Basal Transcription of Human Topoisomerase II

A thesis presented to Massey University in partial fulfillment of the requirements for the degree of Master of Science in Biochemistry

Natisha Magan

2002

Acknowledgments

Foremost I would like to express my thanks to Dr Kathryn Stowell who has been a phenomenal supervisor, in every sense of the word. Your enthusiasm, support and understanding have been much appreciated and will never be forgotten.

Thanks to all the people in the past who have worked on topoisomerase IIα and made it every bit as interesting as it is, Agnieszka Szremska, Patricia Hintz, Sam McLenachan and Dr Richard Isaacs.

I extend my thanks to all the members of the Twilite Zone who have been an irreplaceable source of amusement and made it such fun being in the lab. Extra special thanks to Jennifer Anderson, Christin Down, Robyn Marston, Brenda Belmont, Carole Flyger, Melanie Willingham, Penny Bradbury and Kirsty Allen, all of you have inspired me in so many ways. Many warm thanks, to John Tweedie and Mark Patchett for all their encouraging words and helpful advice.

I am grateful to have such a supportive family, thank you to my Mum, Dad, Mai, Jason and Pradeep who have provided me with much love and patience. Thank you to all my family and friends that have been there for me and given me the reassurances that I needed throughout this project especially, Jennifer Goh and Roshni Nathoo.

Finally, I would like to thank my very best friend, Lawrence.

Abstract

Topoisomerase II is a ubiquitously expressed enzyme, which is required for cell survival. It has the ability to alter the topological states of DNA by introducing transient double-stranded breaks in DNA. Humans have two topoisomerase II isoforms, α and β, and both are differentially expressed and localized. Tissues with rapidly proliferating cells exhibit elevated topoisomerase IIα gene expression whereas the β isoform is ubiquitously expressed amongst tissues.

In addition to a role in cell survival, a number of anti-cancer drugs have been shown to target human topoisomerase II *in vivo*. However, the development of drug resistance is a major clinical problem; for example, approximately 60% of breast cancers treated with the topoisomerase II poison doxorubicin become resistant to this drug. Down-regulation of topoisomerase II is thought to be one of the factors involved in the development of drug resistance, where the relative levels of topoisomerase IIα and topoisomerase IIβ in cells is thought to effect drug efficacy.

The expression of topoisomerase IIα and β is regulated at the transcriptional level, through binding of transcription factors to specific elements within the promoter sequence. Therefore investigating the transcriptional regulation of both isoforms could lead to an understanding of the mechanisms involved in the development of drug resistance. The initial aim of this study was to isolate a fragment of the upstream regulatory sequence of the topoisomerase IIβ gene and carry out systematic analysis of this sequence. However, this could not be pursued, as the clones that were examined did not contain the required topoisomerase IIβ sequence.

This study progressed to examine the relevance of three elements (GC1, ICB1 and GC2) within the topoisomerase IIα minimal promoter and the importance of the cognate transcription factors NF-Y, Sp1 and Sp3 in regulating the expression of the topoisomerase IIα gene. Electrophoretic mobility shift assays and transient transfection assays were used to study protein/DNA interactions and the functional significance of these interactions, respectively. Both NF-Y and Sp1 were shown to activate the transcriptional regulation of topoisomerase IIα by binding to their respective elements; in addition functional interactions between the two proteins bound to the promoter was observed.

Abbreviations

Amp	Ampicillin
AMSA	Topoisomerase II poison
Ap-2	Activator protein 2
ATF	Activating transcription factor
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
β-gal	β-galactosidase
bp	Base pairs (DNA)
BSA	Bovine serum albumin
CAT	Chloramphenicol acetyltransferase
CDE	Cell-cycle dependant element
cDNA	Synthetic DNA, generated from RNA
cpm	counts per minute
DMSO	Dimethyl sulfoxide
Dnase	Deoxyribonuclease
dNTP	Deoxynucleoside triphosphate (dATP, dTTP, dGTP, dCTP)
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra-acetic acid
EMSA	Electrophoretic mobility shift assay
FCS	Foetal calf serum
GCG	Genetics computer group
G segment	Gated segment (DNA)
GUS	β-glucuronidase
IPTG	Isopropyl thiogalactoside
HAT	Histone acetyl transferases
HeLa	Human cervical carcinoma cells
HEPES	N-[2-hydroxyethyl]piperazine-N'-[2-ethane sulfonic acid]
HFM	Histone fold motif
ICB	Inverted CCAAT box
ICBP90	Inverted CCAAT box binding protein Mr 90 kDa
IgG	Immunoglobulin G
IPTG	isopropyl-β-D-thiogalactopyranoside

kb	kilobases (DNA)
KB	Human epidermoid KB cancer cells
KB/VP-2	etoposide resistant KB cells
KB/VM-4	teniposide resistant KB cells
LB	Luria Bertani bacteriological media
MCF-7	Human breast cancer cells
MCS	Multiple cloning site
MDR	Multidrug resistance
MDR1	Multidrug resistance gene
Mnase	Micrococcal nuclease
MRP	Multidrug resistance-associated protein
MEM	Eagle's minimal essential media
mt	mutated/mutant
NEB	New England Biolabs
NF-Y	Nuclear factor Y
ONPG	o-Nitrophenol β-D-Galacto-pyranoside
PAGE	Polyacrylamide gel electrophoresis
$_{P}53$	Tumour suppressor protein
PBS	Phosphate buffered saline
PBSE	Phosphate buffered saline plus EDTA
PEG	Polyethylene glycol
pGL3B	pGL3Basic vector
PIC	Pre-initiation complex
PIPES	Piperazine-n,n'-bis(2-ethane sulfonic acid)
PMSF	Phenylsulfonylmethyl fluoride
Pol II	RNA polymerase II
Q-rich	Glutamine-rich
Rb	Retinoblastoma protein
RNase	Ribonuclease
RT	Room temperature
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
Sp1	Specificity protein 1
Sp3	Specificity protein

STET	Sucrose, Tris, EDTA and triton-X buffer
SV40	Simian virus 40
T segment	Transport segment (DNA)
T12	Human bladder cancer cells
TAE	Tris acetate EDTA buffer
TAFs	TBP associated factors
TATA	TATA box; conserved A/T rich septameter transcription 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
XK469	Topoisomerase IIβ poison (NSC 697887)
UV	Ultra-violet light
VM-26	Teniposide: topoisomerase II poison
VP-16	Etoposide
X-gal	5-bromo-4chromo-3-indolyl-β-D-galactopyranoside
wt	wild type

List of figures

List of tables

Table of Contents

Chapter 1: Introduction.

1.1 Overview.

DNA topoisomerases are essential nuclear enzymes that can adjust the topological states of DNA by introducing transient double-stranded breaks into the DNA helix. Topoisomerases are classified according to their mechanism of action as either type I or type II. Human type II topoisomerase II has two isoforms α and β, where both enzymes appear to share the same catalytic function. However, both Human topoisomerase II isoforms are differentially regulated and localized (Chung *et al*, 1989; Zandvliet *et al*, 1996); topoisomerase IIα is cell-cycle regulated and mainly nuclear (Grue *et al*, 1998), whereas topoisomerase IIβ is not cell-cycle regulated and predominantly nucleolar (Cowell *et al*, 1998).

A number of anti-cancer drugs have been shown to target human topoisomerase II *in vivo*. The drugs act by interfering with the normal action of topoisomerase II, either as topoisomerase II poisons or topoisomerase II inhibitors. Topoisomerase II catalytic inhibitors generally engender DNA topological problems causing cell death, whereas topoisomerase II poisons generate an accumulation of double-stranded breaks causing cell death. The development of drug resistance is a major clinical problem; for example, approximately 60% of breast cancers treated with the topoisomerase II poison doxorubicin become resistant to this drug. Both topoisomerase IIα and topoisomerase IIβ have been shown to be targets for chemotherapeutic agents and changes in expression have been correlated with drug sensitivity.

Several studies have revealed that resistant cell lines and tumors have decreased levels of topoisomerase IIα and topoisomerase IIβ (Son *et al*, 1998; Dingemans *et al*, 1998; Lage *et al*, 2000). Thus, the down-regulation of topoisomerase II is thought to be one of the factors involved in the development of drug resistance. Other cellular events implicated in the development of drug-resistance are; mutations in the catalytic region of the protein (Dingemans *et al*, 1998; Robert and Larsen, 1998), phosphorylation status of the enzyme (Burden and Sullivan, 1994) and topoisomerase II RNA stability (Goswami *et al*, 1996). Interestingly, the ratios of topoisomerase IIα and β present in tumors have been suggested to effect drug efficacy (Padget *et al*, 2000, reviewed in Isaacs *et al*, 1995). To date, the relative contributions of each topoisomerase isoform to drug resistance is not fully understood.

The topoisomerase IIα promoter has been partially characterized in a number of DNA-binding and functional studies (Hochhauser *et al*, 1992; Isaacs *et al*, 1996a, reviewed in Isaacs *et al*, 1998; Bakshi *et al*, 2001) but in comparison little is known about the topoisomerase IIβ promoter, therefore investigating this area is of particular interest. This thesis focuses on the transcriptional regulation of human topoisomerase IIα and topoisomerase IIβ as a means of understanding the mechanisms involved in down-regulation associated with drug resistance.

1.2 Topoisomerases.

Topoisomerases have a general role in relaxing DNA, which entails correcting normal topological problems that are generated during DNA replication, transcription, recombination and chromatin remodeling. They act by introducing transient breaks in the sugar-phosphate backbone of DNA to enable the passage of an intact DNA strand through the break (refer to figure 1.1). DNA cleavage by all topoisomerases is accompanied by the formation of a transient phosphate ester bond between a tyrosine residue in the protein and one of the ends of the broken DNA strand (reviewed in Capranico and Binaschi, 1998). DNA topology is altered whilst the covalent intermediate remains and the topoisomerase enzyme is released as the DNA is religated (Champoux, 2001). Topoisomerases have channels or cavities within their structures to accommodate DNA binding and enable specific catalytic activity (Orphanides and Maxwell, 1994).

Eukaryotes have three types of topoisomerase: I, II and III. Topoisomerase I was the first topoisomerase discovered and it can catalyze the removal of both positive and negative supercoils within DNA by introducing a transient single-stranded DNA break. Topoisomerase I mitigates torsional strain by relieving one supercoil during each catalytic cycle and is independent of ATP (reviewed in Nitiss, 1998). Human topoisomerase III was only recently discovered and has two isoforms, α and β. Topoisomerase III is thought to partially relax negatively supercoiled DNA in an ATP-dependent manner during replication (Nitiss, 1998) with a strong preference for single-stranded DNA (Kim *et al*, 1998). Topoisomerase II is the only topoisomerase that is essential for cell viability. It can relax negative and positive supercoils by making transient double-stranded breaks in DNA at a precise site and religating the nicked strand after the conformational alteration (refer to figure 1.1). It is thought that the process of chromosome condensation and segregation during mitosis and meiosis requires the activity of topoisomerase II, whereas other topological DNA

modifications, such as DNA relaxation can be performed by topoisomerase I (Marsh *et al*, 1997). Topoisomerase II may also have a role in maintaining the structural integrity of chromosomes as a scaffold protein (Kimura *et al*, 1996 and references within).

Mammals have two topoisomerase II isoforms; IIα and IIβ encoded by separate genes on different chromosomes (Tan *et al*, 1992; Isaacs *et al*, 1998). Topoisomerase II α and β display strong amino acid sequence identity and are functionally closely related, however they are regulated differently (Austin *et al*, 1993). Topoisomerase IIα is expressed in a cell-cycle dependent manner and is found predominantly in highly proliferating cells (Woessner *et al*, 1990), whereas topoisomerase IIβ expression remains constant throughout the cell-cycle (Austin *et al*, 1995).

Three functional domains have been recognized within type II topoisomerase enzymes: the ATP-binding domain located near the N-terminus, a central catalytic core domain, followed by the C-terminal domain. The central catalytic core is composed of two constituents, one that is required for the dimerization of topoisomerase II and the other required for DNA binding, cleavage and religation. The DNA cleavage and religation domain encompasses a putative active tyrosine residue involved in DNA break stabilization in addition to amino acid motif(s) that have been implicated in the binding of drugs to the cleavable complex (Lang *et el*, 1998). The binding of ATP to one of the ATP-binding domains has been suggested to orchestrate a conformational change within the enzyme that leads to binding of a second ATP molecule to the other ATP-binding domain (Champoux, 2001). Hydrolysis of these ATPs leads to the presumptive release of ADP instigating the enzyme to be arranged back into its original conformation, where upon DNA capture can reoccur (refer to figure 1.1). The C-terminal end of topoisomerase II contains the largest amino acid sequence divergence between the two isoforms (Austin *et al*, 1995; Cowell *et al*, 1998) and this region is thought to be responsible for distribution and targeting of the protein within cells (Adachi *et al*, 1997). Both isoforms can be phosphorylated, and it has been demonstrated that phosphorylation of the C-terminal domain of topoisomerase IIβ could be involved in controlling protein localization (Cowell *et al*, 1998).

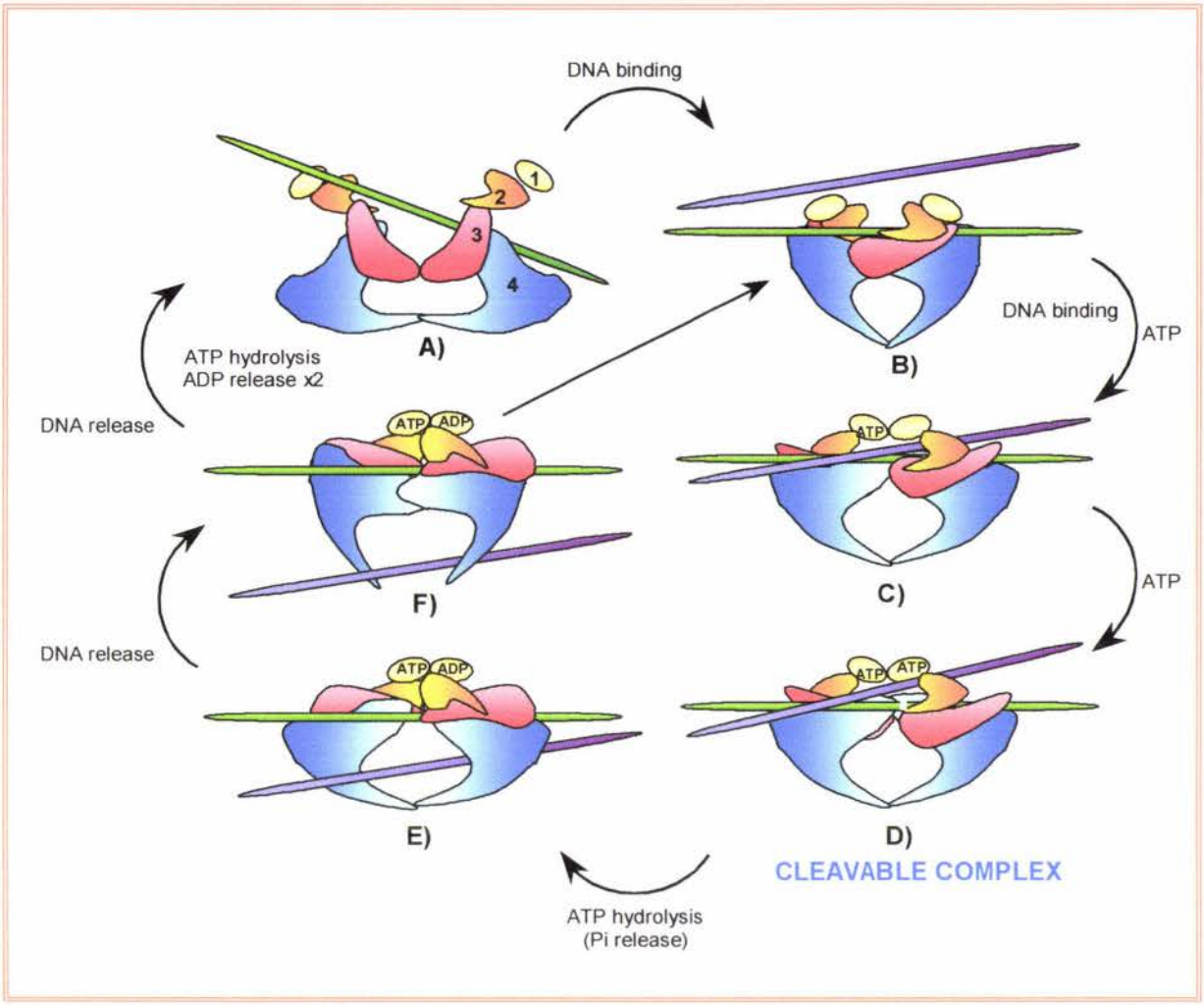


Figure 1.1: Schematic representation of the catalytic cycle of topoisomerase II. **B)** The topoisomerase II homodimer (blue, red, orange and yellow subunits 1-4) binds to a region of double-stranded DNA called the gated (G) segment (green). **C)** ATP binds and dimerization of the ATPase domains (1) occurs, which traps the transfer (T) segment (purple). **D)** A second ATP molecule binds and a double-stranded break is introduced within the G segment. **E)** Passage of the T segment through the G segment and the DNA break is religated. **F)** The T segment is released from the enzyme. The G segment can be released (**A**), or retained (**B**) if further conformational changes are required. **A)** ATP hydrolysis with release of ADP is required to reset the enzyme to an open conformation (Figure created from sources: Baird *et al*, 1999; Champoux, 2001).

1.3 Anti-cancer drug action and the development of resistance.

Anti-cancer drugs that target topoisomerase II generally fall into one of the following two classes: catalytic inhibitors and topoisomerase II poisons. Anti-cancer drugs generally act by stimulating cell death brought about by an accumulation of DNA damage. The drugs themselves do not damage the DNA, but the presence of topoisomerase II-targeting drugs in cells has the ability to convert normal topoisomerase II to a cellular toxin by inhibiting normal topoisomerase II enzyme action (reviewed in Robert and Larsen, 1998). The topoisomerase II poisons have been associated with down-regulation of topoisomerase II expression and consequently the development of drug resistance (reviewed in Larsen and Skladnowski, 1998).

1.3.1 Drugs and topoisomerase II.

Topoisomerase II poisons such as doxorubicin, amsacrines and anthracyclines act by forming a ternary complex of DNA, drug and enzyme, thereby halting the enzyme mid-cycle. These drug-stabilized cleavable complexes (refer to figure 1.2A) are toxic to cells, as normal cellular processes such as replication and transcription convert them into irreversible double-stranded breaks (Austin and Marsh, 1998). Stabilization of this complex is thought to be sufficient to inhibit cell proliferation perhaps by being perceived as a lethal signal, which in turn triggers cells to enter apoptosis (Walker *et al*, 1991). Therefore, the cleavable complex represents a state of intense fragility for the DNA molecule. Drug-induced stabilization of the cleavable complex is actually reversible if the drug is removed from cells before the cellular replication and transcription machinery has a chance to act (reviewed in Baguley and Ferguson, 1998).

The amount of DNA damage accumulated in cells depends on both the amount of drug available and the amount of topoisomerase II present. In cellular conditions where high levels of topoisomerase II is found, there is an accumulation of cleavable complexes hence a greater number of double-strand breaks (refer to figure 1.2B) inducing high cellular toxicity. The ability of the enzyme to form cleavable complexes as well as the degree of interaction between drug and topoisomerase II will also influence cellular sensitivity to the topoisomerase II interfering drugs.

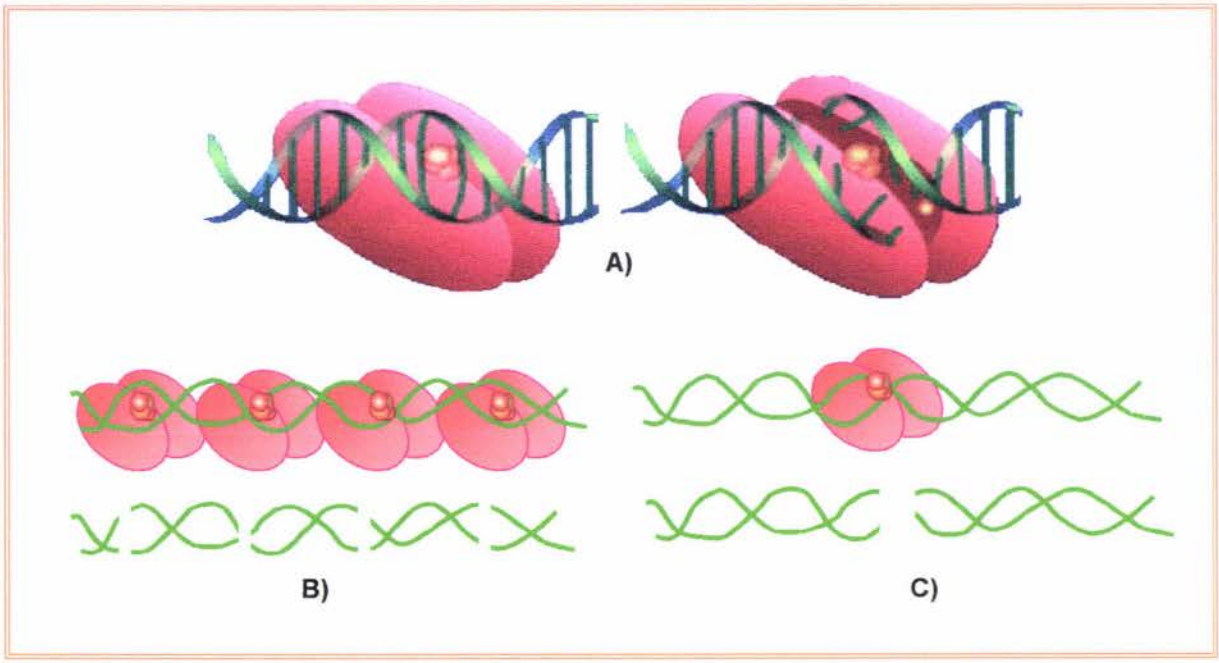


Figure 1.2: Schematic representation of topoisomerase II poison action.
In general, the magnitude of drug-induced DNA strand break production directly correlates with the magnitude of drug-induced cytotoxicity. **A)** Left: Topoisomerase II poison (orange) associates with the (green) DNA-enzyme (pink) complex. Right: A double-stranded break is introduced by topoisomerase II and the drug stabilizes the cleavable complex. **B)** Cells that contain high levels of topoisomerase II in the presence of anti-cancer drug will sustain an accumulation of double-stranded breaks leading to cell death, **C)** whereas cells that have low levels of topoisomerase II in the presence of drug will have fewer breaks (adapted from Slide: Rhone-Poulenc Rorer Pharmaceuticals, 1999).

Topoisomerase II inhibitors function by impeding normal topoisomerase II activity without stabilizing the cleavable complex and they can act in a number of different ways. Some inhibitors act by trapping the DNA-enzyme complex in a closed clamp configuration in the presence of ATP; for example the dioxopiperazines (Roca *et al*, 1994). Novobiocin, another topoisomerase II inhibitor directly interferes with the ATPase domain of topoisomerase II (Reviewed in Isaacs *et al*, 1995) and Suramin inhibits topoisomerase II activity by reducing enzyme phosphorylation. Generally cell death brought about by the presence of topoisomerase II inhibitors is accredited to an accumulation of topological problems in cells, therefore drug efficiency is dependent on its' ability to prevent normal topoisomerase II activity.

While it has been shown that amsacrines may preferentially target topoisomerase IIβ (Dereuddre *et al*, 1997; Gao *et al*, 1999), other drugs such as etoposide can target both isoforms (Willmore *et al*, 1998). It is possible that both topoisomerase II isoforms could vary in relative drug sensitivities, but current knowledge suggests that most topoisomerase IIα poisons also act as topoisomerase IIβ poisons. However, if isoform-specific drugs were developed tumors could be targeted according to the specific expression of either isoform thus increasing efficacy of chemotherapy. Resistance to drugs that target topoisomerase IIα and β can be developed in malignant cells and this has been shown to occur *in vivo* and *in vitro* (Lang *et al*, 1998 and references therein). Consequently, a decrease in topoisomerase II expression or ability to interact with anti-cancer drugs may be responsible for the resistance to these drugs.

1.3.2 Decreased topoisomerase II expression and protein activity.

To date most reported cases of drug resistance have been attributed to a decrease in topoisomerase IIα expression (Robert and Larsen, 1998). However, the α isoform was identified much earlier than the β isoform, therefore the role of topoisomerase IIβ in drug action should not be excluded. Topoisomerase IIβ has been shown to interact with drugs at the same DNA sequences as the α isoform (Austin and Marsh, 1998). In addition, the β isoform has been shown to be decreased or even absent in some resistant cell lines, such as human small lung cancer cells resistant to mitoxantrone (Gao *et al*, 1999 and references therein), Chinese hamster lung cells (DC-3F) resistant to 9-OH-ellipticine (Dereuddre *et al*, 1997) and human leukemia cell line (HL-60) resistant to amsacrine (Herzog *et al*, 1998). Consequently, current knowledge suggests that topoisomerase IIβ is targeted by most of the same drugs that target topoisomerase IIα.

Many different studies have shown that reduced levels of topoisomerase IIα can be found in resistant cell lines and tumors (Wang *et al*, 1997a; Dingemans *et al*, 1998; Takano *et al*, 1999). This reduction in topoisomerase IIα has been directly correlated with a decrease in topoisomerase IIα mRNA, where the mRNA decrease has been accredited to transcriptional activity being repressed and not due to the integrity of mRNA being compromised (Kubo *et al*, 1995; Asano *et al*, 1996). Studies utilizing anti-sense RNA fragments (genetic suppressor elements) have shown that down-regulation of topoisomerase IIα expression can generate drug resistance to etoposide in mammalian cells (reviewed in Isaacs *et al*, 1998). The

presence of anti-sense RNA fragments inhibits normal topoisomerase IIα mRNA processing, which in turn brings about a reduction in the production of cellular topoisomerase IIα protein and consequently generates drug resistance.

Phosphorylation of topoisomerase IIα enzyme has been shown to inhibit the enzymes catalytic activity and influence its vulnerability to topoisomerase II-targeting drugs (reviewed in Isaacs *et al*, 1998). It has been suggested that the different phosphorylation states of the α and β isoforms may affect the enzymes ability to bind to chromosomes, thereby affecting its activity (Burden and Sullivan, 1994; Kimura *et al*, 1996). Many phosphorylation sites have been mapped within the α and β isoforms, with most of the sites residing in the least conserved C-terminal domain of the enzymes where the presumptive localization signal resides (Adachi *et al*, 1997; Cowell *et al*, 1998). One particular study has demonstrated that cellular localization of topoisomerase IIα is important, as trafficking of topoisomerase IIα to the cytoplasm from the nucleus was found to generate resistance to etoposide in human leukemic cells, without changes in the relative amounts of topoisomerase IIα (Valkov *et al*, 2000). Therefore, it is possible that phosphorylation could regulate nuclear import of the topoisomerase II enzymes.

Other studies have shown that topoisomerase IIα down-regulation is due to a number of factors; such as histone deacetylation (Adachi *et al*, 2000), methylation of the gene (Tan *et al*, 1992), mutations in the enzyme (Dingemans *et al*, 1998), RNA stability (Goswami *et al*, 1996), binding of p53 to the promoter (Sandri *et al*, 1996) and altered expression of multidrug resistance gene (Bredel, 2001). In addition, it has been shown that topoisomerase IIα can form heterodimers with topoisomerase IIβ *in vivo*, which has been suggested to occur as a means of maintaining the α isoforms catalytic activity, where the proportion of heterodimers in relation to topoisomerase IIα homodimers is also thought to have a role in drug sensitivity (Gromova *et al*, 1998).

It should be noted that the mechanisms for the development of resistance could vary between cell and cancer type. One such mechanism is that involving the membrane associated chloride pump P-glycoprotein, which causes drug efflux from the cell due to over-expression of the multidrug resistance (MDR1) gene (Bredel, 2001) or altered expression of multidrug resistance-associated protein (MRP). Conversely, some proteins have been found to stimulate

topoisomerase IIα activity. Proteins such as CREB, ATF-2 and c-Jun are thought to interact with the topoisomerase IIα enzyme and increase its ability to decatenate DNA (Bhat *et al*, 1999 and references within).

1.4 Transcription in Eukaryotes.

The major control point of eukaryotic gene expression occurs at the initiation of transcription, through a set of general transcription factors that assemble at a core promoter to form a pre-initiation complex (PIC), which is required for basal transcription. Generally, the promoter region of a gene consists of common core promoter elements that are recognized by general transcription initiation factors (TIFs) as well as gene-specific DNA elements that recruit additional regulatory factors. Interaction between all these components is orchestrated by protein-protein contacts, which then can enable transcription by RNA polymerase II (Pol II). The ability of sequence-specific transcriptional activators to associate with DNA is dependent on a number of factors including DNA organization and accessibility. The regulatory regions of a promoter vary between different genes allowing the gene-specific assembly of transcription factors in a combinatorial fashion.

The TATA box is a specific sequence usually situated about 25 base pairs (bp) upstream of the transcription start site and tends to be surrounded by GC rich sequences. The GC rich sequences are thought to be a factor that assists the transcription machinery to determine the mRNA start site. The TATA-binding protein (TBP) recognizes and binds to DNA at the TATA box followed by recruitment of general TIFs as a multiprotein complex (Hoffmann *et al*, 1997). TBP is essential for the recruitment of Pol II and is thought to orchestrate the bending of DNA at the TATA box (reviewed in Malik and Roeder, 2000) through its exclusive contacts with the minor groove of the DNA double helix. This enables Pol II and general transcription factors to be brought into closer proximity to each other. Other transcription factors are also required and their recruitment to the complex may facilitate the movement of Pol II away from the promoter, or enable access to DNA by having a loosening affect on the chromatin structure.

Any proteins that are required for the initiation of transcription but are not part of the Pol II complex are known as transcription factors. Transcription factors (termed *trans*-acting factors) are either repressors or activators, which act by binding directly to regulatory DNA element(s) or indirectly through protein-protein interactions. Many transcription factors

recognize upstream DNA-binding elements (termed *cis*-acting elements) within the gene promoter region and greatly influence the efficiency and rate of initiation of transcription by interacting in a precise order to build a complex that associates with the basal transcription complex and Pol II.

The same general transcription factors associated with TATA box-containing promoters are required for TATA-less promoters, however other recognition mechanisms are requisite for PIC formation. The TBP-associated factors (TAFs) are required for basal transcription and determining promoter specificity. TATA-less promoters can still recruit TBP, as a wide range of A/T-rich sequences are capable of functioning as TATA boxes. This is consistent with the knowledge that TBP can recognize DNA primarily through sequences independent of minor groove contacts (reviewed in Smale, 1997). It is also possible that TBP can interact indirectly with the DNA by protein-protein interactions (White and Jackson, 1992) in a complex such as TFIID (a TIF), binding with less affinity but still enabling Pol II to be positioned correctly at the start site. TFIID is a multiprotein complex that directs PIC assembly and is composed of TBP and TAFs. As evidence of this, it has been shown that TAFs can interact tightly with TBP and mediate activation as part of the TFIID complex (Bellorini *et al*, 1997) and TFIID is able to interact directly with Pol II (Hoffmann *et al*, 1997).

The promoter of both topoisomerase IIα and topoisomerase IIβ are GC rich and neither contain a TATA box. These characteristics have commonly been found in the promoters of housekeeping genes (Yoon *et al*, 1999). Examining the mechanism(s) of transcriptional regulation will assist in understanding how these genes are regulated.

1.5 Topoisomerase IIα.

Topoisomerase IIα is an essential 170 kDa nuclear enzyme encoded by a gene located on chromosome 17 at 17q21-22 (Tan *et al*, 1992). Expression of topoisomerase IIα is cell-cycle regulated and influenced by the proliferative status of cells, where topoisomerase IIα expression and activity reaches a maximum during cell division however, topoisomerase IIα protein levels decrease as the proliferation of cells decrease (Woessner *et al*, 1990). Highly proliferating tissues such as lymphoid tissue, basal layers of epithelial and glandular tissues of colon, lung and testis exhibit increased topoisomerase IIα expression (reviewed in Isaacs *et al*, 1995). Phosphorylation of topoisomerase II is also cell-cycle dependent and may serve to regulate the activity of the topoisomerase IIα enzyme.

The minimal topoisomerase IIα promoter encompasses a fragment approximately 650 bp upstream of the ATG translation initiation codon (Hochhauser *et al*, 1992). Interestingly, it has been found that a truncated topoisomerase IIα promoter (up to –617 bp) exhibits maximal expression of reporter genes in transient transfection studies. When this activity was compared to full-length (about 2 kb) topoisomerase IIα promoter activity, the full-length promoter exhibited decreased activity (Hochhauser *et al*, 1992). This suggests that there are possible upstream negative regulatory elements, which may serve to down-regulate topoisomerase IIα gene activity. Analysis of the –617 topoisomerase IIα sequence has identified a number of important sequences that have been implicated in transcription of topoisomerase IIα (refer to figure 1.3). These include five inverted CCAAT boxes (ICB1-5), two GC rich regions (GC1 and GC2), a consensus sequence for an activating transcription factor (ATF) binding site, a potential binding site for Myb, a Myc/Max site, an activator protein-2 (AP2) site and a cell-cycle responsive element CDE/CHR, where the major transcription site is numbered as +1. ATF binding sites were originally documented for their role in regulating gene transcription in response to changing cyclic AMP levels (reviewed in Isaacs *et al*, 1998). However the ATF binding site is probably not important for topoisomerase IIα regulation in confluence arrested cells, as a deletion of this element has been shown to have no affect on confluence-regulated topoisomerase IIα transcription (Isaacs *et al*, 1996a). The cell-cycle responsive element CDE/CHR located between ICB2 and ICB3 (in an inverted orientation) has been implicated in repressing glutamine-rich transcription activators. Experiments done using the human *cdc25C* gene, have demonstrated that the cell-cycle regulation of this gene is attributed to the presence of an upstream

CDE/CHR element in conjunction with several upstream regulatory elements; such as binding sites for transcription factors Sp1 and NF-Y (Zwicker *et al*, 1995). The Myb binding site has not been well studied, however it has been shown that binding of the *c-myb* proto-oncogene product (c-Myb) to the Myb binding site can activate the topoisomerase IIα promoter and that this effect is lymphoid and myeloid lineage dependent (Brandt *et al*, 1997). Located at position –21 is the partial consensus Myc/Max site, it is not known if this element has any influence on topoisomerase IIα expression, but this interaction is thought to work in conjunction with a second transcription factor, Max.

Transcriptional control of topoisomerase IIα is thought to be regulated through binding of factors to important elements within the topoisomerase IIα promoter. Identifying the transcription factors that interact with the topoisomerase IIα promoter and examining gene regulation in response to these transcription factors will be a useful approach to study its transcriptional regulation. This in turn should provide insights into the molecular mechanisms involved in the down-regulation of topoisomerase IIα that is associated with drug resistance.

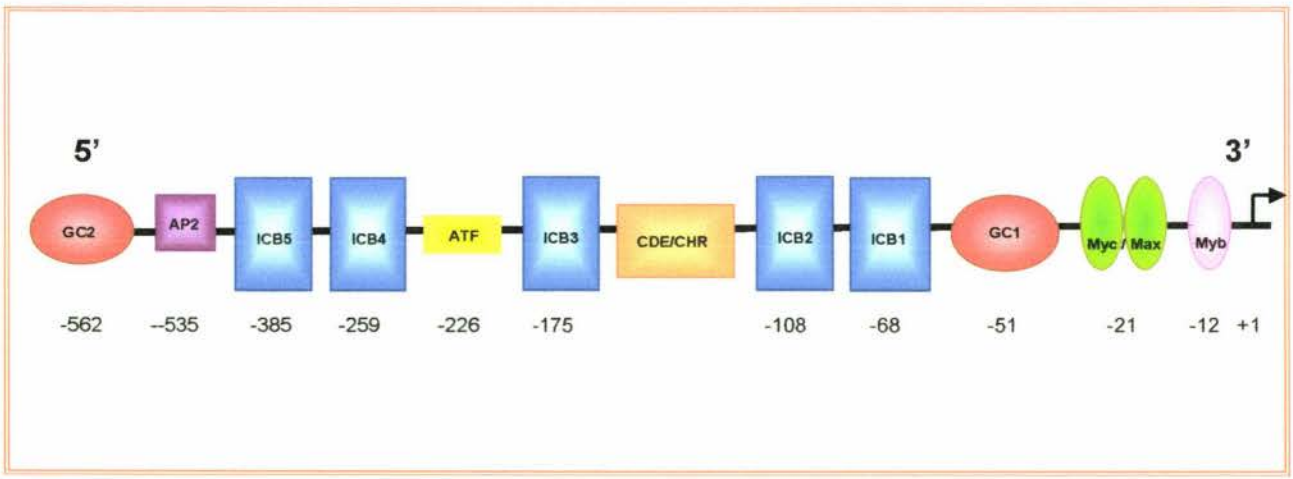


Figure 1.3: Schematic representation of the –617 topoisomerase IIα promoter. Positions of the putative elements are shown and numbered with respect to the major transcription start site +1, indicated by the arrow. There are five inverted CCAAT boxes (blue ICB1-5), 2 GC boxes (red GC1 and 2), as well as sites for Myb (pink), Myc/Max (green), CDE/CHR (orange), ATF (yellow) and AP2 (purple). (Adapted from Isaacs *et al*, 1998).

1.5.1 CCAAT Elements.

The CCAAT box is a conserved regulatory sequence, which tends to be located upstream of the transcription start point and is found in many eukaryotic promoters. Among many upstream elements, the CCAAT box is one of the most common being found in 25% of eukaryotic promoters in the direct or inverse orientation. Many proteins recognize and bind to CCAAT boxes, including ICBP90 (inverted CCAAT box binding protein 90), Y-box binding protein (YB-1), p53 and transcription factor NF-Y (nuclear factor Y). Inverted CCAAT boxes are thought to be important for basal-level transcription and mutations in the CCAAT motif have been shown to alter transcription from a variety of promoters (Yoon *et al,* 1999).

ICBP90 is a recently identified protein, isolated from Jurkat leukemia cell lines and adult thymus, has the ability to bind to CCAAT elements. It is present predominantly in tissues that possess highly proliferating cells, such as adult thymus, fetal thymus, fetal liver and bone marrow. However, in tumor cells such as HeLa, ICBP90 is present even when the cells are confluent. ICBP90 appears to have the ability to stimulate the topoisomerase IIα promoter in COS-1 cells *in vitro,* where electrophoretic mobility shift assays (EMSA) enabled the identification of ICB2 as the preferential target (Hopfner *et al*, 2000; Hopfner *et al*, 2001).

p53 is a protein that is normally expressed at very low levels, it is one of the most important regulators of cell-cycle progression in mammals. However, p53 can be induced and activated in response to DNA damage, where it controls the initiation of cell-cycle arrest and entry into the apoptotic pathway (Reisman and Loging, 1998). Studies using the minimal promoter of topoisomerase IIα have demonstrated that, p53 has the ability down-regulate topoisomerase IIα expression significantly (Sandri *et al*, 1996). Experiments done where all five of the ICBs were deleted within topoisomerase IIα promoter demonstrated that any p53-mediated down-regulation of topoisomerase IIα expression was abolished, suggesting the ICB elements were required for p53 mediated down-regulation (Wang *et al*, 1997b). Interestingly, p53 protein has also been shown to co-precipitate with topoisomerase II proteins, where it is speculated to occur a means of controlling p53-mediated apoptosis (Yuwen *et al*, 1997).

YB-1 may also play a role in topoisomerase IIα expression. YB-1 has been found to specifically bind to ICB elements and it is thought to be responsible for the activation of the

MDR1 promoter in response to various environmental stimuli such as drugs, etoposide and teniposide (reviewed in Kubo *et al*, 1995).

Many experiments have been done to examine the importance of the ICB elements within the topoisomerase IIα promoter. Each ICB element appears to have slightly different functional responsibilities, however still work in a combinatorial fashion to influence topoisomerase IIα expression. Studies exploring the relevance of ICB1 in the topoisomerase IIα promoter have shown that a mutation in ICB1 can abolish the phenomenon of topoisomerase IIα down-regulation in resistant cells (Takano *et al*, 1999), suggesting that the binding of negative regulators to ICB1 could be one of the factors involved in down-regulation of topoisomerase IIα. Other work using a different cell line (NIH3T3) also demonstrated this by showing that enhanced protein binding to ICB1 induced down-regulation of topoisomerase IIα in serum starved cells (Falck *et al*, 1999). Similar conclusions were drawn about ICB1 when Furakawa *et al* (1998) subjected T24 human urinary bladder cancer cells to heat shock. Their experiments showed that the transcriptional activation of the topoisomerase IIα gene by heat shock requires the release of a negative regulatory factor from ICB1; significantly, the protein implicated to bind to ICB1 in this case was shown not to be transcription factor NF-Y.

Isaacs *et al* (1996a) have shown that topoisomerase IIα is down-regulated in confluence-arrested cells, and this is thought to be due to a factor that binds to ICB2 (position –109 to –105). ICB2 was shown to be essential for transcriptional repression in confluence-arrested cells in transient transfection assays using a combination of deletion and mutational analysis, whereas other ICB elements of topoisomerase IIα were not required to induce this effect (reviewed in Isaacs *et al*, 1998). In addition, deletion of ICB2 resulted in down-regulation of topoisomerase IIα promoter activity (Takano *et al*, 1999).

In comparison to the other ICB elements, less work has been done on the relevance of ICB3 ICB4 and ICB5. The first study to implicate ICB3 in the down-regulation of topoisomerase IIα in human leukemic cell lines was carried out by Morgan and Beck (2001). This particular cell line was of interest because topoisomerase IIα expression was up-regulated when drug-resistant. ICB3 was implicated in contributing to this expression pattern and was shown to be able to associate with NF-Y.

Out of each of the ICB elements, ICB5 is thought to have the least influence on topoisomerase IIα expression. Experiments comparing the topoisomerase IIα expression levels, found that there was no difference in topoisomerase IIα expression between two constructs (one containing ICB1-4 and the other containing ICB1-5) in HeLa S3 cells (Hochhauser *et al*, 1992). In addition, EMSA studies using topoisomerase IIα ICB elements found that four of the ICB (1-4) elements were able to compete for binding against each other indicating these elements have the ability to bind similar proteins. ICB5 did not exhibit the same binding patterns as the other four ICB elements even though it did appear to bind protein(s) (Herzog *et al*, 1997). EMSA studies using topoisomerase IIα ICB elements found that four of the ICB (1-4) elements were able to compete for binding against each other indicating these elements have the ability to bind similar proteins. ICB5 did not exhibit the same binding patterns as the other four ICB elements even though it did appear to bind protein(s) (Herzog *et al*, 1997). There was no difference in topoisomerase IIα expression between two constructs (one containing ICB1-4 and the other containing ICB1-5) in HeLa S3 cells, suggesting ICB5 is not important in topoisomerase IIα regulation (Hochhauser *et al*, 1992).

1.5.2 Nuclear Factor Y (NF-Y).

Generally, a mutation in the CCAAT motif causes a several-fold decrease in transcriptional activity both *in vitro* and *in vivo* (Maity and Crombrugghe, 1998). In a number of promoters NF-Y has an absolute requirement for the CCAAT pentanucleotide sequence, which can be recognized in the forward or reverse orientation. Furthermore, NF-Y appears to have a strong preference for specific flanking sequences (Bi *et al*, 1997), making it a bonafide CCAAT box-identifying transcription factor. CCAAT boxes alone cannot activate transcription and it has been suggested that NF-Y may encourage other transcription factors to associate with their targets.

NF-Y has been identified as the CCAAT box activator in over 100 promoters, and is highly conserved throughout evolution where human; mouse and rat share almost 99% similarity in amino acid sequences (Sinha *et al*, 1996). NF-Y, also known as CBF, CP1 or YEBP, is ubiquitously expressed in all mammalian tissues (Coustry *et al*, 2001), however it has been shown to be present in higher amounts in proliferating cells than in confluent cells (Isaacs *et al*, 1996a). In order for NF-Y to bind DNA, all three of its subunits are required (Bi *et al*, 1997); NF-YA, B and C. NF-YB and C interact tightly with one another, and their

association is prerequisite to association with NF-YA, which is thought to be stabilized upon consequent DNA interaction (Laing and Maity, 1998). Significantly, all three subunits of NF-Y have been shown to contact DNA (Sinha *et al*, 1996) at three different regions (Bi *et al*, 1997) and a mutation in one subunit inhibits all contacts on the DNA (Laing and Maity, 1998, Hu and Maity, 2000).

NF-YA domains can be divided into two regions; the N-terminal domain is required for NF-YB/C association and the C-terminal for DNA binding (reviewed in Mantovani, 1999). NF-Y has two separate glutamine (Q)-rich activation domains; they are located on NF-YA and NF-YC and become redundant in the trimer (di Silvio *et al*, 1999). When both of the Q-rich domains of NF-YA and C are deleted a truncated version of NF-Y forms that can still bind DNA, however it can no longer activate transcription of topoisomerase IIα (Coustry *et al*, 2001). This indicates these Q-rich domains are important for transcriptional activation. The Q-rich domains are considered to be crucial in coordinating protein-protein interactions between transcriptional activators and other components of the transcription machinery, for example to direct binding of TAFs (Mantovani *et al*, 1992). This may be of particular significance in TATA-less promoters where NF-Y is closer to the transcription start site, and so could direct the formation of the PIC.

NF-YB/NF-YC dimerize tightly by virtue of their histone fold motifs (HFM), which are 65 amino acid stretches common to all histones. It has been speculated that sequence specificity of NF-Y stems from the HFMs within its subunits (reviewed in Mantovani, 1999). The HFM is also shared by some of the TBP-associated TAFs that mediate activation as part of the TFIID complex (Burley and Roeder, 1997). NF-YB can associate with TFIID in the absence of NF-YA and it has been proposed that two pools of NF-YB and NF-YC exist, where one set associates with each other forming NF-YB/NF-YC complexes and the other set associates with TBP (Bellorini *et al*, 1997). This provides evidence to suggest that NF-Y may have a role in transcription initiation. It has been suggested by some that NF-Y may act more as a general promoter organizer than a powerful activator of transcription (di Silvio *et al*, 1999).

Histones are required for chromatin formation, where histone association with DNA occurs in a highly ordered and condensed manner that results in the nucleosome structure. The binding of transcription factors to promoter sequences is generally inhibitory when DNA is organized in this manner, and therefore needs to be remodelled to enable access. Coustry *et*

al (2001) have shown that NF-Y can disrupt nucleosomal organization. These authors used cloned linearized topoisomerase IIα promoter DNA (wild type and mutated CCAAT elements), which was induced to assemble in a regular nucleosomal array. This was then treated with (and without) NF-Y and digested with *Micrococcal nuclease* (Mnase). DNA is coiled in regular arrays of nucleosomes, and individual nucleosomes can be separated by digestion with Mnase, as Mnase can sever DNA at the junction between nucleosomes, which is not protected by association with histones. If DNA remained organized around histones the Mnase digestion patterns would remain in a regular array, however if the nucleosomal array was disrupted the regular Mnase digestion patterns would be disrupted. The patterns of digestion indicated that specific binding of NF-Y to the CCAAT elements disrupted the regular nucleosomal structure over the nucleosomal promoter (refer to figure 1.4) with increased disruption at the transcription initiation site +1. Nucleosome disruption was less when only one ICB1 was present compared to when four ICBs (1-4) were present, establishing that multiple CCAAT boxes may be functionally important. In addition, NF-Y was able to activate transcription from topoisomerase IIα completely repressed by addition of histone H1. These results demonstrate that binding multiple NF-Y complexes generate greater nucleosome disruption and hence activate transcription. Significantly, NF-Y has been shown to interact with preformed histones and inhibit the formation of nucleosomes in a reconstituted nucleosome system (Motta *et al*, 1999). Therefore, NF-Y is thought to be able to disrupt the chromatin organization of promoters by competing for histone binding.

Acetylation of histones in the topoisomerase IIα promoter has also been shown to affect transcription. Histone acetylation has the ability to weaken histone interactions with DNA, which results in the loosening of chromatin and consequently enable the access of transcription factors. A recent study has suggested that NF-Y can also interact with or recruit histone acetyl transferases (HAT) to the promoter, thereby stimulating histone acetylation and regulating cell-cycle dependent transcription of the topoisomerase IIα gene (Adachi *et al*, 2000). Deacetylation of histones has the opposite effect of acetylation.

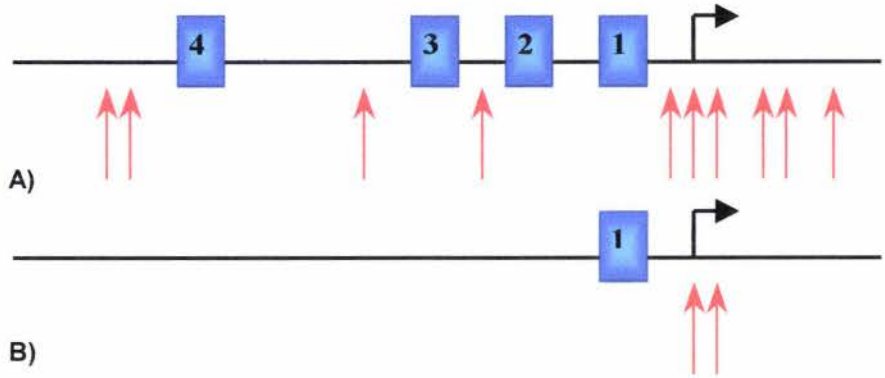


Figure 1.4: Schematic representation of NF-Y mediated chromatin disruption of the topoisomerase IIα promoter.

Red arrows show the Mnase sensitive regions of the promoter assembled into chromatin and blue depicts the relative ICB elements. This figure illustrates that multiple NF-Y binding sites can increase the comparative amounts of nucleosome disruption. **A)** Four ICB elements present, many Mnase sensitive sites due to NF-Y mediated nucleosome disruption. **B)** Only ICB1 present in the topoisomerase IIα promoter, only two cleavage sites therefore less NF-Y mediated nucleosome disruption (adapted from Coustry *et al*, 2001).

The binding of NF-Y to DNA has also been shown to induce distortion of the double helix, for which the Q-rich domains of NF-YA and C have been implicated (Liberati *et al*, 1999). Actually, NF-Y is thought to be capable of bending DNA at CCAAT sites by 62-82° and twisting DNA rotationally by 100° (Ronchi *et al*, 1995), which is reminiscent of the bending of DNA induced by TBP. Experiments carried out using DNA with two CCAAT boxes demonstrated the phenomenon of DNA bending, and found that the amplitude of angle distortion induced by NF-Y is dependent on the distance between CCAAT boxes (Liberati *et al*, 1999) and the CCAAT flanking sequences (Ronchi *et al*, 1995). DNA bending by transcription factors is thought to contribute significantly to transcriptional regulation, as it can bring elements closer together. It has been suggested that in the absence of TBP-TATA interactions at promoters, NF-Y could function as a pivotal factor in connecting upstream activators with the general transcription machinery, helping Pol II to focus on the start site

(Mantovani, 1998). Antibodies against NF-Y were used to demonstrate that NF-Y may not solely serve to orchestrate transcriptional initiation but is probably required to maintain basal transcription (Mantovani *et al*, 1992).

NF-Y has a strong preference for the –80 and –100 positions in relation to the transcription start site (Mantovani, 1998), which covers the region enclosing ICB1 and 2 of the topoisomerase IIα promoter. EMSAs show that NF-Y can bind to three of the proximal ICB (1-3) elements in topoisomerase IIα and that the extent of the NF-Y binding is greater at all three sites in extracts from proliferating cells than those from growth-arrested cells (Isaacs *et al*, 1996a). However, others have shown that NF-Y can interact with ICB1-4 (Coustry *et al*, 2001). In some cases, a decrease in NF-Y activity is observed when topoisomerase IIα is down-regulated. In doxorubicin resistant human multiple myeloma RPMI 8226 cells reduced levels of topoisomerase IIα mRNA has been attributed to a decrease in NF-Y interacting with the topoisomerase IIα promoter (Wang *et al*, 1997a). Consequently, NF-Y can not be the sole transcription factor involved in down-regulation of topoisomerase IIα, because confluence-arrested cells exhibit complete repression of topoisomerase IIα expression even though NF-Y is present (Isaacs *et al*, 1996a). In addition, the topoisomerase IIα promoter is not completely repressed in NF-Y depleted nuclear HeLa cell extracts *in vitro* (Coustry *et al*, 2001). Thus, other transcription factors must also be important in regulating topoisomerase IIα transcription and these may be cell-type specific.

1.5.3 GC Rich Regions.

The CCAAT box is invariably flanked by at least one functionally important promoter element, such as the GC boxes in the case of human topoisomerase IIα. GC boxes are the second most frequent element in promoters and tend to be present in multiple copies, where each GC box is proposed to have different functions. One such example are the GC boxes present in the human transcobalamin II promoter. This promoter has a proximal and distal GC box similar to the topoisomerase IIα promoter, where the proximal GC box of the human transcobalamin II promoter appears to be negative acting and the distal GC box positive acting (Li *et al*, 1998).

Human topoisomerase IIα promoter has two GC boxes; GC1 is contained in the defined minimal human topoisomerase IIα promoter whereas GC2 is not. Generally, mutations in

GC1 do not result in a decrease in topoisomerase IIα activity, however in resistant cells a mutation in GC1 can generate an increase in promoter activity (reviewed in Isaacs *et al*, 1998). GC2 is considered to have a lesser effect on topoisomerase IIα regulation and is thought to be functional under certain growth conditions (Hochhauser *et al*, 1992). Yoon *et al* (1999) have shown that in the rat topoisomerase IIα promoter Sp1 binds specifically to GC2 (equivalent to GC1 in the human topoisomerase IIα promoter) and that its binding activity increases after growth stimulation. Therefore, binding of Sp1 has been proposed to up-regulate gene expression of topoisomerase II α in growing rat cells.

1.5.4 Sp1 (Specificity Protein 1).

Sp1 was originally isolated from HeLa cells as a protein that bound to multiple GGGCGGG sequences (GC boxes) in the SV40 early promoter (Dynan and Tjian, 1983b). Since then studies have revealed that Sp1 can bind to GGGGCGGGGC and other closely related sequences (reviewed in Suske, 1999). Some promoters have multiple GC boxes, however it has been shown that a single Sp1 binding site is sufficient to stimulate promoter activity (Kadonaga *et al*, 1987). Purified Sp1 has been shown to require Zn (II) for sequence specific binding to DNA (Chen *et al*, 1994) and can bind to its target sequence in assembled nucleosomes (Li *et al*, 1994).

Sp1 is a typical Q-rich activator, commonly known as a transcriptional activator shown to be able to up-regulate transcription in a variety of promoters (Ge *et al*, 2001; Keates *et al*, 2001; reviewed in Suske, 1999). Sp1 mediated transcriptional regulation depends on three zinc finger structures responsible for DNA binding, and at least one of two Q-rich regions (activation motifs) that are required for transcriptional activation (Courey and Tjian, 1988). The Q-rich domain of Sp1 is thought to target one of the TFIID components (Chen *et al*, 1994), and consequently activate promoters that are transcribed by Pol II (Dynan and Tjian, 1983a; Kadonaga *et al*, 1986).

When two or more Sp1 sites are found in a promoter, synergistic activation can be observed (Kadonaga *et al*, 1987), which has been attributed to multiple intra and inter-molecular protein-protein interactions between different Q-rich domains. Su *et al* (1991) have shown that Sp1 can self-associate, which in turn can bring together distant DNA elements. This was demonstrated using the simian virus 40 (SV40) promoter and a modified thymidine kinase (TK) promoter. The TK promoter was modified so that it carried up-stream and down-stream

Sp1 binding sites. It was shown that the DNA directly looped *in vitro* and this correlated with synergistic activation of transcription *in vivo* due to the presence of Sp1. The looping phenomenon has been visualized using conventional and scanning electron microscopy (Mastrangelo *et al*, 1991). Sp1 initially forms a tetramer and subsequently assembles multiple tetramers stacked in register at the DNA loop junctions, where each monomer is capable of interacting with a single GC box. The formation of homo-multimeric structures was also demonstrated by Sp1-Sp1 two hybrid assays (Su *et al*, 1991). The formation of higher order complexes, in the absence of DNA-binding domains, is necessary for synergistic activation of Sp1 molecules bound at proximal and distal sites in the promoter (reviewed in Liberati *et al*, 1999). The fact that there are two GC boxes at either end of the topoisomerase IIα promoter may signify that DNA looping is required to regulate transcription.

Sp1 phosphorylation is thought to increase binding to the GC box and facilitate promoter activation for a number of genes. It has been demonstrated that when Sp1 is phosphorylated, it can bind DNA more tightly than dephosphorylated Sp1 (Ge *et al*, 2001).

1.5.5 Interaction between Sp1 and NF-Y.

Several studies have shown that parallels between NF-Y and Sp1 exist. Both have Q-rich domains and can interact with particular TAFs within TFIID (Coustry *et al*, 1998). In addition, elements that exhibit an affinity for NF-Y and Sp1 are often found nearby in promoters, however an interaction between NF-Y and Sp1 can also be detected in the absence of DNA (Roder *et al*, 1999). Hence, it is no surprise that NF-Y and Sp1 have been implicated to interact on a variety of promoters, such as the promoters of the following genes; rat fatty acid synthase (Roder *et al*, 1999), human multidrug resistance gene (Sundseth *et al*, 1997), human cystathionine-β-synthase (CBS) (Ge *et al*, 2001), major histocompatibility complex (MHC) class II-associated invariant chain (Ii) gene (Wright *et al*, 1995) and many others (Inoue *et al*, 1999). Several lines of study have been exploited to observe an interaction between NF-Y and Sp1, including the two-hybrid system, EMSA, co-immunoprecipitation and co-transfections.

Roder *et al* (1999) have demonstrated an *in vivo* interaction between NF-Y and Sp1 using the yeast two-hybrid system, where the domains required for the interaction were located between amino acids 55 and 139 of NF-YA and amino acids 139 and 344 of Sp1; the Q-rich domains. In the case of the MHC II promoter, both NF-Y and Sp1 elements were required to

obtain full activation, as NF-Y mediated transcriptional activation was only observed when Sp1 was bound to DNA nearby (Wright *et al*, 1995). A CBS reporter gene was co-transfected with NF-Y and/or Sp1 in *Drosophila* Sl2 cells. *Drosophila* Sl2 cells were an excellent cell line to use as they provide a null background for Sp family of transcription factors and NF-Y, which the CBS promoter requires for activation. Their experiments demonstrated a significant increase in CBS promoter activity when both NF-Y and Sp1 were present compared with Sp1 alone, indicating synergism may exist between NF-Y and Sp1 to regulate transcription (Ge *et al*, 2001). The same synergism was observed when a similar experiment was carried out using the rat pyruvate kinase M (PKM) promoter (Yamada *et al*, 2000). Therefore, it has been suggested that NF-Y could be stabilized by the presence of adjacently bound Sp1, as the CCAAT box binds NF-Y and activates transcription only when it is stabilized by a neighboring factor, Sp1.

1.5.6 Sp3.

Sp1 is member of the Sp family of small proteins, consisting of; Sp1, Sp2, Sp3 and Sp4 all of which have similar structures. Sp3 is similar to Sp1 in many ways; it is ubiquitously expressed in mammalian cells, can be phosphorylated to increase DNA binding affinity (Ge *et al*, 2001) and can compete with the same target sequences with similar binding affinities due to their similar DNA binding domains. However, Sp3 is different from Sp1 in that it is bifunctional. Sp3 is able to activate or repress transcription (Noti, 1997). The repressive function of Sp3 has been attributed to a region near the N-terminus, which acts as a transcriptional repressor of several activators (de Luca *et al*, 1996). Three isoforms of Sp3 exist, due to three internal translation initiation sites within the mRNA and this has made examining its transcriptional role difficult. It is thought that promoters that contain a single binding site tend to be activated by Sp3, but promoters with multiple binding sites respond weakly to Sp3 or not at all (reviewed in Suske, 1999). The relative levels of Sp1 and Sp3 have been implicated in the response of a promoter to these transcription factors. If Sp3 is expressed at the same level as Sp1, Sp3 may not act as an activator as it can compete for binding with Sp1. This is consistent with investigations that suggest that Sp3 represses Sp1-mediated transcription (de Luca *et al*, 1996). Thus, Sp3 activity appears to be dependent on the cell and promoter context (Majello *et al*, 1997).

It has been demonstrated that Sp3 is up-regulated in etoposide/teniposide-resistant human cancer cells (KB/VP-2 and KB/VM-4 respectively), which correlates with a decrease in

topoisomerase IIα gene expression (Kubo *et al*, 1995). In a separate study, Sp3 was shown to interact with NF-YA and synergistically alter gene expression (Yamada *et al*, 2000). Therefore, Sp3 may function as a repressor by protein-protein interaction with components of the general transcription complex or by inhibiting other transcription factors. Interestingly, when the repressor region of Sp3 was tethered to nascent RNA it also inhibited translation of the RNA (de Luca *et al*, 1996) indicating repression does not always require an interaction with the promoter.

1.6 Topoisomerase IIβ.

Topoisomerase IIβ has been shown to be a dispensable protein, but it is required for normal mammalian development. Cultured rodent cell lines that lack topoisomerase IIβ are viable, but mice that carry a homozygous targeted deletion exhibit neural development defects resulting in pups that die at birth or shortly after (Yang *et al*, 2000). The gene for human topoisomerase IIβ is situated on chromosome 3 at 3p24 (Tan *et al*, 1992) and is comprised of 36 exons (Lang *et al*, 1998), to generate a protein 180 kDa in size.

Topoisomerase IIβ is ubiquitously expressed, independent of the cell-cycle unlike topoisomerase IIα, which is cell-cycle regulated and found mainly in proliferating tissues (Isaacs *et al*, 1995). Fundamentally, topoisomerase IIβ appears to be localized to different compartments during cell-cycle progression. During interphase, topoisomerase IIβ has been found to be predominantly nucleoplasmic (Adachi *et al*, 1997; Meyer *et al*, 1997), but a proportion has also been found in the nucleoli (Cowell *et al*, 1998). At mitosis the level of topoisomerase IIβ appears to increase but the enzyme was detected less in mitotic nucleus (Adachi *et al*, 1997) and shown to diffuse into the cytosol. The bulk of topoisomerase IIβ is not detected in any condensed chromatin during mitosis, even though it is capable of interacting with the condensed metaphase chromatin (Meyer *et al*, 1997). Whereas, topoisomerase IIα is observed to bind to the chromosome scaffold, where it is thought to be required for mitotic chromosome condensation and disjunction during metaphase (Grue *et al*, 1998). Due to the different localization patterns of each topoisomerase II isoform, it is suggested that even though their catalytic mechanisms are similar, they have different functions during the cell-cycle. Thus, these patterns suggest that topoisomerase IIβ may not have a role in mitotic chromosomes possibly due to altered catalytic activity, however it may be required for a minor role in chromosome condensation and disjunction as it can bind to chromosomes and has been found in nucleoli (Kimura *et al*, 1996).

The topoisomerase IIβ minimal promoter was isolated from a human lung fibroblast genomic DNA library by a single group of scientists (Ng *et al*, 1997). The minimal promoter was proposed to be within nucleotides –14 and +326. The published sequence has no TATA, nor any CCAAT boxes (refer to figure 1.5), but it is rich in GC content with two perfect GC boxes located at position +285 and +315. There is a putative CTF/NF1 site mapped to the reverse of nucleotides +267 to +280, which is thought to be important for transcriptional

regulation as CTF/NFI has been shown to interact directly with TBP (Xiao *et al*, 1994). Finally, there is a putative AP2 binding site at the sequence CCGCGGGC, which is the reverse strand of nucleotides +305 to +312. EMSAs demonstrated that protein factors can bind to downstream elements of the topoisomerase IIβ promoter, however any investigations of upstream factors were not reported. Therefore, the functional significance of these putative sequence elements remains to be determined.

Figure 1.5: The published 5'-flanking sequence of topoisomerase IIβ.

The ATG codon for the start of exon 1 (underlined) is in red. The two GC boxes are shown in blue. The major transcription start site is indicated by the arrow (+1). The putative AP2 site is complementary to the sequence shown in orange. The consensus CTF/NF1 site maps to the complementary sequence of that shown in pink (Ng *et al*, 1997).

The β isoform differs from the α isoform in many ways, however relatively little work has been carried out on the β isoform as it was discovered after topoisomerase IIα. In addition to having different functional roles, the damaged induced by drugs acting on the two isoforms may be sensed differently within the cell inducing different signal transduction pathways and cell-cycle checkpoints (Gao *et al*, 1999). Recent advances have enabled the cellular quantification of topoisomerase IIα and topoisomerase IIβ. Levels of topoisomerase IIβ exceed those of topoisomerase IIα in plateau phase cells (Padget *et al*, 2000) and the levels of topoisomerase IIβ are comparable to topoisomerase IIα even in proliferating cells, which contradicts previous studies of these isoforms. This is significant because it indicates that the

β isoform is present in notable levels, which may be important when investigating the action of anti-cancer drugs. Consequently, it has been shown that the proportion of tumor cells expressing topoisomerase IIβ at a high level is greater than that of cells expressing topoisomerase IIα (Robert and Larsen, 1998). These observations demonstrate that proliferation alone is unlikely to be the basis for the up-regulation of topoisomerase IIβ, which may prove useful when developing and administering topoisomerase II-interacting drugs.

Some anti-cancer drugs have been found to preferentially target topoisomerase IIβ, such as XK469 (Gao *et al*, 1999). This synthetic drug has an unusual selectivity for solid tumors and activity against multidrug-resistant cancer cells. It has been suggested that the difference in cell-cycle regulation of topoisomerase IIα and topoisomerase IIβ may be the key to this selectivity. Solid tumors, unlike leukemias, often have large populations of cells in the G_1/G_o phase of the cell-cycle in which topoisomerase IIβ expression is high whereas topoisomerase IIα expression is low. However, the presumptive primary target of many leukemia selective drugs is topoisomerase IIα. This suggests there may be an advantage to using topoisomerase IIβ as a target, as it is ubiquitously expressed at higher levels than topoisomerase IIα. Even though most cases of drug-resistance is attributed to changes in topoisomerase IIα expression, topoisomerase IIβ should not be excluded. It has been shown that a Chinese hamster lung cell line (DC-3F/9-OH-E) made resistant to 9-OH-ellipticine and cross-resistant to other topoisomerase II inhibitors has no topoisomerase IIβ expression and this is due to a mutation in the topoisomerase IIβ gene, at position 1710 in the cDNA (Dereuddre *et al*, 1997). Therefore, loss of topoisomerase IIβ activity is also thought to be in part responsible for the development of resistance to topoisomerase II inhibitors.

While the two topoisomerase II isoforms may vary in expression, relative drug sensitivities and localization, their differences in expression patterns could be utilized to create isoform-specific drugs. In order to do so, a detailed knowledge of how topoisomerase IIα and topoisomerase IIβ expression is required. This can be achieved by examining the promoter region in a series of DNA-binding and functional assays.

1.7 Research Aims.

Both topoisomerase II isoforms are important as targets for many anti-cancer drugs. A major problem with understanding the mechanism by which topoisomerase IIβ acts is that little is known about its regulation and expression. Because a cloned promoter construct for topoisomerase IIβ was not available, the initial aim of this research was to isolate genomic DNA upstream of exon I of the topoisomerase IIβ gene and carry out a systematic analysis of this region. This would have enabled a direct comparison of the mechanisms that regulate topoisomerase IIα and β expression. Several putative topoisomerase IIβ promoter clones were obtained but sequence analysis revealed that none of them contained the upstream regulatory region. It may have been possible to isolate the upstream region of topoisomerase IIβ from a human genomic library but such a library was not readily available at the time this was carried out. Since the publication of the human genome sequence was also imminent when the project began, it was not considered worthwhile to pursue the topoisomerase IIβ promoter at this time. Hence, a study of topoisomerase IIα promoter was carried out.

Compared to topoisomerase IIβ, the regulation of topoisomerase IIα is relatively well established although there is no definite mechanism that can be proposed to account for the drug-resistance that develops. In order to understand this phenomenon, an enhanced understanding of the relationship between DNA consensus elements and the cognate DNA-binding proteins of the topoisomerase IIα promoter is required. To this end, the regulation of topoisomerase IIα expression was studied by focusing on the two most proximal promoter elements, ICB1 and GC1 and the most distal element, GC2.

The specific objectives of this research were as follows:

- Carry out a series of DNA-binding assays using oligonucleotides representing elements GC1 and ICB1 of the topoisomerase IIα promoter.
- Generate a series of mutant promoter constructs containing single, double or triple mutations in topoisomerase IIα promoter elements GC2, GC1 and ICB1.
- Analyze these topoisomerase IIα promoter constructs for transcription in the presence and absence of Sp1 or Sp3.

Chapter 2: Materials and methods.

2.1 Materials.

Restriction endonucleases and buffers, DNA modifying enzymes, buffers, calf alkaline phosphatase (1 unit/μL), herring sperm DNA, nitrocellulose filter, DNA quantification standards and agarose LE powder for electrophoresis were from a number of different sources: New England Biolabs, MA, USA; Boehringer Mannheim, Germany and Roche, Mt. Wellington, Auckland.

DMSO, DTT, IPTG, TEMED, PIPES, EDTA, X-gal, PEG, dNTPs, triton-X-100, ampicillin, lysozyme, SDS, BSA, mineral oil, ficoll, oligonucleotides for gel shifts, primers RV3, GL2 and topoisomerase IIβ primers were purchased from Sigma Chemical Company, St Louis, MO, USA.

CONCERT™ gel cleaning kit and maxi preparation kit, 10x trypsin, penicillin-streptomycin (5,000 units/mL), MEM with Earle's salts (L-glutamine and non-essential amino acids), bacteriological agar, 1kb plus ladder, Luria Bertani (LB) broth base, foetal calf serum were purchased from GIBCOBRL, Invitrogen Corporation, Invitrogen NZ limited, Penrose, Auckland, New Zealand.

PCR purification kit QIAquick was purchased from QIAGEN, New Zealand distributors: Biolab Scientific Ltd, Albany, Auckland, New Zealand.

Bradford protein assay kit and Quantum® Prep Plasmid Miniprep Kit were purchased from BioRad Laboratories, CA, USA.

40% (w/v) solution of acrylamide and NN'-methylenebisacrylamide in water (Acrylogel) was from BDH Laboratory Supplies, Poole, England.

α^{32}P-[dCTP] and γ^{32}P-[ATP] were purchased from Perkin Elmer Life Sciences Inc, Boston, MA, USA.

All primary antibodies (NF-YA, NF-YB, Sp1 and Sp3) were purchased from Santa Cruz Biotechnology, CA, USA.

GCG Version 9.1 was purchased from Wisconsin Genetics Computer Group, USA. FLUOstar galaxy system and software was purchased from BMG Labtechnologies Pty. Ltd, Melbourne, Australia.

DE-81 and 3MM paper were purchased from Whatman, Maidstone, England.

Poly (dIdC), Ready to Go DNA Labelling Beads (-dCTP) and ProbeQuant™ G-50 Micro Columns were purchased from Amersham Pharmacia Biotech AB, Uppsala, Sweden.

Plasmids pGL3Basic, pSVβ-galactosidase, pGEM®-T vector system, Packagene® Lambda DNA Packaging System and the Luciferase Assay System were purchased from Promega Corporation, WI, USA.

The *Escherichia coli* XL-1 blue strain and LE392 strain were purchased from Stratagene, La Jolla, CA, USA.

All sterile tissue culture flasks, plates, cell scrapers and tubes were purchased from Nunc Inc, Naperville, IL, USA.

0.8 μM sterilization double layer filters were purchased from Drummond, USA.

FuGENE™ 6 transfection reagent, ONPG, Complete™ Mini EDTA-free Protease inhibitor cocktail tablets, T4 ligase and *Taq* polymerase were purchased from Roche Molecular Biochemicals, IN, USA.

