ENHANCED TARGETING OF TRIPLEX FORMING OLIGONUCLEOTIDES

The University of Waikato
Te Whare Wānanga o Waikato

A thesis submitted in partial fulfilment of
the requirements for the Degree of
Doctor of Philosophy in Biological Sciences
at the University of Waikato
by

LINDA MARIE PETERS

JULY 1999

ABSTRACT

Oligonucleotides offer the potential to inhibit the expression of specific genes within living cells by binding to the major groove of the DNA helix in a sequence-specific manner to form a DNA triple helix. The ultimate application of such oligonucleotides is to act as selective therapeutic tools to suppress transcription, inhibit replication and induce site specific mutations in mammalian genes. However, to date, most triplex experiments have utilised simple *in vitro* systems and it has proved rather more difficult to demonstrate selective actions *in vivo*.

In this study, a series of psoralen-conjugated oligonucleotides (10-17 nt) were designed to target specific bovine repetitive sequences in the male Y chromosome. Repetitive Y chromosome sequences described by Reed *et al.*, 1995 and Grobet *et al.*, 1996 were investigated for efficient DNA triplex formation. The rationale in this approach was that the presence of several hundred copies rather than just one target per genome would greatly enhance the efficiency of DNA triplex formation *in vivo*. Psoralen-conjugated oligonucleotides designed to these sequences were screened by performing DNA triplex binding assays and analysing photoadduct formation by electrophoretic mobility shift assays (EMSA).

Six psoralen-conjugated oligonucleotides were found to selectively and efficiently intercalate to their bovine dsDNA target sequences (38-49 bp) *in vitro* after irradiation with UV light (360 nm), forming covalent DNA photoadducts. Stable photoadducts were also detected when these oligonucleotides were targeted to larger dsDNA targets, ranging from 0.1 to 3 Kb in size. In general, a small UV dose (0.03 J/cm^2) was sufficient to form monoadducts, with crosslinks forming after fifteen minutes of UV irradiation (1 J/cm^2). The efficiency of bovine monoadduct formation was similar to that in the well-publicised *sup*F DNA triplex system.

PCR and Southern blot analyses also showed that the repetitive Y-chromosomal DNA region (100-400 bp) was indeed male specific and present in multiple copies.

However, when Southern blot analyses were carried out using oligonucleotide probes (16-49 nt) which corresponded to the triplex target sequence within this region, relatively weak hybridisation was observed, suggesting that these specific sequences were not present in such high copies. Although this data indicates that the male specific repeat regions must contain segments of consensual rather than exact sequences, high quality non-degenerate DNA sequencing data in the triplex targeting region was obtained from the male specific PCR products. Clearly, the exact nature of these male specific repeats requires further investigation to establish the effects of targeting psoralen-conjugated oligonucleotides to male genomic DNA as opposed to female genomic DNA.

ACKNOWLEDGEMENTS

I'd like to take the opportunity to thank the following people for assisting me throughout my thesis and keeping my spirits high when things got tough:

- Mum, Dad, Helen, Steven and Ma. The food parcels, phone calls, e-mails and letters never went amiss.

- Dick Wilkins (PhD supervisor); your advice, proofreading skills and brain storming ideas were greatly appreciated.

- Members of C.203 Laboratory (1996-1999); Quanah Hudson, Marcel Dinger, Seumas McCroskery, Kathy Speed, Jo Haines, Tao Xie, Greg Jacobson, Greg Baillie, Julie Ayton, Mahina Gunatillake, Anandan Manicks, and Shona Duncan.

- Raewyn Towers, Bridget Douglas, Dr Ray Cursons and Dr Dave Musgrave; your proofreading skills were much appreciated.

- Funding sponsor: New Zealand Marsden Fund (Grant N# VOW52)

- Funding contributions from NZ Marsden Fund, Biological Sciences (Waikato University), Basden Fund, and NZ Biochemistry & Molecular Biology Society gave me the opportunity to present posters at the 48th American society of Genetics meeting in Denver (1998) and also at the Queenstown Molecular Biology Meeting, NZ (1998).

- Dr Helen Davy, Michael McLachlan and Brett Langely (AgResearch, Hamilton).

- Vivienne Robson, Mary Dalbeth and Lisa Holmes (SBS Secretaires).

- My flatmates, especially Mark Manhire and Anke Schirp.

- My Nana and Aunty Judy who passed away while I was completing this research thesis. My thoughts will always be with you.

- Finally, a special thank you to Seumas McCroskery, the guy who sat next to me in my lab and never said a word (only at the beginning).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

α	Alpha
β	Beta
°C	Degree Celsius
$\equiv$	Equivalent
γ	Gamma
$>$	Greater than
λ	Lambda
$<$	Less than
μF	Microfarads
μL	Microlitre (10^{-6} L)
μm	Micrometre
μM	Micromolar
Ω	Ohms
ϕ	Phi
π	Pi
$\therefore$	Therefore
A	Adenine
Ac	Acetate
AML	Acute Myelogenous Leukemia
aprt	Adenine phosphoribosyltransferase
APS	Ammonium persulphate
ATP	Adenine triphosphate
BePi	Benzopyridoindole
Bov	Bovine
bp	Base pairs
BPB	Bromophenol blue
BPQ	Benzopyridoquionoxaline
C	Cytosine
$CaCl_2$	Calcium chloride
CAP	Calf Intestinal Alkaline Phosphatase
Ci	Curie
C&G	Church & Gilbert
CMV	Cytomegalovirus
CsCl	Caesium chloride
CTP	Cytosine triphosphate
DHFR	Dihydrofolate reductase
dH_20	Distilled water

dNTP	Deoxynucleotide
dsDNA	Double stranded DNA
DTE	Dithioerythritol
DTT	Dithiothreitol
Du	Duplex
EDTA	Ethylenediamine Tetraacetic Acid
EMSA	Electrophoretic Mobility Shift Assay
EtBr	Ethidium bromide
F	Forward
FACS	Fluorescence-activated cell sorting
FC	Final concentration
FTIR	Fourier Transform Infrared Spectroscopy
g	Relative Centrifugal Field
G	Guanine
ΔG	Free energy
gDNA	Genomic DNA
GTP	Guanine triphosphate
H	Hydrogen
HIV	Human Immunodeficiency Virus
Hprt	Hypoxanthine phosphoribosyltransferase
IPTG	Isopropyl-β-D-thiogalactoside
ISC	Intersystem crosslinking
J	Joule
K_d	Equilibrium dissociation constant
K^+	Potassium ions
KAc	Potassium acetate
Kb	Kilo base pairs
kV	Kilo volts
LB	Luria Broth
LMP	Low melting point
mμ	Millimicron (≡ nanometre, nm)
M^+	Monovalent ions
MA	Monoadduct
MgAc	Magnesium acetate
$MgCl_2$	Magnesium chloride
mL	Millilitre (10^{-3} L)
MOP	Methoxypsoralen
mRNA	Messenger RNA
NaCl	Sodium chloride

NaOH	Sodium hydroxide
NER	Nuclear excision repair
NH_4Ac	Ammonium acetate
nm	Nanometre
NMR	Nuclear Magnetic Resonance
nts	Nucleotides
oligo	Oligonucleotide
P	Phosphate
P≈	Psoralen
pers. comm	Personal communication
Pr	Primer
PAGE	Polyacrylamide Gel Electrophoresis
PEG	Polyethylene glycol
PCR	Polymerase chain reaction
pK_a	Equilibrium constant for an acid
PNA	Peptide Nucleic Acids
Pso	Psoralen
PUVA	Psoralen Ultraviolet A
R	Reverse
RE	Restriction enzyme
RT	Room temperature
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
sry	Sex Determining Region Y
ssDNA	Single stranded DNA
T	Thymine
TAE	Tris-acetate
TBE	Tris-borate
TE	Tris-EDTA
TEMED	N, N, N', N'-tetramethylethylenediamine
TEN	Tris-EDTA-NaCl
TFO	Triplex forming oligonucleotide
T_M	Melting temperature
TMP	4, 5', 8 trimethylpsoralen
TPA	Target Protection Assay
TSPY	Testis Specific Protein Y
TTP	Thymine triphosphate
U	Uracil or Units
UV	Ultraviolet

V	Volts
v/v	Volume per volume
W	Watts
w/v	Weight per volume
XC	Xylene Cyanole FF
Xgal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
XL	Crosslink

There must be a beginning of any great matter, but the continuing unto the end until it be thoroughly finished yields the true glory.

Francis Drake, 1587

Chapter 1

LITERATURE REVIEW

1.1 INTRODUCTION

Oligonucleotides are short, synthetic single strands of DNA or RNA, traditionally 15-30 bases long, which are complementary to specific messenger RNA (mRNA) or DNA sequences. They offer the potential to directly inhibit the expression of specific genes within living cells. Certain criteria must be met in order to use oligonucleotides to inhibit gene expression. The oligonucleotides must be easily synthesised, be stable *in vivo*, be able to enter and be retained by the target cell and interact with their cellular targets (mRNA, premRNA, genomic DNA) without non-sequence specific interactions with other macromolecules (Stein & Cheng, 1993).

The inhibition of gene expression using oligonucleotides is currently achieved by one of three strategies; antisense, antigene or aptamer technologies (Figure 1.1). Although in theory, any gene can be targeted by oligonucleotides, a number of considerations play a role in the actual target selection. These are the half life of the protein, the unique biological role of the protein, the existence of closely related proteins capable of carrying out similar functions and the levels of mRNA present.

This *Doctor of Philosophy* research thesis was based on using the antigene strategy to target psoralen-conjugated oligonucleotides to dsDNA targets, forming stable photoadducts upon UV irradiation (360 nm).

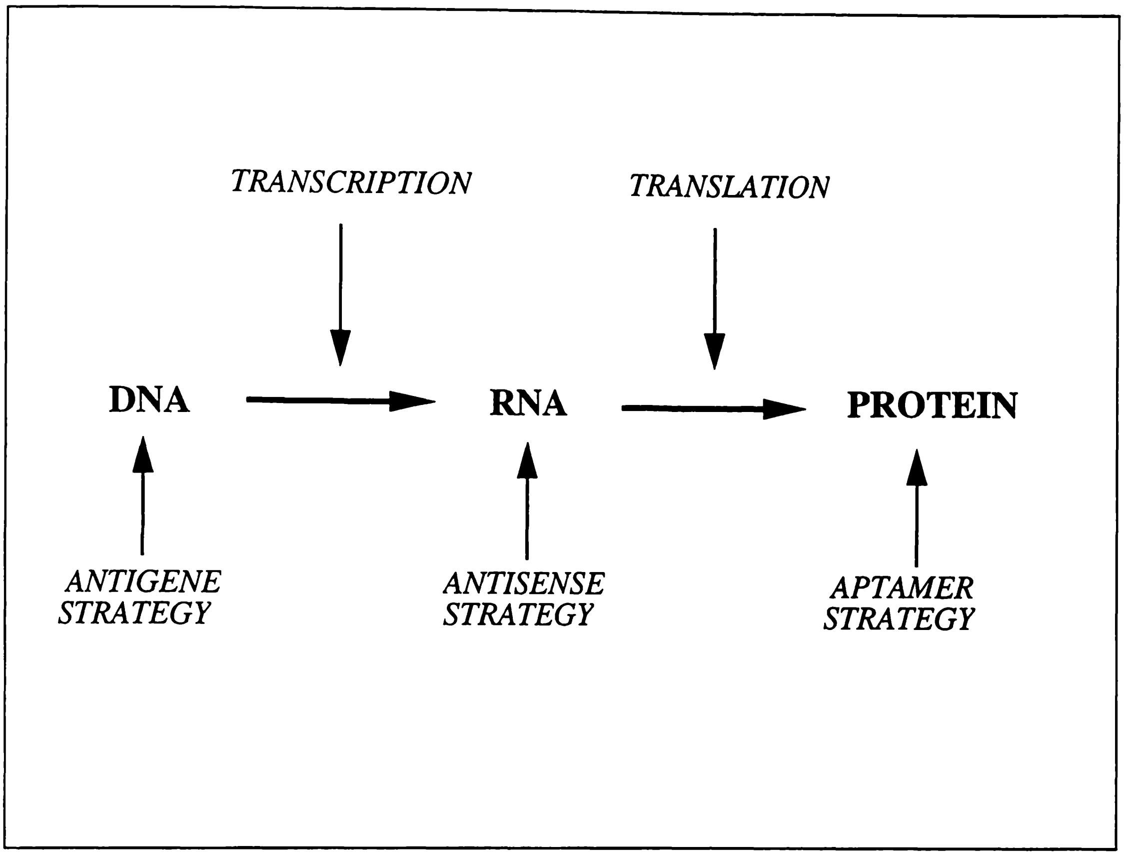


Figure 1.1: A Flow diagram showing the inhibition of gene expression using oligonucleotides.

1.1.1 ANTISENSE STRATEGY

The antisense strategy involves binding antisense oligonucleotides to ssDNA (the sense strand) or RNA via Watson-Crick binding, or using molecules that bind to and cleave the target RNA, such as bioengineered ribozymes. In most cases, antisense oligonucleotides are targeted to part of the ssDNA or promoter region of RNA. Gene expression is inhibited because transcription factors cannot bind to the promoter and initiate transcription or because the mRNA is degraded by the induction of RNase action. Any part of the mRNA molecule can be targeted, for example the 5' and 3' untranslated regions, the start codon, or exon and intron boundaries. In practise, no accurate method has been yet devised to predict which oligonucleotide will most effectively achieve the desired effect.

Zamecnik and Stephenson first demonstrated that synthetic oligonucleotides could block gene expression in 1978 by using a 13 nucleotide (nt) strand of DNA (anti-sense strand) complementary to the RNA (sense strand) of the Rous sarcoma virus causing an enzyme to destroy the mRNA.

Antisense therapy is currently being explored for the treatment of medical disorders such as hypertension, cardiovascular disease, and leukaemia. It is also potentially an attractive treatment for many viral diseases such as Human Immunodeficiency Virus (HIV), Cytomegalovirus (CMV) and Human Papillamovirus as they have genomes which are distinct from that of the host cell and thus offer unique DNA targets. (Wagner & Flanagan, 1997). Gene inhibition by the antisense strategy has been investigated in a variety of genes such as the neuropeptide YY1 receptor (Hanze *et al.*, 1997), *cdc*2 (Furukawa *et al.*, 1990), rabbit *β*-globin (Blake *et al.*, 1985), and a number of cancer-related targets such as *myc* (Cooney *et al.*, 1988) and *bcl*-2 (Olivas & Maher, 1996).

The first clinical trial using the antisense strategy was conducted in 1994 by Bayever *et al.*, for Acute Myelogenous Leukemia (AML). However, only five AML patients were tested with a 20 nt antisense oligonucleotide (OL(1)p53) that was complementary to a portion of exon 10 of p53 mRNA. Western blot analysis showed decreased p53 levels in leukemic cells. Further clinical trials showed that OL(1)p53 could be administered systematically to sixteen AML patients for ten days without complications and degradation (Bishop *et al.*, 1996). Phase 3 clinical trials are currently in progress for a number of medical disorders. Second generation antisense drugs are also being developed (Wagner & Flanagan, 1997).

Various problems with the antisense strategy have been encountered, such as the degradation of oligonucleotides by DNAses which are ubiquitous in the body (both extra and intracellular), the large dose required for a therapeutic response and thus the large scale synthesising equipment needed to produce the compound, the difficulty in directing oligonucleotides to particular cells, and the need for parental administration (Askari *et al.*, 1996).

1.1.2 APTAMER STRATEGY

Aptamers are oligonucleotide ligands that carry molecule-specific binding sites. They bind tightly and specifically to a variety of proteins and small molecules and directly interfere with their function (Kenan *et al.*, 1994). They can be targeted to proteins involved at any step of gene expression. Helene *et al.*, 1994 have shown that proteins involved in several stages of gene expression can be targeted by aptamers.

A number of aptamers have been described that can bind to proteins encoded by the HIV virus. Hence they have the potential to act as therapeutically useful "decoys" (Conrad *et al.*, 1996).

1.1.3 ANTIGENE STRATEGY

The antigene strategy involves using triplex forming oligonucleotides (TFO) which bind in the major groove of the double stranded DNA helix to form a triple helix, more commonly referred to as a triplex (Figure 1.2). DNA triplexes can be formed within a single DNA molecule (intramolecular triplex) or by different molecules (intermolecular triplex) such as a TFO binding to a dsDNA target.

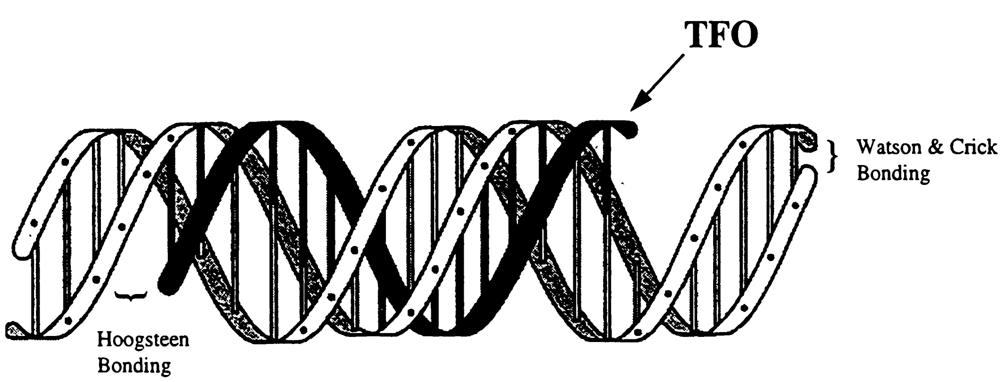


Figure 1.2: A schematic representation of a DNA triple helix showing the third strand (darker) binding to a DNA duplex in the major groove (Kool, 1996).

DNA triplexes may also be formed by a strand-switch mechanism when the third strand binds to part of the sense and antisense strand of the dsDNA helix (alternate strand triplex, Frank-Kamenetskii & Mirkin, 1995).

A DNA triplex inhibits transcription of a specific gene either by preventing the binding of regulatory transcription factors (Cooney *et al.*, 1988) or by directly interfering with the RNA Polymerase action at either the initiation or elongation steps (Young *et al.*, 1991). DNA triplex formation may also inhibit progression of DNA Polymerase and affect DNA replication *in vitro* (Samadashwily *et al.*, 1993).

The effect of a DNA triplex on transcription initiation reflects one or more negative interactions such as steric occlusion of a transcription factor or RNA polymerase binding, stiffening or bending of the promoter DNA, recruitment of inhibitory activity that recognises the third stranded structure, and/or generation of a repressive chromatin configuration at the promoter (Maher *et al.*, 1992).

The advantage of the antigene over the antisense strategy is that there are fewer target molecules per cell, the regeneration rate of DNA is far lower than that of RNA, and higher selectivity due to the specific binding of the third strand (Section 1.2, Kochetkova *et al.*, 1996).

1.2 THE DNA TRIPLEX

1.2.1 HISTORY OF DNA TRIPLEX RESEARCH

The DNA triplex structure was first demonstrated by Felsenfeld, Davies and Rich who used two synthetic polyribonucleotides, polyadenylic acid (polyA) and polyuridylic acid (polyU) to form a polyA-polyU-polyU triplex (Felsenfeld *et al.*, 1957). Seventeen years later it was shown that in the analogous DNA triplex (polyT-polyA-polyT), one of the polypyrimidine strands was wrapped around the duplex in the major groove (Arnott & Selsing, 1974).

Major interest in DNA triplexes was initiated by two independent developments in 1987. Firstly, it was observed that H-DNA, an intramolecular triplex, was formed in homopurine-homopyrimidine mirror repeats under negative superhelical stress and/or acid pH (Figure 1.3). The formation of H-DNA is believed to be of potential importance in gene regulation and recombination (Lyamichev *et al.*, 1991). In H-DNA, the pyrimidine strand forms the pyrimidine-purine-pyrimidine triplex with one half of the purine strand while the other half of the purine strand remains unstructured (Htun & Dahlberg, 1989).

Secondly, homopyrimidine and homopurine oligonucleotides were found to target homopurine sequences in the DNA duplex and bind in a sequence specific manner, forming a DNA triplex. Although DNA triplexes can only be formed at homopurine-homopyrimidine sequences, these sequences are found in or close to many transcriptional units and recombination hot spots in eukaryotes (Wells *et al.*, 1988).

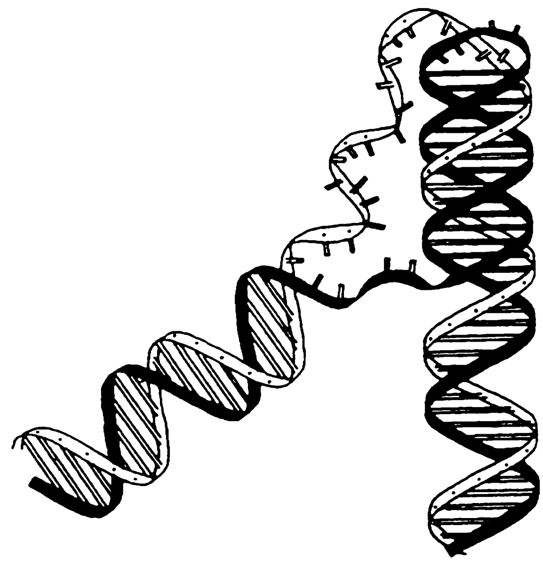


Figure 1.3: Diagram depicting the structure of H-DNA (Wells *et al*, 1988).

Alternating $[d(G\text{-}A)\text{-}d(T\text{-}C)]_{>20}$ sequences constitute up to 0.4% of the total mammalian genome and have functional roles in various biological processes in eukaryotic organisms (Manor *et al.*, 1988). Also, relatively long stretches ($n \geq 10$) of homopurine-homopyrimidine sequences are known to be 3.5-4 times over represented (above random expectation) in higher eukaroytes and eukaryotic viruses (Behe, 1987). Homopurine-homopyrimidine sequences are virtually absent from bacterial and phage genomes (Rippe *et al.*, 1992).

There has been little precise information regarding the structure adopted by various triple helical motifs. In general, the DNA triplex maintains the right-handed turn of DNA duplex without distorting it greatly (Arnott & Selsing, 1974). Initial Fibre diffraction data indicated that the underlying duplex of a DNA triplex adopted an A-DNA structure (Arnott & Selsing, 1974) whereas recent studies have suggested that it may closely resemble a B-DNA structure (Howard *et al.*, 1992).

The structure and sequence requirements for DNA triplex formation have been investigated *in vitro* under both physiological and nonphysiological conditions. Two types of DNA triple helix binding motifs have been described, differing in the sequence composition of the third strand, the relative orientation and position of the backbone of the three strands, and the base triad interactions.

1.2.2 HOOGSTEEN HYDROGEN BONDING

Hoogsteen hydrogen bonds were originally described by Hoogsteen in 1963. Essentially they are very similar to Watson and Crick hydrogen bonding. For example, the base pairing between adenine and thymine nucleotides involves two hydrogen bond interactions. Watson and Crick hydrogen bonding of adenine bases use N1 and C6 amino groups whereas the Hoogsteen hydrogen bonding utilizes the N7 and C6 amino groups (Figure 1.4). The Watson and Crick hydrogen bonding of thymine bases use the same atoms as the Hoogsteen hydrogen bonding between the N3 hydrogen and C4 oxygen (Frey *et al.*, 1973).

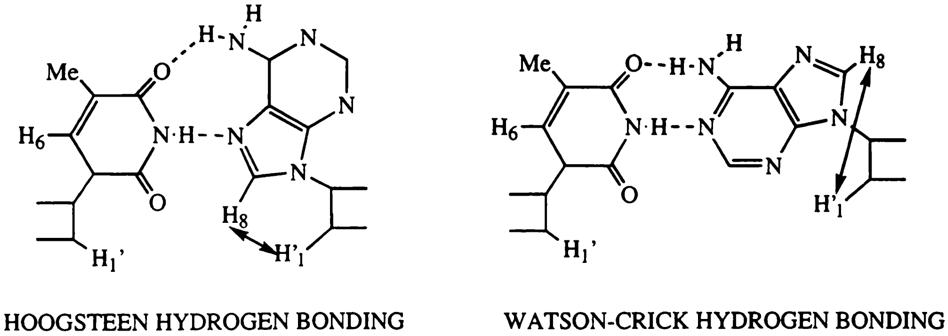


Figure 1.4: Comparison of Watson and Crick hydrogen bonding and Hoogsteen hydrogen bonding between thymine and adenine bases (Searle & Wickham, 1990).

Hoogsteen hydrogen bonding requires that one strand in the helix must contain a stretch of purine bases and the other strand contains pyrimidine bases, whereas in Watson-Crick hydrogen bonding a mixture of both is permitted. The conformational space is also highly restricted between a Hoogsteen paired double helix compared to a Watson-Crick paired double helix, whose conformational space is largely variable (Raghunathan *et al.*, 1994).

1.2.3 THE PYRIMIDINE MOTIF

In the pyrimidine motif, the thymine (T) in the third strand recognises Watson-Crick base pairs adenine(A)-thymine(T) forming the DNA triplex T·AT[1], and the protonated N3 of cytosine (C^+) in the third strand recognises Watson-Crick pairs guanine(G)-cytosine(C) to form the DNA triplex C^+·GC (Moser & Dervan, 1987; Le Doan *et al.*, 1987). The pyrimidine rich third strand oligonucleotides bind in a parallel orientation with respect to the homopurine target strand forming isomorphous Hoogsteen hydrogen bonding (Figures 1.5, 1.6).

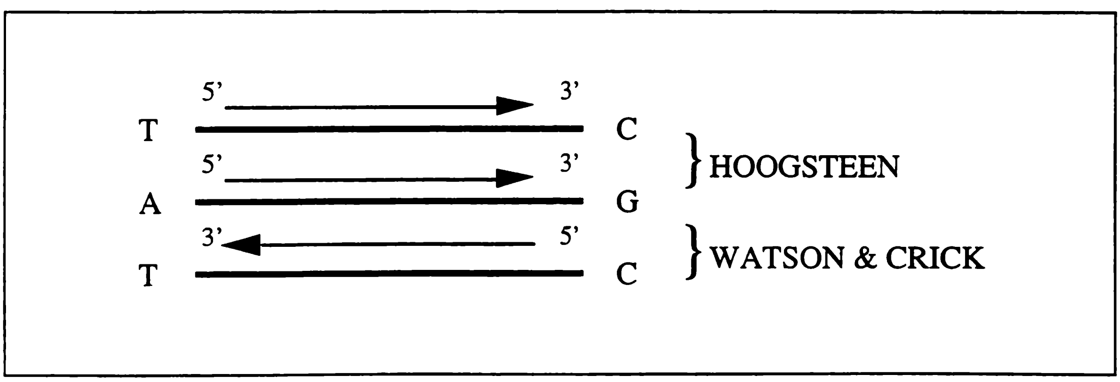
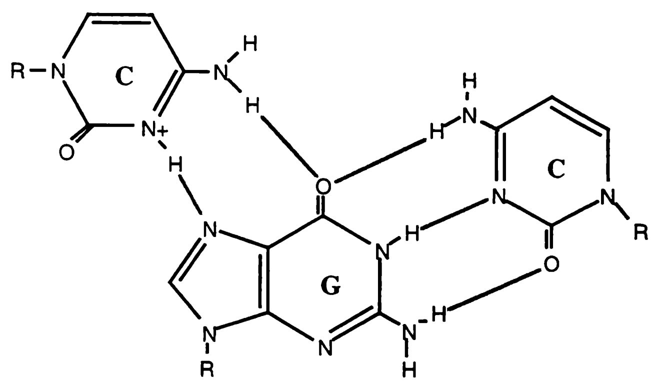
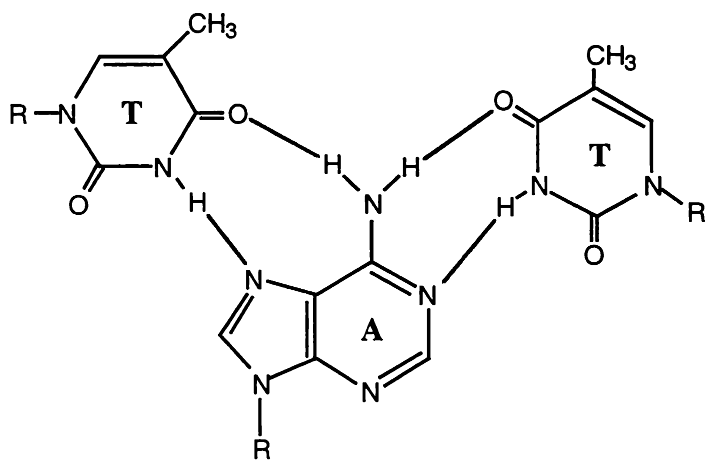


Figure 1.5: Strand orientation of the pyrimidine motif (Kool., 1996).

[1] T· represents the third strand; AT represents the Watson & Crick hydrogen bonding between adenine and thymine base pairs. This format will be standard for all DNA triplexes described in this current study.



C^{+}·GC BASE TRIAD



T·AT BASE TRIAD

Figure 1.6: DNA composition of the pyrimidine motif (Kool, 1996).

C^{+}·GC DNA triplexes are destabilised at physiological pH due to the requirement for base protonation of cytosine on the Nitrogen 3 atom in order that two Hoogsteen hydrogen bonds are formed. Therefore, DNA triplex formation by the oligopyrimidines is pH dependent and of relatively low affinity at neutral pH, since the Equilibrium constant of an acid (pK_a) is 4.3 for cytosine (Bates *et al.*, 1996).

This creates problems for designing DNA triplexes that will work under physiological conditions, but these can be overcome by replacing the third strand cytosine with 5-methylcytosine residues (5-Me-dC) as this compound has a higher pK_a value, 4.8. However, this may not be sufficient to alter the pH requirement for recognition of contiguous blocks of GC base pairs (Koh & Dervan, 1992). Other analogues have been explored such as pseudoisocytosine, 6-oxocytosine α- and β-amino-5-(2-deoxy-D-ribofuranosyl)pyrimidines but, so far, none have been used regularly in the antigene studies because they decrease DNA triplex stability. The design of such readily synthesised cytosine analogues still remains a challenge (Bates *et al.*, 1996).

Organic solvents have also been used, with moderate success, to stabilise the pyrimidine motif (Moser & Dervan, 1987).

1.2.4 THE PURINE MOTIF

The purine motif involves third strand purine rich oligonucleotides binding antiparallel with respect to the homopurine target, forming reverse Hoogsteen hydrogen bonds and generating purine-purine-pyrimidine DNA triplexes A·AT, G·GC, and T·AT (Dervan & Beal, 1991; Durland *et al.*, 1991).

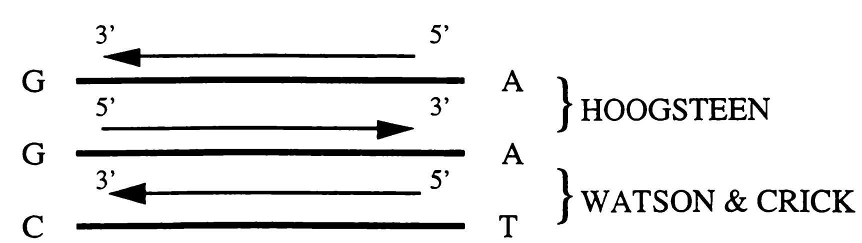


Figure 1.7: Strand orientation of the purine motif (Kool, 1996).

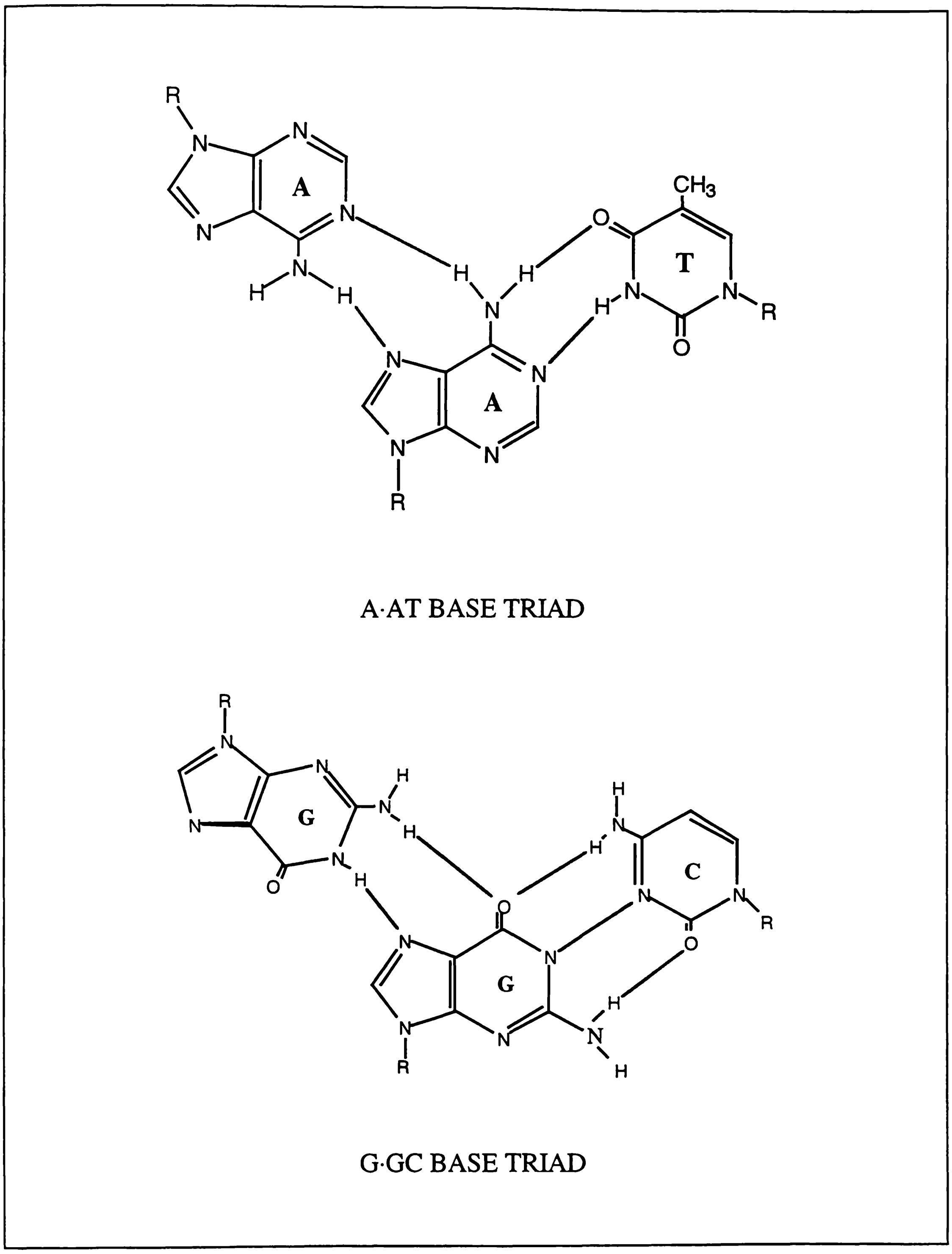


Figure 1.8: DNA composition of the purine motif (Kool., 1996).

The purine motif is largely insensitive to pH and is stabilised by metal ions and/or polyamines under neutral pH conditions (Helene, 1991). Hence it is more versatile and useful under physiological conditions than pyrimidine motif DNA triplexes.

Purine motif DNA triplexes are stabilised by bivalent cations, especially Manganese (Mn^{2+}), Cadmium (Cd^{2+}), Cobalt (Co^{2+}), Nickel (Ni^{2+}) and Zinc (Zn^{2+}) (Lyamichev *et al.*, 1991). Transition metal ions stabilise the purine motif DNA triplex whereas the ions of the alkali earth metals do not (Malkov *et al.*, 1993). Barium (Ba^{2+}) and Mercury (Hg^{2+}) ions do not stabilise the purine motif DNA triplex either, probably due to the cations being too bulky to be incorporated into the DNA triplex (Malkov *et al.*, 1993).

Combtype polycations such as poly(L-lysine) and poly(L-arginine) have been designed to thermally stabilise the DNA triplex (Marugama *et al.*, 1997). Poly(L-lysine) graft dextran copolymers also stabilise purine and pyrimidine motif DNA triplexes within the rat $\alpha(1)$ collagen gene promoter under physiological pH and ionic conditions (Ferdous *et al.*, 1998).

Several reports have shown that purine motif DNA triplex formation is diminished by physiological concentrations of monovalent cations (M^+) such as potassium (K^+) (Cheng & Van Dyke., 1993). The presence of K^+ ions does not substantially decrease the stability of the DNA triplex but decreases the kinetics of their formation (Debin *et al.*, 1997). The M^+ inhibition of triplex formation exhibits a concentration and ionic radius dependence that correlates with the ability of M^+ to stabilise guanine quartet structures (Olivas & Maher, 1995).

Guanine quartets are formed by the aggregation of four guanine residues into a square planar configuration (Figure 1.9). Each guanine interacts with its two neighbours via Hoogsteen like hydrogen bonding, resulting in a very tightly bound quartet (Williamson *et al.*, 1989). Stable guanine quartets are found in telomeric G-strands of eukaryotic chromosomes and may contribute to the regulation of telomeric length *in vivo* (Zahler *et al.*, 1991).

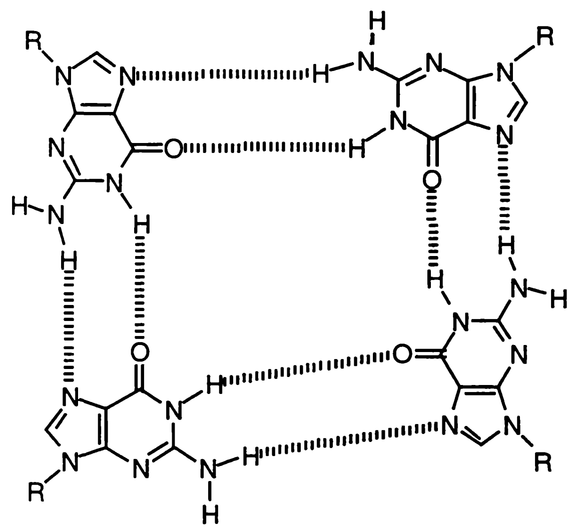


Figure 1.9: Structure of the guanine quartet (Shafer, 1998).

Intermolecular GA homoduplexes for oligonucleotides containing (GA) repeats (Rippe *et al.*, 1992) and guanine quartets (Olivas & Maher, 1995) are two major competing self-aggregating structures which affect DNA triplex formation. Purine oligonucleotides have a strong tendency to self-aggregate at temperatures less than 37°C (Cheng & Van Dyke., 1993). This problem has been successfully overcome *in vitro* by replacing guanine with 6-thioguanine (Olivas & Mather, 1995) or 7-deazaguanine (Fawad Faruqi *et al.*, 1997), using zipper structures (Svinarchuk *et al.*, 1996) or counteracting the stabilisation of the quartet with divalent transition metal cations (Blume *et al.*, 1997).

1.2.5 ALTERNATIVE DNA TRIPLEX SEQUENCES

DNA triplexes exhibit a lower tolerance for unfavourable sequence mismatches than do DNA duplexes. The base sequence requirements for Hoogsteen hydrogen bonding are also more stringent. The melting temperature change (ΔT_m) per mismatch for DNA triplexes ranges between 12.7 and 20°C, whereas it varies only between 6.0 and 7.7°C for DNA duplexes (Fossella *et al.*, 1993). Mismatches destabilise the DNA triplex but some sequence variations such as G·TA (Griffin & Dervan, 1989) and T·CG (Yoon *et al.*, 1992) allow DNA triplex formation (Figure 1.10, Table 1.1). Third strands containing guanine and thymine can adopt a parallel or an antiparallel orientation with respect to the dsDNA target sequence (Sun *et al.*, 1991).

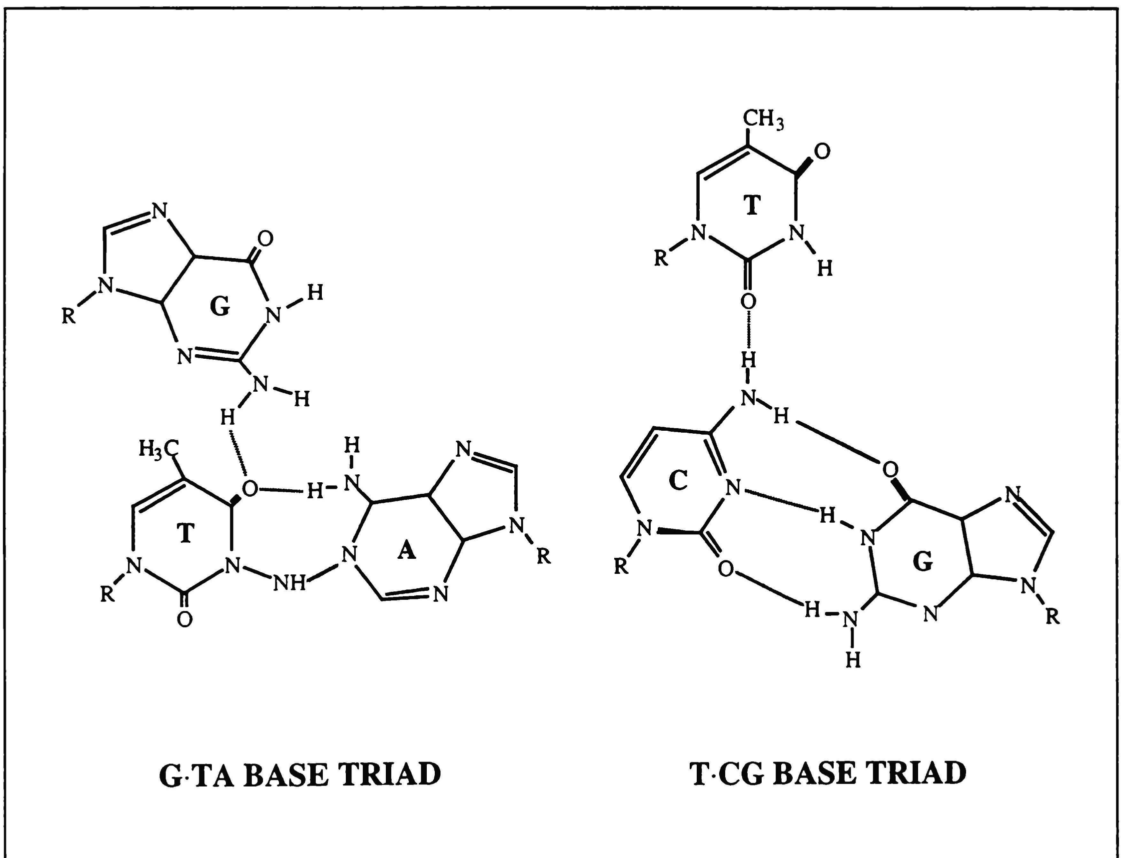


Figure 1.10: Structure of mismatched DNA triads; G·TA (Griffin & Dervan, 1989) and T·CG (Yoon *et al.*, 1992).

	BASE IN TFO	
BASE IN PURINE RICH STRAND OF THE TARGET DNA DUPLEX	PYRIMIDINE MOTIF	PURINE MOTIF
G A T C	C/(G) T G T/C	G T/A C T

Table 1.1: The binding of unmodified TFOs to the purine rich strand of the target DNA Duplex.

Under certain conditions, specific DNA triplexes at a $(AT)_n$ tract with alternating G·TA and T·AT triplets are possible. DNA triplex stability increases with the length of the T·AT anchor and stretch of alternating G·TA and T·AT triplets (Gowers and Fox, 1998).

Every change of base in the TFO leads to a distortion of the third strand and thus to an energetic penalty. It has been shown that a 15 nt purine TFO can be interrupted by four central pyrimidines and still be able to form a stable DNA triplex. However, addition of each pyrimidine residue leads to a thirty fold decrease in 3rd strand affinity (Gowers & Fox, 1997). Replacement of a single base pair in the centre of a homopurine sequence is highly detrimental to the third strand binding and formation of a DNA triplex. However a mismatch at the end of the homopurine-pyrimidine-purine sequence is less destabilising (Mergny *et al.*, 1991).

Purine-rich oligonucleotides targeted to the human dihydrofolate reductase promoter with a combination of A·GC, G·CG, and C·GC mismatches within the purine-purine-pyrimidine DNA triplex was destabilising, with the G·AT triplet being extremely unfavourable for DNA triplex formation. This is caused by the alignment of the phosphodiester backbone of the third strand in relation to the dsDNA target (Blume *et al.*, 1992).

1.2.6 THIRD STRAND BINDING SPECIFICITY

The haploid mammalian genome consists of 3 X 10^9 bp and given the same *random* genome size, a *random* 17 nt DNA sequence would have a high chance of being unique. However, for absolute specificity one would want to be able to distinguish between a DNA triplex formed with a 17 bp perfect match and one with a single base mismatch (Branch, 1998).

DNA triplex formation is highly specific, for example, a 16 nt polypyrimidine TFO was able to selectively bind and form a triplex to a single 16 nt target site *in vitro* among more than 3x 10^9 bp genomic DNA (Strobel *et al.*, 1991). A 13 nt TFO is expected to be unique in the mammalian genome, however, given the size of the genome, there could be multiple shorter nucleotide matches with other genomic sequences under certain conditions (Simons *et al.*, 1996).

The specificity of the DNA triplex is mostly due to the Hoogsteen hydrogen binding between the TFO and the duplex DNA (Figure 1.6). However, not every TFO will bind with high affinity due to an inappropriate DNA sequence content. DNA triplex formation is 10^6 times more specific than DNA recognition by restriction enzymes (Moser & Dervan, 1987).

1.2.7 DETECTION OF DNA TRIPLEX FORMATION

Detection of DNA triplexes has utilized many techniques such as comigration of oligonucleotides with DNA (Lyamichev *et al.*, 1988), photofootprinting (Lyamichev *et al.*, 1990, 1991), NMR (Rajagopal & Feigon, 1989), cleavage of DNA by chemical groups attached to oligonucleotides (Moser & Dervan, 1987), DNA footprinting using DNAse I (Cooney *et al.*, 1988), Fourier Transform Infrared (FTIR) Spectroscopy (Liquier *et al.*, 1991), Immunofluorescent staining with monoclonal antibodies to DNA triplex (Burkholder *et al.*, 1988), Gel retardation assays (Lyamichev *et al.*, 1988), and a Fiber optic biosensor (Uddin *et al.*, 1997).

Several of these techniques are indirect or have low time resolution, while others require large amounts of oligonucleotide samples (Yang *et al.*, 1994).

1.2.8 APPLICATIONS OF THE ANTIGENE STRATEGY

DNA triplex formation has several potential applications such as acting as therapeutic tools to determine the role played by any given gene in a living cell, to generate unique cleavage sites in DNA, or to act as sequence-specific artificial nucleases, modulators of DNA binding proteins, and regulators of gene expression. Most antisense and triplex oligonucleotides are currently being developed for treatment of cancer and viral diseases.

Inhibition by the antigene strategy has been investigated in various genes such as *c-myc* (Cooney *et al.*, 1988), IL2Rα (Orson *et al.*, 1991), β-lactamase promoter (Dual-Valentin *et al* ., 1992), murine *c-pim-1* proto-oncogene (Svinarchuk *et al.*, 1994), rat *neu* oncogene promoter (Gee *et al.*, 1994, 1995), murine Ki-*ras* proto-oncogene (Xodo, 1995), rat α1(I) Collagen Promoter (Kovacs *et al.*, 1996), human *aprt* (Faucon *et al.*, 1996), and rat aquaporin-5-gene (Delporte *et al.*, 1997). The mouse metallothionein I promoter was targeted by a TFO, with the resulting DNA triplex successfully blocking the binding of the Sp1 transcription factor *in vitro* (Maher *et al.*, 1989). Inhibition of HIV-1 replication in human cells (McShan *et al.*, 1992), and of Klenow Fragment of DNA polymerase (Hacia *et al.*, 1994) has also been reported. Each experiment differed in the TFO target sequene and the conditions of the DNA triplex binding assay.

The antigene strategy also has been used to suppress transcription, to inhibit replication and to induce site-specific mutations in an extrachromosomal vector transfected into mammalian cells (Wang *et al.*, 1995). *In vitro* targeting of a TFO within the genome of transgenic mice also caused site-specific modification of the mouse genomic DNA (Gunther *et al.*, 1996).

1.2.9 KINETICS OF DNA TRIPLEX FORMATION

Proteins bind to their substrates with biomolecular association rate constants of 10^7-10^{10} $M^{-1}s^{-1}$ (Berg *et al.*, 1982) whereas TFOs bind to their dsDNA targets with bimolecular association rate constants of 10^2-10^5 $M^{-1}s^{-1}$ (Paes & Fox, 1997). Hence DNA triplex formation is slow, and has also been reported to be three orders of magnitude slower than DNA duplex formation (Rougee *et al.*, 1992). It was observed that the poly(rA)·2poly(rU) DNA triplex formed one hundred times more slowly than the corresponding DNA duplex (Blake *et al.*, 1966, 1967).

The association and disassociation rates of DNA triplex formation in the case of DNA is also slower than RNA (Blasko *et al.*, 1996). This slow association rate may be due to the oligonucleotides self-aggregating in solution and having to disrupt their unusual structures before binding to their target, or alternatively the double helix target has to switch from B-DNA to A-DNA before hosting the oligonucleotide into its major groove (Alunni-Fabbroni *et al.*, 1996). The slow rate may be overcome by increasing the oligonucleotide concentration (Maher *et al.*, 1990).

The kinetics of triplex formation will ultimately depend on many factors such as the pH, temperature, ionic conditions, length and sequence of the TFO used. A filter binding assay showed that the dissociation constant of a pyrimidine-purine-pyrimidine DNA triplex containing a 19 nt polypyrimidine TFO was virtually temperature independent at low pH while it become strongly temperature dependent as the pH increased with decreased association at higher temperatures. Also, the formation of the DNA triplex was a second-order reaction at low pH, whereas a third-order reaction occurred at neutral pH. The pH dependence of the dissociation constants of the DNA triplex is a result of the rapid acid-base equilibrium of pyrimidine single strands (Shindo *et al.*, 1993).

1.3 *IN VIVO* DNA TRIPLEX FORMATION

1.3.1 CELLULAR UPTAKE OF OLIGONUCLEOTIDES

The uptake of oligonucleotides by cells depends on their length, overall charge, hydrophilicity, lipophilicity and extracellular concentration. In particular, while increasing the length of the oligonucleotide improves binding to, and hybrid stability with DNA, uptake may become less efficient and increased binding to nonspecific target sequences may occur. The uptake of oligonucleotides into the cell requires energy and is maximal at 37°C, presumably due to cells dying at higher temperatures (Crooke, 1991).

1.3.2 CELLULAR DELIVERY OF OLIGONUCLEOTIDES

Oligonucleotides are polyanionic and cannot passively diffuse across the cell membrane. Therefore delivery of the oligonucleotides to the target cell can only be achieved by artificial means such as electroporation (Bergan *et al.*, 1993), microinjection (Wagner *et al.*, 1993), transfection using cationic liposomes (Bennett *et al.*, 1992) or viruses (adenovirus-polylysine complexes), or addition of Streptolysin O to the culture medium (Bonn, 1996). Each method is characterised by a series of limitations and drawbacks. No consensus has been reached as to which method is the best, due to differences in cell lines and the use of various TFOs in the individual experiments. The efficiency of delivery depends on many factors and sometimes is decreased due to the oligonucleotides becoming separated into different cellular compartments from the targeted plasmid or alternatively, trapped by nuclear proteins (Stein & Cheng, 1993).

One method of improving the delivery of oligonucleotides to the desired target cell is encapsulation of the oligonucleotides in liposomes.

Liposomes have the advantage over the viral gene delivery system of being biologically inert and therefore do not constitute a disease risk (Lasic et al., 1998). However, liposomes are sensitive to serum (Lewis *et al.*, 1996) and they are less efficient than viruses for transferring genes to cells (Lasic, 1997). Their efficiency also depends on the tissue type (Jarnagin *et al.*, 1992), lipid membrane composition (Gould-Fogerite *et al.*, 1990) and the relative mass ratio of DNA to lipid (Felgner *et al.*, 1987).

1.3.3 ACCESSIBILITY OF TARGET DNA WITHIN CELLS

Successful *in vivo* DNA triplex formation will ultimately depend on the accessibility of the target DNA sequence within the chromatin structure in the cell nucleus. Limited information is available regarding the TFOs ability to bind to its dsDNA target in a cellular environment, as it is difficult to detect such a non-covalent complex by *in situ* footprinting methods currently available (Giovannangeli *et al.*, 1997). Potential DNA triplex binding sites may not be accessible, as the genomic DNA in its intact supranucleosomal structure in the cell nuclei is wrapped around the protein core. Recently, modified TFOs have been targeted to the Hamster hypoxanthine phosphoribosyltransferase gene (*Hprt*) and phenotypically *Hprt*-deficient clones were recovered, suggesting that the chromatin structure in the target triplex region is more dynamic than predicted by current *in vitro* experiments (Majumdar *et al.*, 1998).

DNA triplex formation has been shown to be inhibited within the nucleosome. However, intermolecular triplexes are able to form towards the ends of the nucleosomal fragments, in regions where the DNA is less tightly associated with the protein core. This may be due to fraying or breathing of the duplex ends (Brown & Fox, 1996). Brown *et al.*, 1998 showed that DNA triplex formation is only possible on the first 40 or so base pairs from the ends of the nucleosome-bound DNA.

In general, the narrow minor groove of AT rich sequences are orientated towards the protein whereas the GC rich regions, with their wider than average minor grooves, face away from the protein (Drew & Travers, 1985). Intermolecular DNA triplexes have been observed on nucleosome-bound DNA fragments containing $A_n{\cdot}T_n$ tracts but the stability of the complex depends on the orientation of the A tract with respect to the protein surface (Brown & Fox, 1998). DNA triplex formation may cause some nucleosome remodeling, resulting in a change in the translational position of the DNA (Brown *et al.*, 1998). Recent studies have also shown that the most important factor affecting DNA triplex formation is the position of the target site within the nucleosome bound DNA fragment (Brown & Fox, 1999).

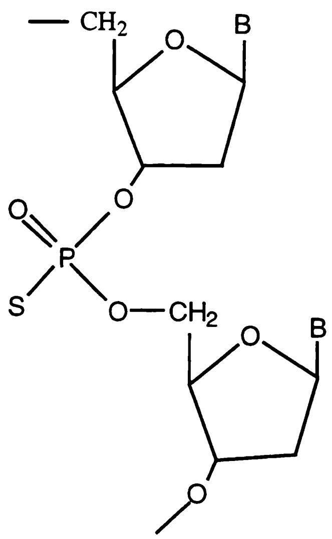
The physiological effect of DNA triplex formation during the cell cycle has not been investigated and documented. Further research is required to determine the effect of targeting TFOs to dividing cells as opposed to resting cells.

1.3.4 DEGRADATION OF OLIGONUCLEOTIDES

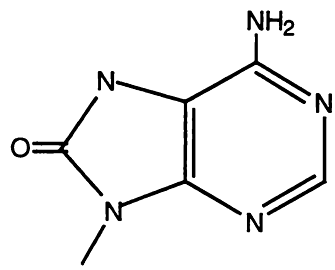
Oligonucleotides are degraded by intracellular nucleases with unmodified oligonucleotides in biological fluids having a half life of about five minutes in serum and thirty minutes in living cells (Monia *et al.*, 1996). It has also been observed that the SV40 large T-antigen helicase can unwind pyrimidine motif DNA triplexes (Kopel *et al.*, 1996). Degradation of oligonucleotides by intracellular nucleases can be successfully overcome by modifying the oligonucleotide heterocyclic backbone or ribose moiety.

A variety of oligonucleotide analogues such as the phosphorothioates (Akhtar & Agrawal, 1997) and the 8-oxoadenine analogues (Miller *et al.*, 1992) have been found to extend the half life of the oligonucleotides (Figure 1.11). The 8-oxoadenine analogs are easy to prepare. The nucleotide base adopts the syn conformation, which enables the analog to bind to guanine with two hydrogen bonds (Miller *et al.*, 1996).

Phosphorothioate oligonucleotides were first synthesised in 1984 with each phosphorus in the oligonucleotide bound to a sulphur atom (Stec *et al.*, 1984). Research has shown that cellular uptake by phosphorothioates in tissue culture is high and involves several active mechanisms (Crooke, 1991). Several antiviral and anticancer phosphorothioate oligonucleotides have now entered clinical trials (Wagner & Flanagan, 1997).



PHOSPHOROTHIOATE



8-OXOADENINE

Figure 1.11: The chemical structure of a phosphorothioate and 8-oxoadenine. The nucleotide bases adenine, thymine, cytosine and guanine are represented by the letter "B."

Alternatively, Peptide Nucleic Acids (PNA) can be used as an oligonucleotide analogue. These compounds mimic DNA by having a pseudopeptide backbone composed of achiral and uncharged *N*-(2 aminoethyl)glycine units (Nielson *et al.*, 1991) (Figure 1.12). These analogues hybridise specifically and selectively with high affinity to complementary sequences in single stranded and double stranded DNA. PNA is uncharged and stable at physiological pHs; however, it has poor H_20 solubility, a tendency to self aggregate, the inability to activate RNAse H and poor cellular uptake (Nielson *et al.*, 1996).

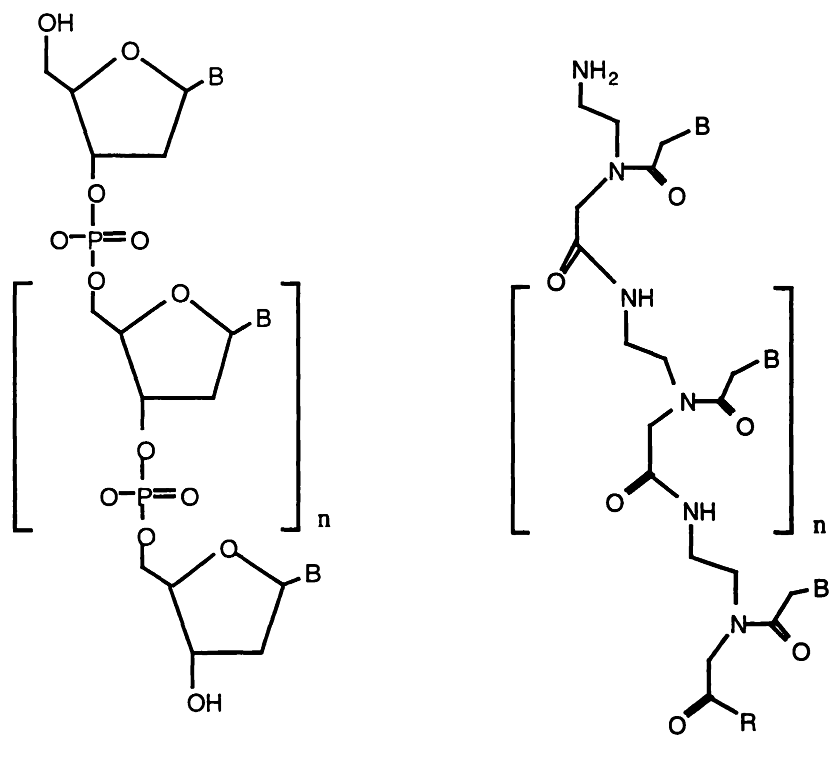


Figure 1.12: A diagram showing the chemical structures of PNA and DNA. B represents a nucleoside base such as adenine, guanine, cytosine or thymine (Nielson, 1996).

1.4 STABILITY OF THE DNA TRIPLEX

Third strand binding is dependent on a variety of parameters such as base composition, temperature and pH, so altering these conditions may improve DNA triplex stability (Moser & Dervan, 1987). Alternatively, optimisation of DNA triplex stability might be achieved by variations in oligonucleotide length and sequence alignment of various triplet combinations (Kandimalla *et al.*, 1995). It was observed that the binding of an oligonucleotide homologous to a 12 nt target duplex was more important at the 3' end than the complementary 5' end due to the strong stacking interactions of guanine at the 3' end (Cheng & Van Dyke, 1994). Clarenc *et al.*, (1994) suggested that purine TFO should contain at least 65% guanines to enhance formation of a stable DNA triplex.

The affinities with which GT rich oligonucleotides form stable DNA triplexes can be high; Equilibrium dissociation constant (K_d) values of less than 10^{-9} M have been observed, and the mean bound lifetimes for these DNA triplexes can be many hours (Vasquez *et al.*, 1995). A co-migration assay showed that a polypurine oligonucleotide targeted to the promoter region of *c-pim*-1 gene was extremely stable at 37°C, and only began disassociating at 65°C. Indeed, even after twenty four hours at 37°C no DNA triplex dissociation was observed (Svinarchuk *et al.*, 1996). Recently, it has been observed that a 19 nt G/T rich oligonucleotide can inhibit the expression of a luciferase reporter plasmid *in vivo* for up to seventy two hours (Musso *et al.*, 1996).

1.4.1 POLYAMINES

Polyamines increase the yield of functional DNA triplex interactions at least ten fold when present in physiological conditions, presumably by stabilising the DNA triplex once formed. This stabilisation occurs because the polyamines are positively charged and target the phosphate groups on the DNA, therefore neutralising the high negative charge density imposed by the association of the three DNA strands (Thomas & Bloomfield, 1984). Millimolar concentrations of the natural polyamines, spermine and spermidine, promote efficient purine-purine-pyrimidine triplex formation under *in vitro* conditions. Their effect can be considerably decreased under physiological conditions due to the presence of coexisting cations. The effectiveness of these polyamines is directly proportional to the number of potential positive charges they carry (Musso & Van Dyke, 1995).

1.4.2 SMALL LIGANDS

A variety of small ligands, such as DNA binding drugs and intercalating agents, have been used to increase the stability of the DNA triplex and also prevent *in vivo* degradation.

A) DNA Binding Drugs

Combined exposure of the TFO and specific DNA triplex binding drugs, such as minor groove ligands may increase their ability to inhibit specific genes by DNA triplex formation or alternatively they may dissociate DNA triplexes. The formation of purine-purine-pyrimidine DNA triplexes on the human c-ki-Ras promoter has shown that AT specific minor groove ligands Distamycin and Berenil have no effect on the DNA triplex, whereas the GC specific minor groove ligand Mithramycin abolished the DNA triplex (Vigneswaran *et al.*, 1996). Berenil preferentially binds to DNA triplexes in low salt concentrations but destabilises the DNA triplex under high salt concentrations (Pilch *et al.*, 1995).

B) Intercalating Agents

It was first shown that the thermodynamic interaction of oligonucleotides with their nucleic acid targets was stabilised several fold by the covalent linkage of acridine orange, an intercalating agent, at the 5' terminus of the oligonucleotide (Helene *et al.*, 1985).

A variety of intercalating agents have now been employed such as coralyne, ethidium bromide (EtBr) and psoralen (Figure 1.13, 1.14). The addition of an intercalating agent at the 5' or 3' terminus of a TFO still preserves the properties of the original TFO, but significantly increases the stability of the DNA triplex.

Coralyne is a yellow fluorescent antitumour antibiotic which has convenient absorption bands at 330 and 450 nm, well removed from DNA absorption in the 250-280 nm range. It is a planar acromatic ligand with four fused aromatic rings (Moraru-Allen *et al.*, 1997). Coralyne is positively charged and stabilises poly d(CT)·poly d(GA)·poly d(C^+T) at low ionic strength and pH 7.0, possibly by intercalating between every base triad. However, coralyne has been reported to bind to DNA triplexes with low affinity (Lee *et al.*, 1993).

