Transcription Factor AP-2 Regulatory Signatures in Breast Cancer

Williams, Christopher MJ

Transcription Factor AP-2 Regulatory Signatures in Breast Cancer

Christopher M J Williams

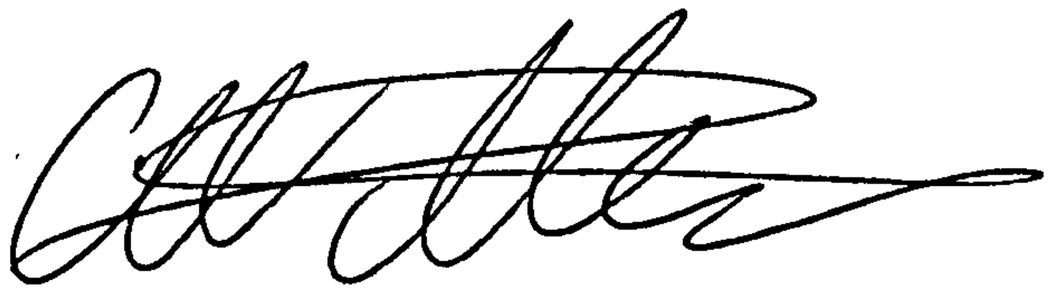
PhD Thesis

CANCER RESEARCH UK

Barts and The London
Queen Mary's School of Medicine and Dentistry

DECLARATION

I declare that the material presented in this thesis is my own, unless where appropriately stated. The work was carried out between October 2003 and May 2007 in the Gene Transcription Laboratory at the Cancer Research UK Clinical Centre, Molecular Oncology Unit, Barts and The London, Queen Mary's School of Medicine and Dentistry.



Christopher M J Williams

May 2007

ABSTRACT

AP-2 transcription factors are highly conserved basic helix-span-helix proteins whose members (α, β, γ, δ and ε) are crucial regulators of embryonic development. They also play an important role in human neoplasia. Immuohistochemical studies have detected high levels of AP-2γ expression in primary tumours of breast cancer patients. This high expression has been correlated with reduced survival in all patients and reduced survival in an ERα positive subset treated with hormone therapy. In breast cancer cell lines, AP-2 factors have been implicated in the regulation of the *ERBB2* proto-oncogene and ERα. In an effort to further understand the role of AP-2γ in breast carcinoma, this study has sought to identify additional AP-2 activated cellular pathways and ultimately novel transcriptional targets for AP-2 through the use of gene expression profiling.

RNAi using three independent AP-2γ targeting sequences, has been used to deplete AP-2γ levels in the ERα positive MCF-7 breast carcinoma cell line, chosen as it exclusively expresses the AP-2γ family member. Microarrays were then utilised to create an AP-2γ dependent transcription profile. Statistical comparisons between non-silencing control siRNA and AP-2γ targeting siRNA groups identified a total of 162 gene expression changes ($p<0.01$). These changes implicate AP-2γ in the control of cell cycle progression and developmental signalling. Indeed a role for AP-2γ in the control of cell cycle, in particular at the G1/S transition, has been verified using flow cytometry. Several of these gene expression changes, including IGFBP3, Transgelin and KIAA1324, have been confirmed using qPCR and immunoblotting.

Finally, elevated levels of p21 mRNA and protein have been observed following AP-2γ silencing in MCF-7 cells. Additionally, the activity of a *p21* promoter reporter is repressed following transfection with an AP-2γ expression construct in HepG2 cells. These results coupled with ChIP experiments showing AP-2γ occupancy at the proximal promoter region of *p21* in cycling MCF-7 cells, implicate AP-2γ in the repression of *p21* transcription and suggest a role for AP2γ in the control of cell cycle in breast carcinoma in part through the transcriptional repression of *p21*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Ac	Acetylation
AP-2	Activator Protein-2
BLAST	Basic Logical Alignment Search Tool
BrdU	Bromodeoxyuridine
CBP	CREB Binding Protein
CDK	Cyclin Dependant Kinase
CDKI	Cyclin Dependant Kinase Inhibitor
CH1	Cysteine-, Histidine-Rich Domain 1
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
ChIP	Chromatin Immunoprecipitation
CITED	CBP/p300 Interacting Transactivator with ED-Rich Tail
CMV	Cauliflower Mosaic Virus
CREB	cAMP Regulatory Element-Binding Protein
cRNA	"copy" RNA
CRUK	Cancer Research UK
DAPI	4', 6-diamino-2-phenylindole, dilactate
DCIS	Ductal Carcinoma *in situ*
DMEM	Dubellcos Modified Eagles Medium
DMSO	Dimethyl Sulphoxide
dpc	Days Post Coitum
dsRNA	double stranded RNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth Factor Receptor
EMSA	Electromobility Shift Assay
ERα	Estrogen Receptor α
EST	Expressed Sequence Tag
FACS	Fluorescent Activated Cell Sorting
FCS	Foetal Calf Serum
FDR	False Discovery Rate
GFP	Green Fluorescent Protein
GO	Gene Ontology
GST	Glutathione-S-Transferase
H4	Histone 4
HAT	Histone Acetyl Transferase
HDAC	Histone Deacetyltase
HRP	Horse Radish Peroxidase

IGFBP	Insulin-like Growth Factor Binding Protein
IP	Immunoprecipitation
MEF	Mouse Embryonic Fibroblast
miRNA	micro RNA
MM	mismatch
MMP	Matrix Metalloproteinase
MOPS	morpholinepropanesulfonic acid
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NFAT	nuclear factor of activated T-cells
PAGE	Poly Acrylimide Gel Electrophoresis
PI	Propidium Iodide
PIC	Pre Initiation Complex
PIPES	1,4-piperazinediethanesulfonic acid
PM	perfect match
Pol	Polymerase
PSA	Protein A Sepharose
qPCR	qPCR
RA	Retinoic Acid
Rb	Retinoblastoma
RMA	Robust Multi-array Average
RNAi	RNA Interference
RNase	Ribonuclease
RSV	Respiratory syncytial virus
RT	Room Temperature
SDS	Sodium Dodecyl (lauryl) Sulphate
SERM	Selective oestrogen receptor modulator
shRNA	short hairpin RNA
siRNA	short interfering RNA
Sp	Specificity Protein
SUMO	Small Ubiquitin-Related Modifier
SV40	Simian Virus 40
TBP	TATA-box Binding Protein
TEB	Terminal End Bud
TFBS	Transcription Factor Binding Site
TNF	Tumour Necrosis Factor
url	Uniform resource locater
VEGF	Vascular endothelial growth factor

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge Helen Hurst for her excellent supervision. Without her support and encouragement none of this work would have been possible. Special thanks to Kartsen Friedrich for collaborating on the Chromatin Immunoprecipritation experiments presented in this thesis. Thanks to Claire Ibbitt for all her assistance and for managing the most organised lab in the world!! Thank you to Tracy Chaplin for her assistance with the microarray hybridisations and the members of the CRUK FACS Lab for their help with the flow cytometry. I would also like to thank various members of the Molecular Oncology Unit, both past and present, including Charlotte Moss, Amelie Morin, Jamie Meredith, Safia Ali, Sarah Jezzard, Kayi Chan, Cheng Yeoh, Pompa Bhattacharya and Zoe Leech.

Finally, I would like to thank Jodi. She is surely the most gorgeous and amazing person I could ever have met and has supported me through all of the ups and downs over the last three and half years.

CHAPTER 1: INTRODUCTION

Changes in the abundance and activity of specific transcription factors are fundamental to faithful execution of the processes which control growth and differentiation in the life of an organism. Of further interest and greater clinical importance is their involvement in growth related diseases such as cancer. Research presented herein will focus upon the AP-2 family of transcription factors and their involvement in a disease that causes the premature death of more than 14,000 women a year in the UK, breast cancer.

1.1. The AP-2 family of transcription factors

Activator protein-2 (AP-2) is the term given to a family of five related DNA-binding transcription factors: AP-2α, AP-2β, AP-2γ, AP-2δ and AP-2ε (Williams *et al.*, 1988; Moser *et al.*, 1995; Bosher *et al.*, 1995; Zhao *et al.*, 2001; Feng *et al.*, 2003). These closely related mammalian family members are transcribed from different genes and each has a molecular weight of approximately 50kDa. A generalised protein structure is outlined in Figure 1.1.

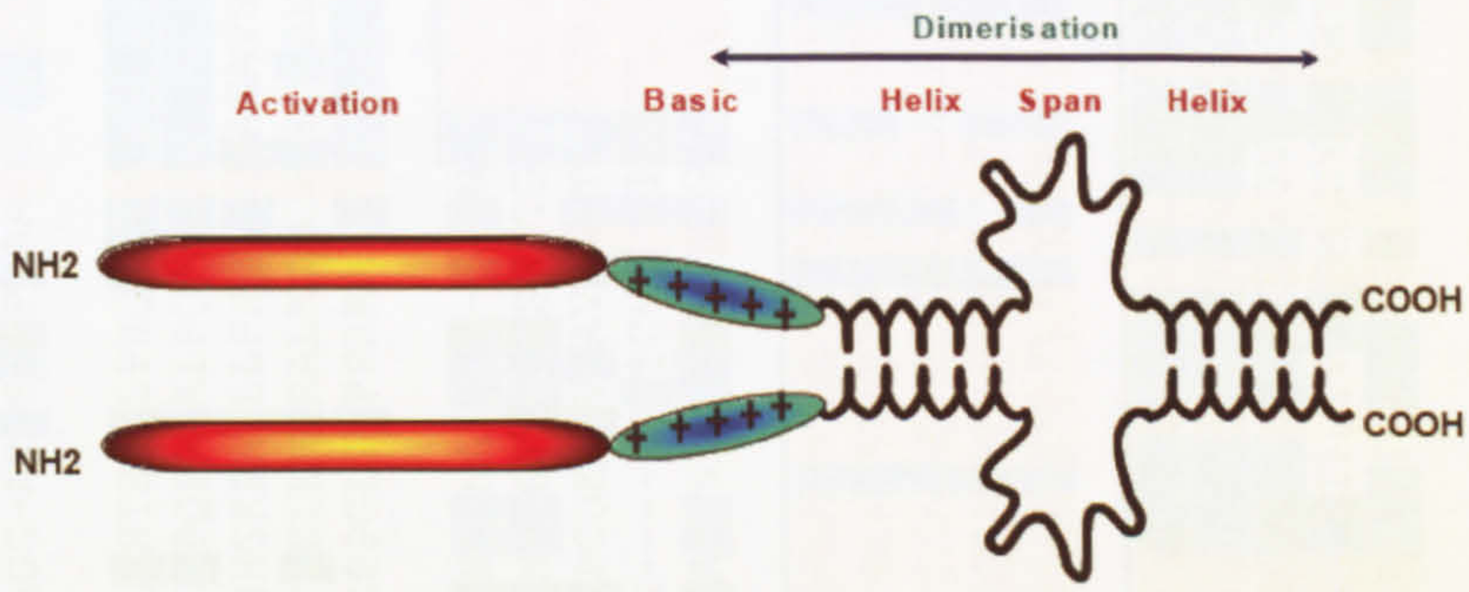


Figure 1.1. A representation of AP-2 protein structure. AP-2 proteins share highly conserved C-terminal basic DNA binding and helix-span-helix dimerisation domains. The less conserved proline rich transactivation domain is located toward the N-terminus. Detailed amino acid sequence alignments are shown in Figure1.2.

All AP-2 proteins can homo- and hetero- dimerise through a unique C-terminal helix-span-helix motif and bind DNA through a basic domain that lies immediately N-terminal of the dimerisation motif (Mohibullah *et al.*, 1999). A third less conserved region toward the N-terminus contains residues critical for modulation of AP-2 activity. The remarkable degree of amino acid conservation across these regions is shown for the human AP-2 family members in Figure 1.2. Very high amino acid sequence

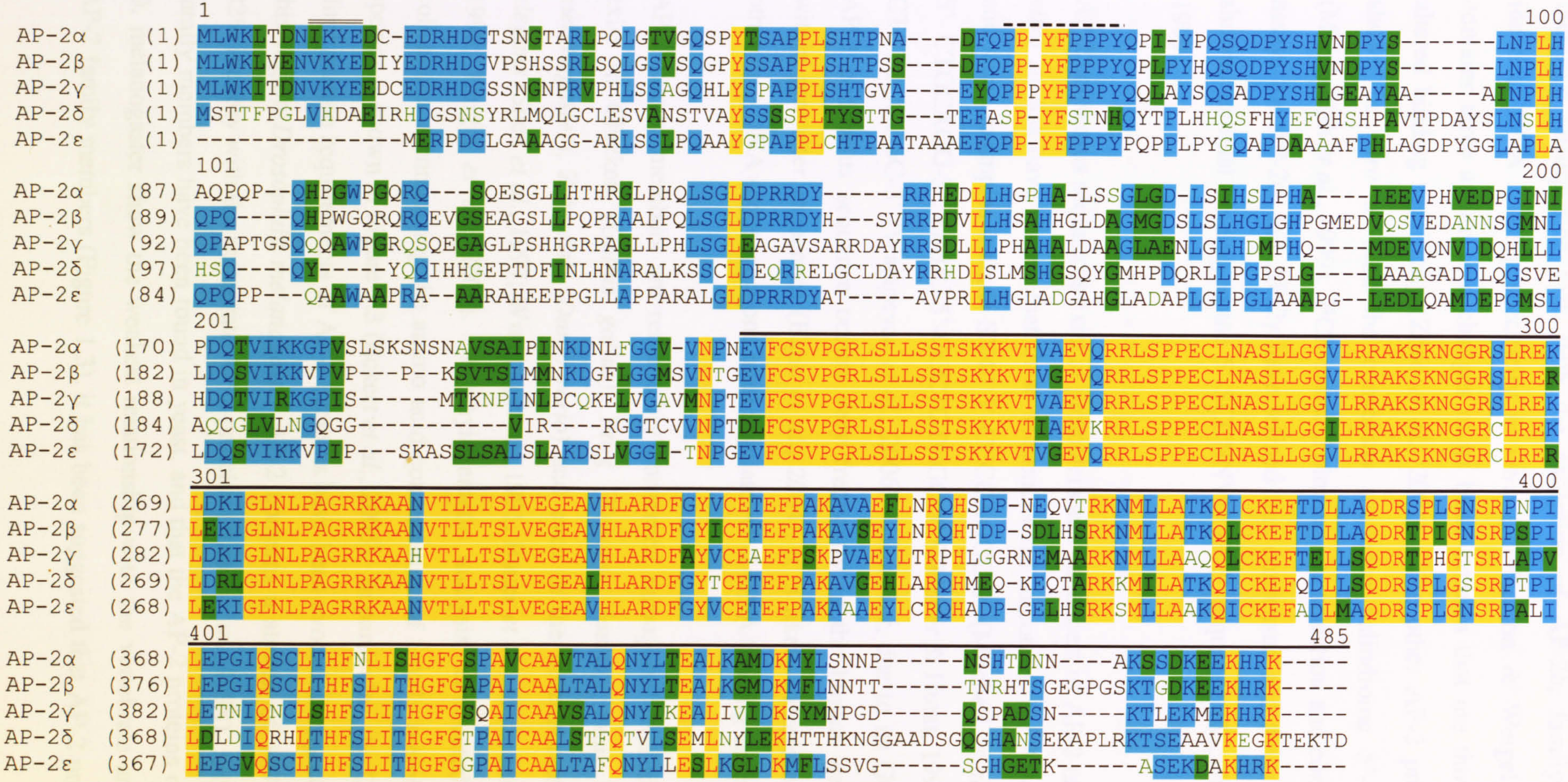


Figure 1.2 Amino acid conservation between human AP-2 family members. Amino acid sequence identity is shown by red text with yellow background. Conservative changes are indicated by blue text with cyan background. Amino acid sequence similarity is shown with a green background. A bold line indicates residues important for DNA binding and dimerisation, a double underline indicates residues important for sumo modification and dashed line indicates PY (xPPxY) motif. ClustalW amino acid sequence alignments and implemented in Vector NTI (Invitrogen).

conservation is observed across the DNA binding and dimerisation regions; it is therefore unsurprising that, with the exception of AP-2δ, family members show identical DNA binding specificity *in vitro* (McPherson & Weigel, 1999). AP-2δ contains three amino acid changes in its basic domain that are thought to cause an altered binding specificity (Zhao *et al.*, 2001). The other AP-2 proteins have been shown *in vitro* to bind the consensus sequence palindrome 5'-GCCN$_{3-4}$GGC-3' (Mohibullah *et al.*, 1999). PCR-assisted binding site selection and competitive gel shift assays for AP-2α and AP-2γ have refined this consensus and these family members showed optimal binding to the 5'-SCCNNVRGB-3' sequence (McPherson & Weigel, 1999).

AP-2 consensus sequences exist in promoters and enhancers of AP-2 target genes and *in vitro* studies have shown that these sequences are important for AP-2 transactivation and some examples include SV40 5'-CCCCAGGC-3' (Mitchell *et al.*, 1987), *ERBB2* 5'- CTGCAGG-3', *ERBB3* 5'-GCCTCTGGC-3' (Skinner & Hurst, 1993), and Bcl-2 5'-CTCTCCCCGC-3' (Wajapeyee *et al.*, 2006). However, due to its GC-rich nature the AP-2 consensus sequence occurs very frequently in the human genome especially within promoter sequences (Bajic *et al.*, 2004), it is likely therefore, that constraints other than DNA sequence alone must regulate AP-2 DNA binding specificity.

AP-2 family members are remarkably conserved throughout evolution, homologues exist in Chicken (*Gallus gallus*), Frog (*Xenopus laevis)*, Fruit Fly (*Drosophila melanogaster*), Zebrafish (*Danio rerio*) and even in the invertebrate *Caenorhabditis elegans* (Shen *et al.*, 1997; Winning *et al.*, 1991; Bauer *et al.*, 1998; Monge & Mitchell, 1998; Knight *et al.*, 2003). A phylogenetic analysis, conducted by Eckert and colleagues, summarising the amino acid conservation of family members across these species is shown in Figure 1.3 (Eckert *et al.*, 2005). In particular, the DNA-binding and dimerisation requirements of AP-2 proteins have been evolutionarily conserved. Within this region *Drosophila melanogaster* AP-2 contains a stretch of 107 amino showing 92% identity with the vertebrate AP-2 proteins (Bauer *et al.*, 1998). Given that no AP-2 family members have been found in yeast, and that the AP-2 proteins of C. elegans and D. melanogaster show the weakest relationship with the vertebrate and protochordate AP-2 family members (Figure 1.3), it has been suggested that AP-2 proteins originated

at this stage in evolution and subsequently expanded to give the five members known in higher species (Eckert *et al.*, 2005). In both human and mouse the genes are spatially located in a similar pattern on their respective chromosomes, shown in Table 1.1.

	AP-2α	**AP-2β**	**AP-2γ**	**AP-2δ**	**AP-2ε**
Human	*TFAP2A* 6p24	*TFAP2B* 6p12	*TFAP2C* 20q13.2	*TFAP2D* 6p12	*TFAP2E* 1p34.3
Mouse	*tfap2a* 13A5-B1	*tfap2b* 1 A2-4	*tfap2c* 2H3-4	*tfap2d* 1 A2-4	*tfap2e* 4D2-2

Table 1.1 Gene Names and chromosomal locations of each of the AP-2 family members in human and mouse. Table adapted from Feng & Williams, 2003.

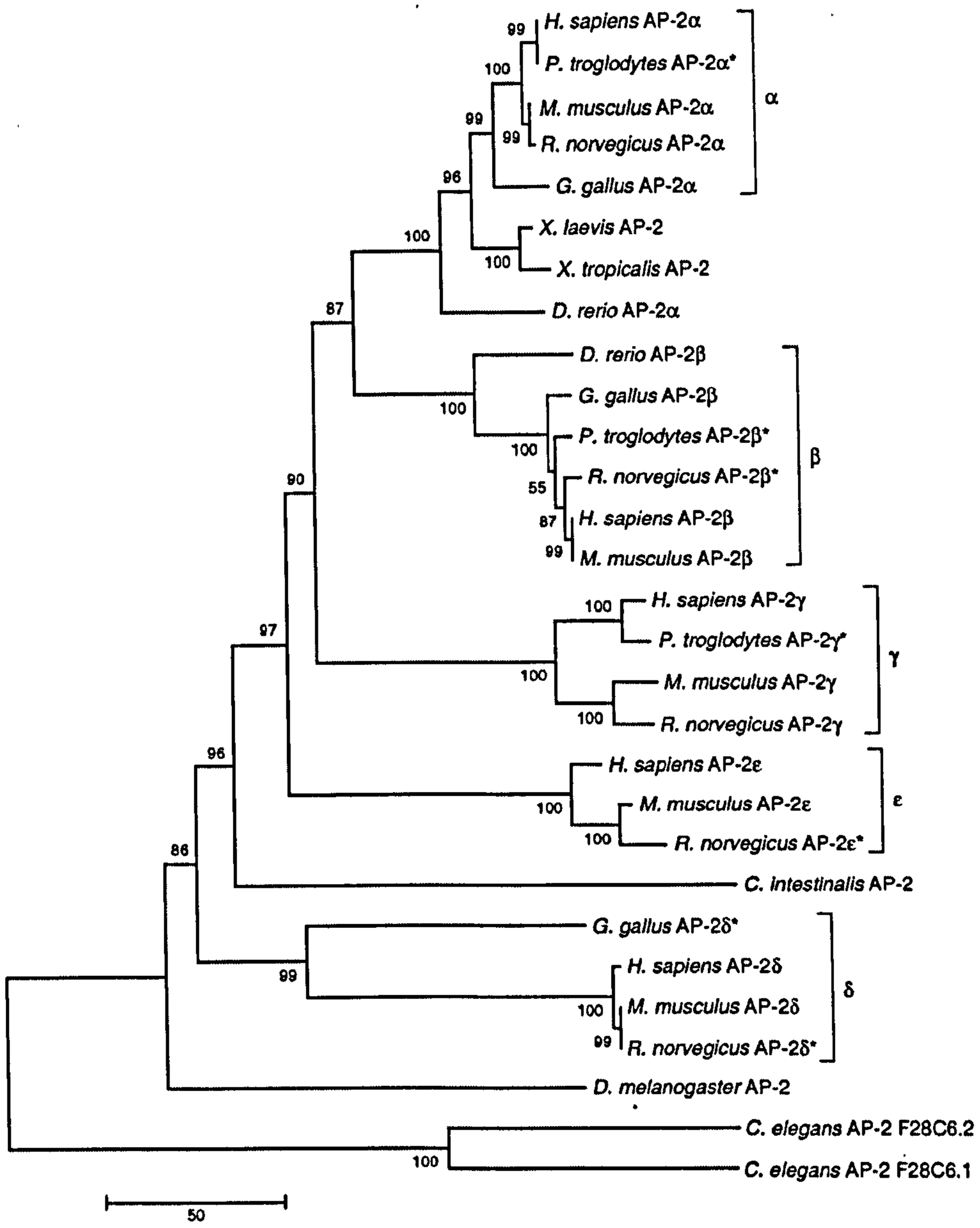


Figure 1.3. A phylogenetic tree of the AP-2 family (Adapted from Eckert *et al.*, 2005). This phylogenetic analysis was based on amino acid sequence alignments. A scale bar indicates the number of amino acid residue changes and asterisks indicate where predicted proteins have been included.

1.2. AP-2 and Gene Transcription

As described in the section above, AP-2 transcription factors can bind specific DNA sequences within promoter or enhancer regions of their target genes through their highly conserved basic domains. However, until relatively recently it was unclear how AP-2 factors were able to regulate gene transcription.

1.2.1. Mechanisms of Gene Transcription

In eukaryotes, genomic DNA exists as highly compacted repeating nucleosome arrays called chromatin. Each nucleosome is comprised of 147bp stretches of DNA wrapped around an octamer of the four basic histone proteins (H2A, H2B, H3 and H4). The linker histone (H1) and other non-histone proteins bind between nucleosomes and facilitate the higher organisation and further compaction of the chromatin fibre within the nucleus. The packaging of the DNA in this manner results in a stable chromatin state, where DNA is inaccessible to the transcriptional machinery. Thus, it has become apparent that chromatin structure forms the physiological substrate for all DNA dependent pathways and not just the DNA sequence itself. Chromatin has been shown to affect transcription through multiple mechanisms including: histone modification, chromatin remodelling, histone variant incorporation and histone eviction (reviewed by the following: Li *et al.*, 2007; Kouzarides, 2007; Bernstein *et al.*, 2007).

1.2.1.1. Histone modifications

The majority of histone modifications occur in the N-terminal tail domains which protrude from the nucleosome surface where modifications at approximately 60 different residues have been described to date. Some of the better understood include the acetylation (ac) and methylation (me) of lysine (K) residues. Acetylation of lysines on H3 and H4, facilitated by histone acetyltransferase (HAT) enzymes, is commonly associated with chromatin in the process of active transcription. Acetylation occurs on multiple lysines and is thought to act cumulatively to alter the net electrostatic charge of histones, through the neutralisation of the positive lysine charges. A consequence of this is a destabilisation of histone-DNA interactions and an increased propensity of

histones to be displaced by chromatin remodelling enzymes (described below), thus making the DNA more accessible for transcription. Histone deacetylatases (HDACs) can reverse this process and their activity is therefore often associated with transcriptional repression.

Histone modifications can also recruit specific non-histone proteins to chromatin. For example, Heterochromatin Protein 1 (HP1) is specifically recruited to methylated H3K9, via its chromodomain. HP1 binding is associated with deacetlyase activity leading to transcriptional repression. HP1 binding also facilitates further H3K9 methlyating activity and is thought to transmit H3K9me to adjacent nucleosomes and in this way is implied in the establishment and maintenance of a global heterochromatin environment (reviewed by Grewal & Jia, 2007).

1.2.1.2. Chromatin remodelling

In vitro studies have shown that nucleosomes exist in equilibrium between a fully wrapped state and a set of partially unwrapped states. The propensity of nucleosomes to rewrap is negatively regulated by ATP-dependant chromatin remodelling complexes, which actively alter the histone-DNA contacts. This activity is also necessary for displacement (or sliding) of nucleosomes away from conformationally stable positions on the DNA.

1.2.1.3. Transcription factor recruitment and Pre Initiation Complex (PIC) Assembly.

It is generally thought that spontaneous unwrapping of nucleosomal associated DNA is sufficient to reveal buried DNA consensus-binding sequences required for the recruitment of specific transcription factors like AP-2. Histone modifications like those discussed above affect the inclination of a nucleosome to remodel and thus specific transcription factors may only be able to access their DNA consensus binding sites given an appropriate chromatin context. Indeed, recently a genome-wide study showed that high levels of acH3 and H3K4/79me were prerequisite for recognition of any target site by the Myc transcription factor (Guccione *et al.*, 2006). Additional genome-wide studies in yeast have shown that nucleosome density at promoter regions is typically lower than that at adjacent coding regions (reviewed by Li *et al.*, 2007). This would

suggest that DNA consensus binding sequences within promoter regions are more amenable to specific transcription factors. Conversely, increased nucleosome density outside of promoter regions might serve to mask spurious recognition sequences that occur coincidently in large genomes.

Once bound to a promoter or enhancer, sequence specific transcription factors have been suggested to stimulate transcription via multiple mechanisms. Firstly, this might involve recruitment of coactivator complexes that include the HAT and chromatin remodelling activities described above. These facilitate the further binding of specific transcription factors but also act to adjust chromatin structure sufficiently to accommodate the recruitment and positioning of the PIC. Given a favourable chromatin environment, a second mechanism might involve transcription factors directly interacting with components the of PIC stimulating its positioning within the core promoter or indeed stimulating its activity once positioned.

1.2.2. Mechanisms for transcriptional regulation by AP-2 family members

1.2.2.1 AP-2, CITED and CBP/p300

AP-2 proteins were generally considered to be weak activators of transcription when analysed in reporter gene assays. However, it is now known that AP-2 proteins can be co-activated by the recruitment of the p300 and CREB-binding protein (CBP) to their target gene promoter and that this recruitment is mediated by the CITED (CBP/p300 interacting transactivators with ED-rich termini), family of proteins (Bamforth *et al.*, 2001; Braganca *et al.*, 2002; Braganca *et al.*, 2003).

CBP and p300 are two homologous proteins that were first identified through their ability to activate gene transcription (reviewed in Chan & La Thangue, 2001; Kalkhoven, 2004). Subsequent studies have shown that this activity is dependant on the coactivation of a wide variety of gene specific transcription factors, including AP-2 family members. CBP and p300 have been shown to interact with gene specific transcription factors, which then guide them to their target genes. Transcriptional activation by CBP and p300, involves multiple domains of these proteins (shown for

p300 in Figure 1.4). Once the target gene has been accessed the intrinsic HAT activity is required for the acetylation of target lysine residues on the histone tails. Within the region required for p300 HAT activity, two functionally important domains have been characterised. A PHD type zinc finger (CH2) is important for HAT activity whilst the preceding bromodomain (B) has been shown to tether p300 to chromatin through binding to acetylated lysines, which is thought to stabilise its interaction with chromatin. In addition to histones, CBP/p300 can also acetylate other non-histone chromatin associated proteins, including transcription factors. Interestingly, auto-acetylation has been shown to be important in controlling p300 activity. The basal catalytic activity of p300 was shown to be stimulated by autoacetylation of several key lysine sites within an activation loop motif (Thompson *et al.*, 2004). In addition to this activity, recent work has shown that the autoacetylation activity of p300 is required for ordered PIC assembly. It was shown that TFIID and p300 compete for binding to Mediator (a general coactivator complex) at promoters, with p300 autoacetylation tipping the balance in favour of TFIID (reviewed by Pugh, 2006). TFIID is a general transcription factor which, in cooperation with Mediator directs rapid and efficient PIC assembly. Both CBP and p300 have the ability to bind TBP and TFIIB and/or form a complex with RNA pol II, through N- and C-Terminal activation domains (reviewed in Kalkhoven 2004). In summary, it has been suggested that CBP and p300 form a bridge between gene specific transcription factors and the general transcription machinery. Their intrinsic HAT activity is important for establishing a favourable chromatin environment for both specific and general transcription machinery, indeed p300 can interact with components of this machinery and this interaction is important for the subsequent ordered assembly of the PIC.

The majority of DNA binding transcription factors that bind CBP/p300 do so by direct interaction. AP-2 factors however need an adapter protein, CITED. AP-2 transcription factors were first linked to CITED proteins after the observation that the embryonic lethal phenotype of CITED2 knockout mice displayed cardiac malformation and neural crest defects remarkably similar to those seen in AP-2α null mice (described below) and indeed those seen in the human condition Rubinstein-Taybi syndrome caused by *CBP* mutations (Bamforth *et al.*, 2001). In parallel, a Far-Western screen identified CITED2 as an interacting partner for AP-2α, AP-2β and AP-2γ (Bamforth, *et al.* 2001). In order

to determine the functional significance of the CITED2/AP-2 interaction, the reporter activity of an AP-2 dependant reporter construct (p3XAP2-Bluc) was compared in *Cited2*+/+ or *Cited2*-/- mouse embryonic fibroblast cells following transfection of AP-2α, AP-2β or AP-2γ. A considerable decrease in reporter activity was observed in the CITED2 null cells and this decrease could not be rescued by increasing levels AP-2 family member expression. However, expression of exogenous CITED2 was able to restore the activating potential of AP-2 (Bamforth, *et al.* 2001). Subsequent work has shown that CITED4 can also potentate the activation of AP-2 reporter constructs (Braganca *et al.*, 2002; Braganca *et al.*, 2003). These studies suggested that CITED proteins might act as adapter proteins bringing CBP/p300 to AP-2 factors and thus allowing co-activation of AP-2 target genes. Indeed all three proteins were immunoprecipitated together from U2-OS cells, indicating that they interact with each other *in vivo*. The nature of the AP-2, CITED and CBP/p300 interaction has been further characterised to understand how these proteins interact to mediate transcription activation and is summarised in Figure 1.4 (Braganca *et al.*, 2003).

Recent Chromatin Immuno Precipitation (ChIP) experiments in our lab, conducted by Karsten Friedrich, have shown that *ERBB2* transcriptional activation by endogenous AP-2γ is accompanied by an increase in AP-2γ occupancy of the *ERBB2* enhancer region and concurrently, an increase in occupancy of the AP-2 associated transcriptional machinery: CITED2 and CBP/p300. These data suggests that AP-2γ can also recruit these factors *in vivo* (Helen Hurst, *personal communication*).

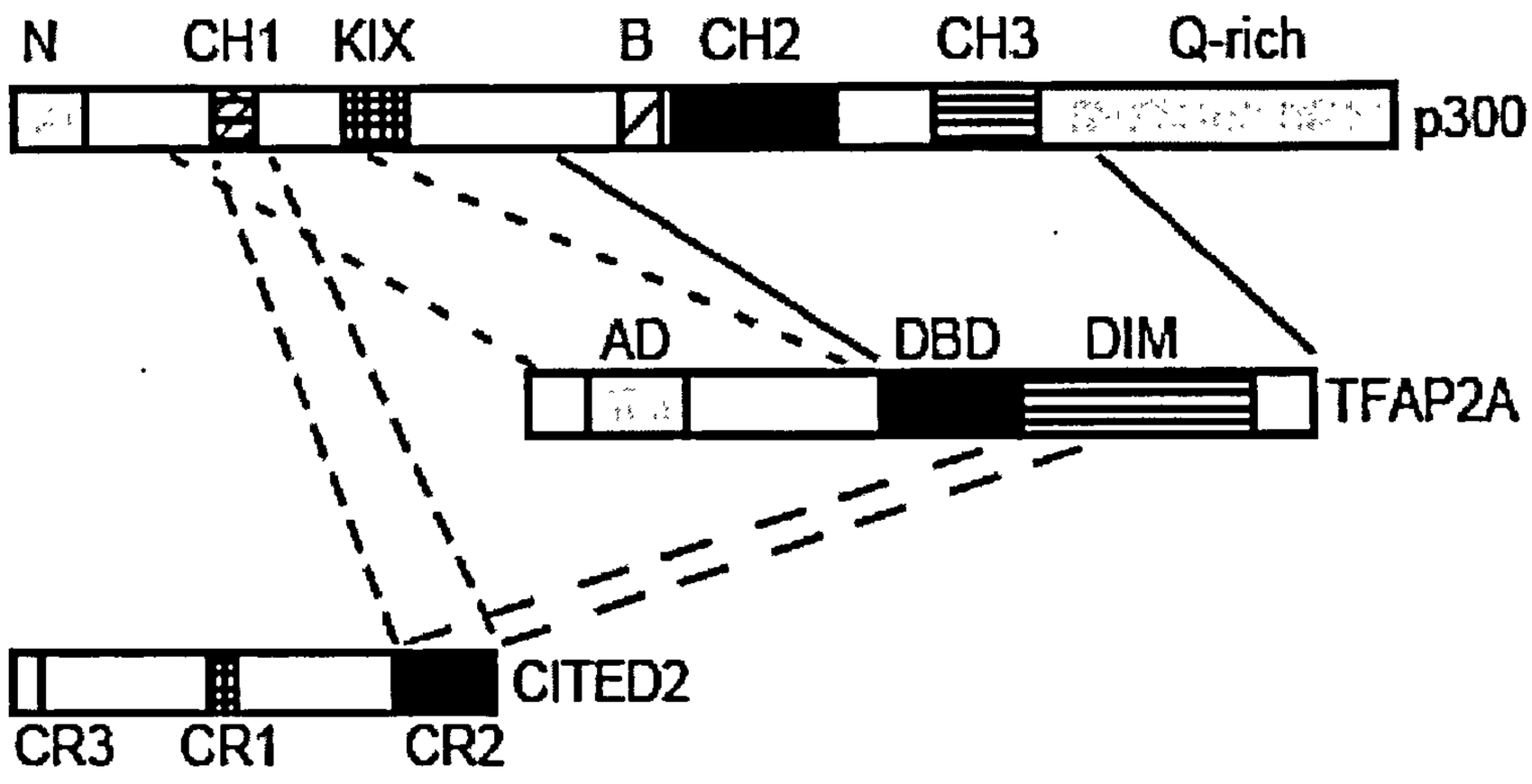


Figure 1.4. Interactions between AP-2, CITED and CBP/p300 (from Braganca *et al.*, 2003). The interactions between AP-2, CITED and CBP/p300 are have been determined using co-immunoprecipitation and yeast two-hybrid techniques (Bamforth *et al.*, 2001; Braganca *et al.*, 2002; Braganca *et al.*, 2003). CoIP experiments have demonstrated that AP-2 family members, CITED2 and p300 can interact physically *in vivo*. This interaction required the presence of the CR2 domain of CITED2 and an intact binding site on p300 for CITED proteins (the CH1 domain). CITED2 interacts with the residues contained within the first dimerisation helix of AP-2α and AP-2γ (DIM). AP-2α was also shown to interact with p300 independently of CITED2, via the N-terminal residues that make up its activation domain. Therefore, it was suggested that that CITED proteins act as a bridge between AP-2 family members and p300 permitting their proximity and thus allowing them to interact via the N-terminal domain of AP-2. Evidence suggests that some of the AP-2/p300 interaction may be direct, although the N terminus of AP-2 has also been shown to recruit other co-activators and these could also be important for this interaction. Finally a HAT deficient p300 mutant (D1399Y) was unable to activate full length AP-2 mediated transcription despite being able to interact with CITED2 suggesting that this activity is necessary for AP-2 transcriptional activation.

Protein domains are labelled as follows:

p300: CH1-3, Cysteine- Histidine-rich 1-3; B, Bromodomain; Q-Rich, Glutamine-rich.

AP-2: AD, Activation Domain; DBD, DNA Binding Domain; DIM, Dimerisation Domain.

CITED: CR1-3, Conserved Region 1-3.

Several other proteins have been reported to interact with AP-2 factors and the more intriguing are summarised below. However, many of these are highly abundant proteins and the functional significance of their interaction with AP-2 remains to be established. The AP-2, CITED, p300 interaction is therefore the only example with a genetically verified functional role.

1.2.2.2 AP-2, PC4 and PARP1

Work by Kannan and Tainsky has identified two further potential transcriptional cofactors of AP-2 using a GST-AP-2 fusion protein to identify AP-2α interacting proteins in human teratocarcinoma cells (Kannan & Tainsky, 1999). One of these PC4 was shown to interact with the N-terminal domain of AP-2 in yeast two hybrid assays and increase transcriptional activation in luciferase reporter assays. The second AP-2 interacting protein identified in this study was poly-ADP-ribose polymerase 1 or PARP-1 (Kannan *et al.*, 1999). PARP-1 was shown to interact with the C-terminal domains of AP-2α and was also implicated in coactivation of AP-2α in luciferase reporter assays. Although this study requires further validation, the interaction of these cofactors with AP-2 provides evidence for additional mechanisms that AP-2 family members may employ to activate transcription. PC4 is thought to promote transcription through direct interactions with components of the general transcription machinery (Fukuda *et al.*, 2004), whereas PARP-1 plays multiple and diverse roles in transcription, including the ADP-ribosylation of core histones and chromatin associated proteins, which acts to promote chromatin decondensation and nucleosome dissociation (Reviewed in Schreiber *et al.*, 2006).

1.2.2.3 AP-2 and GAS41

A recent study has identified a physical and functional interaction of AP-2β with an additional known transcriptional coactivator, GAS41 (also known as YEATS4) (Ding *et al.*, 2006). An *in vivo* interaction was demonstrated by their coimmunoprecipition in HeLa cells and GST pull downs showed that the C terminal portion, encompassing the DNA binding and dimerisation domains of AP-2β, was required for this interaction *in vitro*. Although the functional significance of this interaction requires further elaboration, in cotransfection experiments GAS41 acted to enhance transcriptional

activation by AP-2β in reporter gene assays (Ding *et al.*, 2006). Interestingly, GAS41 has been shown to interact with INI1 a component of the human SWI/SNF chromatin remodelling complex (Debernardi *et al.*, 2002), suggesting that via GAS41, AP-2β could function by recruiting these complexes to enhance gene transcription.

1.2.2.4 AP-2 and SUMO modification

As well as activating gene transcription, AP-2 family members have also been reported to act as repressors of transcription, notably AP-2α represses EGFR expression during epidermal differentiation in mice (Wang *et al.*, 2006) and AP-2α represses transcriptional of Bcl-2 in colon carcinoma cells (Wajapeyee *et al.*, 2006).

It is thought that the conjugation of the small ubiquitin-like modifier, SUMO to AP-2 proteins might lead to such modification of transcriptional activity. Indeed the AP-2α, AP-2β and AP-2γ family members have been shown to be modified by SUMO polypeptides. Work in our lab has shown that AP-2α, AP-2β and AP-2γ can interact with the E2 SUMO conjugating enzyme UBC9. This interaction results in SUMO conjugation at the conserved lysine residue (lys 10) in the N-terminal region of AP-2, and was shown to reduce the ability of AP-2 to activate transcription when analysed in reporter gene assays (Eloranta *et al.*, 2002). This lysine is part of a SUMO modification motif φKXE (where φ is any hydrophobic amino acid), conserved between AP-2α, AP-2β and AP-2γ (shown in Figure 1.2.). The transcriptional mechanism of this effect is at present uncertain but SUMO modification of other transcription factors is known to modulate their function. In particular, SUMO modification of CBP and p300 allows them to recruit histone deacetylases which lead to transcriptional repression (Girdwood, *et al.* 2003). Whilst sumolation reduces the activating potential of AP-2, there are no reports that this modification alters the function of AP-2 to a transcriptional repressor, such that it recruits known co-repressors or repression complexes.

1.2.2.5 AP-2 and WW domain containing proteins

Another mechanism of regulating the transcriptional potential of AP-2 family members is thought to be through the interaction with WW domain containing proteins. With the exception of AP-2δ, AP-2 family members possess the highly conserved PY (XPPXY)

motif required for such interactions, these residues are marked by a dotted underline in Figure 1.2. Mutation of the PY motif from FPPPY to FAPPY reduced the transactivating potential of AP-2α by approximately 35% (Wankhade *et al.* 2000). Early work showed that the WW domain containing YES associated protein (YAP) was able to interact with peptides corresponding to the PY motif of AP-2α (Yagi *et al.*, 1999). More recently, the WW domain containing protein Wwox was shown to directly interact with AP2γ via the highly conserved PY motif. Interestingly, cotransfection of AP-2γ and Wwox reduced the ability of AP-2γ to activate transcription in reporter assays, and confocal microscopy suggested that the AP-2γ-Wwox interaction facilitated the redistribution of AP-2γ to the cytoplasm limiting its transcriptional potential (Aqeilan *et al.*, 2004). However, these experiments involved the over expression of tagged AP-2γ and WWox in a cell line where both proteins are already expressed endogenously at high levels. Therefore the physiological relevance of this interaction and the functional consequence of redistribution of AP-2γ are difficult to interpret. Subsequent work in our laboratory by Amelie Morin, has shown that endogenous AP-2γ and WWox can interact in breast cancer cell lines, although the functional consequences of this interaction still remain unclear. Additionally an interaction was also observed for endogenous AP-2γ and other WW domain proteins including YAP and ITCH (Helen Hurst, *personal communication*). Although further work is required, YAP and ITCH appear to affect the protein stability of AP-2γ. YAP promotes AP-2γ stability whereas ITCH promotes AP-2γ degradation. In reporter assays YAP clearly enhances AP-2γ transactivation and ITCH appears to antagonise the effect of YAP. A role for WWox in this fashion was not determined. Interestingly, YAP and ITCH have also been shown to effect the stability of the p73 transcription factor. YAP binds p73 on the same epitope as ITCH and prevents ITCH mediated ubiquitination (Levy *et al*, 2007). A similar mechanism could be controlling stability of AP-2γ.

1.2.2.6 AP-2 and phosphorylation

AP-2α has been reported to be phosphorylated both *in vitro* and *in vivo* by cAMP-dependent protein kinase A (PKA) (Park & Kim 1993; Garcia, *et al.* 1999) at Ser239. Mutation of the Ser239 to Ala abolished phosphorylation of AP-2 by PKA *in vitro*, however, the protein was still able to bind to DNA. Luciferase reporter assays using the

APOE promoter, a potential AP-2α target gene in neuroectodermal cells, were utilised to look at the effects of phosphorylation of AP-2α on gene activation. HepG2 cells, which do not express AP-2α, were co-transfected with the catalytic subunit of PKA and either wild type AP-2α or AP-2α S239A. The results showed that phosphorylation of AP-2α appreciably enhanced the reporter activity, suggesting that cAMP is able to modulate the activity of AP-2α through phosphorylation by PKA, and not through increased transcription of the endogenous AP-2α gene (Garcia *et al.* 1999). The significance of these observations to AP-2 function however remains uncertain as these data have never been repeated.

1.2.2.7 AP-2 and other gene specific transcription factors

As well as the interactions described above, AP-2 has been shown to cooperate or antagonise the activity of other sequence specific transcription factors through direct protein-protein interactions. These include p53, Myc, and SP1 and are discussed with reference to their cellular contexts later in the study.

1.3. AP-2 and embryogenesis

AP-2 family members are widely expressed during vertebrate embryogenesis, showing overlapping and distinct patterns of expression. In particular AP-2 proteins are expressed in neural crest, neural tube and epidermal structures which all develop from the primitive ectoderm during embryogenesis. Table 1.2 summarises the expression of AP-2α, AP-2β and AP-2γ family members during early mouse embryogenesis (Mitchell *et al.*, 1991; Chazaud *et al.*, 1996; Moser *et al.*, 1997a). In particular, AP-2α, AP-2β and AP-2γ are recognised markers of the migratory neural crest and developing neural tube in mice. These family members are also co-expressed in the central and peripheral nervous systems (derivatives of the neural crest), as well in limb buds, epidermal, facial, kidney and placental tissues (Mitchell *et al.*, 1991; Chazaud *et al.*, 1996; Moser *et al.*, 1997a). Considering the significant homology between these proteins, particularly within their DNA binding and dimerisation domains (Figure 1.3) and the remarkable degree of overlap in expression patterns, early studies of AP-2 mutants in mice, zebrafish, frog and chicken revealed surprisingly little redundancy between family members in their overall function.

1.3.1 AP-2α

Mice containing homologous gene disruptions of *tfap2a* die just before birth from multiple defects in neural tube and body wall closure, in limb morphogenesis and in the development of neural crest-derived tissues of the cranio-facial skeleton, cranial ganglia (Schorle *et al.*, 1996, Zhang *et al.*, 1996, Nottoli *et al.*, 1998; Brewer & Williams, 2004; Brewer *et al.*, 2004) and the outflow tract of the heart (Brewer *et al.*, 2002). Interestingly, early neural crest cell migration appeared largely unaffected (Schorle *et al.*, 1996).

This role is supported by analyses of the lockjaw mutation in Zebrafish. Caused by a truncation of the DNA binding domain of AP-2α, lockjaw zebrafish display defects in cranio-facial skeleton, pigment cells and peripheral nervous system development, a phenotype that could be recapitulated using anti-sense directed to AP-2α (Holzschuh, *et*

Days post coitum	Tissue	AP-2α	AP-2β	AP-2γ
8	Neural fold	+++	+++	+++
	Trophoectoderm	++	++	++
	Lateral Head Mesenchyme	+++*	+++*	+++*
	Neural Tube	+	+	+++
9 - 10	Forebrain	-	-	+++
	Midbrain	+++	+++	-
	Hindbrain	++	++	-
	Dorsal Root Ganglia	+++	+	-
	Limb Buds	-	-	+++
	Facial Mesenchyme	++*	+++*	+*
	Cranial ganglia	+++	+++	-
	Epidermis	++	++	n.d
11 - 12	Limb Buds	++	-	+++
	Cranial ganglia	+++	+	-
	Metencephelon	++	++	-
	Myelencephelon	++	++	+
	Spinal Cord	++	++	+
	Facial Mesenchyme	++*	+++*	++*
	Dorsal Root Ganglia	+++	+/-	-
	Dental Lamina	++	-	-
	Kidney	++	++	n.d
	Adrenal Medulla	-	++	n.d
	Cornea	-	++	n.d
	Epidermis	+++	+++	+++

Table 1.2. Expression patterns of three different AP-2 transcription factors during murine embryogenesis. *In situ* hybridisation was scored on the strength of signal as follows; no signal -, weak +/-, +, moderate ++, strong +++, n.d not determined. An asterisk indicates distinct spatial distribution within this tissue type. Figure adapted from Chazaud *et al.*1996 and Moser *et al.* 1997a.

al. 2003; Knight *et al.*, 2003; Knight *et al.*, 2004; Knight *et al.*, 2005). Similarly to mice lacking AP-2α, neural crest cell migration in zebrafish also appeared largely unaffected (Knight *et al.*, 2005).

Ablation of AP-2α during chicken embryogenesis also produced embryos showing critical defects in cranio-facial structures and limb bud formation (Shen *et al.*, 1997). Supporting this data, AP-2α has been shown to negatively regulate chondrocyte differentiation (Haung *et al.*, 2004), a process that occurs in the limb buds of developing embryos and is the first step in bone formation. Levels of AP-2α were shown to decline during normal chondrocyte differentiation. The forced expression of AP-2α inhibited differentiation, via the suppression of key chondrocyte matrix genes including type II collagen, aggrecan, and type X collagen and the suppression of transcription factors Sox5 and Sox6. The multiple skeletal/cartilage defects found in the AP-2α knockout mice are postulated to be a result of premature chondrogenic differentiation in the absence of AP-2α regulation.

Another interesting feature of AP-2α expression with respect to embryonic development is the activation of its expression following Retinoic Acid (RA) treatment (Lüscher *et al.*, 1989; Mitchell *et al.*, 1991; Holzschuh *et al.*, 2003). Congenital malformations resulting from aberrant levels of RA show similar phenotypes to those described above and it has been suggested that AP-2α may be an integral component of the response of an organism to RA signalling (Nottoli *et al.*, 1998; Holzschuh *et al.*, 2003).

In an effort to identify potential genes that are controlled by AP-2α during this developmental programme a subtractive cloning and differential screening approach was conducted to compare expression changes in an *tfap2a* knockout mouse embryo heads compared to embryo heads from a control animal at 8.75 d.p.c. (Pfisterer, *et al.* 2002). Although these results require validation and follow up, a set of genes repressed by AP-2α were identified with known roles in the inhibition of cellular proliferation, activation of differentiation and apoptosis. These included KLF4 (Kruppel-like factor 4), mEFEMP1 (epidermal growth factor-containing fibulin-like extracellular matrix

protein 1), Mtd (matador) and Stra13 (stimulated by retinoic acid 13) (Pfisterer *et al.* 2002).

1.3.2 AP-2β

Mice containing homologous gene disruptions of *tfap2b* displayed fewer gross phenotypic defects than those observed in mice lacking *tfap2a*. Nevertheless, mice lacking *tfap2b* died shortly after birth due to kidney failure (Moser *et al.*, 1997b). In the normal developing mouse AP-2α is expressed in proximal tubular kidney epithelial cells compared to the distal tubular kidney epithelial cell expression observed for AP-2β. AP-2α is down regulated after birth while AP-2β expression persists (Moser *et al.*, 1997a). The study of *tfap2b* null mice showed that this factor is required only for normal adult kidney function. At the end of a largely normal embryonic renal development, distal tubular kidney epithelial cells of *tfap2b* null mice undergo enhanced apoptosis preceded by down-regulation of anti-apoptotic genes of the *BCL* gene family. This leads to failure to form vital structures for water retention and homeostasis and death of *tfap2b* null mice after birth. Of direct relevance to this phenotype is recent work in human cell lines implicating AP-2α in the direct transcriptional repression of *BCL2* (Wajapeyee *et al.*, 2006). It is tempting to speculate that *BCL2* expression may also be influenced by AP-2β, perhaps the imbalance of AP-2α with respect to AP-2β expression may contribute to *BCL2* down regulation and to the enhanced apoptotic phenotype seen in these mice.

Mutations of *TFAP2B* have been implicated in the human condition Char Syndrome (Satoda *et al.*, 2000; Zhao *et al.*, 2001). The main characteristics of this condition include abnormal heart development and facial dysmorphism. Interestingly they share similarities with phenotypes observed following *tfap2a* ablation in mice (described above). This suggests that, rather than producing a non-functional AP-2β protein, these mutations are able to modulate the activity of other AP-2 family members, perhaps through their capacity to form non-productive heterodimers (so-called dominant negative factors).

1.3.3 AP-2γ

Generation of mice lacking *tfap2c* showed it is vital for embryonic survival during early post-implantation development (Auman *et al.*, 2002 and Werling & Schorle, 2002). AP-2γ deficient mice died at 7 to 9 d.p.c. due to failure of the cells derived from trophoectoderm (the first specialised cells which arise from the fertilised egg) and the ectoplacental cone to proliferate. Failure of the placenta to develop and subsequent nutrient deprivation caused the early embryonic lethality (Werling and Schorle, 2002). In contrast to AP-2α and AP-2β, it was suggested that AP-2γ does not appear to play an autonomous role within the embryo proper (Auman *et al.*, 2002). Chimeras made form *tfap2c* null ES cells injected in to wild type blastocytes survived beyond birth with no obvious defects, although the authors admit a detailed histological analysis was not conducted (Auman *et al.*, 2002; Koster *et al.*, 2006). Consistent with these findings AP-2γ has been reported to positively regulate genes involved in early placental development *in vitro*. These include amongst others: adenosine deaminase, chorionic gonadotrophin beta, placental lactogen gene and placental leucine aminopeptidase (Shi & Kellens 1998, Jin *et al.*, 2004, Richardson *et al.*, 2000 and Ito *et al.*, 2002).

1.3.4 AP-2δ and AP-2ε

The more recently described, AP-2δ and AP-2ε genes seem be more temporally and spatially restricted in their expression patterns. AP-2δ is expressed in the central nervous system, retina, and the developing heart (Zhao *et al.*, 2003) whereas AP-2ε is expressed in the developing olfactory bulb (Feng & Williams, 2003). Compared to the other family members these expression patterns are distinct and non-overlapping. This is perhaps reflective of the reduced sequence homology of AP-2δ and AP-2ε compared to the other members.

1.3.5 Drosophila AP-2 protein

Expression pattern analysis of the single *Drosophila* AP-2 protein shows expression in discrete regions of neuroectoderm, brain, ventral nerve cord, and maxillary segment

during *Drosophila* embryogenesis. In the third instar larvae, expression is seen in the brain, optic lobes, ventral nerve cord, antennal-maxillary complex, and antennal and leg imaginal disks (Monge & Mitchell, 1998). This is consistent with the work described above and implies that AP-2 has been structurally and functionally conserved during metazoan evolution.

1.3.6 Redundant activities of AP-2 Family Members

The studies of mutant AP-2 family members described above suggest that there is little redundancy in their overall function. However, studies examining mouse and zebrafish with mutations in two AP-2 family members provide evidence that at least some redundant activity of these proteins is occurring. For example, AP-2α and AP-2γ are co-expressed in early extra-embryonic placental tissues, but only *tfap2c* knockout mice display placental defects dying at 8.5 d.p.c., whereas *tfap2a* knockout mice show no placental defects and die shortly before birth. Recent work by Winger and colleagues, however, highlights their redundant activity by generating mutants knocked out for both *tfap2c* and *tfap2a*. These mice die at 3.5 d.p.c., even earlier than *tfap2c* knockout mice, demonstrating that a level of genetic redundancy of these factors exists within these extra embryonic tissues (Winger *et al.*, 2006).

An additional example of redundancy has been shown in zebrafish embryos. Despite the co-expression of AP-2α and AP-2β in the hindbrain of zebrafish embryos, anti-sense directed to AP-2α caused severe craniofacial defects, whereas anti-sense directed to AP-2β had no visible effect. However, simultaneous anti-sense directed to AP-2α and AP-2β resulted in far more severe defects in craniofacial cartilage than for AP-2α alone. Again this suggests a redundant activity of these factors in cranial epidermal development (Knight *et al.*, 2005; Li & Cornell, 2007).

Most recently a redundant activity of AP-2 family members has been shown to be important for their role in neural crest induction. Neural crest induction refers to the specification of ectoderm cells to the neural crest lineage. Despite AP-2 expression in early stages of neural crest induction, the single mutant studies above describe defects only in cells derived from the neural crest, with early stages of neural crest cell

migration largely unaffected (Schorle *et al.*, 1996; Knight *et al.*, 2005). It is interesting therefore that recent work by Li and Cornell showed the failure of neural crest induction altogether in zebrafish embryos deficient in both AP-2α and AP-2γ, but no induction defect in embryos singly deficient in either gene (Li & Cornell, 2007). Additional phenotypes observed in the double AP-2α and AP-2γ mutants but not single mutants included abnormal epidermal, lens, ear, and olfactory placode development and highlight further redundant activities for these two proteins.

Together these three examples suggest AP-2 family members can compensate for each other's absence in many contexts during mouse and zebrafish development and when considered with single AP-2 mutant phenotypes, suggest that AP-2 family members function both separately and in tandem to coordinate the development of ectoderm derivatives during embryogenesis.

1.4 AP-2 and tissue specific development

The embryological defects described in *tfap2* null mice clearly implicate a role for AP-2 in the regulation of differentiation and proliferation during embryogenesis. However, the severe perinatal defects observed in *tfap2* null mice have precluded much of the analysis of their involvement in postnatal tissues. Recent studies have revealed an important role for AP-2 family members in other ectoderm lineages particularly in the self-renewing tissue of mammalian skin epidermis, and differentiating mammary gland tissue but also during gonocyte development.

1.4.1 AP-2 and gonocyte development

Initial work by Jager and colleagues showed that targeted over expression of AP-2γ to the testis of male mice resulted in increased cell proliferation counter balanced by enhanced apoptosis in the epithelial structures of the seminal vesicles (Jager *et al.*, 2003). Differentiation was also impaired in these cells marked by the lack of androgen receptor expression. A more detailed analysis of AP-2γ expression in human seminal vesicles revealed high AP-2γ levels in immature undifferentiated germ foetal cells (Pauls *et al* 2005), and gradual down regulation of AP-2γ was observed with increasing differentiation of foetal testis. Authors suggested that AP-2γ expression might be required for the suppression of differentiation and the promotion of proliferation in germ cells (Pauls *et al* 2005). In support of this role, AP-2γ expression is found in germ cell tumours which share characteristics with early germ cells (Hoei-Hansen *et al.*, 2004 and Pauls *et al.*, 2005).

1.4.2 AP-2 and skin epidermis

Normal development of skin epidermis is first characterised by a single-layered surface ectoderm, which eventually forms the basal regenerating layer in developed skin. Subsequent commitment to stratification is dictated largely by the balance of p63 transcription factor isoforms and results in multiple layers of supra-basal differentiated epithelial cells characterised by the expression of specific keratins (reviewed by Koster

et al., 2004a). Importantly, AP-2 family members have been implicated in the regulation of transcription from keratin promoters in cultured keratinocytes, outlining a role for these factors in differentiated cells of the skin (Leask *et al.*, 1991; Magnaldo *et al.*, 1993; Wanner *et al.*, 1997; Chen *et al.*, 1997; Dunn *et al.*, 1998; Talbot *et al.*, 1999; Maytin *et al.*, 1999; Kaufman *et al.*, 2002). Furthermore, frog embryos show a loss of epidermal character following injection of AP-2α anti-sense oligonucleotides (Luo *et al.*, 2002). Recent work by Wang and colleagues has shown that, with the exception of AP-2δ, expression of all the AP-2 family members was observed in normal mouse skin epidermis. In particular, AP-2α was expressed at high levels in the supra-basal nuclei and to a lesser extent in the basal nuclei of skin epidermis, whereas high levels of AP-2γ were restricted to basal nuclei (Wang *et al.*, 2006).

A conditional *tfap2a* null mouse, targeting AP-2α ablation to the skin epidermis resulted in hyperproliferation of the suprabasal cells caused in part by an increase in EGFR expression (Wang *et al.*, 2006). The effect of AP-2α ablation was less dramatic in the basal cells of the epidermis where high levels of AP-2γ were present, suggesting AP-2α independent mechanisms control proliferation in these cells. EGFR promoter occupancy by AP-2α in wildtype cultured primary keratinocytes but not keratinocytes null for AP-2α, suggested that AP-2α functions in the epidermis by repressing EGFR gene transcription as cells exit the basal layer and commit to terminal differentiation.

In parallel to these findings for AP-2α, two recent studies have identified a role for AP-2γ in earlier, p63-dependant events of epidermal morphogenesis. Studies in p63 null mice have shown that the transactivating (TA) isoforms of p63 are required for the commitment to initiate epithelial stratification in surface ectoderm (Koster *et al.*, 2004a; Koster *et al.*, 2004b). Subsequent work by Koster and colleagues showed, by Chromatin IP, that transcriptional activation of AP-2γ by TAp63 was fundamental to this process (Koster *et al.*, 2006). In the absence of AP-2γ, TAp63 failed to induce Keratin 14 gene transcription, highlighting a role for AP-2γ in the transcription of this early epidermal marker (Koster *et al.*, 2006). As epidermis differentiates, the levels of ΔN p63 isoforms (forms without transactivation domains that act to inhibit transcription activation), increase with respect to TAp63 isoforms. This p63 isoform balance has been suggested to be important for commitment to terminal differentiation in epidermis

(Koster *et al.*, 2004a). ΔNp63 was shown to inhibit AP-2γ transcriptional activation, in a concentration dependent manner and this may be important for the loss of AP-2γ expression observed in differentiated supra-basal cells of the epidermis (Koster *et al.*, 2006). A further level of complication has been revealed by the fact that AP-2 family members together with p63 are important in the maintenance of transcription from the p63 promoter, via an epidermis specific distal enhancer region (Antonini *et al.*, 2006). Together these two studies suggest the presence of a possible control loop between AP-2 family members and p63. Based on these findings a role for AP-2 factors in epidermis is suggested in Figure 1.5.

In summary AP-2α is important for the control of signalling and proliferation in differentiating supra-basal cells of the skin epidermis. Although the function of AP-2γ requires further elaboration, its carefully controlled and restricted basal expression suggests it may be of some importance in the aetiology of the basal cell type.

1.4.3 AP-2 and mammary morphogenesis.

Two major tissue types constitute the mammary gland. Firstly, the stroma, which predominately consists of adipocytes but also contains fibroblasts, vasculature and haemoatopoiteic cells. The second compartment is the mammary epithelium, which forms the milk producing alveoli and associated ducts for milk transport. Following the initial embryonic stages of development these tissues undergo a sequence of hormonally controlled developmental stages termed puberty, pregnancy, lactation and regression or involution (Reviewed by Hennighausen & Robinson, 2001, Robinson, 2004 and Hennighausen & Robinson, 2005). These stages are summarised in Figure 1.6. The binding of ligands such as EGF, TGFα and Amphiregulin to members of the epidermal growth factor receptor family (EGFR, ErbB2 and ErbB3), activates EGFR endogenous tyrosine kinase activity and further stimulates growth and branching of the epithelium in the virgin and pregnant mouse mammary gland (Sternlicht *et al.*, 2006).