Topoisomerase IIβ in Bluescript and Topoisomerase IIβ in lambda T2-E4 phage were gifts from Sally Davies (ICRF, Oxford, UK). Original HeLa cells were a gift from Dr. Rachel Page (Department of Biochemistry, University of Cardiff). The various –617 topoisomerase IIα promoters in pGL3Basic vector were generated by Samuel M^{c}Lenachan and Agnieszka Szremska (Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand).

All other chemicals and reagents used were of analytical grade or better.

2.2 Methods.

2.2.1 Agarose Gel Electrophoresis.

Agarose gel electrophoresis is a method that can rapidly resolve mixtures of nucleic acid molecules giving information about the size and nature of DNA (deoxyribose nucleic acid), also providing some level of purification if required. In an electrical field and near neutrality (pH 8.0), linear DNA is negatively charged and will migrate towards the positive electrode, where mobility is primarily dependent on the DNA fragment size.

Generally, electrophoresis was carried out using 1% agarose in 1xTAE (40 mM Tris-acetate, 2 mM ethylene diamine tetra-acetic acid (EDTA), pH 8.0) containing ethidium bromide (0.5 µg/mL agarose). Ethidium bromide forms a complex with DNA, which absorbs UV irradiation at 260 nm and fluorescence can be detected. Thus, enabling visualization of DNA within the gel by exposure to UV light. The agarose gel was melted in a microwave and cooled to 55°C in a water bath, then poured into a gel-casting tray. Occasionally when fragments were expected to be very large or very small, a different percentage of agarose was used (refer to table 2.1).

Percentage Agarose	Linear DNA (kb)
0.3	60 - 5.0
0.6	20 - 1.0
0.7	10 - 0.8
0.9	7.0 - 0.5
1.2	6.0 - 0.4
1.5	4.0 - 0.2
2.0	3.0 - 0.1

Table 2.1: Percentage agarose for optimal separation of linear DNA (kb).

This table was used as a guide to estimate the percentage agarose used in separation of DNA fragments (Berger and Kimmel, 1987).

DNA samples were loaded into agarose wells by mixing with approximately 10% loading dye (40% w/v sucrose, 0.25% bromophenol blue). The loading dye results in the DNA

solution being heavier than the buffer causing it to sink to the bottom of the wells. The bromophenol blue enables the visualization of the travel front. Electrophoresis was carried out at 80 V for approximately 1 hour in 1xTAE buffer and DNA visualized by exposure to UV light, using a UV transilluminator. The sizes of DNA bands were determined by comparing them to a known DNA size standard marker (1 kb plus ladder, Invitrogen).

2.2.2 Oligonucleotides.

Oligonucleotides were synthesized by the Sigma Chemical Company and provided as a dried stock. Each was rehydrated in TE buffer (10 mM Tris-HCl pH 8, 0.1 mM EDTA) to a concentration of 10 μg/μL and stored at -20°C. Oligonucleotides were further diluted when required in TE buffer to a concentration of 50 ng/μL for PCR reactions, 100 ng/μL for electrophoretic mobility shift assays and 1 μg/μL for competitor oligonucleotides.

2.2.3 DNA Purification by Gel Electrophoresis.

Purification of DNA fragments was achieved by using agarose gel electrophoresis. The required DNA was excised from the gel and extracted using CONCERT™ gel purification kit (GIBCOBRL, Invitrogen) as described by the manufacturers. This kit uses a system of spin cartridges containing a silica membrane. The agarose gel is dissolved by sodium perchlorate and DNA adheres to the silica support, DNA is released upon elution with TE buffer after washes with ethanol-based buffers.

2.2.4 Phenol-Chloroform Extraction and Ethanol Precipitation.

Phenol-chloroform extraction was generally used to purify DNA from contaminants in solution. Contaminants could be RNA or proteins. When chloroform is used in conjunction with phenol it improves the efficiency of nucleic acid extractions, due to its ability to denature proteins, remove lipids and enhance separation of the two phases (due to its high density). Typically, the DNA solution was purified by sequential extractions with equal volumes of phenol, phenol:chloroform and chloroform as described by Berger and Kimmel (1987).

Phenol-Chloroform extraction was commonly followed by ethanol precipitation to concentrate DNA in solution. A precipitate was formed by mixing the DNA sample with salt (1/10 volume of 3 M sodium acetate at pH 4.8) and ethanol (2 ½ volumes of 95% ice cold ethanol) at low temperatures. DNA was pelleted by centrifugation at 12,000 rpm and the

pellet was dried and resuspended in (usually 25 μL) TE or sterile H_2O (Sambrook and Russell, 2001). Samples that contained EDTA or phosphate were not ethanol precipitated, as these materials will also precipitate out with the DNA and affect purification.

2.2.5 DNA Quantification.

(i) Quantification using DNA Standards and Gel Electrophoresis.

The concentration of many DNA samples was estimated using gel electrophoresis. 1 μL of the query DNA (or varying dilutions in TE, 1:5, 1:10, 1:100) with 10% loading dye was loaded on a 1% agarose gel (incorporated with 0.5 μg ethidium bromide per mL of agarose), against 5 μL of each quantification standard (5 ng/5 μL, 10 ng/5 μL, 20 ng/5 μL, 50 ng/5 μL and 100 ng/5 μL) and electrophoresed for about 1 hour in 1xTAE at 80V. DNA was visualized by exposure to UV light and the intensity of fluorescence omitted from the query DNA sample was compared to that of the known standards thereby indicating what concentration the query DNA possessed. The DNA standards were generated from linearized plasmid DNA (pBluescript SK-II).

(ii) Quantification using UV Spectrophotometry.

The plasmids used for transfections were quantified using UV spectrophotometry. At a wavelength of 260 nm the absorbance of double-stranded DNA (50 μg/mL) generates an absorbance of 1.0. Plasmid DNA was diluted 1:100 in TE and 1 mL was placed into quartz cuvettes. Cuvettes were placed into the Ultraspec 300 UV/Visible Spectrophotometer (Pharmacia Biotech, Cambridge, England) and absorbances were measured using the Nucleic Acid Scan programme. The Nucleic Acid Scan programme records the absorbance of a sample at intervals (200 nm-350 nm), which enabled the concentration of DNA to be calculated, using the relationship: A_{260} x 50 μg/mL x dilution factor. The purity of DNA could also be analyzed by calculating the ratio of A_{260}/A_{280}. A ratio of 1.8 indicates pure DNA; a ratio is less than 1.8 is due to protein contamination and a ratio greater than 1.8 indicates RNA (ribose nucleic acid) contamination.

2.2.6 Restriction Endonuclease Digests.

Restriction endonucleases were used to characterise plasmid DNA or prepare DNA for cloning. Generally 500-1000 ng of DNA was digested with 5-10 units of restriction endonuclease along with 1x of the appropriate buffer (as specified by the manufacturer).

For digesting plasmid DNA prepared by the rapid boil method, 10 μL of DNA was used with 1 μL (8-10 units) of restriction enzyme along with 3 μL of appropriate (10x) buffer (as specified by the manufacturer) in a total volume of 30 μL. When large volumes of DNA were required, the total volume of the digestion mixture was made up to 50 μL. Each digest was incubated at 37°C for 1-2 hours and samples were analyzed for completeness of digestion using agarose gel electrophoresis. Human genomic DNA (50 ng/μL) digests and lambda DNA preparations (approximately 20 ng/μL) were incubated overnight at 37°C, to facilitate complete digestion. Rapid boil plasmid DNA samples were also treated with 2 μL of RNase (10 mg/mL) for 2 minutes at 37°C after undergoing restriction endonuclease digestion. For digesting plasmid DNA prepared using the Quantum® Prep Plasmid Miniprep (BioRad) method (30-50 ng/μL), 5-10 μL of DNA was digested depending on the concentration of each preparation. Only 1 μL of plasmid DNA (~1 μg/μL) prepared by the CONCERT Rapid Plasmid Maxiprep Kit (GIBCOBRL, Invitrogen) was used in restriction digests (in 30 μL total volume) due to its' high concentration and RNase treatment was not necessary.

2.2.7 Alkaline Phosphatase Treatment of Vectors.

Alkaline phosphatase treatment was necessary when setting up ligations. It is important to remove 5'-phosphates from the vector DNA ends after restriction endonuclease digestion to prevent the occurrence of the vector self-ligating.

1μl of calf alkaline phosphatase (1 unit/μL, Roche) was added to 1-5 μg of digested vector and incubated for exactly 15 minutes at 37°C. This mixture was never left for any longer than 20 minutes as the calf alkaline phosphatase is impure. After incubation 2 μL of proteinase K (10 mg/mL) and 2 μL of 20% sodium dodecyl sulfate (SDS) was added, mixed and incubated for 1 hour at 37°C to inactivate the phosphatase. The vector DNA was then phenol-chloroform extracted, ethanol precipitated and quantified using agarose gel electrophoresis and quantification standards.

2.2.8 Sequencing of DNA.

DNA sequencing was carried out by Lorraine Berry (MUSeq facility, Institute of Molecular BioSciences, Massey University) using dideoxy cycle sequencing with Big Dye terminators (Version 1.0, PE Biosystems, Foster City, CA, USA) and analyzed on an ABI 377-18/36 automated DNA sequencer according to the manufacturers' instructions. Primers and

templates were diluted to 0.8 picomol/μL and 200 ng/μL respectively (in TE buffer) for sequencing.

2.2.9 Primer Walking.

Primer walking was used to sequence larger DNA fragments (500+ bp), as one sequencing reaction usually generates only about 500 bp of sequence. This technique is only useful if a start point for sequencing is known, such as a primer site within a vector. The initial primer is used to sequence the first region of the DNA template, and then primers are designed consecutively according to the sequences obtained. Each individual sequence is then organized according to the overlapping regions of DNA between primers, thereby generating a contiguous DNA sequence.

Each custom primer was designed to be 18-19 bases in size and approximately 50% G/C, as primers that have too high an A/T composition tend to prime ineffectively (Berger and Kimmel, 1987). Consequently, primers that are high in G/C content should also be avoided, as templates that have high G/C regions may cause inaccurate sequencing reactions due to primers adhering to multiple sites.

2.2.10 Contig Assembly using GCG.

Short overlapping DNA sequences can be organized into a longer sequence (contig) using the fragment assemble system (FAS) within GCG (Version 9.1; Wisconsin Genetics Computer Group, USA). Individual programs within FAS have to be activated sequentially to successfully create a contig.

Firstly, a directory was created within the program using the command line 'Gelstart'. This command initializes the FAS and enables the GCG network to recognize and build individual projects. Then each short sequence was entered individually into the directory using the 'Gelenter' command, enabling each sequence to be stored within the created directory. Each of the sequences were then assembled into contigs using the 'Gelmerge' program. Gelmerge automatically recognizes overlapping sequences between individual fragments and creates aligned assemblies. Finally, the actual contig sequence was retrieved from the 'Consensus' directory, which is within the directory that was created earlier. In order to view a schematic representation of the organization of the shorter sequences within the generated contig, the 'Gelview' program was used.

2.2.11 Analysis of DNA Sequences.

Generally, DNA sequences were analyzed using sequence analysis software provided by GCG (Version 9.1; Wisconsin Genetics Computer Group, USA). Various programs within this suite of software were used to accomplish particular tasks, such as: restriction maps (map), generate contigs of individual sequences (GelMerge), find sequence similarity (bestfit), execute database searches (stringsearch) and find putative binding factors by matching consensus sequences within a sequence (findpatterns).

Database searches were carried out to find any putative transcription factor consensus sequences. The TRANSFAC database relies on a compilation of information about *cis*-regulatory sequences and *trans*-acting factors. This database is intrinsically linked to other databases such as the TRRD (transcription regulatory region database), which enables an extensive database search (*in silico*) on transcription factors, their binding sites and DNA-binding profiles (Quandt *et al*, 1995; Wingender *et al*, 2000). The TRANSFAC database (MatInspector V2.2, http://transfac.gbf.de/cgi-bin/matSearch.pl) was used and one occasion, the software package BioNavigator™ (ENTIGEN, Sunnyvale, CA, USA, www.BioNavigator.com) was used to analyze sequences. The outputs were analyzed manually for any relevant sequence motifs.

2.2.12 General PCR Reactions.

Each PCR reaction contained 5 µL (50 ng/µL) of each primer, 5 µL (3 mM) dNTPs (a mixture of dATP, dTTP, dCTP and dGTP), 5 µL 10 x PCR buffer (100 mM Tris-HCl, 500 mM KCl pH 8.3, Roche), 0.25 µL *Taq* polymerase (5 U/µL, Roche), 1 µL template (approximately 1 ng/µL) and dH_2O to make the total volume up to 50 µL. The *Taq* polymerase was always added last, and reactions were mixed thoroughly before a drop of mineral oil was added. PCR reactions were carried out in a thermal cycler (HYBAID Omn-E, SciTech (NZ) Ltd) programmed as shown below. A negative control (no DNA template) was also included with each PCR reaction to check for the possibility of contamination.

Thermocycling conditions for PCR

95°C for 5 minutes.		
Denaturing	95°C for 1 minute.	30x cycles.
Annealing	45°C for 1 minute.	
Extension	72°C for 1 minute.	

10% of each PCR reaction was analyzed for completion by agarose gel electrophoresis.

2.2.13 PCR Purification.

PCR products were purified before they could be used as a probe or in a ligation, i.e. the oil, primers, salts and excess dNTPs had to be removed. This was achieved using QIAquick PCR purification kit (QIAGEN). This system utilizes a silica-based spin cartridge to purify double-stranded DNA. PCR products were purified following the manufacturers instructions and then quantified using agarose gel electrophoresis.

2.2.14 Cloning PCR products into pGEM®-T.

PCR products were cloned into pGEM®-T vectors (Promega) to enable the production of amplified insert DNA without the need for further PCR. pGEM-T vector kits are equipped with 2x rapid ligation buffer (60 mM Tris-HCl pH 7.8, 20 mM $MgCl_2$, 20 mM dithiothreitol (DTT), 2 mM adenosine triphosphate (ATP) and 10% polyethylene glycol (PEG)) and control insert DNA. The vector is already linearized (digested with *Eco* RV) and contains 3' terminal thymidines attached to each cut end (T-tailed); this is to enable efficient ligation of PCR products into the plasmids and preventing the recircularization of the vector.

PCR products were subcloned into pGEM-T, following the manufacturers instructions. However, ligations were carried out using half the recommended reaction, thus 25 ng control vector was used with a 1:1 (vector:insert) molar ratio of control DNA and a 3:1 (insert:vector) molar ratio of PCR DNA in a total volume of 10 µL. Ligation mixes were incubated overnight at 4°C and transformation of *Escherichia coli (E.coli)* XL-1 blue cells performed the following day, utilizing blue/white selection (Berger and Kimmel, 1987).

2.2.15 Ligations.

In order for the ligation to be successful, the DNA fragments must have compatible ends. Therefore, insert and vector DNA were digested with the same restriction endonuclease to generate compatible ends. Ligation efficiency is dependent on the concentration of DNA ends in the reaction. To ensure complementary ends of vector and insert are joined, the concentration of insert must be higher than that of the vector (Berger and Kimmel, 1987).

The amount of vector DNA used in ligations was 50 ng, to which the insert DNA was added at either a 1:1, 2:1 or 3:1 (insert:vector) molar ratio of vector. The amount of insert DNA that was added to ligations was calculated using the following formula.

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

Ligation mixes also contained 4 µl (5x) T4 ligase buffer (250 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25% w/v PEG), 1 µL T4 ligase (1 U/µL, Roche) and dH_2O to make the final volume up to 20 µL. Ligation mixes were prepared, mixed thoroughly and incubated overnight at 16°C. The following day, transformation of *E.coli* XL-1 blue competent cells was attempted.

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

E.coli ultra-competent XL-1 blue cells were prepared by Carole Flyger (Institute of Molecular BioSciences, Massey University) as follows.

300 mL of SOB broth (20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 2.5 mL 1 M KCl all autoclaved, then 5 mL 2 M $MgCl_2$ added just before use) was inoculated with 5 mL of an overnight culture of XL-1 cells and the culture was grown to A_{600} of 0.6 at 16°C. The culture was chilled on ice for 10 minutes and pelleted in a refrigerated centrifuge at 4,000 rpm for 10 minutes. The pellet was resuspended in 4 mL ice cold TB buffer (10 mM piperazine-N, N'-bis[2-ethanesulphonic acid] (PIPES), 15 mM $CaCl_2$, 250 mM KCl pH 6.7, 55 mM $MnCl_2.4H_2O$) and centrifuged again at 4,000 rpm for 10 minutes. The pellet was washed in 4 mL TB and centrifuged again at 4,000 rpm for 10 minutes, supernatant removed and resuspended in 4 mL ice cold TB buffer and then 1-2 mL of cold dimethyl sulfoxide (DMSO) was added drop-wise to the cells. The ultra-competent XL-1 cells were then aliquoted into 500 µL volumes, snap frozen using liquid nitrogen and stored at -70°C.

Typically, half of a ligation mix (10 µL) was added to 100 µL of XL-1 blue competent cells and incubated on ice for 15 minutes. The samples were then diluted appropriately (10^2 and 10^3) in LB broth (20 mg/mL) and 50 µl of each dilution was plated onto LB agar plates (15 g/L bacteriological agar, 20 g/L LB broth mix; 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) containing ampicillin (100 mg/mL). Plates were incubated overnight at 37°C, where colonies were scored the following day. Ampicillin was added to plates as a selection control, for plasmids that confer resistance. XL-1 blue cells also have the capacity for blue/white selection; to achieve this transformants were grown on LB plus amp plates supplemented

with 0.5 mM isopropyl thiogalactoside (IPTG) and 80 μg/mL X- gal (5-bromo-4-chloro-3-indoyl-β-D-galactoside).

Blue/white selection was utilized when selecting for pGEM-T transformants. Colonies that appeared blue were those with empty plasmids and colonies whitish in colour carry plasmids with an insert. This occurs because the *lacZ'* gene was disrupted and α-complementation could not occur. Blue/white selection could not be utilized for transformation using the plasmid pGL3B, as these vectors do not contain a *lacZ'* gene.

2.2.17 Isolation of Plasmid DNA from E.coli.

All bacteriological media was sterilized by autoclaving before use and aseptic techniques utilized. 5 mL overnight cultures were prepared by inoculating 5 mL of Luria Bertani (LB) broth plus 5 μL ampicillin (100 mg/mL) with one single colony from LB plus ampicillin agar plates and grown overnight at 37°C with vigorous shaking to enable sufficient aeration for growth. Different plasmid isolation techniques were used as they each give different DNA qualities and quantities after each preparation.

(i) Rapid Boil Plasmid Preparations.

Rapid boil plasmid preparations (Holmes and Quigley, 1981) generate impure DNA extractions, but this method is sufficient to screen XL-1 blue colonies for plasmid DNA. 1.5 mL of a 5 mL overnight culture was pelleted by centrifugation at 12,000 rpm for 1 minute. The supernatant was removed and the pellet was resuspended in 350 μL of STET (8% (w/v) sucrose, 50 mM EDTA, 50 mM Tris pH 8.0, 5% (v/v) Triton X-100) with 25 μL of freshly prepared lysozyme (10 mg/mL). The tubes were then placed into a boiling water bath for exactly 40 seconds, followed by a 10 minute centrifugation at 12,000 rpm. The gelatinous pellet was then removed and an equal volume of isopropanol was added to the supernatant, which was mixed and placed at -70°C for at least 30 minutes, to precipitate the DNA. The DNA was pelleted by centrifugation for 20 minutes at 12,000 rpm at 4°C. The supernatant was removed by aspiration and the pellet washed with 500 μL of ice cold 95% ethanol, and centrifuged again for 10 minutes (12,000 rpm at 4°C). Finally, all the supernatant was removed and the pellet dried and resuspended in 50 μL TE. Samples were stored at -20°C until required.

(ii) Small-scale Plasmid DNA Preparations.

Small-scale plasmid DNA preparation was used when a small quantity of high quality plasmid DNA was required for PCR or sequencing. Generally, 2 mL of a 5 mL overnight culture was pelleted and plasmid DNA isolated using a Quantum® Prep Plasmid Miniprep Kit (BioRad) according to the manufacturers instructions. This system uses an alkaline lysis (containing SDS) method (Birnboim and Doly, 1979) to release plasmid DNA from the cell. Purification of the plasmid DNA is achieved by a patented Quantum prep matrix, which contains a diatomaceous earth that binds DNA.

(iii) Large-scale Plasmid DNA Preparations.

Large-scale plasmid DNA preparations were carried out to obtain high quality and quantity plasmid DNA required for transfections. The required cultures were prepared by inoculating 500 mL LB plus 500 µL ampicillin (100 mg/mL) with 5 mL of a fresh overnight culture, and grown overnight at 37°C with shaking. Plasmid DNA was isolated using CONCERT™ High Purity Plasmid Maxiprep Kit (GIBCOBRL®, Invitrogen) according to the manufacturers instructions. This system is based on a modified alkaline lysis method (Birnboim and Doly, 1979), where plasmid DNA is purified by anion exchange resin.

2.2.18 Glycerol stocks of transformed XL-1 blue cells.

Glycerol stocks were made of selected transformed XL-1 blue lines by mixing 500 µL of an overnight culture (5 mL LB broth plus 5 µL (100 mg/mL) ampicillin) with 500 µL sterile 40% glycerol in 1 mL sterile cryotubes (Nunc) and storing at -70°C.

2.2.19 Lambda (λ) phage manipulations.

Bacteriophage λ infects *E.coli* cells, and is released as virions (λ DNA packaged into viral heads), which can then be used to isolate replicated DNA.

(i) Ligating the λ DNA for packaging.

Firstly, the λ phage DNA must be concatamerised before it can be packaged. This was done by ligation with enzyme T4 ligase. Approximately 100 ng of each λ clone was combined with 1 µL T4 ligase (Roche), 2 µL (5x) T4 ligation buffer (250 mM Tris-HCl pH 7.6, 50 mM $MgCl_2$, 5 mM ATP, 5 mM DTT, 25% w/v PEG) and enough dH_2O to make the final volume up to 10 µL. Each mix was incubated overnight at 16°C to ensure DNA was ligated.

(ii) Packaging ligated λ DNA.

In order for the bacteriophage λ DNA to be infectious, it must be packaged into viral capsids (virions). 5 μL of ligated λ DNA or 5 μL (0.25 μg) of control λ DNA (supplied in the kit) was packaged using 25 μL of packaging extract (Packagene® Lambda DNA Packaging System, Promega), following the manufacturers instructions.

After packaging, each sample was stored at 4°C with 25 μL chloroform and 445 μL phage buffer (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 10 mM $MgSO_4$). For long-term storage, 200 μL packaged λ DNA was added to 2 μL (1%) gelatin and 14 μL (7% v/v) DMSO then aliquoted into separate tubes and stored at –70°C. The packaging efficiency was calculated using the following formula:

$$\text{Packaging efficiency (recombinants/μg DNA packaged)} = \frac{\text{\# of plaques} \times \text{dilution factor / volume plated (μL)} \times \text{total volume (475μL)}}{\text{Amount of DNA packaged (μg)}}$$

(iii) Preparing LE392 (E.coli) plating cells.

Bacteriophage λ adsorbs to receptors in the outer membrane of *E.coli* that are encoded by the *lamB* gene (maltose-binding protein). These receptors are normally used to transport maltose into the cell, however the synthesis of these receptors is suppressed by glucose. Therefore, plating bacteria were grown in broth containing maltose to maximize adsorption. Adsorption of bacteriophage particles to the receptors is also influenced by the presence of magnesium ions. Magnesium ions are necessary for adsorption and phage integrity and therefore, the maltose-grown cultures were resuspended in 0.01 M $MgSO_4$ (Berger and Kimmel, 1987).

5 mL of overnight cultures of LE392 *E.coli* cells were prepared by inoculating 5 mL of LB broth containing 50 μL (20%) maltose and 50 μL 1 M $MgSO_4$ with one colony and growing with shaking at 20°C. The next day, the cultures were transferred into 10 mL centrifuge tubes and centrifuged at 6000 rpm for 5 min at 4°C, supernatant was removed and cells resuspended in 2.5 mL 10 mM $MgSO_4$. These LE392 plating cells were stored at 4°C, until required and were viable for only one week.

(iv) Plating for plaques.

The λ DNA packaged into heads is released from the host cell (*E.coli*) by destroying the host cell. This shows up as a cleared area of growth in a lawn of *E.coli* cells, known as a plaque.

10 μL of each packaged λ phage was added to 100 μL of LE392 (enough to grow a lawn of *E.coli*) plating cells. Adsorption of bacteriophage particles to the receptors is facilitated by magnesium ions, which can occur both at room temperature and at 37°C. However, penetration of the cell by the bacteriophage DNA and the subsequent events in the lytic cycle do not occur efficiently at room temperature, therefore samples were incubated at 37°C for 20- 30 minutes. The LE392 and packaged phage mixture was added to 3 mL of molten top LB agar (37°C), mixed well and poured onto pre-warmed (37°C) LB plates. Plates were incubated overnight at 37°C, and the numbers of plaques formed were determined the following day.

(v) Phage lysates.

Individual plaques were picked using a sterile pasture pipette. The agarose plug was then placed into 0.5 mL of SM buffer (0.01% gelatin, 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 8 mM $MgSO_4$) and left overnight at 4°C to enable phage to diffuse into the SM. Samples were then centrifuged for 5 minutes at 4°C at 12,000 rpm, and supernatant used as a source of phage particles. The phage lysate had to be titered before use and were stored at 4°C, where they lasted about 7-10 days before a significant drop in titer was observed.

(vi) Titering phage particles.

In order to establish how many plaque forming units (pfu) were present in λ phage lysates, LE392 plating cells were infected and the number of plaques that formed were scored.

Lambda phage suspension (lysate) was diluted (generally, a 1:100 dilution) in sterile SM buffer. 10 μL of the diluted phage was then added to 100 μL of LE392 plating cells, samples were incubated at 37°C for 20 minutes, then mixed with 3 mL of top LB agar and poured onto pre-warmed (37°C) LB plates. Plates were incubated overnight at 37°C, and the numbers of plaques formed were scored and the titer calculated as follows:

$$\text{CALCULATING TITER (pfu/mL):} \quad \frac{\text{Number of plaques scored x dilution factor}}{\text{Volume of } \lambda \text{ phage suspension plated (mL)}}$$

(vii) Medium scale λ preparations.

Medium scale λ preparations were carried out to isolate the amplified λ DNA. LB agarose plates were used to grow phage plaques, as agarose contains fewer contaminants than LB agar.

Appropriate volumes of phage lysate/suspension (diluted to generate about 10^5 pfu/plate) were added to 200 µL of LE392 plating cells and incubated for 20 minutes at 37°C. The mixture was then added to 3 mL of melted top agarose (40°C), mixed well and poured onto pre-warmed (37°C) agarose plates. Plates were incubated overnight at 37°C to induce confluent lysis. The following day plates were placed at 4°C for 10 minutes, then flooded with 9 mL of ice-cold SM buffer and left at 4°C overnight. After phage elution, the SM was removed and centrifuged for 15 minutes at 3,500 rpm. 5 mL of the supernatant was removed and 1 µL DNase I (10 mg/mL), 3 µL RNase (2mg/mL) added, mixed well and left at 37°C for 30 minutes; this step enabled the nucleic acids (released from the lysed bacterial cells) to be digested, without which adequate DNA isolation would not occur. The phage was then precipitated by the addition of 5 mL (20%) PEG-8000, which was left on ice for 1 hour. Phage DNA was pelleted by centrifugation for 40 minutes at 5,000 rpm (4°C) and the supernatant discarded. The pellet was resuspended in 500 µL SM where 5 µL 10% SDS and 5 µL 0.5 M EDTA (pH 8.0) was added to each phage preparation and then incubated in a waterbath at 68°C for 15 minutes. This was then followed by phenol-chloroform extraction to purify the DNA, and ethanol precipitation to concentrate the DNA. The final phage DNA pellet was resuspended in 50 µL TE and stored at 4°C, until required.

2.2.20 Southern Blotting.

Southern blotting (Southern, 1975) was used to determine whether a clone contained a DNA fragment of interest.

(i) Radiolabelling the Probe.

Probes were radioactively labelled by incorporation of [α-^{32}P] dCTP during random primer extension (Feinberg and Vogelstein, 1983). Double-stranded DNA was denatured by heating (95-100°C for 2-3 minutes) and random hexanucleotide oligonucleotides served as primers for the replication of the template DNA. The primer extension was carried out using Klenow fragment of DNA polymerase I which catalyzed the addition of nucleotide triphosphate, from a mixture that contained [α-^{32}P] dCTP. Ready To Go DNA Labelling Beads (-dCTP) (Amersham Pharmacia Biotech) were used to radioactively label 25 µL (2 ng/ µL) of PCR product with [α-^{32}P], following the manufacturers instructions.

The radioactive probe was purified using ProbeQuant™ G-50 Micro Columns (Amersham Pharmacia Biotech) following manufacturers instructions. These columns contain Sephadex

G-50 DNA (grade F), which removes any unincorporated labelled nucleotides from the DNA labelling reaction by size exclusion. Finally, 1 μL of the probe was analyzed for incorporation of radioactivity using Cerenkov counting (Beckman LS3801 Scintillation Counter) in counts per minute (cpm). The specific activity of the probe was calculated using the following method:

$$\text{Specific activity of the labelled probe (cpm/μg)} = \frac{\text{cpm of labelled probe (A) x Dilution factor}}{\text{μg of DNA labelled}}$$

(ii) Transferring DNA to nitrocellulose filter (blotting stand).

DNA samples were digested and heated to 65°C for 2-3 minutes to disrupt base pairing. DNA fragments were separated by gel electrophoresis on a 0.7% agarose gel at 30-40 V for 12-18 hours. The gel was agitated for 15 minutes in 0.25 M HCl and then with 0.5 M NaOH/ 0.5 M NaCl (to neutralize) for 30 minutes. Then the gel was washed with 2 x SSC (from 20 x stock of 3 M NaCl, 0.3 M trisodium citrate). The blotting stand was constructed as described by Southern (1975). The blot was disassembled after 48 hours and the DNA crosslinked to the nylon membrane by exposure to UV light for about 1 minute. Finally, the membrane was dried and stored dry until required.

(iii) Prehybridisation.

The membrane was soaked by capillary action with 2 x SSC and placed in pre-hybridization solution (6 x SSC, 2 x Denhardts reagent (0.02% ficoll, 0.02% polyvinylpyrolidone, 0.02% bovine serum albumen (BSA)) at 68°C for 3 hours, with shaking.

(iv) Hybridisation.

Hybridization was carried out in glass rotary tubes, the pre-hybridization fluid removed and 25 mL hybridization solution (1 M NaCl, 50 mM phosphate, 2 mM EDTA, 0.5% SDS, 1 x Denhardts, 10 mg herring sperm DNA) was placed in the tube, followed by the membrane. Herring sperm DNA was boiled for 5 minutes and cooled before use. This was required as a blocking agent to inhibit nonspecific binding of the probe. 25 μL of labelled probe was added directly to the hybridization solution in the tube and left overnight in the rotary oven at 68°C.

(v) Washes.

The membrane was washed to remove any excess probe and reduce nonspecific binding. This was done by removing the hybridization solution and replacing it with 6 x SSC and 1% SDS and placing it back in the rotary oven for 1 hour at 68°C. The membrane was then transferred into 6 x SSC plus 0.5% SDS and washed twice with shaking at 68°C for 30 minutes. This was followed by the high stringency wash, which was an exact 30 minute wash in 1 x SSC at 68°C.

The membrane was wrapped in Saran wrap to keep moist and exposed to X-ray film overnight, using a radioactive safe cassette. The autoradiograph was developed using a 100Plus™ Automatic X-ray (Kodak) film processor in a dark room.

2.2.21 Electrophoretic Mobility Shift Assays (EMSA).

EMSA are used to examine protein-DNA interactions that occur within particular promoter elements and protein transcription factors. The basis for EMSA is that the mobility a protein-DNA complex is retarded during non-denaturing polyacrylamide gel electrophoresis (PAGE).

(i)Radioactively labelling the oligonucleotide probe.

T4 polynucleotide kinase (PNK) was used to catalyse the transfer of the γ-phosphate of ATP to the 5'-hydroxyl terminus of the DNA strand. Labelling of oligonucleotides was carried out as indicated below in 1.5 mL microcentrifuge tubes, mixed then incubated at 37°C for 45 minutes.

Oligonucleotide (100 ng/μL)	1 μL
10 x 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^{32}P-[ATP] (10 μCi/μL, PerkinElmer)	4 μL
T4 polynucleotide kinase (10 u/μL, Roche)	1 μL
H_2O	3 μL

The labelled oligonucleotide was then annealled to its unlabelled counterpart by adding the complementary oligonucleotide in 6x excess. 6 μL complementary oligonucleotide (100

ng/μL) was added to the labelled oligonucleotide, along with 2.5 μL 1 M KCl and 31.5 μL of water. This mixture was heated to 95°C for 5 minutes in a boiling water bath, the water bath turned off and oligonucleotides were left to gradually cool to room temperature in the water bath. The cooling process took approximately 1 hour and was an important step as it enables correct annealing to occur. An equal volume of gel shift buffer (40 mM Tris pH 7.6, 16% ficoll, 100 mM KCl, 0.4 mM EDTA, 1 mM DTT) was added to the cooled double-stranded oligonucleotides, which were gel purified immediately.

(ii)Purifying the labeled Oligonucleotides.

To remove any residual single-stranded oligonucleotides, unincorporated label or other contaminants, the labelled oligonucleotides (total volume 50 μL) were purified by electrophoresis on a 10% polyacrylamide gel (37 cm long with 0.4 mm spacers) in 1x TBE (0.09 M Tris, 0.09 M boric acid, 0.02 M EDTA pH 8.0) at 30 W for 1 ½ hours. The gel was wrapped in saran wrap and exposed to X-ray film for approximately 1 minute, which enabled the location of the double-stranded labelled oligonucleotides within the gel. The appropriate band(s) were excised from the gel and DNA eluted by placing each gel slice into 500 μL 50 mM KCl overnight at 37°C. The following day, the gel suspension was mixed by vortexing, centrifuged at 12,000 g for 5 minutes and the supernatant was transferred to a fresh 1.5 mL microcentrifuge tube. 1 μL of each probe was analyzed for incorporation of radioactivity using Cerenkov (Beckman LS3801 Scintillation Counter) counting. The samples were stored in a radioactively safe Perspex container at 4°C until required.

(iii)Preparing Double-Stranded Competitors.

Unlabelled double-stranded competitors were generated by adding 5 μL of each complementary single-stranded oligonucleotide (1 μg/μL) together and heating to 95°C for 5 minutes in a Hybaid Omn-E thermal cycler. Samples were cooled stepwise by 10°C every 5 minutes until they reached 25°C. No other reagents were required to enable annealing due to the high concentration of single stranded oligonucleotides. Double-stranded competitors were stored at -20°C until required.

(iv)Electrophoretic Mobility Shift Assays.

Poly dI.dC is a non-biological polymeric nucleic acid that acts as competitor DNA in EMSAs. Excess poly dI.dC is added to the reactions to minimize the occurrence of nonspecific DNA-protein interactions.

HeLa extract (about 6 μg protein) was added to the following premix; 10 μL gel shift buffer, 1 μL poly dI.dC (1 mg/mL in $MgCl_2$) and water to make the total volume up to 20 μL (25 μL for antibody supershifts), then incubated on ice for 10 minutes. Competitors were added before addition of probe. Approximately 0.5 ng of labeled oligonucleotide (1 μL) was added to the premixes and left at room temperature for 15 minutes. As a control 2 μL of loading dye was loaded in an exterior lane to monitor the travel front; 10 μL of each reaction was loaded onto a non-denaturing 4% polyacrylamide gel in 0.25 x TBE and electrophoresed at 200 V for about 1 hour (BRL V15.17 apparatus fitted with 0.75 mm spacers). The gels were transferred onto DE-81 paper and dried using Bio-Rad Gel Dryer 583 for about 15 minutes at 80°C. The gels were then exposed to X-ray film for at least 20 hours at -70°C using a radioactive safe cassette with intensifying screens. The X-ray (Kodak) films were developed using a 100Plus™ Automatic X-ray film processor in a dark room.

2.2.22 Tissue Culture.

HeLa cells were used for transient transfections and HeLa cell extracts for electrophoretic mobility shift assays. All tissue culture operations prior to harvesting were done aseptically and in a laminar flow hood (Crossflow 1800 with HEPA filter, Westinghouse). Cells were incubated at all times in a 37°C, 5% CO_2 incubator (Jouan IG150, France) in humid conditions.

(i)Media for HeLa Cells.

HeLa cells were grown in complete Eagle's minimal essential medium (MEM) containing non-essential amino acids (GIBCOBRL, Invitrogen) which was prepared as instructed by the manufacturers and filter sterilized through a 0.2 μm filter. MEM was dispensed into sterile bottles in 180 mL volumes, which was supplemented as required with 10% (20 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). Media was stored at 4°C, but warmed to room temperature before use.

(ii)Starting HeLa Cell Cultures.

Cell cultures were started from 1 mL frozen HeLa cell stocks (prepared by Kirsty Allen, Institute of Molecular BioSciences, Massey University, Palmerston North). A frozen aliquot of HeLa cells (in 10% DMSO in FCS, stored under liquid nitrogen) was thawed and the cells transferred to 5 mL of supplemented MEM, then mixed thoroughly. The cells were pelleted

by centrifugation for 5 minutes at approximately 100x g and the supernatant removed. The pellet was resuspended in 2 mL of supplemented MEM and 1 mL used in each monolayer T80 flask (Nunc) containing 14 mL of supplemented MEM. HeLa cells were left to grow for about 2-3 days in the 37°C incubator before passaging.

(iii)Maintenance of HeLa Cells.

HeLa cells were grown in T80 flasks to 80-90% confluence before passaging the cells into new flasks with fresh media. Passaging the cells initially involves removing the old media and then washing the cells twice with 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) plus 1 mL 10x trypsin (GIBCOBRL). The cells were left lying flat for a few minutes to dislodge them from the flask surface; further detachment was done by abruptly tapping the side of the flask. The cells were then thoroughly resuspended by aspiration in 5 mL of supplemented MEM, where 1 mL of resuspended cells was used to seed a new T80 flask containing 14 mL of fresh supplemented MEM and the remainder of the HeLa cells was used to seed plates for transfections. Cells were placed back into the 37°C incubator and were passaged every 2-3 days.

(iv)Preparing HeLa Cells for Freezing.

Stocks of HeLa cells were frozen after the first passage. Cells were grown to 80% confluence, passaged and resuspended in 5 mL of FCS containing 10% DMSO. The cells were then dispensed into sterile 1 mL cryotube (Nunc) and frozen slowly (by wrapping cryotubes in copious layers of tissue paper) at -70°C, to avoid disruption of the cell membrane. Once frozen the HeLa cells were stored under liquid nitrogen.

2.2.23 Preparing HeLa Cell Extracts for EMSA.

HeLa cells were grown to 80-90% confluence in 14 mL supplemented MEM in 150 mm diameter plates (Nunc) in a 37°C CO_2 incubator. The media was removed from the HeLa cells and washed twice with 2 mL PBS (0.14 M NaCl, 2.7 mM KCl, 4.3 mM $NaHPO_42H_2O$ pH 7.2). Then 1 mL of TEN buffer (40 mM Tris-HCl pH 7.4, 1 mM EDTA, 0.15 M NaCl) was added to each plate and cells scraped off the plate and placed into 1.5 mL microcentrifuge tubes. The cell suspension was centrifuged at 12,000 g for 5 minutes and the supernatant discarded. The pellet was 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). To disrupt cells, they were treated with 3 freeze-thaw

cycles using liquid nitrogen. The cell extract was centrifuged again at 12,000 g for 5 minutes (at 4°C) and the supernatant dispensed into 30 μL volumes, then immediately snap frozen in liquid nitrogen and stored at -70°C. The proteins present in the HeLa cell extract were quantified using the Bradford Protein Assay.

2.2.24 *Bradford Protein Assay.*

The protein concentration of HeLa extracts was determined using the Bradford protein assay dye reagent concentrate (BioRad) according to manufacturers instructions. The concentrated reagent was diluted 1:5 in water, then 200 μL was added to BSA (1mg/μL) standards (diluted to a range of 0-2.5 μg in water) and 1 or 2 μL of HeLa cell extract (undiluted, diluted 1:5 and 1:10 in water). Samples were left to develop colour at room temperature for a minimum of 10 minutes. Reactions were carried out in a 96-well microplate (Nunc) and absorbance read at 595 nm. A protein standard curve was constructed using the standard amounts of BSA and the amount of protein present in HeLa extract was estimated.

2.2.25 *Transient Transfections.*

HeLa cells were transiently transfected with FuGENE^{TM}6; a multi-component lipid-based reagent that complexes with and transports DNA into the cell during transfection. FuGENE^{TM}6 requires that cells are 50-80% confluent on the day of transfection. Thus, the day prior to transfection 800 μL of supplemented MEM (in 12 well plates) was seeded with 100 μL (approximately 2 drops) of HeLa cell suspension, which had been taken from freshly resuspended (in 5 mL) 80-90% confluent HeLa cell stocks. Each transfection was performed in triplicate using 3x as much FuGENE^{TM}6 as DNA (3 μL FuGENE^{TM}6:1 μg DNA) according to the manufacturers instructions. HeLa cells were harvested 60-65 hours after transfection and assayed for luciferase and β-galactosidase activity as soon as possible.

2.2.26 *Harvesting HeLa Extracts for Luciferase and β-Galactosidase Assays.*

The media was removed from the cells, which were rinsed twice with 500 μL of PBS. 80 μL of cell lysis buffer (25 mM Tris-phosphate pH 7.8, 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid, 10% glycerol, 1% TritonX-100) was added to each well containing cells and left at room temperature for 15 minutes. The cells were then scraped off the surface of wells using a cell scraper and transferred to 1.5 mL microcentrifuge tubes. Cell extracts were centrifuged at 12,000 rpm for 30 seconds to remove

any cellular debris and the supernatant transferred into fresh microcentrifuge tubes. The cell extracts were stored at 4°C until required.

2.2.27 *β*-Galactosidase Assays.

The plasmid pSVβGal was used as an internal control for transfections, which enabled the β-galactosidase assay to be performed to check the efficiency of transfection. Cell extract (5 μl) was incubated with 50 μL ONPG (*o*-nitrophenyl-β-D-galactopyranside, 2 mg/mL in 60 mM NaH_2PO_4, 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 $MgCl_2$) in microtitre (96 well, Nunc) plates at 37°C for 3-8 hours. After incubation, 50 μL 1 M Na_2CO_3 was added to stop the reaction and develop full yellow colour. Absorbances were measured at 405 nm against a blank of 5 μL cell lysis buffer using the 96 well plate reader (anthos reader HT2 type 12 500, anthos labtech instruments, Salsburg, Austria).

2.2.28 Luciferase Assays.

The reporter gene luciferase enabled the measurement of promoter activity. 5 μL of HeLa cell extract was mixed with 20 μL of luciferase reagent (Promega) and analyzed using a FLUOstar galaxy microplate reader (BMG Labtechnologies Pty. Ltd, Melbourne, Australia). Luciferase activity was measured by the amount of light that was emitted from HeLa extracts and fibre optics within the microplate reader detected and conveyed this information as actual photon counts. This information was accessed through Excel (Microsoft 97), which is connected to the microplate reader. The maximum photon count obtained (within the 3 minute measuring period) was the value retained for data analysis. 5 μL cell lysis buffer was used as a blank.

2.3 GMO approval codes.

This work was carried out with ERMA regulatory authority approval as follows:

- Topoisomerase IIβ clones (pBluescript and λ clones): GMO 99/MU 25
- Expression plasmids (Sp1 and Sp3): GMO 00/MU 40
- Topoisomerase IIα promoter plasmids (pGL3B): GMO 98/MU 53

Chapter 3: Topoisomerase IIβ

3.1 Introduction.

Topoisomerase IIβ has been shown to be an important target for anti-cancer drugs, with the observation that some drugs may preferentially target topoisomerase IIβ rather than topoisomerase IIα (Gao *et al*, 1999). One of the mechanisms thought to be involved in the development of drug resistance is the down-regulation of expression of both topoisomerase IIα and β (Robert and Larsen, 1998; Dereuddre *et al*, 1997) and the actual ratios of topoisomerase II α and β have been implicated as an important factor (Padget *et al*, 2000).

The initial aim was to isolate and investigate the 5'-flanking sequence of the human topoisomerase IIβ gene. This would have allowed a comparison of the regulatory mechanisms between the α and β gene. The partial topoisomerase IIβ promoter sequence had been published in 1997 (Ng *et al*, 1997) however this sequence did not include an extensive upstream regulatory region. The search for the topoisomerase IIβ promoter started with two clones that had been isolated from a human genomic library screened with a topoisomerase IIβ exon II probe (personal communications Sally Davies, Institute of Molecular Medicine, University of Oxford). Both of the putative topoisomerase IIβ promoter clones had been subcloned into pBluescript SK (+) and had been named TopoB3 and TopoB3.2. The original objective for analysis of these clones was to perform Southern blots using a short topoisomerase IIβ oligonucleotide designed from the sequence already reported by Ng *et al* (1997). However, difficulties obtaining a topoisomerase IIβ PCR product from human genomic DNA meant that sequencing of the topoisomerase IIβ clones had been completed before the probes were even prepared.

3.2 Generating probes for topoisomerase IIβ.

Short topoisomerase IIβ promoter double-stranded oligonucleotides were required as probes to analyze the topoisomerase IIβ clones via Southern blots. These were generated using PCR and this exercise was done concurrently with the sequence analysis of the topoisomerase IIβ pBluescript (+) clones.

3.2.1 Topoisomerase IIβ PCR.

PCR reactions were carried out using *Taq* polymerase, using a human genomic DNA template as described in table3.1. Two primer sets were designed using the 5'-flanking sequence of topoisomerase IIβ reported by Ng *et al* (1997, refer to figure 3.1).

The first primers, Ng1 & 2 were used in PCR to amplify a 380 bp product. Several different concentrations of $MgCl_2$ were tested in an attempt to obtain the desired fragment (topoIIβ1, refer to figure 3.2a). This reaction was eventually successful using 2.5 mM $MgCl_2$ but unfortunately, the PCR product could not be reproduced. Thus, a second set of primers was designed (Ng3 & 4). PCR with these primers was also carried out with a range of $MgCl_2$ concentrations. After several attempts, a small amount of specific PCR product (300 bp) was obtained using 1.25 mM $MgCl_2$ (lane 3, figure 3.2B). In order to ensure that the topoIIβ2 DNA fragment would be readily available, it was cloned into a pGEM-T vector (Promega, refer to appendix 2 for vector map). PCR reactions are described in table 3.1.

Figure 3.1: The 5'-flanking sequence of the topoisomerase IIβ promoter published by Ng *et al* (1997) used to generate topoisomerase IIβ probes.

The above sequence was used to design two sets of primers, Ng1/2 (red) and Ng3/4 (blue). PCR was used to amplify short regions of the topoisomerase IIβ promoter. The coding region for exon 1 within the sequence is shown in bold. The regions where primers anneal are illustrated by the arrows. Primers Ng1/2 and Ng3/4 generate oligonucleotides approximately 380 bp in size and 300 bp in size, respectively. Primer sequences can be found in appendix 1.1.

topoIIβ1	topoIIβ2
95°C for 5 minutes. Denaturing 95°C for 1 minute. Annealing 50°C for 1 minute. Extension 72°C for 1 minute. } 30x cycles.	95°C for 5 minutes. Denaturing 95°C for 1 minute. Annealing 57°C for 1 minute. Extension 72°C for 1 minute. } 30x cycles.
The PCR mix contained 5 μL (50 ng/μL) of each primer Ng1 and Ng2, 5 μL (3 mM) dNTPs, 5 μL 10x PCR buffer (100 mM Tris-HCl, 500 mM KCl pH 8.3, Sigma), varying amounts 2.5, 5 or 7.5 μL $MgCl_2$ (25 mM, Sigma), 1 μL *Taq* Polymerase (5 U/μL, Sigma), 1 μL template plasmid (1 ng/μL) and dH_2O to make up the total volume up to 50 μL.	The PCR mix contained 5 μL (50 ng/μL) of each primer Ng3 and Ng4, 5 μL (3 mM) dNTPs, 5 μL 10x PCR buffer (100 mM Tris-HCl, 500 mM KCl pH 8.3, Sigma), varying amounts 2.5, 5, 7.5 or 10 μL $MgCl_2$ (25 mM, Sigma), 1 μL *Taq* Polymerase (5 U/μL, Sigma), 1 μL template plasmid (1 ng/μL) and dH_2O to make up the total volume up to 50 μL.

Table 3.1: PCR protocols for topoisomerase IIβ probes.

The PCR reactions involved in producing topoisomerase IIβ probes. The PCR reactions were carried out in thermal cyclers (HYBAID Omn-E, SciTech (NZ) Ltd or GeneAmp®PCR System 2700, Applied Biosystems, Foster City, USA). Primer sequences can be found in appendix 1.1.

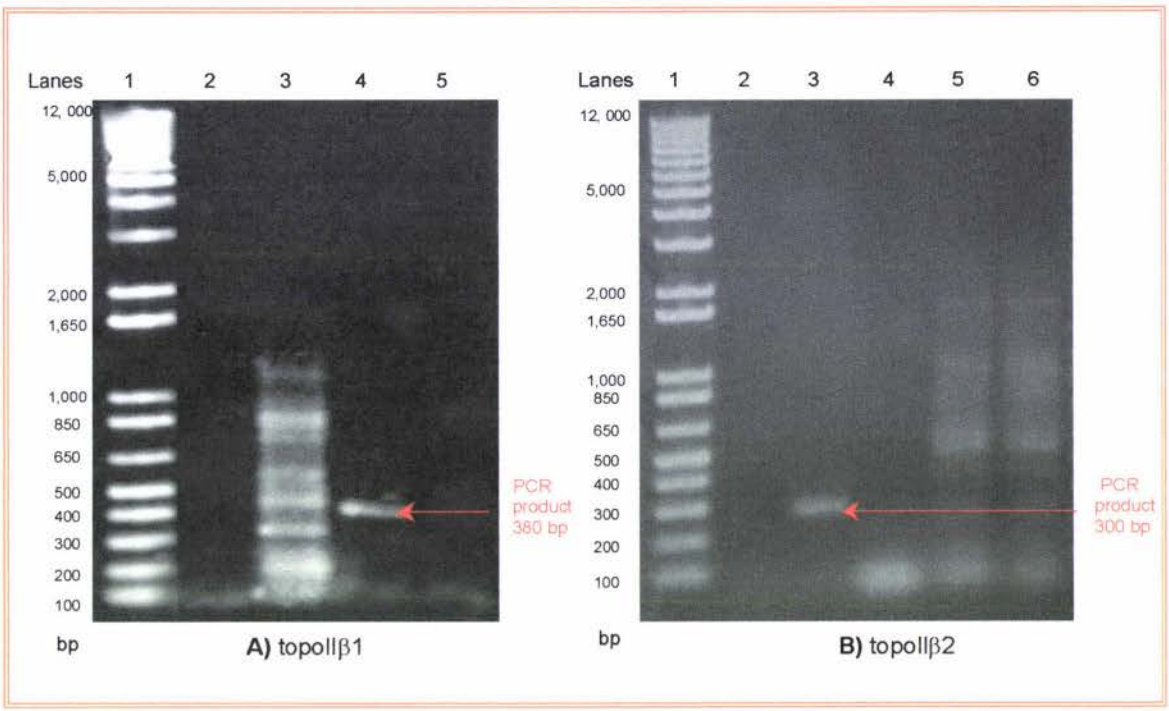


Figure 3.2: PCR reactions to generate topoisomerase IIβ oligonucleotides.

10% of each PCR reaction was electrophoresed on a 1% agarose gel in 1xTAE buffer for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 μg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp). 10 μL 1Kb plus ladder is present in lane 1.

A) PCR products using primers Ng1 and Ng2:

Lanes 2 5 μL negative control (PCR reaction with no template): No PCR product

3 5 μL standard PCR reaction with 1.25 mM $MgCl_2$: Non-specific PCR products

4 5 μL standard PCR reaction with 2.5 mM $MgCl_2$: desired PCR product (topoIIβ1)

5 5 μL standard PCR reaction with 3.75 mM $MgCl_2$: No PCR product

Lane 4 contains the desired topoisomerase IIβ PCR product, 380 bp in size.

B) PCR products using primers Ng3 and Ng4:

Lanes 2 5 μL negative control (PCR reaction with no template): No PCR product

3 5 μL standard PCR reaction with 1.25 mM $MgCl_2$: desired PCR product (topoIIβ2)

4 5 μL standard PCR reaction with 2.5 mM $MgCl_2$: No PCR product

5 5 μL standard PCR reaction with 3.75 mM $MgCl_2$: Non-specific PCR products

6 5 μL standard PCR reaction with 5.0 mM $MgCl_2$: Non-specific PCR products

Lane 3 contains the desired topoisomerase IIβ PCR product, 300 bp in size.

3.2.2 Cloning PCR product TopoIIβ.2 into pGEM-T vector.

The topoisomerase IIβ PCR product topoIIβ2, was purified using QIAquick PCR purification kit (QIAGEN) and the concentration determined by gel electrophoresis using quantification standards. Ligations were carried out as described in section 2.2.15 and are illustrated in table 3.2 below. Reactions were incubated overnight at 4°C before transformations were attempted. pGEM-T vectors (appendix 2, Promega) are supplied linearized and ready for ligation as they have already been digested with *Eco* RV. They also contain a 3' terminal thymidine attached to either cut end, to facilitate efficient ligation of PCR products. Transformations were carried out as described in section 2.2.16, utilizing blue/white selection (refer to table 3.3 for results).

Reaction	2x buffer	Vector (25 ng)	PCR product (~8 ng)	Control insert DNA	T4 ligase (1 U/μL)	H_2O	Total volume
a) Vector minus ligase (control)	2.5 μL	0.5 μL	-	-	-	2 μL	5 μL
b) Vector plus control DNA (positive control)	2.5 μL	0.5 μL	-	1 μL (25 ng)	0.5 μL	0.5 μL	5 μL
c) Vector plus ligase (background control)	2.5 μL	0.5 μL	-	-	0.5 μL	1.5 μL	5 μL
d) Vector plus PCR DNA (3:1 insert:vector molar ratio)	2.5 μL	0.5 μL	4 μL (dry)	-	0.5 μL	1.5 μL	5 μL

Table 3.2: Ligation reactions to clone into pGEM-T Vector.

Ligation reactions were carried out to clone the 300 bp PCR product topoIIβ2 into pGEM-T. Reactions were incubated overnight at 4°C, and then half of each ligation was used to transform XL-1 blue (*E.coli*) competent cells.

The concentration of the pGEM-T vector was 50 ng/μL, whereas the concentration of the PCR product was about 2 ng/μL. 4 μL of PCR product was aliquoted and dried down using the SpeedVac (speed vacuum concentrator), then resuspended in 1.5 μL H_2O. For PCR DNA, the ligation was carried out using a 3:1 (insert:vector) molar ratio and for the control DNA the molar ratio was 1:1.

Reaction	Ligation mixture	XL-1 competent cells	Dilution	Volume plated	Number of colonies White	Blue
1) XL-1 competent cells (control)	- -	10 µL	10^6 10^7	50 µL 50 µL LB plates	TMC 301	- -
2) XL-1 competent cells (negative control)	-	10 µL	Undiluted	50 µL LB+amp	0	-
3.a) Vector minus ligase (control)	3 µL	100 µL	Undiluted 1:5 1:10	50 µL 50 µL 50 µL	0 0 0	50 6 7
4.b) Vector plus control DNA (positive control)	3 µL	100 µL	Undiluted 1:5 1:10	50 µL 50 µL 50 µL	341 53 20	48 4 3
5.c) Vector plus ligase (background control)	3 µL	100 µL	Undiluted 1:5 1:10	50 µL 50 µL 50 µL	0 0 0	TMC 6 6
6.d) Vector plus PCR DNA	3 µL	100 µL	Undiluted 1:5 1:10	30 µL 50 µL 50 µL	150 14* 2	136 18 17

Table 3.3: Results of XL-1 transformations using pGEM-T ligations.

Transformation of XL-1 competent cells, using plasmids generated from ligation (table 3.2) of pGEM-T vector and topoIIβ2 DNA or control DNA. Unless indicated, all samples were plated onto LB plus ampicillin plates with 20 µL IPTG (20 mg/mL in water) and 50 µL X-gal (20 mg/mL). Plates were incubated overnight at 37°C and colonies scored the following day, some plates had too many colonies to count (TMC). Colonies were creamy white in colour or blue. Appropriate white colonies (*) were selected and cultured to isolate plasmid DNA for analysis.

XL-1 competent cells cannot grow in the presence of ampicillin, however competent cells transformed with pGEM-T plasmids are ampicillin resistant. Each transformation was spread onto LB plates containing ampicillin, IPTG and X-gal to distinguish which plasmids had the insert DNA. XL-1 colonies that appear blue are those with empty plasmids, whereas the white colonies carry plasmids with an insert in the correct location. The variation in colour is generated as the *lacZ'* gene is disrupted by the presence of the insert DNA, so α-complementation cannot occur within the bacteria.

The control plates that contained XL-1 cells on LB only (reaction 1) grew only white colonies, as expected. This, in conjunction with the control plate of XL-1 cells with ampicillin (reaction 2), where no colonies grew, provided evidence that the XL-1 cells were viable and that no contamination or spontaneous ampicillin resistance could be attributed to any of the white colonies observed in other reactions in these experiments.

The ligation reaction with the control DNA (reaction 4.b) was set up to assess the efficiency of the ligation. Typically, about 10-40% of colonies should be blue and no less than 60% of colonies should be white. If no white colonies occurred it could have been due to a failed ligation. However, if no colonies grew then it could have been due to a failed transformation. In addition, if fewer than 50% of colonies were white then the ligations were likely to have been carried out using suboptimal conditions (Promega pGEM-T technical manual, No. 042).

The results in table 3.3 illustrate that approximately 10% of colonies in reaction 4.b were blue, and the majority of colonies were white, indicating that the ligation efficiency was satisfactory, and that the transformation worked well. The blue colonies that arose were probably due to undigested pGEM-T vectors or the absence of a 3' terminal thymidine attached to ends of the vector (which meant the vector could then religate). The background blue colonies observed in control reactions 5.c and 3.a, indicate that the majority of blue colonies were due to vector self-ligation.

The ligation of pGEM-T and the 300 bp PCR product topoIIβ2, gave rise to about 50% each of white and blue colonies. Fortunately, there were many white colonies to choose from and 8 white colonies were isolated to analyze for the presence of the insert DNA. Each colony was maintained in the presence of ampicillin and DNA extracted using the Quantum® Prep Plasmid Miniprep Kit (BioRad) as described in section 2.2.17(ii). Each pGEM-T plasmid was checked for insert presence by enzymatic digestion, using *Sph* I and *Nde* I and shown in figure 3.3.

The results from figure 3.3 illustrate that miniprep DNA isolated from colonies 1, 3 and 5 appeared to have the correct insert; ie. the topoisomerase IIβ PCR proximal promoter region. Interestingly the sample seen in lanes 3 and 4 appears to generate an insert approximately 600 bp in size, which is probably due to the incorporation of two PCR products ligated

together into the pGEM-T vector. Bands that exhibit lower mobility can be seen in many of the lanes, these are due to incomplete digestion of pGEM-T.

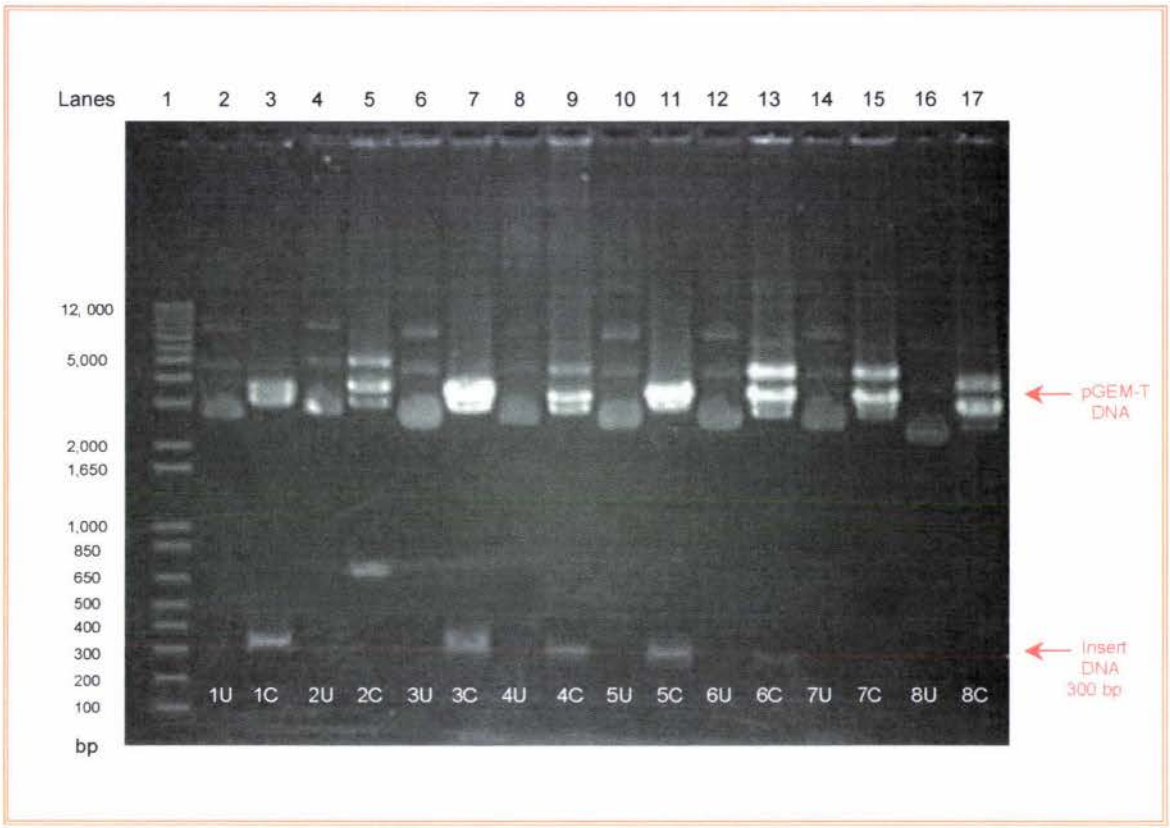


Figure 3.3: Diagnostic digests of various pGEM-T plasmids.

10 µL (50 ng/µL) of each mini-preparation (1-8) of pGEM-T topoIIβ2 colonies was digested with enzymes *Sph* I and *Nde* I to release the insert DNA. Digested (C for cut) samples and uncut (U) samples were analyzed on a 1% agarose gel in 1xTAE buffer for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 µg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp). 10 µL 1Kb plus ladder is present in lane 1. Enzymatic digestions were performed as described in section 2.2.6.

Even numbered lanes	Relate to 2 µL (U) uncut miniprep plasmid DNA (50 ng/µL)
Odd numbered lanes	Relate to 15 µL (C) cut/digested miniprep plasmid DNA

Insert DNA can be seen at the bottom of the gel approximately 300 bp in size, whereas the bands of lower mobility are the pGEM-T plasmid. In the uncut lanes, three bands of high mobility can be seen, these are different states of the circular plasmid; ie relaxed or supercoiled.

The topoisomerase IIβ promoter has a unique restriction site for the enzyme *Nco* I, which can be used to positively identify the PCR product. In addition, the pGEM-T vector has restriction sites for enzymes *Sph* I, *Nde* I and *Nco* I within the multiple cloning site (MCS). Thus, mini-preparation plasmid DNA from sample 1 was subjected to further analysis (via PCR and enzymatic digestion), to confirm that the insert DNA was indeed the desired topoisomerase IIβ sequence.

To amplify the insert within the pGEM-T vector, PCR was carried out using *Taq* polymerase with standard pGEM-T PCR primers, T7 and SP6 (described in section 2.2.12). The size of the insert after PCR was approximately 420 bp (lane 2, figure 3.4). The larger size was due to extra bases attributed to the MCS of the vector. When the PCR product was digested with *Sph* I and *Nde* I (to remove most of the sequences from the MCS), the size of the TopoIIβ2 was shown to be about 350 bp (lane 3, figure 3.4). The PCR product was then digested with *Nco* I, *Sph* I and *Nde* I resulting in the production of two major fragments of about 100 bp and 280 bp confirming the identity of the insert, there would have been a few smaller bands also but they were not resolved in the gel (lane 4, figure 3.4). Mini-preparation plasmid DNA was digested with *Sph* I and *Nde* I (lane 6, figure 3.4) and the size of the fragment matched that observed in lane 3. In addition, when the mini-preparation plasmid DNA was digested with all three of the enzymes (lane 7, figure 3.4) it clearly showed that the insert DNA was digested with *Nco* I, which also enabled positive identification. The insert DNA was indeed the desired topoisomerase IIβ oligonucleotide.

Therefore, a region of the topoisomerase IIβ promoter had been cloned, which could be excised and labelled for use as a probe in Southern blotting and hybridization analysis of either cloned or genomic DNA. However, due to several technical difficulties encountered with obtaining reproducible PCR products, sequence analysis of the two pBluescript clones (TopoB3.2 and TopoB3) was completed by the time this cloning had been completed. Analysis of the two pBluescript clones is described in the following section.

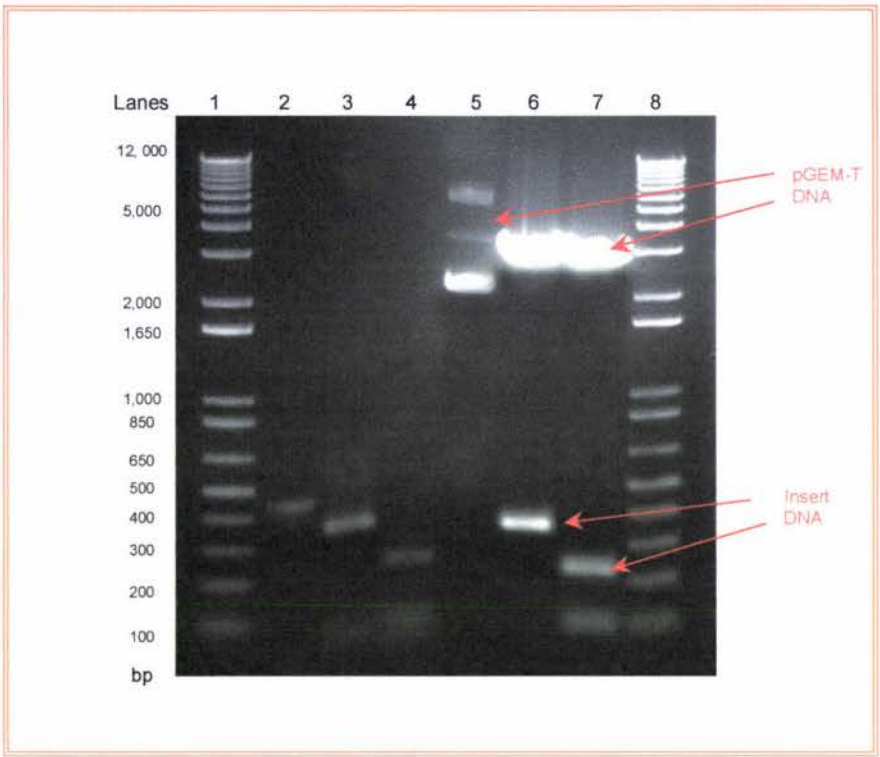


Figure 3.4: Analysis of pGEM-T vector containing TopoIIβ.2.

10 µL (30 ng/µL) of mini-preparation DNA or 10 µL (20 ng/µL) of PCR DNA was digested with 1 µL (10 U/µL) of each restriction enzyme. Samples were electrophoresed on a 1% agarose gel in 1x TAE buffer for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 µg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	1 & 8	10 µL 1kb plus ladder
	2	2 µL PCR reaction
	3	20 µL PCR product digested with *Sph* I and *Nde* I
	4	20 µL PCR product digested with *Sph* I, *Nde* I and *Nco* I
	5	2 µL uncut plasmid DNA
	6	20 µL plasmid DNA digested with *Sph* I and *Nde* I
	7	20 µL plasmid DNA digested with *Sph* I, *Nde* I and *Nco* I

The pGEM-T clone was confirmed to have a fragment of the topoisomerase IIβ promoter sequence, approximately 350 bp in size.

3.3 pBluescript SK (+) topoisomerase IIβ clones.

pBluescript SK (+) is a plasmid that is approximately 3 kb in size and can carry up to 10 kb of insert DNA. The insert DNA is flanked by two primer sites representing the T3 and T7 promoter elements, which can enable insert DNA amplification using PCR (refer to appendix 2 for vector map).

Two pBluescript SK (+) clones (TopoB3 and TopoB3.2) were initially sequenced from the T3 end of the MCS by dideoxy chain termination using Big Dye chemistry (Lorraine Berry; MUSeq facility, Institute of Molecular BioSciences, Massey University). Each sequence was compared with the topoisomerase IIβ cDNA sequences (accession number AF087143) for insert identity and orientation with sequence analysis software GCG (Version 9.1; Wisconsin Genetics Computer Group, USA), using the bestfit programme. Bestfit takes two sequences and makes matched sequence comparisons, so that any similarities between the two sequences can be observed.

Of the two clones, TopoB3.2 was chosen for further analysis as the topoisomerase IIβ insert sequence appeared to continue in the required direction, away from exon 2 of the human topoisomerase IIβ gene (refer to figure 3.5). The exon organization of the human topoisomerase IIβ gene in relation to the functional domains of the topoisomerase IIβ protein is illustrated in figure 3.6. Therefore, the 5' topoisomerase IIβ promoter region would be upstream of exon 1 (Ng *et al*, 1997). Restriction endonuclease digests were carried out to generate a restriction map and primer walking was used to sequence the entire topoisomerase IIβ insert, the results for these are described in the following sections.

The other clone TopoB3, was orientated in the opposite direction to the TopoB3.2 clone, such that it appeared to sequence into exon 2 and lead into intron 2 away from the promoter region. Further analysis of the TopoB3 clone was consequently abandoned.

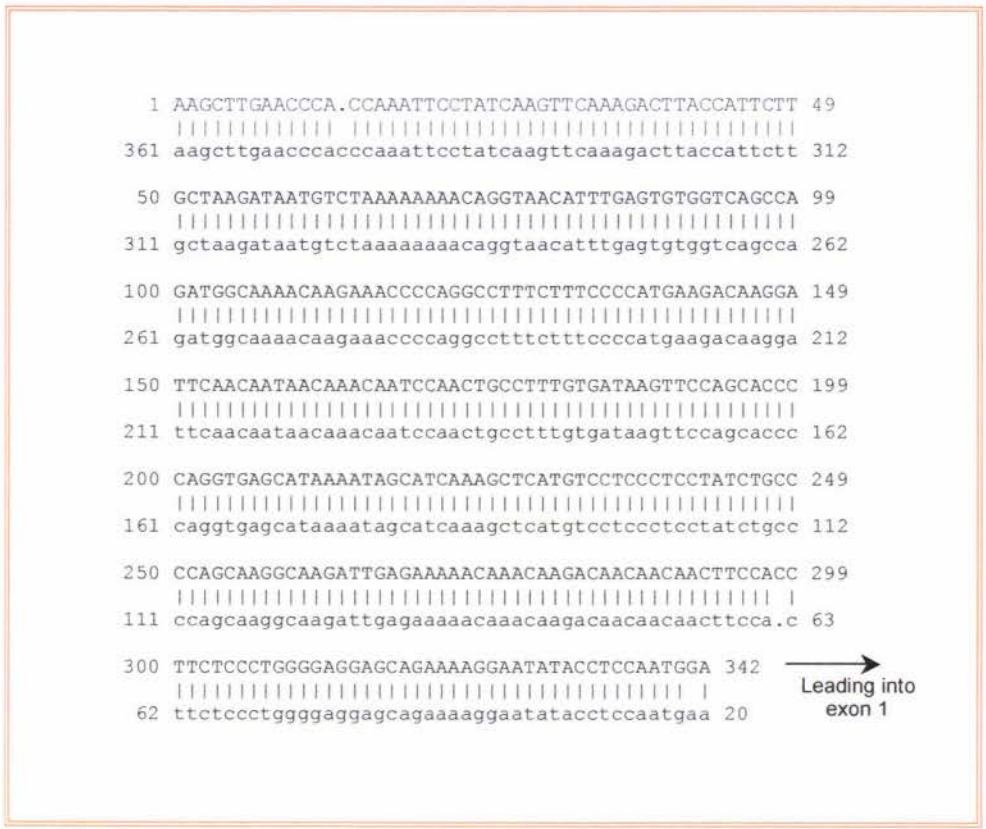


Figure 3.5: GCG analysis of topoisomerase IIβ clone TopoB3.2.

