

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

# Protein Interactions at the Human Topoisomerase II $\alpha$ Promoter

A thesis presented to Massey University in partial fulfilment of the requirement  
for the degree Doctor of Philosophy in Biochemistry.

**Natisha Magan**

2009

*If these written words are an indication of  
what I have learnt and who I have become.....  
then let a blank page represent  
what I have yet to learn and who I am yet to become.....*

# Acknowledgements

Many, many, many, many thanks to my wonderful supervisor Associate Professor Dr. Kathryn Stowell. Thank you for all your help, your guidance, your honesty, your support, your time and enthusiasm. I have truly enjoyed my time in your lab, I am fortunate to have had the chance to work with a group of talented, brilliant and beautiful people over the years, thank you for that opportunity.

I would like to thank Massey University for the financial support of a Doctoral scholarship. I would have never been able to take on such a task nor survive the extended period without additional financial support from The Todd Foundation, as well as the Edward and Isobel Kidson scholarship. It was greatly appreciated, thank you.

I am not going to list all the people that have helped me. There are just far too many of you. My friends, my family, my colleagues, all of you have influenced my life and helped me in so many different ways. I am incredibly lucky to have so many people who look out for me, who are there to support me, to guide me, who make me laugh and see a brighter day. I appreciate all of you, I truly, truly do.

Mum, who through it all, believes in me the most. It is because of you, I always try my best. It is because of you, I have learnt my strengths. It is because of you, I'm not afraid to fall...because I know, you will always catch me if I do. Thank you for just being mum. Finally, Dad....your Natsu has finally done it. I wish you were here to see this one.

# Abstract

Among women in the 45 to 64 age group, over half of the recorded deaths are from cancer, breast cancer being the most common. Just over 30% of all deaths in New Zealand women is caused by breast cancer. Treatment of cancer is difficult, not only due to the physiological and immunological similarities between a cancer cell and a normal cell, but also due to the high cardiotoxicity of many treatments, and also the problems related with the development of resistance. Approximately 40% of the cancer cells treated with the chemotherapy drug doxorubicin will become resistant to treatment. Drug efficacy is strongly associated with the proliferation status of a cell, as cancer cells divide rapidly, this can often be the defining factor between effective treatments or the development of resistance. Central to this proliferation status is an enzyme known as topoisomerase II $\alpha$ . This essential enzyme is expressed in all cells and is required to relieve the torsional stress in DNA that is created during normal cellular processes. A number of commonly used anti-cancer drugs have been found to target topoisomerase II $\alpha$  in cancer cells and significantly, during the development of drug resistance levels of topoisomerase II $\alpha$  enzyme have been found to be reduced in some cell lines and tumours. There are a number of factors that can modulate the amount of topoisomerase II $\alpha$  enzyme found in a cell, and one of the ways to understand this is to examine the regulation of the topoisomerase II $\alpha$  gene, most importantly the proteins that interact with the promoter region to direct transcription.

The human topoisomerase II $\alpha$  promoter has been found to be regulated by a number of transcription factors that can bind to their cognate sequences. The introduction of mutations within specific sequences of the topoisomerase II $\alpha$  promoter has enabled the identification of a key regulatory region within the promoter, a sequence of DNA that encompasses both the ICB1 and GC1 regulatory elements. Transcription factor NF-Y is found to bind to ICB1 element, whereas transcription factors Sp1 and Sp3 have been found to associate with the GC element. However this region of the promoter was also found to bind a fourth uncharacterised component. This research aims to further define the protein components that are found to bind to this important ICB1/GC1 regulatory region and distinguish the protein-protein and protein-DNA interactions that are important for the regulation of the human topoisomerase II $\alpha$  promoter.

# Abbreviations.

1°	Primary
2°	Secondary
APAF	Australian Proteome Analysis Facility
APS	Ammonium persulphate
AP-2	Activator protein 2
ATF	Activating transcription factor
ATP	Adenosine triphosphate
bp	Base Pair
β-gal	β-galactosidase
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
Buffer E	Equilibration buffer
CDE	Cell-cycle dependent element
CDTA	1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid
CEM	Human leukemic cell line
CENins	ICB1/GC1 altered to have an inserted sequence within the intervening region
CENdel	ICB1/GC1 altered to have a deletion within the intervening region
CENmt	ICB1/GC1 altered to have a mutation within the intervening region
ChIP	Chromatin immunoprecipitation
CHR	Cell-cycle gene homology region
COS-1	SV40 transformed African green monkey kidney cells
ddH <sub>2</sub> O	Double distilled water
DNA	Deoxyribose Nucleic Acid
DNA-PK	DNA-dependant protein kinase
DNA-PKcs	DNA-dependant protein kinase catalytic subunit
dNTP	Deoxyribonucleotide triphosphate
DSB	DNA double-strand break
DTT	dithiothreitol
EDTA	Ethylene diamine tetra-acetic acid
EMSA	Electrophoretic Mobility Shift Assay
GCF	GC rich sequence binding factor

GSB	Gel shift buffer
GZP1	GC box-binding protein-1
H209/VP	Human small-cell lung cancer cell line resistant to etoposide (VP-16)
HBT20	Human brain tumour cell line
HCl	Hydrochloric acid
HeLa	Human cervical cancer cell line
HEPES	N-[2-hydroxyethyl]piperazine-N'-[2-ethane sulfonic acid]
Her-2	Human epidermal growth factor 2
HL-60	Human promyelocytic leukemia cells
HL-60/MX2	Human promyelocytic leukemia cells resistant to mitoxantrone
HMGB	High mobility group B proteins
HRP	Horse-radish peroxidase
ICB	Inverted CCAAT box
ICBP90	Inverted CCAAT box binding protein of 90 kDa
IgG	Immunoglobulin G
IPTG	Isopropyl thiogalactosidase
KB	Human epidermoid cancer cell line
KCl	Potassium chloride
kDa	Kilo Dalton
Ku86	Ku autoantigen 86
Ku70	Ku autoantigen 70
Ku	heterodimeric protein consisting of Ku86 and Ku70
LB	Luria Bertani bacteriological media
MCF7	Human breast cancer cells
MCF12A	Normal human breast cells
mt	Mutant
MgAc	Magnesium acetate
Mw	Molecular Weight
NaCl	Sodium chloride
NIH3T3	Cells established from NIH swiss mouse embryo, highly contact inhibited
PARP	Poly (ADP-ribose) polymerase
PBS	Phosphate buffered saline solution
PCR	Polymerase chain reaction
pGL3B	pGL3B vector (for luciferase reporter gene assays)

PMSF	Phenylsulfonylmethyl fluoride
PSF	Polypyrimidine tract binding protein
PVDF	Polyvinylidene fluoride
NF-Y	Nuclear factor Y
mM	milli molar
$\mu$ L	Micro Litre
$\mu$ g	Micro Gram
NF90	Nuclear family 90
ng	nano gram
pRb	retinoblastoma protein
RNA pol II	RNA polymerase II
SDS-PAGE	Sodium dodecylsulphate polyacrylamide gel electrophoresis
Sp1	Specificity factor 1
SUMO	Small ubiquitin-like modifier
SV40	Simian virus 40
TBE	Tris borate EDTA
TBP	TATA-binding protein
TBST	Tris-buffered saline tween-20
TE	Tris-EDTA buffer
TEMED	Tetramethylethylenediamine
Topo2 $\alpha$	Topoisomerase II $\alpha$
Topo2 $\beta$	Topoisomerase II $\beta$
TPA	12-O-tetradecanoylphorbol-13-acetate
Tris	Tris (2-amino-2-hydromethyl-1,3-propandiol)
TSS	Transcription start site
U937	Human monocytic leukaemia cell line
UV	Ultra Violet
V	Volts
VM-26	Teniposide
VP-16	Etoposide
wt	Wild type
X-gal	5-bromo-4-chloro-3-indoyl- $\beta$ -D-galactosidase.
ZBP89	Zinc finger binding protein of 89 kDa

# List of Figures

Figure 1.1: Schematic representation of topoisomerase II enzyme activity. ....	3
Figure 1.2: Schematic representation of topoisomerase II-targeting poisons. ....	6
Figure 1.3: Multifactorial regulation of topoisomerase II $\alpha$ . ....	8
Figure 1.4: Schematic representation of topoisomerase II $\beta$ and PARP-1 mediated gene expression. .	11
Figure 1.5: Representation of the interactions that may occur at the human Topo2 $\alpha$ promoter. ....	19
Figure 1.6: DNA-binding proteins that bind to the ICB1/GC1 composite oligonucleotide. ....	21
Figure 3.1: DNA binding proteins are nuclear proteins. ....	44
Figure 3.2: Antibody supershifts to verify protein-DNA complexes. ....	46
Figure 3.3: Increasing amount of MCF7 extract added to different ICB1/GC1 oligonucleotides. ....	48
Figure 3.4: EMSA using increasing amount of MCF12A extract. ....	49
Figure 3.5: Investigating the intervening sequence for protein binding. ....	51
Figure 3.6: Complex three is the same using the ICB1 wt/CENins/GC1 wt oligonucleotide. ....	53
Figure 3.7: EMSA using a short GC1 oligonucleotide as a competitor. ....	56
Figure 3.8: EMSA using a ICB1 wt/CENdel/GC1 wt competitor. ....	57
Figure 3.9: Transient transfections using reporter gene constructs that have alterations in the -617 minimal topoisomerase II $\alpha$ promoter. ....	60
Figure 4.1: Analysis of proteins in fractions purified using heparin-sepharose column. ....	64
Figure 4.2: EMSA using different fractions from heparin-sepharose column. ....	65
Figure 4.3: EMSAs using F 0.35 M KCl and different topoisomerase II $\alpha$ oligonucleotides. ....	66
Figure 4.4: DNA-binding proteins purified using a biotinylated-ICB1 mt/GC1 wt oligonucleotide and the proteins are eluted using different salt concentrations. ....	71
Figure 4.5: Extracts purified using biotinylated-ICB1 mt/GC1 wt oligonucleotide with the addition of increasing amounts of competitors. ....	74
Figure 4.6: Using a competitor to generate biotinylated-ICB1 mt/GC1 wt purified extracts. ....	76
Figure 4.7: EMSA using extracts that were purified using a biotinylated-ICB1 mt/GC1 wt oligonucleotide in the presence of competitors. ....	78
Figure 4.8: Silver stained SDS-PAGE of electro-eluted protein samples from EMSA gels. ....	80
Figure 4.9: Immunoblot of electro-eluted proteins. ....	82
Figure 5.1: Schematic representation of a cross-linking assay. ....	86
Figure 5.2: ICB1 mt/GC1 wt BrdU oligonucleotides tested for protein binding. ....	89
Figure 5.3: Different BrdU oligonucleotides used in cross-linking assays. ....	90
Figure 5.4: Biotinylated-oligonucleotide purified extract contains DNA-binding proteins. ....	93
Figure 5.5: Confirming the identity of DNA-binding proteins found in extract purified using the biotinylated-ICB1 mt/GC1 wt oligonucleotide. ....	98

Figure 6.1: Transient transfection to show the effects of a PARP-1 inhibitor (PJ34) on wt -617topo2 $\alpha$ promoter activity.....	102
Figure 6.2: Transient transfection to show the effects of a PARP-1 inhibitor (PJ34) on CENins - 617topo2 $\alpha$ promoter activity.....	103
Figure 6.3: Transient transfection to show the effects of a PARP-1 inhibitor (PJ34) on -617topo2 $\alpha$ promoter activity with mutations in GC1 and GC2.....	105
Figure 6.4: Sp1 has increased affinity for DNA upon inhibition of ribosylation.....	107
Figure 6.5: Immunodetection of proteins present in electro-eluted extracts.....	109
Figure 6.6: Schematic representation of DNA-termini binding assay.....	114
Figure 6.7: Schematic representation for the generation of the ICB1mt/GC1wt concatemers.....	115
Figure 6.8: Immunodetection DNA-termini binding assay. ....	117
Figure 7.1: Summary of the different DNA-binding proteins found in each of the protein-DNA complexes that are seen to form using the topoisomerase II $\alpha$ ICB1 wt/GC1wt oligonucleotide. ....	120
Figure 7.2: Schematic representation of the DNA-binding proteins observed using EMSA and the corresponding promoter activities observed using transient transfections. ....	128
Figure 7.3: Increased cytotoxicity of chemotherapy in the presence of PARP-1 inhibitors.....	134
Figure 7.4: Schematic representation of fluorescence resonance energy transfer for detection of DNA-binding proteins. ....	140

## List of Tables

Table 2.1: Processing $\beta$ galactosidase data.....	40
Table 2.2: Processing Luciferase data. ....	40
Table 2.3: Generating normalised luciferase activities. ....	40
Table 3.1: Sequences of oligonucleotides used in EMSAs .....	50
Table 4.1: Summary of the molecular weights of purified DNA-binding proteins using different techniques.....	83
Table 5.1: Sequences of oligonucleotides used in EMSAs and cross-linking. ....	88
Table 5.2: Summary of the sizes of isolated DNA-binding proteins .....	95
Table 5.3: Summary of the results from protein analysis.....	96

# Table of Contents

<b>ABSTRACT</b> .....	<b>I</b>
<b>ABBREVIATIONS</b> .....	<b>II</b>
<b>LIST OF FIGURES</b> .....	<b>V</b>
<b>LIST OF TABLES</b> .....	<b>VI</b>
<b>1 INTRODUCTION</b> .....	<b>1</b>
<b>1.1 Cancer</b> .....	<b>1</b>
<b>1.2 Topoisomerases</b> .....	<b>2</b>
1.2.1 <i>Topoisomerase II<math>\alpha</math> and Her-2 genes</i> .....	4
1.2.2 <i>Topoisomerase II-targeting anti-cancer drugs</i> .....	5
1.2.3 <i>Drug resistance</i> .....	7
<b>1.3 Gene regulation</b> .....	<b>10</b>
1.3.1 <i>Chromatin remodelling for transcriptional regulation</i> .....	11
1.3.2 <i>Initiation of transcription</i> .....	12
<b>1.4 Human topoisomerase II<math>\alpha</math> promoter</b> .....	<b>12</b>
1.4.1 <i>GC elements</i> .....	14
1.4.2 <i>CCAAT sequences/ICB elements</i> .....	16
1.4.3 <i>Central regulatory region ICB1/GC1</i> .....	20
<b>1.5 Research outline</b> .....	<b>22</b>
<b>2 MATERIALS AND METHODS</b> .....	<b>23</b>
<b>2.1 Materials</b> .....	<b>23</b>
<b>2.2 Methods</b> .....	<b>25</b>
2.2.1 <i>DNA agarose gel electrophoresis</i> .....	25
2.2.2 <i>Restriction endonuclease digestion</i> .....	25
2.2.3 <i>Oligonucleotides</i> .....	25
2.2.4 <i>Preparing double-stranded oligonucleotides</i> .....	25
2.2.5 <i>Polymerase chain reactions (PCR)</i> .....	26
2.2.6 <i>DNA purification by gel electrophoresis</i> .....	26
2.2.7 <i>Phosphatase treatment of plasmids</i> .....	26
2.2.8 <i>Addition of 5' phosphate groups to oligonucleotides</i> .....	27

2.2.9 Ligation reactions. ....	27
2.2.10 Transformation of Escherichia coli (E.coli). ....	27
2.2.11 Purification of plasmids. ....	28
2.2.12 Tissue culture. ....	28
2.2.13 Treatment with PJ34. ....	29
2.2.14 Whole cell HeLa extracts. ....	29
2.2.15 HeLa cell nuclear extracts. ....	30
2.2.16 Quantification of total protein in extracts. ....	31
2.2.17 Sodium dodecylsulphate polyacrylamide gel electrophoresis. ....	31
2.2.18 Coomassie-blue stain for protein gels. ....	32
2.2.19 Silver staining for protein gels. ....	32
2.2.20 Immunodetection/Western Blotting. ....	32
2.2.21 Ion-exchange chromatography (heparin-sepharose column). ....	33
2.2.22 Oligonucleotide-affinity chromatography. ....	34
2.2.23 <sup>32</sup> P-labelling oligonucleotides. ....	35
2.2.24 Electrophoretic mobility shift assays (EMSA). ....	36
2.2.25 Electro-elution of proteins from EMSA gels. ....	37
2.2.26 BrdU cross-linking. ....	37
2.2.27 Transient transfections. ....	38
2.2.28 $\beta$ -galactosidase activity. ....	38
2.2.29 Luciferase assays. ....	39
2.2.30 Calculations for the normalised luciferase activities. ....	39
2.2.31 Database searches. ....	41
2.2.32 DNA sequencing reactions. ....	41
<b>3 DNA BINDING PROTEINS. ....</b>	<b>42</b>
3.1 Synopsis. ....	42
3.2 Protein complexes on the ICB1/GC1 composite element. ....	42
3.2.1 Whole cell and HeLa nuclear extracts. ....	43
3.2.1.1 Antibody Supershifts. ....	45
3.2.2 Protein complexes in other cell lines. ....	47
3.3 Investigating binding patterns of complex three. ....	49
3.3.1 Examining the intervening sequence between ICB1 and GC1. ....	50
3.3.2 Does ICB1wt/CENins/GC1wt bind complex three? ....	52
3.3.2.1 Using a mutated ICB1wt/CENins/GC1wt oligonucleotide. ....	53
3.3.2.2 Using a short GC1 competitor. ....	54
3.3.2.3 Using an ICB1wt/CENdel/GC1wt competitor. ....	56
3.4 Functional Significance of Complex three. ....	58
3.5 Chapter summary. ....	61

<b>4 PURIFICATION OF DNA BINDING PROTEINS.....</b>	<b>62</b>
<b>4.1 Synopsis.....</b>	<b>62</b>
<b>4.2 Ion-exchange Chromatography.....</b>	<b>62</b>
4.2.1 Purification of proteins from whole cell HeLa extract using heparin-sepharose column chromatography.....	62
4.2.2 Proteins purified using heparin-sepharose column.....	63
4.2.2.1 DNA-binding proteins in heparin-sepharose fractions.....	65
<b>4.3 Oligonucleotide-affinity Chromatography.....</b>	<b>68</b>
4.3.1 Purification of ICB1mt/GC1wt-binding proteins.....	70
4.3.1.1 Using a competitor to reduce non-specific protein-DNA interactions.....	72
4.3.1.2 EMSA with the extract purified using oligonucleotide-affinity chromatography.....	77
<b>4.4 Isolating protein from EMSA gels.....</b>	<b>79</b>
<b>4.5 Comparing the molecular weights of all the different purified DNA-binding proteins.....</b>	<b>82</b>
<b>4.6 Chapter summary.....</b>	<b>84</b>
<b>5 IDENTIFICATION OF DNA-BINDING PROTEINS.....</b>	<b>85</b>
<b>5.1 Synopsis.....</b>	<b>85</b>
<b>5.2 Using BrdU to cross-link DNA to protein.....</b>	<b>85</b>
5.2.1 Incorporating BrdU at different locations within the ICB1mt/GC1wt oligonucleotide.....	87
5.2.1.1 EMSA using BrdU oligonucleotides.....	88
5.2.1.2 Cross-linking using BrdU oligonucleotides.....	89
5.2.1.3 Cross-linking nuclear extract & biotinylated extract.....	92
<b>5.3 Analysis of potential DNA-binding proteins.....</b>	<b>94</b>
5.3.1 Results of protein analysis.....	95
<b>5.4 Confirming identity of DNA-binding proteins.....</b>	<b>96</b>
<b>5.5 Chapter summary.....</b>	<b>99</b>
<b>6 ANALYSIS OF DNA BINDING PROTEINS.....</b>	<b>100</b>
<b>6.1 Synopsis.....</b>	<b>100</b>
<b>6.2 Inhibiting Poly (ADP-ribose) polymerase activity.....</b>	<b>100</b>
6.2.1 PARP-1 & the topoisomerase II $\alpha$ promoter.....	101
6.2.2 PARP-1 ribosylates Sp1.....	106
<b>6.3 Identifying proteins in electro-eluted extracts.....</b>	<b>108</b>
<b>6.4 Ku proteins.....</b>	<b>113</b>
6.4.1 Are the proteins binding to the ends of DNA?.....	114
6.4.1.1 Generating a concatemered ICB1mt/GC1wt oligonucleotide.....	115
6.4.1.2 DNA-termini binding assays.....	116
<b>6.5 Chapter summary.....</b>	<b>118</b>

<b>7 DISCUSSION AND FUTURE WORK</b> .....	<b>119</b>
<b>7.1 Synopsis</b> .....	<b>119</b>
<b>7.2 Protein-ICB1/GC1 interactions</b> .....	<b>120</b>
7.2.1 <i>Complex 1/Sp1 interacting with the GC1 element</i> .....	121
7.2.2 <i>Complex 2/NF-Y interacting with the ICB1 element</i> .....	122
7.2.3 <i>Complex 4/Sp3 interacting with the GC1 element</i> .....	123
7.2.4 <i>Complex 3/proteins and the intervening sequence</i> .....	125
<b>7.3 Ku Proteins</b> .....	<b>129</b>
<b>7.4 PARP-1</b> .....	<b>131</b>
7.4.1 <i>PARP-1 inhibitors</i> .....	133
<b>7.5 PSF</b> .....	<b>136</b>
<b>7.6 Improving methods &amp; further research</b> .....	<b>137</b>
7.6.1 <i>Affinity chromatography</i> .....	137
7.6.2 <i>Electro-eluting proteins</i> .....	138
7.6.3 <i>Immunoprecipitation</i> .....	138
7.6.4 <i>Other methods to examine protein-DNA interactions</i> .....	139
7.6.5 <i>Functional assays</i> .....	140
7.6.5.1 <i>ChIP assays</i> .....	141
7.6.5.2 <i>Protein-protein interactions</i> .....	141
<b>7.7 Significance of topoisomerase II<math>\alpha</math> expression</b> .....	<b>142</b>
<b>7.8 Conclusion</b> .....	<b>142</b>
<b>REFERENCES</b> .....	<b>144</b>
<b>APPENDIX 1</b> .....	<b>169</b>
<b>APPENDIX 2</b> .....	<b>170</b>
<b>APPENDIX 3</b> .....	<b>172</b>
<b>APPENDIX 4</b> .....	<b>174</b>

# 1 Introduction.

## 1.1 Cancer.

Cancer is commonly described as a mass of abnormal cells that have the ability to divide uncontrollably. Normal cells grow and divide a limited number of times in the body, where the regulation of cell division is an intricate process requiring many cellular check-points, many different enzymes and a variety of cellular processes. Generally these mechanisms are tightly regulated and dividing cells require enzymes to maintain the constant changes in DNA topology required for these processes. These enzymes are known as topoisomerases. All of the mechanisms that cause cells to divide uncontrollably are not fully understood, in some cases a series of genetic mutations involving tumour suppressors or oncogenes can turn a normal cell into a cancerous cell. Often growth factors bind to receptors on the cells surface and trigger unregulated cell division. In some cases, cancer cells can have an excess of these cellular receptors, others produce their own growth hormones to stimulate their own growth causing them to be able to divide autonomously. These abnormal cells have the ability to aggregate within tissues, form tumours and even induce the development of blood vessels to provide nutrients and stimulate further growth, a process known as angiogenesis. In more severe cases, the cancer cells can spread to nearby tissues and throughout the bloodstream into the lymphatic system.

There are over 100 different types of cancer where traditional prognostic factors include age, tumour size, tissue type, lymph node status, histological type, grade, human epidermal growth factor receptor 2 (Her-2) status and hormone receptor status. For women in the 25-44 age group the leading cause of death is cancer, with 1 in 3 of these deaths being from breast cancer ([www.wrongdiagnosis.com/b/breast\\_cancer/stats.htm](http://www.wrongdiagnosis.com/b/breast_cancer/stats.htm)). New Zealand has the third highest death rate from breast cancer in the western world ([www.stats.govt.nz/analytical-reports/women-in-nz/cancer-leading-cause-death.htm](http://www.stats.govt.nz/analytical-reports/women-in-nz/cancer-leading-cause-death.htm)). Treatments for cancers can range from surgical removal, radiation therapy, monoclonal antibodies (to target cellular receptors), chemotherapy and even vaccines against viruses that have been shown to increase the chance of specific cancers. Amongst the most commonly used chemotherapy agents are drugs that target the topoisomerases.

## 1.2 Topoisomerases.

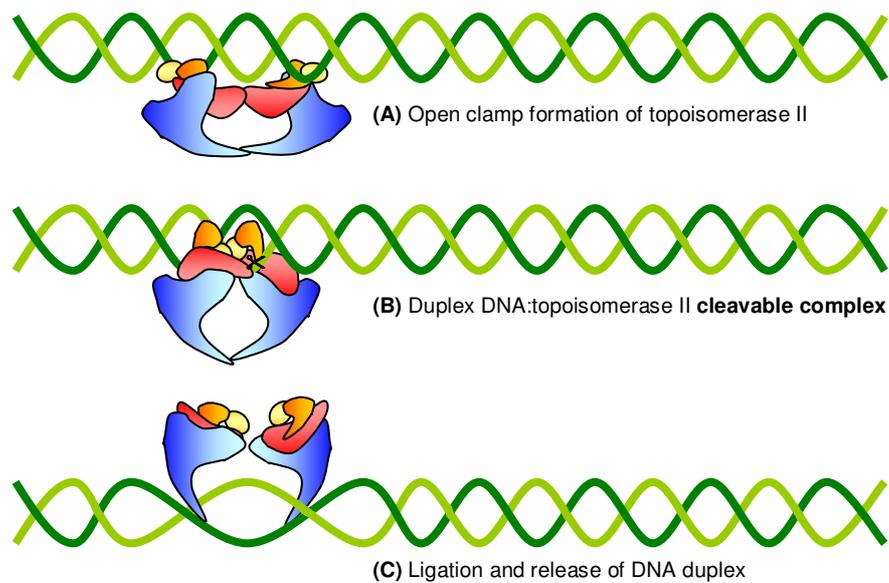
Normal cellular processes such as DNA replication, transcription, recombination, DNA repair, DNA compaction and chromatin remodelling generate torsional stress in chromosomes. This stress must either be accommodated by conformational changes in DNA structure (e.g supercoils) or dissipated to correct DNA topology. These topological changes occur based on the individual requirements of the cell; if DNA can not be modified high levels of torsional stress can halt transcription machinery and deform chromosome structure with serious cellular consequences. DNA topoisomerases are essential nuclear enzymes that can alter the topological states of DNA, and are capable of rotating DNA about a DNA break. Humans have four types of topoisomerase I, II, III and V (Kellner *et al.*, 2000) which are classified based on their catalytic activity. The ability of topoisomerases to create breaks in the backbone of DNA is fundamental to their catalytic functions.

The topoisomerase II enzyme is the only topoisomerase essential for cell viability. This enzyme forms a heart-shaped homodimer to bind duplex DNA (Wei *et al.*, 2005) and functions by introducing a transient double-stranded break in DNA, enabling the safe passage of a second duplex DNA strand through the break (figure 1.1). This action relieves two supercoils during each catalytic cycle of the enzyme and requires ATP hydrolysis. During this process, the topoisomerase II enzyme remains bound to the DNA stabilising and protecting the DNA break that it creates (cleavable complex), as this complex represents a state of intense fragility for the DNA molecule. Double-stranded DNA breaks are considered amongst the most cytotoxic DNA lesions that can occur within a cell. In response to double-stranded breaks cells must react, arrest the cell cycle to repair the lesion, or in cases where damage is too extensive initiate apoptosis and cell death. If no further alterations in DNA topology are required, the topoisomerase II enzyme religates the double-stranded break and releases itself from DNA (Roca, 2009).

Humans have two different topoisomerase II isoforms, topoisomerase II $\alpha$  (Topo2 $\alpha$ ) and topoisomerase II $\beta$  (Topo2 $\beta$ ). They are similar in protein structure and have almost identical catalytic properties *in vitro* (Austin and Marsh, 1998; Jenkins *et al.*, 1992). They have the ability to resolve problems that occur during cellular DNA replication such as over-winding, under-winding and catenation of DNA. In cells that lack normal topoisomerase II $\alpha$  activity, it has been shown that topoisomerase II $\beta$  can partially substitute for the topoisomerase II $\alpha$

activity suggesting that a functional compatibility may exist between the two isoforms (Sakaguchi and Kikuchi, 2004). However, each of these isoforms are localised differently during the cell cycle and each are differentially regulated. Topoisomerase II $\alpha$  levels fluctuate during the cell cycle whereas topoisomerase II $\beta$  levels remain constant (Zandvliet *et al.*, 1996; Adachi *et al.*, 1997; Mirski *et al.*, 2003) indicating that these enzymes may also have different biological roles *in vivo*. Topoisomerase II $\alpha$  is important for cell proliferation as it is essential during replication however topoisomerase II $\beta$  is not (Grue *et al.*, 1998) and it is thought that the divergence in the C-terminal domains between the two enzymes may attribute to these differences (Linka *et al.*, 2007).

### Topoisomerase II alters DNA topology



**Figure 1.1: Schematic representation of topoisomerase II enzyme activity.**

Through a series of conformational changes requiring ATP (not shown in figure) the topoisomerase II enzyme can bind and alter DNA topology. **(A)** Before interacting with DNA, the topoisomerase II homodimer has an open clamp formation. **(B)** The topoisomerase II enzyme binds to a DNA duplex (G-segment) where binding of a second DNA duplex (T-segment) stimulates the introduction of a transient double-stranded break (✂) in the first DNA duplex (G-segment). The enzyme remains bound to the broken DNA (cleavable complex) whilst the intact duplex (T-segment) is passed through the break. **(C)** If no further alterations are required, the double-stranded break is religated and both DNA duplexes are released by the enzyme.

### 1.2.1 Topoisomerase II $\alpha$ and Her-2 genes.

Breast cancer is considered to be relatively sensitive to chemotherapy compared to other solid tumours (Nagasaki and Miki, 2006). Adjuvant therapies can incorporate a number of different drugs (refer to appendix 1 for summary) to minimise the cardiotoxicity of chemotherapy, decrease the development of resistance and increase overall efficiency of the treatment regime. The most effective combination regimens for treatment of breast cancers include anthracyclines (epirubicin or doxorubicin) which are topoisomerase II poisons. However as yet there are no definitive clinically applied predictive markers for a response to chemotherapy. Therefore most patients who receive chemotherapy generally receive the same treatment, even though de novo drug resistance can result in treatment failure in many patients. There is a current expansion in research examining the possibility of predictive biomarkers for cancer types, in attempts to deliver more responsive treatments based on the profile of the individual tumours. In solid tumours, such as breast, gastric and ovarian cancers, the predominant genetic mechanism for oncogene activation is gene amplification (Järvinen and Liu, 2003). The genes encoding Her-2 and topoisomerase II $\alpha$  have both been implicated in this process.

Her-2 protein (or receptor) is commonly found on the surface of breast epithelial cells to help maintain normal growth. Tumour samples from various sources have verified that gastric, ovarian and breast cancers can have extra copies of the Her-2 gene which is thought to be associated with more aggressive or resistant cancers (reviewed in Ross *et al.*, 2003; Gurel *et al.*, 1999; Varis *et al.*, 2002). Her-2 over-expression occurs in about 30-50% of patients with breast cancer and this can occur by gene amplification or by increased promoter activity as single copy genes can still over-express Her-2 protein (Hurst, 2001; Slamon *et al.*, 1989). The human chromosomal locus 17q12-q21 harbours a number of genes found to be amplified in cancers; of most importance are the Her-2 and topoisomerase II $\alpha$  genes which are found adjacent to each other in this region. Co-amplification of the Her-2 and topoisomerase II $\alpha$  gene has been described in breast cancers, however this does not always occur. It has also been shown that these genes can amplify as separate amplicons (Järvinen *et al.*, 1996) and often both genes have separate copy numbers within the same tumour (Bhargava *et al.*, 2005). It does appear as though there is a slightly higher possibility that when the topoisomerase II $\alpha$  gene is amplified the Her-2 gene is also amplified, but often when Her-2 is amplified topoisomerase II $\alpha$  is not (reviewed in Mano *et al.*, 2007). Herceptin

(trastuzumab) is an anti-cancer drug that is specifically designed for the treatment of Her-2 positive breast cancers; it is based on a targeted biologic therapy of monoclonal antibodies which directly block the Her-2 receptors on the cells surface. Recently Her-2 positive (Her-2+) tumours have also been found to have increased sensitivity to the topoisomerase II inhibitors, anthracyclines (Dressler *et al.*, 2005; Pritchard *et al.*, 2006), which is useful as the benefits of herceptin are in addition to a patients' normal treatment regimen.

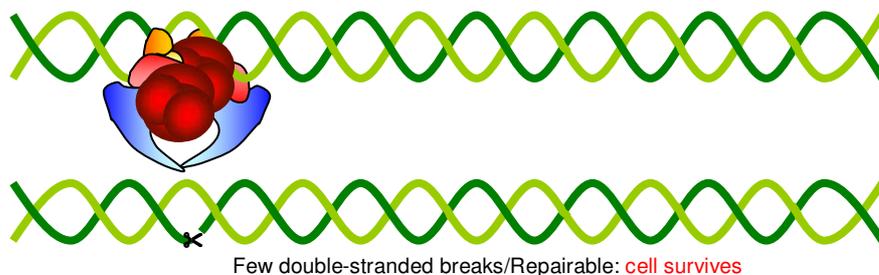
### 1.2.2 Topoisomerase II-targeting anti-cancer drugs.

In highly proliferating cells where DNA exists in a dynamic state constantly replicating and being processed for transcription, higher levels of topoisomerase II $\alpha$  can be found (Turley *et al.*, 1997). Many studies have found a relationship between topoisomerase II $\alpha$  expression and the proliferative state of a tumour, with higher levels of topoisomerase II $\alpha$  protein being found in proliferating tumour types (reviewed in Isaacs *et al.*, 1998; O'Malley *et al.*, 2009). Topoisomerases have now been established as a major target for chemotherapeutic attack, where a majority of the common anticancer drugs appear to target topoisomerase II $\alpha$  *in vivo*, however it is also thought that some drugs could preferentially target the  $\beta$  isoform (Dereuddre *et al.*, 1997; Gao *et al.*, 1999; Errington *et al.*, 2004). Topoisomerase II-targeting drugs can fall into one of two categories, topoisomerase II inhibitors or topoisomerase II poisons (reviewed in Fortune and Osheroff, 2000).

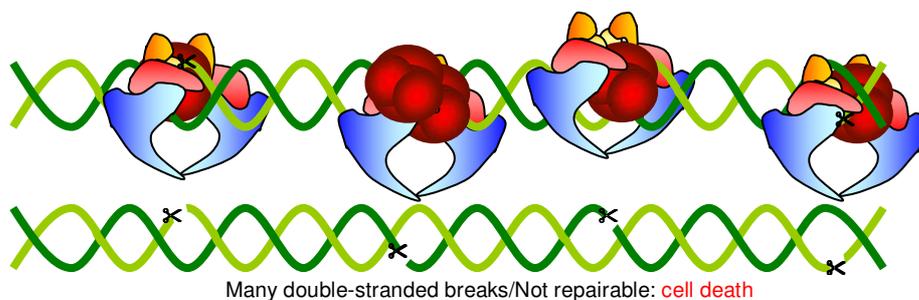
Commonly used topoisomerase II poisons are amsacrine, etoposide, teniposide and the anthracyclines; doxorubicin and epirubicin. Topoisomerase II-targeting poisons act by converting normal cellular topoisomerase II into a cellular toxin by stabilising the topoisomerase II-DNA cleavable complex, this is the conformation where DNA is held in a broken state (figure 1.2). Normal cellular machinery converts these DNA breaks into permanent double-stranded lesions when they try to traverse this stabilised ternary complex, this causes irreparable DNA damage and cell death. Drug-induced stabilisation of this cleavable-complex is thought to be reversible if the drug is removed from the cells before the cellular replication and transcription machinery has a chance to act (reviewed in Baguley and Ferguson, 1998). Some drugs act by inhibiting the DNA ligation step and others act by enhancing the ability of topoisomerase II to create double-stranded breaks. Significantly, drug efficacy for these types of topoisomerase II-targeting drugs has been found to be

directly correlated with the amount of active topoisomerase II $\alpha$  protein found in tumours and cancerous cell lines (Cardoso *et al.*, 2004; Tanner *et al.*, 2006; Uesaka *et al.*, 2007).

**(A) Topoisomerase II-targeting poisons create double-stranded DNA breaks.**



**(B) Topoisomerase II-targeting poisons in cancer cells create many double-stranded DNA breaks.**



**Figure 1.2: Schematic representation of topoisomerase II-targeting poisons.**

Topoisomerase II poisons act by stabilising the cleavable complex of DNA:topoisomerase II within a cell. Each topoisomerase II enzyme in the homodimer is thought to contain a binding site for the topoisomerase II poison, when drug concentrations are higher both sites can be occupied. (A) In normal cells where levels of topoisomerase II $\alpha$  are low, these drug stabilised complexes generate few double-stranded DNA breaks which can be repaired by the cell normal cellular repair machinery and the cell survives. However in (B) cancerous cells that are highly proliferating, high levels of topoisomerase II $\alpha$  protein can be found. The number of drug stabilised DNA-topoisomerase complexes are greater, consequently the number of double-stranded breaks accumulated by the cell are much greater. These can not be repaired and lead to cell death.

Commonly used topoisomerase II inhibitors are daunorubicin, aclarubicin, bisdioxopiperazines and novobiocin. Drugs that inhibit topoisomerase II function without targeting the cleavable complexes are termed catalytic inhibitors (reviewed in Larsen *et al.*, 2003). Topoisomerase II-targeting inhibitors do not increase the number of covalent topoisomerase II-DNA cleavable complexes but act rather by inhibiting overall normal topoisomerase II activity (refer to appendix 1 for summary). Cell death by these types of

drugs is generally thought to occur due to an inability of the cell to correct torsional stress and related topological problems that may arise in DNA.

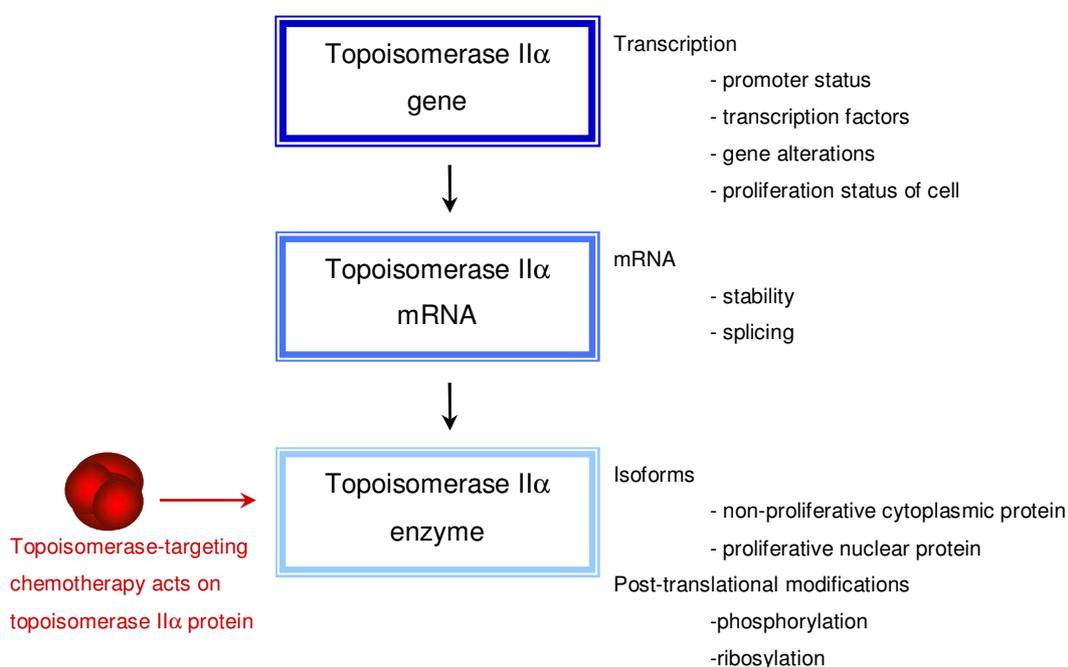
### 1.2.3 Drug resistance.

Chemotherapy is one of the most potent forms of treatment for cancer, however drug toxicity and the development of resistance to these drugs is a major clinical problem; approximately 40% of breast cancers become resistant to treatment with topoisomerase II poison doxorubicin.

The development of drug resistance can be either intrinsic or acquired. One of the most common mechanisms for intrinsic resistance is due to the altered expression of a transmembrane glycoprotein called Pgp or Mdr1 (Grant *et al.*, 1994; reviewed in Larsen and Skladanowski, 1998), which is a product of the multi-drug resistance gene (MDR1). Many topoisomerase-targeting drugs are natural products for which cellular efflux mechanisms already exist within the cell. This classical multi-drug resistance (MDR) phenotype increases drug efflux from the cell thereby decreasing drug efficiency (Bredel, 2001). In addition to Pgp there are a number of other proteins which are commonly over-expressed in MDR cells; such as MRP (multi-drug resistance associated proteins), MXR (mitoxantrone resistance associated) and LRP (lung resistance associated protein). All of these proteins belong to a class of proteins known as the ABC transporter proteins that control traffic of biological molecules across cell membranes. In line with acquired resistance, studies have demonstrated that down-regulation of the topoisomerase II $\alpha$  gene is correlated with the development of drug-resistance (Saxena *et al.*, 2004). Often the drug resistant cell lines used in studies have been created by repeated exposure to anti-cancer drugs, where the surviving cells are propagated and used for further analysis (Kubo *et al.*, 1995; Takano *et al.*, 1999).

Cellular events that have been implicated in the development of drug-resistance include point mutations, amplification or deletions in the topoisomerase II $\alpha$  gene (some are summarised in appendix 2). Alternative splicing or gene deletions have been shown to contribute to reports of two different mRNA species which result in a catalytically active cytoplasmic topoisomerase II $\alpha$  variants, such as those found in etoposide resistant lung (H209/VP) cancer cells, H209/VP (Mirski *et al.*, 2000) or the leukaemia cell line resistant to mitoxantrone, HL-60/MX2 (Harker *et al.*, 1995). These cytoplasmic variants of the topoisomerase II $\alpha$  protein

are thought to be associated with increased drug-resistance, due to a concomitant reduction in the amount of wild-type nuclear topoisomerase II $\alpha$  protein. Changes in transcription factors as well as post-translational modifications of the topoisomerase II $\alpha$  protein can also regulate gene expression (reviewed in Robert and Larsen, 1998; reviewed in Isaacs *et al.*, 1998). As anti-cancer drugs target the topoisomerase II $\alpha$  protein *in vivo*, often the process of developing resistance to these drugs is multifactorial (figure 1.3). Each of the major regulatory processes may contribute to the final outcome of topoisomerase II $\alpha$  protein activity and interaction with anti-cancer drugs.



**Figure 1.3: Multifactorial regulation of topoisomerase II $\alpha$ .**

Changes in the amount and activity of topoisomerase II $\alpha$  protein *in vivo* can be attributed to a number of different regulatory mechanisms. Gene regulation via the topoisomerase II $\alpha$  promoter can be a direct measure of the amount of topoisomerase II $\alpha$  protein found in cells.

In experiments using human brain tumour cells (HBT20) resistant to etoposide, it was thought that resistance could be overcome by expressing exogenous topoisomerase II $\alpha$  protein in these cells. Surprisingly, only a short term increase in topoisomerase II $\alpha$  expression was observed with a concomitant transient increase in sensitivity to etoposide. Even though the levels of topoisomerase II $\alpha$  protein went back to normal within 24 hours,

this demonstrated the possibility that drug sensitivity can be altered by increasing the amount of topoisomerase II $\alpha$  protein present in a cell (Asano *et al.*, 1996). The levels of topoisomerase II $\alpha$  protein fluctuate within the cell-cycle, and it is thought that the mechanisms that regulate this fluctuation may also influence the development of drug resistance. Both ubiquitination and proteasome-dependent protein degradation have been implicated in this (Ogiso *et al.*, 2000; Salmena *et al.*, 2001). There are mixed reports on whether phosphorylation of the topoisomerase II $\alpha$  enzyme can affect catalytic activity and drug resistance (Takano *et al.*, 1991; Kimura *et al.*, 1996; Chen and Beck, 1995). Most of the phosphorylation sites reside within the C-terminal region of the protein and it is more likely that phosphorylation affects localisation of the topoisomerase II $\alpha$  protein (Ishida *et al.*, 1996), thereby affecting its activity. It has been shown that trafficking of topoisomerase II $\alpha$  to the cytoplasm from the nucleus can generate resistance to etoposide in human leukaemia cells (Valkov *et al.*, 2000). Ribosylation of the topoisomerase II protein has been shown to decrease catalytic activity *in vitro* (Darby *et al.*, 1985). The topoisomerase II $\alpha$  protein can be ribosylated in HeLa cells (Scovassi *et al.*, 1993) but the *in vivo* effects of this are still undetermined.

As a direct consequence of increased promoter activity or greater mRNA stability within a cell, the levels of mRNA transcript can be increased. In various resistant cell lines it has been demonstrated that a reduction in topoisomerase II $\alpha$  protein is correlated with a decrease in the amount of topoisomerase II $\alpha$  mRNA which is thought to be due to altered transcriptional activity and not altered mRNA stability (Kubo *et al.*, 1995; Asano *et al.*, 1996; Harker *et al.*, 1995). However, experiments using a system that inhibits endogenous mRNA expression in human monocytic leukaemia cells (U937) demonstrated that inhibition of topoisomerase II $\alpha$  mRNA was sufficient to confer drug resistance to etoposide in these cells (Towatari *et al.*, 1998).

Alterations in the topoisomerase II $\alpha$  gene and promoter status are at the core of topoisomerase II $\alpha$  protein production. Currently there are a few studies that examine gene amplification, deletions and mutations in the topoisomerase II $\alpha$  gene for use as a prognostic marker to predict sensitivity to anthracycline treatment in breast cancers; however this is with limited success. Apart from some studies having small sample sizes, some studies also fail to recognise the progression by which a gene alteration can relate to active nuclear

topoisomerase II $\alpha$  protein. Although it has been suggested that over-expression of the topoisomerase II $\alpha$  protein rarely occurs without topoisomerase II $\alpha$  gene amplification (Bhargava *et al.*, 2005), this can not be the only mechanism that elevates the amount of topoisomerase II $\alpha$  protein in tumours. This is because it has been found in many cases that topoisomerase II $\alpha$  gene amplification is not always related to an increase in topoisomerase II $\alpha$  protein (Mueller *et al.*, 2004; reviewed in Mano *et al.*, 2007) and often over-expression is independent of gene amplification (Varis *et al.*, 2002; Durbecq *et al.*, 2004A), suggesting that gene regulation and protein integrity are influential.

Different stress conditions can lead to a decrease in the expression of the topoisomerase II $\alpha$  gene, for example in the case of confluence-arrested cells (Isaacs *et al.*, 1996) or the presence of oncogene suppressor protein p53 which is activated in response to DNA breakage (Millau *et al.*, 2009). Glucose-regulated stresses, such as hypoxia and nutrient deprivation often lead to decreased amounts of topoisomerase II $\alpha$  protein (Yun *et al.*, 1995). All these conditions represent situations where cellular stresses could trigger cellular responses which perhaps alter topoisomerase II $\alpha$  expression due to changes in transcription factors that modulate the amount of topoisomerase II $\alpha$  protein rather than a gene alteration or protein degradation process.

Understanding drug-resistance will also shed light on drug-sensitivity and why cancer cells are so prolific. Key to this is recognising that topoisomerase II $\alpha$  expression is multifactorial. It should be considered that perhaps examining the proliferation status of a tumour, quantification of the amount of active topoisomerase II $\alpha$  enzyme, determining the intracellular localisation of the topoisomerase II $\alpha$  protein, and defining the transcription factor status or promoter status may be more clinically relevant than the current method of prognosis using only gene alterations and amplifications.

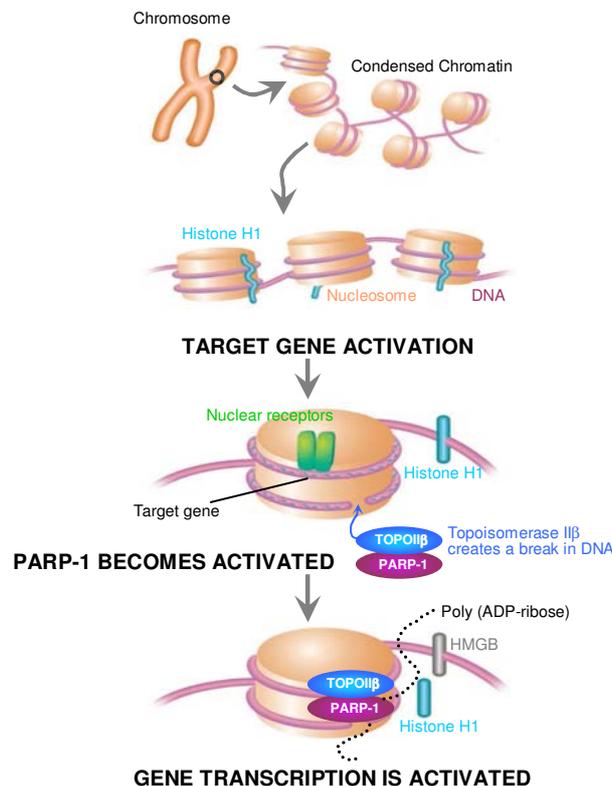
### **1.3 Gene regulation.**

Epigenetic mechanisms (DNA methylation and histone modifications) which operate in advance of actual gene transcription events provide a signature to specific regions of the genome, allowing the cell to recognise the genes targeted for active transcription. DNA is organised into extremely compact nuclear structures of higher order called chromatin. This

condensation naturally restricts accessibility of proteins including transcription factors and polymerases to the DNA.

### 1.3.1 Chromatin remodelling for transcriptional regulation.

Chromatin remodelling, describes the process by which chromatin that is tightly wound around nucleosomes is disrupted to allow accessibility to the DNA (reviewed in Workman, 2006). Structural transitions in chromatin also depend on the core domains of histone proteins which are responsible for the compaction of DNA and for establishing protein-protein interactions. An interesting recent development that highlights a role for the topoisomerase II $\beta$  enzyme involves a novel DNA strand-break model for transcriptional activation (Ju *et al.*, 2006).



**Figure 1.4: Schematic representation of topoisomerase II $\beta$  and PARP-1 mediated gene expression.** DNA strand-break dependent gene activation involves a complex that contains both topoisomerase II $\beta$  and PARP-1. Upon binding of the nuclear receptor to the target gene, a site-directed double-stranded break is introduced by topoisomerase II $\beta$ . This in turn activates PARP-1 which is thought to ribosylate chromatin-associated proteins, such as histone H1. H1 is a known repressor for transcription, exchange of H1 for high mobility group B (HMGB) proteins appears to be dependent on the activation of PARP-1 during the gene-specific transcriptional activation. This process is thought to allow chromatin to relax, which is required to initiate gene expression. Figure taken from Haince *et al.*, 2006.

In response to a nuclear signal topoisomerase II $\beta$  is thought to introduce a nucleosome-specific site-directed double-stranded break in the target gene, which in turn activates an enzyme known for its role in double-stranded break repair (PARP-1). The activation of PARP-1 may have multiple implications, but one of them in this instance is thought to bring about a nucleosome-specific exchange of histone H1 for high mobility group B (HMGB) proteins (summarised in figure 1.4), which in turn promotes a loosening effect on chromatin (reviewed in Haince *et al.*, 2006; Lis and Kraus, 2006; Ju and Rosenfeld, 2006). HMGB proteins are chromatin-associated nucleoproteins that bend DNA and facilitate the binding of various transcription factors to their DNA targets, both HMGB1 and HMGB2 have recently been shown to interact with transcription factor NF-Y and modulate transcription from the topoisomerase II $\alpha$  promoter (Stros *et al.*, 2009).

### **1.3.2 Initiation of transcription.**

Eukaryotic transcription is mediated by a series of orchestrated protein-protein and protein-DNA interactions that enable transcription by RNA polymerase II. Although transcription requires RNA polymerase II, this enzyme itself is required to interact with a number of different proteins to be active. Regulatory sequences of DNA located at the 5' end of a gene cooperate in the control of transcription; basal or core elements support the formation of pre-initiation complex, direct the binding of RNA polymerase II, start transcription at the correct site, as well as *cis*-elements that are required to interact with sequence-specific DNA binding proteins (transcription factors). These bind to either RNA polymerase II or other transcription factors which ultimately leads to different degrees of activation or repression of transcriptional initiation.

## **1.4 Human topoisomerase II $\alpha$ promoter.**

The minimal promoter of the human topoisomerase II $\alpha$  gene is well established and is found to encompass a region 617bp upstream of the transcription start site (Isaacs *et al.*, 1996). It contains five inverted CCAAT boxes (ICB1-5), two GC elements (GC1 and GC2), consensus sequences for binding of an activator protein-2 (AP2) as well as an activating transcription factor (ATF), putative binding sequences for both Myb and Myc/Max, as well as a CDE/CHR element (figure 1.5). A number of different transcription factors have been shown to bind to these regulatory elements within the topoisomerase II $\alpha$  promoter (reviewed in

Bronner *et al.*, 2002), where changes in the transcription factors that bind are considered to be able to direct changes in the amount of topoisomerase II $\alpha$  protein found in a cell in different cellular contexts, such as during the development of drug resistance and a change in the proliferation status of a cell.

Very little research has been carried out on the role of the ATF binding motif in the human topoisomerase II $\alpha$  promoter. Earlier work demonstrated that this element was not important in confluent-arrested cells (Isaacs *et al.*, 1996). Since then it has been shown that protein(s) are capable of specifically binding to the ATF sequence isolated from the topoisomerase II $\alpha$  promoter, and that this interaction may be involved in cell-cycle regulation of HeLa cells (Son *et al.*, 2002) or down-regulation of topoisomerase II $\alpha$  when HL-60 cells differentiate (Lim *et al.*, 1998). These investigators however did not identify these proteins or sufficiently demonstrate their significance in both accounts of these protein-DNA interaction. The AP2 site is a potential binding site for the activator protein-2, however little is known about the regulatory consequences of this in regards to the topoisomerase II $\alpha$  promoter.

Exogenous expression of the proto-oncogene product (c-Myb protein) has been shown to activate the human topoisomerase II $\alpha$  promoter in HL-60 leukemia cells in a dose-dependent manner. By specifically mutating the Myb binding site this effect was abolished and also greatly reduced basal topoisomerase II $\alpha$  activity. However none of these trends were observed when experiments were repeated using HeLa cells, suggesting that c-Myb regulation is cell-type specific (Brandt *et al.*, 1997).

The Myc/Max protein complex binds to E-box sequences with a consensus 5'-CACGTG-3' or variations thereof (Luscher and Larsson, 1999); there is a partial consensus sequence within the human topoisomerase II $\alpha$  promoter located at position -21 (MYC/MAX). Tumour cells generally exhibit higher levels of Myc protein compared to their normal counterparts (Wu *et al.*, 2007). However, it is not known if topoisomerase II $\alpha$  promoter activity is influenced by the binding of the Myc/Max heterodimer, or if it is even capable of binding to the putative consensus site. It has been demonstrated that even low affinity sites become occupied with Myc/Max when levels of Myc are elevated (Fernandez *et al.*, 2003). The roles of Myc proteins in promoting proliferation and cell cycle progression are well established (Adhikary *et al.*, 2005), therefore it is possible that these proteins may regulate the

topoisomerase II $\alpha$  gene, but this still needs to be established. Little is known about the CDE/CHR element except that it is thought to be involved in cell cycle regulation of genes (reviewed in Isaacs *et al.*, 1998).

### 1.4.1 GC elements.

GC elements are amongst the most frequent regulatory element to be found within promoters, quite often present in multiple copies. GC elements have been shown to bind and recruit a number of proteins such as the Sp family of transcription factors (Sp), GC box-binding protein-1 (GZP1), zinc finger binding protein of 89 kDa (ZBP-89) and GC rich sequence binding factor (GCF) (reviewed in Hapgood *et al.*, 2001). Transcription factors Sp1 and Sp3 are known to bind to GC elements and are considered to be important for the recruitment and formation of complexes with a variety of proteins, this in turn can cause region-specific changes in chromatin structure (histone acetylation) or enable RNA polymerase II recruitment within promoters (Davie *et al.*, 2008). In promoters that have more than one GC element it is quite common for each GC element to have opposing functions; one may act as a repressive element and the other as an activator element (reviewed in Li *et al.*, 2004). The human topoisomerase II $\alpha$  promoter has two GC elements, one close to the transcription start site (GC1) and the other further upstream (GC2). Transient transfections carried out using the minimal -617topo2 $\alpha$  promoter in luciferase reporter gene assays have demonstrated that mutations within the GC1 and/or GC2 elements carry little consequence on promoter activity, though GC2 may have the tendency to play a repressive role (Magan N., MSc Thesis, 2002). Other studies have also shown that a mutation in GC1 does not greatly affect promoter activity in proliferating cells; however this same mutation in confluent cells appears to increase promoter activity (Isaacs *et al.*, 1996). Therefore the GC elements should not be disregarded as non-functional. Changes in the transcription factors that are found to bind to these elements have also been shown to vary during the development of drug resistance (Kubo *et al.*, 1995; Allen *et al.*, 2004).

Sp1 is considered to up-regulate topoisomerase II $\alpha$  expression in proliferating cells. In a previous study it was demonstrated that by exogenously increasing the amount of Sp1 in HeLa cells, it is possible to get an increase in topoisomerase II $\alpha$  promoter activity, demonstrating that Sp1 can activate topoisomerase II $\alpha$  gene expression (Magan *et al.*, 2003). Sp1 is capable of self-association which is thought to be able to bring proximal and distal GC

elements together, inducing synergistic activation of promoters (Su *et al.*, 1991). Using a system that captures interacting regions of DNA *in vivo*, it has now also been shown that both GC1 and GC2 can interact *in vivo* which is reminiscent of DNA-bending to bring both elements close together (Williams *et al.*, 2007). It has long been considered that both Sp1 and Sp3 proteins were capable of interacting, as both these proteins do co-immunoprecipitate, are often found to super-shift with opposing antibodies in electrophoretic mobility shift (EMSA) assays (Magan N., MSc Thesis, 2002) and are thought to interact through their zinc finger DNA-binding domains. However this observation may be due to a common co-factor that may bind both Sp1 and/or Sp3, and it is now considered that Sp1 and Sp3 do not directly interact with each other (Reviewed in Davie *et al.*, 2008). Therefore it may be possible that other proteins are also found to associate with the GC elements *in vivo*. Sp1 levels have been found to be elevated in breast carcinomas (Zannetti *et al.*, 2000), human thyroid tumours (Chiefari *et al.*, 2002) and human gastric cancers (Jiang *et al.*, 2004).