AP-2 family members were first implicated in the control of growth and differentiation of cells of the mammary gland through early studies showing the association of AP-2α

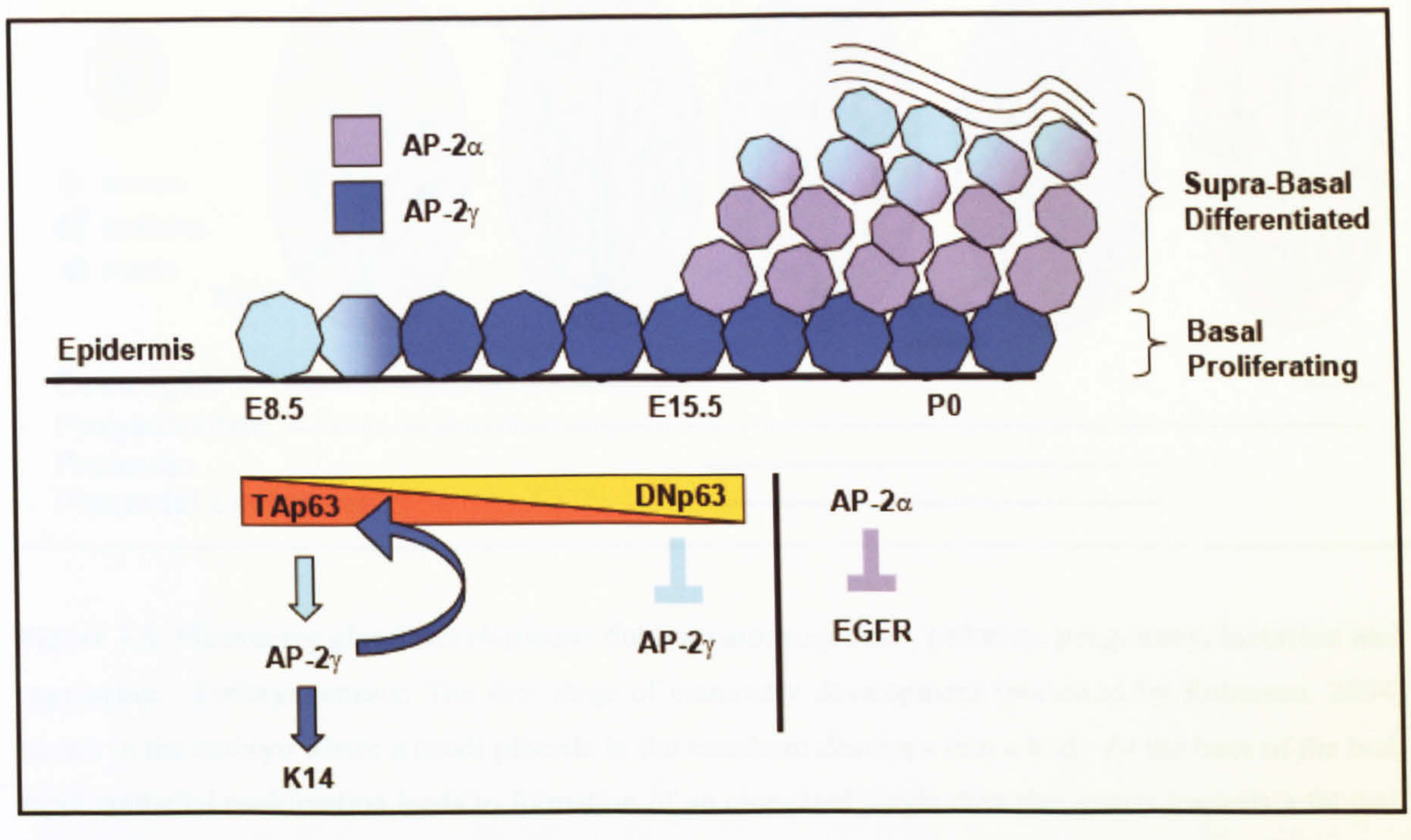


Figure 1.5. A proposed role for AP-2 and p63 in the differentiation of skin epidermis. First expression of TA isoforms of p63 initiate the differentiation of the surface ectoderm and up-regulate AP-2γ. AP-2γ and p63 can then act together to maintain p63 transcription and thus commit the surface ectoderm to differentiate. As differentiation progresses the ratio of ΔNp63 isoforms increases with respect to TAp63 isoforms, perhaps inhibiting AP-2γ and TAp63 expression and committing cells to terminally differentiate. An accompanying increase in AP-2α expression acts to repress EGFR expression and prevents proliferation in these differentiating cells. The time line of p63 expression was adapted from Koster *et al.*, 2006 (E8.5 and E15.5 are d.p.c. and P0 corresponds to birth).

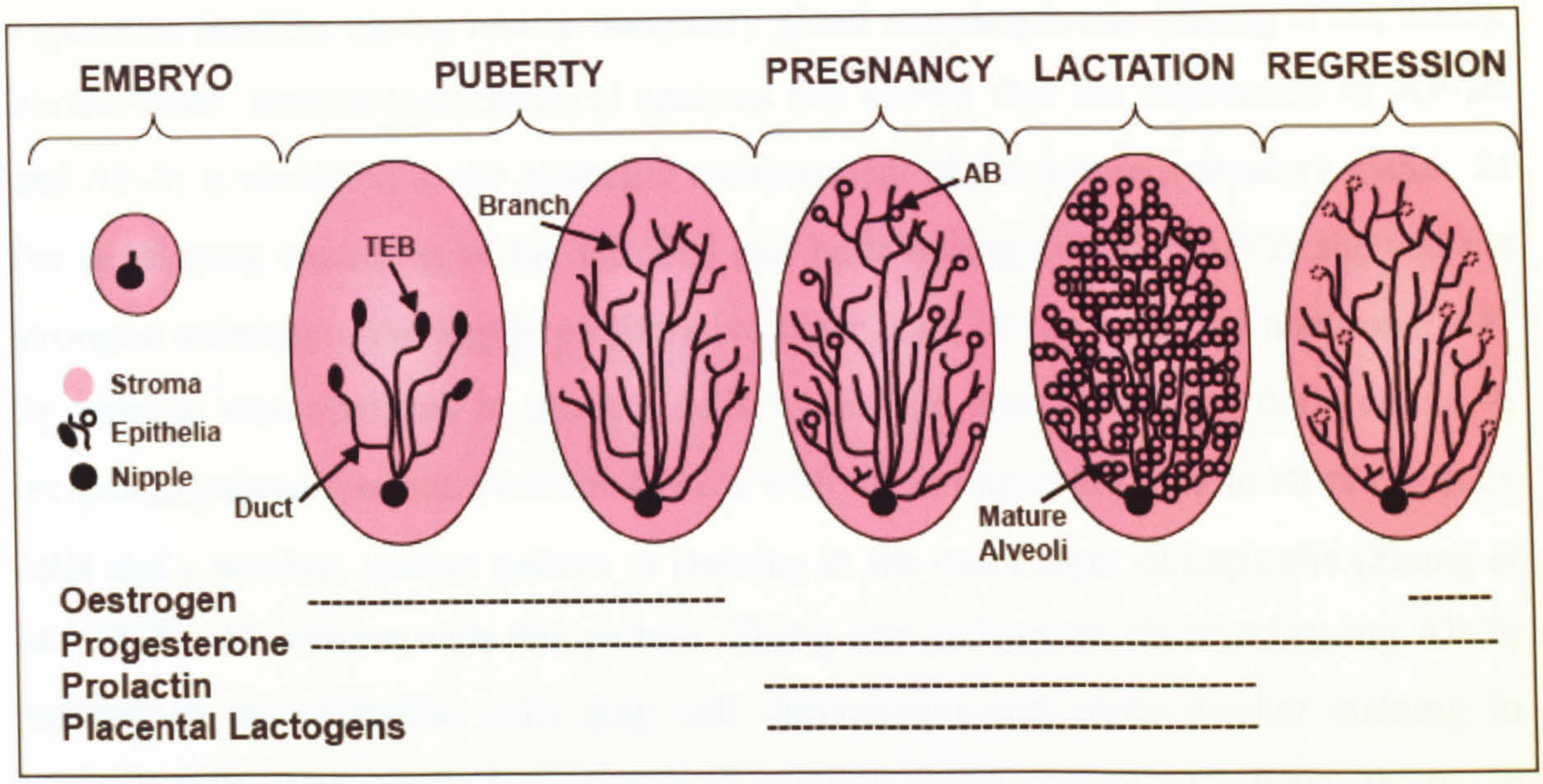


Figure 1.6. Mammary gland development during embryogenesis, puberty, pregnancy, lactation and regression. Embryogenesis: The first stage of mammary development (reviewed by Robinson, 2004) occurs in the embryo where a small placode in the ectoderm develops into a bud. At the base of the bud, rapid epithelial proliferation leads to formation of an elongated single duct that grows towards a fat pad. Finally, mesenchyme derived structures form the nipple. **Puberty:** During puberty cyclical production of ovarian oestrogen and progesterone accelerate epithelial duct outgrowth. Terminal end buds (TEB), consisting of an outer layer of cap cell (myoepithelial progenitors) and an inner layer of body cells (luminal progenitors) are rapidly cycling and facilitate ductal outgrowth. In the mature gland ductal side branches form and disappear with each oestrogen cycle. **Pregnancy**: At pregnancy placental lactogens, prolactin, and progesterone stimulate cell proliferation, alveolar bud (AB) formation and alveolar expansion. **Lactation:** During lactation luminal cells of mature alveoli synthesise milk (characterised by the expression of milk genes and milk fat production), which is transported through the ducts to the nipple for feeding. **Regression or Involution:** Upon cessation of feeding, apoptosis occurs in the secretory epithelium and the surround stroma is remodelled to replace apoptosed cells. Finally, a cyclical production of ovarian oestrogen and progesterone returns. This diagram is based on information reviewed by Hennighausen and Robinson (Hennighausen & Robinson, 2001; Hennighausen & Robinson, 2005).

and AP-2γ expression with breast carcinoma (discussed in detail in Section 1.5, Bosher *et al.*, 1995; Turner *et al.*, 1998). Subsequent work by Zhang and colleagues, displayed in Figure 1.7 A, has shown that AP-2α, AP-2β and AP-2γ share similar temporal expression patterns during mouse mammary gland morphogenesis (Zhang *et al.*, 2003). Furthermore, immunohistochemical analysis has shown that the expression of AP-2α and AP-2γ is restricted to the epithelial compartment of the mouse mammary gland. In the developing structures of the terminal end buds during puberty, AP-2γ showed the strongest staining in the highly proliferative outer layer of cap cells and additional AP-2γ staining was observed in discrete cells within the area containing body cells. A reciprocal pattern was observed for AP-2α, with the strongest staining in all of the body cells and a weaker, sparser pattern of staining in the outer layer of cap cells (Zhang *et al.*, 2003). Consistent with this pattern, Zhang and colleagues observed strong AP-2γ staining in myoepithelial cells (cap cell derivatives) and some weaker staining in luminal cells (body cell derivatives) of mature virgin mouse mammary ducts. In contrast, AP-2α was seen in both cell types. A similar pattern of expression was also seen in the alveolar tissue of pregnant mice (Zhang *et al.*, 2003).

Subsequent studies looking at AP-2α and AP-2γ staining mature in human mammary ducts tend to support the findings in mouse. In particular, work from our lab (shown in Figure 1.7 B) and others shows strong myoepithelial AP-2γ staining with additional weaker luminal staining (Hurst, *personal communication* and Hoei-Hansen *et al.*, 2004). However, work by Fredrichs and colleagues showed more restricted staining: AP-2α stained only the luminal cells of mature human mammary ducts while AP-2γ staining was present only in discrete cells within the myoepithelial layer (Friedrichs *et al* 2005). All of these studies demonstrated the AP-2 family member specificity of the antibodies used for immunohistochemistry. Therefore, the slight differences between the data from Fredrichs and colleagues compared to others might reflect a difference in antibody avidity.

Unfortunately, gene knockout studies, in mice described in the previous sections could not provide a direct analysis of the role of AP-2 factors in breast development, due to the perinatal death caused by the severe embryological phenotypes. However, separate transgenic mouse studies targeted AP-2α or AP-2γ over expression to the mouse

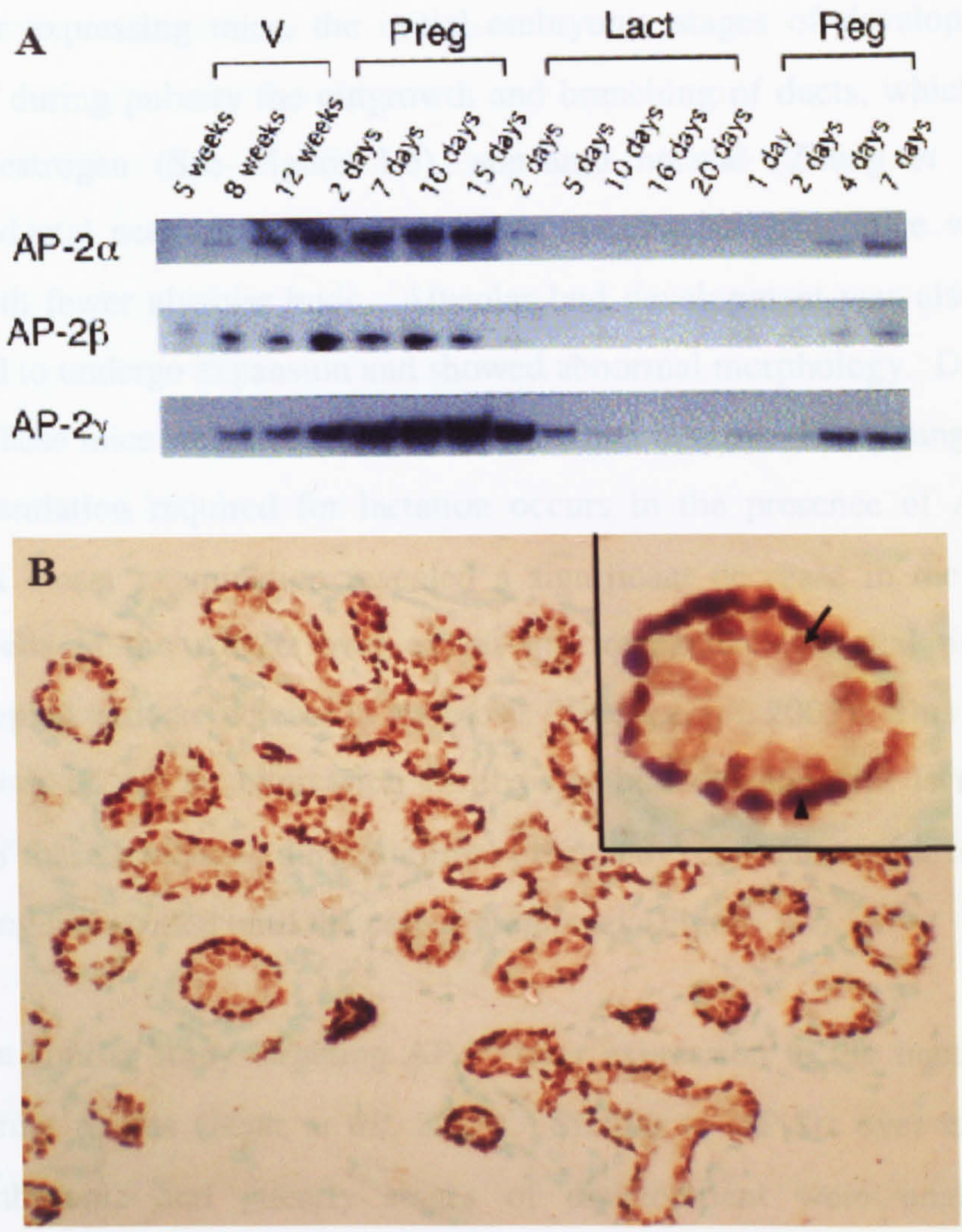


Figure 1.7. Expression of AP-2 family members in the mammary gland. (A) Figure from Zhang *et al.*, 2003. An RNase protection analysis of AP-2 family members during various stages of mouse mammary gland maturation (as indicated V, virgin; Preg, pregnant; Lact, lactation; Reg, regression or involution). **(B)** Immunohistochemical analyses of AP-2γ protein levels in transverse sections of normal adult human mammary tissue. Inset: Enlargement of a single duct shows strong staining of AP-2γ in myoepithilial cells (arrow head) of the outer (basal) layer. A weaker staining in some of the luminal epithelial cells (arrow), the inner layer of the duct lining is also observed. Staining is absent from the surrounding stroma. (Helen Hurst, *personal communication*).

mammary gland, using a transgene under the control of the long terminal repeat of the mouse mammary tumour virus (Zhang *et al.*, 2003; Jäger *et al.*, 2003).

In AP-2α over expressing mice, the initial embryonic stages of development were unaffected and during puberty the outgrowth and branching of ducts, which occurs in response to oestrogen (See Figure 1.6), appeared normal (Zhang *et al.*, 2003). However, the ductal network of AP-2α over expressing pregnant mice was slightly sparser and with fewer alveolar buds. Alveolar bud development was also impaired. The buds failed to undergo expansion and showed abnormal morphology. Despite these abnormalities these mice were able to produce milk and to nurse their young, indicating that the differentiation required for lactation occurs in the presence of AP-2α over expression. A closer examination revealed a significant decrease in the number of proliferating cells of the AP-2α over expressing transgenic epithelial tissue during pregnancy, coupled with an increase in apoptosis (Zhang *et al.*, 2003). This implies that loss or reduction of AP-2α expression in alveolar bud development is required for proliferation of these structures during normal pregnancy, despite some level of AP-2α expression being maintained until the end of pregnancy (Figure 1.4).

Interestingly, a similar study targeting AP-2γ over expression to the mammary gland revealed different results (Jäger *et al.*, 2003). Similar to AP-2α over expression in mice, the embryonic and puberty stages of development were unaffected. But contrastingly, in AP-2γ over expressing pregnant mice, the ductal network and number of alveolar buds appeared similar to those of wild type mice. However, at later stages of pregnancy, these alveolar buds failed to develop into mature alveoli in AP-2γ over expressing mice and these mice were unable to nurse their young. A closer examination revealed that impaired alveolar development was due to hyperproliferation of cells within these structures, counterbalanced by enhanced apoptosis. Analysis of differentiation markers (Whey Acidic Protein and β-casein) showed that the forced expression of AP-2γ prevented these cells form terminally differentiating. This implies that loss or reduction of AP-2γ expression in later stages of alveolar bud development is required for cessation of proliferation and for terminal differentiation.

In summary, these studies emphasise the apparent opposite effects of the two family members on cell proliferation in the mammary gland. Forced AP-2α expression inhibits epidermal proliferation in developing alveolar buds whereas forced AP-2γ expression appears to promote proliferation in the same structures. Based on these observations, it is reasonable to consider AP-2 family members may also play a role in the balance of proliferation and differentiation of TEBs, developing structures in which they are also expressed. Indeed, recent unpublished work by Hubert Schorle and colleagues examining the phenotype of targeted ablation of AP-2γ to mammary epithelial tissue supports this hypothesis (Helen Hurst, *personal communication*). In *tfap2c* null mammary glands a severe impairment in ductal elongation was observed during puberty despite the presence of TEB structures.

It is important to stress that despite the hyperproliferative effect observed in AP-2γ over expressing mice, no mammary tumourigenesis was observed, but the observed epithelial hyperplasia is thought to be a typical initiating event in breast tumourigenesis (Beckmann *et al.*, 1997). The authors postulated that overcoming the pro-apoptotic mechanisms in these cells, a common occurrence during tumourigenesis (section 1.4), might permit tumours to form in these animals.

Indeed both AP-2α and AP-2γ have been implicated in tumourigenesis and their proposed roles in human cancers reflect their apparently opposing roles in post-embryonic mammalian development. AP-2α expression suppresses proliferation and promotes differentiation and therefore is often lost in cancer. Contrastingly, AP-2γ expression has often been found in tumours, and although further research is necessary, evidence suggests AP-2γ might be driving proliferation and suppressing differentiation in these tumours. Before the association of AP-2 and cancer is discussed, I will first provide a brief introduction into the common mechanisms that are thought to promote carcinogenesis in humans. Particular attention is placed on the regulation of cell proliferation in which AP-2 factors have been suggested to play a role.

1.5 Human Cancer

The evolution of cancer is a considerably complex process, since specific interactions between tumour cells and host tissues are necessary for tumour angiogenesis, tissue invasion, and metastasis. Yet, despite the existence of many forms of cancer it has been suggested that relatively few essential alterations are common to most if not all tumours, shown in Figure 1.8 (reviewed by Hanahan & Weinberg, 2000).

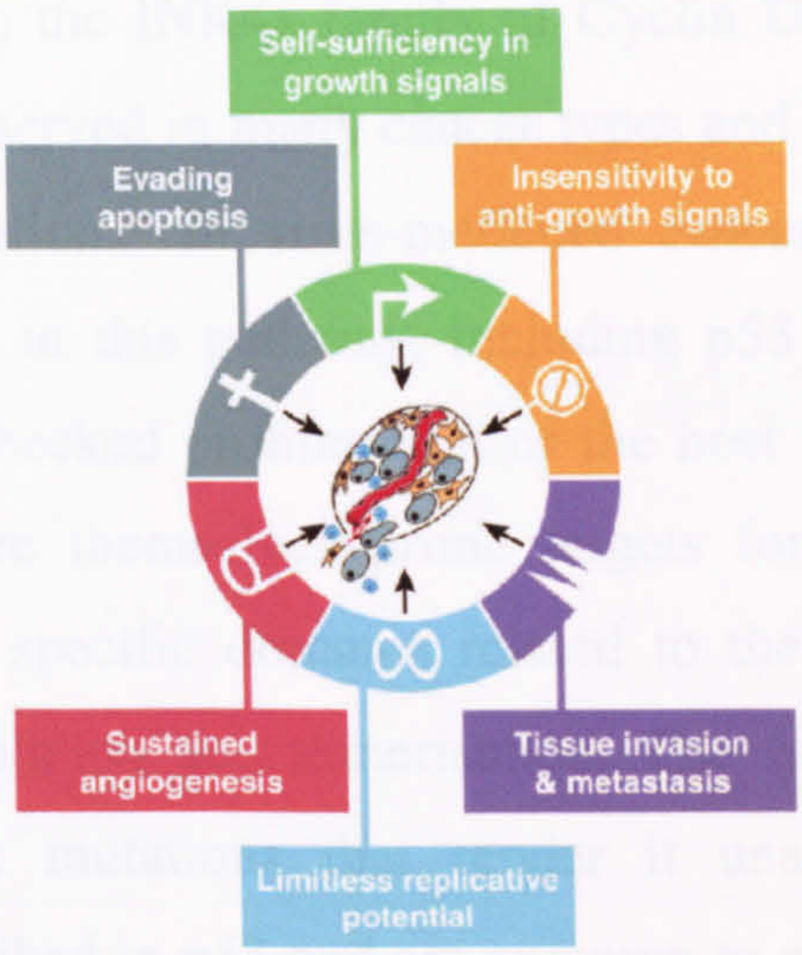


Figure 1.8 The Hallmarks of Cancer (from Hanahan & Weinberg 2000). Hanahan and Weinberg have suggested that all human cancers are a symptom of six essential alterations in cell physiology that collectively dictate tumour progression.

Over expression or inappropriate activation of growth promoting signals, by so-called oncogenes is significant in promoting a cancerous state. Cell surface receptors that normally relay extra cellular growth stimulatory signals into the cell are oncogenic targets for deregulation in cancer. For example, receptor over expression, such as occurs for ErbB2 in mammary carcinomas (Slamon *et al.*, 1998) enable a cancer cell to become over-responsive to ambient levels of growth factor that might not normally trigger cell proliferation. Equally, mutations in growth factors receptors can alter them to become constitutively activated in the absence of ligands.

A second important event is insensitivity to anti-growth signals. Within a normal developed tissue the potential of a cell to proliferate is held in check by the absence of growth signals, via soluble growth inhibitors or inhibitors present on the surfaces of

contacting cells and in the extra cellular matrix. Cells can be held in a quiescent state and are prevented from cell cycle entry or they can undergo terminal differentiation into as state where they permanently loose there potential to proliferate. Most of the machinery that facilitates a response to antigrowth signals in a normal cell is associated with the mitotic cell cycle; in particular these pathways tend to converge on transition from G1 phase to S phase (shown in Figure 1.9.). Disruption of this pathway can render cells insensitive to anti-growth signals that normally operate at this checkpoint and thus unleash the activity of E2F transcription factors and allow inappropriate cell proliferation. Mutations in the INK4a family of Cyclin Dependant Kinase Inhibitors (CDKIs) are commonly observed in many cancer types and are thought to be important early events in transformation. In virus-mediated cancers, viral proteins bind and sequester specific proteins in this pathway, including p53 and Rb, inhibiting normal function and allowing unchecked proliferation of the host cell. The key transcription factors in this pathway are themselves prime targets for mutation given that their structures consist of very specific domains related to their functions, such as DNA binding, transcription regulation and dimerisation. The function of RB can become disrupted through genetic mutations that render it unable to interact with E2F. Mutations have been described in p53 and are common to all types of cancer cells. It a normal cell p53 becomes stabilised in response to upstream kinases that are activated by DNA damage in the cell (see Figure 1.9.). This in turn leads to cell cycle arrest and apoptosis via the transcriptional activation of the CDKI, p21 and additional pro-apoptotic genes. Mutations in p53 render cells unresponsive to DNA damage signals which in turn will exacerbate the cancer phenotype.

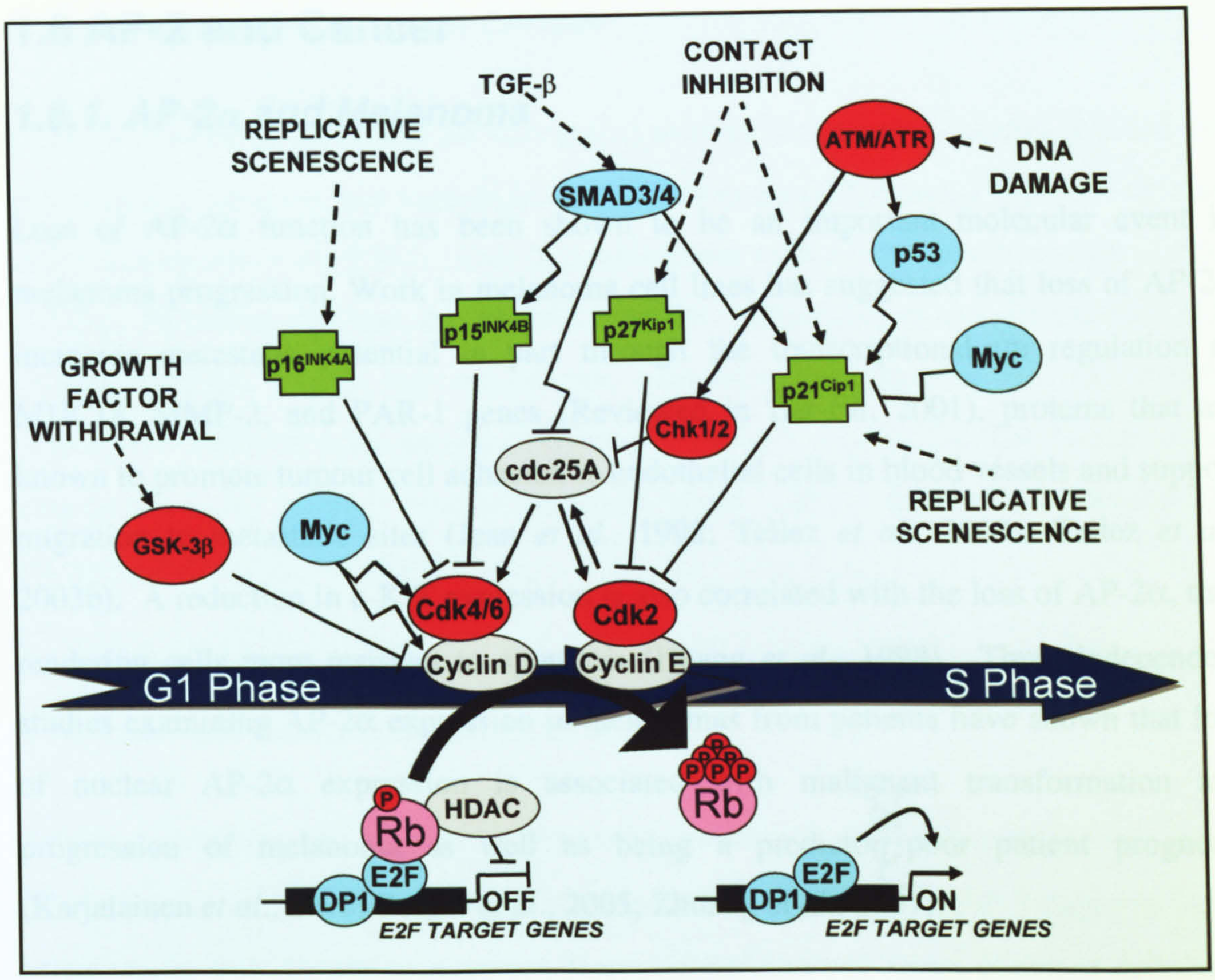


Figure 1.9. Overview of the control of cell cycle progression from G1 phase to S phase. Central to the progression from G1 phase to S phase of the cell cycle is the control of phosphorylation of Retinoblastoma Protein (Rb). In its hypophosphorylated state, Rb and its associated HDAC repressor complex bind to the E2F-DP1 transcription factors, inhibiting downstream transcription of E2F target genes required for DNA synthesis and subsequent cell cycle phases. In response to upstream signals, the Cyclin Dependant Kinases (CDKs) and their associated Cyclin proteins sequentially phosphorylate Rb. In turn this leads to dissociation of the RB repressor complex permitting the transcription of E2F target genes and thus promoting cell cycle progression. Cyclin/CDK complexes are regulated at the level of transcription by the Myc transcription factor, or via the inhibition of their activity by upstream INK4 or Kip/Cip families of CDKIs, shown in green. TGFβ signalling, DNA damage, contact inhibition and replicative senescence all act to inhibit cell cycle progression by inducing members of the INK4 or Kip/Cip families of CDKIs. TGFβ also inhibits the transcription of cdc25A, a phosphatase required for CDK activation. Growth factor withdrawal can activate GSK-3β, which in turn phosphorylates cyclinD, leading to its rapid ubiquitination and proteosomal degradation (Reviewed by Sherr & Roberts 1999; Caldon *et al.*, 2006). Solid lines indicate direct interactions, Dashed lines indicated indirect interactions and crooked lines indicate transcriptional interactions. Kinases are shown in Red, CKI in Green and transcription factors in blue.

1.6 AP-2 and Cancer

1.6.1. AP-2α and Melanoma

Loss of AP-2α function has been shown to be an important molecular event in melanoma progression. Work in melanoma cell lines has suggested that loss of AP-2α increases metastatic potential in part through the transcriptional up regulation of MUC18, MMP-2, and PAR-1 genes (Reviewed in Bar-Eli, 2001), proteins that are known to promote tumour cell adhesion to endothelial cells in blood vessels and support migration to metastatic sites (Jean *et al.*, 1998; Tellez *et al.*, 2003a; Tellez *et al.*, 2003b). A reduction in c-KIT expression is also correlated with the loss of AP-2α, thus rendering cells more resistant to apoptosis (Huang *et al.*, 1998). Three independent studies examining AP-2α expression in melanomas from patients have shown that loss of nuclear AP-2α expression is associated with malignant transformation and progression of melanoma as well as being a predictor of poor patient prognosis (Karjalainen *et al.*, 1998; Berger *et al.*, 2005; Zhuang *et al.*, 2007).

1.6.2. AP-2α and prostate cancer

A similar role for AP-2α has been suggested in prostate carcinoma. AP-2α is expressed in normal differentiating prostate epithelium but expression is lost early in prostate cancer development (Ruiz *et al.*, 2001). Interestingly, forced AP-2α expression in AP-2-negative prostate cancer cell lines, caused reduced tumourigenicity and was correlated with down regulation of the growth promoting VEGF (Ruiz *et al.*, 2004). This study suggested that AP-2α can repress VEGF gene transcription via direct interaction with its promoter (Ruiz *et al.*, 2004). Interestingly, work by others in keratinocytes implicates AP-2α in the transcriptional activation of VEGF (Gille *et al.*, 1997; Brenneisen *et al.*, 2003). This highlights the importance of cellular context in the regulation of AP-2 target genes.

1.6.3 AP-2α and ovarian cancer

Over expression of AP-2α in AP-2-negative ovarian cancer cells, caused reduced cell proliferation, invasion and tumourigenicity when assessed following injection in nude mice (Sumigama *et al.*, 2004). This is in agreement with a tumour suppressing role for AP-2α. However, this was inconsistent with the examination of AP-2α expression in ovarian tumours, where low cytoplasmic expression accompanied by high nuclear expression predicted poor patient prognosis (Anttila *et al.*, 2000). The inconsistency may be due to the antibody used in this study, which failed to discriminate between AP-2α and AP-2γ family members.

1.6.4 AP-2α and colon cancers

AP-2α expression has also been inversely correlated with colon carcinoma grade (Ropponen *et al.*, 2001). Recently, the expression of AP-2α in AP-2-negative colon cancer cells was shown to suppress their tumourigenicity when assessed following injection in nude mice (Schwartz *et al.*, 2007). In the same study, RNAi mediated silencing of endogenous AP-2α in AP-2α-positive colon cancer cells, was shown to increase migration in invasion assays. The transcriptional regulation of E-cadherin and MMP-9, by AP-2α, via direct interaction with their promoters, was implicated in the observed AP-2α dependant phenotypes (Schwartz *et al.*, 2007).

1.6.4 AP-2α and Gliomas

In addition to the examples described above, expression of AP-2α in gliomas associated with MMP-2 and VEGF expression has been inversely correlated with glioma grade (Ropponen *et al.*, 2001).

Broadly, in the examples described above, the reduction or loss of nuclear AP-2α expression in tumours appears to be correlated with increased proliferation, increased disease progression and poor patient prognosis.

1.6.5 AP-2γ and cancer

In contrast to AP-2α, high levels of AP-2γ have often been found in tumours and although further research is necessary AP-2γ in some cases have been correlated with increased disease progression. High levels of AP-2γ are associated with testicular cancer (Hoei-Hansen *et al.*, 2004), germ cell tumours (Hoei-Hansen *et al.*, 2004; Pauls *et al.*, 2005), advanced stages of ovarian cancer (Anttila *et al.*, 2000) and certain stages of melanoma (Nyormoi & Bar-Eli, 2003). A recent study advocated the detection of AP-2γ expression as a useful diagnostic tool in the detection of testicular cancer (Hoei-Hansen *et al.*, 2005).

1.7. AP-2 and Breast Cancer

In addition to the aberrant AP-2 expression associated with the cancers described above, AP-2 proteins were originally linked to carcinogenesis via their association with breast carcinoma.

1.7.1 Breast Cancer

Breast cancer causes the premature death of more than 14,000 women a year in the UK. About 10% of these cases are thought to result from a hereditary predisposition to the disease, caused by mutations in genes including *BRCA1*, *BRCA2* and *CHEK2* thus leading to the malfunction of DNA damage checkpoints (Wooster & Weber, 2003). The remaining 90% of cases are sporadic. Clinical evidence suggests that breast cancer starts with atypical epithelial hyperplasia, followed by ductal carcinoma *in situ* (DCIS) and then invasive carcinomas leading to metastatic tumours (Beckmann *et al.*, 1997). Figure 1.10 illustrates the typical mammary gland histology at these stages. The most important markers discovered to date in breast cancer are the hormone receptors: oestrogen receptor (ERα) and progesterone receptor (PrR) and the receptor tyrosine kinase ErbB2 (or Her2) expressed in 75–80% and 15–20% respectively, of all cases. These proteins are intrinsically liked to the aetiology of these tumours and importantly are amenable targets for therapeutic intervention. Selective oestrogen receptor modulators (SERMs) bind oestrogen receptor and exert anti-oestrogen activity in breast tumour tissue. ERα positive patients generally respond well to SERMs and therefore have a good survival prognosis, although prolonged treatment is often associated with the recurrence of hormone independent tumours and therefore resistance to therapy. Aromatase inhibitors are used in conjunction with SERMs to inhibit local oestrogen synthesis and can help reduce resistance to therapy. In contrast, ErbB2 over expression in breast cancer correlates with aggressive disease and poor patient prognosis (Slamon *et al.*, 1998; Siegel *et al.*, 1999). ErbB2 positive breast tumours are often ERα negative and therefore resistant to treatment with SERMs. Recent developments have shown that these tumours may respond to ErbB2 directed monoclonal antibody therapy (reviewed by Cleator *et al.*, 2007) and it is hoped that this treatment will lead to improved

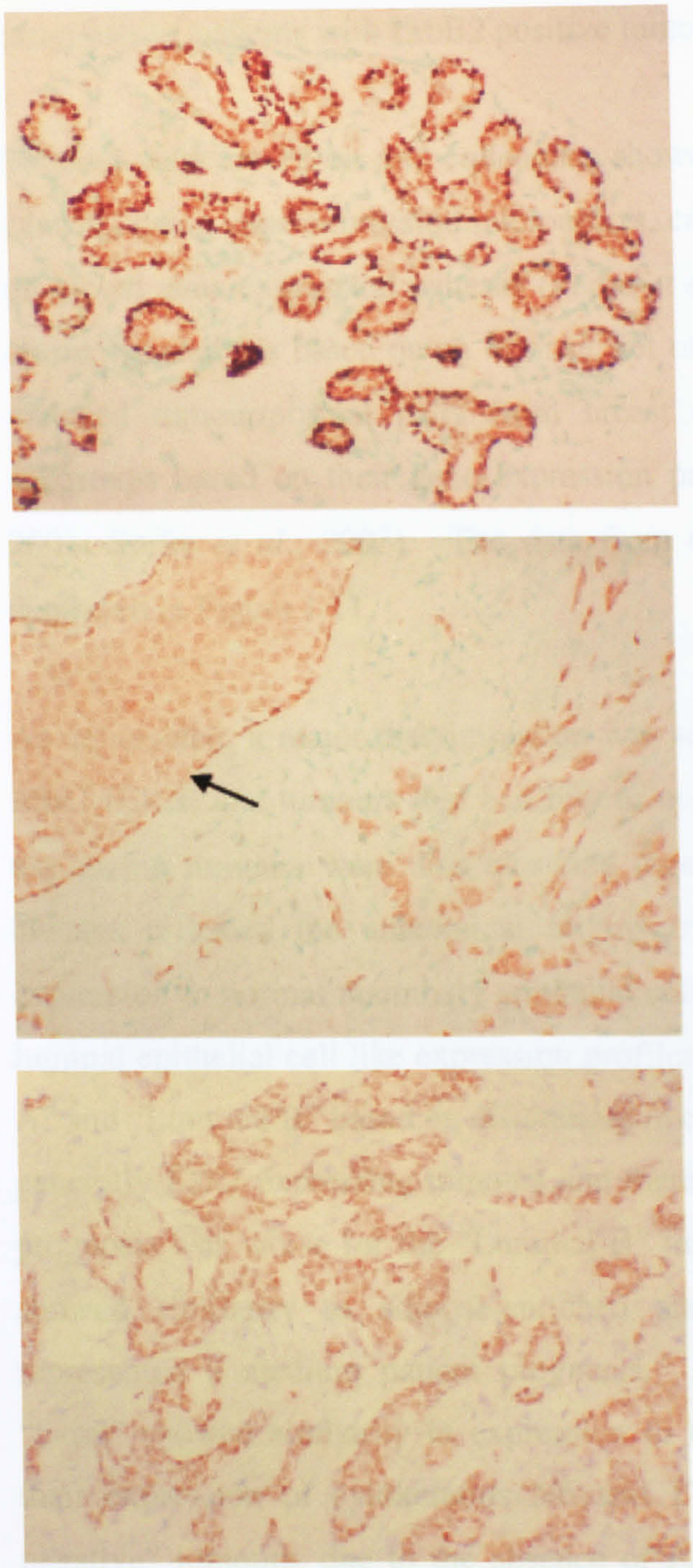


Figure 1.10. Typical stages of breast cancer. The top panel shows the typical morphology of normal human breast tissue (described in detail in section 1.4.3). Clinical evidence suggests that breast cancer starts with atypical epithelial hyperplasia, followed by ductal carcinoma *in situ* (DCIS) shown in the middle panel indicated by an arrow and then primary invasive carcinoma (bottom panel) characterised by a loss of recognisable morphology. Images show immunohistochemical analysis of AP-2γ protein levels in transverse sections of adult human mammary tissue. AP-2γ staining corresponds to myoepithelial and luminal epithelial cells within the normal mammary gland, in DCIS and in primary invasive carcinoma (Helen Hurst, *personal communication*).

prognosis of patients with ErbB2 positive tumours.

Seminal work by Perou and colleagues showed that the molecular profiling of breast cancers, using gene expression microarrays, can further classify tumour subtypes based on global gene expression patterns (Perou *et al.*, 2000). An unsupervised hierarchical clustering analysis based on an intrinsic set of genes showing high variability between matched tumour/normal pairs from breast cancer patients, identified five distinct subgroups based on their gene expression profiles (Perou *et al.*, 2000, Sorlie *et al.*, 2001; Sorlie *et al.*, 2003). The data from experiments by Sorlie and colleagues is displayed in Figure 1.11.

As anticipated, a major discrimination was seen between tumours that expressed high *ESR1* (ERα) and tumours that had low or negative expression for this factor. *ERBB2* expressing tumours were also clustered together. Interestingly, these and additional clusters reflected the anatomical origins of tumours, based on their similarity in expression to normal mammary epithelial cells. The largest cluster of tumours showed luminal epithelial cell like expression profiles. These further subdivided into "Luminal A" and "Luminal B" based on differences in expression profile. Luminal subtypes were generally *ESR1* expressing tumours and therefore represented good prognosis although prognosis was worse for the "Luminal B" tumours. The "Normal" cluster of tumours showed similarity to adipose-enriched and basal mammary epithelial cells and represented a medium patient prognosis. A final group of breast tumours termed "basal" showed similarity in expression to basal mammary epithelial cells, based on their expression of cytokeratins 5/6 and 17. These have also been called "triple-negative" tumours, due to the absence of *ESR1, PGR1* (PgR) and *ERBB2* expression and additional validation studies showed they represent very poor patient prognosis due to the lack of expression of known targets for therapeutic intervention (van de Rijn *et al.*, 2002; reviewed by Cleator *et al.*, 2007). Further analysis of the gene expression profile of these basal like tumours has revealed that they share similarities to *BRCA1* mutated familial breast cancers implicating BRCA1 dysfunction in the aetiology of this subtype.

Importantly, identifying new and distinct subgroups of breast tumours, which show

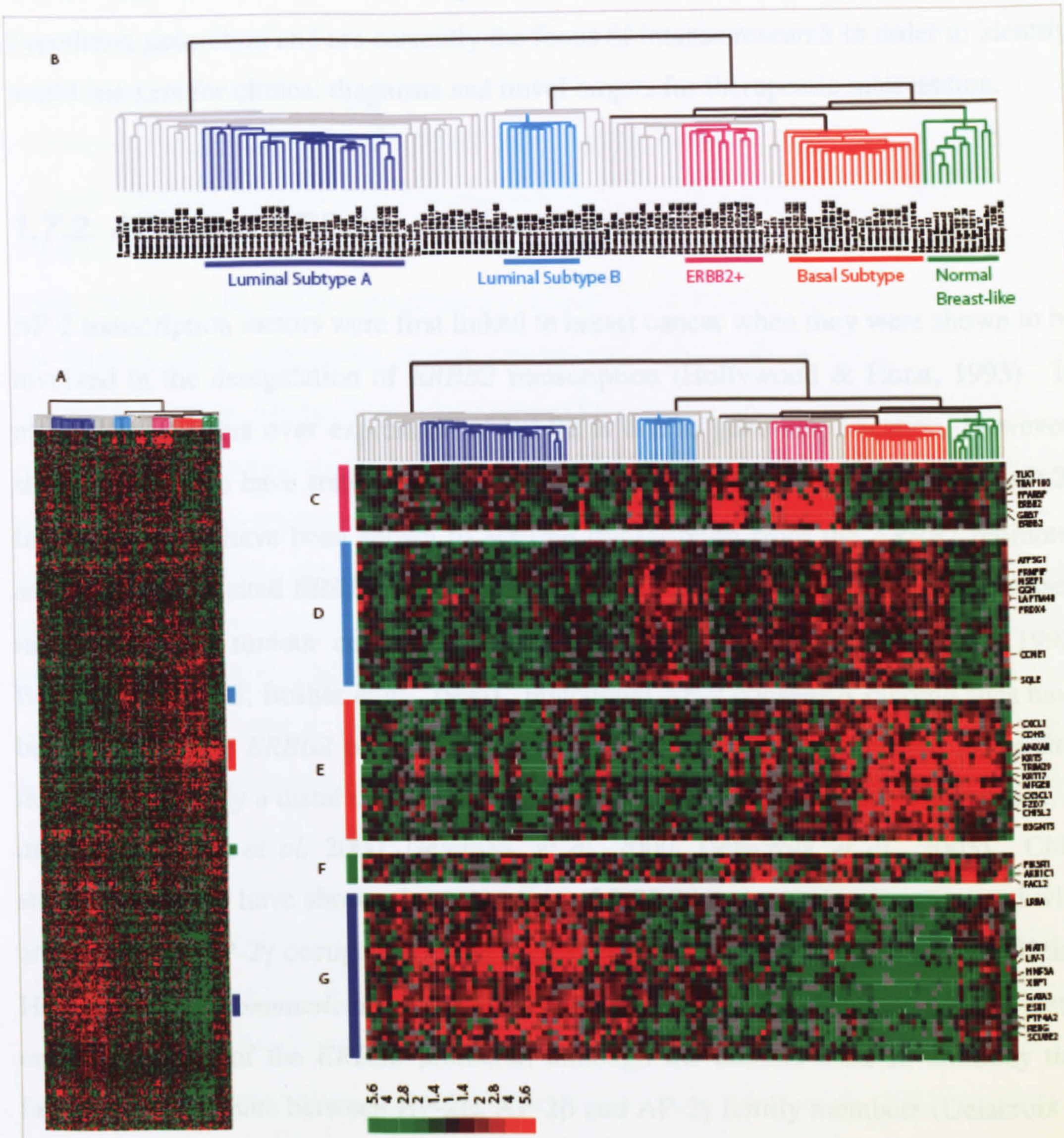


Figure 1.11. Hierarchical clustering of breast-cancer samples on basis of microarray gene expression data (Figure from Sorlie *et al.*, 2003). Heat map of all samples clustered according to intrinsic gene set (534 genes) used for analysis (A) and dendrogram (B) shows segregation of breast cancers into five major subgroups, including a tightly clustered basal-like cluster. Red colouration for a given gene (listed vertically along right hand side of heat map) denotes relative up-regulation of expression, and green colouration, relative down-regulation in tumours compared to normal controls. Heat maps of gene clusters define the *ERBB2* (C), luminal subtype B (D), basal-like (E), normal breast-like (F), and luminal subtype A (G) subgroups of breast cancer.

homogeneity in transcriptional programs, suggests common mechanisms which might initiate and sustain breast tumourigenesis. These studies have proved useful in hypothesis generation and are currently the focus of intense research in order to identify useful markers for clinical diagnosis and novel targets for therapeutic intervention.

1.7.2. AP-2 and ERBB2: cell line studies

AP-2 transcription factors were first linked to breast cancer when they were shown to be involved in the deregulation of *ERBB2* transcription (Hollywood & Hurst, 1993). In most breast cancers over expression of ErbB2 is due to gene amplification. However, single copies also have transforming potential (Hurst, 2001). Both AP-2α and AP-2γ family members have been shown to activate transcription from the *ERBB2* promoter and the closely related *ERBB3* promoter *in vitro* and AP-2γ, in particular, was expressed in all mammary tumour cell lines over expressing ErbB2 (Skinner & Hurst, 1993, Bosher *et al.*, 1995, Bosher *et al.*, 1996). Functional AP-2 consensus binding sites have been found in the *ERBB2* proximal promoter, an enhancer element located in the first intron and recently a distal enhancer region (Hollywood & Hurst, 1993; Vernimmen, *et al.* 2003, Perissi, *et al.* 2000, Newman, *et al.* 2000; Delacroix *et al.*, 2005). ChIP studies in our lab have shown that activation of *ERBB2* transcription is associated with an increase in AP-2γ occupancy in the *ERBB2* intronic enhancer region *in vivo* (Helen Hurst, *Personal communication*). AP-2 occupancy has also been observed at the distal enhancer region of the *ERBB2* promoter, although the authors used an antibody that failed to discriminate between AP-2α, AP-2β and AP-2γ family members (Delacroix *et al.*, 2005).

1.7.3. AP-2, ErbB2 and ERα: cell line studies

AP-2 family members have also been implicated in the interplay between ErbB2 and ERα in breast carcinoma. ErbB2 and ERα show an inverse correlation in breast cancer, that is, when ErbB2 is over expressed, ERα signalling pathways are down regulated. Conversely, when ERα is activated by oestrogen, ErbB2 levels are reduced. This repression can be relieved using anti-oestrogens such as Tamoxifen or Fasolodex

(ICI182,780) (Bates & Hurst, 1997; Perissi, *et al.* 2000; Orso *et al.*, 2004). The promoter of human *ERBB2* does not however contain a binding site for ERα, therefore, the oestrogenic suppression must be mediated through another indirect mechanism. Indeed, the oestrogenic regulation of *ERBB2* expression was mapped to the AP-2 consensus binding sites and additional transcription factor binding sites in the *ERBB2* intronic enhancer element and over expression of AP-2γ but not AP-2α was shown to overcome the repressive effects of oestrogen (Bates & Hurst, 1997; Newman *et al.*, 2000; Perissi, *et al.*, 2000). Interestingly, following oestrogen treatment in hormonally manipulated ERα expressing breast cancer cell lines, AP-2γ mRNA levels were increased whereas AP-2α levels decreased (Orso *et al.*, 2004). ChIP analysis of the endogenous AP-2γ promoter region observed ERα occupancy in response to oestrogen (Orso *et al.*, 2004), indicating that the balance of AP-2 family members in response to oestrogen may allow AP-2 proteins to mediate some oestrogenic responses, such as *ERRB2* repression. Indeed, following oestrogen treatment in hormonally manipulated MCF-7 cells, *ERRB2* mRNA down regulation is accompanied by a decrease in AP-2γ occupancy of the *ERBB2* enhancer region, and concurrently, a reduction in the AP-2 associated transcriptional machinery, CITED2 and CBP/p300 (Helen Hurst, *personal communication*).

Of additional interest is the ability of AP-2γ to maintain ERα transcription (reviewed in Hilger-Eversheim *et al.*, 2000). *In vitro* reporter assays have shown that AP-2γ and AP-2α can activate ERα transcription via interaction with AP-2 consensus sites within the untranslated region of exon 1 of the ERα promoter (DeConinck *et al.*, 1995; McPherson & Weigel, 1999). Adenovirus mediated over expression of AP-2α or AP-2γ was also shown to induce the formation of HS1, a DNase I hypersensitive site found in all ERα expressing breast cancer cell lines and tumours over the *ESR1* locus (Schuur, *et al.* 2001). However, although this result confirms the ability of AP-2 family members to remodel chromatin structure at the ERα locus, the expression of AP-2γ in ERα negative cell lines did not yield a detectable increase in ERα transcript levels. This was postulated to be due to other more potent repressive mechanisms acting on the ERα promoter (Schuur, *et al.* 2001). Although the role of AP-2 family members in ERα regulation requires further evidence, these data suggest the presence of a possible control loop between these factors.

1.7.4. AP-2 and ErbB2 and ERα: human breast cancer

Immunohistochemical and gene expression microarray studies performed in parallel to the cell line studies described above have looked at the relationship between AP-2, ERα and ErbB2 in tumours from breast cancer patients. AP-2α and AP-2γ were found to be expressed at high levels in 17-35% and 84% of cases respectively and, in tumours co-expressing AP-2α and AP-2γ, a positive correlation with ErbB2 status was also shown (Turner *et al.*, 1998). Two separate studies (Turner, *et al.* 1998; Gee, *et al.* 1999) showed a direct correlation between AP-2α expression and ERα expression, in primary breast cancer and in invasive ductal carcinoma. Expression profiling studies, in agreement with these data, have shown that when *TFAP2A* expression is recorded it was in ERα positive breast tumours rather than ERα negative tumours ($p<0.004$, Wang *et al.*, 2005 and $p<0.00092$, Sotiriou *et al.*, 2006). AP-2α and ERα co-expression might reflect a tumour cell origin as both are predominantly expressed in luminal epithelia of normal breast. Additionally it might be related to the potential control loop between these factors. Immunohistochemical studies examining AP-2α only in clinical breast tumour samples showed a negative correlation with ErbB2 status but a positive correlation with staining for the universal cell cycle inhibitor p21WAF1/CIP1 and a lower rate of proliferation (Gee *et al.*, 1999). AP-2α was also reduced in level from normal breast through to DCIS and primary invasive cancers. Others have shown that reduced nuclear AP-2α staining was associated with more aggressive breast cancers (Pellikainen *et al* 2002) and *TFAP2A* mRNA expression was inversely correlated with increased tumour grade ($p<0.01$, Sotiriou *et al.*, 2006). In summary, these studies indicate AP-2α expression in breast cancer is associated with favourable prognostic markers, namely ERα expression and ErbB2 negativity and reduced proliferation. Considered together with observations of AP-2α expression in normal breast (Zhang *et al.*, 2003), these data further suggest that AP-2α expression observed in breast tumours may be an attempt by the cells to arrest their growth and terminally differentiate. AP-2α may therefore be acting in a tumour suppressive role, consistent with the observation that its expression is reduced with advancing stages of breast cancer and also with its suggested role in other cancer types. The loss of AP-2 expression in invasive breast cancer has been significantly correlated with hypermethylation of a CpG island in exon 1 of AP-2α gene. Hypermethylation of this region does not occur in normal breast

epithelium and occurred in only 16% of DCIS lesions, but was present in 75% of primary invasive breast tumours. Tumours unmethylated for this region expressed AP-2α protein, whereas tumours with hypermethylation showed considerable loss, describing a mechanism of suppression of AP-2α expression *in vivo* (Douglas *et al.*, 2004).

An increasing body of evidence suggests that AP-2γ expression might play an opposite role to that of AP-2α in breast carcinomas. Although AP-2γ over expressing transgenic mice did not develop tumours, epithelial hyperplasia was observed, implicating AP-2γ in the promotion of proliferation. Studies in ErbB2/AP-2γ over expressing transgenic mice support a role for AP-2γ expression in the promotion of tumour proliferation (Jäger *et al.*, 2005). Whilst a small reduction in the overall number of tumours was observed in AP-2γ/ErbB2 transgenic mice, an increased proportion of advanced stage carcinomas were seen. Closer examination revealed AP-2γ expression was significantly correlated with increased proliferation rates in all stages of mouse mammary carcinoma.

Indeed, a recent study has shown that elevated levels of AP-2γ in advanced primary breast tumours correlate with reduced patient survival. Additionally, the relationship between AP-2γ expression and ERα or ErbB2 expression was also explored in this series of advanced primary invasive tumours (Helen Hurst, *personal communication*). A wide range of AP-2γ expression levels were found across the samples with a highly statistically significant correlation between high levels of staining and shortened survival, shown in Figure 1.12. (A). These data are supported by a second clinical study where a similar link between high levels of *TFAP2C* mRNA and reduced disease-free survival was observed (Zhoa *et al.*, 2003). Additionally, *TFAP2C* mRNA expression has been associated with advancing clinical grade in gene expression profiling studies ($p<0.002$, Sotiriou *et al.*, 2006).

Figure 1.12. (B), shows high AP-2γ was also linked to poor patient prognosis in patients positive for ERα expression, a typically good prognosis group due to their responsiveness to hormone therapy (Helen Hurst, *personal communication*). This suggests that one of the consequences of high level AP-2γ expression in tumours may be an increased propensity towards resistance to anti-oestrogen therapy. As described

earlier, AP-2γ transcription is activated by oestrogen bound ERα and anti-oestrogen abrogates this ERα mediated transcriptional activation (Orso *et al.*, 2004). Based on these observations a reduction in AP-2γ levels might be expected in ERα positive tumours undergoing anti oestrogen therapy. High AP-2γ expression therefore represents a challenge to this mechanism and thus implicates AP-2γ expression in the aetiology of anti-oestrogen resistance. A number of microarray studies have shown that *TFAP2C* mRNA is also expressed in ERα negative tumours and that this expression in general is higher in these tumours compared to ERα positive tumours ($p<0.002$, van de Vijver, *et al.*, 2002; $p<0.005$, Gruvberger *et al.*, 2001 and $p<0.00012$, Sotiriou *et al.*, 2006).

High AP-2γ expression was also linked to poor patient prognosis in patients negative for ErbB2 expression, Figure 1.12. (C). No significant prognostic outcome was contributed by high AP-2γ in patients positive for ErbB2 expression, since these patients already represented a poor prognosis subgroup (Helen Hurst, *personal communication*).

In summary, this immunohistochemical analysis (Figure 1.12), links high AP-2γ expression to poor survival, both independently and in the context of the generally considered indicators of good prognosis: namely ERα expression and ErbB2 negativity. These findings indicate that AP-2γ may be a useful prognostic indicator in breast cancer and the AP-2γ regulation of target genes other than *ERRB2* and ERα must also be important. Additionally, the poor prognosis in high AP-2γ expressing, ERα positive tumours casually links AP-2γ expression with resistance to hormone therapy. A reasonable hypothesis, given the consequences of AP-2γ over expression in normal mouse mammary glands (Jäger *et al.*, 2003) and mouse ErbB2/AP-2γ over expressing mammary gland tumours (Jäger *et al.*, 2005) is that AP-2 may promote proliferation in human breast carcinomas. Interestingly, the AP-2γ genomic locus, 20q13.2, is frequently amplified in breast cancer cell lines and breast carcinoma suggesting a possible mechanism for this aberrant expression (Williamson *et al.*, 1996), providing further evidence that high AP-2γ might be important in the aetiology of this disease.

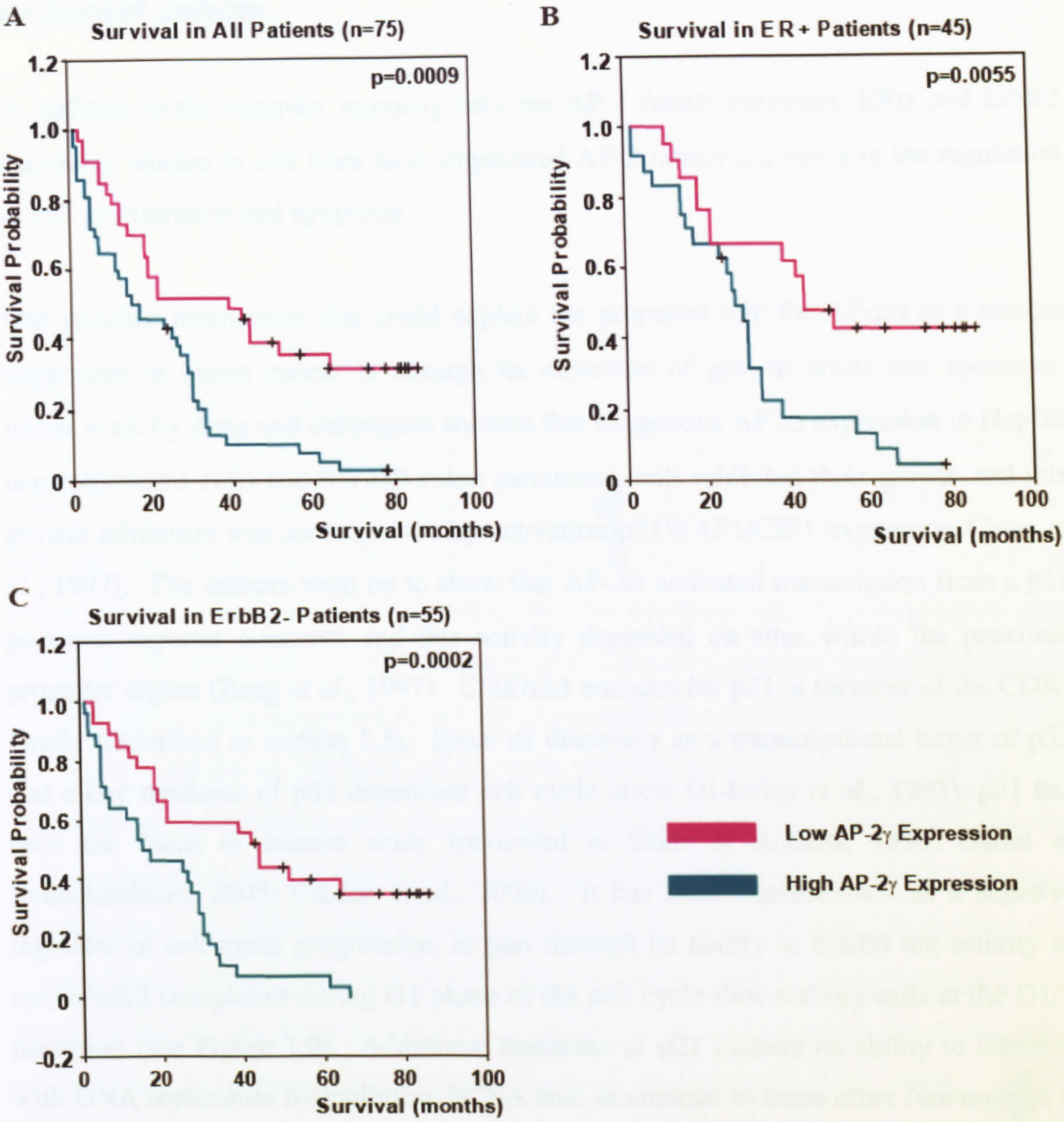


Figure 1.12. Correlation between AP-2 expression and poor prognosis in advanced primary breast tumours. A series of 75 patients with primary invasive breast cancer showed a range of intensity of nuclear AP-2γ immunostaining across the samples. In a cohort of 75 patients who presented with advanced primary tumours and were treated with surgery followed by anti-oestrogen therapy. Each sample was assigned a histopathology score (H-score) from 1-300 based on the intensity of the staining and the percentage of tumour cells stained. A statistically defined cut-off point, the median (H-score > 120), was used to define "positivity" for subsequent analysis. **(A)** For all patients, higher AP-2γ expression correlated with shortened survival time with clear curve separation after 20 months (p=0.0009). **(B)** For all ERα positive patients, higher AP-2γ expression correlated with shortened survival time (p=0.0055). **(C)** For all ErbB2 negative patients, higher AP-2γ expression correlated with shortened survival time (p=0.0002). (Helen Hurst, *personal communication*).

1.7.5. AP-2 and cell proliferation and apoptosis with reference to breast cancer.

In addition to the complex interplay between AP-2 family members, ERα and ErbB2 additional studies in cell lines have implicated AP-2 family members in the regulation of cell proliferation and apoptosis.

One possible mechanism that could explain the proposed role for AP-2α as a tumour suppressor in breast cancer is through its induction of growth arrest and apoptosis. Initial work by Zeng and colleagues showed that exogenous AP2α expression in HepG2 hepatoblastoma cells and SW480 colon carcinoma cells inhibited their growth and this growth inhibition was accompanied by activation of p21WAF1/CIP1 expression (Zeng *et al.*, 1997). The authors went on to show that AP-2α activated transcription from a p21 promoter reporter construct and this activity depended on sites within the proximal promoter region (Zeng *et al.*, 1997). *CDKN1A* encodes for p21, a member of the CDKI family (described in section 1.5). Since its discovery as a transcriptional target of p53 and a key mediator of p53 dependent cell cycle arrest (el-Deiry *et al.*, 1993), p21 has been the focus of intense study (reviewed in Sherr & Roberts, 1999; Gartel & Radhakrishnan, 2005; Caldon *et al.*, 2006). It has been characterised as a negative regulator of cell cycle progression, in part through its ability to inhibit the activity of cyclin/cdk2 complexes during G1 phase of the cell cycle thus stalling cells at the G1/S transition (see Figure 1.9). Additional functions of p21 include its ability to interfere with DNA replication by inhibiting PCNA and, in contrast to these other functions, it is also required for the stabilisation of Cyclin/CDK complexes.

A separate study has suggested that the ability of AP-2α to activate *CDKN1A* transcription is dependant on a direct interaction with p53, although the results require cautious interpretation (McPherson *et al.*, 2002). AP-2α and AP-2γ were shown to interact with p53 *in vitro* using yeast two hybrid and GST pull-down assays. The authors then showed a very small amount of endogenous AP-2 activity could be immunoprecipitated using a p53 specific antibody in MCF-7 cells. AP-2 activity was measured using DNA binding assays and controlled using an AP-2 antibody, but the antibody used cannot discriminate between AP-2α, AP-2β and AP-2γ. The AP-2

activity in these assays was attributed to AP-2α, however AP-2γ is the predominant family member expressed in MCF-7 cells whereas AP-2α levels are barely detectable (see Results Figure 3.1). In order to confirm an interaction between endogenous AP-2α and/or AP-2γ and p53, and indeed the functional relevance of this interaction, these experiments need to be revisited with appropriate immunological reagents in alternative cell lines.

