model explained how antigene therapy had silenced *c-myc* in experiments that had attributed the effects to triplex formation by polypurine oligomers in HeLa cell extracts (Cooney *et al.*, 1988) and *in vivo* in HeLa cells (Postel *et al.*, 1991).

1.15 Use of stabilising agent to target G-quadruplex in G-rich regions

With the acceptance that quadruplex formation can inhibit telomerase action, and so provide a novel molecular target, came the need for new quadruplex-specific ligands. Many quadruplex ligands have now been described, which selectively recognise the peculiar geometry of the G-tetrad. These are principally planar or extended-aromatic molecules, which are presumed to stack either within stacked G-tetrads or between the terminal G-tetrads within the loop. Opening of an intercalation site within stacked G-quartets is less likely than terminal stacking, and the sidechains of these ligands allow

interactions within the grooves of the quadruplex or with the loop. The various quadruplex topologies possess specific groove sizes and loop orientations that provide an under exploited opportunity for differential recognition by G4 ligands (Han *et al.*, 2001). Such ligands are required to possess lower telomerase inhibition than general cytotoxicity (i.e. $^{tel}IC_{50} << IC_{50}$), and higher quadruplex affinity than affinity for duplex, triplex or RNA (i.e. K_d in the high nanomolar to low micromolar range).

It is important to appreciate the standard methods, that are used to characterise the potencies of such novel quadruplex-binding agents. One widely used technique is Telomeric Repeat Amplification Protocol (TRAP) that measures telomerase activity (Kim *et al.*, 1994; Krupp *et al.*, 1997). It is based on telomerase-mediated extension of a forward primer, followed by PCR amplification of the extension products with added reverse primer and *Taq* polymerase. The PCR products can be quantified with and without ligand and $^{tel}IC_{50}$ values (the concentration required to inhibit 50% telomerase) determined. TRAP is useful as a quantitative cell-free assay of the inhibitory activities of anti-telomerase compounds. However, it must be shown that the inhibition is a direct result of sequestering the telomeric primer into quadruplex, rather than by non-specific effects on the *Taq* polymerase. Molecular modelling is often useful and many ligands have theoretical quadruplex affinity which correlate well with their experimental telomerase inhibition. Melting studies of quadruplexes using UV absorbance or Fluorescence Resonance Energy Transfer (FRET) probe the secondary structure of a wide range of oligonucleotides in the presence of drug molecules (Mergny *et al.*, 2001). An alternative approach is to follow changes to the absorption or emission spectra of the ligand upon quadruplex-binding. Also electrophoresis (bandshift) has been used to separate and quantitate the mole fraction of folded quadruplex and single strand induced by drug-binding (Han *et al.*, 1999). Typically, NMR structural data on quadruplex-ligand complexes can be obtained but are often limited by extensive drug-induced line broadening (Fedoorff *et al.*, 1998). A competition (or equilibrium) dialysis assay can also be used to rapidly screen specific ligand binding to a variety of nucleic acid structures (Ren & Chaires, 2001). Here, a ligand solution is incubated overnight with dialysis units containing different nucleic acid structures (various single-stranded, duplex, triplex, quadruplex DNA). At the end of the equilibration period, drug concentration in each dialysis unit is determined and compared.

1.15.1 Anthraquinones, Fluorenones & Acridines (Figure 1.6)

2,6-di-amido-**anthraquinones** (AQs) were first shown to be selective triplex-interactive compounds with low duplex affinities (Fox *et al.*, 1995). Their bind similarly to quadruplex DNA was therefore also considered. An early 2,6-anthraquinone was predicted by modelling to intercalate within a quadruplex and was found to give an $^{tel}IC_{50}$ of 23 μM (Sun *et al.*, 1997). Intercalation of the molecule was also supported by NMR data. 1,4- and 2,6-anthraquinones, with amino-alkyl-amido sidechains terminating with basic groups such as piperidine, produced $^{tel}IC_{50}$ from 4 to 11 μM, comparable to cytotoxicity *in vitro* in several tumour lines (Perry *et al.*, 1998a). Furthermore, the structure-activity studies of 1,5- 1,8- and 2,7-anthraquinone isomers produced $^{tel}IC_{50}$ from 1.3 to 17.3 μM, with little regio-selectivity, suggesting no strict positional requirements (Perry *et al.*, 1998b). Physical studies suggested that these all had a 1:1 binding stoichiometry, and three distinct modes of intercalative binding were predicted determined by the positions of substituents. Failed attempts to co-crystallise a 1,4-bis-piperidino-anthraquinone did however, give fiber diffraction patterns for a $[d(TG_4T)]_4$ quadruplex that suggested that two ligands were bound by intercalation or end-stacking. Based on this result, molecular modelling predicted external stacking was energetically favoured (Read & Neidle, 2000).

Neidle and co-workers synthesised a series of **fluorenone** derivatives (Figure 1.6B). The fluorenone moiety has a five- rather than a six-membered central ring as in quinone, and just one carbonyl group. Molecular modelling predicted that the crescent-shaped chromophore would create distortion in the binding site and that the loss of one carbonyl group would reduce the electron deficiency of the ring system, reducing π-stacking with the G-tetrad. When fluorenones were compared to their equivalent anthraquinone analogues decreases in undesirable cytotoxicities were not associated with reduced telomerase inhibition. Simple alkyl-amino-, piperidinyl- and piperazinyl- derivatives attached via propion-amido- substituents gave optimal activity, the most potent $^{tel}IC_{50}$ were from 8 to 12 μM. NMR indicated that intercalation at the ApG step (Perry *et al.*, 1999).

A rationally designed series of 3,6-disubstituted **acridines** (ACIs) (Figure 1.6C) was somewhat successful at inhibiting telomerase (Harrison *et al.*, 1999). It was predicted by molecular modelling that the acridine moiety would be at least comparable to the anthraquinone moiety in terms of quadruplex affinity and would similiarly end-stack to the

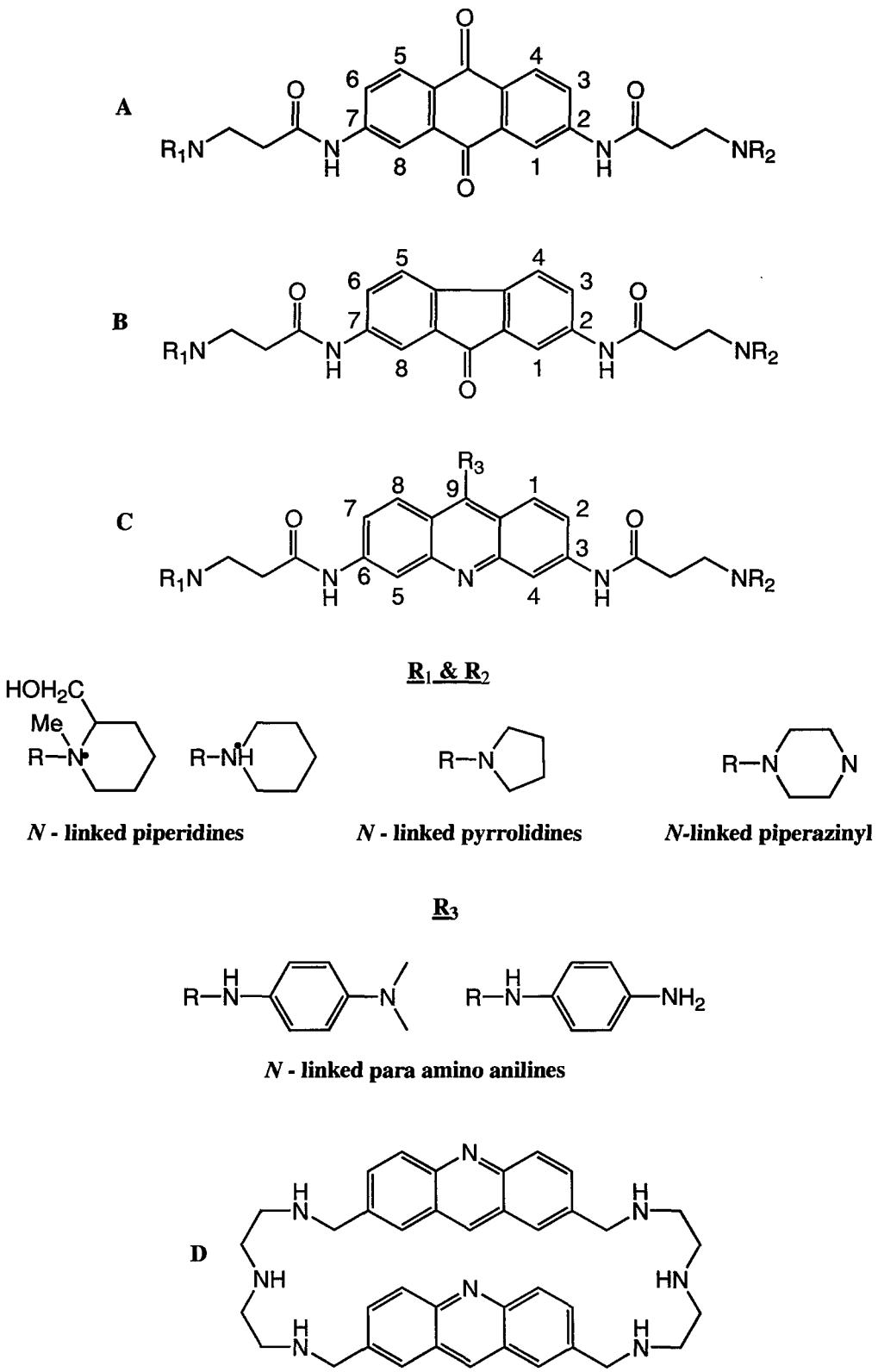


Figure 1.6. Structures of (A) 2,7-amidoalkyl-anthraquinones, (B) 2,7-amidoalkyl-fluorenones, (C) 3,6,9- trisubstituted acridines, with the common *N* - linked R substituents, and (D)Bis-acridine.

G-tetrads (Read *et al.*, 1999). Acridine contains a heterocyclic nitrogen atom capable of protonation at physiological pH. This heteroatom increases electron deficiency in the chromophore, enabling stronger quadruplex interactions over anthraquinones. A recent crystal structure (Figure 1.7) has shown that protonation of the nitrogen of acridine allows hydrogen bond interactions with thymines within the loop (Haider *et al.*, 2003). The same structure also showed that the sidechains do not interact within the grooves but with the loop. Predicted improvements to binding energy with various substituents for human quadruplex showed a close relationship with telomerase inhibition (Read *et al.*, 1999).

3,6-substituted acridines showed similar activities to those of the 2,6-anthraquinones, with $^{tel}IC_{50}$ values of 1.3 to 8 μM. Molecular modelling predicted that tri-substitutions at the 3,6,9-positions would benefit from additional interactions within a third groove of quadruplex , producing enhancing relative binding energies and selectivity over 3,6-di-substituted analogues. This was validated by a $^{tel}IC_{50}$ of 60 nM for the anilino-substitution at position-9 of acridine (Read *et al.*, 2001). Recently, these tri-substituted acridines were shown to be potent inhibitors of Bloom's and Werner's syndrome helicases which act on both G-quadruplex and duplex DNA. This inhibition was associated with reduced levels of G-quadruplex binding and a reduced DNA-dependent ATPase activity by the helicases (Li *et al.*, 2001).

Another novel acridine developed for quadruplex interactions was the acridine dimer, Bis-A, shown in Figure 1.6D (Alberti *et al.*, 2001). Bis-A was initially designed to bind hairpin rather than duplex targets as its linker arms prevent intercalation. As such it was hoped it might interact with quadruplexes at the loops or stack at the terminal G-quartet. Bis-A increased the T_m of melting profiles (probed by FRET melting studies) of intramolecular G- and intercalated C- (i-motif) quadruplexes adopted by a 21-mer human telomeric G-rich sequence and its complementary C-rich sequence. Bis-A produced a ΔT_m of 15 and 33°C at 1 μM for G- and C-quadruplexes respectively. However, it may not be appropriate to compare the increases to the stability of G4 to C4, as their association-dissociation equilibria differ, and the C-quadruplex is kinetically trapped. Bis-A also delayed duplex formation significantly, equilibrium dialysis showed that it preferred quadruplex DNAs over an array of disparate DNA forms, and it produced Bis-A produced a $^{tel}IC_{50}$ of 0.74 μM.

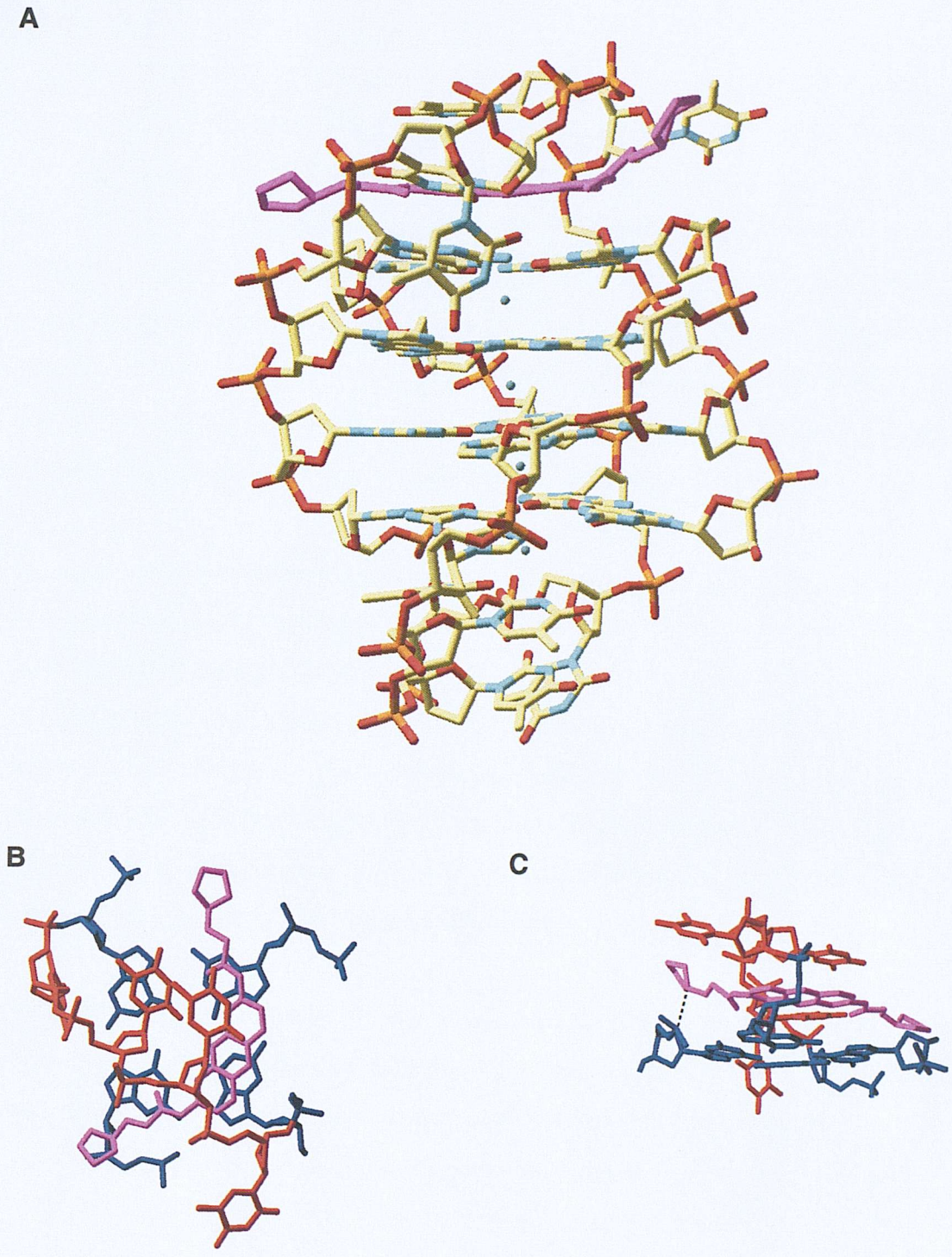


Figure 1.7. (A) The crystal structure of the complex between the acridine BSU6039 (coloured magenta) and the dimeric $d[G_4T_4]_2$ quadruplex (taken from Haider *et al.*, 2003). (B) A view onto the plane of the acridine chromophore (magenta), showing its stacking onto the adjacent G-quartet (blue). (C) A view of the bound acridine molecule, sandwiched between a G-quartet and the thymine loop (red).

1.15.2 Cationic porphyrins (Figure 1.8)

Porphyrins (TMPyPs) have previously been utilised for photodynamic cancer therapy. The planar aromatic ring structure suggests the possible interactive stacking with the G-tetrads, and G-quadruplexes have been shown to be stabilised by the cationic porphyrin, TMPyP4 (Anantha *et al.*, 1998; Wheelhouse *et al.* 1998). Studies of visible absorption, circular dichroism and FRET between the chromophore and DNA, all supported end-stacking with a 1:1 binding mode and a high binding affinity for quadruplex DNA ($K_d = 3.7 \times 10^{-8}$ M), which was twice that for duplex DNA (Anantha *et al.*, 1998). Quadruplex binding by the porphyrins has also been supported by photocleavage and the interaction was shown to inhibit telomerase with a $^{tel}IC_{50}$ of 6.5 μM (Wheelhouse *et al.*, 1998).

The interaction of porphyrins by end-stacking to stacked G-tetrads has been supported by the location of oxidative cleavage of a human telomeric sequence by the metallo-porphyrin (Mn-TMPyP4/$KHSO_5$) (Han *et al.*, 1999a; Han *et al.*, 2001). However, the stoichiometry of binding determined by Haq *et al.* (1999), using ITC and spectrophotometry, seemed to imply the intercalation of porphyrins at each GpG site (even though this violates the nearest neighbour exclusion principle). Molecular modelling suggests that ΔG for external stacking will be driven by a large ΔH with a small unfavourable $T\Delta S$, whereas intercalation will be driven by a small ΔH and large $T\Delta S$ (Han *et al.*, 2001).

TMPyP4 inhibited telomerase whereas its isomer TMPyP2 did not produce any inhibition. This possibly reflects an inability of TMPyP2 to stack externally to the G-tetrad at the GpT step, while TMPyP4 can bind externally in the TTA loop (Han *et al.*, 1999a). The TMPyP3 and 4 isomers seem to differentiate between different quadruplex topologies. TMPyP3 interacts more strongly with tetrameric quadruplexes than TMPyP4 (measured by inhibition of *Sgsp1* helicase), while the converse was true for antiparallel quadruplexes. TMPyP2 did not interact with either form (Han *et al.*, 2001). Tumour cell lines showed slower cell growth in TMPyP4 than in TMPyP2, and TMPyP4 alone induced the formation of anaphase chromosomal bridges. That the interactions with quadruplex *in vitro* mirrored their biological effects supports the idea of targeting the telomeres in cells (Izbicka *et al.*, 1999). Furthermore, along with the trisubstituted acridines, TMPyP4 has also been shown to inhibit Bloom's and Werner's syndrome helicases (Li *et al.*, 2001).

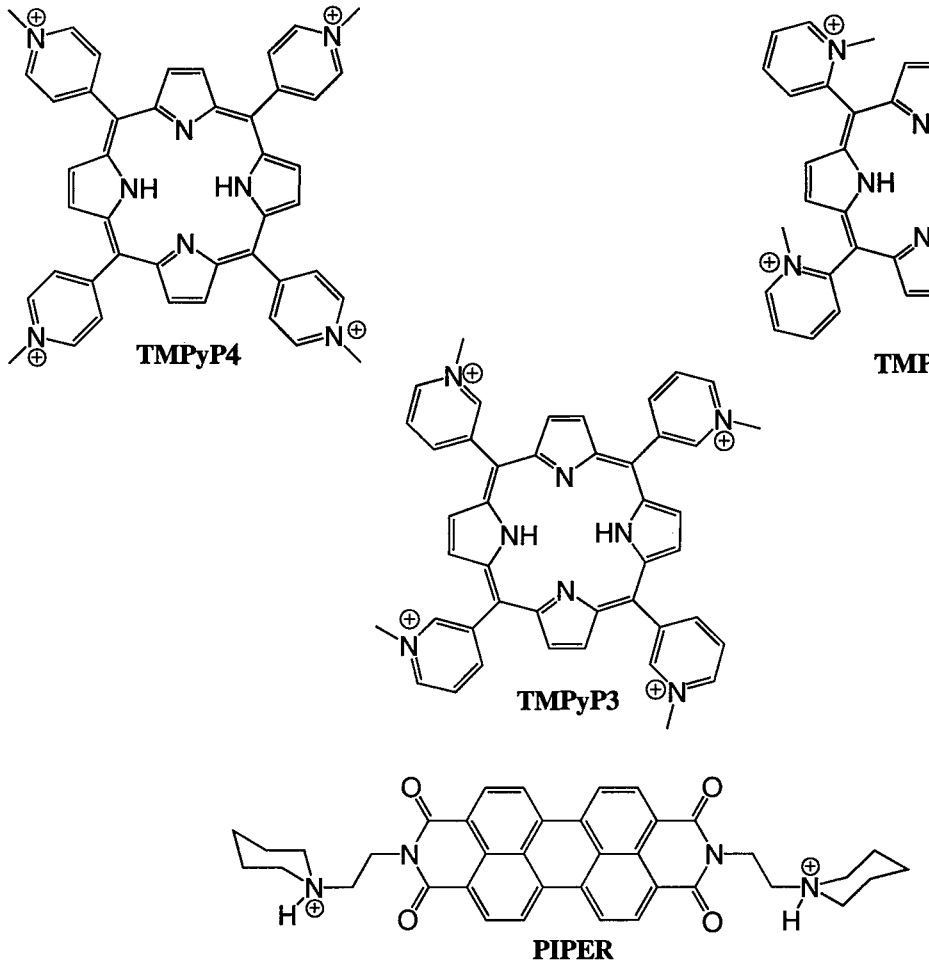


Figure 1.8. Structures of the the three isomeric cationic porphyrins, TMPyP4 TMPyP3 and TMPyP2; and the perylene, PIPER.

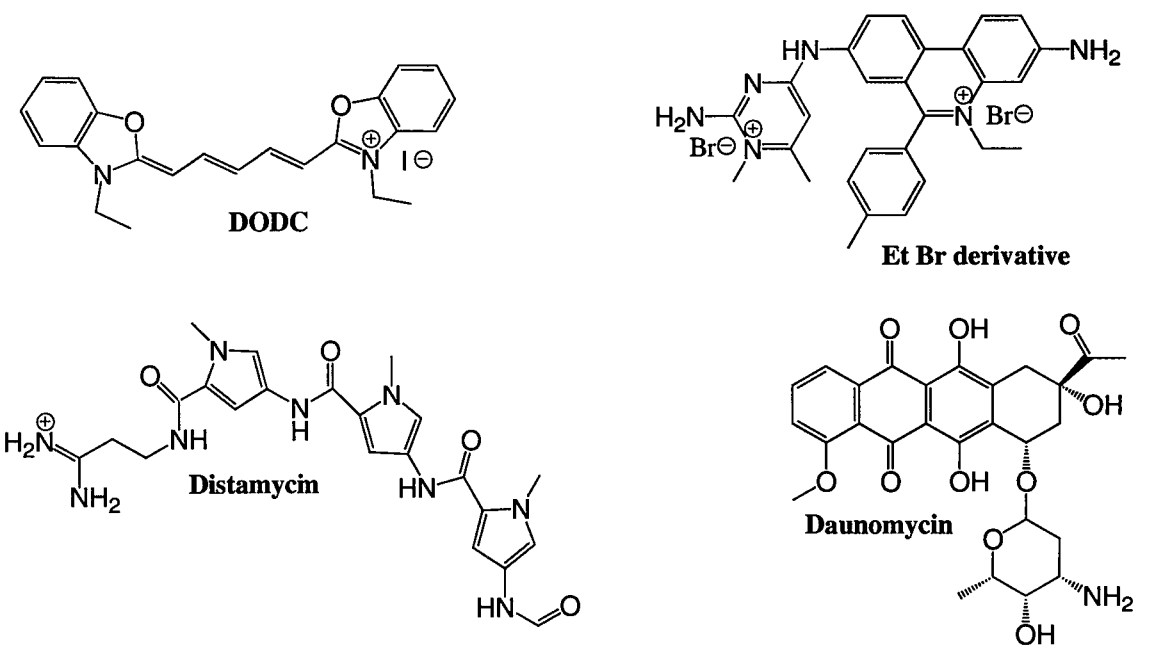


Figure 1.9. Structures of the carbocyanine derivative, DODC; an ethidium bromide derivative; distamycin; and daunomycin.

1.15.3 Perylenes (Figure 1.8)

PIPER, a perylene derivative, was designed to bind strongly and specifically to quadruplex DNA. NMR data were consistent with the formation of 1:1 complexes between a tetrameric quadruplex and PIPER. External stacking of the ligand to G-tetrad was also determined, and PIPER was shown to be either sandwiched between the terminal G-tetrads of two tail-to-tail quadruplexes that lacked TTA-3' tails, or intercalate at the GpT step of longer sequences (Fedoroff *et al.*, 1998). Photocleavge by a PIPER-Fe-EDTA conjugate at 3' G-tetrad also confirms end-stacking (Tuntiwechapikul & Salazar, 2001).

The formation of an intermolecular quadruplex was accelerated 100-fold by 10 μM PIPER, providing the first evidence of a small molecule driving polynucleotide assembly (Han *et al.*, 1999b). This ability was later extended to duplex targets, where PIPER promoted the transition of duplex telomeric sequences and the NHE sequence of the *c-myc* promoter into quadruplex structures (Rangan *et al.*, 2001). Further studies have shown that PIPER specifically interferes with the unwinding of quadruplex but not duplex DNA by yeast *Sgs1* helicase, a member of the RecQ DNA helicase family. Competition experiments eliminated the possibility of interaction with the enzyme rather than the quadruplex (Han *et al.*, 2000). *Sgs1p* may be relevant to telomere maintenance that is associated with telomerase independent elongation.

1.15.4 Known duplex interactive agents (Figure 1.9)

Several duplex interactive agents have also been found to interact with quadruplex DNA. An early report suggested this was the case for the dye, **ethidium bromide** (EtBr) (Guo *et al.*, 1992). EtBr associated tightly to the tetrameric quadruplex $[d(T_4G_4)]_4$, enhancing its thermal stability, and shifted the equilibrium between single-strands and tetramer in favour of tetramer. Chemical cleavage patterns suggested that EtBr selectively interacted at the end of the G-region. The spectroscopic characteristics of EtBr binding to quadruplex and duplex DNA were similar, consistent with an intercalatory mechanism (suggesting some degree of insertion of the phenanthridine ring between 5'-distal Gs). The evidence suggested that one EtBr was bound per tetramer with an affinity comparable to that for normal duplex intercalation (Guo *et al.*, 1992). A series of EtBr derivatives have been described that stabilise intramolecular G-quadruplexes with anti-telomerase activities ($^{tel}IC_{50}$) of 18 to 100 nM. Under quadruplex-favouring conditions, addition of G4-DNA

produced bathochromic shifts in their emission spectra that were not evident with an oligomer that was not able to form quadruplex or duplex DNA, indicating interaction only with the folded form. The enhanced fluorescence of these compounds indicated a K_d of $9x10^{-8}$ M. However, EtBr had little effect upon the quadruplex melting profiles, though its derivatives gave ΔT_m of 7.9 to 10.6°C (Koeppel *et al.*, 2001). FRET between the DNA bases and EtBr supports stacking as the mode of binding.

Chen and co-workers have described the **carbocyanine**, DODC, which exhibited unique spectroscopic signatures for interaction with dimeric quadruplexes. These included the appearance of a new absorbance peak at the λ_{max} of 534 nm, an induced circular dichroism, a quenching of dye fluorescence and a strong energy transfer from DNA. As such, the molecule could be used as a quadruplex probe *in vivo* or *in vitro* (Chen *et al.*, 1996). These properties are thought to be due to a different mode of G2 quadruplex binding by DODC, which may involve groove-binding with possible intercalation of one group in the TTA loop.

The 1:1 binding of DODC to quadruplex DNA displayed a K_d of 1.1×10^{-7} M, which is 5-fold higher than its affinity for duplex DNA. However, it preferentially bound to triplex DNA. A series of five DODC derivatives were therefore examined, from which only DTC inhibited human telomerase, reaching 35 % at 50 μM. Scatchard analysis of its visible spectrum gave a K_d of 2.2×10^{-6} M and a n of 2 (Kerwin *et al.*, 2001). DTC also had improved the selectivity for quadruplex over triplex DNA measured by equilibrium dialysis. Recently, a heterocycle-peptide conjugate has been designed (Schouten *et al.*, 2003) where the heterocyclic core was derived from DODC and the best tetrapeptide was that selected (by a combinatorial technique) for optimal quadruplex affinity. This produced a compound with a K_d of 1.5×10^{-5} M, which was a 40-fold discrimination over duplex. This bound with a 1:1 stoichiometry, via loop and groove interactions, but without π-stacking to the bases.

A recent NMR study has reported the interaction of the groove binding agents **distamycin A** and **netropsin** with the quadruplex $[d(TG_4T)]_4$ (Randazzo *et al.*, 2001). The results showed that two distamycin A molecules simultaneously bound to the same groove, and that each dimeric pair was in close contact with another pair in an antiparallel orientation.

Initially, a 2:1 complex is formed with one pair binding in one groove. At higher drug:DNA ratios, a second distamycin A dimer tightly bound giving a 4:1 complex. This behaviour is explained by suggesting that binding of the second ligand is more favourable due to a cooperative process. However, a fully saturated complex was not seen, possibly as the binding of the first dimer expanded the groove, narrowing the two adjacent grooves and so preventing interaction with further dimers. The netropsin – quadruplex complex consisted of one molecule per groove, as the doubly charged ends of netropsin caused repulsion that prevented a side-by-side arrangement. Contrary to these findings, other NMR study has suggested binding of distamycin by stacking on the terminal G-quartet, contacting the flanking bases, and inhibiting specific protein interactions (Cocco *et al.*, 2003).

The first structure for a drug bound to a parallel quadruplex was recently reported for the anticancer drug, **daunomycin**. No drug molecules intercalated within the guanine core, but two layers of three coplanar molecules were formed that were sandwiched between two quadruplexes, tightly stacking to end of quadruplex. This produced a ligands:quadruplex ratio of 6:2. Each layer interacted within three grooves (Clark *et al.*, 2003).

1.15.5 New generation of G4 ligands (Figure 1.10)

A novel series, the **pentacyclic acridines** (termed RHPS4), shown in Figure 1.10A, which have high affinity for tetraplex and triplex structures compared to duplex and single-stranded DNA, as assessed by differential equilibrium dialysis. RHPS4 possesses a promising $^{tel}IC_{50}$ of 0.3 μM, far less than its cytotoxicity of IC_{50} at 7 μM. A concentration of 200 nM was sufficiently potent for antiproliferation in several human tumour cells, with the exception of a cell line with substantial telomeres (Gowan *et al.*, 2001). NMR has provided structural details of the RHPS4-quadruplex (Gavathiotis *et al.*, 2001). Titration of RHPS4 resulted in some line broadening attributed to exchange of the drug between a number of possible bound conformations and/or a number of similar affinity sites. The ApG step was the primary intercalation site. The acridine 13-N seemed to act as a pseudo-cation in the pocket of a low energy complex.

The **dibenzophenanthrolines** are another family of quadruplex-binding ligands (see Figure 1.10B), that have been selected by a FRET melting assay (Mergny *et al.*, 2001).

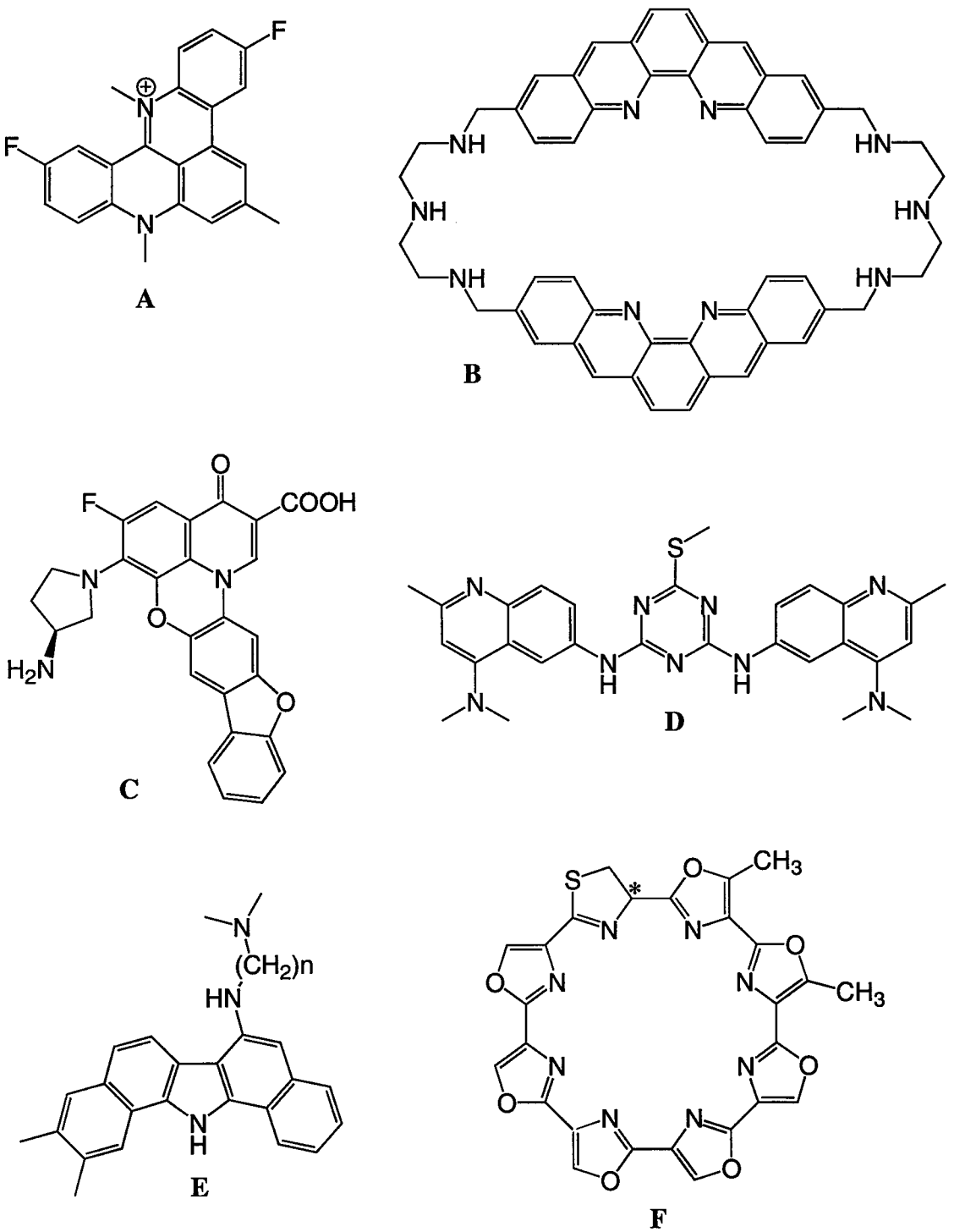


Figure 1.10. Structures of a number of new G4 ligands: (A) a pentacyclic acridine, (B) a dibenzophenanthroline dimer, (C) a fluoroquinolone, (D) a triazine, (E) a benzoindoloquinoline, and (F) the natural compound telomestatin (asterisk denotes chiral centre).

These derivatives are a family of pentacyclic crescent-shaped planar molecules. They produced ΔT_m values that are much greater than those previously reported for other G4 ligands, that range from 2 to 20°C at 1 μM concentrations, and K_d values were measured in the submicromolar range. The most active member possessed a $^{tel}IC_{50}$ of 28 nM (measured by TRAP assay). FRET measurements of DNA melting have shown that ligand induced stabilisation correlated well with increased inhibition of telomerase activity. The authors were unsure whether these properties were based on intercalation, end-stacking or loop and groove interactions of the drugs. Recently, BOQ1 has been described which consists of two dibenzophenanthroline subunits (Teulade-Fichou *et al.*, 2003). It has a high affinity for quadruplex DNA ($K_d < 1 \times 10^{-7}$ M) which the authors suggested was achieved by binding of one drug molecule at each end of the quadruplex. Thermal denaturation studies with this ligand gave a large ΔT_m of 28°C, while a bandshift assay indicated that it efficiently induced quadruplex formation, and a $^{tel}IC_{50}$ of 0.13 μM was produced.

The **fluoroquinolones** (see Figure 1.10C) are well-known anti-bacterial inhibitors of DNA gyrase, which possess weak anti-telomerase activities. The drug A-62176 had a $^{tel}IC_{50}$ of 58 μM, and was chosen as a lead compound to develop better analogues. The derivative QQ58 produced an improved $^{tel}IC_{50}$ of 28 μM, and NMR showed that similar to PIPER – this ligand was bound to the terminal G-tetrads of tetrameric or intramolecular quadruplexes. Further modifications gave the two derivatives, QQ27 and QQ28. Modelling predicted their binding energies for tetraplex which was reduced for QQ27 and even worse for QQ28. This correlated with their activities determined by a *Taq* stop assay; their cytotoxicities in cancer cell lines followed QQ58<QQ27<QQ28; and their antiproliferative effects *in vivo* at 1 μM on developing sea urchin embryos (Duan *et al.*, 2001).

The promising quinoline-based G-quadruplex ligands termed the **triazines** (see Figure 1.10D) were initially identified by FRET melting and were predicted to end-stack to the terminal G-quartet. At 1 μM the best member gave a ΔT_m of 20°C. The TRAP assay gave very effective $^{tel}IC_{50}$ of 41 to 130 nM. Long term treatment of tumour cell lines induced growth arrest associated with telomere erosion measured by FISH (fluorescence *in situ* hybridization) where 9 to 25% of chromosomes showed fluorescent telomeric spots. In another tumor cell line an anaphase bridge was induced leading to cell senescence (also observed for TMPyP4). Equilibrium dialysis also confirmed all members had quadruplex

specificity (Riou *et al.*, 2002).

Benzoindoloquinolines (Figure 1.10E) have also been examined as quadruplex-binding ligands. A dimethylamino-sidechain, capable of protonation was predicted to give good stabilisation. Initial results verified this concept, and this series of compounds increased the T_m of a foldback quadruplex by 3 to 11°C (Alberti *et al.*, 2002), and quadruplex binding was also confirmed by spectroscopy. Equilibrium dialysis experiments identified one candidate drug that accumulated (and thus preferred) quadruplex above any other nucleic acid forms. External stacking was proposed but other modes such as groove or loop binding were not excluded.

The natural product, **telomestatin** (Figure 1.10F) was isolated from *Streptomyces anulatus*. Telomerase inhibition, estimated using the TRAP assay gave a $^{tel}IC_{50}$ of 5 nM, significantly greater than all other classes of quadruplex stabilisers to-date, and it showed a several thousand fold specificity for telomerase inhibition over the other DNA polymerases and reverse transcriptases (Shin-ya *et al.*, 2001). This is presumably due to its high structural similarity to a G-tetrad. Incubation of a 29-mer quadruplex with 0 to 30 μM telomestatin at 20°C for 30 min produced a modest stabilisation as shown by bandshift experiments (Kim *et al.*, 2002). Computational chemistry has supported a model where telomestatin binds in a 2:1 ratio by end-stacking. Simulations also indicated that the telomestatin molecule may reorientate in the quadruplex structure, and that it has a higher selectivity for intramolecular rather than intermolecular quadruplexes.

1.16 Aims of quadruplex research

Chapters 3-5 of this thesis examine the interaction of a range of related putative quadruplex binding ligands with inter- and intramolecular quadruplex-forming oligonucleotides. These studies used a fluorescence melting assay and bandshift analysis to explore the features that produce good quadruplex binding in a series of substituted anthraquinones. These studies were extended to examine the formation of intramolecular quadruplexes based on the human and *Oxytricha* telomeric repeats.

1.17 Triplexes

The discovery of triple helical DNA was first made in 1957 by Felsenfeld *et al.* from UV melting experiments on synthetic polynucleotides in which 2:1 mixtures of polyU and polyA were found to form a specific three stranded structure (Felsenfeld & Davies, 1957; Felsenfeld *et al.*, 1957). At the time this was considered little more than a laboratory curiosity and received little interest for many years (see Figure 1.11). It was not until the late 1980s that renewed interest in such structures was inspired by the work of Hélène & co-workers and Moser & Dervan, who independently demonstrated that the formation of a local triple helix by an oligonucleotide at an oligopurine tract could specifically inhibit transcription of particular gene sequences (Le Doen *et al.*, 1987; Moser & Dervan, 1987). It was realised that oligonucleotide-directed triple helix formation presented a highly specific approach to duplex DNA recognition and artificial gene-repressors could be designed that were of value as antiviral or anticancer therapies. This was termed the antigene approach.

1.18 Triple helical structure

Triple helix formation arises from the binding of a synthetic oligonucleotide in the major groove of duplex DNA. This forms specific hydrogen bonds with the purine bases of the duplex. Since the original studies on triple-stranded nucleic acids, two main classes of triple helix have been characterised. It is now well accepted that pyrimidine-rich third strands bind in a parallel orientation with respect to the duplex purine rich strand and include T·AT, C^+·GC and G·GC triplets, whereas purine-rich oligonucleotides bind antiparallel and include G·GC, A·AT and T·AT (see Figure 1.12).

All triplets involve the formation of two hydrogen bonds with the purine of the AT and GC base pairs. For cytosine to achieve this with GC, it must be protonated at its N3 position. The hydrogen bonding between the third strand base and the purine strand of parallel triplexes corresponds to the scheme described by Hoogsteen (1959). The pattern in antiparallel triplex represents an 180^o rotation of the third strand bases, and is termed reverse-Hoogsteen hydrogen bonding. G C, and T can all interact in either orientation (though C disfavours an antiparallel orientation), while A can only form reverse Hoogsteen interactions. The details of triplex formation and its structure and sequence dependency are

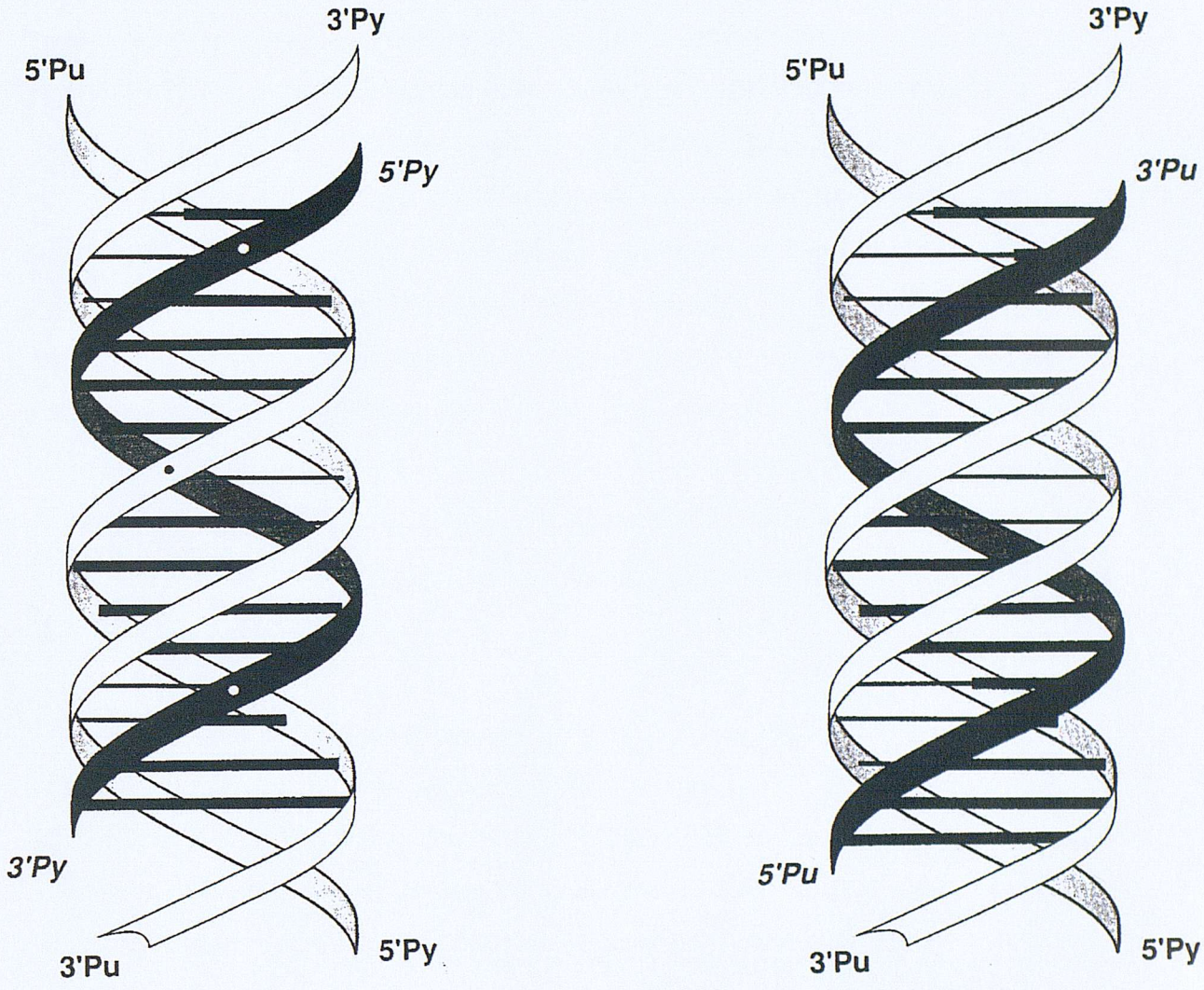


Figure 1.11. Ribbon models of a parallel (left) and antiparallel (right) intermolecular triplexes. The black strand represents the third strand located in the major groove.

considered in section 1.22. First, I will outline the uses of triplexes which have so attracted research.

1.19 Triplex applications

There is a high frequency of polypurine tracts in all eukaryotic genomes, and particularly those found in upstream promoter regions. These can be targeted by the **antigene approach**. The primary goal of triple helix technology is its therapeutic applications. An ever-increasing number of reports support the potential of triplex formation for modifying levels of target gene expression, while illustrating the varied possible applications of the antigene approach. Several possible mechanisms are proposed to explain triplex-mediated transcription (and replication) inhibition. Triplex formation at a promoter can block the binding of various transcription factors (plus replication machinery). Prevention of transcription elongation (whether arrest is adjacent or further downstream) is also possible. A further possibility is that triplex formation causes a conformational change that disrupts transcription factor binding. It is noteworthy that interfering with transcription factors that down- rather than upregulate, could enhance rather than suppress transcription of some gene targets.

The triplex-mediated antigene approach has been used to target a variety of promoters *in vitro* from human *c-myc* (Cooney *et al.*, 1988), human Ha-*ras* (Mayfield *et al.*, 1994), dihydrofolate reductase (Blume *et al.*, 1992), murine c-*pim*-1 (Svinarchuk *et al.*, 1994), rat *HER2/neu* (Gee *et al.*, 1994), human EGF and murine insulin receptor (Durland *et al.*, 1991). Examples of the targeted transcriptional inhibition of plasmid-borne promoters reported include rat α1(I) collagen (Kovacs *et al.*, 1996), rat androgen receptor (Song *et al.*, 1995), a progesterone-responsive gene (Ing *et al.*, 1993), an interferon inducible gene (Roy 1994), HIV, and the interleukin-2 receptor (Grigoriev *et al.*, 1992). Endogenous genes that have been targeted include aldehyde dehydrogenase (Tu *et al.*, 1995), interleukin-2 receptor (Orson *et al.*, 1991), *c-myc* (Postel *et al.*, 1991; Thomas *et al.*, 1995), and the human *HER2/neu* gene (Porumb *et al.*, 1996). Most examples target plasmid-borne genes rather than endogenous genes, reflecting the current limitations of oligonucleotide delivery. However, these experiments which involve pre-formed triplex prior to transformation are of limited relevance. Non-specific effects also need to be ruled out. Good controls are those where the same triplex forming oligonucleotides (TFOs) are used upon scrambled

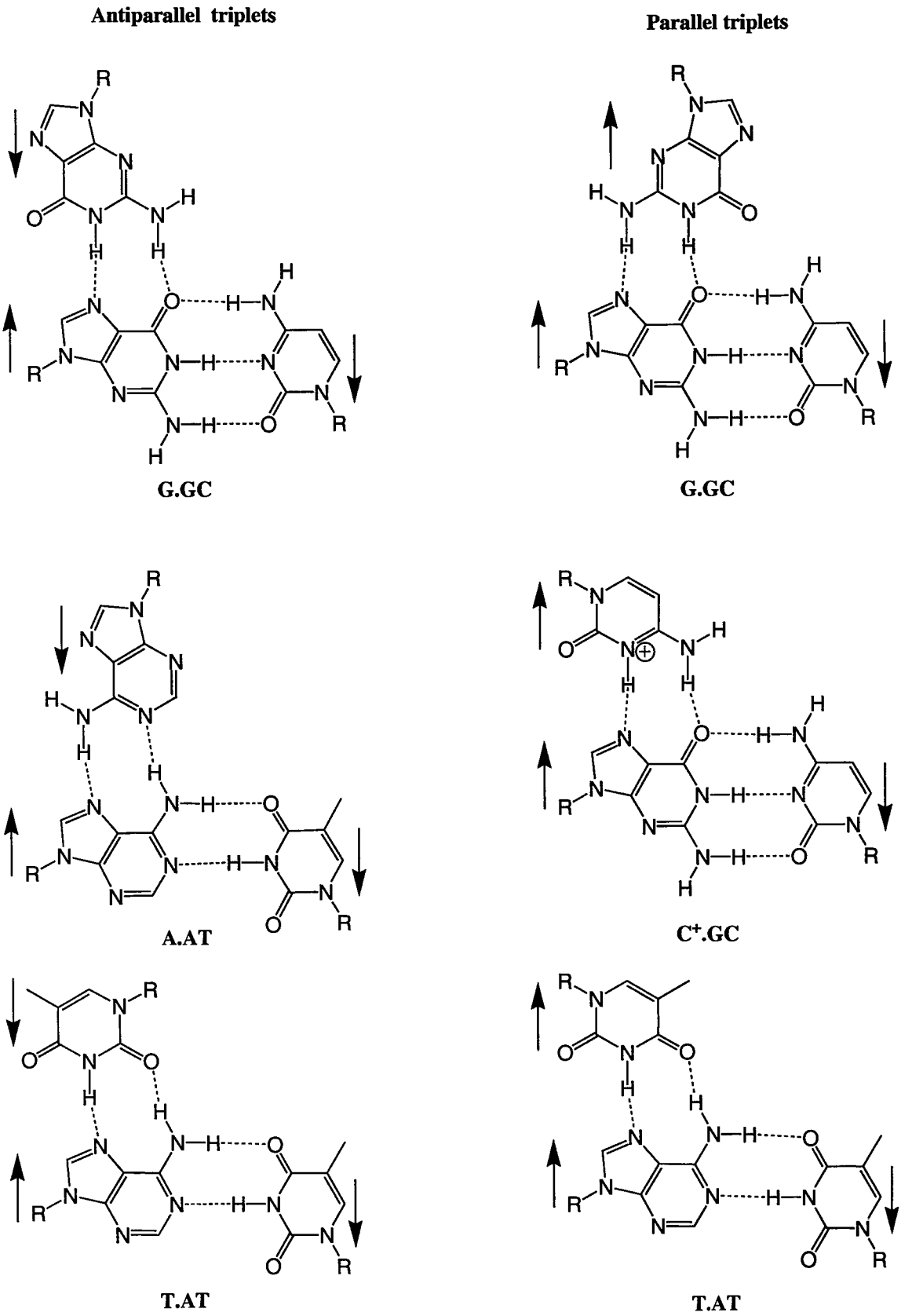


Figure 1.12. Structures of the antiparallel (left) and parallel triplets (right).

target sites (that do not affect gene expression) in a control gene, rather than simply using scrambled TFOs.

Triplex formation by TFO conjugates can be used as **site-specific cleavage agents** that could be potentially useful in antiviral or anticancer therapy (Moser & Dervan, 1987; Perroualt *et al.*, 1990), and TFO-nuclease conjugates have also been developed (Landgraf *et al.*, 1994; Pei *et al.*, 1990). Triplex formation also permits tethered mutagens to be placed next to specific targets, allowing **site-directed mutagenisis** (Takasugi *et al.*, 1991; Wang *et al.*, 1995). and recombination (Faruqi *et al.*, 1996; Wang *et al.*, 1996;Vasquez *et al.*, 2000). Thus triplex-induced mutagenesis can repair point mutations or knockout gene function. **Oligonucleotide clamps** can specifically bind ssDNA and ssRNA targets by wrapping around them to form both Watson-Crick and Hoogsteen interactions (Giovannangeli *et al.*, 1993; Knudsen & Nielsen, 1996). Triplex-mediated **DNA purification** has been shown to be amenable to isolation and enrichment of (very large) DNAs from heterogeneous mixtures (Ito *et al.*, 1992; Nishikawa *et al.*, 1995).

1.20 Advantages of triplex targeting

Using a DNA reading agent that is itself a nucleic acid is inherently advantageous. Any reading agent needs to interact with the base constituents of the DNA that are helically arranged. Flexibility is needed in the reading molecule to follow this pattern around the helix and nucleic acids themselves are ideally suited to achieve this. Furthermore, using a polymer as a reading agent allows versatility in the range of recognition targets. The selectivity exhibited by nucleic acids is such, that it far exceeds that of any other DNA ligand available, and is sufficiently high to recognise a single sequence in the double helical DNA within a living cell. This is essential, as genome size varies from a few million bps (4.5×10^6 in *E.coli*) to several billion in eukaryotes (humans ~3×10^9, which is not the largest). Viruses have smaller genomes (of single- or double-stranded RNA or DNA), HIV ssRNA genome consists of 9200 nts. This makes it highly probable that an 11 bp sequence in *E.coli* or 17 bp in humans will be the minimum required to be unique. Lastly, only two copies (two alleles) of target are present per cell, so only low levels of oligonucleotide of the appropriate sequence are theoretically needed to down regulate transcription and produce a long lasting effect. In the analogous antisense strategy, which targets the many thousands of mRNA molecules, sustained levels of oligonucleotide are

needed as the mRNA pool is continually replenished. Though triple helix gene targeting is essentially a simple idea that has attracted much attention, several outstanding problems make this difficult to achieve in practice.

1.21 Biological considerations

Oligonucleotides are relatively poorly taken up by cells (the extent and mechanism are cell dependant). They also tend to be concentrated in particular regions (a possible advantage) necessitating adequate compartmentalisation (Clarenc *et al.*, 1993). Rapid nuclease degradation also represents a problem. Another difficulty is target accessibility within the context of the chromatin, though a targeted genetic sequence is likely to be biologically active and as such might be open to interact with protein factors and enzymes. Finally, third strand association from the duplex is invariably weak necessitating a high local concentration at the target sequence within the cell.

1.22 Structural influences on triplex formation

The formation of triplex is influenced by a number of major factors, these range from the need for cytosine protonation, limited sequence recognition that excludes pyrimidine inversions, TFO length, secondary and self associated structures, sequence composition and stability.

1.22.1 Cytosine protonation

The need for cytosine protonation means that the formation of C^+·GC triplet is pH-dependent, and parallel triple helices are generally stable at pH < 5.5 and unstable at physiological pH. However, the T_m of triplex to duplex transitions increases with C^+·GC content, suggesting that the free energy for C^+·GC formation is greater than for T·AT (Roberts & Crothers, 1996). Quantitative DNase I cleavage protection has also shown that the dissociation constants for several TFOs at pH 5 increase with C content (Keppler & Fox, 1997). As both triplets are isohelical and are held in place by two hydrogen bonds, these observations may reflect the contribution that protonation makes to stabilisation. It is likely this improved stability represents the electrostatic interactions of the positive charge overcoming the repulsion of the three negative backbones.

Although triplex stability increases with increasing C^+·GC content, at low pH, runs of contiguous C^+·GC triplets are less stable (Kiessling *et al.*,1992), as a result of anticooperative effects on the pK_a of adjacent cytosines. The most stable triplexes at low pH consist of alternating $(CT)_n$ (Roberts & Crothers, 1996). The pK_a of the third strand cytosines is also influenced by their position and terminal cytosines are harder to protonate (Asensio *et al.*, 1998; Leitner *et al.*, 2000) though isolated internal cytosines can have pK_a as high as 9.

Potential solutions

A number of possible base analogues have been synthesised in oligonucleotides to try to overcome the above problem. These may have higher pK_a allowing easier protonation under less acidic conditions, an early example is 5-methyl C (Lee *et al.*, 1984). Other analogues do not require protonation to form two hydrogen bonds, such as pseudo isoC (Ono *et al.*, 1992); and 6-oxoC (Xiang *et al.*, 1994). However these can suffer from decreased triplex stability resulting from the loss of the beneficial positive charge that reduces strand repulsion. Other similarly rationalised purine analogues have been used in place of C including pyrazine (von Krosigk *et al.*, 1995), 2-aminopyridine (Bates *et al.*, 1996), 8-oxo-2-deoxyA (Krawczyk *et al.*, 1992), N7-G (Brunar & Dervan, 1996), and P1 (Koh & Dervan, 1992). These purine analogues can cause distortion of the TFO backbone that also reduces triplex stability. The structures of these compounds are shown in Figure 1.13.

1.22.2 Other triplets at pyrimidine interruptions

Triple helices are most effective against perfect homopurine·homopyrimidine sequences and though these tracts are over-represented in the genome, a universal recognition code is needed. Pyrimidine bases only present one hydrogen bonding site within the major groove, in contrast to the two provided by purines. Thus recognition of TA and CG is inherently weaker. Several natural base combinations have been used to target pyrimidine interruptions and their structures are shown in Figure 1.14.

G·TA triplet

TA inversion can only be stably recognised by G·TA, which is only tolerated in the parallel motif. TA in antiparallel complexes are therefore more difficult to accommodate. Any

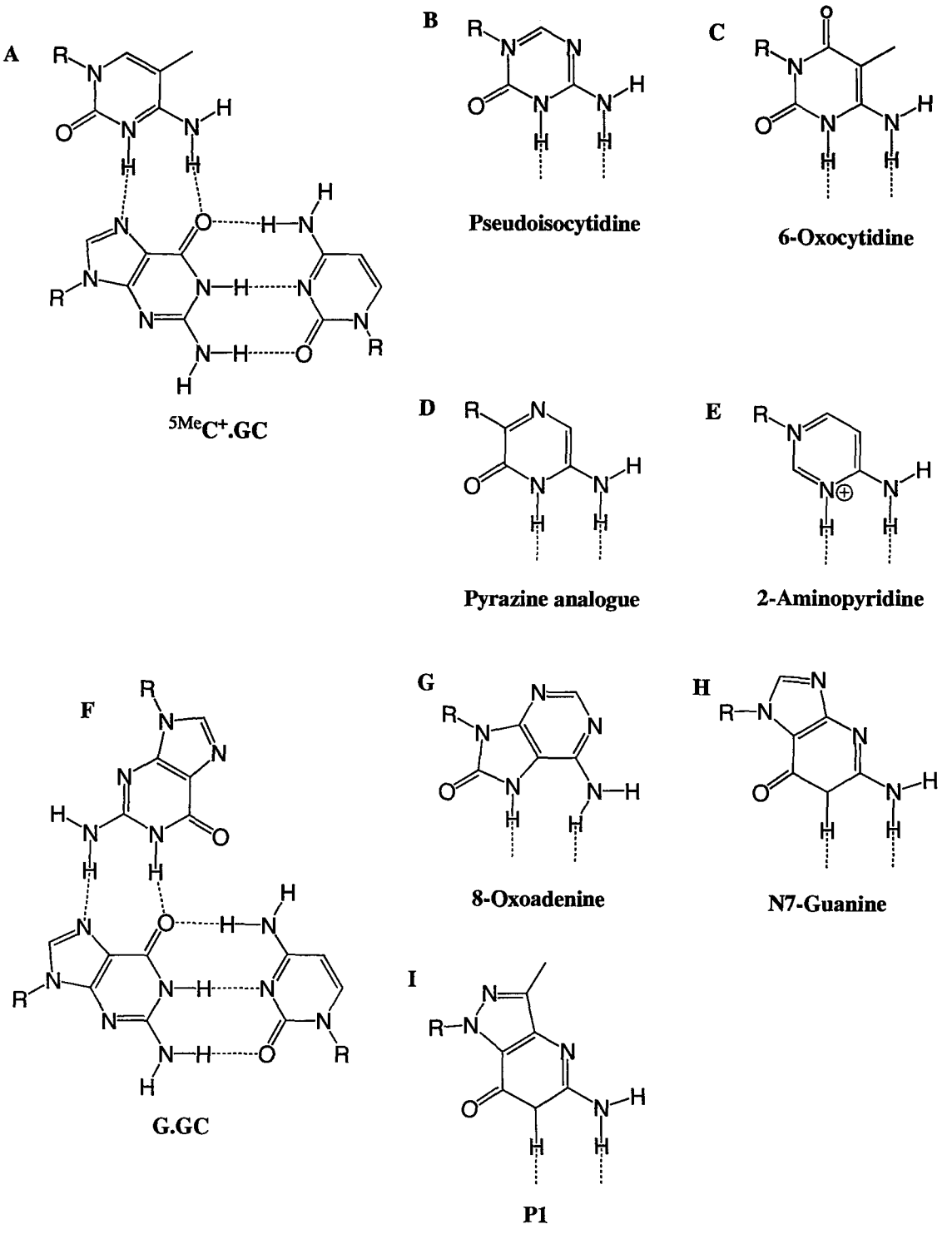


Figure 1.13. Base analogues for the pH independent recognition of GC base pairs in a parallel triplex. (A) the $^{5Me}C^{+}$·GC triplet, (B) pseudoisocytosine, (C) 6-oxocytosine, (D) a pyrazine analogue, (E) 2-aminopyridine, (F) the parallel G·GC triplet, (G)8-oxoadenine, (H)N7-guanine, and (I)P1.

other mismatches cause a large decrease in stability. DNaseI footprinting provided the first evidence that G·TA could be recognised within a parallel triplex (Griffin & Dervan, 1989). Of all the natural bases used to target TA within an otherwise exclusively T·AT triplex, G·TA produced the most specific and stable recognition as measured by footprinting studies, while a weaker complex was possible with a T·TA triplet. (Chandler & Fox, 1993). Melting studies have also shown that G·TA produces the least destabilised triplex (Mergny *et al.*, 1991; Fossella *et al.*, 1993).

Three adjacent G·TA were possible within a parallel triplex of C^+·GC if a triplex ligand was used (Gowers & Fox, 1997). Another footprinting study showed that though a triplex containing alternating G·TA and T·AT triplets was unable to form, tethering an adjacent block of T·AT triplets allowed triplex formation (Chandler & Fox, 1995). Further studies showed that these complexes can be extended to longer $(AT)_n$ tracts with shorter T·AT tracts, and that the stability can be enhanced by including C^+·CG triplets in the anchoring tail (Gowers & Fox, 1998). This is consistent with the observation that C^+·CG triplets impart greater stability than T·AT triplets.

Other DNaseI footprinting studies showed that the G·TA triplet was more stable when flanked by two T·AT triplets rather than 5'-C^+·GC (Kiessling *et al.*, 1992). It was suggested that this might reflect an additional hydrogen bond formed from the other amino proton of guanine to the adjacent O4 group of the duplex thymine in the next T·AT (Radakrishnan & Patel, 1994a). The closest proton of the guanine N3 is favoured for the single hydrogen bond as this optimises stacking and minimises backbone distortion. A major factor influencing G·TA stability is the presence of guanine in an otherwise pyrimidine TFO that produces local distortion in the backbone. Moving the guanine out of the average plane of the TA target avoids steric clash with the 5-methyl of thymine (Radhakrishnan & Patel, 1994a). TpG steps reduce base stacking as they undergo overwinding, whereas the reverse occurs at GpT steps (Radhakrishnan & Patel, 1994b).

T·CG triplet

T·CG (and C·CG) are the best natural base combinations for recognising CG interruptions, and can participate in both parallel and antiparallel triplexes. The stable formation of T·CG in a parallel triplex was first proposed by Yoon *et al.*, (1992) using bandshift analysis. This

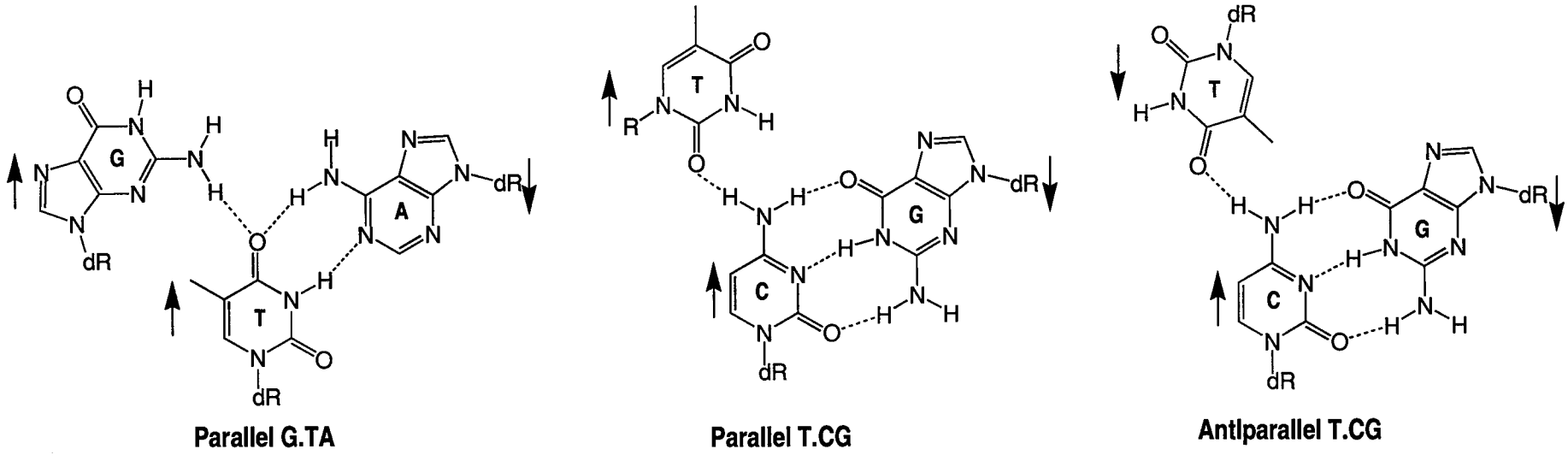


Figure 1.14. Structures of non-canonical mismatched triplets.

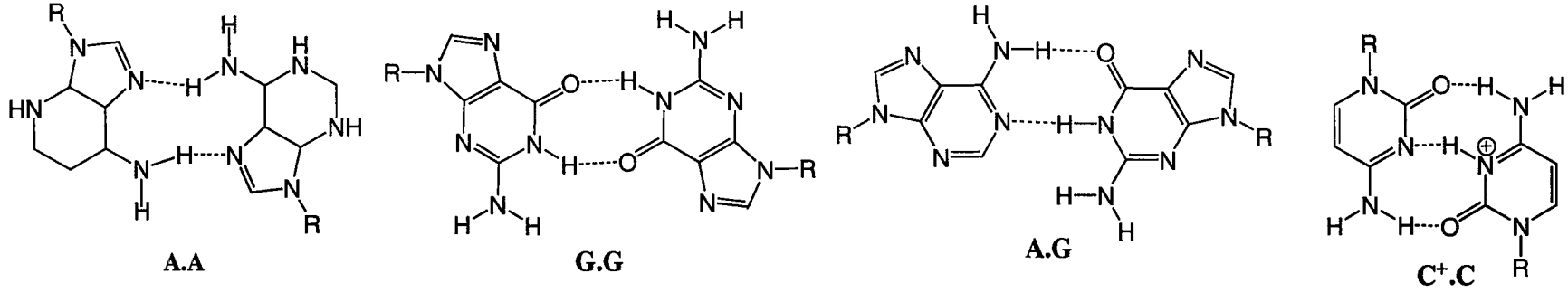


Figure 1.15 Structures of non-canonical base pairs that give rise to self-associated structures.

agrees with footprinting experiments that showed the T·CG triplet (followed by the C·CG triplet) gave the most stable recognition at a CG interruption within a parallel triplex (Chandler & Fox, 1993). As with G·TA, up to three T·CG were tolerated in a triplex if a ligand was added (Gowers & Fox, 1997). A single CG inversion was most stably targeted by T·CG, though C·CG and A·CG were also well accommodated in both a GT- and GA-TFO (Chandler & Fox, 1996).

T·CG and the similar hydrogen bonding pattern of C·CG, are less destabilising at CG than purines that would encounter steric clash with the exocyclic amino of C (Mergny *et al.*, 1991). Structurally, the alignments of both T·C in T·CG, and G·T in G·TA in parallel triplex are similar, this presumably produces the same over- and under-twisting that effects the base stacking (Radhakrishnan & Patel, 1994b).

Location

Interestingly, a mismatch near the centre of a triplex might logically be thought to be more destabilising than near the ends, as it would result in shorter contiguous tracts. However, mismatches near the ends of a 12-mer triplex were found to be more destabilising, particularly when at the 3' end, this did not reflect intrinsic binding as the same TFOs could bind with high affinities to their correct duplex target (Cheng & Van Dyke, 1994). It was speculated that the results represented an asymmetry in GT-motif triplex formation that must proceed along the TFO from 3' to 5'. The REPSA selection protocol described by Hardenbol and Van Dyke (1996) (discussed in detail in chapter 6) produced stable triplex at a 19-mer target site, with more exact matches for the 13-mers at the 5'-end, while more mismatches were observed at the 5'-end. Triplex formation by a $(AGG)_4$ TFO was shown to be 6-times greater than for a $(GGA)_4$ TFO for their respective targets (though each would contain an identical number of each triplet) implying a 3'-guanine was superior to a 3'-adenine (Arimondo *et al.*, 1998). The greater effect of 3'-end mismatches was also seen in footprinting experiments with $(GGA)_n$ for which slippage was observed (Chandler & Fox, 1996).

Potential solutions

Various base analogues have also been designed to recognise pyrimidine inversions. The benzoimidazole, D3 was able to bind to all four base pairs by intercalation (Sasaki *et al.*,

1995), 2-deoxynebularine (Stilz & Dervan, 1993), and a dihydropyrazine (Eldrup *et al.*, 1997) were both tolerated at pyrimidine inversions but can only form a single hydrogen bond, while hypoxanthine can recognise UA (Marfut *et al.*, 1998). Alternatively, a range of analogues have been studied which either interact with both bases or stretch across the major groove to bind to the purine. These include 2-deoxyformycin (Rao *et al.*, 1994), the β-anomer of guanidinocytidine (Doronina & Behr, 1997), a N4-C derivative (Huang *et al.*, 1996), a naphthyl-imidazole (Zimmerman *et al.*, 1995), and L1 (Lehmann *et al.*, 1997), these are presented in Figures 1.16 and 1.17. Triplex ligands (for review see Fox, 2000) have also been shown to sufficiently stabilise triplex containing pyrimidine inversions to permit their formation (Gowers & Fox, 1996).

1.22.3 TFO length

Almost every triplex studied shows increasing binding affinity with length, while increased length also increases specificity. However, one study using GT-oligonucleotides has uniquely proposed that optimal antiparallel triplex formation was achieved by a short 12-mer (Cheng & Van Dyke, 1994). It was also noted that additional bases at the 5' end of the TFOs were less stable than if they were added at the 3'end. A possible explanation is that the runs of purines produced stronger stacking interactions with adjacent purines, which placed the TFO in a conformation unfavourable to reverse-Hoogsteen triplex formation. Loss of base stacking may not be fully compensated for by the additional hydrogen bonding of guanine, making long stretches less favourable. Shorter sequences also seemed to form stronger triplexes within a population of TFOs selected by REPSA for high affinity triplexes (Hardenbol & Van Dyke, 1996).

1.22.4 Secondary & self-associated structures

The formation of triplexes can be severely impaired if the third strand oligonucleotide forms competing self-associated DNA structures. G-rich triplexes are inhibited by physiological concentrations of monovalent cations that increase triplex dissociation and sequester the TFOs into stable quadruplex structures comprised of G-quartets (Olivas & Maher, 1995a). Unusually, triplex formation with GA-TFO was observed to increase at higher temperatures, which is in contrast to the usual situation where affinity decreases with temperature. This triplex formation was insensitive to the presence of monovalent cations. This was shown to be due to the formation of competing intermolecular

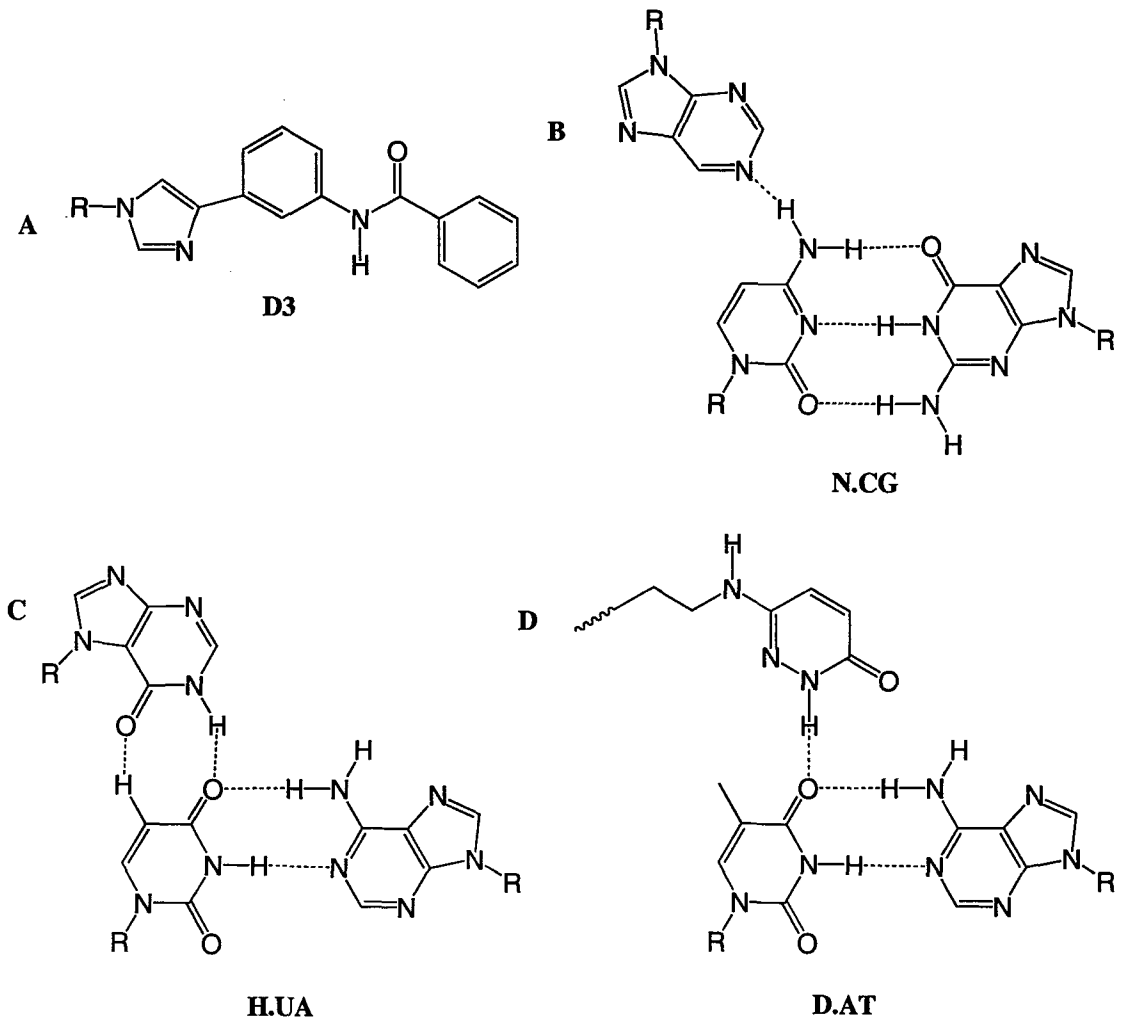


Figure 1.16. Base analogues that extend base pair recognition to pyrimidine inversions, (A) a functionalised benzoimidazole , (B) 2-deoxynebularine, (C) hypoxanthine, and (D) a dihydropyrazine derivative.

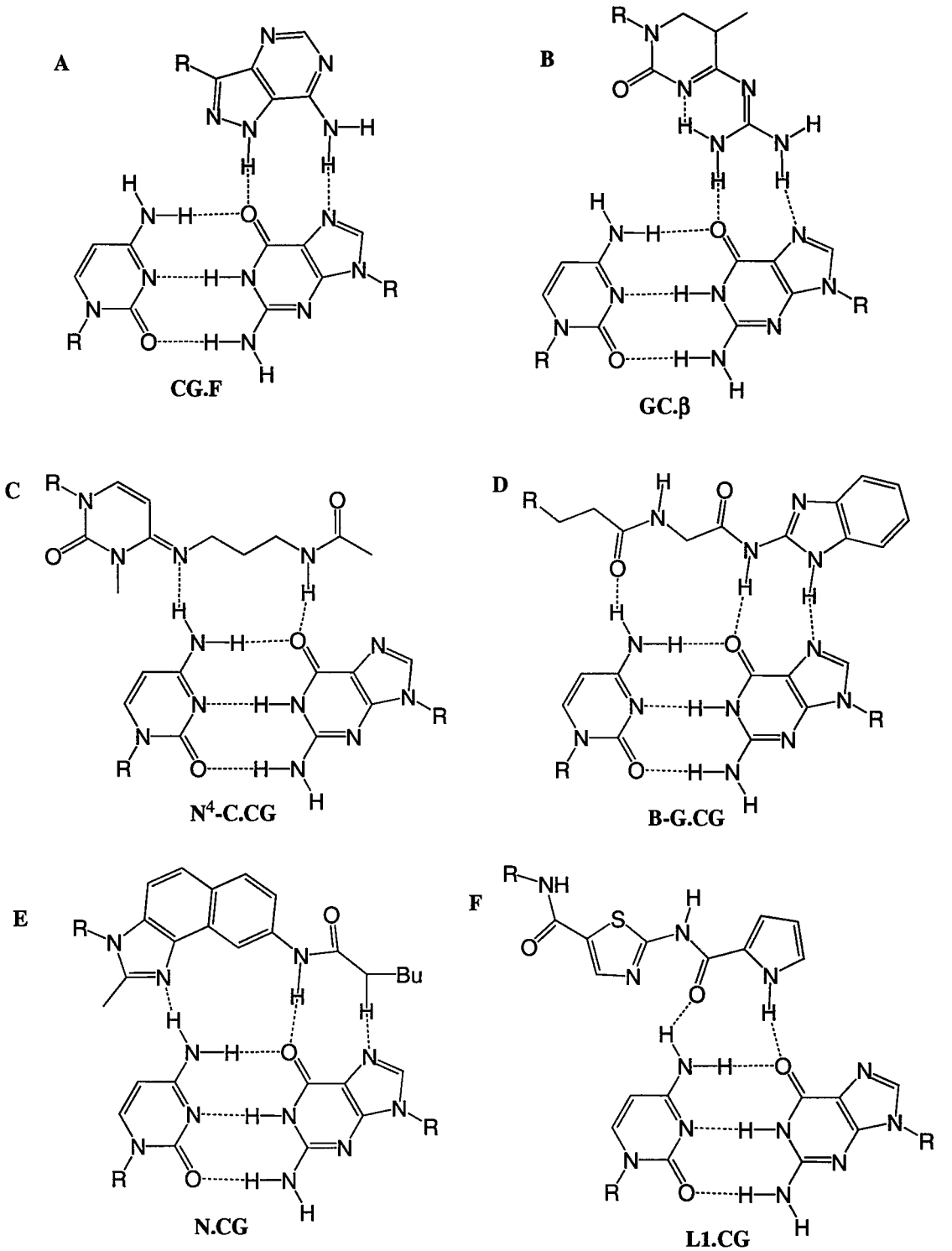


Figure 1.17. Base analogues that reach across the major groove or interact with both bases, (A) 2-deoxyformycin, (B) B-anomer of guanidinocytidine. (C) a N4-cytosine derivative, (D) benzoimadazole-glycyl, (E) 2-methyl-8-(N'-n-butylureido)naphth[1,2-d]imidazole)and (F) L1.

homoduplexes stabilised by d(GA) repeats. The proposed base-pairing schemes include G·G A·A G·A pairs (Noonberg *et al.*, 1995) and are shown in Figure 1.15. Also, poly dC and d(C,T) oligomers can form C^+·C base pairs at low pH, that may be involved in the i-motif (Gehring *et al.*, 1993).