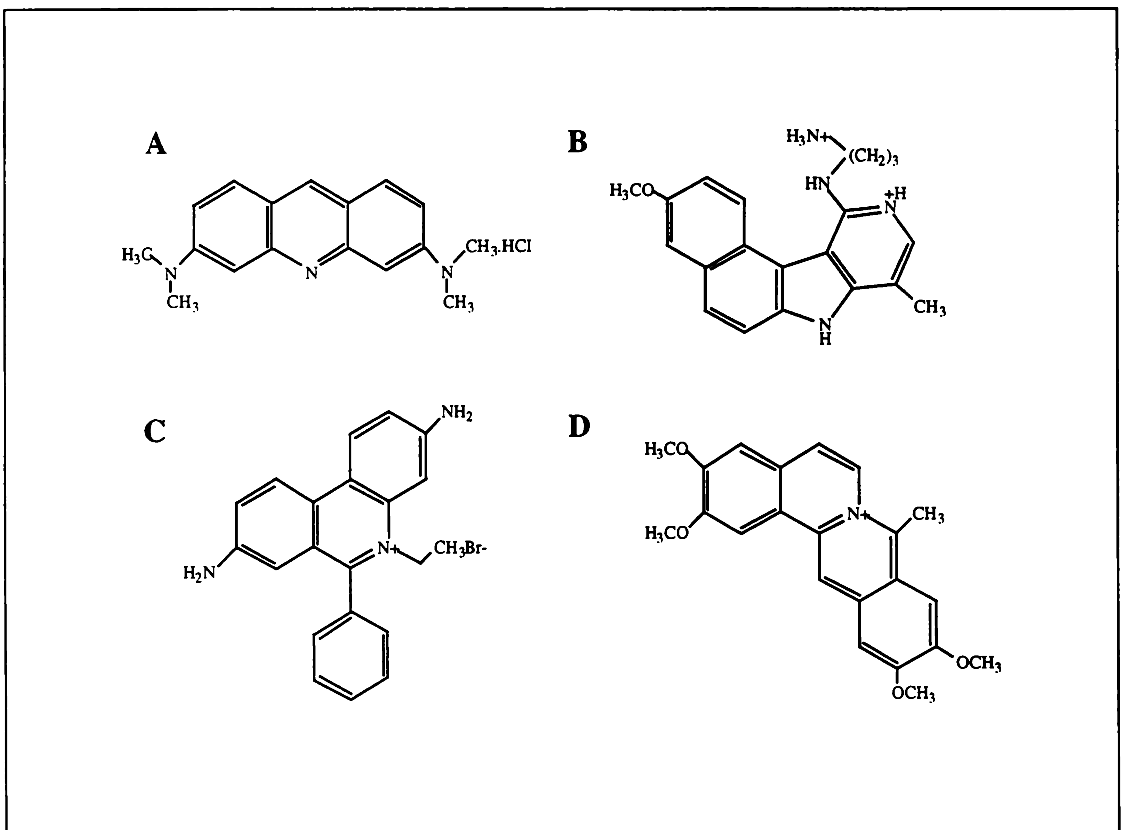


Figure 1.13: Chemical structure of various intercalating agents. A represents Acridine orange, B represents BePI (Mergny *et al.*, 1992), C represents Ethidium bromide, D represents Coralyne (Lee *et al.*, 1993).

Other tetracyclic acromatic compounds such as the quinoline derivatives [Benzopyridoindoles (BePI), and Benzopyridoquinoxaline (BPQ)] also strongly stabilise the DNA Triplex (Figure 1.13). BePI stabilises DNA triplexes that are T·AT rich whereas coralyne exhibits little preferential sequence binding. BePI can be used to stabilise a DNA triplex when the oligonucleotide pyrimidine-purine target sequence contains a single or double base pair inversion (Kukreti *et al.*, 1998). TFO-BePI conjugates may be able to recognise base pair inversions, therefore extending the current range of DNA target sequences available for DNA triplex formation.

Ethidium bromide can bind to DNA triplexes and stabilise them except for C^{+}·GC triplexes (Scaria & Shafer, 1991; Mergny *et al.*, 1991). Earlier studies showed that ethidium bromide destabilised the poly U·AU DNA triplex (Le Pecq & Paoletti, 1965) yet stabilised the T·AT DNA triplex by an intercalative mechanism (Scaria & Shafer, 1991).

Psoralen-conjugated TFOs have been successfully targeted to dsDNA, and with UV exposure, stable intermolecular DNA triplexes are formed with irreversible crosslinks (Takasugi *et al.*, 1991). Hence, psoralen-conjugated TFOs are very promising for *in vivo* gene targeting due to the photobiological and intercalating properties of psoralen. These properties will be discussed in the next section (Chapter 1.5).

1.5 PSORALEN

1.5.1 STRUCTURE OF PSORALEN

Psoralens are a class of furocoumarins which form covalent bonds with the pyrimidine bases of nucleic acids when irradiated with UV light (320-400 nm). They occur naturally at relatively high concentrations in a number of plants and consist of a linear furan ring joined to a coumarin ring (Figure 1.14).

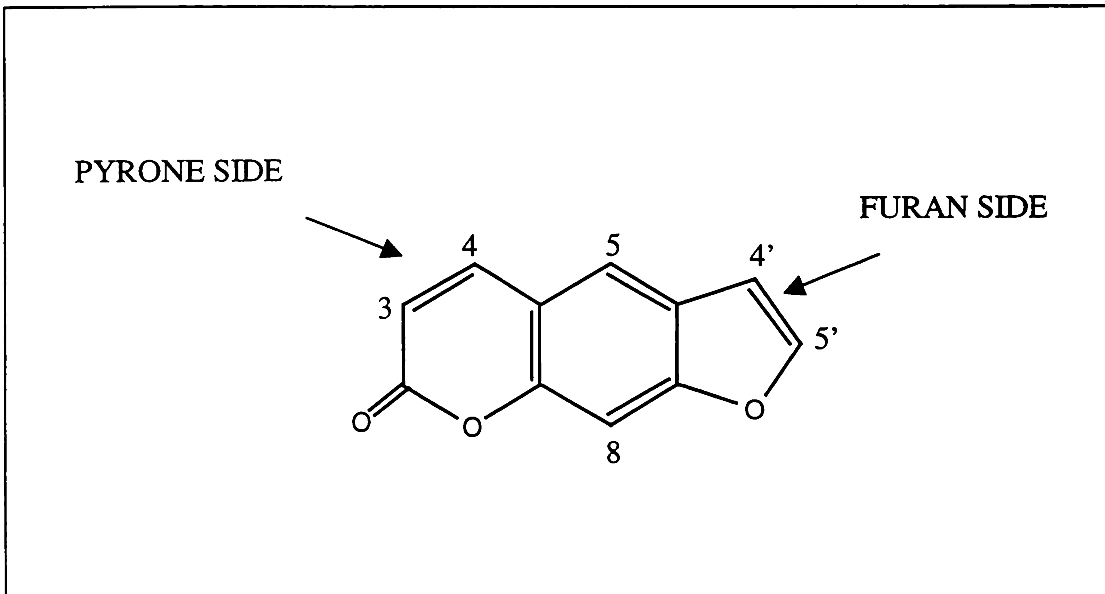


Figure 1.14: Molecular structure and numbering system of psoralen (Bates *et al.*, 1995).

The two most widely used naturally occurring psoralens are 8-methyloxypsoralen (8-MOP, Xanthotoxin) and 4, 5', 8-trimethylpsoralen (TMP or Trioxsalen). The photoreaction of psoralens with DNA was first reported by Musajo in 1969. The crystal and molecular structure of 8-methyloxypsoralen (8-MOP) was elucidated by Stemple and Watson in 1972.

1.5.2 PSORALEN INTERCALATION

The interaction of psoralen with DNA is similar to that of other polycyclic aromatic hydrocarbons such as ethidium bromide, and involves aromatic ring stacking, hydrophobic interactions and van der Waals forces (Musajo *et al.*, 1969). Psoralen primarily associates with thymine in DNA, and uracil in RNA by a two step reaction catalysed by UV light (360 nm) to form interstrand crosslinks (Figure 1.15). A minor reaction occurs with cytosine (Cimino *et al.*, 1985). A crosslink can only occur if an adjacent pyrimidine is present in the opposite strand of the DNA duplex. The non-covalent association of psoralen with DNA is UV-independent whereas the crosslinking reaction is UV-dependent. The formation of psoralen crosslinks disrupts normal replication, transcription and recombination events in DNA molecules (Hearst *et al.*, 1981).

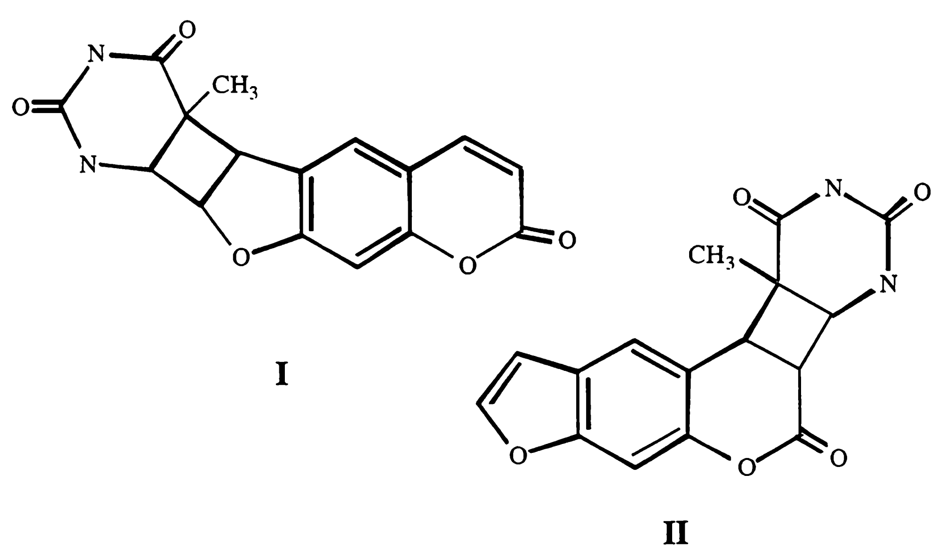


Figure 1.15: Molecular structure of the psoralen-thymine photoadduct. Two photoadduct structures are possible between the psoralen and thymine (Cole, 1970).

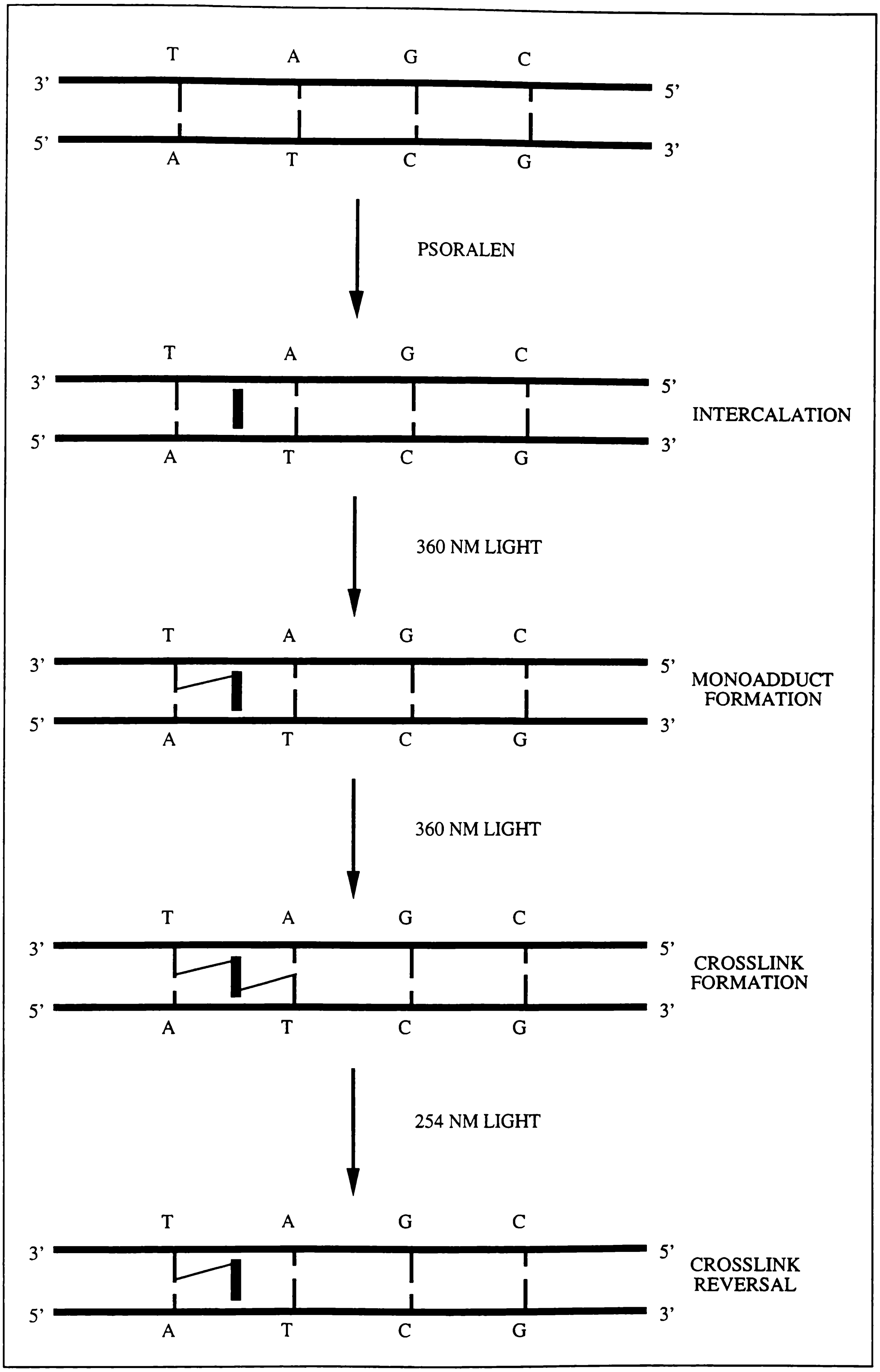


Figure 1.16: A Flow diagram showing the reaction between psoralen and the thymine nucleic acid of the double stranded DNA helix (Ussery *et al.*, 1992).

Initially psoralen intercalates between the adjacent pyrimidine base pairs in the DNA, and after irradiation with long UV light (320-410 nm), one UV photon is absorbed, forming furan-side or pyrone-side monoadducts (Kanne *et al.*, 1982). A wavelength of 360 nm is most common for UV crosslinking studies due to a convenient light source (mercury line emission). Recent investigations have determined that 320-335 nm may be more effective for inducing crosslinks, however, these shorter wavelengths maybe more harmful (Cripps *et al.*, 1982). Absorption of a second photon results in the conversion of furan-side monoadduct to a crosslink (diadduct) via photoreaction with the adjacent pyrimidine base on the opposite strand of the duplex DNA (Figure 1.16). The pyrone-side monoadducts do not absorb light of wavelengths greater than 320 nm and therefore cannot be converted to a crosslink (Cimino *et al.*, 1985). It has been shown also that the crosslink can be photo-reversed by exposure to 254 nm UV light or chemically reversed by treatment with alkali (Shi *et al.*, 1988). However psoralen is largely photo-destroyed during the photo-reversal of the adducts (Sastry *et al.*, 1993).

Pyrone monoadducts have two possible conformations; an opened or closed pyrone ring. The pyrone ring will open at pH 10 and close at pH 4 (Cimino *et al.*, 1986). Each pyrone monoadduct has a different mobility on a denaturing polyacrylamide gel and is interconvertable. The stereochemistry of both monoadducts and crosslinks is always cis-syn (Kanne *et al.*, 1982).

Several models for the intercalation of psoralen with DNA have been proposed. One possibility is that the dsDNA helix within the crystal DNA-psoralen complex is bent by 46.5° and unwound by 88° upon crosslinking to psoralen (Kim *et al.*, 1985). In an alternative model, the psoralen crosslinks do not generate bends in the DNA helix but lengthen the helix by approximately one base pair (Sinden & Hagerman, 1984). A third model proposes that psoralen crosslinks stiffen and bend the DNA double helix by less than 10°, and unwind the DNA by 1 bp (Haran & Crothers, 1988).

1.5.3 PSORALEN KINETICS

The Jablonski energy diagram shows that the kinetics of psoralen addition are affected by the separation of various energy levels which alter the lifetimes and distribution of the singlet and triplet states (Figure 1.17). The diagram shows that there are two transitions for psoralen which are localised in the C-2 carbonyl ($n \rightarrow \pi^*$ transition) and ring system ($\pi \rightarrow \pi^*$ transition). The lowest excited singlet and triplet states are regarded as a $\pi \rightarrow \pi^*$ transition and are implicated in the (2+2) cycloaddition with the pyrimidines of nucleic acids (Song & Tapley, 1979).

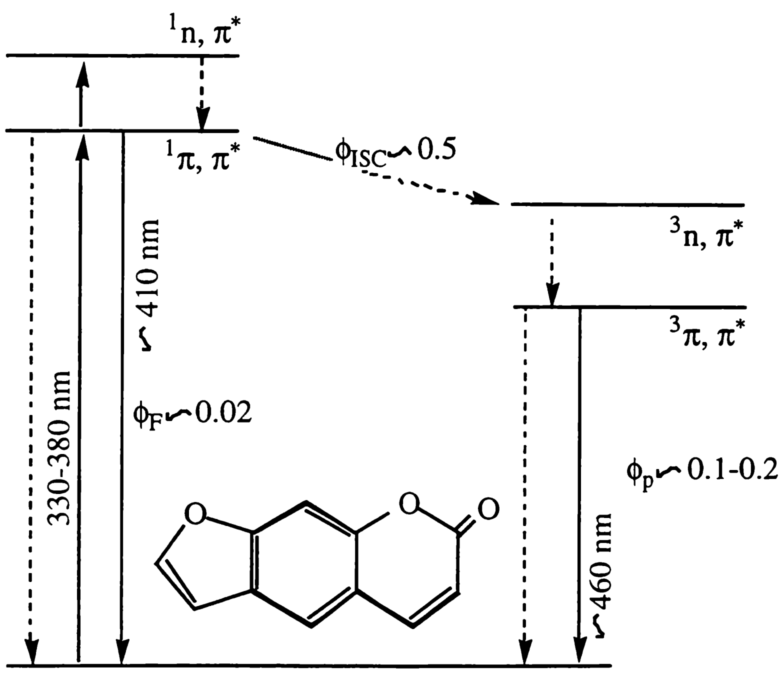


Figure 1.17: A Jablonski energy diagram for the known excited states of psoralen (Song & Tapley, 1979). F = Fluorescence, P = Phosphorescence, ISC = Intersystem crosslinking.

The absorption spectra of psoralen, 5-methoxypsoralen (5-MOP), 8-MOP and 5,8-dimethoxypsoralen are shown in Figure 1.18. There are three regions in which one or more absorption maxima occur; the less than 225 nm region, the 230-270 nm region and the 290-330 nm regions (Fowlks, 1959).

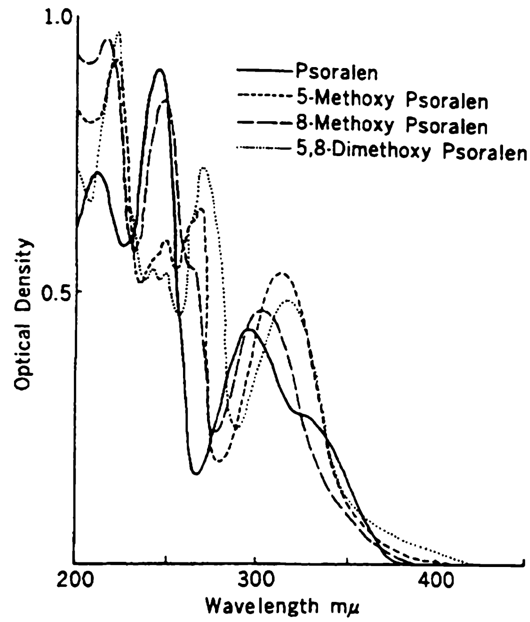


Figure 1.18: Absorption spectra of psoralens (Fowlks, 1959).

1.5.4 PSORALEN APPLICATIONS

Photosensitization of intracellular psoralen molecules can cause mutagenesis (Ashwood-Swith *et al.*, 1980) and inactivation of viruses (Musajo *et al.*, 1965), microorganisms (Grossweiner & Smith, 1981) and mammalian cell cultures (Nielsen & Bohr, 1983). Photoinduced psoralen interstrand crosslinks can induce recombination in bacterial, yeast and mammalian cells (Oroskar *et al.*, 1993). Psoralen cross-linking also provides the means to distinguish between active and inactive copies of ribosomal genes in both vertebrate and yeast cells (Conconi *et al.*, 1989). This can be achieved by crosslinking psoralen to DNA in chromatin under *in vivo* and *in vitro* conditions and analysing the electrophoretic mobility of the psoralen-crosslinked restriction fragments.

Psoralen interacts with unsaturated fatty acids (Kittler *et al.*, 1980), lipid membranes (Kittler *et al.*, 1980), ssDNA (Thompson *et al.*, 1982), and proteins (Megaw *et al.*, 1980) which can lead to enzyme inactivation (Verones *et al.*, 1982). Cellular and nuclear membranes are also permeable to psoralens (Hanson *et al.*, 1978). However, psoralen cannot cross link DNA that is associated with nucleosomes (Brown & Fox, 1996) as 60% of the DNA is protected by the protein core (Hanson *et al.*, 1976). The binding of psoralen is also sensitive to superhelical density, alternative DNA conformations and protein association (Ussery *et al.*, 1992).

Psoralens are used clinically in the treatment of dermatological disorders such as psoriasis (Parrish *et al.*, 1974), vitiligo (El Mofty, 1948), mycosis fungoides (Gilchrest *et al.*, 1976), and cutaneous T cell lymphoma (Edelson *et al.*, 1987). Psoriasis causes an increased turnover of epidermal tissue and approximately 1-3% of the worldwide population is affected by this non contagious disease (Baker *et al.*, 1972).

Plants containing high levels of psoralen such as celery and parsley leaves have been used for hundreds of years for the treatment of psoriasis (Scott *et al.*, 1976). They were also thought to act as natural insecticides in plants. Ancient Hindus, Turks and Egyptians used psoralen as medicine 3000 years ago. Psoralen was not extracted until 1834 when Kalbruner isolated 5-MOP from bergamot oil (Fowlks, 1959). Psoralen can also be isolated from microorganisms including fungi (Cimino *et al.*, 1985).

In 1974, treatment of psoriasis was introduced using oral 8-MOP and long-wave ultraviolet-A radiation (PUVA) (Parrish *et al.*, 1974). The treatment is very efficient with 80-90% clearance in 8-12 weeks upon oral or topical administration, in the presence of UVA light, although side effects may include nausea, erythema, and dizziness (Melski *et al.*, 1977). Several European and American studies have shown that PUVA therapy causes an increased risk of squamous cell carcinomas in humans; however, the aetiology is unknown (Stern *et al.*, 1984, 1988, 1998).

Psoralens are also used as probes to study DNA or RNA structure and DNA-protein structure (Ussery *et al.*, 1992). It is believed that the repair of psoralen adducts is via the Nucleotide Excision Repair (NER) pathway through sequential excision, gap filling and ligation (Hoeijmakers & Bootsma, 1994), although the pathways involved are not completely understood. The repair of psoralen crosslinks in an expressed mammalian gene may depend upon its mode of transcription. One study showed that 90% of the interstrand cross links were removed from the human DHFR gene within forty eight hours, while less than 25% were repaired in rRNA genes (Vos *et al.*, 1991).

1.5.5 PSORALEN MUTAGENICITY

DNA triplexes which involve psoralen crosslinks are more mutagenic than monoadducts (Sage & Bredberg, 1991). The majority of mutations are single base substitutions and their clastogenic and carcinogenic effects have been well documented. There has been much debate over the mutagenicity of psoralen crosslinks. However, it still remains a complex issue as several factors play a role. Mutations photoinduced at endogenous genes in human and rodent cells by psoralen targeted DNA triplex formation appear to be specific and highly targeted to sites of photoaddition (Laquerbe *et al.*, 1995). Gene modification can be extended beyond the site of third strand binding but is accompanied by a decrease in the precision of the targeting. This was shown by testing TFO conjugated to psoralen via linker arms of lengths varying from 2, 6, 26, and 86 bonds, allowing the psoralen to be delivered at varying distances from the third strand binding site. The frequency and distribution of the mutation(s) depend on the length of the linker arm with 2 (3%) and 6 (14%) giving the highest mutation frequencies (Raha *et al.*, 1998).

1.6 *supF* DNA TRIPLEX SYSTEM

A significant development for improving the use of the antigene strategy was achieved by Wang *et al.*, (1995), who showed that psoralen-conjugated oligonucleotides could be used to target specific mutations within viral genomes replicating in monkey COS cells. The experimental approach that was used is summarized in Figure 1.19. The *supF* gene in the shuttle vector pSP189 is 4944 bp (Paris & Seidman, 1992) and encodes an amber suppressor tyrosine tRNA of *Escherichia coli* (Brown *et al.*, 1979). Researchers favour the use of the *sup*F gene for mutagenesis studies because of its small size and easy detection of forward mutations to the nonsuppressor phenotype. This is achieved by a colorimetric assay in host bacteria carrying a *lac*Z nonsense mutation where blue colonies represent the *sup*F^{+} gene and white colonies represent the mutant *sup*F^{-} gene. Once *sup*F genes are identified as mutant, they can be further analysed by DNA sequencing methods. The *sup*F gene is also not expected to exhibit any phenotype in mammalian cells, only when the gene is retrieved from mammalian cells and introduced into the appropriate *E. coli* host (Glazer *et al.*, 1986).

A mutagenesis rate of 1-2%, mostly T·A to A·T transversions at the targeted psoralen intercalation site were observed when psoralen-conjugated TFOs were targeted to the *supF* mutation reporter gene in bacteriophage vector DNA or Simian Virus 40 vector DNA *in vitro* (Wang *et al.*, 1995). Targeted mutagenesis depends on the specificity of the TFO for the target site and the strength of oligonucleotide binding (Wang *et al.*, 1995).

DNA triplex directed DNA damage within mammalian cells has also been used to induce specific recombination and gene conversion in a sequence specific manner (Wang *et al.*, 1996). Targeted mutagenesis of a *supF* mutation reporter gene in bacteriophage vector DNA and Simian Virus 40 vector DNA was achieved by allowing psoralen oligonucleotides to enter cells, form a DNA triplex and stimulate intramolecular recombination with an extrachromosomal substrate.

Wang *et al.*, (1996) also demonstrated that there is a multiple pathway of recombination and repair due to the targeted psoralen monoadduct and crosslinks having different recombination products.

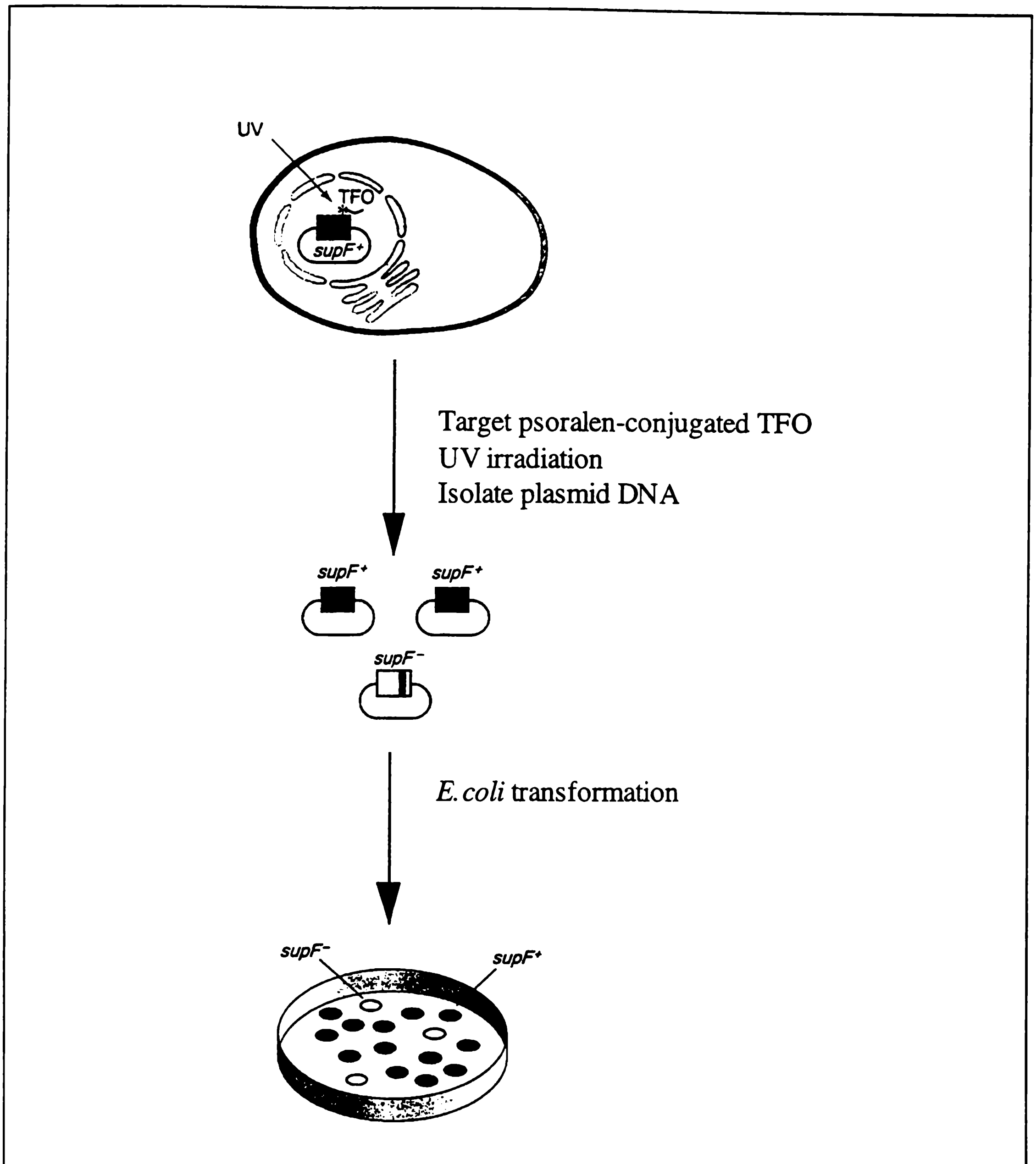


Figure 1.19: Strategy for targeted mutagenesis by TFOs. A plasmid and a psoralen conjugated TFO are transfected into monkey COS-7 cells to allow site-specific DNA triplex formation at the *supF* gene target. The cells are exposed to UVA irradiation to photoactivate the psoralen and generate photoadducts. Following vector replication, plasmid DNA is extracted and transformed into *E. coli* to allow genetic analysis of the *supF* gene. A blue colony represents the $supF^{+}$ gene whereas a white colony represents a mutant $supF^{-}$gene (Wang *et al.*, 1995, diagram modified from Vasquez & Wilson, 1998).

1.7 CURRENT OBJECTIVES

The ultimate aim of the current research is to develop psoralen-conjugated oligonucleotides that are capable of efficiently targeting specific sequences in mammalian DNA, forming DNA triplexes which are detectable by physio-chemical means. Normally, it is difficult to directly observe the low levels of DNA triplexes which occur in a specific mammalian sequence and most investigators choose, as their end-point, enhanced mutation rates of genes (Majumdar *et al.*, 1998).

The experimental approach in the current research is to target mammalian DNA sequences which are highly repetitive, the rationale being that if a sequence is present in 500 to 1,000 copies per genome, one might well expect a comparable increase in triplexing over that occurring in a single copy sequence.

For a number of reasons, repetitive sequences in the bovine Y chromosome were chosen for the current study. Firstly, large numbers of bovine repetitive DNA sequences have been sequenced. Secondly, large amounts of bovine tissue are readily available, including sperm. Thirdly, in the latter stages of experimentation it would be possible to carry out a range of spermogenic, embryological and animal manipulations that would not be feasible in humans. Fourth, there is considerable economic interest in techniques which might differentiate between bovine male and female cells.

The main objectives in the present study are the synthesis and testing of a series of psoralen-conjugated oligonucleotides targeted to specific regions of the bovine male Y chromosome, the determination of optimal reaction and UV photoadduction conditions, and the development of assay conditions which will enable the efficiency of triplexing to be measured *in vitro*, both on short bovine Y chromosome specific sequence fragments and, ultimately on bovine Y chromosome sequences in bull DNA samples.

1.7.1 THE Y CHROMOSOME

The size of the bovine Y chromosome is approximately 7 X 10^9 bp (Lewin, 1974). The presence or absence of an entire or functional portion of the Y chromosome determines the primary sex of a mammal (McCarrey & Abbott, 1979). Various gene(s) on the Y-chromosome play important roles in the development of the testes and male phenotype (Gordon & Ruddle, 1981).

Several specific Y-chromosomal DNA sequences have been discovered, many of which are repetitive and conserved among the human, pig, cattle and mouse species (Matthews & Reed 1991). These sequences hybridise preferentially or exclusively to male DNA. The function of these sequences are unknown but they have been very important tools in the study of genetic evolution, the diagnosis of foetal and embryonic sex in humans where there is a risk of X-linked recessive disease (Vergnaud *et al.*, 1984), the predetermination of genetic sex in livestock species, and they also provide an insight into the molecular organisation of the Y chromosome.

A) *Bos taurus* Sex Determining Region Y gene (SRY)

The sex determining region Y gene (*sry*) was identified in 1990 as being implicated in the sex determination of males (Sinclair *et al.*, 1990). It is required to initiate the cascade of steps necessary to form a testis from an undifferentiated gonad. This was established by injecting a transgene into a chromosomal female (XX) mouse embryo which later developed into a male (Koopman *et al.*, 1991). The *sry* gene is a small, intronless gene and is conserved on the Y-chromosome in many mammalian species such as the human, cow, pig and mouse (Payen & Cotionot, 1993; Daneau *et al.*, 1995). The bovine *sry* gene is 1800 bp in size and shows greater resemblance to the human *sry* gene than to the mouse *sry* gene, as do most bovine genes (Daneau *et al.*, 1994).

Expression of the SRY protein only occurs between 10.5 and 12.5 days in the mouse, and is localized in the somatic cells of the genital ridge (Gubbay *et al.*, 1990; Koopman *et al.*, 1990). Limited transcriptional data is available for the *sry* gene but it is known that the SRY transcript is present in the testicles of adult bulls (Daneau *et al.*, 1995).

The *sry* gene encodes a member of the high mobility group protein family (HMG) with an amino acid pattern of 79 which represents the HMG Box. The HMG protein may act as a transcription factor by intercalating with the DNA double helix across the minor groove. This DNA binding capacity is critical to SRY function (van de Wetering & Clevers, 1992).

Nonsense, missense and frameshift mutations have been identified in the *sry* gene of females with 46 XY pure gonadal dysgenesis (Berta *et al.*, 1990; Affara *et al.*, 1993).

B) Bovine Repetitive Y chromosome DNA Sequences

The first Y-associated element known to be repeated exclusively on the male Y-chromosome of a number of different species was discovered in 1991 (Matthews & Reed, 1991). This element, a 307 bp *Sau3A1* DNA probe was isolated from the bovine Y chromosome and named bovine repeat Y-associated 1 (BRY-1). It was found to have a high copy number (up to 500 copies per cell) and be interspersed with other sequences rather than being tandemly repeated. BRY-1 has no significant open reading frames but features short A/T rich tracts (Matthews & Reed, 1991). The DNA sequence was not present in female domestic species but is conserved within male bovine, ovine and caprine species. The primers for the amplification of the BRY-1 region were used to determine the sex of 6-7 day old cattle embryos (Schroder *et al.*, 1990).

A variety of DNA probes such as BTU39891, BTRDYCR, and BTU39887 have now been characterised that specifically hybridise to the Y chromosome (Table 1.2).

BOVINE REPETITIVE Y CHROMOSOME DNA SEQUENCE	SIZE (bp)	SEQUENCE HOMOLOGY	REFERENCE
BTU39891	306	130 bp has homology to many other bovine DNA sequences.	Grobet *et al.*, (1996)
BTRDYCR	489	Similar homology to many human DNA sequences.	Antoniou *et al.*, (1996)
BRY4a	545	Repetitive Y chromosome DNA sequence homology to OY11.1, GRY.1, and BCY11a.	Reed *et al.*, (1995)
BTU39887	561	95% homology to part of the *Bos taurus* DNA repeat region BTRPTDNAF.	GenBank

Table: 1.2: Classification of bovine repetitive Y-chromosomal DNA probes. Refer to Appendix III for DNA target sequence.

C) Testis Specific Protein Y gene (TSPY)

The *Bos taurus* Testis Specific Protein Y gene (TSPY) is specific to the male Y chromosome and is a homologue of the human TSPY gene with respect to almost all aspects of structure and expression (Vogel *et al.*, 1997). It is organised as a repetitive gene family and is composed of a variable number of family members, 20-40 in humans, 50-200 in cattle (Arnemann *et al.*, 1987; Jakubiczka *et al.*, 1993). It is a member of the larger protein superfamily TSNN which includes SET, NAPI and TSPY like genes (Ishimi & Kikuchi, 1991; von Lindern *et al.*, 1992; Schnieders *et al.*, 1996).

The gene is 3179 bp in length and conserved in cattle, ovine and caprine species. The bovine TSPY gene is related to BRY-2, a probe to which BRY-1 binds (Matthews & Reed, 1992).

The function of the TSPY gene remains unknown; however, the cellular pattern of expression suggests a role in spermatogenesis (Schnieders *et al.*, 1996). The expression of TSPY is observed in adult and fetal testes and is apparently restricted to germ cells and their precursors during fetal development (Vogel *et al.*, 1997).

Chapter 2

MATERIALS AND METHODS

2.1 MATERIALS

Glassware:

All glassware was cleaned using Cleanaid detergent (Intermed Scientific Ltd, NZ), rinsed with dH_2O, and autoclaved where possible.

Solutions:

All solutions were autoclaved unless this process would degrade reagents. In these cases solutions were filter sterilised using 0.2 µm filters. All solutions were prepared using double-distilled, deionised water (18 megohm-cm) that was obtained from the Barnstead Water System. Appendix I and II gives information regarding solutions and suppliers.

Bacterial Strains:

Escherichia coli XL1-Blue MRF' cells were obtained from the School of Biological Sciences (Waikato University, NZ).

Plasmid DNA:

The shuttle vector, pSP189 was kindly donated by Dr Peter Glazer (Yale University, USA) and pBluescript was obtained from the School of Biological Sciences (Waikato University, NZ). Appendix III lists all the plasmid DNA sequences.

Bovine Genomic DNA:

Bovine male and female genomic DNA was kindly donated by Professor Richard Wilkins (University of Waikato, NZ) or extracted from Bovine liver or testicle tissue samples supplied by the AgResearch abattoir (Ruakura Agricultural Research Centre, Hamilton, NZ) or Dr Ray Cursons (University of Waikato, NZ). Appendix III lists all the bovine Y-chromosomal DNA sequences.

Primers:

All the primers used in these experiments were synthesized by Life Technologies™ (USA). The lyophilized primers were resuspended in TE buffer (10 mM Tris-HCl, 1 mM ETDA, pH 8.0) to a final concentration (FC) of 200 μM. The stock solution was stored at -20°C and the working solution at 4°C.

Psoralen-conjugated Oligonucletoides:

Psoralen-conjugated oligonucleotides were obtained from Oligo's Etc. (USA). The psoralen was incorporated into the synthesised oligonucleotides as a psoralen phosphoramidite, resulting in an oligonucleotide linked at its 5' end via a 2-carbon arm to psoralen (Figure 2.1). The freeze-dried psoralen-conjugated oligonucleotide was resuspended in TE buffer (pH 8.0), FC = 200 μM. The tube was wrapped in Aluminium foil to exclude light and stored at 4°C. Appendix IV lists the DNA sequence of the psoralen-conjugated oligonucleotides.

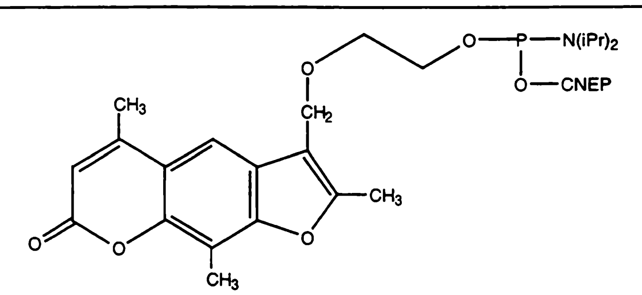


Figure 2.1: Chemical structure of psoralen C2 phosphoramidite. CNEP represents a N, N-diisopropyl-β-cyanoethyl phosphoramidite.

Modified Triplex-forming Oligonucleotides:

Modified TFOs were obtained from Glen Research (USA). These modified TFOs incorporated psoralen at the 5' end as previously described and three phosphorothioate linkages at the 3' end. Cytosine bases were also replaced with 5-methylcytosine residues. The freeze-dried modified TFO was resuspended in TE buffer (pH 8.0), FC = 200 μM. The tube was wrapped in Aluminium foil to exclude light and stored at 4°C. Appendix IV lists the DNA sequence of the modified TFOs.

2.2 LABELLING THE 5' END OF SINGLE STRANDED DNA

The sense strand of the target DNA duplex was 5' end-labelled with T4 polynucleotide kinase and [γ^{33}P] ATP by a forward reaction. The ^{33}P was transferred to a hydroxyl group created by the removal of phosphate from the 5' end of DNA by T4 polynucleotide kinase.

A 20 μL reaction was prepared with the following final concentrations: 5 μM single stranded oligonucleotide primer (24-49 nt), 1X polynucleotide buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT, 0.1 mM spermidine, pH 8.0), 0.5 μCi [γ^{33}P]ATP (3000 Ci/mmol, Amersham, UK), 10 units of T4 polynucleotide kinase (Boehringer Mannheim, Germany). The reaction was incubated at 37°C for one hour followed by heat inactivation for 10 minutes at 65°C. The sample was stored at 4°C for up to two weeks.

2.3 DOUBLE STRANDED DNA FORMATION

The synthetic dsDNA target representing a bovine Y specific repetitive region or gene was obtained by annealing a primer representing the sense strand (38-49 nt) to a small complementary antisense primer (10 nt, refer to Appendix IV). The two synthetic primers were annealed together and the partial duplex was completed by end-filling using the large fragment of *E. coli* DNA polymerase I (Klenow enzyme).

The DNA duplex was radioactively labelled by 5' end-labelling of the sense strand with [γ^{33}P]ATP (Section 2.3.1) or alternatively, incorporating [α^{32}P]dCTP or [α^{32}P]dATP into the antisense strand (Section 2.3.2).

2.3.1 DNA DUPLEX FORMATION USING [γ^{33}P]ATP

A 1 μM solution of γ^{33}P 5' end-labelled sense strand oligomer was mixed with 40 μM of the complementary 10 nt synthetic primer in a final volume of 30 μL and incubated at 65°C for 10 minutes, 37°C for 10 minutes and then cooled to room temperature (RT) for 10 minutes. The annealed fragment was incubated in 1X Triplex buffer (10 mM $MgCl_2$, 10 mM Tris-Cl, pH 7.1), 50 μM dNTP, 2 units of Klenow enzyme (Boehringer Mannhein, Germany) in a final volume of 100 μL for one hour at 37°C (the Klenow enzyme is active under a variety of buffer conditions and does not require a specific buffer). Following incubation, the Klenow enzyme was heat inactivated for 10 minutes at 65°C and then the DNA duplex stored at 4°C for up to two weeks.

The formation of dsDNA was confirmed by electrophoresing a sample of ssDNA against dsDNA on a native 10% PAGE gel (Section 2.3.3).

Alternatively, DNA duplexes were prepared using [γ^{32}P]ATP. This was achieved by using the same protocol as previously described in Chapter 2.2 and 2.3.1 but replacing [γ^{33}P]ATP with [γ^{32}P]ATP.

2.3.2 DNA DUPLEX FORMATION USING [α^{32}P]dCTP

A 2 μM solution of sense strand was mixed with 6 μM of the complementary 10 nt synthetic primer in a final volume of 8 μL, and incubated at 65°C for 10 minutes, 37°C for 10 minutes and then cooled to RT for 10 minutes. The annealed fragment was incubated in 1X Triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1), 0.5 μCi [α^{32}P]dCTP (3000 Ci/mmol, Amersham, UK), 100 μM dATP, 100 μM dGTP, 100 μM dTTP, 2 units of Klenow enzyme (Boehringer Mannhein, Germany) in a final volume of 20 μL for 15 minutes at RT.

Finally, 10 μM dCTP was added and incubation continued for a further two hours at RT to ensure complete synthesis of the DNA duplex. Following incubation, the Klenow enzyme was heat inactivated for 10 minutes at 65°C and then the DNA duplex was stored at 4°C for up to two weeks.

The formation of dsDNA was confirmed by electrophoresing a sample of ssDNA against dsDNA on a native 10% PAGE gel (Section 2.3.3).

2.3.3 PAGE ELECTROPHORESIS

PAGE electrophoresis was performed using the Model S2 Sequencing Gel Electrophoresis System (Life Technologies, USA) for large polyacrylamide gels (330 cm X 410 cm X 0.4 mm), or alternatively the Penguin™ Water-cooled Dual-Gel Electrophoresis System (OWL Scientific Inc, USA) for mini polyacrylamide gels (10 cm X 10 cm X 0.8 mm). Preparation of native and denaturing TBE polyacrylamide gels is described in Appendix II. One times gel loading dye (Type III: 40% glycerol, 0.25% bromophenol blue (BPB), 0.25% xylene cyanole FF (XC), 59.5% Tris-50 mM) was added to DNA samples before loading onto native polyacrylamide gels. Alternatively 2X Formamide DNA sequencing dye (98% deionized Formamide, 0.025% XC, 0.025% BPB, 10 mM EDTA pH 8.0) was used for denaturing polyacrylamide gels.

In general, PAGE electrophoresis was performed at constant power, 30 W for native gels and 60 W for denaturing gels. Following PAGE electrophoresis, the gel was removed from the glass plates and dried under vacuum (80°C for denaturing gels, 50°C for native gels). The dried radioactive gel was exposed to Kodak X-OMAT film, and stored in an autoradiograph cassette at RT (up to one week) and then developed by autoradiography. Radioactive gels containing $\alpha^{32}P$ were exposed to an intensifying screen and stored at -70°C (for up to 1 week).

2.3.4 AUTORADIOGRAPHY

A manual Kodak GBX X-Ray Film Developing System (Kodak, Australia) was used in the Dark Room to develop autoradiograph films. The protocol was followed as recommended by the manufacturer.

2.4 PURIFICATION OF OLIGONUCLEOTIDES

Unincorporated [α^{32}P]ATP, [γ^{33}P]ATP and dNTPs were removed from the radioactively labelled oligonucleotide by using a G25 Sephadex column or alternatively by isolating the dsDNA fragment from a polyacrylamide gel using a "Crush and soak" method.

2.4.1 G25 SEPHADEX COLUMN

This method was modified from the G50 Sephadex column method of Maniatis *et al.*, (1982).

A) Preparation of G25 Sephadex Slurry

One gram of G25 Fine Sephadex was carefully combined with 10 mL TE Buffer (pH 8.0) in a sterile 50 mL universal tube. The mixture was incubated overnight at RT to allow the beads to swell. The tube was then equilibrated three times with 1X TEN buffer (10 mM Tris (pH 8.0), 1 mM EDTA (pH 8.0), 100 mM NaCl), autoclaved, and stored at RT for future use.

B) Preparation of G25 Spin Column

A small hole was pierced in the base of a 0.6 mL microcentrifuge tube which was then placed inside a 1.5 mL microcentrifuge tube. Glass wool was placed inside the smaller tube. Sterile G25 Sephadex slurry was added and the tubes centrifuged in a C5145 centrifuge for 1000 g for 2 minutes at RT.

More sterile G25 slurry was added and centrifuged until the Sephadex was compacted and 0.5 cm from the top of the tube. A 100 μL volume of 1X TE Buffer (pH 8.0) was then added to the column and centrifuged as before. This was repeated three times to equilibrate the column. Finally the column was placed in a sterile 1.5 mL microcentrifuge tube, wrapped with parafilm, and stored at RT until required for DNA purification.

C) DNA Purification using a G25 Sephadex Spin Column

A 100 μL sample of the labelling reaction was placed into the G25 column and centrifuged in a C5145 centrifuge at 1000 g for 5 minutes, at RT. Most of the labelled oligonucleotide was recovered in the 100 μL eluate (in TE buffer). This was stored at 4°C. The extent of purification was checked by native PAGE electrophoresis (Section 2.3.3) to confirm that unincorporated [γ^{33}P]ATP or [α^{32}P]ATP and dNTPs were removed from the radioactively labelled DNA.

2.4.2 "CRUSH AND SOAK METHOD"

This method was originally described by Maxam and Gilbert (1977) and modified by Maniatis *et al.*, (1989).

A 5 μL sample of radioactive DNA was mixed with 5 μL of 1X gel loading dye (Type III), loaded onto a 6% native PAGE gel and electrophoresed for one hour with pre-chilled 1X TBE buffer at 1000 volts, 30 watts, RT. Following electrophoresis, the gel was covered with plastic wrap and exposed to Kodak X-OMAT AR film for one hour at RT.

The DNA band of interest was located by autoradiography (Section 2.3.4) and was cut from the gel using a sharp scalpel blade. The gel was transferred to a 0.6 mL microcentrifuge tube and crushed using a sterile disposable pipette tip. Two volumes of elution buffer (0.5 M NH_4Ac, 10 mM MgAc, 1 mM EDTA [pH 8.0.], 0.1% SDS) were then added to the crushed gel, followed by overnight rotation at 37°C.

Following the overnight incubation, the sample was centrifuged for 1 minute, 12,000 g at RT to pellet the polyacrylamide. The supernatant was transferred to a new 1.5 mL microcentrifuge tube and two volumes of absolute ethanol were added. The solution was stored on ice for 30 minutes and the DNA was recovered by centrifugation as before but for 15 minutes. The DNA pellet was redissolved in 20 µL TE buffer (pH 8.0). The amount and quality of the purified DNA was confirmed by native PAGE gel electrophoresis (Section 2.3.3).

2.5 RADIOACTIVE LABELLING OF THE 3' END OF PSORALEN-CONJUGATED OLIGONUCLEOTIDES

This method was originally described by Deng & Wu (1983) and was modified to radioactively label 5' psoralen-conjugated oligonucleotides at the 3' end using terminal transferase. A 50 µL reaction was prepared containing 1X terminal transferase buffer (200 µM sodium cacodylate, 0.25 mg/mL Bovine Serum Albumin, 25 mM Tris-HCl, pH 6.6), 50 µM psoralen-conjugated oligonucleotide, 2.5 mM $CoCl_2$, 0.5 µCi [α^{32}P] dCTP (3000 Ci/mmol, Amersham, UK), and 0.5 unit of terminal transferase (Boehringer Mannheim, Germany).

The reaction was covered with Aluminium foil to minimize the amount of light exposure to the psoralen-conjugated oligonucleotide. The reaction was incubated for 30 minutes in the dark at 30°C followed by heat inactivation at 65°C for 10 minutes. The reaction was stored in the dark at 4°C for up to 2 weeks. The labelling of the psoralen-conjugated TFO at the 3' end was confirmed by native PAGE electrophoresis (Section 2.3.3). The final product only contained one incorporated base, either C, A or G. The base chosen depended on the DNA sequence of the oligonucleotide and its dsDNA target.

2.6 *IN VITRO* DNA TRIPLEX BINDING ASSAYS USING SMALL SYNTHETIC DNA FRAGMENTS

DNA triplex binding assays were performed in 200 μL UltraFlux Flat Cap PCR tubes. A master mix was prepared containing: approximately 100 nM radioactive duplex DNA, 100 μM psoralen-conjugated oligonucleotide, 1X Triplex buffer (10 mM $MgCl_2$, 100 mM Tris-HCl, pH 7.1), 1 mM spermidine. In the control tubes the psoralen-conjugated oligonucleotide was replaced with the same volume of H_2O. Two microlitre samples were aliquoted into 200 μL tubes and centrifuged for 10 seconds, 18,000 g at RT. The tubes were incubated for two hours at 37°C in the dark (wrapped in Aluminium foil).

Following the two hour incubation period, the tubes containing the 2 μL DNA triplex samples were placed in the cold room (4°C) on top of a plastic petri dish filled with ice and covered with a sheet of black paper to minimize the amount of reflection (3 cm away from a 360 nm Fluorescent lamp emitting UVA light, 320-400 nm, Osram L18W/73).

A plastic OHP mylar sheet (Visual Systems Products, France) was wrapped around the UV Fluorescent lamp to remove the UVB light (Figure 2.2). The samples were irradiated for a specified time to generate photoadducts (0 - 30 minutes, maximum UV dose of 2.0 J/cm^2). The samples were then stored in the dark at 4°C until required for further analysis (up to 2 days).

The level of photoadduct formation was measured by electrophoretic gel mobility shift assays (EMSA). A 4 μL volume of Formamide DNA sequencing dye was mixed with the 2 μL DNA triplex samples and then the samples were denatured for 5 minutes at 90°C in a Techne Progene PCR machine before loading a 4 μL sample onto a pre-electrophoresed (constant power, 90 W, approx 1700 V for 15 minutes) TBE denaturing 8% PAGE sequencing gel.

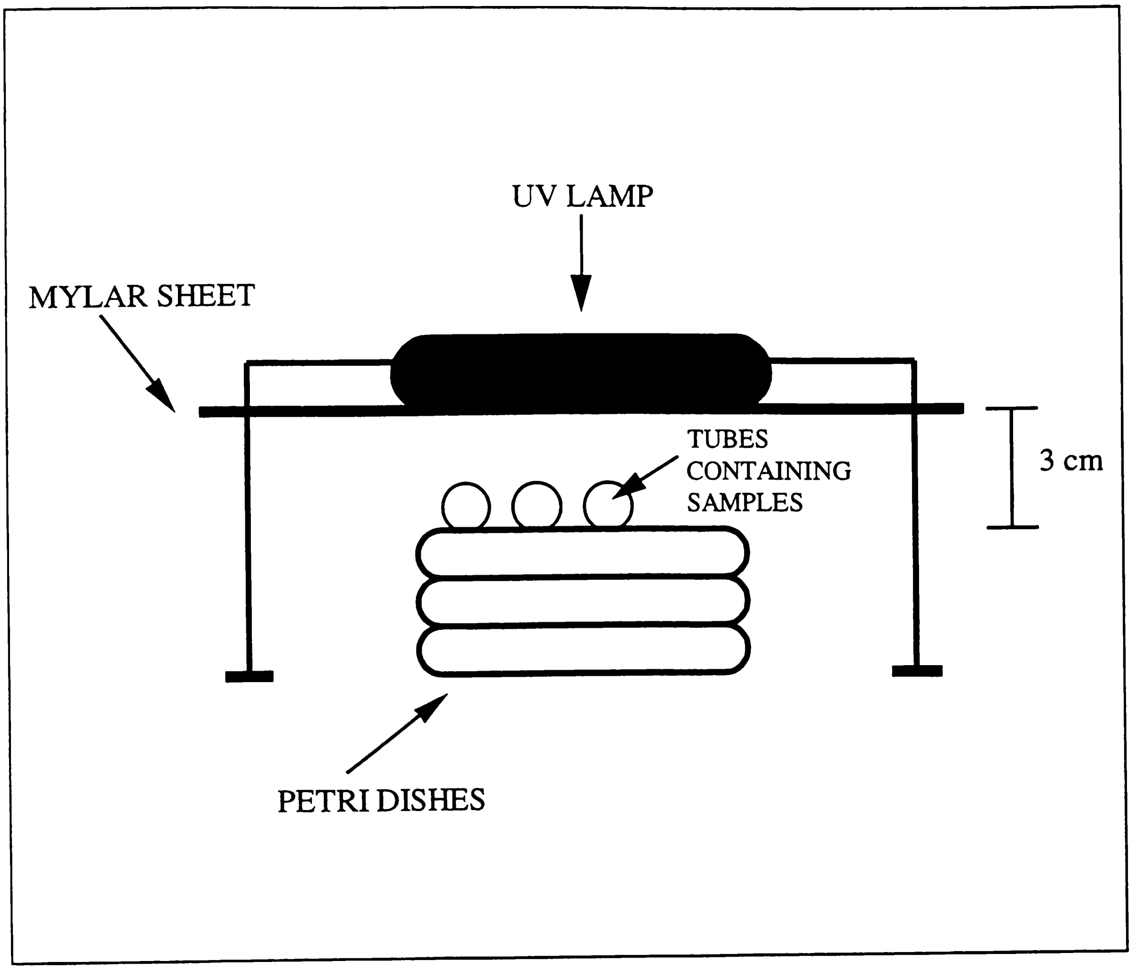


Figure 2.2: A schematic diagram showing the procedure used to irradiate DNA samples to generate photoadducts (Diagram is not to scale).

PAGE electrophoresis was performed at constant power (approx 60 W) until the bromophenol blue dye (indicates the distance a 19 nt fragment would travel) was 15 cm away from the bottom of the gel. The gel was removed from the glass plates, dried under vacuum at 80°C, exposed to Kodak X-OMAT AR film at RT and developed by autoradiography (Section 2.3.4).

2.7 STAINING POLYACRYLAMIDE GELS

A mini native 20% PAGE gel was electrophoresed and stained with Stains-All (Eastman, USA) to visualise small radioactive and nonradioactive oligonucleotides (8-40 nt). This method was obtained from Applied Biosystems User Bulletin, 1987.

A 5 µL sample containing radioactive or non radioactive oligonucleotides was diluted to a final concentration of 100 µM with 1X gel loading dye (Type III) and loaded onto a mini native 20% PAGE gel containing 6 mM Formic acid, 0.168 mM Tris (pH 9.0). In the case of 5' psoralen-conjugated oligonucleotides, these were diluted to a final concentration of 0.3 µM. Electrophoresis was performed using a Tris-tricine electrophoresis buffer (0.5 M Tris, 100 mM Tricine, pH 8.6) at 120 volts for one hour at RT. Following electrophoresis, the gel was removed from the glass plates and soaked in Stains-All solution (50 mg Stains-All dye in 400 mL of 1:1 v/v deionized H_2O and deionized Formamide) for 45 minutes, in the dark at RT. The gel was destained by rocking the gel in deionized H_2O until the purple background diminished revealing stained bands. The stained gel was photographed and then vacuum dried between two pieces of Whatman N#4 paper at 37°C to minimise cracking. The gel was then transferred to a plastic bag and sealed. Radioactive gels were exposed to Kodak X-OMAT AR film, stored at RT (up to 1 week) and developed by autoradiography as described in Section 2.3.4.

2.8 SOUTHERN BLOTTING

2.8.1 PREPARATION OF DNA FROM EUKARYOTIC TISSUE

This method was used to prepare high purity genomic DNA from eukaryotic tissue. One gram of fresh tissue (liver, testicle) was cut into lumps using a sterile blade and stored for a few hours in Aluminium foil at -70°C. Alternatively, the tissues were frozen using liquid nitrogen.

The frozen tissue was placed into a pouch of Mylar sheet and crushed using a hammer. One millilitre of ETS buffer solution (10 mM EDTA [pH 8.0], 25 mM Tris-HCl [pH 8.0], 0.5% SDS) was added to the pouch and mixed with the tissue using a sterile pipette tip. The solution was transferred to a 50 mL Falcon tube and made up to a final volume of 15 mL with ETS buffer. Proteinase K was added to a final concentration of 100 µg/mL before incubation with rotation overnight at 60°C in a Mini MK11 Hybaid Hybridisation oven. Any large lumps remaining after the Proteinase K step were removed by centrifugation for 5 minutes, 1000 g at RT. The supernatant was removed and extracted with an equal volume of Tris-saturated phenol (pH 8.0). The sample was mixed by rocking for 10 minutes at RT and then the phases were separated by spinning the samples for 10 minutes, 2000 g at RT. The procedure was repeated for one phenol:chloroform extraction, one chloroform extraction, and finally, one diethyl ether extraction. Following the diethyl ether extraction the top phase was discarded and the tube was left open to evaporate the ether at 37°C for 5 minutes.

Two and a half volumes of absolute ethanol was gently layered on top of the lysate and a sealed glass Pasteur pipette was used to wind and collect the white gelatinous DNA at the interphase. The DNA on the Pasteur pipette was then washed two times by dipping it into a tube containing 5 mL of ice cold 70% ethanol for one minute. The ethanol was washed off by dipping the Pasteur pipette into 5 mL of TE buffer (pH 8.0) for 5 seconds. The Pasteur pipette end containing the moist DNA was broken off into a sterile 1.5 mL tube.

The genomic DNA was resuspended in 1 mL of TE buffer (pH 8.0) and dissolved into solution, overnight at RT, and then stored at 4°C until further use. The quantity and purity of genomic DNA isolated was confirmed by agarose gel electrophoresis using a DNA mass ladder as a reference and/or spectrophotometry (Refer to Appendix VI). Typically, 1.5 mg of pure DNA was extracted from 1.0 g of bovine tissue.

2.8.2 RESTRICTION DIGEST OF GENOMIC BOVINE DNA

Genomic DNA was digested overnight at 37°C with a restriction enzyme that did not cut the Hybridisation probe. Each restriction digest reaction was set up in a 600 μL microcentrifuge tube containing 1 μg bovine male or female genomic DNA, 1X restriction enzyme (RE) buffer H (50 mM Tris-HCl, 10 mM MgAc, 100 mM NaCl, 1 mM DTE, pH 7.5), 10 units of *Eco*RI, and dH_20 to a final volume of 50 μL. The control reaction contained no restriction enzyme and was replaced with the same volume of H_20. After digestion the reactions were inactivated by heat for 10 minutes at 65°C and then stored at 4°C until required for agarose gel electrophoresis.

2.8.3 AGAROSE GEL ELECTROPHORESIS

A 25 μL sample of digested bovine genomic DNA was mixed with 5 μL of 6X gel loading dye and loaded onto a 1% mini TAE agarose gel containing 0.5 μg/mL Ethidium bromide (EtBr). The samples were electrophoresed in the Horizon 11.14 GibcoBRL Horizontal Gel Electrophoresis Apparatus (Life Technologies, USA) at 20 volts, overnight at RT alongside molecular weight DNA markers. The gel was visualized by UV illumination and photographed using *Scion Image*, an image processing and analysis program on an IBM computer.

2.8.4 BLOTTING OF THE AGAROSE GEL

Following agarose gel electrophoresis, the gel was denatured in 0.5 M NaOH, 1.5 M NaCl for 10 minutes while rocking at RT. The solution was changed so that total denaturing time was 30 minutes, and then the gel was then placed on top of an upside down plastic gel plate. Hybond™-N+Nylon membrane (Amersham, UK) was cut to the appropriate gel size, soaked in salt solution and carefully placed on top of the gel to avoid air bubbles.

Eight pieces of Whatman 4 paper (11 x 14 cm) soaked in the salt solution were placed on top of the membrane followed by eight pieces of Beermat (11 x 14 cm), a stack of paper towels, and a 50 gram paper weight (Figure 2.3). A wick was not used in the Southern blot as the liquid from the gel and paper was sufficient for capillary action. The Southern blot was left overnight at RT to allow maximum transfer of DNA onto the membrane.