GCG analysis of the sequence from the T3 end of clone TopoB3.2. The TopoB3.2 sequence (top rows) was aligned to the exon 2 sequence (bottom rows) from human topoisomerase IIβ cDNA (accession number AF087143), using the 'bestfit' program within GCG.

Number 1 on the top row designates position 1 on the TopoB3.2 sequence, which corresponds to position 361 of the human topoisomerase IIβ exon 2 sequence. The TopoB3.2 insert DNA sequence was orientated in the vector such that exon2 should lead into intron 1, followed by exon 1 and consequently the promoter region of the human topoisomerase IIβ gene.

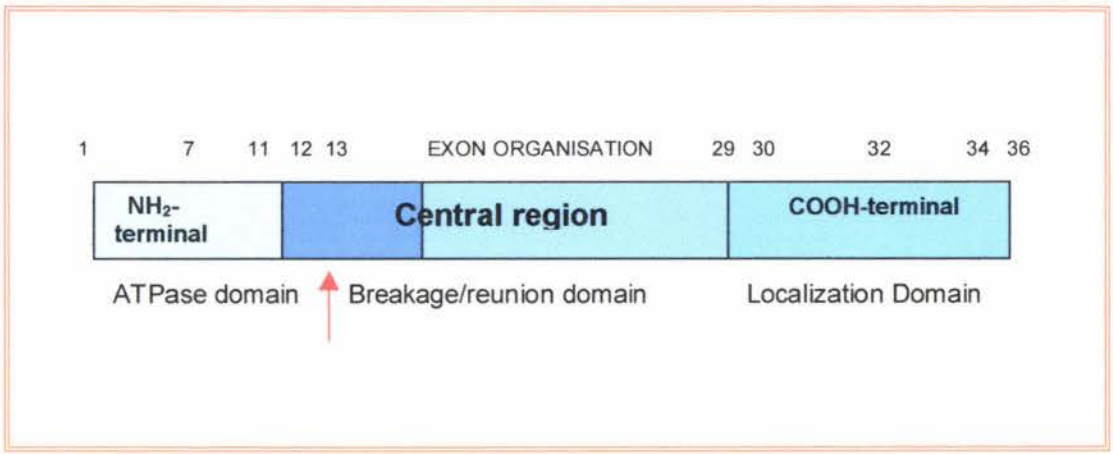


Figure 3.6: Schematic representation of the exon organization of the human topoisomerase IIβ gene.

The human topoisomerase IIβ protein (blue) is encoded by the various regions of the topoisomerase IIβ gene. Each domain of the topoisomerase IIβ protein is encoded by specific exons of the topoisomerase IIβ gene (numbers above the blue domains), where each exon has been mapped to particular regions of the topoisomerase IIβ protein. Domain organization is crucial for correct enzyme function. The second domain (indicated with the red arrow) is thought to play a role in coordinating the actions of the ATPase domain with the DNA binding and cleavage domain within the second division of the central region. (Figure created from information drawn from: Lang *et al*, 1998; Austin *et al*, 1993; Marsh *et al*, 1997; Cowell *et al*, 1998 and Sng *et al*, 1999).

3.3.1 Restriction map of TopoB3.2.

In order to determine the size of the insert of TopoB3.2 and identify possible restriction sites within the insert, approximately 1 μg of TopoB3.2 DNA was digested with enzymes *Hind* III, *Eco* RI, *Xho* I, *Pst* I, *Bam* HI *and Sac* I (as described in section 2.2.6). Each sample was loaded onto a 1% agarose gel, electrophoresed for about 1 hour at 80 V, then stained with ethidium bromide where bands were visualized by exposure to UV light. The size of each band was determined using the 1 kb plus ladder (refer to appendix 3.1) and the information gathered in conjunction with the sequencing data from the T3 primer, was used to construct the restriction map illustrated in figure 3.7.

The *Hind* III digest revealed that the TopoB3.2 insert was about 2.5 kb rather than 3.2 kb as originally indicated, and there were two *Pst* I sites, two *Sac* I sites and one *Eco* RI restriction site within the insert. The size of the fragment between the two *Pst* I sites within the insert was approximately 600 bp, whereas the distance between the two *Sac* I sites was only about 400 bp (refer to appendix 3.1, for fragment sizes). The insert was flanked by the restriction sites for *Hind* III, *Pst* I, *Bam* HI, *Eco* RI *and Xho* I and *Sac* I.

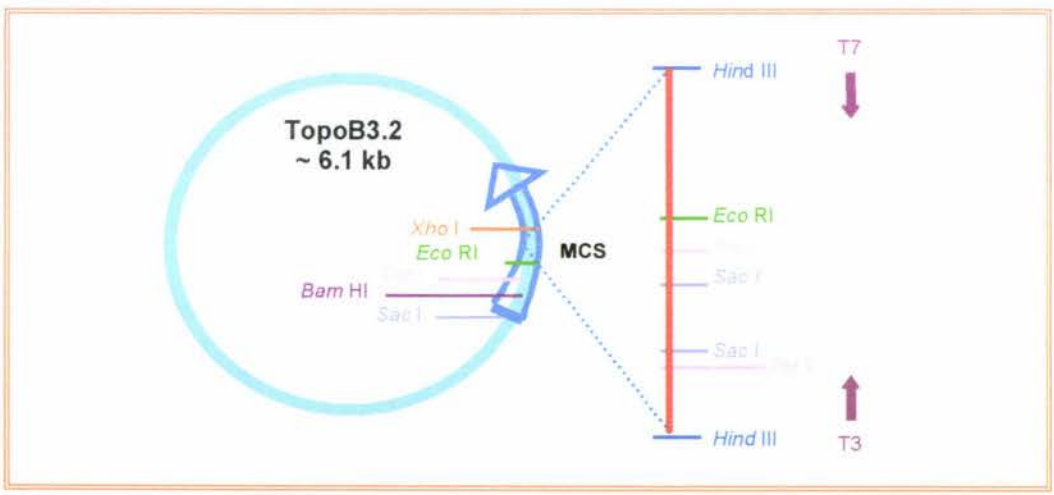


Figure 3.7: Schematic representation of the restriction map of TopoB3.2.
TopoB3.2 was digested with various restriction endonucleases and the resulting fragment sizes were used to generate a restriction map. The insert (red) had been cloned into pBluescript SK (+/-) using the *Hind* III site within the multiple cloning site (MCS, blue arrow), therefore the insert was flanked by two *Hind* III sites (blue). The TopoB3.2 insert contains two *Pst* I and *Sac* I sites along with one *Eco* RI site. The map of the MCS of pBluescript (appendix 2) and T3 sequencing data also helped generate the above map.

3.3.2 Primer walking to sequence topoisomerase IIβ clone TopoB3.2.

As the TopoB3.2 vector contained approximately 2.5 kb of insert DNA, primer walking was utilized to sequence the entire insert. Primer walking is a valuable tool when attempting to sequence DNA fragments that are larger than the standard 500 bp that can be sequenced from one template, using automated DNA sequencing.

To begin with, sequence data was obtained using a standard primer, which hybridizes to the vector sequence situated upstream of the insert DNA. Once the initial sequence was obtained, it could then be used to design a custom oligonucleotide. This custom oligonucleotide was synthesized so that it can prime the second sequencing reaction by annealing and enabling extension in the desired direction. By repeating the cycle of designing oligonucleotides based on the previous sequence, all the acquired DNA fragments were organized and assembled based on the overlapping regions of DNA between each primer used (as described in section 2.2.9). Thus, generating a contig of individual sequence data. Beginning with the second standard primer that primes in the opposite direction and following the same strategy, a set of primers were designed to enable sequencing of the complementary strand. Thus the 2.5 kb insert was sequenced on both strands using an ABI 377-18/36 automated DNA sequencer with Big Dye chemistry (Version 1.0, PE Biosystems), carried out at the Massey University sequencing service (Lorraine Berry; MUSeq facility, Institute of Molecular BioSciences, Massey University).

The initial standard primer used to sequence Topo3.2 was T3 and each sequencing reaction generally gave rise to approximately 500-600 bases of good quality sequence (refer to appendix 3.2 for sequences). Accordingly, for the 2.5 kb insert 2 primers were designed consecutively used in sequencing reactions with the TopoB3.2 template (Tbeta1 and Tbeta2 see appendix 1.1 for primer details). The final sequence obtained from TopoB3.2 was generated by sequencing from the opposite pBluescript primer, T7.

3.3.3 Generating a contig of topoisomerase IIβ sequences.

Any ambiguous sequence data from the TopoB3.2 sequencing reactions was manually removed from either end and was then assembled into a contig using the fragment assembly system (FAS) within GCG (Version 9.1; Wisconsin Genetics Computer Group, USA). The fragment assembly system recognizes overlapping sequences from individual sequence data

and can generate an aligned assembly of each sequence. Individual programs within the FAS enable this system to function.

Each of the four TopoB3.2 fragments were entered into the FAS (as described in section 2.2.10), and a contig 2452 bases in size was generated. Any discrepancies in bases were corrected manually, by carefully examining the sequence chromatograms, any discrepancies in bases were corrected manually, and this contig (TopoB3.2corrected, refer to appendix 3.2 for sequence) was analyzed using other programs within GCG or alternative sequence analysis software.

3.3.4 Analysis of TopoB3.2corrected sequence.

Once the entire TopoB3.2 insert had been sequenced, it was analyzed to confirm its identity and locate the putative 5' region of the topoisomerase IIβ promoter.

In order to establish the identity of the TopoB3.2corrected sequence, it was subjected to a BLAST (Basic Local Alignment Search Tool) through the NCBI website (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/BLAST/). There was only one significant match identified, which was to the human topoisomerase IIβ gene (exon 2), and turned out to be the match that was identified earlier to exon 2 (refer to figure 3.5). GCG (Version 9.1; Wisconsin Genetics Computer Group, USA) was used to examine whether exon 1 from the human topoisomerase IIβ gene (accession number U65315) was contained within the TopoB3.2corrected sequence. The bestfit program revealed that exon 1 did not appear within the contig sequence.

Due to the fact that exon 2 was located within the TopoB3.2corrected sequence, but not exon 1 the assumption was made that the TopoB3.2corrected sequence could not contain the 5' region of topoisomerase IIβ promoter and the sequence was likely to be mainly intron 1. Using GCG, a bestfit search against the published 5' topoisomerase IIβ sequence (Ng *et al*, 1997, accession number U65315) was also carried out, and this demonstrated that the topoisomerase IIβ promoter sequence was not contained within the TopoB3.2 insert (refer to figure 3.8).

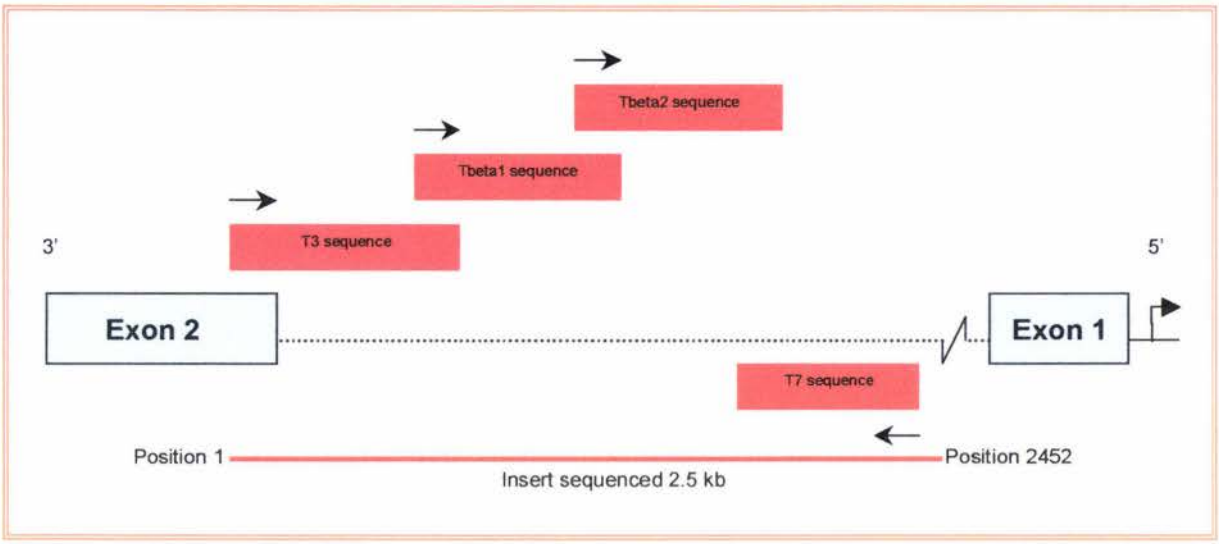


Figure 3.8: Schematic representation of the TopoB3.2 insert DNA.

The 2452 bases sequenced from the insert of TopoB3.2, was aligned with the human topoisomerase IIβ exons. The 'Gelview' program (within FAS, GCG) provided a schematic representation of each TopoB3.2 fragment in correlation to each other, thereby establishing the way each sequence stacked into the contig. The insert was sequenced from the 3' end going into the 5' direction, but the sequence fell short of exon 1 and the promoter (arrow) region. Even though the sequence mapped to a significant region of exon 2, most of the insert DNA appeared to be intronic DNA (dotted black line). ⁄ Signifies a region of intronic DNA that is undefined in length.

In order to confirm that the TopoB3.2corrected sequence had in fact been sequenced correctly, GCG was used to generate a map of theoretical restriction sites (refer to appendix 3.3). This was done using the Map program and the enzymes that were chosen were those used in the restriction digest(s) from section 3.3.1. A representation of the results can be seen in figure 3.9. Significantly, each theoretical restriction site would generate fragments that matched those that were observed in the actual restriction digests and consequently the restriction map that was constructed (figure 3.7) was confirmed to be correct. This analysis suggested that there was only one *Eco* RI restriction site within the insert, two *Pst* I sites and two *Sac* I sites. Therefore, the TopoB3.2corrected sequence was representative of the original clone.

Further work with this clone was discontinued and the search for the topoisomerase IIβ promoter was continued with lambda clones that had been isolated from a genomic library. This work is described in the following section.

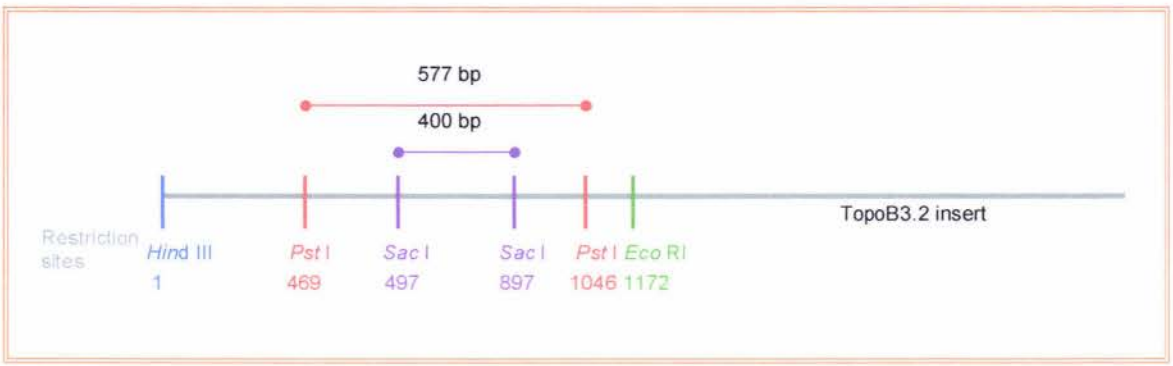
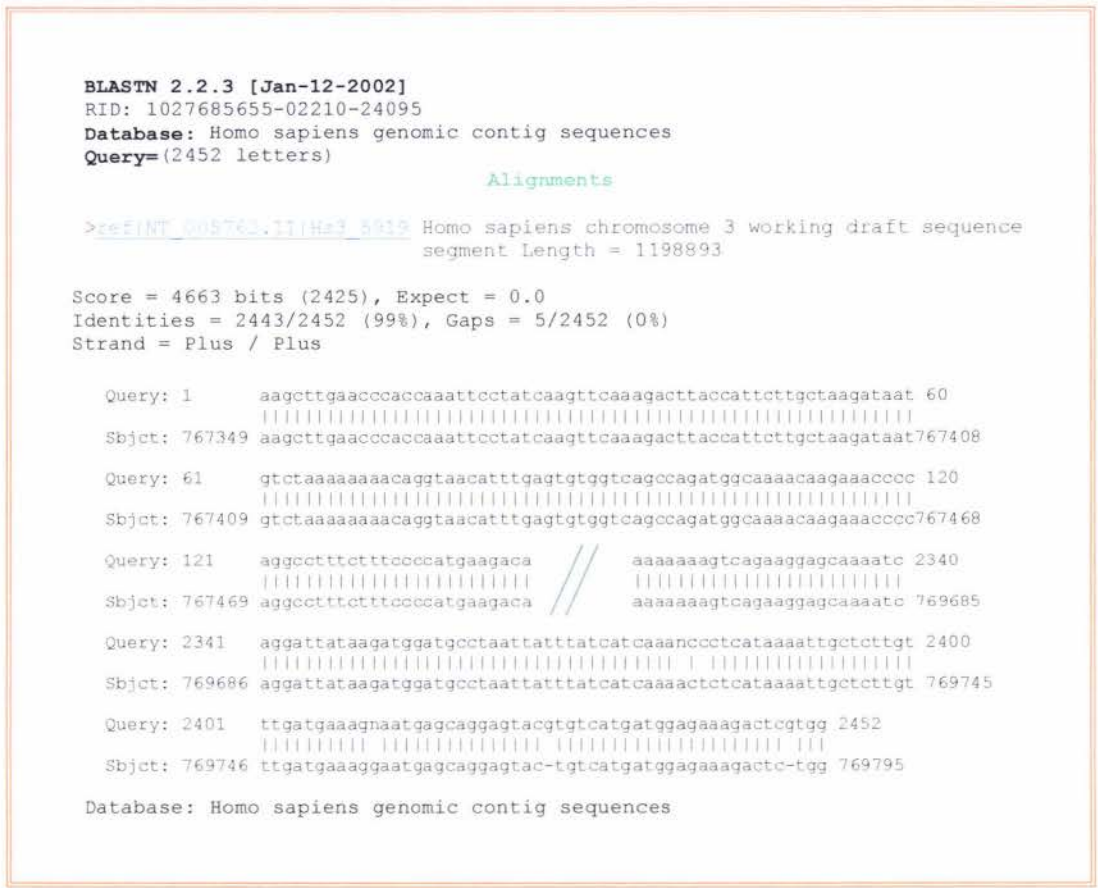


Figure 3.9: Schematic representation of restriction sites within TopoB3.2 insert. Theoretical restriction sites (coloured lines) within TopoB3.2corrected were estimated using the Map program within GCG (Version 9.1; Wisconsin Genetics Computer Group, USA). The *Hin*d III restriction site represents the T3 primer side of the MCS within pBluescript SK (+). Each restriction site would generate fragments of identical sizes to those observed in the actual restriction digests (section 3.3.1, appendix 3.1).

Well after the work with topoisomerase IIβ had been abandoned, the draft human genome sequence was published (NCBI database). Therefore, the TopoB3.2corrected sequence was subjected to a BLAST search (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html), but this time against the working draft sequence of the human genome. The entire 2452 base sequence matched with high identity to the human genome sequence that mapped to chromosome 3, within the region that contains the topoisomerase IIβ gene. It was found that the TopoB3.2corrected sequence spanned a region of exon 2 to intron 1 of the human topoisomerase IIβ sequence, which corresponded to nucleotides 767349 to 769795 (contig ID: NT 005762.11) of the working draft human genome sequence (refer to figure 3.10). This confirmed that the TopoB3.2corrected sequence was indeed a sequence from the topoisomerase IIβ gene, but unfortunately was not the 5' promoter region required.



```
BLASTN 2.2.3 [Jan-12-2002]
RID: 1027685655-02210-24095
Database: Homo sapiens genomic contig sequences
Query=(2452 letters)
                                Alignments

>ref|NT_005763.3|Hs3_5919 Homo sapiens chromosome 3 working draft sequence
                          segment Length = 1198893

Score = 4663 bits (2425), Expect = 0.0
Identities = 2443/2452 (99%), Gaps = 5/2452 (0%)
Strand = Plus / Plus

  Query: 1      aagcttgaacccaccaaattcctatcaagttcaaagacttaccattcttgctaagataat 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Sbjct: 767349 aagcttgaacccaccaaattcctatcaagttcaaagacttaccattcttgctaagataat767408

  Query: 61     gtctaaaaaaaacaggtaacatttgagtgtggtcagccagatggcaaaacaagaaacccc 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Sbjct: 767409 gtctaaaaaaaacaggtaacatttgagtgtggtcagccagatggcaaaacaagaaacccc767468

  Query: 121    aggcctttctttccccatgaagaca  //  aaaaaaagtcagaaggagcaaaatc 2340
                |||||||||||||||||||||||||      ||||||||||||||||||||||||
  Sbjct: 767469 aggcctttctttccccatgaagaca  //  aaaaaaagtcagaaggagcaaaatc 769685

  Query: 2341   aggattataagatggatgcctaattatttatcatcaaanccctcataaaattgctcttgt 2400
                |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
  Sbjct: 769686 aggattataagatggatgcctaattatttatcatcaaaactctcataaaattgctcttgt 769745

  Query: 2401   ttgatgaaagnaatgagcaggagtacgtgtcatgatggagaaagactcgtgg 2452
                ||||||||| |||||||||||||| |||||||||||||||||||| ||| |
  Sbjct: 769746 ttgatgaaaggaatgagcaggagtac-tgtcatgatggagaaagactc-tgg 769795

Database: Homo sapiens genomic contig sequences
```

Figure 3.10: Summary of the BLAST alignment using the human genome database.

The TopoB3.2corrected sequence was analyzed against the human genome using NCBI, human genome BLAST resources. The top rows (query) are the TopoB3.2corrected sequence and the bottom rows (subject) are the human genome references. This region of the human genome is mapped to chromosome 3, within the vicinity of the topoisomerase IIβ gene 3p24. // Represent sequences that were identical, but were not shown for simplicity.

3.4 Topoisomerase IIβ lambda clones.

Prior to the publication of the draft human genome sequence, a second attempt was made at isolating the topoisomerase IIβ promoter. Four lambda (λ) clones isolated from an original genomic library, which had been screened with a topoisomerase IIβ exon 2 probe were obtained from Dr. Sally Davies (Institute of Molecular Medicine, University of Oxford). As these clones were reputed to contain larger inserts than the pBluescript clones initially investigated, it seemed likely that one of these four λ clones might contain the topoisomerase IIβ promoter region required.

Very little information about the clones was provided; all that was known was that there were four topoisomerase IIβ clones, each in bacteriophage lambda (λ) T2-E4 phage (personal communication Sally Davies, Institute of Molecular Medicine, University of Oxford). It was assumed that this particular λ derivative was a standard library vector and that the *Eco* RI site had been used for cloning.

Each λ phage clone was quantified by gel electrophoresis, and found to have very low DNA concentrations. In fact the DNA concentrations were so low they were difficult to accurately gel quantify. λclone1 appeared to have a DNA concentration less than 5ng/μL, whereas λclone2, 3 and 4 were even lower than that (data not shown). As there was insufficient DNA to begin analysis, each λ DNA sample was packaged and then used to reinfect *E.coli*. This allowed sufficient quantities of DNA to be prepared from λ lysates (described in section 3.4.2), which enabled subsequent analysis of these clones.

3.4.1 Bacteriophage lambda.

Bacteriophage λ is a virus that infects bacterial cells such as *E.coli*. They have the ability to destroy bacterial cells which is demonstrated by a cleared area (plaque) in a colony (lawn) of bacterial growth. Bacteriophage λ consists of a head in which the viral chromosomal DNA resides and a tail by which the virus attaches to the host bacterial cell.

Virion DNA is linear and double-stranded (approximately 50 kb), divided into 3 regions; the left arm, the right arm and a central region. The central region within the λ phage genome is considered non-essential for phage growth and can be deleted or replaced; this is where the DNA of interest is placed (Chauthaiwale *et al*, 1992). Each arm is associated with single-stranded complimentary termini (cohesive (*cos*) termini) 12 nucleotides in length (refer to

figure 3.11A). Most of the essential genes required for λ phage structure and function reside within the left and right arms of the genome. The left arm generally codes for head and tail proteins, whereas the right arm contains genes involved in gene regulation. The head acts as a protective coating, in which the nucleic acid resides and the tail is a hollow tube through which the nucleic acid passes during infection (refer to figure 3.11B).

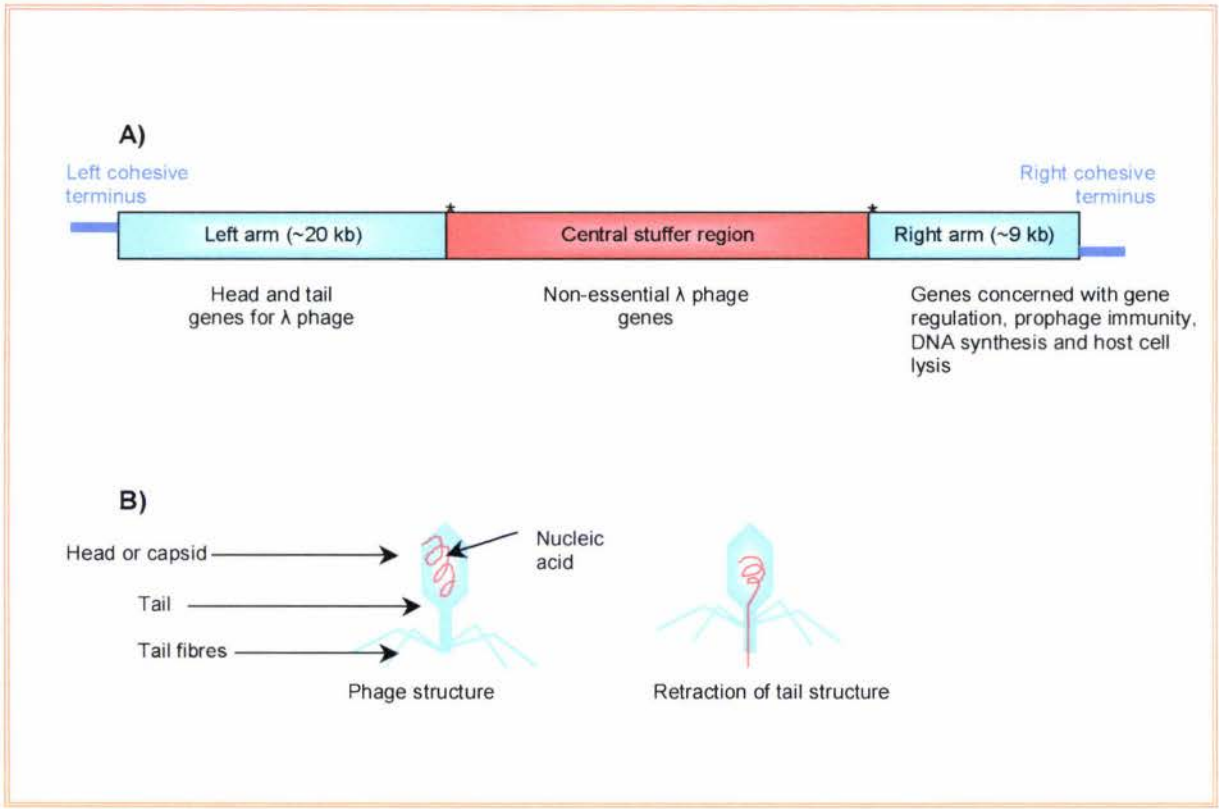


Figure 3.11: Schematic representation of bacteriophage λ cloning vector.

A) The double-stranded virion DNA: The central region is replaced with the DNA insert in a recombinant phage using restriction sites on either side of the central region (*). The left and right arms serve different purposes and both are essential for phage function. λ phage termini are flanked by short single-stranded (12 nucleotides, blue line) projections, which are complimentary in sequence. **B)** Schematic representation of λ phage structure: λ phages are comprised of two main structures, the head within which the nucleic acid (red) resides and the tail structure, which feeds the nucleic acid through to the host bacterium. The tail structure retracts once the phage attaches to the host cell, as the nucleic acid is inserted into the host cell. Tail fibers are involved in the binding of the phage to the bacterial host cell.

Bacteriophage λ can adopt one of two replication pathways, lytic or lysogenic (refer to figure 3.12). In the lysogenic pathway, phage gene expression is repressed and the circular chromosome inserts into the bacterial chromosome by site-specific recombination. In the lytic pathway, phage genes encoding replication, lysis and virion proteins are expressed, the chromosome replicates and the replicas are cleaved and packaged into progeny phage particles, which are then released by destroying the host cell. Exploiting the lytic pathway was the means used to amplify each of the T2-E4 phage samples provided. Understanding the bacteriophage replicative cycle was imperative in understanding each of the procedures used to amplify the topoisomerase IIβ insert DNA.

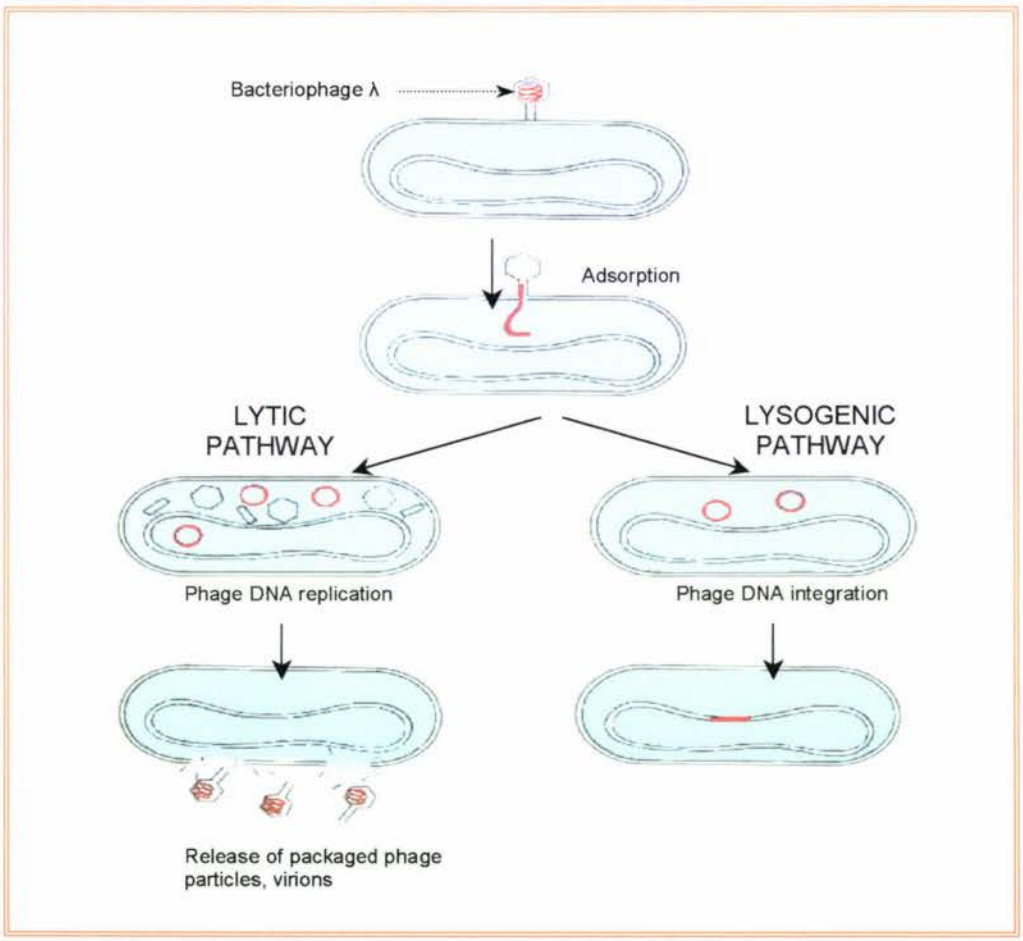


Figure 3.12: Schematic representation of the bacteriophage λ life cycle.
Bacteriophage λ attaches to its host cell using its tail fibers. The host cell is infected with linear viral λ DNA, where it circularizes and replicates inside the host cell. The lytic pathway leads to death of the host cell due to the release of packaged viral DNA, whereas in the lysogenic pathway the host cell remains intact.

3.4.2 Packaging phage and isolating phage DNA.

In order to generate infectious λ phage particles, lambda DNA must be concatamerised prior to packaging into viral capsids *in vitro*. Bacterial cells are then infected by newly created infectious particles and λ phage DNA can be extracted.

Each λ phage DNA sample was concatamerised using T4 ligase as described in section 2.2.19(i) and shown in table 3.4 below. Due to the inability to accurately quantify λ clones approximate amounts of DNA (about 50 ng) were used in ligations.

Sample	T4 Ligase	Ligation Buffer	Amount λ DNA (dried)	H_20	Total Volume
λclone1	1 μL	2 μL	20 μL	7 μL	10 μL
λclone2	1 μL	2 μL	40 μL	7 μL	10 μL
λclone3	1 μL	2 μL	40 μL	7 μL	10 μL
λclone4	1 μL	2 μL	80 μL	7 μL	10 μL

Table 3.4: Ligations of λ phage topoisomerase IIβ clones.

Approximately 50 ng of each λ clone was ligated with T4 ligase to enable packaging of the phage DNA. Due to the low concentrations of each λ clone, the liquid volume of each λ clone was dried to a pellet in a speedvac (speed vacuum concentrator). The exact concentration of each λ clone was not known; therefore, each sample was ligated and packaged consecutively, starting with λclone1.

Each sample was prepared as shown in the table above and incubated overnight at 16°C to ensure ligation. 5 μL (about 25 ng) of each ligation was packaged with Packagene® Lambda DNA Packaging System (Promega).

Each ligated λ clone was then subjected to packaging using Packagene® Lambda DNA Packaging System (Promega) and plated as described in section 2.2.19(iv). The packaging efficiency could not be calculated for the λ DNA samples, as the exact amount of DNA that was packaged was not known. However, the packaging efficiency for the control DNA was about 1.4×10^7 recombinants/μg DNA packaged (refer to table 3.5). This packaging efficiency was low compared to the expected packaging efficiency (5×10^7- 5×10^8 recombinants/μg DNA, Promega product guide). Never the less, it was obvious that the

packaging efficiencies of the λ samples were very low, as very few plaques formed. This may have been due to the low DNA concentrations of the λ DNA samples.

Sample	Volume plated (dilution)	Number of Plaques formed
λclone1	100 μL	4
λclone2	100 μL	85
λclone3	100 μL	3
λclone4	100 μL	4
Control DNA	100 μL (10^4)	748

Table 3.5: Packaging λ phage topoisomerase IIβ clones.

5 μL of each ligation was packaged with half a reaction (25 μL) of the Packagene® Lambda DNA Packaging Extract (Promega). 25 ng of control (λ c1857) DNA was also packaged to monitor packaging efficiency. 100 μL of extract was plated with 100 μL of plating cells and incubated overnight at 37°C.

The following day the number of plaques formed were scored and used to calculate the packaging efficiency or phage lysates. Individual plaques were used to generate phage lysates, used in λ phage preparations.

Individual plaques were isolated and placed into 500 μL of SM buffer to enable phage particles to diffuse into the buffer. This phage lysate was then titered (described in section 2.2.19(vi)) to determine the number of plaque forming units (pfu)/mL. In order to isolate the λ phage DNA, medium-scale λ preparations (described in section 2.2.19(vii)) were done where confluent-lysis plates were required (refer to table 3.6).

Generally, medium-scale λ preparations produced approximately 1-2 μg DNA, which provided sufficient DNA to use in restriction endonuclease digests and Southern blots. The λ lysates provided a rich source of phage particles and were used again to generate more λ DNA if required. Accordingly, the titer of λ phage lysates was determined prior to each subsequent medium-scale λ preparation.

Sample	Dilution	Volume plated	Number of plaques	Phage lysate pfu/ml	Volume plated for λ preparation (1x 10^5 pfu/plate)	DNA concentration after λ preparation	Yield of λ DNA
λclone1	10^2	0.01 mL	64	6.4 x 10^5	156 μL	50 ng/μL	2.5 μg
λclone2	10^2	0.01 mL	185	1.9 x 10^6	54 μL	30 ng/μL	1.5 μg
λclone3	10^2	0.01 mL	106	1.06 x 10^6	94 μL	20 ng/μL	1.0 μg
λclone4	10^2	0.01 mL	270	2.7 x 10^6	37 μL	40 ng/μL	2.0 μg

Table 3.6: Titering phage particles from λ phage lysates.

Representation of the pfu/mL generated from a selection of phage lysates and the corresponding amounts of λ DNA extracted from concomitant medium-scale λ preparations.

Phage lysates were diluted and 10 μL plated with 100 μL of plating cells (LE392), incubated overnight at 37°C and the pfu/mL calculated. The phage lysate (with 100 μL LE392 plating cells) was then used to plate for confluent lysis (1 x 10^5 pfu/plate) on agarose plates and DNA extracted using medium-scale λ preparation.

3.4.3 Topoisomerase IIβ insert DNA within the λ clones.

Following packaging and λ DNA extraction, each clone was analyzed for the presence of insert DNA. As nothing specific was known about the type of λ vector, the insert could not be analyzed easily by sequencing or PCR. Using the assumption that most derivatives of bacteriophage λ are generated to have all sites for *Eco* RI eliminated from the essential part of the viral genome (Sambrook and Russell, 2001) and that this enzyme is commonly used for cloning, each λ clone was digested with *Eco* RI. Approximately 200 ng of λ DNA was digested and examined by gel electrophoresis (figure 3.13). The identity of each λ DNA fragment was deduced using information commonly known about λ phage structure, digests and gel electrophoresis.

In order to estimate the sizes of the larger DNA fragments, the samples were also electrophoresed against a λ ladder (λ DNA digested with *Hin*d III). The preliminary digest using λclone2 shows that there are 5 bands (labelled *b-f*, lane 4, figure 3.13A) after digestion with *Eco* RI. According to the sizes of the fragments, band *c* (about 20 kb) was most likely to

be the left arm of the λ phage and band *e* (about 9 kb) was probably the right arm of the λ phage (refer to figure 3.11A). Therefore, band *b* (about 30 kb) was most likely to be right arm and left arm joined through the *cos* sites. The smallest band, band *f* (about 3kb) was probably the insert DNA, therefore band *d* (about 12 kb) was probably the right arm and insert DNA joined together. Due to the sizes of each related λ DNA fragment, the uncut λclone2 DNA (band *a*) was estimated to be about 40 kb, however due to the lack of larger size markers this could not be clearly established from the gel.

The fragments identified from the digest described above were indicative of incomplete digestion. Consequently, when each of the λ DNA samples were digested with more enzyme, digests remained the same, thereby suggesting that the λ DNA must be religating after enzymatic digestion. λ DNA ends are flanked by *cos* sites, which could enable λ arms to re-associate after digestion. To improve the resolution of the larger λ DNA fragments, samples were analyzed on a 0.5% agarose gel at a lower voltage over a much longer period. The results of this are shown in figure 3.13B, with clearer separation of bands *b* and *c*. Each digest clearly shows that an insert DNA fragment (approximately 3 kb in size) was released from each λ clone and that the same 5 bands are seen in each digest (lanes 3, 4, 5 and 6, figure 3.13B). Bands in each of lanes 4-6 are identical, however there is an additional band present in lane 3 (λclone1) which was presumed to be uncut λ DNA (band *a*), as it has the same mobility as the uncut control (lane 2). The results of both λ digests are summarized in table 3.7 below.

Band	Size of band	Identity of band
a	About 40 kb	Uncut λ DNA
b	About 30 kb	Left arm + right arm of λ phage
c	About 20 kb	Left arm of λ phage
d	About 12 kb	Right arm of λ phage + insert DNA
e	About 9 kb	Right arm of λ phage
f	About 3 kb	Insert DNA

Table 3.7: Summary of the identity of DNA bands represented in λ DNA digests. The sizes of the DNA fragments were determined by comparing them to standard DNA size markers, which were electrophoresed with the λ DNA samples (seen in figure 3.13). The identity of the bands was determined by comparing them with the sizes of the right and left arms of λ phage as described in figure 3.11A.

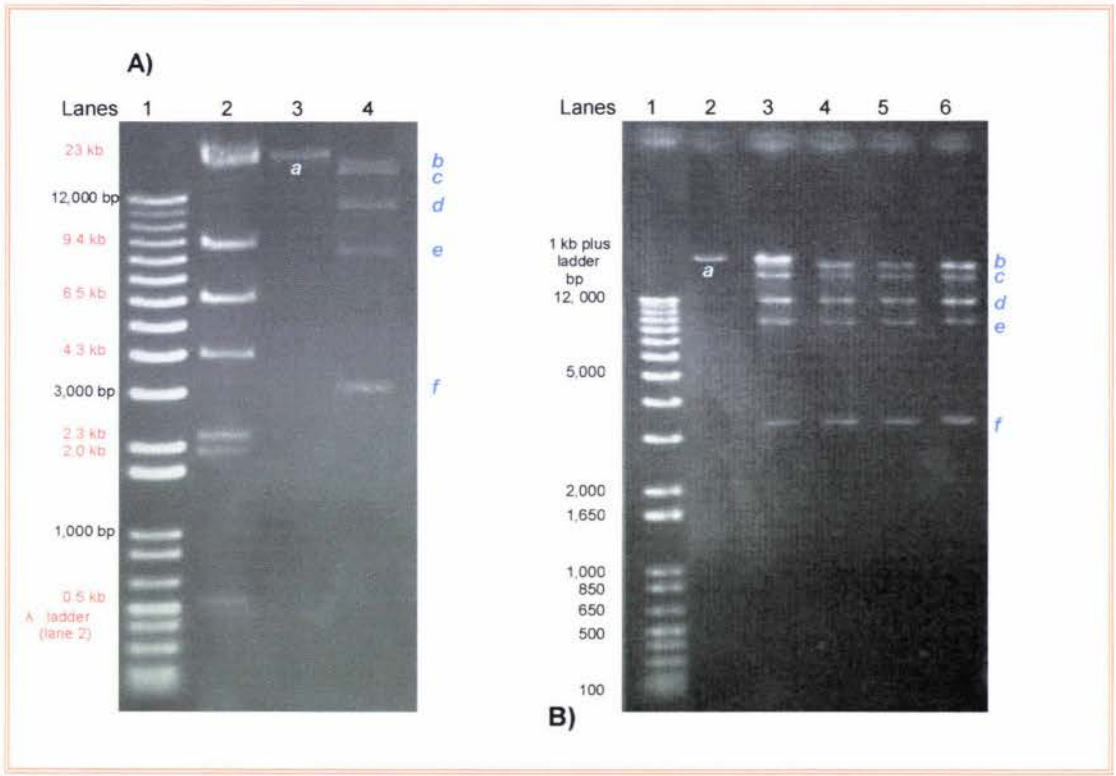


Figure 3.13: Topoisomerase IIβ lambda clones digested with *Eco* RI.

200 ng of each λ clone was digested with *Eco* RI and analyzed by gel electrophoresis. DNA was visualized by incorporating ethidium bromide (0.5 μg/mL) into the gel and exposure to UV light. Molecular sizes of bands illustrated by the ladders are indicated on the left of each figure; λ ladder is in red, 1 kb plus ladder is in black. Each of the λ DNA bands are labelled *a-f*.

A) λclone2 DNA samples were electrophoresed on a 1% agarose gel in 1xTAE buffer at 80V for about 1 hour.

Lanes	1	10 μL 1kb plus ladder
	2	10 μL λ DNA ladder (λ DNA digested with *Hin*d III)
	3	2 μL (20 ng) undigested λclone2 DNA
	4	10 μL *Eco* RI digested λclone2 DNA

B) All four lambda samples were electrophoresed on a 0.5% agarose gel in 1xTAE buffer at 40V for about 9 hours.

Lanes	1	10 μL 1kb plus ladder
	2	2 μL (40 ng) undigested λclone1 DNA
	3	10 μL *Eco* RI digested λclone1 DNA
	4	10 μL *Eco* RI digested λclone2 DNA
	5	10 μL *Eco* RI digested λclone3 DNA
	6	10 μL *Eco* RI digested λclone4 DNA

3.4.4 Southern Blots of topoisomerase IIβ λ clones.

Due to the size of the insert DNA and the fact that sequencing the insert could not be achieved without subcloning, it was decided to perform a Southern blot and hybridization of each λ clone (details are shown in figure 3.14). The PCR-generated topoisomerase IIβ promoter region topoIIβ1 (described in section 3.2.1) was used as a probe to detect for the presence of any putative topoisomerase IIβ promoter sequences within the insert DNA of the four λ clones. Southern blots were performed concurrently with subcloning the λ topoisomerase IIβ insert into pBluescript (section 3.4.5).

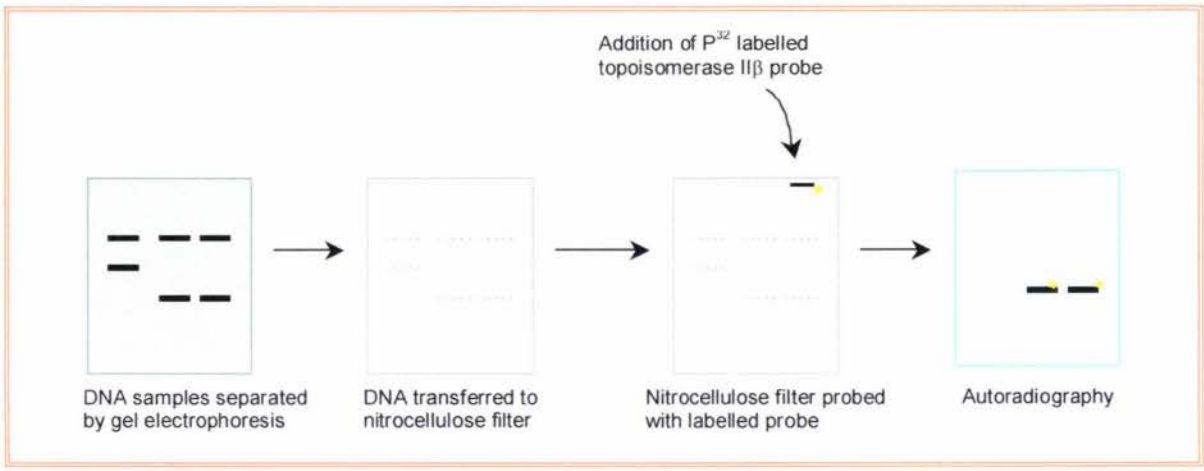


Figure 3.14: Schematic representation of topoisomerase IIβ Southern Blots.

A short topoisomerase IIβ promoter region was generated by PCR and used as a probe against DNA samples that were immobilized on a nitrocellulose membrane. If the DNA sample contained sequences similar to that of the probe, it will anneal and this interaction can be visualized by exposure to X-ray film.

Each λ clone was digested with *Eco* RI, subjected to electrophoresis (refer to figure 3.15A) and the DNA transferred to a nitrocellulose membrane as described in section 2.2.20(ii) and hybridized with a radioactively labelled probe. 25 µL of topoIIβ1 (50 ng) was labelled with [α^{32}P] as described in section 2.2.20(i) and incorporation of radioactivity was measured using Cerenkov counting. 1 µL of [α^{32}P]-labelled topoisomerase IIβ probe (diluted 1:100) generated a Cerenkov count of about 9,534 counts per minute (cpm), which correlates to a specific activity of 1.9 x 10^7 cpm/µL. This was sufficient labelling to carry on with the Southern hybridization. If any sequence similarities between the probe and λ insert DNA were present then hybridization would occur, which would be detected upon exposure to X-ray film (refer to figure 3.15B). Human genomic DNA was also digested with *Eco* RI and

used in the blots, as a positive control, to confirm that the topoisomerase IIβ probe was part of the human genome.

The bands in figure 3.15A and B were labelled *a-f* as in figure 3.13. These bands were identical to those identified in figure 3.13 signifying incomplete digests as before however, there is an additional faint band (labelled X) that can be seen between bands *c* and *d* in lane 2 only. This band X is very faint and is about 15 kb in size, which suggests that it could be composed of two 3 kb inserts and a right arm (9 kb) joined together. Band *g* is approximately 300 bp in size and is used as a control to detect for self-association of the topoIIβ1 probe. The human genomic digest (lane 6) showed that this sample digested well, giving a smear of smaller bands.

Figure 3.15B illustrates the autoradiograph of the hybridization blot. The probe hybridized to each of the bands that contained insert DNA (refer to table 3.7) in each of the λ clones, bands *a*, X, *d* and *f*. This indicated that each of the λ clones might have the putative topoisomerase IIβ promoter region. Unfortunately, the probe also showed non-specific hybridization to the human genomic digest (lane 6), producing a smear on the autoradiograph. This collectively suggests that the 300 bp probe may have been too small to efficiently identify putative topoisomerase IIβ promoter DNA in larger DNA fragments such as the 3 kb insert and fragments within the human genomic digest, or that the stringency of the hybridization was less than optimal. Thus, the topoisomerase IIβ probe may not have been specifically binding to putative topoisomerase IIβ promoter DNA.

From the digests shown in figure 3.13 and 3.15, it was surprising to find that the inserts were all about 3 kb and each hybridized to the topoIIβ1 probe. This size is generally considered small for an isolate from a human genomic library (usually about 9 kb), and it was odd that each insert was identical in size. At this point concerns were raised about the identity of these clones, therefore subsequent analysis was carried out one clone only, λ clone1. This clone was not chosen for any specific reason, it simply had a higher DNA concentration than the rest of the clones and appeared to digest consistently in reactions. Due to the inconclusiveness of the Southern blot, but the possibility that each of the λ clones may contain putative topoisomerase IIβ promoter DNA, the insert from λclone1 was subcloned into pBluescript SK (+) to facilitate sequencing of the insert.

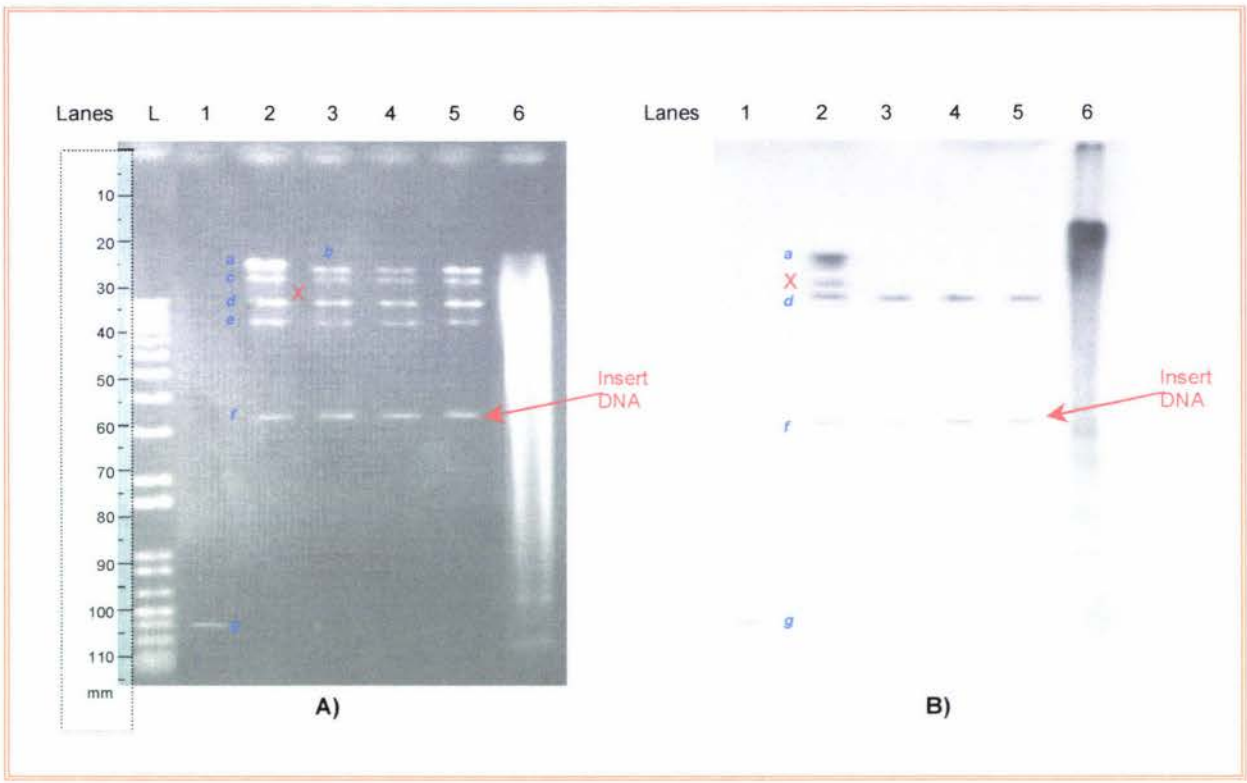


Figure 3.15: Southern Blotting λ clone insert DNA.

10 µL (200 ng) of each λ clone was digested with *Eco* RI and electrophoresed on a 0.5% agarose gel in 1xTAE buffer at 30V for about 12 hours. Each of the λ DNA bands are labelled *a-g*.

Lanes	L	10 µL 1kb plus ladder
	1	5 µL (2 ng/µL) topoIIβ1 probe (control)
	2	10 µL *Eco* RI digested λclone1 DNA
	3	10 µL *Eco* RI digested λclone2 DNA
	4	10 µL *Eco* RI digested λclone3 DNA
	5	10 µL *Eco* RI digested λclone4 DNA
	6	10 µL *Eco* RI digested human genomic DNA

A) Picture of the 0.5% agarose gel used to separate fragments of the λ DNA *Eco* RI digests and blot onto nitrocellulose membrane. Lane L, 10 µL 1 kb plus ladder. DNA was visualized by incorporating ethidium bromide (0.5 µg/mL) into the gel and exposure to UV light. Ruler (left of figure) was used to mark the distance traveled by each band, so that the corresponding bands on the autoradiograph could be mapped. An uncut λ DNA control was inadvertently omitted from the gel. **B)** Picture of the corresponding autoradiograph of the hybridized blot. Probe was [α^{32}P] labelled topoIIβ1. X-ray film was overlaid on the blot for approximately 12 hours at room temperature and developed using an automated X-ray film processor.

3.4.5 Insert from λclone1 placed into pBluescript SK (+).

The insert DNA from the λ clones was subcloned into a conventional vector to facilitate DNA sequencing. Using medium-scale λ preparation DNA, the insert of λclone1 was isolated by digesting approximately 500 ng of λ clone1 with *Eco* RI, gel extracting the 3 kb insert and quantified using gel electrophoresis. The concentration of insert DNA was found to be about 15 ng/μL.

500 ng of pBluescript SK (+) was also digested with *Eco* RI, phosphatased then quantified using gel electrophoresis to prepare it for ligations. The concentration of the phosphatased pBluescript SK (+) vector was found to be 10 ng/μL. Ligation reactions were performed (as illustrated in table 3.8) and *E.coli* XL-1 blue competent cells were transformed the following day as described in section 2.2.16. The results of the transformations are shown in table 3.9.

Sample	ng vector	ng insert	μL vector	μL insert	T4 buffer μL	H_20 μL	Total volume
1) Vector + ligase	50	-	5	-	4	10	20 μL
2) Vector – ligase	50	-	5	-	4	11	20 μL
3) 3:1 ratio insert:vector	50	150	5	10	4	0	20 μL

Table 3.8: Ligation reactions carried out to subclone presumptive topoisomerase IIβ promoter DNA into pBluescript SK (+).
The concentration of the phosphatased pBluescript vector was 10 ng/μL and the insert DNA was 15 ng/μL. Ligation reactions were carried out as described in section 2.2.14 and incubated overnight at 16°C. The following day *E.coli* XL-1 blue competent cells were transformed with half of each ligation.

Transformations appeared to be successful with no contamination and sufficient white colonies were generated to analyze for the presence of plasmid DNA. Blue colonies most likely were the result of vector self-ligation. One of the white colonies from the transformation was used to isolate plasmid DNA using a Quantum® Prep Plasmid Miniprep Kit (BioRad), and the identity of the plasmid (pBSTopoB1.26) was investigated, by restriction endonuclease digest with *Eco* RI (figure 3.16).

Reaction	Ligation mixture	XL-1 competent cells	Dilution	Volume plated	Number of colonies	
					White	Blue
XL-1 competent cells (control) LB plates	-	10 μL	10^7	50 μL	231	-
XL-1 competent cells (negative control) LB+amp	-	10 μL	Undiluted	50 μL	0	-
1) Vector minus ligase (control) LB plates	10 μL	100 μL	Undiluted	50 μL	0	50
2) Vector plus ligase (background control) LB plates	10 μL	100 μL	Undiluted	50 μL	0	13
3) Vector plus insert DNA (1:3) LB plates	10 μL	100 μL	Undiluted 1:10	100 μL 100 μL	105 1*	185 3

Table 3.9: Results of XL-1 transformations using pBluescript SK (+) ligations.
Transformation of XL-1 blue competent cells, using plasmids generated from ligation (table 3.8) of pBluescript SK (+) vector and λclone1 insert DNA, or control DNA. Unless indicated, all samples were plated onto LB plus ampicillin plates with 20 μL IPTG (20 mg/mL in water) and 50 μL X-gal (20 mg/mL). Plates were incubated overnight at 37°C and colonies scored the following day. Colonies were creamy white in colour or blue. One colony was selected and cultured to isolate plasmid DNA for analysis (*).

The plasmid was confirmed to contain a 3.2 kb insert (lane 2, figure 3.16). Sequencing of the insert was accomplished using both the T3 and T7 primers of the vector (Lorraine Berry, MUSeq facility, Institute of Molecular BioSciences, Massey University). The sequence obtained from T7-primed sequencing reaction was of very poor quality; even when the sequencing reaction was repeated the result did not improve. Therefore, the sequence produced from the T3 primer was analyzed using both GCG and NCBI resources (refer to appendix 3.4, for T3 sequence).

No significant matches were found between the λclone1 insert and either exon 2 or exon 1 of the human topoisomerase IIβ gene, nor the sequence reported by Ng *et al* (1997). A BLAST search against the human genome sequence revealed that the putative topoisomerase IIβ insert was in fact a sequence that belonged to a region of chromosome 16. It has been well documented that the topoisomerase IIβ gene resides on chromosome 3 (Tan *et al*, 1992), therefore the λ insert DNA did not even belong to any region of the topoisomerase IIβ gene. All subsequent examinations of the topoisomerase IIβ λ clones were abandoned.

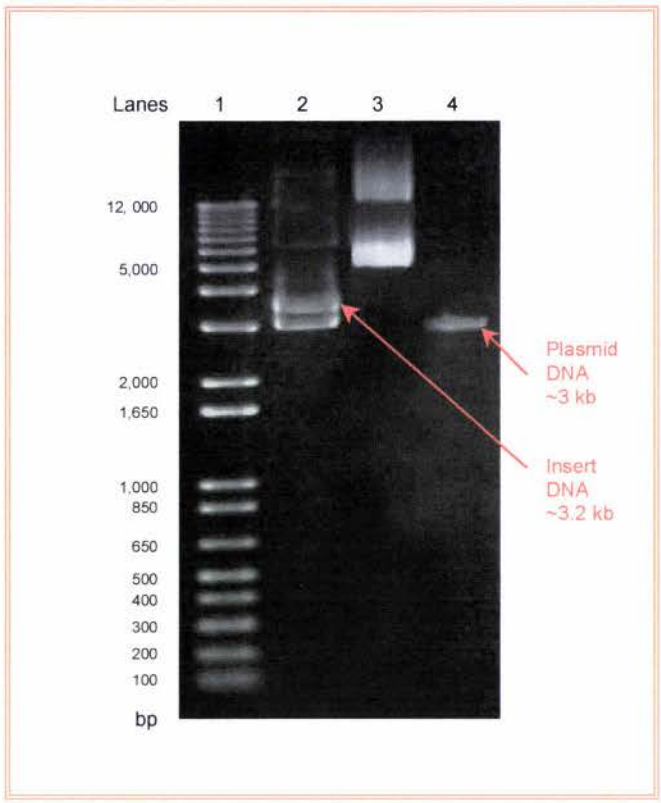


Figure 3.16: Digesting pBSTopoB1.26 with *Eco* RI.

5 µL of pBSTopoB1.26 (50 ng/µL) was digested with 1 µL (1U/µL) *Eco* RI and electrophoresed on a 1% agarose gel in 1xTAE buffer for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 µg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

Lanes	1	10 µL 1kb plus ladder
	2	10 µL pBSTopoB1.26 digested with *Eco* RI
	3	2 µL uncut pBSTopoB1.26
	4	1 µL PCR pBluescript SK (+) digested with *Eco* RI

Digesting the pBluescript TopoB1.26 clone with *Eco* RI produced two fragments, each approximately 3 kb in size. The smaller fragment was the vector DNA, as it is identical to the band seen in lane 4. Therefore, the larger band in lane 2 was insert DNA. Lane 3 shows the different conformational states of the circular uncut plasmid.

3.5 Acquisition and analysis of the human topoisomerase IIβ promoter sequence.

Due to failings in obtaining an actual DNA fragment containing the topoisomerase IIβ 5'-flanking region, an *in silico* approach was taken. With the advent of the human genome sequence being available online, approximately 2-3 kb of 5'-flanking topoisomerase IIβ sequence was analyzed, using bioinformatics tools (described in section 2.2.11).

3.5.1 Obtaining a 5'-flanking sequence for the topoisomerase IIβ gene.

In order to generate the required sequence, firstly, a BLAST search of the human genome (http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html) was carried out using the published topoisomerase IIβ sequences (Ng *et al*, 1997). The resulting sequence match was to a region of human chromosome 3. The links were followed and an appropriate sized sequence was downloaded (the source of this sequence is outlined in figure 3.17).

```
LOCUS       NT_005762              3556 bp    DNA     linear
DEFINITION  Homo sapiens chromosome 3 reference genomic contig.
ACCESSION   NT_005762 REGION: 859185..862740
VERSION     NT_005762.12  GI:22044654
  ORGANISM  Homo sapiens
  AUTHORS   NCBI Annotation Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2002) National Center for Biotechnology
            Information, NIH, Bethesda, MD 20894, USA
COMMENT     GENOME ANNOTATION REFSEQ:  NCBI contigs are derived from
            assembled genomic sequence data. They may include both draft and
            finished sequence.
     source          /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-659P16"
                     /note="Accession: AC093416 sequenced by University of
                     Washington Genome Center"
```

Figure 3.17: Origin of the topoisomerase IIβ sequence obtained from NCBI.

The published 5'-flanking sequence of topoisomerase IIβ (Ng et al, 1997) was used as 'bait' to identify an appropriate 5' region of the topoisomerase IIβ gene. Above are the relevant details of the sequence downloaded. The accession number relating the human genome sequence is shown in red, and the corresponding contig that the sequence originated from is shown in blue.

(Adapted from: http://www.ncbi.nlm.nih.gov/genome/seq/HsBlast.html).

This human topoisomerase IIβ sequence was truncated at about 2500 bases upstream of the start codon (RIDtopoIIB) and analyzed using a range of sequence analysis tools as follows.

The topoisomerase IIβ 5'-flanking sequence reported by Ng *et al* (1997) was mapped to the RIDtopoIIB sequences using the 'bestfit' command within GCG, revealing that the RIDtopoIIB sequence was the minus strand (refer to figure 3.18A). The RIDtopoIIB sequence was inverted using the FAS programs within GCG (Version 9.1; Wisconsin Genetics Computer Group, USA). This was achieved by using the Gel assemble programme (GCG) with both the topoisomerase IIβ sequence reported by Ng *et al* (1997) and the RIDtopoIIB sequence and generating a contig of the two fragments, this in turn forced the orientation of the RIDtopoIIB sequence to be altered (refer to figure 3.18B). The newly altered sequence (RIDtopoIIB+) was then used for subsequent analysis (refer to figure 3.19).

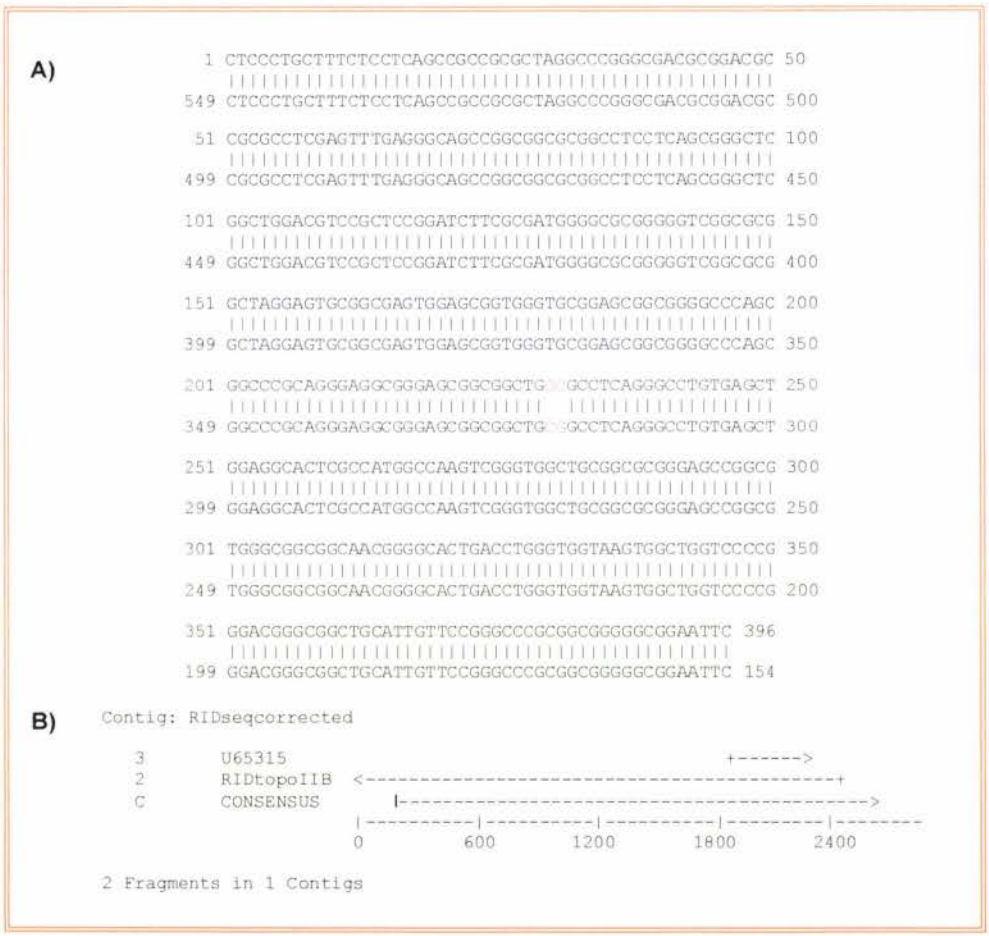


Figure 3.18: Mapping the sequence reported Ng *et al* (1997) to the RIDtopoIIB sequence.
A) The sequence reported by Ng *et al* (1997) (top rows) was mapped to the RIDtopoIIB sequence (bottom rows) using GCG (bestfit program). The above alignment illustrates that the RIDtopoIIB is in the 3' to 5' orientation. **B)** The orientation of RIDtopoIIB was corrected using GCG (FAS programs). The GELVIEW Fragment Assembly display of the contig generated from the sequence reported Ng *et al* (1997) (accession number U65315) and the RIDtopoIIB sequence, to create the inverted consensus sequence (RIDtopoIIB+).

3.5.2 Analysis of the RIDtopoIIB+ sequence.

The RIDtopoIIB+ sequence identity and orientation was checked once again, by using the bestfit program (GCG) to map against the topoisomerase IIβ promoter sequence (accession number U65315) and carrying out a BLAST search using NCBI resources. Both exercises confirmed that the RIDtopoIIB+ sequence was most likely topoisomerase IIβ 5' flanking sequence, however there was a discrepancy with the topoisomerase IIβ promoter sequence deposited by Ng *et al* (1997). In that, two bases G and C were actually C and G (refer to sequences in red in figure 3.18) in the RIDtopoIIB+ sequence, this discrepancy can only be addressed once actual sequencing of the topoisomerase IIβ promoter region is carried out. The above sequence analyses also enable the major transcription start site, exon1 and ATG codon to be mapped to the RIDtopoIIB+ sequence (refer to figure 3.19).

Following GCG analysis, the RIDtopoIIB+ sequence was entered into databases that can search for possible transcription factor binding sites within the sequence. These databases work by looking for sequence similarities and patterns. Three separate searches were carried out; the first was using the findpatterns program within GCG, secondly a search using the transcription factor sites accessed by Bionavigator (Entigen, www.BioNavigator.com) and finally a search for putative transcription factor binding sites using TRANSFAC database (MatInspector V2.2 (Quandt *et al*, 1995), http://transfac.gbf.de/cgi-bin/matSearch.pl). Each search came back with copious possibilities, which were sorted manually to identify any relevant sites (refer to appendix 3.5 for TRANSFAC database results).

The RIDtopoIIB+ sequence appears to be highly GC rich (about 80%) with many possible transcription factor sites. A total of 12 putative NF-Y binding sites exist, through the presence of five CAAT boxes sequences, 3 perfect inverted CAAT boxes and 4 more related CAAT sequences. Most of which appear to cluster around 500-750 nucleotides away from the major transcription start site. There are many possible Sp1 binding sites (GC boxes), which appear to congregate closer to the ATG start codon and transcriptional start site (position 1942 on sequence) than further upstream (refer to figure 3.19). These findings are merely speculations, as without functional assays the important regulatory elements of topoisomerase IIβ cannot be assessed.

Figure 3.19: The RIDtopoIIB+ sequence.

This sequence was analyzed for putative transcription factor sites and the significant results are illustrated above. The sequence is highly GC rich within which, there are five CAAT boxes sequences (pink), 3 perfect inverted CCAAT boxes (red) along with 4 more related CAAT sequences (purple) and several putative Sp1 binding sites (blue). The ATG codon is underlined, exon 1 (bold) of the topoisomerase IIβ gene maps to nucleotides 2135-2203 and the major transcription start site maps to 1941 (green). The Ng *et al* (1997) topoisomerase IIβ promoter sequence maps to nucleotides 1872-2267.

3.6 Chapter summary.

Two pBluescript and 4 λ clones isolated by screening a human genomic library with an exon 2 human topoisomerase IIβ probe were analyzed for the presence of topoisomerase IIβ promoter DNA. Unfortunately, none of these clones contained the 5'-flanking sequence for topoisomerase IIβ. However, the 5'-flanking sequence was finally identified using the human genome sequence that was available through NCBI and analyzed for putative transcription factor binding sites.

The 5'-flanking sequence of human topoisomerase IIβ revealed many putative transcription factor binding sites, which appeared to cluster together in regions upstream and downstream of the major transcription start site. However, the functional relevance of these putative transcription factor binding sites can only be established by carrying out functional assays, using deletion series of the 5'-flanking sequence of human topoisomerase IIβ. The publication and analysis of this sequence could now be a starting point for a detailed functional analysis of the topoisomerase IIβ promoter (see discussion, section 6.1.2).

Chapter 4: ICB1 and GC1 binding assays of the topoisomerase IIα promoter.

4.1 Introduction.

It has been suggested that the transcription factors Sp1 and NF-Y bind to two proximal elements of the topoisomerase IIα promoter, GC1 and ICB1 respectively. In order to characterise the binding of transcription factors to these elements, electrophoretic mobility shift assays (EMSA) were carried out using sequence specific probes and HeLa cell extracts.

