Regulation of Topoisomerase IIα Expression in Humans

A thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Biochemistry at Massey University, Palmerston North, New Zealand.

Kelly Senior

2006

Acknowledgments

I am extremely grateful for having the most supportive and knowledgeable supervisor, Dr. Kathryn Stowell. Her encouragement and open door policy, made this work possible and also enjoyable. Thanks for believing in me and always staying so positive and optimistic. I feel very privileged to have been one of your students. Also, thanks to Richard Isaacs, John Tweedie and Mark Patchett, for your helpful advice and support, and for always managing to find the time to help me.

I would like to express my eternal gratitude to Mum and Dad for being so encouraging, understanding, tolerant, and very generous throughout these years. Without your support I would not have come to an achievement of this extent. Thank you with all my heart!

Thanks to everyone in the Twilite Zone laboratory. You have all been great help and good entertainment. Thanks to Robyn Marston for being a great listener and for all the advice on life! Exceptional thanks to all my friends for all your support and good times, especially Tim Byrne and Troy Makan for keeping life in perspective, heaps of fun times and enduring all these years at Massey with me! Also thanks to Bex Grierson and Lili Griffiths for endless conversations and singstar battles! And of course thanks to Tracy MacKenzie for being my best friend for so many years, I hope our friendship is endless.

I would also like to thank Massey University, the Freemasons of New Zealand, and the Allan Wilson Centre for their financial support on this endeavour. Without your support, this year and years to come would have been a lot more difficult.

I would finally like to dedicate my work to Aunty Lyl, who sadly passed away due to cancer well before her time, and Nana Lotte, who endured Alzheimers disease. They are the inspiration for my work. I wish I had had more time to get to know them better, and I wish I could thank them in person.

Abstract

In mammalian cells, the loss or down-regulation of tumour-suppressor genes and/or the mutation or overexpression of proto-oncogenes, whose products promote unregulated proliferation in cells, characterise the process of malignant transformation. This generates mitogenic signals that promote abnormal cell growth resulting in tumour progression. Topoisomerase IIα (topo IIα) is an enzyme present in elevated concentrations in highly proliferating cells due to the requirement for untwisting and unknotting of the DNA which is essential for replication. Because of this requirement, a number of anti-cancer drugs have been designed with topo IIα as their primary target. The effectiveness of these drugs however is limited by the development of resistance. One factor linked to drug resistance is the down-regulation of topo IIα at the transcription level. Expression of topo IIα appears to be regulated through various transcription factors with members of the Sp1 family having a major contribution.

Previous work has shown down regulation of topo IIα can occur at the level of transcription. Nucleotide sequencing of the topo IIα promoter in drug-resistant cell lines has not revealed any mutations thus far. Three known proteins and one uncharacterised protein are capable of interacting with the proximal topo IIα promoter region. The uncharacterised protein may act as a co-activator or a co-repressor depending on the complement of transcription factors associated with the DNA in this region. Because drug resistant cell lines showed modulated expression of these transcription factors, it is important to identify the unknown protein and characterise its role in regulating topo IIα expression.

This research aimed to identify the minimal binding site and DNA elements required for the uncharacterised protein to bind, as well as introduce mutations into this proximal region and examine their functional significance. The results of this study could provide insights into the molecular mechanisms responsible for the development of drug resistance, contributing to more efficient and effective methods for the treatment of cancer.

Abbreviations

Amp	Ampicillin
ATP	Adenosine triphosphate
β-gal	β-galactosidase
bp	Base pairs (DNA)
BSA	Bovine serum albumin
CDE	Cell-cycle dependent element
cpm	counts per minute
CTD	C-terminal domain
DMSO	Dimethyl sulfoxide
DNA	deoxyribose nucleic acid
dNTP	deoxynucleoside triphosphate (dATP, dTTP, dGTP, dCTP)
EDTA	Ethylene diamine tetra-acetic acid
EMSA	Electrophoretic mobility shift assay
FCS	Fetal calf serum
GC1	GGGCGGG box
GC2	GGGGCGGGG box
GCG	Genetics computer group
GFP	Green fluorescent protein
G segment	Gated segment (DNA)
HAT	Histone acetyl transferases
HeLa	Human cervical carcinoma cells (Helen Lane)
HTETOP	Human fibrosarcoma cell line HT180, tTA-expressing topo IIα
ICB	Inverted CCAAT box
kb	kilobases
LB	Luria Bertani bacteriological media
MCS	Multiple cloning site
MDR1	Multidrug resistance gene
MEM	Eagle's minimal essential media
mt	mutated/mutant

NES	nuclear export signal
NLS	nuclear localisation signal
NF-Y	nuclear factor Y
ONPG	o-Nitrophenyl β-D-Galacto-pyranoside
PAGE	Polyacrylimide gel electrophoresis
p53	Tumour suppressor protein
PBS	Phosphate buffered saline
PBSE	Phosphate buffered saline plus EDTA
PCR	Polymerase chain reaction
pGL3B	pGL3 Basic vector
PIC	preinitiation complex
RNase	Ribonuclease
RT	Room temperature
Sp1	Specificty protein 1
Sp3	Specificity protein 3
STET	Sucrose, Tris, EDTA and triton-X buffer
SV40	Simian virus 40
T segment	Transport segment (DNA)
TAE	Tris acetate EDTA buffer
TAFs	TBP associated factors
TATA	TATA box; conserved A/T rich septameter transcriptional sequence
TBE	Tris borate EDTA
TBP	TATA binding protein
TE	Tris-EDTA buffer
TEMED	N,N,N',N'-Tetramethylethylenediamine
TEN	Tris-EDTA buffer with sodium
TIFs	Transcription initiation factors
TFIID	Transcription initiation factor complex; TBP and TAFs
TF	Transcription factor
UV	Ultra-violet light
wt	wild type

List of Figures

List of Tables

Table of Contents

Chapter 1: Introduction.

1.1 Cancer.

DNA replication and cell division involves a series of coordinated events. Each step in the cell division cycle is carefully regulated and responsive to the needs of the organism. Normal human somatic cells are mortal, dividing into approximately 50 generations before they enter a period of senescence, and then finally die. Occasionally, the mechanisms which regulate cell division break down, allowing cells to become immortal and proliferate uncontrollably. This results in the development of tumours, otherwise known as cancer.

Cancer is a disease of inappropriate cell proliferation, fundamentally permitting the existence of too many cells. This cell number excess is linked in a vicious cycle with a reduction in sensitivity to differentiation, or apoptotic signals (Collins *et al.*, 1997). Tumours are unlimited in size as they induce the formation of new blood vessels from surrounding capillaries; this is termed angiogenesis, which provides nourishment throughout the tumour. Angiogenesis is indispensable to growth, invasion and metastasis (Du *et al.*, 2004). Without it tumours would be limited in size, as the rate of cell proliferation on the outside would be equal to the rate of cell death within the centre of the tumour, due to inadaquate supply of nutrients (Folkman, 1996). Unfortunately most tumours undergo angiogenesis, which then allows them to secret hormones and metastasize (the spreading of tumour cells).

Cancer is generally caused by mutations or translocations in DNA that result in the loss or down-regulation of tumour-suppressor genes and/or the mutation or overexpression of proto-oncogenes. Due to the high proliferation rate of cancer cells, enzymes essential for cell growth and replication are up-regulated. One such enzyme is topoisomerase IIα (topo IIα), which is required to facilitate the untwisting and separation of DNA (Robert and Larsen, 1998).

1.2 Topoisomerases.

DNA topoisomerases are essential enzymes required for chromosome segregation during mitosis. Each cell has DNA organised in a three-dimensional structure, the topology of which influences the expression of the genes contained within. Therefore enzymes which alter DNA topology are crucial for correct gene expression or replication. Utilisation of DNA for normal mitogenic functions causes knotting, twisting, and interlinking of the helical strands to occur. The general role of topoisomerases is to relax DNA, preventing such topological problems that will inhibit normal cell functions, ultimately causing apoptosis.

There are three types of topoisomerase found in eukaryotes, topo I, topo II, and topo III. Each isoform has a special cavity in its structure which accommodates DNA and allows passage through the enzyme after the specific catalytic activity has taken place. This catalytic activity varies depending on the class of topoisomerase, but the mechanism of each is essentially the same. These enzymes introduce a break in the sugar-phosphate backbone of DNA, and form a phosphate ester bond between a tyrosine residue in the catalytic site of the protein and either a 3` or 5` phosphate of the broken DNA end (Baird *et al.*, 1999). Once this has occurred, intact DNA is then passed through the break relaxing a supercoil. The DNA is religated and the topoisomerase is then released.

The first topoisomerase to be discovered was *E.coli* topoisomerase I (topo I), which has roles in transcription and DNA repair. By creating a single-strand break in DNA, topo I from *E.coli* relaxes negative supercoils, whilst in eukaryote cells it can relax both positive and negative supercoils. Type one topoisomerases are divided into two subfamilies, IA and IB based on structural and mechanistic differences (Redinbo *et al.*, 1998). Depending on the type, the torsional strain of the DNA is relieved by one supercoil each catalytic cycle with topo IA or several supercoils each catalytic cycle with topo IB. Topo I does not require the hydrolysis of ATP or any other external energy source for catalysis (Wang, 1996).

Eukaryotic topoisomerase II (topo II) is a ubiquitous enzyme essential for cell survival. Without topoisomerase II cells perish, whereas, without topo I cells can remain viable. This is because topo II is critical for the process of chromosomal condensation and segregation, DNA supercoiling, decatenation, and unlinking of intertwined chromosomes during mitosis and meiosis. Topo II achieves these important roles by utilising ATP, which enables it to cut both strands of double-stranded DNA, pass another double stranded portion through and then reseal the break. Depending on the DNA substrate this effectively turns a positive supercoil into a negative, or increases the number of negative supercoils. Bacterial topo II (DNA gyrase) comprises two subunits, GyrA and GyrB which have sequence similarity to eukaryotic topo II. In mammals, topo II has two isoforms α and β which are differentially expressed (discussed Section 1.2.1), while invertebrates possess only a single form of the enzyme.

Topoisomerase III (topo III) was discovered more recently and is categorised as a subfamily of type IA topoisomerases. Like topo II, it also has two isoforms α and β encoded by different genes which have distinct tissue specificities (Ng *et al.*, 1999). Topo III has been shown to have a strong preference for single-stranded DNA (Kim *et al.*, 1998). It partially relaxes negatively supercoiled DNA during replication, requiring ATP hydrolysis to catalyze this activity. Mutation of topo III in *Schizosaccharomyces pombe* showed it is essential for cell viability as it has a role in nuclear division (Goodwin *et al.*, 1999).

1.2.1 Topoisomerase IIα and β.

Topoisomerase II is a homodimeric nuclear enzyme. Each isoform, topo IIα and topo IIβ, is encoded by separate genes located on different chromosomes. Topo IIα and β share strong amino acid sequence identity and are functionally closely related, however they are regulated differently. The topo IIα gene encodes a protein of 170 kDa (Chung *et al.*, 1989b) and is located on chromosome 17q21-22; it's expression is controlled and found to be differentially expressed in a cell-cycle dependent manner. Maximum expression occurs during the G_2/M phase, predominately in the nuclei of highly proliferating cells

(Chung *et al.*, 1989b; Giaccone *et al.*, 1995; Lang *et al.*, 1998) The topo IIβ gene is found on chromosome 3p24, is constitutively expressed by virtually all cells in the body, and the protein (approximately 180 kDa in size (Chung *et al.*, 1989b; Satherley *et al.*, 2000)) is found predominantly in the nucleolus (Giaccone *et al.*, 1995).

The human TOP2B gene (encoding topo IIβ) is alternatively spliced in different cell types, however it is not known if the resultant proteins have different physiological roles (Davies *et al.*, 1993). Lang *et al.* (1998) performed sequence alignments of the TOP2A and TOP2B genes which revealed high conservation, except at the ends of N-terminal and C-terminal regions. This suggests the isoforms evolved by a duplication of an ancestral gene (Lang *et al.*, 1998). The additional intron in the TOP2B C-terminal region is the most divergent of the two isoforms, and may be responsible for the subnuclear localisation and possibly other regulatory functions (Meyer *et al.*, 1997).

The topo II protein contains three functional domains. These are the N-terminal domain, the central catalytic domain, and the C-terminal domain. The N-terminal domain is homologous to *Escherichia coli* gyrase B (Jensen *et al.*, 1996) and contains an ATPase active site. The central catalytic core domain is homologous to gyrase A (Jensen *et al.*, 1996) and contains the active site tyrosine for DNA binding, cleavage and religation, as well as the primary dimerization interface (Baird *et al.*, 1999). The C-terminal domain is required for nuclear localisation, and may be involved in regulation of topo II activity but not for catalytic activity (Adachi *et al.*, 1997; Jensen *et al.*, 1996; Mirski and Cole, 1995). Adachi *et al.* (1997) fused the C-terminal regions of topo IIα and topo IIβ to GFP (green fluorescent protein) and found the GFP signals localised to the nucleus irrespective of which isoform CTD (C-terminal domain) was used. When the upstream region of the CTD was added the localisation became cell-cycle dependent, with topo IIα-GFP present in the mitotic nucleus, but absent from the interphase nucleus. The localisation of topo IIβ-GFP was the opposite, being predominantly seen in the interphase, and less in the mitotic nucleus. This shows cells can distinguish between the two isoforms and localise them appropriately. Taken together, these observations suggest different functions for the two isoforms.

Mirski *et al.* (2003) carried out similar experiments fusing a potential NES (nuclear export signal) with GFP, and found both topo IIα and β have an active NES NH_2-terminal proximal to the NLS (nuclear localisation signal). The positions of the functional NES in topo IIα and β are conserved between all known vertebrate sequences, signifying a functionally important role (Mirski *et al.*, 2003). Recently, Turner *et al.* (2004) not only confirmed Mirski's results, but discovered topo IIα actually has two functional NES in the full-length protein. The implications of the NES will be discussed further in section 1.4.

Potential phosphorylation sites are also found within the C-terminal region, which may have regulatory roles on enzyme activity. This regulation however, cannot be essential as the phosphorylation sites can be deleted with no significant effect. This indicates the phosphorylation may be to fine-tune activity or to alleviate negative regulation (Jensen *et al.*, 1996).

The mechanism of topoisomerase II is known to involve several steps and associated protein conformational changes (Figure 1.1). The upper jaws of the topoisomerase II enzyme binds the DNA G-segment, inducing a conformational change which brings active-site tyrosines into position. ATP binds the ATPase domains, and a series of conformational changes traps the T-segment DNA. The DNA G-segment is cleaved by a transesterification between a pair of tyrosyl residues and a pair of DNA phosphodiester bonds four base pairs apart (Berger *et al.*, 1996). These tyrosine residues become covalently linked to the phosphoryl groups at the 5` ends of the broken G-segment DNA. The base pairs then separate and the tyrosine-linked 5` DNA ends move away from each other, forming the 'gate' for the transport of another double-stranded DNA segment. The ATPase domains dimerise, and the T-segment is passed through the break in the G-segment into the lower jaws. The G-segment is religated, and the lower jaws open to release the T-segment. The ATPs are hydrolysed and ADP + P_i is released resetting the enzyme to its starting state The G-segment may be released unless it is required for another reaction, in which case it is retained in the complex.

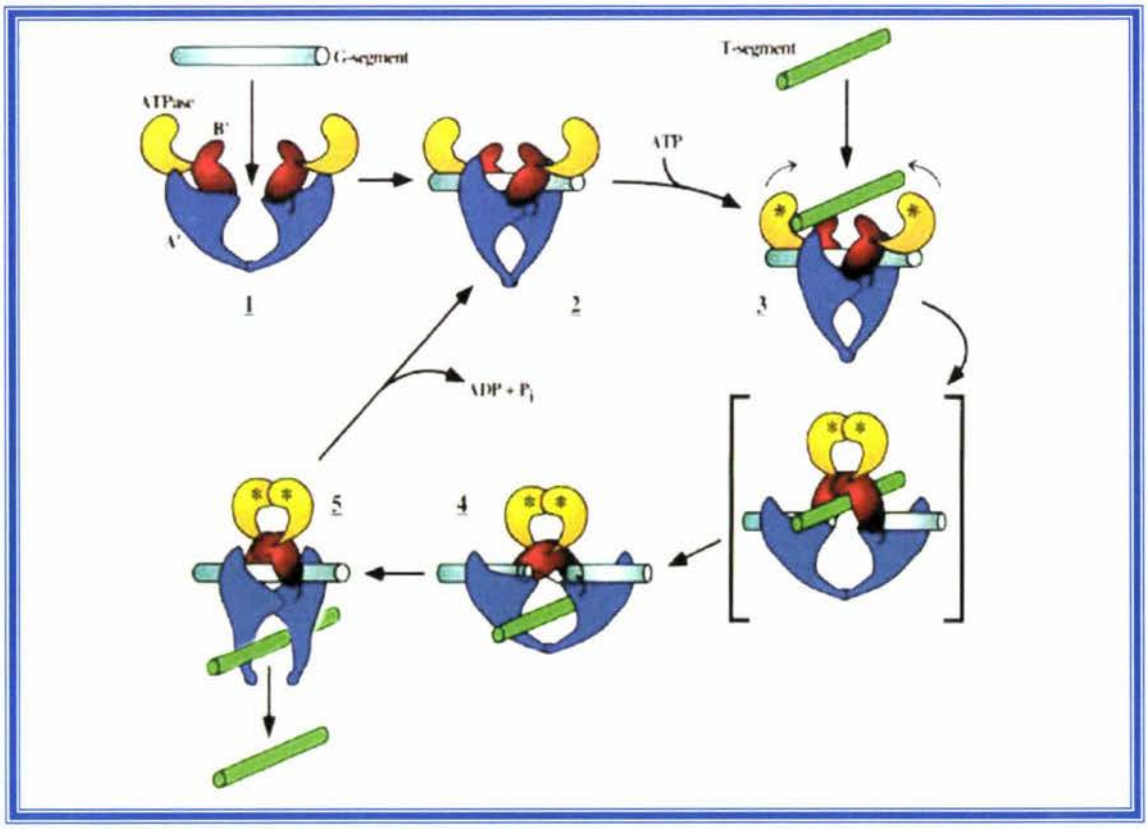


Figure 1.1: Model of topoisomerase II activity.

The ATPase domain is shown in yellow. A' and B' subfragments are shown in blue and red respectively. **(1)** Topoisomerase II enzyme binds the DNA G-segment, inducing a conformational change which brings active-site tyrosines (B' subfragment) in to attack the DNA as shown in **(2)**. **(3)** ATP (*) binds the ATPase domains, and a series of conformational changes traps the T-segment DNA. The active site splits the G-segment, the ATPase domains dimerise, and the T-segment is passed through the break into the lower jaws (A' subfragment) shown in **(4)**. For clarity the DNA transport step was shown with a hypothetical intermediate, the cleavable complex (brackets). **(5)** The G-segment is resealed, and the lower jaws open to release the T-segment. The ATPs are hydrolysed and ADP + P_i is released resetting the enzyme to its starting state (Figure from Berger *et al.*, 1996).

A conditional-lethal DNA topoisomerase IIα mutant human cell line has been generated by Carpenter and Porter (2004), under the control of tetracycline which disrupts the endogenous topo IIα by gene targeting from a transfected plasmid. The topo IIα-depleted cells still enter mitosis and go through chromosome condensation with delayed kinetics, however normal anaphases and cytokineses are inhibited, which results in cell death. These cells can be rescued by expression of a GFP fused topo IIα. This supplies new evidence that topo IIα has an essential catalytic role in chromosome segregation, but is redundant in chromosome condensation (Carpenter and Porter, 2004).

1.3 Anti-cancer Drugs & Topoisomerase IIα.

Because topoisomerase IIα is present at high levels in proliferating cells, and is essential for cell survival, a number of anti-cancer drugs have been designed with topo II as their primary target (Cornarotti *et al.*, 1996; Isaacs *et al.*, 1995). These anti-cancer drugs act as either catalytic inhibitors or poisons. Inhibitors interfere with normal topoisomerase II function causing topological problems resulting in apoptosis. The poisons function by stabilizing the topoisomerase enzyme in the cleavable complex form (the transient intermediate of double-strand nicked DNA covalently bound to topo IIα) (Figure 1.1 in brackets). This produces many double-strand breaks in the DNA of the cell, which leads to apoptosis. The efficacy of drugs is limited however by toxicity and the development of resistance.

There are many topo II targeting drugs, such as: epipodophyllotoxins etoposide and teniposide, anthracyclines doxorubicin and mitoxantrone, and amsacrines, which act to stabilise the cleavable complex. Stabilization of this complex is either sufficient to inhibit cell proliferation, or is perceived as a lethal signal for cells which enter apoptosis as a response. The drugs function by converting topo II into a cellular poison unable to religate the breaks it creates in the DNA (Figure 1.2) (Rorer, 1999). Therefore, the cellular levels of topo IIα correlate with the cellular sensitivity to topo IIα targeting drugs. The content of topo IIα in tissues and tumours can be used as a prognostic factor for drug treatment, facilitating targeted therapy.

Inhibitory topo II drugs function by halting normal topo II activity in a number of ways. Some trap the DNA-enzyme complex in a closed clamp form, others interfere with the ATPase domain or phosphorylation of the enzyme. The result of this is the accumulation of topological problems that result in cell death. The drug efficiency relies on the ability to block normal topoisomerase II activity.

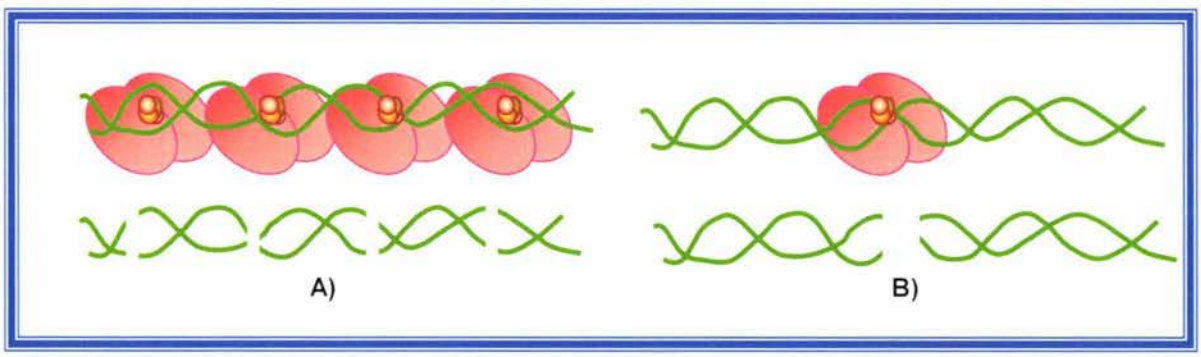


Figure 1.2: Topoisomerase II poison action.

A) Cells with high levels of topo II result in many breaks in the DNA when treated with anti-cancer drugs, which will lead to apoptosis. B) Cells with low levels of topo II have fewer breaks in the presence of the anti-cancer drugs, and therefore may survive treatment (Figure adapted from slide, Rorer, 1999).

1.4 Development of Resistance.

Resistant cell lines have been shown to have a lower amount of topo IIα, which correlates with decreased drug-induced DNA damage (Hazlehurst *et al.*, 1999; Hazlehurst *et al.*, 2001; Matsumoto *et al.*, 1997; Saleem *et al.*, 1997; Takano *et al.*, 1999; Wang *et al.*, 1997a). Therefore, the down-regulation of topo IIα activity has been proposed to circumvent the cytotoxic effect of chemotherapeutic agents, resulting in resistance (Kim *et al.*, 1996). Possible mechanisms of topo IIα down-regulation are a reduction in topo II RNA stability, mutations in the promoter region or catalytic region of the protein, change in cellular localisation or the phosphorylation status of the protein, or an alteration in protein interactions at the promoter to reduce transcription (Kubo *et al.*, 1995; Magan *et al.*, 2003).

RNA isolation and northern blot assays have shown that the topo IIα mRNA levels were decreased in resistant cell lines, and that this decrease was not due to changes in mRNA stability (Asano *et al.*, 1996; Kubo *et al.*, 1995). Analysis of the promoter sequence found no mutations present, however transcription of the gene occurred at a level of one fifth of the parent cell line. This demonstrates that the topo IIα gene is repressed in resistant cell lines at the transcriptional level (Kubo *et al.*, 1995). In 1996 however, Kubo *et al.* were the first to discover point mutations in clinical tumours. Mutations occur preferentially in two domains, either the nucleotide binding domain, or the active site tyrosine domain involved in DNA binding (Khélifa *et al.*, 1999; Renodon-Cornière *et al.*, 2003; Suda *et al.*, 2004). There have been no mutations reported to date in the promoter region.

Phosphorylation has been shown to both increase or decrease topo IIα activity depending on the differential effects at divergent phosphorylation sites (Wells *et al.*, 1995). This is reflected in work carried out in resistant cells lines with some shown to have an increase of topo IIα phosphorylation (Takano *et al.*, 1991), and others a decrease (Ganapathi *et al.*, 1996) which may slow topo IIα activity for resistant cells in the absence of reduced enzyme expression. It is known that both topo IIα and topo IIβ have many phosphorylation sites, with the majority residing in the C-terminal domain of each enzyme where the presumptive nuclear localisation signal resides. Phosphorylation may affect the cellular localisation of topo IIα, possibly transporting it from the nucleus to the cytoplasm and influencing the response to topo IIα drugs.

The distribution of topo IIα could be due to either nuclear export or decreased nuclear import. Engel *et al.* (2004) however, used leptomycin B to effectively block cytosolic accumulation of topo IIα, which proves it is exported from the nucleus. Topo IIα has been found to contain two functional NES (Turner *et al.*, 2004), the implications of this are large in regards to resistance to anti-cancer drugs. Transportation of topo IIα out of the nucleus was found to generate resistance to etoposide in human leukemic cells (Valkov *et al.*, 2000). The mechanisms of this resistance could be due to either the decrease of target enzyme in the nucleus, or the topo IIα acting as a cytosolic drug sink.

In vitro topo IIα has been shown to bind etoposide in the absence of DNA, so having topo IIα in the cytosol could cause a buffer effect, limiting the amount of drug which reaches topo IIα in the nucleus (Burden *et al.*, 1996). This would then decrease the amount of drug-induced DNA damage. Therefore, blocking the export of topo IIα with drugs such as leptomycin B may increase sensitivity to topo IIα poisons (Turner *et al.*, 2004).

Interestingly, an investigation in cell-cell adhesion found that cell-adhesion mediated protection from drug-induced apoptosis by topo II inhibitors as a result of decreased drug-induced damage, possibly caused by changes in the nuclear localisation (Hazlehurst *et al.*, 2001). This coincides with the observation that topo IIα expression is decreased in confluent cells (Isaacs *et al.*, 1996).

All these studies indicate drug resistance is achieved by cells using more than one mechanism. Mechanisms vary between different cells and types of cancer, however reduced expression and activity of topo IIα has usually been the first change observed in the process. Other adaptations follow this, contributing to a higher degree of resistance. Therefore, a more thorough understanding of the mechanisms involved in regulating topo IIα expression may lead to the development of intervention strategies to prevent drug resistance.

1.5 Eukaryotic Transcription.

Transcription of eukaryotic genes proceeds through multiple steps referred to as preinitiation, initiation, elongation, and termination. The initiation of transcription is the primary control point of gene expression. Basal transcription begins with formation of the preinitiation complex (PIC), which involves the binding of RNA polymerase II (Pol II) and a number of general transcription initiation factors (TIFs) to the transcription-control region (promoter) of the gene.

Eukaryotic cells have three types of RNA polymerases which each have different roles in transcription. RNA polymerase I and III are involved in the transcription of rRNAs and tRNAs respectively. RNA polymerase II recognises class II promoters and synthesises mRNAs of these eukaryotic genes. Class II promoters may contain any of the following elements: a TATA box, an upstream element (GC box or CCAAT box), an initiator, and a downstream element. Many promoters lack one or two of these elements. Other gene-specific DNA elements may also be found, recruiting additional transcription factors (TFs). Components bound to these elements tightly control expression levels of the genes through protein-protein interactions.

TATA boxes are highly conserved sequences found approximately 25-30 bp upstream of the transcription start site. This is the most common element in most promoters, with the consensus sequence TATAAAA. An aptly named protein, the TATA-binding protein (TBP) binds to the minor groove of the TATA box, bending the DNA to allow the recruitment of general transcription initiation factors forming a multiprotein complex (Goodrich *et al.*, 1996; Kim *et al.*, 1993). This complex phosphorylates the C-terminal domain of Pol II allowing the initiation of transcription (Cramer *et al.*, 2000).

TATA-less promoters do exist however, in housekeeping genes and developmentally regulated genes. Housekeeping genes are constitutively active as they control common biochemical pathways essential for cell survival. These promoters have GC-rich regions which compensate for the lack of a TATA box. These require the same general TFs as the promoters which contain TATA boxes utilise (Goodrich *et al.*, 1996). To form the PIC however, it is necessary for TBP-associated factors (TAFs) to assist TBP to bind TATA-less promoters (Verrijzer *et al.*, 1995). This is achieved through protein-protein interactions with proteins bound on upstream elements (GC box or CCAAT box) or initiator elements.

GC boxes and CCAAT boxes are found in a variety of promoters, usually upstream of the TATA box (termed *cis*-acting elements). Each has its own specific TFs (termed *trans*-acting factors) which bind to these elements controlling transcription initiation. TFs can

be either activators or repressors which can bind DNA elements directly or through protein-protein interactions. These proteins are not part of the PIC but are required to promote or inhibit its formation, therefore influencing transcription initiation.

Initiators are conserved sequences around transcription start sites. Additional elements are occasionally found downstream of the transcription start site. Both initiators and downstream elements are required for optimal transcription (Verrijzer *et al.*, 1995). TAFs are required to endorse TBP binding to an initiator.

1.6 Topoisomerase IIα Transcription.

Topo IIα transcription is under the control of a TATA-less promoter which is moderately GC-rich and contains a high frequency of CpG dinucleotides, all characteristic of the promoter of a housekeeping gene (Hochhauser *et al.*, 1992). A number of potential sequence motifs were identified within a 650 bp region 5' to the ATG translation start codon. Compared to the full-length (2 kb) topo IIα promoter however, maximal expression is observed with a truncated version of -617 bp, indicating upstream negative regulatory elements may be present (Hochhauser *et al.*, 1992). Analysis of the -617 bp region revealed 5 inverted CCAAT boxes (ICB1-5), two GC-rich regions (GC1 and GC2), an activator protein-2 (AP2) site, a consensus sequence for an activating transcription factor (ATF) binding site, a cell-cycle responsive element CDE/CHR, a Myc/Max site, and a Myb binding site (Figure 1.3). The Myb binding site is the most proximal to the transcription initiation site, however it is not well characterised. It is known that the binding of the proto-oncogene product c-Myb causes trans-activation of the topo IIα promoter, which is lymphoid and myeloidlineage dependent (Brandt and Kroll, 1997); however, it does not play a role in the cell cycle regulation of topo IIα expression (Adachi *et al.*, 2000). Just upstream from the Myb site is the Myc/Max site, which may bind the proto-oncogene c-Myc as a heterodimer with the transcription factor Max. Together these have been implicated in activation of genes involved in cell cycle regulation, however it is not known if they are functional in the topo IIα promoter. The

ATF and AP2 sites are also undefined in their role in topo IIα transcription. Although, Isaacs *et al.* (1996) deleted the ATF binding site, and found it has no affect on confluence-regulated topo IIα transcription. The cell-cycle responsive element CDE/CHR lies between the second and third inverted CCAAT boxes also in an inverted orientation. This is known to control expression of genes required in late S-phase of the cell cycle, in conjunction with TFs bound to regulatory elements such as GC boxes and CCAAT boxes (Isaacs *et al.*, 1998).

Transcriptional control of topoisomerase IIα is regulated through binding of TFs to these important elements (Isaacs *et al.*, 1996). TFs such as NF-Y and members of the Specificity protein family, in particular Sp1 and Sp3, are known to play a role.

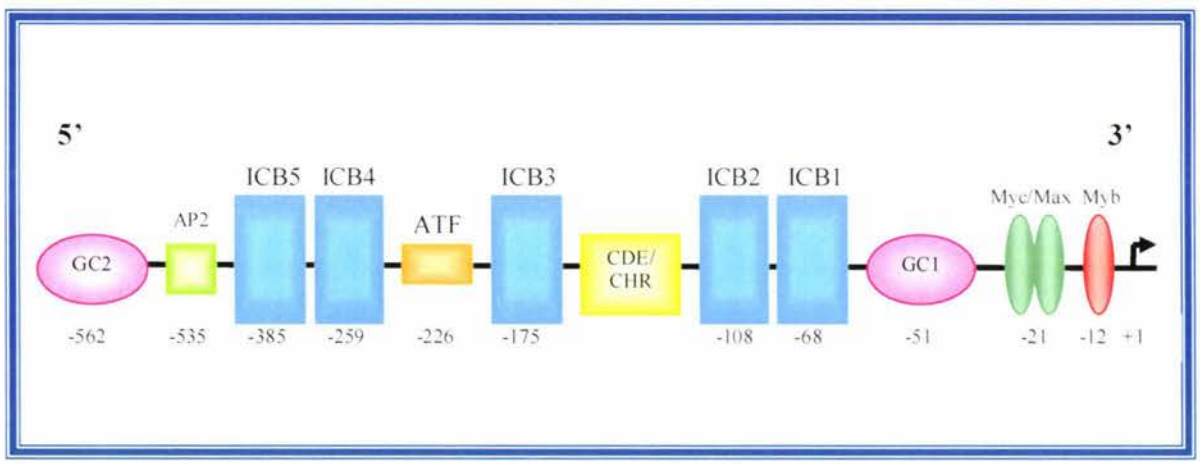


Figure 1.3: Schematic representation of the topoisomerase IIα promoter region. The major transcription start site is indicated by the arrow at position +1. Putative *cis*-acting elements are indicated by the boxes and their positions relative to the transcription start site is indicated underneath each. There are 5 ICBs represented by the blue boxes, 2 GC boxes in pink, an AP2 site in bright green, an ATF site in orange, the CDE/CHR is in yellow, the Myc/Max sites are represented in dark green, and next to them is the Myb site in red (Figure adapted from Isaacs *et al.*, 1998).

1.6.1 CCAAT Elements.

CCAAT boxes are the most common conserved sequence regulatory elements found in approximately one third of eukaryotic housekeeping and lineage-specific gene promoters. CCAAT boxes are found in the direct or inverse orientation ATTGG (termed an inverted CCAAT box (ICB)), and are important for basal-level transcription. They are usually located upstream of the transcription start site and often found in multiple copies or in conjunction with GC boxes. Specific proteins known to have the ability to bind CCAAT boxes are p53, Y-box binding protein (YB-1), ICBP90, and NF-Y.
Although a number of studies show CCAAT boxes have a positive role in gene activation, this is not always the case. The role of CCAAT boxes can be dependent on the stage of development (Fang *et al.*, 2004; Isaacs *et al.*, 1996), cell cycle phase (Adachi *et al.*, 2000), and environmental stresses (Furukawa *et al.*, 1998).

The topo IIα promoter contains five ICBs (Figure 1.3). Electrophoretic mobility shift assays using probes with each of ICB1-5, from both exponential and confluent cells, have shown there is binding of complexes to each ICB (Tolner *et al.*, 2001). However, ICB5 was not able to compete for the same proteins as those that bind ICB1-4, whereas ICB1-4 all competed for binding against each other, indicating ICB1-4 bind similar proteins but ICB5 binds alternative proteins (Herzog and Zwelling, 1997). The most distal ICBs, ICB4 and ICB5, are the least characterised out of the five, but it is known that these are dispensable for cell cycle regulation of topo IIα expression (Adachi *et al.*, 2000). In addition, ICB5 is not necessary for topo IIα regulation as constructs containing only ICB1-4 had the same level of topo IIα expression as those with all five ICB1-5 (Hochhauser *et al.*, 1992).

Isaacs *et al.* (1996) found ICB2 is the element necessary for topo IIα down-regulation in confluence-arrested cells. Site-directed mutagenesis converting ATTGG to ATTCC completely obliterated the down-regulation normally seen in the confluent cells, demonstrating this regulation is not simply due to the lack of binding of a positive regulator. Mutation of the ICB1 element did not have the same effect, indicating they are

functionally distinct. ICB1 is known to have a role in heat shock induced topo IIα expression. This was shown by mutating the ICB1 which eliminated the increased topo IIα expression normally seen after heat shock treatment (Furukawa *et al.*, 1998). Therefore, Furukawa *et al.* (1998) proposed a model in which under normal conditions a negative regulator was found associated with ICB1. After heat shock stress this regulator disassociated, allowing a positive regulator to bind and thus increase transcription. More recent studies with ICB1 mutants however, found this element was essential to achieve normal basal regulation levels (Magan *et al.*, 2003), indicative of a positive factor associated with ICB1 under normal conditions.

The p53 protein is a transcription factor involved in regulating cell-cycle progression and cell survival. It is normally expressed at low levels, but upon DNA damage it accumulates in the nucleus where it induces cell-cycle arrest to allow DNA repair, or trigger apoptosis (Ullrich *et al.*, 1992). The p53 protein is a sequence-specific DNA binding protein which is able to positively regulate transcription from promoters which contain a p53 consensus element (Farmer *et al.*, 1992). Promoters which lack this consensus element are usually negatively regulated by p53 (Mercer *et al.*, 1991). The topoisomerase IIα promoter does not harbour any known p53 consensus elements, and has therefore been found to be negatively regulated by p53 (Sandri *et al.*, 1996). The suppression of the topo IIα promoter was alleviated by consecutive deletion of the five ICBs, indicating there may be a functional interaction with specific ICBs, resulting in p53 mediated down-regulation (Wang *et al.*, 1997b). It is know known that p53 can negatively regulate transcription through a specific DNA-protein complex containing the NF-Y transcription factor which forms at a CCAAT element (Matsui *et al.*, 2004; Yun *et al.*, 1999).

The transcription factor Y-box binding protein (YB-1) is also known to bind ICB elements. It is expressed in a range of different cell types and has been implicated in the expression of genes involved in cell proliferation. YB-1 targets include genes such as the multidrug resistance gene (MDR1), proliferating cell nuclear antigen (PCNA), P-glycoprotein (P-gp), and topo IIα. YB-1 expression has been found to significantly

correlate with topo IIα expression (Gu *et al.*, 2001; Oda *et al.*, 2003). Co-transfection of an antisense YB-1 expression construct and a topo IIα promoter-luciferase reporter resulted in a reduction of promoter activity (Shibao *et al.*, 1999). This indicates that YB-1 normally plays a positive regulatory role in topo IIα expression.

Another novel protein recently identified is the inverted CCAAT binding protein of 90 kDa (ICBP90). It is predominantly found in highly proliferating cells, although in tumours such as HeLa, it is present even when confluence has been reached. Hopfner *et al.* (2000) identified this protein using a one-hybrid system, and found it's expression is concomitant with topo IIα expression. Competitor assays have shown it binds specifically to ICB2 in the topo IIα promoter. Two unidentified proteins of 90 kDa and 140 kDa have been shown to bind the ICBs in the topo IIα promoter (Herzog and Zwelling, 1997), therefore one may possibly be ICBP90 (Hopfner *et al.*, 2000). Overexpression of ICBP90 increases topo IIα expression and the rate of cell proliferation (Hopfner *et al.*, 2002). Overexpression in confluent cells prevented the normal down-regulation of topo IIα, suggestive of a role in carcinogenesis. Recently it was found phosphorylation of ICBP90 enhances binding to topo IIα and therefore increases expression (Trotzier *et al.*, 2004). ICBP90 expression is cell cycle regulated and is essential for S phase entry; because of this it is one of the targets at the DNA-damage checkpoint whose activity is down-regulated, important for cell cycle arrest (Arima *et al.*, 2004). The model proposed by Arima *et al.* (2004) implies that upon DNA-damage p53 is stabilised and activates $p21^{Cip1/WAF}$ which causes ICBP90 protein degradation through the ubiquitin/proteasome-mediated pathway. Therefore, deregulation of ICBP90 may impair control of the G_1/S transition, leading to genomic instability or carcinogenesis (Arima *et al.*, 2004).

Adachi *et al.* (2000) performed mutational analysis of ICB1-3 and found both ICB1 and ICB2 are required for full activation of topo IIα expression in the G_2/M phase of the cell cycle, as when both are mutated this activation is completely abolished. Interestingly, ICB3 is able to fully compensate when either ICB1 or ICB2 is mutated. This indicates that they have equivalent activation functions, and that either two of the three are

sufficient for this activation. These authors also suggest that the normal G_0/G_1-specific repression seen is most likely from deacetylation activity. By using trichostatin A (TSA) a histone deacetylase (HDAC) inhibitor, they saw an increase in expression in the G_0/G_1 phase, which signifies the existence of histone acetyl transferase (HAT) activity. This activity may be mediated by a protein found associated with ICB1 and ICB2, as when these were mutated there was a dramatic decrease in the TSA-induced activation seen. The transcription factor nuclear factor Y (NF-Y) has been found to bind these ICB elements and has also been shown to interact with p300, a transcriptional coactivator with intrinsic HAT activity (Li *et al.*, 1998). This suggests a model in which NF-Y binds the ICBs and recruits the HAT activating transcription in G_2/M phase, and then HDACs repress expression in G_0/G_1 (Adachi *et al.*, 2000).

Overall, the ICBs of the topo IIα promoter have functionally distinct roles in regulating topo IIα during the cell cycle, but work with each other to produce basal transcription. A variety of different proteins can bind these elements, and a number of other proteins have been seen in complexes with them. So the function of ICBs in different conditions will vary according to the associations occurring. Although in the case of topo IIα, the transcription factor NF-Y is thought to play the primary role of binding ICBs and influencing topo IIα transcription (as discussed 1.6.2).

1.6.2 NF-Y.

Nuclear factor Y (NF-Y) also termed CBF, CP1, YEBP and ACF is ubiquitously expressed in all mammalian tissues, with higher amounts in proliferating cells (Isaacs et al., 1996). They are trimeric proteins with the 3 subunits, NF-YA, NF-YB and NF-YC, all required to bind DNA (Maity and de Crombrugghe, 1998). They bind DNA at specific CCAAT pentanucleotide sequences (section 1.6.1) either in the forward or reverse orientation, which are usually located upstream of transcription start points in many eukaryotic promoters including topo IIα (Figure 1.3 and Figure 1.4) (Herzog and Zwelling, 1997; Isaacs et al., 1996; Wang et al., 1997a). NF-YB and NF-YC dimerise tightly, a prerequisite to enable NF-YA to bind (Laing and Maity, 1998). This complex can then bind DNA. The sequences flanking the CCAAT box are also important for NF-

Y binding, as mutation or deletion of these abolishes NF-Y binding (Bi et al., 1997). NF-Y cannot activate transcription alone however, and so it has been suggested that NF-Y may bind to the CCAAT box and encourage other transcription factors to associate with their targets (Hu et al., 2002).