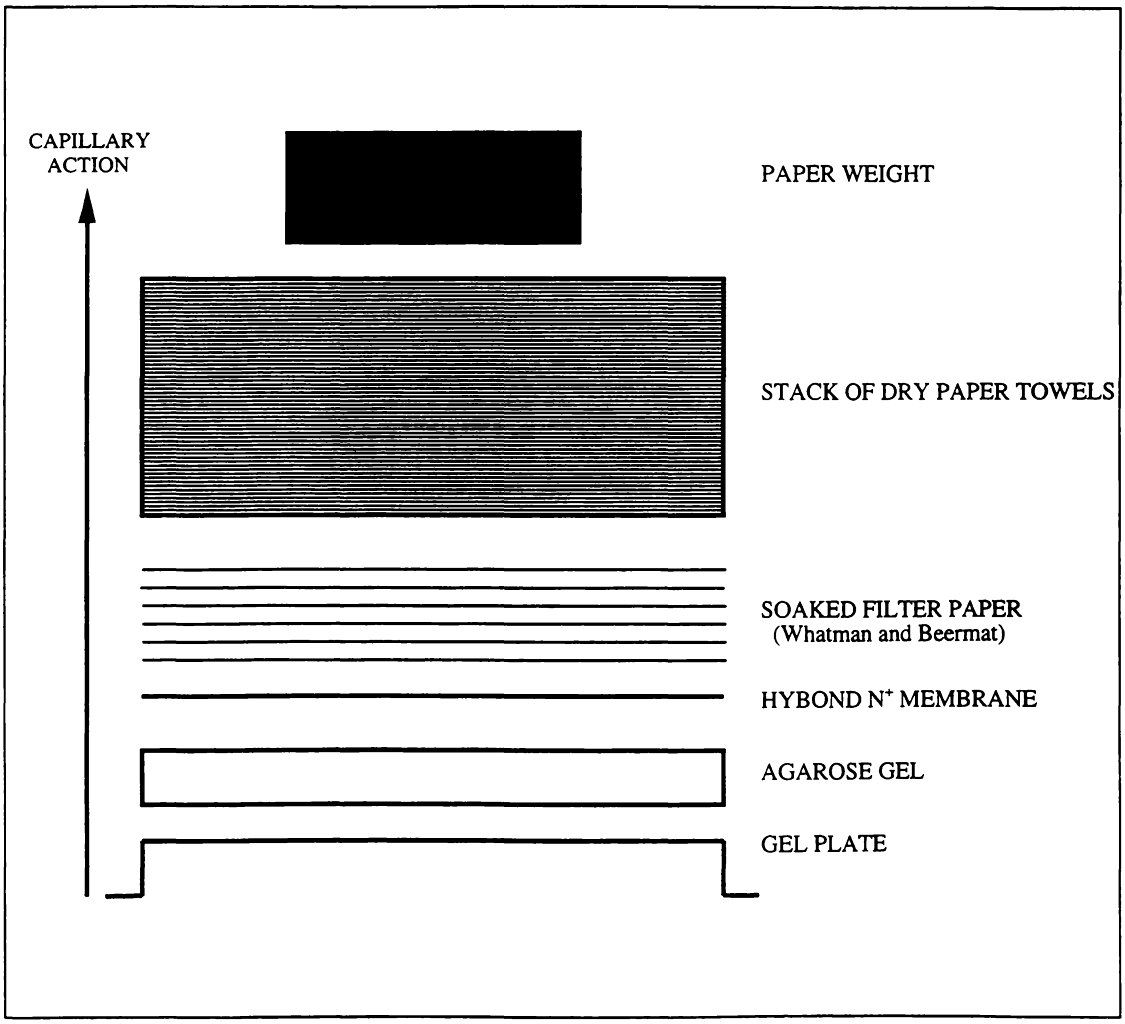


Figure 2.3: A schematic diagram showing the organization of layered materials for a Southern blot.

The next day the paperweight, Beermat and Whatman paper were removed from the membrane. The membrane was carefully removed from the gel and soaked in Church and Gilbert Hybridisation solution (0.5 M Na_2PO_4, 7% SDS, Church & Gilbert, 1984) for 5 minutes at RT. The membrane was then placed onto a thin plastic sheet and placed on top of a UV transilluminator (GIBCO BRL TFX-35M) with the DNA side facing upwards and irradiated for thirty seconds to irreversibly fix the ssDNA to the membrane surface.

Alternatively, the membrane was placed inside a UV crosslinker (Life Technologies, Model BLX 254, USA) and irradiated with 120 mJ/cm^2 of 254 nm UV light. The membrane was then transferred to a 200 mL Hybridisation tube containing 10 mL Church & Gilbert (C&G) Hybridisation solution. The Hybridisation tube was rotated in the Hybaid Mini MKII Hybridisation oven at a specified hybridisation temperature for one hour (Table 2.1).

2.8.5 PREPARATION OF RADIOACTIVE DNA PROBES BY PCR

The following method was modified from Stuerzl & Roth (1990) by Richard Wilkins (pers.comm).

Ten microlitres of a 2X concentrated reaction mix was prepared containing the normal PCR constituents (PCR Buffer (20 mM Tris-HCl [pH 8.0], 3 mM $MgCl_2$, 100 mM KCl), 400 μM dATP, dGTP, dTTP and 10 μM of dCTP, 0.5 U *Taq* DNA Polymerase (Boehringer Mannheim, Germany), 1.6 μM forward primer (F), 1.6 μM reverse primer (R), 20 ng template DNA) from a previous PCR reaction (Section 2.9).

A 2.5 μL sample of this 2X reaction mix was carefully combined with 2.5 μL of [α^{32}P]dCTP (3000 Ci/mmol, Amersham, UK) in a 200 μL UltraFlux Dome Cap PCR tube. Fifteen microlitres of mineral oil (Sigma, USA) was carefully layered on top of the 5 μL sample and PCR was performed as for the original PCR but for 15 cycles only.

Alternatively, the Southern blot probe was prepared using the Rediprime™ DNA labelling system (Amersham Life Science, UK). The protocol recommended by the manufacturers was followed exactly.

2.8.6 PREPARATION OF SMALL OLIGONUCLEOTIDE PROBES

Small oligonucleotide probes were 5' end-labelled with T4 polynucleotide kinase and [γ^{32}P]ATP. A 20 µL reaction was prepared with the following final concentrations: 1 µM single stranded oligonucleotide primer (16-49 nt, Table 2.1, Appendix V), 1X polynucleotide buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT, 0.1 mM spermidine, pH 8.0), 12.5 µCi [γ^{32}P]ATP (3000 Ci/mmol, Amersham, UK), 10 units of T4 polynucleotide kinase (Boehringer Mannheim, Germany). The reaction was incubated at 37°C for 2 hours followed by heat inactivation for 10 minutes at 65°C. The sample was stored at 4°C for up to two weeks.

2.8.7 HYBRIDISATION

Radioactive probes prepared as outlined in Section 2.8.5 were denatured with NaOH (to a final concentration of 1 M) and then left at RT for 5 minutes. The C&G Hybridisation solution was removed and replaced with 4 mL of fresh C&G and the tube was rotated in the Hybaid Mini MKII Hyaid Hybridisation oven at the specified temperature. One millilitre of C&G solution was mixed with the denatured probe and this solution was then added to the tube containing the prewarmed 4 mL C&G solution and membrane. The Hybridisation tube was rotated overnight in the Hybaid Mini MKII Hybridisation oven (Table 2.1 for hybridisation temperature).

RADIOACTIVE PROBE	HYBRIDISATION TEMPERATURE (°C)	PROBE SIZE (nt)
Bov*REP3*	60	100
Bov*SRY7*	65	200
Bov*TSPY*	65	500
Bov*REED*	65	400
Bov*USPI*	68	173
USPI	RT, 37	22
38Bov*REP3*	RT, 37	38
30Bov*REP3*	RT, 37	30
16Bov*REP3*	RT	16
49Bov*REED*	RT, 37	49
30Bov*REED*	RT, 37	30
16Bov*REED*	RT	16
16Bov*SRY*	RT	16

Table 2.1: Southern blot hybridisation temperatures and nucleotide size of each radioactive probe. RT = room temperature. Appendix III and V lists the complete DNA sequence of all the above radioactive probes.

The next day the probe was removed into a 50 mL Falcon tube and stored at 4°C for future use (up to 2 weeks). The membrane was washed three times with 3X SSC (0.45 M NaCl, 0.045 M sodium citrate) containing 0.1% SDS at an optimised temperature (Table 2.1). Alternatively, the membrane was washed until low radioactive counts were detected by the Geiger-Muller detector.

The membrane was placed into a thin plastic bag, sealed, and transferred to a cassette. A small adhesive label containing a 0.5 µL radioactive sample was placed near the membrane for orientation. The membrane was then exposed to Kodak X-OMAT AR film using an intensifying screen at -70°C and developed by autoradiography (Section 2.3.4).

2.9 POLYMERASE CHAIN REACTION (PCR)

A forward and reverse primer (Life Technologies, USA) were constructed to amplify a specific region of the bovine Y chromosome containing the DNA triplex site. The size of the PCR product was limited to less than 500 bp, thereby allowing photoadducts to be detected by PAGE electrophoresis. The primer DNA sequence and availability of restriction enzyme sites were also considered.

A 25 µL PCR reaction mixture was set up in sterile UltraFlux Flat Cap PCR Tubes on ice with the following final concentrations: 200 µM dNTPs; 0.8 µM of the specified F and R primer (Table 2.2, 2.3); 10 ng genomic male bovine DNA; 1X PCR buffer containing 10 mM Tris-HCl [pH 8.0], 1.5 mM $MgCl_2$, 50 mM KCl; 1 unit of *Taq* DNA polymerase (Boehringer Mannheim, Germany) and sterile dH_2O. A negative control was set up which contained no bovine genomic DNA as well as a positive control containing 10 ng amplified bovine male genomic DNA. A control containing bovine female genomic DNA was also set up. The tubes were then sealed and centrifuged 10 seconds, 12,000 g at RT.

The bovine DNA template was replaced with the shuttle vector pSP189 DNA (Yale University, USA) when PCR was performed to amplify the *sup*F PCR fragment.

The tubes were placed in the Techne Progene PCR machine and PCR was performed with the following optimal conditions: 94°C for 2 minutes to denature the template, 25 cycles of PCR amplification were performed with denaturation at 94°C for 30 seconds, annealing at a specified temperature for one minute (Table 2.2), and DNA synthesis at 72°C for one minute.

A final extension step was carried out at 72°C for 10 minutes to allow additional DNA synthesis and then the tubes were cooled to 4°C. Thirty cycles of PCR were performed for Bov*SRY* and Bov*REP3*. The amplified DNA was stored at - 20°C.

PCR amplification was confirmed by agarose gel electrophoresis (Appendix VI). Once the optimal conditions for PCR had been determined, [α^{32}P]dCTP was included in the PCR reaction for future *in vitro* DNA triplex binding assays. For PCR reactions with "cold" dNTPs and [α^{32}P]dCTP, a final concentration of 200 μM dGTP, dTTP, dATP, 10 μM dCTP and 0.4 μCi [α^{32}P]dCTP (3000 Ci/mmol, Amersham, UK) was used. PCR reactions were performed as before but the annealing temperature was adjusted for optimal amplification as shown in Table 2.2.

	ANNEALING TEMPERATURE		PRIMER NAME	
DNA	NONRADIOACTIVE PCR (°C)	RADIOACTIVE PCR (°C)	FORWARD PRIMER	REVERSE PRIMER
Bov*REP3*	53	56	FBov*REP3*	RBov*REP3*
Bov*SRY*	57	63	FBov*SRY*	RBov*SRY*
Bov*REED*	56	58	FBov*REED*	USP2
Bov*TSPY*	56	58	FBov*TSPY*	RBov*TSPY*
*sup*F	56	N/A	F*sup*F	R*sup*F
Bov*USPI*	57	N/A	USP1	USP2

Table 2.2: Annealing temperatures of the PCR primers during radioactive and nonradioactive PCR.

PRIMER NAME	PRIMER DNA SEQUENCE
F*sup*F	5' GAATTCGAGAGCCCTGCTC 3'
R*sup*F	5' GGAGTCAGTCGAGCTCACG 3'
FBov*REP3*	5' GATTCCAGGGCTGCATATCATA 3'
RBov*REP3*	5' CGAGGGAAGAATCAAGTAAGGC 3'
FBov*SRY*	5' GAACGACGATGTTTACAGTCCAGCT 3'
RBov*SRY*	5' CCTTTCGTCTTCGTTCACGAGACCA 3'
FBov*REED*	5' CCCACAGAAGGCAATCTGATTG 3'
USP1	5' CCCTTCCAGCTGCAGTGTCA 3'
USP2	5' GATCTGTAACTGCAAACCTGGC 3'
FBov*TSPY*	5' ATGCCCTGCAGGCTGTGGCCG 3'
RBov*TSPY*	5' TCACCCACATTTACACCAATA 3'

Table 2.3: DNA sequence of forward and reverse PCR primers.

Following PCR, psoralen-conjugated oligonucleotides were targeted to the amplified radioactive dsDNA product, generating photoadducts upon UV irradiation (360 nm). This was the same procedure that was described in Section 2.7 except that the radioactive DNA duplex was replaced with radioactive amplified DNA.

2.10 PURIFICATION OF AMPLIFED DNA

The Promega Wizard™ PCR Preps DNA purification system was used for rapid purification of double stranded amplified DNA. The manufacturer's protocol was followed exactly. Alternatively, PCR products were purified using G25 Sephadex columns (Section 2.4.1).

2.11 PCR CLONING

Two cloning methods were employed to ligate PCR products into the Plasmid Bluescript® II SK^+ shuttle vector (Stratagene, USA). This vector is a 2961 bp phagemid with an appropriate multiple cloning site for genetic manipulation (Refer to Appendix III).

2.11.1 CLONING PCR PRODUCTS USING A T-TAILED VECTOR

A) Preparation of PCR Products for Cloning

A 25 µL PCR sample was mixed with 5 µL of 6X gel loading dye and loaded onto a 0.8% mini LMP (SeaPlaque®, FMC Bioproducts) TAE agarose gel containing 0.5 µg/mL EtBr. The sample was electrophoresed for one hour, 120 volts at RT. Following electrophoresis, the bands were visualised on a UV transilluminator that was covered with plastic wrap to minimise UV damage (Gibco BRL, model TFX-335M). The desired band was cut out in the smallest possible volume of agarose using a clean razor blade, transferred to a 1.5 mL tube, and stored at 4°C.

B) Preparation of T-Tailed Vector for Cloning

Ten micrograms of Plasmid Bluescript®II SK^+ was digested overnight at 37°C with 100 units *Eco*RV (Boehringer Mannheim, Germany), 1X RE buffer B (10 mM Tris-HCl, 5 mM $MgCl_2$, 100 mm NaCl, 1 mM 2-mercaptoethanol) and dH_20 to a final volume of 85 µL.

The reaction was then heat inactivated at 65°C for 10 minutes and extracted with an equal volume of Tris-saturated phenol (pH 8.0). The aqueous phase was then extracted with an equal volume of chloroform. Plasmid DNA was precipitated with 0.1 volume of 7.5 M ammonium acetate and 2.5 volumes of ice cold absolute ethanol at -20°C for 15 minutes. The DNA was recovered by centrifugation for 15 minutes, 12,000 g at RT.

The DNA precipitate was resuspended in 10 µL 10X PCR buffer (10 mM Tris-HCl [pH 8.0], 1.5 mM $MgCl_2$, 50 mM KCl), 86.5 µL H_2O, 1 µL 100 mM dTTP, 2.5 µL 1U/µL *Taq* DNA polymerase (Boehringer Mannheim, Germany), and incubated for 13 hours at 73°C. Following the incubation, 15 µL of the reaction was mixed with 5 µL 6X gel loading dye and loaded onto a 0.7% LMP (SeaPlaque®, FMC Bioproducts) TAE mini agarose gel containing 0.5 µg/mL EtBr. The remaining electrophoresis and band removal protocol was continued as described earlier in Section 2.11A. The agarose band containing the T-tailed vector was stored at 4°C.

C) Cloning PCR Products into a T-Tailed Vector

The LMP agarose *Eco*RV digested T-tailed plasmid and LMP agarose PCR product were melted at 65°C for 10 minutes. Five microlitres of each were mixed in a sterile 600 µL microcentrifuge tube and transferred to a 37°C water bath for 5 minutes. 10 µL of a 2X T4 DNA ligase reaction mix (2X T4 DNA ligase buffer [20 mM Tris-HCl, 1 mM EDTA, 60 mM DTT, 60 mM KCl, 50% glycerol], 2 units T4 DNA ligase [Boehringer Mannheim, Germany], dH_2O to 20 µL) was then added to each tube. The contents were mixed and incubated overnight at RT to allow the ligation of the PCR product and vector DNA. A control reaction contained 5 µL of melted agarose containing *Eco*RV digested T-tailed plasmid, 5 µL dH_2O and 10 µL of 2X T4 DNA ligase reaction mix.

2.11.2 BLUNT-ENDED CLONING

A) Preparation of PCR Fragments for Blunt-ended Cloning

The PCR products were first phosphorylated to enable ligation to a dephosphorylated Bluescript®II SK$^+$ shuttle vector. Fifteen microlitres of purified PCR DNA (Section 2.11), 2 µL of 10X T4 polynucleotide kinase buffer (50 mM DTT, 700 mM Tri-HCl, and 100 mM $MgCl_2$) and 2 µL of 10 mM ATP were incubated at 37°C for 5 minutes. Ten units of T4 polynucleotide kinase (Boehringer Mannheim, Germany) was added and the reaction incubated for 30 minutes at 37°C. The reaction was heat inactivated for 15 minutes at 70°C, then cooled to RT, purified using the Promega Wizard PCR purification system (Section 2.10), and stored at 4°C till further use.

B) Preparation of Vector DNA for Blunt-ended Cloning

A 50 µL restriction digest containing 5 - 20 µg Bluescript®II SK$^+$ plasmid DNA, 30 µL sterile dH_2O, and 7.5 µL 10X RE buffer A (330 mM Tris-acetate, 100 mM MgAc, 660 mM KAc, 50 mM DTT) and 25 units of *Sma*I (Boehringer Mannheim, Germany) was incubated overnight at RT.

The extent of *Sma*I digestion was checked by agarose gel electrophoresis (Appendix VI). When digestion was complete, the restriction digest was incubated for 10 minutes at 70°C to inactivate *Sma*I, and then stored on ice for 10 minutes. One tenth volume of 3M sodium acetate (pH 5.0) and 2.5 volumes of absolute ethanol were added to precipitate the DNA. The sample was stored at -20°C for 30 minutes and then the DNA was recovered by centrifugation at 14,000 g for 15 minutes at RT. The DNA was resuspended in 44 µL sterile dH_2O and stored on ice for 10 minutes.

Five microlitres of 10X dephosphorylation buffer (0.5 M Tris-HCl, 1 mM EDTA, pH 8.5) and 1 µL of 1 U/µL calf intestine alkaline phosphatase (CAP, Boehringer Mannheim, Germany) were added to the purified *Sma*I digested plasmid DNA. The sample was incubated at 30°C for 30 minutes, and then 50°C for 30 minutes to dephosphorylate the linear vector.

The dephosphorylated vector was then extracted twice with an equal volume of Tris-saturated phenol (pH 8.0) and centrifuged at 5,000 g for 5 minutes at RT.

The aqueous phase was extracted with an equal volume of chloroform and centrifuged as before. The aqueous phase was transferred to a sterile 1.5 mL microcentrifuge tube and 0.1 volumes of 3 M sodium acetate (pH 5.0) and 2.5 volumes of absolute ethanol were added. The sample was incubated at -20°C for 1 hour. The DNA was recovered by centrifugation at 15,000 g for 15 minutes at RT. The DNA pellet was washed with 1 mL of 70% ethanol and recentrifuged as before, but for 5 minutes only. The DNA pellet was air dried by placing the tube in a 37°C heating block and then resuspended in 25 µL sterile dH_20. The quantity and purity of DNA isolated was confirmed by agarose gel electorphoresis (Refer to Appendix VI) and the remaining DNA was stored at -20°C.

C) Ligation of PCR Products and Plasmid DNA

A 5 µL ligation reaction was prepared containing the following: 1 µL 5X T4 DNA ligase buffer (250 mM Tris-HCl [pH 7.6], 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25% w/v PEG 8000), 1 µL dephosphorylated plasmid DNA, 2.5 µL phosphorylated amplified DNA, and 0.5 units T4 DNA ligase (Boehringer Mannheim, Germany). Two ligation controls were also prepared; one reaction had no amplified DNA and the other reaction had no plasmid DNA present.

The reactions were incubated overnight at RT and then 1.25 µL of 1 mg/mL Proteinase K containing 10 mM EDTA (pH 8.0) and 10 mM Tris.HCl (pH 8.0) was added. The reactions were incubated for 15 minutes at 70°C to inactivate T4 DNA ligase, and then stored on ice. A small aliquot was electrophoresed against a molecular weight standard (Refer to Appendix VI) to check the extent of ligation. The ligation reaction was then transformed using either a Calcium chloride protocol or by electroporation.

A 2-4 μL sample of the ligation reaction was combined with 200 μL of competent *E. coli* cells for transformation (Section 2.12.1), and the remaining protocol of Section 2.14.1 was followed exactly. Alternatively the ligation reaction was diluted five fold with sterile H_20 and 1 μL was used with 40 μL electro-competent *E. coli* cells for transformation (Section 2.14).

2.12 TRANSFORMATION USING CALCIUM CHLORIDE

2.12.1 PREPARATION OF COMPETENT *E. coli* CELLS

A LB agar plate was inoculated with an *E. coli* host strain (XL1-Blue MRF'[2]) from a frozen glycerol stock (School of Biological Sciences, Waikato University, NZ) by scraping a few microlitres from the surface with a sterile pipette and streaking it onto a LB agar plate containing 10 μg/mL tetracycline. The plate was grown overnight at 37°C. The remainder of the glycerol stock was returned to the -80°C freezer without thawing.

A single colony from the overnight plate was isolated and inoculated in 50 mL of SOB broth (Refer to Appendix I). The cells were incubated at 37°C with shaking at 250 rpm until the cells had grown to an absorbance of 0.4-0.5 (A_{600}). The cells were transferred to sterile 50 mL centrifugation tubes, incubated on ice for 10 minutes, and recovered by centrifugation in a Beckman J2-21M Centrifuge at 2,000 g for 10 minutes at 4°C. The supernatant was removed and the tubes were inverted to remove all the media from the pellet. The pellet was resuspended in 10 mL of ice cold 0.1 M $CaCl_2$ and stored on ice for 10 minutes. The sample was centrifuged at 2,000 g for 10 minutes at 4°C and the supernatant was carefully removed. The pellet was resuspended in 2 mL of ice cold 0.1 M $CaCl_2$. The competent *E. coli* cells were then stored on ice.

[2] XL1- Blue MRF' cells carry an F' episome that confers antibiotic resistance to tetracycline.

2.12.2 TRANSFORMATION

Ten microlitres of ligated plasmid DNA was added to 200 μL of fresh competent *E. coli* cells. Each tube was gently swirled and incubated on ice for 30 minutes, heat shocked for 90 seconds in a 42°C water bath and then chilled on ice for 2 minutes.

A 200 μL sample of transformed cells was added to a sterile 1.5 mL microcentrifuge tube containing 800 μL of SOC medium (Appendix I) which had been prewarmed to 37°C. The cells were grown for 45 minutes at 37°C with vigorous shaking at 250 rpm to allow the bacteria to recover, and express the antibiotic marker encoded by the plasmid.

Various volumes of the cells were plated out onto LB medium plates containing 75 μg/mL ampicillin, 20 mg/mL Xgal and 200 mg/mL IPTG and spread evenly using a sterile glass rod. Fifty microlitres of untransformed, competent *E. coli* host cells were also plated out as a negative control. The plates were left at RT until the liquid had been absorbed. The plates were inverted and grown overnight at 37°C for 12-16 hours.

2.13 TRANSFORMATION USING ELECTROPORATION

2.13.1 PREPARATION OF COMPETENT *E. coli* CELLS

This method was developed by BioRad Laboratories (1990).

Fresh streaks of XL1-Blue cells from a frozen glycerol stock were grown overnight at 37°C on LB agar plates containing 10 μg/mL tetracycline. One colony was dispersed in 5 mL LB medium containing 10 μg/mL tetracycline. The bacteria cells were grown overnight at 37°C, 250 rpm shaking.

Five microlitres of overnight culture was inoculated in 500 mL of LB medium (Appendix I). The cells were grown at 37°C, shaking at 250 rpm until they reached an A_{600} of 0.5-0.7. The bacteria were then chilled on ice for 30 minutes followed by centrifugation at 4,000 g for 15 minutes, at 4°C. The pellet was resuspended in 500 mL of ice-cold sterile 10% glycerol and recentrifuged as before. The pellet was then resuspended in 10 mL of ice-cold sterile 10% glycerol and recentrifuged as before, with the pellet being resuspended in a final volume of 2.5 mL ice-cold 10% glycerol (final cell concentration; 1-3 x 10^{10} cells/mL). This final suspension was aliquoted into 40 μL volumes and then immediately flash-frozen in a dry ice/ethanol bath. The frozen electro-competent cells were stored at -70°C for up to six months.

2.13.2 ELECTRO-TRANSFORMATION

The white sample chamber slide and sterile Gene Pulser®/E.coli Pulser™ cuvette (0.1 cm electrode gap) were chilled on ice. The BioRad Gene Pulser apparatus was set to a capacitance of 25 μF, and a voltage of 1.8 kV. The BioRad Pulse Controller was then set to 200 Ω. DNA samples were diluted 10 fold with sterile dH_20 and a 1 μL sample of the diluted DNA transferred to a 1.5 mL microcentrifuge tube of thawed electro-competent cells (40 μL). The sample was gently mixed, and then transferred to the chilled Gene Pulser®/E.coli Pulser™ cuvette, pulsed, and then 1 mL of SOC medium was quickly added to the cuvette. The cells were gently resuspended using the pipette tip and transferred to a 15 mL polypropylene tube and incubated at 37°C for 1 hour with shaking at 250 rpm. The pulse length during the electroporation was recorded.

Various volumes (25-100 μL) of the cells were plated out onto LB medium plates containing 75 μg/mL ampicillin, 20 mg/mL Xgal and 200 mg/mL IPTG. The plates were left at RT until the liquid had been absorbed then inverted and incubated at 37°C for 12-16 hours.

2.14 SMALL SCALE ISOLATION OF PLASMID DNA

Single white colonies of putative transformed bacterium were used to inoculate 5 mL of LB medium containing ampicillin with a final concentration of 75 µg/mL. The cells were grown overnight at 37°C with vigorous shaking at 250 rpm. A 2 mL sample of overnight culture was stored at 4°C in case glycerol stocks and agar plate samples were required for future DNA manipulation (Glycerol stocks were prepared by mixing 400 µL of sterile 2X Freezing media (Appendix I) and 400 µL of overnight culture in a sterile 1.5 mL microcentrifuge tube and storing at -20°C).

The overnight culture was also streaked onto an ampicillin LB agar plate and grown overnight at 37°C. The remaining overnight culture was used for the plasmid DNA "mini prep." One and a half microlitres of overnight culture of transformed *E. coli* was transferred to a sterile 1.5 mL microcentrifuge tube and centrifuged at 12,000 g for 30 seconds at RT to pellet the cells in preparation for plasmid isolation by the alkaline lysis technique.

The supernatant was removed and the pellet was resuspended in 200 µL of Solution I (100 mM Tris-HCl, 10 mM EDTA, 400 µg of heat-treated RNaseI/mL) and vortexed vigorously for 1 minute to lyse the bacterial cell wall. Two hundred microlitres of Solution II (0.2 M NaOH, 1% (w/v) SDS) was added and mixed by inversion several times to disrupt protein-protein and lipid-protein complexes. The tubes were incubated at RT for 5 minutes. Two hundred microlitres of Solution III (3 M potassium, 5 M acetate) was added and mixed by inversion several times to precipitate chromosomal DNA, high molecular weight RNA and potassium/protein/membrane complexes. The tubes were then centrifuged at 12,000 g for 5 minutes at RT. The supernatant, which contains plasmid DNA, was transferred to a sterile 1.5 mL microcentrifuge tube and mixed with 0.7 volumes of isopropanol at ambient temperature. The sample was incubated for 5 minutes at RT and the DNA pelleted by centrifugation at 15,000 g for 10 minutes at RT. The DNA pellet was resuspended in 200 µL of TE buffer (pH 8.0) and extracted with 200 µL of Tris-saturated phenol (pH 8.0).

The phases were separated by centrifugation at 5,000 g for 5 minutes at RT. The aqueous phase was then extracted with an equal volume of chloroform:isoamyl alcohol (24:1) and the phases were separated as before.

The upper aqueous phase was transferred to a sterile 1.5 mL microcentrifuge tube and 0.1 volumes of 3 M sodium acetate (pH 5.0) and 2.5 volumes of absolute ethanol was added. The tubes were incubated at -20°C for 15 minutes and the plasmid DNA was recovered by centrifugation at 15,000 g for 15 minutes at RT. The DNA pellet was washed with 1 mL of 70% ethanol and centrifuged as before but for 5 minutes only. The DNA pellet was air dried in a 37°C heating block for 10 minutes and resuspended in 20 µL of TE buffer (pH 8.0). The plasmid DNA was stored at 4°C.

Plasmid DNA samples were analysed for the presence of a cloned insert by performing double restriction digests. A 20 µL digestion reaction was prepared containing isolated plasmid DNA, 1X RE buffer and two different restriction enzymes to generate detectable fragments on a 2% mini TAE agarose gel containing 0.5 µg/mL EtBr (Table 2.4, Appendix VI). Electrophoresis was performed at 100 volts for one hour at RT.

Alternatively, the cloned PCR fragment could be positively identified by amplifying the isolated plasmid DNA with the original PCR primers and confirming a positive band of the size of the cloned insert on the agarose gel (Section 2.9).

Triplex Region	1st Restriction Enzyme (5')	2nd Restriction Enzyme (3')	Size of insert (bp)	Restriction Enzyme Buffer
Bov*REP3*	*Xba*I	*Hind*III	150	React 2
Bov*SRY*	*Xba*I	*Hind*III	250	React 2
*sup*F	*Eco*RI	*MLu*I	161	Buffer H
Bov*REED*	*Xba*I	*Eco*RI	450	React 2

Table 2.4: Restriction enzymes used to confirm the presence of a cloned insert. React 2 1X RE buffer contains 50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, pH 8.0. 1X RE buffer H contains 50 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTE, pH 7.5.

Once the cloned fragment was isolated, the DNA concentration was determined by spectroscopy. The DNA was then sequenced by the Waikato DNA Sequencing Facility on a ABI Prism 377 Automated DNA Sequencer to ensure that the insert and DNA triplex region were not mutated during PCR or DNA cloning. During 1996 ssDNA was used for DNA sequencing rather than dsDNA as the results were found to be better. The forward and reverse PCR primers and/or M13 primer were used for the DNA sequencing reaction.

2.15 SINGLE STRANDED DNA RESCUE METHOD

The following method was provided by Helen Davey (AgResearch, Ruakura, Hamilton, NZ) to produce ssDNA from the Bluescript®II SK^+ plasmid for DNA sequencing. All centrifugation was performed in the Centrifuge 5415C.

A 10.4 mL sample of 2xYT broth containing XL1-Blue MRF' cells and cloned Bluescript®II SK^+ plasmid was inoculated with 30 µL 10^{11} pfu/mL VCSM13 helper phage (Stratagene, USA) and 10.4 µL 50 mg/mL ampicillin. The culture was grown for 3 hours at 37°C with vigorous shaking and then 2.8 µL of 25mg/mL kanamycin was added to promote selection of the infected cells. The culture was grown overnight, at 37°C, with shaking at 250 rpm.

Overnight cultured cells were transferred to 1.5 mL microcentrifuge tubes and centrifuged for 5 minutes, 12,000 g at RT. The supernatant (1.3 mL) was transferred to a sterile 1.5 mL microcentrifuge tube and mixed with 195 µL 20% PEG 8000, 2.5 M NaCl solution. The tube was left on ice for 15 minutes to precipitate phage particles then centrifuged as before with the pellet resuspended in 100 µL TE buffer (pH 8.0). The sample was vortexed for 30 seconds and then extracted with 0.5 volume of Tris-saturated phenol (pH 8.0), vortexed for 30 seconds and then extracted with 0.5 volume of chloroform. The sample was vortexed for a further 30 seconds and centrifuged as before to separate the phases.

The aqueous phase was transferred to a sterile 1.5 mL microcentrifuge tube and 0.1 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of 95% ethanol were added to precipitate the DNA. The tube was left on ice for 5 minutes and then the DNA was pelleted by centrifugation at 12,000 g for 10 minutes at RT. The DNA pellet was resuspended in 10 μL TE buffer (pH 8.0) and stored at 4°C. The purity and quantity of ssDNA recovered was analysed by agarose gel electrophoresis and spectrophotometry (Refer to Appendix VI).

2.16 DNA SEQUENCING

DNA samples were sequenced by the Waikato DNA Sequencing Facility which uses an ABI Prism 377 Automated DNA Sequencer with dye terminator chemistry. The following DNA concentrations are required for sequencing: 200 ng/μL dsDNA, 50 ng/μL ssDNA, 20-45 ng/μL PCR product, 0.8 pmol/μL primer.

2.17 LARGE SCALE PLASMID DNA ISOLATION USING ALKALINE LYSIS

This method was obtained from Robert S. Ronimus (1993) which was modified from Maniatis *et al.*, (1989). All centrifugation was carried out in the Beckman Model J2-21M Centrifuge unless otherwise stated.

A single white colony from the agar plates containing the desired plasmid was used to inoculate 500 mL of LB medium (supplemented with 0.2% glucose) containing ampicillin (FC = 75 μg/mL). The cells were grown overnight at 37°C with vigorous shaking at 250 rpm. The overnight culture was transferred to a sterile G23 bottle and centrifuged at 11,000 g for 5 minutes at 4°C to harvest the cells.

The supernatant was removed and the cells resuspended in 18 mL of ice-cold Solution I (50 mM glucose, 10 mM EDTA [pH 8.0], 25 mm Tris-HCl [pH 8.0]) and left on ice for 15 minutes. Two millilitres of ice cold Lysozyme (10mg/mL in Solution I) was added.

The cells were then transferred to a sterile 100 mL centrifugation tube and 40 mL of freshly made Solution II (0.2 M NaOH, 1% SDS) was added and mixed by inversion several times to disrupt the protein-protein and lipid-protein interactions. The sample was incubated at RT for 5 minutes and 20 mL Solution III (3 M KAc, glacial acetic acid, pH 5.5) was added then mixed by vigorous shaking and incubated on ice for 10 minutes.

The sample was then centrifuged at 4,000 g, 15 minutes at 4°C without braking while decelerating. The supernatant was filtered through two layers of sterile cheesecloth and precipitated with 0.6 volumes of isopropyl alcohol. The sample was incubated at RT for 15 minutes and then centrifuged for 15 minutes at 3,000 g, 20°C. The DNA pellet was washed with 5 mL 80% ethanol and allowed to airdry at RT for 5 minutes. The DNA pellet was recovered by centrifugation as before and resuspended in 2.5 mL TE buffer (pH 8.0).

Two and a half grams of caesium chloride (CsCl) was added to the total plasmid DNA sample and the total volume was determined. Ethidium bromide (10 mg/mL) was added to the sample in the ratio 0.8 mL EtBr:10mL DNA/CsCl. The sample was mixed and centrifuged at 5,000 g for 5 minutes at RT. The supernatant was transferred to open-top Beckman 3.0 mL thick-walled polycarbonate tubes which were weighed out and balanced within 1 mg. The tubes were then placed in a TLA100.3 fixed angle rotor and centrifuged in a Beckman benchtop ultracentrifuge (Model TC-100) for 24 hours, 70 000 rpm, 15°C.

After centrifugation, the band representing the supercoiled plasmid DNA was identified by visualising the EtBr bands (Figure 2.4), and was removed using a pipette tip and placed in a sterile 1.5 microcentrifuge tube. The EtBr was removed by adding an equal volume of 1-butanol saturated with H_2O and mixing by vortexing for 30 seconds. The mixture was centrifuged at 11,000 g for 2 minutes at RT in the 5415 C Centrifuge. The top pink layer was discarded and the lower aqueous phase was transferred to a sterile 1.5 mL microcentrifuge tube. The extraction with 1-butanol saturated with H_2O was repeated until no pink colour was observed in either the aqueous or organic phase.

CsCl was removed by diluting the DNA with 3 volumes of H_2O and precipitating the DNA solution with 2 volumes of ethanol for 15 minutes at 4°C followed by centrifugation at 10,000 g for 15 minutes at RT. The precipitated DNA was resuspended in 500 µL of TE buffer (pH 8.0). The purity and quantity of DNA recovered was determined by agarose gel electrophoresis and spectrophotometry (Appendix VI).

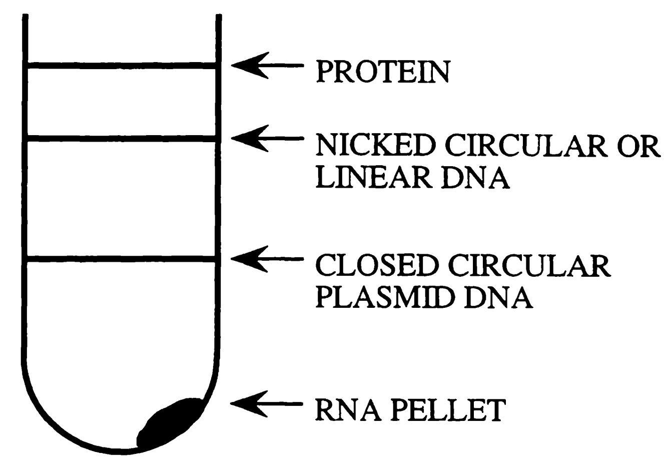


Figure 2.4: A schematic diagram showing the banding pattern of DNA in a Polycarbonate centrifugation tube (51mm by 13 mm) after equilibrium centrifugation in CsCl-EtBr gradients.

2.18 *IN VITRO* DNA TRIPLEX BINDING ASSAYS USING PLASMID DNA

The following method was modified from Musso *et al.*, (1996). A restriction digest reaction was prepared to digest the 3' end of the plasmid DNA so the sticky end could be end-labelled with [α^{32}P]dATP using Klenow enzyme and then end-filled with "cold" dNTP (Figure 2.5). The radioactively labelled plasmid DNA was then incubated with a specific psoralen-conjugated TFO and UV irradiated (360 nm) to generate photoadducts. The plasmid DNA was then digested with another restriction enzyme at the 5' end to release the DNA fragment containing the DNA triplex region. The photoadducts were analysed by EMSA.

A 20 µL restriction reaction containing 100 ng plasmid DNA (contains the cloned bovine DNA triplex), 1X React 2 RE buffer (50 mM Tris-HCl, 10 mm $MgCl_2$, 50 mM NaCl, pH 8.0), 10 units *Hind*III (Boehringer Mannheim, Germany) and dH_20 was incubated for 2 hr at 37°C. Following digestion, the sample was incubated for 10 minutes at 65°C to heat inactivate the restriction enzyme. The digested plasmid DNA was then end-labelled with 10 µCi [α^{32}P]dATP (3000 Ci/mmol, Amersham, UK) by adding one unit of Klenow enzyme, and incubating at RT for 15 minutes. One microlitre of a 2 mM dNTP mixture containing dGTP, dCTP, and dTTP was then added to the samples and incubated for a further 15 minutes at RT to completely fill the 3' recessed end. The sample was incubated at 65°C for 10 minutes to heat inactivate the Klenow enzyme.

A 20 µL DNA triplex binding assay was prepared containing 20 ng of *Hind*III end-labelled plasmid DNA (PL-Bov*SRY*[3] or PL-Bov*REP3*[4]), 5 µM of a specified psoralen-conjugated oligonucleotide, 40 mM Tris-HCl (pH 8.0) and 100 mM $MgCl_2$. The assay was incubated in the dark for 2 hours at 37°C. The control reactions contained no psoralen-conjugated TFOs or unmodified TFOs (ie. absence of psoralen at the 5' end).

[3] PL-Bov*SRY* represents plasmid DNA that contains the cloned 200 bp Bov*SRY* PCR insert.
[4] PL-Bov*REP3* represents plasmid DNA that contains the cloned 100 bp Bov*REP3* PCR insert.

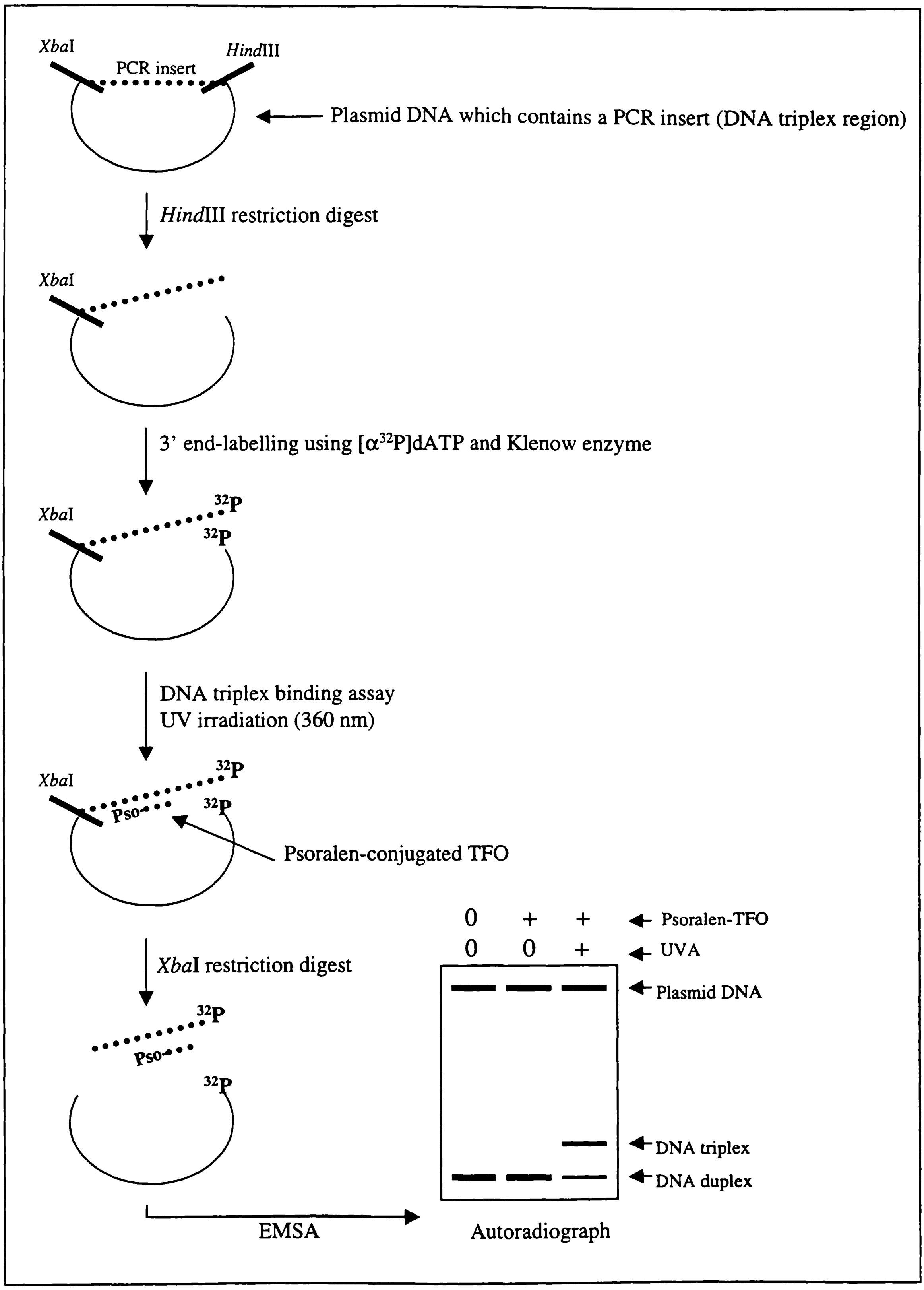


Figure 2.5: A Flow diagram showing the experimental approach used to target psoralen-conjugated oligonucleotides to double stranded plasmid DNA (contains the cloned bovine DNA triplex region). The PCR insert is represented by dotted lines. Two restriction enyzmes, *Xba*I and *Hind*III were used in the process.

Following the two hour incubation period, the tube containing the 20 μL DNA triplex binding assay sample was placed in the cold room (4°C) on top of a plastic petri dish filled with ice and UV irradiated (as previously described in Section 2.6) to generate photoadducts.

The DNA triplex binding sample was incubated at 37°C for 5 minutes and then the 5' end was cleaved off by digestion with 10 units of *Xba*I (Life Technologies, USA) for 2 hours in the dark at 37°C. Following digestion, the restriction enzyme was heat inactivated for 10 minutes at 65°C.

After these steps, PAGE electrophoresis was performed to analyse DNA triplex formation. Each sample was diluted 10 fold with 2X Formamide DNA sequencing dye, and a 4 μL volume loaded onto a 4% denaturing PAGE gel which had been prerun for 15 minutes at 90 W. Electrophoresis was performed at 60 watts until the xylene cyanole dye had migrated half way down the PAGE gel, indicating the distance at which a 19 nt fragment would run. The gel was dried under vacuum at 80°C, and exposed to Kodak X-OMAT AR film using an intensifying screen at -70°C and developed by autoradiography (Section 2.3.4).

2.19 PREPARATION OF RADIOACTIVE MOLECULAR WEIGHT DNA MARKERS

A *Hind*III restriction digest of λDNA was prepared to generate the following size fragments: 23130, 9416, 6557, 4361, 2322, 2027, 564, and 125 bp. A 20 μL restriction reaction containing 50 ng λDNA (cl857Sam7, Boehringer Mannhein, Germany), 1X RE buffer B (10 mM Tris-HCl, 5 mM $MgCl_2$, 100 mM NaCl, 1 mM 2-mercaptoethanol, pH 8.0), 50 units of *Hind*III and dH_20 was incubated for two hours at 37°C. The reaction was then end-filled by adding 1 μL of [α^{32}P]dCTP (10 mCi/mL, 3000 Ci/mmol, Amersham, UK) and 1 μL of Klenow enzyme (2 U/μL, Boehringer Mannheim, Germany) with incubation for 15 minutes at RT followed by the addition of 1 μL of 1 mM [dATP, dGTP, dTTP]. The sample was incubated for 15 minutes at RT and then 10 minutes at 65°C to heat inactivate the Klenow enzyme.

Twenty microlitres of 1X gel loading dye (Appendix I) was added to the λDNA radioactive molecular weight marker and stored at -4°C for up to two weeks. A 5 μL aliquot was loaded onto an agarose gel for Southern blot analysis (Section 2.8.3).

Chapter 3

IN VITRO DNA TRIPLEXES

In this study, a series of psoralen-conjugated TFOs have been designed to target specific repetitive regions on the male bovine Y chromosome such as Bov*REED*[7] and Bov*REP3*. Other genes, such as the Testis Specific Protein Y (TSPY) and Sex Determining Region Y (SRY), were also analysed for potential DNA triplex formation. In all cases, putative triplexes were irradiated with UV light (360 nm) to induce stable *in vitro* photoadduct formation.

3.1 GENERAL EXPERIMENTAL PROCEDURE

The ability of the psoralen-conjugated TFO to covalently crosslink to dsDNA targets upon UVA irradiation was tested by an electrophoretic mobility shift assay (EMSA). Specific details of the procedures used in each experiment are given in other sections of this research thesis, but in general the following methods were used. Radioactively end-labelled dsDNA targets were prepared by two different procedures. In the first procedure, a single strand of synthetic DNA containing the homopurine region of the DNA triplex site was 5' end-labelled with [γ^{33}P]ATP, and a DNA duplex was formed by annealing with a complementary short primer and end-filling using the large fragment of *E. coli* DNA polymerase 1 (Klenow enzyme) and "cold" dNTPs. The second method was essentially the same as the first except that the labelling step was performed during the end-filling stage using [α^{32}P]dCTP or [α^{32}P]dATP followed by a "cold" chase with dNTPs. These radioactive DNA duplexes were incubated with specific psoralen-conjugated TFOs, spermidine, and DNA triplex buffer containing Tris-HCl and $MgCl_2$ under various conditions. The samples were then irradiated with UV light (360 nm) to generate stable photoadducts. Photoadduct formation was detected by EMSA.

---

[7] Bov*REED* and Bov*REP3* refer to the male bovine fragments that contain a specific repetitive Y chromosome DNA triplex region (Appendix III).

The EMSA can detect two photoadduct structures; monoadducts (MA) and crosslinks (XL). In the present discussion, a monoadduct refers to the photoaddition of the psoralen to just one strand of the DNA duplex, yielding a structure in which the TFO is covalently linked to that strand. A crosslink (bisadduct) refers to a psoralen molecule that has photoreacted with both strands of the DNA duplex, yielding a structure in which the TFO covalently crosslinks to both strands of the target DNA. Under denaturing conditions, psoralen derived photoadducts will persist but non-covalent interactions will be disrupted.

The detection of photoadducts on PAGE gels depends on the fact that DNA triplexes migrate more slowly on polyacrylamide gels compared to DNA duplexes (Cooney *et al.*, 1988). TFOs which are more negatively charged will produce triplexes with mobilities closer to those of the parent duplex. As the negative charge is either neutralised or converted to a positive charge, the resulting triplex is expected to migrate more slowly (Dagle *et al.*, 1996).

The level of photoadduct formation detected by EMSA was measured by scanning the autoradiograph results into the *Personal Densitometer*™ Computing Densitometer (Molecular Dynamics) and using the computer software program *ImageQuant*™ (version 3.2) to select the bands for quantification. Once the photoadduct(s) and DNA duplex (reference marker) were selected, and the level of background noise defined, the software automatically calculated the percentage of duplex and triplex bands by measuring the total volume of integration.

3.2 *sup*F DNA TRIPLEX EXPERIMENTS

A number of experimental features were considered for the development of an efficient *in vitro* photobiological system; in particular, the UV dose, UV light source, and the constituents of the DNA triplex binding assay. Initially, the well-documented *sup*F DNA triplex model was used to analyse the efficiency of my UV photobiological system before testing for bovine DNA triplex formation with specific repetitive Y chromosome DNA target sequences.

Many researchers have used the *sup*F gene as a model for DNA triplex formation (Chapter 1). Gasparro *et al.*, (1994) and Wang *et al.*, (1995) report that psoralen-conjugated TFOs efficiently target the *sup*F gene. However, Gasparro *et al.*, (1994) used a somewhat shorter duplex target, 24 bp compared to the 38 bp used by Wang *et al.*, (1995). Thus, as a first step it was decided to investigate the advantages and disadvantages of various duplex target sizes in the *sup*F system before designing duplex targets for the bovine targeting experiments.

3'TGGTGGTGGGGGAAGGA≈**Pso** 5'	PsoAGT17
3'GGGGGAAGGA≈**Pso** 5'	PsoAG10
5'TGGTGGTGGGGGAAGGA**T**TCGAAC3' 3'ACCACCACCCCCTTCC**T**AAGCTTG5'	*sup*F DNA DUPLEX (24 bp)

Figure 3.1: *sup*F DNA triplex sequence with psoralen-conjugated TFOs, PsoAG10 and PsoAGT17 (Gasparro *et al.*, 1994). Psoralen intercalates in between the two thymines on the sense and antisense strand, at the 5' end of the triplex-duplex junction of the DNA duplex as shown in bold lettering. Note that ≈Pso represents psoralen. The DNA sequence of TFO-*sup*F is 5'AGGAAGGGGG3' and this binds antiparallel to the DNA duplex.

Gasparro *et al.*, (1994) showed that a 10 nt psoralen-conjugated oligonucleotide, PsoAG10, composed of seven guanines and three adenines, was able to form an antiparallel DNA triplex directed crosslink with a 24 bp unmodified DNA duplex (Figure 3.1). A TpA site for psoralen intercalation was also available at the 5' end of the triplex-duplex junction, allowing a high percentage of photoadducts to form under UV light (67% crosslinks with a high UVA dose, 1.71 J/cm^2).

My duplication of these experimental conditions resulted in 22% monoadduct formation after thirty seconds of UV light (Figure 3.2A). A higher UV dose (32 minutes, 2 J/cm^2), generated 37% crosslink and 22% monoadduct formation. Crosslinking could be detected under a range of UV irradiation conditions, and it was obvious that the induction of photoadducts was very efficient, even at 15 cm from the lamp and with exposure as short as 8 minutes (0.15 J/cm^2). Similar high efficiencies have been reported in the literature for this particular triplex (Wang *et al.*, 1996).

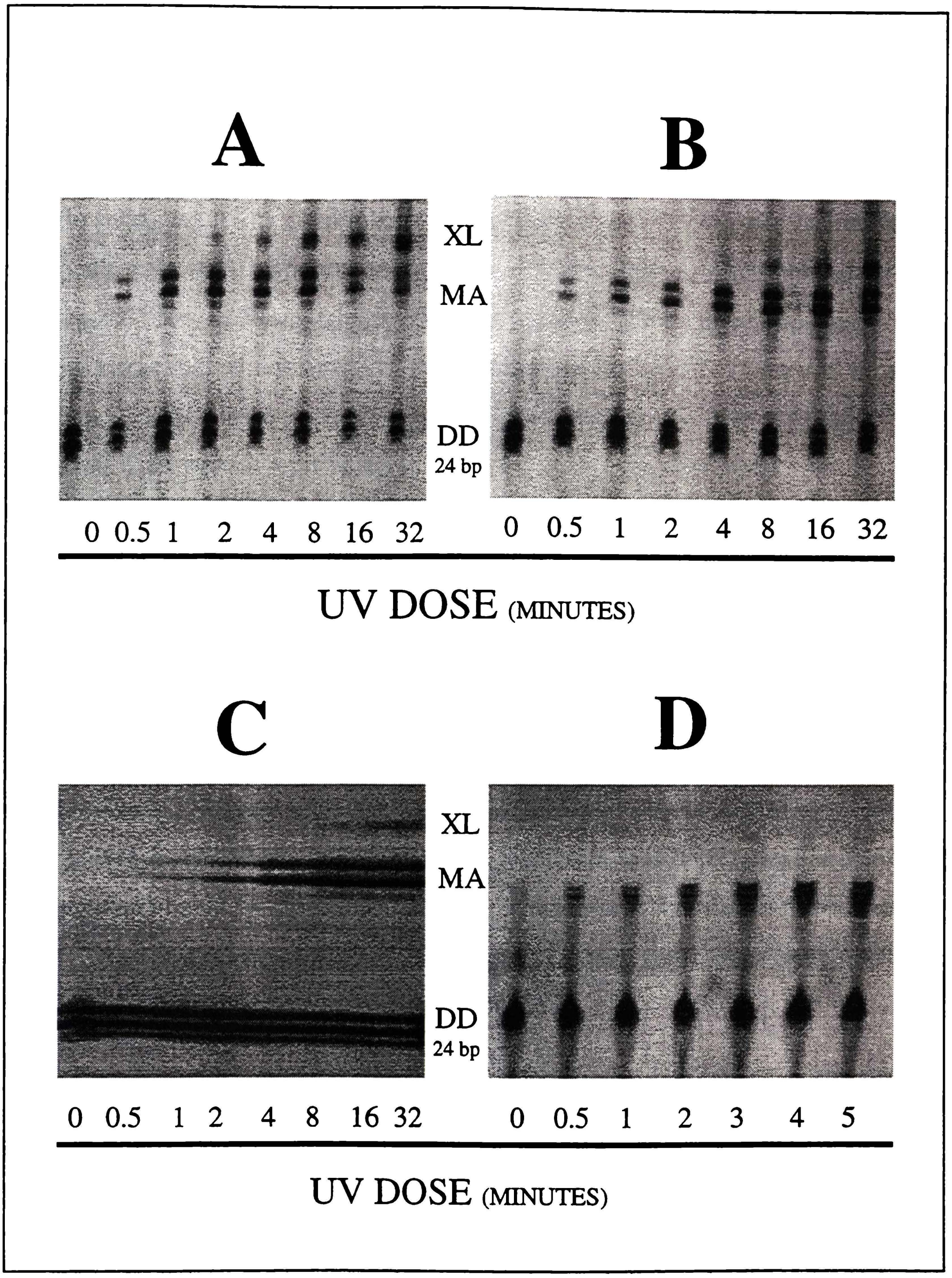


Figure 3.2: Photo-induced crosslinking of a 10 nt psoralen-conjugated TFO to a 24 bp *sup*F dsDNA target. The γ-^{33}P end-labelled duplex was incubated with 80 µM PsoAG10, 1 mM spermidine, and 1X DNA triplex buffer (10 mM $MgC1_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and irradiated with UV light (360 nm) on ice for 0-32 minutes (maximum UV dose of 2 J/cm^2). The samples were then electrophoresed on a TBE 8% denaturing PAGE gel. (**A**) 1 cm distance from the UV lamp (maximum UV dose of 4 J/cm^2), (**B**) 5 cm distance from the UV lamp (maximum UV dose of 1.5 J/cm^2), (**C**) 15 cm distance from the UV lamp (maximum UV dose of 0.6 J/cm^2). (**D**) *sup*F dsDNA incubated with 80 µM PsoAGT17 and irradiated on ice for 0-5 minutes, 5 cm distance from the UV light (maximum UV dose of 0.3 J/cm^2). XL = crosslink, MA = monoadduct, DD = amplified dsDNA.

Samples were irradiated as close as 1 cm from the UV lamp (Figure 3.2A, B, C) but in most of the subsequent experiments it was found convenient to use a distance of 3 cm (maximum UV dose of 2 J/cm^2).

PsoAG10 was then targeted to a longer dsDNA *sup*F target (38 bp). Using experimental conditions similar to those of Wang *et al.*, (1995), similar results were obtained for the dose response, and the efficiency of triplex formation was similar to that reported for the Wang triplex (data not shown). Thus, there appeared to be no obvious experimental advantage of the 38 over the 24 bp dsDNA target; however it was more desirable to target longer dsDNA targets, so the synthetic dsDNA fragments used were at least 38 bp in size.

These *sup*F experiments utilised 10 nt TFOs that were solely purines. The use of a much longer TFO (17 nt, PsoAGT17) which also contained pyrimidines was next investigated. Unfortunately, it was not possible to use the same TFO as that published by Wang *et al.*, (1995) as they used a modified *sup*F gene; instead a TFO was designed to the *sup*F gene in the pSP189 shuttle vector which was available in our laboratory. Figure 3.2D shows that after 5 minutes of UVA irradiation (0.5 J/cm^2), 41% monoadduct and 2% crosslink formation was detected when PsoAGT17 was targeted to a 24 bp *sup*F DNA duplex. Therefore, it was shown to be possible to target longer psoralen-conjugated TFOs to the *sup*F gene, even with pyrimidine mismatches in the homopurine triplex region.

I also labelled the homopyrimidine strand of the DNA duplex with [α^{32}P]dATP. Figure 3.3 shows that a crosslink formed after 15 minutes of UV irradiation (1 J/cm^2) with 4 μM of PsoAG10. Previous studies have shown that psoralen preferentially intercalates into the homopyrimidine strand of the DNA duplex; therefore, the major photoadduct in the ^{32}P UV crosslinking studies is the crosslink. In the γ^{33}P UV crosslinking studies the homopurine strand of the DNA duplex was end-labelled and therefore monoadduct and crosslink formation was observed.

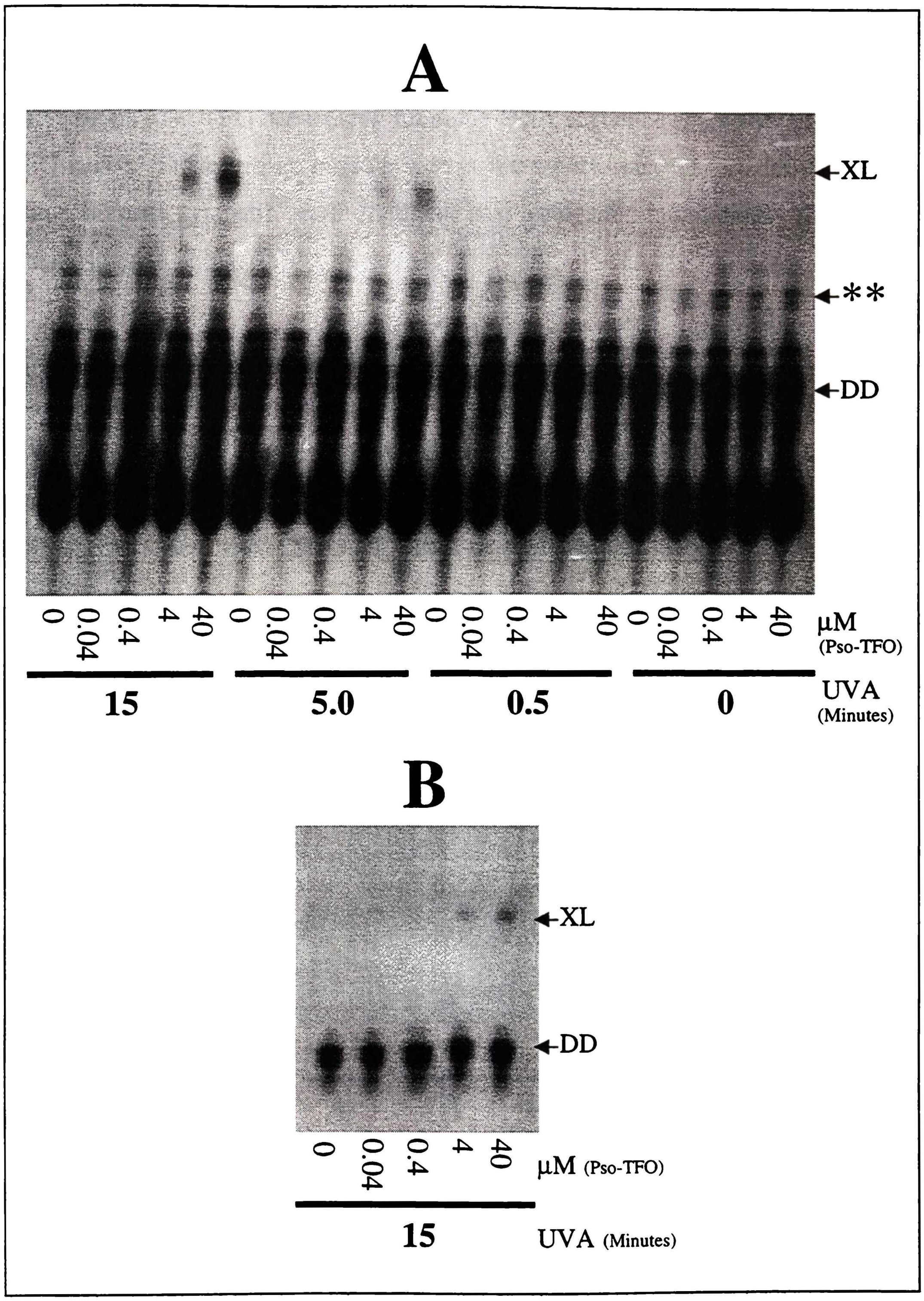


Figure 3.3: *In vitro* targeting of PsoAG10 to a $\alpha^{32}P$ unpurified (**A**) and purified (**B**) *sup*F synthetic dsDNA duplex. The radioactive *sup*F dsDNA was incubated with PsoAG10 (0-40 μM), 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated with UV light (360 nm) on ice for 0-15 minutes, 3 cm distance from the UV lamp (maximum UV dose - 1 J/cm^2). The samples were electrophoresed on a 8% TBE denaturing PAGE gel. T = crosslink, DD = DNA duplex, Pso-TFO = psoralen-conjugated TFO, ** = bands detected due to the increased sensitivity of the end-filling reaction with ^{32}P and long exposure of the X-ray film

I next investigated the effect of using unmodified TFOs where the psoralen molecule was absent at the 5' end of the TFO to determine the implications of UV-induced photoadducts in producing detectable levels of monoadducts and crosslinks. The unmodified 10 nt TFO had exactly the same DNA sequence as PsoAG10 (Figure 3.1, TFO-*sup*F). DNA triplex binding assays were performed as previously described using a small synthetic *sup*F fragment and analysing photoadduct formation by denaturing and native EMSA analysis.

Photoadducts were detected under both denaturing and native EMSA conditions, when PsoAG10 was targeted to the 24 bp *sup*F dsDNA fragment and irradiated with UV light (360 nm). However, none were detected when the unmodified TFO-*sup*F was targeted to the same *sup*F target (Figure 3.4). No experiments have been published to show that TFO-*sup*F binds to *sup*F dsDNA target and forms a stable DNA triplex.

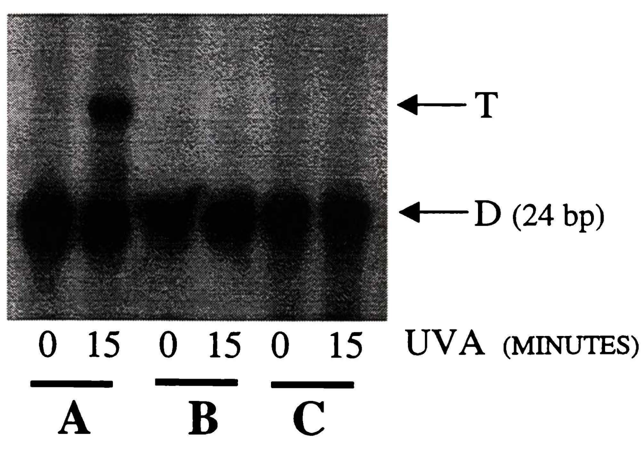


Figure 3.4: $\alpha^{32}P$ *sup*F DNA triplex binding assay with a psoralen-conjugated TFO and/or an unmodified TFO. A 10 µL assay containing $\alpha^{32}P$ end-labelled duplex (24 bp) was incubated with 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1), 1 µg/µL tRNA and 40 µM PsoAG10 (**A**) or 0-40 µM TFO-*supF* (**B**), overnight at 37°C, in the dark. PsoAG10 and TFO-*supF* were absent from Sample C (control). Following incubation, samples were irradiated with UV light (360 nm) on ice for 0-15 minutes, 3 cm away from the UV lamp (maximum UV dose-1 J/cm^2). Samples were then diluted with 10 µL 1X Loading Dye (Type III) and 5 µL volumes were loaded onto a 15% large Tris-acetate-magnesium native PAGE gel. Electrophoresis was performed at 30 W for 2 hr at RT. D = DNA duplex, T = DNA triplex.