The importance of minimising the potential for undesired secondary binding sites such as loops and bulges has been highlighted by the recognition of a 53 bp target duplex by a 17-mer, as examined by DNaseI footprinting (Fox *et al.*, 2000). The target contained oligopurine tracts with several pyrimidine interruptions, and produced a second previously unexpected binding site with a 20-fold lower affinity. This site arose from each end of the TFO forming a 7-mer triplex with adjacent regions on the duplex generating a contiguous 14 base triplex. Subsequently this was explored by TFOs designed to contain loops of varying length and various composition (bulges, loops, mismatches, or frayed ends). It was found that single base bulges were generally more stable than dinucleotide ones, though loops of up to 9 thymines still allowed triplex formation. Not only was the linker influential, but also base composition seemed to influence triplex stability, for single base bulges stability followed C>T>G=A. Such data implies that great care is needed when designing TFOs for triplex targeting to minimise these artifacts. Such factors might be important in determining strong triplex sites, the aim of the triplex research within this thesis.

Potential solutions

Self-association can be overcome by using guanine analogues such as 6-thio-G (Olivas & Maher, 1995b), 7-deazaG (Milligan *et al.*, 1993), or 7-deazaxanthine (Faruqi *et al.*, 1997), see Figure 1.18. Also, studies have sought to limit the formation of these self-associated structures by using base or backbone analogues (outlined in section 1.22.6).

1.22.5 Orientation

As mentioned previously, triplexes can be formed in either the parallel or antiparallel orientation relative to the target duplex strand. Pyrimidine containing oligonucleotides bind in the parallel orientation forming T·AT and C^+·GC triplets, while purine containing third strands bind antiparallel forming G·GC and A·AT triplets. However, G·GC and T·AT triplets can form in either orientation and the polarity of GT-containing third strands is

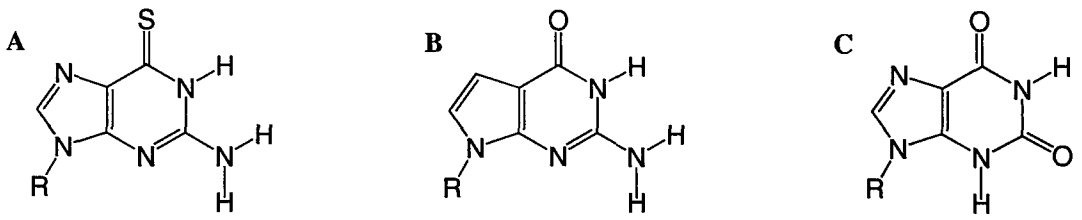


Figure 1.18. Base analogues that reduce self association (A) 6-thioguanine (B) 7-deazaguanine (C) 7-deazaxanthine.

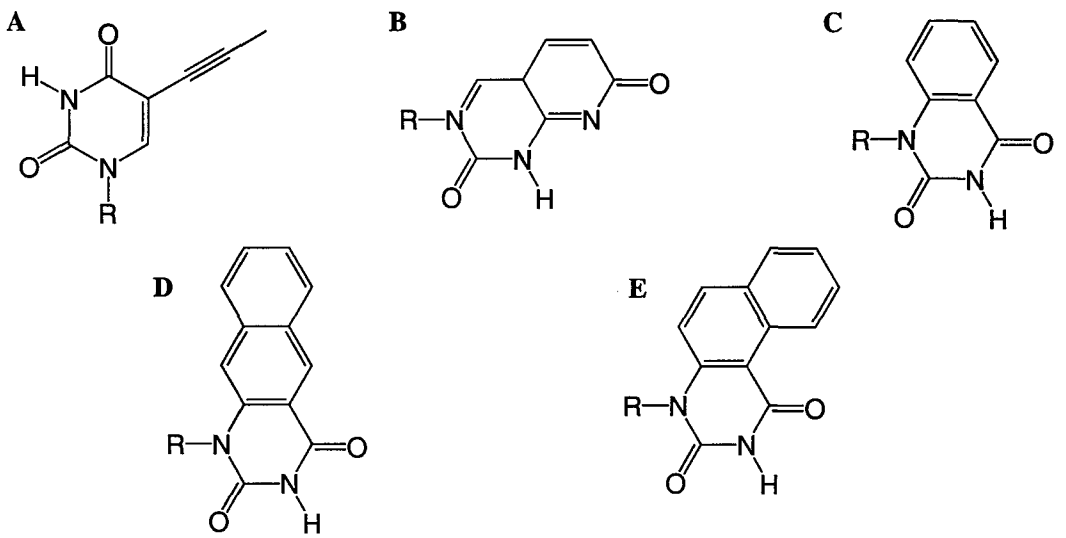


Figure 1.19. Base analogues aimed at improving base stacking, (A) Propynyl-dU (B) pyrido[2,3-d]pyrimidine (C)quinazoline-2,4 (1H,3H)-dione (D)benzo[f] (E) benzo[g].

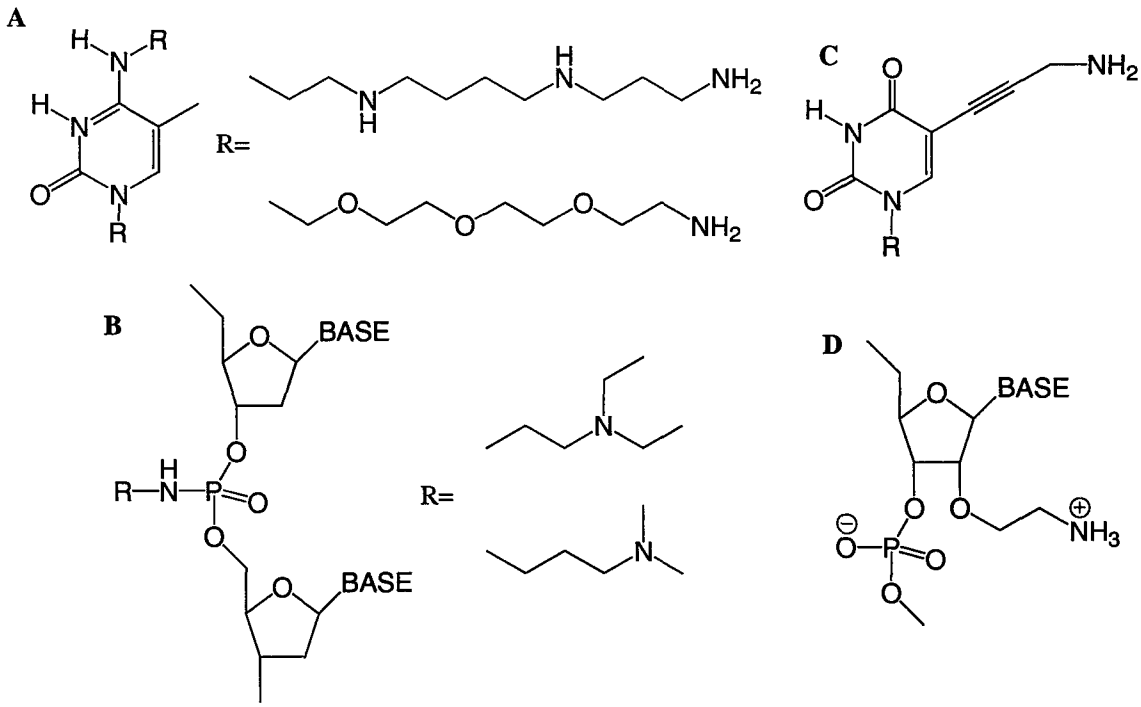


Figure 1.20. Examples of cationic groups and their possible linkage, (A) N4-spermine derivatives of methylcytosine, (B) phosphoramidate backbone, (C) propynylamino-dU, and (D) 2'-aminoethoxy moiety attached to 2'C of sugar.

ambiguous. An important factor for determining the orientation (and stability of several base analogues) is the position of the backbone. Unlike all the other the triplets, C^+·GC and T·AT are the only triplets which are isohelical; all other triplets produce distortions of the backbone between different base steps. Since there is less distortion between G·GC and T·AT in the antiparallel orientation, long GT oligonucleotides with several GpT and TpG steps tend to bend in an antiparallel orientation while long blocks of G and T can bind in the parallel form.

1.22.6 Sequence arrangement & composition

Since antiparallel triplexes form at physiological pH it might be expected that purine-containing oligonucleotides would form superior triplexes. The stability of antiparallel triplexes however varies widely from one study to another and appears to depend on the exact sequence of the target site. The reported affinities of antiparallel triplexes differ widely between different studies even under similar conditions. In some studies GA-containing third strands were shown to have very high affinities, short oligonucleotides exhibiting nanomolar binding constants (Svinarchuk *et al.*, 1994; 1995). While the apparent affinity was much weaker for triplexes in other studies, failing to show binding at micromolar concentrations (Chandler & Fox, 1996; Brown *et al.*, 1998). These differences most probably arise from the high sequence dependency of formation of antiparallel triplexes. This may be caused by the different positions of the third strand backbone at G·GC, A·AT and T·AT triplets. The most stable complexes are therefore likely to be those with the lowest number of ApG and GpA steps in the homopurine target site.

The relative stability of each triplet is an important factor and the antiparallel triplex motif is dominated by the G·GC triplet which seems to be more stable than T·AT or A·AT (Fox, 1994). Thus, G·GC dominates sequence composition, usually constituting over 60% of the triplets present in any stable antiparallel triplex. This must however be balanced by the propensity of G-rich oligonucleotides to form self-associated structures, which compete for and reduce the effective concentration of the oligonucleotide available for triplex formation. Self-association of these oligonucleotides compete with triple helix formation, especially at high concentrations. A study by Debin *et al.* (1999) found strong triplexes were produced by those sequences in a 13-mer population that possessed a high G content, and it was noted that at any given G-content, superior stability was afforded if the Gs lay in

runs. Though high G-content results in a concomitant increase in self-association, the best antiparallel triplexes may therefore be those in which the third strand has a high G-content, but in which the bases are arranged in such a way as to prevent the formation of unusual structures.

Sequence considerations for the parallel motif are dominated by cytosine protonation (as discussed in section 1.22.1). The parallel C^+·GC triplet is stronger than the T·AT triplets, especially under acidic conditions (Roberts & Crothers, 1996) and at pH 5 TFOs with higher C content are increasingly superior (Keppler & Fox, 1997). This is thought to represent improved electrostatic interactions of the positive charge on C reducing the repulsion of the three negatively charged backbones. The disadvantage of high C content is the concomitant reduction in the ability for the Cs to be protonated (Asensio *et al.*, 1998; Leitner *et al.*, 2000), and the decreased stability of C^+·GCs when adjacent (Kiessling *et al.*,1992). Thus, the most stable triplex will be produced by pyrimidine-TFOs where the number of $C^{(+)}$ is balanced against too many being contiguous (which is extenuated at low pH).

1.22.7 Affinity

To improve the binding affinity of triplexes, base analogues are incorporated into TFO which contain bulky substituents, aimed at improving base stacking. Examples of this approach include propynyl-U (Sollogoub *et al.*, 2002), pyrido[2,3-d]pyrimidine (Staubli & Dervan, 1994), and benzoquinazoline-diones (Michel *et al.*, 1997) (see Figure 1.19). Also cationic groups can be linked either to the base, backbone or sugar moiety to reduce backbone repulsion, examples include polyamines on the N4 of C (Ganesh *et al.*, 1997), propargylamino-U (Bijapur *et al.*, 1999), and an aminoethoxy moiety attached to the C2' of the ribose (Cuenoud *et al.*, 1998), see Figures 1.19 and 1.20. Modification of the sugar-phosphate backbone itself such as methylphosphonates (Trapane *et al.*, 1996), phosphoamidates (Zhou *et al.*, 1997), phosphothioates (Lacoste *et al.*, 1997), 2'-O methyl ODNs (Kandimalla *et al.*, 1997) can reduce backbone repulsion and also reduce the problem of nuclease attack *in vivo* (see Figure 1.21). Addition of a variety of ligands can also help to improve triplex stability (for review see Fox, 2000).

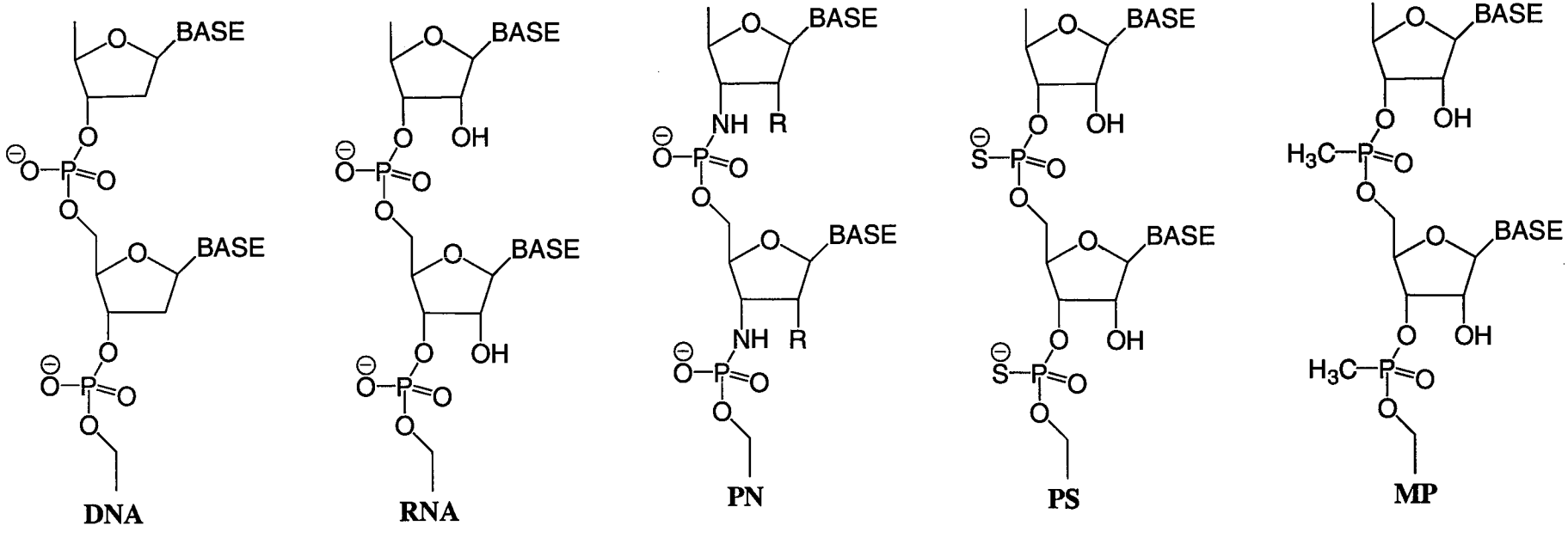
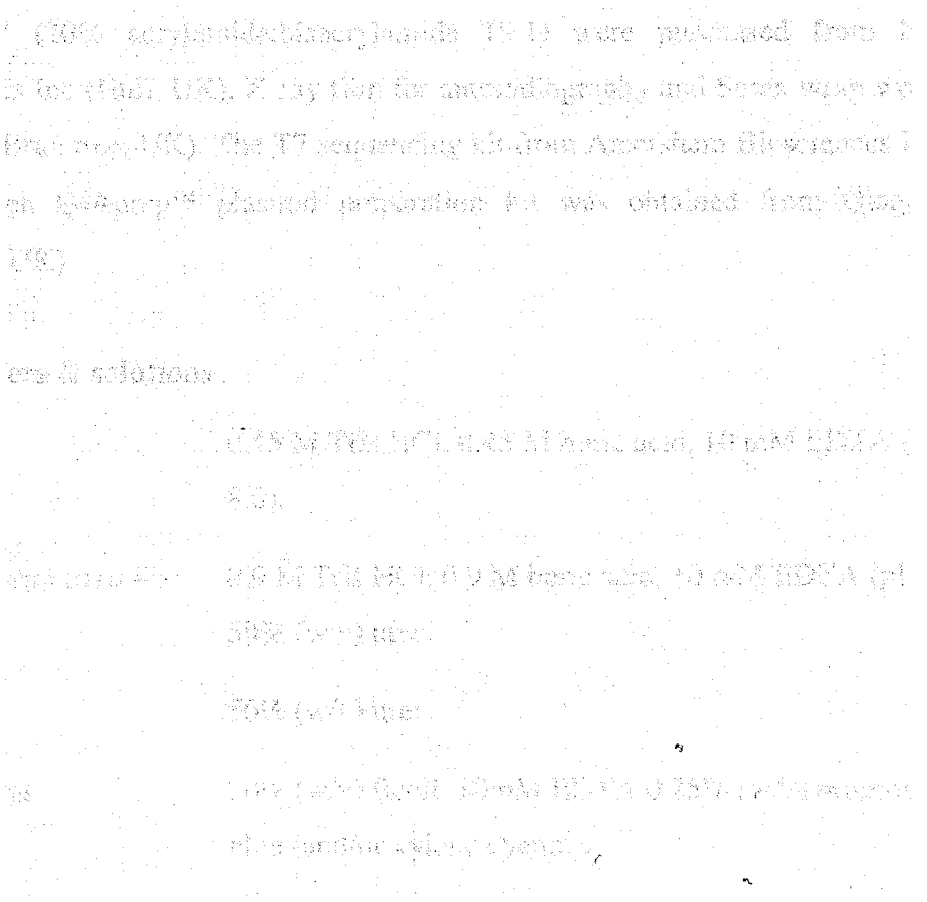


Figure 1.21. Backbone modifications, DNA and RNA possess a phosphodiester backbone. (PN) Phosphoamidates (PS) Phosphothioates. (MP) Methylphosphonates.

1.23 Aims of triplex research

The aim of the chapter concerning triplex formation is to identify the best antiparallel (GA and GT-containing) triplexes which are able to form with high affinity under physiological conditions. Similar studies are also performed to determine the optimum arrangement of bases for parallel (CT-containing) triplexes. These studies will define which sequences are the best candidates for targeting with triplex formation. The optimal triplex target sites are isolated from a mixed pool of oligonucleotides using a modification of the Restriction Enzyme Protection, Selection and Amplification (REPSA) assay.



2 Materials & Methods

2.1 Reagents

2.1.1 Chemicals & enzymes

The majority of chemicals were purchased from Sigma-Aldrich Inc (Poole, UK) or Fisher Inc (Loughborough, UK). T4-polynucleotide kinase and *Bsg*I were purchased from New England Biolabs Inc (Herts, UK) *Bam*HI, DNA ligase, *Sac*I and *Hin*dIII were purchased from Promega Inc (Southampton, UK). AMV reverse transcriptase and DNaseI were purchased from Sigma-Aldrich Inc *Taq* polymerase was obtained from Bioline Ltd (London, UK). pUC18 (*Bam*HI and BAP digested) and radioactivity as redivue [α^{32}P-] dATP and [γ^{32}P-] dATP at an initial specific activity of 3000 Ci/mmol (110 TBq/mmol) were purchased from Amersham Biosciences UK Ltd (Bucks, UK). *E.coli* strain TG2 was used for bacterial transformations. Polyacrylamide concentrates for gel electrophoresis as 'Accugel' (40% acrylamide:bis-acrylamide 19:1) and 'Sequagel' (20% acrylamide:bisacrylamide 19:1) were purchased from National Diagnostics Inc (Hull, UK). X-ray film for autoradiography and Saran wrap were from GRI Ltd (Braintree, UK). The T7 sequencing kit from Amersham Biosciences UK Ltd. The Qiagen QIAprep™ plasmid preparation kit was obtained from Qiagen Ltd. (Crawley, UK).

2.2.1 Buffers & solutions

5 x TBE –	0.45 M Tris HCl, 0.45 M boric acid, 10 mM EDTA (pH 8.0).
10 x TBE with urea –	0.9 M Tris HCl, 0.9 M boric acid, 10 mM EDTA (pH 8.0), 50% (w/v) urea.
Diluent –	50% (w/v) urea.
Loading dye –	20% (w/v) ficoll, 10mM EDTA 0.25% (w/v) bromophenol blue (and/or xylene cyanol).
Stop solution –	80% formamide, 10 mM NaOH, 10 mM EDTA, 0.25% (w/v) bromophenol blue.
1 x TE –	100 mM Tris.HCl, 0.1 mM EDTA (pH 7.5).

1 x TN – 10 mM Tris HCl, 50 mM NaCl (pH 7.5).

10 x kinase buffer – 700 mM Tris.HCl, 100 mM $MgCl_2$, 50 mM dithiothrietol (pH 7.6).

10 x NEB4 – 500 mM KOAc, 200 mM Tris-Acetate, 100 mM MgOAc, 10 mM dithiothrietol (pH 7.5).

10 x PCR buffer – 670 mM Tris-HCl, 160 mM $(NH_4)_2SO_4$ (pH 8.8).

10 x Multicore buffer – 250 mM Tris-OAc (pH 7.8), 1 M KOAc, 100 mM MgOAc, 10 mM DTT (pH 7.8).

10 x Ligase buffer – 300 mM Tris-HCl , 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP (pH 7.8).

Transformation buffer – 50 mM $CaCl_2$, 10 mM Tris-HCl (pH 7.4).

2YT media – 16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl.

KP buffer – 80.2% K_2HPO_4 / 19.8% KH_2PO_4 (pH 7.4).

NaP buffer – 77.4% Na_2HPO_4 / 22.6% NaH_2PO_4 (pH 7.4).

2.1.4 Oligonucleotides

Oligonucleotides were all synthesised by Oswel DNA Service (Southampton, UK). All unmodified oligonucleotides were purified by gel-filtration, and fluorescently-labelled oligonucleotides were purified by HPLC where indicated. On occasions fluorescently-labelled oligonucleotides were purified by denaturing 12.5% PAGE. For these fluorescent oligonucleotides, fluorescein (FAM) was incorporated at the 5'-end and methyl red (MeRed) was attached at the 3'-end using dR·FAM and dR·MeRed (Darby *et al.*, 2002). The structures of these groups are shown later in Figure 3.1. The nucleotide sequence of the oligonucleotides (5' – 3') were as follows:

REPSA oligonucleotides –

ssST GGTAGGATCCAGAGTGCAGACTGCTTRRRRRRRRRRCTCCTGAGA GGATCCATGG

RPR GGTAGGATCCAGAGTGCAG

LPR CTGGACTCTCCTAGGTACC

K KKKKKKKKKK

R RRRRRRRRR

Y YYYYYYYYY

K = G or T; R = G or A; Y = C or T.

Bandshift oligonucleotides –

TR2 TACAGATAGTTAGGGTTAGGGTTA

TR2A TACAGATAGTTAGAGTTAGAGTTA

TR2I TACAGATAGTTAIIITTAIIITTA

TR4 T_{10}TTAGGGTTAGGGTTAGGGTTAGGGT_{10}

TR4A T_{10}TTAGAGTTAGAGTTAGAGTTAGAGT_{10}

TR4I T_{10}TTAIIITTAIIITTAIIITTAIIIT$_{10}$

I = inosine.

Fluorescent oligonucleotides –

HT1.5 **Q**TGGGTTAGGG**F**

HT3.5 **Q**AGGGTTAGGGTTAGGGTTAGGGT**F**

OT3.5 **Q**TGGGGTTTTGGGGTTTTGGGGTTTTGGGGT**F**

HT1-3 AGGGT**Q**AGGGTTAGGGT**F**AGGGTTAGGG

Q = dR·MeRed; F = dR·FAM.

2.2 Polyacrylamide gel electrophoresis (PAGE)

For denaturing conditions, samples were heated to 95°C for 2 min with ~20% stop solution, whereas for non-denaturing conditions ~20% loading dye was added immediately before loading. The polyacrylamide gels were 40 cm long and 0.3 mm thick, and for denaturing conditions gels were run at 1500 V and 45 W for 1.5 h, the gel and running buffer contained 1 x TBE buffer with 5% urea. For native conditions, gels were run at 800 V and 20W for 2 h in 1 x TBE buffer. Both denaturing and non-denaturing gels were at room temperature. For bandshifts examining quadruplex

folding of samples, the native conditions were the same but 1 x TBE was supplemented with 20 mM K^+, and electrophoresis was run for 6 h at 4°C, so as to perturb higher oligonucleotide structures as little as possible. Following preparative PAGE, bands were simply visualised by autoradiography on X-ray film and excised and eluted. For quantitative analysis in either footprinting or band-shifting, gels were fixed in 20% acetic acid and dried on Whatman 3 mm paper for 1 h at 80°C. Band intensities were then determined by overnight exposure to a phosphoimager screen.

2.3 Polynucleotide kinase 5'-end ^{32}P-radiolabelling

2 µl aliquots of oligonucleotides were ^{32}P-labelled by the addition of 2 µl 10 x kinase buffer, 2 µl [γ^{32}P-] ATP, 10 µl deionised water and 1 µl T4-polynucleotide kinase (10 units). Reactions were incubated for 1-2 h at 37°C. 5 µl denaturing loading dye was added and samples were heated for 2 min at 95°C. Samples were subsequently purified by denaturing 12.5% PAGE. Desired bands identified by autoradiography were then excised from the gel and eluted in 400 µl 1 x TE (pH 7.5) and ethanol precipitated and stored in 1x TE (pH 7.5).

2.4 REPSA (Hardenbol & Van Dyke, 1996a)

The experimental details are outlined below, while the rationale of this technique is discussed further in chapter 7.

2.4.1 Preparation of duplex selection template (dsST)

The double-stranded selection template (dsST) for REPSA was produced by a single round of PCR using ^{32}P-labelled RPR. PCR reactions were prepared as 2 µl ssST, >50 nM ^{32}P-RPR, equimolar ~0.25 mM each dNTPs, 1.5 mM $MgCl_2$, and 1 µl *Taq* polymerase (3 units). PCR conditions used were 94°C (5 min) initially to denature; 60°C (10 min) for a single extension; then 4°C (∞). The ^{32}P-labelled dsST was purified by native 8% PAGE. The band corresponding to dsST was excised from the polyacrylamide gel and eluted in 1x TE overnight. Eluted dsST was ethanol precipitated and dissolved in 10 µl 1 x TE.

2.3.2 Protection

To 10 µl ^{32}P-dsST or the sample from previous rounds were added 4 µl of one of the degenerate TFO 10-mers (K10, R10 or Y10) at a final concentration of 100 µM, 2 µl 10

x NEB4 and 3 µl deionised water. This provided conditions favourable to triplex formation. Incubation proceeded overnight at room temperature.

2.4.3 Selection

After triplex formation, 1 µl 1.6 mM S-adenosylmethionine (80 µM) and 2 µl *Bsg*I (6 units) were added to the 19 µl *protection*-samples that were incubated for a further 60 min at 37°C. *Bsg*I cleavage was stopped by boiling at 65°C for 20 min.

2.4.4 Amplification

The *selection*-samples are a mixture of undigested *Bsg*I-cleavage resistant (*i.e.*triplex-competent) and digested dsST sequences. 22 µl selection-samples were amplified by PCR. PCR reactions were prepared as ~100 pmol LPR, ~100 pmol RPR, equimolar ~0.25 mM each dNTPs, 1.5 mM $MgCl_2$, >50 nM ^{32}P-RPR , and 1 µl *Taq* polymerase (3 units). PCR conditions used were 94°C (3 min) initially to denature; then a cycle of 94°C (1.5 min) and 60°C (1 min) for extension; then 4°C (∞). 9 cycles were performed in initial REPSA followed by a further 3 and 6 additional rounds. 35 µl was withdrawn after each further 3 cycles. Amplification in subsequent REPSA was for 12 rounds only.

Following PCR, the two distinct dsST subpopulations were separated by 8% native PAGE. Bands representing (amplified) dsST were excised from the polyacrylamide gel and eluted in 1x TE overnight. Eluted DNAs were ethanol precipitated and dissolved in 10 µl 1 x TE. The dsST populations that have undergone selection and amplification were then subject to subsequent rounds of REPSA.

2.5 Cloning of dsST products into plasmid vector

2.5.1 *Bam*HI digestion

The emergent (triplex-forming) dsSTs were redissolved in 20 µl 1x TE. 2 µl 10 x buffer E and 1.5 µl *Bam*HI (15 units) were added. *Bam*HI digestion proceeded for between 30 min and 2 h at 37°C. *Bam*HI was then heat-inactivated at 80°C for 20 min. Half of each sample was retained.

2.5.2 Plasmid ligation

12 µl (*ca.* half) *Bam*HI digested samples were ligated to 0.5 µg linear pUC18 plasmid, which had been similarly cut with *Bam*HI and treated with bacterial alkaline phosphatase in 1 x ligase buffer, this was to prevent self-ligation. 2 µl DNA ligase (2 units) was added in a total reaction mixture of 20 µl. Ligation proceeded for 3 h to overnight at room temperature. Again half of each sample was retained.

2.5.3 Transformation

Samples of the ligation mixtures (*i.e.* plasmid containing insert) were added to TG2 competent cells, and left on ice for 30 min. The cells were then heat shocked at 45°C for 1 min and replaced on ice. Using sterile technique 20 µl transformation mix was plated on an agar plate while the remainder was plated on another. Agar plates contained 1 mM X-gal, 0.02% (w/v) IPTG, 0.1 mg / ml carbenicillin. Transformants were selected by ampicillin resistance conferred by the pUC18 plasmid. Those with the DNA insertion into the multiple cloning site were selected by inactivation of the *lacZ'* gene (blue/white screening) on X-gal media. Transformants were grown overnight in at 37°C.

2.5.4 Growth & collection of transformants

Individual white colonies were picked-off the agar plates and used to inoculate 5 ml 2YT media to which had been added 5 µl 100 mg/ml carbenicillin. Cultures were then grown for 6 to 16 h on a shaker at 37°C.

2.5.5 TG2 competent cells

A 5 ml culture of 2YT was inoculated with *E.coli* TG2 cells from a stock agar plate stored at 4°C. The culture was grown for 6 and 16 h and then 0.5 ml was transferred to 50 ml 2YT media. The exponential phase of growth was reached when optical density at 600 nm reached 0.4 to 0.6. The culture was then centrifuged at 5000 rpm, at 4°C for 5 min. The supernatant was discarded and the pellet resuspended in 20 ml transformation buffer (50 mM $CaCl_2$, 10 mM Tris-HCl pH 7.4) and left on ice for 30 min. The sample was then, centrifuged as before and the pellet resuspended in 3 ml transformation buffer and retained at 4°C for use within 14 days.

2.6 Plasmid recovery & purification (Qiagen QIAprep)

Transformed TG2 cultures were grown overnight in 5 ml 2YT media as described in section 2.5.4. 400 µl 50% sterile glycerol was added to 600 µl of each culture and retained as a stock. The remainder was divided between two 1.5 ml Eppendorf tubes and centrifuged at 7000 rpm for 5 min. The pelleted cells were then resuspended in 250 µl buffer P1 and lysed with 250 µl buffer P2 then gently mixed. 350 µl neutralising buffer N2 was added to the lysed cell suspension that was again gently mixed, before centrifuging at 14000 rpm for 10 min. The supernatant was transferred to a Qiaprep spin column in a 2ml collection tube then spun for 1min at 14000 rpm and the flow-through discarded. The plasmid, captured in the column was then washed and spun as before with 750 µl buffer PB and 500 µl PE. Finally, the column was transferred to a 1.5 ml Eppendorf tube and the plasmid eluted by adding 40 µl water for 1 min and separated by spinning at 14000 rpm.

2.7 Dideoxy sequencing

2.7.1 Denaturation of DNA

To 40 µl plasmid pUC18 (containing insert) purified as described above in section 2.6, 10 µl 2 M NaOH was added and left at room temperature for 10 min. Subsequently, 15 µl 3 M NaOAc (pH 4.8), 35 µl deionised water and 300 µl EtOH was added and the sample placed on dry ice for 10 min to precipitate the DNA. The sample was then centrifuged at 14000 rpm, at 4°C for 10 min, and the pellet washed in 70% ethanol, before drying samples under vacuum.

2.7.2 Annealing of primer

The DNA pellet was redissolved in 10 µl deionised water, and 2 µl annealing buffer and 2 µl universal primer were added (supplied in T7 sequencing kit). The sample was allowed to anneal for 20 min at 37°C, then for at least 10 min at room temperature.

2.7.3 Sequencing reactions

For each sequence a mixture of 3 µl label mix A, 1.5 µl deionised water and 0.5 µl [α^{32}P-] ATP was prepared on ice. 2 µl T7 polymerase (T7 sequencing kit) diluted 4-fold was added to this chilled mixture. The above 6 µl mixture was then added to each 10 µl sequence and left at room temperature for 5 min, while 4 samples were also placed of 2.5 µl of each of the dideoxy G, C, T, A mixes for each sequence in the 37°C block. 4.5

µl of the sequence reaction mixture was added after 5 min to each of the four dideoxy samples to initiate elongation. After 5 min the reaction was terminated by the addition of 5 µl stop solution and heated at 95°C for 2 min, then finally crash cooled on ice. Half the samples were loaded onto a 10% denaturing PAGE.

2.8 DNaseI footprinting

2.8.1 Recovery of sequence inserts from plasmid and ^{32}P- labelling

35 µl plasmid purified (as in section 2.6), was co-digested by addition of 4.5 µl 10 x multicore buffer, 1.5 µl deionised water, 2 µl *Sac*I and 2 µl *Hin*dIII, for 2 h at 37°C. Fragments generated were subsequently radiolabelled by addition of 1 µl [α^{32}P] dATP and 1 µl AMV reverse transcriptase and incubated for 30 min at 37°C. Labelling was stopped by addition of stop solution and samples were purified by native 8% PAGE. The band corresponding to the insert was excised and eluted in 400 µl 1 x TE, then ethanol precipitated by addition of 40 µl 3 M NaOAc (pH 8.0) and 1 ml ethanol. Samples were redissolve to an activity of 10 cps/µl in 1 x TE.

2.8.2 Footprinting reactions

Individual TFOs for interaction with the selected sequences were assumed to be those that produce triplexes consisting of canonical triplets **only**. Where selection was based on degenerate GT-TFOs – for oligonucleotides which can bind in both parallel and antiparallel orientations both third strand oligonucleotides were examined. The affinity of these synthesised TFOs was examined by DNase I footprinting.

Usually, 1.5 µl of the ^{32}P-labelled DNA fragment was mixed with 3 µl of the TFO. The TFO was serially diluted over an appropriate concentration range in 1 x TN plus 10 mM $MgCl_2$or $MnCl_2$ for experiments at pH 7.5 and a buffer of 50 mM NaOAc and 10 mM $MgCl_2$ for acidic pH.

Samples were equilibrated for 3 h at room temperature to allow triplex formation, then digested by DNase I. Single hit kinetics were achieved by limiting DNaseI digestion . This was produced by using 2 µl DNase I, diluted from 7200 units/ml to ~0.03 units/ml in 20 mM NaCl, 2 mM MgCl, and 2 mM $MnCl_2$. Reactions were allowed to proceed for 1 min before adding 4.5 µl stop solution and heating to 95°C for 2 min, then crash cooled in ice and resolved by an denaturing 9% PAGE. A Maxam-Gilbert GA specific-

marker lane was also prepared by adding 20 µl deionised water and 5 µl stop solution to 1.5 µl of labelled DNA, and boiling down for 30 min.

Once fixed and dried, any footprints were quantified on a Phosphoimager™ (Molecular Dynamics Model Storm 860) and analysed using ImageQuant™ software. Triplex formation was examined over a narrow concentration range, allowing approximate binding constants (K_d) for an oligonucleotide to be calculated. This would be expressed as the concentration producing a 50% intensity reduction (C_{50}), this can be described by:

$$\frac{I}{I_0} = \frac{C_{50}}{C_{50} + [L]} \qquad 2.1$$

Where I and I_0 are relative band intensities in the presence and absence of TFO and $[L]$ is the TFO concentration.

2.9 Fluorescence-quenching studies of quadruplex

2.9.1 Fluorescence melting curves

Fluorescence melting curves were monitored in a Roche LightCycler™, using a total reaction volume of 20 µl. 10 µl of 0.5 µM fluorescently labelled oligonucleotide and 10 µl G-quadruplex ligand at a given concentration in an appropriate buffer (typically phosphate buffer) were added to capillary tubes and briefly spun to collect at the bottom. The final oligonucleotide concentration was 0.25 µM, diluted in an appropriate buffer.

Experiments were then performed using various thermal programmes in which the samples were melted and annealed at a heating and cooling rate of 0.1°C sec^{-1} in the 30 to 95°C range. Excitation was at 488 nm while fluorescence changes were continuous measured at 520 nm during both melting and annealing steps. This was modified where very slow kinetics were observed so that single readings were recorded following one degree steps every 2 min to 20 min (0.5 to 0.05°C sec^{-1}) again over the same 30 to 95°C range. Where indicated samples were re-used in subsequent melts and anneals at differing rates of heating and cooling and schemes.

T_m values were determined from first derivatives of the melting profiles, or from van't Hoff analysis of the melting profile using FigP™ or SigmaPlot™, the fitting the curves (based on an unimolecular two-state reaction) was derived as follows:

$$K = \frac{[q]}{[s]} \quad 2.2$$

$$[q] = K[s] \quad 2.3$$

$$[t] = [q]+[s] \quad 2.4$$

Where K is the equilibrium constant, $[q]$ is the concentration of quadruplex, $[s]$ is the concentration of single strand and $[t]$ is the total oligonucleotide concentration. Substituting equation 2.3 into 2.4 gives

$$[t] = K[s]+[s]$$

$$[t] = [s](K+1) \quad 2.5$$

$$[s] = \frac{[t]}{K+1} \quad 2.6$$

$$F = \varepsilon_q[q]+\varepsilon_s[s] \quad 2.7$$

Where F is the fluorescence of a sample, ε_q is the fluorescence coefficient of the quadruplex species and ε_s is that of the single strand. Substituting equation 2.3 into 2.7 gives

$$F = \varepsilon_q K[s]+\varepsilon_s[s]$$

$$F = [s](\varepsilon_q K+\varepsilon_s) \quad 2.8$$

substitute equation 2.6 into 2.8 gives:

$$F = \frac{[t](\varepsilon_q K+\varepsilon_s)}{K+1} \quad 2.9$$

Assuming initially $[t] \approx [q]$ and finally $[t] \approx [s]$, then $[t]\ \varepsilon_q \approx F_{init}$ and $[t]\ \varepsilon_s \approx F_{final}$ then:

$$F = \frac{F_{init}K+F_{final}}{K+1} \quad 2.10$$

The van't Hoff equation describes the relationship between K and temperature:

$$\Delta G = -RT\ln K \quad 2.11$$

$$\Delta G = \Delta H - T\Delta S \quad 2.12$$

Where ΔG is Gibbs free energy, ΔH is enthalpy change and T temperature. Rearranging equations 2.11 and 2.12 gives:

$$K = e^{-\left(\frac{\Delta G}{RT}\right)} = e^{-\left(\frac{\Delta H}{RT} - \frac{\Delta S}{R}\right)} \quad 2.13$$

The change in K between T_0 and T_1 is therefore expressed as:

$$\Delta K = K_1 - K_0 = e^{-\left[\left(\frac{\Delta H}{RT1} - \frac{\Delta S}{R}\right) + \left(\frac{\Delta H}{RT0} - \frac{\Delta S}{R}\right)\right]} = e^{\left[\frac{\Delta H}{R}\left(\frac{1}{T0} - \frac{1}{T1}\right)\right]} \quad 2.14$$

$$K_1 = K_0 e^{\left(\frac{\Delta H}{R}\frac{1}{T0-T1}\right)} \quad 2.15$$

Taking K_0 as 1 at $T_0 = T_m$ where 50% population is in either form and substituting equation 2.15 into 2.10 gives the theoretical curve to which experimental data is fitted by FigP software:

$$F_{app} = \frac{F_{final} + F_{init}\, e^{[(\Delta H/R)(1/Tm - 1/T)]}}{1 + e^{[(\Delta H/R)(1/Tm - 1/T)]}} \quad 2.16$$

In some instances the initial or final fluorescence was not constant. On these occasions a simple linear variable in F_{init} or F_{final} was included in the calculation.

2.9.2 Temperature jump kinetics

The kinetics of quadruplex dissociation were determined by measuring the rate of change of fluorescence, after rapid increasing temperature, in a manner similar to that of temperature-jump relaxation kinetics. Each sample contained 0.25 to 2.5 μM quadruplex-forming oligonucleotide (as indicated). In order to ascertain the relevance of proper assembly of quadruplex, for HT3.5 oligonucloeotide, the mixtures were first, melted by heating to 95°C and re-annealed by cooling to 30°C at 0.1°C sec^{-1}. However, for the OT3.5 sequence this annealing was performed by denaturing the samples in a PCR machine at 95°C, then slowly re-annealing in the chosen buffer in 5 °C decreases every 100 min to 30°C.

HT3.5 complexes were then equilibrated in the LightCycler™ for 5 min at 30°C. The temperature was then rapidly increased at 20°C sec^{-1} by 5°C in consecutive steps up to 90°C and the time dependent change in fluorescence was recorded over the next minute. OT3.5 complexes underwent similar 5°C temperature jumps from 30°C up to 95°C, but each step was for a period of 60 min during which readings were taken every 30 sec. This increase in temperature caused some of the quadruplex to dissociate, producing an increase in fluorescence. Theoretically, a dead-time of 0.25 sec is expected and as such the first 1 sec during equilibration to new temperatures was

ignored. Successive temperature jumps were performed on the same sample. The observed relaxation rate constant ($1/\tau$) should be independent of concentration as it is unimolecular reaction, and be the sum of the dissociation (k_{-1}) and association (k_1) rate constants. Values for $1/\tau$ were determined from the relaxation curves, by fitting the time-dependent changes in fluorescence using SigmaPlot™ to a single or double exponential function:

$$F_T = F_F\left(1 - e^{-t/\tau}\right) + F_0 \qquad 2.17$$

$$F_T = F_1\left(1 - e^{-t/\tau_1}\right) + F_2\left(1 - e^{-t/\tau_2}\right) + F_0 \qquad 2.18$$

where F_T is fluorescence at time t, for the single exponential fit, F_f and F_0 are the final the initial fluorescence respectively; while for the bi-exponential fit, F_1 and F_2 are two floating final fluorescences and $1/\tau_1$ and $1/\tau_2$ are the two relaxation times.

2.10 Bandshift studies of quadruplex

The bandshift protocol was adapted from Han *et al.* (1999). The oligonucleotide stock prepared contained equal volumes of 20 μM unlabelled and ^{32}P-labelled oligonucleotide producing a concentration of 10 μM, this was then denatured at 95°C for 5 min before crash cooling. Quadruplex ligand solutions of an appropriate range were also prepared (as discussed in chapter 6) at 2-times the final solution. 3 μl 10 μM unlabelled / ^{32}P-labelled oligo aliquots were mixed with 3 μl quadruplex ligand in 1 x TE plus 200 mM KCl. This produced the final the reaction: 5 μM quadruplex forming oligonucleotide, various ligand (typically 0 to 20 μM) and 0.5 x TE + 100 mM KCl. Samples were incubated for 4 h at room temperature, before adding 3 μl ficoll loading dye and placing on ice before loading. The samples were loaded onto a 16% native PAGE that was run at 4°C for 6 h, the gel and running buffer contained 0.5 x TBE supplemented with 20 mM KCl.

As before, fixed and dried gels were quantified on a Phosphoimager™ (Molecular Dynamics Model Storm 860) and analysed using ImageQuant™ software. The ^{32}P-intensity of each distinct band within the lanes were determined. It was assumed that higher and lower mobility bands corresponded to single-strand and quadruplex structures, respectively. The percentage of the intensity within the quadruplex bands at specific ligand concentrations (q) was calculated from the single strand band (I_{ss}) and quadruplex band (I_q) intensities, according to the simple equation:

$$q = \frac{I_q}{I_q + I_{ss}} \times 100 \qquad 2.19$$

Plotting q against (varying) drug concentration gave a sigmoidal curve suggesting a co-operativity for drug interactions, as such these plots were also fitted to the following equation:

$$\frac{q}{q_{max}} = \frac{[L]^n}{C_{50} + [L]^n} \qquad 2.20$$

Where q_{max} is the maximum percentage of quadruplex achieved following the 4 h incubation, C_{50} is a parameter that represents binding affinity, L is ligand concentration and n is the number of ligands bound.

3 Effects of G4 Ligands on Quadruplex Stability Studied by Fluorescence Melting and Bandshift Analysis

3.1 Introduction

3.1.1 Quadruplex as a molecular target

The telomeres are acknowledged as potential targets for chemotherapy in cancer treatment. Quadruplex formation is an accepted means to inhibit telomere extension by telomerase *in vitro* (Zahler *et al.*, 1991), presenting quadruplexes as novel molecular targets. As discussed in chapter 1, these studies are motivated by the clinical potential of stabilisation of telomeric DNA folded into a quadruplex by drug-binding. Many small molecules can interact with (unusual) nucleic acid structures, the challenge is to design ligands that selectively stabilise quadruplex DNA in preference to the vast excess of duplex DNA.

3.1.2 Studies of quadruplex-drug interactions

Though native quadruplex structures have been extensively characterised, at present there is limited structural data on their complexes with drugs. The first high resolution structure of a quadruplex-ligand complex was recently described (though some previous NMR data is available) (Haider *et al.*, 2003). This comprised of a dimeric quadruplex formed by two *Oxytricha* repeats in a 1:1 complex with a di-substituted acridine (BSU6039). Several key insights into ligand binding were made by comparison with an earlier crystal structure of the same oligonucleotides (Haider *et al.*, 2002). It was shown that binding of a single drug molecule within one loop displaced K^+ from within the loop. The ligand sidechains did not penetrate the grooves, but formed specific interactions with the loop, as did the amide of the acridine ring. Ligand-binding extensively altered the native structure, particularly in the loop regions.

Studies on quadruplex drug-binding have used several techniques. (i) Molecular modelling can predict compounds with strong interaction. (ii) Screening libraries of compounds determines their cytotoxicities in tumour cell lines. (iii) The inhibition of telomerase activity *in vitro* can be assayed by TRAP in cell extracts. (iv) Quadruplex-drug interactions is also sometimes used to examine changes in ligand absorption or emission spectra. (v) Electrophoresis (bandshifts) has been used to in a limited number

of studies to quantify the fraction of folded quadruplex induced by drug-binding. (vi) The stability of quadruplexes and their interactions with ligands can also be analysed by oligonucleotide melting studies.

Melting studies are valuable in assessing drugs whose therapeutic effects are assigned to the stabilisation of structured nucleic acids (Wilson *et al.*, 1997). This follows the thermally-induced melting of the quadruplex into its separate strands. The midpoint of the transition (T_m) is the temperature at which half the strands are in the native state and half are in the random coil state, and this serves as a measure of the stability of the (quadruplex) structure. Changes in this value (ΔT_m) allow the effect of experimental conditions, and drug-binding to be compared. However, it should be remembered that T_m values give no specific information about the structure of the nucleic acid-drug complex.

The most commonly used method for following thermal denaturation is UV absorbance. This technique is simple to use and easily interpreted qualitatively. It measures changes in absorbance at 260 nm which are indicative of changes in base stacking and solvation of the bases. However, these absorbance changes are typically small, typical increases are less than 25%. The absorbance change upon quadruplex formation is small (Mergny *et al.*, 1998). Relatively large volumes (1-3 ml) of a solution with an OD_{260} of at least 0.2 (*i.e.* total of ~20 nmol) are required and this is especially problematic if studying a short oligonucleotide where concentrations need to be raised higher. Interpretation can be complicated if there are several overlapping components to the melting profiles, making it difficult to resolve the different transitions. Fluorescence resonance energy transfer (FRET) between a donor (fluorescein) and an acceptor (rhodamine) is another technique that has been used to follow the folding of an oligonucleotide into a quadruplex structure (Mergny & Maurizot, 2001).

3.1.3 Fluorescence-quenching of molecular beacons.

Molecular beacon methods were initially developed for use in *in situ* hybridization, real-time monitoring of PCR, and for gene detection and mRNA probes within living cells (Tyagi & Kramer, 1996). In this study we have developed the use of molecular beacons for melting studies with folded nucleic acid.

The molecular beacons approach uses fluorophore and quencher molecules that are located so as to produce a fluorescence change when there are changes in the nucleic acid structure (Darby *et al.*, 2002). A fluorophore (fluorescein) and a quencher (methyl red) are covalently attached to the oligonucleotide of study. When the oligonucleotide is folded, the fluorophore-quencher pair are in close proximity and cannot fluoresce. This results from collisional quenching whereby the energy received by the fluorophore is transferred to the quencher and dissipated as heat, rather than emitted as light (as in FRET). As the structure melts, the fluorophore and quencher are separated, which produces a large increase in fluorescence. This provides an excellent way of directly following the assembly and/or association of oligonucleotides. Valuable thermodynamic and kinetic data on the unfolding and the formation of quadruplex-forming oligonucleotides can be provided from these melting and annealing profiles.

Importantly, this method overcomes some of the limitations of UV absorbance for studying nucleic acid melting profiles, and their interactions with ligands (Darby *et al.*, 2002). It benefits from higher sensitivity, reducing the amount of nucleic acid required. Another advantage is that by carefully selecting the location of the fluorescent molecules, different transitions can be measured independently if desired.

3.1.4 Bandshift analysis of quadruplex assembly

In parallel with the fluorescence-quenching experiments, simple bandshift assays were carried out to assess the effects of ligands on quadruplex assembly. These have been previously used as a qualitative tool for exploring a range of influences on quadruplex behaviour. This includes such as the effects of (mono- and divalent) cations (Sen & Gilbert, 1990; Venczel & Sen, 1993), the effects of varying the length of G tract(s) or intervening thymine tract (Guo *et al.*, 1993), and for demonstrating the stoichiometry of specific complexes (Sen & Gilbert 1990). There are however, only a few examples in which bandshifts have been used for quantitative analysis of quadruplex stability and dynamics, and these have been discussed in chapter 1 (Raghuraman & Cech, 1990; Sen & Gilbert, 1990; Fang & Cech, 1993; Guo *et al.*, 1993; Han *et al.*, 1999). Another study used bandshifts to demonstrate the unusual effect of one oligonucleotide on quadruplex formation by another (Marco-Haviv *et al.*, 1999).

3.1.5 Aims

The work described in this chapter uses the molecular beacons technique to examine ligand-induced quadruplex formation with oligonucleotides containing human or *Oxytricha* telomeric repeats. In addition, a bandshift assay was adapted from similar studies (for example Han *et al.*, 1999) and was used to examine intermolecular (TR2) and intramolecular (TR4) quadruplex-forming oligonucleotides, composed of two and four human telomeric repeats respectively. This assay differentiates between the unfolded strands and folded quadruplexes, and allows an estimate of the single-strand to quadruplex equilibrium and the influences of ligands upon this equilibrium. Although quadruplexes are thermodynamically stable, their formation is extremely slow, and under the reaction conditions used in this work, there is very little quadruplex formation in the absence of the G4 ligands over the incubation time. Their addition therefore acts as a catalyst driving the reaction more rapidly towards equilibrium by increasing the rate of quadruplex formation.

These studies were performed to compare the effects of a wide range of novel quadruplex-binding ligands, and to provide further information on drug-quadruplex interactions. Thus it is hoped that the structure-activity relationships needed to help design improved quadruplex ligands can be elucidated.

3.2 Methodology

3.2.1 Fluorescence melting studies

Oligonucleotides containing either two or four guanine-tracts were used in this work. A 23-mer (HT3.5) containing 3.5 human telomeric repeats was designed to fold into an intramolecular quadruplex. Also a 9-mer (HT1.5) containing 1.5 repeats was expected to dimerise to produce an intermolecular quadruplex. A 30-mer (OT3.5) was also produced containing 3.5 repeats of the *Oxytricha* telomeric repeat. All these oligonucleotides were labelled with dR-FAM at the 3'-end and dR-methyl red at the 5'-end, enabling the fluorescence changes to be measured during thermal denaturation. The sequences of these fluorescent oligonucleotides are given in Materials & Methods and the fluorophore and quencher structures are shown in Figure 3.1.

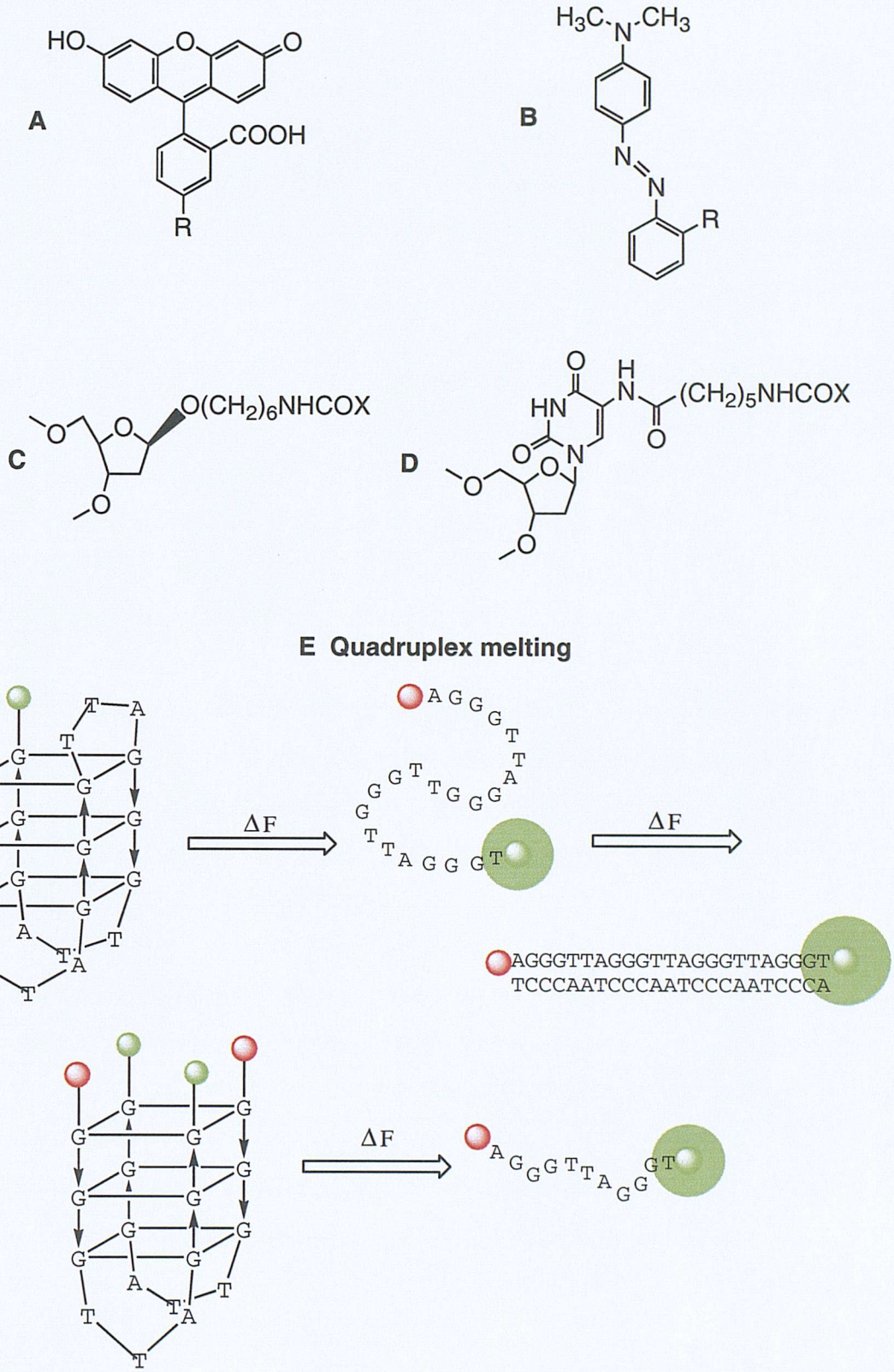


Figure 3.1. Structures of fluorescent molecules and their oligonucleotide linkage. (A) Fluorescein, and (B) methyl red (R = ribose or uracil); (C) linkage to abasic ribose moiety, and (D) linkage to uracil of uridine (X = fluorophore or quencher).(E) Schematic representation of oligonucleotides designed to show the melting of intra- and intermolecular quadruplexes (for simplicity only one possible quadruplex topology shown).

3.2.2 Bandshift studies

Oligonucleotide TR2 has the same sequence as that used by Han *et al.* (1999), while a similar oligonucleotide containing four G-tracts that self-folds have been used to test other G4 ligands (Kim *et al.*, 2002), for sequences see Materials & Methods. It is particularly important to specify the experimental conditions used in this work, these are detailed below.

An oligonucleotide concentration of 5 μM was selected, since this is comparable with that used in other studies with similar oligonucleotides (Han *et al.*, 1999; Han *et al.*, 2001; Koeppel *et al.*, 2001). In these reactions a small fraction of the DNA was radiolabelled with ^{32}P, while the majority consisted of unlabelled DNA. The reactions were presumed to be bimolecular for TR2, and unimolecular for TR4. As such strand concentrations is more important for reactions with TR2.

The presence of a moderate millimolar concentration of K^+ (or Na^+) is essential for any study on quadruplex formation. Therefore 100 mM KCl was chosen for these experiments. This is both physiologically relevant and is similar to that used in most ligand-quadruplex studies (Han *et al.*, 1999; Rangan *et al.*, 2001; Han *et al.*, 2001; Koeppel *et al.*, 2001; Kim *et al.*, 2002).

Low micromolar levels of drug were used in the fluorescence-quenching studies, representing a 40-fold excess over the oligonucleotide concentration. A similar excess of ligand over oligonucleotide is not needed for bandshift analysis, as the oligonucleotide concentration is much higher. Ligand concentrations up to 20 μM were generally used, though this was increased to 100 μM in some instances. Ligands which are only active at greater concentrations will be of little relevance. These concentrations are similar to those used with other compounds in the literature, allowing comparison with other characterised G4 ligands.

The association rate constant for the TR2 oligonucleotide has previously been reported to be 10^2 $M^{-1}h^{-1}$ (Han *et al.*, 1999). As such, quadruplex formation under these conditions will be far from equilibrium at any reasonable experimental time-frame. The results presented in this chapter were all taken after 4 h incubation and therefore

represent a 'snapshot' of the progression of quadruplex formation. Again this is similar to conditions that have been described in other studies of G4 ligands (Han *et al.*, 1999; Rangan *et al.*, 2001).

Samples were subsequently run on native polyacrylamide gels maintaining non-denaturing conditions by running at low voltage and wattage, and at 4°C, typically for over 6 h. The K^+ concentration was 20 mM in both the gel and running buffer as used in previous studies (Han *et al.*, 1999; Han *et al.*, 2000; Rangan *et al.*, 2001; Han *et al.*, 2001; Koeppel *et al.*,2001; Kim *et al.*, 2002).

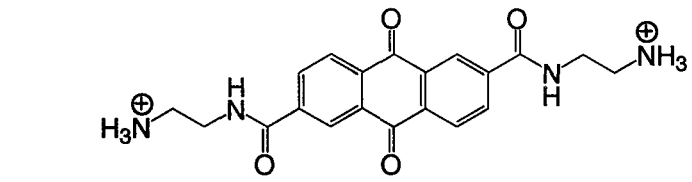
3.2.3 Quadruplex-interactive ligands

A new series of novel substituted anthraquinone have recently been prepared as quadruplex-binding ligands, shown in Figures 3.2 to 3.5. For consistency all amines are shown with all secondary and tertiary amines protonated, their exact pattern of protonation is however unclear. These were supplied by Prof. M. Palumbo and Prof. G. Zagotta (Dept of Chemistry, University of Padua). These compounds were arbitrarily assigned codes based upon the order of their synthesis. All are di-substituted at the 2,6- or 2,7- positions with a range of propionamido-amino acyl or simple alkyl-amine sidechains. The compounds allow us to compare the nature by which the substituents are attached. The propionamides are either linked via NH or carbonyl to the anthraquinone moiety *i.e.* the ligands are either amides (RCONH-AQ) or reverse-amides (RNHCO-AQ). These compounds were compared with well characterised di- or tri-substituted anthraquinones, acridines and acridones (see Figure 3.6 & 3.7), which were provided by Prof. S. Neidle (London School of Pharmacy). By studying the effect of these ligands on the stability of several quadruplexes, it was hoped that the effects of each modification could be systematically examined. These included the nature, functionality and position of the terminal amino substituents. This will help to elucidate the structure-activity relationship required for efficient quadruplex targeting.

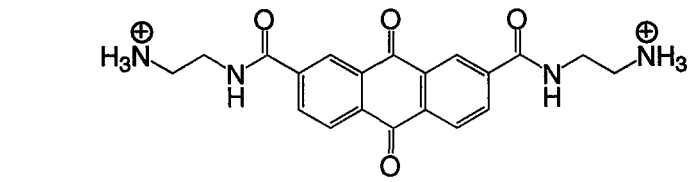
3.3 Results

3.3.1 Preliminary fluorescence melting experiments

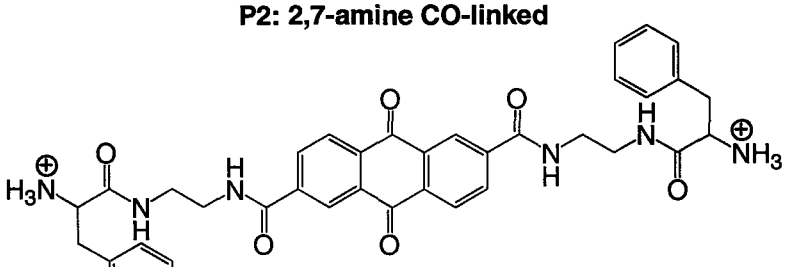
Initial fluorescence melting studies examined the melting and annealing profiles of HT1.5. This oligonucleotide can potentially associate into either dimeric or tetrameric quadruplex complexes. The bimolecular reaction is assumed to be more likely, as at



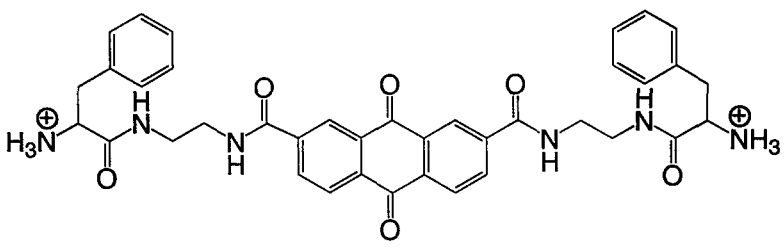
P1: 2,6-amine CO-linked



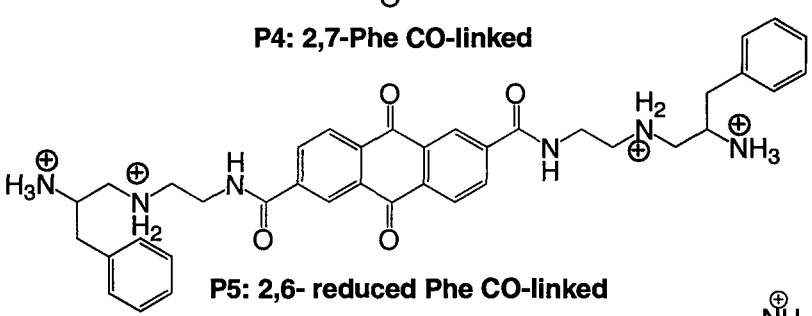
P2: 2,7-amine CO-linked



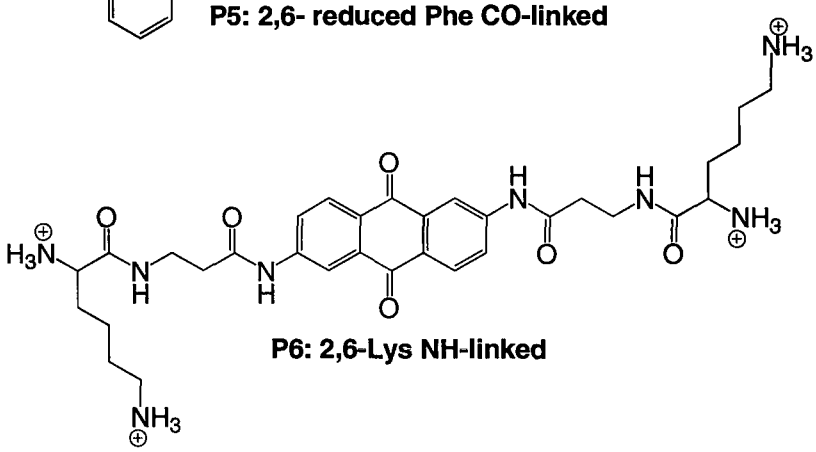
P3: 2,6-Phe CO-linked



P4: 2,7-Phe CO-linked

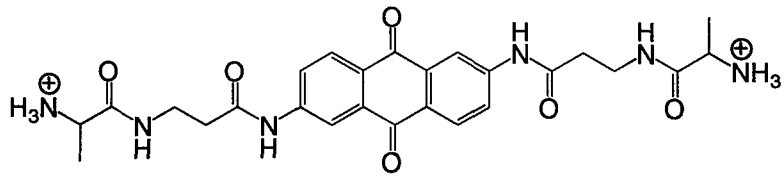


P5: 2,6- reduced Phe CO-linked

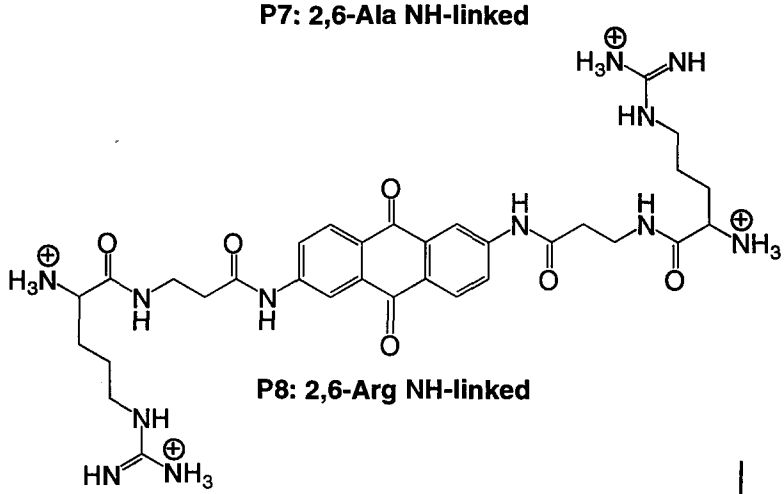


P6: 2,6-Lys NH-linked

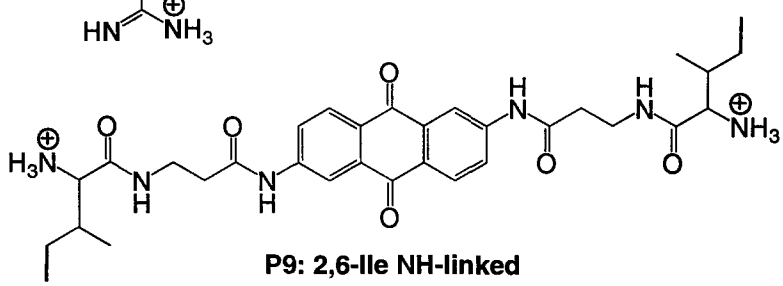
Figure 3.2. Structures of the anthraquinones P1 - P6 (primary and secondary amines are all shown as protonated forms). Positions, nature and position of propionamide sidechains are noted.



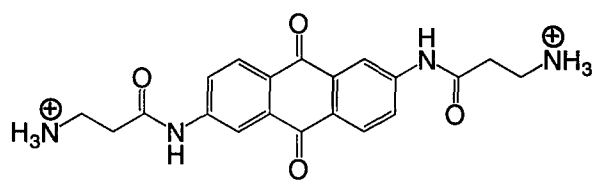
P7: 2,6-Ala NH-linked



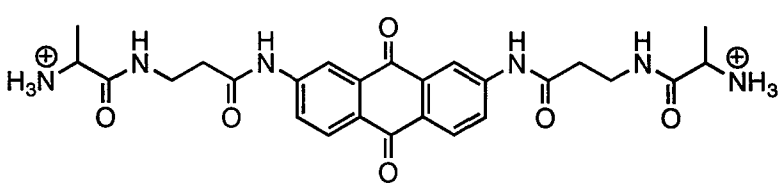
P8: 2,6-Arg NH-linked



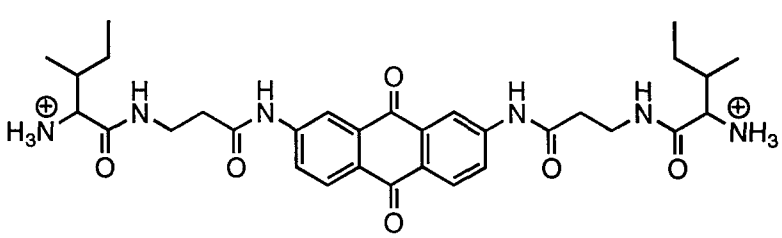
P9: 2,6-Ile NH-linked



P10: 2,6-amine NH-linked

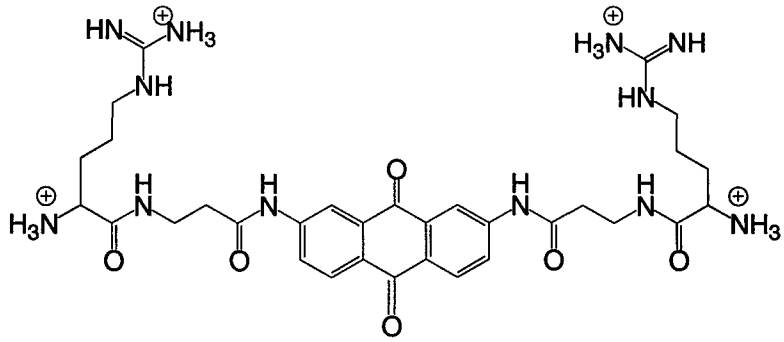


P11: 2,6-Ala NH-linked

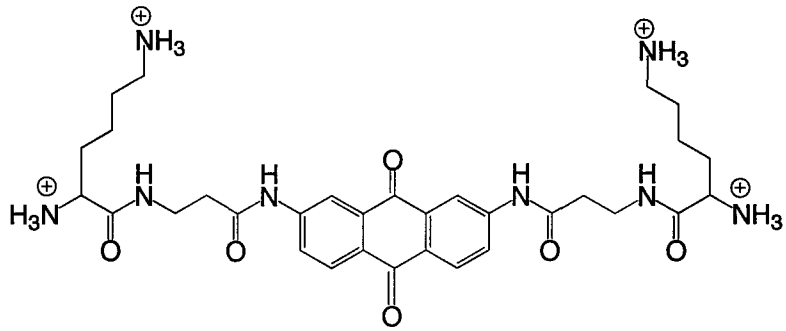


P12: 2,7-Ile NH-linked

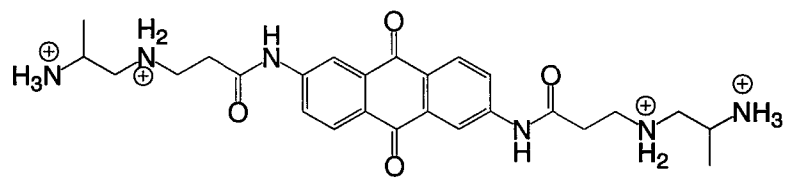
Figure 3.3. Structures of the anthraquinones P7 - P12. (primary and secondary amines are all shown as protonated forms). Positions, nature and position of propionamide sidechains are noted.



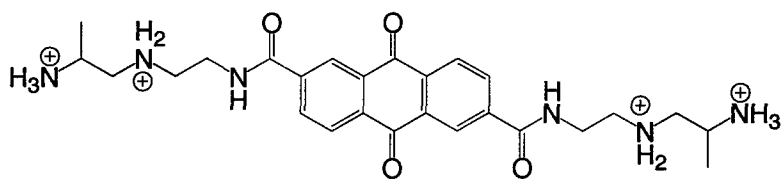
P13: 2,7-Arg NH-linked



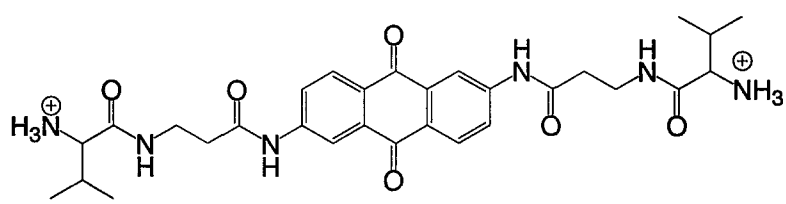
P14: 2,7-Lys NH-linked



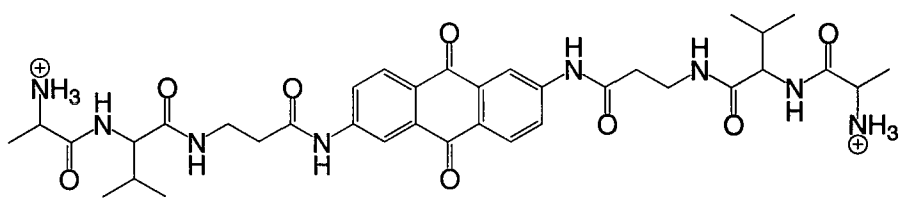
P15: 2,7-Reduced Ala NH-linked



P16: 2,7-Reduced Ala CO-linked

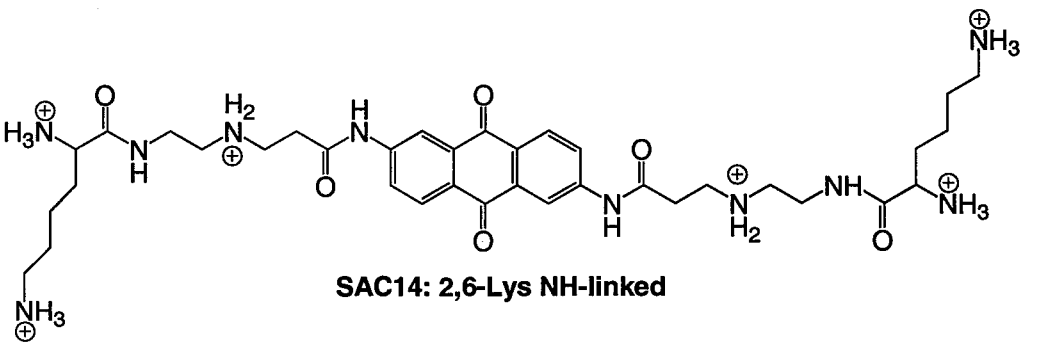


D1: 2,7-Val NH-linked

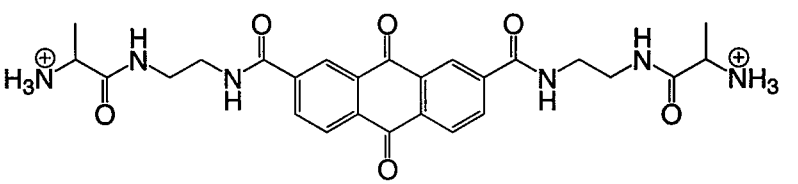


D2: 2,7-Val-Ala NH-linked

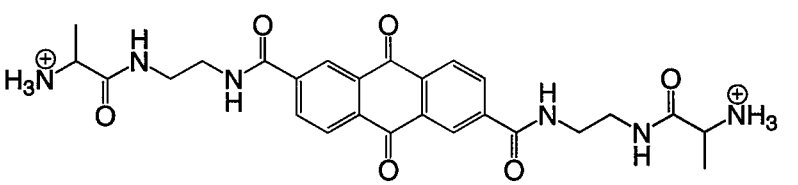
Figure 3.4. Structures of the anthraquinones P11 - P16, D1 and D2 (primary and secondary amines are all shown as protonated forms). Position, nature and linkage of propionamide sidechains are noted.



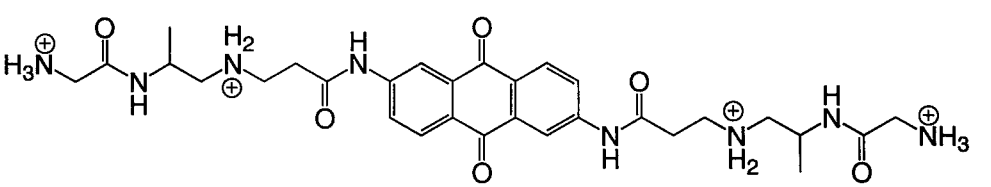
SAC14: 2,6-Lys NH-linked



SAC27: 2,7-Ala CO-linked

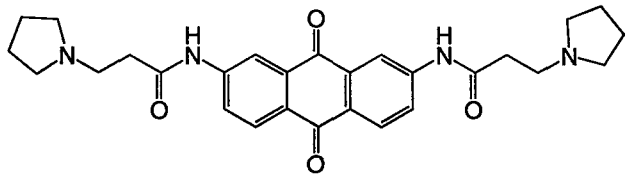


SAC28: 2,6-Ala CO-linked

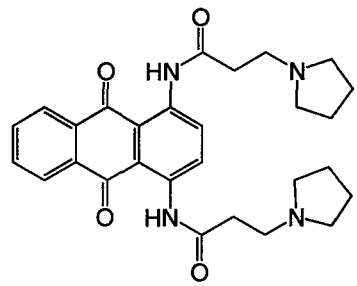


SAC49: 2,6-Gly NH-linked

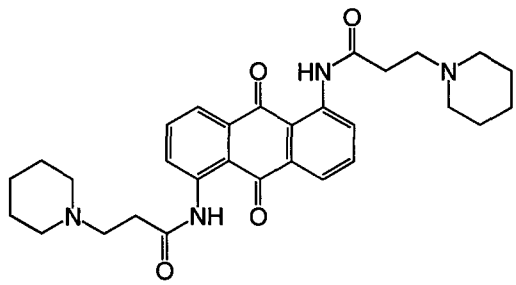
Figure 3.5. Structures of the anthraquinones SAC 14, SAC27, SAC28, SAC49 (primary and secondary amines are all shown as protonated forms).Positions, nature and linkage of propionamide sidechains are noted.