Adachi *et al.* (2000) identified NF-Y as the inverted CCAAT box (ICB)-binding transcription factor using antibody supershift experiments. The NF-Y proteins have been shown to positively regulate topo IIα expression (Joshi *et al.*, 2003). They may achieve this positive affect by being able to distort/bend DNA upon binding by angles up to 62-82° and twist rotationally by 100° (Ronchi *et al.*, 1995). Thus facilitating interactions of proteins associated on distal elements with those associated with more proximal elements. Topo IIα expression is repressed upon inhibition of NF-Y binding the topo IIα promoter, either by mutations in the ICB sites (Magan *et al.*, 2003), expression of a dominant negative NF-YB mutant (Hu and Maity, 2000), or by p53 (Joshi *et al.*, 2003).

It is essential for NF-Y to bind for cell cycle-dependent activation of the topo IIα promoter (Hu *et al.*, 2002). The expression profile and binding capacity of NF-Y however is constant throughout the cell cycle, suggesting a more global mechanism in topo IIα promoter regulation (Adachi *et al.*, 2000). p53 can inhibit NF-Y binding the ICB in topo IIα resulting in down-regulation (Joshi *et al.*, 2003). NF-Y is also the target of p53 mediated repression of the Chk2 gene, which like topo IIα is cell cycle regulated, has a CCAAT box, and does not contain a p53 consensus (Matsui *et al.*, 2004). p53 also inhibits NF-Y binding the ICBs in the topo IIα promoter after treatment with topo IIα targeting anticancer drugs (Joshi *et al.*, 2003). In cells deficient in p53 however, NF-Y binding was still inhibited upon drug treatment, indicating a p53-independent mechanism may exist, possibly involving p53 transcriptional targets such as $p21^{Waf1/Cip}$ (Joshi *et al.*, 2003). This may in part explain some of the instances where resistance has developed through a down-regulation of topo IIα. It also highlights the fact that NF-Y may be a common target of p53. Therefore, the importance and reliance of NF-Y binding for topo IIα transcription needs to be deciphered for a better understanding of the factors that modulate topo IIα expression.

Coustry *et al.* (2001), performed *Micrococcal nuclease* (Mnase) analysis on cloned linear topo IIα promoter DNA (wild type and mutated CCAAT elements) assembled in a regular nucleosomal array. These arrays were treated either with or without NF-Y, and subjected to Mnase digestion. The results showed different patterns of digestion when NF-Y was present, indicating it had facilitated nucleosomal rearrangement. This result was lost when the mutant ICB elements were used, and enhanced when multiple ICB boxes were present, demonstrating that multiple CCAAT boxes are functionally important. Therefore, NF-Y is thought to be able to disrupt the chromatin organisation of promoters, where the specific promoter structure may have a role in NF-Y-mediated transcription activation of topo IIα (Coustry *et al.*, 2001). Coustry *et al.* (2001) also deleted the glutamine (Q)-rich domains and found DNA binding and nucleosomal disruption still occurred, however transcription was not activated. These results indicate NF-Y controls topo IIα promoter activity with two mechanisms, nucleosomal disruption and direct transcription activation. Also, the glutamine (Q)-rich domains must be necessary for transcription activation, possibly through recruitment of other TFs.

As shown by Coustry *et al.* (2001), the glutamine (Q)-rich activation domains may be important for interactions with additional TFs. NF-Y may recruit histone acetyltransferases (HATs) with these domains to the promoter region. HATs have been suggested to play a role in which NF-Y recruits them to the promoter region, where they stimulate histone acetylation and activate transcription (Adachi *et al.*, 2000; Li *et al.*, 1998). Histone acetylation activates transcription by adding an acetyl group to lysines in the histone proteins, which causes them to relax their hold on the DNA enabling binding of activating transcription factors. It is known NF-Y physically interacts with HAT homologues in both yeast and *Arabidopsis* (Stockinger *et al.*, 2001) and the HAT protein p300 (Salsi *et al.*, 2003). Therefore, it was proposed NF-Y stimulates transcription through the recruitment of HAT chromatin-modifying complexes to the promoters of target genes.

Collectively, these results show NF-Y does not act alone in topo IIα transcriptional regulation, and it is mainly a promoter organising protein, interacting with the likes of p300 or TFIID to help initiate transcription (Bellorini *et al.*, 1997). Interactions of NF-Y and p300 have also been seen to involve the transcription factor Specificity protein 1 (Sp1) (Xiao *et al.*, 2000). In accordance with this, NF-Y and Sp1 binding sites (CCAAT boxes and GC boxes) are located in close proximity to each other in a number of genes, including topo IIα and topo IIβ.

1.6.3 GC boxes.

GC boxes are commonly found flanking ICB elements in many TATA-less promoters, and are often located in close proximity within the proximal promoter. Their general consensus sequence is GGGCGGG. This is the case in the topo IIα promoter (Figure 1.3), which contains two GC boxes (termed GC1 and GC2), GC1 is located within the minimal promoter region. They are the second most frequent element found in promoters and are often present in multiple copies, where each box is proposed to have different functions. As they are so common and identified in all types of promoters, it is thought their function is not linked to any specific cellular process or mechanism of regulation (Hapgood *et al.*, 2001). This implies they may be essential for basal transcription, and have been shown to be involved in both positive and negative regulation.

The topo IIα GC2 element in rats (equivalent to GC1 in humans) has been shown to have a positive affect on transcription (Yoon *et al.*, 1999). Mutations in either GC1 or GC2 do not reduce transcription levels significantly in human topo IIα; mutations in GC2 actually increased promoter activity (Magan *et al.*, 2003). Co-transfection of Sp1 increases expression with a wt topo IIα promoter, however if GC1 or GC2 is mutated this increase is lost (Magan *et al.*, 2003). These results confirm Sp1 can be an activator which binds to GC boxes, and indicates each element has a different role. Whether the element acts positively or negatively on expression may be dependent on mutually exclusive binding of different proteins to the GC box. Both Sp1 and Sp3 are known to bind the GC boxes in human topo IIα (discussed section 1.6.4 and 1.6.5 respectively).

1.6.4 Specificity protein 1.

Specificity protein 1 (Sp1) was found to bind multiple GGGCGGG sequences (GC boxes) in the SV40 early promoter (Dynan and Tjian, 1983) and stimulate transcription from promoters containing this element (Kadonaga et al., 1987). Sp1 is ubiquitously expressed, and shown to up-regulate transcription in a variety of promoters, including topo IIα (Yoon et al., 1999). EMSA have shown Sp1 binds specifically to the GC1 box (Figure 1.4) (Magan et al., 2003) and the GC2 box (Yoon et al., 1999) of topo IIα. Sp1 contains three zinc finger structures responsible for DNA binding, and Q-rich regions (activation motifs) which are required for activating transcription. Synergistic activation can be observed either with itself or other members of the Sp1 protein family, due to protein-protein interactions between the Q-rich domains (Kadonaga et al., 1987).

The Sp1 family proteins can also distort the DNA, inducing a bend from the 3' end of the GC box towards the major groove (Sjostrom *et al.*, 1997). DNA bending however was not sufficient by itself for transcriptional activation, consistent with the Sp1-family protein Sp3, also inducing distortion but not acting as an activator. This is evidence DNA distortion may only enhance the transcription potential of Sp1, and specific protein-protein contacts need to be established through activation domains for a combinatorial effect (Sjostrom *et al.*, 1997). The topo IIα GC1 and GC2 box are at either end of the promoter, therefore DNA looping may be necessary for basal transcription. Also, GC1 is close to ICB1 which binds NF-Y and can distort DNA like Sp1, so their combined effect may bend DNA allowing proteins to act synergistically for normal levels of transcription.

Sp1's general role is as an activator of transcription. Recently however, it has been found Sp1 interacts with p53 to repress cyclin B1 expression. Mutation of both the Sp1 binding sites (GC boxes) eliminated the p53 mediated repression, and co-immunoprecipitation revealed Sp1 is bound to p53 (Innocente and Lee, 2005). Sp1 can also target downstream TFIID general transcription components through it's Q-rich domain (Chen *et al.*, 1994), therefore activating promoters transcribed by RNA polymerase II by assisting in transcription initiation complex positioning. It is thought p53 repression acts by binding Sp1 preventing this interaction with the TFIID components, therefore resulting in

repression of the promoter (Innocente and Lee, 2005). Interestingly, the Sp1 gene is positively autoregulated by its own gene product (Nicolas *et al.*, 2001).

Confirming the work of Yoon *et al.* (1999), Sp1 has a positive affect on topo IIα transcription as cotransfection of Sp1 with the topo IIα promoter increases expression, and this induction is lost when the Sp1-binding sites are mutated (Magan *et al.*, 2003). Further confirmation came from research on drug resistant cancer cell lines, which have decreased levels of both topo IIα levels and Sp1 levels (Allen *et al.*, 2004).

As Sp1 is ubiquitously expressed, post-translational modifications may be important mechanisms to control it's regulatory function. Phosphorylation of Sp1 can have either a positive or negative affect on transcription. It is thought it positively affects transcription by increasing the binding affinity to the GC box, therefore facilitating promoter activation for a number of genes (Ge *et al.*, 2001).

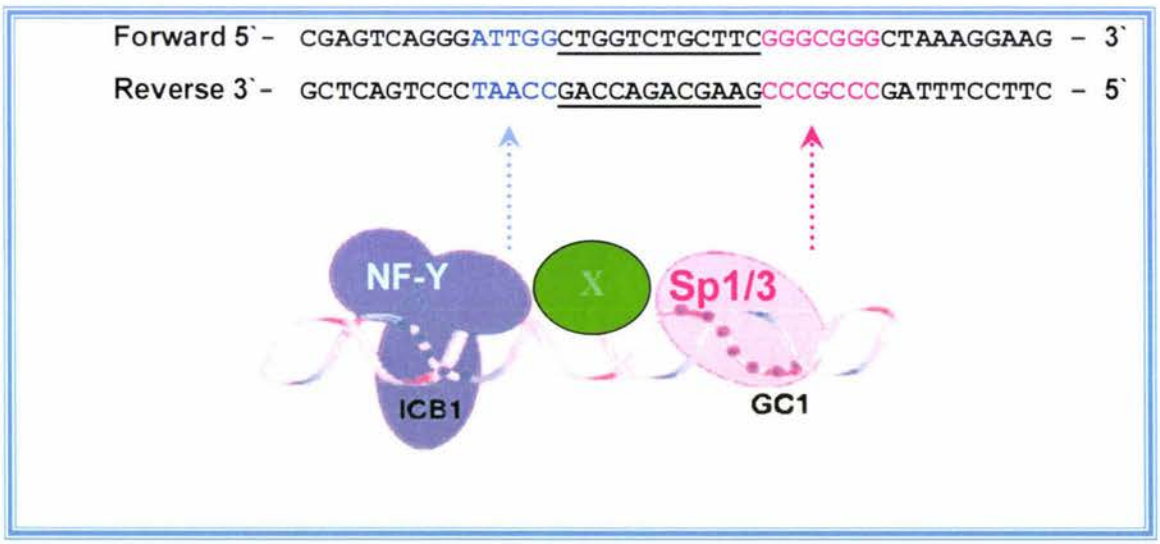


Figure 1.4: Schematic representation of ICB1, GC1 and putative transcription factors.

NF-Y is shown as the blue trimeric complex on the left associating with the ICB1 element, the GC1 element is shown to associate with Sp1 (pink) on the right. The uncharacterised protein is shown as X (green) in the centre region. The sequence of the double-stranded wild type 44 bp oligonucleotide probe used in electrophoretic mobility shift assay is shown at the top. The ICB1 sequence is in blue and GC1 sequence is in pink (adapted from Isaacs *et al.*, 1996).

1.6.5 Specificty protein 3.

Sp3 is a protein in the Sp1-family which is similar to Sp1 in many ways; it is a zinc finger protein ubiquitously expressed in mammalian cells, that can cause DNA distortion upon binding (Sjostrom *et al.*, 1997), can be phosphorylated to increase DNA binding affinity (Ge *et al.*, 2001), and is bifunctional being able to repress or activate transcription. There are three isoforms of Sp3 due to internal translation initiation sites, which has made investigations into its function difficult.

The majority of evidence to date indicates Sp3 functions as a repressor at the topo IIα promoter. Sp3 is also able to specifically bind the same sequences as Sp1 does. Therefore, the relative levels of Sp1 and Sp3 are thought to play a role in the response of a promoter to these transcription factors. If Sp3 is the same level as Sp1, it may not act as an activator as it can compete for binding with Sp1, consistent with suggestions that Sp3 represses Sp1-mediated transcription (Hagen *et al.*, 1994). Birnbaum *et al.* (1995) performed coexpression assays which confirmed Sp3 does repress Sp1 trans-activation, in a manner responsive to the Sp1/Sp3 ratio, and also dependent on the number of GC boxes present in the promoter region. Promoters containing two GC boxes are much more sensitive to Sp3 repression of Sp1 activation, than promoters with only a single GC box (Birnbaum *et al.*, 1995). Topo IIα contains two GC boxes, therefore Sp3 repression of Sp1 activation is likely to occur. In support of these suggestions, resistant cancer cell lines have an up-regulated expression of Sp3 (Kubo *et al.*, 1995) (and a decrease in Sp1 (Allen *et al.*, 2004)), correlating with a decrease in topo IIα gene expression.

1.6.6 Protein Interactions.

NF-Y and Sp1 both have Q-rich domains and can interact with particular TAFs within TFIID (Coustry *et al.*, 1998). They have been shown to interact on a variety of promoters (Ge *et al.*, 2001; Inoue *et al.*, 1999), and this may explain why the specific DNA elements associated with these two proteins are found in close proximity to each other. *In vivo* interactions between NF-YA of NF-Y and Sp1 have been shown to occur in the Q-rich domains using the yeast two-hybrid system (Roder *et al.*, 1999).

Recently, Magan *et al.* (2003) carried out an investigation of the interactions of the proteins associated with topo IIα transcription. A number of electrophoretic mobility shift assays were performed using wild type ICB1/GC1 topo IIα promoter element oligonucleotides, or mutated versions of these oligonucleotides. The EMSAs clearly showed four specific protein complexes can form with the ICB1/GC1 elements of the topo IIα promoter. Antibody supershift assays were carried out to identify which proteins were forming these complexes with the specific promoter elements. Only 3 could be identified however, these were: NF-Y, Sp1 and Sp3 (Figure 1.4). The fourth uncharacterised protein was found to specifically bind to the oligonucleotide containing the ICB1 and GC box (discussed section 1.6.7) (Magan, 2002).

Competitor assays with a wild type ICB1/mutant GC1 oligonucleotide showed that Sp1 and Sp3 were still associated with the oligonucleotide despite the GC box being mutated. Therefore, it was suggested that there is a functional interaction occurring between the NF-Y protein bound at the ICB1 element, and the Sp1 and Sp3 proteins associated together with the GC box. This correlates with other evidence that NF-Y recruits other transcription proteins to promoters. Also, the proximity of the GC1 box to the ICB1 box suggests it may be important in allowing interactions between proteins associated with these elements, as this occurs in many other promoters (Jean *et al.*, 2002).

Transient transfection assays confirmed NF-Y's role as an activator of basal transcription at the ICB1 element. Cotransfections with Sp1 showed activated transcription from binding GC1 and possible cross talk with NF-Y bound at ICB1 may occur. Models of the possible interactions and roles of the proteins associated with specific DNA elements were proposed (Figure 1.5). Figure 1.5 A) shows normal basal transcription of the topo IIα promoter. NF-Y binds the ICB1 element and interacts with Sp1/Sp3 bound at GC1. It also binds ICB2 possibly bending the DNA allowing Sp1/Sp3 bound at GC2 to interact with the other proteins bound closer to the transcription start point. Figure 1.5 B) shows that when GC2 is mutated there is an increase in transcription. This suggests the interaction of proteins normally bound to this element play a negative regulatory role in basal transcription. When ICB1 is mutated (Figure 1.5 C) there is a decrease in

expression, demonstrating the importance of this element and the associated NF-Y for normal transcription. Overexpression of Sp1 increased the activity of the promoter (Figure 1.5 D). These models demonstrate the possible interactions between these four proteins bound close to the transcription start site and other transcription factors bound to distal elements are critical for both basal and activated transcription of topo IIα (Magan *et al.*, 2003).

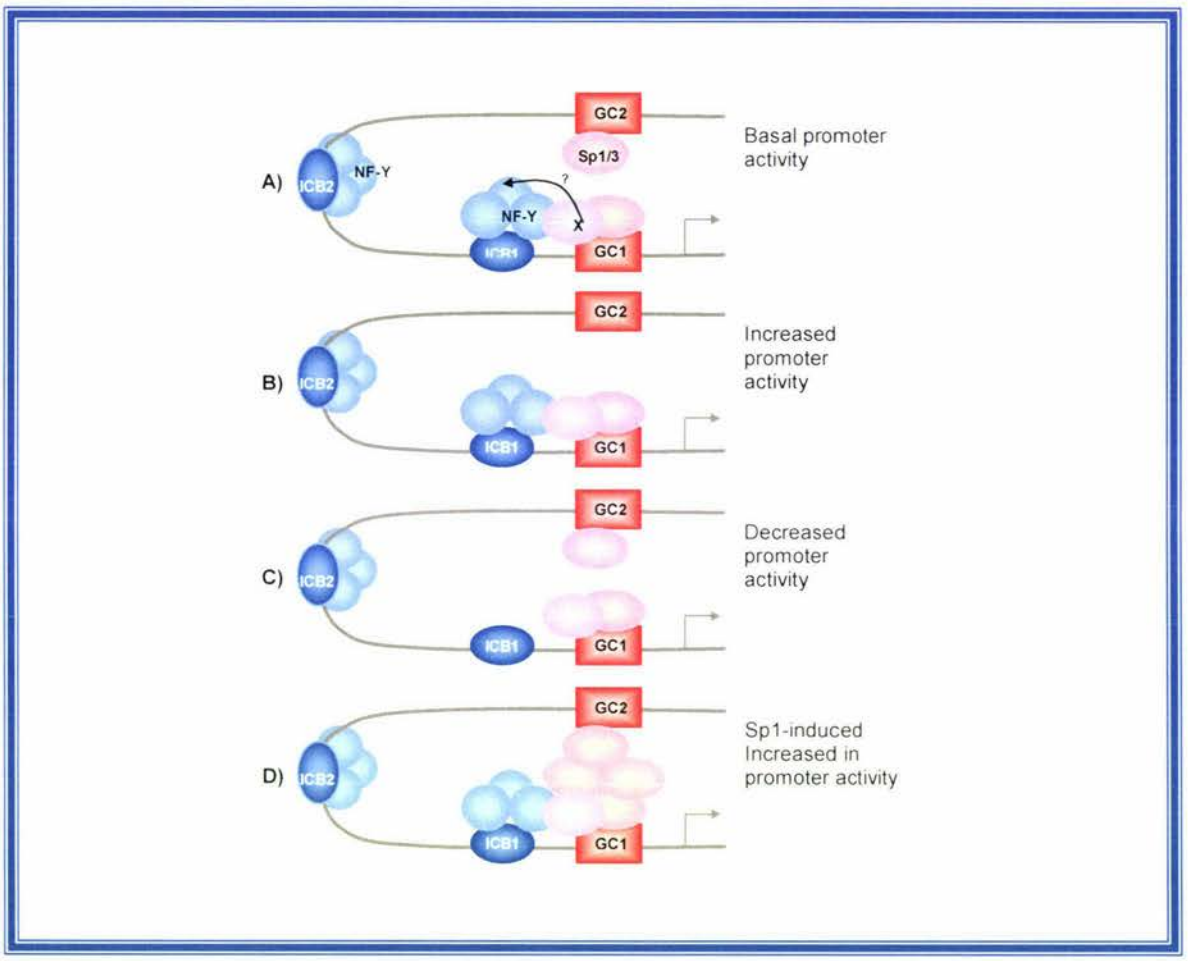


Figure 1.5: Models of transcriptional regulation of topoisomerase IIα.

X is the uncharacterised protein. **A)** Interaction between NF-Y and X is still undetermined. When Sp1 3 levels are low, it is possible that Sp1 3 preferentially binds to GC1 due to an interaction with NF-Y. **B)** A mutation in GC2 generated an increase in promoter activity; functional interaction between NF-Y and Sp1 is observed. **C)** ICB1 is an important regulatory element; mutation in ICB1 caused a decrease in promoter activity. **D)** Increase in Sp1, generally caused an increase in promoter activity; Sp1 mediated interactions between GC1 and GC2 (Figure courtesy of Magan *et al.*, 2003).

1.6.7 The Uncharacterised Protein.

The uncharacterised protein found by Magan *et al.* (2003) predominantly associated with the topo IIα promoter in a GC1-dependent manner, but required both ICB1 and GC1 elements to bind to GC1, which suggests both NF-Y and Sp1/3 may be required for binding. Alternatively, the sequence in between ICB1 and GC1 may recruit the protein.

1.7 Research Aims.

Human topoisomerase IIα regulation is relatively well understood, except for the down-regulation mechanism seen in drug resistant cancer cells. A recently discovered uncharacterised protein possibly mediates a functional interaction between several known transcription factors bound at proximal and distal elements of the topo IIα promoter. This functional interaction may have direct relevance to the development of drug resistance seen in cancer cells treated with topoisomerase targeting anti-cancer drugs.

Characterisation of the protein with respect to its function in modulating the transcription of human topo IIα, was investigated by performing electrophoretic mobility shift assays and transient transfection assays with topo IIα promoter constructs that had been mutated. This gave an indication of the requirement for the uncharacterised protein to associate with that region, and show the affect on transcription from the promoter. By deducing the activity of these transcription factors at their specific DNA regulatory elements and the role of the uncharacterised protein, a better understanding of topo IIα transcription will be gained. This in turn, could assist with improving treatment for cancer patients and preventing drug resistance developing.

The specific objectives of this research were:

- Carry out EMSAs to identify a minimal region within the topo IIα ICB1-GC1 region required for specific binding of the uncharacterised protein.
- To systematically mutate this region within an existing topo IIα promoter-reporter gene construct and test for functional significance in reporter gene assays.

Chapter 2: Materials and Methods.

2.1 Materials.

Oligonucleotides and Plasmids

Oligonucleotides were all made by SIGMA® Genosys Australia Pty. Ltd, NSW, Australia. pGL3 Basic, pSVβ-galactosidase plasmids were purchased from Promega Corporation, WI, USA. pCMV-Sport-β-galactosidase was purchased from Invitrogen Corporation, Invitrogen NZ limited, Penrose, Auckland, New Zealand. pBluescript was purchased from Stratagene, LA, USA. The -617 topoisomerase IIα in pGL3 Basic vector was originally generated by Samuel M^{c}Lenachan and Agnieszka Szremska (Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand).

EMSAs

γ^{33}P-[ATP] radioactive isotope and poly (dI.dC) was purchased from Amersham Pharmacia Biotech, UK Ltd, England. Polyacrylamide (40% w/v) and NN`-methylenebisacrylamide in water (Acrylogel) was purchased from BDH Laboratory Supplies, Poole, England. TEMED, EDTA, DTT and ficoll were all obtained from Sigma Chemical Company, St Louis, MO, USA. Antibodies NF-YA, Sp1, and Sp3 were purchased from Santa Cruz Biotechnology, CA, USA. Scintillation 2.5 cm filters, DE-81 and 3MM paper were purchased from Whatman, Maidstone, England.

DNA Manipulation and Purification.

T4 DNA polymerase, *Taq* and *Pwo* polymerase, restriction endonucleases (*Sac*I, *Hind*III, *Kpn*I, *Nco*I, *Hae*III) and buffers, DNA quantification standards, were purchased from either New England Biolabs, MA, USA or Roche, Mt. Wellington, New Zealand. Primers RV3 and GL2, dNTPs, and BSA were all purchased from Sigma Chemical Company. Agarose, 1 kb plus DNA ladder, and T4 DNA ligase were purchased from Invitrogen Corporation. Thermosensitive alkaline phosphatase (TsAP) was purchased from Life Technologies GIBCOBRL, Auckland. NuSeive agarose was obtained from BioWhittaker Molecular Applications, ME, USA. Parafilm M® was from Pechiney Plastic Packaging, WI, USA. QIAquick PCR purification kits and Midiprep plasmid kits

were both purchased from QIAGEN, New Zealand distributers: Biolab Scientific Ltd, Albany, New Zealand. NuSieve® GTG® Agarose was from Cambrex Bio Science Rockland, Inc., Rockland, USA.

Transformations

The *Escherichia coli* XL-1 blue strain was purchased from Stratagene, La Jolla, CA, USA. The genotype for this strain is F'::Tn*10 pro*A$^+$B$^+$ *lacI*^{q} Δ(*lacZ*)*M15*/*recA1 endA1 gyrA96* (Nal^r) *thi hsdR17* ($r_K^-m_K^+$) *glnV44 relA1 lac* (NEB catalogue). Bacteriological agar, and Luria Bertani broth base were purchased from Invitrogen Corporation. Ampicillin, tetracycline, RNase, lysozyme, triton X-100, PEG, and SDS were all purchased from Sigma Chemical Company. Quantum® Prep Plasmid Miniprep Kits were purchased from BioRad Laboratories, CA, USA.

Tissue Culture

Original HeLa cells were donated from Dr. Rachel Page (Department of Biochemistry, University of Cardiff). All sterile tissue culture plasticware (flasks, cell scrapers, tubes, plates) were purchased from Nunc Inc, Naperville, IL, USA. Opti-MEM, trypsin, penicillin-streptomycin, and foetal calf serum were purchased from Invitrogen Corporation. Acrocap Filter Unit (0.2 μm) sterilisation filters for cell media and Acrodisk Syringe Filters (0.45 μm and 0.2 μm) were from Pall Corporation, MI, USA. Complete™ Mini EDTA-free protease inhibitor cocktail tablets were purchased from Roche Molecular Biochemicals, IN, USA. Bradford protein assay dye reagent was purchased from BioRad Laboratories, CA, USA. DMSO was purchased from Sigma Chemical Company.

Transient Transfections

FuGene™6 transfection reagent and ONPG were purchased from Roche Diagnostics. The luciferase assay system was from Promega Corporation. FLUOstar galaxy system and computer software was purchased from BMG Labtechnologies Pty. Ltd, Melbourne, Australia.

All other materials and reagents were of an analytical grade or better.

2.2 DNA modification, purification and binding assays.

2.2.1 Oligonucleotides.

All oligonucleotides were synthesised by Sigma Genosys Australia Pty. Ltd. The oligonucleotides were provided in a dry powder form, which was then resuspended in 1x TE buffer (10 mM Tris, pH 7.5-8.0, 1 mM EDTA) to produce a final concentration of 10 µg/µL (0.65-1.4 nmol/µL), which was then stored at -20°C. From these original stocks, working stocks of 1 µg/µL, 100 ng/µL and 50 ng/µL were prepared, for use in competitor assays, EMSAs, and PCR reactions. These working stocks were then stored in either -20°C or -4°C according to the frequency of their use.

2.2.2 Electrophoretic Mobility Shift Assays.

i) Radioactively Labelling Oligonucleotide Probes.

The oligonucleotide probes were labelled with γ^{33}P-[ATP] radioactive isotope. The γ-phosphate was attached to the 5'-hydroxyl terminus of each DNA strand using T4 polynucleotide kinase (PNK). The reagents listed below were mixed together in 1.5 mL microcentrifuge tubes, and incubated at 37°C for 45 minutes.

Forward oligonucleotide 100 ng/µL ($6.5x10^{-3}$-$1.4x10^{-2}$ nmol)	1 µL
10x PNK buffer (0.5 M Tris-HCl, pH 8.0,	
0.1 M $MgCl_2$, 50 mM DDT, 0.5 mg/mL BSA)	1 µL
γ^{33}P-[ATP] (10 µCi/µL, Amersham Biosciences)	5 µL
T4 PNK (10 U/µL, Roche)	1 µL
H_20	2 µL

The labelled forward oligonucleotide was then added to a 6x excess of unlabelled reverse oligonucleotide as follows.

Reverse oligonucleotide 100 ng/μL (0.039-0.084 nmol)	6 μL
KCl (1 M)	2.5 μL
H_2O	31.5 μL

The resultant mixture was heated in a water bath at 95°C for 5 minutes. The water bath was then switched off; oligonucleotides were left to remain in the bath for ~1 hour cooling slowly to room temperature, which enables annealing to occur.
Gel shift buffer (GSB, 40 mM Tris pH 7.6, 16% ficoll, 100 mM KCl, 0.4 mM EDTA, and 1 mM DTT) was added in an equal volume (50 μL) to each of the annealed oligonucleotides. These were then gel purified immediately on a 10% polyacrylamide gel.

ii) Purification of Radioactively Labelled Oligonucleotides.

To purify the double-stranded labelled oligonucleotides from contaminants and any unlabelled, or single-stranded oligonucleotides, the total reaction volumes (50 μL annealed oligonucleotide plus 50 μL GSB) were loaded into a 10% polyacrylamide gel (37 cm long, 0.4 mm spacers) in 1x TBE (0.09 M Tris, 0.09 M boric acid, 0.02 M EDTA pH 8.0), and electrophoresed at 30 W for 1 ½ hours. After electrophoresis, the gel was wrapped in Saran wrap and exposed to X-ray (Kodak) film for 15 minutes. The labelled oligonucleotide bands were visualised on the X-ray film, and matched up with the gel, and the corresponding fragments excised. The excised gel bands were immersed in 300 μL of 50 mM KCl and incubated overnight at 37°C to elute the DNA. The microcentrifuge tube was then mixed using a vortex and centrifuged at 12,000 g for 5 min after which the supernatant was transferred to a new 1.5 mL microcentrifuge tube. Aliquots of 1 μL were removed from each for scintillation counting, and the remaining samples were stored at 4°C in a radioactive safe Perspex container until required.

iii) Scintillation Counting

The degree of labelling for each probe was estimated by scintillation counting. This was necessary to ensure approximately equal amounts of the probes were used in each assay, so the required X-ray exposure times were identical, for visualising consistent gel bands.

An amount of 1 µL removed from each probe was applied to the centre of a glass microfibre filter (Whatman® 2.5 cm, GF/A), and left to dry. The filters were then placed in a scintillation vial with 3 mL of scintillation fluid (OptiPhase 'SuperMix', Wallac Scintillation Products, Liquid Scintillation Cocktail). The scintillation fluid contains fluors which are excited by the β-rays causing them to emit light after this excitation. The light was measured using a scintillation counter (Beckman LS3801 Scintillation Counter) as counts per minutes (cpm). If necessary, probes were diluted accordingly in 50 mM KCl.

iv) Double-stranded Competitor Oligonucleotides.

Competitors were sometimes used in gel shift assays to identify specific and non-specific DNA-protein interactions. The competitor oligonucleotides were produced by adding 5 µL (0.32-0.7 nmol) of each of the complementary single-stranded oligonucleotides (1 µg/µL), heating to 95°C for 5 minutes in a Hybaid Omn-E thermal cycler, and allowing them to anneal by cooling them stepwise by 10°C every 5 minutes. It was not necessary to use any other reagents to assist in the annealing step, as the single-stranded oligonucleotides were present in relatively high concentrations, so annealing would occur without the requirement of additional salts. Once the samples reached 25°C the double-stranded competitors were stored at -20°C until required.

v) Electrophoretic Mobility Shift Assays.

Electrophoretic mobility shift assay reactions were set up in a 0.4 mL microcentrifuge tube in a total volume of 20 µL. A non-biological polymeric nucleic acid, poly dI.dC, was added in excess to each reaction. This acts as a competitor to reduce non-specific protein binding to the oligonucleotide probes.

The following components were preincubated on ice for 10 minutes:

GSB (40 mM Tris pH 7.6, 16% ficoll, 100 mM KCl, 0.4 mM EDTA, and 1 mM DTT)	10 µL
Poly dI.dC (1 mg/mL)	1 µL
H_2O (to a total volume of 20 µL)	x µL
HeLa nuclear extract (~10 µg protein)	2-4 µL

Competitor double-stranded DNA (5-100 ng) or antibodies (2-3 µL of 200 µg/mL) were added at this stage depending on the type of assay being performed. After incubation, 1-4 µL of labelled oligonucleotide (~100000 cpm, ~1 ng DNA) was added, and the mixture was left at RT for 15 minutes. 20 µL of each reaction was loaded onto a non-denaturing 4% polyacrylamide gel in 0.25x TBE and electrophoresed at 200 V for approximately 1 hour (BRL V15.17 apparatus with fitted 0.75 mm spacers). Loading dye (~5 µL) was added to a control lane to visual the distance travelled through the gel during electrophoresis. The gel was then transferred onto DE-81 paper and dried for approximately 20 minutes at 80°C in a Bio-Rad Gel Dryer 583, and exposed to a phosphoimager plate overnight. The phosphoimager plate was then scanned in a Fujifilm FLA-5000 phosphoimager. The gels were then also exposed to X-ray film (Kodak) for at least 60 hours at -70°C within a radioactive safe cassette. The X-ray films were then developed using a 100Plus™ Automatic X-ray film processor in a dark room.

vi) EMSA Photo Stimulated Luminescence Quantification Profiles.

To measure the signal intensity of bands observed in EMSAs, PSL (photo stimulated luminescence) profiles were performed using the Fujifilm FLA-5000 phosphoimager Image Gauge 4.0 software. Within the lane of interest, an area containing the desired bands for quantification was boxed off and PSL counts were collected per mm^2. A blank area the same size was also measured and subtracted from the PSL counts, to eliminate background noise. This was repeated for other lanes using the exact same area sizes to allow comparisons between the profiles. Quantification results were then graphed using the average PSL counts across the width of the boxed area on the y-axis and the length of the boxed area (mm) on the x-axis.

2.2.3 Agarose Gel Electrophoresis.

Agarose gels were made with 1% agarose and 1x TAE (40 mM Tris-acetate, 2 mM ethylene diamine tetra-acetic acid (EDTA), pH 8.0) according to the methods in (Sambrook and Russell, 2001). Electrophoresis was carried out for ~1 hour at 80 V. Ethidium bromide (0.5 µg/mL) was incorporated into the gel to allow visualization of the

DNA bands. The gel was then placed in a transilluminator where the DNA bands were visualised under UV light.

2.2.4 Purification of DNA Fragments by Freeze-squeeze.

DNA fragments were separated by electrophoresis in 1% agarose gel (section 2.2.3). The bands of interest were visualised under UV light and then precisely excised with a scalpel blade. The agarose was then wrapped in parafilm and frozen at -70°C for at least 1 hour. The edges of the parafilm were wrapped over and placed on a hard service where it was then squeezed, to squash all the agarose. The liquid that came out was collected and 50 µL TAE was added to the agarose, which was then refrozen for 30 min. This was then squeezed again and the remaining liquid which was expelled was pooled with the first lot. Traces of agarose were removed by centrifugation at 12,000 g for ~10 min; the supernatant was then transferred to a fresh microcentrifuge tube. Ethanol precipitation was then carried out to increase the DNA concentration in solution.

2.2.5 Ethanol Precipitation.

DNA was ethanol precipitated according to the methods in (Sambrook and Russell, 2001). The pellets were resuspended in either TE buffer or water.

2.2.6 Polymerase Chain Reaction

PCR (polymerase chain reaction) is a method used to amplify pieces of DNA between predetermined sites. Oligonucleotides complementary to each site are designed and used as primers for DNA synthesis. Each PCR cycle doubles the amount of DNA present, so a large quantity is produced, which can then be used for a number of different assays.

The following components were mixed together in a 0.2 mL microcentrifuge tube:

Forward primer (50 ng/μL, ~30 pmol)	5 μL
Reverse primer (50 ng/μL, ~30 pmol)	5 μL
dNTPs (3 mM)	
(a mixture of dATP, dTTP, dGTP, and dCTP)	5 μL
10x PCR buffer	
(100 mM Tris-HCl, 500 mM KCl, pH 8.3, Roche)	5 μL
Polymerase (5 U/μL, Roche)	0.25 μL
Template (1 ng/μL)	1 μL
H_2O (to make a total volume of 50 μL)	28.75 μL

The polymerase was added last to each of the tubes and then mixed well. The polymerase used in this work was *Pwo* (Roche) polymerase as it has a higher fidelity than others. The tubes were then put into the Applied Biosystems, GeneAmp® PCR System 2700 PCR machine for thermocycling. Annealing temperatures were specific to the primers used (Appendix 4).

The PCR conditions were:

95°C for 5 min

Denaturing	95°C for 1 min.	30x cycles.
Annealing	50-60°C for 1 min.	
Extension	72°C for 1 min.	

2.2.7 PCR Mutagenesis.

PCR mutagenesis (Ho *et al.*, 1989) was used to create mutations in a wild-type DNA sequence. This was achieved by designing internal oligonucleotide primers containing specific mutations. Separate PCR reactions were used to amplify DNA on either side of the mutations. The products from each were gel purified (2.2.4), then pooled together and used as templates with only the external primers to amplify the whole length of DNA. This resulted in the whole length of DNA being amplified with the mutations created internally.

2.2.8 PCR Purification.

PCR products were purified using QIAquick PCR purification kits (QIAGEN) according to the manufacturer's instructions. These have silica-based spin cartridges which bind double-stranded DNA whilst the unwanted products, such as excess dNTPs, primers and buffer salts are washed out. The PCR products were then quantified using either agarose gel electrophoresis or spectrophotometry.

2.2.9 DNA Quantification.

(i) Gel Electrophoresis DNA Quantification.

Gel electrophoresis can be used to quantify DNA by running samples along side known DNA quantification standards. Varying dilutions of the query sample were made up with TE buffer, and 1 µL of each was mixed with 10% loading dye and loaded into a 1% agarose gel (section 2.2.3). Along side the query samples, 5 µL of each of the standards were loaded (5 ng/µL, 10 ng/µL, 20 ng/µL, 50 ng/µL, and 100 ng/µL) into the gel. These DNA standards were produced from linearised plasmid DNA (pBluescript SK-II). After electrophoresis, the fluorescence intensity of the query samples were compared to the known quantification standards, whereby the concentration of the original sample was estimated.

(ii) Spectrophotometric DNA Quantification.

The concentration of DNA samples was also measured using the Nucleic Acid Sample program in the Nanodrop® ND-1000, Nanodrop Technologies computer software. Without any dilutions, 2 µL of the query sample was loaded directly onto the pedestal, and absorbances were read at intervals between 230-350 nm. The concentration was then automatically calculated using the relationship: A_{260} x 50 µg/mL (at 260 nm the absorbance of 50 µg/mL double-stranded DNA is 1.0). The purity of the DNA was also analysed by calculating the ratio of A_{260}/A_{280}. A ratio of 1.8 indicates pure DNA, <1.8 indicates protein contamination, >1.8 indicates RNA (ribonucleic acid) contamination.

2.2.10 DNA sequencing.

DNA sequencing was carried out on an Applied Biosystems 3730 automated DNA analyzer by Lorraine Berry from the Allan Wilson Centre at Massey University (Institute of Molecular BioSciences). This used dideoxy cycle sequencing with BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems Inc., USA) according to the manufacturers protocol. Plasmids were sequenced using 300 ng/µL of template, whilst PCR products were sequenced using 1 ng/100 bp of template. RV3 and GL2 primers were used at a concentration of 0.8 pmol/µL.

2.2.11 Analysis of DNA Sequences.

Analyses of DNA sequences and restriction endonuclease maps were carried out on GCG software (Version 9.1; Wisconsin Genetics Computer Group, USA). The sequence alignments were performed using the ClustalW database (Version 1.82; European Bioinformatics Institute, UK; www.ebi.ac.uk/clustalw/index.html).

The TRANSFAC® database (http://its-gcg1.massey.ac.nz/gcg-bin/seqweb.cgi) analyses eukaryotic transcription factors, their genomic binding sites and DNA-binding profiles *in silico*. This database was used to analyse the new mutations generated in the DNA sequences, to ensure no new protein-binding motifs were produced.

2.2.12 Restriction Endonuclease Digests

Restriction endonuclease digests prepared on ice in 1.5 mL tubes.

Sterile water (to make total volume 30 µL)	x µL
10x buffer (specified by manufacturer)	3 µL
DNA (500-1000 ng)	x µL
Enzyme (10 U/µL)	1-10 U

Each sample was then vortexed, pulse centrifuged at 13,000 rpm and incubated at 37°C in a dry block for 1-2 hours. After digestion, rapid boil plasmid DNA samples had 2 µL of RNase (10 mg/mL) added and were incubated at 37°C for 2 min.

2.2.13 NuSeive Gel Electrophoresis.

NuSieve agarose gels were made with 3.5% NuSieve® GTG® Agarose and 1x TAE (40 mM Tris-acetate, 2 mM ethylene diamine tetra-acetic acid (EDTA), pH 8.0). This gel has high resolution and was used to visualise DNA fragments from 50-1000 bp. Electrophoresis was carried out for ~6 hours at 30 V. Ethidium bromide (0.5 µg/mL) was incorporated into the gel to allow visualization of the DNA bands. The gel was then placed in a transilluminator where the DNA bands are visualised under UV light.

2.2.14 Thermosensitive Alkaline Phosphatase Treatment of Vectors.

After restriction endonuclease digestion, vectors were treated with thermosensitive alkaline phosphatase (TsAP) to hydrolyse 5'-phosphates from the vectors' ends, which prohibits the vector from self-ligating. The vector treatment was carried out according to the TsAP manufacturer's instructions (GIBCOBRL). As TsAP contains no contaminating endodeoxyribonuclease, or exodeoxyribonuclease it is not necessary to purify the DNA again before ligations.

2.2.15 Ligations.

Both vector (pGL3B) and insert DNA ends were cut with the same restriction endonucleases to produce compatible and sticky ends, which are able to ligate together. Ligations rely on the ratio of insert ends to be higher than the vectors ends (Berger and Kimmel, 1987), so the concentration and calculation of these ends needs to be accurate. This increases the chances of the vector ligating with an insert rather than one vector ligating with another.

The amount of vector used in the ligations was 40 ng, to which insert DNA was added at a 3:1 or 4:1 molar ratio of insert to vector. The amount of insert DNA was calculated with the following formula:

$$\text{Amount of insert DNA(ng)} = \frac{\text{Amount of vector DNA(ng)} \times \text{size of insert(bp)}}{\text{Size of vector(bp)}} \times \frac{\text{insert}}{\text{vector}} \left(\text{molar ratio}\right)$$

Ligations were carried out using T4 DNA ligase and the manufacturers rapid ligation protocol (GIBCOBRL). After incubation on ice for 5 min, half the final volume of the ligation was added to *E.coli* XL-1 blue competent cells for transformation.

2.3 Transformation of E.coli XL-1 blue cells.

2.3.1 Transformation of Escherichia coli (E.coli) XL-1 cells.

All bacterial work was carried out using strict aseptic technique and autoclaving all media and materials prior to their use. A sample of *E.coli* XL-1 blue bacteria from a glycerol stock, was streaked onto Luria Bertani (LB) agar (15 g/L bacteriological agar, 20 g/L LB broth mix; 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) plates containing tetracycline (10 mg/mL) antibiotic to produce single colonies. This was repeated again using a single colony to re-streak another plate, to ensure the colonies were originally from a single cell. A single colony from the second streak plate was then used to inoculate a 5 mL LB broth (20 g/L LB broth mix) containing tetracycline (10 mg/mL), this was incubated with shaking at 37°C overnight. From this broth, *E.coli* "Grow-on-the-Day" competent cells (Chung *et al.*, 1989a) were prepared on the day of the transformation as follows.