Cell growth inhibition accompanied by p21 up regulation was also observed in colon carcinoma cells following adenovirus mediated over expression of AP-2α compared to the over expression of a LacZ control virus (Wajapeyee & Somasundaram, 2003). This growth arrest was dramatically reduced in HCT116 p21-/- cells compared to the wildtype counterparts suggesting that activation of *CDKN1A* transcription by AP-2α was essential for this arrest. In support of the data described above, p21 induction was shown to be enhanced by p53 as measured following AP-2α over-expression in HCT116 p53-/- cells compared to the HCT116 wildtype cells. However, p21 was still induced by AP-2α in HCT116 p53-/- cells and the level of cell cycle arrest was equivalent to the wildtype cells. These data suggest therefore that AP-2α can activate p21 independently of p53, although cooperation with p53, similarly to previous findings (McPherson *et al.*, 2002), might slightly enhance this activation. Importantly, these cell line studies complement the *in vivo* data described above which correlate expression of AP-2α in the mouse mammary gland and in human breast cancers to the cessation of proliferation and an up regulation of p21 (Zhang *et al.*, 2003; Gee *et al.*, 1999).

As well as cell cycle arrest, Wajapeyee and Somasundaram also observed an induction of apoptosis in AP-2α over expressing cells (Wajapeyee & Somasundaram, 2003). AP-2α over expression was also correlated to increased sensitivity to chemotherapeutic drugs in breast and colon cancer cell lines (Wajapeyee *et al.*, 2003; Wajapeyee *et al.*, 2005). Also, apoptosis in response to cisplatin (a DNA damaging agent), and taxol (an inhibitor of microtubule depolymerisation) in colon carcinoma cells was accompanied by an increase in AP-2α levels. This apoptotic response was shown to be inhibited by siRNA mediated AP-2α silencing, suggesting that AP-2α expression is intrinsic to the chemosensitivity of these cells (Wajapeyee *et al.*, 2003; Wajapeyee *et al.*, 2005). Recent work has shown that AP-2α can mediate apoptosis through the transcriptional

repression of the anti-apoptotic protein Bcl-2 in colon carcinoma cells (Wajapeyee *et al.*, 2006). Apoptosis in AP-2α over expressing cells was accompanied by exogenous AP-2α occupancy of the proximal promoter region of *BCL2*. Exogenous expression of Bcl-2 inhibited AP-2α induced apoptosis suggesting that *BCL2* repression is the principal mechanism for mediation of apoptosis by AP-2α (Wajapeyee *et al.*, 2006). These data contrast with the observations of AP-2 in TNFα induced apoptosis in breast carcinoma cells that express both AP-2α and AP-2γ (Nyormoi *et al.*, 2001). TNFα induced apoptosis was accompanied by a decrease in levels of AP-2, detected using the non discriminating AP-2 antibody. Furthermore, it was suggested that caspase mediated cleavage of AP-2 (which resulted in loss of its transcriptional activity) was critical for this TNFα induced apoptotic pathway (Nyormoi *et al.*, 2001). The differences in the roles for AP-2 in these studies may reflect mechanistic differences in the intrinsic apoptotic pathways activated by chemotherapy or TNFα. Equally, Wajapeyee and colleagues concentrated on the AP-2α family member whereas the other studies did not discriminate between family members. Additional studies are required to further elucidate the role of specific family members with respect apoptosis. Although a role for AP-2α expression in the induction of apoptosis is clearly supported *in vivo* by the enhanced cell death observed in AP-2α over expressing mouse mammary glands (Zhang *et al.*, 2003). Targeting AP-2γ over expression to the mammary gland also resulted in increased apoptosis (Jäger *et al.*, 2003). Interestingly work in our laboratory by Jamie Meredith, has shown that adenovirus forced over expression of AP-2γ in the AP-2 negative normal breast epithelial cell line (MTSV1.7neo) resulted in increased cell death, although the mechanism behind this phenotype could not be identified (Meredith J, 2006).

In mammary carcinoma cell lines, 5' regulatory sequences for TGFα have been shown to be positively regulated by AP-2 in luciferase reporter assays (Wang *et al.*, 1997, Gille *et al.*, 1997, and Kim & Rho, 2002). Expression of TGFα is known to be growth promoting in mammary gland development and mammary neoplasia, its effects largely exerted via EGFR signalling (reviewed by Sternlicht *et al.*, 2006). Interestingly, convincing evidence exists showing a role for AP-2 in EGFR family member transcription (described above), although a role in TGFα transcriptional regulation requires further analysis.

In vitro reporter assays also implicate AP-2α in the activation of transcription from the *IGFBP5* promoter via AP-2 DNA recognition sites (Duan & Clemmons, 1995; Erclik & Mitchell, 2005). Interestingly, over expression of AP-2γ targeted to the mouse mammary gland was accompanied by induction of IGFBP5 (Jager *et al.*, 2003). Reports have suggested that IGFBP5 might be involved in the induction of apoptosis in alveolar cells during postlactational mammary gland involution (reviewed in Beattie *et al.*, 2006). Indeed, enforced expression of AP-2γ was associated with enhanced apoptosis and it was suggested that this may have been induced via transcriptional activation of *IGFBP5* by AP-2γ. Work in our lab supports this and using ChIP assays it has been shown that AP-2γ is bound to the *IGFBP5* promoter in MCF-7 cells (Karsten Friedrich, *personal communication*), although a functional role of IGFPB5 with regards to breast cancer requires further clarification.

In summary, the cell line studies described above do not always provide a clear picture of the role of AP-2 proteins in breast cancer. Some general threads however can be drawn together. In support of a tumour suppressor-like role for AP-2α, it has been shown to up-regulate *CDKN1A* transcription and repress transcription of the anti-apoptotic gene *BCL2*. It is reasonable to consider that AP-2α and AP-2γ might share some overlapping functions in relation to breast cancer and breast development, indeed they are both expressed in luminal and myoepithelial cells albeit at varying levels (Zhang *et al.*, 2003) and over expression of both AP-2α and AP-2γ has been associated with enhanced apoptosis.

The molecular mechanisms behind the association of high AP-2γ expression and increased breast tumour progression leading to poor patient prognosis remain elusive, as does the transcriptional program initiated by AP-2γ in the promotion of cell proliferation and inhibition of differentiation in breast tissues during development. AP-2γ expression is clearly associated with the regulation of *ERBB2* and ERα in breast cancer cell lines, however high levels of AP-2γ predict poor prognosis independently of these factors in patients with advanced primary invasive breast carcinomas, suggesting other molecular mechanisms may be involved. In order to try to define the additional pathways activated by AP-2γ specifically in breast cancer, the research presented in this

thesis aims to use Affymetrix GeneChip high-density oligonucleotide microarrays to define AP-2γ dependant changes in gene expression in breast epithelial cells.

1.8. AP-2 and Gene Expression Profiling

Gene expression profiling using microarrays is fast becoming an established technology since its conception more than 10 years ago, pioneered by Brown and colleagues who established cDNA microarrays (Schena *et al.*, 1995) and Affymetrix (Santa Clara, California, USA) who have built a considerable business around their oligonucleotide microarray technology. A microarray is an array of nucleotide features, representing tens of thousands of known genes and uncharacterised partial gene sequences, immobilised to a solid substrate. Labelled nucleic acid material derived from an RNA sample of interest can be hybridised to the nucleotide features. The amount of label detected at each feature corresponds to the level of each specific RNA message expressed in the sample. This technology enables thousands of nucleic acid hybridisation events to be investigated in a single experiment.

The major challenge, in terms of AP-2 expression profiling in breast cancer, has been to determine the best strategy to derive two RNA populations to compare to identify AP-2 regulated genes. Initial efforts concentrated on comparing cells with low AP-2 expression with derivatives engineered to express AP-2γ at levels similar to those observed in over-expressing breast tumour lines. This approach identified some interesting genes potentially regulated by AP-2 (Helen Hurst, *personal communication*). However, it is likely that the overexpression of AP-2γ in a null line may not effectively activate all the key genes since DNA methylation and epigenetic chromatin modification will silence many genes in a manner that might not be reversed by over expression of AP-2γ alone. Tissues derived from knockout mice have been used to examine expression profiles due to AP-2α (Werling & Schorle, 2002), but this has so far not been possible for AP-2γ due to the early embryological death of the null mice. Expression profiles of wild type and AP-2α null mouse embryo fibroblasts were disappointingly similar (Helen Hurst, *Personal communication*), probably because AP-2 factors are not significantly expressed in fibroblasts.

Importantly, the ability of short interfering RNA (siRNA) to mediate sequence-specific gene silencing in mammalian cells (Elbashir *et al.*, 2001) opens up the possibility of circumventing the majority of these problems by knocking down AP-2γ expression in AP-2γ over expressing breast tumour lines.

1.9. Short Interfering RNA (siRNA)

The biological process of RNA interference (RNAi) was first discovered by Fire and colleagues who showed that the injection of long dsRNA species into *C. elegans* led to sequence-specific degradation of the corresponding mRNAs (Fire *et al.*, 1998). This process was subsequently found to be conserved across eukaryotes and has greatly influenced our understanding of gene regulation in animals, plants and many fungi (Reviewed in Dykxhoorn *et al.*, 2003). However, initially this process could not be applied to mammalian cells, as the introduction of long dsRNAs (>30 nucleotides) were known to elicit an interferon response (Minks *et al.*, 1979).

The interferon system is an innate antiviral immune response, which can be triggered by dsRNA, a common viral replicative intermediate (reviewed in Samuel *et al.*, 2001). The response is mediated by the JAK-STAT signalling cascade and culminates in the induction of interferon-stimulated genes. These include 2'-5' Oligoadenylate Synthase, which in turn activates RNase L and triggers global degradation of mRNA, and the protein kinase, PKR, which can phosphorylate the translation initiation factor eIF2 leading to a global inhibition of mRNA translation. The ultimate aim of this response is to prevent viral multiplication but in the process it can affect growth and cause damage to the host cell.

The discovery that long dsRNA enabled the effective silencing of gene expression by presenting various short, 19-23 nucleotide, RNA cleavage fragments to the target mRNA was exploited by Elbashir and colleagues. They showed that direct introduction of these short RNAs into mammalian cells, could cause target mRNA degradation without triggering the interferon response (Elbashir *et al.*, 2001). They went on to show that these short interfering RNAs (siRNAs) allowed the study of gene-specific phenotypes in mammalian cultured cells (Harborth *et al.*, 2001). Since these early

experiments more about the mechanism of RNAi has been ascertained. It is thought to occur by one of two pathways in the cell, summarised in Figure 1.13: the siRNA pathway or the micro RNA pathway.

MicroRNAs are endogenous 22 nucleotide RNAs that elicit either translational repression or mRNA cleavage (reviewed in Bartel, 2004). Typically encoded in non-protein coding regions of plant and animal genomes, microRNAs are emerging as an important post-transcriptional gene silencing mechanism in eukaryotes and have been shown to, amongst other processes, regulate neuronal patterning in nematodes (Johnston & Hobert, 2003) and modulate haematopoietic lineage differentiation in mammals (Chen *et al.*, 2004). A third mechanism of transcriptional gene silencing via RNAi has been observed. In fission yeast, expressed genes can be direct targets of RNAi dependent heterochromatin formation (Schramke & Allshire, 2003). It has also been shown that siRNA induced transcriptional gene silencing can occur in mammalian cells (reviewed in Kawasaki *et al.*, 2005). siRNAs targeted to CpG islands within a gene promoter can induce RNA-directed DNA methylation.

It is important to stress the differences between the siRNA and microRNA processes. siRNAs are thought to require a 100% identity with the targeted transcript to cause mRNA degradation; a single base pair change has been shown to abolish ability of an siRNA to bind and degrade its specific target (Brummelkamp *et al.*, 2002a). On the other hand microRNAs tolerate more sequence diversity and therefore can silence related sequences, primarily by translational repression, leading to loss of protein levels while mRNA levels are maintained (reviewed in Bartel, 2004).

1.9.1 Short Interfering RNA Delivery

Two main methods have developed to utilise RNAi via the siRNA pathway in mammalian cells. Firstly, siRNAs have been introduced into mammalian cells by transient transfection of synthetic double-stranded RNA (Elbashir *et al.*, 2001). The major drawback of this method is the transient nature of the silencing effect. This becomes important when trying to reduce the level of proteins with a long half-life or when an RNAi specific phenotype only appears after an extended period of time

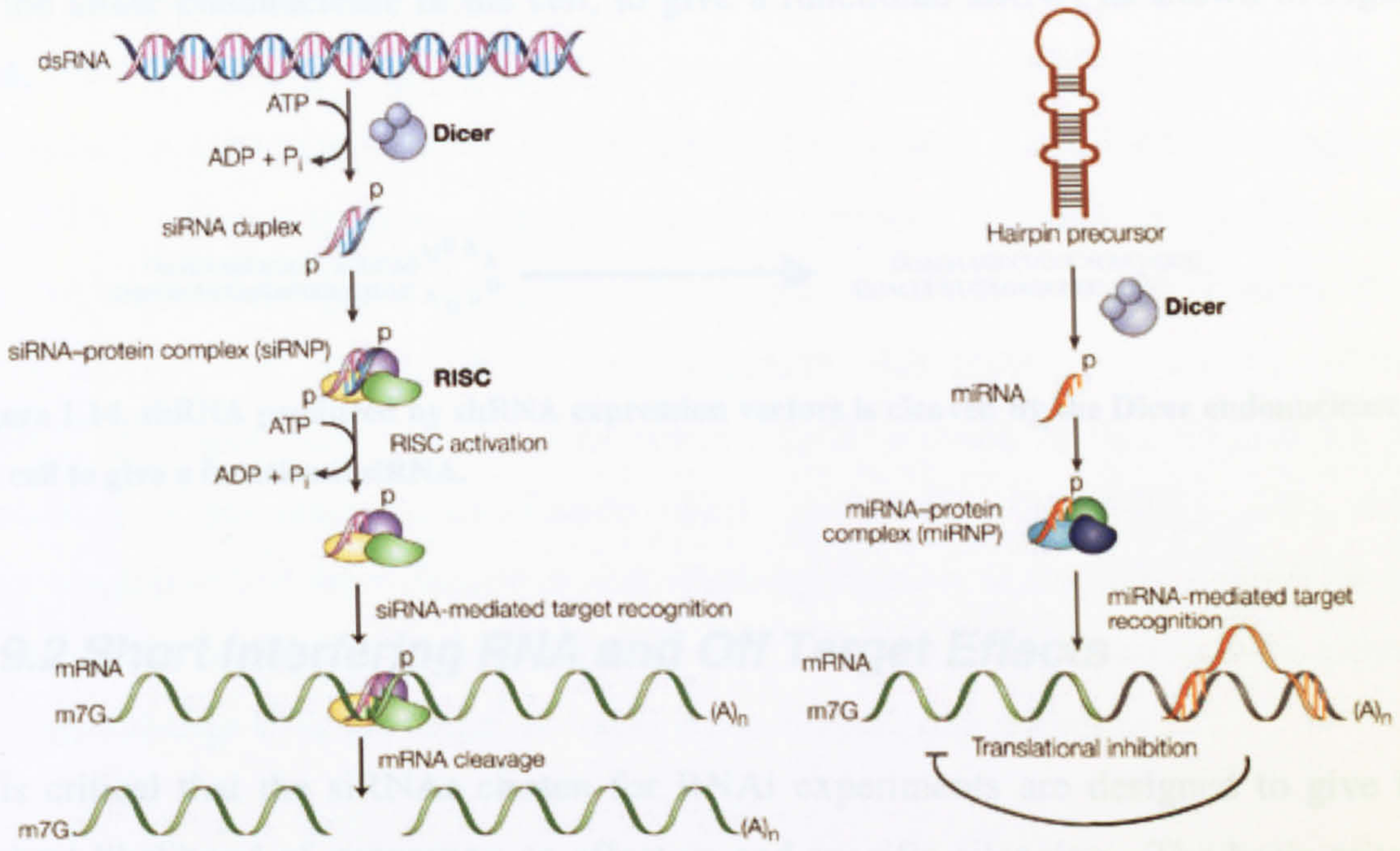


Figure 1.13. The proposed mechanisms of RNA interference in mammalian cells (Adapted From Dykxhoorn *et al.*, 2003). The siRNA pathway. Long double-stranded RNA is cleaved by the endoribonuclease Dicer, into siRNAs in an ATP-dependent reaction. siRNAs are then incorporated into the RNA-inducing silencing complex (RISC). The siRNA duplex is unwound in an ATP dependant manner. The single-stranded antisense strand guides RISC to messenger RNA that has a complementary sequence, resulting in the endonucleolytic cleavage of the target mRNA. **The miRNA pathway.** The endoribonuclease Drosha first processes the endogenous pre-miRNA transcripts into 70-nucleotide hairpin pre-miRNAs inside the nucleus. Once exported to the cytoplasm Dicer then cleaves these precursors to produce single stranded 22 nucleotide miRNA. These are incorporated into a ribonucleoprotein complex and are guided to messenger RNA with a partial complementary sequence and result in a repression of mRNA translation via obstruction of the translational machinery.

(Reviewed in Huppi *et al.*, 2005). To overcome these problems siRNA expression vectors have been developed that can be stably introduced into cells either as selectable plasmids or as viral vectors (Brummelkamp *et al.*, 2002a; Brummelkamp *et al.*, 2002b; Paddison *et al.*, 2002). These vectors utilise the RNA polymerase III promoters of the H1 or U6 genes to produce a short RNA hairpin. The loop of the hairpin is then cleaved by the Dicer endonuclease in the cell, to give a functional siRNA as shown in Figure 1.14.

Figure 1.14. shRNA produced by shRNA expression vectors is cleaved by the Dicer endonuclease in the cell to give a functional siRNA.

1.9.2 Short Interfering RNA and Off Target Effects

It is critical that the siRNAs chosen for RNAi experiments are designed to give the highest likelihood of generating an effective and specific silencing. The basic criteria outlined by Tuschl and colleagues suggested: searching the open reading frame of the target sequence for 23 nucleotide motifs $AA_{(N19)}TT$ (approx. 50% GC content). It was emphasised that selected sequences should be subject to BLAST analysis against EST libraries to ensure that only one gene is targeted (Tuschl *et al.*, 1999). More recently there have been reports that specific nucleotide combinations produce more effective siRNAs, although the mechanistic significance of these findings has not been established (Reynolds *et al.*, 2004; Ui-Tei *et al.*, 2004).

Despite the reported sequence-specific nature of the siRNA-mediated RNA interference many reports have highlighted the potential for non-specific effects. Two forms of off-target effects are proposed to exist: sequence-specific and sequence-independent. Firstly, a sequence-specific effect of degradation of mRNA species other than the target, due to cross-hybridisation has been reported. Systematic studies using microarrays showed that similarities as low as 11 contiguous bases could cause off-target effects (Jackson *et al.*, 2003). However, a similar study exploring the same question saw no such effects (Chi *et al.*, 2003). Other postulated sequence-specific effects include: the

binding to cellular proteins in a sequence-specific manner via an RNA aptamer effect, effecting protein function (Shi *et al.*, 1999) and inhibition of mRNA translation through a micro RNA effect (Doench *et al.*, 2003). Also, unexpectedly, it has been reported in the literature that both shRNA vectors and transiently transfected siRNAs can cause sequence-independent responses in cells via the interferon and PKR pathways described in section 1.9. (Sledz *et al.*, 2003; Sledz *et al.*, 2004; Bridge *et al.*, 2003; Persengiev *et al.*, 2004). Genome wide studies have shown that these off-target effects can be minimised in a concentration dependant manner (Semizarov *et al.*, 2003; Persengiev *et al.*, 2004). Also interferon induction can be monitored by looking for changes in interferon-stimulated genes during siRNA experiments (Sledz *et al.*, 2003; Sledz *et al.*, 2004; Bridge *et al.*, 2003). It has been suggested that using a number of different siRNA sequences targeting the same gene should provide further control for the specificity of any observed knockdown effect (Downward *et al.*, 2004). Also, it has been suggested that the inclusion of non-silencing siRNAs as a further control for off-target effects is imperative (Huppi *et al.*, 2005). These latter could contain either a base-pair change from the target, or target a non-mammalian sequence.

1.10. Aims of this study

In order to establish which cellular pathways have become activated by AP-2γ in breast cancer this project aims to generate an AP-2γ dependant expression profile using oligonucleotide microarrays and RNA interference technologies. This study will particularly concentrate on the AP-2γ family member, a protein that has been linked with tumour progression and poor prognosis in patients. My aim was to do this in the following ways:

- Establish a robust system for the selective knock-down of AP-2γ using RNAi technology, including the identification of a number of distinct targeting sequences.

- Initial work to focus on the MCF-7 mammary carcinoma cell line, chosen as it exclusively expresses AP-2γ at high levels. Upon RNAi mediated depletion of AP-2γ in MCF-7 cells, effectively all AP-2 activity would be removed from the

cells thus avoiding a further level of complication that might be encountered in other cell lines where AP-2α and AP-2γ are expressed together. MCF-7 cells also express high levels of ERα together with AP-2γ, a combination best thought to reflect the clinical data showing high levels of AP-2γ correlated with poor patient prognosis in ERα positive primary breast carcinomas.

- Use Affymetrix GeneChip microarrays to compare the wild type over-expressing lines with derivatives "knocked down" for AP-2γ expression, ultimately identifying genes whose expression levels changed significantly in the knock down derivatives.

- Validate selected genes in MCF-7 cells and, where possible, additional mammary cell lines (using real time PCR, Western blotting and ChIP).

CHAPTER 2: MATERIALS AND METHODS

2.1. Cell Culture

2.1.2. Mammary Lines

Cell lines were cultured as described in Table 2.1. All cells were regularly passaged to maintain exponential growth. In addition to mammary cell lines, the non AP-2γ expressing HEPG2 hepatic carcinoma cell line was also used in this study and maintained in DMEM with 10% Foetal Calf Serum at 10% CO_2. All cells were regularly screened for *Mycoplasma sp.* contamination.

Cell Line	Description	Media
MCF-7	Mammary Carcinoma	**DMEM, 10% Foetal Calf Serum, 10μg/ml Insulin.**
ZR75-1	Mammary Carcinoma	RPMI, 10% Foetal Calf Serum.
T47D	Mammary Carcinoma	DMEM, 10% Foetal Calf Serum.
MCF10A	Immortalised Normal Mammary Epithelial Cells	DMEM: HAMS F12 (1:1), 5% Horse Serum, 10μg/ml Insulin, 5μg/ml Hydrocortisone, 100ng/ml Cholera Toxin, 20ng/ml Epidermal Growth Factor.
MCF-7 EcoR	Mammary Carcinoma. Stably expressing mouse retroviral receptor, EcoR.	As parental Line. 500μg/ml neomycin (G418)

Table 2.1. List of cell lines and cell culture conditions. MCF-7, T47D, and MCF10A were grown in 10% CO_2 at 37°C. ZR75-1 were grown in 5% CO_2 at 37°C. MCF-7 EcoR cell lines were kindly provided by Dr Subham Basu.

2.1.3 Phoenix (ϕNX) Ecotropic Cells

Phoenix ecotropic cells are a high titre, helper virus free, retrovirus producing line based on 293T cells (Pear *et al.*, 1993). Phoenix cells stably carry a Maloney Gag-Pol (Gag Pol-IRES-Lyt2) construct under the control of an RSV promoter and separately the Maloney ecotropic envelope gene driven by a CMV promoter. Phoenix cells were grown in DMEM supplemented with 10% foetal calf serum in 10% CO_2 at 37°C. Phoenix cells were regularly passaged to maintain exponential growth.

2.2. RNA Interference

2.2.1. siRNA Selection

Initially sequences specific to AP-2γ were selected for the generation of siRNA oligonucleotides based on published suggestions (Tuschl *et al.*, 1999). Additional sequences specific to AP-2γ were based on published patterns (Reynolds *et al.*, 2004; Ui-Tei *et al.*, 2004). BLAST searches were performed to verify the AP-2γ specificity. A non-specific random sequence was used for a non-silencing control (Qiagen) These siRNA target sequences are detailed in Table 2.2. Four additional sequences specific to AP-2γ were ordered from Dharmacon and are detailed in are detailed in Table 2.3.

Target	Name	Target Sequence 5' - 3'
AP-2γ	AP-2γ_1	AATGAGATGGCAGCTAGGAAG
	AP-2γ_2	GCGGCCCAGCAACTGTGTAAA
	AP-2γ_3	CCACACTGGAGTCGCCGAATA
	AP-2γ_4	GCTCATGTGACTCTCCTGACA
	AP-2γ_5	CTGCACGATCAGACAGTCATT
AP-2γ Mismatch	Mismatch1	TGAGATGGCAGCTAAGAAG
	Mismatch1	GGCCGAGCAACTGTGTAAA
	Mismatch3	ACACTGGAGTCACCGAATA
Non Silencing Control siRNA	-	AATTCTCCGAACGTGTCACGT

Table 2.2. siRNA Target Sequences. The non-specific random sequence (Qiagen) was used for a non-silencing control for transient transfection. Those labelled mismatch have single base changes (underlined) compared to AP-2γ_1, AP-2γ_2 and AP-2γ_3, therefore should not knock down the AP-2γ target mRNA and were only used as non-silencing controls in pRetroSuper vectors (Figure 3.1).

Catalogue Number	Name	Target Sequence 5' - 3'
J-005238-05	#05	AACCGATAATGTCAAGTACGA
J-005238-06	#06	GGACACTGGAGTCGCCGAATA
J-005238-07	#07	TAGTAAACCAGTGGCAGAATA
J-005238-08	#08	TTGGACAAGATTGGGTTGAAT

Table 2.3. AP-2γ siRNA targeting sequences from Dharmacon. The #06 sequence is identical to AP-2γ_2 shown in Table 2.2.

2.2.2. Transient Transfection of siRNA oligonucleotides

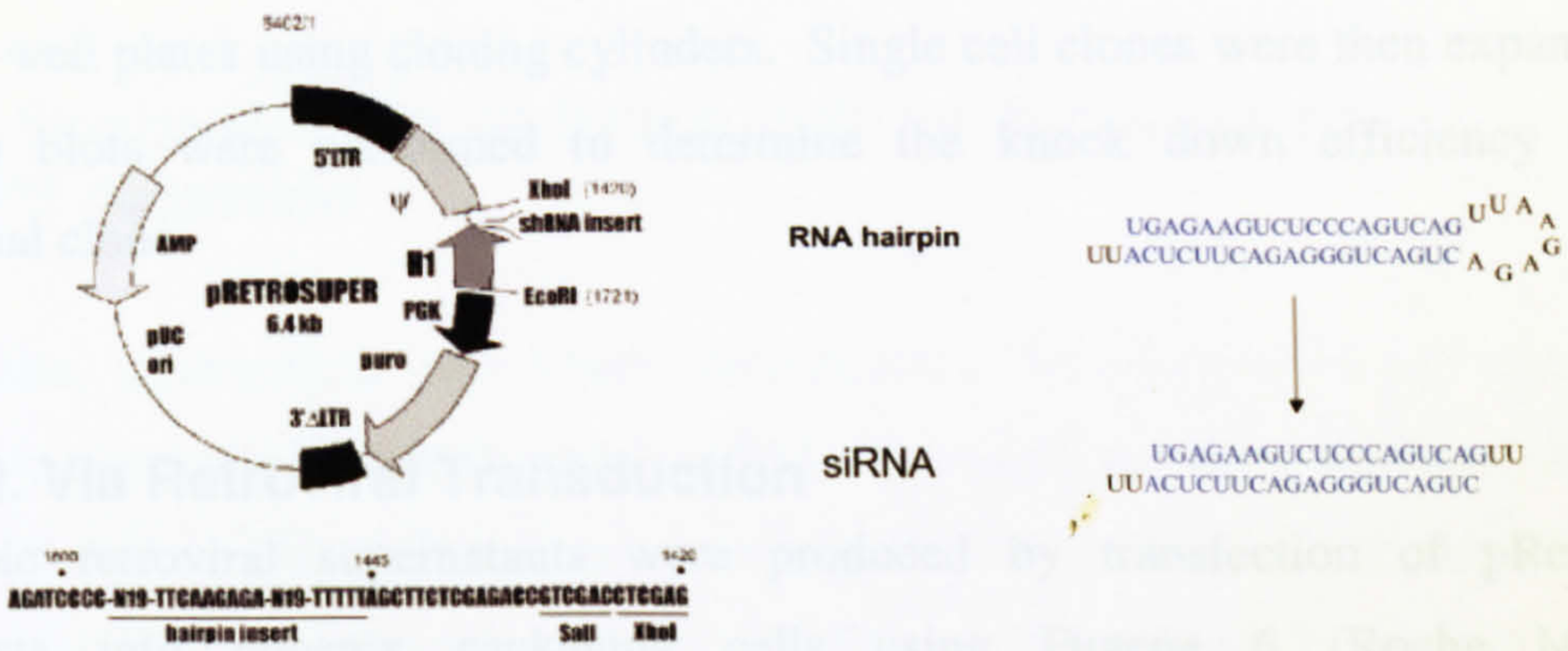
siRNAs were obtained from Qiagen or Dharmacon. Synthetic siRNAs consisted of a duplex with two 19-nucleotide RNA stands, each containing a two nucleotide overhand on the 3' end. The 21-nucleotide target sequences beginning with AA had dTdT overhangs on both strands and those starting other bases (XY) had a dTdT on the sense strand and dXdY overhangs on the antisense strand. siRNAs were diluted in the manufacturer's buffer and annealed by a 1 minute incubation at 90°C followed a 60 minute incubation at 37°C, before storage at -20 °C.

MCF-7 and ZR75-1 cells at 30 – 40 % confluency were transfected in six well plates using Oligofectamine as per the manufacturer's instructions (Invitrogen). Briefly the siRNA was diluted in DMEM to a volume of 185µl. At the same time 3µl of Oligofectamine reagent was diluted in 12 µl of DMEM, and allowed to equilibrate for 5-10 minutes. The diluted transfection reagent was then added the diluted siRNA, mixed by inversion and incubated at room temperature for 20 minutes. Following two washes of cells in DMEM, 800µl of DMEM was added per well. The transfection complexes were then added to the cells and incubated at 37 °C. Four hours later 500µl of 3x Complete Medium (DMEM + 30% FCS + 30µg/ml Insulin) was added per well, and AP-2γ knock down was assayed at time points thereafter. All the siRNA concentrations referred to in this report represent the final siRNA concentration in a 1ml, the volume in which transfection complexes were incubated on the cells for 4 hours.

MCF10A cells at 30 – 40 % confluency were transfected in six well plates using Interferin as per the manufacturer's instructions (PolyPlus). Briefly, the siRNA was diluted in DMEM:HAMS F12 (1:1) to a volume of 200µl. 8µl of Interferin transfection reagent was added to the diluted siRNA and incubated at room temperature for 15 minutes. The transfection mix was then aliquoted onto the cells which were supplemented with 2ml of their normal complete media and AP-2γ knock down was assayed at time points thereafter.

2.2.3. Short Hairpin RNA Vector Construction

Short hairpin RNA expressing constructs were based on the pRetroSuper vector and details are shown in Figure 2.1. (Brummelkamp *et al.*, 2002a, 2002b). pRetroSuper constructs were made for AP-2γ using the AP-2γ_1, AP-2γ_2 or AP-2γ_3 target sequences and the corresponding mismatch non-silencing control constructs (mismatch1, mismatch2 or mismatch3). The correct sequence of inserts was confirmed before use. The resulting siRNAs were identical to the RNA oligonucleotides used for transient transfection.



AP-2γ shRNA_1

5'GATCCCCTGAGATGGCAGCTAGGAAGTTCAAGAGACTTCCTAGCTGCCATCTCATTTTTGGAAA 3'

5'AGCTTTTCCAAAAATGAGATGGCAGCTAGGAAGTCTCTTGAACTTCCTAGCTGCCATCTCAGGG 3'

AP-2γ shRNA_2

5'GATCCCCGGCCCAGCAACTGTGTAAATTCAAGAGATTTACACAGTTGCTGGGCCTTTTTGGAAA 3'

5'AGCTTTTCCAAAAAGGCCCAGCAACTGTGTAAATCTCTTGAATTTACACAGTTGCTGGGCCGGG 3'

AP-2γ shRNA_3

5'GATCCCCACACTGGAGTCGCCGAATATTCAAGAGATATTCGGCGACTCCAGTGTTTTTTGGAAA 3'

5'AGCTTTTCCAAAAAACACTGGAGTCGCCGAATATCTCTTGAATATTCGGCGACTCCAGTGTGGG 3'

Figure 2.1. The pRetroSuper vector for expression of short interfering RNA. Custom oligonucleotide pairs were annealed then ligated into the BglII/HindIII sites of the vector. Each inserted region contained a 19nt sequence derived from the target sequence, separated by a short spacer from the reverse complement of the same 19nt sequence. The H1 promoter transcribes a short RNA hairpin, which is cleaved to give a siRNA. In order to make the mismatch shRNA controls, Oligos were designed to contain the appropriate base changes within the 19nt target sequences (see underlined in Table 2.2).

2.2.4. Generation of shRNA Stable Lines

2.2.4.1. Via Transfection

AP-2γ targeting constructs (shRNA1, shRNA2 or shRNA3), corresponding mismatch non-silencing control constructs (mismatch1, mismatch2 or mismatch3) and an empty pRetroSuper vector were linearised using the Sap1 restriction endonuclase before transfection. All cell lines were transfected using FuGENE 6 following the manufacturer's instructions. Forty-Eight hours after transfection, cells were split one in three into selection medium containing 2.5μg/ml puromycin. Media was changed every 2-3 days, and cells were maintained until all cells in the untransfected control had died (approximately 2 weeks). Single colonies of sufficient cell numbers were transferred onto 96-well plates using cloning cylinders. Single cell clones were then expanded and Western blots were performed to determine the knock down efficiency of each individual clone.

2.2.4.2. Via Retroviral Transduction

Ecotropic retroviral supernatants were produced by transfection of pRetroSuper constructs into phoenix packaging cells using Fugene 6 (Roche Molecular Biochemicals). Phoenix cells were split into selection medium containing 2μg/ml puromycin 48 hours following transfection and surviving cells were pooled for the long-term stable production of retrovirus. A day prior to the harvest of retroviral supernatants the media was changed to media without puromycin. The following day the virus-containing medium was removed from the Phoenix cells and passed through a 0.45μm filter. Supernatants were then concentrated for virus particles using Centriplus (YM-100) centrifugal filter devices. An initial volume of 20ml virus containing supernatant was applied to the filter device and centrifuged using a swinging-bucket rotor adapter at 2,500g for one and half hours. A final volume of 4ml virus containing supernatant was then snap-frozen in liquid nitrogen and stored at -80°C. The viral supernatant was used for infection of cells after addition of 5μg/ml polybrene. Cells at 30 – 40 % confluency on 6cm plates were incubated with 2 ml of viral supernatant. After 6 hours incubation cells were allowed to recover for 24 hr with fresh medium. Infected cells were selected with puromycin 2.5μg/ml and single colonies expanded as described above.

2.3. Western Blotting

2.3.1. Whole Cell Extracts

Cells were washed twice with ice-cold PBS and lysed on the plate in Urea whole cell extract buffer (8M Urea, 1M Thiourea, 0.5% CHAPS, 24mM Spermine, 50mM DTT). Extracts were collected using cell scrapers. The approximate protein content of each extract was determined using the Bradford Assay (Biorad Reagent) and a BSA standard curve ranging from 0 – 20µg/ul. Typically 5µg of whole cell extracts were used in SDS PAGE.

2.3.2. Concentration of tissue culture supernatant for secreted IGFBP3 detection

Conditioned supernatants were briefly centrifuged at 4°C, to remove cell material and then 5x concentrated using Centriplus (YM-10) centrifugal filter devices. An initial volume of 2ml conditioned supernatant was applied to the filter device and centrifuged at using a swinging-bucket rotor adapter at 2,500g for approximately one and half hours at to a final volume of 400µl. 10µl of this 5x concentrated supernatant was used in SDS PAGE.

2.3.3. SDS PAGE and Western Blotting

Samples were separated on a 6-12% SDS PAGE gels (National Diagnostics) prepared with a BioRad Mini Protean system. Full range molecular weight Rainbow Markers (Amersham) were loaded to allow size determination of detected proteins. Separated proteins were transferred onto a nitrocellulose membrane (Amersham) using a semi-dry blotting apparatus (Biorad). The membrane was blocked (5% Marvel 0.1% Tween-20 in PBS) for an hour at room temperature. The primary antibody was diluted in blocking solution and incubated with the membrane for an hour at room temperature. Three 10-minute washes with blocking solution, removed excess primary antibody before incubation with the appropriate secondary (1 hour at room temperature). The membranes were washed in blocking solution twice and a further 3 times in PBS / 0.1%

Tween-20 to remove excess secondary antibody and Marvel. Supersignal WestFemto reagents (Pierce) were used for chemiluminescence and the membrane was exposed to Kodak BMR autoradiograph film. Where necessary, membranes were then stripped for 15 minutes at 55°C in strip buffer (1M Tris HCl pH6.7, 2% SDS, 100mM β-Mercaptoethanol), re-blocked for 1 hour at RT and finally re-probed for Ku-70 as a loading control. A summary of the antibodies used and the appropriate dilutions is shown in Table 2.4.

2.4. Immunofluorescence

Coverslips were cleaned in 40% concentrated HCl / 60% ethanol for 30 minutes, then washed three times in distilled water and stored in 70% ethanol. One coverslip was placed in each well of a 6 well plate, and $3x10^5$ cells were plated for analyses 24 hours later. Cells were washed twice in PBS then fixed in 4% paraformaldehyde solution for 20 minutes at room temperature. Cells were permeablised in 0.1% TritonX 100, PBS for 5 minutes at 4°C. The Kiaa1324 antibody was diluted 1:500 in PBS and added to the plated cells for 30 minutes at room temperature. Three 10-minute washes with PBS removed excess primary antibody. Goat anti-rabbit 564nm Alexaflour conjugate was diluted 1:1000 in 5% heat inactivated goat serum PBS and incubated with the plated cells for 30 minutes in the dark. Three 10-minute washes with PBS followed to remove the excess secondary antibody. 20μg/ml 4'6-diamidino-2-phenylindole, dilactate (DAPI, Molecular Probes) in PBS was added to the cells, incubated at room temperature for 5 minutes and washed twice with PBS. Stained cells were then put on glass slides using Permafluor mounting medium (Thermo Shandon). The slides were left in the dark overnight at 4°C to let the mounting medium set. Photos were taken with an Olympus BX51 with a 20x objective.

Antigen	Animal origin	Source	Dilution
AP-2γ (6E4/4)	Mouse	Helen Hurst	1:1000
AP-2α (3B5)	Mouse	Santa Cruz	1:1000
BAX (N20)	Rabbit	Santa Cruz	1:500
CITED2 (MRG1)	Mouse	Santa Cruz	1:500
CITED4 (KH104)	Mouse	S. Bhattacharya	1:500
CBP (A-22)	Rabbit	Santa Cruz	1:500
CDK2	Mouse	Santa Cruz	1:500
CDK4(DCS156)	Mouse	Cell Signalling	1:1000
CDK6 (DCS83)	Mouse	Cell Signalling	1:1000
Cyclin D1 (DCS6)	Mouse	Cell Signalling	1:1000
Cyclin D3 (DCS22)	Mouse	Cell Signalling	1:1000
Cyclin E2 (HE12)	Mouse	Santa Cruz	1:1000
ERα (HC-20)	Rabbit	Santa Cruz	1:500
IGFBP3 (#06108)	Rabbit	Upstate	1:200
Inhibin βA	Mouse	R&D systems	1:10
Keratin 8 (35βH11)	Mouse	DAKO	1:1000
Keratin 18 (DC10	Mouse	DAKO	1:1000
Kiaa1324	Rabbit	L. Deng	1:1000
Ku-70 (C19)	Goat	Santa Cruz	1:1000
Myc (9E10)	Mouse	Santa Cruz	1:500
p21 (DCS60)	Mouse	Cell Signalling	1:1000
p27 (# 2552)	Rabbit	Cell Signalling	1:1000
p53 (DO1)	Mouse	Santa Cruz	1:1000
p300 (N-15)	Rabbit	Santa Cruz	1:500
phospho-p27 (thr187) (2B10B7)	Mouse	Invitrogen	1:1000
phospho-Rb (ser807/811)	Rabbit	Cell Signalling	1:1000
phospho-Rb (ser795)	Rabbit	Cell Signalling	1:1000
SP1 (PEP2)	Rabbit	Santa Cruz	1:500
STAT-1 (C136)	Mouse	Santa Cruz	1:1000
Transgelin (ab14106)	Rabbit	Abcam	1:1000
HRP linked Mouse IgG	-	Amersham Pharmacia	1:2500
HRP linked Rabbit IgG	-	Amersham Pharmacia	1:2500
HRP linked Goat IgG	-	Santa Cruz	1:2500

Table 2.4. List of antibodies and their dilutions for western blotting.

2.5. Flow Cytometry

Flow cytometry analysis was performed with the assistance of Derek Davies and colleagues at the CRUK FACS Laboratory, London Research Institute, Lincoln's Inn Fields.

2.5.2. Bromodeoxyuridine Staining

Cells were treated with 10μM BrdU for 8 hours. Cells were removed using trypsin, washed with PBS, and the pellet was resuspended in ice cold 70% ethanol to fix the cells. Ethanol was discarded after centrifugation at 2000rpm and cells were treated with 2M HCl for 30 minutes at room temperature. Cells were then washed twice in PBS then once in PBS / 0.1% BSA 0.1% Tween-20. Anti-BrdU antibody (Becton Dickinson) was added directly to the cell pellet and incubated in the dark, for 20 minutes at room temperature. Cells were then washed twice in PBS / 0.1% BSA 0.1% Tween-20. FITC-conjugated rabbit anti-mouse F(ab')2 fragments (DAKO) were incubated for 20 minutes at room temperature in the dark, followed by one wash in PBS. Treatment with 100μl of RNase (100μg/ml) for 5 minutes at room temperature ensured only DNA would be stained. Propidium Iodide (PI) (50μg/ml) was then added to the cells and left for 15 minutes at room temperature. Flow cytometry was performed using a FACS Calibur (Becton Dickinson). Fluorochromes were excited by a 488nm laser and fluorescence was collected between 515 and 545nm for FITC and above 580nm for PI. Forward and right angle scatter was used to define the cellular population and pulse processing of the PI signal was used to distinguish true G2 cells from G1 doublets.

2.5.3. Propidium Iodide staining for Sub-G1 apoptosis analysis

Cells were removed using trypsin, washed with PBS, and the pellet was resuspended in ice cold 70% ethanol to fix the cells. The media was collected to capture any floating cells for the analysis. Ethanol was discarded after centrifugation at 2000rpm and cells were washed twice with phosphate citrate buffer (0.2M Na2HPO4 / 0.1M Citric acid). Treatment with 100μl of RNase (100μg/ml) for 5 minutes at room temperature ensured only DNA would be stained. PI (50μg/ml) was then added to the cells and left for 15 minutes at room temperature. Flow cytometry was performed as out lined in 2.5.2.

2.5.4. Propidium Iodide staining for Cell Cycle analysis

An identical procedure as the PI staining for Sub-G1 apoptosis analysis was followed but omitting the phosphate citrate buffer washes. Flow cytometry was performed as out lined in 2.5.2.

2.6. RNA extraction

Cells were washed twice with ice cold PBS. In each well of a six well plate 250µl TRIZOL reagent (Invitrogen) was added to the cells and incubated for 5 minutes at room temperature. TRIZOL extracts were collected using cell scrapers. 50µl chloroform was added per 250µl TRIZOL, vortexed for 15 seconds, and then incubated for 2-3 minutes at room temperature. Samples were then centrifuged at 12,000g for 15 minutes at 4°C and the aqueous upper phase was transferred to a new microfuge tube. An equal volume of 70% ethanol was added and mixed by pipetting. The mixture was applied to an RNeasy Mini column (QIAGEN) then centrifuged for 15 seconds at 8000g. RNA clean up was followed using the manufactures instructions, including the DNAse I digestion step. The RNA concentration was estimated by spectrophotometer and formaldehyde-MOPS denaturing agarose gel electrophoresis was then used to assess RNA quality.

2.7. Gene Expression Microarrays

2.7.1. The Affymetrix GeneChip

All experiments were performed using Human Genome U133 Plus 2.0 high-density oligonucleotide arrays (Affymetrix, *url*). Oligonucleotides of 25 base pairs in length are used to probe ~~message~~ mRNA levels in the samples. Each gene of interest is represented by a set of oligonucleotides comprised of 11 probe pairs. Each probe pair is composed of a perfect match (PM) probe against a section of the mRNA molecule of interest, and a mismatch (MM) probe that is created by changing the middle (13th) base of the PM

with the intention of measuring non-specific binding. The HG-U133-Plus2 array contains 54,000 probe sets, representing an estimated 47,000 human transcripts.

2.7.2. Target Preparation, Microarray Hybridisation, Staining and Scanning

High quality RNA from each sample was used to prepare biotinylated target RNA, according to the manufacturer's recommendations (Affymetrix, *url*). An overview of this procedure is shown in Figure 2.2. Briefly, 8μg of total RNA was used to generate first-strand cDNA by using a T7-linked oligo(dT) primer. After second-strand synthesis, *in vitro* transcription was performed with biotinylated UTP and CTP, resulting in approximately 100-fold amplification of RNA. This labelled cRNA target was quantified then fragmented before preparation of the hybridisation cocktail. Spike in controls were added to fragmented cRNA (15μg per HG-U133Plus2 array), before overnight hybridisation. Arrays were then washed and stained with streptavidin-phycoerythrin, before being scanned on an Affymetrix GeneChip 3000 scanner. After scanning, array images were assessed by eye to confirm scanner alignment, the absence of significant bubbles or scratches on the chip surface, and the absence of slides with very high background (scanning and image analysis was performed by Tracy Chaplin, Institute of Cancer, Charterhouse Square).

2.7.3. Data Analysis

Data analysis was performed using BioConductor (Gentleman *et al.*, 2004) and Genespring (Genespring GX 7.3, Agilent Technologies) software. BioConductor is an open development software project, providing access to a wide range of statistical and graphical approaches for the analysis of genomic data. It works through the R open source programming language (Ihaka and Gentleman, 1996). Data was processed from the *.CEL file format which contains information on background values and perfect match and mismatch intensities. Data processing was done using "affy" (Gautier *et al.*, 2004) and "simpleaffy" (Wilson & Miller, 2005) packages in BioConductor.

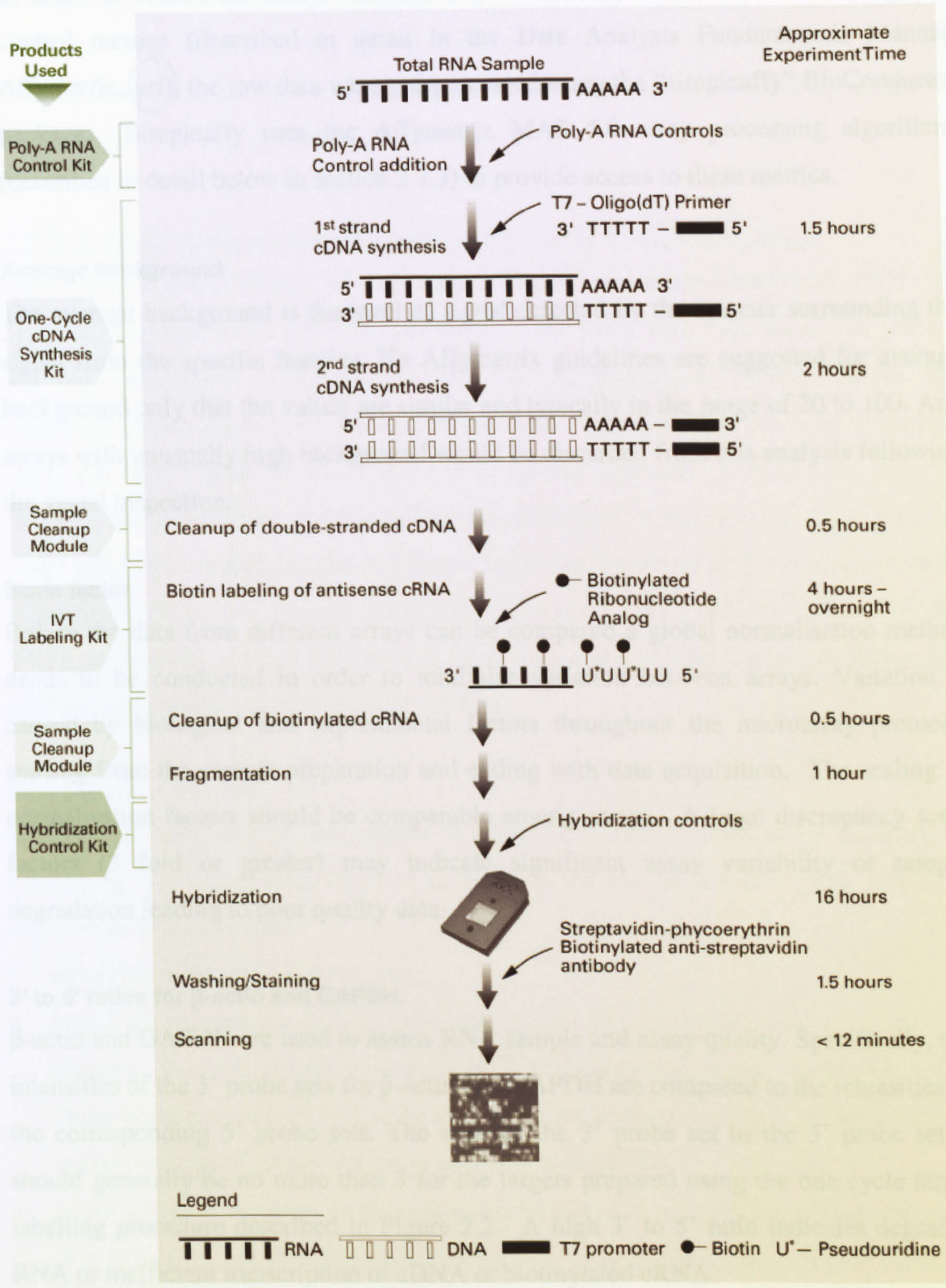


Figure 2.2 Affymetrix GeneChip Eukaryotic Sample and Array Processing. A more detailed description of this procedure can be found in Section 3: Eukaryotic Sample and Array Processing of the GeneChip Expression Analysis Technical Manual (Affymetrix, *url*).

2.7.3.1. Quality Control

In order to ensure the arrays fulfilled a series of Affymetrix recommended quality control metrics (described in detail in the Data Analysis Fundamentals Manual, Affymetrix, *(url)*) the raw data were pre-processed using the "simpleaffy" BioConductor package. Simpleaffy uses the Affymetrix MAS 5.0 array processing algorithms (described in detail below in section 2.7.3) to provide access to these metrics.

Average background

The average background is the level of signal detected by the scanner surrounding the signal from the specific features. No Affymetrix guidelines are suggested for average background only that the values are similar and typically in the range of 20 to 100. Any arrays with unusually high background would be discarded from this analysis following the visual inspection.

Scale factor

Before the data from different arrays can be compared a global normalisation method needs to be conducted in order to minimise variation between arrays. Variation is caused by biological and experimental factors throughout the microarray protocol, starting from the sample preparation and ending with data acquisition. The scaling or normalisation factors should be comparable among arrays. A large discrepancy scale factors (3 fold or greater) may indicate significant assay variability or sample degradation leading to poor quality data.

3' to 5' ratios for β-actin and GAPDH.

β-actin and GAPDH are used to assess RNA sample and assay quality. Specifically, the intensities of the 3' probe sets for β-actin and GAPDH are compared to the intensities of the corresponding 5' probe sets. The ratio of the 3' probe set to the 5' probe set is should generally be no more than 3 for the targets prepared using the one cycle target labelling procedure described in Figure 2.2. A high 3′ to 5′ ratio indicates degraded RNA or inefficient transcription of cDNA or biotinylated cRNA.

In addition to the "simpleaffy" quality control measures described above, target cRNA quality was also assessed using the array-by-array cRNA digestion plot produced using the "affy" BioConductor *plotAffyRNAdeg* function. This averages individual probe

intensities by their location in each probe set, following a scaling transformation an average can then be taken over all probe sets on the array. A side-by-side plots of these averages, then illustrates any global patterns of 5' to 3' probe intensity. Any abnormally low levels of 3' intensity would illustrate a degraded RNA or inefficient transcription of cDNA or biotinylated cRNA.

2.7.3.2. Data Correction, Normalisation and Transformation

After careful consideration of the quality control measures described above the correction, normalisation and transformation of raw array data took place, following a four-step process of Background Correction, Normalisation, Perfect Match(PM)/ Mismatch (MM) Correction and Summarisation. The data analysis was performed using the "affy" BioConductor package using the *expresso* function. In order to establish the most appropriate method for the microarray experiments performed in this study, the Affymetrix recommended *mas5* method (Affymetrix, *url*) was assessed along with variations of the Robust Multi-array Average (RMA) process (Irizarry *et al.*, 2003), specifically altering Background Correction and PM/MM Correction steps. The detail of each method used is described below.

Background correction

Background correction is the process of correcting probe intensities on an array by subtracting the background level of signal detected by the scanner. The following approaches were assessed.

- ***none*** – no background correction.
- ***rma*** - Developed by Irizarry and colleagues, this correction uses a model that assumes observed intensity is the sum of an exponential signal component and a linear noise component. PM probe intensities are corrected using a global model for the distribution of probe intensities (Irizarry *et al.*, 2003).
- ***mas5*** – This is the method recommended by Affymetrix (Affymetrix, *url*), where a chip is broken into subgrids, and background is calculated for each region based on the lowest 2% of probe intensities. For each region, a weighted average background value is calculated using the distances of the probe location and the areas surrounding the probes of the different regions. Individual probe intensity is then adjusted based upon the average background for each region.

Normalisation Methods

Normalisation is the process of removing non-biological variability between arrays. All global normalisation methods work on the assumption that there is no variability between different microarrays.

quantiles - Developed by Bolstad and colleagues (Bolstad *et al.*, 2003), quantiles aims to give each chip the same empirical distribution. The method assumes that there is an underlying common distribution of intensities across the different experimental samples. The principle of the quantiles algorithm is to sort the probe intensities into ranks of similar intensity across all the microarrays. A median intensity for each rank is then calculated and reattributed to the all the probesets in the originating rank.

mas5 - Developed by Affymetrix mas5 uses a global scaling factor in order to normalise the data between each microarray. A scaling factor is calculated based on the average of all the intensities, after removing the intensities in the lowest 2% and highest 2%. This factor is then used to correct the intensities across all the probe sets on all the arrays.

Perfect Match / Mismatch Correction

PM correction is the process of adjusting PM intensities based on information from the MM intensity values.

pmonly - No PM/MM correction was performed and only PM values were used for analyses. It is widely reported that MM probesets may be detecting signal as well as non-specific binding and therefore including the MM parameter will contribute to the overall noise in the data analysis (Naef *et al.*, 2002; Irizarry *et al.*, 2003).

mas5 – Recommended by Affymetrix, in this method an ideal MM value is subtracted from the PM intensity value, always leaving a positive value. An "artificial" mismatch value is computed when the MM intensity is greater than or equal to the PM and results in a PM-MM that is close to zero.

Summarisation Method

In order to combine the pre-processed probe intensities together in order to compute a single expression measure for each probe set on the array, a summarisation method was employed

medianpolish – Described by Irizarry and colleagues (Irizarry *et al.*, 2003), median polish uses a multi chip linear model fitted to the data from each probe set and the result value is in log2 scale.

mas5 – Recommended by Affymetrix uses a robust average using 1-step Tukey bi-weight on log2 scale.

2.6.3.1 Differential Gene Expression

Filtering

Expression profiling experiments are interested in identifying genes that change their expression between two groups. Therefore, it is important to first to filter the data to include only those probe sets that change in expression. A commonly used method is to filter genes based on the fold change between the *test* and reference groups. However, when filtering on fold change there is a risk of ignoring genes that change significantly but are below the arbitrary fold change threshold. A more biologically sensitive method is to order genes from low to high standard deviation, in order to identify the most variable genes across the conditions analysed. The top 2500 probe sets with highest variance can then be used in clustering and statistical testing.

Hierarchical Clustering

Hierarchical clustering analysis allows the monitoring of overall patterns of gene expression between the normalised arrays, and uses standard statistical algorithms to arrange the genes according to a similarity in gene expression patters. The hierarchical clustering was performed using GeneSpring, based on the approach used by Eisen and colleagues (Eisen *et al.*, 1998). The hierarchical clustering analysis aimed to produce a map of results where probe sets were grouped together based on similarities in their patterns of normalised expression all of the microarrays. The similarity or dissimilarity between a pair of objects in the data set was found by evaluating a distance measure and assuming a normal distribution of gene expression values, it is appropriate to use the Pearson correlation coefficient, which calculates the similarity measure based on a linear model. The objects are then grouped into a hierarchical cluster tree (dendrogram) by linking newly formed clusters. The same algorithms can then be applied to cluster the experimental samples for similarities in there overall patterns of gene expression.

Hierarchical clustering analysis was preformed on the normalised and filtered gene expression data.

Statistical Analyses

This process involves the identification of deferentially expressed genes between different experimental conditions. The Welch's t-test was applied on filtered data using the one way ANOVA setting in GeneSpring. This is a parametric test that works on the assumption that microarray data are normally distributed. Another common assumption of parametric statistical analyses, such as Student's t-test, is that the variability of a gene is constant across treatment types. This is difficult to assess for microarray data so it is safest to assume that variance may differ between treatment and control. Welch's t-test corrects for difference in variability, and does not detect it, therefore is more suitable for microarray data differential gene expression analysis. In a conjunction with this test a False Discovery Rate (FDR) multiple test correction was applied across the significant genes using GeneSpring. Multiple testing corrections adjust p-values to correct for occurrence of false positives. False positives are genes that are identified as significant changes following statistical tests, when their true state is unchanged. A False Discovery Rate of 5% (p-value <0.05) on an array of 54,000 reporters would mean that on any size gene list, 2700 genes would be expect to be false leads. The Benjamini and Hochberg FDR correction was applied across the significant genes (Benjamini & Hochberg, 1995). This test reduces the number of false positives without enriching the number of false negatives which can be the case for other types of correction (Benjamini & Hochberg, 1995). Briefly, the p-values are ranked from the smallest to the largest. The largest p-value remains as standard. The second largest p-value is multiplied by the total number of genes in the gene list of differentially expressed genes divided by its rank. The same approach is repeated with the second largest p-value and so on, until no gene is found to be significant. The resulting FDR corrected values mean that a FDR of 5% (FDR corrected p-value <0.05) on a gene list of 500 would expect 25 to be false leads, regardless of the number of reporters on the array.

2.7.4.Gene Ontology

2.7.4.1. Introduction to Gene Ontology

Developed by the Gene Ontology Project, Gene Ontology (GO) is a controlled vocabulary to describe gene and gene product attributes in any organism (Ashburner *et al.*, 2000; The Gene Ontology project, *url*). In particular, each gene is assigned a term based on its Molecular Function (e.g. DNA Binding Transcription Factor), Biological Process (e.g. Regulation of Transcription from RNA polymerase II promoter), and Cellular Component (e.g. Nucleus). The annotations are a mixture of those that are manually assigned, based on primary and review literature and some automatically assigned based on predicted gene functions. Each GO term is additionally assigned a code based on the strength of the associated evidence and more than one term can be assigned to each gene. It should be noted that the value of GO is limited by the quality of the associated research that is used to define each annotation. GO terms are organised into a hierarchical structure, an example for AP-2γ is shown in Figure 2.3. A GO path is where a GO term is connected to several other GO terms higher in the GO hierarchy. An added level of complication is that these paths intersect or split and also several paths may lead to an individual gene.

GO:0008150 : biological_process [141300]
  └GO:0065007 : biological regulation [20429]
    └GO:0050789 : regulation of biological process [18593]
      └GO:0050794 : regulation of cellular process [14845]
        └GO:0031323 : regulation of cellular metabolic process [9642]
          └GO:0019219 : regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process [8738]
            └GO:0045449 : regulation of transcription [8305]
              └GO:0006355 : regulation of transcription, DNA-dependent [6764]
                └GO:0006357 : regulation of transcription from RNA polymerase II promoter [1747]

Figure 2.3. Biological Process Gene Ontology Hierarchy for *TFAP2C*. In brackets is the total number of genes assigned the particular GO term.

In order to assess any patterns in groups of genes within a differential gene expression data set, data can be sorted into broad functional groups based on their assigned

Biological Process GO terms. This can be done by hand which offers the advantage of deciding the most appropriate terms for each gene where two or more differing terms are described. Or this analysis can be automated. For automated analysis the GOSTAT and Panther web based interfaces were used.

The GOstat interface

This analysis was conducted using the GOstat interface developed by Beissbarth and Speed (GOstat *url*; Beissbarth & Speed 2004), in order to establish with statistical confidence which GO terms were over-represented or under-represented in a list of significantly regulated transcripts. Briefly, for all of the transcripts analysed, GOstat determines the associated annotated GO terms and all branches on their connection path. The program then counts the number of appearances of each GO term for the transcripts in the list being analysed, as well as in a reference list. For each GO term, a Chi^2 p-value is calculated representing the probability that the observed numbers of counts could have resulted from randomly distributing this GO term between the tested and the reference lists.

The PANTHER classification system

The PANTHER classification system is very similar to the Gene Ontology system, based on interconnecting hierarchies of terms, but greatly abbreviated and simplified (Mi *et al,* 2005; PANTHER Classification System, *url*). This facilitates automated high-throughput analyses and accessible graphical outputs. In order to establish with statistical confidence which PANTHER Biological Process annotations were over-represented or under represented in a list of significantly regulated transcripts, they can be compared to all of the PANTHER Biological Process represented by all of the 25431 genes within the NCBI Refseq Database. Similarly to the GOstat program described above, this allows an objective and statistically controlled measure of enriched GO terms from with the data set.

2.8. Quantitative PCR

2.8.1. Reverse transcription reaction

A microgram of total RNA was used to generate cDNA. Reverse transcription reactions were performed using sterile plastic ware throughout and aerosol filter tips to reduce contamination. All materials were from Applied Biosystems. The reactions were prepared on ice as follows, 1µg total RNA, 5.5mM $MgCl_2$, 2.5mM dNTP mix, 2.5µM Random Hexamers, RNase Inhibitor (0.4U/µl), MultiScribe Reverse Transcriptase (1.25U/µl) and RNAse free water to make the reaction volume up to 50µl. Samples were then transferred to a thermocycler and reactions incubated at 25°C for 10 minutes, 48°C for 30 minutes followed by 95°C for 5 minutes to inactivate the enzyme. After the RT reaction one can assume that 1µg of total RNA corresponds to 1µg of cDNA.

2.8.2. Quantitative "Real Time" PCR reaction

Pre-designed transcript specific primer-probe sets for use in for "real time" PCR reactions were purchased from Applied Biosystems, the details of these are outlined in Table 2.5. Briefly, a probe binds the DNA between the two primers and is dye labelled at the 5' end whose fluorescence is controlled by a quencher at the 3'. As the reaction proceeds, the dye is cleaved from the probe and thus released from the quencher. Therefore as the reaction progresses the fluorescent emission from the dye increases allowing accurate quantification of the target sequence. PCR reactions (25µl) were prepared in triplicate on a 96-well plate using 15ng cDNA per reaction. 2X Universal PCR master mix (Applied Biosystems) was combined with the primer-probe mix to minimise pipetting error between each well of the plate. In addition to the samples, a no template control was included to look at DNA contamination of the samples and a no reverse transcriptase control was included to control for potential genomic DNA contamination in the RNA extraction. This consisted of an equivalent amount of a reverse transcriptase reaction in which the reverse transcriptase was omitted. In order to assess the efficiency of the PCR reaction and to allow relative quantification, a standard curve was run alongside the samples. The standard curve used four separate dilutions of cDNA per reaction and was prepared in triplicate; 25ng, 6.25ng, 1.5625ng and 0.39ng.