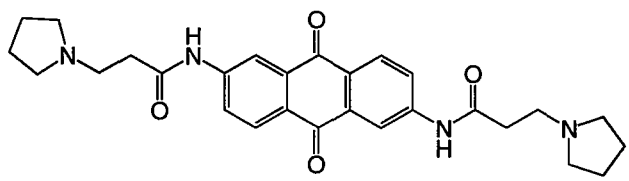
BSU9060: 2,7-AQ



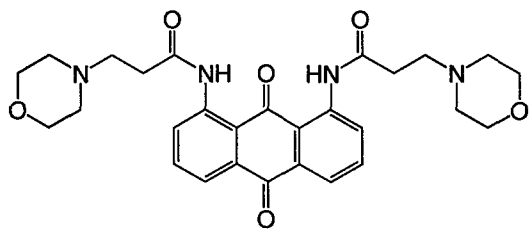
BSU1074: 1,4-AQ



BSU9010: 1,5-AQ

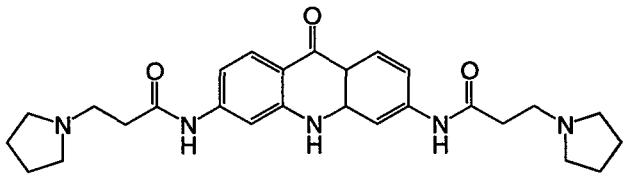


BSU1078: 2,6- AQ

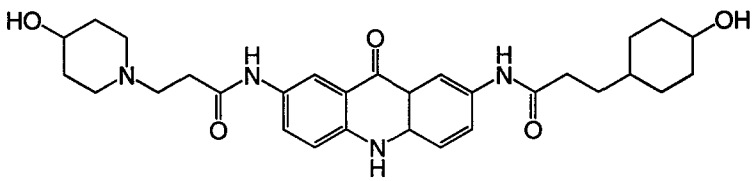


BSU9048: 1,8-AQ

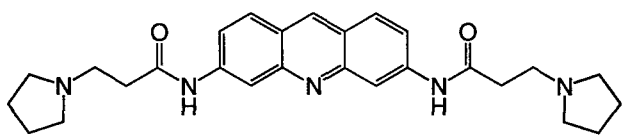
Figure 3.6. Structures of the ligands BSU9060, BSU1074, BSU9010, BSU1078, BSU9048. Position, nature of end groups of propionamide sidechains are noted.



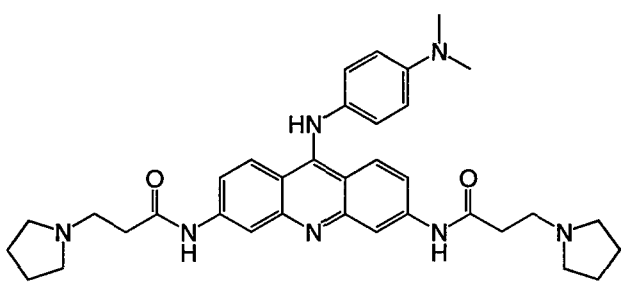
BRACO4: 3,6-ACO



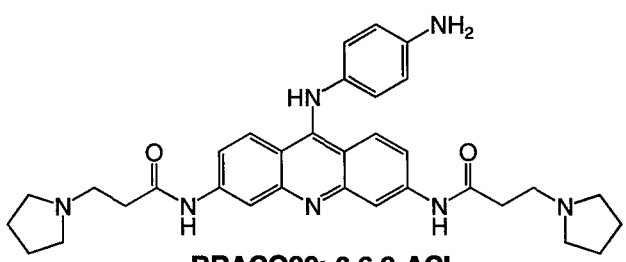
JMACO9: 2,7-ACO



BSU6039: 3,6-ACI



BRACO19: 3,6,9-ACI



BRACO20: 3,6,9-ACI

Figure 3.7. Structures of known ACOs BRACO4, JMACO9; ACIs BSU6039, BRACO19 and BRACO20. Position and nature of endgroups of propionamide sidechains are noted.

0.25 μM oligonucleotide, a second rather than fourth order reaction (which is known to be extremely slow) is likely to be more favourable.

Examples of initial annealing and melting profiles for this oligonucleotide are shown in Figure 3.8. It can be seen that these do not follow the expected sigmoidal annealing and melting curves in either sodium or potassium-containing buffers. Inspection of the stock solutions revealed that some of the oligonucleotide had precipitated while stored. Although we cannot be sure about the origin of this precipitate, it may be related to the formation of a very stable aggregate of the oligonucleotide, so-called 'G-wires' (Marsh & Henderson, 1994). Subsequent oligonucleotide preparations (see below) were always checked for the presence of any precipitate and were filtered through a 0.2 μM filter. Since these results are also not consistent with literature reports using similar non-fluorescent oligonucleotide, we reconsidered the oligonucleotide synthesis, as deprotection of guanines in long G-tracts can be a problem.

Inadequate deprotection of the exocylic amine group of guanine during the synthesis of HT1.5 could possibly explain the absence of quadruplex melting and annealing. The lower panels of Figure 3.8 show that addition of the quadruplex-binding ligand BRACO19 altered the unusual profiles. Increasing concentrations of drug shifted the broad transitions to higher temperatures. This suggests that though likely to be far less favourable, Hoogsteen interactions between unprotected Gs may give rise to G-tetrad structures (and quadruplex folding), if the addition of a active drug be sufficiently stabilising. After discussion with Oswel DNA service, oligonucleotides were prepared therefore using a different protecting group (fast-G), and were deprotected under harsher conditions (55°C in ammonia overnight). Indeed, these changes generated oligonucleotides which produced more typical melting profiles, and these are used in the experiments described below.

Re-synthesis of the HT1.5 oligonucleotide (by Oswel DNA service, Southampton) produced melting with improvement. Melting studies were carried out first in either high concentrations of potassium or sodium phosphate buffers. Other buffers were tested to try to increase the usability of the fluorescence data produced. Buffers used included the low and high sodium and potassium forms of phosphate, HEPES, and tris-

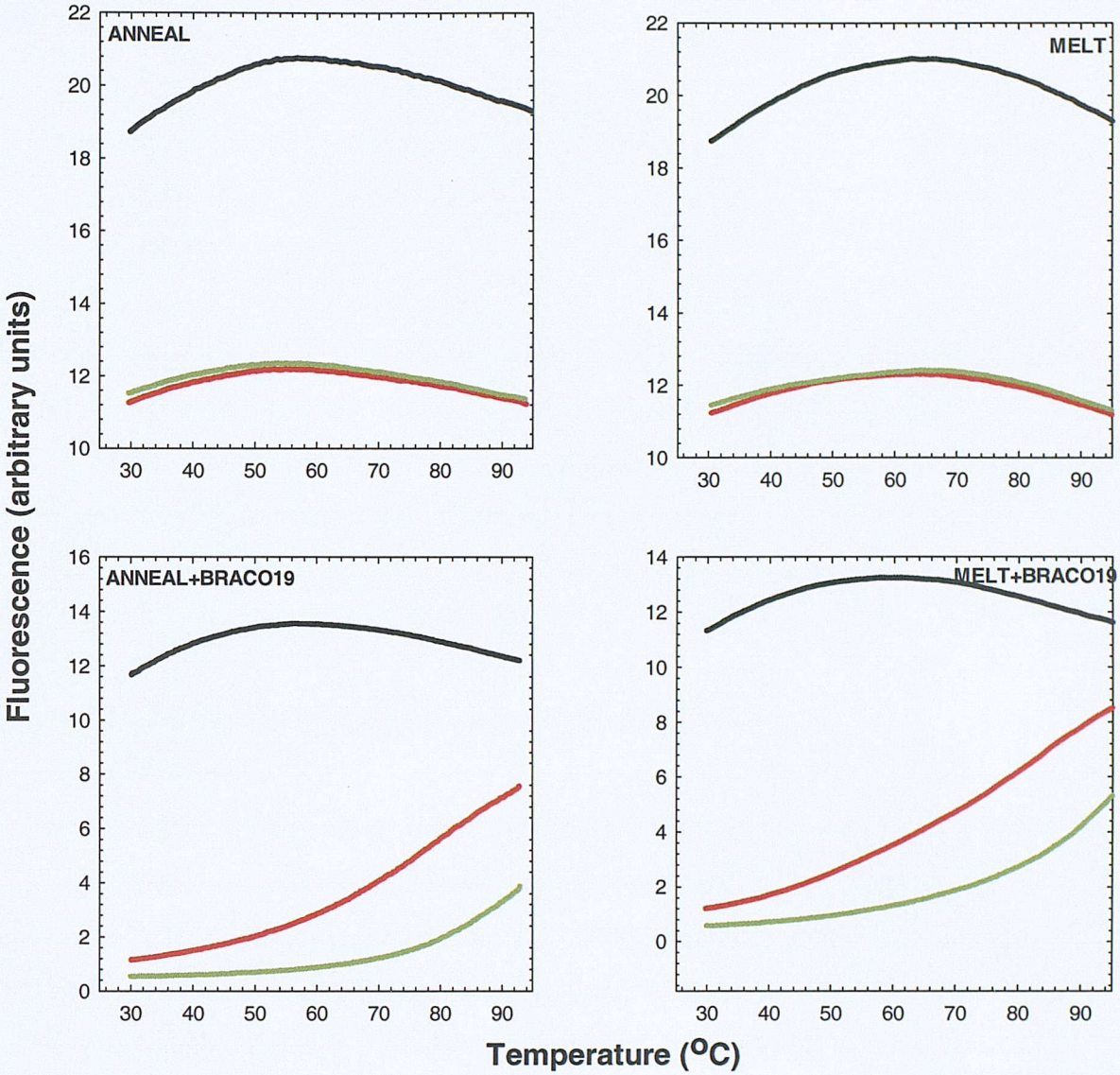


Figure 3.8. Examples of annealing (left) and melting (right) profiles obtained with HT1.5 in 10 mM KP (black), 50 mM KP (red) and 100 mM KP (green) are shown in the upper panels. Lower panels show the effect of BRACO19 on HT1.5 in 50 mM KP profiles. Plots show BRACO19 concentrations as control (black), 5 μM (red) and 10 μM (green).

EDTA buffers, all with additional NaCl or KCl. Typical profiles produced under the various conditions are represented in Figure 3.9. Finally, none were decided to be superior to 50 mM potassium phosphate buffer (*i.e.* ~ 90 mM K^+), which also has the most stable pK_a over a wide temperature range, and is close to intracellular K^+ concentration.

In these final experiments HT1.5 was used as supplied by Oswel DNA service, without further purification. HPLC purification of short oligonucleotides such as HT1.5 is difficult, HT1.5 was re-purified by polyacrylamide gel electrophoresis. This reduced the noise in the melting curves but did not improve the transition. Concentrating HT1.5 or adding non-fluorescent HT1.5 DNA to promote quadruplex formation did not produce any improvement to the profiles.

3.3.2 Fluorescence melting in the absence of ligand

Choice of buffer

Having overcome the initial difficulties with HT1.5, the KP buffer was chosen as the best for quadruplex-ligand studies with this sequence and HT3.5. Each sample was first melted to 95°C before annealing to 30°C, then repeatedly melted and annealed. The first melting curves were disregarded for interpretation, though it is noteworthy that these were often different from the second melting curve where the oligonucleotide was more properly annealed. These typically showed abnormal biphasic melts and asymmetry in presence of some ligands (data not shown). All data analysed is for the annealing profile first, followed by a melting profile.

HT1.5 profiles

As can be seen in upper panels of Figure 3.10 (upper panels) the annealing and melting curves produced by HT1.5 in 50 mM KP are shallow and broad, this is especially true for the annealing profile. The melting and annealing profiles also fail to coincide, the annealing curve is significantly shifted to lower temperatures. This results from hysteresis at the rate of heating and cooling of 6°C min^{-1}, due to the slow kinetics of quadruplex formation by HT1.5 (detailed in the discussion). The T_m values for the annealing and melting curves do not change on adding to excess, increasing concentrations of non-fluorescent oligonucleotide (data not shown). The lack of any effect is puzzling and the reasons are unclear.

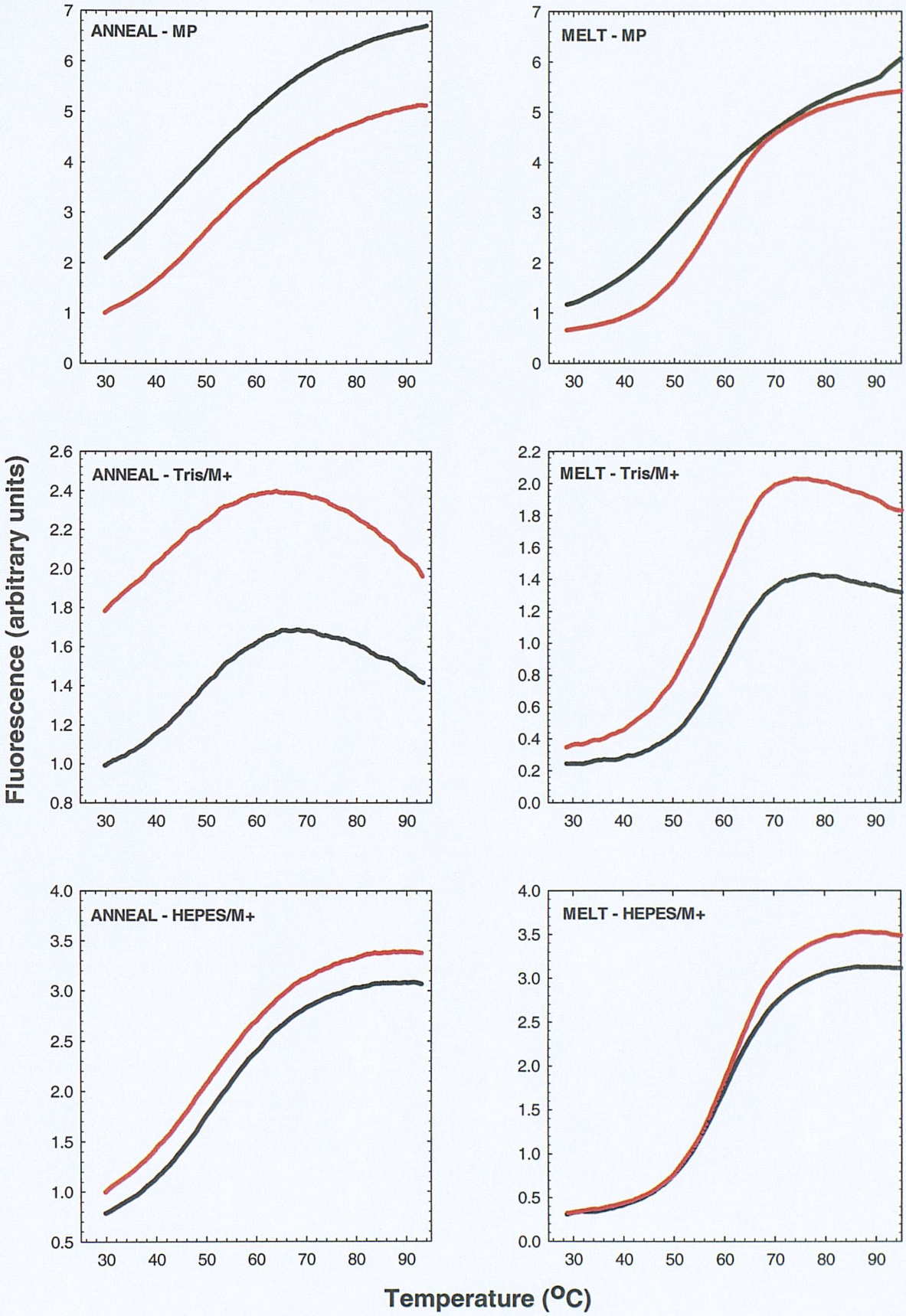


Figure 3.9. Examples of annealing and melting profiles obtained with HT1.5 under different buffer conditions. For each buffer 50 mM K+ are shown in red and 50 mM Na+ are in black.

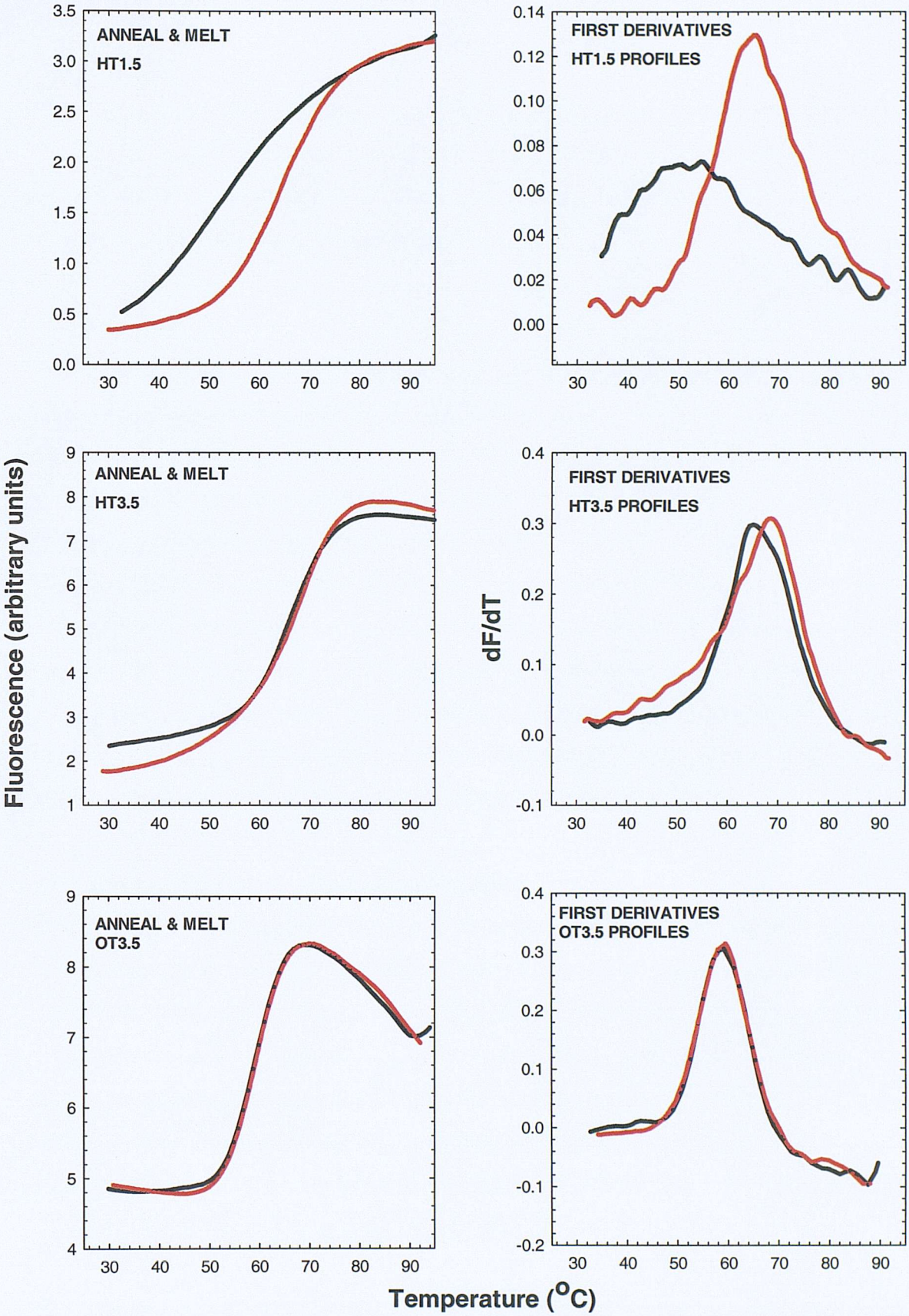


Figure 3.10. Examples of annealing (black) and melting (red) profiles (left panels), accompanied by first derivatives of the curves (right panels). Profiles obtained with HT1.5 in 50 mM KP at 6 °C min^{-1} are in the upper panels; and those with HT3.5 also in 50 mM KP at 6 °C min^{-1} are in the middle panels. OT3.5 profiles were obtained in 10mM NaP + 90mM NaCl at the **slower rate** of 0.05 °C min^{-1}, these are in the lower panels.

HT3.5 profiles

The annealing and melting curves produced by HT3.5 under the standard conditions (50 mM KP, $dF/dT = 6°C\ min^{-1}$) are far more useful. Annealing and melting curves (see middle panels of Figure 3.10) effectively coincide, suggesting very little or no hysteresis and transitions are much sharper.

OT3.5 profiles

This sequence, like HT3.5, most probably forms intramolecular quadruplex. As shown in the lower panels of Figure 3.10, quadruplex formation by OT3.5 produces annealing and melting profiles which also show very sharp transitions. However, these profiles are more greatly affected by hysteresis than either HT3.5 or HT1.5. A much slower rate of heating and cooling is therefore needed, as such the OT3.5 profiles shown are not at $6°C\ min^{-1}$ but $0.05°C\ min^{-1}$. This reflects the very slow kinetics of quadruplex formation by the *Oxytricha* telomeric repeat (further discussed in chapter 4). Also, these results suggest that the hysteresis and the shallowness of the profiles are not related to a similar property of the sequences. To overcome this exceptional hysteresis and the high stability of OT3.5 quadruplex, the cationic buffer used, the ionic strength of the buffer, and the rate of heating were all systematically studied (see chapter 4).

3.3.3 Effects of ligands on fluorescence melting

In the experiments described above it is clear that hysteresis between annealing and melting curves is a problem for HT1.5 and OT3.5, when melting experiments are performed under standard conditions ($6°C\ min^{-1}$). Nonetheless, it was decided to investigate the effect of ligands on the melting profiles under these conditions. This allowed a high throughput comparison of the relative quadruplex affinities for a range of compounds. The hysteresis itself also gives some information on ligand binding and as described below, may vary between different ligands. In addition, some experiments with OT3.5 were performed at much slower rates of heating and cooling.

The effects of each member of the series of novel anthraquinone conjugates were systematically investigated. These studies were first performed with the more relevant human (HT) sequences. In each case, drug concentrations of between 0.5 to 10 μM were added to 0.25 μM of the fluorescent oligonucleotides. Higher concentrations will

be inappropriate as they are unlikely to be physiologically relevant. The results described below focus on the effects of the ligands on the annealing and melting T_m values. However, two other effects should also be noted. Firstly, the ligands sometimes change the shape (slope) of the melting transitions, making them broader at higher ligand concentrations. This effect will be considered further in the discussion, but may be caused by multiple drug-binding sites and/or heterogeneity in the quadruplex structures. Secondly, the ligands often affected the total fluorescence signal, quenching the fluorescence throughout the melting profiles. This effects is also considered in the discussion. The fluorescence profiles have not been normalised in order to emphasise this effect.

3.3.4 HT1.5 fluorescence melting experiments

Figures 3.11 to 3.14 show representative annealing and melting curves for HT1.5 in the presence of different concentrations of the a range of G4 ligands with varying activities and the T_m values for all compounds (at 1 and 10 μM) are summarised in Table 3.1. As described above, the oligonucleotide displays considerable hysteresis between the melting ($T_m \approx 64$°C) and annealing ($T_m \approx 51$°C) profiles. However, this hysteresis is still evident in the presence of all the ligands, and active compounds such as P15 and P16, increase the annealing temperature more than the melting (thereby reducing the amount of hysteresis). This can be seen from Figure 3.15, which shows the ΔT_m obtained with HT1.5 under increasing concentrations of several active ligands. This may suggest that these ligands facilitate quadruplex assembly and increase the slow rate of quadruplex formation. Further examples of this can be seen in Table 3.1, which show that many compounds including P8, P10, P15 and P16 have clear effects and stabilise the quadruplex at concentration as low as 1 μM. These have even greater effects at 10 μM, at which concentration for example, P15 produces a ΔT_m of over 25°C with no hysteresis between the melting and annealing profiles. Several other compounds also have good activity at this concentration producing ΔT_m (anneal) greater than 10°C (including P5, P6, P16, BRACO19, BRACO20). Other ligands are devoid of appreciable activity (P4, P12, P14, D2, BSU9048, BRACO4, JMACO9, BSU6039).

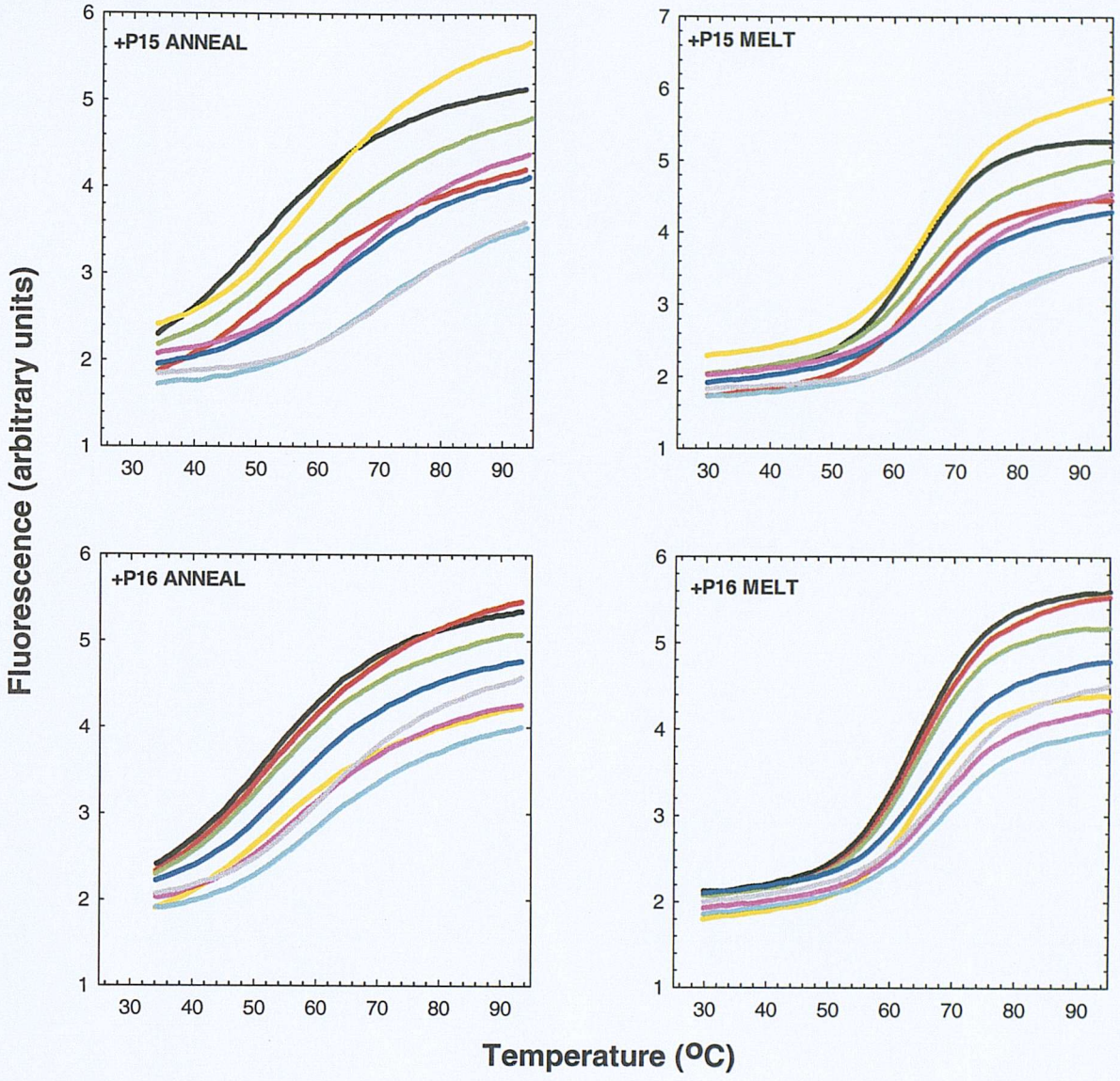


Figure 3.11. Examples of addition of active G4 ligands on anneal (left panels) and melt profiles (right panels) obtained with HT1.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

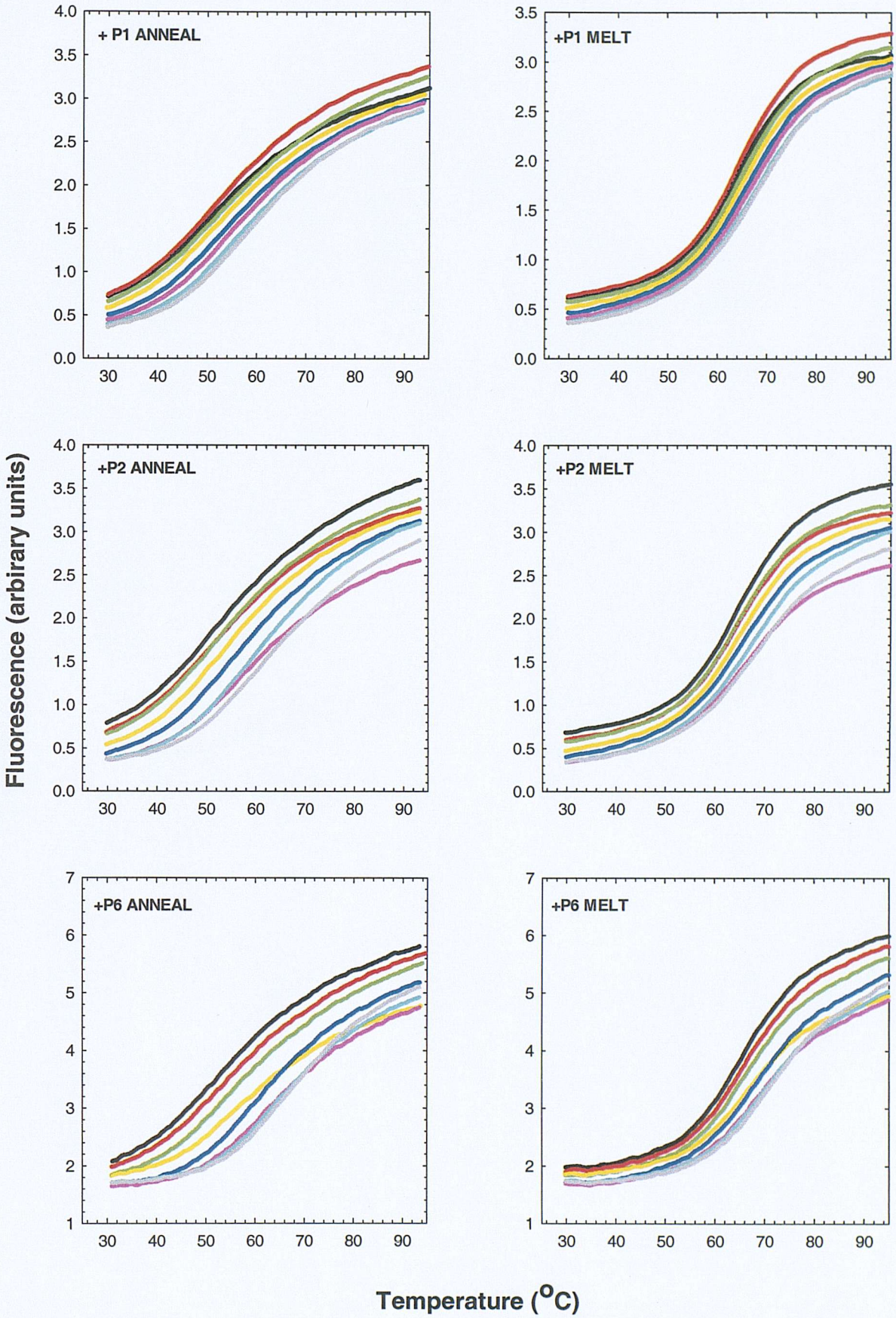


Figure 3.12. Examples of addition of modest G4 ligands on annealing (left panels) and melting profiles (right panels) obtained with HT1.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

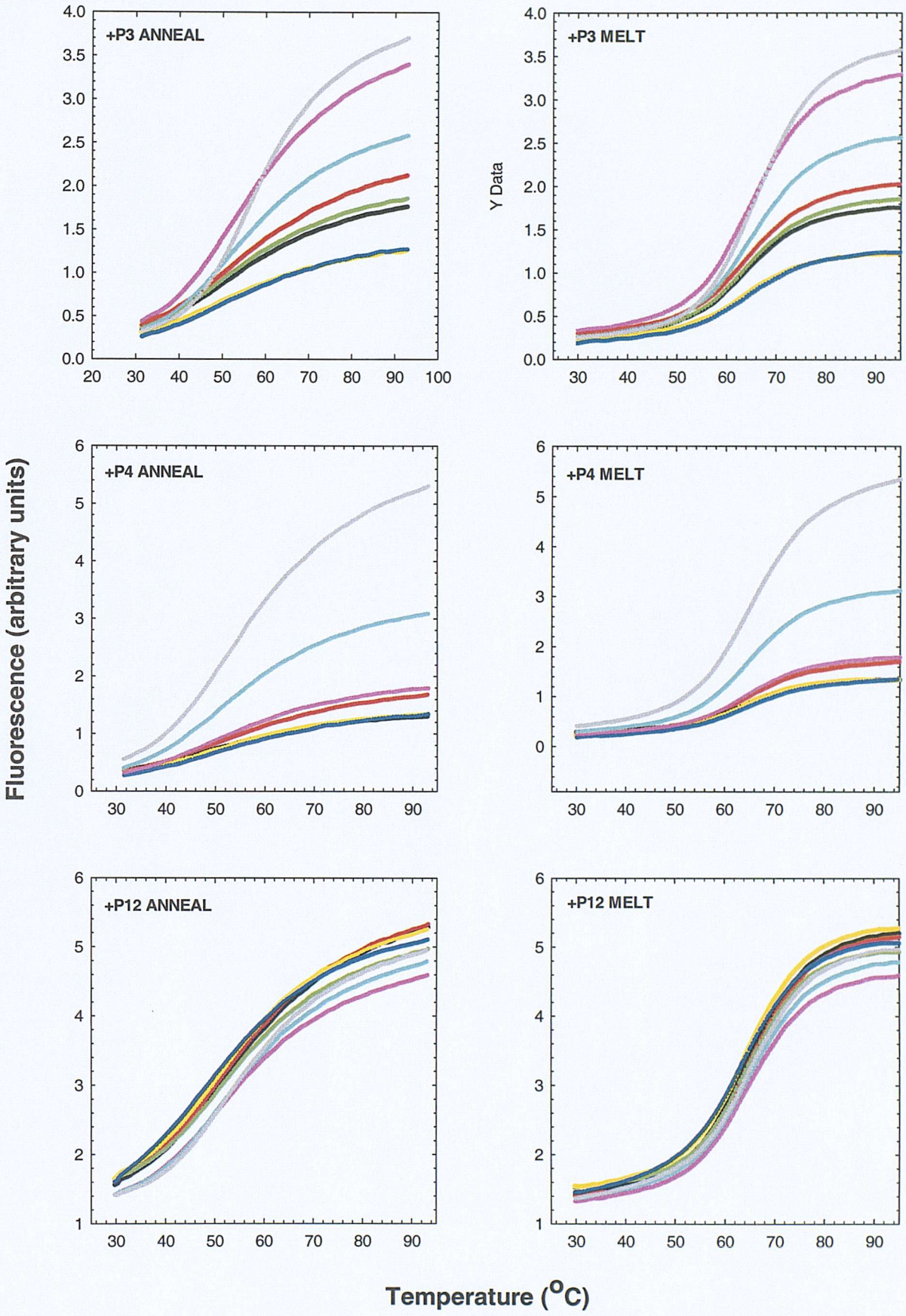


Figure 3.13. Examples of addition of poor G4 ligands on annealing (left panels) and melting profiles (right panels) obtained with HT1.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

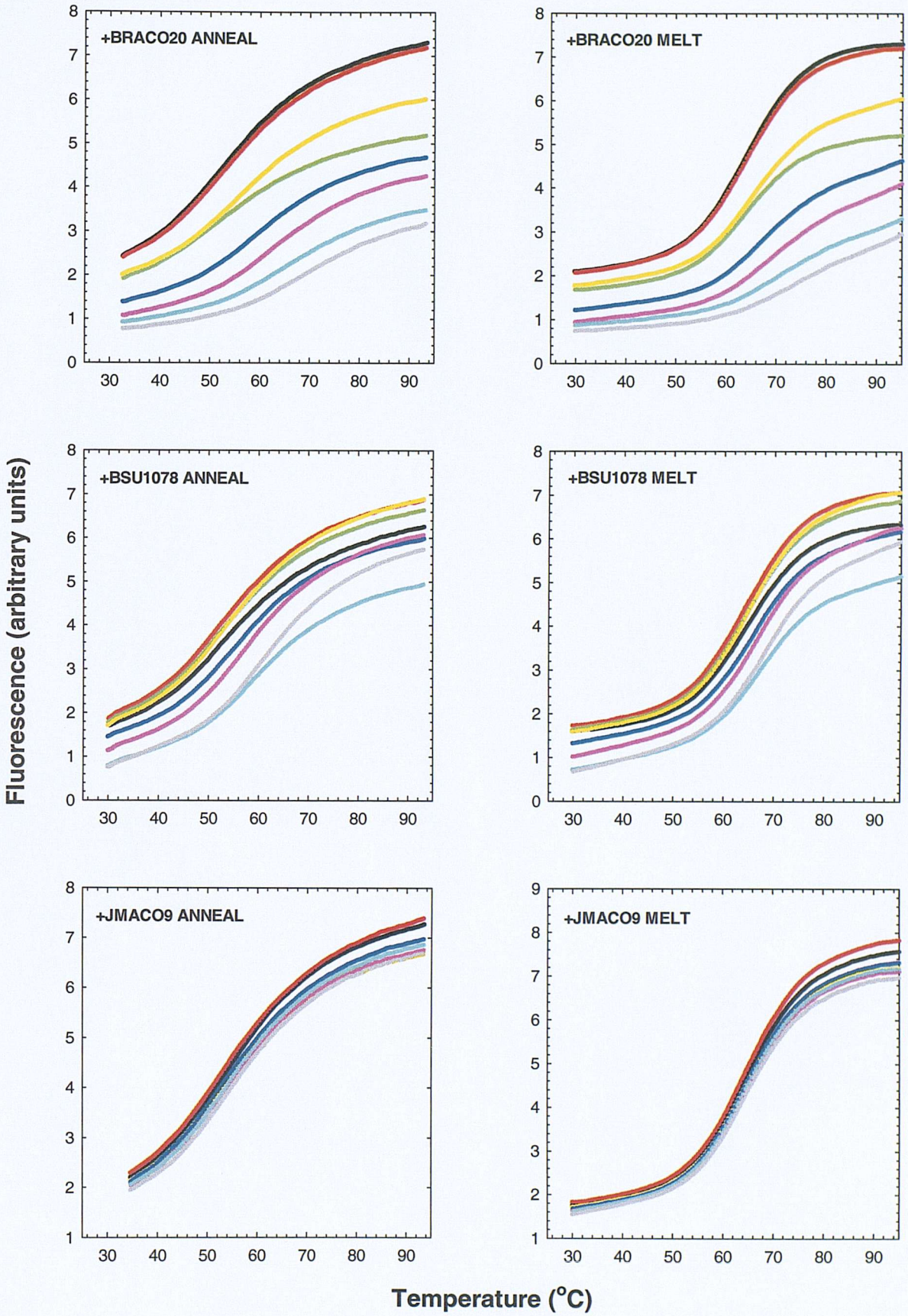


Figure 3.14. Examples of addition of known nucleic acid ligands on annealing (left panels) and melting profiles (right panels) obtained with HT1.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

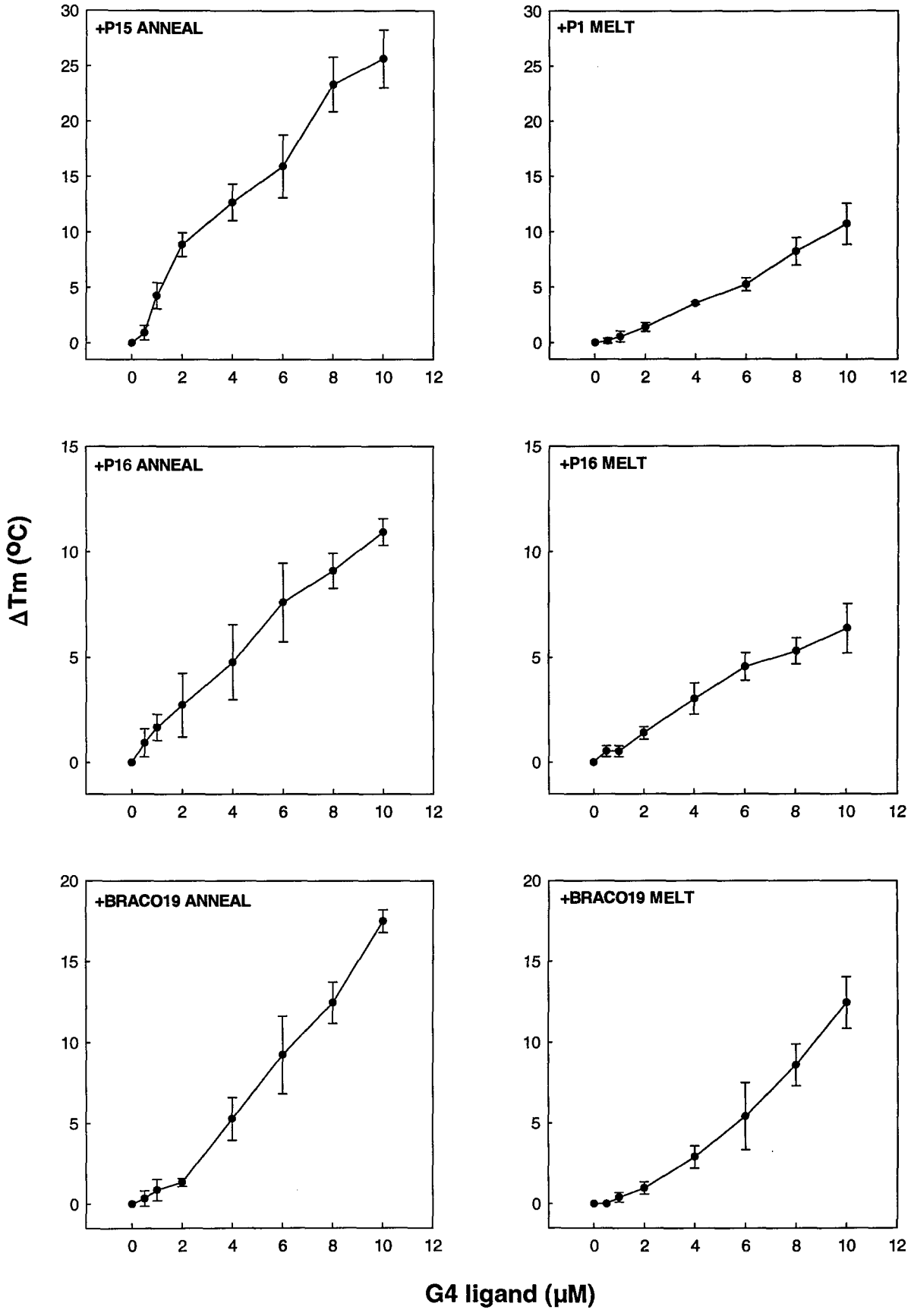


Figure 3.15. ΔT_m values from annealing (left) and melting (right) profiles with HT1.5 produced by several G4 ligands. Both axes are to the same scale.

Table 3.1. Summary of ligand-induced increases in the transitions of fluorescence (annealing and melting) profiles representing intermolecular quadruplex formation by HT1.5 oligonucleotide.

Ligand	ΔT_m anneal (°C)		ΔT_m melt (°C)		Hysterisis (°C)[a]	
	1 μM	10 μM	1 μM	10 μM	1 μM	10 μM
Control	51.3 ± 0.7		64.0 ± 0.5		12.7	
P1	1.2 ± 0.9	7.5 ± 1.1	0.4 ± 0.5	4.1 ± 0.6	12.0	9.3
P2	0.3 ± 0.4	9.2 ± 1.0	0.2 ± 0.2	4.8 ± 1.1	12.6	8.3
P3	0.6 ± 0.6	3.8 ± 1.3	1.5 ± 0.3	2.4 ± 0.6	0.9	11.3
P4	0.6 ± 0.6	1.0 ± 1.0	0	1.0 ± 0.1	12.1	12.7
P5	0	10.3 ± 0.5	1.3 ± 0.2	5.6 ± 1.8	14.0	7.9
P6	0.8 ± 1.1	13.9 ± 3.0	0.1 ± 0.1	5.6 ± 0.1	12.0	4.4
P7	0.3 ± 0.5	9.6 ± 3.7	0.2 ± 0.2	2.2 ± 1.5	12.5	5.3
P8	2.3 ± 1.6	12.6 ± 5.1	0.4 ± 2.3	4.3 ± 0.2	10.8	4.4
P9	1.0 ± 0.7	4.8 ± 1.2	0.4 ± 0.5	2.1 ± 0.7	12.1	10.0
P10	2.1 ± 1.5	7.9 ± 3.5	0.5 ± 0.4	3.5 ± 0.7	11.1	8.3
P11	1.1 ± 0.1	6.8 ± 0.7	0.1 ± 0.2	1.6 ± 0.1	11.8	7.5
P12	0.6 ± 0.8	1.9 ± 0.2	0.6 ± 0.5	1.3 ± 0.7	12.8	12.1
P13	0.1 ± 0.1	3.6 ± 2.4	0	1.8 ± 1.2	12.6	10.9
P14	0.6 ± 0.5	1.7 ± 0.7	0.2 ± 0.3	0.2 ± 0.3	12.3	11.2
P15	4.3 ± 1.2	25.7 ± 2.6	0.6 ± 0.5	10.8 ± 1.9	9.0	-2.2
P16	1.7 ± 0.6	10.9 ± 0.6	0.5 ± 0.3	6.4 ± 1.2	11.6	8.1
D1	0.2 ± 0.3	5.9 ± 2.1	0.2 ± 0.2	1.6 ± 0.7	12.7	8.5
D2	0.3 ± 0.5	0.3 ± 0.5	0	0	12.4	12.4
BSU9060	0.8 ± 0.7	7.6 ± 1.6	1.0 ± 0.6	3.5 ± 1.6	12.9	8.6
BSU1074	1.2 ± 0.9	8.8 ± 0.8	0.6 ± 0.2	4.8 ± 0.2	12.1	8.7
BSU9010	0.7 ± 0.3	5.2 ± 1.2	0.3 ± 0.5	2.1 ± 0.6	12.3	9.6
BSU1078	0.7 ± 0.3	8.8 ± 1.6	0.1 ± 0.1	4.4 ± 1.0	12.1	8.2
BSU9048	1.2 ± 0.4	1.7 ± 0.9	0.8 ± 0	0.8 ± 0	12.3	12.6
BRACO4	1.6 ± 1.6	1.8 ± 1.8	0.5 ± 0.5	0.5 ± 0.5	12.6	11.4
JMACO9	0	1.1 ± 1.1	0.2 ± 0.2	0.2 ± 0.2	12.9	11.9
BSU6039	0.3 ± 0.3	1.0 ± 0.2	0.4 ± 0.4	1.4 ± 0.9	12.9	13.1
BRACO19	0.9 ± 0.7	17.5 ± 0.7	0.4 ± 0.3	12.5 ± 1.6	12.2	7.7
BRACO20	1.0 ± 0.6	16.1 ± 3.0	0.4 ± 0.3	9.5 ± 2.0	12.2	6.1

Control values are the average ± stardard error to 1d.p from all experiments, while ligand values are those from three independent experiments. **a** – values for the hysterisis is defined as the *T*m melt minus *T*m anneal.

3.3.5 HT3.5 fluorescence melting experiments

Figures 3.16 to 3.19 show the results of similar experiments with HT3.5. Figure 3.20 shows the ΔT_m derived from these for several active compounds over a range of concentrations, while ΔT_m values are summarised in Table 3.2. It can be seen that the ligands have larger effects on the T_m of this sequence than HT1.5. In addition, the ligands appear to affect melting more than annealing. This presumably occurs because these intramolecular complexes fold up relatively quickly and the ligands slow the rate of dissociation. The greatest activity is seen with compounds P8, P15, SAC14, BRACO19 and BRACO20, for which 10 μM ligand stabilises the quadruplex so to such an extent that no melting is observed even at the highest temperatures used (for example, see SAC14 in the bottom panel of Figure 3.16). With 1 μM ligand the greatest effects are produced by P10 and BRACO19. In contrast, P3, P4, P7, P9, P14, D1, D2, SAC27, SAC28 SAC49 and BSU9048 all have very little effect on the melting profiles.

3.3.6 OT3.5 fluorescence melting experiments

The novel anthraquinones P6, P8, P15, and P16 which showed the best activity on the human HT sequences were selected to determine their effects on the *Oxytricha* oligonucleotide, OT3.5. Whereas the previous work was carried out in 50 mM potassium phosphate and at rate of temperature change of 6°C min^{-1}, experiments in this case were performed in 10mM sodium phosphate and 90 mM sodium chloride and at several rates of heating and cooling (0.05, 0.5 and 6°C min^{-1}). These changes to the protocol were introduced because of the exceptionally slow rates of association and dissociation of this quadruplex structure and the substantial hysteresis that is observed, especially at low ionic strength (as described in chapter 4). Also, quadruplexes containing runs of four guanines will be more stable than those with three guanines.

Annealing and melting profiles for each of the four active G4 ligands under different rates of heating and cooling are represented in Figures 3.21 to 3.24, while T_m values are given in Tables 3.3 to 3.5. Looking first at the T_m values for the oligonucleotide alone, it can be seen that there is considerable hysteresis between the annealing and melting curves at the fast rates of heating. The hysteresis is still apparent when the samples are heated at 0.5°C min^{-1}, but the annealing and melting curves coincide at 0.05°C min^{-1}. Addition of the ligands increases the T_m, as expected, though the changes are different at the different rates of heating and cooling. At the fastest rate of heating, the ligands

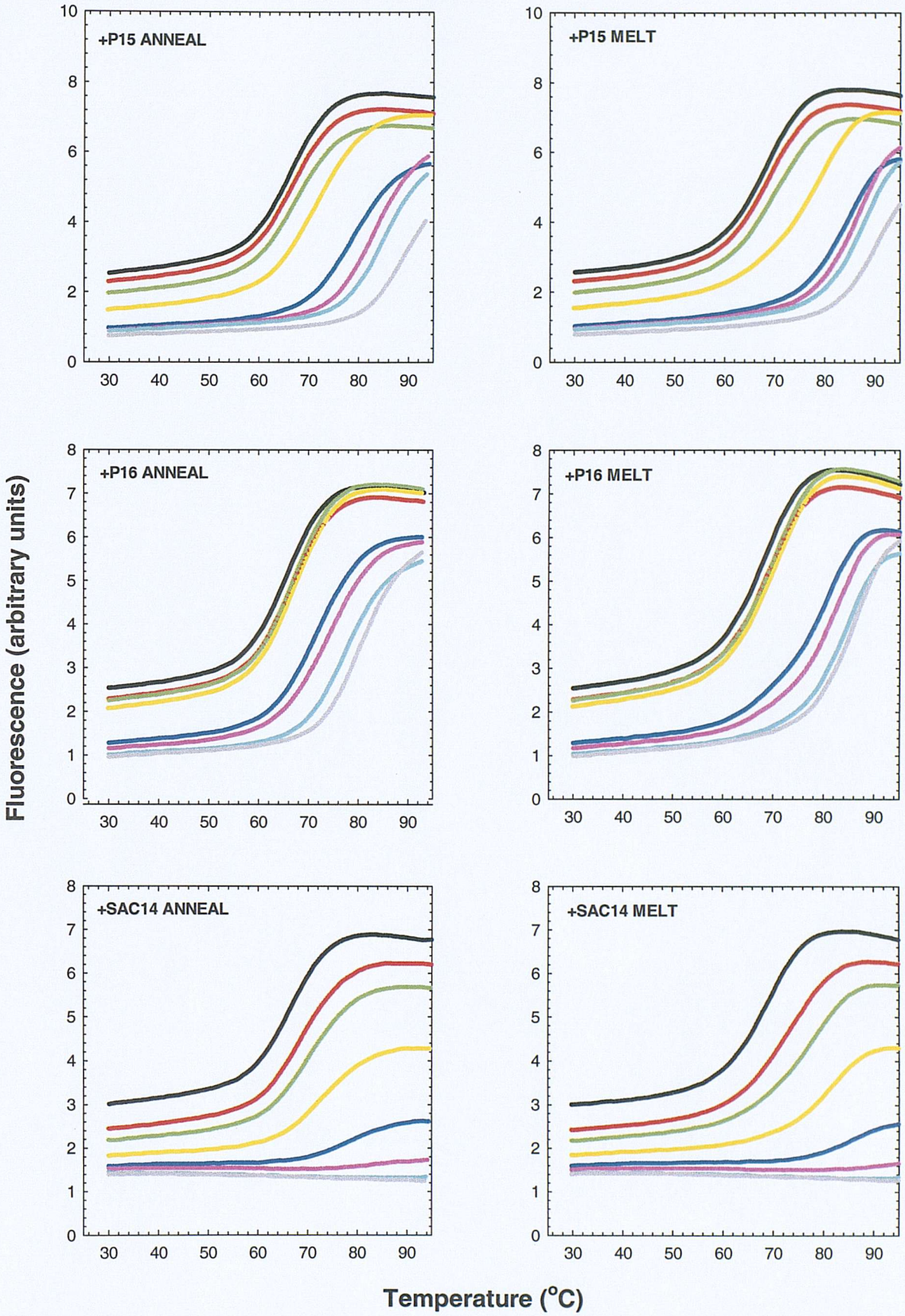


Figure 3.16. Examples of addition of active G4 ligands on annealing (left panels) and melting profiles (right panels) obtained with HT3.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

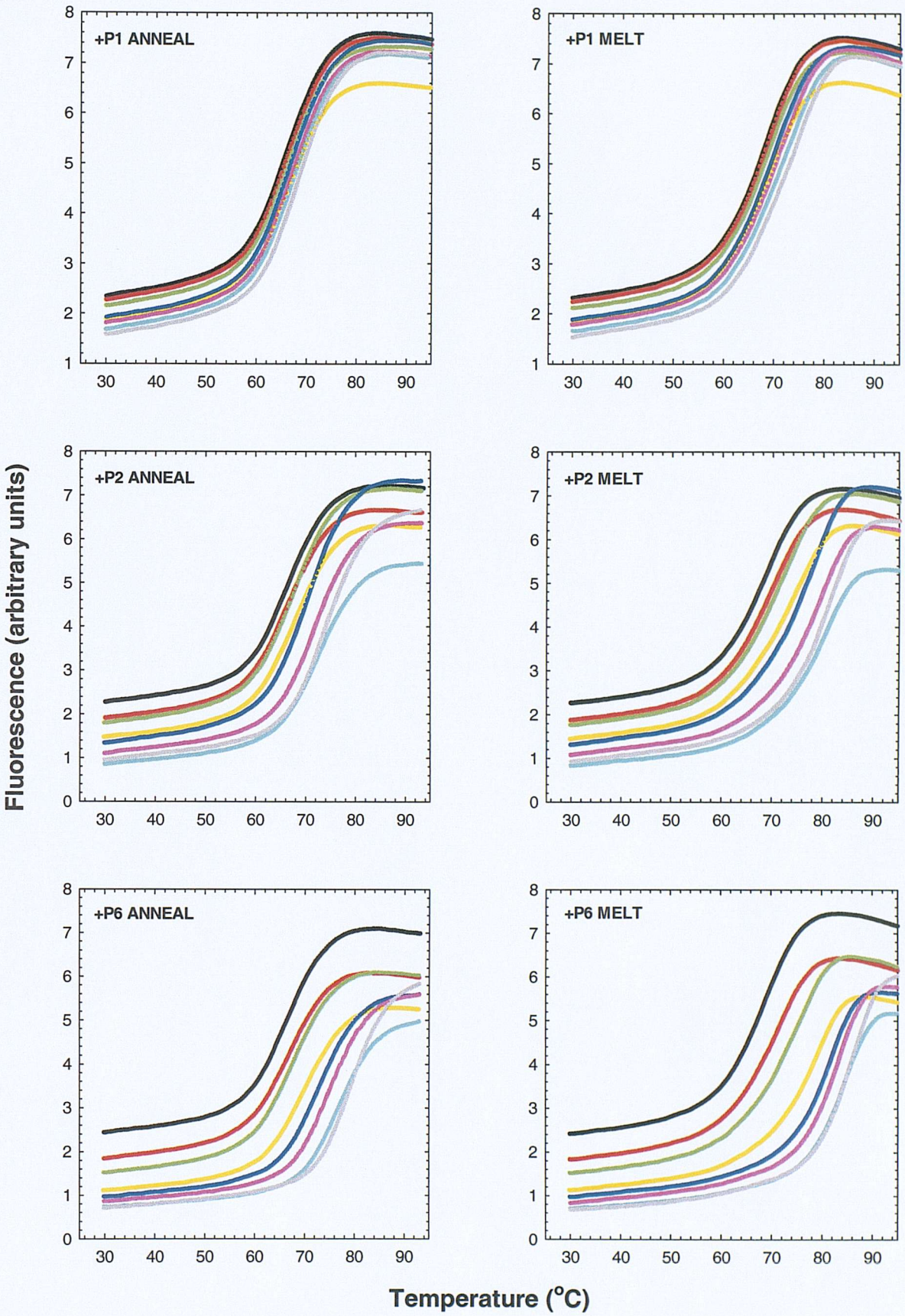


Figure 3.17. Examples of addition of modest G4 ligands on annealing (left panels) and melting profiles (right panels) obtained with HT3.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

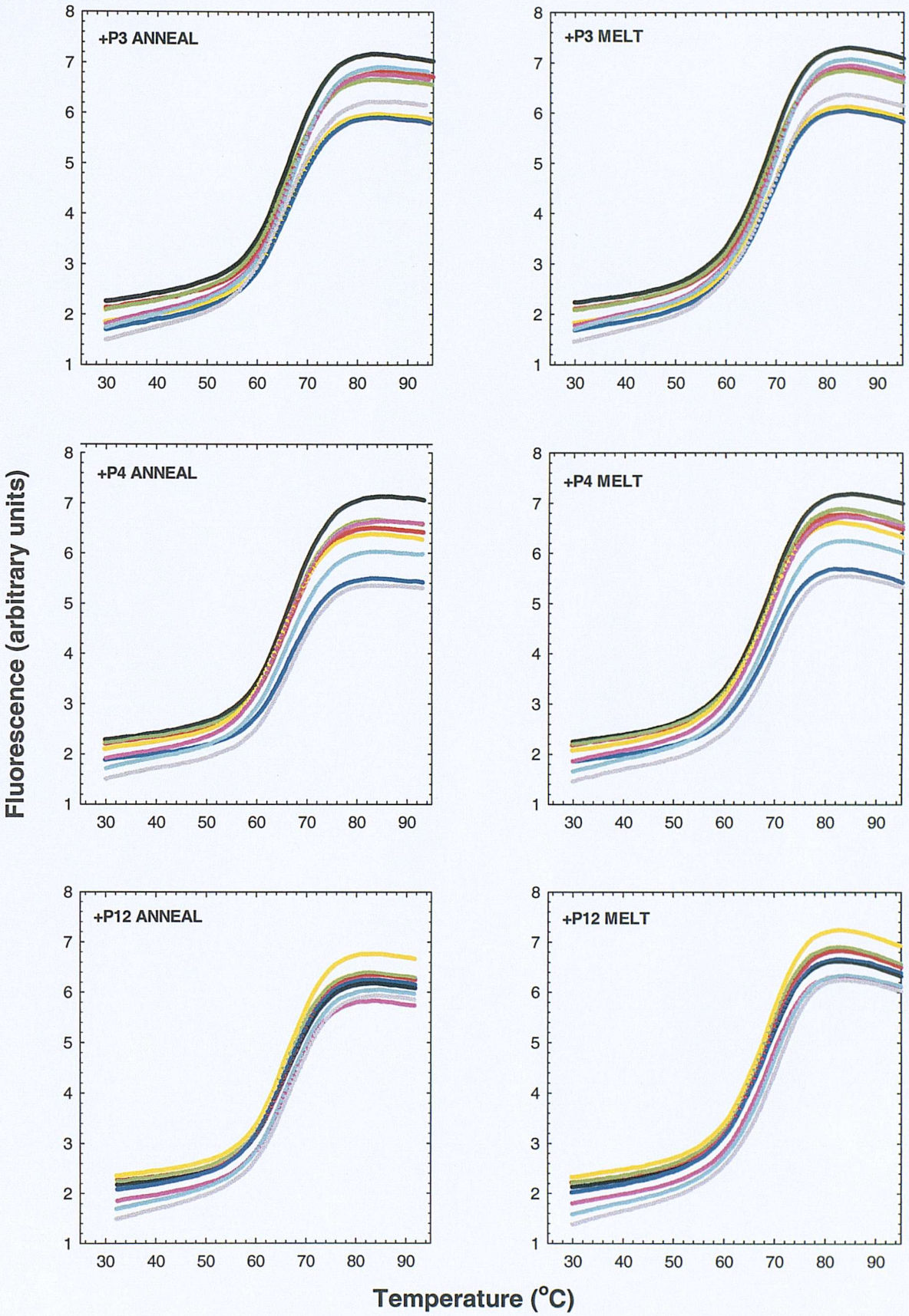


Figure 3.18. Examples of addition of poor G4 ligands on annealing (left panels) and melting profiles (right panels) obtained with HT3.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

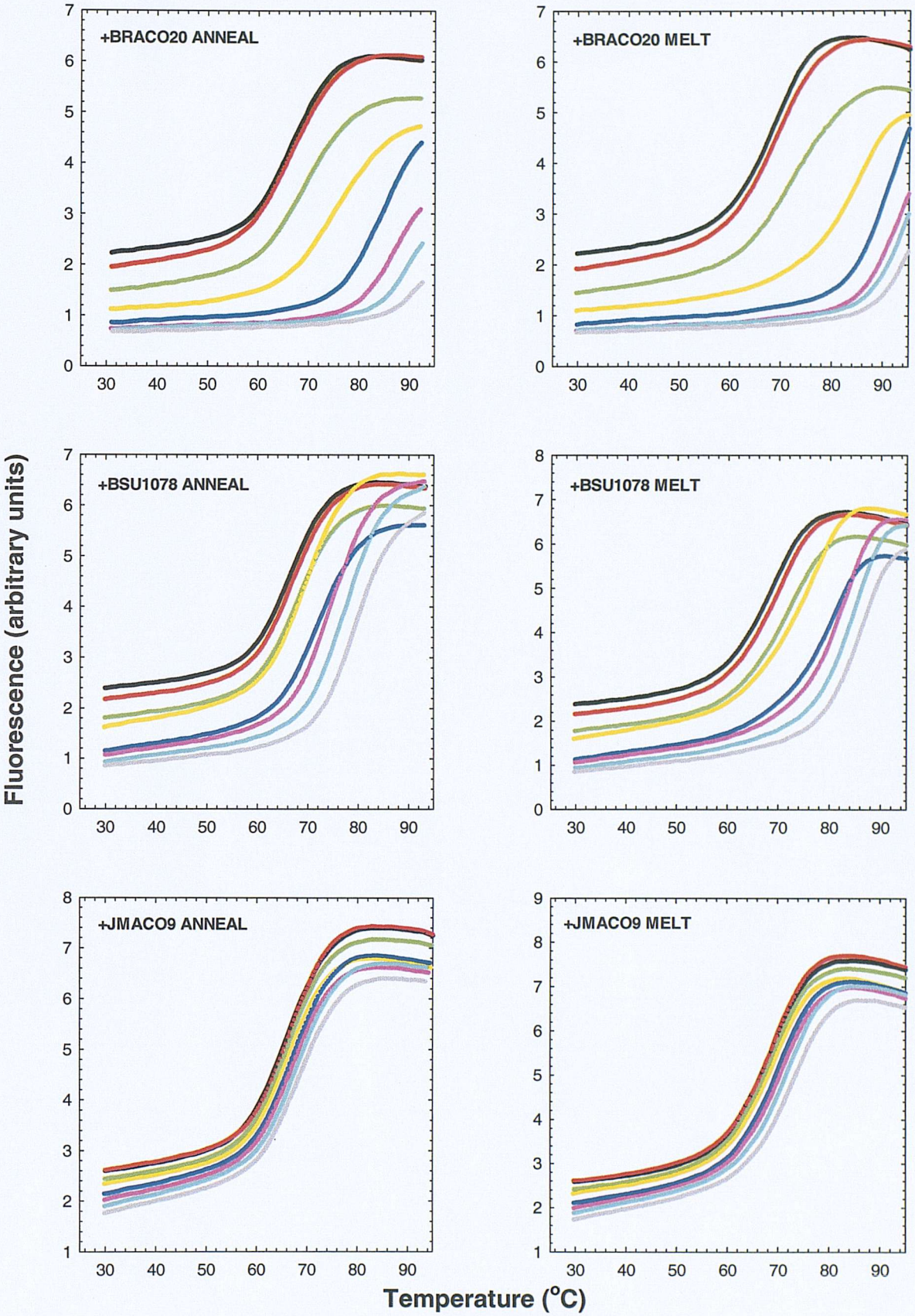


Figure 3.19. Examples of addition of nucleic acid ligands, on annealing (left panels) and melting profiles (right panels) obtained with HT3.5. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

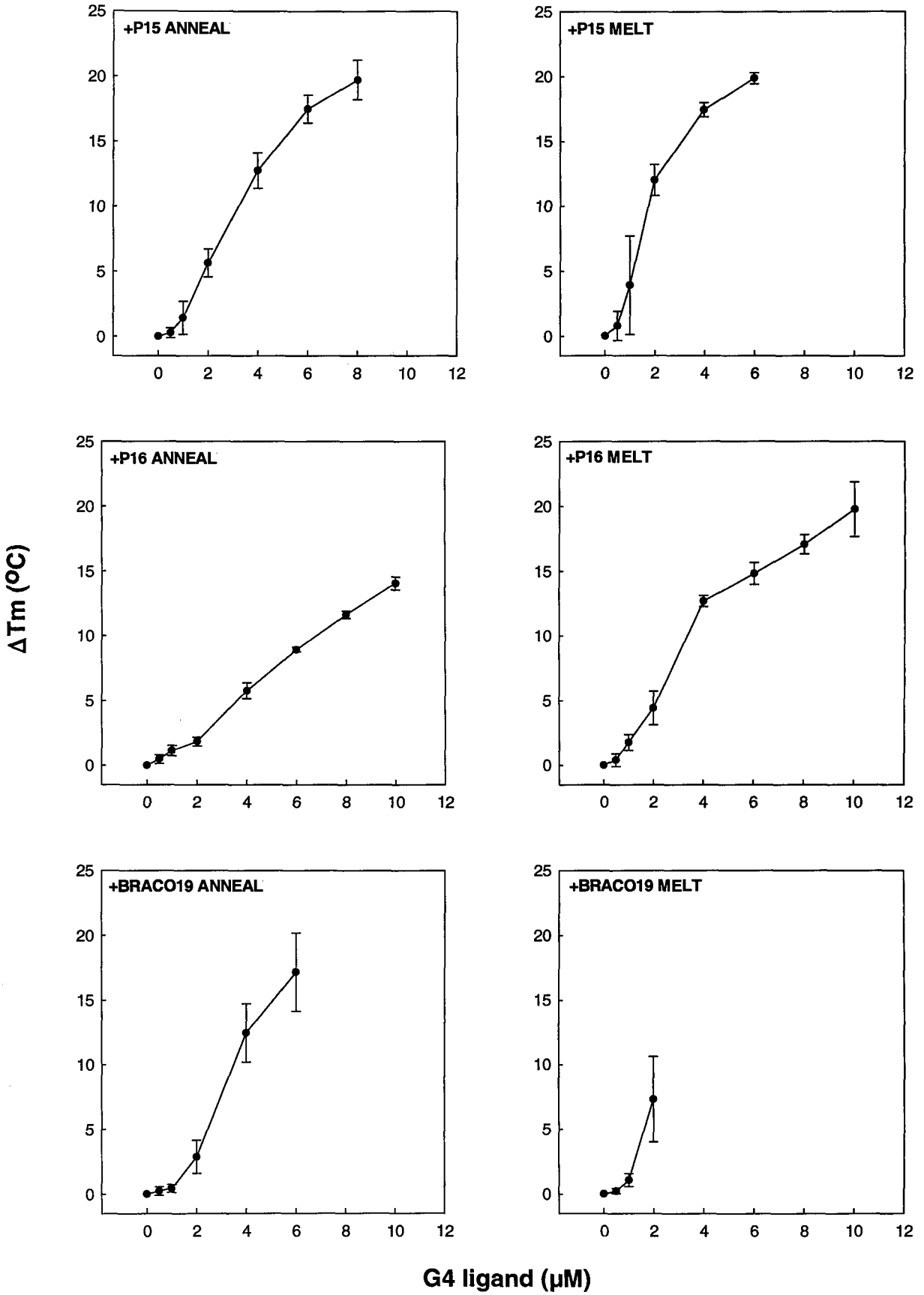


Figure 3.20. ΔT_m values from annealing (left) and melting (right) profiles with HT3.5 produced by several G4 ligands. Both axes are to the same scale, where ligand concentrations are not shown to 10 μM, the ΔT_m values are too great (>25°C) to be determined.

Table 3.2. Summary of the ligand-induced increases in the transitions of fluorescence (annealing and melting) profiles representing formation of intramolecular quadruplex by HT3.5 oligonucleotide.

Ligand	ΔT_m anneal (°C)		ΔT_m melt (°C)		Hysterisis (°C) [a]	
	1 μM	10 μM	1 μM	10 μM	1 μM	10 μM
Control	66.4 ± 0.3		68.8 ± 0.3		2.4	
P1	0.9 ± 0.7	5.3 ± 1.8	2.0 ± 1.1	9.0 ± 4.0	3.5	6.0
P2	0.6 ± 0.2	7.4 ± 0.3	3.3 ± 0.2	12.9 ± 0.1	5.1	7.9
P3	0.2 ± 0.2	0.5 ± 0.2	0.2 ± 0.2	1.1 ± 0.4	2.5	3.0
P4	0	0.3 ± 0.3	0.1 ± 0.1	1.0 ± 0.5	2.5	3.1
P5	0.5 ± 0.5	4.8 ± 2.3	0.7 ± 0.5	10.1 ± 2.2	2.6	7.7
P6	1.4 ± 0.5	12.5 ± 0.7	6.6 ± 0.4	17.2 ± 0.3	7.6	7.1
P7	0	2.9 ± 0.1	0.2 ± 0.3	8.2 ± 0.3	2.6	7.7
P8	2.0 ± 1.1	19.2 ± 1.8	4.8 ± 2.3	>25 (>8 μM) [b]	5.2	n.d.
P9	0	1.7 ± 0.3	0	5.7 ± 0.2	2.4	6.4
P10	3.7 ± 0.9	14.3 ± 1.9	9.1 ± 0.6	18.1 ± 0.3	7.7	6.2
P11	0.4 ± 0.3	3.7 ± 0.6	0.3 ± 0.2	9.0 ± 0.6	2.3	7.8
P12	0.4 ± 0.3	0.6 ± 0.2	0.2 ± 0.2	2.8 ± 0.6	2.2	4.6
P13	0.4 ± 0.5	7.4 ± 1.7	0.6 ± 0.7	13.3 ± 1.7	2.6	8.2
P14	0	1.7 ± 0.3	0	7.0 ± 1.2	2.4	7.7
P15	1.4 ± 1.3	>25 (>8 μM) [b]	3.9 ± 3.8	>25 (>6 μM) [b]	4.9	n.d.
P16	1.1 ± 0.4	14.1 ± 0.5	1.8 ± 0.6	19.8 ± 2.1	3.0	8.1
D1	0	1.5 ± 0.4	0.1 ± 0.1	6.0 ± 0.7	2.5	6.9
D2	-0.1 ± 0.3	0	0.3 ± 0.5	1.2 ± 0.2	2.9	3.7
SAC14	2.9 ± 0.6	>25 (>4 μM) [b]	6.1 ± 2.1	>25 (>4 μM) [b]	5.6	n.d.
SAC27	0.3 ± 0.3	1.0 ± 0.2	0.5 ± 0.3	3.9 ± 0.6	2.6	5.3
SAC28	0.2 ± 0.3	1.4 ± 0.3	0.4 ± 0.5	4.4 ± 0.9	2.6	5.5
SAC49	0.1 ± 0.1	0.8 ± 0.5	0	0.8 ± 0.5	2.3	4.4
BSU9060	0	6.5 ± 0.4	0.3 ± 0.5	13.3 ± 0.3	2.7	9.2
BSU1074	2.4 ± 0.4		5.9 ± 0.3		5.9	
BSU9010	0.8 ± 0.4	7.0 ± 0.1	1.8 ± 1.0	>25 (>6 μM) [b]	3.4	n.d.
BSU1078	1.3 ± 0.8	11.5 ± 1.5	2.3 ± 0.9	16.2 ± 1.2	3.4	7.1
BSU9048	0.3 ± 0.2	0.8 ± 0.5	0.1 ± 0.1	2.0 ± 1.7	2.2	3.7
BRACO4	0.4 ± 0.1	3.3 ± 0.6	0.6 ± 0.1	5.9 ± 1.0	2.6	5.0
JMACO9	0.6 ± 0.3	2.6 ± 0.5	0.2 ± 0.1	4.1 ± 0.5	2.8	3.9
BSU6039	1.1 ± 0.5	6.9 ± 1.2	0.7 ± 0.7	12.1 ± 1.9	2.7	7.6
BRACO19	0.4 ± 0.3	>25 (>6 μM) [b]	1.1 ± 0.5	>25 (>2 μM) [b]	3.0	n.d.
BRACO20	3.6 ± 0.8	>25 (>4 μM)	8.1 ± 5.0	>25 (>2 μM)	6.9	n.d.

Control values are the average ± stardard error to 1d.p from all experiments, while ligand values are those from three independent experiments. **a** – values for the hysterisis is defined as the *T*m melt minus *T*m anneal. **b** – ΔTm were too high (>25 °C) to be determined above the ligand concentration indicated, as such hystersis could not be determined (n.d.)

increase the T_m of melting by more than that of annealing, thereby increasing the hysteresis. This may suggest that, as noted with HT1.5, the ligands have a greater effect on the dissociation of these intramolecular complexes than on their rate of folding. It should also be noted that, at the fastest rates of heating, the analysis is complicated as the melting curve are biphasic, and that this persists in the presences of low ligand concentrations. This is not evident with the slower rates of heating, and so must reflect some kinetic aspect of the process, which will be considered further in chapter 4.

Despite these problems it is clear that all four ligands stabilise the OT3.5 quadruplex. The results consistently show that P8 and P15 are the most active, and that P16 is the least active. It should also be noted that the effects of these ligands on the intramolecular quadruplex formed by *Oxytricha* telomeric repeats are considerably lower than the effects on complexes formed by the human HT sequences (compare ΔT_m values from Tables 3.1 and 3.2 to T_m at different rates in Table 3.3). This may be related to the length of the G-stack, or the ionic conditions, and will be considered further in the discussion.

3.3.7 TR2 (intermolecular) quadruplex bandshifts.

The progression of quadruplex formation by TR2 (which should be a dimerisation reaction) was followed for 30 h (see Figure 3.25). As a very slow second order reaction, the concentration, an oligonucleotide concentration of 25 μM was used (higher than that used in bandshifts analyses of ligand activity). From this time course a second order rate constant of 72 M^{-1} h^{-1} was determined. The novel anthraquinones and the other established G4-binding ligands were subsequently studied with oligonucleotide TR2. Examples of such bandshift analyses, showing the effects of several anthraquinones on the formation of a dimeric complex can be found in Figures 3.26 to 3.29. Anthraquinones have already been reported to bind strongly to quadruplexes, and increasing concentrations of these ligands can be seen to generate a retarded species which is not present in the control lane. The accompanying plots show how the retarded species, expressed as the percent of total oligonucleotide, varies with ligand concentration. In this assay and those described below, the amount of material in the retarded band is always less than 50%, and the data were analysed by estimating the percent quadruplex formation with 20 μM ligand.