Within the topoisomerase II $\alpha$  promoter both GC elements are known to bind transcription factors Sp1 and Sp3 *in vivo*. However *in vitro* DNA-binding assays have demonstrated that in order for Sp3 to bind to the GC1 element additional flanking sequences are required, suggesting that the region between ICB1 and GC1 may be recruiting other factors that enhance the binding of Sp3 to GC1 (Magan *et al.*, 2003). Sp1 and Sp3 are very similar in protein structure, although they can support different roles in regulation of the human topoisomerase II $\alpha$  promoter. Sp1 has been shown on several accounts to be able to activate topoisomerase II $\alpha$  promoter activity, in both cervical (HeLa) and breast cancer (MDA-MB-231) cell lines (Magan *et al.*, 2003; Allen *et al.*, 2004). The Sp3 protein is known to have an inhibitory domain, a highly charged region of about 13 amino acids located immediately in front of the DNA-binding (zinc finger) domain (Dennig *et al.*, 1996). This is thought to be able to allow this protein to support both repressor and activator roles. Sp3 is a transcriptional repressor or activator, depending on promoter context and cell type. Sp3-mediated gene activation is thought to be repressed when this protein is SUMO-modified (Ross *et al.*, 2002; Valin and Gill, 2007). It has been shown that both Sp1 and Sp3 are capable of competing for binding to the same GC sequences, therefore each protein may be able to displace each other when binding to their cognate sequence. Significantly, Sp3 has been shown to be able to repress Sp1-mediated gene activation in a dose-dependent manner by associating with the GC elements within the human topoisomerase II $\alpha$  promoter (Williams *et al.*, 2007). Using *in*

*in vivo* chromatin immunoprecipitation (ChIP) assays this study also was able to demonstrate that the GC1 element appears to bind Sp1 and Sp3 with the same affinity, however the GC2 element preferentially binds Sp1. After exposure to doxorubicin, the relative levels of occupancy of these transcription factors changed such that both GC1 and GC2 were preferentially binding Sp3, this was most likely due to an increase in the amount of this transcription factor present in these cells which carried a concomitant reduction in the amount of topoisomerase II $\alpha$  protein.

#### 1.4.2 CCAAT sequences/ICB elements.

The CCAAT sequences are often located upstream of the transcription start site, and can bind a number of factors, such as nuclear factor-Y (NF-Y), inverted CCAAT box binding protein of 90 kDa (ICBP90), Y-box binding protein (YB-1) and CCAAT/enhancer binding proteins (C/EBPs). Of these, both NF-Y and ICBP90 have been found to bind to the topoisomerase II $\alpha$  promoter so far. The ICB elements are composed of an inverted CCAAT motif that is generally flanked by short sequences which are also required for the recruitment and binding of transcription factor NF-Y. NF-Y is composed of three small subunits NF-YA, NF-YB and NF-YC, where NF-YB and NF-YC tightly associate with each other which is prerequisite to an association with NF-YA and DNA (Liang and Maity, 1998). NF-Y has been extensively studied for its role in transcriptional regulation (Reviewed in Mantovani, 1999). It is capable of interacting with general proteins of the transcription machinery and other transcription factors as well as having a direct role in chromatin reorganisation (Motta *et al.*, 1999). Specifically it has been shown that NF-Y binding to the ICB elements can disrupt nucleosomal organisation of the topoisomerase II $\alpha$  promoter (Cousty *et al.*, 2001). NF-Y also has a functional interaction with Sp1 bound at nearby GC elements (Roder *et al.*, 1999; Magan *et al.*, 2003). *In vitro* binding assays (EMSA) have shown that ICB1, ICB2, ICB3, and ICB4 essentially bind the same components and each of these ICB elements can compete for binding of this specific factor (Herzog *et al.*, 1997). This has now been identified as NF-Y. ICB5 was shown to bind different protein components (which have yet to be identified), consequently this element is unable to compete for binding of the factors that bind to ICB1-4 (Herzog *et al.*, 1997). Gene regulation and the development of drug resistance is a complicated process, often requiring alterations in transcription factors and multiple changes in protein-protein interactions. One particular study has shown that topoisomerase II $\alpha$  expression can be up-regulated when resistance to bisdioxopiperazines is acquired in human

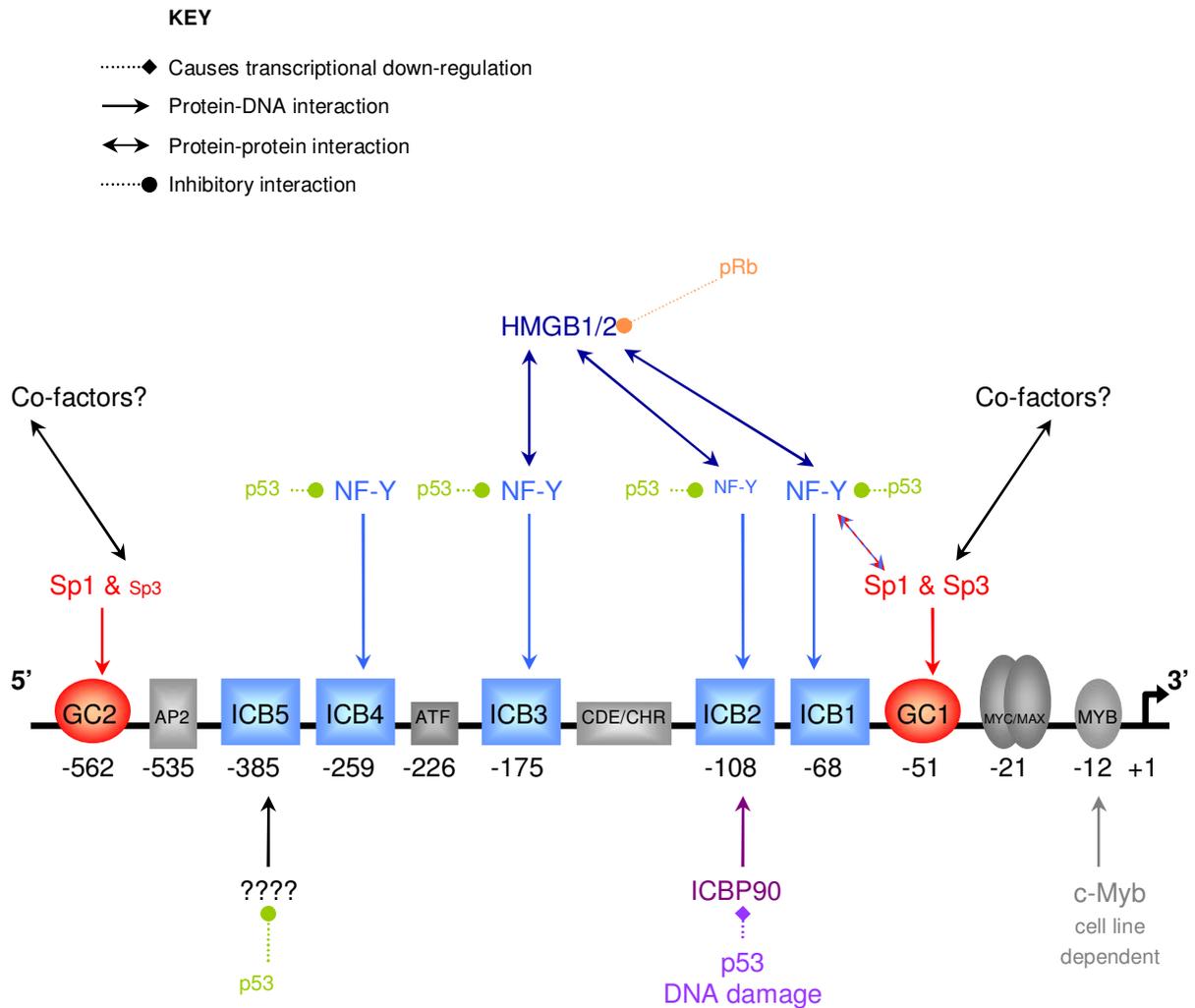
leukemic cells lines, and this was found to be related to changes in proteins that bound to the ICB3 element (Morgan and Beck, 2001). Subsequent research has shown that ICB2 binds NF-Y with a lower affinity than the other ICB elements. ICB2 is also implicated with observations that the topoisomerase II $\alpha$  promoter is down-regulated in confluence arrested cells (Isaacs *et al.*, 1996; Joshi *et al.*, 2003). Interestingly, another factor ICBP90 has also been shown to bind to ICB2. ICBP90 is a nuclear protein that is capable of activating transcription of the human topoisomerase II $\alpha$  promoter by specifically binding to the ICB2 element (Hopfner *et al.*, 2000). Using *in vitro* DNA-binding assays it has been demonstrated that phosphorylated ICBP90 can bind more efficiently to ICB2, which in turn corresponds with an increase in both topoisomerase II $\alpha$  mRNA and protein in COS-1 cells (Trotzler *et al.*, 2004). Normally ICBP90 protein is cell-cycle regulated. It is however, found at consistently higher levels in proliferating cell lines and tumours in conjunction with elevated levels of topoisomerase II $\alpha$  protein (Mousli *et al.*, 2003). When ICBP90 protein is transiently over-expressed, it even has the ability to overcome cell contact inhibition of cell growth in primary cultures of human lung fibroblasts due to a forced mechanism of topoisomerase II $\alpha$  expression (Hopfner *et al.*, 2002). An interesting feature of ICBP90 is that it exhibits anti-apoptotic properties, ICBP90 down-regulation has been observed as a prerequisite for apoptosis (Abbadly *et al.*, 2005). Recently it has been demonstrated that ICBP90 is down-regulated in response to DNA damage induced by adriamycin/doxorubicin (Arima *et al.*, 2004).

Tumour suppressor protein p53 is normally expressed at low levels within cells, however in response to cellular stresses such as DNA damage, hypoxia, nutrient deprivation and even damage caused by various topoisomerase II targeting drugs, cellular levels of p53 protein are seen to rise (reviewed in Millau *et al.*, 2009). This in turn induces a cell-signal cascade mediating a range of cellular responses, one of them being apoptosis and cell death. Loss of p53 activity increases genome instability and susceptibility to malignant transformations. Approximately 50% of human cancers are reported to have mutated or inactive p53 protein (Lacroix *et al.*, 2006). It has been shown that wild-type p53 can inhibit the binding of NF-Y to ICB1, ICB2, ICB3, and ICB4 as well as the protein that is found to bind to ICB5 within the topoisomerase II $\alpha$  promoter, resulting in a marked decrease in topoisomerase II $\alpha$  promoter activity, mRNA and protein (Wang *et al.*, 1997A; Joshi *et al.*, 2003). p53 is also a potential transcription factor capable of sequence-specific binding to regions within

promoters (El-Deiry *et al.*, 1992). Even though there are no exact p53-consensus sequences within the topoisomerase II $\alpha$  promoter it is reported that it may associate with a region between -90 to -30 bp to activate gene transcription in the absence of any ICB elements (Wang *et al.*, 1997A). This also recognises that there are still regulatory elements within the topoisomerase II $\alpha$  promoter that have not yet been defined and demonstrates that the human topoisomerase II $\alpha$  promoter may support both p53-dependent and p53-independent gene regulation.

Using experiments that over-expressed HMGB proteins, both HMGB1 and HMGB2 were shown to enhance binding of NF-Y to ICB1, ICB2 and ICB3, with a concomitant increase in topoisomerase II $\alpha$  promoter activity and protein (Stros *et al.*, 2009). However, this effect was abolished in the presence of retinoblastoma susceptibility protein (pRb). pRb represents a situation similar to that of p53, where many tumours have mutant variants of this protein and may serve to deregulate normal cellular functions in cancer cells. Deletion constructs of the minimal topoisomerase II $\alpha$  promoter have revealed that ICB5 exhibits the least influence on topoisomerase II $\alpha$  gene regulation (Hochhauser *et al.*, 1992), whereas ICB1 is regarded as the most important ICB element in this regard as mutations in this element generate the greatest reduction in topoisomerase II $\alpha$  promoter activity (Magan *et al.*, 2003).

The different binding affinities and protein-protein interactions that are observed at each of the regulatory elements is of significance. Each ICB element is stipulated to have slightly different functional responsibilities *in vivo* which will affect the interactions that occur at nearby elements within the promoter, representing the complexity involved in the regulation of gene expression. A summary of these interactions can be seen in figure 1.5, but they only represent a small understanding of these mechanisms. Novel transcription factors and regulatory regions are being identified constantly. Further defining the interactions that occur to regulate the human topoisomerase II $\alpha$  promoter can be improved by examining the two most proximal elements, ICB1 and GC1.



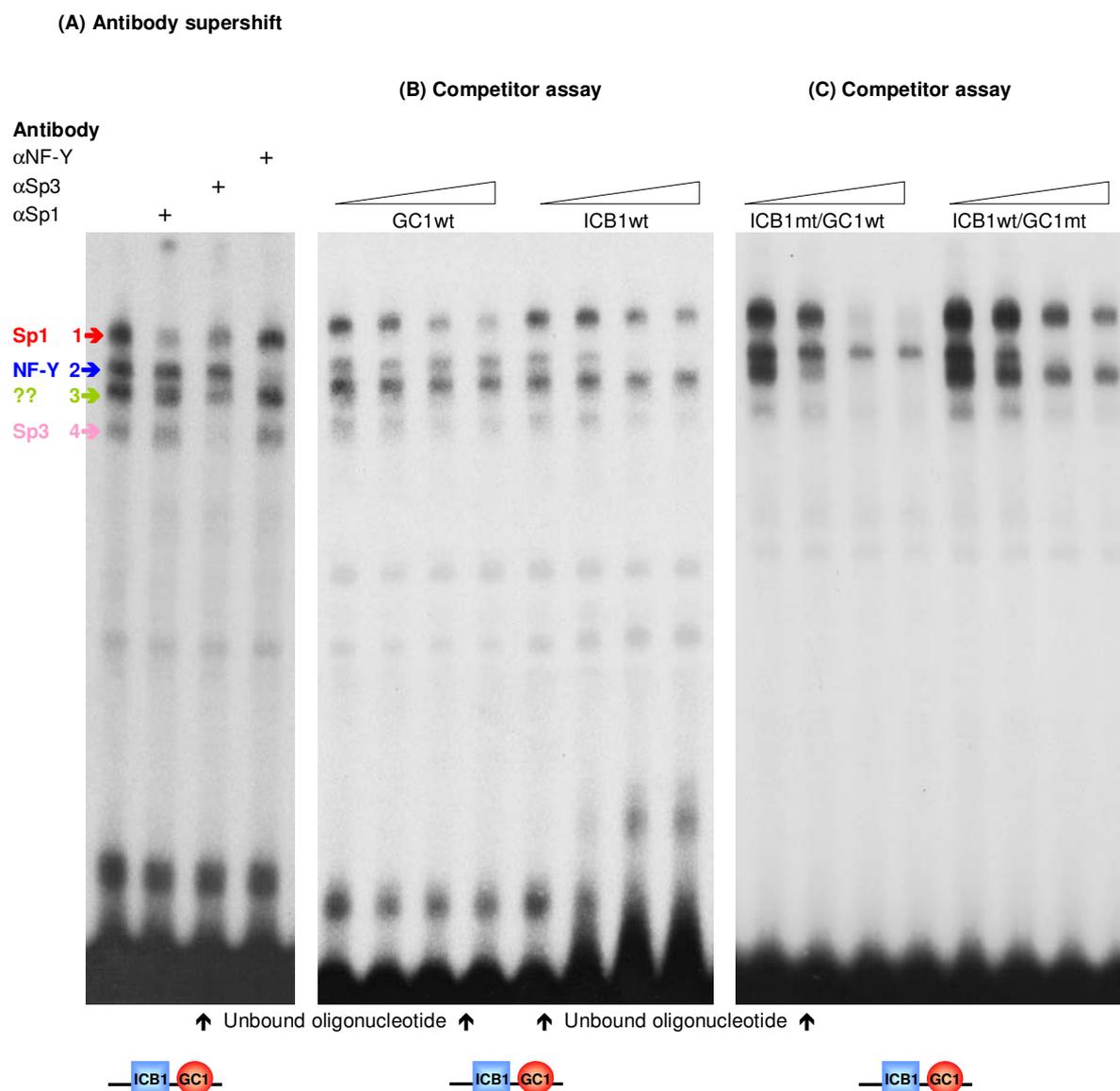
**Figure 1.5: Representation of the interactions that may occur at the human Topo2 $\alpha$  promoter.**

The human topoisomerase II $\alpha$  minimal promoter (Hochhauser *et al.*, 1992; Isaacs *et al.*, 1996) is contained within 617 bp of the 5' regulatory sequence. Positions of the different regulatory elements are mapped and are numbered with respect to the major transcription start site (+1). There are 5 inverted CCAAT boxes (ICB1-5), two GC elements (GC1 and GC2) and consensus sequences for AP2, ATF, CDE/CHR, MYC/MAX and MYB. A summary of the protein-protein and protein-DNA interactions that are thought to regulate gene expression are displayed above. Sp1 activates transcription whereas Sp3 represses Sp1-mediated activation. Sp3 has a lower affinity for GC2 than GC1 (Williams *et al.*, 2007). Co-factors may be required to enhance the binding of Sp protein to GC elements (Davie *et al.*, 2008). Binding of ICBP90 to ICB2 enhances topoisomerase II $\alpha$  promoter activity and can circumvent cell cycle arrest; however p53 can inhibit the binding of ICBP90 (Hopfner *et al.*, 2002). ICB2 has lower affinity for NF-Y than the other ICB elements (Joshi *et al.*, 2003). NF-Y binding to the ICB elements enhances promoter activity, whereas p53 inhibits NF-Y from binding and HMGB1 can enhance NF-Y binding. pRb inhibits the HMGB1-enhanced binding of NF-Y (Stros *et al.*, 2009). c-Myb can activate topoisomerase II $\alpha$  expression in lymphoid and myeloid cell lines (Brandt *et al.*, 1997).

### 1.4.3 Central regulatory region ICB1/GC1.

Earlier work carried out using the minimal -617 topoisomerase II $\alpha$  promoter has demonstrated that both the ICB1 and GC1 elements carry the greatest ability to activate transcription (Takano *et al.*, 1999; Magan *et al.*, 2003). A mutation in the ICB1 element has been observed to reduce basal promoter activity by approximately 70% and enhancement of transcription factor Sp1 has been shown to be able to activate transcription in a dose-dependent manner (Magan N., MSc Thesis, 2002). ICB1 sits adjacent to GC1 within the topoisomerase II $\alpha$  promoter and the two elements are thought to work synergistically to regulate transcription through the recruitment of basal transcription factors to their respective elements, via mechanisms that allow chromatin reorganisation, DNA bending and an interaction with RNA polymerase II.

Using an oligonucleotide encompassing both ICB1 and GC1, *in vitro* binding assays (figure 1.6) demonstrated that four distinct protein-DNA complexes were capable of forming on this segment of DNA (Magan N., MSc Thesis, 2002). Transcription factors Sp1, Sp3 and NF-Y were confirmed to be binding, specific to their cognate sequence, however one of the protein-DNA complexes remained uncharacterised (figure 1.6, complex 3 green arrow). This uncharacterised protein(s) exhibited properties of being a GC-associating factor, however it was unable to bind efficiently on a shorter GC1 oligonucleotide, nor did it require an intact ICB1 element to bind. Collectively this suggests that proteins may be recruited to the intervening sequence between both ICB1 and GC1 and perhaps represents a novel regulatory sequence within the topoisomerase II $\alpha$  promoter. There has been a reported regulatory sequence implicated with the binding of p53 protein located -30 to -90 relative to the transcription start site (Wang *et al.*, 1997A) which actually overlaps with the intervening sequence between ICB1 and GC1 (-53 to -65). However p53 is functionally inactivated by human papilloma virus-E6 in HeLa cells (Arima *et al.*, 2004), therefore it is unlikely that the uncharacterised protein in complex 3 is p53. With the apparent affinity for GC elements, it could be that the uncharacterised protein is another GC binding protein. Literature suggests that it is unlikely to be another member of the Sp transcription factor family, Sp2 has affinity for a slightly different sequence (Moorefield *et al.*, 2004), Sp4 is only expressed in tissues of the central nervous system and Sp7 is expressed in bone derived cells (Milona *et al.*, 2003). Further work will need to be carried out to define the properties of this uncharacterised protein-DNA complex.



**Figure 1.6: DNA-binding proteins that bind to the ICB1/GC1 composite oligonucleotide.** EMSA reactions contained a total of 5  $\mu$ g of whole cell HeLa extracts, along with approximately 0.5 ng of ICB1wt/GC1wt  $^{32}$ P-oligonucleotide (A) Antibody supershifts contained a total of 1  $\mu$ g of Sp1, Sp3 or NF-Y antibody in each EMSA reactions. (B) and (C) Competitor assays (refer to table 3.1 for oligonucleotide sequences) contained an increasing amount of unlabelled competitor (0, 5, 50 and 100 ng) in each EMSA reaction, as indicated in figure above. Each gel was dried using DE-81 paper and exposed to X-ray film for (A) 15 hours (B) 10 hours or (C) 20 hours at  $-80^{\circ}\text{C}$ . The coloured arrows define the specific protein-DNA complexes found to associate with the ICB1 and GC1 elements. The free/unbound oligonucleotide is located at the bottom of the gel. Figure replicated from Magan N., MSc Thesis, 2002.

## 1.5 Research outline.

Gene expression outcomes are determined in part by a number of factors, the composition of the promoter, organism, cell type, chromatin organisation, transcription factors, and post-translational modification of each of the associated factors. The objectives of this research was to examine the protein-DNA interactions that occur at the human topoisomerase II $\alpha$  promoter, specifically within the region that encompasses both the ICB1 and GC1 elements, which maps to the region -36 to -80 upstream of the transcription start site.

Objectives:



Further delineate the binding parameters required to efficiently induce binding of complex 3 to the ICB1/GC1 composite sequence.



Determine if complex 3 is functionally required for the transcriptional regulation of human topoisomerase II $\alpha$ .



Purify and identify DNA-binding proteins that associate with the ICB1/GC1 composite sequence.



Carry out functional assays based on results of DNA-binding experiments.

Understanding the mechanisms by which the topoisomerase II $\alpha$  promoter can be activated may, in part, provide insight into the molecular mechanisms that are involved in the development of drug resistance. The fundamental problem with developing therapies for cancer treatment is that the cancer cell and normal cell are remarkably similar both immunologically and physiologically. This similarity is what makes it difficult to develop therapies that can selectively destroy cancer cells without also causing damage to normal healthy cells. The main difference between these types of cells is the proliferation status. Key to this proliferation is the amount of topoisomerase II $\alpha$  enzyme which is essential for cell survival.

# 2 Materials and Methods.

## 2.1 Materials.

Product	Company	Distributor
Na <sub>2</sub> HPO <sub>4</sub> , KH <sub>2</sub> PO <sub>4</sub>	Ajax chemicals, Auburn, Australia	Ajax Finechem NZ Ltd
ATP, poly dl.dC Heparin-sepharose 6 fast-flow affinity resin γ <sup>32</sup> P [ATP] (Redivue [gamma <sup>32</sup> P] ATP)	Amersham/GE healthcare	GE Healthcare, Auckland, NZ
Centricon protein concentrator columns	Amicon/Millipore, MA, USA	BioLab Ltd, Auckland, NZ
ICBP90 (612264) Monoclonal antibody NF90 (612154) anti-DRBP76 antibody	BD transduction laboratories	BD Biosciences, Becton Dickson Ltd, Auckland, NZ
Nonidet P40	BDH Laboratory supplies, England	BioLab Ltd, Auckland, NZ
Precision plus dual colour protein standards Quantum <sup>®</sup> Prep plasmid miniprep kit Bio-Rad protein assay reagent	Bio-Rad, Hercules, CA, USA	Bio-Rad Laboratories Pty Ltd, Life Science Group, Auckland, NZ
PJ34 Protein G /protein A agarose beads	Calbiochem, EMD chemicals, Darmstadt, Germany	Merck Ltd, Palmerston North, NZ
CONCERT <sup>™</sup> gel purification kit OPTI-MEM <sup>®</sup> I, Penicillin/streptomycin mix CONCERT <sup>™</sup> high purity plasmid maxiprep kit Thermosensitive calf alkaline phosphatase TrypLE <sup>™</sup> Express, Fetal calf serum	GIBCOBRL/Invitrogen Corporation	Invitrogen NZ Ltd, Penrose, Auckland, NZ
12 well tissue culture plates 15mm tissue culture plates 96 well clear flat bottom microplates 96 well white/opaque flat bottom microplates	Griener bio-one, Germany	Raylabs, Glendene, Auckland, NZ
X-ray film (X-omat BT)	Eastman Kodak, NY, USA	Radiographic supplies, NZ
Dynabeads M-280 streptavidin beads	Dynal Biotech, Norway	Invitrogen life technologies, Penrose, NZ
T4 DNA ligase 1 kb plus DNA ladder CONCERT <sup>™</sup> gel purification kit	Invitrogen, Carlsbad, USA	Invitrogen, Penrose, NZ
Coomassie brilliant blue, boric acid , EDTA, Sucrose, Tris-acetate, CaCl <sub>2</sub> , PMSF Bromophenol blue, MgCl <sub>2</sub> , LB broth mix KCl, NaCl, HEPES, dithiothreitol, MgAc 40% acrylamide-bis solution Sodium thiosulphate, triton X-100, glycerol 37% formaldehyde solution	Merck, Darmstadt, Germany	Merck Ltd, Palmerston North, NZ
MagPrep streptavidin beads D-Tube <sup>™</sup> dialysers (midi)	Novagen, Madison, USA	Merck Ltd, Palmerston North, NZ
T75 tissue culture flasks	Nunc <sup>™</sup> , Roskilde, Denmark	In Vitro technologies, Mt Wellington, NZ

Product	Company	Distributor
PGEMT easy vector Luciferase assay reagents pCMV SPORT-βGal Thermosensitive alkaline phosphatase	Promega Corporation, Madison, USA	In Vitro technologies, Mt Wellington, NZ
QIAquick PCR purification kit	QIAGEN, Hilden, Germany	BioLab Ltd, Auckland, NZ
Taq DNA polymerase, agarose, dNTP mix T4 DNA ligase, FuGENE®6 BM chemiluminescence blotting substrate Protease inhibitor cocktail (complete mini EDTA-free) Positively charged nylon membrane	Roche Applied Science, Mannheim, Germany	Roche Diagnostics Auckland, NZ Ltd
Protein A /G agarose beads Ku86 (H-300) polyclonal antibody Ku70 (H-308) polyclonal antibody NF-YA (H-209) polyclonal antibody PARP-1 (F-2) monoclonal antibody PSF (H-80) polyclonal antibody Sp1 (PEP 2) polyclonal antibody Sp3 (D-20) polyclonal antibody	Santa Cruz Biotechnology, Santa Cruz, USA	Global Science, Lower Hutt, NZ
Oligonucleotides	SigmaGenosys	Sigma life sciences, Sigma- Aldrich, Auckland, NZ Ltd
EDTA, ficoll, tween 20, ethidium bromide IPTG, ampicillin, X-Gal, DMSO, SDS Ammonium persulfate, TEMED, tetracycline HRP-conjugated anti-mouse or anti-rabbit IgG Tubulin (104K4800) monoclonal antibody	Sigma chemical company, St Louis, USA	Sigma-Aldrich, Auckland, NZ Ltd
<i>E. coli</i> XL-1 blue	Stratagene, CA, USA	Pacific Laboratory Products, Auckland, NZ
OptiKinase™ T4 polynucleotide kinase	usb, Cleveland, Ohio, USA	GE Healthcare, Auckland, NZ
Vivaspin ultrafiltration devices	Vivascience	Sigma-Aldrich NZ Ltd, Auckland, NZ
DE-81 chromatography paper 3MM Chr chromatography paper	Whatman International, Maidstone, England	Global Science, Lower Hutt, NZ

All chemicals or reagents were of analytical grade or better. Original HeLa cells were a gift from Dr. Rachel Page (Department of Biochemistry, University of Cardiff, Wales). The original -617topo2α pGL3B plasmids were created by Samuel M<sup>c</sup>Lenachan (BSc (Hons) project, Massey University, 1998) and Agnieszka Szremska (BSc (Hons) project, Massey University, 2000). The pGL3B plasmids used in transient transfections containing the -617topo2α minimal promoter with alterations CENins, CENmt and CENdel were created by Kelly Senior (Senior K., MSc Thesis, Massey University, 2006) and the GC1-/GC2-construct was created by Natisha Magan (Magan N., MSc Thesis, Massey University, 2002). Research was carried out under GMO approval code GMO98/MU/53.

## **2.2 Methods.**

### **2.2.1 DNA agarose gel electrophoresis.**

Generally, DNA agarose gel electrophoresis was carried out using 1% agarose (7cm x 10 cm gel) containing ethidium bromide (0.5 µg/mL) in 1x TAE (40 mM Tris-acetate, 2 mM ethylene diamine tetra-acetic acid (EDTA), pH 8). DNA samples were loaded into agarose wells by mixing with approximately 10% v/v DNA loading dye (40% w/v sucrose, 0.25% bromophenol blue) and electrophoresis carried out at 100 V for ¾-1 hour for sufficient separation. DNA was visualised by exposure to UV light and sizes of bands compared with a DNA size standard.

### **2.2.2 Restriction endonuclease digestion.**

Generally 100-500 ng of plasmid DNA was digested with 5-10 units of restriction enzyme along with 10% v/v of the recommended 10x buffer (specified by the manufacturer) in a total volume of 50 µL per reaction. Digests were incubated at 37°C for 1-2 hours and then analysed by agarose gel electrophoresis.

### **2.2.3 Oligonucleotides.**

Each oligonucleotide was reconstituted in TE buffer (10 mM Tris-HCl, pH 8; 0.1 mM EDTA) to generate stock solutions of 10 µg/µL and stored at -20°C. Additional dilutions were carried out as required using TE buffer: 100 ng/µL for probes used in EMSA assays, 50 ng/µL for PCR primers and 1 µg/µL for preparing competitors for use in EMSA and oligonucleotide-affinity chromatography.

### **2.2.4 Preparing double-stranded oligonucleotides.**

The competitors and biotinylated oligonucleotides were stored as single-stranded stock solutions (10 µg/µL). Double-stranded oligonucleotides were generated by adding 10 µL (equal volumes) of each complimentary single-stranded oligonucleotide (1 µg/µL) together and placed in a boiling water-bath (95-100°C) for 5 minutes. Samples were then left in the water-bath to cool gradually for about 1 hour (cooling approximately 10°C every 5 minutes), then stored at -20°C until required.

## 2.2.5 Polymerase chain reactions (PCR).

Each PCR reaction contained 250 ng of each primer, 1.5 mM dNTP mix, 1-2 units Taq DNA polymerase (5 U/ $\mu$ L), 10% v/v of 10x PCR buffer (100 mM Tris-HCl; pH 8.3, 500 mM KCl and 15 mM  $MgCl_2$ ) and 1 ng of template DNA in a total volume of 50  $\mu$ L. PCR samples were amplified using a thermocycler (GeneAmp<sup>R</sup> PCRsystem2700, AB, Applied Biosystem). Standard PCR conditions are as follows

	95°C for 5 minutes	
Denaturing	95°C for 1 minute	} 25 cycles
Annealing	45-55°C for 30 seconds	
Extension	72°C for 40 seconds	

Approximately 10% of the final PCR product was analysed by gel electrophoresis and PCR products were either gel purified or purified using the QIAquick PCR purification kit according to the manufacturers' instructions. This purification kit utilises a silica-based spin cartridge that purifies double-stranded DNA and removes excess dNTPs, Taq polymerase and primers from the sample.

## 2.2.6 DNA purification by gel electrophoresis.

Standard gel electrophoresis was used to separate DNA fragments in a complex mixture of DNA products. The required fragment was visualised under long wave length UV light and excised from the gel using a scalpel blade. DNA was extracted from the gel slice using a CONCERT<sup>TM</sup> gel purification kit according to the manufacturers' instructions. This kit uses a sodium perchlorate solution to first dissolve the agarose, and then the double-stranded DNA is purified as it is added to the silica-based cartridge, and eluted in TE buffer.

## 2.2.7 Phosphatase treatment of plasmids.

Removal of phosphate groups from linearised plasmid DNA was performed by incubation of 1  $\mu$ g of digested plasmid DNA with 1 unit thermosensitive alkaline phosphatase (1 U/ $\mu$ L), 10%v/v of a 10x phosphatase buffer (250 mM Tris-acetate; pH 7.8, 1 M potassium acetate, 100 mM magnesium acetate, 10 mM dithiothreitol) in a total volume of 40  $\mu$ L. the phosphatase was deactivated by heating reactions to 74°C for 15 minutes.

### 2.2.8 Addition of 5' phosphate groups to oligonucleotides.

The addition of phosphate groups onto the 5'-termini of synthesized oligonucleotides in preparation for a ligation reaction was carried out by incubation at 37°C for 50 minutes with 10 units T4 polynucleotide kinase (30 U/  $\mu$ L), 10% v/v 10x polynucleotide kinase buffer (0.5 M Tris-HCl; pH 7.6, 100 mM MgCl<sub>2</sub>, 100 mM 2-mercaptoethanol), 0.2 mM ATP (adenosine triphosphate) and 500 ng of oligonucleotide in a total volume of 10  $\mu$ L. Reactions were inactivated by heating to 65°C for 10 minutes.

### 2.2.9 Ligation reactions.

Insert DNA was added at a 3:1 molar ratio compared to vector DNA per ligation reaction with 1  $\mu$ L T4 DNA ligase (1 U/ $\mu$ L) and 10% v/v 10x reaction buffer (660 mM Tris-HCl, 50 mM MgCl<sub>2</sub>, 10 mM dithioerythritol, 10 mM ATP; pH 7.5). Individual ligation reactions were made up with a maximum of 150 ng total DNA in a total reaction volume of 20  $\mu$ L. Ligation reactions were incubated overnight at 16°C. The following day the T4 DNA ligase was inactivated by heating to 65°C for 10 minutes and transformations were performed.

### 2.2.10 Transformation of *Escherichia coli* (*E.coli*).

All bacterial work was carried out following standard aseptic techniques. *E.coli* XL-1 blue competent cells were prepared by growing a fresh 5 mL overnight culture of cells in Luria Bertani (LB) broth (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) supplemented with ampicillin (100  $\mu$ g/mL) and using this to inoculate 200 mL LB broth supplemented with 2 mg tetracycline. The cells were left to grow at 37°C until an optical density of approximately 0.5 at 600 nm was reached. After this, cells were harvested by centrifugation at 4°C for 15 minutes at 2000 g and the supernatant was removed. The cell pellet was resuspended in 25 mL of ice-cold 100 mM CaCl<sub>2</sub>, followed by 20 minute incubation on ice. Cells were pelleted again by centrifugation at 4°C for 15 minutes at 2000 g and all the supernatant was removed. The final cell pellet was resuspended in 5 mL ice-cold 100 mM CaCl<sub>2</sub> and left to stand on ice for 1 hour. The cell suspension was supplemented with 15% v/v glycerol and mixed well to resuspend cells. Competent XL-1 cells were then split into 200  $\mu$ L aliquots, snap frozen using liquid nitrogen, and then stored at -80°C until required.

Typically half a ligation reaction (10  $\mu$ L) was added to 100  $\mu$ L of XL-1 competent cells and incubated on ice for 30 minutes. Reactions were heat shocked by heating to 42°C for 2 minutes, then 800  $\mu$ L of LB broth was added to the reaction and incubated at 37°C for 1 hour (with shaking) to recover before plating. Transformed cells were pelleted by centrifugation at 1500 g for 5 minutes, then resuspended in 100  $\mu$ L LB supplemented with ampicillin (100 mg/mL). LB (with ampicillin) bacteriological agar plates (20 g/L LB broth mix, 15 g/L bacteriological agar, ampicillin 100 mg/mL) were prepared by spreading plates with 0.5 mM isopropyl thiogalactosidase (IPTG) and 80  $\mu$ g/mL X-gal (5-bromo-4-chloro-3-indoyl- $\beta$ -D-galactosidase, before use. A proportion of each transformation (5-50  $\mu$ L) was plated and incubated overnight at 37°C to develop colonies.

### **2.2.11 Purification of plasmids.**

Large scale plasmid preparations were carried out to obtain plasmid DNA of high quality and quantity for transient transfections. The required cultures were prepared by inoculating 500 mL LB broth supplemented with ampicillin (100 mg/mL) and incubating overnight at 37°C with shaking. Plasmid DNA was purified using CONCERT™ high purity plasmid maxiprep kit according to the manufacturers' instructions. This kit uses a system of anion-exchanged-based columns/tips to purify high grade plasmid DNA (QIAGEN handbook).

Small scale plasmid preparations were used when a small quantity of DNA was required for PCR or sequencing. Generally 2 mL of a 5 mL overnight culture was used and plasmid DNA was isolated using a Quantum® Prep plasmid mini-prep kit according to the manufacturers' instructions. This kit uses the silicon dioxide exoskeleton of diatoms as the DNA binding matrix (Bio-Rad product sheet).

### **2.2.12 Tissue culture.**

All tissue culture operations were carried out aseptically in a class II biohazard hood (LA2-4A1, ESCO) and cells were incubated at all times with 5% CO<sub>2</sub> at 37°C in humid conditions (HERAcell 150, Heraeus, Kendro). HeLa cells were maintained in a modified Eagle's minimal essential reduced serum media, OPTI-MEM® I that was completed by the addition of 2% v/v fetal-calf serum and 1% v/v penicillin/streptomycin (5000 U/mL penicillin and 5mg/mL streptomycin sulphate in 0.85% saline).

All cells were revived from 1 mL frozen stocks (10% v/v Dimethyl sulfoxide (DMSO) in fetal calf serum, stored under liquid nitrogen) defrosted and resuspended in 5 mL complete OPTI-MEM<sup>®</sup>I and mixed briefly but thoroughly. The cells were pelleted by centrifugation for 5 minutes at 100 g and the supernatant removed to be replaced with 5 mL fresh complete OPTI-MEM<sup>®</sup>I media. The total volume of cells was placed in T75 tissue culture flask containing 9 mL complete OPTI-MEM<sup>®</sup>I and left to grow in incubator.

Cells were passaged once they reached 80-90% confluence. First the old media was removed from the cells and then cells were washed twice with 5 mL 1x TrypLE<sup>™</sup> Express. Cells were left to stand for up to 5 minutes to round-up and detach from the surface of the flask. Further detachment was achieved by abruptly tapping the flask several times. The cells were then resuspended in 5 mL complete OPTI-MEM<sup>®</sup>I and new T75 flasks were seeded with 1 mL of cell suspension and 12 mL of fresh complete OPTI-MEM<sup>®</sup>I then placed back into the 37°C incubator.

### **2.2.13 Treatment with PJ34.**

PJ34 (PARP inhibitor VIII) was provided as a dry powder and was reconstituted with water as working stock solution of 500 µM. PJ34 was dispensed into aliquots in dark amber tubes and stored at -20°C until required.

HeLa cells were seeded in 150 mm tissue culture plates containing 15 mL complete OPTI-MEM<sup>®</sup>I to obtain a cell density of approximately 60% confluence prior to PJ34 treatment. Cells were treated for 48 hours with 0-2 µM PJ34 diluted directly into the medium and harvested as described in 2.2.14. For transient transfections PJ34 was added 2 hours after transfection (as described in section 2.2.27).

### **2.2.14 Whole cell HeLa extracts.**

HeLa cells were grown to 80-90 % confluence in 15 mL complete OPTI-MEM<sup>®</sup>I in 150 mm diameter tissue culture plates. The media was removed from plates and cells were washed twice with 4 mL PBS (0.14 M NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>; pH 7.2-7.4). One 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 were scraped and collected in 1.5 mL microfuge tubes. Cells

were pelleted by centrifugation at 12,000 g for 5 minutes at 4°C and the supernatant discarded. Then the cells were resuspended in 300 µL (for each 150 mm plate) extraction buffer (40 mM HEPES; pH 7.9, 0.4 M KCl, 1 mM dithiothreitol, 10 % v/v glycerol and protease inhibitor cocktail) and treated with three freeze-thaw cycles using liquid nitrogen. Extracts were centrifuged at 12,000 g for 15 min at 4°C to pellet cell debris and the supernatant retained. Extracts were dispensed into aliquots, snap frozen immediately in liquid nitrogen and stored at -80°C until required.

### **2.2.15 HeLa cell nuclear extracts.**

All buffers used for this method were made fresh on the day of extraction and kept stored on ice until use, PSMF and dithiothreitol were added just prior to use. Samples were kept chilled where possible and cell counts were approximate.

HeLa cells were grown in 15 mL complete OPTI-MEM<sup>®</sup>I to approximately 90% confluence in 150 mm tissue culture plates. Each plate of cells was washed twice with 5 mL PBS followed by harvesting, by scraping cells off plates, in 1 mL of PBS. Cells were pelleted by centrifugation at 500 g for 5 minutes; generally each plate of 90-100% confluent HeLa cells gave rise to approximately  $2-3 \times 10^7$  cells. PBS was removed and cells were resuspended in 100 µL sucrose buffer (10 mM Tris-HCl; pH 8, 0.32 M sucrose, 3 mM CaCl<sub>2</sub>, 2 mM MgAc, 5% v/v Nonidet P40, 1 mM dithiothreitol and 0.5 mM PMSF) per  $1 \times 10^7$  cells. This was usually 300 µL per 150 mm plate of HeLa cells, depending on initial cell density. The cells and sucrose buffer were mixed gently by inverting tubes several times to promote cell lysis. Nuclei were pelleted using a refrigerated centrifuge (4°C) 500 g for 5 minutes, then the supernatant removed (this can be used as a cytoplasmic fraction of proteins if required). The pellets were washed a further two times in 1 mL wash buffer (10 mM Tris-HCl; pH 8, 0.32 M sucrose, 3 mM CaCl<sub>2</sub>, 2 mM MgAc, 1 mM dithiothreitol and 0.5 mM PMSF), and centrifuged for 5 minutes at 500 g at 4°C to pellet cells each time between washes. All traces of wash buffer was removed by aspiration, then pellets were resuspended in 90 µL (for every plate of cells) low salt buffer (20 mM HEPES; pH 7.9, 1.5 mM MgCl<sub>2</sub>, 20 mM KCl, 0.2 mM EDTA, 25 % v/v glycerol, 0.5 mM dithiothreitol and 0.5 mM PMSF) and mixed gently to completely resuspend nuclei. An equal volume of high salt buffer (20 mM HEPES; pH 7.9, 1.5 mM MgCl<sub>2</sub>, 800 mM KCl, 0.2 mM EDTA, 25 % v/v glycerol, 1% v/v Nonidet P40, 0.5 mM dithiothreitol, 0.5 mM PMSF and protease inhibitor cocktail) was added very slowly

while mixing gently by pipetting. Samples were then incubated for 45 minutes at 4°C with gentle shaking. Cellular debris was removed by centrifuging samples at 14,000 g for 20 minutes at 4°C and the supernatant retained. Nuclear extracts were dispensed into aliquots and immediately snap frozen in liquid nitrogen. Nuclear extracts were stored at -80°C until required.

### **2.2.16 Quantification of total protein in extracts.**

Quantification was carried out using either a Bradford protein assay or directly reading the absorbance at 260 nm using the nano-drop (ND-100) spectrophotometer.

Protein concentrations were determined using the Bio-Rad protein assay, according to the manufacturers' instructions. This system is based on the Bradford method using coomassie brilliant blue G-250 dye. A protein standard curve was constructed using bovine serum albumin (BSA) ranging from 0-2.5 µg of protein. The Bio-Rad protein reagent concentrate was diluted 1:5 in ddH<sub>2</sub>O, where 180 µL was added to 1-2 µL of each diluted protein sample (diluted in triplicate) or BSA protein standard (done in duplicate) in a total volume of 200 µL in a 96 well microplate. Samples were allowed to develop blue colour at room temperature for about 5 minutes and then absorbances were read at 595 nm using 96 well plate reader (anthos reader HT2 type 12 500, anthos labtech instruments, Salsburg, Austria).

### **2.2.17 Sodium dodecylsulphate polyacrylamide gel electrophoresis.**

For visualisation and determination of the molecular weight of proteins, samples were separated using sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE), according to the method of Laemmli (1970). Denaturing protein gels were prepared using 6.5% polyacrylamide stacking gel and 8% polyacrylamide resolving gel.

Protein samples were denatured by heating with 5% v/v 5 x treatment buffer (0.06 M Tris-HCl; pH 6.8, 5% w/v SDS, 5% v/v β mercaptoethanol, 25% v/v glycerol and 0.001% w/v bromophenol blue) in a boiling waterbath for 5 minutes, followed by separation in a 8% SDS-PAGE for 1-1½ hours at 110-130 V utilising a mini-protean 3 (Bio-Rad) system. A protein size standard (precision plus dual colour protein standards) was always included

along side protein samples: 5-10  $\mu\text{L}$  for coomassie-blue stained gels or 0.5-1  $\mu\text{L}$  for silver stained gels. Electrophoresis was carried out for 1-2 hrs at 110-130 V.

### **2.2.18 Coomassie–blue stain for protein gels.**

In general, Coomassie-blue was used to visualise  $>0.1 \mu\text{g}$  protein per band after SDS-PAGE. Each SDS-PAGE was submerged in Coomassie-blue staining solution (2.5 mM Coomassie brilliant blue (R-250), 50% methanol, 10% acetic acid) for 5-20 minutes at room temperature with shaking. Then destained in (15 % v/v methanol, 6% v/v acetic acid) at room temperature with shaking and successive changes of destain solution for 4-16 hours.

### **2.2.19 Silver staining for protein gels.**

Silver staining was used when protein concentrations were low as this technique is approximately 100 times more sensitive than Coomassie-blue staining. This method allows as little as 1 ng of protein per band to be detected. The following protocol was adapted to suit a single small SDS-PAGE (0.75 mm thick, 9.5cm x 6.5 cm gel). All steps were carried out at room temperature with constant shaking and solutions were made immediately before to use. Gels were fixed in 100 mL fixing solution (50% v/v ethanol and 10% acetic acid) for approximately 1 hour, followed by three washes of 5 minutes in 100 mL ddH<sub>2</sub>O. The gel was then placed in 100 mL 0.8 mM sodium thiosulphate solution for 2 minutes, followed by a washing step of 5 minutes in 100 mL ddH<sub>2</sub>O. Next the gel was placed in 100 mL 2g/L silver nitrate solution for 20 minutes followed by two very brief 5-10 second washes in ddH<sub>2</sub>O. Bands were developed by immersing the gel in 100 mL developing solution (566 mM Na<sub>2</sub>CO<sub>3</sub>, 0.016 mM sodium thiosulphate and 0.05% v/v formaldehyde solution), with constant supervision, bands can develop quite quickly (30 seconds) if procedure works well and there is substantial amounts of protein. Otherwise staining can take up to 15 minutes. The staining reaction was stopped by placing gel in 5% ice cold acetic acid for 5-10 minutes and the gel was stored in ddH<sub>2</sub>O with 10% v/v glycerol.

### **2.2.20 Immunodetection/Western Blotting.**

Standard SDS-PAGE was carried out to separate the protein samples as described in section 2.2.17. Then using the same system (mini-protean 3), proteins were transferred to a positively charged nylon membrane for 1 ½ hours at 450 mA submerged in transfer buffer

(15 mM Tris, 192 mM glycine; pH 8.3). The blotting apparatus was disassembled and the membrane was blocked overnight at 4°C in 1% v/v blocking solution (supplied with BM chemiluminescence blotting substrate) diluted in TBST (0.15 M NaCl, 50 mM Tris-HCl; pH 7.5 plus 1 mL/L Tween 20). The remaining procedure was carried out at room temperature and antibodies were diluted immediately before use in 0.5% v/v blocking solution in TBST. The membrane was incubated in primary (1°) antibody for 2 ½ hours, followed by four 20 minute washes in TBST. Next the membrane was placed in horse-radish peroxidase (HRP) conjugated secondary (2°) antibodies for 1 hour, followed by three 15 minute washes in TBST before detection. Chemiluminescence detection was carried out using BM chemiluminescence blotting substrate (POD) according to manufacturers' instructions and the image was captured by exposure to X-ray film. The protein size standard was visualised by using fluorescent paint (Polymark, natural glow/dimensional fabric paint) spotted onto marker positions indicated on the membrane.

### **2.2.21 Ion-exchange chromatography (heparin-sepharose column).**

During this procedure different salt concentrations were required at different steps, each was prepared in a standard equilibration buffer; buffer E (25 mM HEPES; pH 7.9, 12.5 mM MgCl<sub>2</sub>, 10% v/v glycerol, 1 mM EDTA, 1mM dithiothreitol plus protease inhibitor cocktail) containing the appropriate amount of KCl as specified throughout the protocol. All buffers used for this method were also filter sterilised and stored at 4°C until required.

To prepare a sample suitable for loading onto the heparin-sepharose column, whole cell HeLa extract was dialysed 12-15 hours in 50 times the extract volume of buffer E containing 0.1 M KCl, changing the buffer every 8 hours. Cellular debris was removed by centrifugation for 15 minutes at 700 g at 4°C. The cleared extract was then filter sterilised through an 8 µm filter and quantified before loading onto the column. The 5 mL heparin-sepharose column was equilibrated prior to use by washing with 20 times the column volume of buffer E containing 0.1 M KCl. Approximately 20 mg of total protein was loaded onto the column and any unbound proteins were washed away using 10 times the column volume of buffer E containing 0.1 M KCl. A gradient was set up using buffer E containing 1 M KCl, to range from 0.15 M-0.7 M KCl. Fractions were collected between 0.1-0.6 M KCl and the column was washed with 10 times the column volume of buffer E containing 2 M KCl to remove any

residual bound proteins. Finally, the column was restored with 40 times the column volume of 0.05 M sodium acetate containing 20% v/v ethanol before storage at 4°C.

### **2.2.22 Oligonucleotide-affinity chromatography.**

The magnetic Dynabeads (Invitrogen, Dynal Dynabeads M-280) are a polystyrene bead coated with a monolayer of streptavidin. A magnetic microfuge tube rack was used to immobilise the dynabeads during washing and elution steps, microfuge tubes were placed in the rack for 1-2 minutes and the supernatant was removed with a pipette while the sample remained drawn to the magnet. Dynabeads were supplied at a concentration of 10 mg/mL and were used according to the manufacturers' instructions, with modifications as detailed below. During this procedure different salt concentrations were required at different washing and elution steps, each was prepared in a standard equilibration buffer; buffer E (25 mM HEPES; pH 7.9, 12.5 mM MgCl<sub>2</sub>, 10% v/v glycerol, 1 mM EDTA, 1mM dithiothreitol plus protease inhibitor cocktail) containing the appropriate amount of KCl as specified in throughout the protocol.

To prepare magnetic beads for use 500 µL dynabeads were first washed three times with 1 mL 2x binding and wash (B&W) buffer (10 mM Tris-HCl; pH 7.5, 1 mM EDTA, 2 M NaCl). The next step was to bind the biotinylated-oligonucleotide to the streptavidin beads, this was achieved by adding 200 µL (100 ng/µL, to allow DNA to be in excess) of double-stranded ICB1mt/GC1wt-biotinylated oligonucleotide (refer to table 3.1 for oligonucleotide sequence) with 300 µL ddH<sub>2</sub>O and 500 µL 2x B&W buffer to the washed beads, followed by a 30 minute incubation at room temperature on a orbital shaker. To remove any unbound biotinylated-oligonucleotide, the bead mixture was washed 3 times in 1 mL 1x B&W buffer (5 mM Tris-HCl; pH 7.5, 0.5 mM EDTA, 1 M NaCl).

In experiments where a competitor was added to this assay the double-stranded competitor DNA (1 µg/µL) was pre-mixed with the protein binding reaction 30 minutes prior to the addition of the beads and kept on ice. A standard protein binding reaction contained 600 µL GSB (gel shift buffer: 40 mM Tris-HCl; pH 7.6, 100 mM KCl, 0.4 mM EDTA, 1 mM dithiothreitol and 16% w/v ficoll) with 300 µL HeLa whole cell or nuclear extract (2 µg/µL) and 300 µL ddH<sub>2</sub>O all of which was added to the bead-oligonucleotide mixture and incubated at 4°C for ½-1 hour with shaking. To remove unbound and excess proteins, the

bead-oligonucleotide-protein complexes were washed 3 times in 1 mL buffer E containing 0.025 M KCl, then 3 times in 1 mL buffer E containing 0.05 M KCl and finally 2 times in 1 mL buffer E containing 0.1 M KCl. Finally, oligonucleotide-bound proteins were eluted from the bead-oligonucleotide complexes with 250  $\mu$ L buffer E containing 0.5 M KCl and incubated at 4°C for 1 hour with shaking. Final purified extracts were supplemented with 20  $\mu$ L 7x protease inhibitor cocktail before dispensing into aliquots and storage at -80°C.

Alternatively, when creating fractionated aliquots of oligonucleotide-affinity purified extract; proteins were eluted step-wise from the bead-oligonucleotide complex by increasing the salt concentration at each step. The oligonucleotide-binding and protein-binding reactions remained the same however the bead-oligonucleotide-protein complexes were washed 4 times in 1 mL buffer E containing 0.025 M KCl, then 4 times in 1 mL buffer E containing 0.05 M KCl. Oligonucleotide-bound proteins were eluted step-wise, first in 80  $\mu$ L buffer E containing 0.1 M KCl and incubating at 4°C for 30 minutes with shaking before magnetic separation, then repeated with 80  $\mu$ L buffer E increasing the KCl concentrations by 0.1 M each time. This created a succession of proteins that eluted from the oligonucleotide in a range of 0.1 M- 0.7 M KCl. Each final fraction was supplemented with 5  $\mu$ L 7x protease inhibitor cocktail before dispensing into aliquots and storage in -20°C.

### **2.2.23 <sup>32</sup>P-labelling oligonucleotides.**

Each oligonucleotide used as a probe in EMSA or cross-linking assays was <sup>32</sup>P-labelled at the 5'-end and gel purified before use. Labelling reactions were carried out using 1.5  $\mu$ L (10 U/ $\mu$ L) optiKinase<sup>TM</sup>, 2.5  $\mu$ L 10x optiKinase reaction buffer (500 mM Tris-HCl; pH 7.5, 100 mM MgCl<sub>2</sub>, 50 mM dithiothreitol), 5  $\mu$ L  $\gamma$ <sup>32</sup>P [ATP] with 100 ng of single-stranded (forward) oligonucleotide in a total volume of 25  $\mu$ L. Labelling reactions were incubated at 37°C for 40 minutes and the reaction was terminated by heating to 65°C for 10 minutes. After completion of the labelling reaction, the complimentary strand was annealed to its <sup>32</sup>P-labelled counterpart by adding 600 ng (6x excess) of the reverse single-stranded oligonucleotide in the presence of 50 mM KCl in a total volume of 50  $\mu$ L, samples were boiled for 5 minutes in a waterbath then left to cool slowly over an hour to enhance annealing.

An equal volume (50  $\mu\text{L}$ ) of GSB was added to each reaction and the total volume was loaded into a 10% polyacrylamide gel (37 cm long with 0.4 mm spacers) in 1x TBE (0.89 M Tris, 0.89 M boric acid and 0.02 mM EDTA; pH 8) and electrophoresis carried out for approximately 1 ½ hours at 30 W (1500 V). The wet gel was wrapped in Saran wrap and exposed to X-ray film for approximately 1 minute and the location of the double-stranded  $^{32}\text{P}$ -labelled oligonucleotide was mapped to a region within the polyacrylamide gel. Each oligonucleotide was excised from the gel and the oligonucleotides eluted by adding 400  $\mu\text{L}$  of 50 mM KCl to each gel slice, followed by an overnight incubation at 37°C. The following day the gel slices were centrifuged for 5 minutes at 12,000 g and the supernatant transferred into new 1.5 mL microfuge tubes.

Each labelled oligonucleotide was analysed for the incorporation of radioactivity by scintillation counting and individual oligonucleotides were diluted to approximately 0.5 ng/ $\mu\text{L}$  (about 20,000 counts per minute) to ensure a consistent amount of  $^{32}\text{P}$ -labelled oligonucleotide was used in each assay.

#### **2.2.24 Electrophoretic mobility shift assays (EMSA).**

Each EMSA reaction contained 1  $\mu\text{g}$  of poly dI.dC (1  $\mu\text{g}/\mu\text{L}$  in 50 mM  $\text{MgCl}_2$ ), 50 % v/v GSB, 2-10  $\mu\text{g}$  of protein along with 0.5 ng of  $^{32}\text{P}$  labelled oligonucleotide in a total volume of 20  $\mu\text{L}$ . Reactions were pre-mixed without the  $^{32}\text{P}$  labelled oligonucleotide and if antibody or double-stranded competitor were required these components were added at this stage. After a 10 minute incubation on ice 0.5 ng of labelled oligonucleotide was added to the reaction and incubated for 15 minutes at room temperature. Half (10  $\mu\text{L}$ ) of each EMSA reaction was loaded into a non-denaturing 4% polyacrylamide gel (BRL V15.17 apparatus fitted with 0.75 mm spacers) in 0.25x TBE and electrophoresis carried out at 200 V for approximately 1-1½ hours. Gels were then dried onto DE-81 paper using a gel drier (model 583, Bio-Rad) and exposed to X-ray film for 10-20 hours using a radioactive safe cassette with double intensifying screens at -80°C. The X-ray films were developed using a 100Plus<sup>TM</sup> automatic X-ray film processor in a dark room.

The list of antibodies used can be found in section 2.1. In the case of NF-Y, antibodies against the NF-YA subunit were specifically chosen as this subunit is essential for an

association with DNA. The NF-Y trimeric protein contains NF-YA, NF-YB and NF-YC which are tightly associated together.

### **2.2.25 Electro-elution of proteins from EMSA gels.**

EMSA reactions were prepared as described for EMSA assays, however the total volume was increased to 40  $\mu\text{L}$  with 50% v/v GSB, 2  $\mu\text{g}$  poly dI.dC, 10-20  $\mu\text{g}$  HeLa extract and 2 ng  $^{32}\text{P}$  labelled ICB1mt/GC1wt oligonucleotide. The total volume (40  $\mu\text{L}$ ) of each EMSA reaction was loaded into each well of a 5% non-denaturing polyacrylamide gel (BRL V15.17 apparatus fitted with 1.5 mm spacers) in 0.25x TBE. Electrophoresis was carried out for 1½ - 1¾ hours in 0.25x TBE at 200 V. After electrophoresis, the wet gel was wrapped in Saran-wrap and exposed to X-ray film for 5-8 hours. The X-ray film and gel were marked to line up exactly, so that the bands that corresponded with specific protein-DNA complexes in the wet gel could be aligned. The band of interest was carefully excised from the gel using a clean scalpel blade. A single gel slice, approximately 2 cm x 1 cm was placed inside a single D-tube dialyser (Midi, MWCO 5.3 kDa) and filled with about 700  $\mu\text{L}$  0.25x TBE and 25  $\mu\text{L}$  7x protease inhibitor cocktail. Four gel slices, therefore four D-tubes, were prepared for each electro-elution procedure. The D-tube dialysers were placed in their floatation rack and electrophoresis was carried out submerged in 0.25x TBE for 12-15 hours at 100 V at 4°C. The polarity was reversed for 2 minutes to release any protein that may have bound to the membranes within the tubes before disassembling the electrophoresis equipment. The gel slice was removed from inside each D-tube and the buffer was collected, pooled and concentrated using ultrafiltration devices (centricon-10, Amicon). During the ultrafiltration process, the TBE buffer was exchanged into extraction buffer (40 mM HEPES pH 7.9, 0.4 M KCl, 1 mM dithiothreitol, 10% v/v glycerol and protease inhibitor cocktail). A final volume of approximately 200  $\mu\text{L}$  was obtained from this procedure and protein samples were dispensed into aliquots and stored at -80°C until required.