The PCR reaction was carried out on the 7700 Sequence Detection System (Applied Biosystems) using the following program as standard: 50°C for 2 minutes (AmpErase UNG step), 95°C for 10 minutes to activate the AmpliTaq Gold, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute.

Transcript	Assay ID	Probe Dye Layer
GAPDH	4319413E	VIC
TFAP2C	Hs00231476_m1	FAM
TIMP2	Hs00234278_m1	FAM
PERP	Hs00751717_s1	FAM
SPDEF	Hs00171942_m1	FAM
DKK1	Hs00183740_m1	FAM
IGFBP3	Hs00181211_m1	FAM
CXCL12	Hs00171022_m1	FAM
INHBA	Hs00170130_m1	FAM
KIAA1324	Hs00331399_m1	FAM
CELSR2	Hs00154903_m1	FAM
TGLN	Hs00162558_m1	FAM

Table 2.5. Details of primer-probe sets used in qPCR analyses. All PCR reactions were analysed by agarose gel electrophoresis to ensure the presence of a single product under standard PCR conditions.

Results were initially analysed using the Sequence Detection software version 1.9.1. (Applied Biosystems). The amplification plots were observed in both linear and semi-log plots, with the background corrected and the threshold cycles determined. Following this the standard curve was plotted showing the slope (PCR efficiency) and the correlation coefficient. A slope of –3.3 relates a 100% efficient PCR reaction, a ten fold increase in PCR product every 3.3 cycles. The PCR efficiency was considered satisfactory above 98% and only if all the samples fell within the points of the standard curve. Data was analysed according to the Standard Curve Method for relative quantification (Applied Biosystems; *url*). Standard curves were prepared for both the target (e.g. AP-2γ) and the endogenous reference (e.g. *GAPDH*). For each experimental sample, the relative quantity of target and endogenous reference levels was determined from the appropriate standard curve. The target amount was then divided by the endogenous reference amount to obtain a normalised target value. The transfection

control sample was used as a calibrator and each of the normalised target values was divided by the calibrator normalised target value to generate the relative expression levels. Triplicate samples were used to generate standard errors.

The *OAS1* real time PCR reactions were performed using a SYBR green PCR reaction to detect PCR products, in an identical manner as the qPCR described for ChIP (see section 2.9.3), the data was normalised to endogenous reference levels as described above. The primers used for *OAS1* amplification were 5'-AGGTGGTAAAGGGTGGCTCC-3' and 5'-ACAACCAGGTCAGCGTCAGAT-3' (Bridge *et al.*, 2003).

2.9 Chromatin ImmunoPrecipitation Assays

All of the ChIP assays presented in this thesis were performed in collaboration with Karsten Friedrich, who performed the chromatin preparation and IPs.

2.9.1. Preparation of soluble chromatin

Cells were analysed at approximately 70-80% confluency on a 15cm plate. Chromatin was cross-linked by the addition of formaldehyde to a final concentration of 1% and incubated at 37°C for 10 minutes. Cross-linking was stopped by the addition of glycine to a final concentration of 0.125M and incubated at RT for 5 minutes with gentle rocking. Cells were washed twice with ice cold PBS then detached using a cell scraper in 5ml PBS with Complete Proteinase Inhibitors (Roche) and pelleted at 3,000g for 5 minutes at 4°C. The supernatant was removed and the resultant cell pellet was gently resuspended in 3ml (for 3 plates) of cell lysis buffer (5mM PIPES pH 8.0, 85mM KCl, 0.5% NP40, and Complete Proteinase Inhibitors). Cell lysis was performed on ice for 10 minutes and followed by centrifugation to pellet the crude nuclei fraction (5,000g for 5 minutes, 4 °C). The pellet from 3 plates was resuspended in 500µl of nuclei lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1) and incubated for 10 minutes at 0-4 °C (care was taken to prevent SDS precipitation). The chromatin was sheared to an average length of 500bp by sonication. This was achieved using a 10 second pulse followed by a 30 second rest for a total of 24 cycles at 30% amplitude on the Sonics

Vibracell VCX500 using a 0.3cm tip. Sonications were performed on ice and a multiple tipped probe allowed simultaneous sonication of up to four microfuge tubes, maximising the homogeneity of sonication between experimental conditions. After sonication, debris was pelleted by centrifugation at 14,000g for 10 minutes at 10°C and the supernatant transferred to a clean tube. At this point chromatin could be snap frozen in liquid nitrogen and stored at –80°C for several months. Typically, 3-4 plates (15cm) yielded enough chromatin to perform 10 IP reactions including the controls.

2.9.2. Immuno Precipitation

Chromatin was diluted approximately 2.5 fold in IP buffer (0.5 % Titron X100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 100 mM NaCl, plus protease inhibitors) before being bulk pre-cleared using of 50% (w/v) blocked Protein A Sepharose (PSA), in order to reduce the level of background in IPs.

To block the PSA beads two washes, followed by a 2 hour gentle rotation at 4°C, in a 2X volume of IP buffer with 0.1 mg/ml sonicated herring sperm DNA (Sigma B8894) and 0.2 mg/ml BSA (Invitrogen 15634-017) was conducted, followed by a final wash in 1x IP buffer prior to use. 500µl of 50% (w/v) blocked PSA beads were added to the 1250µl of diluted chromatin (50µl of beads per IP reaction) and incubated by gentle rotation at 4°C for 1 hour. PSA was removed by centrifugation at 2,000g for 3 minutes and the supernatant transferred to a clean tube.

For the ChIP experiments documented in this thesis looking at the differences in promoter occupancy between AP-2γ expressing and non-expressing cell lines, it was important to ensure that the amount of starting chromatin was kept consistent between experimental conditions. Therefore, at this stage a 125µl aliquot of the pre-cleared chromatin was taken for DNA isolation and quantification.

DNA isolation was performed using Chelex 100 beads based on the protocol described by Nelson and colleagues (Nelson *et al.*, 2006). The DNA from this aliquot was precipitated in 2.5 volumes of ethanol and then washed in 70% ethanol. The dissolved pellet was then resuspended in 100µl of a 10% (w/v) Chelex 100 suspension and boiled

for 10 minutes. After cooling, 1µl of proteinase K (10 mg/ml) was added and incubated on a thermal mixer at 55°C for 30minus (1000rpm). Samples were boiled for a further 10 minutes to inactivate the protenase K and then centrifuged at 12,000g for 1 minute (4°C). The DNA supernatant (approximately 80µl) was then transferred to a new tube. A further 120µl of water was added to the Chelex beads, vortex mixed and then centrifuged at 12,000g for 1 minute (4°C). The 120µl supernatant was then pooled with the previous supernatant and the DNA concentration calculated using a spectrophotometer. This sample was kept on ice and used as the total input chromatin.

An equivalent amounts of chromatin as determined by DNA concentration were then aliquoted for each IP reaction. A no antibody control was included as a background comparator for IP experiments and in addition a mock sample was set up containing no chromatin to give an indication of DNA contamination during the reaction. The details of the antibodies using in ChIP are outlined in Table 2.6. The amounts of each specific antibody were empirically determined by Karsten Friedrich. IP reactions were incubated over night at 4°C with gentle mixing.

Antibody	Animal Origin	Amount of Antibody	Source (Cat#)
AP-2γ (KF3)	Rabbit	10µl	Lab Reagent
p300 (N-15)	Rabbit	5µl	Santa Cruz (sc-584)
Acetylated-H4	Rabbit	1µl	Upstate (06-866)

Table 2.6 Antibodies used in ChIP assays in this study. Amount (µl) corresponds to the volume of antibody added to a tenth of the chromatin extracted from 3-4, 70-80% confluent, 15cm plates of MCF-7 cells (Approximately 5µg of antibody per 10^6 cells).

The immune complexes were collected by adding 50 µl of blocked 50 % (w/v) PSA to each sample and incubating on a rotating platform at 4°C for 1 hr. The PSA beads were pelleted by centrifugation at 2,000g for 3 minutes at RT and the supernatant discarded. PSA beads were washed for three minutes on a rotating platform using 1 ml of wash solution and then centrifuged for 3 minutes at 2,000g at RT using the following wash conditions:

- Wash One: 2 mM EDTA , 20 mM Tris-HCl [pH 8.0], 0.1 % SDS, 1 % Titron X100, 150 mM NaCl

- Wash Two: 2 mM EDTA , 20 mM Tris-HCl [pH 8.0], 0.1 % SDS, 1 % Titron X100, 500 mM NaCl
- Wash Three: 1 mM EDTA , 10 mM Tris-HCl [pH 8.0], 1 % NP-40, 1 % Deoxycholate, 0.25 M LiCl
- TE wash repeated twice: 1 mM EDTA , 10 mM Tris-HCl [pH 8.0],

After the centrifugation, as much of the remaining TE as possible was removed without disrupting the PSA beads. The antibody/chromatin complexes were eluted from the PSA by adding 150µl elution buffer (1% SDS, 50mM NaHCO3) and shaking at 1000rpm for at least 30minutes at RT. Elutes were centrifuged at 14,000g for 1 minute to pellet PSA and supernatant was removed to a clean tube. DNA was then isolated using Chelex 100 beads as described above and 5µl was used per reaction in a quantitative real time PCR reaction described below.

2.9.3. Quantitative PCR for ChIP

Quantitative "real time" PCR for ChIP was performed using promoter specific primer pairs and measured using the fluorescent SYBR Green dye as a detector. Briefly, SYBR Green intercalates into double-stranded DNA and produces a fluorescent signal. The intensity of the signal is proportional to the amount of dsDNA present in the reaction. Therefore, at each step of the PCR reaction, the signal intensity increases as the amount of specific product increases.

Promoter specific primer pairs, except where stated were designed using the Primer Express 3.0 software (Applied Biosystems, *url*) following recommended guidelines. In order to optimise the PCR reaction, a titration of primer concentration was performed. Combinations of the final primer concentrations: 50nM, 300nM, 900nM, were assessed for each primer pair under standard PCR conditions. All PCR products were visualised by agarose gel electrophoresis in order to confirm the presence of a specific product. Details of the primers used in ChIP are shown in Table 2.7

Genomic Region	Amplicon Location	Primer Pairs 5' – 3'	Final Primer Concentrations	Reference
CDKN1A Promoter	-2290 -2185	GTGGCTCTGATTGGCTTTCTG CTGAAAACAGGCAGCCCAAG	50nM 300nM	Kaeser & Iggo, 2001
	-21 +44	TTGTATATCAGGGCCGCGCT CGAATCCGCGCCCAGCTC	300nM 300nM	Shan *et al.*, 2003
	-4005 -3931	TGGCCCCTCTGTGAAAACAT TTCCTGTTCCTGGCTCTAACAAC	300nM 300nM	This Study
	-935 -864	TGATCTGAGTTAGGTCACCAGACTTC TCCCCACATAGCCCGTATACA	300nM 900nM	This Study
Sat2	Chr1 satellite repeat	CATCGATGGAAATGAAAGGAGTC ACCATTGGATGATTGCAGTCAA	300nM 300nM	Jiang *et al.*, 2004

Table 2.7. The detail of the primers used for SYBR Green Quantitative PCR for analysis of ChIP elutes. Amplicon Location are relative to the start of transcription.

PCR reactions (25µl) were prepared in triplicate on a 96-well plate using 5µl of eluted DNA per reaction. 2X SYBR green master mix (Applied Biosystems) was combined with the primers to the indicated concretion (Table 2.7). In addition to the samples, a no template control was included to look at DNA contamination of the samples. In order to assess the efficiency of the PCR reaction and to allow relative quantification, a standard curve was run alongside the samples. The standard curve used four separate dilutions of input chromatin and was prepared in triplicate; 1/100, 1/500, 1/2000 and 1/5000. The PCR reaction was and initial data analysis was carried out as described in Section 2.8. The relative quantities of each PCR reaction were calculated from the standard dilution of input chromatin. The "no antibody control" quantity was then subtracted from the specific antibody quantity and expressed relative to the 1/1000 dilution of the total input chromatin.

2.10. Luciferase Reporter assays

2.10.1. Transient Transfection of Luciferase Reporter Constructs in HEPG2 cells.

HEPG2 cells were seeded in 24-well plates at 1×10^5 cells per well and were transfected in triplicate the following day using FuGENE 6 (Roche Molecular Biochemicals). Briefly, equal total amounts of plasmid DNA for triplicate reactions was aliquoted into tubes, 100µl media without serum was added to the DNA and Fugene was added at a ratio of 3:1 (µl Fugene: µg DNA) and incubated at RT for 15 minutes. The reaction is then aliquoted onto the cells which were supplemented with media containing 10% FCS. To accurately show changes in luciferase reporter activity, cells were transfected with both the luciferase reporter construct and the phRG-TK Renilla plasmid (Promega). The phRG-TK Renilla utilises a constitutively expressing promoter to drive the expression of Renilla luciferase, therefore allowing luciferase levels to be normalised between reactions. After transfection cells were incubated for 48 hours before being assessed for luciferase reporter activity. Cells were washed twice with ice cold PBS, then lysed with 100µl passive lysis buffer (Promega) at RT for 15 minutes with shaking. Lysis buffer and cell debris was transferred to a clean microfuge tube and the debris was pelleted by centrifugation at 14,000g for 1 minute. 40µl of cleared lysates were aliqouted onto a 96-well plate. Luciferase levels were determined using Promega's dual luciferase kit according to the manufacturer's instructions on the Luminoskan Ascent (Thermo Labsystems) luminometer.

2.10.2 Transient Transfection of Luciferase Reporter Constructs in MCF-7 cells in conjunction with siRNA transfections.

MCF-7 cells were transfected with siRNAs in triplicate using six well plates as described in Section 2.2.2. Twenty four hours later a second transfection with both the luciferase reporter construct (1µg) and the phRG-TK Renilla plasmid (0.25µg) was carried out using FuGENE 6 in an identical manner to that described above. At 48 hours, cells were passaged 1 in 2 and finally at 72 hours luciferase activity was assayed as described above except cells were lysed in 200µl of passive lysis buffer.

CHAPTER 3: RESULTS AP-2γ RNA INTERFERENCE

3.1. Assessment of AP-2γ expressing cell lines.

With the aim to establish which cellular pathways have become activated in AP-2γ expressing breast tumours we planned to determine the AP-2γ dependent gene expression profile, initially in breast derived cell lines. It was decided to compare wild type over-expressing lines with derivatives "knocked down" for AP-2γ expression using RNAi technology.

In order to confirm the suitability of some known AP-2 expressing breast cell lines levels of AP-2α and AP-2γ were assessed by western blot, shown in Figure 3.1. Table 3.1 summarises this data together with the status of key breast cancer related proteins in these cell lines. Initial work focused on the MCF-7 mammary carcinoma cell line, chosen as it exclusively expresses AP-2γ at high levels. Upon RNAi mediated depletion of AP-2γ in MCF-7 cells all AP-2 would be removed from this cells thus avoiding a further level of complication that might be encountered in other cell lines where AP-2α and AP-2γ are expressed together. MCF-7 cells also express high levels of ERα together with AP-2γ, a combination best thought to reflect the clinical data showing high levels of AP-2γ correlated with poor patient prognosis in ERα positive primary breast carcinomas.

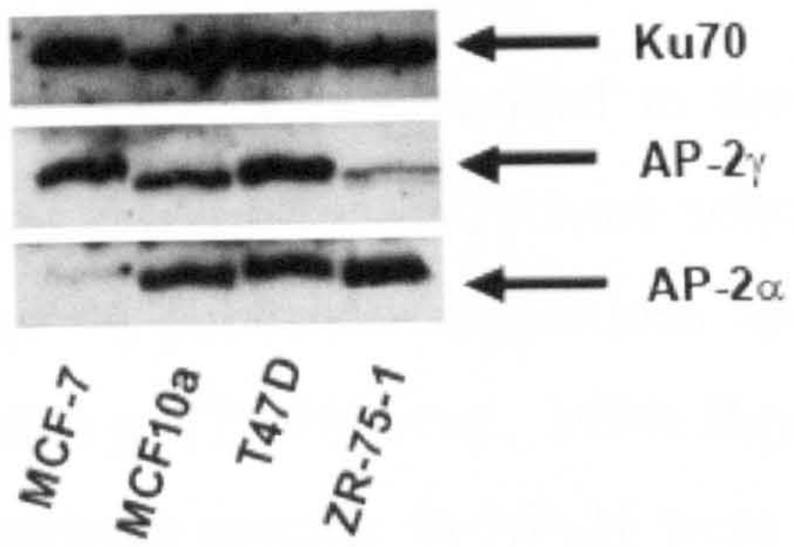


Figure 3.1. AP-2γ and AP-2α protein levels in a panel of breast cell lines. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ or AP-2α before being reprobed for Ku-70 as a loading control. A representative loading control is shown. See Materials and Methods 2.3 for antibody details.

	MCF-7	MCF10A	T47D	ZR-75-1
AP-2γ	+++	++	+++	+
AP-2α	+/-	++	++	+++
AP-2β	-	nd	nd	+/-
CITED2	++	+	+	+
CITED4	+++	+++	+++	+
CBP	++	+	++	+/-
p300	++	+	++	++
ERα	+++	-	++	++
ERBB2	-/+	+	+	++
p53	wt	wt	mut	wt

Table 3.1. Summary of AP-2 and cofactor expression data and in addition the status of key cancer related proteins in a panel of breast cell lines. Table summarises protein levels as determined by western blot, except for p53 where status as wildtype (wt) or mutant (mut) is indicated. Key: Negative (-), barely detectable (-/+), low expression (+), expression readily detectable (++), very high level of expression (+++), or not determined (nd). p53 status from Bartek *et al.,* 1990. Others determined by Western Blot (Helen Hurst, *personal communication*).

3.2. Identification of AP-2γ targeting siRNAs.

Selection of AP-2γ targeting sequences is described in the Materials and Methods, Section 2.2. Briefly sequences specific to AP-2γ were selected for the generation of siRNA oligonucleotides based on published suggestions and BLAST searches were performed to verify the specificity. (Tuschl *et al.*, 1999; Reynolds *et al.*, 2004; Ui-Tei *et al.*, 2004). Additional sequences specific to AP-2γ were ordered from Dharmacon and consisted of four sequences recommended for use in SmartPool transfections (see Section 2.2). A non-specific random sequence was used for a non-silencing control (Qiagen). Figure 3.2 shows efficient knock down of AP-2γ protein levels by three dsRNA oligonucleotide sequences after transient transfection in MCF-7 cells.

For later work in this study sequences from Dharmacon were tested in transient transfections in MCF-7 cells. Figure 3.3 shows western blots of MCF-7 cells following transient transfection of AP-2γ targeting siRNAs in order to assess the additional Dharmacon sequences and a SMARTPool of the four Dharmacon sequences alongside two of the siRNAs already determined to be effective at AP-γ silencing. AP-2γ siRNA3 and Dharmacon siRNA #06 target an identical region of the AP-2γ mRNA and show an expected similar efficiency of knock down. Dharmacon siRNA #05 and #08 are not very efficient showing approximately 50% knockdown. Dharmacon siRNA #07 or the SMARTpool of #05, #06, #07 and #08 show knock down at levels equivalent to AP-2γ siRNA2 and AP-2γ siRNA3.

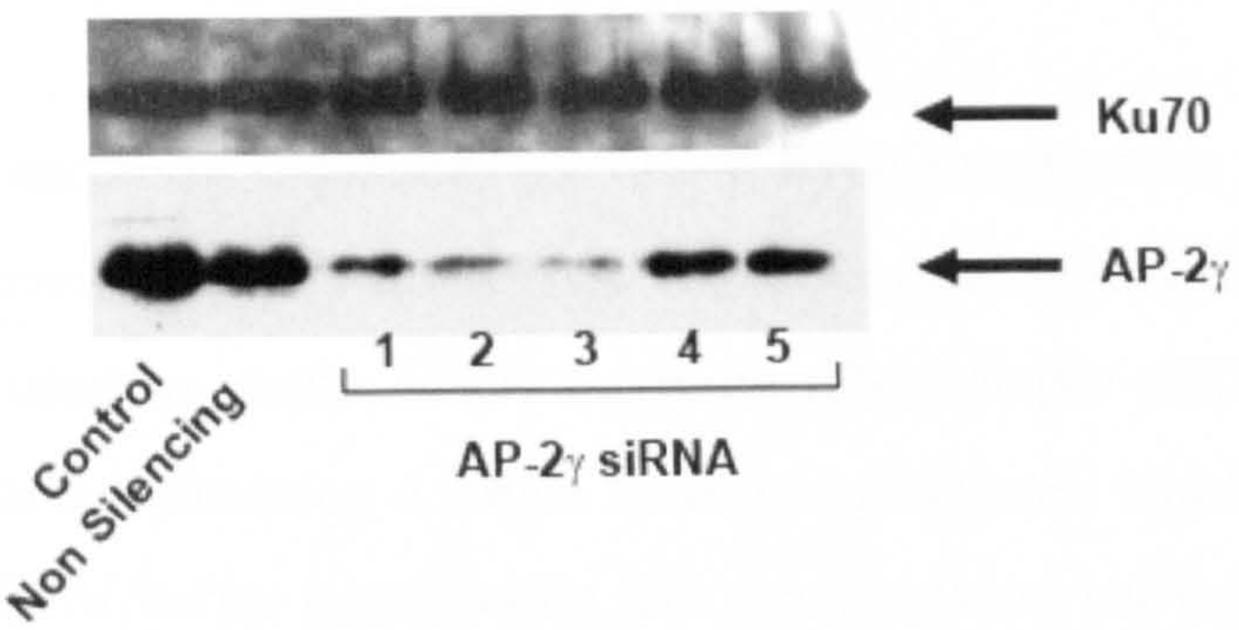


Figure 3.2 Reduction of AP-2γ protein in MCF-7 cells with different siRNAs. MCF-7 cells were transfected with Transfection reagent only (Control), Non-silencing control siRNA or siRNA targeted to AP-2γ: siRNA1, siRNA2, siRNA3, siRNA4 or siRNA5 (Lanes 2-6). Cells were harvested at the 96 hours following transfection with 25nM of siRNA per well of a six well plate. Whole Cell extracts (10µg/lane) from MCF-7 cells were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ before being reprobed for Ku-70 as a loading control.

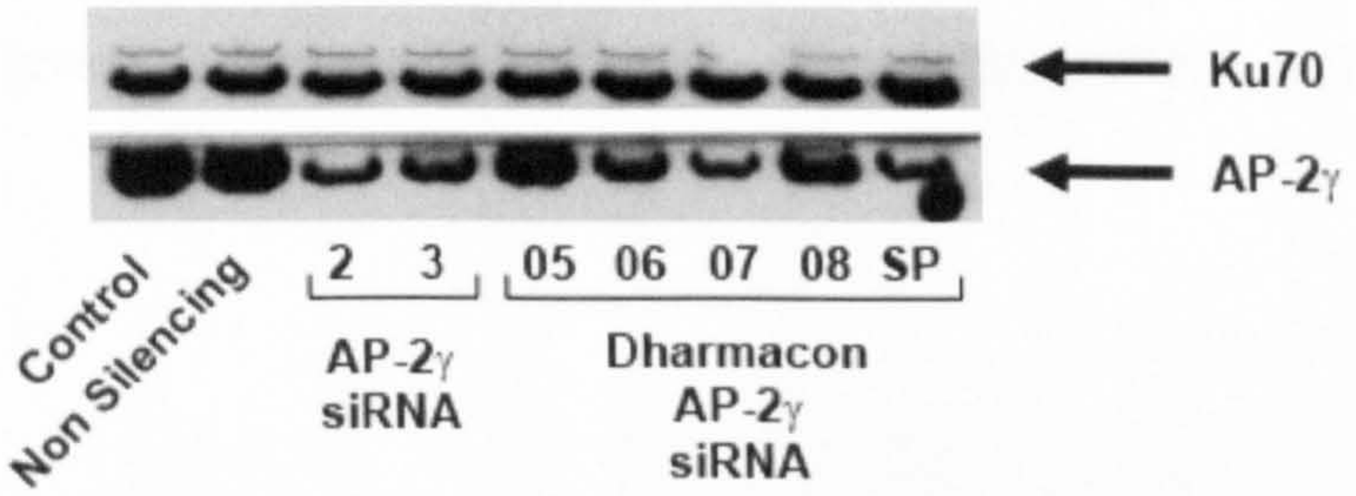


Figure 3.3 Reduction of AP-2γ protein in MCF-7 cells with Dharmacon siRNAs. MCF-7 cells were transfected with Transfection reagent only (Control), Non-silencing control siRNA, siRNA targeted to AP-2γ (siRNA1, siRNA2), individual Dharmacon AP-2γ siRNAs (siRNAs 05-08) or a SMARTPool of siRNAs 05-08 (SP). Dharmacon siRNAs 05,06, 07 and 08 correspond to Dharmacon product codes J-005238-05, J-005238-06, J-005238-07, and J-005238-08 respectively. Cells were harvested at the 96 hours following transfection with 25nM of siRNA per well of a six well plate or for the Smart Pool 6.25nM of each 6, 7, 8, and 9 to give a total of 25nM of siRNA per well of a six well plate. Whole Cell extracts (5µg/lane) from MCF-7 cells were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ before being reprobed for Ku-70 as a loading control.

3.3. shRNA Vectors to make stable AP-2γ silenced lines.

Effective AP-2γ targeting sequences for use in RNA interference have been identified using transient transfections with siRNA duplexes (see Figures 3.2 and 3.3.). These sequences were used to create AP-2γ targeting pRetroSuper shRNA constructs along with control vectors containing a single base pair mismatch within the AP-2 targeting sequence. Details of vector construction and sequences are shown in the Materials and Methods 2.2. Figure 3.4 shows that the three AP-2γ shRNA constructs were able to slightly reduce AP-2γ levels in MCF-7 cell transient transfections compared to control vectors containing a single mismatch in the AP-2 targeting sequence. The degree of knock down is small, but it is thought that this reflects sub-optimal transfection efficiency and not the ability of these sequences to reduce AP-2γ levels. Indeed, in parallel transfections with a pMIG-GFP expression construct (derived from the same vector back bone) showed only 20-30% transfection efficiency. Figure 3.5 shows that the three AP-2γ targeting pRetroSuper shRNA constructs also reduced AP-2γ transactivation from the 3xAP-2 luciferase reporter construct when cotransfected with AP-2γ and CITED2 expression constructs in HepG2 cells, where the mismatch controls had no effect. Together these data suggested that the three shRNA constructs were able to reduce AP-2γ levels and might be used to generate stable shRNA cell lines in MCF-7 cells.

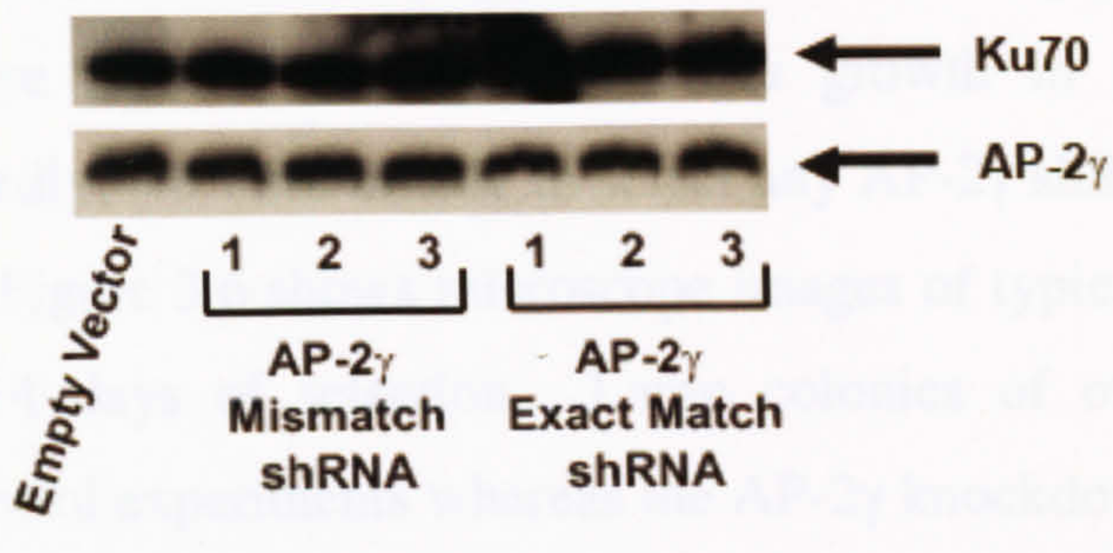


Figure 3.4. Transient transfection of AP-2γ targeting shRNA constructs in MCF-7 cells. MCF-7 cells were transfected as indicated and harvested 72 hours later. Whole Cell extracts (5µg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ and Ku-70 as a loading control.

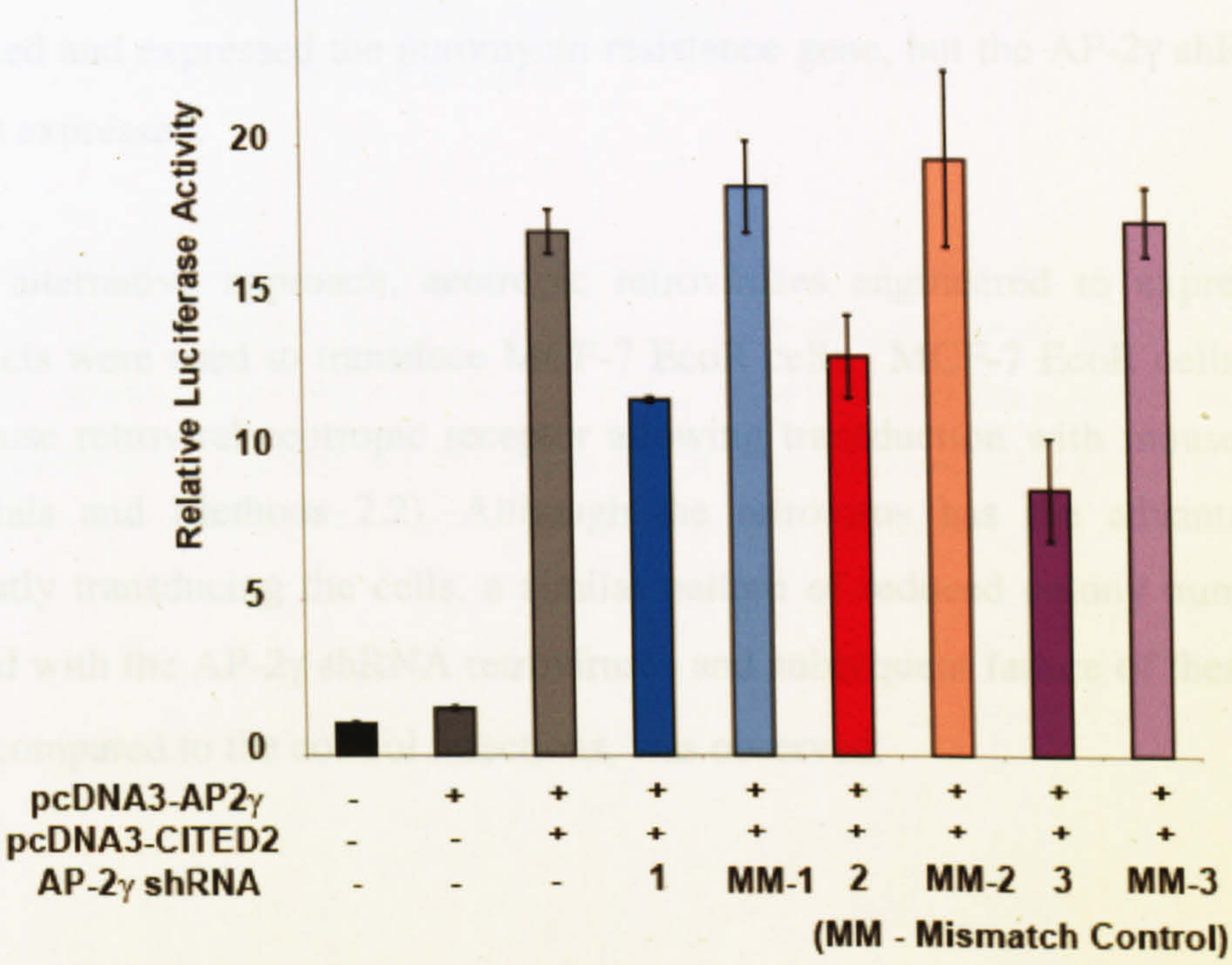


Figure 3.5. Reduction of AP-2γ transactivation of the 3xAP-2Luc reporter upon cotransfection of AP-2γ targeting shRNA constructs. HepG2 cells were cotransfected with a 3xAP-2Luc reporter and the constructs indicated. (phRG-TK) was co-transfected in each experiment. Luciferase activity was assayed after 48 hours. Results (mean ± S.E. of three independent experiments) are presented as relative luciferase activity, corrected for Renilla luciferase activity. The reporter alone transfection value was set at 1.

MCF-7 cells were transfected with each of the three linearised constructs targeting AP-2γ (shRNA1, shRNA2 or shRNA3), the corresponding mismatch non-silencing control constructs (mismatch1, mismatch2 or mismatch3) or the empty pRetroSuper vector. Stable colonies were selected for by continuous growth in puromycin-containing medium. Unexpectedly, we were unable to select any AP-2γ shRNA stable puromycin resistant sub lines. Figure 3.6 shows microscope images of typical colonies from these experiments after 14 days of selection. Large colonies of over 100 cells in size developed in the control experiments whereas the AP-2γ knockdown colonies were only 5-15 cells in size and failed to thrive any further. Colonies from triplicate transfections were counted after 21 days under selection conditions. Figures 3.7 A&B show AP-2γ targeting shRNA transfected cells formed significantly fewer colonies than the corresponding control or empty vector experiments. This data was consistent for the three independent targeting shRNA sequences. Furthermore, surviving single cell clones derived from the AP-2γ targeting shRNA transfections showed no reduction in levels of AP-2 (Figure 3.7 C), indicating that although these cells have successfully integrated and expressed the puromycin resistance gene, but the AP-2γ shRNA cassette was not expressed.

As an alternative approach, ecotropic retroviruses engineered to express the same constructs were used to transduce MCF-7 EcoR cells. MCF-7 EcoR cells stably carry the mouse retroviral ecotropic receptor allowing transduction with mouse retroviruses (Materials and Methods 2.2). Although the retrovirus has the advantage of more efficiently transducing the cells, a similar pattern of reduced colony numbers in cells infected with the AP-2γ shRNA retroviruses and subsequent failure of these colonies to thrive compared to the control infections, was observed.

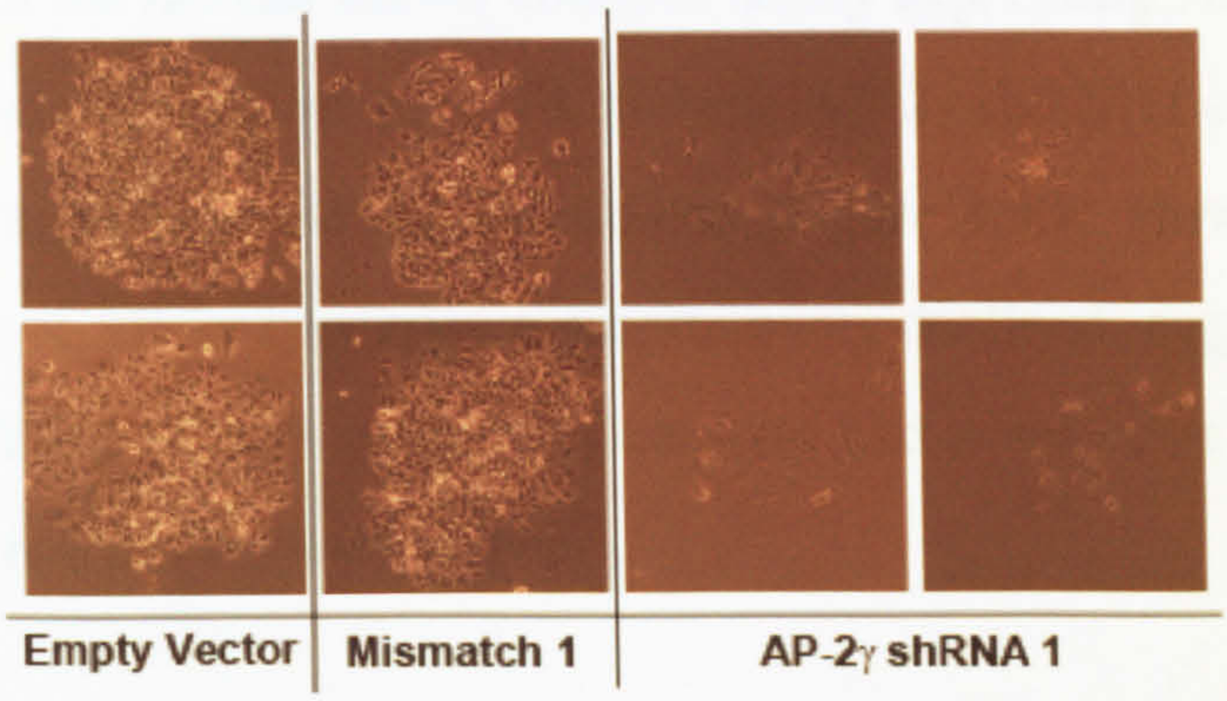


Figure 3.6. Colony growth of MCF-7 cells stably expressing pRetroSuper constructs. MCF-7 were transfected with the shRNA constructs indicated. Untransfected cells were eliminated by continuous growth in puromycin-containing medium. Images show representative colonies of cells 14 days following selection. Large colonies of over 100 cells in size have grown in the control plates. The AP-2γ knockdown colonies are only 5-15 cells in size and failed to grow any further.

The same plasmid constructs have also been used to try to stably reduce AP-2γ in three further breast cell lines, MCF10A, ZR75-1 and T47D, that express both AP-2α and AP-2γ family members. MCF10A cells were transfected with three linearised constructs targeting AP-2γ (shRNA1, shRNA2 or shRNA3), corresponding mismatch non-silencing control constructs (mismatch1, mismatch2 or mismatch3) or an empty pRetroSuper vector. Stable colonies were selected for by continuous growth in puromycin-containing medium. Reduced colony numbers in cells transfected with shRNA constructs targeting AP-2γ and subsequent failure of these colonies to thrive compared to the control infections was again observed (see Figure 3.8). Work by Charlotte Moss in the lab has shown colonies stably expressing AP-2γ shRNA1 did not survive in ZR75-1 cells, however in T47D cells AP-2γ shRNA1 colonies survived and initially showed significant AP-2γ knockdown (Figure 3.9.). However, during passage the effectiveness of AP-2γ knockdown progressively declined, suggesting that cells able to silence the AP-2γ shRNA expression cassette had a selective advantage in culture (Helen Hurst, *personal communication*).

A summary of the ability to stably reduce AP-2γ levels in four different AP-2γ expressing lines is shown in Table 3.2. These results were unexpected and imply that reducing the level of AP-2γ in breast epithelial lines which express this factor to significant levels, (i.e. readily detectable by western blot Figure 3.1) forces them to arrest their growth.

	AP-2γ shRNA1	**AP-2γ shRNA2**	**AP-2γ shRNA3**
MCF-7	No	No	No
MCF10A	No	No	No
ZR75-1	No	nd	nd
T47D	Yes	nd	nd

Table 3.2. Table indicating survival of colonies stably expressing the AP-2γ shRNA pRetrosuper constructs. Stable colonies were selected for by continuous growth in puromycin-containing medium. Failure of AP-2γ shRNA stable colonies to form or survive was controlled by comparing colony formation in empty vector and mismatch control transfected cells. (nd = not determined).

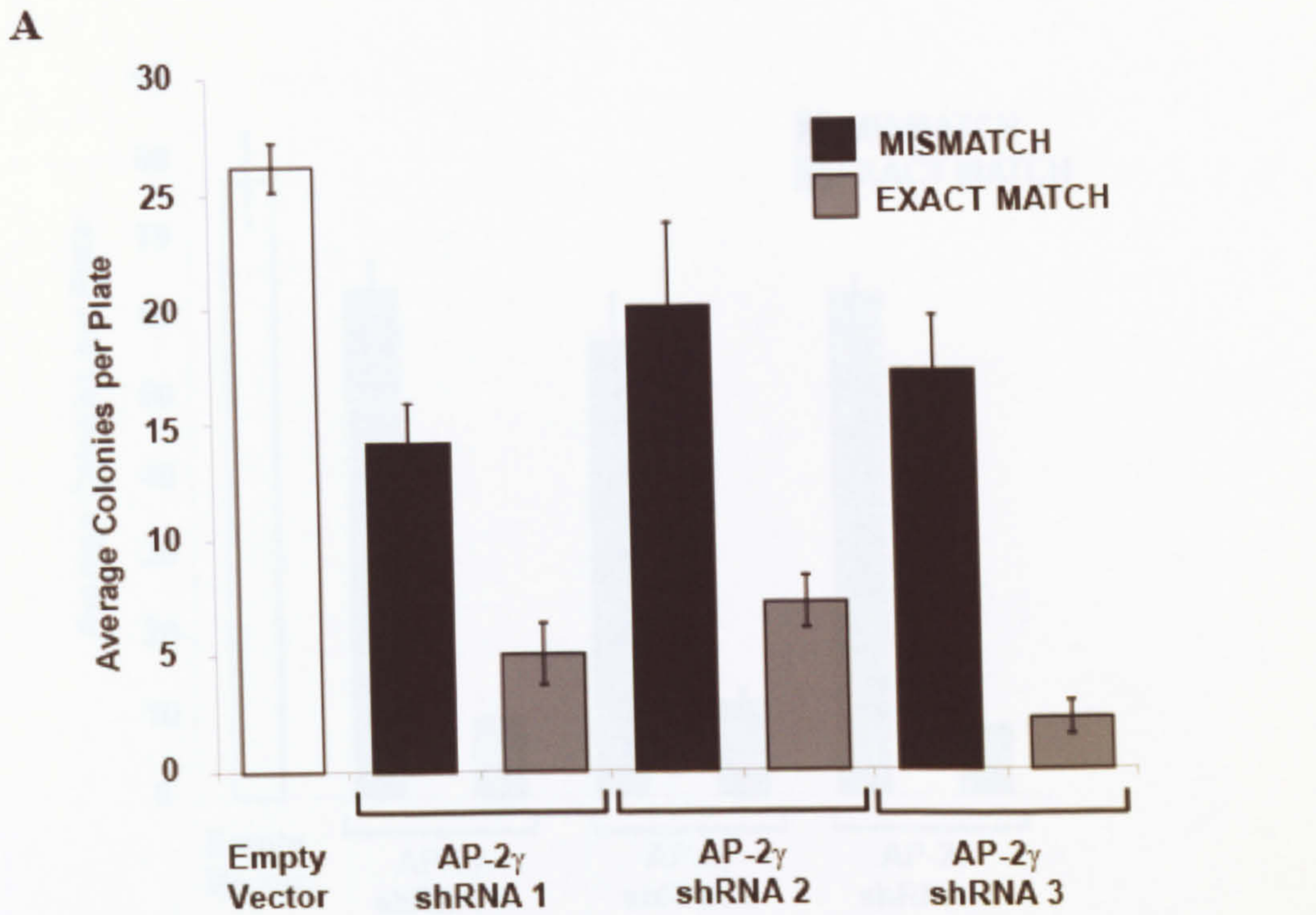


Figure 3.7. AP-2γ silencing results in reduced colony formation in MCF-7 cells. MCF-7 cells were transfected with the shRNA constructs indicated. Untransfected cells were eliminated by continuous growth in puromycin-containing medium and colonies were counted after 21 days under selection conditions. **(A)** Graphical representation of average colony numbers per plate (three plates per condition). **(B)** Representative experiment showing colonies stained with MTT. **(C)** Whole Cell extracts (5μg/lane) from surviving colonies were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ and Ku-70 as a loading control.

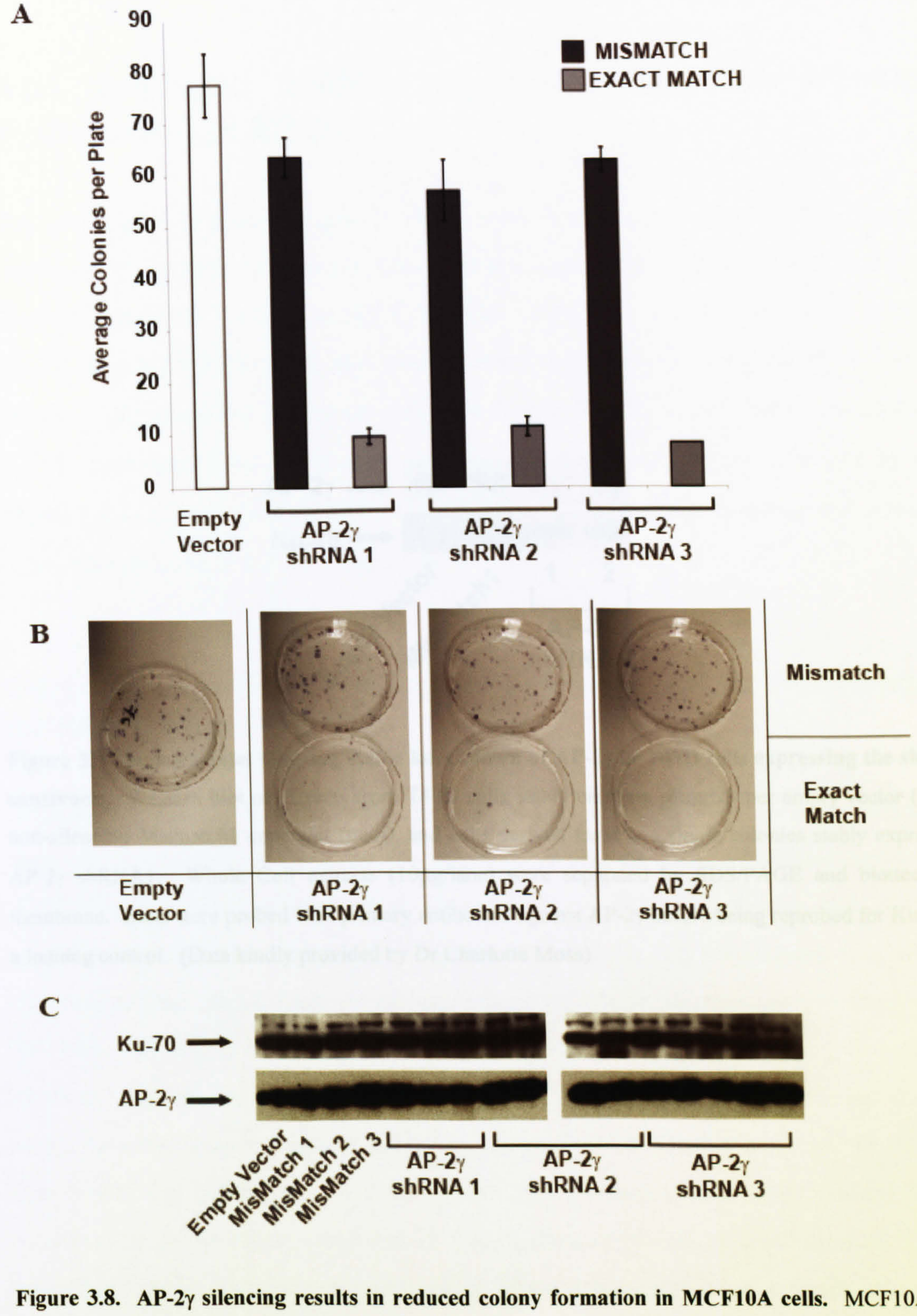


Figure 3.8. AP-2γ silencing results in reduced colony formation in MCF10A cells. MCF10A cells were transfected with the shRNA constructs indicated. Untransfected cells were eliminated by continuous growth in puromycin-containing medium and colonies were counted after 21 days under selection conditions. **(A)** Graphical representation of average colony numbers per plate (three plates per condition). **(B)** Representative experiment showing colonies stained with MTT. **(C)** Whole Cell extracts (10μg/lane) from surviving colonies were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ and Ku-70 as a loading control.

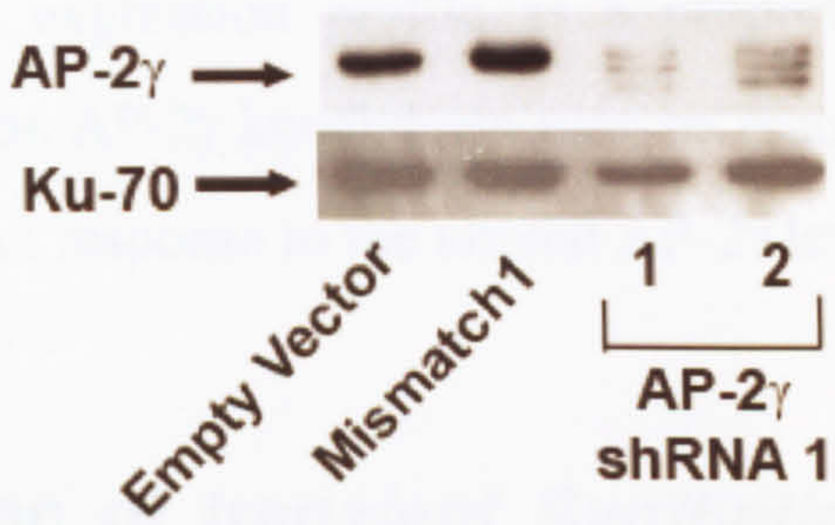


Figure 3.9 Single colonies showing stable knockdown of AP-2γ in T47D cells expressing the shRNA constructs. Western blot of extracts from T47D cells stably carrying pRetroSuper empty vector (pool), non-silencing Mismatch1 construct (pool), and cells derived from two single colonies stably expressing AP-2γ shRNA1. Whole Cell extracts (10μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ before being reprobed for Ku-70 as a loading control. (Data kindly provided by Dr Charlotte Moss)

3.4. Synthetic siRNA oligonuleotides for transient knockdown of AP-2γ.

The observed effect on the growth of the cells stably expressing shRNA constructs targeting AP-2γ made gene expression profiling impossible by this method, so transient knockdowns were pursued in MCF-7 cells. Transient knockdowns with synthetic siRNA oligonucleotides were also used to investigate how silencing AP-2γ effects cell growth. The observed effects on cell growth might make it difficult to determine an AP-2γ dependent gene expression profile as a proportion of the changes in gene expression observed upon AP-2γ knockdown may be secondary to the altered cell cycle rather than due to a direct response to the altered AP-2γ level.

3.4.1. Optimisation of transient Synthetic siRNA transfections for knockdown of AP-2γ.

Originally I planned to use double transfection experiments, that is cells transfected once then re-transfected 72 hours later after a 1 in 3 split, using a 200nM concentration of siRNA, to achieve a longer-term knock down of AP-2γ. However, this approach served to highlight the limitations of RNAi. After the second transfection non-specific effects of siRNA transfection were observed and an interferon response was triggered in the cells. Figure 3.10 shows the efficient knockdown of AP-2γ using this double transient transfection of AP-2γ siRNA1. Efficient knock down was achieved after 96 hours, however this was accompanied by an increase in STAT1 levels at the 96 hour time point and thereafter. STAT1 is the major transcription factor central to the induction of an interferon response in mammalian cells and up-regulation of STAT1 following siRNA transfection is indicative of such non-specific effects (Bridge *et al.*, 2003). Interestingly, a pilot gene expression microarray experiment comparing the expression profiles of the non-silencing control siRNA condition with AP-2γ siRNA1 using RNA harvested from the 96 hour time point clearly demonstrated an induction of additional interferon stimulated genes (see Table 3.3.)

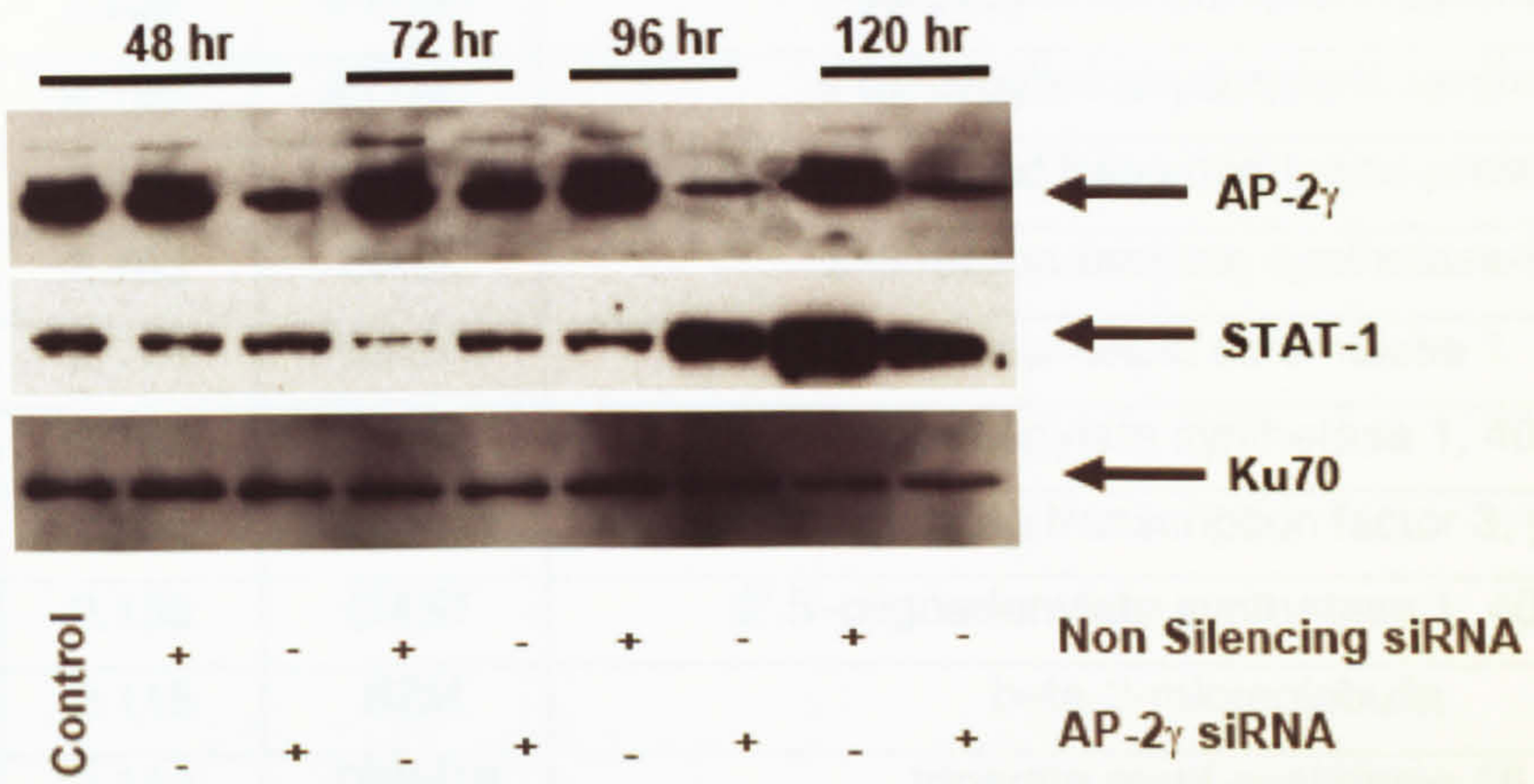


Figure 3.10. STAT-1 induction in MCF-7 cells transfected with siRNAs. MCF-7 cells were transfected with Transfection reagent only (T), Non Silencing control siRNA (1) or siRNA1 targeted to AP-2γ (2). Cells were first transfected at time point 0 hours then re-transfected 72 hours later after a 1 in 3 split. Cells were harvested at the indicated time points across this 120 hour period. Whole Cell extracts (10μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ or STAT-1 before being reprobed for Ku-70 as a loading control.

Probe ID	Fold Change	Name	Description
205483_s_at	2.350	*G1P2*	interferon, alpha-inducible protein (clone IFI-15K)
203153_at	2.309	*IFIT1*	interferon-induced protein with tetratricopeptide repeats 1
221577_x_at	2.261	*PLAB*	prostate differentiation factor
201601_x_at	2.202	*IFITM1*	6-pyruvoyltetrahydropterin synthase
214022_s_at	2.196	*IFITM1*	6-pyruvoyltetrahydropterin synthase
212203_x_at	2.185	*IFITM3*	interferon induced transmembrane protein 2 (1-8D)
205660_at	2.183	*OASL*	2'-5'-oligoadenylate synthetase-like
202446_s_at	2.176	*PLSCR1*	phospholipid scramblase 1
205552_s_at	2.168	*OAS1*	2',5'-oligoadenylate synthetase 1, 40/46kDa
203882_at	2.136	*ISGF3G*	interferon-stimulated transcription factor 3, gamma 48kDa
202869_at	2.133	*OAS1*	2',5'-oligoadenylate synthetase 1, 40/46kDa
216231_s_at	2.115	*B2M*	beta-2-microglobulin
204341_at	2.114	*TRIM16*	tripartite motif-containing 16
204415_at	2.112	*G1P3*	interferon, alpha-inducible protein (clone IFI-6-16)
209140_x_at	2.111	*HLA-B*	major histocompatibility complex, class I, B
201315_x_at	2.110	*IFITM2*	interferon induced transmembrane protein 2 (1-8D)
200887_s_at	2.110	*STAT1*	signal transducer and activator of transcription 1, 91kDa

Table 3.3. Interferon stimulated genes up-regulated 96 hours following a double transfection with AP-2γ siRNA1 compared to non-silencing control siRNA. Biotin labelled cRNA from cells harvested at 96 hours following a double transfection with nonsilencing control siRNA or AP-2γ siRNA1 was hybridised in duplicate onto Affymetrix U133A chips. (Sledz *et al.*, 2003, 2004; Bridge *et al.*, 2003; Persengiev *et al.*, 2004)

In order to avoid such an interferon response it was necessary to optimise the transient siRNA transfections. An interferon response was only triggered in the cells after the second transfection, therefore I removed the second transfection step and also aimed to reduce the concentration of siRNA transfected to the lowest effective concentration.

Figure 3.11. shows AP-2γ and STAT-1 protein levels 96 hours following a single transfection with 5nM to 200nM siRNAs. Efficient knockdown was achieved with siRNA concentrations as low as 5nM (80x less than early experiments). No STAT-1 induction was seen at any concentration tested. Biological replicates siRNA1 and siRNA2 showed similar results. Figure 3.12. shows AP-2γ and *OAS1* relative mRNA levels, measured by real time PCR (Materials and Methods, Section 2.6), 96 hours following a single transfection with siRNAs. OAS1 is an interferon induced gene that was also up-regulated in the pilot gene expression studies (Table 3.3). Again, efficient AP-2γ knockdown was achieved with siRNA concentrations as low as 5nM. No OAS1 induction was seen at any of the concentrations tested and replicates siRNA1 and siRNA2 showed similar results (see Figure 3.11). Taken together, data from Figures 3.11 and 3.12 show effective silencing of AP-2γ can be achieved using short-term transfections and without induction of the interferon pathway.

There are several other non-specific effects that could be induced by siRNAs. Firstly the degradation of mRNA species other than the target, due to cross-hybridization. Secondly the binding to cellular proteins in a sequence-specific manner via an RNA aptamer effect (Shi *et al.*, 1999). Finally a translational silencing through a micro RNA effect (Doench *et al.*, 2003). Unlike the interferon response it is difficult to monitor these "off target" effects. However, it has been reported that they too can be minimised in a concentration dependant manner (Semizarov *et al.*, 2003; Persengiev *et al.*, 2004). Figure 3.12. shows the reduction of AP-2γ mRNA levels, the primary target of dsRNA oligonucleotides AP-2γ siRNA1 and AP-2γ siRNA2. This supports a direct effect of the dsRNA oligonucleotides on the AP-2γ mRNA via a siRNA mechanism rather than a translational effect preventing production of the AP-2γ protein via a microRNA mechanism.

In summary three effective AP-2γ targeting sequences, AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3 have been characterised. It was hoped that using siRNAs targeting distinct sequences within the AP-2γ mRNA would provide biological replicates and control for "off target" effects. I planned to derive gene expression profiles comparing two control conditions (the nonsilencing control siRNA and cells treated with transfection reagent only) with the AP-2γ targeting siRNAs (AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3). Data from Figures 3.11 and 3.12 showed that using a final concentration of 25nM of siRNA gave the most effective knockdown (greater than 90% knockdown) in these transient transfection experiments and therefore this would used in subsequent experiments. Transfections performed under these conditions were also assessed using a 3' Alexa Fluor 488 labelled non-silencing control siRNA and observed via fluorescence microscopy. Fluorescence was reproducibly detected in greater than 90% of transfected cells reflecting the level of AP-2γ knockdown observed following transfection with the AP-2γ targeting siRNAs.

3.5. Levels of the established AP-2 coactivators: CITED2, CITED4, CBP and p300 following AP-2γ silencing in MCF-7 cells

Work introduced in detail in section 1.2.2, has shown that AP-2 proteins can be co-activated by the recruitment of p300/CBP to their target genes and that this recruitment is mediated by the CITED adaptor proteins CITED2 and CITED4 (Bamforth *et al.*, 2001; Braganca *et al.*, 2002; Braganca *et al.*, 2003). As the abundance of each of these proteins is likely to influence the ability of AP-2 factors to activate their target genes, it was important to assess if the level of these key co-activators was perturbed following AP-2γ silencing in MCF-7 cells. Figure 3.13 shows the levels of CITED2, CITED4, CBP and p300 72 hours following AP-2γ silencing in MCF-7 cells. Compared with the controls, no change in the level of CITED2 or CITED4 was observed for any of the AP-2γ targeting siRNAs. However, a moderate increase in the abundance of both CBP and p300 was observed following AP-2γ silencing. This was unexpected but may be a consequence of the observed cell cycle phenotype.