Fresh 5 mL LB broths containing tetracycline were inoculated with 100 µL each of the overnight *E.coli* XL-1 culture. These were then incubated with shaking at 37°C until the optical density of the culture measured between 0.4-0.5 at 600 nm. A single spare LB broth was used to monitor the growth as it can take anywhere from 3-6 hours depending on the strain. Once optimal density was reached, the culture was dispensed in 1 mL aliquots into 1.5 mL microcentrifuge tubes, and centrifuged at 12,000 rpm for 1 min. The supernatant was removed, and cells in the pellet were gently resuspended in 1/10 volume of ice cold Transformation and Storage Solution (TSS - 1.0 g Tryptone, 0.5 g Yeast extract, 0.5 g NaCl, 10.0 g PEG-4000, 5 mL DMSO, 5 mL $MgCl_2$, 70 mL H_2O, adjust pH to 6.5 with HCl or NaOH, then filter sterilised). The resuspended cells were then placed on ice for up to a maximum of 3 hours, to await the addition of the ligation reaction.

Half of the total ligation mix (10 µL) was added to each 100 µL of *E.coli* XL-1 competent cells, mixed gently and incubated on ice for 10 min. The transformed cells (50-100 µL) were then spread onto LB agar plates containing ampicillin (100 mg/mL), and incubated overnight at 37°C. Ampicillin was added to the agar as the selection control, which allows only those cells to grow that have been successfully transformed with plasmids that confer resistance. The agar plates were observed the next day and scored for any colony growth, indicating potentially transformed cells. Plasmids were then isolated for characterisation and sequencing, to check if the ligation and transformation was a success.

2.3.2 Plasmid Isolation.

Depending on the quantity and quality of DNA plasmids required, one of two techniques was used to isolate plasmids from the transformed *E.coli* XL-1 blue cells. Both techniques begin with inoculating 5 mL LB broths containing 5 µL ampicillin (100 mg/mL) with one single colony from the selected agar plate. These are incubated overnight at 37°C with shaking.

(i) Rapid Boil Plasmid Preparations.

Rapid boil plasmid preparations were used to isolate plasmids from *E.coli* XL-1 blue cells to preliminarily check if the ligation and transformation seemed successful. These were carried out following methods developed by (Holmes and Quigley, 1981). The DNA samples were then resuspended in 50 µL TE buffer and stored at -20°C until required.

(ii) Small Scale Plasmid Preparations.

Before using large-scale plasmid preparation kits, small-scale kits were used to isolate plasmid DNA for sequencing, to establish if the ligation and transformations had been successful. Plasmid DNA purification was achieved by pelleting 2 mL of the overnight culture to use in a Quantum® Prep Plasmid Miniprep Kit (Biorad) according to the manufacturer's instructions. These kits use an alkaline lysis (containing SDS) method to

release plasmid DNA from the cell. Quantum has a patented matrix which contains a diatomaceous earth that binds DNA whilst other impurities are washed out.

(iii) Large-scale Plasmid Preparations.

Large-scale plasmid preparations were carried out to produce high quality and quantity of DNA to use in transient transfections. The 5 mL culture was used to inoculate a 100 mL LB broth containing 100 µL ampicillin (100 mg/mL), which was incubated overnight at 37°C with shaking. Plasmid DNA purification was achieved by pelleting 50 mL of the culture to use in a QIAGEN Plasmid Midiprep Kit according to the manufacturers instructions. This system is based on an alkaline lysis method similar to the small-scale plasmid prep. The DNA is bound by an anion exchange resin whilst other impurities are washed out.

2.3.3 Glycerol Stocks of Transformed XL-1 blue cells.

Glycerol stocks of the transformed XL-1 blue cells were prepared for long term storage and use. 5 mL of LB broth (containing 5 µL ampicillin 10 mg/mL) was inoculated with a single colony and incubated overnight at 37°C with shaking. 800 µL of this culture was then mixed with 200 µL of 40% glycerol in 1 mL sterile cryotubes (Nunc) and stored at -70°C.

2.4 Tissue Culture.

All tissue culture was performed in a laminar flow hood (Crossflow 1800 with HEPA filter, Westinghouse) with aseptic technique. Cell extracts for EMSAs and cell lines used for transient transfections were HeLa cells or derived from HeLa cells. These were grown in a 37°C incubator (Jouan IG150, France) with 5% CO_2.

2.4.1 HeLa Cell Media.

HeLa cells were grown in Opti-minimal essential media (Opti-MEM) containing non-essential amino acids (GIBCOBRL, Invitrogen) which was made according to the manufacturers protocol and filter (0.2 μm) sterilized into sterile bottles in 180 mL aliquots. Before use, 2% (4 mL) foetal calf serum (FCS) and 1% (2 mL) penicillin/streptomycin (Pen/Strep, 5000 U/mL penicillin G sodium and 5 mg/mL streptomycin sulfate in 0.85% saline) were added to supplement the media. Media was stored at 4°C, but prior to use it was warmed to room temperature. FCS and Pen/Strep were stored at -20°C.

2.4.2 Beginning HeLa Cell Cultures.

HeLa cells were started from frozen stocks (10% DMSO in FCS, prepared by Natisha Magan, Institute of Molecular BioSciences) kept in liquid nitrogen. A 1 mL aliquot of HeLa cell stock was thawed and mixed with 5 mL of supplemented Opti-MEM. Cells were then pelleted by centrifugation at 2,500 rpm for 10 min, and the supernatant was removed. Another 2 mL of supplemented Opti-MEM was added to resuspend the cells. 1 mL of this was added to each of two monolayer T80 flasks containing 14 mL supplemented Opti-MEM. Cells were left to grow for 2-3 days depending on confluence before passaging.

2.4.3 Maintaining HeLa Cell Cultures.

HeLa cells were grown in T80 flasks until they reached 80-90% confluence, after which the cells were passaged into a new flask with fresh media. Cells were passaged by first

removing the old media by aspiration, and washing cells twice with 2 mL of 1x trypsin (1 mL 10x trypsin (GIBCOBRL) in 9 mL phosphate buffered saline (PBSE) (0.14 M NaCl, 2.7 mM KCl, 4.3 mM $NaHPO_4.2H_2O$ pH 7.2 plus 0.5 mM EDTA)). Flasks were left flat whilst cells became dislodged from the surface; after a few minutes, dislodgement was encouraged further by sharply tapping the side of the flask. The cells were then resuspended in 5 mL of supplemented Opti-MEM, 2 mL of this was either used to seed fresh T80 flasks, or 150 mm plates for production of cell extracts both containing 14 mL supplemented Opti-MEM. 100 µL of the resuspended cells was also used to seed each well in a 12-well transient transfection plate containing 800 µL supplemented Opti-MEM.

2.4.4 Preparing HeLa Cells for Freezing.

HeLa cell stocks were prepared after the first passage by resuspending 80% confluent cells in 5 mL FCS containing 10% DMSO. Aliquots of 1 mL were then dispensed into sterile 1 mL cryotubes and frozen slowly (by wrapping the cryotubes in plenty of tissue paper) at -70°C, to avoid damage to the cell membrane. Once they were frozen, the HeLa cell stocks were transferred into liquid nitrogen storage until required.

2.4.5 Preparation of HeLa Cell Extracts.

HeLa cell extracts were used in EMSA as the source of proteins which could potentially bind to the DNA. HeLa cells were seeded into 150 mm plates with 14 mL of supplemented Opti-MEM. Once the necessary confluence (80-90%) was reached, the media was removed and cells were washed twice quickly with 5 mL PBS (0.14 M NaCl, 2.7 mM KCl, 4.3 mM $NaHPO_4.2H_2O$ pH 7.2). 1 mL TEN buffer (40 mM Tris-HCl pH7.4, 1 mM EDTA, 0.15 M NaCl) was then added and cells were scraped off the plate using a cell scraper, and transferred into 1.5 mL microcentrifuge tubes. These were centrifuged at 12,000 g for 1 min to pellet the cells. The supernatant was discarded, and cells were resuspended in 300 µL of extraction buffer (containing Complete Mini EDTA-free Protease inhibitor cocktail (Roche), 40 mM Hepes pH 7.9, 0.4 M KCl, 1 mM DTT and 10% glycerol). The tubes were then freeze-thawed 3 times in liquid nitrogen which

completely disrupts the cell membranes releasing the contents. To remove debris, the tubes were centrifuged again at 12,000 g for 5 min at 4°C. The supernatant was then dispensed in 40 µL aliquots into 1.5 mL microcentrifuge tubes, and immediately snap frozen in liquid nitrogen. Extracts were then stored at -70°C until required.

2.4.6 Bradford Protein Assay.

The Bradford protein assay dye reagent concentrate (Biorad), was used to quantify the protein concentration in HeLa cell extracts. The reagent was diluted 1:5 in H_2O and 200 µL was added to each well of a 96-well microplate (Nunc). Each well contained either BSA protein standards (1 mg/mL diluted in H_2O to produce a range from 0-2.5 µg) or 10 µL HeLa extract (undiluted, 1:2, 1:5, 1:10, 1:20, 1:50, and 1:75 dilutions in H_2O). After the Bradford reagent was added, the samples were left at room temperature for at least 10 min to allow the colour to develop. The plate was then read in the 96 well plate spectrophotometer (anthos reader HT2 type 12 500, anthos labtech instruments, Australia) at 595 nm. A protein standard curve was constructed using the absorbance readings of the BSA protein standards (Appendix 2), and the amount of protein in the HeLa cell extracts calculated.

2.5 Transient Transfections.

2.5.1 Transient Transfections.

Transient transfections were carried out in 12-well plates. Each well contained 800 µL of supplemented Opti-MEM and was seeded with 100 µL HeLa cell suspension (from 5 mL 80-90% confluent cell stocks). These were grown overnight so they reached 50-80% confluent on the day of transfection. Cells were transfected using FuGene^{TM}6 which is a multi-component lipid-based reagent that complexes with DNA and transports it into the cell. Each transfection was performed in triplicate using a ratio of 3:1 FuGene^{TM}6: DNA according to the manufacturer's instructions. After 48 hours, HeLa cells were harvested and the cell extracts were assayed for luciferase and β-galactosidase activity.

2.5.2 Harvesting HeLa Cell Extracts for Luciferase and β-galactosidase Assays.

Media was removed from each well by aspiration, and cells were washed twice with 1 mL PBS. 160 µL of cell lysis buffer (25 mM Tris-phosphate pH 7.8, 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N,N,N',N'-tertaacetic acid, 10% glycerol, 1% TritonX-100) was added to each well, and plates were left at room temperature for 15 min. Cells were then scraped off the surface using a cell scraper and transferred into wells of a 96-well plate. Plates were then centrifuged at 2000 rpm for 5 min to remove cell debris. The supernatant was then stored in a 96-well plate at 4°C until required for luciferase and β-galactosidase assays.

2.5.3 β-galactosidase Assays.

The plasmid pCMV Sport was used as the β-galactosidase internal control, which enabled the transfection efficiency to be checked. 10 µL of transfected HeLa cell extract was transferred into a well of a 96-well flat-bottom microtitre plate, with 50 µL ONPG (*o*-nitrophenyl-β-D-galactopyranoside, 2 mg/mL in 60 mM NaH_2PO_2, 40 mM Na_2HPO_2) and 100 µL β-galactosidase assay buffer (60 mM NaH_2PO_4, 40 mM Na_2HPO_2, 10 mM KCl, 1 mM $MgCl_2$). The plate was then incubated at 37°C for 0.5-2 hours until yellow colour developed. To stop the reaction and enhance the yellow colour, 50 µL of 1 M

Na_2CO_3 was added. Absorbances of each well were measured in the 96-well plate spectrophotometer at 405 nm against a blank prepared using 10 µL cell lysis buffer.

2.5.4 Luciferase Assays.

The reporter gene used to measure the transcriptional effects of the DNA alterations, was firefly luciferase. To measure promoter activity the amount of luciferase expression was measured using a FLUOstar galaxy microplate reader (BMG Labtechnologies Pty. Ltd, Melbourne, Australia). 10 µL of transfected HeLa cell extract was transferred into a well of a white 96-well microtitre plate, with 20 µL of luciferase reagent (Promega). The amount of light emitted from each extract was detected by fibre optics in the FLUOstar reader and conveyed as actual photon counts. A flash of light proportional to the quantity of enzyme is released upon addition of the luciferase reagent, this is followed by a rapid decay giving a period of low-intensity emission. The level of luminescence was measured every second for 3 minutes. Data was accessed directly from an Excel (Microsoft 97) spreadsheet to find the maximum reading for each well. This represents the maximum amount of expression from the promoter during the 3 minute assay, and was therefore used for subsequent analyses. The blank contained 10 µL cell lysis buffer.

2.6 GMO approval codes

Approval for this work was granted by the ERMA regulatory authority.

The GMO approval codes were:

- Topoisomerase IIα promoter plasmids (pGL3B): GMO 98/MU 53
- Expression plasmids: GMO 00/MU 40

Chapter 3: Protein Binding Assays of the ICB1-GC1 Region from Topoisomerase IIα.

3.1 Introduction.

The down-regulation of topoisomerase IIα transcription (topo IIα) is known to confer drug resistance to standard chemotherapy regimens. The mechanisms responsible for this down-regulation are of great importance in design of new more effective treatments. Regulation of topo IIα has been relatively well documented, but the actual molecular determinants of transcriptional modulation in response to drug treatment are unclear. Recently, an uncharacterised protein with the potential to interact with components of the transcription machinery was found associated with known transcription factors NF-Y, Sp1 and Sp3 bound at the topo IIα promoter (Magan, 2003). These proteins bind the topo IIα promoter within the minimal promoter region (-617 bp), which contains the specific regulatory sequences ICB1 and GC1. A number of sequence specific oligonucleotide probes were designed, and used in electrophoretic mobility shift assays (EMSAs) to characterise the specific region of DNA or the protein requirements which allow binding of this uncharacterised protein. A knowledge of these requirements was necessary to physically and functionally characterise this protein.

An EMSA is a technique used to examine protein-DNA interactions depending on the electrophoretic mobility of DNA fragments under non-denaturing conditions. When proteins associate with DNA and specific protein-DNA complexes are allowed to form prior to electrophoresis, a mobility shift of the original DNA band will be seen, as the complex is now larger than the DNA oligonucleotide alone (refer Figure 3.1). Addition of antibody against a specific protein results in an even greater shift of the band. This is termed an antibody supershift assay and is useful in identifying proteins that bind specifically to DNA sequences. Another modification of the EMSA, called a competitor assay, can also aid in the identification of proteins and their specific DNA binding elements. This is acheived by the addition of excess unlabelled oligonucleotides, which effectively compete for available proteins in the reaction mixture, resulting in the disappearance of specific bands. Competitor assays are also useful for determining

relative affinities for proteins and DNA binding sites.

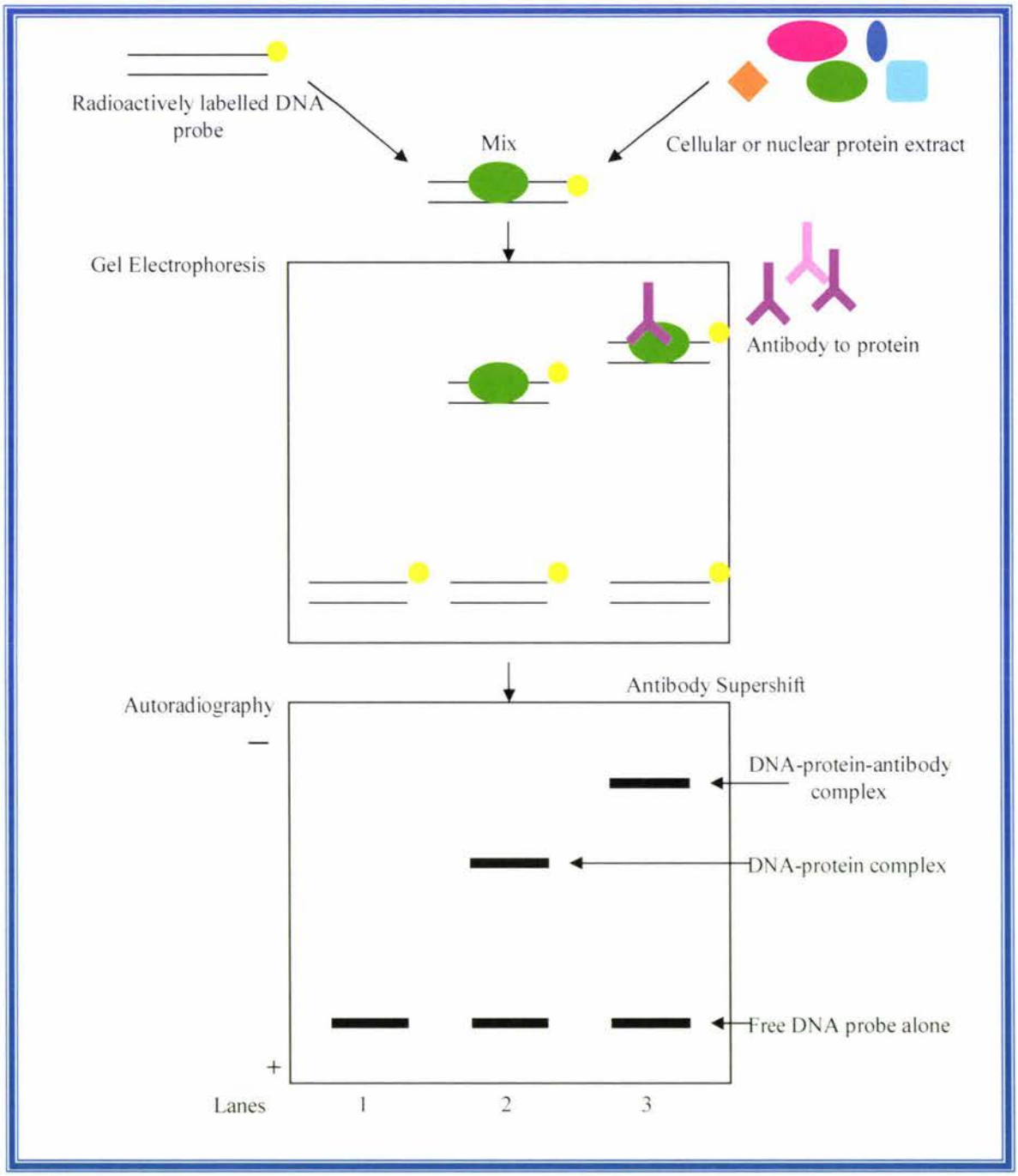


Figure 3.1: Schematic representation of electrophoretic mobility shift assays. Radioactively labelled DNA runs down the gel towards the positive electrode (lane 1). When protein extract is added, specific DNA-protein complexes can form (lane 2), hindering DNA mobility and producing a band higher in the gel. Antibodies identify proteins by binding the complex shifting the band higher, termed an antibody supershift (adapted from Magan, 2002).

3.2 DNA-protein binding assays of the ICB1-GC1 region.

The ICB1-GC1 region of the topo IIα minimal promoter has been shown to bind the transcription factors NF-Y, Sp1, Sp3 and an uncharacterised protein (Figure 1.4). NF-Y associates specifically with the ICB1 box, and Sp1 and Sp3 were found associated with the GC1 box (Magan *et al.*, 2003). This study did not differentiate between the ability of the protein to bind DNA directly or as a result of recruitment by NF-Y, Sp1 or Sp3. A number of different oligonucleotide probes were designed for this specific region of DNA (Figure 3.2 and Figure 3.3), to study the binding patterns of the uncharacterised protein.

The oligonucleotide probes were double-stranded and approximately 44 bp in length, with the exception of the deletion, insertion and centre probes (refer to Appendix 1.1 for full sequences). The mutant oligonucleotide sequences were analysed using the TRANSFAC® database (http://its-gcg1.massey.ac.nz/gcg-bin/seqweb.cgi) to ensure no new protein-binding motifs were produced (section 2.2.11). The first set of oligonucleotides were designed to alter the centre region between the ICB1 and GC1 elements (Figure 3.2), as this was a potential binding site for the uncharacterised protein. Therefore, these probes specifically tested where and whether the uncharacterised protein could bind to the DNA, they were aptly named the "uncharacterised protein binding probes". A second set of oligonucleotides was designed to test whether the uncharacterised protein was possibly being recruited by another transcription factor bound at an adjacent site (Figure 3.3); these were termed the "uncharacterised protein recruitment probes".

The TRANSFAC® database was also used to search for putative transcription factor binding sites within the 12 bp centre region between the ICB1 and GC1 elements, however none were identified.

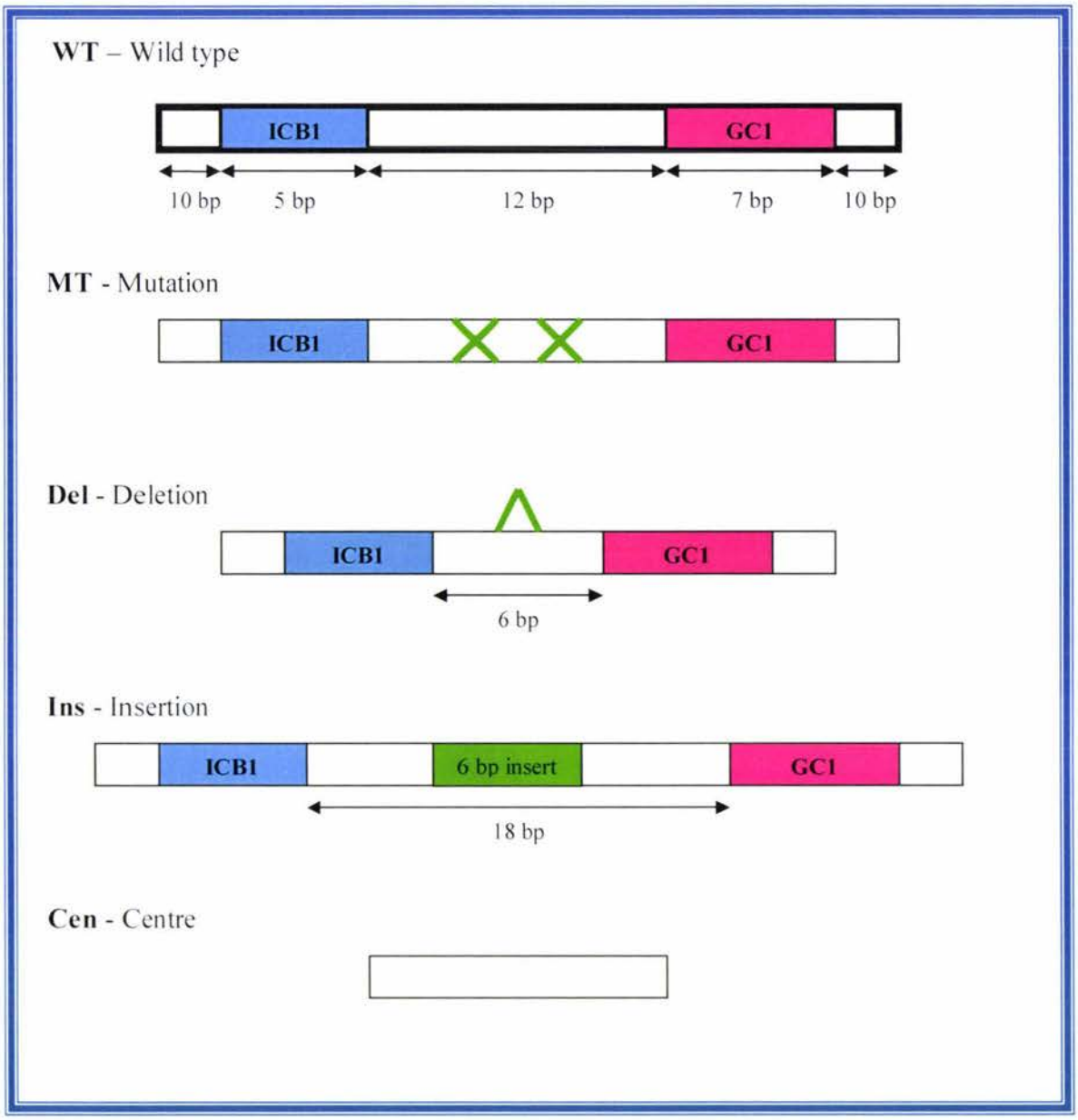


Figure 3.2: Schematic representation of the uncharacterised protein binding probes. Double-stranded oligonucleotide probes were designed to test where the uncharacterised protein binds the DNA within the ICB1-GC1 region. Indicated in bold is the **WT** (wild-type) oligonucleotide, which was 44 bp in length, encompassing the ICB1 and GC1 elements and 10 bp either side, and spanning the 12 bp region between. The **MT** (mutant) oligonucleotide had a two base pair substitution within the centre region, the **Del** (deletion) oligonucleotide had a 6 bp deletion within the centre region, and the **Ins** (insertion) oligonucleotide had a 6 bp insertion within the centre region. The **Cen** (centre) oligonucleotide was the 12 bp centre region alone.

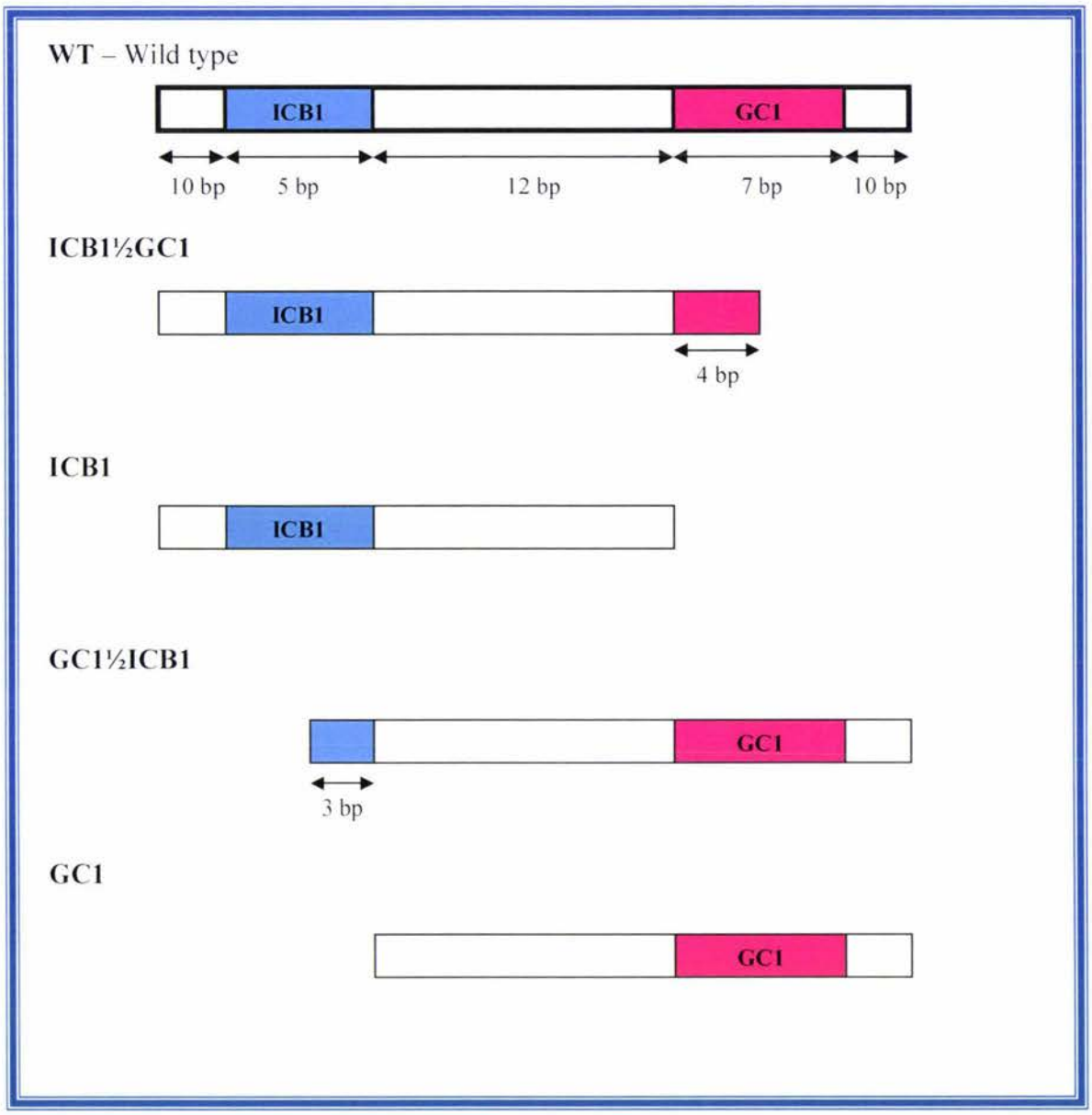


Figure 3.3: Schematic representation of uncharacterised protein recruitment probes.

Double-stranded oligonucleotide probes were designed to test if the uncharacterised protein was being recruited to the ICB1-GC1 region by another transcription factor. Indicated in bold is the **WT** (wild-type) oligonucleotide, which was 44 bp in length, encompassing the ICB1 and GC1 elements and 10 bp either side, and spanning the 12 bp centre region between these. The **ICB1½GC1** oligonucleotide had half the GC1 box deleted, and the **ICB1** oligonucleotide had the entire GC1 box deleted. The opposite of these were the **GC1½ICB1** oligonucleotide, which had half the ICB1 box deleted, and the **GC1** oligonucleotide which had the entire ICB1 box deleted.

3.2.1 Producing the ^{33}P radioactively Labelled Oligonucleotide Probes.

Each DNA oligonucleotide was commercially synthesised. The single-stranded forward oligonucleotides were radioactively end-labelled with γ^{33}P-[ATP] (refer 2.2.2(i)), and annealed to their unlabelled complementary reverse strands. These double-stranded oligonucleotide probes were then gel purified using PAGE (polyacrylamide gel electrophoresis) (refer 2.2.2(ii)), to remove any contaminating single-stranded oligonucleotides or double-stranded oligonucleotides resulting from failed synthesis (Figure 3.4 & Figure 3.5). There are unmeasurable losses which occur during the gel purification process, so the amount of DNA probe can only be estimated. The amount of radioactivity incorporated in to the probes could be measured using scintillation counting (refer 2.2.2(iii)) with the counts ranging from 20,000-200,000 cpm (Table 3.1). Before their use in EMSA's each probe was diluted to give a scintillation count of approximately 100,000 cpm/µL (approximately 1-2 ng DNA) so the autoradiography exposure time was equivalent for all probes.

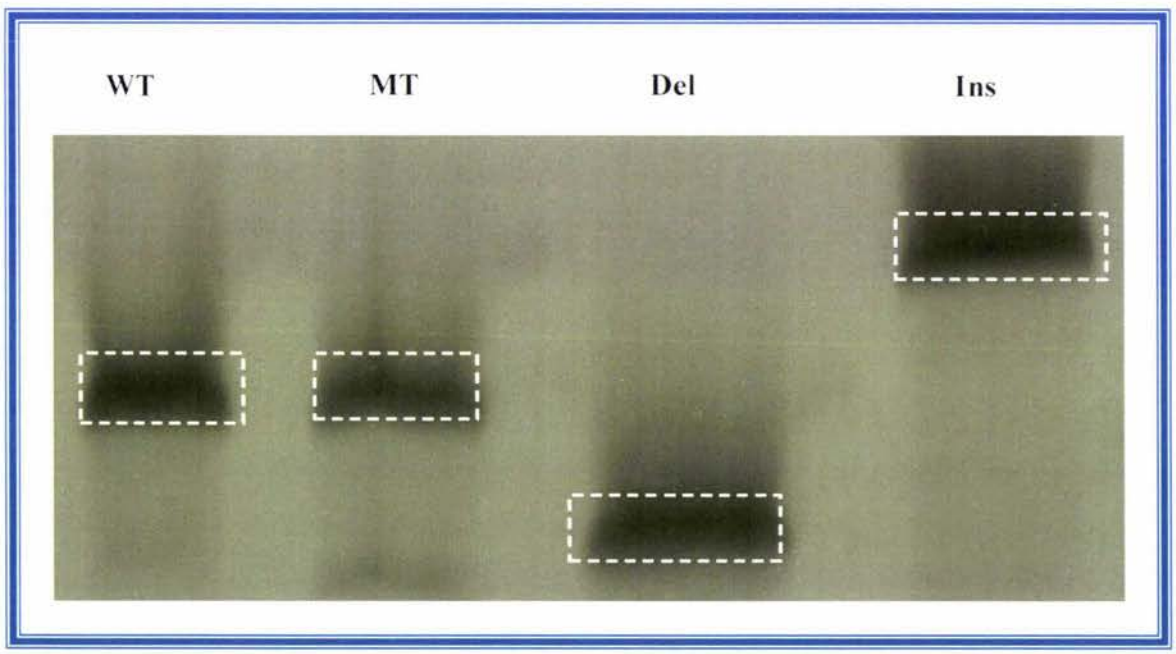


Figure 3.4: Gel purification X-ray film of ^{33}P labelled binding probes for EMSAs.

The forward strand of each oligonucleotide was end-labelled with γ^{33}P-[ATP] using T4 polynucleotide kinase. After annealing to the reverse strand, the double-stranded probes were purified in a 10% non-denaturing PAGE with 1 x TBE at 30 W for ~1 hour. Probes were visualised in the gel by exposure to X-ray film for 10 mins. Dashed lines indicate regions excised, the DNA probes were eluted from these regions in 300 µL 50 mM KCl overnight.

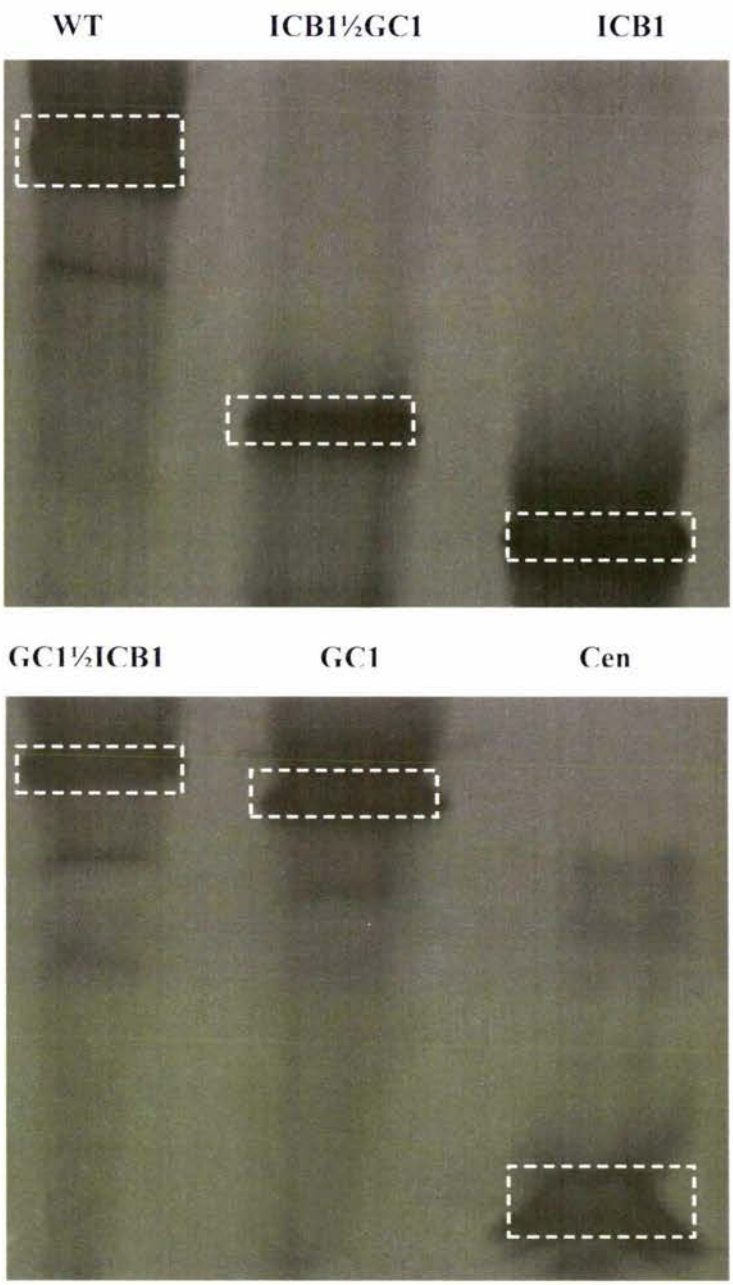


Figure 3.5: Gel purification X-ray film of ^{33}P labelled recruitment probes.

The forward strand of each oligonuleotide was end-labelled with γ^{33}P-[ATP] using T4 polynucleotide kinase. After annealing to the reverse strand, the double-stranded probes were purified in a 10% non-denaturing PAGE with 1 x TBE at 30 W for ~1 hour. Probes were visualised in the gel by exposure to X-ray film for 10 mins. Dashed lines indicate regions excised, the DNA probes were eluted from these regions in 300 µL 50 mM KCl overnight.

DNA Probe	Counts per minute (1 µL)
Blank	41
WT	110,634
MT	50,831
Del	109,252
Ins	127,836
Centre	223,390
ICB1½GC1	20,038
ICB1	150,520
GC1½ICB1	125,402
GC1	146,133

Table 3.1: Scintillation counts of ^{33}P radioactivity incorporation.

Scintillation counting was carried out using 1 µL of each probe in a Beckman LS3801 Scintillation Counter to determine the incorporation of γ^{33}P-[ATP]. Probes were used at 100,000 cpm for EMSAs (probes were diluted in 50 mM KCl where necessary), which was approximately 1-2 ng DNA.

3.2.2 HeLa Cell Extract Preparation.

HeLa cells were used for all experiments as they are known to contain the transcription factors NF-Y, Sp1 and Sp3 (Hagen *et al.*, 1994), and were also the original cells used when the uncharacterised protein was first discovered (Magan *et al.*, 2003). HeLa extracts were prepared as described in section 2.4.5, and quantified using the Bradford protein assay as described in section 2.4.6. The approximate protein concentrations used were 5.5 µg protein/µL with the binding probes (Figure 3.2), and 3.7 µg protein/µL with the recruitment probes (Figure 3.3). HeLa extract was stored at -70°C prior to use.

3.2.3 Competitor Assays.

To verify the results shown by Magan *et al.* (2003) single and double competitor EMSAs were performed using the wild-type probe, WT, which spans the ICB1-GC1 region of the topo IIα minimal promoter. Competitor assays involved the addition of unlabelled double-stranded DNA oligonucleotides at increasing concentrations (0-100 ng) to the standard EMSA reactions (2.2.2(v)). The competitors were prepared as described in section 2.2.2(iv), and used with approximately 10 µg protein extract in conjunction with 1-2 ng of radioactively labelled WT DNA probe in each reaction.

(i) Double Element Competitors.

The double-stranded double element competitors to oligonucleotides were exactly the same length as the WT probe, 44 bp, encompassing the ICB1-GC1 elements and the centre region between them (refer to Appendix 1.2 for sequences). The mutated versions of this oligonucleotide contained point mutations in either the ICB1 or GC1 elements or both (refer to Appendix 1 for each sequence). These mutations have previously been shown to prevent NF-Y and Sp1/Sp3 interactions with the DNA (Magan, 2002).

The results shown in Figure 3.6 confirm those found by Magan (2002). Four specific DNA-protein complex bands were seen using the WT probe (lanes 2, 7, 12, 17). These were most likely Sp1, NF-Y, the uncharacterised protein, and Sp3 (in descending order), and were labelled as such after confirmational results from Figure 3.10. The homologous wtICB1/wtGC1 competitor almost totally abolished all four of the DNA-protein complex bands seen in the normal EMSAs (lanes 2-5). A decrease in band intensity was seen when as little as 5 ng of competitor DNA was used (lane 3). The reverse of this was seen when the double mutant mtICB1/mtGC1 competitor was used (lanes 17-20). Thus indicating, the proteins binding were specific for the wt DNA sequence. If they were binding non-specifically, the mt competitors also would have decreased the bands' intensities. Non-specific protein binding is seen below the dashed red line, as the intensity of the bands diminish regardless of the competitors used.

The band thought to represent the transcription factor NF-Y remains relatively intense compared with the other three bands upon addition of the mtICB1/wtGC1 competitor (lanes 7-10). The strong competition for the other 3 bands indicates those proteins (Sp1, the uncharacterised protein, and Sp3) have a strong affinity for the wt GC1 sequence or the flanking sequences. The remaining NF-Y band signifies and confirms it has a specific binding affinity for the wt ICB1 sequence. When a competitor containing a wt ICB1 and mt GC1 was added, the only band strongly competed for was the second band down, again corresponding to the NF-Y band (lanes 12-15). Comparing lanes 12 and 15 it is clear some competition has occurred for the top and bottom bands (Sp1 and Sp3), as the intensity has slightly decreased. Magan (2002) also observed this occurance, and suggested it may be due to an interaction between NF-Y bound at ICB1 and Sp1/Sp3 bound at GC1.

Competition for the uncharacterised protein only occurred when a competitor was used which contained the wild-type GC1 element (lanes 3-5, 8-10). For that reason, it was thought the uncharacterised protein associates with Sp1/Sp3, the GC1 element, or it's flanking sequences.

This experiment confirms the results reported by Magan (2002) and illustrates that ^{33}P can be used as a radiolabel as all previous experiments were carried out using ^{32}P.