### **2.2.26 BrdU cross-linking.**

The success of UV cross-linking was highly dependent on the amount of BrdU oligonucleotide used in each assay. Therefore,  $^{32}\text{P}$ -labelled oligonucleotides were initially titrated with 2  $\mu\text{L}$  of HeLa extract to optimise each reaction. Generally, 2-4 ng (0.5 ng/ $\mu\text{L}$ ) of  $^{32}\text{P}$ -labelled BrdU oligonucleotide was required along with 1-5  $\mu\text{g}$  of protein in a total

volume of no more than 10  $\mu\text{L}$  for every cross-linking reaction. Protein and  $^{32}\text{P}$ -labelled oligonucleotide were mixed together in 0.2  $\mu\text{L}$  clear (thin walled) PCR tubes for 15 minutes at room temperature to allow proteins to associate with DNA, samples were centrifuged briefly to ensure all the liquid remained at the bottom of the tubes. The cross-linking reactions were carried out by inverting open PCR tubes on a flatbed transilluminator, so the samples were exposed to UVB light (290-320 nm) for 30 minutes at room temperature. After cross-linking, each sample was denatured by heating with 5% v/v 5 x treatment buffer (0.06 M Tris-HCl; pH 6.8, 5% w/v SDS, 5% v/v  $\beta$  mercaptoethanol, 50% v/v glycerol and 0.001% w/v bromophenol blue) and boiled for 5 minutes in a waterbath. The total volume of each reaction was subjected to SDS-PAGE (section 2.2.17) for protein separation and gels were either silver or Coomassie-blue stained then dried on DE-81 paper. Fluorescent paint (Polymark) was spotted onto the marker positions on the dried gel and exposed to X-ray film for 5-15 hours using a radioactive safe cassette with double intensifying screens at  $-80^{\circ}\text{C}$ .

### **2.2.27 Transient transfections.**

HeLa cells were seeded in a total of 1 mL of complete OPTI-MEM<sup>®</sup>I 24 hours prior to transfection in 12 well tissue culture plates, to obtain a cell density of approximately 70% confluence at the time of transfection. Each of the plasmids to be transfected was quantified and diluted to exactly 0.5  $\mu\text{g}/\mu\text{L}$  in TE and stored at  $-20^{\circ}\text{C}$  until required. A total of 0.25  $\mu\text{g}$  pCMV SPORT  $\beta$ -gal ( $\beta$ -Galactosidase control vector) and 0.5  $\mu\text{g}$  of pGL3B (luciferase reporter gene) -617topo2 $\alpha$  plasmid was transiently transfected using FuGENE<sup>®</sup>6 using a 3:1 ratio of FuGENE<sup>®</sup>6 per  $\mu\text{g}$  of total plasmid DNA according to the manufacturers' instructions.

Each transfection was performed in triplicate and harvested 48 hours after transfection in 100-150  $\mu\text{L}$  of cell lysis buffer (25 mM Tris-HCl; pH 7.8, 2 mM dithiothreitol, 2 mM 1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid (CDTA), 10 % v/v glycerol and 5 % v/v triton X-100) and analysed for  $\beta$ -galactosidase and luciferase activity immediately.

### **2.2.28 $\beta$ -galactosidase activity.**

The plasmid pCMV SPORT- $\beta$ gal was used as an internal transfection control to be able to correct for variations due to transfection efficiency, cell density and plasmid integrity. After

harvesting each transient transfection in cell lysis buffer, 5  $\mu\text{L}$  of this was added to 50  $\mu\text{L}$  ONPG buffer (o-nitrophenyl- $\beta$ -D- galactopyranoside, 2 mg/mL in 60 mM  $\text{NaH}_2\text{PO}_4$ , 40 mM  $\text{Na}_2\text{HPO}_2$ ) and 100  $\mu\text{L}$   $\beta$ -galactosidase assay buffer (60 mM  $\text{NaH}_2\text{PO}_4$ , 40 mM  $\text{Na}_2\text{HPO}_2$ , 10 mM KCl, 1 mM  $\text{MgCl}_2$ ) in 96 well microtitre plates. Samples were incubated at 37°C for approximately 10-30 minutes to develop faint yellow colour then 50  $\mu\text{L}$  1 M  $\text{Na}_2\text{CO}_3$  was added to terminate the reaction. Absorbancies were measured at 405 nm against a blank of 5  $\mu\text{L}$  cell lysis buffer using a 96 well plate reader (anthos reader HT2 type 12 500, anthos labtech instruments, Salsburg, Austria).

### **2.2.29 Luciferase assays.**

The pGL3B vector provides a luciferase reporter gene which is under the control of the human -617topo2 $\alpha$  promoter. Promoter strength dictates how much luciferase is synthesized within the transfected cells, hence different pGL3B -617topo2 $\alpha$  promoter constructs may generate different luciferase activities depending on the promoter elements that are altered. Luciferase catalyses the reaction which converts luciferin into oxyluciferin and light, this yellow-green light was detected using a 96 well plate reader (FLUOstar galaxy, BMG labtechnologies, Melbourne, Australia). After harvesting each transient transfection in cell lysis buffer, 10-20  $\mu\text{L}$  of this was added to 20  $\mu\text{L}$  luciferase assay reagent and the FLUOstar galaxy was programmed to take a photon count every second over three minutes. An equivalent amount of cell lysis buffer was used as a blank for this assay. Data was recorded and displayed in Excel (Microsoft Office, 1997), where the maximum luciferase readings were recorded and used for further calculations.

### **2.2.30 Calculations for the normalised luciferase activities.**

As each transfection is highly dependent on the transfection efficiency, cell density and plasmid integrity a number of controls were required to maintain experimental integrity. Each individual transient transfection was performed in triplicate and this was repeated a minimum of three times on separate occasions. This created a large amount of data that was processed as described in table 2.1, 2.2 and 2.3.

Statistical analysis was performed using software within Excel (Microsoft office, 2003). Errors were calculated using the AVGDEV feature and the significance testing was carried

out using a (one tailed distribution) paired students *t*-test. The p-value generated from the students *t*-test describes whether both data sets originate from distributions with equal population means. A p-value of less than 0.1 indicates that there is a less than 10% chance that the two data sets are similar, therefore are significantly different at a 90% confidence interval.

	Triplicates			
$\beta$ -galactosidase	Absorbance 405 nm	Absorbance 405 nm	Absorbance 405 nm	Averages
$\beta$ gal	$\beta a$	$\beta b$	$\beta c$	
Cell lysis buffer blank	Ba	Bb	Bc	AvgB
Corrected $\beta$ gal	$\beta a - \text{AvgB} = c\beta a$	$\beta b - \text{AvgB} = c\beta b$	$\beta c - \text{AvgB} = c\beta c$	AvgC $\beta$ 1

**Table 2.1: Processing  $\beta$  galactosidase data.**

Each transient transfection was performed in triplicate.  $\beta$ -galactosidase activity was corrected by subtracting the average blank  $\beta$ -galactosidase absorbance read at 405 nm (AvgB).

	Triplicates			
Luciferase Activity	Luciferase Maxima	Luciferase Maxima	Luciferase Maxima	Averages
Luciferase	La	Lb	Lc	
Cell lysis buffer blank	LBa	LBb	LBc	AvgLB
Corrected luciferase	$La - \text{AvgLB} = cLa$	$Lb - \text{AvgLB} = cLb$	$Lc - \text{AvgLB} = cLc$	AvgCL1

**Table 2.2: Processing luciferase data.**

Each transient transfection was performed in triplicate. The luciferase maxima were recorded as actual photons released from each reaction using the FLUOstar galaxy 96 well plate reader fitted with a chemiluminescence detection system. Luciferase activity was corrected by subtracting the average blank luciferase activity (AvgCL).

	Triplicates			
Luciferase Activity	Luciferase Maxima	Luciferase Maxima	Luciferase Maxima	Averages
Corrected $\beta$ gal	AvgC $\beta$ 1	AvgC $\beta$ 2	AvgC $\beta$ 3	
Corrected luciferase	AvgCL1	AvgCL2	AvgCL3	
Normalised luciferase	$\text{AvgCL1}/\text{AvgC}\beta$ 1	$\text{AvgCL2}/\text{AvgC}\beta$ 2	$\text{AvgCL3}/\text{AvgC}\beta$ 3	AvgNL

**Table 2.3: Generating normalised luciferase activities.**

Normalised luciferase activity was calculated by dividing the corrected luciferase values by the corrected  $\beta$  galactosidase values for every transient transfection that was performed. Each transient transfection was performed in triplicate (a-c) on the day of transfection and this was repeated a minimum of a further three more times (1-3) on separate occasions then averaged as shown in table above.

### **2.2.31 Database searches.**

To identify putative transcription factors and their binding sites, DNA sequences were submitted through ALGEN PROMO restricting searches to Homo sapiens factors only ([http://alggen.lsi.upc.es/cgi-bin/promo\\_v3/promo/promoinit.cgi?dirDB=TF\\_8.3](http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3)). This site is a division of TRANSFAC, a database on transcription factors (Wingender *et al.*, 2000).

To confirm protein identity based on N-terminal sequencing results. Amino acid sequences were submitted through BLAST (<http://www.ncbi.nih.gov/BLAST/>) limiting the search for proteins of “short nearly exact matches” and restricting the parameters to Homo sapiens only.

The final data from MALDI TOF/TOF analysis was submitted to the database search program Mascot (Matrix Science Ltd, London, UK) and the best match(es) were provided by the operator from Australian Proteome Analysis Facility (APAF, Macquarie University, Sydney, NSW, Australia).

### **2.2.32 DNA sequencing reactions.**

DNA sequencing reactions were performed at the Allan Wilson Centre (Massey University, Palmerston North, New Zealand) using Applied Biosystems BigDye terminator (v3.1) chemistry. Samples were provided in a total volume of 15  $\mu$ L, premixed with DNA and primer in concentrations specified by the operator.

# 3 DNA binding Proteins.

## 3.1 Synopsis.

As previously discussed, there are a number of proteins important for the transcriptional regulation of topoisomerase II $\alpha$ . In order to investigate the binding properties and the significance of these interactions, EMSAs were used to investigate the DNA sequences that are required for protein binding *in vitro* and transient transfections using a luciferase reporter gene driven by these DNA sequences were used to ascertain their functional significance *in vivo*. It has been established that NF-Y has a strict requirement for ICB1 whereas Sp1 and Sp3 bind to the GC1 element (Magan *et al.*, 2003). In addition to this, there is an unidentified protein complex that has been shown to bind to the region between the ICB1 and GC1 elements. This protein complex appears to have the properties of a GC element-associating factor but does not bind to the isolated GC1 element alone. In addition, the centre region between ICB1 and GC1 while being important for protein binding is insufficient on its own to support protein binding *in vitro* (Senior K., MSc Thesis, 2006). This chapter addresses the issue of this unidentified complex, whether it is a bonafide DNA-binding interaction and whether it is significant for the transcriptional regulation of human topoisomerase II $\alpha$ .

## 3.2 Protein complexes on the ICB1/GC1 composite element.

Sp1 has been shown to be able to self associate through its Q-rich domain, capable of forming tetramers and homo-multimeric structures which can associate with a single GC box (Mastrangelo *et al.*, 1991; Su *et al.*, 1991). Both NF-Y and Sp3 also have Q-rich domains (Li *et al.*, 2004; Mantovani, 1999) thus it is probable that protein-protein interactions between these three transcription factors may occur. In fact, it has been shown that Sp1, Sp3 and NF-YA can interact *in vitro* and *in vivo* (Yamada *et al.*, 2000) and there are several reports that NF-Y and Sp1 can physically interact (Liang *et al.*, 2001, Roder *et al.*, 1999 and Yamada *et al.*, 2000). Due to the non-denaturing properties of the gels used for EMSA experiments it is possible to demonstrate that protein complexes are formed *in vitro* at each of the important elements. Thereby each regulatory element may be responsible *in vivo* for recruitment of a set of proteins by means of protein-protein interactions and DNA-protein interactions. Antibody supershifts together with competitor assays have enabled the identification of specific Sp1, Sp3 and NF-Y interactions with the topoisomerase II $\alpha$  promoter, specifically at

the region encompassing ICB1 and GC1. However, even with excess antibody a complete shift in the respective bands in an EMSA gel has not been demonstrated. This suggests the formation of multi-protein complexes at regulatory elements of the topoisomerase II $\alpha$  promoter. These multi-protein complexes appear to contain Sp1, Sp3 and NF-Y as well as at least one more other unidentified protein.

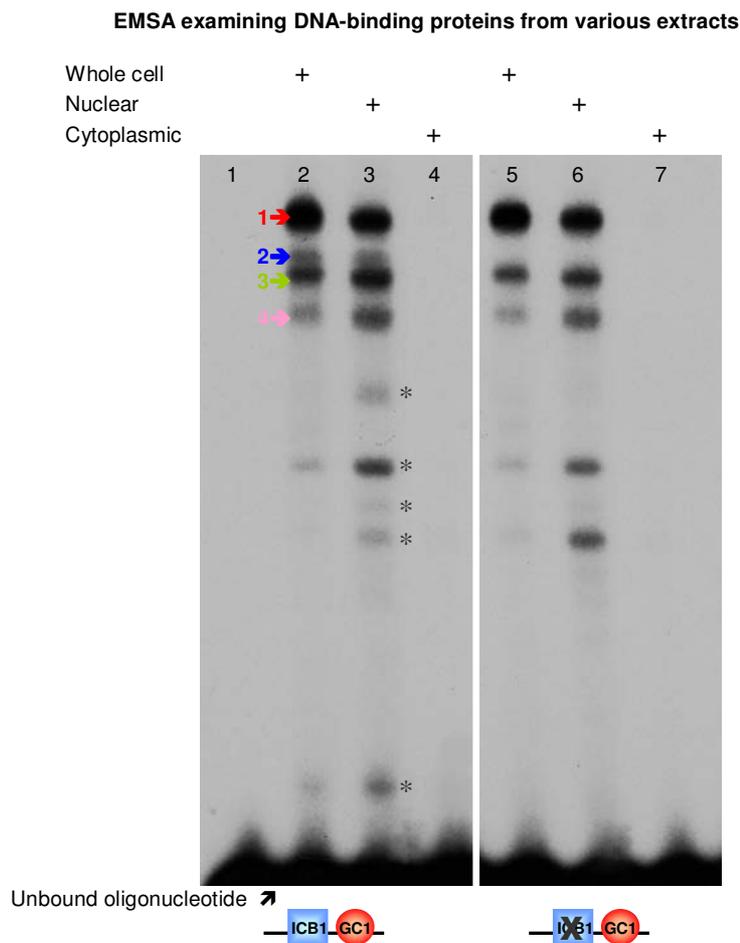
### **3.2.1 Whole cell and HeLa nuclear extracts.**

Many of the EMSA reactions done for this study and subsequent purification experiments that follow in chapter 4, 5 and 6 have been performed using whole cell HeLa extracts; therefore it was of particular importance to establish that the DNA-binding proteins were actually nuclear proteins (refer to figure 3.1). Preliminary analysis using western blots of nuclear extracts found that relative concentrations of Sp1, Sp3 and NF-Y compared to other proteins were higher in nuclear extracts than whole cell extracts (data not shown). These nuclear proteins were found to exhibit identical binding properties as those seen when using whole cell extracts from HeLa cells.

Figure 3.1 (left) represents a typical EMSA gel, where at least four protein-DNA complexes (indicated by the coloured arrows) are formed in the presence of the wild-type composite ICB1/GC1 oligonucleotide (refer to table 3.1 for oligonucleotide sequence). In the absence of any extract (lane 1, figure 3.1), a mobility shift is not observed and the free oligonucleotide that has not bound any protein can be seen at the bottom of the gel. When the ICB1 element is mutated within the oligonucleotide (figure 3.1, right) there are only three visible protein complexes, as the mutated oligonucleotide has lost the ability to bind NF-Y without any subsequent loss in binding of any of the other proteins. A mobility shift is not observed with cytoplasmic HeLa extracts (figure 3.1, lanes 4 and 7) indicating the DNA binding proteins were indeed nuclear, as expected.

In some EMSA gels, there are several high mobility bands (indicated with \* in figure 3.1 and 3.2). These high mobility bands often vary in intensity and mobility as different batches of extracts are used in EMSA. Quite commonly, as more extract is added to each assay these high mobility bands increase in frequency and intensity, and sometimes a trend can be seen where the mobility of these bands decreases as more extract is added (figure 3.3, circled in red). These high mobility complexes may represent proteins that are smaller in size compared

to the four major complexes under examination, but the variability in these high mobility bands makes it difficult to draw any real conclusions on their composition or their presence. Due to the observation that when antibody is added to EMSA (refer to bands \* compare lane 1 with lanes 2-5, figure 3.2A), there is a super-shift in these high mobility bands, regardless of which antibody is added. It is possible that these high mobility bands represent non-specific interactions, but the presence of them can not be completely ignored as they still represent protein-DNA complexes.



**Figure 3.1: DNA binding proteins are nuclear proteins.**

EMSAs were carried out as described in sections 2.2.23 and 2.2.24. Approximately 2 µg of whole cell, nuclear and cytoplasmic HeLa extract was added to approximately 0.5 ng of each <sup>32</sup>P-oligonucleotide. The oligonucleotides used in each assay are annotated below the figure (from left to right) ICB1wt/GC1wt and ICB1mt/GC1wt. The gel was dried onto DE-81 paper and exposed to X-ray film for about 10 hours at -80°C. Each of the coloured arrows illustrates specific protein-DNA complexes analysed in this study. The free oligonucleotide with no protein bound is located at the bottom of the gel. \* Denotes the high mobility protein-DNA complexes thought to be non-specific interactions. This figure is representative of at least three experiments.

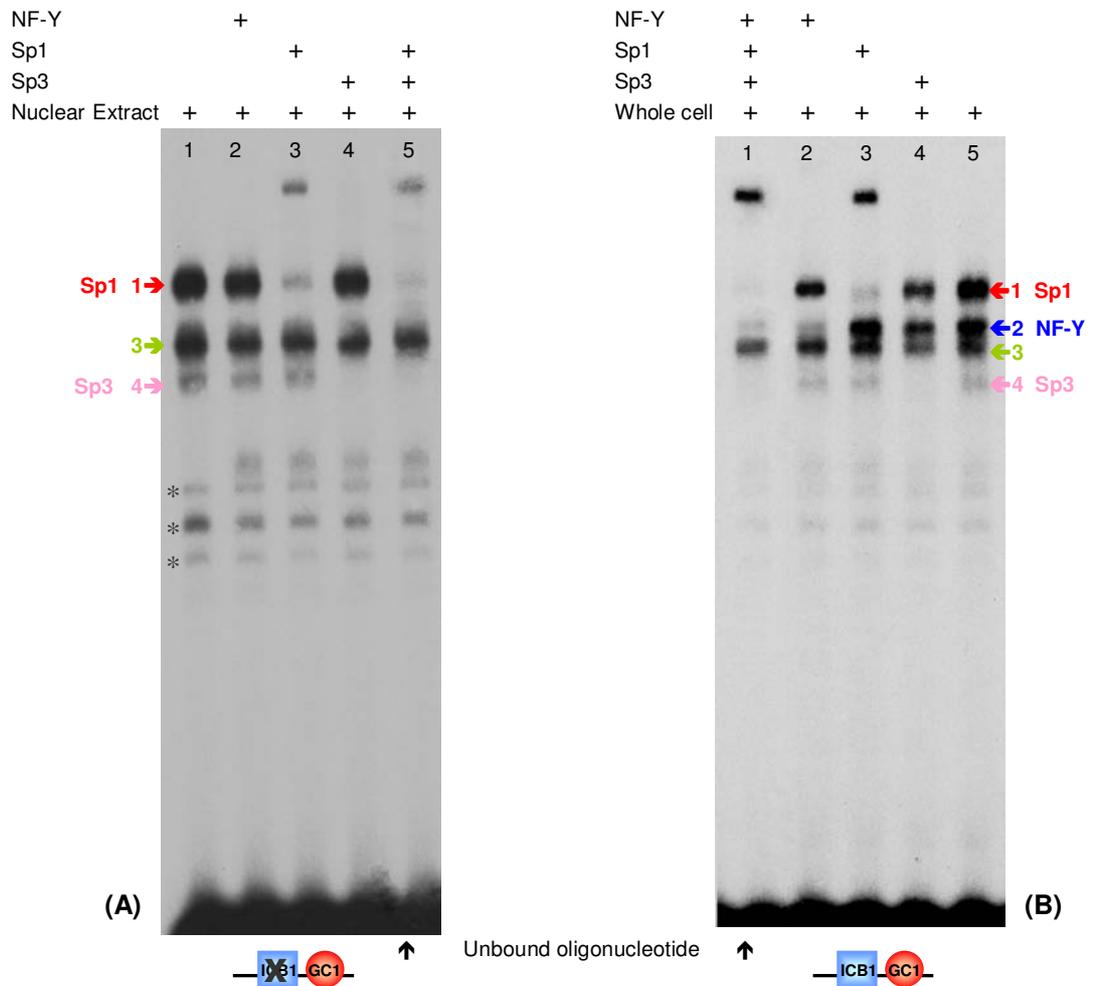
### 3.2.1.1 Antibody Supershifts.

The addition of antigen specific antibody to an EMSA reaction is a very useful technique that enables the identification of protein(s) that are present in a protein-DNA complex. As the antibody associates with its respective protein, the observed protein-DNA complex becomes heavier and causes a further mobility shift (super-shift), thereby making identification possible. In some cases the addition of antibody inhibits binding of protein to DNA as it can mask the DNA binding site, but the overall outcome is still a reduction in intensity of the respective band representing the protein-DNA complex.

When antibodies were added to EMSA reactions, it was confirmed that the same protein-DNA complexes produced in whole cell HeLa extracts were also present in HeLa nuclear extracts. Figure 3.2 shows that Sp1 is present in band 1 (red arrow, figure 3.2A and B lanes 3), Sp3 is present in the fourth band (pink arrow, figure 3.2A and B lanes 4) and NF-Y is present in the second band (blue arrow, figure 3.2B lane 2). Of particular interest is the common trend that is observed in many of these antibody supershift gels; where even with excess antibody a complete shift is not observed in the respective band. This is most obvious in lane 3 (figure 3.2A) where a total of 1.4  $\mu\text{g}$  of Sp1 antibody and only 1  $\mu\text{g}$  of total protein was used, a shift in band 1 (red arrow) is observed but a faint band remains representing residual protein-DNA interaction. This residual band was reduced further upon the addition of Sp3 as well as Sp1 antibody (lane 5, figure 3.2A) and almost completely disappeared when all three antibodies were added (lane 1, figure 3.2B). This suggests a complex of all three proteins may be present in band 1, although it is predominantly made up of Sp1. The addition of antibody against NF-Y causes a marked decrease in band 2 (lane 2, figure 3.2B). This band is predominantly composed of NF-Y, however presence of NF-Y cannot be completely excluded from complex 1 as a decrease in band 1 is also observed (lane 2 compared to lane 1 figure 3.2A, or figure 3.2B lane 2 compared to lane 5) no antibody present). Sp3 appears to be the only protein present in band 4 (pink arrow), as antibody against Sp3 completely removes this band (lane 4 figure 3.2A and B), however, there is also a (faint but reproducible) reduction in band 3 (green arrow, figure 3.2B, lane 4), as well band 1 (red arrow, lane 4 figure 3.2A and B) suggesting the presence of Sp3 in complexes one and three even though it is the major component in band four. This outlines the possibility that all three identified transcription factors may be interacting at multiple elements within the topoisomerase II $\alpha$  promoter. Taken together, these data suggests the presence of multi-

protein complexes forming on the composite ICB1/GC1 oligonucleotide *in vitro*. These protein-protein interactions may be crucial in regulation of the topoisomerase II $\alpha$  promoter *in vivo*.

**Antibody supershifts to show protein-DNA complexes are made up of multiple proteins.**



**Figure 3.2: Antibody supershifts to verify protein-DNA complexes.**

A total of 1.4  $\mu\text{g}$  antibody was added to 1  $\mu\text{g}$  of nuclear (A) or 1.5  $\mu\text{g}$  of whole cell (B) HeLa extracts, along with approximately 0.5ng of each  $^{32}\text{P}$ -oligonucleotide in EMSA reactions. The oligonucleotides used in each assay are annotated below the figure. Each gel was dried using DE-81 paper and exposed to X-ray film for 12-18 hours at  $-80^{\circ}\text{C}$ . The coloured arrows indicate specific protein-DNA complexes analysed in this study and the major protein present in the complex is labelled. The free oligonucleotide is located at the bottom of the gel. Non-specific protein-DNA complexes of high mobility are highlighted with an \*. These gels are representative of duplicate experiments.

There are also several high mobility bands (indicated with \*) seen in figure 3.2A. These have been discussed in section 3.1.1, and may represent smaller proteins or complexes forming with the oligonucleotide. The observed super-shift seen in these high mobility bands (lanes 2-5) appears to be non-discriminate of the antibody used in each assay, which provides further evidence that these complexes may be non-specific interactions.

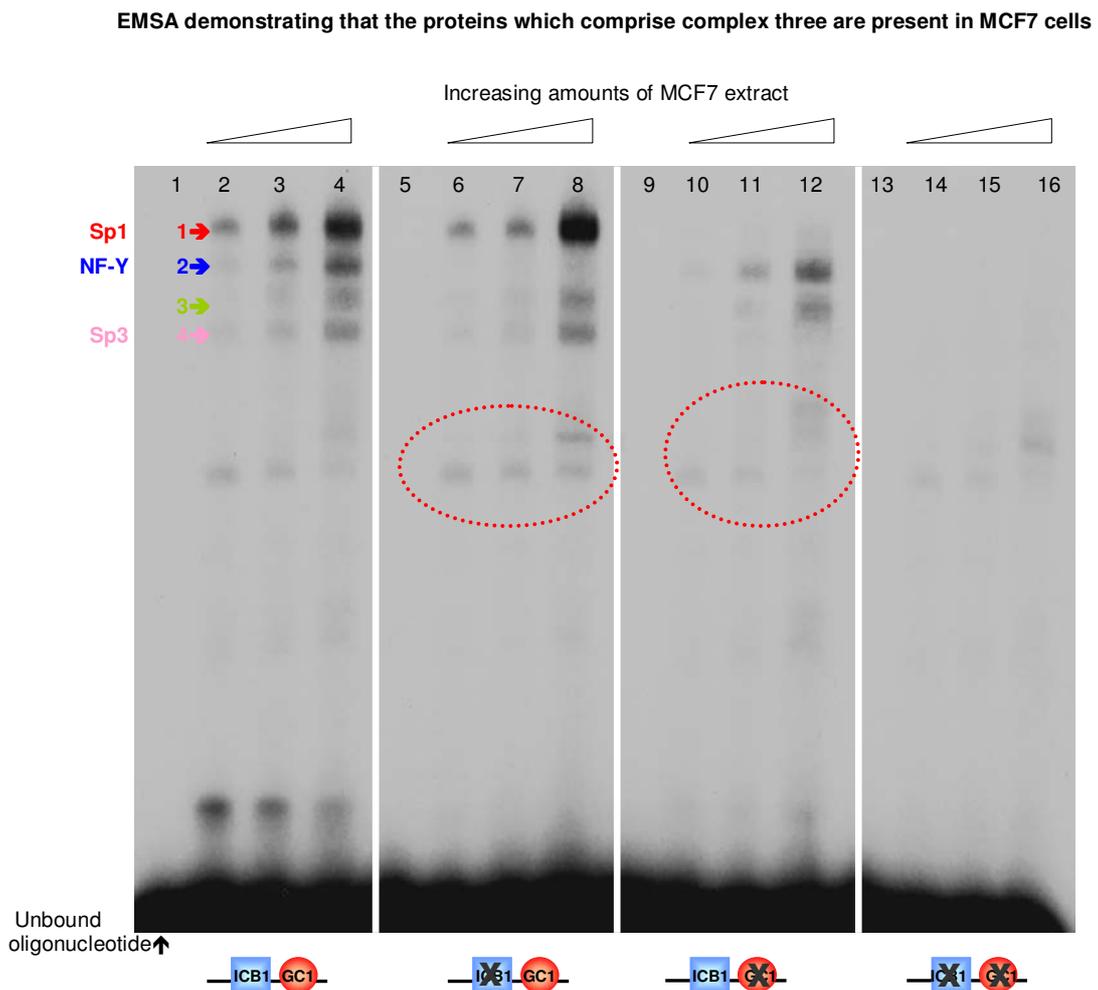
### 3.2.2 Protein complexes in other cell lines.

Part of establishing the functional importance of these DNA binding proteins was to investigate their presence and binding to the ICB1/GC1 topoisomerase II $\alpha$  sequence in other cell lines. Therefore DNA-binding proteins were examined using whole cell extracts from MCF7 (human mammary adenocarcinoma) and MCF12A (normal breast epithelial) cell lines.

Figure 3.3 represents an EMSA with increasing amounts of MCF7 extract using various oligonucleotides; these binding patterns are very typical and mimic the patterns previously established using HeLa extracts (Magan *et al.*, 2003). Four protein complexes are seen when the oligonucleotide is not mutated (lanes 2-4) and complex two (blue arrow) is missing when the ICB1 element is mutated (lanes 6-8). Bands one (red arrow) and four (pink arrow), complex one and four respectively, representing the GC-associating proteins Sp1 and Sp3 are diminished when the GC1 element is mutated (lanes 10-12). Complex 3 (green arrow) is also present when the GC1 element is mutated (compare ICB1wt/GC1wt oligonucleotide in lane 4 with ICB1wt/GC1mt oligonucleotide in lanes 11 and 12). Although it is not obvious in figure 3.3, often less complex 3 is seen to form with the ICB1wt/GC1mt oligonucleotide compared with the ICB1wt/GC1wt oligonucleotide (data not shown). Finally, no specific protein-DNA interactions are observed when both ICB1 and GC1 elements are mutated (lanes 14-16). Significantly, complex three is observed in MCF7 extracts, albeit in relatively lower abundance than in HeLa cells and it behaves as a GC associating protein with the composite oligonucleotide (lanes 6-8) as previously observed in HeLa cell extracts.

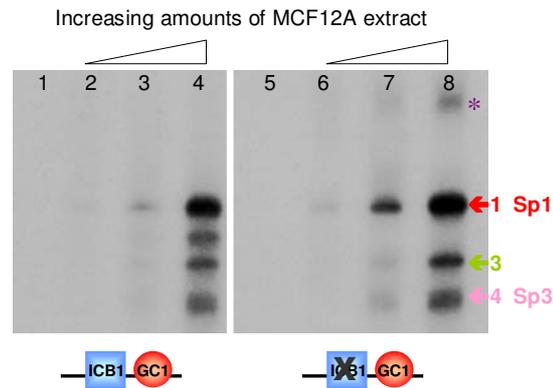
Using MCF12A cells (figure 3.4) the same binding patterns were also observed as already described with HeLa and MCF7 extract. Mutation of the ICB1 element results in a loss of NF-Y binding (lanes 5-8) without subsequent reduction in binding of the other transcription factors. Again complex three is present and binds consistently when the GC1 element is

intact. It is interesting to note the presence of an additional lower mobility band (\*) in lane 8 (figure 3.4), this band is quite often seen when extract is added in excess in EMSA reactions. It is not simply specific to MCF12A cell extracts, but has also been observed with HeLa (Magan N., MSc Thesis, 2002) and MCF7 cell extracts (data not shown) and most likely represents the formation of a larger heavier protein-DNA complex. Sp1 can bind as multimers, so it is possible that this heavier complex contains multiple Sp1 components, attributing to the observed decrease in mobility.



**Figure 3.3: Increasing amount of MCF7 extract added to different ICB1/GC1 oligonucleotides.** Increasing amounts (0, 1, 2 or 4 µg) of MCF7 whole cell extract was added to approximately 0.5ng of each <sup>32</sup>P-oligonucleotide in EMSA reactions. The oligonucleotides used in each assay are annotated below the figure (from left to right) ICB1mt/GC1wt, ICB1wt/GC1wt, ICB1wt/GC1mt and ICB1mt/GC1mt. The gel was dried onto DE-81 paper and exposed to X-ray film for about 25 hours at -80°C. The coloured arrows indicate specific protein-DNA complexes and the major protein in each complex is annotated. The non-specific, high mobility protein-DNA complexes are circled in red. The free oligonucleotide with no protein bound is located at the bottom of the gel. This figure is representative of duplicate experiments.

## EMSA demonstrating that the proteins which comprise complex three are present in MCF12A cells



**Figure 3.4: EMSA using increasing amount of MCF12A extract.**

For simplicity, only the top half of an EMSA gel is shown above. Increasing amounts (0, 1, 2 or 4  $\mu\text{L}$ ) of 0.5  $\mu\text{g}/\mu\text{L}$  of MCF12A extract was added to approximately 0.5 ng of each  $^{32}\text{P}$ -topoisomerase II $\alpha$  oligonucleotide in EMSA reactions. The oligonucleotides used in each assay are annotated below the figure (from left to right) ICB1wt/GC1wt and ICB1mt/GC1wt. Each gel was dried using DE-81 paper and exposed to X-ray film for about 48 hours at  $-80^{\circ}\text{C}$ . The coloured arrows indicate specific protein-DNA complexes and the major protein found in each complex is described. The \* highlights an additional low mobility band seen when excess extract is added to the assay. This gel is representative of duplicate experiments.

Antibody supershifts were also carried out using extracts from MCF7 and MCF12A cell lines (data not shown), confirming that the composition of the bands were as described in HeLa extracts. However the proportion of each of the protein-DNA complexes was found to vary between cell types, indicating that there are some differences between these cell lines. Collectively, these EMSAs demonstrate that both MCF7 and MCF12A extracts can form the same overall protein-DNA complexes on the ICB1/GC1 composite oligonucleotide *in vitro*. Of major significance is the presence of complex three in both MCF7 and MCF12A cells, and its binding properties are consistent across the cells lines, indicating this is not an interaction which is an artefact belonging to HeLa cells only.

### 3.3 Investigating binding patterns of complex three.

The region between ICB1 and GC1 is exactly 12bp in size which represents slightly more than a complete turn of the DNA double helix. Therefore sufficient space should be available for a transcription factor to bind to the intervening sequence. However, EMSA experiments

using only the intervening sequence between the ICB1 and GC1 elements have demonstrated that no proteins are capable of binding to this region alone *in vitro* (Senior K., MSc Thesis 2006). Significantly, complex three has been shown to be dependent on the presence of an intact GC box, as the ICB1wt/GC1mt composite oligonucleotide has reduced ability to bind complex three meanwhile, a single, shorter oligonucleotide with only the GC1 or ICB1 element is unable to bind complex three (Magan *et al.*, 2003). This suggests that the proximal elements are important, not only for their ability to recruit transcription factors, but also to enable protein-protein interactions that may engage other proteins at other sites along the promoter.

### 3.3.1 Examining the intervening sequence between ICB1 and GC1.

To address the issue of where complex three is forming on the oligonucleotide, the composite ICB1/GC1 oligonucleotide was modified (see table 3.1 for details) to accommodate a larger centre region. The 6 bp sequence “GTC TGC” within the intervening region of the ICB1wt/GC1wt oligonucleotide is important, as previous work has shown that when this sequence is mutated or removed there is almost a complete loss of binding of complex three *in vitro* (Senior K., MSc Thesis 2006). This suggests that this 6 bp sequence may act as the core to initiate binding of complex three and was therefore used to lengthen the intervening sequence at two different locations.

Oligonucleotide	Sequence (5'-3')
ICB1wt/GC1wt	CGA GTC AGG GAT TGG CTG GTC TGC TTC GGG CGG GCT AAA GGA AG
ICB1mt/GC1wt	CGA GTC AGG GAT Tcc CTG GTC TGC TTC GGG CGG GCT AAA GGA AG
ICB1wt/CENins/GC1wt	CGA GTC AGG GAT TGG CTG GTC* gtc tgc TGC TTC GGG CGG GCT AAA GGA AG
ICB1mt/CENins/GC1wt	CGA GTC AGG GAT Tcc CTG GTC gtc tgc TGC TTC GGG CGG GCT AAA GGA AG
ICB1wt/CENins/GC1mt	CGA GTC AGG GAT TGG CTG GTC gtc tgc TGC TTC GtG Cgt GCT AAA GGA AG
ICB1wt/repeatseq/GC1wt	CGA GTC AGG GAT TGG CTG GTC TGC* gtc tgc TTC GGG CGG GCT AAA GGA AG
Short GC1	C TGC TTC GGG CGG GCT AAA G
ICB1wt/CENdel/GC1wt	CGA GTC AGG GAT TGG CTG TTC GGG CGG GCT AAA GGA AG

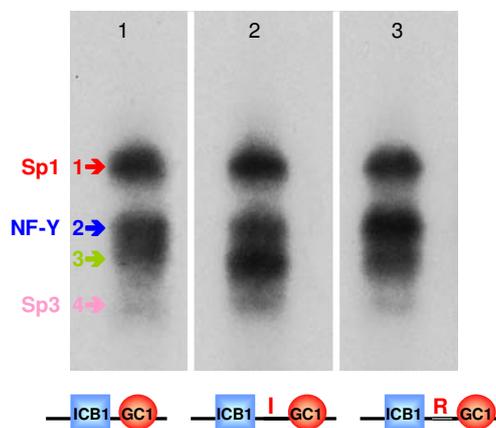
**Table 3.1: Sequences of oligonucleotides used in EMSAs**

The different oligonucleotide sequences are shown in the table above. For simplicity, only the top strand of DNA is shown for each of the oligonucleotide pairs. The ICB1 box is highlighted in blue and the GC1 element is pink. Alterations in the sequences are indicated in lower case font and the repeated sequences in the intervening region are indicated in green. The major difference between ICB1wt/CENins/GC1wt and ICB1wt/repeatseq/GC1wt is the location of the repeated sequence indicated by the \*.

Three different oligonucleotides were compared for protein-binding in vitro using EMSA; a standard ICB1wt/GC1wt, a lengthened ICB1wt/CENins/GC1wt and a lengthened ICB1wt/repeatseq/GC1wt. The difference between the two longer oligonucleotides is the location of the inserted sequence; ICB1wt/CENins/GC1wt has the 6 bp insertion closer to the ICB1 box and ICB1wt/repeatseq/GC1wt has the 6 bp insertion placed closer to the GC1 element (oligonucleotide sequences are listed in table 3.1).

In attempts to identify any possible transcription factors that may be binding to the original intervening sequence and the extended intervening sequences, the oligonucleotide sequences were analysed for putative transcription factor binding sites using ALGEN PROMO ([http://algggen.lsi.upc.es/cgi-bin/promo\\_v3/promo/promoinit.cgi?dirDB=TF\\_8.3](http://algggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3)). This site is a division of TRANSFAC, a database on transcription factors (Wingender *et al.*, 2000). The search was restricted to incorporate only human transcription factor sites and human factors and each of the three intervening sequences from ICB1wt/GC1wt, ICB1wt/CENins/GC1wt and ICB1wt/repeatseq/GC1wt were tested. Unfortunately, no possible human transcription factor binding sites were identified.

**EMSA to show increased binding of complex 3 to CENins oligonucleotide**



**Figure 3.5: Investigating the intervening sequence for protein binding.**

For simplicity, only the top half of an EMSA gel is shown. EMSA were carried out as described in 2.2.23 and 2.2.24 using 3  $\mu\text{g}$  of whole cell HeLa extract with approximately 0.5 ng of each  $^{32}\text{P}$ -oligonucleotide in EMSA reactions. Each oligonucleotide used is indicated at the bottom of gel picture (from left to right) ICB1wt/GC1wt, ICB1wt/CENins/GC1wt and ICB1wt/repeatseq/GC1wt. Each gel was dried using DE-81 paper and exposed to X-ray film for about 15 hours at  $-80^{\circ}\text{C}$ . Protein-DNA complexes are indicated by the coloured arrows and the major protein found in each complex is described. This gel is representative of triplicate experiments.

The results of an *in vitro* DNA-binding assay are shown in figure 3.5. The same standard four protein-DNA complexes are observed with each oligonucleotide; however there do appear to be some differences in the amount of complex 3 that is seen to form. Figure 3.5 shows that the ICB1wt/CENins/GC1wt produces a more intense band in the position of complex three (green arrow, lane 2) than do the other two oligonucleotides. The mobility of complex three is slightly different to that seen with the ICB1wt/GC1wt oligonucleotide (lane 1); however this is most likely due to the size differences of the oligonucleotides, as the modified oligonucleotides are 6bp longer. Mobility may also be compromised as more protein binds to the oligonucleotide making the complex larger. Examining lane 3 (figure 3.5) it appears that the ICB1wt/repeatseq/GC1wt oligonucleotide binds more of complex two (containing NF-Y) than the other two oligonucleotides, however it still binds complex three and complex one at the same relative intensity as the ICB1wt/GC1wt oligonucleotide (lane 1, figure 3.5). This demonstrates that the ICB1wt/repeatseq/GC1wt oligonucleotide binds proteins in approximately the same manner as the ICB1wt/GC1wt oligonucleotide. These results suggest that the location of the inserted sequence is important for protein binding and this may be due to factors such as its proximity to the GC or ICB elements, or the sequences flanking the inserted “GTC TGC” sequence. Taken together the data suggests that simply lengthening the intervening region does not increase complex 3 binding, but that the binding is both sequence specific and location specific. Overall, this EMSA clearly shows that ICB1wt/CENins/GC1wt is capable of binding more of complex three than any of the other oligonucleotides tested.

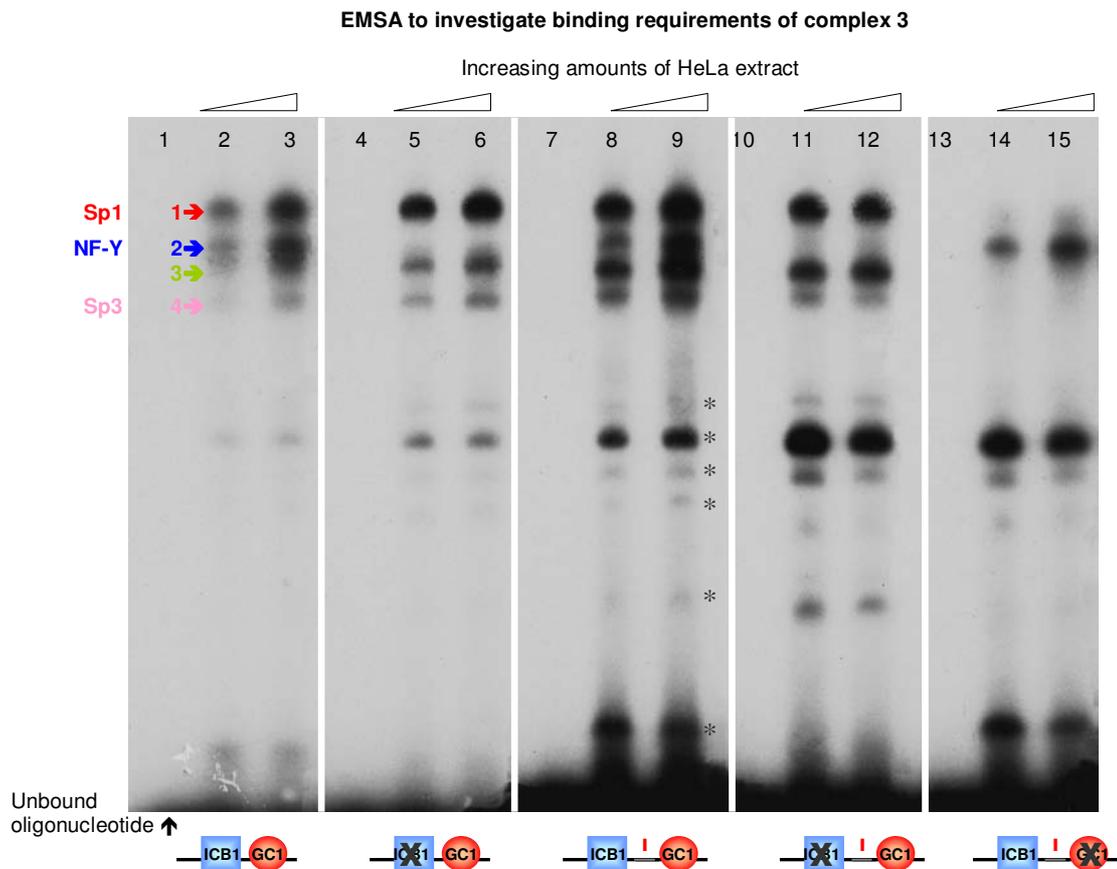
### **3.3.2 Does ICB1wt/CENins/GC1wt bind complex three?**

Due to the mobility of complex three being slightly different with the longer ICB1wt/CENins/GC1wt oligonucleotide, it was important to demonstrate that the proteins in this complex had the same binding properties as those previously observed with complex three using the shorter ICB1wt/GC1wt oligonucleotide. Antibody supershifts using the ICB1wt/CENins/GC1wt oligonucleotide (Senior K., MSc Thesis 2006), have indicated that Sp1, Sp3 and NF-Y bound to the ICB1wt/CENins/GC1wt oligonucleotide in the same complexes. However, as the proteins that compose complex 3 are not currently identified, individual proteins in this complex can not be positively identified in this manner. Therefore, variations of the ICB1wt/CENins/GC1wt oligonucleotide were created to investigate the

binding parameters of this complex that appeared in the same vicinity as complex 3 in EMSA gels.

### 3.3.2.1 Using a mutated ICB1wt/CENins/GC1wt oligonucleotide.

The ICB1wt/CENins/GC1wt oligonucleotide was mutated within the ICB1 or GC1 elements to observe the binding patterns (table 3.1 for oligonucleotide sequences) and specificity of complex 3 to these oligonucleotides.



**Figure 3.6: Complex three is the same using the ICB1wt/CENins/GC1wt oligonucleotide.**

Increasing amounts of HeLa extract (0, 2, and 4 µg HeLa whole cell extract) was added to approximately 0.5 ng of each <sup>32</sup>P-oligonucleotide in EMSA reactions. The oligonucleotides used in each assay are annotated below the figure (left to right) ICB1wt/GC1wt, ICB1mt/GC1wt, ICB1wt/CENins/GC1wt, ICB1mt/CENins/GC1wt, ICB1wt/CENins/GC1mt. The gel was dried using DE-81 paper and exposed to X-ray film for about 18 hours at -80°C. Protein complexes are indicated by the coloured arrows and the major protein found in each complex is described. \* Represents some high mobility protein-DNA complexes thought to be non-specific interactions and any unbound oligonucleotide can be seen at the bottom of the gel. This gel is representative of triplicate experiments.

Figure 3.6 shows that complex three binding to the longer ICB1wt/CENins/GC1wt oligonucleotide exhibits that same binding patterns that have been seen with the shorter original ICB1wt/GC1wt oligonucleotide (Magan N., MSc Thesis, 2002). Strict requirement for the GC1 element is observed; as when the GC1 element is mutated (lanes 14 and 15, figure 3.6) there is a complete loss of complexes 1, 3 and 4. Both MCF7 and MCF12A extracts were also tested with the ICB1wt/CENins/GC1wt oligonucleotide and the binding patterns were identical as seen in figure 3.6 (data not shown).

In figure 3.6 it is interesting to note again the presence of the high mobility bands (highlighted with \*) looking across from lane 2 through to lane 15, it almost appears as though the bands increase in intensity and frequency as less specific binding is observed (due to mutations in the functional binding sequences of the oligonucleotides) over the four complexes that are examined for this study. In addition, more of these high mobility bands (and in greater intensity) appear with the oligonucleotides that bind more overall protein such as the CENins oligonucleotides, indicating that these longer oligonucleotides bind more overall protein than their shorter ICB1/GC1 counterparts.

### **3.3.2.2 Using a short GC1 competitor.**

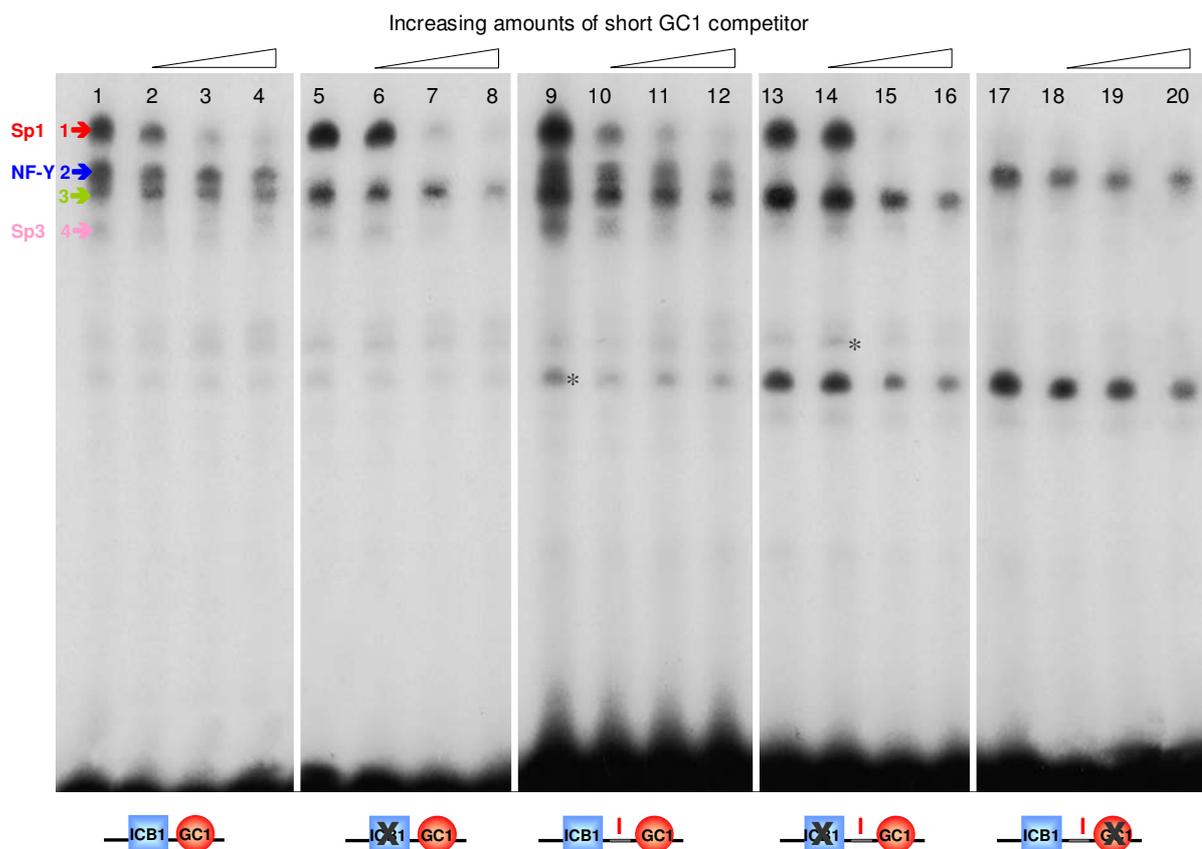
Commonly, unlabeled double-stranded oligonucleotides can be used as competitors in EMSA. In these assays, a decrease in band intensity is observed as the competitor is added in excess and sequesters the proteins from binding to the <sup>32</sup>P-labelled oligonucleotide. As complex three has been previously shown to have a weak affinity for the short GC1 competitor in EMSA, this was again used as a competitor to confirm binding and specificity of complex 3 to the longer oligonucleotide.

Figure 3.7 shows a competitor assay using various labelled oligonucleotides, with increasing amount of the short GC1 competitor (table 3.1 for oligonucleotide sequence). Lanes 1, 5, 9, 13 and 17 contain no competitor and therefore depict the “normal” amount of protein-DNA binding for each oligonucleotide and are used as controls for comparison. It is again obvious that the ICB1wt/CENins/GC1wt oligonucleotide binds more of complex three (and the mobility is also slightly different) than of the ICB1wt/GC1wt oligonucleotide (comparing lanes 9 and 1). In addition, when the ICB1 element is mutated a standard binding pattern is observed (comparing lanes 5 and 13), where complex 2 (containing NF-Y) is absent but

complex 1, 3 and 4 remain. Complex three does not bind very strongly to the short GC1 competitor, having a greater preference for the wild-type versions of the composite oligonucleotides. The evidence for this is in lanes 1-4, and lanes 9-12 where bands 1 and 4 are greatly reduced however bands 2 and 3 remain. However, complex 3 is competed out with the GC1 short competitor when using both the oligonucleotides with a mutated ICB1 element (compare lanes 5-8 and lanes 13-16). Overall, the GC1 competitor is capable of binding Sp1 (complex 1, red arrow) and Sp3 (complex 4, pink arrow) with high affinity, as these bands are greatly reduced in the presence of increasing amounts of GC1 competitor, independent of which labelled oligonucleotide is used. Interestingly, the GC1 competitor does appear to bind NF-Y (blue arrow, complex 2) to some degree, as a reduction in band two can be seen in lanes 17-20. This is most likely due to protein-protein interactions and the fact that competitor in a large excess can have an inhibitory role on binding in some circumstances. Taken together this suggests that complex three, may contain some GC associating factors, such as Sp3 (also indicated in antibody supershifts) but the proteins in complex three still have a greater preference for the longer intact oligonucleotide than the shorter GC1 competitor. This provides further evidence that the proteins in complex three appear the same, regardless of whether the ICB1wt/CENins/GC1wt or ICB1wt/GC1wt oligonucleotide is used.

Again there is the presence of multiple high mobility bands in this EMSA gel (figure 3.7, highlighted with \*). However, they appear to reduce in intensity with increasing amounts of GC1 competitor. But they do not appear to form in exactly the same vicinity or intensity as other observed non-specific high mobility bands from EMSA gels. With the addition of a non-specific competitor (designed with neither an ICB nor GC element), there is a slight reduction in only these high mobility bands (data not shown), further highlighting that these bands are most likely non-specific protein-DNA interactions.

### EMSA with GC1 competitor and CENins oligonucleotide.



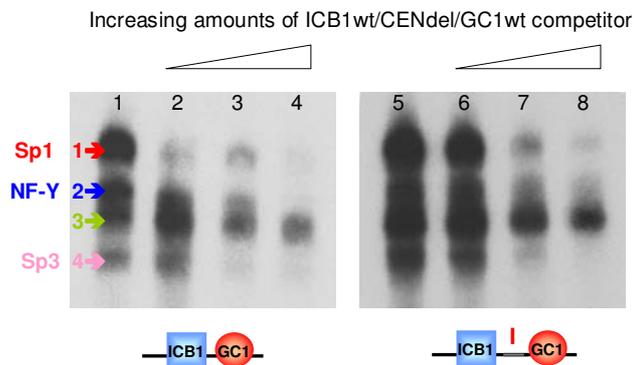
**Figure 3.7: EMSA using a short GC1 oligonucleotide as a competitor.**

Increasing amounts of unlabelled short GC1wt competitor (5, 50 and 100 ng) was added to 2 µg of HeLa whole cell extract, along with approximately 0.5 ng of each <sup>32</sup>P-oligonucleotide in EMSA reactions. The oligonucleotides used in each assay are annotated below the figure (from left to right) ICB1wt/GC1wt, ICB1mt/GC1wt, ICB1wt/CENins/GC1wt, ICB1mt/CENins/GC1wt and ICB1wt/CENins/GC1mt. Each gel was dried using DE-81 paper and exposed to X-ray film for about 15 hours at -80°C. \* illustrates the high mobility protein-DNA complexes thought to be non-specific interactions. These results are representative of duplicate experiments.

#### 3.3.2.3 Using an ICB1wt/CENdel/GC1wt competitor.

Previous EMSAs using an oligonucleotide with a deletion in the intervening sequence (table 3.1, ICB1wt/CENdel/GC1wt) have indicated that this oligonucleotide was unable to efficiently bind complex 3 (Senior K., MSc Thesis 2006), therefore this oligonucleotide was used as a competitor to distinguish that the intervening region, specifically the “GTC TGC” 6 bp sequence was important for the binding of complex 3 to a region within the human topoisomerase II $\alpha$  promoter.

### EMSA with competitor that does not bind complex 3 efficiently



**Figure 3.8: EMSA using a ICB1wt/CENdel/GC1wt competitor.**

For simplicity, only the top half of an EMSA gel is shown. Increasing amounts of unlabelled ICB1wt/CENdel/GC1wt competitor (5, 50 and 100 ng) was added to 3 µg of HeLa whole cell extract, along with approximately 0.5 ng of each <sup>32</sup>P-oligonucleotide in EMSA reactions. The oligonucleotides used in each assay are annotated below the figure (from left to right) ICB1wt/GC1wt and ICB1wt/CENins/GC1wt. Each gel was dried using DE-81 paper and exposed to X-ray film for about 18 hours at -80°C. Protein complexes investigated in this study are indicated with the coloured arrows and the major protein found in each complex is described. This figure is representative of duplicate gels.

Figure 3.8 demonstrates the effects of adding increasing amounts of ICB1wt/CENdel/GC1wt oligonucleotide as a competitor to protein-DNA binding reactions. As before, greater binding of complex 3 was observed using the ICB1wt/CENins/GC1wt oligonucleotide (lane 5, figure 3.8) than the ICB1wt/GC1wt oligonucleotide (lane 1, figure 3.8). Actually, in this figure it also appears as though the ICB1wt/CENins/GC1wt oligonucleotide binds overall more protein than the shorter counterpart as most of the bands are in greater intensity in lane 5 and 6 than in lane 1 (ICB1wt/GC1wt oligonucleotide). In the presence of the ICB1wt/CENdel/GC1wt competitor, there is an overall reduction in intensities of the bands representing complex 1, 2 and 4 indicating that Sp1, Sp3 and NF-Y are capable of binding to the ICB1wt/CENdel/GC1wt oligonucleotide, as expected. There is also a decrease in band 3 (complex 3, green arrow) in the presence of the ICB1wt/CENins/GC1wt competitor, however the decrease is not as significant as that seen with complex 1, complex 2 and

complex 4. In order to achieve a significant decrease in complex 3 the competitor has to be present in the equivalent of approximately 100-200 times excess (lanes 3-4 and 7-8, respectively) of the labelled oligonucleotide, suggesting that the competitor does not efficiently bind the proteins that are required to form complex 3. This is in agreement with other results (Senior K., Msc Thesis, 2006) and demonstrates the importance of the “GTC TGC” sequence. Together with the results of both the competitor assays (figure 3.7 and 3.8) this collectively suggests that complex 3 is most likely binding to the intervening sequence and not the GC1 element of the ICB1wt/GC1wt oligonucleotide *in vitro*.

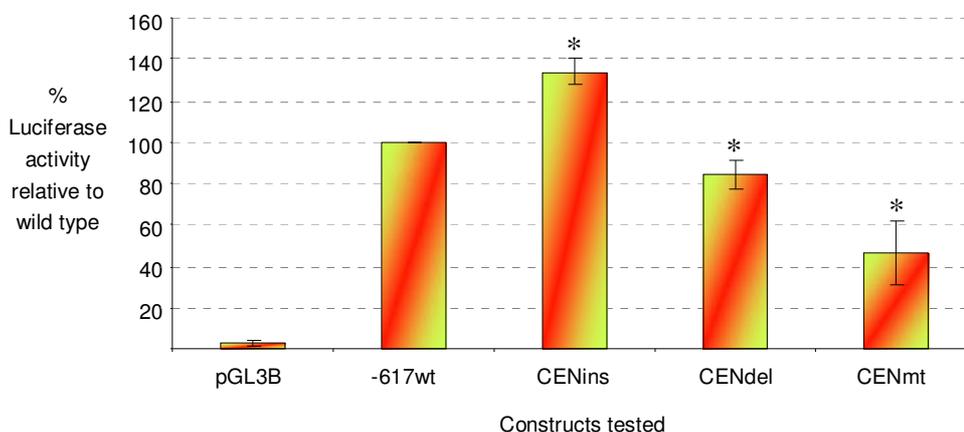
### 3.4 Functional Significance of Complex three.