3.3 BOVINE DNA TRIPLEX FORMATION

3.3.1 Experimental Design for Bovine Photoadduct Formation

The Genbank computer database program was used to search for sites within repetitive bovine Y chromosome DNA sequences that could bind a homopurine or a homopyrimidine motif TFO. DNA sequences were screened for those regions containing at least ten continuous homologous purines or pyrimidines, and also a nearby pyrimidine at the 5' end at the triplex-duplex junction, preferably a 5'TpA site for psoralen intercalation. The size, nucleotide composition and availability of restriction enzyme sites of the bovine Y chromosome DNA sequence were also considered for potential use in *in vivo* DNA triplex formation with bovine spermatozoa.

BLAST computer database searches were also performed to detect any other DNA sequences that contained all or part of the repetitive bovine Y chromosome sequence. This was investigated to determine the specificity of the male bovine repeat within bovine genomic DNA and in comparison with other species. However, these results do not represent all the DNA sequences, as mammalian genomes have not been completely sequenced.

3.3.2 Bov*REP3* DNA Triplex Formation

A 38 bp DNA fragment of the 306 bp *Bos taurus* Y chromosome repeat (GenBank accession n# U39887, Appendix III) was employed as a dsDNA target for potential DNA triplex formation and UV crosslinking studies. This fragment contains a consecutive 10 nt run of purine bases, enabling the targeting of psoralen-conjugated TFOs Puv6a, Puv6b, and Puv6c. The sequences of the target duplex and reactive psoralen-conjugated TFOs are shown in Figure 3.5.

3′ GAGGAGGGGG≈Pso 5′ Puv6a

3′ GAGGAGGGGGTTGGGG≈Pso 5′ Puv6b

3′ GTGGTGGGGGTTGGGGTT≈**Pso 5′** **Puv6c**

5′ GAGGAGGGGGCCTGGGCAT**CTGAAAGTTCAATGTTCGC** 3′
3′ CTCCTCCCC**C**GGACCC**CGT**AGACTTTCAAGTTACAAGCG 5′] Bov*REP3* DNA DUPLEX (38 bp)

Figure 3.5: A schematic representation of Bov*REP3* DNA triplex formation with psoralen-conjugated TFOs Puv6a, b and c. Psoralen intercalates into the pyrimidine bases on the sense or antisense strand, at the 5' end of the triplex-duplex junction of the DNA duplex, as shown in colour; GREEN for Puv6a, BLUE for Puv6b and RED for Puv6c. Note that ≈Pso represents psoralen. The DNA sequence for TFO-*REP3* is 5'GGGGGAGGAG3' and this binds antiparallel to the DNA duplex. The DNA sequence for Puv5 is 5'Pso≈GAGGAGGGGG3' and this binds parallel to the DNA duplex.

Puv6a[6] is a 10 nt AG rich psoralen-conjugated TFO that is designed to bind antiparallel to the dsDNA homopurine target forming a DNA triplex via Hoogsteen hydrogen bonding. The psoralen can intercalate between the two pyrimidines present at the 5' end of the DNA triplex-duplex junction, forming a cytosine-psoralen-cytosine crosslink after UV exposure (360 nm). Figure 3.6A shows that after 32 minutes of UV irradiation (2 J/cm^2), 55% monoadduct formation occurred. However no crosslinks were detected. The intensity of the monoadduct band increased as UV dosage increased. Under the conditions of this assay monoadducts were only detectable if the Puv6a concentration was at least 8 μM (results not shown).

I next investigated the temperature at which the DNA was incubated with the TFO before UV irradiation. Previous research has shown that this can have a significant effect on the stability of the DNA triplex. Monoadduct formation was detected at temperatures as diverse as 4, 18 and 55°C (Figure 3.6B).

[6] Puv refers to a psoralen-conjugated TFO

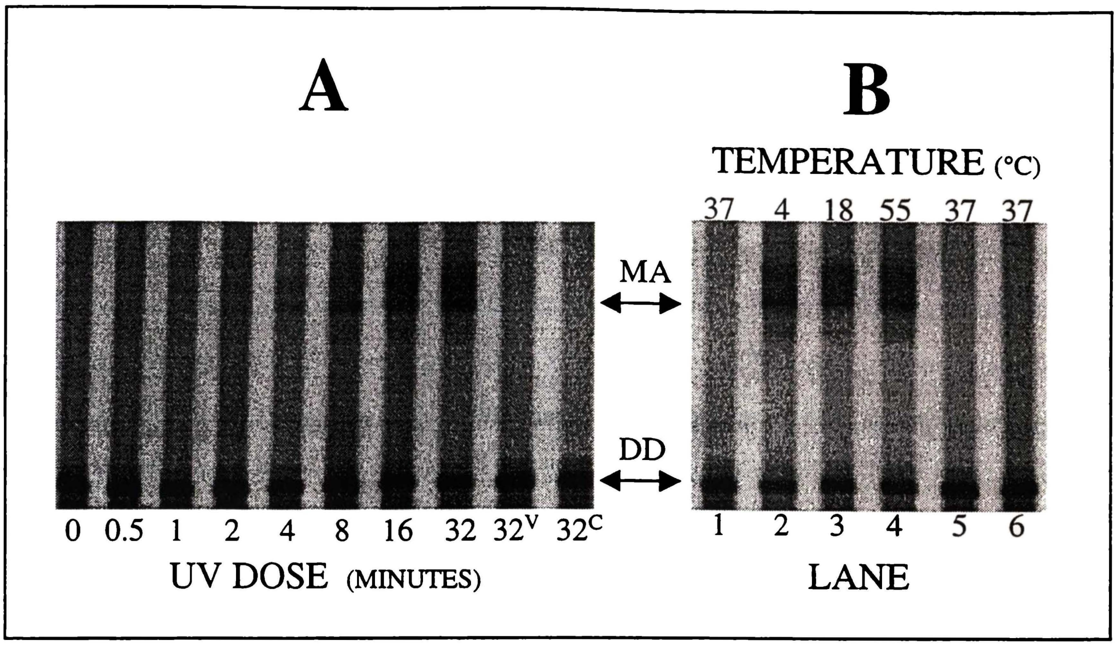


Figure 3.6: Photo-induced crosslinking of 10 nt purine rich psoralen conjugated TFO to a 38 bp dsDNA Bov*REP3* fragment. **A)** The γ-^{33}P end-labelled duplex was incubated with psoralen 40 µM Puv6a for 2 hr, 37°C, pH 7.0 and then irradiated with UV light (360 nm) on ice for 0-32 minutes, 3 cm away from the UV lamp (maximum UV dose-2 J/cm^2). The samples were electrophoresed on a 8% TBE denaturing PAGE gel. 32^V represents 32 minutes of visible light and 32^C represents 32 minutes of UV irradiation with the psoralen-conjugated TFO absent. **B)** Comparison of DNA triplex binding assay incubation temperatures. The γ-^{33}P end-labelled 38 bp Bov*REP3* duplex was incubated with 40 µM Puv6a for 2 hr on ice (Lane **2**), 18 hr at RT (**Lane 3**), 2 hr at 55°C (**Lane 4**), pH 7.0 and then irradiated with UV light (360 nm) on ice for 32 minutes, 3 cm away from the UV lamp (maximum UV dose-2 J/cm^2). The samples were electrophoresed on a 8% TBE denaturing PAGE gel. **Lane 1** represents 32 minutes of UV irradiation TFO-*REP*3 (psoralen is absent at the 5' end). Other psoralen-conjugated TFOs were tested under the same conditions as **Lane 2** except that incubation before UV irradiation was carried out at 2 hr, 37°C; Puv6b (**Lane 5**) and Puv5 (**Lane 6**).
MA = monoadduct, DD = DNA duplex.

A longer psoralen-conjugated TFO was then designed to bind to the 38 bp synthetic dsDNA target (Puv6b). This psoralen-conjugated TFO is 16 nt in length consisting of adenine (18%), guanine (64%) and thymine (18%) bases which bind antiparallel to the dsDNA homopurine target, forming a DNA triplex via Hoogsteen hydrogen bonding. There were two cytosine bases present in the homopurine strand of the DNA, therefore thymine bases were incorporated into the TFO allowing a stable T·CG triplet to form.

The psoralen can also intercalate at the 5' end of the DNA triplex-duplex junction, forming a cytosine-cytosine crosslink after UV exposure (360 nm). No photoadducts were detected by EMSA analysis, presumably due to the pyrimidine bases on the third strand causing DNA triplex instability.

Another psoralen-conjugated TFO was designed to overcome the pyrimidine mismatches within the homopurine region. An 18 nt psoralen-conjugated TFO (Puv6c) consisting of 67% guanines and 33% thymine was designed to bind antiparallel to the 38 bp synthetic dsDNA target. The TFO sequence was extended two nucleotides from the 5' end of Puv6b, as an ideal psoralen intercalation site at the 5' end of the DNA triplex-duplex junction was available, allowing a thymine-psoralen-thymine crosslink to form after UV exposure (360 nm). Four pyrimidine bases, three cytosines and one thymine base were present in the homopurine region of the dsDNA target. A guanine base was incorporated into the TFO in a position designed to bind to the thymine in the purine strand of the DNA triplex site, forming a G·TA triplet. In addition, the thymine in the TFO was predicted to bind to the cytosine in the purine strand of the DNA triplex site, forming a T·CG triplet. Unfortunately, no photoadducts were detected when Puv6c was targeted to the 38 bp synthetic Bov*REP3* dsDNA target (results not shown).

Three controls were included in the current experiments to show that the 3rd strand orientation and psoralen intercalation was important for stable DNA triplex formation. Puv5, a ten base purine rich psoralen-conjugated TFO with the inverse DNA sequence to Puv6, was targeted to bind parallel to the dsDNA homopurine target, forming a DNA triplex via reverse Hoogsteen hydrogen bonding. The purine motif TFO can only bind antiparallel to the dsDNA target. Figure 3.6B shows that Puv5 does not form a DNA triplex. These findings are consistent with previous results showing that the purine motif TFO confers antiparallel third strand orientation preferences (Sun *et al.*, 1991).

A 10 nt purine rich TFO (TFO-*REP3*) with identical DNA sequence to Puv6 except that the psoralen was absent at the 5' end, was targeted to form an antiparallel DNA triplex. Figure 3.6A shows that stable DNA triplex formation did not occur with this TFO on UV activation, showing that the presence of a psoralen group was a necessary criterion to generate stable monoadducts. Under native EMSA conditions, no DNA triplexes were observed with this TFO (results not shown).

An unrelated psoralen-conjugated TFO (psoAG10, a 10 nt purine rich TFO, used in the *sup*F experiments) was also tested in these experiments to exclude non-specific crosslinking. This TFO was not crosslinked to the 38 bp duplex, as would be expected, given the non-related sequence (results not shown).

3.3.3 Bov*SRY* DNA Triplex Formation

In theory, the *Bos taurus* sex determining region Y (SRY) protein gene (GenBank accession n# BTU15569, Appendix III) contains a single copy of a specific DNA triplex site whereas Bov*REP3* contains hundreds of copies of its triplex site. Photoadduct formation in the SRY gene would thus serve as a control to test the theory that the efficiency of *in vivo* DNA triplex formation significantly increases by targeting repetitive, rather than single copy, Y chromosome DNA targets for photoadduct formation. This theory was described in Chapter 1.

A 40 bp DNA fragment of the SRY gene was employed as a dsDNA target for DNA triplex formation and UV crosslinking studies. This 40 bp fragment contains a consecutive run of 12 purine bases (Figure 3.7). A homopurine psoralen-conjugated TFO (Puv12a), which consists of 50% guanines, was designed to bind antiparallel to the dsDNA target via Hoogsteen hydrogen bonding to form a DNA triplex. A psoralen intercalation site was available at the 5' end of the DNA triplex-duplex junction, allowing the formation of a thymine-psoralen-cytosine crosslink after UV exposure (360 nm).

Other psoralen-conjugated TFOs such as Puv12b, Puv12c and Puv12d, were targeted to the 40 bp fragment (Figure 3.7). Puv12b and Puv12c were designed to bind antiparallel to the dsDNA homopurine target via Hoogsteen hydrogen bonding. Due to the presence of pyrimidine bases in the purine strand of the DNA target, Puv12b and Puv12c were modified to avoid destabilisation of the DNA triplex. Puv12b is 13 nt in length and consists of 46% guanines. There are three thymine bases in this homopurine TFO, which are designed to bind to two thymines or one cytosine in the purine strand of the dsDNA target to form T·TA or T·CG triplets.

```
        3′ ³²GGAAAGGGGAGAA≈Pso 5′                      ³²PUV12a
          3′ GAAAGGGGAGAA≈Pso 5′                         Puv12a
                         3′ ATGTTAGGGAGAG≈Pso 5          Puv12b
          3′ GAAAGGGGAGAATAGGGGAGGGAGAG≈Pso 5            Puv12c
        5′Pso≈CTTTCCCCTCTT 3′                            Puv12d
5′ ATGATCAGTGTGAAAGGGGAGAACATGTTAGGGAGAGCAG 3′   BovSRY
3′ TACTAGTCACACTTTCCCCTCTTGTACAATCCCTCTCGTC 5′   DNA DUPLEX
                                                 (40 bp)
```

Figure 3.7: A schematic representation of Bov*SRY* DNA triplex formation with psoralen-conjugated TFOs, Puv12a, b, c, and d. Psoralen intercalates into the pyrimidine bases on the sense or antisense strand, at the 5' end of the triplex-duplex junction of the DNA duplex, as shown in colour: GREEN for Puv12d, BLUE for Puv12a and ^{32}Puv12a, RED for Puv12b and Puv12c. ≈Pso represents psoralen, ^{32}Puv12a represents the α^{32}P end-labelled Puv12a. The DNA sequence for TFO-*SRY* is 5'AAGAGGGGAAG3' and this binds antiparallel to the DNA duplex.

Puv12c is 25 nt in length and consists of 64% guanines. This psoralen-conjugated TFO is targeted to a homopurine region of the DNA duplex that contains four pyrimidine bases, three thymines and one cytosine. Thus, the guanine in the TFO was designed to bind to the thymine in the purine strand of the dsDNA target, forming a G·TA triplet or alternatively bind to a cytosine, forming a T·CG triplet.

A 12 nt homopyrimidine psoralen-conjugated TFO, Puv12d, which consists of 64% cytosines, was designed to bind parallel to the dsDNA homopurine target via reverse Hoogsteen hydrogen bonding. A psoralen intercalation site was available at the 5' end of the DNA triplex-duplex junction, allowing a thymine-psoralen-cytosine crosslink to form after UV exposure (360 nm).

In the γ^{33}P UV crosslinking studies, Puv12a was the only TFO that was successfully targeted to a region of the SRY gene, forming stable photoadducts under a variety of experimental conditions (Figure 3.8). A low UV dose (30 seconds, 0.03 J/cm^2) generated 10% monoadducts and a higher UV dose (8 minutes, 0.5 J/cm^2) generated 39% monoadduct, and 4% crosslink formation. The presence of photoadducts increased as the dose increased. No DNA band shifts were detected when Puv12a was absent from control reactions.

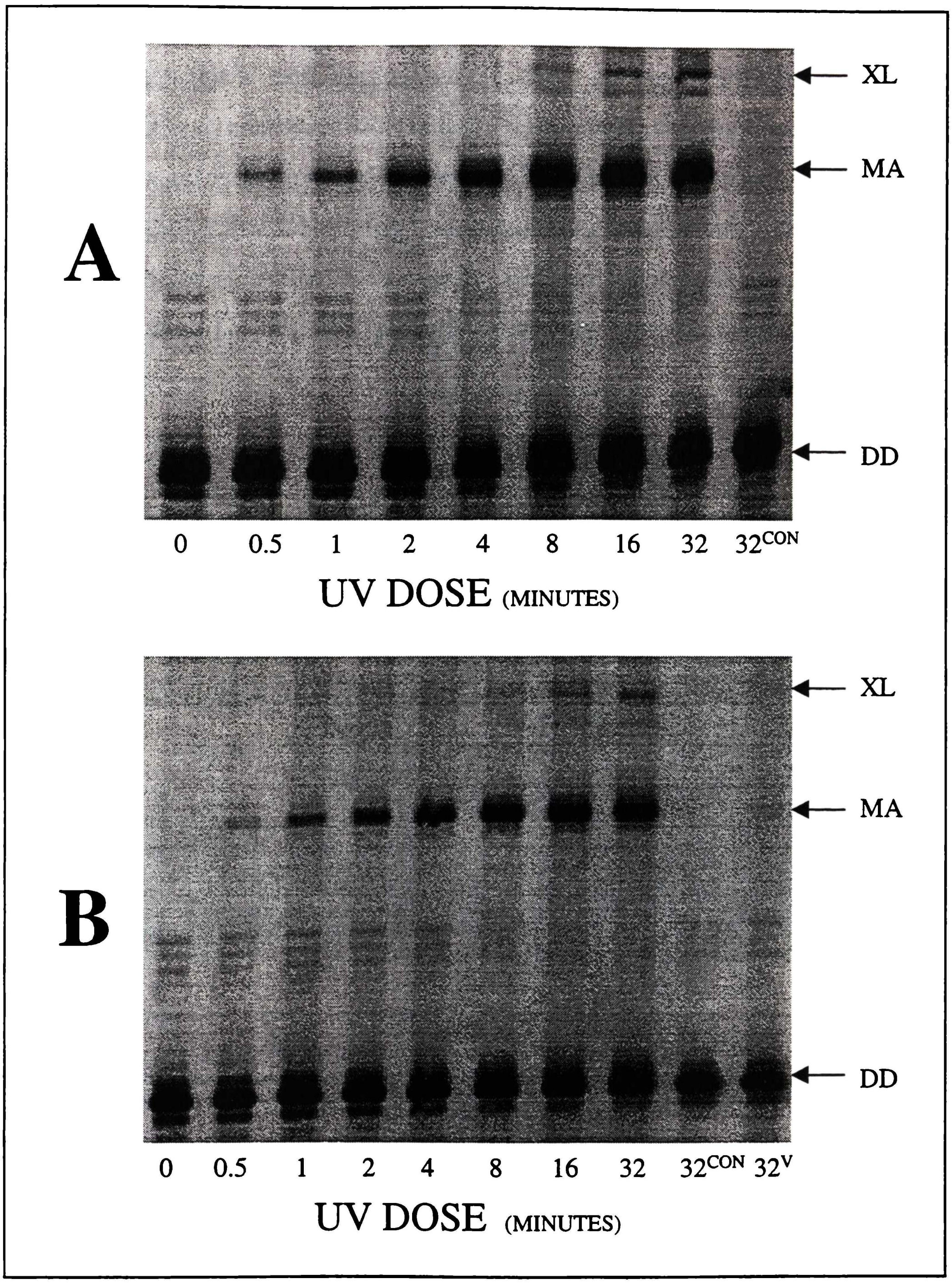


Figure 3.8: Photo-induced crosslinking of a 12 nt conjugated purine rich psoralen-conjugated TFO to a 40 bp Bov*SRY* dsDNA target. **A)** The γ-^{33}P end labelled duplex was incubated with 40 µM Puv12a, 1 mM spermidine, 1X DNA triplex binding assay (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated with UV light (360 nm) on ice for 0-32 minutes, 3 cm distance from the UV lamp (maximum UV dose-2 J/cm^2). The samples were then electrophoresed on a 8% TBE denaturing PAGE gel. Two controls were included; 32^{CON}, psoralen-TFO absent with 32 minutes of UV irradiation, and 32^{V}, psoralen-TFO present with 32 minutes of visible light exposure. **B)** The same conditions as A except that the DNA triplex binding assay was incubated for 2 hr on ice before UV exposure.
XL = crosslink, MA = monoadduct, DD = DNA duplex.

In order to test the specificity and stability of Puv12a, a 12 nt purine rich unmodified TFO (TFO-*SRY*) with identical DNA sequence to Puv12a except for the absence of psoralen from the 5' end, was targeted to form an antiparallel DNA triplex. This control TFO was incubated with radioactively labelled target duplex, irradiated and subjected to gel electrophoresis in parallel with equivalent reactions using Puv12a. Figure 3.9 shows that DNA triplex formation was unsuccessful with this TFO under native and denaturing EMSA analyses. However, photoadducts were detected with Puv12a, indicating that the psoralen molecule was necessary for photoadduct formation.

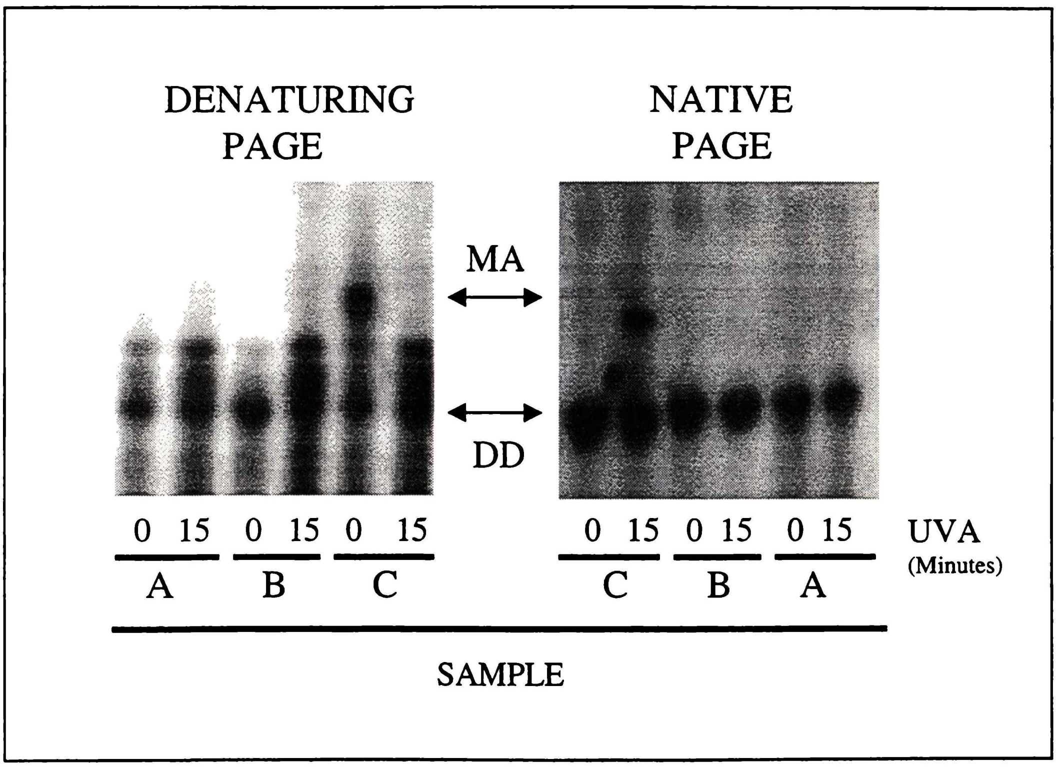


Figure 3.9: Denaturing and native EMSA analyses of DNA triplex formation with Puv12a and unmodified TFO-*SRY*. The sense strand of the dsDNA Bov*SRY* target (38 bp) was 5' end-labelled with [γ-^{32}P]dCTP and incubated with 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1), 0.5 µg/mL tRNA and either no TFO present (A), 40 µM TFO-*SRY* (**B**), or 40 µM Puv12a (C), overnight at 37°C. The samples were then irradiated with UV light (360 nm) on ice for zero and fifteen minutes, 3 cm distance from the UV lamp (maximum UV dose-1 J/cm^2). Samples were electrophoresed on either a 10% Tris-borate-magnesium native or denaturing PAGE gel. DD = DNA duplex, MA = monoadduct.

In a further series of experiments, the 5' psoralen conjugated TFO, Puv12a, was 3' end-labelled with [α^{32}P]dGTP and terminal transferase. A technique devised by Deng and Wu (1983) was employed to ensure the addition of just one guanine base onto the 3' end of the TFO. This would ensure the formation of a G·TA triplet and avoid transferase action, which would result in the addition of further guanines at the 3' end. The incorporation of many guanines at the 3' end would destabilise the DNA triplex due to unsuitable triplet combinations such as G·AT, G·CG, and G·GC (refer to Figure 3.7). The 3' end-labelling of 5' psoralen-conjugated TFOs was confirmed by native PAGE electrophoresis (results not shown).

Initially, the 3' end-labelled psoralen conjugated TFO (^{32}Puv12a) was incubated in the presence and absence of synthetic Bov*SRY* DNA duplex (40 nt). The DNA triplex binding samples were incubated for two hours at 37°C and then UV irradiated on ice for 0-30 minutes to form stable photoadducts. Figure 3.10 shows that 10% monoadduct formation occurs when ^{32}Puv12a was targeted to the synthetic 40 nt dsDNA duplex and subjected to five minutes of UV irradiation (0.3 J/cm^2). Further irradiation does not generate any more photoadducts. Figure 3.10 also shows that the use of ^{32}Puv12a labelled TFO was a very insensitive means of detecting photoadducts in comparison to targeting Puv12a to radioactive synthetic dsDNA duplexes.

In addition, 3' end-labelled ^{32}Puv12a was targeted to bovine male and female genomic DNA. This investigation was designed to determine whether or not photoadduct formation could be detected specifically with male DNA. No photoadduct formation was observed by EMSA or agarose electrophoresis when ^{32}Puv12a was targeted to bovine male and female gDNA. The targeting of end-labelled ^{32}Puv12a to *Hae*III or *Eco*RI digested bovine male and female DNA also failed to produce any detectable photoadduct formation.

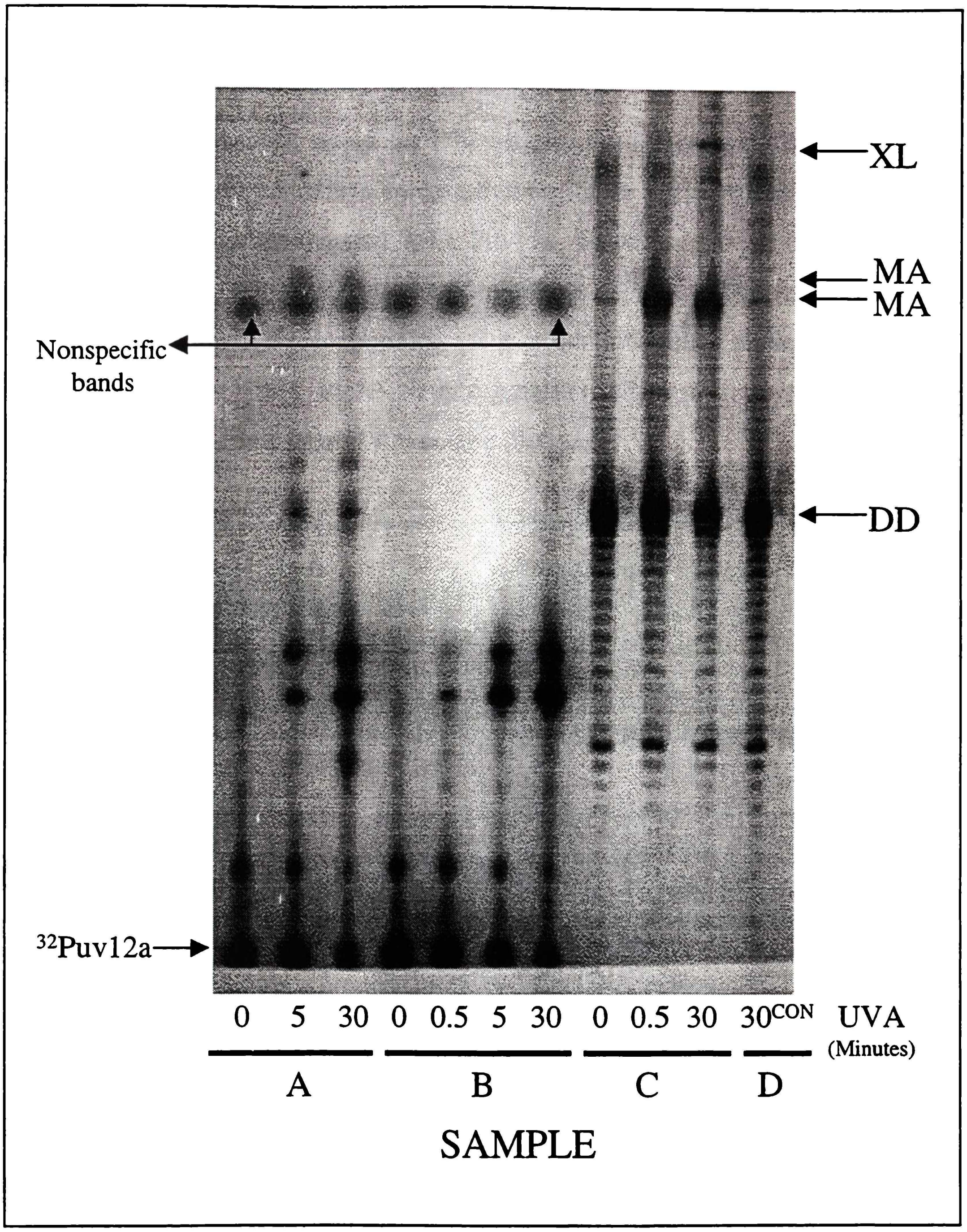


Figure 3.10: *In vitro* targeting of radioactively labelled psoralen-conjugated TFOs (^{32}Puv12a). Puv12a was radioactively 3' end-labelled with [α^{32}P]dGTP and terminal transferase. Radioactive Puv12a (FC = 20 μM) was incubated in 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, Tris-HCl pH 7.1), and 1 μM non-radioactive Bov*SRY* duplex (40 bp) for 2 hr, 37°C (Sample **A**). The control sample contained ^{32}Puv12a only (**B**). Bov*SRY* duplex (1 μM) was also radioactively 5' end-labelled with γ^{33}PATP and incubated in the presence (**C**) and absence (**D**) of non-radioactive Puv12a (FC=20 μM). The samples were then irradiated with UV light (360 nm) on ice for 0-30 minutes, 3 cm away from the UV lamp (maximum UV dose-2 J/cm^2). Following irradiation, the samples were electrophoresed on a 8% TBE denaturing PAGE gel. DD =DNA duplex, MA = monoadduct, XL = crosslink. The non specific bands present by the monoadducts maybe secondary structures that occur following the 3' end-labelling of Puv12a.

3.3.4 Bov*REED* DNA Triplex Formation

A 49 bp DNA fragment of the BRY11a sequence (Patent n# 5459038, Appendix III) was employed as a dsDNA target for DNA triplex formation and UV crosslinking studies. This fragment contains a consecutive run of 16 purine bases (Figure 3.11). A homopurine psoralen-conjugated TFO, Puv15a, which consists of 56% guanines, was designed to bind antiparallel to the homopurine dsDNA target, forming a DNA triplex via Hoogsteen hydrogen bonding. A psoralen intercalation site was available at the 5' end of the DNA triplex-duplex junction, allowing a thymine-psoralen-cytosine crosslink to form after UV exposure (360 nm).

5′Pso≈AAAGGAGGAGGGAAGG 3′	Puv15a
5′Pso≈GGTGGGTTGG 3′	Puv15b
3′CGGT**C**AAAGGA**T**GAGGGAAGGTCGACGTCACAGTGTAACTTTCACAGGT5′	Bov*REED*
5′GCCAG**T**TTCC**T**ACTCCCTTCCAGCTGCAGTGTCACATTGAAAGTGTCCA3′	DUPLEX (49 bp)

Figure 3.11: A schematic representation of the Bov*REED* DNA triplex formation with psoralen-conjugated TFOs, Puv15a and Puv15b. Psoralen intercalates into the pyrimidine bases on the sense or antisense strand, at the 5' end of the triplex-duplex junction of the DNA duplex, as shown in colour; BLUE for Puv15a and RED for Puv15b. Note the Pso≈ represents psoralen. The DNA sequence of TFO-REED is 5'AAAGGAGGAGGGAAGG3' and this binds antiparallel to the DNA duplex.

A smaller psoralen-conjugated TFO was also designed to bind to the 49 bp fragment so the psoralen could intercalate between the two thymine bases at the 5' end of the DNA triplex-duplex after UV exposure (360 nm). This 10 nt psoralen-conjugated TFO, Puv15b, contained seven guanine and three thymines (Figure 3.11). The thymine bases were designed to bind antiparallel to adenine bases on the purine strand of the DNA duplex forming T·AT triplets whereas the guanine bases bind antiparallel to guanine bases on the purine strand of the DNA duplex forming G·GC triplets. There is one pyrimidine mismatch within this homopurine region, a thymine which is near the 5' end of the DNA triplex site. To overcome this mismatch, a guanine base was incorporated into the TFO allowing a G·TC triplet to form.

Research has shown that this triplet combination prevents destabilisation of the DNA triplex (Griffin & Dervan, 1989).

Puv15a, Puv15b and TFO-*REED* were targeted to the 49 bp Bov*REED* fragment with Puv15a being the only TFO that successfully formed stable photoadducts under a variety of experimental conditions. Figure 3.12 shows that a low UV dose (30 seconds, 0.03 J/cm^2) generated monoadducts (1%), which increased with intensity as the dose increased. After 32 minutes of UV irradiation (2 J/cm^2), 39% monoadduct formation was detected, however no crosslinks were detected. No DNA band shifts were detected when Puv15a was absent from the control reactions.

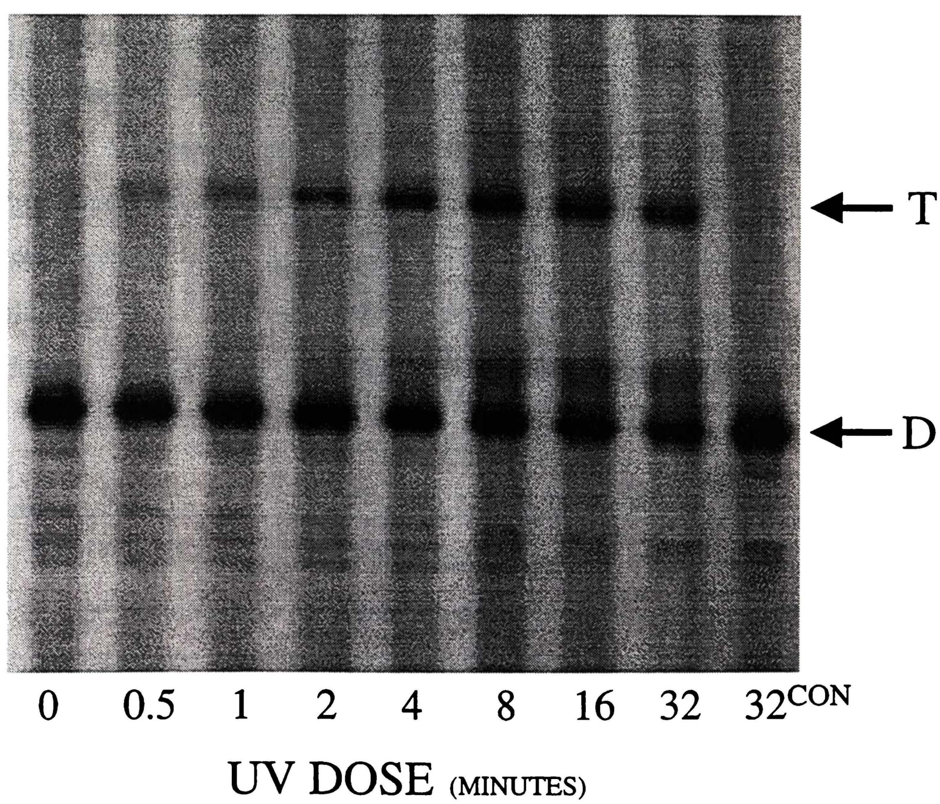


Figure 3.12: Electrophoretic mobility shift assay showing photo-induced crosslinking of a 16 nt purine rich psoralen-TFO to a 49 bp Bov*REED* dsDNA target. The γ-^{33}P end-labelled duplex was incubated with 40 μM Puv15a, 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated on ice, 3 cm distance from a UV lamp (360 nm) for 0-32 minutes (maximum UV dose - 2 J/cm^2). The samples were electrophoresed on a 8% TBE denaturing PAGE gel. 32^{CON} represents the control reaction (psoralen-TFO absent, 32 minutes of UV radiation).
T = DNA triplex, D = DNA duplex.

A control TFO was included in these UV crosslinking assays, TFO-*REED*, which has the identical DNA sequence to Puv15a but with no psoralen moiety at the 5' end. This control TFO was incubated with radioactively labelled target duplex, irradiated and subjected to denaturing gel electrophoresis in parallel with equivalent reactions using Puv15a. Under the conditions of this assay, it was not possible to detect any DNA triplex formation as non-covalent DNA triplexes would be disrupted during denaturing PAGE electrophoresis (results not shown).

3.3.5 Bov*TSPY* DNA Triplex Formation

A 39 bp DNA fragment of the *Bos taurus* TSPY gene (GenBank accession n# X74028) was employed as a dsDNA target for DNA triplex formation and UV crosslinking studies. This fragment contains a consecutive run of 17 purine bases (Figure 3.13). A homopurine psoralen-conjugated TFO, Puv14a, which is composed of 65% guanines, was designed to bind to the purine strand of the DNA duplex via Hoogsteen hydrogen bonding, forming a DNA triplex. A psoralen intercalation site was available at the 5' end of the DNA triplex-duplex junction, forming a thymine monoadduct or a cytosine monoadduct after UV exposure (360 nm).

3′ GAGAAGGGGAGAAGGGG≈Pso 5′	Puv14
5′ CAGGAGGACGAGAAGGGGAGAAGGGG**C**AGGAGCAGGCGA 3′ 3′ GTCCTCCTCCTCTTCCCCTCTTCCC**C**GTCCTCGTCCGCT 5′	Bov*TSPY* DNA DUPLEX (39 bp)

Figure 3.13: A schematic representation of Bov*TSPY* DNA triplex formation with psoralen-conjugated TFO Puv14. Psoralen intercalates with either the cytosine or thymine base on the antisense strand, at the 5' end of the triplex-duplex junction of the DNA duplex, as shown in bold lettering. Note that ≈Pso represents psoralen.

Figure 3.14 shows that 8 minutes of UV irradiation (0.5 J/cm^2) gave detectable monoadducts (10%), however, the level of formation did not increase after thirty minutes of UV irradiation (2 J/cm^2). As expected, no crosslinks were detected. The psoralen could only intercalate into the pyrimidine bases on the antisense strand of the DNA duplex, forming monoadducts (Figure 3.13). Hence Puv14a was not targeted to larger DNA fragments.

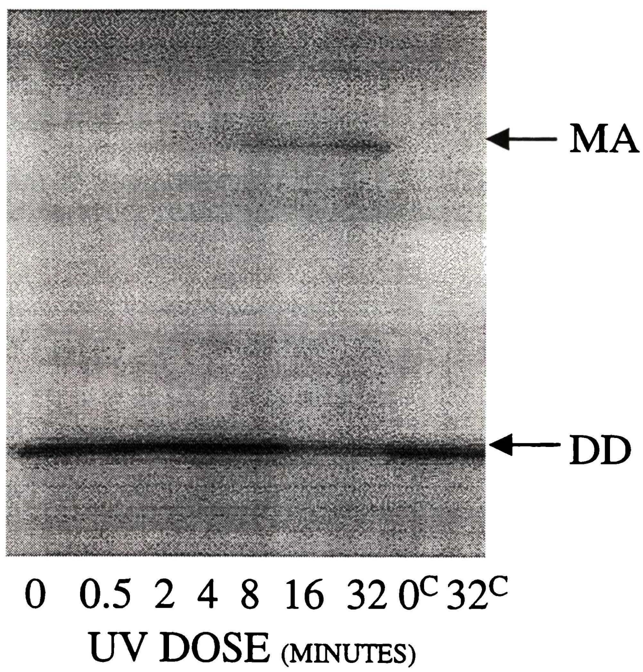


Figure 3.14: Effect of UV dose on Bov*TSPY* DNA triplex formation. γ-^{33}P end-labelled Bov*TSPY* duplex (39 bp) was incubated with 40 μM Puv14a, 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated on ice for 0- 32 minutes. 0^C and 32^C are control samples that contains no psoralen-conjugated TFO and were subjected to zero and thirty two minutes of UV irradiation (maximum UV dose- 2 J/cm^2), 3 cm distance from the UV lamp (360 nm). The samples were then electrophoresed on a 8% TBE denaturing PAGE gel. MA = monoadduct, DD = DNA duplex.

3.4 UNSUCCESSFUL TARGETING OF TFOs

A variety of psoralen-conjugated TFOs failed to form photoadducts when targeted *in vitro* to synthetic dsDNA targets Bov*Grobet*, Bov*REED2* and Bov*REED3*. These dsDNA targets represent a small region of bovine repetitive Y-chromosomal DNA sequences (Figure 3.15).

Bov*Grobet* DNA triplex:

3′ GGAGAAAAAGGAAGGGAG≈Pso 5′	Puv1
5′ Pso≈CCTCTTTTTCTTTCCCTC 3′	Puv2
5′ Pso≈CCTCTTTTTC 3′	Puv3
3′ GGAGAAAAAG≈Pso 5′	Puv4
5′ **C**GGAGAAAAAG**C**AAGGGAG**T**TCAAGATATACATCTACAT 3′ 3′ G**C**CTCTTTTT**C**GTTCCCT**C**AAGTTCTATATGTAGATGTA 5′	DNA duplex (40 bp)

Bov*Reed2* DNA triplex:

3′ AAAGAGAAAAGGGGAG≈Pso 5′	Puv7a
5′ AAAGAGAAAAAGGGAG**C**ACAAGACC 3′ 3′ TTTCTCTTTTTCCCT**C**GTGTTCTGG 5′	DNA duplex (25 bp)

Bov*Reed3* DNA triplex:

3′ ATTGAAAGTG≈Pso 5′	Puv8a
3′ AGGGAAAGGG≈Pso 5′	Puv8b
3′ AGGGGTAGAGGGAAAGGG≈Pso 5′	Puv8c
5′ GCAGTGTCATATTGAAAGTG**T**CCA 3′ 3′ CGTCACAGTATAACTTTCA**C**AGGT 5′	DNA duplex (24 bp)

Figure 3.15: A schematic representation of unsuccessful targeting of psoralen-conjugated TFO to synthetic dsDNA fragments (represents part of a repetitive bovine Y-chromosomal DNA sequence). The psoralen intercalation sites are shown in bold lettering. Note that ≈Pso represents psoralen. The complete bovine DNA sequences of Bov*Grobet*, Bov*Reed2* and Bov*Reed3* are listed in Appendix III.

The TFO and/or target site may have been unsuitable due to a number of reasons such as the presence of nucleotide mismatch(s) within the homopurine region of the triplex site, a cytosine nucleotide for psoralen intercalation rather than a thymine, and/or a high adenine content. Alternatively the dissociation kinetics of the DNA triplex may have been too fast.

Previous studies have shown that psoralen-conjugated TFOs have to be modified to increase the stability of the DNA triplex and prevent nuclease degradation *in vivo*. Pyrimidine psoralen-conjugated TFOs; Puv3, Puv4 and Puv12d failed to form detectable photoadducts upon UV irradiation. Based on these results, modified pyrimidine psoralen-conjugated TFOs were designed and synthesised which incorporated a psoralen moiety at the 5' end and three phosphorothioate linkages at the 3' end. This modification should increase the stability of the DNA triplex and prevent nuclease degradation *in vivo*. The cytosine nucleotides were also methylated to increase the pK_a, thus increasing the DNA triplex stability at physiological pHs.

A strict criterion was imposed for using modified TFOs, as these TFOs are expensive to synthesise. There was at least one 5' TpA site for psoralen intercalation and a homopurine run of twelve bases. Hence Puv3, Puv4 and Puv12d were not ideally suitable and new TFOs were designed to bind and form triplexes to alternative regions within the REED and REP3 DNA sequences (Appendix III).

3.5 EFFECT OF USING MODIFIED TFOs

3.5.1 Bov*REP4* DNA TRIPLEX FORMATION

A 40 bp DNA fragment of the 306 bp *Bos taurus* Y chromosome repeat (GenBank accession n# U39887, Appendix III) was employed as a dsDNA target for potential DNA triplex formation and UV crosslinking studies. This fragment (Bov*Rep4*) contains a consecutive run of 13 nt purine bases, enabling the targeting of a modified psoralen-conjugated TFO *m*Puv6d[9]. This 13 nt pyrimidine rich TFO binds parallel to the dsDNA duplex. The sequences of the target duplex and reactive psoralen-conjugated TFOs are shown in Figure 3.16. The psoralen can intercalate in between a cytosine and a thymine base at the 5' end of the triplex-duplex junction. There are three nucleotide mismatches at the 3' end of the dsDNA target; one thymine and two cytosine bases. Hence, a thymine and a cytosine have been incorporated into *m*Puv6d, forming a T·CG and C·TA triplet respectively.

5' Pso≈CTCCTCCCCCTTC 3' *m*Puv6d

5' TTGCC**T**GAGGAGGGGGCCTGGGCATCTGAAAGTTCAATGT 3' Bov*REP4*
3' AACGGA**C**TCCTCCCCCGGACCCGTAGACTTTCAAGTTACA 5' (40 bp)

Figure 3.16: A schematic representation of Bov*REP4* DNA triplex formation with a modified psoralen-conjugated TFO, *m*Puv6d. Psoralen intercalates into the pyrimidine bases on the sense or antisense strand, at the 5' end of the triplex-duplex junction of the Bov*REP4* DNA duplex, as shown in bold lettering. The cytosine bases of *m*Puv6d are methylated and also, the last three nucleotides at the 3' end have phosphorothioate linkages. Note that ≈Pso represents psoralen.

[9] *m*Puv6d is a modified TFO, i.e psoralen is present at the 5' end, the cytosine bases are methylated and the last three nucleotides at the 3' end have phosphorothioate linkages.

$\gamma^{33}P$ UV crosslinking studies were performed using this modified TFO and dsDNA target as previously described. Figure 3.17 shows that stable photoadducts are formed when 10 µM *m*Puv6d is targeted to Bov*REP4* and irradiated with UV light (360 nm). Following thirty seconds of UV irradiation (0.03 J/cm^2), 1% monoadduct formation was detected. As the UV dose increased to fifteen minutes (1 J/cm^2), 50% monoadduct and 20% crosslink formation was detected. Under the conditions of this assay it was not possible to detect monoadducts with *m*Puv6d concentrations of less than 0.1 µM. A small degree of crosslinking (<1%) was observed when the DNA triplex assay was performed with 1 µM *m*Puv6d after exposure to thirty minutes of UV light (2 J/cm^2).

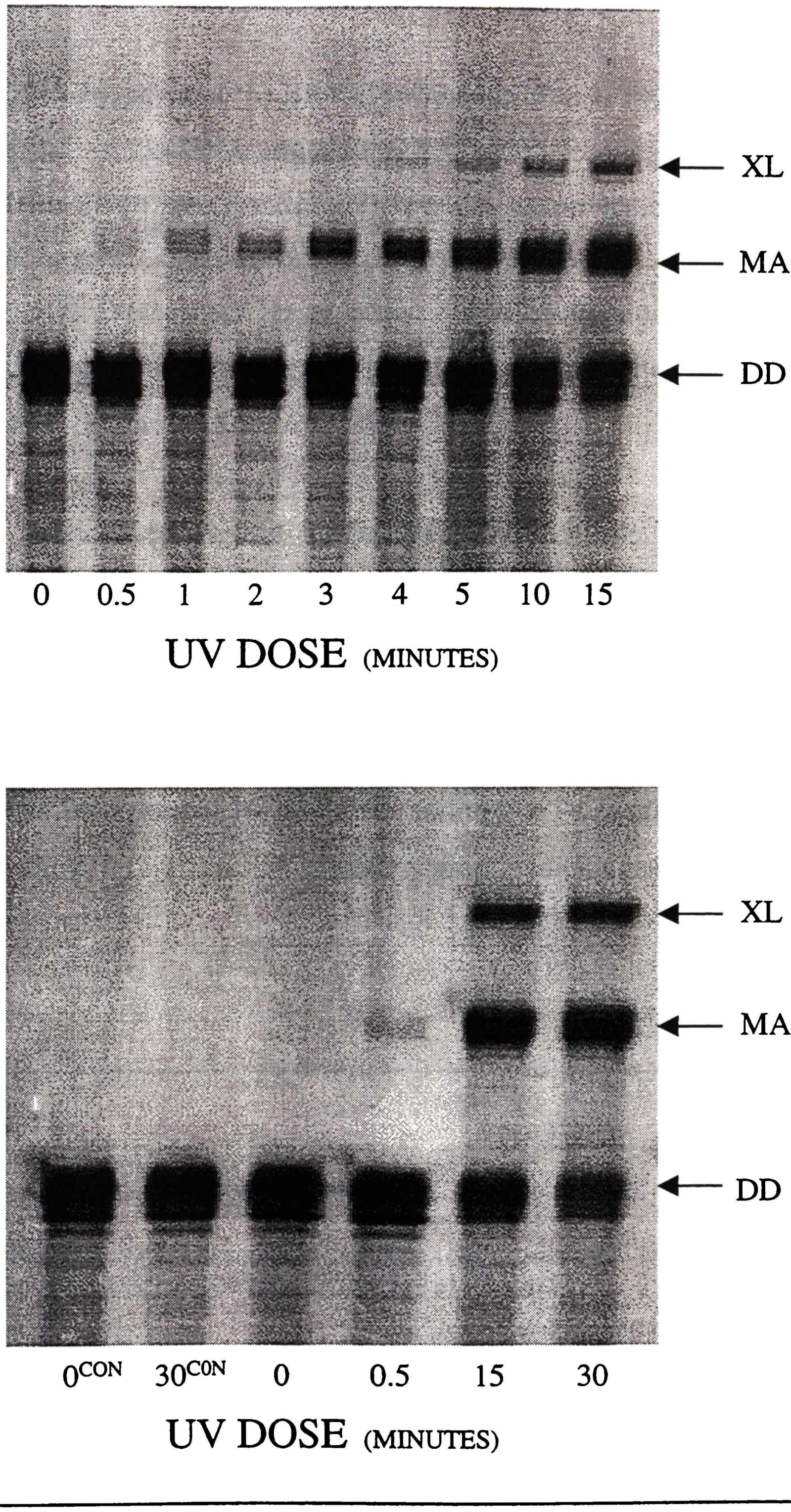


Figure 3.17: Electrophoretic mobility shift assay showing photo-induced crosslinking of a 16 nt pyrimidine rich modified psoralen-TFO to a 40 bp Bov*Rep4* dsDNA target. The γ-^{33}P end-labelled duplex was incubated with 10 µM *m*Puv6d, 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated on ice, 3 cm distance from the UV lamp (360 nm) for 0 - 30 minutes (maximum UV dose – 2 J/cm^2). The samples were electrophoresed on an 8% TBE denaturing gel. The control samples, O^{CON} and 30^{CON} had *m*Puv6d absent from the DNA triplex assay. XL = crosslink, MA = monoadduct, DD = 40 bp DNA duplex.

3.5.2 Bov*REED2* DNA TRIPLEX FORMATION

A 40 bp DNA fragment of the BRY11a sequence (Patent n# 5459038, Appendix III) was employed as a dsDNA target for DNA triplex formation and UV crosslinking studies. This fragment (Bov*REED2*) contains a consecutive run of 12 nt purine bases, enabling the targeting of modified psoralen-conjugated TFOs, *m*Puv7b and *m*Puv7c. An unmodified 12 nt purine psoralen-conjugated TFO (Puv7a) was previously targeted to this triplex region and failed to form any detectable photoadducts. (Figure 3.18).

Modified psoralen-conjugated TFOs, *m*Puv7b and *m*Puv7c both consisted of pyrimidine bases only and were designed to bind parallel to the dsDNA duplex. *m*Puv7b was 12 nt in length with no mismatches for third strand binding whereas *m*Puv7c was 15 nt in length with the presence of one nucleotide mismatch. There was a cytosine mismatch at the 3' end of the homopurine triplex region, thus a thymine was incorporated into *m*Puv7c, forming a T·CG triplet. The sequences of the target duplex and reactive modified psoralen-conjugated TFOs are shown in Figure 3.18.

5'Pso≈TTTCTCTTTTCT	*m*Puv7b
5'Pso≈TTTCTCTTTTCTTTC	*m*Puv7c
5' TTAACTTT**T**AAAGAGAAAAGACAGCACAAGACCCAGTGTG 3' 3' AATTGAAAA**T**TTCTCTTTTCTGTCGTGTTCTGGGTCACAC 5'	Bov*REED2* (40 bp)

Figure 3.18: A schematic representation of Bov*REED2* DNA triplex formation with modified psoralen-conjugated TFOs, *m*Puv7b and *m*Puv7c. Psoralen intercalates into the pyrimidine bases on the sense or antisense strand, at the 5' end of the triplex-duplex junction of the DNA duplex, as shown in bold lettering. The cytosine bases present in Puv6d are methylated in Puv6d and also, the last three nucleotides at the 3' end have phosphorothioate linkages present. Puv7a (5' Pso≈GAAAAGAGAAA 3') binds antiparallel to the DNA duplex and is an unmodified psoralen-conjugated TFO. Note that ≈Pso represents psoralen.

$\gamma^{33}P$ UV crosslinking studies were performed using this modified TFO and dsDNA target as previously described. Figure 3.19 shows that stable photoadducts were formed when 10 μM *m*Puv7b and *m*Puv7c were targeted to Bov*REED2* and irradiated with UV light (360 nm). Thirty seconds of UV irradiation (360 nm, 0.03 J/cm^2) caused 50% monoadduct formation with *m*Puv7b as compared to 40% with *m*Puv7c. However, as the UV dose increased to fifteen minutes (1 J/cm^2), a significant number of crosslinked bands were detected in both cases, due to a series of thymine bases available for psoralen intercalation at the 5' end of the triplex-duplex junction (Figure 3.31). After thirty minutes of UV irradiation (2 J/cm^2), most of the DNA duplex was converted to a crosslink. Under the conditions of this assay it was not possible to detect photoadducts in the presence of *m*Puv7b and *m*Puv7c concentrations of less than 0.2 μM (results not shown).

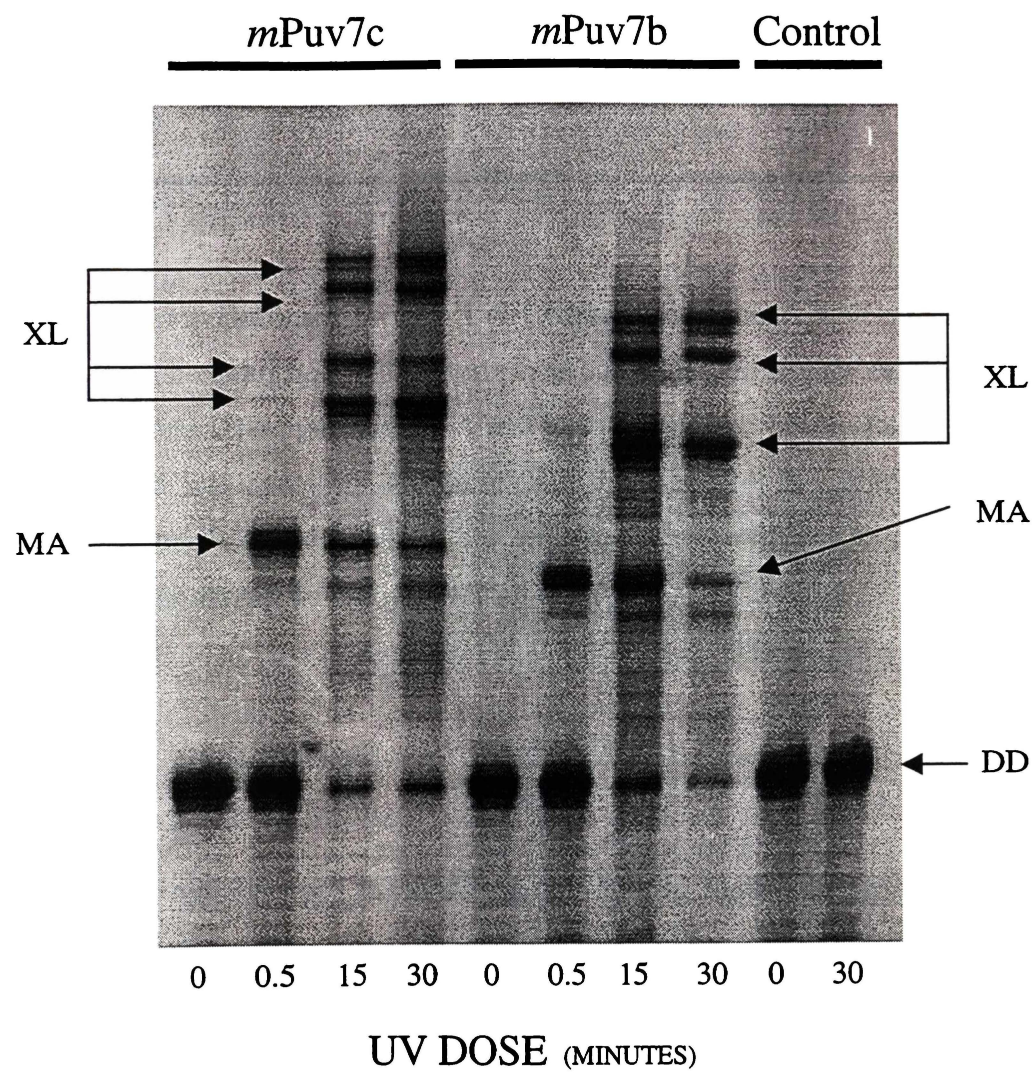


Figure 3.19: Photo-induced crosslinking of modified pyrimidine rich psoralen-conjugated TFOs to a 40 nt dsDNA target. The γ-^{33}P end-labelled Bov*Reed2* duplex was incubated with 10 μM *m*Puv7b or *m*Puv7c, 1 mM spermidine, 1X DNA triplex binding assay (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated with UV light (360 nm) on ice for 0-30 minutes, 3 cm distance from the UV lamp (maximum UV dose-2 J/cm^2). The samples were then electrophoresed on a 8% TBE denaturing PAGE gel. *m*Puv7b and *m*Puv7c were absent from the control samples. XL = crosslink, MA = monoadduct, DD = DNA duplex.

Chapter 4

REPETITIVE Y CHROMOSOME DNA SEQUENCES

Following these preliminary *in vitro* DNA triplex experiments with small dsDNA targets, psoralen-conjugated TFOs were tested with larger bovine dsDNA targets such as amplified DNA, plasmid DNA and bovine genomic DNA. As a prelude to these studies it was necessary to characterise the repetitive Y chromosome DNA sequences. Thus PCR and Southern blot analyses were performed to determine the specificity of the chosen bovine dsDNA target.

4.1 PCR ANALYSIS

PCR primers were designed to amplify a 100-500 bp region containing the DNA triplex site (Appendix III). The PCR was performed in the presence of either male or female bovine genomic DNA (gDNA). Figure 4.1 shows that amplification of male genomic DNA with FBov*REED*[8] and USP2[9] primers resulted in a single 400 bp PCR product that contains the selected Bov*REED* target triplex region. A single band was also observed for the following bovine male amplifications with each of their specific primer pairs: a 500 bp Bov*TSPY* PCR product, a 200 bp Bov*SRY* PCR product, and a 100 bp Bov*REP3* PCR product. In all these cases, amplification of bovine female genomic DNA failed to generate a PCR product.

Figure 4.1 also shows that the amplification of the *sup*F target from the pS189 shuttle vector results in a single band, approximately 200 bp in size.

[8] FBov*REED* refers to the forward primer that amplifies the Bov*REED* DNA triplex region.
[9] USP2 refers to the reverse primer that amplifies the Bov*REED* DNA triplex region.

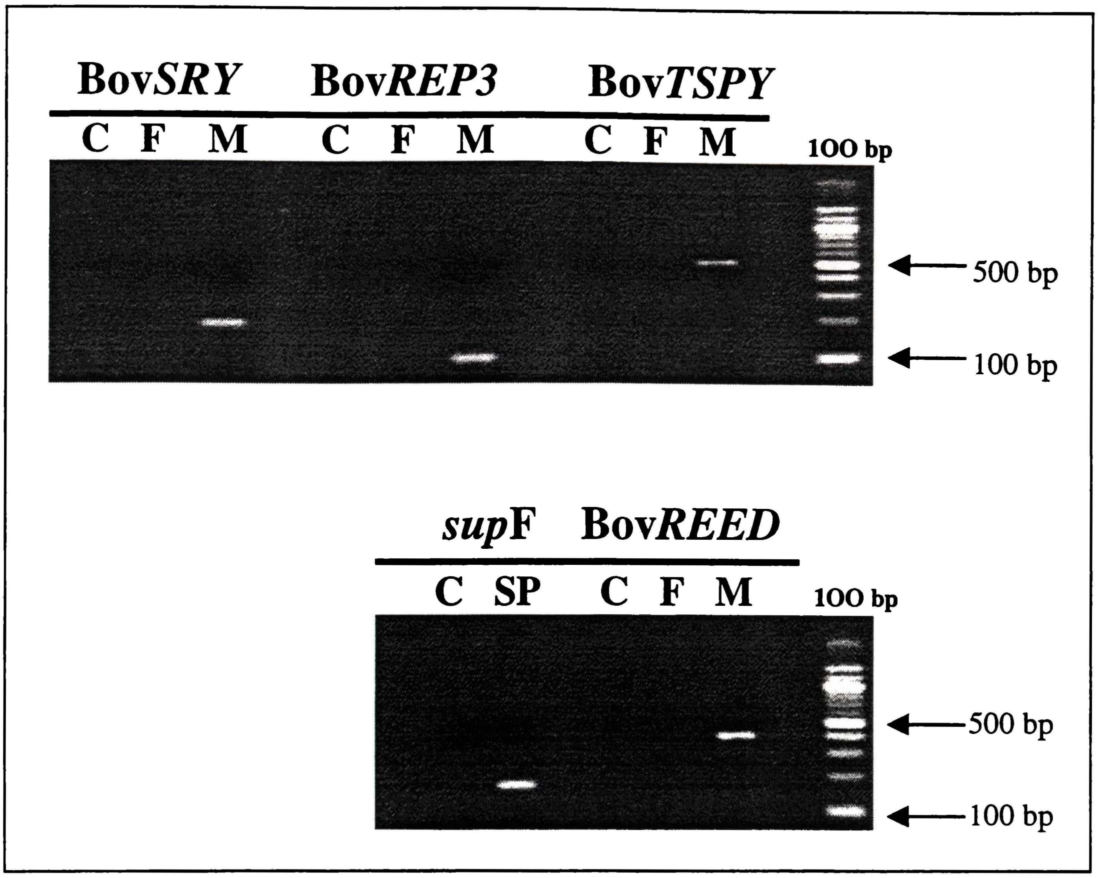


Figure 4.1: Amplification of bovine genomic DNA that contains a specific bovine Y chromosome DNA triplex site. The Polymerase chain reaction was performed as stated in Section 2.9 in the presence and absence of 10 ng/µL bovine male gDNA (M) or female gDNA (F). In the case of *sup*F amplifications, pS189 shuttle vector was used as the DNA template (SP). Note that "C" represents the control where no DNA template was present in the PCR reaction. The PCR samples were diluted with 1X Loading Dye (Type II) and 10 µL volumes were loaded onto a 2% TAE agarose gel containing 0.5 µg/mL EtBr (Agarose-1000, Life Technologies, USA). Electrophoresis was performed for one hour at 100 volts. The 100 bp DNA ladder (New England Biolabs, UK) generated the following fragment sizes: 1500, 1200, 1000, 900, 800, 700, 600, 500, 400, 300, 200, and 100 bp.