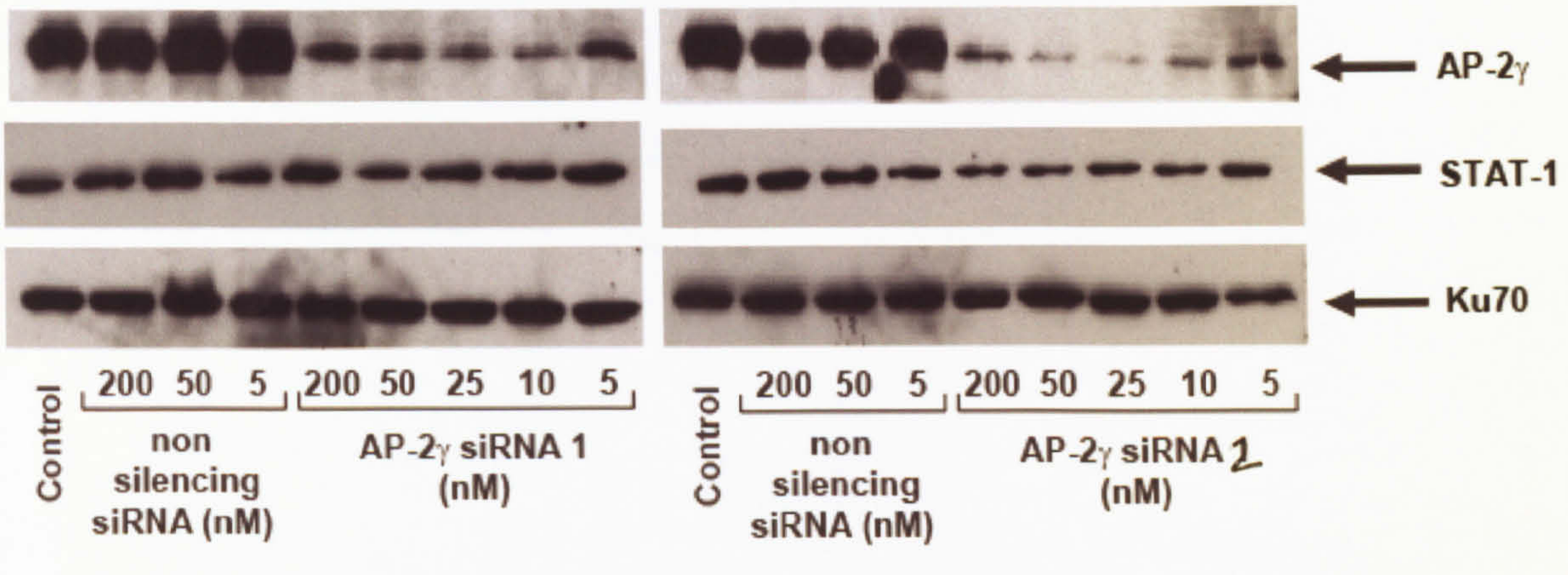


Figure 3.11. Monitoring STAT-1 induction over a range of siRNA concentrations. MCF-7 cells were transfected with siRNAs at the indicated concentrations. Cells were transfected once and harvested 96 hours later. Whole Cell extracts (10μg/lane) from MCF-7 cells were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ or STAT-1 before being reprobed for Ku-70 as a loading control. (Control - Transfection reagent only)

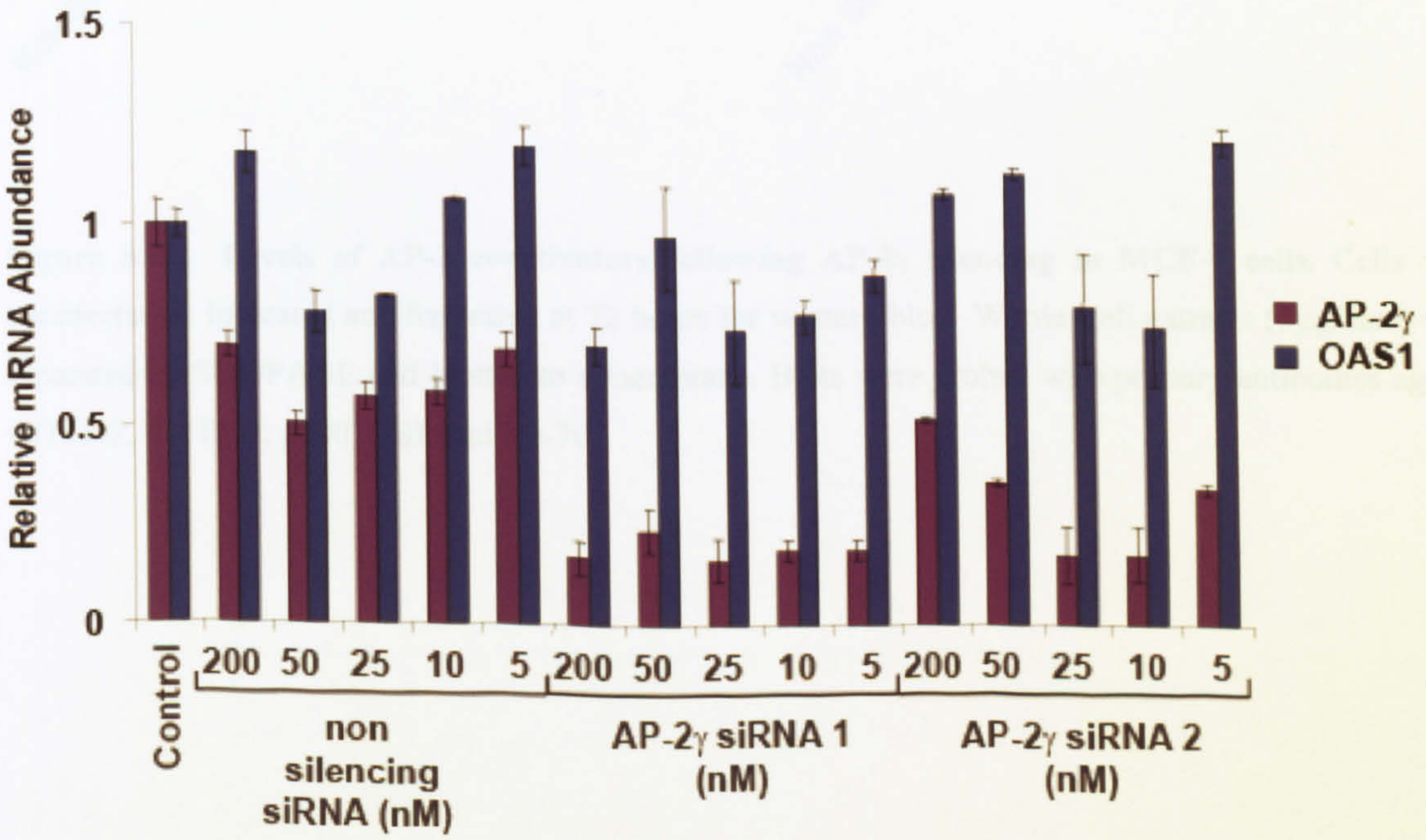


Figure 3.12. Monitoring OAS 1 mRNA induction over a range of siRNA concentrations. MCF-7 cells were transfected with siRNAs at the indicated concentrations. Cells were transfected once and total RNA harvested 96 hours later. Quantitative Taqman RT–PCR analysis of AP-2γ and OAS1 mRNA levels was performed. Data were normalised to the GAPDH RNA levels. The transfection reagent only value was set at 1 to generate relative mRNA levels. The graph represents data averaged from three replicates ± standard deviation.

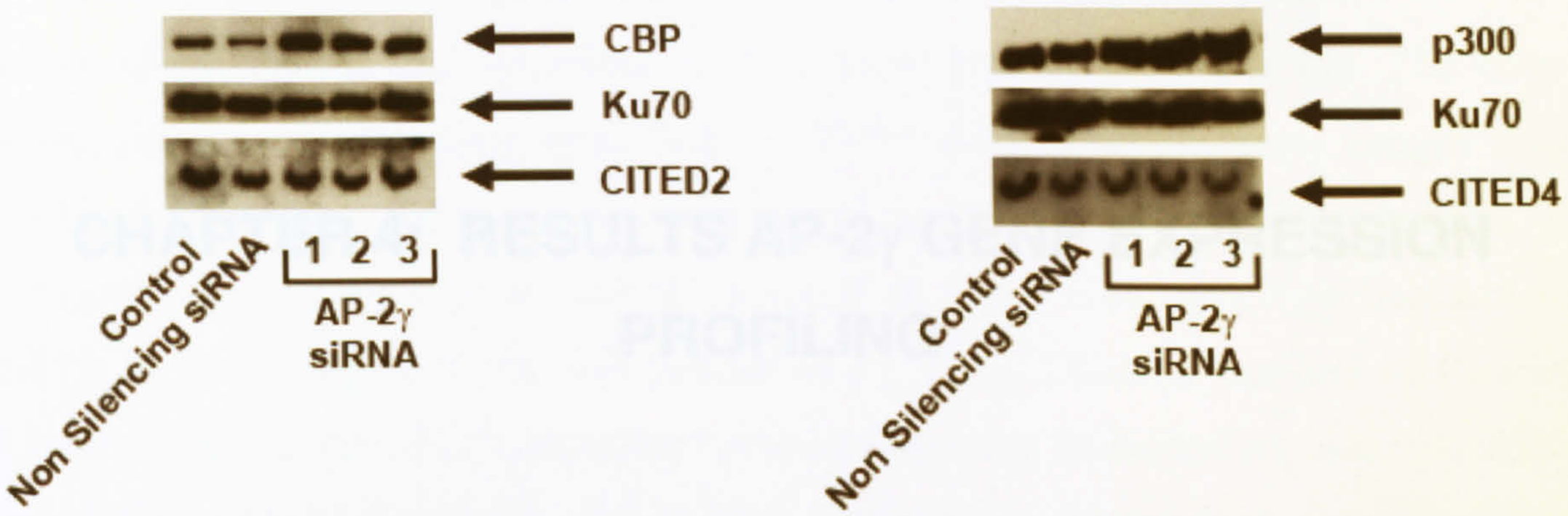


Figure 3.13. Levels of AP-2 coactivators following AP-2γ silencing in MCF-7 cells. Cells were transfected as indicated and harvested at 72 hours for western blot. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against CITED2, CITED4, p300, CBP and Ku-70.

CHAPTER 4: RESULTS AP-2γ GENE EXPRESSION PROFILING

4.1. Array processing and quality control

A more detailed description of the array procedure is given in Materials and Methods section 2.7. Briefly, total RNA was extracted from MCF-7 cells 96 hours following transfection with AP-2γ targeting siRNAs (AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3), a non silencing control siRNA, and cells treated with transfection reagent alone. The RNA was treated with DNase to remove any genomic DNA contaminants and quality was assessed via a formaldehyde-MOPS denaturing agarose gel. High yield biotinalyted cRNA (A260/A280 ratio between 1.8 and 2.1) was prepared for each experimental sample and hybridised to Affymetrix HG-U133Plus2 arrays. The chips were scanned using the GeneArray Scanner 3000. After scanning, array images were assessed by eye to confirm scanner alignment, the absence of significant bubbles or scratches on the chip surface, and the absence of slides with very high background (scanning and image analysis was performed by Tracy Chaplin, Institute of Cancer, Charterhouse Square). Each experiment was performed in biological triplicates to allow appropriate statistical analysis. A summary of the array experiments is shown in Table 4.1

Raw data were analysed using the BioConductor statistical package (Gentleman *et al.*, 2004; BioConductor, *url*) as described in Materials and Methods 2.7. The raw data, acquired from the scanner, includes information on background intensity values, perfect match, (PM) and mismatch, (MM) values and is contained in *.CEL files.

As well as visually assessing RNA quality before cRNA target preparation, information from the raw array data also provided a measure of the quality of hybridised material. Figure 4.1 shows an array-by-array cRNA digestion plot produced using the "affy" BioConductor *plotAffyRNAdeg* function (Gautier *et al.*, 2004). This plot illustrates that no severe degradation of the targeted cRNA occurred, except for a small degree of degradation at the 3' of the molecule; importantly the plots are consistent for all fifteen samples.

Experimental Condition	Replicate	File Name	Sample Type
MCF-7 Transfection Reagent Alone	1	MCF7TC.CEL	*reference*
	2	MCF7TC2.CEL	*reference*
	3	MCF7TC3.CEL	*reference*
MCF-7 Non-Silencing siRNA	1	MCF7NS.CEL	*reference*
	2	MCF7NS2.CEL	*reference*
	3	MCF7NS3.CEL	*reference*
MCF-7 AP-2γ siRNA1	1	MCF7GB.CEL	*test*
	2	MCF7GB2.CEL	*test*
	3	MCF7GB3.CEL	*test*
MCF-7 AP-2γ siRNA2	1	MCF7GE.CEL	*test*
	2	MCF7GE2.CEL	*test*
	3	MCF7GE3.CEL	*test*
MCF-7 AP-2γ siRNA3	1	MCF7GF.CEL	*test*
	2	MCF7GF2.CEL	*test*
	3	MCF7GF3.CEL	*test*

Table 4.1: Affymetrix gene expression profiling: experimental design. Table shows each experimental condition with details of replicates, CEL files and sample type.

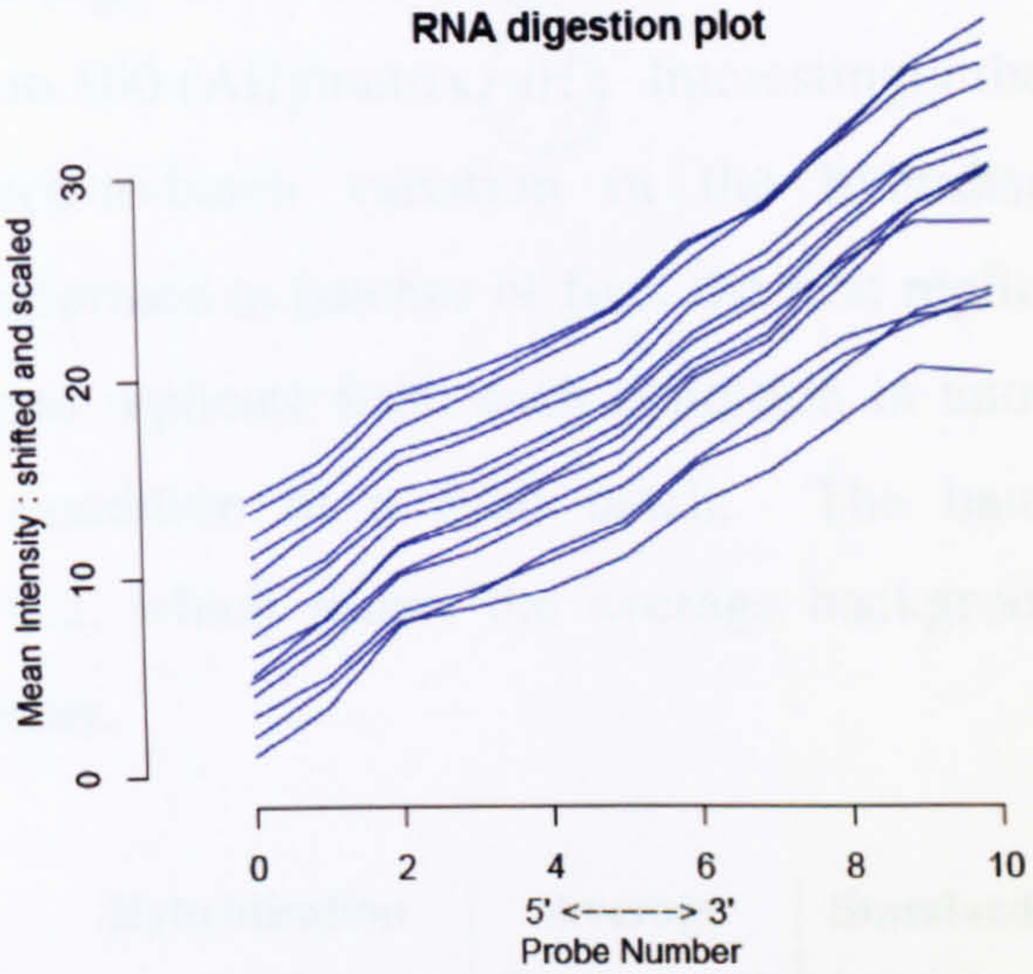


Figure 4.1 An assessment of target cRNA quality. Side-by-side cRNA digestion plot produced using the "affy" BioConductor *plotAffyRNAdeg* function. Individual probes in a probe set were ordered by location relative to the 5' end of the targeted RNA molecules. Typically RNA degradation starts from the 5' end of the molecule. Therefore probe intensities were expected to be systematically lowered at the 5' end compared to the 3' end. This plot analysed each array: probe intensities were averaged by location in the probe set, with an average then taken over all probe sets. Then side-by-side plots of these means exhibited the 5' to 3' trend and illustrate severity of target cRNA degradation. The plot illustrates that there was no severe degradation of the targeted cRNA except for a small degree of degradation at the 3' of the molecule, which is consistent in all the samples.

In order to ensure the arrays fulfil a series of Affymetrix recommended quality control metrics (described in detail in the Expression Analysis Fundamentals Manual, Affymetrix, *url*), raw data were pre-processed using "simpleaffy" BioConductor package. Simpleaffy uses the Affymetrix MAS 5.0 array analysis algorithms to provide access to these metrics (Wilson & Miller, 2005; Materials and Methods 2.7.3). Figure 4.2 shows an array-by-array quality control plot generated using "simpleaffy" BioConductor *plotQC* function. This plot highlights variability in background across the fifteen arrays although all the background values falls within the Affymetrix suggested range of 20 to 100 (Affymetrix, *url*). Interestingly the background variability observed reflects batch-to-batch variation in the hybridisation process. Array hybridisations were performed in batches of five: the first replicate from each condition in one batch, the second replicate from each condition in another batch and the third replicate from each condition in a final batch. The batch-to-batch variation is highlighted in Table 4.2, which shows the average background for each set of five arrays hybridised together.

Hybridisation Batch	Average Background	Standard Deviation
First Replicates	64	6
Second Replicates	77	7
Third Replicates	91	9

Table 4.2 shows the average background for each batch of five arrays hybridised together. First Replicates: MCF7TC.CEL, MCF7NS.CEL, MCF7B.CEL, MCF7E.CEL, MCF7F.CEL. Second Replicates: MCF7TC2.CEL, MCF7NS2.CEL, MCF7B2.CEL, MCF7E2.CEL, MCF7F2.CEL. Third Replicates: MCF7TC3.CEL, MCF7NS3.CEL, MCF7B3.CEL, MCF7E3.CEL, MCF7F3.CEL.

It is thought that this variability in background, although slightly high, is within acceptable limits for this experimental design and importantly it occurs equally across the experimental conditions and therefore its effect will not impair or skew the following downstream data analysis towards a particular condition. Figure 4.2 also shows that this observed variability in background has no effect on % present value and mas5 scale factors which both remain within acceptable bounds. None of the arrays stand out as having high *GAPDH* or β-actin 3'/5' ratios confirming the data in Figure 4.1 showing no severe degradation of the targeted cRNA.

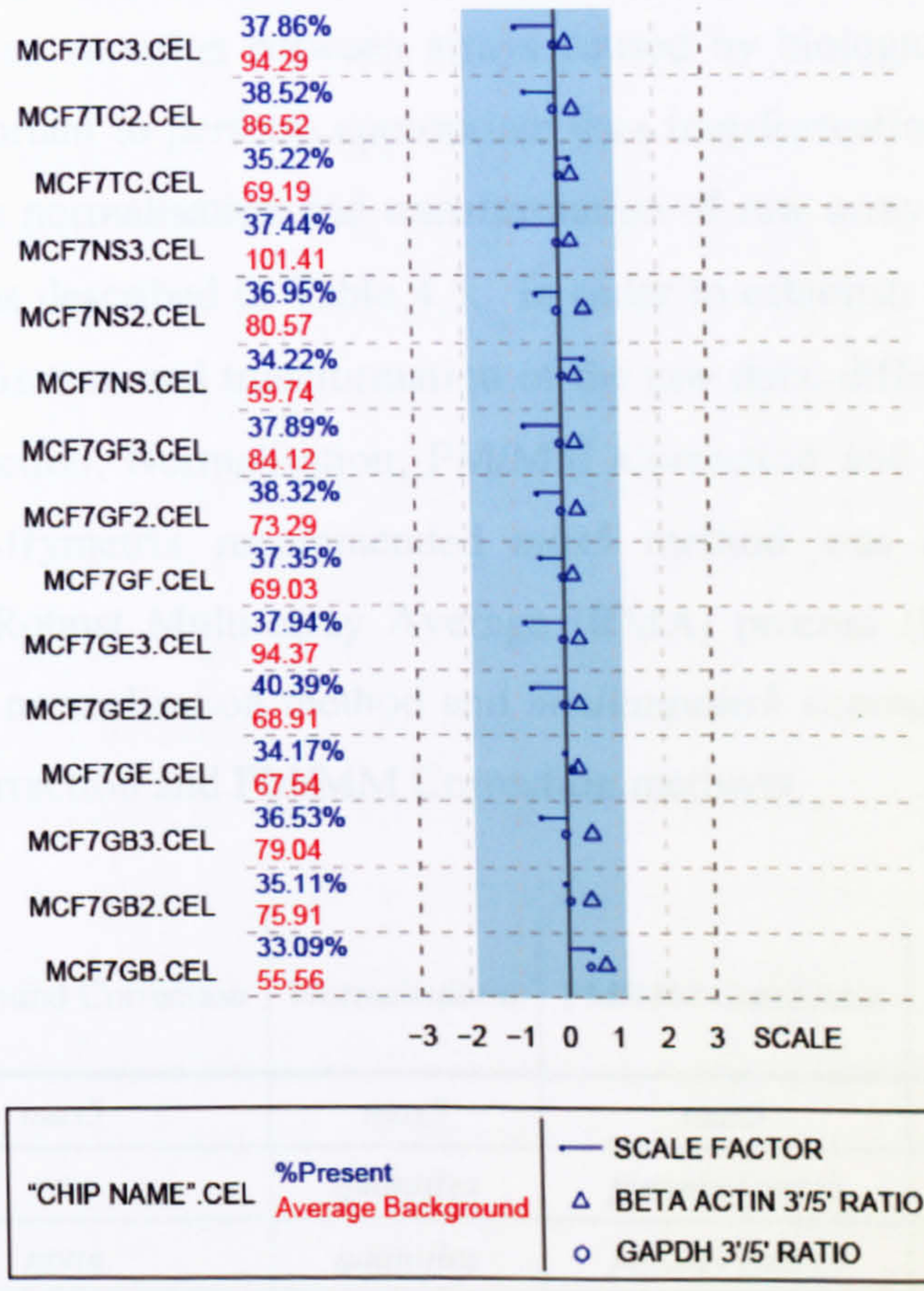


Figure 4.2 Affymetrix Quality Control Plot. Affymetrix recommended quality control metrics were produced using the "simpleaffy" BioConductor *plotQC* function in order to assess the average background, scale factor, percentage of genes called present and 3' to 5' ratios for β-actin and GAPDH. Each line represents a *mas5* normalised array. For each line, top value % present call; bottom value, average background; circle, *GAPDH* ratio; triangle, beta actin ratio; line, scale factor. The shaded region represents the range where scale factors are within 3-fold of each other (centred around the mean). Although variability in background is observed across the arrays the % present value and mas5 scale factors remain with acceptable bounds. No arrays stand out as having high GAPDH and β-actin ratios supporting the data in Figure 4.1 showing no severe degradation of the targeted cRNA.

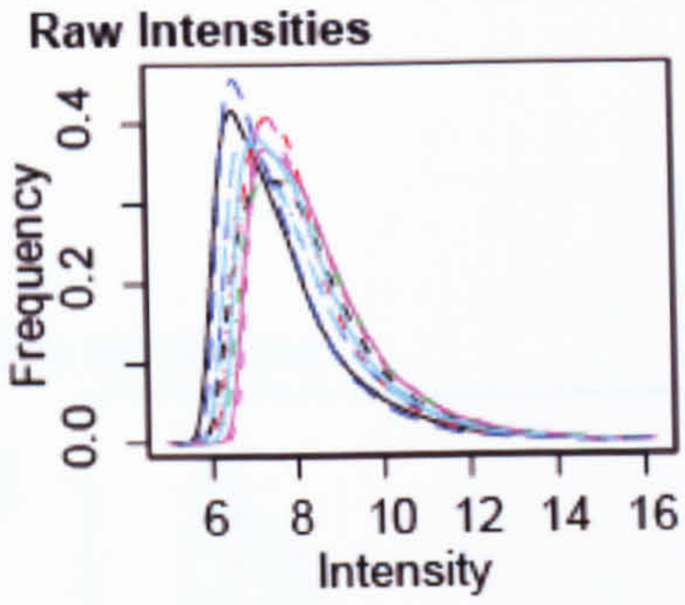
4.2. Normalisation and transformation of Raw Array Data

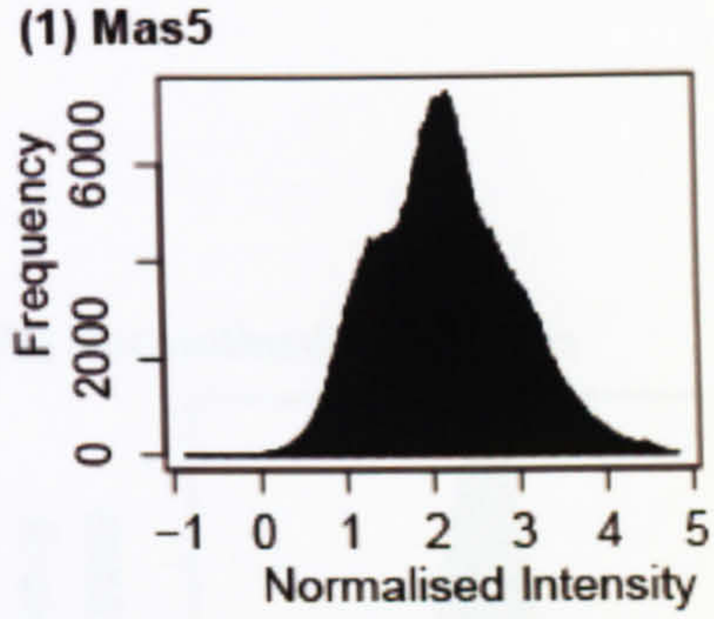
The aim of these analyses was to identify deferentially expressed genes in MCF-7 cells that were transfected with AP-2γ targeting siRNAs compared to the control conditions. In order to minimise variation between arrays caused by biological and experimental factors it was important to perform appropriate data transformation and normalisation. The analysis of the normalisation and transformation of raw array data took place in a four-step process as described in Table 4.3. In order to establish the most appropriate method for normalisation and transformation of the raw data, different combinations of Background Correction, Normalisation, PM/MM Correction and Summarisation were assessed. The Affymetrix recommended *mas5* method was assessed along with variations of the Robust Multi-array Average (RMA) process (Irizarry *et al.*, 2003) using the *quantile* normalisation method and *medianpolish* summarisation, but altering the background correction and PM/MM Correction methods.

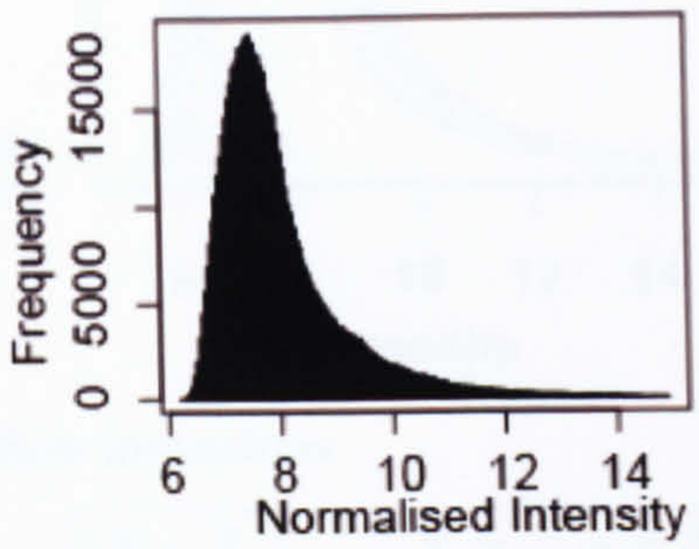
	Background Correction	Normalisation	PM/MM Correction	Summarisation
1	*mas5*	*mas5*	*mas5*	*mas5*
2	*rma*	*quantiles*	*pmonly* (none)	*medianpolish*
3	*none*	*quantiles*	*pmonly* (none)	*medianpolish*
4	*mas5*	*quantiles*	*pmonly* (none)	*medianpolish*
5	*none*	*quantiles*	*mas5*	*medianpolish*
6	*mas5*	*quantiles*	*mas5*	*medianpolish*

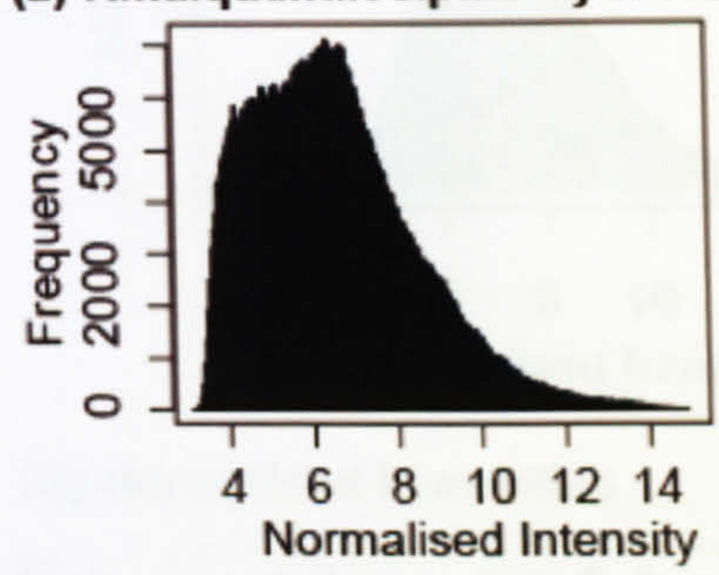
Table 4.3 Overview of the Bioconductor data analysis. Data analysis was performed using the "affy" BioConductor package using the *expresso* function in six indicated combinations for Background Correction, Normalisation, PM/MM Correction and Summarisation. Background corrections were performed using either the Affymetrix *mas5* method (average background estimates from defined array regions), or the *rma* method (PM corrected values as described by (Irizarry *et al.*, 2003)) or *none* (no background correction). Normalisation was performed using the Affymetrix recommended *mas5* method or the *quantile* normalisation method. PM correction was applied either as *pmonly*, (using only the PM normalised values), or as *mas5* (using the ideal mismatch method as suggested by Affymetrix). Summarisation analysis combined the pre-processed probe intensities together to compute an expression measure for each probe set on the array. *Medianpolish* was applied as described by (Irizarry *et al.*, 2003) and *mas5* using a robust average 1-step Tukey bi-weight on $\log_2$ scale.

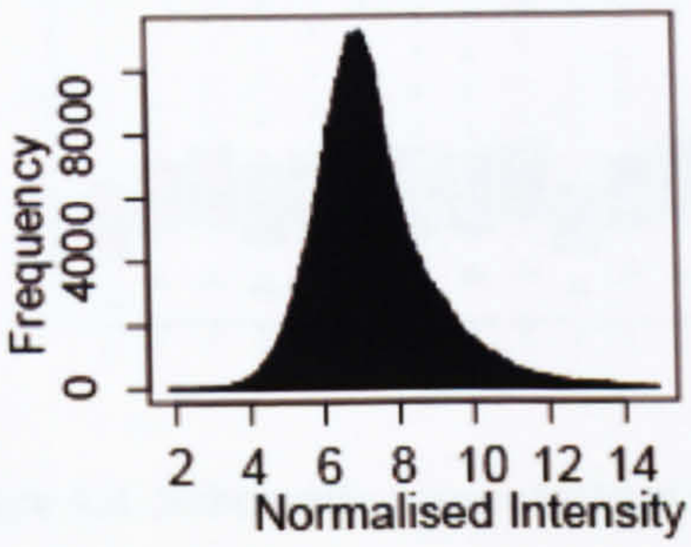
Figure 4.3 shows frequency histograms for the HG-U133Plus2 raw expression data and expression data processed by the six different methods outlined in Table 4.3. It is apparent from the frequency histograms of raw intensity values that variation between arrays is present. Box plots of all of the normalisation methods confirm that transformation of the data is such that the distribution of probe intensities for every array in the set of the fifteen arrays used in this experiment are the same and fit the same distribution (data not shown). However it is clear from the histograms that this distribution is not necessarily a normal distribution, which is an assumption for any downstream parametric statistical analysis. For example in histogram (1), the Affymetrix recommended *mas5* method, gives an *almost* bell shaped curve which has a shoulder to the left of the median values. This is most likely due to the PM/MM correction. It is widely reported that MM may be detecting signal in addition to non-specific binding and therefore, including the MM parameter, contributes to the overall noise in the data analysis (Naef *et al.*, 2002; Irizarry *et al.*, 2003; Landis *et al.*, 2004). Similar double peaked distributions are observed in histograms (5) and (6) where the *mas5* PM/MM correction is also employed. Histograms (2), (3) and (4) show normalisation methods that considered PM values only but varied in the background correction method used. All three show smooth distributions, but (4) employing *mas5* background correction is most similar to a normal distribution. Therefore it was considered that *mas5* background correction, *quantile* normalisation, PM correction using PM values only (*pmonly*) and *medianpolish* summarisation method were the most effective processing steps for intra- and inter-array normalisation, and data processed this way would fulfil the assumptions made for parametric statistical analysis. Figure 4.4 shows schematic presentations of the HG-U133plus2 raw and *mas5.quaniltes.pmonly.medianpolish* normalised expression data. This confirms the data appears to be transformed in such a way that the distribution of probe intensities for every array in the set of the fifteen arrays used in this experiment are the same and the data resembles a normal distribution.

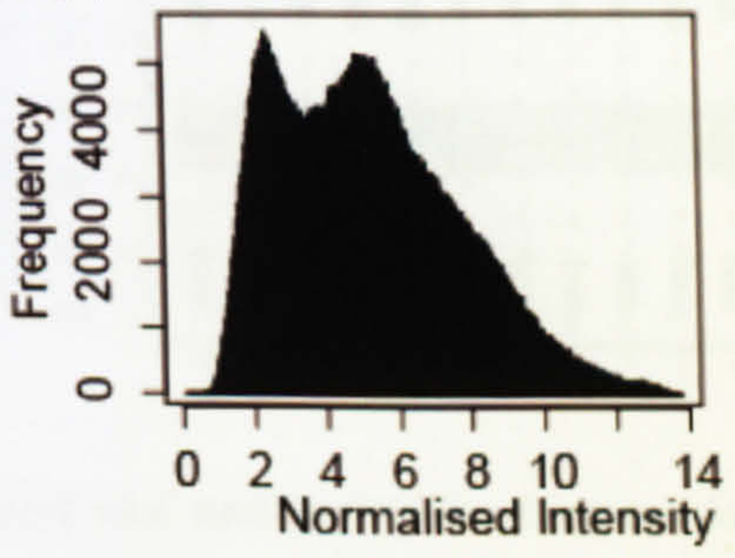












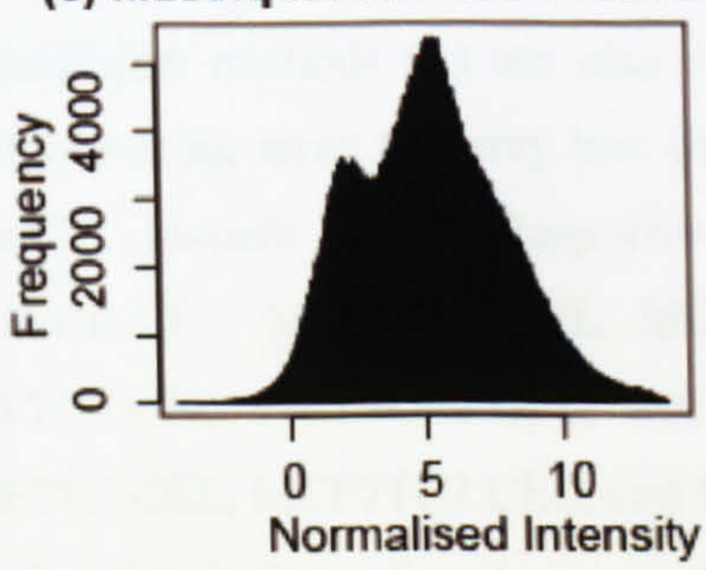


Figure 4.3. Frequency Histograms of the HG-U133Plus2 raw and normalised expression data. This Figure shows plots of the densities for probe intensities of the fifteen human arrays datasets, with their distributions before normalisation (Raw Intensities) and after normalisation using the methods outlined in Table 4.3 (1-6). Histograms were produced using the "affy" BioConductor *Hist* function.

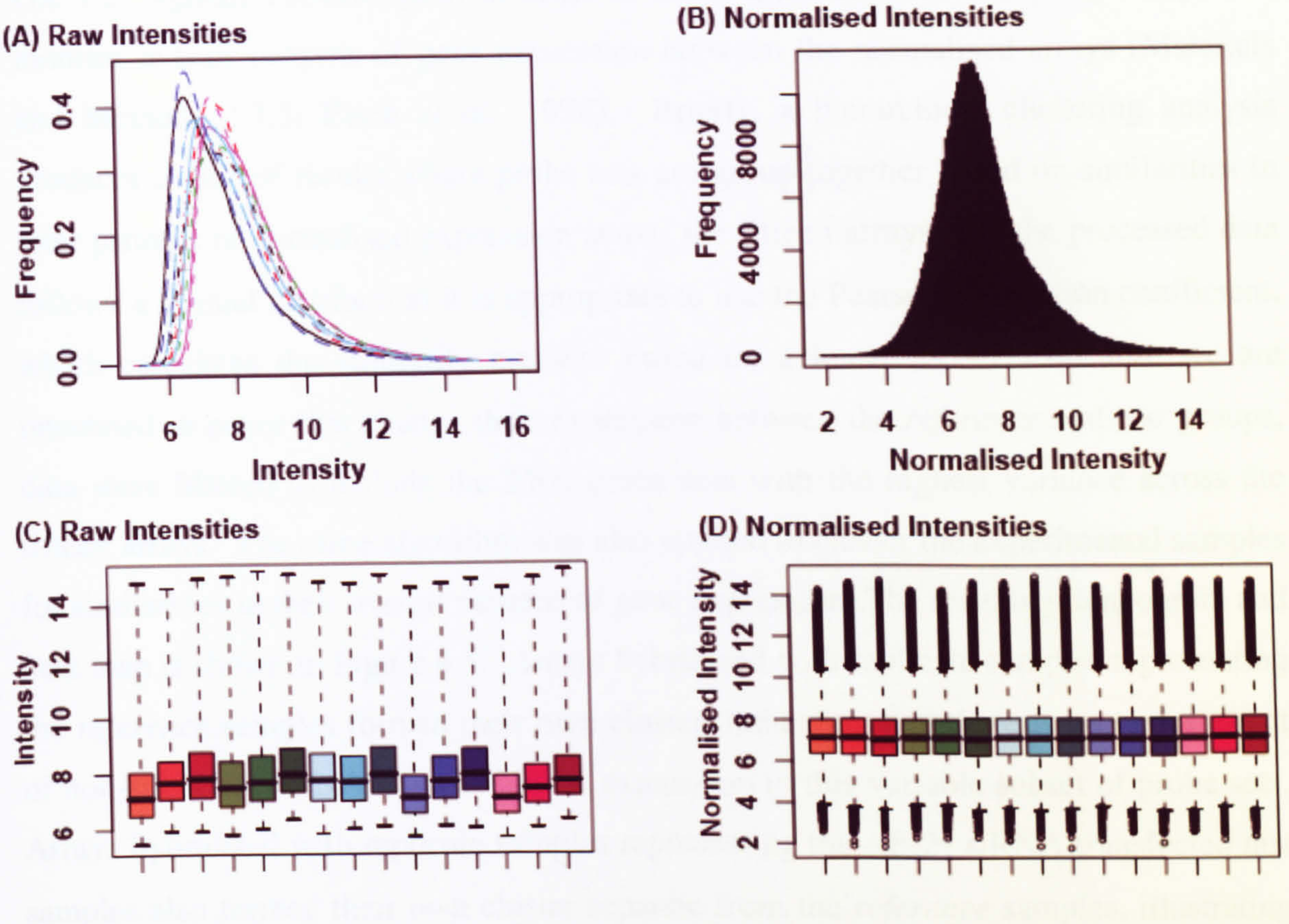


Figure 4.4. Schematic presentations of the HG-U133plus2 raw and normalised expression data. The Figure shows plots of the densities for probe intensities of the fifteen human arrays datasets, with distribution before normalisation (A) and after normalisation using the *mas5, quantiles, pmonly* and *medianpolish* methods (B) are also displayed. The distribution of the average expression measure is displayed using array by array box plots of raw data (C) and data after normalisation using the *mas5, quantiles, pmonly* and *medianpolish* methods (D). Box plots array by array from left to right: MCF7GB.CEL. MCF7GB2.CEL, MCF7GB3.CEL, MCF7GE.CEL, MCF7GE2.CEL, MCF7GE3.CEL, MCF7GF.CEL MCF7GF2.CEL, MCF7GF3.CEL, MCF7NS.CEL, MCF7NS2.CEL, MCF7NS3.CEL, MCF7TC.CEL, MCF7TC2.CEL, and MCF7TC3.CEL. The data appears to be transformed in such a way that the distribution of probe intensities for every array in the set of the fifteen arrays used in this experiment are the same and fit a normal distribution. The plots were produced using the "affy" BioConductor *hist* and *boxplot* functions.

4.3. Hierarchical Clustering Analysis

Bioconductor normalised data were imported into the Genespring package (Genespring GX 7.3, Agilent Technologies) in order to allow a hierarchical clustering analysis to monitor overall patterns of gene expression between the normalised arrays (Materials and Methods 2.7.3; Eisen *et al.,* 1998). Briefly, a hierarchical clustering analysis produces a map of results where probe sets are group together based on similarities in their patterns of normalised expression across the fifteen arrays. As the processed data follows a normal distribution it is appropriate to use the Pearson correlation coefficient, which calculates the similarity measure based on a linear model. Because we are interested in genes that change their expression between the *reference* and *test* groups, data were filtered to include the 2500 probe sets with the highest variance across the fifteen arrays. The same algorithm was also applied to cluster the experimental samples for similarities in their overall patterns of gene expression. The resulting dendogram and heat map is shown in Figure 4.5. Arrays hybridised with replicate samples representing the *reference* samples formed their own cluster, indicating that there is no major effect of non-silencing siRNA transfection on expression in this variable subset of probe sets. Arrays hybridised with replicate samples representing the AP-2γ siRNA transfected *test* samples also formed their own cluster separate from the *reference* samples, illustrating an effect of AP-2γ siRNA transfection on expression these probe sets. Interestingly within this AP-2γ siRNA cluster the replicates from each AP-2γ targeting siRNA (siRNA1, siRNA2 and siRNA3) form their own sub-clusters together. Particularly noticeable was AP-2γ siRNA2, which formed its own sub-cluster away from AP-2γ siRNA1 and AP-2γ siRNA3 replicates.

These differences in clustering between different AP-2γ targeting siRNAs could be due to differences in efficiency of AP-2γ silencing. Indeed, examination of the *TFAP2C* probesets showed differences in efficiency of silencing of AP-2γ transcript levels between the siRNAs used. As shown in Figure 4.6, AP-2γ siRNA2 silencing appears to have been more efficient than silencing from AP-2γ siRNA1 and AP-2γ siRNA3 replicates.

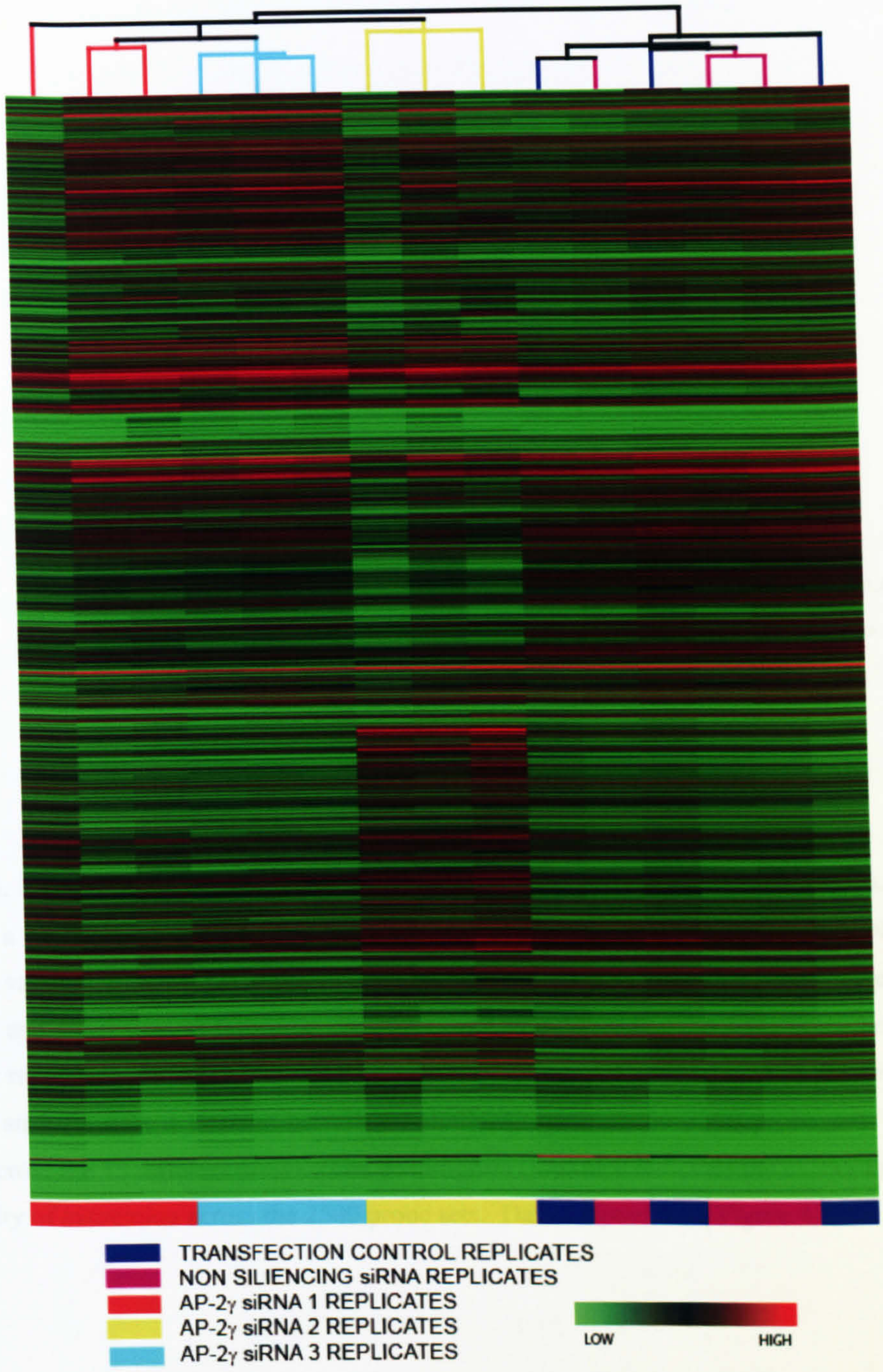


Figure 4.5. Cluster diagram showing 2500 most variable probes sets across the fifteen arrays. Each column represents a single array experiment, and each row represents a single probe set. Normalised gene expression values are shown for arrays hybridised with the samples indicated. Red squares represent high expression and green squares represent low expression as indicated on the scale at the bottom right. The colour bar represents the replicates from different conditions. The Heat map represents a hierarchical clustering analysis with a Pearson correlation similarity measure on 2500 probe sets with the highest variance across the 15 different arrays. The dendrogram (top) also indicates the clustering of arrays based on similarity of expression across the 2500 probe sets.

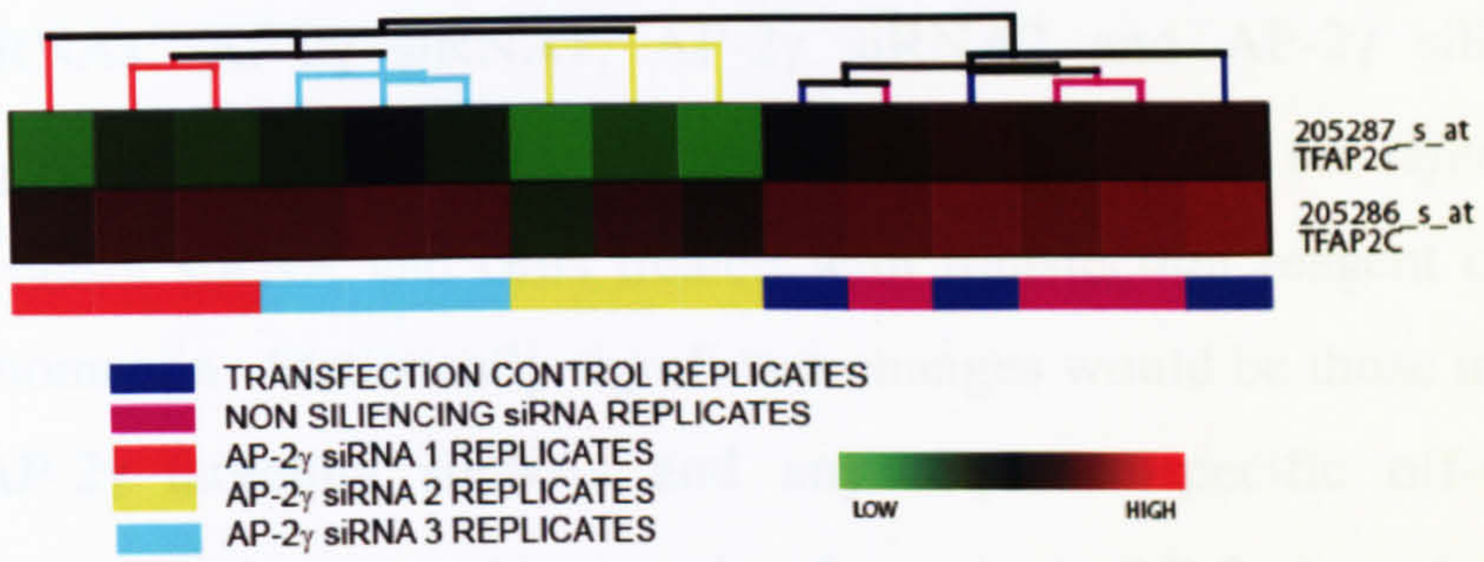


Figure 4.6. Cluster diagram showing AP-2γ probes sets across the fifteen arrays. Each column represents a single array experiment, and each row represents a single probe set. Normalised gene expression values are shown for arrays hybridised with the samples indicated. Red squares represent high expression and green squares represent low expression as indicated on the scale at the bottom right. The colour bar represents the replicates from different conditions. The Heat map represents a hierarchical clustering analysis with a Pearson correlation similarity measure on 2500 probe sets with the highest variance across the 15 different arrays. The dendrogram (top) also indicates the clustering of arrays based on similarity of expression across the 2500 probe sets. Data adapted from Figure 4.5.

Differences in clustering between different AP-2γ targeting siRNAs might also be due to sequence specific off-target effects of the siRNAs. Each AP-2γ targeting siRNA might be directly silencing the transcripts of genes other than AP-2γ, thus causing variations in clustering patterns. Optimisation experiments have served to reduce these potential sequence specific off-target effects (Figures 3.11 & 3.12), which have been shown to be minimised in a concentration dependant manner (Semizarov *et al.*, 2003; Persengiev *et al.*, 2004). Off-target effects have also been addressed in the experimental design, by including three distinct siRNA sequences targeting AP-2γ. It was thought that treating all the AP-2γ targeting siRNAs together as one variable, the *test* group (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) and conducting stringent statistical comparisons between this group and the *reference* group (non silencing control siRNA and cells treated with transfection reagent only) would address this phenomenon. Statistically significant changes would be those in agreement for all three AP-2γ targeting siRNAs and any sequence-specific off-target gene expression changes (i.e. changes only occurring for a single AP-2γ targeting sequence) would thus be eliminated from the final data set.

4.4. Differential Gene Expression Analysis

Bioconductor normalised data were imported into Genespring in order to allow statistical analyses to be carried out. As well as fitting a normal distribution, another assumption of statistical analyses, such as the widely used Student's t-test, is that the variability of a gene is constant across treatment types. However, in the absence of accurate diagnosis methods, it is safest to assume that variability may differ between our *reference* and *test* groups. As we have confirmed that our data fits a normal distribution, it is appropriate to conduct a Welch's t-test, this is a parametric analysis that corrects for difference in variability.

As we are interested in identifying genes that change their expression between the *reference* and *test* groups, it was important first to filter the data to include only those probe sets that change in expression. We chose to order the list from low to high standard deviation. This would enable us to identify the genes that are changed the

most in expression across the fifteen arrays, and remove probe sets from the analysis that change very little. It was felt that this is more informative than filtering on fold change, for which there is a risk of ignoring genes that change significantly but are below the arbitrary fold change threshold.

Genespring was used to calculate a Welch's t-test on the 2500 probe sets with highest variance across the fifteen arrays. A Benjamini and Hochberg False Discovery Rate (FDR) test (Benjamini & Hochberg, 1995) was also applied in order to identify the percentage of false positive genes that could have been included in the final lists of differentially expressed genes. Appendix 1 displays the 578 probe sets that changed their expression significantly (FDR corrected, $p<0.05$) between the *reference* (non silencing control siRNA and cells treated with transfection reagent only) and *test* (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) groups.

4.5. AP-2γ Gene Expression Profile Initial Observations

Significant changes in gene expression were observed after silencing of AP-2γ in MCF-7 cells and are summarised in Table 4.4. Many genes regulated in these experiments were represented by multiple probe sets therefore the actual number of genes regulated is indicated in brackets. A smaller subset of 300 probe sets (254 genes) regulated at a $p<0.01$ probability was used for subsequent Gene Ontology, Pathway Mapping and Transcription Factor Binding Site analysis as it was felt that this provided a more manageable data set and would provide more statistical relevance when conducting these large scale analyses. In the results presented in this thesis, probe sets were assigned to gene symbols and gene descriptions based on the January 2007 release of the HG-U133_Plus2 Affymetrix NetAffx Annotation files (Liu et al, 2003; Affymetrix *url*).

	p < 0.05	p < 0.01
Total probe sets (genes) differentially regulated	578 (469)	300 (254)
Total probe sets (genes) down regulated	389 (315)	225 (190)
Total probe sets (genes) up-regulated	189 (154)	75 (64)

Table 4.4. Significant changes in gene expression following AP-2γ silencing. Changes observed in probe sets between the *reference* (non silencing control siRNA and cells treated with transfection reagent only) and *test* (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) groups at the indicated False Discovery Rate corrected p values. Number of genes represented is indicated in brackets.

Encouragingly within these significant expression changes, probes sets specific to AP-2γ are represented as illustrated in Figure 4.7. As expected, AP-2γ transcript levels are down-regulated in response to AP-2γ targeting siRNA in MCF-7 cells and this change was detected by both AP-2γ HG-U133Plus2 hybridising probe sets (*TFAP2C* 205286_at: Fold Change = -2.09, p=0.00439, *TFAP2C* 205287_s_at: Fold Change = -2.39, p=0.00646). These results compare favourably to the qPCR data presented in Figure 3.12. In MCF-7 cells transfected with 25nM of AP-2γ targeting siRNA or non-silencing control RNA and harvested 96 hours later resulted in a fold down regulation of -3.5 and -3.3 for siRNAs 1 and 2 respectively.

Also within the significant expression changes were genes previously suggested to be transcriptionally regulated by AP-2γ or the family member AP-2α. As AP-2α and AP-2γ has been shown to bind identical DNA recognition sequences *in vitro* (McPherson & Weigel, 1999), genes previously suggested to be regulated by AP-2α are thought to have some relevance to this AP-2γ dependant data set. These changes are summarised in Figure 4.7. Significant down regulation of *MT2A* was supported by previous work (Mitchell *et al.*, 1987), showing AP-2α dependent *in vitro* transcription from the human *MT2A* promoter in HeLa cells. Interestingly paralogs of *MT2A* (*MT1F*, *MT1H*, and *MT1X*) were also significantly down-regulated following AP-2γ knockdown in MCF-7 cells suggesting a common mechanism for their transcriptional regulation. *TGFα* was also significantly regulated following AP-2γ silencing. An increasing body of evidence suggests that AP-2α plays an important role in TGFα transcriptional activation (Gille *et*

al., 1997; Wang *et al.*, 1997; Kim & Rho, 2002). Data from this study shows TGFα up-regulation following AP-2γ silencing. This suggests AP-2γ plays a contrasting role to AP-2α, and could be involved in the transcriptional repression of TGFα. *In vitro* reporter assays implicate AP-2α in the activation of transcription from the *IGFBP5* promoter via AP-2 DNA recognition sites (Duan & Clemmons, 1995; Erclik & Mitchell, 2005). Data from this study shows *IGFBP5* up-regulation following AP-2γ silencing; again suggesting AP-2γ plays a different role, and could be involved in the transcriptional repression of *IGFBP5*. Work in our lab supports this hypothesis and using ChIP assays it has been shown that AP-2γ is bound to the *IGFBP5* promoter in MCF-7 cells (Karsten Friedrich, *personal communication*). AP-2 transcription factors have also been shown to be involved in the regulation of transcription from keratin promoters during epidermal differentiation (See Introduction 1.4.2 for details). Consistent with these observations, three keratin transcripts (*KRT8*, *KRT81*, *KRT86*) are significantly down-regulated following AP-2γ silencing.

A list of the top 30 most significantly changing probe sets is displayed in Figure 4.8. These include several genes that have been selected for follow up and will be discussed further in Results Section 4.9.

PROBE ID	GENE DESCRIPTION		FOLD CHANGE	P VALUE
205287_s_at	TFAP2C	transcription factor AP-2 gamma	-2.09	0.00439
205286_at	TFAP2C	transcription factor AP-2 gamma	-2.39	0.00646
217165_x_at	MT1F	metallothionein 1F (functional)	-1.51	0.00934
208581_x_at	MT1X	metallothionein 1X	-1.72	0.00916
206461_x_at	MT1H	metallothionein 1H	-1.68	0.00793
212185_x_at	MT2A	metallothionein 2A	-1.96	0.00511
204326_x_at	MT1X	metallothionein 1X	-1.49	0.0485
216568_x_at	KRT8	Keratin 8	-1.66	0.000383
213711_at	KRT81	keratin 81	-4.02	0.000275
215189_at	KRT86	keratin 86	-2.12	0.00344
205016_at	TGFA	transforming growth factor, alpha	1.78	0.0074
211959_at	IGFBP5	insulin-like growth factor binding protein 5	2.18	0.00817
1555997_s_at	IGFBP5	insulin-like growth factor binding protein 5	1.46	0.0141

Figure 4.7. Significant changes in gene expression following AP-2γ silencing related to AP-2 biology. Changes observed in probe sets between the *reference* (non silencing control siRNA and cells treated with transfection reagent only) and *test* (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) groups at the indicated False Discovery Rate corrected p values. A full list of significantly regulated probe sets together with normalised expression values, fold change values (between *test* and *reference* groups) and FDR corrected p values is contained in Appendix 1.

PROBE ID	GENE DESCRIPTION		FOLD CHANGE	P VALUE
221874_at	KIAA1324	KIAA1324	-2.57	8.43E-06
224925_at	PREX1	phosphatidylinositol 3,4,5-trisphosphate- RAC exchanger 1	-1.89	8.43E-06
232306_at	CDH26	cadherin-like 26	2.07	8.43E-06
203571_s_at	C10orf116	chromosome 10 open reading frame 116	-2.53	1.09E-05
228176_at	EDG3	endothelial differentiation, sphingolipid G-PCR, 3	-2.49	1.09E-05
204244_s_at	DBF4	DBF4 homolog (S. cerevisiae)	-2.08	1.09E-05
225207_at	PDK4	pyruvate dehydrogenase kinase, isozyme 4	1.96	1.09E-05
235408_x_at	ZNF117	Zinc finger protein 117 (HPF9)	2.01	1.09E-05
201458_s_at	BUB3	BUB3 budding uninhibited by benzimidazoles 3 homolog (yeast)	-1.77	1.56E-05
210511_s_at	INHBA	inhibin, beta A (activin A, activin AB alpha polypeptide)	1.97	1.56E-05
203060_s_at	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	2.18	1.92E-05
210868_s_at	ELOVL6	ELOVL family member 6, elongation of long chain fatty acids	-1.72	2.62E-05
243010_at	MSI2	musashi homolog 2 (Drosophila)	-2.01	3.15E-05
220414_at	CALML5	calmodulin-like 5	-1.90	3.15E-05
220192_x_at	SPDEF	SAM pointed domain containing ets transcription factor	-2.24	3.76E-05
204070_at	RARRES3	retinoic acid receptor responder (tazarotene induced) 3	-2.06	3.76E-05
222557_at	STMN3	stathmin-like 3	-1.97	3.76E-05
209754_s_at	TMPO	thymopoietin	-1.78	3.76E-05
203304_at	BAMBI	BMP and activin membrane-bound inhibitor homolog (X. laevis)	-2.10	4.41E-05
201896_s_at	PSRC1	proline/serine-rich coiled-coil 1	-1.78	5.29E-05
227850_x_at	CDC42EP5	CDC42 effector protein (Rho GTPase binding) 5	-2.16	5.36E-05
212148_at	PBX1	Pre-B-cell leukemia transcription factor 1	-1.79	5.36E-05
227140_at	INHBA	inhibin, beta A (activin A, activin AB alpha polypeptide)	3.06	5.36E-05
205240_at	GPSM2	G-protein signalling modulator 2 (AGS3-like, C. elegans)	-1.78	5.38E-05
213599_at	OIP5	Opa interacting protein 5	-1.72	5.38E-05
239007_at	ZNF616	zinc finger protein 616	1.87	5.38E-05
203058_s_at	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	1.90	5.38E-05
212143_s_at	IGFBP3	insulin-like growth factor binding protein 3	2.30	5.38E-05
204319_s_at	RGS10	regulator of G-protein signalling 10	-1.75	5.90E-05
210095_s_at	IGFBP3	insulin-like growth factor binding protein 3	2.39	5.90E-05

Figure 4.8. The top thirty most significant changes in gene expression following AP-2γ silencing in MCF-7 cells. Changes observed in probe sets between the *reference* (non silencing control siRNA and cells treated with transfection reagent only) and *test* (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) groups at the indicated False Discovery Rate corrected p values. A full list of significantly regulated probe sets together with normalised expression values, fold change values (between *test* and *reference* groups) and FDR corrected p values is contained in Appendix 1.

4.6. Analyses of Associated Gene Ontology Terms

In order to interrogate the significant gene expression changes following AP-2γ silencing in MCF-7 cells further, and to explore any patterns in groups of genes within the data set, additional analyses were conducted based on associated Biological Process Gene Ontology Terms (Ashburner *et al.,* 2000; The Gene Ontology project, *url*). Gene Ontology Terms are organised in a hierarchical structure described in detail in Section 2.7.3. Briefly, each gene can be assigned several associated Biological Process GO terms. A GO path is where a GO term is connected to several other GO terms higher in the GO hierarchy, for example: biological process → cellular process → cellular physiological process → cell cycle. An added level of complication is that these paths intersect or split and also several paths may lead to an individual gene. This hierarchical and interconnected nature of GO terms can make the outputs of such analyses difficult to present graphically. For this reason three separate approaches were employed.

4.6.1. Hand Curated Analysis of Biological Process Ontology Terms

For an initial analysis, the AP-2γ ($p<0.01$) data set was assigned GO terms extracted from the January 2007 release of the HG-U133_Plus2 Affymetrix NetAffx Annotation files (Liu et al, 2003; Affymetrix, *url*). The data set was separated into up-regulated and down-regulated genes, then sorted into broad functional groups based on their assigned Biological Process GO terms. Each gene was assigned to a single GO term. This analysis was done by hand and appropriate terms were first decided based on the most likely biological function where two or more differing terms were described, a decision based on each GO term's associated evidence code. A suitable term higher up the GO path was then assigned to allow the genes to be grouped together into broad functional categories. This data is shown graphically in Figure 4.9. The pie chart represents expression changes for single transcripts (not probe sets) so as to avoid overrepresentation of functional classes where multiple probe sets representing one gene are significantly regulated. From this analysis it is clear that a large proportion of the

probe sets in the AP-2γ data set have no known Biological Process GO term (24% of down-regulated genes and 21% of up-regulated genes). Additionally a large number of genes down-regulated following AP-2γ silencing are classified as being involved in the cell cycle, 34% of down-regulated genes compared to just 3% of up-regulated genes. These are indicated by the "exploded" region of the pie chart in Figure 4.9. Genes assigned the functional categories of mRNA processing and Protein Biosynthesis were also represented within down-regulated changes and not the up-regulated changes. A larger number of genes assigned the functional categories of development and apoptosis were represented in the up-regulated changes compared to the down-regulated changes.

4.6.2. GO STAT Analysis of Biological Process Ontology Terms

A second separate analysis was conducted using the Gostat interface (GOstat, *url*; Beissbarth & Speed, 2004), in order to establish with statistical confidence which GO terms were over-represented or under-represented in our lists of significantly up-regulated and significantly down-regulated transcripts following AP-2γ silencing in MCF-7 cells. This analysis assigned GO Biological Process terms based on the given Affymetrix HG-U133Plus2 identifiers then compared these to all of the GO terms represented by all of the genes on the Affymetrix HG-U133Plus2 microarray. This allows an objective and statistically controlled measure of enriched GO terms from the data set. Table 4.5 shows a sub-set of 29 representative significant GO terms that were significantly enriched in our AP-2γ data set compared to GO terms on the entire Affymetrix HG-U133Plus2 microarray. It is clear for genes down-regulated following AP-2γ silencing, GO terms are highly significantly over-represented in categories related to cell cycle. This is consistent with the analysis shown in Figure 4.9, although the functional categories are more precise. Additionally, for those genes up-regulated following AP-2γ silencing GO terms are significantly over-represented in categories related to development, cell growth and apoptosis. This is also consistent with the analysis shown in Figure 4.9.

