University Of Southampton
Faculty of Medicine, Health and Life Sciences
School of Biological Sciences

Specificity of Triple Helix Formation

By
Antonia Cardew

Thesis for the Degree of Doctor of Philosophy

September 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND LIFE SCIENCES
SCHOOL OF BIOLOGICAL SCIENCES

Doctor of Philosophy

SPECIFICITY OF TRIPLE HELIX FORMATION

By Antonia Cardew

Triplex-forming oligonucleotides (TFOs) have been the subject of extensive research in recent years. They have potential applications in many areas; such as gene-based therapies, site-directed mutation and as biochemical tools. However, triplex technology has been hampered by several problems, including low stability due to electrostatic repulsion between strands. This thesis has investigated combinations of four methods for stabilising triplex DNA; these include incorporation of the positively charged thymine analogues bis-amino-U and propargylamino-dU in TFOs. Also modified TFO's containing anthraquinone derivatives have been tested. Further, the free-intercalating agent naphthylquinoline has been used to modulate TFO binding.

A TFO containing six consecutive BAU molecules has previously been shown to interact with non-target sites. The pH dependence of this TFO was investigated. These experiments showed that considerably higher TFO concentrations were needed to generate a footprint as the pH was increased. The TFO had a high affinity for the exact template (*tyr*T) at pH 5.0 and 6.0 and showed some evidence of binding even at 30 μM at pH 7.0. These gels also showed evidence of the secondary binding seen in previous studies; this was considerably more evident at pH 5.0, however, suggesting that the secondary binding may be more sensitive to pH than the primary binding.

Secondary binding sites for TFOs were examined by 'Restriction Endonuclease Protection, Selection and Amplification' or REPSA. REPSA has been used to select for DNA templates that are bound by the 9mer TFO containing six bis-amino-U residues. Fourteen of the sequences which emerged from REPSA were chosen for footprinting with TFOs containing BAU, propargylamino-dU or T. The BAU-TFO produced clear footprints on all but one of the REPSA templates tested, indicating that the REPSA process was successful in selecting for sequences which are bound by the TFO. Significantly higher concentrations of the P-TFO were required, and magnesium chloride and / or the triplex binding ligand naphthylquinoline were needed to promote binding. Despite the differences in template sequence there does not appear to be a strong pattern in the binding intensities of the TFOs on the different templates. However, all templates do contain a run of four to eight A's. Surprisingly it appears from these data that the BAU TFO discriminates better than the P-TFO against non-exact binding sites

The selectivity of TFOs containing anthraquinone modifications was also investigated. Anthraquinone intercalates between DNA bases in duplex DNA and can be tethered to the end of a TFO to increase stability. The specificity of five TFOs with different anthraquinone modifications was examined by footprinting against fragments containing mismatches. A doubly modified TFO bound with the highest affinity and was most tolerant of mismatches. Mismatches at the centre of the template had a lesser effect on binding affinity than mismatches at the 3' end. The effect of a 3' mismatch was also greater if the anthraquinone was at this end. The presence of an S-base at the 3' end allowing intercalation of the anthraquinone at a YpR step increased the binding affinity on the exact template in comparison to TFO 3 which did not contain the S-base. The TFO containing the S base did not bind quite as well as the doubly modified TFO however.

Acknowledgements

I would like to express my sincerest thanks to Professor Keith Fox for help and guidance throughout my PhD. I would also like to thank post docs, PhD students and technicians past and present; Dave, Phil, Vicki and Dave, from whom I learnt a huge amount and who made the time in and out of the lab extremely enjoyable. Outside of the lab I would like to thank all residents of Cranners over the years (Helen, Giles, Juan, Simon and Pete) also Kath, Kirsty, Tracy and particularly Jo for their constant support and distraction! Finally I would like to thank my parents, without whose amazing support and encouragement I would never have gotten this far. Thank you everyone ☺

Table of Contents

Chapter 1: Introduction

Chapter 2: Materials and methods

Chapter 3: Initial footprinting experiments with *tyr*T DNA show evidence of secondary binding

Chapter 4: Footprinting on REPSA targets

Chapter 5: Selectivity of TFOs containing anthraquinone modifications

Chapter 6: Discussion

List of Figures

List of Tables

DECLARATION OF AUTHORSHIP

I, Antonia Cardew, declare that the thesis entitled Specificity of Triple Helix Formation and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

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

Signed:.....

Date:.....

Abbreviations

A	adenine
AcPrC	N^4-(3-acetamidopropyl)cytosine
AmPyC	N^4-(6-amino-2-pyridinyl)cytosine
^{A}PP	(6-(3-aminopropyl)-7-methyl-3H-pyrrolo[2,3-d]pyrimidin-2(7H0-one))
BAU	bis-amino-U
bp	base pairs
C	cytosine
D_3	(1-(2-deoxy-β-D-ribofuranosyl)-4-(3-benzamidophenyl)imidazole
dATP	deoxyadenosine triphosphate
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease I
ds	double-stranded
E	exact REPSA template
EDTA	ethylenediaminetetracetic acid
G	guanine
IPTG	isopropylthio-β-*D*-galactoside
^{Me}P	3-methyl-2-aminopyridine
mRNA	messenger RNA
N	deoxynebularine
P	propargylamino-dU
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PNA	peptide nucleic acid
R	any purine base
REPSA	restriction endonuclease protection selection and amplification
RNA	ribonucleic acid
S	N-(4-(3-acetamidophenyl)thiazol-2-yl-acetamide
TBE	tris, Borate, EDTA
TFO	triplex-forming oligonucleotide
Tris	tris-hydroxymethyl-aminomethane
T	thymine
U	urasil
Xgal	5-bromo-4-chloro-3-indoyl-β-*D*-galactosidase
Y	any pyrimidine base

Chapter 1: Introduction

The double helical structure of DNA which is so well known today was first elucidated by James Watson and Francis Crick in 1953, supported by Franklin and Gosling's X-ray diffraction image taken the year before (1). DNA has since become the 'central dogma' of molecular biology and has been vital to its progression ever since. Investigations into triplex DNA have opened the door to a new way of manipulating the genome, and this technology is becoming a powerful tool for future applications of molecular biology.

1.1 Triplex DNA

A triple helical DNA structure was first observed in 1957 by Felsenfeld and Rich (2). The idea of triplex DNA had been suggested by Pauling in 1952; however his model was later proven to be incorrect. Felsenfeld and Rich used fibre diffraction and UV melting to show that if polyribonucleotides polyU and polyA were mixed in a 2:1 ratio in the presence of magnesium chloride, a triple helical structure was formed (2). Further experiments showed that a similar structure could be created using polyC and polyG (3). In the years since the initial discovery of triplex DNA, many different structural variants of DNA and RNA triple helices have been shown to exist, including inter- and intra-molecular triplexes in parallel and antiparallel conformations (4-11). The work in this project however will focus on intermolecular parallel DNA triplexes.

Felsenfeld and Rich discovered that a DNA triple helix is created when three strands of DNA come together; two form a classic double helix, consisting of the usual GC and AT base pairs, while the third forms hydrogen bonds with the donor and acceptor groups found in the major grove of the two duplex DNA strands (2).

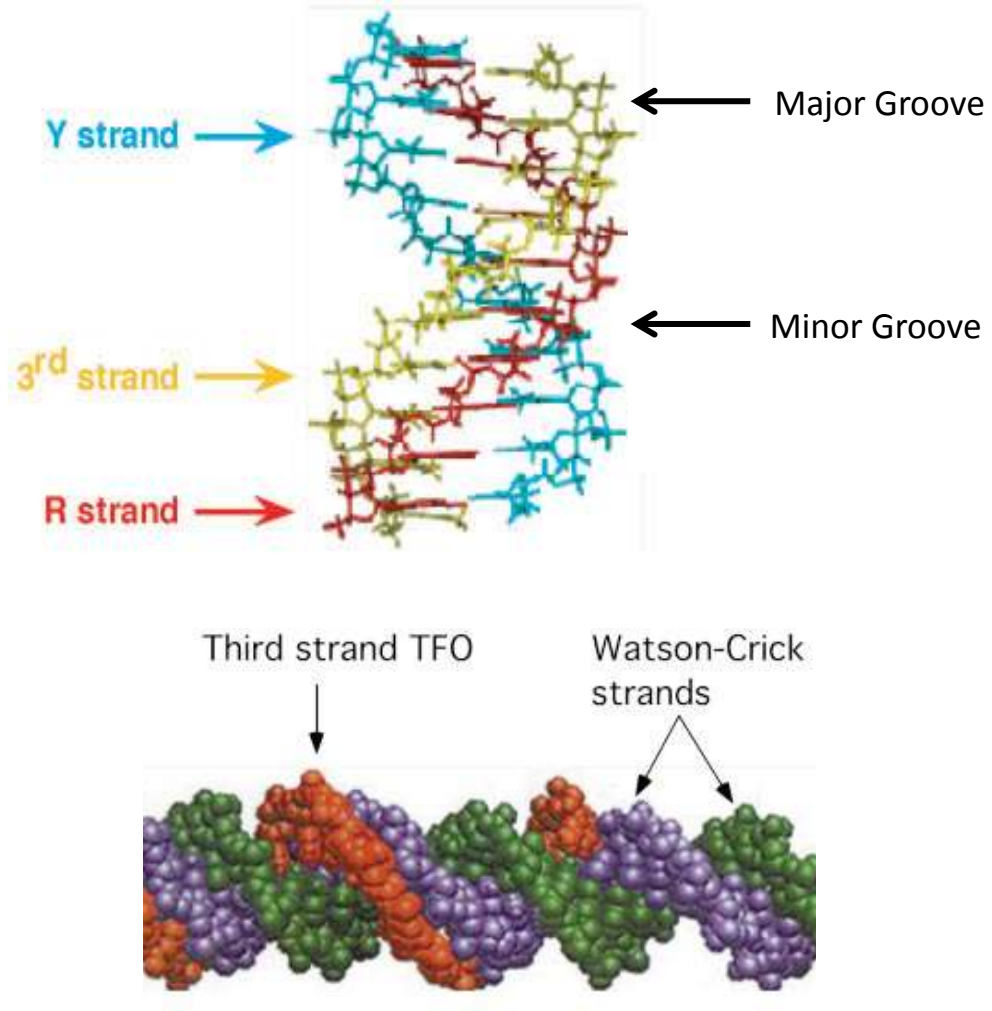


Figure 1.1: Models of a DNA triple helix.
Top: Molecular model of a DNA triple helix. The Y strand is the pyrimidine strand of the duplex DNA template and the R strand is the purine strand. The purine (R) strand is shown bound to the third strand in yellow to form the triplex structure. Adapted from reference (12).
Bottom: Space fill model of an intermolecular triple helix. Adapted from reference (13)

Although DNA triple helices were discovered over 50 years ago it was not until 1987, 30 years after this initial discovery, that the true potential of triplex technology was realised. This was due to the discovery that short oligonucleotides could be used to bind duplex DNA and form a triplex (8). It was immediately obvious that these Triplex Forming Oligonucleotides (TFOs) could be used to inactivate specific genes (14;15). It was also at this point that the idea of H-DNA or intramolecular triplexes first appeared and these were suggested to have a role in gene regulation within cells (16).

Since their discovery it has also been shown that triplexes usually only form with a double helix containing only purines in one strand, with the third strand TFO comprised of either pyrimidines or purines (3;4). The reasons behind this will be covered in more detail later. The base pairs formed between the third strand and the duplex DNA were different from normal Watson-Crick base pairing and these became known as Hoogsteen base pairs (17;18). The pattern of recognition sites in Hoogsteen base pairs differs from Watson-Crick bonding. Hoogsteen bonds form without disrupting the Watson-Crick bonds between the duplex bases, although conformational changes have been shown to occur (13).

1.1.1 Triplex conformations

Triplexes can be divided into two main categories; *inter*molecular and *intra*molecular. An intermolecular DNA triple helix forms when a free strand of DNA (TFO), binds to one strand of a region of duplex DNA (shown in Figure 1.1).

Intramolecular triplexes form when duplex DNA (for example a sequence of genomic DNA) partially unwinds and one strand folds back on itself to form a triple helix leaving an orphaned strand (Figure 1.2 below) (19-21). Intramolecular triplexes can also be generated by synthetic oligonucleotides in which the three strands are connected by loops.

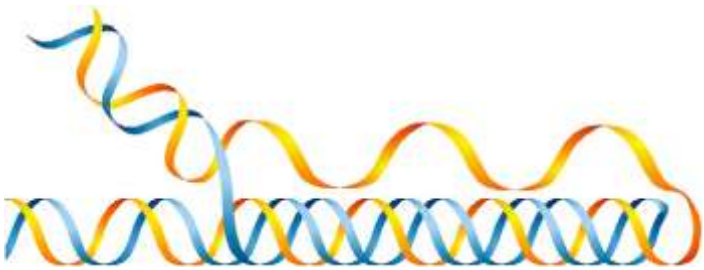


Figure 1.2: An intramolecular triple helix. Taken from reference (22).

Intramolecular triplexes are also known as H-DNA; they are stabilised by H^+ ions, and form most readily at acidic pH and under conditions of superhelical stress (10). This category of triplexes is predominantly restricted to formation within a symmetrical homopurine:homopyrimidine mirror repeat sequence of DNA within plasmids. The most stable natural triplets involve pyrimidines in the third strand binding to purines in the duplex; T binding to an AT pair and C binding to a

GC pair. They were first discovered by Mirkin *et al* in 1987; thirty years after the initial discovery of intermolecular triple helices in 1957 (20).

These triplexes are generated under superhelical stress within plasmids, though they have been postulated to form in genomic DNA where they cause nuclease hypersensitive sites and are often associated with the regulatory regions of genes (16;23-25). As triplex formation requires supercoiling it has been considered possible that intramolecular triplexes could have a role in transcription as RNA polymerase creates regions of supercoiled DNA (26).

When part of the duplex DNA unravels to form the intramolecular triplex it leaves the other strand unpaired and this is then sensitive to S1 endonucleases and other chemical agents (27;28). S1 sensitivity has been used as a tool to study triplexes *in vivo* (19;29;30), and S1 endonuclease hypersensitive sites have been discovered in the upstream regions of several eukaryotic genes, indicating possible triplex formation *in vivo* (22).

There is a growing amount of evidence (31) to show that H-DNA plays an important role in regulation of genes and causes mutations *in vivo*. There are many theories as to how this regulation occurs but it is unfortunately very difficult to studies these processes directly (22). Although there are now several methods for identifying intramolecular triplexes *in vitro* it is very hard to detect them *in vivo* (32). Monoclonal antibodies have been used to probe for intramolecular triplex *in vivo* (33;34). Also the single strand of DNA exposed when an intramolecular triplex forms can potentially be bound by a fluorescently-tagged complementary single strand of DNA (35). Both of these techniques have shown similar results but were not performed under physiological conditions; it remains difficult to detect triplex formation in living cells and it is still an open question as to whether they form *in vivo*.

When a TFO binds to a duplex target it forms hydrogen bonds with one strand of the duplex. This can be achieved in one of two orientations; parallel (14;36) or antiparallel (Figure 1.3) (37). This classification refers to the direction of the TFO in relation to the strand that it’s binding to. This dictates the nature of the triplets that can form in each orientation (14).

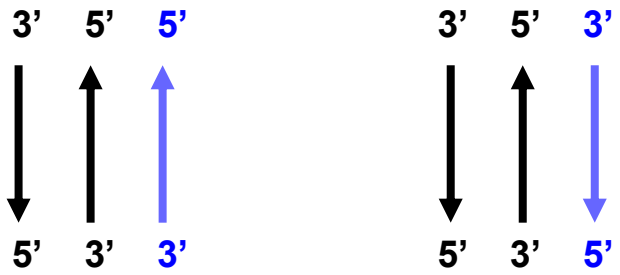
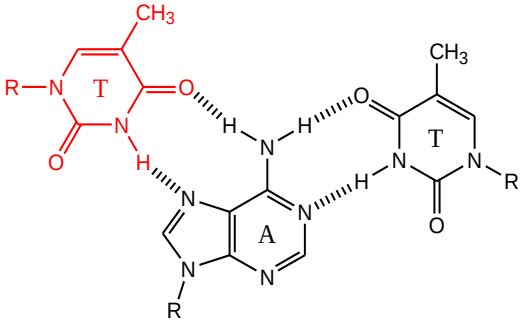
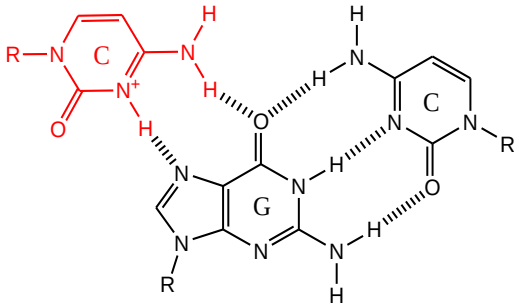


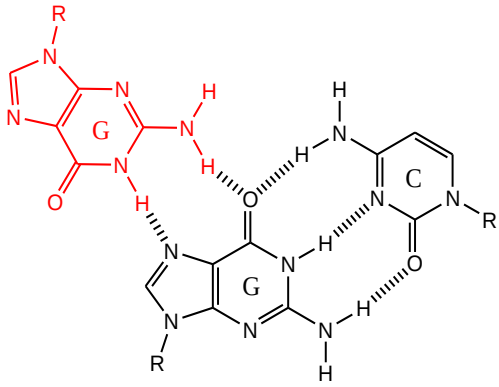
Figure 1.3: Representation of Antiparallel vs. Parallel intermolecular triplexes. The third strand is shown in blue, and binds to the immediately adjacent strand which is purine rich.



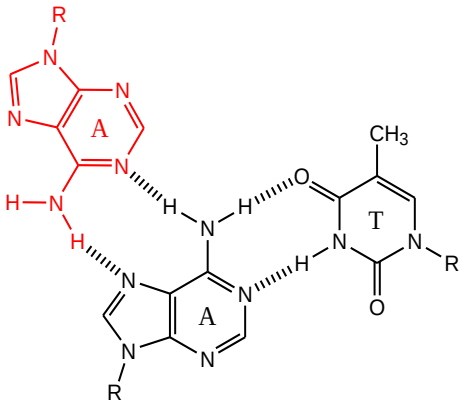
T.AT triplet (parallel)



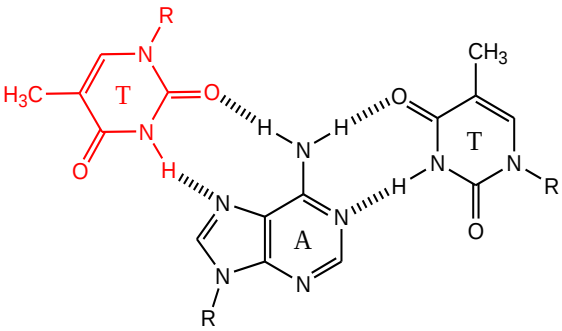
C $^{+}$.GC triplet (parallel)



G.GC triplet (antiparallel)



A.AT triplet (antiparallel)



T.AT triplet (antiparallel)

Figure 1.4: The most stable naturally forming parallel and antiparallel triplets. Third strand bases in red.

As is shown in Figure 1.4 the third strand base only forms hydrogen bonds with one of the two duplex strands, the purine strand. The strand that the TFO binds to therefore becomes the central strand of the triplex, so one side of the base binds the TFO and the other side binds the other duplex strand. Binding to the TFO therefore involves a side of the central base which is not normally involved in hydrogen bonding and thus creates a different hydrogen bonding pattern from that found in duplex DNA.

The most stable naturally-forming parallel triplets are T.AT (or U.AT) and C^+.GC; T (or U in RNA) in the TFO binds to A of AT in a duplex sequence, and C^+ binding to G of GC (5;7;14;38). The third strand cytosine has to be protonated in order to form a second bond with guanine, and this will be covered in more detail later. T.AT formation is favoured by high ionic strength and divalent cations, particularly magnesium; it is a pH neutral triplet (39). Much less stable triplets formed using natural DNA bases include G.TA and T.CG; these have very different bonding patterns to the more stable triplets (See Figure 1.5 below) (40-42).

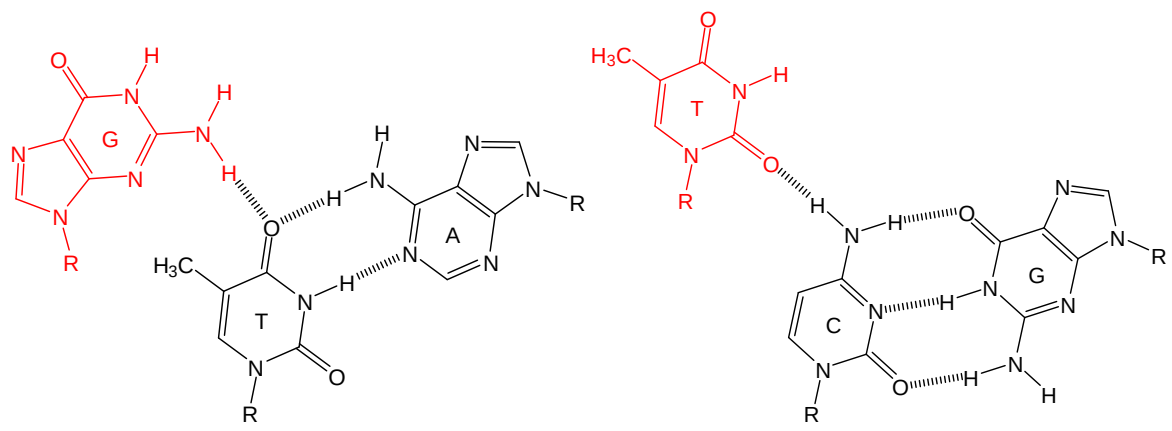


Figure 1.5: Structure of G.TA and T.CG triplets. Third strand bases in red.

G.GC and A.AT are the most stable natural antiparallel triplets, along with T.AT which can also form in this orientation (43;44). In the antiparallel conformation the TFO thymine presents a different face for binding to adenine (45). When intermolecular triplexes form in the antiparallel orientation, the hydrogen bonds are referred to as *reverse-Hoogsteen* bonds.

Unlike parallel triplex formation a TFO composed of G and T bases binds in a relatively pH-independent fashion and the resultant complex is generally more stable than an equivalent parallel triplex, particularly at physiological pH (3;45-49).

The kinetics of antiparallel triplex formation also appears to be faster (50-53). The stability of antiparallel triplexes varies widely however and is highly sequence dependent, partially due to the different triplets not being isohelical, resulting in local distortions in the TFO backbones (45).

G-rich TFOs appear to form the most stable natural antiparallel triplexes, although they have a high tendency to form compact secondary structures such as G quadruplexes (13;54;55). Also GA-rich TFOs can form homoduplexes which are stabilised by magnesium ions (56-58). Both of these alternative structures can inhibit triplex formation (55;57).

In common notation triplets can also be classed by the composition of the third strand; pyrimidine-rich TFOs form Y.RY triplets (usually parallel), while purine-rich TFOs form R.RY triplets (usually antiparallel). In each case the first letter (Y for pyrimidine and R for purine) represents the TFO base, the next is the base in the duplex strand that the TFO binds to (usually a purine), and the third letter is the pyrimidine in the other strand of the duplex. As previously mentioned, triplex formation generally requires a homopurine:homopyrimidine stretch of duplex DNA, either for a pyrimidine TFO to bind to, or to form H-DNA.

1.1.2 Pyrimidine targeting

Pyrimidines in duplex DNA cannot easily be targeted by third strands containing unmodified bases (37). This is because any natural bases binding to pyrimidines in the target duplex can form only one hydrogen bond rather than two:

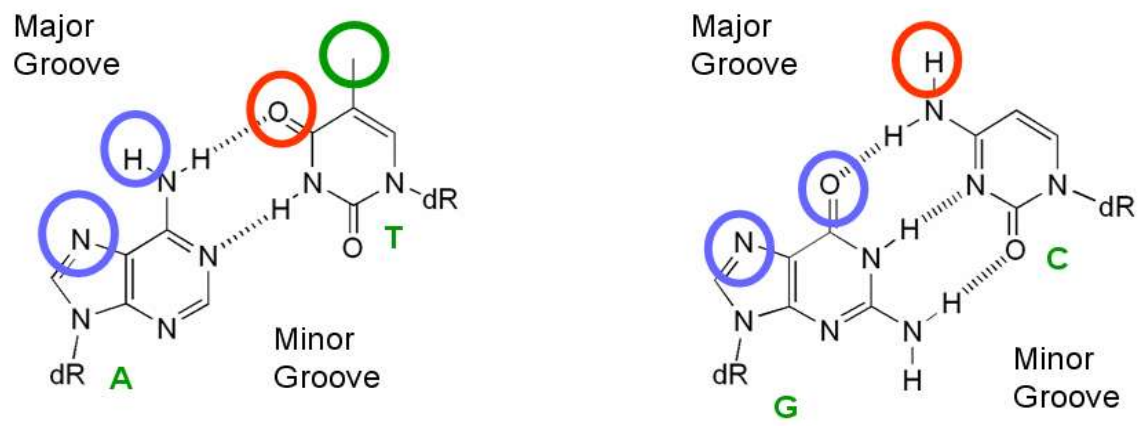


Figure 1.6: Standard duplex AT and GC pairing. Blue circles indicate hydrogen bond acceptors and donors on the purine bases, red show those on the pyrimidine bases. While there are two possible acceptors/donors on both of the purine bases there is only one on each of the pyrimidines. The green circle indicates a methyl group which interferes with recognition of thymidine by a TFO.

As shown in Figure 1.6 above the inability to target pyrimidines in the duplex DNA is due to a lack of hydrogen bond donors and acceptors on the duplex pyrimidine and steric hindrance from the methyl group on T. This has presented problems when using TFOs targeted to mixed purine/pyrimidine sequences (37). The issue of pyrimidine recognition will be covered in more detail later.

1.1.3 Structure

To obtain more knowledge about the exact structure of triplexes so that new nucleotide analogues could be designed, various NMR and X-ray crystallographic studies have been carried out (59-65). The first X-ray crystal structure was finally attained in 1999 (65).

It was initially thought that triplex formation caused the underlying DNA duplex to adopt an A type conformation, in contrast to normal double stranded DNA which is B-form (17). However, it was later shown that triplex DNA has a completely separate conformation, which contains aspects of both A- and B-form, but is distinct from both (Figure 1.7) (65-68).

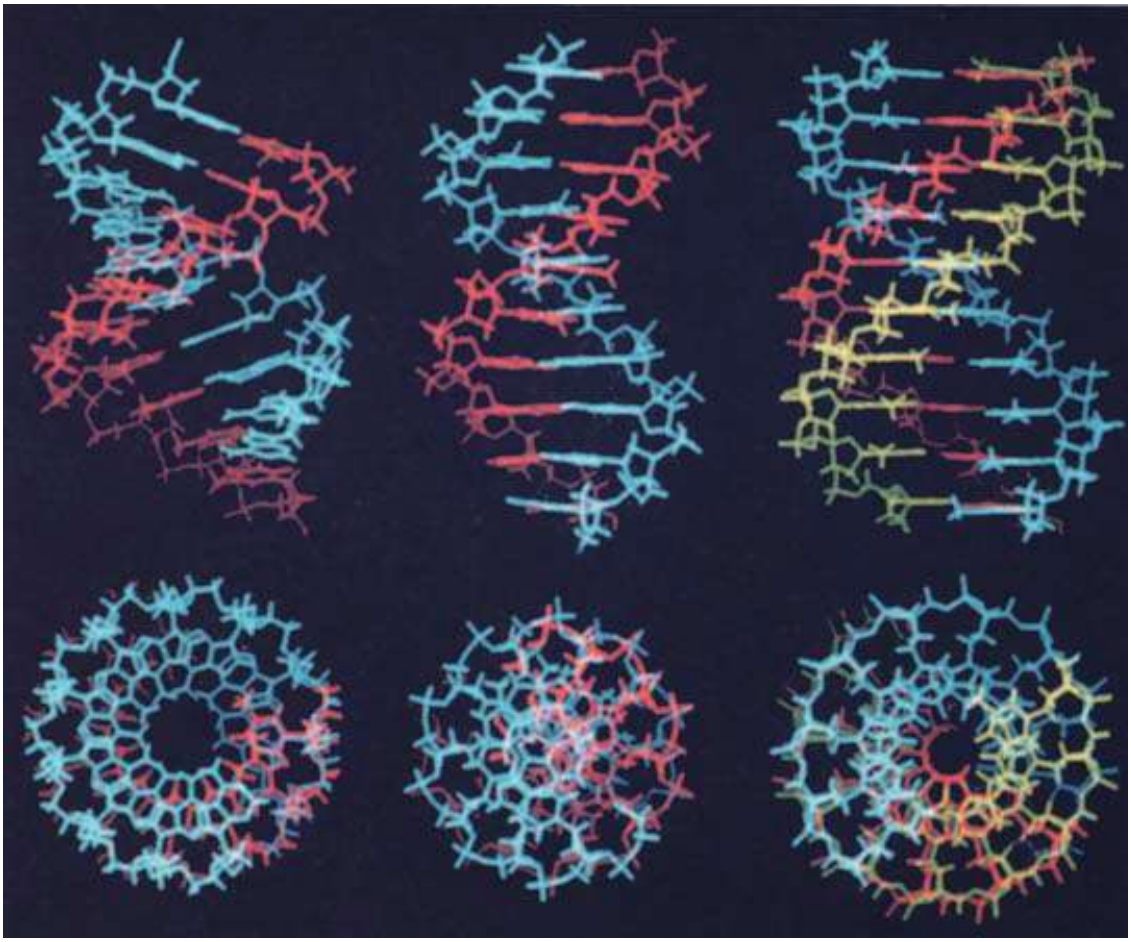


Figure 1.7:
Top: Side on view, Bottom: top down view
Left: A-form DNA (60)
Middle: B-form DNA (60)
Right: Parallel triplex structure, third strand shown in yellow
Based on X-ray diffraction studies, taken from reference (69)

To bind a third strand of DNA, the duplex template must unwind slightly to allow the third strand to fit within the major groove (the increase in helical twist is 31º). This causes the Watson-Crick bases to shift towards the minor groove, leading to nuclease hypersensitive sites appearing at the triplex/duplex junctions (42;70-72). It also causes most of the sugars in the double stranded DNA to adopt an S-type pucker (42). Despite this however the transition from duplex to triplex structure is relatively smooth (65).

1.1.4 DNA grooves

The binding of a third DNA strand into the major grove of a duplex also causes the formation of a third groove in the DNA complex. The three grooves are therefore re-named (see Figure 1.8 below); the minor groove from the original duplex (between the two duplex strands) becomes the Watson-Crick groove, the one between the central purine strand and the TFO becomes the Crick-Hoogsteen groove and between the TFO and the far pyrimidine strand the Watson-Hoogsteen groove (42;73).

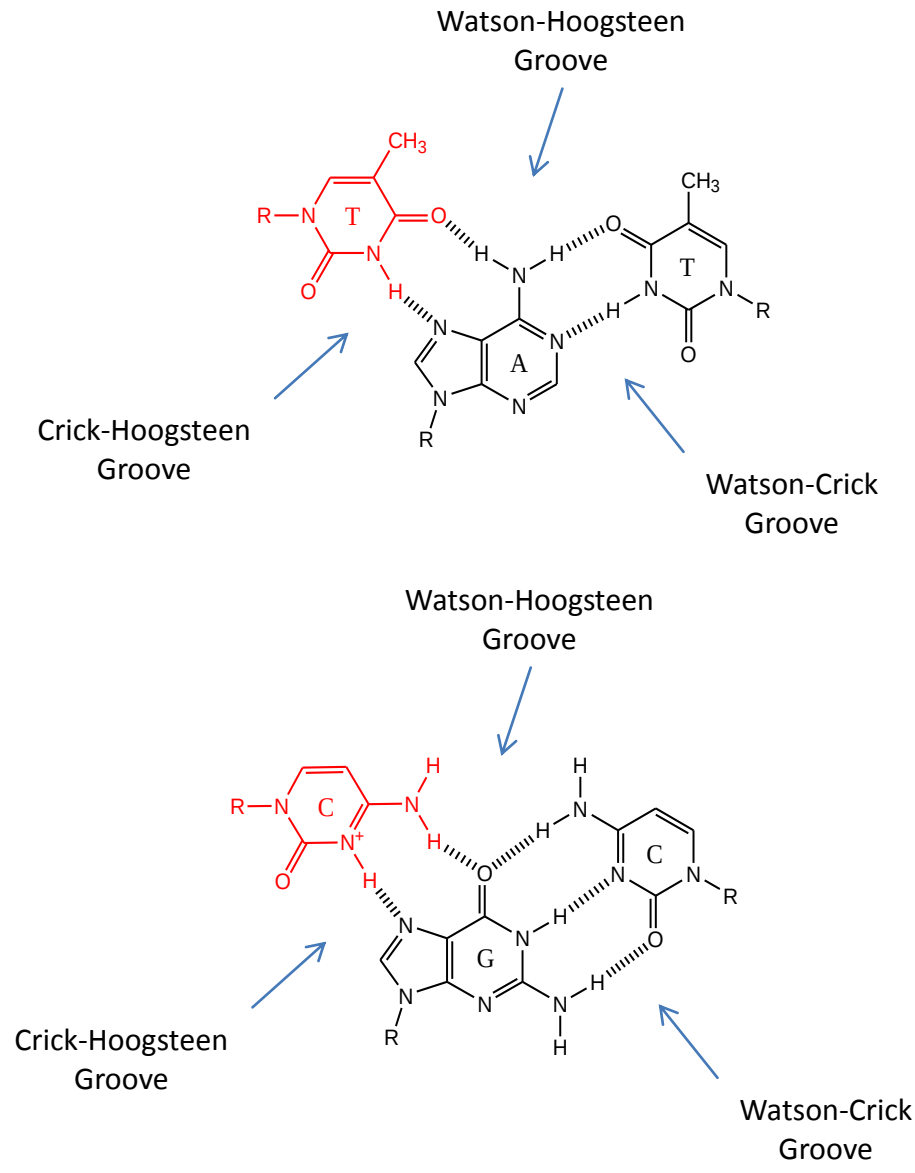


Figure 1.8: T.AT and C^+.GC triplets showing the naming of the three grooves.

In a parallel triplex the Watson-Crick groove becomes slightly narrower upon TFO binding as the duplex major groove becomes wider to accommodate the third strand; in a typical triplex the Watson-Crick groove is around 6-7Å, slightly smaller than the normal minor groove of B-DNA. However, the character of the Watson-Crick groove in triplex DNA has been shown to be very different to that of the normal minor groove, although it is still more similar to B- form than A-form DNA (42). The Crick-Hoogsteen groove is also relatively narrow (2-3Å), but the Watson-Hoogsteen groove is quite wide (7-8 Å) and is highly hydrophobic due to the methyl groups of thymines in the TFO and duplex DNA (42;73).

1.1.5 Phosphate groups

All the triplets mentioned so far are much weaker than natural duplex base pairs. The stability of triplexes is affected by several factors such as pH and ionic conditions, as well as the length of the third strand itself. One of the main contributors to the low stability of triple helices is the negatively charged phosphate backbone of the DNA strands.

In duplex DNA the two strands are spaced such that energetic repulsion between the phosphate groups in the backbones is minimised. The negative charges are also counteracted partially by cationic screening. However, when a TFO binds in the major groove of the duplex it forces the phosphate groups of its backbone into close proximity with those in the duplex. Figure 1.7 on page 9 gives an indication of the different interactions of the duplex and triplex strands. High concentrations of monovalent or divalent cations are often used to stabilise triplexes as they increase the ionic screening around the phosphates (14;39;74).

1.1.6 Base stacking

When considering the stability of triplex DNA molecules it is also important to take into account the effects of base stacking interactions. This type of interaction is due to Van der Waals and hydrophobic bonding between the planar rings of the bases. The alignment of the pyrimidine and purine rings above and below one another has a large effect on stability and any perturbations which disturb this alignment can cause the triplex to become unstable. T.AT and C^+.GC are isohelical in parallel triplexes but the position of the backbone varies from other triplets and so will cause backbone distortions if different bases are used

(75). Stacking interactions are also an important consideration when designing new base analogues as they can dramatically reduce or increase the stability of the overall structure.

Several groups have attempted to increase the base stacking potential of nucleotides by increasing the number of aromatic rings in the base. For example pyrido[2,3-*d*]pyrimidine and its derivatives have been used for AT recognition (76). However it was demonstrated that when incorporated into a TFO, this analogue was no more stable than natural nucleotides (76). Increasing base stacking therefore isn't necessarily the best way to improve triplex stability, as extended ring systems may disrupt the structure or desired base overlap (45).

1.1.7 Length

The chain length of the TFO affects triplex stability. Generally, in parallel triplex formation the longer the third strand the more stable the interaction with the duplex; though this also depends on the base composition of the TFO. A higher proportion of more stable triplets within a DNA triplex will normally lead to more stable overall binding (45;77). For example a short triplex comprised solely of T.AT triplets might be very unstable, whereas a triplex of the same length containing one or two C^+.GC triplets would have a higher stability at low pH. However once too many C^+.GC triplets are introduced it becomes unstable again as contiguous C^+.GC triplets are destabilising (30;78-81).

It has also been shown that longer TFOs are more likely to bind to secondary binding sites as they have a greater potential to form intrachain loops and bulges (82). When considering targeting sites within the human genome, TFOs need to be at least 16-17 bases in length in order to recognise a unique sequence (82). Any TFOs shorter than this are likely to have more than one binding site, due to the length of the human genome (around 3 billion bases). Stable triplexes have been formed with TFOs as short at 9 bases however (83;84).

1.1.8 Mechanisms of triplex formation

The formation of a DNA triple helix most likely occurs via a nucleation-zipper mechanism (85). It is thought that 2 to 3 bases at the 3' end of the TFO first associate and dissociate rapidly with random sites in the duplex DNA due to an unfavourable equilibrium until they find a complementary sequence. At this point if

the TFO has bound to the exact recognition site there is the chance for another nucleotide to bind the template. If this happens then the nucleation point has formed and the rest of the TFO can then bind by 3-D wrapping in a zipper type movement. However, if the sequence is incorrect the TFO dissociates and continues to create nucleation points until the exact site is bound (45;74). This mechanism explains why triplex formation is faster at lower temperatures, since this stabilises the transient intermediates. Triplex formation is about three orders of magnitude slower than duplex formation however (53;74;81).

Electrostatic repulsion between the phosphate groups on the three strands plays a major role in the formation of the nucleation point. Once a nucleation point is formed, base stacking interactions stabilise the complex and allow zippering (74). Triplex formation also involves a conformational change in the duplex, the energetics of which must be overcome in order for the triplex to form (59). This means that more bases are thought to be involved in the nucleation point of triplex formation compared with duplex DNA (74).

1.2 Potential applications of triplexes

Despite the discovery of triple helices by Gary Felsenfeld in 1957 it wasn't until 30 years later that their true potential was realised. Around this time it was found that TFOs could be used to 'read' duplex DNA and this recognition specificity renewed interest in triplex DNA technology (9;45). A potential biological role for triplexes was first observed in 1968; this was an *in vitro* assay looking at inhibiting transcription by RNA polymerase using an RNA TFO (4). In 1994 it was shown that if a homopurine:homopyrimidine sequence that was capable of forming an intramolecular triplex was inserted into SV40 it caused fork stalling *in vivo,* demonstrating the importance of triplexes for biological purposes (86).

Polypurine tracts in the human genome capable of forming intramolecular triplexes have been linked with disease causing genetic mutations; indicating that triplex formation alone could be enough to cause mutation of the target sequence (87-90). It appears that triplex formation may cause double strand breaks (DSBs) in the template DNA and it is during repair of these that mutations occur (91). It is still unclear however exactly how these DSBs are created (22).

Potential binding sites for TFOs have been found to be more common than expected in the human genome and particularly in promoter sequences (25;92;93). They were also found to be more common in genes involved in cell signalling and cell communication (94). The genes containing these sequences are also more susceptible to chromosomal translocations, more likely to undergo alternative splicing and are expressed at relatively low levels (94). It's therefore thought that intramolecular triplexes might play a role in transcription (21;47).

The use of TFOs in gene therapy has several advantages over other methods; they can be extremely specific and easily designed providing the DNA sequence of interest is known and there are fewer binding sites per cell than if RNA or protein species were being targeted, allowing lower concentrations of the TFO to be used (45). The high specificity and potentially low concentrations needed for action also make toxic side effects *in vivo* less likely. Potential triplex formation sites are also often associated with high numbers of SNPs, opening up the possibility for individually-tailored therapies (25). It has been repeatedly suggested that triplexes could be used to target one specific site in the entire human genome (95;96).

This section looks at the different areas where triplex DNA is being utilised, as well as the many challenges involved with this technology.

1.2.1 Anti-gene technology

Following the realisation that triplex forming oligonucleotides could interact with DNA *in vivo*, possibilities opened up for using TFOs to block the activation of specific genes by inhibiting protein-DNA binding (49;97-99). Work has been carried out with the aim of blocking transcription, and hence protein production, this is called the anti-gene method. In anti-gene technology a TFO is used to bind part of a gene such as the promoter region and thus prevent transcription (see Figure 1.9 below) (100). For example a TFO could be targeted to part of a promoter or operator sequence containing the binding site for a protein involved in transcription, such as RNA polymerase or a transcription factor. When the TFO binds to this sequence it would compete for and possibly inhibit binding of the protein and therefore prevent transcription of the gene (47;101-103). There can be several problems with this method, however. For example, the DNA ,may be inaccessible or the TFO may be degraded before reaching the target (see section 1.2.4) (25). This method was used in a study in 2004 to reduce the size of tumours in mice by injecting TFOs targeted to the p53 gene (104).

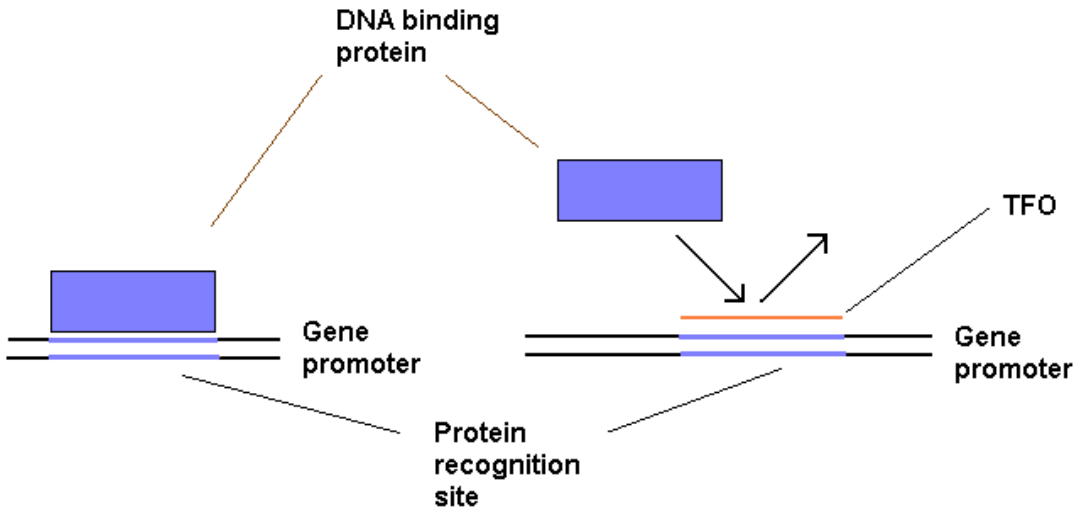


Figure 1.9: DNA/protein binding inhibition via the formation of a DNA triple helix. The left side shows a protein binding to DNA in the normal way, on the right a TFO with a sequence complementary to the protein recognition site has bound to the promoter and prevented the protein from interacting with the DNA.

As well as preventing the binding of proteins to DNA, TFOs can also be used to inhibit movement of proteins along DNA; for example inhibiting replication

by blocking the progression of polymerases (105). TFO binding has also been shown to provide a very strong block to Taq DNA polymerase and RNA Polymerases; this technique has been used to successfully inhibit the production of the c-MYC oncogene *in vivo* (106;107).

1.2.2 Oligonucleotide-clamp technology

A further method for gene inhibition using TFOs is a variation of anti-sense technology known as the oligo-clamp method. A single-stranded homopurine DNA or RNA molecule is targeted by a TFO comprised of 2 strands which binds to both sides of the single strand forming a local, short triple helix (69). This single stranded target could potentially be an mRNA molecule or part of a viral genome (108). The targeted sequence becomes the central strand of the triplex, forming both Watson-Crick and Hoogsteen bonds with the 2 strands of the TFO; see Figure 1.10 below. The linker between these 2 strands can either be a short nucleotide sequence or a flexible synthetic linker such as hexaethylene glycol (69). The triplex formed is more stable than a duplex sequence of the same length and has been shown to successfully inhibit peptide chain elongation (108)

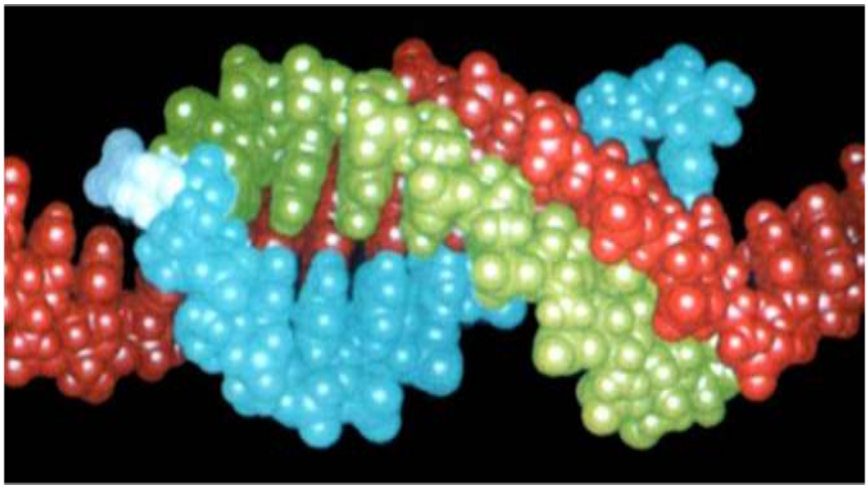


Figure 1.10: Space-filling model of an oligo-clamp TFO (blue and green strands joined by a hexaethylene glycol linker in white) binding to a single strand (red). First the blue strand forms Watson-Crick base pairs with the target single strand while the green strand becomes the third strand and forms Hoogsteen bonds with the red central target strand. Taken from reference (69).

This type of triplex formation is more sensitive to mismatches than anti-gene technology as each target base is recognised on both sides by the TFO, rather than just one. Joining the 2 strands of the TFO together rather than using separate strands can help to increase the stability of the complex, joining them at both ends to form a circular molecule increases stability even further, although this is more difficult to synthesise (69). The covalent attachment of ligands to these complexes can also be used to further increase triplex stability. Psoralen has been used in this way to 'lock' the triplex structure as it cross-links with both the target and TFO strands (108). Additionally these types of TFOs are less prone to aggregation within cells, which has be found to be a problem with standard TFOs (109).

1.2.3 Other applications

As well as blocking transcription directly or using targeted oligo-clamp TFOs, there are several other ways triplex technology may be used. To improve the likelihood of inactivating a gene various ligands and proteins can be attached to TFOs to target them to a particular site on a duplex DNA template. In this way artificial endonucleases can be created by attaching nucleases to TFOs; these complexes can then be used to cleave at very specific sequences and this cleavage can lead to a high frequency of mutations during the repair process (25;36;95;96;109-112). Unlike simply using a TFO to block transcription, this technique offers the advantage of permanently inactivating a gene and passing on the mutation.

Addition of chemical groups such as psoralen allows careful timing of the mutagenic effect; i.e. after the TFO has bound its target, to minimise non-specific DNA damage. Psoralen intercalates with the duplex DNA and causes mutations, again at very specific sites (113). In 2000 this technique was used to introduce mutations into the genomic DNA of mice, demonstrating the effectiveness of this technology in whole organisms (114). Anti-cancer drugs such as daunomycin have also been attached to TFOs. Daunomycin intercalates with the DNA increasing stability although it does not appear to retain any anti-cancer activity (115;116).

A different approach is to take advantage of potential intramolecular triplex formation sites in genes of interest. These sites are common in promoter regions and are inherently recombinogenic and mutagenic; they may be able to be

exploited in a wide variety of ways to control gene expression (22). The frataxin gene which is involved in Friedreich's ataxia contains an H-DNA formation site which is known to affect protein production when a triplex is formed (117). Ataxia is caused by transcriptional silencing of this gene which codes for the 210 amino acid frataxin (118). This silencing is caused by expansion of a 5' – GAA – 3' triplet repeat which can then form a disease promoting intramolecular triplex and prevent transcription (119). Destabilising this triplex could therefore potentially allow transcription to resume, providing an attractive therapy for this disease (118).

Recently it has been shown that potential triplex DNA sequences are over represented in genes expressed in the brain, which has led to the theory that intramolecular triplex formation may play a role in the development of schizophrenia (94;120). Interfering in some way with these triplex structures could therefore have an effect on the brain.

1.2.4 Problems

There are several problems that need to be overcome before the above triplex technology can become viable. The main problems are: delivery and uptake to cells; stability within cells; secondary structure formation of TFOs; target binding affinity; displacement by cell components and inaccessibility of the target DNA. TFOs have also been found to have unpredictable effects within , such as binding transcription factors and therefore not binding the target DNA (121).

In vitro studies have shown that TFO binding can prevent transcription and that a triplex has comparable – and in some cases greater – stability than the transcription factor/DNA complex (12;122). This can also be used to activate genes, for example if an inhibitor binding site is targeted. Another approach is to attach proteins to the TFO such as transcription factors or inhibitor which are then targeted to a particular gene (123). Some success at applying triplex technology *in vivo* has been documented, such as successfully injecting TFOs into mice, causing mutagenesis in specific genes (114).

Endonucleases that act from ends or within DNA chains are prevalent in cells and these can rapidly degrade unmodified TFOs. Oligonucleotides are also vulnerable to extracellular degradation (124;125). Many nucleotide analogues have been synthesised, in order to increase specificity or selectivity whist making the TFO more nuclease resistant (45). Some of these modifications such as backbone charge reduction also help to improve uptake of the TFO by cells

(13;126;127). Other modifications have been investigated to help prevent TFOs from adopting secondary structures such as G-quadruplexes within cells (55;128;129).

Because TFOs are polyanions they cannot easily penetrate the lipid bilayer and enter cells, although there is evidence that oligonucleotides can reach their targets within the nucleus, possibly by receptor-mediated endocytosis (130;131). This process is very inefficient however and may be subject to cell-type specificity (124). Viral delivery has been a popular technique used in clinical trials to overcome this problem although the body's immune response to these is still an issue (132). There are several reagents which have been found to improve cellular uptake of TFOs by cells. These include: cationic lipids, cationic polymers, cell penetrating peptides and polycations such as polyethylenimine. However these are not usually suitable for *in vivo* work (12;22;133-135). One technique which has shown to be successful *in vitro* and *in vivo* is coupling TFOs to hydrophobic molecules like cholesterol since this aids passage through the nuclear membrane (136).

Nucleotide analogues have also been synthesised to help with this problem such as polyamine analogues, cholesterol derivatives, nuclear targeting peptide conjugates, 6-phosphate-bovine serum albumin (BSA) and polypropylenimine dendrimers (136-140). Some of these analogues may also allow the TFO to compete with chromatin and therefore access previously unavailable DNA targets. DNA packaging is a major obstacle for TFO technology to address as the target may be inaccessible when bound to chromatin proteins (12;141;142).

One potential solution which has been investigated to improve TFO delivery is using in-vivo cell transcription to synthesis the TFO. This would overcome problems involved with immunogenicity, uptake and dose, as well as the expense of synthesising large quantities of oligonucleotides (124). PNA (peptide nucleic acid) and LNA (locked nucleic acid) structures have also been successfully used to overcome many of these problems (143-154). These will be covered in more detail later.

1.2.5 TFOs as biochemical tools

As well as looking at targeting genes triplex technology has also been used to create tools for molecular biology. They can also be used for labelling DNA or as sensors or to target particular molecules to DNA sites for example. TFOs have

been used in various new biochemical assays that for example look at translocation of proteins along DNA (155) and investigating topoisomerase activity (156). A TFO coupled to a topoisomerase I inhibitor has been used to target the topoisomerase action to a particular DNA sequence (157;158). Another application is using triplex structures to recognise and purify specific DNA sequences from a pool. The TFO can be attached to beads or an affinity column and if parallel triplexes are used binding can be controlled by pH (159-161). Because TFOs can compete with proteins for binding sites on a duplex target fluorescently labelled oligonucleotides can be used to investigate the interaction between DNA and protein (162). This technique was used to identify the binding site of topoisomerase II (163). It has also been used to study translocation of Type I restriction enzymes on DNA (164;165).

1.3 Problems and solutions

The investigation of DNA triple helix formation has brought to light various problems many mentioned above, which need to be considered when contemplating the use of TFOs in gene therapy. The main problems that have been encountered alongside in-vivo challenges are pyrimidine recognition, pH dependence and stability. Each of these problems will be examined in more detail along with examples of the methods which have been developed to overcome them.

Much of the investigation into possible solutions has examined the use of modified nucleotides, where alterations have been made to the sugar, phosphate or base moieties. Base modifications in particular have proved to be especially effective and they also display some unexpected benefits. Some analogues have been shown to improve uptake by cells and help protect the TFO from degradation, others can be used to induce chemical reactions with the DNA and potentially be used as DNA probes (13).

1.3.1 Pyrimidine recognition

The most stable parallel forming natural triplets are T.AT and C^+.GC. Both of these involve the binding of a third strand pyrimidine to a duplex purine base and so rely on the presence of a purine tract in the target sequence. For the targeting of biologically relevant sequences this limitation is impractical, although there does appear to be a higher than expected propensity of purines in many promoter sequences (166).

Synthetic TFOs could be a much more potent biological tool if it were possible to target any sequence; however there is a lack of potential hydrogen bond donor and acceptor sites offered by pyrimidine bases in the duplex target (Figure 1.6 page 8). There is also the possibility of a steric clash between a TFO and the methyl group of thymine which protrudes into the major groove (37).

The natural bases with the highest affinity for pyrimidines in the duplex are G, which binds TA via its 2-amino group, and T in the TFO binding to CG (see Figure 1.5 page 6) (40;41;167-170). Although these natural bases can be used to recognise pyrimidines in duplex DNA they only form one hydrogen bond with the duplex base pair so the interaction is significantly weaker than T.AT and C^+.GC triplets which can form two bonds (170). They also cause local backbone

distortions, since they are not isomorphous with T.AT and C^+.GC (167). The more pyrimidines that need to be recognised the worse the stability becomes. For example, a TFO with two guanines for T recognition is 30 times less stable than a TFO with just one (171). However, the stability of a G.TA triplet can be improved by having T.AT flanking triplets as these are more stabilising in this context than C^+.GC triplets (172). This is due to the formation of an extra hydrogen bond with thymine in the 5' adjacent triplet (167). Recognition of alternating duplex AT tracts by GT-containing TFOs, forming alternating T.AT and G.TA triplets is also possible if these are tethered to a long region of T.AT triplets (41;173;174). The stability can be increased further by adding a few C^+.GC triplets to the T.AT region (173).

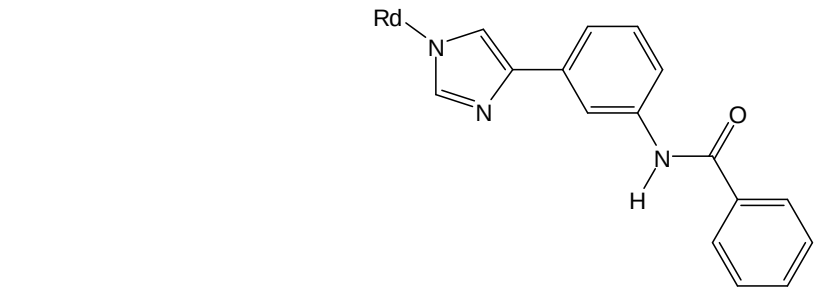
The T.CG triplet is even weaker than G.TA. Thymine in the TFO forms one hydrogen bond between the C2 oxygen from T and the N4 of C, but again if multiple substitutions are made the stability is greatly reduced (37;175). Like the T.AT triplet, T.CG can also form in a parallel or antiparallel conformation (170). In contrast to the G.TA triplet T.CG is stabilised by a 3' C^+.GC triplet, but like G.TA is destabilised by C^+.GC in the 5' position.

1.3.1.1 Analogues

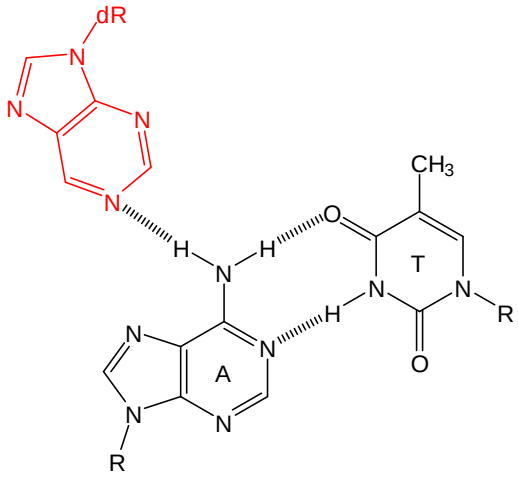
The most promising solution to this problem is to design base analogues which can bind with higher stability and affinity to pyrimidines. This can be achieved through the formation of additional hydrogen bonds. Both purines and pyrimidine have been adapted to form these analogues, although there have been fewer successful nucleotides synthesised for TA recognition compared to CG (37). Some of the most effective analogues are shown in Figure 1.11 on the following page.

D_3

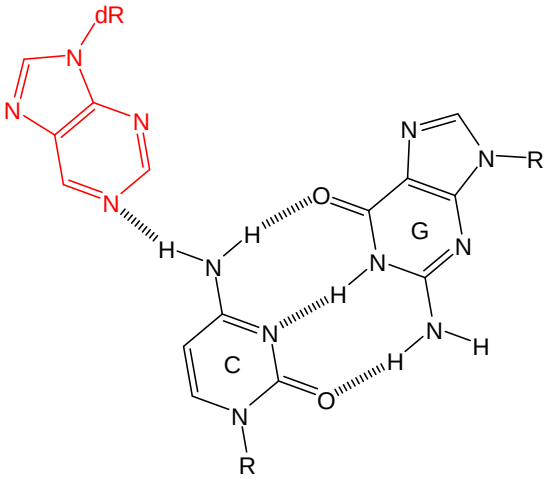
D_3 was one of the first synthetic bases designed to recognise pyrimidines (172). This analogue has a functionalised benzimidazole central group and is selective for pyrimidines over purines in the duplex with a slight preference for CG over TA (69). It will also bind at physiological pH and temperature and is known to stabilise mismatches (172). However it was shown by NMR studies that it actually intercalates between bases rather than binding directly in the major groove (37;176;177).



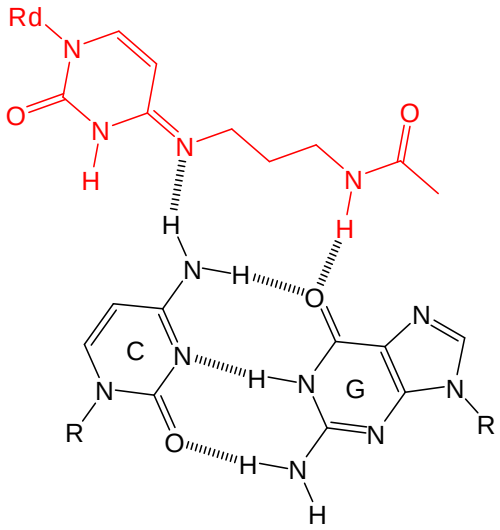
(i) D_3



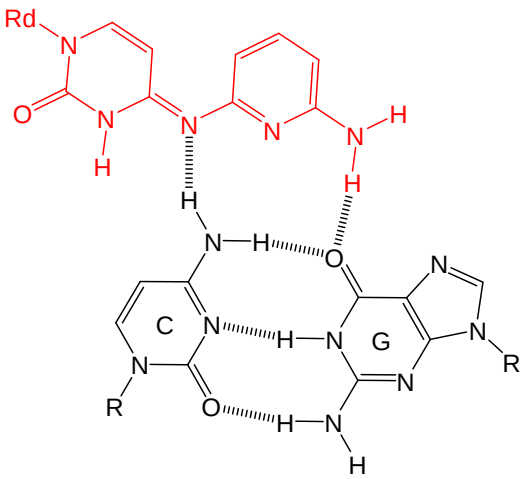
(ii) N.AT



(iii) N.CG



(iv) AcPrC.CG



(v) AmPyC.CG

Figure 1.11: Structures of some synthetic bases bound to their targets.

(i) D_3 (1-(2-deoxy-β-D-ribofuranosyl)-4-(3-benzamidophenyl)imidazole.

(ii) N (deoxynebularine) bound to CG

(iii) N (deoxynebularine) bound to AT.

(iv) AcPrC (N^4-(3-acetamidopropyl)cytosine)

(v) AmPyC (N^4-(6-amino-2-pyridinyl)cytosine)

Adapted from reference (37).

D_3 prefers to intercalate adjacent to TA or CG base pairs, i.e. into YpR steps which have been found to be the favoured site for intercalation. In this conformation the three rings of D_3 stack onto the bases of each of the three DNA strands (177).

N and N^4 derivatives of cytosine

N (deoxynebularine) binds to CG in the duplex in an *anti*-conformation and is more stabilising than a mis-match at the same position (178). It discriminates well against TA and GC base pairs, but is not particularly specific for CG as it will bind AT with similar affinity (37).

The N^4 derivatives of cytosine have been much more successful than D_3 or N and are among the best analogues for recognition of CG base pairs (179-181). AcPrC for example has cytosine at its core and forms 2 bonds with the duplex bases. However only one of these is to the duplex cytosine, the other is thought to form between the amine hydrogen on AcPrC and the O6 carbonyl of G (37). This base is not stable opposite TA unlike D_3 and N, but is compatible with GC, albeit with a lower affinity than for CG (37). It has a flexible side chain as part of its structure and it has been suggested that a nucleotide with a more rigid structure might form a more stable triplex (37).

AmPyC is similar to AcPrC but has an additional ring in place of the flexible linker (181). Like AcPrC it forms 2 hydrogen bonds; one with C and one with G (181). Modelling studies indicate that this monomer spans the major groove placing the 6-amino group of the AmPyC near O6 group of guanine; this is confirmed by experimental evidence which has shown that removing the 6-amino group prevents formation of the second bond (37;181). This N analogue proved to be more stable than AcPrC and binds even at physiological pH, possibly due to increased stacking with the second ring. AmPyC can also form a triplex with AT which is more stable than T.AT. This is thought to be due to a third hydrogen bond being formed with the N7 group of adenine in the duplex (37). Although this is useful in terms of binding strength, the loss of specificity is the down side.

^{A}PP, S and S derivatives

Two of the most stable analogues for pyrimidine recognition synthesised to date are $^{\mathbf{A}}$**PP** which binds C in duplex DNA, and **S** and its derivatives which bind T (182;183).

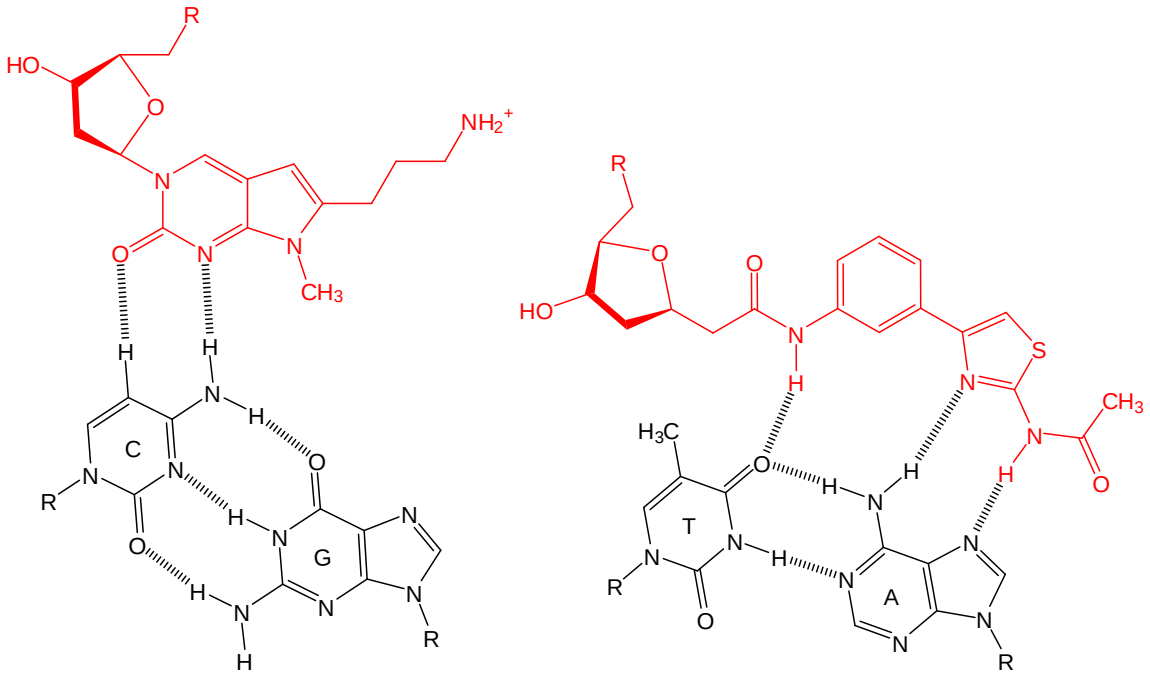


Figure 1.12:
(i) ^{A}PP (6-(3-aminopropyl)-7-methyl-3H-pyrrolo[2,3-d]pyrimidin-2(7H0-one)).
(ii) S (N-(4-(3-acetamidophenyl)thiazol-2-yl-acetamide)).
Third strand bases shown in red.
Adapted from reference (182)

Both S and ^{A}PP have been shown to form stable triplets, and have even been combined with other analogues to achieve recognition of all four bases by one TFO (168;182-184;184). The S base has been used in this thesis during the comparison of anthraquinone intercalation at different sites (see section 1.5.2).

The S analogue is comprised of 2 unfused aromatic rings linked to the sugar via an acetamide group. A single thymidine interruption in a duplex sequence, if targeted by S, does not significantly reduce the stability of the complex even at physiological pH (168). S.TA triplets have been shown to have a melting temperature 5-8ºC better than any natural base for T recognition (166;168). It is interesting to note that S is thought to form 3 base pairs with the duplex, but only one is with thymidine. This bond forms with the 4-oxo group of T in a co-planar structure. The other two bonds are with the N7 and 6-amino groups of adenine in the other duplex strand, forming an unusual triplex structure (166). It has also been shown that this analogue is more stable if targeted to U rather than T, this is due to steric clash of S with the methyl group of T (166).

Other derivatives of this nucleotide have also been designed with greater affinity for TA interruptions, such as ^{2AE}S (183). The addition of an aminoethoxy group to the 2' position of the ribose ring has been particularly beneficial to this and other analogues due to interactions of the positive charge with the phosphate backbones (185-188). However in terms of specificity, S and its derivatives will tolerate CG interruptions in the duplex and at low pH S actually binds CG preferentially to TA, so could potentially reduce the stringency of a TFO (168;183).

^{A}PP shows good selectivity for CG over TA or purine bases, but again these triplets are less stable than T.AT and C^{+}.GC triplets (182). This is another example of an analogue containing a positive amino group; which goes some way to explaining its increased stability: the positive charge may interact with the negative phosphate backbone. The bonding pattern of the ^{A}PP analogue is unusual when compared with most other nucleotides. As shown in Figure 1.12 on the previous page it is hypothesised that ^{A}PP forms 2 hydrogen bonds with T in the duplex. However, one of these bonds, rather than being a usual NH-O bond, is **C**H-O instead. Although the stability of this complex suggests a second hydrogen bond may form, more evidence is needed to confirm this theory (184).

Although these two synthetic nucleotides have gone a long way towards the aim of pyrimidine recognition they are overall still less stable than T.AT or C^{+}.GC. They also are not isomorphic with these triplets and so other methods have been investigated to overcome the problem of pyrimidine recognition (182).

1.3.1.2. Other methods

Rather than carefully designing specific nucleotides to recognise pyrimidines, another approach has been to use abasic linkers to effectively 'skip' these bases. 1,2-dideoxy-D-ribose is one of these. However, this reduces the stability of the triplex compared to one without pyrimidines, especially if multiple substitutions are used, due mainly to loss of base stacking (189;190). Placing a ligand such as acridine next to the abasic site can compensate somewhat for the loss of stability (191;192). However, there is also the problem of loss of stringency with abasic sites. The least destabilising abasic linker found to date is propanediol (193;194).

Another technique which has had more success for targeting pyrimidines is the alternate strand approach. This involves switching which strand of the duplex is targeted in order to avoid a pyrimidine region, see Figure 1.13.

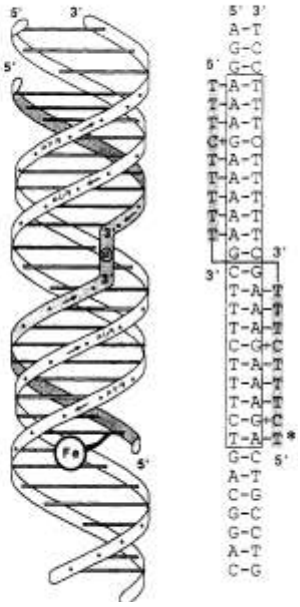


Figure 1.13: Example of alternate strand targeting, taken from reference (195)

There are 2 conformations for this strategy, in which the two halves of the TFO chain have the same or opposite 5'-3' polarities. If the two halves are in the same polarity then they must alternate between a parallel and an antiparallel triplex (196). TC-rich TFOs are used in this way to recognise one strand of the duplex in a parallel orientation; this is joined via a linker to a GT- or GA-rich TFO which binds in an antiparallel orientation to recognise the other strand of the adjacent sequence. There are still sequence limitations: 2 or more consecutive polypurine tracks are needed in opposite orientation; however these don't each need to be as long.