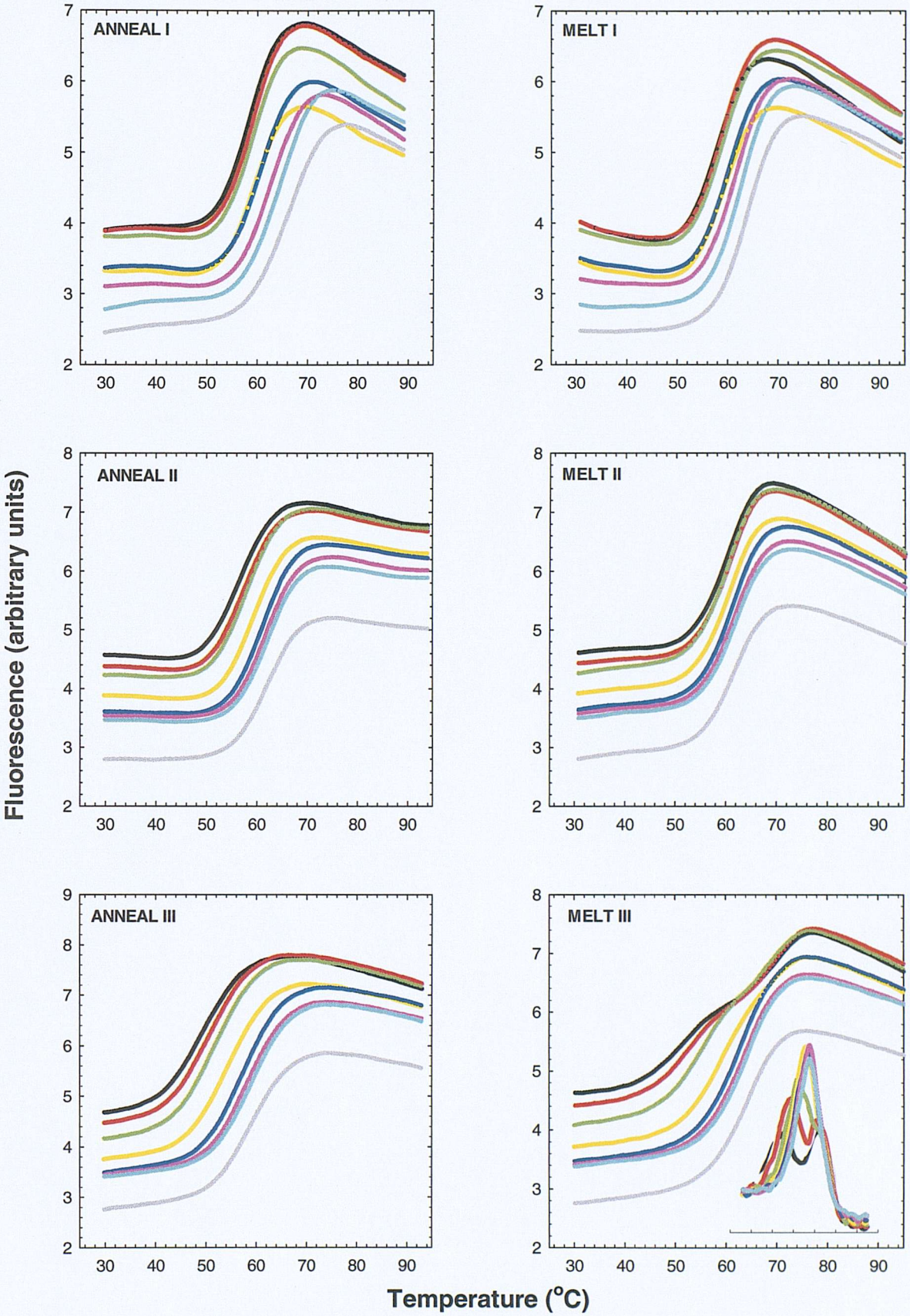


Figure 3.21. Examples of addition of P6 on annealing (left panels) and melting profiles (right panels) obtained OT3.5 at slow (I), intermediate (II) and fast (III) rates of heating and cooling. Inset within melt III is its first derivative. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

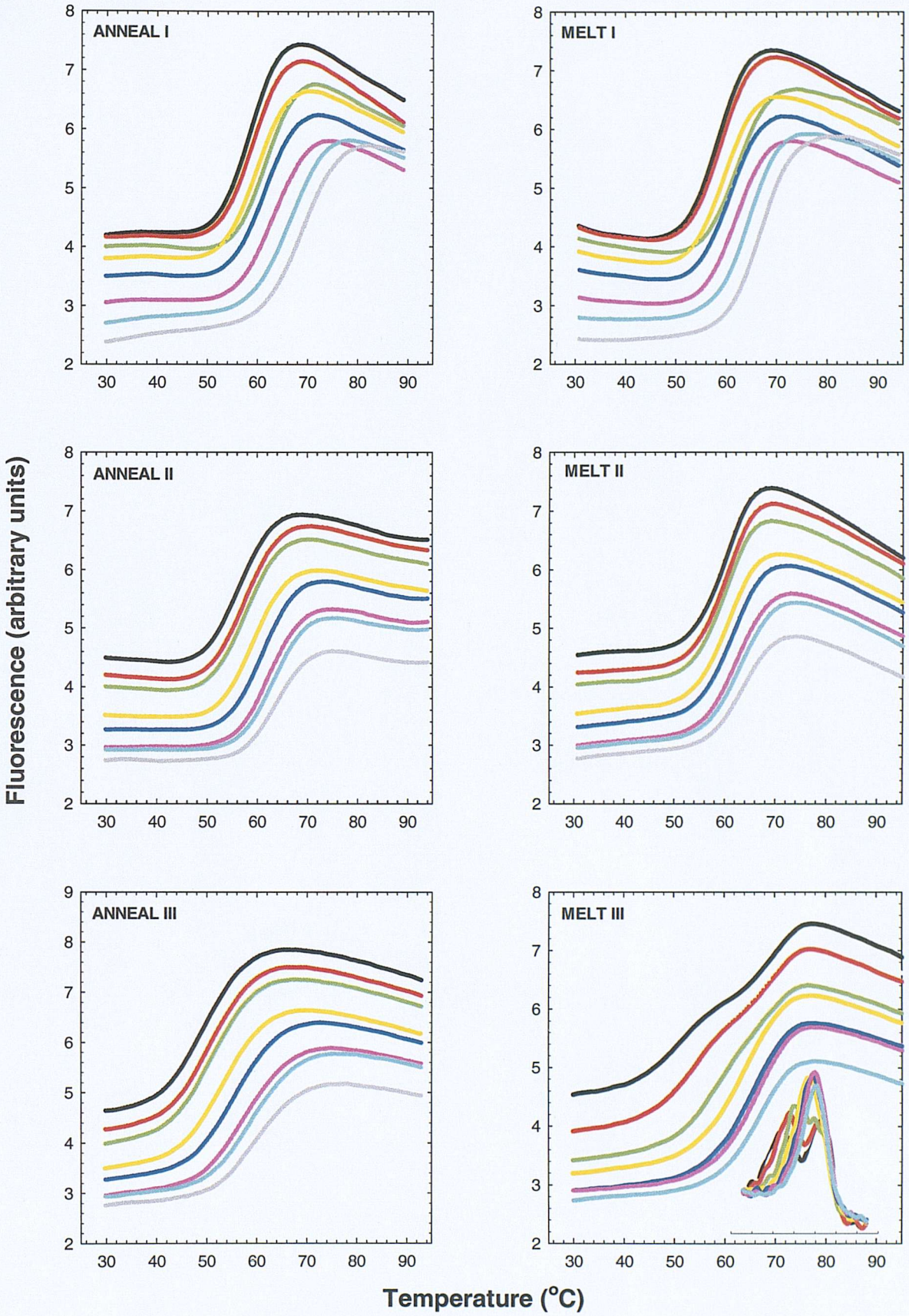


Figure 3.22. Examples of addition of P8 on annealing (left panels) and melting profiles (right panels) obtained with OT3.5 at slow (I, intermediate (II) and fast (III) rates of heating and cooling. Inset within melt III is its first derivative. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

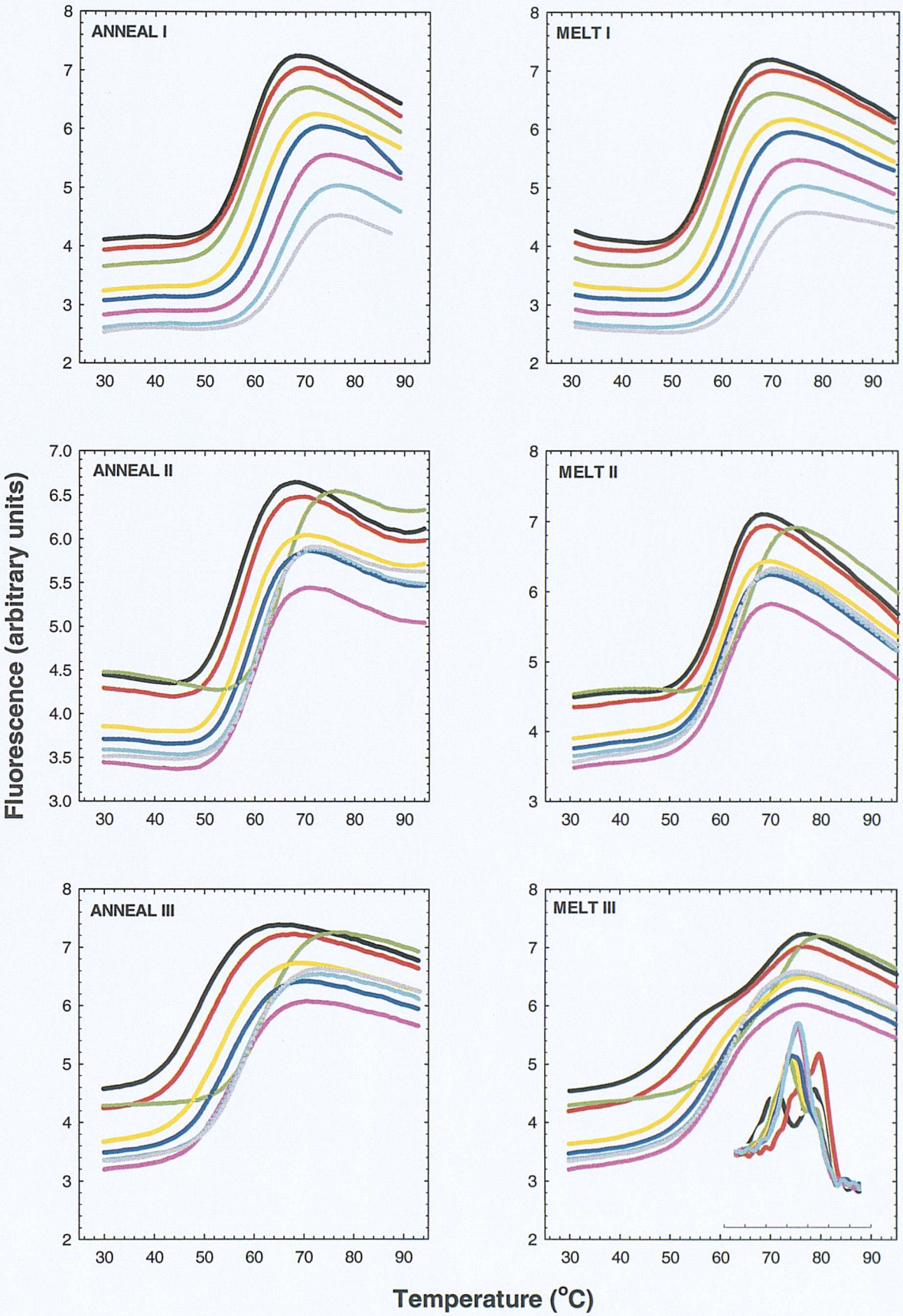


Figure 3.23. Examples of addition of P15 on annealing (left panels) and melting profiles (right panels) obtained OT3.5 at slow (I), intermediate (II) and fast (III) rates of heating and cooling. Inset within melt III is its first derivative. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

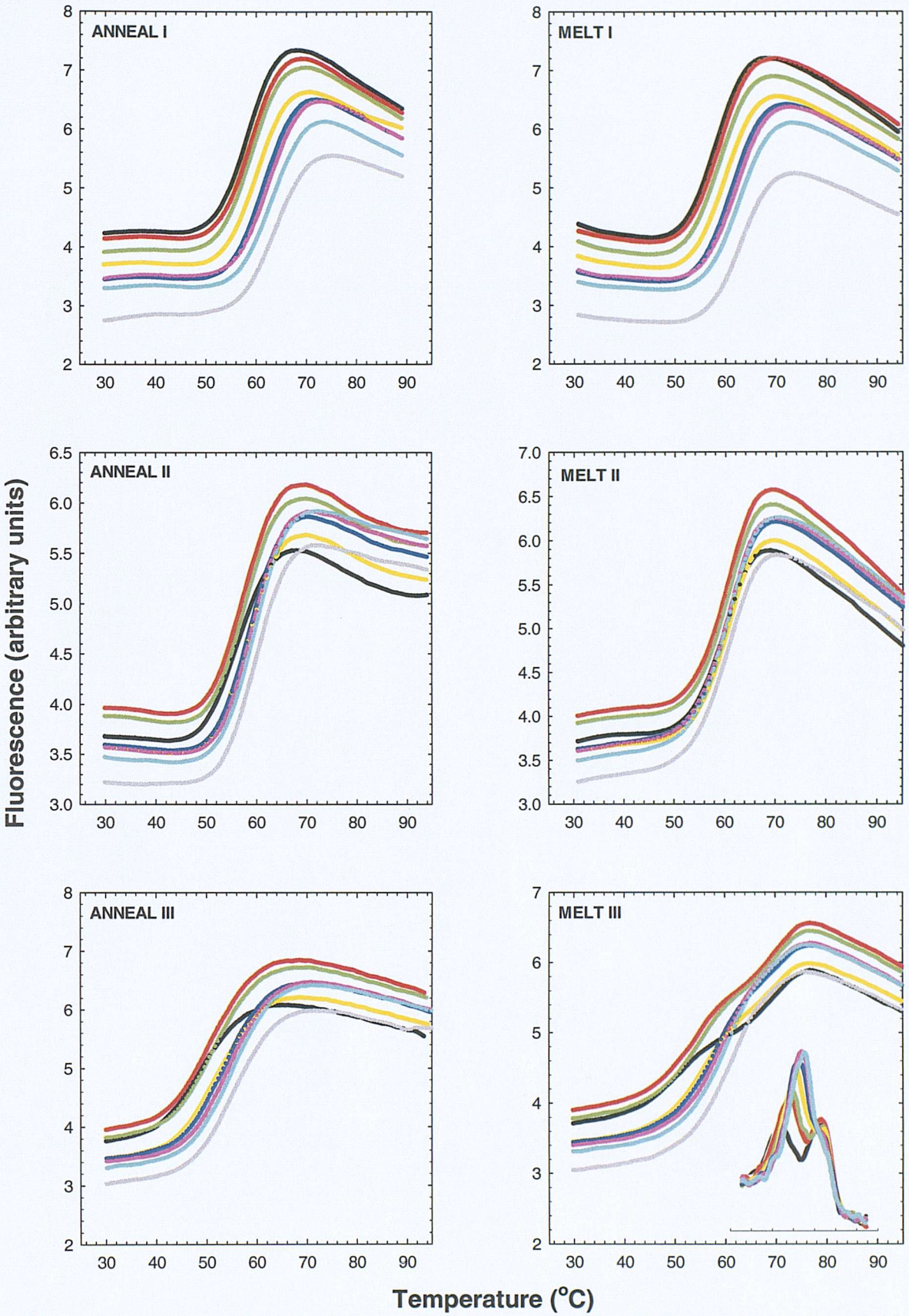


Figure 3.24. Examples of addition of P16 on annealing (left panels) and melting profiles (right panels) obtained OT3.5 at slow (I), intermediate (II) and fast (III) rates of heating and cooling. Inset within melt III is its first derivative. Plots show ligand concentrations as control (black), 0.5 μM (red), 1 μM (green) 2 μM (yellow), 4 μM (blue), 6 μM (pink), 8 μM (light blue) and 10μM (grey).

Table 3.3. Summary of the ligand-induced increases in the transitions of fluorescence (annealing and melting) profiles representing formation of intramolecular quadruplex by OT3.5 oligonucleotide at slow of temperature change (0.05°C min^{-1}).

[AQ] μM	T_m anneal (°C)	T_m melt (°C)	Hysterisis (°C)	ΔT_m anneal (°C)	ΔT_m melt (°C)	Δ Hyst[a] (°C)
Control	58.5 ± 0.1	58.5 ± 0.1	0	n.a.	n.a.	n.a.
P6						
1	58.9 ±0.2	59.1 ± 0.3	0.2	0.4	0.7	0.3
10	63.6 ±2.1	62.3 ± 1.4	-1.3	5.1	3.9	-1.2
P8						
1	60.0 ±1.1	60.2 ±1.7	0.2	1.5	1.7	0.2
10	65.2 ±2.7	63.9 ±2.3	-1.3	6.7	5.4	-1.3
P15						
1	58.7 ±0.1	58.9 ±0.4	0.2	0.2	0.3	0.1
10	65.2 ±1.0	65.1 ±0.8	-0.1	6.7	6.5	-0.2
P16						
1	58.9 ±0.2	58.9 ±0.3	0	0.5	0.4	-0.1
10	63.1 ±0.5	62.9 ±0.4	-0.3	4.7	4.4	-0.3

Control values are the average ± stardard error to 1d.p from all experiments, while ligand values are those from three independent experiments. a – values for the relative change in hysterisis (Δ hysterisis) is defined as the ΔTm melt minus ΔTm anneal.

Table 3.4. Summary of the ligand-induced increases in the transitions of fluorescence (annealing and melting) profiles representing formation of intramolecular quadruplex by OT3.5 oligonucleotide at an intermediate rate of temperature change (0.5°C min^{-1}).

[AQ] μM	T_m anneal (°C)	T_m melt (°C)	Hysterisis (°C)	ΔT_m anneal (°C)	ΔT_m melt (°C)	Δ Hyst[a] (°C)
Control	56.5 ± 0.1	60.3 ± 0.1	3.8	n.a.	n.a.	n.a.
P6						
1	57.7 ±0.3	60.6 ±0.3	2.9	1.4	0.5	-0.9
10	64.1 ±4.4	64.6 ±4.7	0.4	7.8	4.5	-3.4
P8						
1	57.9 ±0.4	60.7 ±0.2	2.8	1.4	0.3	-1.1
10	65.8 ±5.1	65.9 ±5.0	0.1	9.3	5.5	-3.8
P15						
1	60.0 ±3.6	62.0 ±2.3	1.9	3.4	1.6	-1.8
10	65.3 ±4.4	65.3 ±4.2	0	8.7	4.9	-3.8
P16						
1	58.6 ±1.3	61.2 ±0.8	2.6	2.1	0.9	-1.2
10	62.5 ±2.5	62.7 ±1.9	0.3	6.0	2.4	-3.5

Control values are the average ± stardard error to 1d.p from all experiments, while ligand values are those from three independent experiments. **a** – values for the relative change in hysterisis (Δ hysterisis) is defined as the ΔTm melt minus ΔTm anneal.

Table 3.5. Summary of the ligand-induced increases in the transitions of fluorescence (annealing and melting) profiles representing formation of intramolecular quadruplex by OT3.5 oligonucleotide at a fast rate of temperature change (6°C min^{-1}).

[AQ] μM	T_m anneal (°C)	T_m melt (°C)	Hysterisis (°C)	ΔT_m anneal (°C)	ΔT_m melt (°C)	Δ Hyst[a] (°C)
Control	49.4 ± 0.1	51.8 ± 0.2 69.8 ± 0.2	?	n.a.	n.a.	n.a.
P6						
1	51.7 ± 0.9	55.9 ± 1.3 69.0 ± 0.9	?	2.4	4.3 -0.7	?
10	60.3 ± 4.2	65.2 ± 3.5	?	11.0	13.6	?
P8						
1	51.4 ± 0.8	55.6 ± 0.6 69.3 ± 0.4	?	1.9	3.5 -0.3	?
10	61.9 ± 5.3	66.8 ± 3.4	?	12.5	14.7	?
P15						
1	54.3 ± 5.4	53.7 ± 1.3 69.6 ± 2.0	?	4.8	2.1 -0.4	?
10	61.0 ± 3.1	66.0 ± 2.5	?	11.5	14.4	?
P16						
1	52.7 ± 2.0	55.8 ± 1.5 69.3 ± 0.7	?	3.3	4.0 -0.4	?
10	58.5 ± 3.3	64.2 ± 1.5	?	9.1	12.4	?

Control values are the average ± stardard error to 1d.p from all experiments, while ligand values are those from three independent experiments. i – values for the relative change in hysterisis (Δ hysterisis) is defined as the ΔTm melt minus ΔTm anneal.

The results of TR2 bandshift studies on all the ligands tested are summarised in Table 3.6. These data show that in each case the amount of material in the retarded species increases with ligand concentration. The formation of specific quadruplex structures are substantiated by studies with analogues of TR2 in which the G_3 tracts were substituted with inosines (I_3) or GAG. No retarded bands were observed with either of these oligonucleotides, even at high ligand concentrations (see Figure 3.29).

3.3.8 TR4 (intramolecular) quadruplex bandshifts

Several of the drugs were also tested in bandshift studies with oligonucleotide TR4. The TR4 oligonucleotide was composed of 4 human telomeric repeats $(TTAGGG)_4$ flanked at the 3' and 5' ends by T_{10}. These tails were added in order to increase the overall length and to facilitate studies with chemical probing. These experiments were performed in order to confirm that the ligands affect the intramolecular folding of the human telomeric repeat, in addition to their observed effect on dimeric association into an intermolecular quadruplex.

The results of such an experiment with BRACO19 are shown in Figure 3.30. It can be seen that addition of the ligand increases the electrophoretic mobility. This cannot be simply the result of interaction with the ligand as this would increase the molecular weight and decrease the overall negative charge. Both these effects should decrease, rather than increase the mobility. It therefore seems that the ligand has induced the formation of a more compact quadruplex structure. This is again supported by experiments with sequences that are unable to form quadruplexes (replacing guanines with inosines or GAG repeats). The ligand can be seen to have no effect on the mobility of these oligonucleotides (see Figure 3.30). The increased mobility on forming an intramolecular complex is consistent with previous observations with similar sequences (Sen & Gilbert, 1991).

The conditions used in the bandshifts with TR4 have not been optimised and the results show only limited success. The small difference in mobility between the unfolded single strand and the folded complex makes differentiation of the two species difficult. A 10 h running time was also required to separate the oligonucleotide into two bands, and even then they were very close and often indistinct. Bandshifts with adequately separated bands (necessary for quantification) were not easily obtained. Data were

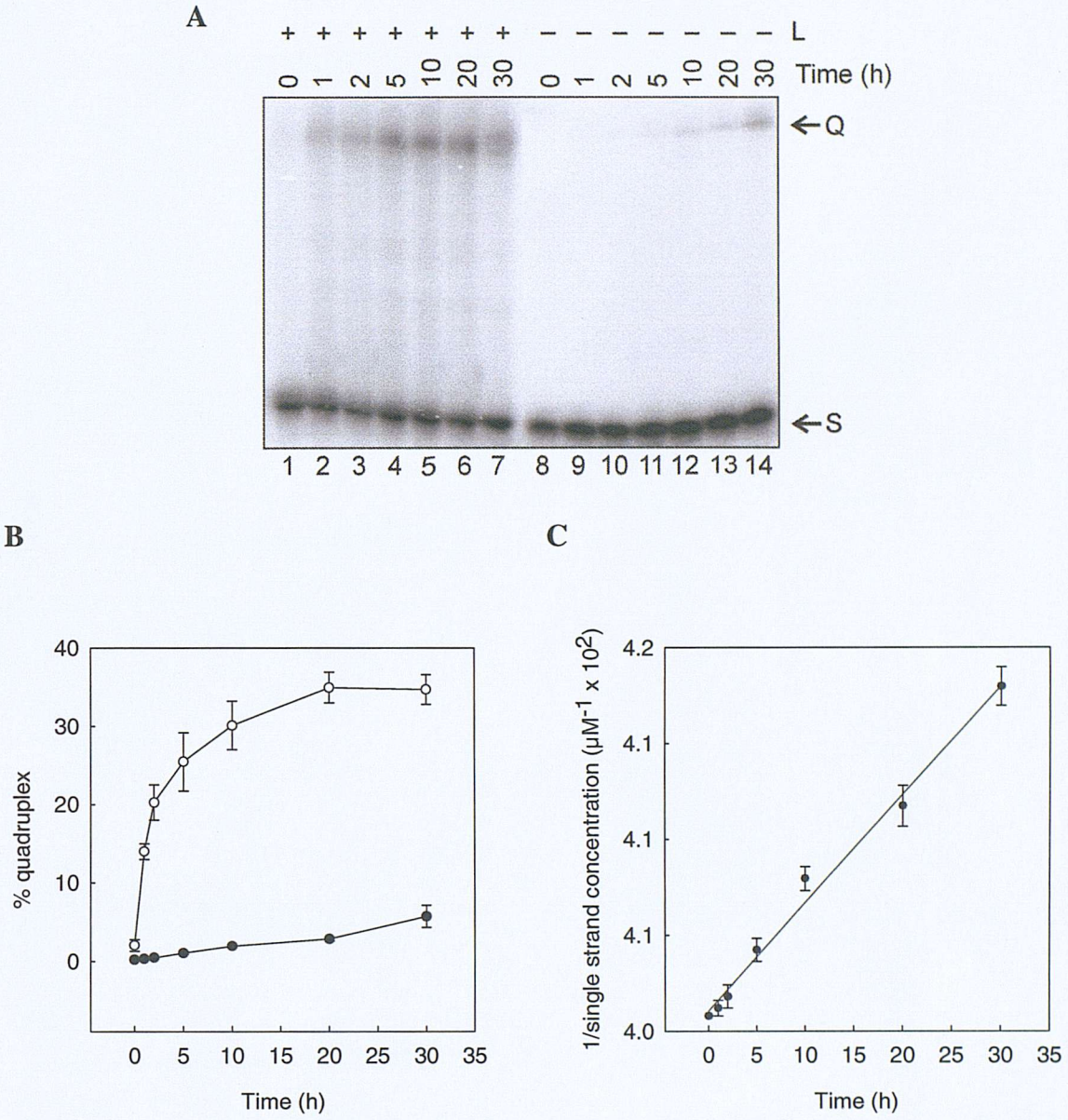


Figure 3.25. (A) Examples of a timecourse showing bandshifts obtained with 25 μM TR2 over 30 h in the absence (-) and presence (+) of the 25 μM BRACO20 (L). Time is shown above each lane. Q and S indicate the dimeric quadruplex and single stranded forms respectively. (B) Quantification of the percentage of quadruplex DNA over time. Filled circles represent quadruplex formation in the absence of ligand, and open circles indicate the presence of ligand. (C) Rate onstant for formation of dimeric quadruplex was derived by ploting 1 / single strand concentration against time.

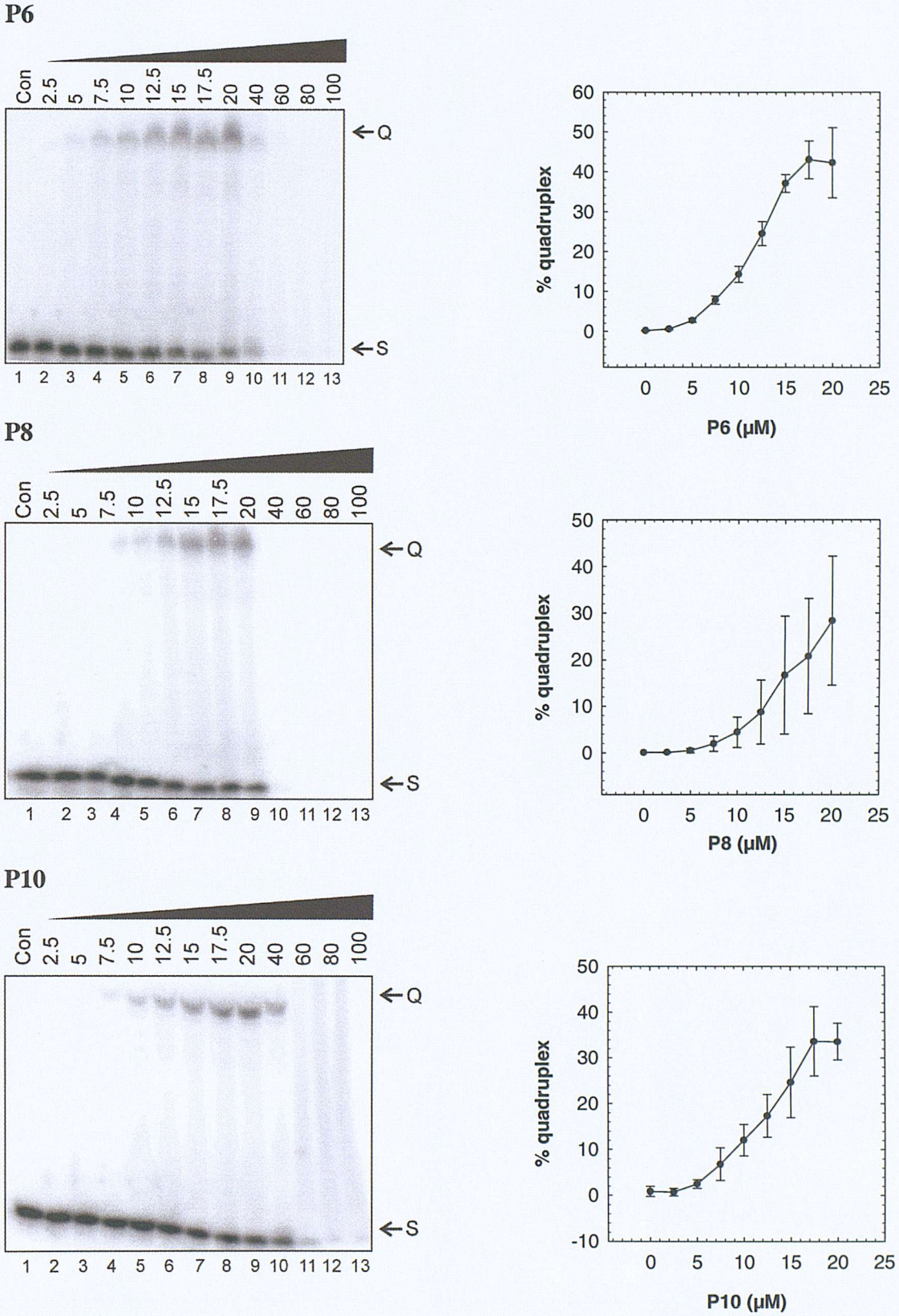


Figure 3.26. Examples of addition of active G4 ligands on bandshifts obtained with TR2. Ligand concentrations (μM) are shown at the top of each lane. Con indicates the banshift in absence of ligand. Q and S represent the dimeric quadruplex and single stranded forms, respectively. Accompanying plots show quantification of the percentage of quadruplex DNA as a function of ligand concentration.

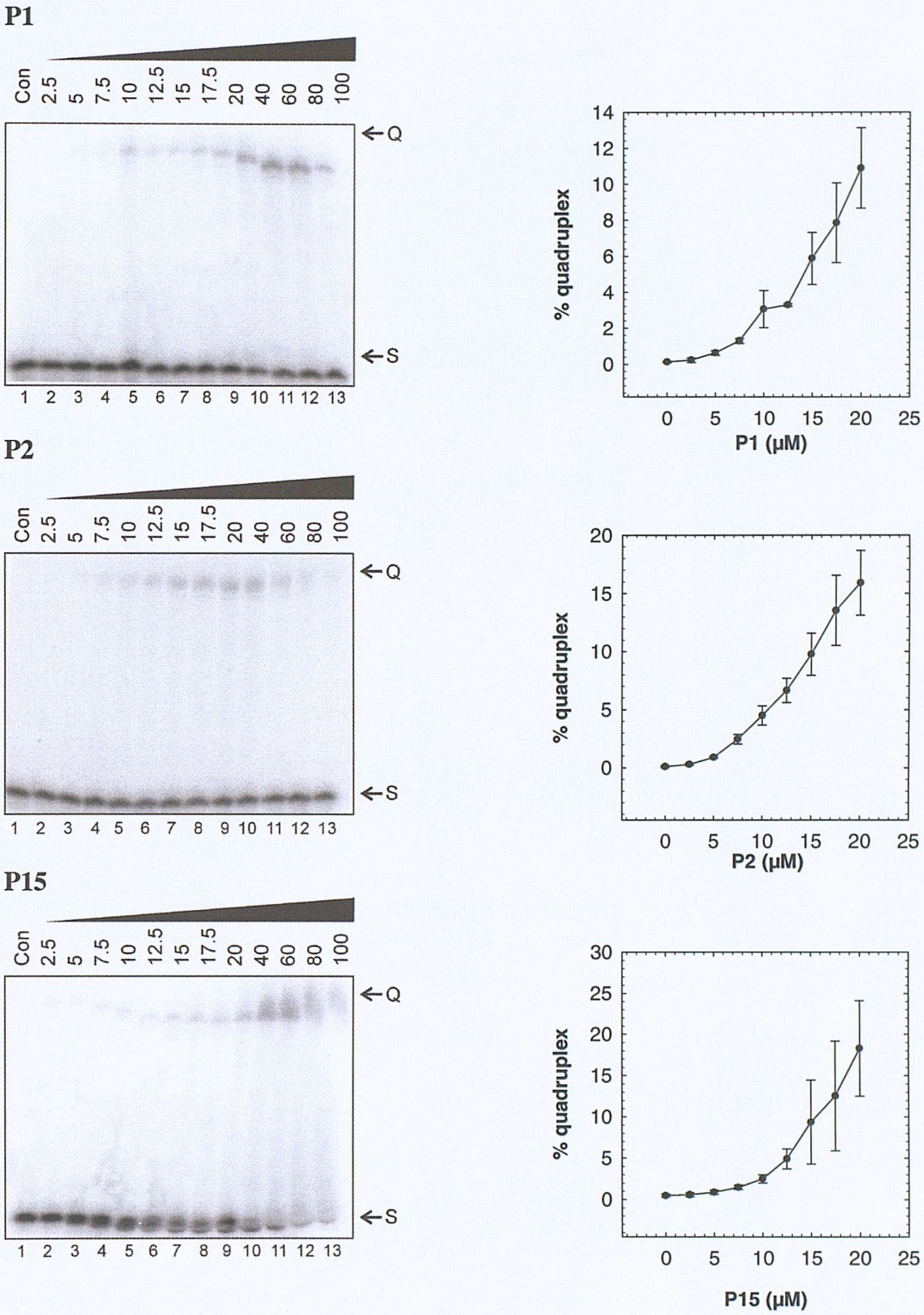


Figure 3.27. Examples bandshifts with TR2 on addition of modest G4 ligands. Ligand concentrations (μM) are shown at the top of each lane. 'Con' indicates the banshift in absence of ligand. 'Q' and 'S' represent the dimeric quadruplex and single stranded forms, respectively. Accompanying plots show quantification of the percentage of quadruplex DNA as a function of ligand concentration.

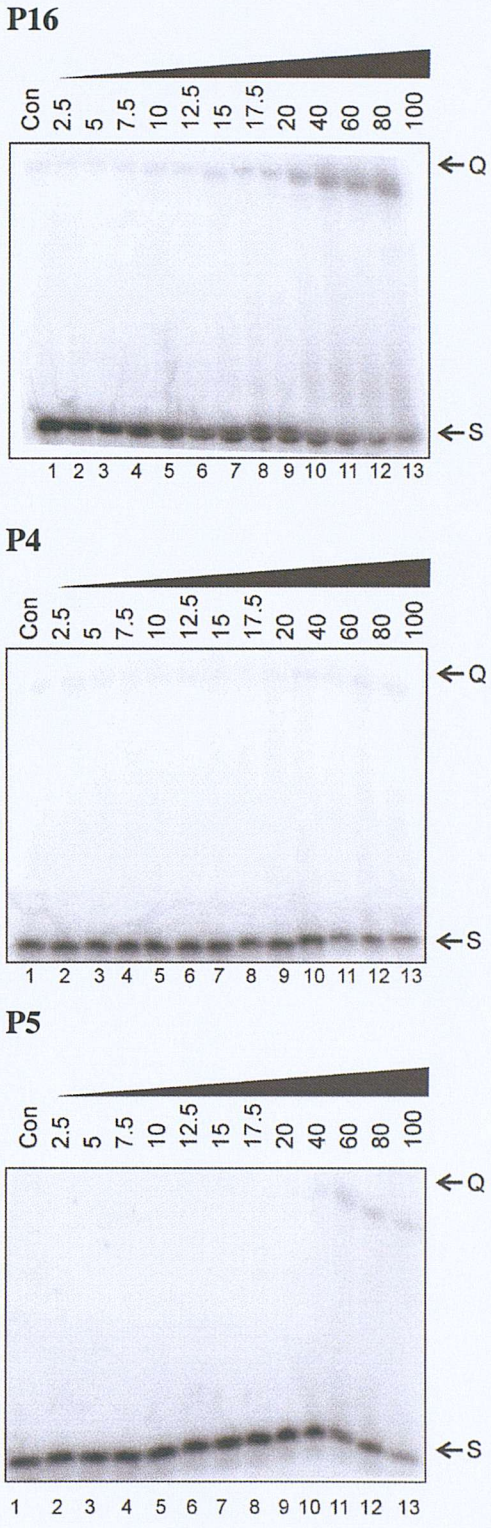


Figure 3.28. Examples of bandshifts with TR2 on addition of poor G4 ligands. No quantification of the percentage of quadruplex DNA was possible.

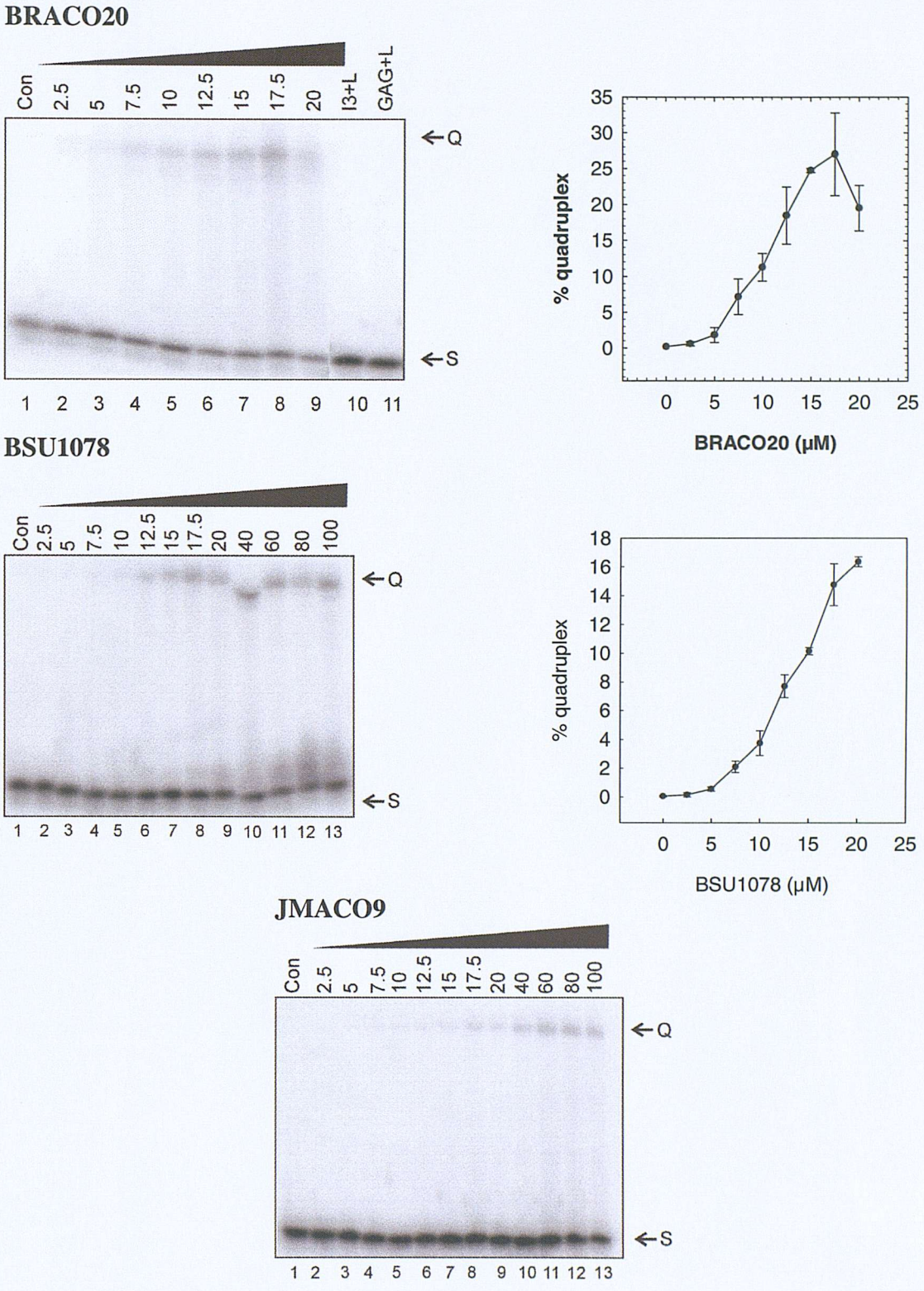


Figure 3.29. Examples of bandshifts obtained with TR2 on addition of known nucleic acid ligands. Ligand concentrations (μM) are indicated at the top of each lane. 'Con' indicates the banshift in absence of ligand. 'I3+L' and 'GAG+L' represent bandshifts with the oligonucleotides TR2I and TR2A on addition of 100 μM ligand. 'Q' and 'S' represent the dimeric quadruplex and single stranded forms, respectively. Where possible, quantification of the percentage of quadruplex DNA as a function of ligand concentration accompany the bandshifts.

Table 3.6. Results of bandshift analysis of quadruplex formation by TR2 on addition of ligands, determined in 100 mM K^+ over 4 h.

Ligand	Activity at 20 μM		2-site cooperative binding model	
	$q + 20\ \mu M\ L$ (%)	k_1' ($M^{-1}\ h^{-1}$)	q_{max} (%)	C_{50} (μM)
P1	10.9 ± 2.2	6100	34.0 ± 11.1	31.7 ± 8.1
P2	15.90 ± 2.8	9500	25.5 ± 9.1	18.0 ± 3.6
P3	trace	-	-	-
P4	trace	-	-	-
P5	0.20 ± 0	100	20.1 ± 9.7	77.6 ± 36.4
P6	42.3 ± 8.8	37000	100	21.0 ± 1.1
P7	17.5 ± 0.8	11000	100	50.0 ± 1.7
P8	28.4 ± 13.8	20000	76.2 ± 20.9	33.3 ± 15.2
P9	2.3 ± 1.3	1200	37.5 ± 5.6	62.2 ± 11.6
P10	33.6 ± 4.0	25000	86.4 ± 19.3	24.0 ± 6.8
P11	1.7 ± 1.7	860	57.5 ± 42.5	97.1 ± 104.8
P12	trace	-	-	-
P13	3.6 ± 0.9	1900	26.9 ± 7.3	58.6 ± 13.4
P14	1.1 ± 0.4	560	6.9 ± 4.1	30.1 ± 26.1
P15	18.3 ± 10.8	11000	48.9 ± 5.3	31.0 ± 8.2
P16	7.4 ± 0.9	4000	49.3 ± 11.6	50.0 ± 11.3
D1	3.1 ± 2.0	1600	34.1 ± 16.6	51.7 ± 5.7
D2	1.0 ± 0.8	500	30.7 ± 29.5	106.0 ± 68.0
SAC14	14.3 ± 5.0	8300	56.3 ± 19.8	38.2 ± 9.5
SAC27	7.0 ± 0.2	3800	23.6 ± 6.9	35.2 ± 4.4
SAC28	5.4 ± 1.3	2900	22.6 ± 0.3	38.1 ± 7.4
SAC47	1.4 ± 0.6	710	12.9 ± 9.9	60.5 ± 44.8
SAC49	10.4 ± 1.2	5800	62.2 ± 3.2	46.6 ± 7.3
BSU9060	trace	-	-	-
BSU1074	trace	-	-	-
BSU9010	trace	-	-	-
BSU1078	16.4 ± 0.4	9800	88.5 ± 11.6	41.1 ± 3.2
BSU9048	trace	-	-	-
BRACO4	3.0 ± 0.3	1500	8.8 ± 1.1	29.1 ± 5.6
JMACO9	2.3 ± 0.4	1200	8.3 ± 2.4	36.0 ± 1.7
BSU6039	5.1 ± 1.7	2700	13.1 ± 3.3	26.4 ± 6.2
BRACO19	25.4 ± 0.2	17000	100	32.1 ± 0.8
BRACO20	19.6 ± 3.2	12000	100	13.4 ± 0.6

therefore only obtained for a limited number of the ligands. Examples of the bandshift produced by the effective G4 ligands BRACO20, P8 and P10 are shown in Figures 3.30, and a summary of those ligands initially tested on the intramolecular bandshifts is given in Table 3.7.

It can be seen that the extent of quadruplex formation is generally much higher for intra- than intermolecular complexes. This is also evident in the TR4 controls in which the quadruplex already accounts for around ~ 8% of the material (Figure 3.30), in contrast to a minimal (~ 0.5%) quadruplex formation for the dimeric complexes with TR2, seen in the absence of added ligand (see plots in Figures 3.26 to 3.29). The 4 h incubation therefore allows folding to be nearer equilibrium for TR4 than for dimerisation of TR2. This faster rate is to be expected for an intramolecular process as the guanine-tracts are in close proximity on the same strand and do not require a bimolecular collision.

The bandshift experiments with TR4 were initially carried-out at the same strand concentration as used for TR2 (5 μM). However, this is not essential, if quadruplex formation is an intramolecular process and the extent of the reaction showed, will only depend on the ligand concentration, as this far exceeds its K_d for DNA.

3.4 Discussion

3.4.1 Fluorescence melting profiles in the absence of ligand

Full details of the profiles obtained with both the human and *Oxytricha* telomeric repeats can be found in chapter 4. The sigmoidal shape of the melting curves is a hallmark of the cooperative process, which is typical of melting of a structured nucleic acid. Thermodynamic parameters can be derived from van't Hoff analysis of the melting profiles providing there is no hysteresis, and a two-state model is assumed where melting is all-or-none, and the strands only exist as single stranded or fully-intact structures with no partially melted intermediates.

Hysteresis is a major consideration when analysing the fluorescence profiles. The profiles only represent the reaction at equilibrium if the rate of heating and cooling is sufficiently slow to allow the reaction to fully relax to a new equilibrium at all temperatures throughout the profiles. If not, hysteresis occurs where the annealing and

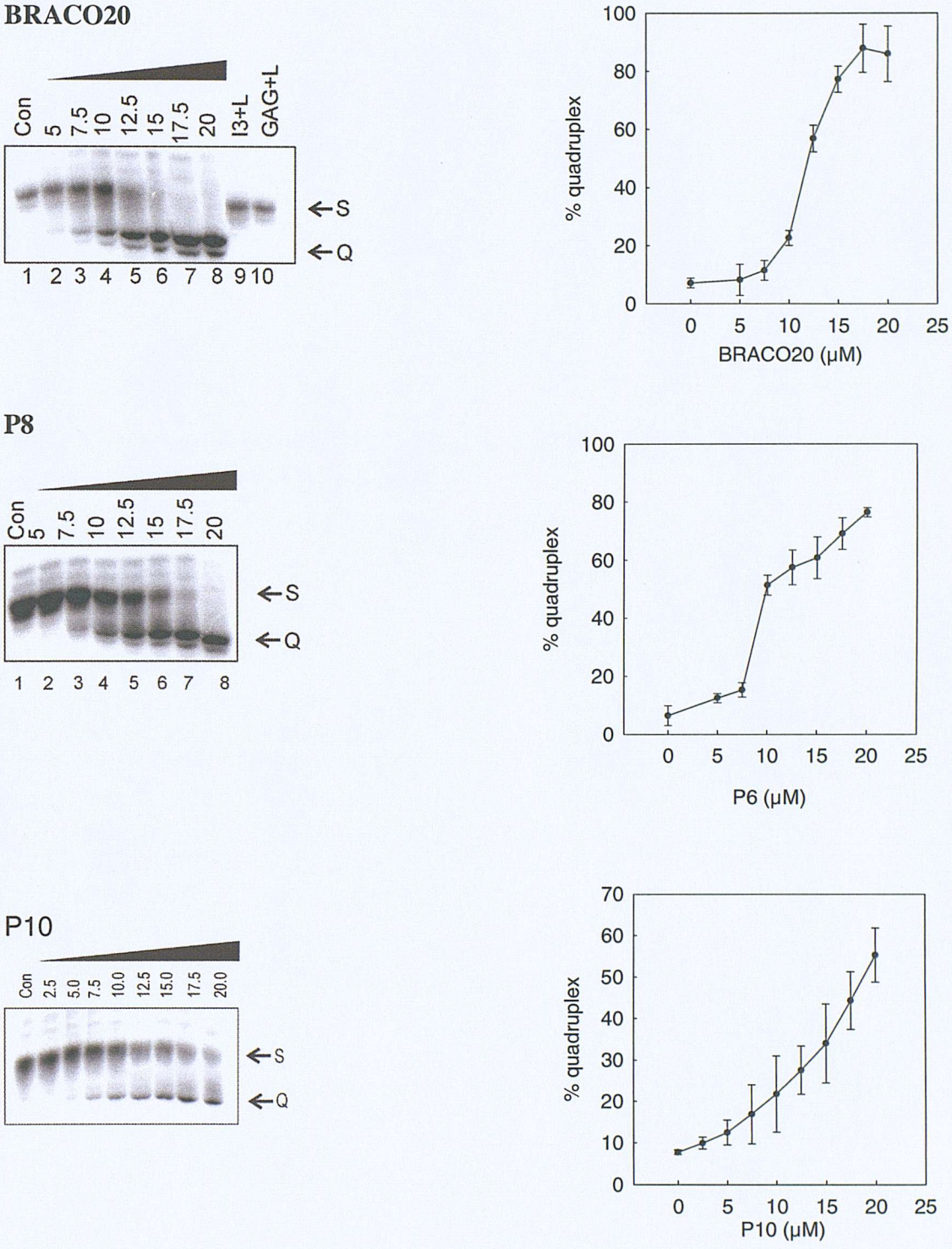


Figure 3.30. Examples of bandshifts with TR4 on addition of G4 ligands. Ligand concentrations (μM) are indicated at the top of each lane. 'Con' indicates the bandshift in absence of ligand. 'I3+L' and 'GAG+L' represent bandshifts with the oligonucleotides TR4I and TR4A on addition of 100 μM ligand. 'Q' and 'S' represent the monomeric quadruplex and single stranded forms, respectively. Accompanying plots show quantification of the percentage of quadruplex DNA as a function of ligand concentration.

Table 3.7. Results of bandshift analysis of quadruplex formation by TR4 on addition of ligands, determined in 100 mM K^+ over 4 h

ligand	activity at 20 μM	
	q + 20 μM L (%)	k_1' (h^{-1})
P1	46.6 ± 9.1	0.16
P2	46.8 ± 1.0	0.16
P3	trace	
P4	trace	
P5	59.6 ± 6.0	0.23
P6	74.9 ± 1.6	0.35
P7	48.8 ± 4.0	0.17
P8	48.3 ± 6.3	0.16
P9	12.1 ± 2.0	0.03
P10	55.3 ± 6.5	0.20
BRACO19	76.1 ± 11.0	0.36
BRACO20	85.5 ± 6.7	0.48

melting profiles do not coincide, giving a false measure of the fraction of oligonucleotide as single strand and quadruplex. The amount of hysteresis indicates how much slower the melting or annealing reactions are than the rate of heating and cooling. A slow rate of association will depress the apparent annealing T_m, while a slow rate of dissociation will elevate the melting T_m. Experimental conditions such as the addition of ligands (see below) and cations (see chapter 4) influence association and/or dissociation as can be seen from the hysteresis.

The hysteresis is greatest for OT3.5, so much so that it was necessary to revise the standard protocol for fluorescence-quenching studies (one subject of chapter 4). In the absence of ligands, HT1.5 is significantly more affected by hysteresis than HT3.5 (for which this is negligible). The profiles obtained with HT1.5 are also broad and shallow, whereas those with the HT3.5 and OT3.5 are much sharper and defined. This may result from the polymorphic nature of dimeric quadruplexes formed by HT1.5, for which there are more possibilities than for an intramolecular complex (detailed in chapter 1). Intramolecular quadruplexes, like dimeric quadruplexes, can form edge-loop and diagonal-loop topologies with different groove dimensions and loop orientations. However, more positions for the loops are possible for the dimers. Since interactions within the loops may be an important factor for ligand-binding, a mixed population of quadruplex structural variants may produce subtly different melting and annealing profiles. The technique nonetheless, proved to be a useful way of obtaining data on the effects of ligands and ionic conditions on quadruplex stability.

3.4.2 Fluorescence melting on addition of ligands

The ligands were characterised based on the ΔT_m they produce (relative to the native T_m). This reflects both the higher binding affinity of the drugs for quadruplex over random coiled DNA and the strength of the resulting complex. However, addition of ligands also produced several other effects on the profiles obtained with each sequence.

Firstly, the effects of the ligands on the annealing and melting profiles of all three sequences significantly affected the hysteresis. Increases in hysteresis can be seen for HT3.5 where the ligands induce greater ΔT_m in the melting profiles than the annealing profiles. This indicates the melting is further removed than annealing from the (new) true equilibrium profile. Such an effect suggests that the ligands interact with

quadruplexes that are already folded, causing a decrease of the rate of dissociation. Alternatively, decreases in hysteresis can be seen for HT1.5 and OT3.5 (the two sequences with much slower kinetics) where the opposite trend in enhancement of the annealing and melting is observed. Here, the results suggest that the ligand interactions with the random coil are more pronounced, and as the rate of association is much slower in the HT1.5 experiments, a substantial increase in association is produced. The different effects of ligands on the association or dissociation rates will both result in an increased K_{eq}.

Secondly, the G4 ligands noticeably affect the shape of the profiles produced by all the oligonucleotides. This characteristic is already pronounced for the annealing and melting profiles of HT1.5. In general, the ligands increase the width of the melting transitions of the profiles. This is most evident for the most active compounds.

Thirdly, the ligands also produce asymmetric transitions *i.e.* the curve slopes more prominently above than below the transition. This possibly reflects stronger drug-binding and stabilisation of specific regions of the quadruplex structure. Selective interactions are more likely to represent specific end-stacking of the G4-ligands between the terminal G-quartets and the loops. This is a mode common to most ligands cited in the literature (discussed in chapter 1) rather than intercalation at multiple sites within the G-stack. Importantly, such drug-binding may complicate a two-state model.

Finally, the fluorescence signals throughout the profiles (especially at higher temperatures) are reduced by the compounds. This could be caused by ligand-fluorophore quenching, or by absorbance by the ligands competing with that by the fluorophore at the excitatory wavelength. As the effect is more evident at the high end of the temperatures range, it is more likely to represent ligand-fluorophore quenching. This is more likely to be enhanced at high temperature where the oligonucleotide is more unstructured, exposing a higher number of fluorophore molecules to the ligands.

The ΔT_m values are generally greater with HT3.5 than HT1.5, while the most active ligands added to OT3.5 seem to be weakly activity in comparison with the HT sequences. A recently published structure has shown that a G4-ligand can form specific interactions with a dimeric $d[(G_4T_4G_4)]_2$ quadruplex (Haider *et al.*, 2003). However, it

may be inappropriate to compare of the effects of the ligands on the two. The activities of the ligands are those produced upon a Na^+-induced quadruplex for *Oxytricha* telomeric sequence rather than a K^+-induced structure for the human telomeric sequence. These two quadruplexes may adopt different topologies, and the decreased ligand activities may therefore represent the interactions of the ligands with different quadruplex geometries. Although, the structure of these two telomeric repeats have been examined under a range of different cationic conditions, which indicate that the same topologies are formed with either cation. It is unclear if the differences seen in this work can be attributed only to use of Na^+ rather than K^+.

3.4.3 SAR from fluorescence melting experiments

It is often difficult to distinguish between the effects of the sidechains, their position on the ring, or the linkage upon activity. Nevertheless, studying an array of differing substituents attached via the amide or reverse-amide linkage at the 2,6 and 2,7 positions of anthraquinone moiety has helped to illustrate several general trend. The structure of each anthraquinone conjugate discussed below are shown in Figures 3.2 to 3.7, where they are noted with their code.

Position of substituent: 2,6- versus 2,7- anthraquinones

The differences in activity produced by 2,6 or 2,7 attachment of various sidechains to the anthraquinone chromophore were first examined. The ΔT_m values (shown in Tables 3.1 and 3.2) for annealing and melting of the HT sequences produced by a number of 2,6- to 2,7- anthraquinones were compared. This was achieved by contrasting P1 and P2, P3 and P4, P6 and P14, P7 and P11, P8 and P13, P9 and P12, SAC27 and SAC28 (for HT3.5 only).

The majority of the data show that the 2,6 position produces better stabilisation of quadruplex than the 2,7 position. P1 and P2 have similar effects on HT1.5, but P1 is clearly more active against HT3.5. P3 appears to be more active than P4 on HT1.5, however their activities on HT3.5 are similar. All the other compounds show greater activity on HT1.5 and HT3.5 with sidechains at the 2,6 position, namely P6, P7, P8, and P9. SAC27 and SAC28 give similar ΔT_m with HT3.5.

The 2,6 substitution produces an anthraquinone where the sidechains diagonally traverse the short axis of the aromatic system. This possibly extends the reach of these sidechains, whereas substituents at the 2,7 positions reside along the long axis. These positions on opposite short sides of the polycyclic system may affect where and how the chains interact depending on whether the ring stacks within or on top of the stacked G-tetrads.

Amide versus "reverse"-amide

The effect of the type of linkage of the propionamide chain to the anthraquinone ring was also examined. Again ΔT_m as shown in Tables 3.1 and 3.2 were used to examine the activity of anthraquinone isomers with identical substituents, varying only by the linkage to the anthraquinone moiety. The two classes were those linked via the CO or the NH of the amide. The ΔT_m values on HT1.5 and HT3.5 were compared for P1 and P10, P15 and P16. Similarly, P7 and SAC28, and P11 and SAC27 were studied on HT3.5.

The results suggest that the CO linkage is better than the NH linkage. This is seen for P1 with HT1.5 but is not seen with HT3.5. It is also true for P15 with both HT1.5 and HT3.5. The compounds P7 and P11 also produce much higher activity than their respective isomers. A possible explanation for the superiority of the CO linkage is that it provides better sidechain interactions with the guanines in the terminal G-tetrads near to where the anthraquinone polycycle is stacked, rather than interactions in the grooves or within the loop.

Nature of sidechains

Longer sidechains seem to produce better activity, but this may be negated by other factors. P15 produces the best ΔT_m values for annealing and melting profiles obtained with both HT1.5 and HT3.5, while P6, P8 and P16 are equipotent. P15 and P16 both possess reduced alanyl-propionamide sidechains, eight atoms in length. Their sidechains lack reduced amide COs at atom-6, which are found in the less active compounds P7, P11, SAC27, and SAC28. As such the sidechains of P15 and P16 are capable of accepting one hydrogen bond via a remaining amide CO and donating three

hydrogens linked to either amide or amine N. The active ligand P6 possesses twelve-atom backboned lysyl-propionamides, which provide three hydrogen bond donor amides or amines and only one CO amide hydrogen bond acceptor. The activity of P8 is conferred by two argyl-propionamide sidechains with thirteen-atom backbones comprised of six possible hydrogen bond donor NH amides or amines and two amide CO hydrogen bond acceptors. SAC14 also produces an excellent ΔT_m among the anthraquinones in studies of HT3.5, it stabilises the annealing and melting to temperatures above 90°C at a concentration of 10 μM. SAC14 possesses the longest sidechains within the series, it has two fifteen-atom backbones with five potential hydrogen bond donor NH amide or amine groups. D1 has eleven-atom valyl-propionamide sidechains, providing two possible hydrogen bond acceptor oxygens. The nine-atom valyl-alanyl-propionamide sidechains of D2 contain three possible hydrogen bond acceptor amide CO. Though both D1 and D2 contain these long sidechains, they are weakly active on HT3.5 oligonucleotide.

The number of amines (secondary and tertiary) which can be protonated is also a possible major determinant of the activity of the ligands studied. These groups are hydrogen bond donors, but more significantly if protonated they could be involved in stronger electrostatic interactions which may increase activity. Which amines in the sidechains are protonated under the experimental conditions (pH 7.4) is unclear and is a matter of speculation. However, while the majority of the anthraquinones have only two amines (one per sidechain), several of the most active compounds including P5, P6, P15, and P16 have four amines, while SAC14 has six amines.

The presence of bulky groups which may decrease activity is another important feature of the sidechain. P7 has two eight-atom backbone alanyl-propionamide substituents, yet it is more active than D1 and D2 that have longer sidechains and are identical up to atom-6. The methyl group at atom-7 of the P7 sidechains replaces the isopropyl group on atom-7 of the D1 and D2 sidechains. The amino-propionamide sidechains of P10 even though short, do not result in good activity within the anthraquinone series. Furthermore, bulky electronegative groups such as the phenylalanyl-propionamides on P3 and P4 significantly impair their activity. Although P5 has similar substituents but for a reduced amide CO at atom-6 to P3 and P4, it is more active.

The importance of the sidechains on the activity of the anthraquinones relates to the possible ways in which substituted polycyclic G4 ligands interact with quadruplexes. The aromatic ring (anthraquinone moiety) can form π-π stacking interactions with the G-tetrads, and the central ring oxygen atoms can hydrogen bond to thymines in loop (Haider *et al.*, 2003). The sidechains also play a crucial role in binding. If protonated the secondary or tertiary amines may electrostatic interact with the phosphate-sugar backbone. The nitrogen and oxygen atoms of the amide or amine in the sidechains may also interact with the grooves or the loops by hydrogen bond interactions with nucleobases or phosphate oxygen atoms. These interactions may also be mediated by networks of water molecules. Longer backbones with more groups capable of interactions may aid binding by allowing the substituents to stretch further down the grooves of a quadruplex and make more specific interactions. Longer chains will also be more flexible and able to adopt positions that allow optimal interactions. It is also conceivable that such chains would impair intercalation within the stacked G-quartets, and favour end pasting between terminal G-quartet and loops. Cationic groups in the sidechains may allow electrostatic or hydrogen bonding interactions, while bulky groups may hinder binding with the quadruplex grooves by steric clashes. Electronegative atoms, such as amide CO in the sidechains may reduce repulsion allowing interactions with the groove. However, the nitrogen of an amide cannot easily be protonated which might increase available hydrogen bonding or electrostatic interactions.

Other quadruplex-binding ligands

Examining other well-characterised drugs provides a way of comparing their effects to their established properties. These compounds included difunctionalised acridines (ACIs), acridones (ACOs) and other anthraquinones (AQs). Five di-substituted anthraquinones were studied, (2,7-AQ) BSU9060, (1,4-AQ) BSU1074, (1,5AQ) BSU9010, (2,6-AQ) BSU1078, and (1,8-AQ) BSU9048. The anthraquinones are a class of compounds that have previously been studied as triplex binding ligands (Fox *et al.*, 1995). Three acridines were also studied, (3,6,9-ACI) BRACO19, (3,6,9-ACI) BRACO20 and (3,6-ACI) BSU6039. The acridine moiety contains a heterocyclic nitrogen atom whose protonation has been implicated in forming specific interactions and should enhance electron deficiency and so π-stacking on G-tetrads (Haider *et al.*, 2003). The tri-substituted acridines BRACO20 and BRACO19 are among those ligands

at the forefront of quadruplex drug-binding. Predictions of their strong interaction with quadruplexes by molecular modelling are consistent with telomerase inhibition (Read *et al.*, 1999; Read *et al.*, 2001) and inhibition of human *SgsI* family of helicases that unwind quadruplexes (Li *et al.*, 2001). The two acridones (3,6-ACO) BRACO4 and (2,7-ACO) JMACO9 were also investigated. The acridone moiety also has a nitrogen atom capable of enhancing π-stacking and forming a hydrogen bond.

As expected the activities of BRACO19 and BRACO20, determined with HT1.5 and HT3.5, are the greatest produced by **any** of the compounds tested (including the novel anthraquinones). This shows that this technique is consistent with other studies that determine G4-activity. Their activity has been attributed to the interaction of the three sidechains with three (rather than two) grooves of a quadruplex, which provides extra quadruplex stability. This again indicates that sidechain interactions are important for activity. The activity of four of the five established anthraquinones are similar to that of the novel ligands with modest activity determined with HT1.5 and HT3.5. They are not however, as active as the good novel drugs P5, P6, P8, P15, and P16. The only inactive anthraquinone, BSU9048 had no effective on melting of HT1.5 and HT3.5. The acridones (3,6-ACO) BRACO4 and (3,6-ACO) JMACO9 also show little activity against both HT1.5 and HT3.5. The acridine (3,6-ACI) BSU6039 is ineffective on HT1.5, but is moderately active on HT3.5, like the other anthraquinones.

P6, P8, P15 & P16 ligand activity on OT3.5

Melting experiments were performed to determine the ability of several of the most active anthraquinones (as determined with the HT sequences) to stabilise the *Oxytricha* telomeric repeats. The annealing and melting curves obtained at different ligand concentrations showed no significant hysteresis, only when the temperature change was 0.05°C min^{-1}. Decreased hysteresis is observed with increased ligand concentration, as seen with HT1.5, but not HT3.5. Larger ΔT_m values are produced at a rate of heating and cooling of 0.5°C min^{-1} than at 0.05°C min^{-1}. This may result from the rate of heating and cooling approaching a rate similar to that of ligand-induced association but not ligand-induced dissociation. Therefore greater effects on association than dissociation would result. The reduced hysteresis observed with all four anthraquinones

implies they are primarily increasing the rate of quadruplex association (as discussed earlier).

Summary

The data on all the ligands indicate that the activity on the three oligonucleotides is different. The ligands have the greatest effects on HT3.5 with lower activity with HT1.5, and little activity with OT3.5. However, the relative order of the compounds is similar for each sequence. The major difference between the oligonucleotides is the relative effects of ligands on the annealing and melting profiles. The ΔT_m for melting are greater than ΔT_m for annealing with HT1.5 and OT3.5 in all cases, while the opposite is seen with HT3.5. These effects on the hysteresis mirror the activity *i.e.* the ligands which produce the greatest ΔT_m also produce the greatest increases in hysteresis for HT3.5, and the greatest decreases in hysteresis for HT1.5 and OT3.5. These are likely to indicate effects on association or dissociation, respectively (as detailed earlier).

The larger overall ΔT_m values (for both annealing and melting) with HT3.5 than HT1.5 may reflect improved ligand interactions with intramolecular rather than dimeric quadruplexes. Monomeric and dimeric quadruplexes are thought to fold into similar diagonal loop and edge loop complexes. However, the orientation of the opposite ends can only vary in dimeric complexes. Thus it is possible that the difference in the relative activities of the ligands reflect interactions with loop structure in the fold-back structure.

The ΔT_m values produced by the ligands are sometimes small, particularly those obtained with HT1.5, but are larger with HT3.5. This makes it difficult to classify the drugs according to their activity. The novel anthraquinones can however, be bracketed into groups that reflect their relative activities. Ligand activities determined by HT1.5 melting (which are greater in the annealing profiles) indicate that P15 is by far the best G4 ligand. P5, P6, P8, and P16 also appear to produce relatively high ΔT_m values. P1, P2, P7, and P10 give moderate ΔT_m values. P9, P10, P11, and P13 give small ΔT_m values, whereas P3, P4, P12, and P14 all produce very low ΔT_m values.

The ΔT_m values observed with HT3.5 are similar to those seen with HT1.5. The drugs SAC14, P8, and P15 give the greatest ΔT_m values. Less, but still highly active are the drugs P1, P2, P5, P6, P7, P10, P11, P13, and P16. Low ΔT_m values are produced by P9, P14, and D1. ΔT_m values of less than a degree are obtained for P3, P4, P12, D2, and SAC27, SAC28, SAC49.

The most notable differences in ΔT_m values for HT1.5 and HT3.5 are seen for P5, P6, P10, and P14. P5 and P6 are more active with HT1.5, P10 is more effective with HT3.5, while P14 is ineffective against HT1.5 but more effective against HT3.5. It is unclear whether these differences relate to any preference for different quadruplex structures. Though both forms can adopt similar topologies, only the dimeric forms can vary the location of their loops between the opposite or same end of the stacked G-quartets.

3.4.4 Limitations of fluorescence melting

Melting using this technique can be difficult to interprete when the profiles are shallow and broad. This is most significant with HT1.5. Trying to fit ideal curves to deduce the T_m and ΔH values to such experimental fluorescence data can be difficult. Moreover, such a van't Hoff analysis is only possible if the reaction is a simple two-state process which is at dynamic equilibrium. Quantification is also complicated by the changes in fluorescence at high and low temperatures. In such cases a simple linear component can be added into the equation (2.16). These temperature-dependent effects are presumably a property of the fluorophore. The point of inflection in the melting profiles, however, is generally taken as an indication of the midpoint for the transition. In contrast, Rougée *et al.* (1992) used the shape of the hysteresis curves (non-equilibrium melting and annealing) to determine the individual rates of association and dissociation, and the associated thermodynamic parameters for an intermolecular triplex .

It is also worth considering whether attachment of the fluorophore has any effect on quadruplex stability. Their presence will presumably alter the entropy of the system, adding disorder to the quadruplex. In addition, they may stack on the ends of the folded structures, thereby stabilising the complex or hindering the association of small molecules. Alternatively, these groups could sterically hinder the folding of the

complexes. The T_m values determined may therefore not accurately reflect the stabilities of the native structures. Such effects have previously been noted in FRET studies of quadruplexes (Mergny & Maurizot, 2001). However, all reactions are subject to these same restrictions, and the comparative studies of the relative effects of ligands or conditions is therefore still valid.

3.4.5 Bandshift analysis & interpretation

The analysis of bandshift experiments assumes that the equilibrium is preserved during the electrophoresis and that the intensities of the bands reflect the true population distribution. Although this should be a simple process, some consideration of the analysis is important so that the limitations of these data can be understood.

Firstly, the quantification of the raw data is not straight-forward. Some studies with dimeric quadruplexes such as TR2 have reported the presence of multiple bands, corresponding to different quadruplex species (Han *et al.*, 1999: Han *et al.*, 2001). However, only **two** bands are seen in all the experiments in this work, indicating that the oligonucleotides exist as only **two** structures, single strand and quadruplex. Increasing drug concentrations lead to the progressive appearance of the band of slower mobility (for TR2) and faster mobility (for TR4). These can confidently be assigned as bimolecular and intramolecular quadruplex structures, as their appearance is ligand dependent, and is not observed with oligonucleotide analogues which are unable to form G-quartet structures.

In several instances the two bands become smeared and indistinct at high drug concentrations (above 20 μM), and some of the ^{32}P-labelled DNA remained in or near the wells (see Figure 3.26). This is especially noticeable with active compounds such as P1, P2 or P15. Similar results can be seen for several other compounds. This may result from the formation of higher-order structures, generally structures with multiple aggregated forms such as G-wires (Marsh & Henderson, 1994). Non-specific drug interactions with the DNA may also be possible, though this is not observed with oligonucleotides that are unable to form quadruplexes. This smearing makes the measurement of the relative ^{32}P band intensities increasingly difficult and less reliable, as well as complicating the analysis.

Secondly, estimating the intensity of the **two** bands resolved on the gel may overlook other effects. This ignores **any** ^{32}P-DNA which may be smeared between the bands or trapped in the wells. The smear can be seen to progress up from the higher mobility (single strand) band, whereas the dissociation of a quadruplex during electrophoresis would be expected to produce a smear progressing down from the upper band. It is possible that this smearing corresponds to further assembly of the quadruplex dimer within the gel-matrix.

Both these problems may be overcome by calculating the appearance of the quadruplex (or disappearance of the single strands) as a function of **total band intensity** (*i.e.* **all** radiolabel present within the lane). Another possibility is that the smearing represents non-specific electrostatic interactions between the K^+ ions or cationic drug and the polyanionic backbone of the oligonucleotide within the gel-matrix, causing charge neutralisation and reducing the rate of migration through the gel.

The data presented in Tables 3.6 and 3.7 for the effect of the various ligand were calculated as follows:

(i) Simply as ***q*** **values** (at specific ligand concentrations) determined by:

$$q = \frac{I_q}{I_q + I_{ss}} \times 100 \qquad 3.1$$

where q is the percent quadruplex, I_q and I_{ss} are the ^{32}P-intensity within the quadruplex and single strand bands respectively, at specific ligand concentrations (q). The values obtained in the presence of 20 μM ligands are shown in Table 3.6 and 3.7.

(ii) ***k*** **values**: since quadruplex formation is extremely slow, the bandshift experiments are almost certainly not at thermodynamic equilibrium, and therefore reflect a snapshot along the kinetic pathway. It is possible to estimate the apparent rate constant for formation, by making a number of assumptions. The integrated rate equations for the second order dimerisation of TR2 and the first order folding of TR4 are:

First order: $\ln[S]_t = -kt + \ln[S]_0$ 3.2

Second order: $\frac{1}{[S]_t} = -kt + \frac{1}{[S]_0}$ 3.3

where $[S]_0$ and $[S]_t$ are the concentrations of single strand at time zero and time t, and k is the association rate constant. If q is the fraction of quadruplex (determined by bandshift), then these can then be expressed as:

First order: $\ln(1-q)[S]_0 = -kt + \ln[S]_0$ 3.4

$$k = -\frac{\ln(1-q)}{t}$$ 3.5

Second order: $\frac{1}{(1-q)[S]_0} = -kt + \frac{1}{[S]_0}$ 3.6

$$k = -\frac{q}{(1-q)[S]_0 t}$$ 3.7

Since the bandshift experiments were performed after a t of 4 h, with $[S]_0$ of 5 μM, these equations can be used to estimate apparent rate constants for the assembly. The bandshifts with TR2 show that, in absence of any ligand, about 0.5% of the DNA is present as the dimeric quadruplex after 4 h incubation. This would suggest an association rate constant of 0.05 / (0.995 x 4 x 5 x 10^{-6}) = ~ 250 $M^{-1}h^{-1}$. A more thorough estimate of k from a time-course of up to 30 h, estimated a second order rate constant of 72 M^{-1} h^{-1} (Figure 3.25). These compare with a value of 10^2 M^{-1} h^{-1} which has previously been reported (Han *et al.*, 1999). This emphasises the slow kinetics of formation, suggesting a half life ($t_{1/2}$) for the single strand TR2 (*i.e.* time for 50% conversion to quadruplex) of ~33 days. The reaction is therefore extremely slow, and the 4 h incubation in these experiments represents a small fraction along the approach to equilibrium. The q for the control bandshifts for TR4 in the absence of any ligand is ~ 7.9% which gives an apparent k of 0.02 h^{-1} and a $t_{1/2}$ of about ~ 1.4 days. The apparent association rate constant k for TR2 and TR4 in the presence of 20 μM ligand

are shown in Tables 3.6 and 3.7. These indicate the apparent enhancement in the association reactions that is generated by the ligands.

(iii) It is noticeable that plots of q (band intensity) against ligand concentration, shown in Figures 3.26 to 3.30, are sigmoidal rather than the expected hyperbolic shape. All bandshift data show this relationship, which suggests some form of **co-operative binding**; possibly indicating the simultaneous interaction of more than one ligand molecule with each quadruplex. This is highly unusual and such quadruplex binding has not previously reported for any G4 ligand. It is unlikely that this represents the binding of two anthraquinones within one loop. The opening of two intercalation sites within the G-stack is also very unlikely. However, the interaction of two ligands with the symmetrical antiparallel dimer formed with TR2 seems reasonable, since the two ends are equivalent. It is also very likely that simultaneous binding of two ligands will have a much greater effect on stability. Although the data presents the interesting possibly of co-operativity, how this may occur is difficult to conceive, such an effect requires the binding of one anthraquinone to somehow affect a major change to the stable quadruplex structure to alter binding of a second ligand. Assuming cooperative binding does occur, the experimental data in the presence of the ligands for TR2 was fitted to the Hill equation:

$$\frac{q}{q_{max}} = \frac{[L]^n}{C_{50} + [L]^n} \qquad 3.8$$

where q and q_{max} are the amount of quadruplex formed in the presence of ligand concentration $[L]$ and the maximum amount of quadruplex formed (expressed as percentages), $[L]$ is ligand concentration; n is Hill coefficient (*i.e.* ligand stoichiometry), and C_{50} is a parameter which represents the square of the apparent dissociation constant, $(K_d)^2$. The unknown variables C_{50} (thus K_d) and q_{max} were found by fitting this equation to the data, assuming a value of 2 for n.

The value of q may be considered as a kinetic parameter, which indicates the maximum extent of the reaction after the 4 h incubation; active ligands will allow the reaction to be nearer equilibrium in a shorter period so will yield a higher maximum quadruplex formation. The apparent K_d may indicate the stability of the quadruplex-drug complex.

3.4.6 Structural inductions

It is notable that in bandshift studies with TR2 and TR4, none of the ligands induce any additional bands corresponding to higher molecular weight species with lower mobility. This has previously been reported on addition of the perylene, PIPER (Han *et al.*, 1999). In theory, TR2 may form a tetramer, in addition to a dimer, while TR4 may form a dimer and a tetramer, as well as the intramolecular species. The rates of formation of quadruplexes of higher molecularity will be slower, and limited by oligonucleotide concentration. Thus, lower strand concentration disfavours higher stoichiometries.

The problem of aggregation observed at high ligand concentrations may be due to the ability to induce the formation of high-order structures. For example, the ligands may act by inducing a two-step process whereby single strand TR2 goes to a $[TR2]_2$ quadruplex dimer which then forms $[TR2]_n$ polymeric quadruplex. This is more likely if the ligands act by enhancing the rates of strand association, rather than by kinetically trapping the quadruplex once it is formed and slowing the rate of dissociation. This aggregation may obscure the product of bandshifts by some drugs.

3.4.7 SAR from TR2 bandshifts

As with the structure-activity relationships determined by fluorescence-quenching, the activities of these ligands are affected by a combination of factors including the exact nature of the sidechains, their positions and the type of linkage to the anthraquinone moiety.

Position of substituents: 2,6- versus 2,7-anthraquinones

The data obtained for the G4 ligands, were used to compare anthraquinone isomers that differ only with respect to the position of substitution. Looking at the data shown in Table 3.6, it seems that anthraquinones substituted at the 2,6 position are generally more effective in promoting intermolecular quadruplex formation than their 2,7 isomers (which agrees with the findings of the fluorescence-quenching experiments). P6 gave greater activity than P14. This is also true for P7 over P11, and P9 over P12. SAC27 is only slightly more active than SAC28. However, this relationship does not appear to hold for P1 and P2, where the 2,7 isomer is more active.

Amide versus 'reverse'-amide

The benefits of the NH linkage versus CO linkage of the propionamide sidechains are far less obvious than suggested by the fluorescence-quenching experiments, which indicated that the CO-linkage is better. Looking at isomers that differ in this respect the results give no clear conclusion. NH-linked anthraquinones seem better for P10 over P1, and SAC27 over P11. Conversely, CO-linked anthraquinones are seen to be superior for P7 over SAC28, and P15 over P16.

Nature of sidechains

As seen with the fluorescence denaturation studies, the different sidechains affect the activities as a result of their length; electronegativity and bulkiness; and the (number and position) of possible positive charges. These factors influence their (hydrogen bond) interactions with either the phosphodiester backbone and/or the nucleobases of the nucleotides within the loop and/or grooves of the quadruplex. P6, P8, and P10 possess twelve-atom, thirteen-atom and five-atom backboned sidechains, capable of forming four, eight and three hydrogen bonds, respectively. Other anthraquinones with shorter or bulkier sidechains such as P3 and P4 produce less quadruplex stabilisation. These interactions affect quadruplex formation either by decreasing the rate of dissociation and/or by increasing the rate of association (which for TR2 must be at least second order). Enhancement of association may occur if ligands facilitate the correct assembly of two strands when they collide in solution, or stabilise transient complexes that would normally dissociate. Such an effect would be more pronounced at lower concentrations where fewer collisions occur.

The effects of the ligands on the TR2 bandshifts are generally in good agreement with the results of fluorescence melting studies. P6 and P8, both with long lysyl- and argyl-propionamide sidechains, capable of forming many hydrogen bonding and ionic interactions, are amongst the most active compounds. P10 appears to be very active from the bandshift experiments, and though its sidechains are only short amino-propionamides, its linkage may be significant. Unexpectedly, P15 and P16, which contain reduced alanyl- and lysyl- propionamide sidechains respectively, do not produce substantial enhancement in the bandshifts, while fluorescence melting data suggests they are active compounds. SAC14 produces good stabilisation to thermal

denaturation, however it only moderately enhances the percent quadruplex formation in the bandshift experiments. More modest activity is seen for P1 and P2 which have short sidechains that do not provide substantial activity. P3, P4, and P12 do not produce retarded bands but generate aggregates in the bandshifts. These compounds had little activity measured by thermal denaturation.

Other quadruplex-binding ligands

Looking at the bandshift experiments with all the other acridines, acridones and previously studied anthraquinones, the trisubstituted acridines BRACO19 and BRACO20 produce the greatest activity among these ligands. However, this is not significantly better than the activity seen for the best novel anthraquinones (P6, P8 and P10), as also noted in the melting studies. The activity of the 2,6-AQ, BSU1078 was comparable to moderately active members of the novel series. However, the other anthraquinones did not produce bandshifts, possibly due to the formation of aggregated structures in the wells (as discussed earlier). The acridones BRACO4 and JMACO9 both show little ability to promote quadruplex formation.

Summary

The bandshift assay gives good **qualitative** information on the relative potencies of different quadruplex-binding ligands, though a rigorous quantitative analysis is difficult. As with the fluorescence results, these data do not warrant determining an absolute order of activity, the ligand have been grouped into different classes based on the TR2 experiments.

P6, P8, and P10 can be seen to produce the greatest activities, promoting the highest percent quadruplex under the experimental conditions. 20 μM concentrations of these compounds allow over 20% of TR2 to form quadruplex, which is a substantial increase over folding in the absence of ligand. The same concentrations of P1, P2, P7, P15, SAC14, and SAC49 produce more modest quadruplex association of between 10% and 20%. P9, P11, P13, P14 P16, D1, D2, SAC27, SAC28, and SAC47 all give low activity, allowing under 10% of TR2 to form a quadruplex. P3, P4, P5, and P12 failed to produce noticeable increases in the amount of quadruplex. P5, P15, and P16 seem to be relatively more active from the fluorescence melting studies than from the bandshift

experiments. However, this assay cannot distinguish between ligand-induced promotion of association or stabilisation of the quadruplex itself reducing dissociation.