When using these male specific products as radioactive probes, the PCRs were performed after radioactively labelling the 5' end of the forward or reverse primer with [γ^{33}P]ATP and polynucleotide kinase, or alternatively, labelling the PCR product with [α^{32}P]dCTP during the PCR synthesis step. Following radioactive PCR, the amplified products were used as Southern blot probes or, alternatively, as dsDNA targets for *in vitro* UV crosslinking assays.

4.2 DNA SEQUENCING

Non-radioactive PCR fragments were DNA sequenced by the Waikato DNA Sequencing Facility to confirm that the correct DNA sequence was amplified and that the triplex region agreed with the published sequence. The forward and reverse primers that were used in each PCR reaction were also used for the DNA sequencing reaction (Table 2.3). DNA sequencing analysis of amplified fragments, *sup*F, Bov*REED*, and Bov*REP3*, showed that there were no significant nucleotide changes with respect to the published GenBank sequences (Appendix III). This was also observed for Bov*SRY* (results not shown). Each electropherogram was of good technical quality, indicating the amplification of one DNA template (Figure 4.2, 4.3, 4.4). There was also a lack of degenerate bases (represented by an N on the electropherogram). Therefore, these results suggest that a single sequence was present in those regions of the Y-chromosomal repeats rather than degenerate sequences containing variations of a basic consensual sequence.

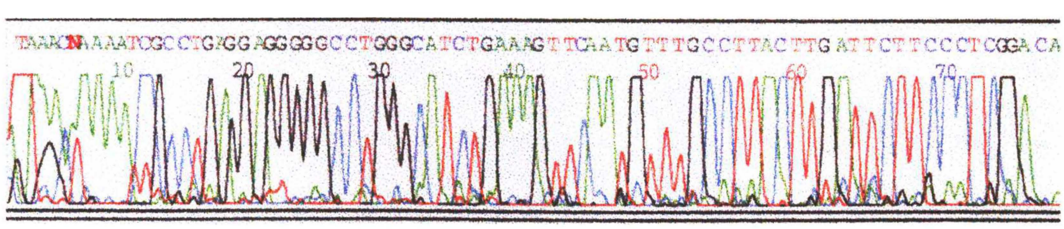


Figure 4.2: An Electropherogram showing the DNA triplex region of the sequenced Bov*REP3* PCR product. The FBov*REP3* primer (Table 2.3) was used for the DNA sequencing reaction. The homopurine DNA triplex region is from nucleotides 16 to 26. The N at position 6 is probably a technical artefact.

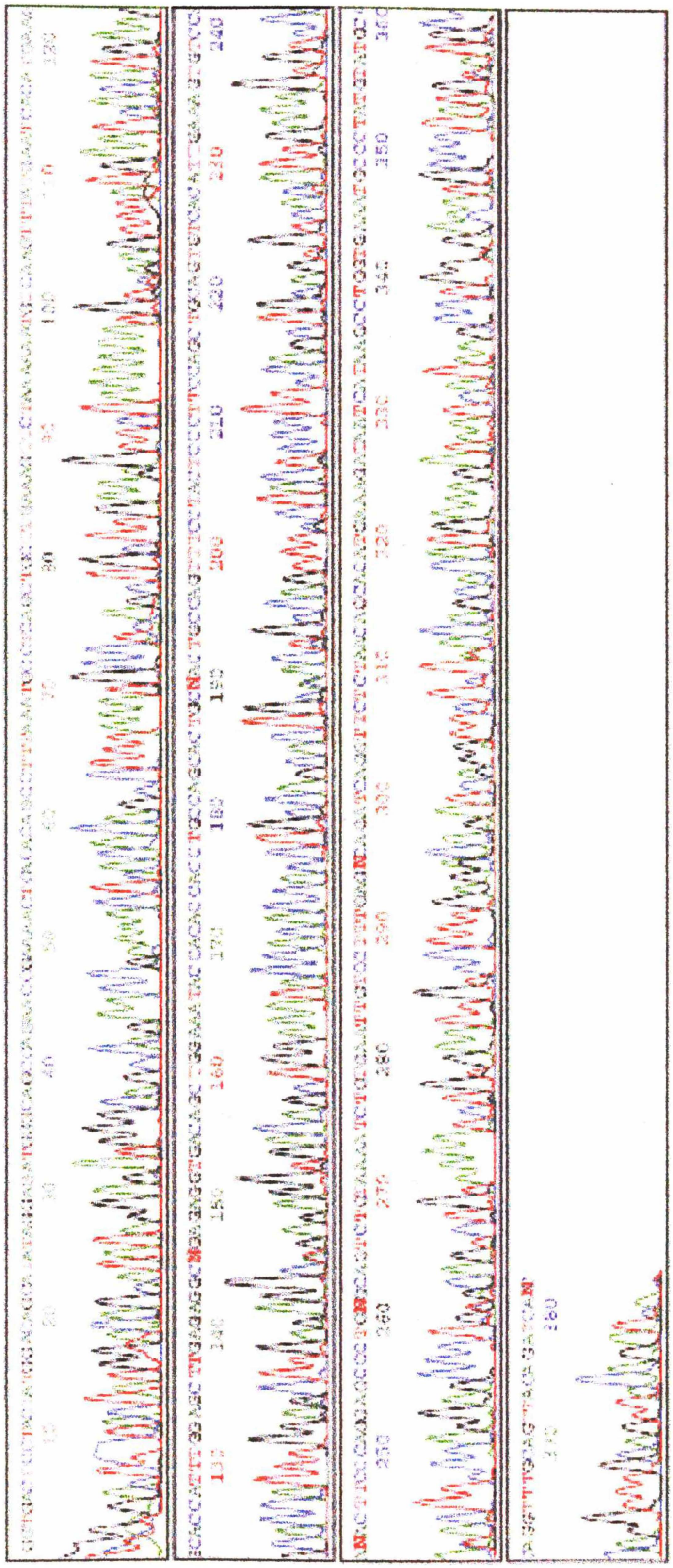


Figure 4.3: An Electropherogram showing the DNA triplex region of the sequenced Bov*REED* PCR product. The FBov*REED* primer (Table 2.3) was used for the DNA sequencing reaction. The homopyrimidine DNA triplex region is from nucleotides 199 to 214.

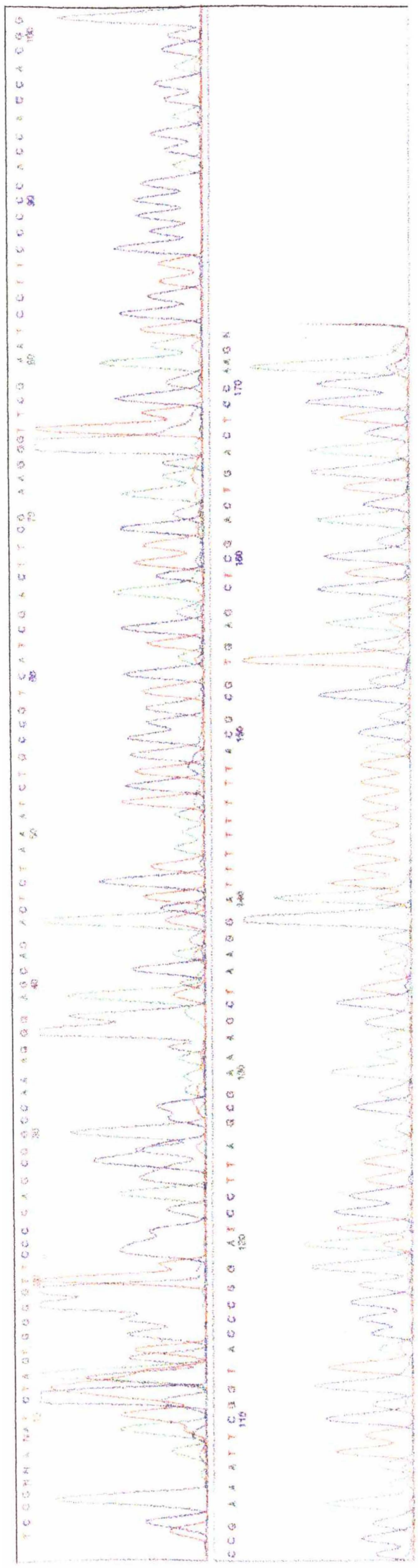


Figure 4.4: An Electropherogram showing the DNA triplex region of the sequenced *sup*F PCR product. The F*sup*F primer (Table 2.3) was used for the DNA sequencing reaction. The homopyrimidine DNA Triplex region is from nucleotides 82 to 91.

In order to investigate even larger DNA targets for UV crosslinking studies, the amplified DNA containing the specific DNA triplex region was cloned into a Bluescript II SK^+ vector. Once the PCR products were cloned into the Bluescript plasmid, positive identification of cloned inserts was determined by restriction digestion analysis and also by PCR. Plasmids containing the PCR insert were then DNA sequenced to ensure that the PCR product was indeed cloned and identical to the PCR product and GenBank sequence. The T7 primer[10] (22 nt, Stratagene, USA), M13 primer[11] (17 nt, Stratagene, USA) or alternatively the PCR forward or reverse primers (Table 2.3) were used for the DNA sequencing reaction. Electropherogram analysis of plasmid-Bov*REED*, plasmid-Bov*SRY* and plasmid-Bov*REP3* showed well-defined sharp peaks, lack of degenerate bases (N) and no significant nucleotide changes compared to the cloned PCR product and published GenBank sequences (Figure 4.5, 4.6, 4.7, Appendix III).

[10] The T7 primer has the following nucleotide sequence: 5' GTAATACGACTCACTATAGGGC 3'

[11] The M13 primer has the following nucleotide sequence: 5' GTAAAACGACGGCCAGT 3'

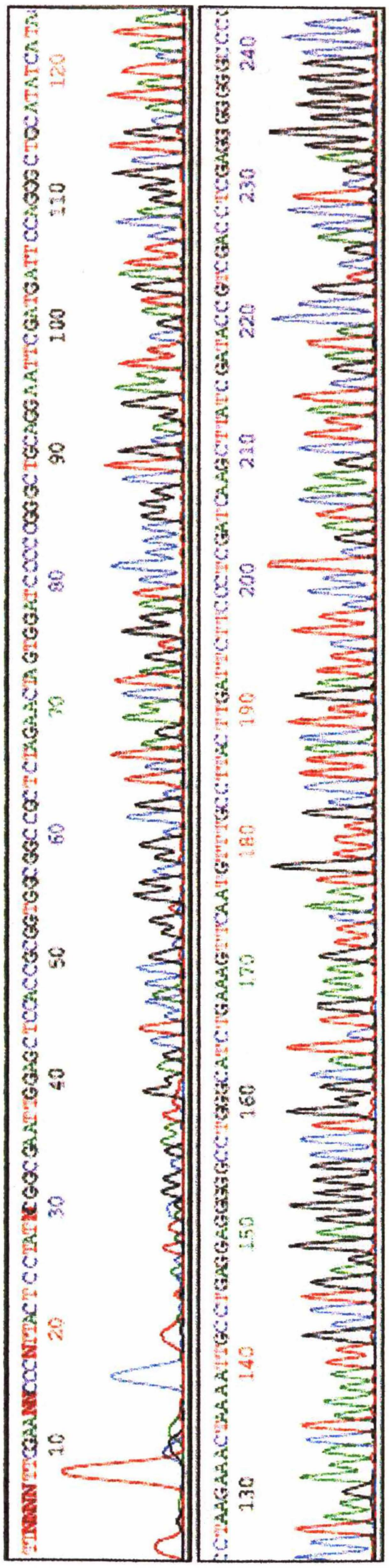


Figure 4.5: An Electropherogram showing the DNA sequence of the Bluescript II SK+ vector that contains the cloned Bov*REP3* PCR insert. The M13 primer was used for the DNA sequencing reaction. The FBov*REP3* primer sequence is from nucleotides 103 to 124 and the RBov*REP3* primer is from nucleotides 182 to 203. The homopurine DNA triplex region is from nucleotides 146 to 155.

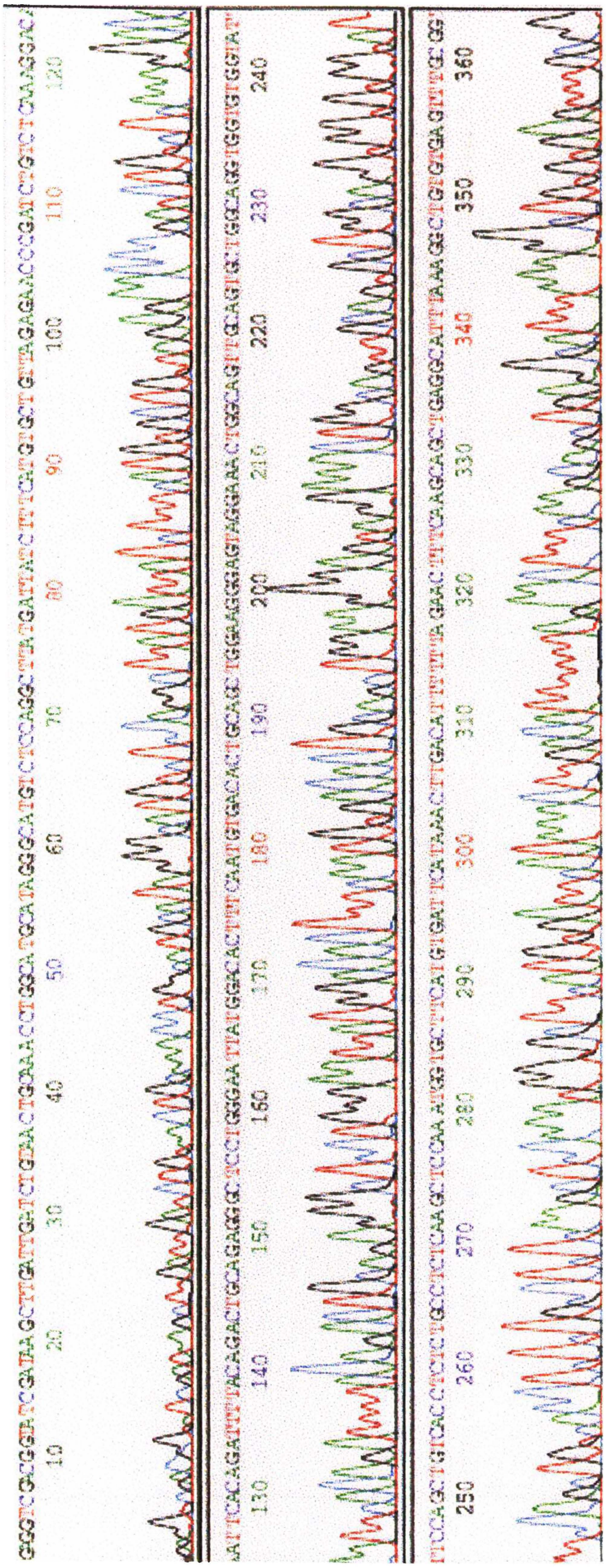


Figure 4.6: An Electropherogram showing the DNA sequence of the Bluescript II SK^{+} vector that contains the cloned Bov*REED* PCR insert. The T7 primer was used for the DNA sequencing reaction. The RBov*REED* primer sequence is from nucleotides 29 to 50 and the homopurine DNA triplex region is from nucleotides 195 to 210.

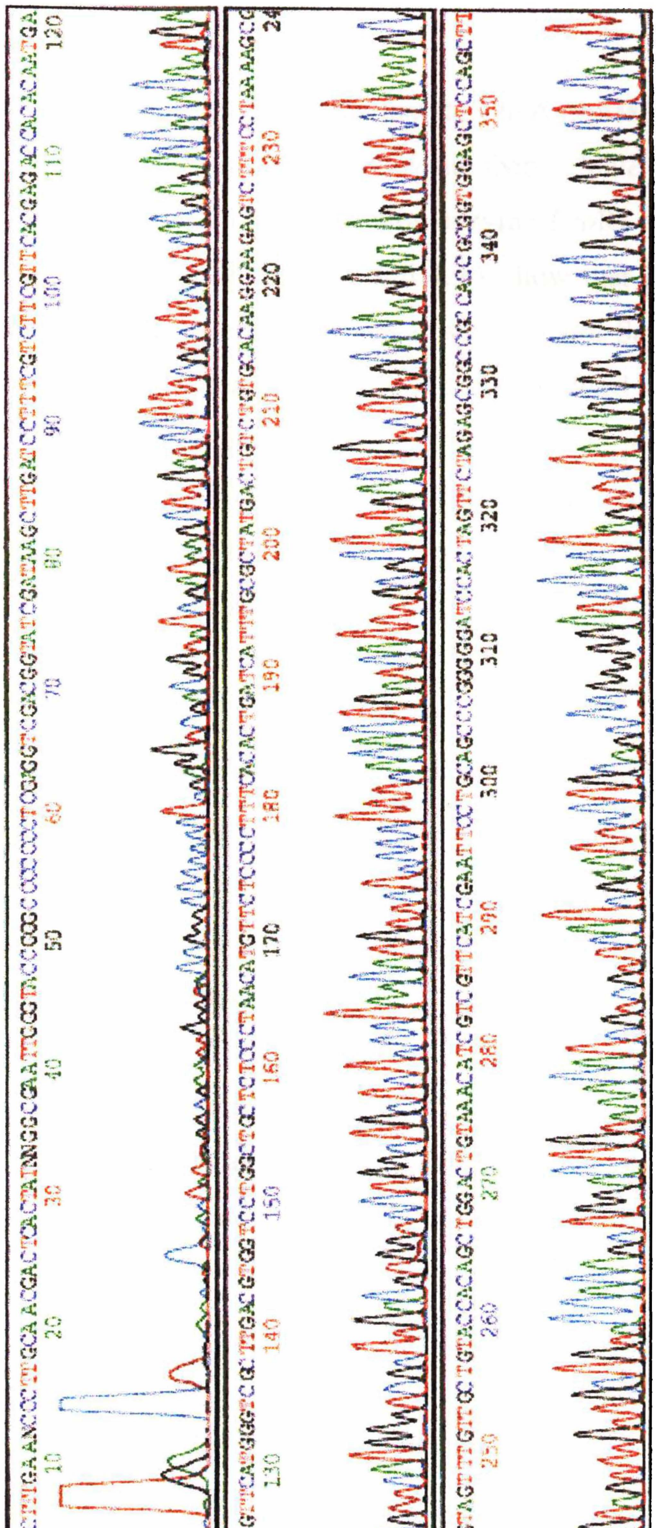


Figure 4.7: An Electropherogram showing the DNA sequence of the Bluescript II SK^{+} vector that contains the cloned Bov*SRY* PCR insert. The M13 primer was used for the DNA sequencing reaction. The FBov*SRY* primer sequence is from nucleotides 264 to 288 and the RBov*SRY3* primer is from nucleotides 89 to 113. The hompyrimidine DNA triplex region is from nucleotides 171 to 192.

4.3 SOUTHERN BLOT ANALYSIS

Southern blots were performed to determine that the DNA triplex region was specific to bovine male DNA and present in high copy numbers. Male and female bovine genomic DNA was digested with the restriction enzyme *Eco*RI and electrophoresed on a 1% TAE agarose gel. Figures 4.8A and 4.9A show that genomic DNA was completely digested and generated a long smear with a major Satellite I band at 1.4 Kb. The agarose gel was then denatured, blotted and hybridised with a male specific PCR probe. Theoretically, the probes should at least bind to the following bovine *Eco*RI digested fragments; 250 bp (Bov*REP3* probe), 1800 bp (Bov*SRY* probe), 3200 bp (Bov*TSPY* probe), and 400 bp (Bov*REED* probe). The size of the fragment ultimately depends on where the second unknown *Eco*RI restriction site is within the bovine genome.

Control Southern blots were performed using plasmid DNA that containing a single copy of the bovine DNA triplex target. These were compared against bovine genomic DNA Southern blots. Single and double restriction digests were performed to show that the PCR probe could bind to the undigested plasmid and also to the PCR insert of the double digested plasmid. A Southern blot using a 173 bp Bov*USP* probe, as previously investigated by Reed *et al.*, (1995), was also repeated as a control to demonstrate male specificity and repetitiveness of the Reed *et al.*, (1995) DNA sequence. This probe was amplified using the universal sex primers USP1 and USP2 (Table 2.3). The USP1 primer contains half the DNA triplex target sequence of Bov*REED* (Figure 3.11).

Following hybridisation, the membranes were washed and subjected to autoradiography. Analysis of the resulting autoradiographs showed no evidence of a smear ("ladder effect"), which is often associated with tandemly repeated sequences (Smith *et al.*, 1987). Figure 4.8B shows that Bov*REED* probe binds to male gDNA with strong bands observed at 4 and 5 Kb with a extra band observed at 9 Kb for testicular tissue. Possibly, the testicular DNA is methylated. Thus, these results suggest that this probe is specific to bovine male DNA.

In the plasmid control, Bov*REED* probe hybridised to digested and undigested Plasmid-Bov*REED* (results not shown). The Bov*REED* probe hybridised significantly to the digested bovine genomic DNA compared to the plasmid DNA. Thus, these results suggest that the copy number of the male specific repeats was less than 500.

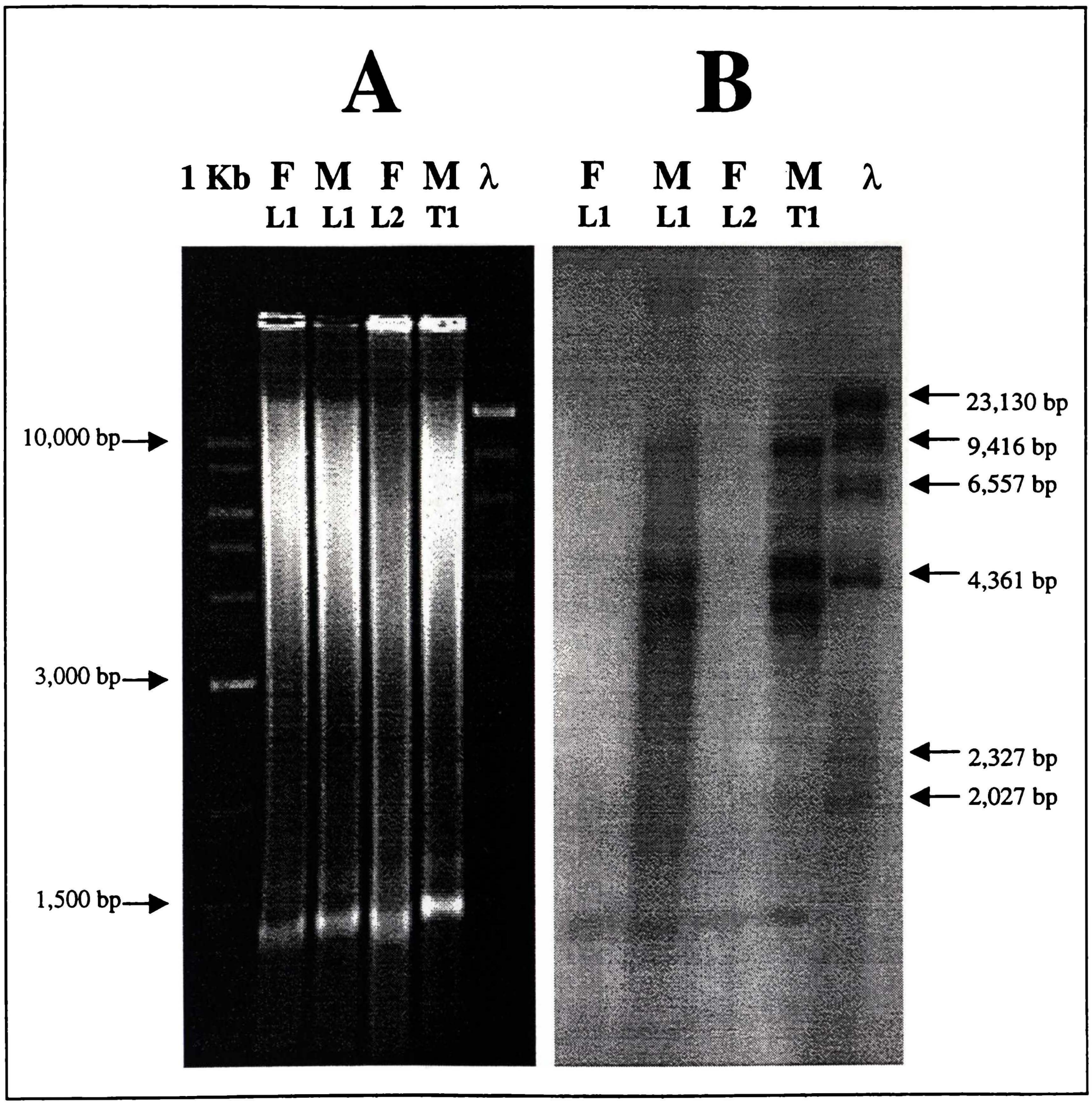


Figure 4.8: Southern Blot analyses of male and female bovine DNA. Half a microgram of Testicle (T) and Liver (L) bovine male (M) genomic DNA, and Liver female bovine (F1, F2) genomic DNA was digested with *Eco*RI, electrophoresed and transferred to a Hybond-N+ membrane under alkaline conditions. The membrane was hybridised with a [α^{32}P]dCTP labelled probe. A 1 Kb DNA ladder (New England Biolabs, UK), and λ DNA/*Hind*III radioactive fragments were used as molecular markers. **A)** Ethidium bromide (0.5 µg/mL) stained 1% TAE agarose gel of *Eco*RI digested male and female genomic DNA. The 1 Kb ladder (0.5 µg) generates the following sized fragments: 10, 8, 6, 5, 4, 3, 2, 1.5 and 1 Kb. **B)** Southern blot hybridisation with Bov*Reed* probe at 65°C. High molecular weight bands were observed in the male DNA samples only.

In comparison, the smaller Bov*USP* probe binds to the same bands as Bov*REED* probe as shown in Figure 4.9B.

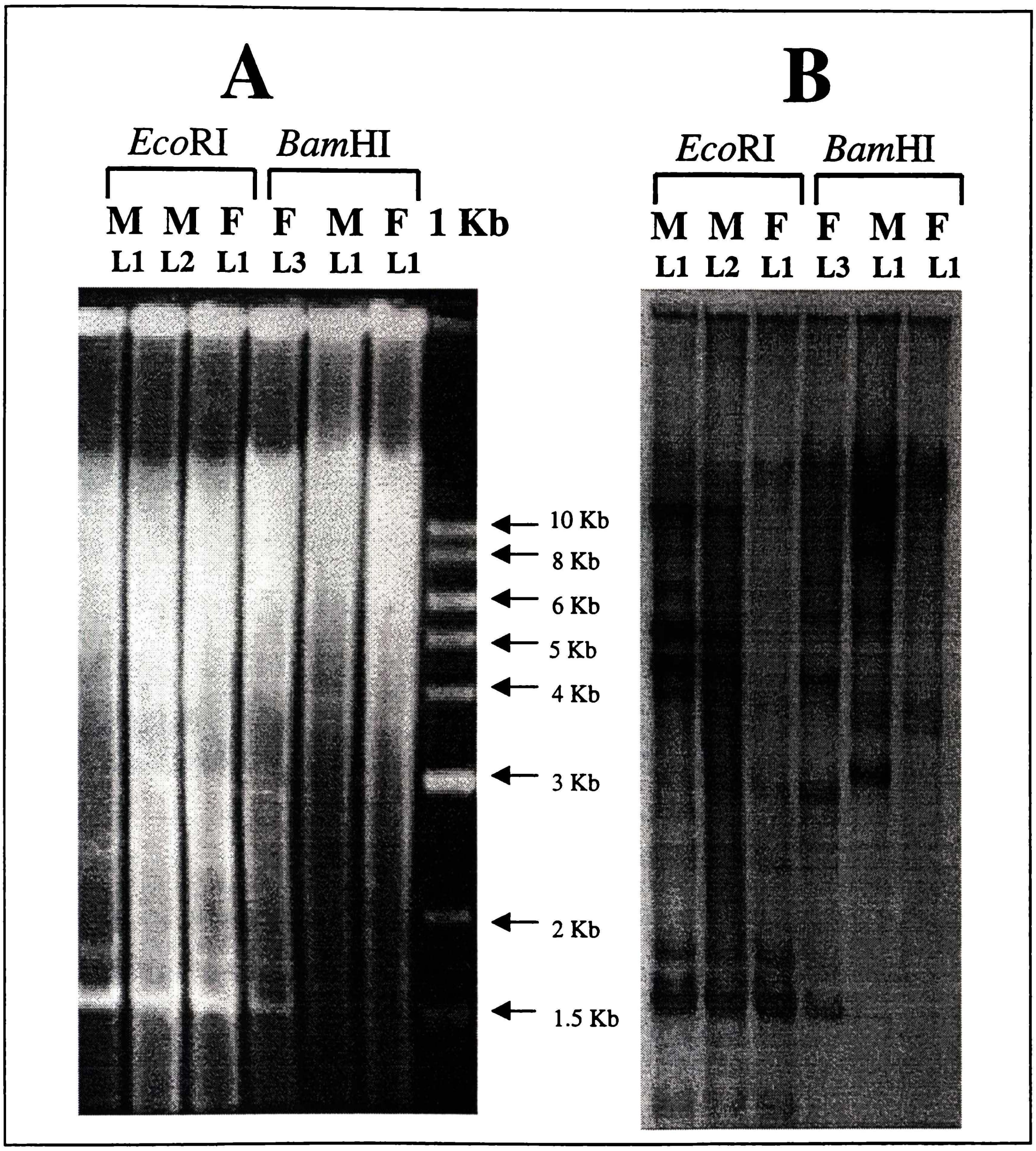


Figure 4.9: Southern Blot analyses of male and female bovine DNA. One microgram of Testicle (T) and Liver (L) bovine male (M) genomic DNA, and Liver female bovine (F1, F3) genomic DNA was digested with *Eco*RI or *Bam*HI, electrophoresed and transferred to a Hybond-N$^+$ membrane under alkaline conditions. The membrane was hybridised with a [α^{32}P]dCTP labelled probe. A 1 kb DNA ladder (0.5 µg, New England Biolabs, UK) was used as a molecular marker. **A)** Ethidium bromide (0.5 µg/mL) stained 1% TAE agarose gel of *Eco*RI or *Bam*HI digested bovine genomic DNA. **B)** Hybridisation with Bov*USP* probe at 68°C. High molecular weight bands were observed in the male DNA samples only.

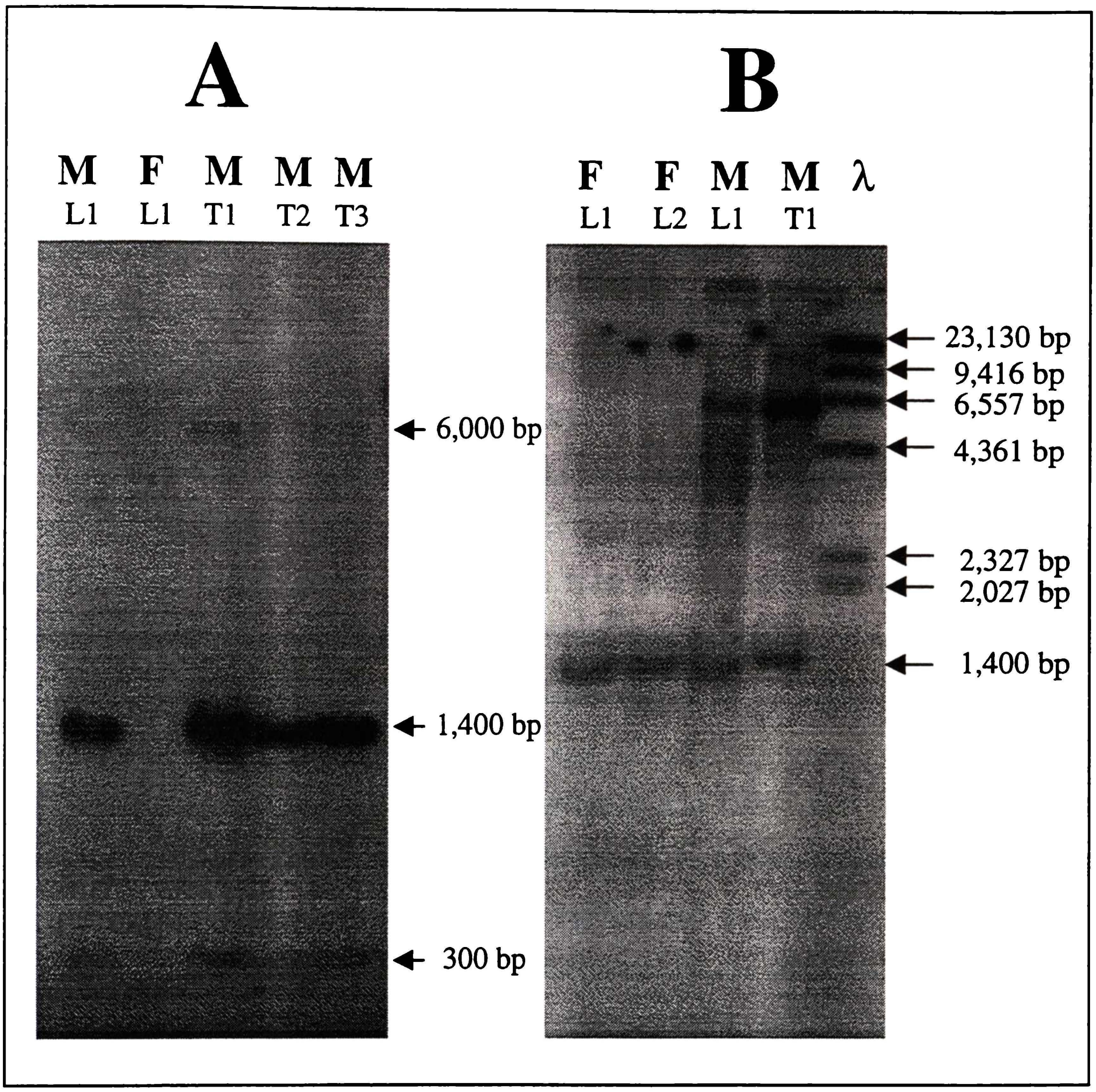


Figure 4.10: Southern Blot analyses of male and female bovine DNA. 1 μg of Testicle (T) and Liver (L) bovine male (M) genomic DNA, and Liver female bovine (F1, F2) genomic DNA was digested with *Eco*RI, electrophoresed and transferred to a Hybond-N^+ membrane under alkaline conditions. The membrane was hybridised with a $\alpha^{32}P$ dCTP labelled probe. A 1 kb DNA ladder (New England Biolabs, UK), and λ DNA/*Hind*III radioactive fragments were used as molecular markers. **A)** Hybridisation with Bov*REP3* PCR probe at 60 °C. This probe binds to three male specific DNA fragments; approximately 0.3, 1.4, and 6 Kb in size. **B)** Hybridisation with Bov*TSPY* probe at 65°C. This probe binds to a 6 Kb male DNA fragment.

Figure 4.10A shows that Bov*REP*3 probe binds to male gDNA only, with a major band at 1.4 Kb and a minor band at 300 bp. The 1.4 Kb band represents a bovine satellite I DNA whereas the 300 bp represents the repetitive bovine fragment. Hence this sequence is specific to male DNA. In the plasmid control, the Bov*REP3* probe also bound to undigested and digested Plasmid-Bov*REP3* (results not shown). Figure 4.10B shows that the Bov*TSPY* probe binds to male and female DNA, however there is a male specific band at 6 Kb.

The Bov*SRY* probe binds to both male and female DNA, with a major band at 1.4 Kb (bovine satellite I band) and a 2.4 Kb band (SRY gene). Therefore this suggests that this probe is not sequence specific (results not shown).

The question of how specific and repetitive the small DNA triplex region (10-16 nt) was to male gDNA as opposed to female gDNA was also investigated by Southern blot analysis. Small oligonucleotide probes ranging from 16 to 49 nt in size were 5' end-labelled with [γ^{32}P]ATP and hybridised overnight to a membrane containing DNA molecular weight markers, male and female bovine *Eco*RI digested gDNA and plasmid DNA containing a single copy of the appropriate DNA triplex site. The small oligonucleotide probes contained at least 50% guanine; however, significant hairpin or dimer structures that would interfere with hybridisation were not detected using the computer software program *OLIGO4*.

Hybridisation was performed at room temperature and followed by a low stringency wash (6X SSC). The resulting autoradiographs showed that all probes tested bound non-specifically to the molecular weight ladders and also to the major male and female *Eco*RI digested bovine gDNA bands present in the ethidium bromide stained agarose gel including the *Eco*RI 1.4 Kb satellite band (results not shown). The oligonucleotide probe hybridised 100 times more strongly to the bovine digested genomic DNA than the plasmid DNA (contains the DNA triplex site). The membranes were washed further with 6X SSC to reduce the high background and also then exposed for a longer period. This also failed to detect any male specific bands. High stringency washes were also investigated but this tended to remove most of the probe.

Hybridisation was also performed at higher temperatures (37°C) but the results were the same for hybridising at room temperature. Thus, autoradiograph analysis failed to detect any high molecular weight male specific bands as seen earlier in Figures 3.8 and 3.9. Hence these results suggest that the DNA triplex region of the chosen bovine sequences are not male specific and repetitive as previously reported by Reed *et al.*, (1995) and Grobet *et al.*, (1996).

An alternative strategy to determine male specificity and the repetitive nature of the Bov*REED* DNA sequence was also investigated. The Bov*REED* PCR product was digested with restriction enzymes, generating smaller fragments that could be isolated from a LMP agarose gel and radioactively labelled. This approach was hindered by the fact that the restriction enzyme sites available within the Bov*REED* PCR product were either rare or cut DNA fragments that were approximately the same size, and therefore could not be isolated by agarose electrophoresis.

4.4 DNA TRIPLEX BINDING ASSAYS WITH LARGE TARGETS

In order to investigate large DNA targets, the PCR products containing the specific DNA triplex region was cloned into the Bluescript II SK^+ vector. The DNA was sequenced to verify the correct clone and reconfirm the DNA triplex sequence as previously described in Section 4.2. Once verified, DNA triplex crosslinking assays were performed, initially with psoralen-conjugated TFOs which were known to form monoadducts in the synthetic duplex and PCR triplex assays.

4.4.1 *sup*F DNA Triplex Experiments

Initially, two *sup*F experiments, those of Wang *et al.*, (1995) and Gunther *et al.*, (1996) were repeated so that a small and large DNA target could be compared. As well as the size of the dsDNA target (38 and 200 bp), the main differences between the two published experiments was the pH of the Tris buffer (7.1 and 8.0) and the use of spermidine (Wang *et al.*, 1995) or sucrose (Gunther *et al.*, 1996). Duplicating these conditions, 75% photoadduct formation (presumably a monoadduct) was observed after 5 minutes of UV light (0.5 J/cm^2) exposure when PsoAG10 was targeted to the γ^{33}P-labelled *sup*F DNA (Figure 4.11A, B). Only one photoadduct was detected when both the forward and reverse primer were end-labelled with [γ^{33}P]ATP at the 5' end which would support the theory that they were both monoadducts. EMSA analysis showed they both migrated at the same distance (results not shown).

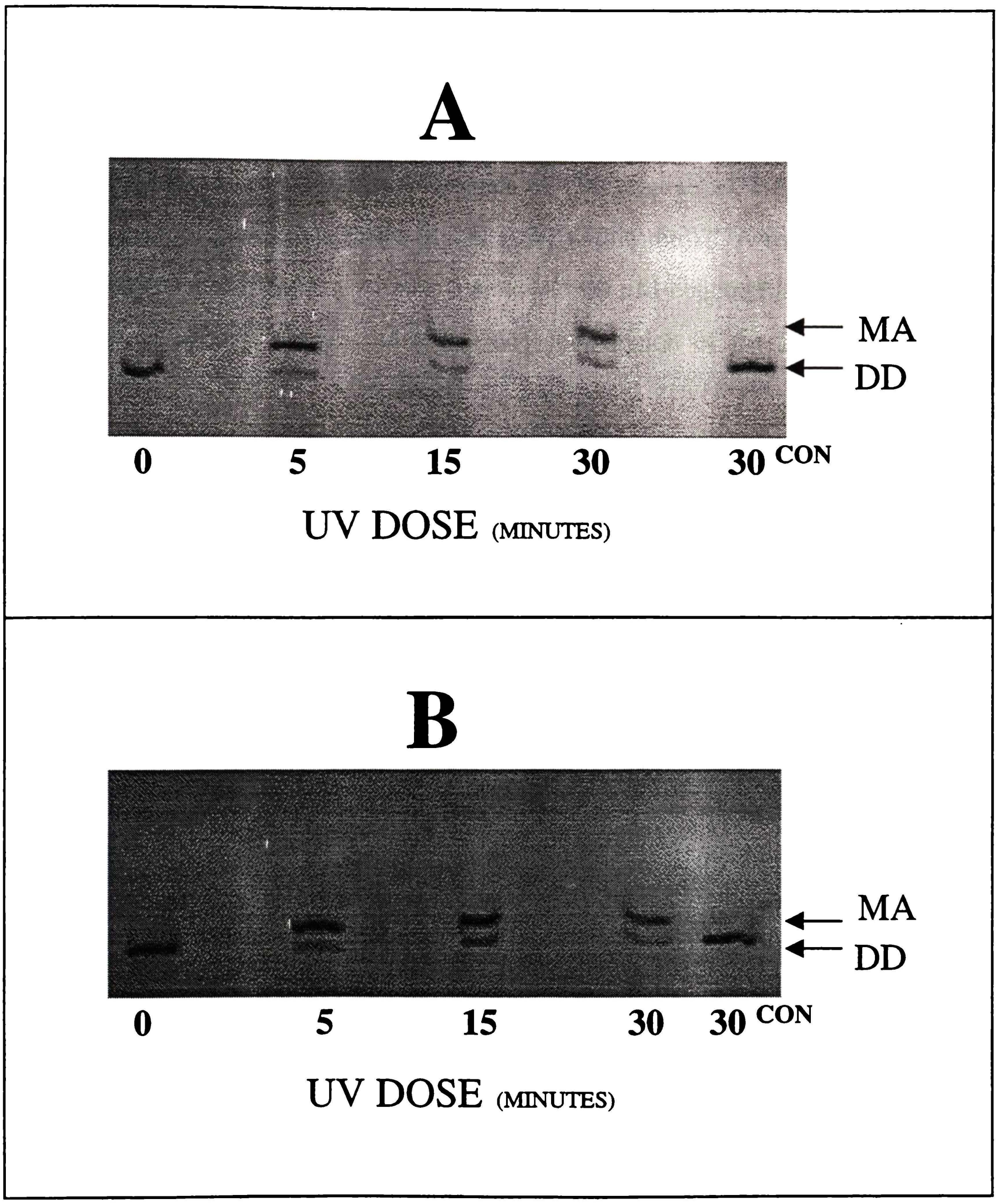


Figure 4.11: *In vitro* targeting of PsoAG10 to amplified dsDNA *sup*F target. **A)** PCR Triplex assay using conditions similar to Gunther *et al.*, 1996. The reverse primer was 5' end-labelled with [γ^{33}P]ATP. Following PCR, the 182 bp radioactive amplified DNA was incubated with 10% sucrose, 20 mM $MgCl_2$, 10 mM Tris (pH 8.0), and 80 µM PsoAG10. The samples were incubated for 2 hr, 37°C, and then irradiated on ice for 0-30 minutes (maximum UV dose of 2 J/cm^2), 3 cm away from the 360 nm UV lamp. PsoAG10 was absent in the control sample (30^{CON}) and was subjected to 30 minutes of 360 nm UV light. The samples were electrophoresed on a 6% TBE denaturing PAGE gel. **B)** PCR Triplex assay using conditions similar to Wang *et al.*, 1995. The reverse primer was 5' end-labelled with [γ^{33}P]ATP. Following PCR, the 182 bp radioactive amplified DNA was incubated with 20 mM $MgCl_2$, 1 mM spermidine, 10 mM Tris (pH 7.4), and 80 µM PsoAG10. The exact conditions were then followed as stated in Figure 4.11A. MA = monoadduct, DD = amplified dsDNA.

Figure 4.12 shows that *in vitro* targeting of PsoAG10 to the α^{32}P-labelled *supF* DNA fragment containing the *supF* DNA triplex region also results in photoadduct formation. DNA triplex formation increased when either the concentration of psoralen-conjugated TFO (0.4-80 μM) or the UV dose (0.5-32 minutes, 0.03-2 J/cm^2) was increased. After thirty seconds of UV irradiation, 0.4 μM of PsoAG10 was sufficient to generate a DNA triplex (2%) within the DNA fragment. After 15 minutes of UV irradiation (1 J/cm^2), 0.4, 4, and 40 μM concentrations of PsoAG10 generated, respectively, 15%, 44% and 64% conversion of the DNA duplex to a triplex.

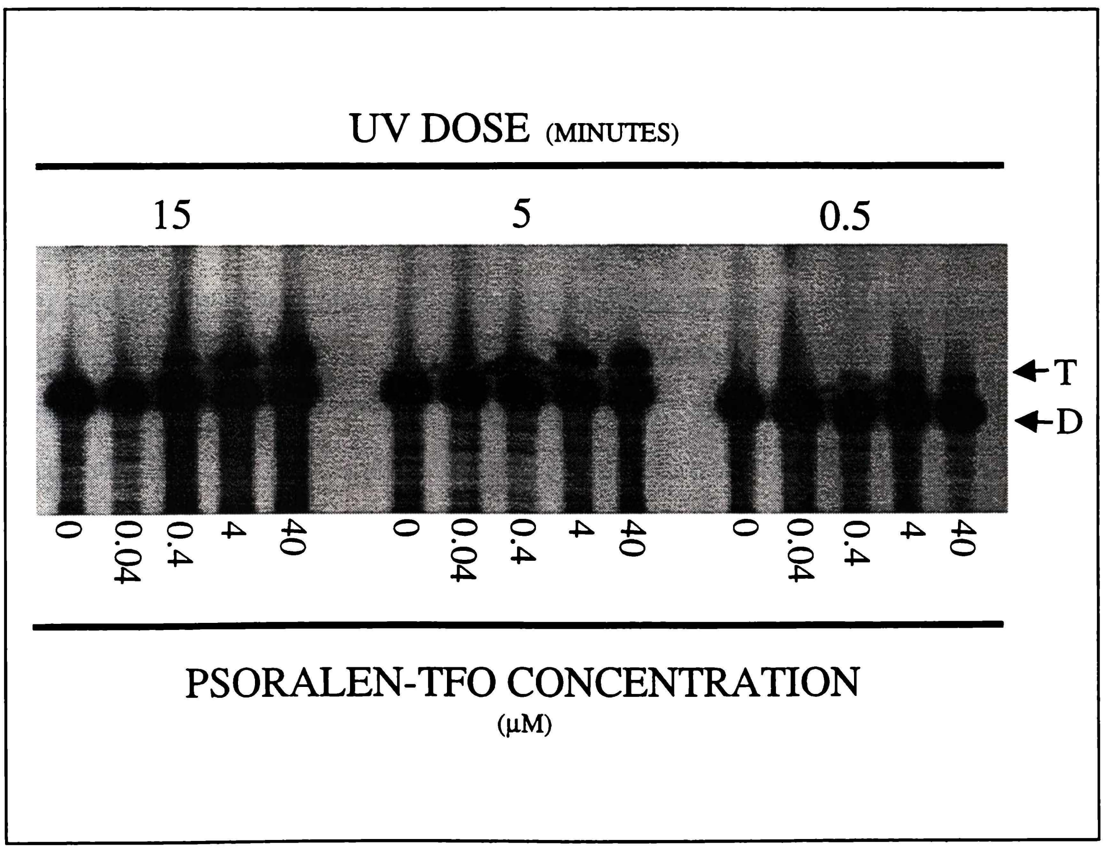


Figure 4.12: Photoinduced crosslinking of a 10 nt purine rich psoralen-conjugated TFO to a 200 bp amplified α^{32}P *supF* dsDNA target. The α^{32}P end-labelled duplex was incubated with 0-40 μM PsoAG10, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl pH 7.1), 1 mM spermidine for 2hr, 37°C, and then irradiated with UV light (360 nm) on ice for 0.5-15 minutes, 3 cm away from the UV lamp (maximum UV dose of 1 J/cm^2). The samples were then electrophoresed on a 8% PAGE gel. D = DNA duplex, T = DNA triplex.

Potassium ions were present in the PCR buffer and these ions have been previously shown to inhibit DNA triplex formation (Cheng & Van Dyke, 1993). In this study, the presence of 100 mM KCl in the DNA triplex binding assay inhibited photoadduct formation by 20% when targeting PsoAG10 to the 24 bp *sup*F duplex *in vitro* (results not shown). Therefore a variety of methods were employed to remove K^+ ions following the PCR reaction with the aim of increasing the formation and stability of PCR photoadducts. These included using a Promega mini Wizard Prep, G25 Sephadex columns, ethanol precipitation and phenol extractions. No significant difference was detected when using this purified amplified DNA in the UV crosslinking assays (results not shown).

4.4.2 Bov*REP3* DNA Triplex Experiments

Following these initial *sup*F amplified DNA triplex experiments, psoralen-conjugated TFOs were tested with the larger bovine dsDNA fragments *in vitro*. Forward and reverse primers (22 nt) were used to amplify a 100 bp fragment from the nucleotide region 239-248 of the published bovine repetitive DNA sequence (Figure 4.1). The DNA triplex site is from nucleotides 239 to 248 as shown in Appendix III.

Figure 4.13 shows that *in vitro* targeting of 40 μM of Puv6a to a γ^{33}P radioactive 100 bp PCR fragment that contained the 10 nt DNA triplex site, formed a detectable monoadduct (50%) after fifteen minutes of UVA irradiation (360 nm, 1 J/cm^2) when the sense strand of the PCR product was radioactively labelled. Monoadduct formation was also detected with Puv6c to a lesser degree (10%). Puv6b failed to form any photoadducts, presumably due to the pyrimidine mismatches within the homopurine region. PsoAG10 failed to form any photoadducts, which shows that the psoralen-conjugated TFOs target and crosslink to specific sequences. Puv5 also failed to form any photoadducts, as it could not bind parallel to the amplified dsDNA.

Under the conditions of this assay, monoadduct formation was not detected in the presence of Puv6a concentrations of less than 8 μM (results not shown). No photoadducts were detected when the antisense strand of the PCR product was radioactively labelled, targeted with Puv6a, Puv6b, Puv6c and subjected to thirty minutes of UV irradiation (2 J/cm^2).

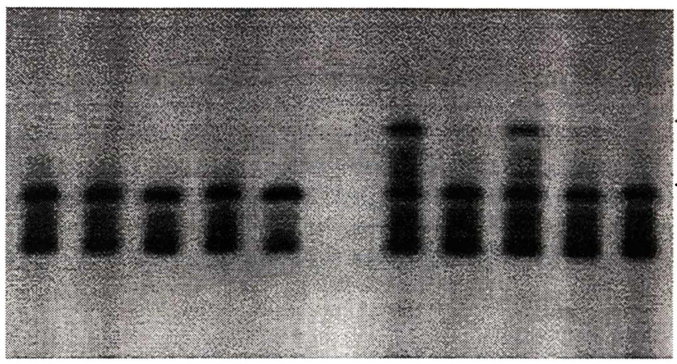


Figure 4.13: *In vitro* targeting of psoralen-conjugated TFOs to a ^{33}P radioactive 100 bp PCR fragment that contains the Bov*REP3* DNA triplex region. The reverse (A) or forward (B) primer were 5' end-labelled with [γ^{33}P]ATP and then PCR was performed using one radioactive primer per reaction as stated in Chapter 2. Following PCR, the amplified DNA was incubated for 2 hr, 37°C, with 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) and 40 μM of a specific psoralen-conjugated TFO; Puv6a, Puv6b, Puv6c, PsoAG10 or absence of a psoralen-conjugated TFO (control). The samples were irradiated with UV light (360 nm) on ice for 15 minutes, 3 cm away from the UV lamp (maximum UV dose-1 J/cm^2), and then electrophoresed on a 6% TBE denaturing PAGE gel.
MA = monoadduct, DD = amplified dsDNA.

Figure 4.14 shows that a detectable amount of monoadduct (1%) was formed when Puv6a was targeted to a ^{32}P radioactive 100 bp PCR fragment *in vitro* after 5 minutes of UVA irradiation (0.5 J/cm^2), with some 50% stronger monoadduct formation after 30 minutes (2 J/cm^2). No crosslinks were detected even after the high UV dose. Under the conditions of this assay, no monoadduct formation was detected in the presence of Puv6a at concentrations of less than 0.4 μM, even after 30 minutes of UV irradiation (2 J/cm^2).

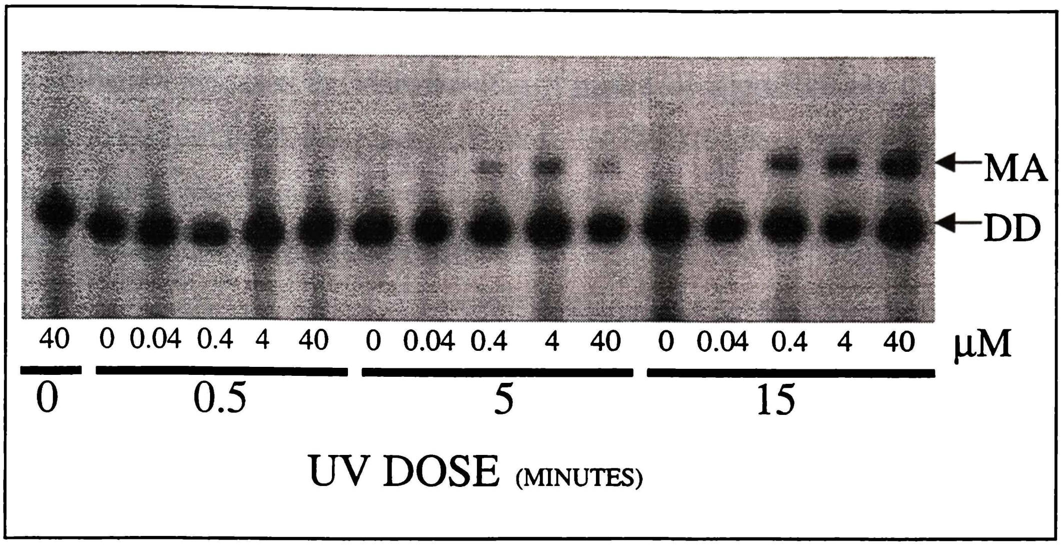


Figure 4.14: *In vitro* targeting of Puv6a to a ^{32}P radioactive 100 bp PCR fragment that contains the Bov*REP3* DNA triplex region. The [$\alpha^{32}P$]dCTP was incorporated into the amplified DNA during the PCR synthesis step. The amplified $\alpha^{32}P$ Bov*REP3* DNA was incubated with Puv6a (0-40 μM), 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated with UV light (360 nm) on ice for 0-15 minutes, 3 cm away from the UV lamp (maximum UV dose-2 J/cm^2). The samples were then electrophoresed on a 4% TBE denaturing PAGE gel. MA = monoadduct, DD =amplified dsDNA.

4.4.3 Plasmid DNA Triplex Experiments

The 100 bp amplified DNA containing the male bovine DNA triplex site was cloned into Bluescript II SK$^+$ vector generating a 3061 bp plasmid. The DNA was sequenced to verify the insert identify and reconfirm the DNA sequence (Figure 4.5). Once verified, DNA triplex crosslinking assays were performed, initially with Puv6a and then with other specific psoralen-conjugated TFOs.

In general, the following method was used for the plasmid UV crosslinking studies (Figure 2.5). The plasmid containing the male bovine DNA triplex site was digested with the restriction enzyme *Hind*III, 5' end-labelled with a specific [$\alpha^{32}P$]dNTP, incubated with a specific psoralen-conjugated TFO and Triplex binding buffer (contains a higher concentration of magnesium to enhance DNA triplex stability), irradiated with 360 nm UV light (2 J/cm^2), digested with the restriction enzyme *Eco*RI to release the cloned PCR insert and then electrophoresed on a 8% denaturing PAGE gel.

Figure 4.15 shows that non-irradiated EMSA samples migrated as two bands. The lower band represents the released cloned insert, of approximately 130 bp in size, while the upper band the plasmid vector. A DNA triplex, presumably a cytosine monoadduct, was formed after ten and thirty minutes of UV irradiation with 33% and 77% monoadduct formation respectively.

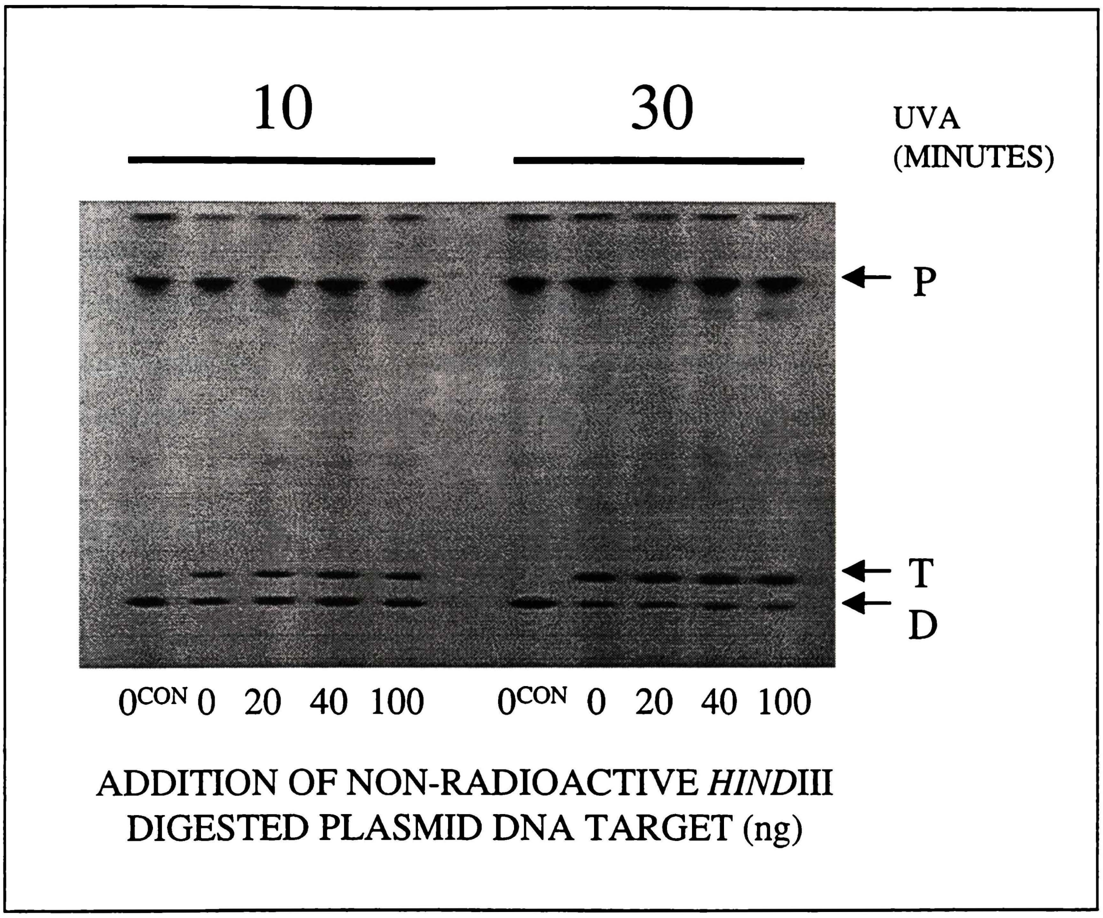


Figure 4.15: *In vitro* targeting of psoralen-conjugated TFO to radioactive and non-radioactive plasmid-Bov*REP3* DNA. $\alpha^{32}P$ end-labelled *Hind*III digested plasmid containing the Bov*REP3* triplex region was incubated with increasing amounts of non-radioactive *Hind*III digested plasmid-Bov*REP3* (0, 20, 40, 100 ng), 1 mM spermidine, 1X DNA triplex buffer (100 mM $MgCl_2$, 40 mM Tris-HCl, pH 8.0), 1 μM Puv6a for 2 hr, 37°C. The samples were irradiated with UV light (360 nm) on ice for 10 or 30 minutes, 3 cm distance from the UV lamp (maximum UV dose-2 J/cm^2), and then digested with *Eco*RI to release the DNA insert (D) and DNA triplex (T). The samples were electrophoresed on a 4% denaturing PAGE gel. P represents *Hind*III/*Eco*RI plasmid vector.

An important aspect of this investigation is competition type experiments in which one would predict that DNA triplex formation in the radioactive target would decrease as increasing amounts of non-radioactive target DNA were added to the assay. However, Figure 4.15 shows that no such competition was detected, even when up to 100 ng of competing DNA was added. Although disappointing, this result is perhaps predictable. Namely, the large molar amounts of Puv6a (1000 nM) that had to be used in the assay for photoadduct detection. The equivalent molar concentration of radioactive target in the assay was 20 nM and the highest concentration of competing DNA was 500 nM. Hence, there was a vast excess of competing non-radioactive competing DNA, but this was still much lower than the molar concentration of triplex. The psoralen concentration (1-100 µM), dsDNA concentration (0-500 nM), and UV dose (0.03-4.0 J/cm^2) were adjusted, but no significant difference was achieved in terms of DNA triplex competition between the psoralen-conjugated TFO and plasmid DNA. The reasons for the high "threshold" levels of DNA triplex required to initiate the reaction are not clear.

Competitive DNA triplex assays were performed to determine whether or not Puv6a DNA triplex formation would be preferentially competed out by male bovine genomic DNA which contains highly repetitive Y-specific gene sequences as opposed to female genomic DNA. One would also expect to see a decrease in the photoadduct formation between the plasmid DNA and Puv6a in the presence of male, but not female, bovine genomic DNA.

The plasmid containing the male bovine DNA triplex site was digested with the restriction enzyme *Hind*III, end-filled with a specific [α^{32}P]dNTP, incubated with Puv6a (1-80 µM), bovine male or female genomic DNA (5-500 ng), and Triplex binding buffer in a total volume of 10 µL, irradiated with 360 nm UV light (2 J/cm^2), digested with the restriction enzyme *Eco*RI to release the cloned PCR insert, and then electrophoresed on a 4% denaturing PAGE gel.

Relatively high concentrations of genomic DNA failed to suppress triplex formation in the plasmid DNA. Figure 4.16 shows that 75% of the plasmid DNA was triplexed, presumably to a monoadduct, when Puv6a was targeted to plasmid-Bov*REP*3 and subjected to 30 minutes of UV irradiation (2 J/cm^2). The intensity of the monoadduct band remained constant as the amount of male and female genomic DNA increased. Therefore, under the conditions of this assay no competition was detected. The psoralen and DNA concentration, and UV dose were adjusted but no significant difference was achieved in terms of DNA triplex competition between the plasmid and genomic DNA, even though there are believed to be hundreds of copies of Y-chromosomal repeats in the 500 ng of male DNA.