If the TFOs are in the opposite polarity then only parallel triplexes are formed (See Figure 1.13 above) (195;197). These need a central linker to make the 3' 3' or 5' 5' connection. 1,3-propanediol, 1,2-dideoxyribose and a xylose moiety have all been used as linkers for alternate strand technology (195;197;198).

Another way this approach has been adapted is to use much longer, flexible linkers to join two TFOs which bind to the same strand of the duplex, 'skipping' a region of pyrimidines (199). All of these types of triplex formation improve affinity relative to that of two separate TFOs due to cooperative binding of the two strands, and have much greater stability compared with triplexes which contain

mismatches. They can also improve specificity as longer stretches of the duplex can be targeted without the use of low affinity analogues or abasic sites.

1.3.2 Low-pH Dependence

Another major problem that has arisen during the investigation of triplexes is the need for a low pH in order to create 2 hydrogen bonds between a third strand cytosine and duplex guanine. This is necessary because a third strand cytosine can only form one hydrogen bond with guanine at physiological pH and requires protonation of the N3 atom in order to form a second bond and create a stable structure (Figure 1.14 below) (9).

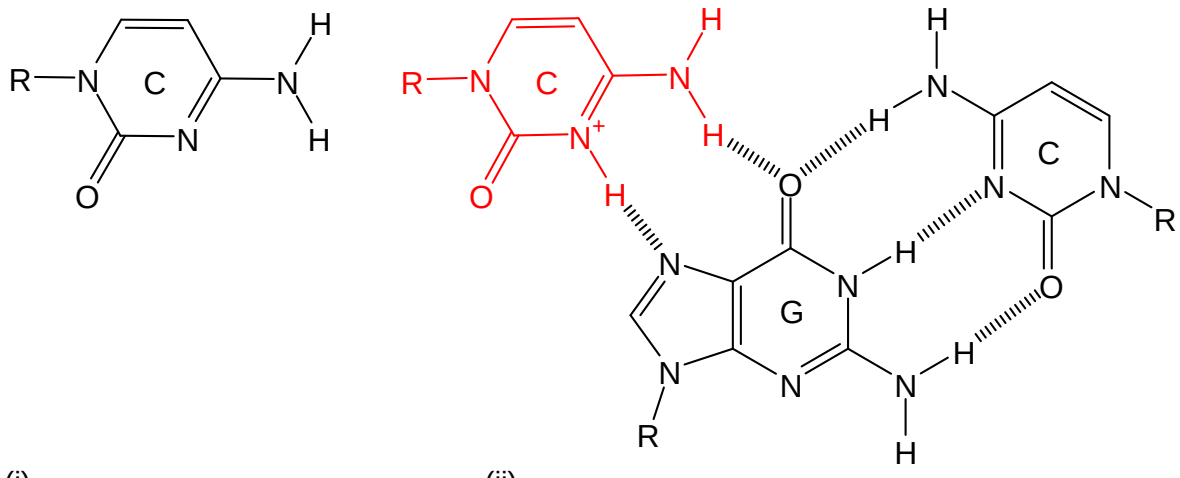


Figure 1.14:
(i) Natural unprotonated cytosine molecule.
(ii) Protonated cytosine molecule in a C^+.GC triplet at low pH.

1.3.2.1 Cytosine

The pK_a of free cytosine is around 4.5 but may be greatly elevated within a triplex, particularly if positioned at the centre. Since protonation of cytosine only occurs at low pH, this is impractical if TFOs are to be used *in vivo* (78). The stability of TFOs containing C^+ therefore decreases as the pH increases, since without protonation the interaction of cytosine with guanine is drastically destabilised (7;79).

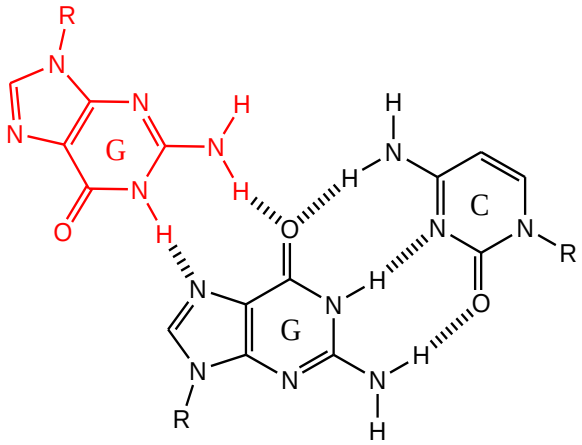
It is also interesting to note that not all of the increase in affinity that comes from protonation can be attributed to the extra hydrogen bond; having an extra positive charge in the triplex also helps to negate some of the repulsion between the phosphate backbones. X-ray crystallography has even shown that the Crick-Hoogsteen groove (between the TFO and central purine strand) is slightly narrower at C^+.GC triplets, presumably because of the attraction between the positive cytosine and the negative phosphates on the purine strand (65). So protonated cytosine actually helps to increase the stability of triplex by reducing the repulsion between phosphates (78). However, runs of C^+.GC triplets have been shown to be destabilising due to repulsion between the positive charges, so this base is most stabilising when dispersed within a T-rich TFO (200).

1.3.2.2 Analogues

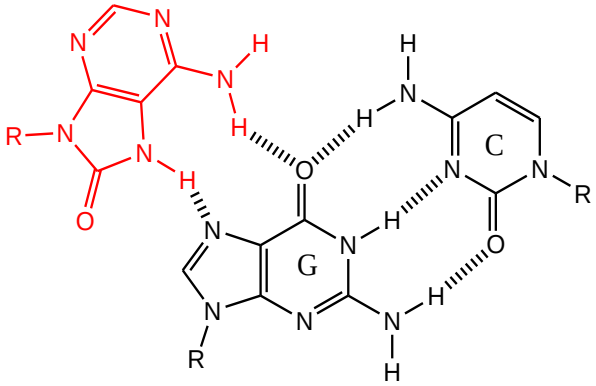
In order to use TFOs in cells, cytosine needs to be able to bind guanine at physiological pH. In order to achieve this goal several base analogues have been designed, some based on a cytosine-like pyrimidine ring, others on a purine ring. These purine analogues are often based on guanine as G.GC triplets are stable and pH independent (75).

Purine-based analogues

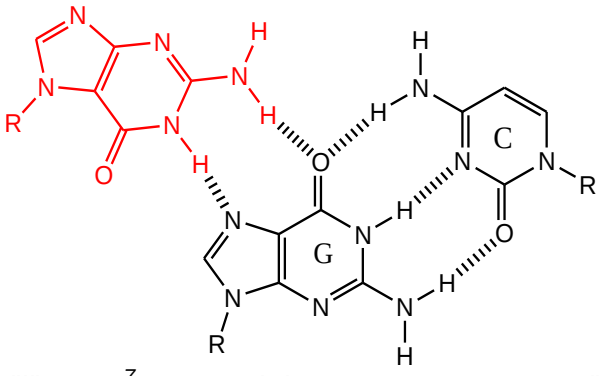
8-oxoadenine (Figure 1.15 ii on the following page) and its derivatives form triplexes in a pH independent manner, with a similar affinity to protonated cytosine. They also have a larger surface area which may help to increase stability by utilising greater base stacking potential. Unlike cytosine, these nucleotides are not destabilising when multiple substitutions are used as they don't contain a positive charge, and can help make a triplex less sensitive to changes in pH (201;202). Many derivatives of 8-oxoadenine with improved binding characteristics have been synthesised, such as N^6-methyl-8-oxo-2-deoxyadenosine and 7,8-dihydro-8-oxoadenine (202;203).



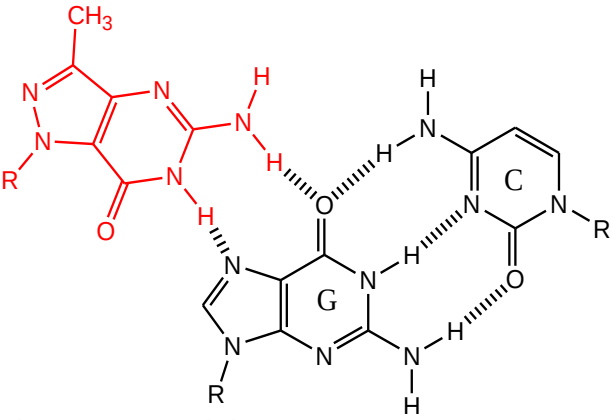
(i) *G.GC base pair*



(ii) 8-oxoadenine.GC triplet



(iii) *N^7-G.GC triplet*



(iv) P1.GC triplet.

Figure 1.15: Purine analogues used to recognise guanine in a duplex GC pair.

N^7 (Figure 1.15 iii) also shows higher stability than methylcytosine for GC recognition but only when contiguous runs are used as it is even less stable than C^+.GC or T.AT when isolated. N^7 is a very selective nucleotide, binding almost exclusively to GC and only in the antiparallel motif (204-206). P1 (3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]-pyrimidine-7-one) is another promising analogue for binding GC as it will bind at neutral pH (Figure 1.15 iv). It has been shown by NMR spectroscopy that P1 also binds to GC in an antiparallel conformation by mimicking the hydrogen bonding pattering in C^+.GC (64;200;207).

Purine analogues generally work well when multiple clustered substitutions are used but are destabilising when interspersed with other bases such as T.AT, probably because these triplets are not isohelical with C^+.GC and T.AT and so cause distortions in the backbone. Although these analogues have an increased pK_a and therefore bind with higher stability at higher pH they have not been as useful as initially hoped as they only bind in the antiparallel conformation and

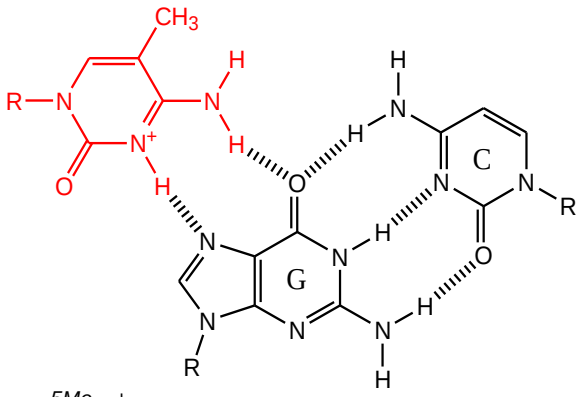
therefore cannot be used in conjunction with the promising thymidine analogues (described below) as these can only be used in the parallel orientation (75).

Pyrimidine-based analogues

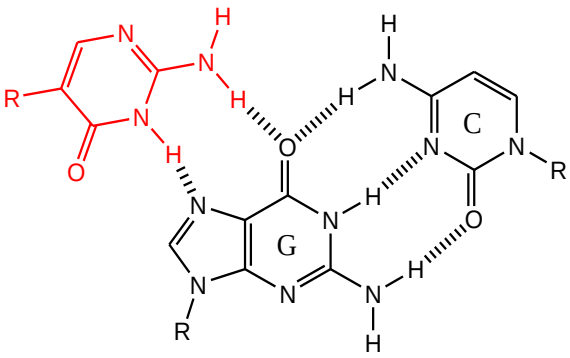
Many pyrimidine analogues have also been designed to help overcome the problem of pH dependence. The first studied nucleotide was 5'-methylcytosine ($^{5Me}C^+$), a naturally occurring cytosine analogue which is now widely used (Figure 1.16 i on the following page) (30;71). $^{5Me}C^+$ has a slightly higher pK_a than naturally occurring cytosine and is therefore protonated at a higher pH (208). The increase in stability however is thought to be partially due to base stacking interactions due to the spine of methyl groups in the major groove (30;208). These methyl groups can also interact with the methyl groups of thymine, increasing the stacking interactions (209). They may also act to displace water molecules creating a positive entropy change (45).

Many of the other analogues in Figure 1.16 are based on this original structure with alterations to attempt to increase the pK_a and stability even further. 6-oxocytosine (Figure1.16 iii) demonstrates pH independent triplex formation, but is less stable than cytosine or $^{5Me}C^+$ at low pH, possibly due to the loss of the positive charge or less favourable base stacking interactions. This nucleotide also reduces triplex stability if long runs are used, despite not having the charge repulsion problems of protonated cytosine (210;211). Some analogues of this base have been synthesised and show more favourable characteristics such as 5-methyl-6-oxocytidine (212;213).

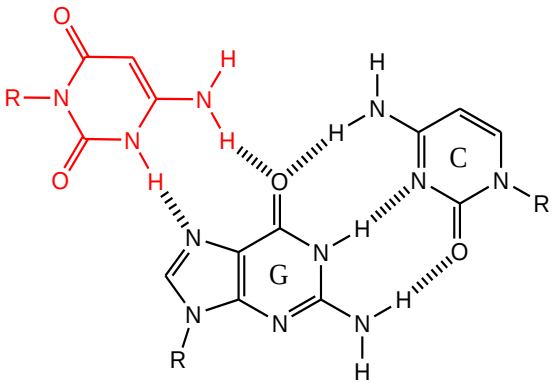
Pseudoisocytosine (ψisoC) and its derivatives including 2'-O-methyl-pseudoisocytidine have been successfully used for stable triplex formation, independent of pH (Figure 1.16 ii) (214). They achieve this by having a permanent hydrogen atom to form the second bond, hence creating the same bonding pattern as protonated cytosine. These analogues are also some of the only uncharged cytosine analogues that form stable triplexes when contiguous bases are used (214;215). Unfortunately these nucleotides have not been widely used as they are difficult and expensive to synthesise (216;217).



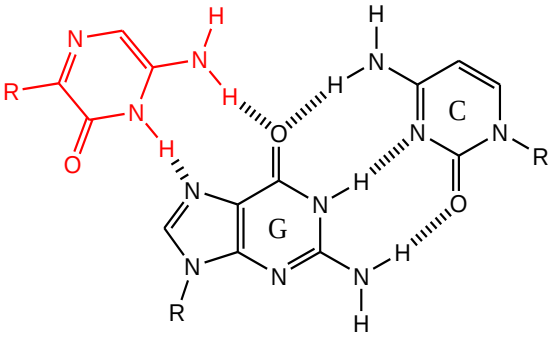
(i) $^{5Me}C^{+}$.GC



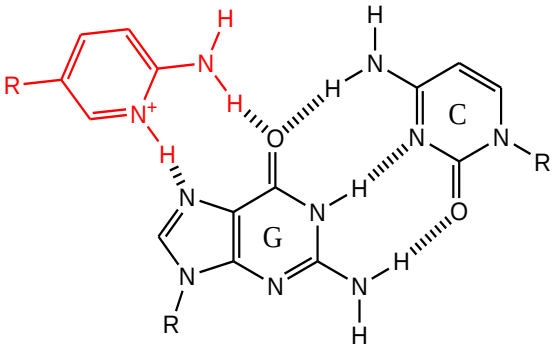
(ii) ψisoC.GC



(iii) 6oxoC.GC



(iv) PyDDA.GC



(v) P.GC

Figure 1.16: Pyrimidine analogues of cytosine.

Some of the most successful analogues created so far are 2-aminopyridine (P) and its derivatives (Figure 1.16 v) (45). P has a pK of around 6.8 and so forms stable triplexes at higher pH than any of the previously mentioned nucleotides, but is also stable at low pH, demonstrating its chemical stability (218;219). Despite its positive charge P is stable in contiguous blocks even at pH 7.0, unlike $^{5Me}C^{+}$ (220).

In more recent work $^{5Me}C^{+}$ and P, the two most promising analogues, have been combined to create a new analogue; 3-methyl-2-aminopyridine (^{Me}P) shown below (220).

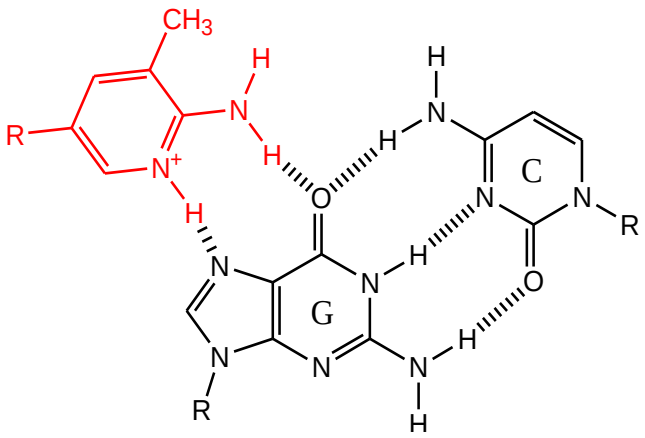


Figure 1.17: ^{Me}P bound to GC

^{Me}P has several advantages over the other analogues mentioned including: no tendency to self-associate, greater triplet stability at higher pH, increased specificity for GC and an inability to form standard Watson-Crick duplexes with guanine, making it ideal for triplex formation (218). However it is still pH dependent, although not to as much of an extent as other protonated cytosine analogues (182).

With this analogue the goal of stable guanine recognition at physical pH has been achieved; the positive charge on the N3 group allows formation of the second hydrogen bond while extra stability is produced by using the methyl group to create extra base stacking interactions (218).

1.3.3 Stability

When a TFO binds to its duplex target the three phosphodiester backbones are brought into close proximity, destabilising the complex. Although this can be counteracted somewhat by the addition of cations such as magnesium it would be useful to be able to form more stable triplexes under physiological conditions (221). The low stability of triplexes creates problems when considering the binding of a TFO to an *in vivo* target. Without a high affinity, that is, a low kD, for its duplex target the TFO will not be able to compete against other DNA binding

molecules such as transcription factors. Other methods for increasing the stability of triplexes have therefore been investigated.

1.3.3.1 Ions

The concentration of various monovalent and divalent metal cations can dramatically alter triplex thermal stability (13;42). They help to stabilise triplexes by screening the negative repulsion between phosphates in adjacent DNA strands. Magnesium salts in particular have been widely used to stabilise parallel triplexes (222). Calcium and zinc ions are also promising parallel and antiparallel triplex stabilisers (109). Potassium ions however have been shown to inhibit the formation of antiparallel triplexes, as they promote quadruplex formation in G-rich TFOs.

Although cations have been useful for *in vitro* experiments to increase triplex stability it is important to develop nucleotide analogues which are less cation dependent for use *in vivo.* Magnesium ions in particular are thought to have much lower concentrations *in vivo* than have been regularly used to increase TFO affinity (223). With this aim, nucleotide analogues as already described have been designed containing positive groups to mimic the screening effect of cations (186;187;224;225).

1.3.3.2 Modifications

Many different methods for increasing the stability of triplexes have been investigated, mainly revolving around alterations to third strand nucleotides to give them higher affinity for their duplex targets. Sugar, phosphate and base analogues have been created in an attempt to increase stability, some of which are described briefly below (183;225-228). Much of the research into increased triplex stability has focused on either attaching positive groups to a TFO, or increasing base stacking interactions (76;186;187;224;225;229). Two important modifications have been used in this report: bis-amino-U and propargylamino-dU. These are discussed separately in sections 1.4 and 1.5.

Backbone

Most backbone modifications involve replacing all or part of the phosphate groups with other moieties in order to make the TFO less negatively charged and hence reduce repulsion between it and the duplex (13). Unfortunately TFOs with neutral backbones tend to be difficult to solubilise and synthesise (13). Cationic phosphoramidates, where the O3' is replaced by NH, have shown some potential, as the positive charge helps to screen the negative repulsion between the strands; they also inhibit quadruplex formation (227). This modification appears to be particularly effective with cytosine and allows triplex formation at a site containing six consecutive guanines at pH 7.0 (230). A TFO containing this modification has also been used *in vivo* to inhibit transcription (231). There have been several other modifications to the phosphate group of DNA backbones including phosphorothioates and methylphosphonates but neither of these is as stable as the phosphoramidate modification (232-234).

One type of backbone modification which has been very successful is PNA (peptide nucleic acid); where the backbone resembles a peptide backbone rather than phosphodiester backbone (235). This means that the molecule is neutral rather than negatively charged and therefore does not repel the two duplex strands, making a much more stable triplex. PNA has a pseudopeptide backbone comprised of N-(2-aminoethyl)glycine with the base attached to the nitrogen of glycine via a carbonyl methylene linker (236).

Rather than forming a PNA.DNA.DNA structure, thymidine-rich PNAs prefer to invade double stranded DNA forming a PNA.PNA.DNA complex – similar to the oligo-clamp method described in section 1.2.2. C-rich PNA molecules tend to form PNA.DNA.DNA structures (237;238). TG- or homopurine- PNAs also use strand invasion but form PNA.DNA duplexes with the purine DNA strand, displacing the pyrimidine DNA strand (237). PNA can in fact form duplexes with DNA and RNA which are more stable than natural duplexes, and PNA monomers are very stable *in vivo* making them attractive tools for gene therapy techniques (236). They also have a lessened tendency to bind to proteins, which can be a problem with traditional TFOs, presumably due to the PNAs lack of charge (148).

Sugar

2'-methoxy groups have been added to the sugars of several nucleotides and nucleotide analogues. This modification, which was introduced in order to prevent nuclease digestion of antisense oligonucleotides, has been shown to stabilise triplex formation as it stabilises the RNA-like backbone conformation (12;239). The 2'-aminoethoxy modification has also been investigated with the hope of positioning a positive charge so as to interact with the negative phosphate groups (225;240). This analogue has been shown to increase association rate and affinity of TFOs due to close proximity with a phosphate group on the purine duplex strand (240). The importance of this interaction has been emphasised by the large decrease in stability if the ethoxy arm is replaced with propoxy. This would position the amino group differently, probably too far away,and break the ionic interaction with the phosphate group (225). An NMR structure later confirmed the interaction between the amino group and phosphates in the purine duplex strand (240). This modification has also been incorporated into the analogue bis-amino-U which is discussed in detail in section 1.6.

Using RNA instead of DNA has also been investigated; the very first triplexes investigated were in fact RNA rather than DNA. RNA TFOs form more stable triplexes than DNA but are much more easily degraded by cells. Also, antiparallel triplexes cannot form if any strand is RNA, and in the parallel conformation the central purine strand cannot be RNA (241). Several RNA- type analogues have been successfully created. These include 2'-methoxylation of the RNA TFO which stabilises the C3'-endo conformation of the sugar and provides protection from nuclease degradation (239;242). 2'-aminoethylribose combines this modification with a positive charge to increase stability (243).

Locked Nucleic Acid (LNA) TFOs increase triplex stability because the sugar is 'locked' into the C3'-endo conformation, which creates an entropic advantage (152;244;245). However TFOs synthesised completely of LNA do not form stable triplexes, presumably because they are too rigid, so must be combined with natural or other modified monomers to allow flexibility (151). Various groups have been attached to LNA monomers and investigated, a 2'-amino modification being among the most effective due to introduction of a positive charge (143). LNA molecules are also nuclease resistant and non-toxic, making them attractive for *in vivo* use (151;246).

Base

The most successful base modifications involve the addition of positive groups to help counteract the negative phosphate repulsion (225-227;247). The initial idea for this type of alteration stemmed from the discovery that the protonation of cytosine to create C^+.GC triplets increases the stability not just because of the extra hydrogen bond formed, but also because of the positive charge (80). In fact protonated cytosine is more stable in a triplex than T.AT, so it was thought that the addition of similar positive groups to thymine might increase its affinity for the duplex as well (79;80;248).

It has been demonstrated that predominantly homothymidine tracts fail to bind their target even at high concentrations (as T.AT is less stable than C^+.GC), but when a positive charge is incorporated, stable binding to a duplex is observed (80). Various positions for the positive charge have been investigated with mixed results (167;249). Propargylamino-dU is a particularly stable positive base analogue and has been investigated in this report. The positive charge on this analogue virtually eliminates the need for the addition of divalent cations, but it does still demonstrate some pH dependence (12). Propargylamino-dU and the sugar analogue 2'-aminoethoxy have been combined to form bis-amino-U, an extremely stable thymidine analogue which is discussed in more detail in section 1.5.

1.4 Positively charged thymidine analogues

Many of the analogues discussed in the previous chapter contain positive groups to help counteract the negative repulsion between the phosphates in a DNA triple helix. Two positively charged analogues in particular have been the focus of this thesis; propargylamino-dU and bis-amino-U.

1.4.1 Propargylamino-dU

Propargylamino-dU is a positively charged analogue of thymine with an amine group on the 5' position of the base. The amino group is designed to interact with the phosphates in the backbone of the TFO strand, helping to screen the negative interactions between the three strands (224). Because of the electron rich triple bond, the amino group should be easily protonated below about pH 8.0; so the important positive charge on the amine will be present at physiological pH (250).

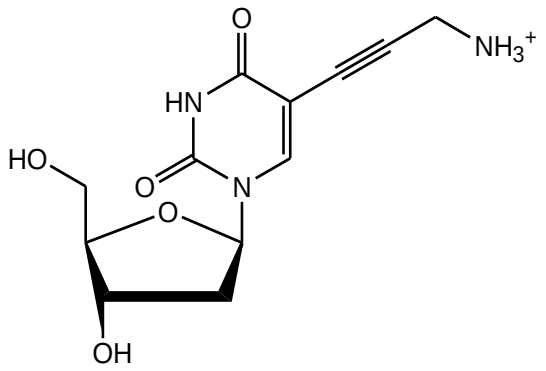


Figure 1.18: 5' - Propargylamino-dU (protonated form).

Triplets formed with propargylamino-dU are pH dependent due to the protonation of the amino group, but it does confers considerable stabilisation compared with thymine (224). Magnesium ions are still needed for triplex formation, although at around a ten times lower concentration (167).

Experiments were first performed with 5-(1-propynyl)-2'-deoxyuridine (pdU); this is similar to propargylamino-dU but without the amino group (Figure 1.19 ii) (251). Several other analogues have also been investigated and are shown below;

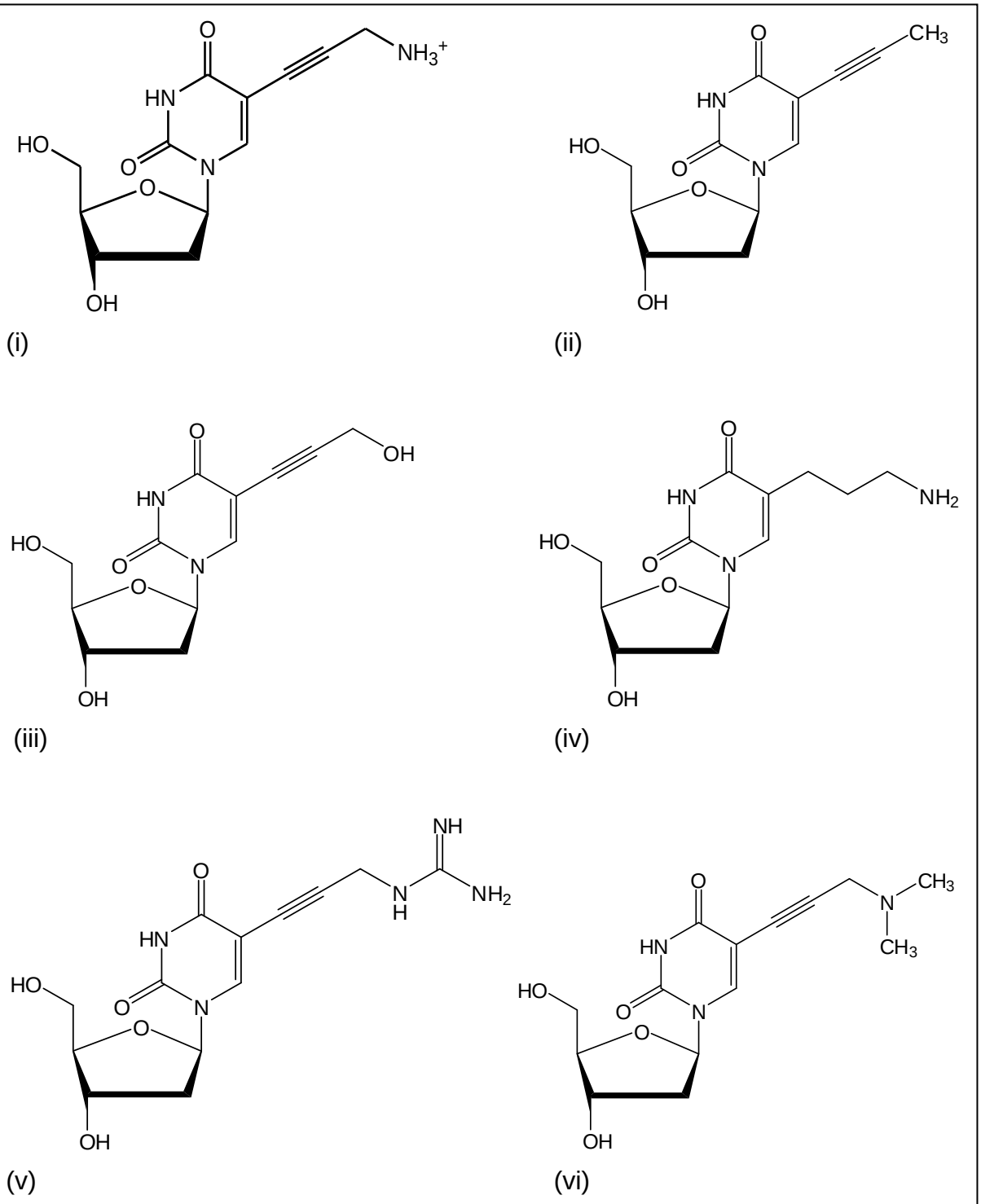


Figure 1.19: Structure of a variety of thymidine analogues with 6 modifications to the 5' position:

(i) Propargylamino-dU (protonated)

(ii) 5-(1-propynyl)-2'-deoxyuridine (pdU)

(iii) 5-(3-hydroxyprop-1-ynyl)-2'-deoxyuridine

(iv) Aminopropyl-dU (unprotonated)

(v) 5-guanidinopropargyl-dU (GPdU)

(vi) 5-dimethylaminopropargyl-dU (DMAPdU)

Like propargylamino-dU, pdU binds at neutral pH, decreases Mg^{2+} dependence compared with thymine and increases intracellular targeting (166). It is not as effective at stabilising triplex formation as propargylamino-dU however, indicating that both the propyne group and the charged amino group play a role in stabilisation (224). In 1998 the NMR structure of a triplex containing pdU was solved (252). This and other experiments have demonstrated that the propynyl group on pdU has increased entropy compared with the methyl group of thymidine, it contributes to base stacking, and has increased hydrophobicity which allows TFOs containing pdU to pass into cells more easily (228;251-253). When used in conjunction with a phosphorothiate backbone, susceptibility to degradation by nucleases within cells was also reduced (254).

In 1996 pdU was used to increase potency of an anti-sense oligonucleotide (ON) designed to inhibit luciferase expression in HeLa cells (253). When pdU was incorporated, a much shorter ON could be used with the same potency; only 9 nucleotides compared with a minimum of 15 with an unmodified ON (253). The ON also retained its sensitivity to mismatches; one mismatch caused a 43% reduction in potency and two mismatches abolished inhibition (253).

Three years later, pdU was used in a TFO coupled with psoralen to study the ability of TFOs to cause mutation (166). This investigation found that a TFO containing pdU caused mutations at physiologically low Mg^{2+} concentration, unlike one comprised solely of natural nucleotides (166). It also showed that the pdU containing TFO caused mutation at a rate four times higher than the natural TFO (166).

Another important analogue for studying propargylamino-dU is 5-(3-hydroxyprop-1-ynyl)-2'-deoxyuridine (Figure 1.19 iii). This has the same structure as propargylamino-dU but with a hydroxyl group as an uncharged hydrogen donor instead of an amino group. It has a slightly lower stability than propargylamino-dU indicating that the high stability of propargylamino-dU may arise from formation of a salt bridge rather than a simple electrostatic interaction (250). The increase in stability from a hydroxyl to amino group is also consistent with a charge rather than electrostatic interaction. This analogue is much more stable than pdU as the hydroxyl group can interact electrostatically with the phosphate backbone (250).

Experiments have also been performed on aminopropyl-dU. Although this does contain an amino group like propargylamino-dU the linker is different (see Figure 1.19 iv) (250). Unlike the other analogues mentioned, TFOs containing

aminopropyl-dU become less stable the more substitutions are made (250). This is presumably due to the increased flexibility of this type of linker (255).

In 2009 a new analogue called 5-dimethylaminopropargyl-dU (DMAPdU) was synthesised (see Figure 1.19 vi) (256). This analogue has a slightly lower stabilisation effect than propargylamino-dU; however it is more stabilising than pdU and does not require amino protection during synthesis like the more powerful modifications (256;257). It can therefore be used with other modifications as well as psoralen and other triplex ligands which require much milder deprotection conditions (257). The lower stability of DMAPdU may be due to steric hindrance caused by the addition of two methyl groups at this position (257). However, it may also be due to loss of the electrostatic interaction between the amino group and phosphate backbone; this is consistent with the good stabilisation produced by 5-(3-hydroxyprop-1-ynyl)-2'-dU (257).

Another interesting analogue is 5-guanidinopropargyl-dU (GPdU), this is again similar to propargylamino-dU but with a guanidine group in place of the amine (see Figure 1.19 v) (257). It has approximately the same stability as propargylamino-dU (257). Propargylamino-dU, GPdU and DMPdU all increase triplex stability when targeted to GC compared with thymidine, although the stability of these triplets is still much lower than when targeted to AT (257). The difference in stability of TFOs containing dispersed or clustered modifications has also been investigated; DMAPdU is less stable if the modifications are clustered, unlike PdU, propargylamino-dU and GPdU which are unaffected by placement (257). This is possibly due to a steric clash between adjacent DMAPdU residues. However, GPdU also has a bulky 5' substitution but is not affected by its dispersal within a TFO; therefore the preference of DMAPdU for a dispersed arrangement may be due to competition of protonation between adjacent DMAPdU nucleotides (257). Many of these analogues have also been investigated for use in anti-sense technology with similar results (255).

Each substitution of T to propargylamino-dU leads to an approximately 12 K stability increase and its use in TFOs does not appear to affect the stringency of binding compared with thymine (224). Propargylamino-dU has been shown to interact with two phosphates in the backbone of the third strand (see Figure 1.23 on page 47) (186). The interaction between the amino group of propargylamino-dU and phosphates on the same strand of DNA may reduce the flexibility of the molecule and therefore help to pre-arrange the TFO for duplex binding. This might help to explain why aminopropyl-dU is less stable (186).

Unfortunately it appears that placing propargylamino-dU adjacent to C^+.GC is unfavourable, probably due to repulsion between the positive charges (167). Unlike contiguous C^+.GC triplets however, runs of propargylamino-dU are not destabilising even though both containing positive groups; the positive charges on the propargylamino-dU residues are screened from each other by their interaction with the negative phosphates, whereas C^+ residues are not (167). Adjacent propargylamino-dU residues are also stabilised by stacking interactions between the propynyl arms (224).

Several of these analogues have been used in antisense oligonucleotides with similar advantages as well as increased protection against nuclease digestion (254). Only propargylamino-dU has been investigated in this thesis.

1.4.2 Aminoethoxy-U

Although 2'-aminoethoxy-U has not been used in this thesis it has been combined with propargylamino-dU to form bis-amino-U which is discussed in section 1.4.3. Like propargylamino-dU this analogue has a positive group specifically designed to interact with the phosphate in the duplex purine strand (240). The positive group has a pK_a of around 8.5-9.0 and so should be protonated at physiological pH (258).

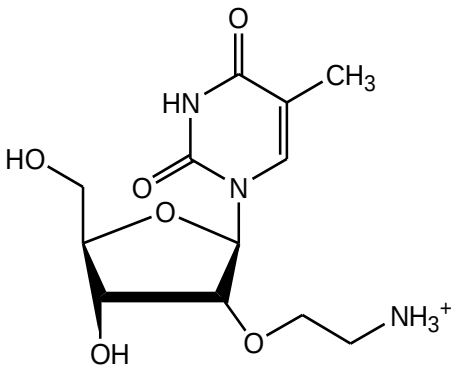


Figure 1.20: 2'-aminoethoxy-U

As can be seen in Figure 1.20 above this analogue has an oxygen atom at the 2' position of the sugar, making it a ribose rather than deoxyribose sugar. This means that unlike the other analogues discussed in section 1.4, 2'-aminoethoxy has an *N*- (RNA-like) rather than *S-type* (DNA-like) sugar pucker (240). This changes the structure of the TFO slightly, making it more rigid and requires less

distortion of the purine strand upon triplex formation (240;259). Another difference is a reduction of inter-phosphate distances within the third strand (240).

Modifications at the ribose 2' position were first investigated in 1998 after the observation that the 2' hydroxyl groups of RNA TFOs appeared to be in close proximity with the purine strand phosphates (260). It was therefore hoped that a positive group at this position could interact with the negative phosphates and reduce the repulsion between strands (225). Before analogues with 2' positive charges were synthesised however a methyl derivative was first investigated. 2'-methyl-U forms a more stable triplex as part of a TFO, indicating that the carbon chain of 2'-aminoethoxy-U is also likely involved in stabilisation, possibly due to the altered sugar pucker pre-organising the TFO for triplex formation (242).

Other 2' modifications have also been investigated. Extension of the chain from ethoxy to propoxy and beyond caused a distinct loss of stability, demonstrating that the amino group needs to be positioned in a specific way to interact with the backbone phosphates in the purine strand (225). Like propargylamino-dU replacement with a simple hydrogen donor (hydroxyl) decreased affinity compared with 2'aminoethoxy-U, but is still more stable than thymine or a simple methyl substitution (225;261).

2'-aminoethoxy-U increases triplex stability by up to 3.5°C per base substitution and TFOs containing this modification are much more effective *in vivo* (225;259;261). It also retains sensitivity to mismatches in the target duplex and is unaffected by magnesium cation concentration (225;259). TFOs containing 2'-aminoethoxy-U have association rates around 1000 times faster than unmodified TFOs (225). This confirms the presence of a specific interaction which not only reduces the negative repulsion between strands but also contributes to the nucleation-zipper mechanism (225;240). This may also increase the bio-activity of TFOs containing 2'-aminoethoxy-U (188). These TFOs also dissociate 40 times more slowly than natural ones (225).

The interaction of the amino group with the phosphate of the purine strand has been confirmed by modelling studies, as shown in Figure 1.21 below:

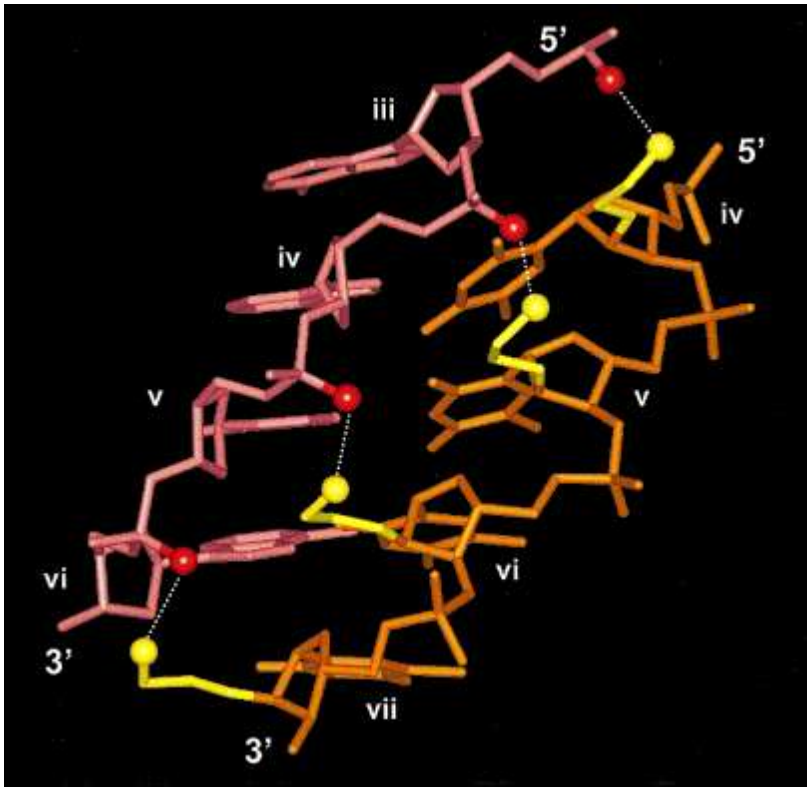


Figure 1.21: Molecular model showing the interaction of the 2'-aminoethoxy side chains (yellow) on the TFO (orange) with the phosphate oxygen atoms (red) of the purine strand (pink). Taken from reference (225).

Another benefit is that 2'-aminoethoxy-U has been shown to have a high level of nuclease resistance (188). Several studies have used 2'-aminoethoxy-U-modified TFOs conjugated to psoralen for gene inhibition *in vivo* (188;262;263). A similar study demonstrated the ability of TFOs containing 2'-aminoethoxy-U to displace proteins from their duplex DNA target (264). Although studies have shown that the more 2'-aminoethoxy-U resides are included in a TFO the more stable it becomes, it has also been shown that having too high a proportion of 2'-aminoethoxy can have a detrimental effect on bioactivity (265). It is thought this might be due to the TFOs being too 'sticky' and adhering to proteins, other molecules or even non-specific DNA sequences (265).

The addition of a 2'-aminoethoxy group to C also stabilises triplex formation, though a C5-propargylamino group decreases the pK_a of this nucleotide (225). 2-aminoethoxy substituent's are even more stabilising if long tracts are used, and this effect is also seen with bis-amino-U (188;225).

1.4.3 Bis-amino-U

The synthesis of 2'-aminoethoxy-5-(3-aminoprop-1-ynyl)uridine (bis-amino-U / BAU) came about by combining two thymine analogues which had shown promise in previous experiments: 2'-aminoethoxy-U and propargylamino-dU (sections 1.4.1 and 1.4.2) (186;224;266). BAU was first synthesised in 2000 and was the first example of an analogue with two positively charged groups (249). The two positive groups were chosen specifically to act at different positions within the triplex (249). Combining the two charged modifications in BAU increases the stability of the BAU.AT triplet more than the combined stability of the two separate analogues and its use in TFOs has been a big step towards more stable triple helical structures which bind with high affinity *in vivo*.

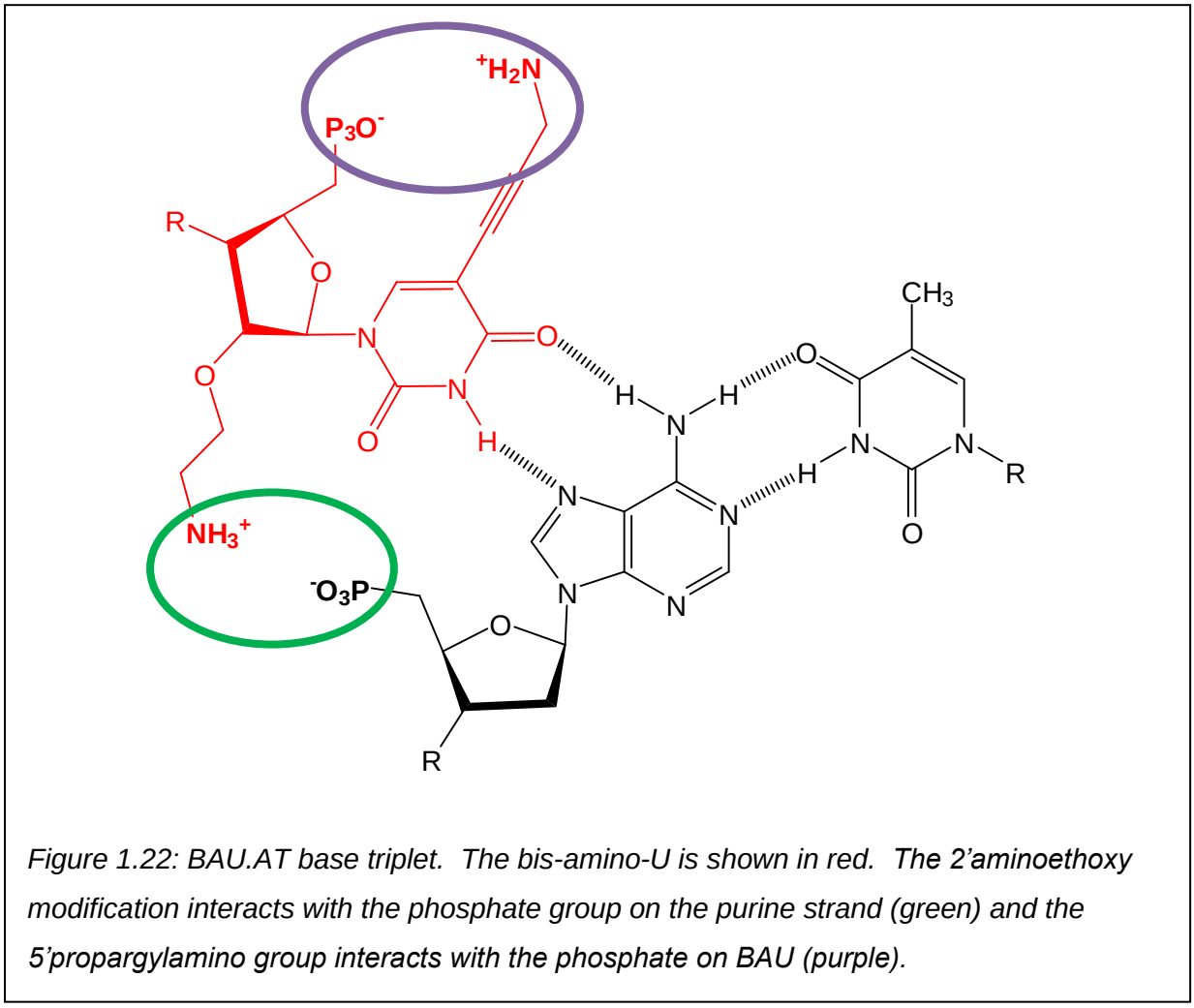


Figure 1.22: BAU.AT base triplet. The bis-amino-U is shown in red. The 2'aminoethoxy modification interacts with the phosphate group on the purine strand (green) and the 5'propargylamino group interacts with the phosphate on BAU (purple).

BAU has two positive charges. One is attached to the sugar (2'-aminoethoxy) and one to the pyrimidine base (5-propargylamino) (see Figure 1.22 above). These positively charged groups appear to interact with specific phosphates in the DNA backbones of the TFO and the duplex purine strands as

described in the previous two sections. Experimental evidence supports the suggestion that there are favourable interactions between both amino groups and separate phosphate groups (186;250). A study carried out in 2002 compared the stabilising effects of propargylamino-dU, aminoethoxy-U and bis-amino-U (186). When the two positively charged groups are combined in BAU, an additive effect is observed indicating that each positive group is contacting a different phosphate (186). TFOs with the same sequence and containing the same number of each modification were compared as follows, using melting curve analysis. At pH 7.0 propargylamino-dU and aminoethoxy-U produced ΔT_m values of 27.1 and 12.5 K respectively, while the ΔT_m value of BAU was 42.4 K, slightly higher than the two separate modifications combined (186).

Bis-amino-U is also less pH dependent than the separate analogues, producing stable binding at pH 7.5, although stability decreases above pH 8.0 due to lack of protonation (186). Like 2'-aminoethoxy-U BAU also appears to dissociate much more slowly than thymidine (187). Unlike propargylamino-dU, where multiple substitutions are generally required before stabilisation is observed, a single BAU substitution enables triplex formation in the absence of magnesium ions; with two substitutions the complex is stable at pH 6.0 (267).

Unlike 2'-aminoethoxy-U, substitution with BAU does not appear to enhance the association rate. However, the dissociation rate of TFOs containing BAU are much slower than thymine (243). This is unfortunate as fast association rates are ideal *in vivo*, although slow dissociation rates are also beneficial. There is some evidence to suggest that if multiple substitutions of BAU are used the association rate is increased slightly (243).

There is also compelling modelling evidence to confirm the interaction between the amino and phosphate groups (186). NMR studies have already shown the close interaction of the 2'aminoethoxy group and a phosphate in the purine strand, with an N-O distance of approximately 2.8 Å (240;268). The 5'-propargylamino-dU group was then manually added to this structure to give an approximation of the amino-phosphate distances for this extra group (see Figure 1.23 on the following page) (186). This modelling showed a N-O distance of 3.0 Å between propargylamino-dU and the closest phosphate in the third strand, and 4.7 Å to the next third strand phosphate suggesting a possible interaction with both of these phosphate groups, although the second interaction may not be as strong (186). Additionally, the propyne chain of propargylamino-dU is 2.4 Å from the methyl group of the adjacent thymine suggesting a hydrophobic interaction

between these chains (186). The interaction of the propargylamino-dU group with phosphates on the TFO itself may help to pre-organise the TFO for triplex formation (186).

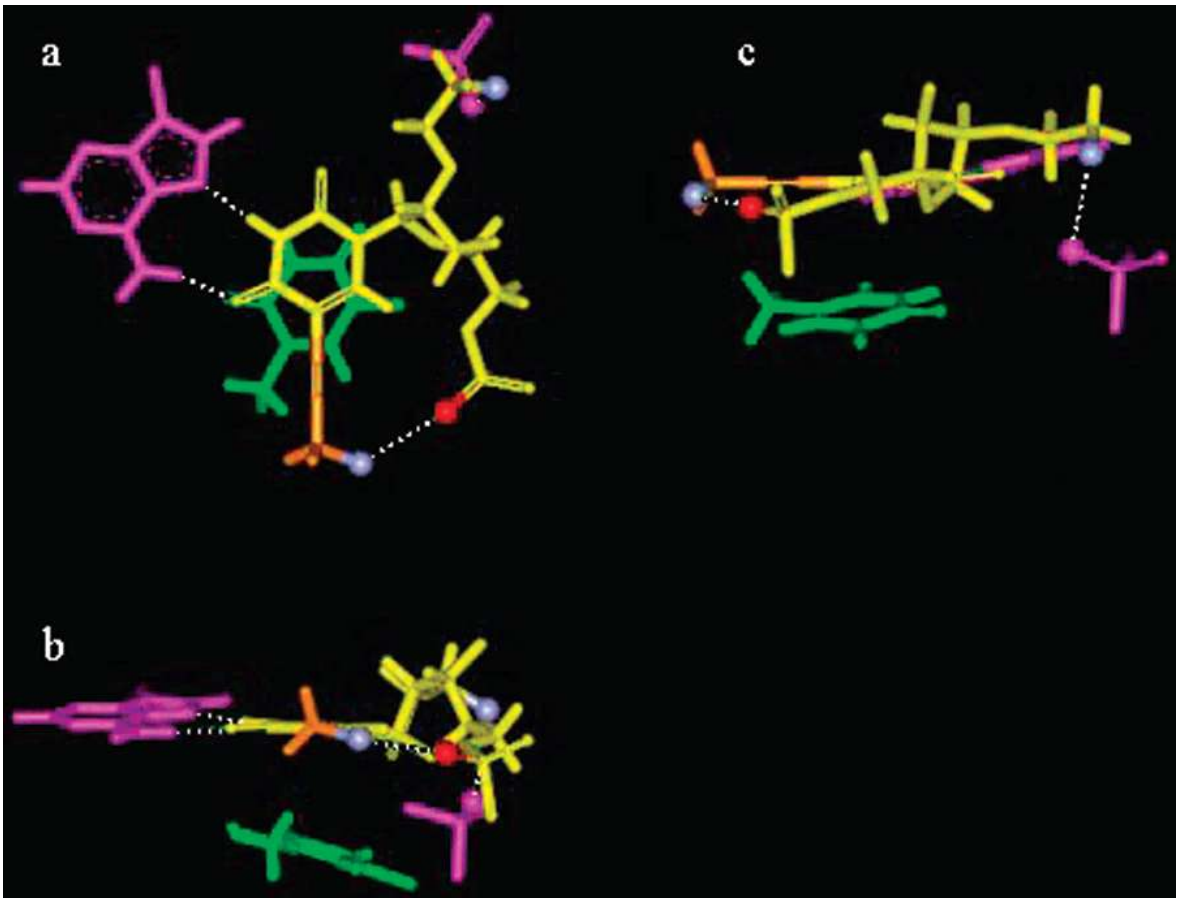


Figure 1.23: Model of possible interactions of the amino groups of BAU based on the NMR structure of 2'aminoethoxy-U in a triplex. (a) top down view, (b) and (c) side-on views. The BAU residue is in yellow, the propargylamino group in orange with the amino group (blue) interacting with the phosphate (red). The amino of the 2'-aminoethoxy group is also shown in blue interacting with a phosphate (not shown in red) in the purine strand (purple). The thymine residue on the 5' side of BAU is in green. Possible hydrogen bonds are shown by white dotted lines.
Taken from reference (186).

BAU has been used in conjunction with several other nucleotide analogues (^{Me}P, S and ^{A}PP) to recognise all four base pairs at physiological pH and without magnesium ions, a huge step forwards for the use of TFOs *in vivo* (182). An interesting advantage of TFOs which interact with the backbone of the target as well as the bases is that they may compete better with DNA binding proteins as these generally interact with the DNA backbone (240).

There is a growing body of evidence showing that the positioning of charged residues within a TFO can affect its binding affinity. 2'-aminoethoxy-U has been shown to be more stabilising if multiple clustered substitutions are used, although the use of too many substitutions in a TFO appears to reduce *in vivo* effectiveness even when it is very stable *in vitro* (188;265). The hydroxyl version of 2'-aminoethoxy-U also appears to be detrimental *in vivo* and has unusual 'sticky' properties *in vitro* if too many substitutions are used, indicating that perhaps the positive charge of 2'-aminoethoxy-U is not to blame for the reduction in effectiveness and it is a result of the extra hydrogen bonding potential of these nucleotides (261). Propargylamino-dU is also more stable when multiple substitutions are used and TFOs containing only a few modifications show only a slight increase in binding affinity compared to TFOs comprised of natural nucleotides (167;224).

Initial work on BAU-containing TFOs suggested that adjacent BAU residues were not destabilising, despite the large numbers of positive charges. This is possibly because unlike C^+, which is destabilising if contiguous residues are used, the positive charges in BAU interact directly with phosphate groups, not only screening the negative repulsion but also preventing the large numbers of positive charges from destabilising the complex (186).

Later work conflicts with this theory however, showing that BAU residues are more effective if dispersed throughout the TFO, and clearly demonstrates that clustered residues can cause dramatic loss of binding affinity, particularly at physiological pH (267). The TFOs used in this study also contained the cytosine analogue ^{Me}P however, so the decrease in affinity may have been due to clustering of these residues rather than BAU (267).

Previous work has been carried out using two particular TFO sequences, and these are also used in this thesis, containing a run of six propargylamino-dU (P) or BAU residues. The propargylamino-dU-containing TFO increases binding affinity by three orders of magnitude at pH 5.0 compared with an unmodified TFO, showing binding at nanomolar concentrations (224). A TFO of the same sequence but containing only three substitutions rather than six failed to produce a footprint. In contrast, a TFO containing three BAU substitutions bound at nanomolar concentrations and was stable at pH 7.0 (186).

1.5 Ligands

Another method which has been widely used to increase the stability of triplexes is to use ligands which selectively bind to triplex (not duplex) DNA. These have been shown to increase the strength of TFO binding by up to 1000 fold (37). Duplex or triplex ligands can also be covalently attached to the TFO, while triplex binding ligands can be added in solution. Ligands can also be used to promote the formation of triplexes which may not form under normal conditions; they can also help to stabilise mismatches and reduce cation dependence (37;171;269).

Naturally occurring polyamines like spermine, spermidine and putrescine have been used to stabilise triplexes (270-273). They are positively charged and so probably help to reduce the repulsion between the strands. Some groups have also attached spermine to the N4 position of methylcytosine, which then becomes stable at pH 7.4 even though at this pH the N3 of methylcytosine is not protonated (226;247;271). Unlike other cytosine analogues this monomer is actually less stable at lower pH and much less dependent on magnesium ions (45). Spermine can also be attached to the 5'-end of a TFO molecule and this terminal substituent is more effective than when spermine is placed at the centre of a TFO (45;273).

Positively charged benzopyridoindole derivatives such as BePI and BgPI are triplex-specific intercalating ligands which have been used to stabilise triplexes, particularly those comprising T.AT (274). These ligands have large fused aromatic ring systems and can increase the melting temperature of short triplexes by over 20ºC; however they cannot bind near to C^+.GC due to ionic repulsion between the positive charges (274-276). Acridine and coralyne are two other positively charged intercalating ligands which have often been tethered to TFOs (192). When tethered to the end of a TFO acridine binds at the triplex/duplex junction and can increase stability around 100 fold. Coralyne is another triplex-specific ligand which has slight preference for T.AT triplets over C^+.GC, though it can also bind to duplex DNA (274;274;277-279). A few planar polycyclic aromatic chromophores such as pyrene, anthracene and dansyl have also been investigated (280).

The intercalating ligands described above generally have large planar ring systems and it is probable that the size of these rings dictates whether a ligand is duplex or triplex specific, depending on how well they stack with the bases in a double or triple stranded DNA structure. These ligands do not generally affect

binding stringency but some have been used specifically to overcome pyrimidine recognition by increasing stability of the surrounding triplets (191;192;281). Two ligands of particular interest have been used in this project and are covered in more detail below.

1.5.1 Naphthylquinoline

Naphthylquinoline is a triplex-specific ligand which can promote triplex formation at sites containing up to 3 consecutive inversions (171;269;282). It is an intercalating ligand with a large aromatic area, which is unfused and so has torsional flexibility which allows it to intercalate between triplex nucleotides as it can adapt to any propeller twist. Although antiparallel triplexes are pH independent, naphthylquinoline has a greater stabilising effect at low pH, suggesting that the active compound needs to be protonated (283). It has also been shown to reduce the concentration of TFO needed to produce a footprint by up to 200-fold (269;284).

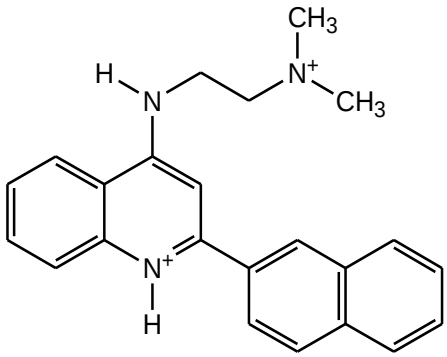


Figure 1.24: Naphthylquinoline

Naphthylquinoline is moderately selective for T.AT over C^+.GC, and stabilises parallel triplexes more effectively than antiparallel ones (269;283;284). As base stacking and base overlap are very similar in both the T.AT and C^+.GC triplets it is likely that the selectivity for T.AT over C^+.GC is due to charge repulsion between the ligand and the positively charged cytosines (283). There are two nitrogen atoms in naphthylquinoline which can become protonated; one in the ring system and one on the aminoalkyl side chain. The side chain nitrogen should be fully protonated at all pHs, but the ring nitrogen has a pK of 7.1. It is therefore

likely that the ring nitrogen is the cause of the apparent pH dependence, and that this protonation has an important effect on triplex stabilisation (283).

Several derivatives of naphthylquinoline have been prepared which also effectively stabilise triplexes (269). In the work described in this thesis naphthylquinoline has been used free in solution, though it can also be tethered to a TFO with slightly greater stabilising effect (285).

Of crucial interest to this investigation, naphthylquinoline has been show to stabilise triplex mismatches even when the TFO would not normally bind the template, although higher TFO concentrations are required (269). This stabilisation is limited to a few mismatches however, and unrelated TFO sequences cannot be forced to bind the template simply by addition of ligand (269). One study indicated that if a TFO is comprised of two sections, one to bind a perfect site and one to bind an adjacent mismatch site, naphthylquinoline can induce triplex formation by anchoring the first section of the TFO (173).

1.5.2 Anthraquinone

Bis-substituted anthraquinone molecules have been investigated as TFO-tethered ligands to increase triplex stability (286;287). A 2,6 disubstituted amidoanthraquinone derivative has been shown to bind triplex DNA preferentially over duplex DNA due to its extended planar surface, and can enhance triplex stability up to 200 fold (288). However a similar 1,4 substituted analogue stabilises duplex DNA and so can be used to destabilise the triplex and cause dissociation of the third strand as it competes with the TFO for the major groove (45). More recently another investigation has looked at a variety of anthraquinone substitutions and found the following order of stabilisation; 2,7 > 1,8 = 1,5 > 2,6, although there is only a six-fold difference between them (289). They also examined mono-substituted anthraquinones and found these to be less effective than their disubstituted counterparts. The variation between these derivatives is thought to arise from differences in DNA groove accessibility and base stacking (290). The positively charged side chains of these derivatives seem to have a relatively minor role in stabilisation in terms of electrostatic interactions. The stacking ability and position of the side chains is more relevant and evidently leads to the differences in binding preferences (288).

It is also interesting to note the difference in anthraquinone stabilisation between parallel and antiparallel triplexes; all the stabilisation properties discussed

above are for parallel triplexes. In the antiparallel conformation, anthraquinone stabilisation depends on the composition of the third strand. TG-containing TFOs are stabilised by 1,8 and 2,7 derivatives but not 1,5 or 2,6. AG-containing TFOs however show no stabilisation by 1,8 or 2,7 substitutions but are stabilised by 1,5 and 2,6.

Anthraquinone molecules can also be tethered directly to TFOs where they act to staple the triplex together, greatly increasing the overall TFO stability. Tethered anthraquinones have been used in this investigation to examine their effect on stringency (see Figure 1.25 below).

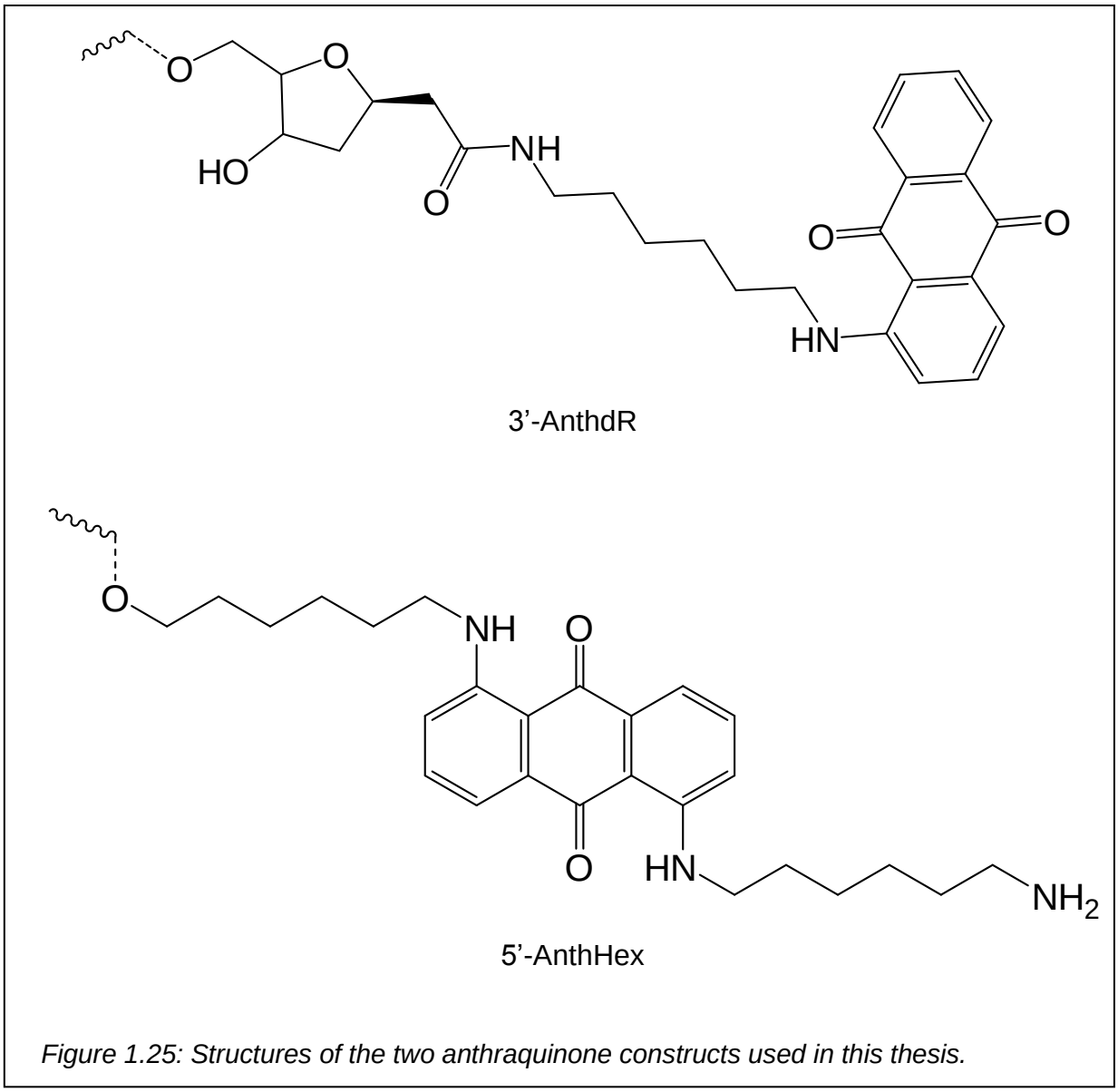


Figure 1.25: Structures of the two anthraquinone constructs used in this thesis.

The 3' anthraquinone modification has no substitutions and is tethered to the DNA backbone with a hexamethylenediamine linker. The 5' anthraquinone has a hexamethylenediamine group on the 5 carbon atom and is bound to the DNA backbone with a hexamethylenamine linker.

1.6 Secondary binding sites

The ability of different modifications to either stabilise or destabilise mismatches within a triplex is of particular importance in this thesis. The binding of TFOs containing multiple substitutions of propargylamino-dU or BAU to secondary binding sites is examined and compared. The high positive charge of these TFOs appears to allow them to bind at these secondary target sites.

Within a triplex comprised of purely natural nucleotides just one mismatched triplet can cause destabilisation of the complex (168;291). Mismatched bases in triplex DNA have been shown to be as destabilising as mismatches within duplex DNA, if not more so (45;292;293). They appear to reduce the stability of triplexes by increasing the dissociation rate rather than altering the rate of association (74). The effect of a mismatch on triplex stability is subject to positional effects; they have a greater destabilising effect at the centre of a triplex than near the triplex-duplex junction (74;172;291). They are also affected by the flanking bases, and different mismatches have different destabilising strengths (172;291).

Many of the analogues mentioned in this chapter have been tested against various mismatches. For example spermine conjugated to cytosine residues within a TFO appears to allow the cytosine residue to bind AT as strongly as GC, although it still discriminates against pyrimidine bases (247). The spermidine compensates for the loss of cytosine protonation at high pH but this reduces specificity (247).

Several investigations using a propargylamino-dU substitution in an anti-sense oligonucleotide found that it showed similar sensitivity to mismatches compared with an unmodified oligonucleotide (253;255).

Although early experiments found that 2'-aminoethoxy-U was sensitive to mismatches in the target DNA (225), a later study on bioactivity showed that TFOs extensively modified with this substitution have low bioactivity despite appearing to be very effective in *in vitro* experiments (262). The authors of this paper have speculated that this may be due to the high positive charge making the TFO 'sticky' and allowing it to bind not only to non-target sequences but also to other molecules within the cell such as proteins. They found that the TFO was difficult to manipulate *in vitro* and adhered to glass columns and tubes etc, substantiating this hypothesis (262). A more recent paper has support the idea that TFOs highly

modified with 2'-aminoethoxy-U show high activity *in vitro* but very low activity *in vivo* (265).

Studies on BAU triplex formation have generally shown similar results to those for other positively charged analogues. BAU appears to be more sensitive opposite pyrimidine mismatches than thymidine is, particularly at high pH (187;266). Although BAU is highly selective against pyrimidine bases in the duplex it is no better than thymidine at discriminating against guanine (182;187). Also because it generally binds the duplex much more tightly than naturally comprised TFOs it can stabilise mismatches at other points in the triplex (182). A TFO containing six substitutions of BAU was shown to be stable at unusual secondary sites which at first glance show very little relation to the TFOs expected target site (294). It was this discovery which became the basis for the experiments discussed in this thesis.

1.7 Methods used in this thesis for studying triplex specificity

1.7.1 Footprinting

Footprinting was first used in 1978 for studying the interaction between proteins and DNA (295). It has since been developed as a versatile method for identifying the sequence-specific interaction of many drugs, proteins and nucleic acids with their DNA targets. Footprinting is a method for determining the sequence selectivity and affinity of DNA-binding compounds. Ligands such as TFOs are used to protect radiolabelled DNA from cleavage by agents like DNase I, which has been used in this thesis.

The DNA template is cleaved in the presence and absence of a ligand such as a TFO, and the regions to which the ligand is bound are protected from digestion, creating a gap or 'footprint' in the ladder of cleavage products when these are resolved on denaturing polyacrylamide gels (see Figure 1.26 below). The reaction requires that each DNA molecule is only cleaved once (single hit kinetics).