3.4.8 Limitations of electrophoresis test

It is important to appreciate the limitations of the bandshift technique used in this chapter. (i) Quadruplex assembly reactions are generally sensitive to the reaction conditions (cation and strand concentration changes), and this has significant implications for their study by electrophoresis. As discussed earlier, this can result in perturbations to the populations during the process of electrophoresis. (ii) Electrophoresis is not performed under exactly the same conditions as those used for forming the complexes, and this might disturb the equilibrium. Such re-equilibration might arise during the transfer of reaction contents to the gel matrix. However, this may be a minimal effect as the rate of re-equilibration will be long when compared to the electrophoresis process, as quadruplexes are known to dissociate very slowly. The changes were minimised by ensuring that K^+ concentration during electrophoresis **was** similar to the incubation conditions, and re-equilibration was minimised by reducing the temperature to 4°C. Thus, we assume that the drug-DNA equilibrium or time-dependent reaction was **not** altered during electrophoresis. (iii) Another possible complication arises from the physical process of electrophoresis. Although both the gel and running buffer contained 20 mM KCl, higher salt would perturb the process by 'carrying' the charge instead of the macromolecules (though higher salt conditions were not required) (Hardin *et al.*, 2001).

3.4.9 Conflicting implications

The melting studies indicate that quadruplex folding and unfolding can be followed during successive melting and annealing of HT1.5, HT3.5 and OT3.5, though hysteresis is seen if the rates of heating and cooling are too rapid. This suggests that rates of association and dissociation are not excessively slow. However, the bandshift experiments which assay extent of quadruplex formation by TR2 at room temperature over a 4 h incubation find very little quadruplex forms in this time in the absence of G4 ligands; suggesting a very slow rate of formation.

These seemingly conflicting conclusions for the quadruplex stability may simply arise form the fact that the bandshift experiments assayed the rate of formation at room

temperature; while the fluorescence melting experiments monitored quadruplex formation over a large and much higher temperature range, over which the kinetics will be faster. Much published research on quadruplex has used UV melting or FRET to study kinetics and stability of quadruplex formation (Mergny *et al.*, 1998; Mergny & Maurizot, 2001; Darby *et al.*, 2002). Such techniques have also been applied to determine ligand stabilsation (Riou *et al.*, 2001; Mergny *et al.*, 2001; Han *et al.*, 1999a). Typically only the melting phase is measured for convenience, though it is assumed that annealing is similar. The bandshift experiments are also in good agreement with other reports using similar oligonucleotide sequences for studying sequence-dependent and ligand-induced quadruplex formation (Sen & Gilbert 1990; Han *et al.*, 1999b; Han *et al.*, 2001; Koeppel *et al.*, 2001; Teulade-Fichou *et al.*, 2003).

4 Effects of Na^+ and K^+ on Quadruplex Formation Studied by Fluorescence Melting Studies

4.1 Introduction

4.1.1 Monovalent cation interactions

Though a variety of nucleic acid sequences form quadruplex topologies, these all contain stacked G-quartets, which produce an internal cleft of eight carbonyl oxygen atoms. As a result, all quadruplex complexes share a requirement for the chelation of monovalent cations within this pocket. K^+ ions generally lie between the G-quartets (Kang *et al.*, 1992; Schultze *et al.*, 1999; Parkinson *et al.*, 2002; Haider *et al.*, 2002), while Na^+ is has often found to be coplanar (Laughlan *et al.*, 1994; Phillips *et al.*, 1997; Schultze *et al.*, 1994). Furthermore, quadruplex stability is unique among nucleic acid structures in showing a specificity for different monovalent cations. Na^+ and K^+ are the most active, and provide good stability at physiological concentrations. Generally, the effects of K^+ ions are superior, producing the most stable complexes (Sen & Gilbert, 1992). Importantly, Na^+ and K^+ may promote the formation of different complexes (Sen & Gilbert, 1990; Muira *et al.*, 1995). This may result from their different sizes, their hydration energies (Hud *et al.*, 1996) and their interaction in the loops (Jing *et al.*, 1997). A detailed discussion of cation interactions can be found in chapter 1.

4.1.2 *Oxytricha* & human telomeric sequences

Though most quadruplexes are G-rich, the number of guanines and the nature of the loops vary. Small changes in the sequences of quadruplex complexes can have profound effects on their folding and may affect their thermodynamic stability (Jing *et al.*, 1997; Marathias & Bolton, 1999). The human and *Oxytricha* telomeric sequences consist of stacks of three and four G-quartets, respectively, with TTA and TTTT in the connecting loops. All quadruplex-forming sequences form extremely stable complexes that dissociate slowly.

The *Oxytricha* telomeric repeat produces kinetics that are far slower than most other sequences such as the human repeat. The dissociation of a pre-formed quadruplex of $d(T_4G_4T_4G_4T_4G_4T_4G_4)$ along with bandshifts estimated k_a and k_d values of $0.0017s^{-1}$ and $4.8 \times 10^{-5}\ s^{-1}$ ($K_{eq} \approx 35$) in 50 mM Na^+, while these values were $0.023\ s^{-1}$ and $1.1 \times$

10^{-5} s^{-1} in 50 mM K^+ ($K_{eq} \approx 2100$) (Raghuraman & Cech, 1990). Other estimates of k_a for dimeric quadruplex formation by the *Oxytricha* telomeric repeat include 2.6 x 10^{-3} M^{-1} h^{-1} for $[d(T_4G_4T_4G_4)]_2$ in 80 mM Na^+ (Fang & Cech, 1993), 5.8 $M^{-1}s^{-1}$ for the same process in 50 mM Na^+ (Guo *et al.*, 1993), and 4.5 x 10^{-4} $M^{-1}s^{-1}$ at 60°C in 1 M Na^+ for a 49-mer containing two 3' *Oxytricha* repeats ($T_4G_4T_4G_4$) (Sen & Gilbert, 1990). This compares to a k_a of 10^2 M^{-1} h^{-1} that was determined by bandshift for a 24-mer containing two human repeats in 100 mM K^+ (Han *et al.*, 1999b).

Both dimeric and monomeric quadruplex complexes formed by *Oxytricha* telomeric repeats have been extensively characterised by NMR (Smith & Feigon, 1992; Schultze *et al.*, 1999; Wang & Patel, 1995; Smith *et al.*, 1995) and X-ray crystallography (Kang *et al.*, 1992; Schultze *et al.*, 1999; Haider *et al.*, 2002). Similar human telomeric sequences have also been studied (Parkinson *et al.*, 2002). Several oligonucleotides that form tetrameric quadruplex have been examined, although these lack intervening loops that may influence the structure. A variety of cationic conditions were used in these structural investigations.

4.1.3 Aims

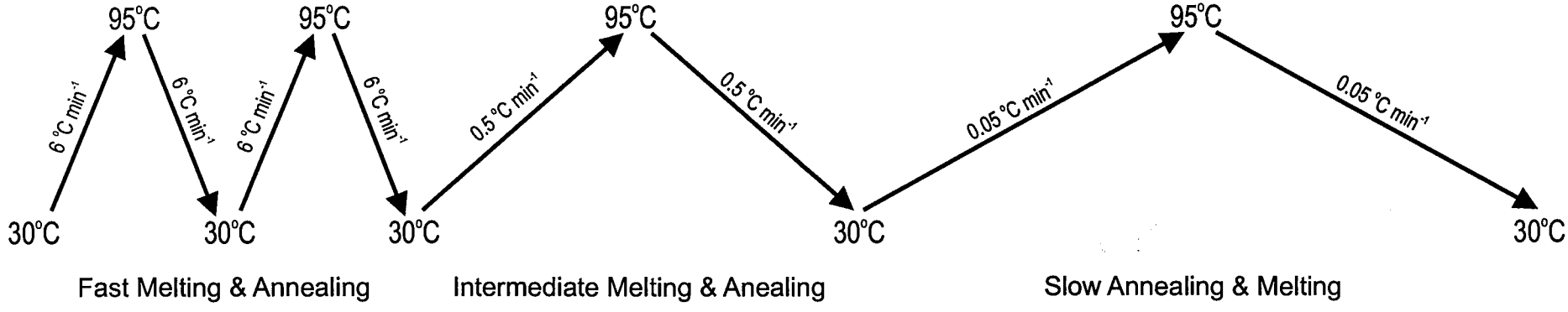
The aim of this chapter is to examine the stability of the human and *Oxytricha* telomeric sequences and to compare their properties in the presence of Na^+ and K^+. For this, HT3.5 and OT3.5 oligonucleotides were used. Within the cell, most G-rich sequences will be presented along with their complementary C-rich strands, which will generate competing duplex structures. The effect of adding an excess of complementary strand was studied, in order to compare quadruplex and duplex stability. Fluorescence melting experiments were carried out at varying rates of heating and cooling, and using two different schemes (see Figure 4.1). The reasons and implications of this will be considered in more detail in the results and discussion.

4.2 Results

4.2.1 Inter- or intramolecular complexes?

Although these oligonucleotides are expected to form intramolecular quadruplexes, they could in theory, also associate to form intermolecular structures. This possibility was examined by determining the quadruplex melting profiles at different DNA concentrations. The T_m values for an unimolecular reaction should not vary with

Scheme 1



Scheme 2

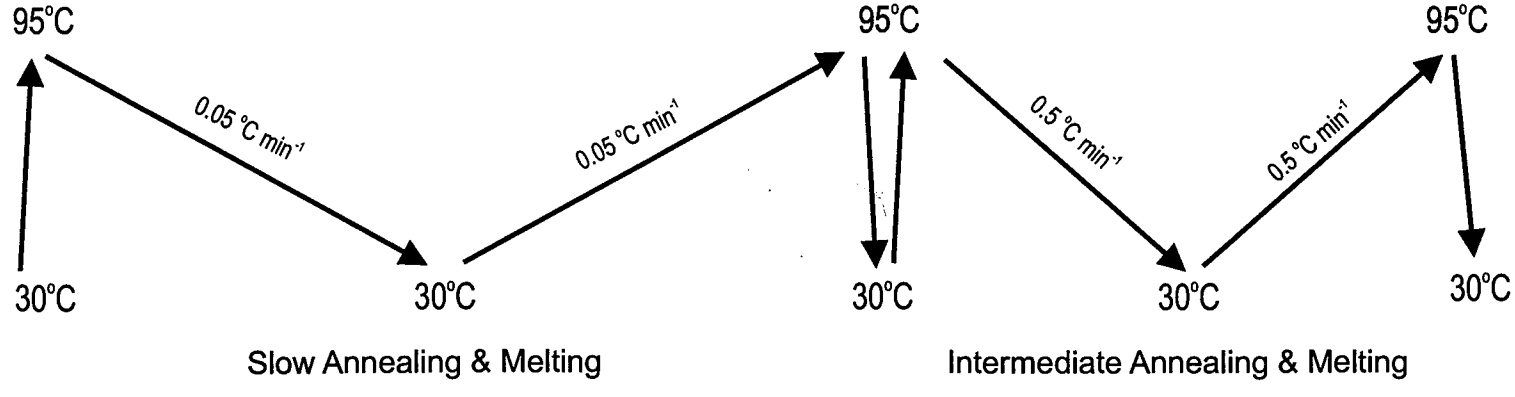


Figure 4.1. A representation of the schemes for heating and cooling fluorescently labelled oligonucleotides during melting studies.

concentration, while higher order reactions will increase with increasing DNA concentration. These control experiments were not performed for all conditions studied, but representative curves for OT3.5 and HT3.5 are shown in Figure 4.2. It can be seen that, as expected these transitions are independent of concentration and confirm that it is intramolecular quadruplex formation which is observed.

4.2.2 *Oxytricha* telomeric sequence, OT3.5

On the basis of previous kinetic analyses of *Oxytricha* telomeric repeats (Raghuraman & Cech, 1990; Fang & Cech, 1993; Guo *et al.*, 1993; Sen & Gilbert, 1990) it was anticipated that melting experiments on OT3.5 would be hampered by its slow kinetics of folding and unfolding, producing considerable hysteresis. Hysteresis is indeed observed as described below, showing that the rate of temperature change is greater than the rate of association and/or dissociation, which means that the melting profile is not at thermodynamic equilibrium.

Fast rate of heating and cooling (6 °C min^{-1})

20 µl samples of 0.25 µM OT3.5 were prepared in the presence of different concentrations of monovalent cation between 10 and 100 mM. The samples were heated to 95°C, maintained at this temperature for 5 min before annealing to 30°C, where the temperature was held for a further 5 min. This cycle was then repeated by melting to 95°C, followed by annealing to 30°C. Continuous fluorescence measurements were taken throughout both the melting and annealing.

Typical melting and annealing profiles from one experiment at several Na^+ and K^+ concentrations are presented in Figure 4.3 and the T_m values from a range of experiments are summarised in Table 4.1. The T_m values are determined from the inflection point, taken from first derivative plots. Figure 4.3 shows that the first and second annealing profiles are very similar, as expected, producing T_m values which vary by only about 0.5°C. However, these annealing curves are very different to the melting curves, which are different from each other.

The first melting profiles may represent the melting of aberrant structures formed before the oligonucleotide is properly annealed in each ionic environment. These initial

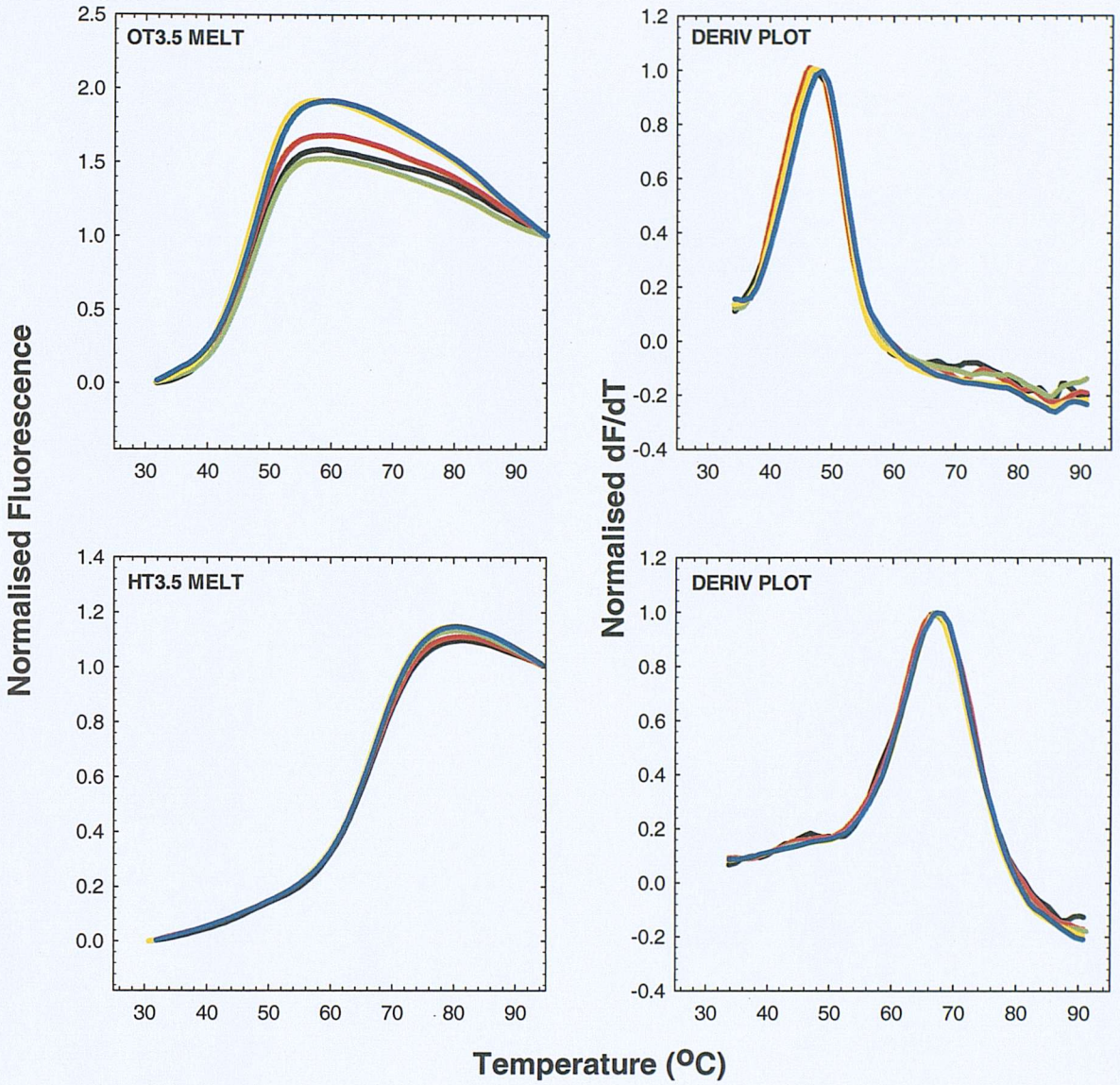


Figure 4.2. Examples of the effect of increasing oligonucleotide concentration on normalised melting profiles (left panels), accompanied by normalised first derivatives of the profiles (right panels). Those obtained with HT3.5 in 50 mM KP at 0.5 °C min^{-1}, are in the upper panels; and OT3.5 in 10 mM NaP + 40mM NaCl at 0.05 °C min^{-1}. Plots show oligonucleotide concentration as 0.0625 µM (black), 0.125 µM (red), 0.25 µM (green), 0.5 µM (yellow) and 1.25 µM (blue).

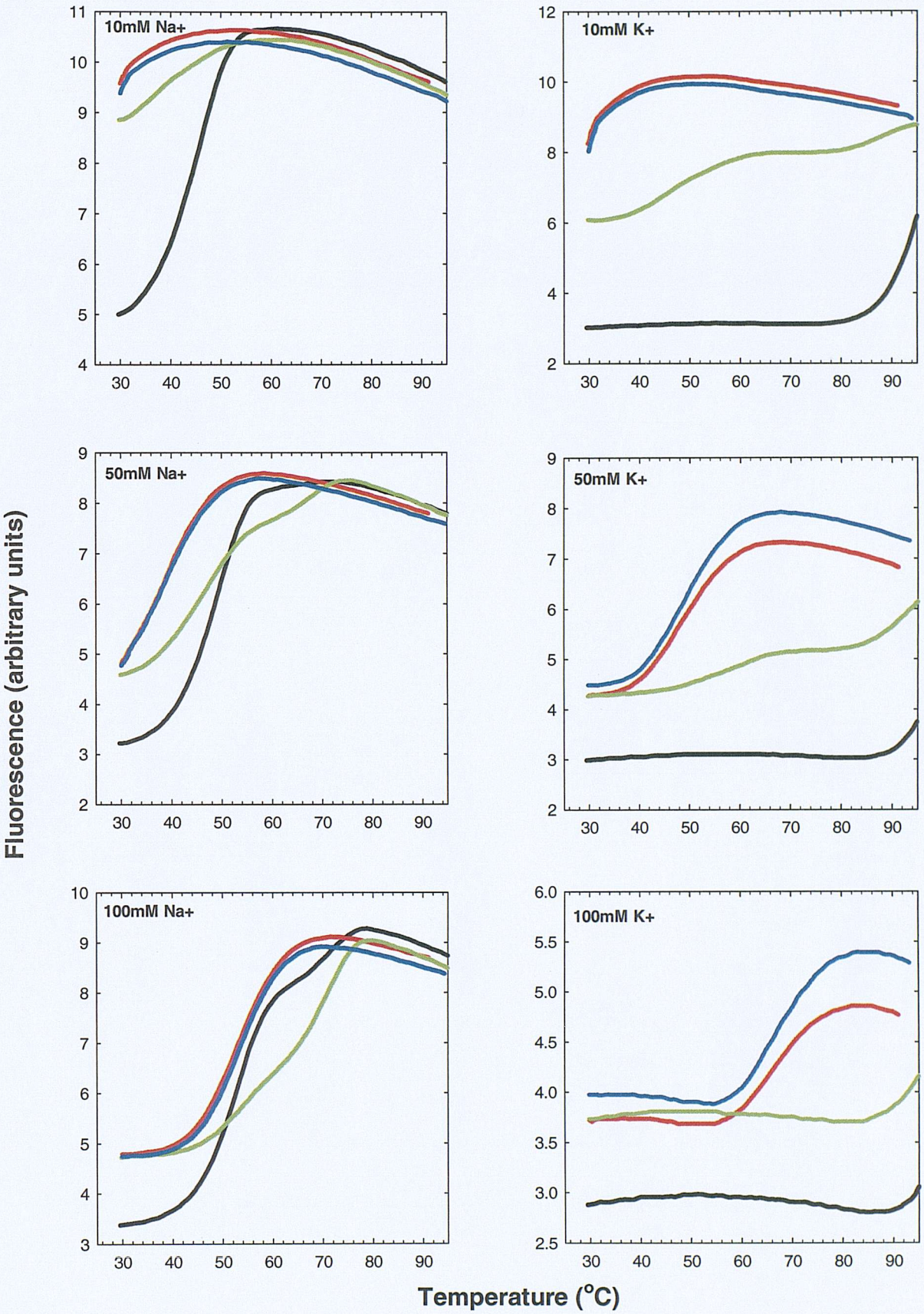


Figure 4.3. Examples of melting and annealing profiles for OT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at 6°C min^{-1}. Profiles show the first melt (black), first anneal (red), second melt (green) and second anneal (blue).

Table 4.1. Apparent T_m values for OT3.5 determined from successive fast heating and cooling cycles at different concentrations of Na^+ and K^+, at the rate of heating and cooling of 6°C min^{-1}.

$[M^+]$ (mM)	T_m *1st* melting profile (°C)	T_m annealing profile (°C)	T_m *2nd* melting profile (°C)
Na^+			
10	46.1 ± 0.6	< 35 [a]	< 35 [a]
20	47.0 ± 0.3	< 35 [a]	< 35 [a]
30	47.8 ± 0.8	< 35 [a]	39.9 ± 0.4 65.8 ± 0.5
50	49.7 ± 1.3	38.7 ± 0.6	47.1 ± 0.7 67.5 ± 0.5
75	51.5 ± 0.9	45.7 ± 0.4	51.0 ± 0.7 69.0 ± 0.2
100	53.6 ±1.0 71.1 ± 0.6	52.7 ± 1.4	54.9 ± 1.0 70.8 ± 0.8
K^+			
10	> 90 [a]	< 35 [a]	46.3 ± 0.9 88.3 ± 0.3
20	> 90 [a]	< 35 [a]	48.4 ± 0.61 > 90 [a]
30	> 90 [a]	39.1 ± 0.9	52.4 ± 2.1 > 90 [a]
50	> 90 [a]	49.5 ± 0.8	58.73 ± 1.3 > 90 [a]
75	> 90 [a]	58.0 ± 0.5	62.3 ± 0.4 > 90 [a]
100	> 90 [a]	65.9 ± 1.8	> 90 [a]

Values are presented as the average ± standard errors from three independent experiments.
a – indicates that the T_m was too high (> 90°C) or too low (< 35°C) to be measured.

melts are assumed to be less relevant than the second melts. It can be seen that these second melting profiles have an unusual biphasic nature under all conditions. This effect is probably a result of the fact that these experiments are not at thermodynamic equilibrium, and suggests that more than one quadruplex structure is generated during the fast annealing reactions. This effect is seen in the presence of both Na^+ and K^+. In each case the higher transition is close to that seen in the first melting profiles, consistent with the suggestion that this represents improperly folded quadruplex. Although hysteresis is observed under all these conditions, it is most pronounced at low cation concentrations.

To reduce this hysteresis, the experiments were repeated at slower rates of heating and cooling (0.5°C min^{-1} and 0.05°C min^{-1}). Since the previous experiments had shown biphasic melting curves after annealing at 6°C min^{-1}, these experiments were performed in two different ways, and the results compared. In the first method (scheme 1, Figure 4.1) the samples were heated and cooled at 6°C min^{-1}, as in the previous section, before melting and annealing at the slower rate. Under these conditions, the slow melt may contain improperly folded structures generated during the rapid annealing. In the second method (scheme 2, Figure 4.1) the samples were first rapidly denatured and then slowly annealed and melted again. In this case, the slower annealing is more likely to generate properly folded complexes, which are examined in the subsequent melting profiles.

Intermediate rate of heating and cooling (0.5 °C min^{-1})

Figures 4.4 and 4.5 show representative melting and annealing curves for OT3.5, determined using scheme 1 or scheme 2 respectively at a heating and cooling rate of 0.5°C min^{-1}. The T_m values determined over a range of ionic conditions are summarised in Table 4.2. Examination of these data reveal that the annealing profiles generally show the usual sigmoidal profiles, while the melting curves often have unusual shapes. Looking first at the annealing curves it can be seen that similar T_m values are obtained for scheme 1 and scheme 2 at all concentrations of cations. As expected, these complexes are more stable in K^+ than Na^+ and the melting temperatures increase with ionic strength.

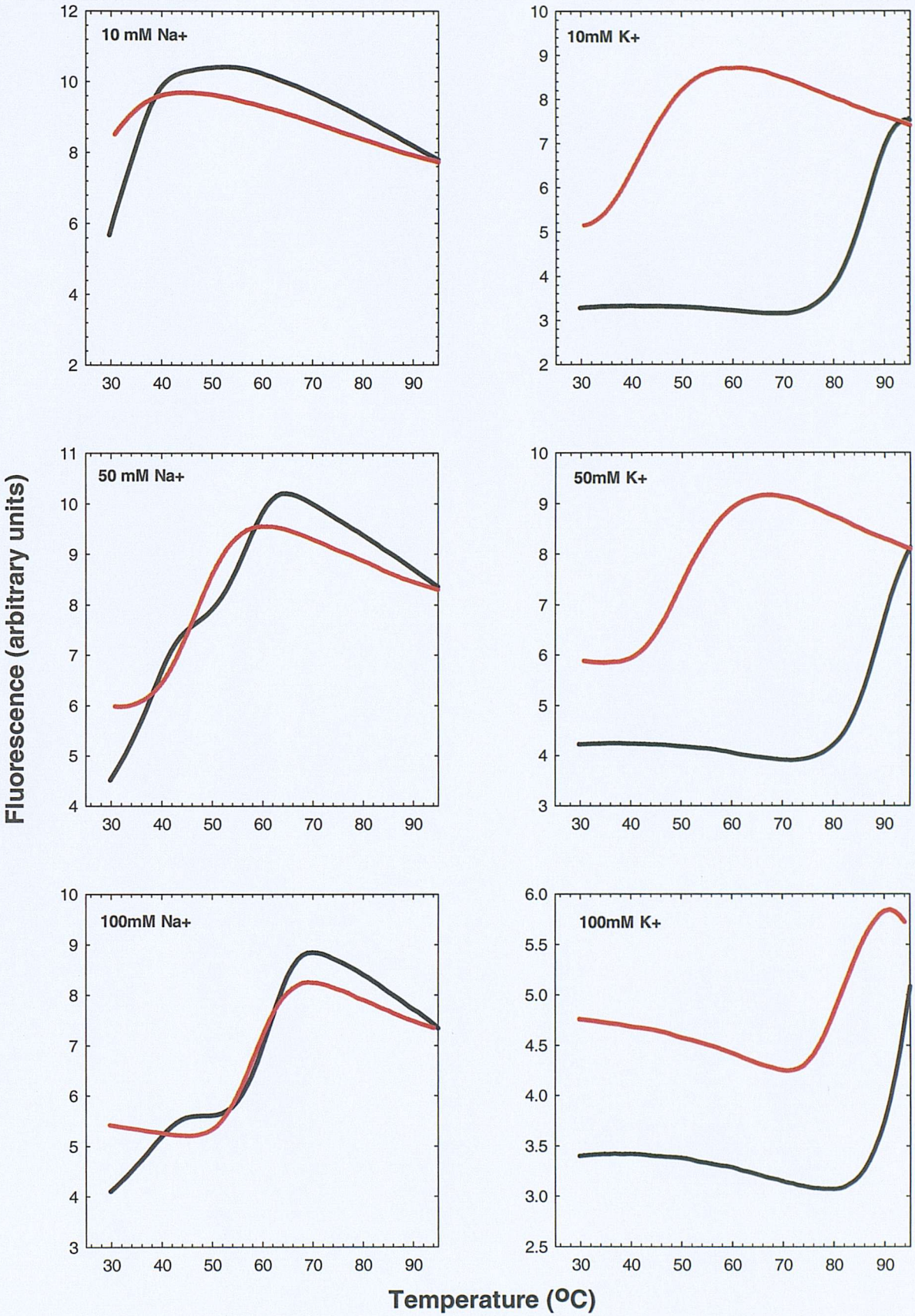


Figure 4.4. Examples of melting (black) and annealing (red) profiles for OT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at 0.5 °C min^{-1}, following scheme 1 (Figure 4.1).

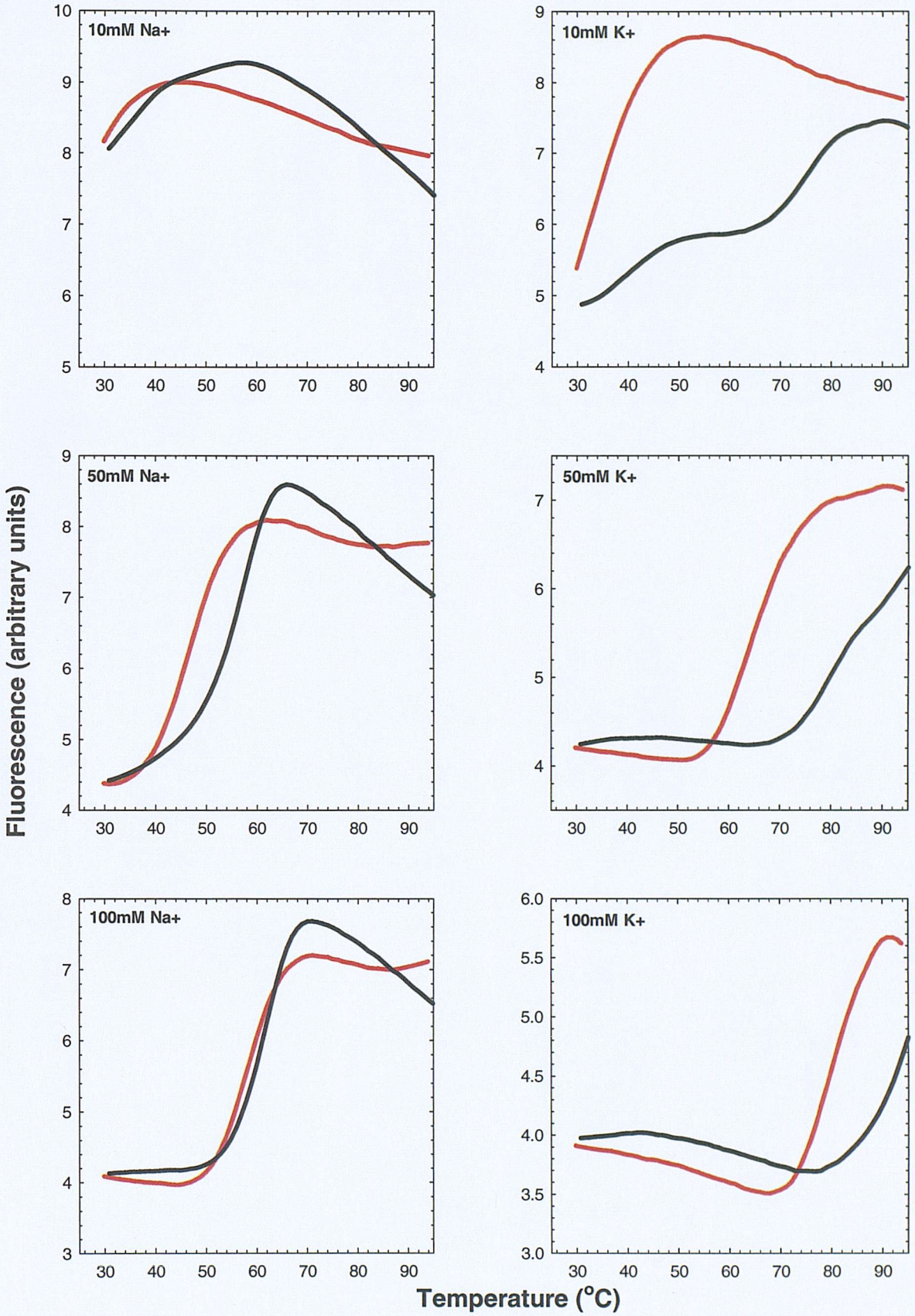


Figure 4.5. Eamples of melting (black) and anealing (red) profiles for OT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at 0.5 °C min^{-1}, following scheme 2 (Figure 4.1).

Table 4.2. Apparent T_m values for OT3.5 determined from successive fast heating and cooling cycles at different concentrations of Na^+ and K^+, at the rate of heating and cooling of 0.5°C min^{-1}.

	Scheme 1		*Scheme 2*	
$[M^+]$ (mM)	T_m melt (°C)	T_m anneal (°C)	T_m anneal(°C)	T_m melt (°C)
Na^+				
10	35.8	< 35 [a]	< 35 [a]	< 35 [a]
20	36.3 ± 0 52.2 ± 0.1	< 35 [a]	< 35 [a]	36.8 ± 0.1 54.2 ± 0
30	37.1 ± 0.1 54.7 ± 0.1	38.1 ± 0.6	37.3 ± 0.1	37.7 ± 0.4 55.5 ± 0.2
50	38.6 ± 0.1 57.1 ± 0.1	46.9 ± 0.4	46.1 ± 0.5	57.5 ± 0.3
75	38.9 ± 0.4 59.4 ± 0.1	53.3 ± 0.6	53.1 ± 0.5	59.7 ± 0.2
100	39.1 ± 0.3 61.6 ± 0.3	57.9 ± 0.6	58.1 ± 0.25	61.9 ± 0.4
K^+				
10	87.2 ± 0	41.2 ± 0.9	< 35 [a]	41.2 ± 0.9 75.5 ± 0.3
20	89.25 ± 0.5	49.6 ± 0.3	45.1 ± 0.3	77.3 ± 0.5
30	> 90 [a]	57.4 ± 1.4	51.3 ± 3.0	78.3 ± 0
50	> 90 [a]	66.9 ± 1.4	65.3 ± 1.7	80.5 ± 0.3
75	> 90 [a]	75.4 ± 1.2	76.1 ± 2.4	> 90 [a]
100	> 90 [a]	81.5 ± 0.8	80.4 ± 0.5	> 90 [a]

Values are presented as the average ± standard errors from three independent experiments.
a – indicates that the T_m was too high (> 90°C) or too low (< 35°C) to be measured.

In contrast, the **melting** curves have unusual shapes and often reveal two transitions. In Na^+-containing buffers, biphasic **melting** curves are always observed when the complexes are prepared by scheme 1. Similar results are obtained with scheme 2 at low Na^+ concentrations, though higher concentrations (50 mM and above) produced a single melting transition, which occurs at the same temperature as the second transition when the complexes are prepared by scheme 1. It should also be noted that the first transition only changes by 3°C on increasing from 10 to 100 mM Na^+, while the second transition shows a much greater dependence on ionic strength. At low ionic strengths the annealing T_m appear to correspond to the first melting transition, though at Na^+ concentrations above 30 mM this similarity is no longer apparent. At the highest Na^+ concentration (100 mM), the annealing T_m is much closer to the second melting transition.

The differences between the annealing and melting profiles are even more pronounced in K^+ containing buffers. The annealing profiles for scheme 1 and scheme 2 are similar at high K^+ concentrations, but these diverge at lower concentrations where the T_m values from scheme 2 are lower than that from scheme 1. It can be seen that the melting profiles occur at much higher temperatures than the annealing curves and this is most pronounced for scheme 1. Once again, two transitions are often observed for the melting curves by scheme 2, though this is not evident for scheme 1. Presumably this may simply be because the complex is too stable and most of the melting profile is not accessible within the temperature range studied.

Slow rate of heating and cooling (0.05 °C min^{-1})

Since the results in the previous section show that there is considerable hysteresis between the melting and annealing profiles, these experiments were repeated at even slower rates of heating and cooling (0.05°C min^{-1} = *i.e.* 3°C h^{-1}). It was hoped that these exceptionally slow conditions would resolve the apparent differences. Once again the complexes were treated by either scheme 1 (melting after fast annealing) or scheme 2 (melting after slow annealing). Representative melting and annealing profiles for these two schemes are shown in Figures 4.6 and 4.7 and T_m values for all the data are summarised in Table 4.3.

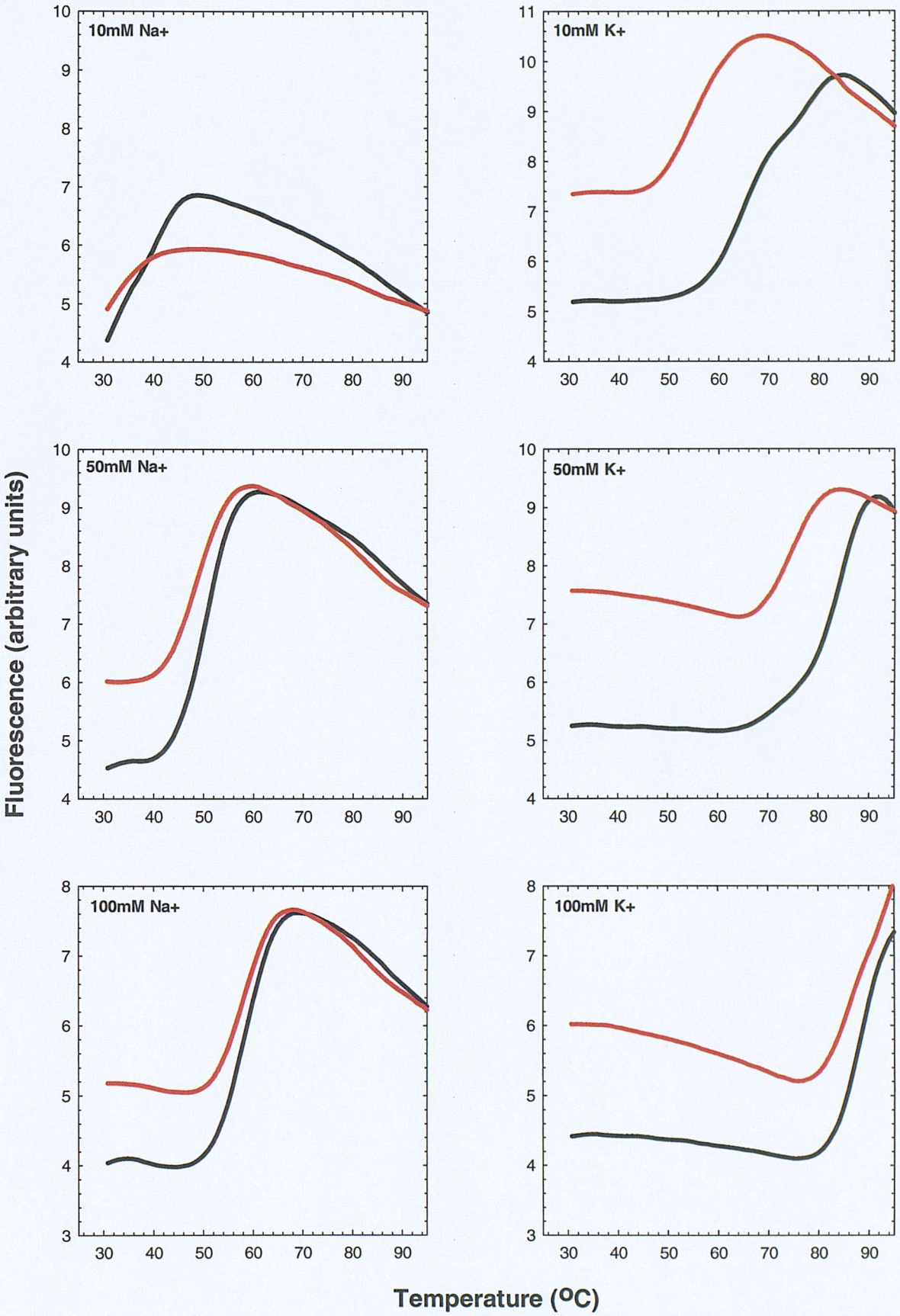


Figure 4.6. examples of melting (black) and annealing (red) profiles for OT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at 0.05 °C min^{-1}, following scheme 1 (Figure 4.1).

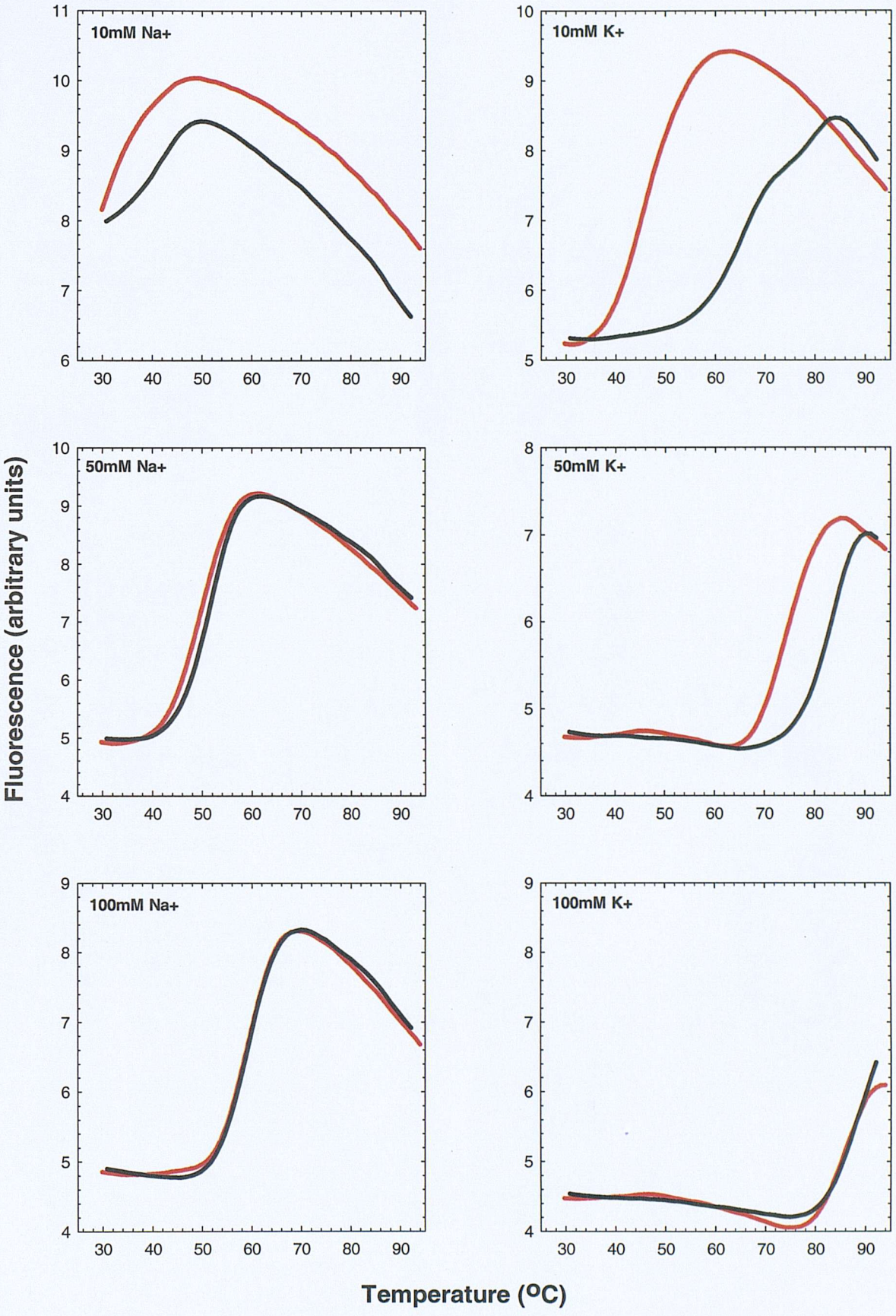


Figure 4.7. Examples of melting (black) and annealing (red) profiles for OT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at 0.05 °C min^{-1}, following scheme 2 (Figure 4.1).

Table 4.3. Apparent T_m values for OT3.5 determined from successive fast heating and cooling cycles at different concentrations of Na^+ and K^+, at the rate of heating and cooling of 0.05°C min^{-1}.

	Scheme 1		*Scheme 2*	
[M^+] (mM)	T_m melt (°C)	T_m anneal (°C)	T_m anneal (°C)	T_m melt (°C)
Na^+				
10	40.4 ± 0.1	< 35 [a]	< 35 [a]	42.2 ± 0.1
20	43.9 ± 0.4	37.7 ± 0	36.3 ± 0.1	45.0 ± 0
30	47.0 ± 0.2	42.0 ± 0.2	42.3 ± 0.3	47.5 ± 0.2
50	51.2 ± 0.4	46.5 ± 0.5	49.8 ± 0	51.9 ± 0.2
75	55.7 ± 0.9	54.9 ± 0.2	55.4 ± 0.1	55.7 ± 0.1
100	60.1 ± 1.4	59.4 ± 1.1	59.5 ± 0.1	59.3 ± 0.1
K^+				
10	65.5 ± 0.5 78.6 ± 0.1	54.9 ± 0.4	45.5 ± 0.1	65.6 ± 0.4 79.7 ± 0.2
20	67.8 ± 1.0 80.4 ± 0.6	61.7 ± 0.5	58.0 ± 0.	67.7 ± 0.1 81.2 ± 0.3
30	69.7 ± 0.9 82.0 ± 0.8	67.8 ± 0.6	64.9 ± 0.4	70.8 ± 0.4 82.3 ± 0.5
50	84.8 ± 0.5	75.5 ± 1.0	75.0 ± 0.5	84.3 ± 0.5
75	87.2 ± 0.9	82.1 ± 0.6	81.6 ± 0.3	86.2 ± 0.5
100	>90 [a]	> 90 [a]	> 90 [a]	> 90 [a]

Values are presented as the average ± standard errors from three independent experiments.
a – indicates that the T_m was too high (> 90°C) or too low (< 35°C) to be measured.

Looking first at the data in the presence of Na^+, it can be seen that as expected the annealing profiles are very similar for scheme 1 or scheme 2. Under these conditions the melting profiles are now single sigmoidal transitions and the T_m values are similar for the two schemes. At higher Na^+ concentrations produce melting and annealing profiles which are very similar and there is no hysteresis. As the Na^+ ion concentration is lowered, some hysteresis is evident (6°C with 20 mM Na^+), though this is much less than that seen with the higher rates of heating.

In the presence of K^+ the curves still show considerable hysteresis, and depend on the route of preparation (scheme 1 or scheme 2), though these profiles are improved relative to the faster rates of heating. The annealing curves always showed monophasic transitions, though differences are apparent between scheme 1 and scheme 2 at the lower K^+ concentrations, in which the complexes anneal at lower temperatures for scheme 2. The annealing profiles for these two schemes would be expected to be identical, as they are both derived from slow annealing of melted complexes. The differences therefore suggest that the starting DNA is not fully single stranded when annealed by scheme 1. Surprisingly, we find that the melting profiles from scheme 1 and scheme 2 are very similar, though biphasic curves are again evident at low K^+ concentrations. However, these two transitions are much closer together than when the experiments were performed at faster rates of heating and cooling. Nonetheless, there are still considerable differences between the melting and annealing curves, emphasising the exceptionally slow kinetics of this reaction.

It can also be seen that in many instances the fluorescence signal decreases at high and low temperatures, where the oligonucleotide should be in the fully folded or single stranded conformations. This could be due to the degradation of the DNA with time, as a result of the long exposure to high temperatures and light, or it may reflect the formation of other (aggregated?) DNA structures. These will be considered further in the discussion.

OT3.5 plus complementary strand

Fluorescence melting experiments were also used to compare the relative stability of the Watson – Crick duplex and the intramolecular quadruplexes formed by OT3.5 as described by Risitano and Fox (2003). This was achieved by adding a 10-fold excess of the complementary C-rich strand to OT3.5 under the Na^+ and K^+ conditions already described above for the oligonucleotide alone. The C-rich complementary strand can bind to the G-rich strand to form a duplex that will compete with quadruplex formation. Since a duplex is more rigid than a random coil, the time-averaged distance between the fluorophore and quencher will be greatest for duplex DNA, resulting in a larger fluorescence signal. The relative order of fluorescence signals for the different structures will therefore be duplex (F_{ds}) > single strand (F_{ss})>> quadruplex (F_q). These experiments were performed in Na^+ or K^+ at different rates of heating and cooling, and the results are presented in Figures 4.8 to 4.10. The T_m values derived from these curves are presented in Tables 4.4 to 4.6. In each case the first melting transitions are disregarded as these correspond to the melting of the pre-formed quadruplex, and do not represent an equilibrium between the quadruplex and duplex forms.

Figure 4.8 shows the annealing and melting profiles for OT3.5 in the presence of varying concentrations of Na^+ and K^+ produced by heating and cooling at the fast rate (6°C min^{-1}). In the presence of 10 mM cations these curves are unusual, and no clear melting transitions are observed. However, at cation concentrations of 20 mM and above, single transitions are observed for both melting and annealing curves in which the fluorescence is lower at the higher temperature. These curves therefore seem to correspond to the simple transition between duplex (higher fluorescence) and random coil forms. The T_m values for these transitions are summarised in Table 4.4. It can be seen that there is no hysteresis for these curves; and the T_m values for melting and annealing are the same. In addition, these transitions do not depend on the nature of the cation, and are the same in both K^+ and Na^+ buffers. Each of these observations is consistent with the suggestion that this represents a duplex to single strand transition, with no contribution from a quadruplex. Since duplex formation may be much faster than quadruplex formation it is possible that these curves do not represent true thermodynamic equilibrium, but are dominated by duplex formation before the quadruplex has started to fold. These experiments were therefore repeated at slower rates of heating and cooling.

Fluorescent melting curves in the presence of the complementary oligonucleotide at the intermediate rate of temperature change are shown in Figure 4.9, and the T_m values are summarised in Table 4.5. In the presence of Na^+ the profiles are very similar to those obtained at the faster rate of heating and again show no hysteresis between the annealing and melting curves. These again appear to represent duplex to single strand transitions, with no evidence for quadruplex formation. However, in the presence of K^+ the profiles are different. As the temperature decreases, there is a reduction in the fluorescence signal, consistent with quadruplex formation. At lower temperatures, the fluorescence increases again, as expected if there is some duplex formation. These curves are therefore characterised by two T_m values, which are presumed to correspond to quadruplex to single strand, and single strand to duplex. It should be noted, that in contrast to the reactions with Na^+, the fluorescence at low temperature is lower than that of the single strand. This may result from the persistence of some quadruplex in the presence of the complementary strand, or it may represent the formation of overlapping duplexes which place the fluorophore and quencher in close proximity (discussed further later). The lower transition is similar to that seen with Na^+, as expected for a duplex melt, and shows no hysteresis between annealing and melting profiles. The higher transition shows considerable hysteresis and is consistent with quadruplex formation, though these values are not the same as those observed with the quadruplex alone (Table 4.2).

4.2.3 Human telomeric sequence, HT3.5

Fast rate of heating and cooling (6 °C min^{-1})

Chapter 3 described the effects of novel ligands on melting and annealing profiles of HT3.5, HT1.5 and OT3.5. These experiments were performed in 50 mM K^+ phosphate buffer pH 7.4, representing a total of 90.1 mM K^+, and at a rate of heating and cooling of 6°C min^{-1}. Under these experimental conditions there is very little hysteresis (under 2°C) between the annealing and melting curves of the intramolecular quadruplex formed by HT3.5. However, the second order assembly of HT1.5 oligonucleotide into an intermolecular quadruplex does show slower kinetics and significant hysteresis (~15°C).

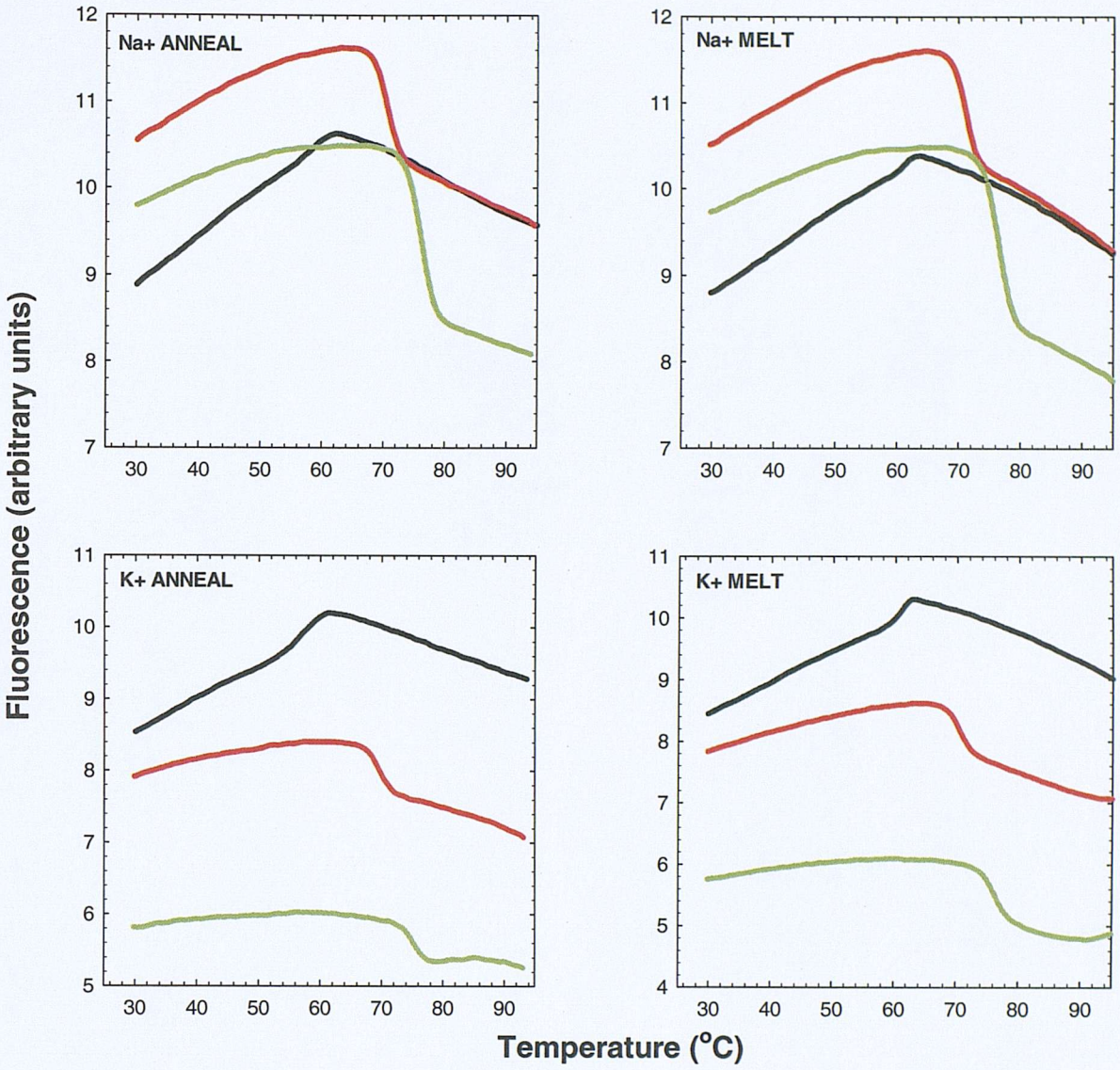


Figure 4.8. Examples of melting and annealing curves for OT3.5 in the presence of 10-fold excess of the complementsary C-rich strand. The reactions were performed in different concentrations of Na^+ and K^+ as indicated (10 mM black, 50 mM red, 100 mM green). The temperature was changed at a rate of 6 °C min^{-1}.

Table 4.4. Apparent T_m values for the melting and annealing observed with OT3.5 in the presence of the complementary oligonucleotide at different concentrations of Na^+ and K^+, determined at the rate of heating and cooling of 6°C min^{-1}.

$[M^+]$ mM	Annealing transition T_m (°C)	Melting transition T_m (°C)
Na^+		
10	-	-
20	64.9 ± 0.4	65.6 ± 0.4
30	66.9 ± 0.3	67.4 ± 0.2
50	70.3 ± 0.1	70.8 ± 0.4
75	73.2 ± 0.4	73.4 ± 0.4
100	75.7 ± 0.5	75.9 ± 0.8
K^+		
10	-	-
20	65.1 ± 0.4	65.4 ± 0.2
30	66.7 ± 0.3	67.2 ± 0.2
50	69.9 ± 0.3	70.5 ± 0.2
75	72.8 ± 0.2	73.2 ± 0.2
100	75.2 ± 0.1	75.8 ± 0.5

Values are the average ± standard error from three independent experiments.

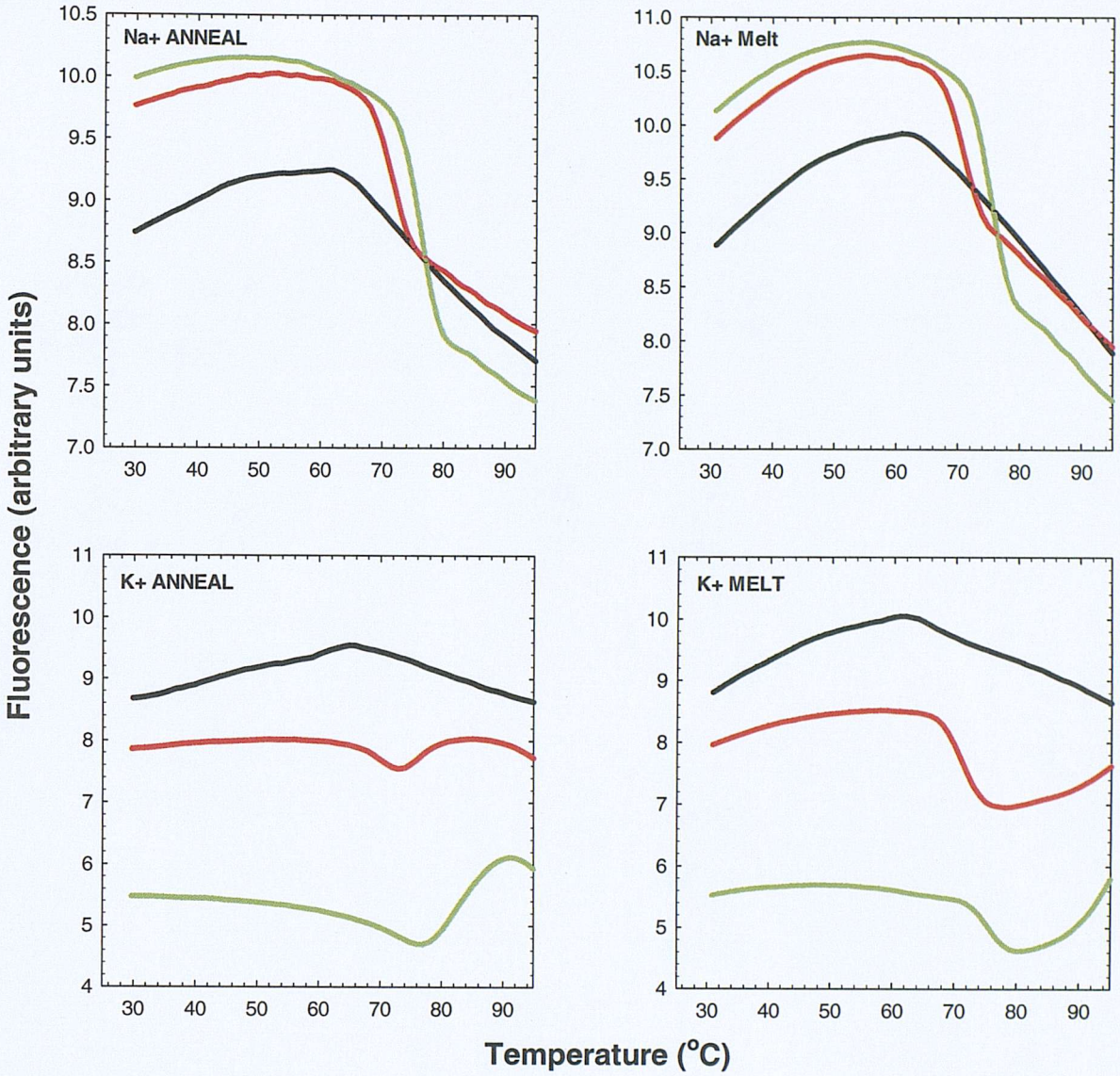


Figure 4.9. Examples of melting and annealing profiles for OT3.5 in the presence of 10-fold excess of the complementsary C-rich strand. The reactions were performed in different concentrations of Na^+ and K^+ as indicated (10 mM black, 50 mM red, 100 mM green). The temperature was changed at a rate of 0.5 °C min^{-1}.

Table 4.5. Apparent T_m values for the melting and annealing observed with OT3.5 in the presence of the complementary oligonucleotide at different concentrations of Na^+ and K^+, determined at the rate of heating and cooling of 0.5°C min^{-1}.

[M^+] mM	Anneal lower [a] transition T_m (°C)	Anneal T_Q [b]	Anneal higher [c] transition T_m (°C)	Melt lower [a] transition T_m (°C)	Melt T_Q [b]	Melt higher [c] transition T_m (°C)
Na^+						
10	-	-	-	-	-	-
20	66.5 ± 0.1	-	-	66.6 ± 0.1	-	-
30	68.5 ± 0.2	-	-	68.2 ± 0.1	-	-
50	71.7 ± 0.1	-	-	71.4 ± 0.1	-	-
75	74.6 ± 0.1	-	-	74.4 ± 0.3	-	-
100	76.1 ± 0.1	-	-	76.0 ± 0.4	-	-
K^+						
10	-	-	-	-	-	-
20	67.0 ± 0.4	71.4 ± 0.1	73.5 ± 0.1	67.2 ± 0.6	-	-
30	67.6 ± 0.4	71.5 ± 0.1	73.5 ± 0.3	68.3 ± 0.1	-	-
50	70.3 ± 0.2	73.4 ± 0.2	76.8 ± 0.4	71.4 ± 0.4	78.5 ± 0.6	-
75	72.0 ± 0.3	74.8 ± 0.2	79.3 ± 0.1	74.0 ± 0.2	79.3 ± 1.2	-
100	73.4 ± 0.3	76.4 ± 0.2	82.1 ± 0.3	75.6 ± 0.2	80.5 ± 1.8	-

Values are the average ± stardard error from three independent experiments.
a – lower transition represents duplex to single strand.
b – T_Q values are the temperatures at which fluorescence is lowest and dF / dT = 0.
c – higher transition represents quadruplex to single strand.

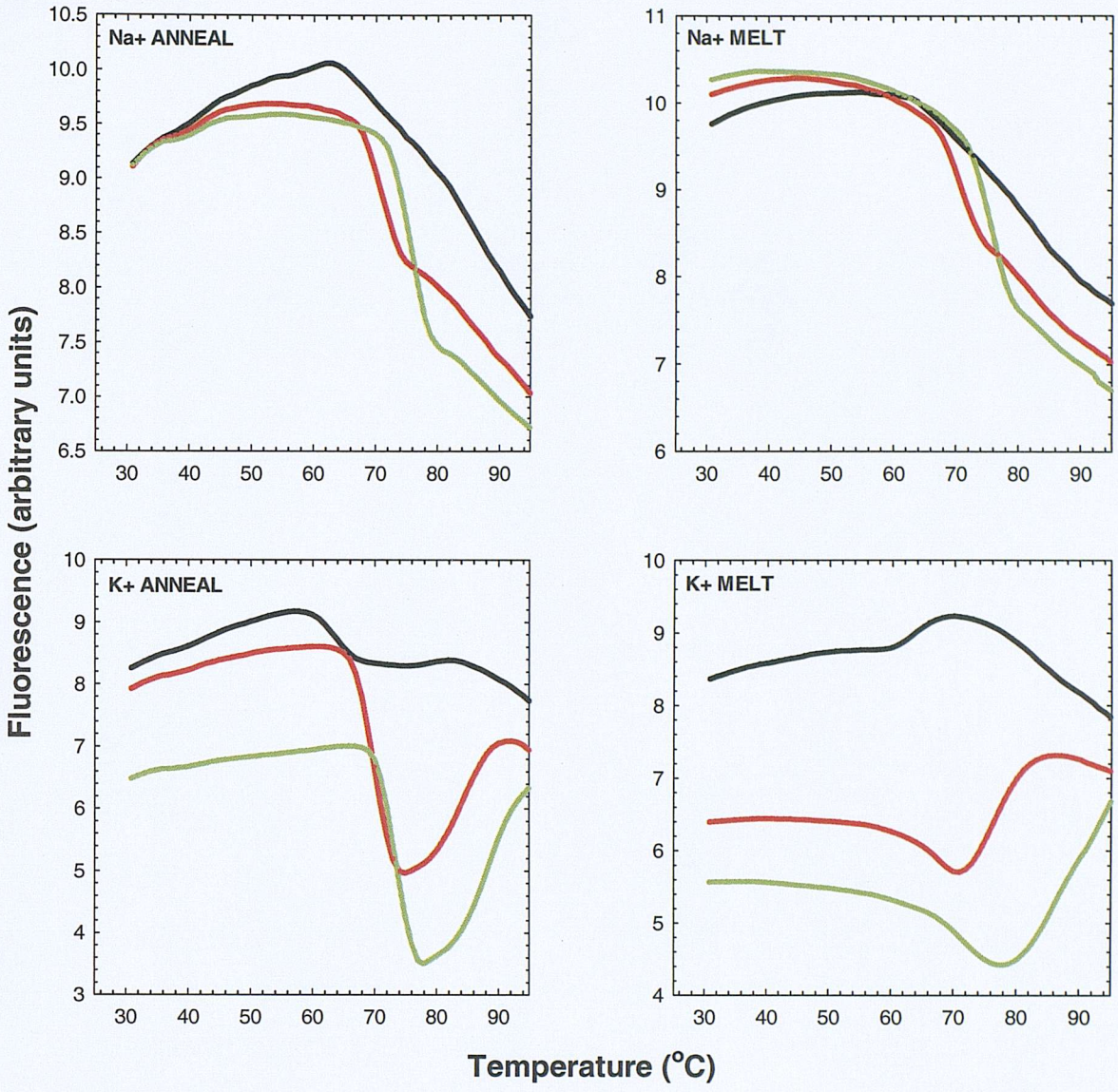


Figure 4.10. Examples of melting and annealing profiles for OT3.5 in the presence of 10-fold excess of the complementsary C-rich strand. The reactions were performed in different concentrations of Na^+ and K^+ as indicated (10 mM black, 50 mM red, 100 mM green). The temperature was changed at a rate of 0.05 °C min^{-1}.

Table 4.6. Apparent T_m values for the melting and annealing observed with OT3.5 in the presence of the complementary oligonucleotide at different concentrations of Na^+ and K^+, determined at the rate of heating and cooling of 0.05°C min^{-1}.

[M+] mM	Anneal lower [a] transition T_m (°C)	Anneal T_Q [b]	Anneal higher [c] transition T_m (°C)	Melt lower [a] transition T_m (°C)	Melt T_Q [b]	Melt higher [c] transition T_m (°C)
Na+						
10	-	-	-	-	-	-
20	66.2 ± 0.5	-	-	66.5 ± 0.9	-	-
30	68.2 ± 0.3	-	-	68.4 ± 0.5	-	-
50	71.0 ± 0.3	-	-	71.7 ± 0.6	-	-
75	74.0 ± 0.6	-	-	74.3 ± 0.4	-	-
100	76.2 ± 0.3	-	-	76.5 ± 0.7	-	-
K+						
10	58.4 ± 0.1	57.8 ± 0.5	62.6 ± 1.5	64.1 ± 0.7	69.5 ± 0.2	79.0 ± 0.6
20	60.9 ± 0.4	63.9 ± 0.3	68.7 ± 0.5	65.6 ± 0.1	72.3 ± 0.6	81.6 ± 0.3
30	64.0 ± 0.6	67.0 ± 0.3	71.7 ± 0.1	67.6 ± 0.3	73.2 ± 0.2	82.8 ± 0.6
50	67.8 ± 0.4	70.9 ± 0.3	76.3 ± 0.1	70.3 ± 0.4	75.2 ± 0.2	84.5 ± 0.8
75	70.5 ± 0.4	74.8 ± 0.3	82.0 ± 0.6	72.5 ± 0.5	77.2 ± 0.2	86.9 ± 0.4
100	72.7 ± 0.5	78.4 ± 0.6	>90	74.2 ± 0.5	78.9 ± 0.5	88.6 ± 0.2

Values are the average ± stardard error from three independent experiments.
a – lower transition represents duplex to single strand.
b – T_Q values are the temperatures at which fluorescence is lowest and dF / dT = 0.
c – higher transition represents quadruplex to single strand.

This section examines the effects of varying Na^+ and K^+ concentrations (10 to 100 mM) on melting and annealing of HT3.5. Initial experiments were again performed at the same fast rate as in the ligand studies. Figure 4.11 shows representative melting and annealing profiles determined under different cation concentrations. The T_m values obtained are summarised in Table 4.7 for all conditions. It can be seen that, in contrast to OT3.5 the melting and annealing profiles are all simple sigmoidal transitions. As expected this intramolecular quadruplex is more stable in K^+ than Na^+ and the T_m values increase with increasing ionic strength. There is very little difference between the annealing and melting profiles at the higher cation concentrations of both ions. This effect is more pronounced in the presence of K^+, and is similar to that observed by Risitano & Fox (2003) using a closely related oligonucleotide. It is clear that this intramolecular quadruplex has much faster dynamics than OT3.5.

Intermediate rate of heating and cooling (0.5 °C min^{-1})

Repeated melting and annealing cycles were performed on the same samples at the reduced rate of temperature change of 0.5°C min^{-1} and the results are presented in Figure 4.12. It can be seen that there is no hysteresis and the melting and annealing profiles are very similar. A small difference of unknown origin occurs between these profiles, wherein that the melting profiles appear to show a greater temperature dependent change in fluorescence signal at high and low temperatures on either side of the melting transition.

Since these curves show little hysteresis, we can assume that the samples are at thermodynamic equilibrium and therefore the data can be used to determine ΔH values, by van't Hoff analysis, assuming a simple two-state equilibrium. The results of this analysis are shown in Table 4.8. There is generally good agreement between the ΔH values for melting and annealing, though larger differences are apparent at the lower ionic conditions (especially Na^+). This discrepancy may be an artefact of the analysis, which is caused by the sloping baselines, as previously described. It can be seen that these ΔH values increase with ionic strength, as previously reported (Risitano & Fox, 2003). These effect will be considered further in the discussion.

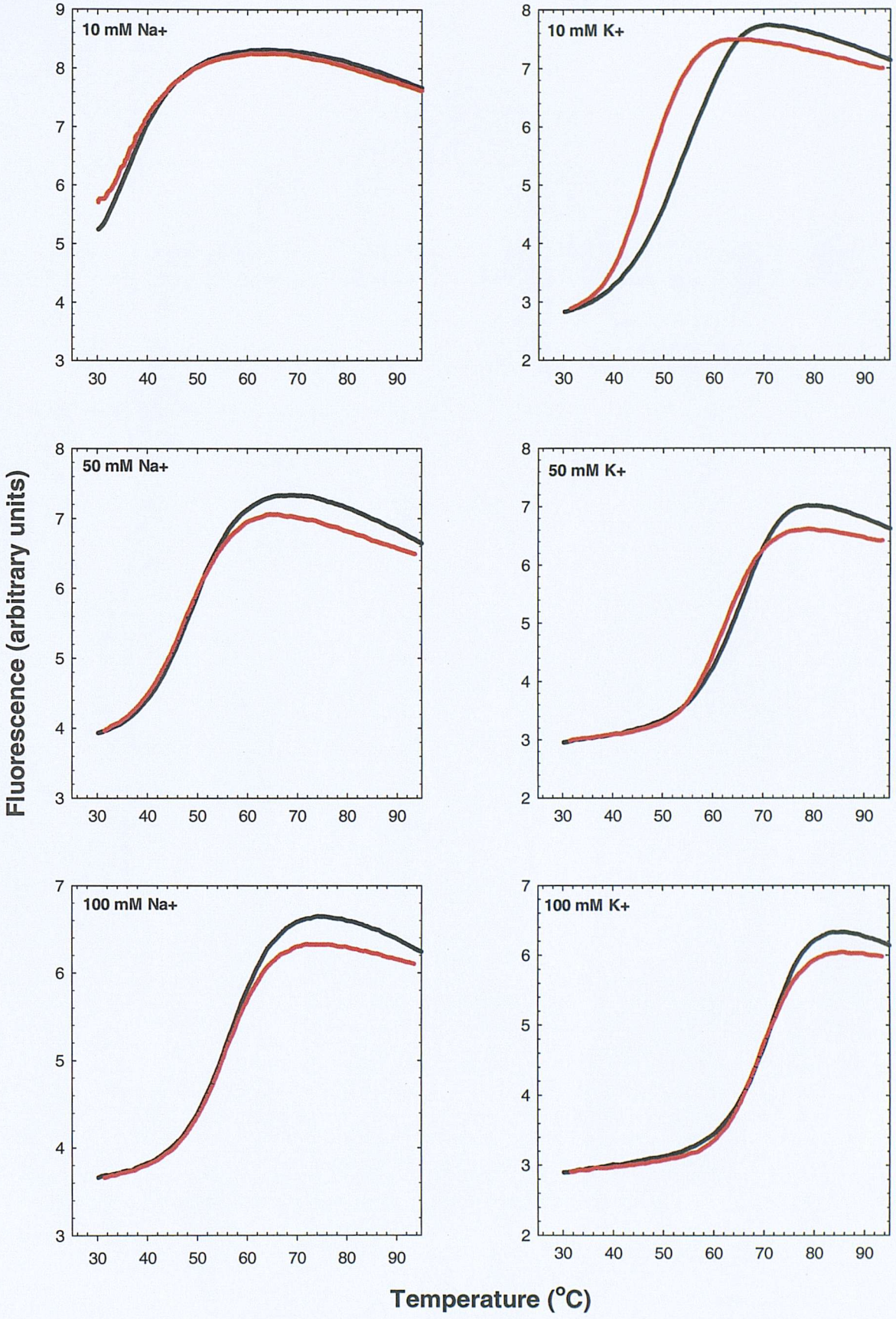


Figure 4.11. Examples of melting (black) and annealing (red) profiles for HT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at a rate of 6 °C min^{-1}.

Table 4.7. Apparent T_m values for HT3.5 determined from successive fast heating and cooling cycles at different concentrations of Na^+ and K^+, at the rate of heating and cooling of 6°C min^{-1}.

$[M^+]$ (mM)	T_m (°C)	T_m (°C)
Na^+		
10	<35[a]	36.7 ± 0.3
20	37.7 ± 0.6	40.2 ± 0.1
30	42.0 ± 0.2	43.8 ± 0.2
50	47.8 ± 0.5	48.7 ± 0.3
75	52.9 ± 0.6	53.1 ± 0.3
100	55.5 ± 0.2	56.9 ±0.1
K^+		
10	46.9 ± 0.3	54.0 ± 0.5
20	52.9 ± 0.5	59.6 ± 0.4
30	56.6 ±0.1	62.6 ± 0.1
50	62.1 ± 0.1	65.8 ± 0.2
75	66.7 ± 0.1	69.0 ± 0.2
100	69.1 ± 0.1	71.4 ± 0.3

Values are presented as the average ± standard errors from three independent experiments.
a – indicates that the T_m was too high (> 90°C) or too low (< 35°C) to be measured.

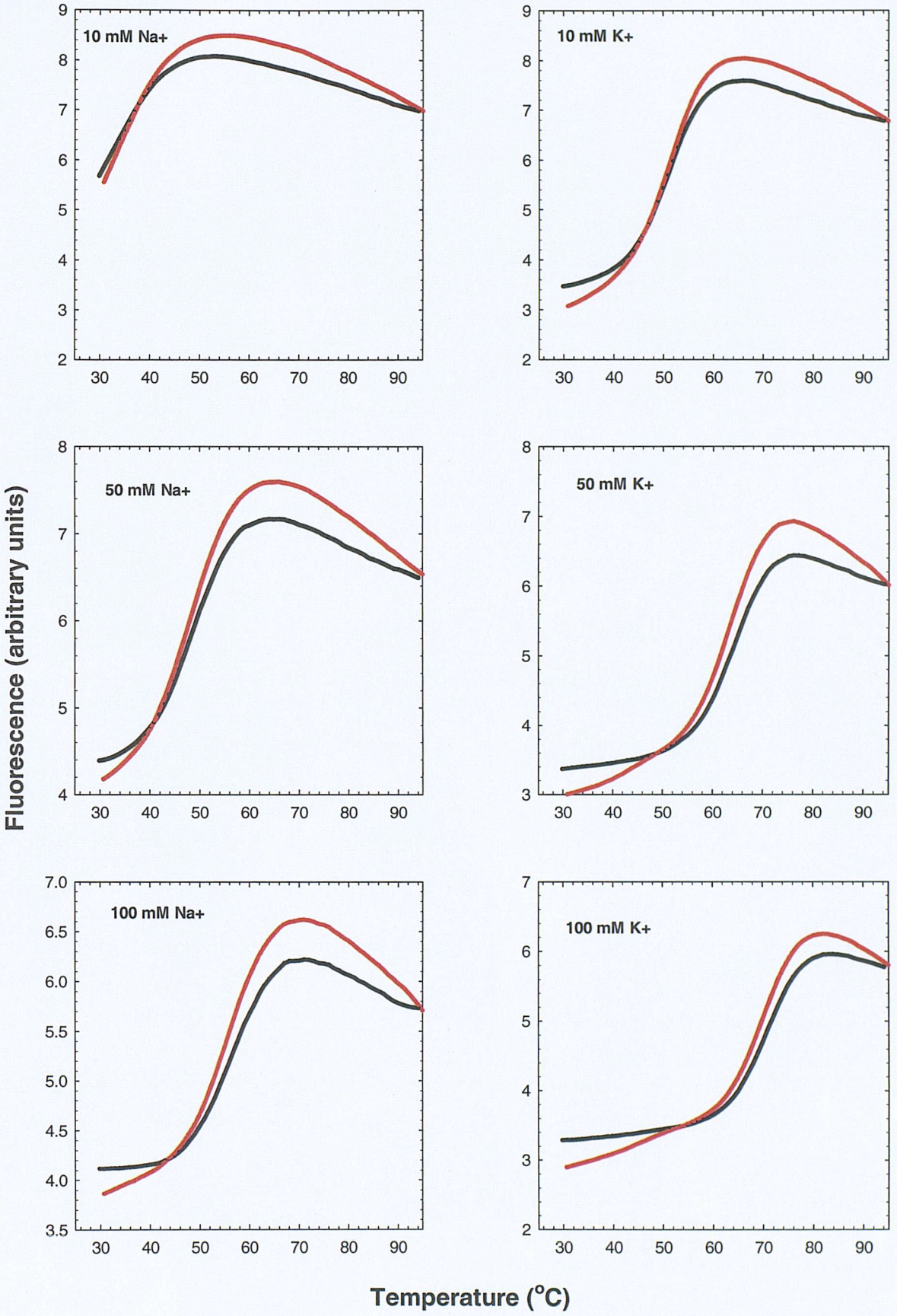


Figure 4.12. Examples of melting (black) and annealing (red) profiles for HT3.5 in the presence of different concentrations of Na^+ and K^+. The temperature was changed at a rate of 0.5 °C min^{-1}.

Table 4.8. Apparent T_m and ΔH values for HT3.5 estimated by van't Hoff analysis of the melting and annealing profiles at different concentrations of Na^+ and K^+, at a heating and cooling rate of 0.5°C min^{-1}.