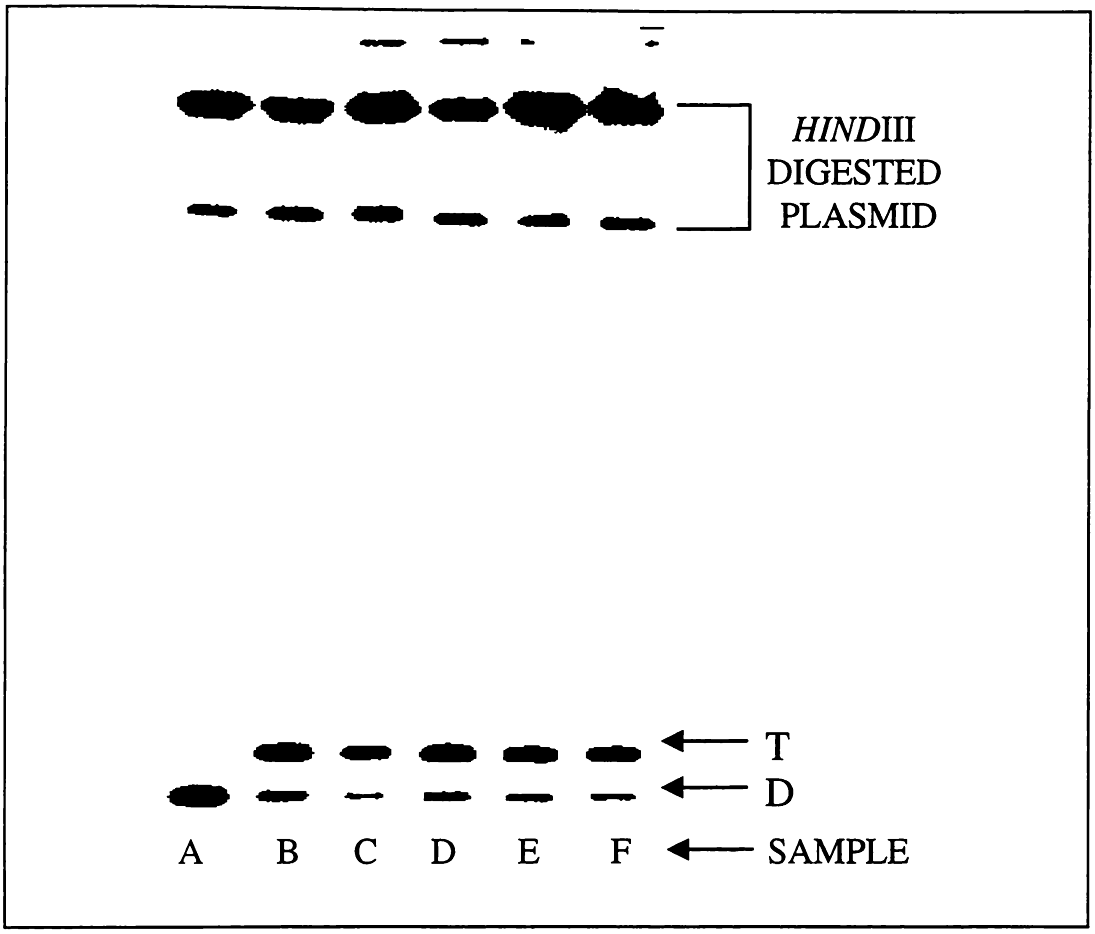


Figure 4.16: Competitive DNA triplex assay between plasmid-Bov*REP3* DNA, bovine male or female genomic DNA, and Puv6a. $\alpha^{32}P$ end-labelled *Hind*III digested plasmid that contains the Bov*REP3* triplex region was incubated with 80 μM Puv6a, 1 mM spermidine, 1X DNA triplex buffer (100 mM $MgCl_2$, 40 mM Tris-HCl, pH 8.0) for 2 hr, 37°C, and then irradiated with UV light (360 nm) on ice for thirty minutes, 3 cm distance from the UV lamp (maximum UV dose of 2 J/cm^2). The cloned insert containing the DNA triplex region was released by *Eco*RI digestion. Samples A-F were then electrophoresed on a 4% TBE denaturing PAGE gel. Table 3.1 lists the constituents of samples A-F in the DNA triplex binding assay.
D = DNA duplex, T = DNA triplex.

Sample	Plasmid-Bov*REP3* ^{32}P *Hind*III digested	Triplex buffer, pH 8.0	Puv12a	Male bovine gDNA	Female bovine gDNA
A	50 nM	1X	—	—	—
B	50 nM	1X	80 μM	—	—
C	50 nM	1X	80 μM	500 ng/μL	—
D	50 nM	1X	80 μM	50 ng/μL	—
E	50 nM	1X	80 μM	5 ng/μL	—
F	50 nM	1X	80 μM	—	500 ng/μL

Table 4.1: Final concentration of samples A-F which were used in the competitive DNA triplex binding assay. The Triplex buffer contains 100 $MgCl_2$ and 40 mM Tris-HCl, pH 8.0.

4.4.4 Bov*SRY* DNA Triplex Experiments

Forward and reverse primers (25 nt) were constructed to amplify a 200 bp fragment from region 926-1125 of the published *Bos taurus* SRY gene sequence (Figure 4.1). The DNA triplex site is from nucleotides 1032-1043 as shown in Appendix III. Figure 4.17 shows that when the ^{33}P label was used, Puv12a forms a detectable monoadduct after fifteen minutes of UVA irradiation (360 nm, 1 J/cm^2). A monoadduct was observed when the sense and antisense strand of the PCR product were radioactively labelled at the 5' ends with γ-^{33}P. Puv12b, Puv12c, and Puv12d failed to form detectable levels of photoadducts. Under the conditions of this assay, it was not possible to detect monoadducts with Puv12a at concentrations of less than 10 µM.

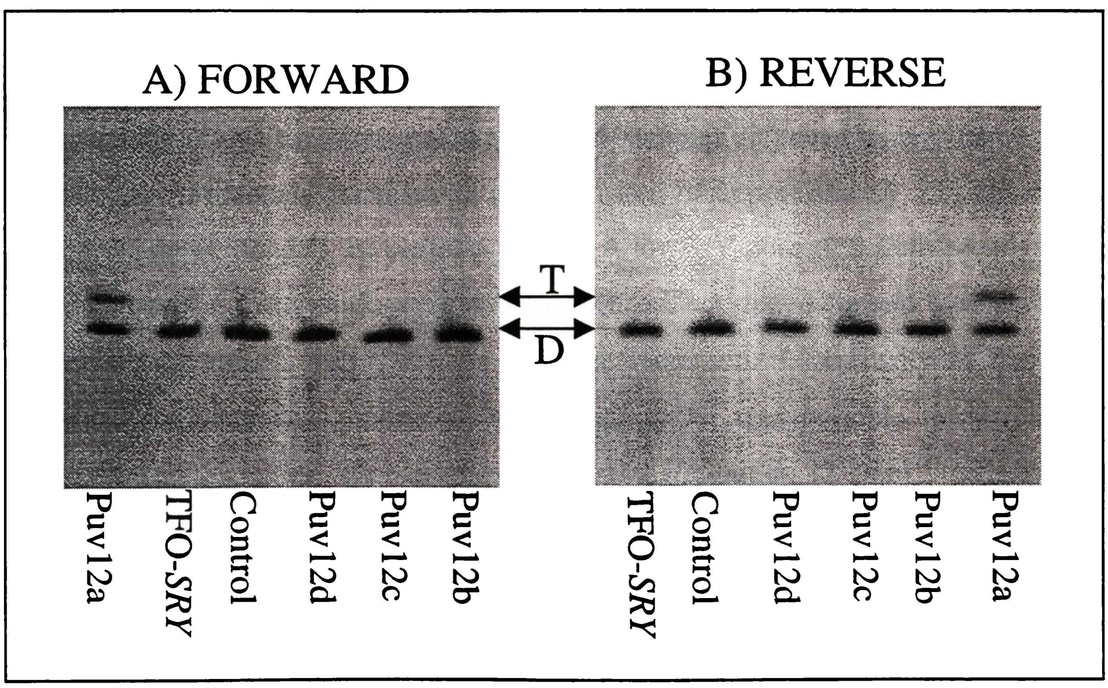


Figure 4.17: *In vitro* targeting of psoralen-conjugated TFOs to $\gamma^{33}P$ radioactive amplified DNA that contains the Bov*SRY* DNA triplex region. The forward (A) and reverse (B) primer were 5' end-labelled with [$\gamma^{33}P$]ATP, and then PCR was performed using one radioactive primer in each reaction. Following PCR, the 200 bp radioactive amplified DNA was incubated with 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1), 40 µM of a specific psoralen-conjugated TFO; Puv12a, Puv12b, Puv12c, Puv12d, no psoralen-conjugated TFOs (control) or absence of psoralen at the 5' end (TFO-*SRY*). The samples were then irradiated with UV light (360 nm) on ice for 15 minutes, 3 cm distance from the UV lamp (maximum UV dose of 1 J/cm^2), and then electrophoresed on a 6% TBE denaturing PAGE gel.
T = DNA triplex, D = amplified dsDNA.

Figure 4.17 also shows that psoralen is necessary for the generation of photoadducts as TFO-*SRY* (identical DNA sequence to Puv12a but no psoralen moiety is on the 5' end) only gave a single duplex band.

The 200 bp amplified DNA containing the male bovine DNA triplex site was cloned into Bluescript II SK^+ vector generating a 3161 bp plasmid. The DNA was sequenced to verify the insert identity and reconfirm the DNA triplex sequence (Figure 4.7). Once verified, DNA triplex crosslinking assays were performed, initially with Puv12a and then with other specific psoralen-conjugated TFOs. The plasmid containing the male bovine DNA triplex site was digested with the restriction enzyme *Hind*III, 5' end-labelled with a specific [α^{32}P]dNTP, incubated with Puv12a and triplex binding buffer (contains a higher concentration of magnesium to stabilise the DNA triplex), irradiated with 360 nm UV light, digested with the restriction enzyme *Eco*RI or *Xba*I to release the cloned PCR insert and then electrophoresed on a 8% denaturing PAGE gel.

Figure 4.18 shows that non-irradiated EMSA samples migrated as two bands. The lower band represents the released cloned insert, of approximately 250 bp in size, while the upper band is the plasmid vector. A low UV dose (0.5 J/cm^2) leads to the appearance of a third band, presumably a monoadduct that is formed over a Puv12a concentration range 1-80 μM. As the dose increases, the intensity of this band increases. Under the conditions of these assays it was possible to detect 50% monoadduct formation after 30 minutes of UV irradiation (2 J/cm^2) with Puv12a concentrations of at least 1 μM (results not shown).

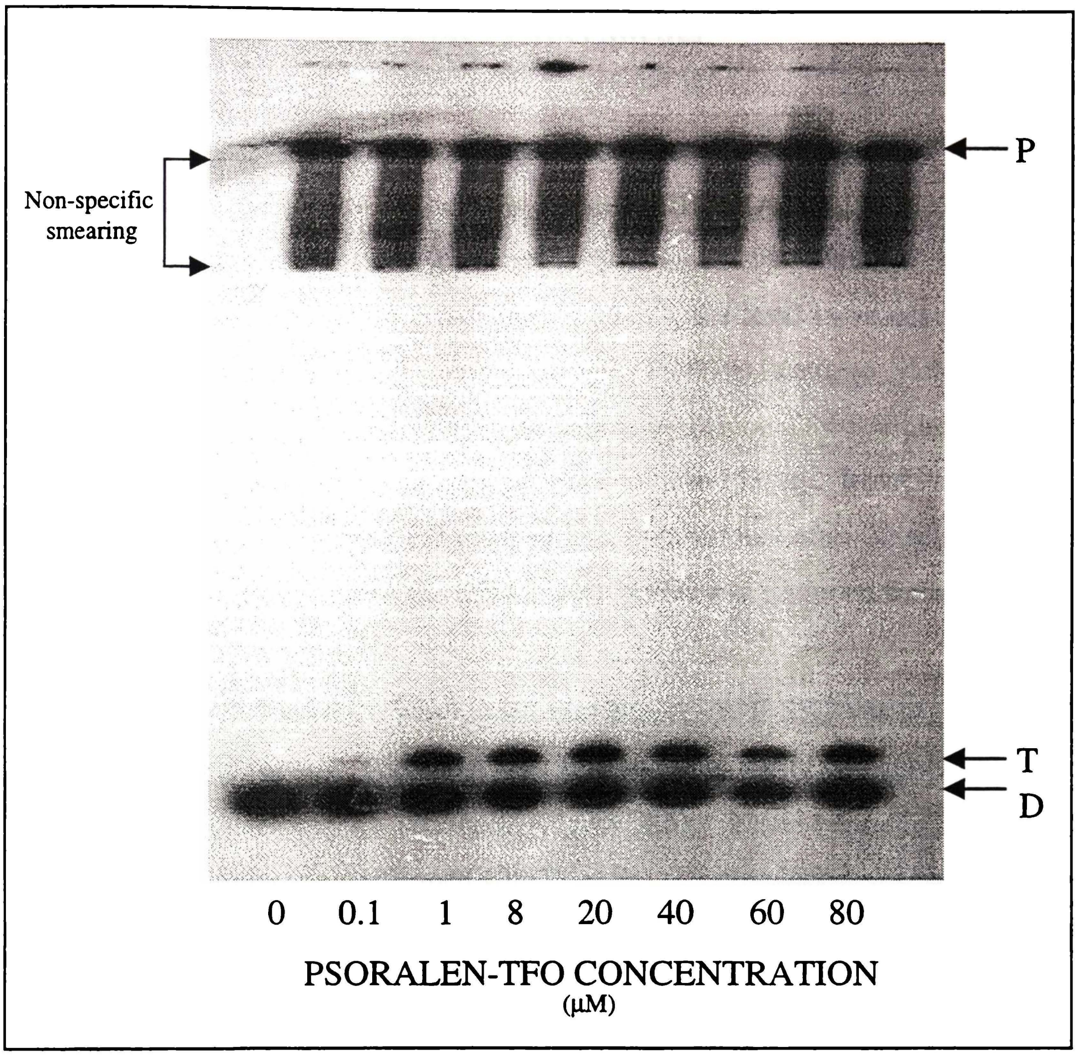


Figure 4.18: Effect of psoralen-conjugated TFO concentration on plasmid-Bov*SRY* DNA triplex formation. $\alpha^{32}P$ *Hind*III digested plasmid-Bov*SRY* (contains bovine SRY triplex region) was incubated with 0-80 μM Puv12a, 1 mM spermidine, 1X DNA triplex buffer (100 mM $MgCl_2$, 40 mM Tris-HCl, pH 8.0) for 2 hr, 37°C, and then irradiated with UV light (360 nm) on ice for 5 minutes, 3 cm distance from the UV lamp (maximum UV dose-0.5 J/cm^2). The DNA triplex region was released by further digestion with *Xba*I for 2 hr at 37°C. The samples were electrophoresed on a 4% TBE denaturing PAGE gel. P = *Hind*III digested plasmid-BovSRY vector, T = DNA triplex, D = DNA insert (250 bp). The non-specific smearing was detected due to the increased sensitivity of the end-labelling reaction with ^{32}P and long exposure of the X-ray film.

Competitive DNA triplex assays were also performed by targeting Puv12a to radioactive and non-radioactive plasmid DNA and also to male and female bovine gDNA (results not shown). No competition was detected in either case due to the high molar ratio of TFO with respect to the dsDNA target and also, the poor sensitivity of detecting photoadduct formation with low concentrations of Puv12a (Refer to Section 4.4.3). The psoralen concentration, dsDNA concentration and UV dose were adjusted but no significant difference was achieved in terms of DNA triplex competition between the plasmid and genomic DNA.

4.4.5 Bov*REED* DNA Triplex Experiments

Forward and reverse primers (22 nt) were constructed to amplify a 500 bp fragment from the BRY11a sequence as shown in Appendix III. Figure 4.19A shows that after 2 minutes of UVA irradiation (0.1 J/cm^2), 5% monoadduct formation was detected, and after 30 minutes of UVA irradiation (2 J/cm^2), the level increases to 10%. No crosslinks were detected. Experiments with Puv15b failed to yield detectable photoadducts. Figure 4.19B shows the effect of pH on DNA triplex formation. DNA triplex binding assays in the presence of Puv15a and Bov*REED* synthetic dsDNA target at pH 7.1 and 8.0 resulted in a monoadduct after 32 minutes of UV irradiation (2 J/cm^2) but not if the DNA triplex buffer was excluded from the assay.

The 400 bp amplified DNA was cloned into Bluescript II SK^+ vector, generating a 3461 bp plasmid vector. The DNA was sequenced to verify the correct clone and reconfirm the DNA triplex sequence (Figure 4.6). Once verified, DNA triplex crosslinking assays were performed with Puv15a. UV crosslinking experiments were repeated with the same method as previously described in Section 4.4. Difficulty was experienced in detecting DNA triplex formation due to inefficient plasmid digestion by restriction enzymes, lack of common restriction enzyme sites and poor separation of large DNA fragments on the PAGE gel (data not shown).

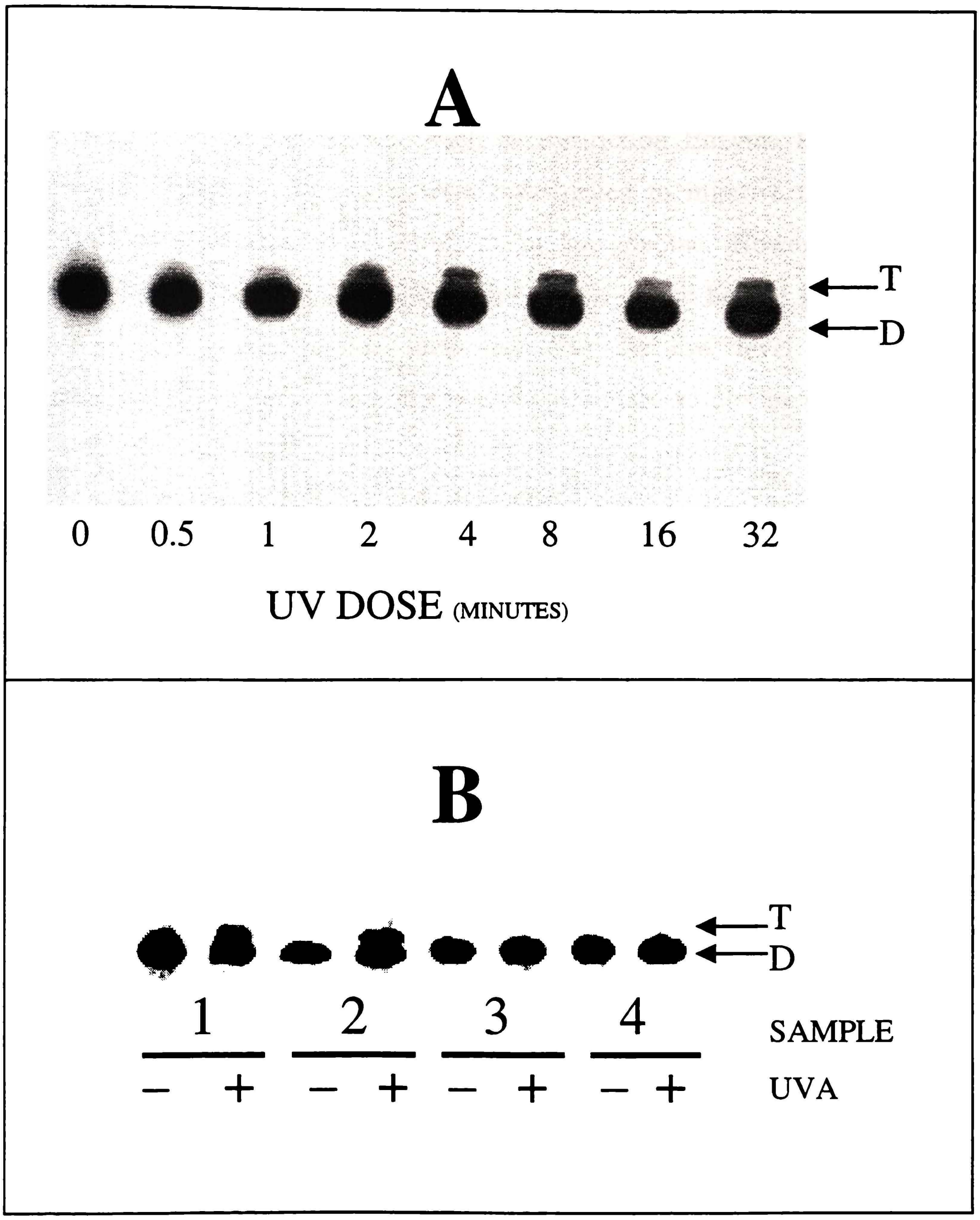


Figure 4.19: *In vitro* targeting of Puv15a to a 500 bp PCR product that contains the Bov*REED* triplex region. **A)** Effect of UV dose on DNA triplex formation. $\alpha^{32}P$ amplified dsDNA containing Bov*REED* triplex region was incubated with 40 μM Puv15a, 1 mM spermidine, 1X DNA triplex buffer (10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.1) for 2 hr, 37°C, and then irradiated on ice for 0-32 minutes, 3 cm distance from the UV lamp (360 nm, maximum UV dose of 2 J/cm^2). The samples were then electrophoresed on a 4% TBE denaturing PAGE gel. **B)** Effect of pH on DNA triplex formation. $\alpha^{32}P$ amplified DNA containing Bov*REED* triplex region was incubated with 1 mM spermidine, 10 mM $MgCl_2$, 10 mM Tris-HCl, 40 μM Puv15a for 2 hr, 37°C, under pH 7.1 (Sample 1), pH 8.0 (Sample 2) and then irradiated on ice for 32 minutes (2 J/cm^2). Sample 3 was incubated in the absence of DNA triplex buffer, pH 7.0 and Sample 4 had Puv15a absent, pH 7.1. "–" represents no UV exposure and "+" represents thirty minutes of UV irradiation (2 J/cm^2), 3 cm distance from the UV lamp (360 nm). The samples were then electrophoresed on a 4% TBE denaturing PAGE gel. T = DNA triplex, D = DNA duplex.

4.5 ALTERNATIVE METHODS FOR DETECTION OF COVALENT DNA TRIPLEXES

A variety of alternative methods, such as Restriction Enzyme Protection Assay (REPA) and DNase Footprinting, were investigated to improve the efficiency of detecting DNA triplex formation within the bovine genome. However, this research was unsuccessful due to the large complex nature of the bovine genome, the low sensitivity of detecting photoadducts and also the low level of photoadduct formation. For the record, each of these methods and the results are briefly described below.

4.5.1 Restriction Enzyme Protection Assay

A restriction enzyme protection assay (REPA) involves targeting a TFO to a target DNA site that contains a restriction enzyme site. Formation of a DNA triplex at this restriction enzyme site prevents the restriction enzyme from cleaving its specific DNA target, resulting in a slower migrating band on the agarose gel (Maher *et al.*, 1989). This common method of analysing DNA triplexes has been successfully employed by many researchers to detect DNA triplex formation such as the *sup*F gene. However the application of REPA to this research was not possible due to the absence of restriction enzyme sites within the screened bovine DNA triplex sequences.

4.5.2 DNase Footprinting

DNase (Deoxyribonuclease) footprinting experiments have been previously used to locate the positions where proteins and TFOs bind to DNA targets. This method is based on the theory that the TFO binds to its dsDNA target and protects the DNA duplex from DNase digestion. PAGE analysis shows a staggered banding pattern and the DNA triplex is located in the region where there are no bands ("a footprint"). This type of analysis is not easy to use in UV crosslinking studies as DNase cuts the DNA minor groove while the TFO binds to the major groove (Chandler & Fox, 1996).

Initial experiments were carried out to target Puv6a and Puv12a to their $\gamma^{33}P$ amplified dsDNA targets, forming stable photoadducts following UV irradiation. The triplex samples were then subjected to DNase I digestion and analysed by denaturing PAGE electrophoresis. The results obtained were not clear and it was extremely difficult to confirm footprints. This was probably due to the intercalating properties of psoralen. Most researchers use unmodified TFOs for DNase Footprinting experiments. Therefore this form of detection was not investigated any further.

4.5.3 3' End-labelling of Psoralen-conjugated TFOs

An alternative approach to detect photoadduct formation was to target 5' psoralen-conjugated TFOs that were end-labelled at the 3' end with terminal transferase and [α^{32}P]dNTP. Wang *et al.*, (1995) used a similar protocol to target 3' end-labelled psoAG10 to the pSupFG1 vector using [α^{32}P] dideoxyATP. The selected dNTP for 3'labelling depended on the adjacent nucleotide at the 3' end of the triplex target. In the case of Puv6a and Puv12a, a thymine nucleotide followed the Bov*REP3* and Bov*SRY* duplex-triplex region at the 3' end. Therefore, 5' psoralen-conjugated TFOs were 3' end-labelled with [α^{32}P]dGTP allowing a G·TA or A·TA triplet to form at the 3' end.

A method devised by Deng and Wu (1983) was performed to attach one nucleotide rather than a chain at the 3' end. This was achieved by using a low concentration of terminal transferase and a short incubation. The 3' end-labelling of 5' psoralen-conjugated TFOs were confirmed by native PAGE electrophoresis (results not shown).

The 3' end-labelling of 5' psoralen-conjugated TFO could have been performed using a α^{32}P dideoxynucleotide as prepared by Wang *et al.*, (1995). This would ensure the addition of one nucleotide to the 3' end. However, only adenine dideoxynucleotide ([α^{32}P]ddATP) was commercially available as a radioactive product at the time; its use would result in a A·TA triplet at the 3' end of the triplex target when 3' end-labelling Puv12a and Puv6b.

Previous studies have shown that a triplex combination of G·TA would be more suitable for DNA triplex stability; therefore ^{32}PddATP was not experimentally tested. 3' end-labelling of Puv12a and Puv6a with [α^{32}P]dATP and [α^{32}P]dCTP was also investigated, with no significant improvement in photoadduct formation (results not shown).

EMSA analysis showed that targeting of [α^{32}P]dGTP 3' end-labelled Puv12a to a non-radioactive 38 bp synthetic dsDNA target resulted in some monoadduct formation. No crosslinks were detected. The same results were observed for 3' end-labelled Puv6a (results not shown). The sensitivity in both experiments was very poor in comparison with that achieved in the γ^{33}P UV crosslinking studies, suggesting that it is better to radioactively label the dsDNA target than the psoralen-conjugated TFO.

Further DNA triplex binding assay studies were performed to determine if 3' end-labelled 5' psoralen-conjugated TFOs could be targeted to bovine genomic DNA. No significant difference was detected when the radioactive psoralen-conjugated oligonucleotide was incubated in the presence of male or female genomic DNA (both digested and undigested). The presence of many inhibitory ions (such as K^+) due to unpurified genomic DNA.

Potassium ions were present in the terminal transferase 3' end-labelling reaction mix. Thus, to overcome any potential K^+ inhibitory effects, the potassium cacodylate in the end-labelling reaction was replaced with sodium cacodylate buffer. Native PAGE electrophoresis showed that Puv6a labelled more efficiently with sodium based buffers (results not shown). However, no significant improvement in photoadduct formation was detected.

4.5.4 Mixed Psoralen-conjugated TFOs

A "combination method" was investigated to determine the most efficient psoralen-conjugated TFO for photoadduct formation within the male bovine genome. Initially two psoralen-conjugated TFOs were designed, containing a mixture of purine bases (Table 4.1). These two 15-mer psoralen-conjugated TFOs, PsoA1 and PsoB1, statistically allowed at least 2^{11} combinations of the third strand DNA sequence. It was thought that it maybe possible to detect photoadduct formation within the bovine genome using a gradient of mixed purine psoralen-conjugated TFOs, ie. a 14 nt TFO followed by a 13 nt, 12 nt, 11 nt and so on. Hence the resulting EMSA banding pattern could be compared, indicating the most efficient TFO sequence that would generate a photoadduct.

PsoA1: 5' Pso≈RRRRRRRRRRRAGGG 3' or 5' Pso≈R^{11}AGGG 3'

PsoB1: 5' Pso≈RRRRRRRRRRRGGGG 3'or 5' Pso≈R^{11}GGGG 3'

Figure 4.20: DNA sequence of mixed purine psoralen-conjugated TFOs.
R represents guanine or adenine bases, Pso≈ represents psoralen.

Two experimental methods were used to target PsoA1 and PsoB1 to bovine genomic DNA. The first method utilised a modified Southern blot procedure where a membrane containing digested and undigested female and male bovine genomic DNA was probed with a 3' end-labelled [α^{32}P]dATP PsoA1 or PsoB1. The second method involved performing DNA triplex binding assays with genomic DNA and 3' end-labelled [α^{32}P]dATP PsoA1 or PsoB1. No photoadducts were detected using either of these experimental methods. The level of crosslinking may have been too low for detection using current procedures or alternatively internal secondary structures such as guanine quartets within the TFO could have formed, therefore decreasing DNA triplex formation. Hence the "combination method" was not used with other psoralen-conjugated TFO combinations.

4.5.5 Exonuclease III

A Target Protection Assay (TPA) using triplex molecules and exonuclease III was developed by CyGene Inc, USA 1998 to solve the problem of locating disease-related genes in DNA and RNA when only a few copies of the gene are present. This assay involves digesting genomic DNA with one or two restriction enzymes to isolate the target region, and targeting a specific TFO to form a DNA triplex. The DNA is then digested with exonuclease III to degrade dsDNA in the 3' to 5' direction. The presence of a triplex structure interferes with the digestion and prevents further nuclease action. The target region is then hybridised with a capture probe, recaptured using a magnetic bead (antibody against capture probe) and then the captured complex is visualised by targeting a fluorescent reporter probe.

The first three steps of the TPA were performed to determine the efficiency of photoadduct formation within PCR and plasmid DNA. Initially the DNA target was end-labelled with $\alpha^{32}P$ and incubated in the presence and absence of a specific psoralen-conjugated TFO. The DNA triplex binding sample was subject to UV irradiation (360 nm) and then digested with exonuclease III to degrade the DNA in the 3' to 5' direction. The sample was then analysed by EMSA. One would expect to observe a difference between the DNA duplex and DNA triplex. However, no significant difference was detected between UV irradiated and non-UV samples, in the presence or absence of psoralen-conjugated TFOs. Therefore this method was not sensitive enough for detecting photoadducts.

4.5.6 Agarose Gel Electrophoresis

The EMSA analysis of photoadducts on PAGE gels were limited to studies of small dsDNA targets less than 200 bp in size, due to the poor separation and migration of larger triplexes and duplexes. Agarose gel electrophoresis was investigated as an alternative for the detection of photoadducts when targeting psoralen-conjugated oligonucleotides to larger dsDNA targets such as bovine genomic DNA.

Initially, small radioactive synthetic DNA duplexes, PCR products and plasmid DNA containing the triplex region were targeted with a specific psoralen-conjugated TFO, UV irradiated, electrophoresed by PAGE and agarose procedures and finally exposed to X-ray film. Alternatively, the agarose gel was blotted to a membrane and exposed to X-ray film. Agarose and PAGE gels both clearly showed bands representing dsDNA but only PAGE gels detected photoadducts (results not shown).

Increasing the percentage of agarose in the gel (1-4%), the duration of agarose gel electrophoresis (1-16 hr), and using a variety of agarose products such as Metaphor® (FMC Bioproducts, USA), SeaKem® LE Agarose (FMC Bioproducts, USA), and Agarose-1000 (GIBCOBRL, USA) all failed to lead to the detection of photoadducts on agarose gels. Alternative agarose gel staining techniques such as Ethidium bromide and GelStar® nucleic acid gel stain (FMC Bioproducts, USA) were also tried unsuccessfully.

Chapter 5

DISCUSSION

5.1 GENERAL OVERVIEW

The main objective of this study was to develop psoralen-conjugated TFOs that were capable of efficiently targeting specific regions of the bovine male Y chromosome. A photobiological system was established to determine the optimal UV-induced photoadduct conditions for *in vitro* DNA triplex formation with specific sequences within the bovine Y chromosome. These target sequences ranged in size from 38 to 3500 bp.

The specific aim of this study was to target repetitive bovine specific Y chromosome DNA sequences in an effort to significantly increase the efficiency of DNA triplex formation. Current antigene research is based on targeting TFOs to a single gene copy to directly interfere with gene expression. However, a major problem is that a TFO with relatively high specificity but low affinity for a specific target sequence may be ineffective in inactivating the one copy of the target sequence present in the genome. The efficiency may increase substantially if 500-1000 copies of this target DNA triplex site are present.

Potential target sequences within repetitive Y-chromosomal DNA were restricted to homopurine-homopyrimidine stretches of at least 10 bases in length, as stable *in vitro* DNA triplex formation is only achieved by TFOs recognising the homopurine strand of the DNA duplex. This crititeria was investigated using the GenBank database and Blast search programmes. Once suitable bovine DNA sequences were found, synthetic dsDNA targets were radioactively labelled and DNA triplex binding assays were then performed to test the targeting of psoralen-conjugated TFOs to specific dsDNA sequences. Following UV irradiation, electrophoretic mobility shift assays (EMSA) were performed to detect photoadduct formation.

A number of psoralen-conjugated TFOs were screened and, after the optimisation of a number of parameters (base composition, pH, temperature, magnesium and potassium concentration), seven psoralen-conjugated TFOs were found to selectively and specifically bind to their dsDNA bovine target forming covalent DNA triplexes upon photoactivation with UV light (360 nm).

5.2 THE DEVELOPMENT OF A PHOTOBIOLOGICAL SYSTEM

A number of experimental conditions such as the psoralen linker length, the constituents of the DNA triplex binding assay, and the UV light source and dose were investigated during the development of an efficient photobiological system. Initially, the well-documented *sup*F DNA triplex model was used to analyse the efficiency of the UV photobiological system before testing for bovine DNA triplexes with specific repetitive Y chromosome dsDNA target sequences.

The *sup*F photoadduct experiments originally published by Wang *et al.*, (1995) and Gasparro *et al.*, (1994) were repeated by targeting a 10 nt psoralen-conjugated TFO (PsoAG10) to both a 24 and a 38 bp synthetic *sup*F dsDNA target. The original authors used a linker length of C2 between the psoralen and the 5' of the TFO. Bates *et al.*, (1995) compared the linker lengths of psoralen C2 and C6, and showed that both have the same pattern of adduct formation suggesting a similar site and orientation of the psoralen, however C2 formed crosslinks more slowly than psoralen C6. In the present study the C2 *sup*F length was used with EMSA analysis showing that photoadduct formation was very efficient and consistent with previously published *sup*F data with 60% crosslink and 25% monoadduct induced by 32 minutes of 360 nm UV irradiation (2 J/cm^2). Given this high efficiency, it was decided to use a C2 linkage for all subsequent work.

The DNA triplex binding assay contained a dsDNA fragment, psoralen-conjugated TFO, spermidine, Tris-HCl buffer, and magnesium chloride. Magnesium ions are known to inhibit TFO dissociation and therefore promote DNA triplex formation. Hence, a high magnesium ion concentration of 10-100 μM was included in the DNA triplex binding assays.

Vasquez *et al.*, (1995) showed that a magnesium ion concentration of 10-20 μM did not enhance the rate of DNA triplex formation but did increase the stability of the DNA triplex itself. The DNA triplex was stable for 17 hours when a TFO containing guanines and thymine bases was targeted to intron 1 of the hamster *aprt* gene in the presence of at least 5 μM $MgCl_2$. A high spermidine concentration of 1 mM was included in the current experiments as recommended by other investigators to stabilise the DNA triplex (Krawczyk *et al.*, 1992; Gasparro *et al.*, 1994 and Havre *et al.*, 1993).

The efficiency of the UV irradiation was investigated to determine if changing the UV source and dose could significantly enhance *sup*F photoadduct formation. The psoralen reactions are versatile in that the photochemistry can be controlled by carefully timing UV delivery and by varying the wavelength and power of the UV source (Tessman *et al.*, 1985; Chatterjee *et al.*, 1978). A variety of additional UV sources including a 312 nm UV transilluminator (TFX-35M GIBCOBRL, Life Technologies), hand held UV light (Birtcher UV filter N#627) and a 254 nm UV light cabinet (15W/G15.T8 Phillips UV Bulb) were analysed for the efficiency of *sup*F photoadduct formation. A mylar sheet (3M Visual Systems Products, France) that was 0.5 mm in thickness, was wrapped around each UV lamp to prevent the emission of UVB rays.

Most investigators use a UV dose range of 0-20 J/cm^2 for generating *in vitro* photoadducts. Bates *et al.*, (1995) showed that ten minutes of UV irradiation using a 366 nm UV lamp (2.3 J/cm^2) resulted in 60-70% of the targeted psoralen-conjugated TFO forming stable crosslinked *sup*F photoadducts with the dsDNA target. These authors also showed that the use of a more intense UV light source such as a long wave UV transilluminator resulted in at least 90% crosslinked photoadduct formation. However, a high UVA dose can lead to photoiosmerization and photodegradation of the psoralen (Gasparro *et al.*, 1994). In the present study, the blacklight tube (360 nm, OSRAM L18 W/73) was the most efficient for *sup*F photoadduct formation (60% crosslink formation after 32 minutes of UV light, 2 J/cm^2, Appendix VII) and causes minimal DNA damage. Therefore this UV light source was used for subsequent DNA triplex assays.

The heat generated by the UV light source during UV irradiation was investigated as this may induce instability of the DNA triplex. The temperature of the tube containing the DNA triplex binding sample was recorded before, during, and after UV exposure (360 nm). Following thirty minutes of UV irradiation (2 J/cm^2) in the coldroom (4°C), the temperature increased by two degrees. It has been shown that the extent of crosslinking remains essentially unchanged over temperatures of between 4 and 30°C but begins to rapidly decrease at higher temperatures (Esposito *et al.*, 1988).

A variety of temperatures and incubation periods were investigated for their possible effects on DNA triplex formation. The following incubations were performed before UV irradiation of the DNA triplex binding sample: two hours at 37°C; two hours on ice; overnight and room temperature. No significant difference was detected after UV exposure when screening all psoralen-conjugated TFOs with their specific dsDNA target. Most of the DNA triplex binding assays were incubated at 37°C for two hours as this was a time period found to be sufficient to approach equilibrium for third strand binding by purine TFOs in the antiparallel motif (Gunther *et al.*, 1996).

5.3 PSORALEN-CONJUGATED TFOs

DNA triplex binding assays were performed using psoralen-conjugated TFOs, thus utilising the photobiological and intercalating properties of psoralen. The covalent linkage of psoralen at the 5' end of the TFO increases DNA triplex stability and also decreases nuclease degradation. The psoralen intercalates between adjacent pyrimidines at the 5' end of the triplex-duplex junction upon UV irradiation (360 nm), forming stable photoadducts. In this way, the DNA triplex is converted from a relatively unstable non-covalent complex to a stable covalent complex (Kanne *et al.*, 1982).

The chemistry of photoreaction has been studied extensively but detailed information regarding the structure of nucleic acid-psoralen adducts is lacking (Kanne *et al.*, 1982). Indeed, it is often difficult to predict the sequence specificity of the reactivity of psoralen with DNA. Studies of psoralen intercalation have mostly utilised free synthetic psoralen and calf thymus DNA, rather than psoralen-conjugated oligonucleotide (Gasparro *et al.*, 1994).

Most studies concentrate on thymine as the target for photomodification as they are the most abundant photoadducts found with psoralen. Ideally, a thymine-thymine crosslink is preferred over thymine-cytosine or a cytosine-cytosine crosslink (Hearst *et al.*, 1981; Kanne *et al.*, 1982). However, the reactivity of psoralen is sequence dependent with the preferred order of 5'-TpA > 5'-ApT >> 5'-TpG > 5'-GpT (Esposito *et al.*, 1988). The distance of 5'-TpA sites from the DNA duplex-triplex junction is the most important determinant of both the efficiency of photoadduct formation and the proportion of monoadduct and crosslinks (Vasquez *et al.*, 1996). The presence of a 5'-TpA site at the triplex-duplex junction is not an absolute requisite for psoralen cross linking; however, there is a requirement for a neighbouring 5'-TpA site close to the junction with at least one extra base at the corresponding end of the TFO (Faucon *et al.*, 1996).

The rate of cross-linking of psoralen to a 5'TpA sequence can vary by at least a factor of three depending on the base composition flanking 5'TpA (Espositio *et al.*, 1988). The photobinding of psoralen increases with increasing (A, T) content of the DNA and the number of monoadducts and crosslink adducts is highest when the (A, T) content of the DNA is 50-60% (Dall'Acqua *et al.*, 1978). The degree of crosslinking depends on the ratio of oligonucleotides bound to the target DNA and this is directly affected by the binding equilibrium constant, the rate of crosslinking, and the rate of photochemical formation of the inactive oligonucleotide. The rate of crosslinking also depends on the neighbouring bases and the long range DNA sequence content (Esposito *et al.*, 1988).

In the present work, a small number of 5'-TpA sites for psoralen intercalation, followed by at least ten continuous purine bases, were available for stable psoralen-conjugated TFO binding. In most cases, a cytosine rather than a thymine base was accessible for 5' psoralen intercalation following the homopurine region. Despite the lack of 5' TpA sites and/or non-homopurine bases for triplex targeting, I have shown that stable DNA triplexes can be formed within short bovine DNA fragments (38-49 bp), amplified DNA (199-500 bp) and plasmid DNA (3.2 Kb) containing a specified

DNA triplex region. Further DNA sequencing of the bovine genome may find potential DNA triplex regions for future psoralen-conjugated TFO targeting.

5.4 DISCUSSION OF RESULTS

An efficient photobiological system was established in which small psoralen-conjugated TFOs could be targeted *in vitro* to specific regions of the bovine Y chromosome. Psoralen-conjugated TFOs were targeted to synthetic dsDNA fragments, PCR products and plasmid DNA that contained a specific DNA triplex region. Following the DNA triplex assay and UV irradiation (360 nm) the photoadducts were detected by EMSA.

5.4.1 TFO Targeting to Small Synthetic dsDNA Fragments

In general, a small UV dose caused rapid monoadduct formation within *sup*F and bovine dsDNA targets. In contrast to monoadduct formation, crosslink formation was slow, presumably due to a requirement for sufficient monoadducts to accumulate and serve as precursors for the formation of crosslinks. The 40 bp Bov*SRY* fragment resulted in 39% cytosine-monoadduct and 4% crosslink formation (presumably a cytosine-thymine crosslink, respectively) after 8 minutes of UV exposure (0.5 J/cm^2) when targeted with 40 μM Puv12a. The 38 bp Bov*REP3* fragment resulted in 25% cytosine-monoadduct formation after 8 minutes of UV exposure (0.5 J/cm^2) when targeted with 40 μM Puv6a. The 49 bp Bov*REED* fragment resulted in 30% thymine-monoadduct formation after 8 minutes of UV exposure when targeted with 40 μM Puv15a. No crosslinks were observed with the Bov*REED* and Bov*REP3* dsDNA targets. Bovine monoadduct formation was as efficient as *sup*F monoadduct formation.

Targeting of unmodified TFOs (identical DNA sequence but with psoralen absent from the 5' end) did not form detectable levels of triplexes, with or without exposure to UV light. Two experimental procedures were used to determine whether unmodified TFOs formed non-covalent DNA triplexes when targeted to their synthetic dsDNA targets (Figure 3.6 and 3.9).

The first procedure involved targeting $\gamma^{32}P$ end-labelled TFOs to specific "cold" small dsDNA targets. The second procedure involved targeting "cold" TFOs to small $\alpha^{32}P$ and/or γ-^{33}P labelled dsDNA duplexes. All unmodified TFOs failed to form detectable levels of non-covalent DNA triplexes with their dsDNA targets when analysed by native EMSA.

A number of experimental conditions were tested in an effort to form non-covalent DNA triplexes; for example, the incubation time and temperature of the DNA triplex binding assay were increased. PAGE electrophoresis conditions were also adjusted by modifying the acrylamide concentration of the native PAGE gel (8-20%), electrophoresis time (1-4 hr), and electrophoresis temperature (4-30°C). A variety of electrophoresis buffers such as Tris-acetate-EDTA (pH 6.0, 8.0), Tris-borate-EDTA (pH 8.0), Tris-acetate-magnesium (pH 8.0), Tris-tricine (pH 9.0), Tris-glycine (pH 9.0) and cryogenic[12] (pH 8.0, Bain *et al.*, 1998) were also tried. In most cases, covalent DNA triplex formation was detected with psoralen-conjugated TFOs in the presence of UV light (360 nm) under all the above conditions (Sections 3.3.2 and 3.3.3). The fact that non-covalent DNA triplex formation was not detected was presumably due to the dynamics of DNA triplex formation and dissociation. Thus, the intercalating properties of psoralen-conjugated oligonucleotides are important for the formation of a stable DNA triplex. Psoralen-conjugated TFOs bind to the DNA duplex and upon UV irradiation form a covalent DNA triplex rather than a non-covalent triplex that can dissociate.

5.4.2 TFO Targeting to PCR Products

Limited research has been published in which investigators have targeted psoralen-conjugated TFOs to large DNA targets such as PCR products and plasmid DNA that are at least 100 bp in size. Most researchers used small synthetic fragments for DNA targeting due to convenience and simplicity.

[12] Cryogenic buffer consists of Tris-glycine, ethanol, ethylene glycol and is stored at -4°C.

In this work, targeting psoralen-conjugated TFOs to small PCR products, ranging from 100-200 bp in size was initially investigated to confirm that DNA triplex formation was possible with larger DNA targets. In this work, psoralen-conjugated TFOs, Puv6a, Puv12a and PsoAG10 formed monoadducts when targeted to the amplified DNA target that contained the DNA triplex site. In general, thirty seconds of 360 nm UV light (0.03 J/cm^2) was sufficient to generate monoadducts. No crosslinks were detected. Psoralen-conjugated TFOs, Puv12b, Puv12c, Puv12d, Puvb, and Puv6c failed to form photoadducts. This was presumably due to the presence of pyrimidine bases within the homopurine region of the triplex region.

Following the success found with targeting psoralen-conjugated TFOs to amplified Bov*REP3*, Bov*SRY* and sup*F* targets, it was decided to increase the Bov*REED* target size to 400 bp. Hence, increasing the selectivity of the TFO binding to its dsDNA target. Monoadduct formation was detected with Puv15a following two minutes of UV irradiation (0.2 J/cm^2). It was observed that the absence of $MgCl_2$ prevented monoadduct formation. Hence, this observation confirms that magnesium ions stabilise the DNA triplex as previously stated in Section 1.2.4.

The Bov*REED* PCR product was unsuitable for TFO targeting as it migrated slowly on a 4% PAGE gel during electrophoresis and the triplex and duplex bands were very close together and poorly resolved. Hence for optimum discrimination, PCR products should be 100-200 bp for psoralen-conjugated targeting. The effect of *in vitro* targeting to a 200-400 bp sized PCR product was not determined as the amplified DNA would be poorly resolved on the 4% PAGE gel, making analysis of the resulting photoadducts difficult.

Interestingly, no crosslinks were detected in any of the all four PCR triplex experiments using PCR fragments (including the *sup*F targets) despite the fact that crosslinks were detected in all the previous synthetic DNA triplex experiments. Possibly, the concentration of monoadducts may not have been sufficient for the conversion into detectable crosslinks. Alternatively, it is known that several movalent cations, including K^+ ions inhibit DNA triplex formation (Cheng & Van Dyke, 1993). The presence of potassium chloride in the PCR reaction buffer (FC = 5 mM) could have contaminated the DNA triplex binding assay with K^+ ions.

In this study, 100 mM KCl inhibited *sup*F photoadduct formation by 20%. Previous TFO targeting research has shown that 50-95% triplex inhibition occurs when 10-100 mM K^+ ions are present in the DNA triplex binding assay (Cheng & Van Dyke, 1993; Musso *et al.*, 1996).

Cheng *et al.*, (1993) suggested that K^+ ions stabilise an oligonucleotide species that is refractory to triplex formation, such as a guanine quartet, or possibly the oligonucleotides assume a particular conformation prior to their association into triplexes, with monovalent ions (M^+) inhibiting this isomerization. Alternatively, M^+ ions may directly interfere with TFO-DNA triplex association through competition with divalent cations (Maher *et al.*, 1990). The presence of K^+ ions does not substantially decrease the stability of triplexes but decreases the kinetics of their formation (Debin *et al.*, 997).

A variety of methods were employed to remove K^+ ions after the PCR reaction and prior to the DNA triplex binding assay, with the aim of increasing the formation and stability of the PCR photoadducts. No significant difference was detected when using purified amplified DNA in the UV crosslinking assays. There are now a variety of new PCR purification kits on the market such as the High Pure PCR Product Purification Kit (Boehringer Mannheim, Germany) that removes K^+ ions from the purification protocol. Future research should determine whether or not these products have any affect on improving PCR photoadduct formation.

The behaviour of triple helices in mixed-valency solutions, ie. in solutions containing monovalent, divalent, and multivalent cations as found in the cell, is more complex. Further research is required to prevent any possible K^+ ion inhibition, as this will ultimately affect future *in vivo* experiments due to the high intracellular K^+ ion concentrations within cells.

5.4.3 TFO Targeting to Plasmid DNA

Psoralen-conjugated TFOs were targeted *in vitro* to *Hind*III digested 5' end-labelled plasmid DNA that contained the appropriate bovine triplex region. Following UV irradiation, the cloned PCR insert and PCR triplex was released by *Eco*RI digestion. The ability to detect triplex formation to the intact plasmid vector was limited due to the poor electrophoretic resolution of photoadducts and plasmid DNA. Alternative methods were investigated to address this particular issue such as 3' end-labelling of psoralen-conjugated TFOs and DNase Footprinting but to no avail.

Monoadducts were detected when psoralen-conjugated TFOs, Puv6a and Puv12a were targeted to plasmid DNA that contained the appropriate bovine triplex region. Thus, photoadduct formation was very selective as the psoralen-conjugated TFO did not bind non-specifically to the plasmid that was approximately 3 Kb in size. *In vitro* targeting of Puv6a to a radioactive ^{32}P plasmid DNA containing the Bov*REP3* triplex region resulted in 77% monoadducts after 30 minutes of UV irradiation (2 J/cm^2). Monoadducts could also be detected using concentrations of psoralen-conjugated TFOs as low as 1 μM and as little as 5 minutes of UV irradiation (0.3 J/cm^2). As the UV dose increased the intensity of these monoadducts increased.

Competition assay with excess non-radioactive plasmid DNA, genomic male and/or female bovine DNA did not suppress the formation of monoadducts. This was due to the high molar ratio of TFO with respect to the dsDNA target and therefore an assay where the TFO was competing for radioactive dsDNA targets over non-radioactive targets was not achieved. Adjusting this ratio did not significantly improve this competition assay due to the high concentration of TFO required for EMSA photoadduct detection.

5.4.4 The Nature of the Bovine Y-chromosomal Repeat Regions

One complication in these experiments was not fully appreciated at the beginning of this investigation, was the exact nature of the sequences contained within the published bovine Y-chromosomal repeat regions. In the case of Reed *et al.*, (1995), the published sequence was obtained from a lambda genomic library and various experiments involving Southern blotting, dot blots and PCR reactions indicated that this sequence was indeed male specific and present in high copies (less than 500). All the initial experiments performed in this current investigation indicated that this was the case. Thus, PCR reactions using primers FBov*REED* and USP2 confirmed these results. Southern blot analysis with this PCR probe gave strong male-specific bands and sequence analysis of the PCR product gave clear non-degenerate sequence which agreed with that published by Reed *et al.*, (1995).

It was only when DNA triplex experiments were attempted with bovine genomic DNA that some reservations surfaced about the exact nature of these repetitive Y-chromosomal sequences. The first doubt arose from the fact that small oligonucleotide probes to the triplex region within the repetitive sequence did not give strong hybridisation signals on Southern blots of bovine digested male genomic DNA. Indeed, in comparison with controls of plasmid DNA, I estimate that the copy number for the "Reed" sequence must be less than 500. The most likely explanation for this discrepancy in the copy number between the "Reed" sequence overall and the triplex region within it, is that the "Reed" sequence is actually a consensus type sequence and considerable variations occur in certain internal regions. There are several pieces of evidence, which suggest that this might be the case.

Firstly, the sequences originally published by "Reed" actually do consist of a number of consensus repetitive sequences, with some regions and primer sequences that are specific to bovine and others of a more universal nature, for example, USP1 and USP2 primers bind to cattle caprine and ovine species. However, all of these are sufficiently similar and bind to the published Reed DNA sequence.

Secondly, it is well documented that when a genomic library is created, unless specific precautions are taken with the λ phage and host bacterial strains, many repetitive type sequences do not clone and there may be selective enrichment for certain sequences.

Thirdly, when the FBov*REED* and USP2 primers were used, the yield of PCR product, although high, reflects a much lower copy number when one does a series of dilution experiments. This last discrepancy appears to result from poor PCR efficiency, possibly due to the GC rich sequences. This can be partially overcome by using a hot start Qiagen *Taq* DNA polymerase. However, the resulting PCR product was not significant enough to suggest that this DNA sequence was highly repetitive (Ray Cursons, pers. comm).

Taken together, these considerations point to a family of similar rather than identical sequences constituting each Y-chromosomal repeat. It would appear that Reed *et al.*, (1995) has selected the more high conserved regions for their primers (some of which are species specific, some more universal) and it is also likely that specific subgroups of sequences are preferentially recovered during PCR and, possibly by the generation of λ libraries.

Furthermore, it would appear that the DNA triplex target regions I have chosen may not be highly conserved, as evidently shown by the low signal on Southern blots of bovine male genomic DNA. Perhaps the simplest method of choosing conserved sequences would be to select potential target regions, make oligonucleotide probes for these regions and check that they gave strong signals on Southern blots under high stringency conditions. Thus, further research is required to determine whether or not the psoralen-conjugated TFOs are sequence-specific to the DNA triplex site within bovine genomic DNA and also the effect of having multiple copies of the dsDNA target. To date, a large percentage of DNA triplex research has failed to show that the TFO is indeed specific to a single copy DNA target *in vivo.*

Targeting TFOs to bovine repetitive Y chromosome sequences will promote third strand binding specificity, and also selectivity with respect to male genomic DNA as opposed to female genomic DNA. One would also expect the presence of multiple copies to increase third strand binding and therefore enhance DNA triplex stability.

5.5 PYRIMIDINE MOTIF DNA TRIPLEXES

The pyrimidine motif TFO binds parallel to the dsDNA target forming a DNA triplex via reverse Hoogsteen hydrogen bonding. Previous studies have shown that C^{+}·GC DNA triplexes are destabilised at physiological pH due to the requirement of base protonation for cytosine (Bates *et al.*, 1996). Research presented here showed that three unmodified pyrimidine psoralen-conjugated TFOs, Puv2, Puv3, Puv12d failed to form a DNA triplex in the γ-^{33}P radioactive UV crosslinking studies. α^{32}P crosslinking studies showed some crosslink formation, but this was very poor in comparison with the purine psoralen-conjugated TFOs. Most of the pyrimidine psoralen-conjugated TFOs reported here had a higher ratio of cytosine to thymine bases.

Pyrimidine psoralen-conjugated TFOs were tested in a variety of conditions such as adjusting the incubation temperature before UV irradiation (37°C, 4°C, RT), incubation time before UV irradiation (2 hr, O/N), pH (5.0, 7.0, 8.0) and UV dose (0-4 J/cm^2). A range of magnesium (0-100 μM), spermidine (0-10 mM), tRNA (0.1-0.5 μg/μL), and NaCl (10-100 mM) concentrations were also tested. The effect of addition of 10% sucrose, 10% glycerol, and/or 1 mM DTT in the DNA triplex binding assay was also analysed. The adjustment of buffer and UV conditions did not result in the formation of a bovine pyrimidine motif DNA triplex.

Cytosine requires structural modification, such as methylation, to enable TFO targeting in physiological conditions. Majumdar *et al.*, (1998) modified a pyrimidine TFO by incorporating a 2'-O methyl sugar modification, methylating cytosines at the 5' position, and thioating three residues at the 3' end to increase DNA triplex stability. Cytosine methylation stabilises the DNA triplex due to the extensive protonation at the N3 position and is absolutely essential for *in vivo* targeting with a minichromosome model (Majumdar *et al.*, 1998). Thus, for future *in vivo* triplex targeting TFOs have to be significantly modified to increase DNA triplex stability and decrease nuclease degradation.

Modified pyrimidine motif TFOs were designed to bind to synthetic fragments representing part of a bovine repetitive Y-chromosomal DNA sequence. These modified TFOs incorporated psoralen at the 5' end and three phosphorothioate linkages at the 3' end. The cytosine bases were also substituted with 5'-methylcytosine residues. Results presented in Section 3.5 show that modified TFOs, *m*Puv6d, *m*Puv7b and *m*Puv7c target their dsDNA targets with high efficiency. A high UV dose of thirty minutes (2 J/cm^2) was sufficient to produce 50% monoadduct and 20% crosslink formation with 10 μM *m*Puv6d.

In vitro targeting of *m*Puv7b and *m*Puv7c was significantly enhanced due to the presence of two 5'TpA sites for psoralen intercalation and a homopurine run of 12 bases within the triplex region. A UV dose of thirty seconds (0.03 J/cm^2) was sufficient for 40-50% monoadduct formation. As the UV dose increased, a number of crosslinks were observed, mainly due to the run of thymine bases by the 5' end of the triplex-duplex junction. Repeated TA sequences are "hot spots" for the photochemical reaction of psoralen and DNA (Sage & Moustacchi, 1987). Thus *m*Puv7b and *m*Puv7c are very efficient for generating photoadducts.

5.6 PURINE MOTIF DNA TRIPLEX

The purine motif TFO binds antiparallel to the dsDNA target, forming a DNA triplex via Hoogsteen hydrogen bonding. The purine motif is preferred over the pyrimidine motif due to pH independence with respect to DNA triplex formation. For this reason, mostly purine psoralen-conjugated oligonucleotides were targeted to bovine Y-chromosomal DNA sequences, and in some cases, there was at least one pyrimidine mismatch within the homopurine strand of the dsDNA target. Therefore the purine TFO had to be modified to avoid instability of the DNA triplex. Extensive research has been aimed at increasing the third strand binding code for triplet combinations to overcome limitations inherent in using homopurine-homopyrimidine targets for DNA triplex formation. So far, moderate success has been achieved by using the following mismatched triplets within the purine motif DNA triplex: T·CG, T·TA and G·TA.

Bates *et al.*, (1995) showed that the following non-canonical DNA triplex could form T·CG(>)C·CG>>G·CG>A·CG. A purine base targeted to a CG duplex was highly destabilising and photoadduct formation was almost abolished when such a triplet was introduced.

The position of the mismatch within the TFO is critical to DNA triplex formation. A mismatch at the 5' end or middle of the TFO can effect the stability and formation of the DNA triplex. In this work, psoralen-conjugated TFOs, Puv6b, Puv6c, Puv12b, Puv12c, and Puv15b all had at least one pyrimidine mismatch present within the homopurine region of the triplex site. The mismatch was present in the middle and/or 3' or 5' end of the TFO.

Purine psoralen-conjugated oligonucleotides, Puv7a and Puv4 failed to form photoadducts even though there were no pyrimidine mismatches and a pyrimidine site available for psoralen intercalation at the 5' end of the triplex-duplex junction. No explanation is yet available as to why some GA triplexes are difficult to induce in some sequences yet very stable for others (Giovannangeli *et al.*, 1997).

My results show that a high guanine content (at least 50%) within the TFO is very important for high affinity DNA triplex formation. These findings are consistent with previous research. Clarenc *et al.*, (1994) detected no DNA triplexes when there was a low guanine content (22-50%) in the TFO, with Perkins *et al.*, (1998) recommending that at least 54% guanine bases be present in the third strand. However, de Bizemont *et al.*, (1996) showed that a high guanine content in GT containing TFOs was not a prerequisite for DNA triplex formation. However, the major disadvantage of using GT or GA oligonucleotides is the tendency of guanines to self-associate, forming stable quadruplex structures. This can be overcome by modifying guanosine as described previously in Section 1.2.4.

5.7 THE POTENTIAL FOR PHOTOADDUCT FORMATION WITHIN BOVINE SPERM CELLS

The delivery of TFOs to their *in vivo* cellular targets has been investigated for the last decade. The use of bovine Y-specific chromosomal DNA sequences provides a potential non-invasive delivery system for passing psoralen-conjugated TFOs into spermatozoa and forming DNA triplexes following UV irradiation (360 nm). It also has potential for a range of future spermogenic, embryological and animal manipulations such as knocking out the Y chromosome. Brackett *et al.*, (1971) was the first person to report the capacity of sperm cells to capture foreign DNA. In this pioneering report, rabbit sperm cells were exposed to [^{3}H]-thymidine labelled SV40 DNA and 30-35% radioactive material was detected in the postacrosomal area of the exposed sperm cells. In 1989, binding of exogenous DNA was shown in sea urchin (Arezzo, 1989) and mouse sperm (Lavitrano *et al.*, 1989).

The ability of sperm cells to take up foreign DNA is not passive and random; it is actually a complex biological process, during which both the binding of DNA to the sperm head and its nuclear internalisation are modulated by the activity of specific factors (Lavitrano *et al.*, 1992; Sperandio *et al.*, 1996). Two factors have been identified so far; a class of proteins acting as DNA binding substrates and Inhibitory Factor 1 (Zani *et al.*, 1995). The presence of these inhibitory factors in the seminal plasma might prevent foreign DNA from gaining access to the spermatozoa by binding to the DNA molecules or competing with the DNA for the same binding region on the sperm surface (Camaioni *et al.*, 1992).

Siracusa (1989) showed that the exogenous uptake of DNA can only occur in live sperm, with the capacity of sperm binding to DNA being astonishingly high; 0.3% plasmid DNA in the sperm head (unpublished but quoted by Birnstiel *et al.*, 1989). The sperm-DNA association kinetics are rapid with maximum values reached within 20-40 minutes in rabbit, mouse, boar, bull and blowfly sperm (Camaioni *et al.*, 1992).

The uptake of small TFOs maybe a problem, as mammalian sperm cells are more efficient in taking up large DNA molecules (7 Kb) compared to small molecules (150-750 bp). However a 80 bp DNA fragment has been reported to be taken up by sperm (Lavitrano *et al.*, 1992; Chan *et al.*, 1995). Sperm cells also able to take up negatively charged macromolecules such as heparin, dextran sulphates, and proteins with isoelectric points of less than 7 (Lavitrano *et al.*, 1992). The amount of negative charge on the DNA molecules may be important for the interaction between the nucleic acid and sperm cells (Lavitrano *et al.*, 1992). To date no experiments have been published showing that sperm have taken up TFOs less than 20 nt in size.

Many delivery systems to efficiently introduce DNA into sperm cells have been investigated. One study showed that sperm retained exogenous DNA introduced via electroporation (59%), liposomes (52%) or passive diffusion (6%). However, electroporation has been shown to affect fertilisation *in vivo* (Nakanishi *et al.*, 1993).

5.8 THE FUTURE

The problems encountered with the antigene strategy must be solved in order for this technique to remain a viable option for future research or applications. The optimisation of TFO uptake into cells and tissue, increased stability of the DNA triplex *in vitro*, demonstration that specific binding occurs to DNA targets within the chromosome and no where else in the genome, development of new DNA damaging agents, and expansion of the restricted DNA triplex binding code are some of the areas where further work is necessary.

Psoralen applications will ultimately be limited due to the difficulty of UV activation of the psoralen moiety in an *in vivo* situation and also the removal of the DNA triplex directed psoralen adducts by cellular repair mechanisms. UV is a known mutagen; studies have shown that the frequency of chromosome aberrations increase remarkably with UVA irradiation (Matsuda *et al.*, 1988; Libbus *et al.*, 1987) although not to the extent induced by UVB and UVC. However, the use of sperm as a delivery system may overcome the problems of UV activation. UVA lasers have been used for many years in flow cytometric sorting to accurately analyse the DNA of X and Y chromosome bearing sperm. This clinical technique has resulted in several hundred normal births in animals and at least one human birth (Catt *et al.*, 1997), suggesting that sperm can be exposed to UV light without sustaining damage. Fluorescence-activated cell sorting (FACS) has also been used for many years to analyse and sort viable sperm cells according to their fluorescent levels following 488 nm laser beam excitation (Ericsson *et al.*, 1989).

Psoralen and UVA light has been used for twenty five years for the treatment of psoriasis. In most cases, the UV doses that psoriasis patients receive ranges from 3 to 15 J/cm^2 (Gasparro *et al.*, 1996). This UV dose is considerably higher than the doses tested here. Hence, the prospects for *in vivo* bovine sperm triplex targeting look encouraging and represent an interesting area of work for future studies. Psoralen-conjugated TFOs will require further modification, especially to the 3' end and the internal backbone, to ensure they are stable within the spermatozoa.

Chapter VI

CONCLUSION

The antigene strategy has become well established since the discovery of the DNA triplex more than 40 years ago. So far, three major parameters have been defined for designing and targeting synthetic psoralen-conjugated oligonucleotides to the major groove of dsDNA, forming stable photoadducts upon UV irradiation. Firstly, the sequence composition of the dsDNA target must have at least ten purine bases. Secondly, the third strand binds parallel (purine motif) or antiparallel (pyrimidine motif) with respect to the purine strand of the DNA duplex. Finally, the psoralen moiety preferentially intercalates into pyrimidine bases at the 5' end of the triplex-duplex junction, forming stable covalent photoadducts upon UV irradiation.