The EMSA technique is used to investigate the binding reaction(s) between proteins and DNA or RNA *in vitro*. It utilizes the phenomenon that a DNA probe/fragment will be bulkier and less mobile when bound to protein, than just DNA alone. To assess the binding of DNA to protein, one strand of a double-stranded oligonucleotide probe was radioactively labelled with γ-^{32}-P-[ATP] and mixed with cell extract. After the binding reaction, the sample was subjected to electrophoresis on a low ionic strength polyacrylamide gel and exposed to X-ray film. In the presence of a DNA-protein interaction the oligonucleotide was retarded (refer to figure 4.1) due to its lowered mobility. The specific protein(s) involved were identified with antibodies and sequence specificity of the interactions was elucidated by generating mutations in the double-stranded oligonucleotide probes, which were used in competitor assays.

Competitor assays involved the addition of unlabelled DNA oligonucleotides in several times molar excess, which competed for protein binding with the labelled DNA. If the competitor preferentially bound to a protein, less protein was available to bind to the labelled probe and a decrease in band(s) intensity was observed.

The binding reactions are sensitive to factors such as DNA concentration, salt concentration, protein concentration, temperature and stability of the DNA-protein complex itself. Therefore, appropriate binding conditions were required to ensure optimal and consistent results. Each experiment was carried out in duplicate using different preparations of HeLa extract.

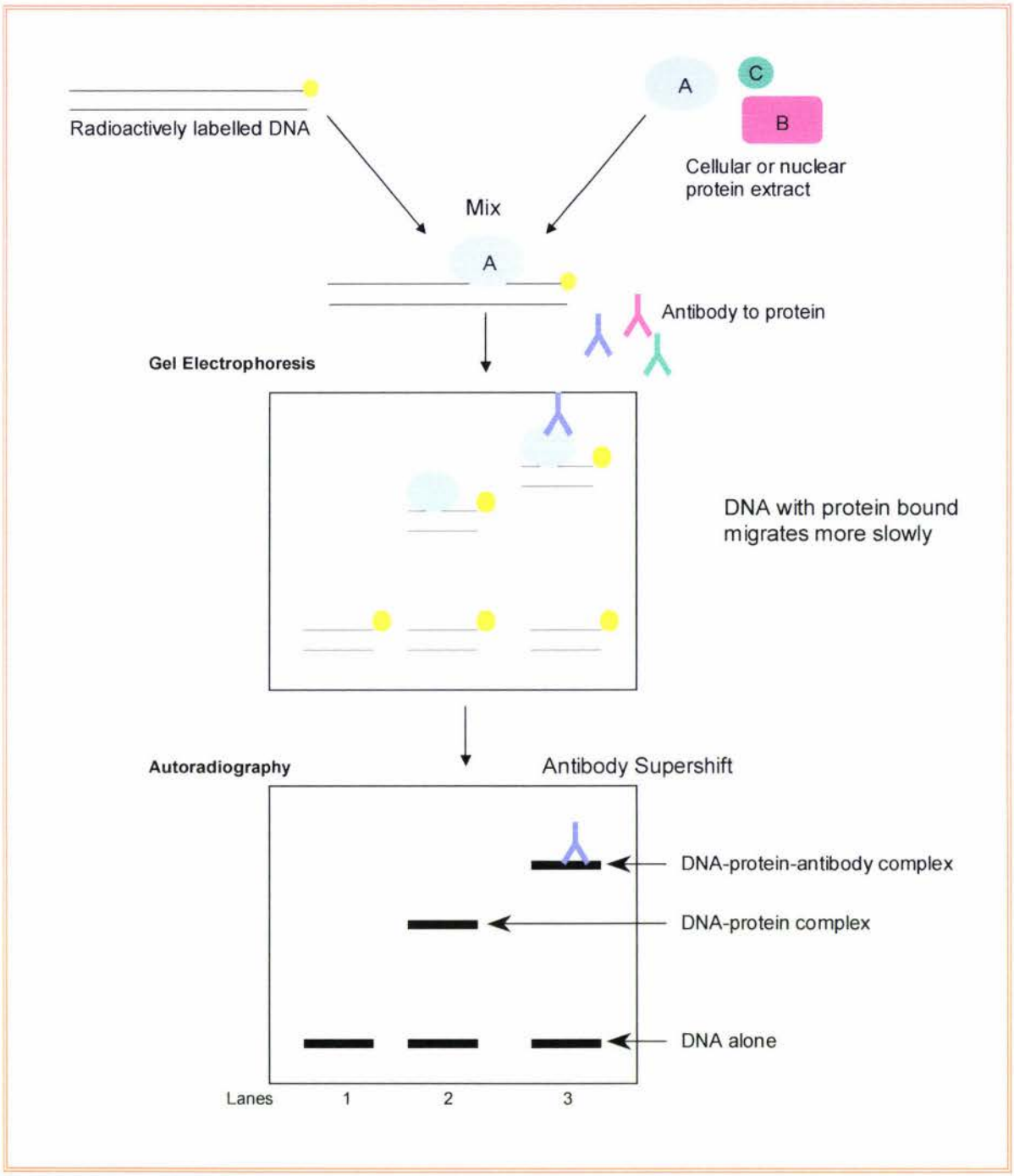


Figure 4.1: Schematic representation of Electrophoretic Mobility Shift Assays. The radioactively labelled DNA probe (lane 1) will have less mobility if a protein is bound to it (lane 2). This can be visualized by exposure to X-ray film. To identify the proteins that are bound an antibody supershift can be performed (lane 3) (adapted from Latchman, 1993).

4.2 DNA-protein binding studies using the topoisomerase IIα promoter.

The topoisomerase IIα promoter has been shown to recruit NF-Y to ICB1-4 (Herzog and Zwelling, 1997), and previous work carried out by Agnieszka Szremska (BSc honours project, 2000) suggested that Sp1 could bind ICB1. Since ICB1 does not contain a GC element, it is possible that Sp1 could be recruited to ICB1 via an interaction with NF-Y bound at ICB1. This suggests that interplay between Sp1 and NF-Y bound at GC1 and ICB1 respectively may have a functional significance for the transcriptional regulation of topoisomerase IIα. Therefore, oligonucleotide probes were designed to encompass both the ICB1 and GC1 elements of the topoisomerase IIα promoter, where each probe was 44 bp in size (refer to figure 4.2). In order to identify sequence specific binding a number of competitor oligonucleotides were designed containing mutations in either or both GC1 and ICB1.

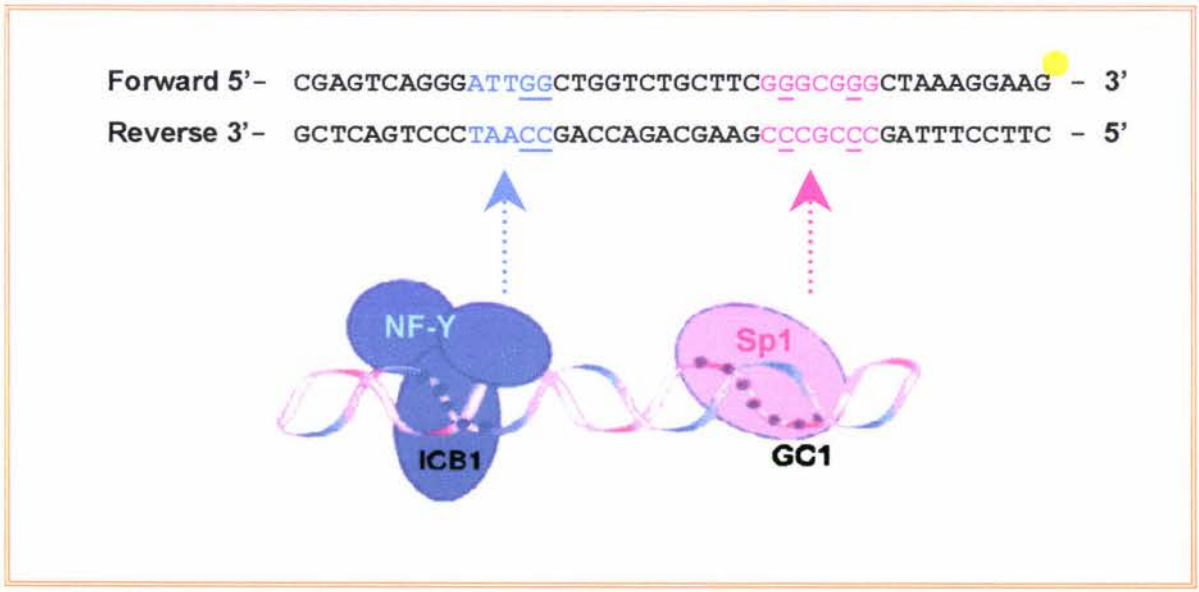


Figure 4.2: Schematic representation of GC1, ICB1 and putative transcription factors. NF-Y is shown as the blue trimeric complex on the left associating with the ICB1 element, the GC1 element is shown to associate with Sp1 (pink) on the right. The sequence of the double-stranded wild type 44 bp oligonucleotide probe used in electrophoretic mobility shift assays, is shown at the top. The ICB1 sequence is in blue and GC1 sequence is in pink. The underlined sequences are those that were mutated to inhibit protein binding and yellow signifies which strand was labelled with γ-^{32}P[ATP] (adapted from Isaacs, 1996b).

Mutant sequences (see appendix 1.2 for full sequences) were analyzed using the Findpatterns programme of GCG (Wisconsin Genetics Computer Group, USA) and the MatInspector V2.2 programme of the TRANSFAC4.0 (http://transfac.gbf.decreases/cgibin/matSearch /matsearch.pl). Both of these programs were used to search the transcription factor database for possible binding motifs. This exercise demonstrated that the mutations introduced into GC1 and ICB1 would be sufficient to abolish the Sp1 and NF-Y consensus sequences (respectively).

4.2.1 Preparation of HeLa extracts.

HeLa cells have been shown to contain NF-Y, Sp1 and Sp3 (Hagen *et al*, 1994) therefore it was not necessary to prepare cell extracts enriched for these proteins. HeLa extracts were prepared as described in section 2.2.23 and quantified as described in section 2.2.24. Using the Bradford protein assay, the protein concentrations of HeLa extracts 1 and 2 were estimated to be 1.64 μg protein/μL and 3.1 μg protein/μL respectively (refer to appendix 4 for details). The HeLa extract was stored at -70°C until required.

4.2.2 Preparation of the ^{32}P labelled probes.

The oligonucleotides were synthesized commercially and the forward oligonucleotides (refer to figure 4.2) were labelled with γ-^{32}P[ATP] as described in section 2.2.21(i). Once the labelling reaction was complete, the single-stranded oligonucleotides were annealed to their complementary unlabelled oligonucleotides. Each double-stranded oligonucleotide was purified using polyacrylamide gel electrophoresis (PAGE) to remove any contaminating shorter or single-stranded sequences (refer to figure 4.3). Due to immeasurable losses during the DNA purification procedure, the concentration of each DNA probe could only be estimated based on the amount of oligonucleotides added in the labelling reaction and the likely recovery. The amount of radioactivity incorporated into the purified oligonucleotides was determined using Cerenkov counting (refer to table 4.1).

The recovery of each DNA probe was estimated to be approximately half to a quarter of the initial amount labelled. A total of 200 ng of oligonucleotide was labelled, where 1 μL of labelled oligonucleotide incorporated varying amounts of radioactivity, ranging from about 10,000-40,000 counts per minute (cpm). Before carrying out each EMSA, the probes were diluted to generate about 9-10,000 cpm/μL, which correlates to a very approximate DNA concentration of about 0.1-0.2 ng DNA/μL.

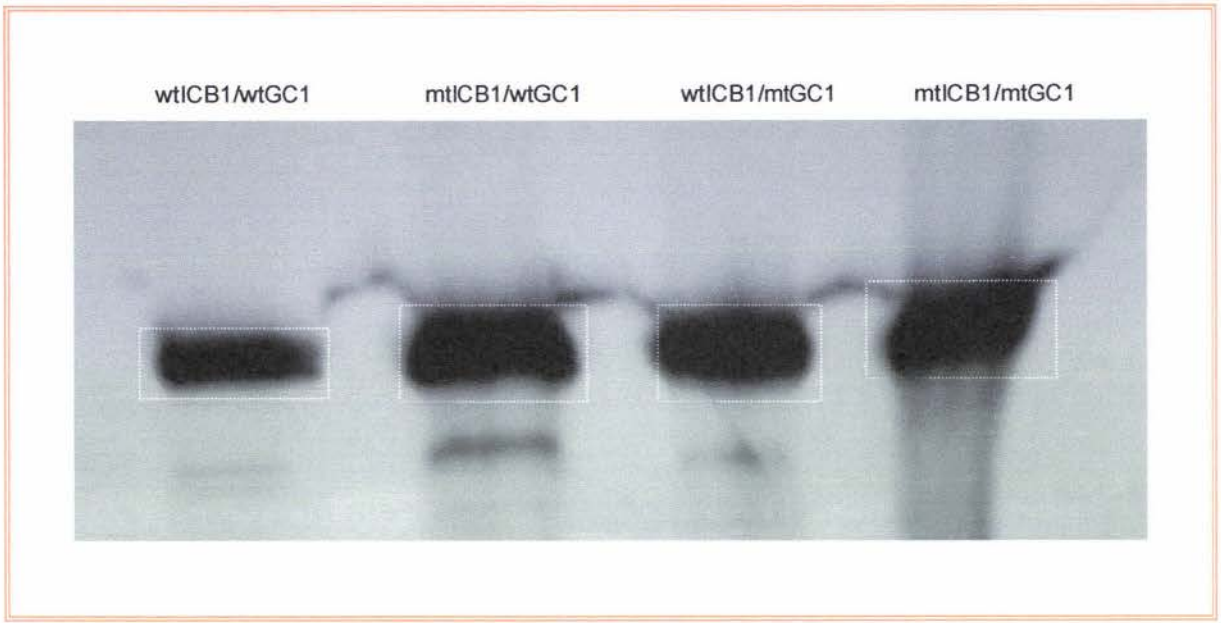


Figure 4.3: Gel purification of the ^{32}P labelled oligonucleotide probes for EMSA.
The oligonucleotides were endlabelled with γ-^{32}P[ATP] using T4 polynucleotide kinase. The double-stranded oligonucleotides were purified by electrophoresis on a 10% non-denaturing polyacrylamide gel in 1 x TBE for about 1 hour at 30 W.

The probes were located within the gel by exposure to X-ray film for about 1 minute, then excised from the gel (as shown) and the DNA was eluted overnight in 500 µL 50 mM KCl.

Probe	Counts per minute (1 µL)
Blank	39
wtICB1/wtGC1	33565
mtICB1/wtGC1	30192
wtICB1/mtGC1	41938
mtICB1/mtGC1	9159

Table 4.1: Incorporation of γ-^{32}P[ATP] into oligonucleotide probes.
Cerenkov counting (Beckman LS3801 Scintillation Counter) was used to determine the radioactivity incorporated into each double-stranded oligonucleotide probe. Approximately 0.1-0.2 ng of DNA was used for EMSA.

4.2.3 HeLa extract titration.

The precision of EMSA experiments relies on an equilibrium between bound and unbound DNA to protein, however, the amount of probe DNA recovered after purification could only be estimated. Therefore, it was necessary to determine the amount of HeLa extract required for each EMSA to generate optimal binding conditions. HeLa extract titrations were carried out using a standard 1 μL of radiolabelled probe (about 10,000 cpm) and varying amounts of HeLa extract from 0-4 μL (0-12 μg protein depending on the extract used). Accordingly, the binding reactions consisted of HeLa extract, poly dI.dC, buffer, competitor or antibody if included and finally the labelled probe. Poly dI.dC is a synthetic polymeric nucleic acid that is added to hinder nonspecific DNA-protein interactions.

Figure 4.4 illustrates the mobility shifts observed when increasing amounts of HeLa extract was added to four different oligonucleotide probes, representing either wild-type of mutated ICB1 and GC1 elements designed from the topoisomerase IIα promoter sequence. In the control lanes (lanes 1, 5, 9 and 13), no HeLa extract was added and consequently only free probe was detected, as expected. With the addition of 1 μL of extract, distinct banding patterns were observed that increased in intensity as more extract was added. Some of the bands appeared smeared or distorted, which could be due to dissociation of protein from DNA during electrophoresis. Each of the EMSA was carried out twice using two different sets of HeLa extract. Figure 4.4 is representative of the images obtained using both sets of HeLa extracts. The optimal amount of HeLa extract for this assay appears to be 4 μL (1.64 μg/μL) as the bands are more distinct. In each round of EMSAs, approximately 6-7 μg of protein was shown to be sufficient to generate optimal binding conditions with 1 μL of probe DNA (lane 4).

There are four distinctive protein/DNA complexes (labelled 1-4) present when the probe was wtICB1/wtGC1 (lanes 2, 3 and 4), however all four bands disappear when both ICB1 and GC1 were mutated (lanes 14, 15 and 16) indicating that the mutant probe has lost the ability to form those particular interactions. In conjunction, the appearance of these four bands varies depending on the mutational status of ICB1 and GC1 (lanes 6-12), suggesting that protein-binding specificity may exist. Other bands with higher mobility (shown below the dotted blue line figure 4.4) appear to occur more randomly, depending on which probe was used. Therefore, it was likely that these random bands represented non-specific interactions between probe and protein from the HeLa extracts.

When ICB1 was mutated (lanes 6-8), only bands 1, 3 and 4 remained and bands 1, 2 and 4 remained when GC1 was mutated (lanes 10-12). However, with the mtICB1/wtGC1 probe the intensity of bands 1 and 4 was significantly less when compared with the intensity of the same bands with the wt probe (lanes 2-4) and wtICB1/mtGC1 probe (lanes 6-8). This suggests that perhaps only band 2 may be specific for the ICB1 element, while bands 1 and 4 were possibly present due to residual binding. Residual binding of proteins (bands 1 and 4, lanes 11 and 12) was only observed when the probe was wtICB1/mtGC1. Bands 1, 3 and 4 appear to be specific for the GC1 element (lanes 6-9), where band 3 appears to be highly specific as a mutation in GC1 completely inhibits this interaction.

These differences in banding patterns provided important preliminary evidence to suggest that ICB1 and GC1 may bind different proteins, perhaps due to sequence specificity. To further investigate the interactions of proteins with the ICB1/GC1 oligonucleotide, a series of EMSAs were carried out using antibody supershifts and competitor oligonucleotides.

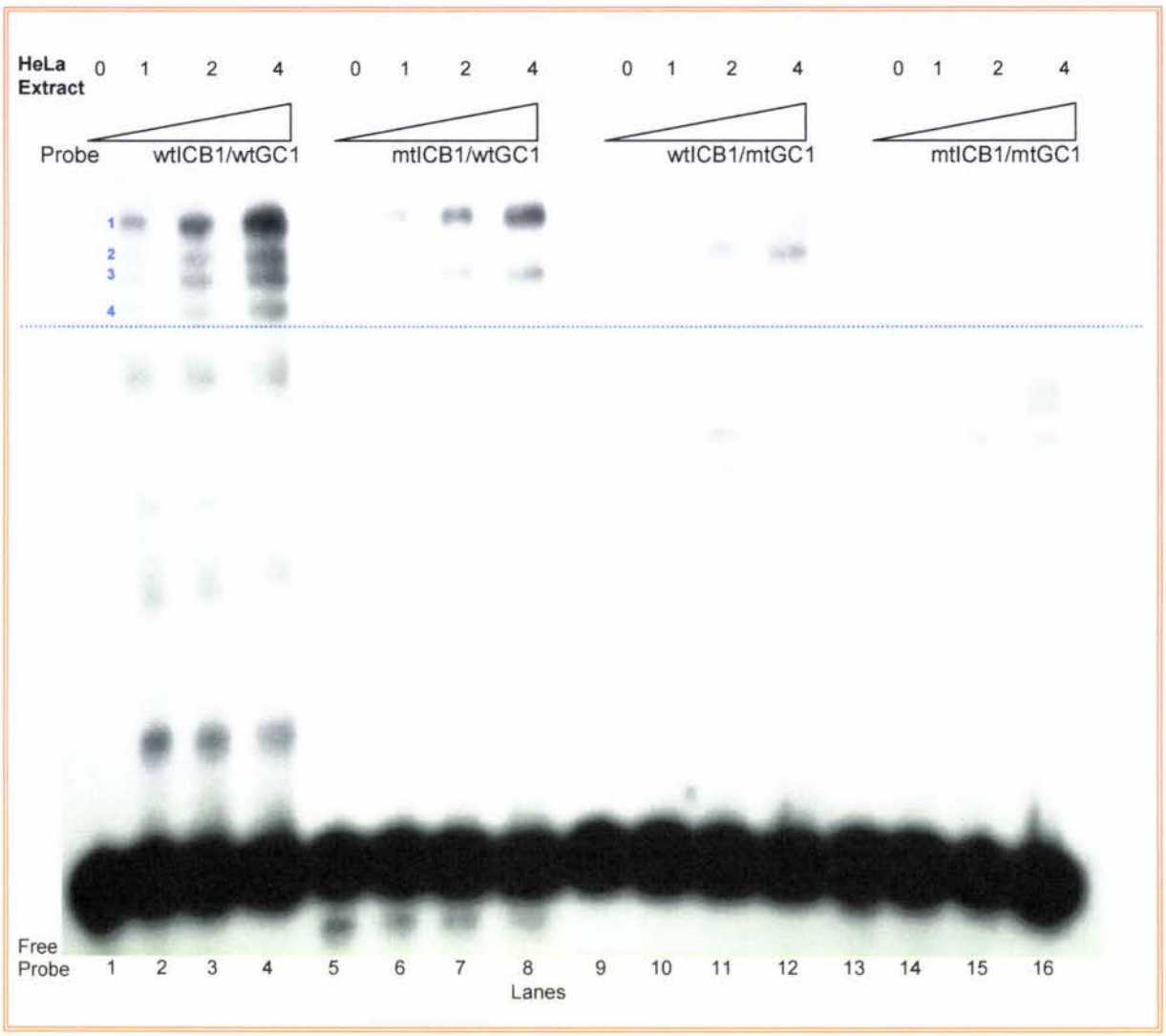


Figure 4.4: HeLa extract titration.

Increasing amounts (0, 1, 2 and 4 µL) of HeLa extract (1.64 µg protein/µL) was added to 1 µL of each probe. Half (10 µL) of each binding reaction was loaded onto a 4% polyacrylamide gel in 0.25 x TBE and electrophoresis was carried out for about 1 hour at 200 V. The gel was dried onto DE-81 paper and exposed to X-ray film for about 18 hours at -70°C in a radioactive safe cassette with intensifying screens. EMSAs were carried out as described in section 2.2.21.

Control lanes with no extract are lanes 1, 5, 9 and 13. The blue dotted line divides putative specific and non-specific interactions. Free probe with no protein bound is shown at the bottom of the gel. This figure is representative of duplicate experiments.

4.2.4 Antibody supershifts.

The addition of antibody to the binding reactions can assist in the identification of proteins present in a DNA-protein interaction. The antibody can interact with protein in one of two ways; the first is by binding to the DNA-protein complex through the protein, which results in further reduction of band mobility (refer to figure 4.1). The second is by directly binding to the protein and masking its DNA binding site, which inhibits the protein from binding to the DNA resulting in a reduction in band intensity. Antibodies against Sp1, Sp3 and NF-Y (using antibodies against subunits NF-YA and NF-YB) were added to a standard EMSA containing the wtICB1 /wtGC1 probe and HeLa cell extract in an attempt to identify the proteins involved in the four DNA-protein complexes observed in figure 4.4. The results are shown in figure 4.5.

The control lanes (1 and 8) in figure 4.5 have the same four bands that were observed in figure 4.4, lanes 2-4. Antibody supershifts occurred with each of the antibodies tested and three of the four protein/DNA complexes could be identified. When antibody against Sp1 was added to the assay (lanes 2 and 3) a definite shift in band 1 was observed, suggesting that Sp1 was present in the DNA-protein complex represented by this band. Upon the addition of antibody against NF-Y (lanes 4-7) the only change observed was a strong shift in band 2, suggesting that band 2 contained NF-Y. The same band (2) shifted regardless of which NF-Y antibody was used in the assay, which was to be expected since NF-Y is a trimer of three subunits A, B and C. In lanes 9 and 10, antibody against Sp3 was added and a shift was observed only in band 4, illustrating that Sp3 was present in that particular DNA-protein complex. The mobility of band 3 did not change upon addition of any of the antibodies tested; therefore, the third band remains an unknown protein-DNA interaction. In addition, no changes were observed in the bands of higher mobility, suggesting that these were the result of non-specific interactions.

Overall, the data shown in figure 4.5 suggests that four distinct DNA-protein complexes can be formed with the wtICB1/wtGC1 oligonucleotide from the topoisomerase IIα promoter. These DNA-protein complexes observed appear to contain Sp1, Sp3, NF-Y and an unidentified protein. The different amounts of antibody added to assays do not appear to induce any deviance in band patterns. Competitor assays were carried out to characterise any sequence-specific binding.

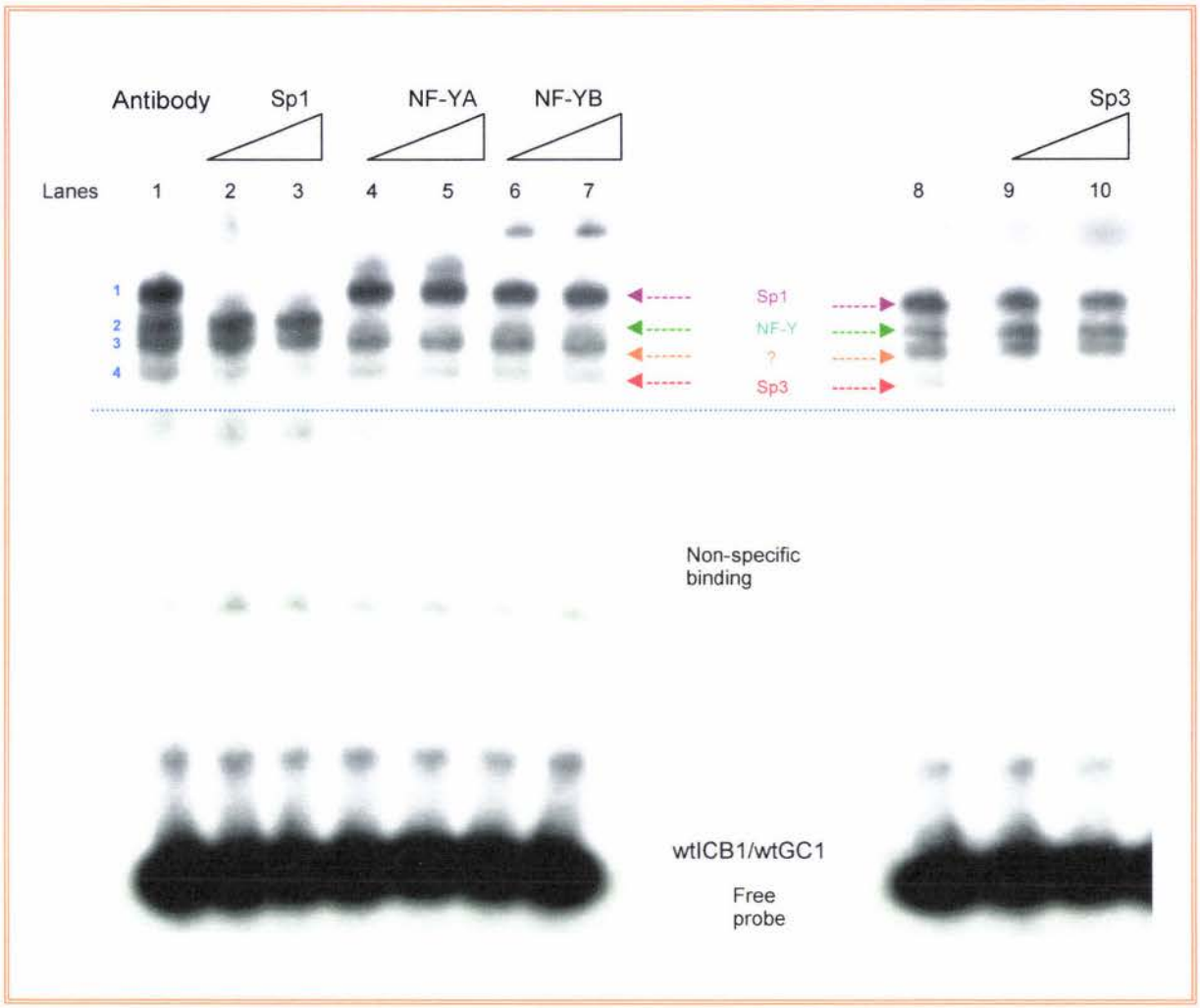


Figure 4.5: Antibody Supershift.

4 µL HeLa cell extract (1.64 µg protein/µL) with 1 µL of wtICB1/wtGC1 probe and increasing amounts of antibody (1or 2 µL of antibodies, which were at a concentration of 200 µg/mL). Half (12.5 µL) of each binding reaction was loaded onto a 4% polyacrylamide gel in 0.25 x TBE and electrophoresis was carried out for about 1 hour at 200 V. The gel was then dried and exposed to X-ray film for 24 hours at -70°C. Lanes 1 and 8 are the control lanes with no added antibody.

Specific and non-specific binding is indicated relative to the blue line. The arrows locate the bands that identify specific protein(s) in the DNA/protein complexes; purple: Sp1, green: NF-Y, orange: unknown and red: Sp3. This figure is representative of duplicate experiments.

4.2.5 Competitor assays.

Competitors were prepared as described in section 2.2.21(iii). Increasing amounts (0-100 ng) of unlabelled cold competitor (double or single element) were added to standard EMSA reactions containing wtICB1/wtGC1 probe and HeLa extract. In these assays, a decrease in band intensity indicated that the unlabelled competitor preferentially bound the protein that would normally be present in the DNA-protein complex. The sequence specificity of the four DNA-protein complexes was investigated using competitor oligonucleotides containing specific point mutations that were known to hinder NF-Y and Sp1 interactions.

4.2.5 (i) Double element competitors.

The competitors used in this assay were 44 bp in size and contained both ICB1 and GC1 elements with mutations in both or either element. The results are shown in figure 4.6.

The homologous competitor wtICB1/wtGC1 was able to compete for binding with all of the protein in the four complexes observed (lanes 1-4), as a uniform reduction in band intensity was observed as more competitor was added. In contrast, the competitor oligonucleotide containing mutations in both ICB1 and GC1 did not show any significant competition for any of the proteins (lanes 13-16), even though a difference in lane 16 is observed. The weaker intensity of the bands in lane 16 is likely to be due to a loading error.

Strong competition was observed for bands 1, 3 and 4 using the competitor oligonucleotide containing a mutation in the ICB1 sequence (lanes 5-8). This result confirms that the protein in band 2 (identified as NF-Y in figure 4.5) preferentially binds to ICB1, even though slightly weaker binding is observed in lanes 7 and 8. Some residual binding is apparent for band 1 (comparing lanes 3 and 4 with lanes 7 and 8) which suggests that NF-Y may be able to recruit a small amount of Sp1 to ICB1 and that binding of NF-Y to ICB1 may be assisted by Sp1 bound at GC1. Using competitors that contained a mutated GC1 (lanes 9-12), the only significant competition observed was for the protein in band 2. Therefore, the proteins identified in band 1, 3 and 4 (Sp1, unknown and Sp3, respectively from figure 4.5) appear to preferentially bind to GC1. However, the intensity of bands 1 and 4 also appear reduced, when comparing the intensity of band 3 across (which appears uniform) lanes 9-12. The reduced binding of Sp1 and Sp3 (bands 1 and 4, lanes 11 and 12) could be due to an interaction between NF-Y bound at ICB1. Taken together these data suggest that an interaction occurs between NF-Y bound at ICB1 and Sp1 (and/or Sp3) bound at GC1.

4.2.5 (ii) Single element competitors.

A second set of competitor assays was carried out using oligonucleotide competitors that consisted of a single element, either ICB1 or GC1 (refer to appendix 1.2 for sequence details). This experiment was designed to investigate the importance of the flanking sequences in protein binding and to further investigate any possible synergism between the two elements. Standard EMSAs were carried out using the labelled wtICB1/ wtGC1 probe (1 μL), HeLa extract (6 μg) and competitor oligonucleotides (0-100 ng). The results are shown in figure 4.7.

If protein bound at ICB1 was independent of protein bound at GC1, then an ICB1 competitor should simply compete for binding with only an ICB1 associating protein (ie NF-Y) and vice versa. However, if synergism occurs between the ICB1 and GC1 elements, then a single site oligonucleotide may have less effect as a competitor. If a protein preferentially binds to the longer oligonucleotide of the probe rather than the shorter oligonucleotide of the competitor, then it can be assumed that the flanking sequences between ICB1 and GC1 are of significance for protein binding.

The same four DNA/protein complexes were observed (figure 4.7) as for the other standard EMSA reactions. Lanes 1-4 show that the ICB1 competitor has the ability to compete strongly for the protein(s) associated with band 2 (NF-Y) and to a much lesser extent the protein(s) present in bands 1 and 4, suggesting that synergism between the proteins may exist or perhaps the longer probe preferentially binds Sp1, the unknown protein and Sp3. The mutated ICB1 competitor did not compete for binding, with any of the bands (lanes 5-8). The GC1 competitor (lanes 9-12) exhibits the ability to compete weakly for binding with the proteins present in bands 1 and 4 only, which were identified as Sp1 and Sp3 in figure 4.5, respectively. This weak competition suggests that Sp1 and Sp3 may preferentially bind to the longer wtICB1/wtGC1 probe, possible due to the presence of bound NF-Y. The mutated GC1 element competitor does not have the ability to compete for any of the bands observed (lanes 13-16).

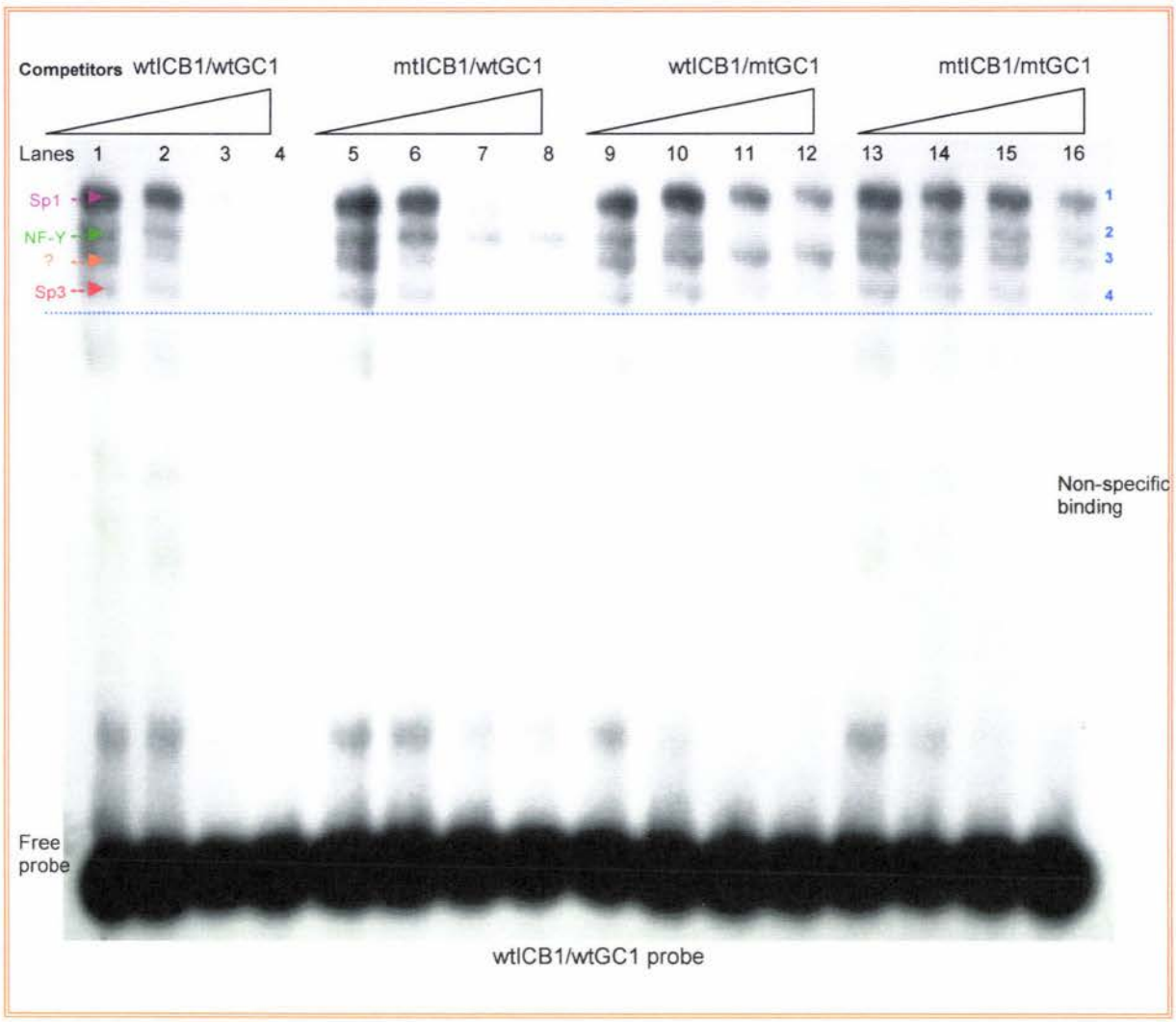


Figure 4.6: Competitor Assay (Double element competitors).

Increasing amounts (0, 5, 50, 100 ng DNA) of double-stranded competitor was added to 1 μL of wtICB1/wtGC1 probe with 4 μL HeLa extract (1.64 μg protein/μL). Half (10 μL) of each binding reaction was loaded onto a 4% polyacrylamide gel and electrophoresis carried out for about 1 hour at 200 V. The gel was dried onto DE-81 paper and exposed to X-ray film for approximately 20 hours at -70°C. The differences in intensity between lanes 13-16 are due to error whilst loading the samples in the wells.

Specific and non-specific binding is indicated relative to the blue line. The arrows locate the bands that identify specific protein(s) in the DNA/protein complexes; purple: Sp1, green: NF-Y, orange: unknown and red: Sp3. This figure is representative of duplicate experiments.

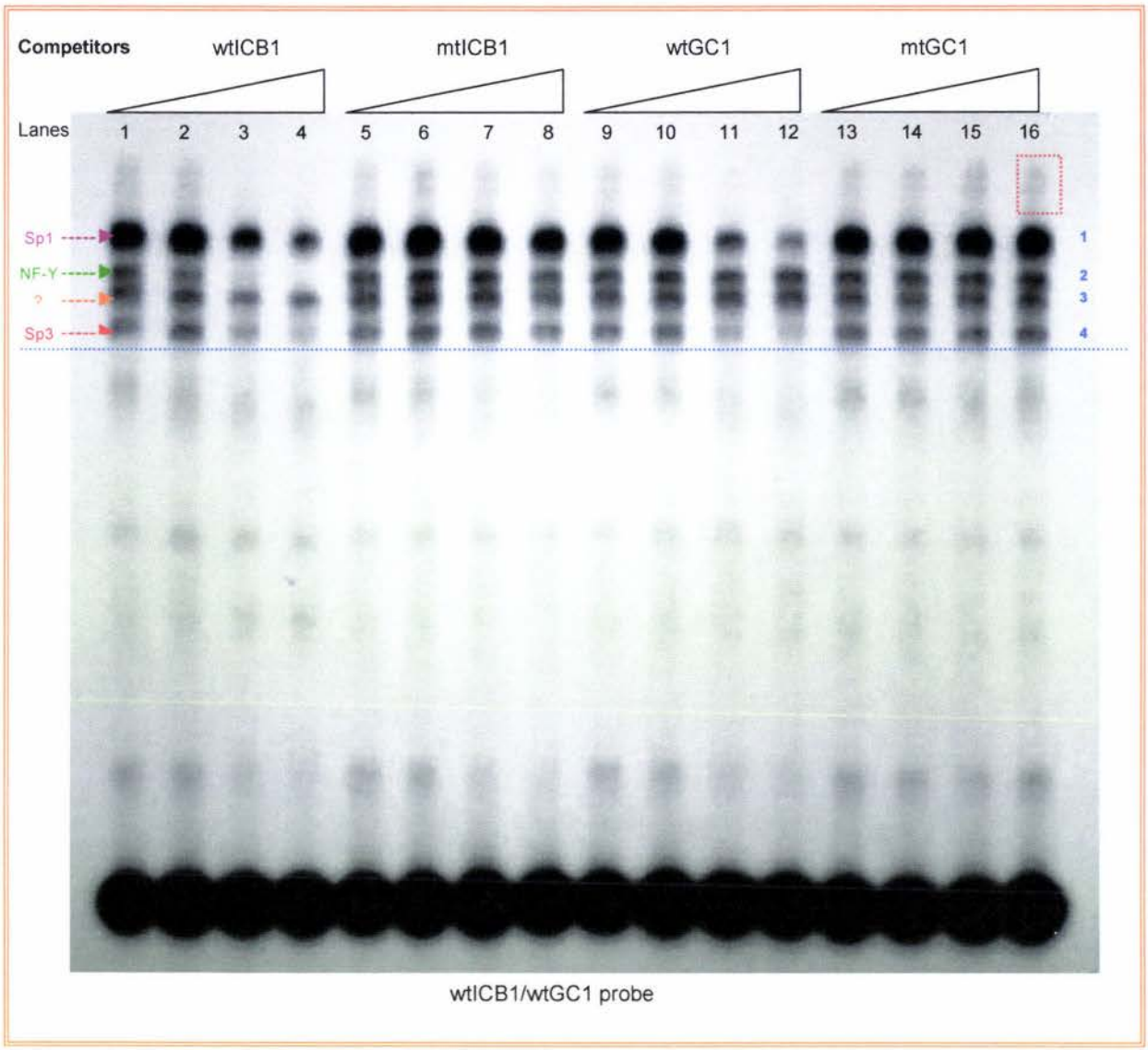


Figure 4.7: Competitor Assays (Single element competitors).

Increasing amounts (0, 5, 50, 100 ng DNA) of double-stranded competitor was added to 1 µL of probe wtICB1/wtGC1 with 2 µL HeLa extract (3.1µg protein/µL). Half (10 µL) of each binding reaction was loaded onto a 4% polyacrylamide gel and electrophoresis carried out for about 1 hour at 200 V. The gel was dried onto DE-81 paper and exposed to X-ray film for approximately 22 hours at -70°C.

Specific and non-specific binding is indicated relative to the blue line. The arrows locate the bands that identify specific protein(s) in the DNA/protein complexes; purple: Sp1, green: NF-Y, orange: unknown and red: Sp3. This figure is representative of duplicate experiments. Within the red box is a representation of a putative ICB1/NF-Y/Sp1 or Sp3 complex, this multi-protein complex can also be seen in lanes 1, 2, 5-10 and 13-16.

The collective results of figures 4.4, 4.5, 4.6 and 4.7 suggest that four proteins can associate with the wtICB1/wtGC1 sequence of the topoisomerase IIα promoter. Band 1 contains Sp1 and it appears that this protein binds to the GC1 region of the oligonucleotide with high sequence specificity, since this band is not observed when the probe contains a mutated GC1 sequence (figure 4.4, lanes 9-16). Using the same rationale band 4 contains Sp3 and this protein also binds to the GC1 region (figure 4.4 lanes 11 and 12). Band 2 contains the NF-Y complex, which binds to the ICB1 region of the oligonucleotide, as this band is not observed when the probe contains a mutated ICB1 element (figure 4.4, lanes 5-8).

Even though both ICB1 and GC1 were shown to associate with different proteins, it appears that ICB1 has the ability to recruit Sp1 and Sp3 (figure 4.4, lanes 11 and 12). In conjunction with this finding, the single element competitor assays demonstrate that ICB1 can recruit NF-Y and at the same time decrease the binding affinities of Sp1 and Sp3 (figure 4.7, lanes1-4). Therefore, it is likely that Sp1 and Sp3 are recruited to ICB1 via a putative interaction between NF-Y and Sp1/Sp3 (lanes 9-12, bands 1 and four of figure 4.6). Interestingly, the opposite scenario of Sp1 or Sp3 bound to GC1 being able to recruit NF-Y to GC1 does not appear to be the case in figures 4.4 (lanes 7 and 8) and 4.7 (lanes 11 and 12). This suggests that it may be possible that NF-Y first binds to ICB1 and enables the recruitment of Sp1/Sp3 to the complex. When trying to examine the sequence specificity of the proteins binding it was found that the shorter GC1 competitor oligonucleotide had the ability to recruit a portion of Sp1 and Sp3, although a substantial amount remains when compared with the competition exhibited by the double element wtICB1/wtGC1 probe (compare lanes 9-12, figure 4.7 and lanes 1-4, figure 4.6). This suggests that Sp1 and Sp3 preferentially bind to the longer ICB1/GC1 sequence.

Taken together these binding patterns suggest that the ICB1 sequence is not essential for Sp1, Sp3 and the unknown protein to bind, nor is the GC1 sequence required to bind NF-Y. However, the interaction between GC1, Sp1 and Sp3 is enhanced if the ICB1 element is intact. By inference, this suggests that NF-Y bound at ICB1 is essential to maintain strong interactions between GC1 and Sp1 and Sp3. It could be speculated that NF-Y binds to ICB1, which in turn recruits Sp1 and Sp3 to ICB1 possibly via a putative interaction between the two proteins. In fact, both NF-Y and Sp1 have been shown to contain Q-rich domains, which have been suggested to mediate protein-protein interactions (Roder *et al*, 1997). Therefore, it is likely that Sp1 and Sp3 are recruited to ICB1 via a putative interaction between NF-Y and Sp1/Sp3 (lanes 9-12, bands 1 and four of figure 4.6). If this were the case then there should

be some evidence of a less mobile DNA/NF-Y/Sp1 or Sp3 complex forming, an example of this can be seen clearly in figure 4.7 enclosed in a red box. This heavier complex also appears to diminish when NF-Y, Sp1 or Sp3 are removed by competitors (figure 4.7 lanes 3, 4 and 11, 12). Other EMSA did have evidence of this complex but were not as clear. Generally, HeLa extracts that had a higher protein concentration (3.1 μg protein/μL) showed the formation of this very low mobility band.

The unknown protein also displayed properties of being a GC1 associating protein, but it did not compete for binding with a single element competitor (figure 4.6, lanes 9-12). In addition, binding of the unknown protein to the ICB1/GC1 element appears to be independent of having either NF-Y or Sp1/Sp3 bound. The EMSA observations suggest that the sequence between ICB1 and GC1 is important for binding of the unknown protein. Alternatively, the unknown protein may require Sp1 or Sp3 to be bound at GC1 before strong binding can occur (figure 4.6, compare lanes 5-8 with lanes 9-12). Further experiments will be necessary to determine the functional significance of the unknown protein binding to ICB1/GC1.

4.4 Chapter summary.

A DNA fragment encompassing both the ICB1 element and GC1 element of the topoisomerase IIα promoter was shown to form four distinct DNA-protein complexes, and antibody supershift assays enabled the positive identification of proteins Sp1, Sp3 and NF-Y binding to these elements.

Sp1 and Sp3 appear to bind preferentially to GC1. The unknown protein also displayed properties of being a GC1 associating protein, however it may bind to sequences between ICB1 and GC1. NF-Y was shown to bind to ICB1 exhibiting strict sequence specificity. In addition, the presence of NF-Y may enable Sp1 to be recruited to GC1, but in order to do so both ICB1 and GC1 must be intact to facilitate a strong interaction between these two proteins. It is possible that a protein complex consisting of NF-Y/Sp1/Sp3 can form at DNA elements within the topoisomerase IIα promoter, where NF-Y binds first and enables the binding of the other transcription factors.

The functional significance of Sp1, Sp3 and NF-Y binding to ICB1 and GC1 in topoisomerase IIα expression was addressed by carrying out transient transfection experiments using reporter gene assays in chapter 5.

Chapter 5: Transient Transfections.

5.1 Introduction.

The previous chapter illustrated that NF-Y, Sp1 and Sp3 are all capable of interacting with topoisomerase IIα elements GC1 and ICB1 *in vitro*. In addition, previous studies have shown that GC2 is also capable of interacting with Sp1 and Sp3 (Szremska, 2000). The procedure whereby exogenous DNA is introduced into recipient cells is known as transfection. Transient transfections and reporter gene assays were carried out to examine the functional effects of these interactions on topoisomerase IIα expression *in vivo*.

Reporter genes have proven to be widely applicable in animal and plant cell biology and are commonly used to examine eukaryotic gene expression *in vivo*. Reporter gene assays involve a vector (eg. pGL3B), which contains a promoter-less gene (reporter gene) that is adjacent to a multiple cloning site (refer to figure 5.1) in which the promoter of interest is placed. The promoter is then exploited to drive the expression of the reporter gene within transfected cells. In many experiments, a 'minimal' promoter is attached to a reporter gene, such as firefly luciferase. The minimal promoter generally consists of a short promoter sequence containing a TATA box and a transcriptional start site, which enables a low level of 'background' transcription. Information can be gathered about promoter strength and activity by measuring the enzymatic activity of the reporter gene product in cell extracts.

The minimal promoter (-617 bp upstream of the 5' transcriptional start site) for topoisomerase IIα expression has already been defined (Isaacs, 1996b) and previously cloned into a pGL3Basic vector (M^{c}Lenachan, 1998). The transient expression of luciferase under the control of the topoisomerase IIα promoter (pGL3B –617 topoisomerase IIα) has also been monitored successfully in HeLa cells (M^{c}Lenachan, 1998; Szremska, 2000), therefore the same approach was used in this study. The significance of these elements and the functional relevance of putative transcription factors that bind to them were investigated *in vivo* by generating mutations in GC2, GC1 and ICB1 and carrying out functional assays.

5.2 Generating –617 (GC2mt) topoisomerase IIα promoter constructs.

Topoisomerase IIα –617 promoter fragments (in pGL3B) with mutations in GC2, GC1, ICB1 and one with a mutation in both GC1 and ICB1 had been prepared previously (M^{c}Lenachan, 1998; Szremska, 2000). However, in this study, constructs that harboured mutations in GC2 as well as GC1 and/or ICB1 were also required. These new plasmids were generated by ligating existing fragments of the –617 topoisomerase IIα promoter and pGL3B. Plasmid identity was checked by restriction endonuclease digestion and sequencing.

5.2.1 Generating pGL3B (GC2mt) Vector and Insert DNA.

Upon analysis, the pGL3B wt –617topoisomerase IIα sequence revealed that two *Hin*d III sites exist, one site was close to the GC2 element and the other was within the multiple cloning site of the pGL3B vector (refer to figure 5.1 and appendix 5.1). Therefore, a *Hin*d III digest of pGL3B –617 topoisomerase IIα plasmids would produce fragments of approximately 5 kb and 400 bp in size. The larger 5 kb fragment would contain most of the vector plus the GC2 portion of the topoisomerase IIα promoter and the remaining 400 bp fragment would contain the rest of the topoisomerase IIα promoter from position –259. Promoter constructs containing double and triple mutations were prepared by religating appropriate *Hin*d III fragments (refer to figure 5.2).

A pGL3B –617 topoisomerase IIα plasmid that was known to have a mutation in GC2 (prepared by Agnieszka Szremska, Department of Molecular BioSciences, Massey University, Palmerston North; appendix 5.2) was digested with *Hin*d III and the 5 kb fragment gel purified and quantified, in order to generate a pGL3B vector that had GC2mt attached (refer to figure 5.2). To generate the insert DNA, known pGL3B –617 topoisomerase IIα vectors were also digested with *Hin*d III and the 400 bp fragment released from this digest was gel purified and quantified. Figures 5.3 and 5.4 are representative of the manipulations involved in cloning each of the required –617 topoisomerase IIα pGL3B reporter constructs.

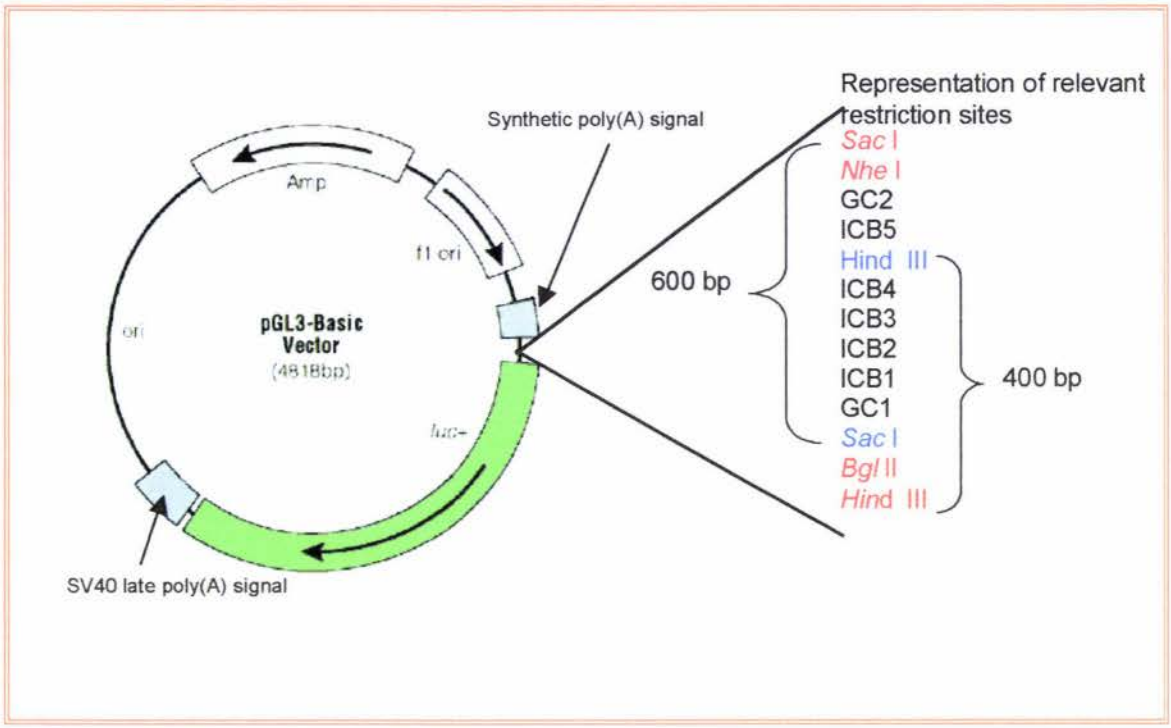


Figure 5.1: Schematic representation of the restriction sites within the topoisomerase IIα -617 promoter and the pGL3Basic vector.

The pGL3B vector is on the left and the restriction sites flanking the –617 topoisomerase IIα elements are shown on the right. pGL3B vector can confer ampicillin (Amp^R) resistance on transformed *E.coli* cells. The synthetic poly(A) signal is required to terminate transcription, and the SV40 late poly(A) signal is for RNA processing in mammalian cells.

The insert DNA (-617 topoisomerase IIα) is adjacent to the luciferase reporter gene (green). Red represents the restriction sites that belong to the vectors' multiple cloning site (MCS), whereas blue designates the restriction sites within the topoisomerase IIα promoter. If the topoisomerase IIα insert is in the correct orientation then a *Sac* I digest will release fragments of approximately 600 bp (insert) and 5 kb (vector) however, if the insert is in the incorrect orientation then the *Sac* I fragment will be smaller (approximately 450 bp). A *Hin*d III digest will generate fragments approximately 400 bp (partial insert DNA) and 5 kb (pGL3B vector with GC2 attached) in size, and will therefore confirm the presence of the insert.

(Figure adapted from www.Promega.com/vectors).

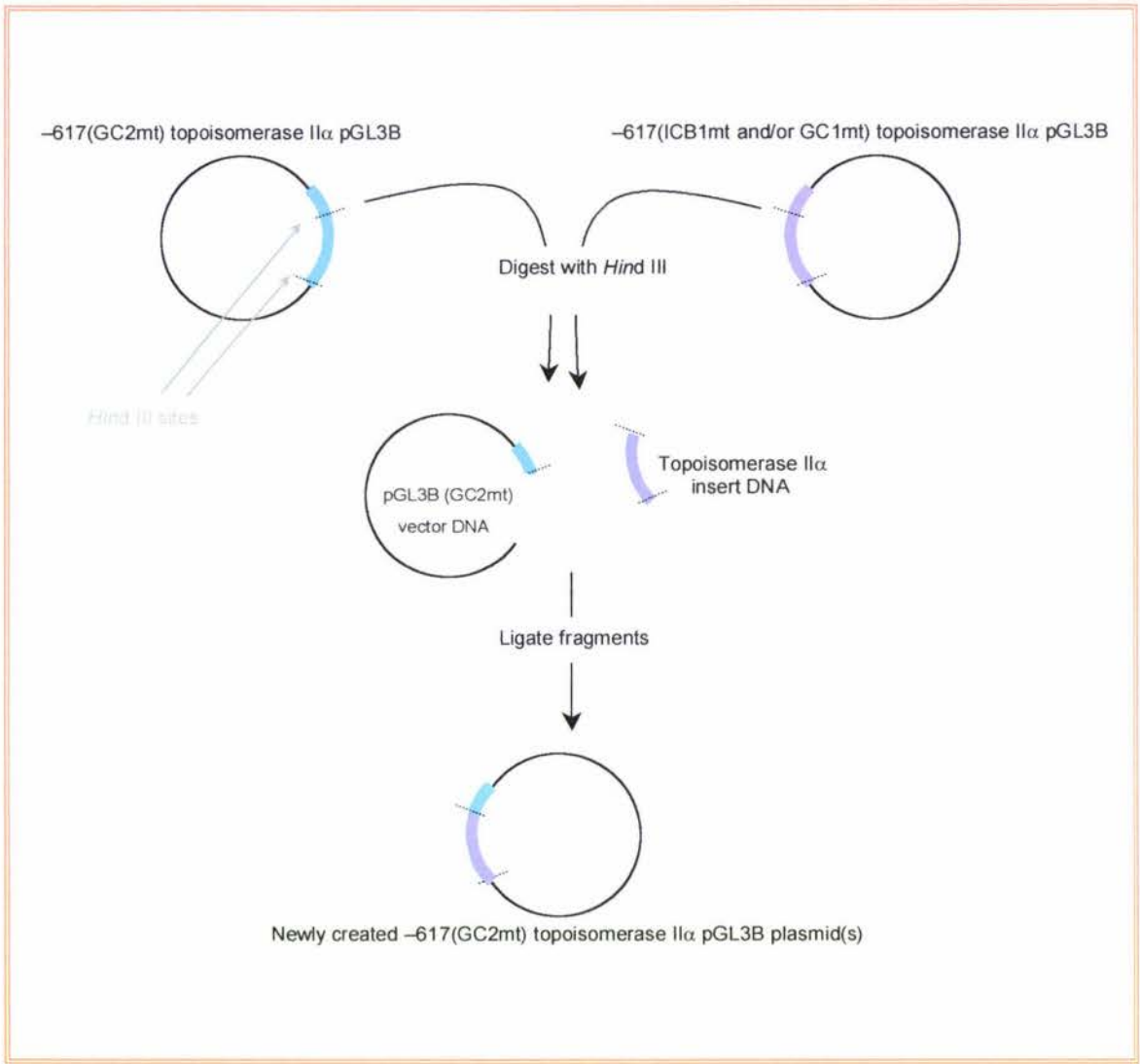


Figure 5.2: Schematic representation of the cloning strategy to create recombinant –617 topoisomerase IIα pGL3B plasmids.

Each of the GC2mt, ICB1mt and GC1mt –617 topoisomerase IIα pGL3B vectors were digested with *Hind* III to generate the required vectors and inserts. Each was then religated to create new –617topoisomerase IIα pGL3B plasmids that harbour mutations in GC2 and ICB1 and/or GC1.

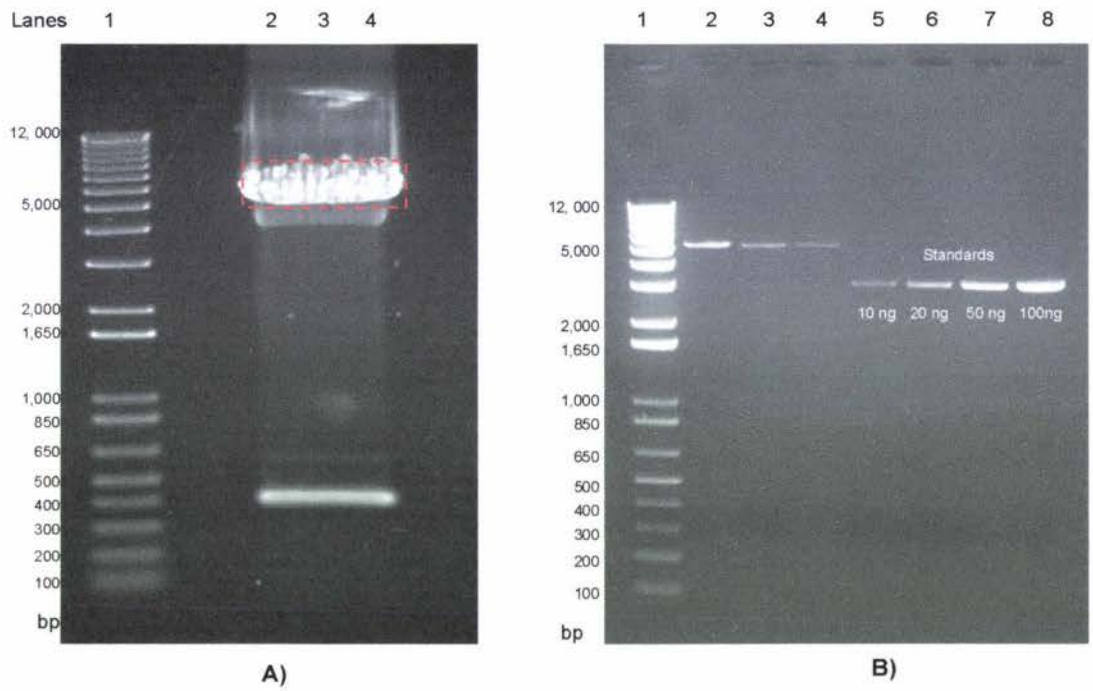


Figure 5.3: pGL3B –617 topoisomerase IIα (GC2mt) digest with enzyme *Hind* III.
Samples were analyzed on a 1% agarose gel in 1xTAE buffer by electrophoresis for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 μg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

A) a total of 35 μg of pGL3B –617 topoisomerase IIα (GC2mt) was digested with *Hind* III for about 2 hours at 37°C. Digested vector was then pooled, phosphatased, ethanol precipitated, and the total volume electrophoresed. The pGL3B vector with GC2mt incorporated was excised from the gel (indicated by the red box) and DNA purified using the CONCERT™ DNA extraction kit.

B) Quantification of pGL3B GC2mt vector:

lanes	1	10 μL 1kb plus ladder
	2	1 μL undiluted GC2mt pGL3B
	3	2 μL diluted 1:5 GC2mt pGL3B
	4	2 μL diluted 1:10
	5-8	quantification standards

The intensity of the samples was compared to the standards and the pGL3B GC2mt vector DNA was estimated to be at a concentration of ~ 50 ng/μL.

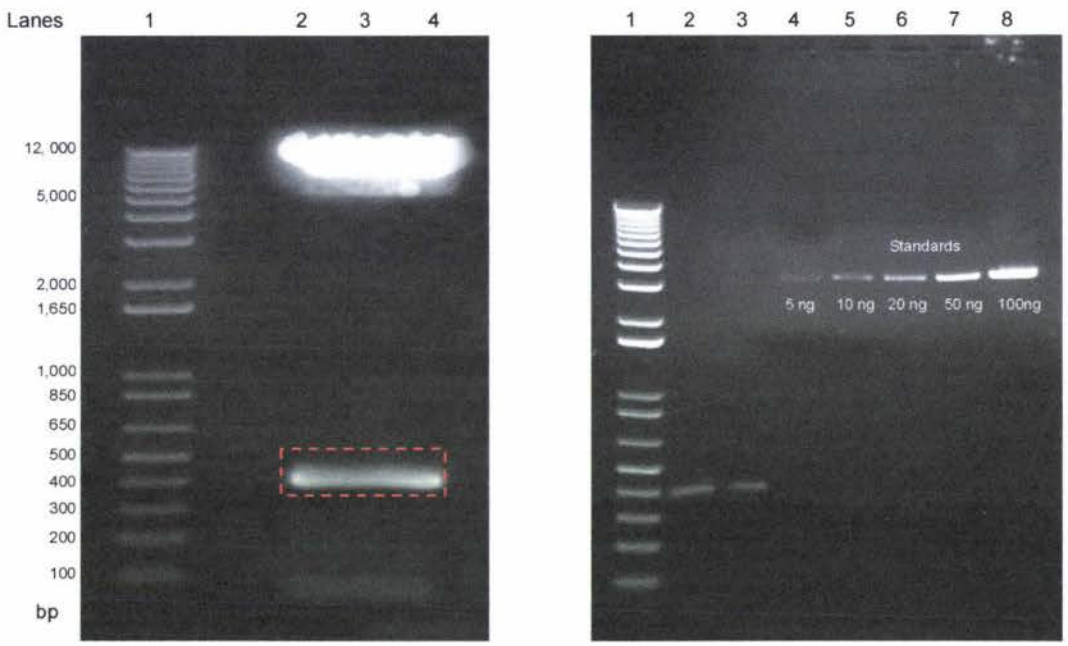


Figure 5.4: pGL3B –617 topoisomerase IIα (ICB1mt/GC1mt) digest with enzyme *Hind* III to generate insert DNA.

Samples were run analyzed on a 1% agarose gel in 1xTAE buffer by electrophoresis for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 μg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

A) a total of 40 μg of –617 topoisomerase IIα (ICB1mt/GC1mt) pGL3B plasmid was digested with *Hind* III for 1 3-4 hours at 37°C. Digested samples were pooled, ethanol precipitated and the total volume electrophoresed. The insert DNA was isolated from the gel (indicated by the red box) and purified using the CONCERT™ DNA extraction kit.

B) Quantification of insert ICB1mt/GC1mt DNA:

Lanes	1	10 μL 1kb plus ladder
	2	1 μL undiluted ICB1mt/ GC1mt DNA
	3	1 μL undiluted ICB1mt/ GC1mt DNA
	5-8	quantification standards

The intensity of the samples was compared to the standards and the ICB1mt/GC1mt insert DNA was estimated to be at a concentration of about 10 ng/μL.

5.2.2 Ligations of pGL3B GC2mt Vector and Insert DNA.

Once the appropriate mutated –617 topoisomerase IIα fragments and pGL3B GC2mt vector were prepared, the DNA fragments were ligated together. Ligation efficiency is dependent on the concentration of DNA ends in the reaction. To favour inter-molecular over intra-molecular ligation, the concentration of insert must be higher than that of the vector. The insert was added at either a 2:1 or 3:1 (insert:vector) molar ratio, where the amount of vector DNA used in each ligation was 50 ng in a total volume of 20 μL. Table 5.1 describes the ligation reactions carried out when generating pGL3B –617topoisomerase IIα GC2mt/ICB1mt/GC1mt (See appendix 6.1 for full set of ligation reactions). Ligations were carried out as described in section 2.2.15.

Sample	ng vector	ng insert	μL vector	μL insert	T_4 buffer	T_4 ligase	H_2O	Total volume
Vector plus ligase (Control)	50 ng	-	1 μ L	-	4 μ L	1 μ L	14 μ L	20 μ L
Vector minus ligase (Control)	50 ng	-	1 μ L	-	4 μ L	-	15 μ L	20 μ L
Vector plus insert DNA 1:3	50 ng	12 ng	1 μ L	1.2 μ L	4 μ L	1 μ L	12.8 μL	20 μ L
Vector plus insert DNA 1:2	50 ng	8 ng	1 μ L	0.8 μ L	4 μ L	1 μ L	13.2 μ L	20 μ L

Table 5.1: A typical ligation reaction; generating –617 topoisomerase IIα pGL3B with mutations in GC2, ICB1 and GC1.

Vector GC2mt pGL3B –617 topoisomerase IIα was at concentration of 50 ng/μL. Insert DNA ICB1mt/GC1mt was at concentration of 10 ng/μL. Ligations were carried out using either a 2:1 or 3:1 (insert:vector) molar ratio.

Two controls were included; the vector plus ligase control was useful in determining the amount of vector re-ligation and the vector minus ligase control, for an indication of the amount of uncut vector. Ligations were incubated overnight at 4°C and transformation of XL-1 competent cells was carried out the following day.

5.2.3 Transformation of XL-1 competent cells using ligation reactions.

Once foreign DNA is inserted into the cloning vector, the newly formed plasmid can be amplified by introducing it into a host cell, such as bacteria. Transformation of bacteria (*E.coli* XL-1 competent cells) with exogenous DNA (pGL3B –617 topoisomerase IIα plasmids) was performed as described in section 2.2.16. A typical transformation reaction is illustrated in table 5.2 (see appendix 6.2 for full set of transformation results).

Sample	Ligation mixture	XL-1 Competent cells	Dilution	Volume plated	Number of colonies
XL-1 (control)	10 µL 10 µL	100 µL 100 µL	10^7 10^8	50 µL 50 µL LB plates	289 35
XL-1 (control)	10 µL	100 µL	Undiluted	50 µL LB+Amp plates	0
Vector plus ligase (control)	10 µL 10 µL	100 µL 100 µL	Undiluted 10	50 µL 50 µL LB+Amp plates	27 3
Vector minus ligase (control)	10 µL 10 µL	100 µL 100 µL	Undiluted 10	50 µL 50 µL LB+Amp plates	5 0
Vector plus insert 1:3 (GC2mt/ICB1mt/GC1mt pGL3B)	10 µL 10 µL 10 µL 10 µL	100 µL 100 µL 100 µL 100 µL	Undiluted Undiluted 10 10^2	50 µL 30 µL 50 µL 50 µL LB+Amp plates	412 202 3* 0
Vector plus insert 1:2 (GC2mt/ICB1mt/GC1mt pGL3B)	10 µL 10 µL 10 µL 10 µL	100 µL 100 µL 100 µL 100 µL	Undiluted Undiluted 10 10^2	50 µL 30 µL 50 µL 50 µL LB+Amp plates	472 119 11* 1

Table 5.2: Results of XL-1 transformation using pGL3B ligations.

Transformation of XL-1 competent cells, involving plasmids generated from ligation of pGL3B –617 topoisomerase IIα GC2mt and insert DNA (ICB1mt/GC1mt). Half of the total volume of ligations were used to transform XL-1 competent cells. Unless indicated, all samples were plated onto LB plus ampicillin plates and incubated overnight at 37°C. Colonies were creamy white in colour, and appropriate colonies (*) were selected and cultured to isolate plasmid DNA for analysis.

Competent cells are able to take up DNA, which is thought to occur due to the presence of cell-surface proteins or enzymes that enable DNA binding or uptake. XL-1 blue competent cells alone cannot grow on LB plates with ampicillin; therefore, only successfully transformed XL-1 colonies will grow on LB plates with ampicillin. The fact that no colonies grew on the control plate of cells with ampicillin demonstrates that there was no contamination and no spontaneous ampicillin resistance generated in these experiments.

The vector plus ligase control was useful in determining the amount of vector re-ligation, whereas vector minus ligase control to give an indication of the amount of XL-1 blue cells transformed by digested vector alone. Both these controls had comparatively low growth, indicating a small amount of religatable vector and uncut vector (respectively) in the vector preparation.

The controls with XL-1 blue competent cells on LB only plates, enabled the calculation of viable XL-1 blue cells; approximately 6.4 X 10^7 colonies grew for every μL of XL-1 blue competent cells. Of that number, only 7.6 transformants arise demonstrating that these competent cells had low transformation efficiency. The positive colonies that grew on the vector plus insert plates were screened for the presence of pGL3B –617 topoisomerase IIα plasmid.

5.2.4 Confirming plasmid identity.

Plasmid DNA was extracted from nine positive XL-1 colonies (indicated by * in table 5.2) using the rapid boil technique described in section 2.2.17(i). Plasmid identity was checked by enzymatic digestion of rapid boil DNA followed by agarose gel electrophoresis and the insert was checked by automated DNA sequencing.