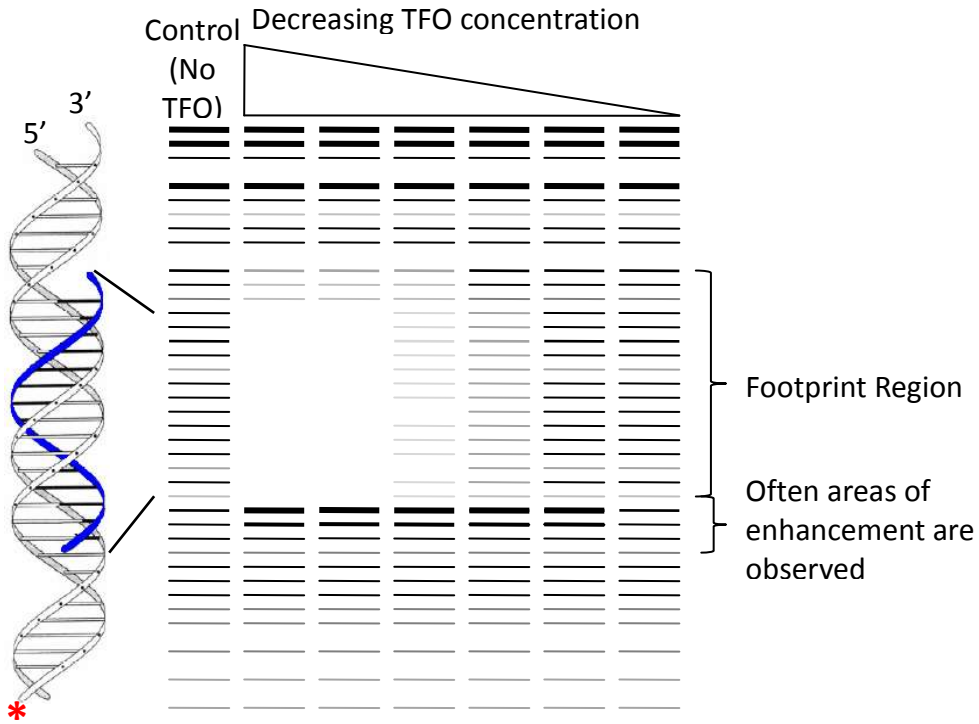


Figure 1.26: Footprinting basics. The left side image is a representation of a duplex DNA template bound by a TFO (blue), the red star represents the radiolabel. On the right is a schematic of a footprinting gel. The region of duplex DNA bound by the TFO is protected from DNase I cleavage. Once digested, the DNA is run on a denaturing polyacrylamide gel where missing bands show the area bound by the TFO.

DNase I is a monomeric glycoprotein with a molecular weight of about 30,400 Da. It is a double strand-specific endonuclease, though it will cut single-stranded DNA, albeit at a much lower rate. It requires the presence of divalent cations and makes single strand nicks in the DNA backbone, cleaving the O3'-P bond. Optimal cleavage is obtained using calcium or magnesium ions, while manganese ions enhance the activity.

1.7.2 REPSA

A relatively novel selection method, REPSA (restriction endonuclease protection, selection and amplification), has been utilised in this study to identify the binding sites of the TFO of interest. This technique was first devised by Hardenbol and Van Dyke in 1996 and has been used to study interactions of DNA with proteins, TFOs and small molecules (77;296-299). Unlike previous selection techniques REPSA does not depend on physical separation of bound from unbound template. This makes it ideal for non-covalent interactions such as DNA-DNA and protein-DNA interactions.

Like a gel retardation assay REPSA is a method by which a population of possible target sequences can be separated into those that bind a TFO and those that don't. It also contains aspects of a protection assay, where a ligand is used to protect the DNA sequence it binds to from restriction enzyme digestion (3).

In REPSA a synthetic population of random templates is tested with the ligand. This template contains a randomised central region adjacent to a binding site for a Type IIS restriction enzyme, such as Fok I (which is used in this study). Type IIS restriction enzymes recognise a specific sequence but cut at a fixed distance from their binding site and so can be used to cleave a region with an unknown sequence. This makes REPSA superior to other methods as no prior knowledge of the target sequence if required.

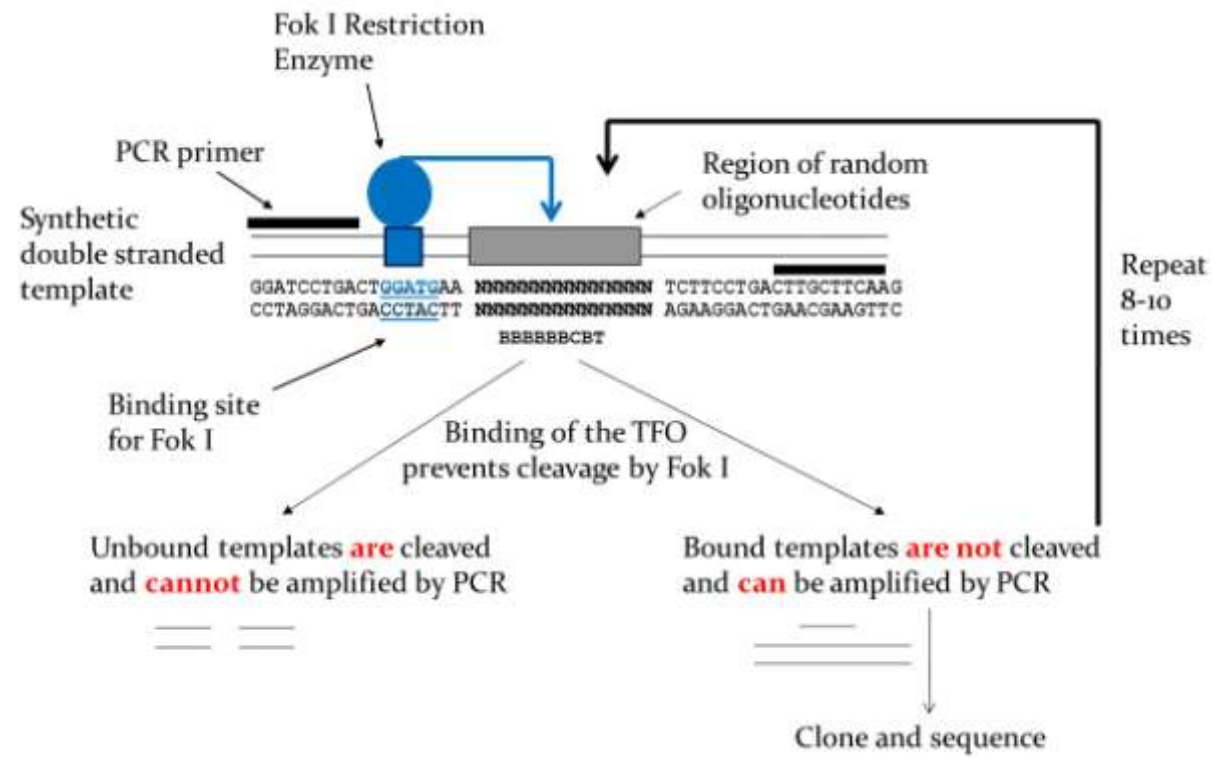


Figure 1.27: Typical REPSA experiment, using a double stranded template, primers and Type IIS restriction enzyme.

The duplex targets are incubated with the ligand. A small number of the randomised templates will be bound by the ligand, but the majority will not. A type IIS restriction enzyme is then used to cleave both strands of any unbound templates, which are protected from cleavage by the TFO, and can then be amplified by PCR. The amplified DNA is then purified and the protocol repeated around 10 times, each round of REPSA enriching for ligand-bound targets. Once REPSA has been completed the selected templates can be cloned into plasmids, transformed into competent cells and plated out. The colonies that grow should each contain a different template which can then be sequenced.

1.7.3 Band shift

A gel retardation assay involving the separation of triplex from duplex DNA in order to obtain the target sequences for a TFO was also investigated. It relies on the principle that a DNA triple helix will have a marginally different mobility to that of duplex DNA when run on a non-denaturing polyacrylamide gel. The purpose of this technique is the same as REPSA; to determine the binding site preferences of a TFO which exhibits some non-specific binding. It was run alongside REPSA with the same TFOs in the hope of corroborating the REPSA results.

1.8 Aims and objectives

The main aim of this work was to examine how positively charged stabilising modifications affect triplex specificity, especially when multiple substitutions are included. The first part of this thesis (Chapters 3 and 4) looks at the specificity of a TFO containing a high number of positively charged bis-amino-U residues and uses REPSA and footprinting to elucidate and examine the primary and secondary binding sites of this TFO. The second part (chapter 5) uses footprinting to examine the specificity of TFOs containing the positively charged intercalating ligand anthraquinone at different positions. The preliminary optimisation experiments for REPSA are detailed in chapter 2 as well as an alternative band shift technique which was investigated but failed to yield results.

Chapter 2: Materials and Methods

This thesis investigates the specificity of several TFOs containing positively charged nucleotides. The first part of this thesis (chapters 3 and 4) looks at the specificity of a TFO containing a large proportion of positively charged bis-amino-U residues and uses REPSA and footprinting to elucidate and examine its secondary binding sites. The second part (chapter 5) examines the specificity of TFOs containing the positively charged intercalating ligand anthraquinone at different positions using footprinting. The protocols for generating targets, preparing and labeling DNA, footprinting and REPSA can be found in section 2.2. In sections 2.3 and 2.4 the preparatory work prior to REPSA and a secondary band shifting method are detailed.

2.1 Materials

2.1.1 Footprinting templates

Several DNA fragments have been used as footprinting templates in this thesis. One of the main templates used throughout is the *tyr*T(43-59) sequence shown below:

```
5′ -AATTCCGGTTACCTTTAATCCGTTACGGATGAAAATTACGCAACCAGTTCTTTTTTTC
3′ ---AAGGCCAATGGAAATTAGGCAATGCCTACTTTTAATGCGTTGGTCAAGAAAAAAG

TCTTCCTAACACTTTACAGCGGCGCGTCATTTGATATGAAGCGCCCCGCTTCC----- 3′
AGAAGGATTGTGAAATGTCGCCGCGCAGTAAACTATACTTCGCGGGGCGAAGGGCTC- 5′
```

Figure 2.1: Sequence of the tyr*T template, the binding site for the TFOs used in chapters 3 and 4 is shown in red (the binding site for the TFOs used in chapter 5 will be discussed later) the nucleotides bearing the ^{32}P are underlined.*

This template was originally created by cloning into the *tyr*T fragment between the *Eco*RI and *Sma*I sites of pUC18. The 17 base pair oligopurine tract used in footprinting was then introduced into the sequence by mutation (142).

A series of DNA fragments which emerged from REPSA selection have been used as footprinting templates in Chapter 4 and are detailed there. The *tyr*T fragment shown above has also been mutated to create a range of similar

templates which are used in Chapter 5. The technique and mutated fragments are detailed in section 2.2.8.

2.1.2 Oligonucleotides

Apart from noted exceptions all the oligonucleotides and TFOs listed below were synthesised by Prof. Tom Brown and his group (Department of Chemistry, University of Southampton). The oligonucleotides were dissolved in water and stored at -20 ºC and at stock concentration on 20 mM.

2.1.2.1 Specificity of bis-amino-U (chapters 3 and 4)

Three TFOs have been used in this thesis to determine the specificity of BAU and to compare this with T and propargylamino-dU:

i) 5′ – TTTTTTCTT – 3′

ii) 5′ – BBBBBBCBT – 3′

iii) 5′ – PPPPPPCPT – 3′

Figure 2.2: Sequence of TFOs used to study the specificity of BAU and propargylamino-dU compared to thymine.
B = Bis-amino-U
P = Propargylamino-dU
T = Thymine

The TFO in Figure 2.2 i is a control TFO for reference; Figures 2.2 ii and iii show TFOs of the same sequence containing either propargylamino-dU or bis-amino-U to compare the effect of one or two positive modifications respectively.

A single stranded oligonucleotide was synthesised for use as the template in REPSA experiments:

Template:

5′-GTAGGATCCTGACTGGATGAA**NNNNNNNNNNNNNNN**TCTTCCTGACAAGCTTCAG-3′

Primer:

5′-*CTGAAGCTTGTCAGGAAGA-3′

Figure 2.3: Above: Sequence of randomised single stranded template synthesised for REPSA. N represents any base inserted at random during synthesis (in red). Below: Sequence of the primer used in primer extension, the red star indicates the position of the radiolabel.

In this sequence N represents any base inserted randomly during synthesis thus creating a theoretical random pool of template sequences. Primer extension was used to create double stranded DNA for use in REPSA from these single stranded synthetic templates (see sections 2.2.6 and 2.3.1).

A similar sequence of the same length containing the exact TFO binding site was also synthesised along with its complement so that a double stranded template could be generated:

5′-GTAGGATCCTGACTGGATGAACGC**AAAAAAGAA**GCGTCTTCCTGACAAGCTTCAG-3′

3′-CATCCTAGGACTGACCTACTTGCG**TTTTTTCTT**CGCAGAAGGACTGTTCGAAGTC-5′

Figure 2.4: Complementary single stranded synthetic templates containing the exact binding site for the TFOs shown in red.

This was used for some of the preliminary REPA and band-shift experiments. The sequence was subsequently cloned between the *Bam*HI and *Hin*dIII sites of pUC19 to generate a fragment for use in footprinting experiments.

Many other sequences have also been used for footprinting; these were selected from the random pool of oligonucleotides using REPSA and the sequences are shown in Chapter 4. See sections 2.2.7 and 2.3 for more detail on REPSA.

2.1.2.2 Specificity of anthraquinone (chapter 5)

Five TFOs were used for this part of the thesis, shown below in Figure 2.5;

i	5′ – TCCTTCTCTTTTTTCTTT – 3′
ii	5′ – XTCCTTCTCTTTTTTCTTT – 3′
iii	5′ – TCCTTCTCTTTTTTCTTTX – 3′
iv	5′ – XTCCTTCTCTTTTTTCTTTX – 3′
v	5′ – TCCTTCTCTTTTTTCTTSX – 3′

Figure 2.5: TFOs A-E. Sequence of TFOs used to study specificity of anthraquinone.
X = Anthraquinone
S = S base (N-(4-(3-acetamidophenyl)thiazol-2-yl-acetamide)

Figure 2.5 i is a control (unmodified) TFO; ii-iv are TFOs with the same sequence but with an anthraquinone modification at the 5', 3' or 5' and 3' ends respectively. TFO v has a similar sequence but with an S base at the 3' end.

These TFOs were footprinted on several templates to investigate their selectivity. The generation and sequence of these templates is described in section 2.2.8.

2.1.3 Enzymes and chemicals

All enzymes, buffers and dNTPs used in this thesis were purchased from Promega. Exceptions were AMV reverse transcriptase and DNase I which were purchased from Sigma-Aldrich. Polynucleotide kinase, PNK buffer, Fok I, Fok I buffer and pUC19 were purchased from New England Biolabs. The majority of chemicals were purchased from either Sigma-Aldrich. However, sodium chloride, sodium hydroxide and ethanol were purchased from Fisher Scientific.

Radionucleotides were purchased initially from Amersham and later from Perkin Elmer as Amersham discontinued production of short life isotopes. All reagents for preparing polyacrylamide gels were purchased from National Diagnostics. Hydrophobic 'non-stick' microcentrifuge tubes were from Alpha Laboratories. Tryptone, carbenicillin, IPTG and Xgal were purchased from Melford Labs, and yeast extract and blood agar from Difco (Becton, Dickinson & Co.). Agarose and Ficoll were from Sigma-Aldrich.

The Naphthylquinoline triplex binding ligand was a gift from Dr. L. Strekowski (Dept. of Chemistry, Georgia State University) and was stored at a concentration of 20 mM in dimethylsulphoxide at -20°C. The source of any other materials is stated in the text. All reagents were stored as recommended by the manufacturer.

2.2 Protocols

2.2.1 Transformation

Competent cells were made for use in transformations using the following protocol. A colony of *E.coli* TG2 cells was gown overnight at 37ºC in sterile 2YT media (16 g tryptone, 10 g yeast extract and 5 g sodium chloride per litre). 1 ml of this culture was then transferred to 100 ml of 2YT media and grown until an optimal density of 0.6-0.8 at 600 nm was reached. The cells were harvested by centrifugation at 3000 rpm for 10 minutes in a sterile 25 ml tube. The media was then decanted from the pelleted cells, which were then resuspended in approximately 20 ml of sterile transformation buffer (50 mM calcium chloride, 10 mM Tris-HCl pH 7.5). The resuspended cells were then placed on ice for at least 30 minutes before being pelleted again and finally resuspended in approximately 5 ml of transformation buffer and stored at 4ºC for up to two weeks.

In transformations the exact quantities of cells and plasmid were varied over the course of this investigation in order to obtain different densities of cells. However, the general procedure is outlined below.

1-5 µl of plasmid DNA (typically 100 µg/ml) was added to 20-200 µl of competent cells in a sterile 1.5 ml micro-centrifuge tube and placed on ice for at least 30 minutes. This was heat shocked at 45 ºC for 1 minute to allow the plasmid to enter the cells, then placed back on ice for a few minutes. The transformed cells were plated onto sterile 20 ml agar plates containing 100 µg/ml carbenicillin to select for cells containing plasmid DNA. Transformants containing clones from REPSA experiments or other fragments were detected by blue-white selection on agar plates containing 1 mM IPTG and 5 µM Xgal in addition to the carbenicillin. This allows selection between ‘empty’ plasmids and recombinant plasmids (which have had a new fragment ligated in). The plates were incubated overnight at 37 ºC and then sealed with parafilm and stored at 4 ºC.

2.2.2 Plasmid Preparation

Colonies were picked from the agar plates using either a sterile metal loop or sterile 2-200 µl pipette tips and grown overnight at 37 ºC in 5ml of 2YT media containing 100 µg/ml carbenicillin. The cultures were then split into two or three 1.5 ml micro-centrifuge tubes and centrifuged at 6000 rpm for 5 minutes to pellet the cells. The supernatant was discarded and the plasmids were purified using a Qiagen QIAprep kit. First the pellets were re-suspended in 250 µl Buffer P1 containing RNase A and then lysed by addition of 250 µl Buffer P2. The lysis was stopped by addition of 350 µl of Buffer P3 and the solution centrifuged at 13000 rpm for 10 minutes. The supernatant was then transferred to a Qiagen spin column and centrifuged three times for 30-60 seconds at 13000 rpm, discarding the flow through each time. After the first and second centrifugation steps 500 µl of Buffer PB and 750 µl of Buffer PE were added respectively. After the third centrifugation the column was placed in a 1.5 ml microcentrifuge tube and 50 µl of Buffer EB was added, the tube was then spun for 30-60 seconds to elute the plasmid. Purified plasmids were stored in Buffer EB at -20 ºC; in some experiments water was used to elute the plasmid, for example in REPSA.

2.2.3 Radiolabeling

Many of the techniques used during this project involve radiolabelled DNA fragments of varying lengths in which the radioactive isotope ^{32}P is attached to double or single stranded DNA at one or other end.

Two different radiolabeling methods were used in this project; one attached [α-^{32}P]-dATP to the 3' end of a section of double stranded DNA, and the second to attach a radioactive phosphate group from [γ-^{32}P]-ATP to the 5' end of a single stranded DNA sequence. Alpha and Gamma refer to the placement of the radioactive phosphate in the ATP;

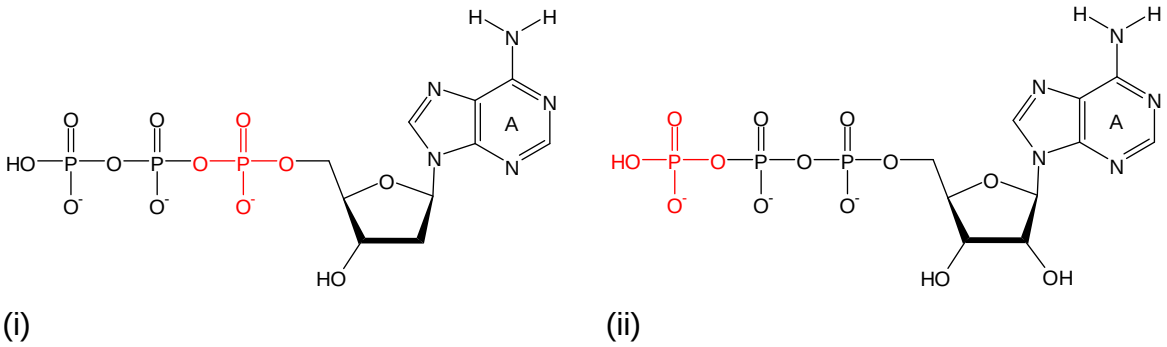


Figure 2.6: Diagram showing alpha and gamma labelling positions; the red phosphate indicates a radioactive phosphorus. (i) [α-^{32}P]-dATP (ii) [γ-^{32}P]-ATP

2.2.3.1 3' labelling

3' labelling is the method used to attach a dAMP molecule containing a radioactive phosphate, to the 3' end of one strand of duplex DNA, by filling in sticky ends left after restriction digestion using [α-^{32}P]-dATP and a DNA polymerase. The double stranded *tyr*T and REPSA fragments were labelled in this way.

The plasmid DNA for labelling was obtained as described in section 2.2.2. For plasmids containing the *tyr*T sequence, 40 µl of purified plasmid was combined with 4 µl of Promega buffer B, 18 units of *Eco*RI and 15 units of *Ava*I. This mix was then incubated at 37 °C for one hour. 1 µl of [α-^{32}P]-dATP, 10 units of AMV reverse transcriptase and 5 µl of reverse transcriptase buffer were added and incubated at 37 ºC for a further hour to fill in the sticky ends at the *Eco*RI site. The labelled fragment was then purified as described in section 2.2.3.3.

The REPSA fragments were labelled in a similar way but the restriction enzymes *Eco*RI and *Hin*dIII were used to release the fragment of the desired length. As these both leave exposed Ts in the 5' overhang the plasmid was first cut with *Eco*RI and then labelled with α[^{32}P]-dATP using reverse transcriptase. The reverse transcriptase was then deactivated by heating at 65 ºC for 10 minutes before cutting with 15 units of *Hin*dIII for one hour.

2.2.3.2 5' labelling

This method was used to attach a phosphate from [γ-^{32}P]-ATP to the 5'-end of a single stranded DNA fragment. The primer used to create the mixed sequence REPSA fragments and the band shift templates, as well as the TFO were labelled in this way.

2 µl of 5 µM oligonucleotide was incubated with 2 µl [γ-^{32}P]-ATP, 2 µl T4 polynucleotide kinase buffer (NEB), 20 units of T4 polynucleotide kinase (NEB) and 13 µl water at 37 ºC for one hour. The kinase attaches the radioactive terminal phosphate from the ATP onto the 5' end of the DNA. The radioactive DNA was then purified as described below.

2.2.3.3 Purification

After labelling the DNA was run on a 0.3 mm-thick polyacrylamide gel to purify it from the enzymes and other components. For 3'-labelling, 20 µl of loading dye (20% Ficoll, 10 mM EDTA and 0.005% w/v bromophenol blue) was added and the sample loaded onto an 8% non-denaturing polyacrylamide gel (40 cm long) and run at 800 V for at least 90 minutes (until the dye reached the bottom of the gel). For 5'-labelled single stranded DNA fragments 10 µl of DNase I stop solution (80% formamide, 10 mM EDTA, 10 mM NaOH and bromophenol blue) was added and the sample was heated to 95 ºC for 3 minutes. It was then loaded onto a 14% denaturing polyacrylamide gel (containing 8 M urea) and run at 1500 V for around 90 minutes.

After running, the wet gel was exposed to X-ray film for about 3 minutes to locate the position of the radioactive DNA and the band was excised with a razor blade. The DNA was eluted from the gel slice by submerging it in 300 µl of elution buffer containing 10 mM Tris-HCl pH 7.5 and 10 mM EDTA. This was agitated overnight and the buffer containing the labelled DNA extracted. Approximately 1.3 ml of ethanol was added and the sample left on dry ice for 10 minutes to allow the DNA to precipitate, then centrifuged at 15000 rpm for 10 minutes. In early experiments 30 µl of 3 M sodium acetate was used in addition to ethanol to aid precipitation. However this occasionally produced a pellet of excess salt during the centrifugation and it was discovered that omitting the sodium acetate did not unduly affect the efficiency of the precipitation. After centrifugation the supernatant was removed and checked to ensure it contained no radioactive DNA

which has not precipitated. The purified DNA pellet was washed with 70% ethanol and dried. The DNA was re-dissolved in 10 mM Tris-HCl pH 7.5 containing 0.1 mM EDTA, at a concentration of about 10 cps per microlitre as estimated using a hand held Geiger counter.

When labelling the primer and the TFO some problems were encountered during purification. As a result of their short length the DNA did not precipitate well and the supernatant still contained radiolabelled DNA. To overcome this, the DNA was split into two or more reaction tubes before the addition of ethanol so that more ethanol could be added than normal at this point. The DNA was then left on dry ice for at least 30 minutes (rather than 10), and spun for 15 minutes instead of 10. If any of the supernatant still contained radioactivity it was placed in a separate tube with more ethanol and the process was repeated.

2.2.4 Footprinting

DNase I footprinting is a technique which can be used to detect the binding sites and binding affinity of TFOs. This technique was used in the initial stages of the project to investigate the binding of the bis-amino-U containing TFO on the *tyr*T fragment (chapter 3). Footprinting experiments were also carried out on sequences isolated from REPSA (chapter 4) and to investigate the specificity of the anthraquinone containing TFOS (chapter 5).

2.2.4.1 Incubation

The TFO of interest was diluted in an appropriate buffer to give a range of six different concentrations which were 150% of the desired final concentration. 3 µl of each of the TFO dilutions was combined with 1.5 µl of radiolabelled DNA and incubated to allow binding (usually overnight) at 20 ºC, although 1 hour incubations were also tested as described in the results chapters. The majority of experiments were performed at pH 5.0 in 50 mM sodium acetate. Any other buffers used are detailed with the results. Magnesium chloride was added to the buffer to improve the binding for some experiments; concentrations are detailed with the results. The naphthylquinoline triplex-binding ligand was also used for some experiments to enhance the affinity (see section 1.5.1) (282-284). When this ligand was used the reaction mix was as follows; 1.5 µl of TFO at 300% of the

desired concentration, 1.5 µl of ligand at 300% desired concentration and 1.5 µl of duplex target. This was then incubated and digested in the same way.

2.2.4.2 Digestion

After incubation the template DNA was cleaved using the non-specific enzyme DNase I, in order to create fragments of DNA of different lengths. 2 µl of 0.01 units/ml DNase I in DNase I buffer (20 mM NaCl, 2 mM $MgCl_2$ and 2 mM $MnCl_2$) was added to each of the incubation mixtures, and the reaction was terminated after 2 minutes by adding 4 µl of DNase I stop solution. By using a low concentration of enzyme and limiting the digestion time, on average each duplex strand was cleaved only once (single hit kinetics). In this way a population of radioactive template DNA strands cut are at every possible length should be created.

2.2.4.3 GA tract and control lane

In addition to the 6 reaction tubes containing different concentrations of TFO, a GA marker lane and control lane were included in each footprint. The GA marker was prepared by mixing 1.5 µl of the radioactive DNA fragment, 5 µl of DNase I stop solution and 20 µl of dH_2O. This was heated at 100 ºC for 30 minutes, with the Eppendorf cap open, before cooling on ice. This process causes the DNA template to be cleaved at purine sites only. The control lane (in the absence of added oligonucleotide) was prepared by combining 1.5 µl of radiolabelled DNA fragment with 3 µl of buffer. This was digested in the same way as the TFO-containing mixtures.

2.2.4.4 Polyacrylamide gel

The products of the DNase I digestion were run on 40 cm long, 0.3 mm thick denaturing polyacrylamide gels containing 8 M urea. Footprints with the *tyr*T and *tyr*T-based templates were run on 9% polyacrylamide gels (chapters 3 and 5), while those using the REPSA fragments on 12.5% as they were shorter (chapter 4). After cleavage the samples were heated to 100 ºC for 3 minutes then cooled on ice before loading onto the gel. The gels were run at 1500 V for between 1.5 and 2 hours, depending on percentage acrylamide. They were then fixed in 10%

(v/v) acetic acid and transferred to 3 MM Whatman paper for drying under vacuum at 86 ºC for 1-2 hours. Once dried the gels were imaged using a Molecular Dynamics Storm 860 Phosphorimager.

2.2.5 pH jump experiments

The footprinting method detailed in 2.2.3 was altered slightly to carry out experiments in which the pH of the sample was rapidly increased (pH jump experiments - chapter 3). Rather than using a range of TFO concentrations the same concentration was used in every lane. The TFO and template were incubated together at pH 5.0 to allow binding, either overnight or for one hour. 1 µl of Tris-HCl pH 8.0 was then added to jump the pH to around pH 7.0. Samples were taken at various time points after the jump and were digested with DNase I as described in section 2.2.4.2. One sample was digested before the pH jump.

2.2.6 Site directed mutagenesis

As mentioned in section 2.1.1 several new templates were generated from the *tyr*T(43-59) fragment by site directed mutagenesis for use in Chapter 5. The position and nature of these mutations is shown in Figure 2.7 on the following page. *Tyr*T42A, *tyr*T42A50T and *tyr*T41AT were generated in the course of this thesis by site directed mutagenesis. The *tyr*T50T template was generated previously by D.A.Rusling *et al* by site directed mutagenesis of *tyr*T(43-59) and was used as a template for some of the mutagenesis experiments in this thesis (294).

Quick-Change site directed mutagenesis was used to mutate the *tyr*T(43-59) and *tyr*T50T templates to create three additional templates for footprinting of the anthraquinone TFOs used in chapter 5, these were *tyr*T42A, *tyr*T42A50T and *tyr*T41AT. This technique involves designing primers which contain the desired mutation(s) then using PCR to synthesise the desired template. To mutate the plasmids, 125 ng of each pair of primers was mixed with 1 µl of the start plasmid (*tyr*T(43-59) or *tyr*T50T) diluted 1 in 10 from a standard mini-prep, 1 µl of 25 mM dNTPs, 5 µl 10x *Pfu* buffer and 2-3 units of *Pfu* DNA polymerase (Promega). This mixture was made up to 50 µl with distilled water and cycled in a PCR block under the following conditions;

Step 1	94°C	5 minutes
Step 2	94°C	30 seconds
	55°C	1 minute
	68°C	3 minutes
Cycle step 2 16 times		
Step 3	4°C	hold

10 units of *Dpn*I was then added to digest the methylated template DNA and the sample was incubated at 37°C for 90 minutes. The open-circle mutated plasmids were then transformed into competent *E.*coli TG2 cells and plated onto carbenicillin-containing plates as described in section 2.2.1. Colonies were picked and sequenced as described in section 2.2.7.2 to confirm the presence of the desired mutation.

(i)

```
tyrT            5′ – AACCAGTTCTTTTTTCTCTTCCTAACA – 3′
                3′ – TTGGTCAAGAAAAAAGAGAAGGATTGT – 5′

tyrT50T         5′ – AACCAGTTCTTTATTCTCTTCCTAACA – 3′
                3′ – TTGGTCAAGAAATAAGAGAAGGATTGT – 5′

tyrT42A         5′ – AACCATTTCTTTTTTCTCTTCCTAACA – 3′
                3′ – TTGGTAAAGAAAAAAGAGAAGGATTGT – 5′

tyrT42A50T      5′ – AACCATTTCTTTATTCTCTTCCTAACA – 3′
                3′ – TTGGTAAAGAAATAAGAGAAGGATTGT – 5′

tyrT41AT        5′ – AACCTATTCTTTTTTCTCTTCCTAACA – 3′
                3′ – TTGGATAAGAAAAAAGAGAAGGATTGT – 5′
```

(ii)

Template	Primers
tyrT42A	5′ AGAGAAAAAAGAAATGGTTGCGTAATTTTC 3′ 3′ TCTCTTTTTTCTTTACCAACGCATTAAAAG 5′
tyrT42A50T	5′ AGAGAATAAAGAAATGGTTGCGTAATTTTC 3′ 3′ TCTCTTATTTCTTTACCAACGCATTAAAAG 5′
tyrT41AT	5′ GTTAGGAAGAGAAAAAAGAATAGGTTGCGTAATTTTCATC 3′ 3′ CAATCCTTCTCTTTTTTCTTATCCAACGCATTAAAAGTAG 5′

(iii)

```
5′ -AATTCCGGTTACCTTTAATCCGTTACGGATGAAAATTACGCAACCUWTTCTTTYTTC
3′ ---AAGGCCAATGGAAATTAGGCAATGCCTACTTTTAATGCGTTGGVXAAGAAAZAAG

TCTTCCTAACACTTTACAGCGGCGCGTCATTTGATATGAAGCGCCCCGCTTCC----- 3′
AGAAGGATTGTGAAATGTCGCCGCGCAGTAAACTATACTTCGCGGGGCGAAGGGCTC- 5′
```

Figure 2.7: (i) - Sequences of the five footprinting templates used in Chapter 5, with TFO binding sites shown in red.
(ii) – Sequence of primers used to generate mutated templates with the desired mutation shown in red. tyr*T42A and* tyr*T41AT were generated using* tyr*T(43-59) as the start template, while* tyr*T42A50T used* tyr*T50T as the start template.*
(iii) – Sequence of the tyr*T(43-59) fragment showing the positions of mutations (red) to generate the footprinting templates used in Chapter 5. A TA base pair was introduced at position UV (to generate template tyrT41AT). TA (tyrT41AT) or AT (tyrT42A and tyrT42A50T) base pairs were introduced at WX and a TA base pair had previously been introduced at YZ (tyrT50T and tyrT42A50T) (294).*

2.2.7 Primer extension

In many of the experiments carried out in this thesis it was necessary to generate duplex DNA by primer extension using a chemically synthesised single stranded DNA template and primer to create the second strand. For template and primer sequences see Figure 2.3 on page 61. Because this contains a random region, simply annealing two complementary strands together was not possible as this would generate many duplexes that contain one or more mismatched base pairs. It was therefore necessary to copy the single stranded template in order to generate a perfectly paired duplex.

A complementary primer was allowed to anneal to the single stranded template (2.4 µM) and 5 units of *Taq* polymerase with *Taq* Buffer was used to extend from the primer using dNTPs (1 mM) to create a complementary strand of DNA. The primer concentration was varied and was not quantified as it was radiolabelled prior to primer extension.

The mixture was heated in a PCR block to 94 ºC for six minutes then cooled to 50 ºC for 30 minutes to allow annealing of the primer and extension by *Taq* polymerase before finally cooling to 4 ºC. 20 µl of loading dye (20% Ficoll, 10 mM EDTA 0.02 g bromophenol blue) was added and the mixture was run on a 6.4% non-denaturing polyacrylamide gel and purified as in section 2.2.3.3. When a radiolabelled duplex template was required the primer was first labelled with [γ-^{32}P]-ATP using 5' labelling as described in section 2.2.3.2. The exact reaction conditions for primer extension were optimised after the first set of REPSA experiments. This optimisation is described in detail in section 2.3.1.

One of the key concepts for REPSA is that the population of starting template sequences should contain every possible sequence combination, so that every possible binding site for the TFO is represented. Therefore some calculations have been done to ensure that this is true.

First the number of different sequences possible in a 15 base pair long random sequence was calculated;

$4^{15} = 1.07 \times 10^{9}$

This means that a minimum of a billion template molecules should be used in order to have one copy of every possible sequence present for REPSA selection. Next the number of template molecules in 1 µl of template (47 µM) was calculated;

$(1x10^{-6}) \times (47x10^{-6}) = 47$ pmoles

$(47x10^{-12}) \times (6.023x10^{23}) = 2.8x10^{13}$ molecules

So there are over 10000 times more molecules than needed for every possible sequence to be represented. This means that even if some DNA is lost during the purification process after primer extension every possible template sequence should still be represented when the TFO is introduced and every binding site should be represented if DNA synthesis was random.

2.3 REPSA

Two REPSA protocols have been used in this thesis; an initial un-optimised method was used for the first set of REPSA experiments. Three sequences selected using this initial protocol (shown in Figure 2.8 below) were footprinted with the BAU containing TFO (Figure 2.2 on page 60). The TFO failed to produce a footprint on any of the templates, indicating that this REPSA protocol had not worked.

5′-GTAGGATCCTGACTGGATGAA**TTAGGGCCATGTTAC**TCTTCCTGACTTGCTTCAG-3′
3′-CATCCTAGGACTGACCTACTT**AATCCCGGTACAATG**AGAAGGACTGAACGAAGTC-5′

5′-GTAGGATCCTGACTGGATGAA**ATCGATTGTAGGGA**TCTTCCTGACTTGCTTCAG-3′
3′-CATCCTAGGACTGACCTACTT**TTGGTAACATCCCT**AGAAGGACTGAACGAAGTC-5′

5′-GTAGGATCCTGACTGGATGA**TCGATCCCTGA**TCTTCCTGACTTGCTTCAG-3′
3′-CATCCTAGGACTGACCTACT**AGCTAGGGACT**AGAAGGACTGAACGAAGTC-5′

Figure 2.8: Sequences of three templates from the first set of REPSA experiments, original random region shown in bold.

After this REPSA protocol showed no obvious selection, experiments were carried out to optimise the *Fok* I cleavage step of the protocol. This type of experiment is generally referred to as REPA (restriction endonuclease protection assay). The second protocol is based on the results of these optimisation experiments. More detail about the optimisation of REPSA reaction conditions can be found in section 2.3.3. The method detailed below is the improved method

based on the optimisation experiments. Differences between this method and the original are described in section 2.3.2.

2.3.1 Optimised REPSA Protocol

After primer extension (see section 2.2.6) with radiolabeled primer the double stranded template containing the random sequence of bases was separated from the other elements of the reaction mixture on a 6.4% non-denaturing polyacrylamide gel. The DNA was then extracted from the gel slice as described in section 2.2.3.3 and resuspended in 10 µl 1x *Fok* I Buffer (50 mM KAc, 20 mM Tris-Ac, 10 mM MgAc, 1 mM DTT). 10 µl of 10 µM TFO was then added and the mixture was left for an hour at 20°C to equilibrate. Any unbound templates were then cleaved with 8 units of *Fok* I (NEB) at 37°C for 5 minutes after which the enzyme was deactivated by heating to 90°C for 3 minutes. Theoretically sequences which were protected by binding to the TFO were not cleaved by the enzyme, these templates were then amplified by PCR; whereas cleaved templates (those to which the TFO did not bind) were not amplified. To amplify the un-cleaved templates 10 µl of the cleavage mixture was combined with 15 µl of labelled forward primer, 3 µM unlabelled reverse primer, 1 µl dNTP mix (25 mM), 5 µl of 10x *Taq* Buffer and 5 units of *Taq* polymerase, the solution was made up to 50 µl with water. This mixture was then cycled in a PCR block under the following conditions;

Step 1	92°C	30 seconds
	50°C	3 minutes
Cycle step 1 15 times		
Step 2	4°C	hold

At this point a further 1 µl of dNTPs and 5 units of *Taq* were added along with 2.5 µl of unlabelled forward primer. This procedure ensures maximum utilisation of the labelled primer while preventing PCR products from annealing together (which could generate mismatched duplexes from the mixed population). Steps 1 and 2 were then repeated for two further rounds to ensure that all templates were double stranded.

Once amplified, 20 µl of loading dye (20% Ficoll, 10 mM EDTA, bromophenol blue) was added and the DNA was run on an 8% non-denaturing

polyacrylamide gel. The full length regenerated templates were then purified as described for preparing radiolabelled DNA in section 2.2.3.3.

The purified templates were re-dissolved in 10 µl 1x *Fok* I buffer and 10 µl of 10 µM TFO was added, the whole process was then repeated 10 times to obtain optimum selection. If too many rounds of REPSA are performed the template can start to lose bases from the random region; samples were taken after seven rounds to ascertain if enough selection had taken place and it was decided to carry out an additional three rounds. Once REPSA had been completed, the selected templates were cloned into plasmids, transformed into competent cells (see below) and plated onto blue/white selection plates (section 2.2.7.1). The colonies that grew each contained a different selected template which was then sequenced.

2.3.1.1 Cloning and Transformation

After the final round of selection the remaining templates were re-dissolved in 20 µl of dH_2O. At the same time *Bam*HI-cut pUC19 was prepared by mixing 1 µg of pUC19 with 19 µl of dH_2O. The REPSA products and pUC19 were each mixed with 2 µl of buffer E (Promega) 10 units *Bam*HI and 10 units *Hin*d III at 37 °C for 1 hour before the DNA from both digests was precipitated with ethanol (as described in section 2.2.3.3). The plasmid and REPSA fragments were each re-dissolved in 10 µl dH_2O and were mixed together, giving a final volume of 20 µl. 2 µl of 10x ligase buffer (Promega) and 2 units of DNA ligase (promega) were then added and the DNA left to ligate overnight. The ligated plasmid was mixed with competent TG2 cells (section 2.2.1), placed on ice for 30 minutes and then heat-shocked at 45 °C for 1 minute to allow the plasmids to enter the cells. The cells were then plated out onto blue-white selection plates and placed in at 37 ºC overnight for colonies to grow (section 2.2.1). Only cells which have taken up the plasmid are able to grow on the carbenicillin containing plates and only those plasmids which contain an insert appear as white colonies. Bacteria containing unmodified pUC19 will be blue. The white colonies were carefully picked off and grown overnight in liquid culture before purifying the DNA (section 2.2.2) and sequencing.

2.3.1.2 Sequencing

Once the DNA from a particular white colony had been grown and purified it was sequenced using a USB T7 sequencing kit. 40 µl of DNA from the purification was denatured using 10 µl of 2 M NaOH at room temperature for 20 minutes. The denatured DNA was then precipitated by addition of 15 µl 3 M NaOAc and 1 ml ethanol. This was left on dry ice for at least 10 minutes then spun at 13000 rpm for 15 minutes. The supernatant was removed and the pellet washed with 100 µl of 70% ethanol. The supernatant was removed and the samples were dried in a speed vac for 3 minutes before being re-dissolved in 10 µl dH_2O. 2 µl each of annealing buffer and universal primer were added (USB T7 kit). The primer was left to bind for 20 minutes at 37 °C then 10 minutes at room temperature. For each DNA sample 4 microcentrifuge tubes were prepared, each containing 2.5 µl of the 4 different dideoxynucleotides mixtures (kit). A polymerase mix was also made up containing 1 µl [α-^{32}P]-dATP, 7 µl of dH_2O and 12 µl of label mix A (kit). 12 units of T7 polymerase (kit) and 6.5 µl of enzyme dilution buffer (kit) were mixed and added to the polymerase mix. 6 µl of the polymerase mixture was then added to each of the four annealed DNA samples and left at room temperature for 5 minutes. 4.5 µl of this mixture was added to each of the dideoxynucleotide mixtures and left for 5 minutes at 37 °C before the reaction was stopped with 5 µl of the stop solution. The DNA was then denatured by heating to 95 °C for 3 minutes before running on a 9% denaturing polyacrylamide gel. This was then fixed in 10% (v/v) acetic acid, dried and imaged (section 2.2.4.4).

Each of the dideoxynucleotides causes the polymerase reaction to stop at a different nucleotide and by carefully controlling the incubation time and concentrations of the various components the reaction can be stopped at every point where a particular base is present. The position of the different bases can then be read off the sequencing gel, from 5'-3', bottom to top of gel.

2.3.1.3 Footprinting

Footprinting experiments were carried out on DNA fragments that were derived from some of the sequences isolated from the REPSA experiments. The method for this was essentially the same as that discussed earlier (section 2.2.4) with a few differences. After purification, the plasmid containing the sequence for footprinting was cut with *Eco*RI and labelled with [α-^{32}P]-dATP using reverse

transcriptase (section 2.2.3.1). The *Eco*RI and reverse transcriptase were then inactivated by heating at 65°C for 5 minutes, before releasing the fragment of interest by cutting with *Hin*dIII. The labelled DNA was then run on a non-denaturing polyacrylamide gel and the fragment purified as previously described (section 2.2.3.3).

When footprinting with these labelled sequences it was also necessary to increase the percentage of the gel as the sequences are shorter and a 12.5% denaturing gel gave optimum separation of the bands.

2.3.2 Original REPSA protocol

The REPSA method previously described incorporated the improved Fok I digestion protocol which resulted from the Fok I optimisation experiments (section 2.3.3.3). The original Fok I digestion protocol and other minor differences between the two protocols are described here. Anything not discussed was the same for both protocols.

To ensure a high yield of radiolabelled template from primer extension the concentration of radiolabelled primer was increased in the second set of experiments. The original Fok I protocol was as follows; rather than resuspending the double-strand template DNA in 10 µl of 1x Fok I buffer, 18 µl of pH 5.0 buffer (50 mM Sodium Acetate) was used. 2 µl of 5 µM TFO was added to this and the mixture left for 1 hour at room temperature to equilibrate. 2 µl of 10x Fok I Buffer was then added with 3 units of Fok I and heated at 37°C for 5 minutes. The Fok I was then deactivated by heating to 90°C for 3 minutes. As for primer extension a higher concentration of radiolabelled primer was used for PCR.

2.3.3 REPSA optimisation

2.3.3.1 Primer extension

Primer extension was used to generate a duplex DNA template from a single strand containing a random region of nucleotides (Figure 2.3 page 61) using a primer, as described in section 2.2.6. Some optimisation was carried out before any REPSA experiments were carried out ensure generation of a high yield of duplex template.

Primer extension is similar to PCR and consists of binding a complementary primer to the non-random region of the single stranded template, then using Taq polymerase to extend from the primer with dNTPs to create a complementary strand of DNA. For REPSA the duplex template needed to be radioactive, so the primer was first labelled with [γ-^{32}P]-ATP using 5' labeling as described in section 2.2.3.2. The sequences of the single stranded random template and the radiolabelled primer are shown in Figure 2.3 on page 61.

Initial attempts to synthesise the second DNA strand used one round of a standard PCR heating and cooling protocol, but these produced a poor yield of duplex DNA when run on an agarose gel. It was thought that this might be due to the primer dissociating at the higher extension temperature. It was decided to omit the extension step and simply have a longer anneal step to allow the primer to bind and the Taq to extend. The following adapted protocol was designed to test this theory, with three different anneal/extension temperatures:

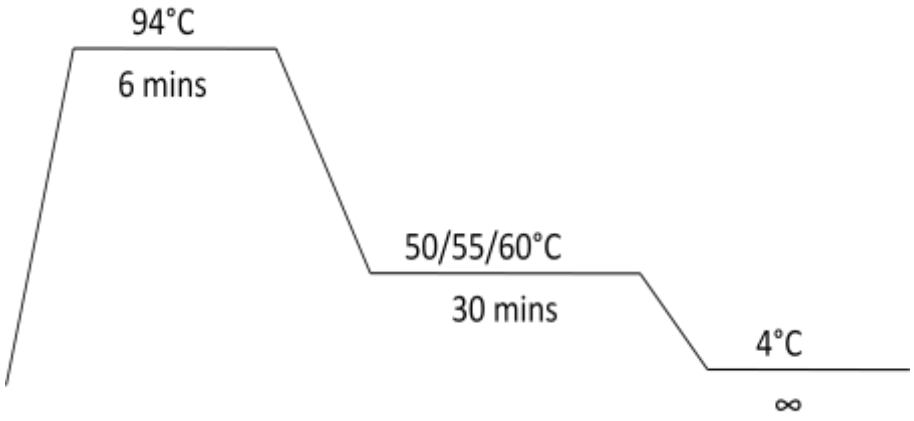


Figure 2.9: Primer extension protocol; an initial denaturing stage was followed by 30 minutes of combined annealing and extension followed by cooling.

Three versions of this protocol were tested with different annealing and extension temperatures indicated, based on the Tm of the primer (around 55°C). The products of these experiments were run on an 8% non-denaturing polyacrylamide gel along with two controls. The first control contained template and radiolabelled primer which had been allowed to anneal using the protocol above (55°C); but without *Taq* polymerase present; this therefore represents annealed primer and template. The other control contained only radiolabelled primer. The results are shown in Figure 2.10 below:

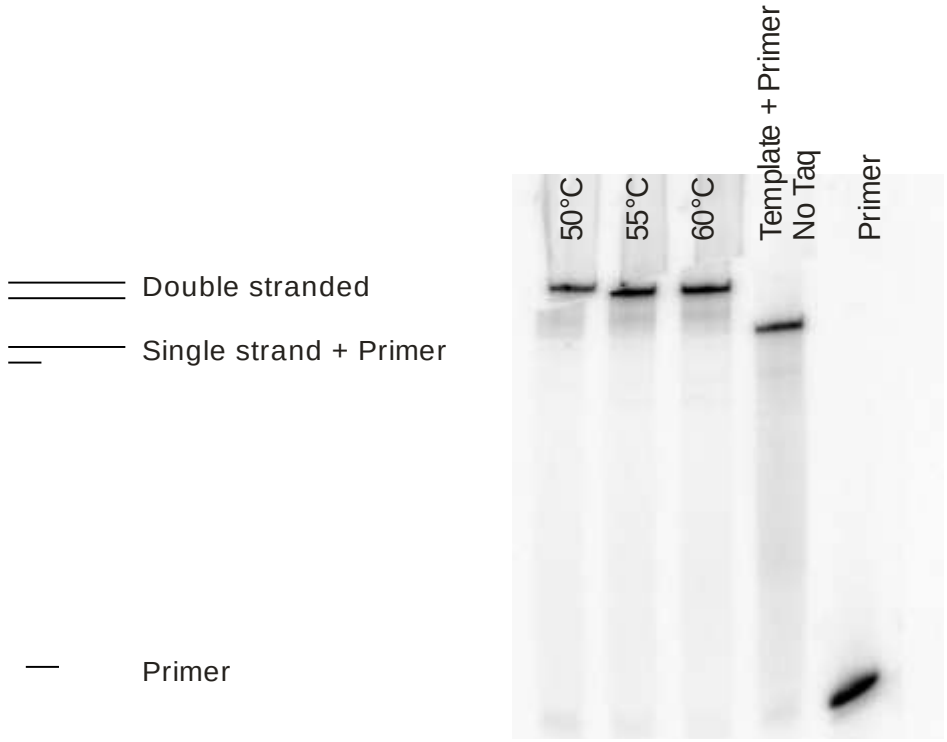


Figure 2.10: Phosphorimage of non-denaturing polyacrylamide gel showing results from the three different protocols shown in Figure 2.9.

From this gel it was evident that the primer extension protocol was working with high efficiency and that the three different anneal/extension temperatures gave identical results. The products of primer extension ran with lower mobility than just the template and primer annealed together, confirming that the *Taq* polymerase had synthesised a second strand of DNA. For a detailed protocol see section 2.2.6.

2.3.3.2 PCR

As with primer extension this optimisation was carried out before the first set of REPSA experiments. Several different PCR protocols were tested to ensure a good yield of amplification after the REPSA cleavage. Three different temperature profiles were tested, each with 6, 10 or 15 cycles. The first (A) is based on the primer extension profile shown in Figure 2.10 as the same template and primers were used. Profiles B and C are based on previously used PCR protocols, with different annealing times.

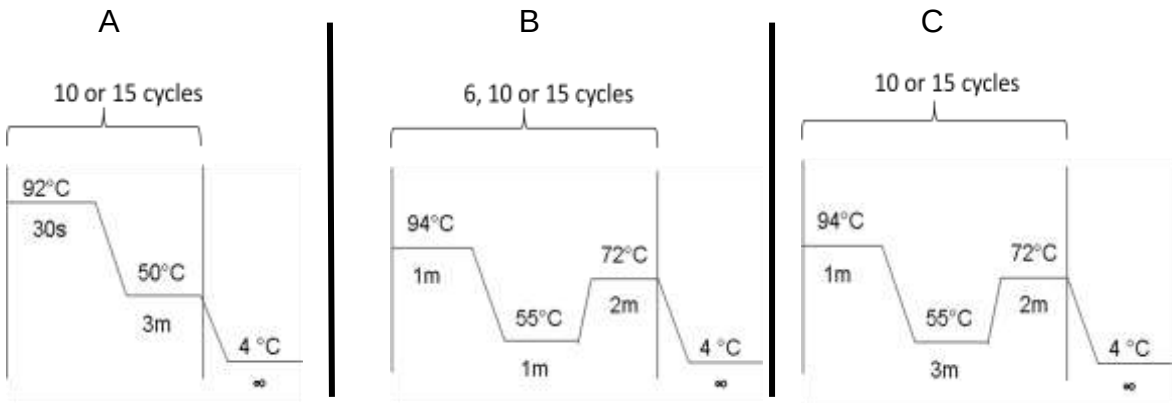


Figure 2.11: PCR protocols tested (results shown in Figure 2.12 below).

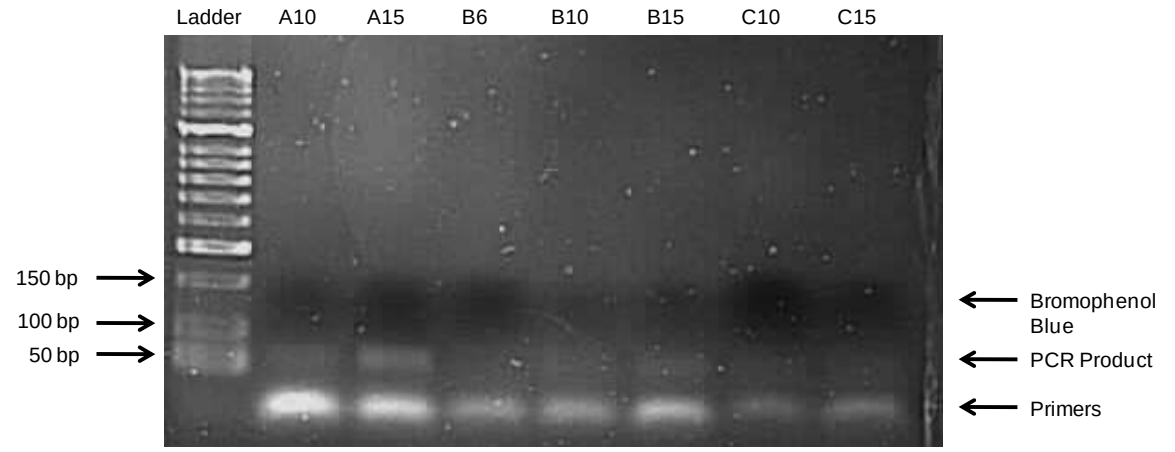


Figure 2.12: 1% Agarose gel showing DNA from different PCR protocols, the left hand lane contained a 50 bp DNA ladder. Labels at the top indicate which PCR protocol each sample was from.

From the agarose gel in Figure 2.12 above, it is clear that lane three (profile A with 15 cycles) gives the highest yield, with 10 cycles giving a slightly lower yield. Therefore this PCR protocol with 15 cycles was used in REPSA to amplify the uncut templates after *Fok* I digestion. It is important not to over-amplify the products of *Fok* I digestion as this may generate mismatched duplexes if product anneals to product, therefore a protocol with more than 15 cycles was not tested and the last step of the protocol was an extension rather than annealing step. For the exact protocol used in REPSA see section 2.2.7.

2.3.3.3 Fok I

A key aspect of REPSA was the *Fok* I cleavage of the template DNA. This needed to be optimised so that the enzyme cleaved a high percentage of unbound templates (but not the bound ones) in both strands of the target. Also the TFO concentration needed to be optimised so that it would bind target sites with sufficient affinity to protect from *Fok* I cleavage. These optimisation experiments were carried out after the first set of REPSA experiments. For these latter experiments a target sequence was used, which contained the exact binding site for the TFO. The sequence is shown in Figure 2.4 on page 61.

Both strands of this template were synthesised separately and a double strand template created by annealing of the two complementary strands. The TFO used in these experiments was the bis-amino-U containing TFO shown in Figure 2.2 on page 60.

The first experiment examined the time course of *Fok* I cleavage of the target DNA in the same conditions used for the first set of REPSA experiments. The exact DNA target was radiolabelled and cut with *Fok* I, samples were taken at different times after the start of the reaction and the reaction was stopped by addition of 5 µl of DNase I stop solution. The results of this experiment are shown below.

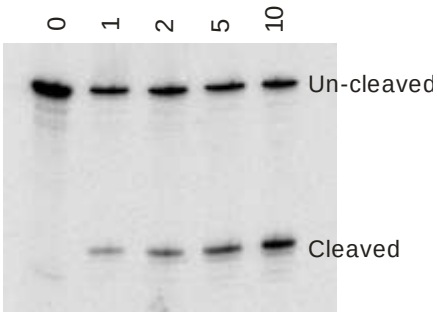


Figure 2.13: 9% denaturing polyacrylamide gel showing samples of Fok *I digested DNA, digestion times in minutes at the top. Left hand lane contains a control of uncut DNA.*

From this experiment it can be seen that *Fok* I had begun to cleave the DNA after just one minute, however even after 10 minutes only around 50% of the targets had been cleaved. Therefore different reaction conditions needed to be tested in order to obtain optimal cleavage.

Different buffers and incubation times were tested in an attempt to discover conditions in which the TFO would bind to the DNA with high affinity, and the *Fok* I

would still cleave a high percentage of unbound templates. In ideal conditions > 90% of the targets would be cleaved without TFO present, and < 10% cleaved when TFO was added. For these experiments an exact version of the random template was generated by primer extension on a single strand in the same way as the random template, in order to obtain a similar concentration after gel purification.

Five different buffers were tested;

Buffer	Composition
1	50 mM Sodium Acetate, pH 5.0
2	10 mM PIPES 50 mM NaCl, pH 6.5
3	10 mM Tris-HCl 50 mM NaCl, pH 7.0
4	0.1 mM EDTA, 10 mM Tris
5	50 mM KAc, 20 mM Tris-Ac, 10 mM MgAc, 1 mM DTT

Table 2.1: Buffers tested to optimise Fok I cleavage. Buffers 1, 2 and 3 are buffers commonly used in footprinting, Buffer 4 is the DNA template re-suspension buffer and Buffer 5 is the same as the New England Biolabs Fok I Buffer (1x).

The footprinting buffers (1, 2 and 3) and buffer 1 in particular should allow good TFO association, however they may not be optimal for *Fok* I cleavage. Buffer 5 should be optimal for *Fok* I cleavage but may not be suitable for triplex formation. For each buffer tested two experiments were carried out. The first consisted of just the template and *Fok* I incubated together for different lengths of time to test *Fok* I cleavage of unbound template. In the second, the template was first allowed to associate with the TFO for one hour in the buffer being tested before cleavage with *Fok* I. The template and 5 µM TFO / buffer were incubated for one hour at 20°C before *Fok* I cleavage.

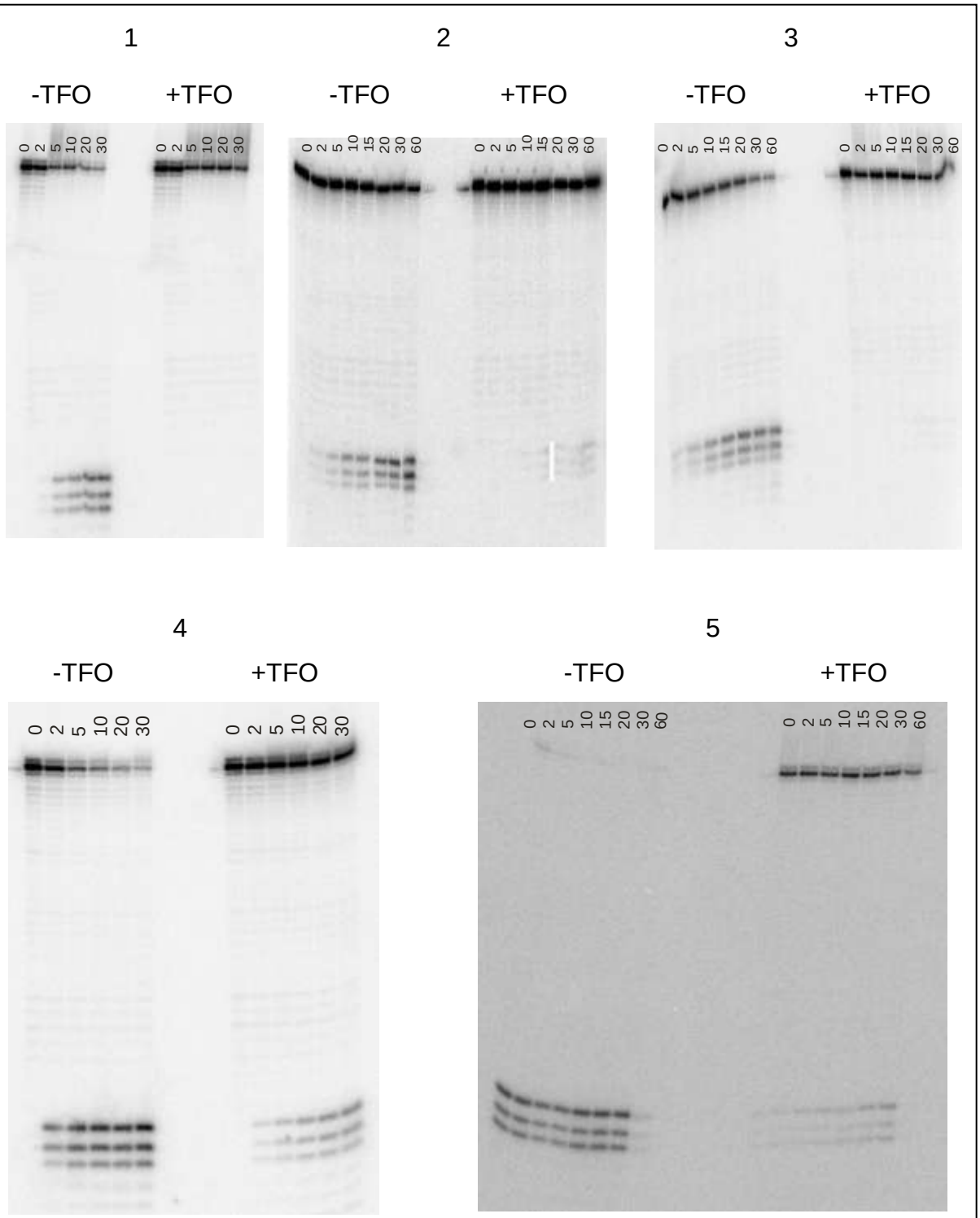


Figure 2.14: Fok *I cleavage patterns of the exact template with or without the 9mer TFO.* Fok *I digestion time in minutes is shown at the top. The buffer used is indicated above the gels.*

The three sets of cleavage gels with buffers 1, 2 and 3 all shown similar results; good protection by the TFO for the whole hour, but only some cleavage (generally < 50%) in the absence of TFO. This indicates that the triplex is forming but the *Fok* I enzyme is not working efficiently under these conditions. With buffer 4 cleavage in the absence of TFO was much higher, about 70-80% after 60 minutes; however the TFO failed to protect the template from cleavage even after only two minutes.

As expected, the Magnesium ions in buffer 5 gave the best cleavage of the template alone, and in the presence of the TFO there was only a small amount of cleavage after 5 minutes. It was therefore decided to use this buffer in the next set of REPSA experiments; although some bound templates may be cleaved it was felt that it was more important to eliminate a high proportion of unbound templates.

A footprinting experiment was also carried out to test different TFO concentrations in order to find the optimal concentration for preventing cleavage by *Fok* I. This footprinting experiment was carried out in buffer 5 using the *tyr*T fragment and five different TFO concentrations

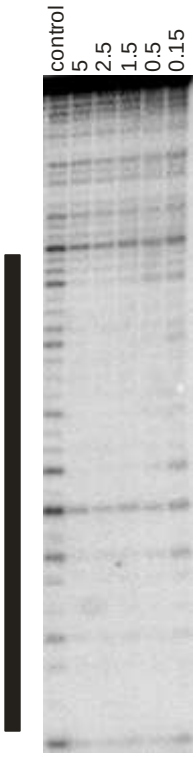


Figure 2.15: DNase I cleavage pattern of exact REPSA template with the 9mer TFO. TFO concentrations in μM are shown at the top of each lane. A control lane shows the DNase I cleavage pattern in the absence of TFO. The filled box shows the position of TFO binding site.

From this footprint it is clear that the TFO binds down to concentration as low as 0.15 µM in this buffer. In order to ensure that secondary binding sites were also selected (which occur at higher concentrations) a 5 µM concentration was used for selected experiments with REPSA. For the exact reaction conditions used in REPSA see section 2.2.7.

2.4 Band Shifts

As a secondary project to REPSA, an assay method based on gel retardation was also attempted. This involved the separation of triplex from duplex DNA on a polyacrylamide gel in order to obtain the sequences to which the TFO had bound. It relies on the principle that a DNA triple helix will have different gel mobility to that of duplex DNA. The purpose of this technique was the same as REPSA; to determine the (secondary) binding sites of a TFO. It was performed with the same TFOs in the hope of confirming the REPSA results. Unfortunately, despite many attempts to optimise the conditions the two DNA species did not separate sufficiently to obtain reliable sequences. The method and different reaction conditions tested are discussed in this section.

A population of sequences was created using a template containing a random central region, as with REPSA. Every possible target sequence was represented within this population, so a TFO would bind to only a small fraction of the targets. These sequences were chemically synthesised as single strands and primer extension was used with a radiolabelled primer to create duplex DNA (as with REPSA). The TFO was mixed with this population of sequences and allowed to bind overnight before being run on a non-denaturing polyacrylamide gel to separate the triplex from duplex DNA. The percentage of the gel and the buffer were varied during the course of the investigation in an attempt to obtain maximum separation. In the preliminary experiments the gel was fixed and dried, as described for footprinting experiments, then imaged using a Molecular Dynamics Storm 860 Phosphorimager. In the final experiment the band containing the triplex DNA was excised from the wet gel and purified for cloning and sequencing. The initial optimisation experiments all used the exact TFO template rather than the random one so that the degree of separation could be easily visualised.

For the initial experiment to test this method the exact template was radiolabelled and incubated with 3 µM TFO overnight at room temperature to allow triplex formation. After addition of loading dye (20% Ficoll, 10 mM EDTA and

bromophenol blue) the mix was loaded onto a 4.5% non-denaturing polyacrylamide gel and run until the dye front reached the bottom. The result of this gel is shown in Figure 2.16 I on page 88.

It is clear from this gel that a triplex had been formed and had a lower mobility than the duplex DNA; however the separation between the duplex and triplex was not very large. The percentage of the gel was therefore increased to 11% to attempt to get better separation. For this experiment the TFO was also radiolabelled to better visualise how the duplex and TFO were interacting. The labelled TFO was used as a tracer and the reaction was performed in the presence of 3 µM unlabelled TFO. The results of this experiment are shown in Figure 2.16 ii on the following page.

Unfortunately it appeared that increasing the gel percentage caused the triplex to disassociate in the gel as no triplex population was visible. In order to increase triplex stability within the gel several changes were made to the protocol. The TFO concentration was increased to 10 µM and 10 µM of the triplex-binding-ligand naphthylquinoline was added before incubation at 20°C overnight. The gel percentage was also decreased to 8% and the gel buffer was altered from TBE (pH 8.3) to Tris-acetate pH 6.0. This produced the result shown in Figure 2.16 iii on the following page.

There is a clear triplex population which is well separated from the position of the duplex. However the sample appears to have smeared on the gel. Several alterations were made to the method in attempt to reduce this smearing effect including filtering all gel components, spinning the samples to reduce the effect of excess salt and using hydrophobic ‘no-stick’ microcentrifuge tubes. None of these produced any significant improvements.

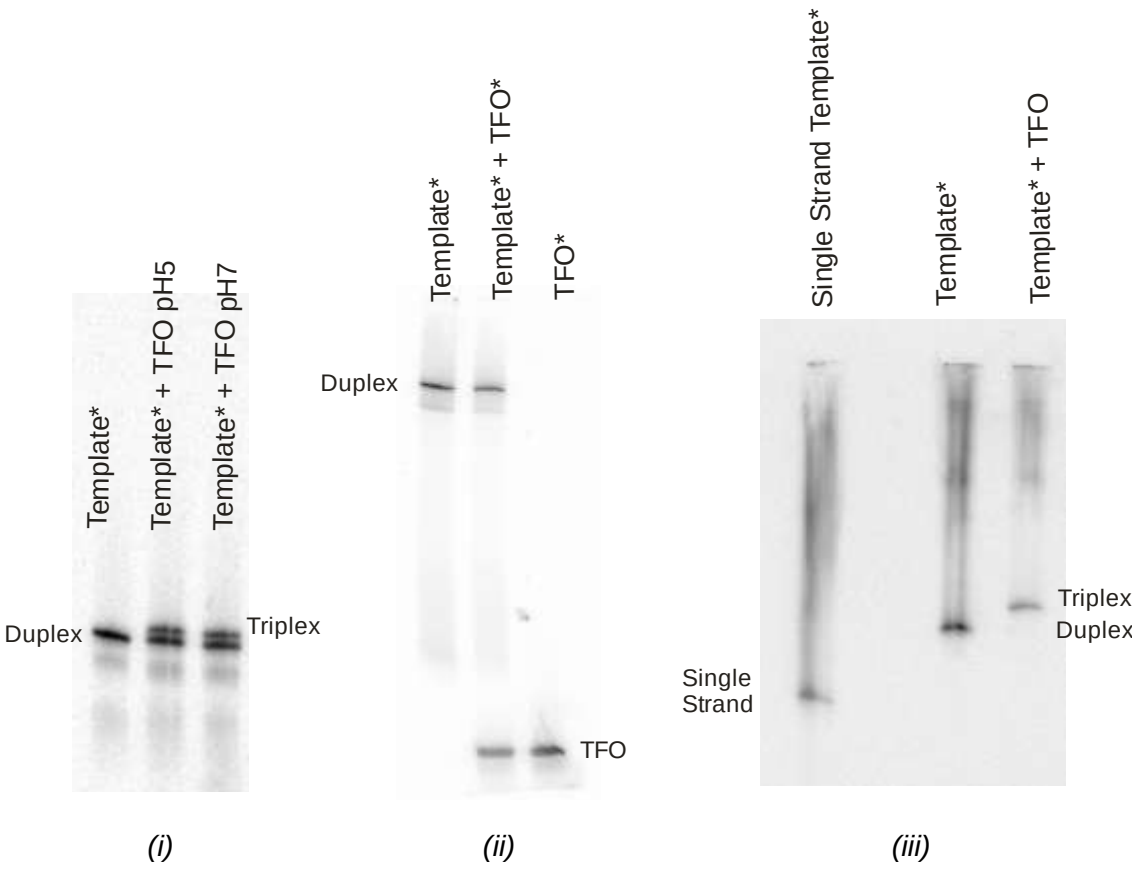


*Figure 2.16: Band shift gels i, ii and iii. Samples in each lane are labelled at the top, the * indicates that the DNA was radiolabelled.*

It was eventually decided to use the experiment C conditions and see if it was possible to isolate the bound templates from the gel. The template containing a random region was incubated with the TFO (10 µM) and naphthylquinoline ligand (10 µM) overnight at 20°C. This was then run on an 8% non-denaturing polyacrylamide gel containing Tris-acetate buffer pH 6.0. A control of the radiolabelled exact template and the TFO bound together under the same conditions was run alongside, so that the position of the triplex population could be determined (the fraction of the random sequence that would be bound by the TFO would be too low to visualise directly). Another gel was run using the same samples but in the cold room at 4°C in an attempt to further stabilise triplex formation. Two different TFOs were used for this final experiment; the TFO containing 6 bis-amino-U residues used in REPSA (Figure 2.17 i on the following page) and a second slightly longer TFO containing three bis-amino-U residues and three 2'-aminopyridine residues (Figure 2.17 ii);

i) 5′ – BBBBBBCBT – 3′

ii) 5′ – T2T2TBTBTB2T – 3′

2 = 2'-aminopyridine

B = BAU

Figure 2.17: Sequence of TFOs used in band shifts.