Previous research has shown that a mutation or deletion in the “GTC TGC” sequence within the intervening region can significantly hinder binding of complex 3 (Senior K., MSc Thesis, 2006). So far, this study has demonstrated that by altering the ICB1/GC1 composite oligonucleotide to have a site-specific sequence insertion within the intervening region of these two elements (ICB1wt/CENins/GC1wt) it is possible to enhance for the binding of complex 3 *in vitro*. Even though at this stage it is not known which proteins may be involved in the formation of this complex, it was important to ascertain the effect this extra binding may have on topoisomerase II $\alpha$  promoter activity *in vitro*. The promoter containing most of the major regulatory sites (-617 bp upstream of the transcriptional start site) of human topoisomerase II $\alpha$  has already been defined and cloned (Isaacs *et al.*, 1996). Using this -617 human topoisomerase II $\alpha$  promoter sequence, the intervening sequence between ICB1 and GC1 was extended by inserting an additional 6bp sequence “GTC TGC” (CENins), or deleted by removing the “GTC TGC” 6bp sequence (CENdel), or mutated by changing two bases within the “GTC TGC” sequence to “GCC TAC”(CENmt). These altered -617topo2 $\alpha$  promoter constructs were created by PCR mutagenesis (Senior K., Msc Thesis, 2006) and cloned into a pGL3B vector upstream of a luciferase reporter gene. Before using these reporter gene constructs in transient transfections, plasmid identity was confirmed by sequencing using standard primers GL2 and RV3 which bind to sequences flanking the multiple cloning site within the pGL3B plasmid.

Luciferase activity is directly correlated to the strength of the topoisomerase II $\alpha$  promoter used in the reporter gene construct. Figure 3.9 summarises the results of the transient transfections and illustrates the changes in luciferase activity when alterations are made in

the topoisomerase II $\alpha$  promoter. Each data set is expressed as relative luciferase activity, to be able to monitor any changes relative to wt -617topo2 $\alpha$  promoter. The CENins construct has elevated promoter activity, whereas the CENdel and CENmt constructs have decreased promoter activity compared to the wt -617topo2 $\alpha$  construct. Together with the results from EMSA experiments, this suggests that when there is more complex three binding to the topoisomerase II $\alpha$  promoter, activity increases about 30%. However as complex three does not efficiently bind to either CENdel or CENmt, topoisomerase II $\alpha$  promoter activity decreases by 20-50%. The statistical significance of the differences observed between two data sets was verified using the statistical *t*-test, signifying that there is more than 95% confidence that each of the altered -617topo2 $\alpha$  constructs (CENins, CENdel and CENmt) is different to the wt -617topo2 $\alpha$  construct. Taken together this suggests that complex three binding to the topoisomerase II $\alpha$  promoter may have an *in vivo* role in gene activation and is therefore of functional significance.

**THE EFFECTS OF CHANGING THE INTERVENING SEQUENCE BETWEEN ICB1 AND GC1 ON TOPOISOMERASE II $\alpha$  PROMOTER ACTIVITY**



Construct	Luciferase activity relative to WT (%)	p-values (% similarity to wt)	
pGL3B	2.7 ± 1.5%		
-617topo2 $\alpha$ wt	100 %		
-617topo2 $\alpha$ CENins	134 ± 6 %	0.002	0.2%
-617topo2 $\alpha$ CENdel	84 ± 7 %	0.021	2%
-617topo2 $\alpha$ CENmt	46 ± 15 %	0.007	0.7%

**Figure 3.9: Transient transfections using reporter gene constructs that have alterations in the -617 minimal topoisomerase II $\alpha$  promoter.**

HeLa cells were transfected with 0.25  $\mu$ g of pCMV SPORT- $\beta$ -gal plasmid along with 0.5  $\mu$ g of the different -617topo2 $\alpha$  pGL3B reporter constructs as described in section 2.2.27-2.2.30. An empty pGL3B plasmid was used as a control to show no luciferase activity was possible in the absence of a cloned promoter. The normalised luciferase activities are displayed as a percentage relative to WT activity. The results are averages of a set of four experiments all carried out in triplicate on each occasion. The errors bars are calculated as average deviations between data sets and the asterix indicates that there is a significant difference between WT and the different constructs tested. \* specifies that there is a less than 5% chance of the data set being similar to WT or >95% confidence that these two data sets are different.

### 3.5 Chapter summary.

Previous work has indicated that transcriptional regulators form functional complexes at different elements along the topoisomerase II $\alpha$  promoter (Morgan and Beck, 2001; Coustry *et al.*, 2001; Williams *et al.*, 2007). The work discussed in this chapter is an extension of previous work (Magan *et al.*, 2003) confirming the existence of an additional protein complex that can bind *in vitro* to the topoisomerase II $\alpha$  promoter. Data from EMSA indicate that this protein complex is likely to contain recruitment factors that interact with Sp1, Sp3 and NF-Y bound at the nearby GC1 and ICB1 elements. The region between ICB1 and GC1 (to which complex three cannot independently bind) is involved in the formation of a protein complex and the GC1 element also has a critical role in formation of this complex. Interestingly, Sp3 appears to be a component of complex three; however other proteins within this complex are yet to be elucidated. Preliminary DNA-binding analysis demonstrates that these unidentified components appear to associate with DNA in a sequence-specific manner, with a preference for the “GTC TGC” sequence. The proteins that are associated with complex three are also present in other cell lines and are most likely to be nuclear proteins, providing evidence that they could be transcription factors and may be ubiquitously expressed. Finally, complex three is clearly an important regulatory component of the topoisomerase II $\alpha$  promoter as a mutation in the region responsible for binding of these proteins results in a reduction in promoter activity.

# 4 Purification of DNA Binding proteins.

## 4.1 Synopsis.

The data presented in the previous chapter demonstrates that a protein complex (called complex three) could be important for the transcriptional regulation of topoisomerase II $\alpha$ . Therefore, attempts were made to purify the proteins in this complex. This was approached a number of ways including ion-exchange chromatography, oligonucleotide -affinity chromatography and eluting proteins directly from EMSA gels. The final objective in the purification of these proteins was to identify some of the DNA-binding proteins present in complex 3 (chapter 5).

## 4.2 Ion-exchange Chromatography.

Heparin is a naturally occurring glycosaminoglycan that is negatively charged. DNA-binding proteins have been shown to be able to associate with heparin as it mimics the negative charge normally associated with DNA. Heparin and DNA appear to compete non-specifically for the same site on DNA-binding proteins (Gadgil and Jarrett, 1999). Heparin is covalently attached to a sepharose matrix, thereby immobilising the heparin, which can be used as an ion exchange ligand in protein purification. By binding the proteins at a low salt concentration, proteins can be eluted sequentially by increasing the salt concentration in either a step-wise or continuous gradient. As the salt concentration increases, the salt shields the charge on the heparin; thereby the DNA-binding proteins are released from the heparin when the salt concentration (positive ions) reaches a critical level. This is a common method, often used as a first step to purify DNA-binding proteins from whole cell or nuclear extracts (Briggs *et al.*, 1986).

### 4.2.1 Purification of proteins from whole cell HeLa extract using heparin-sepharose column chromatography.

To prepare the HeLa extract for chromatography; 30 mL (2 mg/mL) of whole cell HeLa extract was equilibrated by dialysis overnight in 0.1 M KCl buffer E (refer to section 2.2.21 for method details and buffer specifications). The precipitated cell debris was removed by centrifugation and the cleared extract was filter sterilised through a 0.8  $\mu$ m filter and

quantified (2.5 mg/mL). The residual HeLa precipitate was retained and resuspended in 3 mL extraction buffer, quantified (1 mg/mL) and retained for subsequent analysis.

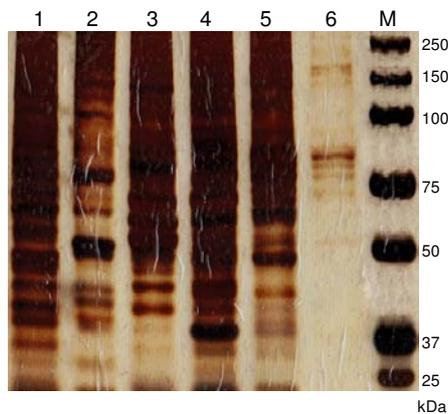
Approximately 20 mg of equilibrated, filtered, whole cell HeLa extract was loaded onto a 5 mL heparin-sepharose column for fractionation. A gradient was set up from 0.1 M KCl to 1 M KCl and 20 fractions (5 mL each) were collected between the 0.15-0.6 M KCl ranges. Most DNA-binding proteins are reported to dissociate from positively charged supports at 0.4 M KCl (Briggs, 1986; Gadgil *et al.*, 1999). Only four fractions (15 mL) were collected within the 0.6 M KCl to 1 M KCl range, in case some protein remained bound at KCl concentrations greater than 0.6. Finally the column was washed with 50 mL 2 M KCl to remove any remaining bound proteins. Each fraction was very dilute after fractionation (0.05-0.1 mg/mL), therefore they were pooled, concentrated and buffer exchanged into extraction buffer (refer to section 2.2.14 for buffer recipe) using centricon 10 (Amicon) ultrafiltration devices to give rise to four fractions according to salt concentrations; 0.25 M KCl, 0.35 M KCl, 0.45 M KCl and 0.55 M KCl for further analysis. After ultrafiltration, the protein content of each of the fractions was determined and each sample was diluted to 1 mg/mL in extraction buffer, with the exception of the 0.55 M KCl fraction which was less than 0.01 µg/µL even after ultrafiltration. The absorbance at 280 nm was determined for each of the dilute samples, which indicated that most of the proteins eluted off the column between 0.25 and 0.5 M KCl, therefore the low protein concentration obtained for the 0.55 M KCl sample was probably not due to protein loss during the ultrafiltration and buffer exchange process.

#### **4.2.2 Proteins purified using heparin-sepharose column.**

To determine if the ion-exchange process had been successful, concentrated and quantified samples were tested for the presence of known DNA-binding proteins (Sp1, Sp3 and NF-Y) by immunoblotting. Silver staining of the SDS-PAGE gel showed that many proteins were present in each sample (figure 4.1A). Preliminary analysis of the different fractions prepared from the heparin-sepharose column showed that the 0.35 M KCl fraction had enhanced amounts of Sp1, Sp3 and NF-Y compared with unfractionated HeLa extract that had been loaded on the column (compare lanes 1 and 5 then with lanes 8 and 9 in figure 4.1B) or any of the other purified fractions (lanes 3, 4, 6 and 7, figure 4.1B). A very small amount of Sp1 was also detected in the 0.45 M KCl fraction (lane 6, figure 4.1B). Taken together, these

results indicate that DNA-binding proteins had been successfully purified from whole cell HeLa extracts using heparin-sepharose column chromatography. Most of these DNA-binding proteins dissociated from heparin within 0.25-0.45 M KCl. This is consistent with findings that Sp1 can elute from a heparin-sepharose column using 0.3 M KCl (Briggs *et al.*, 1986). The fractions were then tested for the presence of complex three using <sup>32</sup>P-ICB1wt/GC1wt oligonucleotide in EMSAs as described in chapter 3.

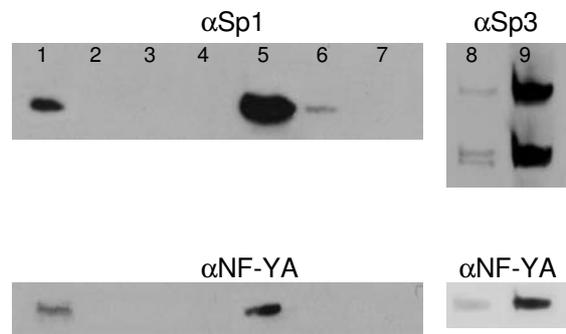
**Silver stained SDS-PAGE of protein fractions**



**(A)**

- Lane 1) 3 µg HeLa extract loaded onto column
- Lane 2) 1 µg HeLa precipitate after dialysis
- Lane 3) 5 µg 0.25 M KCl proteins
- Lane 4) 5 µg 0.35 M KCl proteins
- Lane 5) 5 µg 0.45 M KCl proteins
- Lane 6) 30 ng 0.55 M KCl proteins
- M) Protein size standard

**Immunoblot of protein fractions**



**(B)**

- Lane 1 & 8) 20 µg HeLa extract loaded onto column
- Lane 2) 20 µg HeLa precipitate after dialysis
- Lane 3) 20 µg 0.1 M KCl (Wash) proteins
- Lane 4) 20 µg 0.25 M KCl proteins
- Lane 5 & 9) 20 µg 0.35 M KCl proteins
- Lane 6) 20 µg 0.45 M KCl proteins
- Lane 7) 0.5 µg 0.55 M KCl proteins

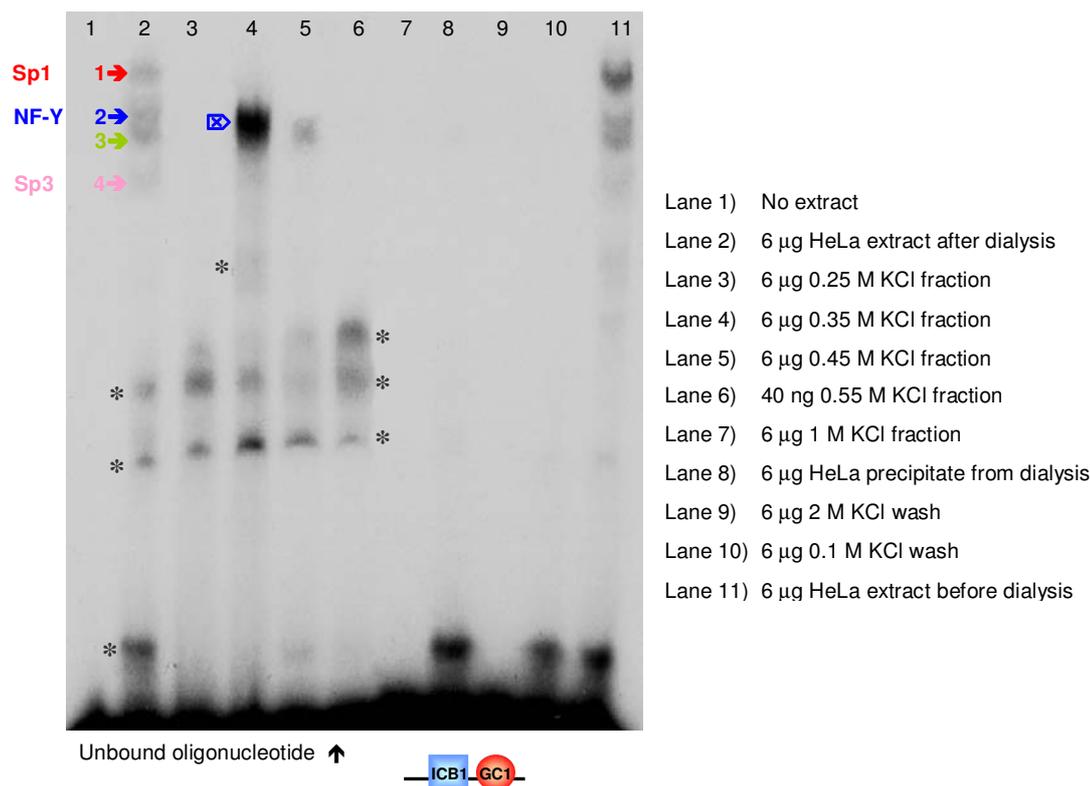
**Figure 4.1: Analysis of proteins in fractions purified using heparin-sepharose column.**

Proteins were separated by 8% SDS-PAGE and **(A)** silver stained to visualise the protein composition of the fractions purified using a heparin-sepharose column. **(B)** Immunoblot of protein fractions purified using a heparin-sepharose column. 20 µg of total protein was separated by SDS-PAGE using 8% polyacrylamide and transferred onto a positively charged nylon membrane. Primary antibodies against Sp1, Sp3 and NF-Y were diluted 1:500 and 2<sup>o</sup> HRP-conjugated antibodies were diluted 1/5000. Chemiluminescence was visualised by exposure to X-ray film. Three bands can be seen for Sp3: there are four reported isoforms of this transcription factor, the smaller forms of Sp3 arise from translation initiation at internal methionines.

#### 4.2.2.1 DNA-binding proteins in heparin-sepharose fractions.

Even though the 0.35 M KCl fraction contained Sp1, Sp3 and NF-Y the other fractions may contain other proteins that may be present in complex three, therefore all fractions were tested for the presence of any DNA-binding proteins using EMSA.

##### DNA-binding proteins in fractions purified using heparin-sepharose chromatography

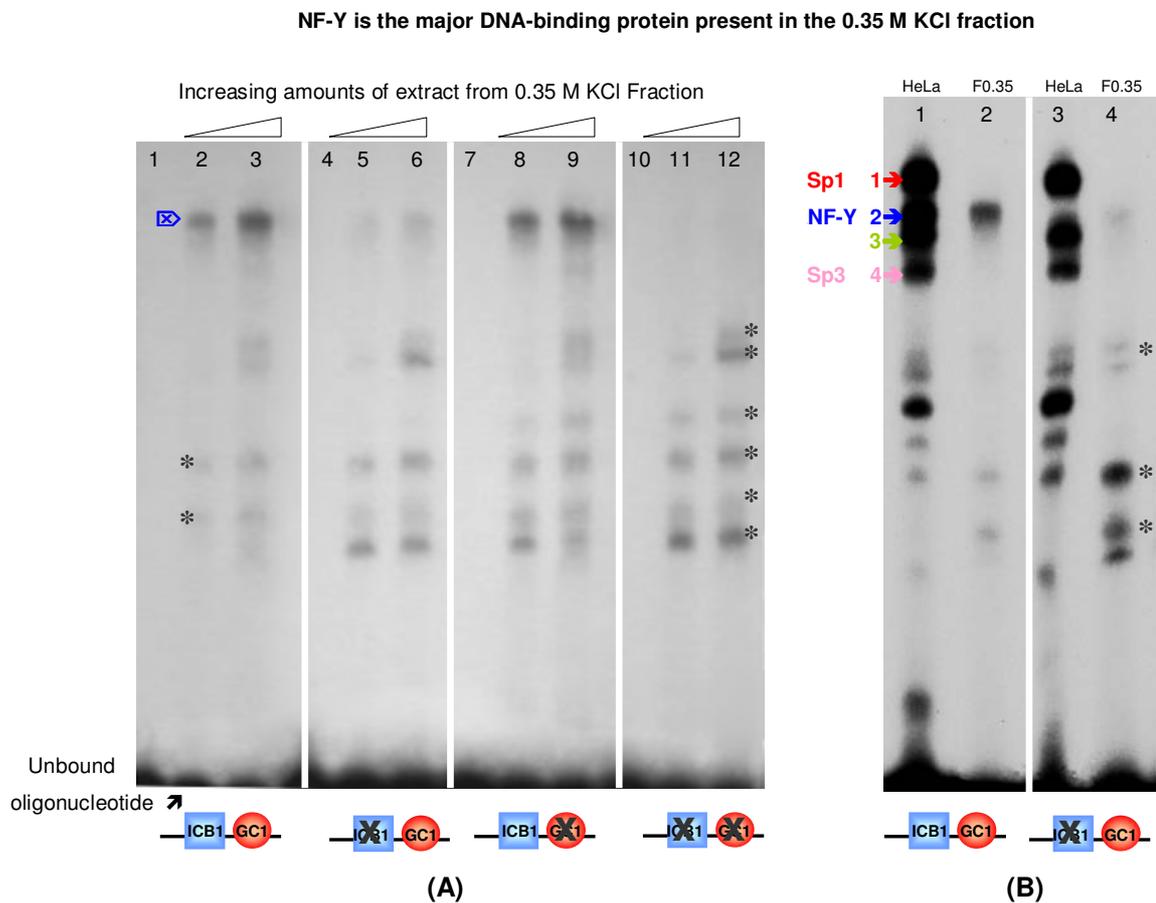


**Figure 4.2: EMSA using different fractions from heparin-sepharose column.**

6 µg of proteins from fractions (except for the 0.55 M KCl fraction)) purified using heparin-sepharose chromatography or of HeLa extract was added to approximately 0.5 ng <sup>32</sup>P-ICB1wt/GC1wt oligonucleotide in each EMSA reaction. Gels were dried onto DE-81 paper and exposed to X-ray film for about 19 hours. This gel is representative of duplicate experiments. Only the 0.35 M KCl fraction has produced a distinct mobility shift, an intense low mobility band ⊗. The coloured arrows show each of the four protein-DNA complexes and describe the major protein found in each complex. \* highlights some of the high mobility bands thought to be non-specific interactions.

Figure 4.2 shows the mobility shift of the different fractions in the presence of the <sup>32</sup>P-ICB1wt/GC1wt oligonucleotide. Lane 2 contains the HeLa extract that was loaded onto the column (after dialysis), four protein complexes can be seen (coloured arrows). This demonstrates that the dialysed HeLa extract retained the same typical DNA-binding mobility

shift patterns observed with whole cell HeLa extract (lane 11), albeit the bands appear less intense in lane 2 than lane 11, this is probably due to some protein loss during dialysis. There is only one low mobility band observed in lanes 4 and 5, the major band present in lane 4 (0.35 M KCl fraction, ) which appears to have a slightly different mobility to that of any of the four complexes under examination. This may be due to the large amount of protein bound to the oligonucleotide, or it may be comprised of other proteins yet to be identified. Therefore, the binding specificity of the proteins in the 0.35 M KCl fraction was challenged using different topoisomerase II $\alpha$  oligonucleotides. To gain a clearer picture of the mobility shift seen in figure 4.2, less extract was used so any underlying protein-DNA binding would not be masked.



**Figure 4.3: EMSAs using F 0.35 M KCl and different topoisomerase II $\alpha$  oligonucleotides.** (A) Increasing amounts (0, 1, and 2  $\mu$ g) of the 0.35 M KCl fraction or (B) 5  $\mu$ g of HeLa whole cell extract or 2  $\mu$ g 0.35 M KCl fraction (F0.35), was added to approximately 0.5 ng of each different  $^{32}$ P-labelled oligonucleotide. Gels were dried onto DE-81 paper and exposed to X-ray film for about 25 hours. The four protein-DNA complexes under investigation for this study are indicated with the coloured arrows and the major protein present is described. Some of the high mobility bands are represented with \* and the low mobility band present in fraction 0.35 M KCl is shown with . These gels are representative of duplicate experiments.

Figure 4.3A demonstrates that the major protein-DNA complex (⊗) was not capable of forming unless an intact ICB1 element was present (compare lane 6 with lane 3), as a mutation in the ICB1 element (lanes 4-6) almost completely abolishes the corresponding mobility shift. It is possible that the residual protein-binding observed in lane 6 may be due to protein-protein or protein-DNA interactions at other sequences (other than the ICB1 element) in the oligonucleotide. Including a HeLa extract control (figure 4.3B, lanes 1 and 3) confirmed that the majority of protein from the 0.35 M KCl fraction shown to bind to the oligonucleotide was ICB1 dependent. Antibody supershifts were also carried out several times using the 0.35 M KCl fraction and antibodies against Sp1, Sp3 or NF-Y. These revealed that one of the proteins present in this complex was NF-Y. Sp1 and Sp3 were also present to a much lesser extent (data not shown). These antibody supershifts however, were not conclusive as they gave weak and variable super shifts in the presence of antibody. Overall, the antibody supershifts raised the possibility that another DNA-binding protein was present in fraction 0.35 M KCl and it was most likely something other than those proteins already recognised. It is interesting to note, that a mutation in the GC1 element (lanes 7-9 compared to lanes 2 and 3) does not appear have any deleterious effects on the formation of this complex (⊗). This suggests that Sp1 or Sp3 may not be present in this protein-DNA complex, which is odd as this extract was shown to contain these factors in immunoblots (figure 4.1B).

There are no low mobility bands (that line up with any of the four complexes under examination) in lanes 3, 6, 7, 8, 9, and 10 (figure 4.2) suggesting that none of the transcription factors Sp1, Sp3 or NF-Y are present or capable of binding to the oligonucleotide used in any of these samples. This is consistent with the results from the immunoblot blots (figure 4.1B), that showed only the 0.35 M KCl fraction (and to a lesser extent the 0.45 M KCl fraction) contained Sp1, Sp3 or NF-Y. However there are some higher mobility bands (\*) present in these lanes, suggesting the presence of some DNA-binding proteins. In addition, some of these bands are similar in mobility to those seen in lane 2 (figure 4.2) containing dialysed HeLa extract. When both the ICB1 and GC1 elements are mutated (figure 4.3A, lanes 11 and 12), only the high mobility bands can be seen (\*), suggesting that these protein-DNA interactions are not forming at either of these elements. It is possible that these protein-DNA interactions are forming at other regions within the oligonucleotide, but it is more likely that these interactions are non-specific. This is

supported by the inconsistencies observed in mobility shifts across several EMSA gels (compare \*s from figure 4.1, figure 4.3A and B).

Even though the 0.35 M KCl fraction appears to contain both Sp1 and Sp3 by immunoblot (figure 4.1B) it is unclear why these were not evident in DNA-binding assays (figure 4.2, lane 4 and 5). There are a number of possible explanations; the proteins may have degraded as the procedure from crude whole cell extract to concentrated fractions from the heparin-sepharose column required many steps. Although care was taken to keep samples at 4°C throughout this process some protein loss and degradation is likely to occur. As the concentration of transcription factors increased in fractions (especially Sp1 and Sp3 in the 0.35 M KCl fraction, figure 4.1B), it is possible that the concentration of potential inhibitors or co-factors that could hinder DNA-association may have also increased. The DNA-binding proteins may preferentially associate with other protein factors instead of DNA. On the other hand, it is also possible that a co-factor which both Sp1 and Sp3 require to bind DNA was separated out or lost during the process. It has been shown that Sp1 and Sp3 are highly sensitive to salt concentration in DNA binding reactions, as little as 150 mM NaCl (Briggs *et al.*, 1986) can inhibit binding. It is possible that not enough of the KCl was removed from fractions, even though attempts were made to reduce KCl from the concentrated fractions. In addition to this, Sp1 has been shown to require Zn (II) for sequence-specific binding to DNA, EDTA can also inhibit binding to DNA (Kadonaga *et al.*, 1987), so if these factors were not in the correct balance in the final concentrated fractions DNA-binding would also be restricted when examined using EMSA. Even though this method of purifying DNA-binding proteins may have successfully enhanced for the presence of some DNA-binding proteins from whole cell HeLa extract, as EMSA was the only method available for detecting complex three and EMSA assays were inconclusive this purification method was abandoned.

### 4.3 Oligonucleotide–affinity Chromatography.

Due to the fact that ion-exchange chromatography required vast amounts of extract, the proteins appeared to be present in a complex mix of other proteins and preliminary results were inconclusive, an alternative method was attempted. Of all the methods currently in use for the purification of DNA-binding proteins, sequence-specific DNA affinity chromatography potentially offers the highest selectivity. Transcription factors usually bind

their cognate DNA sequences with  $10^3$ - $10^5$  higher affinity for their DNA sequences than for any other sequence (Gadgil *et al.*, 2001).

Instead of using a  $^{32}\text{P}$ -ICB1mt/GC1wt (as in EMSA reactions), the same topoisomerase II $\alpha$  composite oligonucleotide was biotinylated at the 3' end of one of the strands of DNA. With the biotin at the 3' end, the 5' end was still free to be  $^{32}\text{P}$ -labelled if required. When the biotinylated-ICB1mt/GC1wt oligonucleotide was tested in EMSA versus the non-biotinylated ICB1mt/GC1wt oligonucleotide, there were no differences in the resulting band-shift profiles (data not shown). Therefore the addition of biotin to this oligonucleotide did not alter protein-binding profiles. The biotinylated-ICB1mt/GC1wt oligonucleotide that was used in DNA-affinity chromatography, as a mutation in ICB1 did not inhibit the formation of complex three, hence NF-Y was not required for binding of complex three. By using streptavidin magnetic beads (DynaL, Dynabeads M-280) in conjunction with biotinylated-ICB1mt/GC1wt oligonucleotide, the protein-DNA complexes could be immobilised and separated, then DNA-bound proteins could be sequentially washed off DNA in a purification step (described in section 2.2.22). This method was pursued as a means of producing an aliquot of HeLa extract that was enriched for the factors that would normally bind to this oligonucleotide.

Each streptavidin dynabead has 3-4 biotin binding sites and the manufacturers' indicate that the binding capacity of the dynabeads is dependent on the length fragment of DNA. It is recommended that twice as many copies of 500 bp DNA can bind as 1000 bp DNA where reduced binding of larger DNA fragments is thought to occur due to steric hinderance. Therefore according to the manufacturers' instructions, approximately 5  $\mu\text{g}$  of double-stranded 44 bp biotinylated oligonucleotide should bind 100  $\mu\text{L}$  of Dynabeads. To optimise the amount of biotinylated oligonucleotide used in binding reactions, increasing amounts of double-stranded ICB1mt/GC1wt biotinylated oligonucleotide was titrated (50, 100 and 150  $\mu\text{L}$  of 100 ng/ $\mu\text{L}$  stock solutions) against 100  $\mu\text{L}$  of dynabeads (10 mg/mL). The amount of biotinylated oligonucleotide that bound to the streptavidin dynabeads was tracked by reading the absorbance at 260 nm using the nano-drop (ND-100) spectrophotometer, and it was found that 100  $\mu\text{L}$  (1 mg) of streptavidin dynabeads was capable of binding only 2.5  $\mu\text{g}$  of double-stranded ICB1mt/GC1wt-biotinylated oligonucleotide. This was found to be only half the amount recommended by the manufacturers, nevertheless the streptavidin dynabeads still

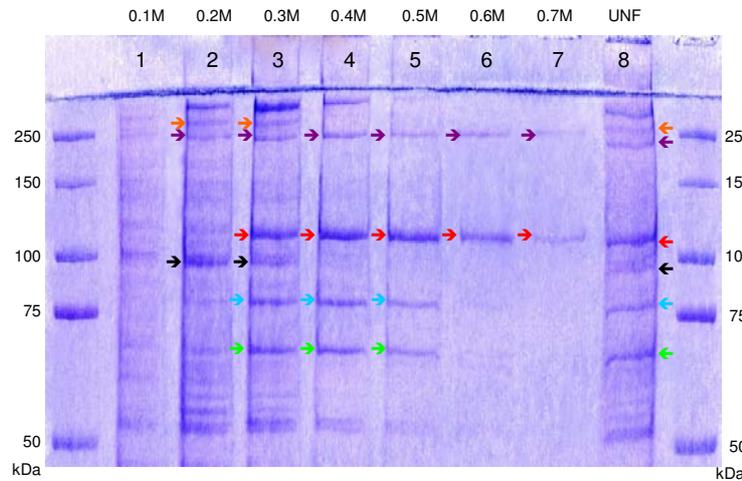
bound the biotinylated oligonucleotide and thus the assay was optimised and carried out as described in section 2.2.22.

### **4.3.1 Purification of ICB1mt/GC1wt-binding proteins.**

A number of different conditions were trialed to optimise protein binding, protein washing and protein elution steps from the biotinylated-ICB1mt/GC1wt oligonucleotide. Binding reactions were favoured by maintaining a low salt concentration (25-50 mM KCl) in 50% GSB (gel shift buffer) and greater yields of protein were produced at elution steps using buffers containing >0.4 M KCl. Proteins were eluted either in a step-wise process to fractionate the different oligonucleotide-bound proteins, achieved by increasing the KCl concentration in buffer E sequentially (fractionated) or unfractionated (by simply eluting in buffer E containing 0.5 M KCl). This is accordance with recommendations that a higher salt concentration (0.4-0.6 M KCl) is required to elute proteins from specific DNA sequences (Kadonaga and Tjian, 1986; Gadgil and Jarrett, 2001). The affinity of biotin to streptavidin is exceptionally high ( $K_D = 10^{-15}$  M), making this interaction extremely resistant to high salt concentration.

Proteins were initially fractionated starting at a 0.1 M KCl through to 0.7 M KCl to determine the optimum KCl concentration for elution of proteins from DNA. Resolution of proteins in SDS-PAGE and DNA-binding in EMSA was improved if the salt concentration of samples were kept to 0.4 M KCl or lower. Excess salt was removed from protein samples either by dialysis in D-tube dialysers (Novagen) in buffer E (containing 0.1 M KCl) or by buffer exchange using protein ultrafiltration devices (centricon-10, Amicon). Figure 4.4 illustrates a typical Coomassie blue stained gel of proteins purified using this method. During the purification process a large number of proteins were washed away using 0.05 M KCl buffer E, these were proteins that did not bind to the oligonucleotide or perhaps bound with low stringency and were not required (data not shown). Lane 8 (figure 4.4) represents the total population of proteins that bind to the ICB1mt/GC1wt oligonucleotide, where all the DNA-binding proteins were eluted in one step (unfractionated) using 0.5 M KCl buffer E. Even though there are numerous DNA-binding proteins some of them may arise due to non-specific DNA-protein or protein-protein interactions, over many consecutive experiments this set of DNA-binding proteins proved to be extraordinarily reproducible.

**Coomassie blue stained SDS-PAGE of proteins purified using DNA affinity chromatography.**



- Lane 1 ) 25  $\mu$ L of 0.1 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 2 ) 25  $\mu$ L of 0.2 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 3 ) 25  $\mu$ L of 0.3 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 4 ) 25  $\mu$ L of 0.4 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 5 ) 25  $\mu$ L of 0.5 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 6 ) 25  $\mu$ L of 0.6 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 7 ) 25  $\mu$ L of 0.7 M KCl elution of extract purified using biotinylated-ICB1mt/GC1wt
- Lane 8 ) 20  $\mu$ L of unfractionated extract purified using biotinylated-ICB1mt/GC1wt

**Figure 4.4: DNA-binding proteins purified using a biotinylated-ICB1mt/GC1wt oligonucleotide and the proteins are eluted using different salt concentrations.**

25  $\mu$ L of extract purified using a biotinylated-ICB1mt/GC1wt oligonucleotide was denatured and separated using 8% SDS-PAGE. The extract in each lane represents about 50% of the total amount of protein eluted at each of the protein elution steps using different salt concentrations. Proteins were either fractionated by eluting with increasing salt concentrations or unfractionated representing the total population of proteins that are found to bind to the biotinylated-ICB1mt/GC1wt oligonucleotide. This result is representative of several similar experiments. The coloured arrows illustrate DNA-binding proteins of the same molecular weight that are present in different samples.

Of the proteins purified using oligonucleotide-affinity chromatography, some proteins preferentially eluted at low salt concentrations (figure 4.4, lane 2 and 3) such as the as the 95 kDa protein (indicated with the black arrows) and the 250+ kDa protein (indicated with the orange arrows). Upon closer examination the 95 kDa protein actually represents two bands similar in size (lane 3, black arrow). Other proteins consistently appear to elute together (light blue 85 kDa protein and light green arrow 70 kDa protein, lanes 3, 4 and 5) within the same salt range (0.3-0.5 M KCl). This may suggest that these two proteins (85 and 70 kDa) may have a functional interaction; possibly they are subunits for a protein which are both

required to bind to DNA. Some proteins preferentially eluted over a range of higher salt concentrations (0.3-0.6 M) such as the 115 kDa protein (red arrows, lanes 3-7 in figure 4.4). This 115 kDa protein still remained partially bound to the oligonucleotide even at 0.7 M KCl (lane 7, figure 4.4). And other proteins appeared to elute over a broader range of salt concentrations (protein approximately 250 kDa, lanes 2-6 purple arrows), showing little preference at all. It was difficult to decipher and analyse every single protein purified using this process, not only due to the resolution but also as some of the higher intensity bands were often composed of several bands similar in size (for example the 95, 115 and 250 kDa bands). Nevertheless, these results demonstrate that the different DNA-binding proteins could be separated from each other using differing salt concentrations (if required) and that it is possible that some of these proteins could be interacting with each other as well as associating with the oligonucleotide used to purify these proteins in this assay. However, the overall composition of the proteins always remained the same, suggesting that a number of proteins bound to the oligonucleotide with quite high specificity and that the biotinylated-ICB1mt/GC1wt oligonucleotide was capable of being used to purify specific DNA-binding proteins.

Typically, this purification method created a population of about 15 proteins that appeared in a relatively high concentration compared to the rest, as well as 20-30 other minor protein components. This still represented a large number of putative DNA-binding proteins that could be involved in the formation of complex three as well as Sp1 and Sp3, which have been shown to bind to this oligonucleotide. It was also possible that some of the proteins in the mixture arose from non-specific protein-DNA interactions or protein-protein interactions. Therefore it was important to confirm that this mixture actually contained proteins known to bind to this oligonucleotide, such as Sp1 and Sp3, as well as try to minimise the amount of non-specific proteins that bound. Therefore a series of double-stranded competitors were used in conjunction with this purification method to increase specificity of this technique.

#### **4.3.1.1 Using a competitor to reduce non-specific protein-DNA interactions.**

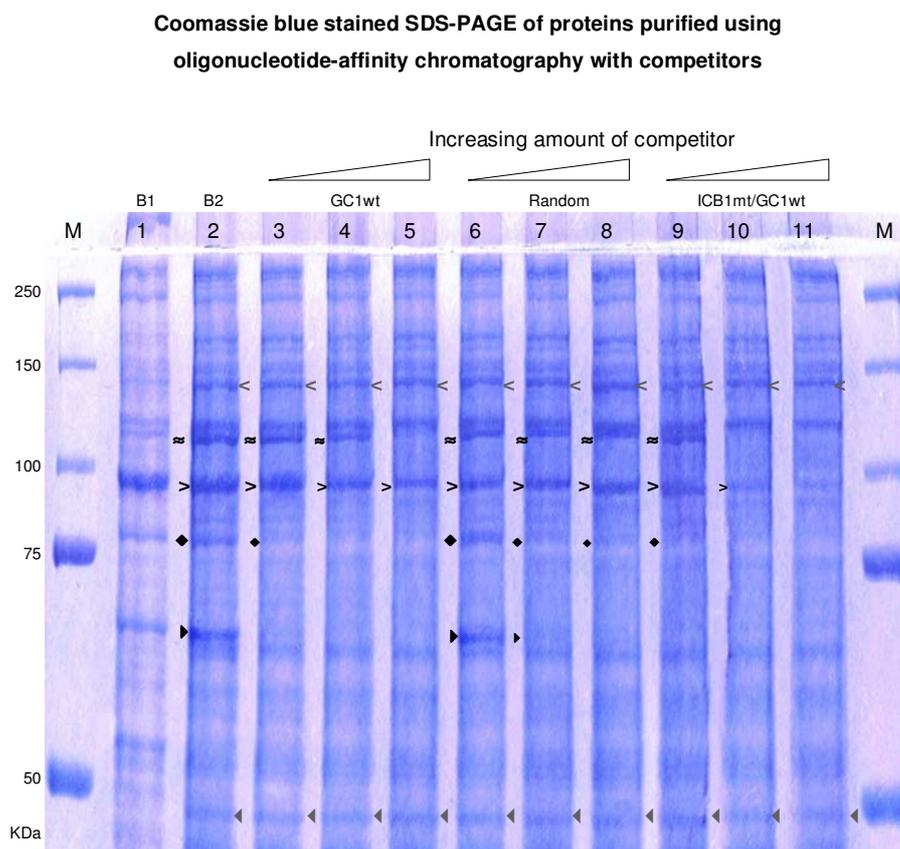
In order to determine if the biotinylated-ICB1mt/GC1wt oligonucleotide was purifying the required protein components, specific double-stranded competitors were added to the protein-DNA binding process. The addition of a short GC1wt competitor should reduce the amount of Sp1 and Sp3 in the final purified extract. A random competitor was also used to

reduce any non-specific DNA-binding proteins. The ICB1wt/GC1wt competitor should reduce the overall amount of all specific protein-DNA interactions. The random competitor was designed to be the same length as the ICB1mt/GC1wt oligonucleotide and was tested as a competitor in EMSA. The random competitor was found to be unable to compete for binding of any of the four complexes under examination in EMSA (data not shown).

Figure 4.5 shows the results of adding increasing amounts of GC1wt, random and ICB1mt/GC1wt competitors during the oligonucleotide-affinity purification process. Lanes 1 and 2 consist of extract purified using the biotinylated-ICB1mt/GC1wt oligonucleotide made on two separate occasions and demonstrates that the overall population of proteins that are purified in this manner are similar in composition. If all the proteins were specific protein-DNA interacting proteins it would be expected that an overall reduction in all bands would be observed with the addition of the ICB1mt/GC1wt competitor (lanes 9-11). However this was not the case, even though there is a marked reduction in some of the bands ( $\approx$ ) other bands do not change at all (140 kDa band  $\blacktriangleleft$  or 45 kDa band  $\blacktriangleright$ , lanes 2-11). It is possible that insufficient amounts of competitor was added to the assay to compete for binding relative to the amount of the DNA-binding protein present and the total amount of biotinylated-oligonucleotide used in the purification process, therefore a marked reduction may not be observed in some instances. It is possible that some trends will not be seen as they are below the limit of detection for coomassie blue staining. The level of sensitivity for coomassie blue is approximately 0.1  $\mu$ g of protein per band therefore only the proteins which are in greater abundance will be visualised with coomassie blue staining and there will be underlying proteins that have not stained in this circumstance due to lower abundance. Generally transcription factors are in low abundance therefore the presence of these proteins should not be forgotten.

There are two bands (figure 4.5,  $\blacklozenge$  and  $\blacktriangleright$ , approximately 80 kDa and 65 kDa respectively) that appear to decrease in the presence of each competitor (compare lane 2 with lanes 3-11). Both these bands are completely abolished in the presence of both the GC1wt and ICB1mt/GC1wt competitors (lanes 3-5 and 9-11 respectively). However, both proteins do appear to be competed for to a lesser degree with the random competitor (lane 6). In addition, the 80 kDa protein ( $\blacklozenge$ ) appears to be less affected by the presence of random competitor (lanes 6-8) than the 65 kDa protein, as some of the 80 kDa band ( $\blacklozenge$ ) still remains in lanes 7

and 8 compared to the 65 kDa band ( ▶ ) in lanes 7 and 8. Never the less, the overall binding patterns observed by these two proteins suggests that these could be non-specifically associating with the biotinylated-ICB1mt/GC1wt oligonucleotide. It is possible that these proteins are non-specifically binding to the ends of DNA rather than a specific internal sequence.



**Figure 4.5: Extracts purified using biotinylated-ICB1mt/GC1wt oligonucleotide with the addition of increasing amounts of competitors.**

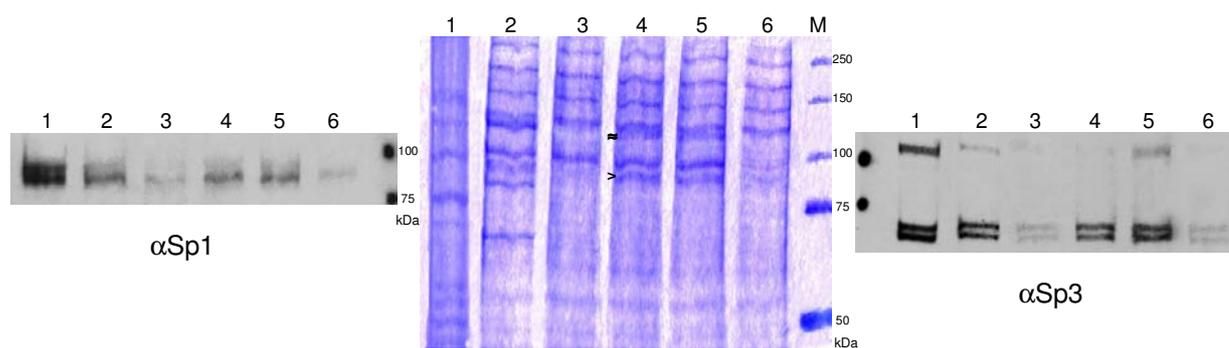
150  $\mu$ L of HeLa cell nuclear extract (5  $\mu$ g/ $\mu$ L) was premixed with increasing amounts of double-stranded competitors (0, 1, 5 and 10  $\mu$ g) and incubated on ice for 40 minutes. The binding reactions were carried out as described in section 2.2.22, using only 250  $\mu$ L dynabeads and 100  $\mu$ L of biotinylated-ICB1mt/GC1wt (100ng/ $\mu$ L). The protein-DNA complexes were washed 3x with 0.025 M KCl then 3x with 0.05 M KCl and finally 3x with 0.1 M KCl before elution in 40  $\mu$ L 0.5 M KCl buffer E. The total volume of each reaction was denatured and separated on an 8% SDS-PAGE (large gel, 140mm x 160mm x 1.5mm thick). B1 (lane 1) and B2 (lane 2) are extracts purified using the biotinylated-ICB1mt/GC1wt oligonucleotide prepared from two separate experiments. B2 is the control with no competitor added in lane 2, whereas lanes 3-5, 6-8 and 9-11 all have increasing amounts of a double-stranded competitor added to the purification process. The symbols  $\approx$ , >,  $\blacklozenge$ , and  $\blacktriangleright$  highlight visible changes in protein composition with the addition of competitors. Proteins marked with  $\blacktriangleleft$  or  $\blacktriangleright$  represent those that have not changed with the addition of competitors. These results are representative of duplicate experiments.

There are two proteins that appear to exhibit a reduction in intensity with the ICB1mt/GC1wt competitor (compare lanes 9, 10 and 11, figure 4.5), one is approximately 115 kDa ( $\approx$ ) and the other is approximately 95 kDa ( $>$ ). Both these bands do not appear to decrease in the presence of a random competitor (lanes 6-8) however they appear to be slightly reduced in the presence of the GC1wt competitor (lanes 4 and 5,). Examining lanes 3-5 (with GC1 competitor) the 115 kDa band ( $\approx$ ) appears to have a stronger affinity for the GC1 competitor than the 95 kDa band, as the 115 kDa band is absent from lane 5 even though the 95 kDa band still remains. This binding pattern suggests that both these proteins may be specifically interacting with the biotinylated-ICB1mt/GC1wt oligonucleotide and they are most likely GC1-associating factors. This observed binding pattern in the presence of competitor is similar to that observed with those proteins in complex three in EMSA reactions.

Some of the proteins do not appear to change in appearance upon the addition of any of the competitors ( $<$  and  $\blacktriangleleft$ , lanes 2-11, Figure 4.5) and there are so many proteins it is difficult to examine all of them. Obviously, only the proteins that are in greater abundance ( $>0.1 \mu\text{g}$ ) will be visualised with coomassie blue staining, and only these trends can be compared in this circumstance. Therefore antibodies against Sp1 and Sp3 were used in immunoblots as an indication of the efficacy of this approach to detect transcription factors and their DNA-binding properties in the presence and absence of competitors.

The immunoblots (figures 4.6A and C, lanes 2) clearly demonstrates that the biotinylated-ICB1mt/GC1wt oligonucleotide is able to bind specific transcription factors as both Sp1 and Sp3 are present in extracts prepared using this oligonucleotide-affinity purification method. Figure 4.6 (A and C) also demonstrates that addition of the GC1wt (lanes 3) and ICB1mt/GC1wt (lanes 6) competitor both decrease the amount of Sp1 and Sp3 that can be purified using this oligonucleotide-affinity method. In addition, both competitors seem to inhibit binding to biotinylated-ICB1mt/GC1wt oligonucleotide to the same degree, suggesting that both GC1 and ICB1mt/GC1wt are capable of recruiting similar amounts of Sp1 and/or Sp3. There is a minor reduction in the amount both Sp1 and Sp3 when the GC1mt competitor is added (lanes 4), suggesting this oligonucleotide is still capable of binding small amounts of both Sp1 and Sp3 (which has also been observed in EMSAs; Magan N., MSC Thesis, 2002). But as expected, the random oligonucleotide does not hinder the amount of Sp1 and Sp3 being purified (compare lanes 2 and 5 in both Figure 4.6A and C respectively).

**Immunoblot of proteins purified using oligonucleotide-affinity chromatography with competitors**



(A)

(B)

(C)

- Lane 1) 10 µg Whole cell HeLa extract
- Lane 2) 15 µL extract purified using biotinylated-ICB1mt/GC1wt
- Lane 3) 15 µL extract purified using biotinylated-ICB1mt/GC1wt with GC1wt short competitor
- Lane 4) 15 µL extract purified using biotinylated-ICB1mt/GC1wt with GC1mt short competitor
- Lane 5) 15 µL extract purified using biotinylated-ICB1mt/GC1wt with Random competitor
- Lane 6) 15 µL extract purified using biotinylated-ICB1mt/GC1wt with ICB1mt/GC1wt competitor
- M) Protein size standard

**Figure 4.6: Using a competitor to generate biotinylated-ICB1mt/GC1wt purified extracts.**

200 µL of HeLa cell nuclear extract (5 µg/µL) was pre-treated with 30 µg of double-stranded competitor before purification using 500 µL streptavidin dynabeads (10 mg/mL) and 200 µL biotinylated-ICB1mt/GC1wt oligonucleotide (100 ng/µL) as described in section 2.2.22. DNA-binding proteins were eluted from the biotinylated-ICB1mt/GC1wt oligonucleotide in a total of 150 µL 0.5 M KCl. 15 µL of this final volume was used in each of the above experiments. (A & C) Immunoblot of proteins purified using oligonucleotide-affinity chromatography in the presence of competitors. Primary antibody (A) αSp1 diluted 1/200 (C) αSp3 diluted 1/300. 2<sup>o</sup> antibody was dilute 1/4000. Chemiluminescence was visualised by exposure to X-ray film and the protein size standard was visualised by using fluorescent paint. (B) Coomassie blue stained (control) gel of protein samples. Observable changes in binding of a ≈ 115 kDa protein and a > 90 kDa protein are indicated. These results are representative of duplicate experiments.

There are four reported isoforms of Sp3 (Sapetschnig *et al.*, 2004), two separate bands close together about 115 kDa and another set of two bands about 70 kDa, although only three bands can clearly be seen in figure 4.6C. The smaller two isoforms are generated from two alternative internal translation initiation sites within the Sp3 mRNA and each of the isoforms is capable of binding GC elements specifically (Kennett *et al.*, 1997, Sapetschnig *et al.*, 2004). Interestingly, the 115 kDa Sp3 band(s) corresponds with the previously noted reduction in band of the same size seen in figure 4.5 (≈, lanes 5, 10 and 11). This can also be seen in figure 4.6B (annotated with ≈), therefore it is possible that a substantial amount of Sp3 is being purified using this affinity oligonucleotide method. The molecular weight of

Sp1 is approximately 90 kDa, and this is in agreement with the size seen in the immunoblot figure 4.6A. Interestingly, there is also a marked reduction in a band of similar size in the coomassie blue stained gel (figure 4.6B annotated with >, compare lane 2 with lanes 3 and 6), it is possible that this 90 kDa band is the same band that is mentioned as the 95 kDa band in figure 4.5 as they do appear to follow the same binding patterns. It should also be noted that antibody against Sp1 (approximately 90 kDa) quite commonly detects two bands, there is a slightly larger variant (approximately 95 kDa) that is regarded to be the phosphorylated version of Sp1 (Sapetschnig *et al.*, 2004). Both these bands can be seen clearly in lane 1 (figure 4.6A) and faintly in lanes 2, 4, and 5 suggesting that the phosphorylated version of Sp1 is also important in DNA-binding. This is in agreement with reports that phosphorylated Sp1 has an increased affinity for DNA (Reviewed in Chu and Ferro, 2005).

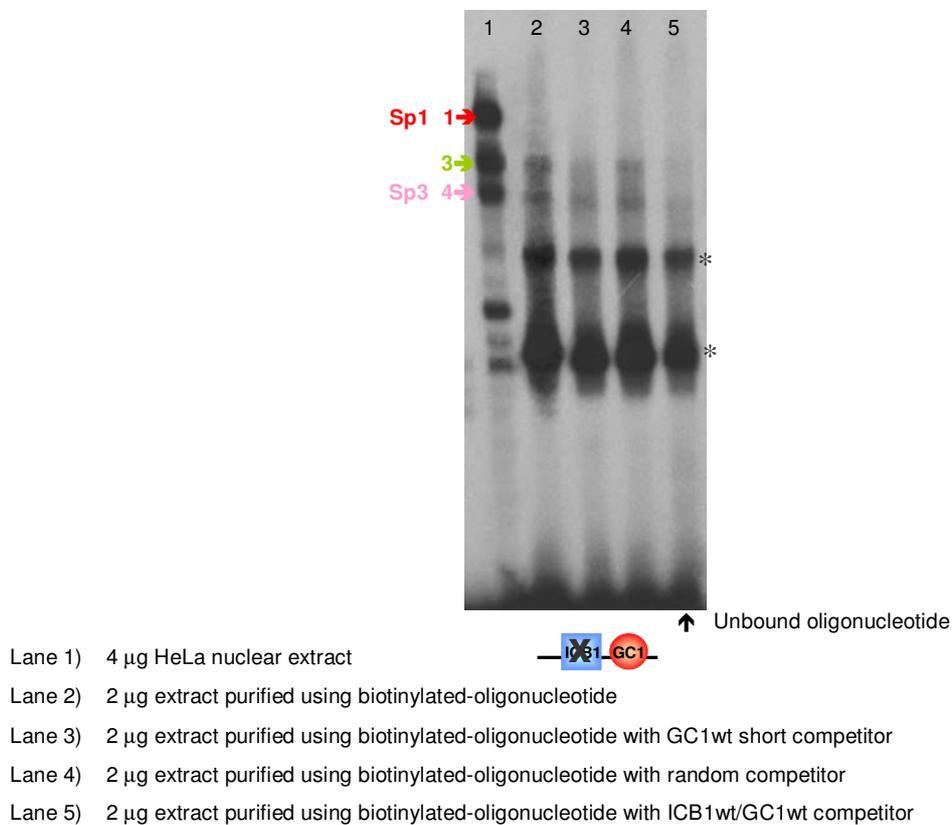
#### **4.3.1.2 EMSA with the extract purified using oligonucleotide–affinity chromatography.**

Generally, when extracts were purified using oligonucleotide-affinity chromatography and dynabeads the final concentrations of total protein were very low (ranging from 0.07-0.1 µg/µL). These extracts were routinely tested for DNA-binding using EMSA and they usually did not give rise to any “typical” mobility shifts patterns. However, on occasion, when the extracts were able to be generated at protein concentrations of 0.3-0.5 µg/µL through pooling and ultrafiltration, these extracts were capable of producing a faint mobility shift similar to those under examination.

Figure 4.7 shows there are two bands present in oligonucleotide-affinity purified extracts and they are similar in mobility to complex three (green arrow, lanes 2 and 4) and four (pink arrow, containing mainly Sp3 lanes 2 and 4). Of particular significance was the observed reduction in these two bands when EMSAs were carried out using extracts that were made with GC1wt and ICB1wt/GC1wt competitor (figure 4.7, lanes 3 and 5 respectively). Unfortunately, the band respective to complex 1 (consisting mainly of Sp1) was absent in lanes 2-4 suggesting that Sp1 was not greatly present in these oligonucleotide-affinity purified extracts, however the immunoblot (figure 4.6A) clearly demonstrates that Sp1 is present in these extracts. It is possible that the lack of binding observed in EMSA assays is due to the reduced abundance of Sp1 in these extracts, seeing that the immunoblot also demonstrates that there is most likely less Sp1 than Sp3 in these oligonucleotide-affinity

purified extracts. Sp1 has been shown to bind to DNA in multimeric arrays (Mastrangelo *et al.*, 1991), and without the optimal concentration of this transcription factor it is also possible that this could hinder binding to DNA. It is also possible that in order for Sp1 to bind with high affinity, it may require a number of other factors that may have been purified out or are also in low abundance, reducing the typical Sp1-DNA interaction. Surprisingly, there are at least two new high intensity bands seen in lanes 2-5 (\*), which are also observed with oligonucleotide-affinity purified extracts of low concentration. The composition of these high mobility bands is unknown, but they are quite prominent and do not appear to change greatly in the presence of competitors during the purification process.

**EMSA of proteins purified using oligonucleotide-affinity chromatography with competitors**



**Figure 4.7: EMSA using extracts that were purified using a biotinylated-ICB1mt/GC1wt oligonucleotide in the presence of competitors.**

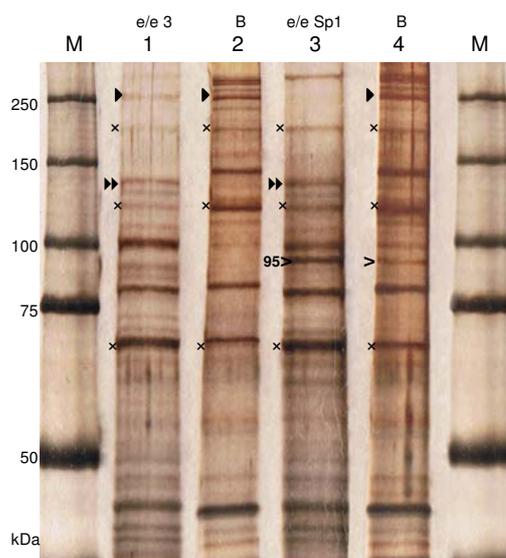
4 µg of HeLa cell nuclear extract or 2 µg of extract purified using a biotinylated-ICB1mt/GC1wt oligonucleotide (and competitors) was added to approximately 0.5 ng of <sup>32</sup>P-ICB1mt/GC1wt oligonucleotide. Gels were dried onto DE-81 paper and exposed to X-ray film for about 24 hours. The protein-DNA complexes under investigation for this study are indicated with the coloured arrows and the major protein found to be present in each complex is labelled. High mobility bands are indicated with \* and are thought to be non-specific interactions. The unbound oligonucleotide can be seen at the bottom of the gel. This gel is representative of triplicate EMSA gels.

Taken together, the data from figure 4.4, 4.5, 4.6 and 4.7 demonstrate that this method of using a biotinylated-ICB1mt/GC1wt oligonucleotide in conjunction with magnetic streptavidin-coated dynabeads could be used to generate an extract that is enriched with specific DNA-binding proteins. Sp1 and Sp3 were identified in the oligonucleotide-affinity purified extract and a band shift (albeit of low intensity) containing a complex of similar mobility to complex 3 was observed. There are however major changes in the band shift profile (figure 4.7, \*) using this method, the significance of this remains obscure. The addition of competitors and differing salt elution steps, are practical tools that could aid the purification process if required. This process has its advantages and disadvantages, it appears to generate extracts that are usually a mixture of about 15 major proteins as well as an additional 30 others. There are still many non-specific proteins present, but far fewer than those isolated using the heparin-sepharose method. Nevertheless these extracts could be used in downstream applications to define the size of the proteins present in complex three and any other proteins that may be interacting *in vitro* with the topoisomerase II $\alpha$  promoter. Cross-linking assays were carried out (chapter 5) in attempts to determine the molecular weight of these specific DNA-binding proteins, as well as direct elution of the proteins from EMSA gels.

#### **4.4 Isolating protein from EMSA gels.**

As the original observation of complex 3 has been in EMSA gels, it was rationalised that it might be possible to isolate proteins from these gels directly. This may provide an indication of the appropriate sizes of proteins involved in each complex in EMSA gels. By performing multiple EMSA reactions and running multiple EMSA gels, the appropriate band(s) were cut out of several polyacrylamide gels and the respective proteins eluted from these gel slices. The first method of elution using the 'crush and soak' process was unsuccessful. The second method was electro-elution, where gel slices were subjected to electrophoresis to release the proteins from within the polyacrylamide gel into a reservoir for collection (refer to section 2.2.25 for details). Proteins were eluted from the band that corresponds with complex 3 as well the band that corresponds with complex 1 (containing mainly Sp1) from EMSA gels, for comparison and verification of the success of the method. The pooled concentrated samples were examined by SDS-PAGE. Due to the low protein concentration of these samples (ranging from 0.01-0.05  $\mu\text{g}/\mu\text{L}$ ) polyacrylamide gels were silver stained (figure 4.8), where the level of sensitivity is approximately 1-5 ng protein per band.