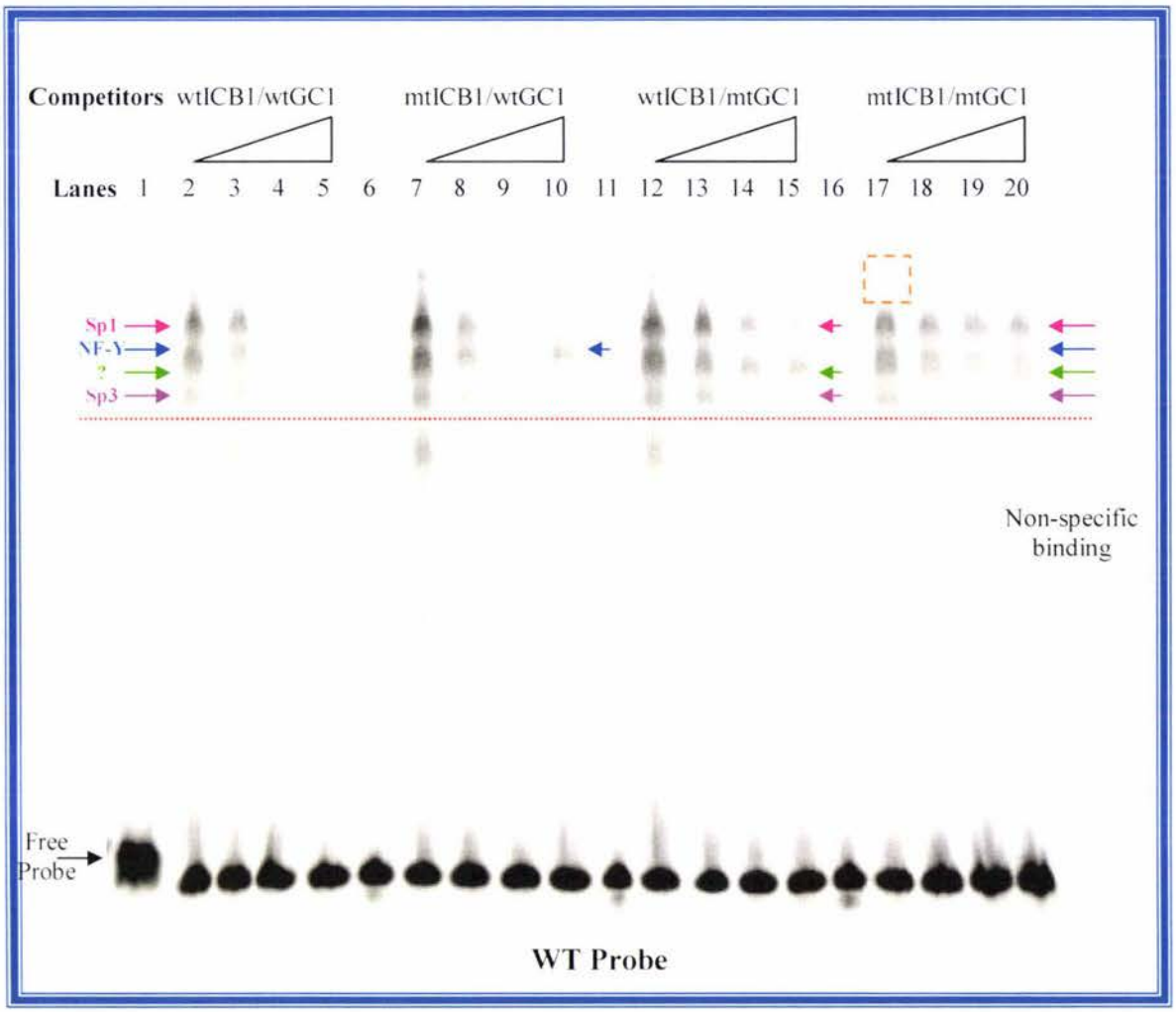


Figure 3.6: Double element competitor assay.

Double-stranded competitor DNA was added in increasing amounts (0, 5, 50, 100 ng indicated by ◢) with 10 µg HeLa extract, and 1-2 ng WT probe. Competitors used are indicated across the top. Lanes 1, 6, 11, and 16 contain WT probe alone. Each reaction was loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight.

The specific DNA-protein complexes are indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. Non-specific binding is seen below the red line and free probe across the bottom. Within the orange box is a putative multi-protein NF-Y/Sp1/Sp3 complex, which is also observed in lanes 2, 7, and 12. This figure is representative of triplicate experiments.

(ii) Single Element Competitors.

Single element competitors were ICB1 or GC1 in either wt or mt form (refer to Appendix 1.2 for each sequence). These were designed to test protein binding affinity for the single elements themselves and the dependence on flanking DNA sequences. The results shown in Figure 3.7 concur with those of Magan (2002).

As with the previous EMSA, Figure 3.7 has four specific DNA-protein complexes observed in the same locations, and non-specific binding is shown below the dashed red line. The wtICB1 competitor showed competition for the NF-Y band (lanes 2-5), however, this competition was lost with the mtICB1 competitor (lanes 7-10) indicating NF-Y has specific affinity for the ICB1 element. It is interesting to note the Sp1 and Sp3 bands were also weakly competed by the wtICB1 competitor (lanes 2-5). This coincides with results from the double competitor assays (Figure 3.6, lanes 12-15), supporting the notion of an interaction or synergism between NF-Y and Sp1/Sp3.

Weak competition was observed for the Sp1 and Sp3 bands upon the addition of the wtGC1 competitor (lanes 12-15). Consequently, Sp1 and Sp3 must preferentially bind the longer probe possibly relying on the flanking sequences. Alternatively NF-Y or the uncharacterised protein may be required to increase their binding affinity for the GC1 element. The mtGC1 competitor was unable to compete for any protein binding (lanes 16-20). The intensity of the uncharacterised protein's bands remained the same regardless of competitors used (lanes 2-5, 7-10, 12-15, 17-20).

The fact that the uncharacterised protein was not competed for by any of the single element competitors indicates it may be binding to the flanking sequences or is possibly recruited to the region through protein-protein interactions.

Both these experiments confirmed results obtained by Magan (2002), no new information was obtained.

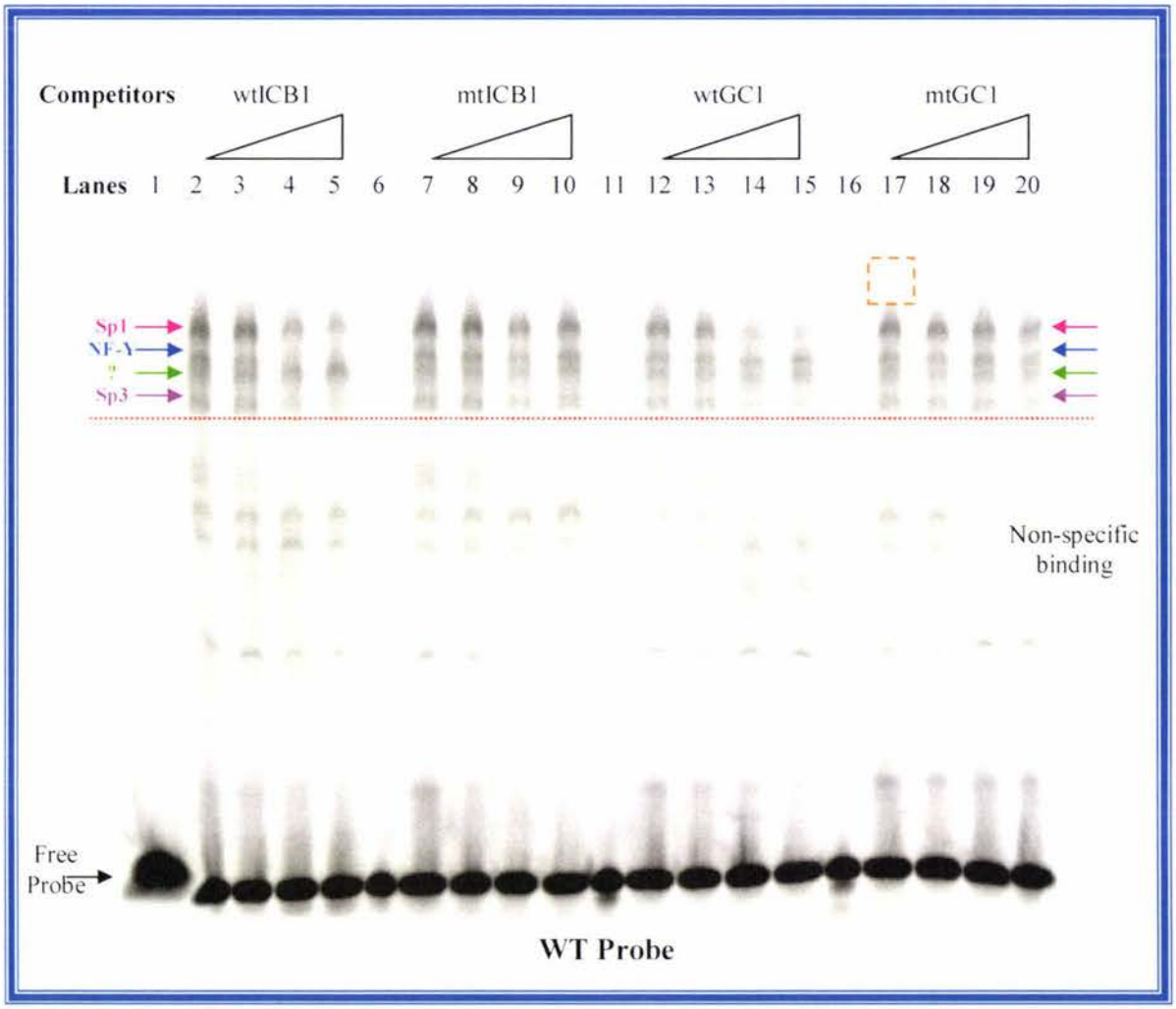


Figure 3.7: Single element competitor assay.

Double-stranded competitor DNA was added in increasing amounts (0, 5, 50, 100 ng indicated by ◢) with 10 µg HeLa extract, and 1-2 ng WT probe. Competitors used are indicated across the top. Lanes 1, 6, 11, and 16 contain WT probe alone. Each reaction was loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight.

The specific DNA-protein complexes are indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. Non-specific binding is seen below the red line and free probe across the bottom. Within the orange box is a putative multi-protein NF-Y/Sp1/Sp3 complex, which is also observed in lanes 2, 3, 7-10, 12, 13, and 17-20. This figure is representative of triplicate experiments.

3.2.4 EMSAs of the Uncharacterised Protein Binding Set of Oligonucleotide Probes.

Results of the double and single competitor assays indicated there are protein-protein interactions occurring between the proteins bound at the ICB1 and GC1 elements. The uncharacterised protein does not seem to bind specifically to either of these elements, but does tend to have an association with Sp1 and/or Sp3 bound to the full-length double element probes containing a wt GC1 (Figure 3.6, lanes 2-5 and 7-10). The single element wtGC1 competitor was unable to compete for the uncharacterised protein at all (Figure 3.7, lanes 12-15). It is therefore possible the uncharacterised protein is binding the flanking sequence in the centre region between the two elements, or may be recruited to the region by Sp1 and/or Sp3. The interplay suggested between NF-Y and Sp1/Sp3 may be mediated by the uncharacterised protein by increasing their binding affinities through correctly positioning them.

The first set of probes were designed to delineate the cognate DNA sequence for the uncharacterised protein, and to investigate the DNA sequence requirement for interactions between proteins bound at ICB1 and GC1. These probes had the centre region altered either by two base pair substitutions (MT) to test the significance of the sequence; shortened by a six base pair deletion (Del) to test the importance of the spacing between the two elements; or extended by a six base pair insertion to test the requirement for interactions across the centre region. A small (12 bp) probe of the centre region alone was produced to test if the uncharacterised protein could bind specifically to this region (refer Figure 3.2) without requiring additional flanking sequence.

Figure 3.8 shows the results of an EMSA carried out using these probes. Lane 2 shows the four DNA-protein complexes normally seen with WT probe, thought to be Sp1, NF-Y, uncharacterised protein, and Sp3 (in descending order). It is interesting to see the third band down (which is thought to be the uncharacterised protein) is absent using both the MT and Del probes (compare lane 2 with 4 & 6 red circles). This indicates the centre region sequence is important for binding, however, the centre region alone is not sufficient for binding (lane 10). The Ins probe has all four DNA-protein complexes, with a relative increase in binding of the uncharacterised protein (lane 8).

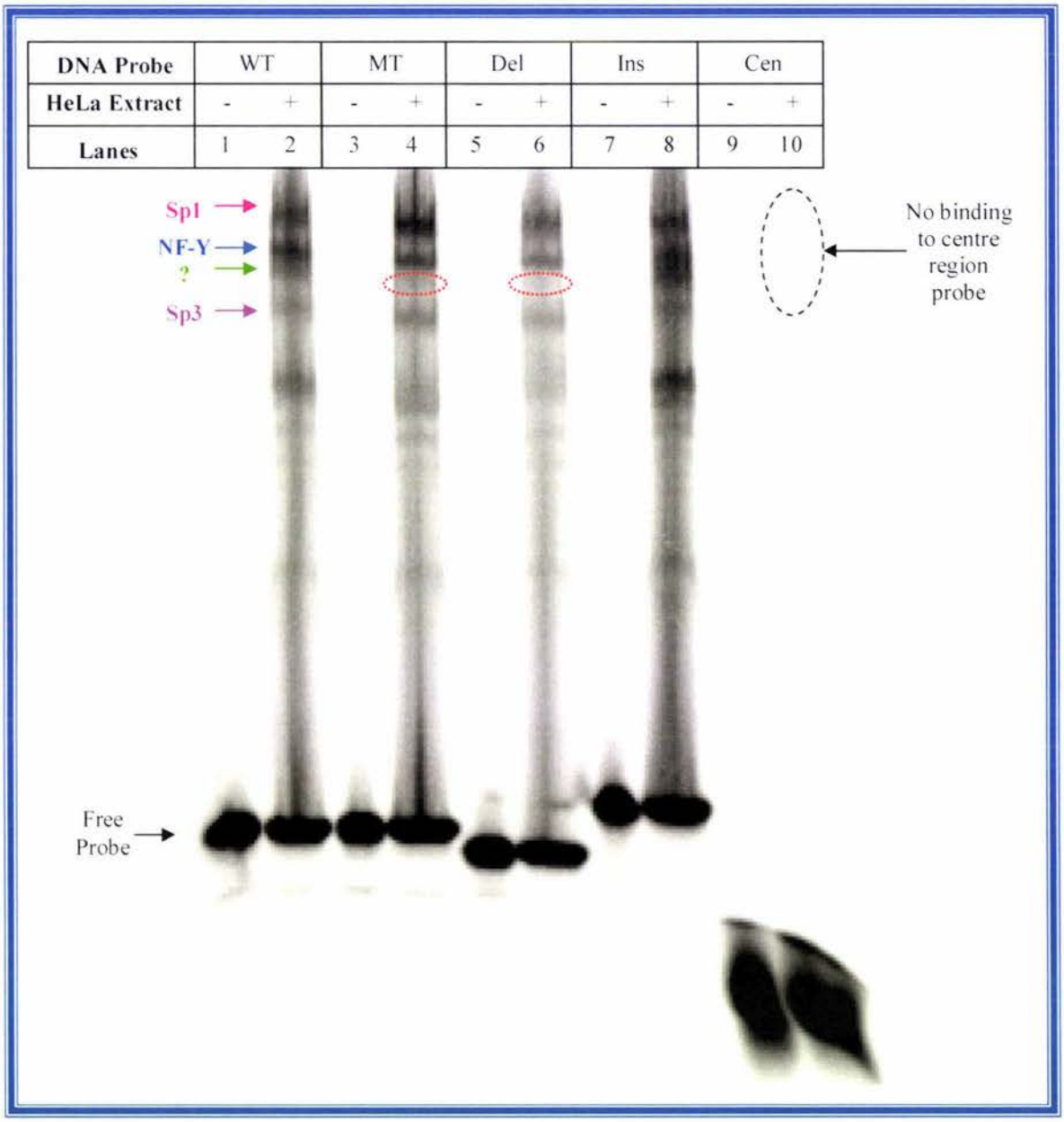


Figure 3.8: EMSA of the uncharacterised protein DNA binding probes.

Each reaction contains 1-2 ng DNA probe, and even numbered lanes contain 10 µg HeLa protein extract. These were loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight.

Four specific DNA-protein complexes seen with WT and Ins probes indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. MT and Del probes are missing a band (red ovals) no apparent binding to the Cen probe (black oval).

Quantification of EMSA signal intensities from each lane were carried out using phosphoimager analysis (refer to section 2.2.2(vi)). After background subtraction, the PSL values (photo stimulated luminescence) for each lane were plotted over distance to give the overall profile, and an indication of whether bands (representing proteins bound to the specific DNA probes) are present or absent (refer to Figure 3.9).

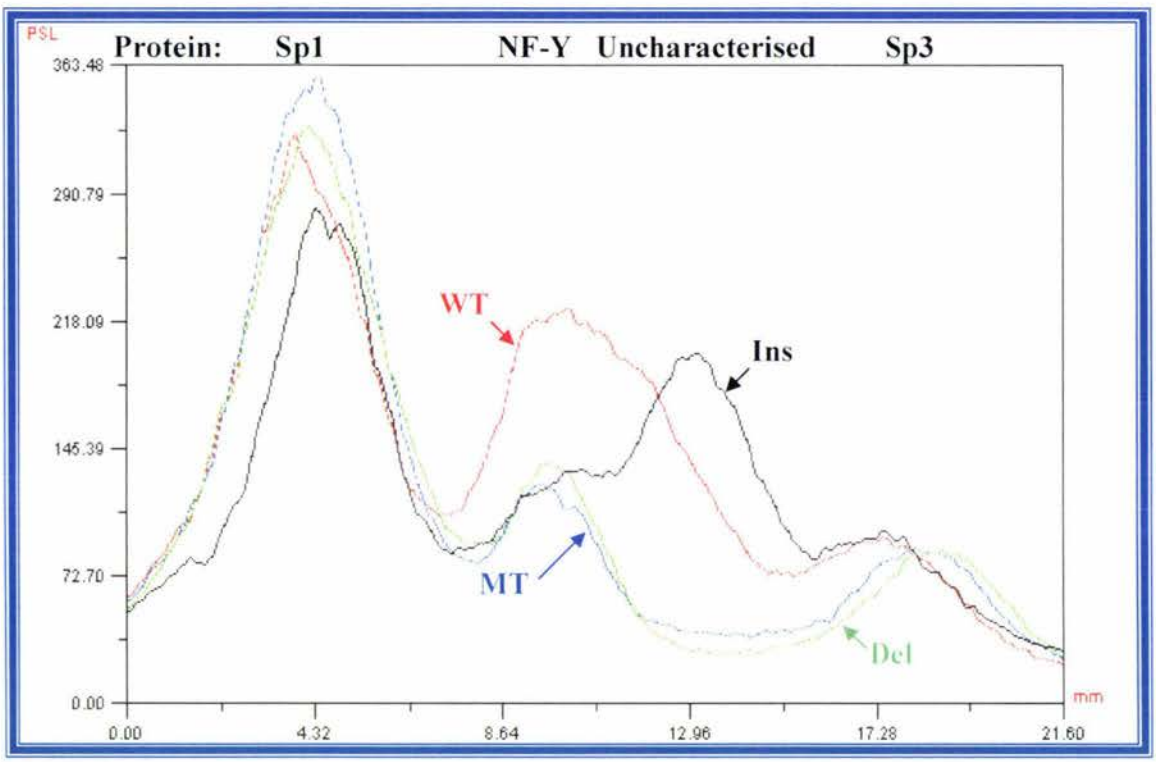


Figure 3.9: PSL profiles of proteins binding respective DNA probes in EMSAs. Phosphoimager software (Image Gauge 4.0) was used to produce a protein binding profile for each of the 4 DNA binding probes used in EMSAs Figure 3.8 and Figure 3.10. The WT DNA protein binding profile is represented by the red line, the MT is represented by the blue line, Del is in green and the Ins profile is in black. Each peak corresponds to a band in the gel which represents a protein bound to the DNA. Specific proteins are indicated across the top of the graph.

The profile for the WT probe indicated in red has four peaks produced by the PSL of the bands in the EMSA gel lane, representing Sp1, NF-Y, the uncharacterised protein and Sp3. The NF-Y and uncharacterised protein peaks are not very distinct as the bands are quite close together on the actual gel. The MT (blue) and Del (green) profiles lack the

uncharacterised protein peak and have a small NF-Y peak, thus indicating the uncharacterised protein is not present and less NF-Y protein has bound than with the WT probe. Sp1 and Sp3 appear to bind all 4 of the DNA probes with similar affinity regardless of the DNA alterations. The Ins profile shows a higher and more distinct peak for the uncharacterised protein, indicating more of the protein has been able to bind. The NF-Y peak is low as for the MT and Del profiles, indicating less NF-Y binds to each of these altered DNA sequences.

3.2.5 Antibody Supershift Assay of the Uncharacterised Protein Binding Probes.

The uncharacterised protein showed an affinity for the specific DNA sequence within the centre region, however, this was not sufficient itself to allow the uncharacterised protein to bind. As observed in Figure 3.8 and Figure 3.9 binding of the uncharacterised protein seemed to be lost when the DNA sequence within the centre region was mutated or deleted (lanes 4 & 6), and no binding was seen with the centre region alone (lane 10). To confirm these results antibody supershifts were performed to confirm the identity of the proteins present in the band seen in the EMSAs (refer to 2.2.2(v)). The antibodies added bind to their respective target proteins specifically, increasing the mass of the DNA-protein complex, which in turn causes the relative band to appear higher in the gel, therefore diminishing the intensity of the original band.

The antibody supershift (Figure 3.10) confirmed the results seen in Figure 3.8 and Figure 3.9. The probes used and the antibodies added are indicated across the top. Control lanes are 1, 6, 11, and 16 which contain probe alone, and lanes 2, 7, 12, and 17 which have probe and HeLa extract, used as controls to show the standard EMSA bands before the addition of any antibodies. The bands in these lanes are the same as what was seen in the original EMSAs (Figure 3.8) (the Cen probe was not used as it does not appear to bind protein). The remaining lanes have had 2-3 µL of the respective antibody added each at a concentration of 200 µg/mL.

The WT probe produces four distinct DNA-protein complexes (lane 2); upon addition of the NF-YA antibody the intensity of the second band has diminished and a faint DNA-

protein-antibody complex is seen higher in the gel (lane 3, blue circle). This result is also evident with MT, Del, and Ins probes (compare lanes 3, 8, 13, and 18), verifying this as the NF-Y protein band. The Sp1 antibody resulted in a shift of the top band and the appearance of a DNA-protein-antibody complex (lane 4, pink circle). Again, this was evident with all four probes used (lanes 4, 9, 14, and 19), confirming the top band as Sp1. The Sp3 protein was shown to be the fourth band down (lanes 5, 10, 15 and 20), as the presence of the Sp3 antibody almost totally abolished this band (compare lanes 2 & 5, 7 & 10, 12 & 15, and 17 & 20), and a faint DNA-protein-antibody complex appears just above the Sp1 band (lane 5, purple circle).

The results from Figure 3.8 and Figure 3.10 show NF-Y, Sp1 and Sp3 can bind the ICB1 and GC1 elements in the absence of the uncharacterised protein (Figure 3.10, lanes 7 & 12). The uncharacterised protein can not bind the centre region alone, however the DNA sequence is important for it to associate with the region. When the Ins probe is used, there appears to be more binding of the uncharacterised protein (Figure 3.8, compare lanes 2 & 8, Figure 3.10, compare lanes 2 & 17). This suggests the spacing between the elements may also be a factor, or the protein preferentially binds the inserted sequence (which was a repeat of the original sequence within the centre region).

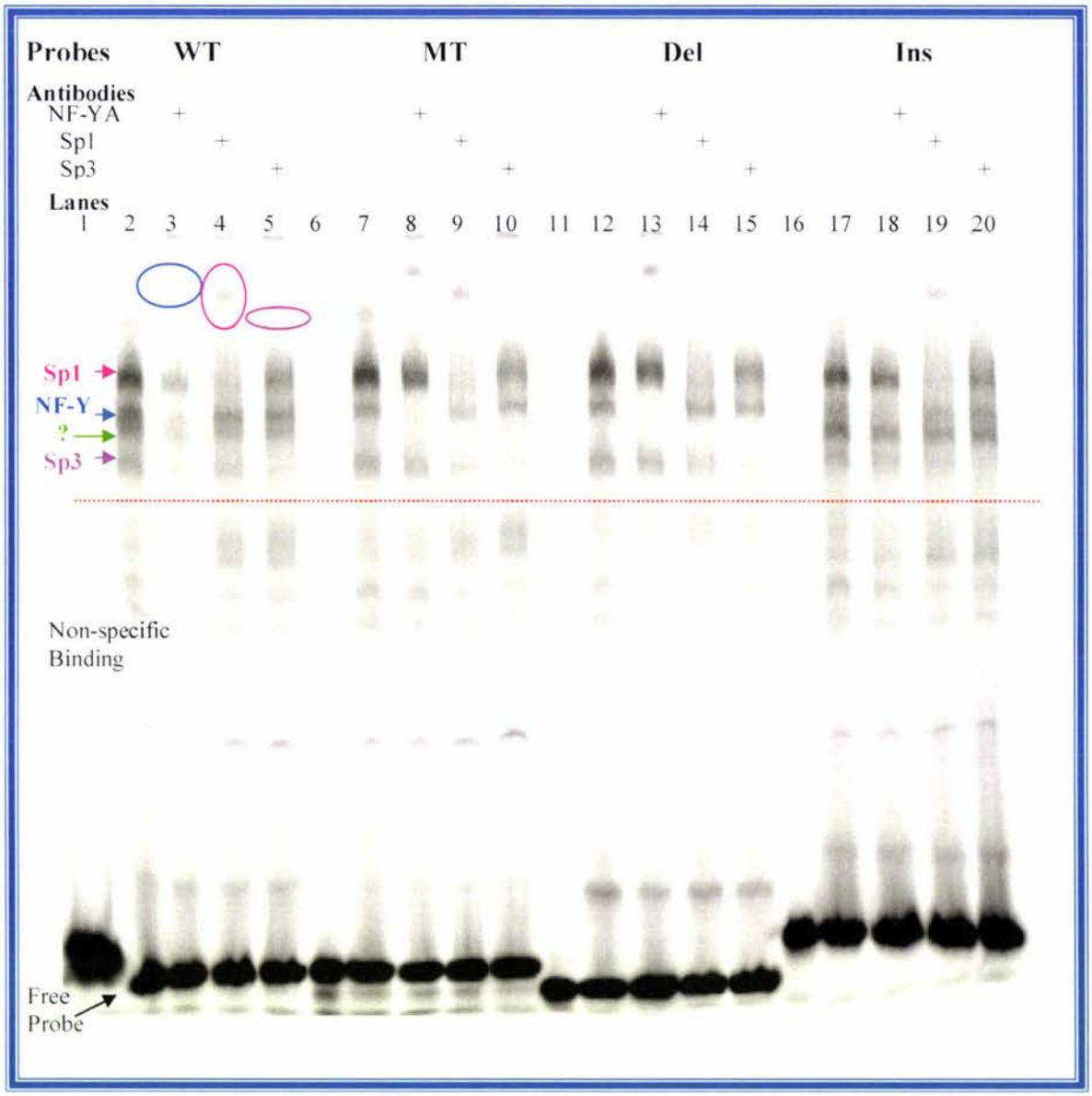


Figure 3.10: Antibody supershift with uncharacterised protein binding probes.

Control lanes 1, 6, 11, and 16 contain probe alone, and lanes 2, 7, 12, and 17 have probe and 10 µg HeLa extract. The remaining lanes have had 2-3 µL of the respective antibody added at a concentration of 200 µg/mL. Each reaction was loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight.

The specific DNA-protein complexes are indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. Supershift bands are within the coloured circles. This figure is a representative of triplicate experiments.

3.2.6 EMSAs of the Uncharacterised Protein Recruitment Set of Oligonucleotide Probes.

From the information gained from Figure 3.6-Figure 3.10, a second set of DNA oligonucleotide probes were designed (refer to Figure 3.2). These were termed the "uncharacterised protein recruitment probes", intended for identifying any specific recruitment of the uncharacterised protein, or any reliance on specific DNA elements.

Although the resolution of the bands is poor, it can be seen in Figure 3.11 there is a difference in binding between the WT and the new probes (compare lanes 3, 6, 9, 12, 15). There is no difference however in binding patterns between ICB1½GC1 and ICB1 probes (lanes 6 & 9), or any difference between GC1 ½ ICB1 and GC1 probes (lanes 12 & 15). Both the ICB1 probes seem to bind only NF-Y and may possibly have a small amount of Sp1 associated (lanes 6 & 9). The GC1 probes seem to bind only 3 proteins Sp1, Sp3 and NF-Y.

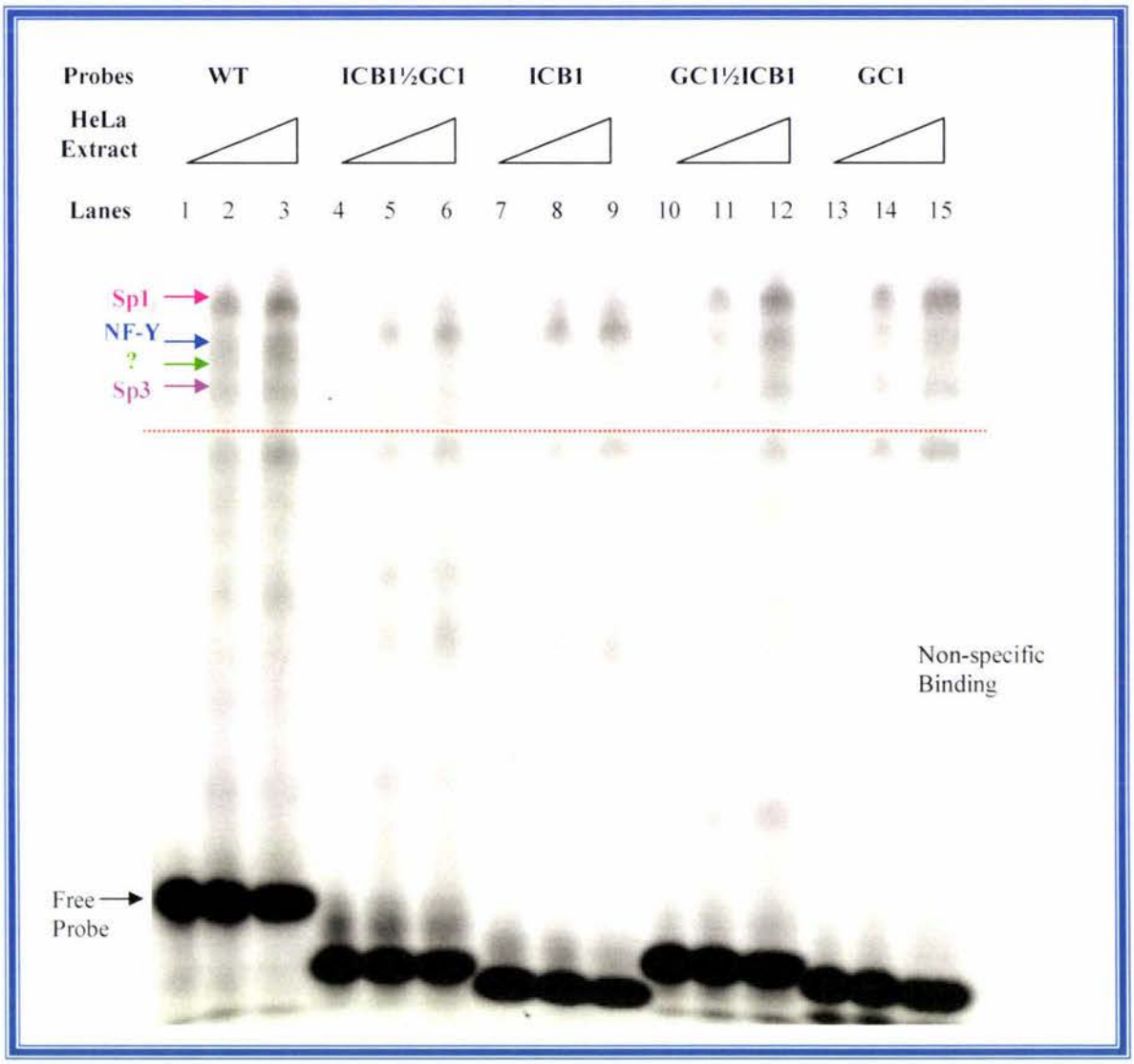


Figure 3.11: EMSA of the uncharacterised protein recruitment probes.

Each reaction contains 1-2 ng DNA probe, and increasing amounts of HeLa extract (0, 5, 10 µg). These were loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight.

Four specific DNA-protein complexes seen with WT probe are indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. Non-specific binding is indicated under the red line and free probe at the bottom.

3.2.7 Antibody Supershift Assay of the Uncharacterised Protein Recruitment Probes.

Antibody supershifts were performed as described (2.2.2(v)) to clarify results observed in Figure 3.11. The Figure 3.12 shows results from the ICB1 probes (refer to Figure 3.3). The WT probe shows Sp1, NF-Y, the uncharacterised protein, and Sp3 (lane 2). Addition of the respective antibodies confirmed the identity of the Sp1, NF-Y and Sp3 bands (lanes 3-5). Both the ICB1½GC1 and ICB1 probe showed a clear supershift with only the NF-Y antibody (compare lanes 7 & 8 blue circle, and 12 & 13 blue circle). Although no other supershifted bands were observed, NF-Y seems to be able to recruit Sp1 and Sp3 to the region as faint bands were observed at the correct mobilities for these proteins (see coloured arrows, lanes 7 & 12). Upon addition of the respective antibodies these bands disappeared, which is similar to the supershifts seen for Sp1 and Sp3 in Figure 3.13. The uncharacterised protein does not seem to be associated with NF-Y, neither is it able to bind when only half the GC1 element is present or when there is only small amounts of Sp1 or Sp3. There is a faint band however, found in between where the NF-Y and the uncharacterised protein bands are usually located (lanes 8 & 13 orange arrows). This band could be residual NF-Y protein not shifted by anti-NF-Y, although this was not observed in Figure 3.8. The mobilities do not completely correspond to either NF-Y or the uncharacterised protein, therefore it may possibly be another uncharacterised protein.

Figure 3.13 is the antibody supershift performed with the GC1½ICB1 and GC1 probes (refer Figure 3.3). Again the WT probe clearly shows the four protein bands (lane 2), and the relative supershifts with Sp1, NF-Y, and Sp3 antibodies (lanes 3-5). The GC1½ICB1 and GC1 probes both showed a supershift with Sp1 and Sp3 antibodies (lanes 7-10 arrows and 12-15 arrows respectively). No shift was seen with the addition of the NF-YA antibody, suggesting that NF-Y does not bind these sequences. The middle band seen with these probes is most likely to be the uncharacterised protein. These results are in accordance with the combined results from Figure 3.6-Figure 3.12, all of which suggest the uncharacterised protein is a GC1 associated protein, possibly recruited to the region by Sp1/Sp3.

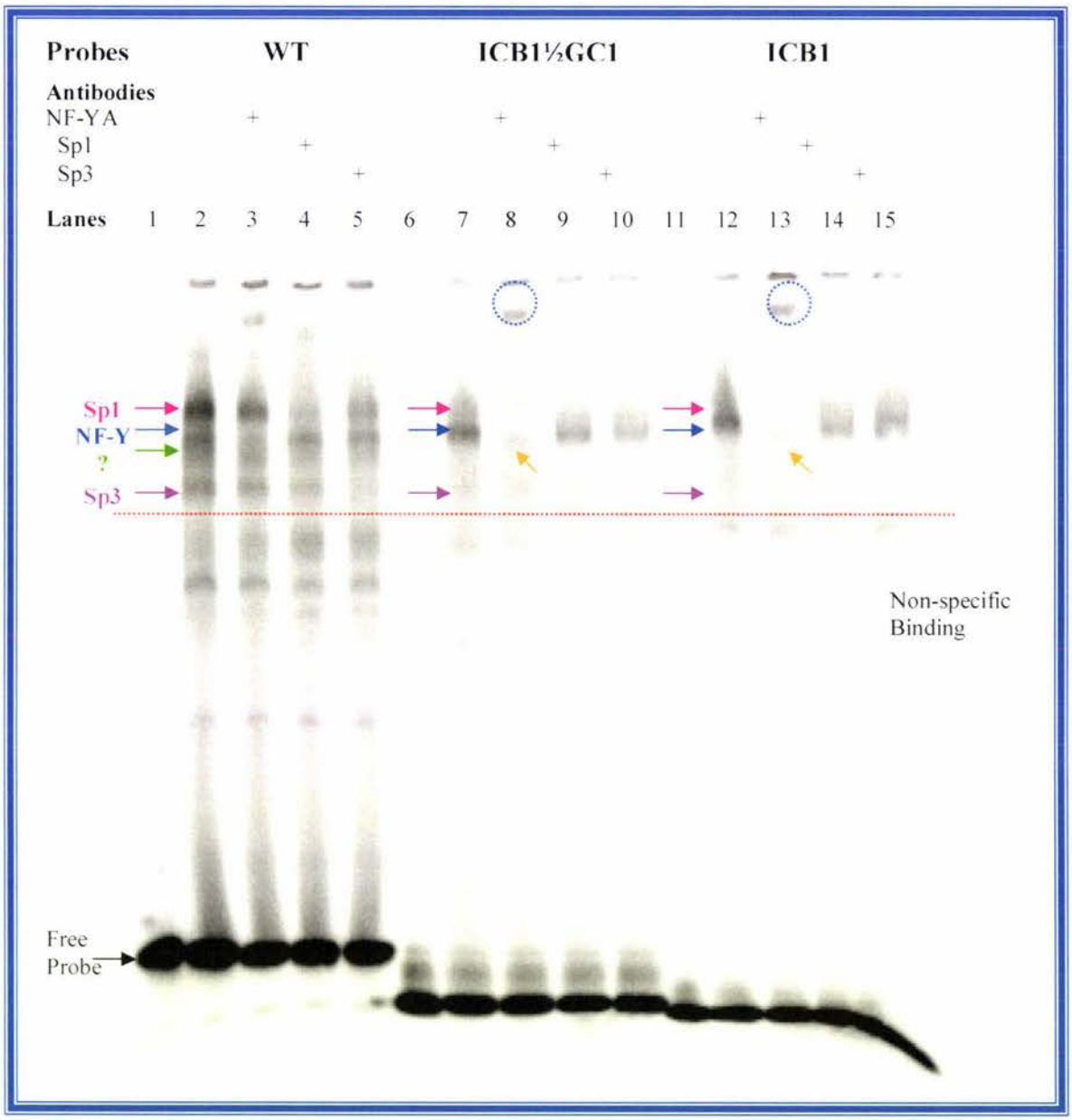


Figure 3.12: Antibody supershift of uncharacterised protein recruitment probes.

Control lanes 1, 6, & 11 contain probe alone, and lanes 2, 7, & 12 have probe and 10 µg HeLa extract. The remaining lanes have had 2-3 µL of the respective antibody added at a concentration of 200 µg/mL. Each reaction was loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight. The specific DNA-protein complexes are indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. Blue circles indicate NF-Y supershift. Orange arrows point to residual protein or another uncharacterised protein.

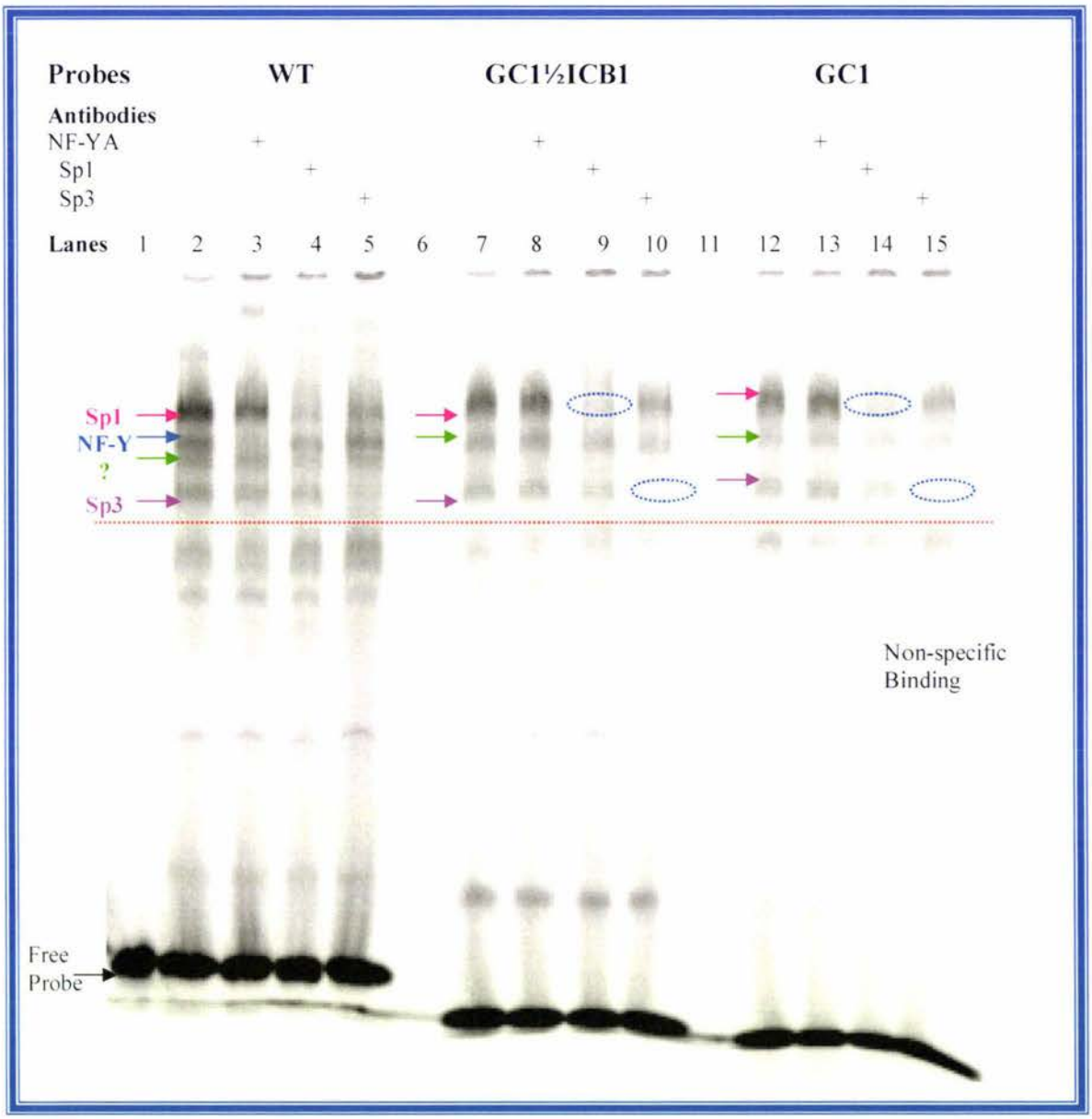


Figure 3.13: Antibody supershift of uncharacterised protein recruitment probes.

Control lanes 1, 6, & 11 contain probe alone, and lanes 2, 7, & 12 have probe and 10 µg HeLa extract. The remaining lanes have had 2-3 µL of the respective antibody added at a concentration of 200 µg/mL. Each reaction was loaded into a 4% PAGE and electrophoresis was carried out at 200 V for ~1 hour. The gel was dried onto DE-81 paper and exposed to a phosphoimager plate overnight.

The specific DNA-protein complexes are indicated by coloured arrows: pink: **Sp1**, blue: **NF-Y**, green: uncharacterised protein (labelled **?**), and purple: **Sp3**. Ovals indicate shifted bands.

3.3 Chapter Summary.

A range of different DNA oligonucleotide probes were designed to test protein binding behaviour and/or recruitment to the ICB1-GC1 region of the topo IIα minimal promoter. As previously shown by Magan *et al.* (2003), four distinct DNA-protein complexes are formed in EMSAs, with Sp1, NF-Y, and Sp3 identified in antibody supershift assays.