After running overnight the gels were exposed to X-ray film so that the position of the bands could be seen; these were developed and scanned; the results are shown in Figure 2.18 below.

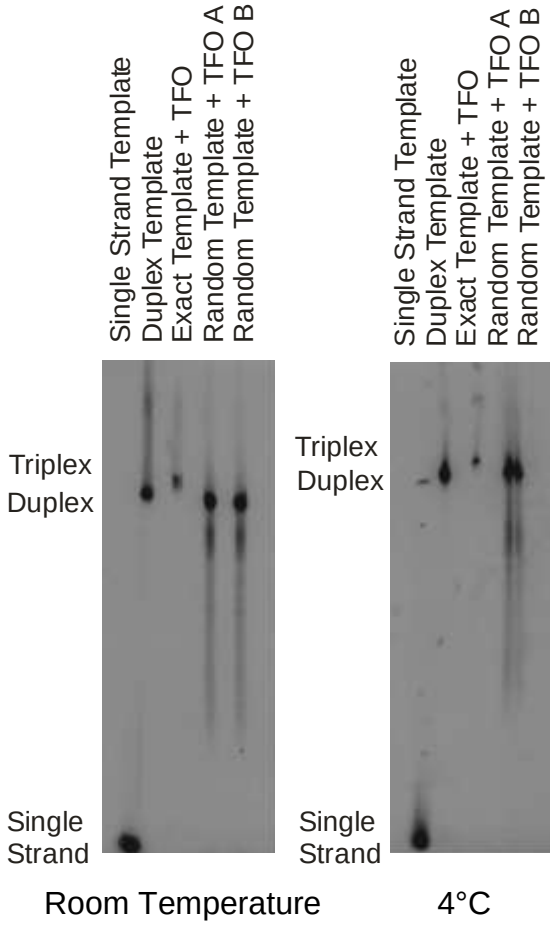


Figure 2.18: Scanned X-ray films from final Band Shift experiment. The single strand or duplex template was radiolabelled for all samples.

A variety of bands were excised from each of these gels, as shown in the diagram below.

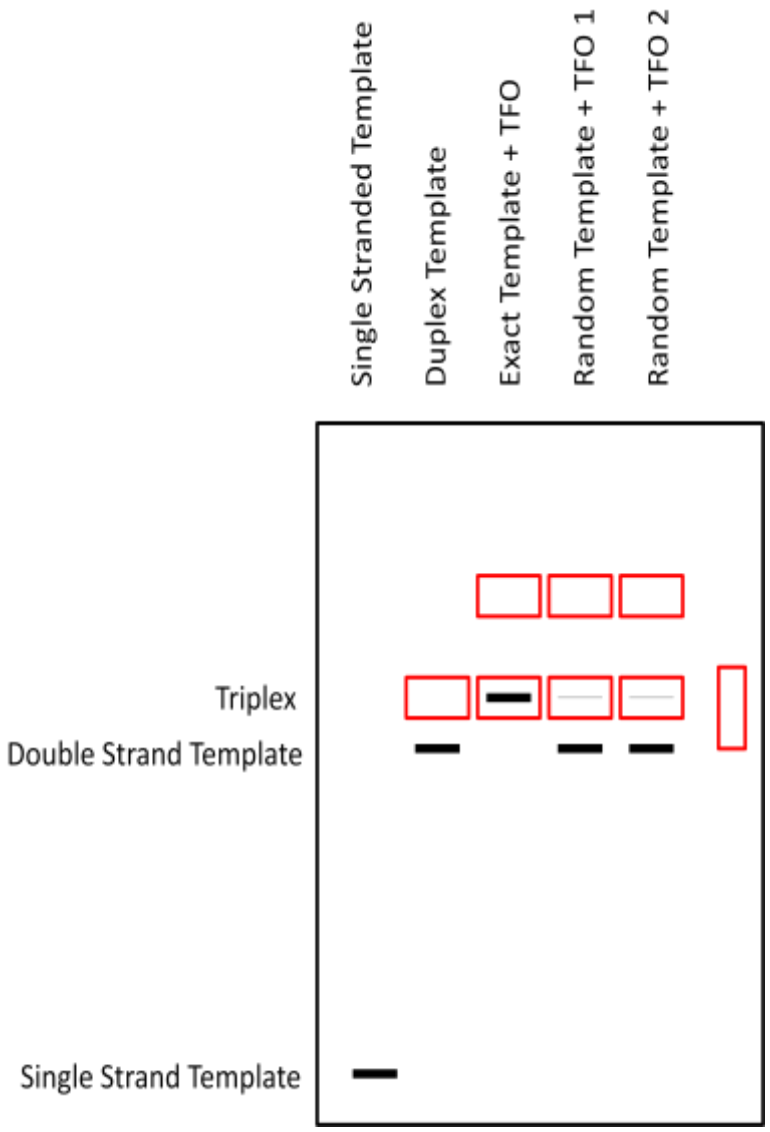


Figure 2.19: Diagram of positions of gel slices, black bands indicate radioactive DNA, red boxes show the position of gel slices taken.

The DNA from all of these slices was purified as in section 2.2.3.3 and then cloned. All slices produced clones, indicating that there was DNA present in every part of the gel. Sequencing of these clones showed no correlation or indication that the triplex and duplex populations had been separated successfully.

Chapter 3: Initial footprinting experiments with *tyr*T DNA show evidence of secondary binding

3.1 Introduction

This is a preliminary chapter covering the initial footprinting studies carried out on the *tyr*T template to examine the binding strength and specificity of the 9mer TFO which contains 6 consecutive bis-amino-U substitutions.

i)

5′ BBBBBBCBT 3′ B = Bis-amino-U

ii)

```
5′ -CTCGGGAAGCGGGGCGCTTCATATCAAATGACGCGCCGCTGTAAAGTGTTAGGAAGA
3′ -----CCTTCGCCCCGCGAAGTATAGTTTACTGCGCGGCGACATTTCACAATCCTTCT

GAAAAAAGAACTGGTTGCGTAATTTTCATCCGTAACGGATTAAAGGTAACCGGAA--- 3′
CTTTTTTCTTGACCAACGCATTAAAAGTAGGCATTGCCTAATTTCCATTGGCCTTAA- 5′
```

Figure 3.1:

i) - Sequence of the BAU containing TFO.

ii) - Sequence of the tyr*T(43-59) fragment used in footprinting studies. Primary TFO binding site is shown in red, radiolabel is shown in blue.*

Bis-amino-U (BAU) is a nucleotide analogue developed to help counteract the negative repulsion between the three strands in a DNA triplex. Each BAU nucleotide contains two positively charged amino groups which act to screen the negative phosphates in much the same way as divalent cations such as magnesium, and therefore increase the binding affinity of the TFO (see section 1.4.3). The 9mer TFO used in this investigation was first designed as part of a range of TFOs to investigate whether clustered or dispersed BAU residues are more effective at stabilising triplexes. It was found in this study that BAU-containing TFOs formed more stable triplexes if the nucleotides were clustered within the TFO. However, one of these initial footprinting experiments also revealed the presence of secondary binding sites for the clustered 9mer TFO (294). The footprinting gel below has been taken from this initial study.

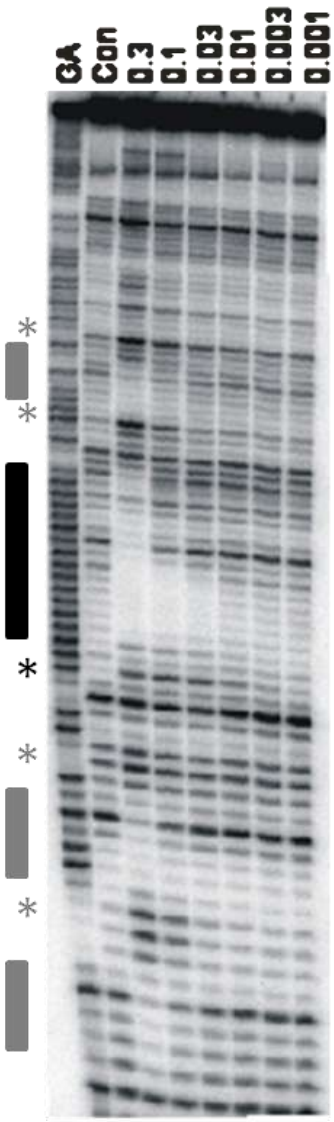


Figure 3.2: DNase I cleavage pattern of tyrT(43-59) with the BAU containing TFO at pH 5.0 (50mM sodium acetate) with 10 mM magnesium chloride. The complexes were left overnight at 20°C to equilibrate. TFO concentrations in μM are shown at the top of each lane. Lane labelled GA is a marker for purines (section 2.2.4.3) and control lanes show the DNaseI cleavage pattern in the absence of TFO. The black filled box shows the primary TFO binding site and the black asterisks shows the primary enhancement. Grey boxes and asterisks show the positions of secondary footprints and enhancements. Taken from (294).

The largest footprint in this gel is at the expected TFO binding site within the purine tract. However at the highest TFO concentration there are also three faint secondary footprints accompanied by several additional enhancements. All of these secondary footprints are smaller than the primary footprint and are only evident at the highest TFO concentration. The primary footprint persists down to 0.03 µM, indicating that these secondary interactions are significantly weaker than the primary one. The primary footprint at 0.3 µM is also larger than the binding site of the TFO; it extends in the 5' direction above the TFO binding site which is usual for triplex footprints. However, when compared to footprints generated by other TFOs of the same length this footprint appears to extend further than expected (294). There are also several secondary enhancements between the footprints in this gel, as indicated by asterisks, some of which also persist to 0.1 µM.

It appeared that the run of contiguous BAU residues within the TFO was adversely affecting the TFOs specificity, however it was unclear what the exact nature of these secondary binding sites was and why they were occurring. The sequences of the secondary binding sites on this template are examined further in the discussion and shown in Figure 3.5 on page 99. Using the *tyr*T template alone it was not possible to determine a consensus sequence for these secondary binding sites, so to fully investigate the binding preference of this TFO, REPSA experiments were carried out to select the preferred sequences from a pool of random templates (see Chapter 4).

Before performing REPSA experiments however it was necessary to investigate the pH dependence of this secondary binding further, as the selecting type IIS restriction enzyme required higher pH. This has been examined in this chapter.

3.2 Experimental design

The sequences of the TFO and the *tyr*T target duplex are shown in Figure 3.1 on page 91 above. *Tyr*T was radiolabelled and purified as described in section 2.2.3.1 then incubated overnight at 20ºC with the 9mer BAU TFO to allow association. No Magnesium ions were used in any of the buffers in this chapter. The template was then digested with DNaseI to give a ladder of bands as described in section 2.2.4. PH jump experiments were also performed using the same TFO and template, following the protocol from section 2.2.5.

3.3 Results

The gels shown in Figure 3.3 on the following page confirm the initial footprinting experiment from Figure 3.2 (page 92) and extend the results to show the footprinting patterns of the BAU-containing TFO on *tyr*T at three pHs. For these experiments the TFO was incubated with the *tyr*T template overnight in three different buffers; pH 5.0 (50mM sodium acetate), pH 6.0 (10mM PIPES 50mM NaCl) and pH 7.0 (10mM Tris-HCl 50mM NaCl) before digestion with DNase I. An additional footprinting experiment was carried out in the same pH 5.0 buffer but with only a one hour incubation time instead of an overnight incubation.

Figure 3.3 shows that there is a clear footprint at the intended target for the experiment at pH 5.0 and 6.0 and at pH 5.0 after 1 hour incubation. At pH 7.0 the footprint is less clear, although the enhancement in the surrounding sites suggest that the TFO has bound in the same position as at lower pH, only not as strongly. At pH 5.0 this primary footprint and the enhancement below it are visible at concentrations down to 0.1 µM. A secondary binding site with enhanced cleavage and a weak footprint can also be seen in this gel; the enhancement is visible down to 0.1 µM, while the two regions of attenuated cleavage above and below this only persist to about 0.3 µM.

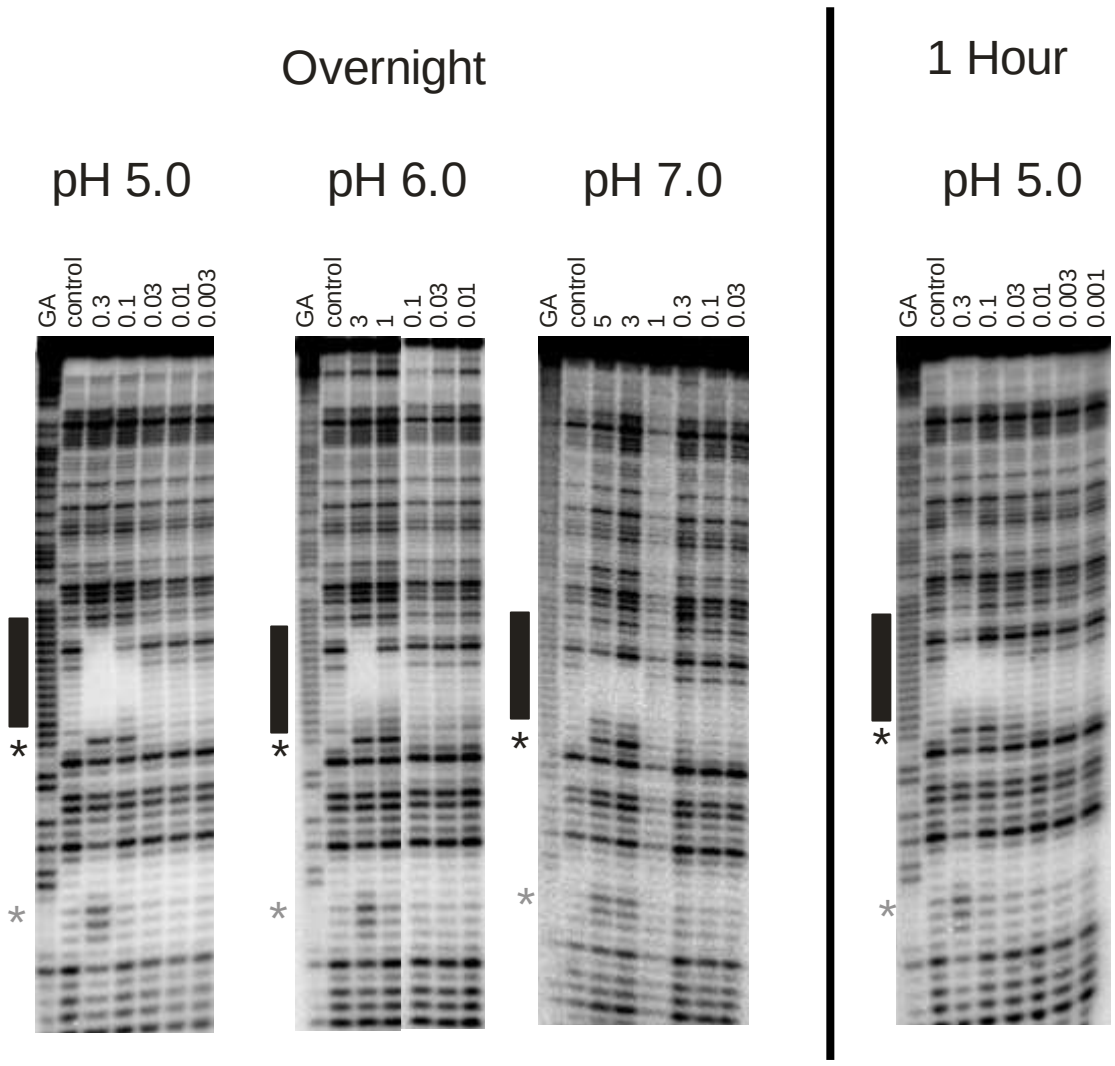


Figure 3.3: DNase I cleavage pattern of tyr*T(43-59) with the BAU containing TFO at three different pHs, shown above gels. Two different incubation times were used; either overnight or only one hour. TFO concentrations in μM are shown at the top of each lane. Lanes labelled GA are specific purine markers (section 2.2.4.3) and control lanes show the DNaseI cleavage pattern in the absence of TFO. Filled boxes show the position of TFO binding sites and black asterisks show the primary position of on enhanced DNase I cleavage. Grey asterisks show the position of secondary enhancements.*

At pH 6.0 higher TFO concentrations were used, as the triplex is less stable at this higher pH. The main footprint persists to 3 µM and the enhancement below can still be seen at 1 µM. As with the pH 5.0 footprint a secondary enhancement can be seen, also down to 1 µM. However there are no obvious regions of attenuated cleavage to indicate secondary binding sites. At pH 7.0 the main footprint is much less clear and attenuated cleavage is only seen at the highest TFO concentration (5 µM), though a region of enhanced cleavage is visible to 3 µM. Again there is a secondary enhancement down to 3 µM but no evidence of secondary footprints.

When the template and TFO were incubated at pH 5.0 for only one hour, the footprinting pattern is very similar to that seen with overnight incubation. The primary footprint is less clear than after long incubation, presumably because the shorter incubation time has not been sufficient to allow equilibrium binding.

When the overnight pH 5.0 gel is compared with the one in figure 3.2 (page 92) from the original study, fewer secondary binding sites can be seen and the primary binding site has not been enlarged at the highest oligonucleotide concentration. It is likely that this difference is due to the presence of magnesium chloride in the original experiment, which will have enhanced TFO binding. This was not included in the experiments performed in this chapter.

An interesting observation is that, although the cleavage pattern around the primary footprint is similar at all three pH's (requiring higher concentrations at higher pH's) the secondary binding is more obvious at pH 5.0. Only the pH 5.0 experiment has lighter bands indicating the position of secondary binding sites. This suggests that the TFOs propensity for binding these secondary sites is more sensitive to pH than binding to the primary site.

Because of this observation footprinting experiments were carried out to investigate the effect of sudden changes in pH on TFO binding (Figure 3.4 on the following page). These were also necessary as the REPSA experiments, described in the following chapters, require *Fok* I digestion of the complexes at a higher pH than the incubation at pH 5.0 necessary to form the triplexes.

As before the TFO and *tyr*T template were incubated together overnight at 20ºC in the pH 5.0 buffer (50 mM sodium acetate). Before digestion with DNase I however the pH was jumped from pH 5.0 to around pH 7.0 by adding 1 µl of 1 M Tris-HCl pH 8.0. Samples were then taken at various time points after the pH jump and digested to give the digestion pattern shown below. The detailed protocol for pH jumping is described in section 2.2.5.

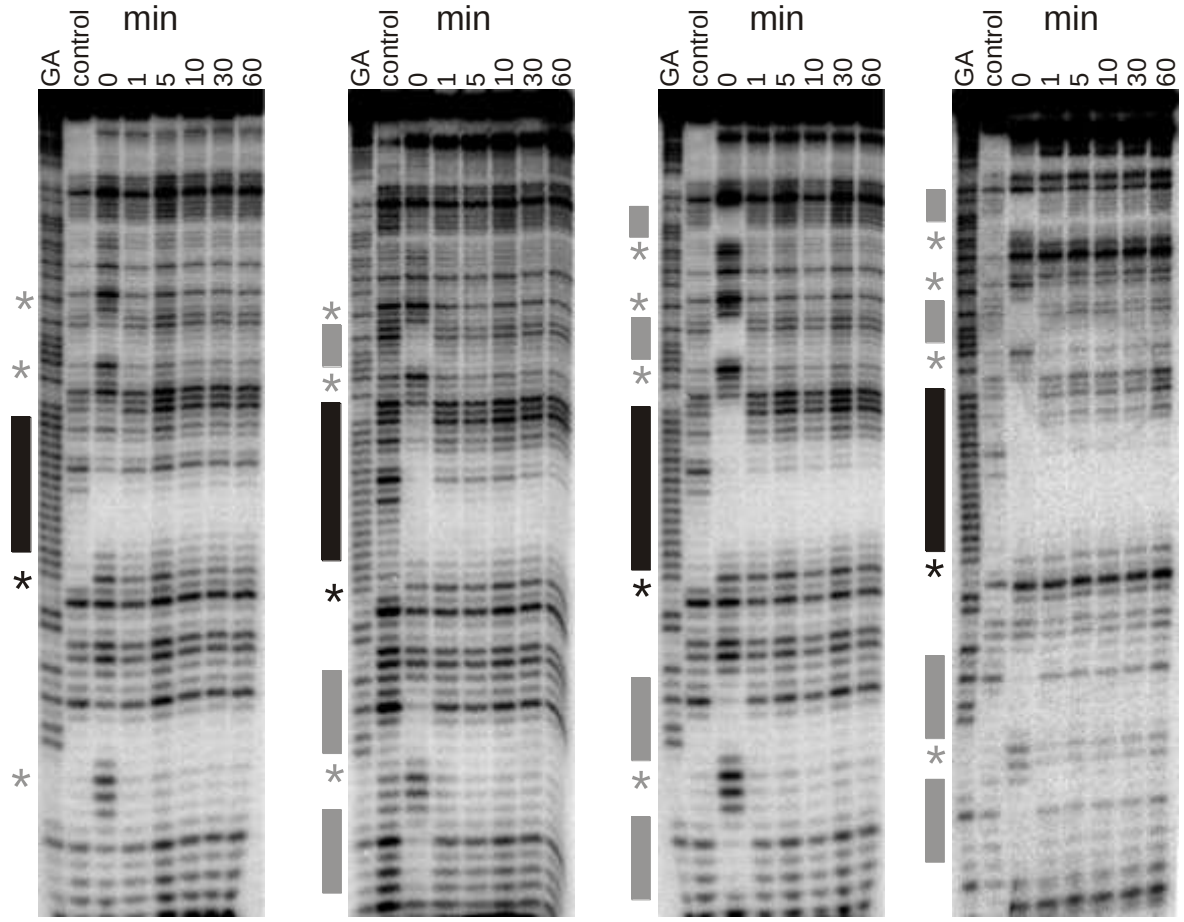


Figure 3.4: DNaseI cleavage patterns of tyr*T(43-59) with the BAU containing TFO. All footprint incubations were carried out overnight at pH 5.0 and four different TFO concentrations were used, as shown above the gels. A pH jump from pH 5.0 to around pH 7.0 was carried out after the overnight incubation, numbers above each lane indicate minutes after the jump before digestion with DNase I. Lanes labelled GA are markers specific for purines (section 2.2.4.3) and control lanes show the DNaseI cleavage pattern in the absence of TFO. Black filled boxes and black asterisks show the position of primary TFO binding sites and primary enhancements, grey boxes and asterisks show the positions of secondary footprints and enhancements.*

All of these footprinting experiments were performed after overnight incubation at pH 5.0 (before the pH jump) but use higher TFO concentrations than the experiments shown in Figure 3.3 (page 95). The primary and secondary binding sites are therefore clearer before the pH jump (lane 0) at these higher TFO concentrations. It can be seen that after the pH jump the primary footprint persists at the higher TFO concentrations but all of the secondary footprints and enhancements have disappeared before the first time point after the jump.

As with the pH 5.0 footprint shown in Figure 3.3, the cleavage pattern at 0.3 µM in Figure 3.4 (previous page) shows a primary footprint and enhancement before the pH jump, as well as three secondary enhancements with some evidence of secondary footprints. However, after the pH jump all footprints and secondary enhancements have disappeared by the 1 minute time point, consistent with the pH 7.0 experiment shown in Figure 3.3 which showed no footprint with 0.3 µM TFO. The primary enhancement below the primary target site remains for the duration of the time course.

All three experiments at the higher TFO concentrations (1, 3 and 5 µM TFO) show an enlarged primary footprint and several clear secondary footprints and enhancements before the pH jump. With 1 µM TFO three secondary footprints and enhancements can be seen before the pH jump, but all of these have disappeared by 1 minute after the jump. The primary footprint has also disappeared by this point, although the primary enhancement persists for the course of the experiment, as with the 0.3 µM footprint.

The footprints with 3 and 5 µM TFO retain a primary footprint and enhancement for the duration of the time course, although after the jump the footprint shrinks back to the expected size. This is again consistent with the pH 7.0 experiment shown in Figure 3.3, which showed a primary footprint and enhancement at 3 and 5 µM TFO. These two gels also show an additional secondary footprint and enhancement before the pH jump, making four of each in total, although all secondary binding is again eliminated by 1 minute after the jump.

The sequences of the primary and secondary footprints taken from Figure 3.4 are shown on the following page. The sequence of the TFO is also shown for comparison.

i)
```
5′ – TTCATATCAAATGACGCGCCGCTGTAAAGTGTTAGGAAGAGAAAAAAGAA
3′ – AAGTATAGTTTACTGCGCGGCGACATTTCACAATCCTTCTCTTTTTTCTT
                *** *           ****                  *
     CTGGTTGCGTAATTTTCATCCGTAACG – 3′
     GACCAACGAATTAAAAGTAGGCATTGC – 5′
        ****     ****
```

ii)
5′ AAATGA 3′

5′ AAAG 3′

5′ AAAAAAGAA 3′ (main site)

5′ GGATGAAAA 3′

iii) 5′ BBBBBBCBT 3′

Figure 3.5;

i) Sequence of the tyr*T template. For footprinting gels in this chapter the upper strand (hence referred to as the purine strand) was radiolabelled at the 3' end. The positions of footprints from the gels in this chapter are shown in red, and enhancements are indicated by asterisks under the sequence. The exact primary binding site of the TFO is indicated by the double underline and possible secondary binding sites of the TFO are indicated by the single underlines.*

ii) Summary of the possible TFO binding sites including the primary binding site, all shown from 5' to 3'

iii) Sequence of the TFO, B = Bis-amino-U. Note: TFO binds in a parallel *orientation to the template.*

By comparing the size and position of the primary footprint with the known TFO binding site (double underline), we can make some suggestions about the nature of the secondary sites that are represented by the secondary footprints. These possible secondary binding sites are shown by the singly underlined sequences. The two secondary sites to the left (in the 5' direction on the purine strand) of the primary footprint contain a purine-rich region to the right (5') of each footprint, consistent with the primary site, which also extends in the 5'-direction beyond the target. However, for the two footprints close together to the right (in the 3' direction on the purine strand) of the primary footprint there are no obvious purine-rich regions on the purine strand. There is a purine rich region on the lower (pyrimidine) strand between the two footprints, which extends into the far right footprint (underlined). This site may have bound a molecule of the TFO and resulted in both footprints.

3.4 Discussion

TFOs containing BAU substitutions generate extremely stable triplexes and individual substitutions have been shown to be highly selective for AT (187). The overall specificity of BAU is AT>GC>CG=TA. BAU also has enhanced discrimination against YR bases in the duplex compared with T. Although BAU is most stabilising against AT, BAU and T have similar stabilities at YR base pairs (187) It is likely that the specific contacts between the positive charges on BAU and phosphate groups in the triplex do not occur when this modification is placed opposition a YR base pair. BAU also shows high selectivity for AT over GC bases pairs in the duplex, although BAU.GC triplets are more stable than T.GC or BAU.YR triplets (187). TFOs with the same sequence as the one used in this thesis but containing 3 or 4 BAU residues showed no evidence of secondary binding on *tyr*T (186). It is therefore likely that BAU reduces the stringency of TFOs only when there are large numbers of consecutive substitutions.

The footprinting experiments shown in Figure 3.3 (page 95) reveal clear pH dependence for binding of this 9-mer BAU-containing TFO. Higher TFO concentrations were needed to generate a footprint as the pH was increased. The TFO appears to have high affinity for the template at pH 5.0 and 6.0 producing clear footprints, though there is still some evidence of binding at the highest concentration (5 μM) at pH 7.0. All of the gels show some evidence for secondary binding sites at the highest concentrations; this is seen by either attenuated DNase I cleavage or the presence of enhanced cleavage.

Although the primary footprint shows a similar cleavage pattern at all three pHs the secondary binding is more pronounced at pH 5.0. While regions of secondary enhanced cleavage are seen at all pHs, the additional footprints (or attenuated cleavage) are only evident at pH 5.0. This suggests that the TFO's ability to bind these secondary sites may be more sensitive to pH than binding to the primary site.

One reason for carrying out a footprint after only one hour incubation was to see if this shorter incubation would be appropriate for use in the REPSA experiments. From the experiment in Figure 3.3 it can be seen that there is very little difference in the DNase I cleavage pattern between the overnight and one hour incubations. However, the footprint is clearer with the longer incubation and the primary footprint persists to lower concentrations. This is consistent with the

known slow rate of triplex formation and indicates that TFO-binding has not reached equilibrium within this shorter time.

In the pH-jump footprinting experiments (Figure 3.4 page 97) the secondary binding of the TFO can been seen more clearly as higher TFO concentrations were used. Before the pH increase (lane 0) there is a clear footprint in the expected region and several secondary footprints and enhancements at the higher TFO concentrations. But it is clear that at low concentrations the TFO dissociates rapidly from the secondary binding sites when the pH is increased as all the secondary footprints and enhancements have disappeared by the first (1 minute) time point for all concentrations of TFO. However, with the two highest TFO concentrations the primary footprint persists after the pH jump for the length of the experiment. The three highest TFO concentrations produce a primary footprint which is larger than the expected nine bases before the pH jump, but this decreases back to the expected size immediately after the jump.

This enlarged primary footprint can also be seen in the original experiment (Figure 3.2 page 92). This enlargement of the primary footprint may be caused by 'slippage' of the TFO allowing it to cover more of the purine rich region of the template. Alternately TFO molecules bound at the primary and secondary sites may interact in such a way as to extend the apparent binding site, by restricting the access of DNase I to the DNA in this region. In all gels the primary footprint shrinks back to its expected size as soon as the secondary footprints and enhancements disappear, lending weight to this theory. Also enlarged footprints are not seen with similar TFOs which do not bind to secondary sites, even at very high concentrations (294).

In general the footprints in Figures 3.3 and 3.4 show the same binding pattern and pH dependence. There is one small difference; in Figure 3.4 the primary enhancement persists at all TFO concentrations whereas in Figure 3.3 at pH 7.0 no enhancement is visible with 0.3 µM TFO. This is most likely because in the pH jump experiments the TFO was already bound to the template before the pH was increased, and therefore may have been able to affect DNase I cleavage enough to cause the enhancement even after the increase in pH.

This set of experiments supports the suggestion that the TFO molecules bound at secondary sites dissociate more rapidly at high pH, in comparison with TFOs bound to primary sites.

Although these footprinting experiments give some useful information about the interaction of the BAU-rich TFO with secondary sites, the results are limited to

this particular duplex DNA fragment and the sequences that it contains. Therefore the next stage in the investigation was to use REPSA to find the preferred binding sites of the TFO. The proposed position and sequence of the secondary footprints and enhancements described in this chapter are summarised in Figure 3.5 (page 99) and allow us to make some suggestions about the essential features of the secondary sites. They appear to be purine tracts of various lengths that share little similarity with the exact recognition site of the TFO. To be able to draw meaningful conclusions about the specificity of this TFO it would be more useful to examine binding sites that differ only slightly from the exact recognition sequence, and this has been done in Chapter 4. As noted earlier the triplexes formed on the secondary binding sites of the *tyr*T template are more sensitive to pH than the primary triplex. By keeping the pH relatively high during REPSA we may therefore be able to screen for only the most tightly bound secondary sites, which will presumably be the sites which are most similar to the exact recognition sequence.

Chapter 4: Footprinting on REPSA targets

4.1 Introduction

This chapter farther examines the sequences that were isolated from the REPSA experiments. 14 different templates were chosen from these REPSA-selected templates and their interaction with the BAU-containing TFO was analysed by DNase I footprinting. These results were compared with experiments using related TFOs with the same sequence but containing propargylamino-dU and T instead of BAU. Propargylamino-dU is similar to BAU but contains only one of the positive groups found on BAU (see section 1.4 for more details). It was thought that reducing the positive charge on the TFO would make it less likely to bind secondary sites. A TFO with the same sequence but containing only T, with no modified bases, was also tested as a control.

As well as footprinting these TFOs on the sequences selected by REPSA, two sets of control footprinting experiments were also carried out to compare the TFOs affinity for the primary and secondary binding sites under similar conditions. These two templates both contain the exact recognition site for the TFOs but in a different context; one is the *tyr*T template used in Chapter 3, and the other is the REPSA template with the exact sequence in place of the random region (see Figure 4.2 on the following page).

Early summary of Results:

The BAU substituted TFO was able to bind to all but one of the sequences isolated from the REPSA experiments, even though they do not contain the exact target site, down to concentrations between 1-0.1 µM. The propargylamino-dU (P) and T containing TFOs failed to produce a footprint on any template even at 30 µM, although some enhancements are seen. In order to promote binding of these TFOs magnesium chloride and the naphthylquinoline triplex binding ligand were added to the incubation buffer. Under these conditions the propargylamino-dU TFO was able to bind almost all of the templates. However, the control TFO containing only T and C did not bind to the REPSA templates, even in the presence of the triplex binding ligand.

4.2 Experimental design

The original and revised REPSA methods, the optimisation experiments and the results of the first set of REPSA experiments are described in section 2.3. Template sequences selected by REPSA are shown in the relevant sections of this chapter. The sequences of the TFOs used for footprinting are shown below:

```
B    5′ – BBBBBBCBT – 3′
P    5′ – PPPPPPCPT – 3′
T    5′ – TTTTTTCTT – 3′
```

Figure 4.1: Sequence of TFOs used in footprinting, abbreviations B, P and T used throughout this chapter.
B = Bis-amino-U
P = Propargylamino-dU
T = Thymine

The two control templates used were the *tyr*T fragment and the exact target for the oligo (contained within the REPSA template), both of which were cloned into pUC19 and labelled in the same way as the sequence derived from REPSA selection.

i)
```
5′ –CTCGGGAAGCGGGGCGCTTCATATCAAATGACGCGCCGCTGTAAAGTGTTAGGAAGA
3′ -----CCTTCGCCCCGCGAAGTATAGTTTACTGCGCGGCGACATTTCACAATCCTTCT

GAAAAAAGAACTGGTTGCGTAATTTTCATCCGTAACGGATTAAAGGTAACCGGAA--- 3′
CTTTTTTCTTGACCAACGCATTAAAAGTAGGCATTGCCTAATTTCCATTGGCCTTAA- 5′
```

ii)
```
5′-AATTCGAGCTCGGTACCCGGGGATCCTGACTGGATGAACGCAAAAAAGAAGCGTCTTC
3′-TTAAGCTCGAGCCATGGGCCCCTAGGACTGACCTACTTGCGTTTTTTCTTCGCAGAAG

CTGACAAGCTTCAG-3′
GACTGTTCGAAGTC-5′
```

Figure 4.2: Sequence of the two control templates used in this chapter, TFO binding site shown in red and radiolabel in blue.
i) Sequence of the tyr*T template*
ii) Sequence of the synthesised exact REPSA template cloned into pUC19.

The templates were radiolabelled as described in section 2.2.3 and the footprinting experiments were carried out as described in section 2.2.4. For separating the products of DNaseI digestion a 9% denaturing polyacrylamide gel was used for the *tyr*T sequence but for all REPSA templates and the exact template a 12.5% gel was used as these fragments are shorter. For all footprinting experiments the template and TFO were incubated in pH 5.0 buffer (50 mM sodium acetate) overnight at 20°C. For some footprints 5 mM magnesium chloride or 5 or 10 µM naphthylquinoline triplex binding ligand were added to the incubation buffer.

4.3 Results

4.3.1 REPSA selection

Since the first set of REPSA experiments produced templates which showed no obvious selection (Section 2.3) and failed to bind the BAU containing TFO, optimisation experiments were then carried out (section 2.3.3) and REPSA was then repeated with the new protocol (section 2.3.1). The templates sequenced at the end of this set of REPSA experiments are shown in Tables 4.1 i and ii on the following page. The original sequencing gels and files are shown in Appendix 1 and 2.

Out of 114 colonies picked for sequencing 37 contained no insert even though the colonies were white, possibly due to a mutation elsewhere in the plasmid. Eight of the colonies were unidentifiable and 37 contained an insert which couldn't be fully sequenced (possibly because they adopt unusual structures). The other 32 colonies picked contained plasmids which were fully sequenced and these are shown in the tables on the following page. 14 of these sequences were chosen for footprinting (Table 4.1 i).

An interesting pattern emerged during sequencing; the selected templates were all in the same orientation, with the purine-rich strand visible on the sequencing gel (and therefore on the footprinting fragments as well). Since the TFO binds to duplex DNA we would have expected an equal distribution of pyrimidine- and purine-rich strands. This unexpected phenomenon will be considered further in the Discussion.

Due to constraints of time and materials it was not possible to study every template sequence isolated by REPSA. As the aim of this work was to compare secondary binding sites of the TFOs with each other and the exact binding site it was decided to examine templates which were most similar to the exact target sequence. These sequences are shown in Table 4.1 ii; they all contain a run of at least four consecutive A residues. The only exception to this is sequence 6, which was originally thought to have a run of six As but was re-sequenced when the BAU TFO failed to produce a footprint and was found to contain a pyrimidine interruption.

(i)

Number	Sequence
1	CAATC**AAAAAAA**CAA
2	**AAAAAAAA**GTTGCTCT
3	CCCCGAAC**AAAAAAA**
4	C**AAAAAAA**GAGAGAC
5	TATTGCT**AAAAAA**GG
6	**AAA**T**AAA**CGGCCATTG
7	CCAGAGCAC**AAAAAA**CAA
8	GGT**AAAAAA**CCGGCA
9	CAACTC**AAAAAA**TCA
10	CCCT**AAAAA**TAACAA
11	GCAGT**AAAAAAA**CTA
12	AAG**AAAA**GCAATGGGATCT
13	CACACGC**AAAAAA**G
14	AAG**AAAAAA**CAGCTCCTAC

(ii)

Number	Sequence
15	GCTACAAGGTCGCAC
16	GCCCAAAGAGACACA
17	ATTTATCGGACAGAA
18	GACAACCGAAAGCAC
19	TATCGCGAGCACAAC
20	GCGGGTGGGATAGCATA
21	CGAGCTAAAGATCCA
22	GCCTAACCAGATAGA
23	CCTAAAACCCATA
24	TATCGTAAGGCAAGT
25	GCACAACCATGCAAGA
26	CTTTAGCCCAGGAAT
27	ACCACCTAAAATCAT
28	ACCTACCATGACTATA
29	GTGCCCTCGGAGAAA
30	AGATATCAGATAC
31	GAATCCATACGACTT
32	ACCAACCGATATAGC

Table 4.1;
i) Table of template sequences selected in the second set of REPSA experiments that have been used in footprinting experiments.
ii) Sequences which were not used in footprinting

The sequences isolated by REPSA which were not investigated further are shown in Table 4.1 ii. We considered that these were less likely to bind the TFO as they do not contain a run of As or purines. These fragments may contain sequences which were very weakly bound by the TFO so were selected by

REPSA, or they may be random sequences which were not screened out by REPSA as it is purposefully not 100% efficient; sequences which have not bound the TFO may still elude cleavage by Fok I for other reasons, such as formation of a secondary structure which prevents access by the enzyme (299). Multiple rounds of REPSA are carried out for this reason, in order to give the enzyme multiple opportunities to cleave unbound templates. REPSA cannot be continued indefinitely however; if too many rounds are carried out it can lead to mutations in the template sequences. Indeed it can be seen that some of the sequences in Table 4.1 are not the same length as the random template with which REPSA was started. This is an acknowledged limitation of the REPSA technique (299).

4.3.2 Footprinting controls

As an initial control experiment the three TFOs shown in Figure 4.1 (page 104) were footprinted on the two control templates shown in Figure 4.2 (page 104); the *tyr*T fragment and the exact REPSA template. Several different conditions were used for the experiments with the propargylamino-dU and thymidine-containing oligos in order to potentiate any possible interactions with the target binding site. These included the addition of magnesium chloride and/or the naphthylquinoline triplex binding ligand. It is important to note that the BAU footprint appears at a 100 times lower concentration than the footprints with all the other oligonucleotides in most of these Figures.

The following abbreviations are used to indicate the TFOs used and whether magnesium chloride or naphthylquinoline were added for all of the footprinting Figures presented in this chapter;

Abbreviation	**Meaning**	**TFO Sequence**
BAU	Bis-amino-U-containing TFO	5’ – BBBBBBCBT – 3’
P	Propargylamino-dU-containing TFO	5’ – PPPPPPCPT – 5’
P+M	P + 5 mM magnesium chloride	5’ – PPPPPPCPT – 5’
P+5L	P + 5 µM naphthylquinoline	5’ – PPPPPPCPT – 5’
P+10L	P + 10 µM naphthylquinoline	5’ – PPPPPPCPT – 5’
P+M+5L	P + 5 mM $MgCl_2$ + 5 µM naphthylquinoline	5’ – PPPPPPCPT – 5’
P+M+10L	P + 5 mM $MgCl_2$ + 10 µM naphthylquinoline	5’ – PPPPPPCPT – 5’
T	Thymidine-containing TFO	5’ – TTTTTTCTT – 3’
T+10L	T + 10 µM naphthylquinoline	5’ – TTTTTTCTT – 3’

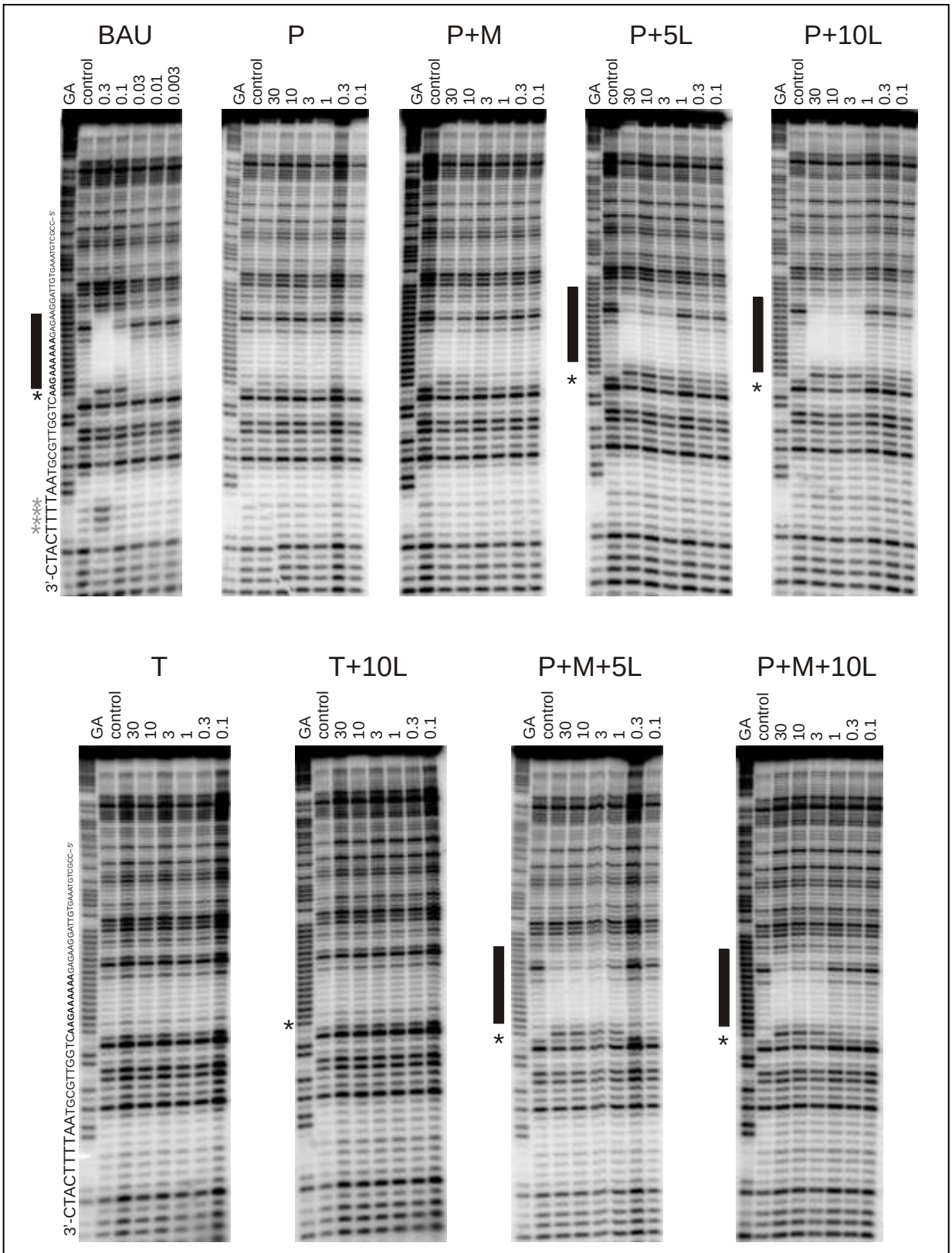


Figure 4.3: DNaseI cleavage patterns of Tyr*T(43-59) with the three TFOs shown in Figure 4.1. The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO. Black rectangles show the position of the footprints, black asterisks indicate enhancements and grey asterisks indicate enhancements remote from the primary binding site*

DNase I cleavage patterns of the *tyr*T fragment after incubation with various TFOs are shown in Figure 4.3 on the previous page. It can be seen that the BAU-containing TFO produces a footprint on this fragment down to concentrations of about 0.1 µM, as shown in Chapter 3. This footprint is accompanied by enhanced cleavage at the 3'- (lower) end of the target site at the triplex-duplex junction, also down to 0.1 µM. At the highest TFO concentration (0.3 µM) a secondary enhancement can also be seen further down the gel and there are some slightly lighter bands, again as seen in the footprints in Chapter 3. The position of the footprint and enhancements are shown below;

```
                    5′ – BBBBBBCBT – 3′
5′ – AAAGTGTTAGGAAGAGAAAAAAGAACTGGTTGCGTAATTTTCATCC – 3′
3′ – TTTCACAATCCTTCTCTTTTTTCTTGACCAACGAATTAAAAGTAGG – 5′
                                *             ****
```

Figure 4.4: Tyr*T sequence showing the binding of the BAU substituted TFO. The upper / purine-rich strand was radiolabelled at the 3' end in all the footprinting experiments described in this thesis. The position of the footprint from the BAU gel in Figure 4.3 is shown in red and the enhancements are indicated by asterisks below the sequence. The expected binding site of the TFO is underlined and the TFO is shown above the sequence in its anticipated binding position.*

The footprint extends for several bases beyond the target site in the 5'-direction on the purine strand, as noted in many other triplex footprinting studies(300). This extension is due to the size of DNase I and is also related to the fact that DNase I cuts from the minor, while the TFO is located in the major groove. The footprint is accompanied by enhanced cleavage at the 3'-end of the purine strand at the triplex-duplex junction, as often noted in other studies(300).

The cleavage patterns of *tyr*T with the propargylamino-dU TFO alone or with addition of 5 mM magnesium chloride show a slight enhancement at the same position as the enhancement with the BAU TFO. The thymidine TFO does not produce a footprint but shows enhanced cleavage at the triplex-duplex junction in the presence of the naphthylquinoline triplex binding ligand. There is also no evidence of secondary footprints or enhancements in any of these experiments with the P or T TFOs.

All templates incubated with the propargylamino-dU TFO and naphthylquinoline show a footprint down to 3 µM TFO, at both ligand concentrations and with or without magnesium chloride. They also show an

enhancement at the same position as the previous gels, which persists down to around 1 µM TFO. However there is also some (weaker) enhanced cleavage one band above this (in the 5'-diretion). This may indicate that the terminal bases of the TFO are fraying from the target so that, for some of the time, the triplex-duplex junction is located one base pair higher.

From these footprints on the *tyr*T template it can be seen that the BAU substituted TFO binds at a 30-fold lower concentration than the propargylamino-dU containing TFO in the presence of the triplex binding naphthylquinoline. Without the ligand only a slight enhancement can be seen with this TFO, even at 300 times the concentration that produces a footprint with the BAU TFO. The control TFO containing only thymidine also shows only a very slight enhancement at this concentration, and then only with the addition of naphthylquinoline.

Additional enhancements are evident when the BAU-containing TFO is footprinted on the *tyr*T fragment, which indicate secondary binding sites. Although much higher concentrations of the propargylamino-dU containing TFO are required to affect the DNase I cleavage pattern, there is no evidence of secondary binding at these elevated concentrations. The secondary enhancement also appears at a three times higher concentration than the primary footprint on the BAU gel, but there is no secondary binding even at 10 times the concentration of the primary footprint on the P+M+10L gel. It is clear that the BAU-substituted TFO binds more strongly to these secondary sites and that this secondary (weaker) interaction is a feature of BAU-substitution.

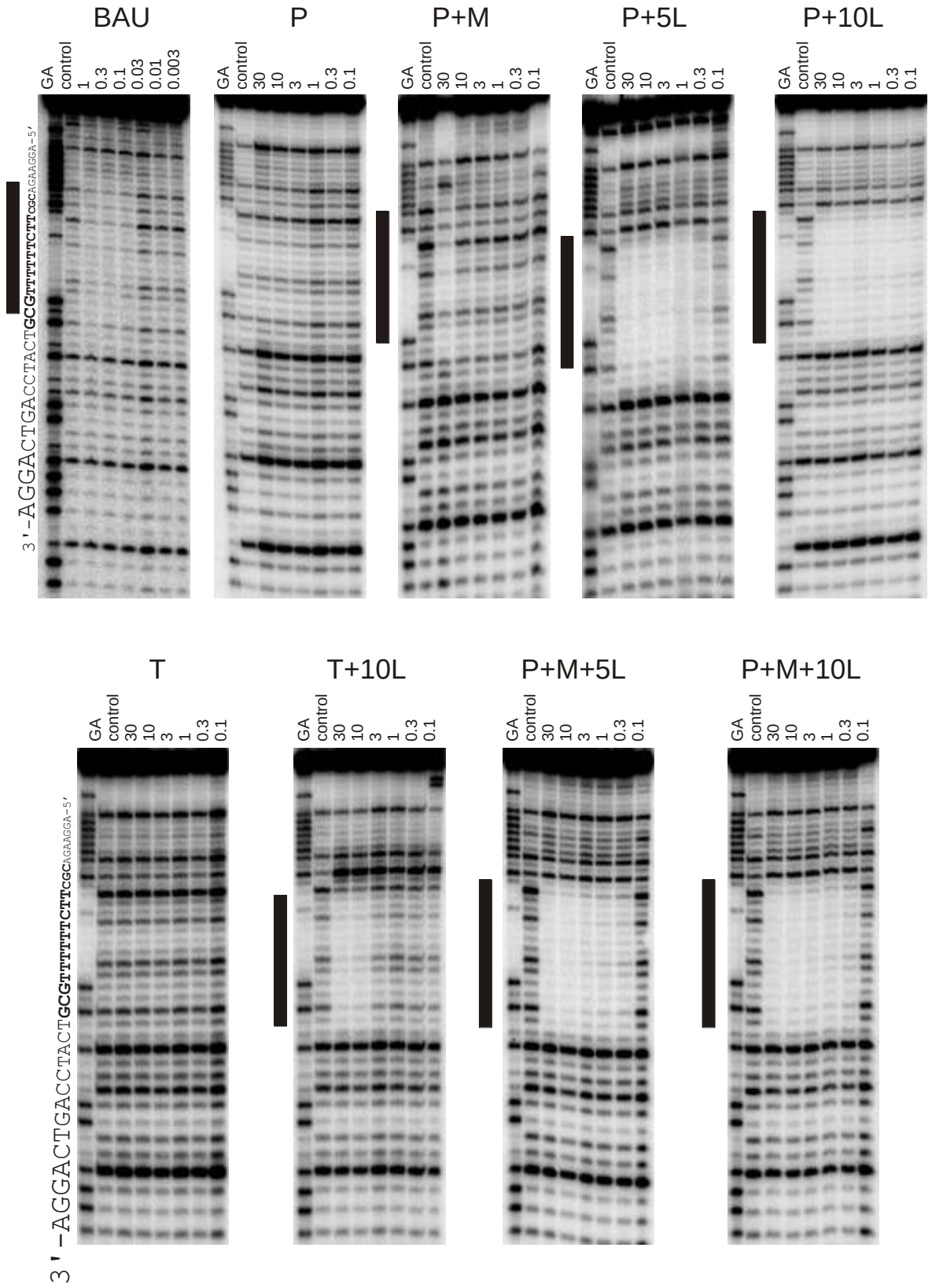


Figure 4.5: DNaseI cleavage pattern of the exact target (Figure 4.2 ii page 104), cloned into pUC19, footprinted with the three TFOs shown in Figure 4.1 (page 104). The experiments were performed in 50 mM sodium acetate. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates the DNaseI cleavage pattern in the absence of TFO.

The results of similar footprinting experiments with a fragment containing the exact target site within the context of the rest of the REPSA template are shown in Figure 4.5 on the previous page. The complementary strands containing this sequence were synthesised, annealed and then cloned into the pUC19 polylinker (sequence shown in Figure 4.2ii page 104). One important thing to note is that in this fragment the sequence has been cloned so that the pyrimidine-rich strand is labelled and visualised in the gels. In contrast, all the sequences isolated from the REPSA experiments were orientated so that the purine-rich strand was visualised. This unexpected bias in the REPSA sequence, akin to directional cloning, is considered in the Discussion.

An important consequence of this difference is that no enhancements are visible on these gels in Figure 4.5. This is often the case when looking at the pyrimidine-rich strand in footprinting experiments(284). Unfortunately, this makes the results on this template slightly more difficult to compare with footprints on the REPSA templates or the *tyr*T fragment. However this synthetic template contains the exact binding site for the TFOs, in the context of the same flanking sequences as used in the REPSA experiments.

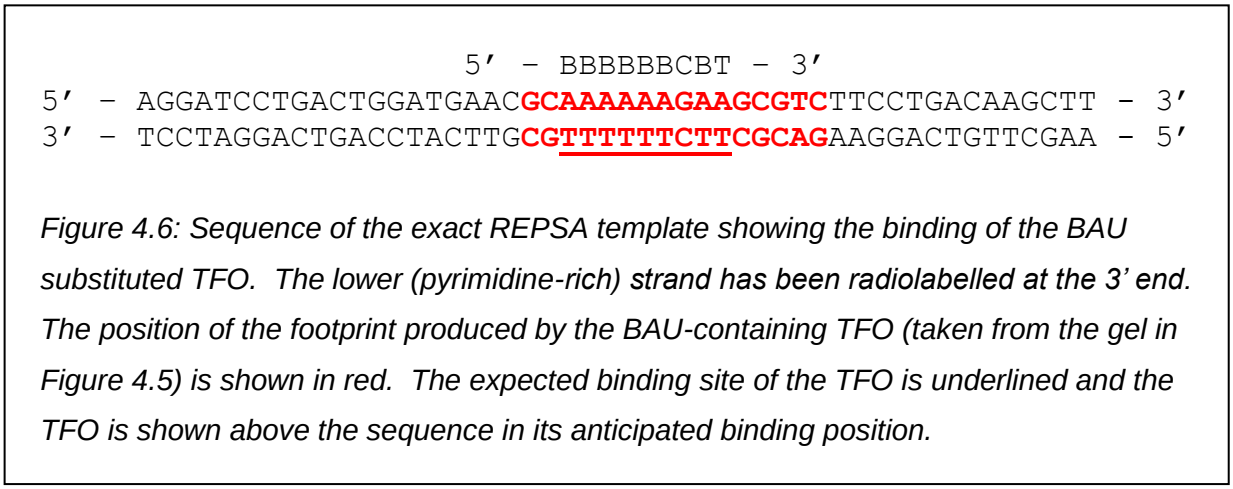


Figure 4.6: Sequence of the exact REPSA template showing the binding of the BAU substituted TFO. The lower (pyrimidine-rich) strand has been radiolabelled at the 3' end. The position of the footprint produced by the BAU-containing TFO (taken from the gel in Figure 4.5) is shown in red. The expected binding site of the TFO is underlined and the TFO is shown above the sequence in its anticipated binding position.

The footprint produced by the BAU substituted TFO extends beyond the anticipated binding site by several bases at both ends. This footprint is very unclear however; the bands within the footprint are fainter but still visible and the edges of the footprint are therefore difficult to determine. Unfortunately this is not a particularly clear gel so the faint edges may simply be an artefact. This footprint persists down to around 0.1 µM TFO. Unlike with the *tyr*T fragment there is no evidence of secondary binding on this template.

There is no footprint visible with the propargylamino-dU TFO, but when magnesium chloride is added to the incubation a footprint can be seen at the

highest concentration (30 μM). This footprint is four bases shorter at the 5' (upper) than the footprint with BAU where it appears to terminate at this upper triplex-duplex junction; the lower (3') of this footprint extends a few bases beyond (below) the target site, terminating at approximately the same position as the footprint generated by the BAU-containing TFO.

For all of the gels with P and ligand (with or without magnesium chloride) there are strong footprints which persist to 0.3 / 0.1 μM. These footprints are four bases shorter than the BAU footprint at the 5' (upper) end, as seen on the P+M gel. However they are also two bases longer at the 5' end than the footprint on the BAU or P+M gels. This is atypical of triplexes with pyrimidine targets and naphthylquinoline; footprints are usually shorter at the 5' end in the presence of the ligand(167).

As with *tyr*T there is no footprint produced by the thymidine TFO alone, but when ligand is added there is a clear footprint down to 10 μM. Again the position of this footprint is slightly different to any of the others. It extends to the same point as the P+L footprints at the 3' end but is two bases shorter at the 5' end.

Figure 4.7: Sequence of the exact REPSA template showing the position of footprints and enhancements generated by the P- and T- TFOs under various conditions. The lower/pyrimidine strand has been radiolabelled at the 3' end. The bases in red and blue indicate the extent of the footprint on the P+M gel, the bases in all three colours show the size of the footprint in all experiments involving the P-TFO and the ligand, and red and green bases indicate the size of the footprint seen on the T+10L gel. The bases in red are therefore the only ones found in all footprints, including the one generated by the BAU-substituted TFO. The TFOs are shown above the sequence in the anticipated binding position.

4.3.3 REPSA footprints

Fragments containing the 14 chosen sequences (Table 4.1 i) were footprinted with the three TFOs shown in Figure 4.1 (page 104) to examine the effect of different substitutions on TFO binding to these punitive secondary sites. In each case the radiolabelled DNA fragment and the TFO were incubated overnight at 20°C in pH 5.0 buffer (50 mM sodium acetate). The sequences of the relevant parts of the fragments are shown along the side of the first gel in each Figure. In each case the purine-containing strand is visualised on the footprinting gel.

Magnesium chloride or naphthylquinoline were added to the incubation buffer in some instances in order to enhance any weak binding and this is indicated above the gel. In each case the positions of footprints are indicated by a black rectangle. All incubation conditions (9 gels) are shown for the first template (Figure 4.8 on the following page). For ease of comparison only four gels are shown for the remainder of the fragments (B, P, P+10L and P+M+10L).

The templates are presented in the order of roughly how closely they resemble the exact target sequence starting with the one with the fewest mismatches.

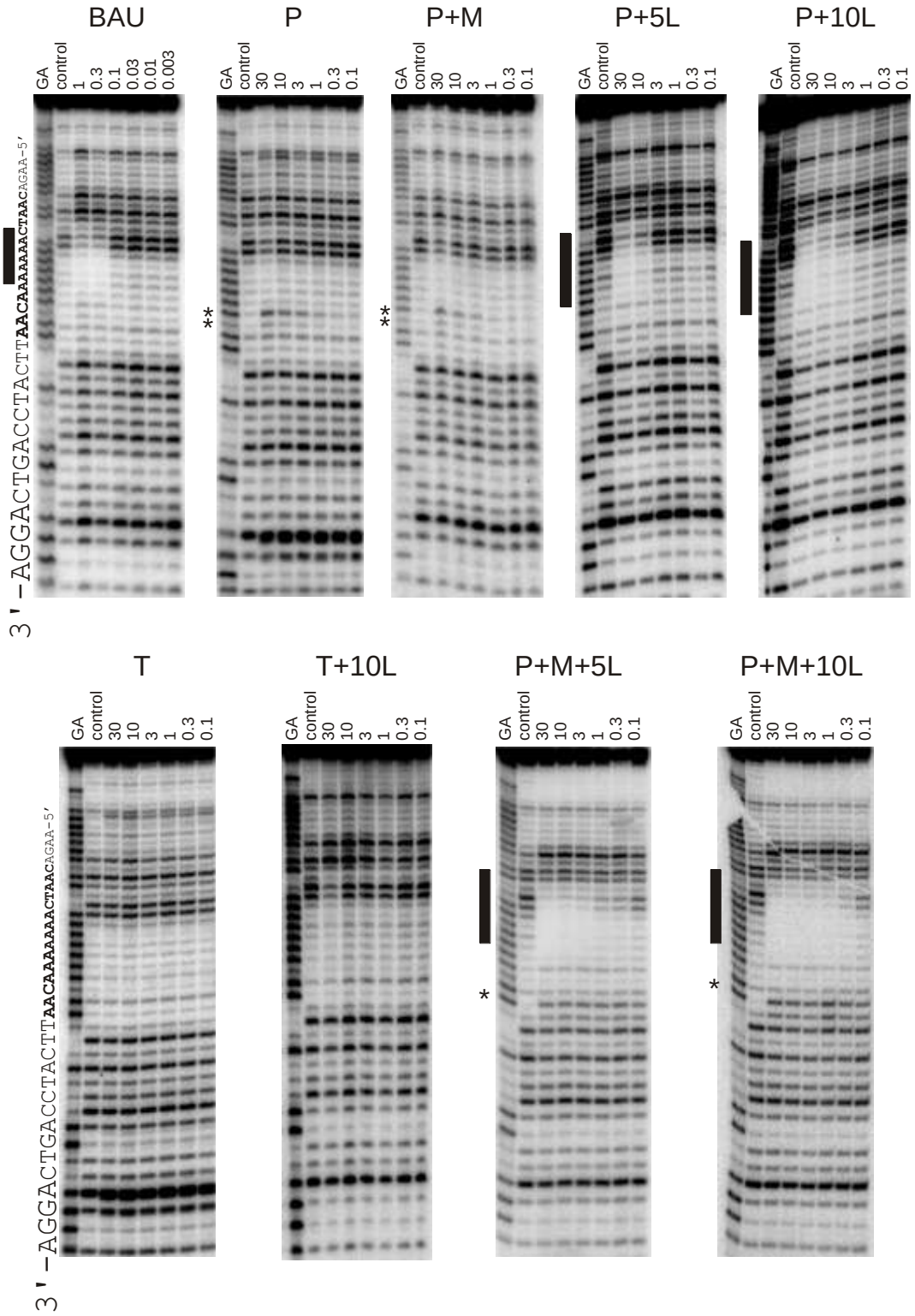


Figure 4.8: DNaseI cleavage patterns of template 1 from the REPSA experiment with the 3 TFOs shown in Figure 4.1. Template and TFOs were incubated overnight at 20°C in a pH 5.0 buffer containing 50 µM sodium acetate. TFO concentrations (µM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side; footprints are indicated by black boxes and enhancements by asterisks.

The DNase I cleavage patterns of template 1 after incubation with various TFOs are shown in Figure 4.8 on the previous page. The BAU substituted TFO produces a footprint down to a concentration of around 0.3 µM, even though this template does not contain its correct binding site, confirming that REPSA had successfully selected for TFO-binding. No enhancement is visible on this gel. The position of the footprint (which is shorter than the TFO) is shown below;

```
            5′ – BBBBBBCBT – 3′
            5′ – PPPPPPCPT – 3′
5′ – AAGACAATCAAAAAAACAATTCA – 3′
3′ – TTCTGTTAGTTTTTTTGTTAAGT – 5′
                          *
```

Figure 4.9: Template 1 sequence showing the locations of the footprints and enhancements produced by the TFOs. Note that the upper (purine-rich) strand is visualised in Figure 4.8 and has been radiolabelled at the 3' end. The bases in red show the BAO-TFO generated footprint and the red and blue bases show the extended footprint in the P+L gels. The enhancement was also only seen on these gels, not with the BAU-TFO. A possible location for the TFOs on this template is shown above the sequence and the complementary bases in the sequence are underlined.

Since triplex footprints are often shifted several bases in the 5' direction compared with the actual binding site it seems possible that the six consecutive BAU substitutions have bound to six of the run of seven adenines in the template, perhaps excluding the one at the 5' end.

In contrast, the propargylamino-dU-containing TFO does not produce a footprint in this region, and addition of magnesium chloride had no visible effect on the binding. This indicates that this isolated secondary binding site is a feature of the multiple BAU modifications and is not produced by all oligos. Two bands show a slight enhancement on both of these gels with P at the 3' end of the A tract, which persists down to around 3 µM TFO, maybe indicating that there is some weak interaction.

This double enhancement will be considered further in the Discussion and may indicate that the oligo has more than one binding location. It is also interesting to note that there was no enhancement produced by the BAU-substituted TFO. On addition of 5 or 10 µM triplex-binding-ligand a clear footprint is evident with the propargylamino-dU substituted TFO (this is evident both with

and without magnesium chloride), in a similar position to that produced by the BAU-containing TFO, but one base shorter at the 5' end.

With 5 µM ligand the TFO binds down to a concentration of around 10 µM, while with 10 µM ligand the footprint persists to around 3µM. The gels with ligand and magnesium chloride both show footprints at lower concentrations, down to around 0.3 µM of the P-containing TFO. There is also an enhancement on the P+M+5L and P+M+10L gels at all concentrations, but not on the P+5L or P+10L gels. The position of this enhancement is slightly different to the P or P+M gels; it is shifted three bases to the right. This may indicate that under these conditions the TFO is binding at multiple sites; perhaps slipping along the A tract.

No clear footprints are produced by the T-containing TFO even in the presence of the naphthylquinoline ligand. However on close examination of the T+10L gel it can be seen that, at the highest TFO concentration (30µM), the intensity of the bands is attenuated towards the top of the region where the BAU and P TFOs generated footprints; this may indicate a weak interaction with this TFO. None of the other templates generated footprints or enhancements with this TFO, and the results for T and T+10L are therefore not shown for any of the other templates presented below.

As seen with the *tyr*T DNA fragment gels the footprint generated by the propargylamino-dU substituted TFO is shifted in the 5' direction compared with the BAU footprint. In this case however the footprint with the P-TFO is one base longer in the 5' direction, whereas on *tyr*T it was one base shorter at the 3' end.

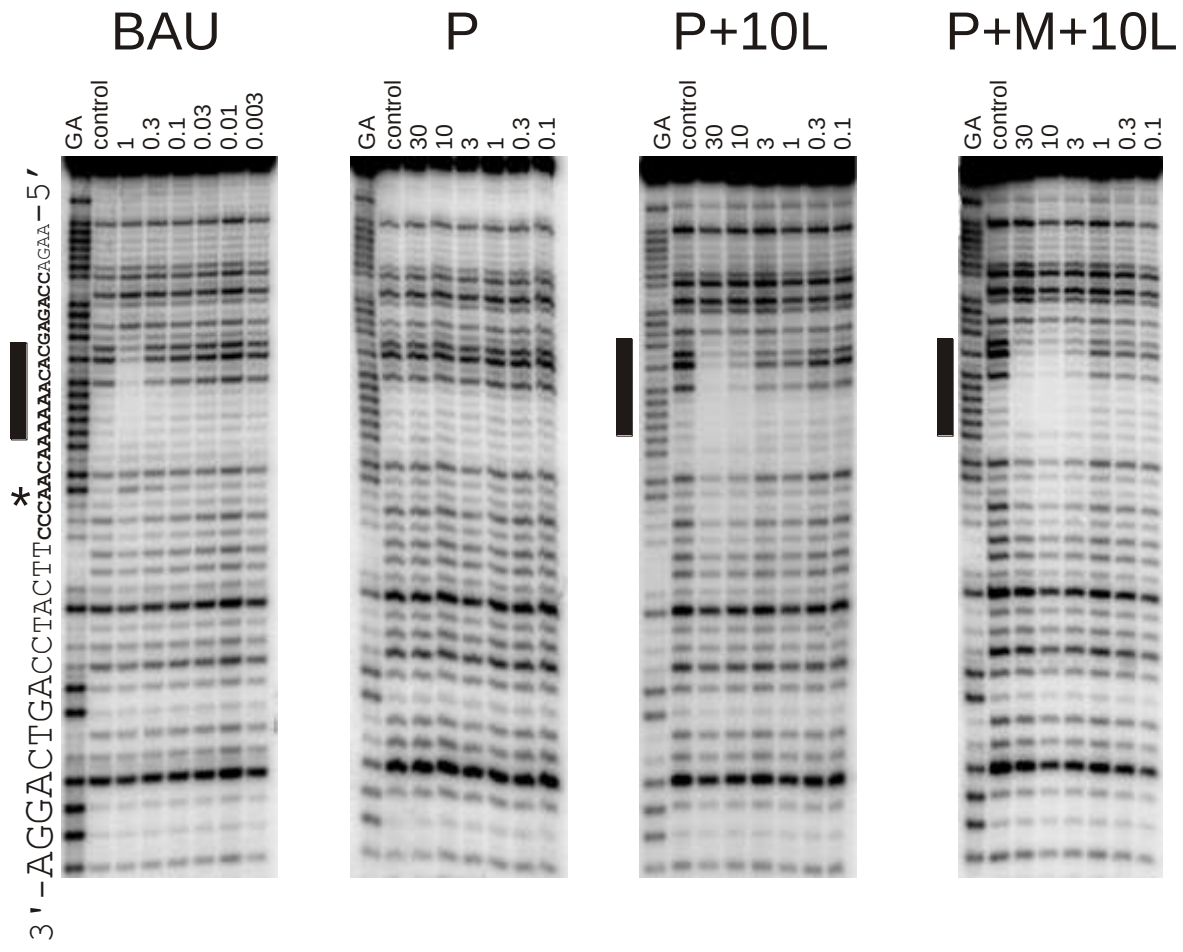


Figure 4.10: DNaseI cleavage patterns of template 7 with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

The results of similar footprinting experiments with template 7 are shown in Figure 4.10 on the previous page. Once again the BAU-containing oligonucleotide produces a footprint within the REPSA-selected region, even though it does not contain a canonical binding site for this TFO. However this footprint is only seen at the highest TFO concentration (1 µM), though it is accompanied by a slight enhancement, which is visible down to 0.3µM.

```
                5′ – BBBBBBCBT – 3′
                5′ – PPPPPPCPT – 3′
5′ – AAGACCAGAGCACAAAAAACAACCCTTCA – 3′
3′ – TTCTGGTCTCGTGTTTTTTGTTGGGCCGT – 5′
                              *
```

Figure 4.11; Sequence of template 7 showing the positions of the footprints and enhancements generated by the TFOs. In Figure 4.10 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The red bases show the footprint with the BAU-TFO and the red and blue bases shown the extended footprint in the P+L gels. The enhancement was only seen with the BAU-TFO. The probable binding site of the TFOs is underlined and the TFO sequences are shown above.