### Silver stained SDS-PAGE of proteins electro-eluted from EMSA gels



- Lane 1) 10  $\mu$ L concentrated electro-eluted complex three (e/e 3)
- Lane 2) 4  $\mu$ L Biotinylated-oligonucleotide purified extract using random competitor (B)
- Lane 3) 10  $\mu$ L concentrated electro-eluted complex one (e/e Sp1)
- Lane 4) 5  $\mu$ L Biotinylated-oligonucleotide purified extract using random competitor (B)
- M) Protein size standards

#### Figure 4.8: Silver stained SDS-PAGE of electro-eluted protein samples from EMSA gels.

DNA-binding proteins associated with complex one and complex three from EMSA gels were eluted, concentrated and analysed by silver staining of SDS-PAGE. Protein concentrations of electro-eluted samples were too low to accurately quantify, but are estimated to be about 10-30 ng/ $\mu$ L. The 95 kDa protein thought to represent Sp1 is annotated with 95>.  $\blacktriangleright$  Represents a protein that is possibly unique to the sample electro-eluted from complex 3.  $\blacktriangleright\blacktriangleright$  Represents a protein present in both electro-eluted samples.  $\times$  Represents examples of proteins common to all three protein samples. These results are representative triplicate experiments.

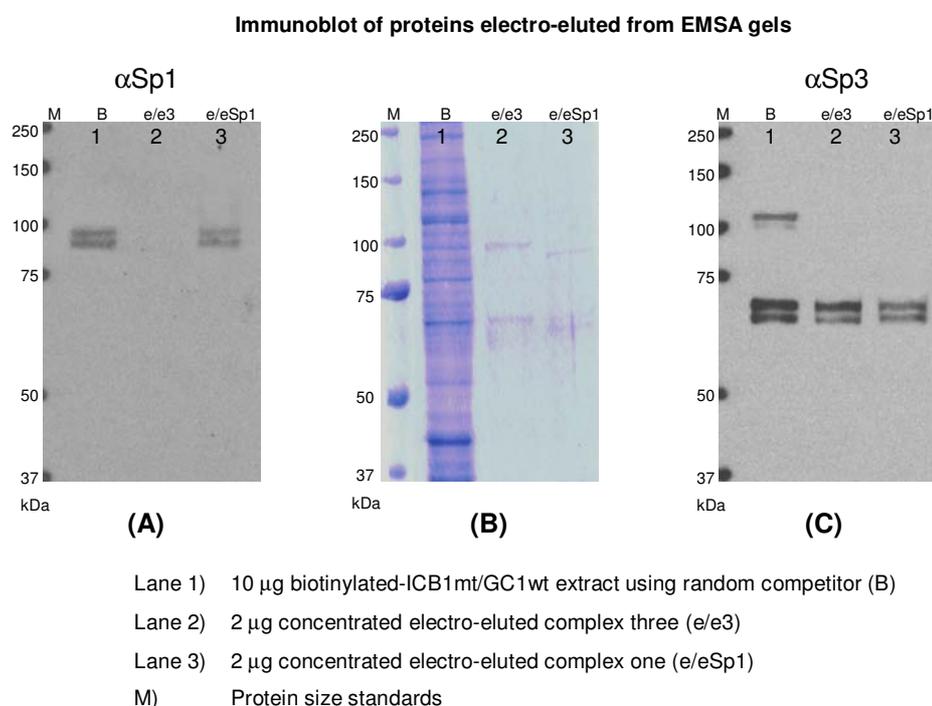
Figure 4.8 represents the results of a typical electro-elution procedure. Most of the proteins present in both of the electro-eluted samples appear to be of the same apparent molecular weight (comparing lanes 1 and 3 refer to  $\blacktriangleright\blacktriangleright$  for an example) however, some bands do appear in differing intensities. This suggests that the overall protein composition of both complex 1 and complex 3 are quite similar. There are however a few differences that can be observed in figure 4.8. The most marked difference, is a 95 kDa band (indicated in figure 4.8 with >) that is missing from the electro-eluted sample from complex 3 (lane 1) but is present in the electro-eluted sample from complex 1 (95>, lane 3). Complex 1 was found by EMSA to be predominantly made up of Sp1 which has a molecular weight of 95 kDa. It is significant to

note that this 95 kDa band is also seen in lane 4 (>), the sample prepared by oligonucleotide-affinity chromatography.

Other observed differences in the electro-eluted samples is a faint band approximately 250 kDa (▶, figure 4.8) in size that appears in the electro-eluted proteins from complex 3 (lane 1), but not in the sample electro-eluted from complex 1 (lane 3). The lack of this 250 kDa band in lane 3 suggests that this protein may be unique to the proteins found in complex 3. This 250 kDa band is notably also present in the protein sample prepared with the biotinylated-ICB1mt/GC1wt oligonucleotide (lanes 2 and 4). Within the population of proteins isolated using oligonucleotide-affinity chromatography (lanes 2 and 4), there are proteins that share the same molecular weight with those found in complex 3 (lane 1) and complex 1 (lane 3, refer to ✕ for examples). Overall, this demonstrates that the biotinylated-ICB1mt/GC1wt oligonucleotide is capable of purifying factors that are specific to either complex 1 or complex 3, confirming the validity of this oligonucleotide-affinity purification approach.

The electro-eluted samples were further analysed for protein composition by using antibodies against proteins that are known to be present in these samples. Immunoblotting results (figure 4.9A) revealed that Sp1 is found in both lane 1 (extract prepared using the biotinylated-ICB1mt/GC1wt oligonucleotide) and lane 3 (electro-eluted proteins from complex 1), but not in lane 2 (electro-eluted proteins from complex 3). This not only indicated that Sp1 is absent from complex 3, but also that the electro-elution process was successfully eluting proteins specific to the region that was excised from EMSA gels. It was interesting to find that Sp3 was present in all three of the protein samples tested (figure 4.9C). All four Sp3 isoforms were present in the sample derived from the proteins purified by oligonucleotide-affinity chromatography (lane 1), however only the two lower molecular weight isoforms were evident in the samples derived from electro-eluted samples (lanes 2 and 3). Even though all four isoforms are capable of binding to their specific DNA sequence, only the larger 117 kDa isoforms are considered to be able to activate transcription from certain promoters and the shorter forms are thought to repress transcription (Kennett *et al.*, 1997, Sapetschnig *et al.*, 2004). The relative proportions of each Sp3 isoform are never homogeneous; therefore it is possible that the larger isoforms are in lower abundance and below the level of detection in this experiment (figure 4.9C, lanes 2 and 3). Nevertheless, the data from electro-eluted protein samples obtained from complex 3 suggests that the lower molecular weight isoforms

of Sp3 are a component of this protein-DNA interaction *in vitro*. Sp1 does not appear to be a component of complex 3. Taken together, these results support the data obtained from the original EMSA experiments described in chapter 3.



**Figure 4.9: Immunoblot of electro-eluted proteins.**

Each protein sample was separated using 8% SDS-PAGE. (A & C) Immunoblot blot of protein samples using antibody against Sp1 and Sp3. Proteins were transferred onto a positively charged nylon membrane for 2 hours at 500 mA. (A) Sp1 diluted 1:300 and (C) Sp3 diluted 1/400 in 0.5% blocking solution. HRP conjugated secondary antibody, anti-rabbit was dilute 1/4000. Chemiluminescence was visualised by exposure to X-ray film. Protein size standards were visualised by using fluorescent paint. (B) Coomassie blue stained (control) gel of protein samples before transfer to nylon membrane for immunoblotting. These results are representative of triplicate experiments

## 4.5 Comparing the molecular weights of all the different purified DNA-binding proteins.

Each of the various techniques used to purify DNA-binding proteins appeared to be able to purify both Sp1 and Sp3, suggesting that they were specifically enhancing for the factors that were under investigation. However, all the different approaches purified a number of different DNA-binding proteins some similar and others different in molecular weight. The results of many purification experiments and numerous SDS-PAGE analyses were combined

and a consensus of the sizes of the proteins that were purified by each technique were summarised, as shown in table 4.1.

	e/e3	e/e1	Biotinylated- ICB1mt/GC1wt
		<b>250+</b>	<b>250+</b>
→	250		<b>250</b>
→	200	200	200
			<b>145</b>
→	140	<b>140</b>	
	128	128	
→	120	<b>120</b>	<b>120</b>
→	115	115	115
			100/105
→	<b>95/98</b>	<b>95/98</b>	95/98
		<b>90</b>	90
	85	85	
→	80	80	<b>80</b>
→	<b>78</b>	<b>78</b>	
	74	74	
	72	72	
→	<b>70</b>	<b>70</b>	<b>70</b>
	65	65	65
	62	62	62

**Table 4.1: Summary of the molecular weights of purified DNA-binding proteins using different techniques.**

The results from different purification experiments and subsequent SDS-PAGE analysis are summarised in the table above. There are a total of 19 putative DNA-binding proteins that appear to be significant for this study. This can be simplified down to 9 (→) potential candidates for MS/MS or N-terminal sequencing. e/e1 are the proteins generated from electro-elution of proteins associated with complex 1 in EMSA gels. e/e3 are the proteins generated from electro-elution of proteins associated with complex 3 in EMSA gels. The proteins that appeared in high abundance in protein gels are represented in bold font.

There are potentially 19 different DNA-binding proteins that appear to be significant, and they range from 250-75 kDa in size. Of these 19 proteins, only 6 of them appear in all three protein samples and so there is a certain degree of similarity for each of the purification method used. Specifically, 9 proteins (indicated with the blue arrows in table 4.1) were found to be similar between the electro-eluted proteins from complex 3 and the oligonucleotide-affinity purified extracts, these seemed suitable for identification using MS/MS analysis or N-terminal sequencing.

## 4.6 Chapter summary.

Each of the methods trialed had advantages and disadvantages for the purification of the proteins under investigation. The heparin-sepharose column provided proteins in the most complicated mixture of components, which could be purified further if required. Whereas, oligonucleotide-affinity chromatography provided a quick and effective way of recovering DNA-bound proteins, which was improved by the addition of non-specific competitors. The electro-elution method required multiple reactions and proteins were recovered in very small quantities (picograms of protein), which needed to be pooled and concentrated in order to even visualise these proteins. However this method was useful as a direct measure of the proteins that could be found in specific protein-DNA complexes.

These methods were never designed to be able to purify an exact protein to homogeneity, but rather enrich for the protein components that are capable of specifically binding to the ICB1mt/GC1wt sequence of topoisomerase II $\alpha$  *in vitro*. The combination of each technique provides valuable information about the size of the proteins that are relevant, and a means to partially purify them. The final objective for this purification was to be able to identify these proteins, to obtain enough protein for MS/MS analysis or N-terminal sequencing. In order to further define the sizes of proteins that do specifically bind to the composite ICB1mt/GC1wt sequence, cross-linking assays were carried out (chapter 5) before the relevant proteins were analysed for identity.

# 5 Identification of DNA-binding proteins.

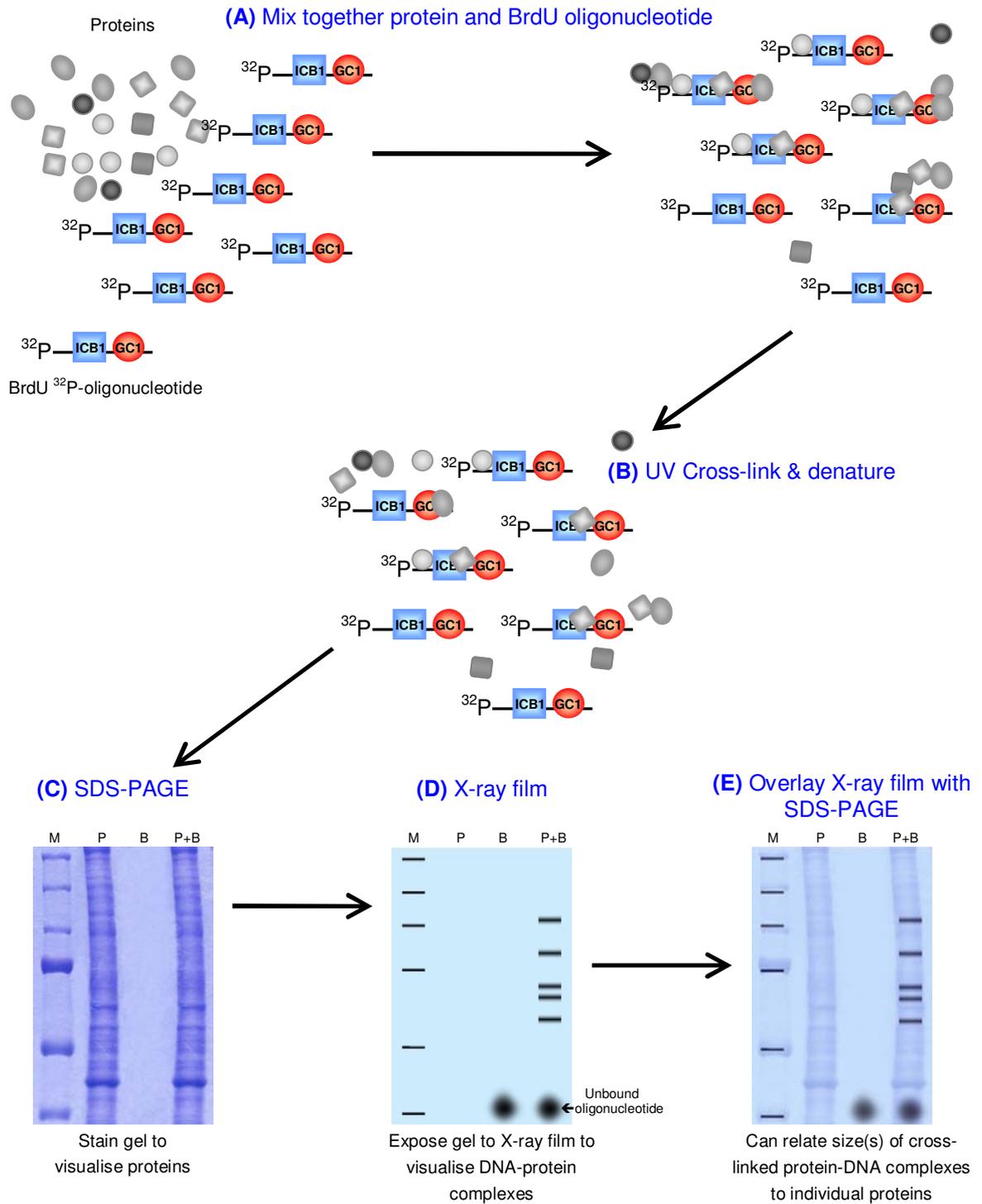
## 5.1 Synopsis.

The previous chapter outlined the methods used in attempts to purify DNA-binding proteins from HeLa extracts. However these methods still generated a large number of proteins, many of which could be interacting with DNA non-specifically. Other methods were therefore required to be able to elucidate which of these DNA-binding proteins were viable for sequence analysis. Up to this point, the sole method for visualisation of DNA-protein interactions was EMSA and due to the non-denaturing properties of EMSA gels the molecular weight of DNA-bound protein complexes could not be estimated. Cross-linking the DNA-binding proteins to a modified ICB1mt/GC1wt oligonucleotide followed by SDS-PAGE was an additional method used to investigate the apparent molecular weight of DNA-bound protein(s).

Using the information provided from these cross-linking assays in conjunction with the sizes of proteins obtained from directly electro-eluting proteins from EMSA gels, a total of 12 proteins were further analysed for sequence identity. This was achieved using either matrix assisted laser desorption ionisation (MALDI) mass spectrometry (MS) or N-terminal sequencing. N-terminal sequencing utilises standard Edman degradative chemistry, sequentially removing amino acid residues from the N-terminus of the protein. Protein identity is confirmed by matching amino acids to the N-terminal sequences of known proteins. Whereas MALDI involves digesting the protein and separating the components based on mass ( $m$ ) and charge ( $z$ ), generating data based on the individual fragmented  $m/z$  ratios. Each protein will generate unique  $m/z$  arrays which can then be used to assist in protein identification.

## 5.2 Using BrdU to cross-link DNA to protein.

Cross-linking assays (a schematic representation of this method is shown in figure 5.1) were carried out using a modified variant of the same composite ICB1mt/GC1wt topoisomerase II $\alpha$  sequence used in EMSA (refer to table 3.1, chapter 3) and DNA-affinity chromatography. Within the intervening sequence of the ICB1mt/GC1wt oligonucleotide each thymidine (T) was replaced with bromodeoxyuridine (BrdU) to aid in UV cross-linking.



**Figure 5.1: Schematic representation of a cross-linking assay.**

(A) BrdU oligonucleotide is mixed with proteins and (B) cross-linked using UV light. The proteins are denatured and SDS-PAGE (C) carried out to separate the proteins based on the molecular weight of each complex. DNA-bound proteins are visualised by exposure to X-ray film (D) and by overlaying the X-ray film over the SDS-PAGE (E) the molecular weight of some of the individual DNA-binding proteins can be established. Any excess oligonucleotide not bound to protein is located at the bottom of the gel. M = protein marker, P = protein, B = BrdU oligonucleotide.

Even though a standard unmodified oligonucleotide can be used for UV cross-linking, BrdU was incorporated in place of thymidine as the bromide atom is more susceptible to free radical formation in the presence of UV radiation. Exposure of the DNA-protein complexes to UV light causes the formation of covalent bonds between DNA and nearby amino acids of the bound protein; thereby the protein and DNA remain irreversibly cross-linked together. Amino acids that can participate in this reaction are arginine, cysteine, histidine, lysine, methionine, serine, phenylalanine, tryptophan, threonine and tyrosine (Molnar *et al.*, 1995). Once the cross-linking has taken place the proteins are denatured by treatment with SDS and  $\beta$ -mercaptoethanol then boiling, followed by standard SDS-PAGE. Any proteins that do not make direct contact with DNA will be denatured and separate from the oligonucleotide however any proteins that are in direct contact with DNA will remain attached to the oligonucleotide due to the cross-linking. All proteins can be visualised by normal staining of SDS-PAGE, however only those proteins that are cross-linked to the  $^{32}\text{P}$ -oligonucleotide can be detected by exposure to X-ray film. Typically UV cross-linking has a success rate of between 0.1 and 20% (Chodosh, 1996; Verdine and Norman, 2003), which means that even though it is possible to cross-link multiple proteins to DNA, the reaction is favoured to generate single protein-DNA complexes in any one single cross-linking event.

To optimise these cross-linking assays a number of parameters were trialed, such as distance from UV light, temperature, incubation and cross-linking times, total volume of the reaction, amount of oligonucleotide and extract, as well as the presence of poly dI.dC or other competitors. These assays were problematic, in that cross-linking was successful in about 50% of experiments. Assays were highly dependent on the quality of the  $^{32}\text{P}$ -labelled BrdU oligonucleotide and volumes of less than 10  $\mu\text{L}$  always cross-linked well (refer to section 2.2.26 materials and methods for details). Nevertheless, the results obtained were reproducible even though they were somewhat tedious to obtain.

### **5.2.1 Incorporating BrdU at different locations within the ICB1mt/GC1wt oligonucleotide.**

The results outlined in chapter 3 demonstrated that an uncharacterised protein complex formed in the region between the ICB1 and GC1 element within the topoisomerase II $\alpha$  promoter, therefore BrdU was incorporated into this intervening sequence. As there are a total of 5 thymidine bases within this sequence, each thymidine was sequentially replaced

with BrdU (summarised in table 5.1) to further delineate where proteins may contact DNA within this intervening sequence. As ICB1 is not required for the formation of complex 3, the ICB1mt/GC1wt oligonucleotide was used in cross-linking assays.

Oligonucleotide	Sequence (5'-3')
ICB1wt/GC1wt	CGA GTC AGG GAT TGG CTG GTC TGC TTC GGG CGG GCT AAA GGA AG
ICB1mt/GC1wt	CGA GTC AGG GAT Tcc CTG GTC TGC TTC GGG CGG GCT AAA GGA AG
BrdU1 (BrdU at position 1)	CGA GTC AGG GAT Tcc CB <sup>1</sup> G GTC TGC TTC GGG CGG GCT AAA GGA AG
BrdU2 (BrdU at position 2)	CGA GTC AGG GAT Tcc CTG GB <sup>2</sup> C TGC TTC GGG CGG GCT AAA GGA AG
BrdU3 (BrdU at position 3)	CGA GTC AGG GAT Tcc CTG GTC B <sup>3</sup> GC TTC GGG CGG GCT AAA GGA AG
BrdU4 (BrdU at position 4)	CGA GTC AGG GAT Tcc CTG GTC TGC B <sup>4</sup> TC GGG CGG GCT AAA GGA AG
BrdU5 (BrdU at position 5)	CGA GTC AGG GAT Tcc CTG GTC TGC TB <sup>5</sup> C GGG CGG GCT AAA GGA AG
BrdUF (forward strand only)	CGA GTC AGG GAT Tcc CB <sup>1</sup> G GB <sup>2</sup> C B <sup>3</sup> GC B <sup>4</sup> B <sup>5</sup> C GGG CGG GCT AAA GGA AG
BrdUall (both strands)	CGA GTC AGG GAT Tcc CB <sup>1</sup> G GB <sup>2</sup> C B <sup>3</sup> GC B <sup>4</sup> B <sup>5</sup> C GGG CGG GCT AAA GGA AG

**Table 5.1: Sequences of oligonucleotides used in EMSAs and cross-linking.**

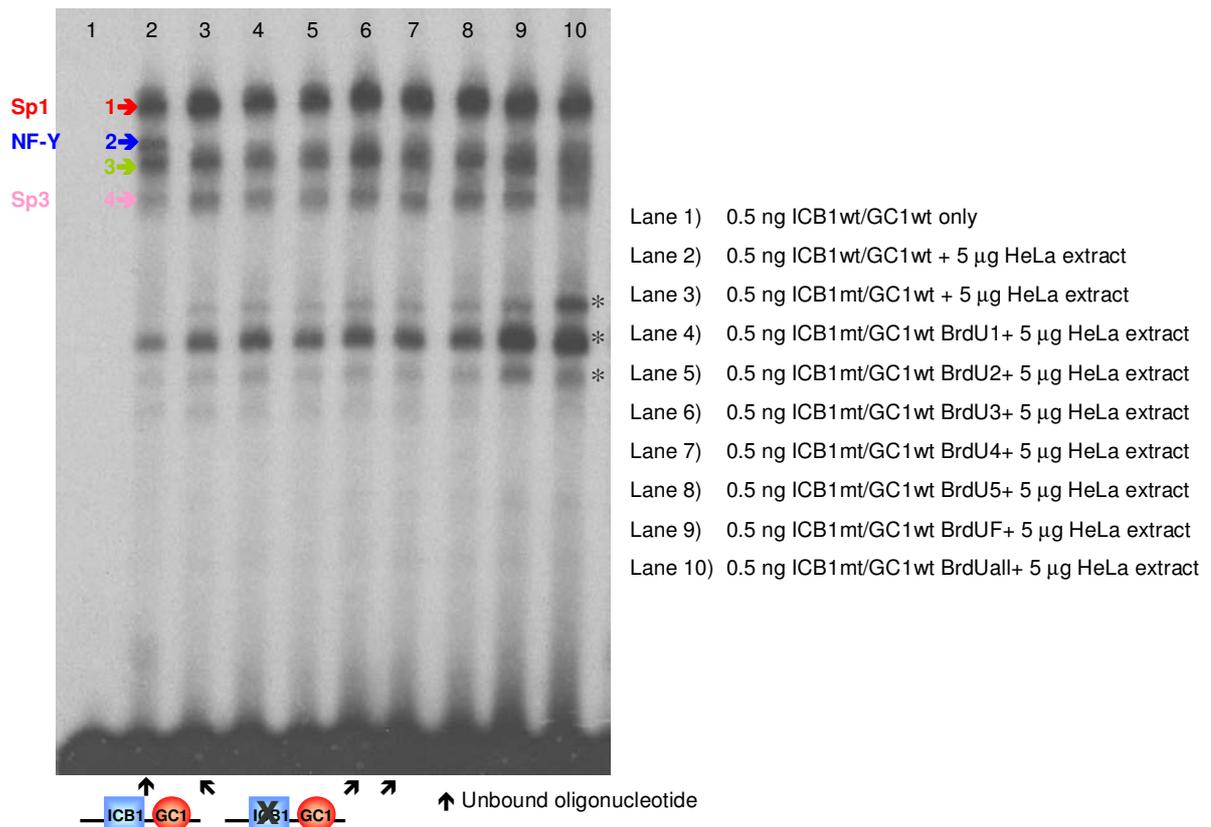
For simplicity, only the top strand of DNA is shown for each of the oligonucleotide pairs. The ICB1 box is highlighted in blue and the GC1 element is pink. Mutations in the sequences are indicated in lower case font and the thymidine that is replaced by BrdU is annotated by B<sup>1-5</sup> within the intervening sequence.

### 5.2.1.1 EMSA using BrdU oligonucleotides.

Individually the BrdU oligonucleotides were examined for standard protein-DNA binding *in vitro* using EMSA. Figure 5.2 shows that each of the BrdU oligonucleotides (lanes 4-10) was capable of generating three protein-DNA complexes (complexes 1, 3 and 4), as would normally occur using the ICB1mt/GC1wt oligonucleotide (lane 3). It should be noted that the BrdUall oligonucleotide (lane 10) appears to have reduced capacity to resolve the band that relates to complex 3 (green arrow), which suggests that the integrity of this oligonucleotide may not be as robust as the other oligonucleotides. It is also possible that this BrdUall oligonucleotide is capable of binding more proteins, thereby resulting in the observed lack of resolution.

There are a number of high mobility protein-DNA complexes (indicated with \*) which appear to increase in intensity within lanes 9 and 10. This suggests that both oligonucleotides containing BrdU at multiple sites may be able to recruit additional proteins that may not normally be required; these high mobility complexes are thought to bind non-specifically to the oligonucleotide. This also provides evidence that the BrdUall oligonucleotide is binding overall more protein than the other BrdU oligonucleotides tested.

### Testing different BrdU oligonucleotides using EMSA



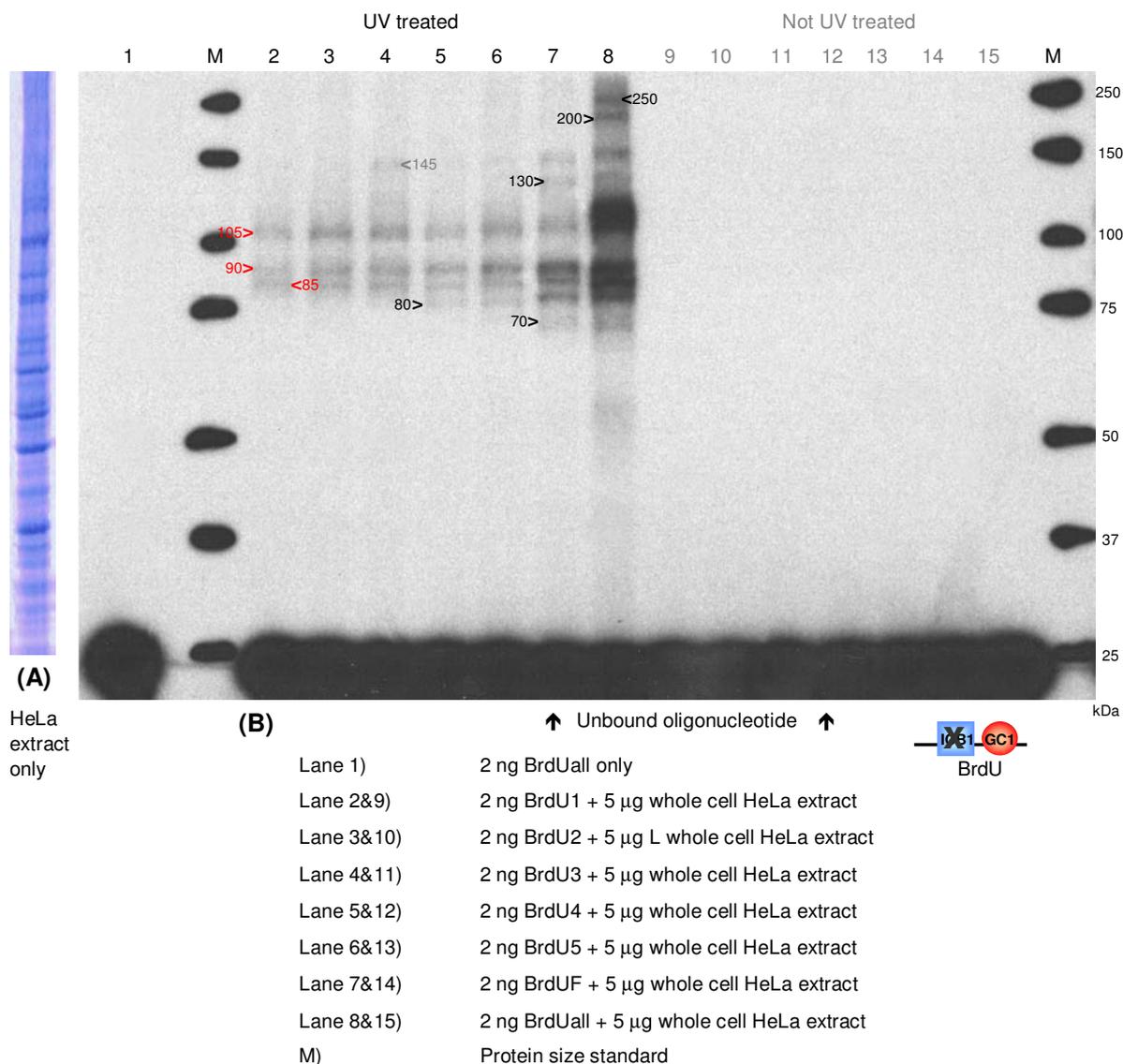
**Figure 5.2: ICB1mt/GC1wt BrdU oligonucleotides tested for protein binding.**

5 µg of whole cell HeLa extract was added to approximately 0.5 ng of the  $^{32}\text{P}$ -oligonucleotide in EMSA reactions. Half of each EMSA reaction (10 µL) was cross-linked by exposure to UVB for 30 minutes before loading onto a 4% polyacrylamide gel and standard EMSA was carried out. The gel was dried onto DE-81 paper and exposed to X-ray film for about 16 hours at  $-80^{\circ}\text{C}$ . The coloured arrows indicate the specific protein-DNA complexes analysed in this study the major protein found in each complex is described. \* denotes the high mobility protein-DNA complexes thought to be non-specific interactions. This figure is representative of triplicate experiments.

#### 5.2.1.2 Cross-linking using BrdU oligonucleotides.

Whole cell HeLa extract was used in conjunction with each of the  $^{32}\text{P}$ -labelled BrdU oligonucleotides in cross-linking assays to investigate the binding parameters of individual proteins within the intervening sequence. Figure 5.3B demonstrates that each BrdU oligonucleotide is capable of forming the same core DNA-protein complexes (illustrated with > in lane 2), one approximately 105 kDa, another approximately 90 kDa and the third approximately 85 kDa in size. In other experiments where the resolution was better, the 105 kDa complex actually separated into two bands of similar size (data not shown), however this is not clear in figure 5.3.

SDS-PAGE of cross-linking experiment using different BrdU oligonucleotides



**Figure 5.3: Different BrdU oligonucleotides used in cross-linking assays.**

(B) 2 ng of each BrdU <sup>32</sup>P-oligonucleotide was mixed with 5 µg whole cell HeLa extract, then either cross-linked using UVB or kept at room temperature for 30 minutes. Samples were denatured and SDS-PAGE carried out using an 8% polyacrylamide gel. The gel was coomassie stained then dried onto DE-81 paper and exposed to X-ray film for about 6 hours at -80°C. > denotes the protein-DNA complexes that appear to be common with all the BrdU oligonucleotides, < denotes other protein-DNA complexes and approximate sizes of these bands are as annotated. Protein size standard is visualised using fluorescent paint. These results are representative of triplicate experiments. (A) Is a representation of a Coomassie blue stained SDS-PAGE of whole cell HeLa extract, to demonstrate the protein complexity of this extract.

Based on the molecular weight of some of the bands seen in figure 5.3, it could be speculated that the 105 kDa complex is composed of Sp3-oligonucleotide, and that the 90 kDa complex is made up of Sp1 associating with the oligonucleotide. In other cross-linking assays there is

occasionally the appearance of a 95 kDa band (data not shown) which could be the phosphorylated version of Sp1. However, it should not be forgotten that there are no “T” bases (therefore no incorporated BrdU) in the GC element to which both Sp1 and Sp3 are found to bind. This does not eliminate the possibility that Sp1 or Sp3 could be recruited to the intervening sequence through an association with proteins bound nearby. In addition, non-BrdU oligonucleotides will still cross-link proteins to a certain degree (data not shown) therefore it is possible that Sp1 and Sp3 may be cross-linked to the oligonucleotide.

When examining lanes 2 through to 6 there appears to be an increasing number of protein-DNA complexes forming (shown with >), suggesting that the position where BrdU is incorporated may play a role in the cross-linking of different protein components. There is an 80 kDa protein-DNA complex which only appears in lanes 5-8, in these lanes BrdU is incorporated in positions close to the GC1 element. Therefore it is possible a particular protein component is preferentially binding to this region within the intervening sequence. In other lanes there is a larger complex approximately 145 kDa in size that only appears with BrdU3 (lane 4), BrdUF (lane 7) and BrdUall (lane 8). This 145 kDa complex also appears with weak affinity for the other BrdU oligonucleotides when more oligonucleotide is used (data not shown), suggesting this band is not just an artefact of BrdU3, BrdUF and BrdUall. However, this 145 kDa (>) complex is readily formed when BrdU is in position 3, than any of the other locations. This may demonstrate that a particular protein component is binding directly at position 3 in the intervening sequence between ICB1 and GC1. With the oligonucleotides that have BrdU incorporated at all positions (lanes 7 and 8) there is an additional band approximately 70 kDa in size which can also resolve as two bands close together when separation is better (data not shown). These 70 kDa bands are usually quite faint, suggesting that the particular protein(s) involved in the formation of these complexes are either in low abundance or that the protein-DNA interactions are fragile or possibly even transient.

There are two high molecular weight complexes, approximately 250 kDa and 200 kDa in size (lane 8). The 200 kDa complex is often observed with the other BrdU oligonucleotides (data not shown) when the oligonucleotide is used in great excess and cross-linking has worked well, but the 250 kDa complex is only observed with BrdUF and BrdUall. It is possible these larger complexes are made up of multiple proteins associating with the oligonucleotide at different regions of the intervening sequence as both these oligonucleotides have BrdU

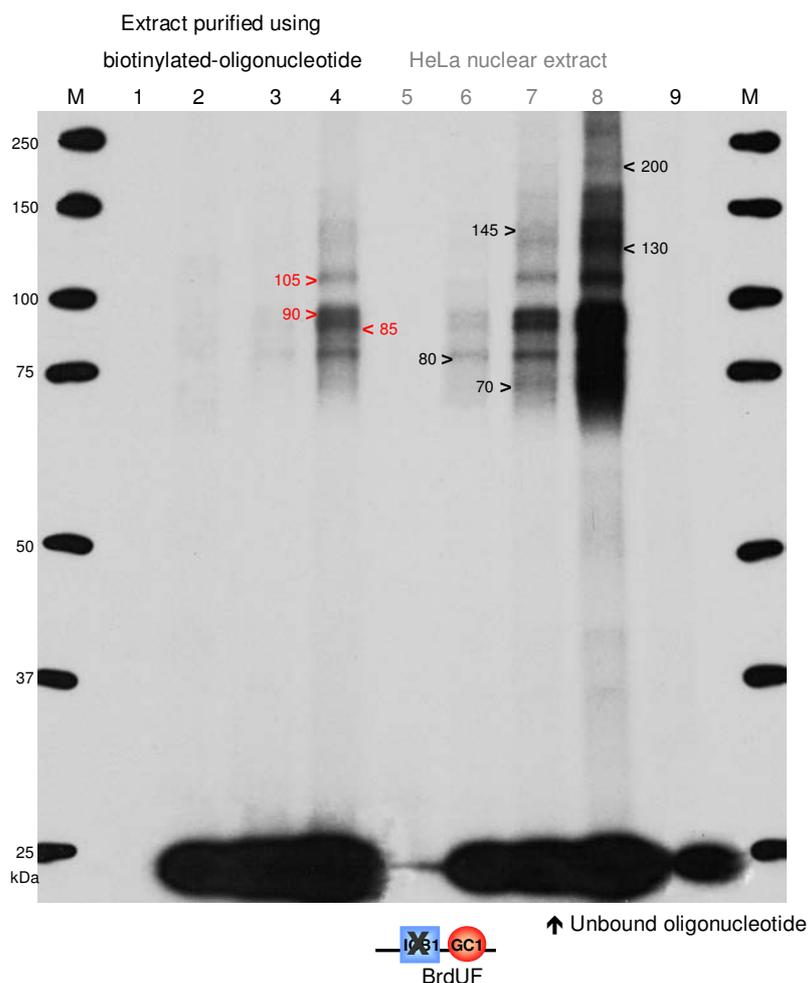
incorporated at multiple sites, although this is considered unlikely to occur due to the low frequency of cross-linking events. There is also a 130 kDa complex that appears only when the oligonucleotides have BrdU incorporated at all possible locations (lane 7 and 8 figure 5.3). Overall, the oligonucleotide that exhibits the greatest number of protein-DNA complexes was BrdUall. This is most likely due to the greatest amount of BrdU being incorporated into this oligonucleotide therefore more overall cross-linking would be expected, possibly even at multiple sites. However, it was also the BrdUall oligonucleotide that gave questionable mobility for complex 3 in EMSA (figure 5.2, lane 10) as the bands did not clearly resolve, therefore cross-linking results obtained using this particular oligonucleotide were treated tentatively.

Lane 1 was included as a control to demonstrate that in the absence of protein no bands are observed. The BrdU oligonucleotides are not capable of forming complexes on their own. In the absence of exposure to UV light (lanes 9-15) no protein-DNA complexes can be seen, indicating that these protein-DNA complexes are dependent on the successful cross-linking of protein to <sup>32</sup>P-labelled oligonucleotide. Generally, whole cell HeLa extract is a complicated mixture of proteins (figure 5.3A). The fact that so few proteins from this mixture were found to cross-link to the BrdU oligonucleotides provides an indication of how specific these protein-DNA interactions are and also that they are in low abundance, as would be expected for transcription factors.

### **5.2.1.3 Cross-linking nuclear extract & biotinylated extract.**

Electro-eluted samples were useful in demonstrating individual sizes of proteins that could be present in complex 3; however the electro-eluted samples were too dilute to use directly for N-terminal sequencing or MALDI-MS. Therefore the extract purified using the biotinylated-ICB1mt/GC1wt oligonucleotide (oligonucleotide-affinity chromatography) was used as it could be obtained in higher concentrations. The previous chapter demonstrated that both these extracts did contain proteins of similar size and both were enriched for factors that were known to bind to the ICB1mt/GC1wt sequence. Unfortunately, EMSA could not be used to show conclusively that these extracts contained all the required DNA-binding proteins. Therefore cross-linking assays were now used to investigate the proteins present in the extract purified using oligonucleotide-affinity chromatography. Increasing amounts of BrdUF oligonucleotide was added to cross-linking reactions and compared with cross-linked proteins from HeLa nuclear extract, the results can be seen in figure 5.4.

### SDS-PAGE of cross-linking experiment using increasing amounts of BRdUF



- Lane 1) 500 ng extract purified using biotinylated-ICB1mt/GC1wt oligonucleotide only
- Lane 2) 1 ng BrdUF + 500 ng extract purified using biotinylated-ICB1mt/GC1wt oligonucleotide
- Lane 3) 2 ng BrdUF + 500 ng extract purified using biotinylated-ICB1mt/GC1wt oligonucleotide
- Lane 4) 3 ng BrdUF + 500 ng extract purified using biotinylated-ICB1mt/GC1wt oligonucleotide
- Lane 5) 2 µg HeLa nuclear extract only
- Lane 6) 1 ng BrdUF + 2 µg HeLa nuclear extract
- Lane 7) 2 ng BrdUF + 2 µg HeLa nuclear extract
- Lane 8) 3 ng BrdUF + 2 µg HeLa nuclear extract
- Lane 9) 1 ng BrdUF only
- M) Protein size standard

**Figure 5.4: Biotinylated-oligonucleotide purified extract contains DNA-binding proteins.** Increasing amounts of BrdUF <sup>32</sup>P-oligonucleotide was mixed with extract purified using biotinylated-oligonucleotide or HeLa nuclear extract and then cross-linked using UVB. Samples were denatured before SDS-PAGE using an 8% polyacrylamide gel. The gel was dried onto DE-81 paper and exposed to X-ray film for about 20 hours at -80°C. > denotes the core protein-DNA complexes, > denotes other protein-DNA complexes. Protein size standards are visualised using fluorescent paint. These results are representative of duplicate experiments.

Figure 5.4 shows that in the absence of any BrdUF oligonucleotide (lanes 1 and 5) or protein (lane 9) no DNA-binding patterns were observed. With increasing amounts of oligonucleotide (comparing lanes 6-8) increased intensity and frequency of protein-DNA interactions were observed. This illustrates how these assays were highly dependent on the amount of oligonucleotide that was used in each cross-linking reaction. When too much oligonucleotide was used (lane 8) the resolution of individual bands was poor, however when too little oligonucleotide was used (lane 6), detection was compromised.

Overall (figure 5.4), the same common three protein-DNA complexes were observed with both sets of extracts (compare > lane 4 purified and lane 7 HeLa nuclear extract). These were 105, 90 and 85 kDa in size as well as the other previously observed protein-DNA complexes of 80 and 145 kDa. Generally, HeLa nuclear extract had enhanced protein binding (lane 7) resulting in more intense bands than those seen with extract purified using biotinylated-ICB1mt/GC1wt oligonucleotide (lane 4). This is most likely due to the higher concentration of proteins in the nuclear extract, but it is also possible that this extract also contains other factors that are required to enhance the recruitment of proteins to their cognate sequence. Nevertheless, both extracts yielded similar binding patterns when cross-linked and accordingly should contain the same population of DNA-binding proteins. Extract purified using oligonucleotide-affinity chromatography was therefore deemed suitable for isolation of proteins for amino acid sequencing or peptide analysis.

### **5.3 Analysis of potential DNA-binding proteins.**

Using the information gathered from electro-eluted extracts, purified extracts and cross-linking assays a total of 12 proteins were further analysed for protein identity (summarised in Table 5.2). Extracts purified using biotinylated-ICB1mt/GC1wt oligonucleotide were subjected to SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membrane prior to sequence analysis. Three of these proteins were sent for N-terminal sequencing (highlighted in blue in table 5.2). The remainder were sent for analysis using mass spectrometry, where proteins of the appropriate sizes were carefully excised from Coomassie stained SDS-PAGE gels. All samples were sent to the Australian Proteome Analysis Facility (APAF), Sydney, Australia.

Sizes of proteins in e/e3 extract	Sizes of proteins in Biotinylated-ICB1mt/GC1wt	Sizes of proteins from cross-linking assays	Proteins sent for identification
	<b>250+</b>		
250	<b>250</b>	250	250
200	200	200	
	<b>150</b>		150
145	145	145	145
130		130	
120	<b>120</b>		
115	115		115
		108	
105	105	105	105
<b>95</b>	95	95	98 & 95
	90	90	89&87
85	85	85	85
80	<b>80</b>	80	
<b>78</b>			
74			
72		73	
<b>70</b>	<b>70</b>	70	70
65	65		
62	62		

**Table 5.2: Summary of the sizes of isolated DNA-binding proteins**

The results of different purification techniques and cross-linking assays were compared to determine which proteins should be sent for analysis. Most of the DNA-binding proteins are within the 85-150 kDa size range. The proteins highlighted in blue are those that were analysed by N-terminal sequencing.

### 5.3.1 Results of protein analysis.

The proteins analysed by N-terminal sequencing were subjected to 12 cycles of automated Edman degradation (494 Procise protein sequencing system, Applied Biosystems). The 85 kDa protein resolved 9 amino acids and the 115 kDa protein resulted in 10 N-terminal amino acids (refer to appendix 3). These sequences were used to search for protein identity utilising BLAST (<http://www.ncbi.nih.gov/BLAST/>) limiting the search for proteins of “short nearly exact matches” and also restricting the parameters to Homo sapiens only. The results are summarised in Table 5.3. Unfortunately the 75 kDa protein returned no sequence information, most likely due to insufficient amounts of protein.

The proteins subjected to mass spectrometry analysis were treated as specified by APAF and the best match(es) were provided by the operator (refer to appendix 4 for summary of data). These protein identities were scrutinized and many were found to have a number of aliases. Table 5.3 summarises the most probable identity for each protein based on the probability based Mowse score (provided by APAF), putative function of the protein, molecular weight and name of the protein originated from a Homo sapien source.

<b>Proteins sent for identification (kDa)</b>	<b>Proposed identity</b>
250	Human Myosin-9 (MYH9)
150	Human Splicing factor 3B subunit 1(SF3B1)
145	Human ATP-dependent RNA helicase A (DXH9)
120	No ID
115	Poly (ADP-ribose) polymerase (PARP-1)
105	No ID
98	Polypyrimidine tract-binding protein (PSF)
95	Polypyrimidine tract-binding protein (PSF)
89	Human putative pre-mRNA splicing factor ATP-dependent RNA helicase (DHX15)
87	Nuclear family 90 (NF90)
85	Ku autoantigen 86 (Ku86)
75	No ID

**Table 5.3: Summary of the results from protein analysis.**

Each protein was analysed for identity using either N-terminal sequencing (in blue) or mass spectrometry (in black). Three of the protein samples were unable to be identified.

## 5.4 Confirming identity of DNA-binding proteins.

Although great care was taken whilst preparing and excising proteins for analysis from SDS-PAGE it was not always possible to be certain of the homogeneity of the excised band. A higher than normal concentration of proteins was required for this process which meant that resolution within SDS-PAGE was compromised, especially when two or more bands were similar in size. Therefore it is likely that some of the proteins analysed may have been contaminated by proteins from nearby bands. Although protein identity in many cases was successful, this may not necessarily represent specific DNA-binding proteins that associate with the topoisomerase II $\alpha$  promoter. Some of these proteins may be functionally redundant

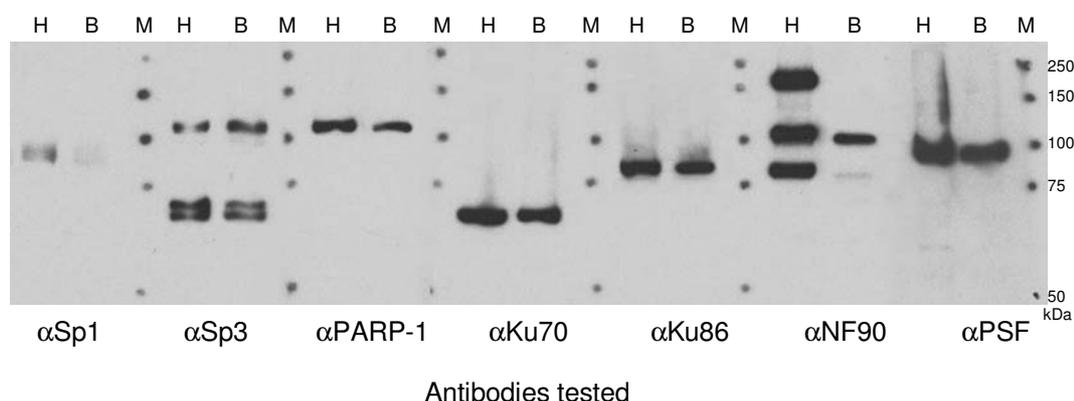
and have no role in gene regulation. For example, it is difficult to reconcile a role for transcriptional regulation for the protein identified as MYH9. All the other proteins however appear to have functions with some relevance to the regulation of gene expression. Some are typically known to be involved in double-strand break repair (Ku86 and PARP-1), whilst others appear to be important for pre-mRNA processing (PSF, SF3B1 and DHX9).

Ku86 and PARP-1 have been extensively studied for their roles in double-stranded break repair, but there is an ever expanding amount of evidence that connects these proteins with roles in gene regulation as well. Unfortunately, N-terminal sequencing failed (twice) for the 70 kDa protein. It was rationalised as Ku86 and Ku70 form heterodimers *in vivo* the 70 kDa protein could be Ku70. PSF has been found to have a heterodimeric partner p54<sup>nrb</sup> which have been shown to interact with both Ku86 and Ku70 (Bladen *et al.*, 2005) and the C-terminal domain of RNA polymerase II (Emili *et al.*, 2002; Rosonina *et al.*, 2005). Some studies have indicated that PSF is capable of repressing transcription by binding directly to a promoter (Mathur *et al.*, 2001; Song *et al.*, 2005) and perhaps acting as a tumour suppressor protein (Garen and Song, 2008). NF90 consists of a number of proteins which belong to this family with two 110 kDa and three 90 kDa variants (Reichman *et al.*, 2003). NF90 has also been shown to interact with NF-Y (Zhao *et al.*, 2005) as well as Ku86 and Ku70 (Shi *et al.*, 2007), therefore it is possible it has a regulatory role. The following chapter (chapter 6) addresses the potential functions and binding specificity of some of these proteins in transcriptional regulation with more detail.

As a number of proteins were identified by N-terminal sequencing or MALDI it was first important to confirm the presence of these in the purified extracts. This was carried out using immunoblotting (figure 5.5) performed with antibodies against some of these newly identified proteins. The extract purified using oligonucleotide-affinity chromatography was tested for the presence of these proteins, as it was the source for N-terminal sequencing and MALDI analysis. Whole cell HeLa extract was included as the positive control, as well as antibodies against Sp1 and Sp3 as these are known to be present in the oligonucleotide-affinity purified extract. Antibody against tubulin was also used as a negative control. Tubulin was not present within any of the extracts purified using oligonucleotide-affinity chromatography (data not shown). Each of the newly identified proteins PARP-1 (110-115 kDa), Ku70 (70-75 kDa), Ku86 (85-90 kDa), NF90 (90 or 110 kDa) and PSF (100 kDa) were

positively identified to be present in the extract purified using oligonucleotide-affinity chromatography (figure 5.5).

**Immunodetection of proteins present in extracts purified using biotinylated-ICB1mt/GC1wt oligonucleotide**



**Figure 5.5: Confirming the identity of DNA-binding proteins found in extract purified using the biotinylated-ICB1mt/GC1wt oligonucleotide.**

20  $\mu$ g of whole cell HeLa extract (H) or 5  $\mu$ g of extract purified using oligonucleotide-affinity chromatography (B) was used in each immunoreaction. SDS-PAGE was carried out on a 8% polyacrylamide and proteins were transferred onto a positively charged nylon membrane for 1 ½ hours at 450 mA.  $\alpha$ Sp1,  $\alpha$ Sp3 and  $\alpha$ PSF were diluted 1/500,  $\alpha$ Ku70,  $\alpha$ Ku86 and  $\alpha$ PARP-1 were diluted 1/600,  $\alpha$ NF90 was diluted 1/400. HRP conjugated 2<sup>o</sup> antibodies,  $\alpha$ rabbit was diluted 1/3500 and  $\alpha$ mouse was diluted 1/4000. Chemiluminescence was visualised by exposure to X-ray film. M = protein size standard visualised with the use of fluorescent paint.

Antibody against NF90 recognised a number of proteins in HeLa extract, which were not specified by the manufacturer, it is possible that they appear due to the different number of variants for this protein. NF90 also has a heterodimeric partner, NF45 (Reichman *et al.*, 2003) therefore it is possible the high molecular weight band observed in HeLa extract is due to insufficient denaturation before carrying out SDS-PAGE. Alternatively, some non-specific cross-reactivity may have occurred giving rise to several non-specific bands within the immunoblot.

## 5.5 Chapter summary.

Overall, the proteins PARP-1, Ku86, Ku70, NF90 and PSF were positively identified as potential DNA-binding proteins, they were all confirmed to be present in the extract purified using the biotinylated-ICB1mt/GC1wt oligonucleotide. This suggests that these proteins are capable of binding to sequences within the human topoisomerase II $\alpha$  promoter.

The cross-linking assays enabled the identification of the molecular weight of individual DNA-binding proteins that possibly bind to the intervening sequences between ICB1 and GC1. Three main protein-DNA complexes were observed to bind to the ICB1mt/GC1wt BrdU oligonucleotides; they were 105, 90 and 85 kDa in size. The 85 kDa protein-DNA complex could possibly be Ku86 and the 105 kDa protein-DNA complex may be composed of Sp3. Unfortunately this 105 kDa protein was unable to be identified via MALDI analysis as no identity could be assigned to the data by the operator. The 90 kDa protein-DNA complex could be composed of NF90 or Sp1 based on size, or perhaps even PSF as this protein is known to run autonomously due to being highly basic. There was an 80 kDa protein that was found to bind preferentially to oligonucleotides which had BrdU incorporated close to GC1 element and this may indicate an association with either Sp1 or Sp3. Unfortunately this 80 kDa protein was overlooked when selecting proteins for sequence analysis. A 145 kDa protein appears to bind preferentially at position 3 (right in the centre) of the intervening sequence but it was not possible to positively identify this protein as DXH9, as specific antibodies were not available at this time. Even though protein sequencing failed for the 70 kDa protein, the prediction that this protein was Ku70 appears to be correct and it does appear to cross-link to BrdU oligonucleotides. Although not all the proteins that were analysed provided conclusive identification, several candidates are of potential interest and were examined further for binding and functionality as described in the following chapter.

# 6 Analysis of DNA binding proteins.

## 6.1 Synopsis.

The previous chapters described the purification and identification of DNA-binding proteins, with the potential to both, interact with, and regulate the topoisomerase II $\alpha$  promoter. A selection of these (Ku86, Ku70, PARP-1, NF90 and PSF) were examined further to determine which may have a significant role in gene regulation. Antibodies against these proteins were used in EMSA (antibody supershifts) in an attempt to confirm specific interactions with the composite ICB1/GC1 sequence. Unfortunately, no distinctive mobility shift(s) were detected in any of the DNA-protein complexes making it difficult to confirm specific interactions. It was possible that multiple protein-protein interactions were blocking the antigenic sites for the antibodies used in the assays. Alternatively, these proteins may not directly interact with the oligonucleotide, or may only have been minor components in the protein-DNA complexes and thus supershifts would be difficult to detect. Therefore alternative methods were adopted. Ribosylation was inhibited as a method for assessing a function for PARP-1 in regulation of topoisomerase II $\alpha$ . Immunodetection was used to verify the presence of proteins in extracts that were generated by directly eluting proteins from DNA-protein complexes in EMSA gels. In addition, a DNA-termini binding assay was developed to establish the binding parameters of these proteins.

## 6.2 Inhibiting Poly (ADP-ribose) polymerase activity.

Poly (ADP-ribose) polymerase (PARP-1) has been extensively studied as a protein that is able to bind to DNA as well as being able to catalyse the post-translational ribosylation of nuclear enzymes (D'Amours *et al.*, 1999). PARP-1 has the ability to add mono or poly ADP-ribose groups to acceptor proteins. Although this modification is transient *in vivo* it has the ability to make proteins increasingly negatively charged and thereby alter their affinity for DNA as well as for other proteins. Ribosylation is a rapid process and one of the main targets for ribosylation is PARP-1 itself (Adamietz, 1987). Automodification of PARP-1 can inhibit PARP-1 from binding to DNA, inhibit protein interactions and inhibit the ADP-ribosyltransferase activity of PARP-1 (Althaus, 1992; Yung *et al.*, 2004; Mendoza-Alvarez & Alvarez-Gonzalez, 1993). A number of potent PARP-1 inhibitors are commercially available and are regularly used as a means of inhibiting ribosylation and hence assaying PARP-1

activity. In this study HeLa cells were treated with the PARP-1 inhibitor PJ34 for 48 hours in a range of concentrations that cover those that have been used in other studies 0.5  $\mu$ M (Luo *et al.*, 2003) - 1  $\mu$ M (Zaniolo *et al.*, 2007). PJ34 treatment did not appear to affect HeLa cell density, viability or growth.

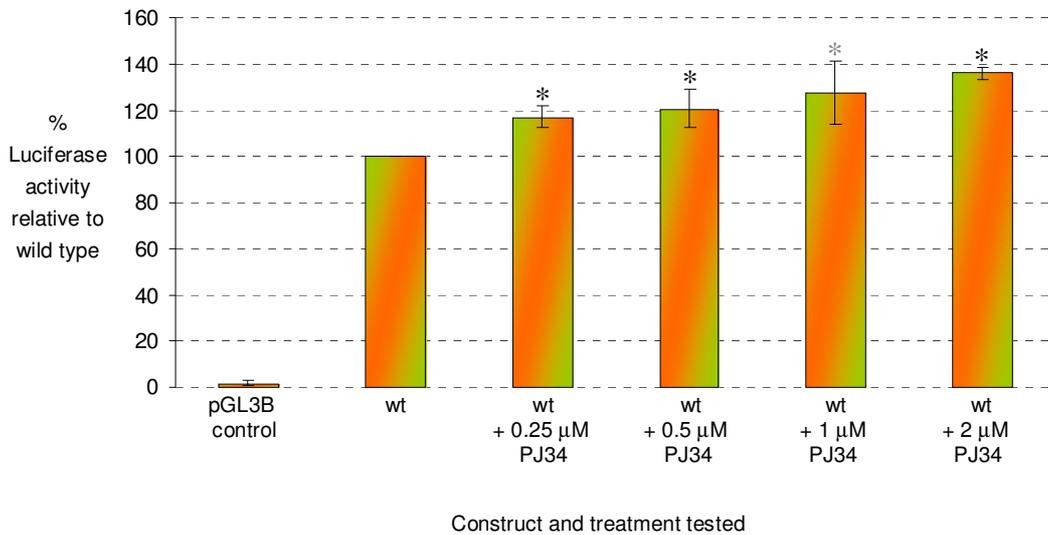
### **6.2.1 PARP-1 & the topoisomerase II $\alpha$ promoter.**

The -617topo2 $\alpha$  promoter had previously been cloned into a pGL3B vector upstream of a luciferase reporter gene to monitor promoter activity. Mutations in the GC1 and GC2 elements were created with site directed mutagenesis (Magan *et al.*, 2003) and the intervening sequence between ICB1 and GC1 was extended by PCR mutagenesis (Senior K., Msc Thesis, 2006). Plasmid identity was confirmed by sequencing using standard primers GL2 and RV3, before using these reporter gene constructs in transient transfections.

To determine if PARP-1 activity had any effect on basal topoisomerase II $\alpha$  promoter activity, transient transfections were carried out in HeLa cells using the wt -617topo2 $\alpha$  reporter gene construct in conjunction with increasing amounts of PARP-1 inhibitor, PJ34. The results of which are shown in figure 6.1. With the addition of 0.25  $\mu$ M PJ34, luciferase activity increases by approximately 15% and this gradually increases up to 30% when PJ34 treatment was increased to 2  $\mu$ M. Each increase was found to be statistically different to the wt -617topo2 $\alpha$  reporter gene construct that was not treated with PJ34. This suggests that PARP-1 inhibition may result in a slight activation of the topoisomerase II $\alpha$  promoter.

As described in chapter 3 when the intervening sequence between ICB1 and GC1 was extended by inserting an additional 6 bp sequence, more of complex three was found to bind to this oligonucleotide (refer to figure 3.5). This increase in binding was also shown to correspond with an increase in topoisomerase II $\alpha$  promoter activity (refer to the CENins construct in figure 3.9). Therefore the CENins construct was used in conjunction with PJ34 treatment to investigate if the proteins involved in the formation of complex three were influenced by PARP-1 activity. Figure 6.2 shows that there is an overall increase in luciferase activity when treated with PJ34. However, this increase is no greater than that which was seen when the same experiment was carried out using the wt -617topo2 $\alpha$  reporter gene construct (figure 6.1).