With these considerations in mind, a photobiological system was developed to target psoralen-conjugated TFOs to bovine repetitive Y chromosome sequences (Bov*REED*, Bov*REP3*, Bov*GROBET*) and also to bovine Y chromosome genes such as SRY and TSPY. Initially, experiments were performed using the well-established *sup*F DNA triplex system, thus allowing the photoadducts to be compared against the bovine DNA triplex system. *sup*F photoadduct formation was detected by EMSA, with high efficiency as previously reported by Wang *et al.*, (1996) and Gasparro *et al.*, (1994).

Twenty three psoralen-conjugated TFOs (10-25 nt) were targeted to various dsDNA fragments that ranged from 24 bp to 3.2 Kb in size. Based on the results of these experiments, the following conclusions were made:

1. Only six purine rich psoralen-conjugated TFOs formed photoadducts *in vitro*; PsoAG10, PsoAGT17, Puv6a, Puv12a, Puv14a, and Puv15a. These oligonucleotides contained a high guanine content, which supports the theory that this is a requirement for stable DNA triplex formation. However, the major drawback is that guanines have a tendency to self-associate and form guanine quartets.

2. Psoralen intercalation sites were available at the 5' end of the triplex-duplex region. In most cases an adjacent cytosine and thymine in the DNA duplex, allowed a psoralen crosslink to form. Ideally, two thymine bases adjacent to each other would be better, but in the case of the bovine genome it was extremely difficult to find sequences with a 5'TpA site next to a long stretch of purine bases. The bovine repetitive Y-chromosomal sequences that were tested here consisted mainly of guanines with a combination of pyrimidine bases. Despite this, I have shown that cytosine monoadducts can form, and they form at the same efficiency as the *sup*F monoadducts.

3. Psoralen-conjugated TFOS, Puv12b, Puv 12c, had pyrimidine bases within the homopurine strand. Thus, the third strand had to be adjusted to enhance DNA triplex stability. This was achieved by targeting the guanine to thymine in the homopurine region to form the triplet G·TA or alternatively targeting thymine to cytosine in the homopurine region to form the triplet T·CG. Research has shown that these triplet combinations can overcome the instability of the DNA triplex due to the presence of pyrimidine mismatches. However, in the cases described here, these triplet combinations did not significantly improve DNA triplex stability.

4. Psoralen-conjugated TFOs, Puv2, Puv3 and Puv12d, consisted of pyrimidine bases and were designed to bind parallel to the DNA duplex. The γ^{32}P UV crosslinking studies showed that Puv12d produced faint monoadducts, less than 10% following 30 minutes of UV irradiation (2 J/cm^2). Previous research has shown that cytosine requires an acidic pH for DNA triplex formation. Under the pH conditions tested here (pH 5.0, 7.0 and 8.0) it was not possible to detect any γ^{33}P photoadducts. Thus, the cytosine bases in the third strand require structural modification for stable DNA triplex formation. Three pyrimidine modified psoralen-conjugated TFOs which contained three phosphorothioate linkages at the 3' end and also, 5'-demethylcytosine residues, were designed and targeted to Bov*REP4* and Bov*REED2*. Significant photoadduct formation was detected in comparison to other non-modified psoralen-conjugated TFOs.

5. Psoralen-conjugated TFOs, Puv4 and Puv7a consisted of purine bases and were designed to bind antiparallel to the DNA duplex. No pyrimidine bases were present in the homopurine region of the triplex site and there were also pyrimidines available for 5' intercalation at the triplex-duplex junction. Even though these TFOs followed the established DNA triplex rules, no photoadducts were detected. This could be due to unfavourable DNA triplex association and dissociation kinetics and also low guanine content in the TFO.

6. In general, only a small UV dose (0.03 J/cm^2) was required to photoactitivate the psoralen moiety to selectively and efficiently form a monoadduct. As the UV dose increased (2 J/cm^2), the intensity of the monoadduct band increased. Photoadducts were also detected when psoralen-conjugated TFOs were targeted to amplified DNA and plasmid DNA that contained the bovine DNA triplex site. Therefore, these results show that the third strand binds specifically to its dsDNA target (up to 3.2 Kb) to form a DNA triplex.

7. Researchers have used simple *in vitro* systems to detect photoadducts. A variety of methods such as using 3' end-labelled and mixed psoralen-conjugated TFOs were investigated to target psoralen-conjugated TFOs to bovine genomic DNA. A suitable method to detect DNA triplexes within bovine gDNA was not developed due to the low sensitivity of detecting photoadducts with the techniques and resources currently available.

8. PCR analysis and Southern blot analyses showed that Bov*REED*, Bov*REP3* and Bov*TSPY* PCR probes were specific to bovine male genomic DNA as opposed to bovine female genomic DNA. The DNA sequencing results suggested that a single sequence was present in those regions of the Y-chromosomal repeats rather than degenerate sequences containing variations of a basic consensual sequence. However, small oligonucleotide probes containing the DNA triplex region (16-49 nt) failed to hybridise to male genomic DNA of high molecular weights. Thus, the male specificity and the repetitive nature of the DNA triplex region are indeed questionable.

9. In summary, this *Doctor of Philosophy* research thesis has demonstrated that psoralen-conjugated oligonucleotides work efficiently in practice. Future work will require more complex analysis of the bovine Y-chromosomal repetitive sequences to ensure that the TFOs are indeed being targeted to highly repetitive regions. Once this has been achieved, there would seem little doubt, based on the findings presented here, that a high number of extremely efficient covalent DNA triplexes can be induced in male bovine genomic DNA.

Appendix I

SOLUTIONS

1.0% Agarose Gel

1.0 g Agarose
100 mL 1X TAE buffer
Microwave on high until agarose has dissolved and then add 5 µL ethidium bromide (10mg/mL) when the temperature is less than 60°C.

7.5 M Ammonium Acetate

144.53 g ammonium acetate in 250 mL sterile dH_20. Filter sterilise.

10% APS

1.0 g ammonium persulphate in 10 mL dH_20. Store at 4°C for up to one week.

Ampicillin

50 mg/mL ampicillin in sterile dH_20. Filter sterilise and store at -20°C.

Church & Gilbert Solution

Solution A: 89 g $Na_2.HPO_4{\cdot}2H_2O$ in 1 L of H_20. Adjust pH to 7.2 with H_2HPO_4. Autoclave.
Solution B: 70 g SDS in 500 mL H_20.
Heat to 37°C and then add 1 mL 0.5 M EDTA (pH 8.0).
Mix 500 mL of Solution A and 500 mL of Solution B for C & G Solution.
Store at 37°C.

Deionized Formamide

50 mL Formamide
5.0 g Mixed-bed, ion-exchange resin
Mix and stir for 30 minutes at RT. Filter through Whatman N# 1 Filter paper.
Aliquot 5 mL volumes and store at -20°C.

50% Deionized Urea Solution

250 g Urea
500 mL H_2O
5.0 g Mixed-bed, ion-exchange resin
Mix and stir with a magnet for 30 minutes at RT. Filter through Whatman N#1 Filter paper. Store at RT.

0.5 M EDTA (pH 8.0)

186.1 g EDTA in 800 mL H_2O. Adjust pH to 8.0 with NaOH pellets.
Adjust the final volume to 1 L with sterile dH_2O and sterilise by autoclaving.

Elution Buffer

0.5 M NH_4Ac, 10 mM MgAc, 1 mM EDTA (pH 8.0), 0.1% SDS

Ethidium Bromide

10 mg/mL EtBr in sterile dH_2O. Store in the dark at RT.

ETS Buffer

10 mM EDTA (pH 8.0), 25 mM Tris-HCl (pH 8.0), 0.5% SDS

10X Formic Acid Buffer

2.6 mL Formic acid
20.3 g Tris base
Adjust the final volume to 100 mL with sterile dH_2O.

2X Freezing Medium

6.3 g K_2HPO_4
0.45 g Sodium citrate
0.09 g $MgSO_4.7H_2O$
1.8 g $(NH_4)_2SO_4$
44 g Glycerol
Sterile dH_2O to 500 mL. Sterilise by autoclaving.

6X Gel Loading Dye (Type III)

0.25% bromophenol blue, 0.25% xylene cyanole FF, 40% glycerol in H_20, 59.5% 50 mM Tris-HCl buffer (pH 8.0).

Glucose

50% Solution: 50 g glucose in 100 mL sterile dH_20. Filter sterilise.
1 M Solution: 18 g glucose 100 mL sterile dH_20. Filter sterilise.

G25 Sephadex Solution

Carefully combine 10 mL TE buffer (pH 8.0) to 1 g of G25 Fine Sephadex in a sterile tube. Incubate overnight at RT to allow the beads to swell. Equilibrate three times with 1X TEN buffer. Autoclave and store at RT.

IPTG

200 mg/mL IPTG in sterile dH_20. Filter sterilise and store at -20°C.

Kanamycin

25 mg/mL kanamycin in sterile dH_20. Filter sterilise and store at -20°C.

Loading Buffer for Denaturing PAGE Electrophoresis

98% deionized Formamide, 10 mM EDTA (pH 8.0), 0.025% xylene cyanole FF, 0.025% bromophenol blue. Store at 4°C.

1.0% LMP agarose gel

Exactly the same as 1.0% agarose gel but replace the agarose with LMP agarose.

2 M Magnesium Chloride

Dissolve 203.30 g $MgCl_2$ in 500 mL sterile dH_20. Sterilise by autoclaving.

Polyacrylamide Gels for PAGE Electrophoresis

A) Denaturing PAGE gels for large PAGE electrophoresis

	ACRYLAMIDE GEL (%)		
Solution mix	4	6	8
20% Polyacrylamide solution	10 mL	15 mL	20 mL
50% deionized Urea solution	20 mL	20 mL	20 mL
5X TBE buffer	10 mL	10 mL	10 mL
dH_20	10 mL	5 mL	--
10% APS	0.5 mL	0.5 mL	0.5 mL
TEMED	50 µL	50 µL	50 µL

B) Native PAGE gels for mini PAGE electrophoresis

	ACRYLAMIDE GEL (%)			
Solution Mix	6	8	15	20
30% Polyacrylamide solution	2.00 mL	2.67 mL	5.00 mL	6.67 mL
5X TBE buffer	2.00 mL	2.00 mL	2.00 mL	--
Tris-Tricine buffer	--	--	--	--
10X Formic acid buffer	--	--	--	1.00 mL
H_20	6.00 mL	5. 33 mL	3.00 mL	
10% APS	35 µL	35 µL	35 µL	35 µL
TEMED	3.5 µL	3.5 µL	3.5 µL	3.5 µL

PAGE gels were prepared in 50 mL Falcon tubes to a desired acrylamide percentage as shown in the above tables. The gel was mixed thoroughly and poured quickly before the solution polymerized. When large native PAGE gels were required, a 50 mL final volume was prepared instead of the described 10 mL volume.

Plate Agar

Identical to LB or SOC medium except that it contains 15 g Bactoagar.

20% Polyacrylamide Solution (19:1)

190 g Acrylamide
10 g N,N' Methylbisacrylamide
Dissolve the above ingredients in 600 mL 50% deionized Urea solution. Heat at 37°C until dissolved and adjust the final volume to 1 L with 50% deionized Urea solution. Store in the dark at RT.

30% Polyacrylamide Solution (29:1)

29 g Acrylamide
1.0 g N,N' Methylbisacrylamide
Dissolve the above ingredients in 80 mL H_20. Heat at 37°C until dissolved and adjust the final volume to 100 mL with dH_20. Store in the dark at RT.

10X Restriction Enzyme Buffers

a) Buffer A: 330 mM Tris-Ac, 100 mM MgAc, 660 mM KAc, 50 mM DTT, pH 7.9

b) Buffer B: 100 mM Tris-HCl, 50 mM $MgCl_2$, 1 M NaCl, 10 mM 2-mercaptoethanol, pH 8.0.

c) Buffer H: 500 mM Tris-HCl, 100 mM $MgCl_2$, 1000 mM NaCl, 10 mM DTE, pH 7.5.

d) React 2: 500 mM Tris-HCl pH 8.0, 100 mM $MgCl_2$, 500 mM NaCl.

RNaseI

10 mg/mL RNase I in 10 mM Tris-HCl (pH 7.5), 15 mM NaCl. Heat for 15 minutes at 98°C, cool to RT, and aliquot 80 µL volumes. Store at - 20°C.

10% SDS

100g SDS in 900 mL H_2O. Heat to 68°C to assist dissolution. Adjust pH to 7.2 with conc HCl and then adjust the final volume to 1 L with dH_20.

SOB Medium

To 900 mL deionized H_20, add:
20 g Bactotryptone
5.0 g Bacto-yeast extract
0.5 g NaCl
Dissolve the above ingredients and add 10 mL of 20 mM KCl. Adjust the pH to 7.0 with 5N NaOH and adjust the final volume to 1 L with sterile dH_20. After sterilising the solution by autoclaving, add 5 mL of 2M $MgCl_2$.

SOC Medium

Identical to SOB medium except that is contains 20 mM glucose. After SOB medium has been autoclaved, cool to less than 60°C and add 20 mL of sterile 1M solution of glucose.

3 M Sodium Acetate (Ph 5.0)

408.1 g $NaAc.3H_2O$ in 800 mL H_2O. Adjust pH to 5.0 with glacial acetic acid. Adjust the final volume to 1 L with dH_2O and sterilise by autoclaving.

Solution 1

100 mM Tris-HCl (pH 7.5), 10 mM EDTA (pH 8.0), 400 μg heat treated RNase I/mL

Solution II

0.2 M NaOH, 1% v/w SDS. Store up to 1 month at RT.

Solution III

60.0 mL 5 M KAc
11.5 mL Glacial acetic acid
28.5 mL sterile dH_20
Sterilise by autoclaving.

100 mM Spermidine

0.15 g spermidine dissolved in 1 mL H_20. Store at -20°C.

20X SSC

175.3 g NaCl
88.2 g Sodium citrate
Dissolve the above ingredients in 800 mL H_20. Adjust the pH to 7.0 with a few drops of 10N NaOH. Adjust the final volume to 1 L with dH_20 and sterilise by autoclaving.

Stains-all Solution

50 mg Stains-all
400 mL Deionized Formamide
400 mL Deionized H_20
Store in the dark at RT.

STET Buffer

20 mL 40% Sucrose solution
10 mL 0.5 M EDTA (pH 8.0)
5 mL 1M Tris (pH 8.0)
0.5 mL TritonX100
Dissolve and adjust volume to 100 mL with sterile dH_20. Autoclave and store at RT.

Stop Buffer

2.5 mL Glycerol
0.005 g Bromophenol blue
1.0 mL 0.5 M EDTA (pH 8.0)
1.5 mL TE buffer (pH 8.0)

40% Sucrose

40 g sucrose into 100 mL sterile dH_20. Filter sterilise.

Tetracycline

5 mg/mL tetracycline in ethanol. Filter sterilise and store at -20°C.

50X Tris-Acetate Buffer (TAE)

242 g Tris base
57.1 mL Glacial acetic acid
100 mL 0.5 M EDTA (pH 8.0)
Adjust the final volume to 1 L and sterilise by autoclaving.

1X Tris- Acetate-Magnesium Buffer

90 mM Tris-borate
10 mM $MgCl_2$

5X Tris-Borate Buffer (TBE)

54 g Tris base
27.5 g Boric acid
20 mL 0.5 M EDTA (pH 8.0)
Make up to 1 L with sterile dH_20.

Tris-EDTA Buffer (TE)

10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)

1X Tris-EDTA-Sodium Buffer (TEN)

10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 100 mM NaCl

1M Tris-HCl (pH 8.0)

1.21g Tris base in 800 mL H_20. Adjust pH to 8.0 with conc HCl.
Adjust the final volume to 1 L with sterile dH_20 and sterilise by autoclaving.

Tris-Tricine Buffer

60.55 g Tris base
8.95 g Tricine
16 mL conc HCl
Adjust the final volume to 1 L with sterile dH_20.

2×YT Broth

To 900 mL of deionized H_20, add:

16 g Bacto-tryptone

10 g Bacto-yeast extract

0.5 g NaCl

Dissolve above ingredients, adjust pH to 7.0 with 5N NaOH and then adjust the final volume to 1 L with dH_20. Sterilise by autoclaving.

XGal

20 mg/mL in Dimethyl Formamide. Store at -20°C, wrapped in Aluminium foil.

Appendix II

SUPPLIERS

AJAX Chemicals, Australia:	Anhydrous D-glucose Boric acid powder EDTA Formamide Glycerol $MgCl_2$
American National Cam, USA:	Parafilm
Amersham International, UK:	[γ-^{33}P] dATP, 2000 Ci/mmol, 10mCi/ml [α^{32}P] dATP, 10 mCi/ml, 3000 Ci/mmol [α^{32}P] dCTP, 10 mCi/ml, 3000 Ci/mmol [α^{32}P] dGTP, 10 mCi/ml, 3000 Ci/mmol Hybond™-N$^+$ Nylon Transfer Membrane Rediprime Random Primer Labelling Kit
AxyGen Inc, USA:	Eppendorf tubes - 0.6 and 1.7 mL
Barnstead/Thermolyne, USA:	Barnstead E-pure® water system
Beckman, USA:	TL-100 Ultracentrifuge J2-21M Induction Drive Centrifuge
BDH, UK:	Ammonium acetate Bromophenol blue Butan-1-ol $CaCl_2$ Ethanol, 99.7 - 100% KCl NaOH PEG Triton X 100
B& F Papers, NZ:	Beermat

BioLab Scientific Ltd, NZ:	Whatman N#4 paper
BIORAD Laboratories, USA:	Acrylamide, 99.7% 40% Acrylamie/Bis Solution, 19:1 Bis N, N' methylene-bisacrylamide Gene Pulser Apparatus® Gene Pulser®/*E.coli* Pulser cuvettes Molecular Biology Grade AG® 501-X8 Resin Model 3000Xi Computer controlled electrophoresis power supply
BioScience Inc, USA:	MultiFit research pipette tips (in EcoStak racks)
Boehringer Mannheim, Germany:	dNTP (dATP, dCTP, dTTP, dGTP) *Eco*RI *Hae* III *Hind* III Klenow enzyme Polynucleotide kinase SDS *Taq* DNA polymerase Terminal transferase Xgal
Borden Ltd, Australia:	Plastic wrap
Cappelen Laboratory Technics ApS, Denmark:	CAPP advance single channel pipettes
Capral Aluminium Ltd, Australia:	All purpose Aluminium catering foil
Chiltern Scientific	MT19 Deluxe vortex mixer RR91 Rocket unit
Decon Laboratories Ltd, UK:	Decon® 90

DIFCO Laboratories, USA:	Bactoagar Bactotryptone Yeast Extract
Eastman Kodak Co, USA:	Stains-All
FMC Bioproducts, USA:	Metaphor® agarose Seakem® LE agarose SeaPlaque® agarose
GelmanSciences, USA:	Acrodisc 0.2 μm sterile syringe filters
Glen Research, USA:	Psoralen-conjugated oligonucleotides
John Morris Scientific Ltd, NZ:	Techne PCR machine
Kodak, Australia:	GBX Developer & Replenisher, GBXFixer & Replenisher X-OMAT AR film
Labsupply Pierce, NZ:	Centrifuge 5415C
Life Technologies, USA:	Ammonium persulfate BLX-254 UV crosslinker High DNA Mass Ladder Horizon 11.14 GibcoBRL Horizontal Gel electrophoresis apparatus Low DNA Mass Ladder 1 Kb and 100 bp Ladder Model S2 Sequencing Gel Electrophoresis system Model 250 Electrophoresis power supply Primers Tricine TFX-35M UV transilluminator Ultrapure Agarose-1000 *Xba*I
Intermed Scientific Ltd, NZ:	Cleanaid powder

Molecular BioProducts, USA:	ART pipet tips; ART 10, ART 10 Reach, ART 200
Nalge Nunc International, USA:	Nunc Disposable conical tubes; 15 & 50 mL
New Brunswick Scientific Co, USA:	G24 Environmental Incubator Shaker
Oligo's Etc., USA:	Psoralen-conjugated oligonucleotides
Owl Scientific Inc, USA:	Penguin™ Water-cooled Dual-Gel Electrophoresis system
Pharmacia Fine Chemicals, Sweden:	Sephadex® G-25, fine
PROLABO, UK:	Glacial acetic acid Isopropanol Phenol Tri-Sodium-Citrate
Promega, USA:	Wizard® PCR Preps, DNA purification system
Savant Instruments Inc, USA:	Speed Gel™ SG210D Gel Drier
Scientific Specialities Incorporated, USA:	UltraFlux Flat Cap PCR tubes UltraFlux Dome Cap PCR tubes
Sci Tech, NZ:	Mini Hybridisation Oven MK II
Shimadzu Scientific Instruments, Japan:	Shimadzu UV250 Spectrophotometer
SIGMA, USA:	Ampicillin Ethidium bromide

SIGMA, USA:	Glass wool Kanamycin Mineral oil RNase Tetracycline Spermidine TEMED Xylene Cyanole FF
Stratagene, USA:	VCSM13 helper phage pBluescript
Techno-Plas, Australia:	Stackable plastic petri dishes
Thermolyne, USA:	Type 17600 Dry-bath
Unelko Corporation, USA:	RainX
United States Biochemical, USA:	Tris Ultrapure Urea
Visual Systems Products, France:	3M Transparency Film (Mylar sheets)
Whatman Limited, UK:	Full range pH paper Whatman N#4 filter paper (qualitative)

Appendix III

DNA TARGET SEQUENCES

1. Repetitive *Bos taurus* Y-chromosomal DNA Sequences

A) The Complete DNA Sequence of Bov*GROBET*

LOCUS BTU39887 561 bp DNA MAM 29-NOV-1995
DEFINITION *Bos taurus* Y chromosome specific genomic sequence.
ACCESSION U39887, Grobet & Georges, 1995
NID g1079746
FEATURES /organism="*Bos taurus*"
/chromosome="Y"
Unsuccessful DNA triplex formation with Puv1, 2, 3, and 4
Forward primer (F) sequence is from 1-32.
Reverse primer (R) sequence is from 535-561.
Homopurine DNA triplex region is underlined (440-457)

BASE COUNT 197 a 111 c 115 g 138 t
ORIGIN

1 gatcttaatg acccagttag ctgtgatggt gtctttactc acctagaacc agacatcctg
61 gaatgtgaaa tcaagtagcc ctaagtaaat cactatgaac aaagttacta gaggtgatgg
121 gtttccagtc gtgcagtttc aagtcctatc agattatacc attaaaatcc tgcactcagt
181 atgccagaaa ctttggaaaa ctcggcattg gccacagaac tggaaaaggt tagtttaaag
241 tccaatccca aagaaaggct atgccaaaga atgttcaaac tacaacacaa ttacatgtat
301 cccacatgca agcaaagtga tgctcaaagt tctccaagcc aggcttcaac agtacgtgaa
361 cccagaattt ccagatgtta agggtggatt tagaaaagtc agaggaacca gagatcaaat
421 tgccaacatc cattggatcg gagaaaaagc aagggagttc aagatataca tctacatctg
481 atttggacta tgccaaaggt tttgactgtg tggatgacaa caaactatgg agaattcttg
541 aaaagatgca aatatcagat c

B) Complete DNA Sequence of US Patent 5459038, Reed *et al.*, (1995)

Four genomic DNA sequences are listed; BRY4a (Bovine), OY11.1 (Ovine), GRYla (Caprine) and Bcy11a (Bovine testes cDNA). The DNA triplex regions and PCR primer sequences are highlighted in colour.

Figure 13a:

```
BRY4a:    CACCACAAATAGTGCCTAATAGCTTTTC.A..GAGTTATGTTC
                    920        930          940       950
          ................ .. ....... .  ...........
OY11.1:   CACCACAAATAGTGCCTGATGGCTTTTCGAAGGAGTTATGTTC
         830      840        850         860       870

GRY1a:


                            10          20          30
                     CTTCACAAGTTCATTGCAGCATCCAGTAGAAACAGCAG<—— BCY11a
                     |||||||||||||||||||| ||||||||||||||||||
TCCTTTGATTACCTTCACAAGTTCATTGCAGGATCCAGTAGAAACAGCAG
     960       970        980        990       1000
. ..........||||||||| || ||||||||||||||||||||||||||
TTCTTTGATTACCTTCACAAGGTCATTGCAGCATCCAGTAGAAACAGCAG
     880       890        900        910        920


40          50          60          70          80
TTTCTGCAGGAGATATGGATAAAAGTCCTCGTGTATCAAAGCGACTTTCCA
 || |||||||| ||||||||||||||||||| |||||||||| ||||||||
.TTATGCAGGAGCTATGGATAAAAGTCCTCATGTATCAAAGTGACTTTCCA
      1010        1020        1030       1040       1050
||||||||||||||||| |||||||||||| ||||||||||||| |||||
TTTCTGCAGGAGATATGAATAAAAGTCCTCATGTATCAAAGCGAGTTTCCC
    930        940         950        960        970


90         100         110         120         130
CTTCTGAAACAAATAGCACAAGTCCCAAAAGGCCTACACTCAGTGATGG.A
||||||||||||||||||||| |||||||||||| || |||||||||||| |
CTTCTGAAACAAATAGCACAAATCCCAAAAGGCCTAAACTCAGTGATGG.A
       1060       1070        1080       1090       1100
||| |||||||||| ||| || ||||| |||||  || || ||||||||| |
CTTGTGAAACAAATCGCAGAAATCCCAGAAGGCCTAAACTCAGTGATGGCA
     980        990        1000       1010       1020
```

Figure 13b:

```
        150       160       170       180
GTTATAGGAGAACATTCTTTAGCTTTGTCCCTTTCAGGTACTTTTCATAC  <-- BCY11a

ATTATAGGAGAACATTCTTTAGCTTTGTGCCCTCCAGATATTTTTCAAAC  <-- BRY4a
      1110      1120      1130      1140      1150

.TTGTCGGGGAACATTCTTTAGGTTTCTCCCCGTCACGTTTTTTTCATAC  <-- OY11.1
    1030      1040      1050      1060      1070


190       200       210       220       230
  AGTGACTACTCAGCAAAGCACACCAGACAGTGTTCATACCTCATTAAGCC

  AGTGACTACTCAGCAAAGCACACCAGACAGTGTTCATACCTCATTAAGCC
        1160      1170      1180      1190      1200

  AGAGACCACTCAGCAAAGCACACCAGACCGTGTCCATACCTCACTAAGCC
      1080      1090      1100      1110      1120


        250       260       270       280
ATGTTCAGAGTGGAGAACCTTGTCCAACAGCTTTTCCAAAGGAACATGTT

ATGTTCAGAATGGAGAATCTTATCCAACACATTTTCTAAAGGACAATGTC
      1210      1220      1230      1240      1250

ATGTTCAGAATGGAGAACCTTGTCCAACACCTTTTCCAAAGGACAGTGTT
    1130      1140      1150      1160      1170


290       300       310       320       330
  CATTGCAAGCCTGTAAGTCCTTTAGGGGGAAATGTATTGACAAAAACTGA

  CATTGCAAACCTATAAGCCCTTTAGGGTCAAATGGACTGACAAAAATTGA
        1260      1270      1280      1290      1300

  CATTGCAAGCCTGTAAGACCTTTAGGGGAAAGTGGACGGACAAAAACTGA
      1180      1190      1200      1210      1220

        AAGCTTGTAAGACCTTTAGGGGAAAATGGACTGACAAAAACTGA  <-- GRY1a
                10        20        30        40
```

Figure 13c:

FBov*REED* primer

```
       350        360        370        380
CTTTCCAAGTGTATCAAGTCAAAATAAGTTTGCTGATCCCACAGAAGGCA   <-- BCY11a
CTTTCCAAGTGTAGCAAGTCAAAACAAGATTGTTGATCCCACAGAAGGCA   <-- BRY4a
      1310       1320       1330       1340       1350
TTTTCCAAGTTTGGCAAGTCCAAACAAAATTGGTGATCCCACAGAAGGAA   <-- OY11.1
    1230       1240       1250       1260       1270
TTTTCCAAGTTTAGCAAGTCCAAACAAGATTGTTGATCCCACAGAAGGAA   <-- GRY1a
    50         60         70         80         90

390        400        410        420        430
ATCTGATTGTGTTACTTCGTGACTTCTACTATGGACAGCATATAGGAGAT
ATCTGATTGTGTTACTCTGTGACTTCTACTATGGACAGCATATAGGAGAT
      1360       1370       1380       1390       1400
ATCTGATTGTGTTACTTCATGACTTCTACTATGGCGAGCATGGAGGAGTT
    1280       1290       1300       1310       1320
ATCTGATTGTGTTACTTCGTGACTTCTACTATGGAGAGCTTATAGGAGTT
   100        110        120        130        140

         450        460        470        480
GGG.CAGCCAGAACCAAAAACTCACACAGCCTTTAAATGCCACAGCTGCT
GGGGCAGCCAGAACCGAAAACTCACACAGCCTTTAAATGCCTCAGCTGCT
      1410       1420       1430       1440       1450
GGG.CAGCCAGAAGCGAAGACCCACACCGCGTTTAAATGCCTCAGCTGCT
     1330       1340       1350       1360       1370
GGG.CAGCCAGAACCGAAGACCCACACAGCGTTTAAATGCCTCAGCTGCT
     150        160        170        180        190

490        500        510        520        530
TGAAGGTTCTAAAAAATGTTGAGTTTATGAATCACATGAAGCACCATTTG
TGAAAGTTCTAAAAAATGTCAAGTTTATGAATCACATGAAGCACCATTTG
      1460       1470       1480       1490       1500
TGAAAGTTCTAAAAAATGTCAAGTTTATGAATCACATGAAGCACCATTTG
     1380       1390       1400       1410       1420
TGAAAGTTCTAAAAAATGTCAAGTTTATGAATCACATGAAGCGCCATTTG
     200        210        220        230        240
```

Figure 13d:

```
540        550        560        570        580
GAACTTGAGAGGCTGAGAGGTGACAGCTGGAAATAC.ACACCACCTGCCA  <-- BCY11a
GAGCTTGAGAGGCAGAGAGGTGACAGCTGGAAATACCACACCACCTGCCA  <-- BRY4a
      1510       1520       1530       1540       1550
GAACTTGAGAGGCAGAGAGGTGACAGCTGGAAAACCCACACCACCTGCCA  <-- OY11.1
     1430       1440       1450       1460       1470
GAACTTGAGAGGCAGAGAGGTGACAGCTGGAAAAACCACACCACCTGCCA  <-- GRY1a
     250        260        270        280        290

                                          USP1
590        600        610        620        630
GCACTGCCACCGCCAGTTTCCTACGCCCTTCCAGCTGCAGTGTCATATTG      BovREED (triplex region)
GCACTGCAACTGCCAGTTTCCTACTCCCTTCCAGCTGCAGTGTCACATTG
      1560       1570       1580       1590       1600
GCACTGCCTCCGCCAGTTTCCTACTCCCTTCCAGCTGCAGTGTCACATTG
     1480       1490       1500       1510       1520
GCACTGCCTCCGCCAGTTTCCTACTCCCTTCCAGCTGCAGTGTCACATTG
     300        310        320        330        340

640        650        660        670        680
AAAGTGTCCACACTTCCCAGGAGCCCTGCACAGTCTTGTAAAATCTGTGAA     BovREED3 (triplex region)
AAAGTGTCCATACTTCCCAGGAGCCCTCTGCAGTCT.GTAAAATCTGTGAA
      1610       1620       1630       1640       1650
AAAGTGTCCACACGGCCCAGGAGCCCTCCGCAGTCT.GTCACATCTGTGAG
     1530       1540       1550       1560       1570
AAAGTGTCCACACTCCCCAGGGGCCCTCCGCAGTCT.GTCAAATCTGTGAG
     350        360        370        380        390

690        700        710        720        730
TTGTCCTTTGAGGCAGATCAGATTCTCTTACAGCACATGAAAGATAATCA
TTGTCCTTTGAGACAGATCGGGTTCTCTAACAGCACATGAAAGATAATCA
      1660       1670       1680       1690       1700
TTGTCCTTTGAGACAGATCAGGTTCTCTTAGAGCACATGAAAGACAATCA
     1580       1590       1600       1610       1620
TTGTCCTTTGAGACAGATCAGGTTCTCTTAGAGCACATGAAAGACAATCA
     400        410        420        430        440
```

Figure 13e:

USP2

```
740       750       760       770       780
TAAGCCTGGTGAAATGCCCTATGTATGCCAGGTTTGCAGTTACAGATCAT  <-- BCY11a

TAAGCCTGGTGAAATGCCCTATGTATGCCAGGTTTGCAGTTACAGATCAT  <-- BRY4a
     1710      1720      1730      1740      1750

TAAGCCTGGTGAAATGCCCTATGTATGCCAGGTTTGCAGTTACAGATCAT  <-- OY11.a
     1630      1640      1650      1660      1670

TAAGCCTGGTGAAATGCCCTATGTATGCCAGGTTTGCAGTTACAGATCAT  <-- GRY1a
     450       460       470       480       490

790       800       810       820       830
CATTCTTTGCAGATGTGGATGCACATTTCAGAGCATACCATGGTAACACT

CATTCTTTGCAGATGTGGATGCACATTTCAGAGCATACCATGGTAACATT
     1760      1770      1780      1790      1800

CATTTTTTGCAGATGTGGATGCACATTTCAGAGCATACCATGGTAACACC
     1680      1690      1700      1710      1720

CATTTTTTGCAGATGTGGATGCACATTTCCGAGCATGCCATGGTAACACT
   500       510       520       530       540

840       850       860       870       880
AAGAATTTACTTTGCCCCTTTTGTCTCCAAATTTTTAAAACTGAAGCACT

AAGAGTTTACTTTGCCCCTTTTGTCTTCAAATGTTTAAAACTGCAAGACC
     1810      1820      1830      1840      1850

AAGAATTTACTTTGCCCGTTTTGTCTCAAAATTTTTCAAACTGCAACAGC
     1730      1740      1750      1760      1770

AAGAATTTACTTTGCCCGTTTTGTCTCAAAATTTTTAAGACTGCAACAGC
     550       560       570       580       590

890       900       910       920       930       940
ATACATATGTCATTATAAACAGCACTGGGCAAAGCGTTTTCACCAGTGTTC

ATACATATGTCATTATAGACAGCACTGGGAAAAG..TTTTCATCAGTGTTC
     1860      1870      1980      1890      1900

ATACAGACGTCATCATCGAGGGCACTGGGAAAAGAGTTTTCACCAGTGTTC
     1780      1790      1800      1810      1820

ATACAGACGTCATCATAGAGGGCACTGGGAAAAGAGTTTTCACCAGTGTTC
     600       610       620       630       640
```

Figure 13f:

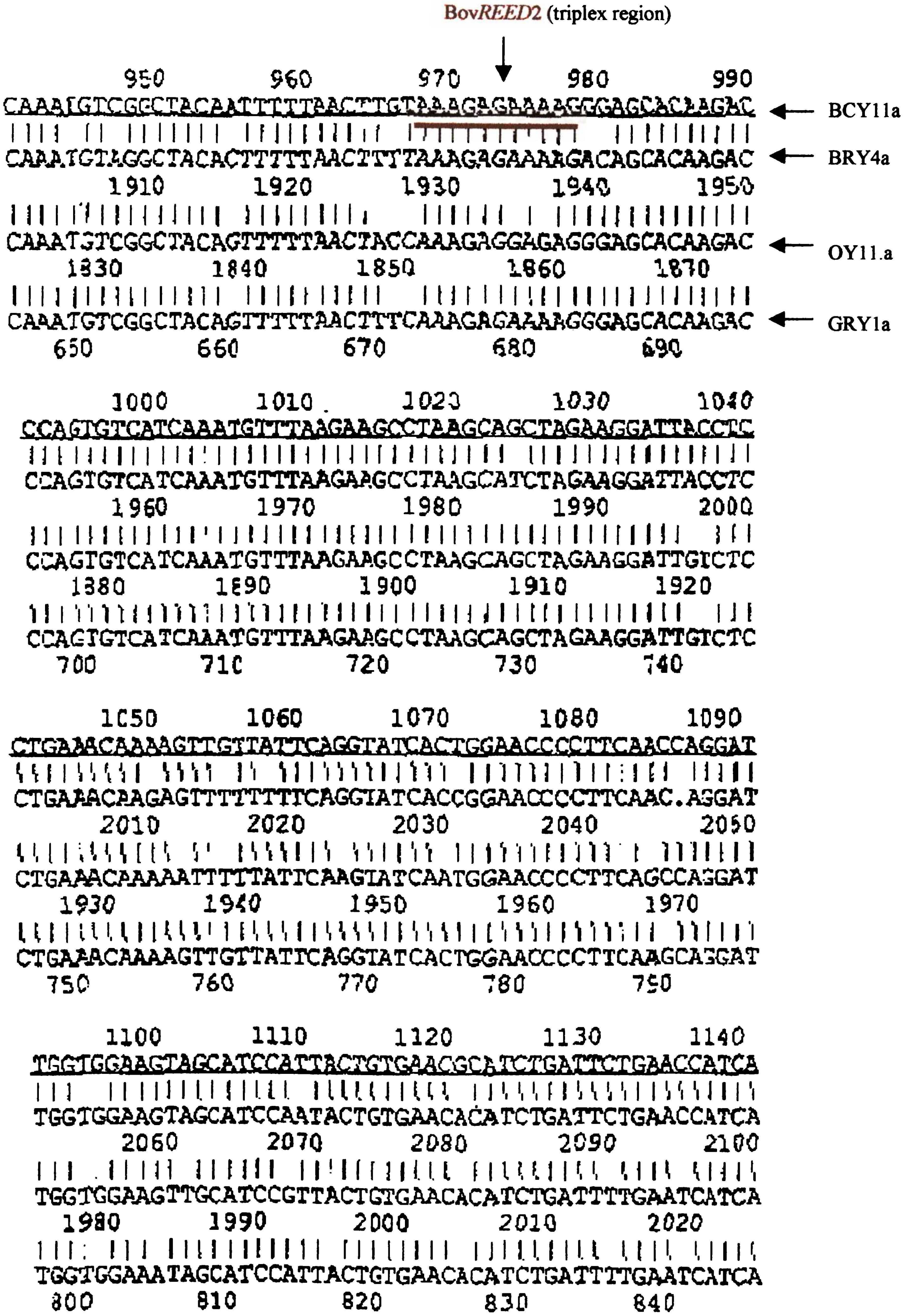

BovREED2 (triplex region)
950 960 970 980 990
CAAATGTCGGCTACAATTTTTAACTTGTAAAGAGAAAAGGGAGCACAAGAC
BCY11a
CAAATGTAGGCTACACTTTTTAACTTTTAAAGAGAAAAGACAGCACAAGAC
BRY4a
1910 1920 1930 1940 1950
CAAATGTCGGCTACAGTTTTTAACTACCAAAGAGGAGAGGGAGCACAAGAC
OY11.a
1830 1840 1850 1860 1870
CAAATGTCGGCTACAGTTTTTAACTTTCAAAGAGAAAAGGGAGCACAAGAC
GRY1a
650 660 670 680 690
1000 1010 1020 1030 1040
CCAGTGTCATCAAATGTTTAAGAAGCCTAAGCAGCTAGAAGGATTACCTC
CCAGTGTCATCAAATGTTTAAGAAGCCTAAGCATCTAGAAGGATTACCTC
1960 1970 1980 1990 2000
CCAGTGTCATCAAATGTTTAAGAAGCCTAAGCAGCTAGAAGGATTGTCTC
1880 1890 1900 1910 1920
CCAGTGTCATCAAATGTTTAAGAAGCCTAAGCAGCTAGAAGGATTGTCTC
700 710 720 730 740
1050 1060 1070 1080 1090
CTGAAACAAAAGTTGTTATTCAGGTATCACTGGAACCCCTTCAACCAGGAT
CTGAAACAAGAGTTTTTTTCAGGTATCACCGGAACCCCTTCAAC.AGGAT
2010 2020 2030 2040 2050
CTGAAACAAAAATTTTTATTCAAGTATCAATGGAACCCCTTCAGCCAGGAT
1930 1940 1950 1960 1970
CTGAAACAAAAGTTGTTATTCAGGTATCACTGGAACCCCTTCAAGCAGGAT
750 760 770 780 790
1100 1110 1120 1130 1140
TGGTGGAAGTAGCATCCATTACTGTGAACGCATCTGATTCTGAACCATCA
TGGTGGAAGTAGCATCCAATACTGTGAACACATCTGATTCTGAACCATCA
2060 2070 2080 2090 2100
TGGTGGAAGTTGCATCCGTTACTGTGAACACATCTGATTTTGAATCATCA
1980 1990 2000 2010 2020
TGGTGGAAATAGCATCCATTACTGTGAACACATCTGATTTTGAATCATCA
800 810 820 830 840

Figure 13g:

```
     1150       1160       1170
CCCCCCAAATCTAAAAGTAGGAGGTCAAAAAAA.................C<--- BCY11a
 | ||||||||||||||||||||||||| ||||||||              |
TCTCCCAAATCTAAAAGTAGGAGGCCAAAAAAAAAAAAAAAACCAAAAAAC<--- BRY4a
     2110       2120       2130       2140       2150
||||||||||||||||| |||||||||||||||||
CCCCCCAAATCTAAAAGGAGGAGGTCAAAAAAAGAAAAAATAATA......<--- OY11.1
  2030       2040       2050       2060
||| |||||||||||||||||||||||||||||||
CCCATCAAATCTAAAAGTAGGAGGTCAAAAAAAGAAAAAAAAA.......<--- GRY1a
 850       860       870       880

 1180       1190       1200       1210          1220
CCGGTTAAGTCTACTTTAAGTAAGTCTGAAGCAAGTATTT....CAAAGT
|| ||||| ||||||| ||||||||||||||||||||||||    ||||||
CCCGTTAATTCTACTTTCAGTAAGTCTGAAGCAAGTATTTCAGTCAAAGT
     2160       2170       2180       2190      2200
   ||||| ||| | || |||||||||||||||||||||||     ||||||
...GTTAATTCTGCCTTCAGTAAGTCTGAAGCAAGTATTTCAGTCAAAGT
 2070       2080       2090       2100       2110
   ||||| ||  |||| ||||||||||||||||||||||     ||||||
...GTTAATTCAGCTTTCAGTAAGTCTGAAGCAAGTATTGCAGTCAAAGT
 890       900       910       920       930

     1230       1240       1250       1260      1270
TATAAACTCCATTAAAAACAAAAAATTCAA.CTATAGCATTATTCAATAA
|| |||| ||||||||||||||||||||||   | ||| ||||||||||||
TAAAAACCCCATTAAAAACAAAAAATTCAAATTTTAGAATTATTCAATAA
      2210       2220       2230       2240      2250
|| |||  ||||||||||| ||||  ||||  ||||||||||||||||||
TAAAAATCCCATTAAAAACCAAAAGATCAACTTATAGCATTATTCAATAA
 2120       2130       2140       2150       2160
|| |||| ||||||||||||||||  ||||  ||||||||||||| |||
TAAAAACCCCATTAAAAACAAAAAGATCAAATTATAGCATTATTCAGTAA
 940       950       960       970       980

TAAA
||
TATCAAATATAGGGCAACCAATGTGTAATTTATGTGTTTTTGTTTTAAAT
      2260       2270       2280       2290      2300
||............ .... ... ............ .............
TATCAAATATAGGGAAACCGATGGGTAATTTATGTGATTTTGTTTTAAAT
 2170       2180       2190       2200       2210
||........... .. . ... ............ .............
TATCAAATATAGGGAAAGCGATGGGTAATTTATGTGATTTTGTTTTAAAT
 990       1000       1010       1020       1030
```

C) The Complete DNA Sequence of Bov*REP3*

LOCUS BTU39891 306 bp DNA MAM 29-NOV-1995
DEFINITION *Bos taurus* Y chromosome specific genomic sequence.
ACCESSION U39891
NID g1079750
SOURCE cow.
ORGANISM *Bos taurus*
REFERENCE 1 (bases 1 to 306)
FEATURES Location/Qualifiers
/source 1..306
/organism="Bos taurus"
/chromosome="Y"
Successful DNA triplex formation with Puv6a
Forward Primer (F4) sequence is from 197-213.
Reverse Primer (R3) sequence is from 275-296.
Homopurine DNA triplex region is underlined (239-248).

BASE COUNT 89 a 66 c 68 g 82 t 1 others
ORIGIN

1 gatctcctgg agaagggaat ggctatccaa ttcagtatgc ttgcctcgag aattccatgg
61 atagaggagc ctggggaact actatccatg ggtttgcaaa gggtcaaaca caaccggatg
121 actaacactt tcattttcat atgcttctaa agattgttca aatctcagaa tcacctggga
181 cattattaaa catagagatt ccagggctgc atatcatacc taagaaacta aattgcctga
241 ggagggggcc tgggcatctg aaagttcaat gttcgcctta cttgattctt ccctcgsata
301 gcacag

2. The Complete DNA Sequence of *Bos taurus* SRY Gene

LOCUS BTU15569 1800 bp DNA MAM
18-OCT-1994
DEFINITION Bos taurus sex determining region Y protein (SRY) gene,
ACCESSION U15569 L36821
NID g558463
ORGANISM *Bos taurus*
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
FEATURES Location/Qualifiers
source 1..1800
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="bSRYBGalB.6; bSRYD.T3; bSRYJ.Adapt"
/chromosome="Y"
/sex="male"
Successful DNA triplex formation with Puv12a.
Forward Primer (F5) sequence is from 926-950
Reverse Primer (R4) sequence is from 1101-1125
Homopurine DNA triplex region is underlined (1021-1060)

"BASE COUNT 549 a 365 c 357 g 529 t
1 aagcttttt ttttgttttg ttctatttag ttttgatcat ctgttaggta aaagtgcaga
61 agagagtgat agatattact tctgcacagc ctacgcatct aagacaccac acaccaaccc
121 ccccctccca cttttcctac tcccaaccg tcagaacaaa aaaataaaaa tcagtccctt
181 tgaagtgtta tattaacacc agggaactgc ttgggtacca aggttatgtg ttttttcttt
241 taatggaaca ggtttttaat ctaattttag ttgtatctga gattgcctgt taaatatgtg
301 ttagtatata ttagcattct gaaagtcgtt agcacagata atagaattac cagttattag
361 ctactggaat acgtacatag atttttcagc tgcactttga gcctaagtgg agaagcaggg
421 ttagtgatta gcatggattg ggtatctgtg tattcaaata tataacagga ttgggctgtt
481 tttcatcttt ttttgtttaa gataacattc tcttacattc ataaccgtag acaatttgct
541 agtatgtctg ctgcaccttc atcctttgaa ataagtatat gaaaataact tcataatgac
601 actttttgta tttttaagca ggtgttagca cattcacaaa ttctgattag atgtaaacaa
661 agaagaaagc agagccttaa tatcctgtta agtagctttg cttgagaaag agtaggttga
721 tgggtttggg ctgactgcca ggacgtattg aggggaggta ttgggggcgg agaaataaat
781 atttcactgt atatattgca ctaagtcagt ctgtggtaag aacaacttat gaatagcacc
841 ataattttta gaacgcttac accgcatatt acttcctccc cttttaaaca gtgcagtcgt
901 atgcttctgc tatgttcaga gtattgaacg acgatgttta cagtccagct gtggtacagc
961 aacaaactac tctcgctttt aggaaagact cttccttgtg cacagacagt catagcgcaa
1021 <u>atgatcagtg tgaaagggga gaacatgtta gggagagcag</u> ccaggaccac gtcaagcgac
1081 ccatgaacgc cttcattgtg tggtctcgtg aacgaagacg aaaggtggct ctagagaatc
1141 ccaaaatgaa aaactcagac atcagcaagc agctgggata tgagtggaaa aggcttacag
1201 atgctgaaaa gcgcccattc tttgaggagg cacagagact actagccata caccgagaca
1261 aatacccggg ctataaatat cgacctcgtc ggagagccaa gaggccacag aaatcgcttc
1321 ctgcagactc ttcaatacta tgcaacccga tgcatgtaga gacattgcac cccttcacat
1381 acagggatgg ttgtgccaag accacatact cacaaatgga aagccaatta agccggtcac
1441 agtccgtgat cataaccaat tcactcctgc aaaaggagca tcacagcagc tggacaagcc
1501 tgggccacaa taaggtaaca ttggctacac ggatttcggc ggactttccc tgtaacaaaa
1561 gcttagagcc tggactttct tgtgcttatt ttcaatattg acttccttac tctcgctaac
1621 aaaggcgctc tttatctcaa ttttactaca gtttcaacct gcgacttaat tttataagaa
1681 gcccaataag taagtacatt taacatgtaa agaattcaga ctttccaata taactggtcc
1741 tctgttaatc agttctttct ataagagtac ttttttgtaa aaaattatct taacagcacc

3. The Complete DNA Sequence of *Bos taurus* TSPY Gene

LOCUS BTTSPYA 3179 bp DNA MAM 03-JAN-1994
DEFINITION *B. taurus* TSPY gene.
ACCESSION X74028
NID g439124
KEYWORDS TSPY gene.
SOURCE cow.
FEATURES Location/Qualifiers
source 1..3179
/organism="Bos taurus"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="p0201"
/clone_lib="bovine genomic phage"
/chromosome="Y"
Forward Primer (F-TSPY) sequence is from 1981-2001.
Reverse Primer (R-TSPY) sequence is from 2460-2480.
Homopurine DNA triplex region is underlined (2190-2207).

BASE COUNT 675 a 955 c 963 g 586 t

```
   1 ggatcctgcc cttgttcatc ccgacagaca ctcccgaggc cacagcacag tccctgccat
  61 gacccaggct gggtcctgcc tttgttcaag cccccgacaa acactcttga gcccacatct
 121 cggtccatgc aactgaatcg ggtcggccct gcccatgccc aagctctttg tagaacgcat
 181 atatcgtgtg tcaggactat aggatggctg gtttagcttg atcctcttgc tgctccaggc
 241 cccacaggct tgggcacagg gcaggctggg atgtgttctc acatggagaa cagaaccact
 301 ttgaaagggc ctccataggg tcgatggcag ccgtgcatgt ctccgacacc ggtgagtcgt
 361 ggaacttaac ggggagctgt ccttagcgtc tccctggccg gcattcgcac aaggttcagg
 421 gggggtccct catgccagca gagcgcgctc gtttctccct ctgcgcacag tgtggccacg
 481 gcacctcagg cttgtgcagc tttcacgtcg gcccagcgag ggagggaggg tcgcccgtcc
 541 ccttctgcgc ttcagcagtg accgcccact gttggctttc aaggtcgaga aagcactctc
 601 cgaacgtgtg gactaaaacc ctctgaactc ccaccggtga aactcactca ggacgagaag
 661 ggcttgcttc acgccaagga taagcgtctc ttctgagcgc tggtgtccac agagttgcag
 721 acggcaacag cactacttag agagacggag acacacagac accagcgcct gctacacacg
 781 cggaaggatt gcccaacaca gtgctatgcc agaagcaaag atacagagac acagagcgag
 841 cgctgtacac taaggcgctg gtgttcccga acagtggaga tccactgtag cactgtaggt
 901 ggggttcttc cctactcccc tagcatccgg acagcagcag gctctctgtc ccttgaagaa
 961 aagtatccag ccccttggca cccctgcgta tcccttggat actcagggag agctctcagt
1021 tattgccagt cccaggggac acaggagtgt gcggttcgga tggcggagca ggaacccgtg
1081 cttggaattt ccagagccag gtcaatgaca gcgcctgggc tagcccggcc acatgcaaaa
1141 ccctgtccct ccctggagca gaatctggcc acgtgggggc gctggagcag cacaggaacc
1201 ccggaagatg cctccggtct gggggcgtgt ccataagagt gccgcccagt acctgtcggc
1261 cattggtcta gatttgggcc agcagtctgt gcatggctca cagcctaggc ccccgcctgt
1321 tcgccggatc tcctagggcc cgttcacccc gtcctgcgtt tctggcaccc ctttcttcgg
1381 cctcacgcca acaccccac acaccagccc ccgccttctg acaatccccc tcactccccg
1441 cccccaaaag agctagtcgg tcccacaggt ttactgcagc ctgggccgcc attacgcccc
1501 gacttgctgt gctcttatgt cgcgtccctt cgcctctgcc ccagctcggg gtcacccgca
1561 aggccaagag gaacgtgaga ggcggtcaga ggaaggcggg agcgtcccgg gaccccgcac
1621 cttccaagtt gtgagcccag gtactcaggg ggcagggccg cccccacccc ccctccccca
1681 ccccgcagcc tgccgtcgtg atgcagtgcg cacgtgggac cttgtcacca gcagttgtca
1741 cgccaggcca agaggcaccc tcttcagggt ggaggcggtg gaggagggca aggcctggtg
1801 gacggagacg tggcgggaat cgggcgggag ttccagctgc ttcgaggaag acatcgtgga
1861 ggaggtggag gttgtggcgg atgaggagca ggaacagcgg ccctcccagg agctagagga
1921 gaagacggtg gaggagcagg ccaggaaagc ccggaggccc ttgtgagcgc caggagctgg
1981 atgccctgca ggctgtggcc gccctgcagg tagaactgag ctctgagcgt gagcaaaacg
```

2041 cagggcttac gttcagttca tgcgcaagaa ccatcagagg aggaagcgtc acttggctcg
2101 gaggagcgcc atcatccagg gcatccctgg cttctgggcc aaagctgtat cctttgcgct
2161 gctccttgag gtccgctgct caggaggacg aggaaggga gaaggggcag gagcaggcgg
2221 agggggtgga ggcgagggca caggaagggc gccggaggcc actgaaggaa atgcggtcct
2281 gggcaattgg caggctccaa agcacagccg agtagggcct gggcagggcc acaggtggac
2341 gtgtcgcagc catgcctttg agtgtctcct tcaggcctgc atctgaggag cgcagccccc
2401 gaggcctcta agtcaacatt gaccacagac ggtacacaag agcagcgttc gcagacagat
2461 attggtgtaa atgtgggtga tcccggcaca ttgctgggac atgtgagtca ttcacacaca
2521 ccgcgcgcgc acacacacac acacacacac acgcacacac acacacgtac acatattcac
2581 acacacacat acccaagaag ccctttactt tcagacacgg tttgtcacaa ggactgcaag
2641 tctgaccgta taagcagtga catgccacta cccagcatac acaggtctgg aggagcagag
2701 tagcggtcag aaacagcccc tgcttgggga atgcagatgc acaccatccc aaggggccca
2761 gagccatgag cgtgtctgat ggtatggcag gtcctgagga tatcagcagt ggccgggcaa
2821 gaggcagtgc ccgccaagct ccttgcagat gcgtgttcca gtatccttgt tccaccccag
2881 actagggctg atgctcctgg ctcttcttgg ccggcatgtg gcctgaatgg ctccttgacc
2941 taaagcagat tatgagccac cctcaagtct ccgtcctgat cagcgaccaa gatcaagact
3001 ttctcggcta catgatggac ttgaaggtca gtgagggaga ctgaggaagg gtgggtatgg
3061 gacctggcgg ggagtgggag gttcgtgttt gccatgttcc tgtgctgtga gactttggag
3121 aggcatcggg gaagctgatc atgcctgcct tggaggcaag ctgaggcccc tagaagctt

4. The Complete DNA Sequence of Shuttle Vector pSP189

S XXU14594 4952 bp DNA SYN 24-MAY-1995
DEFINITION Shuttle vector pSP189, complete sequence.
ACCESSION U14594
NID g550455
ORGANISM Shuttle vector pSP189, artificial sequence; shuttle vectors.
FEATURES organism="Shuttle vector pSP189"
plasmid="pSP189"
lab_host="Escherichia coli"
tRNA 27..111
/note="5' end of mature tRNA (plasmid position 27) is position 99 in the traditional tRNA nomenclature. Bp 1-98, the pre-tRNA region, have been deleted. Promoter is a cryptic promoter in b-lactamase gene. codon recognized: TAG; aa: Tyr"
/citation=[1]
/function="amber codon-tyrosine suppressor"
misc_feature 183..190
/note="DO NOT CLONE THIS PLASMID. Cloning eliminates the sequence variation in the plasmid population"
/function="determination of independence of mutations"
/product="signature (variable) sequence"

BASE COUNT 1343 a 1175 c 974 g 1452 t 8 others
ORIGIN

```
   1 gaattcgaga gccctgctcg agctgtggtg gggttcccga gcggccaaag ggagcagact
  61 ctaaatctgc cgtcatcgac ttcgaaggtt cgaatccttc ccccaccacc acggccgaaa
 121 ttcggtaccc ggatccttag cgaaagctaa gattttttt acgcgtgagc tcgactgact
 181 ccnnnnnnnn gagctcaatt cggtcgaggt cgggccgcgt tgctggcgtt tttccatagg
 241 ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg
 301 acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt
 361 ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt
 421 tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc
 481 tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt
 541 gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt
 601 agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc
 661 tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa
 721 agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt
 781 tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct
 841 acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta
 901 tcaaaaagga tcttcaccta gatccagatc cagacatgat aagatacatt gatgagtttg
 961 gacaaaccac aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta
1021 ttgctttatt tgtaaccatt ataagctgca ataaacaagt taacaacaac aattgcattc
1081 attttatgtt tcaggttcag ggggaggtgt gggaggtttt ttaaagcaag taaaacctct
1141 acaaatgtgg tatggctgat tatgatcatg aacagactgt gaggactgag gggcctgaaa
1201 tgagccttgg gactgtgaat caatgcctgt ttcatgccct gagtcttcca tgttcttctc
1261 cccaccatct tcatttttat cagcattttc ctggctgtct tcatcatcat catcactgtt
1321 tcttagccaa tctaaaactc caattcccat agccacatta aacttcattt tttgatacac
1381 tgacaaacta aactctttgt ccaatctctc tttccactcc acaattctgc tctgaatact
1441 ttgagcaaac tcagccacag gtctgtacca aattaacata agaagcaaag caatgccact
1501 ttgaattatt ctcttttcta acaaaaactc actgcgttcc aggcaatgct ttaaataatc
1561 tttgggccta aaatctattt gttttacaaa tctggcctgc agtgttttag gcacactgta
1621 ctcattcatg gtgactattc caggggggaaa tatttgagtt cttttattta ggtgtttctt
1681 ttctaagttt accttaacac tgccatccaa ataatccctt aaattgtcca ggttattaat
```

1741 tccctgacct gaaggcaaat ctctggactc ccctccagtg ccctttacat cctcaaaaac
1801 tactaaaaac tggtcaatag ctactcctag ctcaaagttc agcctgtcca agggcaaatt
1861 aacatttaaa gctttccccc cacataattc aagcaaagca gctgctaatg tagttttacc
1921 actatcaatt ggtcctttaa acagccagta tctttttta ggaatgttgt acaccatgca
1981 ttttaaaaag tcatacacca ctgaatccat tttgggcaac aaacagtgta gccaagcaac
2041 tccagccatc cattcttcta tgtcagcaga gcctgtagaa ccaaacatta tatccatcct
2101 atccaaaaga tcattaaatc tgtttgttaa catttgttct ctagttaatt gtaggctatc
2161 aacccgcttt ttagctaaaa cagtatcaac agcctgttgg catatggttt tttggttttt
2221 gctgtcagca aatatagcag catttgcata atgcttttca tggtacttat agtggctggg
2281 ctgttctttt ttaatacatt ttaaacacat ttcaaaactg tactgaaatt ccaagtacat
2341 cccaagcaat aacaacacat catcacattt tgtttccatt gcatactctg ttacaagctt
2401 ccaggacact tgtttagttt cctctgcttc ttctggatta aaatcatgct cctttaaccc
2461 acctggcaaa ctttcctcaa taacagaaaa tggatctcta gtcaaggcac tatacatcaa
2521 atattcctta ttaacccctt tacaaattaa aaagctaaag gtacacaatt tttgagcata
2581 gttattaata gcagacactc tatgcctgtg tggagtaaga aaaaacagta tgttatgatt
2641 ataactgtta tgcctactta taaaggttac agaatatttt tccataattt tcttgtatag
2701 cagtgcagct ttttcctttg tggtgtaaat agcaaagcaa gcaagagttc tattactaaa
2761 cacagcatga ctcaaaaaac ttagcaattc tgaaggaaag tccttggggt cttctacctt
2821 tctcttcttt tttggaggag tagaatgttg agagtcagca gtagcctcat catcactaga
2881 tggcatttct tctgagcaaa acaggttttc ctcattaaag gcattccacc actgctccca
2941 ttcatcagtt ccataggttg gaatctaaaa tacacaaaca attagaatca gtagtttaac
3001 acattataca cttaaaaatt ttatatttac cttagagctt taaatctctg taggtagttt
3061 gtccaattat gtcacaccac agaagtaagg ttccttcaca aagatcaagt ccaaaccaca
3121 ttctaaagca atcgaagcag tagcaatcaa cccacacaag tggatctttc ctgtataatt
3181 ttctattttc atgcttcatc ctcagtaagc acagcaagca tatgcagtta gcagacattt
3241 tctttgcaca ctcaggccat tgtttgcagt acattgcatc aacaccagga tttaaggaag
3301 aagcaaatac ctcagttgca tcccagaagc ctccaaagtc aggttgatga gcatatttta
3361 ctccatcttc catttcttg tacagagtat tcattttctt cattttttct tcatctcctc
3421 ctttatcagg atgaaactcc ttgcattttt ttaaatatgc ctttctcatc agaggaatat
3481 tcccccaggc actcctttca agacctagaa ggtccattag ctgcaaagat tcctctctgt
3541 ttaaaacttt atccatcttt gcaaagcttt ttgcaaaagc ctaggcctcc aaaaaagcct
3601 cctcactact tctggaatag ctcagaggcc gaggcggcct cggcctctgc ataaataaaa
3661 aaaattagtc agccatgggg cggagaatgg gcggaactgg gcggagttag gggcgggatg
3721 ggcggagtta ggggcgggac tatggttgct gactaattga gatgcatgct ttgcatactt
3781 ctgcctgctg gggagcctgg ggactttcca cacctggttg ctgactaatt gagatgcatg
3841 ctttgcatac ttctgcctgc tggggagcct ggggactttc cacaccctaa ctgacacaca
3901 ttccacagct ggttctttcc gcctcagaag gtacctaacc aagttcctct ttcagaggtt
3961 atttcaggcc atggtgctgc gccgggatct tttaaattaa aaatgaagtt ttaaatcaat
4021 ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc
4081 tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat
4141 aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc
4201 acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag
4261 aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag
4321 agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgctg caggcatcgt
4381 ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg
4441 agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt
4501 tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc
4561 tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc
4621 attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa cacgggataa
4681 taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg
4741 aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc
4801 caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag
4861 gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt
4921 cctttttcaa tattattgaa gcatttatca gg

5. The Complete DNA Sequence of pBluescript II SK$^+$ Vector

LOCUS ARBL2SKP 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION pBluescript II SK(+) vector DNA, phagemid excised from lambda
ACCESSION X52328
BASE COUNT 746 a 739 c 754 g 722 t
ORIGIN

```
   1 ctaaattgta agcgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc
  61 atttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga
 121 gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc
 181 caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc
 241 ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaagggag
 301 cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa
 361 agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac
 421 cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg
 481 caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg
 541 gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg
 601 taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tgggtaccgg
 661 gccccccctc gaggtcgacg gtatcgataa gcttgatatc gaattcctgc agcccggggg
 721 atccactagt tctagagcgg ccgccaccgc ggtggagctc cagcttttgt tccctttagt
 781 gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt
 841 atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctggggtg
 901 cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct ttccagtcgg
 961 gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc
1021 gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc
1081 ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata
1141 acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg
1201 cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct
1261 caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa
1321 gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc
1381 tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt
1441 aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg
1501 ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg
1561 cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct
1621 tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc
1681 tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg
1741 ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc
1801 aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt
1861 aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa
1921 aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat
1981 gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct
2041 gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg
2101 caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag
2161 ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta
2221 attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg
2281 ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg
2341 gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct
2401 ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta
2461 tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg
2521 gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc
2581 cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg
2641 gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga
2701 tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg
2761 ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg cacggaaat
```

2821 gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc
2881 tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca
2941 catttccccg aaaagtgcca c

Unique Restriction Enzyme Sites:

Restriction enzyme sites		<--	Position.	-->
Dra III	cacnnn/gtg	221	222	2740
BsaA I	yac/gtr	224	225	2737
Nae I	gcc/ggc	327	328	2634
Asp718	g/gtacc	652	653	2309
Kpn I	ggtac/c	652	653	2309
Apa I	gggcc/c	658	659	2303
Bsp120 I	g/ggccc	658	659	2303
EcoO109 I	rg/gnccy	658	659	2303
PaeR7 I	c/tcgag	667	668	2294
Xho I	c/tcgag	667	668	2294
Acc I	gt/mkac	673	674	2288
Hinc II	gty/rac	673	674	2288
Sal I	g/tcgac	673	674	2288
Cla I	at/cgat	682	683	2279
Hind III	a/agctt	688	689	2273
EcoR V	gat/atc	694	695	2267
EcoR I	g/aattc	700	701	2261
Pst I	ctgca/g	706	707	2255
Sma I	ccc/ggg	712	713	2249
Xma I	c/ccggg	712	713	2249
BamH I	g/gatcc	718	719	2243
Spe I	a/ctagt	724	725	2237
Xba I	t/ctaga	730	731	2231
Not I	gc/ggccgc	736	737	2225
Eag I	c/ggccg	737	738	2224
BstX I	ccannnnn/ntgg	743	744	2218
Dsa I	c/crygg	746	747	2215
Sac II	ccgc/gg	746	747	2215
Ecl136 I	gag/ctc	754	755	2207
Sac I	gagct/c	754	755	2207
Afl III	a/crygt	1152	1153	1809
Nsp I	rcatg/y	1152	1153	1809
Nsp7524 I	r/catgy	1152	1153	1809
NspC I	rcatg/y	1152	1153	1809
AlwN I	cagnnn/ctg	1563	1564	1398
Bsa I	ggtctc 1/5	2112	2113	849
Sca I	agt/act	2523	2524	438
Aha II	gr/cgyc	2581	2582	380
Xmn I	gaann/nnttc	2640	2641	321

Appendix IV

DNA TRIPLEX SEQUENCES

	DNA TRIPLEX SEQUENCE
Puv 1	3'GGAGAAAAAGGAAGGGAG≈Pso5'
Puv 2	5'Pso≈CCTCTTTTTCTTTCCCTC3'
Puv 3	5'Pso≈CCTCTTTTTC3'
Puv 4	3'GGAGAAAAAG≈Pso 5'
Bdu*GROBET**	5'**C**GGAGAAAAAG**C**AAGGGAG**T**TCAAGATATACATCTACAT3'
Bpr*GROBET*	3'G**C**CTCTTTTT**C**GTTCCCT**C**AAGTTCTATATGTAGATGTA5'
Puv 7	3'AAAGAGAAAAAGGGAG≈Pso5'
Bdu*REED*2*	5'AAAGAGAAAAAGGGAG**C**ACAAGACC3'
Bpr*REED*2	3'TTTCTCTTTTTCCCT**C**GTGTTCTGG5'
Puv 11	3'AGGGGTAGAGGGAAAGGGG≈Pso5'
Puv 9	3'AGGGAAAGGGG≈Pso5'
Puv8	3'ATTGAAAGTGT≈Pso5'
Bdu*REED*3*	5'GCAGTGTCATATTGAAAGTG**T**CCA3'
Bpr*REED*3	3'CGTCACAGTATAACTTTCA**C**AGGT5'
TFO6	3'GAGGAGGGGG5'
Puv 5	5'Pso≈GAGGAGGGGG3'
Puv 6b	3'GAGGAGGGGGTTGGGG≈Pso5'
Puv 6c	3'GTGGTGGGGGTTGGGGTT≈Pso5'
BDu*REP3**	5'GAGGAGGGGG**C**CTGGG**CAT**CTGAAAGTTCAATGTTCGC3'
BPr*REP3*	3'CTCCTCCCC**C**GGACC**C**G**T**AGACTTTCAAGTTACAAGCG5'
TFO 12	3'GAAAGGGGAGAA5'
Puv 12b	3'ATGTTAGGGAGAG≈Pso5'
Puv 12c	3'GAAAGGGGAGAATAGGGGAGGGAGAG≈Pso5'
Puv 12d	5'Pso≈CTTTCCCCTCTT3'
Bdu*SRY**	5'ATGATCAGTG**T**GAAAGGGGAGAA**C**ATGTTAGGGAGAG**C**AG3'
Bpr*SRY*	3'TACTAGTCACA**C**TTTCCCCTCT**T**GTACAATCCCTCT**C**GTC5'
TFO 15	5'AAAGGAGGAGGGAAGG3'
Puv 15b	5'Pso≈GGTGGGTTGG3'
Bdu*REED*1*	3'CGGT**C**AAAGGA**T**GAGGGAAGGTCGACGTCACAGTGTAACTTTCACAGGT5'
Bpr*REED*1	5'GCCAG**T**TTCC**T**ACTCCCTTCCAGCTGCAGTGTCACATTGTAAGTGTCCA3'

Table A.1: A schematic representation of unsuccessful DNA triplexes formed between double stranded bovine Y chromosome targeted genes and psoralen-conjugated oligonucleotides. The strand radioactively labelled in the mobility shift assay is marked by an asterisk (*) and the 10 nt reverse primer is underlined. The psoralen intercalation site(s) are shown in bold lettering. Note that B = Bovine, Pr = Primer, Du = Duplex (Coding Strand), ≈Pso = Psoralen, TFO = unmodified triplex forming oligonucleotide.