The fourth unidentified protein was originally thought to bind the 12 bp centre region between ICB1 and GC1. Mutating or deleting this region does indeed abolish binding of this protein, however a probe consisting of the centre region sequence alone did not appear to bind any protein (refer to Figure 3.8). Competitor assays with a double element probe possessing a wt GC1 successfully showed competition for the uncharacterised protein, but on the other hand, the single wt GC1 element did not (refer to Figure 3.6 and Figure 3.7). Interestingly, a probe consisting of the ICB1 box, the full centre region and half of the GC1 box did not bind uncharacterised protein. These results indicate the uncharacterised protein requires the centre region sequence in conjunction with the full wt GC1 box sequence to bind, or possibly the presence of Sp1 and/or Sp3. Sp1 and/or Sp3 may help recruit the uncharacterised protein to the DNA or *vice versa*. The relative amount of uncharacterised protein observed bound to the probes remains at the same level regardless of the amount of Sp1 or Sp3 bound, as seen in the single competitor assays and antibody supershifts. Therefore, a small amount of each, or just the presence of either of the two proteins, may be sufficient to induce binding of the uncharacterised protein.

Furthermore, the probe that contains an insert within the centre region appears to bind more strongly to the uncharacterised protein (refer to Figure 3.8 and Figure 3.9); this may be due to the fact the sequence used as the insert was a repeat of the original sequence seen within the centre region. In spite of this, it has been shown that the centre region sequence alone is not sufficient for the protein to bind. For that reason, the increase in binding may be because there is more space between the ICB1 and GC1 elements, possibly increasing access of the protein to the region.

It has been suggested that the NF-Y protein interacts with and possibly recruits Sp1 and/or Sp3 to this region of the topo IIα promoter (Magan, 2002). The probe encompassing the ICB1 box, centre region, and half the GC1 box showed faint Sp1 and Sp3 protein bands which supports this notion. In addition, a residual band close to the position of the NF-Y band is also observed (Figure 3.12). This is unlikely to be the uncharacterised protein as the mobility is slightly less; it could be residual NF-Y protein not shifted by the NF-Y antibody, however this was not observed in Figure 3.10 (lanes 8, 13, and 18), and the mobility does not completely correspond to the normal position of the NF-Y band. These results suggest that yet another uncharacterised protein may be associated with the DNA in this region.

On the other hand, the probe that encompasses the GC1 box, centre region, and half the ICB1 box did not bind NF-Y (refer to Figure 3.13). These two probes consistently show Sp1, Sp3 and the uncharacterised protein do not recruit NF-Y, but NF-Y appears to weakly recruit small amounts of Sp1 or Sp3. The results of the competitor assays also show there is an interaction between NF-Y and Sp1/3, as the single ICB1 element competitor competed slightly for Sp1 and Sp3. This effect was also observed with the double element wtICB1/mtGC1.

To test the functional significance of the differential binding of the uncharacterised protein observed with the binding probes, a series of mutated promoter reporter gene constructs were prepared as described in Chapter 4, for use in transient transfection assays (Chapter 5).

Chapter 4: Generation of Mutant Topoisomerase IIα Promoter Constructs.

4.1 Introduction.

A number of mutant promoter constructs were generated to deduce the functional role in which the uncharacterised protein plays in regards to topoisomerase IIα transcription. These constructs contained an altered 12 bp centre region which lies between the two regulatory elements (ICB1 and GC1), which was a potential binding site for the uncharacterised protein. PCR mutagenesis (section 2.2.6 and 2.2.7) was used to create the base pair substitutions, deletions and insertions within the ICB1-GC1 region of the topo IIα promoter. These mutations resulted in alteration of the distance between the two elements. Distances between regulatory elements has been shown previously to be important in the topoisomerase promoter and also the cyclin B2 promoter (Isaacs, 1996; Lok *et al.*, 2002; Salsi *et al.*, 2003). If alterations in the distance between elements results in changes in transcription levels, this could suggest that the proteins may be interacting with each other in a DNA-dependent manner, and/or causing alteration in DNA topology which affects functional interactions of proteins bound to the promoter. Protein binding assays with oligonucleotides encompassing these same mutations have been shown to either inhibit or increase binding of the uncharacterised protein (Chapter 3). Therefore, potential effects on transcription observed from each construct may help determine the role of the uncharacterised protein in transcriptional regulation of topo IIα.

The aim of the work described in this chapter was to produce DNA constructs of the topo IIα promoter containing mutations known to affect the binding of the uncharacterised protein. By linking these mutated promoter regions to the luciferase reporter gene, functional effects of these mutations could be tested. All the constructs produced were inserted into the luciferase reporter vector pGL3 Basic, for use in functional transient transfection experiments which are described in Chapter 5.

A second set of promoter-reporter constructs were also prepared where each promoter proximal mutation (in the ICB1-GC1 region) was combined with a mutation in the more distal GC2 element. GC2 is known to bind Sp1/Sp3, and has been shown to have a higher affinity for Sp1 than the single GC1 element (Szremska, 2000). This has not been tested in the presence of NF-Y bound at ICB1, which may affect the binding affinity of proteins to GC1. Synergy between the two GC boxes was observed by Magan, *et. al* (2003), when both GC1 and GC2 are mutated transcription levels increased even further than single GC box mutants. The proximal mutations made within the ICB1-GC1 region may affect any interplay occuring between proteins bound at the two GC boxes. Therefore, any additive effects on transcription levels seen when these proximal mutations are combined with a mutant GC2 box will indicate synergism between the GC boxes. These constructs were also used in functional transient transfection experiments which are described in Chapter 5.

4.2 Generation of the ICB1-GC1 Centre Region Mutant Constructs.

4.2.1 PCR Mutagenesis.

The topoisomerase IIα minimal promoter region (-617 bp upstream of the transcription start site) has been defined (Hochhauser *et al.*, 1992; Isaacs, 1996) and previously cloned into a pGL3 Basic vector (M^{c}Lenachan, 1998) (Appendix 3 for map). Three specific internal primers were manually designed (refer to Appendix 4 for full primer sequences) for use in PCR mutagenesis (Ho *et al.*, 1989) to introduce mutated sequences within the topo IIα promoter region. PCR mutagenesis involves two PCR reactions (Figure 4.1); the first PCR reaction was performed with specifically designed internal primers containing the mutations intended for insertion, in two separate reactions (Figure 4.1A). An external primer was added in each reaction in conjunction with the topo IIα -617wt pGL3 Basic vector as the DNA template. The PCR products from both these tubes contain only part of the whole topo IIα promoter sequence with the mutations incorporated at the position of the internal primers. The second PCR reaction (Figure 4.1B) was carried out using the PCR products from both of the first reactions as the DNA templates, together with the

two external primers to amplify the entire -617 topo IIα promoter incorporating the mutation introduced in the first set of reactions.

PCR mutagenesis was carried out as described in sections 2.2.6 and 2.2.7. The PCR products from the first PCR reactions were analysed on a 1% agarose gel (Figure 4.2A) and purified using a PCR purification kit (section 2.2.8). As predicted from sequence analyses, a band of approximately 250 bp was seen when the forward internal primer was used in conjunction with the reverse GL2 primer (Figure 4.2A lanes 3, 5 & 7). A band approximately 650 bp was seen with the internal reverse primer and the forward RV3 primer (Figure 4.2A lanes 4, 6, & 8) as expected. The GL2 and RV3 primers bind to sequences flanking the MCS (multiple cloning site) within the pGL3 Basic vector. The PCR products were gel quantified (section 2.2.9i) by comparisons with quantification standards prior to use in the second round of PCR mutagenesis (Figure 4.2B). The quantities were estimated as; mutant A and deletion A both ~20 ng/μL, insert A ~10 ng/μL (lanes 5, 7 & 9), and mutant, deletion, insert B all ~50 ng/μL (lanes 6, 8 & 10).

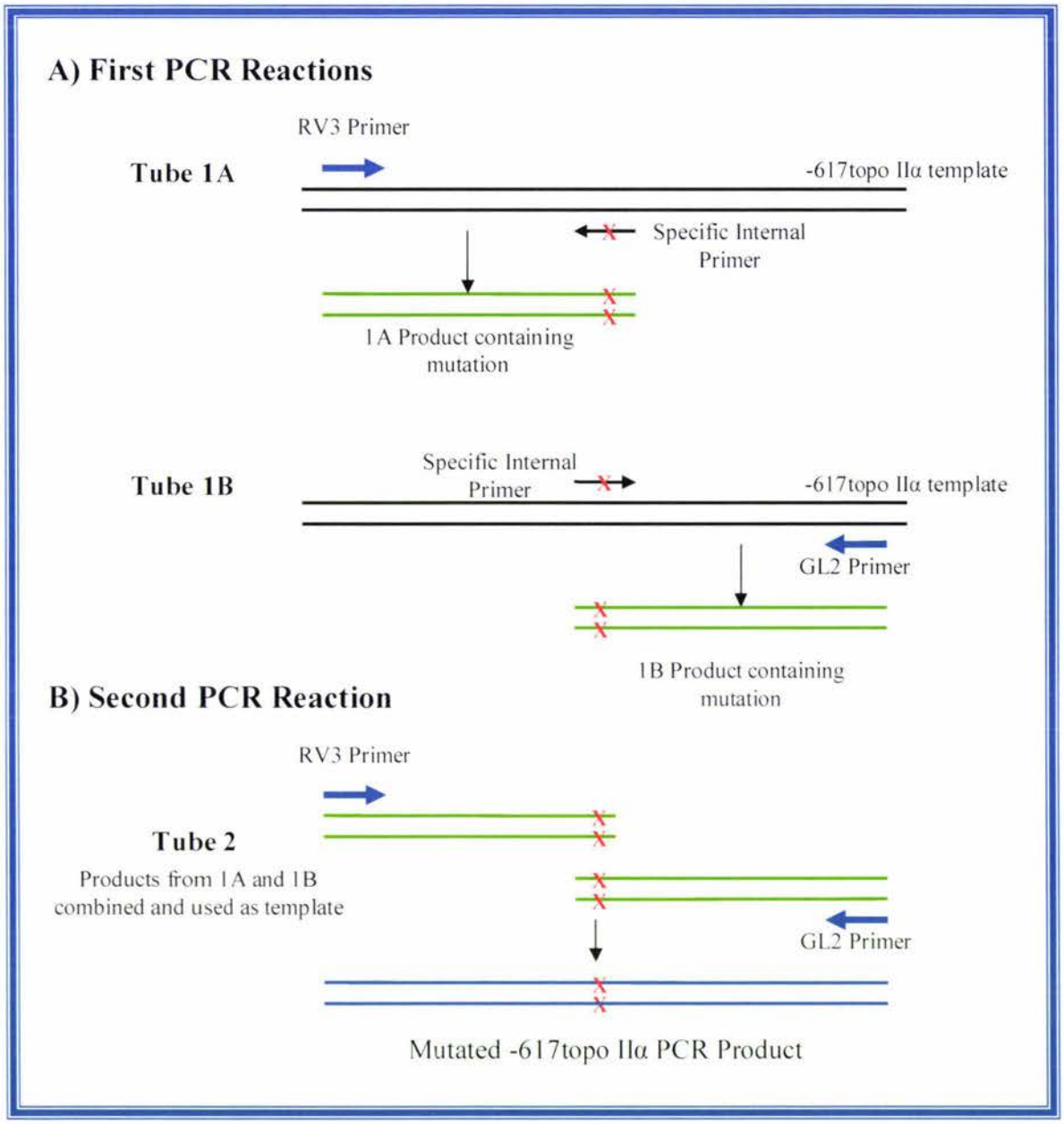


Figure 4.1: Schematic representation of PCR mutagenesis.

PCR mutatgenesis involves two PCR reactions. The first PCR reactions required specifically designed internal primers which contain the mutations intended for insertion. Each complementary internal primer was used in a separate tube (tube 1A and 1B) in conjunction with an external primer and the original -617 topoisomerase IIα promoter sequence as the template DNA (A). B) The PCR products from each of these primary reactions were then combined (tube 2) and used as the DNA templates in a second round of PCR with only the two external primers. The final product is then the whole -617 topo IIα promoter sequence containing the mutations.

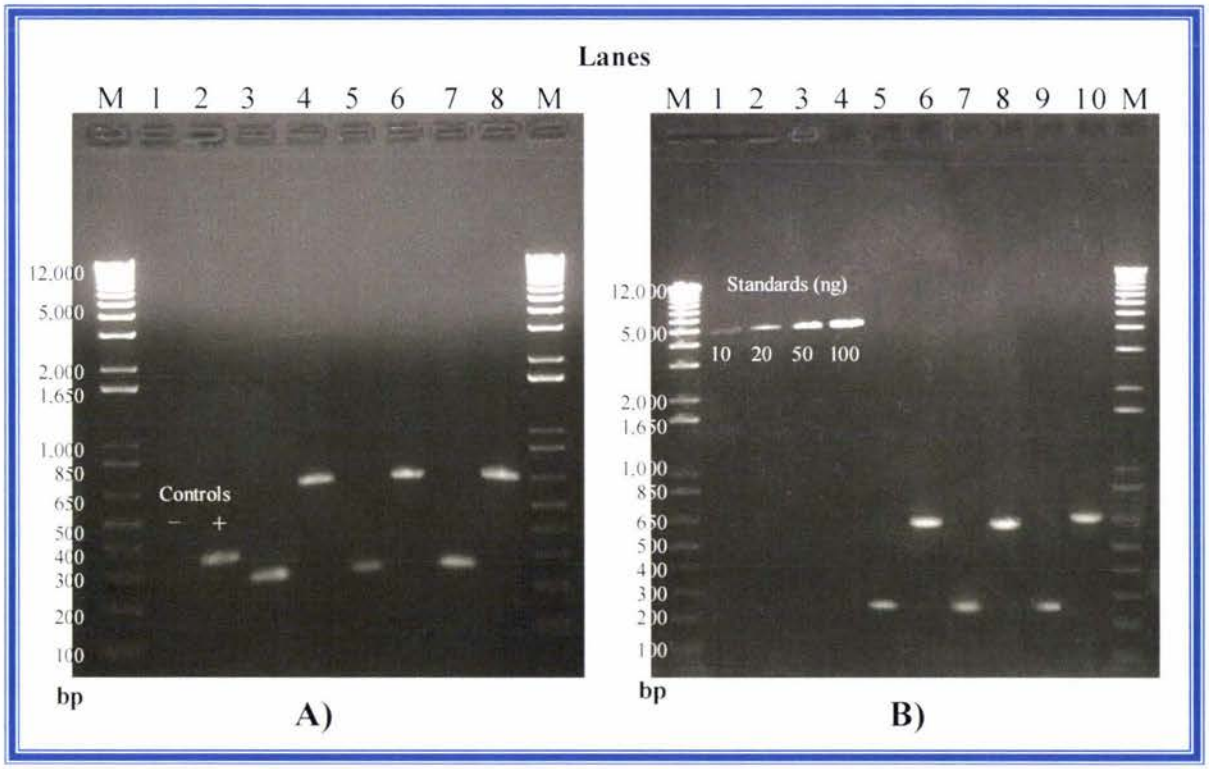


Figure 4.2: First PCR mutagenesis samples and subsequent quantification gel.

DNA samples were analysed on a 1% agarose gel in 1x TAE buffer. Electrophoresis was carried out at 90 V for approximately 1 hour. Ethidium bromide (0.5 µg mL^{-1}) was incorporated into the gel to allow visualization of the DNA bands under UV light. A total volume of 5 µL (1/10 of final volume) of each PCR product, and DNA quantification standard was loaded into each well as indicated. Lane M = 5 µL of 1 Kb plus ladder.

A) PCR products. Lanes 1 & 2: Negative & positive PCR controls respectively.
3: Mutation forward primer & GL2, 250 bp fragment
4: Mutation reverse primer & RV3, 650 bp fragment
5: Deletion forward primer & GL2, 250 bp fragment
6: Deletion reverse primer & RV3, 650 bp fragment
7: Insertion forward primer & GL2, 250 bp fragment
8: Insertion reverse primer & RV3, 650 bp fragment

B) Purification and Quantification Gel.

PCR products were purified using a QIAquick PCR purification kit (Qiagen) and quantified by comparisons with the quantification standards lanes 1-4. A total volume of 5 µL (1/10 of final volume) of the purified PCR products were loaded into each well as indicated. Lanes 5 & 6: Mutation A & B, lanes 7 & 8: Deletion A & B, and lanes 9 & 10: Insertion A & B respectively.

Products from the second PCR reactions were also analysed on a 1% agarose gel (Figure 4.3A). The mutant, deletion and insertion products all appeared as strong bands approximately 860 bp in size (Figure 4.3A lanes 3, 4 & 5 respectively), however, there were two faint bands present in each sample at 1650 bp and 2000 bp. These bands could have been concatamers of each DNA sample, which concurs with their estimated sizes. To eliminate these products prior to cloning, gel extraction and ethanol precipitation was carried out as described in sections 2.2.4 and 2.2.5 respectively. The purified PCR products were then gel quantified (section 2.2.9i) by comparison with quantification standards (Figure 4.3B). The quantities were estimated as mutant ~40 ng/µL, deletion ~3 ng/µL, and insertion ~20 ng/µL (Figure 4.3B lanes 1, 2 & 3 respectively).

4.2.2 Diagnostic Restriction Endonuclease Digests.

The mutated topo IIα promoter sequences produced by PCR mutagenesis in section 4.2.1 were analysed using diagnostic restriction endonuclease digests to see if the promoter did in fact contain the intended mutations. The restriction endonuclease digests were carried out as described in section 2.2.12 with *Hae*III endonuclease, as the mutation (MT) sequence created a new *Hae*III restriction site, therefore enabling it's identification. The deletion (Del) and insertion (Ins) sequences were identified by size comparisons on a high resolution NuSieve® agarose gel (section 2.2.13).

Figure 4.4A is a schematic representation of the estimated fragment sizes resulting from the *Hae*III diagnostic digestion of each different PCR mutated construct. A wild-type promoter was used for size comparisons. The novel *Hae*III restriction site in the MT construct is shown in red. The actual DNA fragments from the digestion were run on a NuSieve® gel for better resolution of bands (refer to Figure 4.4B). A band representing uncut wild-type promoter is shown in the uncut lane (~850 bp), with cut wild-type bands in the WT lane approximately 360, 340, 100 and 70 bp (in descending order). The 340 bp band is absent from the MT lane, but is replaced by a band of approximately 250 bp, indicating the sequence must contain the mutation generating the novel *Hae*III restriction site.

The smaller fragments in the MT lane are harder to define than in the WT lane, which may be due to the existence of the new 90 bp fragment which results in poor resolution by gel electophoresis due to the size similarities. The Del lane contains the bands from the digested deletion promoter. By comparing the two largest bands (360 and 334 bp) with the largest bands of the WT lane (360 and 340 bp), although slight, the difference in size between these two bands can be seen. This becomes more evident when the space between the two bands is compared with one another (larger in the Del lane as the fragment is smaller, thus travelling further). This is evidence the deletion has been introduced in this sequence. The 346 bp band in the Ins lane is shown to be larger than that of the 340 bp WT band, again evident through comparisons with wild-type and more obviously with the Del band. The space between these top two bands is clearly smaller compared with the space between the top two bands of the WT. Therefore the Ins band is larger indicating the insertion was most likely incorporated into the sequence. Although faint, the smaller fragments in the Del and Ins lane are both the same as those in the WT lane (100 and 70 bp) as expected.

These results implied the PCR mutagenesis had been successful, so the products were then prepared for ligation into the pGL3 Basic vector.

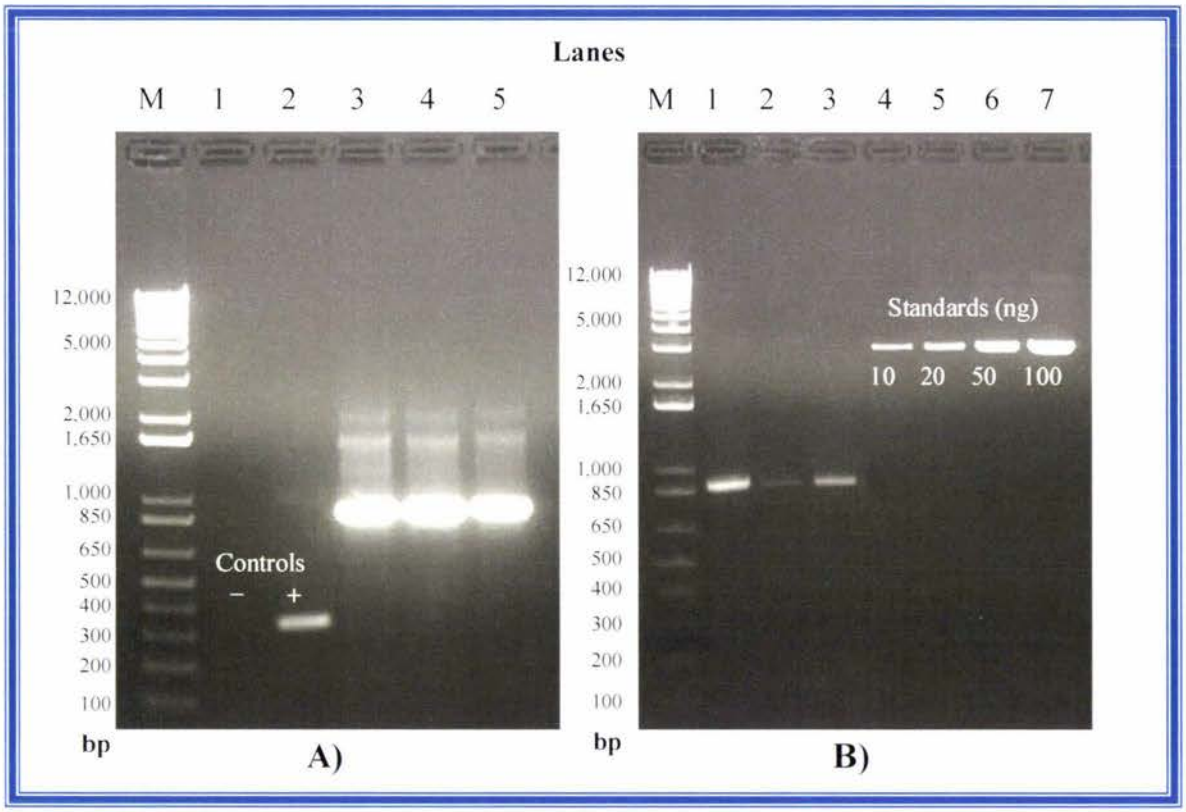


Figure 4.3: Second PCR mutagenesis samples and subsequent quantification gel.

DNA samples were analysed on a 1% agarose gel in 1x TAE buffer. Electrophoresis was carried out at 90 V for approximately 1 hour. Ethidium bromide (0.5 µg/mL) was incorporated into the gel to allow visualization of the DNA bands under UV light. A total amount of 5 µL (1/10 of final volume) of each PCR product, and 5 µL DNA quantification standard was loaded into each well. Lane M = 5 µL of 1 Kb plus ladder.

A) Second PCR Mutagenesis Reaction with First PCR Products used as Templates.

Lanes 1 & 2: Negative & positive PCR controls respectively.

3: -617 Mutant Topo IIα - 850 bp

4: -617 Deletion Topo IIα - 850 bp

5: -617 Insertion Topo IIα - 850 bp

B) Purification and Quantification Gel.

A total volume of 5 µL (1/10 final volume) of the gel purified PCR products was loaded into each well, lane 1: Mutant, lane 2: Deletion, and lane 3: Insertion. These were quantified by comparisons with the quantification standards lanes 4-7 as indicated.

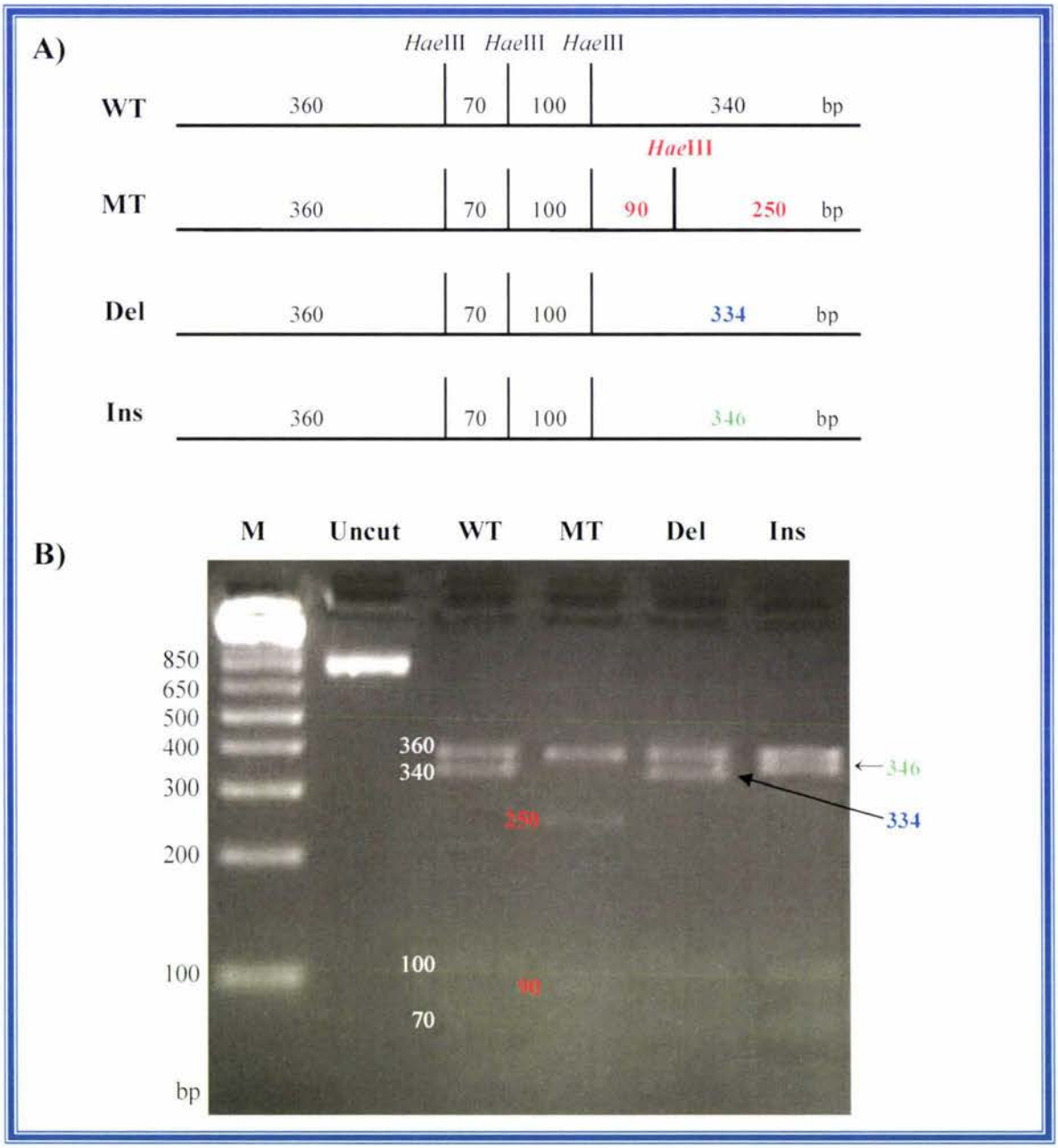


Figure 4.4: *Hae*III diagnostic digest.

A) Schematic diagram of approximate sizes of DNA fragments (base pairs) after *Hae*III endonuclease digestion. The newly created *Hae*III site in the MT sequence is indicated in red. B) DNA fragments represented in (A) from *Hae*III diagnostic digest shown on a NuSieve® gel. DNA samples were analysed on a 3.5% NuSieve® agarose gel in 1x TAE buffer. Electrophoresis was carried out at 30 V for ~6 hours. Ethidium bromide (0.5 µg/mL) was incorporated to allow visualization of the DNA bands. A total of 5 µL of each restriction digest (1/10 final volume), and DNA quantification standard was loaded into each well. Lane M = 5 µL of 1 Kb plus ladder.

4.2.3 Restriction Endonuclease Digests.

The mutated topoisomerase IIα promoter sequences (PCR products) were inserted into the pGL3 Basic luciferase reporter vector, to test functional affects of mutations in future experiments. The pGL3 Basic vector was digested using two different restriction endonucleases, which should prevent self-religation and encourage inserts to ligate in the correct orientation (directional cloning). The enzymes used were *Kpn*I and *Nco*I, both of which recognise and cut within a 6 bp sequence, with the *Kpn*I site located at the start of the multiple cloning site, and the *Nco*I site lying at the start of the luciferase reporter gene (refer to Appendix 3 for vector maps). The gel photo of the vector digestion is shown in Figure 4.5A. Lane 1 is the control uncut pGL3 Basic, with the two bands representing supercoiled and relaxed forms of the plasmid. The vector was first cut with *Kpn*I followed with the addition of *Nco*I and NaCl, as *Nco*I required a higher salt concentration than contained in the buffer for *Kpn*I. Each enzyme was first tested if it could cut under these conditions individually, as shown in Figure 4.5A. Lane 2 *Nco*I alone and lane 3 *Kpn*I alone show that the enzymes were both capable of digesting the vector, producing a 4818 bp linear form. Lanes 4 and 5 both show the vector which has been cut with both *Kpn*I and *Nco*I whose restriction sites were 81 bp away from each other. The band has a slightly higher mobility (4737 bp) than the control band in lane 2 and 3 indicating both have cut correctly; however, this size difference is hard to distinguish as the overall plasmid size is so large. All digests have gone to completion compared to lane 1 as there are no relaxed or supercoiled bands remaining in any of the digestion lanes.

The vector was gel purified as described in section 2.2.4 and 2.2.5. The 5` phosphate groups from the vector were removed using thermosensitive alkaline phosphatase (TsAP) as described in section 2.2.14, which prevents the vector self-ligating. This was then gel quantified as described in section 2.2.9i (Figure 4.5B) at approximately 10 ng/μL.
The PCR products of the topo IIα promoter intended for insertion into the pGL3 Basic vector ends cut by *Kpn*I and *Nco*I were column purified (section 2.2.8) and gel quantified (section 2.2.9i) as shown in Figure 4.5B.

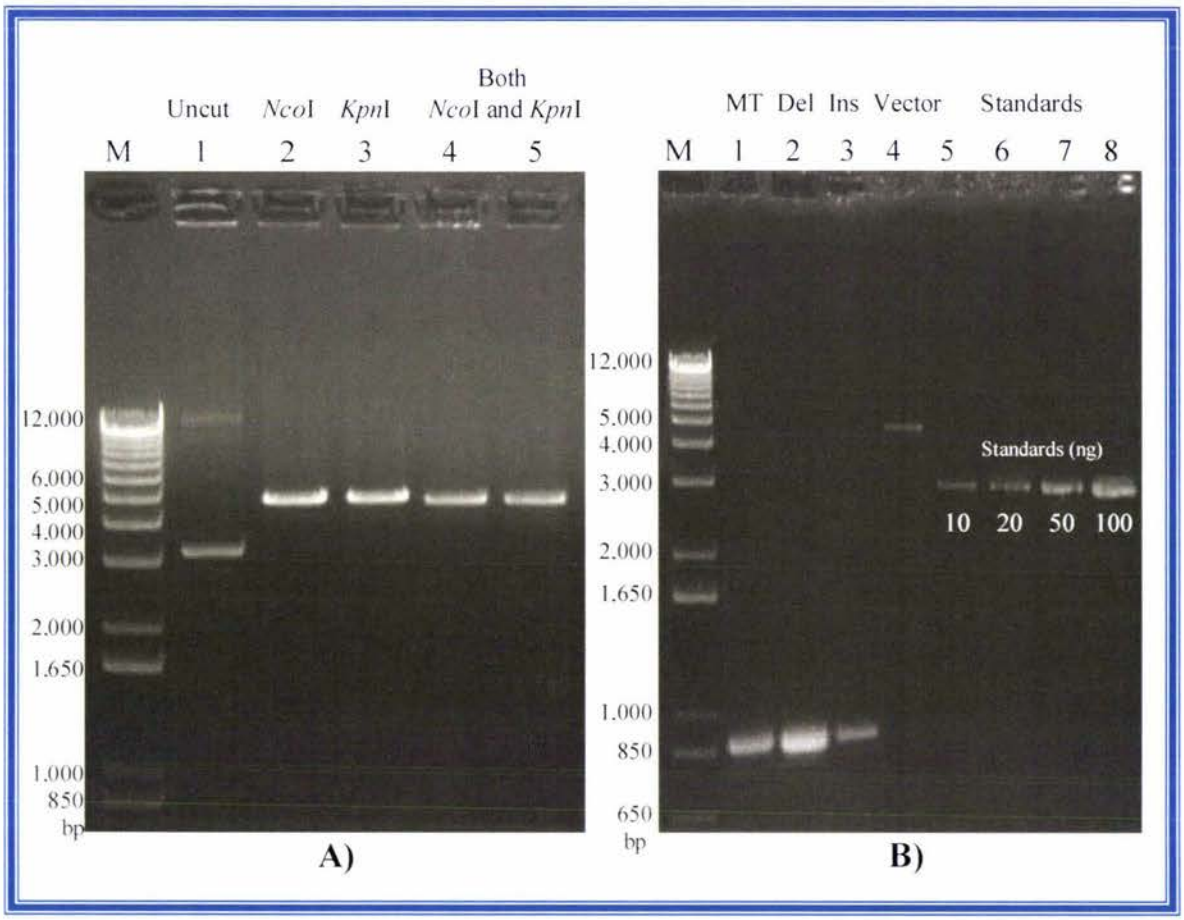


Figure 4.5: *Kpn*I and *Nco*I restriction endonuclease digests and quantification gel.

DNA samples were analysed on a 1% agarose gel in 1x TAE buffer. Electrophoresis was carried out at 90 V for approximately 1 hour. Ethidium bromide (0.5 µg ml^{-1}) was incorporated into the gel to allow visualization of the DNA bands under UV light. A total amount of 5 µL of each digest (1-10 final volume) and DNA quantification standard was loaded into each well. Lane M = 5 µL of 1 Kb plus ladder.

A) Vector Restriction Endonuclease Digestion Gel.

Lane 1: Uncut pGL3 Basic control. Lane 2 & 3: *Nco*I and *Kpn*I single digestions respectively. Lane 4 & 5: Both *Nco*I and *Kpn*I double digestions.

B) Quantification Gel.

Lane 1: Mutation, lane 2: Deletion, lane 3: Insertion, and lane 4: pGL3 Basic Vector. These were quantified by comparisons with the quantification standards lanes 5-8 as indicated. Note: uncut inserts were not included as the size difference would not have resolved well in this particular gel.

4.2.4 DNA Ligations.

The mutant topoisomerase IIα promoter fragments were inserted into the pGL3 Basic vector by a ligation using T4 DNA ligase (section 2.2.15) at the *Nco*I and *Kpn*I restriction sites, which were digested in both the vector and insert DNA sequences prior (refer to Figure 4.6 for schematic diagram). The concentration of DNA ends in the ligation reaction is the important component in ligations. The molar concentration of the insert must be higher than the vector to favour inter-molecular ligations rather than intra-molecular. Ligations were prepared with a 3:1 insert:vector molar ratio, 10 µL of each reaction (refer to Table 4.1 for reaction contents) was used to transform competent *E.coli* XL-1 Blue cells (section 4.2.5).

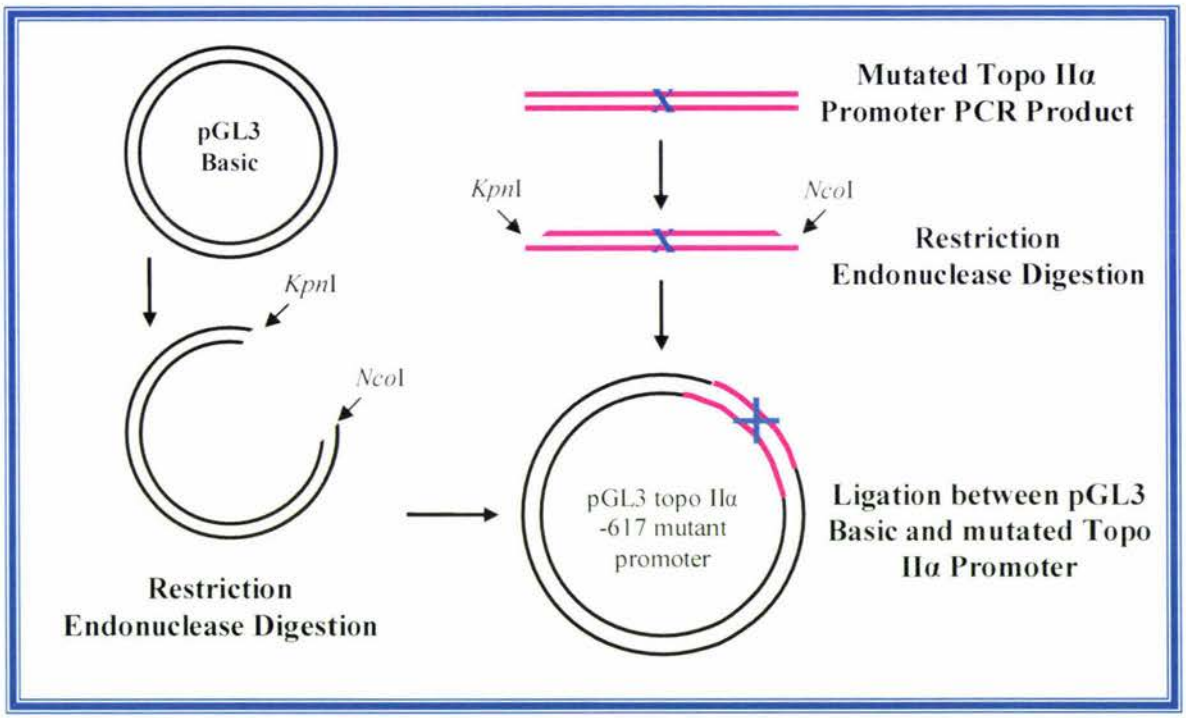


Figure 4.6: Schematic representation of ligation strategy.

pGL3 Basic vector was prepared by digestion with *Kpn*I and *Nco*I restriction endonuclease enzymes, gel purification and TsAP treatment to remove 5' phosphate groups from the vector ends. The mutated topo IIα promoter PCR products were digested by the same restriction enzymes and column purified. Both were gel quanified and ligated together using T4 DNA ligase. *E.coli* XL-1 Blue strain was then transformed with the new DNA constructs, and glycerol stocks were made.

Sample	Vector (ng)	Insert (ng)	T_4 Buffer	T_4 Ligase	H_2O	Volume Cells Plated	Plate	No. Colonies
+ Control	-	-	-	-	-	50 µL	LB	Lawn
+ Control	-	-	4 µL	1 µL	15 µL	50 µL	LB	Lawn
- Control	-	-	-	-	-	100 µL	LB Amp	0
- Control	-	-	4 µL	1 µL	15 µL	100 µL	LB Amp	0
Uncut Vector (+control)	20 ng (1 µL)	-	-	-	19 µL	100 µL	LB Amp	Lawn
Cut Vector -no ligase (-control)	20 ng (1 µL)	-	4 µL	-	15 µL	100 µL 100 µL	LB Amp LB Amp	0 0
Cut Vector + Ligase (-control)	20 ng (1 µL)	-	4 µL	1 µL	14 µL	100 µL 100 µL	LB Amp LB Amp	0 0
MT Insert + Vector (3:1)	20 ng (1 µL)	12 ng (6 µL)	4 µL	1 µL	8 µL	50 µL 50 µL	LB Amp LB Amp	54 47
Del Insert + Vector (3:1)	20 ng (1 µL)	12 ng (6 µL)	4 µL	1 µL	8 µL	50 µL 50 µL	LB Amp LB Amp	42 50
Ins Insert + Vector (3:1)	20 ng (1 µL)	12 ng (6 µL)	4 µL	1 µL	8 µL	50 µL 50 µL	LB Amp LB Amp	101 119

Table 4.1: Ligation reactions and transformation results.

The total volume of each ligation reaction was 20 µL : half of this (10 µL) was used to transform 100 µL of competent *E.coli* XL-1 Blue cells. Cells were plated onto LB agar or LB Ampicillin agar plates and incubated overnight at 37°C. The vector contains an ampicillin resistance gene for selection. Colonies were creamy white in colour. Blue white selection is not available with the pGL3 Basic vector.

4.2.5 Transformation of E.coli XL-1 Blue Cells with Ligation Reactions.

The new topo IIα promoter vectors were introduced into chemically competent *E.coli* XL-1 Blue cells by transformation (refer to section 2.3.1) for amplification. Competent XL-1 cells are thought to be able to take up foreign DNA through cell-surface proteins or enzymes that promote the DNA binding. The pGL3 Basic vector used in the ligations contained an ampicillin resistance gene which was used for selection of bacterial colonies containing an intact plasmid. Untransformed XL-1 cells grow on LB plates, however without the ampicillin resistance gene conferred by the plasmid they were unable to grow on LB Amp plates, as seen in Table 4.1 1st row. These act as positive and negative

controls, showing the cells were both viable, and that there were no spontaneous mutants able to grow on ampicillin plates.

Transformation with undigested vector produced a lawn of growth, compared with no growth on LB Amp plates with linearised vector (Table 4.1 2nd & 3rd row), therefore confirming the intact plasmid was necessary for the XL-1 cells to grow on LB Amp plates, and that there was no uncut vector remaining in the prepared vector stock. Another control carried out was performed with the digested vector and T_4 DNA ligase (Table 4.1 4th row), this tested whether any intramolecular re-ligation was occurring. There was no growth on these plates indicating the vector was unable to re-ligate, consistent with the cloning strategy chosen.

The inserts were each added at a 3:1 insert to vector molecular ratio, which produced reasonable results (Table 4.1 5th-7th row), with 42-119 possible transformant colonies. A number of colonies from each plate were chosen to be screened for each of the pGL3 Basic topo IIα mutated promoter constructs (section 4.2.6).

4.2.6 Identifying Plasmids Sequences.

The *E.coli* cells which successfully grew on the ampicillin plates were screened for the presence of pGL3 Basic containing correctly inserted DNA sequences. The identities of plasmids were checked using a crude extract for a preliminary diagnostic digest screen which was analysed by agarose gel electrophoresis. Upon the presence of the correct plasmid, a refined extraction was carried out to produce purer DNA for automated sequencing to be performed. A larger scale plasmid preparation was then performed to acquire large amounts of the plasmid DNA for functional assays.

(i) Plasmid Identity Screen with Diagnostic Restriction Endonuclease Digest.

Individual colonies representing each of the three different mutant topo IIα promoter clones (MT, Del, and Ins) were selected and grown in LB broth for rapid boil plasmid extractions (refer to section 2.3.2i). The extracts were treated with RNase and then digested with *Sac*I restriction endonuclease and analysed using agarose gel electrophoresis, to check if the plasmid contained an insert (refer to Figure 4.7).