**THE EFFECTS OF PARP INHIBITION ON  
wt -617TOPOISOMERASE II $\alpha$  PROMOTER ACTIVITY**

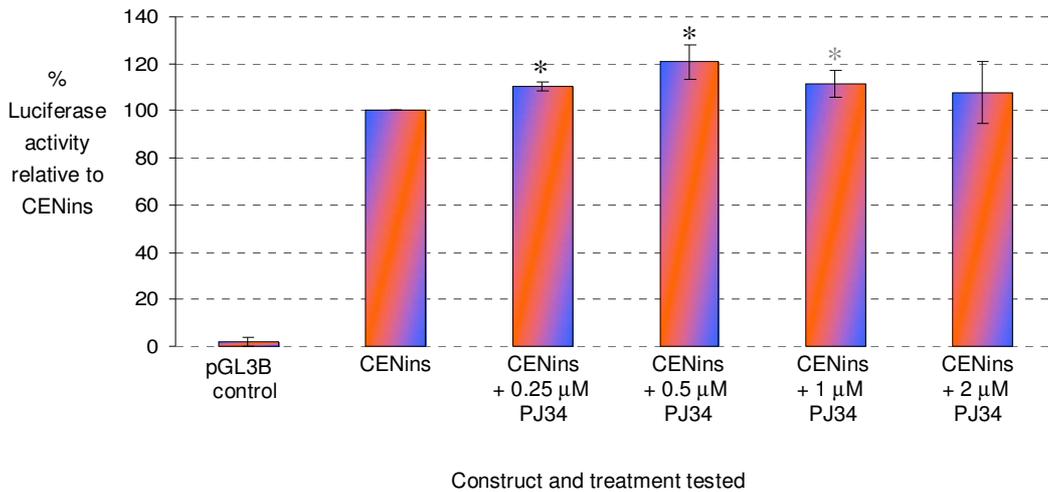


Construct & Treatment	Luciferase activity relative to WT (%)	p-values (% similarity)	
pGL3B	1.7 ± 0.98 %		
wt -617topo2 $\alpha$	100 %		
wt + 0.25 $\mu$ M PJ34	117 ± 4.6 %	0.020	2%
wt + 0.5 $\mu$ M PJ34	120.1 ± 8.2 %	0.048	4.8%
wt + 1 $\mu$ M PJ34	127.3 ± 13.7 %	0.062	6.2%
wt + 2 $\mu$ M PJ34	136 ± 2.7 %	0.002	0.2%

**Figure 6.1: Transient transfection to show the effects of a PARP-1 inhibitor (PJ34) on wt -617topo2 $\alpha$  promoter activity.**

HeLa cells were transfected with 0.25  $\mu$ g of pCMV SPORT- $\beta$ -gal plasmid along with 0.5  $\mu$ g of wt -617topo2 $\alpha$  pGL3B reporter construct then treated with PJ34 after 2 hours. An empty pGL3B plasmid was used as a control to show no luciferase activity was possible in the absence of a cloned promoter. The normalised luciferase activities are displayed as a percentage relative to wt activity with no PJ34 treatment. The results are averages of a set of three experiments all carried out in triplicate on each occasion. The errors bars are calculated as average deviations between data sets and the asterix indicates that there is a significant difference between wt and PJ34 treatment data sets. \* specifies that there is a less than 5% chance of the data set being similar to wt (or >95% chance that they are different). \* specifies there is a 5-10% chance of the data set being similar to wt (or 90-95% chance that they are different).

**THE EFFECTS OF PARP INHIBITION ON  
CENins -617TOPOISOMERASE II $\alpha$  PROMOTER ACTIVITY**



Construct & Treatment	Luciferase activity relative to CENins (%)	p-values (% similarity)	
pGL3B	1.9 ± 1.7 %		
CENins -617topo2 $\alpha$	100 %		
CENins + 0.25 $\mu$ M PJ34	110.3 ± 1.8 %	0.009	0.9%
CENins + 0.5 $\mu$ M PJ34	121 ± 7.3 %	0.032	3.2%
CENins + 1 $\mu$ M PJ34	111.7 ± 5.8 %	0.061	6.1%
CENins + 2 $\mu$ M PJ34	108 ± 13.3 %	0.272	27.2%

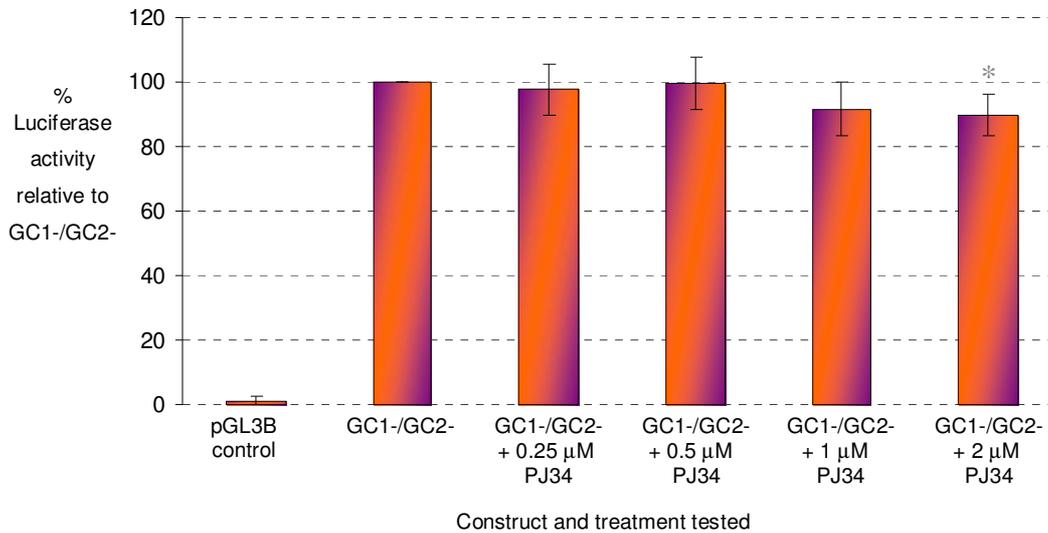
**Figure 6.2: Transient transfection to show the effects of a PARP-1 inhibitor (PJ34) on CENins -617topo2 $\alpha$  promoter activity.**

HeLa cells were transfected with 0.25  $\mu$ g of pCMV SPORT- $\beta$ -gal plasmid along with 0.5  $\mu$ g of CENins -617topo2 $\alpha$  pGL3B reporter construct then treated with PJ34 after 2 hours. An empty pGL3B plasmid was used as a control to show no luciferase activity was possible in the absence of a cloned promoter. The normalised luciferase activities are displayed as a percentage relative to CENins activity with no PJ34 treatment. The results are averages of a set of three experiments all carried out in triplicate on each occasion. The errors bars are calculated as average deviations between data sets and the asterisk indicates that there is a significant difference between CENins and PJ34 treatment data sets. \* specifies that there is a less than 5% chance of the data set being similar to CENins (or >95% chance that they are different). \* specifies there is a 5-10% chance of the data set being similar to CENins (or 90-95% chance that they are different).

Interestingly, there was no significant difference between the CENins construct and when it was treated with 2  $\mu$ M PJ34. In fact the 2  $\mu$ M PJ34 treatment had the most variability in luciferase activity, individual results ranging from 85-120 % of the construct with no PJ34 treatment. This variability in range is mostly likely due to slight variations in cell density and efficacy of PJ34 treatment rather than transfection efficiency. Overall these data suggest that the increase in promoter activity observed with the CENins construct is not greatly affected by treatment with PJ34; hence PARP-1 activity may not specifically affect the formation of complex three and the proteins involved in this process.

It has been demonstrated that PARP-1 can ribosylate Sp1 thereby decreasing its ability to bind to DNA and activate transcription (Zaniolo *et al.*, 2007). In a previous study it was shown that Sp1 has the ability to activate the topoisomerase II $\alpha$  promoter (Magan *et al.*, 2003) through a direct interaction with the GC1 and GC2 elements. To determine if the increase that was observed with PJ34 treatment was dependent on the presence of Sp1, both the GC1 and GC2 elements (GC1-/GC2-) were mutated within the -617topo2 $\alpha$  reporter gene construct. This GC1-/GC2- reporter gene construct was used in transient transfections along with PJ34 treatments. Figure 6.3 demonstrates that luciferase activity does not increase in the presence of PJ34, both 0.25  $\mu$ M and 0.5  $\mu$ M PJ34 treatment had little to no effect on luciferase activity. There is a slight, but not significant decrease in activity with 1  $\mu$ M PJ34. However, with 2  $\mu$ M PJ34 treatment there is a significant 5-10% decrease in luciferase activity suggesting that promoter repression may be occurring when high concentrations of PJ34 was used.

**THE EFFECTS OF PARP INHIBITION ON  
GC1-/GC2- -617TOPOISOMERASE II $\alpha$  PROMOTER ACTIVITY**



Construct & Treatment	Luciferase activity relative to GC1-/GC2- (%)	p-values (% similarity)	
pGL3B	1.3 ± 1.3 %		
GC1-/GC2- -617topo2 $\alpha$	100 %		
GC1-/GC2- + 0.25 μM PJ34	97.8 ± 7.9 %	0.360	36%
GC1-/GC2- + 0.5 μM PJ34	99.5 ± 8.0 %	0.477	48%
GC1-/GC2- + 1 μM PJ34	91.5 ± 8.5 %	0.111	11%
GC1-/GC2- + 2 μM PJ34	89.8 ± 6.6 %	0.056	5.6%

**Figure 6.3: Transient transfection to show the effects of a PARP-1 inhibitor (PJ34) on -617topo2 $\alpha$  promoter activity with mutations in GC1 and GC2.**

HeLa cells were transfected with 0.25  $\mu$ g of pCMV SPORT- $\beta$ -gal plasmid along with 0.5  $\mu$ g of GC1-/GC2- -617topo2 $\alpha$  pGL3B reporter construct then treated with PJ34 after 2 hours. An empty pGL3B plasmid was used as a control to show no luciferase activity was possible in the absence of a cloned promoter. The normalised luciferase activities are displayed as a percentage relative to GC1-/GC2- activity with no PJ34 treatment. The results are averages of a set of three experiments all carried out in triplicate on each occasion. The errors bars are calculated as average deviations between data sets and the asterisk indicates that there is a significant difference between GC1-/GC2- and PJ34 treatment data sets. \* specifies that there is a less than 5% chance of the data set being similar to GC1-/GC2- (or >95% chance that they are different) \* specifies there is a 5-10% chance of the data set being similar to GC1-/GC2- (or 90-95% chance that they are different).

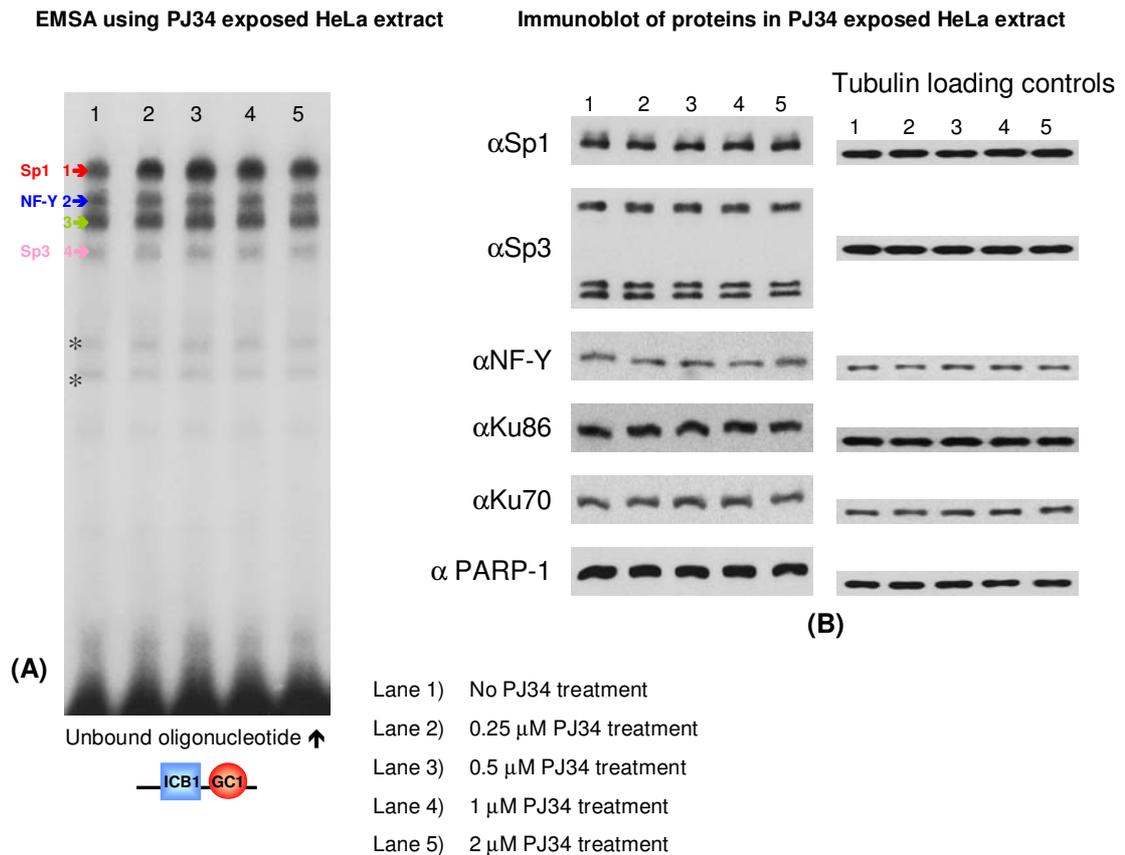
Overall the transient transfection data demonstrates that PJ34 does have an effect on topoisomerase II $\alpha$  promoter activity. Using the wt and CENins reporter gene constructs PJ34 appeared to enhance promoter activity. This suggests that the ribosylation activity of PARP-1 may have an inhibitory role in topoisomerase II $\alpha$  gene regulation. It is possible that PARP-1 may be interacting directly with DNA or acting through protein-protein interactions at the promoter. The results of the transient transfection assays using the GC1-/GC2- constructs suggest that the latter is more likely as the up-regulation caused by PJ34 is not observed when Sp1 or Sp3 are unable to interact with the topoisomerase II $\alpha$  promoter. Both Sp1 and Sp3 bind to the GC2 element (Magan *et al.*, 2003) which implicates either of these proteins as possible targets of PARP-1-ribosylation activity. However, it has been shown that Sp3 does neither bind to, nor get ribosylated by PARP-1 (Zaniolo *et al.*, 2007), making it more likely that Sp1 is involved in the observed PARP-1-induced gene regulation. It is possible this is occurring through a functional interaction between PARP-1 and Sp1, resulting in ribosylation of Sp1.

### 6.2.2 PARP-1 ribosylates Sp1.

Ribosylation involves the attachment of negative poly (ADP-ribose) to substrates which in turn confers a negative charge to the acceptor protein and can lead to electrostatic repulsion between DNA and protein. HeLa cells were exposed to 0.25  $\mu$ M-2  $\mu$ M PJ34 for 48 hours and the whole cell extracts were used in both EMSA and immunodetection to examine any changes in the relative levels of known DNA-binding proteins (Sp1, Sp3, NF-Y, Ku86, Ku70 and PARP-1) that may occur due to inhibition of ribosylation.

Figure 6.4A demonstrates that there is an increase in the amount of complex 1 (red arrow, compare lane 1 with lanes 2-5), containing Sp1, suggesting that there is enhanced binding upon treatment with PJ34. The major protein found in complex 1 is Sp1, therefore this increase in band 1 is most likely due to an increase in the amount of Sp1 binding to the oligonucleotide. In order to eliminate the possibility that this increase was due to changes in the amount of Sp1 present in these PJ34 treated extracts, the same HeLa extracts were subjected to immunoblotting analysis. Figure 6.4B shows that PJ34 treatment does not affect the amount or the proportion of each of the proteins that were investigated. Most importantly, there are no fluctuations in the overall amount of Sp1 when PARP-1 activity was inhibited (compare lane 1 with lanes 2-5), suggesting that the increase observed in the EMSA (figure

6.4A) was most likely to be due to an increased affinity for DNA rather than an increase in the amount of protein. Each of the other complexes (2-4) do not appear to vary when HeLa cells are exposed to PJ34 (lanes 2-5, figure 6.4A), suggesting that ribosylation does not have an effect on the other DNA-binding activities examined in this study. These results do however reflect that the DNA-binding activity of neither NF-Y (complex 2, figure 6.4A) nor Sp3 (complex 4, figure 6.4A) are affected by PARP-1 activity.



**Figure 6.4: Sp1 has increased affinity for DNA upon inhibition of ribosylation.**

HeLa cells were exposed to PJ34 for 48 hours and the extracts subjected to EMSA and immunodetection for analysis. (A) 5 μg of whole cell HeLa extract was added to approximately 0.5 ng of <sup>32</sup>P-ICB1wt/GC1wt oligonucleotide in EMSA reactions. The gel was dried onto DE-81 paper and exposed to X-ray film for about 12 hours. The coloured arrows indicate the specific protein-DNA complexes analysed in this study. \* denotes the high mobility protein-DNA complexes thought to be non-specific interactions. (B) 30 μg of whole cell extract was separated by SDS-PAGE using 8% polyacrylamide and transferred onto a positively charged nylon membrane. Primary antibodies αSp1 and αSp3 were diluted 1/500, αKu70, αKu86 and αPARP-1 were diluted 1/450, αNF-Y was diluted 1/300 and tubulin was diluted 1/5000. HRP conjugated 2° antibodies, αrabbit was diluted 1/4000 and αmouse was diluted 1/4500. The difference in intensity observed between tubulin controls is due to differing exposure times. These results are representative of triplicate experiments.

In order for PARP-1 to ribosylate Sp1, the proteins must be in contact with each other. Immunoprecipitation assays were carried out to determine if there was a physical interaction between these two proteins or PARP-1 and any other proteins. Unfortunately, a high degree of non-specific binding of proteins was observed with the agarose A/G beads that were used in these assays. Even though these preliminary results indicated that PARP-1 interacts with Sp1 and not Sp3, the results that were obtained were inconclusive. Even after several attempts and changing many different conditions the immunoprecipitation assays did not give reliable data, therefore they were abandoned. However it has been shown by others that Sp1 and PARP-1 do physically interact. PARP-1 does not have to be enzymatically active to actually interact with Sp1 (Zaniolo *et al.*, 2007). This observation, in conjunction with the data presented in this study raises the possibility that PARP-1 may be recruited to the topoisomerase II $\alpha$  promoter due to an interaction with Sp1. In attempts to determine the degree of ribosylation before and after treatment with PJ34, an antibody specifically against poly (ADP-ribose),  $\alpha$ PAR, was used in antibody supershifts and immunodetection. Unfortunately this  $\alpha$ PAR antibody was found to bind non-specifically to proteins and therefore it was not possible to interpret any of this data with confidence or continue with these experiments either.

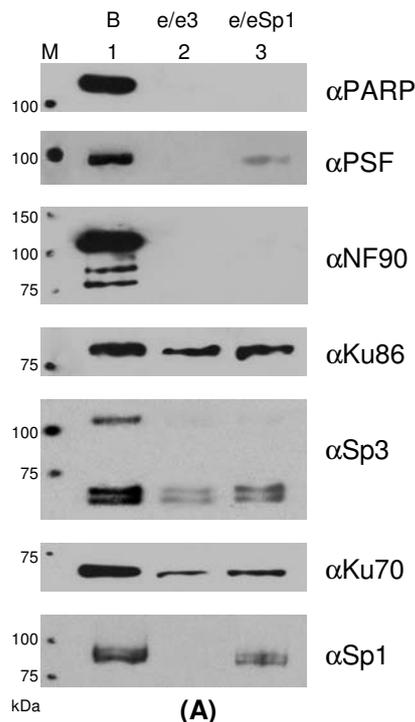
### **6.3 Identifying proteins in electro-eluted extracts.**

DNA-binding proteins were purified and identified on the basis of their ability to bind to the ICB1mt/GC1wt biotinylated-oligonucleotide. Although DNA affinity chromatography can be a robust method used to purify DNA-binding proteins, the proteins Ku86, Ku70, PARP-1, NF90 and PSF identified in this study could not be identified using antibody supershift assays with the topoisomerase II $\alpha$  sequence. As an alternative means of positively identifying these proteins, immunodetection was carried out using proteins that were generated by directly electro-eluting proteins in complex 1 (containing Sp1) and complex three from EMSA gels using the ICB1mt/GC1wt oligonucleotide.

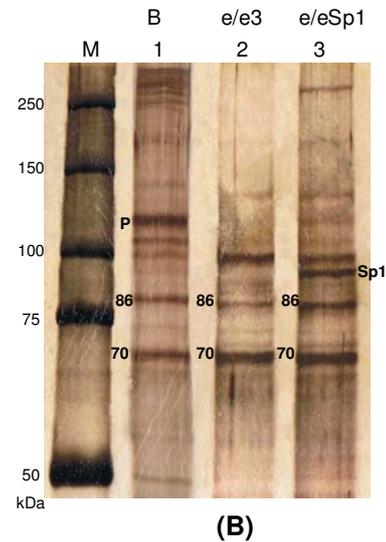
Figure 6.5 shows that each of the antibodies used gave a positive result in lane 1, confirming that PARP-1, PSF, NF90, Ku86, Sp3, Ku70 and Sp1 were all present in extract purified using a biotinylated-ICB1mt/GC1wt oligonucleotide. Sp1 was detected in the sample electro-eluted from the band that represents complex 1 in EMSA, but not in complex 3. This result demonstrated that the electro-elution procedure was valid, as complex 1 has been shown to

be predominantly Sp1 using EMSA antibody supershifts on numerous occasions. However, Sp1 was not identified by antibody supershifts to be present in complex 3 (lane 3 in both figures 3.2A and B). The silver stained SDS-PAGE (Figure 6.5B, lane 3 labelled with Sp1) clearly shows a band approximately 95 kDa in lane 3 (sample eluted from complex 1) but not in lane 2 (sample eluted from complex 3) which corresponds exactly to the same molecular weight for Sp1. Taken together, this strongly suggests that Sp1 is not required for the formation of complex three.

**Immunoblot of proteins in electro-eluted extracts**



**Silver stained SDS-PAGE of purified proteins**



- Lane 1) Extract purified using biotinylated-ICB1mt/GC1wt oligonucleotide (B)
- Lane 2) Electro-eluted proteins from complex three (e/e3)
- Lane 3) Electro-eluted proteins from complex one (e/eSp1)
- M) Protein size standard

**Figure 6.5: Immunodetection of proteins present in electro-eluted extracts.**

(A) 10 µg of protein purified using a biotinylated-ICB1mt/GC1wt oligonucleotide or 5 µg of electro-eluted proteins was separated by SDS-PAGE using an 8% polyacrylamide gel and transferred onto a positively charged nylon membrane. Primary antibodies αSp3 and αSp1 were diluted 1/300. αKu70, αKu86 and αPARP-1 were diluted 1/400. αPSF was diluted 1/500 and NF90 was dilute 1/700. HRP conjugated 2<sup>o</sup> antibodies, αrabbit was diluted 1/3500 and αmouse was diluted 1/4000. Chemiluminescence was visualised by exposure to X-ray film. Exposure times range from ½ second to 1 minute. (B) 2 µg of each protein sample used in immunodetection were visualised by SDS-PAGE and silver staining. The 115 kDa band assumed to be PARP-1 is labelled with a P, Ku86 is labelled with 86, Ku70 with 70 and the 95 kDa Sp1 is also labelled. These results are representative of triplicate experiments.

Sp3 was detected in both of the electro-eluted samples from complex 1 and 3; however there is a decrease in the amount of Sp3 found in the sample from complex 3 (figure 6.5A, lane 2  $\alpha$ Sp3 compared to lane 3  $\alpha$ Sp3). It is possible that overall less Sp3 is present in complex 3 than in complex 1 as the same total amount of protein was used in these immunoreactions (figure 6.5B).

Neither PARP-1 nor NF90 were detected in either of the electro-eluted samples (figure 6.5A, lanes 2 and 3 with  $\alpha$ PARP-1 and  $\alpha$ NF90), indicating that these proteins were not present in either complex 1 or complex 3 from EMSA gels. When examining the silver stained gel (figure 6.5B) there is a protein approximately 115 kDa in size in lane 1 that has been positively identified as PARP-1 (labelled with a P) and this appears to correspond with a faint band of the same size in both electro-eluted samples (lane 2 and 3). Up to this point this 115 kDa band in electro-eluted extracts had been assumed to be PARP-1 based on size and immunoblotting data, but was now confirmed to not be the case. The transient transfection data has indicated that PARP-1 may play a regulatory role in the transcription of topoisomerase II $\alpha$ ; therefore it was surprising to find that PARP-1 was not detected. Even after over-exposure of the immunoblot, up to 30 minutes there was not even a faint band detected for PARP-1 in either of the electro-eluted samples. It is possible that PARP-1 may not be present in high enough concentration in the electro-eluted samples, but this is probably unlikely as Sp3 bands can not clearly be seen in the silver stained gel for the Sp3 proteins, yet immunodetection has detected Sp3. The PARP-1 antibody has worked well in all previous immunoreactions with high sensitivity and specificity. Collectively this leads to the conclusion that PARP-1 does not bind directly to the topoisomerase II $\alpha$  promoter *in vitro*. In this *in vitro* situation the lack of PARP-1 could be explained by assuming that ribosylated Sp1 has lower affinity for DNA, therefore the Sp1 that is purified from EMSA (ie bound to DNA) should not be ribosylated thus will not be associated with PARP-1. This does not exclude the possibility that PARP-1 could be binding to the topoisomerase II $\alpha$  promoter *in vivo* through protein-protein interactions, which may even be transient in nature.

Both Ku86 and 70 are present in both electro-eluted samples (lanes 2 and 3, figure 6.5), indicating that they are present in both complex 1 and complex 3 from EMSA gels. The corresponding bands for these proteins are labelled in the silver stained SDS-PAGE (figure 6.5B labelled 86 and 70), as determined by the molecular weights of these proteins and

overlaying the immunoblot against the stained SDS-PAGE. Within the SDS-PAGE, the band for Ku70 appears to be more intense than the Ku86 band indicating there is possibly more Ku70 in these preparations than Ku86; however the immunoblot does not convey the same result. This is most likely to be due to the 70 kDa band being composed of several proteins of similar molecular weight, when these samples were separated further on a larger SDS-PAGE the 70 kDa band does resolve as several bands (data not shown). The relative proportions of Ku86 and Ku70 appear the same with the extracts prepared using the biotinylated-ICB1mt/GC1wt oligonucleotide (lane 1 immunoblot figure 6.5A) indicating that these Ku proteins may be binding with equal affinity to the oligonucleotide during purification. In general it has been thought that Ku86 and Ku70 form and function as a heterodimers (reviewed in Koike, 2002). This dimerisation is an essential function required for double-strand break repair (DSB repair) where the Ku proteins are in association with the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to form the DNA-PK holoenzyme (Mimori and Hardin, 1986; Jin *et al.*, 1997). In this circumstance where the Ku proteins are apparently involved in gene regulation little is known about sequence specific binding of Ku proteins, even though several lines of evidence support that this does occur (Giffin *et al.*, 1997; Cary *et al.*, 1997; Torrance *et al.*, 1998, Xu *et al.*, 2004).

It has been proposed that sequence independent binding of the Ku proteins to DNA ends may be associated with DSB repair, however sequence-specific binding may be involved in transcriptional regulation (Cary *et al.*, 1997). Emerging evidence suggests that Ku86 and 70 may support multiple functions as monomeric and heterodimeric forms (Wang *et al.*, 1998B; Koike *et al.*, 2001 and reviewed within Koike, 2002) and it is possible that the Ku proteins function as a means of recruiting DNA-PKcs which in turn can phosphorylate transcription factors such as Sp1 (reviewed in Tuteja and Tuteja, 2000). DNA-PKcs is a large protein with a molecular weight of 350 kDa, there is only one distinctive band >250 kDa in the silver stained SDS-PAGE (figure 6.5B and figure 4.8 lane 3) and interestingly it is present in the sample isolated from complex 1 composed mainly of Sp1 which can be phosphorylated to increase its affinity for DNA (Niina *et al.*, 2007).

PSF was only detected in the electro-eluted proteins from complex 1 (lane 3 figure 6.5A  $\alpha$ PSF) albeit at a very low level. Although PSF is reported to be 100 kDa it is unlikely that it is the 100 kDa protein that was isolated from electro-eluted extracts, which can be seen in the silver stained SDS-PAGE (figure 6.5B, lanes 2 and 3, the band just above the labelled Sp1 band). The PSF protein is highly basic and quite often is found to run anomalously within SDS-PAGE (Patton *et al.*, 1993), which may explain why both the 98 kDa and 95 kDa bands analysed by MALDI TOF/TOF resulted with the same protein identity. This unusual nuclear protein has been found to bind to double-stranded and single-stranded DNA in a sequence-independent manner (Akhmedov and Lopez, 2000) and it is thought that more than one PSF can bind to DNA at any one time. However, there have been several reported cases of sequence-specific binding of PSF and in at least two cases the PSF binding site is found to be located 5' to a Sp1-binding GC box within promoters (Urban *et al.*, 2000, Iacobazzi *et al.*, 2005).

Interestingly, within the composite ICB1/GC1 sequence there is a "CTGGTC" sequence just to the right of the ICB1 element (refer to table 3.1 for oligonucleotide sequences). This "CTGGTC" sequence displays high sequence similarity to the reported "CTGAGTC" PSF binding site by Urban *et al.*, 2000 and is also located 5' to the functional GC1 box in the topoisomerase II $\alpha$  promoter. In cross-linking assays (section 5.2.1), it should be noted that the GC1 element to which both Sp1 and Sp3 are known to bind did not contain a BrdU base to aid in cross-linking (discussed in section 5.2.1.2), therefore it may also be possible that the observed 90 kDa protein-DNA complex is PSF and not Sp1. In fact due to the anomalous nature of how PSF protein resolves within SDS-PAGE, it could be attributed to any number of cross-linked bands found within the 90-115 kDa range. However, PSF was not found to be present in complex 3, suggesting that it is not binding to the intervening sequence between both ICB1 and GC1. Alternatively, it may be that PSF is actually present in complex 3, but the weak signal seen in figure 6.5A indicates that PSF is in low abundance in these electro-eluted extracts, and therefore PSF may be below the limit of detection in electro-eluted extract from complex 3.

PSF can inhibit Sp1-driven transcriptional activity (Urban *et al.*, 2000), proposing an explanation for the presence of PSF in extracts electro-eluted from complex 1 (containing Sp1) but not complex three (has no Sp1). While it has not yet been established whether Sp1

and PSF proteins do interact, the fact that both these proteins are found in electro-eluted extracts from complex 1 suggests that they may have a functional interaction. Perhaps PSF is recruited to the ICB1mt/GC1wt sequence due to an interaction with the Ku heterodimer, as these proteins have also been found to cooperate when bound to the same piece of DNA (Bladden *et al.*, 2005). Co-immunoprecipitation experiments were attempted in this study. Unfortunately these could not be continued due to time constraints and experiments were contaminated with high amounts of non-specific protein binding to the agarose A/G beads used in these assays.

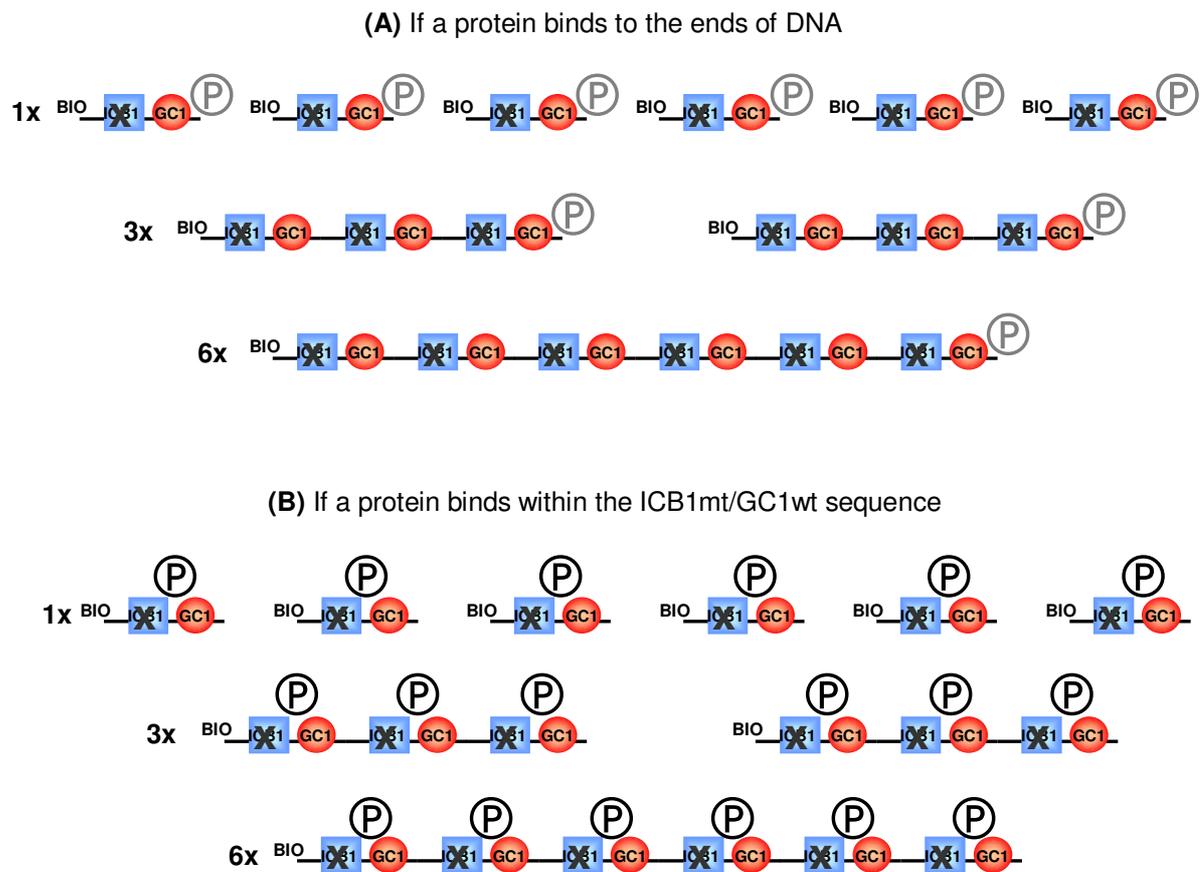
## 6.4 Ku proteins.

There have been reports of Ku contamination amongst DNA affinity chromatography as it is an abundant protein which can bind non-specifically to DNA (reviewed in Tuteja and Tuteja, 2000). It has been suggested that the presence of other contaminating proteins can appear to confer sequence specificity of Ku proteins when purifying through protein-protein interactions (Genersch *et al.*, 1995). Others suggest that it is the nature of the DNA ends that influence Ku binding rather than the internal sequence (Bliss and Lane, 1997).

It is possible that non-specific binding of proteins to DNA-ends may have occurred to a certain degree when purifying proteins using the biotinylated-ICB1mt/GC1wt oligonucleotide, therefore care was taken to reduce the amount of non-specific binding by the addition of random competitors during the purification process. It is recommended that double-stranded competitors are used when performing sequence-specific affinity purification (Kadonaga and Tjian, 1986). Undeniably both Ku86 and Ku70 are present in the extracts obtained by directly eluting proteins from EMSA gels indicating these proteins are components of the *in vitro* complexes with the topoisomerase II $\alpha$  ICB1mt/GC1wt sequence. There is an emerging trend that Ku proteins can actually bind specifically to DNA sequences (Cary *et al.*, 1997; Xu *et al.*, 2004) and thereby may have additional roles to those attributed to the end-binding properties. Therefore it was necessary to establish if both Ku86 and Ku70 were binding to the ends of the ICB1mt/GC1wt oligonucleotide or internally within this sequence.

### 6.4.1 Are the proteins binding to the ends of DNA?

In order to investigate if Ku86 and Ku70 were binding to the ends of DNA or an internal sequence within the ICB1/GC1 composite oligonucleotide a series of DNA-binding assays were carried out. These DNA-termini binding assays utilise the phenomenon that, the number of DNA ends will decrease as the oligonucleotide increases in length, as long as the same amount of DNA is used in each assay. Figure 6.6 better explains the details of these experiments.

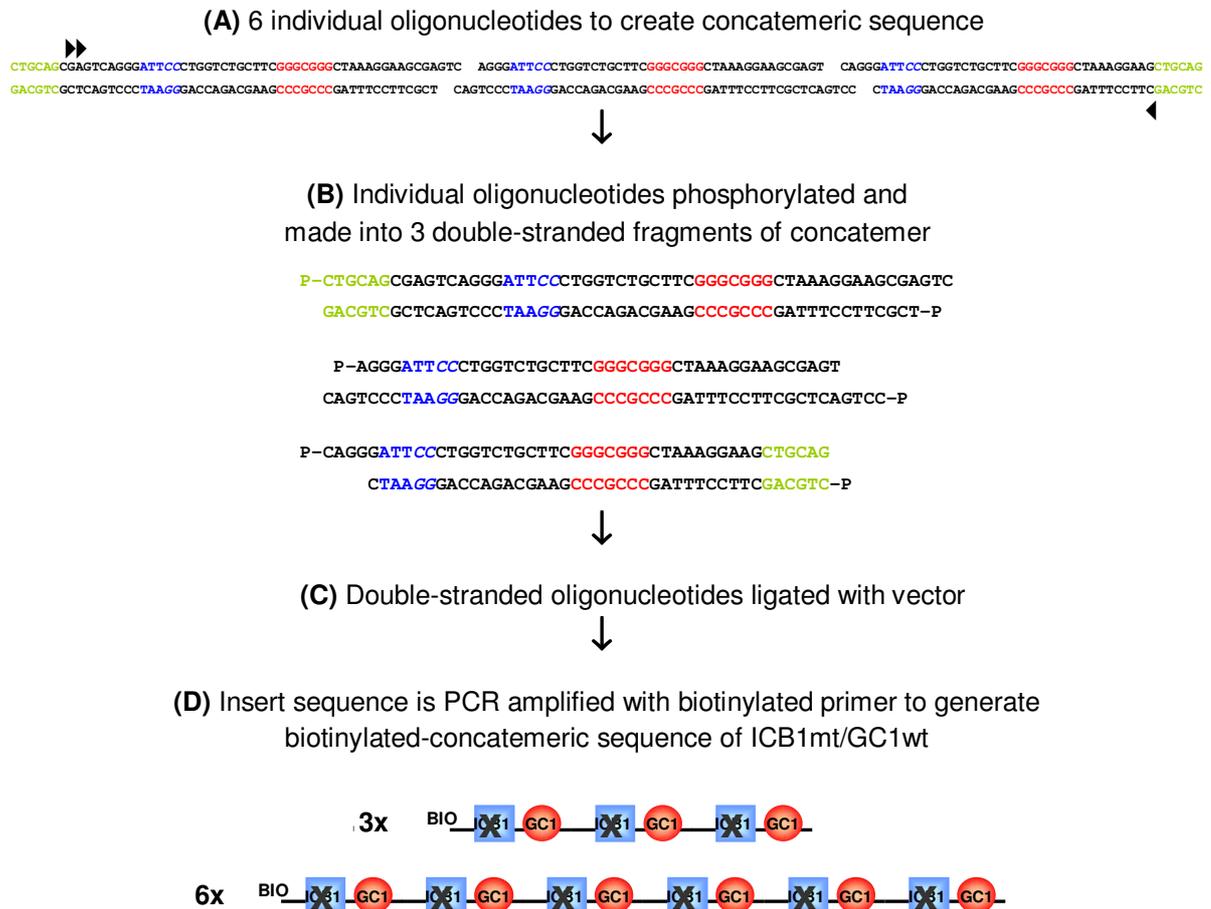


**Figure 6.6: Schematic representation of DNA-termini binding assay.**

Using the same amount of DNA, **(A)** a protein  $\textcircled{P}$  that is found to bind to the ends of DNA will appear to bind less as the length of the oligonucleotide increases and the number of DNA ends decreases. **(B)** Due to the fact that the number of internal ICB1mt/GC1wt sequences will be kept the same, if a protein  $\textcircled{P}$  is found to bind internally within this sequence the amount of protein will remain the same regardless of the length of the DNA used in the assay. 1x, 3x and 6x refers to the number of internal ICB1mt/GC1wt binding sequences.

### 6.4.1.1 Generating a concatemerised ICB1mt/GC1wt oligonucleotide.

In order to carry out these DNA-termini binding assays, a concatemeric sequence of the ICB1mt/GC1wt was required. Figure 6.7 outlines the steps involved in making the biotinylated-ICB1mt/GC1wt concatemeric sequence.



**Figure 6.7: Schematic representation for the generation of the ICB1mt/GC1wt concatemers.**

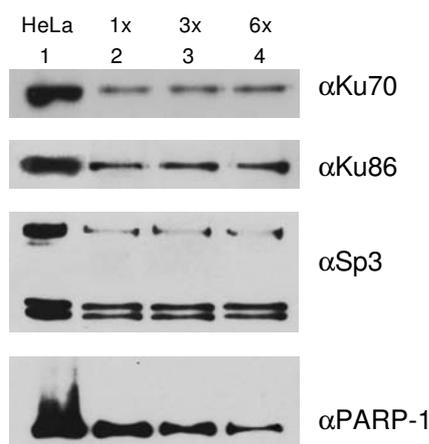
The concatemer was designed to incorporate 3 consecutive ICB1mt (blue)/GC1wt (red) sequences with addition Pst1 restriction sites (shown in green) on either end of the sequence. This oligonucleotide was too long to manufacture, thus (A) the longer sequence was constructed using 6 individual oligonucleotides shown by the spaces in the concatemer sequence. (B) The 6 individual oligonucleotides were phosphorylated and annealed to its complimentary strand to create three overlapping double-stranded fragments. (C) Each double-stranded fragment was added to a ligation with pBluescript SK that had been prepared by digestion with Pst1. XL-1 transformants were selected and plasmids with the correct insert(s) were confirmed by digestion with Pst 1 and sequence analysis. Many plasmids contained a single concatemer, consisting of three consecutive ICB1mt/GC1wt sequences however some were found to contain double concatemeric sequences consisting of six consecutive ICB1mt/GC1wt sequences. (D) To create the biotinylated-concatemers, the appropriate recombinant plasmids were PCR amplified using a ▶▶biotinylated primer and a standard primer ◀ designed to the ICB1mt/GC1wt sequence.

Several attempts were made to individually ligate the three double-stranded oligonucleotides (B, figure 6.7) together, but they were unsuccessful. Therefore a ligation reaction containing all three of the individual double-stranded pieces of the concatemer along with the prepared plasmid (pBluescript SK+/- vector digested with Pst1) was used. This resulted in a small number of transformants, and the identity of each was confirmed by automated sequencing using the standard primers T7 and T3. The plasmids with the insert in the desired orientation were propagated and used as templates in PCR. Primers were designed to the ICB1mt/GC1wt sequence to amplify the concatemeric sequence from the appropriate recombinant plasmid. Using one primer that was biotinylated enabled these PCR amplified concatemers to be used in subsequent purification and binding assays with the magnetic streptavidin beads. The correct biotinylated-PCR products were selected by separation with agarose gel electrophoresis and gel purification.

#### **6.4.1.2 DNA-termini binding assays.**

The 3x and 6x concatemeric ICB1mt/GC1wt biotinylated-oligonucleotides were used in conjunction with the single (1x) ICB1mt/GC1wt biotinylated-oligonucleotide (used in oligonucleotide-affinity chromatography) to compare the DNA end binding patterns of Ku86, Ku70, Sp3 and PARP-1. The success of this assay was highly dependent on maintaining an equal amount of biotinylated-oligonucleotide, protein and streptavidin beads in each of the binding reactions. To ensure that the same number of internal binding sites was kept consistent during these experiments, GC box binding protein Sp3 was used as an internal control. To ensure that the same amount of DNA was bound to the streptavidin beads during the assay, each oligonucleotide was quantified and the amount of oligonucleotide that bound to the streptavidin beads was calculated by monitoring the amount of oligonucleotide left unbound to the beads after each binding reaction. To maintain the same binding conditions for each of the 1x, 3x and 6x reactions, these experiments were all carried out concurrently. Binding reactions were carried out as described in section 2.2.22, with equal amounts of non-specific competitors added to equal amounts of HeLa nuclear extract. Each of the extracts purified using the ICB1mt/GC1wt 1x, 3x and 6x biotinylated-oligonucleotide were then separated by SDS-PAGE and proteins detected using specific antibodies and immunoblotting.

### Immunoblot to determine which proteins are binding to DNA termini



- Lane 1) 30  $\mu$ g HeLa nuclear extract
- Lane 2) 15  $\mu$ g extract purified using (1x) ICB1mt/GC1wt-biotinylated oligonucleotide
- Lane 3) 15  $\mu$ g extract purified using (3x) ICB1mt/GC1wt-biotinylated oligonucleotide
- Lane 4) 15  $\mu$ g extract purified using (6x) ICB1mt/GC1wt-biotinylated oligonucleotide

**Figure 6.8: Immunodetection DNA-termini binding assay.**

DNA binding proteins were purified using 1x, 3x or 6x ICB1mt/GC1wt biotinylated-oligonucleotide to determine which of the proteins may be binding to DNA-termini. Proteins were denatured then separated by SDS-PAGE using an 8% polyacrylamide gel and transferred onto a positively charged nylon membrane. Primary antibodies were diluted as follows,  $\alpha$ Sp3 was diluted 1/300 and  $\alpha$ Ku70,  $\alpha$ Ku86 and  $\alpha$ PARP-1 were diluted 1/450. HRP conjugated 2<sup>o</sup> antibodies,  $\alpha$ rabbit was diluted 1/3500 and  $\alpha$ mouse was diluted 1/4000. Chemiluminescence was visualised by exposure to X-ray film. Exposure times range from 1/2-30 seconds. These results are representative of triplicate experiments.

The results for a typical DNA-termini binding assay are shown in figure 6.8. An equal amount of Sp3 was detected in each of the samples purified using a biotinylated-oligonucleotide (lanes 2-4), regardless of the length of the oligonucleotide. This is a positive result that indicates that there were no fluctuations in the number of internal binding sites and it was most likely that the total amount of DNA used in each assay was consistent. The same amounts of both Ku86 and Ku70 were also detected with each of the biotinylated-oligonucleotides (lanes 2-4  $\alpha$ Ku70 and  $\alpha$ Ku86) suggesting that these proteins may preferentially bind to a sequence within composite ICB1mt/GC1wt oligonucleotide rather than at the DNA-termini. It was interesting to observe that there is a marked decrease in the amount of PARP-1 that was found to bind to the different oligonucleotides. The amount of PARP-1 that was purified decreases as the oligonucleotide increases in length suggesting that the majority of this protein could be binding to DNA-termini. In addition to this there is a

greater amount of PARP-1 purified in comparison to any of the other proteins detected by antibody in this assay, this was a common occurrence observed in many of these oligonucleotide-affinity purification procedures. Even though a majority of PARP-1 is shown to bind to DNA-termini, the possibility that it may also bind internally through protein-protein interactions with Sp1 or other proteins that may be targeted for ribosylation can not be excluded. It is also possible these interactions may be transient *in vivo* and would not necessarily be detected *in vitro*.

## 6.5 Chapter summary.

The transient transfection data clearly demonstrates that PARP-1 is likely to have an inhibitory role on topoisomerase II $\alpha$  promoter activity and this may be directly related to the binding of the transcriptional activator Sp1. When ribosylation was inhibited there was an increase in the amount of Sp1 that bound to the ICB1/GC1 composite oligonucleotide, demonstrating that ribosylation may inhibit the transcriptional activation of topoisomerase II $\alpha$  via ribosylation of Sp1. The fact that PARP-1 is not found in electro-eluted extracts suggests that PARP-1 itself is not bound to the topoisomerase II $\alpha$  promoter *in vitro* or perhaps that the PARP-1-Sp1 interaction is transient in nature.

The proteins that were found to be present in complex one from EMSA reactions were Sp1, Sp3, PSF, Ku86 and Ku70. The proteins found to be present in complex three from EMSA reactions were Sp3, Ku86 and Ku70. Even though the Ku proteins are avid binders of DNA-termini, both Ku86 and Ku70 were shown to bind to internal sequences in this study. This is in agreement with reports that suggest that direct sequence-specific binding of Ku proteins can occur with higher affinity than DNA end binding (Giffin *et al.*, 1997). Immunodetection of the electro-eluted extracts eliminated NF90 and PARP-1 as being proteins that are associated with the composite ICB1/GC1 sequence *in vitro*. However, Ku86, Ku70 and PSF were found as novel components binding to the ICB1wt/GC1wt sequence. PSF was found to bind only in conjunction with Sp1, indicating there is possibly a functional interaction between the two proteins. Unfortunately it was not possible to assign a function for PSF, Ku86 and Ku70 in the context of gene regulation of the topoisomerase II $\alpha$  promoter without further analysis.

# 7 Discussion and future work.

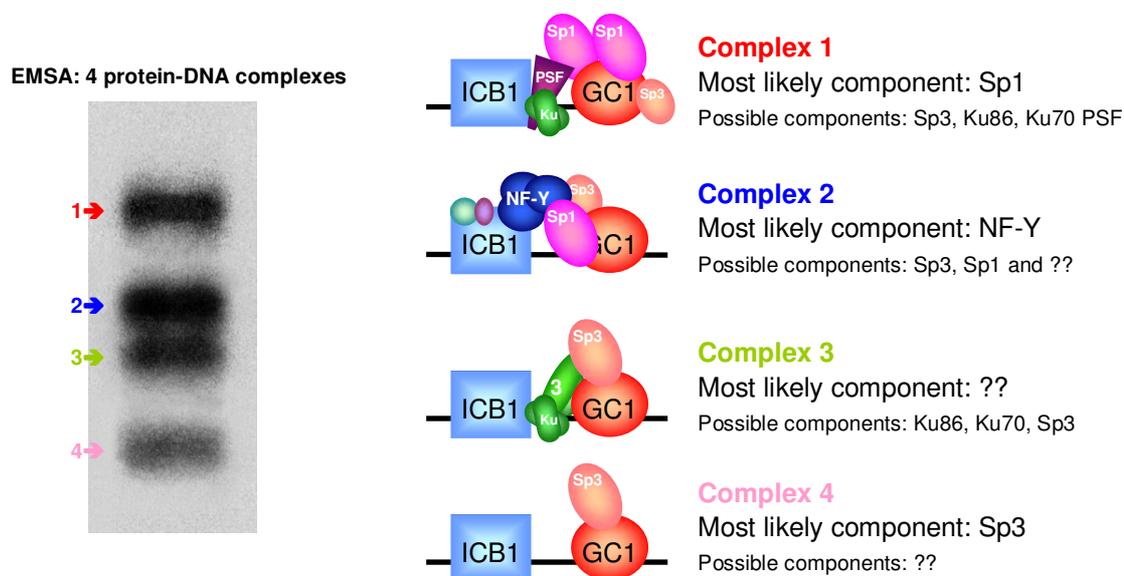
## 7.1 Synopsis.

The regulation of the topoisomerase II $\alpha$  gene is a fundamental part of understanding the mechanisms involved in topoisomerase II $\alpha$  protein production. Topoisomerase II $\alpha$  is currently undergoing extensive investigation for use as a biomarker to predict treatment therapy for patients diagnosed with breast cancer. It has been shown that the efficacy of anthracycline as an anti-tumour agent corresponds with relative levels of topoisomerase II $\alpha$  protein in tumours (Withoff *et al.*, 1996; Jarvinen *et al.*, 2000; Coon *et al.*, 2002). Although it has been suggested that overexpression of the topoisomerase II $\alpha$  protein is related to topoisomerase II $\alpha$  gene amplification (Bhargava *et al.*, 2005), this can not be the only mechanism that regulates the amount of topoisomerase II $\alpha$  protein in tumours and drug-resistant cell lines as there is discordance between topoisomerase II $\alpha$  gene amplification and protein expression (Mueller *et al.*, 2004; Durbecq *et al.*, 2004A). Changes in the levels of transcription factors have also been correlated with concomitant changes in topoisomerase II $\alpha$  expression (Allen *et al.*, 2004) and the development of drug-resistance (Kubo *et al.*, 1995; Wang *et al.*, 1997B). The transcription factors Sp1, Sp3 and NF-Y have been shown to have a role in regulating expression of topoisomerase II $\alpha$  due to interactions between the proximal GC elements and inverted CCAAT boxes respectively (Magan *et al.*, 2003). It is likely however that these do not act alone at the promoter. This study attempted to further define the DNA-binding proteins that may be integral for the regulation of the topoisomerase II $\alpha$  gene. Using a composite ICB1/GC1 topoisomerase II $\alpha$  oligonucleotide in EMSA assays it was evident that multiple protein-DNA interactions occur *in vitro* at the respective elements as well as within the intervening sequence between these elements. The intervening sequence was found to contain an additional regulatory element that appears to enhance topoisomerase II $\alpha$  promoter activity when binding of the respective proteins was enhanced. Oligonucleotide-affinity chromatography was used in an attempt to purify these DNA-binding proteins and identify components of the potential protein-protein and protein-DNA interactions that may be important for transcriptional regulation of the topoisomerase II $\alpha$  gene. Not all of the proteins implicated in these interactions were able to be conclusively determined in the course of this study; however the proteins Ku86, Ku70, PARP-1 and PSF were identified as being able to bind to the topoisomerase II $\alpha$  promoter with potential

functional consequences. This is the first report to demonstrate that the ribosylation activity of PARP-1 has the ability to inhibit topoisomerase II $\alpha$  promoter activity through a mechanism that involves the binding of Sp1 to the GC1 element within the promoter.

## 7.2 Protein-ICB1/GC1 interactions.

It has previously been established that the region encompassing both the ICB1 and GC1 elements from the topoisomerase II $\alpha$  promoter is an integral regulatory region (Magan *et al.*, 2003). Using a composite oligonucleotide representing the ICB1/GC1 region of the topoisomerase II $\alpha$  promoter, EMSA reactions described in the current study have consistently displayed the presence of multiple protein-DNA complexes (summarised in figure 7.1).



**Figure 7.1: Summary of the different DNA-binding proteins found in each of the protein-DNA complexes that are seen to form using the topoisomerase II $\alpha$  ICB1wt/GC1wt oligonucleotide.**

On the left is the top half of an EMSA gel that illustrates each of the four characteristic protein-DNA complexes that are observed in the presence of the ICB1wt/GC1wt topoisomerase II $\alpha$  oligonucleotide. In each case, the schematic representations (right) of the proteins thought to be involved in the formation of these complexes were generated using data collected from EMSA, antibody supershifts, directly eluting proteins from EMSA gels, cross-linking assays and current literature on protein-protein interactions (Roder *et al* 1999; Urban *et al.*, 2000; Davie *et al.*, 2008). Ku proteins (Ku) were found to crosslink within the intervening sequence. There is a putative PSF consensus sequence just to the right of the ICB1 element and PSF may also cross-link to the intervening sequence. PSF is most likely present due to an interaction with Sp1. Components of complex 3 (3) found to bind within the intervening sequence remain unknown.

There are three established protein-DNA interactions involving transcription factors Sp1 (complex 1, red arrow, band 1), NF-Y (complex 2, blue arrow, band 2) and Sp3 (complex 4, pink arrow, band 4), as well as an additional protein (complex 3, green arrow, band 3). Complex 1, 2 and 4 (bands 1, 2 and 4 respectively) have been extensively studied (throughout this study and a previous study) using antibodies against their respective proteins in antibody supershift assays; but it was observed that a complete shift in the respective bands was not always detected. The EMSA data obtained when using combinations of antibodies in assays (refer to figure 3.2) demonstrated that it is possible that multiple proteins could be associating on the same fragment of DNA, perhaps due to protein-protein interactions. However, it should be also noted that an incomplete mobility shift in a single band within an antibody supershift gels may also be due to two or more different protein-DNA complexes that migrate to same (or very similar) location within an EMSA gel, thus providing an additional explanation for an incomplete mobility shift when respective antibodies were used. Therefore the proteins associated within each protein-DNA complex are referred to as major (most likely, as identified by antibody supershift assays) or minor (possible) components and are discussed further throughout this chapter.

### **7.2.1 Complex 1/Sp1 interacting with the GC1 element.**

Using antibodies against Sp1, it was shown that Sp1 is a major component of complex 1 in EMSA gels (refer to figure 3.2). This transcription factor confers its activation by binding to both the GC1 and GC2 elements (Magan *et al.*, 2003; Williams *et al.*, 2007) and enhanced binding of Sp1 has been shown to further promote Sp1-mediated gene activation of the topoisomerase II $\alpha$  promoter (Allen *et al.*, 2004). Sp1 bound at different locations within a promoter are thought to be able to interact thereby bringing two regulatory elements close together to direct synergistic transactivation. This type of synergistic activation through Sp1 has been reported to occur at a range of promoters (Courey *et al.*, 1989; Pascal and Tjian, 1991; Su *et al.* 1991) including the topoisomerase II $\alpha$  promoter (Williams *et al.*, 2007) and is based on the ability of Sp1 molecules to form homo-oligomers (Mastrangelo *et al.*, 1991). When Sp1 is organised as a multimer, it also represents multiple docking sites for interacting proteins. The proteins that were electro-eluted from complex 1 from EMSA gels, indeed show that a number of different proteins are found in conjunction with Sp1 in this protein-DNA complex (refer to figure 4.8, lane 3). Proteins Sp3, Ku86, Ku70 and PSF were positively identified (refer to figure 6.5, lanes 3) and within this population of proteins from

complex 1, there were 7-9 major proteins (stained with higher intensity) as well as 10-15 other minor proteins that (faintly silver stained) were present. This is not surprising as Sp1 has been shown to interact with a vast range of proteins (Ge *et al.*, 2001; Yamada *et al.*, 2000; Santini *et al.*, 2001) and it is possible that some of these proteins may have been isolated due to protein-protein interactions rather than bonafide protein-DNA interactions. It has also been suggested that Sp1 may require mediator proteins to bind to templates *in vitro* (Wierstra, 2008). The exact roles for Ku86, Ku70 and PSF in the transcriptional regulation of topoisomerase II $\alpha$  are still undefined however putative functions will be discussed subsequently throughout this chapter. The presence of Sp3 in combination with Sp1 is not surprising, as both these proteins can bind to the GC1 element and they are thought to be able to interact with each other. Alternatively, it has also been suggested that a mediator protein could also be required to enable Sp1 and Sp3 to interact as opposed to a direct protein-protein interaction (Galvagni *et al.*, 2001). It should be possible to determine if these proteins (Sp1, Sp3, Ku86, Ku70 and PSF) could directly interact with the DNA or if binding required co-factors or mediator proteins by using EMSA with purified versions of these proteins. Proteins could be added either individually or in different combinations sequentially to determine which were required for binding to the oligonucleotide *in vitro*.