When incubated with the P-substituted TFO alone no footprint was evident on this template, though addition of 10 µM ligand produced a clear footprint which persists to 10 µM TFO in the absence of magnesium chloride and about 3 µM in its presence. No enhancements are visible on either of these gels, though the 3' edge of the footprint is unclear; there are lighter bands extending much further down the gels which may indicate some further secondary binding (though there is no evidence for secondary binding with the BAU TFO). Again the footprints generated by the P-TFO are one base longer at the 5' end than the footprint in the BAU gel. No footprint was produced by the T containing TFO, with or without ligand (not shown).

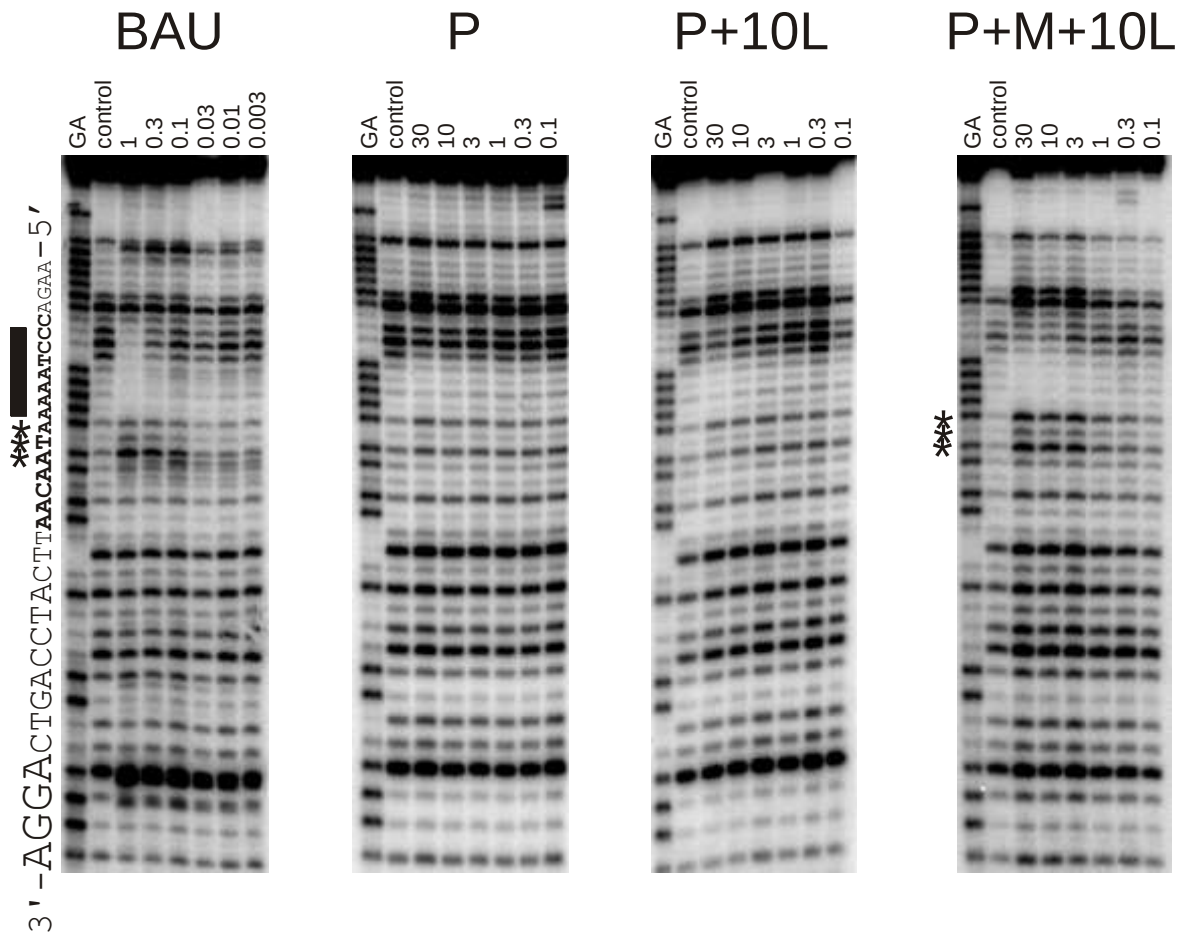


Figure 4.12: DNaseI cleavage patterns of sequence 10 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

The results of footprinting experiments with template 10 are shown in Figure 4.12 on the previous page. With the BAU substituted TFO there is a footprint on this template at 1 µM only. Three bands are enhanced at the 3' end of this footprint, which are also visible down to 0.1 µM TFO.

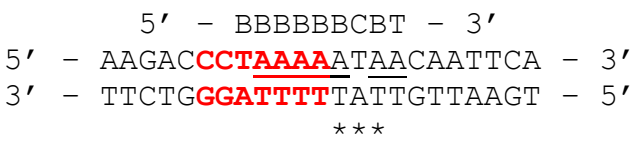


Figure 4.13; Sequence of template 10 showing the position of the footprint and enhancement with the BAU substituted TFO. In Figure 4.12 the upper (purine-rich) strand was radiolabelled at the 3' end. The footprint is shown in red and the enhancements are indicated by an asterisk. The probable binding site of the TFO is underlined and the TFO sequence is shown above.

No footprint or enhancement is visible on the P or P+10L gels, though the intensity of some bands is attenuated on the P+M+10L at 30 and 10 µM TFO. This is at the same position as the BAU footprint, although it is not so well defined. Three bands are also enhanced on this gel at the same position as seen with BAU, which persist to about 3-1 µM. No footprint was seen with the T containing TFO, with or without ligand (not shown).

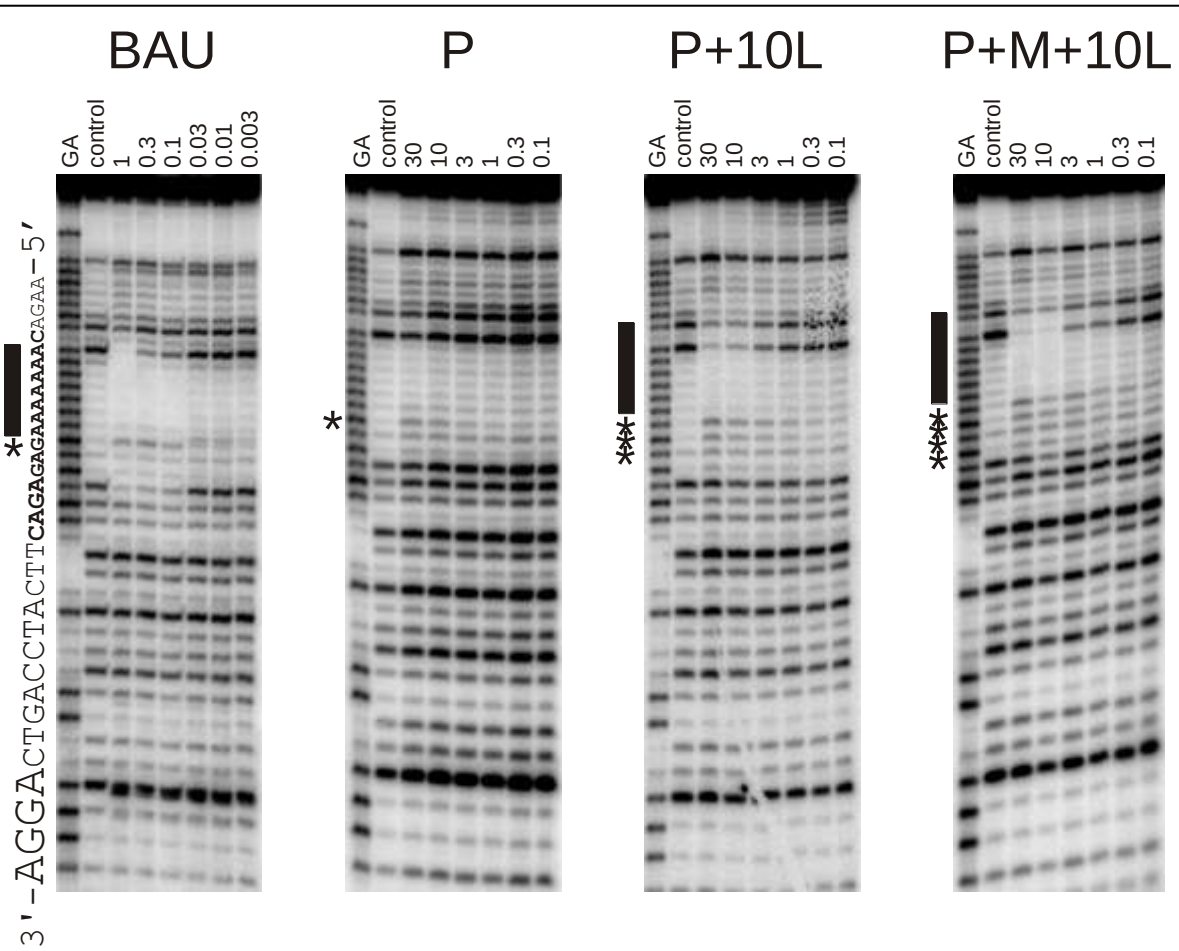


Figure 4.14: DNaseI cleavage patterns of sequence 4 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (µM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

Figure 4.14 on the previous page shows the results of similar experiments with sequence 4. The BAU-substituted TFO produces a very clear footprint on this template at 1 µM and some attenuation is visible down to 0.1 µM. There is also a single enhancement at the 3' end of this footprint, which persists to 0.1 µM. An unusual feature of this gel is that the intensities of some of the bands below the enhancement are also attenuated at 1 - 0.1 µM TFO, indicating that there may be secondary binding of this TFO, this is indicated in green in the Figure below;

```
        5′ – BBBBBBCBT – 3′
        5′ – PPPPPPCBT – 3′
5′ – AAGACAAAAAAAGAGAGACTTCA – 3′
3′ – TTCTGTTTTTTTCTCTCTGAAGT – 5′
                  ****
```

Figure 4.15; Sequence of template 4 showing the positions of the footprints and enhancements generated by the different TFOs. In Figure 4.14 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The bases in red and blue show the footprint in the BAU gel, red and green show position of the footprint on the P+L gels. The black asterisk shows the position of the enhancement with the BAU-TFO, the red asterisks show the additional enhancements on the P+L gel, and the blue asterisk shows a further additional enhancement only seen on the P+M+L gel. The probable binding site of the TFOs is underlined and the TFO sequences are shown above. The base pair in italics indicates a possible secondary binding site.

With P only there is no visible footprint on this template but a slight enhancement can be seen at the same position as on the BAU gel at 30 and 10 µM TFO. Addition of naphthylquinoline produces a clear footprint; without magnesium chloride this extends to 10 µM TFO, with three enhancements at the 3' end which persist to 1 µM. This footprint is also one base shorter at the 3' end than the BAU footprint. With ligand and magnesium chloride the footprint is also visible to 10 µM; it starts at the same position as the previous gel at the 3' end, but extends two bases further than the BAU or P+L gels at the 5' end. There are also four enhancements visible at the 3' end on this gel, down to round 1 µM. No footprint was seen with the T-containing TFO, with or without ligand (not shown).

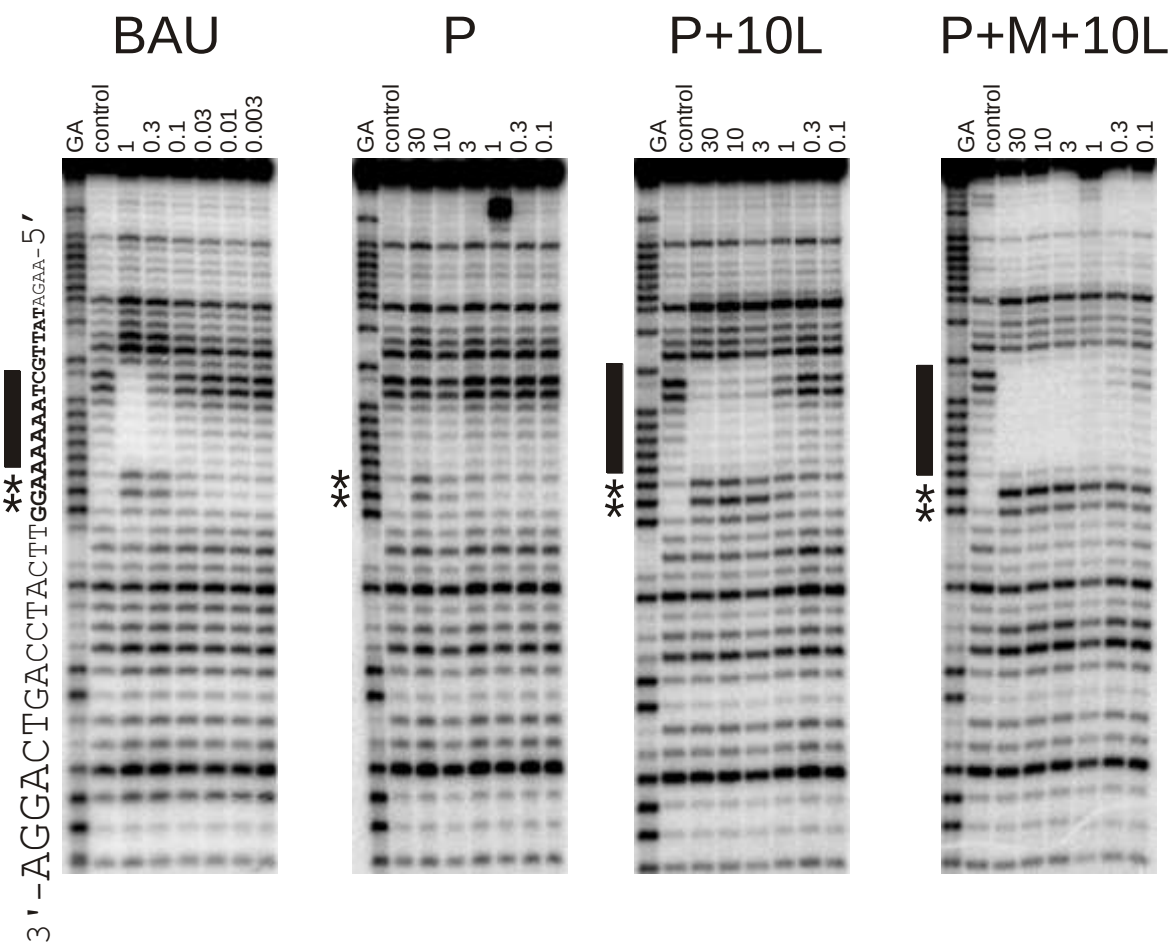


Figure 4.16: DNaseI cleavage patterns of sequence 5 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels, lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

The results of footprinting experiments with template 5 are shown in Figure 4.16 on the previous page. There is a clear footprint with the BAU-substituted TFO down to a concentration of 1 / 0.3 µM, and this is accompanied by enhanced cleavage at the 3'- (lower) end, which persists to around 0.1 µM. In this instance two bands show enhanced DNase I cleavage and this effect is also seen with the other TFOs on this template. The location of the footprint and enhancements from the gels are shown in Figure 4.17 below;

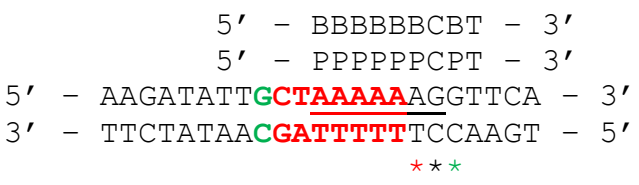


Figure 4.17; Sequence of template 5 showing the positions of footprints and enhancements with different TFOs. In Figure 4.16 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The bases in red show the position of the footprint generated by the BAU-TFO; the green base pair shows how the footprint was extended in the P+L gels. The enhancements are indicated by asterisks, red and black show enhancements on the BAU and P+L gels, black and green show the enhancements on the P+M+L gel only (shifted one base in the 5' direction). The probable binding site of the TFOs is underlined and the TFO sequences are shown above.

The propargylamino-dU-containing oligonucleotide does not produce a footprint on this template, even at a concentration of 30 µM, though enhanced cleavage is evident at the same positions as seen with the BAU-containing TFO at the highest concentration. On addition of the triplex-binding ligand a clear footprint is evident with the propargylamino-dU-containing oligo which appears to extend for one base further in the 5'-(upper) direction compared to the BAU TFO. This footprint is seen with and without magnesium chloride; without magnesium the TFO binds down to around 3 µM and with magnesium to about 0.3 µM. The enhancement is different between these two gels; without magnesium there is a double enhancement at the same position as seen for BAU and P down to 1 µM, with magnesium the enhancement shifts one base in the 3' direction and is seen at all TFO concentrations. This may indicate that the TFO is slipping on the A tract and binding slightly further down the sequence. No footprint can be seen with the T containing TFO, with or without ligand (not shown).

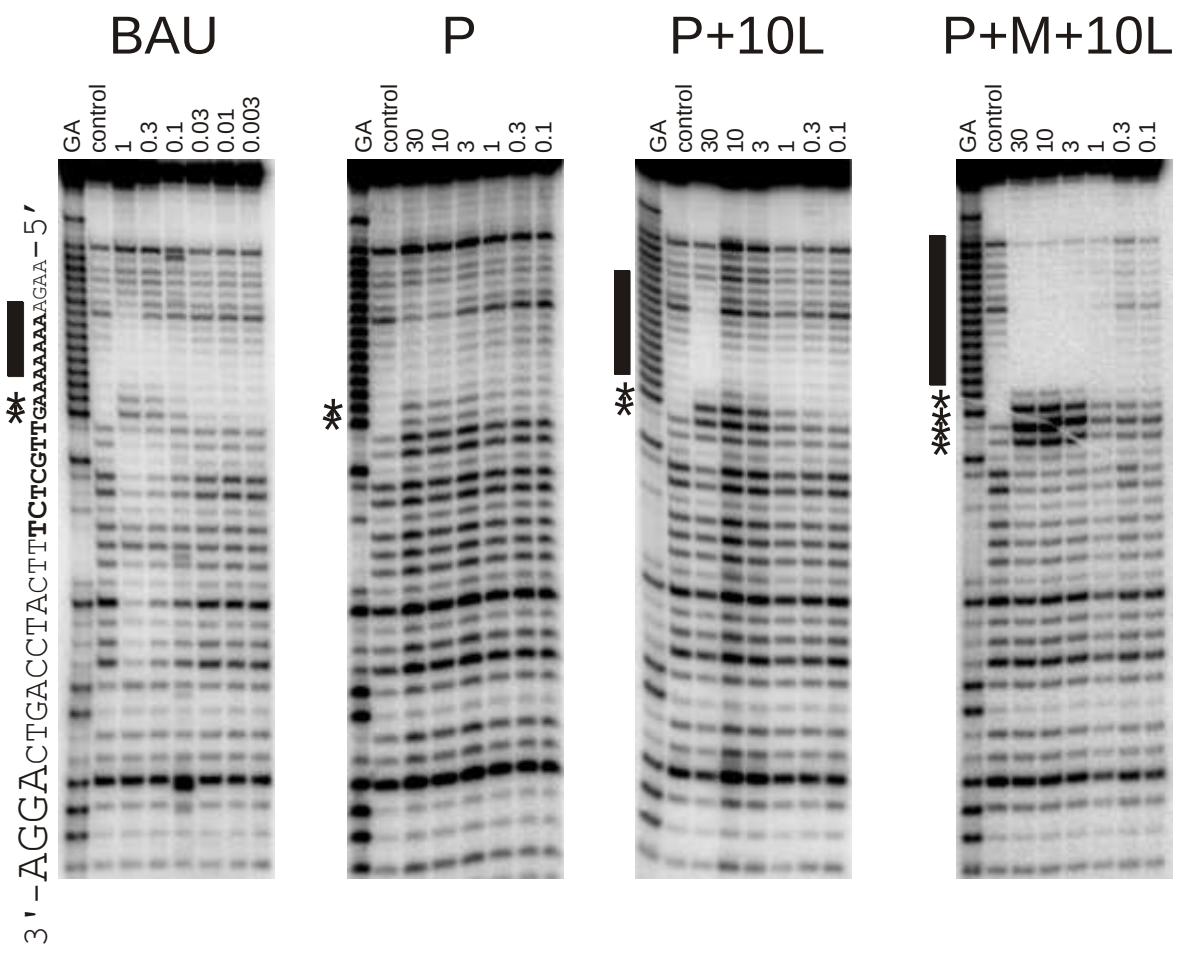


Figure 4.18: DNaseI cleavage patterns of sequence 2 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels, lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

Figure 4.18 on the previous page shows the results of similar footprinting experiments on template 2. With the BAU-substituted TFO a footprint can be seen at 1 µM, which is accompanied by two enhancements at its 3' end, visible down to around 0.1 µM. There are also some lighter bands below the enhancement at 1 and 0.3 µM TFO, indicating the possible presence of a secondary binding site on this template (shown in italics in Figure 4.19 below). With the propargylamino-dU containing TFO the same two enhancements are visible down to around 3 µM but there is only a slight attenuation of bands within the footprinting region at the two highest concentrations.

Figure 4.19; Sequence of template 2 showing the positions of footprints and enhancements with different TFOs and under different conditions. In Figure 4.18 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. Bases in red show the BAU-TFO footprint, the bases from orange to green show how the footprint is extended one base in the 5' and two bases in the 3' direction on the P+L gel. Bases in red, green and blue show the footprint on the P+M+L gel. Enhancements are indicated by asterisks, black and red show the enhancements on the BAU gel, black and green are enhancements on the P+L gel, and all four enhancements are seen on the P+M+L gel. The probable binding site of the TFOs is underlined and the TFO sequences are shown above. A possible secondary footprint on the BAU gel only is shown in italics.

When the triplex binding ligand naphthylquinoline is added to the incubation a footprint becomes visible with the P TFO at 30 µM along with two enhancements at the 3' end; the footprint is also extended by a further two bases in the 3' direction compared with the BAU footprint. The enhancement is also shifted one base down in comparison with the BAU and P gels. The footprint on the P+M+10L gel appears to be completely different; the footprint extends all the way to the top of the gel down to 1 µM and there are now four enhancements at the 3' end at all concentrations. This is likely to be an artefact of the gel, but may indicate that the TFO is binding at multiple sites all the way up the template. This is unlikely as this part of the sequence is the same on all the fragments, as it was not part of the random region of the original REPSA template. The footprint is also one base shorter at the 3' end compared with the P+10L gel. The T containing TFO did not

produce any footprints or enhancements with this fragment, both with or without ligand (not shown).

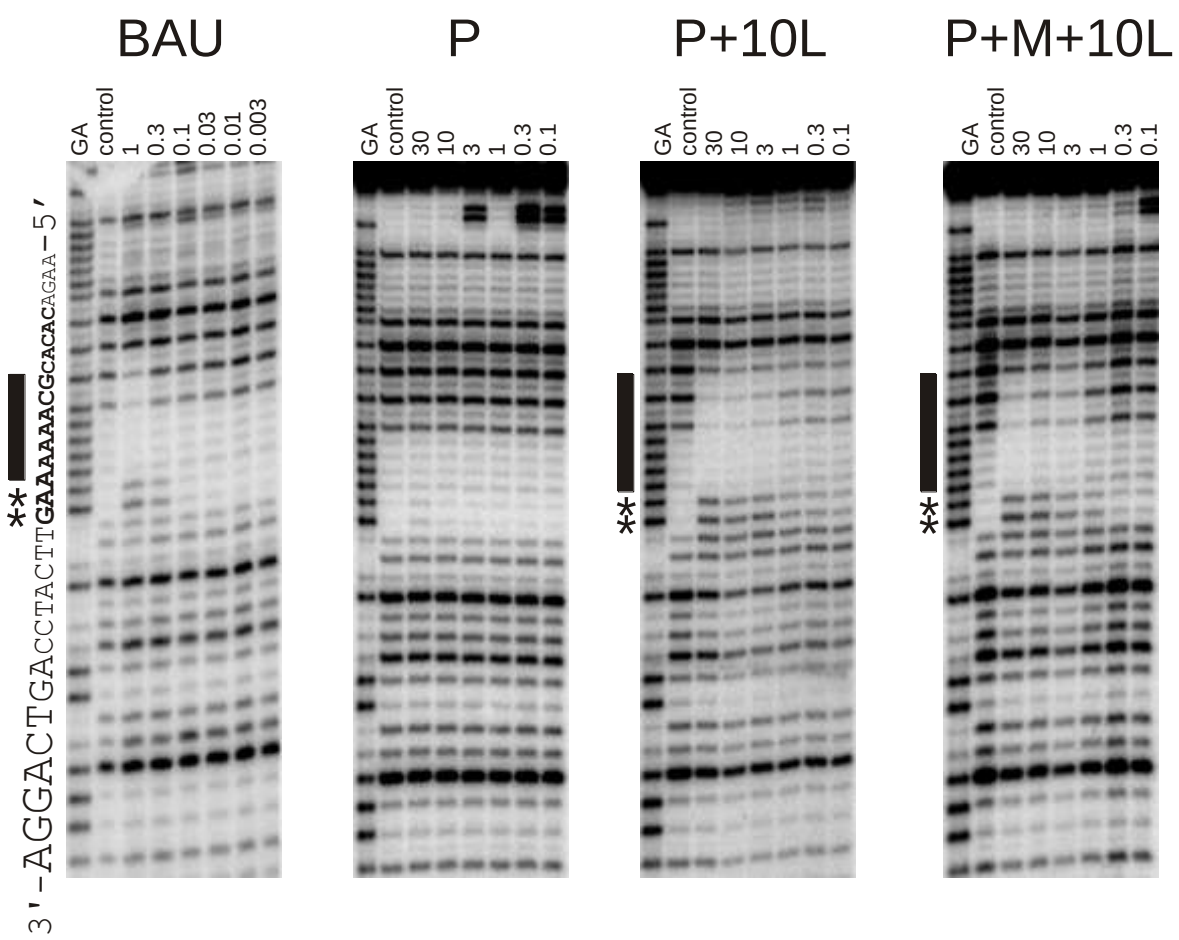


Figure 4.20: DNaseI cleavage patterns of sequence 13 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

Figure 4.20 on the previous page shows the results of similar experiments with sequence 13. The BAU-substituted TFO produces a footprint on this template at 1 µM only, but two bands of enhanced cleavage are visible below this, down to 0.3 µM. With the propargylamino-dU substituted TFO alone there was no evidence of binding, however when in the presence of the triplex binding ligand a footprint is visible to around 3 / 1 µM with or without magnesium chloride. The same two enhancements as seen with BAU are also evident again down to about 3 / 1 µM. As with previous templates the footprint with the P-TFO is slightly different to the BAU containing TFO and it extends one base further in the 5' direction. No footprint was seen with the T-containing TFO, with or without ligand (not shown).

```
                5′ – BBBBBBCBT – 3′
                5′ – PPPPPPCPT – 3′
5′ – AAGACACACGCAAAAAAGTTCA – 3′
3′ – TTCTGTGTGCGTTTTTTCAAGT – 5′
                          **
```

Figure 4.21; Sequence of template 13 showing the positions of footprints and enhancements with different TFOs and under different conditions. In Figure 4.20 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The bases in red show the position of the footprint generated by the BAU-TFO, the base in blue shows an additional base in the footprints on the P+L and P+M+L gels. Enhancements are indicated by asterisks, these were the same on all gels. The probable binding site of the TFOs is underlined and the TFO sequences are shown above.

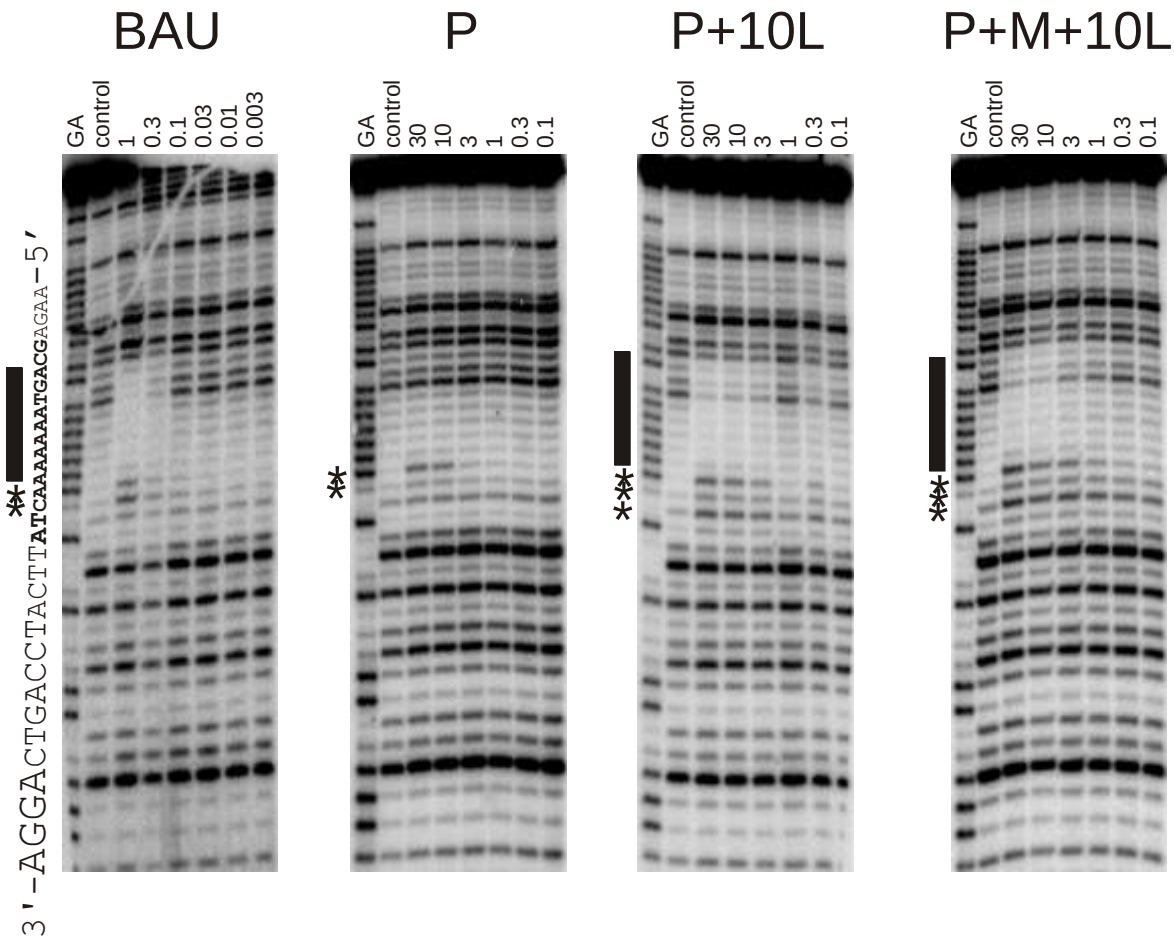


Figure 4.22: DNaseI cleavage patterns of sequence 11 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels, lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

Figure 4.22 on the previous page shows the results of similar experiments with sequence 11. The BAU-substituted TFO produces a footprint on this template down to 0.3 µM with two enhanced bands at the 3' end, visible at the same concentrations. No footprint is visible with the P-TFO, though enhancements can be seen at 30 and 10 µM at the same positions as produced with the B-TFO.

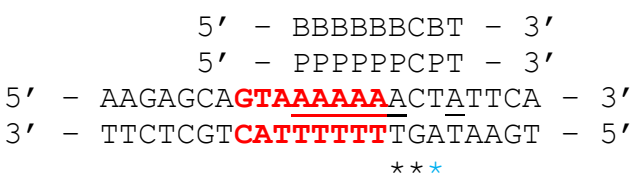


Figure 4.23; Sequence of template 11 showing the positions of footprints and enhancements with different TFOs and under different conditions. In Figure 4.22 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. Footprints are shown in red and enhancements are indicated by asterisks, the blue asterisk indicates an enhancement which was only seen on the P+L and P+M+L gels. The probable binding site of the TFOs is underlined and the TFO sequences are shown above.

When naphthylquinoline is added there is a clear footprint with the P-TFO which covers the same region as the B-TFO, down to 3 µM. Enhanced cleavage is also evident at these concentrations but in this case three enhanced bands can be seen; the two seen on the B and P gels plus an addition enhancement to the 3' (lower) side of these. The P+M+10L gel shows the same footprint and enhancement pattern, again to around 3 µM. Unlike the previous templates the position and size of the footprint has not been affected by the type of substitution; however an extra enhancement can be seen on addition of ligand. With the control T-containing TFO no footprint or enhancement was seen, with or without ligand (not shown).

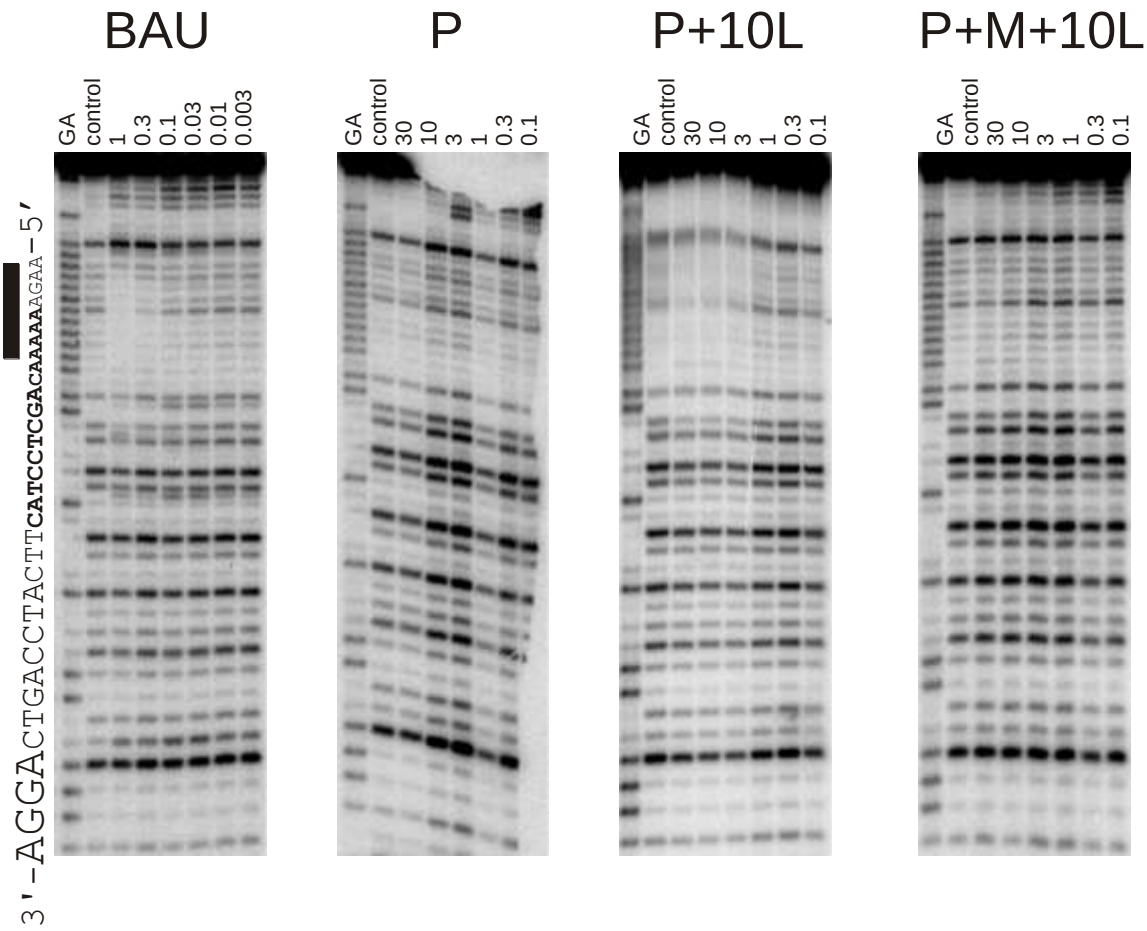


Figure 4.24: DNaseI cleavage patterns of sequences 14 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side and footprints are indicated by black boxes.

The results of similar experiments with sequence 14 are shown in Figure 4.24 on the previous page. On this template there is a clear footprint with the BAU-substituted TFO at 1 µM and also a slight footprint at 0.3 µM. However, no enhancements are evident and the P-TFO does not produce a footprint under any conditions. The T-containing TFO also does not footprint on this template, with or without ligand (not shown).

```
        5′ – BBBBBBCBT – 3′
5′ – AGGAAGAAAAACACAGCTCCTACTTCA – 3′
3′ – TCCTTCTTTTTTGTCGAGGATGAAGT – 5′
```

Figure 4.25; Sequence of template 14 showing the positions of footprints with the BAU substituted TFO. In Figure 4.24 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. Footprints are shown in red, the probable binding site of the TFO is underlined and the TFO sequence is shown above

.

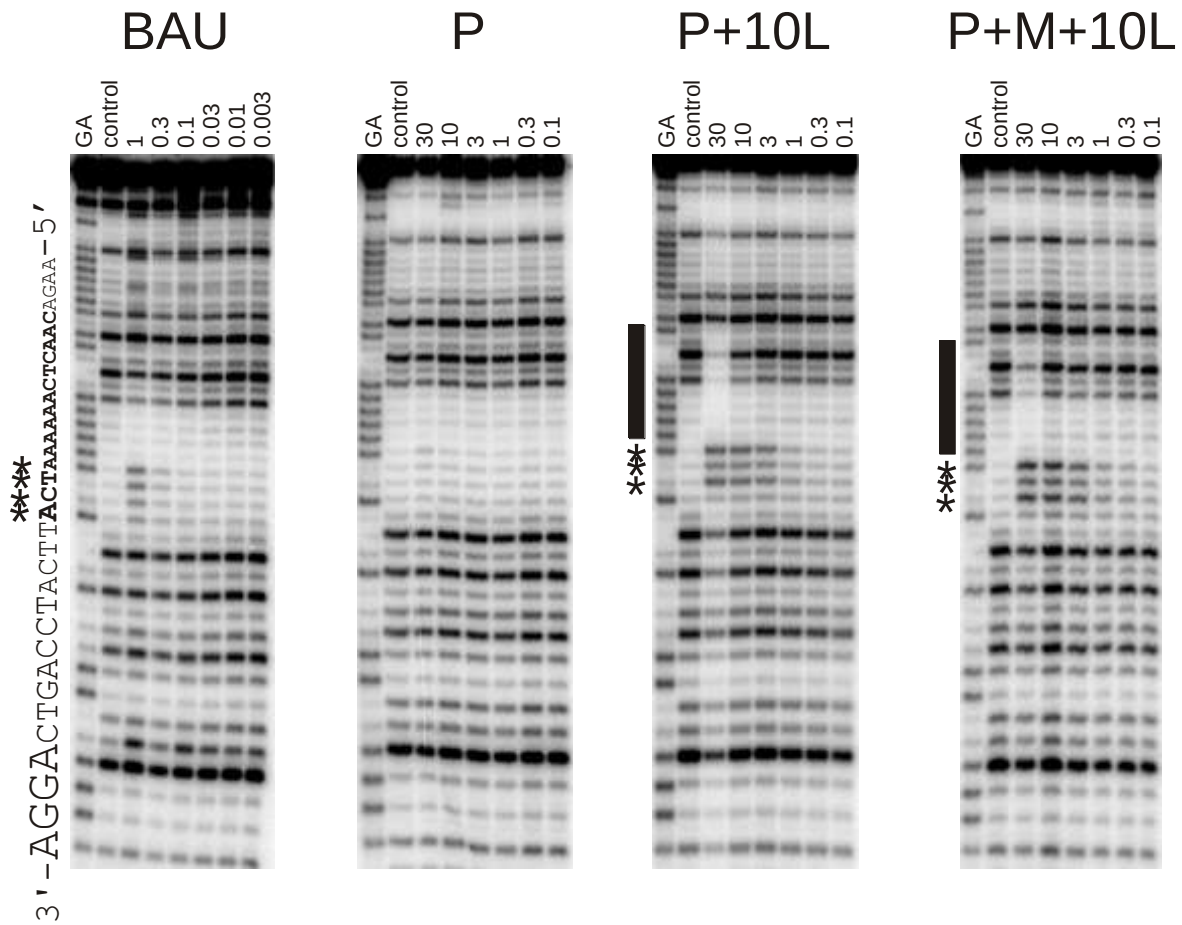


Figure 4.26: DNaseI cleavage patterns of sequence 9 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

The results of similar experiments with sequence 9 are shown in Figure 4.26 on the previous page. The BAU-substituted TFO does not produce a clear footprint on this sequence, though there are four enhancements at the 3' end of the A tract at the highest concentration.

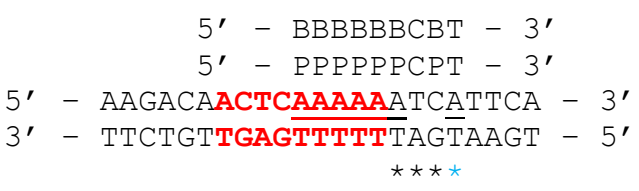


Figure 4.27; Sequence of template 9 showing the positions of footprints and enhancements with different TFOs and under different conditions. In Figure 4.26 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The position of the footprints seen on the P+L and P+M+L gels are shown in red and enhancements on these gels are indicated by black asterisks, these enhancements were also seen on the BAU gel, along with an additional enhancement indicated by the blue asterisk. The probable binding site of the TFOs is underlined and the TFO sequences are shown above.

There is no evidence of binding with the propargylamino-dU containing TFO alone, but a clear footprint is produced when the ligand is added. However, a footprint is seen with P+10L and P+M+10L at the highest TFO concentration (30 μM). These footprints are also accompanied by several enhancements at the 3'-(lower) end, which extend to around 3 μM TFO. No footprint was seen with the T-containing TFO, with or without ligand (not shown).

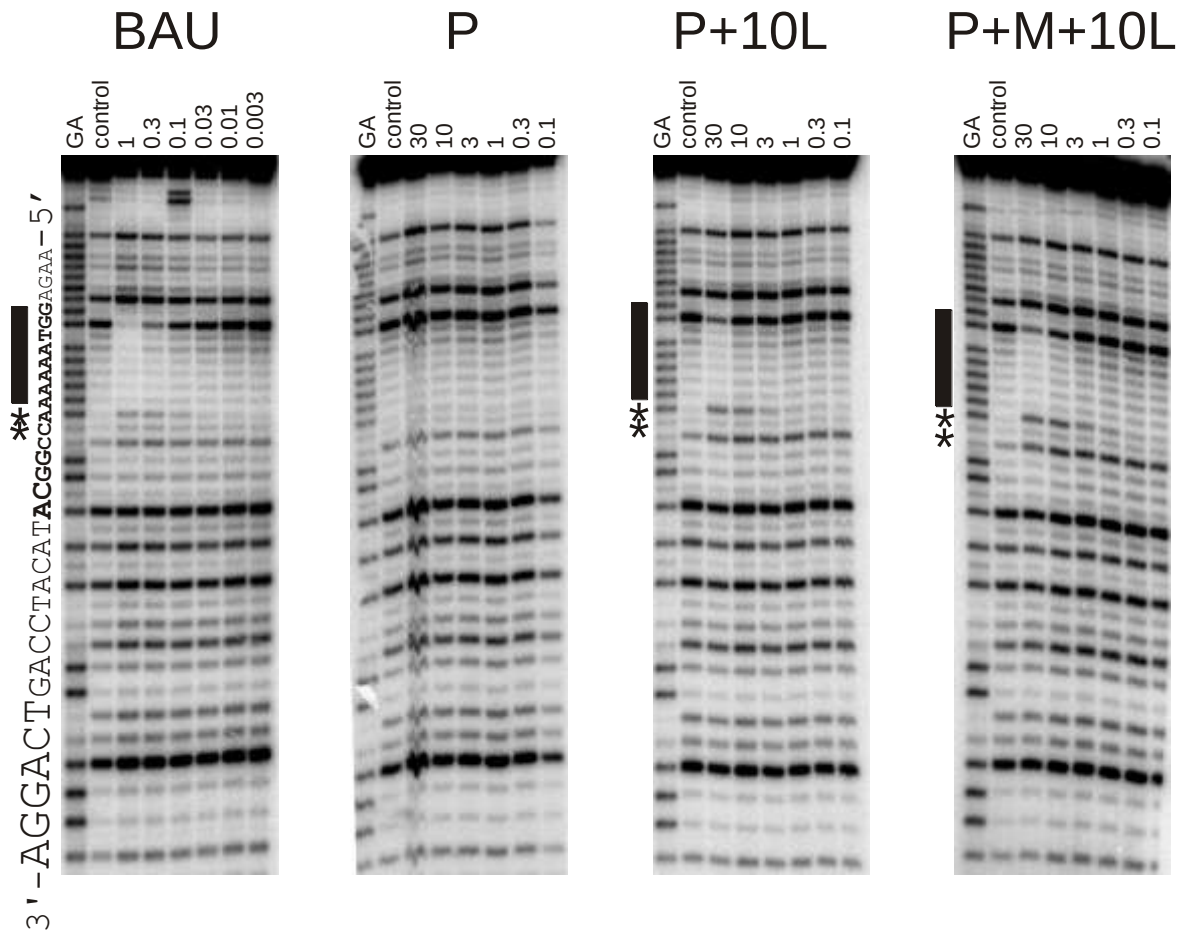


Figure 4.28: DNaseI cleavage patterns of sequence 8 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (µM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

The results of similar experiments with sequence 8 are shown in Figure 4.28 on the previous page. The BAU-substituted TFO produces a footprint at the highest concentration only (1 µM) on this template and produces two enhanced bands at the 3'-end, visible to around 0.1 µM. The enhancements have different intensities; the upper one at the end of the A_6 tract is much stronger than the one located one band further down the gel. This may indicate that the 3'-end of the TFO is either binding in more than one location (slippage) or that it is fraying for a significant proportion of the time. With the P-substituted TFO there is no evidence of a footprint, though the bands are slightly attenuated in the presence of the ligand at the highest concentration (30 µM) in both the presence and absence of magnesium chloride. Again a double enhancement is visible with the top band darker than the lower one. This pattern is also seen on the BAU gel. Without magnesium these enhancements are only visible to around 3 µM, but with magnesium they extend down to 1 µM. Unlike many of the other templates the footprints with BAU or P appear to cover the same area. No footprint was seen with the T-containing TFO, with or without ligand (not shown).

```
          5′ – BBBBBBCBT – 3′
          5′ – PPPPPPCPT – 3′
5′ – AAGAGGTAAAAAACCGGCATTCA – 3′
3′ – TTCTCCATTTTTTGGCCGTAAGT – 5′
                   **
```

Figure 4.29; Sequence of template 8 showing the positions of footprints and enhancements with different TFOs and under different conditions. In Figure 4.28 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The positions of footprints seen on the BAU, P+L and P+M+L gels are shown in red, enhancements are indicated by asterisks, the blue asterisk indicates an enhancement only seen with the BAU-TFO. The probable binding site of the TFO is underlined and the TFO sequence is shown above.

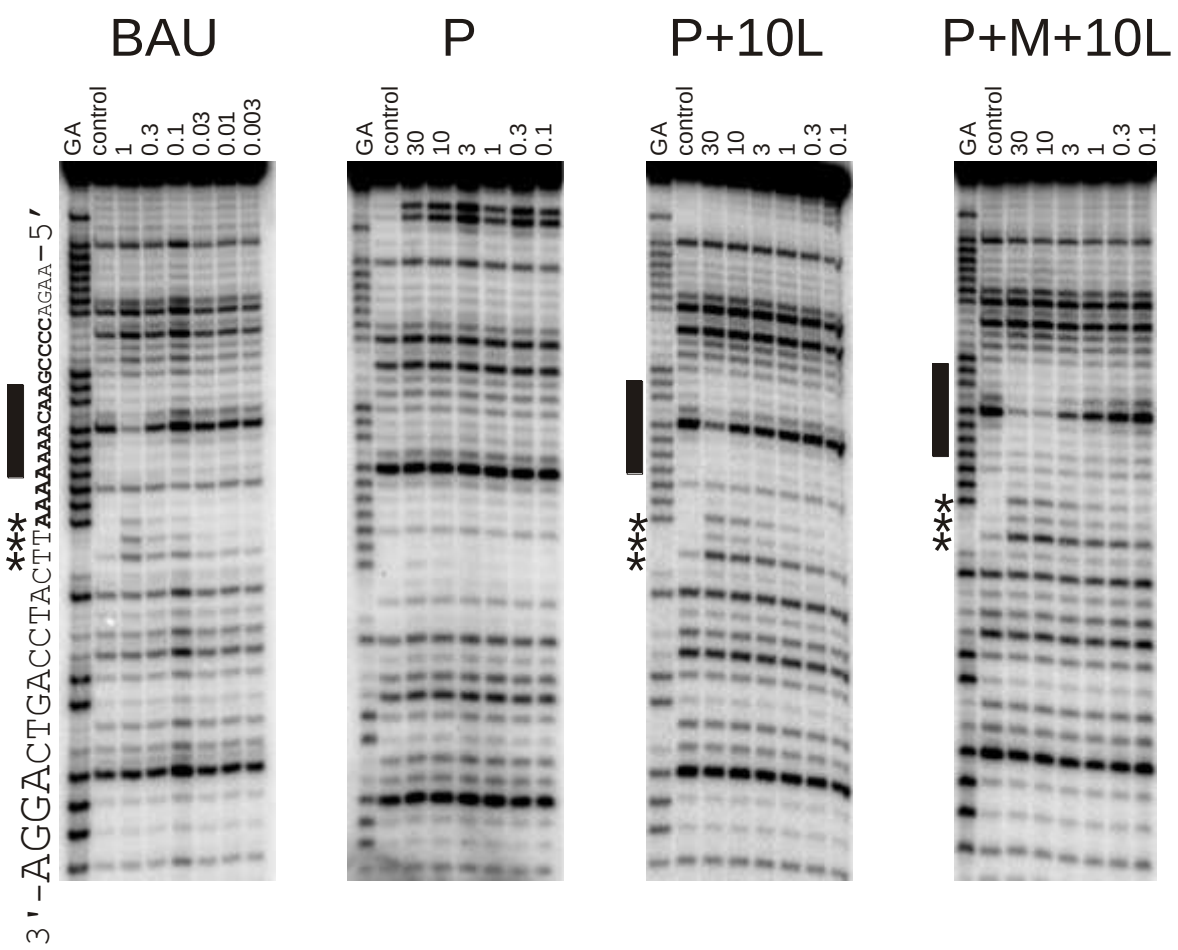


Figure 4.30: DNaseI cleavage patterns of sequence 3 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

Figure 4.30 on the previous page shows the results of similar footprinting experiments with sequence 3. Although there is evidence of TFO binding on this template the footprints are slightly unusual as there is a strong band in the centre of the footprint with all TFOs. With the BAU substituted TFO there is a reasonably clear footprint on the A tract at the highest TFO concentration, which is accompanied by three enhancements, persisting to around 0.1 µM.

```
                5′ – BBBBBBCBT – 3′
                5′ – PPPPPPCPT – 3′
5′ – AAGACCCCGAACAAAAAAATTCA – 3′
3′ – TTCTGGGGCTTGTTTTTTTAAGT – 5′
                           ***
```

Figure 4.31; Sequence of template 3 showing the positions of footprints and enhancements with different TFOs and under different conditions. In Figure 4.30 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. The position of the footprint generated by the BAU-TFO is shown in red; an addition base in the footprints in the P+L and P+M+L gels is shown in blue. Enhancements are indicated by asterisks, these were the same on all gels. The probable binding site of the TFOs is underlined and the TFO sequences are shown above.

With the P TFO alone there is a very faint enhancement at the 3' end at 30 / 10 µM TFO, at the same position as the highest enhancement on the BAU gel. When the naphthylquinoline ligand is added there is again an attenuation of the bands in the expected region at the highest concentration. Three enhancements at the same position as the BAU gel are also visible to around 3 µM. When magnesium chloride are also added to the incubation mix the footprint becomes slightly clearer and it is seen down to 10 µM TFO. Three bands of enhanced cleavage are again visible at the 3'-end of the footprint, which persist down to 3 µM.

The three enhanced bands show a different concentration dependence in each case. The one at the 3' end (bottom on the gel) is the darkest and extends to the lowest concentrations; it is visible at all concentrations on both of the P-TFO and ligand gels, and down to around 0.1 / 0.03 µM on the BAU gel. The two higher enhancements seem to persist to around the same concentration; about 3 / 1 µM on the last two gels and around 0.1 µM with the BAU TFO. No evidence of TFO binding was seen with the T-containing TFO (not shown).

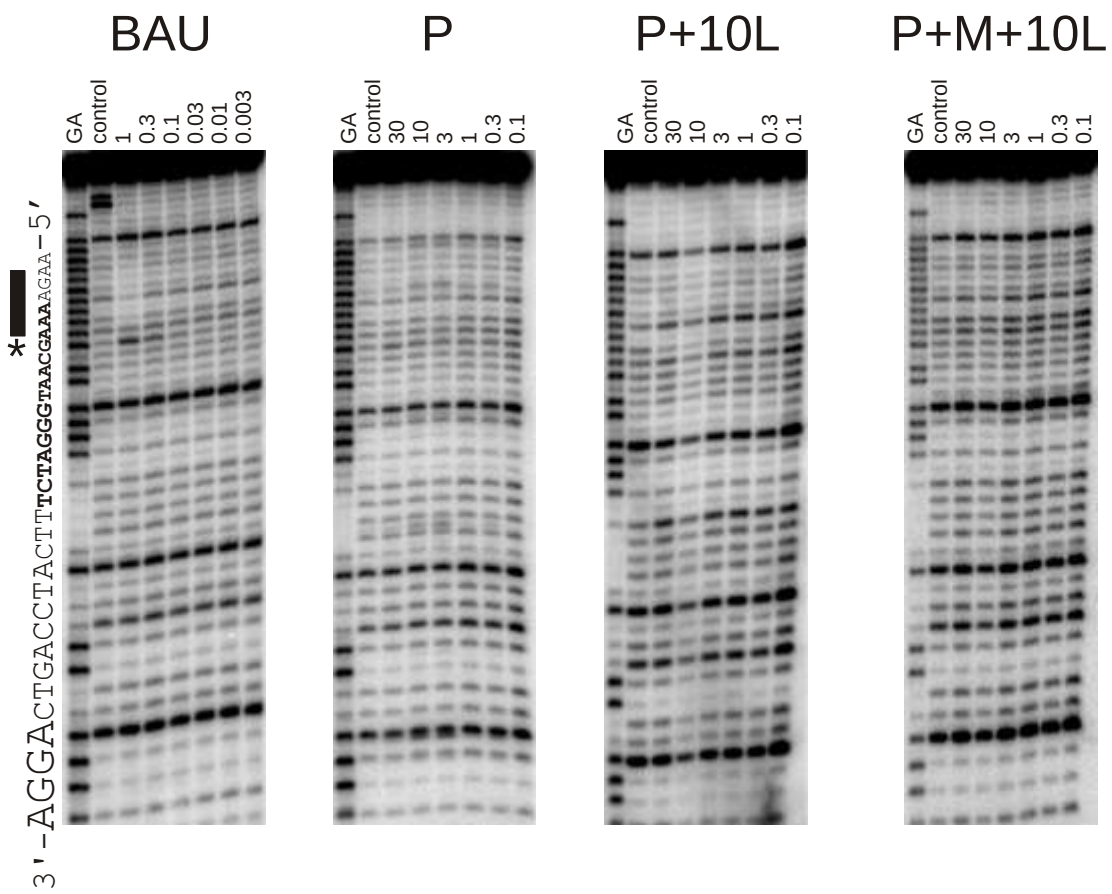


Figure 4.32: DNaseI cleavage patterns of sequences 12 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side, footprints are indicated by black boxes and asterisks indicate enhancements.

Figure 4.32 on the previous page shows the results of similar experiments with sequence 12. On this template the BAU-substituted TFO shows a slight footprint at the highest concentration (1 µM) with an enhancement that persists down to 0.3 µM. Unlike all the other templates almost half of this footprint actually lies outside of the original random region of the REPSA template, meaning that it has actually bound to part of the sequence common to all the templates (the original random region is shown in italics in Figure 4.33). This template contains one of the shortest A-tracts which is only four bases long.

```
      5′ – BBBBBBCBT – 3′
5′ – AGGAAGAAAAGCAATGGGATGGGATCTTTCA – 3′
3′ – TCCTTCTTTTCGTTACCCTACCCTAGAAAGT – 5′
                 *
```

Figure 4.33; Sequence of template 12 showing the positions of footprints and enhancements with the BAU substituted TFO. In Figure 4.32 the upper (purine-rich) strand of this sequence was radiolabelled at the 3' end. Footprints are shown in red and enhancements are indicated by asterisks. The probable binding site of the TFO is underlined and the TFO sequence is shown above. The original random region of the REPSA template is in italics.

The P- and T- containing TFOs failed to produce a footprint or enhancement under any conditions with this sequence.

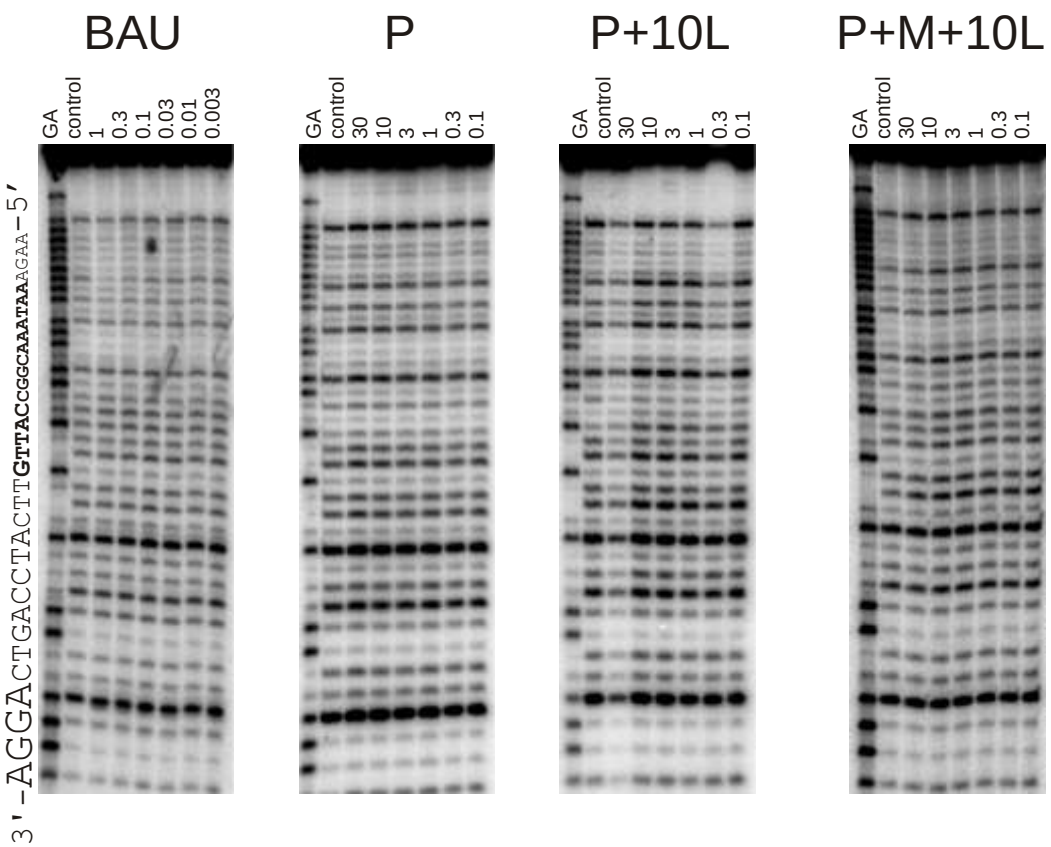


Figure 4.34: DNaseI cleavage patterns of sequence 6 from the REPSA experiment with two of the TFOs shown in Figure 4.1. TFO concentrations (μM) are shown above the gels; lanes labelled GA are markers specific for purines. Control lanes show the DNaseI cleavage pattern in the absence of TFO. The sequence of the DNA fragment is shown along the left side.

The results of similar footprinting experiments with sequence 6 are shown in Figure 4.34 on the previous page. It can be seen that none of the TFOs alter the DNase I cleavage pattern of this template under any conditions. This template contains an A tract like the other templates but with a pyrimidine interruption in the centre (see sequence in Table 4.2 on the following page). It is probably because of this interruption that even the BAU-containing TFO does not produce a footprint on this template.

4.4 Discussion

Binding of TFOs comprised of solely natural nucleotides to secondary sites has been seen previously at high TFO concentrations (82). Binding of TFOs to these sites involved one or two mismatches, fraying of the third strand away from the duplex or loop formation of the TFO (82). TFOs containing BAU substitutions however were shown to generate extremely stable triplexes and individual substitutions have are highly selective for AT (187). BAU has enhanced discrimination against YR bases in the duplex compared with T. BAU also shows high selectivity for AT over GC bases pairs in the duplex, although BAU.GC triplets are more stable than T.GC or BAU.YR triplets (187). TFOs with the same sequence as the one used in this thesis but containing 3 or 4 BAU residues showed no evidence of secondary binding on *tyr*T (186). It is therefore likely that BAU reduces the stringency of TFOs only when there are large numbers of consecutive substitutions. The most likely model for the interaction of the TFOs used in this thesis on secondary sites is binding of the run of modified nucleotides to the run of As and fraying of the 3' end CXT sequence.

The experiments described in this chapter have examined TFO binding to the templates selected by REPSA. From these templates 14 were chosen for footprinting with the TFOs shown in Figure 4.1 on page 104. Magnesium chloride and the triplex binding ligand naphthylquinoline were also added for some of the footprinting experiments in order to promote binding. It was hoped that a pattern of binding affinity might emerge from these experiments to indicate the TFOs preference for differences in template sequence.

On the following page is a table showing the lowest concentration at which the footprints and enhancements are visible on each template, colour coded by

concentration. It is important to note that the BAU TFO was footprinted at a lower concentration range than the propargylamino-dU TFO.

TFO	Sequence of probable binding site	BAU		P		P + 10L		P + M + 10L	
Template		Ftp	Enha	Ftp	Enha	Ftp	Enha	Ftp	Enha
tyrT	AAAAAAGAA	0.1	0.1		10	3	3	3	1
Exact	AAAAAAGAA	0.1				0.3		0.3	
1	AAAAAAACAA	0.3			3	3		0.3	0.1
7	AAAAAACAA	1	0.3			10		3	
10	AAAAATAA	1	0.1					10	1
4	AAAAAAAGAG	0.1	0.1		10	10	1	10	1
5	AAAAAAGGT	0.3	0.1		30	3	1	0.3	0.1
2	AAAAAAAAGTT	1	0.1		3	30	0.3	1	0.1
13	AAAAAAGTT	1	0.3			3	1	3	1
11	AAAAAAACTA	0.3	0.3		10	3	3	3	1
14	AAAAAACAG	0.3							
9	AAAAAATCA		1			30	3	30	3
8	AAAAAACCG	1	0.1				3		1
3	AAAAAAATTC		0.1		30		3	10	3
12	AAGAAAAGC	1	0.3						
6	AAATAAAC								

Table 4.2; Tthe lowest TFO concentration in µM that footprints and enhancements are visible at on different gels. Colour coded by concentration, Ftp = footprint, Enha = enhancement.

The BAU-TFO produces clear footprints and / or enhancements on all but one of the REPSA templates tested, indicating that the REPSA process was successful in selecting for sequences which were bound by the TFO. The highest concentration of the BAU-substituted TFO used for footprinting in this chapter was 1 µM, whereas 5 µM TFO was used for REPSA selection. However the conditions used in REPSA were much less conducive to binding than those used in the footprints in this Chapter, so a higher TFO concentration was needed.

Looking at Table 4.2 there are several general observations that can be drawn. None of the TFOs produced footprints or enhancements on all templates; not all footprints were accompanied by enhancements and on some gels there are only enhancements and no footprints. The enhancement (if present) is always seen at the same or lower concentration than the footprint. The BAU TFO only footprints down to 0.1 µM even on the exact template, although concentrations as low as 0.003 µM were used.

Significantly higher concentrations of the P TFO were used, and magnesium chloride and / or the triplex binding ligand naphthylquinoline were needed to promote binding. Binding of the propargylamino-dU TFO is always improved by addition of ligand and / or magnesium. This TFO does not bind to any template without the ligand, but did produce an enhancement on some templates. The P TFO binds more than half of the templates on addition of 10 µM ligand, and when 5 mM magnesium chloride is also added only four templates remain unbound. Under these conditions this TFO actually binds down to the same concentration as the BAU TFO on some templates (1, 5 and 2).

The BAU-TFO failed to produce a footprint on only three of the templates and enhancements were seen on two of these. Template 6 contains a pyrimidine interruption in the A tract and shows no evidence of binding by any of the TFOs under any conditions. Templates 9 and 3 also show no footprint with the BAU-TFO although there is an enhancement on both of these gels and footprints are produced by the P-TFO in the presence of the ligand. This is difficult to explain as there is little difference between these templates and several of the others. They both contain six or more As and the three bases at the 3' end are similar to several other templates. They do both contain a T and C in these final three bases, but so too does template 11. It may be that the flanking bases of the TFO binding site may have impeded binding in some way, resulting in very weak binding so that only an enhancement was seen at these concentrations.

The BAU-TFO footprints down to the same concentration on the *tyr*T and exact templates, there is no enhancement on the exact template, but this is not unusual for triplex footprints with fragments labelled on the pyrimidine-rich strand (284). The P-TFO does not footprint on either of these control templates, although again there is an enhancement on *tyr*T. Surprisingly, on addition of ligand the P TFO binds more strongly to the exact template (E) than to *tyr*T, even though both of these templates containing the exact recognition sequence for the TFO. This difference in binding intensity must therefore arise from the difference in flanking sequences between the two templates. As there is no difference in the binding intensity of BAU on these two templates, this difference may be limited to the P TFO, or perhaps is only seen on addition of the ligand.

The REPSA sequences in this chapter were presented in the order of how much they resembled the exact binding site of the TFOs, and how well they bind the BAU TFO. Templates 1, 7 and 10 contain a run of 5-7 As followed by a single uncomplementary base, then two more As. This means that all of the nucleotides

in the TFOs, except for the C residue, can potentially bind to a complementary base in the duplex. Templates 1 and 7 contain the same sequence, though the run of As in template 1 is longer than in template 7 and the flanking sequences are different. Template 1 is bound slightly better than template 7 by the BAU TFO, presumably because of these differences. Template 10 contains a very similar sequence to 1 and 7; all have a pyrimidine interruption in place of the G in the exact binding site, but in template 10 this is a T rather than a C as in templates 1 and 7. Template 10 also has a shorter run of As (only 5) which means it cannot bind the whole run of BAU residues in the TFO. Despite this template 10 does bind the BAU TFO with a similar affinity to template 7.

Templates 4, 5, 2 and 13 all contain a run of 6-8 As followed by a G which could bind the C in the TFO. The final two bases are varied and only template 4 contains any complementary bases in this region. Templates 2 and 13 have the same sequence except that the run of As in 13 is one base shorter (7 rather than 8). Out of these four sequences template 4 binds the BAU TFO the strongest, presumably because the G in this sequence is followed by an A to bind the final BAU residue in the TFO, so this template has only one mismatch opposite the 3' T base in the BAU TFO which would be the weakest binding (T.AT) triplet. As might be expected the BAU-TFO binds to this template with the highest affinity out of all of the REPSA templates; producing a footprint at the same concentration as the two control templates.

Templates 11, 14, 8, 9 and 3 all contain runs of 7 or 8 As, but have no similarity to the final three bases of the TFO. Templates 12 and 6 both contain an interruption in the run of As and neither of these bind the P TFO under any conditions. Template 6 is also the only template to show no evidence of binding by the BAU TFO.

Table 4.2 (page 145) shows that despite the differences in template sequence there does not appear to be a strong pattern in the binding intensities of the TFOs on the different templates. However, some conclusions have been draw from these data. The BAU substituted TFO does not bind to the majority of the REPSA templates as well as it binds to *tyr*T or the exact template. Template 4 is the only template that shows a footprint and enhancement down to the same concentration as these two controls. This template also contains the longest run of bases contiguous with the exact binding site of the TFO, only the final 3' base is a mismatch. Templates 1, 7 and 10 also match 8 out of the 9 exact template bases, but the mismatch is opposite the C in the TFO rather than the 3' end T.