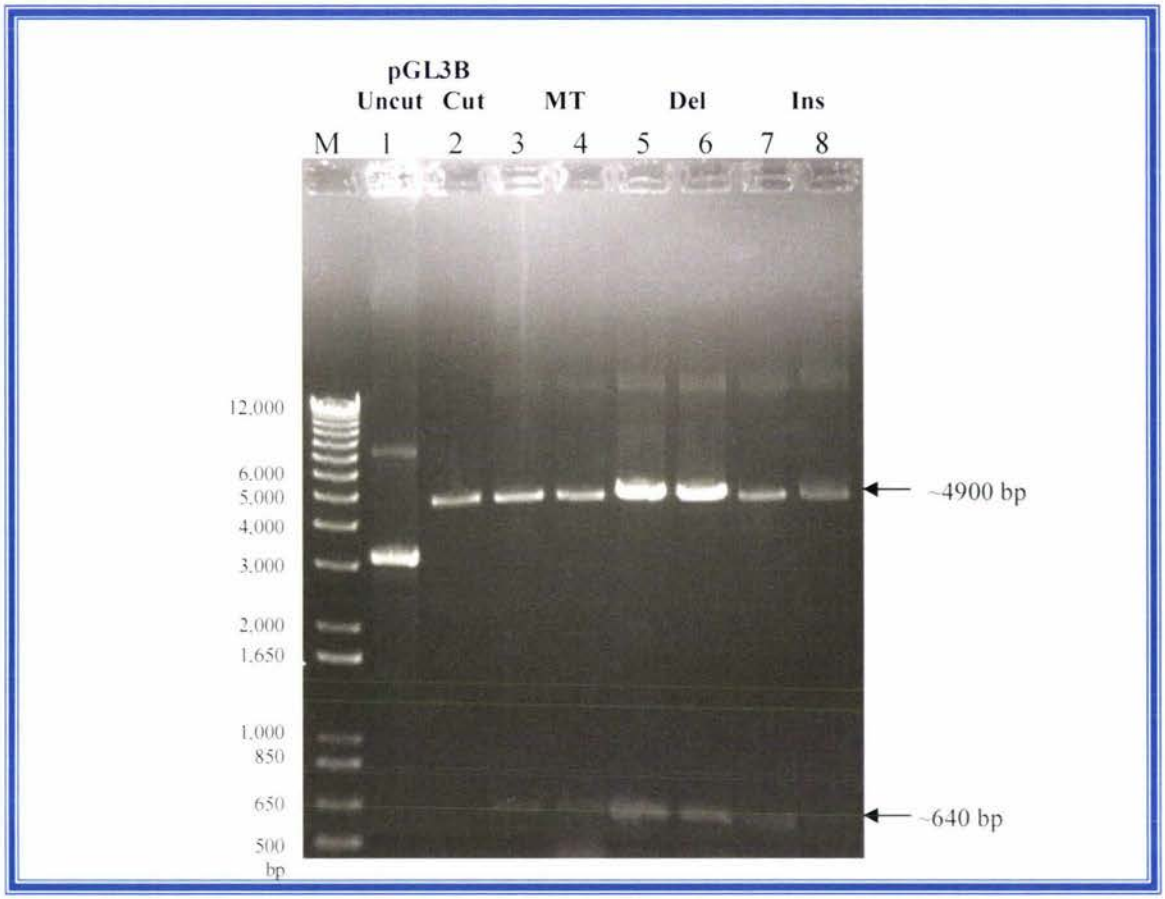


Figure 4.7: Plasmid identity screen with *Sac*I.

DNA samples were analysed on a 1% agarose gel in 1x TAE buffer. Electrophoresis was carried out at 90 V for approximately 1 hour. Ethidium bromide (0.5 µg ml^{-1}) was incorporated into the gel to allow visualization of the DNA bands under UV light. A total amount of 5 µl of each digest (1:10 final volume), and DNA quantification standard was loaded into each well. Lane M = 5 µl of 1 Kb plus ladder.

Lane 1 & 2: Controls – uncut and cut pGL3 Basic empty vector respectively. The digested plasmid produces only a single band as there is only one *Sac*I restriction site in the plasmid. Lane 3 & 4: Topo IIα MT plasmid, lane 5 & 6: Topo IIα Del plasmid, and lane 7 & 8: Topo IIα Ins plasmid. Upon digestion with *Sac*I a 640 bp band is apparent, which indicates each plasmid contains an insert as there must be two *Sac*I restriction sites. The original *Sac*I site was removed during vector preparation; therefore the two sites are both from the insert, which was predicted by sequence analysis. The larger bands seen are ~4900 bp, slightly larger than the band seen in lane 2 due to the *Sac*I restriction sites being located within the insert sequences (refer to Appendix 5 for schematic diagram).

The pGL3 Basic vector contains one *Sac*I restriction site within the multiple cloning site (refer to Appendix 3), therefore digestion with this enzyme produces only a single band at approximately 4800 bp (refer to Figure 4.7 lane 2). The preparation of this vector for the ligation involved digestion at the *Kpn*I and *Nco*I sites, thereby removing the *Sac*I site, however the insertion was PCR-amplified from primers that lie outside of the *Sac*I site in the pGL3 Basic vector, therefore, the *Sac*I site of the vector was also amplified and because *Kpn*I lies outside of the *Sac*I site it was inserted back into the construct along with the new topo IIα sequence which includes it's own *SacI* site (refer to Appendix 5 for schematic diagram). Therefore, digestion of the three plasmids (MT, Del, and Ins) resulted in an approximately 640 bp band and an approximately 4900 bp band (Figure 4.7 lanes 3-8), the large band is slightly longer than the single band seen when the control vector was digested due to the *Sac*I restriction site from the topo IIα sequence being located approximately 100 bp from the end of the insert sequence. The presence of the insert alone was not sufficient to guarantee the sequences were correct, therefore automated sequencing was carried out to confirm the correct mutations had been introduced and no additional mutations had been introduced by PCR.

(ii) Plasmid Sequence Identity Analysis.

Colonies with plasmids found to contain an insert were grown and the plasmids were purified using a Quantum® Prep Plasmid Miniprep Kit (section 2.3.2ii) for purer DNA for use in direct automated sequencing reactions (section 2.2.10). Each sequence was analysed to confirm the presence of the desired mutations (Figure 4.8) and a sequence alignment was carried out against the original wild type sequence (refer to section 2.2.11 and Appendix 6 for the full sequence alignment). Sequence alignment was performed to ensure no random mutations had occurred during the cloning process. Once the sequences were confirmed correct, the colonies were grown at a much larger scale for plasmid extraction using a QIAGEN Plasmid Midiprep Kit (refer to section 2.3.2iii). This was necessary to obtain higher amounts and concentrations for future use in functional assays (Chapter 5). The concentration of each plasmid was measured by spectrophotometry (section 2.2.9ii). Concentrations from 50 mL cultures were as

follows: mutant (MT) topo IIα promoter construct, 1.59 μg/μL; deletion (Del) construct 2.84 μg/μL; and the insert (Ins) construct was 4.18 μg/μL.

Glycerol stocks were produced for long term storage and future production of the plasmids. These were prepared as described in section 2.3.3 and stored at -70^{O}C.

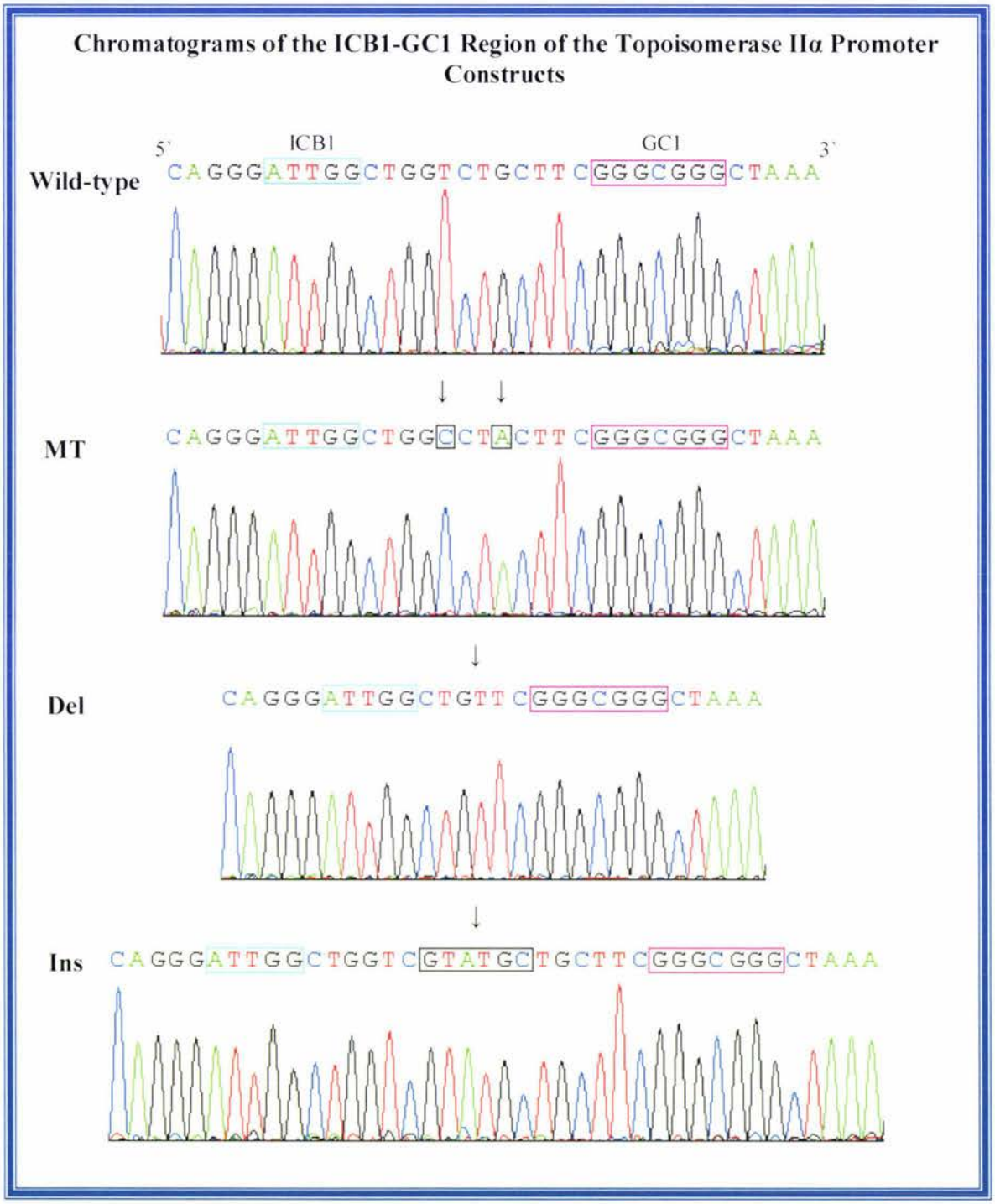


Figure 4.8: Chromatograms of the new topo IIα promoter mutants.

Automated sequence chromatograms of the forward 5'-3' strands encompassing the ICB1-GC1 regions. Comparison between the wild-type chromatogram (top) and the mutants reveal the mutations produced within the topoisomerase IIα promoter (arrows).

4.3 Combining the GC2 Mutation with the ICB1-GC1 Centre Region Mutations.

A topoisomerase IIα GC2 mutant had been previously prepared (M^{c}Lenachan, 1998; Szremska, 2000), and introducing this mutation into the newly generated ICB1-GC1 centre region mutants would allow investigation of any combined affects on transcription. Functional assays using the double mutants may provide insights into putative interactions occurring between proteins bound at this distal element and those associated with the ICB1-GC1 region.

Generation of these double mutants was achieved using the same process as that of the single ICB1-GC1 centre region mutants (refer to section 4.2). The only difference being that the first PCR mutagenesis reaction was carried out using the topo IIα GC2 mutant promoter as the template DNA. The internal centre region primers would therefore incorporate their mutations within this sequence already containing the GC2 mutation, producing the double mutants. The process from here was exactly the same, with the same restriction endonuclease enzymes used and purification methods carried out. The transformation of *E.coli* XL-1 Blue cells with the double mutant plasmids was successful and positive colonies were grown for plasmid extraction and automated sequencing. Results of the automated sequencing reactions showed the GC2 mutant was present in conjunction with the ICB1-GC1 cente region mutations (refer to Figure 4.9 for chromatograms and Appendix 5 for the sequence alignments). Glycerol stocks were also produced with these new clones for long term storage and future production of the plasmids. These were prepared as described in section 2.3.3 and stored at -70^{O}C.

4.4 Chapter Summary.

New topoisomerase IIα mutant promoters were successfully cloned using PCR mutagenesis to introduce mutations, deletions, and insertions into the wild-type promoter construct and GC2 mutant promoter construct. All inserts were completely sequenced to confirm identity of the introduced mutations, and absence of additional errors introduced by PCR due to the inherent infidelity of the polymerase. These plasmids could then be

used in functional assays to investigate the role of these sequences and their associated proteins (refer to Chapter 5).

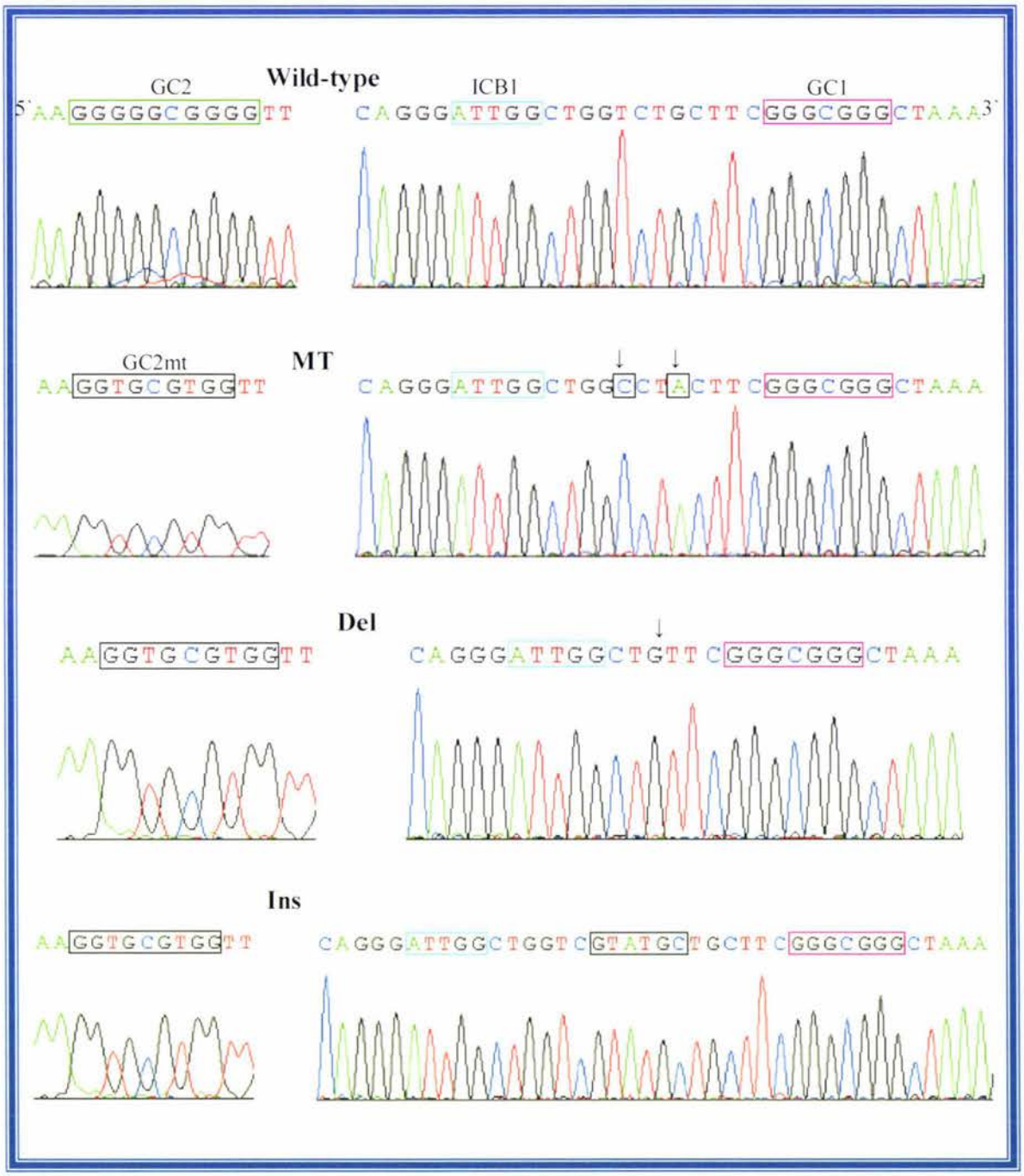


Figure 4.9: Chromatograms of the double GC2 and ICB1-GC1 centre mutants.
Automated sequence chromatograms of the forward 5'-3' strands encompassing the GC2 and ICB1-GC1 regions. Comparison between the wild-type chromatogram (top) and the mutants reveal the mutations produced within the topoisomerase IIα promoter (arrows and black boxes).

Chapter 5: Transient Transfections.

5.1 Introduction.

The protein binding assays performed in Chapter 3 showed that NF-Y, Sp1, Sp3, and the uncharacterised protein all bind within the ICB1-GC1 region of the topoisomerase IIα (topo IIα) promoter. The 12 bp sequence in the centre region between these two DNA binding elements (ICB1 and GC1) was shown to be an important requirement for binding of the uncharacterised protein. Inhibition of binding was apparent when a two base pair mutation or a 6 base pair deletion was generated within this centre region sequence, confirming the requirement of this sequence. A 6 base pair insert in the centre of this sequence actually increased the binding of the uncharacterised protein. An investigation of the functional significance of these alterations would help to deduce the role of the uncharacterised protein in transcription from the topo IIα promoter.

These mutated topo IIα promoter sequences were therefore inserted into the reporter plasmid pGL3 Basic as described in Chapter 4. Double mutants were also created by inserting each of the centre region mutations in conjunction with a topo IIα GC2 mutation, known to prevent binding of both Sp1 and Sp3 to this element. Combining these mutations should enable observations of any additive effects that may occur, as a result of interactions between proteins bound at the distal GC2 element and the more proximal ICB1-GC1 region.

5.2 Transient Transfections.

Transient transfections involve the insertion of exogenous DNA into living cells, which provides a convenient method for measuring transcriptional activity of a promoter in different cell lines. This occurs without DNA integration and the success of the transfection depends on the number of cells which take up the vector and how well it is expressed. The promoter of interest is cloned into the vector directly upstream of a reporter gene, the cell is then transfected with this whole construct and the activity of the promoter can be then measured by a reporter gene assay. Cells are co-transfected with a second reporter vector as a control for transfection efficiency (refer to Figure 5.1). The assumptions in this experimental system are that the cloned promoter controls the reporter gene as it would endogenously, and that the efficiency of the transfection is the same for both the reporter and control vectors. Therefore, the activities of the enzymes encoded by the reporter genes are directly proportional to the efficiency of the promoter.

As described in Chapter 4 the topo IIα promoter constructs generated were cloned directly upstream of the firefly luciferase gene in the pGL3 Basic vector (Promega, Appendix 3), the transfection efficiency control used was pCMV Sport-βgal (Invitrogen, Appendix 3). FuGene^{TM}6 (Roche), a liposomal transfection reagent, was used to transport the DNA into the cells. This method is called lipofection as it utilises liposomes prepared from cationic lipids which spontaneously form cationic-DNA complexes. These are positively charged and fuse with the negatively-charged plasma membrane of cultured animal cells transforming the cells with the exogenous DNA. The transfections were carried out using HeLa cells (refer to section 2.5.1), and cell extracts were harvested (refer to section 2.5.2) for the subsequent assays. This was carried out approximately 48 hours post-transfection as this is when the amount of protein is at a maximum before expression starts to decline due to the loss of gene expression.

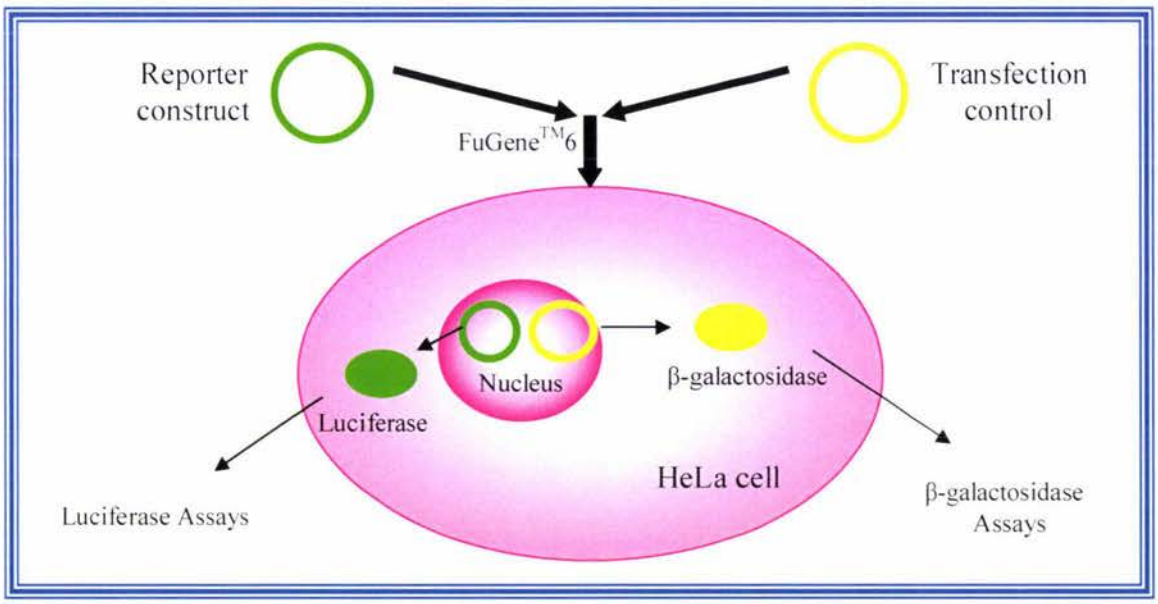


Figure 5.1: Schematic representation of a transient transfection.

The reporter construct containing the luciferase gene and the transfection efficiency control vector were co-transfected into the mammalian cells using the transfection reagent FuGene™6. The vectors enter the nucleus where they are transcribed relative to the efficiency of the cloned promoters. Cell extracts are harvested after 40-48 hours for luciferase and β-galactosidase assays.

5.2.1 β-galactosidase Assays.

An internal transfection control was used to account for differences in DNA uptake and expression, cell density and integrity. The control vector used was pCMV Sport-βgal (Invitrogen, Appendix 3) which constitutively expresses β-galactosidase at high levels. The amount of expression was measured by the cleavage of colourless ONPG (o-nitrophenyl-beta-D-galactopyranoside) to yellow ONP (o-nitrophenol) (Figure 5.2), which was measured spectrophotometrically (refer to section 2.5.3).

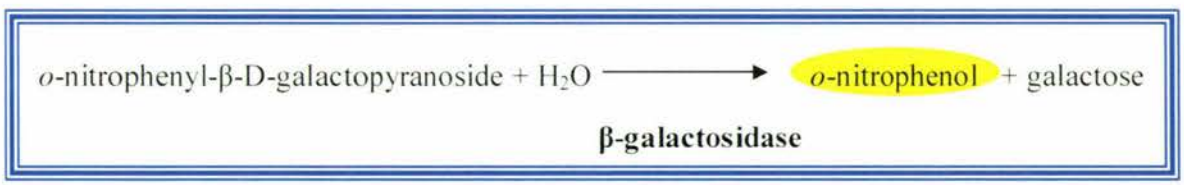


Figure 5.2: ONPG is hydrolysed by β-galactosidase.

Expression levels of β-galactosidase were calculated by measuring the amount of yellow *o*-nitrophenol hydrolysed from ONPG in the assay shown.

5.2.2 Luciferase Assays.

The luciferase gene is from the North American firefly *Photismus pyralis*; the enzyme it encodes can catalyze the oxidation of luciferin into oxyluciferin which requires ATP and oxygen. This reaction emits a yellow-green light as a by-product which can be directly measured as the amount of luciferase activity (Figure 5.3), and therefore relative promoter strength.

Assays were carried out using the Promega luciferase assay kit and the FLUOstar galaxy microplate reader (BMG labtechnologies) (refer to section 2.5.4). This takes a photon count every second over 3 mintues once the luciferase reagent is added. Photons are detected by the reading head which is automatically maneuvered over each microplate well. The counts are then relayed through Excel^{TM} (Microsoft 97) and are actual photon counts over time.

The addition of the luciferase reagent gives a flash of light relative to the quantity of enzyme present which then rapidly decays, therefore the maximum reading over the whole measurement period was used for analyses.

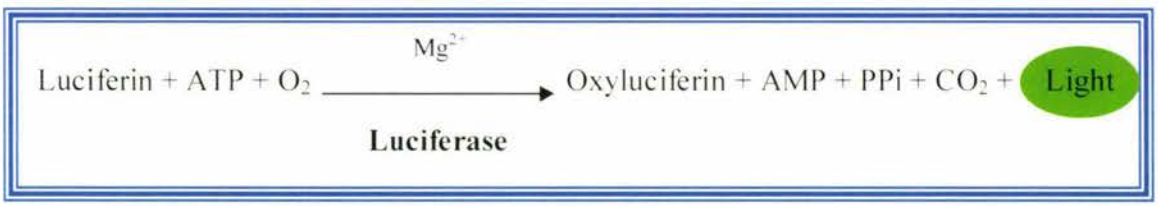


Figure 5.3: Luciferase assay reaction equation.

The oxidation of luciferin into oxyluciferin requires ATP and oxygen. This reaction emits a yellow-green light as a by-product which can be directly measured as the amount of luciferase activity.

5.2.3 Data Analyses.

To account for slight variations seen between transfection assays, each transfection was performed in triplicate and repeated at least 3 times to ensure accuracy. Three blank samples were assayed in both the luciferase and β-galactosidase assays, and the averages

of these were subtracted from each of the actual sample readings to eliminate any background signals (refer to Appendix 7 for example). The level of luminescence from a particular sample was normalised to the β-galactosidase activity of that same sample (refer to Figure 5.4).

$$\text{Normalised Value} = \frac{\text{Luminescence intensity}}{\beta\text{-galactosidase activity } (A_{510nm})}$$

Figure 5.4: Equation used to normalise transient transfection data.
Cell extracts were assayed for both their luminescence and β-galactosidase activity and then normalised against each other to compensate for internal expression variations.

The averages of each triplicate were used for statistical analyses. These could also be expressed as a percentage of the wild type promoter activity by dividing each value by the wild type value and multiplying the result by one hundred (refer to Appendix 7). An increase or decrease relative to the wild type will illustrate if a particular promoter construct is stronger or weaker than the original promoter at initiating expression.

The average deviation was calculated for each triplicate and included in the graphs as error bars. This illustrates the amount of variation within the data sets from the calculated mean (refer to Appendix 8.1). Averages of each triplicate were also analysed using a one tailed, 2 sample equal variance Students T-test (Appendix 8.2 and Table 5.1), to assess and verify the statistical significance of any differences observed between data sets.

Students T-tests were carried out using Microsoft Excel 2002. This test employs a *t*-distribution which accounts for small sample sizes, eliminating the problems associated with using a normal-distribution and a small sample size. It assumes the means of both data sets are equal and calculates the probability of them being identical, called a p-value (Table 5.1).

p-value	Description	Difference
$p > 0.10$	No significant difference between samples	No
$0.05 < p \leq 0.10$	Slight evidence for a significant difference between samples	Yes
$0.01 < p \leq 0.05$	Moderate evidence for a significant difference between samples	Yes*
$0.001 < p \leq 0.01$	Strong evidence for a significant difference between samples	Yes**
$p \leq 0.001$	Very strong evidence for a significant difference between samples	Yes***

Table 5.1: Students t-test p-value descriptions.

The p-values generated by the Students t-test are the probabilities of the data sets being identical. These were used to verify the statistical significance of any differences observed between data sets.

5.3 HeLa Cell Transfections.

5.3.1 Controls.

HeLa cells were transfected in 12-well plates at approximately 70% confluence with 0.25-1.5 μg of the relavant DNA constructs (refer to sections 2.5.1 and 2.5.2). Initially the wild type pGL3B topo IIα -617 promoter was used to establish optimal transfection conditions such as cell harvesting time, DNA amounts, cell confluence levels and the effects of the pCMV Sport β-gal construct on luciferase levels. After harvesting the cells, β-galactosidase and luciferase assays were performed as described in section 2.5.3 and section 2.5.4.

HeLa cells were harvested at 24, 36 and 48 hours post-transfection to uncover the optimal time of incubation for the transfected cells. The highest amount of luciferase activity was recovered after a 48 hour incubation period (refer to Figure 5.5). There was an almost linear relationship seen between cell incubation time and luciferase activity for both confluent and non-confluent cells.

Confluence levels of HeLa cells were thought to be optimal for transfection at approximately 70%, as the topo IIα promoter is known to be down-regulated once confluence is reached (Isaacs, 1996). The level of luciferase activity could therefore decrease as confluence was reached compared with non-confluent cells. This was seen with cells incubated for 24 and 36 hours (refer to Figure 5.5), however the amount of luciferase activity from confluent cells was actually higher after an incubation period of 48 hours compared with that of the 70% confluent cells. Transient transfections can cause cell death, which may have been why there was a higher amount of luciferase activity from the confluent cells as there were more cells to begin with, and so more could have survived the longer incubation period with the transfected DNA. The differences observed between the confluent and non-confluent cells at the relevant harvest times are each statistically significant (refer to Table 5.2), evidence cell confluence is an important factor to keep constant throughout transfection experiments.

The amounts of the DNA constructs used in each transfection were titrated as shown in Figure 5.6. The pCMVSport β-gal vectors alone showed no background luciferase activity as expected. Increasing amounts of the β-gal vector combined with increasing pGL3B topo IIα -617 wild type reporter constructs decreased the amount of luciferase activity. Cells transfected with 0.5 μg of β-gal vector showed a significant decrease in luciferase activity when compared with those transfected with only 0.25 μg β-gal vector, this was the case in co-transfections with 0.5 μg or 1.0 μg of the WT construct (refer to Table 5.3). The amounts of DNA used were all equalised with the empty vector pBluescript SK-II (Stratagene, USA), therefore eliminanting DNA concentration as the cause of differing luciferase levels. The pattern of the luciferase expression is almost linear when only 0.25 μg of β-gal vector is used. The highest luciferase readings were seen with 0.25 μg β-gal vector and 1.0 μg of reporter construct.

From these results, all subsequent transfections were carried out using 70% confluent HeLa cells and an incubation period of 48 hours post-transfection. The amount of control pCMVSport β-gal vector used was 0.25 μg and the amount of reporter constructs used was 1.0 μg.

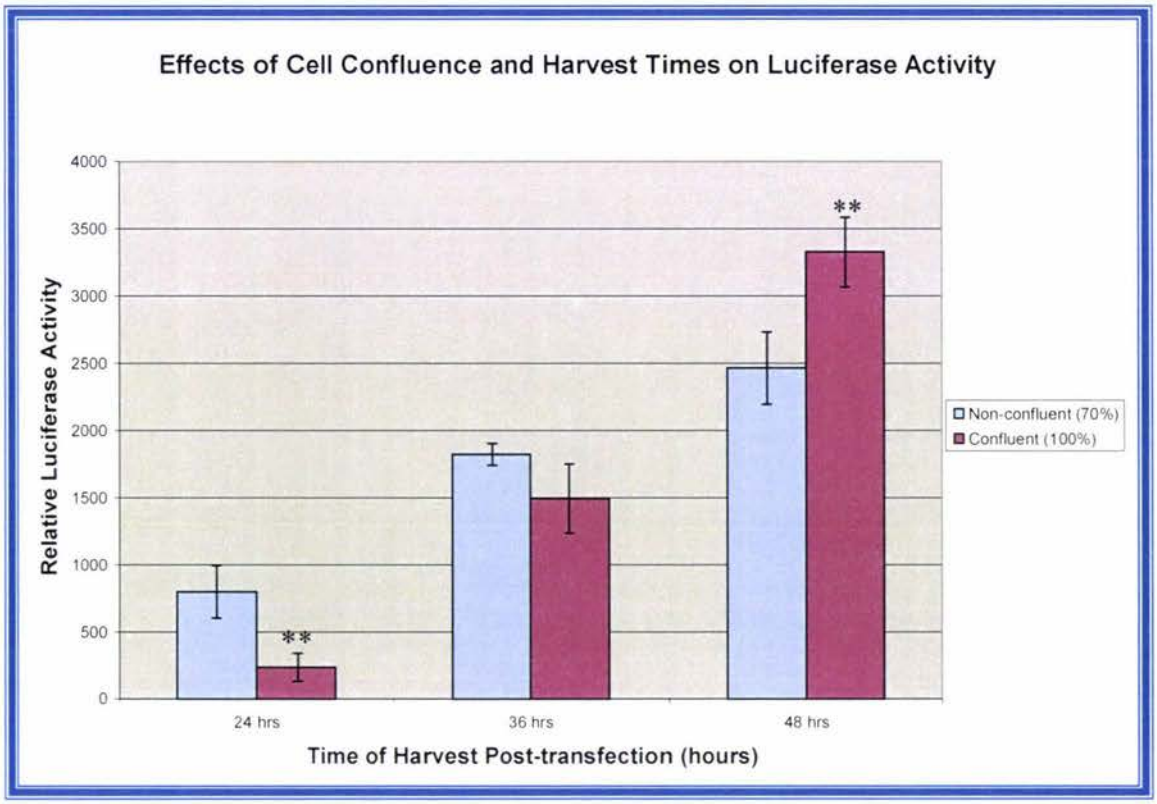


Figure 5.5: The influence of cell confluence and harvest times on luciferase activity.
Non-confluent and confluent HeLa cells were transfected with 0.25 µg pCMVSport β-gal control construct and 0.5 µg pGL3B topo IIα -617 wild type reporter construct. Sets of triplicates were harvested at 24, 36 or 48 hours post-transfection and assayed for β-galactosidase and luciferase activity. Luciferase values were normalised with β-galactosidase values and graphed with error bars which denote the average deviation between the data sets. These are results of at least 3 experiments which were done in triplicate. Asterists indicate significant difference **=0.01 (refer to Table 5.2).

Non-confluent cells vs. Confluent cells		
Harvest Times	**p-value**	**Significant Difference**
24 hours	0.005	Yes**
36 hours	0.070	Yes
48 hours	0.008	Yes**

Table 5.2: Significant differences from Figure 5.5.
Significance represented by Yes or No and *s indicating level of evidence (refer to Table 5.1).

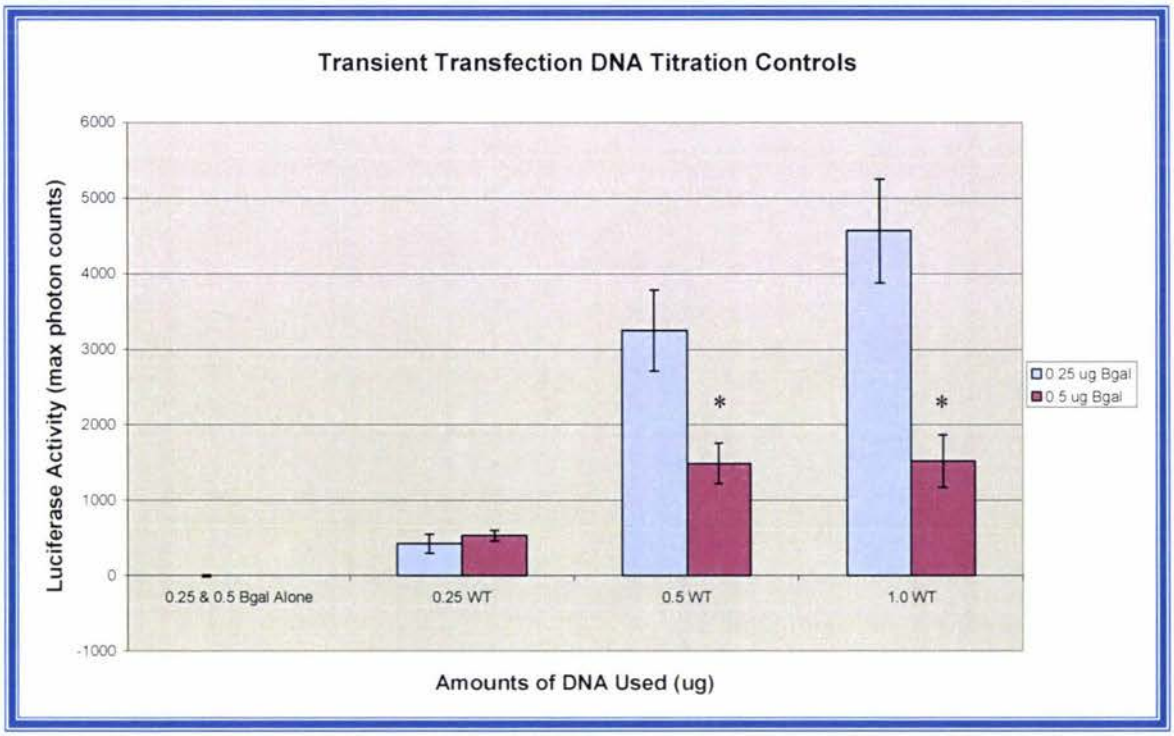


Figure 5.6: Transient transfection titrations of control and reporter DNA construct amounts.

HeLa cells were transfected with 0.25 µg or 0.5 µg of pCMVSport β-gal (Bgal) control vector and either 0.25 µg, 0.5 µg or 1.0 µg of pGL3B topo IIα -617 wild type (WT) reporter construct. Cells were harvested at 48 hours post-transfection and assayed for β-galactosidase and luciferase activity. Maximum luciferase photon counts were graphed with error bars which denote the average deviation between the data sets. These are results of at least 3 experiments which were done in triplicate. Asterists indicate significant difference *<0.05 (refer to Table 5.3).

0.25 µg βgal vs. 0.5 µg βgal		
Amount of WT construct	**p-value**	**Significant Difference**
0.25 µg	0.126	No
0.5 µg	0.018	Yes*
1.0 µg	0.054	Yes

Table 5.3: Significant differences from Figure 5.6.

Significance represented by Yes or No and *s indicating level of evidence (refer to Table 5.1).

5.3.2 Trouble Shooting.

It was noted that the levels of luciferase activity were very low compared with that expected from previous experiments using the wild type topo IIα -617 promoter construct which had relative luciferase values ranging from 8000-50,000 (Magan, 2002). In order to see significant differences between the various promoter constructs luciferase activity needed to be higher than the 4000-6000 relative luciferase values achieved in this current study. A number of trials were carried out in an attempt to gain higher luciferase values including: reproducing all the DNA vectors, requantifying the plasmid DNA being used for the transfections, and preparing new buffers and reagents for both the transfection procedure and subsequent luciferase and β-galactosidase assays.

Transfections were also carried out with and without pCMVSport β-gal control vector to determine any effect on the luciferase values. As shown in Figure 5.7 the β-gal control vector significantly reduces the amount of luciferase expression seen upon addition in conjunction with the pGL3B topo IIα -617 wild type reporter construct (WT) (p-value = 0.07, refer to Table 5.1). Each transfection contained exactly the same amount of DNA (1.25 μg total) to eliminate effects on luciferase activity caused by cell death due to higher DNA concentrations in some transfections. An empty vector (pBluescript SK-II, Stratagene, USA) was used to equalise amounts of DNA for each transfection. The empty pGL3 Basic vector had no luciferase activity as expected. The error bar is extremely large in the transfection with wild type construct alone, showing variation of transfections without a β-galactosidase control vector. The transfection with both the β-gal control vector and the wild type reporter construct has a much smaller margin of error; however luciferase activity is approximately 5 times lower than those transfections without the control vector.

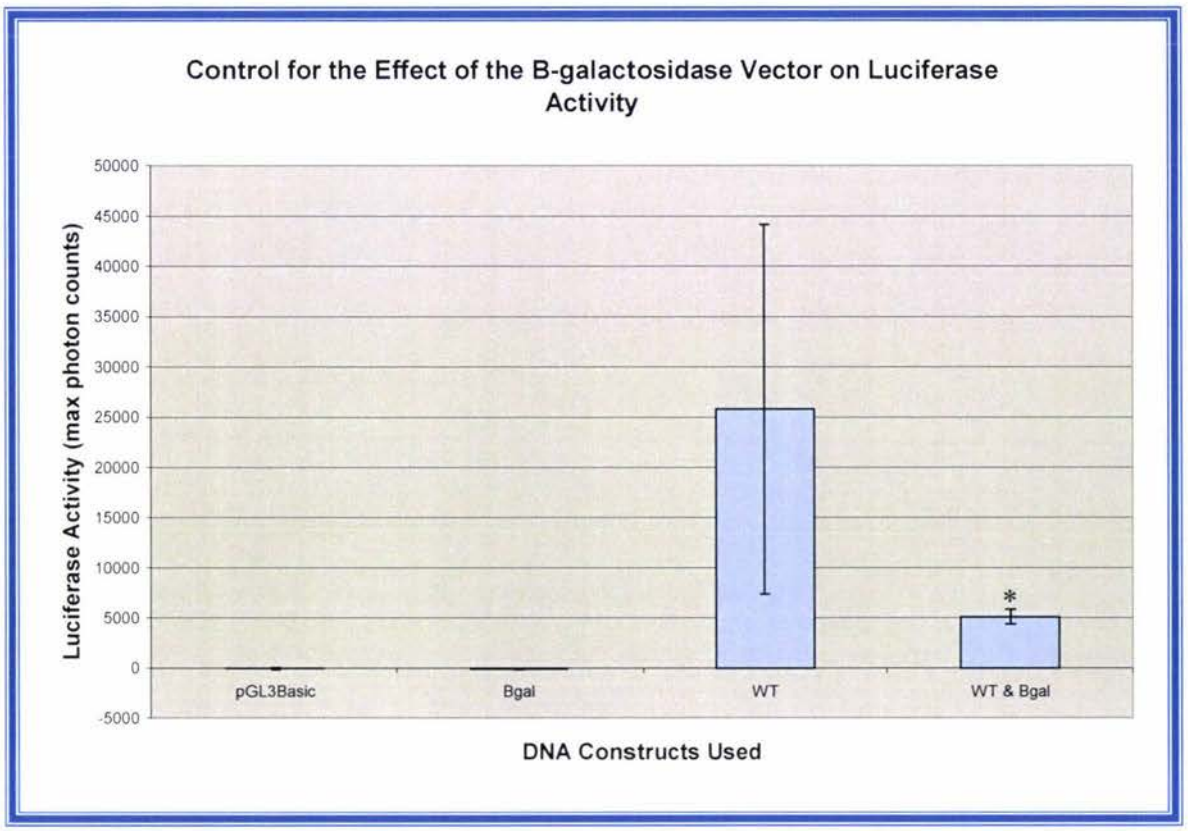


Figure 5.7: Transfection control for the effect of the β-gal vector on the activity of luciferase.