### 7.2.2 Complex 2/NF-Y interacting with the ICB1 element.

It is well documented that the ICB boxes of topoisomerase II $\alpha$  promoter are required for optimal transcriptional activity (Hochhauser *et al.*, 1992; Isaacs *et al.*, 1996; Wang *et al.*, 1997B). In some regards it is the ICB1 element that has been shown to be the most important between the five CCAAT boxes, as a mutation in this element reduces promoter activity by approximately 70% (summarised in figure 7.2, mtICB1 -617topo2 $\alpha$  construct). Previous work has shown that NF-Y has the ability to activate topoisomerase II $\alpha$  gene expression by binding to the ICB1 element (Magan *et al.*, 2003) as well as conveying the same effect when binding to ICB3 and to a lesser extent ICB2 (Joshi *et al.*, 2003). In contrast, NF-Y has also been shown to be a repressor at ICB2 during confluence arrest (Isaacs *et al.*, 1996). Interestingly, ICB2 is also able to bind other factors, such as ICBP90 which appears to stimulate topoisomerase II $\alpha$  promoter activity (Hopfner *et al.*, 2000). Therefore, it is not unusual to find variations in the functions of transcription factors based on the context of the DNA-binding sequences, the cell line from which they were derived from, and the proteins they are found to interact with. In this study it was shown that the GC1 element can compete

for binding of NF-Y to the ICB1wt/CENins/GC1mt oligonucleotide (figure 3.7, lanes 17-20) and in previous work it was shown the ICB1 element can also bind Sp1 (Magan *et al.*, 2003), indicating that the proteins that associate with GC1 may be able to interact with the proteins that bind to ICB1. NF-Y has also been shown to be able to interact with Sp1 (Roder *et al.*, 1999; Yamada *et al.*, 2000; Ge *et al.*, 2001) and NF-Y mediated transcriptional regulation is thought to be enhanced when Sp1 is bound nearby (Wright *et al.*, 1995). This is an indication of the multi-protein complexes required to regulate transcription. In this study, the ICB1 element appears to bind predominantly NF-Y, as an almost complete shift in the bands that relate to complex 2 can be seen when antibodies against NF-YA are added to EMSA reactions (refer to figure 3.2B). Antibody supershift assays that were carried out in this study have shown that ICBP90 is not capable of binding to ICB1 (data not shown), however it may be possible that other proteins may be recruited due to an interaction with NF-Y, such as Sp3 or Sp1 and possibly other CCAAT box binding proteins. The proteins that relate to complex 2 were not electro-eluted from EMSA gels; in most cases the oligonucleotide used in this study was the ICB1mt/GC1wt variation as NF-Y was not required for the binding of proteins in complex 3. Therefore, it might be interesting to attempt this in future studies as the electro-elution method was relatively successful in being able to visualise the components of these types of protein-DNA complexes.

### **7.2.3 Complex 4/Sp3 interacting with the GC1 element.**

Sp3 is known to interact with the GC1 element within the topoisomerase II $\alpha$  promoter. Complex 4 (Figure 7.1) was shown to be mainly composed of transcription factor Sp3 as antibody supershift assays demonstrated there was virtually a complete shift with antibody against Sp3 in band 4 (refer to figure 3.2A and B, lanes 4). When comparing several EMSA gels (figures 3.1, 4.2, 4.3B, 7.1) often the band that relates to complex 4 appears with less intensity than any of the other protein-DNA complexes that form. This suggests that in the total population of all the observed protein-DNA complexes, complex 4 is present in the lowest abundance. It has been shown that the breast cancer cell line T5 has 3-fold more Sp1 protein than Sp3 (Sun *et al.*, 2002) and this was certainly observed using whole cell extracts from HeLa cells in this study, therefore this difference in apparent band intensity may be representative of the relative amounts of Sp3 found in the extracts. Alternatively, it may suggest that the protein-DNA conformations found in the other complexes (other than complex 4) are more favourable *in vitro*. It has been shown that Sp1 has the ability to bind as

a multimer on an individual Sp1-binding site (Mastrangelo *et al.*, 1991) however Sp3 has yet to be shown to be able to form multimers. It is thought that Sp3 is only capable of binding as a monomer and is therefore not capable of synergistic transactivation (Yu *et al.*, 2003; Lomberk and Urrutia, 2005). Complex 4 has the highest mobility of the four standard protein-DNA complexes that form in EMSA reactions, indicating that it is most likely the smallest, most mobile, least complicated mixture of proteins and DNA. Taken together this suggests that complex 4 is most likely composed of a single Sp3-DNA interaction and it is possibly the least favoured protein-ICB1wt/GC1wt conformation *in vitro*. The proteins from this complex were not electro-eluted from EMSA gels for this study, therefore it is not known if any other proteins were also involved in the binding reaction nor was it able to be established which of the Sp3 isoforms were present in the complex. Sp3 has 3 or 4 different isoforms; the two shorter isoforms do not arise from alternative splicing of Sp3 RNA, but are products of differential translational initiation (Sapetschnig *et al.*, 2004). So a construct that enables expression of the longer isoform will yield the two shorter ones as well. Many lines of research do not acknowledge the possibility of differential regulation due to the presence of these different isoforms. Each Sp3 isoform is capable of binding to GC boxes, the larger isoforms are thought to both activate and repress transcription, however the two shorter isoforms are thought only to repress transcription (Kennett *et al.*, 1997, Kennett *et al.*, 2002; Sapetschnig *et al.*, 2004). Repression of transcription via Sp3 is thought to occur through an inherent ability to repress Sp1-mediated activation (Reviewed in Li *et al.*, 2004). Therefore the cellular ratios of Sp1:Sp3 are important and they have been shown to change in different contexts (reviewed in Li *et al.*, 2004). It has been shown that Sp3 can repress Sp1-mediated activation in a dose-dependent manner at the topoisomerase II $\alpha$  promoter (Williams *et al.*, 2007). However, it is not known whether this occurs through Sp3 competing for DNA binding sites (Kennett *et al.*, 2002; Yu *et al.*, 2003) or through a protein-mediated Sp1/Sp3 complex (De Luca *et al.*, 1996). It is not known how Sp3 isoform expression is regulated *in vivo*, nor under which physiological conditions the Sp3 isoform ratio alterations take place. These would be interesting areas of research to follow up especially in the context of the various cancerous cell lines and examining changes due to drug treatment and the acquisition of drug-resistance. It is noteworthy at this point to mention, that in this study there appears to be more of the 117 kDa isoform of Sp3 in whole cell and HeLa nuclear extracts (figure 6.4B) than compared to that found in the purified extracts (figure 4.6C, 4.9C, 6.5A), where there appears to be an abundance of the shorter Sp3 isoforms. The ratios of the Sp3 isoforms do

not appear in equal quantities and may be important in understanding the exact role of Sp3 in topoisomerase II $\alpha$  gene regulation. Therefore it may be of significance to determine which isoform of Sp3 is preferentially bound to the promoter. This could be accomplished by immunoblotting after electro-elution from EMSA gels, and further enhanced by using isoform specific antibodies in antibody supershift assays and immunoblots.

#### **7.2.4 Complex 3/proteins and the intervening sequence.**

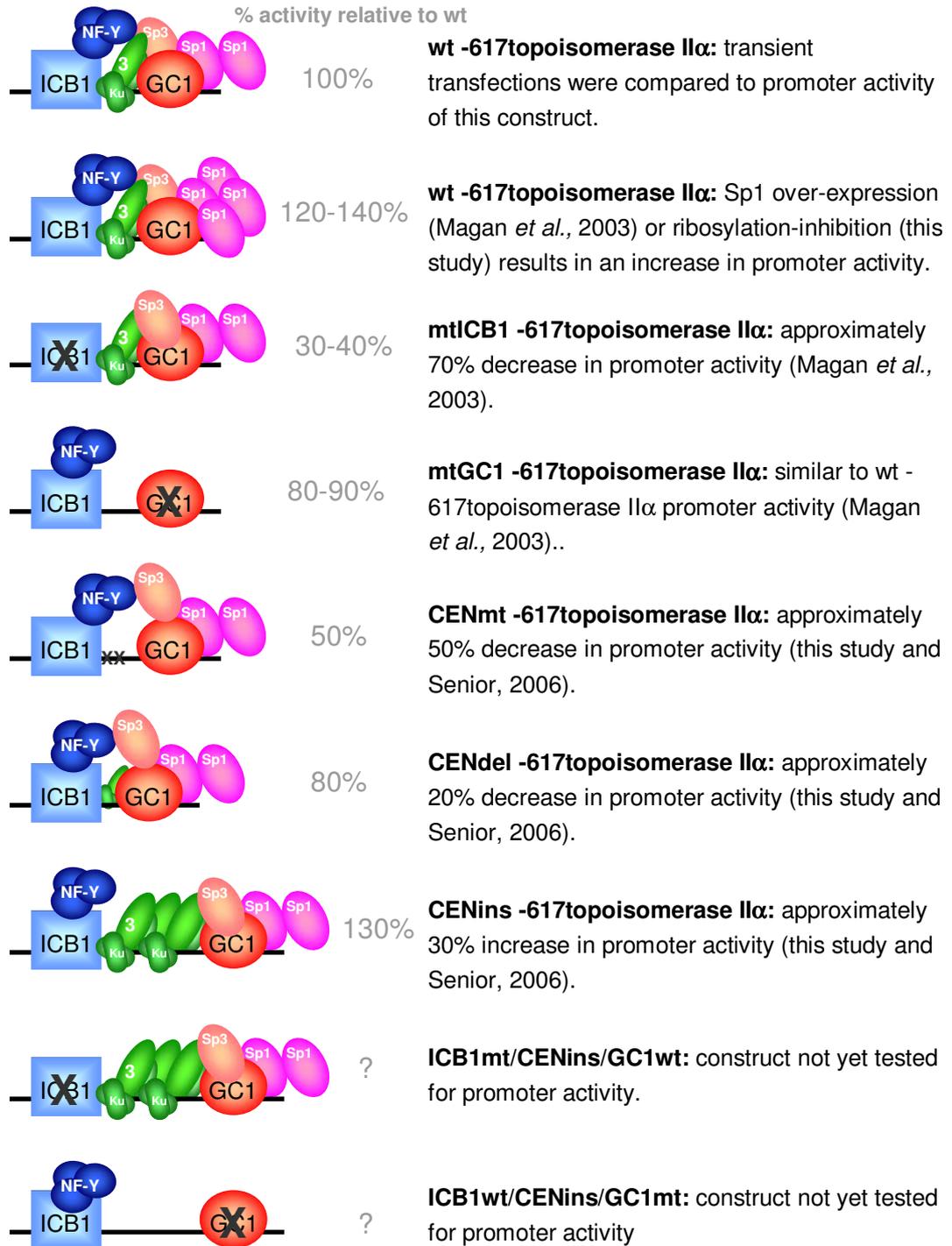
EMSA assays using the ICB1/GC1 oligonucleotide demonstrated the possibility of 4 protein-DNA complexes forming on this oligonucleotide, three of which were identified using antibodies to Sp1, Sp3 and NF-Y. Competitor assays highlighted that additional proteins may bind to the composite ICB1/GC1 oligonucleotide, other than those already known to associate with the ICB1 and GC1 elements. These observations lead to an assumption that the intervening sequence between both ICB1 and GC1 may recruit proteins required for the regulation of topoisomerase II $\alpha$  (complex 3). There appears to be much higher affinity for complex 3 to assemble on the composite ICB1mt/GC1wt oligonucleotide than the shorter GC1 sequence alone. In EMSA assays, the GC1 competitor was shown to reduce binding of these proteins in complex 3 (as discussed in section 3.2.2.3, figure 3.7, lanes 9-16) but not completely reduce protein binding to the same affinity as that seen with the GC1 related transcription factors Sp1 and Sp3. With the observed dependence on the presence of the GC1 element, it could be presumed that complex 3 is composed of Sp1 or Sp3 however when antibodies against Sp1 were used in antibody supershift assays a distinct shift was not observed for complex 3. This suggests that complex 3 does not contain Sp1. Using antibodies against Sp3 in antibody supershift assays a very slight shift in complex 3 was observed (figure 3.2B lane 4), providing preliminary evidence that Sp3 may be a component of complex 3. Proteins were electro-eluted from complex 3 in EMSA gels and it was established that Sp3, Ku70 and Ku86 were among some of the proteins in this complex (figure 6.5A, lane 2). There were many other proteins present in these extracts (figure 6.5B, lane 2) and not all of them were able to be identified through the course of this study. It was interesting to find that many of the proteins present in complex 3 were of the same apparent molecular weight as those proteins electro-eluted from complex 1 (compare lanes 1 and 3 figure 4.8; compare lanes 2 and 3, figure 6.5B), displaying the overall similarities in the two complexes, which would be expected as they do bind to the same oligonucleotide. One significant difference was that complex 3 was missing a protein approximately 95 kDa in size which

was present in the electro-eluted proteins from complex 1 (figure 6.5B compares lanes 2 and 3). This 95 kDa band was rationalised to be Sp1 due to its size and the fact that Sp1 was not detected in protein electro-eluted from complex 3 (figure 6.5A compares lanes 2 and 3). The exclusion of Sp1 from complex 3 may contribute to the difference in mobility that is observed between complex 1 and complex 3 in EMSA. This suggests that complex 3 may not be composed of an additional unidentified protein, but instead is an artefact produced by the lack of a protein component. However, cross-linking assays using BrdU incorporated into the intervening sequence confirmed that there are a number of proteins that could possibly bind to this intervening sequence (figures 5.3 and 5.4), ranging from 70-200 kDa in size, suggesting that another protein(s) may be major component of complex 3.

Complex 3 appears to form within the intervening sequence between the ICB1 and GC1 elements (figure 7.1), however the intervening sequence alone is not sufficient to enable binding of any proteins (Senior K., MSc Thesis, 2006). Although the exact complex 3 binding site was not mapped, it was shown that the “GTC TGC” sequence is important to recruit proteins and binding of these proteins is not a simple spatial requirement within the intervening sequence (figure 3.5, compare binding of ICB1wt/repeatseq/GC1wt and ICB1wt/CENins/GC1wt, lanes 2 and 3 respectively). In addition, there appears to be a strict requirement for an intact GC1 element to enhance binding (figure 3.6, lanes 14 and 15), but the ICB1 element is not required (figure 3.6, lanes 11 and 12). This implies that the proteins that associate with the GC1 element may be important to enhance the DNA-binding properties of the proteins in complex 3. This is substantiated when the following is also considered. EMSA assays (in the context of the topoisomerase II $\alpha$  promoter) have shown that Sp3 is capable of associating with a single GC2 element but not a single GC1 element; however Sp3 can interact with GC1 when it is present in the longer ICB1wt/GC1wt composite oligonucleotide (Magan *et al.*, 2003). This represents a possibility that the longer ICB1wt/GC1wt oligonucleotide may contain regulatory sequence(s) which enables the recruitment of other protein(s) that allow Sp3 to bind to the GC1 element. This is very significant as Sp3 was found to be a component of complex 3, although only the shorter two Sp3 isoforms were detected in these electro-eluted extracts (figures 4.9C and 6.5A lanes 2). As the shorter Sp3 isoforms are thought to represent isoforms of Sp3 that have a repressive role, it is possible that complex 3 represents a protein complex that forms to counter balance any Sp3-mediated repression of the topoisomerase II $\alpha$  gene in the absence of Sp1. Or

conversely, Sp3 may act to repress any aberrant complex 3-mediated promoter activation. Transient transfection data has demonstrated that a construct designed to enhance for the ability to form complex 3 (CENins, figure 3.10) resulted in enhanced topoisomerase II $\alpha$  promoter activity (figure 7.2, CENins -617topo2 $\alpha$  construct). In addition, constructs designed to produce a reduction in complex 3 resulted in diminished topoisomerase II $\alpha$  promoter activity (figure 7.2, CENmt -617topo2 $\alpha$  construct and CENdel-617topo2 $\alpha$  construct). Taken together these results suggest that the proteins which bind to the intervening sequence may be important in the regulation of the topoisomerase II $\alpha$  promoter and that the sequence between ICB1 and GC1 may contain an additional regulatory sequence required to recruit transcriptional regulators to the topoisomerase II $\alpha$  promoter.

Using the composite ICB1/GC1 element from the topoisomerase II $\alpha$  promoter, the proteins that associate to form complex 3 appear to be present in MCF7, MCF12A and HeLa nuclear extracts (figures 3.3, 3.4 and 3.1, respectively) demonstrating that these proteins are potentially important in different cellular contexts. However no significant shift in mobility in any of the bands in these EMSA reactions were observed using antibodies against Ku86, Ku70, PARP-1, NF90, ICBP90 and PSF (data not shown). It is not known why antibody supershift assays did not positively identify Ku86 and Ku70 in complex 3 (or complex 1); it is possible that the presence of multiple protein-protein interactions blocked the antigenic sites for the antibodies in EMSA reactions. Both Ku86 and Ku70 were in high enough abundance in the electro-eluted extracts to be detected with antibodies in immunoblots therefore the antibodies themselves were not at fault. Both Ku86 and Ku70 are thought to be highly abundant proteins in the cell, therefore it may be possible that a huge excess of antibody was required to detect a significant mobility shift in antibody supershift assays. The amount of antibody was titrated against the amount of extract in antibody supershift assays and no significant differences were observed (data not shown). Alternatively it may be useful to use an antibody raised specifically against the Ku86/70 heterodimer in the future.



**Figure 7.2: Schematic representation of the DNA-binding proteins observed using EMSA and the corresponding promoter activities observed using transient transfections.**

The results of different EMSA and transient transfections (from various sources) are summarised in the figure above. **(Left)** The schematic representations of the putative protein-DNA and protein-protein interactions that were observed using different oligonucleotides tested in EMSA. **(Right)** The relative topoisomerase II $\alpha$  promoter activities when proteins are thought to bind in the corresponding configurations. The % promoter activities are displayed relative to the wt -617topo2 $\alpha$  reporter gene construct. PSF was omitted due to insufficient evidence to substantiate a strong interaction with the topoisomerase II $\alpha$  promoter.

### 7.3 Ku Proteins.

The Ku proteins have been implicated in a variety of cellular processes including DNA double-stranded break (DSB) repair, immunoglobulin V(D)J recombination, DNA replication, maintenance of telomere length, progression of cell cycle and regulation of transcription (reviewed in Tuteja and Tuteja, 2000). Traditionally, both Ku86 and Ku70 are thought to have a role in DSB repair. Ku is known to recruit the catalytic subunits (DNA-PKcs) of the DNA-dependent protein kinase (DNA-PK) holoenzyme which initiates a phosphorylation and protein-protein interaction cascade that in turn leads to recruitment of repair enzymes. Dimerisation and binding of the Ku proteins to broken DNA ends is thought to be required to prevent unnecessary DNA degradation and juxtapose the DNA ends. Both Ku86 and Ku70 were found to be novel components that bind *in vitro* to the intervening sequence within the ICB1/GC1 oligonucleotide from the topoisomerase II $\alpha$  promoter. Due to time constraints it was unable to be determined how these proteins may function to regulate gene transcription. Therefore the roles of Ku86 and Ku70 for topoisomerase II $\alpha$  expression can only be speculated by comparison with other studies. For future work it could be useful to use RNAi and/or a dominant-negative expression vector of the Ku86/70 heterodimer in conjunction with the -617topo2 $\alpha$  reporter gene construct in transient co-transfections to determine the effect these proteins might have on transcription. In order to determine if Ku86 and Ku70 bind to the topoisomerase II $\alpha$  promoter *in vivo*, ChIP assays could be carried out using antibodies against these proteins and PCR primers that amplify the relevant protein-binding sites within the topoisomerase II $\alpha$  promoter. It would also be useful to determine if the Ku proteins were capable of interacting with Sp1, Sp3 or NF-Y (in the presence and absence of DNA) by carrying out co-immunoprecipitation assays.

Ku86 and Ku70 are highly abundant nuclear proteins that can bind non-specifically to DNA therefore it is possible that contamination of these proteins may have occurred during purification of the DNA-binding proteins. The use of non-specific competitors with oligonucleotide-affinity chromatography did eliminate some of the Ku protein that was found to purify using this technique (data not shown). This suggests a proportion of the Ku proteins may bind non-specifically to the ICB1mt/GC1wt oligonucleotide. Ku86 requires Ku70 to bind DNA, however Ku70 is thought to be able to bind DNA both with and without Ku86 (Wang *et al.*, 1998A) which suggests Ku70 may have functions independent of Ku86. The oligonucleotide-affinity purification method used in this study demonstrates that both Ku86

and Ku70 co-purify (lanes 3, 4 and 5 figure 4.4; blue arrows and green arrows respectively) with equal affinity, suggesting that in this case both Ku86 and Ku70 may bind to the ICB1mt/GC1wt biotinylated-oligonucleotide with equal affinity therefore it is most likely that Ku86 and Ku70 bind as a heterodimer. This is consistent with others that have found that Ku86 and Ku70 co-purify together (Cary *et al.*, 1997) as heterodimers are required to bind DNA in EMSA assays (Wang *et al.*, 1998B). The fact that both electro-eluted extracts from complex 1 and complex 3 contain substantial amounts of both Ku86 and Ku70 (figure 6.5) may be perceived as an indication that these proteins could be non-specifically binding to DNA ends in these assays. To test if binding of Ku is indeed sequence-specific in EMSA assays, competitors could be made with the binding sequence contained within a closed circular plasmid. If competition is still observed then one could exclude the possibility of binding to DNA ends as there are no ends in closed circular DNA. However, this study provided evidence on two levels that suggest the Ku proteins specifically interact within the intervening sequence of the ICB1mt/GC1wt oligonucleotide. Firstly, the DNA termini-binding assays quite clearly demonstrate that the Ku proteins bind to an internal sequence within the ICB1mt/GC1wt oligonucleotide (figure 6.8). Secondly, cross-linking assays show that both an 85 kDa and 70 kDa protein can cross-link to an internal sequence within the composite ICB1mt/GC1wt-BrdU oligonucleotide (figure 5.4). Many other lines of research maintain that the Ku proteins can bind specifically to internal DNA sequences (Knuth *et al.*, 1990; Genersch *et al.*, 1995; Giffin *et al.*, 1996; Giffin *et al.*, 1997; Cary *et al.*, 1997; Torrance *et al.*, 1998; Xu *et al.*, 2004). A brief overview of these reported Ku-binding sequences did not identify any similarities within the ICB1mt/GC1wt sequence. There have been at least six completely different sequence-specific binding sites described for Ku therefore a consensus sequence is unlikely and additional binding sites may be identified in time. It may be that the interactions of different proteins that are present with the Ku proteins could alter the apparent specificity of Ku binding to DNA, therefore making it difficult to establish a consensus sequence. It has been suggested that sequence-specific binding of Ku86 and Ku70 may specify a role in gene regulation whereas non-specific end-binding of Ku86 and Ku70 may specify a role in double-strand break repair. Ku86 and Ku70 clearly belong to the ever increasing family of proteins described to have dual functional roles.

DNA-PK, along with Ku70 and Ku86 are thought to be required for basal and activated transcription driven by RNA polymerase II (Dvir *et al.*, 1992; Kuhn *et al.*, 1993; Sheppard and Liu, 2000). In addition, DNA-PK has been shown to phosphorylate a number of proteins

such as, p53, c-Myc, RNA polymerase II, TBP, transcription factor IIB (TFIIB), topoisomerases and DNA-PK itself (Dvir *et al.*, 1992; Chibazakura *et al.*, 1997; Bertinato *et al.*, 2003). Although it is thought that the Ku proteins can act as allosteric activators of DNA-PKs, both DNA-PK and the Ku proteins have been shown to be able to bind linear DNA independently of each other and still be catalytically active (Yaneva *et al.*, 1997). The clearest case of where DNA-PK phosphorylation has been shown to be entirely dependent on Ku is for the transcription factor Sp1 (Gottlieb and Jackson, 1993). This suggests a potentially functional role for both Ku86 and Ku70 bound to the ICB1mt/GC1wt oligonucleotide *in vivo*. Ku proteins do not appear to form non-specific aggregates in the absence of DNA, however they are capable of self-association (Cary *et al.*, 1997), which may also support synergistic activation together with Sp1. In this study the presence of DNA-PKs was not investigated, although a large protein (>250 kDa) was seen in electro-eluted extracts isolated from complex 1 (figure 4.8 lane 3 and figure 6.5B lane 3). As has been suggested that DNA-PK can phosphorylate Sp1 (Niina *et al.*, 2007) it may be useful to test these extracts for the presence of DNA-PK by immunoblotting in future work.

## 7.4 PARP-1.

Poly (ADP- ribose) polymerase (PARP-1) is a highly abundant 113-115 kDa predominantly nuclear enzyme. It has low basal enzyme activity which can be rapidly activated by the presence of DNA breaks (Wielckens *et al.*, 1983; Alvarez-Gonzalez and Althaus, 1989), nucleosomes (reviewed in Kim *et al.*, 2005), histones and transcription factors (reviewed in Faraone-Mennella, 2005). This protein is classically implicated in DSB repair however emerging evidence demonstrates that this protein is a major contributor to the post-translational modification (ribosylation) of a range of proteins including histones, DNA polymerases, DNA ligases, p53, Sp1, RNA polymerase II and topoisomerases (reviewed in D'Amours *et al.*, 1999). PARP-1 utilises NAD<sup>+</sup> as a substrate to catalyse the addition of single or multiple poly ADP-ribose polymers (PAR groups) to its donor imparting a negative charge onto the donor relative to the degree of ribosylation, thereby affecting protein-protein or protein-DNA interactions (reviewed in Heeres and Hergenrother, 2007). The major target of ribosylation is PARP-1 itself where the catalytically active species of this enzyme exists as a dimer and the ribosylation reaction is intermolecular (Mendoza-Alvarez and Alvarez-Gonzalez, 1993). Although PARP-1 was found to bind to the biotinylated-ICB1mt/GC1wt oligonucleotide which enabled purification of PARP-1 using oligonucleotide-affinity

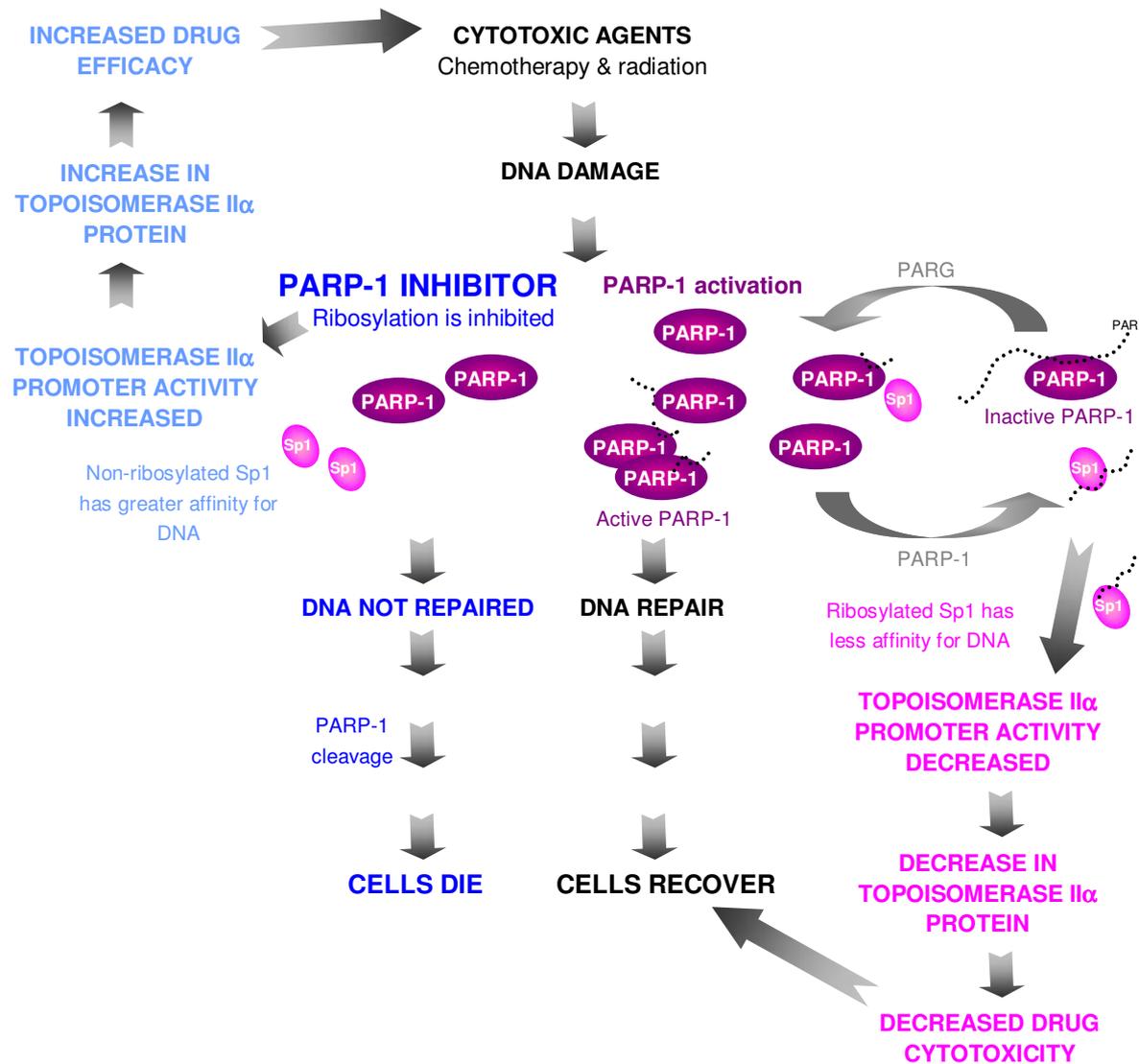
chromatography, DNA termini-binding assays have shown that a large proportion of PARP-1 was found to bind to the ends of the oligonucleotide (figure 6.8). The absence of PARP-1 from electro-eluted extracts and the fact that cross-linking assays did not establish that a protein approximately 113-115 kDa that could bind to the BrdU-oligonucleotides (table 5.2) suggests that this protein probably does not associate with the ICB1mt/GC1wt oligonucleotide in a sequence-specific manner (figure 6.5). This is consistent with observations that PARP can bind electro-statically to DNA ends (Gradwohl *et al.*, 1990). It is most likely PARP-1 was purified due to an association with the DNA termini in oligonucleotide-affinity chromatography as well as protein-protein interactions. PARP-1 has been shown to interact with a range of proteins, including Sp1 and the Ku proteins (Zaniolo *et al.*, 2007; Galande and Kohwi-Shigematsu, 1999; Ju *et al.*, 2006) and it is thought that PARP-1 may specifically be recruited to target promoters by interactions with DNA-binding factors (reviewed in Kraus and Lis, 2003). It is also possible a certain degree of contamination may have occurred due to the relatively high abundance of PARP-1 protein; 1.5-2 million molecules of PARP-1 per cell, which is approximately 5x more than Ku and 200x more than Sp1 in HeLa extracts (Letovsky and Dynan, 1989; Tuteja and Tuteja, 2000; Kraus and Lis, 2003). This apparent abundance of PARP-1 was also observed in this study, as immunoblots (with equal protein loading) were quite often over-exposed and over-loaded with PARP-1 in comparison to Sp1 or Sp3 (data not shown).

Even though PARP-1 was not shown to bind directly to the topoisomerase II $\alpha$  promoter, in the presence of PJ34 (a PARP-1 inhibitor) it was demonstrated that ribosylation may have an inhibitory effect on topoisomerase II $\alpha$  promoter activity (figure 6.1). In transient transfection assays, when ribosylation was inhibited, there was a concomitant increase in topoisomerase II $\alpha$  promoter activity in a dose-dependent manner. This was shown to be reliant on the presence of intact GC elements within the promoter. The EMSA experiments revealed that inhibition of ribosylation may enhance a specific protein-DNA interaction *in vitro* (figure 6.4A) and this is the protein-DNA interaction is known to involve Sp1 (figure 7.1). Immunoblots confirmed that any observed increase in Sp1 binding was not due to an overall increase in the amount of Sp1 protein (figure 6.4B). Collectively, these experiments represent a functional interaction between PARP-1 and Sp1 in the regulation of the topoisomerase II $\alpha$  promoter, where non-ribosylated Sp1 has an enhanced affinity for DNA. This is consistent with observations that ribosylation of Sp1 can inhibit Sp1-DNA interactions and inhibit gene

activation (Zaniolo *et al.*, 2007) and that an increase in Sp1 bound to the topoisomerase II $\alpha$  promoter can increase promoter activity (Magan *et al.*, 2003, Allen *et al.*, 2004; Williams *et al.*, 2007). To confirm that ribosylation can inhibit the binding of Sp1 to GC elements within the topoisomerase II $\alpha$  promoter it was planned to carry out ChIP assays with ribosylation specific antibodies using cells that were exposed to increasing amounts of PJ34. The ribosylation specific antibody that was trialed was found to bind non-specifically to indiscriminate proteins and therefore a number of experiments had to be abandoned at this time.

#### **7.4.1 PARP-1 inhibitors.**

Treatment with radiotherapy and chemotherapies can cause double-stranded DNA breaks in cells, thereby causing cell death. In some cases where cellular DNA damage can be repaired, the cells survive and the resulting population of cells may acquire resistance to these cytotoxic agents. Over 150 DNA repair enzymes are thought to be involved in this process and PARP-1 is implicated in a vast number of different pathways, therefore the way in which PARP-1 inhibitors are effective cannot be narrowed down to a single mechanism (reviewed in Leal *et al.*, 2009). It has been found however that PAR levels in mammalian cells can increase 500-fold after generation of DNA damage and the synthesis of PAR is directly proportional to the number of double-stranded and single-stranded DNA breaks (reviewed in D'Amours *et al.*, 1999). It has also been well documented that PARP-1 inhibitors can potentiate the cytotoxicity of anticancer drugs by essentially impairing the cells natural ability to repair DNA damage (reviewed in Plummer, 2006; Gagne *et al.*, 2006). In the presence of the PARP-1 inhibitor PJ34, this study demonstrated that an increase in topoisomerase II $\alpha$  promoter activity was observed due to an increase in the amount of Sp1 bound at the GC1 element. Taken together these results suggest an additional mechanism by which these PARP-1 inhibitors may improve efficacy of cytotoxic agents in the treatment of cancers (summarised in figure 7.3).



**Figure 7.3: Increased cytotoxicity of chemotherapy in the presence of PARP-1 inhibitors.**

Simplified representation of two processes by which PARP-1 inhibitors may increase the cellular cytotoxicity of anti-cancer drugs. Topoisomerase II $\alpha$  targeting drugs have greater efficacy in highly proliferating cells where topoisomerase II $\alpha$  protein levels are elevated. As more double-stranded breaks are procured, cell death ensues due to an inability to correct all the DNA damage. PARP-1 is activated upon recognition of DNA damage, dimerises and is auto-ribosylated. Excessive ribosylation inactivates PARP-1 by decreasing DNA interactions. The negatively charged poly (ADP-ribose) groups (PAR $\dots$ ) have a short cellular half-life (about 1 minute) due to the presence of poly (ADP-ribose) glycohydrolase (PARG). PARG is activated by an increase in the cellular concentration of PAR. Ribosylated Sp1 has a lower affinity for DNA, thereby decreasing topoisomerase II $\alpha$  promoter activity and the amount of topoisomerase II $\alpha$  protein. PARP-1 inhibitors reduce the amount of ribosylated-Sp1 correlated with an increase in topoisomerase II $\alpha$  promoter activity and topoisomerase II $\alpha$  protein, thereby increasing anti-cancer drug efficacy.

Admittedly, there are a vast range of different PARP-1 inhibitors available (reviewed in Cosi, 2002) and many more in development, that vary in toxicity and in a range of cellular functions in cancer treatment, including inhibiting angiogenesis (Tentori *et al.*, 2007). PARP-1 inhibitors also appear to protect against the cardiotoxicity of doxorubicin (Pacher *et al.*, 2002; De Soto *et al.*, 2006). By combining the two treatments it may be possible to lower the dosage and reduce the toxic side effects commonly seen with many chemotherapeutic agents. It may be useful to carry out a series of experiments using a topoisomerase II $\alpha$  targeting drug, such as doxorubicin in conjunction with PJ34 (and perhaps a range of other PARP-1 inhibitors) in breast cancer cell lines and examine the levels of topoisomerase II $\alpha$  protein and changes in various transcription factors before and after treatment to further test the effects of this type of drug cocktail. Examining the relative cytotoxicity of these drug combinations could also be useful; PARP-1 cleavage could also be used as a marker for cell death as it is commonly shown to be cleaved during apoptosis (Germain *et al.*, 1999).

This study provides evidence on one level that PARP-1 is involved in gene transcription, by means of post-translational modification of transcription factors. PARP-1 has also been found to be important for the transcription-related regulation of chromatin structure where a complex of PARP-1, Ku86/70 and topoisomerase II $\beta$  protein introduces a double-stranded break in DNA that is involved in relaxation of chromatin in response to a nuclear signal (Ju *et al.*, 2006; Lis and Kraus, 2006). PARP-1 can cause relaxation of chromatin by means of ribosylation imparting a negative charge on histones (H1 and H2) thought to cause electrostatic repulsion against bound DNA (Rouleau *et al.*, 2004), this in turn enables access of transcription factors to DNA. Overall, PARP-1 is involved in cell replication, DNA repair, chromatin architecture, cell death and transcriptional regulation, with so many essential pathways it makes it difficult to isolate exact mechanisms by which this protein would act. To examine individual PARP-1 events, experiments would need to be carried out in PARP-1 deficient cells, such as those derived from knockout mice. Alternatively it may be possible to use mammalian cultured cells that have been stably transfected with PARP-1 siRNA, or a mammalian expression vector designed to express PARP-1 siRNA from a temperature responsive promoter, these approaches may be useful as tissue specific cells could then also be utilised if required.

## 7.5 PSF.

PSF is classically known for having a role in RNA splicing, originally identified as a protein component of spliceosomes (Patton *et al.*, 1993; Shav-Tal and Zipori, 2002), in conjunction with its binding counterpart p54<sup>nrb</sup> (Rosonina *et al.*, 2005). PSF has both DNA and RNA binding domains that can repress multiple oncogenic genes in human cancer cells lines and thus has a role as a tumour suppressor protein (Garen and Song, 2008). For example it has been shown that increased expression of PSF protein in MCF7 cells (human breast cancer cell line which normally expresses low levels of PSF) results in decreased cell growth and inhibition of proliferation and colonisation of tumour cells (Song *et al.*, 2005). PSF/p54<sup>nrb</sup> have been implicated in transcriptional control, as it has been shown they bind to various DNA-binding nuclear hormone receptors (Mathur *et al.*, 2001; Urban and Bodenbun, 2002) and the C-terminal domain of RNA polymerase II (Emili *et al.*, 2002). Hyperphosphorylation of PSF is thought to allow the protein to bind to a different set of proteins (Shav-Tal *et al.*, 2001), therefore it would be useful to be able to determine the post-translational modification of this protein when it is bound to the topoisomerase II $\alpha$  promoter. A very small amount of PSF was detected in electro-eluted extract from complex 1, and it is possible that binding of PSF to the biotinylated-ICB1mt/GC1wt oligonucleotide used in oligonucleotide-affinity chromatography was enhanced due to protein-protein interactions with Sp1, as well as with Ku86 and Ku70. PSF has been shown to have both sequence-dependent (Urban *et al.*, 2000; Iacobazzi *et al.*, 2005) and sequence-independent DNA-binding capabilities (Akhmedov and Lopez, 2000). Preliminary data from DNA-termini binding assays has demonstrated that PSF is likely to be associating with an internal region (data not shown) within the ICB1mt/GC1wt oligonucleotide; however these experiments would need to be repeated to confirm this result. There is a putative PSF-binding consensus sequence (“CTGGTC” just to the right of the ICB1 element) and cross-linking assays have demonstrated that there are several proteins 90-105 kDa in size that are capable of binding to the intervening sequence (figures 5.3 and 5.4). These correspond to the observed sizes of PSF. It may be possible that PSF could bind to this region, however most studies regard PSF as an inhibitor of gene expression and therefore this does not relate to the increase in promoter activity that is observed when enhanced binding of complex 3 is induced. Therefore, even though PSF may bind within the intervening sequence it is probably not a component of complex 3, which is in agreement with the observation that PSF was not detected in electro-eluted extracts from complex 3 (figure 6.5A). More work needs to be carried out in order to establish a role for PSF in regulating the topoisomerase II $\alpha$

gene. Firstly, it would need to be confirmed that PSF does bind to the topoisomerase II $\alpha$  promoter as such a small amount was detected in this study. This could be achieved by either using a competitor in EMSA, or introducing a mutation in the ICB1/GC1 oligonucleotide within the putative PSF-binding consensus sequence. Antibody against PSF was used in antibody supershift assays in this study, however no distinctive changes in band mobility were observed (data not shown). Due to the very small amount of PSF found to bind to the ICB1mt/GC1wt sequence, PSF binding may simply be an artefact. HeLa cells have been found to have a short deletion in the region of the PSF gene that codes for the DNA-binding domain for the PSF protein (Song *et al.*, 2005), however the implications of this were not investigated in detail by these investigators. Nevertheless this represents a possibility that HeLa cells have aberrant forms of PSF protein which may exhibit decreased DNA-binding properties. This may be worth investigating as it could represent yet another mechanism in the regulation of the topoisomerase II $\alpha$  gene.

## **7.6 Improving methods & further research.**

It was envisaged at the beginning of this study that it would be possible to purify and identify additional proteins that were found to bind to the composite ICB1/GC1 oligonucleotide; it was not considered that these protein-DNA interactions would involve so many different protein components. Some of the minor components were identified with surprising outcomes. However the lack of a significant antibody supershift against these identified proteins and lack of any further identification of proteins has limited both interpretation of data and further investigation of candidate proteins. Nevertheless, some valid correlations have been made and the intervening sequence was shown to have a regulatory role in expression of the topoisomerase II $\alpha$  gene. Therefore identifying DNA-binding proteins and dissecting potential regulatory roles is still of significance. It is possible that other currently unidentified protein(s) may be in very low abundance, or they may consist of smaller subunits which also make purification and identification difficult. Some possibilities to address these problems are presented in the following sections.

### **7.6.1 Affinity chromatography.**

Oligonucleotide-affinity chromatography was relatively successful in obtaining an extract that was enhanced for the presence of DNA-binding proteins, however this method was not

as effective as originally envisaged. A high degree of non-specific binding occurred with the magnetic beads and proteins used in purification. It has been recommended that the addition of a non-specific competitor will enhance purity of the final product using affinity chromatography (Kadonaga and Tjian, 1986). The addition of non-specific competitors to crude extracts before the purification steps reduced almost all the non-specific binding to the beads; however it appeared as though a significant proportion of proteins still purified due to protein-protein interactions with DNA-bound proteins. Whether any of these were bonafide specific interactions or artefacts of the method remains to be determined. The purification and subsequent identification of NF90 supports the latter. To improve on this method of purification, and to obtain greater specificity, it may be useful to first purify extracts using heparin-sepharose to produce an extract enriched in DNA-binding proteins. Then oligonucleotide-affinity chromatography in conjunction with the non-specific competitors could be used to further isolate the specific factors. A biotinylated-ICB1mt/CENins/GC1wt oligonucleotide or a concatemerised version of the ICB1mt/GC1wt oligonucleotide may produce a greater yield of the proteins that are observed to bind specifically to the intervening sequence.

### **7.6.2 Electro-eluting proteins.**

Electro-elution was successful for the isolation of individual components of each protein-DNA complex within EMSA gels. This was a very tedious process which started with femtomol concentrations of proteins that needed to be constantly pooled and concentrated to achieve amounts that could be visualised using standard protein staining techniques. It is possible a substantial amount of protein was lost due to degradation during the ultrafiltration process. Cutting bands out of several EMSA gels was not always an accurate process. This could be improved by achieving an effective method to stain the proteins that may be present in EMSA gels prior to electro-elution of protein from complex 3. It may also be practical to use the ICB1mt/CENins/GC1wt oligonucleotide in EMSAs to get a higher yield of these proteins.

### **7.6.3 Immunoprecipitation.**

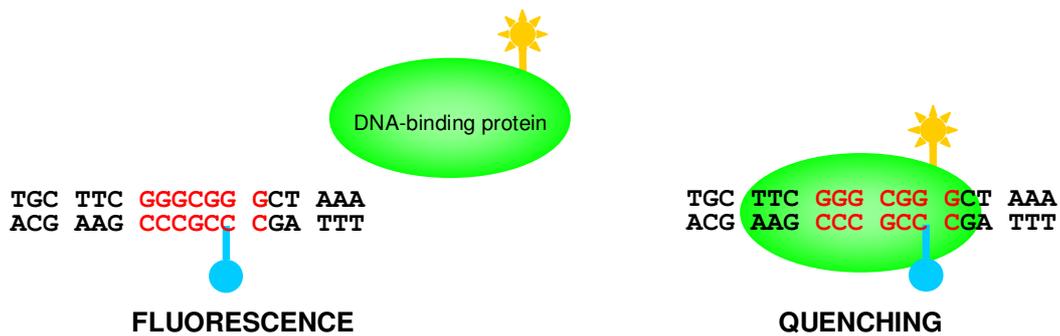
To investigate if protein-protein interactions were occurring between Ku86, Ku70, Sp1, Sp3, PSF and PARP-1, immunoprecipitation assays were attempted during this study but unfortunately due to a large amount of non-specific binding of protein to the agarose-A/G

beads these could not be evaluated. Blocking the agarose-A/G beads with BSA (as recommended by manufacturer) or pre-clearing the extract did not appear to greatly improve this situation. These could be attempted again, using more stringent washing steps, smaller amounts of protein and antibody during the immunoprecipitation steps, as well as using greater dilutions of the secondary antibody during the immunoblot process. This method could also be used to examine the post-translational modifications of proteins; an example would be to use ribosylation specific antibodies for the immunoprecipitation followed by immunodetection using antibodies against PARP-1 and Sp1. The addition of the ICB1mt/GC1wt oligonucleotide to the extract before the immunoprecipitation process may also increase specificity of this assay. As it was not possible to positively identify proteins Ku86, Ku70, PSF and PARP-1 by using antibody supershift, immunodepletion of HeLa extract of individual proteins could be trialled in an attempt to identify these components through reduced binding to the oligonucleotide compared to non-immunodepleted extracts.

#### **7.6.4 Other methods to examine protein–DNA interactions.**

EMSA and cross-linking were used successfully in this study, however it may be useful to use other methods of detecting protein-DNA interactions when the concentrations of proteins are very low. One method that may be practical incorporates the denaturation and separation of proteins by SDS-PAGE, followed by transfer to a positively charged membrane, then the proteins are then slowly re-natured on the membrane and this is probed for DNA-binding proteins by incubation with a <sup>32</sup>P-labelled oligonucleotide.

If possible, using individually purified proteins could enhance the understanding of the protein interactions at the topoisomerase II $\alpha$  promoter. Purified proteins could be used in EMSA reactions or a different technology that utilises fluorescence resonance energy transfer (FRET). This FRET technology is based on the fluorescence transfer between two interacting molecules (Heyduk, 2002), two separate components are individually labelled with a fluorophore or a quencher (explained in figure 7.4). In this case the oligonucleotide and a recombinant protein could be designed to incorporate the specific label. If the purified recombinant DNA-binding protein brings together the two components the overall fluorescent signal will be quenched, however this approach will still only provide information about *in vitro* protein-DNA interactions.



**Figure 7.4: Schematic representation of fluorescence resonance energy transfer for detection of DNA-binding proteins.**

The representation above uses the GC1 sequence from the topoisomerase II $\alpha$  promoter as an example. The double-stranded oligonucleotide is labelled with a specific quencher component. The purified recombinant DNA-binding protein has been modified to incorporate a fluorophore component. If both the oligonucleotide and protein interact, both the fluorophore and quencher will come into close proximity to each other permitting FRET between them and a change in fluorescent signal will be detected.

### 7.6.5 Functional assays.

Throughout the course of this study, a number of experimental techniques were used successfully, such as EMSA, immunoblotting, cross-linking and electro-eluting proteins from EMSA gels. A major consideration with these types of techniques is that they are only an indication of the types of interactions that may be occurring *in vitro*. *In vivo*, inside a cell DNA is organised into a highly compact chromatin structure, where the target gene first has to be remodelled before access of transcription factors and other binding proteins can begin. In order to evaluate the consequences of the protein-protein and protein-DNA interactions described in this study it will be necessary to complete this research with a series of *in vivo* functional assays. Some possible techniques include chromatin immunoprecipitation (ChIP) assays, RNAi, fluorescence *in-situ* hybridisation (FISH) or transient transfections that enable reporter constructs to be monitored in live cells (Gene blazer; Invitrogen).

### 7.6.5.1 ChIP assays.

Chromatin immunoprecipitation works on the basis that interacting proteins and DNA are captured from within the cell, by means of cross-linking proteins to chromatin using formaldehyde. The genomic DNA is fragmented and extracted from the cells, followed by immunoprecipitation with antibodies against the desired protein. The cross-linking is reversed and the DNA is further purified to remove the associating proteins. PCR primers are designed specific to the fragment of DNA to which the protein is thought to interact. Consequently if the protein and DNA are found to interact *in vivo* a PCR product will be observed. These assays could also be useful to monitor any changes in protein-DNA interactions that may occur due to exposure to anti-cancer drugs or PARP-1 inhibitors, as well as changes that may occur with the development of drug-resistance in different cell lines. Using a slightly modified ChIP assay, the post-translational modifications of proteins bound to the topoisomerase II $\alpha$  promoter could also be examined, using antibodies specific for a particular modification such as ribosylation, SUMOylation, phosphorylation or acetylation.

### 7.6.5.2 Protein-protein interactions.

To address the issue of *in vivo* protein-protein interactions it may be possible to express individually labelled recombinant proteins with specific fluorophore(s) and quencher label(s) in mammalian cells and monitor protein-protein interactions in cultured cells. Alternatively, another approach would be to carry out a modified fluorescent *in-situ* hybridisation, where standard primary antibodies are used to identify individual proteins in whole cells but the secondary antibodies would be specifically labelled with fluorophore and quencher components to monitor interacting proteins in whole cells (Heyduk and Heyduk, 2002). With both these techniques that monitor protein-protein interactions *in vivo*, the major limitation would be correlating the results in directly to topoisomerase II $\alpha$  gene regulation. In any context these interactions should also be examined for potential changes in the presence of specific topoisomerase II $\alpha$  targeting drugs, to give an indication of any alterations that may occur upon drug-treatment or acquired drug resistance.

## 7.7 Significance of topoisomerase II $\alpha$ expression.

Predictive biomarkers can be found by screening cancer cell lines or tumours for amplified genes and they are useful as they can help predict the responsiveness to selected therapies. Topoisomerase II $\alpha$  is the target enzyme for anthracyclines, a common chemotherapy used in the treatment of breast and other cancers. It has been well established that increased drug-sensitivity is correlated to the amount of topoisomerase II $\alpha$  protein found in the cell (Burgess *et al.*, 2008; Withoff *et al.*, 1996; Durbecq *et al.*, 2004B). Currently, there is an expansion on the use of the topoisomerase II $\alpha$  gene as a biomarker for the prognosis of patients with breast cancer. This current interest has been fuelled by reports that tumours with an amplification or over-expression of the Her-2 gene respond more favourably to anthracycline-based chemotherapy (Penault-Llorca *et al.*, 2003; Campiglio *et al.*, 2003). The topoisomerase II $\alpha$  gene is adjacent to the Her-2 gene on chromosome 17. It is thought that amplification in both these genes may be correlated in some tumours, although reports on this are conflicting (Bhargava *et al.*, 2005; Beser *et al.*, 2007). In addition, there are inconsistencies in the correlation of topoisomerase II $\alpha$  gene amplification and the relative levels of topoisomerase II $\alpha$  protein (Mueller *et al.*, 2004), suggesting that it may be clinically more relevant to examine the levels of active topoisomerase II $\alpha$  protein rather than gene amplification, deletion or mutations. There is a gap in research as not all studies sufficiently investigate the correlation between gene amplification, gene expression and the amount of active topoisomerase II $\alpha$  protein found within cancer cells and tumours. In recognition of this it would also be useful to examine the relative levels of transcription factors known to regulate the topoisomerase II $\alpha$  gene.

## 7.8 Conclusion.

The amount of active topoisomerase II $\alpha$  protein within a cell is not only important for cell survival but also cellular proliferation. Many of the commonly used anti-cancer drugs are known to target topoisomerase II $\alpha$  protein *in vivo*, where the development of resistance to these drugs is a major clinical problem. Quite often the development of drug resistance is due to an altered expression of the topoisomerase II $\alpha$  gene found in cancerous cells, therefore understanding the mechanisms behind the regulation of the topoisomerase II $\alpha$  gene is of major significance. Transcription factors that bind to the promoter region of a gene can

greatly enhance or suppress the expression of a gene. Thereby additional research still needs to be carried out using additional purification techniques that would enable the identification of candidate proteins that bind to the ICB1/GC1 regulatory region within the topoisomerase II $\alpha$  promoter. There are many factors known to affect topoisomerase II $\alpha$  gene regulation *in vivo*, Sp1, Sp3 and NF-Y have been extensively studied for their role in this. However, throughout the course of this study it is evident that the regulation of the topoisomerase II $\alpha$  gene is more complicated than originally anticipated. This is the first study to recognise that the Ku proteins (Ku86 and Ku70) are capable of an interaction with sequence(s) within the topoisomerase II $\alpha$  promoter. Even though PARP-1 was not found to directly interact with the ICB1/GC1 sequence, it can still affect topoisomerase II $\alpha$  promoter activity through a functional protein-protein interaction with Sp1. PSF may also play a role in transcription regulation, and it may be possible that all these proteins act together with a novel regulatory sequence located between the ICB1 and GC1 elements from within the human topoisomerase II $\alpha$  promoter. To this end, more research needs to be carried out to complete this study, an in depth series of *in vivo* assays would be ideal, incorporating drug-treatment, changes in transcription factors and how they govern topoisomerase II $\alpha$  promoter activity *in vivo*.

# References

- Abbady A. Q., Bronner C., Bathami K., Muller C. D., Jeanblanc M., Mathieu E., Klein J. P., E., C., and Mousli M. (2005). *TCR pathway involves ICBP90 gene down-regulation via E2F binding sites*. *Biochem Pharmacol* 70, 570-579.
- Adachi N., Kobayashi M., and Koyama H. (1997). *Cell cycle-dependent regulation of the mouse DNA topoisomerase IIalpha gene promoter*. *Biochem Biophys Res Commun* 230, 105-109.
- Adamietz P. (1987). *Poly(ADP-ribose) synthase is the major endogenous nonhistone acceptor for poly(ADP-ribose) in alkylated rat hepatoma cells*. *Eur J Biochem* 169, 365-372.
- Adhikary S., and Eilers M. (2005). *Transcriptional regulation and transformation by Myc proteins*. *Nat Rev Mol Cell Biol* 6, 635-645.
- Akhmedov A. T., and Lopez B. S. (2000). *Human 100-kDa homologous DNA-pairing protein is the splicing factor PSF and promotes DNA strand invasion*. *Nucleic Acids Res* 28, 3022-3030.
- Allen K. A., Williams A. O., Isaacs R. J., and Stowell K. M. (2004). *Down-regulation of human topoisomerase IIalpha correlates with altered expression of transcriptional regulators NF-YA and Sp1*. *Anticancer Drugs* 15, 357-362.
- Althaus F. R. (1992). *Poly ADP-ribosylation: a histone shuttle mechanism in DNA excision repair*. *J Cell Sci* 102 ( Pt 4), 663-670.
- Alvarez-Gonzalez R., and Althaus F. R. (1989). *Poly(ADP-ribose) catabolism in mammalian cells exposed to DNA-damaging agents*. *Mutat Res* 218, 67-74.
- Arima Y., Hirota T., Bronner C., Mousli M., Fujiwara T., Niwa S., Ishikawa H., and Saya H. (2004). *Down-regulation of nuclear protein ICBP90 by p53/p21Cip1/WAF1-dependent DNA-damage checkpoint signals contributes to cell cycle arrest at G1/S transition*. *Genes Cells* 9, 131-142.

- Asano T., An T., Mayes J., Zwelling L. A., and Kleinerman E. S. (1996). *Transfection of human topoisomerase II alpha into etoposide-resistant cells: transient increase in sensitivity followed by down-regulation of the endogenous gene*. *Biochem J* 319 ( Pt 1), 307-313.
- Austin C. A., and Marsh K. L. (1998). *Eukaryotic DNA topoisomerase II beta*. *Bioessays* 20, 215-226.
- Baguley B. C., and Ferguson L. R. (1998). *Mutagenic properties of topoisomerase-targeted drugs*. *Biochim Biophys Acta* 1400, 213-222.
- Bertinato J., Tomlinson J. J., Schild-Poulter C., and Hache R. J. (2003). *Evidence implicating Ku antigen as a structural factor in RNA polymerase II-mediated transcription*. *Gene* 302, 53-64.
- Beser A. R., Tuzlali S., Guzey D., Dolek Guler S., Hacıhanefioglu S., and Dalay N. (2007). *HER-2, TOP2A and chromosome 17 alterations in breast cancer*. *Pathol Oncol Res* 13, 180-185.
- Bhargava R., Lal P., and Chen B. (2005). *Chromogenic in situ hybridization for the detection of HER-2/neu gene amplification in breast cancer with an emphasis on tumors with borderline and low-level amplification: does it measure up to fluorescence in situ hybridization?* *Am J Clin Pathol* 123, 237-243.
- Bladen C. L., Udayakumar D., Takeda Y., and Dynan W. S. (2005). *Identification of the polypyrimidine tract binding protein-associated splicing factor.p54(nrb) complex as a candidate DNA double-strand break rejoining factor*. *J Biol Chem* 280, 5205-5210.
- Bliss T. M., and Lane D. P. (1997). *Ku selectively transfers between DNA molecules with homologous ends*. *J Biol Chem* 272, 5765-5773.
- Brandt T. L., and Kroll D. J. (1997). *NF-M trans-activates the human DNA topoisomerase II alpha promoter independently of c-Myb in HL-60 cells*. *Leuk Res* 21, 711-720.
- Bredel M. (2001). *Anticancer drug resistance in primary human brain tumors*. *Brain Res Brain Res Rev* 35, 161-204.

Briggs M. R., Kadonaga J. T., Bell S. P., and Tjian R. (1986). *Purification and biochemical characterization of the promoter-specific transcription factor Sp1*. *Science* 234, 47-52.

Bronner C., Hopfner R., and Mousli M. (2002). *Transcriptional regulation of the human topoisomerase IIalpha gene*. *Anticancer Res* 22, 605-612.

Bugg B. Y., Danks M. K., Beck W. T., and Suttle D. P. (1991). *Expression of a mutant DNA topoisomerase II in CCRF-CEM human leukemic cells selected for resistance to teniposide*. *Proc Natl Acad Sci U S A* 88, 7654-7658.

Burgess D. J., Doles J., Zender L., Xue W., Ma B., McCombie W. R., Hannon G. J., Lowe S. W., and Hemann M. T. (2008). *Topoisomerase levels determine chemotherapy response in vitro and in vivo*. *Proc Natl Acad Sci U S A* 105, 9053-9058.

Campiglio M., Somenzi G., Olgiati C., Beretta G., Balsari A., Zaffaroni N., Valagussa P., and Menard S. (2003). *Role of proliferation in HER2 status predicted response to doxorubicin*. *Int J Cancer* 105, 568-573.

Capranico G., Guano F., Moro S., Zagotto G., Sissi C., Gatto B., Zunino F., Menta E., and Palumbo M. (1998). *Mapping drug interactions at the covalent topoisomerase II-DNA complex by bisantrene/amsacrine congeners*. *J Biol Chem* 273, 12732-12739.

Cardoso F., Durbecq V., Larsimont D., Paesmans M., Leroy J. Y., Rouas G., Sotiriou C., Renard N., Richard V., Piccart M. J., and Di Leo A. (2004). *Correlation between complete response to anthracycline-based chemotherapy and topoisomerase II-alpha gene amplification and protein overexpression in locally advanced/metastatic breast cancer*. *Int J Oncol* 24, 201-209.

Cary R. B., Peterson S. R., Wang J., Bear D. G., Bradbury E. M., and Chen D. J. (1997). *DNA looping by Ku and the DNA-dependent protein kinase*. *Proc Natl Acad Sci U S A* 94, 4267-4272.

Chen M., and Beck W. T. (1995). *DNA topoisomerase II expression stability and phosphorylation two VM-26-resistant human leukemic CEM sublines*. *Oncol Res* 7, 103-111.

Chene P., Rudloff J., Schoepfer J., Furet P., Meier P., Qian Z., Schlaeppli J. M., Schmitz R., and Radimerski T. (2009). *Catalytic inhibition of topoisomerase II by a novel rationally designed ATP-competitive purine analogue*. BMC Chem Biol 9, 1.