Complementary binding of the C base therefore appears to be more important to triplex stability than binding of the terminal T.

As mentioned before none of the templates, including the controls, bind the propargylamino-dU TFO alone, although several gels show enhanced cleavage in similar regions to the BAU-TFO. When 10 µM ligand was added 10 out of the 16 templates were bound by this TFO and two of the others show some enhancements. The exact template is bound at the lowest concentration of all by a factor of 10. The *tyr*T template however is only bound as strongly as several of the other templates. The addition of magnesium chloride as well as ligand does not improve TFO binding to either of the control templates, however many of the other templates are bound at lower concentrations under these conditions. Several templates now appear to bind the TFO more strongly than *tyr*T, even though this is an exact match, and templates 1 and 5 are now bound at the same concentration as the E template.

Surprisingly it appears from these data that the BAU TFO actually discriminates better than P against non-exact binding sites. There is a large difference in BAU-TFO concentrations required to bind the exact sequence compared to the REPSA templates, whereas with the P-TFO the concentrations required to generate footprints (and hence have a measurable dissociation constant) are all similar and many templates are bound with higher or the same affinity as the exact templates. The sequence of the three nucleotides at the 3' end of the binding site appear to have relatively little effect on binding affinity; so long as there is a run of As the TFOs bind at similar concentrations. The run of As does have to be at least 5 long however; the templates containing only three or four consecutive As (12 and 6) do not bind the P TFO, although with four As the BAU TFO still binds relatively well (template 12).

Out of all of the REPSA sequences template 4 is bound with the highest affinity by the BAU-TFO. The only mismatch with this template is against the T at the 3' end of the TFO. Templates 1, 7 and 10 also only contain one mismatch but this is against the C base in the BAU TFO and this appears to be more destabilising. There are six templates which have two mismatches out of the three 3' terminal bases 5, 2, 13, 11, 14 and 9. The sequence of the 3' end of the TFO is CBT, and should bind the sequence GAA. Templates 5, 2 and 13 have a GXX pattern (X being a mismatch in relation to the TFO), 11 and 9 XXA and 14 XAX. There does not seem to be any correlation between the type of mismatch and the

binding intensity from these footprints. The templates with three mismatches at the 3' end are bound the least strongly (templates 8 and 3).

Apart from differences in binding intensity there were several other interesting differences between the BAU and P gels, as well as anomalies such as multiple enhancements and secondary binding sites.

The BAU TFO generated a footprint 11 bases long on *tyr*T and around 16 bases long on the exact template, although this footprint is unclear. This difference may be due to one being labelled on the purine and the other on the pyrimidine strand, rather than an actual difference in the number of bases bound by the TFO. All of the footprints generated by the BAU TFO on REPSA templates were shorter than either of these, generally 7 or 8 bases long; indicating that in all cases the TFO was only partially bound. This is presumably because the 3' end (CBT) of the TFO was not bound, or bound only partially. Template 12 was the template with the most mismatches, including an interruption of the A tract, that still produced a footprint. The footprint generated by the BAU TFO on this template was the shortest of all; only 6 bases long.

On all of the REPSA templates the TFO appears to favour binding at the 3' end of the A tract, leaving any superfluous As at the 5' end. This is possibly due to the TFO sequence at the 3' end being different (CXT with X being B, P or T depending on the TFO). This may bind preferentially to non-A bases.

Only two of the templates showed the same size and position of footprint with the BAU and P TFOs (11 and 8). All but one of the others was either shorter at the 3' end or longer at the 5' end. On *tyr*T and template 4 the footprint with P+10L covered one base less at the 3' end than the footprint with BAU. On templates1, 7, 5, 13 and 3 the footprint is one base longer at the 5'-end with P+10L compared to B. With the exact template the footprint with P+10L with or without magnesium chloride is four bases shorter at the 3' end and 2 bases longer at the 5' end than the BAU TFO generated footprint. However the edges of the footprints on this template are indistinct, especially with the BAU TFO.

Template 2 also shows a different pattern; with the P-containing TFO with ligand the footprint is 1 base longer at the 5' end like many other templates, but unlike any other it is also two bases longer at the 3' end, compared to the BAU TFO generated footprint. This is exaggerated on addition of magnesium chloride; the footprint is now one of the longest seen on any template at 14 bases; at least six bases longer at the 5' end, but now only one base longer at the 3' end compared to the BAU footprint. This template contains the same sequence as

template 13, but template 13 does not produce any of these irregularities in footprinting size. Template 2 is the only template with a run of eight As (template 13 has only 6); the majority of the other templates contain 6 or 7. It's possible that this particularly long run of As allows the TFO to bind at several different positions on different template molecules; producing the unusually large footprint under these conditions. There are also multiple enhancements on these gels indicating that the TFO is binding at multiple sites on the A tract.

Most of the templates show the same footprint size with P+10L and P+M+10L, however in a few the shift in the 5' direction is exaggerated. Template 2 is one of these and has already been discussed, template 4 is the other. With P+10L the footprint is 1 base shorter at the 3' end compared to the footprint generated by the BAU TFO, with P+M+10L however the footprint is also two bases longer at the 5' end.

Templates 14 and 9 show an interesting footprinting pattern with different TFOs. These templates have similar sequences but only the BAU TFO produced a footprint on 14, and the P-TFO only produced a footprint on 9 in the presence of the triplex ligand.

The bands showing enhanced cleavage by DNase I found at the 3' end of many of the footprints are also affected by the sequence of the template and the nature of the TFO. With purine labelled templates a single enhanced band is usually found at the 3' end of the footprint, caused by increased cleavage by DNase I, which is thought to be due to slight unwinding of the DNA strands at this position.

*Tyr*T shows the usual pattern of enhancement with both TFOs; a single enhanced band at the 3' end of the footprint. As the footprint is one base shorter at the 3' end with the P-TFO compared to the BAU-TFO the enhancement is actually one base higher with this TFO, but still at the 3' end of the footprint. This suggests that the terminal T base in the P-TFO is not bound to the template, but the T in the BAU-TFO is, indicating stronger binding. Sequence 13 is the only other template which shows the same enhancement under all conditions; in this case the footprint stops at the same point at the 3' end with both TFOs, so the enhancement is in the same position on all gels.

The exact template, template 14 and template 6 don't show any enhancement with either of the TFOs. This is expected for the exact template as the pyrimidine strand is labelled. Neither of the TFOs produced a footprint on template 6 so the lack of enhancements is not surprising. However, there is a

footprint with the BAU TFO on template 14, but no enhancement. This is seen on two other gels; template 1 with the BAU TFO and template 7 with the P TFO. In all of these a footprint is visible but no enhancement. On template 1 both TFOs produce a footprint but the enhancement is only seen with the P TFO. On template 14 only the BAU-TFO produces a footprint but no enhancement is visible. Template 7 shows the opposite pattern; both TFOs produce a footprint but the enhancement is only visible with the BAU-TFO. With template 7 there is evidence of possible secondary binding with the P-TFO (discussed later in this section) and this may be the reason for the lack of enhancement; if the TFO is binding at multiple overlapping sites on different templates molecules. Despite no secondary binding being obvious on the other templates (1 and 14), this may also be the cause of the lack of enhancements.

Both TFOs produce multiple enhancements on many of the templates. There are four enhancements visible on templates 4 and 2 with the P TFO, magnesium chloride and ligand, and also on template 9 with BAU. Most of the other templates show two or three enhancements. These multiple enhancements are presumably caused by the TFO binding at multiple sites 1 base apart on the template.

On many of the templates the enhancements are not directly at the end of the footprint as would be expected. On some templates the enhancement can be up to four bases from the end of the footprint. The 3' end of the triplex is where the TFO is most likely to be frayed; the terminal three bases of the TFOs are CBT or CPT only containing one modified base. This is also where the majority of the mismatched between the TFO and the duplex occur. Therefore although the footprint does not appear to extend as far as the enhancement on many of the gels, it may be that these three bases are partially bound to the duplex. This interaction may be enough to shift the enhancements in the 3' direction, but not enough to produce a footprint.

As with the footprints there are also variations between enhancements between the two TFOs on each template. The *tyr*T template and template 13 are the only two where the number and position of the enhancements relative to the footprint stays the same with both TFOs and under all conditions. On templates 4, 2 and 11 the number of enhancements increases in the order B, P, P+10L, P+M+10L. With templates 9, 8 and 3 the reverse pattern is seen. On template 5 there are the same number of enhancements on all gels but with P+M+10L the enhancements are shifted two bases away from the end of the footprint. These

differences may be caused by the variation in binding affinity of the TFOs, and the sequence of the templates, however there does not seem to be any clear pattern.

As well as the secondary binding by the BAU TFO on *tyr*T that is discussed in chapter 3, there are also indications of secondary binding by the BAU and P TFOs on three of the REPSA templates. Templates 2 and 4 both show lighter bands below the enhancement at the 3' end of the primary footprint with the BAU TFO which may indicate secondary binding. On template 7 there is no enhancement with the P substituted TFO, but there are lighter bands extending beyond the edge of the primary footprint. Looking at the sequences of these three templates it is not clear why the TFOs should bind at any of these positions. The lightening of bands may perhaps be due to folding of the template caused by binding of the TFO in the primary site. Secondary binding with the BAU-TFO has been seen previously but was not expected with the P-TFO and it is surprising to see a secondary interaction with this TFO but not the BAU-TFO.

An interesting pattern emerged during sequencing; the templates were all in the same orientation and the purine strand always appeared on the sequencing gel to the 3'-side of the asymmetric Fok I cleavage site. This may be due to unintentional selection by the Fok I enzyme for one strand, as shown on the following page.

Also Fok I cuts much more efficiently as a dimer which may influence the direction of cleavage in some way. There is a possibility that the BAU TFO could be binding by strand displacement rather than forming a triplex. This could distort the DNA in such a way that the Fok I enzyme could only cut if the TFO had bound to one strand rather than the other. This theory is considered unlikely however.

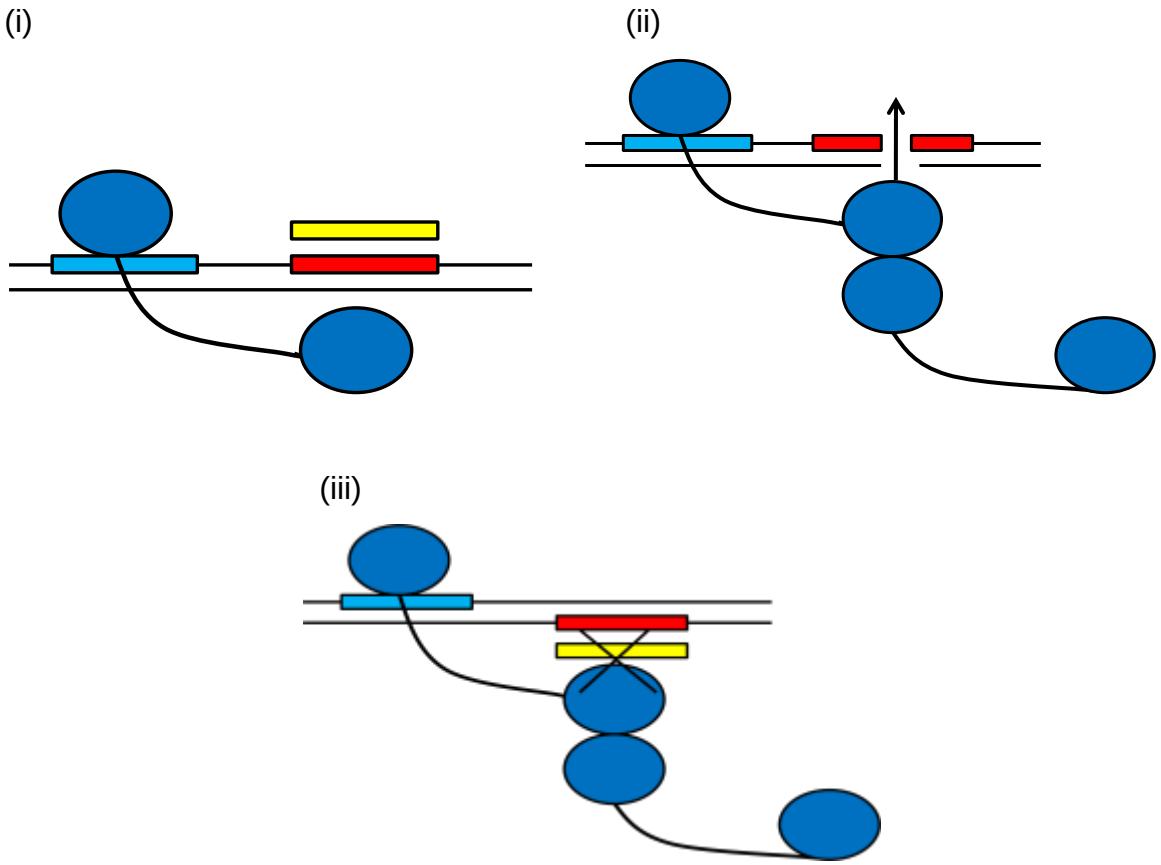


Figure 4.35: Diagram of a possible explanation for why only one strand is seen in the sequenced REPSA products. Straight black lines are the DNA, the blue box is the Fok I binding site, the red box is the TFO binding site and the yellow box is the TFO. The two blue circles are the binding and cleavage domain of the Fok I enzyme.
i and ii: If the TFO binds to the same strand as the Fok I binding site the enzyme can access the other strand and cleave the template.
iii: If the TFO binds to the opposite strand however the structure of the enzyme means that it cannot access the DNA to cut as the TFO is in the way.

Chapter 5: Selectivity of TFOs containing anthraquinone modifications

5.1 Introduction

The previous chapters in this thesis have examined the effect of TFOs containing multiple positively charged nucleotide analogues. In contrast this chapter examines the TFO tethered ligand anthraquinone. Anthraquinone intercalates between DNA bases in duplex DNA and can be tethered to the end of a TFO to increase triplex stability (see Section 1.5.2 for more details). Specificity and stringency of TFOs containing anthraquinone have been examined using DNase I footprinting experiments. Four TFOs with the same sequences have been examined in this chapter; a control TFO with no anthraquinone, TFOs with an anthraquinone molecule at either the 3' or 5' end and a TFO with an anthraquinone tethered to both ends. These TFOs have been footprinted on the *tyr*T template and other templates have been generated by mutating *tyr*T to examine the stringency of the TFOs. Mismatches at the 3' end or in the centre of the triplex were examined. Anthraquinone and other ligands are known to preferentially intercalate at YpR sites; a TFO containing a 3' S-base was therefore also used to allow binding to a YpR step at the 3' end of the triplex (287).

5.2 Experimental Design

The footprinting method used in this investigation is described in Section 2.2.4. Templates and TFOs were incubated overnight at 20°C in sodium acetate pH 5.0 buffer. The TFOs and templates used in footprinting experiments are shown in Figure 5.1 below and the mutagenesis method used to generate the templates from *tyr*T is described in section 2.2.6. The naming of the templates (Figure 5.1ii) reflects the position of the mutation in the *tyr*T sequence and the naming of the TFOs (Figure 5.1i) reflects the position of the anthraquinone or other modifications.

i)

```
T     5′ –  TCCTTCTCTTTTTTCTTT  – 3′
3     5′ –  TCCTTCTCTTTTTTCTTTX – 3′
5     5′ – XTCCTTCTCTTTTTTCTTT  – 3′
3/5   5′ – XTCCTTCTCTTTTTTCTTTX – 3′
S3    5′ –  TCCTTCTCTTTTTTCTTSX – 3′
```

ii)

```
tyrT            5′ – AACCAGTTCTTTTTTCTCTTCCTAACA – 3′
                3′ – TTGGTCAAGAAAAAAGAGAAGGATTGT – 5′
tyrT50T         5′ – AACCAGTTCTTTATTCTCTTCCTAACA – 3′
                3′ – TTGGTCAAGAAATAAGAGAAGGATTGT – 5′
tyrT42A         5′ – AACCATTTCTTTTTTCTCTTCCTAACA – 3′
                3′ – TTGGTAAAGAAAAAAGAGAAGGATTGT – 5′
tyrT42A50T      5′ – AACCATTTCTTTATTCTCTTCCTAACA – 3′
                3′ – TTGGTAAAGAAATAAGAGAAGGATTGT – 5′
tyrT41AT        5′ – AACCTATTCTTTTTTCTCTTCCTAACA – 3′
                3′ – TTGGATAAGAAAAAAGAGAAGGATTGT – 5′
```

Figure 5.1:

i) Sequence of TFOs.

X = Anthraquinone

S = S-base

ii) Sequence of footprinting templates with TFO binding site shown in red.

TFOs T, 5, 3 and 3/5 have the same sequence with different anthraquinone modifications. TFO S3 has an S-base, which binds T in the duplex, in place of the 3' T found in the other TFOs. This allows this TFO to bind to a sequence with a YpR step at the 3' end, the preferred site for anthraquinone intercalation (better than RpY or RpR) (287). The exact binding site for the first four TFOs is TyrT42A, but the exact binding site for the S3 TFO is TyrT41AT. The relationship between the TFOs and each template is reflected in the table below:

	TFOs	
	T, 5, 3, 3/5	S3
tyrT42A	Exact template	3' mismatch
tyrT42A50T	Central mismatch	Central and 3' mismatch
tyrT	3' mismatch	3' mismatch
tyrT50T	Central and 3' mismatch	Central and 3' mismatch
tyrT41AT	3' mismatch	Exact template

Table 5.1: Relationship between each of the TFOs and templates used in this chapter.

The stringency of each of the TFOs when footprinted against templates with different mismatches has been examined.

5.3 Results

Each of the templates has been footprinted in turn with each of the TFOs shown in Figure 5.1 on page 155. The experiments in Figure 5.2 below show the control TFO T (see Figure 5.1) footprinted on all of the templates.

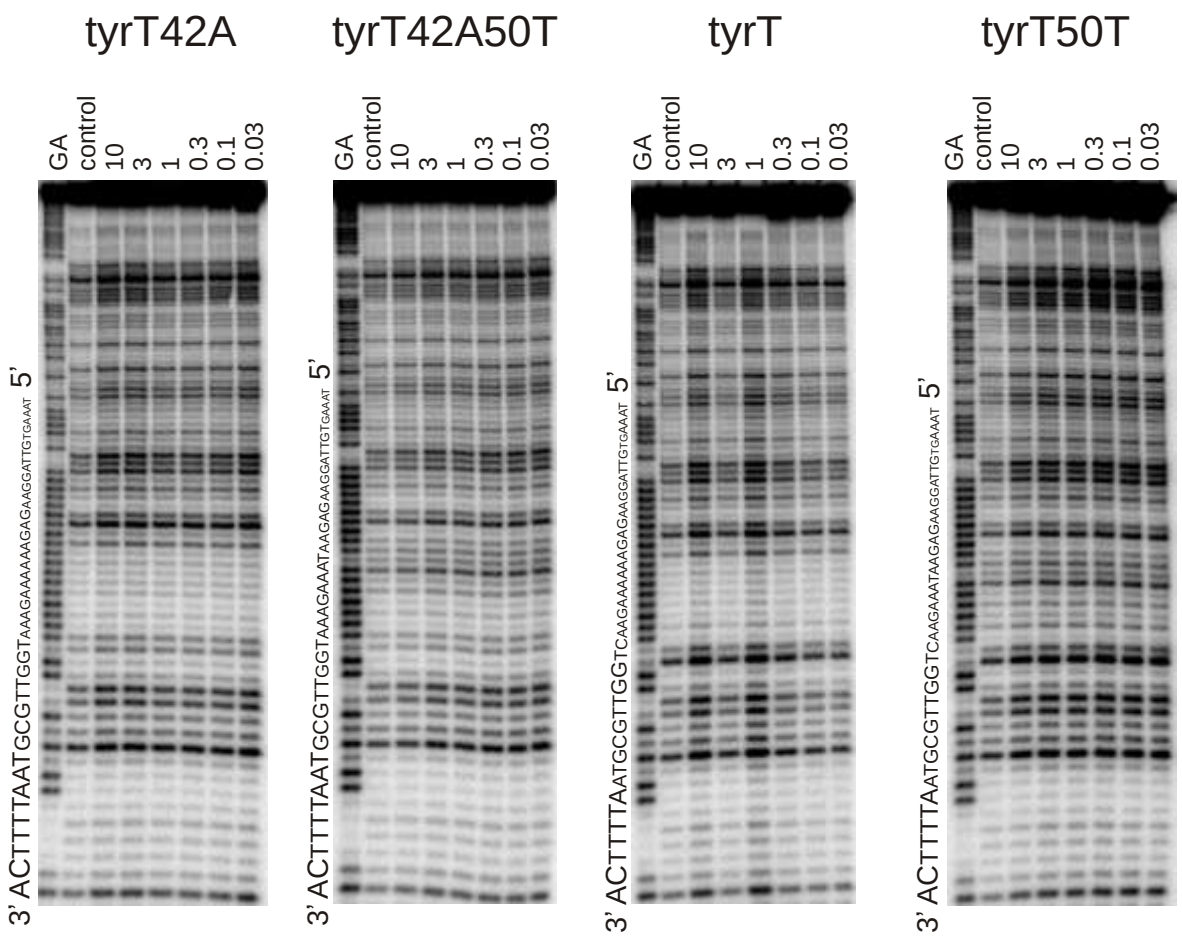


Figure 5.2: DNaseI cleavage patterns of templates TyrT42A, TyrT42A50T, TyrT, TyT50T and TyrT41AT with control TFO T (sequences shown in Figure 5.1). The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO.

There is no evidence of binding on any of these templates, in contrast to previous experiments with similar TFOs (187). As this control was expected to bind to the exact template (tyrT50A) this experiment was repeated multiple times but with the same result. The template was also re-sequenced and the sequencing data is shown in Figure 5.3 on page 159. This confirmed that the sequence of this template was an exact match for the TFO. The TFO was also synthesised multiple times but this also had no effect on binding. There may be something about the sequence of this TFO which prevents binding to the template; for example the formation of secondary structures.

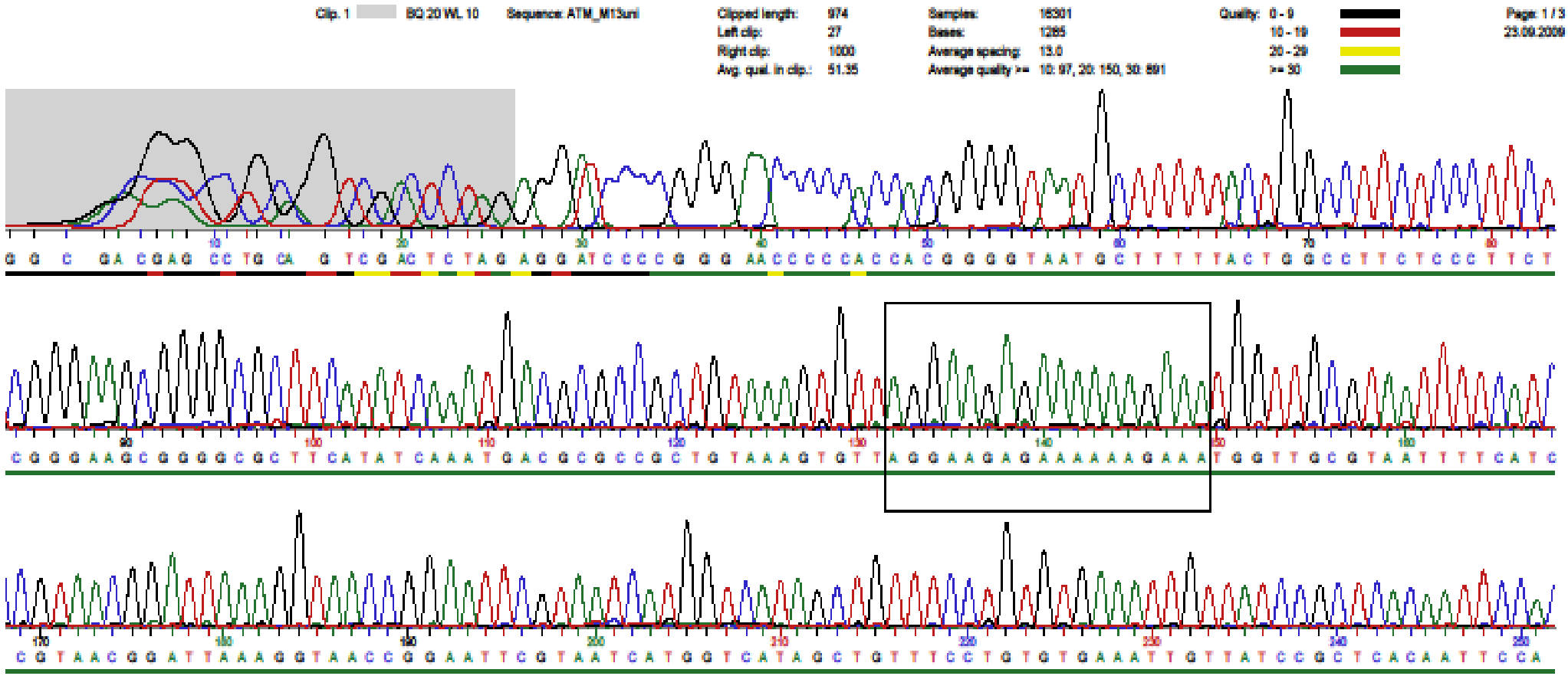


Figure 5.3: Sequence of template TyrT50T.

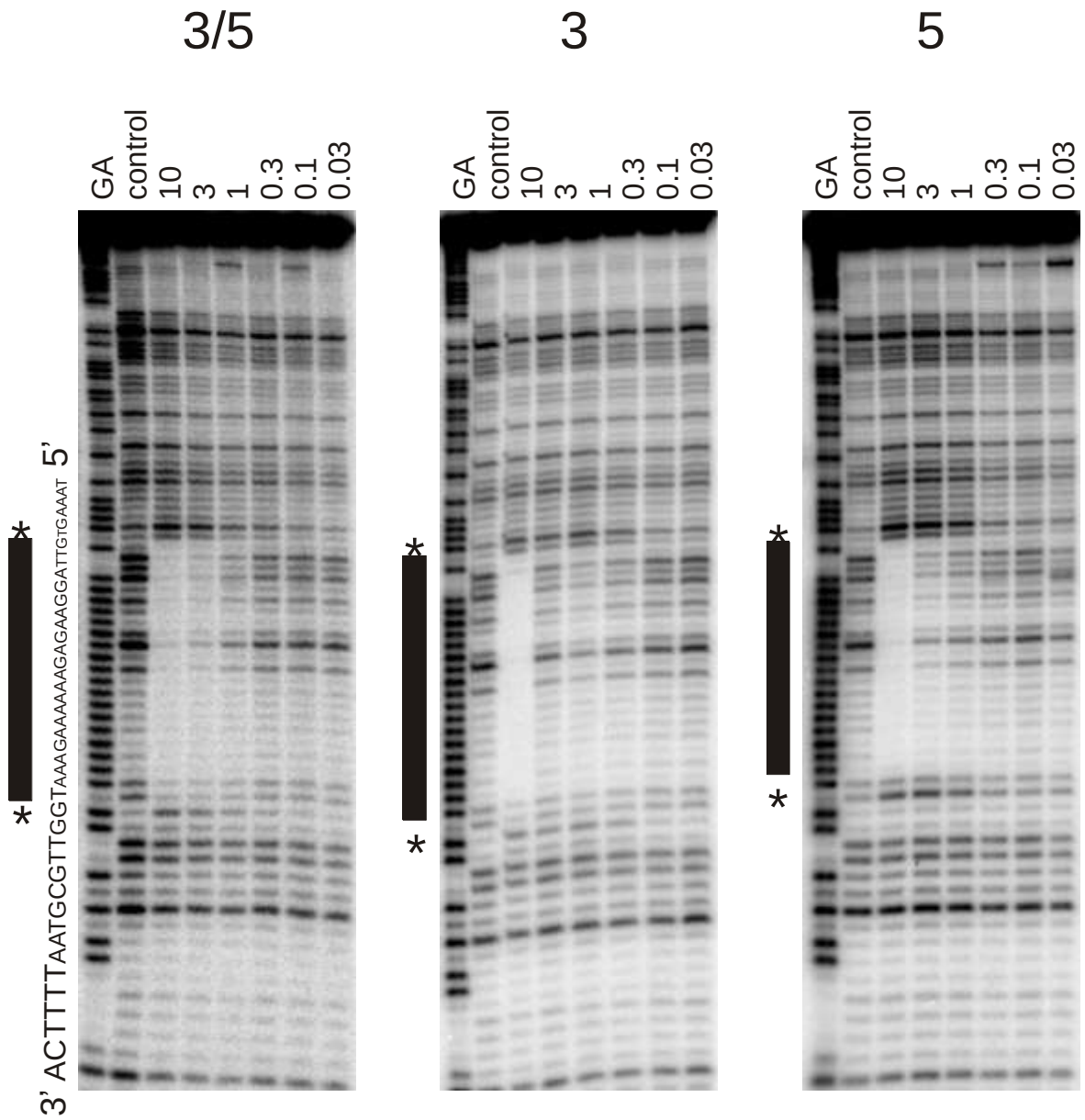


Figure 5.4: DNaseI cleavage patterns of TyrT42A with TFOs 3, 5 and 3/5 (sequences shown in Figure 5.1 page 155). The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO. Rectangles show the position of footprints and asterisks indicate enhancements.

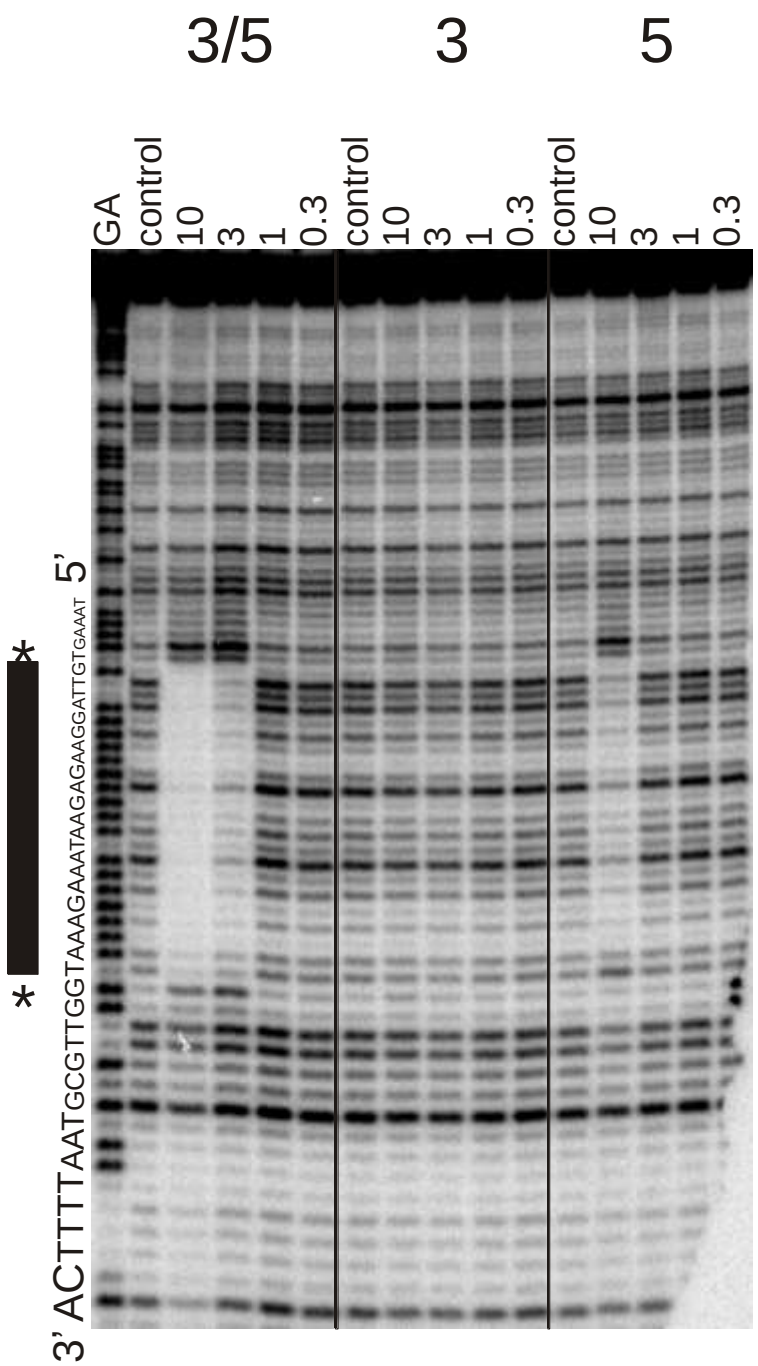


Figure 5.5: DNaseI cleavage patterns of TyrT42A50T with TFOs 3, 5 and 3/5 (sequences shown in Figure 5.1). The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO. Rectangles show the position of footprints and asterisks indicate enhancements.

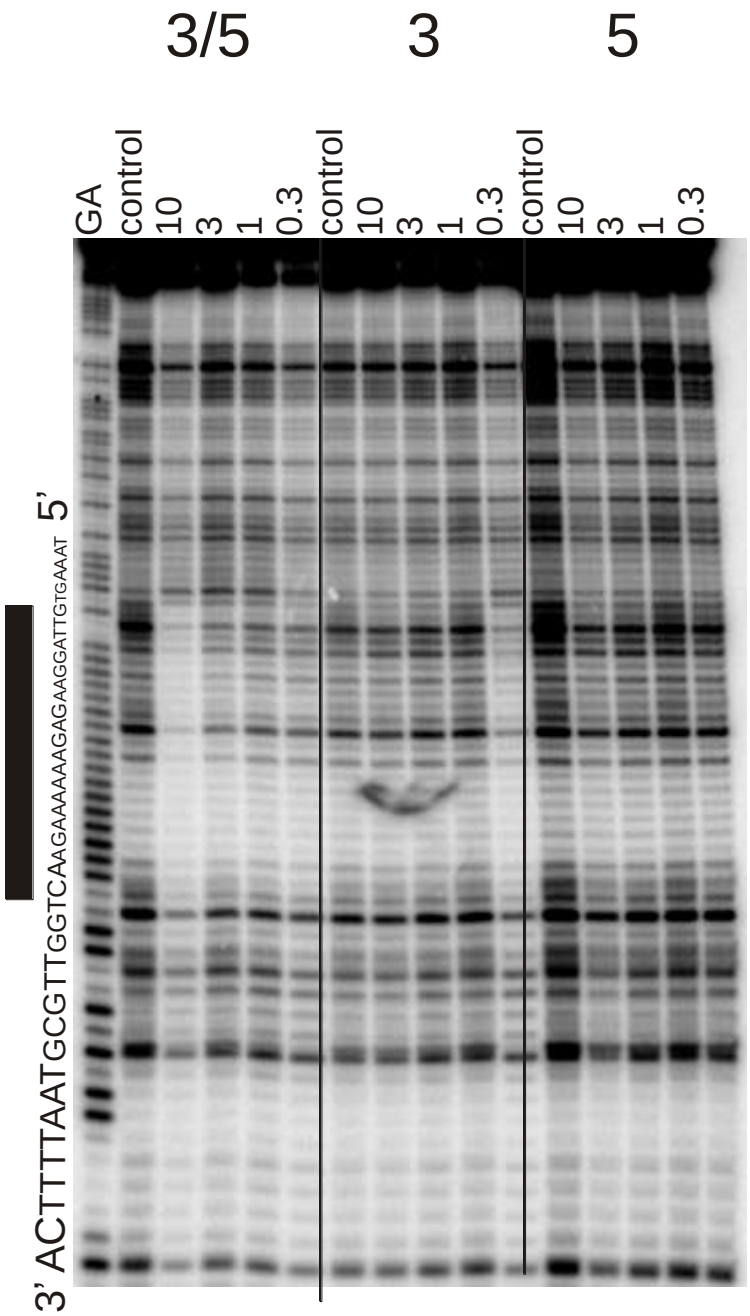


Figure 5.6: DNaseI cleavage patterns of TyrT with TFOs 3, 5 and 3/5 (sequences shown in Figure 5.1). The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO. Rectangles show the position of footprints and asterisks indicate enhancements

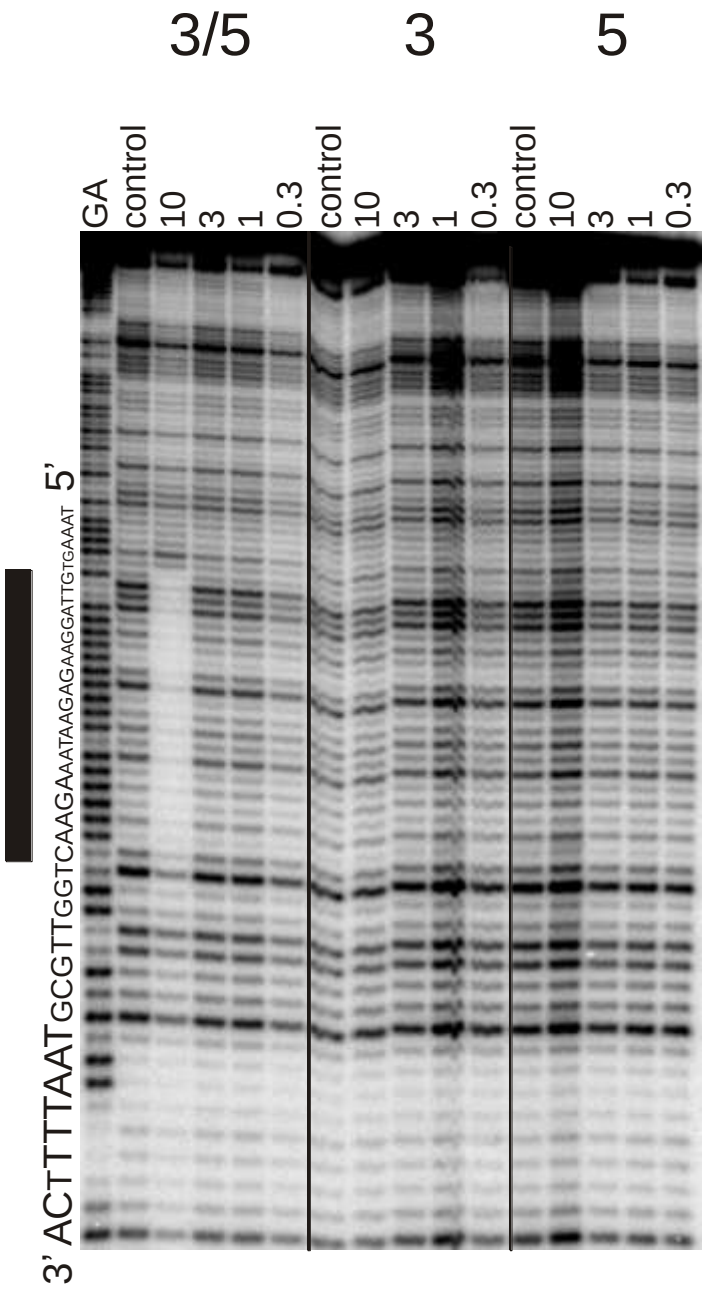


Figure 5.7: DNaseI cleavage patterns of TyrT50T with TFOs 3, 5 and 3/5 (sequences shown in Figure 5.1). The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO. Rectangles show the position of footprints and asterisks indicate enhancements

The experiments in Figure 5.4 on page 160 show footprints of TFOs 3/5, 3 and 5 on template TyrT42A. This is the exact template for these TFOs and clear footprints are visible with all three TFOs. The footprints with TFOs 3 and 5 are similar, both binding strongly at 10 µM TFO. In addition TFO 5 also shows slight attenuating of bands at 3 µM. The double modified TFO (3/5) produces a clear footprint at 3 µM and there is some evidence of binding at 1 µM.

The size of the footprints in this Figure is slightly different for each TFO. This is expected as the addition of an anthraquinone can potentially increase the area of the duplex protected from cleavage by DNaseI (287). The 5' ends of all footprints are similar, with no apparent increase in size caused by the anthraquinone modification on TFOs 5 or 3/5. However with TFOs 3/5 and 5 there is an enhancement at this end of the footprint which is not seen with TFO 3. The 3' ends of the footprints vary between the experiments; with TFO 3/5 or 3 the footprint extends about 2 bases further than the footprint with TFO 5. At the 3' end of the footprint with TFO 5 there is enhanced cleavage which appears one base higher than the enhancement seen with TFOs 3/5 and 3.

The template in Figure 5.5 on page 161 (TyrT42A50T) has a mismatch in the centre of the purine tract, which would be expected to abolish triplex formation with these TFOs. Despite this TFO 3/5 still binds down to 3 µM and TFO 5 produces a slight footprint at 10 µM, though TFO 3 does not show any evidence of binding. There are also enhancements at both ends of the footprints produced by TFOs 3/5 and 5. They are in the same position at the 5' end but there is the same one base difference at the 3' end as seen previously.

The template used for experiments in Figure 5.6 on page 162 (TyrT) has a triplex mismatch at the 3' end for all these oligonucleotides. Only TFO 3/5 produces a footprint on this template, and only at the highest concentration (10 µM). Unlike the previous two sets of experiments there is no enhancement at the 3' end of the footprint, but possibly a slight enhancement at the 5' end.

The template used in Figure 5.7 on the previous page (TyrT50T) combines the two mismatches from Figures 5.5 and 5.6; one at the 3' end and one in the centre of the TFO binding site. As in Figure 5.6 only TFO 3/5 produces a footprint, again only at 10 µM, and the footprint is also fainter than the previous experiment. There is an enhancement at the 5' end of this footprint, but not at the 3' end.

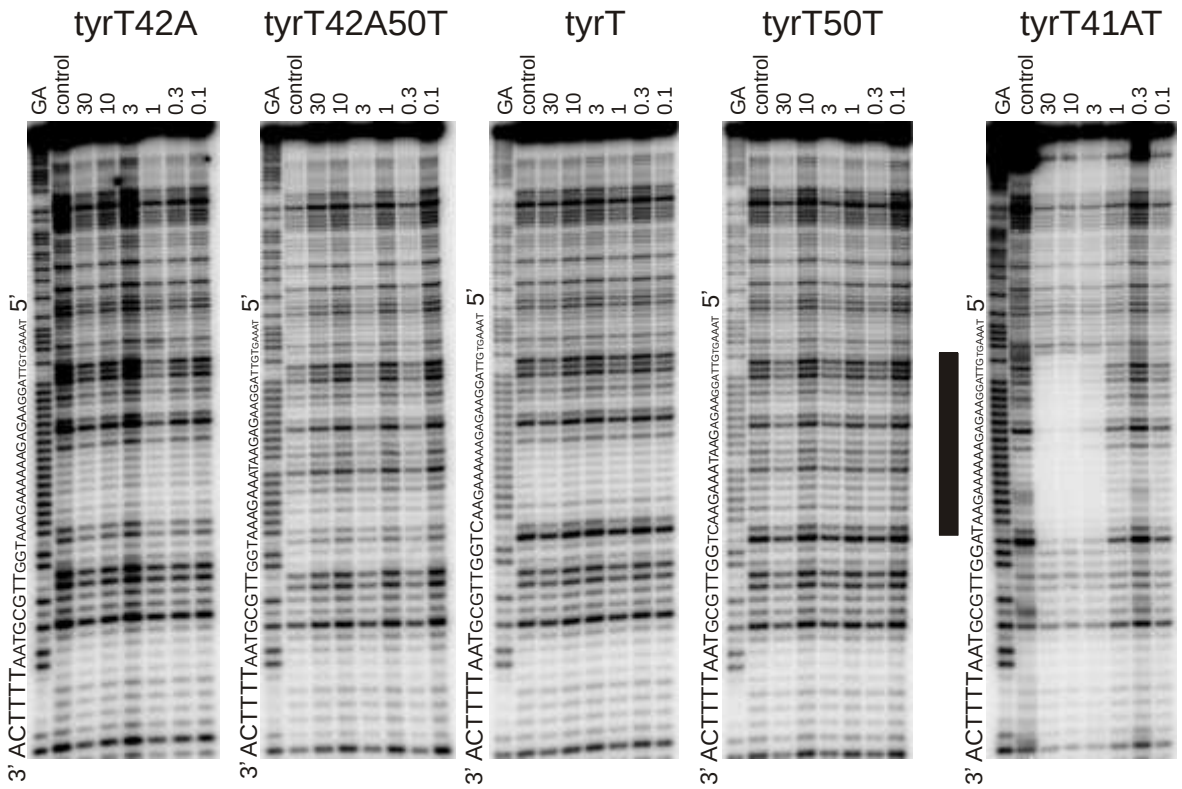


Figure 5.8: DNaseI cleavage patterns of templates TyrT42A, TyrT42A50T, TyrT, TyT50T and TyrT41AT with TFO S3 (sequences shown in Figure 5.1). The experiments were performed in 50 mM sodium acetate pH 5.0. The TFO concentrations (μM) are shown above the gel lanes. Lanes labelled GA are markers specific for purines. Control indicates DNaseI cleavage in the absence of TFO. Rectangles show the position of footprints.

Figure 5.8 on the previous page shows footprinting experiments carried out with TFO S3 which contains an S-base at the 3' end and therefore has a different recognition sequence to the other TFOs used in this chapter. This TFO was designed to investigate the intercalation preference of anthraquinone. It has previously been shown that anthraquinone intercalates preferentially at YpR steps (287). TFO S3 has an S-base at the 3' end followed by a tethered anthraquinone molecule which can then intercalate at a TpG step.

This TFO has been footprinted on all of the templates used in this chapter; templates TyrT42A, TyrT42A50T, TyrT and TyrT50T all contain mismatches in comparison to this TFO, either at the 3' end, in the centre, or both (see Table 5.1 on page 156). TyrT41AT is an addition template containing the exact recognition sequence for this TFO which was generated by site-directed mutagenesis. TFO S3 only produces a footprint on this exact template (down to 3 µM) and does not appear to interact with any of the templates containing mismatches at the concentrations used. Binding of this TFO does not appear to induce enhanced cleavage at either end of the footprint. The size of the footprint is slightly different to the experiments with other TFOs; the 5' end looks the same as previous experiments but the 3' end of the footprint is very indistinct.

5.4 Discussion

The table below summarises the concentration at which footprints were produced by the TFOs on the various templates.

	TFOs				
	T	3/5	3	5	S
tyrT42A	No ftp	1	10	10	No ftp
tyrT42A50T	No ftp	3	No ftp	10	No ftp
tyrT	No ftp	10	No ftp	No ftp	No ftp
tyrT50T	No ftp	10	No ftp	No ftp	No ftp
tyrT41AT	-	-	-	-	3

Table 5.2: Summary of footprinting concentrations of the TFOs on the various templates. Concentration shown in µM.

The five TFOs used in this chapter show different stringencies when footprinted with templates containing mismatches. The control TFO does not

produce a footprint even at 30 μM, despite previous work which indicates that it should (187). This does not appear to be due to the sequence of the template or TFO being incorrect, and may due to the TFO forming a secondary structure that precludes binding to the template. Although this result is unexpected and cannot be explained within this work, it is still possible to compare the footprints with the modified TFOs which clearly do bind to the target.

The doubly modified TFO, 3/5, binds to all four of the templates tested. The strongest binding is seen on the exact template TyrT42A (Figure 5.4 page 160) where the TFO shows evidence of binding down to 1 μM. This TFO binds down to 3 μM on template TyrT42A50T which contains a mismatch in the centre in comparison to the TFO (Figure 5.5 page 161). When footprinted against template TyrT which contains a mismatch at the 3' end, this TFO only binds down to 10 μM TFO (Figure 5.6 page 162). With the final template (TyrT50T Figure 5.7 page 163) which contains both of the previous mismatches, one at the 3' end and one in the centre, this TFO still footprints down to 10 μM, although the footprint is relatively faint. This indicates that for this doubly modified TFO the position of a mismatch has an effect on the strength of binding. Any mismatch reduces the binding affinity, but a mismatch at the centre is less detrimental than a mismatch at the 3' end in this case which is surprising considering previous literature (82). This may be due to the distance from the anthraquinone modification.

TFO 5 which has an anthraquinone molecule at the 5' end binds down to 3 μM on the exact template (TyrT42A, Figure 5.4). It also shows some evidence of binding at 10 μM on the template containing a central mismatch (TyrT42A50T Figure 5.5). There is no evidence of binding on templates TyrT with a mismatch at the 3' end, or TyrT50T with two mismatches (Figures 5.6 and 5.7). This again indicates that mismatches at the end of the binding site have a greater effect than mismatches at the centre.

The TFO with an anthraquinone molecule at the 3' end only binds to the exact binding site in this set of experiments (Figure 5.4). It also does not bind as strongly to the exact template as the 5' modified TFO (10 μM compared to 3 μM with TFO 5). This may be because either the position of the anthraquinone or the intercalation site are affecting the strength of binding to the template. Unlike TFO 5 the TFO with a 3' anthraquinone does not interact with the template containing a 3' mismatch, presumably as the mismatch is at the same position as the anthraquinone and prohibits intercalation.

There are also differences observed in the size of the footprints and the position of enhancements with the TFOs on each of the templates. The footprints all terminate at the same position at the 5' end; however enhanced cleavage is only seen at this end of the footprint with TFOs containing a 5' anthraquinone modification. Modifications at the 3' end of the TFO appear to affect the size of the footprint and the position of the enhancement at this end. The TFOs containing an anthraquinone molecule at this end create footprints around 2 bases longer and push the enhancement 1 base down the gel. With the S base at the 3' end of the TFO the edge of the footprint is indistinct and there is no enhancement visible. In the experiments where there is a mismatch at the 3' end of the footprint no enhancement is seen, this may be because this end of the TFO is fraying away from the duplex because of the mismatch.

Overall it is clear that the doubly modified TFO binds with the highest affinity and is most tolerant of mismatches. Mismatches at the centre of the template have a lesser effect on binding affinity than mismatches at the 3' end. The effect of a 3' mismatch is also greater if the anthraquinone molecule is at this end, this would presumably be the same for a mismatch at the 5'. The presence of an S-base at the 3' end allowing intercalation of the anthraquinone at a YpR step increases the binding affinity on the exact template in comparison to TFO 3 which is the same but does not contain the S-base (3 μM with TFO S3 compared to 10 μM with TFO 3). It does not bind quite as well as the doubly modified TFO however.

Chapter 6: Discussion

Triplex-forming oligonucleotides (TFOs) have been the subject of extensive research in recent years. They have potential applications in many areas such as gene-based therapies, site directed mutation and as biochemical tools (97;98;110;111;155-157;159;162-165). However triplex technology has been hampered by several problems, including low stability due to electrostatic repulsion between strands.

This thesis has investigated two methods for stabilising triplex DNA: the positively charged analogue bis-amino-U, and the triplex stabilising ligand anthraquinone. A TFO containing six consecutive BAU molecules has previously been shown to interact with non-target sites (294). REPSA (Restriction Enzyme Protection, Selection and Amplification) has been used to select for the binding sites of this TFO. These sequences selected by REPSA have then been footprinted with the TFO and two TFOs with the same sequence but containing less heavily modified substitutions. Magnesium chloride and the triplex binding ligand naphthylquinoline have been used to promote binding of the TFOs to the non-specific templates.

The final part of this thesis examines TFOs with anthraquinone molecules tethered at either end to increase binding stability. These have been footprinted against a number of templates containing different triplex mismatches. The emphasis throughout this thesis has been on the stringency of triplex formation when highly modified TFOs are used.

6.1 REPSA

REPSA has been used in this thesis to select for templates which are bound by a 9mer TFO containing six bis-amino-U residues. Two sets of REPSA experiments have been undertaken. The first set was performed using a largely un-optimised protocol and showed no selection for templates bound by the TFO. Optimisation of TFO binding and Fok I cleavage was then carried out and an optimised protocol was generated from the results. Experiments were designed and performed to find buffer conditions optimal for both TFO binding and Fok I cleavage. This presented a problem as both the BAU and cytosine resides are protonated at relatively low pH, but Fok I cleavage is optimal at around pH 8.0. Several different conditions were tested with the exact binding site of the TFO.

The majority of these buffer conditions either only permitted limited binding of the TFO so that some templates were cleaved by Fok I, or were not optimal for Fok I cleavage so that a significant proportion of the templates remained un-cleaved in the absence of the TFO. The standard Fok I buffer was found to give the best results; all templates were cleaved in the absence of TFO within two minutes, but very little of the template was cleaved in the presence of the TFO after 60 minutes. The concentrations of Fok I and the TFO were also optimised for REPSA (see section 2.3 for details).

6.2 Footprinting and pH jumping

Chapter 3 was a preliminary chapter investigating the pH dependence of the BAU containing TFO. These experiments showed that considerably higher TFO concentrations were needed to generate a footprint as the pH was increased. The TFO had a high affinity for the exact template (*tyr*T) at pH 5.0 and 6.0 and showed some evidence of binding at 30 µM at pH 7.0. These gels also showed evidence of the secondary binding evident in previous studies, this was considerably more evident at pH 5.0 however, suggesting that the secondary binding may be more sensitive to pH than the primary binding (294). Different incubation times, either overnight or one hour, were also tested to determine the optimum time for the REPSA experiments. Little difference was observed between the overnight and one hour incubations at pH so a one hour incubation was used in REPSA.

The next step was to carry out pH-jump experiments to determine the effect of the increased pH used in REPSA on binding stability. A range of TFO concentrations were used for these footprinting experiments and the secondary binding on the *tyr*T template was much more pronounced. The TFO and template were incubated at pH 5.0 and the pH was then jumped to around pH 8.0. Samples were taken and digested with DNase I to assess the reaction of the triplex to the change in pH. The TFO molecules bound at secondary sites were seen to dissociate much more rapidly when exposed to high pH, even at high TFO concentrations, suggesting that the secondary binding of this TFO is more sensitive to pH than the primary interaction.

On all of the gels in chapter 3 where high concentrations of TFO have been used the primary footprint is enlarged, possibly caused by slippage of the TFO along the purine-rich region of the template. This enlarged footprint is only seen at

concentrations which also generate secondary interactions so this effect may also be caused by an interaction between the primary and secondary binding sites.

6.3 Footprinting on REPSA templates

14 sequences were chosen for footprinting with TFOs containing BAU, propargylamino-dU or T. Two templates containing the exact recognition sequence of the TFO were also used as controls. Magnesium chloride and the triplex binding ligand naphthylquinoline were added for some of the footprinting experiments in order to promote binding. It was hoped that a pattern of binding affinity might emerge from these experiments to indicate the TFOs preference for differences in template sequence. The BAU-TFO produced clear footprints on all but three of the REPSA templates tested, indicating that the REPSA process was successful in selecting for sequences which were bound by the TFO. On these however the BAU TFO only footprinted down to 0.1 µM even on the exact template, although concentrations as low as 0.003 µM were used. The BAU-TFO failed to produce a footprint on the three remaining templates, although enhancements were seen on two of these.

Significantly higher concentrations of the P-TFO were used, and magnesium chloride and / or the triplex binding ligand naphthylquinoline were needed to promote binding. This TFO does not bind to any template without the ligand, but did produce an enhancement on some templates. The P-TFO bound more than half of the templates on addition of 10 µM ligand, and when 5 mM magnesium chloride was also added only four templates remained unbound. Under these conditions this TFO actually bound down to the same concentration as the BAU TFO on some templates.

The BAU-TFO generated a footprint down to the same concentration on the *tyr*T and exact templates. There was no enhancement on the exact template, but this is not unexpected for triplex footprints with fragments labelled on the pyrimidine-rich strand (284). The P-TFO did not footprint on either of these control templates, although again there was an enhancement on *tyr*T. Surprisingly on addition of ligand the P TFO bound more strongly to E than to *tyr*T, even though both of these templates contained the exact recognition sequence for the TFO. This difference in binding intensity must therefore arise from the difference in flanking sequences between the two templates. As there was no difference in the

binding intensity of BAU on these two templates, this difference may be limited to the P TFO, or perhaps is only seen on addition of the ligand.

Despite the differences in template sequence there does not appear to be a strong pattern in the binding intensities of the TFOs on the different templates. However, some conclusions have been draw from the footprints in Chapter 4. The BAU-substituted TFO did not bind to the majority of the REPSA templates as well as it binds to *tyr*T or the exact template. Template 4 was the only template that showed a footprint and enhancement down to the same concentration as these two controls. This template also contained the longest run of bases contiguous with the exact binding site of the TFO, only the final 3' base was a mismatch. Templates 1, 7 and 10 also matched 8 out of the 9 exact template bases, but the mismatch was opposite the C in the TFO rather than at the 3' end T. Complementary binding of the C base therefore appears to be more important to triplex stability than binding of the terminal T.

None of the templates, including the controls, bound the propargylamino-dU TFO alone, although several gels showed enhanced cleavage in similar regions to the BAU-TFO. When 10 μM ligand was added 10 out of the 16 templates were bound by this TFO and two of the others showed some enhancements. The exact template was bound at the lowest concentration of all, a factor of 10 tighter than the next best REPSA template. The *tyr*T template however was only bound as strongly as several of the other templates. The addition of magnesium chloride as well as ligand did not improve TFO binding to either of the control templates, however many of the other templates were bound at lower concentrations under these conditions. Several templates then appeared to bind the TFO more strongly than *tyr*T, even though this was an exact match, and templates 1 and 5 were bound at the same concentration as the E template under these conditions.

Surprisingly it appears from these data that the BAU TFO actually discriminates better against non-exact binding sites. There was a large difference in BAU-TFO concentrations required to bind the exact sequence compared to the REPSA templates, whereas with the P-TFO and ligand the concentrations required to generate footprints were all similar and many templates were bound with higher or the same affinity as the exact templates. The sequence of the three nucleotides at the 3' end of the binding site appeared to have relatively little effect on binding affinity; so long as there was a run of As the TFOs bound at similar concentrations. The run of As did have to be at least 5 long however; the templates containing only three or four consecutive As (12 and 6) did not bind the

P-TFO, although with four As the BAU TFO still bound relatively well (template 12).

Out of all of the REPSA sequences, template 4 was bound with the highest affinity by the BAU-TFO. The only mismatch with this template was against the T at the 3' end of the TFO. Templates 1, 7 and 10 also only contained one mismatch but this was against the C base in the BAU-TFO and this appears to be more destabilising. There were six templates which had two mismatches out of the three 3' terminal bases 5, 2, 13, 11, 14 and 9. The sequence of the 3' end of the TFO was CBT, and should have bound the sequence GAA. Templates 5, 2 and 13 had a GXX pattern (X being a mismatch in relation to the TFO), 11 and 9 XXA and 14 XAX. There did not seem to be any correlation between the type of mismatch and the binding intensity from these footprints. The templates with three mismatches at the 3' end were bound the least strongly (templates 8 and 3).

On all of the REPSA templates the TFO appeared to favour binding at the 3' end of the A tract, leaving any superfluous As at the 5' end. This is possibly due to the TFO sequence at the 3' end being different (CXT with X being B, P or T depending on the TFO). It may be that the C at this end of the TFO discriminates against binding to an A in the duplex in the context of this TFO.

Apart from differences in binding intensity there were several other interesting differences between the BAU and P gels, as well as anomalies such as multiple enhancements and secondary binding sites. These are discussed in detail in Chapter 4 but did not a show any pattern based on template sequence.

Not all of the templates which emerged from the REPSA selection process were footprinted with the TFOs due to a lack of time and materials. Table 4.1 ii (page 107) shows the sequences which were not footprinted; none of these contain a run of As. It may be that these templates were bound by the TFO at the concentration used in REPSA, or this may be a limitation of the REPSA protocol that not all unbound templates are excluded from selection. An acknowledged limitation of REPSA is that there are factors which can inhibit Fok I cleavage other than TFO binding, for example formation of secondary structures by the template DNA (77). It is interesting that REPSA did not pull out the exact binding site of the TFO, although there are several templates which contain sequences only one base different from the exact template. It is possible that REPSA did pull out this template but that it was not sequenced.

Previous work on BAU-containing TFOs has shown that when single substitutions are made, the TFO retains specificity (187;243). However this is the

first work examining the stringency of TFOs containing multiple, clustered BAU residues. It has been shown that BAU demonstrates more stringent binding than unmodified TFOs (187). This has been confirmed by the experiments presented here; almost all of the templates examined contain a run of six As. Templates which contain fewer than 6 As are not bound as strongly by the TFO, or are not bound at all. Also all of the footprints on the REPSA templates are smaller than the footprint produced by this TFO on the exact templates. This indicates that some or all of the run of six BAU residues are binding to the run of As, and the three other bases (CBT) at the 3' end of the TFO are fraying away from the template. So the BAU residues are not in themselves less stringent; they simply bind so strongly to the template that the specificity of the remaining TFO no longer dominates the binding interaction.

Binding of TFOs comprised of solely natural nucleotides to secondary sites has been shown at high concentrations (82). Binding of TFOs to these sites involved one or two mismatches, fraying of the third strand or loop formation (82). TFOs containing BAU substitutions however were shown to generate extremely stable triplexes and individual substitutions were highly selective for AT (187). BAU has enhanced discrimination against YR bases in the duplex compared with T. BAU also shows high selectivity for AT over GC bases pairs in the duplex, although BAU.GC triplets are more stable than T.GC or BAU.YR triplets (187). TFOs with the same sequence as the one used in this thesis but containing 3 or 4 BAU residues showed no evidence of secondary binding on *tyr*T (186). It is therefore likely that BAU reduces the stringency of TFOs only when there are large numbers of consecutive substitutions. The mostly likely model for the interaction of the TFOs used in this thesis on secondary sites, is binding of the run of modified nucleotides to the run of As and fraying of the 3' end CXT sequence.

6.4 Selectivity of anthraquinone TFOs

Chapter 5 examines footprinting experiments using TFOs containing anthraquinone modifications. Anthraquinone intercalates between DNA bases in duplex DNA and can be tethered to the end of a TFO to increase stability. The specificity of five TFOs with different anthraquinone modifications was examined in this chapter by footprinting against various fragments. Mismatches at the 3' end or in the centre of the triplex were examined. Anthraquinone is known to

preferentially intercalate at YpR sites; a TFO containing a 3' S-base was therefore also used to allow binding to a YpR step at the 3' end of the triplex.

The doubly modified TFO with an anthraquinone modification at either end bound to all four of the templates tested, even with two triplex mismatches. TFO 5 which has an anthraquinone molecule at the 5' end bound down to 3 μM on the exact template and showed some evidence of binding at 10 μM on the template containing a central mismatch. However there was no evidence of binding on the templates with a mismatch at the 3' end or containing two mismatches. This indicates that mismatches at the end of the binding site have a greater effect than mismatches at the centre. The TFO with an anthraquinone molecule at the 3' end only bound to the exact template and did not bind as strongly to this as the 5' modified TFO. This may be because either the position of the anthraquinone or the intercalation site affected the strength of binding to the template. Unlike TFO 5 the TFO with a 3' anthraquinone did not interact with the template containing a 3' mismatch, presumably as the mismatch is at the same position as the anthraquinone and prohibits intercalation. Modifications at the 3' end of the TFO appeared to affect the size of the footprint and the position of the enhancement at this end, and in experiments with a mismatch at the 3' end of the footprint no enhancement was seen, this may be because this end of the TFO was fraying away from the duplex.

Overall it was clear from these experiments that the doubly modified TFO bound with the highest affinity and was most tolerant of mismatches. Mismatches at the centre of the template had a lesser effect on binding affinity than mismatches at the 3' end. The effect of a 3' mismatch was also greater if the anthraquinone was at this end. The presence of an S-base at the 3' end allowing intercalation of the anthraquinone at a YpR step increased the binding affinity on the exact template in comparison to TFO 3 which did not contain the S-base. The TFO containing the S base did not bind quite as well as the doubly modified TFO however.

6.5 Implications

In general although the BAU modification undoubtedly provides a dramatic increase in affinity for the target sequence, it needs to be incorporated carefully to ensure that the TFO does not lose specificity. As TFOs containing single or few BAU substitutions show high fidelity for the target sequence, substituting every

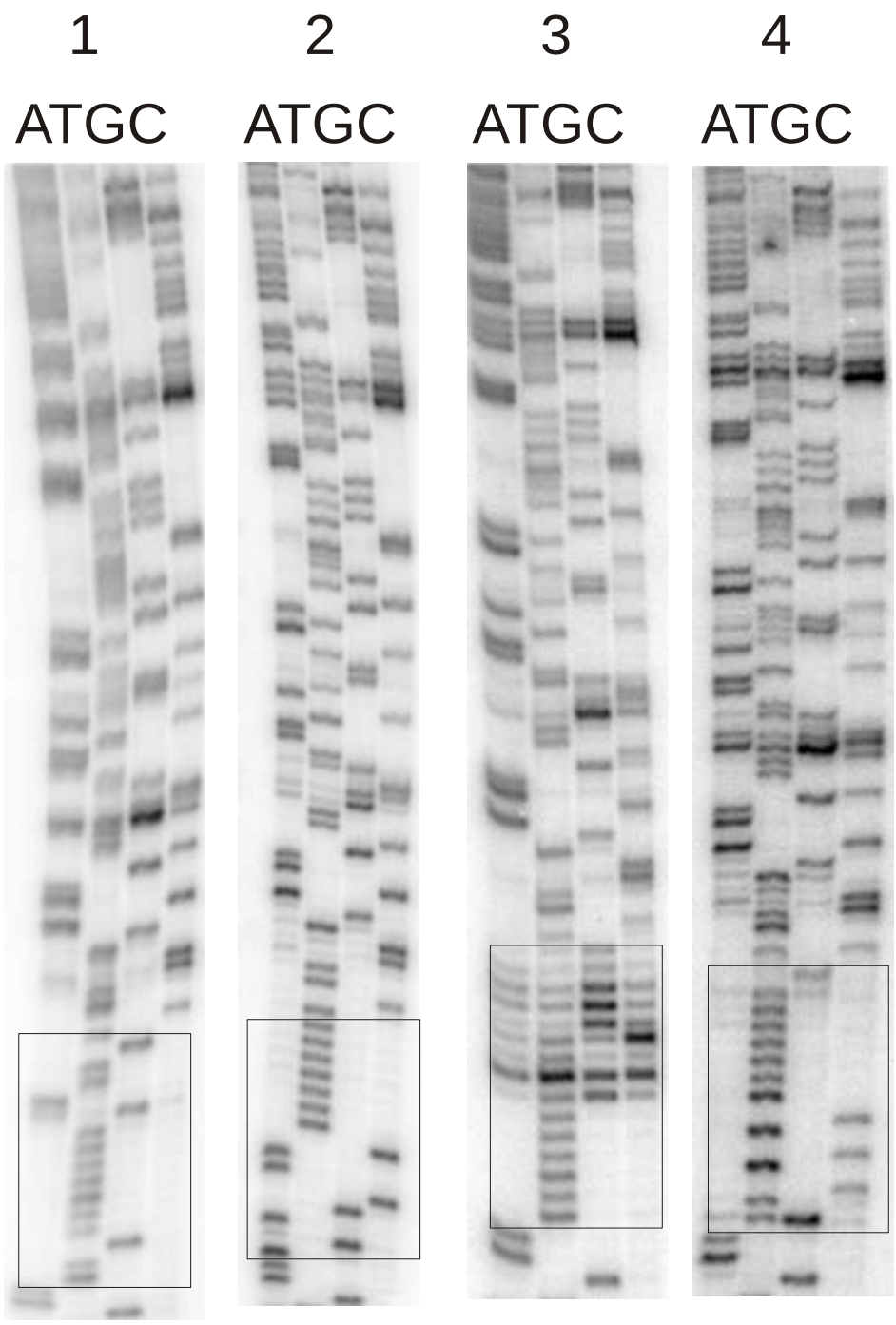
other T with a BAU molecule may be a better way to prevent non-specific binding (187). Using a less heavily modified nucleotide such as the propargylamino-dU examined here does not necessarily alleviate this problem. In the experiments in this thesis the P-TFO in conjunction with the triplex-binding ligand naphthylquinoline actually shows less discrimination between exact and secondary binding sites (see Table 4.2 page 145). This may be partially because of the contribution of the ligand however as it was not possible to quantitate how much the P substitution contributed to specificity with just the TFO alone.

The results from experiments with anthraquinone modified TFOs have two implications for their future incorporation. TFOs containing this modification at both ends are more tolerant of mismatches than TFOs with only one anthraquinone molecule. Also it appears that the closer the mismatch is to the anthraquinone the less stable the triplex becomes. This may be of use if the target for the TFO is one of two similar templates which have only slight differences in sequence for example.

In general the use of positively charged nucleotide analogues and attachment of intercalating moieties to TFOs continue to have great potential in the field of triplex technology, but will have to be used with care to ensure there is no loss of stringency. REPSA also appears to be a very useful tool for examining TFO stringency.