$[M^+]$ (mM)	Melting profile		Annealing profile	
	T_m (°C)	ΔH_{app} (KJ.$mole^{-1}$)	T_m (°C)	ΔH_{app} (KJ.$mole^{-1}$)
Na^+				
10	too unstable [a]		too unstable [a]	
20	too unstable [a]		too unstable [a]	
30	41.8 ± 1.9	−123 ± 5	45.1 ± 0.4	−167 ± 11
50	49.4 ± 0.5	−137 ± 10	50.2 ± 0.4	−182 ± 4
75	54.9 ± 0.5	−161 ± 6	54.9 ± 0.5	−193 ± 5
100	58.6 ± 0.6	−170 ± 4	58.1 ± 0.5	−201 ± 7
K^+				
10	53.0 ± 0.2	−190 ± 11	52.3 ± 0.2	−228 ± 7
20	57.9 ± 0.5	−210 ± 5	57.3 ± 0.3	−237 ± 3
30	61.3 ± 0.5	−223 ± 10	60.7 ± 0.3	−241 ± 4
50	66.0 ± 0.5	−237 ± 12	65.3 ± 0.5	−247 ± 3
75	69.6 ± 0.4	−255 ± 21	69.3 ± 0.4	−251 ± 2
100	72.2 ± 0.9	−258 ± 13	71.9 ± 0.7	−254 ± 1

T_m and ΔH_{app} values are presented as the average ± standard errors from three independent experiments. *a* indicates that the T_m was too high (> 90°C) or too low (< 35°C) to be fitted to a van't Hoff equation..The values presented are apparent ΔH values.

HT3.5 plus complementary strand

The relative stability of the duplex and quadruplex forms of HT3.5 was examined by performing the melting curves in the presence of an excess of the complementary C-rich strand. These experiments were similar to those reported in the previous section for OT3.5. As before, the first melting curves are disregarded as this represents the pre-folded quadruplex in which it is clearly not at equilibrium with the added complementary strand. The results for experiments performed at the fast rate of heating and cooling (6°C min^{-1}) are shown in Figure 4.13, while the T_m values obtained are summarised in Table 4.9.

It can be seen that in the presence of Na^+, the profiles show a simple curve corresponding to the duplex to single strand transition. Under these conditions there is no hysteresis. These results are similar to those obtained with OT3.5 and show no evidence for quadruplex formation under these conditions. In the presence of K^+, complex melting profiles are obtained, similar to those produced by OT3.5 at slow rates of heating. These curves are explained by suggesting that a duplex is formed at low temperatures, when this melts the single strands fold into a quadruplex structure (low fluorescence), which subsequently melts at higher temperatures. The lower melting transition is similar to that seen with Na^+ and shows no hysteresis, as expected for a duplex to single strand transition. The higher transition shows a small amount of hysteresis (~ 4°C) at the lower K^+ concentrations. This hysteresis is no longer evident when the samples are heated more slowly (0.5°C min^{-1}) as shown in Figure 4.14 and Table 4.10.

4.3 Discussion

4.3.1 OT3.5 and HT3.5 kinetics

Many studies have shown that the kinetics of quadruplex formation are slow, and that quadruplex folding by $d(G_4T_4)$ *Oxytricha* telomeric repeats is much slower than that by $d(G_3T_2A)$ repeats of the human sequence (Raghuraman & Cech, 1990; Sen & Gilbert, 1990; Guo et al., 1993). This is confirmed by the present work where hysteresis is seen in the fluorescence melting profiles at faster rates of heating and cooling of OT3.5 and HT3.5. This particularly complicates the fluorescence melting studies with OT3.5,

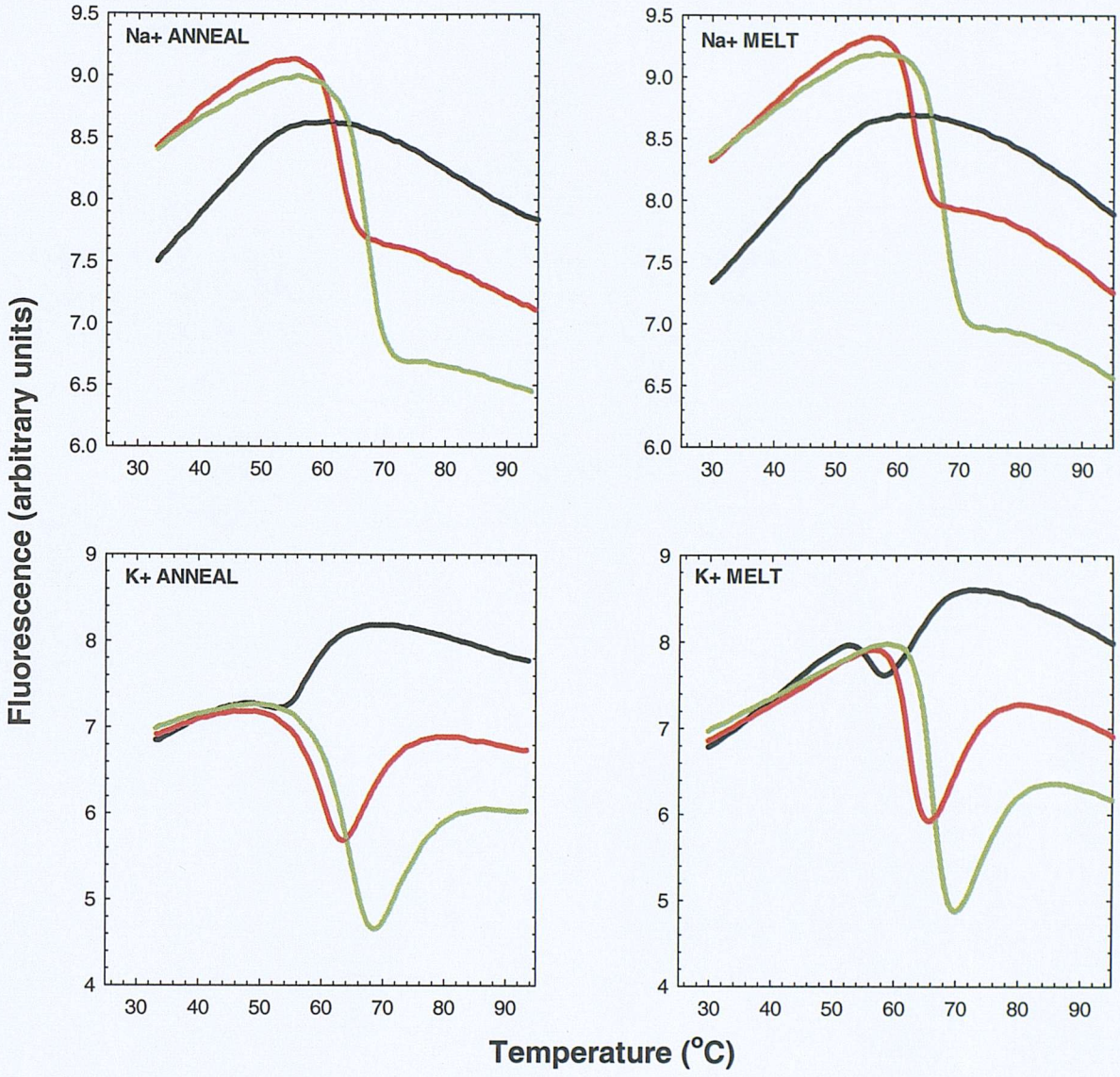


Figure 4.13. Examples of melting and annealing profiles for HT3.5 in the presence of 10-fold excess of the complementary C-rich strand. The reactions were performed in different concentrations of Na^+ and K^+, as indicated (black, 10 mM; red, 50 mM; gree, 100 mM). The temperature was changed at a rate of 6°C min^{-1}.

Table 4.9. Apparent T_m values for the melting and annealing observed with HT3.5 in the presence of the complementary oligonucleotide at different concentrations of Na^+ and K^+, determined at the rate of heating and cooling of 6°C min^{-1}.

	Anneal			Melt		
	lower [a] transition T_m (°C)	T_Q [b]	higher [c] transition T_m (°C)	lower [a] transition T_m (°C)	T_Q [b]	higher [c] transition T_m (°C)
Na^+						
10	-	-	-	-	-	-
20	56.3 ± 0.1	-	-	57.6 ± 0.1	-	-
30	58.8 ± 0.1	-	-	59.7 ± 0.1	-	-
50	62.4 ± 0.1	-	-	62.9 ± 0.1	-	-
75	65.3 ± 0.1	-	-	65.7 ± 0.1	-	-
100	67.3 ± 0.1	-	-	67.7 ± 0.1	-	-
K^+						
10	52.4 ± 1.6	50.3 ± 2.8	-	55.1 ± 0.4	56.8 ± 0.8	62.2 ± 0.3
20	53.0 ± 0.1	55.4 ± 0.1	59.4 ± 0.1	57.1 ± 0.1	59.7 ± 0.1	63.8 ± 0.1
30	56.3 ± 0.1	59.2 ± 0.1	62.6 ± 0.1	59.3 ± 0.1	62.2 ± 0.1	66.6 ± 0.1
50	60.1 ± 0.1	63.5 ± 0.1	66.5 ± 0.1	62.3 ± 0.1	65.6 ± 0.1	69.3 ± 0.1
75	63.0 ± 0.1	66.8 ± 0.1	70.3 ± 0.1	64.6 ± 0.1	68.4 ± 0.1	72.0 ± 0.1
100	64.6 ± 0.1	68.9 ± 0.1	72.4 ± 0.6	66.4 ± 0.1	70.2 ± 0.1	73.9 ± 0.2

Values are the average ± stardard error from three independent experiments.
a – lower transition represents duplex to single strand.
b – T_Q values are the temperatures at which fluorescence is lowest and dF / dT = 0.
c – higher transition represents quadruplex to single strand.

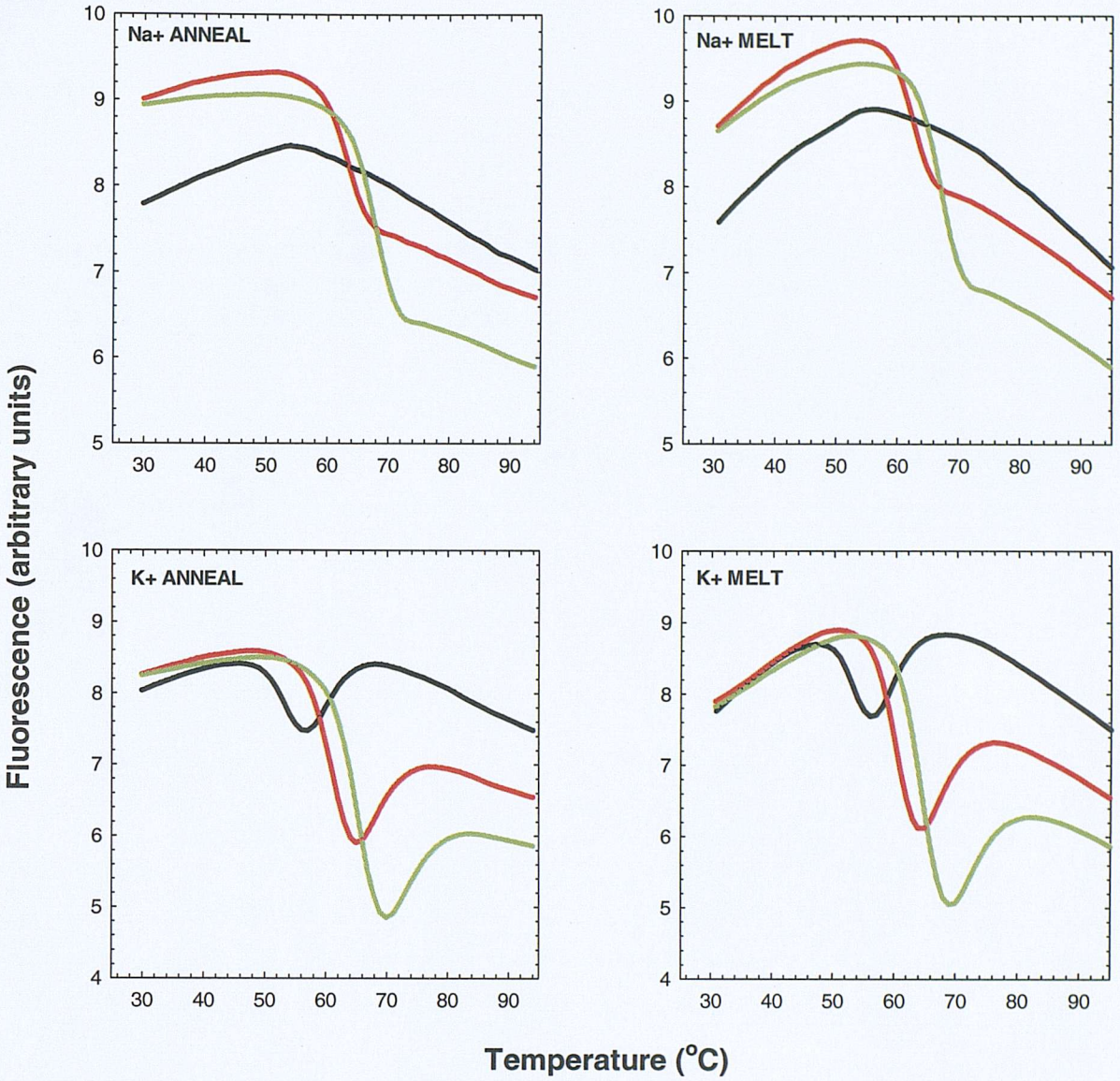


Figure 4.14. Examples of melting and annealing profiles for HT3.5 in the presence of 10-fold excess of the complementary C-rich strand. The reactions were performed in different concentrations of Na^+ and K^+, as indicated (black, 10 mM; red, 50 mM; gree, 100 mM). The temperature was changed at a rate of 0.5 °C min^{-1}.

Table 4.10. Apparent T_m values for the melting and annealing observed with HT3.5 in the presence of the complementary oligonucleotide at different concentrations Na^+ and K^+, determined at the rate of heating and cooling of 0.5°C min^{-1}.

	Anneal			Melt		
	lower [a] transition T_m (°C)	T_Q [b]	higher [c] transition T_m (°C)	lower [a] transition T_m (°C)	T_Q [b]	higher [c] transition T_m (°C)
Na+						
10	-	-	-	-	-	-
20	57.5 ± 0.7	-	-	56.7 ± 0.2	-	-
30	59.5 ± 0.3	-	-	59.2 ± 0.4	-	-
50	63.2 ± 0.2	-	-	62.6 ± 0.3	-	-
75	65.7 ± 0.2	-	-	65.2 ± 0.2	-	-
100	67.6 ± 0.3	-	-	67.3 ± 0.4	-	-
K^+						
10	50.7 ± 1.9	53.7 ± 2.2	57.8 ± 1.8	50.9 ± 1.5	54.9 ± 2.0	57.4 ± 1.6
20	54.0 ± 0.4	57.9 ± 0.5	61.3 ± 0.5	53.9 ± 0.3	57.4 ± 0.4	60.6 ± 0.3
30	57.2 ± 0.2	61.0 ± 0.3	64.3 ± 0.2	56.5 ± 0.4	60.3 ± 0.2	63.6 ± 0.4
50	60.7 ± 0.1	64.9 ± 0.2	67.9 ± 0.2	60.0 ± 0.2	64.2 ± 0.2	67.2 ± 0.1
75	63.4 ± 0.3	67.9 ± 0.3	71.1 ± 0.3	62.5 ± 0.2	67.2 ± 0.4	70.3 ± 0.1
100	65.2 ± 0.3	69.9 ± 0.1	72.9 ± 0.2	64.4 ± 0.3	69.3 ± 0.2	72.6 ± 0.2

Values are the average ± stardard error from three independent experiments.
a – lower transition represents duplex to single strand.
b – T_Q values are the temperatures at which fluorescence is lowest and dF / dT = 0.
c – higher transition represents quadruplex to single strand.

requiring **much** slower rates of temperature change and also a consideration of the scheme for melting and annealing the oligonucleotide. However, OT3.5 is clearly more stable than HT3.5 under all cation concentrations (particular marked in K^+), when compared at rates of heating and cooling where there is little or no hysteresis seen for both sequences (0.05°C min^{-1} and 0.5°C min^{-1} respectively). This can be seen by comparing the T_m values in Tables 4.3 and 4.8. This most likely reflects the extra stability provided by the additional guanine in each G-tract, which allows OT3.5 to form four G-tetrads, while the quadruplex formed by HT3.5 will only consist of three stacked G-tetrads.

4.3.2 Hysteresis

The extent of the hysteresis observed between melting and annealing (discussed in chapter 3) reflects how far quadruplex folding and/or unfolding are from thermodynamic equilibrium at a specific rate of heating and cooling. Since intramolecular quadruplexes are formed by both OT3.5 and HT3.5 (facilitating the folding process), it seems more likely that a slow rate of dissociation is responsible for this hysteresis.

Where the hysteresis is minimised, as for HT3.5 at an intermediate rate of temperature change, it is possible to derive thermodynamic parameters for a range of ionic conditions (see Table 4.8). However, this is not possible for any of the profiles obtained with OT3.5 even at a slow rate of heating and cooling, where hysteresis is still evident. However, the hysteresis does provide some interesting insights into quadruplex folding (as discussed later). A potential solution for overcoming this problem was proposed by Rougée *et al.* (1992) and allows for the individual association and dissociation rates, and their associated thermodynamic parameters to be determined from hysteresis curves. This however cannot be used in the present study with OT3.5, due to the complexity of its profiles, which show biphasic melting profiles and sloping baselines at either temperature extreme.

4.3.3 Biphasic melting curves

Heating and cooling using either scheme 1 or scheme 2 (Figure 4.1) produces different fluorescence melting profiles with OT3.5, under otherwise identical ionic conditions and rates of temperature change (0.5°C min^{-1} and 0.05°C min^{-1}). Two transitions are

sometimes observed in the **melting** profiles, which is particularly noticeable at higher ionic strength and in the presence of K^+. This suggests that the exact order of heating and cooling is important.

These biphasic melting curves presumably reflect the melting of two different quadruplex structures, such as a properly folded and a kinetically-trapped form. In scheme 1 the samples are first melted then annealed at a fast rate, followed by slower rates of heating and cooling. This may cause misannealing and/or incomplete melting of OT3.5 which leads to the formation of additional aberrant structures. Such effects should be minimised by following scheme 2, in which the samples are denatured before annealing and melting at the slowest rate of heating and cooling. We would expect that this scheme with a slower rate of annealing first should avoid the formation of kinetically-trapped structures, resulting from previous incorrect annealing at faster rates of heating and cooling. However, it is unclear why two transitions are still only seen in the melting and not the annealing profiles of scheme 2 as well as scheme 1, and why biphasic curves are even observed at very slow rates of heating and cooling of either scheme, when such kinetic constraints should be minimal. The importance of the heating and cooling scheme to the fluorescence melting curves of OT3.5, but not HT3.5, implies that these unusual profiles are related to the kinetics of melting and/or annealing.

An alternative possibility is that quadruplex melting is a two-step process, a suggestion which is consistent with the unusual behaviour observed in temperature-jump experiments (chapter 5). The (quenched) quadruplex structure may partially melt, resulting in an intermediate fluorescence signal; the intermediate structure then completely melts at higher temperature, producing a second transition as fluorophore and quencher are completely separated.

4.3.4 Monovalent cation interactions

Clearly all the T_m values for both OT3.5 and HT3.5 are higher in the presence of K^+ than Na^+, and increase with cation concentration, regardless of heating and cooling. Many studies have shown that quadruplexes are selectively stabilised by K^+, and this is confirmed by the present studies (Sen & Gilbert, 1992, Hud *et al.*, 1996). This preference for K^+ over Na^+ may reflect differences in its interactions with the G-

quartet, different energies of solvation of the two ions, or formation of different quadruplex topologies (Hud *et al.*, 1996; Sen & Gilbert, 1990; Muira *et al.*, 1995). Increasing the concentration of monovalent cation is known to confer increased stability to quadruplex-cation complexes.

Differences in the structure of the HT3.5 quadruplex are also suggested by the thermodynamic parameters which are more sensitive to K^+ than Na^+ concentration. The ΔH values produced for HT3.5 show a strong dependence on the ionic strength. This is an unusual observation since simple polyelectrolyte theory suggests that there should be no change in enthalpy with salt concentration. However, this has been observed in other studies with quadruplex and is consistent with the presence of specific ion binding sites in the folded structure (Jing *et al.*, 1997).

At all rates of temperature change, the hysteresis is greater in K^+ than Na^+, and increasing Na^+ and K^+ concentrations decrease hysteresis. As discussed in chapter 3 decreased hysteresis reflects a greater effect on association, suggesting the cations are increasing the rate of association rather than decreasing the rate of dissociation. The cations therefore appear to facilitate nucleobase alignment and promote nucleation, rather than by simply reducing the rate of dissociation (though this may also occur). This supports the idea that the hysteresis always observed reflects slow association rather than slow dissociation, which agrees with other reported k_a and k_d values for the *Oxytricha* repeat in both K^+ and Na^+ (Raghuraman & Cech, 1990).

4.3.5 Stability of quadruplex versus duplex

Although the ends of telomeres are single stranded, a large portion is double-stranded. Here, quadruplex formation will be competing with duplex, and the formation of quadruplex or duplex structures will depend on their relative stability. We have used these fluorescence melting studies to compare the stability of the Watson – Crick duplexes to the intramolecular quadruplex formed by OT3.5 and HT3.5.

This was achieved by following the fluorescence melting profiles of the fluorescently-labelled G-rich oligonucleotides in the presence of their C-rich complementary strands. (i) During **annealing**, as the strands are cooled the G-rich strand forms a quadruplex and the fluorescence is quenched; at a temperature at which the duplex form is stable,

the two strands will form a duplex which has higher fluorescence than both single-stranded or quadruplex forms. (ii) When **melting** the duplex, the G-rich strand can fold-back to form a quadruplex which melts at high temperatures.

One important consideration in analysing these fluorescence profiles is the influence of the relative kinetics of quadruplexes and duplexes, since quadruplexes are much slower kinetically (in addition to their relative stabilities). The rate of heating and cooling will therefore impose greater hysteresis on the profile at high temperature which closely resembles quadruplex melting and annealing, than at low temperature where the overall profile is dominated by duplex melting and annealing. Reducing or removing hysteresis completely will depress the annealing of quadruplex less significantly, so that a higher proportion of quadruplex will form. This may result in the kinetic trapping of a larger proportion of the fluorescent strand as quadruplex, preventing its binding to the complementary strand. This is indeed observed during annealing as lower fluorescence at low temperature which represents F_{ds}, and a more significant decrease in fluorescence at intermediate temperature which corresponds to F_q. These two effects are particularly noticeable for OT3.5 (which has slower kinetics than HT3.5), at lower M^+ concentrations (which increases hysteresis of quadruplex melting), and at higher rates of heating and cooling.

The presence of a quenched intermediate and the presence of two opposite transitions indicates duplex to quadruplex conversion. The duplex is more stable at low temperatures, with the quadruplex more stable at intermediate temperatures. A single transition from duplex (high fluorescence) to single strands (slightly lower fluorescence) indicates that the quadruplex is not formed. The difference between the two transitions indicates the relative stability of duplex and quadruplex forms.

Na^+ concentrations above 10 mM result in a single transition for both OT3.5 and HT3.5 which represent the duplex to single strand transition. Increasing concentrations of Na^+ as expected, increase the T_m of this transition. The absence of a quenched form even at Na^+ concentrations of 100 mM implies that Na^+ cannot support quadruplex formation in competition against the duplex form.

K^+ concentrations above 10 mM produce fluorescence melting profiles for OT3.5 and HT3.5 that contain two transitions, corresponding to duplex to quadruplex and quadruplex to single strand. The T_m values of both transitions increase with increasing K^+ concentration. This suggests that K^+ does facilitate quadruplex formation and that this is more stable than the duplex under some conditions.

The duplex to single strand transitions are less affected by hysteresis than the single strand to quadruplex transitions for both OT3.5 and HT3.5 (see Tables 4.5 to 4.7, 4.9 and 4.10). This may account for the presence of only a single transition for OT3.5 in K^+ at the fast rate of heating and cooling, though two transitions are seen at slower rates. As OT3.5 has been shown to have very slow kinetics, it is possible that fast melting and annealing were too rapid to allow significant quadruplex formation. It is clear that, even though duplex reactions are second order, duplex kinetics are faster than quadruplex kinetics.

In the context of the kinetic constraints, quadruplex formation by both sequences occurs in K^+ but not Na^+. Therefore, the profiles appear to be a mixture of quadruplex and duplex at physiological temperatures only in K^+, the physiologically relevant intracellular cation, and emphasise the superiority of K^+ over Na^+ at stabilising quadruplexes. This may be very significant to the importance of quadruplex *in vivo* where K^+ is more abundant than Na^+. These results therefore demonstrate that it is possible to form these higher order structures, even in presence of complementary strands. However, the situation in genomic DNA will be further biased toward duplex formation as the opposite strand will be held in close proximity; however, this also has the potential to fold into structures related to the i-motif (Simonsson *et al.*, 2000).

A possible explanation for the sometimes unusual melting profiles where the fluorescence at low temperatures is seen to be lower than that at high temperature is the formation of duplex structures composed of incorrectly aligned strands. This may arise because of the highly repetitive nature of the duplex sequence, which could allow the two strands to anneal out of register by one or two repeating units. This would produce 'stick ends' which could then anneal to another identical duplex, placing the fluorophores and quenchers of the G-rich strand from each duplex next to one another, therefore producing a quenched signal. This idea is simple to test, running the samples

from the melting experiments on a denaturing polyacrylamide gel will allow the stoichiometry and molecular weight of the DNA to be determined.

4.3.6 Non-specific fluorescence changes

The linear decrease in fluorescence well above the T_m, where the oligonucleotides are predominantly dissociated may reflect the increased kinetic energy of the random coil, increasing the frequency of collisions between the fluorophores and quenchers. The differences in the initial fluorescence after a cycle of melting and annealing by scheme 1 (see Figures 4.4 and 4.6) may reflect the removal of a fraction of the oligonucleotide into a thermally-stable form that cannot re-anneal after melting, photobleaching of the fluorophore, or the fluorophore or quencher becoming labile. Though this complicates the interpretation of the profiles, this factor does not affect the analysis of the remaining quadruplex-competent oligonucleotide that follows the normal re-annealing reaction.

4.3.7 Limitations

Although hysteresis is a constraint for understanding the process at non-equilibrium conditions, it does give an indication of the nature of the effect. As discussed earlier the effects on annealing reflect association, while changes to the melting are indicative of changes to dissociation. Thermodynamic data can only be derived these experiments if there is no hysteresis. The two-state model necessary for such analysis may be complicated by differing cation conditions that may induce the formation of different structures (Sen & Gilbert, 1990). These events may form by several steps and involve different fluorophore-quencher pair distances and/or orientation. Na^+ or K^+ can also affect the fluorescence signal. In one FRET study on the folding of a G-quartet oligonucleotide, Mergny and Maurizot (2001) noted enhanced acceptor emission in Na^+, and quenching of fluorescein emission without concomitant emission from the acceptor in K^+. These effects may reflect differences in the structures. If they do not, they make resulting thermodynamic parameters incorrect. These non-specific cation effects on fluorescence may affect how fluorescent molecules impact on quadruplex stability.

5 Fluorescence-Quenching Studies on the Kinetics of Quadruplex Folding using Temperature Jump Relaxation Experiments

5.1 Introduction

The use of fluorescently-labelled oligonucleotides to obtain quadruplex denaturation curves under varying ligand and cation concentrations has been outlined in earlier chapters. This chapter describes a modification of this technique to measure the kinetics of quadruplex folding in a technique that resembles temperature jump relaxation experiments.

5.1.1 Theory of temperature jumps

Temperature jump relaxation has been used on a vast array of chemical kinetic problems, and it is a useful tool for determining molecular motion, particular protein dynamics (Calender & Dyer, 2002). In the classical temperature jump technique the temperature of a chemical equilibrium is rapidly changed, thereby altering the equilibrium constant, and the concentrations of the different species. The rate at which the system relaxes to the new equilibrium is measured and from this the individual kinetic parameters are estimated. This approach was classically pioneered by Eigen & De Maeyer (1963). Temperature jump relaxation requires that the enthalpy of the reaction (ΔH) is not zero *i.e.* that the equilibrium constant is temperature dependent. This is explicit from the equation (Callender & Dyer, 2002):

$$\frac{\Delta K}{K} = 5.67 \times 10^{-3} \Delta H \, \Delta T \qquad 5.1$$

Where ΔK is the change to the K, the rate constant, ΔH is the enthalpy change and ΔT is the temperature jump.

The technique is usually employed for measuring very fast reaction kinetics in apparatus with a deadtime of a few microseconds. The approach requires a kinetic model to interpret the data, and to relate the observed relaxation times to the individual rate constants. Minimal models assume one or more chemical species which proceed to the new equilibrium position by a series of relaxation times, τ (*i.e.* the reciprocal of the

rate constant). In principle the number of exponentials (relaxation times) will be equal to the number of independent rate equations, corresponding to the number of steps in the reaction. In theory, it should be possible to characterise any reversible reaction, using a suitable method to differentiate between the structures.

For the simplest case of two interconverting species, a single relaxation time is observed which is equal to the sum of the forward and reverse rate constants. It should be noted that this rate constant for approaching the new equilibrium is greater than either of the individual first-order rate constants, k_1 and k_{-1} (Fersht, 1977). This is because the initial velocity in the reversible case is the same as for irreversible, but the reaction does not have to proceed so far:

$$A \longleftrightarrow B$$

$$1/\tau = k_1 + k_{-1} \qquad 5.2$$

For a two-step reversible reaction:

$$A \longleftrightarrow B \longleftrightarrow C$$

the analysis is much more complicated, and the two relaxation times are given by (Callender & Dyer, 2002):

$$1/\tau_1 = 0.5(p+q) \qquad 5.3$$

$$1/\tau_2 = 0.5(p-q) \qquad 5.4$$

$$p = k_1 + k_{-1} + k_2 + k_{-2} \qquad 5.5$$

$$q = \left[p^2 - 4\left(k_1 k_2 + k_{-1} k_{-2} k_1 + k_{-1} k_{-2}\right)\right]^{1/2} \qquad 5.6$$

This may be simplified if one relaxation time is significantly faster, such as in enzymic processes where enzyme – substrate (E – S) binding is usually fast and enzyme

conformational change is much slower. However, the analysis also works if the fast and slow steps are reversed. For a second order reaction the rates will be concentration-dependent, and the possibilities can be distinguished by increasing the substrate concentration: if the fast step is first, this causes $1/\tau_2$ to increase, while if it is second, then a decreased $1/\tau_2$ is observed.

The fluorescence changes in the thermal denaturation curves for HT3.5 and OT3.5 are attributable to unimolecular process(es), corresponding to the folding of the fluorescently-labelled oligonucleotides into intramolecular quadruplex structures. This is supported by the concentration independence of the T_m values. However, folding and unfolding may involve several reactions, for example, the sequential alignment of each guanine-tract with varying kinetics. It is therefore not clear whether any intermediates exist. If intermediates exist, it is conceivable that they have different fluorescence properties. These would then be evident in the kinetic pathway, producing several relaxation times. One problem for deconvoluting these relaxation times is that since each step in the series is unimolecular, it will not be possible to separate k_1 and k_{-1}. Indeed, more than two steps will produce an extremely complicated situation as each set of k_1 and k_{-1} will be related to previous ones.

In this chapter the LightCycler-based technology will be used to produce rapid changes in temperature, affecting the equilibrium between folded and unfolded quadruplex (as discussed above). The temperature changes and rates of reactions that can be measured in this way are clearly much slower than in conventional temperature jump kinetics. The fastest rate of temperature change is 20°C sec^{-1}, so a 5°C increase will take at least 250 msec. In addition the machine measures the fluorescence every 100 msec (if only one sample is inserted). It is therefore obvious that this will only be able to measure reaction times of 1 sec or longer. Previous studies have shown that the folding and unfolding of intermolecular quadruplexes is very slow. Since the melting data presented in chapter 4 showed considerable hysteresis it seems reasonable to suppose that the reaction kinetics are slow enough to be measured in this way.

5.1.2 Aims

The aim of this work was to develop a method for measuring the rate of these slow reactions in the LightCycler by a temperature jump experiment. The temperature of the

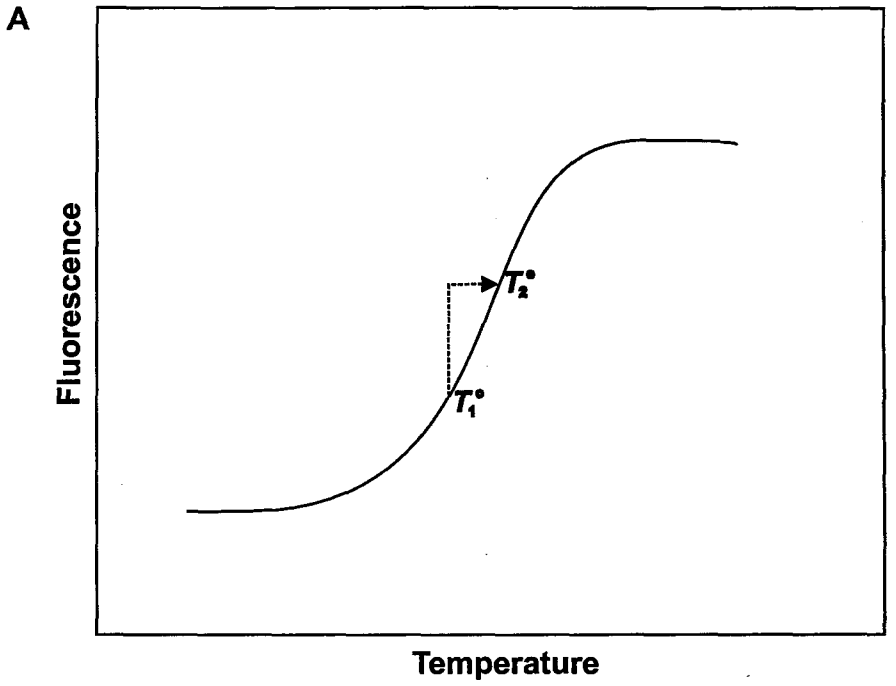
sample is rapidly increased by 5°C. This perturbs the equilibrium between quadruplex and random coil which is reported by monitoring the fluorescence as the system relaxes to a new equilibrium. Figure 5.1 illustrates an ideal temperature jump experiment. Experiments were performed at different temperatures to examine the kinetic parameters of both OT3.5 and HT3.5 under a range of ionic conditions, similar to those used in the previous chapters. The fluorescence values throughout the time course were then fitted to an exponential and/or biexponential curve (where appropriate) to determine the individual relaxation rate constant(s).

5.2 Results

5.2.1 OT3.5

The kinetics of the OT3.5 sequence were already known to be slow, as evidenced by the hysteresis between the melting and annealing curves (chapter 4). It was therefore hoped that its kinetic parameters would be accessible by this temperature jump technique. Initial experiments used 0.25 μM OT3.5 as this was used in melting studies. These experiments showed that for a temperature jump of 5°C, the time-dependent change in fluorescence was extremely slow and in many cases was only effectively complete after 1 h. In these early experiments the fluorescence was measured continuously throughout the time-course. However, a concern was that this long exposure to light might cause photobleaching, and indeed it was observed that the fluorescence decayed steadily over time. This was addressed by preparing the samples with degassed buffers to remove any dissolved O_2, and the samples were then maintained throughout the experiments under mineral oil. For this slow reaction, the fluorescence was also monitored every 30 sec rather than continuously.

The very slow kinetics of OT3.5 limited the number of temperature jumps that could be carried out on each sample, in order to avoid the formation of unusually-folded quadruplexes it was necessary to anneal the samples as slowly as possible. To achieve this, the samples were melted at 95°C then slowly re-annealed in a standard PCR machine which was programmed to decrease the temperature in 5°C steps over 20 h. However, because of the long time for equilibration it is not possible to fully denature and re-anneal the samples after each temperature jump, before performing a further experiment at a new temperature. This was in contrast to similar experiments, where triplexes yielded much faster kinetics ($t_{1/2} < 1$ min), wherein the samples could be



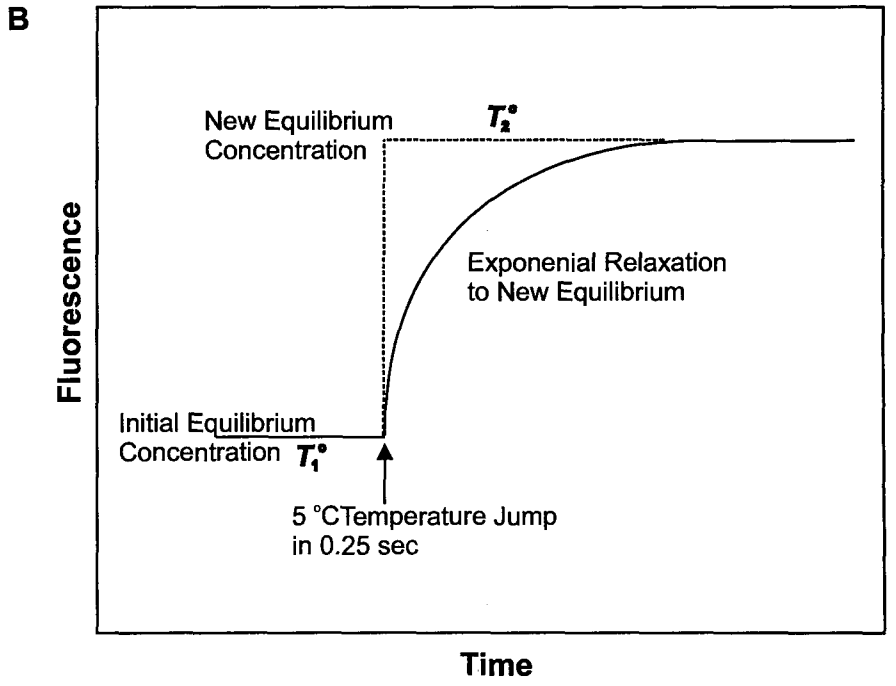


Figure 5.1. (A) .Representation of an ideal melting curve showing the difference in fluorescence between T_1° and T_2° (B) Representation of an ideal temperature jump.

completely denatured and re-annealed after each temperature jump (James *et al.*, 2003). Experiments in this chapter demonstrate temperature jumps that were performed in successive 5°C increments from 30°C to 90°C, and the final temperature after each jump was the starting condition for the next one.

Examples of time-dependent increases in fluorescence over successive temperature jumps from 30°C to 90°C (performed in 50 mM Na^+ and K^+) are shown in Figure 5.2. Similar experiments were performed in concentrations of Na^+ and K^+ between 10 mM and 100 mM. A few points are evident from inspection of these profiles. Firstly, the fluorescence changes appear to be in the wrong direction (decreases) at high temperatures (Na^+) and low temperatures (K^+). The reasons for this are not obvious, but may be related to the phenomenon of photobleaching as discussed above. As a result, useful kinetic information can only be obtained over a fairly narrow temperature range, in which the signal is sufficiently large, and in the correct direction. These temperatures are clustered around the T_m values, at which *dF/dT* is greatest. Secondly, it is apparent that the reactions are indeed slow and that time-dependent changes in fluorescence can clearly be seen in both Na^+ and K^+ containing buffers. Thirdly, K^+ which is known to produce more stable quadruplexes than Na^+, also produces much slower reactions than Na^+. Examples of fits to individual temperature jumps from these profiles (Figure 5.2) are shown in Figures 5.3 and 5.4. The relaxation times which could be successfully resolved and analysed are summarised in Table 5.1. Arrhenius plots for these data are shown in Figures 5.5 and 5.6. It is unclear whether K^+ results in slower association or dissociation, as both are components of the slower relaxation times.. These curves show no simple relationship between the reaction rate and temperature, suggesting that the reaction profile is more complex than anticipated.

In Na^+ containing buffers, relaxation times increase at low temperatures but decrease at high temperatures. The fastest rate of relaxation seems to be around the T_m of the reaction, as determined at the slowest rate of temperature change and scheme 2 (chapter 4). The results of the temperature jumps in K^+ are more confusing are not consistent. However, this may relate to the fact that the kinetics of the OT3.5 sequence are slow, and give rise to two transitions even at the slow rate of heating and cooling (0.05°C min^{-1}) by scheme 2 (again refer to chapter 4).

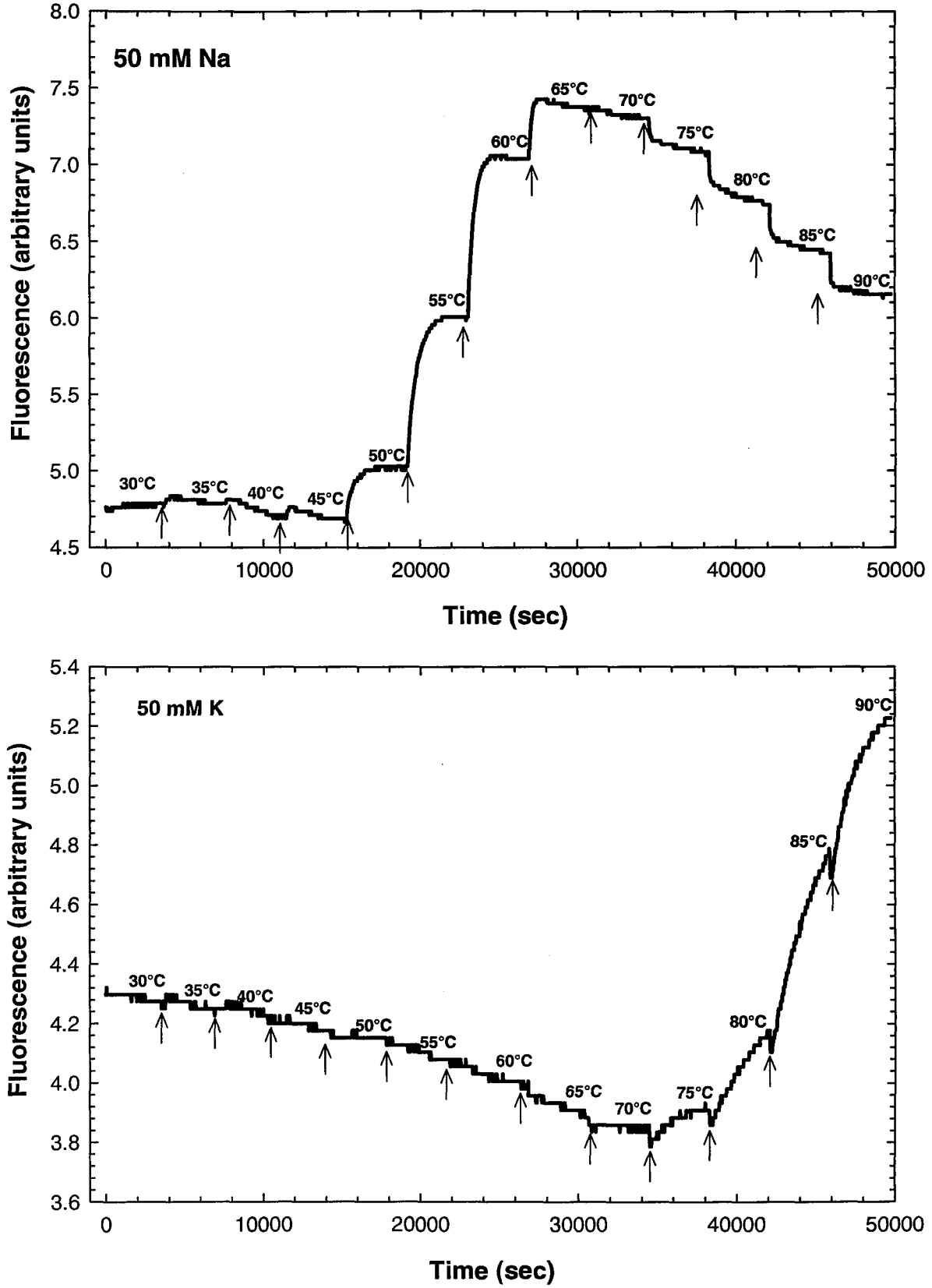


Figure 5.2. Examples of temperature jump traces obtained with OT3.5 in 10 mM NaP / 40mM NaCl (upper panel) and 10 mM KP / 40 mM KCl (lower panel). Arrows indicate the times at which the temperature was rapidly increased by 5°C.

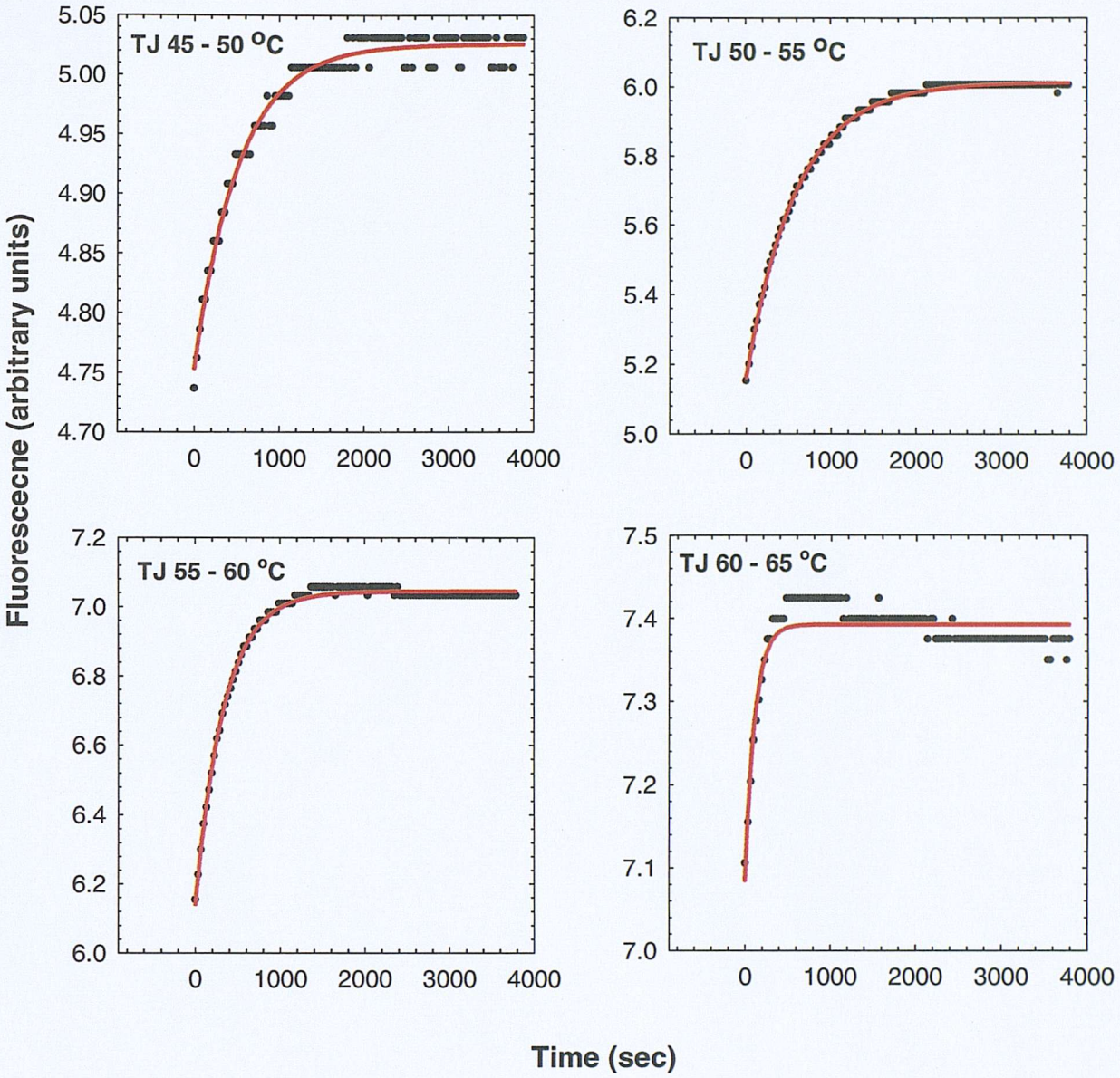


Figure 5.3. Examples of temperature jumps with OT3.5 in 10 mM NaP / 40 mM NaCl taken from the trace shown in Figure 5.2. Each profile is fitted by a single exponenial function. The starting and final temperatures are indicated in each panel.

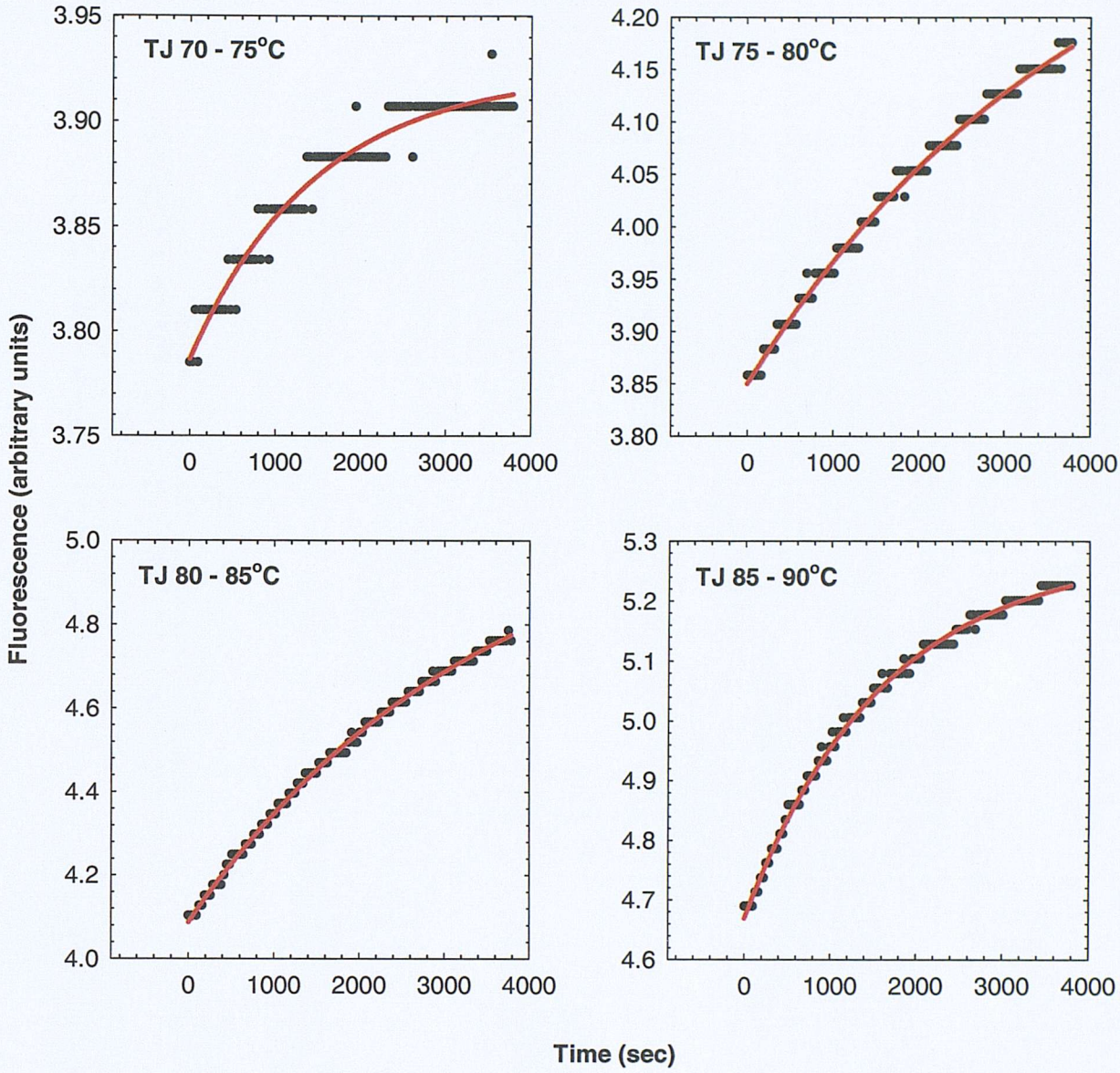


Figure 5.4. Examples of temperature jumps with OT3.5 in 10 mM KP / 40 mM KCl taken from the trace shown in Figure 5.2. Each profile is fitted by a single exponenial function. The starting and final temperatures are indicated in each panel.

Table 5.1. Apparent relaxation constants ($1/\tau$) for OT3.5 at different ionic strengths, derived from single exponential fits to the data. T_m values are shown from Table 4.3 (*scheme 2*, 0.05°C min^{-1}).

T_m*(°C)	42.2	45.0	47.5	51.9	55.7	59.3	65.6 79.7	67.7 81.2	70.8 82.3	84.3	86.2	> 90
TJ (°C)	10 mM Na^+ $1/\tau$ (s^{-1})	20 mM Na^+ $1/\tau$ (s^{-1})	30 mM Na^+ $1/\tau$ (s^{-1})	50 mM Na^+ $1/\tau$ (s^{-1})	75 mM Na^+ $1/\tau$ (s^{-1})	100 mM Na^+ $1/\tau$ (s^{-1})	10 mM K^+ $1/\tau$ (s^{-1})	20 mM K^+ $1/\tau$ (s^{-1})	30 mM K^+ $1/\tau$ (s^{-1})	50 mM K^+ $1/\tau$ (s^{-1})	75 mM K^+ $1/\tau$ (s^{-1})	100 mM K^+ $1/\tau$ (s^{-1})
30 – 35	1.1 ± 0.1 $x10^{-3}$	3.0 ± 1.4 $x10^{-3}$	1.6 ± 0.4 $x10^{-3}$									
35 – 40	4.7 ± 1.3 $x10^{-4}$	9.5 ± 0.1 $x10^{-4}$	5.5 ± 3.4 $x10^{-3}$									
40 – 45	5.7 ± 0.1 $x10^{-4}$	4.2 ± 0.3 $x10^{-4}$	0.6 ± 0.1 $x10^{-4}$	2.5 ± 0.5 $x10^{-3}$								
45 – 50	2.4 ± 1.1 $x10^{-3}$	6.8 ± 0.5 $x10^{-4}$	6.0 ± 0.5 $x10^{-4}$	8.2 ± 0.4 $x10^{-4}$	2.5 ± 0.6 $x10^{-3}$							
50 – 55			1.6 ± 0.2 $x10^{-3}$	1.3 ± 0.1 $x10^{-3}$	1.8 ± 0.1 $x10^{-3}$	4.0 ± 0.1 $x10^{-3}$						
55 – 60				3.6 ± 0.4 $x10^{-3}$	2.9 ± 0.1 $x10^{-3}$	3.9 ± 0.4 $x10^{-3}$	1.2 ± 0.3 $x10^{-4}$					
60 – 65					9.1 ± 0.1 $x10^{-3}$	7.2 ± 0.7 $x10^{-3}$	1.6 ± 0.2 $x10^{-4}$	1.6 ± 0.3 $x10^{-4}$				
65 – 70							3.3 ± 0.1 $x10^{-4}$	3.0 ± 0.1 $x10^{-4}$	3.2 ± 0.3 $x10^{-4}$			
70 – 75							3.6 ± 0.4 $x10^{-4}$	2.9 ± 1.0 $x10^{-4}$	4.4 ± 0.2 $x10^{-4}$	2.5 ± 1.1 $x10^{-3}$		
75 – 80							1.6 ± 0.3 $x10^{-4}$	2.2 ± 1.1 $x10^{-4}$	1.5 ± 0.2 $x10^{-4}$	2.7 ± 0.3 $x10^{-4}$	5.4 ± 1.2 $x10^{-4}$	4.2 ± 0.1 $x10^{-4}$
80 – 85							4.8 ± 0.2 $x10^{-4}$	3.7 ± 0.9 $x10^{-4}$	3.3 ± 0.2 $x10^{-4}$	2.3 ± 0.1 $x10^{-4}$	3.4 ± 0.1 $x10^{-4}$	6.3 ± 0.7 $x10^{-4}$
85 – 90								6.5 ± 0.1 $x10^{-4}$	5.7 ± 3.1 $x10^{-4}$	6.5 ± 0.2 $x10^{-4}$	7.4 ± 1.6 $x10^{-4}$	7.1 ± 0.4 $x10^{-4}$

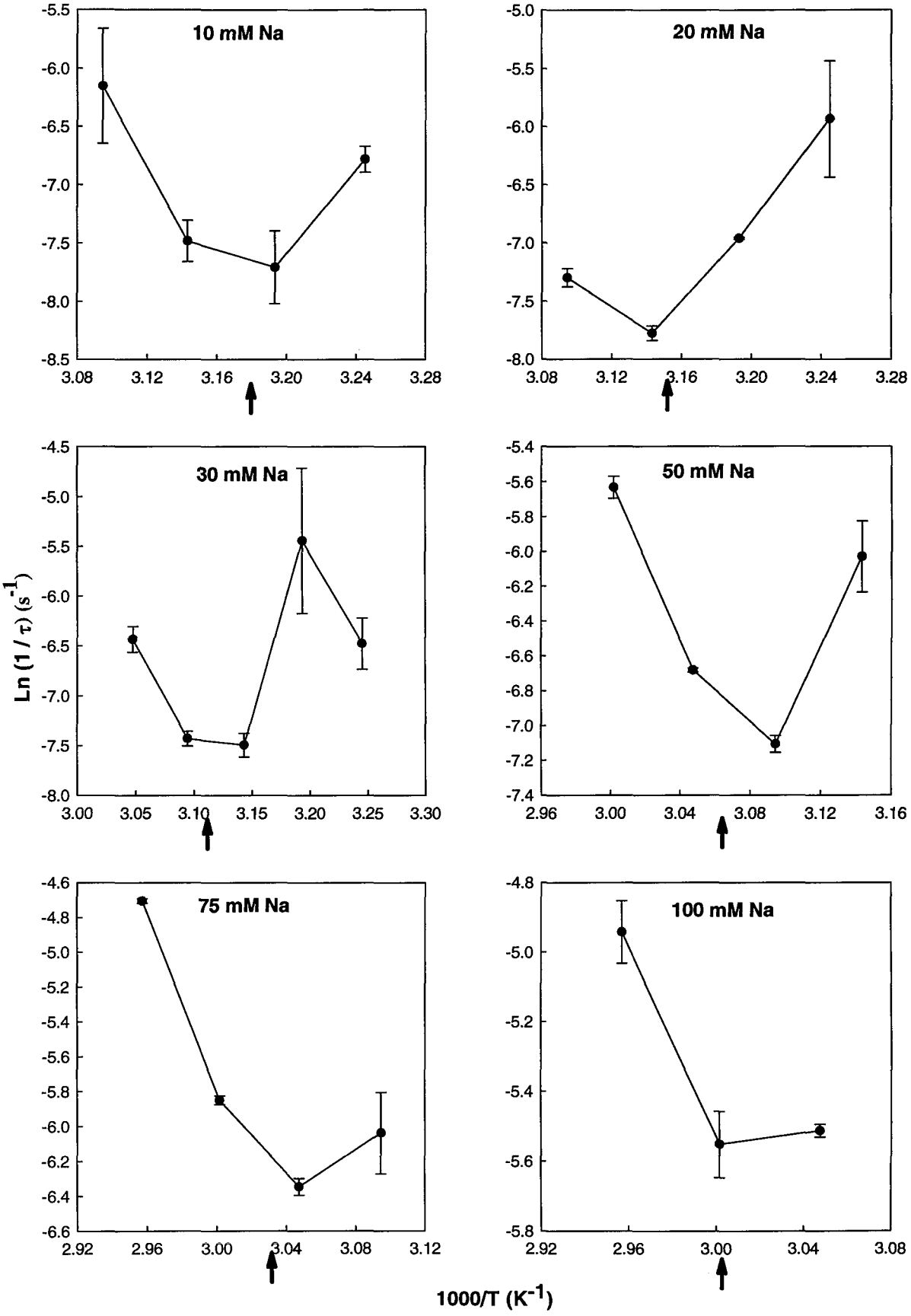


Figure 5.5. Arrhenius plots of relaxation times (1/τ) of OT3.5, derived from a single exponential function against temperature at different Na^+ concentrations. Arrows indicate T_m from Table 4.3.

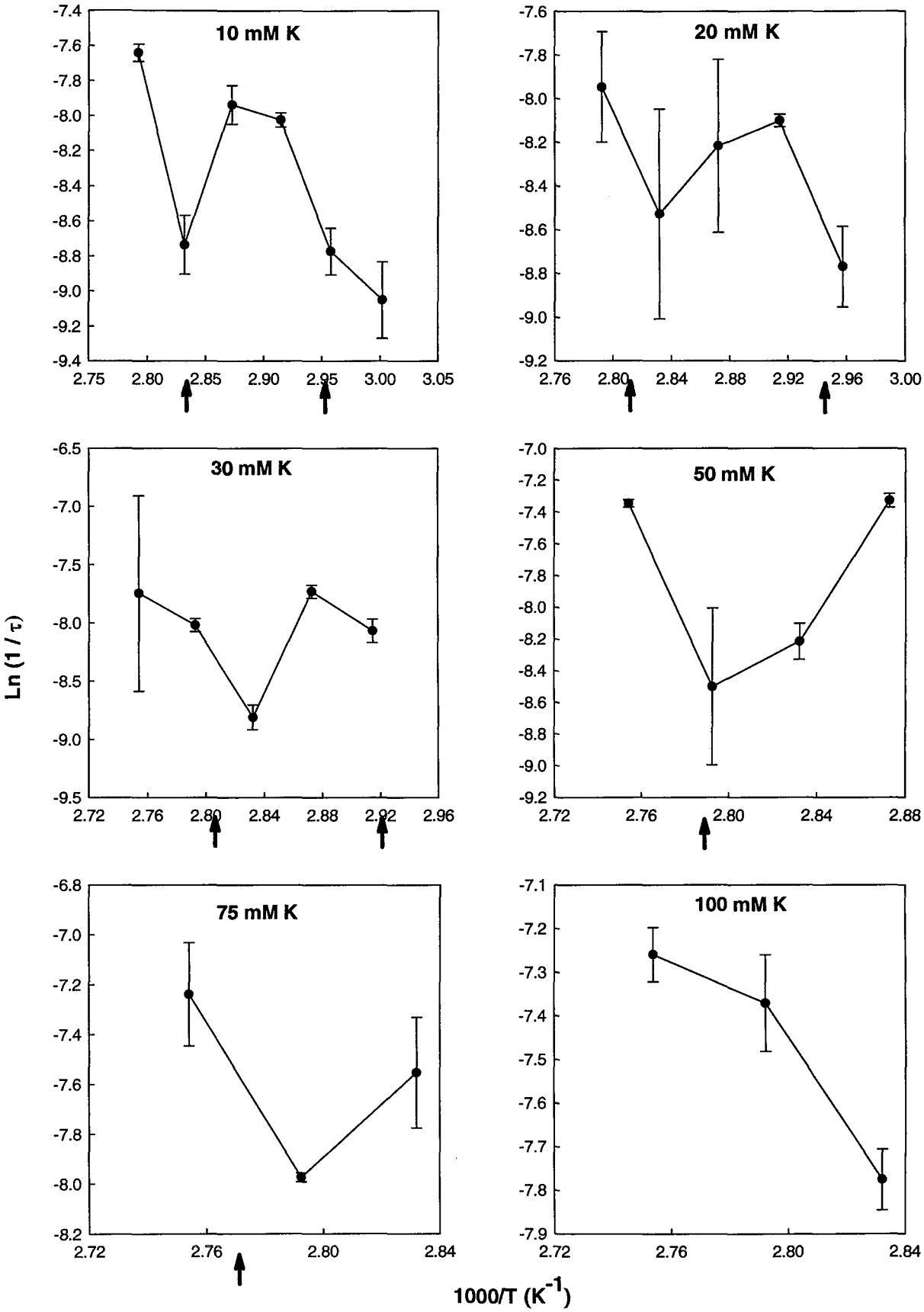


Figure 5.6. Arrhenius plots of relaxation times (1/τ) of OT3.5, derived from a single exponential function against temperature at different K^+ concentrations. Arrows indicate T_m from Table 4.3.

5.2.2 HT3.5

Similar temperature jump experiments were performed with HT3.5, though as a far greater amount of material was available, experiments were carried out using a 5-fold higher concentration of fluorescently-labelled oligonucleotide (1.25 μM) in order to increase both the signal and the signal-to-noise ratio. Even higher concentrations produced a fluorescence signal that was too great for the LightCycler. As HT3.5 has already been shown in chapter 4 to form an intramolecular complex, this change in concentration should not influence the kinetics of the reaction. This produced profiles that were far more amenable for kinetic studies, as the overall amplitude of the fluorescence change was increased, allowing a more accurate set of readings to be taken. The melting profiles in chapter 4 suggested that HT3.5 has faster kinetics than OT3.5, since far less hysteresis was observed. As a result of these faster kinetics, the samples could be readily melted then re-annealed to 30°C at 6°C min^{-1}, before the temperature jumps were performed. An equilibration period of 30 sec following each temperature jump was adequate for complete stabilisation of the fluorescence signal.

Figure 5.7 shows typical fluorescence traces over the course of a complete set of temperature jumps from 30°C to 90°C using OT3.5. This oligonucleotide produced unusual fluorescence changes at high temperatures, at which the fluorescence appears to decrease with time. Representative temperature jump profiles at various temperatures in Na^+ and K^+ are presented in Figures 5.8 and 5.9. Looking at the single exponential fits (left panels) it appears that there is often a systematic deviation between the experimental fluorescence values and the calculated best fit, and that the fluorescence increases have two components. As such double exponential fits were also carried out on these reaction profiles and are shown in the right hand panels. Relaxation time constants, derived from single exponential fits to these data are presented in Table 5.2, while those from the double exponential fits are given in Tables 5.3 and 5.4 for Na^+ and K^+ respectively. The kinetics of quadruplex reactions with HT3.5 are clearly faster by an order of magnitude over those with OT3.5, and again these are slower in K^+ than Na^+.

Single exponential fits to the recorded data are far more convincing at higher temperatures. A two component (double exponential) fit presents a problem at higher temperatures as the fast component becomes too rapid to measure, and therefore the

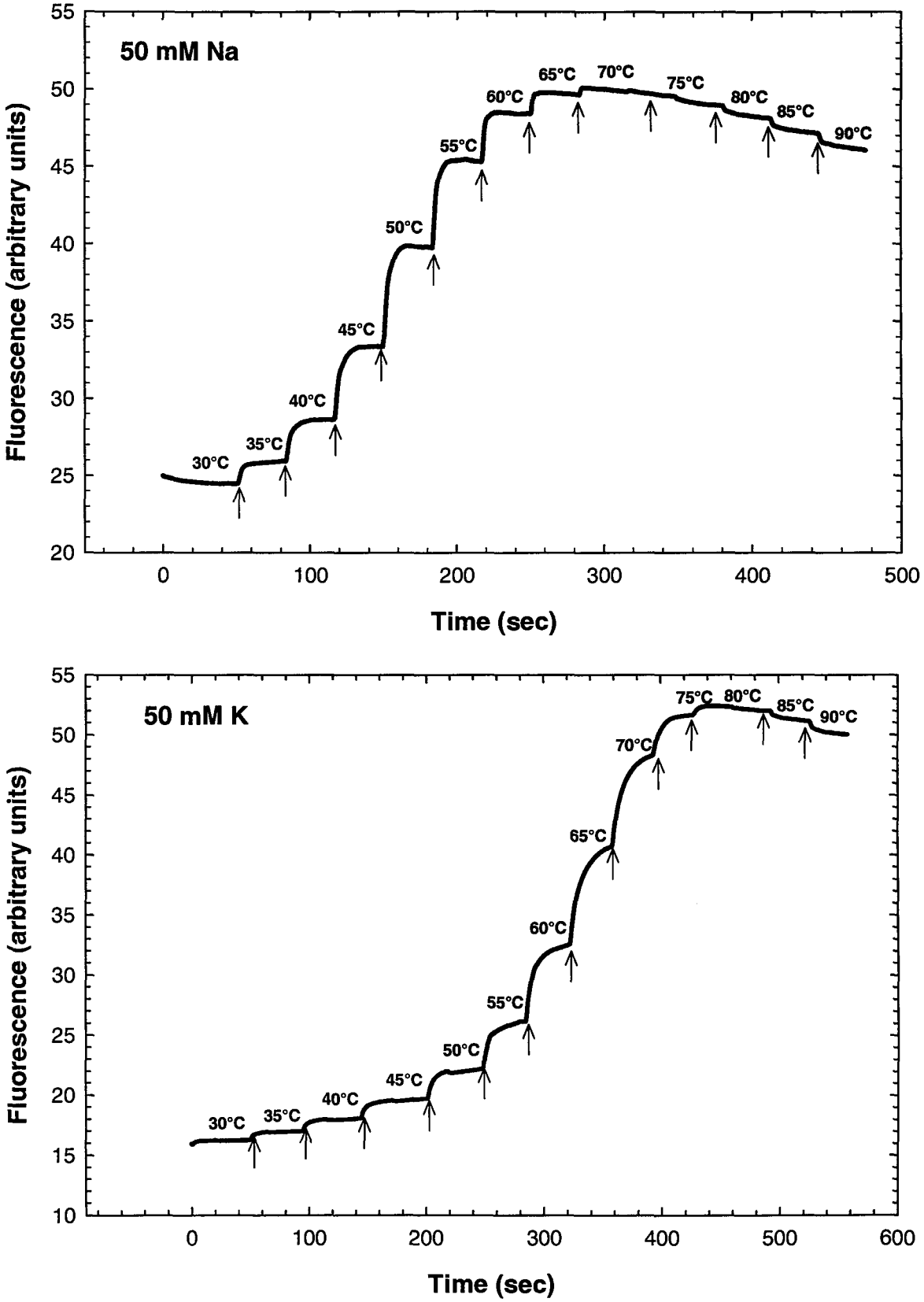


Figure 5.7. Examples of temperature jump traces obtained with HT3.5 in 10 mM NaP / 40mM NaCl (upper panel) and 10 mM KP / 40 mM KCl (lower panel). Arrows indicate the times at which the temperature was rapidly increased by 5°C.

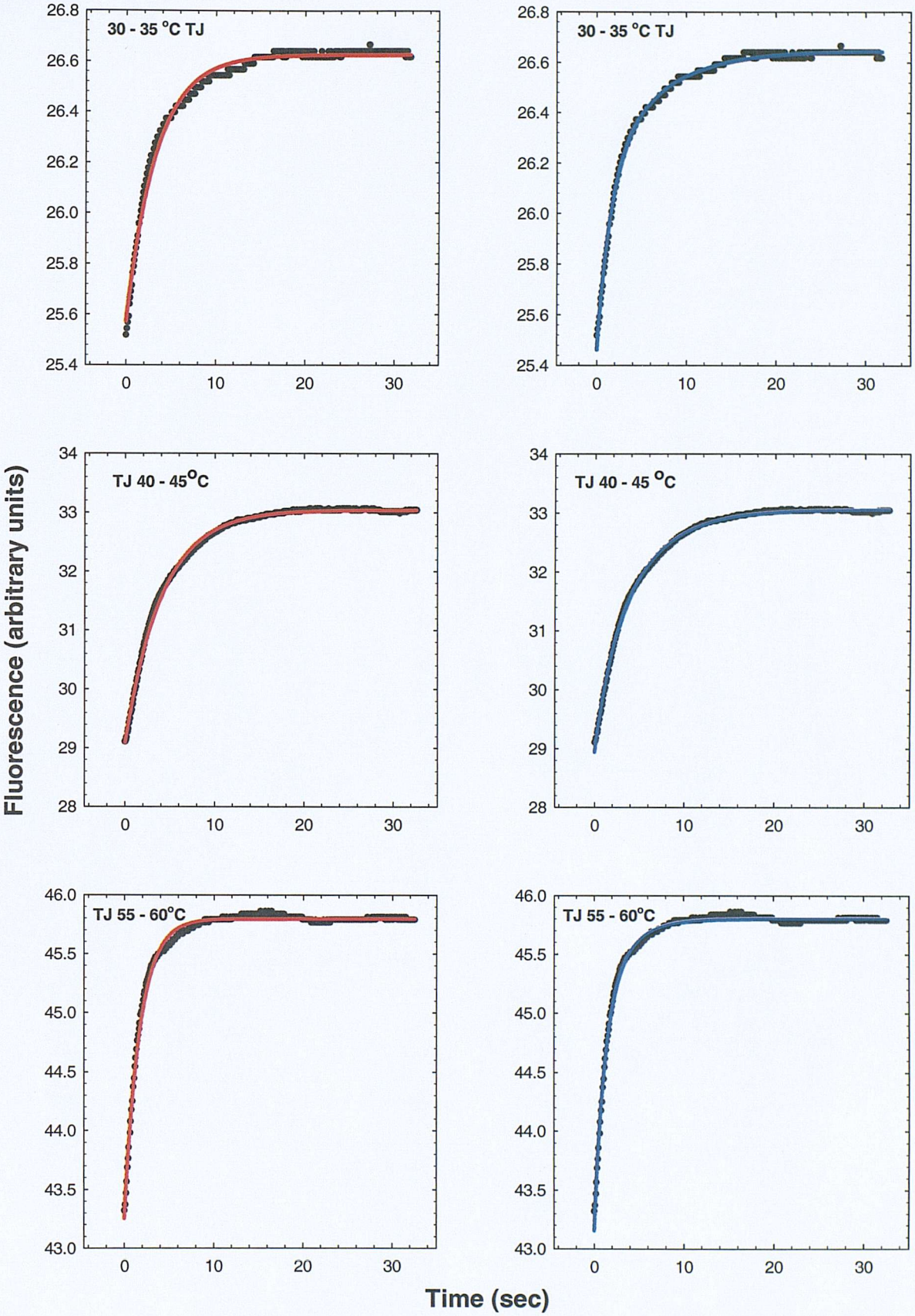


Figure 5.8. Examples of temperature jumps with HT3.5 in 10 mM NaP / 40 mM NaCl taken from the trace shown in Figure 5.5. Each profile is fitted by a single expontenial function on the left (red); and by a double exponential on the right (blue). The starting and final temperatures are indicated in each panel.

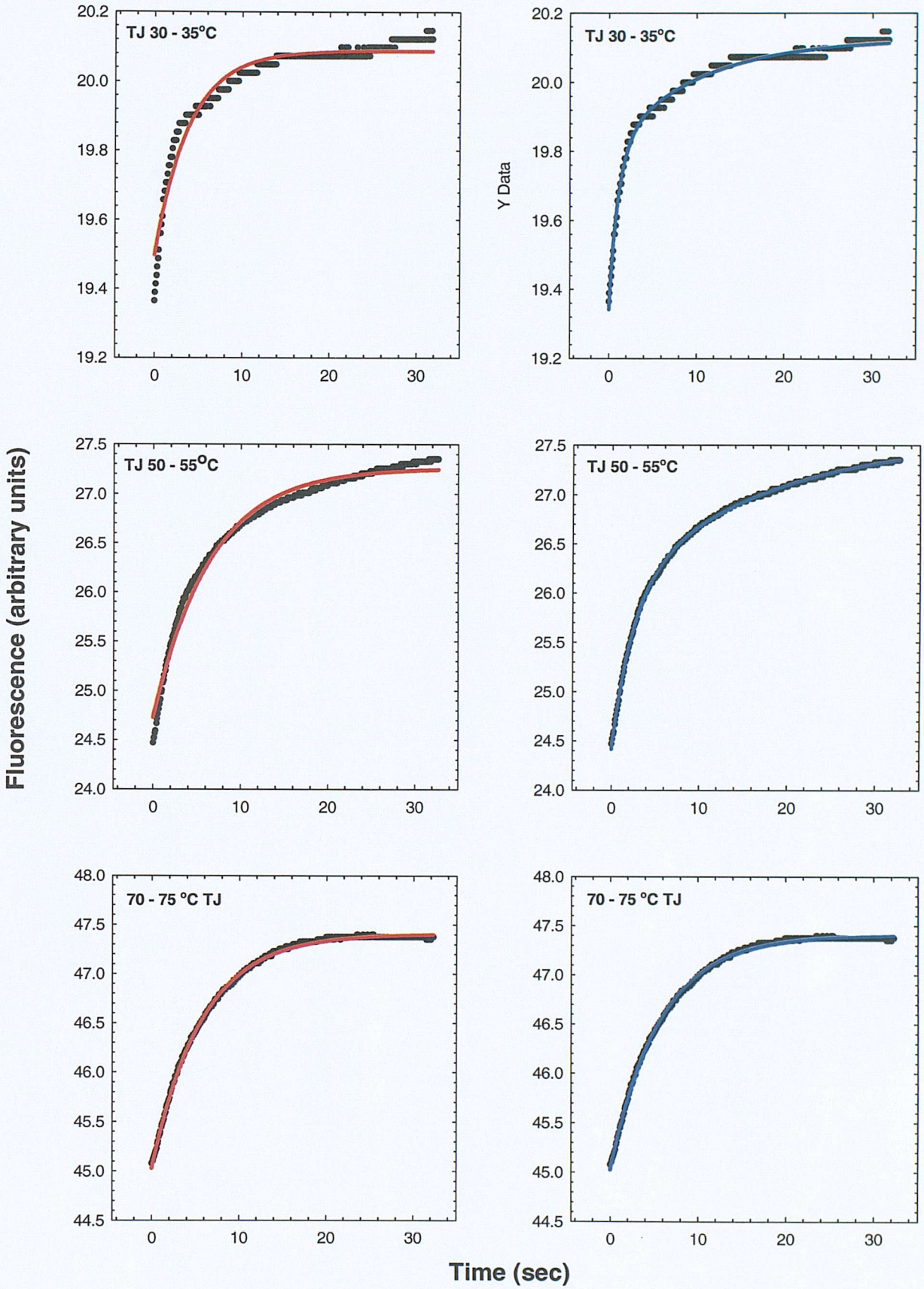


Figure 5.9. Examples of temperature jumps with HT3.5 in 10 mM KP / 40 mM KCl taken from the trace shown in Figure 5.5. Each profile is fitted by a single expontenial function on the left (red); and by a double exponential on the right (blue). The starting and final temperatures are indicated in each panel.

Table 5.2. Apparent relaxation constants ($1/\tau$) for HT3.5 at different ionic strengths, derived from single exponential fits to the data. T_m values are shown from Table 4.8 (0.5°C min^{-1}).