HeLa cells were transfected with a 1.25 µg total amount of DNA, comprising of either pGL3 Basic, pCMVSport β-gal (Bgal) control vector, pGL3B topo Hu -617 wild type (WT) reporter construct or a combination of Bgal & WT constructs (0.25 µg and 1.0 µg respectively). To make sure the DNA amounts were exactly the same an empty vector (pBluescript SK-II) was used. Cells were harvested at 48 hours post-transfection and assayed for β-galactosidase and luciferase activity. Maximum luciferase photon counts were graphed with error bars which denote the average deviation between the data sets. These are results of at least 3 experiments which were done in triplicate. There is a significant difference between WT alone transfections and WT & β-gal co-transfections (p-value = 0.07) (refer to Table 5.1).

To test if the effect of the pCMVSport β-gal vector on luciferase activity was specifically caused by the pCMVSport vector a second control experiment was performed. This involved using an alternative β-gal vector, pSV β-gal (Promega, USA). A maspin promoter (pGL3B Maspin -2551 wild type), which is known to have high luciferase expression in HeLa cells, was also used to investigate whether the effect on luciferase expression from the β-gal vector was specific for the topisomerase IIα promoter construct. The results of these experiments are shown in Figure 5.8.

The pGL3B topo IIα reporter construct (WT) alone had a significantly higher luciferase expression compared to both co-transfections with pCMVSport β-gal (p-value 0.0089) and pSV β-gal (p-value 0.002) (refer to Figure 5.8 and Table 5.4). Both the β-gal vectors alone showed no luciferase activity as expected.

The Mapsin promoter construct alone had 3 times the amount of luciferase activity as the topo IIα promoter construct (WT) alone (refer to Figure 5.8). Luciferase activity was not significantly reduced with the pSV β-gal vector as a transfection control, which is the opposite result seen with the topo IIα promoter. Luciferase activity expressed from the maspin promoter however, was significantly lower when pCMVSport β-gal vector (p-value 0.015) (refer to Table 5.4) was used as a transfection control.

Despite amounts of DNA being exactly the same in the transfections there is less luciferase activity seen when the reporter construct was co-transfected with a β-gal vector. This was interesting as the two vectors do not share similar transcription elements and so should not have a negative effect on one another. Overall luciferase values were lower when compared with experiments carried out previously using the same vectors. Sequencing analyses revealed an absence of mutations from any of the constructs; therefore a more fundamental problem may exist.

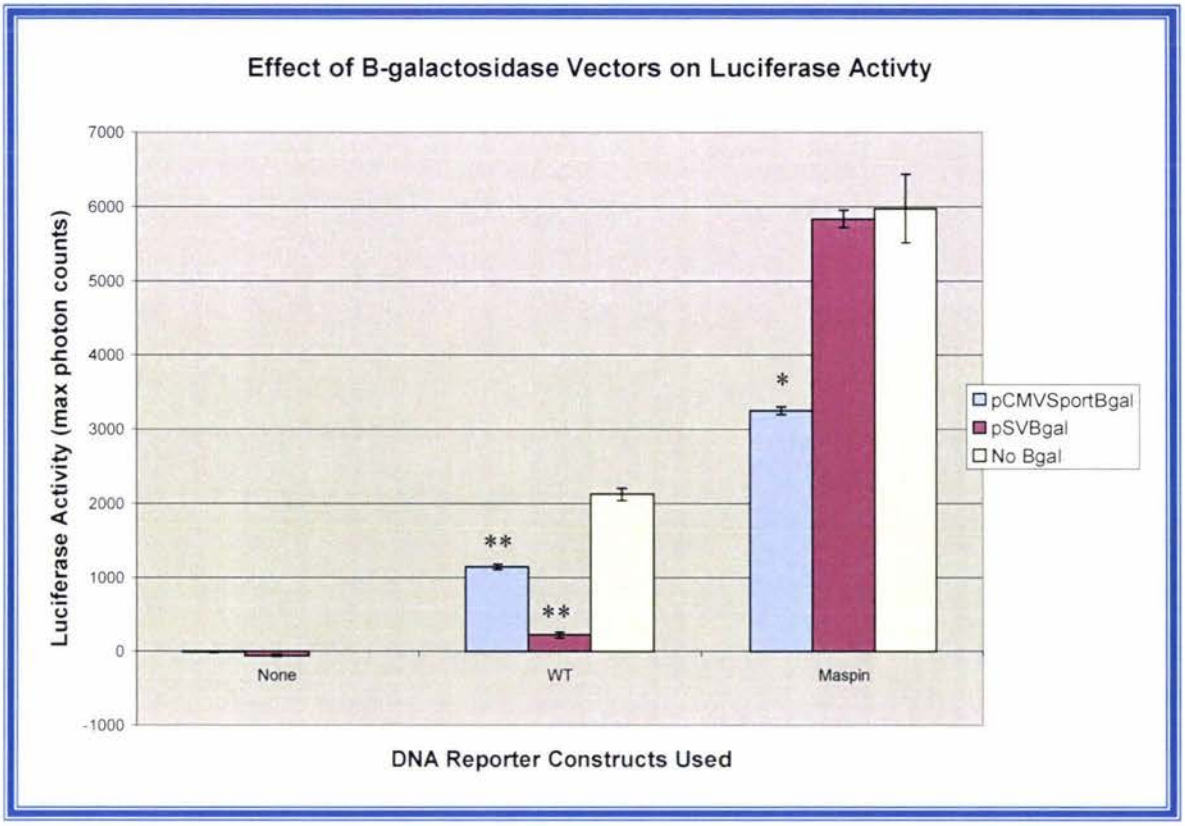


Figure 5.8: Control to test effects of different β-gal vectors on luciferase activity from different promoters.

HeLa cells were transfected with a 1.25 µg total amount of DNA, comprising of 0.25 µg of either pCMVSport β-gal (pCMVSportBgal) vector, or pSVβ-gal (pSVβgal) and co-transfected with 1.0 µg pGL3B topo IIα -617 wild type (WT) reporter construct or each construct alone. To make sure the DNA amounts were exactly the same an empty vector (pBluescript SK-II) was used. Cells were harvested at 48 hours post-transfection and assayed for β-galactosidase and luciferase activity. Maximum luciferase photon counts were graphed with error bars which denote the average deviation between the data sets. These are results of at least 3 experiments which were done in triplicate. Asterists indicate significant difference *<0.05, **<0.01 (refer to Table 5.4).

Promoter Construct	Compared to:	p-value	Significant Difference
WT topo IIα alone	pCMVSportβ-gal & WT	0.0089	Yes**
	pSVβ-gal & WT	0.0020	Yes**
Maspin alone	pCMVSportβ-gal & Maspin	0.0150	Yes*
	pSVβ-gal & Maspin	0.7425	No

Table 5.4: Significant differences from Figure 5.8.

Significance represented by Yes or No and *s indicating level of evidence (refer to Table 5.1).

5.4 Chapter Summary.

Transient transfection assays can be used to test the function of DNA elements or protein interactions *in vivo*. These are often used to support *in vitro* data such as electrophoretic mobility shift assays. In this case, a difference in protein binding patterns was observed upon the introduction of mutations, deletions and insertions in a particular region within the topo IIα minimal promoter (refer to Chapter 3). To test the functional significance of these, transient transfections were to be carried out.

Initial control transfections and titrations gave expected results with linear relationships seen between amounts of DNA and incubation times, however luciferase values were lower than expected. A number of attempts were made to improve the quality of these results and define the cause of the low readings. These included reproducing all the DNA constructs, using new buffers and reagents, sequencing all plasmids, and requantifying plasmid DNA.

Trouble shooting experiments showed the transfection of the β-gal vectors in conjunction with the luciferase reporter vectors significantly compromised the amount of luciferase activity seen. This effect was not only seen with the topo IIα promoter construct, but also with an unrelated Maspin promoter construct. The reason for this phenomenon is, as of yet, unknown. Unfortunately due to time restraints the major issues were not resolved and so the significance of the protein binding behaviour seen in Chapter 3 remains undefined.

Chapter 6: Discussion and Future Research.

6.1 Overview.

Cancer is caused by the breakdown of the mechanisms controlling cell division, allowing cells to become immortal and proliferate uncontrollably which results in the development of tumours. The high proliferation rate of the tumour cells results in an upregulation of enzymes essential for cell growth and division, such as topoisomerase IIα. The topoisomerase IIα (topo IIα) gene encodes a cell cycle regulated 170 kDa nuclear enzyme with maximum expression in the G_2/M phase. This is due to it's critical requirement for chromosome condensation and segregation, relaxing supercoiled DNA, and untwisting and unknotting intertwined DNA (Robert and Larsen, 1998). It achieves these roles by utilizing ATP and introducing a double-stranded break in the DNA, then passing another double-strand of DNA through and resealing the break (Berger *et al.*, 1996).

A number of anti-cancer drugs have been designed to primarily target topo IIα (Cornarotti *et al.*, 1996; Satherley *et al.*, 2000) due to the fact it is essential for cell survival and division, and also found at such high levels in proliferating cells, such as tumour cells. These anti-cancer drugs act as either inhibitors or poisons of topo IIα. Inhibitors interfere with normal topo IIα functions causing topological problems which result in apoptosis. The poisons prevent topo IIα from resealing the DNA breaks it produces making itself a cellular poison, which causes many DNA breaks and leads to apoptosis (Walker *et al.*, 1991). Therefore, the amount of topo IIα correlates with the sensitivity to the topo IIα poisons. The success of these topo IIα targeting drugs is limited by toxicity and the development of resistance.

Resistant cell lines have been shown to exhibit down-regulation of topo IIα gene expression, eliminating the target for the drugs, therefore decreasing the sensitivity of the cells (Wang *et al.*, 1997a). Maximal expression of topo IIα is observed from a -617 bp truncated version of the ~2 kb TATA-less promoter, which has a number of sequence motifs, including 5-inverted CCAAT boxes (ICB1-5) and two GC-rich regions (GC1 and

GC2) (Hochhauser *et al.*, 1992). Transcription factors have been shown to bind to these DNA elements and regulate expression of the topo IIα gene. Nuclear factor Y (NF-Y) is a ubiquitously expressed trimeric protein known to bind the ICB elements in the topo IIα promoter (Isaacs, 1996). The Specificity proteins Sp1 and Sp3 are also ubiquitously expressed and known to bind GC1 and GC2 (Yoon *et al.*, 1999). The results of *in vivo* functional assays have suggested complex protein interactions occuring between these proteins result in modulation of topo IIα expression in response to various external factors. Previous research on the 100 bp region closest to the transcription start site encompassing the ICB1 and GC1 elements, showed NF-Y binds to the ICB1 element and Sp1 binds to GC1 box and both activate transcription of the topo IIα promoter (Magan *et al.*, 2003). Along with Sp1, the Sp3 protein was found also associated with the GC1 box *in vivo*; however it appears to act as a repressor in functional assays. A fourth uncharacterised protein was also found to specifically bind this region of the topo IIα promoter (Magan, 2002).

This uncharacterised protein may have an influence on topo IIα transcription, and may possibly have a significant role in the down-regulation of topo IIα expression seen in drug-resistant cells. It's characterisation and subsequent identification may provide useful information concerning topo IIα expression, and potentially contribute information useful for improving anti-cancer drug treatments. Therefore, the aims of this study were to identify the minimal DNA sequence or proteins required for binding of the uncharacterised protein, and to test the functional significance of its binding within this specific region of the topo IIα promoter.

6.2 Summary of Results.

6.2.1 Results of DNA Binding Assays.

The uncharacterised protein was found, along with transcription factors NF-Y, Sp1 and Sp3, associated with the two most proximal DNA elements within the topo IIα promoter, ICB1 and GC1 (Magan, 2002). It was thought the uncharacterised protein may be

binding to the region between the ICB1 and GC1 elements. The distance between these two elements is 12 bp, which is just over one complete turn of the DNA double helix, sufficient room to accommodate a transcription factor. Unfortunately, a database search using this 12 bp sequence revealed no candidate protein binding motifs. On the other hand, the protein may not require the specific DNA sequence between the elements, but rather the specific distance between the elements, for it to be recruited by either NF-Y or Sp1/3 to the complex. It may play a synergistic role mediating an interaction between NF-Y and Sp1/Sp3, and could be acting as either a co-activator or a co-repressor depending on the complement of specific transcription factors associated with the DNA at each of these sites.

Electrophoretic mobility shift assays were performed using a number of variations of the original ICB1-GC1 region DNA probe, to elucidate the specificity and sequence requirements for the uncharacterised protein to bind (Chapter 3). Probes that could potentially bind the uncharacterised protein were produced by introducing a 2 bp mutation, a 6 bp deletion, and a 6 bp insertion in the centre of the 12 bp region between the ICB1 and GC1 elements (Figure 3.2). A 12 bp probe of the centre region alone was also synthesised to test if the uncharacterised protein could bind this sequence alone.

The results of the DNA binding assays performed with these probes are summarized in Figure 6.1. The 44 bp wild type (WT) probe was shown to bind NF-Y, Sp1, Sp3, and the uncharacterised protein as previously shown by Magan, 2002. The mutant (MT) and deletion (Del) probes bound only NF-Y, Sp1, and Sp3. The insertion (Ins) probe was found to bind all four proteins. The probe encompassing only the 12 bp centre region (Cen) did not exhibit any protein binding. Antibody supershift assays confirmed the identity of the proteins which were binding to each probe (Figure 3.10). Quantification using phosphoimaging of these bands revealed more uncharacterised protein was binding the insert probe, and less NF-Y was bound to the MT, Del and Ins probes (Figure 3.9). The reason for the apparent increase in uncharacterised protein bound to the Ins probe, may be explained by the fact that the sequence inserted was a repeat of the 6 bp sequence in the middle of the 12 bp centre region, except for one base which was altered to prevent the creation of a known sequence motif. This explanation however, does not coincide

with the fact that the sequence alone did not bind proteins. Therefore, the other transcription factors may be required to recruit and/or increase the uncharacterised protein's binding affinity, and the increased space between the ICB1 and GC1 elements may allow more of the protein to associate.

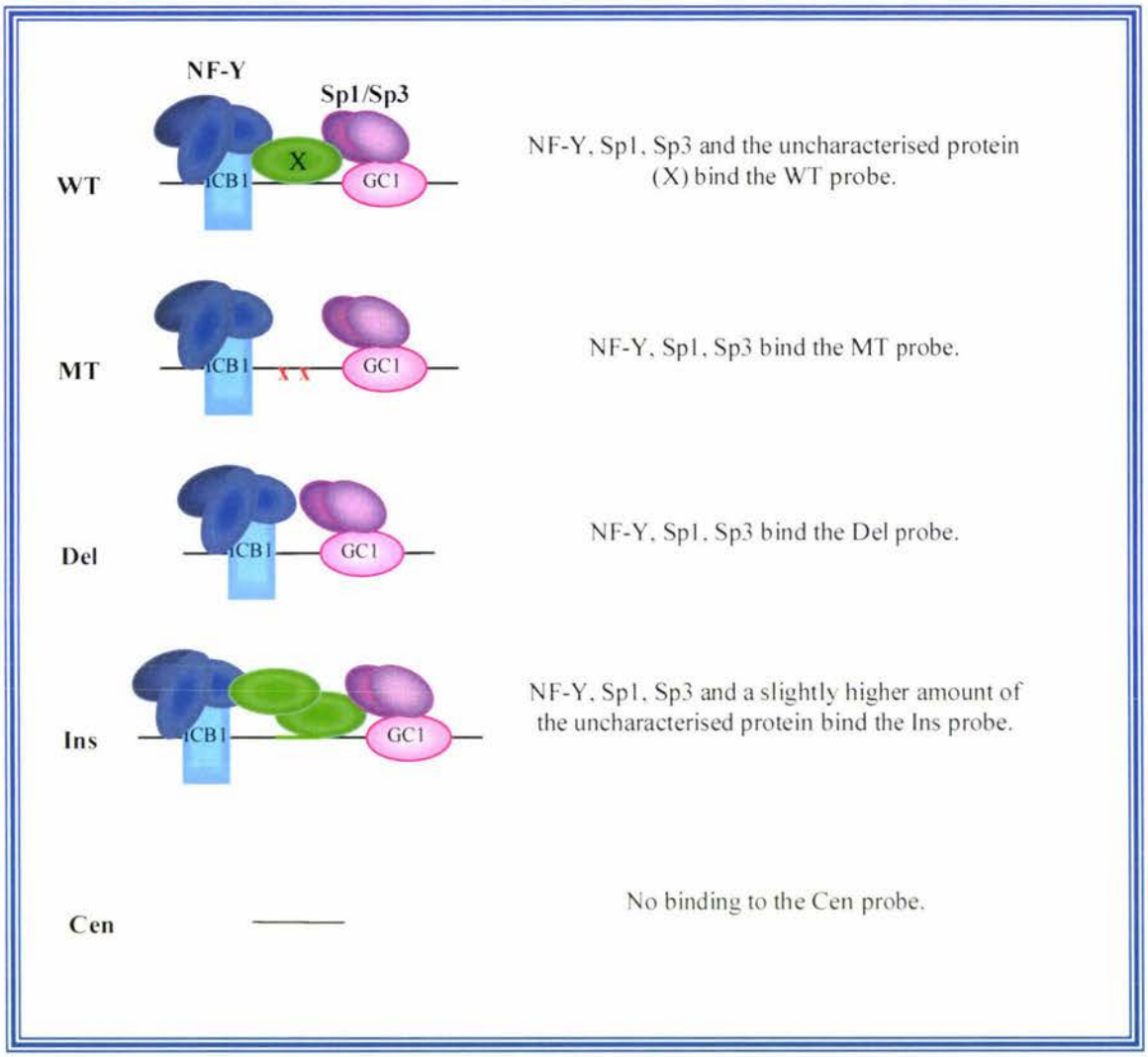


Figure 6.1: Summary of EMSA results with uncharacterised protein binding probes. EMSAs were performed using a number of variations of the original ICB1-GC1 region DNA probe, to elucidate the specificity and sequence requirements for the uncharacterised protein to bind. The uncharacterised protein binding probes were produced by introducing a 2 bp mutation (MT), a 6 bp deletion (Del), and a 6 bp insertion (Ins) in the centre of the 12 bp region between the ICB1 and GC1 elements. A 12 bp probe (Cen) of the centre region alone was also made to test if the uncharacterised protein could bind this sequence alone.

To investigate the possible recruitment of the uncharacterised protein by other transcription factors present or DNA binding elements required, another set of probes were designed. These probes were deleted in half or all of each of ICB1 and GC1 (Figure 3.3). The results of the DNA binding assays using these recruitment probes are summarized in Figure 6.2.

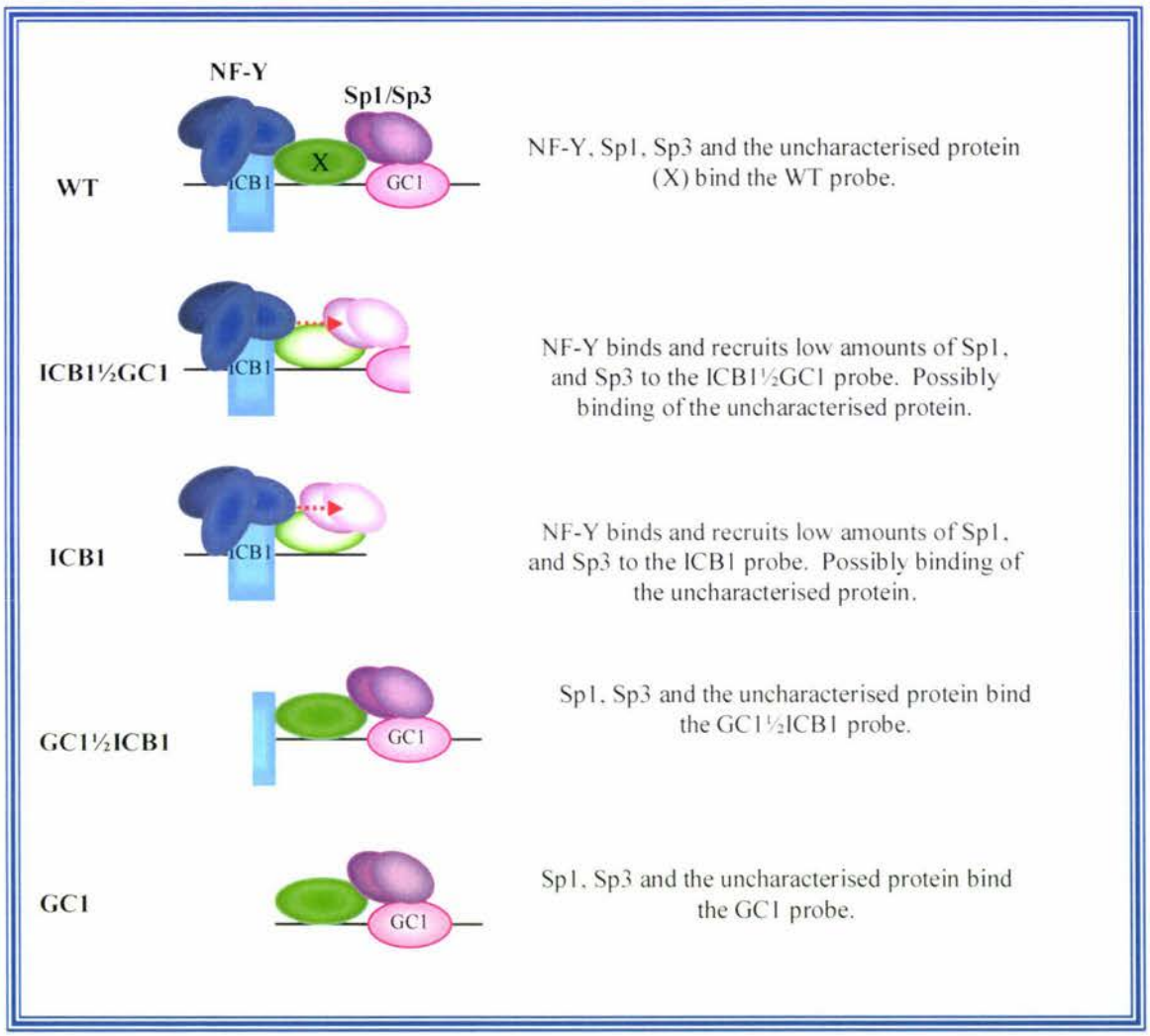


Figure 6.2: Summary of EMSA results with uncharacterised protein recruitment probes.

EMSAs were performed using a number of variations of the original ICB1-GC1 region DNA probe, to elucidate the elements or proteins required for the uncharacterised protein to bind/associate within this region. The recruitment probes were produced by shortening the original probes to remove half or all of either the ICB1 element or the GC1 element whilst leaving the centre region intact.

Again, the 44 bp wild type probe binds all four transcription factors, NF-Y, Sp1, Sp3 and the uncharacterised protein. The identities of the proteins bound to this probe and successive probes were all confirmed by antibody supershift assays (Figures 3.12 and 3.13). The NF-Y protein was thought to possibly recruit Sp1 and/or Sp3 to this region of the promoter. The experiments with the ICB1½GC1 and ICB1 probes suggest that this does occur, as with only half or none of the GC1 box there are faint bands seen with the same mobility as Sp1 and Sp3. Although faint, the addition of their respective antibodies causes the disappearance of these bands, further identifying the presence of these proteins without their DNA binding element. Another faint band is seen when the NF-Y protein has been supershifted by the NF-Y antibody. This may be residual NF-Y uncomplexed with antibody, however the mobility is higher therefore it could be the uncharacterised protein. Although the mobility seems somewhat higher than the band representing the uncharacterised protein, this seems the most likely explanation, as the centre region sequence which was shown to be important for it's binding is intact, and Sp1 and Sp3 are present at low amounts. The presence of Sp1 and Sp3 with those probes is important, as they are shown to possibly recruit the uncharacterised protein in the assays using the GC1½ICB1 and GC1 probes. These probes clearly show the binding of only Sp1, Sp3 and the uncharacterised protein. There is no NF-Y bound even to the GC1½ICB1 probe which has half the ICB1 box. This was expected as the flanking sequences of the ICB boxes have been shown to be important for NF-Y binding (Mantovani, 1998).

6.2.2 Protein Interactions.

These EMSAs suggest that the uncharacterised protein requires the correct centre region sequence and the presence of Sp1 and/or Sp3 transcription factors to associate with the topoisomerase IIα ICB1-GC1 region. This was clear as when the sequence between the two elements had just 2 bp changed, the uncharacterised protein was prevented from binding, despite the presence of each of the other transcription factors and DNA elements. The uncharacterised protein cannot bind the centre region alone however; therefore it must require either the presence of other proteins or DNA elements. It was shown that it does not require the presence of an ICB1 box or NF-Y to bind as it binds

strongly without these, therefore it is most likely to be recruited by either the GC1 box or the Sp1/Sp3 proteins bound to this element. A probe with only the ICB1 box and the full centre region sequence had Sp1 and Sp3 present at low amounts along with NF-Y, and another unidentified protein. This was possibly the uncharacterised protein binding/associating without the presence of a GC1 box. All this evidence highlights the need for Sp1 and/or Sp3 proteins, along with the centre region sequence.

These results correlate with the competitor assays, where a double element probe with a mutated ICB1 and wild type GC1 element competes for all the proteins except the NF-Y protein (Figure 3.6), but the single element wild type GC1 probe only competes for the Sp1 and Sp3 protein (Figure 3.7), confirming the requirement for the centre region sequence for the uncharacterised protein to bind.

The NF-Y protein is able to weakly recruit Sp1 and Sp3 to the region in the absence of a full GC1 box. The uncharacterised protein seemed to be recruited as well; however the recruitment is most likely caused by the presence of Sp1 and/or Sp3. Although NF-Y shows some recruitment of Sp1, Sp3 and therefore possibly the uncharacterised protein, it is not an essential requirement as these proteins bind to this region in the complete absence of NF-Y and an ICB1 box.

6.2.3 Functional Assays.

Transient transfection assays were to be carried out to test the functional affect of the uncharacterised protein on the expression levels from the topo IIα promoter. DNA alterations produced for DNA binding assays showed the uncharacterised protein's ability to bind the ICB1-GC1 region was inhibited when mutations and deletions were introduced between these two elements, and when an insert was added the amount of uncharacterised protein bound increased. These sequence alterations were successfully introduced into a -617 bp topoisomerase IIα promoter and cloned into a reporter vector.

The mutation and deletion reporter constructs, which prevent the uncharacterised protein binding, may have shown if the protein acts to repress or activate transcription depending on the level of the reporter enzyme compared to that produced from the wild type topo IIα promoter. The construct containing the insert sequence may have confirmed these results, as there is more of the protein bound. The deletion and insertion constructs may have shown a combination of transcriptional affects as removing and inserting DNA changes the distances between elements. For example, the NF-Y protein bound to ICB2 is thought to bend the DNA (Ronchi *et al.*, 1995) and bring the Sp1/Sp3 proteins bound to the distal GC2 element into close proximity with the ICB1-GC1 region allowing protein-protein interactions. By decreasing or increasing the distance between ICB1 and GC1, the protein-protein interactions at these and the GC2 element may be prevented. Therefore, reporter constructs were designed with GC2 mutants to observe any combined affect. If transcription levels remained the same between single mutants and GC2 double mutants, this would imply GC2 is important for protein interactions and that these had been prevented by the alteration of the distances between elements.

Unfortunately, the transient transfection assays did not give consistent values for any significant comparisons of the mutants to the wild type topo IIα promoter to be made. A number of controls were performed to deduce the cause of these low values, however, the only clear negative affect was from the β-galactosidase control vectors used. The reason for this is unexplained, and previous research used these same control vectors with reproducible results. It is interesting to note, this previous work was performed on HeLa cells grown in a different media. Due to time restraints, the affect of the different media on the transfections could not be tested, and therefore the cause of the low values remains unresolved.

6.3 Future Research.

6.3.1 Protein-DNA Binding Assays.

A number of other DNA binding probes would be useful to confirm results seen in this research. Schematic representations of these are illustrated in Figure 6.3. A probe encompassing only the centre region and a mutated GC1 box would confirm the necessity of Sp1 and Sp3 for the uncharacterised protein to bind the centre region (Figure 6.3(a)). To investigate if the extra band seen associated with the ICB1½GC1 and ICB1 probes is not just excess NF-Y protein not shifted by the antibody, the supershift needs to be repeated with a titration of NF-Y antibody added. If the band remains upon saturation it will indicate it is not excess NF-Y but another protein. To test whether the extra band is the uncharacterised protein, 2 bp mutations could be introduced into the centre region of those two probes (Figure 6.3(b)). The centre region sequence should be the same as the MT probe which is known to inhibit the uncharacterised protein binding. Therefore, the band disappearing from an EMSA using this probe with the mutated centre region supports the idea that this could be the uncharacterised protein; if the band remains it is most likely another unknown protein.

The increase in uncharacterised protein binding seen on the Ins probe, could be tested to investigate whether this is due to extra space between the elements or because the sequence inserted was nearly an exact repeat of the original sequence. This could be done by designing a probe with an insert of a completely random sequence (Figure 6.3(c)), which had been checked for other protein motifs prior to use. If there is still a higher amount of uncharacterised protein binding to this new Ins probe, it could be concluded that additional space is required for binding rather than the repeated sequence.

The minimal sequence the uncharacterised protein requires for binding is yet to be defined. A set of DNA binding probes designed as a deletion series removing one base at a time may show when the uncharacterised protein is unable to bind. These probes would begin with the full centre region sequence and full GC1 box, one set would begin omitting a base from the centre region end and the other would begin omitting bases from the GC1 box end (Figure 6.3(d)). Eventually the protein should not be able to bind,

therefore revealing the minimal sequence it requires. The results from the dependence the uncharacterised protein has for Sp1 and Sp3 (Figure 6.3(a)) would determine the success of this experiment. This is due to the fact that the protein may not be able to bind without Sp1 and/or Sp3, therefore the sequence identified will be that required by all the proteins combined, not the specific sequence the uncharacterised protein itself recognises.

6.3.2 Functional Assays.

The functional assays are of critical importance for characterising the role the uncharacterised protein plays in topo IIα expression. The constructs designed for these experiments should show it's role relatively easily as two of the constructs inhibit the binding of only the uncharacterised protein. The affect of an increase in the uncharacterised protein's binding and the combined affect of these in conjunction with a GC2 box mutant would paint a very descriptive picture of the type of protein it is. These experiments were performed on HeLa cells grown in Opti-MEM, whereas previous successful transient transfections were performed on HeLa cells grown in Eagle's minimal essential media (MEM) media. Therefore, it would be of interest to repeat these with Eagle's MEM to see if the different media had any affect. The change from MEM to Opti-MEM was driven by financial constraints as Opti-MEM requires considerably less FCS than Eagles MEM.

The β-galactosidase internal control vectors were shown to have a negative affect on the expression levels of the luciferase reporter gene. The reason for this is currently undefined, and the amounts of DNA used in each experiment were equalised so different amounts of DNA were not the cause. In future experiments it would be of interest to use a different internal control such as the *Renilla* luciferase reporter plasmid. Both luciferase levels can be assayed simultaneously using the Dual-Glo™ Luciferase Assay System (Promega) and ratios between the *Renilla* and firefly luciferases can be calculated, providing an alternative control for the normalisation of transfection efficiency.

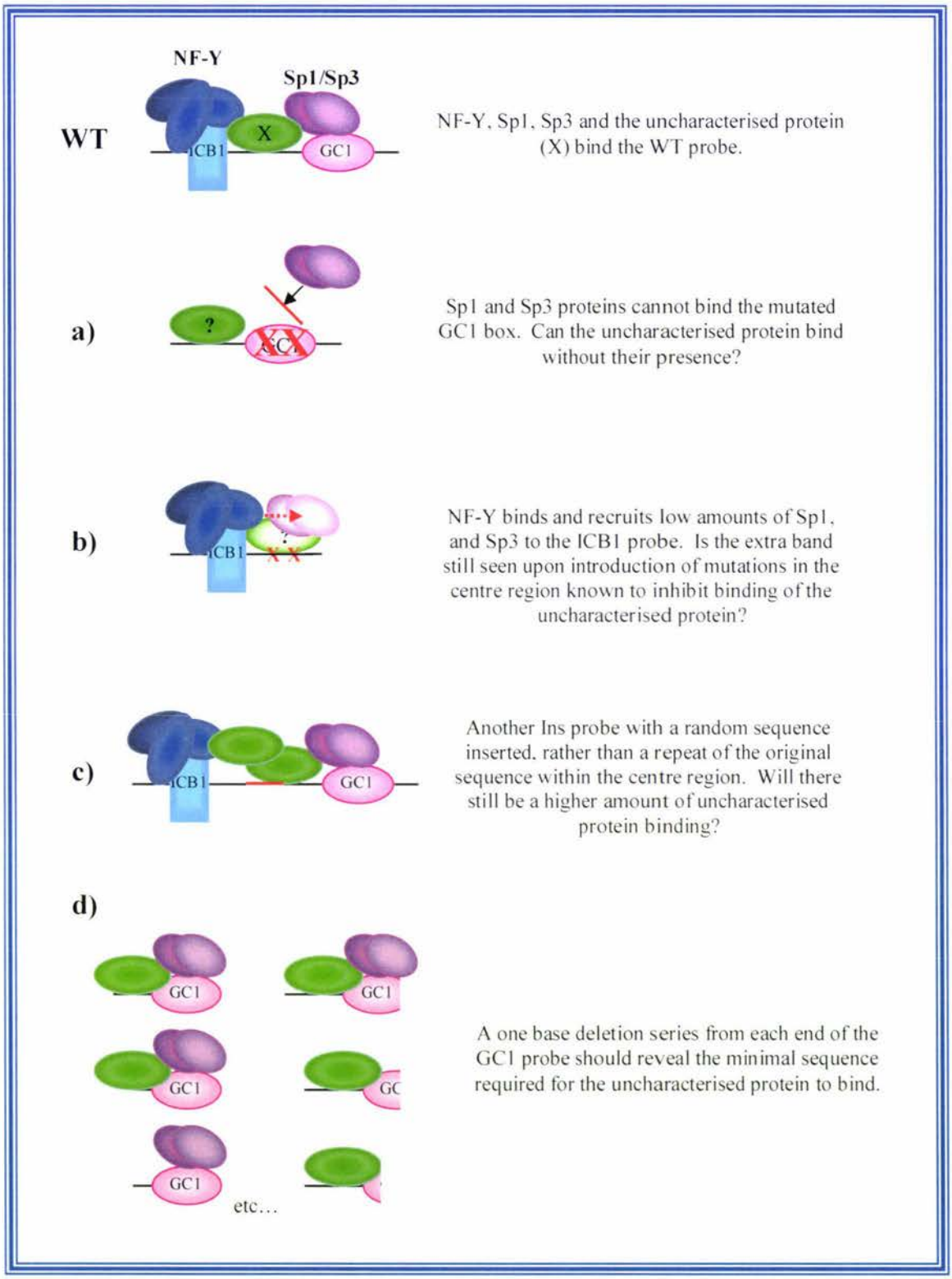


Figure 6.3: Schematic representation of future DNA probes.

6.3.3 Protein Purification.

The ultimate goal of this research would be to eventually purify the uncharacterised protein and determine the amino acid sequence by N-terminal microsequencing. Using databases a search for any sequence comparisons could be done, which may identify the protein or at least speculate it's function or the family of proteins it belongs to. To purify the protein, more information needs to be obtained from experiments such as those described in section 6.3.1. If it is found the protein binds to Sp1 or Sp3, a glutathione-S-transferase pull-down assay could be attempted using GST-Sp1/Sp3 fusion proteins.

Once the minimal binding sequence is determined, a DNA probe incorporating 5-bromo-2'-deoxyuridine (BrdU) could be designed and then used to bind the uncharacterised protein. The protein could then be crosslinked to the DNA using UV light and purified from a gel. The protein could then be run on a SDS-PAGE gel to determine it's molecular weight which will aid in it's identification and designing future purification methods. Once identified a range of functional assays could be carried out to confirm a regulatory role in transcription of topo IIα, such as dominant negative mutant assays.

6.4 Conclusion.

There are a number of complex protein interactions occurring within this small region of the topo IIα promoter. The tentative process is most likely started with the binding of NF-Y to the ICB1 box, as NF-Y is known to bend DNA and rearrange nucleosomes, therefore possibly opening up the DNA to allow other transcription factors access to bind. The NF-Y protein recruits a small amount of Sp1 and/or Sp3 to the region which in turn recruits the uncharacterised protein. These can then bind strongly to the region if the correct sequences are present. Unfortunately, the role of the uncharacterised protein in transcription cannot be inferred due to unforeseen problems with the functional assays. Results from these assays should have given a much clearer understanding of the molecular interactions occurring at the proximal promoter region, and possibly provided insights into the mechanisms causing the down-regulation of transcription seen in drug resistant cells.

References

Adachi, N., Miyaike, M., Kato, S., Kanamaru, R., Koyama, H. and Kikuchi, A. (1997) Cellular distribution of mammalian DNA topoisomerase II is determined by its catalytically dispensable C-terminal domain. *Nucleic Acids Research* **25**, 3135-3142.

Adachi, N., Nomoto, M., Kohno, K. and Koyama, H. (2000) Cell-cycle regulation of the DNA topoisomerase II alpha promoter is mediated by proximal CCAAT boxes: possible involvement of acetylation. *Gene* **245**, 49-57.

Allen, K. A., Williams, A. O., Isaacs, R. J. and Stowell, K. M. (2004) Down-regulation of human topoisomerase II alpha correlates with altered expression of transcriptional regulators NF-YA and Sp1. *Anti-Cancer Drugs* **15**, 357-362.

Arima, Y., Hirota, T., Bronner, C., Mousli, M., Fujiwara, T., Niwa, S., Ishikawa, H. and Saya, H. (2004) Down-regulation of nuclear protein ICBP90 by p53/p21(Cip1/WAF1)-dependent DNA-damage checkpoint signals contributes to cell cycle arrest at G1/S transition. *Genes to Cells* **9**, 131-142.

Asano, T., An, T., Zwelling, L. A., Takano, H., Fojo, A. T. and Kleinerman, E. S. (1996) Transfection of a human topoisomerase II alpha gene into etoposide-resistant human breast tumor cells sensitizes the cells to etoposide. *Oncology Research* **8**, 101-110.

Baird, C. L., Harkins, T. T., Morris, S. K. and Lindsley, J. E. (1999) Topoisomerase II drives DNA transport by hydrolyzing one ATP. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13685-13690.

Bellorini, M., Lee, D. K., Dantonel, J. C., Zemzoumi, K., Tora, L. and Mantovani, R. (1997) CCAAT binding NF-Y-TBP interactions: NF-YB and NF-YC require short domains adjacent to their histone fold motifs for association with TBP basic

resisdues. *Nucleic Acids Research* **25**.

Berger, J. M., Gamblin, S. J., Harrison, S. C. and Wang, J. C. (1996) Structure and mechanism of DNA topoisomerase II. *Nature* **379**, 225-232.

Bi, W., Wu, L., Coustry, F., Crombrugghe, B. and Maity, S. N. (1997) DNA binding specificity of the CCAAT-binding factor CBF/NF-Y. *The Journal of Biological Chemistry* **272**, 26562-26572.

Birnbaum, M. J., van Wijnen, A. J., Odgren, P. R., Last, T. J., Suske, G., Stein, G. S. and Stein, J. L. (1995) Sp1 trans-activation of cell cycle regulated promoters is selectively repressed by Sp3. *Biochemistry* **34**, 16503-16508.

Brandt, T. L. and Kroll, D. J. (1997) NF-M trans-activates the human DNA topoisomerase II alpha promoter independently of c-Myb in HL-60 cells. *Leukemia Research* **21**, 711-720.

Burden, D. A., Kingma, P., Froelich-Ammon, S. J., Bjornsti, M., Patchanm, M., Thompson, R. B. and Osheroff, N. (1996) Topoisomerase II*etoposide interactions direct the formation of drug-induced enzyme-DNA cleavage complexes. *Journal of Biological Chemistry* **271**, 29238-29244.

Carpenter, A. J. and Porter, A. C. G. (2004) Construction, characterization, and complementation of a conditional-lethal DNA topoisomerase II alpha mutant human cell line. *Molecular Biology of the Cell* **15**, 5700-5711.

Chen, G., Jaffrezou, J. P., Fleming, W. H., Duran, G. E. and Sikic, B. I. (1994) Prevalence of Multidrug-Resistance Related to Activation of the Mdr1 Gene in Human Sarcoma Mutants Derived by Single-Step Doxorubicin Selection. *Cancer Research* **54**, 4980-4987.

Chung, C. T., Niemela, S. L. and Miller, R. H. (1989a) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 2172-2175.

Chung, T. D. Y., Drake, F. H., Tan, K. B., Per, S. R. and Crooke, S. T. (1989b) Characterization and immunological identification of cDNA clones encoding two human DNA topoisomerase II isozymes. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 9431-9435.

Collins, K., Jacks, T. and Pavletich, N. P. (1997) The cell cycle and cancer. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 2776-2778.

Cornarotti, M., Capranico, G., Bohm, S., Oriana, S., Spatti, G. B., Mariani, L., Ballabio, G. and Zunino, F. (1996) Gene expression of DNA topoisomerases I, II alpha and II beta and response to cisplatin-based chemotherapy in advanced ovarian carcinoma. *International Journal of Cancer* **67**, 479-484.

Coustry, F., Hu, Q. H., de Crombrugghe, B. and Maity, S. N. (2001) CBF/NF-Y functions both in nucleosomal disruption and transcription activation of the chromatin-assembled topoisomerase II alpha promoter - Transcription activation by CBF/NF-Y in chromatin is dependent on the promoter structure. *Journal of Biological Chemistry* **276**, 40621-40630.

Coustry, F., Sinha, S., Maity, S. N. and Crombrugghe, B. (1998) The two activation domains of the CCAAT-binding factor CBF interact with the dTAF$_{II}$110 component of the *Drosophila* TFIID complex. *Biochemistry Journal* **331**, 291-297.

Cramer, P., Bushnell, D. A., Fu, J. H., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R. and Kornberg, R. D. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640-649.

Davies, S. L., Jenkins, J. R. and Hickson, I. D. (1993) Human cells express two differentially spliced forms of topoisomerase IIB mRNA. *Nucleic Acids Research* **21**, 3719-3723.

Du, W. L., Hattori, Y., Hashiguchi, A., Kondoh, K., Hozumi, N., Ikeda, Y., Sakamoto, M., Hata, J. and Yamada, T. (2004) Tumor angiogenesis in the bone marrow of multiple myeloma patients and its alteration by thalidomide treatment. *Pathology International* **54**, 285-294.

Dynan, W. S. and Tjian, R. (1983) The Promoter-Specific Transcription Factor-Sp1 Binds to Upstream Sequences in the Sv40 Early Promoter. *Cell* **35**, 79-87.

Fang, X., Han, H., Stamatoyannopoulos, G. and Li, Q. (2004) Developmentally specific role of the CCAAT box in regulation of human g-globin gene expression. *Journal of Biological Chemistry* **279**, 5444-5449.

Farmer, G., Bargonetti, J., Zhu, H., Friedman, P., Prywes, R. and Prives, C. (1992) Wild-Type P53 Activates Transcription Invitro. *Nature* **358**, 83-86.