Chibazakura T., Watanabe F., Kitajima S., Tsukada K., Yasukochi Y., and Teraoka H. (1997). *Phosphorylation of human general transcription factors TATA-binding protein and transcription factor IIB by DNA-dependent protein kinase--synergistic stimulation of RNA polymerase II basal transcription in vitro*. Eur J Biochem 247, 1166-1173.

Chiefari E., Brunetti A., Arturi F., Bidart J. M., Russo D., Schlumberger M., and Filetti S. (2002). *Increased expression of AP2 and Sp1 transcription factors in human thyroid tumors: a role in NIS expression regulation?* BMC Cancer 2, 35.

Chodosh L. A. (1996). *UV crosslinking of proteins to nucleic acids*. Curr Protocols in Molecular Biology 12.5.1-12.5.8.

Chu S., and Ferro T. J. (2005). *Sp1: regulation of gene expression by phosphorylation*. Gene 348, 1-11.

Coon J. S., Marcus E., Gupta-Burt S., Seelig S., Jacobson K., Chen S., Renta V., Fronda G., and Preisler H. D. (2002). *Amplification and overexpression of topoisomerase IIalpha predict response to anthracycline-based therapy in locally advanced breast cancer*. Clin Cancer Res 8, 1061-1067.

Cosi C. (2002). *New inhibitors of poly(ADP-ribose) polymerase and their potential therapeutic targets*. Expert Opin Ther Patents 12, 1047-1071.

Courey A. J., Holtzman D. A., Jackson S. P., and Tjian R. (1989). *Synergistic activation by the glutamine-rich domains of human transcription factor Sp1*. Cell 59, 827-836.

Coustry F., Hu Q., de Crombrughe B., and Maity S. N. (2001). *CBF/NF-Y functions both in nucleosomal disruption and transcription activation of the chromatin-assembled topoisomerase IIalpha promoter. Transcription activation by CBF/NF-Y in chromatin is dependent on the promoter structure*. J Biol Chem 276, 40621-40630.

D'Amours D., Desnoyers S., D'Silva I., and Poirier G. G. (1999). *Poly(ADP-ribosylation) reactions in the regulation of nuclear functions*. *Biochem J* 342 ( Pt 2), 249-268.

Darby M. K., Schmitt B., Jongstra-Bilen J., and Vosberg H. P. (1985). *Inhibition of calf thymus type II DNA topoisomerase by poly(ADP-ribosylation)*. *EMBO J* 4, 2129-2134.

Davie J. R., He S., Li L., Sekhavat A., Espino P., Drohic B., Dunn K. L., Sun J. M., Chen H. Y., Yu J., *et al.* (2008). *Nuclear organization and chromatin dynamics--Sp1, Sp3 and histone deacetylases*. *Adv Enzyme Regul* 48, 189-208.

De Luca P., Majello B., and Lania L. (1996). *Sp3 represses transcription when tethered to promoter DNA or targeted to promoter proximal RNA*. *J Biol Chem* 271, 8533-8536.

De Soto J. A., Wang X., Tominaga Y., Wang R. H., Cao L., Qiao W., Li C., Xu X., Skoumbourdis A. P., Prindiville S. A., *et al.* (2006). *The inhibition and treatment of breast cancer with poly (ADP-ribose) polymerase (PARP-1) inhibitors*. *Int J Biol Sci* 2, 179-185.

Dennig J., Beato M., and Suske G. (1996). *An inhibitor domain in Sp3 regulates its glutamine-rich activation domains*. *EMBO J* 15, 5659-5667.

Dereuddre S., Delaporte C., and Jacquemin-Sablon A. (1997). *Role of topoisomerase II beta in the resistance of 9-OH-ellipticine-resistant Chinese hamster fibroblasts to topoisomerase II inhibitors*. *Cancer Res* 57, 4301-4308.

Dinh P., Sotiriou C., and Piccart M. J. (2007). *The evolution of treatment strategies: aiming at the target*. *Breast* 16 Suppl 2, S10-16.

Dressler L. G., Berry D. A., Broadwater G., Cowan D., Cox K., Griffin S., Miller A., Tse J., Novotny D., Persons D. L., *et al.* (2005). *Comparison of HER2 status by fluorescence in situ hybridization and immunohistochemistry to predict benefit from dose escalation of adjuvant doxorubicin-based therapy in node-positive breast cancer patients*. *J Clin Oncol* 23, 4287-4297.

Durbecq V., Desmed C., Paesmans M., Cardoso F., Di Leo A., Mano M., Rouas G., Leroy J. Y., Sotiriou C., Piccart M., and Larsimont D. (2004A). *Correlation between topoisomerase-IIalpha gene amplification and protein expression in HER-2 amplified breast cancer*. *Int J Oncol* 25, 1473-1479.

Durbecq V., Paesmans M., Cardoso F., Desmedt C., Di Leo A., Chan S., Friedrichs K., Pinter T., Van Belle S., Murray E., *et al.* (2004B). *Topoisomerase-II alpha expression as a predictive marker in a population of advanced breast cancer patients randomly treated either with single-agent doxorubicin or single-agent docetaxel*. *Mol Cancer Ther* 3, 1207-1214.

Dvir A., Peterson S. R., Knuth M. W., Lu H., and Dynan W. S. (1992). *Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II*. *Proc Natl Acad Sci U S A* 89, 11920-11924.

el-Deiry W. S., Kern S. E., Pietenpol J. A., Kinzler K. W., and Vogelstein B. (1992). *Definition of a consensus binding site for p53*. *Nat Genet* 1, 45-49.

Emili A., Shales M., McCracken S., Xie W., Tucker P. W., Kobayashi R., Blencowe B. J., and Ingles C. J. (2002). *Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD*. *RNA* 8, 1102-1111.

Errington F., Willmore E., Leontiou C., Tilby M. J., and Austin C. A. (2004). *Differences in the longevity of topo IIalpha and topo IIbeta drug-stabilized cleavable complexes and the relationship to drug sensitivity*. *Cancer Chemother Pharmacol* 53, 155-162.

Faraone-Mennella M. R. (2005). *Chromatin architecture and functions: the role(s) of poly(ADP-RIBOSE) polymerase and poly(ADPribosyl)ation of nuclear proteins*. *Biochem Cell Biol* 83, 396-404.

Fernandez P. C., Frank S. R., Wang L., Schroeder M., Liu S., Greene J., Cocito A., and Amati B. (2003). *Genomic targets of the human c-Myc protein*. *Genes Dev* 17, 1115-1129.

Fortune J. M., and Osheroff N. (2000). *Topoisomerase II as a target for anticancer drugs: when enzymes stop being nice*. *Prog Nucleic Acid Res Mol Biol* 64, 221-253.

Gadgil H., and Jarrett H. W. (1999). *Heparin elution of transcription factors from DNA-Sepharose columns*. J Chromatogr A 848, 131-138.

Gadgil H., Oak S. A., and Jarrett H. W. (2001). *Affinity purification of DNA-binding proteins*. J Biochem Biophys Methods 49, 607-624.

Gagne J. P., Hendzel M. J., Droit A., and Poirier G. G. (2006). *The expanding role of poly(ADP-ribose) metabolism: current challenges and new perspectives*. Curr Opin Cell Biol 18, 145-151.

Galande S., and Kohwi-Shigematsu T. (1999). *Poly(ADP-ribose) polymerase and Ku autoantigen form a complex and synergistically bind to matrix attachment sequences*. J Biol Chem 274, 20521-20528.

Galvagni F., Capo S., and Oliviero S. (2001). *Sp1 and Sp3 physically interact and co-operate with GABP for the activation of the utrophin promoter*. J Mol Biol 306, 985-996.

Gao H., Huang K. C., Yamasaki E. F., Chan K. K., Chohan L., and Snapka R. M. (1999). *XK469, a selective topoisomerase IIbeta poison*. Proc Natl Acad Sci U S A 96, 12168-12173.

Garen A., and Song X. (2008). *Regulatory roles of tumor-suppressor proteins and noncoding RNA in cancer and normal cell functions*. Int J Cancer 122, 1687-1689.

Ge Y., Konrad M. A., Matherly L. H., and Taub J. W. (2001). *Transcriptional regulation of the human cystathionine beta-synthase -1b basal promoter: synergistic transactivation by transcription factors NF-Y and Sp1/Sp3*. Biochem J 357, 97-105.

Geffen D. B., and Man S. (2002). *New drugs for the treatment of cancer 1990-2001*. Isr Med Assoc J 4, 1124-1131.

Genersch E., Eckerskorn C., Lottspeich F., Herzog C., Kuhn K., and Poschl E. (1995). *Purification of the sequence-specific transcription factor CTCBF involved in the control of human collagen IV genes: subunits with homology to Ku antigen*. EMBO J 14, 791-800.

Germain M., Affar E. B., D'Amours D., Dixit V. M., Salvesen G. S., and Poirier G. G. (1999). *Cleavage of automodified poly(ADP-ribose) polymerase during apoptosis. Evidence for involvement of caspase-7.* J Biol Chem 274, 28379-28384.

Giffin W., Kwast-Welfeld J., Rodda D. J., Prefontaine G. G., Traykova-Andonova M., Zhang Y., Weigel N. L., Lefebvre Y. A., and Hache R. J. (1997). *Sequence-specific DNA binding and transcription factor phosphorylation by Ku Autoantigen/DNA-dependent protein kinase. Phosphorylation of Ser-527 of the rat glucocorticoid receptor.* J Biol Chem 272, 5647-5658.

Giffin W., Torrance H., Rodda D. J., Prefontaine G. G., Pope L., and Hache R. J. (1996). *Sequence-specific DNA binding by Ku autoantigen and its effects on transcription.* Nature 380, 265-268.

Gottlieb T. M., and Jackson S. P. (1993). *The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen.* Cell 72, 131-142.

Gradwohl G., Menissier de Murcia J. M., Molinete M., Simonin F., Koken M., Hoeijmakers J. H., and de Murcia G. (1990). *The second zinc-finger domain of poly(ADP-ribose) polymerase determines specificity for single-stranded breaks in DNA.* Proc Natl Acad Sci U S A 87, 2990-2994.

Grant C. E., Valdimarsson G., Hipfner D. R., Almquist K. C., Cole S. P., and Deeley R. G. (1994). *Overexpression of multidrug resistance-associated protein (MRP) increases resistance to natural product drugs.* Cancer Res 54, 357-361.

Grue P., Grasser A., Sehested M., Jensen P. B., Uhse A., Straub T., Ness W., and Boege F. (1998). *Essential mitotic functions of DNA topoisomerase IIalpha are not adopted by topoisomerase IIbeta in human H69 cells.* J Biol Chem 273, 33660-33666.

Gurel S., Dolar E., Yerci O., Samli B., Ozturk H., Nak S. G., Gulden M., and Memik F. (1999). *The relationship between c-erbB-2 oncogene expression and clinicopathological factors in gastric cancer.* J Int Med Res 27, 74-78.

Haince J. F., Rouleau M., and Poirier G. G. (2006). *Transcription. Gene expression needs a break to unwind before carrying on.* Science 312, 1752-1753.

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

Harker W. G., Slade D. L., Drake F. H., and Parr R. L. (1991). *Mitoxantrone resistance in HL-60 leukemia cells: reduced nuclear topoisomerase II catalytic activity and drug-induced DNA cleavage in association with reduced expression of the topoisomerase II beta isoform*. Biochemistry 30, 9953-9961.

Harker W. G., Slade D. L., Parr R. L., Feldhoff P. W., Sullivan D. M., and Holguin M. H. (1995). *Alterations in the topoisomerase II alpha gene messenger RNA and subcellular protein distribution as well as reduced expression of the DNA topoisomerase II beta enzyme in a mitoxantrone-resistant HL-60 human leukemia cell line*. Cancer Res 55, 1707-1716.

Heeres J. T., and Hergenrother P. J. (2007). *Poly(ADP-ribose) makes a date with death*. Curr Opin Chem Biol 11, 644-653.

Herzog C. E., and Zwelling L. A. (1997). *Evaluation of a potential regulatory role for inverted CCAAT boxes in the human topoisomerase II alpha promoter*. Biochem Biophys Res Commun 232, 608-612.

Heyduk T. (2002). *Measuring protein conformational changes by FRET/LRET*. Curr Opin Biotechnol 13, 292-296.

Heyduk T., and Heyduk E. (2002). *Molecular beacons for detecting DNA binding proteins*. Nat Biotechnol 20, 171-176.

Hochhauser D., Stanway C. A., Harris A. L., and Hickson I. D. (1992). *Cloning and characterization of the 5'-flanking region of the human topoisomerase II alpha gene*. J Biol Chem 267, 18961-18965.

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

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

Hurst H. C. (2001). *Update on HER-2 as a target for cancer therapy: the ERBB2 promoter and its exploitation for cancer treatment*. *Breast Cancer Res* 3, 395-398.

Iacobazzi V., Infantino V., Costanzo P., Izzo P., and Palmieri F. (2005). *Functional analysis of the promoter of the mitochondrial phosphate carrier human gene: identification of activator and repressor elements and their transcription factors*. *Biochem J* 391, 613-621.

Isaacs R. J., Davies S. L., Wells N. J., and Harris A. L. (1995). *Topoisomerases II alpha and beta as therapy targets in breast cancer*. *Anticancer Drugs* 6, 195-211.

Isaacs R. J., Harris A. L., and Hickson I. D. (1996). *Regulation of the human topoisomerase IIalpha gene promoter in confluence-arrested cells*. *J Biol Chem* 271, 16741-16747.

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

Ishida R., Iwai M., Marsh K. L., Austin C. A., Yano T., Shibata M., Nozaki N., and Hara A. (1996). *Threonine 1342 in human topoisomerase IIalpha is phosphorylated throughout the cell cycle*. *J Biol Chem* 271, 30077-30082.

Jarvinen T. A., Kononen J., Pelto-Huikko M., and Isola J. (1996). *Expression of topoisomerase IIalpha is associated with rapid cell proliferation aneuploidy and c-erbB2 overexpression in breast cancer*. *Am J Pathol* 148, 2073-2082.

Jarvinen T. A., and Liu E. T. (2003). *Topoisomerase IIalpha gene (TOP2A) amplification and deletion in cancer--more common than anticipated*. *Cytopathology* 14, 309-313.

Jarvinen T. A., Tanner M., Rantanen V., Barlund M., Borg A., Grenman S., and Isola J. (2000). *Amplification, deletion of topoII $\alpha$  associate with ErbB-2 amplification and affect sensitivity topoisomerase II inhibitor doxorubicin in breast cancer*. *Am J Pathol* 156, 839-847.

Jenkins J. R., Ayton P., Jones T., Davies S. L., Simmons D. L., Harris A. L., Sheer D., and Hickson I. D. (1992). *Isolation of cDNA clones encoding the beta isozyme of human DNA topoisomerase II and localisation of the gene to chromosome 3p24*. Nucleic Acids Res 20, 5587-5592.

Jensen L. H., Wessel I., Moller M., Nitiss J. L., Sehested M., and Jensen P. B. (2000). *N-terminal and core-domain random mutations in human topoisomerase II alpha conferring bisdioxopiperazine resistance*. FEBS Lett 480, 201-207.

Jensen L. H., Wessel I., Moller M., Nitiss J. L., Sehested M., and Jensen P. B. (2000). *N-terminal and core-domain random mutations in human topoisomerase II alpha conferring bisdioxopiperazine resistance*. FEBS Lett 480, 201-207.

Jiang Y., Wang L., Gong W., Wei D., Le X., Yao J., Ajani J., Abbruzzese J. L., Huang S., and Xie K. (2004). *A high expression level of insulin-like growth factor I receptor is associated with increased expression of transcription factor Sp1 and regional lymph node metastasis of human gastric cancer*. Clin Exp Metastasis 21, 755-764.

Jin S., Inoue S., and Weaver D. T. (1997). *Functions of the DNA dependent protein kinase*. Cancer Surv 29, 221-261.

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

Ju B. G., Lunyak V. V., Perissi V., Garcia-Bassets I., Rose D. W., Glass C. K., and Rosenfeld M. G. (2006). *A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription*. Science 312, 1798-1802.

Ju B. G., and Rosenfeld M. G. (2006). *A breaking strategy for topoisomerase IIbeta/PARP-1-dependent regulated transcription*. Cell Cycle 5, 2557-2560.

Kadonaga J. T., Carner K. R., Masiarz F. R., and 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., and Tjian R. (1986). *Affinity purification of sequence-specific DNA binding proteins*. Proc Natl Acad Sci U S A 83, 5889-5893.

Kellner U., Rudolph P., and Parwaresch R. (2000). *Human DNA-Topoisomerases - Diagnostic and Therapeutic Implications for Cancer*. Onkologie 23, 424-430.

Kennett S. B., Moorefield K. S., and Horowitz J. M. (2002). *Sp3 represses gene expression via the titration of promoter-specific transcription factors*. J Biol Chem 277, 9780-9789.

Kennett S. B., Udvardi A. J., and Horowitz J. M. (1997). *Sp3 encodes multiple proteins that differ in their capacity to stimulate or repress transcription*. Nucleic Acids Res 25, 3110-3117.

Kim M. Y., Zhang T., and Kraus W. L. (2005). *Poly(ADP-ribosylation) by PARP-1: 'PAR-laying NAD+ into a nuclear signal'*. Genes & Dev 19, 1951-1967.

Kimura K., Nozaki N., Enomoto T., Tanaka M., and Kikuchi A. (1996). *Analysis of M phase-specific phosphorylation of DNA topoisomerase II*. J Biol Chem 271, 21439-21445.

Knuth M. W., Gunderson S. I., Thompson N. E., Strasheim L. A., and Burgess R. R. (1990). *Purification and characterization of proximal sequence element-binding protein 1, a transcription activating protein related to Ku and TREF that binds the proximal sequence element of the human U1 promoter*. J Biol Chem 265, 17911-17920.

Koike M. (2002). *Dimerization translocation and localization of Ku70 and Ku80 proteins*. J Radiat Res (Tokyo) 43, 223-236.

Koike M., Shiomi T., and Koike A. (2001). *Dimerization and nuclear localization of ku proteins*. J Biol Chem 276, 11167-11173.

Kraus W. L., and Lis J. T. (2003). *PARP goes transcription*. Cell 113, 677-683.

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

Kuhn A., Stefanovsky V., and Grummt I. (1993). *The nucleolar transcription activator UBF relieves Ku antigen-mediated repression of mouse ribosomal gene transcription*. Nucleic Acids Res 21, 2057-2063.

Lacroix M., Toillon R. A., and Leclercq G. (2006). *p53 and breast cancer an update*. Endocr Relat Cancer 13, 293-325.

Laemmli U. K., Beguin F., and Gujer-Kellenberger G. (1970). *A factor preventing the major head protein of bacteriophage T4 from random aggregation*. J Mol Biol 47, 69-85.

Larsen A. K., Escargueil A. E., and Skladanowski A. (2003). *Catalytic topoisomerase II inhibitors in cancer therapy*. Pharmacol Ther 99, 167-181.

Larsen A. K., and Skladanowski A. (1998). *Cellular resistance to topoisomerase-targeted drugs: from drug uptake to cell death*. Biochim Biophys Acta 1400, 257-274.

Leal A. P., Vargas J. M. R., Aquilar-Quesada R., I., R. M., L., L. J., Almodovar M. R., and J., O. F. (2009). *PARP inhibitors: New partners in the therapy of cancer and inflammatory diseases*. Free Radic Biol Med XX.

Letovsky J., and Dynan W. S. (1989). *Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence*. Nucleic Acids Res 17, 2639-2653.

Li L., He S., Sun J. M., and Davie J. R. (2004). *Gene regulation by Sp1 and Sp3*. Biochem Cell Biol 82, 460-471.

Liang F., Schaufele F., and Gardner D. G. (2001). *Functional interaction of NF-Y and Sp1 is required for type a natriuretic peptide receptor gene transcription*. J Biol Chem 276, 1516-1522.

Liang S. G., and Maity S. N. (1998). *Pathway of complex formation between DNA and three subunits of CBF/NF-Y. Photocross-linking analysis of DNA-protein interaction and characterization of equilibrium steps of subunit interaction and dna binding*. J Biol Chem 273, 31590-31598.

Lim K., Lee J. I., Yun K. A., Son M. Y., Park J. I., Yoon W. H., and Hwang B. D. (1998). *Reduced level of ATF is correlated with transcriptional repression of DNA topoisomerase II alpha gene during TPA-induced differentiation of HL-60 cells*. *Biochem Mol Biol Int* 46, 35-42.

Linka R. M., Porter A. C., Volkov A., Mielke C., Boege F., and Christensen M. O. (2007). *C-terminal regions of topoisomerase IIalpha and IIbeta determine isoform-specific functioning of the enzymes in vivo*. *Nucleic Acids Res* 35, 3810-3822.

Lis J. T., and Kraus W. L. (2006). *Promoter cleavage: a topoiIbeta and PARP-1 collaboration*. *Cell* 125, 1225-1227.

Liu Z., and Simpson E. R. (1999). *Molecular mechanism for cooperation between Sp1 and steroidogenic factor-1 (SF-1) to regulate bovine CYP11A gene expression*. *Mol Cell Endocrinol* 153, 183-196.

Lomberk G., and Urrutia R. (2005). *The family feud: turning off Sp1 by Sp1-like KLF proteins*. *Biochem J* 392, 1-11.

Luo H. R., Hattori H., Hossain M. A., Hester L., Huang Y., Lee-Kwon W., Donowitz M., Nagata E., and Snyder S. H. (2003). *Akt as a mediator of cell death*. *Proc Natl Acad Sci U S A* 100, 11712-11717.

Luscher B., and Larsson L. G. (1999). *The basic region/helix-loop-helix/leucine zipper domain of Myc proto-oncoproteins: function and regulation*. *Oncogene* 18, 2955-2966.

Magan N. (2002). *Basal transcription of human topoisomerase II*. MSc Thesis Massey University, Palmerston North, New Zealand.

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

Mano M. S., Rosa D. D., De Azambuja E., Ismael G. F., and Durbecq V. (2007). *The 17q12-q21 amplicon: Her2 and topoisomerase-II  $\alpha$  and their importance to the biology of solid tumours*. *Cancer Treat Rev* 33, 64-77.

Mansilla S., Pina B., and Portugal J. (2003). *Daunorubicin-induced variations in gene transcription: commitment to proliferation arrest senescence and apoptosis*. *Biochem J* 372, 703-711.

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

Mastrangelo I. A., Courey A. J., Wall J. S., Jackson S. P., and Hough P. V. (1991). *DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement*. *Proc Natl Acad Sci U S A* 88, 5670-5674.

Mathur M., Tucker P. W., and Samuels H. H. (2001). *PSF is a novel corepressor that mediates its effect through Sin3A and the DNA binding domain of nuclear hormone receptors*. *Mol Cell Biol* 21, 2298-2311.

Matsuo K., Kohno K., Sato S., Uchiumi T., Tanimura H., Yamada Y., and Kuwano M. (1993). *Enhanced expression of the DNA topoisomerase II gene in response to heat shock stress in human epidermoid cancer KB cells*. *Cancer Res* 53, 1085-1090.

McLenachan S. (1998). *Regulation of the human topoisomerase II $\alpha$  promoter*. Honours project Massey University, Palmerston North, New Zealand.

Mendoza-Alvarez H., and Alvarez-Gonzalez R. (1993). *Poly(ADP-ribose) polymerase is a catalytic dimer and the automodification reaction is intermolecular*. *J Biol Chem* 268, 22575-22580.

Millau J. F., Bastien N., and Drouin R. (2009). *P53 transcriptional activities: a general overview and some thoughts*. *Mutat Res* 681, 118-133.

Milona M. A., Gough J. E., and Edgar A. J. (2003). *Expression of alternatively spliced isoforms of human Sp7 in osteoblast-like cells*. *BMC Genomics* 4, 43.

Mimori T., and Hardin J. A. (1986). *Mechanism of interaction between Ku protein and DNA*. J Biol Chem 261, 10375-10379.

Mirski S. E., Bielawski J. C., and Cole S. P. (2003). *Identification of functional nuclear export sequences in human topoisomerase IIalpha and beta*. Biochem Biophys Res Commun 306, 905-911.

Mirski S. E., Voskoglou-Nomikos T., Young L. C., Deeley R. G., Campling B. G., Gerlach J. H., and Cole S. P. (2000). *Simultaneous quantitation of topo II alpha and beta isoform mRNAs in lung tumor cells and normal and malignant lung tissue*. Lab Invest 80, 787-795.

Mo Y. Y., Wang Q., and Beck W. T. (1997). *Down-regulation of topoII $\alpha$  in CEM cells selected for merbarone resistance is associated with reduced expression of Sp3*. Cancer Res 57, 5004-5008.

Molnar G., O'Leary N., Pardee A. B., and W B. D. (1995). *Quantification of DNA-protein interaction by UV cross-linking*. Nucleic Acids Res 23, 3318-3326.

Moorefield K. S., Fry S. J., and Horowitz J. M. (2004). *Sp2 DNA binding activity and trans-activation are negatively regulated in mammalian cells*. J Biol Chem 279, 13911-13924.

Morgan S. E., and Beck W. T. (2001). *Role of an inverted CCAAT element in human topoII $\alpha$  gene expression in ICRF-187-sensitive and -resistant CEM leukemic cells*. Mol Pharmacol 59, 203-211.

Motta M. C., Caretti G., Badaracco G. F., and Mantovani R. (1999). *Interactions of the CCAAT-binding trimer NF-Y with nucleosomes*. J Biol Chem 274, 1326-1333.

Mousli M., Hopfner R., Abbady A. Q., Monte D., Jeanblanc M., Oudet P., Louis B., and Bronner C. (2003). *ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells*. Br J Cancer 89, 120-127.

Mueller R. E., Parkes R. K., Andrulis I., and O'Malley F. P. (2004). *Amplification of the TOP2A gene does not predict high levels of topoisomerase II alpha protein in human breast tumor samples*. Genes Chromosomes Cancer 39, 288-297.

Nagasaki K., and Miki Y. (2006). *Gene expression profiling of breast cancer*. Breast Cancer 13, 2-7.

Niina I., Uchiumi T., Izumi H., Torigoe T., Wakasugi T., Igarashi T., Miyamoto N., Onitsuka T., Shiota M., Okayasu R., et al. (2007). *DNA topoisomerase inhibitor etoposide enhances GC-box-dependent promoter activity via Sp1 phosphorylation*. Cancer Sci 98, 858-863.

Ogiso Y., Tomida A., Lei S., Omura S., and Tsuruo T. (2000). *Proteasome inhibition circumvents solid tumor resistance to topoisomerase II-directed drugs*. Cancer Res 60, 2429-2434.

O'Malley F. P., Chia S., Tu D., Shepherd L. E., Levine M. N., Bramwell V. H., Andrulis I. L., and Pritchard K. I. (2009). *Topoisomerase II alpha and responsiveness of breast cancer to adjuvant chemotherapy*. J Natl Cancer Inst 101, 644-650.

Pacher P., Liaudet L., Bai P., Virag L., Mabley J. G., Hasko G., and Szabo C. (2002). *Activation of poly(ADP-ribose) polymerase contributes to development of doxorubicin-induced heart failure*. J Pharmacol Exp Ther 300, 862-867.

Pascal E., and Tjian R. (1991). *Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism*. Genes Dev 5, 1646-1656.

Patel S., Keller B. A., and Fisher L. M. (2000). *Mutations at arg486 and glu571 in human topoisomerase IIalpha confer resistance to amsacrine: relevance for antitumor drug resistance in human cells*. Mol Pharmacol 57, 784-791.

Patton J. G., Porro E. B., Galceran J., Tempst P., and Nadal-Ginard B. (1993). *Cloning and characterization of PSF a novel pre-mRNA splicing factor*. Genes Dev 7, 393-406.

Penault-Llorca F., Cayre A., Bouchet Mishellany F., Amat S., Feillel V., Le Bouedec G., Ferriere J. P., De Latour M., and Chollet P. (2003). *Induction chemotherapy for breast carcinoma: predictive markers and relation with outcome*. Int J Oncol 22, 1319-1325.

Plummer E. R. (2006). *Inhibition of poly(ADP-ribose) polymerase in cancer*. Curr Opin Pharmacol 6, 364-368.

Pritchard K. I., Shepherd L. E., O'Malley F. P., Andrulis I. L., Tu D., Bramwell V. H., and Levine M. N. (2006). *HER2 and responsiveness of breast cancer to adjuvant chemotherapy*. N Engl J Med 354, 2103-2111.

Reichman T. W., Parrott A. M., Fierro-Monti I., Caron D. J., Kao P. N., Lee C. G., Li H., and Mathews M. B. (2003). *Selective regulation of gene expression by nuclear factor 110, a member of the NF90 family of double-stranded RNA-binding proteins*. J Mol Biol 332, 85-98.

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

Roca J. (2009). *Topoisomerase II: a fitted mechanism for the chromatin landscape*. Nucleic Acids Res 37, 721-730.

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

Rosonina E., Ip J. Y., Calarco J. A., Bakowski M. A., Emili A., McCracken S., Tucker P., Ingles C. J., and Blencowe B. J. (2005). *Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo*. Mol Cell Biol 25, 6734-6746.

Ross J. S., Fletcher J. A., Linette G. P., Stec J., Clark E., Ayers M., Symmans W. F., Pusztai L., and Bloom K. J. (2003). *The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy*. Oncologist 8, 307-325.

Ross S., Best J. L., Zon L. I., and Gill G. (2002). *SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization*. Mol Cell 10, 831-842.

Rouleau M., Aubin R. A., and Poirier G. G. (2004). *Poly(ADP-ribosyl)ated chromatin domains: access granted*. J Cell Sci 117, 815-825.

Sakaguchi A., and Kikuchi A. (2004). *Functional compatibility between isoform alpha and beta of type II DNA topoisomerase*. J Cell Sci 117, 1047-1054.

Salmena L., Lam V., McPherson J. P., and Goldenberg G. J. (2001). *Role of proteasomal degradation in the cell cycle-dependent regulation of DNA topoisomerase II $\alpha$  expression*. *Biochem Pharmacol* 61, 795-802.

Santini M. P., Talora C., Seki T., Bolgan L., and Dotto G. P. (2001). *Cross talk among calcineurin Sp1/Sp3, and NFAT in control of p21(WAF1/CIP1) expression in keratinocyte differentiation*. *Proc Natl Acad Sci U S A* 98, 9575-9580.

Sapetschnig A., Koch F., Rischitor G., Mennenga T., and Suske G. (2004). *Complexity of translationally controlled transcription factor Sp3 isoform expression*. *J Biol Chem* 279, 42095-42105.

Saxena D., Yiu G. K., Ni X., Huang K. C., Mantovani R., Jacquemin-Sablon A. G., and Ng S. W. (2004). *Characterization of promoter elements involved in the down-regulation of topoiI $\alpha$  expression in a drug-resistant cell line*. *Gene* 342, 145-155.

Scovassi A. I., Mariani C., Negroni M., Negri C., and Bertazzoni U. (1993). *ADP-ribosylation of nonhistone proteins in HeLa cells: modification of DNA topo II*. *Exp Cell Res* 206, 177-181.

Senior K. (2006). *Regulation of topoisomerase II alpha expression in humans*. MSc Thesis Massey University, Palmerston North, New Zealand.

Shav-Tal Y., Cohen M., Lapter S., Dye B., Patton J. G., Vandekerckhove J., and Zipori D. (2001). *Nuclear relocalization of the Pre-mRNA splicing factor PSF during apoptosis involves hyperphosphorylation Masking of antigenic epitopes and changes in protein interactions*. *Mol Biol Cell* 12, 2328-2340.

Shav-Tal Y., and Zipori D. (2002). *PSF and p54<sup>nrb</sup>/NonO - multifunctional nuclear proteins*. *FEBS Lett* 531, 109-114.

Sheppard H. M., and Liu X. (2000). *Transcription by RNA polymerase II in DNA-PK deficient scid mouse cells*. *Biochim Biophys Acta* 1493, 41-47.

Shi L., Qiu D., Zhao G., Corthesy B., Lees-Miller S., Reeves W. H., and Kao P. N. (2007). *Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells*. Nucleic Acids Res 35, 2302-2310.

Slamon D. J., Godolphin W., Jones L. A., Holt J. A., Wong S. G., Keith D. E., Levin W. J., Stuart S. G., Udove J., Ullrich A., and et al. (1989). *Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer*. Science 244, 707-712.

Son M. Y., Kim T. J., Kweon K. I., Park J. I., Park C., Lee Y. C., No Z., Ahn J. W., Yoon W. H., Park S. K., et al. (2002). *ATF is important to late S phase-dependent regulation of DNA topoisomerase IIalpha gene expression in HeLa cells*. Cancer Lett 184, 81-88.

Song X., Sun Y., and Garen A. (2005). *Roles of PSF protein and VL30 RNA in reversible gene regulation*. Proc Natl Acad Sci U S A 102, 12189-12193.

Stros M., Polanska E., Struncova S., and Pospisilova S. (2009). *HMGB1 and HMGB2 proteins up-regulate cellular expression of human topoisomerase IIalpha*. Nucleic Acids Res 37, 2070-2086.

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

Sun J. M., Chen H. Y., Moniwa M., Litchfield D. W., Seto E., and Davie J. R. (2002). *The transcriptional repressor Sp3 is associated with CK2-phosphorylated histone deacetylase 2*. J Biol Chem 277, 35783-35786.

Szremska A. (2000). *Sp1, Sp3, NF-Y and the topoisomerase II $\alpha$  promoter*. Honours project Massey University, Palmerston North, New Zealand.

Takano H., Ise T., Nomoto M., Kato K., Murakami T., Ohmori H., Imamura T., Nagatani G., Okamoto T., Ohta R., et al. (1999). *Structural and functional analysis of the control region of the human DNA topoisomerase II alpha gene in drug-resistant cells*. Anticancer Drug Des 14, 87-92.

Tanaka S., Antoniv T. T., Liu K., Wang L., Wells D. J., Ramirez F., and Bou-Gharios G. (2004). *Cooperativity between far upstream enhancer and proximal promoter elements of the human  $\alpha 2(I)$  collagen (COL1A2) gene instructs tissue specificity in transgenic mice.* J Biol Chem 279, 56024-56031.

Tanner M., Isola J., Wiklund T., Erikstein B., Kellokumpu-Lehtinen P., Malmstrom P., Wilking N., Nilsson J., and Bergh J. (2006). *Topoisomerase IIalpha gene amplification predicts favorable treatment response to tailored and dose-escalated anthracycline-based adjuvant chemotherapy in HER-2/neu-amplified breast cancer: Scandinavian Breast Group Trial 9401.* J Clin Oncol 24, 2428-2436.

Tentori L., Lacal P. M., Muzi A., Dorio A. S., Leonetti C., Scarsella M., Ruffini F., Xu W., Min W., Stoppacciaro A., et al. (2007). *Poly(ADP-ribose) polymerase (PARP) inhibition or PARP-1 gene deletion reduces angiogenesis.* Eur J Cancer 43, 2124-2133.

Torrance H., Giffin W., Rodda D. J., Pope L., and Hache R. J. (1998). *Sequence-specific binding of Ku autoantigen to single-stranded DNA.* J Biol Chem 273, 20810-20819.

Towatari M., Adachi K., Marunouchi T., and Saito H. (1998). *Evidence for a critical role of DNA topoisomerase IIalpha in drug sensitivity revealed by inducible antisense RNA in a human leukaemia cell line.* Br J Haematol 101, 548-551.

Trotzler M. A., Bronner C., Bathami K., Mathieu E., Abbady A. Q., Jeanblanc M., Muller C. D., Rochette-Egly C., and Mousli M. (2004). *Phosphorylation of ICBP90 by protein kinase A enhances topoisomerase IIalpha expression.* Biochem Biophys Res Commun 319, 590-595.

Turley H., Comley M., Houlbrook S., Nozaki N., Kikuchi A., Hickson I. D., Gatter K., and Harris A. L. (1997). *The distribution and expression of the two isoforms of DNA topoisomerase II in normal and neoplastic human tissues.* Br J Cancer 75, 1340-1346.

Tuteja R., and Tuteja N. (2000). *Ku autoantigen: a multifunctional DNA-binding protein.* Crit Rev Biochem Mol Biol 35, 1-33.

Uesaka T., Shono T., Kuga D., Suzuki S. O., Niiro H., Miyamoto K., Matsumoto K., Mizoguchi M., Ohta M., Iwaki T., and Sasaki T. (2007). *Enhanced expression of DNA topoisomerase II genes in human medulloblastoma and its possible association with etoposide sensitivity*. J Neurooncol 84, 119-129.

Urban R. J., and Bodenbun Y. (2002). *PTB-associated splicing factor regulates growth factor-stimulated gene expression in mammalian cells*. Am J Physiol Endocrinol Metab 283, E794-798.

Urban R. J., Bodenbun Y., Kurosky A., Wood T. G., and Gasic S. (2000). *Polypyrimidine tract-binding protein-associated splicing factor regulator of transcriptional activity of the porcine p450scc insulin-like growth factor response element*. Mol Endocrinol 14, 774-782.

Valin A., and Gill G. (2007). *Regulation of the dual-function transcription factor Sp3 by SUMO*. Biochem Soc Trans 35, 1393-1396.

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

Varis A., Wolf M., Monni O., Vakkari M. L., Kokkola A., Moskaluk C., Frierson H., Jr., Powell S. M., Knuutila S., Kallioniemi A., and El-Rifai W. (2002). *Targets of gene amplification and overexpression at 17q in gastric cancer*. Cancer Res 62, 2625-2629.

Verdine G. L., and Norman D. P. (2003). *Covalent trapping of protein-DNA complexes*. Annu Rev Biochem 72, 337-366.

Walker J. V., Nitiss K. C., Jensen L. H., Mayne C., Hu T., Jensen P. B., Sehested M., Hsieh T., and Nitiss J. L. (2004). *A mutation in human topoisomerase II alpha whose expression is lethal in DNA repair-deficient yeast cells*. J Biol Chem 279, 25947-25954.

Wang H., Jiang Z., Wong Y. W., Dalton W. S., Futscher B. W., and Chan V. T. (1997B). *Decreased CP-1 (NF-Y) activity results in transcriptional down-regulation of topo II $\alpha$  in a doxorubicin-resistant variant of human multiple myeloma RPMI 8226*. Biochem Biophys Res Commun 237, 217-224.

Wang J., Dong X., Myung K., Hendrickson E. A., and Reeves W. H. (1998A). *Identification of two domains of the p70 Ku protein mediating dimerization with p80 and DNA binding*. J Biol Chem 273, 842-848.

Wang J., Dong X., and Reeves W. H. (1998B). *A model for Ku heterodimer assembly and interaction with DNA. Implications for the function of Ku antigen*. J Biol Chem 273, 31068-31074.

Wang Q., Zambetti G. P., and Suttle D. P. (1997A). *Inhibition of DNA topoisomerase II alpha gene expression by the p53 tumor suppressor*. Mol Cell Biol 17, 389-397.

Wei H., Ruthenburg A. J., Bechis S. K., and Verdine G. L. (2005). *Nucleotide-dependent domain movement in the ATPase domain of a human type IIA DNA topoisomerase*. J Biol Chem 280, 37041-37047.

Wielckens K., George E., Pless T., and Hilz H. (1983). *Stimulation of poly(ADP-ribosylation) during Ehrlich ascites tumor cell "starvation" and suppression of concomitant DNA fragmentation by benzamide*. J Biol Chem 258, 4098-4104.

Wierstra I. (2008). *Sp1: emerging roles--beyond constitutive activation of TATA-less housekeeping genes*. Biochem Biophys Res Commun 372, 1-13.

Williams A. O., Isaacs R. J., and Stowell K. M. (2007). *Down-regulation of human topoisomerase IIalpha expression correlates with relative amounts of specificity factors Sp1 and Sp3 bound at proximal and distal promoter regions*. BMC Mol Biol 8, 36.

Williams A. O., Isaacs R. J., and Stowell K. M. (2007). *Down-regulation of human topoisomerase IIalpha expression correlates with relative amounts of specificity factors Sp1 and Sp3 bound at proximal and distal promoter regions*. BMC Mol Biol 8, 36.

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

Withoff S., Keith W. N., Knol A. J., Coutts J. C., Hoare S. F., Mulder N. H., and de Vries E. G. (1996). *Selection of a subpopulation with fewer DNA topoisomerase II alpha gene copies in a doxorubicin-resistant cell line panel*. Br J Cancer 74, 502-507.

Workman J. L. (2006). *Nucleosome displacement in transcription*. Genes & Dev 20, 2009-2017.

Wright K. L., Terry L. M., J., V. B., and M., B. A. (1995). *Major histocompatibility complex class II-associated invariant chain gene expression is up-regulated by cooperative interactions of Sp1 and NF-Y*. The Am Soc for Bio and Mol Bio 270, 20978-20986.

Wu C. H., van Riggelen J., Yetil A., Fan A. C., Bachireddy P., and Felsher D. W. (2007). *Cellular senescence is an important mechanism of tumor regression upon c-Myc inactivation*. Proc Natl Acad Sci U S A 104, 13028-13033.

Xu P., LaVallee P. A., Lin J. J., and Hoidal J. R. (2004). *Characterization of proteins binding to E-box/Ku86 sites and function of Ku86 in transcriptional regulation of the human xanthine oxidoreductase gene*. J Biol Chem 279, 16057-16063.

Yamada K., Tanaka T., Miyamoto K., and 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*. J Biol Chem 275, 18129-18137.

Yaneva M., Kowalewski T., and Lieber M. R. (1997). *Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies*. EMBO J 16, 5098-5112.

Yu B., Datta P. K., and Bagchi S. (2003). *Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites*. Nucleic Acids Res 31, 5368-5376.

Yun J., Tomida A., Nagata K., and Tsuruo T. (1995). *Glucose-regulated stresses confer resistance to VP-16 in human cancer cells through a decreased expression of DNA topoisomerase II*. Oncol Res 7, 583-590.

Yun J., Tomida A., Nagata K., and Tsuruo T. (1995). *Glucose-regulated stresses confer resistance to VP-16 in human cancer cells through a decreased expression of DNA topoisomerase II*. *Oncol Res* 7, 583-590.

Yung T. M., Sato S., and Satoh M. S. (2004). *Poly(ADP-ribosyl)ation as a DNA damage-induced post-translational modification regulating poly(ADP-ribose) polymerase-1-topoisomerase I interaction*. *J Biol Chem* 279, 39686-39696.

Zandvliet D. W., Hanby A. M., Austin C. A., Marsh K. L., Clark I. B., Wright N. A., and Poulson R. (1996). *Analysis of foetal expression sites of human type II DNA topoisomerase alpha and beta mRNAs by in situ hybridisation*. *Biochim Biophys Acta* 1307, 239-247.

Zaniolo K., Desnoyers S., Leclerc S., and Guerin S. L. (2007). *Regulation of poly(ADP-ribose) polymerase-1 (PARP-1) gene expression through the post-translational modification of Sp1: a nuclear target protein of PARP-1*. *BMC Mol Biol* 8, 96.

Zannetti A., Del Vecchio S., Carriero M. V., Fonti R., Franco P., Botti G., D'Aiuto G., Stoppelli M. P., and Salvatore M. (2000). *Coordinate up-regulation of Sp1 DNA-binding activity and urokinase receptor expression in breast carcinoma*. *Cancer Res* 60, 1546-1551.

Zhao G., Shi L., Qiu D., Hu H., and Kao P. N. (2005). *NF45/ILF2 tissue expression promoter analysis and interleukin-2 transactivating function*. *Exp Cell Res* 305, 312-323.

# Appendix 1

Drug	Classification	Mechanism	Notes	Reference
Tamoxifen	Selective modulator of estrogen receptor (ER). Hormonal therapy	Inhibits binding of estradiol to regulate cell replication and proliferation. Selectively used in ER+ tumours	Works best with chemotherapy	Dinh <i>et al</i> , 2007
Herceptin or trastuzumab	Monoclonal antibody	Binds to extracellular domain of Her-2 receptor Selectively used in Her-2+ breast cancer	Works best with anthracyclines but increases cardiotoxicity	Mano <i>et al</i> , 2007
Amsacrine	Topoisomerase II-poison (intercalating)	Stabilise ternary complex drug/DNA/topoisomerase II	Breast, leukemia, oral carcinoma, small lung cancer	Capranico and Binaschi, 1998
Etoposide or VP16	Topoisomerase II-poison	Stabilise ternary complex drug/DNA/topoisomerase II	Lung carcinoma, melanoma, oral carcinoma	Matsuo <i>et al</i> , 1993
Doxorubicin or adriamycin (anthracycline)	Topoisomerase II-poison (intercalating)	Stabilise ternary complex drug/DNA/topoisomerase II	Treatment of solid tumours	Capranico and Binaschi, 1998
Teniposide or VM26	Topoisomerase II-poison	Stabilise ternary complex drug/DNA/topoisomerase II	Lung carcinoma, leukemia	Matsuo <i>et al</i> , 1993
Epirubicin (anthracycline)	Topoisomerase II-poison (intercalating)	Stabilise ternary complex drug/DNA/topoisomerase II	Breast, lung, ovarian, gastric, soft tissue sarcoma. Less cardiotoxic than doxorubicin	Geffen and Man, 2002
Daunorubicin (anthracycline)	Topoisomerase II-inhibitor (intercalating)	Binds to DNA and can also inhibit transcription	Primarily used in treatment of acute myeloid leukaemia	Mansilla <i>et al</i> , 2003
Aclarubicin (anthracycline)	Topoisomerase II-inhibitor (intercalating)	Intercalates DNA and prevents topoisomerase II $\alpha$ from binding to DNA.	Acute myelocytic leukemia. Antagonistic to a topoisomerase II-poison	Larsen <i>et al</i> , 2003
Bisdioxopiperazines (ICRF-159 and ICRF-193)	Topoisomerase II-inhibitor	Locks topoisomerase II $\alpha$ in a closed clamp formation with DNA. Blocks ATP binding/hydrolysis so topoisomerase II $\alpha$ can not cleave DNA.	Can act as a cardioprotectant	Jensen <i>et al</i> , 2000 Chéne <i>et al</i> , 2009
Novobiocin	Topoisomerase II-inhibitor	Prevents ATP from binding to topoisomerase II $\alpha$ .	Can potentiate the cytotoxic activity of etoposide and teniposide	Larsen <i>et al</i> , 2003

## Appendix 1: Summary of some commonly used chemotherapy agents.

The table above summarises some of the topoisomerase II-targeting agents used for adjuvant chemotherapy. The first two treatments in the list (highlighted in grey) are not chemotherapy agents they are specific treatments as described.

## Appendix 2

Alterations in human topoisomerase II $\alpha$ gene	Cell line and drug	Human topoisomerase II $\alpha$ promoter or protein activity	Notes	Reference
N/A	Human leukemic cell line (CEM) resistant and sensitive to bisdioxopiperazines.	Used topoisomerase II $\alpha$ minimal promoter to look at differences in transcription in these cells. Resistance to drugs caused an increase in topoisomerase II $\alpha$ promoter activity.	Proteins that bind ICB3 appear reduced in resistant cells, but not due to a reduction in NF-Y	Morgan and Beck, 2001
Not tested	Resistance to topoisomerase II inhibitor (merbarone) CEM cell lines	Reduced topoisomerase II $\alpha$ mRNA and protein	Reduced expression of Sp3	Mo <i>et al</i> , 1997
Not tested	Resistance to etoposide in human epidermoid cancer (KB) cell line	Decreased topoisomerase II $\alpha$ promoter activity and mRNA in resistant lines. No change in mRNA stability to wt cell lines	Increased expression of Sp3	Kubo <i>et al</i> , 1995
Not tested	Acquired resistance to doxorubicin in breast cancer cell line (MDA-MB-231)	Decrease in topoisomerase II $\alpha$ protein. Concomitant decrease in Sp1 and increase in NF-Y	Discordant results not all cells exhibited changes	Allen <i>et al</i> , 2004
N/A	Used breast cancer cells (MDA MB 231) and exposed them to doxorubicin. Examined changes over 24 hours after exposure	After exposure to doxorubicin, ChIP assays used to show Sp3 preferentially binds to both GC1 and GC2 with concomitant decrease in cellular topoisomerase II $\alpha$ protein	Sp3 is up-regulated after drug exposure causing repression of Sp1-mediated activation of topoisomerase II $\alpha$ promoter	Williams <i>et al</i> , 2007
None detected	Acquired resistance to doxorubicin in human multiple myeloma (RPMI 8226)	Decreased topoisomerase II $\alpha$ protein activity (2.5fold lower than wt). The same fold decrease in amount of protein	Decreased binding to the ICB elements using extracts from resistance lines, possibly reduced NF-Y in the resistant cell lines	Wang <i>et al</i> , 1997B
Gene Mutation Arg486 $\rightarrow$ Lys	Resistance to etoposide and to a lesser extent to Amsacrine. Used yeast cultures and expressed human topoisomerase II $\alpha$ protein.	Have same catalytic activity as wild-type in absence of drug however less drug-induced DNA cleavage	Has little effect on susceptibility to doxorubicin	Patel <i>et al</i> , 2000
Gene Mutation Arg449 $\rightarrow$ Glu	Acquired resistance to teniposide in human leukemia cell line (CCRF-CEM)	Decreased levels of topoisomerase II catalytic activity. Decreased ability to form drug-stabilised complexes, decreased interaction with ATP	Did not mention if there were changes in the amount of topoisomerase II $\alpha$ protein in these cells.	Bugg <i>et al</i> , 1991
Gene Mutation Asp48 $\rightarrow$ Asn	Used yeast culture and expressed human topoisomerase II $\alpha$ mutant protein. No altered sensitivity to bisdioxopiperazines	Same catalytic ability as wild-type to relax negative supercoiled DNA	Increased DNA cleavage compared to wild type. Enhanced DNA binding	Walker <i>et al</i> , 2004

Alterations in human topoisomerase II $\alpha$ gene	Cell line and drug	Human topoisomerase II $\alpha$ promoter and protein activity	Notes	Reference
Gene Mutation Tyr805→Phe	Used yeast cells to express wt and mt human topoisomerase II $\alpha$ protein. Examined affects of bisdioxopiperazines	Both wt and mt have similar expression of protein. mt version lacks topoisomerase II $\alpha$ activity. The wt topoisomerase II $\alpha$ activity is dominant over the mt topoisomerase II $\alpha$ activity in heterodimers of this protein	Topoisomerase II $\alpha$ protein activity is necessary for sensitivity to bisdioxopiperazines	Jensen <i>et al</i> , 2000
None detected	Acquired resistance to mitoxantrone in human leukemia cell line (HL-60)	4.8 kb topoisomerase II $\alpha$ mRNA, smaller catalytically active topoisomerase II $\alpha$ protein localised to cytosol	Still retains 6.3 kb topoisomerase II $\alpha$ mRNA and normal protein, but at reduced levels	Harker <i>et al</i> , 1995
Deleted amino acids 1423-1531	Etoposide resistant lung cancer cell line (H209/VP)	Truncated topoisomerase II $\alpha$ protein missing C-terminal region, aberrant cytoplasmic localisation and reduced ability to form ternary cleavable complex with DNA/drug/enzyme	Normal topoisomerase II $\alpha$ protein is still found, but is 4 fold less than wild type.	Mirski <i>et al</i> , 2000
None detected in promoter other regions not tested	Stably transfected human bladder cancer cell lines with topoisomerase II $\alpha$ reporter gene constructs, then acquired resistance to etoposide in human bladder cancer cell lines	Reporter gene assays decreased topoisomerase II $\alpha$ expression in resistant cell lines which was abolished when a mutation in ICB1 was used in the topoisomerase II $\alpha$ promoter	Imply a negative regulator binds to ICB1 when cells become resistant to etoposide	Takano <i>et al</i> , 1999
Not tested	Confluence arrested breast cancer MCF7 cells	Decrease in topoisomerase II $\alpha$ mRNA and decrease in topoisomerase II $\alpha$ gene	Change in promoter activity is the primary determinant in decreased protein levels seen in confluent arrested cells	Isaacs <i>et al</i> , 1996
Topoisomerase II $\alpha$ gene amplification	Formalin fixed, paraffin-embedded tumour blocks from invasive breast tumours. Sample size 81	Topoisomerase II $\alpha$ gene amplification is not correlated to the amount of topoisomerase II $\alpha$ protein found in cells	Topoisomerase II $\alpha$ protein expression determined by IHC. Topoisomerase II $\alpha$ gene copy numbers determined by FISH	Mueller <i>et al</i> , 2004
N/A	Human cancer KB cells were enhanced to express heat shock protein HSP-70 under control of a thermosensitive promoter.	Levels of topoisomerase II $\alpha$ protein increased as HSP-70 levels increased 3-6 hours after heat shock treatment. Expression of topoisomerase II $\alpha$ protein was consistent with increases in mRNA levels	Increased catalytic activity of topoisomerase II $\alpha$ protein which correlated to increased sensitive to etoposide	Matsuo <i>et al</i> , 1993

**Appendix 2: Table summarising some of the studies done examining the expression of the human topoisomerase II $\alpha$  gene.**

The results of several studies are summarised in the above table to describe how changes in the topoisomerase II $\alpha$  gene and protein could influence chemotherapy sensitivity. Most of the resistant cell lines were created in these studies by repeated exposure to topoisomerase II-targeting drug, and selecting for the resistant cells.

## Appendix 3.

N-terminal amino acids for the 85 kDa protein

CYCLE #	MAJOR SIGNAL	MINOR SIGNAL	COMMENTS
1	(V)	D, G, S	( ) = tentative call
2	L		
3	-		
4	G		
5	N		
6	(K)		
7	A		
8	(V, A)		
9	(V)		
10	V		
11	-		
12	-		

The results from N-terminal sequencing (APAF) were submitted through BLAST (<http://www.ncbi.nih.gov/BLAST/>) limiting the search for proteins of “short nearly exact matches” and restricting the parameters to Homo sapiens only. The matched amino acids are highlighted in red (below). Protein identity was subsequently confirmed by immunoblot.

Ku86 (Homo sapiens) protein sequence.

```

1  mvrsgnkaav vlcmdvgftm snsipgiesp feqakkvitm fvqrqvfaen kdeialvlfq
61  tdgtdnplsg gdqyqnitvh rhmlpdpfdl ledieskiqp gsqqadfla livsmdviqh
121 etigkkfekr hieiftdlss rfsksqldii ihslkkcdis lqfflpfslg kedgsgdrqd
181 gpfrlgghgp sfplkgiteq qkegleivkm vmislegedg ldeiysfses lrklcvfkki
241 erhsihwpcr ltigsnlisir iaayksilqe rvkktwtvvd aktlkkediq ketvyclndd
301 detevlkedi iqgfrygsdi vpfskvdeeq mkyksegkcf svlgfckssq vqrrffmgng
361 vlkvfaardd eaaavalssl ihalddldmv aivryaydkr anpqvgvafp hikhnyeclv
421 yvqlpfmedl rqymfsslkn skkyapteaq lnavdalids mslakkdekt dtledlfptt
481 kipnprfqrl fqcllhralh preplppiqa hiwnmlnppa evttksqipl skiktlfpli
541 eakkkdqvta qeifqdnhed gptakklkte qggahfsvss laegsvtsvg svnpaenfrv
601 lvkqkkasfe easnqlinhi eqfldtnetp yfmksidcir afreeaikfs eeqrfnflk
661 alqekveikq lnhfweivvq dgitalitkee asgssvtaee akkflapkd psgdaavfe
721 eggdvddlld mi

```

## N-terminal amino acids for the 115 kDa protein

CYCLE #	MAJOR SIGNAL	MINOR SIGNAL	COMMENTS
1	(G)	M	( ) = tentative call
2	A		
3	E	F	
4	S, Y	A, Y	
5	S		
6	-		
7	K		
8	(L)	V, A	
9	E, Y		
10	R-		
11	(V, A)	F	
12	-		
13	-		

The results from N-terminal sequencing (APAF) were submitted through BLAST (<http://www.ncbi.nih.gov/BLAST/>) limiting the search for proteins of “short nearly exact matches” and restricting the parameters to Homo sapiens only. The matched amino acids are highlighted in red (below). Protein identity was subsequently confirmed by immunoblot.

### PARP-1 (Homo sapiens) protein sequence.

```

1  MAESSDKLYR VEYAKSGRAS  CKKCSSESIPK  DSLRMAIMVQ  SPMFDGKVPH  WYHFSCFWKV
61  GHSIRHPDVE  VDFGSELRWD  DQQKVKKTAE  AGGVTGKGQD  GIGSKAEKTL  GDFAAEYAKS
121 NRSTCKGCME  KIEKGQVRLS  KKMVDPEKPQ  LGMIDRWYHP  GCFVKNREEL  GFRPEYSASQ
181  LKGFSLATE  DKEALKKQLP  GVKSEGKRRG  DEVDGVDEVA  KKKSKKEKDK  DSKLEKALKA
241  QNDLIWNIKD  ELKKVCSTND  LKELLIFNKQ  QVPSGESAIL  DRVADGMVFG  ALLPCEECSSG
301  QLVFKSDAYY  CTGDVTAWTK  CMVKTQTPNR  KEWVTPKEFR  EISYLKCLKV  KKQDRIFPPE
361  TSASVAATPP  PSTASAPAAV  NSSASADKPL  SNMKILTLGK  LSRNKDEVKA  MIEKLGKLT
421  GTANKASLCI  STKKEVEKMN  KKMEEVKEAN  IRVVEDFLQ  DVSASTKSLQ  ELFLAHILSP
481  WGAEVKAEPV  EVVAPRGKSG  AALSCKSKGQ  VKEEGINKSE  KRMKLTLLGG  AAVDPDSGLE
541  HSAHVLEKGG  KVFSATLGLV  DIVKGTNSYY  KLQLEDDKE  NRYWIFRSWG  RVGTVIGSNK
601  LEQMPSKEDA  IEQFMKLYEE  KTGNAWHSKN  FTKYPKFFYP  LEIDYGQDEE  AVKCLTVNPG
661  TSKLPPKPVQ  DLIKMFVDE  SMKKAMVEYE  IDLQKMPLGK  LSKRQIQAAAY  SILSEVQQAV
721  SQGSSDSQIL  DLSNRFYTLI  PHDFGMKKPP  LLNNADSVQA  KVEMLDNLLD  IEVAYSLLRG
781  GSDDSSKDPI  DVNYEKLKTD  IKVVDRDSEE  AEIIRKYVKN  THATTHSAYD  LEVIDIFKIE
841  REGECQRYKP  FKQLHNRRL  WHGSRRTNFA  GILSQGLRIA  PPEAPVTGYM  FGKGIYFADM
901  VSKSANYYHT  SQGDPIGLIL  LGEVALGNMY  ELKHASHISR  LPKKGHSVKG  LGKTTDPDSA
961  NISLDGVDVP  LGTGISSGVI  DTSLLYNEYI  VYDIAQVNLK  YLLKLFNFK  TSLW

```

# Appendix 4

## Summary of the results for MS/MS analysis

Size of band excised	Protein identity	Alias	Estimated molecular weight of protein	Sequence coverage	Mowse score
250 kDa	Myosin-9		226 kDa	29%	160
150 kDa	SF3B1	PRP10/SAP155	145 kDa	25%	108
145 kDa	DXH9	RHA/NDH2/DDX9	140 kDa	29%	132
120 kDa	No ID				
105 kDa	No ID				
98 kDa	PSF	SFPQ/POMP100/hPOMp100	100 kDa	35%	301
95 kDa	PSF	SFPQ/POMP100/hPOMp100	100 kDa	41%	156
89 kDa	DHX15	DBP1/DEADH box15/HRH2/RNA helicase 2	93 kDa	28%	110
87 kDa	NF90	ILF3/NFAR/TCP80/NF-AT-90/MPP4	90 kDa	23%	166

The final data was submitted to the database search program Mascot (Matrix Science Ltd, London, UK) and the best match(es) were provided by the operator from Australian Proteome Analysis Facility (APAF), Sydney, Australia.