T_m*(°C)	< 35	< 35	45.1	50.2	54.9	58.1	52.3	57.3	60.7	65.3	69.3	71.9
TJ (°C)	10 mM Na^+ $1/\tau$ (s^{-1})	20 mM Na^+ $1/\tau$ (s^{-1})	30 mM Na^+ $1/\tau$ (s^{-1})	50 mM Na^+ $1/\tau$ (s^{-1})	75 mM Na^+ $1/\tau$ (s^{-1})	100 mM Na^+ $1/\tau$ (s^{-1})	10 mM K^+ $1/\tau$ (s^{-1})	20 mM K^+ $1/\tau$ (s^{-1})	30 mM K^+ $1/\tau$ (s^{-1})	50 mM K^+ $1/\tau$ (s^{-1})	75 mM K^+ $1/\tau$ (s^{-1})	100 mM K^+ $1/\tau$ (s^{-1})
30 – 35	1.8 ± 0.7 $x10^{-1}$	1.7 ± 0.1 $x10^{-1}$	2.2 ± 0.1 $x10^{-1}$	2.9 ± 0.4 $x10^{-1}$	3.1 ± 0.4 $x10^{-1}$	2.5 ± 0.4 $x10^{-1}$	1.9 ± 0.3 $x10^{-1}$	2.5 ± 0.3 $x10^{-1}$	2.8 ± 0.5 $x10^{-1}$	2.4 ± 0.3 $x10^{-1}$	2.5 ± 0.6 $x10^{-1}$	3.1 ± 0.6 $x10^{-1}$
35 – 40	2.0 ± 0.4 $x10^{-1}$	1.7 ± 0.1 $x10^{-1}$	1.9 ± 0.1 $x10^{-1}$	2.6 ± 0.1 $x10^{-1}$	2.7 ± 0.2 $x10^{-1}$	2.4 ± 0.1 $x10^{-1}$	2.0 ± 0.3 $x10^{-1}$	2.2 ± 0.3 $x10^{-1}$	2.1 ± 0.2 $x10^{-1}$	1.8 ± 0.1 $x10^{-1}$	1.6 ± 0.2 $x10^{-1}$	1.7 ± 0.3 $x10^{-1}$
40 – 45	2.8 ± 0.3 $x10^{-1}$	2.4 ± 0.1 $x10^{-1}$	2.3 ± 0.1 $x10^{-1}$	2.6 ± 0.1 $x10^{-1}$	3.0 ± 0.1 $x10^{-1}$	2.9 ± 0.3 $x10^{-1}$	1.6 ± 0.4 $x10^{-1}$	1.9 ± 0.2 $x10^{-1}$	2.0 ± 0.2 $x10^{-1}$	1.7 ± 0.1 $x10^{-1}$	1.1 ± 0.1 $x10^{-1}$	1.1 ± 0.2 $x10^{-1}$
45 – 50	4.5 ± 0.1 $x10^{-1}$	4.1 ± 0.8 $x10^{-1}$	3.1 ± 0.1 $x10^{-1}$	3.0 ± 0.1 $x10^{-1}$	3.1 ± 0.2 $x10^{-1}$	3.0 ± 0.3 $x10^{-1}$	1.1 ± 0.4 $x10^{-1}$	1.3 ± 0.2 $x10^{-1}$	1.7 ± 0.2 $x10^{-1}$	1.7 ± 0.2 $x10^{-1}$	1.1 ± 0.1 $x10^{-1}$	9.4 ± 1.7 $x10^{-2}$
50 – 55	6.5 ± 1.1 $x10^{-1}$	5.2 ± 0.1 $x10^{-1}$	4.6 ± 0.2 $x10^{-1}$	4.0 ± 0.1 $x10^{-1}$	3.8 ± 0.1 $x10^{-1}$	3.6 ± 0.2 $x10^{-1}$	9.3 ± 2.6 $x10^{-2}$	9.2 ± 1.3 $x10^{-2}$	1.2 ± 0.1 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$	1.2 ± 0.3 $x10^{-1}$	1.0 ± 0.2 $x10^{-1}$
55 – 60	1.1 ± 0.2 $x10^{-1}$	8.7 ± 0.4 $x10^{-1}$	7.1 ± 0.4 $x10^{-1}$	6.2 ± 0.3 $x10^{-1}$	5.5 ± 0.1 $x10^{-1}$	4.4 ± 0.6 $x10^{-1}$	8.0 ± 1.4 $x10^{-2}$	7.7 ± 0.5 $x10^{-2}$	9.5 ± 0.7 $x10^{-2}$	1.5 ± 0.1 $x10^{-1}$	1.7 ± 0.2 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$
60 – 65					7.9 ± 0.7 $x10^{-1}$	6.7 ± 0.4 $x10^{-1}$	8.5 ± 1.7 $x10^{-2}$	7.6 ± 0.7 $x10^{-2}$	8.2 ± 0.5 $x10^{-2}$	1.2 ± 0.1 $x10^{-1}$	1.5 ± 0.1 $x10^{-1}$	2.0 ± 0.1 $x10^{-1}$
65 – 70						1.0 ± 0.03	1.7 ± 0.5 $x10^{-1}$	1.1 ± 0.1 $x10^{-1}$	1.1 ± 0.1 $x10^{-1}$	1.2 ± 0.1 $x10^{-1}$	1.3 ± 0.2 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$
70 – 75									1.9 ± 0.3 $x10^{-1}$	1.9 ± 0.2 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$	1.8 ± 0.1 $x10^{-1}$
75 – 80											2.6 ± 0.2 $x10^{-1}$	2.6 ± 0.1 $x10^{-1}$
80 – 85												4.9 $x10^{-1}$

Table 5.3. Apparent relaxation constants ($1/\tau_1$ and $1/\tau_2$) for HT3.5 at different concentrations of Na^+. These values were obtained form double exponential fits to the data, where two constants could not be derived from the data is indicated by 'no fit'.

TJ (°C)	10 mM Na^+		20 mM Na^+		30 mM Na^+		50 mM Na^+		75 mM Na^+		100 mM Na^+	
	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})
30 – 35	1.4 ± 0.4 $x10^{-1}$	6.0 ± 0.3 $x10^{-1}$	1.5 ± 0.1 $x10^{-1}$	6.9 ± 0.9 $x10^{-1}$	1.5 ± 0.3 $x10^{-1}$	6.2 ± 0.9 $x10^{-1}$	1.4 ± 0.2 $x10^{-1}$	6.8 ± 0.8 $x10^{-1}$	1.0 ± 0.3 $x10^{-1}$	7.2 ± 1.6 $x10^{-1}$	7.8 ± 1.4 $x10^{-2}$	8.5 ± 2.3 $x10^{-1}$
35 – 40	1.8 ± 0.4 $x10^{-1}$	8.7 ± 2.0 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$	6.6 ± 1.4 $x10^{-1}$	1.6 ± 0.2 $x10^{-1}$	5.5 ± 1.3 $x10^{-1}$	1.7 ± 0.1 $x10^{-1}$	6.0 ± 0.2 $x10^{-1}$	1.5 ± 0.2 $x10^{-1}$	7.5 ± 1.3 $x10^{-1}$	1.3 ± 0.1 $x10^{-1}$	9.0 ± 0.9 $x10^{-1}$
40 – 45	1.9 ± 0.1 $x10^{-1}$	5.7 ± 0.2 $x10^{-1}$	2.1 ± 0.1 $x10^{-1}$	6.2 ± 1.2 $x10^{-1}$	1.9 ± 0.1 $x10^{-1}$	5.7 ± 0.2 $x10^{-1}$	1.6 ± 0.6 $x10^{-1}$	5.8 ± 1.4 $x10^{-1}$	2.0 ± 0.3 $x10^{-1}$	7.7 ± 0.7 $x10^{-1}$	1.5 ± 0.4 $x10^{-1}$	6.3 ± 0.5 $x10^{-1}$
45 – 50	2.8 ± 0.1 $x10^{-1}$	9.8 ± 3.0 $x10^{-1}$	2.4 ± 1.5 $x10^{-1}$	5.6 ± 1.6 $x10^{-1}$	2.2 ± 0.3 $x10^{-1}$	5.3 ± 0.6 $x10^{-1}$	2.0 ± 0.1 $x10^{-1}$	5.7 ± 1.5 $x10^{-1}$	2.2 ± 0.5 $x10^{-1}$	7.2 ± 1.5 $x10^{-1}$	1.7 ± 0.4 $x10^{-1}$	6.9 ± 1.1 $x10^{-1}$
50 – 55	no fit	no fit	4.0 ± 1.6 $x10^{-1}$	7.9 ± 0.8 $x10^{-1}$	2.4 ± 0.3 $x10^{-1}$	7.8 ± 0.7 $x10^{-1}$	2.0 ± 0.7 $x10^{-1}$	5.9 ± 0.7 $x10^{-1}$	2.4 ± 0.6 $x10^{-1}$	6.9 ± 1.0 $x10^{-1}$	1.8 ± 0.7 $x10^{-1}$	7.3 ± 1.5 $x10^{-1}$
55 – 60	no fit	no fit	no fit	no fit	1.6 ± 0.5 $x10^{-1}$	9.0 ± 0.6 $x10^{-1}$	2.9 ± 0.3 $x10^{-1}$	7.8 ± 0.7 $x10^{-1}$	2.7 ± 0.6 $x10^{-1}$	7.8 ± 0.3 $x10^{-1}$	2.1 ± 0.5 $x10^{-1}$	9.5 ± 2.2 $x10^{-1}$
60 – 65							no fit	no fit	2.9 ± 0.1 $x10^{-1}$	$1.0 \pm$ 0.05	2.8 ± 0.1 $x10^{-1}$	1.1 ± 0.2
65 – 70											3.4 ± 0.3 $x10^{-1}$	1.5 ± 0.1

Table 5.4. Apparent relaxation constants ($1/\tau_1$ and $1/\tau_2$) for HT3.5 at different concentrations of K^+. These values were obtained form double exponential fits to the data, where two constants could not be derived from the data is indicated by 'no fit'.

TJ (°C)	10 mM K^+		20 mM K^+		30 mM K^+		50 mM K^+		75 mM K^+		100 mM K^+	
	$1/\tau_1(s^{-1})$	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})	$1/\tau_1$ (s^{-1})	$1/\tau_2$ (s^{-1})
30 – 35	9.7 ± 2.6 $x10^{-2}$	5.3 ± 1.3 $x10^{-1}$	9.7 ± 2.6 $x10^{-2}$	6.3 ± 0.8 $x10^{-1}$	1.1 ± 0.1 $x10^{-1}$	8.2 ± 0.2 $x10^{-1}$	5.5 ± 0.7 $x10^{-2}$	6.9 ± 2.1 $x10^{-1}$	6.9 ± 1.6 $x10^{-2}$	9.0 ± 1.8 $x10^{-1}$	7.9 ± 0.7 $x10^{-2}$	9.6 ± 1.0 $x10^{-1}$
35 – 40	1.1 ± 0.4 $x10^{-1}$	4.8 ± 2.0 $x10^{-1}$	1.2 ± 0.4 $x10^{-1}$	5.7 ± 0.7 $x10^{-1}$	1.1 ± 0.3 $x10^{-1}$	7.1 ± 1.2 $x10^{-1}$	7.3 ± 1.3 $x10^{-2}$	7.0 ± 0.7 $x10^{-1}$	4.6 ± 2.4 $x10^{-2}$	7.0 ± 2.1 $x10^{-1}$	5.2 ± 0.6 $x10^{-2}$	8.1 ± 0.1 $x10^{-1}$
40 – 45	1.0 ± 0.5 $x10^{-1}$	3.7 ± 1.8 $x10^{-1}$	1.3 ± 0.5 $x10^{-1}$	4.8 ± 1.9 $x10^{-1}$	9.3 ± 1.1 $x10^{-2}$	5.0 ± 0.4 $x10^{-1}$	6.6 ± 0.2 $x10^{-2}$	5.7 ± 1.2 $x10^{-1}$	2.8 ± 1.8 $x10^{-2}$	7.0 ± 1.0 $x10^{-1}$	4.5 ± 0.6 $x10^{-2}$	8.5 ± 0.3 $x10^{-1}$
45 – 50	6.2 ± 2.9 $x10^{-2}$	2.9 ± 0.8 $x10^{-1}$	5.8 ± 3.4 $x10^{-2}$	2.7 ± 0.4 $x10^{-1}$	9.4 ± 3.7 $x10^{-2}$	4.0 ± 0.6 $x10^{-1}$	4.4 ± 2.2 $x10^{-2}$	4.3 ± 1.0 $x10^{-1}$	3.4 ± 1.8 $x10^{-2}$	6.3 ± 0.6 $x10^{-1}$	3.0 ± 0.6 $x10^{-2}$	6.8 ± 0.2 $x10^{-1}$
50 – 55	6.6 ± 1.1 $x10^{-2}$	3.8 ± 0.4 $x10^{-1}$	5.2 ± 1.5 $x10^{-2}$	2.9 ± 0.1 $x10^{-1}$	6.0 ± 0.9 $x10^{-2}$	3.0 ± 0.4 $x10^{-1}$	3.0 ± 0.1 $x10^{-2}$	3.1 ± 0.2 $x10^{-1}$	4.5 ± 1.2 $x10^{-2}$	5.4 ± 0.4 $x10^{-1}$	4.0 ± 0.5 $x10^{-2}$	6.5 ± 0.1 $x10^{-1}$
55 – 60	5.6 ± 0.9 $x10^{-2}$	3.8 ± 0.5 $x10^{-1}$	5.3 ± 0.7 $x10^{-2}$	3.8 ± 0.5 $x10^{-1}$	7.2 ± 1.5 $x10^{-2}$	4.4 ± 0.5 $x10^{-1}$	2.7 ± 0.1 $x10^{-2}$	2.2 ± 0.2 $x10^{-1}$	9.9 ± 2.4 $x10^{-2}$	5.0 ± 1.0 $x10^{-1}$	8.0 ± 0.5 $x10^{-2}$	5.7 ± 0.4 $x10^{-1}$
60 – 65	7.5 ± 1.6 $x10^{-1}$	6.3 ± 1.1 $x10^{-1}$	6.3 ± 0.2 $x10^{-2}$	5.2 ± 1.2 $x10^{-1}$	6.4 $x10^{-1}$	3.9 ± 0.3 $x10^{-1}$	6.2 ± 0.5 $x10^{-2}$	2.6 ± 0.2 $x10^{-1}$	1.1 ± 0.4 $x10^{-1}$	3.7 ± 0.6 $x10^{-1}$	1.6 ± 0.3 $x10^{-1}$	7.1 ± 2.4 $x10^{-1}$
65 – 70			1.0 ± 0.1 $x10^{-1}$	1.4 ± 0.9	9.4 ± 0.7 $x10^{-2}$	5.8 ± 1.5 $x10^{-1}$	8.6 ± 0.6 $x10^{-2}$	3.8 ± 0.8 $x10^{-1}$	1.0 ± 0.2 $x10^{-1}$	3.8 ± 0.8 $x10^{-1}$	1.2 ± 0.2 $x10^{-1}$	6.9 ± 2.7 $x10^{-1}$
70 – 75			no fit	no fit	1.5 $x10^{-1}$	9.7 $x10^{-1}$	1.5 ± 0.1 $x10^{-1}$	7.1 ± 0.2 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$	8.5 ± 0.2 $x10^{-1}$	1.6 ± 0.1 $x10^{-1}$	8.0 ± 2.2 $x10^{-1}$
75 – 80									no fit	no fit	2.2 ± 0.1 $x10^{-1}$	5.6 ± 0.8 $x10^{-1}$

data can be effectively described by a single exponential. In the case of Na^+, single exponential fits increase with increasing temperature, whereas the two component fits are less consistent and did not follow any trend. However, two fits give no clear pattern in K^+.

Arrhenius plots for the single and double fits in the presence of Na^+ are presented in Figures 5.10 and 5.11, and for K^+ in Figures 5.12 and 5.13. Looking at the data in the presence of Na^+, it appears that, the single exponential give values which follow the usual trend, *i.e.* the reaction becomes faster at higher temperatures. Least squares lines fitted to these data show good fits at low ionic strengths. However, these plots show significant curvature at higher concentrations of Na^+ and a simple Arrhenius equation does not adequately describe the data. The two exponential fit in Na^+ produces relaxation times that far less consistent with temperature.

In K^+ conditions, the plots of the values derived from single and double exponential equations are not linear, but do show a definite pattern of increasing rates of relaxation at temperatures below the T_m, and decreasing rates of relaxation above the T_m (determined in chapter 4). Lines of best fit have not been drawn in cases where the data does not seem linear.

As expected the amplitudes were greatest at the T_m, where dF/dT is maximal, and become smaller at higher and lower temperatures. However, though it is unclear why, it does seem that there is a definite pattern to the relative amplitudes of either kinetic component. The slower kinetic component seems to increase with temperature before decreasing again above the T_m (see Figures 5.14 and 5.15).

The time-dependent fluorescence changes with HT3.5 on the **rapid decrease** of temperature by 5°C were performed (see Figure 5.16). These proved unreliable due to the limitations imposed by cooling rather than heating samples. The machine is more effective at rapidly re-adjusting to higher temperatures than to lower temperatures. These also showed two kinetic components similar to the temperature jumps. Interestingly, though a time-dependent decrease in fluorescence should be associated with a rapid temperature decrease, again the fluorescence changes appear to be in the wrong direction at high temperatures.

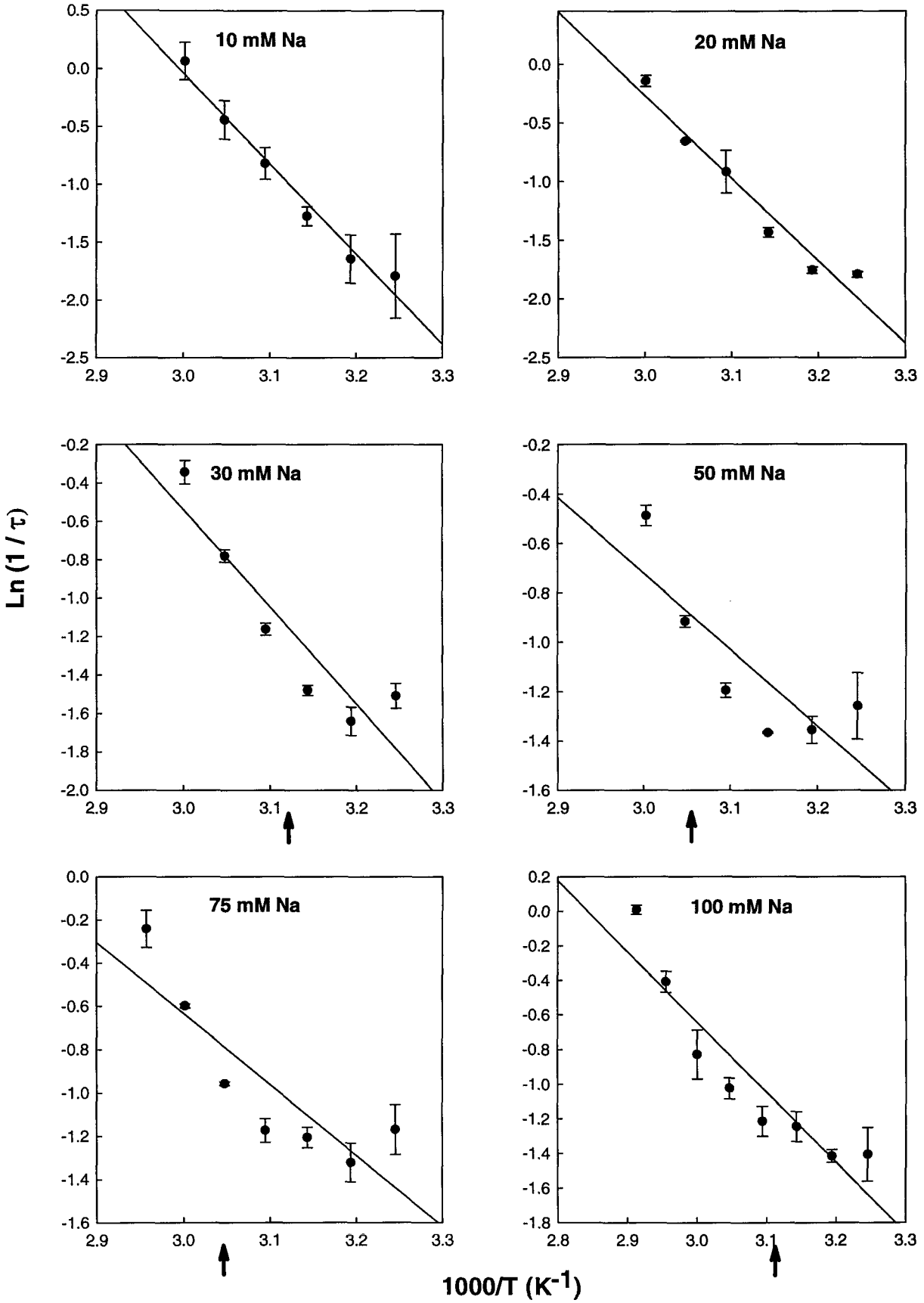


Figure 5.10. Arrhenius plots of relaxation times ($1/\tau$) of HT3.5, derived from a single expontential function against temperature at different Na^+ concentrations. Arrows indicate T_m from Table 4.8.

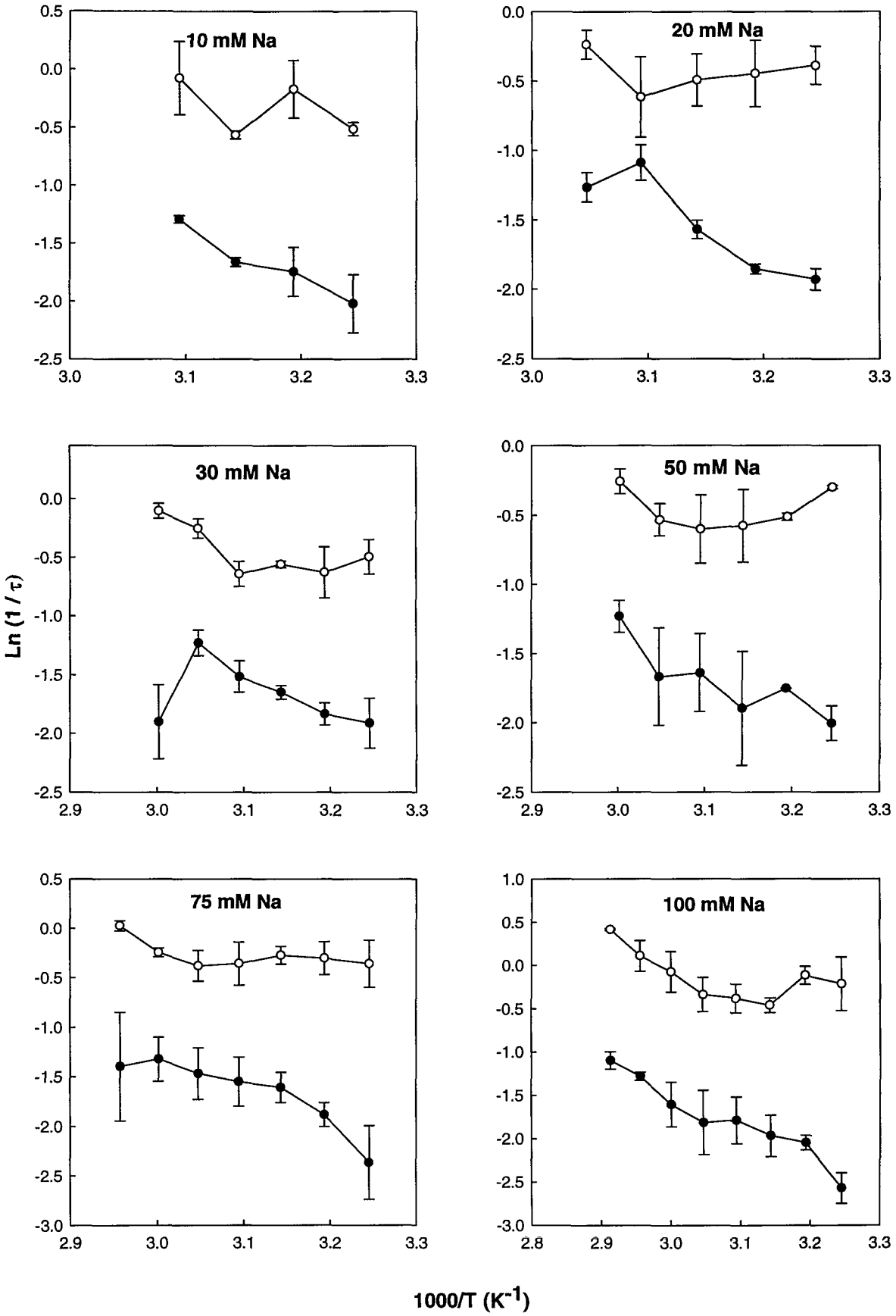


Figure 5.11. Arrhenius plots of relaxation times ($1/\tau_1$ and $1/\tau_2$) of HT3.5, derived from a single exponential function against temperature at different Na^+ concentrations.

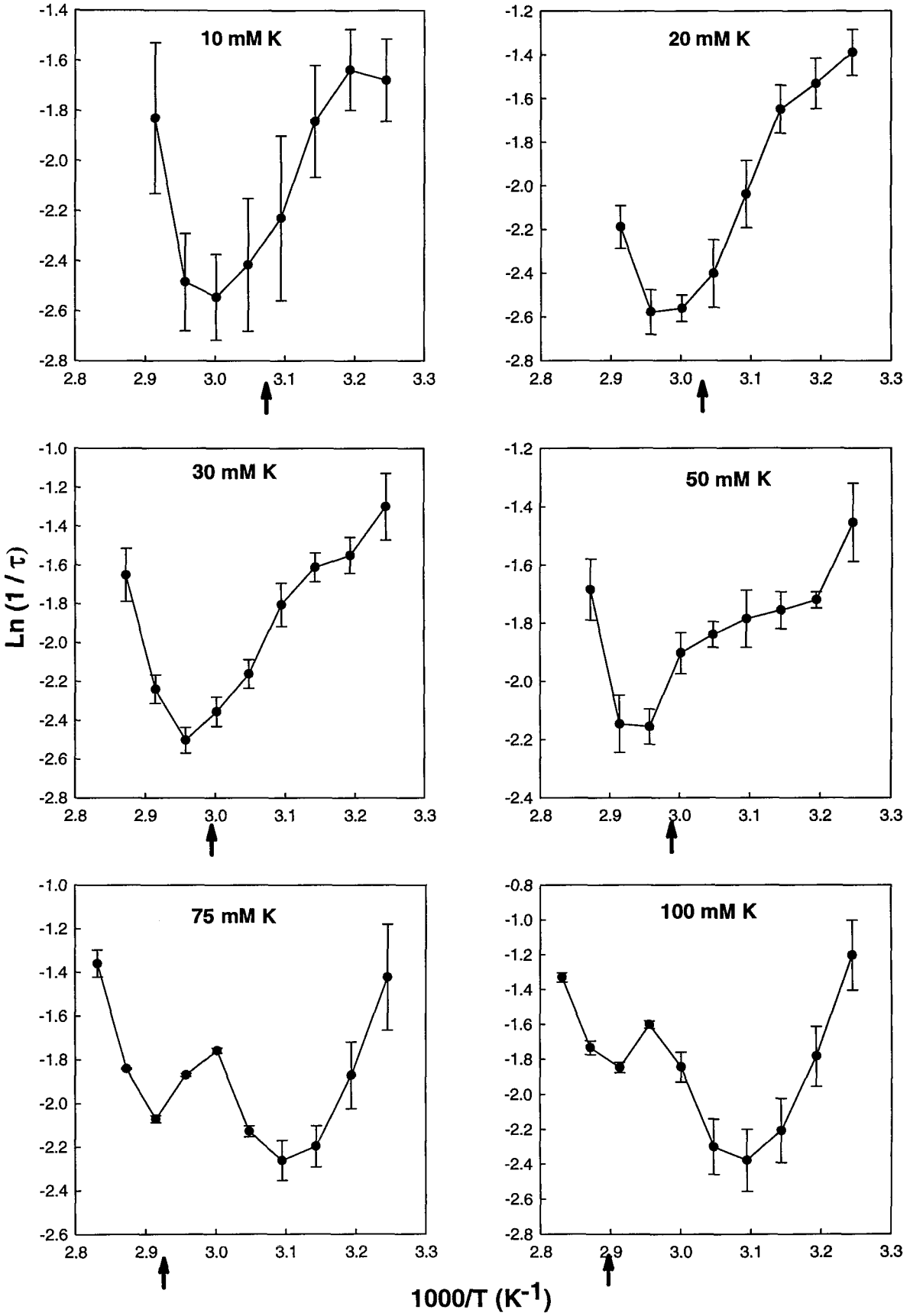


Figure 5.12. Arrhenius plots of relaxation times (1/τ) of HT3.5, derived from a single expontential function against temperature at different Na^+ concentrations. Arrows indicate T_m from Table 4.8.

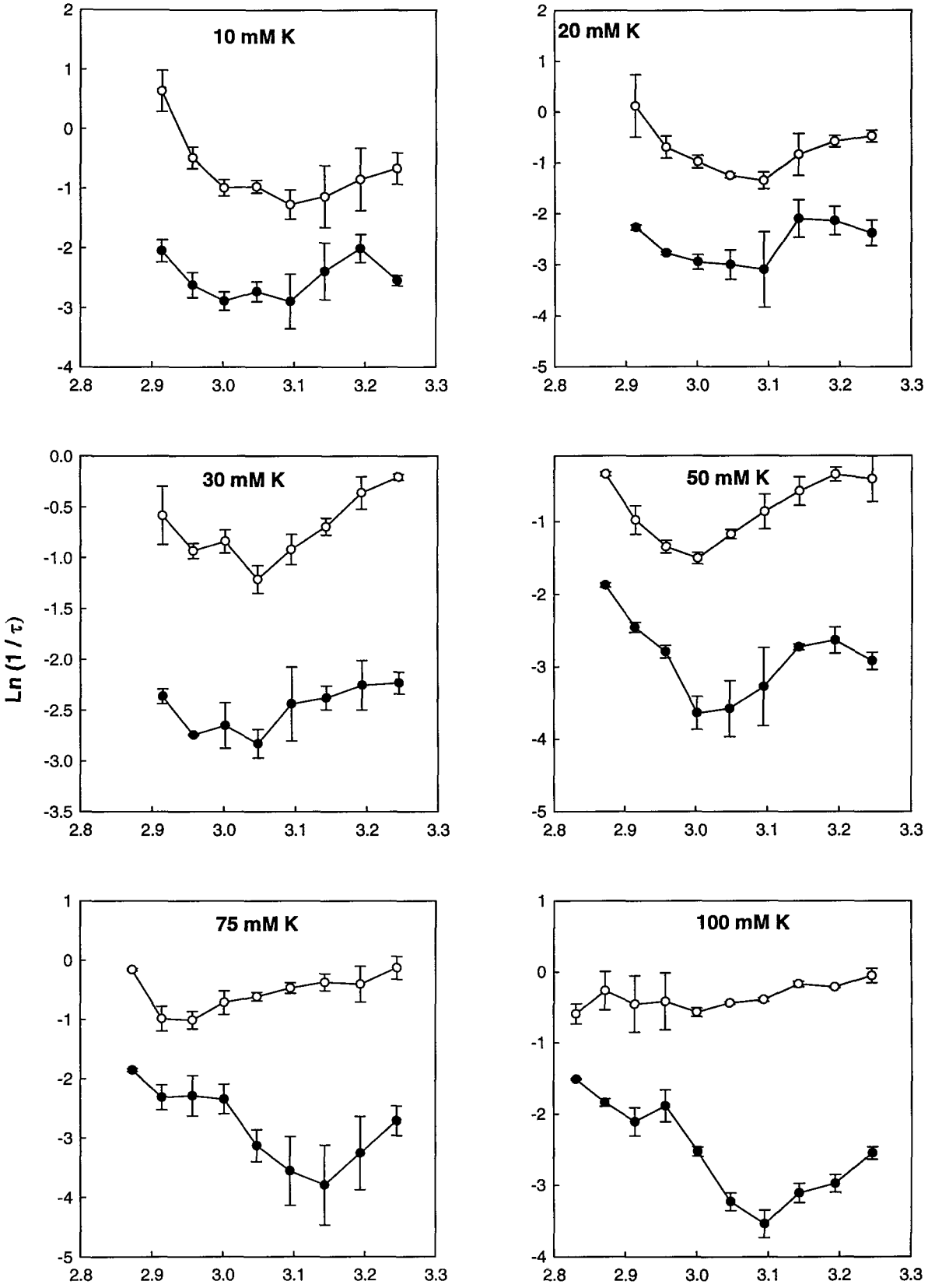


Figure 5.13. Arrhenius plots of relaxation times ($1/\tau_1$ and $1/\tau_2$) of HT3.5, derived from a double exponential function against temperature at different K^+ concentrations.

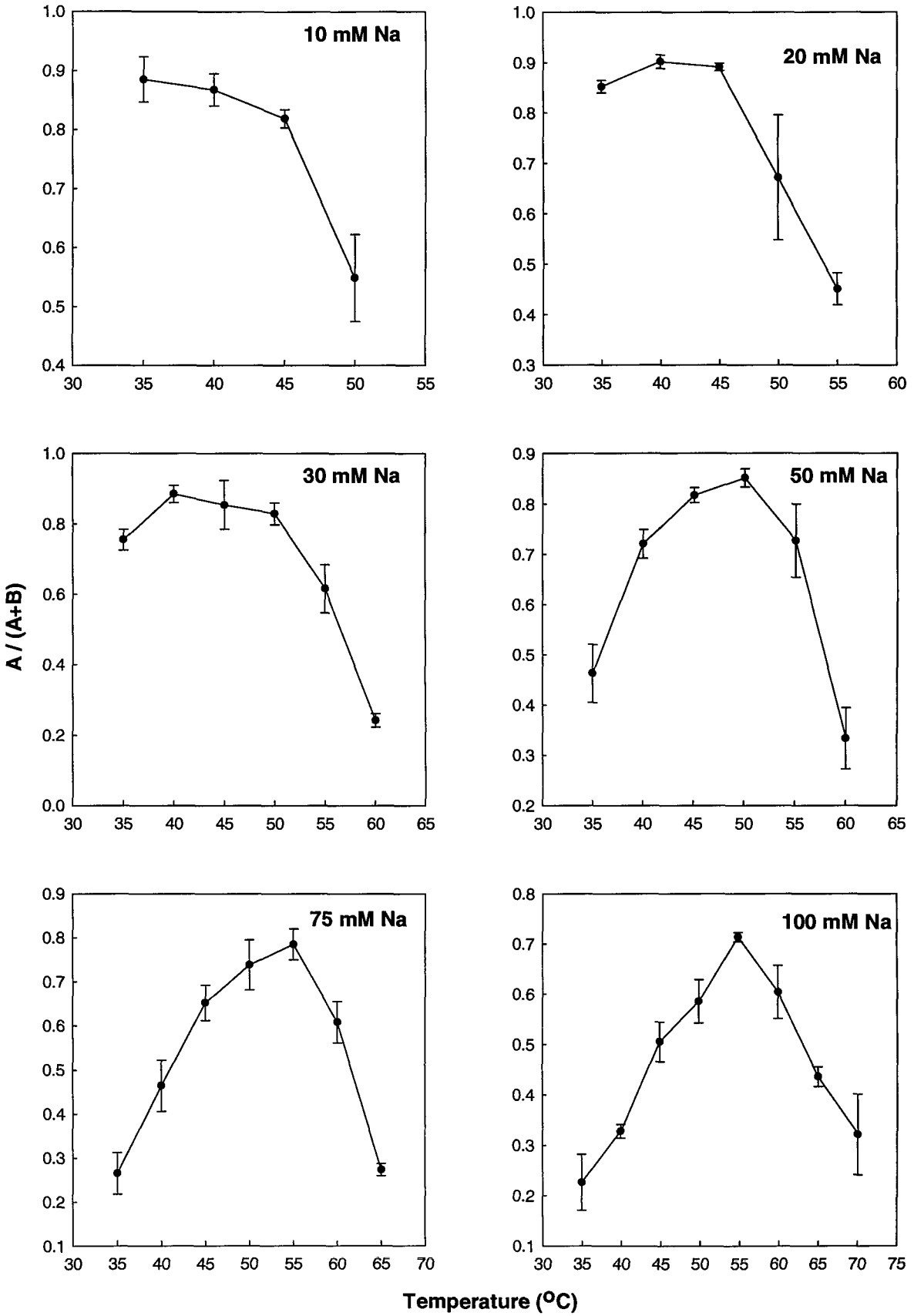


Figure 5.14. Plots of amplitude changes of the slow component (A)relative to those of the fast component (B), derived from double exponential fits to temperature jumps at different Na^+ concentrations.

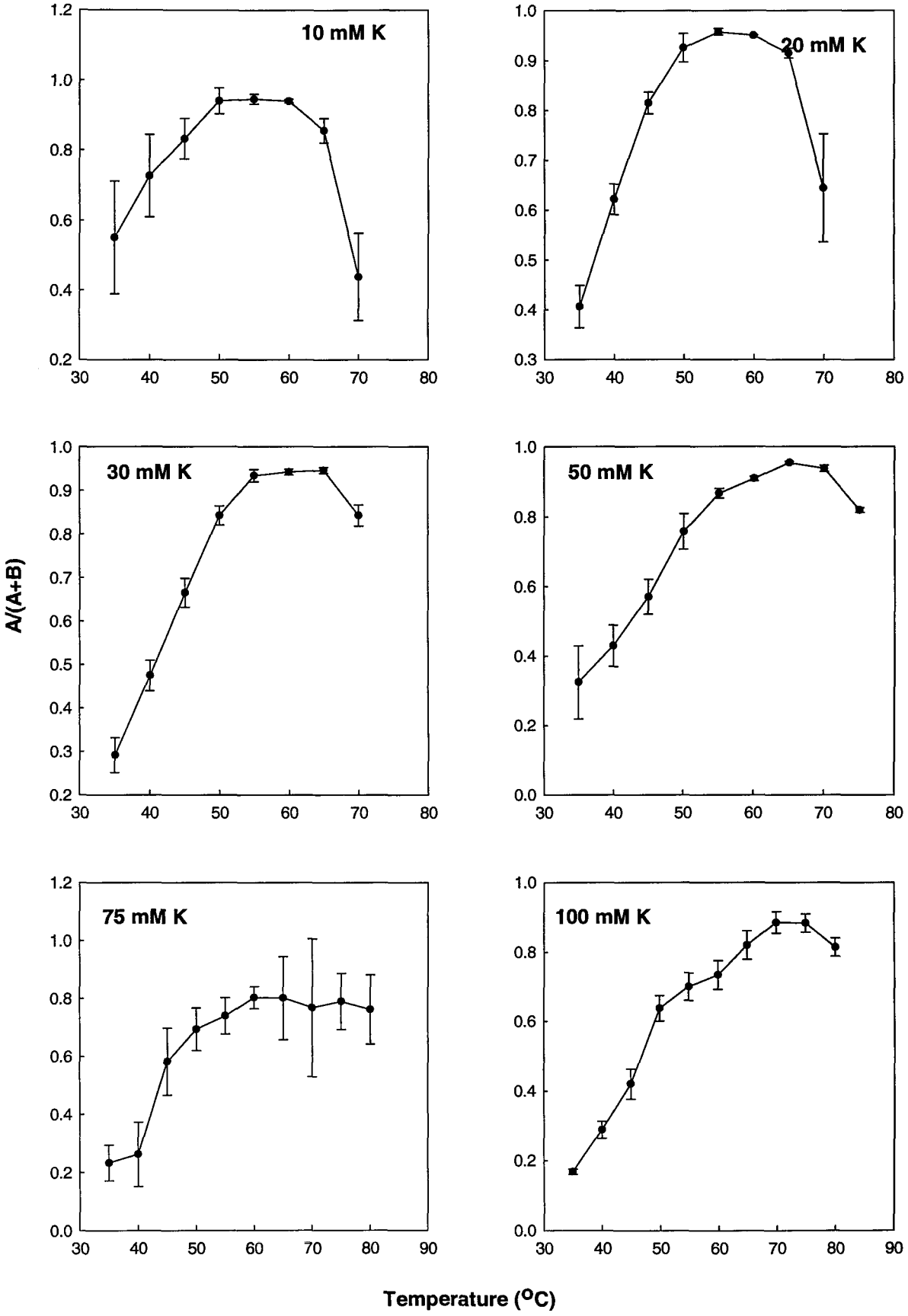


Figure 5.15. Plots of amplitude changes of the slow component (A) relative to those of the fast component (B), derived from double exponential fits to temperature jumps at different K^+ concentrations.

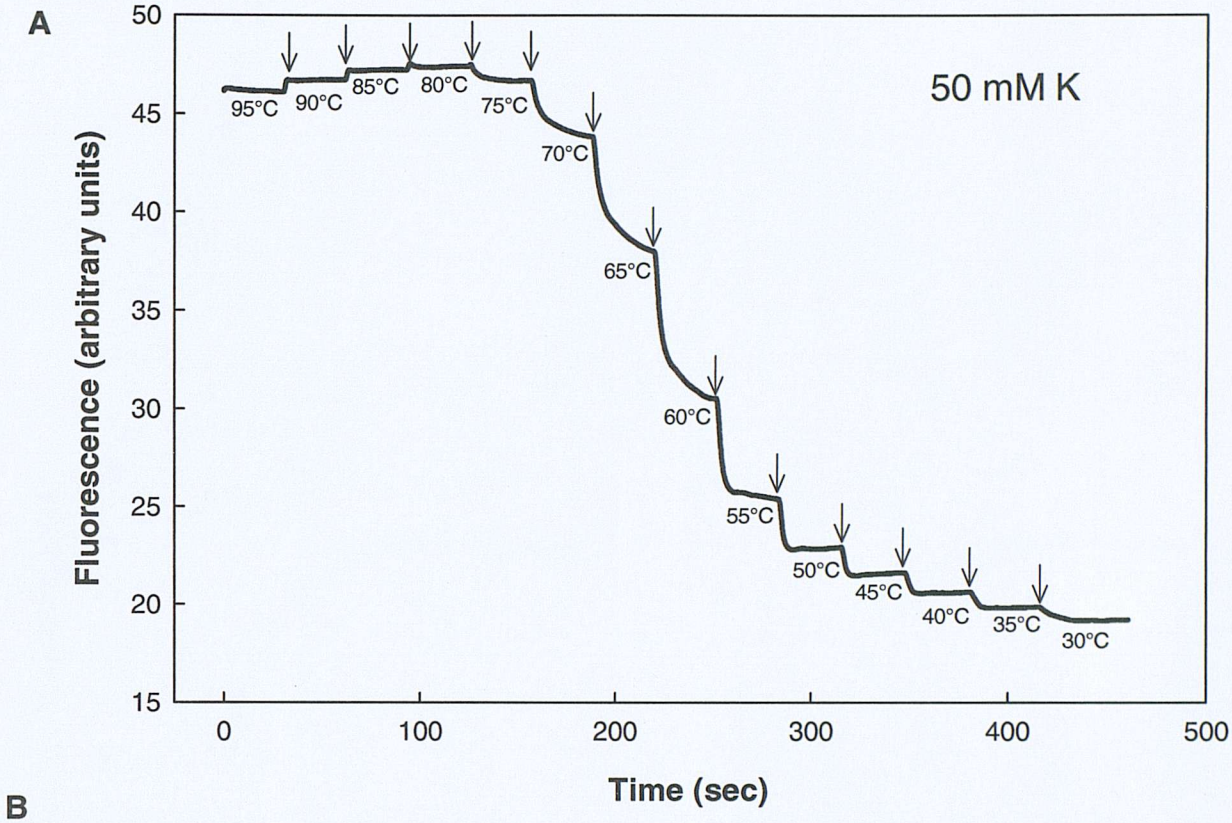
Temperature jumps were also carried out on the oligonucleotide HT1·3 in 10 mM K^+ and 100 mM K^+. This has the same 3.5 human telomeric repeats as HT3.5 and should also fold into an intramolecular quadruplex. However, different to HT3.5, the fluorophore is attached via a dU substituted for T6, while quencher is linked to a dU that replaces T18. Therefore the fluorophore and quencher are located on the central nucleotide of the first and third loops of the folded quadruplex (assuming it folds conventionially). HT1·3 should have similar kinetic properties to HT3.5, and did indeed exhibit two kinetic components to its temperature jump profiles (see Figure 5.17).

5.3 Discussion

The results presented in this chapter show that it is possible to use these fluorescently-labelled oligonucleotides in the LightCycler to measure the slow kinetics of quadruplex assembly. Time-dependent fluorescence changes were measured, following the temperature jumps, which occur on a time-scale of seconds/minutes, confirming the slow dynamics of these structures. The extremely slow nature of these reactions is emphasised by noting that most of the data were acquired at or around the melting temperature, and the reaction is slow even at these elevated temperatures.

Although the precise interpretation of the complex reaction profiles are not obvious (see below), some clear conclusions can be drawn from these data. Firstly, the dynamics of the *Oxytricha* repeat sequence OT3.5 are much slower than the human sequence HT3.5. It is not clear whether this is due to the length of the linker (4 bases instead of 3) or its precise sequence (T_4 instead of TTA) and experiments with other oligonucleotide sequences will be required to resolve this question. One might expect OT3.5 to show slower association kinetics, because of its increased flexibility it may be more likely to form a parallel quadruplex, which is known to have very slow kinetics. It is also possible that these sequences adopt different structures (parallel or antiparallel, with the loops crossed or uncrossed) which will also affect the dynamics.

Secondly, for both sequences, the reactions are much slower in the presence of K^+ than Na^+. This is consistent with the levels of hysteresis seen under similar conditions in melting experiments even at very slow rates of temperature change (chapter 4). Quadruplexes are known to be stabilised by cations, for which K^+ is more active than Na^+ and all other monovalent cations. (Sen & Gilbert, 1992). Therefore slower



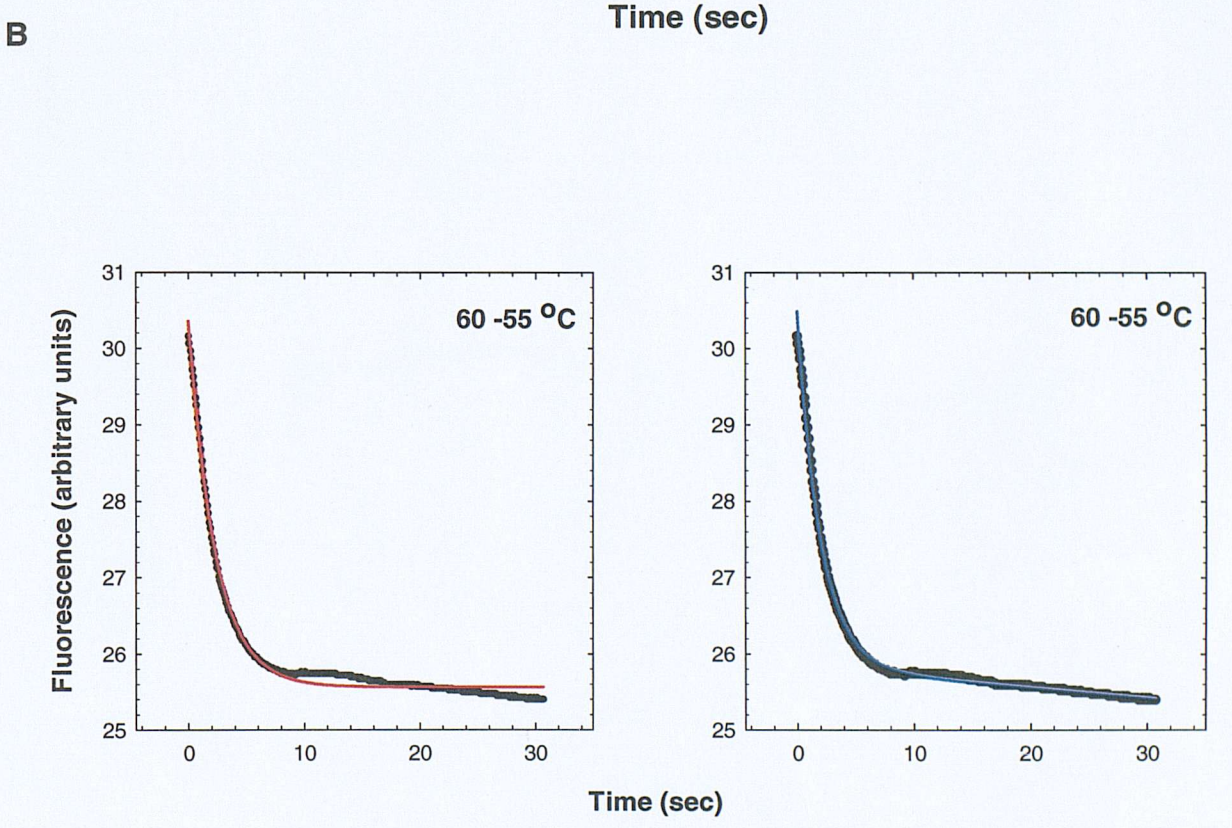


Figure 5.16. (A) Example of temperature decrease jumps obtained with HT3.5 in 10 mM KP/ 40 mM KCl. Arrows idicate the times at which the temperature was rapidly decreased by 5°C. (B) Example of a profile of temperature decrease jumps taken from the jump. A single exponential fit on the left (red) and a double exponential fit on the right (blue).

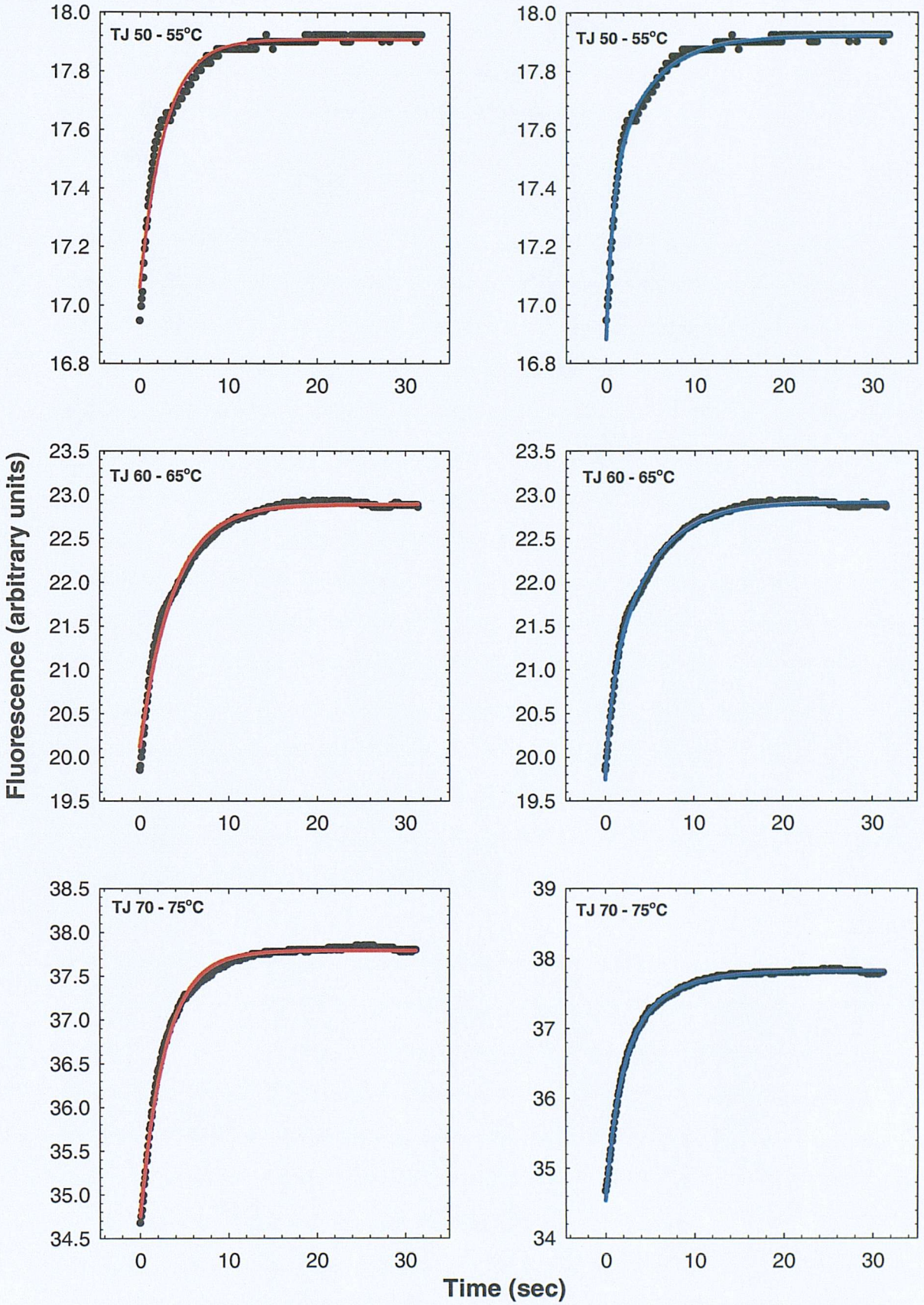


Figure 5.17. Examples of temperature jumps with HT1·3 in 10 mM KP / 90 mM KCl. A single exponential fiis shown on the left (red) and a double exponential fit on the right (blue).

relaxation kinetics with K^+ are consistent with greater stabilisation, but whether this arises from increased association produced by cation-facilitated nucleobase alignment and/or decreased dissociation resulting from internal cation coordination cannot be concluded.

Both HT1·3 and HT3.5 show two component relaxation kinetics. However, it is important to note that the process measured during the temperature jump only shows structural changes which affect the position of the fluorophore and quencher. The profiles therefore may not reflect all possible components. Two kinetic components for quadruplex formation have been previously reported for several quadruplex-forming oligonucleotides similar to the thrombin binding aptamer, as followed by absorbance change upon addition of K^+ over time (Jing *et al.*, 1997; Jing *et al.*, 1998). The authors proposed that the K^+-induced quadruplex folding was a two-step process, where a single slow process was observed at low K^+ concentration, while a second slow kinetic component became apparent and the first became too fast to see at increased K^+ concentrations. Two possible kinetic models can explain our observation of two kinetic components. The first is a two-step process, and the second is the presence of two distinct reactions which involve different quadruplex structures. The similar two component temperature jumps exhibited by HT1·3 implies that quadruplex reactions seen for HT3.5 are unlikely to represent a two-step reaction involving a intermediate formation of a cation-stabilised loop structure.

The relationship between the relaxation times and temperature is unusual but does seem to be related to the T_m of the reactions. The relaxation times appear to be faster at temperatures below the T_m of the reaction (as determined in chapter 4) and slower at temperatures above the T_m. A possible reason for this unusual temperature dependency is that association and dissociation may be unusually affected by temperature. This may arise if quadruplex formation involves transient intermediates (for example, a hairpin) which may be stabilised at low temperatures and thereby facilitate association (analogous to triplex formation via a nucleation-zipper mechanism). This would affect the observed relaxation times for which k_1 is a component (see equation 5.2). This and any other effects also make it difficult to determine the individual contribution of k_1 and k_{-1} to the relaxation times, and such considerations are further complicated if there are two relaxation times.

6 Identification of Strong Triplex Binding Sites by REPSA

6.1 Introduction

6.1.1 SELEX technologies

In the 1990s, so-called SELEX technologies (**s**ystematic **e**volution of **l**igands by **ex**ponential enrichment) were developed by several research groups. These have allowed the identification of many nucleic acid aptamers for biological ligands such as proteins, cofactors and small drug molecules (for reviews see Joyce, 1994; Klug & Famulok, 1994, Famulok & Szostak, 1992).

All such combinatorial methods involve the simultaneous screening of highly diverse (degenerate) pools of different single or double stranded RNA or DNA molecules for a specific property. A random pool of nucleic acids is incubated under conditions that allow the formation of ligand complexes. Functional binding sequences which form complexes are selected and separated from the non-binding pool. Using PCR, the functional subset is then amplified to an amount suitable for further manipulation. This series of steps constitutes a single round. Such pools are highly complex, and ligand-binding sequences constitute only a very small fraction of the overall population. As such any procedure can only produce a limited enrichment of the desired sequences each round, and effective selection requires multiple purification steps for functional sequences to dominate.

6.1.2 REPSA

All such techniques are limited by their reliance on the physical **separation** of bound and unbound nucleic acids, for example by their reduced electrophoretic mobility, increased hydrophobicity or filter binding. This is not always possible if the physical properties of a specific complex are not sufficiently distinct. To overcome this a new combinatorial method termed **r**estriction-**e**ndonuclease **p**rotection, **s**election and **a**mplification (REPSA) was devised by Van Dyke and co-workers which relies on a novel selection process. Selection is achieved by a **protection** step which involves the selective inhibition of cleavage of the nucleic acid aptamers by a type II-S restriction enzyme. These restriction enzymes cleave at a finite point distal to their recognition site in a sequence independent fashion. Ligand-binding protects the cleavage site within the

random section, and non-cleaved sequences alone can then be amplified, allowing selection of sequences with similar physical properties and under mild conditions. (see Figure 6.1).

REPSA was first used to identify a 13 base pair consensus sequence found within 46 sequences isolated after selection by a G,T triplex-forming oligonucleotide. Interestingly, a mismatch triplet G·AT was accommodated within several sequences (Hardenbol & Van Dyke, 1996). Another study using REPSA, extended the known consensus sequence of a human general transcription factor TFIID (Hardenbol *et al.*, 1997a). REPSA also predicted the preferred binding site of distamycin A (Hardenbol *et al.*,1996b) and actinomycin (Shen *et al.*, 2001). Recently, REPSA has been used to determine the nanomolar DNA-binding specificity of two hairpin polyamides. (Gopal & Van Dyke, 2003).

6.1.3 Aims

The sequence arrangement, composition and self-association of triplex-forming oligonucleotides (TFOs) affect their affinity for their target sites as discussed in chapter 1. The complex interplay of these factors make strong triplex formation more difficult to predict, and hinders the goal of producing a generally useful method for recognising any duplex targets which form the best triplex binding sites. This should enable the design of triplexes with superior stability and optimised triplex targeting.

To this aim, REPSA was modified to incorporate a **double aptamer** or **double degeneracy** approach. Such an approach has been successful at identifying some stable purine triplexes and illustrated the greater stability of the G·GC triplet which is widely accepted (Debin *et al.*, 1999). This also confirmed the requirement for contiguous G-tracts for generating strong triplexes and identified some unusual mismatches at specific positions.

The modified protocol involves selection of a sequence from a pool of degenerate duplex targets (detailed in Figure 6.1) which contain a random 10-mer homopurine tract, flanked by the recognition site for the type II-S restriction enzyme *Bsg*I GTGCAG, which cleaves 16 bases 3' to this site. Selection is based on the binding of degenerate pools of 10-mer TFOs which were of three different classes: (i) antiparallel

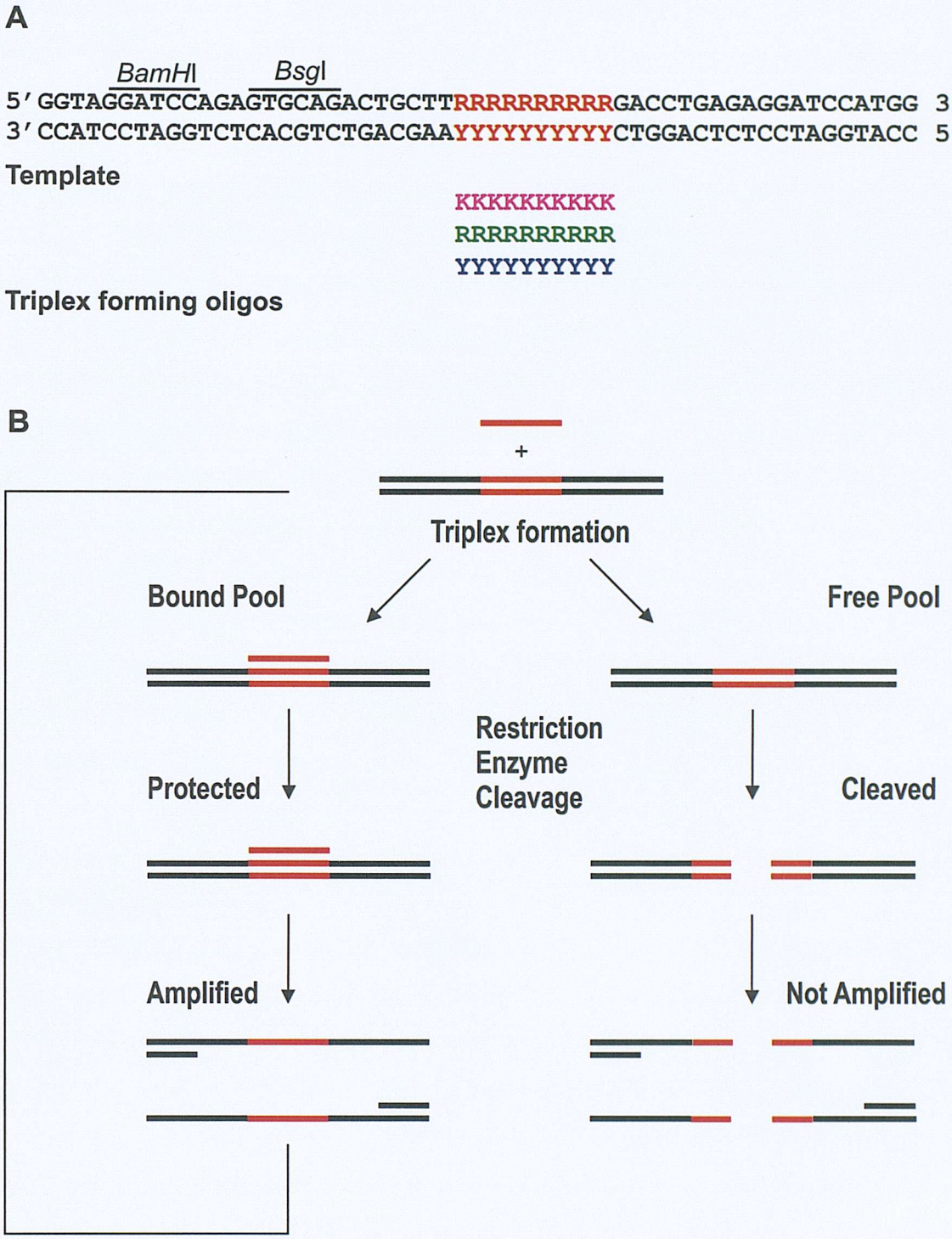


Figure 6.1. (A) Design of the degenerate selection template, used for identificationof duplexes capable of triplex formation with members of a degenerate population of TFOs from one of the 3 motifs *i.e.* K denotes G or T, R denotes G or A, and Ydenotes C or T. (B) Flow diagram of the combinatorial approach for selection by REPSA, the red region indicates the degenerate sequence of random bases.

binding purine decamers (the R_{10}-TFO population), (ii) parallel pyrimidine decamers (the Y_{10}-TFO population), and (iii) mixed G,T containing decamers that may adopt either orientation (the K_{10}-TFO population).

Under conditions known to be favourable for triplex formation, it was hoped that those duplexes capable of forming stable triplexes would encounter their respective third strand. The REPSA protocol would then allow selection of these sequences by removing many of those incapable of triplex formation, and enriching these sequences by amplification. This innovation aims to selectively enrich a range of duplexes within a degenerate duplex pool which form triplexes with a range of oligonucleotide aptamers within the degenerate TFO pool, **rather** than refining the degenerate duplex targets for specific binding of an individual ligand. This should therefore reflect the properties which dictate a strong triplex, in contrast to identifying a strong binding sequence for a particular ligand. In this way, it was hoped that duplexes capable of forming strong triplexes would be identified. This would allow the distinct factors that influence each triplex motif to be elucidated, including sequence constraints of both the duplex and the TFO.

6.2 Results

6.2.1 Preliminary results

Before initially performing REPSA is was important to ensure that the selection which relies on *Bsg*I could occur adequately. Figure 6.2 shows *Bsg*I cleavage of the ^{32}P-labelled random (duplex) template over time, carried out under the same conditions as the REPSA selection step (but without the previous protection step). It can be seen that the majority of the template is digested within 5 min, and that the addition of the random TFO pool immediately before cleavage does not interfere noticeably with the reaction. This indicates that selection is reliable and that the random TFO pool does not non-specifically affect cleavage by *Bsg*I .

Initial experiments were applied to screen the random pool of duplexes for strong binding to a random pool of G,T-third strands. The degenerate K_{10} population at a concentration of 3 to 8 μM, was incubated with nanomolar concentrations of the degenerate duplex population (radiolabelled as outlined in chapter 2) for ~18 h (overnight). The random K_{10} pool should consist of 2^{10} individual TFOs each of which

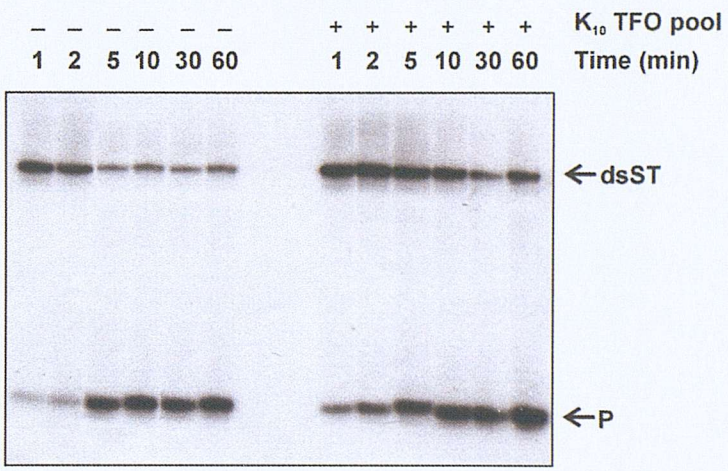


Figure 6.2. Example of *Bsg*I cleavage of the random template pool over time in the presence and absence of random K third strand. Time (min) is indicated above each lane, as is the presence or absence of K_{10} TFO pool.

should be at an effective concentration of ~ 3 to 8 nM. This incubation should be sufficient to facilitate the formation of any strong triplexes, which should then afford protect of those duplexes from subsequent digestion by *Bsg*I. The entire pool consisting of digested triplex-incompetent duplexes and selected undigested triplex-competent duplexes (plus any undigested triplex incompetent sequences) were then amplified by PCR. This amplification should enrich the best sequences by reducing the fraction of the population which does not form strong triplexes. 9, 12 and 15 cycles of PCR were carried out after each round of selection, and the products were separated by 8% native PAGE. This procedure was repeated for a total of 9 rounds of enrichment.

After each round of selection, the resulting sample produced by the lowest number of PCR cycles which could be visualised as a radioactive product band on a native gel was excised, the DNA eluted and used for the subsequent round of selection. Figure 6.3 shows the purification products from several rounds, the desired PCR product can be clearly seen, as can unincorporated ^{32}P-labelled primer in the final selection round 9 (this has run off the gel in the other samples). Primer dimer formation can also be seen as a low intensity band below the desired product in the samples separated following round 6. No digested duplexes can be seen following any step as no radiolabelled primer can be incorporated into these duplexes during the amplification step.

After **4 rounds** of selection, the enriched duplex pool was sampled. This was cloned into pUC18 plasmid, transformed into *E.coli*, and four duplex targets were sequenced. These were tested for triplex formation using DNaseI footprinting. These showed weak footprints for their respective parallel and antiparallel TFOs at high oligonucleotide concentrations. An example of such a footprint by the antiparallel third strand for the duplex target d(AAGGAAGGAGG) is shown in Figure 6.4A. Further selection was continued for a total of **9 rounds**. 15 duplexes were sequenced which are shown in Figure 6.5, and their affinities for third strands appraised. All duplex sequences determined from the pool retain their BsgI recognition site, indicating they did not arise by mutating this sequence. These affinities were determined using DNaseI footprinting with third strands synthesised to produce triplexes which consist of canonical triplets. Again, an example of the footprint produced by the sequence K3(i) with an antiparallel GT-TFO is shown in Figure 6.4B.

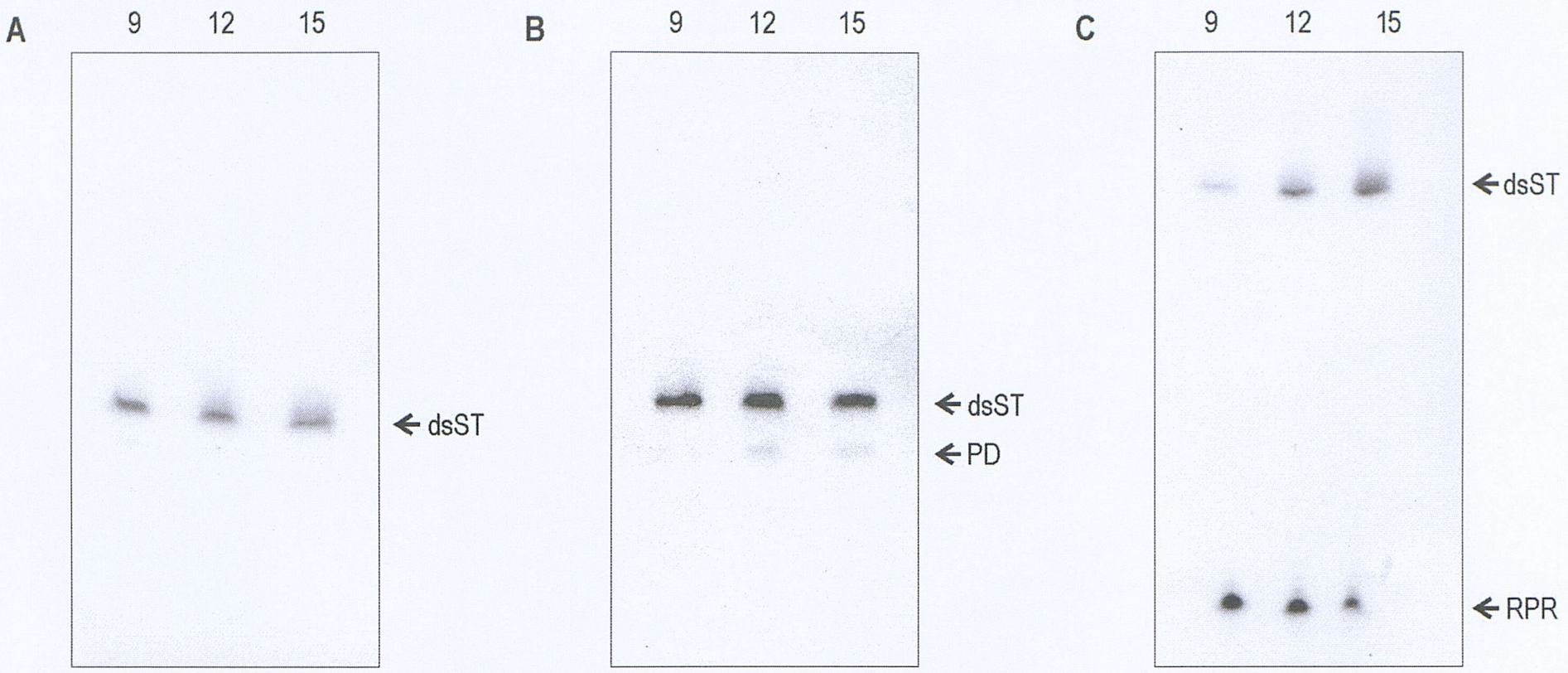


Figure 6.3. Autoradiographs of 8% native polyacrylamide gels, showing purified REPSA products visualised by incorporation of ^{32}P-labelled primer during PCR amplification. (A) Round 3of K10-REPSA,. (B) Round 6 and (C) Round 9. dsST represents the double stranded selection template, PD represents primer dimer and RPR represents labelled right primer. No of cycles of PCR is indicated above the lanes.

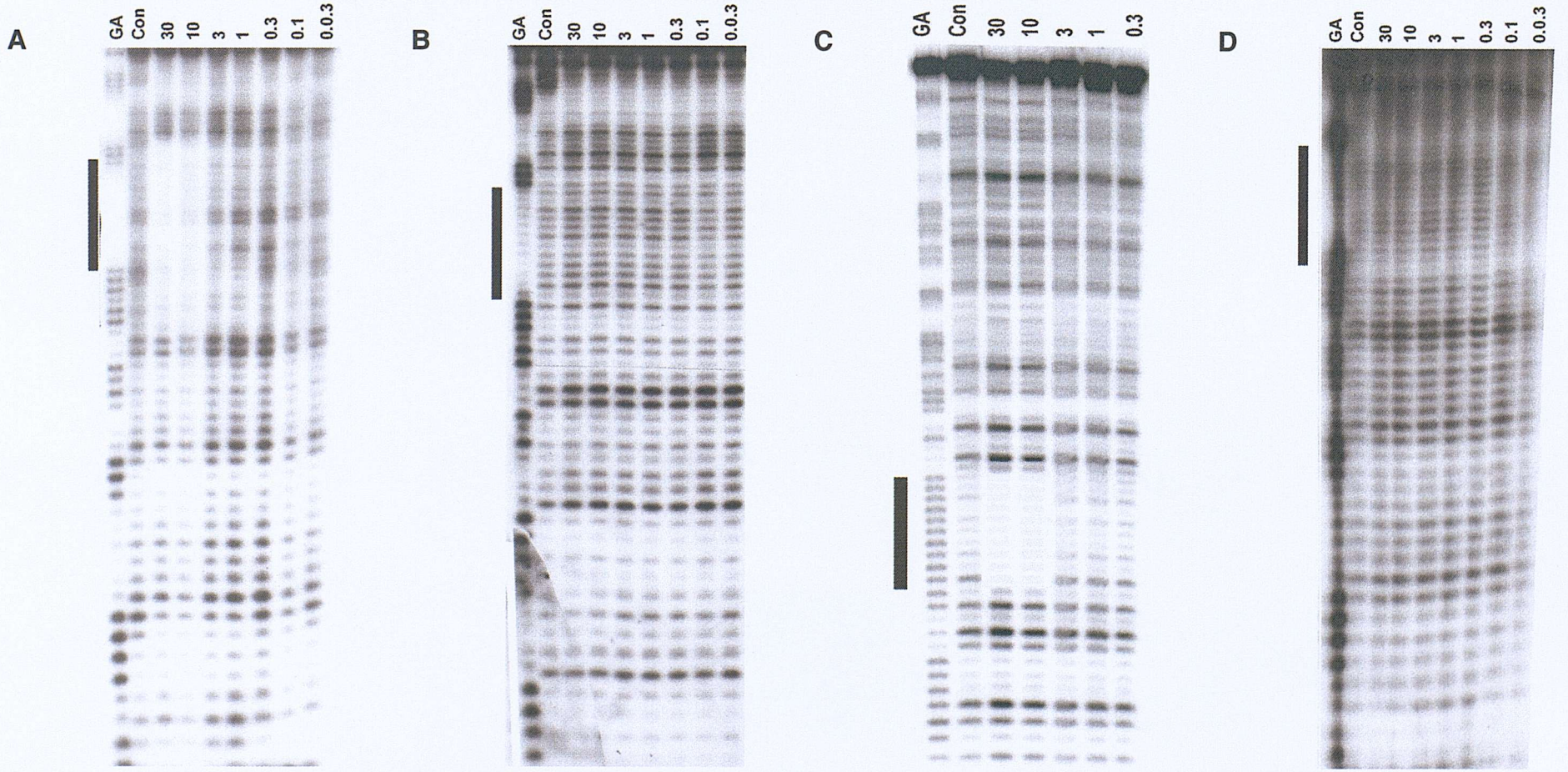


Figure 6.4. DNaseI digestion pattern showing the interaction of specifically synthesised TFOs for specific duplex targets yielded by REPSA. (**A**) the antiparallel G,T-TFO for a putative duplex d(AAGGAGGAGG), produced after 4 rounds of K_{10}-REPSA(I). (**B**) the antiparallel G,T-TFO for the duplex K3(I): d(GAGAGAGGAG), from K_{10}-REPSA(I).(**C**) the antiparallel G,A-TFO with duplex R3(II) d(AAAAGGAGA) produced following 12 rounds of R_{10}-REPSA(II). (**D**) the parallel C,T-TFO with target Y2(III) d(AAGGAGGGAA) produced following 12 rounds of Y_{10}-REPSA(III). Sequences are purine strands only of duplex (5' to 3'). TFO concentrations (μM) are shown at the top of each lane. The lane labelled 'con' shows DNaseI digestion in the absence of TFO. 'GA' represents a Maxam-Gilbert marker specific for purines. The filled box supposed target site and any limited triplex formation.

		G·GC	T·AT
Template	**5′-GGATCCAGAGTGCAGACTGCTTRRRRRRRRRRGACCTGAGAGGATCC-3′**		
Unmutated sequences			
K1 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTGGGGGGGGGAGACCTGAGAGGATCC-3′	9	1
K2 (I)	5′-GGATCCAGA**GTGCAG**ACTGCTTAGGAGAAAAAGACCTGAGAGGATCC-3′	3	7
K3 (I)	5′-GGATCCAGA**GTGCAG**ACTGCTTGAGAGAGGAGGACCTGAGAGGATCC-3′	6	4
K4 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTAAGAAAGAGAGACCTGAGAGGATCC-3′	3	7
K5 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTGGGGGAGAGAGACCTGAGAGGATCC-3′	7	3
K6 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTGAGGAGGGGGGACCTGAGAGGATCC-3′	8	2
K7 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTAAGGGGGGGGAGACCTGAGAGGATCC-3′	7	3
K8 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTGGAAGGGGAAGACCTGAGAGGATCC-3′	6	4
K9 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTAGAAAAGGAAGACCTGAGAGGATCC-3′	3	7
K10 (I)	5′-GGATCC GA**GTGCAG**ACTGCTTGGGGAGGAGGGACCTGAGAGGATCC-3′	8	2
Highly mutated sequences			
K11 (I)	5'-GGATCCAGA**GTGCAG**GACAGGAAGCTCGACTGGACCTGAGAGGATCC-3'	-	-
K12 (I)	5'-GGATCCAGA**GTGCAG**GACAGGATAATCGACTG GACCTGAGAGGATCC-3'	-	-
K13 (I)	5'-GGATCCAGA**GTGCAG**ACAG AGGGGTCGACTGGACCTGAGAGGATCC-3'	-	-
K14 (I)	5'-GGATCCAGA**GTGCAG**GACGGGGTGGTCGAGCTGGACCTGAGAGGATCC-3'	-	-
K15 (I)	5'-GGATCCAGA**GTGCAG**GATAGAAAATTAATTCGACTGGACCTGAGAGGATCC-3'	-	-

Figure 6.5. Initial K_{10}-TFO REPSA I (9 Rounds)**.** The yellow-shading indicates the random homopurine tract. The green-shading highlights mutations to the starting selection template. Bold lettering indicates *Bsg*I recognition site.

Based on an expectation of a modest selection, it was assumed that 9 rounds would be sufficient to yield a population with a high fraction of triplex-competent duplexes. For example, if each round enriches for triplex-competent sequences by as little as 10%, then 9 rounds should produce 61% enrichment. However, these sequences all showed very little or no affinity for their corresponding third strand. It is clear that the refined pool produced in these preliminary experiments has not generated high affinity triplex target sites, and these results were discouraging. Selection was obviously inadequate for isolation of strong triplex candidates from the random pool.

6.2.2 REPSA (II)

It was reasoned that the poor selection could be due to insufficient *Bsg*I cleavage during the selection step. Although *Bsg*I seems to cleave the duplex pool adequately (Figure 6.2), we therefore increased the amount of enzyme (from 3 to 6 units) and the incubation time (from 30 min to 1 h). Also the number of rounds of enrichment were increased to 12, though 9 rounds was adequate for the original studies by Van Dyke and co-workers using REPSA (Hardenbol *et al.*, 1997a; Hardenbol *et al.*,1996b; Shen et al., 2001; Gopal & Van Dyke, 2003).

12 rounds of selection were subsequently performed in two further series of experiments examining the binding of (i) the R_{10} random pool, and (ii) the Y_{10} random pool. The random duplex pool was cloned and the sequences were analysed after completing 12 rounds of selective enrichment. The sequences are summarised in Figure 6.6A and B. Their affinities were again examined and unfortunately no good triplex-forming targets could be found from among these sequenced. For an example see Figure 6.4C, where a footprint of weak affinity can be seen, however, no footprints were observed with any of the other sequences.