(i)Plasmid Identity by Enzyme Digestion

Rapid boil plasmid DNA was digested individually with enzymes *Hin*d III and *Sac* I. If the 400 bp *Hin*d III fragment was in the correct orientation within pGL3B it would generate DNA fragments; 400 bp and 600 bp in size respectively (refer to figure 5.1).

Figure 5.5 illustrates the results of digesting rapid boil plasmid preparations with *Hin*d III and *Sac* I; those plasmids with the insert in the incorrect orientation can been seen in samples 1, 4, 5, and 6 due to the smaller (approximately 450 bp) *Sac* I fragment generated. Those

plasmids with the insert in the correct orientation can been seen in samples 2, 3, 7, 8 and 9 (*Sac* I fragment about 600 bp) however, sample 7 has additional bands present possibly due to pGL3B incorporating two insert –617 topoisomerase IIα fragments which were joined together. The positive colony that was chosen to isolate further plasmid DNA from was sample 8; this colony was cultured and maintained for further analysis and automated DNA sequencing.

(ii)Confirming Plasmid Identity by Sequencing.

Once the colonies were screened, one sample was isolated for sequencing. Sample 8 (figure 5.5) was isolated, cultured and DNA extracted using a plasmid Miniprep Kit (BioRad). The insert DNA was amplified by PCR using primers (GL2 and RV3) that flanked the multiple cloning site (as described in section 2.2.12). The PCR products were checked for size and purity by gel electrophoresis. The correct size for the amplified topoisomerase IIα -617 products from pGL3B was about 700 base pairs (refer to figure 5.6), where the increase in insert size was due to the presence of the multiple cloning site.

PCR was used to amplify the insert for sequencing analysis, as direct sequencing of the PCR products was found to be more efficient and reliable than direct sequencing of plasmid DNA. PCR products were purified using QIAgen PCR clean-up kit, quantified by gel electrophoresis against DNA standards and sent to the Massey University DNA sequencing service for sequencing. The sequencing primers used were GL2 and RV3, both at a concentration of 0.8 picomol/μl.

Analysis of the pGL3B –617 topoisomerase IIα insert sequences was carried out using the bestfit program (GCG Version 9.1; Genetics Computer Group, Madison, Wisconsin) against the human topoisomerase IIα promoter sequence (accession number X66794). This enabled the identification of the particular elements (ICB1-5 and GC1-2) within each insert and the mutations that were present. Each of the pGL3B –617 topoisomerase IIα plasmids were confirmed to have the intended mutations present in the correct location(s) within the insert DNA (refer to appendix 5). The pGL3B –617 topoisomerase IIα plasmids were prepared for transfection using a high purity plasmid preparation CONCERT Maxiprep (GIBCOBRL, life technologies) and quantified using UV Spectrophotometry (section 2.2.5(ii)).

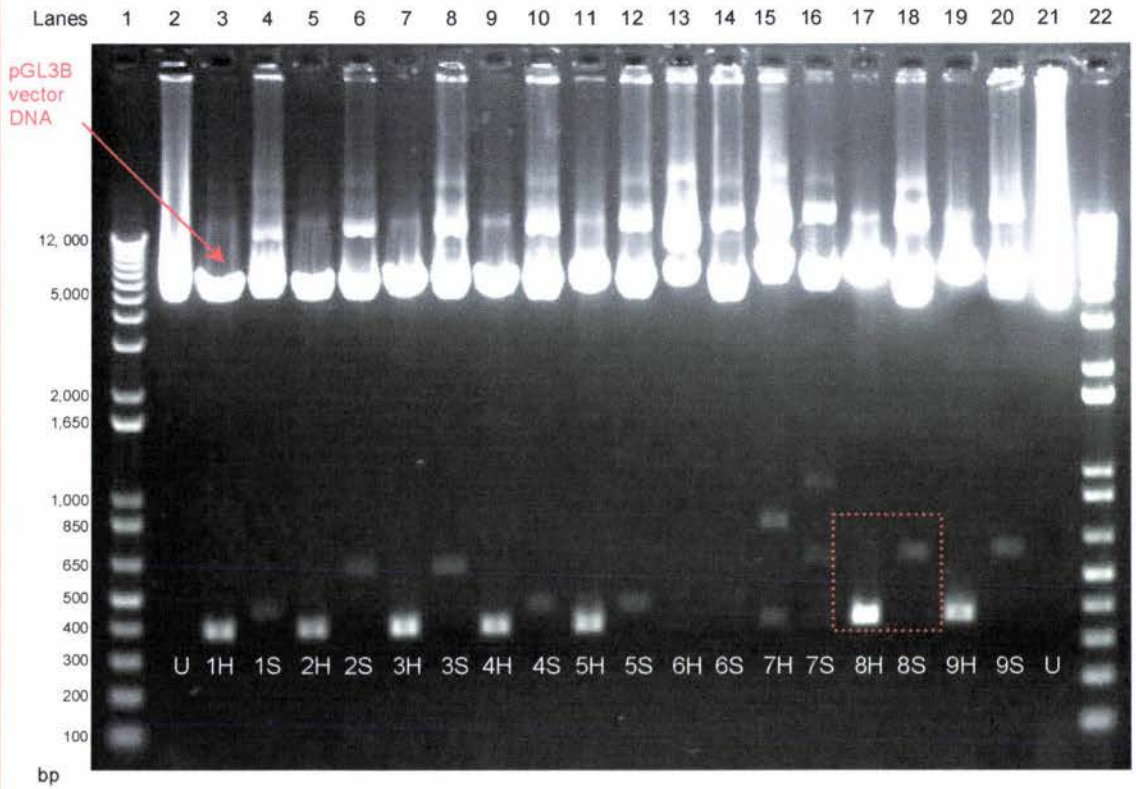


Figure 5.5: Checking plasmid identity by restriction endonuclease digestion. Samples were analyzed on a 1% agarose gel in 1xTAE buffer by electrophoresis for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 μg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp). Individual transformed XL-1 colonies (1-9) were used to extract plasmid DNA. 10 μL of rapid boil DNA was then digested with either *Hin*d III (H) or *Sac* I (S) and samples electrophoresed. 5 μL of uncut (U) plasmid is present in lanes 2 and 21 and 10 μL 1 kb plus ladder in lanes 1 and 22.

Insert orientation was determined by the size of the insert released upon digestion with *Sac* I, a fragment 600 bp in size indicated the insert was in the correct orientation (seen in red box). Insert in the incorrect orientation can be seen in samples 1H & 1S, 4H & 4S, 5H & 5S, 6H & 6S. All *Hin*d III digests revealed 400 bp insert fragments, as expected.

Figure 5.6: PCR products of pGL3B –617 topoisomerase IIα.

Samples were analyzed on a 1% agarose gel in 1xTAE buffer by electrophoresis for about 1 hour at 80 V. DNA was visualized by incorporating ethidium bromide (0.5 μg/mL) into the gel and exposure to UV light. Molecular sizes of bands are indicated on the left in base pairs (bp).

PCR products of the pGL3B –617 topoisomerase IIα insert were about 700 bp.

Lanes	1	10 μL 1 kb plus ladder
	2	5 μL negative control (no template DNA)
	3	5 μL pGL3B GC2mt –617 topoisomerase IIα
	4	5 μL pGL3B GC2mt/ICB1mt –617 topoisomerase IIα
	5	5 μL pGL3B GC2mt/GC1mt –617 topoisomerase IIα
	6	5 μL pGL3B GC2mt/ICB1mt/GC1mt –617 topoisomerase IIα

5.3 Transfections.

Transfections can be either stable or transient. Stable transfections occur as the exogenous DNA is integrated into the host genome. DNA that is stably maintained by cells can often be expressed inefficiently because of position effects or become methylated and rendered transcriptionally inactive. However, transient transfection occurs without DNA integration and is particularly useful for the synthesis of gene products that can demonstrate toxicity at high levels. The level of transient expression of a reporter gene in the transfected cells depends on the number of cells that take up the vector and how well the gene is expressed. Expression of the reporter gene reaches a maximum approximately 48-72 hours post-transfection and then expression declines due to loss of genes. Transfections can be performed a number of different ways, such as lipofection, calcium phosphate precipitation, electroporation and DEAE-dextran-mediated, where each method is valuable for different conditions and cells types.

The method of transfection for this study was lipofection. This transformation system utilizes liposomes prepared from cationic lipids. Liposomes are small unilamellar (single bilayer) vesicles and when DNA in solution is mixed with liposomes, cationic-DNA complexes are spontaneously formed. The positively-charged liposomes not only complex with DNA, but also bind to cultured animal cells (negatively-charged) and are efficient in transforming them, possibly by fusion with the plasma membrane.

There are a number of different reporter genes available such as, β-galactosidase, and chloramphenicol acetyltransferase (CAT). The luciferase reporter gene was chosen to monitor topoisomerase IIα promoter activity in HeLa cells, as it has been used previously with success (M^{c}Lenachan, 1998; Szremska, 2000).

5.3.1 Luciferase assays.

Firefly luciferase from the North American firefly *Photinus pyralis* is an enzyme that consists of a single polypeptide, 550 amino acids long. Luciferase can catalyze the oxidation of luciferin in a reaction requiring both oxygen and ATP, which in turn emits a yellow-green light (refer to figure 5.7). This light can be detected as a direct measurement of the amount of luciferase activity. When HeLa cells were transfected with pGL3B –617 topoisomerase II constructs the luciferase was synthesized within the HeLa cells. The amount of luciferase

protein produced within the cells directly relates to topoisomerase IIα promoter strength, which can be assayed in the HeLa extracts as the light emitted in luciferase assays.

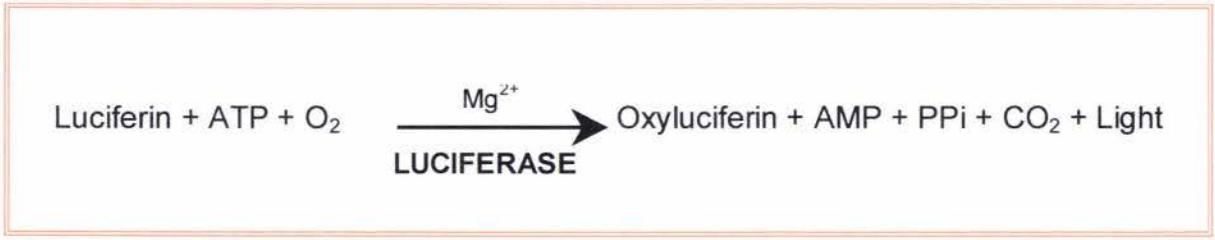


Figure 5.7: Luciferase catalyzed reaction.

Luciferase requires ATP and O_2 to be active, and light is emitted as the substrates are added to the assay. The amount of substrate added to a luciferase reaction directly correlates to the amount of light emitted by the reaction.

The luciferase assays were carried out using the Promega luciferase assay kit and a FLUOstar galaxy (BMG labtechnologies, Germany) detection system. FLUOstar galaxy is a microplate reader that uses a system of fibre optics to detect and relay photon emission. As light is emitted from the luciferase assay (as soon as the luciferase reagent is injected into the assay) photons are released. These photons are detected by the reading head, which is maneuvered to sit directly above each microplate. The reading head is mobile and is automatically positioned above each well before the assay. The reading head relays the photon to detectors (liquid filled fibre optics) within the system and the photon count is recorded according to the assay period stipulated. The readings are relayed through Excel™ (Microsoft Office, 97) and are actual photon counts over time (Stuart Tyler, personal communication; BMG labtechnologies, Melbourne).

The FLUOstar galaxy microplate reader was programmed to take a photon count every second over three minutes, once the luciferase reagent was added. Generally, when luciferase is added to substrates, there is a flash of light proportional to the quantity of enzyme present followed by a rapid decay to give an extended period of low-intensity emission. Therefore, the maximum luciferase reading recorded was used for subsequent calculations. Promega's luciferase chemistry utilizes coenzyme A as a substrate in place of ATP, which results in a slower decay of light, generating longer detection periods. This in conjunction with the accuracy of the FLUOstar detection system enabled data to be recorded with high sensitivity.

5.3.2 β-Galactosidase assays.

In addition to being transfected with various pGL3B –617 topoisomerase IIα constructs, HeLa cells were co-transfected with a pSV-β-galactosidase control vector (Promega, see appendix 2). This was carried out to monitor the interior variations within cells including variations such as DNA uptake and expression, cell density and integrity. The pSV-β-galactosidase control vector enables the expression of bacterial β-galactosidase within transfected cells, which is examined using a spectrophotometric assay.

Figure 5.8: β-galactosidase catalyses the synthetic substrate ONPG.
The exogenous amount of β-galactosidase produced in cells is determined by measuring the amount of *o*-nitrophenol (yellow) production from the assay above.

β-galactosidase generally catalyses the hydrolysis of galactoside sugars such as lactose however it can also cleave a synthetic substrate, *o*-nitrophenyl-β-galactoside (ONPG). ONPG catalysis produces *o*-nitrophenol, which is yellow in colour and can be detected at 405 nm (refer to figure 5.8). Therefore, the exogenous amount of β-galactosidase produced in cells is determined by measuring the amount of ONPG hydrolysis. Cell extracts were assayed for both luciferase and β-galactosidase activity in order to normalize the luciferase values obtained. In order to ensure only exogenous β-galactosidase activity was measured, the assays were conducted at pH 8.0, which inhibits any mammalian β-galactosidase activity but has no effect on bacterial β-galactosidase activity.

5.3.3 Analysis of data.

Due to the sensitive nature of the luciferase assays there tends to be some variation in assay results. For this reason each transfection was performed in triplicate and repeated at least three times to ensure accuracy (see appendix 7). Both luciferase and β-galactosidase assays were carried out on each sample once the cell extracts were prepared.

Luciferase and β-galactosidase activities were corrected by subtracting the average of no less than three appropriate blank samples, to accommodate for any background signals. Luciferase activities were normalized against β-galactosidase activities (as described in figure 5.9) by dividing the observed luciferase value from the HeLa cell extracts with the apposite β-galactosidase value (refer to table 5.3).

Luc/β-gal = Normalized luciferase activity.

Figure 5.9: Normalizing the luciferase activities.
Luciferase activities were normalized to compensate for internal expression variations within HeLa cells.

Constructs	Maximum luciferase values	Corrected maximum luciferase values	β-gal values	Corrected β-gal values	Luc/β-gal ratio	Average Luc/β-gal ratio	Luciferase activity relative to wt %
Blank	841	-	0.139	-	-	-	-
	883	-	0.134	-	-	-	-
	854	-	0.132	-	-	-	-
Avg blank	859	-	0.135	-	-	-	-
wt-617	23157	22298	0.553	0.418	53344	-	-
	22584	21725	0.507	0.372	58401	-	-
	26748	25889	0.640	0.505	51265	54337	100 %
ICB1mt	10557	9698	0.400	0.265	36585	-	-
	9095	8236	0.326	0.191	31079	-	-
	10033	9174	0.389	0.254	36118	34594	64 %

Table 5.3: Calculations involved in analysis of one set of triplicate HeLa cell extracts.
Representative results from a transient transfection experiment. Each transfection was repeated three times, then analyzed as described above, then each set of transfections was repeated a further two times (see appendix 8). The luciferase maxima and β-gal results were corrected by subtracting the relevant averaged blank value.

To normalize the luciferase values each corrected luciferase value was divided by the corresponding corrected β-gal value, then averaged. Most samples were displayed relative to the wt sample, therefore the normalized luciferase values were divided by the wt value, hence wt is arbitrarily set at 100%.

The normalized luciferase activities were expressed as a percentage of wild type –617 topoisomerase II pGL3B luciferase activity, which was determined by dividing by the average of no less than three normalized luciferase activities from transfections of the wt pGL3B –671 topoisomerase IIα plasmid. Therefore, the wild type topoisomerase IIα expression was arbitrarily set at 100%, and any deviation from this is illustrated as an increase or decrease relative to wild type activity.

As in any experiments, a certain amount of error is expected. The average deviation (AVGDEV) is a measure of the variability within a data set and is ideal for smaller data sets. The formula to calculate the average deviation from the mean is as follows:

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

Where n = the number of observations

X = the individual observation

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

Figure 5.10: Formula to calculate the average deviation from the mean.

This formula was used to calculate the fluctuations within data sets. It depicts the amount of variation between samples from the calculated mean. The calculations for the average deviation were performed through Excel (97, Microsoft), this function returns the average of the absolute deviations of data points from their mean.

So using the experiment illustrated in table 5.3, the average deviation of the wt sample would be calculated as follows:

((53344-54337) + (58401-54337) + (51265-54337)) x (1/3) = 2709 luciferase units

and the percentage error: (AVGDEV/AVE) x 100

(2709/54337) x 100 = 4.99% error within the wt triplicates.

The statistical significance of the differences that were observed between two data sets was verified using the statistical *t*-test (refer to figure 5.11 for formula). The *t*-tests deals with the problems associated with using a small sample size, therefore it employs a *t*-distribution as opposed to normal-distribution which is used for larger data sets. *t*-tests (assuming equal variances) were performed using Excel (97, Microsoft) and the probabilities associated with

the *t*-tests (p-value) were generated. This *t*-test assumes that both of the means of both data sets are equal and calculates the probability of them being identical. If both data sets were similar, then the p-value would be greater than 0.10, which means that there is at least a 10% chance that both samples share the same mean. However, if the p-value was less that 0.05, then both data sets are significantly different and they have a less than 5% chance of being similar (refer to table 5.4).

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

Where the pooled variance (s^2) is estimated by:

$$s^2 = \frac{(n_{grp1} - 1)\, s^2_{grp1} + (n_{grp2} - 1)\, s^2_{grp2}}{n_{grp1} + n_{grp2} - 2}$$

Figure 5.11: *t*-test for comparing two populations.

The formula used to calculate the *t*-statistic within Excel (97, Microsoft). The *t*-statistic was then relayed into a p-value. The letter n, stands for number of data points within a population. s^2 is the variance of a population; the *t*-test assumes equal variance between the two populations.

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

Table 5.4: Strength of evidence provided by the p-value.

The p-value, the probability associated with the *t*-test. The p-value can provide evidence of differences in varying degrees.

5.4 Transient transfections using HeLa cells.

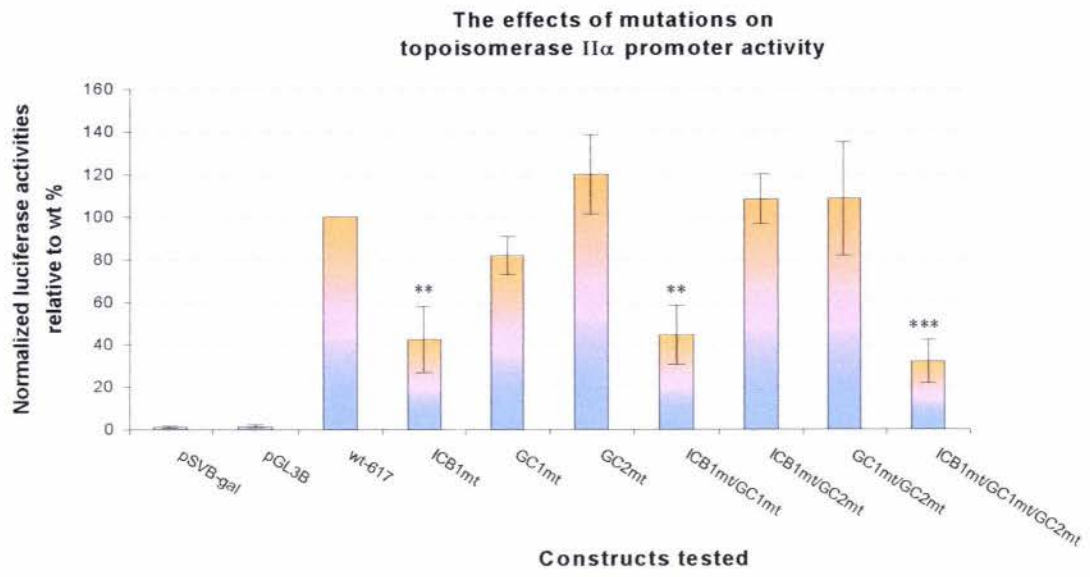
Transfections were carried out as described in section 2.2.25, using the lipofection agent FuGENE^{TM} 6 (Roche). FuGENE^{TM} 6 is unique because it is made from a blend of lipids and other proprietary components that enables high efficiency transfections. In addition, FuGENE^{TM} 6 requires that cells are between 50-80% confluent at the time of transfection. This is of particular significance as the expression of topoisomerase IIα has been shown to change in confluence arrested cells (Isaacs *et al*, 1996a).

The detection of transient reporter gene activity 12-72 hours after transfection allows rapid testing of putative regulatory sequences. HeLa cells were transfected with various reporter gene constructs containing the -617 minimal promoter of topoisomerase IIα as well as the pSV-β-galactosidase control vector.

5.4.1 Expression of topoisomerase IIα with mutations in ICB1, GC1 and GC2.

HeLa cells previously transfected with 0.5-2 μg of plasmid DNA (pGL3B topoisomerase IIα and pSV-β-galactosidase control vector) have shown that this is sufficient DNA for successful reproducible transfections (Szremska, 2000). To minimize the cost but not the efficiency of transfections HeLa cells were transfected with 1 μg of various pGL3B –617 topoisomerase IIα constructs and 1 μg of pSV-β-galactosidase control vector.

HeLa cells were transfected when they were approximately 70% confluent and harvested 60-65 hours after transfection as described in section 2.2.26. β-galactosidase and luciferase activities were determined promptly after harvesting the HeLa extracts (sections 2.2.27 and 2.2.28, respectively); figure 5.12 illustrates the results for these transfections.



Constructs tested	Normalized luciferase activities relative to wt%	p-values	Significant Difference
pSV-β-galactosidase	1.2 ±0.6%	-	-
pGL3B	1.6 ±0.7%	-	-
wt -617	100%	-	-
ICB1mt	42.2 ±15.5%	0.001	Y**
GC1mt	81.7 ±8.9%	0.177	N
GC2mt	119.8 ±18.6%	0.287	N
ICB1mt/GC1mt	44.2 ±13.9%	0.002	Y**
ICB1mt/GC2mt	108.2 ±11.8%	0.632	N
GC1mt/GC2mt	108.3 ±26.4%	0.683	N
ICB1mt/GC1mt/GC2mt	31.8 ±10.1%	0.000	Y***

Figure 5.12 : The effect of mutations in elements of topoisomerase IIα.

Reporter vector expression changes depending on the mutation contained in the –617 topoisomerase IIα promoter. HeLa cells were transfected with 1 μg each of pGL3B and pSV-β-galactosidase plasmid DNA and harvested 65 hours later. The normalized luciferase activities are shown relative to wild type activity and are the results of a set of four experiments each done in triplicate (see appendix 7.1, experiments 1-4).

Each construct tested was analyzed for a significant difference to wild type topoisomerase IIα promoter activity using a *t*-test, the p-values are shown above. Whether each construct tested produced a significant difference to wild type topoisomerase IIα promoter activity is indicated by Y (yes) or N (no), with *s indicating the level of evidence provided by the p-value (table 5.4).

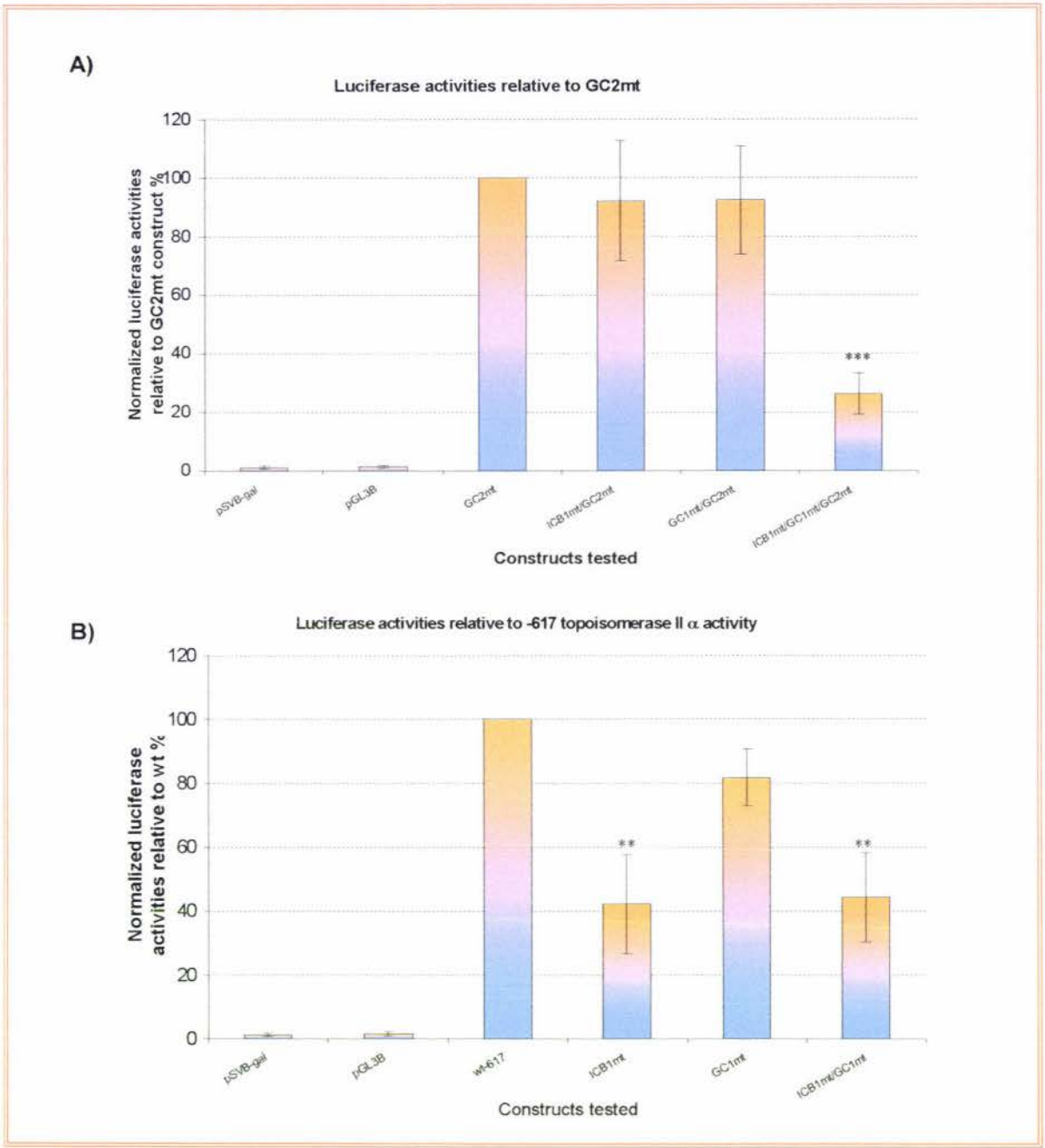


Figure 5.13: Results to experiments 1-4 displayed relative to GC2mt.

Data is the same as presented in figure 5.11, but displayed in this way for clarity and interpretation, to analyze the relevance of elements ICB1, GC1 and GC2. * indicate the level of evidence provided by the p-value (table 5.4) for a significant difference.

A) GC2mt topoisomerase IIα constructs relative to GC2mt. **B)** Corresponding topoisomerase IIα constructs displayed relative to wild type.

Figure 5.12 demonstrates that expression of the reporter gene can vary depending on the mutation in the –617 topoisomerase IIα promoter. Compared to wild type promoter activity, a mutation in ICB1 induced a significant drop in promoter strength (approximately 50%), whereas a mutation in GC1 or GC2 appears to have no effect on promoter strength. However, the construct with mutations in all three elements ICB1, GC1 and GC2 was observed to have the least reporter gene activity as luciferase activity decreases to around 40% of that observed in wild type, suggesting a combination of all three elements is required to maintain wild type topoisomerase IIα promoter strength. When both ICB1 and GC1 were mutated, the promoter activity was only about 45% of wild type. Transcriptional activity from each of the constructs ICB1mt, ICB1mt/GC1mt and ICB1mt/GC1mt/GC2mt were not significantly different.

At first glance, it may appear as though a mutation in GC2 could have the ability to increase wild type promoter activity, which could suggest that the GC2 element may have a repressive role. However, once a *t*-test was performed, this GC2 induced change was not significant. Interestingly, each construct that displayed a significant decrease in wild type promoter activity contained a mutation in the ICB1 element, however this effect appeared to be neutralized by the presence of a mutation in GC2 (construct GC2mt /ICB1mt). In order to examine this phenomenon in more detail, the transfection data was displayed relative to GC2mt (refer to figure 5.13A).

The trends observed in figure 5.13 are similar to those seen in figure 5.12, and clearly show that there is no difference in promoter activity between constructs GC2mt, ICB1mt/GC2mt and GC1mt/GC2mt. Figure 5.13B clearly shows that ICB1 is important for maintaining wild type promoter activity and there is clearly no difference between constructs ICB1mt and ICB1mt/GC1mt promoter activity.

Taken together these data demonstrate that only three of the seven constructs tested produced a significant decrease in expression of the reporter gene, relative to wild type. Each construct that displayed a difference contained a mutation in the ICB1 element, indicating that ICB1 is the most important element for basal transcription. In addition, GC2 may have a role in the ICB1 induced decrease in promoter activity.

5.4.2 Transient co-transfections.

The results described in chapter 4 demonstrated that GC1 has the ability to bind both Sp1 and Sp3, whereas previous work had suggested that only Sp1 was able to associate with GC1 (Szremska, 2000). Transcription factors Sp1 and Sp3 have been implicated in the regulation of topoisomerase IIα expression *in vivo*, however the extent of each interaction has not been defined (Kubo *et al*, 1995; Yoon *et al*, 1999). In order to investigate the roles of transcription factors (Sp1 and Sp3) on topoisomerase IIα transcription *in vivo*, Sp1 and Sp3 expression vectors were used in co-transfection experiments with the topoisomerase IIα reporter constructs (refer to figure 5.14).

5.4.2(i) Sp1 titration.

The effects that transcription factor Sp1 could have on topoisomerase IIα expression *in vivo* were investigated by co-transfecting HeLa cells with an expression vector that enabled Sp1 to be over-expressed. If Sp1 binds to the topoisomerase IIα promoter and influences transcription, then it should be reflected as a change in relative luciferase activity. In order to optimize the amount of Sp1 expression vector used in co-transfections, varying amounts of Sp1 expression vector was added to transient transfection experiments. Topoisomerase IIα constructs chosen for these experiments were the pGL3B wt –617 topoisomerase IIα and pGL3B ICB1mt –617 topoisomerase IIα. The latter pGL3B construct was tested due to its low expression levels comparative to wild type and also because of possible interactions between NF-Y and Sp1 at GC1.

Figure 5.15 illustrates that the addition of Sp1 generally increases luciferase activity; hence, Sp1 is capable of activating the topoisomerase IIα promoter. A maximum of a six-fold increase was observed upon the addition of 0.5 μg of Sp1 expression vector to pGL3B wt –617 topoisomerase IIα, however further addition of Sp1 (1 μg) appeared to have an inhibitory effect. Comparing the p-values, Sp1 appears to have a lesser overall effect on the ICB1mt construct compared to the wild type construct, suggesting that ICB1 may be required to induce strong Sp1-mediated promoter activation. In general, when larger changes were observed in topoisomerase IIα activity the greater the errors became, therefore it was decided that 0.25 μg of Sp1 would be optimal to use in subsequent assays. 0.25 μg Sp1 generally gave rise to a 2-fold induction of topoisomerase IIα promoter activity, this phenomenon was examined using other topoisomerase IIα constructs.

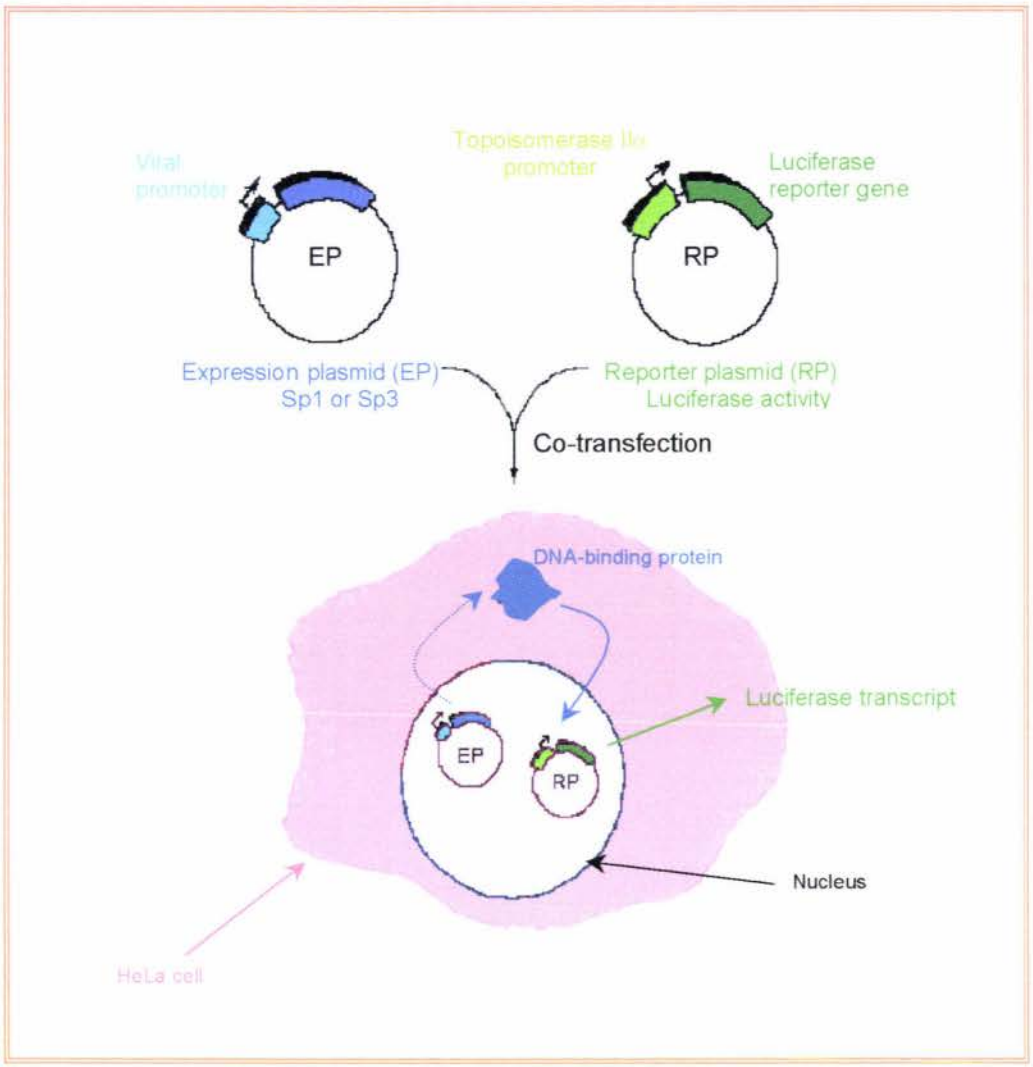
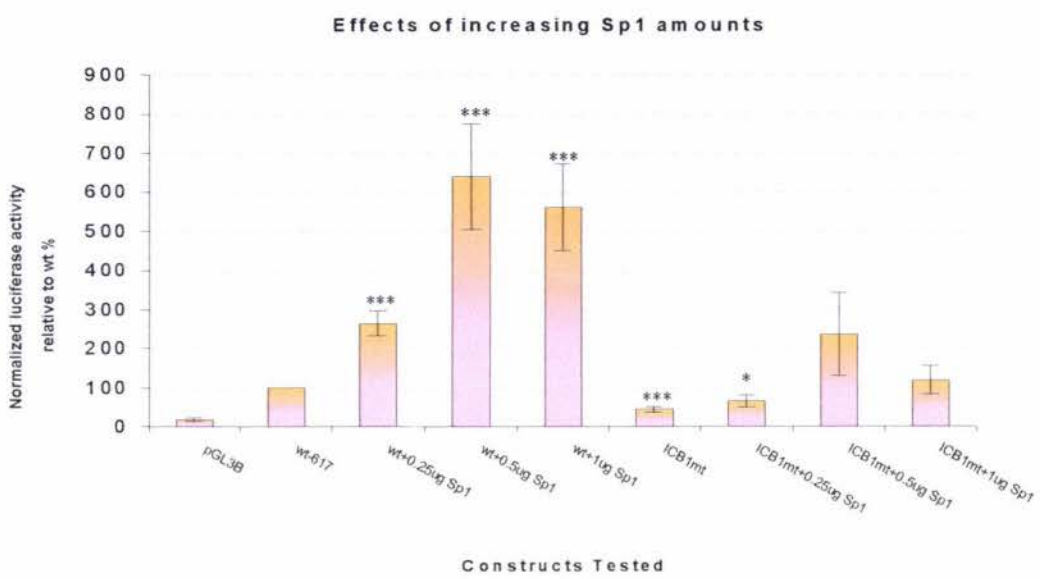


Figure 5.14: Schematic representation of transient co-transfection experiments. HeLa cells (pink) were co-transfected with pGL3B (RP, green) constructs and Sp1 or Sp3 expression vector (EP, blue).

The expression vector enables Sp1or Sp3 to be over-expressed in HeLa cells where it may interact with regions of the topoisomerase IIα promoter, thereby influencing luciferase expression (adapted from: Carey and Smale, 2000).



Constructs tested	Normalized luciferase activities relative to wt %	p-values relative to wt -617	p-values relative to ICB1mt	Significant difference to wt	Significant difference to ICB1mt
pGL3B	17.8 ±5.2%	-	-	-	-
wt –617	100%	-	-	-	-
wt + 0.25 μg Sp1	264.4 ±31.8%	0.002	-	Y**	-
wt + 0.5 μg Sp1	639.1±134.9%	0.009	-	Y**	-
wt + 1.0 μg Sp1	560.8 ±111.9%	0.008	-	Y**	-
ICB1mt	43.61 ±7.1%	0.001	-	Y**	-
ICB1mt+ 0.25 μg Sp1	64.7 ±15.42%	0.049	0.202	Y*	N
ICB1mt+ 0.5 μg Sp1	236.2 ±105.9%	0.182	0.085	N	Y
ICB1mt+ 1.0 μg Sp1	119.2 ±36.9%	0.530	0.570	N	N

Figure 5.15: Sp1 titrations.

Reporter vector expression changes depending on the amount of Sp1 expression vector present in HeLa cells. HeLa cells were transfected with 0.5 μg of each pGL3B and pSV-β-galactosidase in addition to varying amounts (0-1 μg) of Sp1 expression plasmid and harvested about 65 hours after transfection. The normalized luciferase activities are displayed relative to wt activity and are the results of a set of three experiments each carried out in triplicate (see appendix 7.2, experiments 5-7).

Each construct tested was analyzed for a significant difference to wild type topoisomerase IIα promoter activity using a *t*-test; the p-values are shown above. Whether each construct tested produced a significant difference to wild type topoisomerase IIα promoter activity is indicated by Y (yes) or N (no), with *s indicating the level of evidence provided by the p-value (table 5.4).

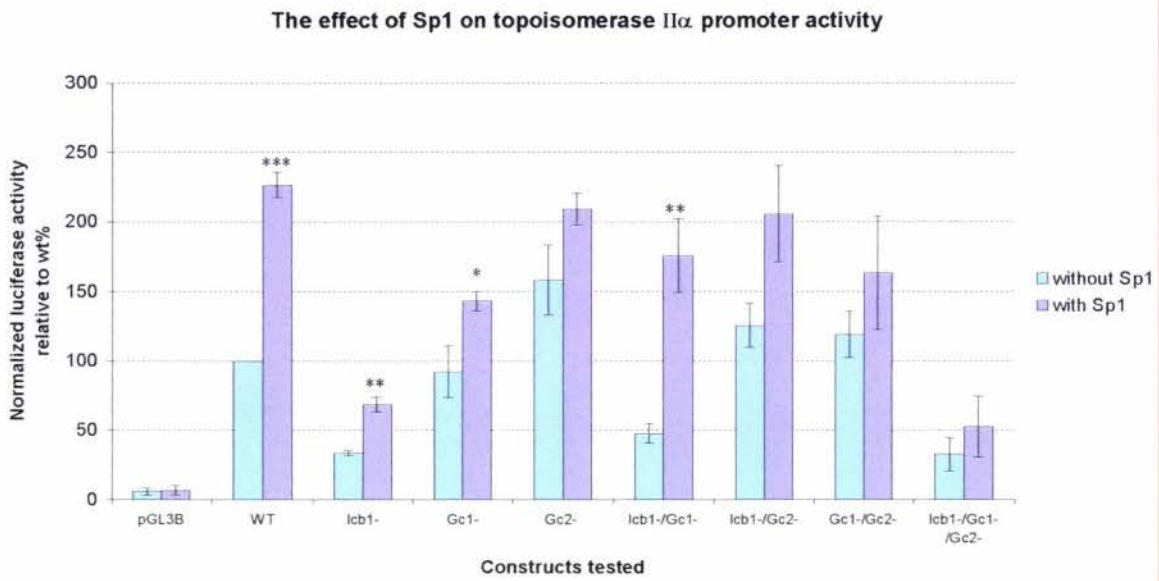
5.4.2(ii) Addition of Sp1 to topoisomerase IIα constructs.

Each topoisomerase IIα construct was tested with the addition of 0.25 μg Sp1 as shown in figure 5.16. The wt topoisomerase IIα construct exhibited a 2-fold increase in luciferase activity, as did the ICB1mt construct, perhaps indicating that even though ICB1 is an important element, it may not be the sole regulatory element within the topoisomerase IIα promoter and that ICB1 is probably not the target element for Sp1-mediated up-regulation of the topoisomerase IIα promoter. Regardless of the degree of Sp1-mediated activation in the ICB1mt, it is still does not exceed the activity of the wild type topoisomerase IIα promoter. Mutations in both GC1 and GC2 abolished any significant Sp1-mediated activation, suggesting that Sp1-mediated induction of topoisomerase IIα promoter activity relies on these two elements.

The greatest induction of luciferase activity (about a 3-fold increase) was observed using the ICB1mt/GC1mt construct in the presence of 0.25 μg Sp1, suggesting that GC2 may be involved in this increase. The GC1mt construct exhibited a reduced Sp1-mediated increase (compared to wild type and ICB1mt/GC1mt) and the GC2mt construct did not generate any significant change in promoter activity in the presence of Sp1. In fact, none of the topoisomerase IIα constructs that harboured a mutation in the GC2 element appeared to be greatly influenced by the presence of Sp1. Taken together, this suggests that GC2 is probably an important element required for Sp1-mediated topoisomerase IIα promoter activation and perhaps some sort of co-operativity between ICB1 and GC1 exists in determining the potency of Sp1-activation via GC2.

In addition, a mutation in GC2 appears to exhibit an increased topoisomerase IIα activity compared to wild type (refer to figure 5.17A). A *t*-test was performed to compare wt and GC2mt promoter activity (p-value = 0.042) and the results indicated that GC2 could be playing a repressive role in topoisomerase IIα promoter activity.

Collectively these experiments show that ICB1 still appears to be the most important element, however GC2 is probably important in Sp1-mediated promoter activity in conjunction with GC1, although GC2 may have a role in the repression of topoisomerase IIα promoter activity.



Constructs tested	Normalized luciferase activities relative to wt% Without Sp1	With Sp1	p-values	Significant Difference
pGL3B	5.95 ±2.5%	4.15 ±3.4%	-	-
wt –617	100%	226.6 ±9.1	0.000	Y***
ICB1mt	33.6 ±1.7%	68.6 ±5.7%	0.001	Y**
GC1mt	92.0 ±18.6%	143 ±7.1%	0.036	Y*
GC2mt	158.3 ±25.5%	209.4 ±11.4%	0.077	N
ICB1mt/GC1mt	47.4 ±7.1%	175.8 ±26.6%	0.004	Y**
ICB1mt/GC2mt	125.3 ±15.7%	205.3 ±34.9%	0.062	N
GC1mt/GC2mt	118.9 ±16.7%	163.4 ±40.9%	0.252	N
ICB1mt/GC1mt/GC2mt	32.5 ±12.2%	52.4 ±21.9%	0.397	N

Figure 5.16: Addition of Sp1 to topoisomerase IIα constructs.

Reporter vector expression changes in the presence of Sp1 in HeLa cells. HeLa cells were transfected with 0.5 μg each of pGL3B and pSV-β-galactosidase plasmid DNA in addition to 0.25 μg of Sp1 expression plasmid and harvested about 60 hours later. The normalized luciferase activities are shown relative to wild type activity and are the results of a set of three experiments each done in triplicate (see appendix 7.3, experiments 8-10).

Each construct tested was analyzed for a significant difference in promoter activity upon addition of Sp1, using a *t*-test; the p-values are shown above. Whether each construct tested produced a significant difference upon addition of Sp1 is indicated by Y (yes) or N (no), with *s indicating the level of evidence provided by the p-value (table 5.4).

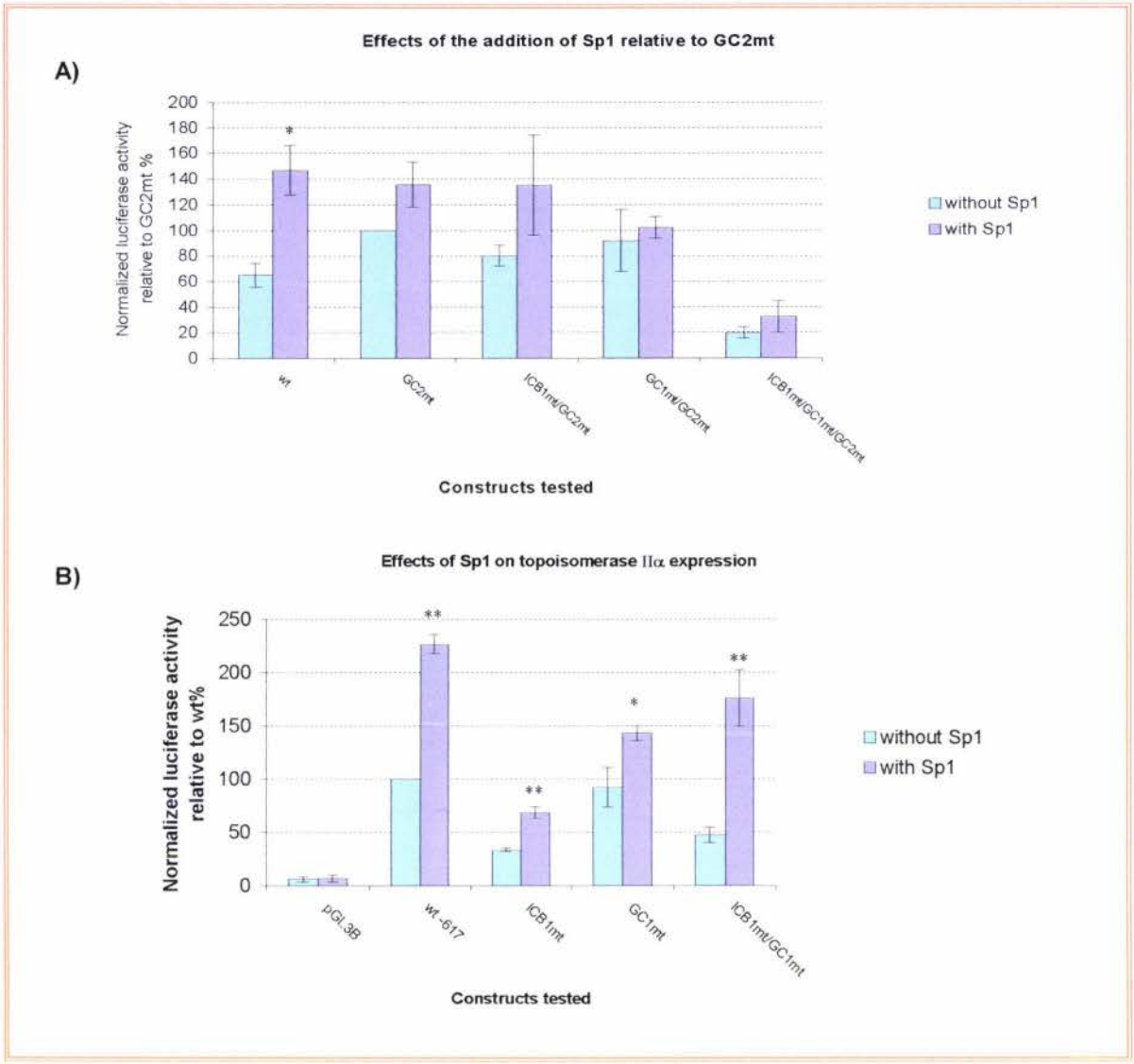


Figure 5.17: Results of experiments 8-10 displayed relative to GC2mt.

Data is the same as presented in figure 5.16, but displayed in this way for clarity and interpretation, to analyze the relevance of elements ICB1, GC1 and GC2. * indicate the level of evidence provided by the p-value (table 5.4) for a significant difference.

A) GC2mt topoisomerase IIα constructs relative to GC2mt with the addition of 0.25 μg Sp1 expression plasmid. No significant differences were observed in any of these constructs with the addition of Sp1. The GC2mt construct does appear to exhibit elevated promoter activity compared to the wild type construct. **B)** Corresponding topoisomerase IIα constructs displayed relative to wild type with the addition of 0.25 μg Sp1 expression plasmid. The largest increase can be observed in ICB1mt/GC1mt when Sp1 is added (approximately 3 fold increase). Mutations in either ICB1 or GC1 appear to inhibit Sp1-mediated topoisomerase IIα promoter activation.

5.4.2(iii) Addition of Sp3 to topoisomerase IIα constructs.

Sp3 is another transcription factor that has been shown to bind to GC1 (chapter 4) and GC2 of the topoisomerase IIα promoter (Szremska, 2000). HeLa cells were co-transfected with an Sp3 expression vector to investigate the *in vivo* effects this transcription factor may have on topoisomerase IIα expression. Increasing amounts of Sp3 expression vector were added to standard transfection experiments and the luciferase activity monitored (results are in appendix 8.4). Unfortunately, these Sp3 co-transfection experiments generated β-galactosidase activities that appeared to be affected by the amount of Sp3 added to each assay (refer to table 5.5, appendix 7.4). As the pSV-β-galactosidase control vector acts as an internal control for transient transfections, the inconsistencies in the β-galactosidase values meant that results for these experiments may not be valid. It was also possible that β-galactosidase activities were influenced by the buffers and reagents used in each assay. The Sp3 titration experiments were repeated using fresh buffers and reagents however, β-galactosidase values were still affected by the amount of Sp3 added to each assay (appendix 7.4, experiments 12 and 13).

It was possible that the amount of pSV-β-galactosidase control vector added to HeLa cells during transfections could have an affect on the level of Sp3-mediated β-galactosidase activities. Therefore, HeLa cells were transfected with 0.5 μg wt –617 topoisomerase IIα pGL3B and increasing amounts of pSV-β-galactosidase control vector. The results from this experiment (refer to appendix 7.5) generated β-galactosidase values that were consistent with the amount of pSV-β-galactosidase control vector added to each assay, suggesting that the pSV-β-galactosidase control vector was not the direct problem. The pSV-β-galactosidase control vector contains the β-galactosidase gene, which is controlled by a SV40 promoter (see appendix 3 for vector map). The SV40 promoter contains multiple GC elements to which Sp3 is thought to be able to interact. Therefore, it is plausible that over-expressed Sp3 in HeLa cells was not only targeting the topoisomerase IIα promoter sequence but also the SV40 promoter, thereby affecting β-galactosidase expression *in vivo*.

LUCIFERASE MAXIMA

pGL3B	wt -617	wt +0.2 µg Sp3	wt +0.5 µg Sp3	wt +1 µg Sp3	ICB1mt	ICB1mt +0.25 µg Sp3	ICB1mt +0.5 µg Sp3	ICB1mt +1 µg Sp3	BLANK
824	40267	8473	15455	6390	17201	7166	15449	2508	536
837	50203	9295	15399	7279	20876	7903	15498	2671	544
814	43904	9591	15857	7438	20311	7576	16269	2897	556
									545.33

BGAL READINGS

0.281	0.403	0.233	0.212	0.196	0.431	0.282	0.239	0.181	0.13
0.308	0.414	0.232	0.218	0.194	0.455	0.294	0.269	0.184	0.125
0.264	0.378	0.234	0.216	0.188	0.491	0.294	0.255	0.183	0.126
									0.127

CORRECTED LUCIFERASE VALUES

278.67	39722	7927.7	14910	5844.7	16656	6620.7	14904	1962.7
291.67	49658	8749.7	14854	6733.7	20331	7357.7	14953	2125.7
268.67	43359	9045.7	15312	6892.7	19766	7030.7	15724	2351.7

CORRECTED BGAL READINGS

0.154	0.276	0.106	0.085	0.069	0.304	0.155	0.112	0.054
0.181	0.287	0.105	0.091	0.067	0.328	0.167	0.142	0.057
0.137	0.251	0.107	0.089	0.061	0.364	0.167	0.128	0.056

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

1809.5	143919	74789	175408	84705	54788	42714	133068	36346
1611.4	173023	83330	163227	100502	61984	44058	105300	37292
1961.1	172744	84539	172041	112995	54301	42100	122841	41994

AVERAGE OF NORMALIZED LUCIFERASE VALUES

1794	163229	80886	170225	99401	57024	42957	120403	38544

NORMALIZED LUCIFERASE RELATIVE TO WT %

1.0991	100	49.554	104.29	60.897	34.935	26.317	73.764	23.614

AVG DEV FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

121.72	12873	4064.6	4665.5	9797	3306.2	733.78	10069	2300
percentage error								
6.785	7.8865	5.025	2.7408	9.856	5.7978	1.7082	8.3624	5.9672

Table 5.5: Results from co-transfection experiment 11, increasing amounts of Sp3. One set of triplicate results from transient co-transfections. Increasing amounts of Sp3 expression vector was added to each experiment, which generated a significant decrease in β-galactosidase activities (highlighted in red).

The differences in β-galactosidase activities were tested using *t*-tests (data not shown) and the differences observed were significant to produce p-values of less than 0.001. Therefore, the luciferase activities could not be interpreted with any confidence.

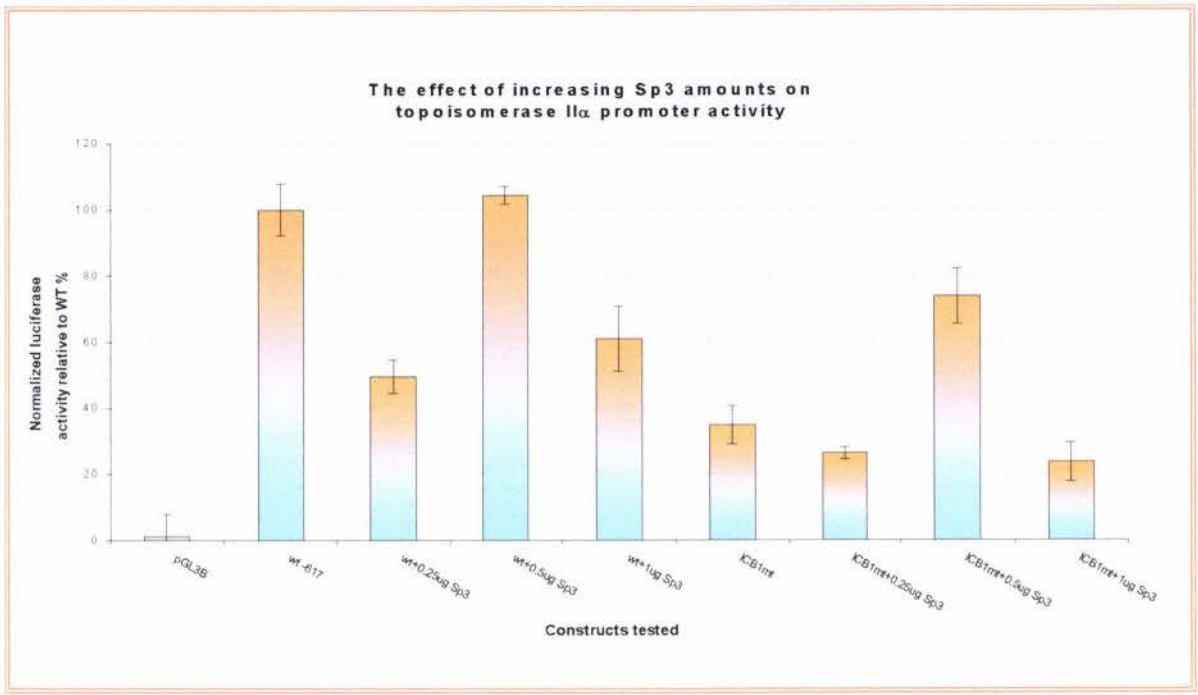


Figure 5.18: Results from experiment 11, Sp3 titrations.

HeLa cells were transfected with 0.5 μg of each pGL3B and pSV-β-galactosidase DNA in addition to varying amounts of Sp3 (0 – 1 μg) expression plasmid. The normalized luciferase activities are shown relative to wild type activity and are the results of one experiment each carried out in triplicate (see appendix 7.4, experiment 11). Even though this graph shows that Sp3 may have an inhibitory effect on topoisomerase IIα expression, other results (see appendix 7.4 experiments 13 and 14) did not exhibit this same trend. Although the trends in this graph cannot be interpreted with any confidence, Sp3 does appear to have some affect on luciferase activity.

Even though the results from any of the Sp3 expression plasmid co-transfection experiments could not be interpreted with any confidence, it is clear that Sp3 could have some affect on luciferase activity. It was unfortunate that Sp3 had such an effect on the β-galactosidase activities, as this meant it could not be further investigated at this stage. Sp3 is generally bifunctional, being able to act as both a repressor and an activator of transcription in some cases (Noti, 1997). It has also been suggested that the relative levels of Sp1 and Sp3 could dictate whether Sp3 acts as a repressor or an activator (de Luca *et al*, 1996). An experiment had been designed to test the effect of different ratios of Sp1 and Sp3 on the topoisomerase IIα promoter, but this could not be carried out due to the problems observed with the β-galactosidase control.

5.5 Chapter summary.

A maximum of 6-fold increase in topoisomerase IIα promoter activity was observed when wt –617 topoisomerase IIα constructs were exposed to 0.5 μg of Sp1 co-expression vector. However, Sp1 in higher quantity (1 μg Sp1 co-expression vector) also appeared to have an inhibitory effect. The –617 topoisomerase IIα construct that exhibited the least amount of promoter activity had mutations in all three elements.

The GC2mt construct seemed to generate luciferase activities that were generally higher than what was observed in the wild type topoisomerase IIα constructs. However, when each of the GC2mt constructs was exposed to Sp1 co-expression vector, there was no significant induction of promoter activity. The transient transfection data suggest that GC2 could normally be a repressive element, however in the presence of elevated levels of Sp1, GC2 may act as an activating element.

Constructs that had a mutation in the GC1 element (unless accompanied by a mutation in ICB1) generally appeared to be influenced less by the presence of Sp1 co-expression vector, where no significant increases in luciferase activity were observed in the presence of 0.25 μg of Sp1 expression vector. This demonstrates that perhaps GC1 is of less significance for the observed Sp1-mediated topoisomerase IIα promoter up-regulation, than GC2.

Mutations in ICB1 appeared to generate the greatest decrease in topoisomerase IIα promoter activity, compared to wild type. Therefore, of the three elements examined ICB1 appears to be the most important regulatory element within the topoisomerase IIα promoter for basal level transcription.

These findings combined with the data from chapter 4 can be interpreted to suggest a model for the regulatory mechanisms involved in topoisomerase IIα transcription and will be discussed in chapter 6.

Chapter 6: Discussion and Future work.

6.1 Topoisomerase IIβ.

The information that is currently available on the regulation and expression of topoisomerase IIβ is limited, therefore it was envisaged that this could be elaborated upon through the examination of the topoisomerase IIβ promoter sequence. Unfortunately, it was not possible to isolate any putative topoisomerase IIβ promoter DNA from clones that had been gifted. An important lesson was learnt from this experience; when uncharacterized clones are obtained from elsewhere they may not yield the expected results, and therefore should be used with caution. This work was abandoned prior to the publication of the human genome sequence. A putative topoisomerase IIβ promoter sequence was identified through a search of the draft human genomic sequence, but no functional work was attempted.

6.1.1 Analysis of the presumptive topoisomerase IIβ clones.

A total of 6 clones were analyzed for the presence of topoisomerase IIβ promoter DNA. Of the two pBluescript clones, sequencing of the TopoB3 pBluescript (+) clone showed that the insert DNA contained part of the human topoisomerase IIβ exon 2 and part of intron 1, which was consistent with findings of Lang *et al* (1998) in that exon 2 is followed by intron 1 within the gene. Sequence analysis of the other pBluescript clone TopoB3.2 indicated that the insert DNA was orientated in the opposite direction, sequencing towards exon 2, therefore it could not contain topoisomerase IIβ promoter DNA either.

Southern blots of the 4 λ clones (figure 3.15B) isolated from a genomic library indicated the presence of putative topoisomerase IIβ promoter, as the PCR-generated topoisomerase IIβ promoter fragment hybridized to each insert DNA. However, when one of the inserts was sequenced (each insert was the same size), it clearly was not topoisomerase IIβ DNA suggesting that hybridization within the Southern blots was not specific. One of the reasons that the topoisomerase IIβ probe could have hybridized to each insert is possibly due to the probe being too small. Therefore, it was not specific enough when hybridizing to larger sequences, such as the λ clone(s) insert DNA (3 kb+). In addition, the oligonucleotide sequence was GC rich (about 65%), which may have increased the chances of the oligonucleotide annealing to multiple sites. To overcome this problem, Southern blotting could be carried out using a larger DNA fragment as a probe. Alternatively, a higher

stringency wash or hybridization could be used to avoid non-specific hybridization. The availability of the human genome (draft) sequence would have allowed the PCR amplification of the topoisomerase IIβ promoter sequence, to generate another probe for Southern blots. However, it was not possible to pursue this approach as no additional clones were available at the time and it was not feasible to obtain a human genomic library for screening.

6.1.2 In silico generated topoisomerase IIβ promoter sequence.

An *in silico* method was used to successfully identify 2.5 kb of the 5'-flanking topoisomerase IIβ sequence (illustrated in figure 3.19). This sequence shows that the majority of the putative transcription factor sites are located around 600-700 bp upstream and 200-400 bp down-stream of the major transcription start site (figure 3.19, around nucleotide 1942) however these are mainly GC boxes, which contribute to the GC rich nature of the sequence (about 80%). There are only three perfect inverted CCAAT boxes, two that are close together (figure 3.19, positions 1420-1424 and 1452-1456) about 500 bp away from the transcription start site and the other about 1 kb upstream from them.

As future work, it should be possible to generate the upstream region of topoisomerase IIβ using a number of different methods. One way would be to purchase a human genomic library and probe the library with a specific topoisomerase IIβ sequence (possibly from exon 1). A quicker alternative would be to use the topoisomerase IIβ promoter sequence (illustrated in figure 3.19) and design primers that will amplify the 5'-flanking sequences, using human genomic DNA as the template. Consequently, primers could be designed to amplify the entire 2-3 kb 5'-flanking sequence or amplify smaller regions of the sequence, which would then have to be ligated together to generate a larger sequence (refer to figure 6.1 for details). The unique restriction sites required for this could be obtained by using the Map program within GCG (Version 9.1; Wisconsin Genetics Computer Group, USA). To ensure that the PCR-generated topoisomerase IIβ promoter DNA was readily accessible, it should be cloned into a suitable cloning vector using restriction sites on either end of the sequence. Due to the GC rich nature of the topoisomerase IIβ promoter sequence, a GC-rich PCR system (Roche) may be useful, which is specifically designed to amplify up to 5 kb of difficult templates such as genomic DNA.

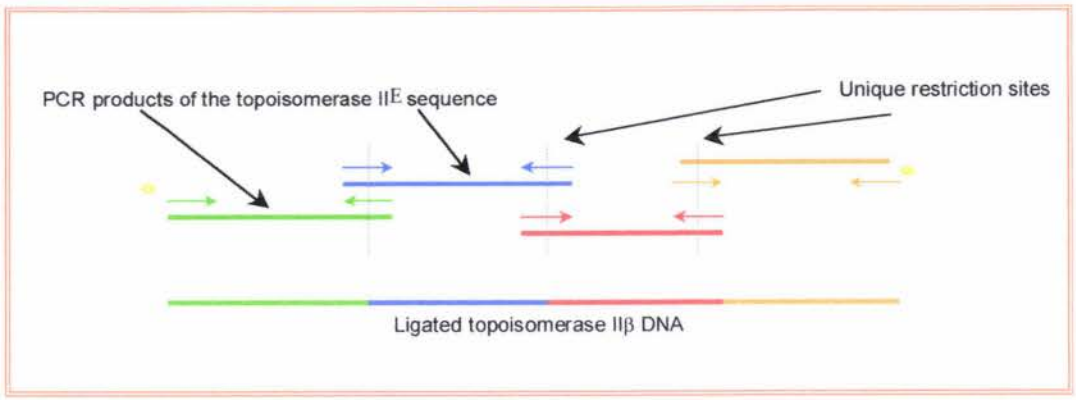


Figure 6.1: Representation of a PCR based method to generate 2- 3 kb of the 5'-flanking sequence of topoisomerase IIβ.

Primers (coloured arrows) could be designed based on the putative topoisomerase IIβ promoter sequence. Each primer would have unique restriction sites (indicated by the dotted lines) incorporated onto the end of their sequence. Thus, the PCR products (coloured lines) will also have these restriction sites, which can then be used to ligate fragments together. The most 3' and 5' primers should include restriction sites that are compatible with those within the multiple cloning site (MCS) of a reporter vector.

Using the TRANSFAC database (Quandt *et al*, 1995), some of the putative transcription factor binding sites within the *in silico*-generated topoisomerase IIβ promoter sequence were predicted using sequence similarities (section 3.5.2). However, the actual regulatory regions within the topoisomerase IIβ promoter cannot be established without experimentation to measure promoter strength in relation to these elements in the promoter. This could be done by generating deletion constructs of the PCR-generated topoisomerase IIβ 5'-flanking sequence and placing them upstream of a promoter-less reporter gene (such as luciferase in pGL3B). The strength of the topoisomerase IIβ promoter would then be examined by transfecting tissue culture cells with the various constructs and examining the relative activities of the reporter gene product (refer to figure 6.2). This would allow the identification of sequences important for transcriptional regulation. It would be feasible to design the deletion series be based on the clustering of the predicted regulatory sequences described earlier. In addition, using knowledge about the importance of ICB boxes within the topoisomerase IIα promoter regulation (Motta *et al*, 2001), one could predict that significant topoisomerase IIβ promoter regulatory elements could be found about 500-600 bp away from the start site. As a starting point, serial deletion studies should be carried out using topoisomerase IIβ promoter sequences that start from about 1 kb upstream of exon 1 and

ending about 100 bases downstream from the ATG codon. In addition to carrying out 5' serial deletions, it would be interesting to examine the effects of 3' serial deletions on topoisomerase IIβ promoter activity.

Once the sequence of topoisomerase IIβ required for minimal promoter activity was determined, it would be fascinating to compare the strength of topoisomerase IIβ promoter activity in comparison to topoisomerase IIα minimal promoter activity as both are differentially regulated *in vivo*. To further investigate the regulatory sequences within the topoisomerase IIβ promoter, internal deletion constructs should also be tested for promoter strength.

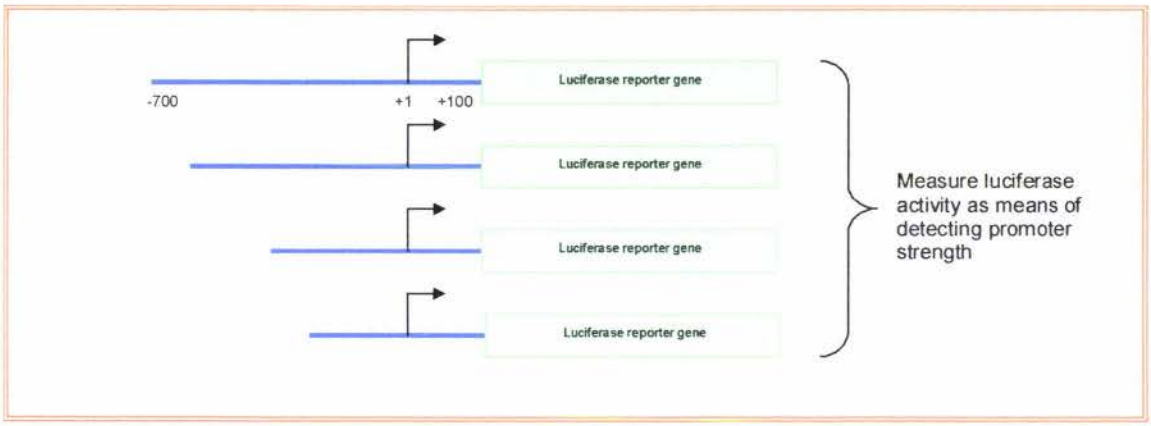


Figure 6.2: Deletion constructs of topoisomerase IIβ promoter sequence.
Deletion constructs could be used to examine important regulatory elements that may be present in the putative promoter sequence. The putative topoisomerase IIβ promoter sequence (blue line) is placed upstream of a promoter-less reporter gene (luciferase) and the resultant plasmid is used to transfect cells. The relative activities of the reporter gene driven by the various topoisomerase IIβ promoter constructs could then be used to identify regions where important regulatory elements may reside within the promoter sequence.

EMSA experiments could be conducted to examine any putative regulatory regions of the topoisomerase IIβ promoter that were identified from the deletion studies. This would enable the identification of putative transcription factors that may have a functional significance.

DNA footprinting assays could also be useful in locating regions of the topoisomerase IIβ promoter that bind proteins. DNA footprinting assays are based on the phenomenon that a

region of DNA that has protein(s) bound will no longer be susceptible to digestion by endonuclease DNase I (refer to figure 6.3). DNase I is used to digest the DNA-protein complex as it cuts DNA molecules randomly, this will result in a mixture of DNA fragments of different lengths but DNA associated with proteins will be protected. Thereby, when the samples are separated on a polyacrylamide gel, the region of the oligonucleotide that binds protein will be identified by a region of missing bands. A sequencing ladder of the same region used in footprinting assays can be used to identify the nucleotide sequence that is binding proteins. The DNA sequence can often be used to identify the cognate transcription factors by making use of such tools as the TRANSFAC database (MatInspector V2.2, http://transfac.gbf.de/cgi-bin/matSearch.pl).

Since the completion of this study a paper was published which identifies two inverted CCAAT boxes and a GC box within the topoisomerase IIβ promoter as being responsible for about 70% of the topoisomerase IIβ promoter activity, which were located between 470-560 bases upstream of the major transcription start site (Lok *et al*, 2002). The CCAAT boxes and GC box are positioned in the opposite order to ICB2, ICB1 and GC1 of the topoisomerase IIα promoter. It may be interesting to carry out a study based in this topoisomerase IIβ sequences similar to the experiments described in this thesis for topoisomerase IIα. A comparison of results from the two promoters may provide insights into the differential regulation observed for topoisomerase IIα and topoisomerase IIβ.