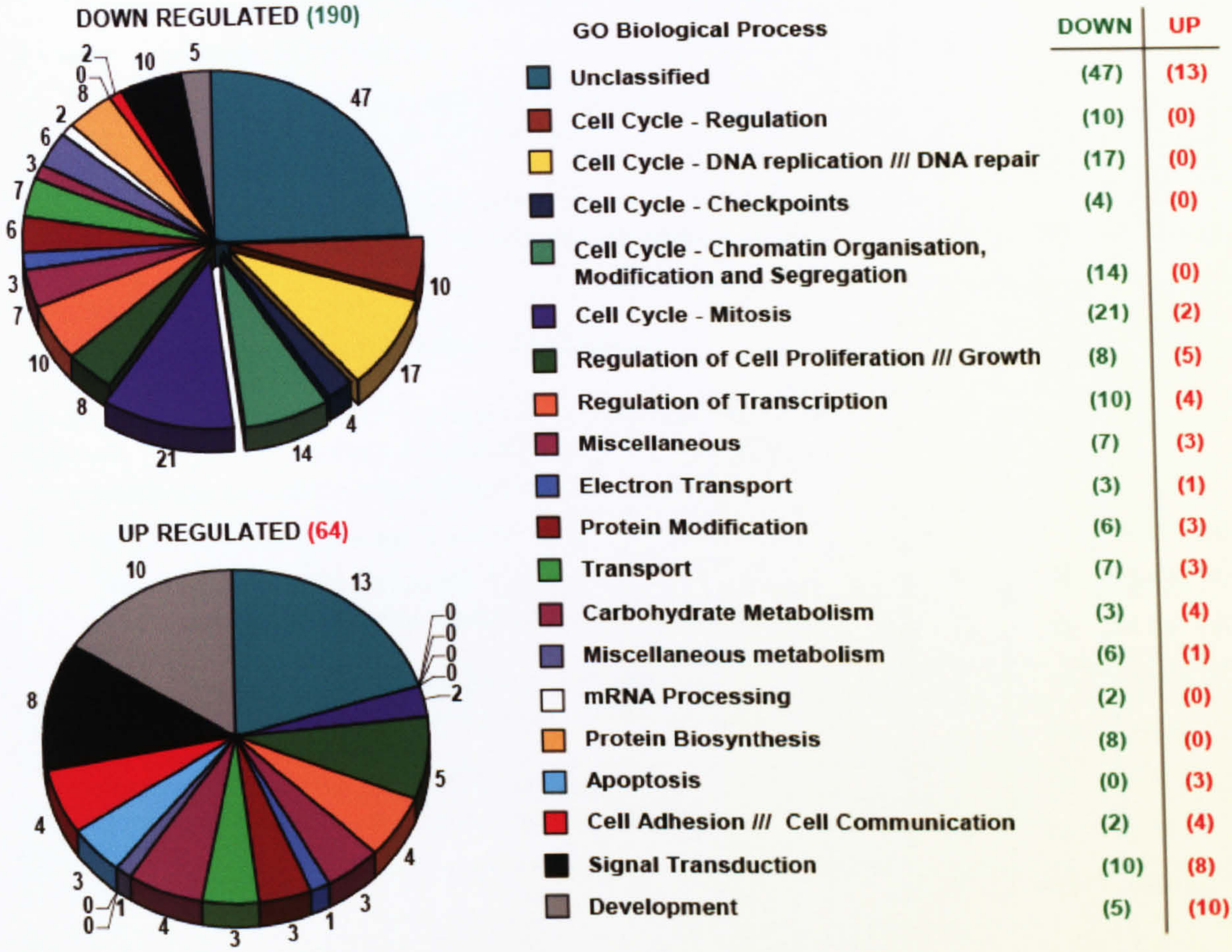


Figure 4.9. A functional classification of genes that are significantly down-regulated or up-regulated ($p < 0.01$) following AP-2γ silencing in MCF-7 cells. Pie charts represent the fraction of genes within functional categories as classified by their associated Gene Ontology Biological Process annotations. Each gene was assigned a single GO term. The numbers of genes are indicated adjacent to the pie chart or with the text (down-regulated genes in green and up-regulated in red). This pie chart represents expression changes for single transcripts (not probe sets) this is to avoid overrepresentation of functional classes where multiple probe sets representing one gene are significantly regulated. Gene ontology terms for each probe set were extracted using the Affymetrix NetAffx web interface (Liu *et al,* 2003, Affymetrix *url*).

Table 4.5. A List of gene ontology biological process terms that were significantly enriched in our set of 254 transcripts found to be significantly regulated following AP-2γ silencing in MCF-7 cells. Shown here is a sub-set of 29 representative significant GO annotations shown in bold type with their associated hierarchy indicated by arrows and indenting. Analysis was performed using the GOstat software (GOstat, *url*; Beissbarth & Speed 2004) which found GO terms that were statistically over-represented within the 254 genes regulated following AP-2γ silencing in MCF-7 cells compared with the all of the genes represented on the Affymetrix HG-U133Plus2 microarray. The Group Count represents the number of genes from the AP-2γ dataset represented within each GO term and the Total Count represents the number of genes from the entire Affymetrix HG-U133Plus2 microarray represented within each GO group. A statistical analysis of the functional annotations overrepresented in the AP-2γ dataset with a Benjamini correction for multiple comparisons was used to calculate P-Value for each GO term. A P-Value cut off of less than 0.05 was used to filter insignificant GO terms. A more detailed description of the GOstat software can be found in the Materials and Methods 2.7.3.

Significant GO Biological Process Hits: Down Regulated	Group Count	Total Count	P Value
cellular process → cellular physiological process → **cell cycle** (GO:0007049)	41	667	5.47E-48
↳**regulation of cell cycle** (GO:0051726)	19	438	2.11E-07
↳**regulation of progression through cell cycle** (GO:0000074)	19	435	2.11E-07
↳**mitotic cell cycle** (GO:0000278)	28	192	1.07E-24
↳**mitotic spindle organization and biogenesis** (GO:0007052)	4	11	3.13E-05
↳**M phase of mitotic cell cycle** (GO:0000087)	25	145	5.01E-24
↳**mitosis** (GO:0007067)	25	143	4.63E-24
cellular process → cellular physiological process → cellular metabolism → nucleobase, nucleoside, nucleotide and nucleic acid metabolism → **DNA metabolism** (GO:0006259)	18	532	1.32E-05
↳**DNA replication** (GO:0006260)	12	142	1.50E-07
↳**DNA-dependent DNA replication** (GO:0006261)	7	70	4.05E-05
↳**DNA repair** (GO:0006281)	9	191	0.000577
cellular process → cellular physiological process → cell organization and biogenesis → **organelle organization and biogenesis** (GO:0006996)	21	750	3.40E-07
↳**cytoskeleton organization and biogenesis** (GO:0007010)	16	322	4.55E-07
↳**microtubule-based process** (GO:0007017)	13	130	4.37E-09
↳**microtubule-based movement** (GO:0007018)	8	76	6.49E-06
↳**microtubule cytoskeleton organization and biogenesis** (GO:0000226)	7	51	6.26E-06
↳**spindle organization and biogenesis** (GO:0007051)	6	18	2.11E-07
cellular process → cellular physiological process → **chromosome segregation** (GO:0007059)	6	34	8.83E-06
↳**sister chromatid segregation** (GO:0000819)	4	24	0.000662
cellular process → cellular physiological process → **cell division** (GO:0051301)	24	162	1.59E-21
↳**cytokinesis** (GO:0000910)	4	24	0.000662

Significant GO Biological Process Hits: Up-regulated	Group Count	Total Count	P Value
development → **organ development** (GO:0048513)	8	508	0.0128
development → **morphogenesis** (GO:0009653)	8	546	0.0131
↳**organ morphogenesis** (GO:0009887)	5	207	0.0161
↳cellular morphogenesis → **regulation of cell size** (GO:0008361)	4	131	0.0181
↳**cell growth** (GO:0016049)	4	131	0.0181
response to stimulus → **response to external stimulus** (GO:0009605)	6	459	0.039
↳**response to wounding** (GO:0009611)	6	354	0.0181
cellular process → cellular physiological process → **cell death** (GO:0008219)	7	551	0.0198
↳**apoptosis** (GO:0006915)	7	526	0.0183

4.6.3. PANTHER Classification System Analysis.

The PANTHER classification system is very similar to the Gene Ontology system, based on interconnecting hierarchies of terms, but greatly abbreviated and simplified (Mi *et al,* 2005; PANTHER Classification System *(url)*. This facilitates automated high-throughput analyses and accessible graphical outputs. For the results presented in this thesis, each Affymetrix ID from the AP-2γ data set was assigned a NCBI RefSeq ID extracted from the January 2007 release of the HG-U133_Plus2 Affymetrix NetAffx Annotation files (Liu et al, 2003; Affymetrix, *url*). After accounting for a number of Affymetrix Ids that had no associated NCBI RefSeq ID and filtering to remove multiple representations of individual RefSeq Ids, the AP-2γ data set were separated into 59 up-regulated and 176 down-regulated genes, then sorted using the PANTHER Classification System interface in to groups based on their assigned PANTHER Biological Process terms. The pie chart shown in Figure 4.10 represents the fraction of genes hits within functional categories against the total gene hits across all categories.

In order to establish with statistical confidence which PANTHER Biological Process annotations were over-represented or under-represented in our lists of significantly up-regulated and significantly down-regulated transcripts, we compared them to all of the PANTHER Biological Process represented by all of the 25431 genes within the NCBI Refseq Database using the PANTHER Classification System interface (Mi *et al,* 2005; PANTHER Classification System *(url)*. Similarly to the GOstat program described above, this allows an objective and statistically controlled measure of enriched GO terms from with the data set. Table 4.6 shows PANTHER Biological Process terms that are significantly enriched in our AP-2γ data set compared to terms represented within the NCBI Refseq Database. It is clear for those genes down-regulated following AP-2γ silencing PANTHER Biological Process terms are highly significantly over-represented in categories related to cell cycle, consistent with the data obtained from the GOstat analysis (Table 4.5) and also reflected in the PANTHER Biological Process pie chart (Figure 4.10). Additionally, for genes up-regulated following AP-2γ silencing, PANTHER Biological Process terms were significantly over-represented in categories including Signal transduction, Cell proliferation and differentiation and Growth factor homeostasis.

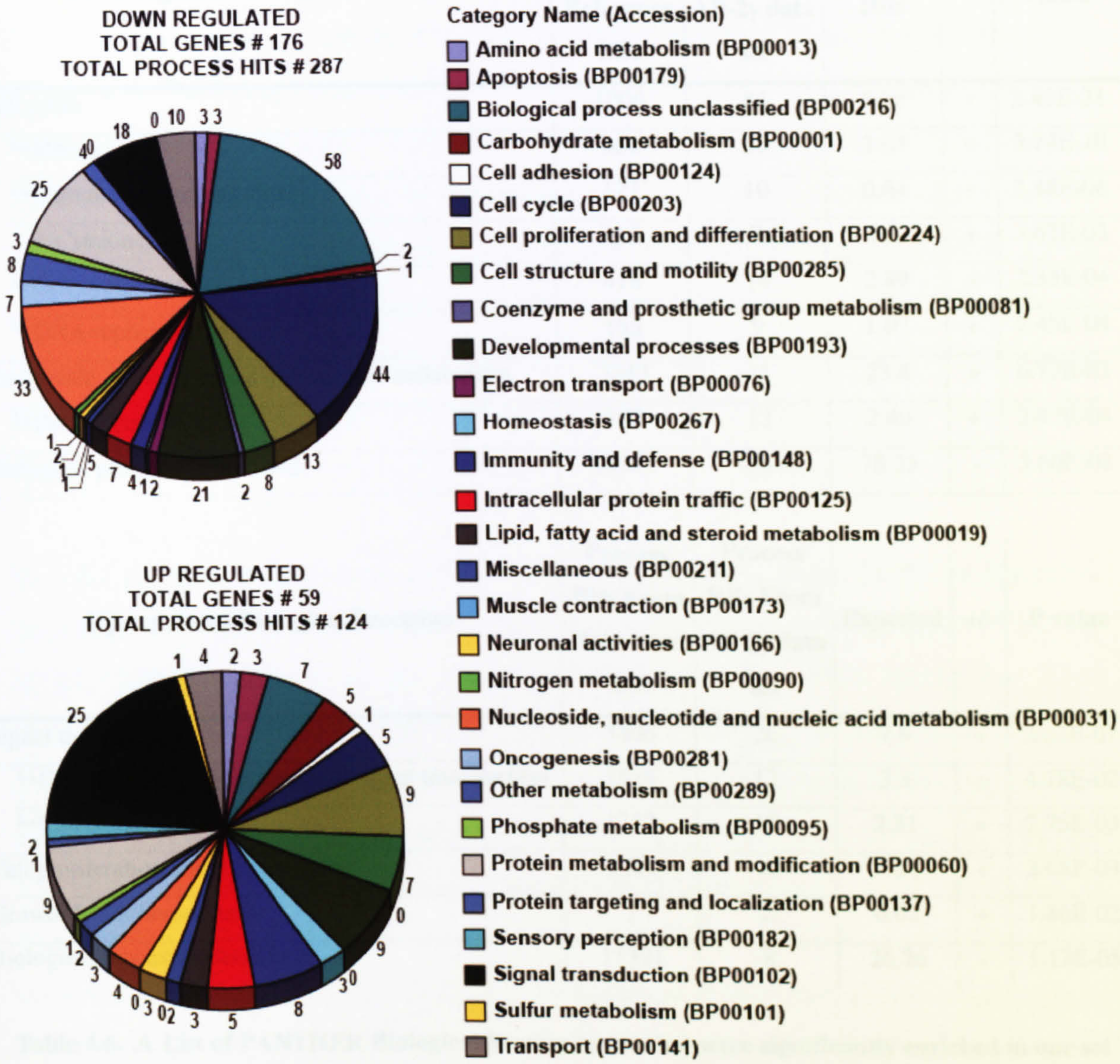


Figure 4.10. A PANTHER functional classification of genes that are significantly down-regulated or up-regulated ($p < 0.01$) following AP-2γ silencing in MCF-7 cells. Pie charts represent the fraction of genes hits within functional categories, as classified by their associated PANTHER Biological Process annotations, against the total gene hits. This pie chart represents expression changes for single transcripts based on their associated gene symbol (not probe sets) this is to avoid overrepresentation of functional classes where multiple probe sets representing one gene are significantly regulated. Gene ontology terms for each probe set analysed using the PANTHER Classification System (Mi *et al,* 2005; PANTHER Classification System, *url*).

Down-regulated Biological Processes	Process Hits From Reference List	Process Hits From AP-2γ data set	Expected Hits	+/-	P value
Cell cycle	1009	44	6.98	+	2.43E-21
↳Mitosis	382	20	2.64	+	5.74E-10
↳Chromosome segregation	121	10	0.84	+	3.48E-06
↳Cytokinesis	115	7	0.8	+	3.62E-03
↳Cell cycle control	418	14	2.89	+	2.33E-04
↳DNA replication	155	9	1.07	+	2.45E-04
Nucleoside, nucleotide and nucleic acid metabolism	3343	41	23.4	+	6.77E-03
↳DNA metabolism	360	13	2.49	+	2.47E-04
Biological process unclassified	11321	58	78.35	-	3.60E-02

Up-regulated Biological Processes	Process Hits From Reference List	Process Hits From AP-2γ data set	Expected	+/-	P value
Signal transduction	3406	26	7.9	+	2.61E-07
↳Cell surface receptor mediated signal transduction	1638	12	3.8	+	4.78E-02
↳Cell communication	1213	12	2.81	+	2.75E-03
Cell proliferation and differentiation	1028	10	2.38	+	3.68E-03
Growth factor homeostasis	7	2	0.02	+	1.86E-02
Biological process unclassified	11321	8	26.26	-	1.17E-05

Table 4.6. A List of PANTHER Biological Process terms that were significantly enriched in our set of 254 transcripts found to be significantly regulated following AP-2γ silencing in MCF-7 cells. Shown here is a sub-set of 14 significant PANTHER Biological Process annotations within their associated hierarchy indicated by indenting. Analysis was performed using the PANTHER Classification System (Mi *et al,* 2005; PANTHER Classification System, *url*). Pair wise analyses between the AP-2γ dataset and genes within the NCBI Refseq Database (Reference List) were performed, using the binomial distribution function with a Benjamini correction for multiple comparisons, in order to identify genes with PANTHER Biological Process annotations statistically overrepresented or underrepresented within the AP-2γ dataset. A P-Value cut off of less than 0.05 was used to filter insignificant terms. A more detailed description of the PANTHER Classification System can be found in the Materials and Methods 2.7.3.

4.7. Analysis of Cell Cycle Ontology Related Gene Expression Changes

The data presented in Section 4.6 clearly illustrates that a large proportion of the significantly down-regulated genes are cell cycle related and also that these cell cycle related transcripts are significantly enriched in the AP-2γ data set compared to the genome as a whole. In order to understand how down regulation of this cell cycle subset might be related to the observed failure to generate long-term stable AP-2γ knockdowns (Table 3.2) these genes were interrogated further. This analysis was restricted to the genes assigned the cell cycle GO terms. An additional miscellaneous group included genes assigned the Regulation of Cell Proliferation /// Growth GO terms, some unclassified genes and AP-2γ itself.

In order to understand at what point in the cell cycle these genes might be regulated, each gene was compared to experimental data from Whitfield and colleagues (Whitfield *et al.*, 2002; Whitfield *et al.*, supplemental data *url*). This microarray study characterised the periodicity of gene expression according to cell cycle phase in chemically synchronised HeLa cells. Figure 4.11 shows the results of this comparison. Interestingly, a large number of genes within this data subset were shown by Whitfield and colleagues to display cell cycle periodicity of expression. It is reasonable to consider that many of the gene expression changes in this AP-2γ data subset may have occurred as a consequence of a cell cycle perturbation as their expression is known to be intrinsic to the cell cycle. For the subset of genes that are known to show cell cycle periodicity the cell cycle phase or transition at which the expression was highest and lowest was also annotated in Figure 4.11 (extracted from Whitfield *et al.*, 2002). With the exception of genes within the DNA Replication /// DNA Repair GO group, all the genes displaying cell cycle periodicity were known to show maximal expression at G2 or G2/M and minimal expression at G1/S. All but one of these genes known to show cell cycle periodicity were down-regulated in response to AP-2γ silencing in MCF-7 cells and it could be postulated that their low expression in AP-2γ silenced cells hints at a cell cycle perturbation at the G1/S transition. However, genes within the DNA Replication /// DNA Repair GO group displaying cell cycle periodicity were known to show maximal expression at G1/S or S phase and minimal expression at G2 or G2/M.

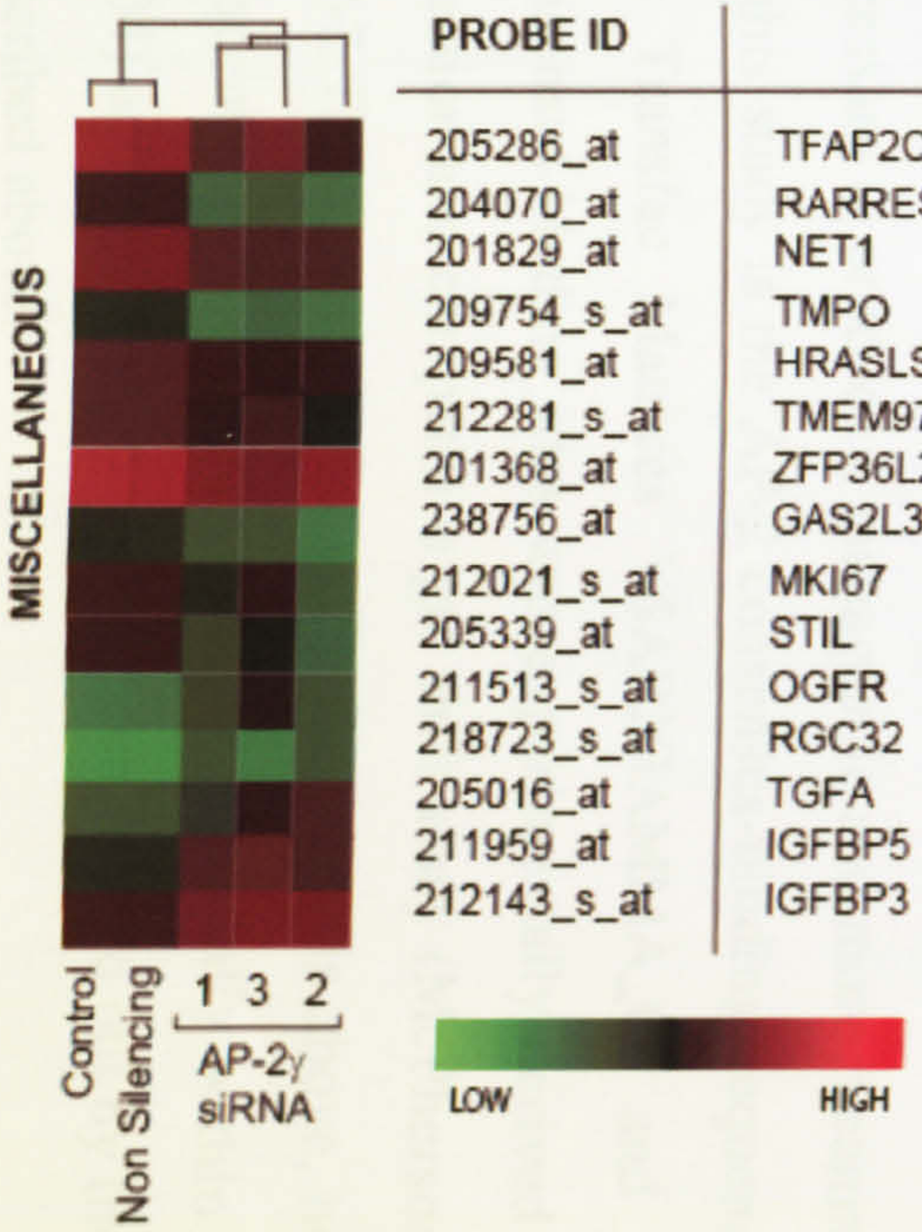
These were also down-regulated in response to AP-2γ silencing in MCF-7 cells and it could be postulated that their low expression in AP-2γ silenced cells hints that the cell cycle perturbation occurred before the cells enter S-phase where expression would normally be expected to be high. Also consistent with this hypothesis is that a large proportion of the genes listed in Figure 4.11 have been shown to be transcriptionally regulated by E2F family members at the G1/S transition (Ishida *et al.*, 2001; Polager *et al.*, 2002; Ren *et al.*, 2002) These genes are annotated in the final column of Figure 4.11.

PROBE ID		GENE DESCRIPTION	FOLD CHANGE	P VALUE	PERIODICY OF CELL CYCLE EXPRESSION LOW	HIGH	E2F TARGET GENE
203213_at	CDC2	cell division cycle 2, G1 to S and G2 to M	-2.25	0.0085	G1/S	G2	yes[1, 2, 3]
209714_s_at	CDKN3	cyclin-dependent kinase inhibitor 3	-2.25	0.000505	G1/S	G2/M	-
228729_at	CCNB1	cyclin B1	-2.10	0.000143	G1/S	G2/M	yes[2, 3]
202870_s_at	CDC20	cell division cycle 20 homolog (S. cerevisiae)	-2.04	0.000125	G1/S	G2	yes[2]
201897_s_at	CKS1B	CDC28 protein kinase regulatory subunit 1B	-2.00	0.000532	-	-	yes[1, 3]
203418_at	CCNA2	cyclin A2	-1.95	0.00149	G1/S	G2	yes[1, 2, 3]
202705_at	CCNB2	cyclin B2	-1.93	0.00117	G1/S	G2/M	yes[2, 3]
208712_at	CCND1	cyclin D1	-1.77	0.00164	S	G2/M	-
202095_s_at	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	-1.54	0.00552	G1/S	G2/M	yes[3]
228033_at	E2F7	E2F transcription factor 7	-1.49	0.00705	-	-	-
1554768_a_at	MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	-2.24	0.00487	G1/S	G2/M	yes[1, 3]
201456_s_at	BUB3	budding uninhibited by benzimidazoles 3 homolog	-1.90	0.000228	G1/S	G2	yes[1, 3]
204822_at	TTK	TTK protein kinase	-1.88	0.0029	G1/S	G2/M	yes[1]
205393_s_at	CHEK1	CHK1 checkpoint homolog (S. pombe)	-1.73	0.00781	no	no	yes[1, 3]
201292_at	TOP2A	topoisomerase (DNA) II alpha 170kDa	-2.28	0.000228	G1/S	G2/M	yes[1, 2, 3]
204244_s_at	DBF4	DBF4 homolog (S. cerevisiae)	-2.08	1.09E-05	-	-	yes[1, 2, 3]
204128_s_at	RFC3	replication factor C (activator 1) 3, 38kDa	-1.95	0.00156	no	no	yes[1, 2, 3]
204146_at	RAD51AP1	RAD51 associated protein 1	-1.92	0.00601	-	-	yes[3]
210983_s_at	MCM7	minichromosome maintenance deficient 7	-1.73	0.00164	no	no	yes[2, 3]
216237_s_at	MCM5	minichromosome maintenance deficient 5	-1.69	0.00487	G2/M	G1/S	yes[1, 3]
1554696_s_at	TYMS	thymidylate synthetase	-1.68	0.0036	G2/M	S	yes[2]
203967_at	CDC6	CDC6 cell division cycle 6 homolog (S. cerevisiae)	-1.67	0.00296	G2/M	S	yes[1, 3]
205909_at	POLE2	polymerase (DNA directed), epsilon 2 (p59 subunit)	-1.65	0.00433	no	no	-
225655_at	UHRF1	ubiquitin-like, PHD and RING finger domains 1	-1.63	0.00847	G2/M	G1/S	-
242560_at	FANCD2	Fanconi anemia, complementation group D2	-1.61	0.00635	G2/M	G1/S	yes[3]
203210_s_at	RFC5	replication factor C (activator 1) 5, 36.5kDa	-1.61	0.0036	-	-	yes[3]
205053_at	PRIM1	primase, polypeptide 1, 49kDa	-1.60	0.00826	G2/M	S	yes[1]
203022_at	RNASEH2A	ribonuclease H2, subunit A	-1.59	0.00635	-	-	-
204510_at	CDC7	CDC7 cell division cycle 7 (S. cerevisiae)	-1.55	0.00786	G2/M	S	yes[3]
210028_s_at	ORC3L	origin recognition complex, subunit 3-like (yeast)	-1.53	0.00504	G2/M	S	yes[3]
202413_s_at	USP1	ubiquitin specific peptidase 1	-1.52	0.00691	M/G1	S	yes[3]

Figure 511.

PROBE ID	GENE DESCRIPTION		FOLD CHANGE	P VALUE	PERIODICY OF CELL CYCLE EXPRESSION LOW	PERIODICY OF CELL CYCLE EXPRESSION HIGH	E2F TARGET GENE
218663_at	HCAP-G	chromosome condensation protein G	-2.11	0.00843	G1/S	G2	yes [3]
203554_x_at	PTTG1	pituitary tumor-transforming 1	-2.02	0.000254	G1/S	G2	yes [1, 3]
223381_at	CDCA1	cell division cycle associated 1	-2.00	0.00413	G1/S	G2	-
201664_at	SMC4	structural maintenance of chromosomes 4	-1.96	0.00283	G1/S	G2	yes [1, 3]
224734_at	HMGB1	high-mobility group box 1	-1.95	7.62E-05	-	-	yes [1]
208808_s_at	HMGB2	high-mobility group box 2	-1.93	0.0029	-	-	yes [2]
204962_s_at	CENPA	centromere protein A	-1.89	0.00177	G1/S	G2	yes [3]
209715_at	CBX5	chromobox homolog 5	-1.87	0.000638	no	no	yes [1, 3]
213253_at	SMC2L1	structural maintenance of chromosomes 2-like 1	-1.87	0.000158	G1/S	G2	yes [1, 3]
209172_s_at	CENPF	centromere protein F, 350/400ka (mitosin)	-1.72	0.00177	G1/S	G2	yes [3]
230165_at	SGOL2	shugoshin-like 2 (S. pombe)	-1.69	0.000594	G1/S	G2	-
219978_s_at	NUSAP1	nucleolar and spindle associated protein 1	-1.66	0.00836	-	-	-
224753_at	CDCA5	cell division cycle associated 5	-1.61	0.00953	-	-	-
219148_at	PBK	PDZ binding kinase	-2.43	0.00635	G1/S	G2/M	-
218755_at	KIF20A	kinesin family member 20A	-2.26	0.000187	-	-	yes [3]
204444_at	KIF11	kinesin family member 11	-2.17	0.00329	-	-	-
203764_at	DLG7	discs, large homolog 7 (Drosophila)	-2.16	0.00245	G1S	G2	-
204641_at	NEK2	NIMA (never in mitosis gene a)-related kinase 2	-2.14	0.000505	G1/S	G2/M	yes [1, 3]
218542_at	CEP55	centrosomal protein 55kDa	-2.03	0.000125	G1/S	G2/M	-
218009_s_at	PRC1	protein regulator of cytokinesis 1	-1.97	0.00775	G1/S	G2/M	yes [1, 3]
206364_at	KIF14	kinesin family member 14	-1.94	0.000574	-	-	-
221258_s_at	KIF18A	kinesin family member 18A	-1.87	0.000158	-	-	yes [3]
202183_s_at	KIF22	kinesin family member 22	-1.86	0.000198	G1/S	G2	yes [3]
204092_s_at	AURKA	aurora kinase A	-1.85	0.00572	G1/S	G2	-
204709_s_at	KIF23	kinesin family member 23	-1.84	0.000113	-	-	-
210052_s_at	TPX2	microtubule-associated, homolog (Xenopus laevis)	-1.82	0.00152	-	-	-
202954_at	PAK3	p21 (CDKN1A)-activated kinase 3	-1.77	0.00344	no	no	-
225297_at	CCDC5	coiled-coil domain containing 5	-1.71	0.00131	-	-	-
222077_s_at	RACGAP1	Rac GTPase activating protein 1	-1.71	0.00489	-	-	-
1552619_a_at	ANLN	anillin, actin binding protein	-1.71	0.0029	G1/S	G2	yes [1, 3]
204886_at	PLK4	polo-like kinase 4 (Drosophila)	-1.66	0.000158	no	no	-
232238_at	ASPM	asp-like, microcephaly associated	-1.63	0.00817	-	-	-
209408_at	KIF2C	kinesin family member 2C	-1.58	0.00226	-	-	yes [3]
209464_at	AURKB	aurora kinase B	-1.52	0.00684	-	-	-
201041_s_at	DUSP1	dual specificity phosphatase 1	1.54	0.00601	no	no	-
204141_at	TUBB2A	tubulin, beta 2A	2.23	0.000358	-	-	yes [3]

Figure 511.



PROBE ID	GENE DESCRIPTION		FOLD CHANGE	P VALUE	PERIODICY OF CELL CYCLE EXPRESSION LOW	PERIODICY OF CELL CYCLE EXPRESSION HIGH	E2F TARGET GENE
205286_at	TFAP2C	transcription factor AP-2 gamma	-2.39	0.00646	no	no	-
204070_at	RARRES3	retinoic acid receptor responder 3	-2.06	3.76E-05	no	no	-
201829_at	NET1	neuroepithelial cell transforming gene 1	-1.79	0.000106	no	no	-
209754_s_at	TMPO	thymopoietin	-1.78	3.76E-05	G1/S	G2/M	yes[1]
209581_at	HRASLS3	HRAS-like suppressor 3	-1.77	0.00112	no	no	-
212281_s_at	TMEM97	transmembrane protein 97	-1.74	0.0012	no	no	-
201368_at	ZFP36L2	zinc finger protein 36, C3H type-like 2	-1.66	0.00222	-	-	-
238756_at	GAS2L3	Growth arrest-specific 2 like 3	-1.62	0.00193	G1/S	G2/M	-
212021_s_at	MKI67	antigen identified by monoclonal antibody Ki-67	-1.61	0.00723	G1/S	G2/M	yes[1]
205339_at	STIL	SCL/TAL1 interrupting locus	-1.60	0.00552	-	-	-
211513_s_at	OGFR	opioid growth factor receptor	1.64	0.00113	no	no	-
218723_s_at	RGC32	response gene to complement 32	1.75	0.00378	no	no	-
205016_at	TGFA	transforming growth factor, alpha	1.78	0.0074	no	no	-
211959_at	IGFBP5	insulin-like growth factor binding protein 5	2.18	0.00817	-	-	-
212143_s_at	IGFBP3	insulin-like growth factor binding protein 3	2.30	5.38E-05	-	-	-

Figure 4.11. Genes from the AP-2γ data set assigned to the cell cycle GO groupings are known to display cell cycle periodicity and are known E2F target genes. Changes observed in probe sets between the *reference* (non silencing control siRNA and cells treated with transfection reagent only) and *test* (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) groups at the indicated False Discovery Rate corrected p values. A full list of significantly regulated probe sets together with normalised expression values, fold change values (between *test* and *reference* groups) and FDR corrected p values is contained in Appendix 1. The same cell cycle GO groups as those used in Figure 4.9 were used to organise the data and a comparison with the experimental data from Whitfield and colleagues (Whitfield *et al.,* 2002; Whitfield *et al.,* supplemental data, *url*) was used to annotate the cell cycle periodicity. Genes were also annotated as known E2F targets based on findings from the literature (Ren *et al.,* 2002[1]; Polager *et al.,* 2002[2]; Cam *et al.,* 2004[3]). A dash indicates where not data was present.

4.8. Analysis of Associated Transcription Factor Binding Sites

In order to explore whether any common regulatory elements existed within the list of significantly changed genes ($p<0.01$) observed following silencing of AP-2γ in MCF-7 cells, transcription factor binding site (TFBS) analysis were carried out. An initial analysis took advantage of data contained within the UCSC Genome Browser (UCSC Genome Browser, *url*; Kent *el al.*, 2002), detailing transcription factor binding sites conserved between the human, mouse and rat genome alignments. Briefly, this consists of transcription factor binding matrices from the Transfac Matrix Database (Transfac Matrix Database (v7.0), *url*; Matys *et al.*, 2003), considered to be conserved across the alignment only if the matrix similarity score meets a given threshold score for its binding matrix in all three species. Using this pre-calculated data, TFBS were searched for in the 2000 base pair region upstream of the start of transcription for transcripts associated with an Affymetrix ID. Table 4.7 lists the most represented TFBS observed within our list with their associated genes. The most represented TFBS are those for cell cycle related transcription factors, in particular binding sites for the E2F family are most represented upstream of transcripts within the AP-2γ dataset. It should be noted that this TFBS analysis was very stringent. Firstly, it was looking for conservation between human, mouse and rat genomes, which may not be biologically appropriate for many of the genes represented within the human genome. Secondly, it includes a statistical comparison that scores each TFBS on the probability ($P<0.01$) of the sequence occurring in the upstream regions of all NCBI RefSeq genes, this will filter out any commonly occurring elements. However, these constraints were considered necessary to produce a manageable and more meaningful dataset. Of particular interest to this study is the AP-2 consensus-binding sequence, SCCNNVRGB, represented by the Transfac Matrices V$AP2GAMMA_01 and V$AP2ALPHA_01. These AP-2 consensus binding sites was experimentally derived using a PCR-assisted binding site selection and competitive gel shift assay (McPherson & Weigel, 1999). No conserved AP-2 TFBS were found using the analysis above, however this sequence occurs fairly frequently in the human genome especially within promoter sequences (Bajic *et al.*, 2004) and will therefore have been filtered out by the significance threshold constraints described above.

TFBS	Transfac Matrix	Description	Number of Sites	2 kb upstream region of genes containing TFBS
E2F	V$E2F_02	E2F family	14	CDC2, CDC2, CDC6, RCF3, RCF3, CHEK1, CHEK1, POLE2, MCM7, BICD2, KIF15, KIF18A, UHRF, C16orf75
FOXD1	V$FREAC4_01	Forkhead box D1	11	VRK1, MCM7, SLC1A1, ID2, PHLDA1, SPDEF, PREX1, PREX1, OTUD1, UNKNOWN, ZNF367
FOXO1, FOXO3	V$FOXO1_01,V$FOXO3_01, V$FOXO4_01	Forkhead box protein O1 and O3	10	BUB3, BUB3, MBNL2, MCM7, ID2, KRT86, PHLDA1, KIF20A, SYNP02, SGOL2
OCT1	V$OCT1_02	POU domain, class 2, transcription factor 1	9	CDC20, MBNL2, CDC6, HMGB2, ID4, TMPO, STX12, ID2, ZNF367
NKX-2.5	V$NKX25_01	NK2 transcription factor related	9	PDZK1, KIAA0746, UHRF1, INHBA, ELOVL7, UNKNOWN, UNKNOWN, ZNF367, TMEM64
PAX5	V$PAX5_02	paired box gene 5	8	C13ORF, dlx1, usp1, igsf1, arf1, fbxl11 kiaa0746, vav3
CREBP1 : cJUN	V$CREBP1CJUN_01	CRE-binding protein 1:c-Jun heterodimer	8	CCNA2, TUBB2A, GEM, PLK4, CCND1, RIN, KIF18A, C6orf173
PPARγ	V$PPARG_01	peroxisome proliferative activated receptor, gamma	8	ME1, GPSM2, TFAP2C, KIF20A, ARF6, C13orf3, RHPN2, SARS
FOXA1	V$HFH1_01	forkhead box A1	7	DUSP1, BUB3, BUB3, ID2, TMPO, E2F7, UNKNOWN
POU3F2	V$POU3F2_01, V$BRN2_01	POU domain, class 3, transcription factor 2	7	DUSP1, MBNL2, TFAP2C, TFAP2C, BUB3, BUB3, ARF6
GATA-1	V$GATA1_05	GATA binding protein 1 (globin transcription factor 1)	7	KIF20A, KIF15, UNKNOWN, UNKNOWN, TKK, HMGB2, SLC1A1
SREBP1	V$SREBP1_01	sterol regulatory element-binding protein 1	6	DUSP1, SMC4, SLC31A2, HRASLS3, DEPDC1, MCM10
MYOG	V$COMP1_01	myogenin (myogenic factor 4)	6	CYB5A, ID2, CYB5A, UNKNOWN, C9orf152, FBXL11
CUTL1	V$CDPCR1_01, V$CDP_01, V$CDPCR3HD_01	cut-like 1, CCAAT displacement protein	6	BAMBI, GEM, ARF6, INHBA, DKK1, CCDC5
CEBP	V$CEBP_01, V$CHOP_01	CCAAT/enhancer binding protein family	6	TMPO, PDK4, INHBA, UNKNOWN, KIF14, MBNL2
YY1	V$YY1_01	YY1 transcription factor	5	HMGB2, FBXL11, KIF20A, UNKNOWN, EGR3
SRF	V$SRF_01	Serum response factor	5	ZNF267, ZNF367, C9orf152, TUFT1, EGR3
ARNT	V$ARNT_02	Hypoxia-inducible factor 1 beta	5	FBXL11, TMPO, IGFBP3, IGFBP3, KRT86
STAT5	V$STAT5A, V$STAT5B	signal transducer and activator of transcription 5A and B	4	WARS, TFAP2C, TFAP2C, PDK4
HOXA3	V$HOXA3_01	homeobox A3	4	TOP2A, TOP2A, INHBA, INHBA
E4BP4	V$E4BP4_01	nuclear factor, interleukin 3 regulated	4	TOP2A, GPSM2, SYNPO2, ERRFI1
TFCP2	V$CP2_01	transcription factor CP2	4	USP1, TUFT1, CENPK, DLX1
RFX1	V$RFX1_02	X-box binding protein RFX1	4	SVIL, NEK2, NEK2, FHL1
TST1	V$TST1_01	POU-factor Tst-1/Oct-6	4	KIF2C, CXCL2, IGFBP5, HMGB1

MMIF	V$MIF1_01	macrophage migration inhibitory factor	4	NEK2, NEK2, SLC1A1, KIAA1324,
TGFBIF	V$TGIF_01	TGFB-induced factor	3	WARS, S100A7, OTUD1
MYC	V$NMYC_01, V$MYCMAX_03	avian myelocytomatosis viral oncogene homolog	3	ZFP35L2, HMMR, OTUD1
UNKNOWN	V$MZF1_02	UNKNOWN	3	ERRBFI1, SMC4, EGR4
SEF1	V$SEF1_C	SL3-3 enhancer factor	3	KRT86, UNKNOWN, NET1
P300	V$P300_01	P300	3	CKS1B, FBXL11, LOC652689
NRF2	V$NRF2_01	nuclear respiratory factor 2	3	CKS1B, AURKA, KIF14
PAX3	V$PAX3_01	paired box gene 3	3	INHBA, UNKNOWN, LOC652689
PAX4	V$PAX4_02	paired box gene 4	3	LOC87763, MBNL2, UNKNOWN
AP4	V$AP4_Q6	Activating enhancer-binding protein 4	3	ID2, FBL1, FBXL11
LHX3	V$LHX3_01	nescient helix loop helix 1	3	KIAA0746, CEP55, MBNL2
P53	V$P53_01	P53	3	VRK1, IGFBP3, IGFBP3
HEN1	V$HEN1_02	nescient helix loop helix 1	3	SLC35E1, C13orf3, CDC42EP5
ER	V$ER_Q6	Oestrogen Receptor α	3	IGSF1, TFAP2C, TFAP2C

Table 4.7. shows the most common conserved TFBS occurring in the 2 kb upstream region of each gene with significant expression change ($p<0.01$) observed following silencing of AP-2γ in MCF-7 cells. The analysis used pre-calculated data contained within the March 2006 assembly of the UCSC Human Genome Browser (UCSC Genome Browser, *url*; Kent *el al.*, 2002), detailing transcription factor binding sites conserved between the human, mouse and rat genome alignments. Briefly, this consists of transcription factor binding matrices from the Transfac Matrix Database (Transfac Matrix Database (v7.0), *url*; Matys *et al.*, 2003), considered to be conserved across the alignment only if the matrix similarity score meets a given threshold score for its binding matrix in all three species. Where each matrix is aligned it assigned a Z score that can be interpreted as the number of standard deviations above the mean raw score for that binding matrix across the upstream regions of all RefSeq genes, a Z score cut-off 2.33 was used which is equivalent to a $p<0.01$ value.

An additional, smaller scale analysis, focusing on the AP-2 consensus-binding site was conducted using the "HumanGenome9999" human promoter sequence data (9999bp up stream sequences of 21787 genes represented in the ENSEMBL database) imported into the Genespring package (Genespring GX 7.3, Agilent Technologies). A subset of 192 genes out of the 254 represented in the AP-2γ dataset was successfully mapped to the ENSEMBL promoter sequences. Searching for the AP-2 consensus binding sequence (SCCNNVRGB) in a –500 to +50 bp upstream region of these sequences showed that 156 sequences contained one or more matches (82.1% of the 192 genes). However, the SCCNNVRGB consensus also matched more than once in the –500 to +50 region upstream of 80.3% of the 21787 sequences contained in the ENSEMBL data set. The –100 to +50 and –200 to +50 regions were also interrogated, in order explore the effect of refining the promoter region analysed on the occurrence of the SCCNNVRGB consensus. The data from this analysis is summarised in Table 4.8 and the entire dataset is listed in Appendix 2. Overall, this analysis shows that there is no enrichment for the presence of AP-2 binding sites in promoter regions of genes within AP-2γ data set. The analysis also highlights how common the AP-2 consensus sequence is in upstream sequences of genes, complementing the data from Bajic and colleagues (Bajic *et al.*, 2004). It is likely therefore, that constraints other than DNA sequence alone must confer AP-2 DNA binding specificity. This analysis is unfortunately insufficient for distinguishing between direct and indirect AP-2γ targets within this data set.

Region Interrogated	AP-2γ Data set (192 genes)	ENSEMBL promoter data set (21787 genes)
-500 to + 50	82	80
-200 to + 50	64	65
-100 to + 50	48	49

Table 4.8. Percentage calls of one or more SCCNNVRGB consensus sequences in the indicated genomic region around the +1 start of transcription of genes within the AP-2γ data set. Sequences were analysed using the "HumanGenome9999" human promoter sequence data (9999bp up stream sequences of 21787 genes represented in the ENSEMBL database) imported into the Genespring package (Genespring GX 7.3, Agilent Technologies).

4.9. Validation of Selected Affymetrix gene expression changes

4.9.1. qPCR Validation

In order to assess the reliability of the microarray gene expression profile generated following AP-2γ silencing in MCF-7 cells, a selection of genes were carried forward for further analysis. Genes were chosen based on the magnitude and statistical significance of their microarray-derived gene expression values and their biological interest with respect to AP-2 and breast cancer. The expression changes for the majority of these genes were included in the top thirty most significant changes in gene expression following AP-2γ silencing displayed in Figure 4.8. Figure 4.12 shows the qPCR validation of selected transcripts following AP-2γ silencing with three independent AP-2γ targeting siRNAs compared to controls assessed 96 hours following transfection. In an effort to examine the concordance between the qPCR data in Figure 4.12 and the primary data generated in the microarray experiment (Appendix 1), the two sets of values were compared graphically, as shown in Figure 4.13. These data fitted well to a line of best fit indicated by an R^2 value of 0.86 and the plot illustrates a good correlation between microarray and qPCR derived gene expression (fold change) data. In general, the fold change was higher for the qPCR derived values compared to the microarray data, indicated by a gradient of 0.81. This level of correlation is consistent with other studies looking at qPCR and Affymetrix microarray gene expression data (Dallas *et al.*, 2005).

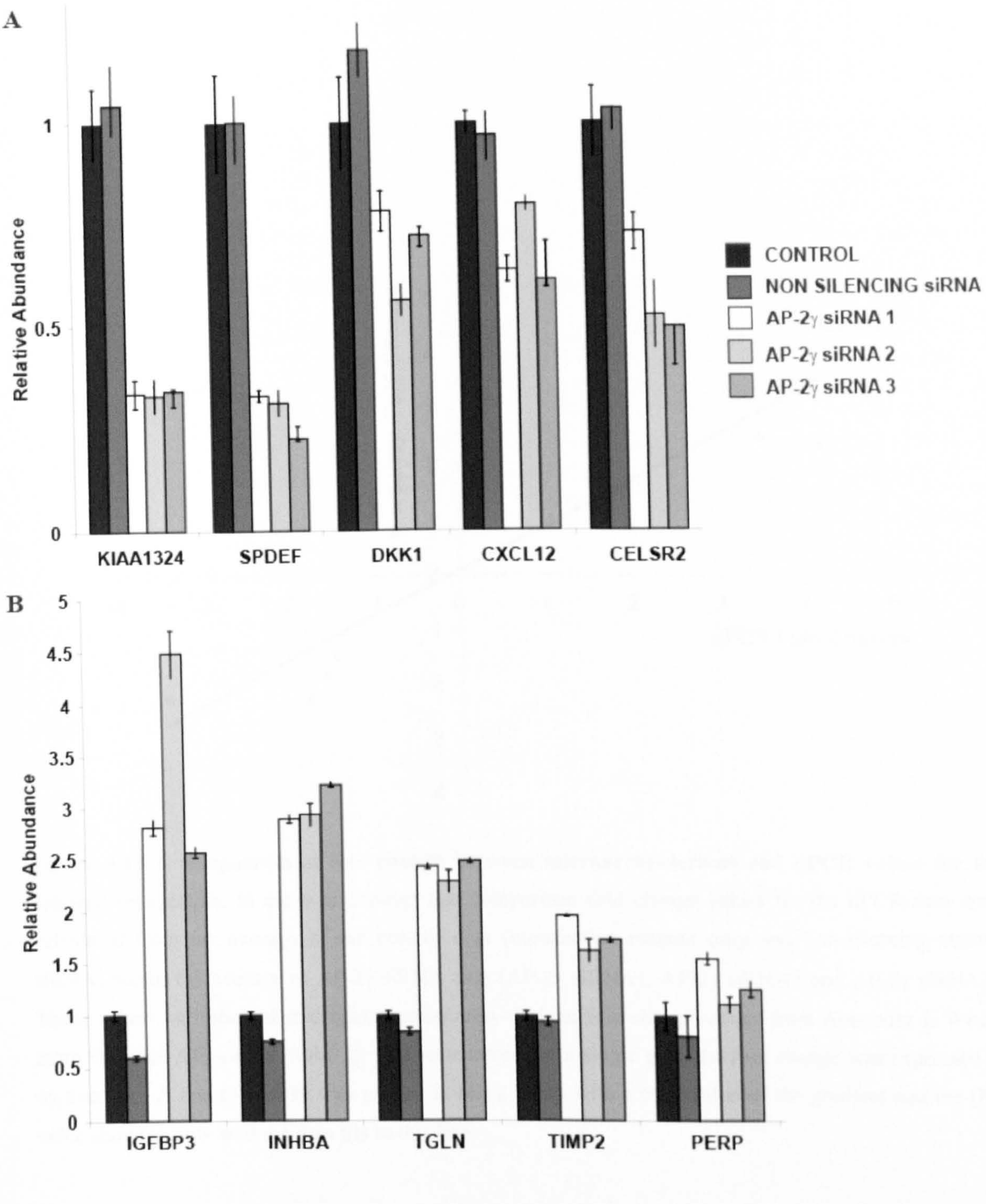


Figure 4.12. Validation of selected microarray gene expression changes by quantitative PCR analysis. MCF-7 cells were transfected as indicated and total RNA harvested 96 hours later. Quantitative Real Time PCR analysis of the indicated mRNA levels was performed. Details of the primer probe sets used for this analysis are shown in the Materials and Methods 2.8. Data were normalised to the *GAPDH* mRNA levels. The transfection reagent only value was set at 1 to generate relative mRNA levels for each transcript analysed. The graph represents data averaged from three PCR replicates ± standard error.

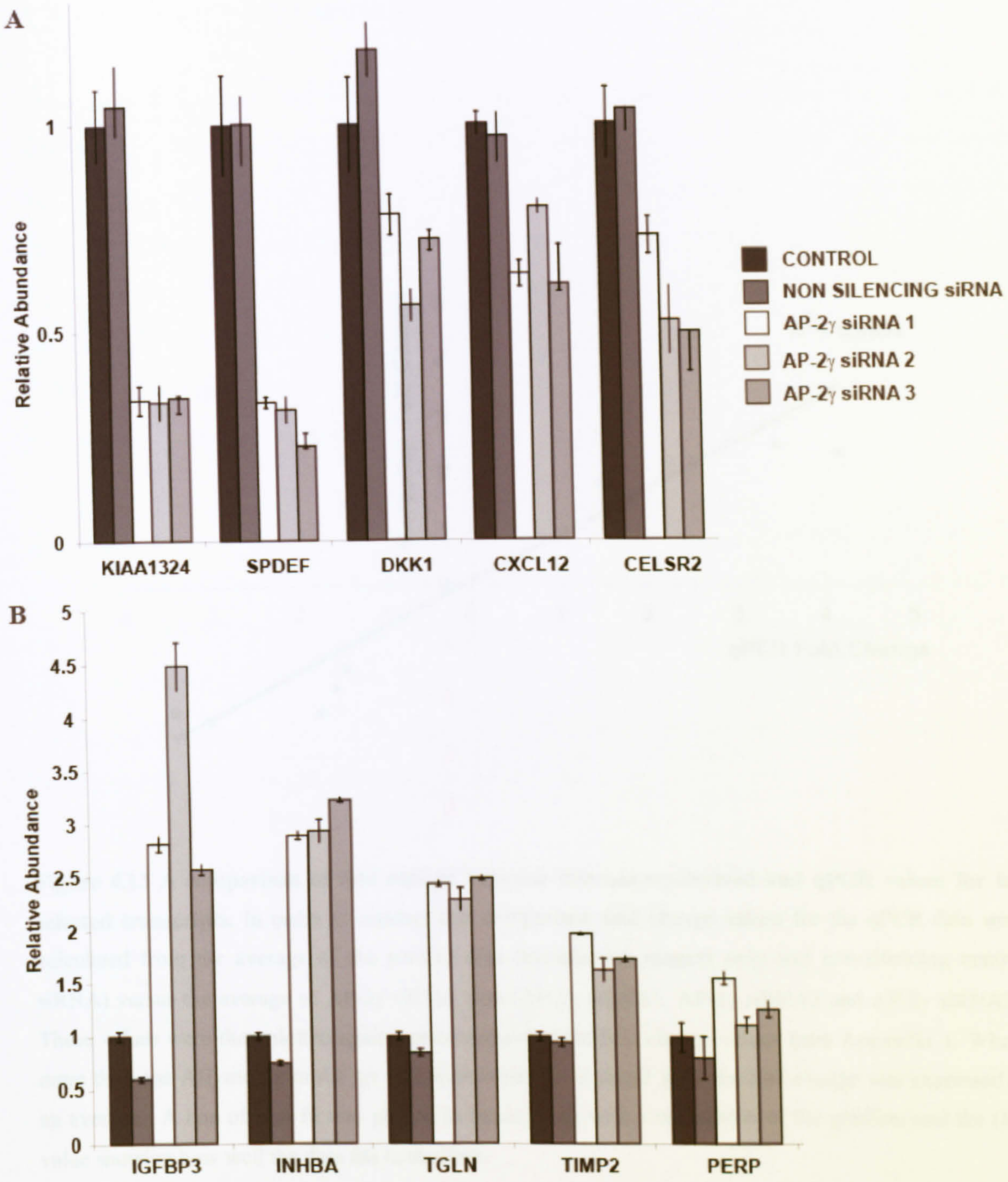


Figure 4.12. Validation of selected microarray gene expression changes by quantitative PCR analysis. MCF-7 cells were transfected as indicated and total RNA harvested 96 hours later. Quantitative Real Time PCR analysis of the indicated mRNA levels was performed. Details of the primer probe sets used for this analysis are shown in the Materials and Methods 2.8. Data were normalised to the *GAPDH* mRNA levels. The transfection reagent only value was set at 1 to generate relative mRNA levels for each transcript analysed. The graph represents data averaged from three PCR replicates ± standard error.

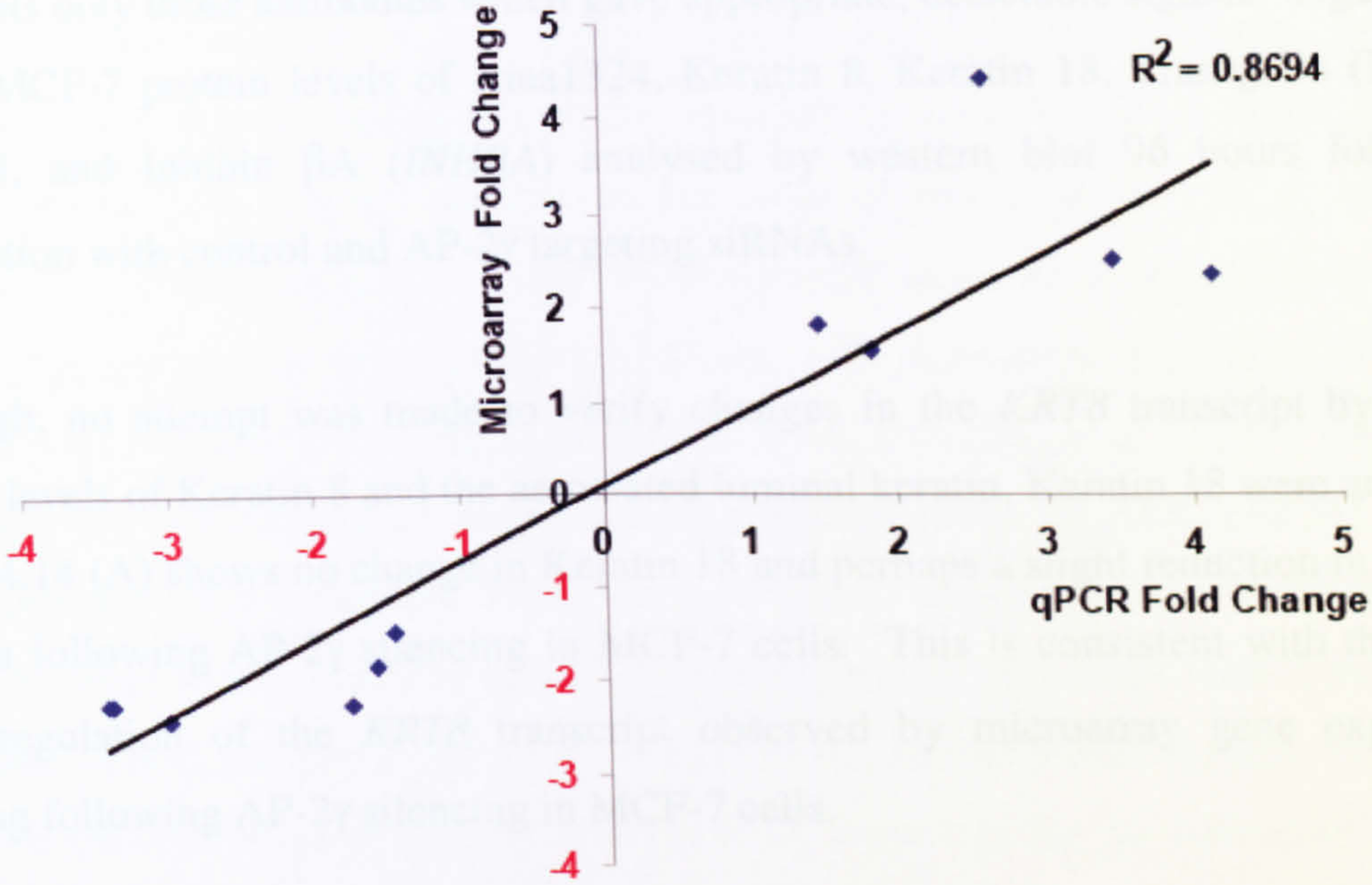


Figure 4.13 A comparison of fold change between microarray-derived and qPCR values for ten selected transcripts. In order to conduct this comparison fold change values for the qPCR data were calculated from the average of the control data (transfection reagent only and non-silencing control siRNA) versus the average of AP-2γ siRNA data (AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3). These values were then plotted against microarray-derived fold change values from Appendix 1. Where more than one Affymetrix probe set was represented for a single gene the fold change was expressed as an average. A line of best fit was plotted in black along with a calculation of the gradient and the (R^2) value showing how well the data fits to this line.

4.9.2. Western blot validation

In order to determine whether any of the microarray gene expression changes could be confirmed at the functional level of protein abundance, western blot analyses were conducted. The selection of genes for validation in this case was largely governed by the availability of suitable commercial antibodies; therefore the data presented represents only those antibodies which gave appropriate, detectable signals. Figure 4.14 shows MCF-7 protein levels of Kiaa1324, Keratin 8, Keratin 18, Transgelin (*TGLN*), IGFBP3, and Inhibin βA (*INHBA*) analysed by western blot 96 hours following transfection with control and AP-2γ targeting siRNAs.

Although, no attempt was made to verify changes in the *KRT8* transcript by qPCR, protein levels of Keratin 8 and the associated luminal keratin, Keratin 18 were analysed. Figure 4.14 (A) shows no change in Keratin 18 and perhaps a slight reduction in Keratin 8 levels following AP-2γ silencing in MCF-7 cells. This is consistent with the slight down regulation of the *KRT8* transcript observed by microarray gene expression profiling following AP-2γ silencing in MCF-7 cells.

Figure 4.14 (A) shows a large increase in the protein levels of Transgelin observed following AP-2γ silencing with each of the AP-2γ targeting siRNAs (although less pronounced for siRNA_3) and again is in support of the observed increase at the mRNA level (Figure 4.12). Transgelin (also known as SM22α) is a 22kDa actin binding protein that was first discovered as an early marker for the early onset of cellular transformation (Shapland *et al.*, 1993).

Figure 4.14 (A) shows a consistent decrease in the protein levels of Kiaa1324 observed following AP-2γ silencing with each of the AP-2γ targeting siRNAs and thus supports the observed decrease at the mRNA level (Figure 4.12). This antibody was a kind gift of Lei Deng (University of Texas) and detects a specific band above the 105 kDa MW marker, consistent with the predicted 121 kDa MW. Kiaa1324 is a protein of unknown function, although it was discovered in a microarray study to be over expressed in endometrial carcinoma (Deng *et al.*, 2005). In this study, Kiaa1324 was selected for further analysis as it was the most statistically significant down-regulated gene on the

microarray, following AP-2γ silencing in MCF-7 cells (Figure 4.8) and was represented by multiple probes sets. In the study in endometrial carcinoma, Kiaa1324 was identified as a novel oestrogen regulated gene (Deng *et al.*, 2005). In order to explore whether Kiaa1324 was similarly regulated in MCF-7 cells, changes in protein levels were monitored following 24 hours treatment with the anti-oestrogen, ICI182,780. Figure 4.14 (B) shows a reduction in the levels of Kiaa1324 protein following anti-oestrogen treatment suggesting that Kiaa1324 expression is modulated in response to oestrogen in this cell line. Immunofluoresence in wild type MCF-7 cells, shown in Figure 4.15 (C), shows a strong staining corresponding to the Kiaa1324 antibody at the cell membrane and discrete patches to one side of the nucleus. In an effort to understand the functional role of Kiaa1324, the protein sequence was interrogated using standard comparative proteomic web based tools. No homologous proteins of known function were revealed by a protein BLAST analysis (Altschul *et al.*, 1997). However, in support the cell membrane compartmentalisation, protein BLAST analysis against known proteins (Non-redundant SwissProt sequences), showed that a discrete region of Kiaa1324, from amino acid 270 to 420, shares significant homology with a number of membrane bound receptors (Altschul *et al.*, 1997), in particular regions containing conserved cysteine residues that are thought to be important for the formation of intra-chain disulphide bonds in the extra-cellular domains of several membrane bound receptors (Bateman *et al.*, 2004).

Figure 4.14 (B) also shows an increase in the protein levels of IGFPB3 observed following AP-2γ silencing with each of the AP-2γ targeting siRNAs and is in support of the observed increase at the mRNA level (Figure 4.12). A specific band between the 35 kDa and 50 kDa markers is consistent with a glycosylated form of IGFPB3 shown by others to run at approximately 40 kDa (Harms & Chen, 2005). IGFBP3 regulation was also explored in ZR75-1 cells following AP-2γ silencing. ZR-75-1 cells were transfected in an identical manner to MCF-7 cells. Figure 4.15 (A) shows the level of AP-2γ knockdown achieved was approximately 80% when compared to the control experiment and this was accompanied by a two fold increase in the relative abundance of *IGFPB3* mRNA for each of the AP-2γ targeting siRNAs.

Figure 4.14 (C) shows very slight increase in the protein levels of Inhibin βA observed following silencing with AP-2γ siRNA_3, although the difference could largely be accounted for by inaccuracy in protein loading marked by a similar increase in the Ku70 signal. This is inconsistent with the observed increase at the mRNA level (Figure 4.12) and suggests that in this cell line additional posttranscriptional mechanisms act to maintain the level of this protein despite transcriptional changes. Inhibins are members of the TGFβ super family. Work by others has shown that Activin A (the Inhibin βA dimerising protein) can mediate growth inhibition and cell cycle arrest through regulation of Smad transcription factor phosphorylation in mammary carcinoma cell lines (Burdette *et al.*, 2005). However, the negligible difference in Inhibin βA protein level observed here suggests it is not responsible for growth abnormalities following AP-2γ silencing in MCF-7 cells.