6.2.3 REPSA(III)

Again the process was revised, it was reasoned that the concentration of TFO might be too low since the individual concentrations of each constituent member within the random TFO pools was less than 10 nM in earlier experiments. Raising the overall concentration of the pool to 100μM should increase the individual sequence concentrations of each to ~100 nM. Although this concentration should facilitate the

		G·GC	A·AT
Template	**5′-GGATCCAGAGTGCAGACTGCTTRRRRRRRRRRGACCTGAGAGGATCC-3′**		
Unmutated sequences			
R1 (II)	5′-GGATCCAGA**GTGCAG**ACTGCTTGAGAAAGGAAGACCTGAGAGGATCC-3′	4	6
R2 (II)	5′-GGATCCAGA**GTGCAG**ACTGCTTAGGGAAGAGAGACCTGAGAGGATCC-3′	5	5
R3 (II)	5′-GGATCCAGA**GTGCAG**TCTGCTTAAAAAGGAGAGACCTGAGAGGATCC-3′	3	7
R4 (II)	5′-GGATCCAGA**GTGCAG**CCTGCTTAGGAGAGAAGGACCTGAGAGGATCC-3′	5	5
Highly mutated sequences			
R5 (II)	5′-GGATCCAGA**GTGCAG**GACAGAGGAATCGACTAGACCTGAGAGGATCC-3′	-	-
R6 (II)	5′-GGATCCAGA**GTGCAG**GACAGAAAAATCGACTAGACCTGAGAGGATCC-3′	-	-
R7 (II)	5′-GGATCCAGA**GTGCAG**GACAGAGAAATCGACTGGACCTGAGAGGATCC-3′	-	-
R8 (II)	5′-GGATCCAGA**GTGCAG**GCACTCAGCTCTAGCTCTAGCCGGACCTGAGAGGATCC	-	-
R9 (II)	5′-GGATCCAGA**GTGCGA**GACAGGAGTGTCGACCAGACCTGAGAGGATCC-3′	-	-

Figure 6.6A. R_{10}-TFO REPSA II (12 Rounds). The yellow-shading indicates the random homopurine tract. The green-shading highlights mutations to the starting selection template. Bold lettering indicates *Bsg*I recognition site.

Figure 6.6B. Y_{10}-TFO REPSA II (12 Rounds). The yellow-shading indicates the random homopurine tract. The green-shading highlights mutations to the starting selection template. Bold lettering indicates *Bsg*I recognition site.

selection of triplexes with lower binding affinity (*i.e.* higher K_d values), a K_d of ≥100 nM for a 10-mer oligonucleotide is still good.

Following these changes, REPSA was again undertaken, all three TFO types (K_{10}, R_{10}, Y_{10}) were used in concurrent experiments. The sequences obtained from these experiments are shown in Figure 6.7. Mutations can be seen in a significant proportion of these sequences, many of which show little similarity to the original duplex template. Irrespective of this fact, the triplex affinities of the less mutated sequences were again examined and were all found to be incapable of triplex formation at concentrations below 30 μM (See Figure 6.4D).

6.4 Discussion

Although this protocol is conceptually novel, the results of this work show that this combinatorial technique has not been successful. Several parameters have been changed without any appreciable improvement to the technique. No emergent population containing high affinity triplex target sites were identified for any of the motifs. Though initial experiments with the K_{10} TFO pool seemed promising (after the first 4 rounds three out of four sequences sampled produced limited DNaseI footprints), no subsequent success was achieved. Therefore, until the underlying limitations of this technique are identified and overcome, no rules can be concluded for the dependency of triplex formation on sequence composition or arrangement.

This double REPSA technique did not identify any strong triplex sites and a significant proportion of the sequences produced were highly mutated, both in the flanking regions, and within the random homopurine tract. Indeed, almost all the sequences have some mutations. Nevertheless as minor changes outside the random cassette should not affect triplex formations, these sequences were not immediately disregarded before determining whether they actually formed triplexes. In total, over half of all the sequences sampled from all the REPSA experiments were highly mutated, of the remainder, a total of 40 sequences showed no significant mutations. Among these were 14 for K_{10}, 14 for R_{10}, and 12 for Y_{10}. Though not every sequence was assayed for triplex affinity, enough were from each set of selections without any positive results to indicate that the technique has not worked as desired. Appraising the technique, there are several possible causes for the lack of success.

Template	**5′-GGATCCAGAGTGCAGACTGCTTRRRRRRRRRRGACCTGAGAGGATCC-3′**	G·GC	T·AT
Largely unmutated sequences			
K1 (III)	5′-GGATCCAGA**GTGCAG**ACTGCTTAGGGAAAAGGAGACCTGAGAGGATCC-3′	5	6
K2 (III)	5′-GGATCCAGA**GTGCAG**ACTGCTTAGAGGAA GACCTGAGAGGATCC-3′	3	4
K3 (III)	5′-GGATCC GA**GTGCAG**ACTGCTTGGAGAGGGAAGACCTGAGAGGATCC-3′	6	4
K4 (III)	5′-GGATCC GA**GTGCAG**ACTGCTTGAAGAAGAG GACCTGAGAGGATCC-3′	3	7
K5 (III)	5′-GGATCC GA**GTGCAG**ACTGCTTAAAGAAGAGAGACCTGAGAGGATCC-3′	3	7
Highly mutated sequences			
K6 (III)	5′-GGATCCAGA**GTGCAG**GACGGGCGGGTCGACTAGACCTGAGAGGATCC-3	-	-
K7 (III)	5′-GGATCCAGA**GTGCAG**GACAGGGGATCGACTG GACCTGAGAGGATCC-3′	-	-
K8 (III)	5′-GGATCCAGA**GTGCAG**GACAGCTTCTCGACTGGACCTGAGAGGATCC-3′	-	-

Figure 6.7A. K-TFO REPSA III (12 Rounds). The yellow-shading indicates the random homopurine tract. The green-shading highlights mutations to the starting selection template. Bold lettering indicates *Bsg*I recognition site.

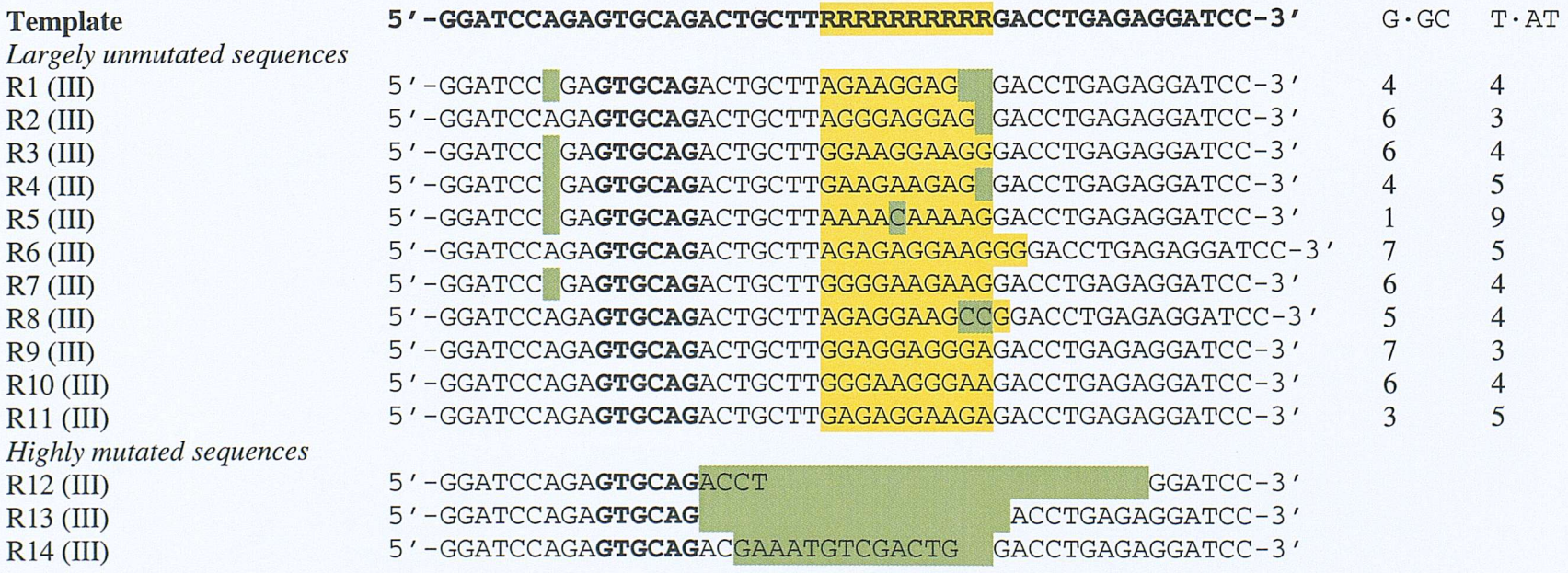


Figure 6.7B. R-TFO REPSA (III) (12 Rounds). The yellow-shading indicates the random homopurine tract. The green-shading highlights mutations to the starting selection template. Bold lettering indicates *Bsg*I recognition site.

		$C^+ \cdot GC$	$T \cdot AT$
Template	**5′-GGATCCAGAGTGCAGACTGCTTRRRRRRRRRRGACCTGAGAGGATCC-3′**		
Largely unmutated sequences			
Y1 (III)	5′-GGATCCGAGTGCAGACTGCTTAGAGAAAAGGACCTGAGAGGATCC-3′	3	6
Y2 (III)	5′-GGATCCAGAGTGCAGACTGCTTAAGGAGGGAAGACCTGAGAGGATCC-3′	5	5
Y3 (III)	5′-GGATCCAGAGTGCAGACTGCTTAGAAGGATAGACCTGAGAGGATCC-3′	5	3
Y4 (III)	5′-GGATCCGAGTGCAGACTGCTTAGGGAGAGGAGACCTGAGAGGATCC-3′	6	4
Y5 (III)	5′-GGATCCAGAGTGCAGACTGCTTAGAAGAGGACCTGAGAGGATCC-3′	3	4
Y6 (III)	5′-GGATCCGAGTGCAGACTGCTTGAATGAGAGGGACCTGAGAGGATCC-3′	5	5
Y7 (III)	5′-GGATCCGAGTGCAGACTGCTTGAAAAGAGACCTGAGAGGATCC-3′	2	5
Y8 (III)	5′-GGATCCAGAGTGCAGACTGCTTGAGAGAGAGGACCTGAGAGGATCC-3′	5	4
Y9 (III)	5′-GGATCCAGAGTGCAGACTGCTTAGAGAGGGACCTGAGAGGATCC-3′	4	3

Figure 6.7C. Y-TFO REPSA III (12 Rounds). The yellow-shading indicates the random homopurine tract. The green-shading highlights mutations to the starting selection template. Bold lettering indicates *Bsg*I recognition site.

Inadequate triplex-mediated protection

REPSA using pyrimidine (Y_{10}) TFOs resulted in the most mutated pool of sequences. This may reflect the use of non-optimal conditions for the formation of these triplexes. Conditions of low pH are required for generating C^+·GC triplets, while these experiments were performed at pH 7.4. If only T·AT and unprotonated C·GC triplets were present then this might prevent the formation of many complexes. Conversely, it could be argued that this should result in sequences which would give rise to triplexes with high T·AT content (though maybe of lower affinity), if the technique was functioning correctly.

The concept of double degenerate REPSA may itself be flawed. Inadequate pairing of individual TFOs with their duplex target might be an important limitation. Double degeneracy requires sufficient time to allow each sequence to encounter its partner and to then form a strong triple helix. An inherent problem might be the fact that levels of triplex-competent duplexes cannot be enriched above the concentrations of their TFO partners in the degenerate population.

Unusual triplex-mediated protection

It is possible (though unlikely) that the incorrect TFO partners were used for the footprinting studies. The sequences of these TFOs were synthesised by assuming that the triplexes only contained canonical triplets. This might arise if the triplexes formed (i) contain mismatched triplets accommodated by strong triplex formation; (ii) are shorter than assumed; or (iii) include base pairs outside the random homopurine tract cassette due to mismatched triplets or mutation into purines. Such events may allow triplex-mediated protection which facilitates selection.

Inadequate selection by the restriction enzyme

Inadequate selection and enrichment may have occurred due to non-optimal *Bsg*I cleavage conditions. This should not have been a major problem as the *Bsg*I is in excess over its substrate, 1-2 µl of *Bsg*I was used at an activity of 3 units/µl where 1 unit digests 1 µg λ-DNA/h (New England Biolabs specifications). While only nanograms of template were present. Conversely, too great an activity might digest some of the triplex- competent sequences, which had not encountered their respective third strands due to a limiting concentration of TFO and/or insufficient incubation time.

In this case more modest cumulative selection over each round might be required. As such, it is difficult to speculate whether (if either) an increased or decreased *Bsg*I activity is necessary for REPSA to work as desired.

The resistance of the template to cleavage due to its secondary structure may also be significant. However, it was hoped that any duplex targets with the ability to adopt unusual structures would be selected if they bound their complementary TFO with high affinity. Certain proprietary buffers supplied contain undesirably high levels of monovalent cations which may favour G-quadruplex formation and lead to sequestering of TFO into these structures. Furthermore, increasing the concentration of the degenerate TFO population will also lead to an increase in the formation of self-associated structures. This may be particularly relevant at the high concentration of the stocks and their repeated freeze-thawing which might encourage the formation of aggregates. The concern that the G-rich sequences might form quadruplexes seems unlikely as these short 10-mers should only be able to form tetrameric quadruplexes. Since this is a fourth order reaction, it will require very high DNA concentrations and a long timescale. Such limitations could be overcome by heating and crash-cooling the TFO pool before addition to the duplex pool. This might remove any self-associated structures in the TFO pool, and allow adequate competition between triplex and self-associated structures during the protection step.

Mutations to the *Bsg*I recognition site might generate cleavage-insensitive sequences, leading to the artificial selective enrichment of such. This is unlikely (though not impossible) as the recognition site is located within the flank that binds the right primer during PCR and is re-introduced during each PCR amplification. However, these possibilities can be discounted as *Bsg*I was able to digest the resulting pools following REPSA.

Improper PCR amplification

Poor copying by the *Taq* polymerase during PCR is another possibility. Although this enzyme does contain a proof-reading activity, it does generate some errors, this should not however account for the apparently high rate of mutations.

7 General Conclusion

7.1 Significance

Quadruplex research

DNA interactive agents are well established as clinically-useful anticancer drugs. However, most drugs presently available are non-specific, producing toxic side-effects. Consequently, much effort has been made to find agents which are more selective. The majority of this thesis looks at G-quadruplexes, which represent an interesting new class of molecular targets that may be relevant for the targeting of telomerase and transcriptional control (Schaffitel *et al.*, 2001; Siddiqui-Jain *et al.*, 2002). Both these processes are very significant to cancer aetiology.

Triplex research

Sequence specificity may also be useful for targeting cancer, viruses and drug resistant bacteria. Triplex-forming oligonucleotides are potentially useful code-reading molecules which can achieve limited sequence discrimination but require improvement before they are generally useful. Chapter 6 examined the sequence dependency of triplex formation which is a factor that limits their potential, using a **modified** REPSA method (Hardenbol *et al.*, 1997a; Hardenbol *et al.*,1996b; Shen et al., 2001; Gopal & Van Dyke, 2003).

7.2 Implications

Quadruplexes

The work presented in chapters 3, 4 and 5 considers the dynamics of quadruplex folding and unfolding by several G-rich sequences. Using fluorescence-melting and bandshift analysis these were examined, under specific cationic conditions that are known to promote quadruplex stability and kinetics (Sen & Gilbert, 1992; Muira *et al.*, 1995). Further, effects were closely studied of the addition of novel ligands from several classes of small molecules which are thought to interact with these secondary DNA structures (Sun *et al.*, 1997; Perry *et al.*, 1998a; Perry *et al.*,1998b).

Chapters 3 and 4 demonstrated the usefulness of the novel fluorescence-quenching technique (Darby *et al.*, 2002) as a means to accurately follow quadruplex formation.

The sequence comprised of four human telomeric repeats which forms an intramolecular quadruplex produced a sharper more defined melting profile than an identical sequence containing two repeats. The former sequence is more amenable to melting studies, possibly because of its kinetic parameters, or because of the greater polymorphic nature of dimeric rather than fold-back quadruplexes. However, intermolecular (dimeric) quadruplex-forming oligonucleotides are more useful for gel-based bandshift analysis than the intramolecular sequence. Quadruplex formation by the former produces a structure with double the molecular weight, which is easily resolved and quantified by electrophoresis, whereas formation of an intramolecular quadruplex has only a small mobility shift relative to its unfolded form.

A few studies have previously used a fluorescence melting technique (FRET) to identify potent leads within several series of G4 ligands (Koeppel *et al.*, 2001; Teulade-Fichou *et al.*, 2003), while most studies identifying new types of G4 ligands and assaying their SAR have used bandshift analysis (Han *et al.*, 1999; Han *et al.*, 2001; Koeppel *et al.*, 2001; Kim *et al.*, 2002). The data presented in chapter 3 assess a novel series of AQs, and show that these small molecules do indeed bind to and stabilise quadruplexes. This was assessed by ΔT_m changes in the melting profiles, and from the extent of quadruplex formation (after 4 h incubation) as measured by bandshifts. The results from these two techniques are generally in agreement. It is clear that the specific nature of the sidechains are an important determinant of their activity, possibly due to interactions with the loop and/or grooves of quadruplexes. It was found that quadruplex stabilisation by AQs substituted at the 2,6 positions seem to be better than at the 2,7 positions. Linkage via the amide nitrogen seems to be better than via the amide carbon, and in general, longer sidechains are better. The length of the substituents alone does not affect the activity, though the size and electronegativity of the backbone of the substituents appears to be significant. The studies included compounds which have been studied by other means, and these results were consistent with published results, giving additional validation of this new method.

Although the fluorescence melting technique provides a powerful means for studying quadruplex-ligand interactions, some compounds can affect the shape of the profile, or increase the breadth and asymmetry of the transition which complicates the analysis. This may occur if the ligands stack against the ends of the quadruplexes, or cause self-

association. In addition, the fluorescence signal can be affected by the ligands own spectral absorbance. Hysteresis during melting studies, resulting from non-equilibrium conditions can also be a constraint. This occurs if the rate of heating and cooling exceeds the kinetics of association and dissociation. This too is affected by the ionic conditions and the addition of ligands. Hysteresis can result from either the depression of annealing-T_m caused by a slow k_a or elevation of the melting-T_m caused by a slow k_d. All the drugs seemed to produce larger ΔT_m for annealing and melting of HT3.5 than HT1.5, though their relative activities are consistent between the two. The decrease in hysteresis obtained with HT1.5, and the increase in hysteresis observed with HT3.5, are both proportional to the ligands' activity. This may represent differences in the ability of ligands to promote formation (increasing association) or stabilise (decreasing dissociation) of the quadruplexes formed by these oligonucleotides.

Another interesting observation is the sigmoidal nature of the bandshift plots which imply cooperative binding. It could be reasoned that, for a bimolecular quadruplex, **two** strong and equivalent binding sites exist for stacking of two drugs at each end of the folded structure. External groove-binding as a mode for stabilising quadruplexes seems unlikely for these AQs as related compounds have been shown to intercalate or end-paste (Sun *et al.*, 1997; Perry *et al.*, 1998a; Perry *et al.*, 1998b; Read & Neidle, 2000). Though most quadruplex-binding compounds stack on only one end or stack more strongly on one end than the other, there is no reason to exclude the possibility of a 2:1 drug:quadruplex complex and there is NMR evidence suggesting that a G4 ligand can bind in this way (Gavathiotis *et al.*, 2001; Teulade-Fichou *et al.*, 2003).

The fluorescence melting experiments described in chapter 4 indicate that the kinetics of the *Oxytricha* telomeric repeat $d(G_4T_4)$ are very slow, much slower than for the human telomeric sequence $d(G_3T_2A)$, in agreement with other reports (Jing *et al.*, 1997; Marathias & Bolton, 1999). This is evident from the large hysteresis in the melting and annealing profiles which is greater for OT3.5 than HT3.5. As a result the melting studies with OT3.5 needed to be performed at much slower rates of temperature change. Some unusual profiles suggest that kinetically-trapped aberrant structures seem to be produced, and these are influenced by the exact order of annealing and melting. These factors resulted in the appearance of biphasic melting profiles, suggesting that the underlying molecular events are complicated and that the conditions do not allow

equilibrium to be achieved. Reducing the hysteresis by slowing the rates of heating and cooling simplifies these events, making profiles more easy to interpret.

Varying the cationic conditions influences quadruplex stability and kinetics (Sen & Gilbert, 1992; Muira *et al.*, 1995). Increased Na^+ or K^+ concentration reduces the hysteresis, suggesting that the increase in kinetics reflects an increased rate of association. Fluorescence melting experiments with HT3.5 and OT3.5 in the presence of their complementary strands show fluorescence profiles which are the sum of both duplex and quadruplex melting curves. Though these are complicated profiles, these give some information on the quadruplex – duplex equilibria. These suggest that Na^+ fails to support quadruplex formation at all temperatures, as duplex DNA is present at low temperatures, with single strands at higher temperatures. In the presence of K^+ the duplex form also predominates at low temperatures, but a quadruplex can form above the duplex melting temperature. These observations are a result of both kinetic and thermodynamic factors.

The results of temperature jump relaxation experiments, using a modification of the fluorescence melting protocol are presented in chapter 5. Similar experiments using molecular beacons to report on the kinetics of triplex formation have recently been reported (James *et al.*, 2003). The results show that the kinetic profiles were often biphasic, complicating the analysis of the component relaxation times and the parameters derived from a single and double exponential fits to the data did not yield the usual increase in rate with temperature as expected. For both the sequences the kinetics of folding are much slower in the presence of K^+ than in Na^+.

7.3 Future work

Quadruplex

Na^+ and K^+ are accepted as the most important monovalent cations for forming stable quadruplexes *in vivo*, with K^+ as the predominant intracellular ion. Most quadruplex studies are performed using one or other cations in isolation, whereas both will be present, at different concentrations *in vivo*. A logical extension of this would be to investigate quadruplex formation in the presence of several ions together, mimicking the intracellular concentrations.

The effects of the novel G4 ligands on quadruplex formation could be extended to determine their influence on quadruplex – duplex equilibrium in experiments similar to those described in chapter 4. These would be a good reflection of the situation *in vivo* in which quadruplex formation often has to compete with the interactions of the complementary strands.

Further sequences worth studying (to determine their thermodynamic and kinetics parameters and the effects of ligands) include the *Tetrahymena* telomeric repeat and other G-rich sequences found within promoter regions, or in the RNA of HIV. The relationship between sequence and stability might also be explored by a rational progression from $d(G_4T_4)$ repeats to various other sequences such as (G_nT_4) or (G_4T_n). For bandshift experiments an oligonucleotide similar to TR4, but lacking the additional thymines may allow greater separation between the bands.

Triplex

Several possible further modifications could be made to the REPSA protocol described in the triplex research. If this can be successfully revised then it could be used to study the stringency of triplex formation, under different conditions such as the presence of monovalent and divalent cations, pH, temperature, sequence composition and arrangement. A possible variation worth exploring might be a series of **partial double degenerate** REPSA experiments. This could be achieved by introducing some degeneracy in the ligand pool with parts of the sequence fixed. This might involve conjugating a triplex-binding ligand to a random oligomer library to increase the affinity while hopefully not affecting the relative triplex strengths. The ligand recognition site could also be located adjacent to the random tract, but sufficiently downstream of the II-S restriction enzyme cleavage site to prevent any interference. The random tract could possibly then be N_{10} rather than R_{10}, allowing the formation of triplexes with some non-canonical triplets. This might also eliminate the possibility of mutations in the template strand by reducing third strand slippage.

8 References

Alberti P., Ren J., Teulade-Fichou M. P., Guittat L., Riou J.-F., Chaires J.B., Hélène C., Vigneron J.-P., Lehn J.-M. & Mergny J.-L. (2001). Interaction of an acridine dimer with DNA quadruplex structures. *J. Biomol. Struct. & Dyn.*, **19**, 505-513.

Alberti P., Schmitt P., Nguyen C.-H., Rivalle C., Hoarau M., Grierson D.S.& Mergny J.-L. (2002). Benzoindoloquinolines interact with DNA tetraplexes and inhibit telomerase.*Biorg. Med. Chem. Lett.*, **12**, 1071-1074.

Allsopp R.C., Chang E., Kashefi-Aazam M., Rogaev E.I., Piatyszek M.A., Shay J.W. & Harley C.B. (1997). Telomere shortening is associated with cell division *in vitro* and *in vivo*. *Exp. Cell Res.*, **220**, 194-200.

Anantha N.V., Azam M. & Sheardy R.D. (1998). Porphryin binding to quadruplexed T_4G_4. *Biochemistry*, **37**, 2709-2714.

Arimondo P.B., Barcelo F., Sun J.-S., Maurizot J.-C., Garestier T. & Hélène C. (1998). Triple helix formation by (G,C)-containing oligonucleotides: assymetric sequence effect. *Biochemistry*, **37**, 16627-16635.

Arimondo P.B., Riou J.-F., Mergny J.-L., Tazi J., Sun J.-S., Garestier T. & Hélène C. (2000). Interaction of human topoisomerase I with G-quartet structures. *Nucleic Acids Res.*, **28**, 4832-4838.

Asensio J.L., Lane A.N., Dhesi J., Bergqvist S. & Brown T.(1998). The contribution of cytosine protonation to the stability of parallel DNA triple helices. *J. Mol. Biol.*, **275**, 811-822.

Awang G. & Sen D. (1993). Mode of dimerization of HIV-1 genomic RNA. *Biochemistry*, **32**, 11453-11457.

Balagurumoorthy P. & Brahmachari S.K. (1994). Structure and stability of human telomeric sequence. *J. Biol. Chem.*, **269**, 21858-21869.

Baran N. Pucshansky L., Marco Y., Benjamin S. & Manor H.(1997). The SV40 large T-antigen helicase can unwind four stranded DNA structures linked by G-quartets. *Nucleic Acids Res.*, **25**, 297-303.

Bates P.J., Laughton C.A., Jenkins T.C., Capaldi D.C., Roselt P.D., Reese C.B. & Neidle S. (1996). Efficient triple helix formation by oligodeoxyribonucleotides containig α- or β-2-amino-5-(2-deoxy-D-ribofuranosyl) pyridine residues. *Nucleic Acids Res.*, **24**, 4176-4184.

Berg J.M., Tymoczko J.L. & Stryer L. (2003). *Biochemistry (5th edition)*, chapter 27. WH Freeman & Co, New York.

Bijapur J., Keppler M.D., Bergqvist S., Brown T. & Fox K.R. (1999). 5-(1-propargylamino)-2'-deoxyuridine(dU^P): a novel thymidine analogue for generating DNA triplexes with increased stability.*Nucleic Acids Res.*, **27**, 1802-1809.

Blume S.W., Gee J.E., Shrestha K. & Miller D.M. (1992). Triple helix formation by purine-rich oligonucleotides targeted to the human dihydrofolate reductase promoter. *Nucleic Acids Res.*, **20**, 1777-1784.

Blume S. W., Guarcello V., Zacharias W. & Miller D. M.(1997). Divalent transition metal cations counteract potassium-induced quadruplex assembly of oligo(dG) sequences. *Nucleic Acids Res.*, **25**, 617–625.

Bock L.C., Griffin L.C., Latham J.A., Vermaas E.H. & Toole J.J. (1992). Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, **355**, 564-566.

Bodnar A.G., Ouellette M., Froklis M., Holt S.E., Chu C.-P., Morin G.B., Harley C.B., Shay J.W., Lichsteiner S. & Wright W.E. (1998). Extension of life-span by

introduction of telomerase into normal human cells. *Science*, **279**, 349-352.
Brown P.M., Madden C.A. & Fox K.R. (1998). Triple-helix formation at different positions on nucleosomal DNA. *Biochemistry*, **37**, 16139-16151.
Burger A.M., Double J.A. & Newell D.R. (1997). Inhibition of telomerase activity by cisplatin in human testicular cancer cells. *Eur. J. Cancer.*, **33**, 638-644.
Brunar & Dervan P.B. (1996). Sequence composition effects on the stability of tripe helix formation by oligonucleotides containing N7-deoxyguanosine. *Nucleic Acids Res.*, **24**, 1987-1991.
Callender R. & Dyer R.B. (2002). Probing protein dynamics using temperature jump relaxation spectroscopy. *Curr. Opin. Struct. Biol.*, **12**, 628-633.
Chandler S.P. & Fox K.R. (1993). Triple helix formation at $A_8XA_8{\cdot}T_8YT_8$. *FEBS lett.*, **332**, 189-192.
Chandler S.P. & Fox K.R. (1995). Extension of DNA triple helix formation to a neighbouring $(AT)_n$ site. *FEBS lett.*, **360**, 21-25.
Chandler S.P. & Fox K.R. (1996). Specificity of antiparallel DNA triple helix formation. *Biochemistry*, **35**, 15038-15048.
Chen Q., Kuntz I.D. & Shafer R.H. (1996). Spectroscopic recognition of guanine dimeric hairpin quadruplexes by a carbocyanine dye. *Proc. Natl. Acad. Sci. USA*, **93**, 2635-2639.
Cheng A.J. & Van Dyke M.W.(1994). Oligodeoxyribonucleotide length and sequence effects on intermolecular purine-purine-pyrimidine triple-helix formation. *Nucleic Acids Res.*, **22**, 4742-4747.
Chung I.K., Mehta V.B., Spitzner J.R. & Muller M.T. (1992). Eukaryotic topoisomerase II cleavage of parallel DNA tetraplexes. *Nucleic Acids Res.*, **20**, 1973-1977.
Clarenc J.P., LeBleu B. & Leonetti J.P. (1993). Characterization of the nuclear binding sites of oligodeoxyribonucleotides and their analogs. *J. Biol. Chem.*, **268**, 5600-5604.
Clark G. R., Pytel P.D., Squire C.J. & Neidel S. (2003). Structure of the first parallel DNA quadruplex-drug complex. *J. Am. Chem. Soc.*, **125**, 4066-4067.
Cocco M.J., Hanakahi L.A., Huber M.D. & Maizels N. (2003). Specific interactions of distamycin with G-quadruplex DNA. *Nucleic Acids Res.*, **31**, 2944-2951.
Counter C.M., Avilion A.A., LeFeuvre C.E., Stewart N.G., Grieder C.W., Harley C.B. & Baccheti S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.*, **11**, 1921-1929.
Cooney M., Czernuszewicz G., Postel E.H., Flint S.J. & Hogan M.E. (1988). Site-specific oligonucleotide binding represses transcription of the human *c-myc* gene *in vitro*. *Science*, **241**, 456-459.
Cuenoud B., Casset F., Husken D., Natt F., Wolf R.M., Altmann K.H., Martin P. & Moser H.E. (1998). Dual recognition of double-stranded DNA by 2'aminoethoxy-modified oligonucleotides. *Angew. Chem. Int. Ed.*, **37**, 1288-1291.
Darby R.A.J., Sollogoub M., McKeen C., Brown L., Risitano A., Brown N.M., Barton C., Brown T. & Fox K.R. (2002). High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. *Nucleic Acids Res.*, **30**, *e39*.
Debin A., Laboulais C., Ouali M., Malvy C., Le Bret M. & Svinarchuk F. (1999). Stability of G,A triple helices. *Nucleic Acids Res.*, **27**, 2699-2707.
Dempsey L.A, Sun H., Hanakah L.A. & Maizels N. (1999). G4 DNA binding by LR1 and its subunits, nucleolin and hnRBP D, a role for G-G pairing in immunoglobulin switch recombination. *J. Biol. Chem.*, **274**, 1066-1071.
Deng H. & Braulin W.H. (1995). Duplex to quadruplex equilibrium of the self-

complimentary oligonucleotide d(GGGGCCCC). *Biopolymer*, **35**, 677-681.
Doronina S.O. & Behr J.-P. (1997). Synthesis of 4-guanidinopyrimidine nucleosides for triple helix mediated guanine and cytosine recogntion. *Tetrahedron Lett.*, **39**, 547-550.
Duan W., Rangan A., Vankayalapati H., Kim M.-Y., Zeng Q., Dun D., Han H., Fedoroff O.Y., Nishioka D., Rha S.Y., Izbicka E., Von Hoff D.D. & Hurley L.H. (2001). Design and synthesis of fluoroquinophenoxazines that interact with human telomeric G-quadruplexes and their biological effects. *Mol. Cancer Ther.*, **1**, 103-120.
Durland R.H., Kessler D.J., Gunell S., Duvic M., Pettitt B.M. & Hogan M.E. (1991). helix-forming oligonucleotides to sites in gene promoters. *Biochemistry*, **30**, 9246-9455.
Eigen M. & De Maeyer L.D. (1963). Relaxation methods. In Friess S.L., Lewis E.S. & Weissberger A. (ed.) *Techniques of organic chemistry*. Interscience , New York. pp 895-1054.
Eldrup A.B. Dahl O. & Nielsen P.E.(1997). A novel peptide nucleic acid monomer for recogntiton of thymine in triple helix structures. *J. Am. Chem. Soc.*, **119**, 1116-1117.
Elenbaas B. Spirio L., Koerner F., Fleming M.D., Zimonjic D.B., Donaher J.L., Popescu N.C. Hahn W.C. & Weinberg R.A. (2001). Human breast cancer cells generated by oncogenic transformation of primary mammalian epithelial cells. *Genes Dev.*, **15**, 50-65.
Erlitzki R. & Fry M. (1997). Sequence-specific binding protein of single-stranded and unimolecular quadruplex telomeric DNA from rat hepatocytes. *J. Biol. Chem.*, **272**, 15881-15890.
Evans T., Schon E., Goramaslak G., Patterson J. & Efstratiadis A. (1984). S1-hypersensitive sites in eukaryotic promoter regions. *Nucleic Acids Res.*, **12**, 8043-8058.
Famulok M. & Szostak J.W. (1992). *In vitro* selection of specific ligand binding nucleic acids. *Angew. Chem. Int. ed. Engl.*, **31**, 979-988.
Fang G. & Cech T.R. (1993). Characterization of a G-quartet formation reaction promoted by the β-subunit of the *Oxytricha* telomere-binding protein. *Biochemistry*, **32**, 11646-11657.
Faruqi A.F., Seidman M.M., Segal D.J., Carroll D. & Glazier P.M. (1996). Recombination induced by triple helix –targeted DNA damage in mammalian cells. *Mol. Cell. Biol.*, **16**, 6820-6828.
Faruqi A.F., Krawczyk S.H., Metteucci M.D. & Glazier P.M. (1997). Potassium-resistant triple helix formation and improved intracellular gene targeting by oligodeoxyribonucleotides containing 7-deazaxanthine. *Nucleic Acids Res.*, **25**, 633-640.
Fedoorff O.Y., Salazar M., Han H., Chemeris V., Kerwin S.M. & Hurley L.H.(1998). NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry*, **36**, 12367-12374.
Felsenfeld G., Davis D.R. & Rich A.(1957). Formation of a three stranded polynucleotide molecule. *J. Am. Chem. Soc.*, **79**, 2023-2024.
Felsenfeld G. & Rich A.(1957). Studies on the formation of two- and three-stranded polyribonucleotides. *Biochim. Biophys. Acta*, **26**, 457-468.
Fersht, A. (1977). *Enzyme structure and mechanism*, chapter 4. WH Freeman & Co, New York.
Fossella J.A., Kim Y.J., Shih H., Richards E.G. & Fresco J.R.(1993).Relative specificities in binding of Watson-Crick base-pairs by 3rd strand residues in a DNA pyrimidine triplex motif. *Nucleic Acids Res.*, **21**, 4511-4515.

Fox K.R. (1994). Formation of DNA triple helices incorporating blocks of G.GC and T.AT triplets using short acridine-linked oligonucleotides.*Nucleic Acids Res.*, **22**, 2016-2021.

Fox K.R., Polucci P., Jenkins T.C. & Neidle S. (1995). A molecular anchor for stabilizing triple helical DNA. *Proc. Natl. Acad. Sci. USA* **92**, 7887-7891.

Fox K.R. (2000). Targeting DNA with triplexes. *Curr. Med. Chem.*, **7**, 17-37.

Fox K.R., Flashman E. & Gowers D.M. (2000). Secondary binding sites for triplex-forming oligonucleotides containing bulges, loops, and mismatches in the third strand. *Biochemistry*, **39**, 6714-6725.

Francis R., West C. & Friedman S.H. (2001). Targeting telomerase via its key RNA/DNA heteroduplex. *Bioorg. Chem.*, **29**, 107-117.

Fry M. & Loeb L.A. (1994). The fragile X syndrome $d(CGG)_n$ nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci. USA*, **91**, 4950-4954.

Fry M. & Loeb L.A. (1999). Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence $d(CGG)_n$. *J. Biol. Chem.*, **274**, 12797-12802.

Ganesh K.N., Rajeev K.G., Pallan P.S., Rana V.S., Barawkar D.A. & Kumar V.A. (1997). Modulation of DNA triplex stability through nucleobase modifications. *Nucleosides Nucleotides*, **16**, 1271-1278.

Gavathiotis E., Gowan S.M., Heald R.A., Stevens M.F.G. & Searle M.S.(2001). Recognition and stabilization of quadruplex DNA by a potent new telomerase inhibitor: NMR studies of the 2:1 complex of a pentacyclic methylacridinium cation with $d(TTAGGGT)_4$. *Angew. Chem. Int. Ed.*, **40**, 4749-4750.

Gee J.E., Yen R.L., Hung M.C. & Hogan M.E. (1994). Triplex formation at the art neu oncogene promoter. *Gene*, **149**, 109-114.

Gehring K., Leroy J.-L. & Guéron M. (1993). A tetrameric DNA structure with protonated cytosine·cytosine base pairs. *Nature*, **363**, 561-565.

Gellert M., Lipsett M. N., Davies D. R. (1962). Helix formation by guanylic acid. *Proc. Natl. Acad. USA*, **48**, 2013-2018.

Giovannangeli C., Thuong N.T. & Hélène C. (1993). Oligonucleotide clamps arrest DNA synthesis on a single-strand DNA target. *Proc. Natl. Acad. Sci. USA*, **90**, 10013-10017.

Giraldo R., Suzuki M., Chapman L. & Rhodes D. (1994). Promotion of parallel DNA quadruplexes by a yeats telomere binding protein: a circular dichroism study. *Proc. Natl. Acad. USA*, **91**, 7658-7662.

Y.N.V. Gopal & M.W. Dyke. (2003). Combinatorial determination of sequence specificity for nanomolar DNA-binding hairpin polyamides. *Biochemistry*, **42**, 6891-6903.

Gowan S.M., Heald R.A., Stevens M.F.G. & Kelland L.R. (2001). Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplex. *Mol. Pharmacol.*, **60**, 981-988.

Gowers D.M. & Fox K.R.(1997). DNA triple helix formation at oligopurine sites containing multiple contiguous pyrimidines. *Nucleic Acids Res.*, **25**, 3787-3794.

Gowers D.M. & Fox K.R.(1998). Triple helix formation at $(AT)_n$ adajcent to an oligopurine tract. *Nucleic Acids Res.*, **26**, 3626-3633.

Griffin L. & Dervan P.B. (1989). Recognition of thymine·adenine base pairs by guanine in a pyrimidine triple helix motif. *Science*, **245**, 967-971.

Griffith J.D., Comeau L., Rosenfield R.M., Stansel R.M., Bianchi A., Moss H. & deLange T. (1999). Mammalian telomeres end in a large duplex loop. *Cell*, **97**, 503 -514.

Grigoriev M., Prauseth D., Robin P., Hemar A., Saison-Behmoaras T., Dautry-Varsat A.,

Thuong N.T., Hélène C. & Harel-Bellan A. (1992). A triple-helix forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor via inhibiton of NKκB binding to interleukin-2 receptor α regulatory sequence. *J. Biol. Chem.*, **267**, 3389-3395.

Guo Q., Lu M., Marky L.A. & Kallenbach N.R. (1992). Interactions of the dye ethidium bromide with DNA containing G repeats. *Biochemistry*, **31**, 2451-2455.

Guo Q., Lu M. & Kallenbach N. R. (1993). Effect of thymine tract length on the structure and stability of model telomeric sequences. *Biochemistry*, **32**, 3596-3603.

Gupta G., Garcia A.E., Guo Q., Lu M. & Kallenbach N. R. (1993). Strucutre of a parallel tetramer of the *Oxytricha* telomeric DNA sequence dT_4G_4. *Biochemistry*, **32**, 7098-7103.

Hahn W.C., Counter C.M., Lundberg A.S., Beijersbergen R.L., Brooks M.W.& Weinberg R.A. (1999a). Creation of human tumour cells with defined genetic elements. *Nature*, **400**, 464-468.

Hahn W.C., Stewart S.A., Brooks M.W., York S.G., Eaton E., Kurachi A., Beijersbergen R.L., Knoll J.H.M., Meyerson M. & Weinberg R.A.(1999b). Inhibition of telomerase limits the growth of human cancer cells. *Nature Med.*, **5**, 1164-1169.

Haider S., Parkinson G.N. & Neidle S. (2002). Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. *J. Mol. Biol.*, **320**, 189-200.

Haider S., Parkinson G.N. & Neidle S. (2003). Structure of a G-quadruplex-ligand complex. *J. Mol. Biol.*, **326**, 117-125.

Hammond-Kosack M.C.U., Dobrinski B., Lurz R., Docherty K. & Kipatrick M.W. (1992). The human insulin gene linked polymorphic region exhibits an altered DNA structure. *Nucleic Acids Res.*, **20**, 231-236.

Han F.X., Wheelhouse R.T. & Hurley L.H. (1999a). Interactions of TMPyP4 and TMPyP2 with quadruplex DNA: structural basis for the differential effects on telomerase inhibition. *J. Am.Chem. Soc.*, **121**, 3561-3570.

Han H., Cliff C.L. & Hurley L.H. (1999b). Accelerated assembly of G-quadruplex structures by a small molecule. *Biochemistry*, **38**, 6981-6986.

Han H., Bennett R.J. & Hurley L.H. (2000). Inhibition of unwinding of G-quadruplex structures by Sgs1 helicase in the presence of PIPER, a G-quadruplex-interactive ligand. *Biochemistry*, **39**, 9311-9316.

Han H., Langley D. R., Rangan A. & Hurley L.H. (2001). Selective interactions of cationic porphyrins with G-quadruplex structures. *J. Am. Chem. Soc.*, 123, 8902-8903.

Hanahan D. & Weinberg R.A. (2000). The hallmarks of cancer. *Cell*, **100**, 57-70.

Haq I., Tent J.O., Chowdhry B.Z & Jenkins T.C. (1999). Intercalative G-tetraplex stabilization of telomeric DNA by a cationic porphyrin. *J. Am. Chem. Soc.*, **121**, 1768-1779.

Hardenbol P. & Van Dyke M.W.(1996). Sequence specificity of triplex DNA formation: Analysis by a combinatorial approach, restriction endonuclease protection selection and amplification. *Proc. Natl. Acad. Sci. USA*, **93**, 2811-2816.

Hardenbol P., Wang J.C. & Van Dyke M.W. (1997a) Identification of preferred hTBP DNA binding sites by the combinatorial method REPSA. *Nucleic Acids Res.*, **25**, 3339-3344.

Hardenbol P., Wang J.C. & Van Dyke M.W. (1997b) Identification of preferred distamycin-DNA binding sites by the combinatorial method REPSA. *Bioconjugate Chem*, **8**, 617-620.

Hardin C.C.,Watson T. , Corregan M.J. & Bailey C. (1992). Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of $d(CGCG_3GCG)$.

Biochemistry, **31**, 833-841.

Hardin C.C., Corregan M.J., Brown B.A. & Frederick L.N.(1993). Cytosine-cytosine$^+$ base pairing stabilizes DNA quadruplexes and cytosine methylation greatly enhances the effect. *Biochemistry*, **32**, 5870-5880.

Hardin C.C., Perry A.G. & White K. (2001). Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers*, **56**, 147-194.

Harrison R.J., Gowan S.M., Kelland L.R. & Neidle S.(1999). Human telomerase inhibition by substituted acridine derivatives. *Bioorg. Med. Chem. Lett.*, **9**, 2463-2468.

Henderson E., Hardin C.C., Walkin S.K., Tinoco I. & Blackburn E.H. (1987). Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine pairs. *Cell*, **51**, 899-908.

Herbert B.S., Pitts A.E., Baker S.I., Hamilton S.E., Wright W.E., Shay J.W., & Corey D.R. (1999). Inhibition of human telomerase in immortal human cells leads to progressive telomere shortening and cell death. *Proc. Natl. Acad. Sci. USA*, **96**, 14276-14281.

Hildenbrandt M., Lacombe M.L., Mesnildrey S. & Veron M. (1995). A human NDP-kinase B specifically binds single-stranded poly-pyrimidine sequences. *Nucleic Acids Res.*, **23**, 3858-3864.

Holt S.E. & Shay J.W. (1999). Role of telomerase in cellular proliferation and cancer. *J. Cell. Physiol.*, **180**, 10-18.

Hoogsteen K. (1959). The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.*, **12**, 822-823.

Howell R.M., Woodford K.J., Weitzmann M.N. & Usdini K. (1996). The chicken β–globin gene promoter forms a novel "clinched" tetrahelical structure. *J. Biol. Chem.*, **271**, 5208-5214.

Huang C.-Y., Bi G. & Miller P.S. (1996). Triplex formation by oligonucleotides containing deoxycytidine derivatives. *Nucleic Acids Res.*, **24**, 2606-2613.

Hud N. V., Smith F.W., Anet F.A.L. & Feigon J. (1996). The selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free enrgies of hydration: a thermodynamic analysis by ^{1}H NMR. *Biochemistry*, **35**, 15383-15390.

Ing N.H., Beekman J.M., Kessler D.J., Murphy M., Jayaraman K., Zendegui J.G., Hogan M.E., O'Malley B.W. & Tsai M.J. (1993). *In vivo* transcription of a progersterone-responsive gene is specifically inhibited by a triplex –forming oligonucleotide. *Nucleic Acids Res.*, **21** 2789-2796.

Ito T., Smith C.L. & Cantor C.R. (1992). Sequence specific DNA purification by triplex affinity capture. *Proc. Natl. Acad. Sci. USA*, **89**, 495-498.

Izbicka E., Wheelhouse R.T., Raymond E., Davidson K.K., Lawrence R.A., Sun D.Y., Windle B.E., Hurley L.H. & Von Hoff D.D. (1999). Effects of cationic porphyrin G-quadruplex interactive agents in human tumor cells. *Cancer Res.*, **59**, 639-644.

James P.L., Brown T. & Fox K.R. (2003). Thermodynamic and kinetic stability of intermolecular triple helices containing different proportions of C^+·GC and T·AT triplets. *Nucleic Acids Res.*, **31**, 5598-5606.

Jin R., Gaffney B. L., Wang C., Jones J. A.& Breslauer K .J. (1992). Thermodynamics and structure of a DNA tetraplex; a spectroscopic and calorimetric study of the tetramolecular complexes of d(TG_3T) and d($TG_3T_2G_3T$). *Proc. Natl Acad. USA*, **89**, 8832-8836.

Jing N., Rando R.F., Pommier Y. & Hogan M.(1997). Ion selective folding of loop domains in a potenty anti-HIV oligonucleotide. *Biochemistry*, **36**, 12498-12505.

Jing N. & Hogan M.E. (1998). Structure-activity of tetrad-forming oligonucleotides as a

potent anti-HIV therapeutic drug. *J. Biol. Chem.*, **273**, 34992-34999.

Joyce G.F. (1994). *In vitro* evolution of nucleic acids. *Curr. Opin. Struct. Biol.*, **4**, 331-336.

Kanazawa Y., Ohkawa K., Ueda K., Mita E., Takehara T., Sasaki Y., Kasahara A. 7 Hayashi N. (1996). Hammerhead ribozyme-mediated inhibition of telomerase activity in extracts of human hepatocellular carcinoma cells. *Biochemistry Biophys. Res. Comm.*, **225**, 570-576.

Kandimalla E.R., Venkataraman G., Sasisekharan V. & Agrawal S.J. (1997). Single stranded DNA and RNA targetd triplex formation: UV, CD and molecular modelling studies of foldback triplexes with different RNA, 2'-OMe-RNA and DNA strand combinations. *J. Biomol. Struct. & Dynam.*, **14**, 715-726.

Kang C., Zhang X., Ratcliff R., Myzis R. & Rich A. (1992). Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature*, **356**, 126-131.

Keniry M.A., Owen E.A. & Shafer R.H. (1997). The contribution of thymine-thymine interactions to the stability of folded dimeric quadruplexes. *Nucleic Acids Res.*, **25**, 4389-4392.

Keppler M.D. & Fox K.R.(1997). Relative stability of triplexes containing different numbers of T.AT and C^+.GC triplets. *Nucleic Acids Res.*, **25**, 4644-4649.

Kerwin S.M., Sun D., Kern J.T., Rangan A. & Thomas P.W. (2001). G-quadruplex DNA binding by a series of carbocyanine dyes. *Bioorg. Med. Chem. Lett.*, **11**, 2411-2414.

Kessler D.J., Pettit B.M., Cheng Y.-K., Smith S.R., Jayaraman K., Vu H.M. & Hogan M.E. (1993). Triple helix formation at distant sites: hybrid oligonucleotides containing a polymeric linker. *Nucleic Acids Res.*, **21**, 4810-4815.

Kiessling L., Griffin L. & Dervan P.B. (1992). Flanking sequence effects within the pyrimidine triple-helix motif characterized by affinity cleavage. *Biochemistry*, **31**, 2829-2834.

Kim N.W., Piatyszek M.A., Prowse K.R., Harley C.B., West M.D., Ho P.L.C., Coivello G.M., Wright W.E., Weinrich S.L. & Shay J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science*, **266**, 2011-2015.

Kinniburgh A.J., Firulli A.B., & Kolluri R. (1994). DNA triplexes and regulation of the *c-myc* gene. Gene, 149, 93-100.

Klug S. & Famulok M. (1994). All you wanted to know about SELEX. *Mol. Biol. Reports*, **20**, 97-107.

Kim M.M., Rivera M.A., Botchkina I.L., Shalaby R., Thor A.D. & Blackburn E.H. (2001). A low threshold level of expression of mutant-template telomerase RNA inhibits human tumour cell proliferation. *Proc. Natl Acad. USA*, **98**, 7982-7987.

Kim M.-Y., Vankayalapati H., Shin-ya K., Wierzba K. & Hurley L.H.(2002). Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J. Am. Chem Soc.*, **124**, 2098-2099.

Knudsen H.& Nielsen P.E. (1996). Antisense properties of duplex- and triplex-forming PNAs. *Nucleic Acids Res.*, **24**, 494-500.

Koeppel F., Riou J.F., Laouii A., Maillet P., Arimondo P.B., Labit D., Petitgenet O., Helene C. & Mergny P.(2001). Ethidium derivatives bind to G-quartets, inhibit telomerase and act as fluorescent probes for quadruplexes. *Nucleic Acids Res.*, **29**, 1087-1096.

Koh & Dervan P.B.(1992). Design of a nonnatural deoxyribonucleoside for recognition of GC base pairs by oligonucleotide-directed triple helix formation. *J. Am. Chem. Soc.*, **114**, 1470-1478.

Kondo S., TanakaY., Kondo Y., Hitomi M., Barnett G.H., IshizikaY., Liu J., Haqqi T.,

Nishiyama A., Villeponteau I., Cowell J.K. & Barna B.P. (1998). Antisense telomerase treatment: induction of two distinct pathways, apoptosis and differentiation. *FASEB J.*, **12**, 801-811.

Kovacs A., Kandala J.C., Weber K.T. & Guntaka R.V. (1996). Triple helix-forming oligonucleotide corresponding to the polypyrimidine sequence in the rat alpha(I) collagen promoter specifically inhibits factor binding and transcription. *J. Biol. Chem.*, **271**, 1805-1812.

Krawczyk S.H., Milligan J.F., Wadwani S., Moulds C., Froehler B.C. & Matteucci M.D. (1992). Oligonucleotide-mediated triple helix formation using an N^3-protonated deoxycytidine analog exhibiting pH-independent binding within the physiological range *Proc. Natl. Acad. Sci. USA*, **89**, 3761-3767.

Ku W.-C., Cheng A.-J. & Wang T.-C.V. (1997). Inhibition of telomerase activity by PKC inhibitors in human nasopharyngeal cancer cells in culture. *Biochemistry Biophys. Res. Comm.*, **241**, 730-736.

Krupp G., Kühne K., Tamm S., Klapper W., Heidorn K., Rott A. & Parwaresch R. (1997). Molecular basis of artifacts in the detection of telomerase activity and a modified primer for a more robust 'TRAP' assay. *Nucleic Acids Res.*, **25**, 919-921.

Lacoste J., Francois J.-C., Hélène C. (1997). Triplex formation with purine-rich phosphorothioate-containing oligonucleotides covalently linked to an acridine derivative. *Nucleic Acids Res.*, **25**, 1991-1998.

Landgraf R., Chen C.B. & Sigman D.S. (1994). Oligonucleotide-directed nucleic acid scission by micrococcal nuclease. *Biochemistry*, **33**, 10607-10615.

Laughton G., Murchie A.I.H., Norman D.G., Moore M.H., Moody P.C.E., Lilley D.M.J. & Luisi B. (1994). The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science*, **265**, 520-524.

LeDoan T., Perrouault L., Prauseth D., Habhoub N., Decout J.L., Thuong N.T., LeHomme L. & Hélène C. (1987). Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-(α)-thymidylate covalently linked to an azidoproflavine derivative. *Nucleic Acids Res.*, **15**, 7749-7761.

Lee J.S., Woodsworth M.L., Latimer L.J.P. & Morgan AR (1984). Poly(purine)·Poly(pyrimidine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. *Nucleic Acids Res.*, **12**, 6602-6614.

Lee J.S. (1990). The stability of polypurine tetraplexes in the presence of mono- and divalent cations. *Nucleic Acids Res.*, **18**, 6057–6061.

Lee H.-W., Blasco M.A., Gottlieb G.J., Horner II J.W., Greider C.W. & dePinho R.A. (1998). Essential role of mouse telomerase in highly proliferative organs. *Nature*, **392**, 569-574.

Lehmann T.E., Greenberg W.A., Liberles D.A., Wada C.K. & Dervan P.B. (1997). Triple-helix formation by pyrimidine oligonucleotides containing nonnatural nucleosides with extended aromatic nucleobases: intercalation from the major groove as a method for recognizing C·G and T·A base pairs. *Helv. Chim. Acta*, **80**, 2002-2022.

Leitner D., Schroder W. & Weisz K.(2000). Influence of sequence-dependant cytosine protonation and methylation on DNA triplex stability. *Biochemistry*, **39**, 5886-5892.

Lew A., Rutter W.J. & Kennedy G.C. (2000). Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effects on transcription by the insulin promoter Pur-1/MAZ. *Proc. Natl. Acad. USA*, **97**, 12508-12512.

Li J.-L., Harrison R.J., Reszka A.P., Brosh R.M., Bohr V.A., Neidle S. & Hickson I.D. (2001). Inhibition of the Bloom's and Werner's syndrome helicases by G-quadruplex interacting ligands. *Biochemistry*, **40**, 15194-15202.

Lin Z.S., Lim S., Viani M.A., Sapp M. & Limm S. (2001). Down-regulation of telomerase activity in malignant lymphomas by radiation and chemotherapeutic agents. *Am. J. Pathol.*, **159**, 711-719.

Liu Z. & Gilbert W. (1994). The yeast KEM1 gene encodes a nuclease specific for G4 tetraplex DNA: implication of *in vivo* functions for this novel DNA structure. *Cell*, **77**, 1083-1092.

Marathias V. M. & Bolton P. H. (1999). Determinants of DNA quadruplex structural type: sequence and potassium binding. *Biochemistry*, **38**, 4355-4364.

Marco-Haviv Y., Baran N. & Manor H. (1999). DNA molecules can drive the assembly of other DNA molecules inot specific four-stranded structures. *J. Mol. Biol.*, **286**, 45-56..

Marfut J., Parel S.P. & Leumann C. J. (1997). Strong, specific, monodentate G-C base pair recognition by N7-inosine derivatives in the pyrimidine.purine-pyrimidine triple-helical binding motif. *Nucleic Acids Res.*, **25**, 1875-1882

Marsh T.C. & Henderson E. (1994). G-wires: self-assembly of a telomeric oligonucleotide, d(GGGGTTTTGGGG), into large superstructures. *Biochemistry*, **33**, 10718-10724.

Matsutomi K., Kaneko S., Hayahi N., Yamashita T., Shirota Y., Kobayashi K. & Murakami S. (2000). Telomerase activity reconstituted *in vitro* with purified human telomerase reverse transcriptase and human telomerase RNA component. *J. Biol. Chem.*, **275**, 22568-22573.

Mazumder A., Neamati N., Ojwang J.O., Sunder S., Rando R.F. & Pommier Y.(1996). Inhibition of the human immunodeficiency virus type 1 integrase by guanosine quartet structures. *Biochemistry*, **35**, 13762-13771.

Mayfield C., Ebbinghaus S., Gee J., Jones D., Rodu B., Squibb M. & Miller D. (1994). Triplex formation by the human Ha-ras promoter inhibits Sp1 binding and *in vitro* transcription. *J. Biol. Chem.*, **269**, 18232-18238.

Mergny J.-L. Sun J.-L., Rougée M., Montenay-Garestier T., Barcelo F., Chomilier J. & Hélène C. (1991). Sequence specificity in triple-helix formation: experimental and theoretical studies of the effect of mismatches on triplex stability. *Biochemistry*, **30**, 9791-9798.

Mergny J.L., Phan A.T. & Lacroix L. (1998). Following G-quartet formation by UV-spectroscopy. *FEBS Letts.*, **435**, 74-78.

Mergny J.L., Lacroix L., Teulade-Fichou M.P. Hounsou C., Guittat, H., Horau M., Arimondo P.B., Vigneron J.P., Lehn J.M., Riou J.F., Garestier, T., Thuong, N.T. & Hélène, C. (2001). Telomerase inhibitors based on quadruplex ligands selected by a fluorescence assay. *Proc. Natl. Acad. USA*, **98**, 3062-3067.

Mergny J.L. & Maurizot J.C. (2001). Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *ChemBioChem*, **2**, 124-132.

McEachern M.J., Krauskopf A. & Blackburn E.H. (2000). Telomeres and their control. *Annu. Rev. Genet.*, **34**, 331-358.

Michel J., Gueguen G., Vercauteren J. & Moreau S. (1997). Triplex stability oligodeoxnucleotides containing substituted quinazoline-2,4(1H,3H)-dione *Tetrahedron*, **53**, 8457-8478.

Michelotti E.F., Tomonaga T., Krutzsch H. & Levens D. (1995). Cellular nucleic acid binding protein regulates the CT element of the human *c-myc* protooncogene. *J. Biol. Chem.*, **270**, 9494-9499.

Michelotti E.F., Michelotti G.A., Aronsohn A.I. & Levens D. (1996). Hetergenous nuclear ribonucleoprotein K is a transcription factor. *Mol. Cell. Biol.*, **16**, 2350-2360.

Michishita E., Nakabayashi K., Ogino H., Suzuki T., Fujii M. & Ayusawa D. (1998). DNA topisomerase inhibitors induce reversible senescence in normal human fibroblasts.

Biochemistry Biophys. Res. Comm., **253**, 1667-671.

Miyoshi D., Nakao A. & Sugimoto N. (2003). Structural transition from antiparallel to parallel G-quadruplex of d($G_4T_4G_4$) induced by Ca^{2+}. *Nucleic Acids Res.*, **31**, 1156-1163.

Morin G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell*, **59**, 521-529.

Moser H.E. & Dervan P.B.(1987). Sequence specific cleavage of double helical DNA by triple-helix formation. *Science*, **238**, 645-650.

Muira T., Benevides J.M. & Thomas G.J. (1995). A phase diagram for sodium and potassium ion control of polymorphism in telomeric DNA. *J. Mol. Biol.*, **248**, 233-238.

Murchie & Lilley (1992). Retinoblasma susceptibility genes contain 5' sequences with a high propensity to form guanine-tetrad structures. *Nucleic Acids Res.*, **20**, 149-53.

Nishikawa N., Oishi M. & Kiyama R.(1995). Construction of a human genomic library of clones containing poly(dG-dA) poly(dT-dC) tracts by Mg^{2+} -dependent triplex affinity capture. *J. Biol. Chem.*, **270**, 9258-9264.

Noonberg S.B., François J.C., Garestier T. & Hélène C. (1995). Effect of competing self-structure on triplex formation with purine-rich oligodeoxynucleotides containing GA repeats. *Nucleic Acids Res.*, **23**, 1956-1963.

Olivas W.M. & Maher L.J. (1995a). Competitive triplex/quadruplex equilibria involving guanine-rich oligonucleotides. *Biochemistry*, **34**, 278-284.

Olivas W.M. & Maher L.J. (1995b). Overcoming potassium-mediated triplex inhibition. *Nucleic Acids Res.*, **23**, 1936-1941.

Ono A., Ts'o P.O.P. & Kan L. (1992). Triplex formation of an oligonucleotide containing 2'-O-methylpseudoisocytidine with a DNA duplex at neutral pH. *J. Org. Chem.*, **57**, 3225-3230.

Orson F.M., Thomas D.W., McShan W.M., Kessler D.J. & Hogan M.E. (1991). Oligonucleotide inhibition of IL2R alpha mRNA transcription by promoter region colinnear triplex formation in lymphocytes. *Nucleic Acids Res.*, **19**, 3435-3441.

Parkinson G.N., Lee M.P.H. & Neidle S. (2002). Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876-880.

Pearson A.M., Rich A. & Krieger M. (1993). Polynucleotide binding to macrophage scavenger receptors depends on the formation of base-quartet-stabilized four-stranded helices. *J. Biol. Chem.*, **268**, 3546-3554.

Pei D., Corey D.R. & Schultz P.G. (1990). Site-specific cleavage of duplex DNA by a semisyntetic nuclease via triple-helix formation. *Proc. Natl. Acad. Sci USA*, **87**, 9858-9862.

Perroualt L., Asseline U., Rivalle C., Thuong N.T., Bisagni E., Giovanngeli C., LeDoan T. & Hélène C. (1990). Sequence-specific artifical photo-induced endonucleases based on triplex-forming oligonucleotides. *Nature*, **344**, 358-360.

Perry P.J., Gowan S.M., Reszka A.P., Polucci P., Jenkins T.C., Kelland L.R. & Neidle S.(1998a). 1,4- and 2,6-Disubstituted amidoanthracene-9,10-dione derivatives as inhibitors of human telomerase. *J. Med. Chem.*, **41**, 3253-3260.

Perry P.J., Reszka A.P., Wood A.A., Read M.A., Gowan S.M., Dosanjh H.S., Trent J.O., Jenkins T.C., Kelland L.R. & Neidle S.(1998b). Human telomerase inhibition by regioisomeric disubstituted amidoanthracene-9,10-diones. *J. Med. Chem.*, **41**, 4873-4884.

Perry P.J., Read M.A., Davies R.T., Gowan S.M., Reszka A.P., Wood A.A., Kelland L.R. & Neidle S. (1999). 2,7-Disubstituted amidofluorenone derivatives as inhibitors of

human telomerase. *J. Med. Chem.*, **42**, 2679-2684.

Phillips K., Dauter Z., Murchie A.I.H., Lilley D.M.J. & Luisi B. (1997). The crystal structure of a parallel-strabded guanine tetraplex at 0.95 Å resolution. *J. Mol. Biol.*, **273**, 171-182.

Porumb H., Gousset H., Letellier R., Salle V., Briane D., Vassy J., Amor-Gueret M., Israel L. & Taillandier E. (1996). Temporary *ex vivo* inhibition of the expression of the human oncogene *HER2 (NEU)* by a triple helix-forming oligonucleotide. *Cancer Res.*, **56**, 515-522.

Postel E.H., Flint S.J., Kessler D.J. & Hogan M.E. (1991). Evidence that a triplex-forming oligodeoxyribonucleotide binds to the *c-myc* promoter in HeLa cells, thereby reducing *c-myc* mRNA levels. *Proc. Natl. Acad. Sci USA*, **88**, 8227-8231.

Radhakrishnan I. & Patel D.J. (1994a). Solution structure of a pyrimidine·purine·pyrimidine DNA triplex containing T·AT, C^+·GC and G·TA triplets. *Structure*, **2**, 17-32.

Radhakrishnan I. & Patel D.J. (1994b). DNA triplexes: solution structures, hydration sites, energetics, interactions, and function. *Biochemistry*, **33**, 11405-11416.

Raghuraman M.K. & Cech T.R. (1990). Effect of monovalent cation-induced telomeric DNA structure on the binding of *Oxytricha* telomeric protein. *Nucleic Acids Res.*, **18**, 4543-4553.

Randazzo A., Galeone A. & Mayol L. (2001).). ^{1}H-NMR study of the interaction of distamycin A and netropsin with the parallel stranded tetraplex $[d(TGGGGT)]_4$. *Chem. Comm.*, 1030-1031.

Rando R.F., Ojwang J., Elbahggar A., Reyes G.R., Tinder R., McGrath M.S. & Hogan M.E.(1995). Suppression of human immunodeficiency virus type 1 activity *in vitro* by oligonucleotides which form intramolecular tetrads. *J. Biol. Chem.*, **270**, 1754-1760.

Rangan A., Fedoroff O.Y. & Hurley L.H. (2001). Induction of duplex to G-quadruplex transition in the *c-myc* promoter region by a small molecule. *J. Biol. Chem.*, **276**, 46400-4646.

Rao T.S., Hogan M.E. & Revankar G.R. (1994). Synthesis of triple helix forming oligonucleotides containing 2—deoxyformycin-A. *Nucleosides Nucleotides*, **13**, 95-107.

Read M.A., Wood A.A., Harrison J.R., Gowan S.M. Kelland L.R., Dosanjh H.S. & Neidle S. (1999). Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: structure-activity relationships. *J. Med. Chem.*, **42**, 4538-4546.

Read M.A. & Neidle S. (2000). Structural characterization of a guanine-quadruplex ligand complex. *Biochemistry*, **39**, 13422-13432.

Read M.A., Harrison J.R., Romagnoli B., Tanious F.A., Gowan S.M. Reszka A.P., W.D. Wilson, Kelland L.R. & Neidle S. ((2001). Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proc. Natl. Acad. Sci. USA*, **98**, 4844-4849.

Ren J. & Chaires J.B. (2001). Rapid screening of structurally selective ligand binding to nucleic acid. *Methods Enzymol.*, **340**, 99-108.

Riou J.-L.,Guittat L., Maillet P., Laoui A., Renou E., Petitgenet O., Mégnin-Chanet F., Hélène C. & Mergny J.-L. (2002). Cell senescence and telomere shortening induced by a series of specific G-quadruplex DNA ligands. *Proc. Natl. Acad. USA.*, **99**, 2672-2677.

Risitano A. & Fox K.R. (2003). Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. *Biochemistry*, **42**, 6507-6513.

Roberts R.W. & Crothers D.M. (1996). Prediction of stability of DNA triplexes. *Proc.*

Natl. Acad. Sci. USA , **93**, 4320-4325.

Rougée M., Faucon B., Mergny J.L., Giovannangeli C., Garestier T. & Hélène C. (1992). Kinetics and thermodynamics of triple-helix formation: effects of ionic strength and mismatches. *Biochemistry*, **31**, 9269-9278.

Roy C. (1994). Triple-helix formation interferes with the transcription and hinged DNA structure of the inerferon-inducible 6-16 gene promoter. *Eur. J. Biochemistry*, **220**, 493-503.

Sarig G., Weismann-Shomer P., Erlitzki R. & Fry M. (1997). Purification and characterization of qTBP42, a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes. *J. Biol. Chem.*, **272**, 4474-4482.

Sasaki S., Nakashima S., Nagatsugi F., Tanaka Y., Hisatome M. & Maeda M. (1995). Design of a novel artificial nucleobase for the selective formation of a triple-complex with a cytosine-guanine base pair. *Tetrahedron Lett.*, **36**, 9521-9524.

Schaffitzel C., Berger I. , Postberg J., Hanes J., Lipps H.J. & Plückthun A.(2001). *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemna* macronuclei. *Proc. Natl. Acad. USA.*, **9**, 8572–8577.

Schouten J.A., Ladame S., Mason S.J., Cooper M.A. & Balasubramanian S. (2003). G-quadruplex-specific peptide – hemicyanine ligands by partial combinatorial selection. *J. Am. Chem. Soc.*, **125**, 5594-5595.

Schultze P., Hud N.V., Smith F.W. & Feigon J. (1999). Refined solution structure of the dimeric quadruplex formed by the *Oxytricha nova* telomere repeat oligonucleotide d($G_4T_4G_4$)*Nucleic. Acids Res.*, **27**, 3018-3028.

Sen D. & Gilbert W. (1988). Formation of parallel four-stranded complexes by guaniine-rich motifs in meiosis. *Nature*, **334**, 364-366.

Sen D. & Gilbert W. (1990). A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature*, **344**, 410-414.

Sen D. & Gilbert W.(1992). Guanine quartet structures. *Curr. Opin. Struct. Biol.*, **1**, 435-438.

Shay J.W.& Bacchetti S. (1997). A survey of telomerase activity in human cancer. *Eur. J. Cancer*, **33**, 787-791.

Shen J., Wang J.C. & Van Dyke M.W. (2001). Identificationof preferred actinomycin-DNA binding sites by the combinatorial method REPSA. *Bioorg. Med. Chem.*, **9**, 2285-2293.

Shin-ya K., wierzba K., Matsuo K., Ohtani T., Yamada Y., Furihata k., Hayakawa Y. & Seto H. (2001). Telomestatin, a novel telomerase inhibitor from *Streptomyces anulatus*. *J. Am. Chem. Soc.*, **123**, 1262-1263.

Siddiqui-Jain A., Grand C.L., Bearss D.J. & Hurley L.H. (2002). Direct evidence of a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA*, **99**, 11593-11598.

Simmonson T., Pecinka P. & Kubista M. (1998). DNA tetraplex formation in the control region of *c-myc*. *Nucl. Acid Res.*, **26**, 1167-1172.

Simmonson T. & Sjöback (1999). DNA tetraplex formation studied with fluoescence resonance energy transfer. *J. Biol. Chem.*, **274**, 17379-17383.

Simmonson T., Pribylova M. & Vorlickova M. (2000). A nuclease hypersensitivity element in the human *c-myc* promoter adopts several distinct i-tetraplex structures. *Biochemistry Biophys. Res. Comm.*, **278**, 158-166.

Simmonson T. & Henriksson M. (2002). *c-myc* suppression in Burkitt's lymphoma cells. *Biochemistry Biophys. Res. Comm.*, **290**, 11-15.

Smith F.W. & Feigon J. (1992). Quadruplex structure of *Oxytricha* telomeric DNA

oligonucleotides. *Nature*, **164**, 164-168.
Smith F.W., Schultze P. & Feigon J. (1995). Solution structures of unimolecular quadruplex formed by oligonucleotides containing *Oxytricha* telomere repeats. *Structure*, **3**, 997-1008.
Smith S. & deLange T. (2000). Tankyrase promotes telomere elongation in human cells. *Curr. Biol.*, **10**, 1299-1302.
Smogorzewska A., van Steensel B., Bianchi A., Oelmann S., Schaefer M.R., Schnapp G., & deLange T. (2000). Control of human telomere length by TRF1 and TRF2. *Mol. Cell. Biol.*, **20**, 1659-1668.
Song C.S., Jung M.H., Supakar P.C., Chen S., Vellanoweth R.L., Chatterjee B. & Roy A.K. (1995). Regulation of androgen action by receptor gene inhibition. *Ann. NY Acad. Sci.*, **761**, 97-108.
Sollogoub M., Darby R.A.J., Cuenound B., Brown T. & Fox K.R.(2002). Stable triple helix formation using oligonucleotides containing 2'-aminoethoxy,5-propargylamino-dU. *Biochemistry*, **41**, 7224-7231.
Stabuli A.B. & Dervan P.B. (1994). Sequence specificity of the nonnatural pyrido[2,3-d] pyrimidine nucleoside in triple helix formation. *Nucleic Acids Res.* **22**, 2637-2642.
Stebbins C.E., Russo A. A., Schneider C., Rosen N., Hartl F. U. & Pavletich N. P. (1997). Crystal structure of an Hsp90–geldanamycin complex: targeting of a protein chaperone by an antitumor agent. *Cell*, **89**, 239-250.
Stilz H.U. & Dervan, P.B. (1993). Specific recognition of CG base pairs by 2-deoxynebularine within the purine·purine·pyrimidine triple-helix motif. *Biochemistry* **32**, 2177-2185.
Sun D., Thompson B., Cathers B.E., Salazar M., Kerwin S.M, Trent J.O., Jenkins T.C., Neidle S. & Hurley L.H. (1997). Inhibition of human telomerase by a G-quadruplex interactive compound. *J. Med. Chem.*, **40**, 2113-2116.
Sun H., Karow J.K., Hickson I.D. & Maizels N. (1998). The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.*, **42**, 27587-27592.
Sun H., Bennett R.J. & Maizels N. (1999). The Saccharomyces cerevisiae *Sgs1* helicase efficiently unwinds G-G paired DNAs. *Nucl. Acids Res.*, **27**, 1978-1984.
Sunquist W.L. & Klug A. (1989). Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature*, **342**, 825-829.
Sunquist W.L. & Heaphy S. (1993). Evidence for interstrand quadruplex formation in the dimerization of the human deficency virus 1 genomic RNA. *Proc. Natl. Acad. USA*, **90**, 3393-3397.
Svinarchuk F., Bertrand J.R. & Malvy C. (1994). A short purine oligonucleotide forms a highly stable triple helix with the promoter of the murine *c-pim-1* proto-oncogene. *Nucleic Acids Res.*, **22**, 3742-3747.
Takasagi M., Guendouz A., Chassignol M., Decoult J.L., Lhomme J., Thuong N.T. & Hélène C. (1991). Sequence-specific photo-induced cross-linking of the two strands of double-helical DNA by a psoralen covalently linked to atriple helix-forming oligonucleotide. *Proc. Natl. Acad. Sci. USA* **88,** 5602-5606.
Teulade-Fichou M.-P., Carrasco C., Guittat L., Bailly C., Alberti P., Mergny J.-P., David A., Lehn J.-M. & Wilson W.D. (2003). Selective recogntion of G-quadruplex telomeric DNA by a bis(quinacridine) macrocycle. *J. Am. Chem. Soc.*, **125**, 4732-4740.
Thomas T.J., Faaland C.A., Gallo M.A. & Thomas T. (1995). Suppression of *c-myc* oncogene expression by a polyamine-complexed triplex forming oligonucleotide in MCF-7 breast cancer cells. *Nucleic Acids Res.*, **23**, 3594-3599.

Tomonaga T. & Levens D. (1996). Activating transcription from single stranded DNA. *Proc. Natl. Acad. USA*, **93**, 5830-5835.

Trapane T.L., Hogrefe R.I., Reynold M.A., Kan L.S., Ts'o P.O.P. (1996). Interstrand complex formation of purine oligonucleotides their nonionic analogs: the model system of $d(AG)_8$ and its complement, $d(CT)_8$. *Biochemistry*, **35**, 5495-5508.

Tu G.C., Cao Q.N. & Israel Y. (1995) Inhibition of gene expression by triple helix formation in heptoma cells. *J. Biol. Chem.*, **270**, 28402-28407.

Tuntiwechapikul W. & Salazar M. (2001). Cleavage of telomeric G-quadruplex DNA with perylene-EDTA·Fe(II). *Biochemistry*, **40**, 13652-13658.

Tyagi S. & Kramer F.R. (1996). Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnol.*, **14**, 303-308.

Vasquez K.M., Narayanan L. & Glazer P.M.(2000). Specific mutations induced by triplex-forming oligonucleotides in mice. *Science*, **290**, 530-533.

Venczel E.A. & Sen D. (1993). Parallel and antiparallel G-DNA structures from a complex telomeric sequence. *Biochemistry*, **32**, 6220-6228.

Von Krosigk U. & Benner S.A. (1995). pH-independent triple helix formation by an ologonucleotide containing a pyrazine donor-donor-acceptor base. *J. Am. Chem. Soc.*, **117**, 5361-5362.

Walsh K. & Gualberto A. (1992).MyoD binds to the guanine tetrad nucleic acid structure. *J. Biol. Chem.*, **267**, 13714-13718.

Wang Y. & Patel D.J. (1993) Solution structure of a parallel-stranded G-quadruplex DNA. *J. Mol. Biol.*, **234**, 1171-1183.

Wang Y. & Patel D. J. (1994). Solution structure of the *Tetrahymena* telomeric repeat $d(T_2G_4)_4$ G-tetraplex. *Structure*, **2**, 1141-1156.

Wang Y. & Patel D. J. (1995). Solution structure of the *Oxytricha* telomeric repeat $d[G_4(T_4G_4)_3]$ G-tetraplex. *J. Mol. Biol.*, **251**, 76-94.

Wang G., Levy D.D., Seidman M.M. & Glazer P.M. (1995). Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation. *Mol. Cell. Biol.*, **15**, 1759-1768.

Wang G., Seidman M.M. & Glazer P.M. (1996). Mutagenesis in mammalian cells induced by triple helix formation and transcription-coupled repair. *Science*, **271**, 802-805.

Weismann-Shomer P. & Fry M. (1993). QUAD, a protein from from hepatocyte chromatin that binds selectively to guanine-rich quadruplex DNA. *J. Biol. Chem.*, **268**, 3306-3312.

Wheelhouse R.T., Sun D., Han H., Han F.X. & Hurley L.H. (1998).Cationic porhyrins as telomerase inhibitors: the interaction of meso-tetra(n-methyl-4- pyridyl)porphine with quadruplex DNA. *J. Am. Chem. Soc.*, **120**, 3261-3262.

Woodford K.J., Howell R.M. & Usdini K. (1996). A novel K^+-dependent DNA synthesis arrest site in a commonly occuring sequence motif in eukaryotes. *J. Biol. Chem.*, **269**, 27029-27035.

Wyatt J.R., Davis P.W. & Freier S.M. (1996). Kinetics of G-quartet-mediated tetramer formation. *Biochemistry*, **35**, 8002-8008.

Xiang G., Soussou W. & McLaughlin L.W. (1994). A new pyrimidine nucleoside ($m^{5oxo}C$) for the pH-independent recognition of G-C base pairs by oligonucleotide-directed triplex formation. *J. Am. Chem. Soc.*, **116**, 11155-11156.

Yoon H.J., Choi I.Y., Kang M.R., Kim S.S., Muller M.T., Spitzner J.R. & Chung I.K. (1998). DNA topoisomerase II cleavage of telomeres *in vitro* and *in vivo*. *Biochim. Biophys. Acta*, **1395**, 110-120.

Zahler A.M., Williamson J.R., Cech T.R. & Prescott D.M. (1991). Inhibition of telomerase

by G-quartet DNA structures.*Nature*, **350**, 718-720.

Zhou-Sun B.W., Sun J.S., Gryaznov S.N., Liquier J., Garestier T., Hélène C. & Tailander E. (1997). A physico-chemical study of triple helix formation by oligodeoxythymidylate with N3'--> P5' phosphoamidate linkages. *Nucleic Acids Res.*, **25**, 1782-1787.

Zimmerman S.C. & Schmitt P. (1995). Model studies directed towards a general triplex DNA recognition system: a novel DNA base that binds a CG base-pair in an organic solvent. *J. Am. Chem. Soc.*, **117**, 10769-10770.