Appendices

Appendix 1: Sequencing Gels



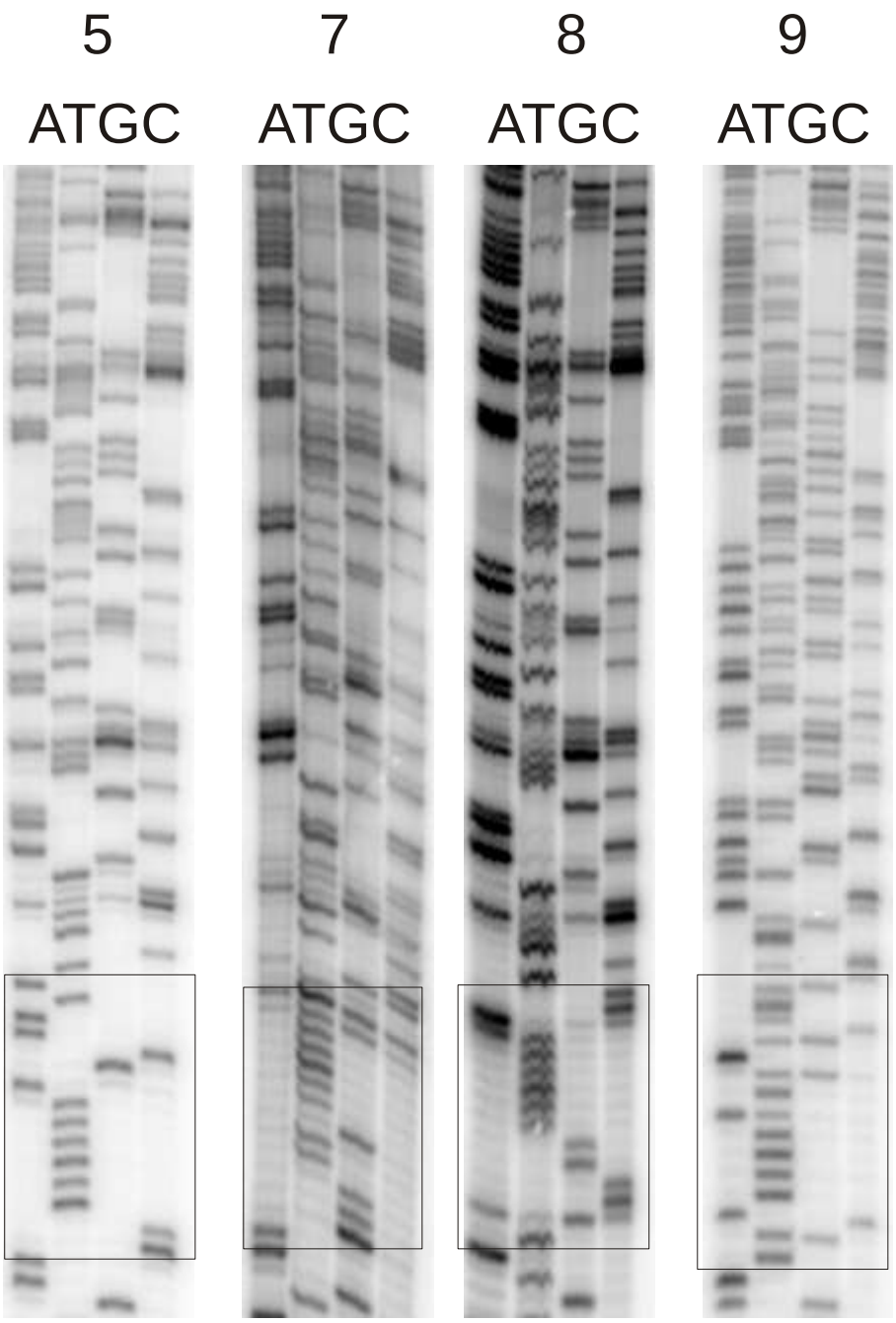
Sequencing gels of templates 1-4.

1: TTGTTTTTTTGATTG

2: AGAGCAACTTTTTTT

3: TTTTTT?????????

4: GTCTCTCTTTTTTTG



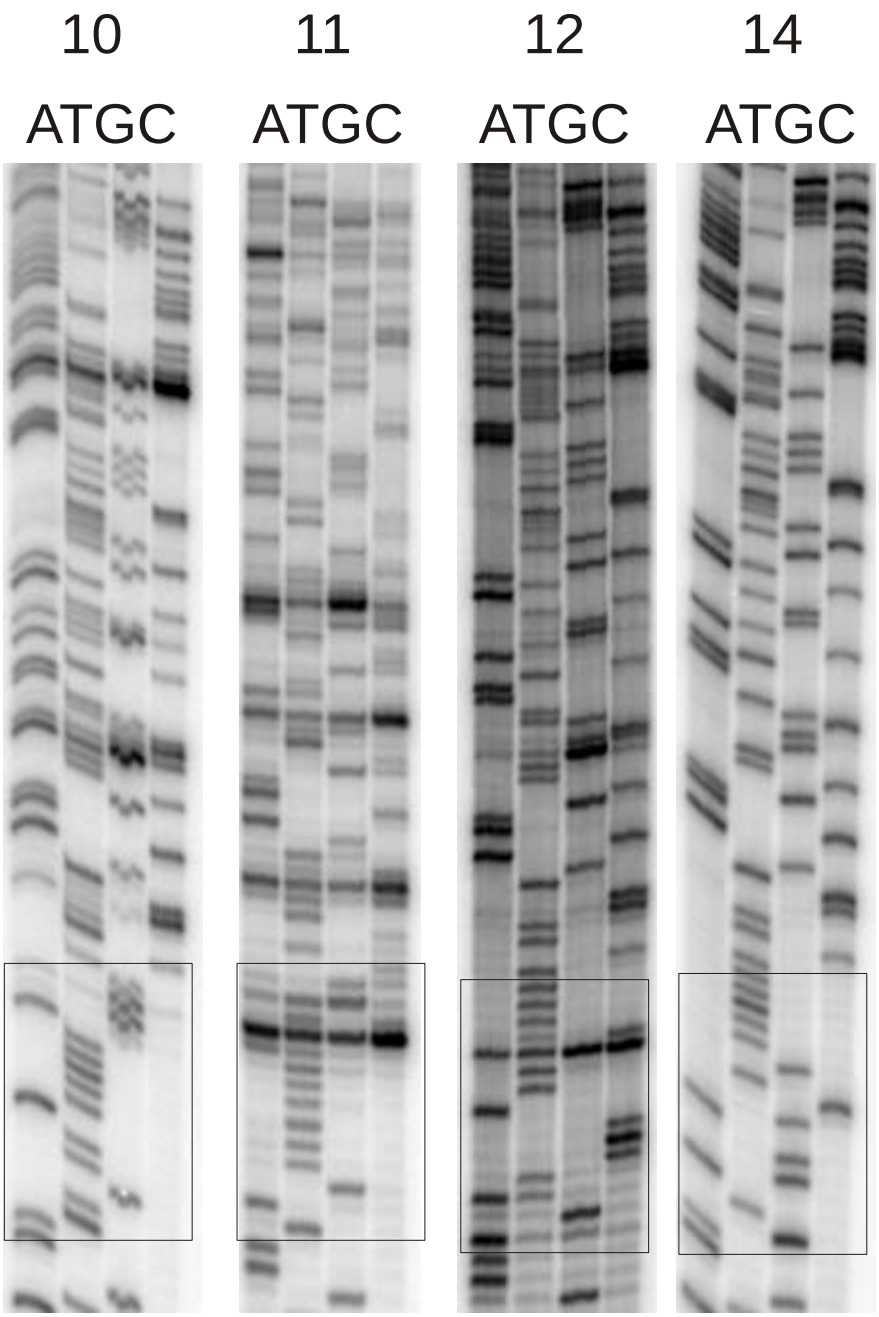
Sequencing gels of templates 5, 7-9. Template 6 would not sequence.

5: CCTTTTTTAGCAATA

7: AGGGTTGTTTTT???

8: T?CCGGTTTTT?ACC

9: T??TTTT?T?A?CTTTG



Sequencing gels of templates 10-12, 14.

10: TTGTTATTTTTAGGG

11: TAGTTTTTT????GC

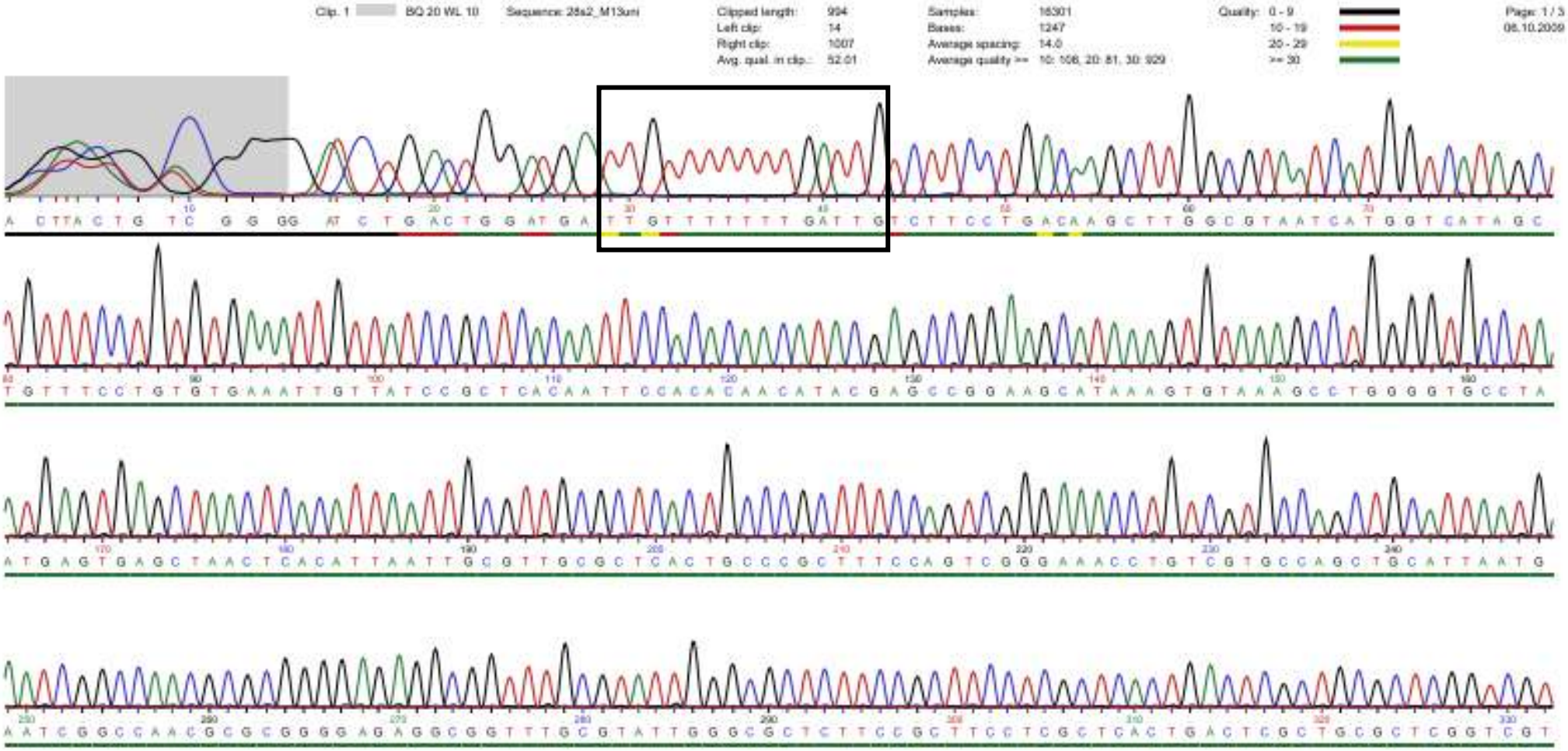
12: AGATCCCATT?TTTT

14: GTAGGAGCTGTTTTT

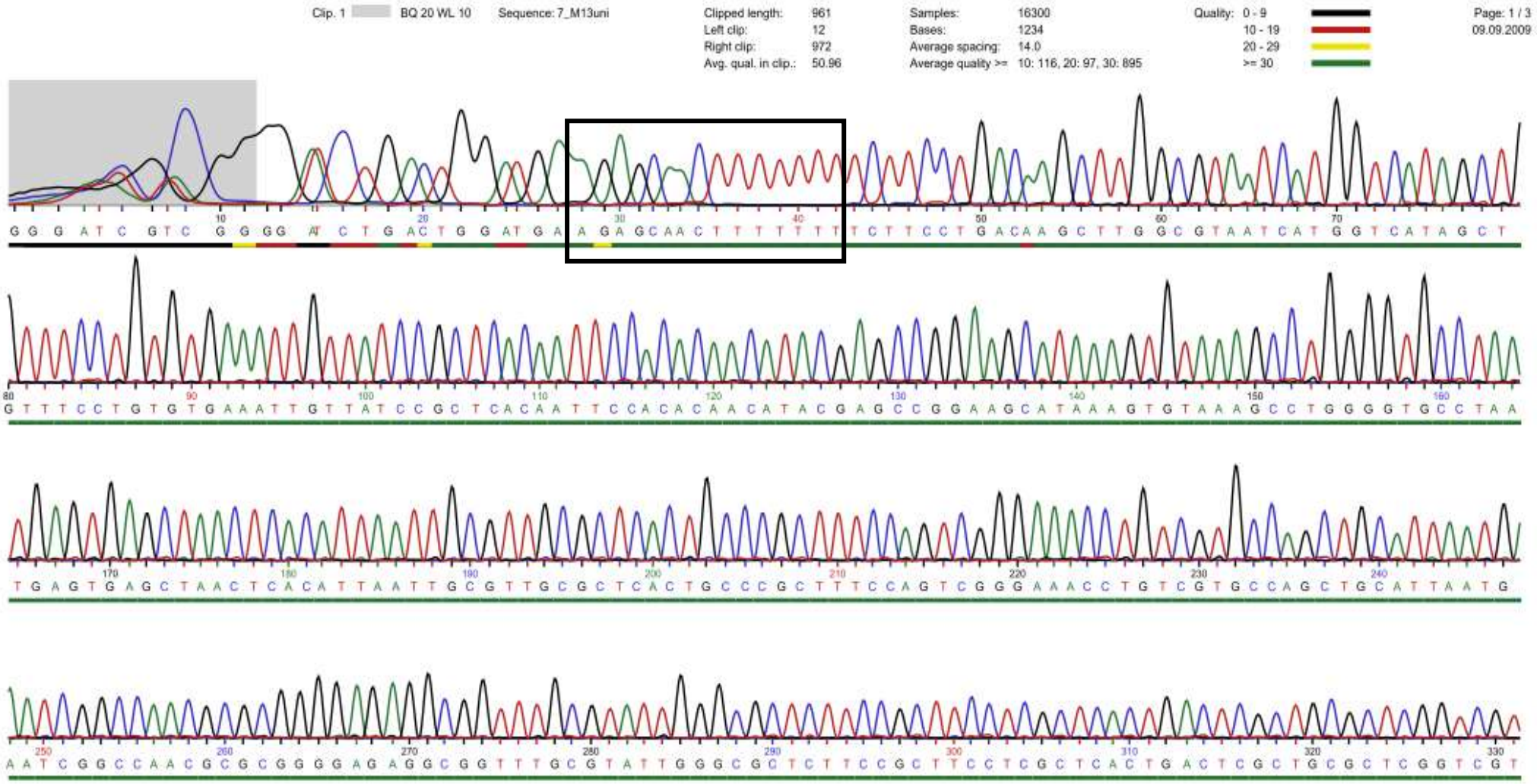
As many of the sequencing gels were unclear but indicated a run of T's plasmid samples were sent to MWG for sequencing. The traces from the sequencing are shown in Appendix 2 below.

Appendix 2: Sequencing Traces

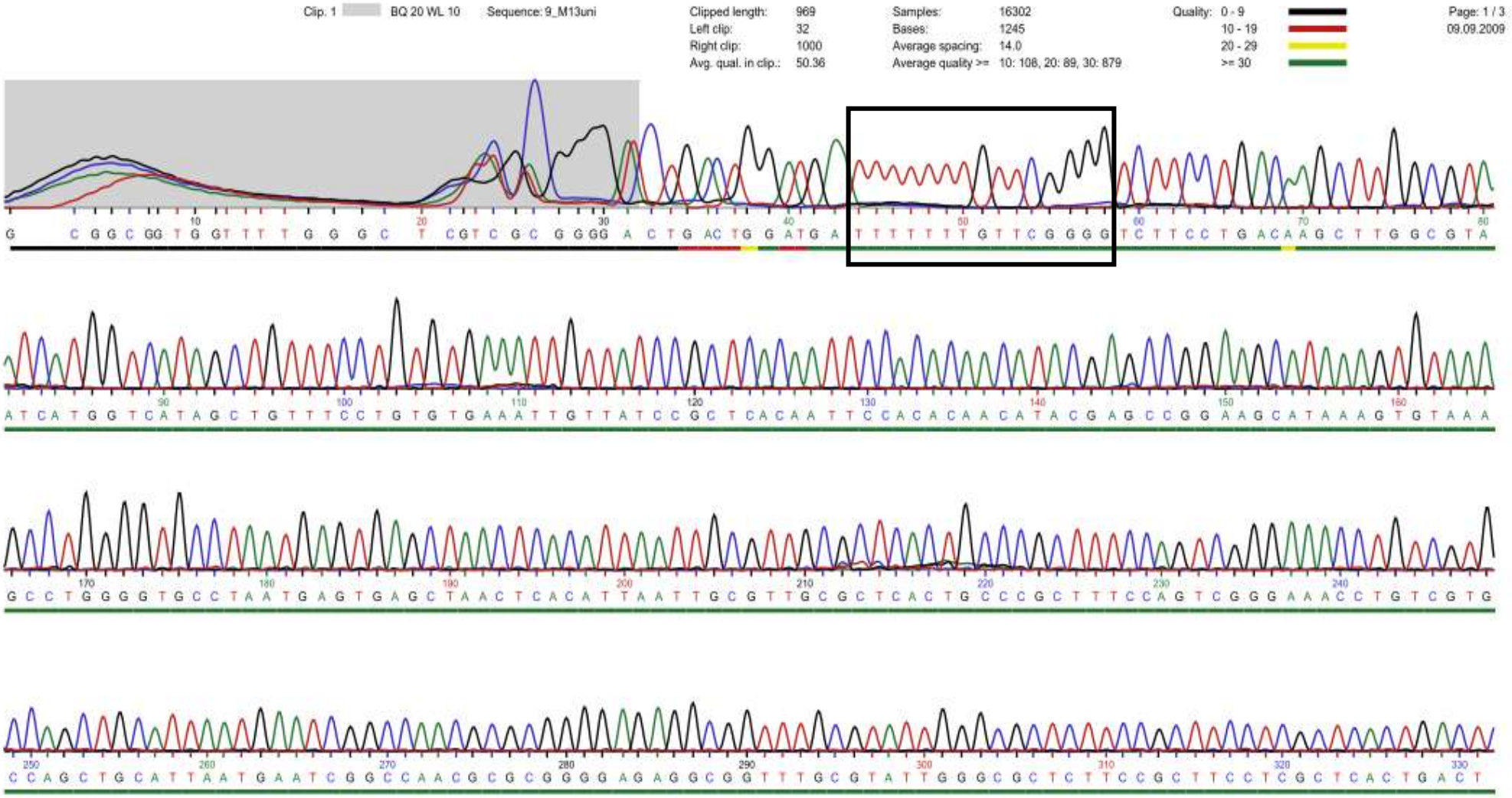
Sequence 1:



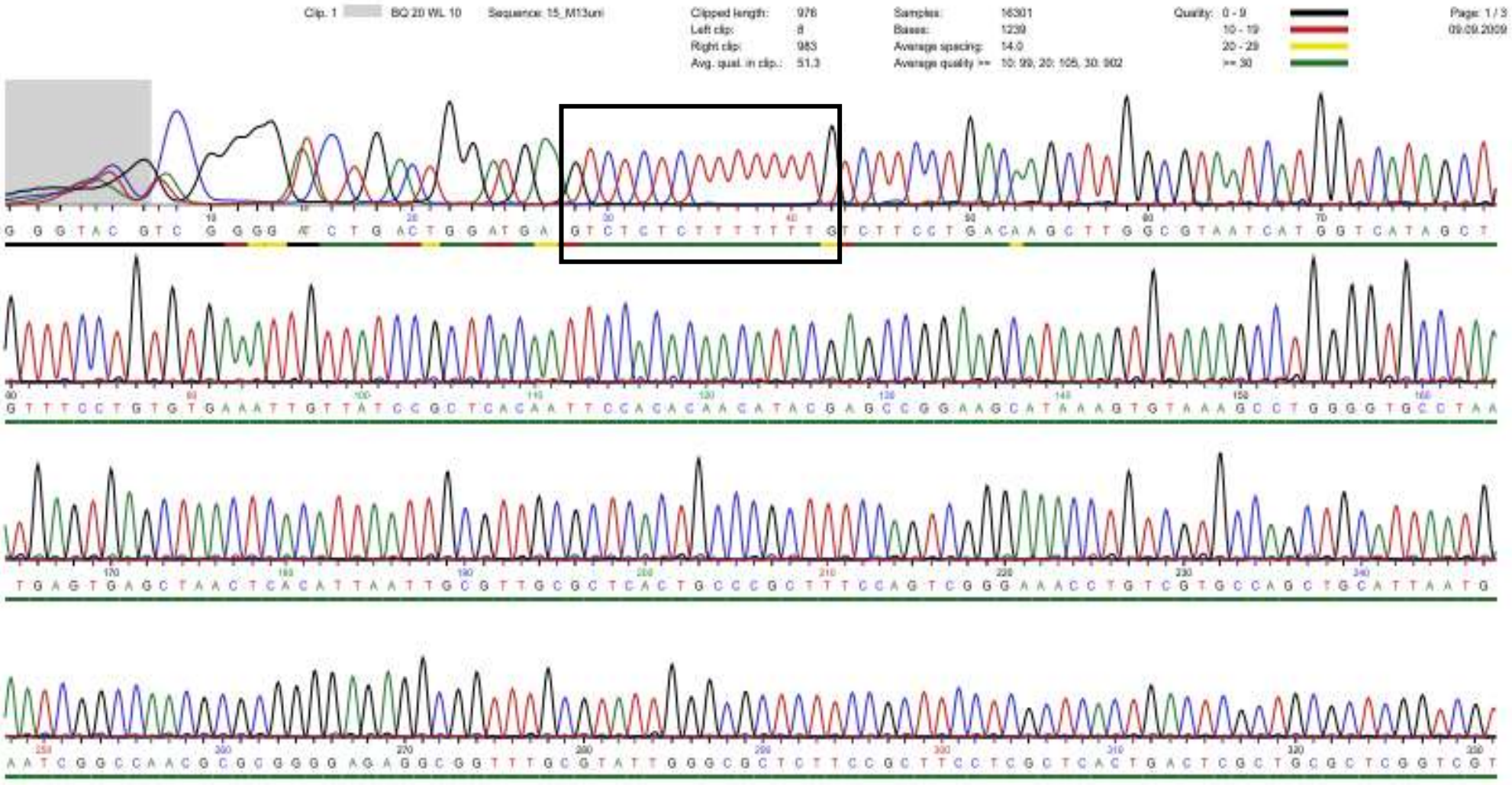
Sequence 2:


Clip. 1
BQ 20 WL 10
Sequence: 7_M13uni
Clipped length: 961
Left clip: 12
Right clip: 972
Avg. qual. in clip.: 50.96
Samples: 16300
Bases: 1234
Average spacing: 14.0
Average quality >= 10: 116, 20: 97, 30: 895
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009

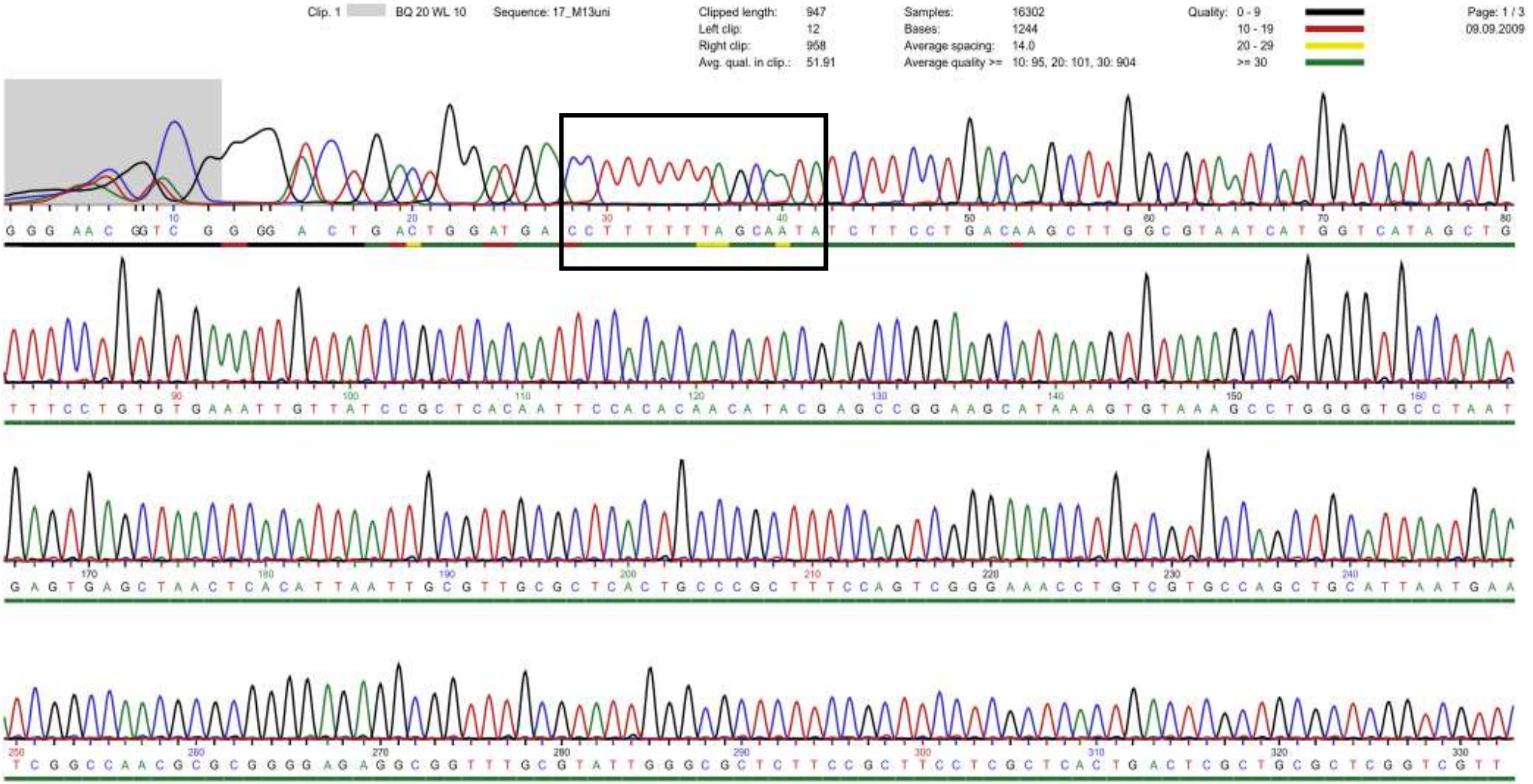
Sequence 3:


Clip. 1 BQ 20 WL 10 Sequence: 9_M13uni
Clipped length: 969
Left clip: 32
Right clip: 1000
Avg. qual. in clip.: 50.36
Samples: 16302
Bases: 1245
Average spacing: 14.0
Average quality >= 10: 108, 20: 89, 30: 879
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009

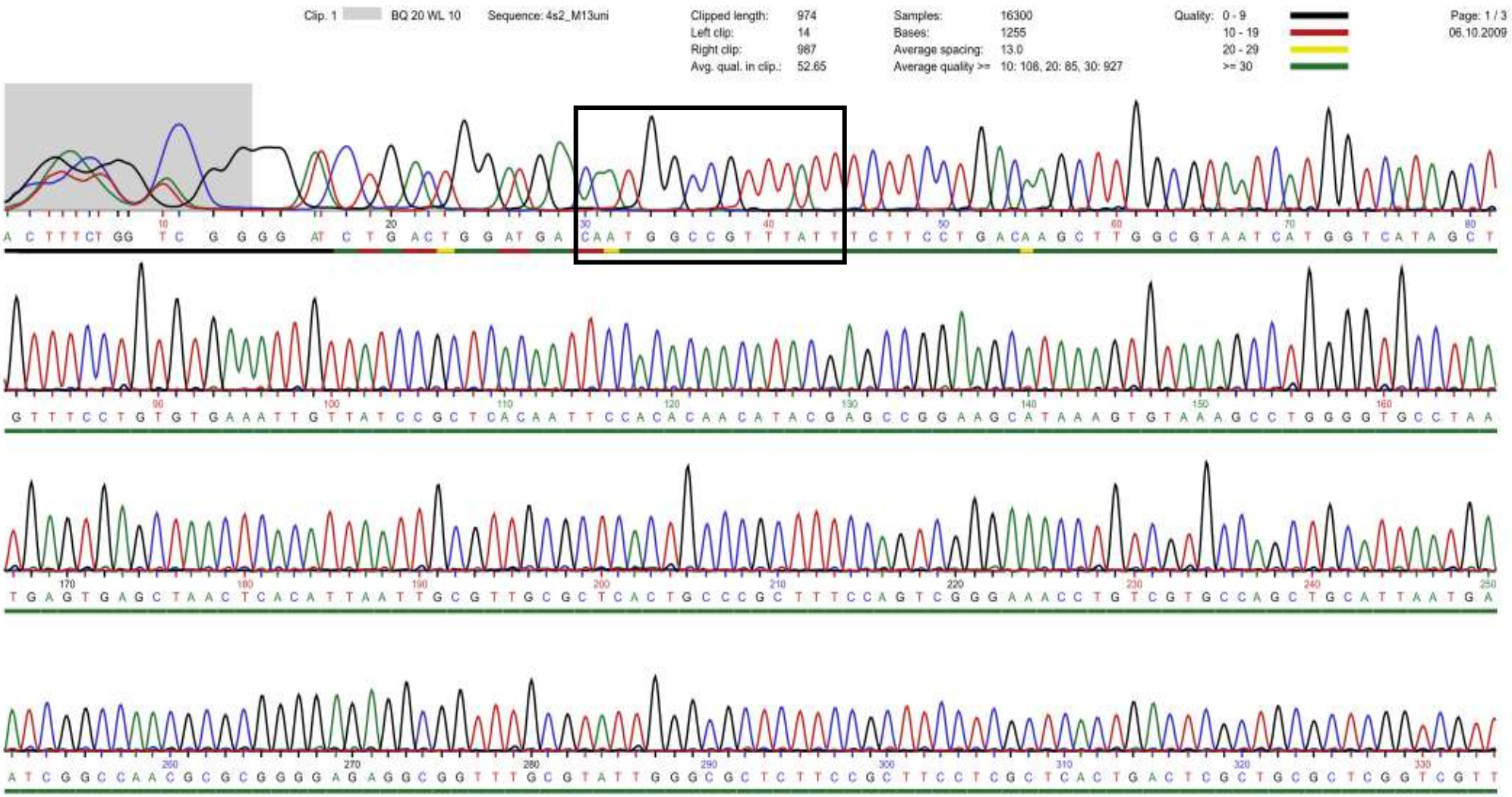
Sequence 4:


Clip. 1
BQ 20 WL 10
Sequence: 15_M13uni
Clipped length: 976
Left clip: 8
Right clip: 983
Avg. qual. in clip.: 51.3
Samples: 16301
Bases: 1239
Average spacing: 14.0
Average quality >=: 10: 99; 20: 105; 30: 902
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009

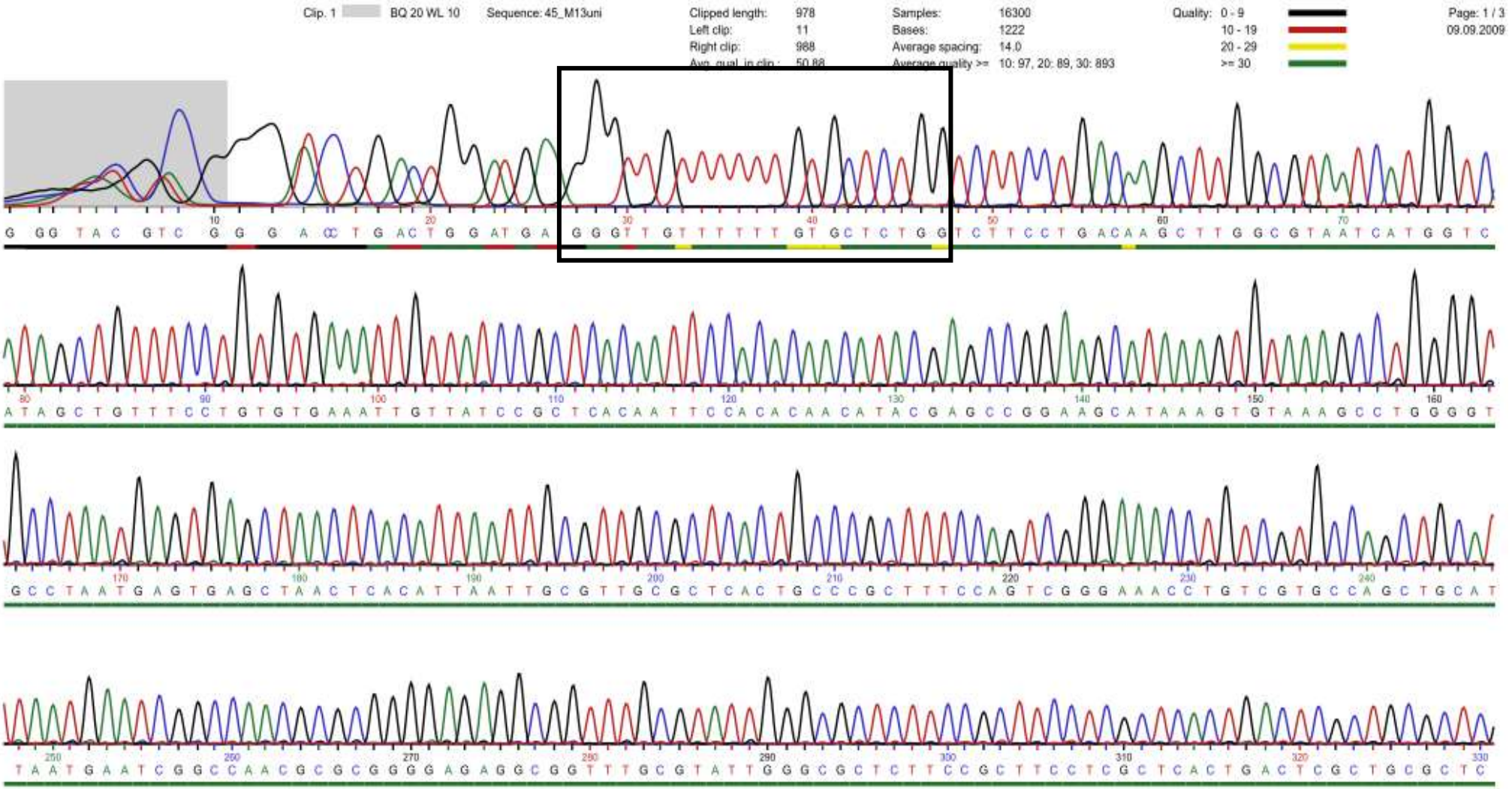
Sequence 5:


Clip. 1
BQ 20 WL 10
Sequence: 17_M13uni
Clipped length: 947
Left clip: 12
Right clip: 958
Avg. qual. in clip.: 51.91
Samples: 16302
Bases: 1244
Average spacing: 14.0
Average quality >= 10: 95, 20: 101, 30: 904
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009

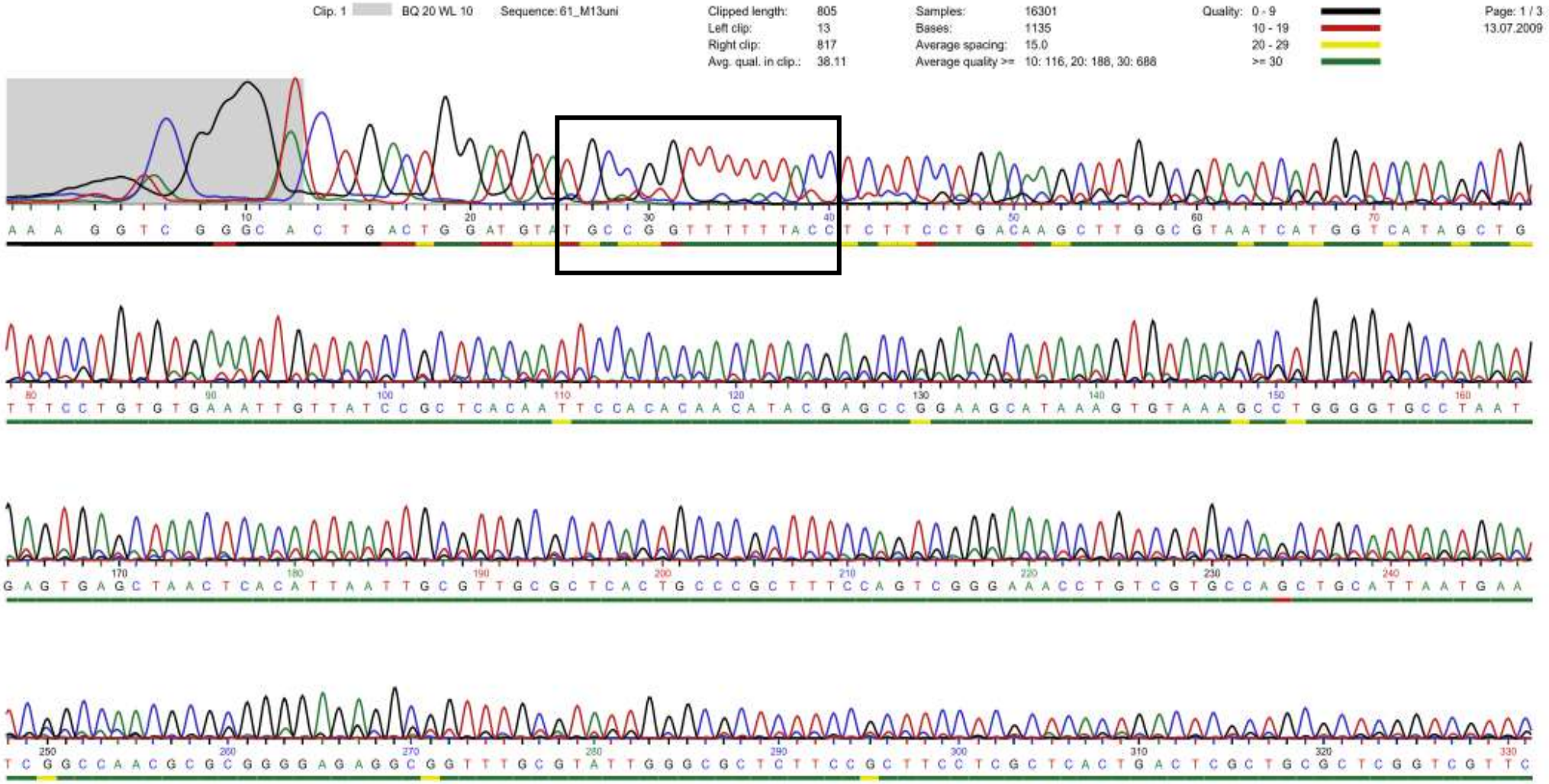
Sequence 6:


Clip. 1
BQ 20 WL 10
Sequence: 4s2_M13uni
Clipped length: 974
Left clip: 14
Right clip: 987
Avg. qual. in clip.: 52.65
Samples: 16300
Bases: 1255
Average spacing: 13.0
Average quality >= 10: 108, 20: 85, 30: 927
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
06.10.2009

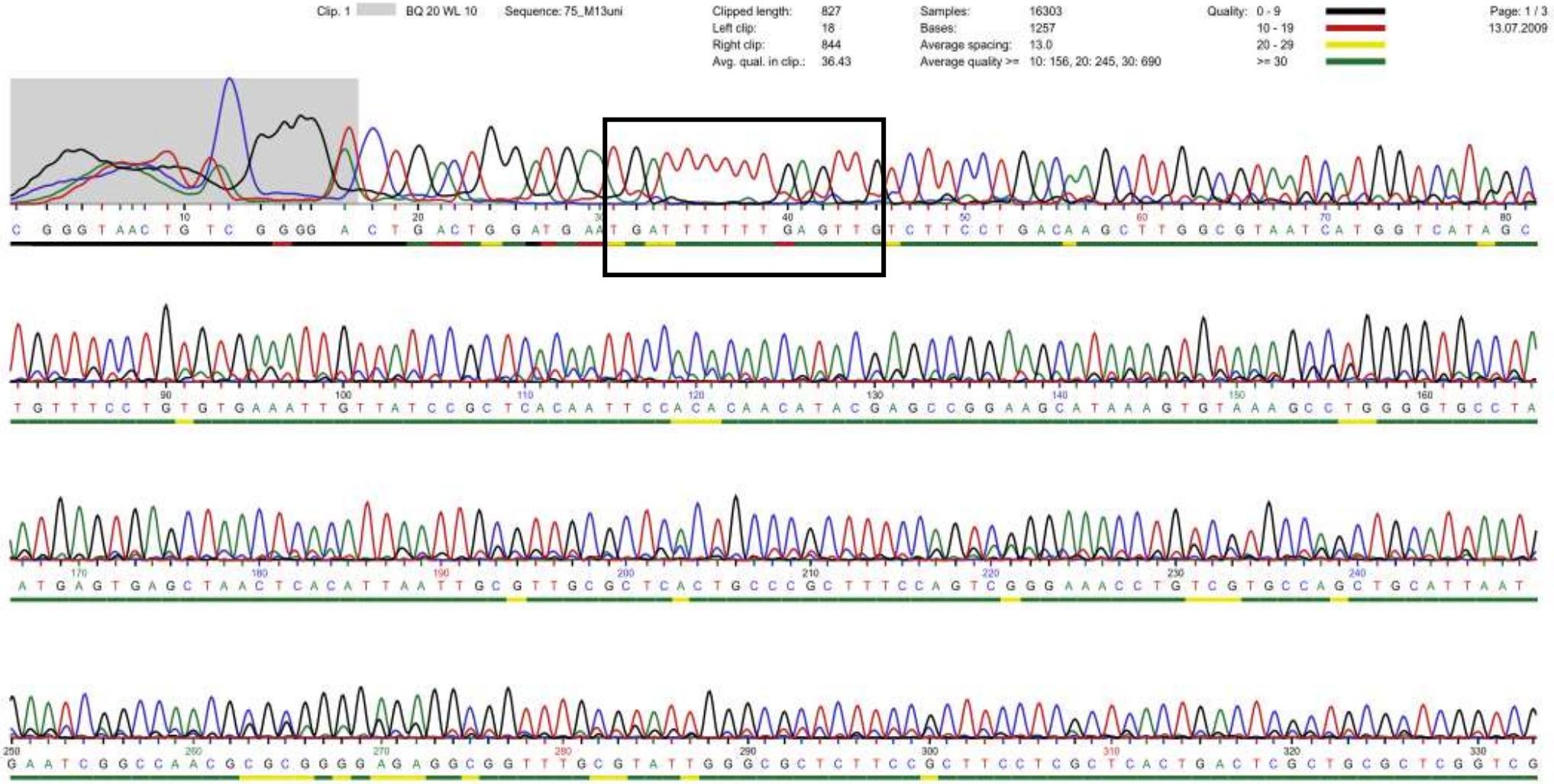
Sequence 7:


Clip. 1
BQ 20 WL 10
Sequence: 45_M13uni
Clipped length: 978
Left clip: 11
Right clip: 988
Samples: 16300
Bases: 1222
Average spacing: 14.0
Average quality >= 10: 97, 20: 89, 30: 893
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009

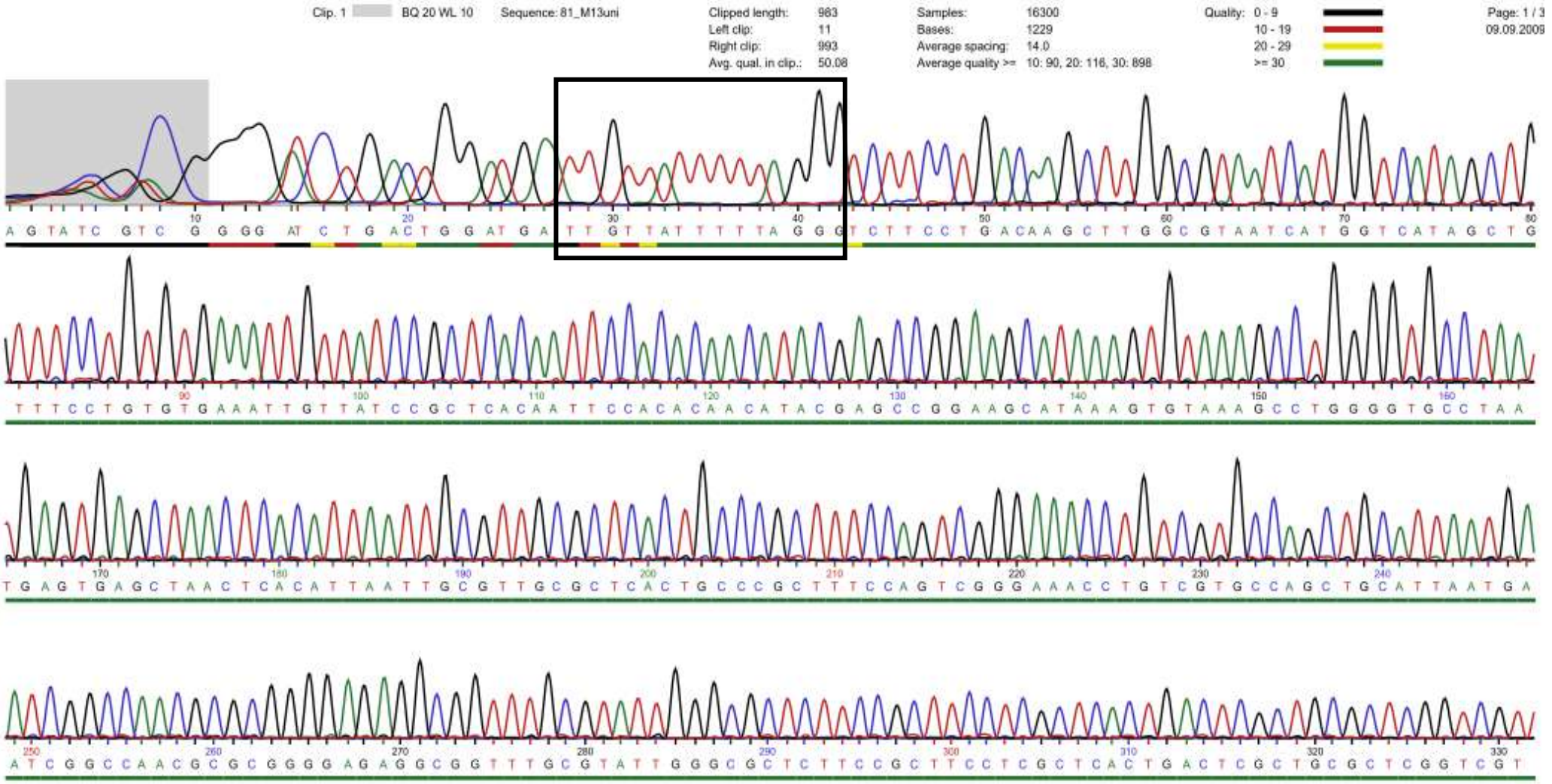
Sequence 8:


Clip. 1
BQ 20 WL 10
Sequence: 61_M13uni
Clipped length: 805
Left clip: 13
Right clip: 817
Avg. qual. in clip.: 38.11
Samples: 16301
Bases: 1135
Average spacing: 15.0
Average quality >= 10: 116, 20: 188, 30: 688
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
13.07.2009
AAA GGTC GGGC A C T G ACTG GAT GTAT GCCG GTTTTTT TACCT CTTCCT GACAAGCTT GGC GTAATCAT GGTCATAGCTG
TTTCCTGTGTGAAATTGTTATCCGC TCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGGTGCCTAAT
GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAA
TCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC

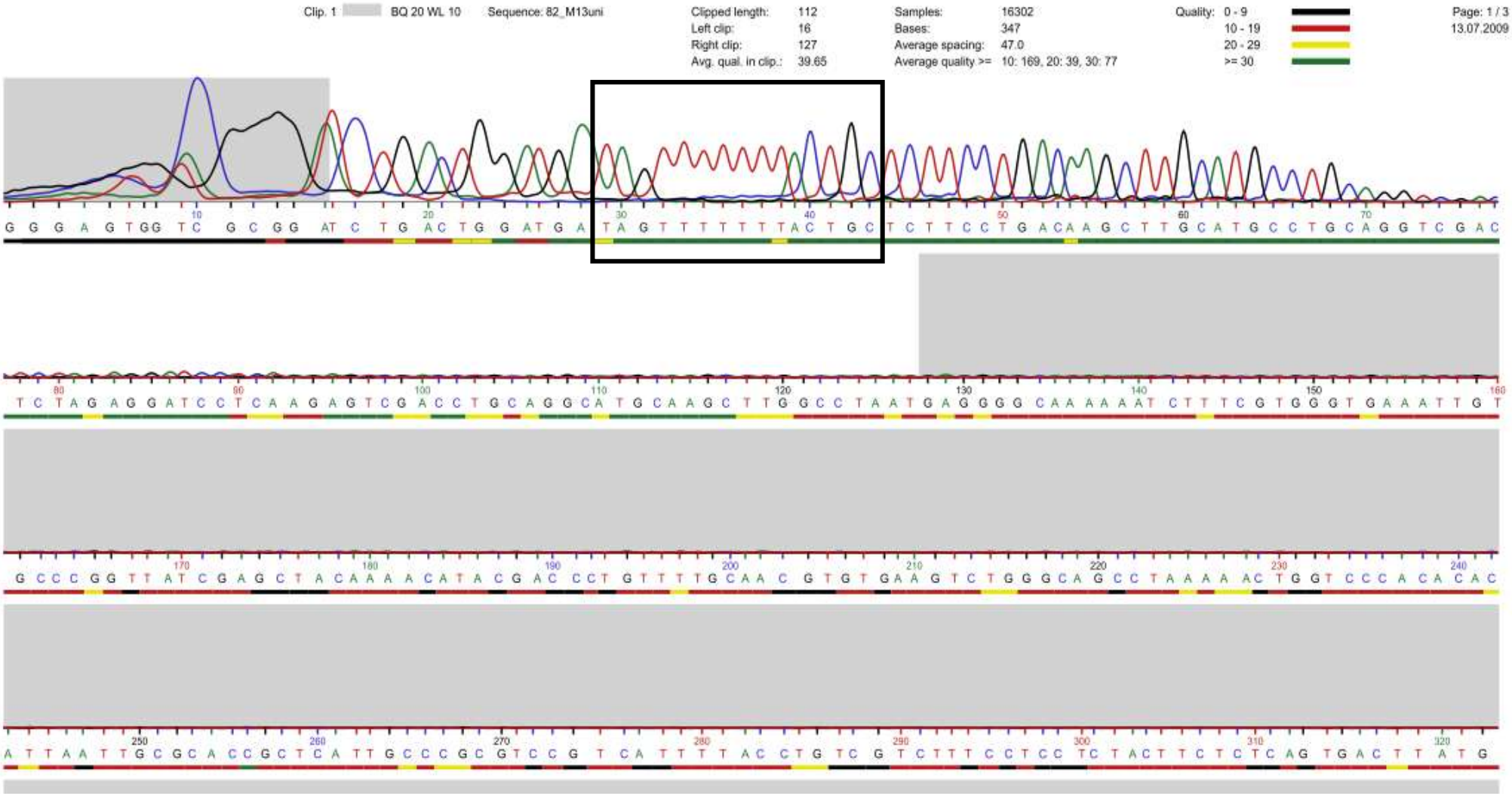
Sequence 9:


Clip. 1
BQ 20 WL 10
Sequence: 75_M13uni
Clipped length: 827
Left clip: 18
Right clip: 844
Avg. qual. in clip.: 36.43
Samples: 16303
Bases: 1257
Average spacing: 13.0
Average quality >= 10: 156, 20: 245, 30: 690
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
13.07.2009

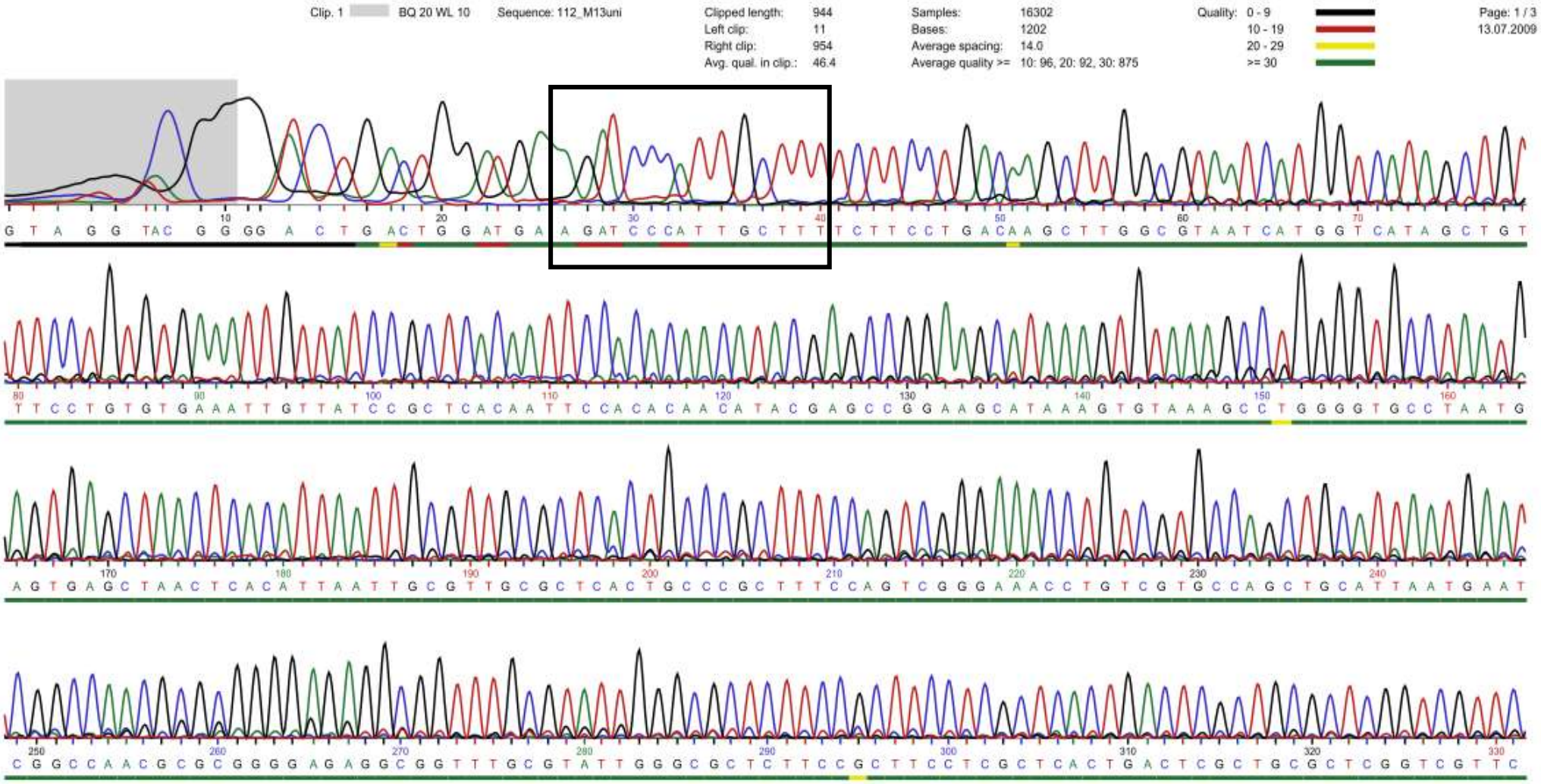
Sequence 10:


Clip. 1
BQ 20 WL 10
Sequence: 81_M13uni
Clipped length: 983
Left clip: 11
Right clip: 993
Avg. qual. in clip.: 50.08
Samples: 16300
Bases: 1229
Average spacing: 14.0
Average quality >= 10: 90, 20: 116, 30: 898
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009
A G T A T C G T C G G G G A T C T G A C T G G A T G A T T G T T A T T T T T A G G G T C T T C C T G A C A A G C T T G G C G T A A T C A T G G T C A T A G C T G
T T T C C T G T G T G A A A T T G T T A T C C G C T C A C A A T T C C A C A C A A A C A T A C G A G C C G G A A G C A T A A A G T G T A A A G C C T G G G G T G C C T A A
T G A G T G A G C T A A C T C A C A T T A A T T G C G T T G C G C T C A C T G C C C G C T T T C C A G T C G G G A A A C C T G T C G T G C C A G C T G C A T T A A T G A
A T C G G C C A A C G C G C G G G G A G A G G C G G T T T G C G T A T T G G G C G C T C T T C C G C T T C C T C G C T C A C T G A C T C G C T G C G C T C G G T C G T

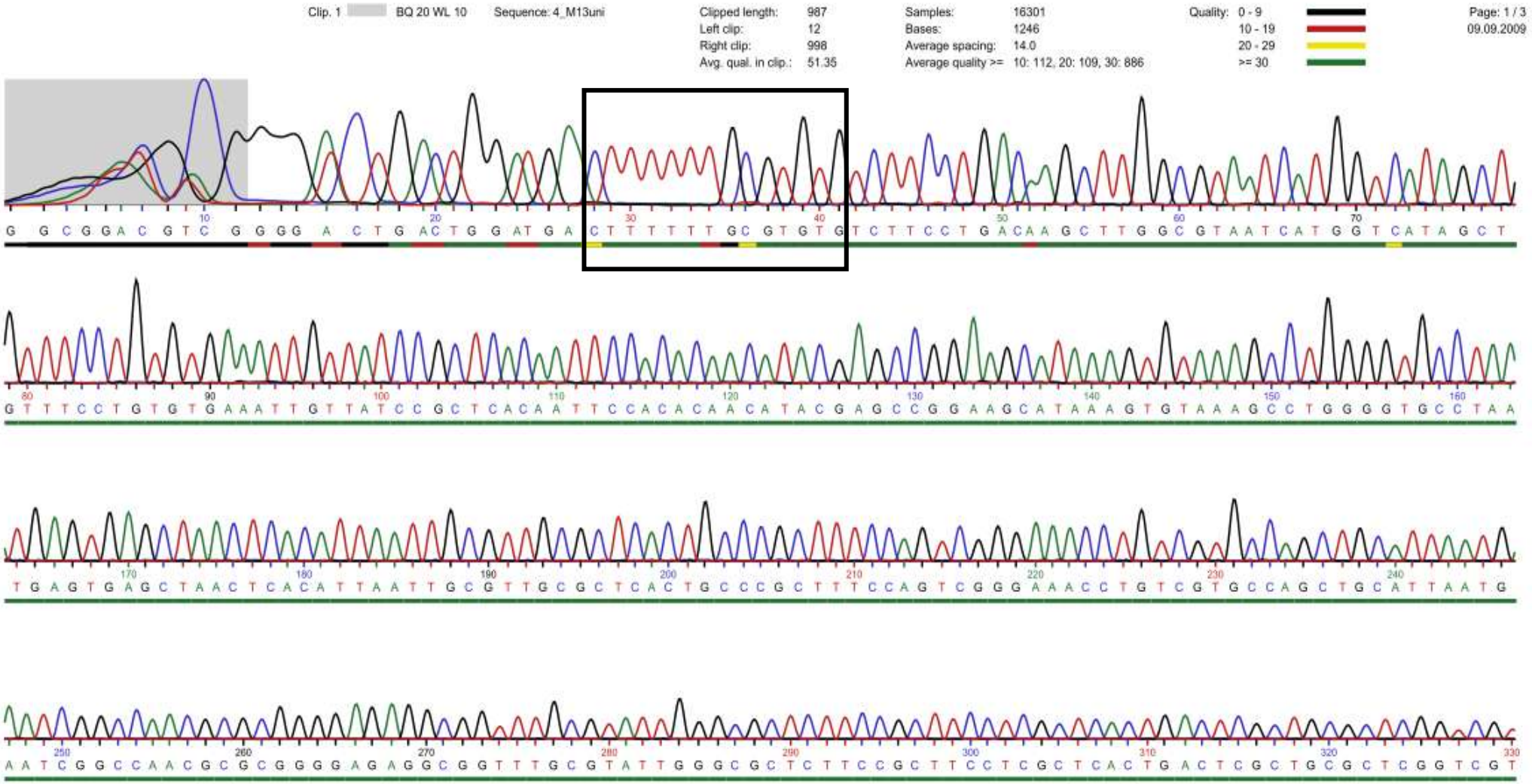
Sequence 11:


Clip. 1
BQ 20 WL 10
Sequence: 82_M13uni
Clipped length: 112
Left clip: 16
Right clip: 127
Avg. qual. in clip.: 39.65
Samples: 16302
Bases: 347
Average spacing: 47.0
Average quality >= 10: 169, 20: 39, 30: 77
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
13.07.2009
G G G A G T G G T C G C G G A T C T G A C T G G A T G A T A G T T T T T T T A C T G C T C T T C C T G A C A A G C T T G C A T G C C T G C A G G T C G A C
T C T A G A G G A T C C C T C A A G A G T C G A C C T G C A G G C A T G C A A G C T T G G C C T A A T G A G G G G C A A A A A A T C T T T C G T G G G T G A A A T T G T
G C C C G G T T A T C G A G C T A C A A A A C A T A C G A C C C T G T T T T G C A A C G T G T G A A G T C T G G G C A G C C T A A A A A C T G G T C C C A C A C A C
A T T A A T T G C G C A C C G C T C A T T G C C C G C G T C C G T C A T T T T T A C C T G T C G T C T T T T C C T C C T C T A C T T C T C T C A G T G A C T T A T G

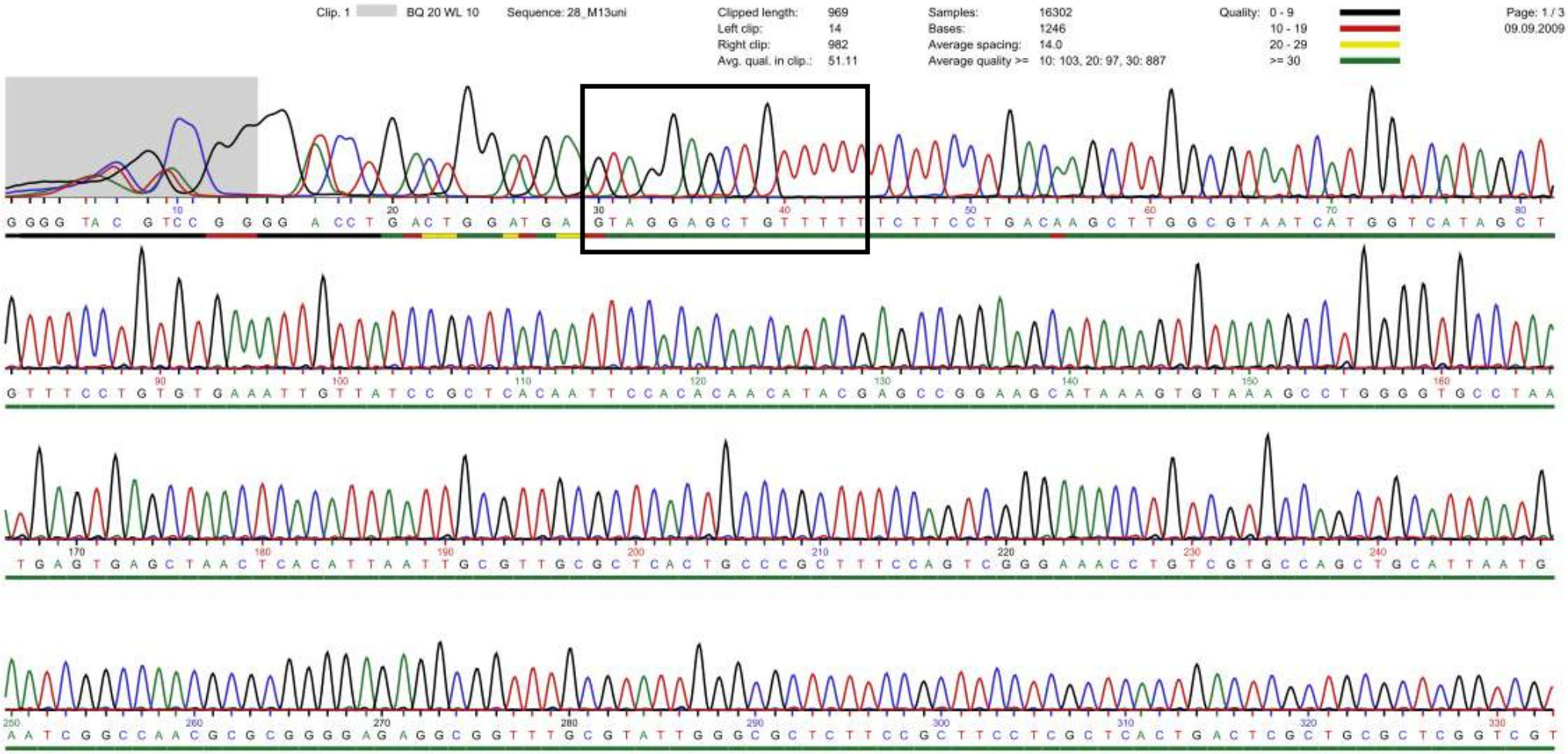
Sequence 12:


Clip. 1 BQ 20 WL 10 Sequence: 112_M13uni
Clipped length: 944
Left clip: 11
Right clip: 954
Avg. qual. in clip.: 46.4
Samples: 16302
Bases: 1202
Average spacing: 14.0
Average quality >= 10: 96, 20: 92, 30: 875
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
13.07.2009
G T A G G TAC G G GG A C T G AC T G G AT G A A GAT C C CA T T G C T T T T C T T C C T G A CAA G C T T G G C G T A A T C A T G G T C A T A G C T G T
T T C C T G T G T G A A A T T G T T A T C C G C T C A C A A T T C C A C A C A A C A T A C G A G C C G G A A G C A T A A A G T G T A A A G C C T G G G G T G C C T A A T G
A G T G A G C T A A C T C A C A T T A A T T G C G T T G C G C T C A C T G C C C G C T T T C C A G T C G G G A A A C C T G T C G T G C C A G C T G C A T T A A T G A A T
C G G C C A A C G C G C G G G G A G A G G C G G T T T G C G T A T T G G G C G C T C T T C C G C T T C C T C G C T C A C T G A C T C G C T G C G C T C G G T C G T T C

Sequence 13:


Clip. 1
BQ 20 WL 10
Sequence: 4_M13uni
Clipped length: 987
Left clip: 12
Right clip: 998
Avg. qual. in clip.: 51.35
Samples: 16301
Bases: 1246
Average spacing: 14.0
Average quality >= 10: 112, 20: 109, 30: 886
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009
G G C G G A C G T C G G G G A C T G A C T G G A T G A C T T T T T T G C G T G T G T C T T C C T G A C A A G C T T G G C G T A A T C A T G G T C A T A G C T
G T T T C C T G T G T G A A A T T G T T A T C C G C T C A C A A T T C C A C A C A A C A T A C G A G C C G G A A G C A T A A A G T G T A A A G C C T G G G G T G C C T A A
T G A G T G A G C T A A C T C A C A T T A A T T G C G T T G C G C T C A C T G C C C G C T T T C C A G T C G G G A A A C C T G T C G T G C C A G C T G C A T T A A T G
A A T C G G C C A A C G C G C G G G G A G A G G C G G T T T G C G T A T T G G G C G C T C T T C C G C T T C C T C G C T C A C T G A C T C G C T G C G C T C G G T C G T

Sequence 14:


Clip. 1
BQ 20 WL 10
Sequence: 28_M13uni
Clipped length: 969
Left clip: 14
Right clip: 982
Avg. qual. in clip.: 51.11
Samples: 16302
Bases: 1246
Average spacing: 14.0
Average quality >= 10: 103, 20: 97, 30: 887
Quality: 0 - 9
10 - 19
20 - 29
>= 30
Page: 1 / 3
09.09.2009
GGGG TA C GTCC G GGG A CCT G ACT G G AT G A GTA G G A G C T G T T T T T T C T T C C T G A C A A G C T T G G C G T A A T C A T G G T C A T A G C T
G T T T C C T G T G T G A A A T T G T T A T C C G C T C A C A A T T C C A C A C A A C A T A C G A G C C G G A A G C A T A A A G T G T A A A G C C T G G G G T G C C T A A
T G A G T G A G C T A A C T C A C A T T A A T T G C G T T G C G C T C A C T G C C C G C T T T C C A G T C G G G A A A C C T G T C G T G C C A G C T G C A T T A A T G
A A T C G G C C A A C G C G C G G G G A G A G G C G G T T T G C G T A T T G G G C G C T C T T C C G C T T C C T C G C T C A C T G A C T C G C T G C G C T C G G T C G T

Reference List

1. Crick,F.H.C. and Watson,J.D. (1954) The Complementary Structure of Deoxyribonucleic Acid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 223, 80-96.

2. Felsenfeld,G. and Rich,A. (1957) Studies on the formation of two- and three-stranded polyribonucleotides. *Biochim. Biophys. Acta*, 26, 457-468.

3. Beal,P.A. and Dervan,P.B. (1991) Second Structural Motif for Recognition of DNA by Oligonucleotide-Directed Triple-Helix Formation. *Science*, 251, 1360-1363.

4. Morgan,A.R. and Wells,R.D. (1968) Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.*, 37, 63-80.

5. Riley,M. and Maling,B. (1966) Physical and chemical characterization of two- and three-stranded adenine-thymine and adenine-uracil homopolymer complexes. *J. Mol. Biol.*, 20, 359-389.

6. Marck,C. and Thiele,D. (1978) Poly(dG).poly(dC) at neutral and alkaline pH: the formation of triple stranded poly(dG).poly(dG).poly(dC). *Nucleic Acids Res.*, 5, 1017-1028.

7. Lee,J.S., Johnson,D.A. and Morgan,A.R. (1979) Complexes formed by (pyrimidine)n . (purine)n DNAs on lowering the pH are three-stranded. *Nucleic Acids Res.*, 6, 3073-3091.

8. Broitman,S.L., Im,D.D. and Fresco,J.R. (1987) Formation of the triple-stranded polynucleotide helix, poly(A.A.U). *Proc. Natl. Acad. Sci. U. S. A*, 84, 5120-5124.