	DNA TRIPLEX SEQUENCE
Puv 6a BDu*REP3** BPr*REP3*	3'GAGGAGGGGG≈Pso5' 5'GAGGAGGGGG**C**CTGGGCATCTGAAAGTTCAATGTTCGC3' 3'CTCCTCCCC**C**GGACCCGTAGACTTTCAAGTTACAAGCG5'
Puv 14 BDu*TSPY** BPr *TSPY*	3'GAGGAAGGGGAGAAGGGG≈Pso5' 5'CAGGAGGAGGAGGAGGGGAGAAGGGGCAGGAGCAGGCGA3' 3'GTCCTCCTCCTCCTCCC**CT**CTTCCCCGTCCTCGTCCGCT5'
Puv 12a Bdu *SRY** Bpr *SRY*	3'GAAAGGGGAGAA≈Pso5' 5'ATGATCAGTGTGAAAGGGGAGAA**C**ATGTTAGGGAGAGCAG3' 3'TACTAGTCACACTTTCCCCTCT**T**GTACAATCCCTCTCGTC5'
Puv 15a Bdu*REED*1* Bpr*REED*1	5'Pso≈AAAGGAGGAGGGAAGG3' 3'CGGT**C**AAAGGATGAGGGAAGGTCGACGTCACAGTGTAATTCACAGGT5' 5'GCCAG**T**TTCCTACTCCCTTCCAGCTGCAGTGTCACATTAAGTGTCCA3'
PsoAGT17 *Sup*F-GDu* *Sup*F-GPr	3'TGGTGGTGGGGGAAGGA≈Pso5' 5'TGGTGGTGGGGGAAGGA**T**TCGAAC3' 3'ACCACCACCCCCTTCC**T**AAGCTTG5'
PsoAG10 *Sup*F-WDu* *Sup*F-WPr	3' GGGGGAAGGA≈Pso5' 5'CCGGGTACCGAATTTCGGCCGTGGTGGTGGGGGAAGGA3' 3'GGCCCATGGCTTAAAGCCGGCACCACCACCCCCTTCC**T**5'

Table A.2: A schematic representation of successful DNA triplexes formed between double stranded bovine Y chromosome targeted genes and psoralen-conjugated oligonucleotides. The strand radioactively labelled in the mobility shift assay is marked by an asterisk (*) and the 10 nt reverse primer is underlined. The psoralen intercalation site(s) are shown in bold lettering. Note that B = Bovine, Pr = Primer, Du = Duplex (Coding Strand), ≈Pso = Psoralen, G = Gasparro, W = Wang.

Appendix V

OLIGONUCLEOTIDE PROBES

PROBE NAME	OLIGONUCLEOTIDE SEQUENCE
49Bov*REED*	5'TGGACACTTTCAATGTGACACTGCAGCTGGAAGGGAGTAGGAAACTGGC3'
30Bov*REED*	5'GTGACACTGCAGCTGGAAGGGAGTAGGAAA3'
16Bov*REED*	5' GGAAGGGAGTAGGAAA3'
38Bov*REP3*	5'GAGGAGGGGGCCTGGGCATCTGAAAGTTCAATGTTCGC3'
30Bov*REP3*	5'CCTAAGAAACTAAATTGCCTGAGGAGGGGG3'
16Bov*REP3*	5'TTGCCTGAGGAGGGGG3'
16Bov*SRY*	5'TGAAAGGGGAGAACAT3'
USP1	5'CCCTTCCAGCTGCAGTGTCA3'

Table A3: A table showing the DNA sequence of the radioactive oligonucleotide probes that were used for Southern blot hybridisations. The DNA triplex region is underlined.

Appendix VI

DNA QUANTIFICATION

The concentration of Plasmid DNA, bovine genomic DNA and/or PCR products were measured using:

1. Spectrophotometry

Nucleic acids have a maximum absorbency at 260 nm in the UV range. Therefore the following data can be used to determine the concentration of DNA or RNA.

OD = 1.0, 50 μg/mL dsDNA
OD = 1.0, 40 μg/mL ssDNA or RNA

The A_{260}/A_{280} ratio is also useful to estimate the purity of the DNA preparation.

$A_{260}/A_{280} = 1.8$, sample is pure
$A_{260}/A_{280} > 1.8$, sample contains impurities
$A_{260}/A_{280} < 1.8$, sample contains protein

All spectrophotometer reading were measured in quartz cuvettes in the Shimadzu UV-250 Spectrophotometer.

2. Agarose Gel Electrophoresis

Quantification of DNA can also be determined by comparing the DNA band of interest against a DNA molecular mass Ladder following agarose gel electrophoresis. For PCR products, a 4 μL sample was combined with 1 μL 6X gel loading dye (Type III) and 4 μL was loaded onto a 2% agarose gel containing 0.5 μg/mL EtBr.

The samples were electrophoresed in the Horizon 11.14 GibcoBRL Horizontal Gel Electrophoresis Apparatus (Life Technologies, USA) at 100 volts, 1 hr at RT alongside a 4 μL volume of Low and/or DNA Molecular Mass Ladder (Table 6.4). The gel was visualised by UV illumination and photographed using *Scion Image*, an image processing and analysis program on the IBM computer.

The same procedure was used to quantify plasmid DNA except that the DNA was initially digested with restriction enzymes. In general, a 20 μL restriction digest was prepared containing 1X RE buffer, 10 units of one or two restriction enzymes, plasmid DNA and dH_20. Controls containing no restriction enzyme were also set up. The restriction digest was incubated for a certain time and temperature, usually at 37°C for 1 hour. The sample was then diluted with 6X gel loading dye as previously described and electrophoresed on a 1-2% agarose gel against DNA Molecular Mass Ladders.

MOLECULAR WEIGHT MARKER	SIZE (bp)
1 Kb DNA Ladder	12,216, 11,198, 10,180, 9162, 8144, 7126, 6108, 5090, 4072, 3054, 2036, 1636, 1018, 506
100 bp DNA Ladder	2072, 1500, 1400, 1300, 1200, 1100. 1000, 900, 800, 700, 600, 400, 300, 200, 100
High DNA Mass Ladder	1000 (20 ng), 2000 (40 ng), 3000 (60 ng), 4000 (80 ng), 6000 (120 ng), 10000 (200 ng)
Low DNA Mass Ladder	100 (10 ng), 200 (20 ng), 400 (40 ng), 800 (80 ng), 1200 (120 ng), 2000 (200 ng)

Table A4: DNA Molecular Weight Markers (Life Technologies, USA) for Agarose Gel Electrophoresis. In the case of the 1 Kb and 100 bp DNA ladder, a 10 μL volume containing 0.5 μg was loaded.

Appendix VII

UV DOSE MEASUREMENT

A variety of parameters such as the radius, length and output of the fluorescent UV black tube light bulb (Osram L18W/73) were measured to calculate the UV dose.

1. Radius of UV lamp (1.3 cm)
   Sample distance from the UV lamp = 3 cm
   $\therefore$ Radius (r) = 3.4

2. Length of UV lamp = 59 cm (l)
   $\therefore$ Circumference: $2\pi r = 27$
   $\therefore$ Area: $27 \text{ X } 59 = 1594 \text{ cm}^2$

3. Total UV output is 3500 mW
   $\therefore 3500 \text{ mW}/1594 \text{ cm}^2 = 2.2 \text{ mW/cm}^2$
   $\therefore$ After 30 min (1800 sec): $2.2 \text{ mW/ cm}^2 \text{ X } 1800 \text{ sec} = 39600 \text{ mW/sec/cm}^2$
   $= 39600 \text{ mJ/cm}^2$
   $= 3.96 \text{ J/ cm}^2$

4. Corrections for loss of energy:

   Estimate that 80% of output was in the 350-380 nm range (0.8)

   Estimate that the Mylar sheet absorbency was 10% (The Mylar sheet was replaced frequently to avoid changes of UV irradiance due to its photocomposition). $\therefore$ Mylar sheet transmissibility is 90% (0.9).

   UV Output: 4°C = 55 (measured by a UV dosimeter)
   20°C = 78

   $\therefore$ Final output after experimental corrections:
   $3.96 \text{ J/ cm}^2 \text{X } 0.8 \text{ X } 0.9 \text{ X } 55/78 = 2 \text{ J/cm}^2$

$\therefore$ UV dose = 2 J/cm^2 after 30 minutes of UV irradiation (350 – 380 nm range).

BIBLIOGRAPHY

Affara, N.A., Chalmers, I.J., Ferguson-Smith, A. (1993). Analysis of the *SRY* gene in 22 sex reversed XY females identifies four new point mutations in the conserved DNA-binding domain. *Hum. Mol. Genet* **2**, 785-789.

Akhtar, S., Agrawal, S. (1997). *In vivo* studies with antisense oligonucleotides. *TIPS* **18**, 12-18.

Alunni-Fabbroni, M., Pirulli, D., Manzini, G., Xodo, L.E. (1996). (A, G)-oligonucleotides from extraordinary stable triple helices with a critical R.Y sequence of the Murine c-Ki-ras promoter inhibit transcription in transfected NIH 3T3 cells. *Biochemistry* **35**, 16361-16369.

Antoniou, E., Oulmouden, A., Skidmore, C.J., Urdaci, M. (1996). Isolation of a Y-specific bovine DNA fragment using minisatellite-related PCR primers. *Anim. Genet* **27**, 214-215.

Applied Biosystems. (1987). The evaluation and purification of synthetic oligonucleotides. *User Bulletin DNA synthesiser Model 380/381*, **13**

Arnemann, J., Epplen, J.T., Cooke, H.J., Sauermann, U., Engel, W., Schmidtke, J. (1987). A human Y-chromosomal DNA sequence expressed in testicular tissue. *Nucleic Acids Research* **15**, 8713-8724.

Arnott, S., Selsing, E. (1974). Structures for the polynucleotide complexes Poly(dT)·Poly(dA)·Poly(dT). *J. Mol. Biol* **88**, 509-521.

Arezzo, F. (1989). Sea urchin sperm as a vector for foreign genetic information. *Cell. Biol. Int. Rep* **13**, 391-404.

Ashwood-Swith, M.J., Poulton, G.A., Barker, M., Mildenberger, M. (1980). 5-Methoxypsoralen, an ingredient in several suntan preparations has lethal, mutagenic and clastogenic properties. *Nature* **285**, 407-409.

Asakari, F.K., McDonnell, W.M. (1996). Antisense-oligonucleotide therapy. *N. Eng. J. Med* **334**, 316-318.

Bain, D.l., Acker, G.K. (1998). A quantitative cryogenic gel-shift technique for analysis of protein-DNA binding. *Ana Biochem* **258**, 240-245.

Baker, H., Wilkinson, D.S. (1972). Psoriasis. *Textbook of Dermatology*, edited by A, Rook, D.S. Wilkinson, F.J.G. Ebling, Oxford Blackwell Scientific Publications, 2nd Edition.

Bates, P.J., Macaulay, V.M., McLean, M.J., Jenkins, T.C., Reszka, A.P., Laughton, C.A., Neidle, S. (1995). Characteristics of triplex-directed photoadduct formation by psoralen-linked oligodeoxynucleotides. *Nucleic Acids Research* **23**, 4283-4289.

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

Bayever, E., Iversen, P., Bishop, M., Sharp, J., Tweary, H., Arneson, M., Pirruccello, S., Ruddon, R., Kessinger, A., Zon, G., Armitage, J. (1994). Systemic administration of a Phosphorothioate oligonucleotide with a sequence complementary to p45 for Acute Myelogenous Leukemia and Myelodyplastic syndrome: Initial results of Phase I trial. *Antisense. Res. Dev* **4**, 383-390.

Behe, M.J. (1987). The DNA sequence of the human beta-globin region is strongly biased in favour of long strings of continuous purine or pyrimidine residues. *Biochemistry* **1**, 7870-7875.

Bennett, C.F., Chiang, M-Y., Chang, H., Ellen, J., Shoemaker, E., Mirabelli, C.K. (1992). Cationic lipids enhance cellular uptake and activity of phosphorothioate antisense oligonucleotides. *Molecular Pharmaocology* **41**, 1023-1033.

Berg, O.G., Winter, R.B., von Hippel, P.H. (1982). How do genome-regulatory proteins locate their DNA target sites. *TIBS* **7**, 52-55.

Bergan, R., Yvette, C., Fahmy, B., Neckers, L. (1993). Electroporation enhances c-myc antisense oligodeoxynucleotide efficacy. *Nucleic Acids Research* **21**, 3567-3573.

Berta, P., Hawkins, J.R., Sinclair, A.H., Taylor, A., Griffiths, B., Goodfellow, P.N., Fellous, M. (1990). Genetic evidence equating SRY and the testis-determining factor. *Nature* **348**, 448-450.

BioRad Laboratories. (1990). Pulse controller operating instructions and application guide. Catalogue number 165-2098.

Birnstiel, M.L., Busslinger, M. (1989). Dangerous liaisons: Spermatozoa as natural vectors for foreign DNA? *Cell* **57**, 701-702.

Bishop, M.R., Iversen, P.L., Bayever, E., Sharp, J.G., Greiner, T.L., Copple, B.L., Ruddon, R., Zon, G., Spinolo, J., Arneson, M., Armitage, J.O., Kessinger, A. (1996). Phase I trial of an antisense oligonucleotide OL(1)p53 in hematologic malignancies. *J. Clin. Oncol* **14**, 1320-1326.

de Bizemont, T., Duval-Valentin, G., Sun, J-S., Bisagni, E., Garestier, T., Helene, C. (1996). Alternative strand recognition of double-helical DNA by (T,G)-containing oligonucleotides in the presence of a triple helix-specific ligand. *Nucleic Acids Research* **24**, 1136-1143.

Blake, K.R., Murakami, A., Miller, P.S. (1985). Inhibition of rabbit Globin mRNA translation by sequence-specific oligodeoxyribonucleotides. *Biochemistry* **24**, 6132-6138.

Blake, R.D., Fresco, J.R. (1966). Polynucleotides VII. Spectrophotometric study of the kinetics of formation of the two-stranded helical complex resulting from the interaction of polyriboadenylate and polyribouridylate. *J. Mol. Biol* **19**, 145-160.

Blake, R.D., Massoulie, J., Fresco, J.R. (1967). Polynucleotide VIII. A spectra approach to the equilbria between polyriboadenylate and polyribouridylate and their complexes. *J. Mol. Biol* **30**, 291-308.

Blasko, A., Dempcy, R.O., Minyat, E.E., Bruice, T.C. (1996). Association of short-strand DNA oligomers with Guanidinium-linked nucleosides. A kinetic and thermodynamic study. *J. Am. Chem. Soc* **118**, 7892-7899.

Blume, S.W., Gee, J.E., Shrestha, K., Miller, D.M. (1992). Triple helix formation by purine-rich oligonucleotides targeted to the human dihydrofolate reductase promoter. *Nucleic Acids Research* **22**, 1777-1784.

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

Bonn, D. (1996). Prospects for antisense therapy are looking brighter. *The Lancet* **347**, 820-820.

Brackett, G.B., Baranska, W., Sawichi, W., Koprowski, H. (1971). Uptake of heterologous genome by mammalian spermatozoa and its transfer to ova through fertilization. *Proc. Natl. Acad. Sci. USA* **68**, 353-357.

Branch, A.D. (1998). A good antisense molecule is hard to find. *TIBS* **23**, 45-50.

Brown, E.L., Belagje, R., Ryan, M.J., Khorana, H.G. (1979). Chemical synthesis and cloning of a tyrosine tRNA gene. *Methods in Enzymology* **68**, 109-151.

Brown, P.M., Drabble, A., Fox, K.R. (1996). Effect of a triplex-binding ligand on triple helix formation at sites within a natural DNA fragment. *Biochemical Journal* **314**, 427-432.

Brown, P.M., Fox, K.R. (1996). Nucleosome core particles inhibit DNA triple helix formation. *Biochemical Journal* **319**, 607-611.

Brown, P.M., Fox, K.R. (1998). DNA triple-helix formation of nucleosome-bound poly(dA)·poly(dT) tracts. *Biochemical Journal* **333**, 259-267.

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

Brown, P.M., Fox, K.R. (1999). DNA triple-helix formation on nucleosome core particles. Effect on length of the oligopurine tract. *Eur J Biochem* **261**, 301-310.

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

Camaioni, A., Russo, M.A., Odorisio, T., Gandolfi, F., Fazio, V.M., Siracusa, G. (1992). Uptake of exogenous DNA by mammalian spermatozoa: Specific localization of DNA by sperm heads. *J. Reprod. Fert* **96**, 203-212.

Catt, S.L., Sakkas, D., Bizzaro, D., Bianchi, P.G., Maxwell, W.M.C., Evans, G. (1997). Hoechst staining and exposure to UV laser during flow cytometric sorting does not affect the frequency of detected endogenous DNA nicks in abnormal and normal human spermatozoa. *Mol. Hum. Reprod* **3**, 821-825.

Chan, P.J., Kalugdan, T., Su, B.C., Whitney, E.A., Perrott, W., Tredway, D.R., King, A. (1995). Sperm as a noninvasive gene delivery system for preimplantation embryos. *Fertility and Sterility* **63**, 1121-1124.

Chandler, S.P., Fox, F.R. (1996). Specificity of antiparallel DNA triple helix formation. *Biochemistry* **35**, 15038-15048.

Chatterjee, P.K., Cantor, C.R. (1978). Photochemical production of psoralen-DNA monoadducts capable of subsequent photocrosslinking.
Nucleic Acids Research **5**, 3619-3633.

Cheng, A.J., Van Dyke, M.W. (1993). Monovalent cation effects on intermolecular purine-purine-pyrimidine triple-helix formation.
Nucleic Acids Research **21**, 5630-5635.

Cheng, A.J., Van Dyke, M.W. (1994). Oligdeoxyribonucleotide length and sequence effects on intermolecular purine-purine-pyrimidine triple-helix formation.
Nucleic Acids Research **22**, 4742-4747.

Church, G.M., Gilbert, W. (1984). Genomic sequencing.
Pro. Natl. Acad. Sci USA **81**, 1991-1995.

Cimino, G.D., Gamper, H.B., Issacs, S.T., Hearst, J.E. (1985). Psoralens as photoactive probes of nucleic acid structure and function: Organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem* **54**, 1151-1193.

Cimino, G.D., Shi, Y.B., Hearst, J.E. (1986). Wavelength dependence for the photoreversal of a psoralen-DNA cross-link. *Biochemistry* **20**, 3013-3020.

Clarenc, J.P., Lebleu, B., Leonetti, J. (1994). Base changes and triple-helix hybridization properties of GT containing third strands: A systematical study.
Nucleoside Nucleotides **13**, 799-809.

Cole, R.S. (1970). Light-induced cross-linking of DNA in the presence of a Furocourmarin (psoralen). *Biochim. Biophys. Acta* **217**, 30-39.

Conconi, A., Widmer, R.M., Koller, T., Sogo, J.M. (1989). Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle.
Cell **2**, 753-761.

Conrad, R.C., Giver, L., Tian, Y., Ellington, A.D. (1996). *In vitro* selection of nucleic acid aptamers that bind proteins. *Methods in Enzymology* **267**, 336-339.

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

Cripps, D.J., Lowe, N.J., Lerner, A.B. (1982). Action spectra of topical psoralens: A re-evaluation. *Br J. Dermatol* **107**, 77-82.

Crooke, R.M. (1991). *In vitro* toxicology and pharmacokinetics of antisense oligonucleotides. *Anti-Cancer Drug Design* **6**, 609-646.

CyGene Inc. (1998). Triplex DNA as a novel diagnostic tool.
Mol. Med. Today, April 140-141.

Dagle, J.M., Weeks, D.L. (1996). Positively charged oligonucleotides overcome potassium-mediated inhibition of triplex DNA formation. *Nucleic Acids Research* **24**, 2143-2149

Dall'Acqua, F., Terbojevich, M., Marciani, S., Vedaldi, D., Recher, M. (1978). Investigation of the dark interaction between furocoumarins and DNA.
Chem. Biol. Interact **21**, 103-115.

Daneau, I., Houde, A., Ethier, J-F., Lussier, J.G., Silversides, D.W. (1994). *Bos taurus* sex determining region Y protein (SRY) gene sequence. Unpublished. Genbank n# BTU15569.

Daneau, I., Houde, A., Ethier, J-F., Lussier, J.G., Silversides, D.W. (1995). Bovine *SRY* gene locus: Cloning and testicular expression. *Bio. Reprod* **52**, 591-599.

Debin, A., Malvy, C., Svinarchuk, F. (1997). Investigation of the formation and intracellular stability of (purine·purine/pyrimidine) triplexes. *Nucleic Acids Research* **25**, 1965-1974

Delporte, C., Panyutin, I.G., Sedelnikova, O.A., Lillibridge, C.D., O'Connell, B.C., Baum, B.J. (1997). Triplex-forming oligonucleotides can modulate aquaporin-5 gene expression in Epithelial cells. *Antisense & Nucleic Acid Drug Dev* **7**, 523-529.

Deng, G-R., Wu, R. (1983). Terminal transferase: Use in the tailing of DNA and for *in vitro* mutagenesis. *Methods in Enzymology.* **100**, 96-116.

Dervan, P.B., Beal, P.A. (1991). Second structural motif for recognition-directed triple-helix formation. *Science* **251**, 1360-1363.

Drew, H.R., Travers, A.A. (1985). DNA bending and its relation to nucleosome positioning. *J. Mol. Biol* **186**, 773-790.

Durland, R.H., Kessler, D.J., Gunnell, S., Duvic, M., Pettitt, B.M., Hogan, M.E. (1991). Binding of triple helix forming oligonucleotides to sites in gene promoters. *Biochemistry* **30**, 9255-9263.

Duval-Valentin, G., Thuong, N.T., Helene, C. (1992). Specific inhibition of transcription by triple-helix-forming oligonucleotides. *Pro. Natl. Acad. Sci. USA* **89**, 504-508.

Edelson, R.L., Berger, C.L., Gasparro, F.P. (1987). Treatment of cutaneous T cell lymphoma by extracorporeal photochemotherapy. *N. Engl. J. Med* **316**, 297-303.

El Mofty, A.M. (1948). A preliminary clinical report on the treatment of leukoderma with *Ammi majus* Linn. *J. R. Egypt. Med. Assoc* **31**, 651-665.

Ericsson, S.A., Garner, D.L., Redelman, D., Ahmad, K. (1989). Assessment of the viability and fertilizing potential of cryopreserved bovine spermatozoa using dual fluorescent staining and two-flow cytometric systems. *Gamet res* **22**, 355-368.

Esposito, F., Brankamp, R.G., Sinden, R.R. (1988). DNA sequence specificity of 4, 5', 8-trimethylpsoralen cross-linking. *J. Biol. Chem* **263**, 11466-11472.

Faucon, B., Mergny, J-L., Helene, C. (1996). Effect of third strand composition of triple helix formation: Purine versus pyrimidine oligodeoxynucleotides. *Nucleic Acids Research* **24**, 3181-3188.

Fawad Faruqi, A., Krawczyk, S.H., Matteucci, M.D., Glazer, P.M. (1997). Potassium- resistant triple helix formation and improved intracellular gene targeting by oligodeoxyribonucleotides containing 7-deazaxanthine. *Nucleic Acids Research* **25**, 633-640.

Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P. Ringold, G.M., Danielsen, M. (1987). Lipofection: A highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. USA* **84**, 7413-7417.

Felsenfeld, F. (1957). Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc* **79**, 2023-2024.

Ferdous, A., Watanabe, H., Akaike, T., Maruyama, A. (1998). Poly(L-lysine)-graft-dextran copolymer: Amazing effects on triplex stablization under physiological pH and ionic conditions (*in vitro*). *Nucleic Acids Research* **26**, 3949-3955.

Ferdous, A., Watanabe, H., Akaike, T., Maruyama, A. (1998). Comb-type copolymer: Stabilization of triplex DNA and possible application in antigene strategy. *J. Pharm. Sci* **87**, 1400-1404.

Fossella, J.A., Kim, Y.J., Shih, H., Richards, E.G., Fresco, J.R. (1993). Relative specificities in binding of Watson-Crick base pairs by third strand residues in a DNA pyrimidine triplex motif. *Nucleic Acids Research* **21**, 4511-4515.

Fowlks, W.L. (1959). The chemistry of psoralens. *J. Invest. Dermatol* **32**, 249-254.

Frank-Kamenetskii, M.D., Mirkin, S.M. (1995). Triplex DNA structures. *Annu. Rev. Biochem* **64**, 65-95.

Frey, M.N., Koetzle, T.F., Lehman, M.S., Hamilton, W.C. (1973). Precision neutron diffraction structure determination of protein and nucleic acid components. XII. A study of H bonding in the purine-pyrimidine base pair 9-methyladenine·1-methylthymine. *J. Chem. Phys* **59**, 915-924.

Furukawa, Y., Piwnica-Worms, H., Kanakura, Y., Ernst, T.J., Griffin, J.D. (1990). cdc2 gene expression at the G1 to S transition in human T lymphocytes. *Science* **250**, 805-808.

Gasparro, F.P., Havre, P.A., Olack, G.A., Guther, E.J., Glazer, P.J. (1994). Site-specific targeting of psoralen photoadducts with a triple helix-forming oligonucleotide: Characterization of psoralen monoadduct and crosslink formation. *Nucleic Acids Research* **22**, 2845-2852

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

Gee, J.E., Revankar, G.R., Rao, T.S., Hogan, M.E. (1995). Triplex formation at the rat *neu* gene utilizing imidazole and 2'-deoxy-6-thioguanosine base substitutions. *Biochemistry* **34**, 2042-2048.

Gilchrest, B.A., Parrish, J.A., Tanenbau, L. (1976). Oral methoxsalen photochemotherapy of mycosis fungicides. *Cancer* **38**, 683-689.

Giovannangeli, C., Dviacco, S., Labrousse, V., Gryaznov, S., Charneau, P., Helene, C. (1997). Accessibility of nuclear DNA to a triple-forming oligonucleotide: The intergrated HIV-1 provirus as a target. *Pro. Natl. Acad. Sci. USA* **94**, 79-84.

Giovannangeli, C., Helene, C. (1997). Progress in developments of triplex-based strategies. *Antisense & Nucleic Acid Drug Dev* **7**, 413-421.

Glazer, P.M., Sarkar, S.N., Summers, W.C. (1986). Detection and analysis of UV-induced mutations in mammalian cell DNA using a λ phage shuttle vector. *Pro. Natl. Acad. Sci. USA* **83**, 1041-1044.

Gordon, J.W., Ruddle, F.H. (1981). Mammalian gonadal determination and gametogenesis. *Science* **211**, 1265-1271.

Gould-Fogerite, S., Mazurkiewicz, J.E., Raska, K., Voelkerding, K., Lehman, J.M., Mannino, R.J. (1989). Chimerasome-mediated gene transfer *in vitro* and *in vivo*. *Gene* **84**, 429-438.

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

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

Griffin, L.C., Dervan, P.B. (1989). Recognition of thymine·adenine base pairs by guanine in a pyrimidine triple helix motif. *Science* **245**, 967-970.

Grigoriev, M., Praseuth, D., Guieysee, A.L., Robin, P., Thuong, N.T., Helene, C., Harel-Bellan, A. (1993). Inhibition of gene expression by triple-helix-directed DNA cross-linking at specific sites. *Pro. Natl. Acad. Sci. USA* **90**, 3501-3505.

Grobet, L., Georges, M. (1996). Cloning seven new bovine Y-repetitive sequences using GDRDA. *Animal Genetics* **27**, 91-91.

Grossweiner, L.I., Smith, K.C. (1981). Sensitivity of DNA repair-deficient strains of *Escherichia coli* K-12 to various furocoumarins and near-ultraviolet radiation. *Photochem. Photobiol* **33**, 317-323.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346**, 245-250.

Gunther, E.J., Havre, P.A., Gasparro, F.P., Glazer, P.M. (1996). Triplex-mediated, *in vitro* targeting of psoralen photoadducts within the genome of a transgenic mouse. *Photochem. Photobiol* **63**, 207-212.

Hacia, J.G., Dervan, P.B., Wold, B.J. (1994). Inhibition of Klenow fragment DNA polymerase on double-helical templates by oligonucleotide-directed triple-helix formation. *Biochemistry* **33**, 6192-6200.

Hanson, C.V., Shen, C.J., Hearst, J.E. (1976). Cross-linking of DNA *in situ* as a probe for chromatin structure. *Science* **193**, 62-64.

Hanson, C.V., Riggs, J.L., Lennette, E.H. (1978). Photochemical inactivation of DNA and RNA viruses by psoralen derivatives. *J. Gen. Virol* **40**, 345-358.

Hanze J., Worgall, S., Rascher, W. (1997). Monitoring of antisense effects of oligonucleotide targeted to the neuropeptide YY1 receptor gene. *Eur. J. Pharmacol.* **330**, 87-92.

Haran, T.E., Crothers, D.M. (1988). Phased psoralen cross-links do not bend the DNA double helix. *Biochemistry* **27**, 6971-6976.

Havre, P.A., Gunther, E.J., Gasparro, F.P., Glazer, P.M. (1993). Targeted mutagenesis of DNA using triple helix-forming oligonucleotides linked to psoralen. *Pro. Natl. Acad. Sci. USA* **90**, 7879-7883.

Hearst, J.E., Straub, K., Kanne, D., Rapoport, H. (1981). Isolation and characterization of pyrimidine-psoralen photoadducts from DNA. *J. Am. Chem. Soc* **103**, 2347-2355.

Helene, C., Montenzy-Garestier, T., Saison, T., Takasugi, M., Toulem, J.J. (1985). Oligodeoxynucleotides covalently linked to intercalating agents: A new class of gene regulatory substances. *Biochimie* **67**, 777-783.

Helene, C., Collier, D.A., Thuong, N.T. (1991). Sequence-specific bifunctional DNA ligands based on triple-helix-forming oligonucleotides inhibit restriction enzyme cleavage under physiological conditions. *J. Am. Chem. Soc* **113**, 1457-1458.

Helene C. (1994). Control of oncogene expression by antisense nucleic acids. *Eur. J. Cancer* **30a**, 1721-1726.

Hoeijmakers, J.H., Bootsma, D. (1994). DNA repairs. Incision for excision. *Nature* **37**, 654-655

Hoogsteen, K. (1963). The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta. Crystallogr* **16**, 907-916.

Howard, F.B., Miles, H.T., Liu, K., Frazier, J., Raghunathan, G., Sasisekharan, V. (1992). Structure of d(T)n.d(A)n.d(T)n: The DNA triple helix has B-form geometry with C2'-endo sugar pucker. *Biochemistry* **10**, 10671-10677.

Htun, H., Dahlberg., J.E. (1989). Topology and formation of triple-stranded H-DNA. *Science* **243**, 1571-1576.

Ishimi, Y., Kikuchi, A. (1991). Identification and molecular cloning of yeast nucleosome assembly protein I which facilitates nucleosome assembly *in vitro*. *J. Biol. Chem* **266**, 7025-7029.

Jakubiczka, S., Schnieders, F., Schmidtke, J. (1993). A bovine homologue of the human TSPY gene. *Genomics* **17**, 732-735.

Jarnagin, W.R., Debs, R.J., Wang, S-S., Bissell, D.M. (1992). Cationic lipid-mediated transfection of liver cells in primary culture. *Nucleic Acids Research* **20**, 4205-4211.

Kandimalla, E.R., Manning, A.N., Venkataraman, G., Sasisekharan, V., Agrawi, S. (1995). Single strand targeted triplex formation: targeting purine-pyrimidine mixed sequences using basic linkers. *Nucleic Acids Research* **23**, 4510-4517.

Kanne, D., Straub, K., Hearst, J.E., Rapoport, H. (1982). Isolation and characterization of pyrimidine-psoralen-pyrimidine photodiadducts from DNA. *J. Am. Chem. Soc* **104**, 6754-6769.

Kanne, D., Strub, K., Rapoport, H., Hearst, J.E. (1982). Psoralen-deoxyribonucleic acid photoreaction. Characterization of the monoaddition products from 8-methoxypsoralen and 4, 5', 8-trimethylpsoralen. *Biochemistry* **21**, 861-871.

Kenan, D.J., Tsai, D.E., Keene, J.D. (1994). Exploring molecular diversity with combinational shape libraries. *TIBS* **19**, 57-64.

Kim, S.H., Pearlman, E.A., Holbrooke, S.R., Pirkle, D. (1985). Structures of DNA containing psoralen crosslink and thymine dimer. *Prog. Clin. Biol. Res* **172A**, 143-152.

Kittler, L., Midden, W.R., Wang, S.Y. (1980). Photoreactions of Lipids sensitized by furocourmarins. *Photochem. Photobiol* **37**, S16

Kochetkova, M., Shannon, M.F. (1996). DNA triplex formation selectively inhibits Granulocyte-macrophage colony-stimulation factor gene expression in human T-cells. *J. Biol. Chem* **217**, 14438-14444.

Koh, J.S., 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.

Kool, E.T. (1996). Circular oligonucleotides: New concepts in oligonucleotide design. *Annu. Rev. Biophys. Biomol. Struct* **25**, 1-28.

Koopman, P., Munsterberg, A., Capel, B., Vivian, N., Lovell-Badge, R. (1990). Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* **348**, 450-452.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for *Sry*. *Nature* **351**, 117-121.

Kopel, V., Pozner, A., Baran, V., Manor, H. (1996). Unwinding of the third strand of a DNA triplex helix, a novel activity of the Sv40 large T-antigen helicase. *Nucleic Acids Research* **24**, 330-335.

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

Krawczyk, S.H., Milligan, J.F., Wadwani, S., Moulds, C., Froehler, B.C., Matteucci, M.D. (1992). Oligonucleotide-mediated triple helix formation using an N3-protonated deoxycytidine analog exhibiting pH-independent binding within the physiological range. *Pro. Natl. Acad. Sci. USA* **89**, 3761-3764.

Kukreti, S., Sun, J-S., Loakes, D., Brown, D.M., Nguyen, C-H., Bisagni, E., Garestier, T., Helene, E. (1998). Triple helices formed at oligopyrimidine-oligopurine sequences with base pair inversions: Effect of a triplex-specific ligand on stability and selectivity. *Nucleic Acids Research* **26**, 2179-2183.

Laquerbe, A., Guillouf, C., Moustacchi, E., Papadopoulo. (1995). The mutagenic processing of psoralen photolesion leaves a highly specific signature at an endogenous human locus. *J. Mol. Biol* **254**, 38-49.

Lasic, D.D. (1997). *Lipsomes in gene delivery*. Boca Raton FL: CRC Press.

Lasic, D.D., D Papahadjopoulous. (1998). *Medical applications of Liposomes*. Elsevier Science, USA.

Lavitrano, M., Camaioni, A., Fazio, V., Dolci, S., Farace, M.G., Spadafora, C. (1989). Sperm cells as vectors for introducing foreign DNA into eggs: Genetic transformation of mice. *Cell* **57**, 717-723.

Lavitrano, M., French, D., Zani, M., Frati, L., Spadafora, C. (1992). The interaction between exogenous DNA and sperm cells. *Mol. Reprod. Dev* **31**, 161-169.

Lee, J.S., Latimer, L.P., Hampel., K.J. (1993). Coralyne binds tightly to both T·A·T and C·G·C$^+$ containing DNA triplexes. *Biochemistry* **32**, 5591-5597.

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

Le Pecq, J.B., Paoletti, C. (1965). Interaction of ethidium hydrobromate with polyribonucleotides. Applications to the study of hybridization reactions. *C.R. Acad. Sc. Paris* **260**, 7033-7036.

Lewin, B. (1974). *Gene expression: Eucaryotic chromosomes* Vol II. John Wiley & Sons.

Lewis, J.G., Lin, K-Y., Kothavale, A., Flanagan, M.W., Matteucci, M.D., DePrince, F.B., Mook, R.A., Hendren, W.R., Wagner, R.W. (1996). A serum-resistant cytofection for cellular delivery of antisense oligonucleotides and plasmid DNA. *Proc. Natl. Acad. Sc. USA* **93**, 3176-3181.

Libbus, B.L., Perreault, S.D., Johnson, L.A., Pinkel, D. (1987). Incidence of chromosome aberrations in mammalian sperm stained with Hoechst 33342 and UV-laser irradiated during flow sorting. *Mut. Res* **182**, 265-274.

von Lindern, M., van Baal, S., Wiegant, J., Rapp, A., Hagemeijer, A., Grosvelt, G. (1992). *can*, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3' half to different genes: Characterization of the *set* gene. *Mol. Cell. Biol* **12**, 3346-3355.

Liquier, J., Coffinier, P., Firon, M., Taillandier, E. (1991). Triple helical polynucleotidic structures: Sugar conformations determined by FTIF spectroscopy. *J. Biomol. Struct. Dyn.* **9**, 437-445.

Lyamichev, V.I., Mirkin, S.M., Frank-Kamenetskii, M.D., Cantor, C.R. (1988). A stable complex between homopyrimidine oligomers and the homologous regions of duplex DNAs. *Nucleic Acids Research* **16**, 2165-2178.

Lyamichev, V.I., Frank-Famenetskii, M.D., Soyer, V.N. (1990). Protection against UV-induced pyrimidine dimerization in DNA by triplex formation. *Nature* **5**, 568-570.

Lyamichev, V.I., Voloshin, O.N., Frank-Kamenetskii, M.D., Soyfer, V.N. (1991). Photofootprinting of DNA triplexes. *Nucleic Acids Research* **19**, 1633-1638.

Maher, L.J., Dervan, P.B., Wold, B.J. (1989). Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation. *Science* **245**, 725-730.

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

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

Majumdar, A., Khorlin, A., Dyatkina, N., Lin, F-L., Powell, J., Liu, J., Fei, A., Khripine, Y., Watanabe, K.A., George, J., Glazer, P.M., Seidman, M.M. (1998). Targeted gene knockout mediated by triple helix forming oligonucleotides. *Nature Genetics* **20**, 212-214.

Malkov, V.A., Voloshin, O.N., Veselkov, A.G., Rostapshov, V.M., Jansen, I., Soyfer, V.N., Frank-Kamenetskii. (1993). Protonated pyrimidine-purine-purine triplex. *Nucleic Acids Research* **21**, 105-111.

Maniatis, T., Sambrook, J., Fritsch, E.S. (1982). Molecular Cloning, a laboratory manual. *Cold Spring Harbor Laboratory Press.* 1st Edition, New York.
Maniatis, T., Sambrook, J., Fritsch, E.S. (1989). Molecular Cloning, a laboratory manual. *Cold Spring Harbor Laboratory Press.* 2nd Edition, New York.

Manor, H., Roa, B.S., Martin, R.G. (1988). Abundance and degree of dispersion of genomic $d(GA)_n{\cdot}d(TC)_n$ sequences. *J. Mol. Evol* **27**, 96-101.

Maruyama, A., Katoh, M., Ishihara, T., Akaike, T. (1997). Comb-type polycations effectively stabilize DNA triplex. *Bioconjuate Chemistry* **8**, 3-6.

Matsuda, Y., Tobari, I. (1988). Chromosomal analysis in mouse eggs fertilized *in vitro* with sperm exposed to ultraviolet and methyl and ethyl methanesulfonate. *Mut. Res* **198**, 131-144.

Matthews, M.E., Reed, K.C. (1991). A DNA sequence that is present in both sexes of Artiodactyla is repeated on the Y chromosome of cattle, sheep, and goats. *Cytogenet. Cell. Genet* **56**, 40-44.

Matthews, M.E., Reed, K.C. (1992). Sequence from a family of bovine Y-chromosomal repeats. *Genomics* **13,** 1267-1273.

Maxam, A.M., Gilbert, W. (1977). A new method for sequencing DNA. *Pro. Natl. Acad. Sci. USA* **74**, 560-564.

McCarrey, J.R., Abbott, U.K. (1979). Mechanisims of genetic sex determination, gonadal sex differentiation, and gem-cell development in animals. *Adv. Genet* **20**, 217-290.

McShan, M.W., Rossen, R.D., Laugher, A.H., Trial, J., Kessler, D.J., Zendegui, J.G., Hogan, M.E., Orson, F.M. (1992). Inhibition of transcription of HIV-1 in infected human cells by oligodeoxynucleotides designed to form DNA triple helices. *J. Biol. Chem* **267**, 5712-5721.

Megaw, J., Lee, J., Lerman, S. (1980). NMR analyses of Tryptophan-8-methoxypsoralen photoreaction products formed in the presence of oxygen. *Photochem. Photobiol* **32**, 265-269.

Melski, J.W., Tanenbaum, L., Parrish, J.A., Fitzpatrick, T.B., Bleich, H.L., and 28 participating investigators. (1977). Oral methoxsalen photochemistry for the treatment of Psoriasis: A cooperative clinical trial. *J. Invest. Dermatol* **68**, 328-335.

Mergny, J.L., Collier, D., Rougee, M., Montenay-Garestier, T., Helene, C. (1991). Intercalation of ethidium bromide into a triplex-stranded oligonucleotide. *Nucleic Acids Research* **19**, 1521-1526.

Mergny, J.L., Sun, J-S., Rougeen, M., Montenay-Garestier, T., Barcelo, F., Chomiller, J., 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.

Mergny, J.L., Duval-Valentin, G., Nguyen, C.H., Perrouault, L., Faucon, B., Rougee, M., Montenay-Garestier, T., Bisagni, E., Helene, C. (1992). Triple helix-specific ligands. *Science* **238**, 1681-1684.

Miller, P.S., Bhan, P., Cushman, C.D., Trapane T.L. (1992). Recognition of a guanine-cytosine base pair by 8-oxoadenine. *Biochemistry* **31**, 6788-6793.

Miller, P.S., Bi, G., Ki, S.A., Fok, V., Delong, R.K. (1996). Triplex formation by a psoralen-conjugated oligodeoxyribonucleotide containing the base analog 8-oxoadenine. *Nucleic Acids Research* **24**, 730-736.

Monia, B.P., Johnston, J.F., Sasmor, H., Cummins, L.L. (1996). Nucleotide resistance and antisense activity of modified oligonucleotides targeted to Ha-*ras*. *J. Biol. Chem* **271**, 14533-14540.

Moraru-Allen, A., Cassidy, S., Asensio Alvarez, J-L., Fox, K.R., Brown, T., Lande, A.N. (1997). Coralyne has a preference for intercalation between TA·T triples in intramolecular DNA triple helices. *Nucleic Acids Research* **25**, 1890-1896.

Moser, H.E., Dervan, P.B. (1987). Sequence specific cleavage of double helical DNA by triple helix formation. *Science* **238**, 646-650.

Musajo, L., Rodighiero, G., Colombo, J., Torlone, Dall'Adqua, F. (1965). Photosensitizing furocoumarins: Interactions with DNA and photoactivation of DNA containing viruses. *Experientia* **21**, 22-24.

Musajo, L., Bordin, F., Caporale, G., Marciani, S., Rigatti, G. (1967). Photoreactions at 3655A° between pyrimidine bases and skin-photosensitizing furocoumarins. *Photochem. Photobiol* **6**, 711

Musajo, L., Rodighiero, G., Dall'Acqua, F., Marciani, S., Caporale, G., Ciavatta, M.L. (1969). A comparison between the photoreactivity of some furocoumarins with native DNA and their skin-photosensitizing activity. *Experientia* **25**, 479-481.

Musso, M., Van Dyke, M.W. (1995). Polyamine effects on purine-purine-pyrimidine triplex helix formation by phosphodiester and phosphorothioate oligodeoxyribonucleotides. *Nucleic Acids Research* **23**, 2320-2327.

Musso, M., Wang, J.C., Van Dyke, M.W. (1996). *In vivo* persistence of DNA triple helices containing psoralen-conjugated oligodeoxyribonucleotides. *Nucleic Acids Research* **24**, 4924-4932.

Nakanishi, A., Iritani, A. (1993). Gene transfer in the chicken by sperm-mediated methods. *Mol. Reprod. Devel* **36**, 258-261.

Nielson, P.E., Bohr, V. (1983). Phototoxic effects of four psoralens on L1210 cells; The correlation with DNA interstrand cross-linking. *Photochem. Photobiol* **38**, 653-657.

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

Nielson, P.E. (1996). Peptide Nucleic Acids: A new dimension to peptide libraries and aptamers. *Methods in Enzymology* **267**, 426-433

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

Olivas, W.M., Maher, L.J. (1996). Binding of DNA oligonucleotides to sequences in the promoter of the human *bcl*-2 gene. *Nucleic Acids Research* **24**, 1758-1764.

Oroskar, A.A., Gasparro, F.P., Peak, M.J. (1993). Relaxation of supercoiled DNA by Aminomethyltrimethylpsoralen and UV photons: Action spectrum. *Photochem. Photobiol* **57**, 648-654.

Orson, R.M., Thomas, D.W., McShan, W.M., Kessler, D.J., Hogan, M.E. (1991). Oligonucleotide inhibition of IL2Rα mRNA transcription by promoter region collinear triplex formation in lymphocytes. *Nucleic Acids Research* **19**, 3435-3441.

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

Paris, C.M., Seidman, M.M. (1992). A signature element distinguishes sibling and independent mutations in a shuttle vector plasmid. *Gene* **117**, 1-5.

Parrish, J.A., Fitzpatrick, T.B., Tanenbaum, L., Pathak, M.A. (1974). Photochemotherapy of psoriasis with oral methoxsalen and longwave ultraviolet light. *N. Engl. J. Med* **5**, 1207-1211.

Payen, E.J., Cotionot, C.Y. (1993). Comparative HMG-box sequences of the *SRY* gene between sheep, cattle, and goats. *Nucleic Acids Research* **21**, 2772

Perkins, B.D., Wilson, J.H., Wensel, T.G., Vasquez, K.M. (1998). Triplex targets in the human rhodopsin gene. *Biochemistry* **37**, 11315-11322.

Pilch, DS., Kirolos, M.A., Plum, G.E., Breslaner, K.J. (1995). Berenil binding to DNA duplexes and to a RNA duplex: Evidence for both intercalative and minor groove binding properties. *Biochemistry* **34**, 9962-9976.

Raghunathan, G., Todd-Miles, H., Sasisekharan, V. (1994). Parallel nucleic acid helices with Hoogsteen base pairing: Symmetry and structure. *Biopolymers* **34**, 1573-1581.

Raha, M., Lacroix, L., Glazer, P.M. (1998). Mutagenesis mediated by triple helix-forming oligonucleotides conjugated to psoralen: Effects of linker arm length and sequence content. *Photochem. Photobiol* **67**, 289-294.

Rajagopal, J., Feigon, J. (1989). Triple-strand formation in the homopurine-homopyrimidine DNA oligonucleotides $d(G\text{-}A)_4$ and $d(T\text{-}C)_4$. *Nature* **339**, 637-640.

Reed, K.C., Lord, E.A., Matthaeri, K.I., Mann, D.A., Beaton, S., Herr, S., Matthews, M.E. (1988). Sex determination in ruminants using Y chromosome specific polynucleotides. *Int Patent Applic* PCT/AU89/00029.

Reed, K.C., Lord, E.A., Matthaei, K.I., Mann, D.A., Beaton, S., Herr, C.M., Matthews, E. (1995). Determination of genetic sex in ruminants using Y chromosome specific oligonucleotides. *Int Patent Applic* 5459038.

Rippe, K., Fritsch, V., Westhof, E., Jovin, T.M. (1992). Alternating d(G-A) sequences form a parallel-stranded DNA homoduplex. *EMBO Journal* **11**, 3777-3786.

Ronimus, R.S. (1993). Investigation into the control of genome topology in the archael isolate ANI. *PhD thesis*. Waikato University.

Rougee, M., Faucon, B., Mergny, J.L., Barcelo, F., Giovannangeli, D., Garestier, T., Helene, C. (1992). Kinetics and thermodynamics of triple-helix formation. Effects of ionic strength and mismatches. *Biochemistry* **31**, 9269-9278.

Sage, E., Moustacchi, E. (1987). Sequence content effects on 8-methoxypsoralen photobinding to defined DNA fragments. *Biochemistry* **16**, 3307-3314.

Sage, E., Bredberg, A. (1991). Damage distribution and mutation spectrum: The case of 8-methyloxypsoralen plus UVA in mammalian cells.
Mutation Research **263**, 217-222.

Samadashwily, G.M., Dayn, A., Mirkin, S.M. (1993). Suicidal nucleotide sequences for DNA polymerization. *EMBO Journal* **12**, 4975-4983.

Sastry, S., Spielmann, P., Hearst, J.E. (1993). Psoralens and their applications to the study of some molecular biological processes. *In Advances in Enzymology*, edited by A. Meister. New York, John Wiley & Sons.

Scaria, P.V., Shafer, R.H. (1991). Binding of ethidium bromide to a DNA triple helix. *J. Biol. Chem* **266**, 5417-5423.

Schnieders, F., Dork, T., Arnemann, J., Vogel, T., Werner, M., Schmidtke, J. (1996). TSPY (testis-specific protein, Y-encoded) expression in testicular tissue.
Human. Mol. Genet **5**, 1801-1807.

Schroder, A., Miller, J.R., Thomsen, P.D., Roschlau, K., Avery, B., Poulsen, P.H., Schmidt, M., Schwerin, M. (1990). Sex determination of bovine embryos using the polymerase chain reaction. *Anim. Biotechnol* **1**, 121-131.

Scott, B.R., Pathak, M.A., Moh, G.R. (1976). Molecular and genetic basis of furocoumarin reactions. *Mutation Research* **39**, 29-74.

Searle, M.S., Wickham, G. (1990). Hoogsteen versus Watson-Crick A-T base pairing in DNA complexes of a new group of 'quinomycin like' antibiotics.
FEBS Letters **272**, 171-174.

Shafer, R.H. (1998). Stability and structure of model DNA triplexes and quadruplexes and their interactions with small ligands.
Prog. Nucleic. Acids. Res. Mol. Biol **59**, 55-94.

Shi, Y-B., Spielmann, P., Hearst, J.E. (1988). Base-catalyzed reversal of a psoralen-DNA cross-link. *Biochemistry* **27**, 5174-5178.

Shindo, H., Torigoe, H., Sarai, A. (1993). Thermodynamic and kinetic studies of DNA triplex formation of an oligohomopyrimidine and a matched duplex by filter binding assay. *Biochemistry* **32**, 8963-8969.

Simons, M. (1996). Antisense approach to restenosis.
Japanese Circulation Journal **60**, 1-9.

Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.M., Lovell-Badge, R., Goodfellow, P.N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346**, 240-245.

Sinden, R.R., Hagerman, P.J. (1984). Interstrand psoralen crosslinks do not produce appreciable bends in DNA. *Biochemistry* **23**, 6299-6303.

Song, P.S., Tapley, K.J. (1979). Photochemistry and photobiology of psoralens. *Photochem. Photobiol* **29**, 1177-1197.

Sperandio, S., Lulli, V., Bacci, M.L., Forni, M., Maione, B., Spadafora, C., Lavittrano, M. (1996). Sperm-mediated DNA transfer in bovine and swine species. *Animal. Biotech* **7**, 59-77.

Stec, W.J., Zon, G., Egan, W., Stec, B. (1984). Automated solid-phase synthesis, separation, and stereochemistry of Phosphorothioate analogues of oligodeoxyribonucleotides. *J. Am. Chem. Soc* **106**, 6077-6079.

Stein, C.A., Cheng, Y-C. (1993). Antisense oligonucleotides as therapeutic agents - is the bullet really magic? *Science* **261**, 1004-1011.

Stemple, N.R., Watson, W.H. (1972). The crystal and molecular structure of Xanthotoxin, $C_{12}H_8O_4$. *Acta. Crystallogr.* **B28**, 2485.

Stern, R.S., Laird, N., Melski, J., Parrish, J.A., Fitzpatrick, T.B., Bleigh, H.L. (1984). Cutaneous squamous-cell carcinoma in patients treated with PUVA. *N. Eng. J. Med* **310**, 115601161.

Stern, R.S., Lange, R. (1988). Non-melanoma skin cancer occurring in patients treated with PUVA five to ten years after first treatment. *J. Invest. Dermat* **91**, 120-124.

Stern, R.S., Liebman, E.J., Vakeva, L. (1998). Oral psoralen and ultraviolet-A light (PUVA) treatment of psorisis and persistent risk of nonmelanoma skin cancer. *J. National Cancer Institute* **90**, 1278-1284.

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

Stuerzl & Roth (1990). "Run-off" synthesis and application of defined single-stranded DNA hybridization probes. *Anal. Biochem* **185**, 164-169.

Sun, J-S., de Bizemont, T., Duval-Valentin, G., Montenay-Garestier, T., Helene, C. (1991). Extension of the range of recognition sequences for triplex helix formation by oligonucleotides containing guanines and thymines. *C. R. Acad. Sci. Paris* **313**, 585-590.

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

Svinarchuk, F., Cherny, D., Debin, A., Delain, E., Malvy, C. (1996). A new approach to overcome potassium-mediated inhibition of triplex formation. *Nucleic Acids Research* **24**, 3858-3865.

Takasugi, M., Guendouz, A., Chassignol, M., Decout, J.L., Lhomme, J., Thuong, N.T., Helene, C. (1991). Sequence-specific photo-induced cross-linking of the two strands of double-helical DNA by a psoralen covalently linked to a triple helix-forming oligonucleotide. *Pro. Natl. Acad. Sci. USA* **88**, 5602-5606.

Tessman, J.W., Isaacs, S.t., Hearst, J.E. (1985). Photochemistry of the Furan-side 8-methoxypsoralen-thymidine monoadduct inside the DNA helix. Conversion to diadduct and to pyrone-side monoadduct. *Biochemistry* **24**, 1669-1676.

Thompson, J.F., Bachellene, J-P., Hall, K., Hearst, J.E. (1982). Dependence of 4'-(hydroxymethyl) 4, 5', 8' trimethylpsoralen photoaddition on the conformation of ribonucleic acid. *Biochemistry* **21**, 1363-1368.

Thomas, T.J., Bloomfield, V.A. (1984). Ionic and structural effects on the thermal helix-coil transition of DNA complexes with natural and synthetic Polyamines. *Biopolymers* **23**, 1295-1306.

Uddin, A., Piunno, P. Hudson, R., Damha, M., Krull, U. (1997) A fiber biosensor for fluorimetric detection of triplex-helical DNA. *Nucleic Acids Research* **25**, 4139-4146.

Ussery, D.W., Hoepfner, R.W., Sinden, R.R. (1992). Probing DNA structure with psoralen *in vitro*. *Methods in Enzymology* **212**, 242-263.

Vasquez, K.M., Wensel, T.G., Hogan, M.E., Wilson, J.H. (1995). High-affinity triplex helix formation by synthetic oligonucleotides at a site within a selectable mammalian gene. *Biochemistry* **32**, 7243-7251.

Vasquez, K.M., Wensel, T.G., Hogan, M.E., 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

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

Vergnaud, G., Kaplan, L., Weissenbach, J., Dumez, Y., Berger, R., Tiollais, P., Guellaen, G. (1984). Rapid and early determination of sex of trophoblast biopsy specimens with Y chromosome specific DNA probes. *Br. Med. J* **289**, 73-76.

Veronese, F.M., Schiavon, O., Bevilacqua, R., Bordin, F., Rodighiero, G. (1982). Photoinactivation of enzymes by linear and angular furocoumarins. *Photochem. Photobiol* **36**, 25-30.

Vigneswaran, N., Mayfield, C.A., Rodu, B., James, R., Kim H-G., Miller, D.M. (1996). Influence of GC and AT specific DNA minor groove binding drugs on intermolecular triplex formation in the human c-Ki-ras promoter. *Biochemistry* **35**, 1106-1114.

Vogel, T., Dechend, F., Manz, E., Jung, C., Jakubiczka, S., Fehr, S., Schmidtke, J., Schnieders, F. (1997). Organization and expression of bovine TSPY. *Mammalian Genome* **8**, 491-496.

Vogel, T., Borgmann, S., Dechend, F., Hecht, W., Schmidtke, J. (1997). Conserved Y-chromosomal location of TSPY in Bovidae. *Chromosome Research* **5**, 182-185.

Vos, J-M., Wauthier, E.L. (1991). Differential introduction of DNA damage and repair in mammalian genes transcribed by RNA polymerases I and II. *Mol. Cell. Biol* **11**, 2245-2252.

Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C., Froehler, B.C. (1993). Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. *Science* **260**, 1510-1513.

Wagner, R.W., Flanagan, W.M. (1997). Antisense technology and prospects for therapy of viral infections and cancer. *Mol. Med. Today* **3**, 31-38.

Wang, S., Friedman, A.E., Kool, E.T. (1995). Origins of high sequence selectivity: a stopped-flow kinetics study of DNA/RNA hybridization by duplex-and triplex-forming oligonucleotides. *Biochemistry* **34**, 9774-9784.

Wang, G., Levy, D.D., Seidman, M.M., Glazer, P.M. (1995). Targeted mutagenesis in mammalian cells mediated by intracellular triple helix formation.
Mol. Cell. Biol **15**, 1759-1768.

Wang, F., Glazer, P.M. (1995). Altered repair of targeted psoralen photoadducts in the context of an oligonucleotide-mediated triple helix.
J. Biol. Chem **270**, 22595-22601.

Wang, Q., Marini, J.C. (1996). Antisense oligodeoxynucleotides selectively suppress expression of the mutant α2(I) collagen allele in Type IV Osteogenesis Imperfecta fibroblasts. *J. Clin. Invest* **95**, 448-454.

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

van de Wetering, M., Clevers, H. (1992). Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *EMBO Journal* **10**, 3039-3044.

Williamson, J.R., Raghuraman, M.K., Cech, T.R. (1989). Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell* **59**, 871-880.

Yang, M., Ghosh, S.S., Millar, D.P. (1994). Direct measurement of thermodynamic and kinetic parameters of DNA triple helix formation by Fluorescence spectroscopy. *Biochemistry* **33**, 15329-15337.

Yoon, K., Hobbs, C.A., Koch, J., Sardaro, M., Kutny, R., Weis, A.L. (1992). Elucidation of the sequence-specific third-strand of four base pairs in a pyrimidine triple-helix motif: T·AT,C·GC, T·CG, G·TA.
Proc. Nat. Acad. Sci, U.S.A **89**, 3840-3844.

Young, S.L., Krawczyk, S.H., Matteuci, M.D., Toole, J.J. (1991). Triple helix formation inhibits transcription elongation *in vitro*.
Proc. Nat. Acad. Sci, U.S.A **88**, 10023-10026.

Xodo, L.E. (1995). Characterization of the DNA triplex formed by d(TGGGTGGGTGGTTGGGTGGG) and a critical R·Y sequence located in the promoter of the murine Ki-*ras*-proto-oncogene. *FEBS Letters* **370**, 153-157.

Zahler, A.M., Williamson, J.R., Cech, T.R., Prescott, D.M. (1991). Inhibition of telomerase by G-quartet DNA structures. *Nature* **350**, 718-720.

Zamecnik, P.C., Stephenson, M.L. (1978). Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide.
Proc. Nat. Acad. Sci. USA **75**, 280-284.

Zani, M., Labitrano, M., French, D., Lulli, V., Maione, B., Sperandio, S., Spadafora, C. (1995). The mechanism of binding of exogenous DNA to sperm cells: Factors controlling the DNA uptake. *Exper. Cell. Rese* **217**, 57-64.