Folkman, J. (1996) Fighting cancer by attacking its blood supply. *Sci. Am.* **272**, 275-278.

Furukawa, M., Uchiumi, T., Nomoto, M., Takano, H., Morimoto, R. I., Naito, S., Kuwano, M. and Kohno, K. (1998) The role of an inverted CCAAT element in transcriptional activation of the human DNA topoisomerase II alpha gene by heat shock. *Journal of Biological Chemistry* **273**, 10550-10555.

Ganapathi, R., Constantinou, A., Kamath, N., Dubyak, G., Grabowski, D. and Krivacic, K. (1996) Resistance to etoposide in human leukemia HL-60 cells: reduction in drug-induced DNA cleavage associated with hypophosphorylation of topoisomerase II phosphopeptides. *Mol Pharmacol* **50**, 243-248.

Ge, Y., Matherly, L. and Taub, J. (2001) Transcriptional regulation of cell-specific expression of the human cystathionine-B-synthase by differential binding of Sp1/Sp3 to the -1B promoter. *Journal of Biological Chemistry* **276**, 43570-43579.

Giaccone, G., vanArkOtte, J., Scagliotti, G., Capranico, G., vanderValk, P., Rubio, G., Dalesio, O., Lopez, R., Zunino, F., Walboomers, J. and Pinedo, H. M. (1995) Differential expression of DNA topoisomerases in non-small cell lung cancer and normal lung. *Biochimica Et Biophysica Acta-Gene Structure and Expression*

1264, 337-346.

Goodrich, J. A., Cutler, G. and Tjian, R. (1996) Contacts in context: Promoter specificity and macromolecular interactions in transcription. *Cell* **84**, 825-830.

Goodwin, A., Wang, S. W., Toda, T., Norbury, C. and Hickson, I. D. (1999) Topoisomerase III is essential for accurate nuclear division in Schizosaccharomyces pombe. *Nucleic Acids Research* **27**, 4050-4058.

Gu, C. D., Oyama, T., Osaki, T., Kohno, K. and Yasumoto, K. (2001) Expression of Y box-binding protein-1 correlates with DNA topoisomerase II alpha and proliferating cell nuclear antigen expression in lung cancer. *Anticancer Research* **21**, 2357-2362.

Hagen, G., Muller, S., Beato, M. and Suske, G. (1994) Sp1-mediated transcriptional activation is repressed by Sp3. *Embo Journal* **13**, 3843-3851.

Hapgood, J. P., Riedemann, J. and Scherer, S. D. (2001) Regulation of gene expression by GC-rich DNA *cis*-elements. *Cell Biology International* **25**, 17-31.

Hazlehurst, L. A., Foley, N. E., Gleason-Guzman, M. C., Hacker, M. P., Cress, A. E., Greenberger, L. W., De Jong, M. C. and Dalton, W. S. (1999) Multiple mechanisms confer drug resistance to mitoxantrone in the human 8226 myeloma cell line. *Cancer Research* **59**, 1021-1028.

Hazlehurst, L. A., Valkov, N., Wisner, L., Storey, J. A., Boulware, D., Sullivan, D. M. and Dalton, W. S. (2001) Reduction in drug-induced DNA double-strand breaks associated with beta 1 integrin-mediated adhesion correlates with drug resistance in U937 cells. *Blood* **98**, 1897-1903.

Herzog, C. and Zwelling, L. A. (1997) Evaluation of a potential regulatory role for inverted CCAAT boxes in the human topoisomerase IIa promoter. *Biochemical and Biophysical Research Communications* **232**, 608-612.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989) Site-Directed

Mutagenesis by Overlap Extension Using the Polymerase Chain-Reaction. *Gene* **77**, 51-59.

Hochhauser, D., Stanway, C. A., Harris, A. L. and Hickson, I. D. (1992) Cloning and Characterization of the 5'-Flanking Region of the Human Topoisomerase-Ii-Alpha Gene. *Journal of Biological Chemistry* **267**, 18961-18965.

Holmes, D. S. and Quigley, M. (1981) A Rapid Boiling Method for the Preparation of Bacterial Plasmids. *Analytical Biochemistry* **114**, 193-197.

Hopfner, R., Mousli, M., Jeltsch, J. M., Voulgaris, A., Lutz, Y., Marin, C., Bellocq, J. P., Oudet, P. and Bronner, C. (2000) ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase II alpha expression. *Cancer Research* **60**, 121-128.

Hopfner, R., Mousli, M., Oudet, P. and Bronner, C. (2002) Overexpression of ICBP90, a novel CCAAT-binding protein, overcomes cell contact inhibition by forcing topoisomerase II alpha expression. *Anticancer Research* **22**, 3165-3170.

Hu, Q. H., Bhattacharya, C. and Maity, S. N. (2002) CCAAT binding factor (CBF) binding mediates cell cycle activation of topoisomerase II alpha. *The Journal of Biological Chemistry* **277**, 37191-37200.

Hu, Q. H. and Maity, S. N. (2000) Stable expression of a dominant negative mutant of CCAAT binding factor/NF-Y in a mouse fibroblast cell resulting in retardation of cell growth and inhibition of transcription of various cellular genes. *The Journal of Biological Chemistry* **275**, 4434-4444.

Innocente, S. A. and Lee, J. M. (2005) P53 is a NF-Y- and p21-independent, Sp1-dependent repressor of cyclin B1 transcription. *Febs Letters* **579**, 1001-1007.

Inoue, T., Kamiyama, J. and Sakai, T. (1999) Sp1 and NF-Y synergistically mediate the effect of vitamin D(3) in the p27(Kip1) gene promoter that lacks vitamin D response elements. *Journal of Biological Chemistry* **274**, 32309-32317.

Isaacs, R. J. (1996) Transcriptional regulation of human topoisomerase IIalpha. University of Oxford, UK.

Isaacs, R. J., Davies, S. L., Sandri, M. I., Redwood, C., Wells, N. J. and Hickson, I. D. (1998) Physiological regulation of eukaryotic topoisomerase II. *Biochimica Et Biophysica Acta-Gene Structure and Expression* **1400**, 121-137.

Isaacs, R. J., Davies, S. L., Wells, N. J. and Harris, A. L. (1995) Topoisomerase-Ii-Alpha and Topoisomerase-Beta as Therapy Targets in Breast-Cancer. *Anti-Cancer Drugs* **6**, 195-211.

Isaacs, R. J., Harris, A. L. and Hickson, I. D. (1996) Regulation of the human topoisomerase II alpha gene promoter in confluence-arrested cells. *Journal of Biological Chemistry* **271**, 16741-16747.

Jean, D., Guillaume, N. and Frade, R. (2002) Characterization of human cathepsin L promoter and identification of binding sites for NF-Y, Sp1 and Sp3 that are essential for its activity. *Biochemical Journal* **361**, 173-184.

Jensen, S., Andersen, A., Kjeldsen, E., Biersack, H., Olsen, E., Andersen, T. and Westergaard, O. (1996) Analysis of functional domain organisation in DNA topoisomerase II from humans and Saccharomyces cerevisiae. *Molecular and Cellular Biology* **16**, 3866-3877.

Joshi, A. A., Wu, Z., Reed, R. F. and Suttle, D. P. (2003) Nuclear factor-Y binding to the topoisomerase II alpha promoter is inhibited by both the p53 tumor suppressor and anticancer drugs. *Molecular Pharmacology* **63**, 359-367.

Kadonaga, J. T., Carner, K. R., Masiarz, F. R. and Tijian, R. (1987) Isolation of Cdna-Encoding Transcription Factor Sp1 and Functional-Analysis of the DNA-Binding Domain. *Cell* **51**, 1079-1090.

Khélifa, T., René, B., Le Mée, S., Lambert, B., Saucier, J.-M., Markovits, J., Jacquemin-Sablon, H. and Jacquemin-Sablon, A. (1999) Transfection of 9-Hydroxyellipticine-resistant Chinese Hamster fibroblasts with Human

Topoisomerase IIa cDNA: Selective restoration of the sensitivity to DNA religation inhibitors. *Cancer Research*, 4927-4936.

Kim, J. C., Yoon, J. B., Koo, H. S. and Chung, I. K. (1998) Cloning and characterization of the 5 '-flanking region for the human topoisomerase III gene. *Journal of Biological Chemistry* **273**, 26130-26137.

Kim, J. L., Nikolov, D. B. and Burley, S. K. (1993) Crystallographic Studies of Complexes of the Tfiid Tata-Box Binding-Protein with DNA. *Journal of Cellular Biochemistry*, 64-64.

Kim, W., Kakehi, Y., Kinoshita, H., Arao, S., Fukumoto, M. and Yoshida, O. (1996) Expression patterns of multi-drug-resistance (MDR1), Multidrug resistance-associated protein (MRP), Glutathione-S-Transferase-phi (GST-phi) and DNA topoisomerase II (topo II) genes in renal carcinomas and normal kidney. *The Journal of Urology* **156**, 506-511.

Kubo, T., Kohno, K., Ohga, T., Taniguchi, K., Kawanami, K., Wada, M. and Kuwano, M. (1995) DNA Topoisomerase Ii-Alpha Gene-Expression under Transcriptional Control in Etoposide Teniposide-Resistant Human Cancer-Cells. *Cancer Research* **55**, 3860-3864.

Laing, S. G. and Maity, S. N. (1998) Pathway of complex formation between DNA and three subunits of CBF/NF-Y. *The Journal of Biological Chemistry* **273**, 31590-31598.

Lang, A. J., Mirski, S. E. L., Cummings, H. J., Yu, Q., Gerlach, J. H. and Cole, S. P. C. (1998) Structural organisation of the human TOP2A and TOP2B genes. *Gene* **221**, 255-266.

Li, Q., Herrler, M., Landsberger, N., Kaludov, N., Ogryzko, V., Nakatani, Y. and Wolffe, A. (1998) *Xenopus* NF-Y pre-sets chromatin to potentiate p300 and acetylation-responsive transcription from the *Xenopus* hsp70 promoter in vivo. *Embo Journal* **17**, 6300-6315.

Lok, C. N., Lang, A. J., Mirski, S. E. L. and Cole, S. P. C. (2002) Characterization of the human topoisomerase II beta (TOP2B) promoter activity: essential roles of the nuclear factor-Y (NF-Y)- and specificity protein-1 (Sp1)-binding sites. *Biochemical Journal* **368**, 741-751.

Magan, N. (2002) Basal transcription of human topoisomerase II. In *Institute of Molecular BioSciences*. Massey University, Palmerston North.

Magan, N. (2003) Basal transcription of human topoisomerase II. In *Institute of Molecular BioSciences*. Massey University, Palmerston North.

Magan, N., Szremska, A. P., Isaacs, R. J. and Stowell, K. M. (2003) Modulation of DNA topoisomerase II alpha promoter activity by members of the Sp (specificity protein) and NF-Y (nuclear factor Y) families of transcription factors. *Biochemical Journal* **374**, 723-729.

Maity, S. N. and de Crombrugghe, B. (1998) Role of CCAAT-binding protein CBF/NF-Y in transcription. *Trends Biochem Sci* **23**, 174-178.

Mantovani, R. (1998) A survey of 178 NF-Y binding CCAAT boxes. *Nucleic Acids Research* **26**, 1135-1143.

Matsui, T., Katsuno, Y., Inoue, T., Fujita, F., Joh, T., Niida, H., Murakami, H., Itoh, M. and Nakanishi, M. (2004) Negative regulation of Chk2 expression by p53 is dependent on the CCAAT-binding transcription factor NF-Y. *Journal of Biological Chemistry* **279**, 25093-25100.

Matsumoto, Y., Takano, H. and Fojo, T. (1997) Celleular adaptation to drug exposure: Evolution of the drug-resistant phenotype. *Cancer Research* **57**, 5086-5092.

M^{c}Lenachan, S. (1998) Regulation of the human topoisomerase IIalpha promoter. In *Institute of Molecular BioSciences*. Massey University, Palmerston North.

Mercer, W. E., Shields, M. T., Lin, D., Appella, E. and Ullrich, S. J. (1991) Growth Suppression Induced by Wild-Type P53-Protein Is Accompanied by Selective

Down-Regulation of Proliferating-Cell Nuclear Antigen Expression. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 1958-1962.

Meyer, K. N., Kjeldsen, E., Straub, T., Knudsen, B. R., Hickson, I. D., Kikuchi, A., Kreipe, H. and Boege, F. (1997) Cell cycle-coupled relocation of types I and II topoisomerases and modulation of catalytic enzyme activities. *Journal of Cell Biology* **136**, 775-788.

Mirski, S. E. L., Bielawski, J. C. and Cole, S. P. C. (2003) Identification of functional nuclear export sequences in human topoisomerase II alpha and beta. *Biochemical and Biophysical Research Communications* **306**, 905-911.

Mirski, S. E. L. and Cole, S. P. C. (1995) Cytoplasmic Localisation of a Mutant M(R)160,000 Topoisomerase-Ii-Alpha Is Associated with the Loss of Putative Bipartite Nuclear-Localisation Signals in a Drug-Resistant Human Lung-Cancer Cell-Line. *Cancer Research* **55**, 2129-2134.

Ng, S. W., Liu, Y., Hasselblatt, K. T., Mok, S. C. and Berkowitz, R. S. (1999) A new human topoisomerase III that interacts with SGS1 protein. *Nucleic Acids Research* **27**, 993-1000.

Nicolas, M., Noe, V., Jensen, K. B. and Ciudad, C. J. (2001) Cloning and characterization of the 5'-flanking region of the human transcription fator *Sp1* gene. *The Journal of Biological Chemistry* **276**, 22126-22132.

Oda, Y., Ohishi, Y., Saito, T., Hinoshita, E., Uchiumi, T., Kinukawa, N., Iwamoto, Y., Kohno, K., Kuwano, M. and Tsuneyoshi, M. (2003) Nuclear expression of Y-box-binding protein-1 correlates with P-glycoprotein and topoisomerase II alpha expression, and with poor prognosis in synovial sarcoma. *Journal of Pathology* **199**, 251-258.

Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J. and Hol, W. G. J. (1998) Crystal structures of topoisomerase I in covalent and noncovalent complexes with DNA.

Science **279**, 1504-1513.

Renodon-Corniere, A., Jensen, L. H., Nitiss, J. L., Jensen, P. B. and Sehested, M. (2003) Analysis of bisdioxopiperazine Dexrazoxane binding to human DNA topoisomerase II alpha: Decreased binding as a mechanism of drug resistance. *Biochemistry* **42**, 9749-9754.

Robert, J. and Larsen, A. K. (1998) Drug resistance to topoisomerase II inhibitors. *Biochimie* **80**, 247-254.

Roder, K., Wolf, S. S., Larkin, K. J. and Schweizer, M. (1999) Interaction between the two ubiquitously expressed transcription factors NF-Y and Sp1. *Gene* **234**, 61-69.

Ronchi, A., Bellorini, M., Mongelli, N. and Mantovani, R. (1995) CCAAT-box binding protein NF-Y (CBP, CP1) recognizes the minor groove and distorts DNA. *Nucleic Acids Research* **23**, 4565-4572.

Rorer, R.-P. (1999) Topoisomerase II poison action. In *Oncology Slide Library*. Rhone-Poulenc Rorer Pharmaceuticals.

Saleem, A., Ibrahim, N., Patel, M., Li, X. G., Gupta, E., Mendoza, J., Pantazis, P. and Rubin, E. H. (1997) Mechanisms of resistance in a human cell line exposed to sequential topoisomerase poisoning. *Cancer Research* **57**, 5100-5106.

Salsi, V., Caretti, G., Wasner, M., Reinhard, W., Haugwitz, U., Engeland, K. and Mantovani, R. (2003) Interactions between p300 and multiple NF-Y trimers govern cyclin B2 promoter function. *Journal of Biological Chemistry* **278**, 6642-6650.

Sambrook, J. and Russell, I. (2001) *Molecular cloning, 3rd edition.* Cold Spring Harbour Laboratory Press, New York.

Sandri, M. I., Isaacs, R. J., Ongkeko, W. M., Harris, A. L., Hickson, I. D., Broggini, M. and Vikhanskaya, F. (1996) p53 regulates the minimal promoter of the human topoisomerase II [alpha] gene. *Nucleic Acids Research* **24**, 4464-4470.

Satherley, K., de Souza, L., Neale, M. H., Alexander, R. A., Myatt, N., Foss, A. J. E., Hungerford, J. L., Hickson, I. D. and Cree, I. A. (2000) Relationship between expression of topoisomerase II isoforms and chemosensitivity in choroidal melanoma. *Journal of Pathology* **192**, 174-181.

Shibao, K., Takano, H., Nakayama, Y., Okazaki, K., Nagata, N., Izumi, H., Uchiumi, T., Kuwano, M., Kohno, K. and Itoh, H. (1999) Enhanced coexpression of YB-1 and DNA topoisomerase II alpha genes in human colorectal carcinomas. *International Journal of Cancer* **83**, 732-737.

Sjostrom, E., Andersen, C. and Johansen, T. (1997) Structural and functional analyses of DNA bending induced by Sp1 family transcription factors. *Journal of Molecular Biology* **267**, 490-504.

Stockinger, E. J., Mao, Y., Regier, M. K., Triezenberg, S. J. and Thomashow, M. F. (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. *Nucleic Acids Research* **29**, 1524-1533.

Suda, N., Ito, Y., Imai, T., Kikumori, T., Kikuchi, A., Nishiyama, Y., Yoshida, S. and Suzuki, M. (2004) The alpha 4 residues of human DNA topoisomerase II alpha function in enzymatic activity and anticancer drug sensitivity. *Nucleic Acids Research* **32**, 1767-1773.

Szremska, A. P. (2000) Sp1, Sp3, NF-Y and the topoisomerase IIalpha promoter. In *Institute of Molecular BioSciences*. Massey University, Palmerston North.

Takano, H., Ise, T., Nomoto, M., Kato, K., Murakami, T., Ohmori, H., Imamura, T., Nagatani, G., Okamoto, T., Ohta, R., Furukawa, M., Shibao, K., Izumi, H., Kuwano, M. and Kohno, K. (1999) Structural and functional analysis of the control region of the human DNA topoisomerase II alpha gene in drug-resistant cells. *Anti-Cancer Drug Design* **14**, 87-92.

Takano, H., Kohno, K., Ono, M., Uchida, Y. and Kuwano, M. (1991) Increased

phosphorylation of DNA topoisomerase II in etoposide-resistant mutants of human cancer KB cells. *Cancer Research* **51**, 3951-3957.

Tolner, B., Hartley, J. and Hochhauser, D. (2001) Transcriptional regulation of topoisomerase IIa at confluence and pharmacological modulation of expression by *bis*-benzimidazole drugs. *Mol Pharmacol* **59**, 699-706.

Trotzier, M. A., Bronner, C., Bathami, K., Mathieu, E., Abbady, A. Q., Jeanblanc, M., Muller, C. D., Rochette-Egly, C. and Mousli, M. (2004) Phosphorylation of ICBP90 by protein kinase A enhances topoisomerase II alpha expression. *Biochemical and Biophysical Research Communications* **319**, 590-595.

Turner, J. G., Engel, R., Derderian, J. A., Jove, R. and Sullivan, D. M. (2004) Human topoisomerase IIa nuclear export is mediated by two CRM-1-dependent nuclear export signals. *Journal of Cell Science* **117**, 3061-3071.

Ullrich, S. J., Anderson, C. W., Mercer, W. E. and Appella, E. (1992) The P53 Tumor Suppressor Protein, a Modulator of Cell-Proliferation. *Journal of Biological Chemistry* **267**, 15259-15262.

Valkov, N. I., Gump, J. L., Engel, R. and Sullivan, D. M. (2000) Cell density-dependent VP-16 sensitivity of leukaemic cells is accompanied by the translocation of topoisomerase II alpha from the nucleus to the cytoplasm. *British Journal of Haematology* **108**, 331-345.

Verrijzer, C. P., Chen, J. L., Yokomori, K. and Tjian, R. (1995) Binding of Tafs to Core Elements Directs Promoter Selectivity by Rna-Polymerase-Ii. *Cell* **81**, 1115-1125.

Walker, P. R., Smith, C., Youdale, T., Leblanc, J., Whitfield, J. F. and Sikorska, M. (1991) Topoisomerase-Ii-Reactive Chemotherapeutic Drugs Induce Apoptosis in Thymocytes. *Cancer Research* **51**, 1078-1085.

Wang, H., Jiang, Z. G., Wong, Y. W., Dalton, W. S., Futscher, B. W. and Chan, V. T. W. (1997a) Decreased CP-1 (NF-Y) activity results in transcriptional down-regulation of topoisomerase II alpha in a doxorubicin-resistant variant of human

multiple myeloma RPMI 8226. *Biochemical and Biophysical Research Communications* **237**, 217-224.

Wang, J. (1996) DNA Topoisomerases. *Ann. Rev. Biochem.* **66**, 635-692.

Wang, Q. J., Zambetti, G. P. and Suttle, D. P. (1997b) Inhibition of DNA topoisomerase II alpha gene expression by the p53 tumor suppressor. *Molecular and Cellular Biology* **17**, 389-397.

Wells, N. J., Fry, A. M., Guano, F., Norbury, C. and Hickson, I. D. (1995) Cell cycle phase-specific phosphorylation of human topoisomerase II_. *Journal of Biological Chemistry* **269**, 28357-28363.

Xiao, H., Hasegawa, A. and Isobe, K. (2000) p300 collaborates with Sp1 and Sp3 in $p21^{(waf/cip1)}$ promoter activation induced by histone deacetylase inhibitor. *Journal of Biological Chemistry* **275**, 1371-1376.

Yoon, J. H., Kim, J. K., Rha, G. B., Oh, M., Park, S. H., Seong, R. H., Hong, S. H. and Park, S. D. (1999) Sp1 mediates cell proliferation-dependent regulation of rat DNA topoisomerase II alpha gene promoter. *Biochemical Journal* **344**, 367-374.

Yun, J., Chae, H. D., Choy, H. E., Chung, J., Yoo, H. S., Han, M. H. and Shin, D. Y. (1999) p53 negatively regulates cdc2 transcription via the CCAAT-binding NF-Y transcription factor. *Journal of Biological Chemistry* **274**, 29677-29682.

Appendices

Appendix 1: Oligonucleotide Sequences.

1.1 Topoisomerase IIα Oligonucleotide Sequences used in EMSAs.

Strand	Probe Name	Sequence
Forward 5'-3'	WT	CGAGTCAGGGATTGGCTGGTCTGCTTCGGGCGGGCTAAAGGAAG
	MT	CGAGTCAGGGATTGGCTGGCCTACTTCGGGCGGGCTAAAGGAAG
	Del	CGAGTCAGGGATTGGCTGTTCGGGCGGGCTAAAGGAAG
	Ins	CGAGTCAGGGATTGGCTGGTCGTATGCTGCTTCGGGCGGGCTAAAGGAAG
	Cent	CTGGTCTGCTTC
	ICB1½GC1	CGAGTCAGGGATTGGCTGGTCTGCTTCGGGC
	ICB1	CGAGTCAGGGATTGGCTGGTCTGCTTC
	GC1½ICB1	TGGCTGGTCTGCTTCGGGCGGGCTAAAGGAAG
	GC1	CTGGTCTGCTTCGGGCGGGCTAAAGGAAG
Reverse 5'-3'	WT	CTTCCTTTAGCCCGCCCGAAGCAGACCAGCCAATCCCTGACTCG
	MT	CTTCCTTTAGCCCGCCCGAAGTAGGCCAGCCAATCCCTGACTCG
	Del	CTTCCTTTAGCCCGCCCGAACAGCCAATCCCTGACTCG
	Ins	CTTCCTTTAGCCCGCCCGAAGCAGCATACGACCAGCCAATCCCTGACTCG
	Cent	GAAGCAGACCAG
	ICB1½GC1	GCCCGAAGCAGACCAGCCAATCCCTGACTCG
	ICB1	GAAGCAGACCAGCCAATCCCTGACTCG
	GC1½ICB1	CTTCCTTTAGCCCGCCCGAAGCAGACCAGCCA
	GC1	CTTCCTTTAGCCCGCCCGAAGCAGACCAG

1.2 Topoisomerase IIα Competitor Oligonucleotide Sequences used in EMSAs.

Single element competitors:

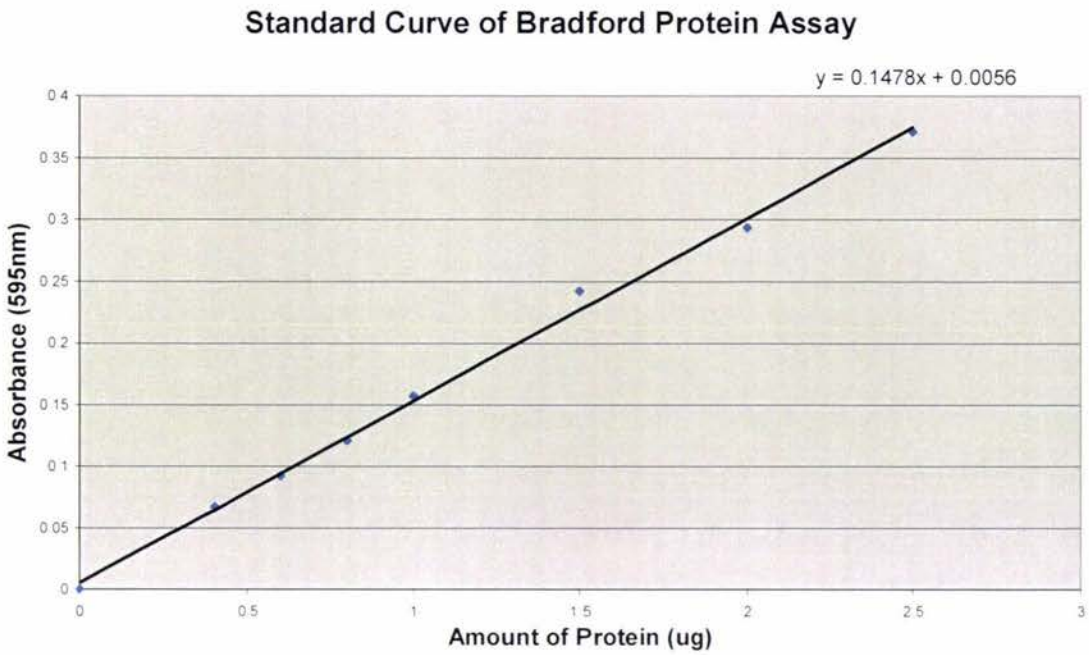
Strand	Probe Name	Sequence
Forward 5'-3'	wtICB1	GAGTCAGGGATTGGCTGGTCTGC
	mtICB1	GAGTCAGGGATTCCCTGGTCTGC
	wtGC1	CTGCTTCGGGCGGGCTAAAG
	mtGC1	CTGCTTCGTGCGTGCTAAAG
Reverse 5'-3'	wtICB1	GCAGACCAGCCAATCCCTGACTC
	mtICB1	GCAGACCAGGGAATCCCTGACTC
	wtGC1	CTTTAGCCCGCCCGAAGCAG
	mtGC1	CTTTAGCACGCACGAAGCAG

Double element competitors:

Strand	Probe Name	Sequence
Forward 5'-3'	wtICB1/wtGC1	CGAGTCAGGGATTGGCTGGTCTGCTTCGGGCGGGCTAAAGGAAG
	mtICB1/wtGC1	CGAGTCAGGGATTCCCTGGTCTGCTTCGGGCGGGCTAAAGGAAG
	wtICB1/mtGC1	CGAGTCAGGGATTGGCTGGTCTGCTTCGTGCGTGCTAAAGGAAG
	mtICB1/mtGC1	CGAGTCAGGGATTCCCTGGTCTGCTTCGTGCGTGCTAAAGGAAG
Reverse 5'-3'	wtICB1/wtGC1	CTTCCTTTAGCCCGCCCGAAGCAGACCAGCCAATCCCTGACTCG
	mtICB1/wtGC1	CTTCCTTTAGCCCGCCCGAAGCAGACCAGGGAATCCCTGACTCG
	wtICB1/mtGC1	CTTCCTTTAGCACGCACGAAGCAGACCAGCCAATCCCTGACTCG
	mtICB1/mtGC1	CTTCCTTTAGCACGCACGAAGCAGACCAGGGAATCCCTGACTCG

Appendix 2: Protein Assays.

2.1 HeLa Extract Protein Quantification Assays.

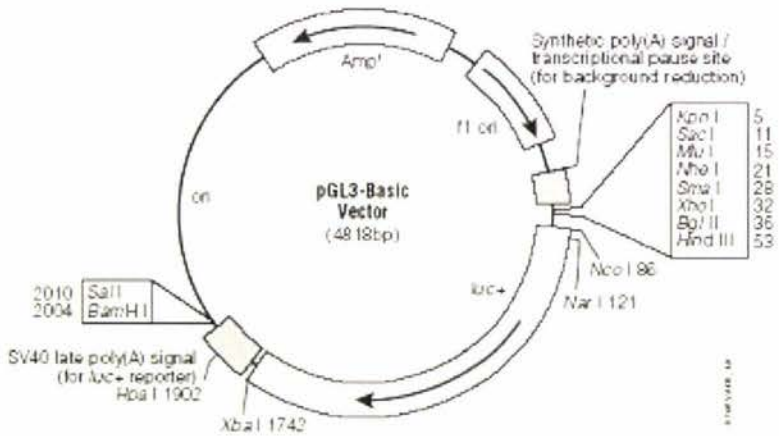


BSA Protein Standard (µg)	Absorbance 595 nm (Averages of Triplicates)
0	0.248
0.4	0.315
0.6	0.341
0.8	0.369
1.0	0.405
1.5	0.490
2.0	0.542
2.5	0.619

First HeLa Protein Extract (Diluted 1:75, 10 µL)		Second HeLa Protein Extract (Diluted 1:5, 10 µL)	
Absorbance 595 nm	Amount of Protein µg (off graph)	Absorbance 595 nm	Amount of Protein µg (off graph)
0.350	0.65	1.118	5.84
0.369	0.78	1.119	5.85
0.365	0.75	1.107	5.76
Average	0.73	**Average**	5.82
Original Conc	**5.47 µg/µL**	**Original Conc**	**3.67 µg/µL**

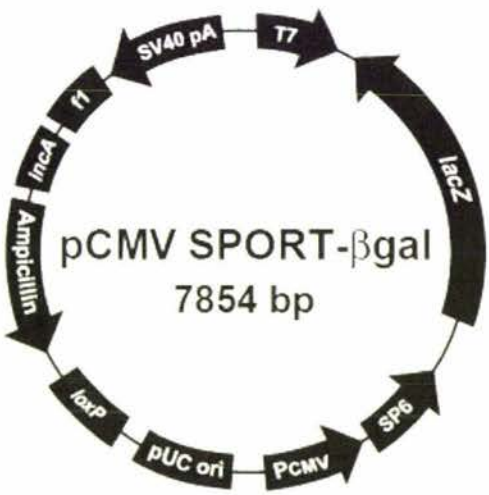
Appendix 3: Vector Maps.

3.1 pGL3-Basic Vector



pGL3 Basic vector (Promega) used for topo IIα promoter functional assays.

3.2 pCMV Sport-βgal Vector



pCMV Sport-βgal (Invitrogen) used as the internal control for transient transfection assays.

Appendix 4: Primer Sequences.

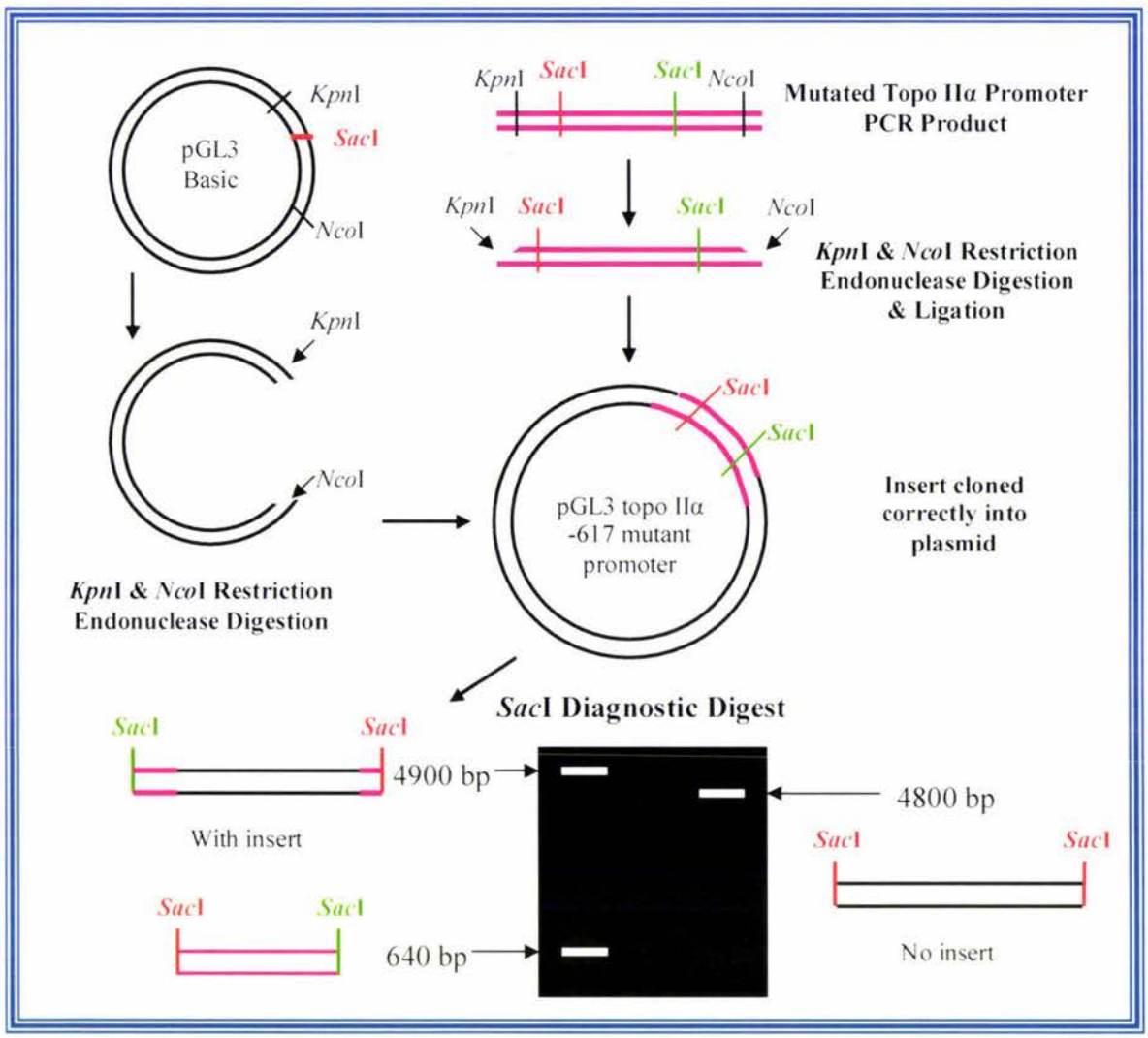
PCR mutagenesis internal primer sequences:

Strand	Probe Name	Sequence
Forward 5'-3'	MT	AGGGATTGGCTGGCCTACTTCGGGCGGGCT
	Del	GTCAGGGATTGGCTGTTCGGGCGGGCTAAA
	Ins	AGGGATTGGCTGGTCGTATGCTGCTTCGGGCGGGC
Reverse 5'-3'	MT	AGCCCGCCCGAAGTAGGCCAGCCAATCCCT
	Del	TTTAGCCCGCCCGAACAGCCAATCCCTGAC
	Ins	GCCCGCCCGAAGCAGCATACGACCAGCCAATCCCT

External primer sequences:

Strand	Probe Name	Sequence
Forward	RV primer 3	CTAGCAAAATAGGCTGTCCC
Reverse	GL primer 2	GAAATACAAAAACCGCAGAAGG

Appendix 5: *Sac*I Diagnostic Digests.



This schematic diagram shows how the *Sac*I diagnostic digests work to identify plasmids which contain an insert correctly cloned into the pGL3 Basic plasmid. The plasmid contains a *Sac*I (red) site which was removed upon *Kpn*I and *Nco*I digestion, but the same site from another plasmid was replaced again within the insert (red), however the insert also contains a *Sac*I (green) site within the topo IIα promoter sequence. Therefore *Sac*I digestion gives a single 4800 bp band if there is no insert, and a 4900 bp and a 640 bp band if there is a insert correctly cloned into the plasmid. Correct orientation of the insert was not an issue as different enzymes were used at each end. Note: this is not to scale.

Appendix 6: Alignment of Sequences.

CLUSTAL W (1.82) multiple sequence alignment (5'-3').
Blue = wildtype sequence, Red = mutations/deletions.
Oligonucleotides used to produce GC2 Mutant constructs:

	Sequence
GC2mt forward	AGGGTAAAGGTG-CGTGGTTGAGGCA (mutated + 1 base deletion)
GC2mt reverse	TGCCTCAACCACG-CACCTTTACCCT (mutated + 1 base deletion)

```
MT       AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGGGGC 60
Del      AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGGGGC 60
WT       AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGGGGC 60
Ins      AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGGGGC 60
GC2-MT   AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGTG-C 59
GC2-Del  AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGTG-C 59
GC2-Ins  AGCCACCGCACACAGCCTACTTTTATTTCTTTGAAAAATGAATTCGAGGGTAAAGGTG-C 59
         ************************************************************ * *

MT       GGGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA120
Del      GGGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA120
WT       GGGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA120
Ins      GGGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA120
GC2-MT   GTGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA119
GC2-Del  GTGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA119
GC2-Ins  GTGGTTGAGGCAGATGCCAGAATCTGTTCGCTTCAACCAAGCAGCCAGGCTGCCTGTCCA119
         * ************************************************************

MT       GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA180
Del      GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA180
WT       GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA180
Ins      GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA180
GC2-MT   GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA179
GC2-Del  GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA179
GC2-Ins  GAAAGCCGGCACTCAGTTTCCTCAGGAAAACGAAGCTAAGGCTCCCATTCCCCTCGCTAA179
         ************************************************************

MT       CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC240
Del      CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC240
WT       CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC240
Ins      CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC240
GC2-MT   CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC239
GC2-Del  CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC239
GC2-Ins  CAACGTCAGAACAGAGGACAGTTTTTAGATTTCAGGGATCTTAAATAGATTGGCAGTTCC239
         ************************************************************

MT       TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC300
Del      TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC300
WT       TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC300
Ins      TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC300
GC2-MT   TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC299
GC2-Del  TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC299
GC2-Ins  TGGAGAATAAACATCCTTTGCTTTTCTCCTGCACACTTTTGCCTCAGGCCACCCCTTCCC299
         ************************************************************
```

```
MT       GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT360
Del      GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT360
WT       GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT360
Ins      GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT360
GC2-MT   GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT359
GC2-Del  GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT359
GC2-Ins  GCTTCCAAAGCCCATCTCTTCCAAGCTTTCCGCACGAGAAAACAAGTGAGCCCTTCTCAT359
         ************************************************************

MT       TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT420
Del      TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT420
WT       TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT420
Ins      TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT420
GC2-MT   TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT419
GC2-Del  TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT419
GC2-Ins  TGGCCAGATTCCCTGTCAATCTCTCCGCTATGACGCCGAGTGGTGCCTTTTGAAGCCTCT419
         ************************************************************

MT       CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT480
Del      CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT480
WT       CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT480
Ins      CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT480
GC2-MT   CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT479
GC2-Del  CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT479
GC2-Ins  CTAGTCCCGCCTCCCTAACCTGATTGGTTTATTCAAACAAACCCCGGCCAACTCAGCCGT479
         ************************************************************

MT       TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC540
Del      TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC540
WT       TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC540
Ins      TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC540
GC2-MT   TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC539
GC2-Del  TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC539
GC2-Ins  TCATAGGTGGATATAAAAGGCAAGCTACGATTGGTTCTTCTGGACGGAGACGGTGAGAGC539
         ************************************************************

MT       GAGTCAGGGATTGGCTGGCC------TACTTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA594
Del      GAGTCAGGGATTGGCTG----------TTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA588
WT       GAGTCAGGGATTGGCTGGTC------TGCTTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA594
Ins      GAGTCAGGGATTGGCTGGTCGTATGCTGCTTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA600
GC2-MT   GAGTCAGGGATTGGCTGGCC------TACTTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA593
GC2-Del  GAGTCAGGGATTGGCTG----------TTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA587
GC2-Ins  GAGTCAGGGATTGGCTGGTCGTATGCTGCTTCGGGCGGGCTAAAGGAAGGTTCAAGTGGA599
         ****************          *********************************

                                  ┌──►~tsp
MT       GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG648
Del      GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG642
WT       GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG654
Ins      GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG655
GC2-MT   GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG648
GC2-Del  GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG642
GC2-Ins  GCTCTCCTAACCGACGCGCGTCTGTGGAGAAGCGGCTTGGTCGGGGGTGGTCTCG654
         ************************************************************
```

Appendix 7: Transient Transfection Data Example.

Construct	Max luc	Luc – blank	β-gal	β-gal - blank	Normalised luc	Avg	Avg Dev	Rel. to WT	Rel. Avg Dev
Blank	109	0	0.067	0					
WT	2873	2764	0.719	0.652	4239.26				
WT	3047	2938	0.649	0.582	5048.11	4679	293.6	100	6.27
WT	3074	2965	0.691	0.624	4751.60				
Ins	3308	3199	0.734	0.667	4796.10				
Ins	3104	2995	0.704	0.637	4701.73	4744	34.3	101	0.72
Ins	3230	3121	0.726	0.659	4735.96				

This table is an example of the kind of data acquired from a transient transfection assay in HeLa cells, with the wild type topo IIα promoter reporter construct and the Ins construct. Luciferase values were normalised by subtracting the blank luciferase reading from each maximum luciferase value. The β-gal blank was subtracted from the β-gal readings. The level of luminescence was then divided by the β-gal to gain normalised values for each well. The average of the triplicates was calculated, as well as the average deviation (refer to Appendix 8.1). The luminescence values of each mutant construct, or in this case the Ins construct, were then divided by the WT luminescence level and multiplied by 100, to produce the expression level as a percentage relative to wild type. These results were then displayed on a bar graph. The relative average deviation was also calculated for use as error bars on graphs. Statistical significance was calculated as described in Appendix 8.2.

Appendix 8: Statistical Formulas.

8.1 Average deviation.

The average deviation was used to calculate fluctuations between data sets, illustrating the amount of variation between samples from the calcutlated mean.

$$\frac{1}{n} \sum X - \overline{x}$$

Where n = the number of observations

X = the individual observation

$\overline{x}$ = the average/mean

8.2 Students T-test.

Students T-tests were carried out using Microsoft Excel 2002, to verify any significant differences between data sets. P-values are shown in Table 5.1.

Formula for Students T-test:

$$t = \frac{Mean_{grp1} - Mean_{grp2}}{\sqrt{s^2 \left[\frac{1}{n_{grp1}} + \frac{1}{n_{grp2}} \right]}}$$

n = number of data points within a population

s^2 = is the vaiance of a population