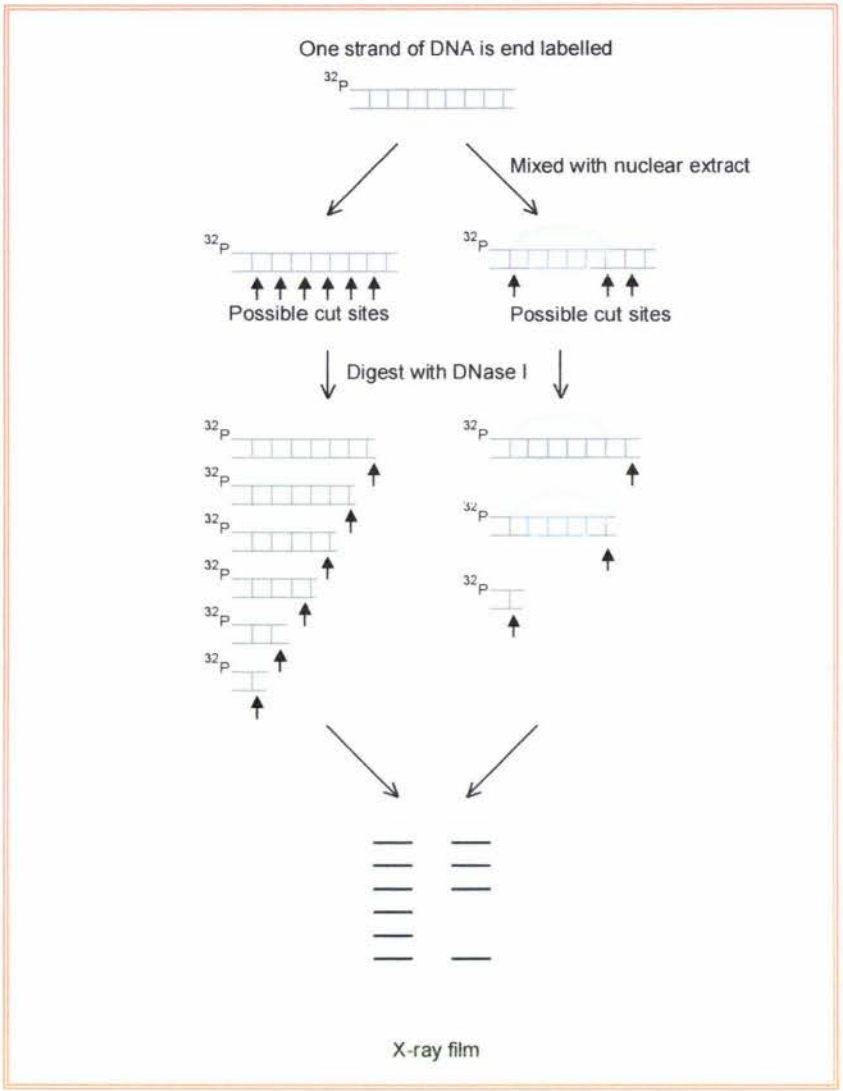


Figure 6.3: Schematic representation of a DNA footprinting assay.

One strand of a double-stranded oligonucleotide is end-labelled with ^{32}P and proteins are allowed to bind to the oligonucleotide, as in EMSA. The DNA-protein complex is then treated with limiting amounts of DNase I, where the result will be a mixture of radioactive fragments of different lengths. The fragments are separated on a polyacrylamide gel and bands are visualized by exposure to X-ray film. The smallest increment is represented by a single nucleotide on the X-ray film, thereby illustrating the region to which the protein binds. A sequencing ladder of the same region used in the footprinting assay will enable the identification of the exact binding site of the protein to the oligonucleotide.

(Figure adapted from: Sambrook and Russell, 2001).

6.2 Topoisomerase IIα.

The regulation of expression of topoisomerase IIα enzyme was investigated through a series of functional assays and EMSAs. Putative binding sites for a variety of transcription factors have been identified in the –617 minimal topoisomerase IIα promoter. ICB1 and GC1 were chosen as the focus of this study because a putative interaction between NF-Y bound at ICB1 and Sp1 bound at GC1 had been implicated by previous work (Szremska, 2001). Preliminary functional assays also suggested that there might be cross-talk by proteins bound distally at GC2 and proximally at GC1 (Szremska, 2001). Therefore, transient transfection experiments were designed to investigate the importance of GC2 as well as GC1 and ICB1 on topoisomerase IIα promoter activity.

6.2.1 Electrophoretic mobility shift assays.

The two most proximal elements of the topoisomerase IIα promoter were studied for protein binding using electrophoretic mobility shift assays, where four proteins were found to bind (chapter four). Three of the four proteins were identified using antibodies against NF-Y, Sp1 and Sp3 (figure 4.5). Significantly, these three transcription factors have been found to be modulated during the development of drug resistance (Kubo *et al*, 1995; Yoon *et al*, 1999, Wang e*t al*, 1997a). Previous work has shown that transcription factor NF-Y can bind to ICB1- 4 elements (Isaacs *et al*, 1996a; Coustry *et al*, 2001; Morgan and Beck, 2001) and transcription factors Sp1 and Sp3 can bind to GC2; however GC1 could associate only with Sp1 of the topoisomerase IIα promoter (Szremska A, 2001).

In this study, using a composite ICB1-GC1 oligonucleotide, transcription factors Sp1 and Sp3 were shown to bind to GC1 and transcription factor NF-Y was shown to bind to ICB1. An uncharacterized protein displayed properties of being a GC1 specific protein (figure 4.6), although it is possible that it contacts DNA between the ICB1 and GC1 elements. The finding that Sp3 can also bind to GC1 was contradictory to the results reported by Szremska (2001). However, Szremska (2001) only used single element oligonucleotides as probes, whereas this study used longer oligonucleotides including both ICB1 and GC1 elements. Therefore, this study provided important evidence for co-operativity between the two proximal elements of the topoisomerase IIα promoter in order to recruit proteins to each element.

It was also shown that NF-Y could be recruited to GC1 in the absence of an intact ICB1 element and that Sp1 and Sp3 were bound at ICB1 in the absence of a complete GC1 element. These patterns are indicative of co-operativity between ICB1 and GC1 through a physical interaction between NF-Y and Sp1/3. Both NF-Y and Sp1/3 have Q-rich domains through which protein-protein interactions are thought to occur. Consequently, NF-Y and Sp1 have been shown to interact *in vivo* using the yeast two-hybrid system (Roder *et al*, 1999) and co-operativity between NF-Y and Sp1/3 has been demonstrated in a number of promoters (Ge *et al*, 2001; Yamada *et al*, 2000). However, this co-operativity had not been established previously within the topoisomerase IIα promoter.

In one of the EMSA experiments, the scenario of Sp1 or Sp3 bound to GC1 did not appear to be able to recruit NF-Y to GC1 (figure 4.4, lanes 7 and 8), suggesting that an intact ICB1 element is required to observe co-operativity between NF-Y and Sp1. NF-Y has been correlated with stimulating histone acetylation, which in turn enables the access of transcription factors to DNA (Adachi *et al*, 2000). Functional analysis has shown that NF-Y could be important in transcription reinitiation as well as activation and it has been implicated to have a role in basal transcription (Mantovani *et al*, 1992). Therefore, it is possible that NF-Y initially binds to ICB1 and enables the recruitment of Sp1/Sp3 to GC1. If this was the case then Sp1/3 should not be able to bind to GC1 in the absence of NF-Y. The results of the EMSA experiments shown in figure 4.6 (lanes 11 and 12, figure 4.7 lanes 3 and 4) do not support this scenario. However, the oligonucleotides used in EMSAs were not organized into nucleosome structures, so transcription factors were not denied access due to structural organization of the DNA.

The distance between ICB1 and GC1 may also be important in maintenance of the interaction observed between NF-Y and Sp1. This could be investigated by carrying out EMSAs with oligonucleotides that have the ICB1 and GC1 elements placed at different distances away from each other and examining the ability of NF-Y to recruit Sp1 to ICB1, and vice-versa. The functional significance of this spacing could also be investigated using reporter gene assays.

Significantly, transcription factors NF-Y, Sp1 and Sp3 have been shown to have a regulatory effect on topoisomerase IIα expression (chapter 5), where NF-Y and Sp1/Sp3 are thought to be able to interact with each other (Roder *et al*, 1999). Down-regulation of topoisomerase IIα

in drug-resistant myeloma cells has also been correlated with a decrease in NF-Y activity (Wang *et al*, 1997a). However, NF-Y cannot be the sole transcription factor involved in down-regulation of topoisomerase IIα, because confluence-arrested cells exhibit complete repression of topoisomerase IIα expression even though NF-Y is present (Isaacs *et al*, 1995). In addition, the topoisomerase IIα promoter is not completely repressed in NF-Y-depleted HeLa cell nuclear extracts *in vitro* (Isaacs *et al*, 1996a). Thus, other transcription factors must also be important in regulating topoisomerase IIα transcription and these may be cell-type specific.

(i) Uncharacterized protein.

Even though the uncharacterized protein displayed properties of being a GC1 specific protein, it required both ICB1 and GC1 elements to bind to GC1 (figure 4.7). This suggests that both NF-Y and Sp1/3 may be required in order for the uncharacterized protein to bind to GC1 or perhaps a sequence between ICB1 and GC1 could be recruiting the uncharacterized protein. The distance between the ICB1 and GC1 elements is 12 bp, which represents a complete turn of the DNA double helix. This is sufficient to independently accommodate a transcription factor. The flanking sequences of CCAAT boxes have been shown to be important for the binding of NF-Y (Mantovani, 1998). Therefore, it is possible that the sequence flanking ICB1 is important, as it may have the ability to recruit an important co-factor such as the uncharacterized protein suggested by the EMSAs.

In order to determine the minimal sequence required for the binding of the uncharacterized protein to the topoisomerase IIα promoter, EMSAs could be conducted using shorter oligonucleotide sequences and sequence specificity examined by sequential mutation of a single base pair within the oligonucleotide. This approach will also indicate whether all of the proteins mentioned are strictly required for the binding of the uncharacterized protein to the topoisomerase IIα promoter DNA. To compare binding of the uncharacterized protein in relation to the presence of these transcription factors, EMSAs could also be conducted using Sp1-enriched HeLa cell extracts and NF-Y depleted cell extracts. This would allow further study of the putative protein-protein interactions involved in the recruitment of the uncharacterized protein. As each EMSA in this study was conducted using HeLa cell extract, it would be of particular interest to examine the cell extracts of other cancer cell lines, for the presence of this uncharacterized protein and to investigate the binding patterns as already

observed for HeLa extracts. Therefore, a wide range of EMSA experiments could be conducted to investigate the binding specificity and properties of the uncharacterized protein.

The fourth uncharacterized protein exhibits properties of being a GC1 associating protein (figure 4.4, lane 4). Due to the nature of the other transcription factors, it is likely that this uncharacterized protein may also have an effect on topoisomerase IIα expression. The functional significance of the uncharacterized protein can only be critically addressed subsequent to its identification and purification.

(ii) Purification of the uncharacterized protein.

Isolation of the uncharacterized protein could be done using the minimal oligonucleotide determined from EMSA experiments. The oligonucleotide could either be labelled with ^{32}P, or biotinylated (refer to figure 6.4), which would aid in either identification or purification of the protein. The oligonucleotide could be mixed with HeLa extract, whereby only the uncharacterized protein should bind and then subjected to separation by carrying out denaturing polyacrylamide gel electrophoresis (SDS-PAGE). The protein could then be isolated from the gel and further analyzed or purified if required. If these protein separation techniques are unsuccessful, then a yeast one-hybrid and/or two-hybrid system could be used to obtain a cDNA clone that represents the uncharacterized protein. It may also be possible to screen an expression library (eg. a λgt11 library) with the appropriate oligonucleotide.

The molecular weight of the uncharacterized protein could be determined by covalently crosslinking the protein to its regulatory sequence using UV light. The oligonucleotide used will not only be ^{32}P-labelled, to enable identification but also have 5-bromo-2'-deoxyuridine (BrdU) incorporated to enable UV-crosslinking. HeLa extract would be mixed with the oligonucleotide, and then UV illuminated at 302 nm to crosslink the protein to its regulatory sequence. However, to ensure that the uncharacterized protein is actually part of an appropriate complex, EMSAs could be done before crosslinking. Following EMSA the gel will be UV illuminated and the region that corresponds to the DNA-protein complex excised from the gel. The protein would then be eluted from the gel and then subjected to separation and molecular weight analysis via SDS-PAGE.

If the preliminary EMSA experiments show that other proteins are required to bind the uncharacterized protein, then those proteins (eg Sp1) could be used to isolate the

uncharacterized protein. A pull-down assay could be attempted using a glutathione-S-transferase (GST)-Sp1 fusion protein (refer to figure 6.5A). Alternatively, a longer oligonucleotide could be used to isolate the protein(s) required for binding of the uncharacterized protein (refer to figure 6.5B). The uncharacterized protein could then be excised from the SDS-PAGE gel according to the estimated size, thereby isolating it from the other proteins.

In order to identify the uncharacterized protein, the amino acid sequence should be determined by N-terminal microsequencing. The sequence can be checked against already existing sequences (such as the NCBI, Entrez protein sequence database) to determine protein identity and speculate protein function. The uncharacterized protein may need to be purified before sequencing can be carried out, this can be done using conventional protein purification techniques such as ion exchange chromatography, size exclusion, affinity chromatography together with EMSA to monitor purification.

As future work, the uncharacterized protein should be identified and examined for the impact it may have on topoisomerase IIα expression. Once the protein has been identified and sequenced, it should be possible to attempt the production of a recombinant protein. The recombinant protein could then be used to explore a range of functional assays aimed at understanding the role of the protein in the modulation of the transcription of topoisomerase IIα. It may be useful to attempt to generate antibodies against the protein so that they could be used to detect for the presence of this protein in different cancer cell lines and to positively identify the protein in EMSA experiments already carried out. The production of a GST-tagged uncharacterized protein could enable the identification of any other proteins it could interact with (see figure 6.5A).

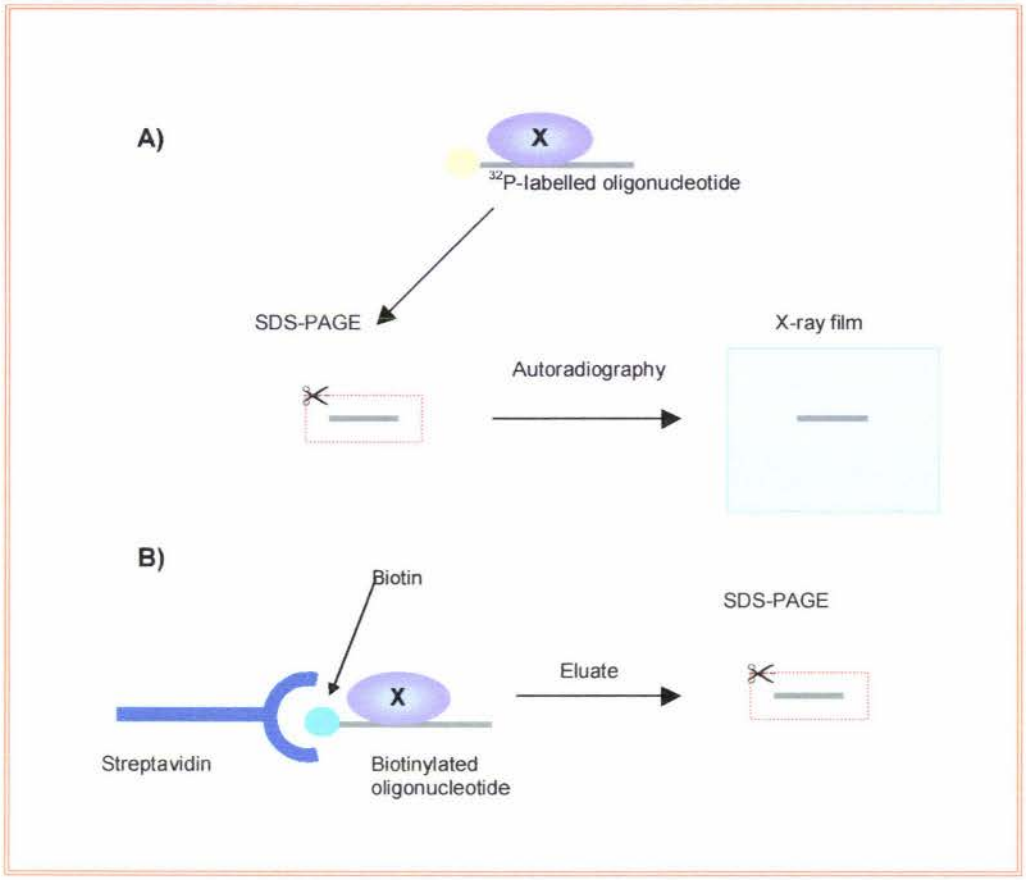


Figure 6.4: Schematic representation of some techniques that could be used to separated the uncharacterized protein from HeLa extract.

Sequence-specific DNA affinity chromatography: The minimal oligonucleotide required to bind the uncharacterized protein would be used, to eliminate chances of binding more than the desired protein from HeLa cell extract. X is the uncharacterized protein.

A) ^{32}P-labelled oligonucleotide used to isolate the uncharacterized protein. The oligonucleotide-protein sample would be run on a SDS-PAGE and the exact location of the protein in the gel could be determined by exposure to X-ray film. The appropriate region can be sliced out of the gel and the protein eluted from the gel slice. **B)** Biotinylated oligonucleotide is used to capture the uncharacterized protein. Biotin will associate with streptavidin, which is immobilized and the uncharacterized protein can be isolated from HeLa extract. Elution of the specifically bound protein can be carried out by using a high salt concentration, and the sample analyzed by SDS-PAGE. The uncharacterized protein could be excised from the gel and the protein could be eluted from the gel slice.

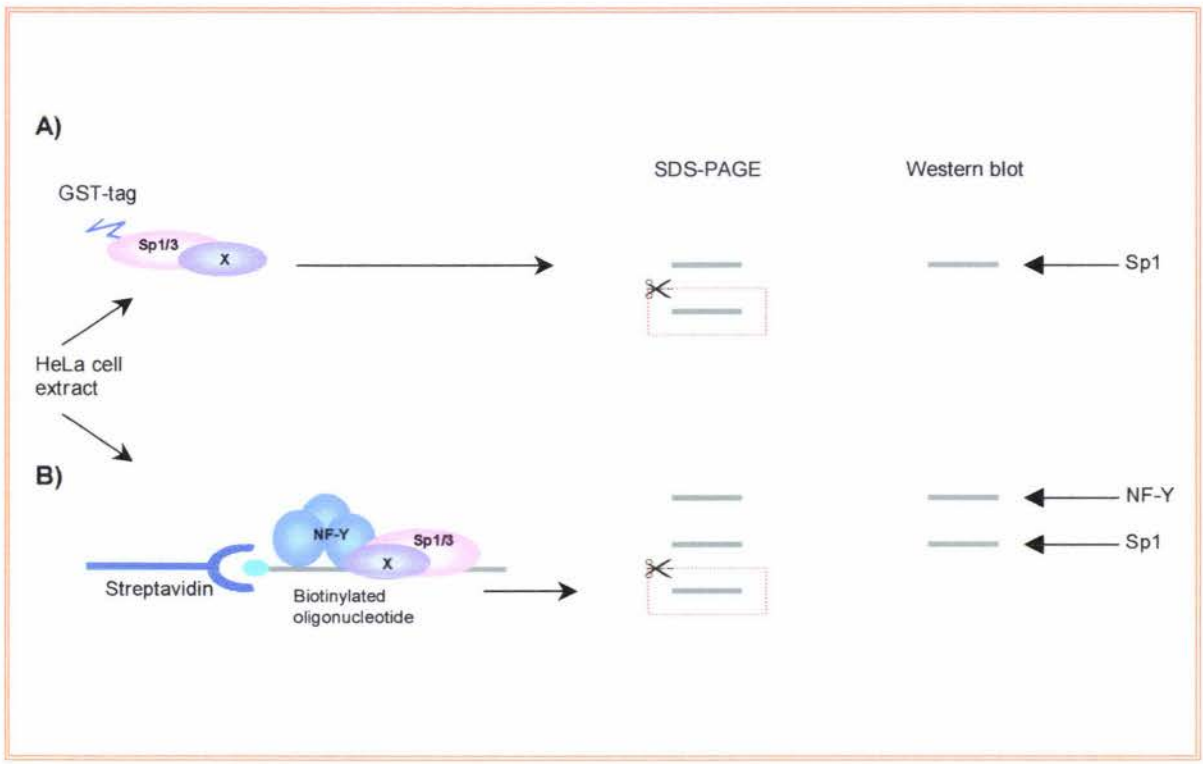


Figure 6.5: Multiple protein complexes to isolate the uncharacterized protein. This figure provides a schematic representation of how multiple protein complexes could be used to isolate the unknown protein. Proteins within the multiple protein complexes can be separated using SDS-PAGE, where sizes can be determined using a molecular marker and proteins can be identified by carrying out western blots with antibodies against NF-Y and Sp1. X is the uncharacterized protein.

A) GST-Sp1 will bind to glutathione-agarose. Separation of the uncharacterized protein from HeLa extract will only occur if there is an interaction between the uncharacterized protein and Sp1. The protein complex can be eluted by adding glutathione, which will compete for binding with the GST-Sp1. **B)** biotinylated oligonucleotide could be used to isolate the proteins required to bind the uncharacterized protein.

6.2.2 Transient transfections.

The –617 minimal topoisomerase IIα promoter was used to examine the functional significance of transcription factors NF-Y and Sp1 on transcriptional regulation. By using constructs that contained mutations in GC1, ICB1 and GC2, information could not only be gained about the importance of these elements, but also the significance each associating transcription factor (identified in EMSAs) had on binding to their respective element.

(i) ICB1 and NF-Y in the transcriptional regulation of topoisomerase IIα.

NF-Y activity is down-regulated in some drug-resistant cells (Wang *et al*, 1997a). It has been suggested that decreased topoisomerase IIα activity is due to a negative regulatory protein (repressor) binding to ICB1, and when this protein can no longer bind to ICB1 topoisomerase IIα activity is no longer down-regulated (Furakawa *et al* 1998; Takano *et al*, 1999). Throughout all of the transient transfections carried out, ICB1mt constructs resulted in approximately a 60% decrease in luciferase activity compared to wild type –617 topoisomerase IIα activity, suggesting that the binding of NF-Y to this element was important in maintaining basal levels of topoisomerase IIα transcriptional activity. Collectively this suggests that the transcriptional regulation of topoisomerase IIα in normal and drug-resistant cells may be modulated by the binding of an activator (NF-Y) and a repressor protein to ICB1. Therefore it could be possible to regulate topoisomerase IIα expression by decreasing the binding of this repressor protein and inducing the binding of NF-Y to ICB1 instead (refer to figure 6.6). A CCAAT box binding protein has been isolated that appears to negatively regulate the binding of NF-Y to CCAAT boxes (Mantovani, 1998). This protein has been aptly named CCAAT displacement protein (CDP). Perhaps future work could involve this CDP protein to see if it can associate with any of the ICBs within the topoisomerase IIα promoter.

The GC1mt/ICB1mt/GC2mt constructs exhibited a 30% decrease in luciferase activity compared to wild type topoisomerase IIα activity. This suggests that the residual activity is probably due to other functional interactions within the –617 region of the topoisomerase IIα promoter, namely ICB2-5. It has been shown that NF-Y can disrupt nucleosomal organization, thereby enabling access of transcription factors to important regulatory elements within promoters (Motta *et al*, 1999). In fact, a study using topoisomerase IIα promoter DNA organized into nucleosomal arrays has shown that multiple ICB elements

generate greater transcriptional activation (Coustry *et al*, 2001). This would explain why topoisomerase IIα promoter activity was not totally inhibited when only ICB1 was mutated in constructs, as other ICB elements can also recruit NF-Y (Isaacs *et al*, 1996a). In addition to this, a small amount of NF-Y was probably recruited to this region even though ICB1 was mutated, through a putative interaction between Sp1/3 and NF-Y (as suggested in EMSA experiments).

Collectively, these results suggest that the ICB1 element is the most important element required for basal transcription, which could be because NF-Y may enable sufficient nucleosome disruption at sites that are close to the transcription initiation site (+1) to allow topoisomerase IIα gene transcription (refer to figure 1.4).

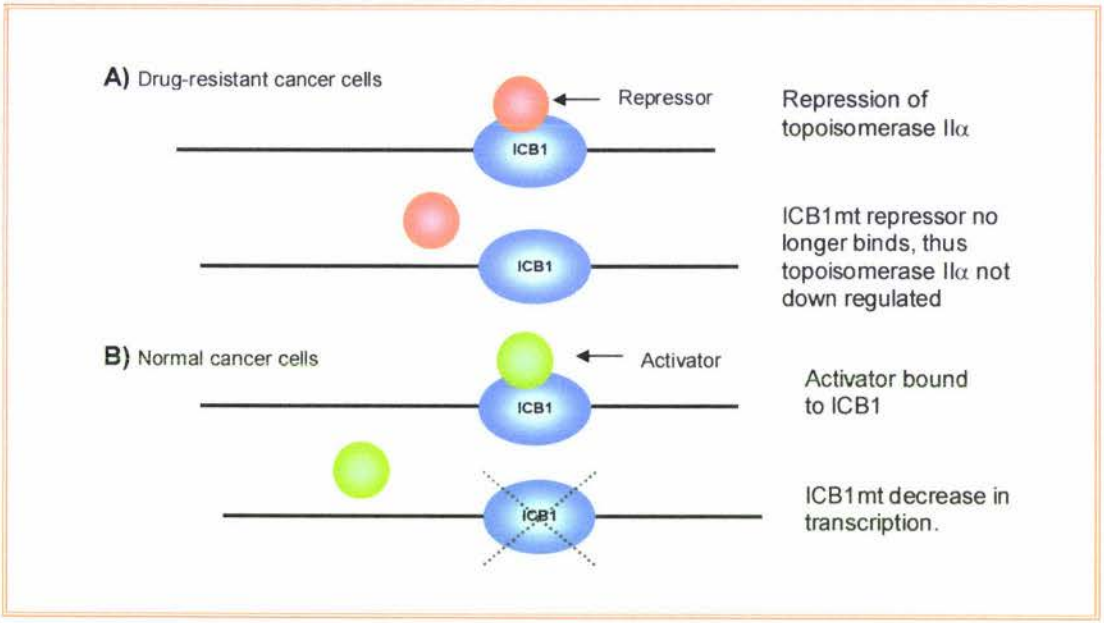


Figure 6.6: Schematic representation of transcriptional regulation of topoisomerase IIα promoter through ICB1.

The transcriptional regulation of topoisomerase IIα could be enabled by reducing the ability of a repressor protein to ICB1 and inducing the binding of NF-Y to ICB1.

A) Drug resistant cells have been shown to exhibit decreased topoisomerase IIα activity. This has been attributed to the binding of a negative regulator (repressor). When the repressor no longer binds to ICB1, drug-resistance is no longer observed. **B)** Normal cancerous cells. When ICB1 is mutated an activator can no longer bind; accordingly, a significant decrease in transcription is observed.

(ii) GC1, GC2 and Sp1 in the transcriptional regulation of topoisomerase IIα.

Sp1 is a commonly recognised transcriptional activator that has been shown to up-regulate transcription in a variety of promoters (Keates *et al*, 2001, Ge *et al*, 2001). In the transient co-transfection experiments the wild type –617 topoisomerase IIα construct exhibited about a 2-fold induction in luciferase activity when co-transfected with to 0.25 μg Sp1 expression vector. Both GC1 and GC2 have been shown to be able to recruit Sp1 and Sp3, and a mutation in both GC1 and GC2 was shown to abolish Sp1-mediated topoisomerase IIα promoter activation.

The GC1mt construct generated luciferase activity relative to that observed from the wild type topoisomerase IIα construct, which is consistent with findings using the rat topoisomerase IIα promoter (Yoon *et al*, 1999). However, when 0.25μg of Sp1 expression vector was added to the transfection with the GC1mt construct, luciferase activity increased only by about half thereby implicating GC2 in this Sp1-mediated increase. A mutation in GC2 resulted in an increase in topoisomerase IIα promoter activity (compared to wild type) possibly because Sp1 and Sp3 can no longer bind (chapter 4). Accordingly, when co-transfected with 0.25 μg of Sp1 expression vector, constructs that had mutations in the GC2 element no longer exhibited a Sp1-mediated increase. Interestingly, the ICB1mt/GC1mt construct generated the greatest increase in luciferase activity (about a 3-fold increase over wt activity) in the presence of 0.25 μg Sp1 expression vector, which suggests that Sp1 bound at GC2 can act as an activator.

Collectively this demonstrates that GC2 can act as a repressor of basal transcription or an activator depending on the status of the proximal ICB1 and GC1 elements, or alternatively the presence or absence of Sp1/3. This is the first study to demonstrate that synergism between proximal and distal GC boxes could exist within the topoisomerase IIα promoter (refer to figure 6.7). Interestingly, it has been observed that Sp1 binds with higher affinity to GC2 than GC1 (Szremska A., 2001). However, this observation has not been tested in the presence of NF-Y bound at ICB1. It has been suggested that when multiple GC boxes exist in a promoter, one of them acts as a repressor and the other as an activator (Li *et al*, 1998). Thus, it is possible that GC2 acts as the repressor element and GC1 as the activator in the topoisomerase IIα promoter.

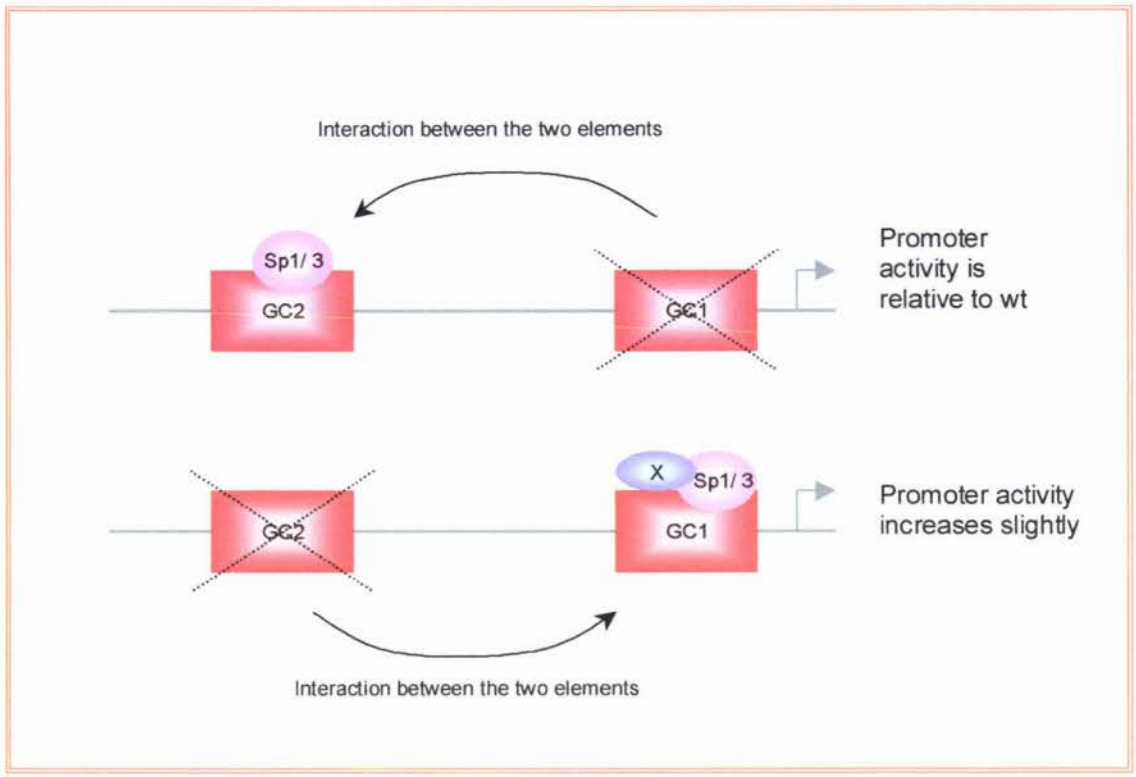


Figure 6.7: Schematic representation of synergism between GC1 and GC2 of the topoisomerase IIα promoter.

The results from the transient transfections suggest that GC1 and GC2 may act co-operatively to regulate the transcription of topoisomerase IIα. GC2 acts as a repressor element of basal topoisomerase IIα transcription, whereas GC1 acts as an activator element. The uncharacterized protein may have an impact on topoisomerase IIα transcriptional activation, by acting at GC1. X is the uncharacterized protein.

It has been observed that synergism between proximal and distal elements of a promoter could be due to interactions between distally and proximally bound Sp1 (Su *et al*, 1991). Sp1 has been shown to form multiple tetramers, which in turn enables DNA to loop thereby bringing proximal and distal elements closer together (Mastrangelo *et al*, 1991; Su *et al*, 1991). Intrinsic DNA bending or curvature is a phenomenon that has been shown to play an important role in a variety of DNA promoters (Sun and Hurley, 1994; Vergeer *et al*, 2000). Since both GC boxes of the topoisomerase IIα promoter recruit Sp1/3, it is possible that Sp1/3 bound at GC1 and Sp1/3 bound at GC2 interact with each other. It has not only been speculated that DNA looping is mediated by Sp1/3 protein-protein interactions but also NF-Y-DNA interactions (Liberati *et al*, 1999). ICB1-4 of the topoisomerase IIα promoter have been shown to bind NF-Y. It is possible that binding of NF-Y to any of these ICBs could

enable DNA bending (see section 1.5.2). The ICB2 element has been shown to have an important regulatory role on topoisomerase IIα expression. Binding of NF-Y to ICB2 has been shown to increase in highly proliferating cells, but diminish in confluence-arrested cells when topoisomerase IIα promoter activity is repressed (Isaacs *et al*, 1996a). Therefore, some sort of regulatory mechanism must exist at ICB2 and it is possible that DNA bending could occur here. DNA bending experiments could be carried out to test this theory, starting with ICB2 then testing the other ICB elements of the topoisomerase IIα promoter.

(iii) GC1, GC2 and Sp3 in the transcriptional regulation of topoisomerase IIα.

Sp3 is different from Sp1 in that it can either activate or repress transcription. One particular study demonstrated that a decrease in topoisomerase IIα gene expression was correlated to Sp3 up-regulation in drug-resistant cell lines (Kubo *et al*, 1995). However, these authors failed to show the position where Sp3 could bind to the topoisomerase IIα promoter. The research described in this thesis established that both Sp1 and Sp3 could bind to either GC1 or GC2 and each has an effect on transcriptional regulation of the topoisomerase IIα gene. Interestingly, others have shown that Sp1 and Sp3 may act in a synergistic manner to regulate the expression of specific genes (Hagen *et al*, 1994), where Sp3 is thought to be able to represse Sp1-mediated transcription (de Luca *et al*, 1996). In any case, the relative levels of Sp1 and Sp3 could be important in the regulation of topoisomerase IIα. As HeLa cells contain endogenous Sp1 and Sp3, it would be interesting to investigate the relative amounts of each transcription factor within these cells. This could be done by carrying out quantitative western blots of HeLa extract (pre and post-confluence) and using antibody to Sp1 and Sp3 to detect the relative levels of each. An appropriate internal control would also need to be used, such as β-tubulin.

The relevance of the binding of these two transcription factors to topoisomerase IIα elements is of particular significance; it was unfortunate that the Sp3 co-transfection experiments could not be interpreted as they could have provided valuable information on the regulation of topoisomerase IIα expression. The problems that occurred using the pSVβ-gal control vector in transient transfections should be corrected by using a different reporter plasmid as an internal control. Preferably, it should be a control plasmid that has a reporter gene driven by a different promoter (not SV40) such as the human cytomegalovirus (CMV) promoter.

One such reporter plasmid is the pCMVSPORT-βgal (Invitrogen), which would still enable transfection efficiency to be monitored through a β-galactosidase assay.

(iv) Problems with transient transfections.

Some of the results from transient transfections generated large variations, which should be taken into consideration. Thus, the activities were normalized and *t*-tests were performed. Errors could have arisen from variables within the experiments during the growth of HeLa cells such as oxygen and CO_2 levels, cell density and integrity, temperature and pH levels. These variables were kept as consistent as possible but inevitably, in a living culture they are difficult to maintain due to non-homogeneous cell growth. The value of the transient transfection experiments could be improved by carrying out more replicates and maintaining consistent cell density. Cell density is an important regulatory factor for topoisomerase IIα expression, as topoisomerase IIα levels are down-regulated in confluence-arrested cells (Isaacs *et al*, 1996a).

HeLa cells contain endogenous Sp1, Sp3 and NF-Y therefore the transient transfections carried out could not distinguish between endogenous and exogenously expressed Sp1. In order to examine topoisomerase IIα activity with relevance to these transcription factors it may be useful to carry them out using *drosophila* Sl2 cells, as they provide a null background for the Sp family of transcription factors and NF-Y. The interactions between NF-Y, Sp1/3 or the uncharacterized protein could be examined in *drosophila* Sl2 cells by sequentially co-transfecting with combinations of each Sp1/3, NF-Y or uncharacterized protein expression vectors, in addition to the –617 topoisomerase IIα pGL3B constructs.

6.3 Examining protein-protein interactions.

It has been suggested that ICB1 and GC1 may act antagonistically to regulate the topoisomerase IIα gene in drug-resistant cells (Takano *et al*, 1999). The information presented in this thesis, suggests that this could be due to an interaction between the transcription factors that bind to these elements. The EMSA experiments revealed that a multiprotein complex might exist comprising of NF-Y, Sp1, Sp3 as well as an uncharacterized protein (figure 4.7, very low mobility band enclosed in a red box). These protein-protein interactions will need to be further investigated.

Due to the binding specificity of the uncharacterized protein, this protein should not be excluded from this complex formation. The results from this study suggest that the uncharacterized protein may mediate an interaction between protein/DNA complexes at the two proximal sites of the topoisomerase IIα promoter and could act as a co-activator or a co-repressor depending on the complement of transcription factors associated with the DNA. Therefore, an important investigation would be to examine the interactions between this uncharacterized protein and other proteins. This could be achieved by generating a GST fusion with the uncharacterized protein, but as this is still not possible, it could be investigated by using GST-Sp1 instead (see figure 6.5A). In addition, using GST-NF-Y may enable the identification of an interaction between the uncharacterized protein and NF-Y. As an alternative, sequence-specific DNA affinity chromatography could also be used, using a biotinylated oligonucleotide that is composed of both ICB1 and GC1 elements of the topoisomerase IIα promoter. Individual proteins within the complex can be separated by SDS-PAGE, and could then be identified by western blotting with antibodies to NF-Y, Sp1 and Sp3 (see figure 6.5B).

As described earlier, eukaryotic transcription is mediated by a series of orchestrated protein-protein and protein-DNA interactions that enable transcription by RNA polymerase II (Pol II). Sp1 has been found to be able to indirectly interact with Pol II through contacts mediated by TBP and TAFs (refer to figure 6.8). It is thought that Sp1 binds to the promoter (GC box) through its carboxyl end but regulates transcriptional activity through its amino terminus (Kadonaga *et al*, 1988). Generally in promoters that have a TATA box, the TATA box enable components of the core promoter to be positioned correctly. However, in TATA-less promoters, such as that of topoisomerase IIα, other sequence motifs must be employed. The location of GC1 in relation to the transcription start site and the knowledge that Sp1 can

interact with specific TAFs indicate that GC1 may play an important role in the basal regulation of the topoisomerase IIα promoter. Thus, the functional significance of a NF-Y-Sp1 interaction could also be important, as NF-Y could recruit Sp1 to GC1 (or ICB1) as a means of locating a position for the transcriptional initiation of topoisomerase IIα promoter. It has also been found that NF-Y is capable of interacting with one of the same TAFs (TAF110) as Sp1, to mediate transcriptional activation (Coustry *et al*, 1998). This observation lends credence to the suggestion the NF-Y and Sp1 have a functional interaction near the transcription start point of the topoisomerase IIα promoter.

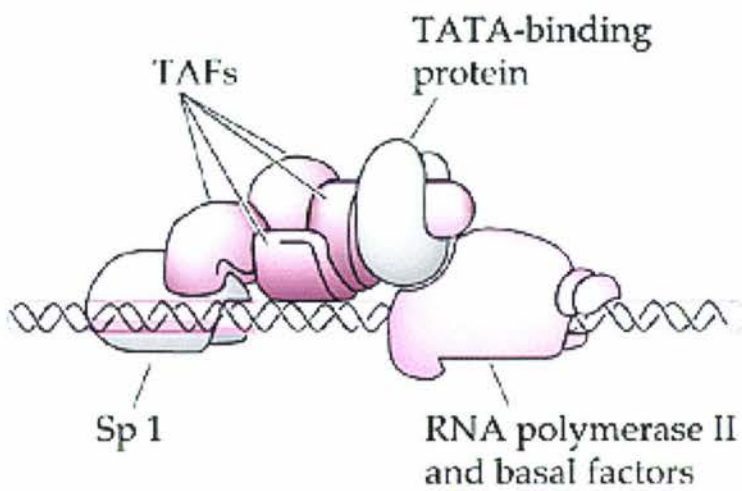


Figure 6.8: Schematic representation of the protein-protein interactions that could implicate Sp1 in mediating transcription .

Organization of the transcription factors that are required to mediate RNA polymerase II regulated transcription in a TATA-less promoter. The Sp1-binding site (GC1 in the case of the topoisomerase IIα promoter) may locate the correct region for RNA polymerase II binding.

(Gilbert S., Developmental Biology 6th edition, 1997).

6.4 Transcriptional regulation of topoisomerase IIα.

Using the information generated from this study, the following models could describe the regulation of topoisomerase IIα gene activity.

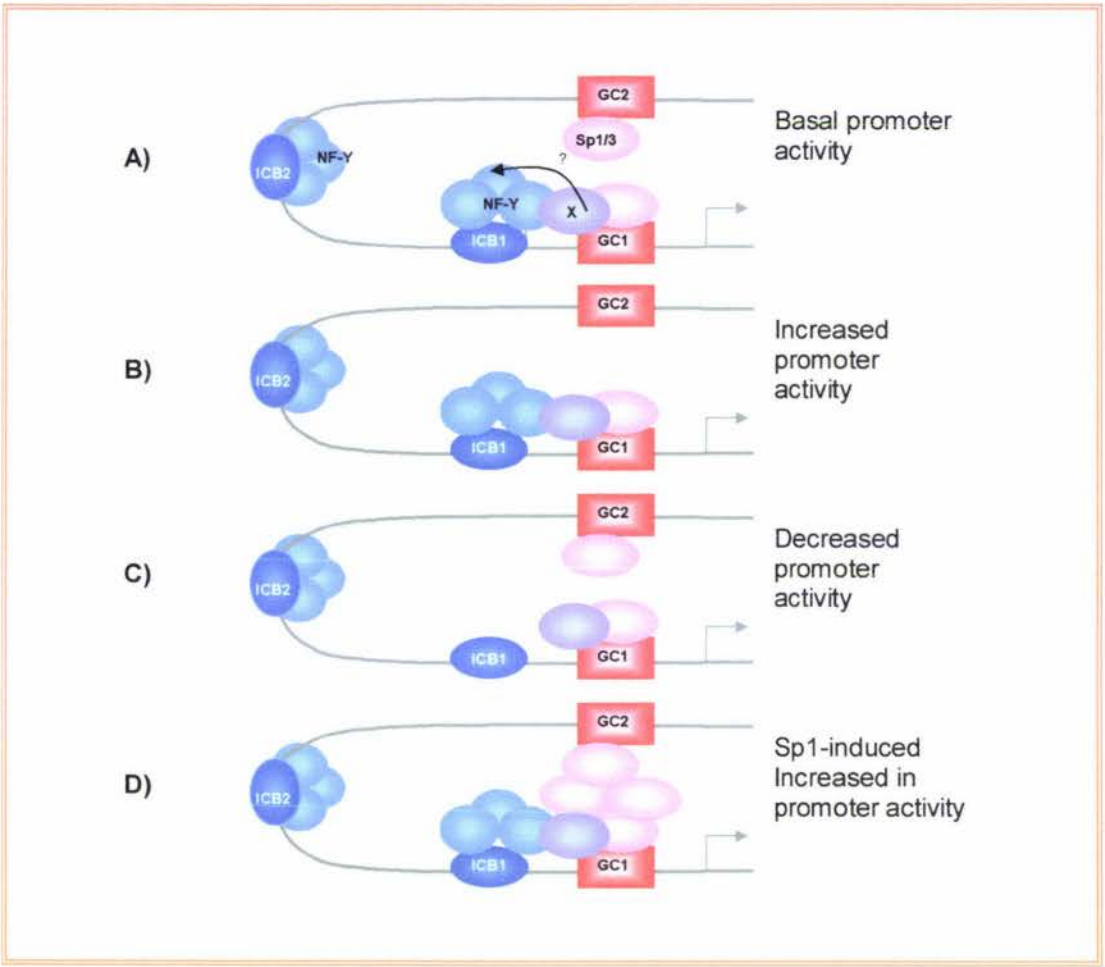


Figure 6.9: Models for the transcriptional regulation of topoisomerase IIα.
Each figure represents putative interactions between transcription factors and the corresponding promoter elements within the topoisomerase IIα promoter. NF-Y is shown to bend DNA at ICB2. X is the uncharacterized protein.

A) An interaction between NF-Y and uncharacterized protein is still undetermined. When Sp1/3 levels are low, it is possible that Sp1/3 preferentially binds to GC1 due to the 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 the most important regulatory element; a mutation in ICB1 causes a decrease in promoter activity. **D)** Increase in Sp1, generally caused an increase in promoter activity; Sp1 mediated interactions between GC1 and GC2.

Figure 6.9 demonstrates that the interactions between these four proteins bound close to the transcription start site and other transcription factors bound at a more distal elements are critical for both basal and activated transcription of topoisomerase IIα. It may even be possible that the uncharacterized protein inhibits the binding of Sp1/3 to GC2 as it interacts with Sp1/3 bound to GC1. However, the consequences of the uncharacterized protein associating with GC1 will not be known until further studies are carried out. This study demonstrates that ICB1 is the most important regulatory element (out of the three elements examined) for topoisomerase IIα promoter activity. The residual activity observed in the GC2mt/ICB1mt/GC1mt construct (about 30%) is most likely to be due to the remaining topoisomerase IIα promoter elements and their cognate proteins. In addition to this, Sp1 has the ability to up-regulate topoisomerase IIα promoter activity *in vivo* and this could be enhanced by an interaction with NF-Y bound at ICB1.

To this end, the transcriptional regulation of the human topoisomerase IIα gene is a complex series of interactions between proximal and distal elements of the promoter. Topoisomerase IIα down-regulation in drug resistant cells may well be a direct consequence of some of the interactions described in this study, therefore further investigations into how these transcription factors interact with the topoisomerase II α and β promoter is required to ascertain the mechanisms involved in the development of drug-resistance. The role that topoisomerase IIβ may have in the development of drug-resistance is not fully understood. By beginning to understand the regulation of the topoisomerase IIβ gene, a better picture of this could also be developed.

References

Adachi N., Miyaike M., Kato K., Kanamaru R., Koyama H. & 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. & Koyama H. (2000): Cell-cycle regulation of the DNA topoisomerase IIα promoter is mediated by proximal CCAAT boxes: possible involvement of acetylation. *Gene* **245**, 49-57.

Asano T., An T., Mayes J., Zwelling L.A. & Kleinerman E.S. (1996): Transfection of human topoisomerase IIα into etoposide-resistant cells: transient increase in sensitivity followed by down-regulation of the endogenous gene. *Biochemistry Journal* **319**, 307-313.

Austin C.A., Marsh K. L., Wasserman R.A., Willmore E., Sayer P.J., Wang J.C. & Fisher L.M. (1995): Expression, domain structure and enzymatic properties of an active recombinant human DNA topoisomerase IIα. *The American Society for Biochemistry and Molecular Biology* **270**, 15739-15746.

Austin C.A. & Marsh K. L. (1998): Eukaryotic DNA topoisomerase II beta. *Bioessays* **20**, 215-226.

Austin C.A., Sng J., Patel S. & Fisher L.M. (1993): Novel HeLa topoisomerase II is the β isoform: complete coding sequence and homology with other type II topoisomerases. *Biochimica et Biophysica Acta* **1172**, 283-291.

Baguley B.C. & Ferguson L.R. (1998): Mutagenic properties of topoisomerase-targeted drugs. *Biochimica et Biophysica Acta* **1400**, 213-222.

Baird R.P., Harkins T.T., Morris S.K. & Lindsley J.E. (1999): Topoisomerase II drives DNA transport by hydrolyzing one ATP. *Proceedings of the National Academy for Science USA* **96**, 13685-13690.

Bakshi R.P., Galande S. & Muniyappa K. (2001): Functional and regulatory characteristics of the eukaryotic type II topoisomerases. *Critical Reviews in Biochemistry and Molecular Biology* **36**, 1-37.

Bellorini M., Lee D.K., Dantonel J.C., Zemzoumi K., Roeder R.G., Tora L. & 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 residues. *Nucleic Acids Research* **25**, 2174-2181.

Berger S.L. & Kimmel A.R. (1987): *Guide to molecular cloning techniques.* Methods in Enzymology, Academic Press.

Bhat U.G., Raychaudhuri P. & Beck W.T. (1999): Functional interaction between human topoisomerase IIα and retinoblastoma protein. *Proceedings of the National Academy for Science USA* **96**, 7859-7864.

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

Birnboim H.C. & Doly J. (1979): A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513.

Brandt T.L, Fraser D.J., Leal S., Halandras P., Kroll A.R. & Kroll D.J. (1997): c-Myb transactivates the human DNA topoisomerase II alpha gene promoter. *Journal of Biological Chemistry* **272**, 6278-6284.

Bredel M. (2001): Anticancer drug resistance in primary human brain tumors. *Brain Research Reviews* **35**, 161-204.

Burden D.A. & Sullivan D.M. (1994): Phosphorylation of the alpha and beta isoforms of DNA topoisomerase II is qualitatively different in interphase and mitosis in Chinese hamster ovary cells. *Biochemistry Journal* **33**, 14651-14655.

Burley S.K. & Roeder R.G. (1997): Biochemistry and structural biology of transcription factor IID (TFIID). *Annual Reviews in Biochemistry* **65,** 769-799.

Capranico G. & Binaschi M. (1998): DNA sequence selectivity of topoisomerase and topoisomerase poisons. *Biochimica et Biophysica Acta* **1400**, 185-194.

Carey M. & Smale S.T. (2000): *Transcriptional regulation in Eukaryotes: concepts, strategies and techniques.* concepts, strategies and techniques, Cold Spring Harbour Laboratory Press, Cold Spring Harbour.

Champoux J.J. (2001): DNA TOPOISOMERASES: Structure, function, and mechanism. *Annual Reviews in Biochemistry* **70**, 369-413.

Chauthaiwale V.M., Therwath A. & Deshpande V.V. (1992): Bacteriophage lambda as a cloning vector. *Microbiology Reviews* **56**(4), 577-591.

Chen A.Y., Attardi L.D., Verrijzer C.P., Yokomori K. & Tijan R. (1994): Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* **79**, 93-105.

Chung T.D.Y. , Drake F.H. , Tan K.B. , Per S.R. , Crooke S.T. & Mirabelli C.K. (1989): Characterization and immunological identification of cDNA clones encoding two human DNA topoisomerase II enzymes. *Proceedings of the National Academy for Science USA* **86**, 9431-9435.

Courey A.J. & Tjian R. (1988): Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine rich activation motif. *Cell* **55**, 887-898.

Coustry F., Hu Q., Crombrugghe B. & Maity S. (2001): CBF/NF-Y functions both in nucleosomal disruption and transcription activation of the chromatin-assembled topoisomerase IIα promoter. *The Journal of Biological Chemistry* **276**, 40621-40630.

Coustry F., Sinha S., Maity S.N. & B., 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.

Cowell I.G., Willmore E., Chalton D., Marsh K. L, Jazrawi E., Fisher L.M. & Austin C.A. (1998): Nuclear distribution of human topoisomerase IIβ: a nuclear targeting signal resides in the 116-residue C-terminal tail. *Experimental Cell Research* **243**, 232-240.

de Luca P., Majello B. & Lania L. (1996): Sp3 represses transcription when tethered to promoter DNA or targeted to promoter proximal RNA. *Journal of Biological Chemistry* **271**, 8533-8536.

Dereudde S., Delaporte C. & Jacquemin-Sablon A. (1997): Role of topoisomerase IIβ in the resistance of 9-OH-elliptine resistant Chinese hamster fibroblasts to topoisomerase inhibitors. *Cancer Research* **57**, 4301-4308.

di Silvio A., Imbriano C. & Mantovani R. (1999): Dissection of the NF-Y transcriptional activation potential. *Nucleic Acids Research* **27**, 2578-2584.

Dingemans A.M., Pinedo H.M. & Giaccone G. (1998): Clinical resistance to topoisomerase-targeted drugs. *Biochimica et Biophysica Acta.* **1400**, 275-288.

Dynan W.S. & Tjian R. (1983a): Isolation of transcription factors that discriminate between different promoters recognised by RNA polymerase II *Cell* **32**, 669-680.

Dynan W.S. & Tjian R. (1983b): The promoter specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**, 79-87.

Falck J., Jensen P.B. & Sehested M. (1999): Evidence for repressional role of an inverted CCAAT box in cell-cycle dependent transcription of the human DNA topoisomerase II alpha gene. *Journal of Biological Chemistry* **274**, 18753-18758.

Feinberg A.P. & Vogelstein B. (1983): A technique for radio-labelling DNA restriction endonuclease fragments to high specific activity. *Annual Reviews in Biochemistry* **132**, 6-13.

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

Gao H., Huang K., Yamasaki E.F., Chan K.K., Chohan L. & Snapka R.M. (1999): XK469, a selective topoisomerase IIβ poison. *Proceedings of the National Academy for Science USA.* **96**(21), 12168-12173.

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

Gilbert S. (1997): *Developmental Biology.* Sixth edition, Sinauer Associates, Inc., Sunderland, MA.

Goswami P.C., Roti Roti J.L. & Hunt C.R. (1996): The cell-cycle coupled expression of topoisomerase II alpha during S phase is regulated by mRNA stability and is disrupted by heat shock or ionizing radiation. *Molecular and Cellular Biology* **16**, 1500-1508.

Gromova I., Biersack H., Jensen P.B., Neilsen O.F., Westergaard O. & Anderson A.H. (1998): Characterisation of DNA topoisomerase II α /β heterodimers in HeLa cells. *Biochemistry Journal* **37**, 16645-16652.

Grue P., Graber A., Sehested M., Jensen P.B., Uhse A., Straub T., Ness W. & Beoge F. (1998): Essential mitotic functions of DNA topoisomerase IIα are not adopted by topoisomerase IIβ in human H69 cells. *The Journal of Biological Chemistry.* **273**, 33660-33666.

Hagen G., Muller S., Beato M. & Suske G. (1994): Sp1-mediated transcriptional activation is repressed by Sp3. *EMBO J* **13**, 3843-3851.

Herzog C.E., Holmes K.A., Tuschong L.M., Ganapathi R. & Zwelling L.A. (1998): Absence of topoisomerase IIβ in an amsacrine-resistant human cell line with mutant topoisomerase IIα. *Cancer Research* **58**, 5298-5300.

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

Hochhauser D., Stanway C.A., Harris A.L. & Hickson I.D. (1992): Cloning and characterization of the 5'-flanking region of the human topoisomerase IIα gene. *The Journal of Biological Chemistry.* **267**, 18961-18965.

Hoffmann A., Oelgeschlager T. & Roeder R.G. (1997): Considerations of transcriptional control mechanisms: Do TFIID-core promoter complexes recapitulate nucleosome-like functions?. *Proceedings of the National Academy for Science USA* **94**, 8928-8935.

Holmes D.S. & Quigley M. (1981): A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* **114**, 193-197.

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

Hopner R., Mousli M., Garnier J., Redon R., Manoir S., Chatton B., Ghyselinck N., Oudet P. & Bronner C. (2001): Genomic structure and chromosomal mapping of the gene coding for ICBP90, a protein involved in the regulation of the topoisomerase IIα gene expression. *Gene* **266**, 15-23.

Hu Q. & Maity S.N. (2000): Stable expression of a dominant negative mutant of CCAAT binding factor/NF-Y in a mouse fibroblast cells resulting in retardation of cell growth and inhibition of transcription of various cellular genes. *The Journal of Biological Chemistry.* **275**, 4435-4444.

Inoue T., Kamiyama J. & 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**(45), 32309-32317.

Isaacs R.J., Davies S.L., Wells & Harris A.L. (1995): Topoisomerase II α and β as targets in breast cancer. *Anti-cancer Drugs* **6**, 195-211.

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

Isaacs R.J. (1996b): Transcriptional regulation of human topoisomerase IIα. Masters thesis, University of Oxford, UK.

Isaacs R.J., Davies S.L., Sandri M.I., Redwood C., Wells N.J. & Hickson I.D. (1998): Physiological regulation of eukaryotic topoismerase II. *Biochimica et Biophysica Acta* **1400**, 121-137.

Kadonaga J.T., Carner K.R., Masiarz F.R. & Tjian R. (1987): Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**, 1079-1090.

Kadonaga J.T., Jones K.A. & Tjian R. (1986): Promoter-specific activation of RNA polymerase II transcription by Sp1. *Trends in Biochemistry* **11**, 20-23.

Keates A.C., Keates S., Kwon J.H., Arseneau K.O., Law D.J., Bai L., Merchant J.L., Wang T.C. & Kelly C.P. (2001): ZBP-89, Sp1 and NF-κB regulate ENA-78 gene expression in Caco-2 human colonic epithelial cells. *Journal of Biological Chemistry* **276**, 43713-43722.

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

Kimura K., Nozaki N., Enomoto T., Tanaka M. & Kikuchi A. (1996): Analysis of M phase-specific phosphorylation of DNA topoisomerase II. *Journal of Biological Chemistry.* **271**, 21439-21445.

Kubo T., Kohno K., Ohga T., Taniguchi K., Kawanami K., Wada M. & Kuwano M. (1995): DNA topoisomerase IIα gene expression under transcriptional control in etoposide/teniposide-resistant human cancer cells. *Cancer Research* **55**, 3860-3864.

Lage H., Helmbach H., Dietel M. & Schadendorf D. (2000): Modulation of DNA topoisomerase II activity and expression in melanoma cells with acquired drug resistance. *British Journal of Cancer* **82**, 488-491.

Laing S.G. & 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. & Cole S.P.C. (1998): Structural organization of the human TOP2α and TOP2β genes. *Gene* **221**, 255-266.

Larsen A.K. & Skladanowski A. (1998): Cellular resistance to topoisomerase-targeted drugs: from drug uptake to cell death. *Biochimica et Biophysica Acta* **1400**, 257-274.

Latchman D.S. (1993): *Transcriptional factors: a practical approach.* The practical approach (Rickwood D. & Hames B.D., Eds.), Oxford University Press.

Li B., Adams C.C. & Workman J.L. (1994): Nucleosome binding by the constitutive transcription factor Sp1. *Journal of Biological Chemistry* **269**, 7756-7763.

Li N., Seetharam S. & Seetharam B. (1998): Characterization of the human transcobalamin II promoter. A proximal GC/GT box is a dominant negative element. *Journal of Biological Chemistry* **273**, 16104-16111.

Liberati C., di Silvio A., Ottolenghi S. & Mantovani R. (1999): NF-Y binding to twin CCAAT boxes: role of Q-rich domains and histone fold helices. *Journal of Molecular Biology* **285**, 1441-1455.

Lok C., Lang A.J., Mirski S.E.L. & Cole S.P.C. (2002): Characterization of the human topoisomerase IIβ *(TOP2B)* promoter activity: Essential roles of the NF-Y and Sp1 binding sites. *Biochemical Journal.* in press manuscript BJ20020791.

Maity S.N. & Crombrugghe B. (1998): Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends in Biochemistry and Science* **23**, 174-178.

Majello B., De Luca P. & Lania L. (1997): Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. *Journal of Biological Chemistry* **272**, 4021-4026.

Malik S. & Roder R.G. (2000): Transcriptional regulation through mediator-like coactivators in yeast and metazoan cells. *Trends in Biochemistry and Science* **294**, 277-283.

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

Mantovani R. (1999): The molecular biology of the CCAAT-binding factor NF-Y. *Gene* **239**, 15-27.

Mantovani R., Pessara U., Tronche F., Li X., Knapp A., Pasquali J., Benoist C. & Mathis D. (1992): Monoclonal antibodies to NF-Y define its function in MHC class II and albumin gene transcription. *The EMBO Journal* **11**(9), 3315-3322.

Marsh K.L., Meczes E.L., Thorn ., Marshall R. & Austin C.A. (1997): Site-directed mutagenesis of human DNA topoisomerase IIβ. *Biochemical Society Transactions* **25**, S638.

Mastrangelo I.A., Courey A.J., Wall J.S., Jackson S.P. & Hough P.V.C. (1991): DNA looping and Sp1 multimeric links: a mechanism for transcriptional synergism and enhancement. *Proceedings of the National Academy for Science USA* **88**, 5670-5674.

M^{c}Lenachan S. (1998): Regulation of the Human topoisomerase IIα promoter. Honours project, Massey University.

Meyer K.N., Kjeldsen E., Straub T., Knudsen B.R., Hickson I.D., Kikuchi A., Kreipe H. & Boege F. (1997): Cell-cycle coupled relocation of types I and II topoisomerases and modulation of catalytic enzyme activities. *The Journal of Cell Biology* **136**(4), 775-788.

Morgan S.E. & Beck W.T. (2001): Role of an inverted CCAAT element in human topoisomerase IIα gene expression in ICRF-187-sensitive and resistant CEM cells. *Molecular Pharmacology* **59**, 203-211.

Motta M.C., Caretti G., Badaracco G.F. & Mantovani R. (1999): Interactions of the CCAAT-trimer NF-Y with nucleosomes. *The Journal of Biological Chemistry.* **274**, 1326-1333.

Ng S. , Liu & Schnipper L. E. (1997): Cloning and characterization of the 5'-flanking sequence for the human DNA topoisomerase II beta gene. *Gene* **203**, 113-119.

Nitiss J.L. (1998): Investigating the biological functions of DNA topoisomerases in eukaryotic cells. *Biochimica et Biophysica Acta* **1400**, 63-81.

Noti J.D. (1997): Sp3 mediates transcriptional activation of the leukocyte integrin genes CDIIC and CDIIB and cooperates with c-Jun to activate CDIIC. *Journal of Biological Chemistry* **272**, 24038-24045.

Orphanides G. & Maxwell A. (1994): In one gate out the other. *Current Biology* **4**, 1006-1010.

Padget K., Pearson A.D.J. & Austin C.A. (2000): Quantitation of DNA topoisomerase II α and β in human leukaemia cells by immunoblotting. *Leukemia* **14**, 1997-2005.

Promega (1999): pGEM-T and pGEM-T Easy Vector Systems. *Technical Manual Number 042.*

Quandt, Karas K., Wingender H., Werner E. & Werner T. (1995): MatInd and MatInspector - New fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Research* **23**, 4878-4884.

Reisman D. & Loging W.T. (1998): Transcriptional regulation of the p53 tumor suppressor gene. *Cancer Biology* **8**, 317-324.

Rhone-Poulenc Rorer Pharmaceuticals. (1999): Oncology Slide Library. *Rhone-Poulenc Rorer Pharmaceuticals.*

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

Roca J., Ishida R., Berger J.M., Andoh T. & Wang J.C. (1994): Anti-tumour bisdioxopiperazines inhibit yeast DNA topoisomerase II by trapping the enzyme in the form of a closed protein clamp. *Proceedings of the National Academy for Science USA* **91**, 1781-1785.

Roder K., Wolf S.S., Larkin K.J. & 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. & Mantovani R. (1995): CCAAT-box binding protein NF-Y (CBP, CP1) recognizes the minor groove and distorts DNA. *Nucleic Acids Research* **23**, 4565-4572.

Sambrook J. & Russell I. (2001): *Molecular Cloning.* 3rd edition. A laboratory manual, Cold Spring Harbor Laboratory Press, New York.

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

Sinha S., Kim S., Sohn B., Crombrugghe B. & Maity N. (1996): Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. *Molecular and Cellular Biology* **16**, 328-337.

Smale S.T. (1997): Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta* **1351**, 73-88.

Sng J., Heaton V.J., Bell M., Maini P., Austin C.A. & Fisher L.M. (1999): Molecular cloning and characterization of the human topoisomerase IIα and IIβ genes: evidence for isoform evolution through gene duplication. *Biochimica et Biophysica Acta* **1444**, 395-406.

Son Y.S., Suh J.M., Ahn S.H., Kim J. C., Yi J.Y., Hur K.C., Hong W.S., Muller M.T. & Chung K. (1998): Reduced activity of topoisomerase II in an Adriamycin-resistant human stomach-adenocarcinoma cell line. *Cancer Chemotherapy and Pharmacology* **41**, 353-360.

Southern E.M. (1975): Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**(503).

Su W., Jackson S., Tijan R. & Echols H. (1991): DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. *Genes and Development* **5**, 820-826.

Sun D. & Hurley L.H. (1994): Co-operative bending of the 21-base-pair repeats of the SV40 viral early promoter by human Sp1. *Biochemistry* **33**(32), 9578-9587.

Sundeth R., MacDonald G., Ting J. & King A.C. (1997): DNA elements recognizing NF-Y and Sp1 regulate the human multidrug-resistance gene promoter. *Molecular Pharmacology* **51**, 963-971.

Suske G. (1999): The Sp-family of transcription factors. *Gene* **238**, 291-300.

Szremska A.P. (2000): Sp1, Sp3, NF-Y and the topoisomerase IIα promoter. Honours Project, Massey University.

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. & Kohno K. (1999): Structural and functional analysis of the control region of the human topoisomerase IIα gene in drug-resistant cells. *Anti-cancer Drug Design* **14**, 87-92.

Tan K.B, Dorman T.E., Falls K.M., Chung T.D.Y., Mirabelli C.K., Crooke S.T. & Mao J. (1992): Topoisomerase IIα and topoisomerase IIβ genes: Characterization and mapping to human chromosomes 17 and 3, respectively. *Cancer Research* **52**, 231-234.

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

Vergeer W.P., Sogo J.M., Pretorius P.J. & W.N., de Vries W.N. (2000): Interaction of Ap1, Ap2 and Sp1 with the regulatory regions of the human pro-alpha 1 (I) collagen gene. *Arch Biochem Biophys* **377**(1), 69-79.

Walker B.P., Smith C., Youndale T., Leblanc J., Whitfeild J.F. & Sikorska M. (1991): Topoisomerase II-reactive chemotherapeutic drugs induce apoptosis in thymocytes. *Cancer Research.* **51**, 1078-1085.

Wang H., Jiang Z., Wong Y.W. & Dalton W.S. (1997a): Decreased CP-1 (NF-Y) activity results in transcriptional down regulation of topoisomerase IIα in a doxorubicin-resistant line of human multiple myeloma RPMI 8226. *Biochimica et Biophysica Reserved Communications* **237**, 217-224.

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

White R.J. & Jackson S.P. (1992): Mechanism of TATA-binding protein recruitment to a TATA-less class III promoter. *Cell* **71**, 1041-1053.

Willmore E., Frank A.J., Padget K., Tilby M.J. & Austin C.A. (1998): Etoposide targets topoisomerase II α and β in leukemic cells: Isoform-specific cleavable complexes visualized and quantified in situ by a novel immunofluorescence technique. *Molecular Pharmacology* **53**, 78-85.

Wingender E., Chen X., Hehl R., Karas H., Liebich I., Matys V., Meinhardt T., Prub M., Reuter I. & Schacherer F. (2000): TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Research* **28**(1), 316-319.

Woessner R.D., Chung T.Y., Hofmann G.A., Mattern M., Mirabelli C.K., Drake F.H. & Johnson R.K. (1990): Differences between normal and *ras*-transformed NIH-3T3 cells in expression of the 170kD and 180kD forms of topoisomerase II. *Cancer Research* **50**, 2901-2908.

Wright K.L., Moore T.L., Vilen B.J. & Brown A. M. (1995): Major histocompatibility complex class II-associated invariant chain gene expression is up-regulated by cooperative interactions of Sp1 and NF-Y. *Journal of Biological Chemistry* **270**(36), 20978-20986.

Xiao H., Lis J.T., Greenblatt J. & Friesen J.D. (1994): The upstream activator CTF/NF1 and RNA polymerase II share a common element involved in transcriptional activation. *Nucleic Acids Research* **22**, 1966-1973.

Yamada K., Tanaka T., Miyamoto K. & Noguchi T. (2000): Sp family members and nuclear factor-Y cooperatively stimulate transcription from the rat pyruvate kinase M gene distal promoter region via their direct interactions. *The Journal of Biological Chemistry* **275**, 18129-18137.

Yang X., Li W., Prescott E.D., Burden S.J. & Wang J.C. (2000): DNA topoisomerase IIβ and neural development. *Science* **287**, 131-134.

Yoon J.H., Kim J.K., Rha G.B., Oh M., Park S., Seong R.H., Hong S.H. & Park S.D. (1999): Sp1 mediates cell proliferation-dependant regulation of rat DNA topoisomerase IIα gene promoter. *Biochemistry Journal* **344**, 367-374.

Yuwen H., Hsia C.C., Nakashima Y., Evangelista A. & Tabor E. (1997): Binding of wild type p53 by topoisomerase II and overexpression of topoisomerase II in human hepatocellular carcinoma. *Biochemical and Biophysical Research Communications* **234**, 194-197.

Zandvliet D.W.J., Hanby A.M., Austin C.A., Marsh K.L., Clark I.B.N., Wright N.A. & Poulsom R. (1996): Analysis of foetal expression sites of human type II DNA topoisomerase α and β mRNA by in situ hybridization. *Biochimica et Biophysica Acta* **1307**, 239-247.

Zwicker J., Gross C., Lucibello F.C., Truss M., Ehlert F., Engeland K. & Muller R. (1995): Cell-cycle regulation of *cdc25C* transcription is mediated by the periodic repression of the glutamine-rich activators, NF-Y and Sp1. *Nucleic Acids Research* **23**, 3822-3831.

Appendix

Appendix 1: Primer and Oligonucleotide sequences.

1.1 Topoisomerase IIβ primer sequences

Primers used in topoisomerase IIβ primer walking: each diluted to 0.8 pmol/μL for sequencing.

T3	ATTAACCCTCACTAAAGGGAA
Tbeta1	TACTACAGCACGAGACGCC
Tbeta2	TAACTGCCATAAGGATGC
T7	TAATACGACTCACTATAGGG

Primers used to produce topoisomerase IIβ probes:

Ng1	GCGCGAATTCCCCTGCTTTCTCCTCAGC(*Eco* RI sites added to each end)
Ng2	GCGCGAATTCAACAATGCAGCCGCCCGT(*Eco* RI sites added to each end)
Ng3	GCCTCGAGTTTGAGGGCA
Ng4	CAGGTCAGTGCCCCGTTG

1.2 Oligonucleotide sequences in EMSA experiments

Sequences of the topoisomerase IIα oligonucleotides used in gel shifts:

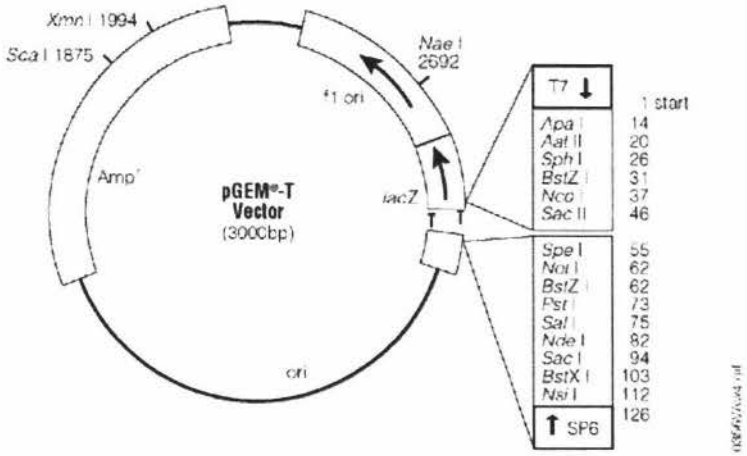
ICB1wt/GC1wt forward	CGAGTCAGGGATTGGCTGGTCTGCTTCGGGCGGGCTAAAGGAAG
ICB1wt/GC1wt reverse	CTTCCTTTAGCCCGCCCGAAGCAGACCAGCCAATCCCTGACTCG
ICB1mt/GC1wt forward	CGAGTCAGGGATTCCCTGGTCTGCTTCGGGCGGGCTAAAGGAAG
ICB1mt/GC1wt reverse	CTTCCTTTAGCCCGCCCGAAGCAGACCAGGGAATCCCTGACTCG
ICB1wt/GC1mt forward	CGAGTCAGGGATTGGCTGGTCTGCTTCGTGCGTGCTAAAGGAAG
ICB1wt/GC1mt reverse	CTTCCTTTAGCACGCACGAAGCAGACCAGCCAATCCCTGACTCG
ICB1mt/GC1mt forward	CGAGTCAGGGATTCCCTGGTCTGCTTCGTGCGTGCTAAAGGAAG
ICB1mt/GC1mt reverse	CTTCCTTTAGCACGCACGAAGCAGACCAGGGAATCCCTGACTCG

Sequences of the single element topoisomerase IIα competitors used in gel shifts:

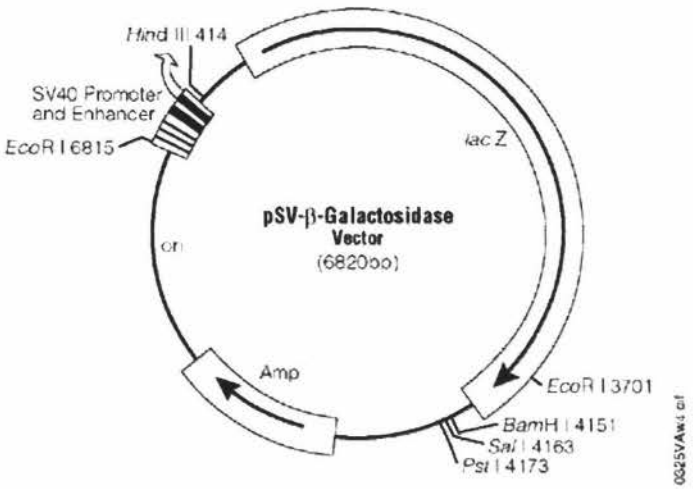
ICB1wt forward	GAGTCAGGGATTGGCTGGTCTGC
ICB1wt reverse	GCAGACCAGCCAATCCCTGACTC
ICB1mt forward	GAGTCAGGGATT**CC**CTGGTCTGC
ICB1mt reverse	GCAGACCAG**GG**AATCCCTGACTC
GC1wt forward	CTGCTTCGGGCGGGCTAAAG
GC1wt reverse	CTTTAGCCCGCCCGAAGCAG
GC1mt forward	CTGCTTCG**T**GCG**T**GCTAAAG
GC1mt reverse	CTTTAGC**A**CGC**A**CGAAGCAG

Appendix 2: Vector maps.