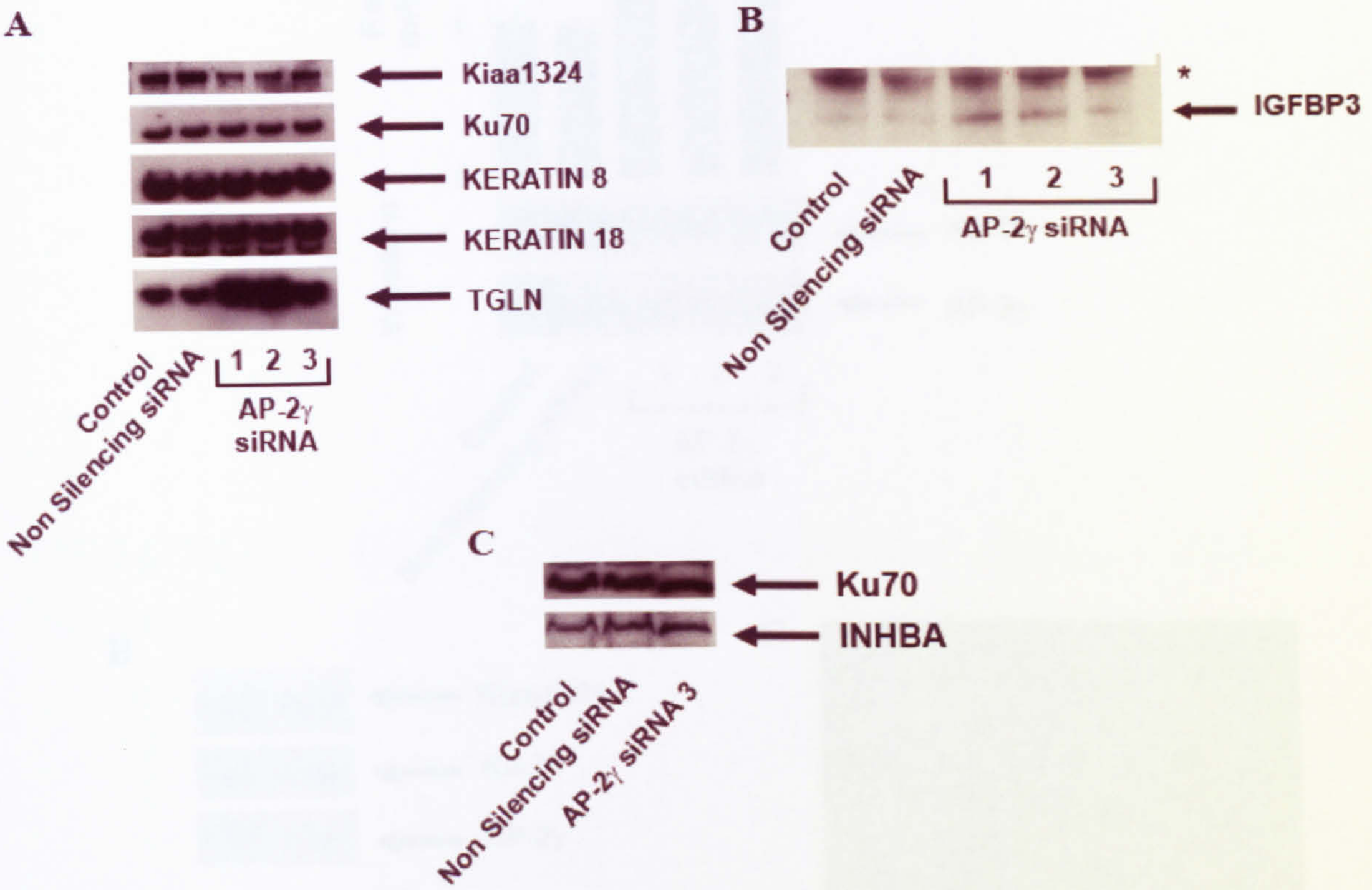


Figure 4.14. Validation of selected microarray gene expression changes by Western Blot analysis. MCF-7 cells were transfected as indicated and whole cell extracts harvested 96 hours later. **(A & C)** Whole Cell extracts (5µg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, Kiaa1324, Keratin 8, Keratin 18, Transgelin, and Inhibin βA. Blots were striped and reprobed where necessary. **(B)** In order to analyse IGFBP3 protein levels, conditioned tissue culture supernatant was collected from MCF7 transfected as indicated. Supernatant was 5x concentrated using a Centricon Centrifugal concentrator Unit with a 10 kDa MW cut off. 10µl of 5x concentrated supernatant was separated by SDS/PAGE and blotted to a membrane. Blots were then probed with primary antibodies IGFBP3, the asterisk shows a non-specific band that serves as a loading control.

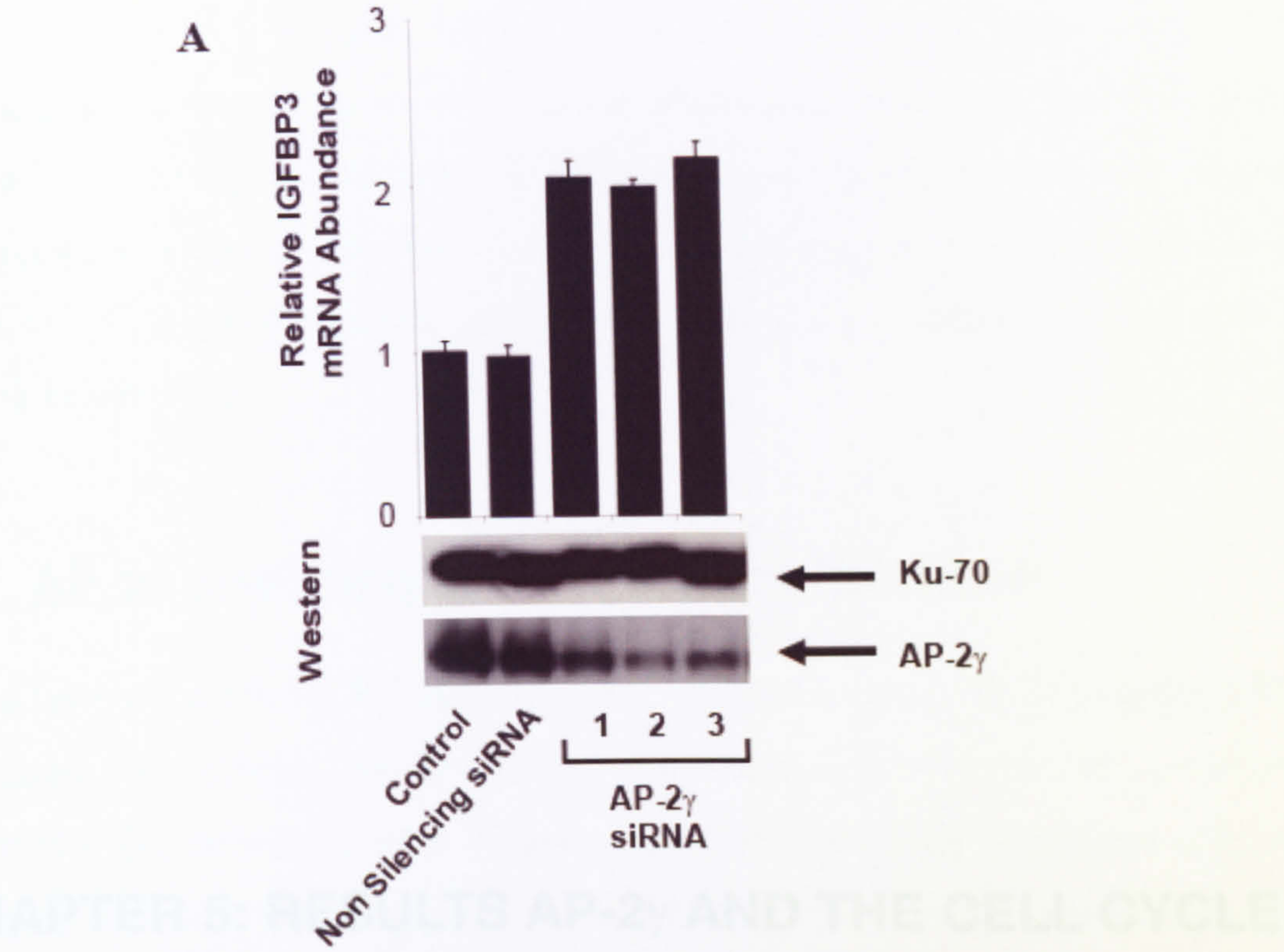


Figure 4.15. (A) *IGBFP3* mRNA up-regulation accompanies AP-2γ silencing in ZR75.1 cells. ZR75-1 cells were transfected as indicated and harvested 96 hours later. qPCR analysis of *IGFBP3* mRNA levels was normalised to the *GAPDH* mRNA levels. The control value was set at 1 to generate relative mRNA levels (average of three PCR replicates ± standard error). Whole Cell extracts (5µg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, AP-2γ. **(B) Western blot showing the effect of ICI182,780 anti-oestrogen treatment on Kiaa1324 levels in MCF-7 cells.** MCF-7 cells were split one in two into ICI182,780 (100nM) containing medium and harvested for western blot 24 h later. Whole Cell extracts (5µg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, AP-2γ, and Kiaa1324. **(C) Immunofluoresence of Kiaa1324 in wild type MCF-7 cells.** Strong Kiaa1324 staining (Red) was observed at the cell membrane and to one side of the nucleus. DAPI staining (Blue) is localised to the nuclear compartment. Details of the Immunofluoresence procedure are described in Section 2.4.

CHAPTER 5: RESULTS AP-2γ AND THE CELL CYCLE

Due to the inability to stably reduce AP-2γ levels in various breast epithelial lines (summarised in Table 3.2) coupled with the observation that genes significantly down-regulated following transient AP-2γ silencing were found to have GO Biological Process terms highly significantly over-represented in categories related to the cell cycle (Section 4.7), it was considered important to explore the affect AP-2γ appears to have on cell proliferation.

5.1. AP-2γ silencing affects cell proliferation.

Initial observations in MCF-7 cells transiently transfected with AP-2γ targeting siRNAs, suggested that reducing AP-2γ levels in these cells was affecting cell proliferation. Figure 5.1 (A) shows that after 72 hours, the cells transiently transfected with AP-2γ targeting siRNA_1 appeared fewer in number than those in the control experiments. Additionally, in cell counts performed on triplicate wells of AP-2γ targeting siRNA_1 and control experiments, reduced numbers of cells were observed in wells transfected with AP-2γ targeting siRNA_1. This is shown graphically in Figure 5.1. (B). These early effects on cell proliferation appeared to be independent of cell death, as numbers of floating cells were not increased in AP-2γ siRNA transfected cells. A SubG1 FACS cell cycle analysis was performed on cells transiently transfected with AP-2γ targeting siRNA_1. Briefly this analysis, as well as analysing DNA content to give a cell cycle profile, also records the subG1 population of cells with fragmented genomic DNA, a typical phenotype of cell death (reviewed in Samejima & Earnshaw, 2005). Figure 5.1 (D) revealed no increase in the SubG1 population after 72 hours and was consistent with our visual observations. Interestingly, an increase in the number of cells with a G1 DNA content was observed, accompanied by a decrease in the proportion of cells in S phase. These data are indicative of a cell cycle perturbation at the G1/S transition.

Figure 5.1. Transient AP-2γ silencing in MCF-7 cells affects cell proliferation. **(A)** MCF-7 cells were transfected as indicated and observed microscopically 72 hours following transfection. **(B)** MCF-7 cells were transfected as indicated and counted at 24 hour intervals following transfection. Counts are averages (± standard error) from triplicate experiments in wells of six-well plates and were performed using a Coulter Counter (Beckman). **(C)** Western blot analysis demonstrating the AP-2γ knock down efficiency 72 hours after transfection. Whole Cell extracts (5µg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ before being reprobed for Ku-70 as a loading control. **(D)** SubG1 FACS analysis for DNA content on control cells and those transfected with an AP-2γ targeting siRNA and harvested at 72 hours following transfection.

5.2. AP-2γ silencing affects cell cycle at the G1/S transition.

A more detailed cell cycle analysis was performed 72 h after transfection to further characterise the observed proliferation differences. These experiments were performed in triplicate and the effects of the three independent AP-2γ targeting siRNAs were examined. Additionally, cells were passaged at the 48-hour time point following transfection. This was to ensure that any observed differences in cell cycle profiling were not due to the cells in the control plates being more confluent than the AP-2γ knock down plates leading to contact inhibition of growth becoming an extra variable in these experiments. A one in two passage ensured that cells remained sub-confluent at the time of harvest and therefore the population consisted of mostly cycling cells. Figure 5.2 (A) displays the cell cycle profiles of cells 72 hours following transfection with the three independent AP-2γ targeting siRNAs. A significant increase in G1 population ($p<0.0.1$) was observed for each AP-2γ targeting siRNA compared to the control G1 populations. Additionally, a moderate but significant decrease in the proportion of cells in S phase ($p<0.0.1$) was observed for each AP-2γ targeting siRNA compared to the control S phase populations. No differences were observed between transfection control populations and cells transfected with a non-silencing siRNA. An additional analysis looking at S-phase activity of the cell population following transient transfection with AP-2γ targeting siRNA_3 is shown in Figure 5.2 (B). To provide a snapshot of the fraction of cells actively synthesising DNA, S-phase activity was measured by labelling transfected cell cultures with BrdU for a 20 minute period immediately prior to harvest (Materials and Methods 2.5). A significant reduction in the fraction of cells actively undergoing S phase is seen in cells transfected with AP-2γ targeting siRNA_3 compared to the control populations. The percentage of cells actively undergoing DNA synthesis was also consistent with the proportion of cells with S Phase DNA content shown in Figure 5.2 (A), this indicated that any cell cycle abnormality was occurring before the cells enter S-phase. Together these data suggest that cells transfected with AP-2γ targeting siRNAs were undergoing a partial cell cycle arrest at the G1/S transition.

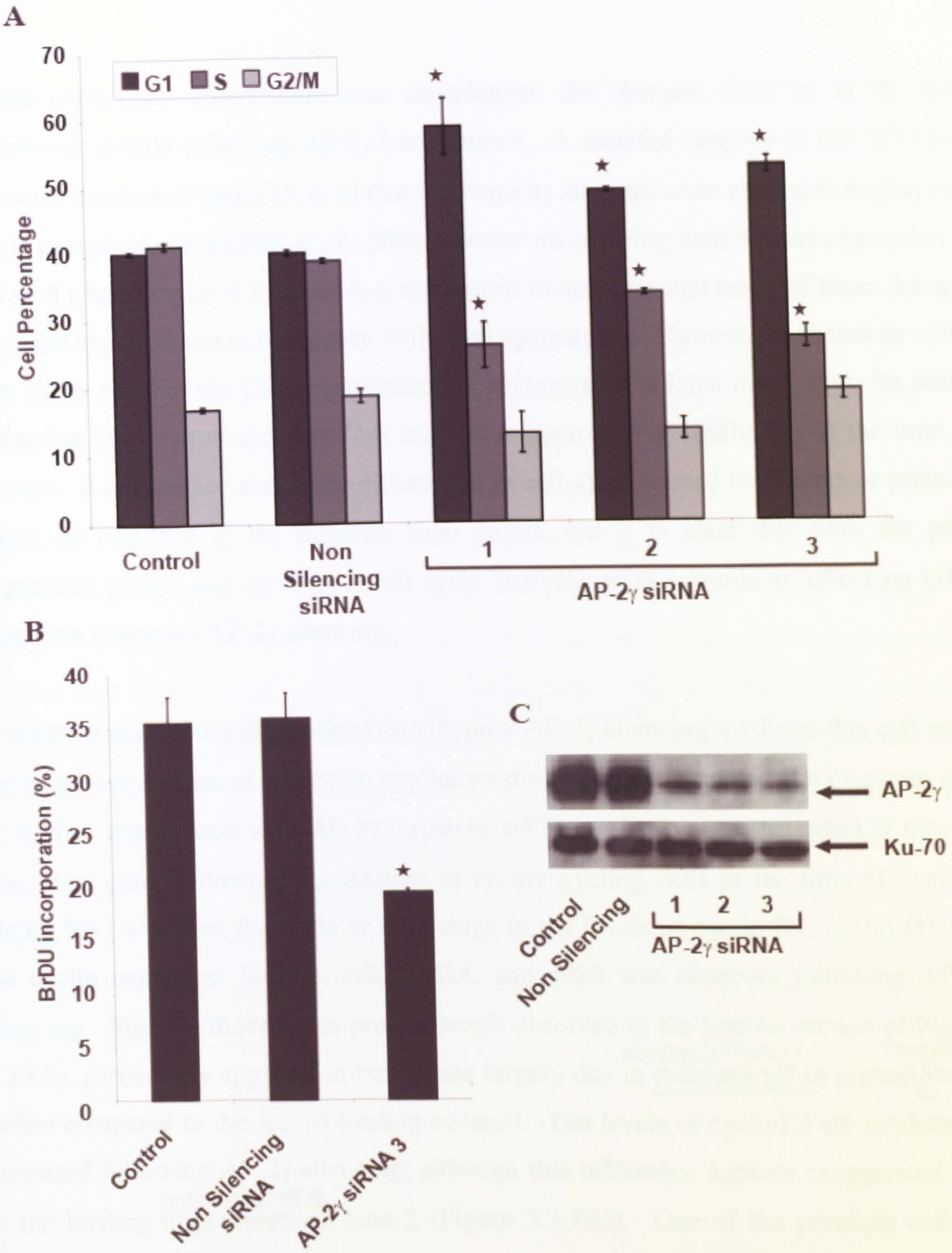


Figure 5.2. MCF-7 cells transiently transfected with AP-2γ targeting siRNAs show altered cell cycle distribution. **(A)** PI FACS analysis for DNA content performed 72 h after transfection. Cells were transfected as indicated, split 1 in 2 at 48 h, then harvested for cell cycle analysis 24 h later. **(B)** FACS analysis for BrDU incorparation performed 72 h after transfection. Cells were transfected as indicated, split 1 in 2 at 48 h, then harvested for cell cycle analysis at 72 h. Cells were labeled for BrDU for 20 mins before harvest. **(C)** Western blot analysis demonstrating the AP-2γ knock down efficiency 72 hours after transfection. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against AP-2γ before being reprobed for Ku-70 as a loading control. Averages from triplicate experiments are shown (± standard error). * $p<0.01$ compared to controls.

These phenotypic observations also complement the changes observed in the gene expression profile following AP-2γ knock down. A detailed analysis of the cell cycle associated subset of genes showed that the majority of these were known to display cell cycle periodicity (Whitfield *et al.*, 2002), almost all showing their lowest expression at G1 or S phase (Figure 4.11), and it is reasonable to consider that many of these changes my have occurred as a consequence of the cell cycle arrest. However, it should be noted that RNA for the microarray analysis was harvested at a later time point, 96 hours following transfection and therefore cells were approaching confluency at the time of harvest. It is expected that some differences in cell cycle related transcripts or proteins might be observed at the different time points, but it is clear that both the gene expression profile and the FACS cell cycle analysis point towards an effect on G1/S transition following AP-2γ silencing.

In order to understand the mechanism whereby AP-2γ silencing mediates this cell cycle arrest, the expression of cell cycle regulatory molecules were analysed by western blot 72 h after transfection with AP-2γ targeting siRNAs. Cells were passaged at the 48-hour time point following transfection to ensure cycling cells at the time of harvest. Figure 5.3 (A) shows that little or no change in the levels of cyclin D1, cyclin D3 and the cyclin dependent kinases, cdk2, cdk4, and cdk6 was observed following AP-2γ silencing. Slight differences in protein levels observed in the bottom section of Figure 5.3 (A), particularly apparent in lane 2, are largely due to inaccuracies in protein levels (when compared to the Ku-70 loading control). The levels of cyclinE2 are moderately increased following AP-2γ silencing, although this difference appears exaggerated due to the loading inaccuracies in lane 2 (Figure 5.3 (A)). One of the principle cellular targets for cyclin/cdk complexes in cell cycle regulation is the retinoblastoma family of pocket proteins. Western blot analysis in Figure 5.3 (B), using antibodies specific to phosphorylated forms of Rb protein, showed a reduction in detected levels following AP-2γ silencing. This indicates that cyclin/cdk activity was reduced following AP-2γ silencing, leading to hypophosphorylation of Rb. A consequence of reduced Rb phosphorylation is the formation of stable Rb/E2F complexes which leads to the repression of E2F target genes. Many known transcriptional targets of E2F family members (Ishida *et al.*, 2001; Polager *et al.*, 2002; Ren *et al.*, 2002). were shown to be

down-regulated in the gene expression profiling experiments following AP-2γ silencing in MCF-7 cells (Figure 4.11), consistent with an increase in Rb/E2F complex stability.

To assess whether up-regulation of Cyclin Dependent Kinase Inhibitor expression was associated with these cell cycle changes, levels of p15, p21 and p27 proteins were also analysed. Levels of p15 were barely detectable by western blot in control and AP-2γ silenced cells (data not shown). p16 was not assessed as it is not expressed in this cell line (Ragione *et al.*, 1996). However, an increase in p21 and p27 levels, as determined by western blot, was observed for each AP-2γ targeting siRNA. The increase in p27 levels is accompanied by a decrease in p27 phosphorylation at Thr187 (Figure 5.3). In cycling cells, p27 is phosphorylated by CyclinE/CDK2 kinase activity and targeted for proteolysis (Sheaff *et al.*, 1997). It is reasonable to consider therefore, that the increase in p27 levels following AP-2γ silencing could be largely due to a reduction in signalling activity promoting p27 proteolysis. An induction of p21 is of particular interest to this study as AP-2α has been implicated in the activation of *CDKN1A* transcription (Zeng *et al.*, 1997; McPherson *et al.*, 2002). It was tempting to speculate that AP-2γ might play an opposing role in *CDKN1A* transcriptional regulation, repressing transcription in wild type MCF-7 cells with AP-2γ silencing relieving this repression, leading to p21 activation and hence a reduction in Cyclin/CDK kinase activity leading to the observed cell cycle perturbation.

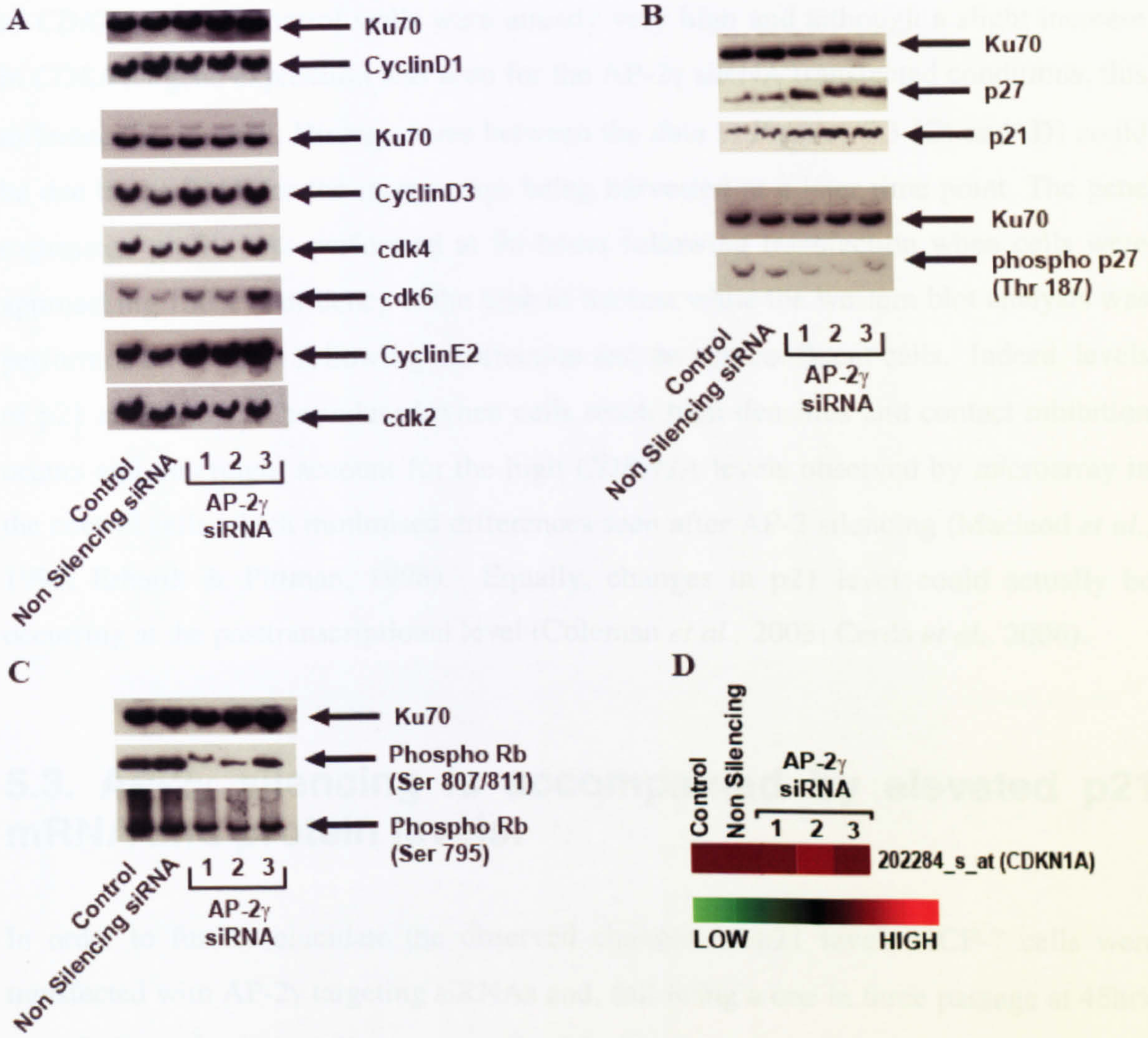


Figure 5.3. Elevated p21 levels following AP-2γ silencing in MCF-7 cells. **(A-C)** Cells were transfected as indicated, split 1 in 2 at 48 h, then harvested at 72 hours for western blot. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, Cyclin D1, Cyclin D3, cdk4, cdk6, cyclinE2, cdk2, p27, p21, phospho p27 (Thr 187) and phosphor RB (Ser 807/811 and Ser 795). Blots were striped and reprobed where necessary. **(D)** *CDKN1A* levels observed on gene expression microarrays 96 hours following transient transfection with AP-2γ targeting siRNAs in MCF-7 cells.

However, no significant gene expression change had been noted on the microarray for the *CDKN1A* (p21) transcript following AP-2γ silencing. Figure 5.3 (D) shows a heat map of normalised expression values for the *CDKN1A* probe set, 202284_at. The levels of *CDKN1A* in the control wells were already very high and although a slight increase in *CDKN1A* gene expression was seen for the AP-2γ siRNA transfected conditions, this difference is minimal. Discrepancies between the data in Figures 5.3 (C) and (D) could be due to the RNA for the microarrays being harvested at a later time point. The gene expression profile was performed at 96 hours following transfection when cells were approaching 100% confluency at the time of harvest while the western blot analysis was performed at 72 hours following transfection and on sub-confluent cells. Indeed, levels of p21 are known to be induced when cells reach high densities and contact inhibition occurs and this might account for the high *CDKN1A* levels observed by microarray in the control cells which minimised differences seen after AP-2 silencing (Macleod *et al.*, 1995; Erhardt & Pittman, 1998). Equally, changes in p21 level could actually be occurring at the posttranscriptional level (Coleman *et al.*, 2003; Cerda *et al.*, 2006).

5.3. AP-2γ silencing is accompanied by elevated p21 mRNA and protein levels.

In order to further elucidate the observed changes in p21 level, MCF-7 cells were transfected with AP-2γ targeting siRNAs and, following a one in three passage at 48hrs, both RNA and whole cell extracts were harvested at 72 and 96 hours. This was to ensure the cells were sub-confluent at the times of harvest. Figure 5.4 shows that using these transfection conditions elevated p21 mRNA (A) and protein (B) were observed at both time points. Additionally, Figure 5.4 (B) shows that the p21 protein level has increased in the control extracts harvested at 96 hours compared to those extracts harvested at 72 hours, suggesting that cell density is also a factor in regulation of p21 levels in MCF-7 cells. Changes in the amount of both p21 mRNA and p21 protein imply that its regulation occurs at the level of transcription upon AP-2γ silencing.

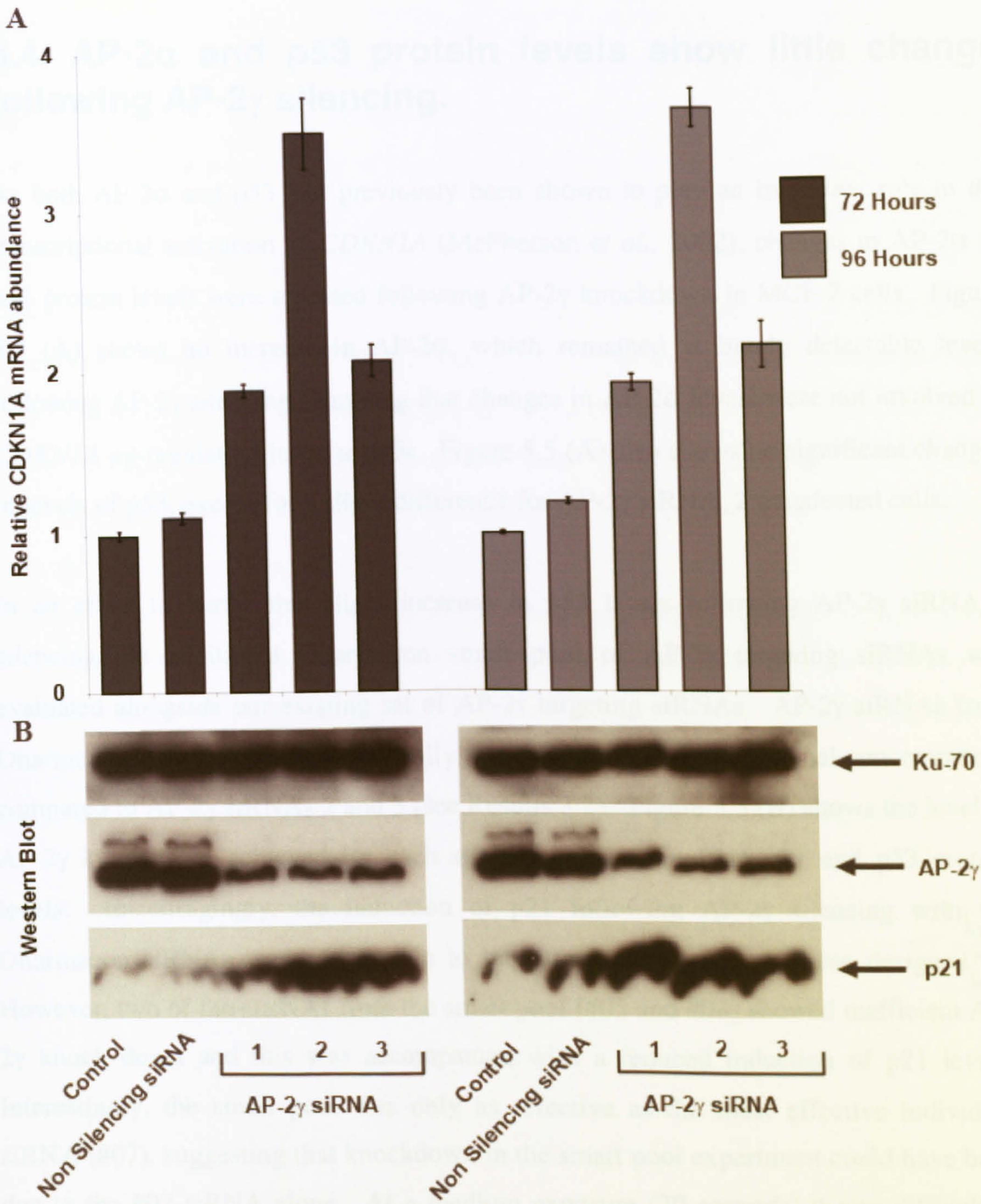


Figure 5.4. Elevated p21 mRNA and protein are observed following transient transfection with AP-2γ targeting siRNAs in cycling MCF-7 cells. Cells were passaged at the 48-hour time point following transfection and harvested at 72 and 96 hours. **(A)** Quantitative PCR analysis of *CDKN1A* mRNA levels. Data were normalised to *GAPDH* mRNA levels and the transfection reagent only (control) values, at 72 and 96 hours, were set at 1 to generate relative mRNA levels. The graph represents data averaged from three PCR replicates, ± standard error. **(B)** Whole Cell extracts (5µg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, p21 and AP-2γ.

5.4. AP-2α and p53 protein levels show little change following AP-2γ silencing.

As both AP-2α and p53 had previously been shown to play an important role in the transcriptional activation of *CDKN1A* (McPherson *et al.*, 2002), changes in AP-2α or p53 protein levels were assessed following AP-2γ knockdown in MCF-7 cells. Figure 5.5 (A) shows no increase in AP-2α, which remained at barely detectable levels following AP-2γ silencing, implying that changes in AP-2α levels were not involved in *CDKN1A* up-regulation in these cells. Figure 5.5 (A) also shows no significant changes in levels of p53, except for a slight difference for AP-2γ siRNA_2 transfected cells.

In an effort to verify this slight increase in p53 levels following AP-2γ siRNA_2 silencing, an additional Dharmacon smart pool of AP-2γ targeting siRNAs was evaluated alongside our existing set of AP-2γ targeting siRNAs. AP-2γ siRNAs from Dharmacon were evaluated individually and as a pool (at the same final concentration) compared to AP-2γ siRNAs 2 and 3 (see Results 3.2). Figure 5.5 (B) shows the level of AP-2γ knockdown achieved for each siRNA and the resulting p21 and p53 protein levels. Encouragingly, the induction of p21 following AP-2γ silencing with the Dharmacon siRNAs was comparable to the effect observed with those designed by us. However, two of the siRNAs from the smart pool (#05 and #08) showed inefficient AP-2γ knock down and this was accompanied with a reduced induction of p21 levels. Interestingly, the smart pool was only as effective as the most effective individual siRNA (#07), suggesting that knockdown in the smart pool experiment could have been due to the #07 siRNA alone. At a medium exposure (30 seconds) it was difficult to determine any differences in p53 levels following AP-2γ knockdown, however following a brief exposure (2 seconds), a very slight increase in p53 levels accompanying AP-2γ knockdown was revealed and this increase was observed for all the siRNAs transfected.

In order to put these slight changes in p53 level in to context we felt it was important to compare them to the changes in p53 levels following DNA damage stress. MCF-7 cells have been shown by others to express wild type p53 and exhibit an intact DNA damage

response pathway (Bartek *et al.*, 1990). Figure 5.6 (A) shows that 24 hours following treatment with cisplatin to induce DNA damage, p53 levels were increased in a dose-dependant manner. This is consistent with the well-documented stabilisation of p53 following DNA damage (Reviewed by Fei & El-Deiry, 2003) and was accompanied by an induction of p21 and BAX protein levels, known p53 transcriptional targets in response to DNA damage (el-Deiry *et al.*, 1993; Miyashita & Reed 1995; Fei & El-Deiry, 2003). This confirms that the wild type p53 protein in our MCF-7 cells behaved predictably following DNA damage stress, and rules out any p53 anomalies in our cells.

The increase in p53 levels seen following AP-2γ silencing (Figure 5.5 (B) Lanes 3-9) is modest compared to the stabilisation of p53 observed in cisplatin treated MCF-7 cells, even at in the 1μM cisplatin concentration (Figure 5.6 (A) Lane 2) and is inconsistent with the observed magnitude of p21 induction. Together these data suggest that this very small increase in p53 is not the major mechanism of p21 induction following AP-2γ silencing. Interestingly, an increase in AP-2α levels and a decrease in AP-2γ levels were also observed following cisplatin treatment (Figure 5.6 (B)). However, the change for AP-2γ is modest, with the apparent reduction in the last lane largely due to inaccuracies in protein levels (loading) (when compared to the Ku-70 signal Figure 5.6 (B)).

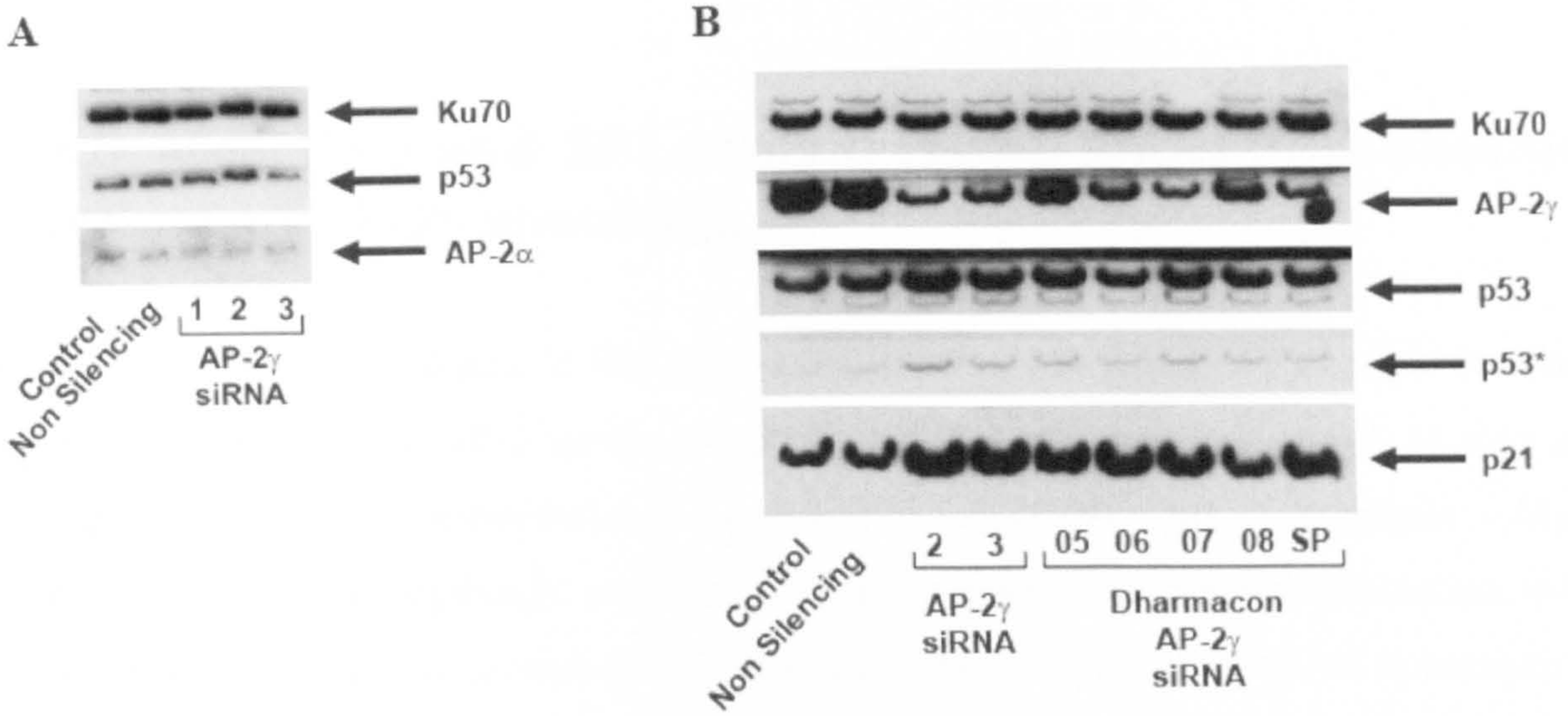


Figure 5.5. Assessment of AP-2α and p53 protein levels 72 hours following AP-2γ silencing in MCF-7 cells. **(A)** Western showing knock down with AP2γ siRNA 1, 2 and 3. **(B)** Comparison of AP2γ siRNA 2 and 3 and with Dharmacon AP-2γ sequences 05, 06, 07, 08 and the "Smart Pool" (SP). A total concentration of 25nM siRNAs was transfected for each experiment. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, p21, AP-2α, p53 and AP-2γ, blots were striped and reprobed where necessary. *Represents films that have undergone a very brief exposure approximately 2 seconds.

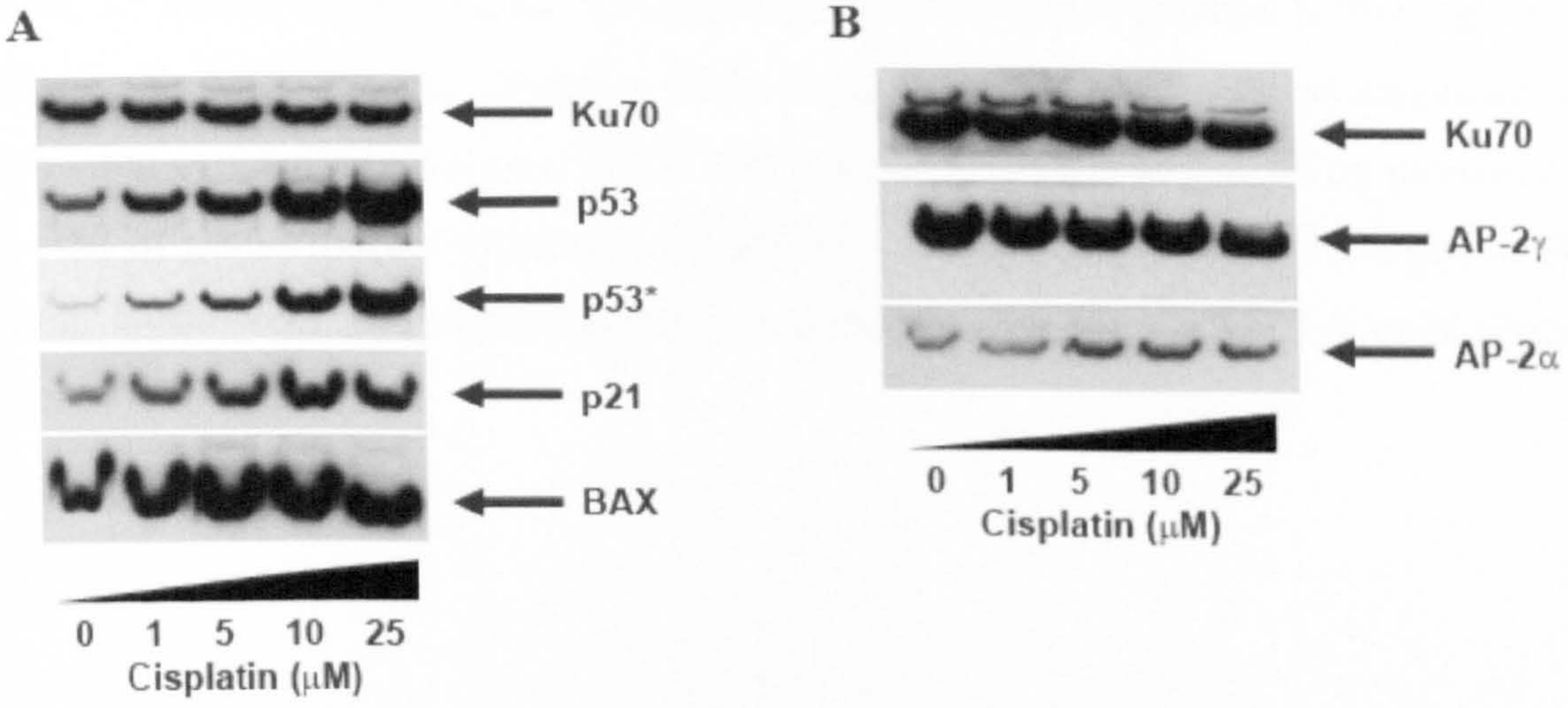


Figure 5.6. Cisplatin mediated DNA damage in MCF-7 cells leads to a stablilisation of p53 and activation of p53 target genes p21 and BAX (A) and modest changes in AP-2 family members (B). MCF-7 cells were treated with cisplatin and harvested 24 hours later. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, p21, AP-2α BAX p53 and AP-2γ, blots were striped and reprobed where necessary. *Represents films that have undergone a very brief exposure approximately 2 seconds.

5.5. Myc, ERα and SP1 protein levels show little change following AP-2γ silencing.

As well as AP-2α and p53, it was considered important to assess the changes in other proteins regulated by AP-2 family members, which have also been shown to play an important role in the transcriptional regulation of *CDKN1A*. One such protein, c-Myc has been shown to negatively regulate *CDKN1A* transcription, through interaction with proximal SP1 sites or via initiator elements down-stream of the start of transcription (reviewed in Gartel & Radhakrishnan 2005). Myc has also been suggested to be under negative control by AP-2 family member, AP-2α (Gaubatz *et al.*, 1995). Because of these findings it was important to assess the contribution of Myc to these observations. Figure 5.7 shows there was no or little change in Myc levels following AP-2γ silencing in MCF-7 cells and suggests that it is not regulated by AP-2γ in this cell line. Additionally, AP-2γ has been shown to maintain ERα transcription in breast cancer cell lines (reviewed in Hilger-Eversheim *et al.*, 2000). In order to assess the contribution of ERα to the observed cell cycle changes, levels were assessed by western blot. However, Figure 5.7 shows that ERα does not contribute to cell cycle regulation following AP-2γ silencing in MCF-7 cells. Finally, many factors have been shown to modulate the expression of *CDKN1A* via interaction with the six SP1 sites in the 100bp proximal to the start of transcription (reviewed in Gartel & Radhakrishnan 2005). Changes in SP1 protein levels were also unaffected following AP-2γ silencing in MCF-7 cells (Figure 5.7).

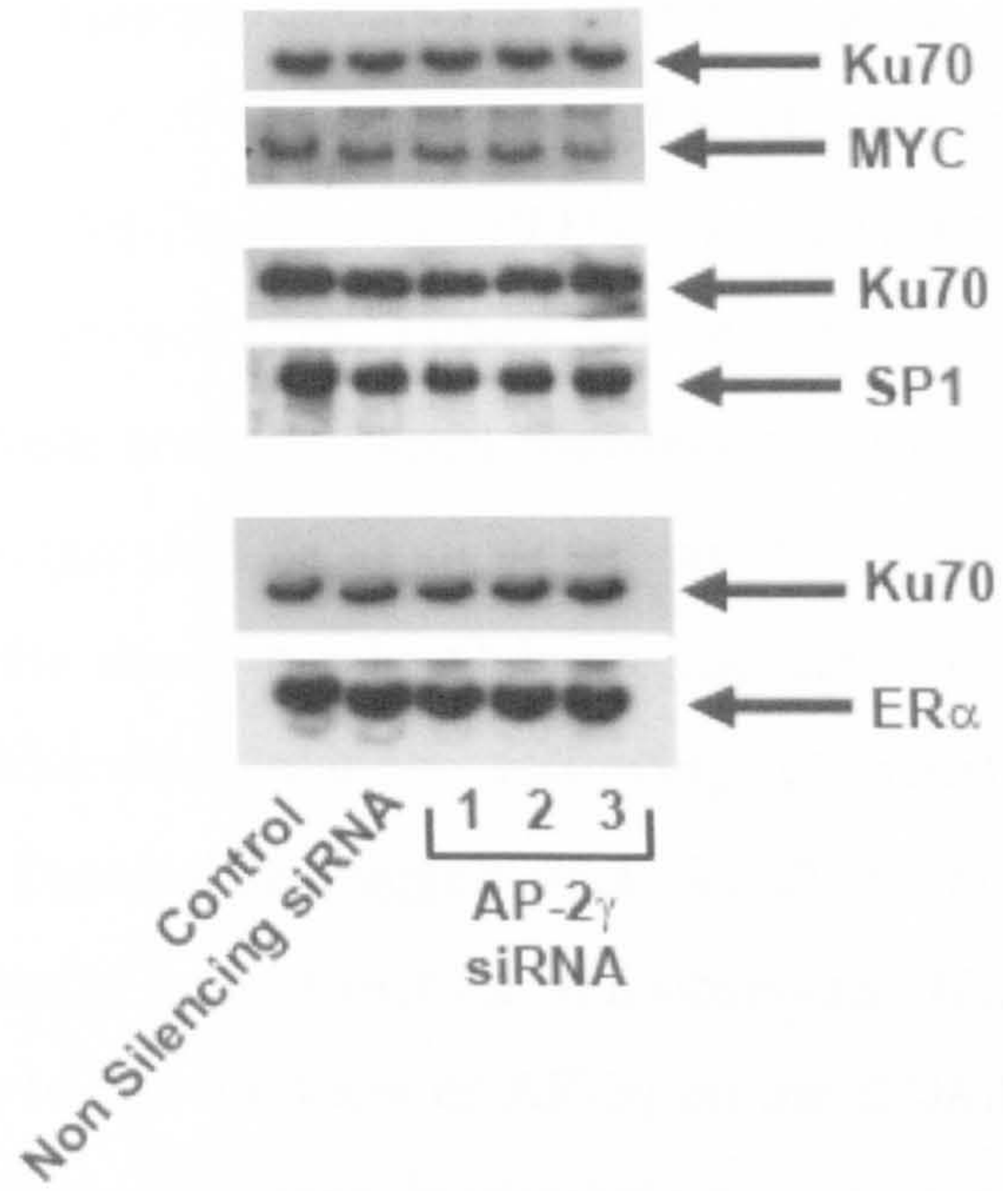


Figure 5.7 Assessment of Myc, SP1, and ERα protein levels 72 hours following AP-2γ silencing in MCF-7 cells. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, Myc, SP1 and ERα, blots were striped and reprobed where necessary.

5.6. AP-2γ expression represses a CDKN1A promoter reporter construct

In order to examine the possible effect of AP-2γ over expression on *CDKN1A* transcription, luciferase reporter assays were conducted to assess *CDKN1A* promoter activity in AP-2 non-expressing HepG2 cells. Genomic sequences from –2325 to +8 relative to the start of transcription of the *CDKN1A* gene have been shown to be important for the regulation of transcription and have been previously cloned into a luciferase reporter vector (el-Deiry *et al.*, 1993). Initially, the reporter construct alone was transfected at different amounts. Figure 5.8 (A) shows that good levels of basal luciferase expression were achieved when transfecting 1μg of reporter, therefore this amount was used in co-transfection studies with an AP-2γ expression construct. The results in Figure 5.8 (B) show that basal luciferase activity from this construct was reduced in a dose-dependant manner following co-transfection with an AP-2γ expression construct. These data suggest that AP-2γ is able to repress *CDKN1A* promoter activity in HepG2 cells. However, the observed effects on luciferase activity are not necessarily due to direct effects of AP-2γ on the *CDKN1A* promoter, and could be due to indirect regulation via other downstream targets of AP-2γ or via sequestration of factors away from the reporter construct.

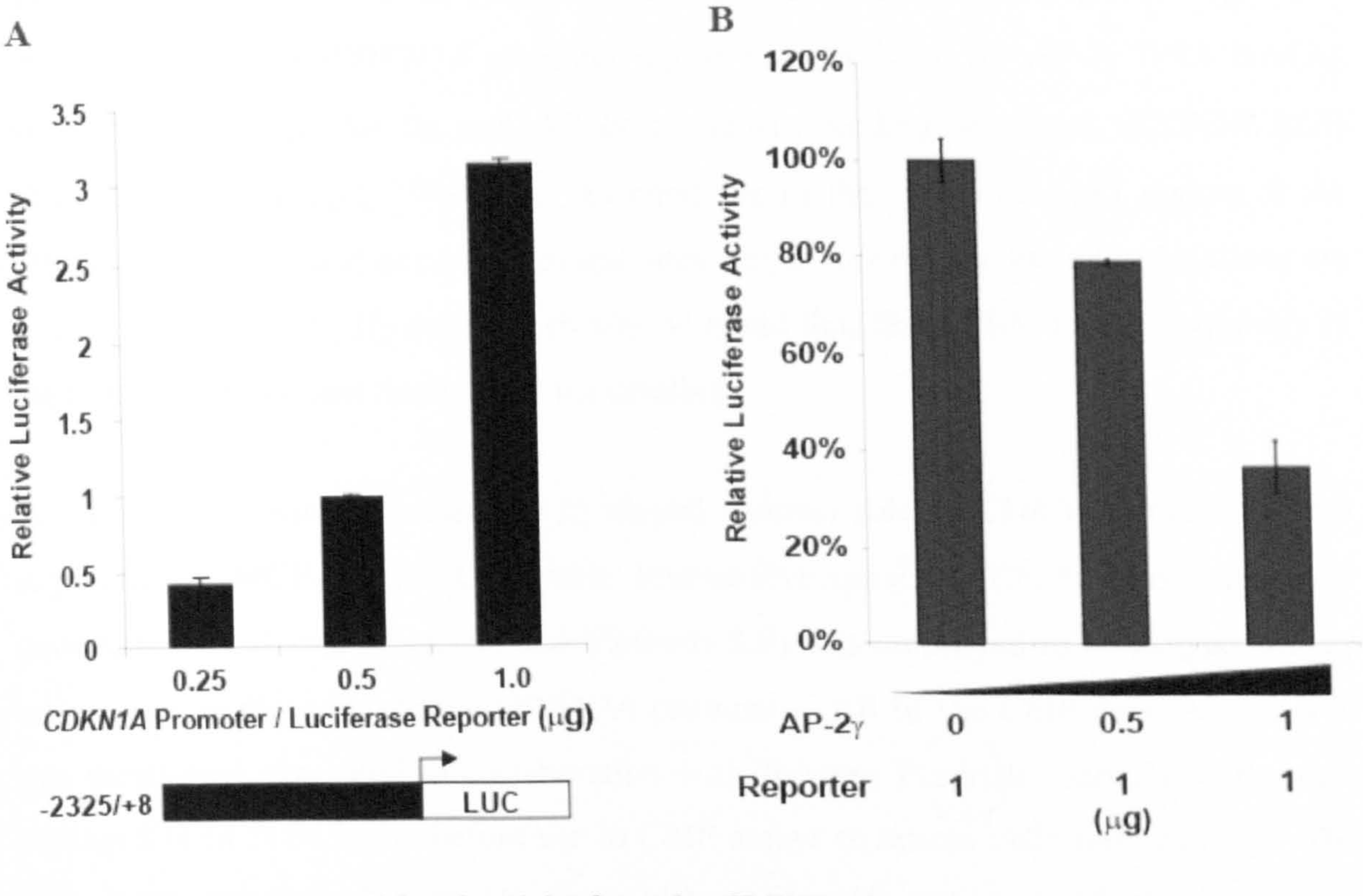


Figure 5.8. AP-2γ repressed basal activity from the *CDKN1A* promoter in HepG2 cells. The AP-2 non-expressing cell line, HepG2 was transiently transfected with 0.25, 0.5 and 1.0μg of the -2325\+8 *CDKN1A* luciferase reporter construct **(A)** or 1.0μg of the -2325\+8 *CDKN1A* luciferase reporter construct and the indicated amounts of the AP-2γ expression construct **(B)**. A renilla luciferase construct (phRG-TK) was co-transfected in each experiment and the differences in DNA amount were compensated for using a promoterless Luciferase construct (A) or pcDNA3 plasmid (B). Forty-eight hours following transfection cells were assayed for luciferase activity. Results (mean ± S.E. of three independent experiments) are presented as relative luciferase activity, corrected for Renilla luciferase activity. Cells transfected with the *CDKN1A* promoter luciferase reporter alone (B) were set at 100%. See materials and methods for plasmid details.

5.7. ChIP reveals occupancy of the CDKN1A promoter by endogenous AP-2γ

To understand whether AP-2γ played a direct role in *CDKN1A* transcriptional regulation in MCF-7 cells, the *CDKN1A* promoter region was analysed for AP-2γ DNA binding sites. The generic AP-2α and AP-2γ consensus-binding sequence SCCNNVRGB (McPherson & Weigel, 1999) was searched for in the -4020 to +185 region of the *CDKN1A* promoter and occurs at several sites across this region, the exact locations are shown in Figure 5.9. However, it should be noted that this TFBS occurs frequently in the human genome (see Section 4.8 for details).

In order to determine whether AP-2γ played a direct role in *CDKN1A* transcriptional regulation in MCF-7 cells, Chromatin ImmunoPrecipitation (ChIP) accompanied by quantitative PCR (see Materials and Methods 2.9) was employed to investigate AP-2γ occupancy on the endogenous *CDKN1A* promoter. All of the ChIP data presented in this thesis was generated in collaboration with Karsten Fredrich. Briefly, cells were passaged (1 in 2) 24 hours before use in ChIP assays to ensure cells were cycling. The cells were crosslinked with 1% formaldehyde for 10 minutes and chromatin was sheared by sonication to an average size of 500bp. Immunoprecipitation reactions were carried out overnight using predetermined amounts of antibody and then washed with optimised wash conditions. Following immunoprecipitation, and the DNA was extracted using Chelex beads. A "no antibody control" was treated in the same manner as the other samples and used to determine the background level to which the qPCR data were normalised. Four sets of qPCR primers were designed and evaluated for their ability to specifically amplify regions across the *CDKN1A* promoter (see Material and Methods 2.9: Table 2.7) Figure 5.10 (A) shows the location of primers relative to known regulatory regions across the *CDKN1A* promoter and summarises the data in Figure 5.9. Primer set -4005 / -3931 was designed across a region of high species conservation, with potential AP-2 binding sites. Primer set -2290 / -2185 was designed against a region flanking the classical p53 binding site (Kaeser & Iggo 2002), also a region previously shown to be occupied by AP-2α (Mcpherson *et al.*, 2002). The primer set -935 / -864 was designed across a region with potential AP-2 binding sites binding sites and finally the -21 / +44 primers were designed across a region encompassing the

start of transcription. This region encompassed several potential AP-2 binding sites and a region previously described to be important for transcriptional activation by AP-2α (Zeng *et al.*, 1997). These sets of four primers were sufficiently spaced across the p21 promoter region that they would not amplify overlapping chromatin fragments, considering the chromatin was sheared to an average size of 500bp. In addition to these p21 promoter specific sequences, primers were designed within a centromeric region consisting of satellite repeat (SAT2) sequences (Jiang *et al.*, 2004). This region of constitutive heterochromatin on Chromosome 1 was used as a negative control to determine the specificity of the antibodies. Binding would be expected to be absent in this region for the factors examined in this study. In addition to examining AP-2γ occupancy across these regions, I also thought it would be informative to examine the occupancy of the known AP-2γ co-factor, p300, and also to examine the acetylation status of the H4 histone protein, a marker of active chromatin.

Figure 5.10 (B) shows the relative enrichment of *CDKN1A* promoter specific sequences measured in AP-2γ, p300 and acH4 ChIP elutes. None of the antibodies used in these ChIP experiments were enriched for amplification from the SAT2 heterochromatin region implying that the activity observed at the *CDKN1A* locus was specific for each protein target. Interestingly these data showed that AP-2γ occupancy within the *CDKN1A* endogenous promoter region was confined to the region amplified by the -21 / +44 primer set. AP-2γ did not occupy any of regions amplified by the other primer sets further upstream and in particular it was not found associated near the classical p53 binding site amplified by primer set -2290 / -2185. It is intriguing that AP-2γ occupancy was associated with p300 occupancy across the *CDKN1A* promoter, p300 also showing greatest enrichment near the region amplified by the -21 / +44 primer set. Although p300 is a known co-factor of AP-2γ, recruited through CITED adapter proteins, it has previously only been described in the context of AP-2 transcriptional activation and it is somewhat surprising that p300 is also associated with AP-2γ when it is potentially mediating transcriptional repression. Strikingly the AP-2γ and p300 occupied region amplified by the -21 / +44 primer set shows a reduction in acH4 is when compared to the acH4 occupancy near regions amplified by the -2290 / -2185 and -935 / -864 regions, suggesting that the -21 / +44 region is less accessible to the general transcription machinery than the regions upstream.