9. Protozanova,E. and Macgregor,R.B., Jr. (1996) Kinetic footprinting of DNA triplex formation. *Anal. Biochem.*, 243, 92-99.

10. Frank-Kamenetskii,M.D. and Mirkin,S.M. (1995) Triplex DNA Structures. *Annual Review of Biochemistry*, 64, 65-95.

11. Wells,R.D. (1988) Unusual DNA structures. *J. Biol. Chem.*, 263, 1095-1098.

12. Duca,M., Vekhoff,P., Oussedik,K., Halby,L. and Arimondo,P.B. (2008) The triple helix: 50 years later, the outcome. *Nucl. Acids Res.*, gkn493.

13. Vasquez,K.M. and Glazer,P.M. (2002) Triplex-forming oligonucleotides: principles and applications. *Q. Rev. Biophys.*, 35, 89-107.

14. Moser,H.E. and Dervan,P.B. (1987) Sequence-specific cleavage of double helical DNA by triple helix formation. *Science*, 238, 645-650.

15. Le Doan,T., Perrouault,L., Praseuth,D., Habhoub,N., Decout,J.L., Thuong,N.T., Lhomme,J. and Heene,C. (1987) Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-(alpha]-thymidylate covalently linked to an azidoproflavine derivative. *Nucl. Acids Res.*, 15, 7749-7760.

16. Wells,R.D., Collier,D.A., Hanvey,J.C., Shimizu,M. and Wohlrab,F. (1988) The chemistry and biology of unusual DNA structures adopted by oligopurine.oligopyrimidine sequences. *FASEB J.*, 2, 2939-2949.

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

18. Hoogsteen,K. (1967) The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallographica*, 16, 907-916.

19. Christophe,D., Cabrer,B., Bacolla,A., Targovnik,H., Pohl,V. and Vassart,G. (1985) An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene. *Nucl. Acids Res.*, 13, 5127-5144.

20. Mirkin,S.M., Lyamichev,V.I., Drushlyak,K.N., Dobrynin,V.N., Filippov,S.A. and Frank-Kamenetskii,M.D. (1987) DNA H form requires a homopurine-homopyrimidine mirror repeat. *Nature*, 330, 495-497.

21. Belotserkovskii,B.P., De Silva,E., Tornaletti,S., Wang,G., Vasquez,K.M. and Hanawalt,P.C. (2007) A Triplex-forming Sequence from the Human c-MYC Promoter Interferes with DNA Transcription. *J. Biol. Chem.*, 282, 32433-32441.

22. Jain,A., Wang,G. and Vasquez,K.M. (2008) DNA triple helices: Biological consequences and therapeutic potential. *Biochimie*, 90, 1117-1130.

23. Burkholder,G.D., Latimer,L.J.P. and Lee,J.S. (1988) Immunofluorescent staining of mammalian nuclei and chromosomes with a monoclonal antibody to triplex DNA. *Chromosoma*, 97, 185-192.

24. Burkholder,G.D., Latimer,L.J.P. and Lee,J.S. (1991) Immunofluorescent localization of triplex DNA in polytene chromosomes of Chironomus and Drosophila. *Chromosoma*, 101, 11-18.

25. Goni,J.R., de la Cruz,X. and Orozco,M. (2004) Triplex-forming oligonucleotide target sequences in the human genome. *Nucl. Acids Res.*, 32, 354-360.

26. Liu,L.F. and Wang,J.C. (1987) Supercoiling of the DNA Template during Transcription. *Proceedings of the National Academy of Sciences of the United States of America*, 84, 7024-7027.

27. Malkov,V.A., Voloshin,O.N., Soyfer,V.N. and Frank-Kamenetskii,M.D. (1993) Cation and sequence effects on stability of intermolecular pyrimidine-purine-purine triplex. *Nucl. Acids Res.*, 21, 585-591.

28. Htun,H. and Dahlberg,J.E. (1988) Single strands, triple strands, and kinks in H-DNA. *Science*, 241, 1791-1796.

29. Larsen,A. and Wells,R.D. (1982) An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. *Cell*, 29, 609-622.

30. Lee,J.S., Woodsworth,M.L., Latimer,L.J. and Morgan,A.R. (1984) Poly(pyrimidine) . poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. *Nucleic Acids Res.*, 12, 6603-6614.

31. Zain,R. and Sun,J.S. (2003) Do natural DNA triple-helical structures occur and function in vivo? *Cellular and Molecular Life Sciences*, 60, 862-870.

32. Mirkin,S.M. and Frank-Kamenetskii,M.D. (1994) H-DNA and Related Structures. *Annual Review of Biophysics and Biomolecular Structure*, 23, 541-576.

33. Lee,J.S., Burkholder,G.D., Latimer,L.J.P., Haug,B.L. and Braun,R.P. (1987) A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes. *Nucl. Acids Res.*, 15, 1047-1061.

34. Agazie,Y.M., Burkholder,G.D. and Lee,J.S. (1996) Triplex DNA in the nucleus: direct binding of triplex-specific antibodies and their effect on transcription, replication and cell growth. *Biochemical Journal*, 316, 461-466.

35. Ohno,M., Fukagawa,T., Lee,J. and Ikemura,T. (2002) Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies. *Chromosoma*, 111, 201-213.

36. Ledoan,T., Perrouault,L., Chassignol,M., Thuong,N.T. and Helene,C. (1987) Sequence-Targeted Chemical Modifications of Nucleic-Acids by Complementary Oligonucleotides Covalently Linked to Porphyrins. *Nucl. Acids Res.*, 15, 8643-8659.

37. Gowers,D.M. and Fox,K.R. (1999) Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res.*, 27, 1569-1577.

38. Felsenfeld,G., Davies,D.R. and Rich,A. (1957) Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.*, 79, 2023-2024.

39. Pilch,D.S., Levenson,C. and Shafer,R.H. (1990) Structural-Analysis of the (dA)10.2(dT)10 Triple Helix. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 1942-1946.

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

41. Yoon,K., Hobbs,C.A., Koch,J., Sardaro,M., Kutny,R. and Weis,A.L. (1992) Elucidation of the sequence-specific third-strand recognition of four Watson-Crick base pairs in a pyrimidine triple-helix motif: T.AT, C.GC, T.CG, and G.TA. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 3840-3844.

42. Radhakrishnan,I. and Patel,D.J. (1994) DNA triplexes: solution structures, hydration sites, energetics, interactions, and function. *Biochemistry*, 33, 11405-11416.

43. Kohwi,Y. and Kohwi-Shigematsu,T. (1988) Magnesium Ion-Dependent Triple-Helix Structure Formed by Homopurine-Homopyrimidine Sequences in Supercoiled Plasmid DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 85, 3781-3785.

44. Beal,P.A. and Dervan,P.B. (1992) The influence of single base triplet changes on the stability of a pur.pur.pyr triple helix determined by affinity cleaving. *Nucl. Acids Res.*, 20, 2773-2776.

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

46. Pilch,D.S., Levenson,C. and Shafer,R.H. (1991) Structure, stability, and thermodynamics of a short intermolecular purine-purine-pyrimidine triple helix. *Biochemistry*, 30, 6081-6087.

47. Cooney,M., Czernuszewicz,G., Postel,E.H., Flint,S.J. and Hogan,M.E. (1988) Site-specific oligonucleotide binding represses transcription of the human c-myc gene in vitro. *Science*, 241, 456-459.

48. Letai,A.G., Palladino,M.A., Fromm,E., Rizzo,V. and Fresco,J.R. (1988) Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns. *Biochemistry*, 27, 9108-9112.

49. Durland,R.H., Kessler,D.J., Gunnell,S., Duvic,M., Hogan,M.E. and Pettitt,B.M. (1991) Binding of triple helix forming oligonucleotides to sites in gene promoters. *Biochemistry*, 30, 9246-9255.

50. Faucon,B., Mergny,J.L. and Helene,C. (1996) Effect of third strand composition on the triple helix formation: purine versus pyrimidine oligodeoxynucleotides. *Nucl. Acids Res.*, 24, 3181-3188.

51. Roy,C. (1993) Inhibition of gene transcription by purine rich triplex forming oligodeoxyribonucleotides. *Nucl. Acids Res.*, 21, 2845-2852.

52. Klysik,J., Kinsey,B.M., Hua,P., Glass,G.A. and Orson,F.M. (1997) A 15-Base Acridine-Conjugated Oligodeoxynucleotide Forms Triplex DNA with Its IL-2Rá Promoter Target with Greatly Improved Avidity. *Bioconjugate Chem.*, 8, 318-326.

53. Paes,H.M. and Fox,K.R. (1997) Kinetic studies on the formation of intermolecular triple helices. *Nucl. Acids Res.*, 25, 3269-3274.

54. Cheng,A.j. and Van Dyke,M.W. (1993) Monovalent cation effects on intermolecular purine-purine-pyrimidine triple-helix formation. *Nucl. Acids Res.*, 21, 5630-5635.

55. Olivas,W.M. and Maher,L.J., III (1995) Overcoming potassium-mediated triplex inhibition. *Nucl. Acids Res.*, 23, 1936-1941.

56. Arimondo,P.B., Barcelo,F., Sun,J.S., Maurizot,J.C., Garestier,T. and Helene,C. (1998) Triple Helix Formation by (G,A)-Containing Oligonucleotides: Asymmetric Sequence Effect. *Biochemistry*, 37, 16627-16635.

57. Noonberg,S.B., Francois,J.C., Garestier,T. and Heiene,C. (1995) Effect of competing self-structure on triplex formation with purine-rich oligodeoxynucleotides containing GA repeats. *Nucl. Acids Res.*, 23, 1956-1963.

58. Noonberg,S.B., Francois,J.C., Praseuth,D., Guieysse-Peugeot,A.L., Lacoste,J., Garestier,T. and Helene,C. (1995) Triplex formation with alpha anomers of purine-rich and pyrimidine-rich oligodeoxynucleotides. *Nucl. Acids Res.*, 23, 4042-4049.

59. Arnott,S. and Selsing,E. (1974) Structures for the polynucleotide complexes poly(dA) À poly(dT) and poly(dT) À poly(dA) À poly(dT). *Journal of Molecular Biology*, 88, 509-512.

60. Arnott,S., Chandrasekaran,R., Hukins,D.W.L., Smith,P.J.C. and Watts,L. (1974) Structural details of a double-helix observed for DNAs containing alternating purine and pyrimidine sequences. *Journal of Molecular Biology*, 88, 523-524.

61. Rajagopal,P. and Feigon,J. (1989) Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides d(G-A)4 and d(T-C)4. *Nature*, 339, 637-640.

62. Rajagopal,P. and Feigon,J. (1989) NMR studies of triple-strand formation from the homopurine-homopyrimidine deoxyribonucleotides d(GA)4 and d(TC)4. *Biochemistry*, 28, 7859-7870.

63. De los Santos,C., Rosen,M. and Patel,D. (1989) NMR studies of DNA (R+)n.cntdot.(Y-)n.cntdot.(Y+)n triple helixes in solution: imino and amino proton markers of T.cntdot.A.cntdot.T and C.cntdot.G.cntdot.C+ base-triple formation. *Biochemistry*, 28, 7282-7289.

64. Radhakrishnan,I., Patel,D.J., Priestley,E.S., Nash,H.M. and Dervan,P.B. (1993) Nmr Structural Studies on A Nonnatural Deoxyribonucleoside Which Mediates Recognition of Gc Base-Pairs in Pyrimidine-Center-Dot-Purine-Center-Dot-Pyrimidine Dna Triplexes. *Biochemistry*, 32, 11228-11234.

65. Rhee,S., Han,Z.j., Liu,K., Miles,H.T. and Davies,D.R. (1999) Structure of a Triple Helical DNA with a Triplex–Duplex Junction. *Biochemistry*, 38, 16810-16815.

66. Ouali,M., Letellier,R., Adnet,F., Liquier,J., Sun,J.s., Lavery,R. and Taillandier,E. (1993) A possible family of B-like triple helix structures: Comparison with the Arnott A-like triple helix. *Biochemistry*, 32, 2098-2103.

67. Radhakrishnan,I. and Patel,D.J. (1994) Solution Structure of A Pyrimidine.Purine.Pyrimidine DNA Triplex Containing T.AT C+.GC and G.TA Triples. *Structure*, 2, 17-32.

68. Radhakrishnan,I. and Patel,D.J. (1994) Solution Structure and Hydration Patterns of a Pyrimidine-Purine-Pyrimidine DNA Triplex Containing a Novel T-CG Base-triple. *Journal of Molecular Biology*, 241, 600-619.

69. Thuong,N.T. and Helene,C. (1993) Sequence-Specific Recognition and Modification of Double-Helical Dna by Oligonucleotides. *Angewandte Chemie-International Edition in English*, 32, 666-690.

70. Hartman,D.A., Kuo,S.R., Broker,T.R., Chow,L.T. and Wells,R.D. (1992) Intermolecular triplex formation distorts the DNA duplex in the regulatory region of human papillomavirus type-11. *J. Biol. Chem.*, 267, 5488-5494.

71. Lyamichev,V.I., Mirkin,S.M., Frank-Kamenetskii,M.D. and Cantor,C.R. (1988) A stable complex between homopyrimidine oligomers and the homologous regions of duplex DNAs. *Nucl. Acids Res.*, 16, 2165-2187.

72. Radhakrishnan,I., De Los Santos,C. and Patel,D.J. (1991) Nuclear magnetic resonance structural studies of intramolecular purine.purine.pyrimidine DNA triplexes in solution. Base triple pairing alignments and strand direction. *J. Mol. Biol.*, 221, 1403-1418.

73. Radhakrishnan,I. and Patel,D.J. (1994) Hydration sites in purine-purine-pyrimidine and pyrimidine-purine-pyrimidine DNA triplexes in aqueous solution. *Structure*, 2, 395-405.

74. Rougee,M., Faucon,B., Mergny,J.L., Barcelo,F., Giovannangeli,C., Garestier,T. and Helene,C. (1992) Kinetics and Thermodynamics of Triple-Helix Formation - Effects of Ionic-Strength and Mismatches. *Biochemistry*, 31, 9269-9278.

75. Giovannangeli,C., Rougee,M., Garestier,T., Thuong,N.T. and Helene,C. (1992) Triple-Helix Formation by Oligonucleotides Containing the 3 Bases Thymine, Cytosine, and Guanine. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 8631-8635.

76. Staubli,A.B. and Dervan,P.B. (1994) Sequence specificity of the non-natural pyrido[2,3-d]pyrimidine nucleoside in triple helix formation. *Nucl. Acids Res.*, 22, 2637-2642.

77. Hardenbol,P. and VanDyke,M.W. (1996) Sequence specificity of triplex DNA formation: Analysis by a combinatorial approach restriction endonuclease protection selection and amplification. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 2811-2816.

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

79. Roberts,R.W. and Crothers,D.M. (1996) Prediction of the stability of DNA triplexes. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 4320-4325.

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

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

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

83. Svinarchuk,F., Paoletti,J. and Malvy,C. (1995) An unusually stable Purine(Purine-Pyrimidine) short triplex. *J. Biol. Chem.*, 270, 14068-14071.

84. Svinarchuk,F., Bertrand,J.R. and Malvy,C. (1994) A short purine oligonucleotide forms a highly stable triple helix with the promoter of the murine c-pim-1 proto-oncogene. *Nucl. Acids Res.*, 22, 3742-3747.

85. Craig,M.E., Crothers,D.M. and Doty,P. (1971) Relaxation kinetics of dimer formation by self complementary oligonucleotides. *Journal of Molecular Biology*, 62, 383-392.

86. Rao,B.J. and Radding,C.M. (1994) Formation of base triplets by non-Watson-Crick bonds mediates homologous recognition in RecA recombination filaments. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 6161-6165.

87. Watnick,T.J., Piontek,K.B., Cordal,T.M., Weber,H., Gandolph,M.A., Qian,F., Lens,X.M., Neumann,H.P. and Germino,G.G. (1997) An unusual pattern of

mutation in the duplicated portion of PKD1 is revealed by use of a novel strategy for mutation detection. *Hum. Mol. Genet.*, 6, 1473-1481.

88. Van Raay,T.J., Burn,T.C., Connors,T.D., Petery,L.R., Germino,G.G., Klinger,K.W. and Landes,G.M. (1996) A 2.5 kb Polypyrimidine Tract in the PKD1 Gene Contains at Least 23 H-DNA-Forming Sequences. *Genome Science and Technology*, 1, 317-327.

89. Blaszak,R.T., Potaman,V., Sinden,R.R. and Bissler,J.J. (1999) DNA structural transitions within the PKD1 gene. *Nucl. Acids Res.*, 27, 2610-2617.

90. Wang,G. and Vasquez,K.M. (2004) Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 13448-13453.

91. Bacolla,A., Jaworski,A., Larson,J.E., Jakupciak,J.P., Chuzhanova,N., Abeysinghe,S.S., O'Connell,C.D., Cooper,D.N. and Wells,R.D. (2004) Breakpoints of gross deletions coincide with non-B DNA conformations. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 14162-14167.

92. Gaddis,S.S., Wu,Q., Thames,H.D., Digiovanni,J., Walborg,E.F., MacLeod,M.C. and Vasquez,K.M. (2006) A web-based search engine for triplex-forming oligonucleotide target sequences. *Oligonucleotides*, 16, 196-201.

93. Goni,J., Vaquerizas,J., Dopazo,J. and Orozco,M. (2006) Exploring the reasons for the large density of triplex-forming oligonucleotide target sequences in the human regulatory regions. *BMC Genomics*, 7, 63.

94. Bacolla,A., Collins,J.R., Gold,B., Chuzhanova,N., Yi,M., Stephens,R.M., Stefanov,S., Olsh,A., Jakupciak,J.P., Dean,M. *et al.* (2006) Long homopurine*homopyrimidine sequences are characteristic of genes expressed in brain and the pseudoautosomal region. *Nucl. Acids Res.*, 34, 2663-2675.

95. Strobel,S.A., Doucette-Stamm,L.A., Riba,L., Housman,D.E. and Dervan,P.B. (1991) Site-specific cleavage of human chromosome 4 mediated by triple-helix formation. *Science*, 254, 1639-1642.

96. Strobel,S.A. and Dervan,P.B. (1991) Single-site enzymatic cleavage of yeast genomic DNA mediated by triple helix formation. *Nature*, 350, 172-174.

97. Maher,L.J., Dervan,P.B. and Wold,B.J. (1990) Kinetic analysis of oligodeoxyribonucleotide-directed triple-helix formation on DNA. *Biochemistry*, 29, 8820-8826.

98. Maher,L.J., Dervan,P.B. and Wold,B. (1992) Analysis of promoter-specific repression by triple-helical DNA complexes in a eukaryotic cell-free transcription system. *Biochemistry*, 31, 70-81.

99. Young,S.L., Krawczyk,S.H., Matteucci,M.D. and Toole,J.J. (1991) Triple helix formation inhibits transcription elongation in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 10023-10026.

100. Knauert,M.P. and Glazer,P.M. (2001) Triplex forming oligonucleotides: sequence-specific tools for gene targeting. *Hum. Mol. Genet.*, 10, 2243-2251.

101. Bonham,M.A., Brown,S., Boyd,A.L., Brown,P.H., Bruckenstein,D.A., Hanvey,J.C., Thomson,S.A., Pipe,A., Hassman,F., Bisi,J.E. *et al.* (1995) An assessment of the antisense properties of RNase H-competent and steric-blocking oligomers. *Nucl. Acids Res.*, 23, 1197-1203.

102. Mayfield,C., Squibb,M. and Miller,D. (1994) Inhibition of Nuclear Protein Binding to the Human Ki-ras Promoter by Triplex-Forming Oligonucleotides. *Biochemistry*, 33, 3358-3363.

103. Kim,H.G., Reddoch,J.F., Mayfield,C., Ebbinghaus,S., Vigneswaran,N., Thomas,S., Jones,J. and Miller,D. (1998) Inhibition of Transcription of the Human c-myc Protooncogene by Intermolecular Triplex. *Biochemistry*, 37, 2299-2304.

104. Re,R.N., Cook,J.L. and Giardina,J.F. (2004) The inhibition of tumor growth by triplex-forming oligonucleotides. *Cancer Lett.*, 209, 51-53.

105. Volkmann,S., Jendis,J., Frauendorf,A. and Moelling,K. (1995) Inhibition of HIV-1 reverse transcription by triple-helix forming oligonucleotides with viral RNA. *Nucl. Acids Res.*, 23, 1204-1212.

106. Hoyne,P.R. and Maher III,L.J. (2002) Functional Studies of Potential Intrastrand Triplex Elements in the Escherichia coli Genome. *Journal of Molecular Biology*, 318, 373-386.

107. Postel,E.H., Flint,S.J., Kessler,D.J. and 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. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 8227-8231.

108. Giovannangeli,C., Thuong,N.T. and Helene,C. (1993) Oligonucleotide clamps arrest DNA synthesis on a single-stranded DNA target. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 10013-10017.

109. Chan,P.P. and Glazer,P.M. (1997) Triplex DNA: Fundamentals, advances, and potential applications for gene therapy. *Journal of Molecular Medicine-Jmm*, 75, 267-282.

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

111. Francois,J.C., Saison-Behmoaras,T., Barbier,C., Chassignol,M., Thuong,N.T. and Helene,C. (1989) Sequence-specific recognition and cleavage of duplex DNA via triple-helix formation by oligonucleotides covalently linked to a phenanthroline-copper chelate. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 9702-9706.

112. Vasquez,K.M. and Wilson,J.H. (1998) Triplex-directed modification of genes and gene activity. *Trends in Biochemical Sciences*, 23, 4-9.

113. Vasquez,K.M., Wensel,T.G., Hogan,M.E. and Wilson,J.H. (1996) High-Efficiency Triple-Helix-Mediated Photo-Cross-Linking at a Targeted Site within a Selectable Mammalian Gene. *Biochemistry*, 35, 10712-10719.

114. Vasquez,K.M., Narayanan,L. and Glazer,P.M. (2000) Specific Mutations Induced by Triplex-Forming Oligonucleotides in Mice. *Science*, 290, 530-533.

115. Carbone,G.M., McGuffie,E., Napoli,S., Flanagan,C.E., Dembech,C., Negri,U., Arcamone,F., Capobianco,M.L. and Catapano,C.V. (2004) DNA binding and antigene activity of a daunomycin-conjugated triplex-forming oligonucleotide targeting the P2 promoter of the human c-myc gene. *Nucl. Acids Res.*, 32, 2396-2410.

116. Napoli,S., Negri,U., Arcamone,F., Capobianco,M.L., Carbone,G.M. and Catapano,C.V. (2006) Growth inhibition and apoptosis induced by daunomycin-conjugated triplex-forming oligonucleotides targeting the c-myc gene in prostate cancer cells. *Nucl. Acids Res.*, 34, 734-744.

117. Potaman,V.N., Oussatcheva,E.A., Lyubchenko,Y.L., Shlyakhtenko,L.S., Bidichandani,S.I., Ashizawa,T. and Sinden,R.R. (2004) Length-dependent structure formation in Friedreich ataxia (GAA)n.(TTC)n repeats at neutral pH. *Nucl. Acids Res.*, 32, 1224-1231.

118. Wells,R.D. (2008) DNA triplexes and Friedreich ataxia. *FASEB J.*, 22, 1625-1634.

119. Ruan,H. and Wang,Y.H. (2008) Friedreich's Ataxia GAAÀTTC Duplex and GAAÀGAAÀTTC Triplex Structures Exclude Nucleosome Assembly. *Journal of Molecular Biology*, 383, 292-300.

120. Walsh,T., McClellan,J.M., McCarthy,S.E., Addington,A.M., Pierce,S.B., Cooper,G.M., Nord,A.S., Kusenda,M., Malhotra,D., Bhandari,A. *et al.* (2008) Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia. *Science*, 320, 539-543.

121. Borgatti,M., Lampronti,I., Romanelli,A., Pedone,C., Saviano,M., Bianchi,N., Mischiati,C. and Gambari,R. (2003) Transcription Factor Decoy Molecules Based on a Peptide Nucleic Acid (PNA)-DNA Chimera Mimicking Sp1 Binding Sites. *J. Biol. Chem.*, 278, 7500-7509.

122. Svinarchuk,F., Nagibneva,I., Cherny,D., it-Si-Ali,S., Pritchard,L.L., Robin,P., Malvy,C., Harel-Bellan,A. and Chern,D. (1997) Recruitment of transcription factors to the target site by triplex- forming oligonucleotides. *Nucl. Acids Res.*, 25, 3459-3464.

123. Kuznetsova,S., it-Si-Ali,S., Nagibneva,I., Troalen,F., Le Villain,J.P., Bellan,A. and Svinarchuk,F. (1999) Gene activation by triplex-forming oligonucleotide coupled to the activating domain of protein VP16. *Nucl. Acids Res.*, 27, 3995-4000.

124. Noonberg,S.B., Scott,G.K., Garovoy,M.R., Benz,C.C. and Hunt,C.A. (1994) In vivo generation of highly abundant sequence-specific oligonucleotides for antisense and triplex gene regulation. *Nucl. Acids Res.*, 22, 2830-2836.

125. Wickstrom,E. (1986) Oligodeoxynucleotide stability in subcellular extracts and culture media. *Journal of Biochemical and Biophysical Methods*, 13, 97-102.

126. Stein,C.A. and Cheng,Y.C. (1993) Antisense oligonucleotides as therapeutic agents--is the bullet really magical? *Science*, 261, 1004-1012.

127. Gewirtz,A.M., Stein,C.A. and Glazer,P.M. (1996) Facilitating oligonucleotide delivery: helping antisense deliver on its promise. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 3161-3163.

128. Rao,T.S., Durland,R.H., Seth,D.M., Myrick,M.A., Bodepudi,V. and Revankar,G.R. (1995) Incorporation of 2'-Deoxy-6-thioguanosine into G-Rich Oligodeoxyribonucleotides Inhibits G-Tetrad Formation and Facilitates Triplex Formation. *Biochemistry*, 34, 765-772.

129. Milligan,J.F., Krawazyk,S.H., Wadwani,S. and Matteucci,M.D. (1993) An anti-parallel triple helix motif with oligodeoxynucleotidescontaining 2'-deoxyguanosine and 7-deaza-2'-deoxy xanthosine. *Nucl. Acids Res.*, 21, 327-333.

130. Loke,S.L., Stein,C.A., Zhang,X.H., Mori,K., Nakanishi,M., Subasinghe,C., Cohen,J.S. and Neckers,L.M. (1989) Characterization of oligonucleotide transport into living cells. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 3474-3478.

131. Krieg,A.M., Gmelig-Meyling,F.R.I.T., Gourley,M.F., Kisch,W.J., Hrisey,L.A. and Steinberg,A.D. (1991) Uptake of Oligodeoxyribonucleotides by Lymphoid Cells Is Heterogeneous and Inducible. *Antisense Research and Development*, 1, 161-171.

132. Dobbelstein,M. (2003) Viruses in therapy--royal road or dead end? *Virus Research*, 92, 219-221.

133. Boletta,A., Benigni,A., Lutz,J., Remuzzi,G., Soria,M.R. and Monaco,L. (1997) Nonviral Gene Delivery to the Rat Kidney with Polyethylenimine. *Human Gene Therapy*, 8, 1243-1251.

134. Boussif,O., Lezoualc'h,F., Zanta,M.A., Mergny,M.D., Scherman,D., Demeneix,B. and Behr,J.P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 7297-7301.

135. Kircheis,R., Wightman,L. and Wagner,E. (2001) Design and gene delivery activity of modified polyethylenimines. *Advanced Drug Delivery Reviews*, 53, 341-358.

136. Cheng,K., Ye,Z., Guntaka,R.V. and Mahato,R.I. (2006) Enhanced Hepatic Uptake and Bioactivity of Type alpha1(I) Collagen Gene Promoter-Specific Triplex-Forming Oligonucleotides after Conjugation with Cholesterol. *Journal of Pharmacology and Experimental Therapeutics*, 317, 797-805.

137. Thomas,R.M., Thomas,T., Wada,M., Sigal,L.H., Shirahata,A. and Thomas,T.J. (1999) Facilitation of the Cellular Uptake of a Triplex-Forming Oligonucleotide by Novel Polyamine Analogues: Structure–Activity Relationships. *Biochemistry*, 38, 13328-13337.

138. Rogers,F.A., Manoharan,M., Rabinovitch,P., Ward,D.C. and Glazer,P.M. (2004) Peptide conjugates for chromosomal gene targeting by triplex-forming oligonucleotides. *Nucl. Acids Res.*, 32, 6595-6604.

139. Ye,Z., Cheng,K., Guntaka,R.V. and Mahato,R.I. (2006) Receptor-Mediated Hepatic Uptake of M6P-BSA-Conjugated Triplex-Forming Oligonucleotides in Rats. *Bioconjugate Chem.*, 17, 823-830.

140. Santhakumaran,L.M., Thomas,T. and Thomas,T.J. (2004) Enhanced cellular uptake of a triplex-forming oligonucleotide by nanoparticle formation in the presence of polypropylenimine dendrimers. *Nucl. Acids Res.*, 32, 2102-2112.

141. Brown,P.M. and Fox,K.R. (1996) Nucleosome core particles inhibit DNA triple helix formation. *Biochem. J.*, 319 (Pt 2), 607-611.

142. Brown,P.M., Madden,C.A. and Fox,K.R. (1998) Triple-helix formation at different positions on nucleosomal DNA. *Biochemistry*, 37, 16139-16151.

143. Hojland,T., Kumar,S., Babu,B.R., Umemoto,T., Albaek,N., Sharma,P.K., Nielsen,P.E. and Wengel,J. (2007) LNA (locked nucleic acid) and analogs as triplex-forming oligonucleotides. *Org. Biomol. Chem. Y1 - 2003///*, 5, 2375-2379.

144. Nielsen,P.E. and Egholm,M. (2001) Strand displacement recognition of mixed adenine-cytosine sequences in double stranded DNA by thymine-Guanine PNA (Peptide nucleic acid). *Bioorganic & Medicinal Chemistry*, 9, 2429-2434.

145. Nielsen,P.E., Egholm,M., Berg,R.H. and Buchardt,O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, 254, 1497-1500.

146. Hanvey,J.C., Peffer,N.J., Bisi,J.E., Thomson,S.A., Cadilla,R., Josey,J.A., Ricca,D.J., Hassman,C.F., Bonham,M.A., Au,K.G. *et al.* (1992) Antisense and antigene properties of peptide nucleic acids. *Science*, 258, 1481-1485.

147. Hu,J. and Corey,D.R. (2007) Inhibiting Gene Expression with Peptide Nucleic Acid (PNA) Peptide Conjugates That Target Chromosomal DNA. *Biochemistry*, 46, 7581-7589.

148. Kaihatsu,K., Huffman,K.E. and Corey,D.R. (2004) Intracellular Uptake and Inhibition of Gene Expression by PNAs and PNA-Peptide Conjugates. *Biochemistry*, 43, 14340-14347.

149. Zhilina,Z.V., Ziemba,A.J., Nielsen,P.E. and Ebbinghaus,S.W. (2006) PNA–Nitrogen Mustard Conjugates Are Effective Suppressors of HER-2/neu and Biological Tools for Recognition of PNA/DNA Interactions. *Bioconjugate Chem.*, 17, 214-222.

150. S°rensen,M.D., Meldgaard,M., Rajwanshi,V.K. and Wengel,J. (2000) Branched oligonucleotides containing bicyclic nucleotides as branching points and DNA or LNA as triplex forming branch. *Bioorganic & Medicinal Chemistry Letters*, 10, 1853-1856.

151. Sun,B.W., Babu,B.R., Sorensen,M.D., Zakrzewska,K., Wengel,J. and Sun,J.s. (2004) Sequence and pH Effects of LNA-Containing Triple Helix-Forming Oligonucleotides: Physical Chemistry, Biochemistry, and Modeling Studies. *Biochemistry*, 43, 4160-4169.

152. Brunet,E., Alberti,P., Perrouault,L.c., Babu,R., Wengel,J. and Giovannangeli,C. (2005) Exploring Cellular Activity of Locked Nucleic Acid-modified Triplex-forming Oligonucleotides and Defining Its Molecular Basis. *J. Biol. Chem.*, 280, 20076-20085.

153. Beane,R.L., Ram,R., Gabillet,S., Arar,K., Monia,B.P. and Corey,D.R. (2007) Inhibiting Gene Expression with Locked Nucleic Acids (LNAs) That Target Chromosomal DNA. *Biochemistry*, 46, 7572-7580.

154. Koizumi,M., Morita,K., Daigo,M., Tsutsumi,S., Abe,K., Obika,S. and Imanishi,T. (2003) Triplex formation with 2'-O,4'-C-ethylene-bridged nucleic acids (ENA) having C3'-endo conformation at physiological pH. *Nucl. Acids Res.*, 31, 3267-3273.

155. Levy,O., Ptacin,J.L., Pease,P.J., Gore,J., Eisen,M.B., Bustamante,C. and Cozzarelli,N.R. (2005) Identification of oligonucleotide sequences that direct the movement of the Escherichia coli FtsK translocase. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 17618-17623.

156. Maxwell,A., Burton,N.P. and O'Hagan,N. (2006) High-throughput assays for DNA gyrase and other topoisomerases. *Nucl. Acids Res.*, 34, e104.

157. Arimondo,P.B. and Helene,C. (2001) Design of New Anti-Cancer Agents Based on Topoisomerase Poisons Targeted to Specific DNA Sequences. *Current Medicinal Chemistry -Anti-Cancer Agents*, 1, 219-235.

158. Arimondo,P.B., Boutorine,A., Baldeyrou,B., Bailly,C., Kuwahara,M., Hecht,S.M., Sun,J.S., Garestier,T. and Helene,C. (2002) Design and Optimization of Camptothecin Conjugates of Triple Helix-forming Oligonucleotides for Sequence-specific DNA Cleavage by Topoisomerase I. *J. Biol. Chem.*, 277, 3132-3140.

159. Ito,T., Smith,C.L. and Cantor,C.R. (1992) Sequence-specific DNA purification by triplex affinity capture. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 495-498.

160. Schluep,T. and Cooney,C.L. (1998) Purification of plasmids by triplex affinity interaction. *Nucl. Acids Res.*, 26, 4524-4528.

161. Fox,K.R. (1992) Wrapping of genomic polydA.polydT tracts around nucleosome core particles. *Nucl. Acids Res.*, 20, 1235-1242.

162. Musso,M., Nelson,L.D. and Van Dyke,M.W. (1998) Characterization of Purine-Motif Triplex DNA-Binding Proteins in HeLa Extracts. *Biochemistry*, 37, 3086-3095.

163. Spitzner,J.R., Chung,I.K. and Muller,M.T. (1995) Determination of 5' and 3' DNA Triplex Interference Boundaries Reveals the Core DNA Binding Sequence for Topoisomerase II. *J. Biol. Chem.*, 270, 5932-5943.

164. Firman,K. and Szczelkun,M.D. (2000) Measuring motion on DNA by the type I restriction endonuclease EcoR124I using triplex displacement. *EMBO J*, 19, 2094-2102.

165. McClelland,S.E., Dryden,D.T.F. and Szczelkun,M.D. (2005) Continuous Assays for DNA Translocation Using Fluorescent Triplex Dissociation: Application to Type I Restriction Endonucleases. *Journal of Molecular Biology*, 348, 895-915.

166. Lacroix,L., Lacoste,J., Reddoch,J.F., Mergny,J.L., Levy,D.D., Seidman,M.M., Matteucci,M.D. and Glazer,P.M. (1999) Triplex formation by oligonucleotides

containing 5-(1-propynyl)-2'-deoxyuridine: decreased magnesium dependence and improved intracellular gene targeting. *Biochemistry*, 38, 1893-1901.

167. Gowers,D.M., Bijapur,J., Brown,T. and Fox,K.R. (1999) DNA triple helix formation at target sites containing several pyrimidine interruptions: stabilization by protonated cytosine or 5-(1-propargylamino)dU. *Biochemistry*, 38, 13747-13758.

168. Guianvarc'h,D., Benhida,R., Fourrey,J.L., Maurisse,R. and Sun,J.S. (2001) Incorporation of a novel nucleobase allows stable oligonucleotide-directed triple helix formation at the target sequence containing a purine.pyrimidine interruption. *Chem. Commun. (Camb.)*, 1814-1815.

169. Chandler,S.P. and Fox,K.R. (1993) Triple helix formation at A8XA8.T8YT8. *FEBS Lett.*, 332, 189-192.

170. Chandler,S.P. and Fox,K.R. (1996) Specificity of antiparallel DNA triple helix formation. *Biochemistry*, 35, 15038-15048.

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

172. Griffin,L.C., Kiessling,L.L., Beal,P.A., Gillespie,P. and Dervan,P.B. (1992) Recognition of All 4 Base-Pairs of Double-Helical DNA by Triple-Helix Formation - Design of Nonnatural Deoxyribonucleosides for Pyrimidine.Purine Base Pair Binding. *J. Am. Chem. Soc.*, 114, 7976-7982.

173. Gowers,D.M. and Fox,K.R. (1998) Triple helix formation at (AT)n adjacent to an oligopurine tract. *Nucleic Acids Res.*, 26, 3626-3633.

174. Chandler,S.P. and Fox,K.R. (1995) Extension of DNA triple helix formation to a neighbouring (AT)n site. *FEBS Lett.*, 360, 21-25.

175. Miller,P.S. and Cushman,C.D. (1993) Triplex formation by oligodeoxyribonucleotides involving the formation of XUA triads. *Biochemistry*, 32, 2999-3004.

176. Koshlap,K.M., Gillespie,P., Dervan,P.B. and Feigon,J. (1993) Nonnatural Deoxyribonucleoside-D(3) Incorporated in An Intramolecular Dna Triplex Binds Sequence-Specifically by Intercalation. *J. Am. Chem. Soc.*, 115, 7908-7909.

177. Wang,E., Koshlap,K.M., Gillespie,P., Dervan,P.B. and Feigon,J. (1996) Solution structure of a pyrimidine-purine-pyrimidine triplex containing the sequence-specific intercalating non-natural base D-3. *Journal of Molecular Biology*, 257, 1052-1069.

178. Stilz,H.U. and Dervan,P.B. (1993) Specific recognition of CG base pairs by 2-deoxynebularine within the purine.cntdot.purine.cntdot.pyrimidine triple-helix motif. *Biochemistry*, 32, 2177-2185.

179. Durland,R.H., Rao,T.S., Bodepudi,V., Seth,D.M., Jayaraman,K. and Revankar,G.R. (1995) Azole Substituted Oligonucleotides Promote Antiparallel Triplex Formation at Non-Homopurine Duplex Targets. *Nucl. Acids Res.*, 23, 647-653.

180. Huang,C.Y., Cushman,C.D. and Miller,P.S. (1993) Triplex Formation by An Oligonucleotide Containing N(4)-(3-Acetamidopropyl)Cytosine. *Journal of Organic Chemistry*, 58, 5048-5049.

181. Huang,C.Y., Bi,G.X. and Miller,P.S. (1996) Triplex formation by oligonucleotides containing novel deoxycytidine derivatives. *Nucl. Acids Res.*, 24, 2606-2613.

182. Rusling,D.A., Powers,V.E., Ranasinghe,R.T., Wang,Y., Osborne,S.D., Brown,T. and Fox,K.R. (2005) Four base recognition by triplex-forming oligonucleotides at physiological pH. *Nucleic Acids Res.*, 33, 3025-3032.

183. Wang,Y., Rusling,D.A., Powers,V.E.C., Lack,O., Osborne,S.D., Fox,K.R. and Brown,T. (2005) Stable recognition of TA interruptions by triplex forming oligonucleotides containing a novel nucleoside. *Biochemistry*, 44, 5884-5892.

184. Ranasinghe,R.T., Rusling,D.A., Powers,V.E., Fox,K.R. and Brown,T. (2005) Recognition of CG inversions in DNA triple helices by methylated 3H-pyrrolo[2,3-d]pyrimidin-2(7H)-one nucleoside analogues. *Chem. Commun. (Camb.)*, 2555-2557.

185. Buchini,S. and Leumann,C.J. (2003) Dual recognition of a C-G pyrimidine-purine inversion site: synthesis and binding properties of triplex forming oligonucleotides containing 2 '-aminoethoxy-5-methyl-1H-pyrimidin-2-one ribonucleosides. *Tetrahedron Letters*, 44, 5065-5068.

186. Sollogoub,M., Darby,R.A., Cuenoud,B., Brown,T. and Fox,K.R. (2002) Stable DNA triple helix formation using oligonucleotides containing 2'-aminoethoxy,5-propargylamino-U. *Biochemistry*, 41, 7224-7231.

187. Osborne,S.D., Powers,V.E., Rusling,D.A., Lack,O., Fox,K.R. and Brown,T. (2004) Selectivity and affinity of triplex-forming oligonucleotides containing 2'-aminoethoxy-5-(3-aminoprop-1-ynyl)uridine for recognizing AT base pairs in duplex DNA. *Nucleic Acids Res.*, 32, 4439-4447.

188. Puri,N., Majumdar,A., Cuenoud,B., Miller,P.S. and Seidman,M.M. (2004) Importance of clustered 2 '-O-(2-aminoethyl) residues for the gene targeting activity of triple helix-forming oligonucleotides. *Biochemistry*, 43, 1343-1351.

189. Vasquez,K.M., Wensel,T.G., Hogan,M.E. and Wilson,J.H. (1995) High-Affinity Triple Helix Formation by Synthetic Oligonucleotides at a Site within a Selectable Mammalian Gene. *Biochemistry*, 34, 7243-7251.

190. Horne,D.A. and Dervan,P.B. (1991) Effects of An Abasic Site on Triple Helix Formation Characterized by Affinity Cleaving. *Nucl. Acids Res.*, 19, 4963-4965.

191. Zhou,B.w., Puga,E., Sun,J.s., Garestier,T. and Helene,C. (1995) Stable Triple Helixes Formed by Acridine-Containing Oligonucleotides with Oligopurine Tracts of DNA Interrupted by One or Two Pyrimidines. *J. Am. Chem. Soc.*, 117, 10425-10428.

192. Kukreti,S., Sun,J.S., Garestier,T. and Helene,C. (1997) Extension of the range of DNA sequences available for triple helix formation: stabilization of mismatched triplexes by acridine-containing oligonucleotides. *Nucl. Acids Res.*, 25, 4264-4270.

193. Mayfield,C., Ebbinghaus,S., Gee,J., Jones,D., Rodu,B., Squibb,M. and Miller,D. (1994) Triplex formation by the human Ha-ras promoter inhibits Sp1 binding and in vitro transcription. *J. Biol. Chem.*, 269, 18232-18238.

194. Mayfield,C. and Miller,D. (1994) Effect of abasic linker substitution on triplex formation, Sp1 binding, and specificity in an oligonucleotide targeted to the human Ha-ras promoter. *Nucl. Acids Res.*, 22, 1909-1916.

195. Horne,D.A. and Dervan,P.B. (1990) Recognition of mixed-sequence duplex DNA by alternate-strand triple-helix formation. *J. Am. Chem. Soc.*, 112, 2435-2437.

196. Beal,P.A. and Dervan,P.B. (1992) Recognition of double helical DNA by alternate strand triple helix formation. *J. Am. Chem. Soc.*, 114, 4976-4982.

197. Ono,A., Chen,C.N. and Kan,L.S. (1991) DNA triplex formation of oligonucleotide analogs consisting of linker groups and octamer segments that have opposite sugar-phosphate backbone polarities. *Biochemistry*, 30, 9914-9921.

198. Froehler,B.C., Terhorst,T., Shaw,J.P. and McCurdy,S.N. (1992) Triple-helix formation and cooperative binding by oligodeoxynucleotides with a 3'-3' internucleotide junction. *Biochemistry*, 31, 1603-1609.

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

200. Koh,J.S. and 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.

201. Miller,P.S., Bhan,P., Cushman,C.D. and Trapane,T.L. (1992) Recognition of A Guanine Cytosine Base Pair by 8-Oxoadenine. *Biochemistry*, 31, 6788-6793.

202. Jetter,M.C. and Hobbs,F.W. (1993) 7,8-Dihydro-8-Oxoadenine As A Replacement for Cytosine in the 3rd Strand of Triple Helices - Triplex Formation Without Hypochromicity. *Biochemistry*, 32, 3249-3254.

203. Krawczyk,S.H., Milligan,J.F., Wadwani,S., Moulds,C., Froehler,B.C. and Matteucci,M.D. (1992) Oligonucleotide-Mediated Triple Helix Formation Using An N3-Protonated Deoxycytidine Analog Exhibiting Ph-Independent Binding Within the Physiological Range. *Proceedings of the National Academy of Sciences of the United States of America*, 89, 3761-3764.

204. Brunar,H. and Dervan,P.B. (1996) Sequence composition effects on the stabilities of triple helix formation by oligonucleotides containing N-7-deoxyguanosine. *Nucl. Acids Res.*, 24, 1987-1991.

205. Koshlap,K.M., Schultze,P., Brunar,H., Dervan,P.B. and Feigon,J. (1997) Solution structure of an intramolecular DNA triplex containing an N-7-glycosylated guanine which mimics a protonated cytosine. *Biochemistry*, 36, 2659-2668.

206. Hunziker,J., Priestley,E.S., Brunar,H. and Dervan,P.B. (1995) Design of An N-7-Glycosylated Purine Nucleoside for Recognition of Gc Base-Pairs by Triple-Helix Formation. *J. Am. Chem. Soc.*, 117, 2661-2662.

207. Priestley,E.S. and Dervan,P.B. (1995) Sequence Composition Effects on the Energetics of Triple-Helix Formation by Oligonucleotides Containing A Designed Mimic of Protonated Cytosine. *J. Am. Chem. Soc.*, 117, 4761-4765.

208. Xodo,L.E., Manzini,G., Quadrifoglio,F., Vandermarel,G.A. and Vanboom,J.H. (1991) Effect of 5-Methylcytosine on the Stability of Triple-Stranded Dna - A Thermodynamic Study. *Nucl. Acids Res.*, 19, 5625-5631.

209. Povsic,T.J. and Dervan,P.B. (1989) Triple Helix Formation by Oligonucleotides on Dna Extended to the Physiological Ph Range. *J. Am. Chem. Soc.*, 111, 3059-3061.

210. Xiang,G.B., Bogacki,R. and Mclaughlin,L.W. (1996) Use of a pyrimidine nucleoside that functions as a bidentate hydrogen bond donor for the recognition of isolated or contiguous G-C base pairs by oligonucleotide-directed triplex formation. *Nucl. Acids Res.*, 24, 1963-1970.

211. Xiang,G.B., Soussou,W. and Mclaughlin,L.W. (1994) A New Pyrimidine Nucleoside (m5oxC) for the Ph-Independent Recognition of G-C Base-Pairs by Oligonucleotide-Directed Tripler Formation. *J. Am. Chem. Soc.*, 116, 11155-11156.

212. Berressem,R. and Engels,J.W. (1995) 6-Oxocytidine A Novel Protonated C-Base Analog for Stable Triple-Helix Formation. *Nucl. Acids Res.*, 23, 3465-3472.

213. Xiang,G.B. and Mclaughlin,L.W. (1998) A cytosine analogue containing a conformationally flexible acyclic linker for triplex formation at sites with contiguous G-C base pair. *Tetrahedron*, 54, 375-392.

214. Ono,A., Tso,P.O.P. and Kan,L.S. (1991) Triplex Formation of Oligonucleotides Containing 2'-O-Methylpseudoisocytidine in Substitution for 2'-Deoxycytidine. *J. Am. Chem. Soc.*, 113, 4032-4033.

215. Ono,A., Tso,P.O.P. and Kan,L.S. (1992) Triplex Formation of An Oligonucleotide Containing 2'-O-Methylpseudoisocytidine with A Dna Duplex at Neutral Ph. *Journal of Organic Chemistry*, 57, 3225-3230.

216. Egholm,M., Christensen,L., Dueholm,K.L., Buchardt,O., Coull,J. and Nielsen,P.E. (1995) Efficient Ph-Independent Sequence-Specific Dna-Binding by Pseudoisocytosine-Containing Bis-Pna. *Nucl. Acids Res.*, 23, 217-222.

217. Kuhn,H., Demidov,V.V., Frank-Kamenetskii,M.D. and Nielsen,P.E. (1998) Kinetic sequence discrimination of cationic bis-PNAs upon targeting of double-stranded DNA. *Nucl. Acids Res.*, 26, 582-587.

218. Hildbrand,S., Blaser,A., Parel,S.P. and Leumann,C.J. (1997) 5-substituted 2-aminopyridine C-nucleosides as protonated cytidine equivalents: Increasing efficiency and selectivity in DNA triple-helix formation. *J. Am. Chem. Soc.*, 119, 5499-5511.

219. Hildbrand,S. and Leumann,C. (1996) Enhancing DNA triple helix stability at neutral pH by the use of oligonucleotides containing a more basic deoxycytidine analog. *Angewandte Chemie-International Edition in English*, 35, 1968-1970.

220. Cassidy,S.A., Slickers,P., Trent,J.O., Capaldi,D.C., Roselt,P.D., Reese,C.B., Neidle,S. and Fox,K.R. (1997) Recognition of GC base pairs by triplex forming

oligonucleotides containing nucleosides derived from 2-aminopyridine. *Nucl. Acids Res.*, 25, 4891-4898.

221. Lyamichev,V.I., Voloshin,O.N., Frankkamenetskii,M.D. and Soyfer,V.N. (1991) Photofootprinting of Dna Triplexes. *Nucl. Acids Res.*, 19, 1633-1638.

222. Blume,S.W., Lebowitz,J., Zacharias,W., Guarcello,V., Mayfield,C.A., Ebbinghaus,S.W., Bates,P., Jones Jr,D.E., Trent,J., Vigneswaran,N. *et al.* (1999) The integral divalent cation within the intermolecular purine*purine. pyrimidine structure: a variable determinant of the potential for and characteristics of the triple helical association. *Nucl. Acids Res.*, 27, 695-702.

223. Hartwig,A. (2001) Role of magnesium in genomic stability. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 475, 113-121.

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

225. Cuenoud,B., Casset,F., Husken,D., Natt,F., Wolf,R.M., Altmann,K.H., Martin,P. and Moser,H.E. (1998) Dual recognition of double-stranded DNA by 2 '-aminoethoxy-modified oligonucleotides. *Angewandte Chemie-International Edition*, 37, 1288-1291.

226. Rajeev,K.G., Jadhav,V.R. and Ganesh,K.N. (1997) Triplex formation at physiological pH: Comparative studies on DNA triplexes containing 5-Me-dC tethered at N-4 with spermine and tetraethyleneoxyamine. *Nucl. Acids Res.*, 25, 4187-4193.

227. Chaturvedi,S., Horn,T. and Letsinger,R.L. (1996) Stabilization of triple-stranded oligonucleotide complexes: Use of probes containing alternating phosphodiester and stereo-uniform cationic phosphoramidate linkages. *Nucl. Acids Res.*, 24, 2318-2323.

228. Colocci,N. and Dervan,P.B. (1994) Cooperative Binding of 8-Mer Oligonucleotides Containing 5-(1-Propynyl)-2'-Deoxyuridine to Adjacent Dna Sites by Triple-Helix Formation. *J. Am. Chem. Soc.*, 116, 785-786.

229. Michel,J., Toulme,J.J., Vercauteren,J. and Moreau,S. (1996) Quinazoline-2,4(1H,3H)-dione as a substitute for thymine in triple- helix forming oligonucleotides: a reassessment. *Nucl. Acids Res.*, 24, 1127-1135.

230. Escude,C., Giovannangeli,C., Sun,J.S., Lloyd,D.H., Chen,J.K., Gryaznov,S.M., Garestier,T. and Helene,C. (1996) Stable triple helices formed by oligonucleotide N3'-->P5' phosphoramidates inhibit transcription elongation. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 4365-4369.

231. Giovannangeli,C., Perrouault,L., Escudθ,C., Gryaznov,S. and HθlΦne,C. (1996) Efficient Inhibition of Transcription Elongationin vitroby Oligonucleotide Phosphoramidates Targeted to Proviral HIV DNA. *Journal of Molecular Biology*, 261, 386-398.

232. Latimer,L.J.P., Hampel,K. and Lee,J.S. (1989) Synthetic repeating sequence DNAs containing plsosphorothioates: nuclease sensitivity and triplex formation. *Nucl. Acids Res.*, 17, 1549-1561.

233. Kibler-Herzog,L., Kell,B., Zon,G., Shinozuka,K., Mizan,S. and Wilson,W.D. (1990) Sequence dependent effects in methylphosphonate deoxyribonucleotide double and triple helical complexes. *Nucl. Acids Res.*, 18, 3545-3555.

234. Kibler-Herzog,L., Zon,G., Uznanski,B., Whittier,G. and Wilson,W.D. (1991) Duplex stabilities of phosphorothioate, methylphosphonate, and RNA analogs of two DNA 14-mers. *Nucl. Acids Res.*, 19, 2979-2986.

235. Egholm,M., Buchardt,O., Nielsen,P.E. and Berg,R.H. (1992) Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone. *J. Am. Chem. Soc.*, 114, 1895-1897.

236. Nielsen,P.E. and Haaima,G. (1997) Peptide nucleic acid (PNA). A DNA mimic with a pseudopeptide backbone. *Chemical Society Reviews*, 73-78.

237. Wittung,P., Nielsen,P. and Norden,B. (1997) Extended DNA-Recognition Repertoire of Peptide Nucleic Acid (PNA): PNA:DNA Triplex Formed with Cytosine-Rich Homopyrimidine PNA. *Biochemistry*, 36, 7973-7979.

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

239. Asensio,J.L., Brown,T. and Lane,A.N. (1999) Solution conformation of a parallel DNA triple helix with 5' and 3' triplex-duplex junctions. *Structure*, 7, 1-11.

240. Blommers,M.J.J., Natt,F., Jahnke,W. and Cuenoud,B. (1998) Dual Recognition of Double-Stranded DNA by 2'-Aminoethoxy-Modified Oligonucleotides: The Solution Structure of an Intramolecular Triplex Obtained by NMR Spectroscopy. *Biochemistry*, 37, 17714-17725.

241. Han,H. and Dervan,P.B. (1993) Sequence-specific recognition of double helical RNA and RNA.DNA by triple helix formation. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 3806-3810.

242. Shimizu,M., Konishi,A., Shimada,Y., Inoue,H. and Ohtsuka,E. (1992) Oligo(2'-O-methyl)ribonucleotides Effective probes for duplex DNA. *Febs Letters*, 302, 155-158.

243. Rusling,D.A., Broughton-Head,V.J., Tuck,A., Khairallah,H., Osborne,S.D., Brown,T. and Fox,K.R. (2008) Kinetic studies on the formation of DNA triplexes containing the nucleoside analogue 2 '-O-(2-aminoethyl)-5-(3-amino-1-propynyl)uridine. *Org. Biomol. Chem. Y1 - 2003///*, 6, 122-129.

244. Petersen,M. and Wengel,J. (2003) LNA: a versatile tool for therapeutics and genomics. *Trends in Biotechnology*, 21, 74-81.

245. Torigoe,H., Hari,Y., Sekiguchi,M., Obika,S. and Imanishi,T. (2001) 2'-O,4'-C-Methylene Bridged Nucleic Acid Modification Promotes Pyrimidine Motif Triplex DNA Formation at Physiological pH. *J. Biol. Chem.*, 276, 2354-2360.

246. Wahlestedt,C., Salmi,P., Good,L., Kela,J., Johnsson,T., H+Âkfelt,T., Broberger,C., Porreca,F., Lai,J., Ren,K. *et al.* (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 5633-5638.

247. Barawkar,D.A., Rajeev,K.G., Kumar,V.A. and Ganesh,K.N. (1996) Triplex formation at physiological pH by 5-Me-dC-N-4-(spermine) [X] oligodeoxynucleotides: Non protonation of N3 in X of X*G:C triad and effect of base mismatch ionic strength on triplex stabilities. *Nucl. Acids Res.*, 24, 1229-1237.

248. Volker,J. and Klump,H.H. (1994) Electrostatic Effects in Dna Triple Helices. *Biochemistry*, 33, 13502-13508.

249. Sollogoub,M., Dominguez,B., Fox,K.R. and Brown,T. (2000) Synthesis of a novel bis-amino-modified thymidine monomer for use in DNA triplex stabilisation. *Physical Chemistry Chemical Physics*, 2, 2315-2316.

250. Brazier,J.A., Shibata,T., Townsley,J., Taylor,B.F., Frary,E., Williams,N.H. and Williams,D.M. (2005) Amino-functionalized DNA: the properties of C5-amino-alkyl substituted 2'-deoxyuridines and their application in DNA triplex formation. *Nucleic Acids Res.*, 33, 1362-1371.

251. Froehler,B.C., Wadwani,S., Terhorst,T.J. and Gerrard,S.R. (1992) Oligodeoxynucleotides Containing C-5 Propyne Analogs of 2'-Deoxyuridine and 2'-Deoxycytidine. *Tetrahedron Letters*, 33, 5307-5310.

252. Phipps,A.K., Tarkoy,M., Schultze,P. and Feigon,J. (1998) Solution structure of an intramolecular DNA triplex containing 5-(1-propynyl)-2 '-deoxyuridine residues in the third strand. *Biochemistry*, 37, 5820-5830.

253. Flanagan,W.M., Kothavale,A. and Wagner,R.W. (1996) Effects of oligonucleotide length, mismatches and mRNA levels on C-5 propyne-modified antisense potency. *Nucl. Acids Res.*, 24, 2936-2941.

254. Moulds,C., Lewis,J.G., Froehler,B.C., Grant,D., Huang,T., Milligan,J.F., Matteucci,M.D. and Wagner,R.W. (1995) Site and Mechanism of Antisense Inhibition by C-5 Propyne Oligonucleotides. *Biochemistry*, 34, 5044-5053.

255. Heystek,L.E., Zhou,H.q., Dande,P. and Gold,B. (1998) Control over the Localization of Positive Charge in DNA: The Effect on Duplex DNA and RNA Stability. *J. Am. Chem. Soc.*, 120, 12165-12166.

256. Brennan,L., Peng,G., Srinivasan,N., Fox,K. and Brown,T. (2007) 2'-O-Dimethylaminoethoxyuridine and 5-Dimethylaminopropargyl Deoxyuridine for at Base Pair Recognition in Triple Helices. *Nucleosides, Nucleotides & Nucleic Acids*, 26, 1283-1286.

257. Rusling,D.A., Peng,G., Srinivasan,N., Fox,K.R. and Brown,T. (2009) DNA triplex formation with 5-dimethylaminopropargyl deoxyuridine. *Nucl. Acids Res.*, 37, 1288-1296.

258. Puri,N., Majumdar,A., Cuenoud,B., Natt,F., Martin,P., Boyd,A., Miller,P.S. and Seidman,M.M. (2002) Minimum Number of 2-O-(2-Aminoethyl) Residues Required for Gene Knockout Activity by Triple Helix Forming Oligonucleotides. *Biochemistry*, 41, 7716-7724.

259. Asensio,J.L., Carr,R., Brown,T. and Lane,A.N. (1999) Conformational and Thermodynamic Properties of Parallel Intramolecular Triple Helices Containing a DNA, RNA, or 2'-OMeDNA Third Strand. *J. Am. Chem. Soc.*, 121, 11063-11070.

260. Escudee,C., Franccois,J.C., Sun,J.s., Ott,G., Sprinzl,M., Garestier,T. and ne,J.C. (1993) Stability of triple helices containing RNA and DNA strands: experimental and molecular modeling studies. *Nucl. Acids Res.*, 21, 5547-5553.

261. Kundu,M., Nagatsugi,F., Majumdar,A., Miller,P.S. and Seidman,M.M. (2001) Enhancement and Inhibition by 2'-O-Hydroxyethyl Residues of Gene Targeting Mediated by Triple Helix Forming Oligonucleotides. *Nucleosides, Nucleotides and Nucleic Acids*, 22, 1927-1938.

262. Puri,N., Majumdar,A., Cuenoud,B., Natt,F., Martin,P., Boyd,A., Miller,P.S. and Seidman,M.M. (2001) Targeted Gene Knockout by 2'-O-Aminoethyl Modified Triplex Forming Oligonucleotides. *J. Biol. Chem.*, 276, 28991-28998.

263. Griffey,R.H., Monia,B.P., Cummins,L.L., Freier,S., Greig,M.J., Guinosso,C.J., Lesnik,E., Manalili,S.M., Mohan,V., Owens,S. *et al.* (1996) 2'-O-Aminopropyl Ribonucleotides: A Zwitterionic Modification That Enhances the Exonuclease Resistance and Biological Activity of Antisense Oligonucleotides. *Journal of Medicinal Chemistry*, 39, 5100-5109.

264. Stutz,A.M., Hoeck,J., Natt,F., Cuenoud,B. and Woisetschlager,M. (2001) Inhibition of Interleukin-4- and CD40-induced IgE Germline Gene Promoter Activity by 2'-Aminoethoxy-modified Triplex-forming Oligonucleotides. *J. Biol. Chem.*, 276, 11759-11765.

265. Alam,M., Majumdar,A., Thazhathveetil,A.K., Liu,S.T., Liu,J.L., Puri,N., Cuenoud,B., Sasaki,S., Miller,P.S. and Seidman,M.M. (2007) Extensive Sugar Modification Improves Triple Helix Forming Oligonucleotide Activity in Vitro but Reduces Activity in Vivo. *Biochemistry*, 46, 10222-10233.

266. Rusling,D.A., Brown,T. and Fox,K.R. (2006) DNA triple-helix formation at target sites containing duplex mismatches. *Biophysical Chemistry*, 123, 134-140.

267. Rusling,D.A., Le,S.L., Powers,V.E., Broughton-Head,V.J., Booth,J., Lack,O., Brown,T. and Fox,K.R. (2005) Combining nucleoside analogues to achieve recognition of oligopurine tracts by triplex-forming oligonucleotides at physiological pH. *FEBS Lett.*, 579, 6616-6620.

268. Carlomagno,T., Blommers,M.J.J., Meiler,J., Cuenoud,B. and Griesinger,C. (2001) Determination of Aliphatic Side-Chain Conformation Using Cross-Correlated Relaxation: Application to an Extraordinarily Stable 2'-aminoethoxy-Modified Oligonucleotide Triplex. *J. Am. Chem. Soc.*, 123, 7364-7370.

269. Chandler,S.P., Strekowski,L., Wilson,W.D. and Fox,K.R. (1995) Footprinting studies on ligands which stabilize DNA triplexes: effects on stringency within a parallel triple helix. *Biochemistry*, 34, 7234-7242.

270. Tung,C.H., Breslauer,K.J. and Stein,S. (1993) Polyamine-Linked Oligonucleotides for Dna Triple-Helix Formation. *Nucl. Acids Res.*, 21, 5489-5494.

271. Barawkar,D.A., Kumar,V.A. and Ganesh,K.N. (1994) Triplex Formation at Physiological pH by Oligonucleotides Incorporating 5-Me-dC-(N4-Spermine). *Biochemical and Biophysical Research Communications*, 205, 1665-1670.

272. Hampel,K.J., Crosson,P. and Lee,J.S. (1991) Polyamines Favor Dna Triplex Formation at Neutral Ph. *Biochemistry*, 30, 4455-4459.

273. Thomas,T. and Thomas,T.J. (1993) Selectivity of Polyamines in Triplex Dna Stabilization. *Biochemistry*, 32, 14068-14074.

274. Mergny,J.L., Duval-Valentin,G., Nguyen,C.H., Perrouault,L., Faucon,B., Rougee,M., Montenay-Garestier,T., Bisagni,E. and Helene,C. (1992) Triple Helix-Specific Ligands. *Science*, 256, 1681-1684.

275. Pilch,D.S., Martin,M.T., Nguyen,C.H., Sun,J.s., Bisagni,E., Garestier,T. and Helene,C. (1993) Self-association and DNA-binding properties of two triple helix-specific ligands: comparison of a benzo[e]pyridoindole and a benzo[g]pyridoindole. *J. Am. Chem. Soc.*, 115, 9942-9951.

276. Escude,C., Nguyen,C.H., Mergny,J.L., Sun,J.s., Bisagni,E., Garestier,T. and Helene,C. (1995) Selective Stabilization of DNA Triple Helixes by Benzopyridoindole Derivatives. *J. Am. Chem. Soc.*, 117, 10212-10219.

277. Grigoriev,M., Praseuth,D., Robin,P., Hemar,A., Saison-Behmoaras,T., utry-Varsat,A., Thuong,N.T., Helene,C. and Harel-Bellan,A. (1992) A triple helix-forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor via inhibition of NF kappa B binding to interleukin-2 receptor alpha-regulatory sequence. *J. Biol. Chem.*, 267, 3389-3395.

278. Stonehouse,T.J. and Fox,K.R. (1994) DNase I footprinting of triple helix formation at polypurine tracts by acridine-linked oligopyrimidines: stringency, structural changes and interaction with minor groove binding ligands. *Biochim. Biophys. Acta*, 1218, 322-330.

279. Moraru-Allen,A.A., Cassidy,S., sensio Alvarez,J.L., Fox,K.R., Brown,T. and Lane,A.N. (1997) Coralyne has a preference for intercalation between TA.T triples in intramolecular DNA triple helices. *Nucleic Acids Res.*, 25, 1890-1896.

280. Filichev,V.V. and Pedersen,E.B. Intercalating nucleic acids (INAs) with insertion of N-(pyren-1-ylmethyl)-(3R,4R)-4-(hydroxymethyl)pyrrolidin-3-ol. DNA (RNA) duplex and DNA three-way junction stabilities. *Org. Biomol. Chem. Y1 - 2003///*, 1, 100-103.

281. Orson,F.M., Klysik,J., Bergstrom,D.E., Ward,B., Glass,G.A., Hua,P. and Kinsey,B.M. (1999) Triple helix formation: binding avidity of acridine-conjugated AG motif third strands containing natural, modified and surrogate bases opposed to pyrimidine interruptions in a polypurine target. *Nucl. Acids Res.*, 27, 810-816.

282. Wilson,W.D., Tanious,F.A., Mizan,S., Yao,S., Kiselyov,A.S., Zon,G. and Strekowski,L. (1993) DNA triple-helix specific intercalators as antigene enhancers: Unfused aromatic cations. *Biochemistry*, 32, 10614-10621.

283. Cassidy,S.A., Strekowski,L. and Fox,K.R. (1996) DNA sequence specificity of a naphthylquinoline triple helix-binding ligand. *Nucleic Acids Res.*, 24, 4133-4138.

284. Cassidy,S.A., Strekowski,L., Wilson,W.D. and Fox,K.R. (1994) Effect of a triplex-binding ligand on parallel and antiparallel DNA triple helices using short unmodified and acridine-linked oligonucleotides. *Biochemistry*, 33, 15338-15347.

285. Keppler,M.D., McKeen,C.M., Zegrocka,O., Strekowski,L., Brown,T. and Fox,K.R. (1999) DNA triple helix stabilisation by covalent attachment of a triplex-specific ligand. *Biochim. Biophys. Acta*, 1447, 137-145.

286. Keppler,M.D., Neidle,S. and Fox,K.R. (2001) Stabilisation of TG- and AG-containing antiparallel DNA triplexes by triplex-binding ligands. *Nucleic Acids Res.*, 29, 1935-1942.

287. Gaied,N.B., Zhao,Z., Gerrard,S.R., Fox,K.R. and Brown,T. (2009) Potent triple helix stabilization by 5', 3' - modified triplex-forming oligonucleotides. *ChemBioChem*, 10, 1839-1851.

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

289. Keppler,M.D., Read,M.A., Perry,P.J., Trent,J.O., Jenkins,T.C., Reszka,A.P., Neidle,S. and Fox,K.R. (1999) Stabilization of DNA triple helices by a series of mono- and disubstituted amidoanthraquinones. *Eur. J. Biochem.*, 263, 817-825.

290. Haq,I., Ladbury,J.E., Chowdhry,B.Z. and Jenkins,T.C. (1996) Molecular Anchoring of Duplex and Triplex DNA by Disubstituted Anthracene-9,10-diones: Calorimetric, UV Melting, and Competition Dialysis Studies. *J. Am. Chem. Soc.*, 118, 10693-10701.

291. Mergny,J.L., Sun,J.S., Rougee,M., Montenaygarestier,T., Barcelo,F., Chomilier,J. and Helene,C. (1991) Sequence Specificity in Triple-Helix Formation - Experimental and Theoretical-Studies of the Effect of Mismatches on Triplex Stability. *Biochemistry*, 30, 9791-9798.

292. Roberts,R.W. and Crothers,D.M. (1991) Specificity and stringency in DNA triplex formation. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 9397-9401.

293. Singleton,S.F. and Dervan,P.B. (1992) Influence of pH on the Equilibrium Association Constants for Oligodeoxyribonucleotide-Directed Triple Helix Formation at Single Dna Sites. *Biochemistry*, 31, 10995-11003.

294. Rusling,D.A. DNA Recognition by Triple Helix Formation. 2006.

295. Galas,D.J. and Schmitz,A. (1978) Dnaase Footprinting - Simple Method for Detection of Protein-Dna Binding Specificity. *Nucl. Acids Res.*, 5, 3157-3170.

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

297. Hardenbol,P., Wang,J.C. and Van Dyke,M.W. (1997) Identification of Preferred Distamycin-DNA Binding Sites by the Combinatorial Method REPSA. *Bioconjugate Chem.*, 8, 617-620.

298. Shen,J., Wang,J.C. and Van Dyke,M.W. (2001) Identification of preferred actinomycin-DNA binding sites by the combinatorial method REPSA. *Bioorganic & Medicinal Chemistry*, 9, 2285-2293.

299. Van Dyke,M.W., Van Dyke,N. and Sunavala-Dossabhoy,G. (2007) REPSA: General combinatorial approach for identifying preferred ligand-DNA binding sequences. *Methods*, 42, 118-127.

300. Cardew,A.S. and Fox,K.R. (2010) DNase I Footprinting. In Fox,K.R. (ed.), *Drug-DNA Interaction Protocols*. Humana Press, pp. 153-172.