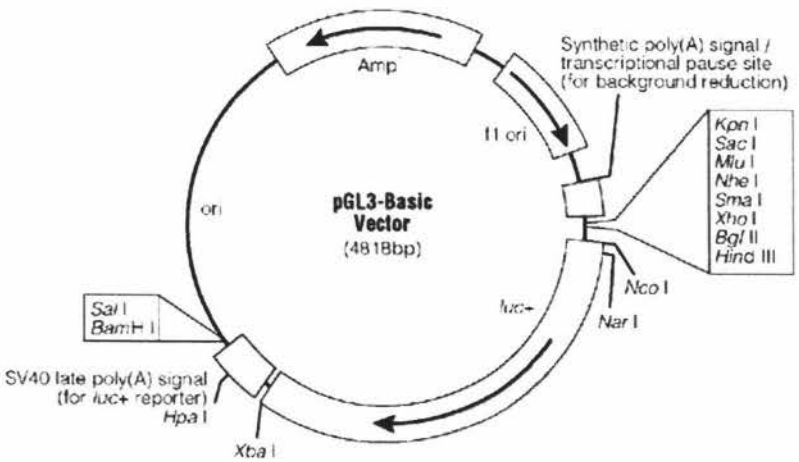
2.1 pGEM-T Vector



2.2 pSV-β-galactosidase Vector

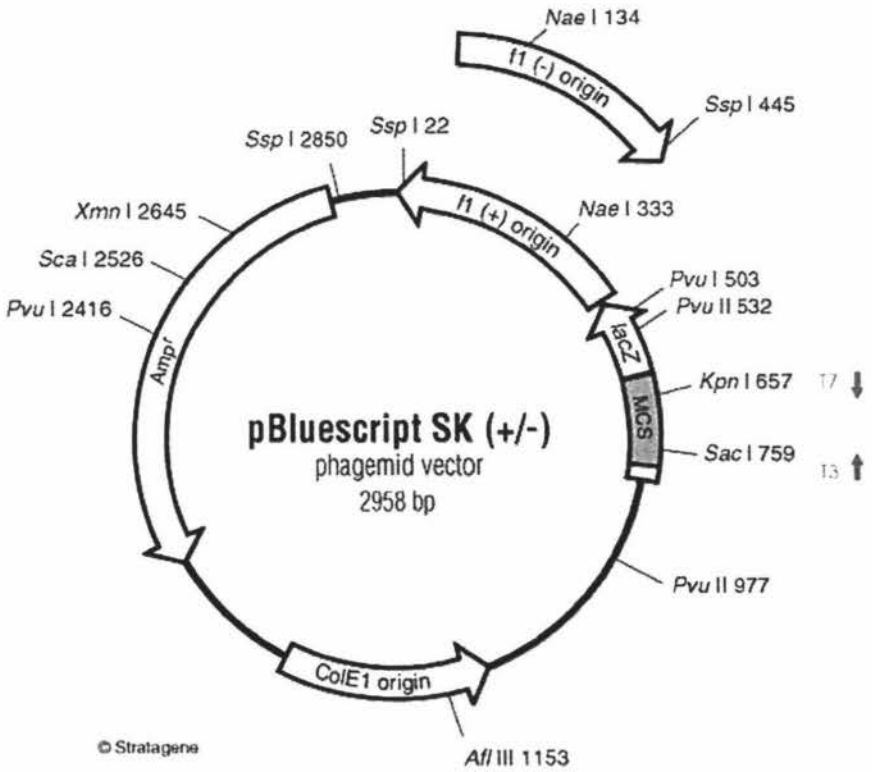


2.3 pGL3B Vector

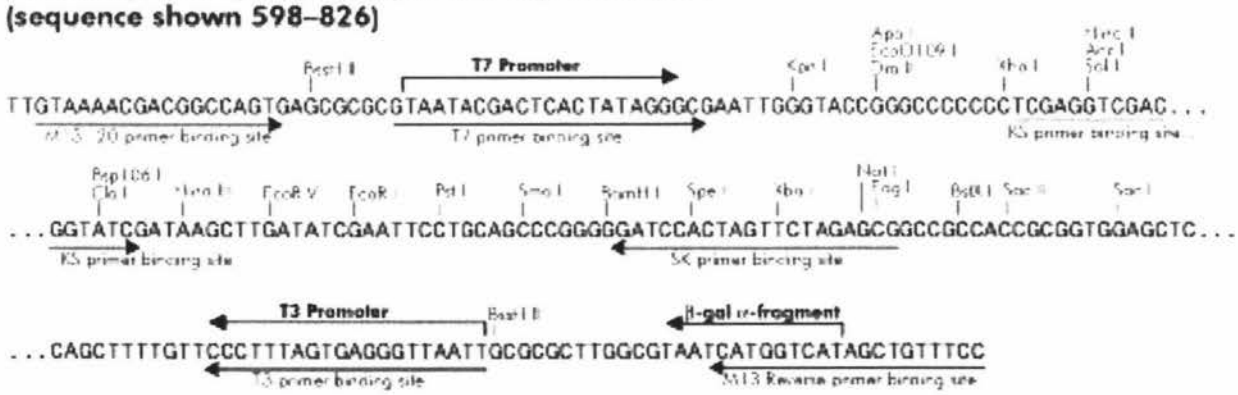


www.promega.com/vectors

2.4 pBluescript SK (+/-) Vector



pBluescript II SK (+/–) Multiple Cloning Site Region (sequence shown 598–826)



Appendix 3:Topoisomerase IIβ sequences and analysis.

3.1 Results of TopoB3.2 restriction digests:

Enzyme(s) used	Approximate sizes of DNA fragments
***Hind* III**	2.5 kb, 3 kb
***Pst* I**	470 bp, 600 bp, 4 kb
***Eco* RI**	1.2 kb, 4 kb
***Bam* HI**	5.5 kb
***Sac* I**	400 bp, 600bp, 4.5 kb
***Xho* I**	5.5 kb
***Hind* III & *Pst* I**	470 bp, 600 bp, 1.4 kb, 3 kb
***Hind* III & *Sac* I**	100 bp, 400 bp, 500 bp, 1.5 kb, 3 kb
***Hind* III & *Eco* RI**	1 kb, 1.5 kb, 3 kb
***Pst* I & *Eco* RI**	150 bp, 550 bp, 600 bp, 4 kb
***Pst* I & *Sac* I**	<100 bp, 100 bp, 400 bp 500 bp, 600 bp, 3.5 kb
Uncut	6 kb

Results of Topo3.2 Restriction Digests.

Approximately 500 ng of TopoB3.2 DNA was digested with 1 μL of various restriction enzymes (1 unit/μL). Each digest was electrophoresed on a 1% agarose gel in 1x TAE at 80V for about 1 hour and the sizes of each band was determined using 1 kb plus ladder.

3.2 Sequencing of topoisomerase IIβ clone TopoB3.2:

Sequence results using primer T3 from clone TopoB3.2
TbetaT3 Length: 581 bases (Tbeta1 primer sequence shown in bold)

```
  1 AAGCTTGAAC CCACCAAATT CCTATCAAGT TCAAAGACTT ACCATTCTTG
 51 CTAAGATAAT GTCTAAAAAA AACAGGTAAC ATTTGAGTGT GGTCAGCCAG
101 ATGGCAAAAC AAGAAACCCC AGGCCTTTCT TTCCCCATGA AGACAAGGAT
151 TCAACAATAA CAAACAATCC AACTGCCTTT GTGATAAGTT CCAGCACCCC
201 AGGTGAGCAT AAAATAGCAT CAAAGCTCAT GTCCTCCCTC CTATCTGCCC
251 CAGCAAGGCA AGATTGAGAA AAACAAACAA GACAACAACA ACTTCCACCT
301 TCTCCCTGGG GAGGAGCAGA AAAGGAATAT ACCTCCAATG GAACAGCTGA
351 GGGCAGTGCT ACCCTAGAAA CTTCTTGACT GAATCTAAGT ACTGACATTG
401 AACAGGGTGC CAGGTTGAGG GCCATTGAAA ACAAAAGCAA AGTGCTTTGG
451 TTTAGCATCA GACCCTGCAG TATCAGAAAC AGACACAAGG NGGAGCTCCA
501 GTACTGTAGT TTACTACAGC ACGAGACGCC ACAAGTAAAG CAGACAAAGA
551 CACACCCAGA NGGAGCAGAC AGATGGACAG A
```

Sequence results using primer Tbeta1 from clone TopoB3.2
Tbeta1 Length: 758 bases (Tbeta2 primer sequence shown in bold)

```
  1 GACACACCAG AGGGAGCAGA CAGATGGACA GAGGCAAACC TCTTCAATTA
 51 AGAAATTACA AACCCAAAGA GGATGCATCC CCATAAAAGA TTTGTGAGGT
101 CTCAGGATTC TCTAGCTAGA CTAACTGGAT GAAGAAAGTC TTCCCAGTAT
151 GAAACCAGAC CGCAAAGACC AGAGAGTTAG CTAAGCCTTC AAATGCCAAA
201 GCCCTAACAG AAAATAACAA TACAAAGAAA CAGGGAAATA CAGCCGATTC
251 AATGTAAAAA ACTAAATTGC AGCCAGGCGC AGTGGCTCAC ACCTGTAATC
301 CCAGCACTTT GGGAGGCTGA GGCAGGCAGA TCCCTTGAGC CCAGAGCTCA
351 AAACCAGTCT GTGCAACATG GCCAAATCTT ATCTCTAGCA AAACAAAACA
401 AAAAAAAATT AGCCAGGTGT GGTAGTGCAC ACTTGTAGTC CCAGCTACTT
451 GGGAGGCTGA GACGGGAGGA TTGCCTGAGC CCAGAAGGTC GAGGCTGCAG
501 TGAGCCATCT GCACACAACT GTACTCCAGC CTGGGTGACA GAGCAAGAGC
551 CTGTCTCAAA AAAATAATAA ATAAATCACC ACAAATTGAG CCTAAAGAAA
601 TGGAAATATA TTGTCATACA AAGAATTCAA ATTAACTGCC ATAAGGATGC
651 TTAATGAAAT AAATGAGTAC AGATGAACAA ACAAAAACTC GGTGTACAAT
701 GCATAAACAA AAGGAAAAAA TCAACAAAAG ATAAAAACTT TTGATGAAGT
751 AGCCAACC
```

Sequence results using primer Tbeta2 from clone TopoB3.2
Tbeta2 Length: 820 bases

```
  1 TGAGTACAGA TGAACAAACN AAAACTCGGT GTAACAATGC ATAAACAAAA
 51 GGAAAAAATC AACAAAGATA AAAACTATGA TGAAGTAGCA AACATTATTT
101 CTGGAACTGA AGCATACAGT AAATGAATTG AAAATACACT AGAGGGATGT
151 AACAGTAGAC TTGAGCAGGC AGAATTTTTT TAAATCAGCA AATTTGAAGA
201 CAGGACTTTC AAAATTGCCA AGGGAAAGGA GCAAAAAGGC AAAGGAATGA
251 ATAAACATAA AAAGAGCTTA AAGAACTTAT GGGCCACTAT CTATCAGAAC
301 AACATATACA TTATGTAAGC CTCAAAGGAG AAGAAAGGGA GAAATGGAGA
351 AAGAGCCTAT TTGAAGAAAT AATAGCTGAA AACTTTCCAA ATCTGAGGAC
401 ATAAATGCAT ATACAAATTA AAGAAGCTCA AAGAACTCCA AGGATAAATC
451 TAAAGAAAAC TACACCAAGA CACATCATAA TTATACTGTC AAAAGTCATA
501 GACAAAGAGA GAATCTTAAA AGCAGCAACA GAAAAGTGAA TCATCAATTA
551 CAAAGGAGCT TCTGTTAGAT TACTAGCCAA TTTCTCAGCA GAAACCTTGC
601 AAGATAGAAG GGATTGGAAG AATATATCTG AAGTACTGAA AGAAAAAAAA
651 AAAAAACACT GCCAACCAAG AACATTGTAT CCAGCAAAAG TGCCCTTCAA
701 AAATGAAGTA GAAATTAATA CTTTCTTAGA TAAACAAACA CTGGGACAGT
751 TCATCACCAC TAGGGTTGCC CACCAGAAAT GATAAAGGGA GTCCCTTAAA
801 NGCAAAGGGT AAAAAATNGC
```

Sequence results using primer T7 from clone TopoB3.2
TbetaT7 Length: 716 bases

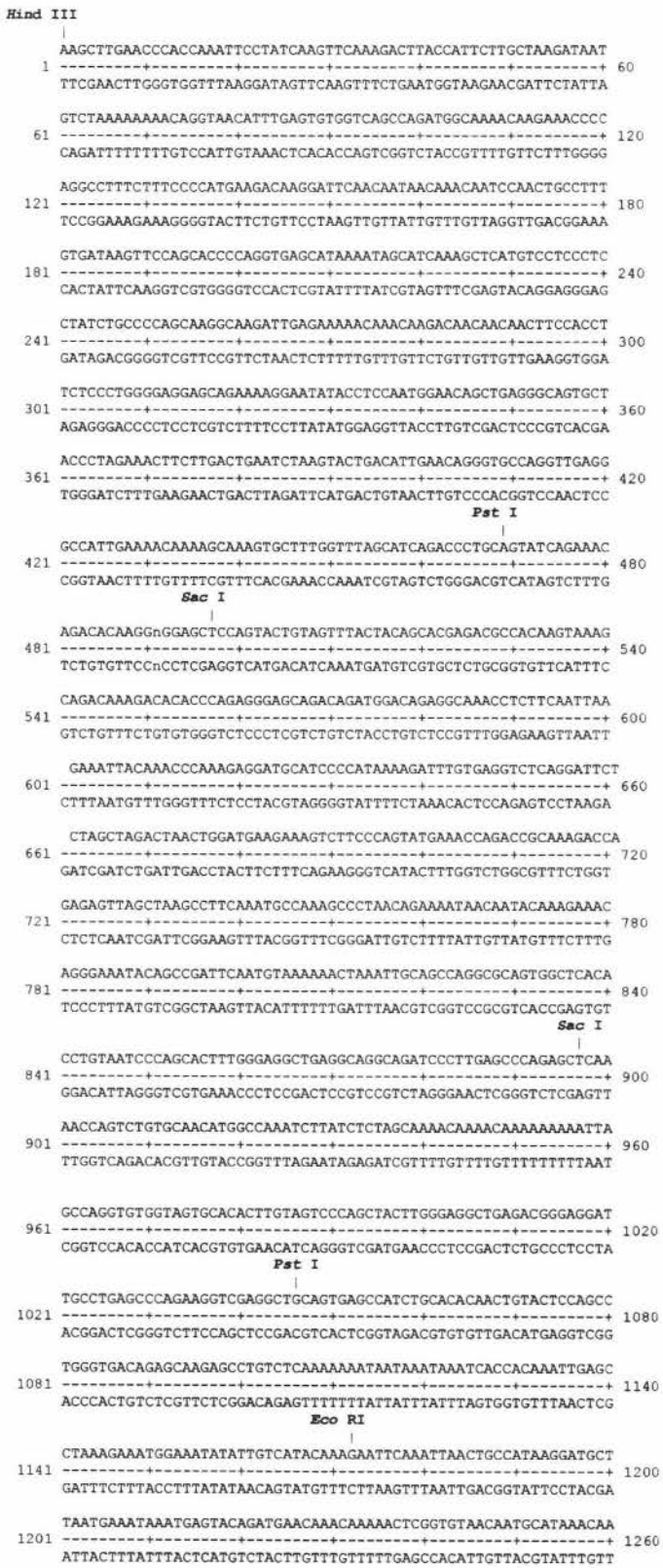
```
  1 CCACGAGTCT TTCTCCATCA TGACACGTAC TCCTGCTCAT TNCTTTCATC
 51 AAACAAGAGC AATTTTATGA GGGNTTTGAT GATAAATAAT TAGGCATCCA
101 TCTTATAATC CTGATTTTGC TCCTTCTGAC TTTTTTTGTT TACTAATCTT
151 AAATCTTTAA AAGGCACCTA ATTTTTTCAG TTGATAATGT AAAAAGACTG
201 CATTGGCATG GTTAAATTCC CAAGACCCTT GGTTCTTCAG AGTTGAGCTA
251 AATGGCTGAT ATCATCACTT ATAAAAGTGT CTTGACCTTT ATGGAGCTTT
301 TGTTGAGAAA TAAAATTTAT ATTTTTTATT TTTATCTTTT GATTCCATTT
351 TTCCAAGAAC TTTTTGGAGT CTCCTTGGAC TTTACCAGGG AGTTATGTAT
401 TTTTGTGTTA CTGTGTTGCT ATTTTTTATC CTTTTGCTTT AAGGGACTCC
451 CTTTATCATT TCTTGTTGGG CAAACCTAGT GGTGATGAAC TGTCCCAGTG
501 TTTGTTTATC TAAGAAAGTA TTAATTTCTA CTTCATTTTT GAAGGGCACT
551 TTTGCTGGAT ACAATGTTCT TGGNTGGCAG TGTTTTTTTT TTTTTTCTTT
601 CAGTACTTCA AGATATATTC TTCCAATCCC TTCTATCTTG CAAGGGTTTC
651 TGCTGAGAAA TTGGCTAGTA ATCTAACAGA AGCCTCCTTT TGTAATTGGA
701 TGATTCACTT TTCTGG
```

Contig of TopoB3.2 sequences TbetaT3, Tbeta1, Tbeta2 and TbetaT7:TopoB3.2corrected sequence: 2452 bases

```
   1 AAGCTTGAAC CCACCAAATT CCTATCAAGT TCAAAGACTT ACCATTCTTG
  51 CTAAGATAAT GTCTAAAAAA AACAGGTAAC ATTTGAGTGT GGTCAGCCAG
 101 ATGGCAAAAC AAGAAACCCC AGGCCTTTCT TTCCCCATGA AGACAAGGAT
 151 TCAACAATAA CAAACAATCC AACTGCCTTT GTGATAAGTT CCAGCACCCC
 201 AGGTGAGCAT AAAATAGCAT CAAAGCTCAT GTCCTCCCTC CTATCTGCCC
 251 CAGCAAGGCA AGATTGAGAA AAACAAACAA GACAACAACA ACTTCCACCT
 301 TCTCCCTGGG GAGGAGCAGA AAAGGAATAT ACCTCCAATG GAACAGCTGA
 351 GGGCAGTGCT ACCCTAGAAA CTTCTTGACT GAATCTAAGT ACTGACATTG
 401 AACAGGGTGC CAGGTTGAGG GCCATTGAAA ACAAAAGCAA AGTGCTTTGG
 451 TTTAGCATCA GACCCTGCAG TATCAGAAAC AGACACAAGG nGGAGCTCCA
 501 GTACTGTAGT TTACTACAGC ACGAGACGCC ACAAGTAAAG CAGACAAAGA
 551 CACACCCAGA GGGAGCAGAC AGATGGACAG AGGCAAACCT CTTCAATTAA
 601 GAAATTACAA ACCCAAAGAG GATGCATCCC CATAAAAGAT TTGTGAGGTC
 651 TCAGGATTCT CTAGCTAGAC TAACTGGATG AAGAAAGTCT TCCCAGTATG
 701 AAACCAGACC GCAAAGACCA GAGAGTTAGC TAAGCCTTCA AATGCCAAAG
 751 CCCTAACAGA AAATAACAAT ACAAAGAAAC AGGGAAATAC AGCCGATTCA
 801 ATGTAAAAAA CTAAATTGCA GCCAGGCGCA GTGGCTCACA CCTGTAATCC
 851 CAGCACTTTG GGAGGCTGAG GCAGGCAGAT CCCTTGAGCC CAGAGCTCAA
 901 AACCAGTCTG TGCAACATGG CCAAATCTTA TCTCTAGCAA AACAAAACAA
 951 AAAAAAATTA GCCAGGTGTG GTAGTGCACA CTTGTAGTCC CAGCTACTTG
1001 GGAGGCTGAG ACGGGAGGAT TGCCTGAGCC CAGAAGGTCG AGGCTGCAGT
1051 GAGCCATCTG CACACAACTG TACTCCAGCC TGGGTGACAG AGCAAGAGCC
1101 TGTCTCAAAA AAATAATAAA TAAATCACCA CAAATTGAGC CTAAAGAAAT
1151 GGAAATATAT TGTCATACAA AGAATTCAAA TTAACTGCCA TAAGGATGCT
1201 TAATGAAATA AATGAGTACA GATGAACAAA CAAAAACTCG GTGTAACAAT
1251 GCATAAACAA AAGGAAAAAA TCAACAAAAG ATAAAAACTA TGATGAAGTA
1301 GCAAACATTA TTTCTGGAAC TGAAGCATAC AGTAAATGAA TTGAAAATAC
1351 ACTAGAGGGA TGTAACAGTA GACTTGAGCA GGCAGAATTT TTTTAAATCA
1401 GCAAATTTGA AGACAGGACT TTCAAAATTG CCAAGGGAAA GGAGCAAAAA
1451 GGCAAAGGAA TGAATAAACA TAAAAAGAGC TTAAAGAACT TATGGGCCAC
1501 TATCTATCAG AACAACATAT ACATTATGTA AGCCTCAAAG GAGAAGAAAG
1551 GGAGAAATGG AGAAAGAGCC TATTTGAAGA AATAATAGCT GAAAACTTTC
1601 CAAATCTGAG GACATAAATG CATATACAAA TTAAAGAAGC TCAAAGAACT
1651 CCAAGGATAA ATCTAAAGAA AACTACACCA AGACACATCA TAATTATACT
1701 GTCAAAAGTC ATAGACAAAG AGAGAATCTT AAAAGCAGCA ACAGAAAAGT
1751 GAATCATCAA TTACAAAAGG AGCTTCTGTT AGATTACTAG CCAATTTCTC
1801 AGCAGAAACC TTGCAAGATA GAAGGGATTG GAAGAATATA TCTGAAGTAC
1851 TGAAAGAAAA AAAAAAAAAA CACTGCCAAC CAAGAACATT GTATCCAGCA
1901 AAAGTGCCCT TCAAAAATGA AGTAGAAATT AATACTTTCT TAGATAAACA
1951 AACACTGGGA CAGTTCATCA CCACTAGGTT TGCCCAACAA GAAATGATAA
2001 AGGGAGTCCC TTAAAGCAAA AGGATAAAAA ATAGCAACAC AGTAACACAA
2051 AAATACATAA CTCCCTGGTA AAGTCCAAGG AGACTCCAAA AAGTTCTTGG
2101 AAAAATGGAA TCAAAAGATA AAAATAAAAA ATATAAATTT TATTTCTCAA
2151 CAAAAGCTCC ATAAAGGTCA AGACACTTTT ATAAGTGATG ATATCAGCCA
2201 TTTAGCTCAA CTCTGAAGAA CCAAGGGTCT TGGGAATTTA ACCATGCCAA
2251 TGCAGTCTTT TTACATTATC AACTGAAAAA ATTAGGTGCC TTTTAAAGAT
2301 TTAAGATTAG TAAACAAAAA AAGTCAGAAG GAGCAAAATC AGGATTATAA
2351 GATGGATGCC TAATTATTTA TCATCAAAnC CCTCATAAAA TTGCTCTTGT
2401 TTGATGAAAG nAATGAGCAG GAGTACGTGT CATGATGGAG AAAGACTCGT
2451 GG
```

3.3 Analysis of topoisomerase IIβ clone TopoB3.2corrected:

Restriction endonuclease MAP (GCG, Wisconsin Package version 9.1) of TopoB3.2corrected from: 1 to: 2452 With enzymes: *Hin*d III, *Sac* I and *Pst* I



```
     AAGGAAAAAATCAACAAAAGATAAAAACTATGATGAAGTAGCAAACATTATTTCTGGAAC
1261 ---------+---------+---------+---------+---------+---------+ 1320
     TTCCTTTTTTAGTTGTTTTCTATTTTTGATACTACTTCATCGTTTGTAATAAAGACCTTG

     TGAAGCATACAGTAAATGAATTGAAAATACACTAGAGGGATGTAACAGTAGACTTGAGCA
1321 ---------+---------+---------+---------+---------+---------+ 1380
     ACTTCGTATGTCATTTACTTAACTTTTATGTGATCTCCCTACATTGTCATCTGAACTCGT

        GGCAGAATTTTTTTAAATCAGCAAATTTGAAGACAGGACTTTCAAAATTGCCAAGGGAAA
1381 ---------+---------+---------+---------+---------+---------+ 1440
     CCGTCTTAAAAAAATTTAGTCGTTTAAACTTCTGTCCTGAAAGTTTTAACGGTTCCCTTT

     GGAGCAAAAAGGCAAAGGAATGAATAAACATAAAAAGAGCTTAAAGAACTTATGGGCCAC
1441 ---------+---------+---------+---------+---------+---------+ 1500
     CCTCGTTTTTCCGTTTCCTTACTTATTTGTATTTTTCTCGAATTTCTTGAATACCCGGTG

     TATCTATCAGAACAACATATACATTATGTAAGCCTCAAAGGAGAAGAAAGGGAGAAATGG
1501 ---------+---------+---------+---------+---------+---------+ 1560
     ATAGATAGTCTTGTTGTATATGTAATACATTCGGAGTTTCCTCTTCTTTCCCTCTTTACC

     AGAAAGAGCCTATTTGAAGAAATAATAGCTGAAAACTTTCCAAATCTGAGGACATAAATG
1561 ---------+---------+---------+---------+---------+---------+ 1620
     TCTTTCTCGGATAAACTTCTTTATTATCGACTTTTGAAAGGTTTAGACTCCTGTATTTAC

     CATATACAAATTAAAGAAGCTCAAAGAACTCCAAGGATAAATCTAAAGAAAACTACACCA
1621 ---------+---------+---------+---------+---------+---------+ 1680
     GTATATGTTTAATTTCTTCGAGTTTCTTGAGGTTCCTATTTAGATTTCTTTTGATGTGGT

     AGACACATCATAATTATACTGTCAAAAGTCATAGACAAAGAGAGAATCTTAAAAGCAGCA
1681 ---------+---------+---------+---------+---------+---------+ 1740
     TCTGTGTAGTATTAATATGACAGTTTTCAGTATCTGTTTCTCTCTTAGAATTTTCGTCGT

     ACAGAAAAGTGAATCATCAATTACAAAAGGAGCTTCTGTTAGATTACTAGCCAATTTCTC
1741 ---------+---------+---------+---------+---------+---------+ 1800
     TGTCTTTTCACTTAGTAGTTAATGTTTTCCTCGAAGACAATCTAATGATCGGTTAAAGAG

     AGCAGAAACCTTGCAAGATAGAAGGGATTGGAAGAATATATCTGAAGTACTGAAAGAAAA
1801 ---------+---------+---------+---------+---------+---------+ 1860
     TCGTCTTTGGAACGTTCTATCTTCCCTAACCTTCTTATATAGACTTCATGACTTTCTTTT

     AAAAAAAAAACACTGCCAACCAAGAACATTGTATCCAGCAAAAGTGCCCTTCAAAAATGA
1861 ---------+---------+---------+---------+---------+---------+ 1920
     TTTTTTTTTTGTGACGGTTGGTTCTTGTAACATAGGTCGTTTTCACGGGAAGTTTTTACT

     AGTAGAAATTAATACTTTCTTAGATAAACAAACACTGGGACAGTTCATCACCACTAGGTT
1921 ---------+---------+---------+---------+---------+---------+ 1980
     TCATCTTTAATTATGAAAGAATCTATTTGTTTGTGACCCTGTCAAGTAGTGGTGATCCAA

     TGCCCAACAAGAAATGATAAAGGGAGTCCCTTAAAGCAAAAGGATAAAAAATAGCAACAC
1981 ---------+---------+---------+---------+---------+---------+ 2040
     ACGGGTTGTTCTTTACTATTTCCCTCAGGGAATTTCGTTTTCCTATTTTTTATCGTTGTG

     AGTAACACAAAAATACATAACTCCCTGGTAAAGTCCAAGGAGACTCCAAAAAGTTCTTGG
2041 ---------+---------+---------+---------+---------+---------+ 2100
     TCATTGTGTTTTTATGTATTGAGGGACCATTTCAGGTTCCTCTGAGGTTTTTCAAGAACC

     AAAAATGGAATCAAAAGATAAAAATAAAAAATATAAATTTTATTTCTCAACAAAAGCTCC
2101 ---------+---------+---------+---------+---------+---------+ 2160
     TTTTTACCTTAGTTTTCTATTTTTATTTTTTATATTTAAAATAAAGAGTTGTTTTCGAGG

     ATAAAGGTCAAGACACTTTTATAAGTGATGATATCAGCCATTTAGCTCAACTCTGAAGAA
2161 ---------+---------+---------+---------+---------+---------+ 2220
     TATTTCCAGTTCTGTGAAAATATTCACTACTATAGTCGGTAAATCGAGTTGAGACTTCTT

     CCAAGGGTCTTGGGAATTTAACCATGCCAATGCAGTCTTTTTACATTATCAACTGAAAAA
2221 ---------+---------+---------+---------+---------+---------+ 2280
     GGTTCCCAGAACCCTTAAATTGGTACGGTTACGTCAGAAAAATGTAATAGTTGACTTTTT

     ATTAGGTGCCTTTTAAAGATTTAAGATTAGTAAACAAAAAAAGTCAGAAGGAGCAAAATC
2281 ---------+---------+---------+---------+---------+---------+ 2340
     TAATCCACGGAAAATTTCTAAATTCTAATCATTTGTTTTTTTCAGTCTTCCTCGTTTTAG

     AGGATTATAAGATGGATGCCTAATTATTTATCATCAAAnCCCTCATAAAATTGCTCTTGT
2341 ---------+---------+---------+---------+---------+---------+ 2400
     TCCTAATATTCTACCTACGGATTAATAAATAGTAGTTTnGGGAGTATTTTAACGAGAACA

     TTGATGAAAGnAATGAGCAGGAGTACGTGTCATGATGGAGAAAGACTCGTGG
2401 ---------+---------+---------+---------+---------+-- 2452
     AACTACTTTCnTTACTCGTCCTCATGCACAGTACTACCTCTTTCTGAGCACC
```

3.4 Sequence of the topoisomerase IIβ lambda insert in pBluescript SK (+):

Sequence of the topoisomerase IIβ λclone1 insert cloned into pBluescript SK(+): pBSTopoB1.26F
Sequenced from the T3 primer end: 500 bases

```
   1  CTTNCTCNGT ACCCGGCGAT CCTCTAGAGT CGACCTGTAG GCATGCNAGC
  51  TTTTTATTCT ATANTNTCAC CTAAATAGCT TGGCGTAATC ATGGTCATNG
 101  GGTGTTTTCT GTGTGAAATT GNTATCCGCT CACAATTCCA CACAACATAC
 151  GAGCCGGGAA GCATAAAGTG TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA
 201  ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG TCGGGAAACC
 251  TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT
 301  TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC
 351  GGTCCGTTCG GCTGCGGCAA GCGGTATCAC NCTCACTCAA AGGCGGTAAT
 401  ACGGTTATNC ACAGAATCAG GGGATAACNC AGGAAAGAAC ATGTGAGCAA
 451  AAGGCCANCA AAAGGCCAGG AACCGTNAAA AGGGNCGCGT TGCTGGCGNT
```

3.5 Search for Potential transcription factor binding sites.

Using MatInspector V2.2

Quandt, K. Frech, K. Karas, H. Wingender, E. and Werner, T.
MatInd and MatInspector - New fast and versatile tools for detection of consensus matches in nucleotide sequence data
Nucleic Acids Research 23, 4878-4884 (1995)

MatInspector Results

MatInspector Release public domain January 2000 Thu Aug 15 15:59:36 2002

Solution parameters:
~~~~~~~~~~~~~~~~~~~~
sequence file: RIDtopoIIB'.seq
core sim: 0.75
matrix sim: 0.85
Explanation for column output:
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

-> Matrix positions correspond to sense strand numbering, but all sequences are given in 5'-3' direction.

-> n/a in column 'core simil.' indicates, that no core search was conducted.

-> Capital letters within the sequence indicate the core string.

Matrix Name	Position(str) of Matrix	Core Simil.	Matrix Simil.	Sequence
Inspecting sequence RIDtopoIIB [?] (1 - 2420):				
V$EVI1_06	6 (+)	1.000	0.900	acaAGATca
V$GATA3_02	7 (+)	0.831	0.852	caaGATCatt
V$GATA3_03	7 (+)	1.000	0.852	caAGATcatt
V$GATA3_03	7 (-)	0.875	0.861	aaTGATcttg
V$TCF11_01	11 (+)	0.807	0.880	ATCAttttaacaa
V$SRY_02	16 (+)	1.000	0.869	tttaACAAgcac
V$CEBPB_01	36 (+)	0.873	0.856	ggctcagGGAAgtt
V$IK1_01	38 (+)	1.000	0.861	ctcaGGGAagtta
V$IK2_01	38 (+)	1.000	0.917	ctcaGGGAagtt
V$HNF3B_01	45 (-)	1.000	0.851	cttgtTATTtaactt
V$E4BP4_01	46 (-)	0.758	0.876	tgttatTTAAct
V$VBP_01	47 (+)	0.785	0.869	gTTAAataac
V$SRY_02	51 (+)	1.000	0.874	aataACAAggct
V$GATA1_03	61 (+)	1.000	0.904	ctatgGATActtgg
V$GATA1_02	61 (+)	1.000	0.922	ctatgGATActtgg
V$CDPCR3HD_01	62 (+)	0.842	0.899	tatgGATAct
V$GATA1_04	62 (+)	1.000	0.862	tatgGATActtgg
V$LMO2COM_02	64 (+)	1.000	0.880	tgGATActt
V$GATA_C	65 (+)	0.875	0.864	gGATACttggt
V$RORA1_01	66 (+)	1.000	0.853	gatacttGGTCag
V$AP1FJ_Q2	69 (-)	1.000	0.930	ccTGACcaagt
V$AP1_Q2	69 (-)	1.000	0.909	ccTGACcaagt
V$AP1_Q4	69 (-)	1.000	0.894	ccTGACcaagt
V$DELTAEF1_01	74 (-)	1.000	0.861	ctgaACCTgac
V$TCF11_01	74 (+)	1.000	0.862	GTCAggttcagaa
V$CEBPB_01	77 (+)	0.986	0.938	aggttcaGAAAagt
V$FREAC2_01	104 (+)	1.000	0.874	tcaaggTAAAgaaagg
V$FREAC3_01	104 (+)	1.000	0.876	tcaagGTAAagaaagg
V$DELTAEF1_01	104 (-)	1.000	0.878	ctttACCTtga
V$XFD1_01	106 (+)	1.000	0.876	aaggTAAAgaaagg
V$GKLF_01	106 (+)	0.937	0.889	aaggtaaagaAAGG
V$CEBPB_01	107 (+)	0.986	0.899	aggtaaaGAAAggc
V$IK2_01	132 (-)	1.000	0.923	actaGGGActtc
V$SRF_Q6	136 (-)	1.000	0.854	tgCCATactaggga
V$CEBPB_01	140 (+)	1.000	0.883	tagtatgGCAAagt
V$NF1_Q6	142 (+)	1.000	0.861	gtaTGGCaaagttacatg
V$OCT1_06	148 (+)	0.833	0.904	caaagttacATGGc
V$RFX1_01	150 (-)	0.945	0.896	tgatggccatGTAActt
V$VBP_01	152 (+)	1.000	0.885	gTTACatggc
V$NF1_Q6	155 (+)	1.000	0.852	acaTGGCcatcagagagc
V$S8_01	167 (+)	1.000	0.971	gagagctaATTAaaaa
V$S8_01	169 (-)	1.000	0.981	ttttttttaATTAgctc
V$NKX25_02	173 (-)	1.000	0.874	ttTAATta
V$HFH3_01	173 (-)	0.838	0.852	gttTTTTtaatta
V$HNF3B_01	173 (-)	0.855	0.856	tagttTTTTtaatta
V$HFH2_01	174 (-)	0.823	0.853	gttTTTTtaatt
V$HNF3B_01	183 (+)	1.000	0.853	aactaTGTTtaaatt
V$XFD2_01	184 (-)	1.000	0.901	aattTAAAcatagt
V$FREAC2_01	184 (-)	1.000	0.880	ataattTAAAcatagt
V$FREAC7_01	184 (-)	1.000	0.850	ataattTAAAcatagt
V$XFD1_01	184 (-)	1.000	0.873	aattTAAAcatagt
V$HFH8_01	185 (+)	1.000	0.874	ctaTGTTtaaatt
V$HFH1_01	185 (+)	1.000	0.877	ctatGTTTaaat
V$HFH3_01	185 (+)	1.000	0.856	ctaTGTTtaaatt
V$S8_01	187 (+)	1.000	0.942	atgtttaaATTAtttc
V$HNF3B_01	192 (+)	1.000	0.885	taaatTATTtccatg
V$OCT1_Q6	195 (-)	1.000	0.911	aaacatggAAATaat
V$OCT1_06	195 (+)	1.000	0.897	attatttccATGTt
V$OCT_C	195 (+)	0.814	0.873	attATTTCcatgt
V$OCT1_05	195 (+)	0.846	0.909	attatttCCATgtt
V$NFAT_Q6	196 (-)	1.000	0.940	acatgGAAAtaa
V$RFX1_01	204 (-)	0.882	0.852	ggagaactctGAAAcat
V$DELTAEF1_01	217 (+)	1.000	0.945	ctccACCTcta
V$ER_Q6	241 (+)	1.000	0.855	gtttaaagatgTGACctga
V$TCF11_01	243 (-)	1.000	0.875	GTCAcatctttaa
V$GATA3_03	245 (+)	1.000	0.934	aaAGATgtga
V$T3R_01	247 (-)	1.000	0.920	gtctcaGGTCacatct
V$DELTAEF1_01	250 (+)	1.000	0.872	tgtgACCTgag
V$AP1FJ_Q2	250 (+)	1.000	0.884	tgTGACctgag
V$RORA1_01	251 (-)	1.000	0.895	tgtctcaGGTCac
V$CREBP1CJUN_01	252 (+)	0.769	0.859	tgACCTga
V$AP1_C	256 (-)	0.841	0.856	aTGTCTCAg
V$HNF3B_01	259 (+)	0.855	0.897	agacaTTTTtgtttt
V$HFH2_01	261 (+)	0.823	0.867	acaTTTTtgttt
V$HFH3_01	261 (+)	0.838	0.915	acaTTTTtgtttt
V$HFH2_01	262 (+)	0.823	0.856	catTTTTgtttt
V$SRY_02	263 (-)	1.000	0.948	gaaaACAAaaat
V$NFAT_Q6	268 (-)	1.000	0.961	caaggGAAAaca
V$IK1_01	268 (-)	1.000	0.893	tcaaGGGAaaaca
V$IK2_01	269 (-)	1.000	0.930	tcaaGGGAaaac
V$NFAT_Q6	279 (+)	1.000	0.945	gagagGAAAtgc
V$CMYB_01	284 (+)	1.000	0.928	gaaatgcacaGTTGgaag
V$LMO2COM_01	289 (+)	0.804	0.889	gcaCAGTtggaa
V$MYOD_Q6	290 (-)	0.872	0.905	tcCAACtgtg
V$VMYB_02	291 (-)	0.820	0.895	tccAACTgt
V$SRY_02	301 (+)	1.000	0.866	gttcACAAaaac
V$HFH3_01	302 (-)	0.838	0.851	gtgTTTTtgtgaa
V$HNF3B_01	304 (-)	1.000	0.853	agaggTGTTtttgtg
V$HFH3_01	304 (-)	1.000	0.894	aggTGTTtttgtg
V$HFH8_01	304 (-)	1.000	0.892	aggTGTTtttgtg
V$DELTAEF1_01	309 (+)	1.000	0.930	aaacACCTctt
V$HFH3_01	319 (-)	1.000	0.877	tacTGTTttttga
V$HNF3B_01	319 (-)	1.000	0.879	attacTGTTttttga
V$HFH8_01	319 (-)	1.000	0.854	tacTGTTttttga

V$HFH2_01	320	(-)	1.000	0.861	tacTGTTttttg
V$VMYB_01	322	(+)	0.820	0.870	aaaAACAgta
V$HNF3B_01	331	(+)	1.000	0.869	aatacTGTTttcatc
V$NFAT_Q6	336	(-)	1.000	0.861	aagatGAAAaca
V$TATA_01	338	(-)	1.000	0.910	ttaTAAAgatgaaaa
V$GATA3_02	339	(-)	1.000	0.893	aaaGATGaaa
V$GATA3_03	339	(-)	1.000	0.924	aaAGATgaaa
V$GATA2_03	339	(-)	0.794	0.871	aaaGATGaaa
V$TCF11_01	340	(+)	0.782	0.863	TTCAtctttataa
V$FREAC7_01	340	(-)	1.000	0.909	ttttttaTAAAgatgaa
V$XFD2_01	340	(-)	1.000	0.902	tttaTAAAgatgaa
V$HFH3_01	341	(+)	0.838	0.853	tcaTCTTtataaa
V$HFH8_01	341	(+)	0.778	0.855	tcaTCTTtataaa
V$FREAC7_01	344	(+)	1.000	0.881	tctttaTAAAagctt
V$TATA_C	346	(+)	1.000	0.926	ttTATAAAAa
V$XFD2_01	346	(+)	1.000	0.868	tttaTAAAagctt
V$HFH8_01	346	(-)	0.778	0.854	agcTTTTtataaa
V$TATA_01	347	(+)	1.000	0.962	ttaTAAAaagctttc
V$BARBIE_01	349	(+)	1.000	0.852	ataaAAAGctttcat
V$CEBPB_01	355	(-)	0.986	0.928	ttgtgatGAAAgct
V$GATA3_02	358	(-)	0.805	0.862	tgtGATGaaa
V$NKX25_01	380	(-)	1.000	0.938	tgAAGTg
V$NKX25_01	392	(+)	1.000	0.900	tcAAGTa
V$TCF11_01	399	(+)	0.798	0.882	CTCAttttggtga
V$CAAT_01	412	(-)	1.000	0.858	gggagCCAAcga
V$NF1_Q6	413	(+)	1.000	0.923	cgtTGGCtccctttagt
V$IK2_01	416	(-)	1.000	0.886	aaaaGGGAgcca
V$GKLF_01	421	(-)	1.000	0.894	acagactaaaAGGG
V$USF_Q6	435	(+)	0.864	0.878	ctCACAagcc
V$CP2_01	436	(+)	1.000	0.879	tcacaagCCAG
V$DELTAEF1_01	442	(-)	1.000	0.854	tataACCTggc
V$TATA_01	448	(+)	1.000	0.930	ttaTAAAtggtttga
V$CAAT_01	451	(-)	0.856	0.887	tcaaaCCATtta
V$GFI1_01	455	(+)	1.000	0.942	tggtttgaAATCaggggccttgtt
V$SRY_02	470	(-)	1.000	0.878	aaaaACAAggcc
V$HFH8_01	472	(+)	1.000	0.876	cctTGTTtttact
V$HNF3B_01	472	(+)	0.855	0.856	ccttgTTTTtacttt
V$GKLF_01	472	(-)	0.937	0.885	aagtaaaaacAAGG
V$XFD1_01	473	(-)	1.000	0.867	aaagTAAAaacaag
V$FREAC2_01	473	(-)	1.000	0.854	tcaaagTAAAaacaag
V$HFH3_01	474	(+)	0.838	0.855	ttgTTTTtacttt
V$AHRARNT_01	503	(+)	1.000	0.853	ctggcacagCGTGgca
V$CAAT_01	524	(+)	1.000	0.905	tgcgcCCAAtaa
V$NFY_01	524	(+)	1.000	0.901	tgcgcCCAAtaaatac
V$NFY_Q6	526	(+)	1.000	0.895	cgcCCAAtaaa
V$FREAC7_01	527	(+)	1.000	0.897	gcccaaTAAAtacttg
V$XFD1_01	529	(+)	1.000	0.870	ccaaTAAAtacttg
V$HNF3B_01	529	(-)	1.000	0.884	tcaagTATTtattgg
V$HFH3_01	529	(-)	0.955	0.886	aagTATTtattgg
V$HFH8_01	529	(-)	0.816	0.873	aagTATTtattgg
V$TATA_01	530	(+)	1.000	0.858	caaTAAAtacttgag
V$NKX25_01	537	(-)	1.000	0.900	tcAAGTa
V$S8_01	541	(+)	1.000	0.938	tgagtggaATTAcact
V$TST1_01	544	(+)	1.000	0.987	gtggAATTacactga
V$VMYB_01	558	(+)	0.820	0.868	aaaAACAgga
V$CETS1P54_01	562	(+)	0.926	0.933	acAGGAagaa
V$NRF2_01	562	(+)	1.000	0.886	acaGGAAgaa
V$YY1_02	572	(+)	1.000	0.872	gtaaagCCATcttccatctg
V$NF1_Q6	580	(-)	1.000	0.921	cttTGGCagatggaagat
V$TAL1ALPHAE47_01	581	(-)	1.000	0.856	tttggCAGAtggaaga
V$TAL1BETAE47_01	581	(-)	1.000	0.857	tttggCAGAtggaaga
V$E47_01	581	(-)	1.000	0.860	ttgGCAGatggaaga
V$LMO2COM_01	583	(-)	0.822	0.895	tggCAGAtggaa
V$MYOD_Q6	584	(+)	0.915	0.978	tcCATCtgcc
V$LYF1_01	588	(-)	0.763	0.856	tttGGCAga
V$OCT1_06	595	(+)	1.000	0.903	aagactgtgATGTt
V$PADS_C	600	(+)	0.904	0.887	tGTGATgtt
V$CEBPB_01	604	(-)	1.000	0.852	caatacaGCAAcat
V$NFY_01	608	(-)	1.000	0.857	aattaCCAAtacagca
V$BRN2_01	611	(+)	1.000	0.880	tgtattggTAATtttc
V$NFY_Q6	611	(-)	1.000	0.889	ttaCCAAtaca
V$S8_01	615	(-)	1.000	0.940	gaatgaaaATTAccaa
V$OCT1_05	618	(+)	0.846	0.899	gtaatttTCATtct
V$OCT1_06	618	(+)	0.889	0.897	gtaattttcATTCt
V$OCT_C	618	(+)	0.847	0.880	gtaATTTTcattc
V$OCT1_Q6	618	(-)	1.000	0.923	aagaatgaAAATtac
V$OCT1_07	620	(-)	1.000	0.861	agaatgaaAATT
V$NFAT_Q6	620	(-)	1.000	0.853	agaatGAAAatt
V$TCF11_01	624	(+)	0.782	0.864	TTCAttcttagaa
V$STAT_01	628	(+)	0.808	0.858	TTCTtagaa
V$OCT1_06	647	(-)	0.889	0.920	aataatttcATTCa
V$S8_01	647	(+)	1.000	0.945	tgaatgaaATTAttcc
V$NFAT_Q6	671	(+)	1.000	0.876	catttGAAAatt
V$CREL_01	676	(-)	0.790	0.852	tggaatTTTC
V$VMYB_01	691	(-)	0.876	0.856	taaAACTgtt
V$TATA_C	695	(-)	0.928	0.872	ttTTTAAAAc
V$TATA_C	695	(+)	0.928	0.876	gtTTTAAAAa
V$S8_01	696	(+)	1.000	0.948	ttttaaaaATTAagca
V$BRN2_01	700	(-)	1.000	0.935	ggcatgctTAATtttt
V$NKX25_02	702	(-)	1.000	0.957	ctTAATtt
V$TAL1BETAE47_01	740	(+)	1.000	0.859	ctggaCAGAtgagaaa
V$TAL1ALPHAE47_01	740	(+)	1.000	0.861	ctggaCAGAtgagaaa
V$TCF11_01	740	(-)	0.798	0.859	CTCAtctgtccag
V$LMO2COM_01	742	(+)	0.822	0.877	ggaCAGAtgaga
V$MYOD_Q6	743	(-)	0.915	0.967	ctCATCtgtc
V$GATA1_06	744	(+)	0.763	0.859	acaGATGaga
V$GATA2_03	744	(+)	0.794	0.895	acaGATGaga
V$GATA3_02	744	(+)	0.805	0.893	acaGATGaga
V$GATA3_03	744	(+)	1.000	0.907	acAGATgaga
V$TCF11_01	749	(-)	1.000	0.901	GTCActtttctca
V$IRF1_01	752	(+)	0.765	0.857	gaaaagtGACAgt
V$IRF2_01	752	(+)	0.750	0.877	gaaaagtGACAgt
V$NKX25_01	753	(+)	1.000	0.884	aaAAGTg
V$TCF11_01	755	(-)	1.000	0.882	GTCActgtcactt
V$AP1_Q4	756	(+)	1.000	0.855	agTGACagtga
V$AP1FJ_Q2	756	(+)	1.000	0.923	agTGACagtga
V$AP1_Q2	756	(+)	1.000	0.910	agTGACagtga
V$OCT1_06	759	(+)	0.833	0.874	gacagtgacATCTc
V$AP1_Q2	762	(+)	1.000	0.888	agTGACatctc
V$AP1FJ_Q2	762	(+)	1.000	0.932	agTGACatctc
V$AP1_Q4	762	(+)	1.000	0.855	agTGACatctc
V$GATA2_03	764	(-)	0.794	0.860	agaGATGtca
V$GATA3_03	764	(-)	1.000	0.911	agAGATgtca
V$SRY_02	773	(-)	1.000	0.926	ctgaACAAtgca
V$SOX5_01	774	(-)	1.000	0.989	tgaaCAATgc
V$STAT_01	780	(-)	0.808	0.871	TTCTctgaa
V$GKLF_01	781	(+)	0.937	0.885	tcagagaagaAAGG
V$EVI1_02	783	(+)	1.000	0.851	agagAAGAaag
V$RORA1_01	786	(+)	1.000	0.938	gaagaaaGGTCaa
V$AP1FJ_Q2	789	(-)	1.000	0.872	gtTGACctttc
V$ER_Q6	790	(-)	1.000	0.858	ctactctttgtTGACcttt
V$XFD3_01	791	(+)	1.000	0.897	aaggtcAACAaaga
V$SRY_02	794	(+)	1.000	0.897	gtcaACAAagag
V$TCFI1_01	794	(+)	1.000	0.874	GTCAacaaagagt
V$HFH3_01	795	(-)	0.838	0.870	tacTCTTtgttga
V$MEF2_02	798	(+)	1.000	0.857	acaaagagtagaAATAgaacac
V$RSRFC4_01	802	(-)	1.000	0.856	gttCTATttctactct
V$PADS_C	812	(-)	0.865	0.887	aGTGTTcta
V$CMYB_01	817	(-)	1.000	0.854	tgcagggtggGTTGagtg
V$GKLF_01	828	(-)	1.000	0.868	gaaagaatgcAGGG

V$HNF3B_01	831	(+)	0.855	0.852	tgcatTCTTtcccta
V$CEBPB_01	835	(-)	0.986	0.867	ttctaggGAAAgaa
V$IK1_01	835	(-)	1.000	0.880	tctaGGGAaagaa
V$NFAT_Q6	835	(-)	1.000	0.920	ctaggGAAAgaa
V$IK2_01	836	(-)	1.000	0.930	tctaGGGAaaga
V$IK2_01	858	(-)	1.000	0.899	ataaGGGAcaca
V$GATA_C	861	(-)	1.000	0.968	tGATAAgggac
V$GATA1_04	862	(-)	1.000	0.968	aaatGATAaggga
V$GATA1_03	862	(-)	1.000	0.957	taaatGATAaggga
V$GATA1_02	862	(-)	1.000	0.980	taaatGATAaggga
V$LMO2COM_02	864	(-)	1.000	0.970	atGATAagg
V$GATA3_02	864	(-)	1.000	0.902	aatGATAagg
V$GATA1_05	864	(-)	1.000	0.919	aatGATAagg
V$GATA1_06	864	(-)	1.000	0.916	aatGATAagg
V$GATA2_02	864	(-)	1.000	0.961	aatGATAagg
V$TCF11_01	868	(+)	0.807	0.881	ATCAtttaagctt
V$BARBIE_01	871	(-)	1.000	0.856	atgaAAAGcttaaat
V$OCT1_06	875	(+)	0.944	0.865	aagcttttcATATg
V$NFAT_Q6	877	(-)	1.000	0.864	catatGAAAagc
V$CEBPB_01	890	(-)	0.986	0.872	gtttcaaGAAAact
V$NFAT_Q6	890	(-)	1.000	0.853	ttcaaGAAAact
V$TCF11_01	895	(-)	1.000	0.865	GTCAgtttcaaga
V$VMYB_01	898	(+)	0.876	0.852	tgaAACTgac
V$AP1_Q2	902	(+)	1.000	0.889	acTGACcatat
V$AP1FJ_Q2	902	(+)	1.000	0.903	acTGACcatat
V$OCT1_02	905	(+)	1.000	0.851	gaccatATGCtcttt
V$MZF1_01	925	(+)	1.000	0.975	tgaGGGGa
V$IK2_01	925	(+)	1.000	0.859	tgagGGGAggag
V$GC_01	926	(+)	0.877	0.871	gaggGGAGgagacg
V$SP1_Q6	926	(+)	0.845	0.861	gaggGGAGgagac
V$NRF2_01	936	(+)	1.000	0.879	gacGGAAgaa
V$GKLF_01	940	(+)	0.817	0.870	gaagaaggagATGG
V$GFI1_01	956	(-)	1.000	0.919	tcgaaaaaAATCatctcagacctt
V$TCF11_01	958	(-)	0.807	0.865	ATCAtctcagacc
V$GATA3_02	962	(+)	0.805	0.861	tgaGATGatt
V$NFAT_Q6	971	(-)	1.000	0.850	ctgtcGAAAaaa
V$CEBPB_01	971	(-)	0.986	0.876	tgctgtcGAAAaaa
V$AP4_Q5	978	(+)	1.000	0.911	gaCAGCagcg
V$CAAT_01	984	(+)	0.827	0.850	agcgtCCACtta
V$NKX25_01	990	(-)	1.000	0.917	ctAAGTg
V$IK2_01	1000	(+)	1.000	0.903	cttgGGGAggcc
V$MZF1_01	1000	(+)	1.000	0.965	cttGGGGa
V$LYF1_01	1001	(+)	1.000	0.877	ttgGGGAgg
V$AP1_Q2	1009	(-)	1.000	0.862	cgTGACccggc
V$AP1FJ_Q2	1009	(-)	1.000	0.908	cgTGACccggc
V$ER_Q6	1010	(-)	1.000	0.898	aggaggccccgTGACccgg
V$NMYC_01	1013	(-)	1.000	0.886	ggcccCGTGacc
V$USF_Q6	1014	(+)	1.000	0.977	gtCACGgggc
V$TCF11_01	1014	(+)	1.000	0.874	GTCAcggggcctc
V$USF_C	1015	(-)	0.848	0.895	cCCCGTga
V$CEBPB_01	1025	(-)	1.000	0.903	cactggtGCAAgga
V$PADS_C	1035	(+)	1.000	0.933	aGTGGTcta
V$CMYB_01	1061	(-)	1.000	0.915	taaaagagcgGTTGtaaa
V$VMYB_02	1063	(+)	0.793	0.872	tacAACCgc
V$TATA_C	1073	(+)	0.928	0.873	ctTTTAAAAt
V$TATA_C	1073	(-)	0.928	0.883	atTTTAAAAg
V$S8_01	1078	(-)	1.000	0.942	aaagggaaATTAtttt
V$IK1_01	1082	(-)	1.000	0.856	caaaGGGAaatta
V$NFAT_Q6	1082	(-)	1.000	0.932	aaaggGAAAtta
V$IK2_01	1083	(-)	1.000	0.916	caaaGGGAaatt
V$IRF1_01	1083	(-)	1.000	0.859	ccaaaggGAAAtt
V$IK2_01	1090	(+)	1.000	0.961	ctttGGGAcggc
V$LYF1_01	1091	(+)	1.000	0.924	tttGGGAcg
V$NFKAPPAB_01	1094	(-)	0.815	0.871	GGGCcgtccc
V$CETS1P54_01	1109	(-)	0.926	0.922	gcAGGAagaa
V$IK2_01	1123	(-)	1.000	0.902	ggtgGGGAtggg
V$GC_01	1125	(-)	0.872	0.894	gcggGGTGgggatg
V$SP1_Q6	1126	(-)	0.819	0.888	gcggGGTGgggat
V$MZF1_01	1127	(-)	1.000	0.996	ggtGGGGa
V$RREB1_01	1128	(+)	1.000	0.862	cCCCAccccgcccc
V$GC_01	1130	(-)	1.000	0.959	tgggGGCGgggtgg
V$SP1_Q6	1131	(-)	1.000	0.980	tgggGGCGgggtg
V$GKLF_01	1139	(-)	0.951	0.861	ggatgaggatGGGG
V$TCF11_01	1146	(+)	0.798	0.873	CTCAtccttagag
V$IK3_01	1176	(-)	1.000	0.854	aaccgGGAAtccg
V$IK1_01	1176	(-)	1.000	0.912	aaccGGGAatccg
V$IK2_01	1177	(-)	1.000	0.953	aaccGGGAatcc
V$VMYB_02	1183	(-)	0.793	0.872	cctAACCgg
V$FREAC4_01	1194	(+)	1.000	0.859	tcagagtaAACAgcac
V$FREAC2_01	1194	(+)	1.000	0.948	tcagagTAAAcagcac
V$XFD1_01	1196	(+)	1.000	0.854	agagTAAAcagcac
V$XFD2_01	1196	(+)	1.000	0.853	agagTAAAcagcac
V$HNF3B_01	1196	(-)	1.000	0.888	cgtgcTGTTtactct
V$HFH8_01	1196	(-)	1.000	0.854	tgcTGTTtactct
V$XFD3_01	1196	(+)	1.000	0.881	agagtaAACAgcac
V$PADS_C	1202	(-)	0.865	0.865	cGTGCTgtt
V$AP4_Q5	1202	(+)	1.000	0.865	aaCAGCacgc
V$AHRARNT_01	1204	(-)	1.000	0.894	tcctccaagCGTGctg
V$SRY_02	1218	(+)	1.000	0.902	gattACAAtcga
V$SOX5_01	1219	(+)	1.000	0.859	attaCAATcg
V$NF1_Q6	1225	(-)	1.000	0.934	aatTGGCctcgggtcgat
V$NFY_01	1232	(+)	1.000	0.902	cgaggCCAAttcgacc
V$CAAT_01	1232	(+)	1.000	0.864	cgaggCCAAttc
V$NFY_Q6	1234	(+)	1.000	0.883	aggCCAAttcg
V$IK1_01	1246	(-)	1.000	0.857	caaaGGGAaaggg
V$GKLF_01	1246	(-)	1.000	0.917	gcaaagggaaAGGG
V$NFAT_Q6	1246	(-)	1.000	0.917	aaaggGAAAggg
V$IRF2_01	1247	(-)	1.000	0.926	gcaaaggGAAAgg
V$GKLF_01	1247	(-)	0.937	0.888	ggcaaagggaAAGG
V$IRF1_01	1247	(-)	1.000	0.877	gcaaaggGAAAgg
V$IK2_01	1247	(-)	1.000	0.919	caaaGGGAaagg
V$E2F_02	1255	(+)	1.000	0.928	TTTGccgc
V$TCF11_01	1261	(-)	1.000	0.864	GTCAattgagggc
V$AP1_Q2	1268	(+)	1.000	0.850	atTGACccttg
V$AP1FJ_Q2	1268	(+)	1.000	0.865	atTGACccttg
V$COUP_01	1270	(+)	1.000	0.871	tGACCcttgaagca
V$AP4_Q5	1280	(+)	1.000	0.851	agCAGCcccc
V$GKLF_01	1285	(-)	0.951	0.865	agggagagcaGGGG
V$GKLF_01	1286	(-)	1.000	0.884	aagggagagcAGGG
V$IK2_01	1290	(-)	1.000	0.895	tgaaGGGAgagc
V$NFAT_Q6	1301	(+)	1.000	0.977	aaatgGAAAacc
V$NFKAPPAB65_01	1305	(-)	1.000	0.946	tgggttTTCC
V$CREL_01	1305	(-)	1.000	0.978	tgggttTTCC
V$FREAC7_01	1321	(+)	0.787	0.863	cacacaCAAAcaaaaa
V$HFH8_01	1323	(-)	1.000	0.917	tttTGTTtgtgtg
V$HFH3_01	1323	(-)	1.000	0.942	tttTGTTtgtgtg
V$HNF3B_01	1323	(-)	1.000	0.871	gttttTGTTtgtgtg
V$HFH2_01	1324	(-)	1.000	0.883	tttTGTTtgtgt
V$SRY_02	1326	(+)	1.000	0.927	acaaACAAaaac
V$HFH3_01	1327	(-)	0.838	0.925	gggTTTTtgtttg
V$HNF3B_01	1327	(-)	0.855	0.869	tggggTTTTtgtttg
V$NRF2_01	1343	(-)	1.000	0.860	aaaGGAAgac
V$CMYB_01	1345	(+)	1.000	0.895	cttcctttcgGTTGctac
V$VMYB_01	1351	(-)	0.809	0.879	agcAACCgaa
V$VMYB_02	1352	(-)	0.793	0.869	agcAACCga
V$TCF11_01	1360	(-)	1.000	0.969	GTCAttgcgggta
V$CREB_02	1365	(+)	1.000	0.903	gcaaTGACgttt
V$ATF_01	1366	(+)	1.000	0.967	caaTGACgtttccc
V$CREB_Q4	1367	(+)	1.000	0.886	aaTGACgtttcc
V$AP1FJ_Q2	1367	(+)	1.000	0.882	aaTGACgtttc

V$AP1_Q2	1367	(+)	1.000	0.858	aaTGACgtttc
V$CREBP1_Q2	1367	(+)	1.000	0.868	aaTGACgtttcc
V$CREB_Q2	1367	(+)	1.000	0.863	aaTGACgtttcc
V$CREB_01	1369	(+)	1.000	0.934	TGACgttt
V$CREBP1CJUN_01	1369	(+)	1.000	0.889	tgACGTtt
V$NFAT_Q6	1371	(-)	1.000	0.910	ggggggGAAAcgt
V$IK2_01	1372	(-)	1.000	0.901	agggGGGAaacg
V$MZF1_01	1376	(-)	1.000	0.991	aggGGGGa
V$IK2_01	1384	(-)	1.000	0.899	gggcGGGAcccg
V$GC_01	1385	(-)	1.000	0.921	gaggGGCGggaccc
V$SP1_Q6	1386	(-)	1.000	0.950	gaggGGCGggacc
V$AP2_Q6	1396	(-)	0.976	0.915	agCCCCctggag
V$STAF_02	1403	(-)	1.000	0.895	caatCCCAaaattccaagccc
V$GFI1_01	1407	(-)	1.000	0.859	tctcggccAATCccaaaattccaa
V$IK2_01	1413	(+)	1.000	0.958	ttttGGGAttgg
V$NFY_01	1414	(-)	1.000	0.926	ctcggCCAAtcccaaa
V$LYF1_01	1414	(+)	1.000	0.877	tttGGGAtt
V$NFY_Q6	1417	(-)	1.000	0.911	cggCCAAtccc
V$CAAT_01	1418	(-)	1.000	0.906	ctcggCCAAtcc
V$NF1_Q6	1419	(+)	1.000	0.943	gatTGGCcgagaggctgt
V$NF1_Q6	1433	(+)	1.000	0.857	ctgTGGCgacaaggcccg
V$GFI1_01	1439	(-)	1.000	0.889	tgctgtccAATCcgggccttgtcg
V$NFY_01	1446	(-)	1.000	0.921	gctgtCCAAtccgggc
V$CETS1P54_01	1447	(+)	1.000	0.853	ccCGGAttgg
V$NFY_Q6	1449	(-)	1.000	0.908	tgtCCAAtccg
V$CAAT_01	1450	(-)	1.000	0.872	gctgtCCAAtcc
V$AP4_Q5	1463	(-)	1.000	0.863	gtCAGCgcca
V$TCF11_01	1464	(-)	1.000	0.873	GTCAgtcagcgcc
V$NFE2_01	1466	(+)	1.000	0.856	cgCTGActgac
V$AP1FJ_Q2	1467	(+)	1.000	0.957	gcTGACtgaca
V$AP1_Q2	1467	(+)	1.000	0.966	gcTGACtgaca
V$AP1_Q4	1467	(+)	1.000	0.953	gcTGACtgaca
V$AP1_C	1468	(-)	0.868	0.892	gTCAGTCAg
V$AP1_Q2	1471	(+)	1.000	0.897	acTGACagcgg
V$AP1FJ_Q2	1471	(+)	1.000	0.911	acTGACagcgg
V$AP4_Q6	1474	(+)	1.000	0.895	gaCAGCgggg
V$AP4_Q5	1474	(+)	1.000	0.922	gaCAGCgggg
V$NGFIC_01	1476	(+)	0.788	0.857	caGCGGgggcgg
V$SP1_Q6	1479	(+)	1.000	0.884	cgggGGCGgccgc
V$GC_01	1488	(-)	1.000	0.856	ggagGGCGcggcgg
V$SP1_Q6	1493	(-)	0.845	0.859	agagGGAGggcgc
V$IK2_01	1495	(-)	1.000	0.905	gagaGGGAgggc
V$GKLF_01	1496	(-)	1.000	0.900	ggagagagggAGGG
V$IK2_01	1503	(-)	1.000	0.871	accgGGGAgaga
V$MZF1_01	1507	(-)	1.000	0.959	accGGGGa
V$CETS1P54_01	1541	(+)	1.000	0.873	gcCGGAcaag
V$GKLF_01	1544	(+)	0.887	0.861	ggacaagaggGAGG
V$ER_Q6	1547	(+)	1.000	0.876	caagagggaggTGACcgtg
V$IK2_01	1548	(+)	1.000	0.880	aagaGGGAggtg
V$TCF11_01	1549	(-)	1.000	0.870	GTCAcctccctct
V$DELTAEF1_01	1552	(-)	1.000	0.933	ggtcACCTccc
V$AP1FJ_Q2	1556	(+)	1.000	0.912	ggTGACcgtgg
V$AP1_Q2	1556	(+)	1.000	0.888	ggTGACcgtgg
V$HNF3B_01	1580	(+)	1.000	0.872	ggctcTGTTtattgt
V$XFD2_01	1581	(-)	1.000	0.926	acaaTAAAcagagc
V$FREAC2_01	1581	(-)	1.000	0.898	ggacaaTAAAcagagc
V$FREAC7_01	1581	(-)	1.000	0.919	ggacaaTAAAcagagc
V$HFH1_01	1582	(+)	1.000	0.864	ctctGTTTattg
V$HFH3_01	1582	(+)	1.000	0.896	ctcTGTTtattgt
V$HFH8_01	1582	(+)	1.000	0.933	ctcTGTTtattgt
V$SRY_02	1587	(-)	1.000	0.863	agggACAAtaaa
V$SOX5_01	1588	(-)	1.000	0.864	gggaCAATaa
V$IK2_01	1590	(-)	1.000	0.899	gagaGGGAcaat
V$CEBPB_01	1611	(+)	0.986	0.942	gtgtgagGAAAtcg
V$GFI1_01	1612	(+)	1.000	0.914	tgtgaggaAATCggggctgcagcg
V$NFAT_Q6	1613	(+)	1.000	0.936	gtgagGAAAtcg
V$AP4_Q6	1655	(+)	1.000	0.854	agCAGCggcg
V$AP4_Q5	1655	(+)	1.000	0.911	agCAGCggcg
V$IK2_01	1667	(+)	1.000	0.926	gaccGGGAcgac
V$CETS1P54_01	1669	(+)	0.833	0.854	ccGGGAcgac
V$USF_Q6	1690	(-)	0.864	0.860	gcCACTcgag
V$CAAT_01	1692	(-)	0.827	0.890	gccagCCACtcg
V$GKLF_01	1716	(+)	0.937	0.859	tgagaaggacAAGG
V$DELTAEF1_01	1727	(+)	1.000	0.928	aggcACCTctg
V$NF1_Q6	1734	(-)	1.000	0.872	acgTGGCgaggacgcaga
V$ARNT_01	1742	(-)	1.000	0.856	gctcggaCGTGgcgag
V$IK2_01	1763	(-)	1.000	0.884	gccgGGGAgccc
V$MZF1_01	1767	(-)	1.000	0.954	gccGGGGa
V$SP1_Q6	1770	(-)	1.000	0.854	cgagGGCGgccgg
V$AHRARNT_01	1787	(-)	1.000	0.867	aagcccgggCGTGcga
V$AP4_Q5	1813	(+)	1.000	0.893	tgCAGCggcg
V$AP2_Q6	1821	(-)	1.000	0.993	cgCCCGcgggcg
V$AP2_Q6	1821	(+)	1.000	0.993	cgCCCGcgggcg
V$AP2_Q6	1825	(-)	1.000	0.922	cgCCCGcccgcg
V$SP1_Q6	1825	(+)	1.000	0.899	cgcgGGCGggcga
V$RFX1_02	1831	(+)	1.000	0.884	cgggcgagaagGCAAcgc
V$IK2_01	1869	(-)	1.000	0.885	agcaGGGAgcga
V$BARBIE_01	1872	(-)	1.000	0.880	ggagAAAGcagggag
V$GKLF_01	1874	(-)	1.000	0.873	aggagaaagcAGGG
V$AP4_Q6	1886	(+)	1.000	0.871	ctCAGCcgcc
V$AP4_Q5	1886	(+)	1.000	0.909	ctCAGCcgcc
V$AP4_Q6	1939	(+)	1.000	0.859	ggCAGCcggc
V$AP4_Q5	1939	(+)	1.000	0.902	ggCAGCcggc
V$AP4_Q6	1960	(+)	1.000	0.923	ctCAGCgggc
V$AP4_Q5	1960	(+)	1.000	0.934	ctCAGCgggc
V$AP4_Q5	1969	(-)	1.000	0.861	tcCAGCcgag
V$CETS1P54_01	1973	(+)	0.852	0.905	gcTGGAcgtc
V$CETS1P54_01	1986	(+)	1.000	0.906	tcCGGAtctt
V$AP2_Q6	1996	(-)	0.976	0.858	cgCCCCatcgcg
V$USF_Q6	2033	(-)	0.864	0.853	tcCACTcgcc
V$SP1_Q6	2055	(+)	1.000	0.887	gagcGGCGgggcc
V$GC_01	2055	(+)	1.000	0.880	gagcGGCGgggccc
V$AP2_Q6	2064	(+)	0.905	0.890	ggCCCAgcggcc
V$AP4_Q6	2066	(+)	1.000	0.871	ccCAGCggcc
V$AP4_Q5	2066	(+)	1.000	0.896	ccCAGCggcc
V$AP2_Q6	2072	(+)	1.000	0.867	ggCCCGcaggga
V$IK2_01	2076	(+)	1.000	0.878	cgcaGGGAggcg
V$SP1_Q6	2080	(+)	1.000	0.866	gggaGGCGggagc
V$IK2_01	2083	(+)	1.000	0.888	aggcGGGAgcgg
V$AP4_Q5	2093	(-)	1.000	0.882	ccCAGCcgcc
V$AP2_Q6	2102	(-)	0.857	0.874	ggCCCTgaggcc
V$USF_Q6	2110	(-)	0.864	0.890	ctCACAggcc
V$AP4_Q5	2115	(-)	1.000	0.873	tcCAGCtcac
V$CETS1P54_01	2119	(+)	0.852	0.885	gcTGGAggca
V$NF1_Q6	2121	(-)	1.000	0.864	ccaTGGCgagtgcctcca
V$USF_Q6	2125	(+)	0.864	0.862	ggCACTcgcc
V$NF1_Q6	2127	(-)	1.000	0.905	actTGGCcatggcgagtg
V$USF_Q6	2143	(-)	0.818	0.880	gcCACCcgac
V$IK2_01	2157	(+)	1.000	0.888	gcgcGGGAgccg
V$EGR3_01	2167	(+)	1.000	0.956	cgGCGTgggcgg
V$EGR2_01	2167	(+)	1.000	0.939	cgGCGTgggcgg
V$EGR1_01	2167	(+)	1.000	0.904	cggcgtgGGCGg
V$NGFIC_01	2167	(+)	1.000	0.936	cgGCGTgggcgg
V$RFX1_02	2170	(+)	1.000	0.871	cgtgggcggcgGCAAcgg
V$GC_01	2170	(+)	1.000	0.859	cgtgGGCGgcggca
V$SP1_Q6	2170	(+)	1.000	0.899	cgtgGGCGgcggc
V$CMYB_01	2178	(-)	1.000	0.911	tcagtgccccGTTGccgc
V$VMYB_01	2180	(+)	1.000	0.898	ggcAACGggg
V$VMYB_02	2180	(+)	1.000	0.954	ggcAACGgg
V$ER_Q6	2182	(+)	1.000	0.887	caacggggcacTGACctgg

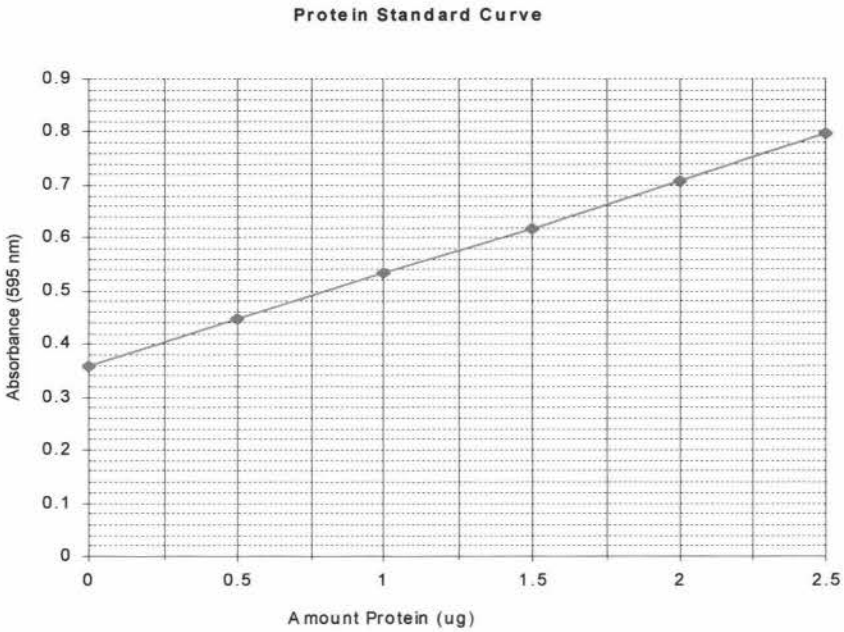
V$AP1FJ_Q2	2191	(+)	1.000	0.891	acTGACctggg
V$AP1_Q2	2191	(+)	1.000	0.877	acTGACctggg
V$DELTAEF1_01	2191	(+)	1.000	0.869	actgACCTggg
V$RORA1_01	2192	(-)	1.000	0.918	ccacccaGGTCag
V$NKX25_01	2204	(+)	1.000	0.930	gtAAGTg
V$NFY_Q6	2204	(-)	0.805	0.853	cagCCACttac
V$CAAT_01	2205	(-)	0.827	0.904	accagCCACtta
V$NF1_Q6	2206	(+)	1.000	0.867	aagTGGCtggtccccggg
V$IK2_01	2212	(-)	1.000	0.874	cccgGGGAccag
V$AP2_Q6	2215	(+)	0.976	0.863	gtCCCCgggacg
V$MZF1_01	2216	(-)	1.000	0.948	cccGGGGa
V$IK2_01	2217	(+)	1.000	0.894	ccccGGGAcggg
V$AP4_Q5	2227	(-)	1.000	0.873	tgCAGCcgcc
V$SRY_02	2233	(-)	1.000	0.911	cggaACAAtgca
V$SOX5_01	2234	(-)	1.000	0.984	ggaaCAATgc
V$CETS1P54_01	2237	(-)	1.000	0.884	ccCGGAacaa
V$SP1_Q6	2248	(+)	1.000	0.864	ccgcGGCGggggc
V$NGFIC_01	2251	(+)	0.788	0.855	cgGCGGgggcgg
V$SP1_Q6	2254	(+)	1.000	0.900	cgggGGCGgaatt
V$GC_01	2254	(+)	1.000	0.869	cgggGGCGgaattc
V$CETS1P54_01	2275	(+)	1.000	0.906	tcCGGAtctt
V$AP2_Q6	2285	(-)	0.976	0.858	cgCCCCatcgcg
V$USF_Q6	2322	(-)	0.864	0.853	tcCACTcgcc
V$SP1_Q6	2344	(+)	1.000	0.887	gagcGGCGgggcc
V$GC_01	2344	(+)	1.000	0.880	gagcGGCGgggccc
V$AP2_Q6	2353	(+)	0.905	0.890	ggCCCAgcggcc
V$AP4_Q6	2355	(+)	1.000	0.871	ccCAGCggcc
V$AP4_Q5	2355	(+)	1.000	0.896	ccCAGCggcc
V$AP2_Q6	2361	(+)	1.000	0.867	ggCCCGcaggga
V$IK2_01	2365	(+)	1.000	0.878	cgcaGGGAggcg
V$SP1_Q6	2369	(+)	1.000	0.866	gggaGGCGggagc
V$IK2_01	2372	(+)	1.000	0.888	aggcGGGAgcgg
V$AP4_Q6	2382	(-)	1.000	0.859	cgCAGCcgcc
V$AP4_Q5	2382	(-)	1.000	0.900	cgCAGCcgcc
V$AP2_Q6	2391	(-)	0.857	0.874	ggCCCTgaggcc
V$USF_Q6	2399	(-)	0.864	0.890	ctCACAggcc
V$AP4_Q5	2404	(-)	1.000	0.873	tcCAGCtcac
V$CETS1P54_01	2408	(+)	0.852	0.885	gcTGGAggca

In 1 seq. 3 matches to V$AHRARNT_01 (re: 0.81) were found.
In 0 seq. 0 matches to V$AHRARNT_02 (re: < 0.01) were found.
In 0 seq. 0 matches to V$AHR_01 (re: < 0.01) were found.
In 1 seq. 13 matches to V$AP1FJ_Q2 (re: 2.45) were found.
In 1 seq. 2 matches to V$AP1_C (re: 1.08) were found.
In 1 seq. 11 matches to V$AP1_Q2 (re: 1.82) were found.
In 1 seq. 4 matches to V$AP1_Q4 (re: 2.48) were found.
In 1 seq. 13 matches to V$AP2_Q6 (re: 4.78) were found.
In 0 seq. 0 matches to V$AP4_01 (re: 0.01) were found.
In 1 seq. 18 matches to V$AP4_Q5 (re: 0.96) were found.
In 1 seq. 8 matches to V$AP4_Q6 (re: 0.50) were found.
In 1 seq. 1 matches to V$ARNT_01 (re: 0.69) were found.
In 0 seq. 0 matches to V$ARP1_01 (re: 0.07) were found.
In 1 seq. 1 matches to V$ATF_01 (re: 0.34) were found.
In 1 seq. 3 matches to V$BARBIE_01 (re: 0.56) were found.
In 0 seq. 0 matches to V$BRACH_01 (re: < 0.01) were found.
In 1 seq. 2 matches to V$BRN2_01 (re: 0.99) were found.
In 1 seq. 9 matches to V$CAAT_01 (re: 2.21) were found.
In 0 seq. 0 matches to V$CAAT_C (re: < 0.01) were found.
In 1 seq. 1 matches to V$CDPCR3HD_01 (re: 3.06) were found.
In 0 seq. 0 matches to V$CDPCR3_01 (re: 0.12) were found.
In 0 seq. 0 matches to V$CDP_01 (re: 0.04) were found.
In 0 seq. 0 matches to V$CDP_02 (re: 0.12) were found.
In 1 seq. 11 matches to V$CEBPB_01 (re: 2.07) were found.
In 0 seq. 0 matches to V$CEBP_C (re: 0.27) were found.
In 1 seq. 11 matches to V$CETS1P54_01 (re: 1.53) were found.
In 0 seq. 0 matches to V$CHOP_01 (re: 1.10) were found.

In 0 seq. 0 matches to V$CLOX_01 (re: 0.12) were found.
In 1 seq. 5 matches to V$CMYB_01 (re: 2.02) were found.
In 0 seq. 0 matches to V$COMP1_01 (re: 0.66) were found.
In 1 seq. 1 matches to V$COUP_01 (re: 0.19) were found.
In 1 seq. 1 matches to V$CP2_01 (re: 0.39) were found.
In 1 seq. 2 matches to V$CREBP1CJUN_01 (re: 0.22) were found.
In 0 seq. 0 matches to V$CREBP1_01 (re: 0.15) were found.
In 1 seq. 1 matches to V$CREBP1_Q2 (re: 0.09) were found.
In 1 seq. 1 matches to V$CREB_01 (re: 0.40) were found.
In 1 seq. 1 matches to V$CREB_02 (re: 1.12) were found.
In 1 seq. 1 matches to V$CREB_Q2 (re: 0.34) were found.
In 1 seq. 1 matches to V$CREB_Q4 (re: 0.34) were found.
In 1 seq. 2 matches to V$CREL_01 (re: 2.74) were found.
In 1 seq. 9 matches to V$DELTAEF1_01 (re: 2.42) were found.
In 0 seq. 0 matches to V$E2F_01 (re: < 0.01) were found.
In 1 seq. 1 matches to V$E2F_02 (re: 0.13) were found.
In 0 seq. 0 matches to V$E2F_Q6 (re: 0.06) were found.
In 0 seq. 0 matches to V$E2_01 (re: 0.29) were found.
In 0 seq. 0 matches to V$E2_Q6 (re: 0.40) were found.
In 1 seq. 1 matches to V$E47_01 (re: 0.11) were found.
In 0 seq. 0 matches to V$E47_02 (re: 0.27) were found.
In 1 seq. 1 matches to V$E4BP4_01 (re: 0.07) were found.
In 1 seq. 1 matches to V$EGR1_01 (re: 0.03) were found.
In 1 seq. 1 matches to V$EGR2_01 (re: 0.03) were found.
In 1 seq. 1 matches to V$EGR3_01 (re: 0.01) were found.
In 0 seq. 0 matches to V$ELK1_01 (re: 0.06) were found.
In 1 seq. 5 matches to V$ER_Q6 (re: 1.73) were found.
In 0 seq. 0 matches to V$EVI1_01 (re: < 0.01) were found.
In 1 seq. 1 matches to V$EVI1_02 (re: 0.02) were found.
In 0 seq. 0 matches to V$EVI1_03 (re: < 0.01) were found.
In 0 seq. 0 matches to V$EVI1_04 (re: 0.92) were found.
In 0 seq. 0 matches to V$EVI1_05 (re: 0.16) were found.
In 1 seq. 1 matches to V$EVI1_06 (re: 0.02) were found.
In 1 seq. 5 matches to V$FREAC2_01 (re: 0.11) were found.
In 1 seq. 1 matches to V$FREAC3_01 (re: 0.02) were found.
In 1 seq. 1 matches to V$FREAC4_01 (re: < 0.01) were found.
In 1 seq. 6 matches to V$FREAC7_01 (re: 0.52) were found.
In 1 seq. 2 matches to V$GATA1_02 (re: 2.27) were found.
In 1 seq. 2 matches to V$GATA1_03 (re: 1.80) were found.
In 1 seq. 2 matches to V$GATA1_04 (re: 1.82) were found.
In 1 seq. 1 matches to V$GATA1_05 (re: 0.42) were found.
In 1 seq. 2 matches to V$GATA1_06 (re: 0.52) were found.
In 1 seq. 1 matches to V$GATA2_02 (re: 1.02) were found.
In 1 seq. 3 matches to V$GATA2_03 (re: 0.74) were found.
In 1 seq. 6 matches to V$GATA3_02 (re: 2.85) were found.
In 1 seq. 6 matches to V$GATA3_03 (re: 2.01) were found.
In 1 seq. 2 matches to V$GATA_C (re: 2.62) were found.
In 1 seq. 9 matches to V$GC_01 (re: 2.12) were found.
In 1 seq. 5 matches to V$GFI1_01 (re: 1.57) were found.
In 1 seq. 15 matches to V$GKLF_01 (re: 4.76) were found.
In 0 seq. 0 matches to V$GRE_C (re: 0.06) were found.
In 0 seq. 0 matches to V$HEN1_01 (re: 0.12) were found.
In 0 seq. 0 matches to V$HEN1_02 (re: 0.07) were found.
In 1 seq. 2 matches to V$HFH1_01 (re: 0.12) were found.
In 1 seq. 5 matches to V$HFH2_01 (re: 2.03) were found.
In 1 seq. 13 matches to V$HFH3_01 (re: 0.31) were found.
In 1 seq. 10 matches to V$HFH8_01 (re: 0.23) were found.
In 0 seq. 0 matches to V$HLF_01 (re: 1.69) were found.
In 0 seq. 0 matches to V$HNF1_01 (re: 0.51) were found.
In 0 seq. 0 matches to V$HNF1_C (re: 0.11) were found.
In 1 seq. 15 matches to V$HNF3B_01 (re: 1.45) were found.
In 0 seq. 0 matches to V$HOX13_01 (re: 0.02) were found.
In 1 seq. 6 matches to V$IK1_01 (re: 0.86) were found.
In 1 seq. 31 matches to V$IK2_01 (re: 3.95) were found.
In 1 seq. 1 matches to V$IK3_01 (re: 0.16) were found.

```
In   1 seq.    3 matches to V$IRF1_01 (re:    0.03) were found.
In   1 seq.    2 matches to V$IRF2_01 (re: < 0.01) were found.
In   0 seq.    0 matches to V$ISRE_01 (re: < 0.01) were found.
In   1 seq.    3 matches to V$LMO2COM_01 (re:    1.11) were found.
In   1 seq.    2 matches to V$LMO2COM_02 (re:    2.02) were found.
In   1 seq.    4 matches to V$LYF1_01 (re:    4.29) were found.
In   0 seq.    0 matches to V$MAX_01 (re:    0.11) were found.
In   0 seq.    0 matches to V$MEF2_01 (re: < 0.01) were found.
In   1 seq.    1 matches to V$MEF2_02 (re:    0.08) were found.
In   0 seq.    0 matches to V$MEF2_03 (re:    0.02) were found.
In   0 seq.    0 matches to V$MEF2_04 (re: < 0.01) were found.
In   0 seq.    0 matches to V$MIF1_01 (re:    0.01) were found.
In   0 seq.    0 matches to V$MYCMAX_01 (re:    0.05) were found.
In   0 seq.    0 matches to V$MYCMAX_02 (re:    1.67) were found.
In   0 seq.    0 matches to V$MYOD_01 (re:    0.21) were found.
In   1 seq.    3 matches to V$MYOD_Q6 (re:    0.96) were found.
In   0 seq.    0 matches to V$MYOGNF1_01 (re: < 0.01) were found.
In   1 seq.    7 matches to V$MZF1_01 (re:    3.84) were found.
In   1 seq.   11 matches to V$NF1_Q6 (re:    4.11) were found.
In   1 seq.   15 matches to V$NFAT_Q6 (re:    1.91) were found.
In   1 seq.    1 matches to V$NFE2_01 (re:    0.12) were found.
In   0 seq.    0 matches to V$NFKAPPAB50_01 (re:    0.05) were found.
In   1 seq.    1 matches to V$NFKAPPAB65_01 (re:    0.09) were found.
In   1 seq.    1 matches to V$NFKAPPAB_01 (re:    0.42) were found.
In   0 seq.    0 matches to V$NFKB_C (re:    0.05) were found.
In   0 seq.    0 matches to V$NFKB_Q6 (re:    0.28) were found.
In   1 seq.    5 matches to V$NFY_01 (re:    0.49) were found.
In   0 seq.    0 matches to V$NFY_C (re: < 0.01) were found.
In   1 seq.    6 matches to V$NFY_Q6 (re:    0.70) were found.
In   1 seq.    3 matches to V$NGFIC_01 (re:    0.03) were found.
In   1 seq.    6 matches to V$NKX25_01 (re:    0.79) were found.
In   1 seq.    2 matches to V$NKX25_02 (re:    0.51) were found.
In   1 seq.    1 matches to V$NMYC_01 (re:    1.64) were found.
In   1 seq.    3 matches to V$NRF2_01 (re:    0.16) were found.
In   0 seq.    0 matches to V$NRSF_01 (re: < 0.01) were found.
In   0 seq.    0 matches to V$OCT1_01 (re:    0.04) were found.
In   1 seq.    1 matches to V$OCT1_02 (re:    2.34) were found.
In   1 seq.    2 matches to V$OCT1_05 (re:    0.03) were found.
In   1 seq.    7 matches to V$OCT1_06 (re:    0.97) were found.
In   1 seq.    1 matches to V$OCT1_07 (re:    0.04) were found.
In   1 seq.    2 matches to V$OCT1_Q6 (re:    4.82) were found.
In   1 seq.    2 matches to V$OCT_C (re:    0.05) were found.
In   0 seq.    0 matches to V$OLF1_01 (re:    0.03) were found.
In   0 seq.    0 matches to V$P53_01 (re: < 0.01) were found.
In   1 seq.    4 matches to V$PADS_C (re:    0.52) were found.
In   0 seq.    0 matches to V$PAX5_01 (re:    0.04) were found.
In   0 seq.    0 matches to V$PAX5_02 (re: < 0.01) were found.
In   0 seq.    0 matches to V$PAX6_01 (re:    1.24) were found.
In   0 seq.    0 matches to V$PBX1_02 (re:    0.05) were found.
In   0 seq.    0 matches to V$POLY_C (re:    0.02) were found.
In   0 seq.    0 matches to V$PPARA_01 (re: < 0.01) were found.
In   1 seq.    2 matches to V$RFX1_01 (re:    0.94) were found.
In   1 seq.    2 matches to V$RFX1_02 (re:    0.95) were found.
In   1 seq.    4 matches to V$RORA1_01 (re:    0.19) were found.
In   0 seq.    0 matches to V$RORA2_01 (re:    0.01) were found.
In   1 seq.    1 matches to V$RREB1_01 (re:    0.03) were found.
In   1 seq.    1 matches to V$RSRFC4_01 (re:    0.02) were found.
In   0 seq.    0 matches to V$R_01 (re:    0.04) were found.
In   1 seq.    8 matches to V$S8_01 (re:    0.67) were found.
In   0 seq.    0 matches to V$SEF1_C (re: < 0.01) were found.
In   1 seq.    4 matches to V$SOX5_01 (re:    1.10) were found.
In   1 seq.   15 matches to V$SP1_Q6 (re:    1.70) were found.
In   0 seq.    0 matches to V$SREBP1_01 (re:    0.21) were found.
In   0 seq.    0 matches to V$SREBP1_02 (re: < 0.01) were found.
In   0 seq.    0 matches to V$SRF_01 (re: < 0.01) were found.
In   0 seq.    0 matches to V$SRF_C (re:    0.08) were found.
In   1 seq.    1 matches to V$SRF_Q6 (re:    0.16) were found.
In   1 seq.   11 matches to V$SRY_02 (re:    3.94) were found.
In   0 seq.    0 matches to V$STAF_01 (re:    0.02) were found.
In   1 seq.    1 matches to V$STAF_02 (re:    0.12) were found.
In   0 seq.    0 matches to V$STAT1_01 (re: < 0.01) were found.
In   0 seq.    0 matches to V$STAT3_01 (re: < 0.01) were found.
In   1 seq.    2 matches to V$STAT_01 (re:    0.48) were found.
In   1 seq.    1 matches to V$T3R_01 (re:    0.03) were found.
In   1 seq.    2 matches to V$TAL1ALPHAE47_01 (re:    0.14) were found.
In   1 seq.    2 matches to V$TAL1BETAE47_01 (re:    0.11) were found.
In   0 seq.    0 matches to V$TAL1BETAITF2_01 (re:    0.06) were found.
In   1 seq.    4 matches to V$TATA_01 (re:    5.33) were found.
In   1 seq.    5 matches to V$TATA_C (re:    0.30) were found.
In   0 seq.    0 matches to V$TAXCREB_01 (re:    0.06) were found.
In   0 seq.    0 matches to V$TAXCREB_02 (re: < 0.01) were found.
In   0 seq.    0 matches to V$TCF11MAFG_01 (re:    0.06) were found.
In   1 seq.   19 matches to V$TCF11_01 (re:    4.63) were found.
In   0 seq.    0 matches to V$TH1E47_01 (re:    1.95) were found.
In   1 seq.    1 matches to V$TST1_01 (re:    4.93) were found.
In   0 seq.    0 matches to V$USF_01 (re:    1.15) were found.
In   1 seq.    1 matches to V$USF_C (re:    2.19) were found.
In   1 seq.    9 matches to V$USF_Q6 (re:    2.52) were found.
In   1 seq.    2 matches to V$VBP_01 (re:    3.78) were found.
In   0 seq.    0 matches to V$VJUN_01 (re: < 0.01) were found.
In   0 seq.    0 matches to V$VMAF_01 (re:    0.99) were found.
In   1 seq.    6 matches to V$VMYB_01 (re:    2.29) were found.
In   1 seq.    5 matches to V$VMYB_02 (re:    1.04) were found.
In   0 seq.    0 matches to V$XBP1_01 (re:    0.34) were found.
In   1 seq.    5 matches to V$XFD1_01 (re:    0.07) were found.
In   1 seq.    5 matches to V$XFD2_01 (re:    0.07) were found.
In   1 seq.    2 matches to V$XFD3_01 (re:    0.04) were found.
In   1 seq.    1 matches to V$YY1_02 (re:    0.13) were found.
In   0 seq.    0 matches to V$ZID_01 (re:    0.33) were found.
```

Appendix 4: Protein standard curve.



BSA standard µg	Absorbance 595 nm
0	0.360
0.5	0.450
1.0	0.535
1.5	0.620
2.0	0.709
2.5	0.798

A)

595 nm Absorbance First HeLa extract 1 µL neat	Amount protein µg First HeLa extract
0.651	1.64
0.655	1.67
0.649	1.62
	Average 1.64
Concentration 1.64 µg/µL	

B)

595 nm Absorbance Second HeLa extract Dil 1:10, 2 µL	Amount protein µg Second HeLa extract
0.464	0.60
0.471	0.63
0.470	0.63
	Average 0.62
Concentration 3.1 µg/µL	

C)

Results of the Protein Standard Curve.

The Bradford protein assay carried out as described in section 2.2.24. The table on the left was used to construct the protein standard curve using the range of absorbances generated when different amounts of BSA protein standards was mixed with Bradford reagent (**A**). HeLa extracts (table **B** or **C**) were also mixed with Bradford reagent and the absorbancies from these samples were used to relate the amount of protein from the graph shown above. The amounts of protein present in HeLa triplicates was averaged and used to calculate the concentration of protein present in HeLa extract.

Appendix 5: pGL3B –617 topoisomerase IIα sequences.

The regulatory regions within the –617 topoisomerase IIα sequence were mapped by carrying out a bestfit (GCG, Wisconsin Package version 9.1) to topoisomerase IIα promoter sequences (accession number X66794).

5.1 pGL3B wt –617 topoisomerase IIα sequence:

Restriction endonuclease MAP (GCG, Wisconsin Package version 9.1) of wt –617topoisomerase IIα pGL3B (from 1 to 853 bp) With 2 enzymes: *Hin*d III and *Sac* I.

5.2 Sequences of ICB1, GC1 and GC2 within the topoisomerase IIα minimal promoter and the corresponding mutations used in pGL3B constructs:

GC1wt	GGGCGGG CCCGCCC	GC1mt	G**T**GCG**T**G C**A**CGC**A**C
ICB1wt	ATTGG TAACC	ICB1mt	ATT**CC** TAA**GG**
GC2wt	GGGGCGGGG CCCCGCCCC	GC2mt	GG**T**GC**T**GGG CC**A**CGC**A**CC

Appendix 6: Ligations and Transformations.

6.1 Ligation reactions to generate –617 topoisomerase IIα pGL3B constructs.

Sample	ng vector	ng insert	μL vector	μL insert	T_4 buffer	T_4 ligase	H_2O	Total volume
Vector (GC2-) plus ligase (Control)	50 ng	-	1 μL	-	4 μL	1 μL	14 μL	20 μL
Vector (GC2-) minus ligase (Control)	50 ng	-	1 μL	-	4 μL	-	15 μL	20 μL
Vector (GC2-) plus insert (ICB1-) DNA								
1:3	50 ng	12 ng	1 μL	2.5 μL	4 μL	1 μL	11.5 μL	20 μL
1:2	50 ng	8 ng	1 μL	2.0 μL	4 μL	1 μL	12 μL	20 μL
Vector (GC2-) plus insert (GC1-) DNA								
1:3	50 ng	12 ng	1 μ L	2.5 μL	4 μL	1 μL	11.5 μL	20 μL
1:2	50 ng	8 ng	1 μL	2.0 μL	4 μL	1 μL	12 μL	20 μL
Vector (GC2-) plus insert (ICB1-/GC1-) DNA								
1:3	50 ng	12 ng	1 μL	1.2 μL	4 μL	1 μL	12.8 μL	20 μL
1:2	50 ng	8 ng	1 μL	0.8 μL	4 μL	1 μL	13.2 μL	20 μL

Components of ligation reactions.

Ligation reactions used to generate various –617 topoisomerase IIα pGL3B with mutations in GC2, ICB1 and/or GC1. Ligations were incubated overnight at 4°C and transformation of XL-1 cells were carried out the following day with half of each ligation mix.

6.2 Transformation Reactions using ligation reactions in appendix 6.1.

Sample	Ligation mix	Competent cells	Dilution	Volume plated	Number of colonies
XL-1 (control) no amp	10 μL 10 μL	100 μL 100 μL	10^7 10^8	50 μL 50 μL	289 35
XL-1 (control) with amp	10 μL	100 μL	Neat	10 μL	0
Vector GC2mt plus ligase (control)	10 μL 10 μL	100 μL 100 μL	Neat 10	50 μL 50 μL	27 3
Vector GC2mt minus ligase (control)	10 μL 10 μL	100 μL 100 μL	Neat 10	50 μL 50 μL	5 0
Vector plus insert (GC2mt/ICB1mt pGL3B) **1:3**	10 μL 10 μL 10 μL 10 μL	100 μL 100 μL 100 μL 100 μL	Neat Neat 10 10^2	50 μL 30 μL 50 μL 50 μL	78 45* 0 0
1:2	10 μL 10 μL 10 μL 10 μL	100 μL 100 μL 100 μL 100 μL	Neat Neat 10 10^2	50 μL 30 μL 50 μL 50 μL	95 75 3* 0
Vector plus insert (GC2mt/GC1mt pGL3B) **1:3**	10 μL 10 μL 10 μL 10 μL	100 μL 100 μL 100 μL 100 μL	Neat Neat 10 10^2	50 μL 30 μL 50 μL 50 μL	80 72* 1* 0
1:2	10 μL 10 μL 10 μL 10 μL	100 μL 100 μL 100 μL 100 μL	Neat Neat 10 10^2	50 μL 30 μL 50 μL 50 μL	114 78* 0 0
Vector plus insert (GC2mt/ICB1mt/GC1mt pGL3B) **1:3**	10 μL 10 μL 10 μL 10 μL	100 μL 100 μL 100 μL 100 μL	Neat Neat 10 10^2	50 μL 30 μL 50 μL 50 μL	412 202 3* 0
1:2	10 μL 10 μL 10 μL 10 μL	100 μL 100 μL 100 μL 100 μL	Neat Neat 10 10^2	50 μL 30 μL 50 μL 50 μL	472 119 11* 1

Results of transformations of *E.coli* XL-1 competent cells with ligation reactions from appendix 6.1.

Unless indicated, all samples were plated onto LB plus ampicillin plates and incubated overnight at 37°C. Colonies were creamy white in colour, and appropriate colonies (*) were selected and cultured to isolate plasmid DNA for analysis.

Appendix 7: Transient transfection data.

7.1 Topoisomerase IIα promoter activity with mutations in ICB1, GC1 and GC2.

Luciferase experiment 1

Luciferase Maxima

pSVB-gal	pGL3B	-617wt	ICB1mt	GC1mt	GC2mt	ICB1mt/ GC1mt	GC2mt/ ICB1mt	GC2/ GC1mt	GC2mt /ICB1mt/ GC1mt	Blank
1134	1296	40902	8016	20795	37030	11696	45410	34770	12329	470
1153	1183	47499	8735	22378	47860	15221	52882	36637	11005	578
1147	1323	46301	7123	26639	42242	24621	50988	41428	13187	559
									avg	535.67
		B-Gal Results								
0.499	0.586	0.477	0.296	0.301	0.423	0.394	0.409	0.391	0.442	0.121
0.429	0.553	0.452	0.28	0.288	0.449	0.386	0.42	0.356	0.389	0.125
0.465	0.611	0.487	0.271	0.331	0.426	0.426	0.367	0.367	0.393	0.116
									avg	0.1207
		luciferase minus blank								
598.33	760.33	40366	7480.3	20259	36494	11160	44874	34234	11793	
617.33	647.33	46963	8199.3	21842	47324	14685	52346	36101	10469	
611.33	787.33	45765	6587.3	26103	41706	24085	50452	40892	12651	
		B-Gal minus blank								
0.3783	0.4653	0.3563	0.1753	0.1803	0.3023	0.2733	0.2883	0.2703	0.3213	
0.3083	0.4323	0.3313	0.1593	0.1673	0.3283	0.2653	0.2993	0.2353	0.2683	
0.3443	0.4903	0.3663	0.1503	0.2103	0.3053	0.3053	0.2463	0.2463	0.2723	
		Normalized: luciferase/B-gal								
1581.5	1634	113283	42663	112344	120709	40830	155634	126637	36701	
2002.2	1497.3	141740	51460	130532	144135	55347	174876	153405	39016	
1775.4	1605.7	124928	43818	124105	136593	78882	204813	166004	46455	
Average										
1786.4	1579	126650	45981	122327	133812	58353	178441	148682	40724	
		Normalized: average relative to wt								
1.4105	1.2467	100	36.305	96.586	105.65	46.074	140.89	117.4	32.155	
				Average deviation from mean						
143.87	54.458	10060	3653.1	6655.3	8735.5	13686	17581	14696	3820.7	
% error										
8.0538	3.4489	7.9432	7.9448	5.4406	6.5282	23.454	9.8528	9.8845	9.3819	

Luciferase Maxima

pSVB-gal	pGL3B	-617wt	ICB1mt	GC1mt	GC2mt	ICB1mt/ GC1mt	GC2mt/ ICB1mt	GC2/ GC1mt	GC2mt/ ICB1mt/ GC1mt	Blank
1295	1310	23157	10557	15699	25094	13835	25110	30437	11005	841
1278	1350	22584	9095	17021	23196	18170	27036	30546	15080	883
1324	1331	26748	10033	16165	22346	16150	25552	21265	12707	854
									avg	859.333

B-Gal Results

0.55	0.463	0.553	0.4	0.412	0.408	0.49	0.474	0.495	0.474	0.139
0.536	0.448	0.507	0.326	0.402	0.402	0.51	0.555	0.423	0.57	0.134
0.472	0.465	0.64	0.389	0.428	0.462	0.553	0.599	0.408	0.66	0.132
									avg	0.135

luciferase minus blank

435.667	450.667	22297.7	9697.67	14839.7	24234.7	12975.7	24250.7	29577.7	10145.7
418.667	490.667	21724.7	8235.67	16161.7	22336.7	17310.7	26176.7	29686.7	14220.7
464.667	471.667	25888.7	9173.67	15305.7	21486.7	15290.7	24692.7	20405.7	11847.7

B-Gal minus blank

0.415	0.328	0.418	0.265	0.277	0.273	0.355	0.339	0.36	0.339
0.401	0.313	0.372	0.191	0.267	0.267	0.375	0.42	0.288	0.435
0.337	0.33	0.505	0.254	0.293	0.327	0.418	0.464	0.273	0.525

Normalized: luciferase/Bgal

1049.8	1373.98	53343.7	36595	53572.8	88771.7	36551.2	71535.9	82160.2	29928.2
1044.06	1567.63	58399.6	43118.7	60530.6	83657.9	46161.8	62325.4	103079	32691.2
1378.83	1429.29	51264.7	36116.8	52237.8	65708.5	36580.5	53217	74746	22567
Averages									
1157.56	1456.97	54336	38610.1	55447.1	79379.4	39764.5	62359.4	86661.6	28395.5

Normalized average relative to wt

2.13038	2.6814	100	71.0581	102.045	146.09	73.1826	114.766	159.492	52.259

average deviation from the mean

147.513	73.7719	2709.09	3005.68	3389.02	9113.93	4264.85	6117.65	10944.7	3885.65
% error									
12.7434	5.06339	4.98581	7.7847	6.11218	11.4815	10.7253	9.81031	12.6292	13.6841

Luciferase Maxima

pSVB-gal	pGL3B	-617wt	ICB1mt	GC1mt	GC2mt	ICB1mt/ GC1mt	GC2mt/ ICB1mt	GC2/ GC1mt	GC2mt/ ICB1mt/ GC1mt	Blank
823	1075	42364	19192	25157	57503	17341	31722	26019	9141	687
837	1048	44478	17735	31780	62925	16385	31921	28971	9770	617
857	1042	42625	14085	29373	43086	15236	31185	33814	9570	730
									avg	678
		B-Gal Results								
0.697	0.67	0.812	0.796	0.633	0.633	0.691	0.492	0.564	0.558	0.134
0.789	0.791	0.747	0.813	0.68	0.68	0.756	0.487	0.65	0.565	0.131
0.698	0.546	0.601	0.635	0.54	0.54	0.653	0.519	0.604	0.596	0.134
									avg	0.133
		luciferase minus blank								
145	397	41686	18514	24479	56825	16663	31044	25341	8463	
159	370	43800	17057	31102	62247	15707	31243	28293	9092	
179	364	41947	13407	28695	42408	14558	30507	33136	8892	
		B-Gal minus blank								
0.564	0.537	0.679	0.663	0.5	0.5	0.558	0.359	0.431	0.425	
0.656	0.658	0.614	0.68	0.547	0.547	0.623	0.354	0.517	0.432	
0.565	0.413	0.468	0.502	0.407	0.407	0.52	0.386	0.471	0.463	
		Normalized: luciferase/B-gal								
257.09	739.29	61393	27925	48958	113650	29862	86474	58796	19913	
242.38	562.31	71336	25084	56859	113797	25212	88257	54725	21046	
316.81	881.36	89630	26707	70504	104197	27996	79034	70352	19205	
Averages										
272.09	727.65	74120	26572	58774	110548	27690	84588	61291	20055	
		Normalized average relative to wt								
0.3671	0.9817	100	35.85	79.296	149.15	37.359	114.12	82.692	27.057	
		average deviation from the mean								
29.813	110.23	10340	992.02	7820	4234.2	1652.1	3702.9	6040.8	660.99	
% error										
10.957	15.149	13.951	3.7334	13.305	3.8302	5.9664	4.3776	9.8559	3.2959	

Luciferase Maxima

pSVB-gal	pGL3B	-617wt	ICB1mt	GC1mt	GC2mt	ICB1mt/ GC1mt	GC2mt/ ICB1mt	GC2/ GC1mt	GC2mt/ ICB1mt/ GC1mt	Blank
2113	2357	52320	15212	22085	41164	14582	30919	39348	11420	1517
2100	2348	57137	16047	21962	41392	13162	36584	35548	12895	1621
2116	2360	60389	16776	24003	45350	12452	36374	39633	10256	1509
									avg	1549
		B-Gal Results								
0.622	0.655	0.6	0.685	0.498	0.552	0.637	0.621	0.579	0.698	0.145
0.69	0.779	0.639	0.627	0.522	0.682	0.698	0.606	0.59	0.751	0.147
0.612	0.674	0.705	0.686	0.601	0.65	0.711	0.644	0.624	0.739	0.141
									avg	0.1443
		luciferase minus blank								
564	808	50771	13663	20536	39615	13033	29370	37799	9871	
551	799	55588	14498	20413	39843	11613	35035	33999	11346	
567	811	58840	15227	22454	43801	10903	34825	38084	8707	
		B-Gal minus blank								
0.4777	0.5107	0.4557	0.5407	0.3537	0.4077	0.4927	0.4767	0.4347	0.5537	
0.5457	0.6347	0.4947	0.4827	0.3777	0.5377	0.5537	0.4617	0.4457	0.6067	
0.4677	0.5297	0.5607	0.5417	0.4567	0.5057	0.5667	0.4997	0.4797	0.5947	
		Normalized: luciferase/B-gal								
1180.7	1582.2	111421	25271	58066	97175	26454	61615	86961	17828	
1009.8	1258.9	112375	30037	54050	74104	20975	75888	76288	18702	
1212.4	1531.2	104946	28111	49169	86620	19241	69696	79397	14642	
Averages										
1134.3	1457.4	109581	27806	53762	85966	22223	69067	80882	17057	
		Normalized average relative to wt								
1.0351	1.33	100	25.375	49.061	78.45	20.28	63.028	73.81	15.566	
		average deviation from the mean								
83.021	132.34	3089.6	1690.5	3061.7	7908.5	2820.6	4967.5	4052.7	1610.4	
% error										
7.3191	9.0804	2.8194	6.0796	5.6949	9.1995	12.692	7.1923	5.0106	9.4413	

COLLATED RESULTS OF LUCIFERASE EXPERIMENTS 1-4:

SUMMARY OF % RELATIVE TO WT

pSVB-gal	pGL3B	wt-617	ICB1mt	GC1mt	GC2mt	ICB1mt/GC 1mt	ICB1mt/GC 2mt	GC1-/GC2-	ICB1mt/GC1mt/GC2mt	
1.41	1.25	100	36.31	96.59	105.66	46.07	140.89	117.39	32.15	**Expt 1**
2.13	2.68	100	71.06	102.04	146.09	73.18	114.77	159.49	52.26	**Expt 2**
0.367	0.98	100	35.85	79.3	149.15	37.36	114.12	82.69	27.06	**Expt 3**
1.04	1.33	100	25.38	49.06	78.45	20.28	63.03	73.81	15.57	**Expt 4**
Averages (graphed)										
1.23675	1.56	100	42.15	81.7475	119.8375	44.2225	108.2025	108.345	31.76	**Overall**
% error(graphed)										
0.623556	0.695556	0	15.54667	8.895556	18.64889	13.98444	11.75333	26.42222	10.06889	**Overall**

7.2 Sp1 titrations

Luciferase experiment 5

LUCIFERASE MAXIMA

pGL3B	wt -617	Wt +0.25μg Sp1	Wt +0.5μg Sp1	Wt +1μg Sp1	ICB1mt	ICB1mt +0.25μg Sp1	ICB1mt +0.5μg Sp1	ICB1mt +1μg Sp1	BLANK
798	4354	11323	25390	12210	2319	3871	9008	2780	489
738	4500	14269	23136	11825	1929	2987	8479	3601	525
709	3703	10351	20029	13713	2245	3496	4169	4005	504
									506

BGAL READINGS

0.308	0.445	0.449	0.5	0.338	0.47	0.56	0.329	0.334	0.145
0.309	0.47	0.512	0.44	0.326	0.433	0.447	0.348	0.397	0.145
0.299	0.4	0.403	0.351	0.318	0.406	0.466	0.435	0.427	0.184
									0.158

CORRECTED LUCIFERASE VALUES

292	3848	10817	24884	11704	1813	3365	8502	2274
232	3994	13763	22630	11319	1423	2481	7973	3095
203	3197	9845	19523	13207	1739	2990	3663	3499

CORRECTED BGAL READINGS

0.15	0.287	0.291	0.342	0.18	0.312	0.402	0.171	0.176
0.151	0.312	0.354	0.282	0.168	0.275	0.289	0.19	0.239
0.141	0.242	0.245	0.193	0.16	0.248	0.308	0.277	0.269

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

1946.7	13408	37172	72760	65022	5810.9	8370.6	49719	12920
1536.4	12801	38879	80248	67375	5174.5	8584.8	41963	12950
1439.7	13211	40184	101155	82544	7012.1	9707.8	13224	13007

AVERAGE OF NORMALIZED LUCIFERASE VALUES

1640.9	13140	38745	84721	71647	5999.2	8887.7	34969	12959

NORMALIZED LUCIFERASE RELATIVE TO WT %

12.488	100	294.86	644.76	545.26	45.656	67.639	266.13	98.625

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

203.82	225.74	1048.6	10956	7264.5	675.28	546.7	14497	32.139

percentage error

12.421	1.718	2.7064	12.932	10.139	11.256	6.1512	41.456	0.248

LUCIFERASE MAXIMA

pGL3B	WT	WT +0.25μg Sp1	WT +0.5μg Sp1	WT+ 1μg Sp1	lcb1-	lcb1- +0.25μg Sp1	lcb1- +0.5μg Sp1	lcb1- +1μg Sp1	BLANK
1588	2628	3810	16530	5571	2088	2583	3448	3654	1240
1523	2704	4389	16883	7422	2185	2842	3728	4673	1304
1566	2989	4608	11467	7682	2582	2847	4062	4211	1244
									1262.7

BGAL READINGS

0.238	0.274	0.249	0.295	0.231	0.387	0.45	0.392	0.402	0.149
0.257	0.278	0.287	0.311	0.29	0.448	0.455	0.457	0.451	0.146
0.246	0.277	0.264	0.255	0.279	0.409	0.458	0.42	0.474	0.147
									0.1473

CORRECTED LUCIFERASE VALUES

325.33	1365.3	2547.3	15267	4308.3	825.33	1320.3	2185.3	2391.3
260.33	1441.3	3126.3	15620	6159.3	922.33	1579.3	2465.3	3410.3
303.33	1726.3	3345.3	10204	6419.3	1319.3	1584.3	2799.3	2948.3

CORRECTED BGAL READINGS

0.0907	0.1267	0.1017	0.1477	0.0837	0.2397	0.3027	0.2447	0.2547
0.1097	0.1307	0.1397	0.1637	0.1427	0.3007	0.3077	0.3097	0.3037
0.0987	0.1297	0.1167	0.1077	0.1317	0.2617	0.3107	0.2727	0.3267

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

3588.2	10779	25056	103391	51494	3443.7	4362.3	8931.9	9390.1
2373.9	11031	22384	95440	43173	3067.6	5133.3	7961.2	11231
3074.3	13314	28674	94777	48754	5042	5099.8	10267	9025.5

AVERAGE OF NORMALIZED LUCIFERASE VALUES

3012.1	11708	25371	97869	47807	3851.1	4865.1	9053.2	9882

NORMALIZED LUCIFERASE RELATIVE TO WT %

25.728	100	216.71	835.94	408.34	32.894	41.555	77.327	84.406

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

425.52	1070.6	2201.9	3680.9	3089.5	793.95	335.19	808.86	898.99

percentage error

14.127	9.1444	8.6787	3.761	6.4624	20.616	6.8897	8.9345	9.0973

LUCIFERASE MAXIMA

pGL3B	WT	wt+ 0.25μg Sp1	wt+ 0.5 μg Sp1	wt+ 1.0μg Sp1	Icb1-	Icb1- +0.25μg Sp1	Icb1- +0.5μg Sp1	Icb1-+ 1μ g Sp1	BLANK
553	1395	4521	5134	6379	1214	1355	4591	1816	310
503	2135	4646	7408	11924	779	1562	3159	1017	302
525	2158	4684	8431	11676	1041	1871	6323	3471	311
									307.67

BGAL READINGS

0.527	0.579	0.491	0.475	0.471	0.48	0.498	0.464	0.393	0.141
0.493	0.54	0.542	0.532	0.461	0.517	0.537	0.415	0.396	0.134
0.474	0.521	0.548	0.557	0.491	0.456	0.529	0.44	0.4	0.131
									0.1353

CORRECTED LUCIFERASE VALUES

245.33	1087.3	4213.3	4826.3	6071.3	906.33	1047.3	4283.3	1508.3
195.33	1827.3	4338.3	7100.3	11616	471.33	1254.3	2851.3	709.33
217.33	1850.3	4376.3	8123.3	11368	733.33	1563.3	6015.3	3163.3

CORRECTED BGAL READINGS

0.3917	0.4437	0.3557	0.3397	0.3357	0.3447	0.3627	0.3287	0.2577
0.3577	0.4047	0.4067	0.3967	0.3257	0.3817	0.4017	0.2797	0.2607
0.3387	0.3857	0.4127	0.4217	0.3557	0.3207	0.3937	0.3047	0.2647

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

626.38	2450.8	11846	14209	18087	2629.6	2887.9	13032	5853.8
546.13	4515.7	10668	17900	35669	1234.9	3122.8	10195	2721.2
641.73	4797.8	10605	19265	31963	2286.9	3971.2	19744	11952

AVERAGE OF NORMALIZED LUCIFERASE VALUES

604.75	3921.4	11040	17125	28573	2050.5	3327.3	14324	6842.4

NORMALIZED LUCIFERASE RELATIVE TO WT %

15.422	100	281.53	436.7	728.65	52.289	84.85	365.28	174.49

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

39.078	980.41	537.68	1943.7	6990.7	543.69	429.27	3613.3	3406.5

percentage error

6.4618	25.001	4.8704	11.35	24.466	26.516	12.902	25.226	49.785

COLLATED RESULTS OF LUCIFERASE EXPERIMENTS 5-7:

PGL3B	wt-617	wt+0.25ug Sp1	wt+0.5ug Sp1	wt+1ug Sp1	ICB1mt	ICB1mt+0.25ug Sp1	ICB1mt+0.5 ug Sp1	ICB1mt+1u g Sp1	
12.49	100	294.86	644.76	545.26	45.66	67.64	266.13	98.63	**Expt5**
25.73	100	216.71	835.94	408.34	32.89	41.55	77.33	84.41	**Expt6**
15.2	100	281.53	436.7	728.65	52.29	84.85	365.28	174.49	**Expt7**
				Averages (graphed)					
17.807	100	264.37	639.13	560.75	43.613	64.68	236.25	119.18	
				average deviation from the mean (error%, graphed)					
5.2822	0	31.771	134.96	111.93	7.1489	15.42	105.94	36.876	

7.3 Addition of 0.25 μg Sp1 expression vector to topoisomerase IIα constructs.

Luciferase experiment 8

0.25 μg Sp1 added to each assay

Luciferase maxima																		
pGL3B	pGL3B +Sp1	WT	WT +Sp1	Icb1-	Icb1- +Sp1	Gc1-	Gc1- +Sp1	Gc2-	Gc2- +Sp1	Icb1- /Gc1-	Icb1- /Gc1- +Sp1	Icb1- /Gc2-	Icb1- /Gc2- +Sp1	Gc1- /Gc2-	Gc1- /Gc2- +Sp1	Triple	Triple +Sp1	blank
1672	1653	6384	19022	3503	6896	3919	10267	10316	11185	3303	16567	8102	9856	9589	9985	2654	2789	1308
1676	1645	9153	15229	2672	6712	7193	8736	9985	14785	3691	15973	8543	17550	9455	9749	2544	2569	1335
1639	1674	5841	14103	3441	6502	4979	9011	9463	15770	2722	12804	9056	8211	9621	9656	2532	2899	1321
																		1321.3
β-gal																		
0.297	0.285	0.249	0.26	0.226	0.286	0.252	0.255	0.236	0.228	0.211	0.286	0.255	0.261	0.256	0.245	0.254	0.265	0.122
0.285	0.281	0.244	0.245	0.259	0.319	0.259	0.245	0.248	0.231	0.232	0.299	0.265	0.265	0.249	0.256	0.256	0.254	0.121
0.274	0.29	0.223	0.238	0.26	0.308	0.234	0.233	0.235	0.249	0.208	0.285	0.256	0.251	0.255	0.255	0.258	0.274	0.122
																		0.1217
Luciferase Minus Blank																		
350.67	331.67	5062.7	17701	2181.7	5574.7	2597.7	8945.7	8994.7	9863.7	1981.7	15246	6780.7	8534.7	8267.7	8663.7	1332.7	1467.7	
354.67	323.67	7831.7	13908	1350.7	5390.7	5871.7	7414.7	8663.7	13464	2369.7	14652	7221.7	16229	8133.7	8427.7	1222.7	1247.7	
317.67	352.67	4519.7	12782	2119.7	5180.7	3657.7	7689.7	8141.7	14449	1400.7	11483	7734.7	6889.7	8299.7	8334.7	1210.7	1577.7	
β-gal Minus Blank																		
0.1753	0.1633	0.1273	0.1383	0.1043	0.1643	0.1303	0.1333	0.1143	0.1063	0.0893	0.1643	0.1333	0.1393	0.1343	0.1233	0.1323	0.1433	
0.1633	0.1593	0.1223	0.1233	0.1373	0.1973	0.1373	0.1233	0.1263	0.1093	0.1103	0.1773	0.1433	0.1433	0.1273	0.1343	0.1343	0.1323	
0.1523	0.1683	0.1013	0.1163	0.1383	0.1863	0.1123	0.1113	0.1133	0.1273	0.0863	0.1633	0.1343	0.1293	0.1333	0.1333	0.1363	0.1523	
Normalized: luc/B-gal																		
2000	2030.6	39759	127957	20911	33923	19931	67093	78671	92762	22183	92773	50855	61254	61546	70246	10071	10240	
2171.4	2031.4	64019	112765	9835	27318	42755	60119	68578	123143	21477	82622	50384	113223	63877	62737	9101.7	9428.2	
2085.3	2095	44602	109871	15323	27803	32561	69069	71838	113471	16224	70302	57578	53271	62248	62510	8880.2	10357	
Averages																		
2085.6	2052.3	49460	116864	15356	29681	31749	65427	73029	109792	19961	81899	52939	75916	62557	65164	9350.8	10008	
Relative to wt %																		
4.2167	4.1495	100	236.28	31.048	60.011	64.191	132.28	147.65	221.98	40.359	165.59	107.03	153.49	126.48	131.75	18.906	20.235	
AVG DEV																		
57.226	28.468	9706	7395	3702.9	2827.8	7878.6	3538.6	3761.1	11354	2491.6	7731.3	3092.8	24872	880.12	3387.8	479.81	386.62	
% error (AVGDEV/avg*100)																		
2.7439	1.3871	19.624	6.3278	24.114	9.5272	24.815	5.4084	5.1502	10.341	12.482	9.4401	5.8422	32.762	1.4069	5.1988	5.1312	3.863	

Luciferase experiment 9

Luciferase Maxima

0.25 μg Sp1 added to each assay

pGL3B	pGL3B +Sp1	WT	WT +Sp1	Icb1-	Icb1- +Sp1	Gc1-	Gc1- +Sp1	Gc2-	Gc2- +Sp1	Icb1- /Gc1-	Icb1- /Gc1- +Sp1	Icb1- /Gc2-	Icb1- /Gc2- +Sp1	Gc1- /Gc2-	Gc1- /Gc2- +Sp1	Triple	Triple +Sp1	blank
2566	2658	6236	11802	3124	7588	7041	9110	7691	11904	2916	13540	8265	16588	6458	11022	3269	5549	1913
2532	2745	7959	10201	4045	7412	7059	7455	7765	12170	4314	12599	8452	15204	7552	8321	3541	4752	1887
2499	2363	5785	10008	4737	7549	6198	7569	7254	10287	4465	12565	8295	14992	7701	7155	3412	5103	1881
																		1893.7

B-Gal Results

0.272	0.265	0.244	0.23	0.227	0.298	0.23	0.224	0.238	0.234	0.225	0.245	0.241	0.254	0.241	0.246	0.253	0.251	0.125
0.264	0.274	0.235	0.231	0.248	0.285	0.231	0.219	0.235	0.245	0.221	0.236	0.256	0.235	0.255	0.231	0.243	0.265	0.128
0.272	0.221	0.212	0.194	0.279	0.298	0.194	0.2	0.198	0.224	0.231	0.231	0.234	0.244	0.254	0.239	0.241	0.277	0.122
																		0.125

luciferase minus blank

672.33	764.33	4342.3	9908.3	1230.3	5694.3	5147.3	7216.3	5797.3	10010	1022.3	11646	6371.3	14694	4564.3	9128.3	1375.3	3655.3
638.33	851.33	6065.3	8307.3	2151.3	5518.3	5165.3	5561.3	5871.3	10276	2420.3	10705	6558.3	13310	5658.3	6427.3	1647.3	2858.3
605.33	469.33	3891.3	8114.3	2843.3	5655.3	4304.3	5675.3	5360.3	8393.3	2571.3	10671	6401.3	13098	5807.3	5261.3	1518.3	3209.3

B-Gal minus blank

0.147	0.14	0.119	0.105	0.102	0.173	0.105	0.099	0.113	0.109	0.1	0.12	0.116	0.129	0.116	0.121	0.128	0.126
0.139	0.149	0.11	0.106	0.123	0.16	0.106	0.094	0.11	0.12	0.096	0.111	0.131	0.11	0.13	0.106	0.118	0.14
0.147	0.096	0.087	0.069	0.154	0.173	0.069	0.075	0.073	0.099	0.106	0.106	0.109	0.119	0.129	0.114	0.116	0.152

Normalized: luciferase/B-gal

4573.7	5459.5	36490	94365	12062	32915	49022	72892	51304	91838	10223	97053	54925	113910	39348	75441	10745	29011
4592.3	5713.6	55139	78371	17491	34490	48730	59163	53376	85636	25212	96444	50064	121003	43526	60635	13960	20417
4117.9	4888.9	44728	117599	18463	32690	62382	75671	73429	84781	24258	100673	58728	110070	45018	46152	13089	21114

Averages

4428	5354	45453	96778	16005	33365	53378	69242	59370	87418	19898	98057	54572	114994	42630	60743	12598	23514

Relative to WT %

9.742	11.779	100	212.92	35.213	73.406	117.44	152.34	130.62	192.33	43.777	215.73	120.06	253	93.791	133.64	27.717	51.733

AVG DEV

206.71	310.09	6457.9	13880	2628.8	749.81	6002.6	6719.4	9373.1	2946.4	6449.6	1744.2	3005.8	4005.9	2188.5	9798.7	1235.5	3664.5

% error (AVGDEV/avg*100)

4.6683	5.7917	14.208	14.342	16.425	2.2473	11.245	9.7041	15.788	3.3704	32.414	1.7787	5.5078	3.4836	5.1337	16.132	9.8074	15.585

Luciferase experiment 10

0.25 μg Sp1 added to each assay

Luciferase Maxima																	
pGL3B	pGL3B +Sp1	WT	WT +Sp1	Icb1-	Icb1- +Sp1	Gc1-	Gc1- +Sp1	Gc2-	Gc2- +Sp1	Icb1- /Gc1-	Icb1- /Gc1- +Sp1	Icb1- /Gc2-	Icb1- /Gc2- +Sp1	Gc1- /Gc2-	Gc1- /Gc2- +Sp1	Triple mt	Triple mt +Sp1	blank
2244	2562	10881	29741	6404	13798	16093	22627	22946	36469	8503	26533	19000	32437	18453	28921	6143	15662	1864
2541	2532	12663	28520	6874	13208	17460	25047	22616	35453	9211	25623	22229	37711	20565	31375	6469	12565	1887
2355	2413	12905	32833	7199	12068	19363	28948	23554	37886	8744	27658	18830	27031	20146	28037	6461	11585	1881
																		1877.3
B-Gal Results																		
0.393	0.428	0.329	0.372	0.41	0.45	0.452	0.457	0.354	0.433	0.337	0.456	0.331	0.401	0.35	0.351	0.279	0.395	0.125
0.378	0.402	0.334	0.341	0.427	0.462	0.432	0.422	0.375	0.445	0.367	0.427	0.389	0.439	0.38	0.376	0.29	0.371	0.128
0.346	0.415	0.301	0.365	0.365	0.356	0.441	0.432	0.284	0.421	0.35	0.454	0.349	0.355	0.39	0.345	0.309	0.364	0.122
																		0.125
luciferase minus blank																		
366.67	684.67	9003.7	27864	4526.7	11921	14216	20750	21069	34592	6625.7	24656	17123	30560	16576	27044	4265.7	13785	
663.67	654.67	10786	26643	4996.7	11331	15583	23170	20739	33576	7333.7	23746	20352	35834	18688	29498	4591.7	10688	
477.67	535.67	11028	30956	5321.7	10191	17486	27071	21677	36009	6866.7	25781	16953	25154	18269	26160	4583.7	9707.7	
B-Gal minus blank																		
0.268	0.303	0.204	0.247	0.285	0.325	0.327	0.332	0.229	0.308	0.212	0.331	0.206	0.276	0.225	0.226	0.154	0.27	
0.253	0.277	0.209	0.216	0.302	0.337	0.307	0.297	0.25	0.32	0.242	0.302	0.264	0.314	0.255	0.251	0.165	0.246	
0.221	0.29	0.176	0.24	0.24	0.231	0.316	0.307	0.159	0.296	0.225	0.329	0.224	0.23	0.265	0.22	0.184	0.239	
Normalized: luciferase/B-gal																		
1368.2	2259.6	44136	112808	15883	36679	43473	62499	92003	112311	31253	74488	83120	110723	73670	119662	27699	51054	
2623.2	2363.4	51606	123346	16545	33622	50758	78012	82955	104924	30304	78628	77090	114120	73285	117521	27828	43446	
2161.4	1847.1	62657	128982	22174	44115	55334	88178	136331	121651	30519	78361	75682	109364	68938	118908	24911	40618	
Averages																		
2050.9	2156.7	52800	121712	18201	38139	49855	76230	103763	112962	30692	77159	78630	111402	71964	118697	26813	45039	
Relative to wt %																		
3.8843	4.0847	100	230.52	34.471	72.233	94.423	144.38	196.52	213.94	58.129	146.14	148.92	210.99	136.3	224.81	50.782	85.302	
AVG DEV																		
455.17	206.4	6571.7	5935.8	2648.7	3984.4	4254.7	9153.9	21712	5792.7	374.08	1780.4	2993	1811.7	2017.3	784.14	1267.8	4010	
% error (AVGDEV/avg*100)																		
22.193	9.57	12.447	4.8769	14.553	10.447	8.5342	12.008	20.925	5.128	1.2188	2.3075	3.8064	1.6263	2.8032	0.6606	4.7282	8.9033	

COLLATED RESULTS OF LUCIFERASE EXPERIMENTS 8-10:

SUMMARY of % relative to wild type

pGL3B	pGL3B +Sp1	WT	WT +Sp1	Icb1-	Icb1- +Sp1	Gc1-	Gc1- +Sp1	Gc2-	Gc2- +Sp1	Icb1-/Gc1-	Icb1-/Gc1- +Sp1	Icb1-/Gc2-	Icb1-/Gc2- +Sp1	Gc1-/Gc2-	Gc1-/Gc2- +Sp1	Icb1-/Gc1-/Gc2-	Icb1-/Gc1-/Gc2- +Sp1
4.22	4.15	100	236.28	31.05	60.01	64.19	132.28	147.65	221.98	40.36	165.59	107.03	153.49	126.48	131.75	18.91	20.23
9.74	11.78	100	212.92	35.21	73.41	117.44	152.34	130.62	192.33	43.78	215.74	120.07	252.99	93.8	133.64	27.72	51.73
3.89	4.08	100	230.52	34.47	72.23	94.42	144.38	196.52	213.95	58.13	146.14	148.92	210.99	136.3	224.81	50.78	85.3
	AVG % (graphed below)																
5.95	6.67	100	226.57	33.577	68.55	92.017	143	158.26	209.42	47.423	175.82	125.34	205.82	118.86	163.4	32.47	52.42
	AVG DEV percentage error % (shown on graph)																
2.5267	3.4067	0	9.1022	1.6844	5.6933	18.551	7.1467	25.504	11.393	7.1378	26.611	15.72	34.889	16.707	40.94	12.207	21.92

7.4 Addition of Sp3 to topoisomerase IIα constructs.

Luciferase experiment 11

SERIES 1A.10 increasing Sp3 amounts

LUCIFERASE MAXIMA

pGL3B	wt -617	wt+0.25µg Sp3	wt+0.5µg Sp3	wt+1µg Sp3	ICB1mt	ICB1mt+0.25 µg Sp3	ICB1mt+0.5 µg Sp3	ICB1mt+1µg Sp3	BLANK
824	40267	8473	15455	6390	17201	7166	15449	2508	536
837	50203	9295	15399	7279	20876	7903	15498	2671	544
814	43904	9591	15857	7438	20311	7576	16269	2897	556
									545.33

BGAL READINGS

0.281	0.403	0.233	0.212	0.196	0.431	0.282	0.239	0.181	0.13
0.308	0.414	0.232	0.218	0.194	0.455	0.294	0.269	0.184	0.125
0.264	0.378	0.234	0.216	0.188	0.491	0.294	0.255	0.183	0.126
									0.127

CORRECTED LUCIFERASE VALUES

278.67	39722	7927.7	14910	5844.7	16656	6620.7	14904	1962.7
291.67	49658	8749.7	14854	6733.7	20331	7357.7	14953	2125.7
268.67	43359	9045.7	15312	6892.7	19766	7030.7	15724	2351.7

CORRECTED BGAL READINGS

0.154	0.276	0.106	0.085	0.069	0.304	0.155	0.112	0.054
0.181	0.287	0.105	0.091	0.067	0.328	0.167	0.142	0.057
0.137	0.251	0.107	0.089	0.061	0.364	0.167	0.128	0.056

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

1809.5	143919	74789	175408	84705	54788	42714	133068	36346
1611.4	173023	83330	163227	100502	61984	44058	105300	37292
1961.1	172744	84539	172041	112995	54301	42100	122841	41994

AVERAGE OF NORMALIZED LUCIFERASE VALUES

1794	163229	80886	170225	99401	57024	42957	120403	38544

NORMALIZED LUCIFERASE RELATIVE TO WT %

1.0991	100	49.554	104.29	60.897	34.935	26.317	73.764	23.614

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

121.72	12873	4064.6	4665.5	9797	3306.2	733.78	10069	2300

percentage error

6.785	7.8865	5.025	2.7408	9.856	5.7978	1.7082	8.3624	5.9672

Luciferase experiment 12: new buffers and reagents used.

SERIES 1A.20 increasing Sp3 amounts

LUCIFERASE MAXIMA

pGL3B	-617wt	WT+0.25µg Sp3	WT+0.5µg Sp3	WT+1µg Sp3	Icb1-	Icb1-+0.25µg Sp3	Icb1-+0.5µg Sp3	Icb1-+1µg Sp3	BLANK
868	12235	8722	4568	3439	10220	2956	4697	1772	797
877	9175	8874	4369	3905	8322	3436	3955	1881	819
884	8451	6840	4273	3997	6735	4107	4292	2287	802
									806

BGAL READINGS

0.243	0.228	0.186	0.16	0.156	0.371	0.221	0.198	0.149	0.121
0.271	0.24	0.188	0.159	0.163	0.376	0.224	0.207	0.156	0.124
0.263	0.226	0.191	0.156	0.173	0.31	0.237	0.202	0.156	0.123
									0.1227

CORRECTED LUCIFERASE VALUES

62	11429	7916	3762	2633	9414	2150	3891	966
71	8369	8068	3563	3099	7516	2630	3149	1075
78	7645	6034	3467	3191	5929	3301	3486	1481

CORRECTED BGAL READINGS

0.1203	0.1053	0.0633	0.0373	0.0333	0.2483	0.0983	0.0753	0.0263
0.1483	0.1173	0.0653	0.0363	0.0403	0.2533	0.1013	0.0843	0.0333
0.1403	0.1033	0.0683	0.0333	0.0503	0.1873	0.1143	0.0793	0.0333

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

515.24	108503	124989	100768	78990	37909	21864	51650	36684
478.65	71327	123490	98064	76835	29668	25954	37340	32250
555.82	73984	88302	104010	63397	31649	28872	43941	44430

AVERAGE OF NORMALIZED LUCIFERASE VALUES

516.57	84605	112261	100947	73074	33076	25563	44311	37788

NORMALIZED LUCIFERASE RELATIVE TO WT %

0.6106	100	132.69	119.32	86.371	39.094	30.215	52.374	44.664

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

26.167	15932	15972	2041.8	6451.1	3222.1	2466	4893.3	4428.1

percentage error

5.0656	18.832	14.228	2.0226	8.8282	9.7417	9.6465	11.043	11.718

Luciferase experiment 13: new buffers and reagents used.

LUCIFERASE MAXIMA

pGL3B	-617wt	WT+0.25µg Sp3	WT+0.5µg Sp3	WT+1µg Sp3	lcb1-	lcb1-+0.25µg Sp3	lcb1-+0.5µg Sp3	lcb1-+1µg Sp3	BLANK
605	58238	34523	59516	64042	21057	27342	34505	14225	207
635	64355	36412	64894	65152	17935	26067	33550	15963	184
650	57051	38486	63010	63581	16294	23145	27756	14648	199
									196.67

BGAL READINGS

0.375	0.418	0.261	0.256	0.227	0.449	0.318	0.258	0.213	0.128
0.376	0.451	0.282	0.264	0.238	0.444	0.327	0.253	0.229	0.124
0.342	0.399	0.281	0.281	0.224	0.398	0.317	0.247	0.209	0.12
									0.124

CORRECTED LUCIFERASE VALUES

408.33	58041	34326	59319	63845	20860	27145	34308	14028
438.33	64158	36215	64697	64955	17738	25870	33353	15766
453.33	56854	38289	62813	63384	16097	22948	27559	14451

CORRECTED BGAL READINGS

0.251	0.294	0.137	0.132	0.103	0.325	0.194	0.134	0.089
0.252	0.327	0.158	0.14	0.114	0.32	0.203	0.129	0.105
0.218	0.275	0.157	0.157	0.1	0.274	0.193	0.123	0.085

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

1626.8	197420	250557	449389	619858	64186	139924	256032	157622
1739.4	196203	229211	462124	569784	55432	127440	258553	150156
2079.5	206743	243881	400085	633843	58749	118903	224060	170016

AVERAGE OF NORMALIZED LUCIFERASE VALUES

1815.3	200122	241216	437199	607828	59456	128756	246215	159264

NORMALIZED LUCIFERASE RELATIVE TO WT %

0.9071	100	120.53	218.47	303.73	29.71	64.339	123.03	79.584

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

176.17	4414.2	8003.6	24743	25363	3153.2	7445.7	14770	7167.6

percentage error

9.7051	2.2057	3.318	5.6594	4.1727	5.3035	5.7828	5.9989	4.5004

LUCIFERASE MAXIMA

pGL3B	-617wt	WT+0.25µg Sp3	WT+0.5µg Sp3	WT+1µg Sp3	lcb1-	lcb1-+0.25µg Sp3	lcb1-+0.5µg Sp3	lcb1-+1µg Sp3	BLANK
347	40626	23149	12063	7151	35008	7162	4761	3074	176
326	46540	25256	16942	7483	24221	7835	4165	3488	177
330	38602	21736	13317	7902	33270	6784	4502	3849	175
									176

BGAL READINGS

0.497	0.476	0.258	0.18	0.162	0.468	0.222	0.185	0.145	0.121
0.526	0.545	0.266	0.21	0.172	0.379	0.243	0.196	0.146	0.12
0.477	0.507	0.268	0.188	0.169	0.451	0.237	0.192	0.151	0.117
									0.1193

CORRECTED LUCIFERASE VALUES

171	40450	22973	11887	6975	34832	6986	4585	2898
150	46364	25080	16766	7307	24045	7659	3989	3312
154	38426	21560	13141	7726	33094	6608	4326	3673

CORRECTED BGAL READINGS

0.3777	0.3567	0.1387	0.0607	0.0427	0.3487	0.1027	0.0657	0.0257
0.4067	0.4257	0.1467	0.0907	0.0527	0.2597	0.1237	0.0767	0.0267
0.3577	0.3877	0.1487	0.0687	0.0497	0.3317	0.1177	0.0727	0.0317

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

452.78	113411	165671	195940	163477	99901	68045	69822	112909
368.85	108921	171000	184919	138741	92599	61933	52030	124200
430.57	99121	145022	191374	155557	99781	56159	59532	115989

AVERAGE OF NORMALIZED LUCIFERASE VALUES

417.4	107151	160564	190744	152591	97427	62046	60462	117700

NORMALIZED LUCIFERASE RELATIVE TO WT %

0.3895	100	149.85	178.01	142.41	90.925	57.905	56.426	109.84

AVERAGE DEVIATION (AVEDEV) FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

32.365	5353.3	10361	3883.4	9233.9	3218.3	3999.9	6240.5	4333.7

percentage error

7.754	4.996	6.453	2.0359	6.0514	3.3033	6.4468	10.321	3.682

7.5 Changing pSV-β-galactosidase control vector amounts.

Luciferase experiment 15.

changing amounts of pSVBgal, with 0.5 μg –617 wt topoisomerase IIα

LUCIFERASE MAXIMA

pGL3B	wt–617 +1.5 μg pSVBgal	0.5 μg pSVBgal + wt–617	1.0 μg pSVBga + wt–617	1.5 μg pSVBgal + wt–617	BLANK
1650	9466	4300	3419	3911	1321
1688	11892	3912	2715	3945	1353
1662	8515	3614	3503	4410	1344
					1339.3

BGAL READINGS

0.163	0.162	0.13	0.142	0.16	0.122
0.165	0.176	0.129	0.128	0.153	0.116
0.164	0.16	0.126	0.138	0.159	0.124
					0.1207

CORRECTED LUCIFERASE VALUES

310.67	8126.7	2960.7	2079.7	2571.7
348.67	10553	2572.7	1375.7	2605.7
322.67	7175.7	2274.7	2163.7	3070.7

CORRECTED BGAL READINGS

0.0423	0.0413	0.0093	0.0213	0.0393
0.0443	0.0553	0.0083	0.0073	0.0323
0.0433	0.0393	0.0053	0.0173	0.0383

NORMALIZED LUCIFERASE VALUES (LUCIFERAS/BGAL)

7338.6	196613	317214	97484	65381
7864.7	190711	308720	187591	80588
7446.2	182432	426500	124827	80104

AVERAGE OF NORMALIZED LUCIFERASE VALUES

7549.8	189919	350811	136634	75358

NORMALIZED LUCIFERASE RELATIVE TO WT %

3.9753	100	184.72	71.943	39.679

AVEDEV FROM THE MEAN OF NORMALIZED LUCIFERASE VALUES

209.91	4991	50459	33971	6650.9

percentage error %

2.7803	2.6279	14.384	24.863	8.8258