-4030 -4020 -4010 -4000 -3990 -3980 -3970 -3960

ATCTCAGTTTGCCCATTAATATTATAGGTCTGGCCCCTCTGTGAAAACATGCCCAGCGGCAAGGCCAACAAGCTGGGGGATGTTGTTA

TAGAGTCAAACGGGTAATTATAATATCCAGACCGGGGAGACACTTTTGTACGGGTCGCCGTTCCGGTTGTTCGACCCCCTACAACAAT

-3940 -3930 -3920 -3910 -3900 -3890 -3880 -3870

GAGCCAGGAACAGGAAGACCATCCAGGTTGGTAACACTGATGCTGAGGGGAGGCTCATACTGGCTGATGCGCTCTGTTACGTGCACAC

CTCGGTCCTTGTCCTTCTGGTAGGTCCAACCATTGTGACTACGACTCCCCTCCGAGTATGACCGACTACGCGAGACAATGCACGTGTG

-3850 -3840 -3830 -3820 -3810 -3800 -3790 -3780

ATTTAACCCGAAGGTCATCCTCAATGCCACCACCTTAACAGGTGTCATAGATGTAGCTTTGGGGTCAGGTGCCACTGGGGTCTTTACC

TAAATTGGGCTTCCAGTAGGAGTTACGGTGGTGGAATTGTCCACAGTATCTACATCGAAACCCCAGTCCACGGTGACCCCAGAAATGG

-3770 -3760 -3750 -3740 -3730 -3720 -3710 -3700 -3690

AATTCATCCTGGCTCTGGAACAAGCTCTTCGAGGCCAGCATTGAAACAGGGGACCGTGTCTGGAGGATGCCTCTCTTCAAACATTGTA

TTAAGTAGGACCGAGACCTTGTTCGAGAAGCTCCGGTCGTAACTTTGTCCCCTGGCACAGACCTCCTACGGAGAGAAGTTTGTAACAT

-3680 -3670 -3660 -3650 -3640 -3630 -3620 -3610 -3600

CAAGACAGGTTGTAGATTGCCAGCTGGCTGATGTTAACAACATTGGAAAATATAGATCTGCGGGAGCATGTACATCTGCGGCATTCCT

GTTCTGTCCAACATCTAACGGTCGACCGACTACAATTGTTGTAACCTTTTATATCTAGACGCCCTCGTACATGTAGACGCCGTAAGGA

-3590 -3580 -3570 -3560 -3550 -3540 -3530 -3520

GAAAGAATTCGTGACTCATCCTAAGTGGGCACATTTAGACATAGCAGGTGTGATGACCAACAAAGATGAGGTTCCCTATCTATGGAAA

CTTTCTTAAGCACTGAGTAGGATTCACCCGTGTAAATCTGTATCGTCCACACTACTGGTTGTTTCTACTCCAAGGGATAGATACCTTT

-3500 -3490 -3480 -3470 -3460 -3450 -3440 -3430

GGCATGACCGGGAGGCCCACAAGGACTCTCATAGAGTTCTTACTTCGTTTCAGTCAAGACAATGCTTAGTTCAGATACTCAAAAATGT

CCGTACTGGCCCTCCGGGTGTTCCTGAGAGTATCTCAAGAATGAAGCAAAGTCAGTTCTGTTACGAATCAAGTCTATGAGTTTTTACA

-3410 -3400 -3390 -3380 -3370 -3360 -3350 -3340

CTTCACTCTATCTTAAATTGGACAGTTGAAGTTAAAAGGTTTTTGAATGAATGGATGAAAATATTTTAAAGGAGGCAATTTATATTTA

GAAGTGAGATAGAATTTAACCTGTCAACTTCAATTTTCCAAAAACTTACTTACCTACTTTTATAAAATTTCCTCCGTTAAATATAAAT

-3330 -3320 -3310 -3300 -3290 -3280 -3270 -3260 -3250

AAAATGTAGAACACAATGAAATTTTTATGCCTTGATTTTTTTTTCATTTTACACAAAGATTTATATATTTTTTTTTTGAGACGGAGTC

TTTTACATCTTGTGTTACTTTAAAAATACGGAACTAAAAAAAAAGTAAAATGTGTTTCTAAATATATAAAAAAAAAACTCTGCCTCAG

-3240 -3230 -3220 -3210 -3200 -3190 -3180 -3170 -3160

TCACTCTGTCACCCAGGCTGGAGTGCAGGTGGCATGATCTCAGCTCACTGCAACCTCCGCCTCCTAGGTTCAAGCGATTCTCCCACCT

AGTGAGACAGTGGGTCCGACCTCACGTCCACCGTACTAGAGTCGAGTGACGTTGGAGGCGGAGGATCCAAGTTCGCTAAGAGGGTGGA

-3150 -3140 -3130 -3120 -3110 -3100 -3090 -3080

CAGCCACCTGAATACCTGGGACTACAGGTGCCCACCACCATGCCCGGCTGATTTTTGTATTTTTAATGGAGACGGGGTTTCACCATAT

GTCGGTGGACTTATGGACCCTGATGTCCACGGGTGGTGGTACGGGCCGACTAAAAACATAAAAATTACCTCTGCCCCAAAGTGGTATA

-3060 -3050 -3040 -3030 -3020 -3010 -3000 -2990

TGGCCAGGCTGGTCTCAAAACTCCTGACCCTGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTAAGCCACCACG

ACCGGTCCGACCAGAGTTTTGAGGACTGGGACACTAGACGGGCGGAGCCGGAGGGTTTCACGACCCTAATGTCCGCATTCGGTGGTGC

-2970 -2960 -2950 -2940 -2930 -2920 -2910 -2900

CCCGGCCAGTATATATTTTTAATTGAGAAGCAAAATTGTACTTCAGATTTGTGATGCTAGGAACATGAGCAAACTGAAAATTACTAAC

GGGCCGGTCATATATAAAAATTAACTCTTCGTTTTAACATGAAGTCTAAACACTACGATCCTTGTACTCGTTTGACTTTTAATGATTG

-2890 -2880 -2870 -2860 -2850 -2840 -2830 -2820 -2810

CACTTGTCAGAAACAATAAATCCAACTTTTTGTGCAAAAAAAAAAATACAAATATTAGCTGGGCATGGTGGTGCATGCCTGTAATCCC

GTGAACAGTCTTTGTTATTTAGGTTGAAAAACACGTTTTTTTTTTTATGTTTATAATCGACCCGTACCACCACGTACGGACATTAGGG

-2800 -2790 -2780 -2770 -2760 -2750 -2740 -2730 -2720

AGCTACTCGGGAGGCTGAGGCAGAATTGCTTGAACCTGGGAGGCGGAGACTGCAGTGAGCTGAGATTGTGCCACTGCTGACTTTGTCT

TCGATGAGCCCTCCGACTCCGTCTTAACGAACTTGGACCCTCCGCCTCTGACGTCACTCGACTCTAACACGGTGACGACTGAAACAGA

-2710 -2700 -2690 -2680 -2670 -2660 -2650 -2640

CAAAAAACAAAACAAAACAAAAAAACAAAATGAAAACAAAAAGCCAGGGCTGCCTCTGCTCAATAATGTTCTATCTTTGTTCCGCCTC

GTTTTTTGTTTTGTTTTGTTTTTTTGTTTTACTTTTGTTTTTCGGTCCCGACGGAGACGAGTTATTACAAGATAGAAACAAGGCGGAG

-2620 -2610 -2600 -2590 -2580 -2570 -2560 -2550

TTCTCTGGGGTCTCACTTCTTGGGAGCCTGTGTGAAGGTGAATTCCTCTGAAAGCTGACTGCCCCTATTTGGGACTCCCCAGTCTCTT

AAGAGACCCCAGAGTGAAGAACCCTCGGACACACTTCCACTTAAGGAGACTTTCGACTGACGGGGATAAACCCTGAGGGGTCAGAGAA

-2530 -2520 -2510 -2500 -2490 -2480 -2470 -2460

TCTGAGAAATGGTGACATTGTTCCCAGCACTTCCTCTCCCTTCCTAGGCAGCTTCTGCAGCCACCACTGAGCCTTCCTCACATCCTCC

AGACTCTTTACCACTGTAACAAGGGTCGTGAAGGAGAGGGAAGGATCCGTCGAAGACGTCGGTGGTGACTCGGAAGGAGTGTAGGAGG

-2450 -2440 -2430 -2420 -2410 -2400 -2390 -2380 -2370

TTCTTCAGGCTTGGGCTTTCCACCTTTCACCATTCCCCTACCCCATGCTGCTCCACCGCACTCTGGGGAGGGGGCTGGACTGGGCACT

AAGAAGTCCGAACCCGAAAGGTGGAAAGTGGTAAGGGGATGGGGTACGACGAGGTGGCGTGAGACCCCTCCCCCGACCTGACCCGTGA

-2360 -2350 -2340 -2330 -2320 -2310 -2300 -2290 -2280

CTTGTCCCCCAGGCTGAGCCTCCCTCCATCCCTATGCTGCCTGCTTCCCAGGAACATGCTTGGGCAGCAGGCTGTGGCTCTGATTGGC

GAACAGGGGGTCCGACTCGGAGGGAGGTAGGGATACGACGGACGAAGGGTCCTTGTACGAACCCGTCGTCCGACACCGAGACTAACCG

-2270 -2260 -2250 -2240 -2230 -2220 -2210 -2200

TTTCTGGCCGTCAGGAACATGTCCCAACATGTTGAGCTCTGGCATAGAAGAGGCTGGTGGCTATTTTGTCCTTGGGCTGCCTGTTTTC

AAAGACCGGCAGTCCTTGTACAGGGTTGTACAACTCGAGACCGTATCTTCTCCGACCACCGATAAAACAGGAACCCGACGGACAAAAG

-2180 -2170 -2160 -2150 -2140 -2130 -2120 -2110

AGGTGAGGAAGGGGATGGTAGGAGACAGGAGACCTCTAAAGACCCCAGGTAAACCTTAGCCTGTTACTCTGAACAGGGTATGTGATCT

TCCACTCCTTCCCCTACCATCCTCTGTCCTCTGGAGATTTCTGGGGTCCATTTGGAATCGGACAATGAGACTTGTCCCATACACTAGA

-2090 -2080 -2070 -2060 -2050 -2040 -2030 -2020

GCCAGCAGATCCTTGCGACAGGGCTGGGATCTGATGCATGTGTGCTTGTGTGAGTGTGTGCTGGGAGTCAGATTCTGTGTGTGACTTT

CGGTCGTCTAGGAACGCTGTCCCGACCCTAGACTACGTACACACGAACACACTCACACACGACCCTCAGTCTAAGACACACACTGAAA

-2010 -2000 -1990 -1980 -1970 -1960 -1950 -1940 -1930

TAACAGCCTGCTCCCTTGCCTTTTTCAGGGCAGAAGTCCTCCCTTAGAGTGTGTCTGGGTACACATTCAAGTGCATGGTTGCAAACTT

ATTGTCGGACGAGGGAACGGAAAAAGTCCCGTCTTCAGGAGGGAATCTCACACAGACCCATGTGTAAGTTCACGTACCAACGTTTGAA

-1920 -1910 -1900 -1890 -1880 -1870 -1860 -1850 -1840

TTTTTTTTAAAGCACTGAATAGTACTAGACACTTAGTAGGTACTTAAGAAATATTGAATGTCGTGGTGGTGGTGAGCTAGAAGTTATA

AAAAAAAATTTCGTGACTTATCATGATCTGTGAATCATCCATGAATTCTTTATAACTTACAGCACCACCACCACTCGATCTTCAATAT

-1830 -1820 -1810 -1800 -1790 -1780 -1770 -1760

AAAAAAATTCTTTCCCAAAAACAACAACAAAAAGAATTATTTCATTGTGAAGCTCAGTACCACAAAAATTTAAATAATTCATTACAAG

TTTTTTTAAGAAAGGGTTTTTGTTGTTGTTTTTCTTAATAAAGTAACACTTCGAGTCATGGTGTTTTTAAATTTATTAAGTAATGTTC

-1740 -1730 -1720 -1710 -1700 -1690 -1680 -1670

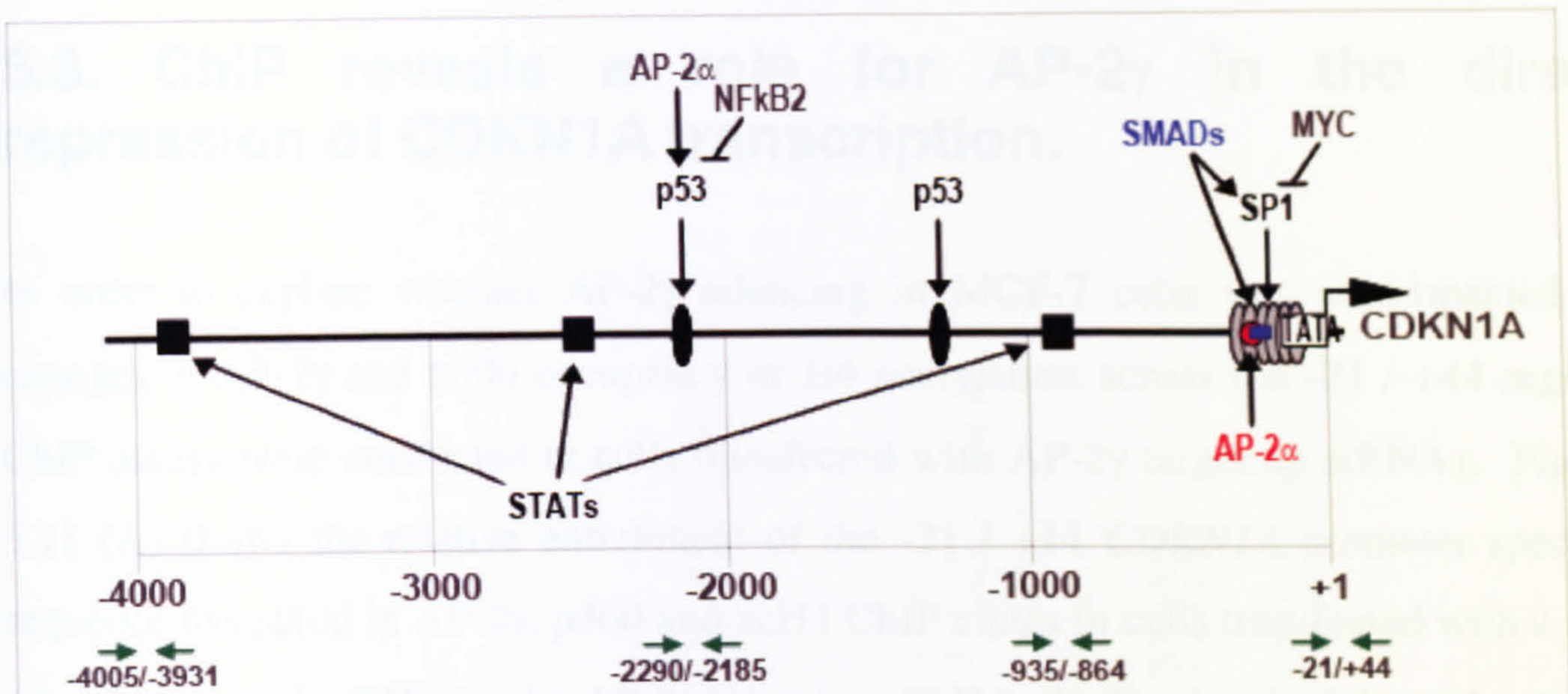
CCTTTATTAAAAAAAATTTTCTCCCCAAAGTAAACAGACAGACAATGTCTAGTCTATTTGAAATGCCTGAAAGCAGAGGGGCTTCAAG

GGAAATAATTTTTTTTAAAAGAGGGGTTTCATTTGTCTGTCTGTTACAGATCAGATAAACTTTACGGACTTTCGTCTCCCCGAAGTTC

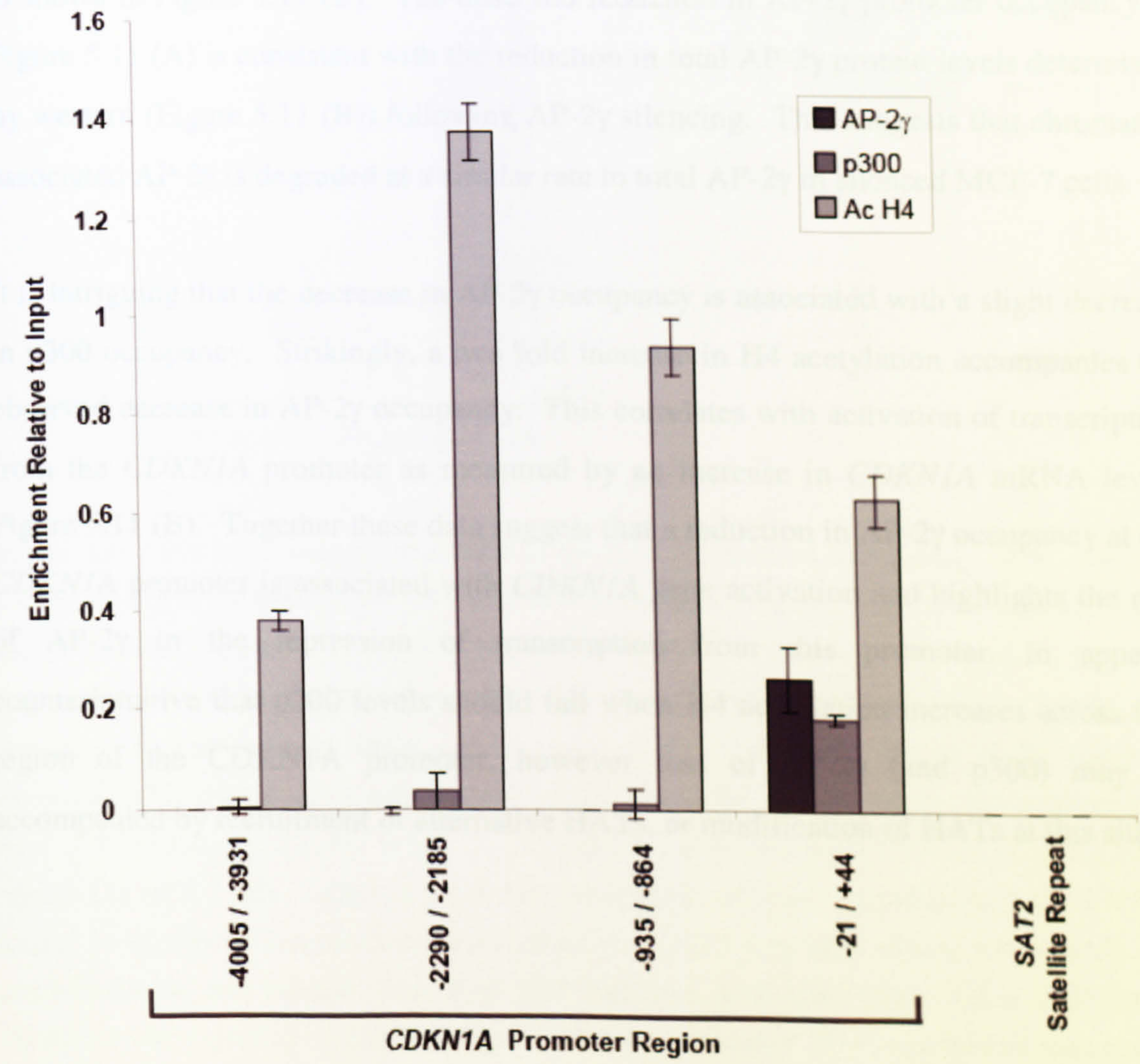
Figure 5.9 CDKN1A promoter region. The CDKN1A promoter from positions -4020 to +185 is shown with potential AP-2 binding sites based on the SCCNNVRGB consensus (McPherson & Weigel, 1999) highlighted in yellow. p53 binding sites are highlighted in green and SP1 sites highlighted in blue. Sequences transcribed into mRNA are marked in red, indicated by the +1 position. p53 sites described previously by (Kaeser & Iggo, 2002). Quantitative PCR Primers used to amplify regions for ChIP assays are in bold and underlined.

Figure 5.10 ChIP assay showing endogenous AP-2γ and p300 occupancy and acH4 levels across the *CDKN1A* promoter region. **(A)** Shows summarises the location of key regulatory elements across the *CDKN1A* promoter in relation to the primers used for the quantitative PCR. **(B)** ChIP assay for AP-2γ, p300 and AcH4 in wild type MCF-7 cells. Cells were passaged (1 in 2) 24 hours before use in ChIP assays to ensure cells were cycling. Results were normalised to a no antibody control and expressed relative to a 1/1000 dilution of the total input chromatin. . Chromatin samples were analysed by qPCR using -4005 / -3931, -2290 / -2185, -935 / -864, and -21/+44 CDKN1A promoter specific primers and SAT2 heterochromatin specific primer. Data were averaged from three PCR replicates, ± standard error.

A


AP-2α
NFkB2
p53
p53
SMADs
MYC
SP1
TATA
CDKN1A
AP-2α
STATs
-4000
-3000
-2000
-1000
+1
-4005/-3931
-2290/-2185
-935/-864
-21/+44

B


1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Enrichment Relative to Input
AP-2γ
p300
Ac H4
-4005 / -3931
-2290 / -2185
-935 / -864
-21 / +44
SAT2
Satellite Repeat
CDKN1A Promoter Region

5.8. ChIP reveals a role for AP-2γ in the direct repression of CDKN1A transcription.

In order to explore whether AP-2γ silencing in MCF-7 cells was accompanied by changes in AP-2γ and p300 occupancy or H4 acetylation across the -21 / +44 region, ChIP assays were conducted in cells transfected with AP-2γ targeting siRNAs. Figure 5.11 (A) shows the relative enrichment of the -21 / +44 *CDKN1A* promoter specific sequence measured in AP-2γ, p300 and acH4 ChIP elutes in cells transfected with a non silencing control siRNA or the AP-2γ targeting siRNA_3. The level of the p21 mRNA and protein up-regulation was also assessed 72 hours following AP-2γ knockdown and is shown in Figure 5.11 (B). The observed reduction in AP-2γ promoter occupancy in Figure 5.11 (A) is consistent with the reduction in total AP-2γ protein levels determined by western (Figure 5.11 (B)) following AP-2γ silencing. This suggests that chromatin-associated AP-2γ is degraded at a similar rate to total AP-2γ in silenced MCF-7 cells.

It is intriguing that the decrease in AP-2γ occupancy is associated with a slight decrease in p300 occupancy. Strikingly, a two fold increase in H4 acetylation accompanies the observed decrease in AP-2γ occupancy. This correlates with activation of transcription from the *CDKN1A* promoter as measured by an increase in *CDKN1A* mRNA levels Figure 5.11 (B). Together these data suggest that a reduction in AP-2γ occupancy at the *CDKN1A* promoter is associated with *CDKN1A* gene activation and highlights the role of AP-2γ in the repression of transcription from this promoter. In appears counterintuitive that p300 levels should fall when H4 acetylation increases across this region of the CDKN1A promoter, however loss of AP-2γ (and p300) may be accompanied by recruitment of alternative HATs, or modification of HATs at this site.

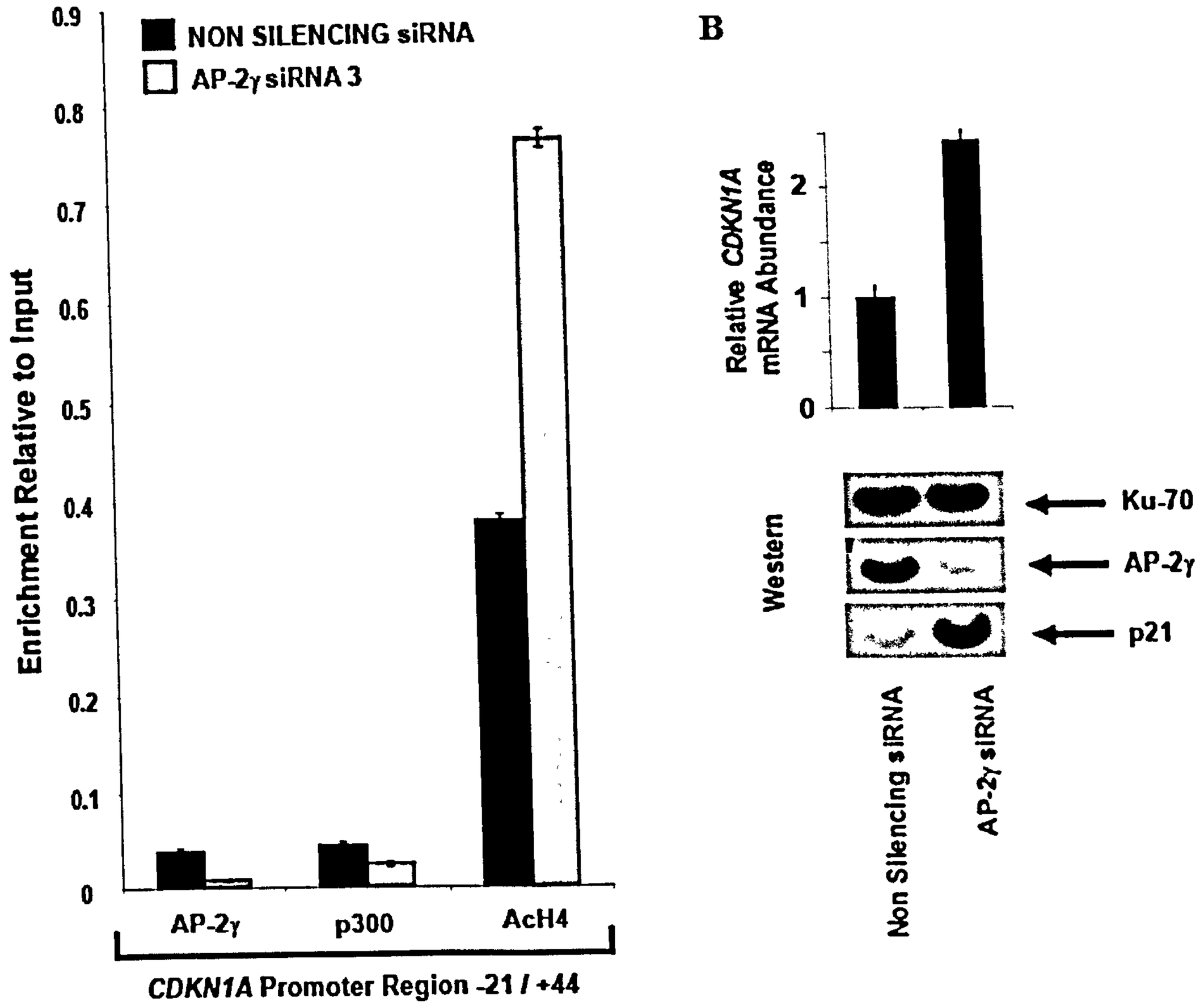


Figure 5.11 ChIP assay showing the effect of AP-2γ knockdown on endogenous AP-2γ and p300 occupancy and acH4 levels at the -21/+44 *CDKN1A* promoter region. MCF-7 cells were transfected with a non silencing control siRNA or AP-2γ siRNA_3, split 1 in 2 at 48 h, then harvested for use in ChIP, qPCR or western blot 24 h later. **(A)** ChIP assay for AP-2γ, p300 and AcH4. Results were normalised to a no antibody control and expressed relative to a 1/1000 dilution of the total input chromatin. Chromatin samples were analysed by qPCR using the -21/+44 primers, data averaged from three PCR replicates, ± standard error. **(B)** p21 mRNA and protein levels 72 hours following AP-2γ silencing in MCF-7 cells. *CDKN1A* levels were analysed by qPCR and normalised to *GAPDH* mRNA levels. The transfection reagent alone (control) value was set at 1 to generate relative mRNA levels. The graph represents data averaged from three PCR replicates, ± standard error. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, p21, and AP-2γ.

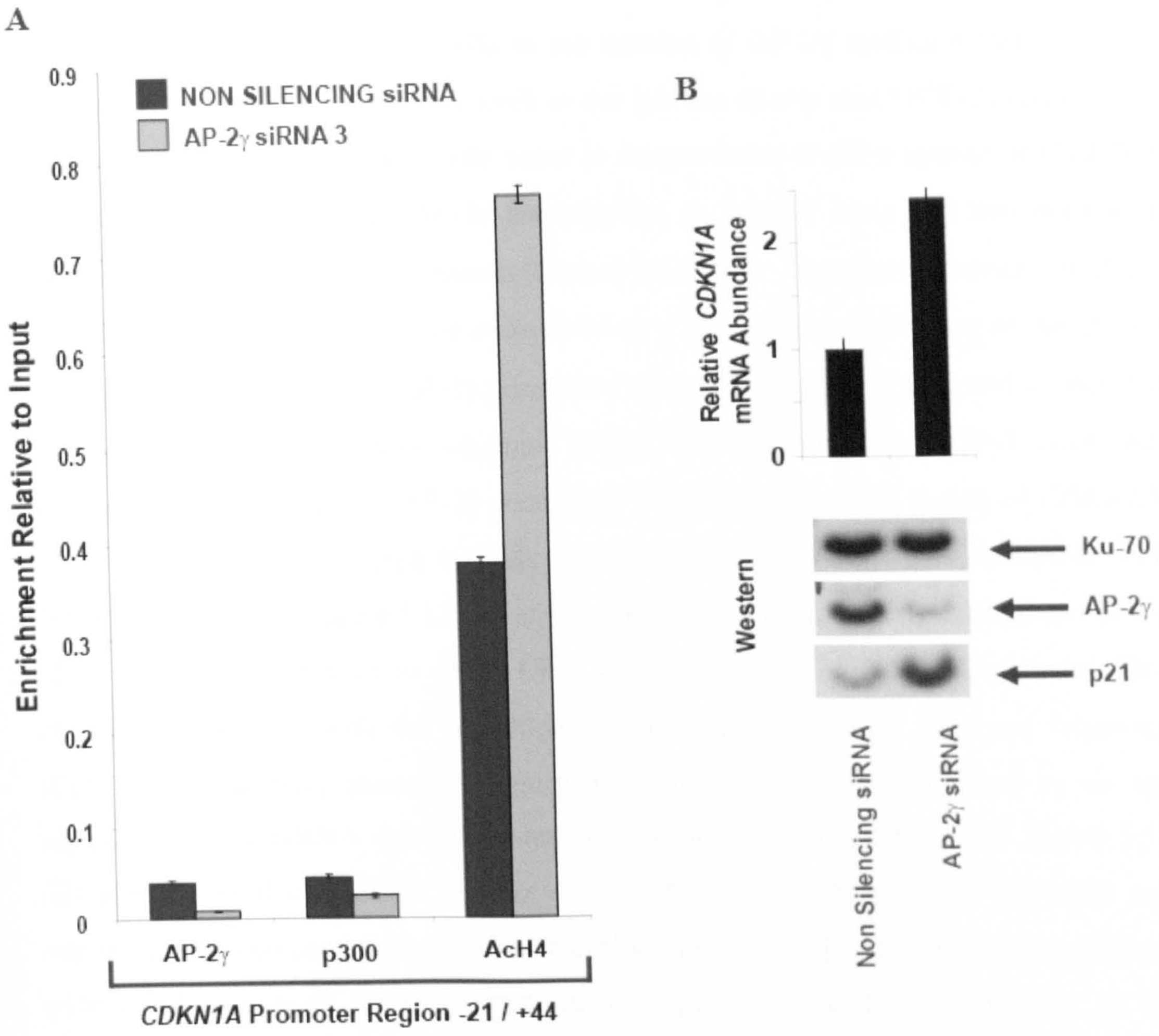


Figure 5.11 ChIP assay showing the effect of AP-2γ knockdown on endogenous AP-2γ and p300 occupancy and acH4 levels at the -21/+44 *CDKN1A* promoter region. MCF-7 cells were transfected with a non silencing control siRNA or AP-2γ siRNA_3, split 1 in 2 at 48 h, then harvested for use in ChIP, qPCR or western blot 24 h later. **(A)** ChIP assay for AP-2γ, p300 and AcH4. Results were normalised to a no antibody control and expressed relative to a 1/1000 dilution of the total input chromatin. Chromatin samples were analysed by qPCR using the -21/+44 primers, data averaged from three PCR replicates, ± standard error. **(B)** p21 mRNA and protein levels 72 hours following AP-2γ silencing in MCF-7 cells. *CDKN1A* levels were analysed by qPCR and normalised to *GAPDH* mRNA levels. The transfection reagent alone (control) value was set at 1 to generate relative mRNA levels. The graph represents data averaged from three PCR replicates, ± standard error. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, p21, and AP-2γ.

Considering these findings it was felt that it was important to assess the role of anti-oestrogen treatment in MCF-7 cells in the context of AP-2γ mediated transcriptional repression of *CDKN1A*. Previous work in our lab has shown that MCF-7 cells respond to oestrogen and undergo cell cycle arrest in the presence of ERα agonist ICI182,780. Additionally ERα regulates AP-2γ transcription in MCF-7 cells and treatment with ICI182,780 reduces AP-2γ expression (Orso *et al.*, 2004). Therefore treatment of MCF-7 cells with ICI182,780 might be considered as a drug-induced silencing of AP-2γ and should also result in loss of AP-2γ occupancy from the *CDKN1A* locus and an increase in p21 levels. To explore whether ICI182,780 treatment in MCF-7 cells was accompanied by changes in AP-2γ occupancy across the -21 / +44 region of *CDKN1A*, ChIP assays were conducted in cells harvested following 24 hours treatment with 100nM ICI182,780. Figure 5.12 (A) shows a reduction in the relative enrichment of the -21 / +44 *CDKN1A* region in AP-2γ ChIP elutes following ICI182,780 treatment. This change is consistent with the reduction in AP-2γ protein levels observed following ICI182,780 treatment, shown in Figure 5.12 (B) and was accompanied by an up-regulation of p21 mRNA and protein levels also shown in Figure 5.12 (B). Figure 5.12 (B) also shows that levels of AP-2α do not change after ICI182,780 treatment and importantly a decrease in phosphorylation of RB is observed confirming these cells are undergoing cell cycle arrest. These data suggest that AP-2γ is important in the transcriptional modulation of *CDKN1A* expression in response to anti-oestrogen treatment.

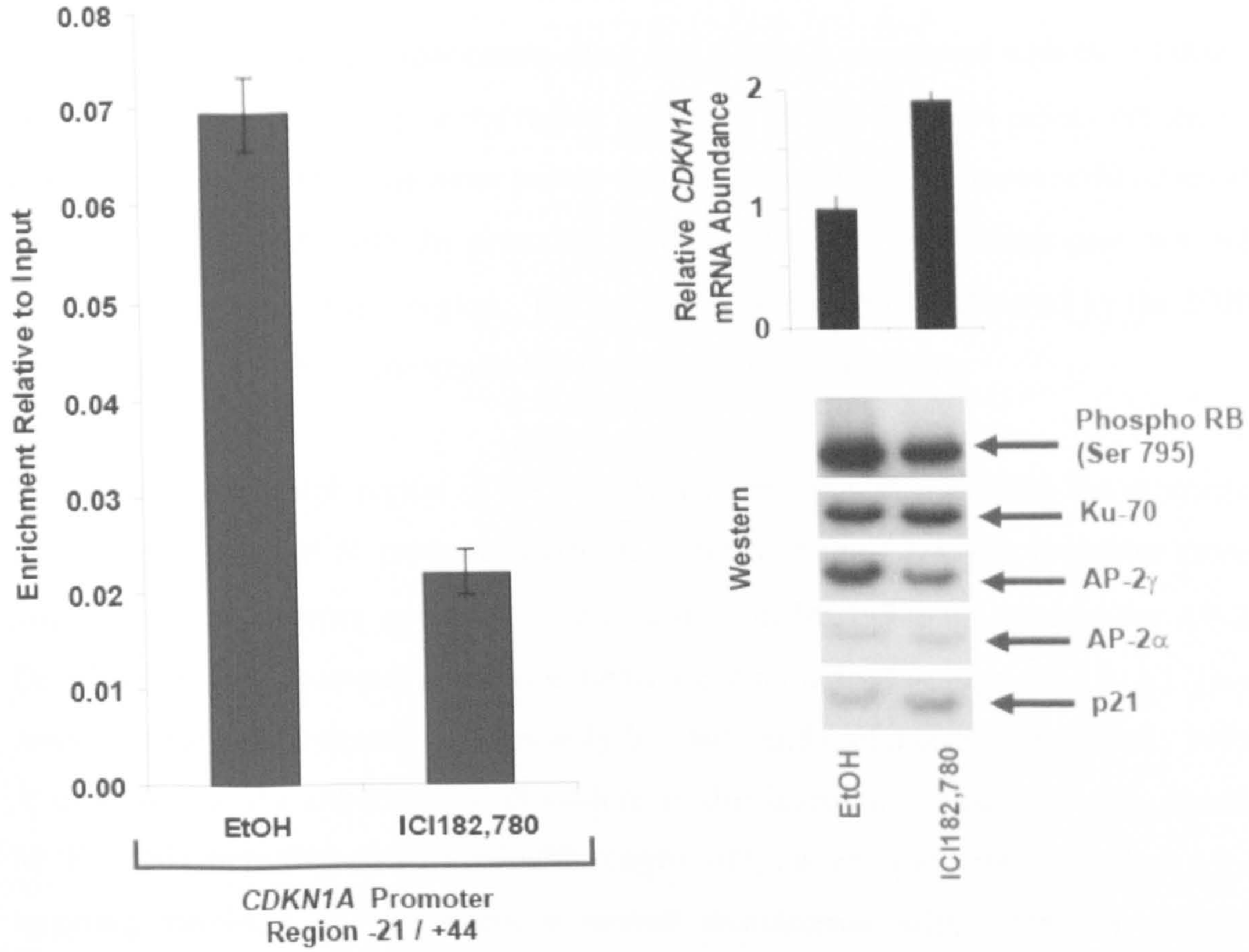


Figure 5.12 ChIP assay showing the effect of ICI182,780 anti-oestrogen treatment on endogenous AP-2γ occupancy at the -21/+44 *CDKN1A* promoter region. MCF-7 cells were split one in two into ICI182,780 (100nM) containing medium and harvested for use in ChIP, qPCR or western blot 24 h later. **(A)** ChIP assay for AP-2γ, p300 and AcH4. Results were normalised to a no antibody control and expressed relative to a 1/1000 dilution of the total input chromatin. Chromatin samples were analysed by qPCR using the -21/+44 primers, data averaged from three PCR replicates, ± standard error. **(B)** p21 mRNA and protein levels following ICI182,780 anti-oestrogen treatment in MCF-7 cells. *CDKN1A* levels were analysed by qPCR and normalised to *GAPDH* mRNA levels. The transfection reagent alone (control) value was set at 1 to generate relative mRNA levels. The graph represents data averaged from three PCR replicates, ± standard error. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, p21, and AP-2γ, AP-2α and phospho Rb (ser 795).

5.9. Reporter assays identify the CDKN1A proximal promoter is important for regulation by AP-2γ.

The data from these ChIP experiments show that AP-2γ is associated with the *CDKN1A* endogenous promoter only near the region amplified by the -21 / +44 primer set and not near regions amplified by the other primer sets further upstream. However this does not imply AP-2γ interacts with the promoter between the -21 / +44 primer pair, but only within the proximity of this region. The resolution of the ChIP is dictated by the 500bp average size at which the chromatin has been sheared by sonication.

To narrow down which region of the *CDKN1A* promoter was mediating the repression by AP-2γ, a series of 5' promoter deletion mutants of the *CDKN1A* promoter cloned into a luciferase reporter construct were evaluated in MCF-7 cells silenced for AP-2γ. Details of these 5' promoter deletion mutants are summarised in Figure 5.13 (A). These constructs have been described previously by Datto and colleagues (Datto *et al.*, 1995). A time line of the experimental procedure is displayed in Figure 5.13 (C). Briefly, MCF-7 cells were first transfected with reagent only, a non-silencing control or AP-2γ targeting siRNA_2. At 24 hours a second transfection with *CDKN1A* promoter luciferase reporter constructs was conducted, at 48 hours cells were passaged 1 in 2 and finally at 72 hours luciferase activity was assayed. The data in Figure 5.13 (B) shows that following AP-2γ silencing in MCF-7 cells an increase in luciferase activity is observed for all of the 5' promoter deletion mutants. These data suggest that the region from –111 to +8 relative to the start of transcription of the *CDKN1A* gene is sufficient for repression by AP-2γ in the MCF-7 cells transfected with reagent only or the non silencing control siRNA and this repression can be relieved upon AP-2γ silencing. All of the 5' promoter deletion mutants of the *CDKN1A* showed a similar pattern of activity to the full length reporter, although the absolute level of activity was lower for the shorter constructs. It has been suggested by others (Maclaren *et al.*, 2003) that this may be due to the fact the 5' promoter deletion mutants all lack the p53 binding region of the of the promoter and this region may account for a large component of promoter activity. However, some of the effect could also be due to differences in the constructs used as the 5' promoter deletion mutants are in a pGL2 (Promega) background where as the -

2325\+8 *CDKN1A* luciferase reporter construct was created by cloning the Firefly Luciferase gene *CDKN1A* promoter into a pBluescript background (pWWP-luc, el-Deiry *et al.*, 1993).

In an effort to narrow down which region of the *CDKN1A* promoter was mediating the repression by AP-2γ further, an additional series of minimal –74/+20 *CDKN1A* luciferase reporter constructs was evaluated, consisting of 10 base pair substitutions across the SP1 sites and flanking sequences. These constructs have been described previously by Datto and colleagues (Datto *et al.*, 1995). Transfections were performed in an identical manner as those described in Figure 5.13 and a time line of the experimental procedure is displayed in Figure 5.14 (C).

The data in Figure 5.14 (B) show that following AP-2γ silencing in MCF-7 cells an increase in luciferase activity is observed for the WT –74/+20 promoter. These data suggest that the region from –74 to +20 relative to the start of transcription of the *CDKN1A* gene is sufficient for repression by AP-2γ in the MCF-7 cells transfected with reagent only or the non silencing control siRNA and this repression can be relieved upon AP-2γ silencing. All of the substitution mutants of the –74/+20 *CDKN1A* promoter showed a similar pattern of activity to the full length reporter, although the absolute level of activity varied. Constructs Mut#2, Mut#3, Mut#4, Mut#5 and Mut#6 show reduced basal levels of luciferase activity in the control experiments compared to the WT –74/+20 promoter control experiments. Upon AP-2γ silencing, the luciferase activity was also lower for the mutants than the WT –74/+20 promoter experiment, although the fold increase in activation remained fairly similar. This suggests the sequences substituted in these five mutants are important for activation of transcription from this promoter. Indeed, these substitutions all affect known SP1 binding sites that are required for this activity (Datto *et al.*, 1995). Substitution mutant Mut#1 showed an increase in basal levels of luciferase activity in the control experiments compared to the WT –74/+20 promoter control experiments, although the magnitude of activation upon AP-2γ silencing, again, remained fairly similar.

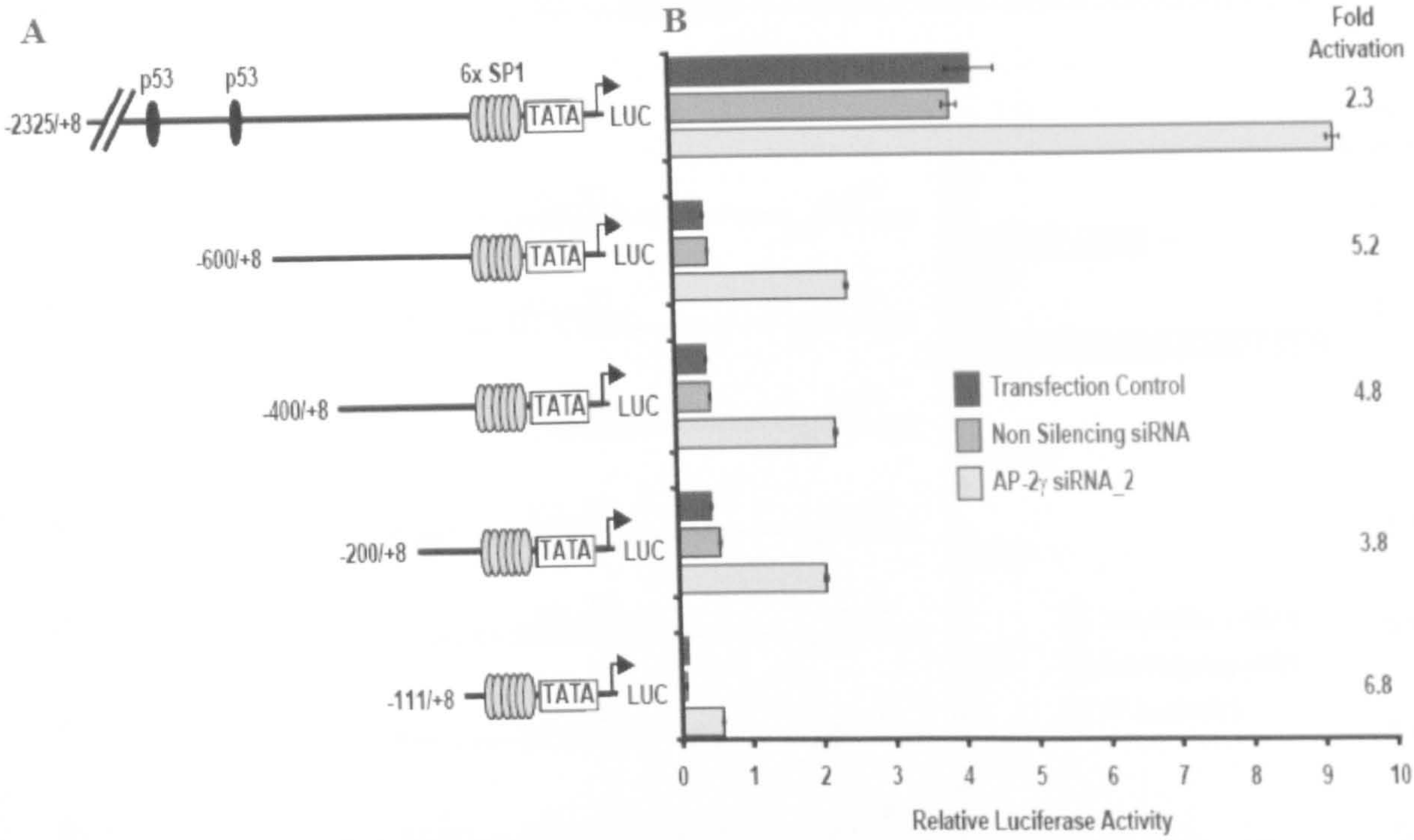


Figure 5.13. MCF-7 cells transfected with reagent only or non-silencing control siRNAs were able to repress transcription from a minimal –111/+8 *CDKN1A* luciferase reporter construct and this repression was relieved upon AP-2γ silencing. MCF-7 cells were first transfected with reagent only, a non-silencing control or AP-2γ targeting siRNA_2, at 24 hours a second transfection with the indicated CDKN1A promoter luciferase reporter constructs was conducted **(A)**, at 48 hours cells were passaged 1 in 2 and finally at 72 hours luciferase activity was assayed **(B)**. A renilla luciferase construct (phRG-TK) was co-transfected in each experiment. Results (mean ± S.E. of three independent experiments) are presented as relative luciferase activity, corrected for Renilla luciferase activity. A time line of the experimental procedure is displayed **(C)**.

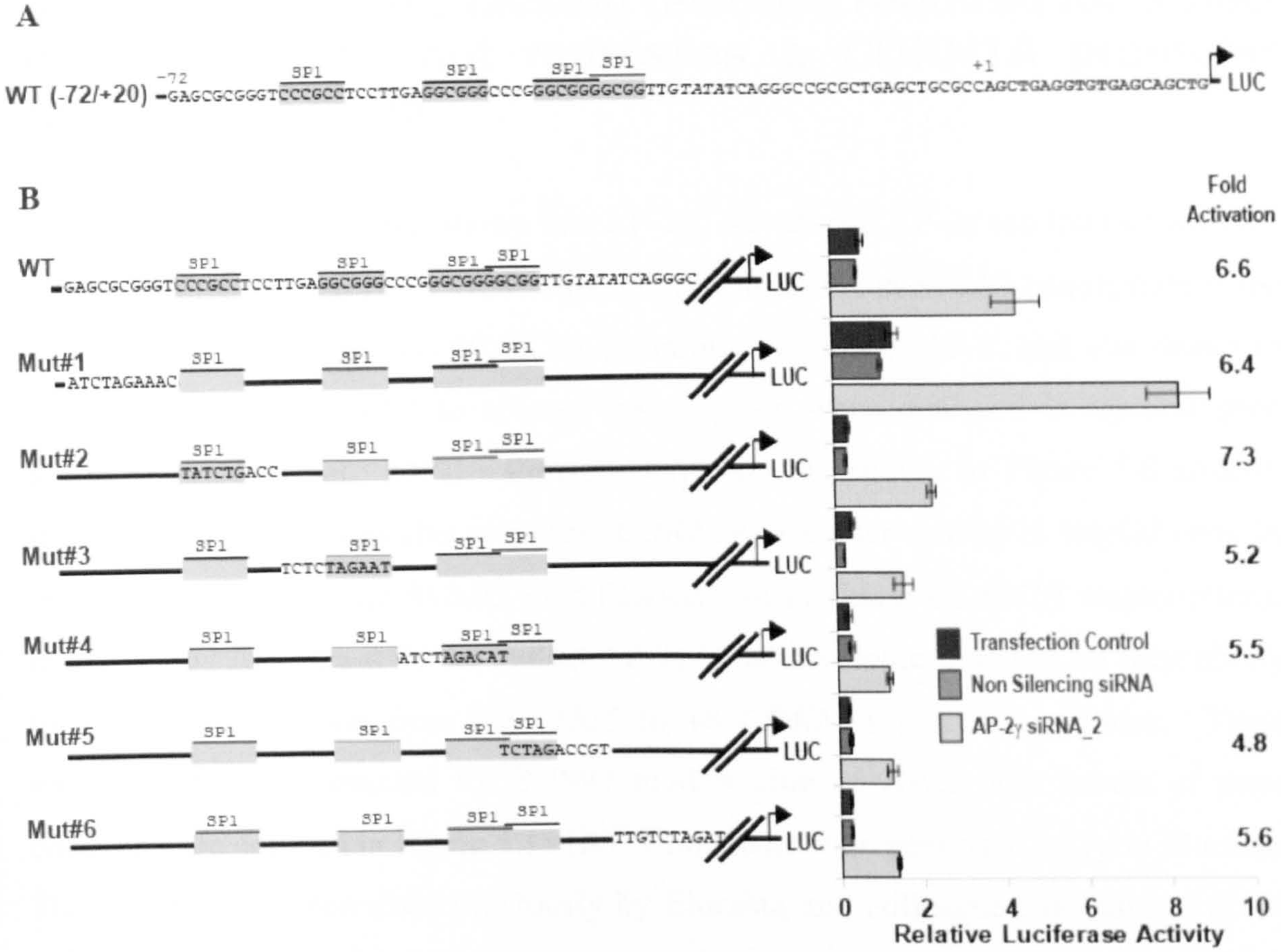


Figure 5.14. Transactivation of the –74/+20 *CDKN1A* luciferase reporter constructs following AP-2γ silencing in MCF-7 cells. MCF-7 cells were first transfected with reagent only, a non-silencing control or AP-2γ targeting siRNA_2, at 24 hours a second transfection with the indicated –74/+20 CDKN1A promoter luciferase reporter construct was conducted, at 48 hours cells were passaged 1 in 2 and finally at 72 hours luciferase activity was assayed **(B)**. A renilla luciferase construct (phRG-TK) was co-transfected in each experiment. Results (mean ± S.E. of three independent experiments) are presented as relative luciferase units (RLU), corrected for Renilla luciferase activity. The sequence of the –74/+20 *CDKN1A* luciferase reporter is displayed in **(A).** A time line of the experimental procedure is displayed **(C)**.

5.10. AP-2γ mutants lacking residues required for SUMO modification do not represses a CDKN1A promoter reporter construct.

Previous work in our lab has shown that AP-2α, AP-2β and AP-2γ can interact with the SUMO conjugating enzyme UBC9. This interaction results in SUMO conjugation at the conserved lysine residue (lys 10) in the N-terminal region of AP-2, and was shown to reduce the ability of AP-2 to activate transcription when analysed in reporter gene assays (Eloranta *et al.*, 2002). Previous experiments, shown in Figure 5.8 (A&B), suggested that AP-2γ was able to repress *CDKN1A* promoter activity in HepG2 cells. In order to explore whether SUMO modification can facilitate *CDKN1A* transcriptional repression by AP-2γ, a series of mutant AP-2γ constructs were assessed for their ability to repress transcription from the –2325 to +8 *CDKN1A* promoter reporter. These mutants effect the potential for SUMO modification of AP-2γ and details of these constructs are outlined in Figure 5.15 (A). Constructs were generated by Jyrki Eloranta. The K10R-AP-2γ, described previously by Eloranta and colleagues, is mutated at the target lysine residue for SUMO modification and is unable to be SUMO modified (Eloranta *et al.*, 2002). The Flag-ΔN10-AP-2γ substitutes a Flag tag for the first 10 N-terminal amino acids. It is therefore also missing the target lysine residue for SUMO modification and is unable to be SUMO modified (Eloranta *et al.*, 2002). In order to assess the affect of artificial SUMO conjugation on AP-2γ activity, two forms of AP-2γ were generated that have the 10 N-terminal amino acids replaced with a consensus SUMO peptide corresponding to either SUMO1 or SUMO3. Only the effect of SUMO3 was assessed as SUMO2 and SUMO3 are almost identical and have been shown to function in the same manner (Reviewed by Hay, 2005).

Figure 5.15 (B) shows that basal luciferase activity from this construct was reduced following co-transfection with a wild-type AP-2γ expression construct, reproducing previous results in Figure 5.8 (B). Co-transfection of the *CDKN1A* promoter reporter with a wild-type AP-2α expression construct was accompanied by an increase in luciferase activity. This is in support of data from others, who have shown AP-2α can activate *CDKN1A* transcription (Zeng *et al.*, 1997; McPherson *et al.*, 2002).

Interestingly, co-transfection of the *CDKN1A* promoter reporter with the K10R-AP-2γ mutant construct was also accompanied by an increase in luciferase activity. This suggests that SUMO modification at lysine 10 of AP-2γ is important for mediating its transcriptional repression at the *CDKN1A* promoter. Indeed this mutation appears to modulate the activity of AP-2γ within this assay from repressor to activator. Co-transfection of the *CDKN1A* promoter reporter with the Flag-ΔN10-AP-2γ mutant construct resulted in luciferase levels similar to those observed in cells transfected with reporter only. Again this suggests residues within the substituted region of AP-2γ, a region including the SUMO modification target lysine, are important for mediating its transcriptional repression at the *CDKN1A* promoter. Interestingly, in contrast to reporter co-transfections with the K10R-AP-2γ, no additional increase in *CDKN1A* promoter reporter activity above basal levels was observed following co-transfection with the Flag-ΔN10-AP-2γ mutant. This suggests that residues required for this additional activity are absent from this construct. Both of the constitutively SUMO modified forms of AP-2γ (SUMO1-ΔN10-AP-2γ and SUMO3-ΔN10-AP-2γ) reduce luciferase activity over and above that of wild-type AP-2γ when co-transfected with the *CDKN1A* promoter reporter. This shows that artificial conjugation of SUMO to AP-2γ can repress *CDKN1A* promoter activity in HepG2 cells.

Together these transient transfection studies indicate that SUMO modification of AP-2γ mediates its ability to repress *CDKN1A* promoter activity in HepG2 cells. However, the observed effects on luciferase activity are not necessarily due to direct effects of AP-2γ and its mutant forms on the *CDKN1A* promoter, and they could be due to indirect regulation via differential sequestration of factors away from the reporter construct by AP-2γ or its mutants.

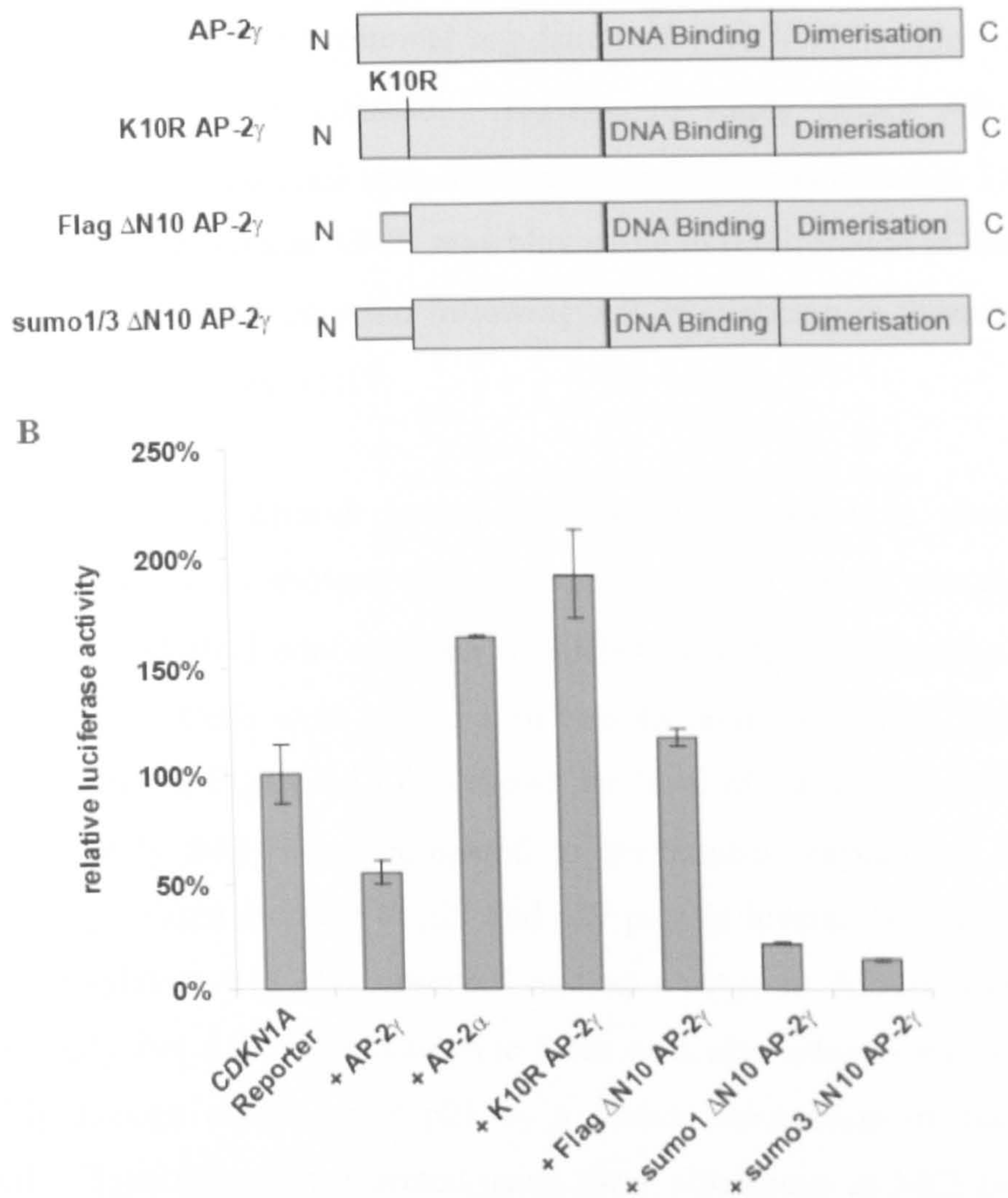


Figure 5.16. The effect of mutant forms of AP-2γ on basal activity from the *CDKN1A* promoter in HepG2 cells. **(A)** Shows mutant forms of AP-2γ in the pcDNA3 vector. K10R AP-2γ and FLAG-ΔN10 AP-2γ are resistant to SUMO modification and SUMO1 and SUMO2 ΔN10 AP-2γ constructs are constitutively SUMO modified. **(B)** The AP-2 non-expressing cell line, HepG2 was transiently transfected with 1.0μg of the -2325\+8 *CDKN1A* luciferase reporter construct and 0.5μg of the indicated AP-2γ expression construct or an AP-2α expression construct. A renilla luciferase construct (phRG-TK) was co-transfected in each experiment and the differences in DNA amount were compensated for using a pcDNA3 plasmid. Forty-eight hours following transfection cells were assayed for luciferase activity. Results (mean ± S.D. of three independent experiments) are presented as relative luciferase activity, corrected for Renilla luciferase activity. Cells transfected with the *CDKN1A* promoter luciferase reporter alone were set at 100%.

5.11. AP-2 silencing in MCF10A and ZR-75-1 cells.

A role for AP-2γ in the transcriptional regulation of *CDKN1A* in other breast cancer derived cell lines was also assessed. Attempts to stably reduce AP-2γ levels in MCF10A and ZR-75-1 cell lines were also unsuccessful (summarised in Table 3.2) and therefore it was postulated that AP-2γ may play a role in proliferation in these cell lines. Levels of p21 protein were assessed following AP-2γ silencing in these cell lines and the results are shown if Figure 5.17.

MCF10A cells proved difficult to transfect when compared to other cell lines. Optimisation experiments showed (not shown) that some AP-2γ silencing could be achieved using 5nM final concentration of siRNA and 8µl of Interferin transfection reagent (PolyPlus). Cells were split one in two 48 hours following transfection and harvested at 72 hours. Figure 5.17 (A) shows the level of AP-2γ knockdown achieved were approximately 50% when compared to the control experiment and this was accompanied by a slight increase in p21 and p27 protein levels. A moderate reduction in Rb phosphorylation was also observed and no change in AP-2α levels was seen. These data imply that AP-2γ knockdown in these cells also effects cell cycle regulation and possibly through regulation of p21 by a similar mechanism to that observed in MCF-7 cells. Transfections performed using these conditions in MCF10A cells were also assessed using a 3' Alexa Fluor 488 labelled non-silencing control siRNA and observed via fluorescence microscopy. Fluorescence was reproducibly detected in 50-60% of transfected cells reflecting the level of AP-2γ knockdown observed following transfection with the AP-2γ targeting siRNAs.

ZR-75-1 cells were transfected in an identical manner to MCF-7 cells. Figure 5.17 (B) shows the level of AP-2γ knockdown achieved were approximately 90% when compared to the control experiment and this was accompanied by a slight increase in p21 but not p27 protein levels. No change in Rb phosphorylation or AP-2α levels was observed. These data imply that AP-2γ knockdown in ZR-75-1 cells may also effect p21 regulation but downstream effects on Rb phosphorylation are not apparent.

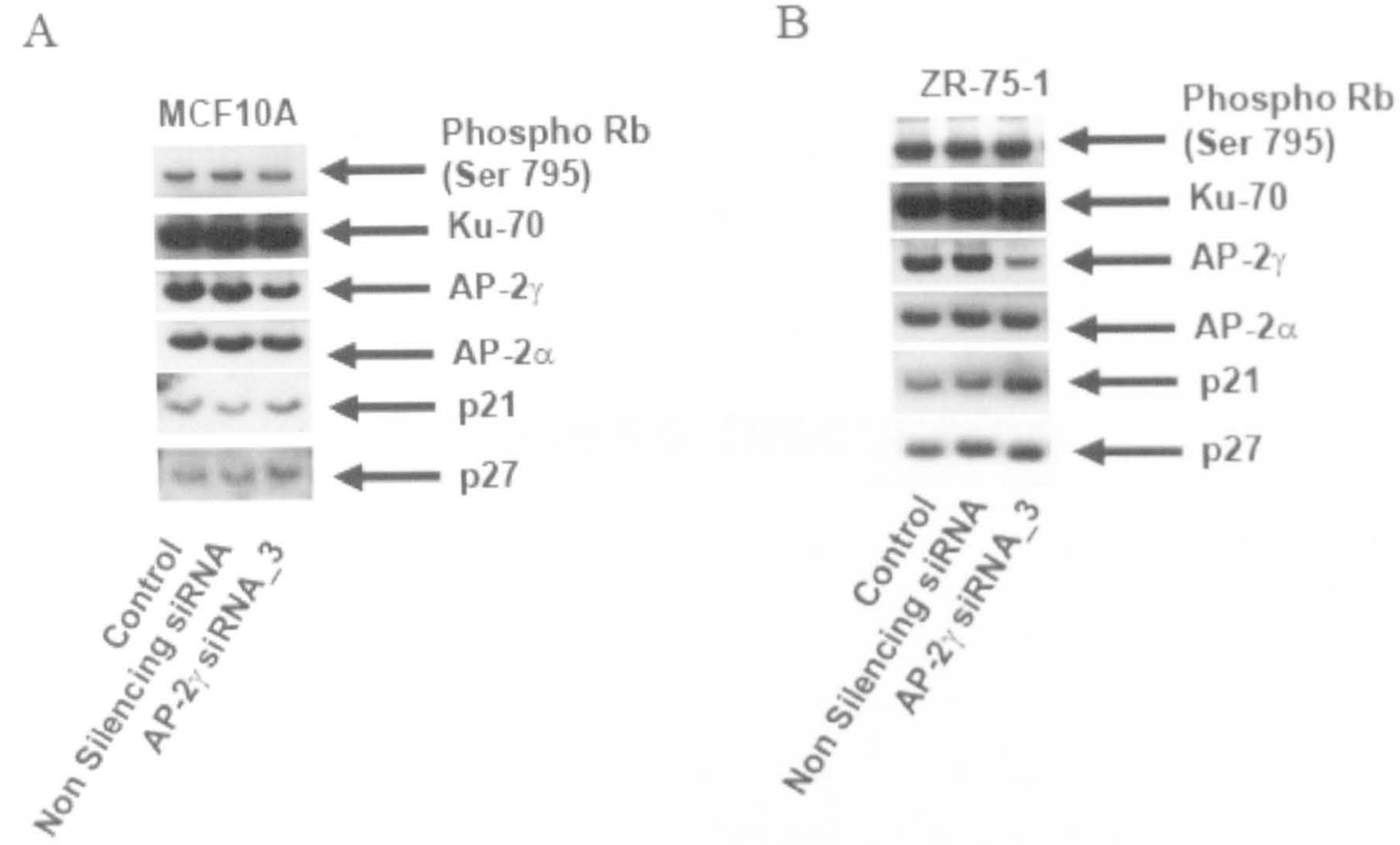


Figure 5.17. Changes in cell cycle proteins following silencing of AP-2γ in MCF10A (A) and ZR-75-1 (B) cells. Cells were transfected as indicated, passaged at the 48-hour time point following transfection and harvested at 72 hours. Whole Cell extracts (5μg/lane) were separated by SDS/PAGE and blotted to a membrane. Blots were probed with primary antibodies against Ku-70, AP-2γ, AP-2α, p21, p27 and Phospho RB (Ser 795). Blots were striped and reprobed where necessary.

CHAPTER 6: DISCUSSION

High levels of the developmentally regulated transcription factor, AP-2γ have been correlated to poor survival in breast cancer. This correlation occurs independently and in the context of the generally considered indicators of good prognosis: namely ERα expression and ErbB2 negativity. Interestingly, AP-2γ has been previously described as a ligand activated target gene of ERα. AP-2γ expression in these tumours therefore, might represent an opposition to anti-oestrogen therapy. This study aimed to understand which pathways are activated in AP-2γ expressing breast cancer cells and how they might contribute to poor patient prognosis.

Several lines of evidence based on the expression of AP-2 family members during development suggest that expression of these proteins might act at a checkpoint between proliferation and differentiation.

1. In regenerating skin epidermis AP-2α expression is required to inhibit proliferation via *EGFR* transcriptional repression as cells exit the basal layer. Without AP-2α, subrabasal cells fail to differentiate and exhibit hyperproliferation (Wang *et al.*, 2006).
2. In gonocyte development AP-2γ expression is high in early proliferating germ cells and must be lost in order for the cells to differentiate. Forced AP-2γ expression in these cells results in hyperproliferation and prevents differentiation (Jager *et al.*, 2003; Pauls *et al* 2005).
3. In the developing mammary gland, targeted over-expression of AP-2α results in inhibition of epithelial proliferation in developing alveolar buds but differentiation is not impaired (Zhang *et al.*, 2003). However, over-expression of AP-2γ results in a hyperproliferation and a failure of the alveolar buds to differentiate into their lactating derivatives (Jager *et al.*, 2003).
4. In early embryogenesis, AP-2γ expression is required for the proliferation of cells derived from the trophoectoderm (Auman *et al.*, 2002; Werling & Schorle, 2002).

Overall this evidence suggests that AP-2α expression is required for inhibition of proliferation. Indeed, previous studies have correlated high AP-2α expression with reduced proliferation and good prognosis in breast cancer patients, whereas loss or reduction of AP-2α expression has been associated with disease progression.

Conversely, AP-2γ expression appears to be required for proliferation to occur. This coordinated control of proliferation and differentiation by AP-2 family members is desirable during development and organogenesis, but aberrant expression of these proteins might lead to hyperplasia, a proliferative and undifferentiated state that could contribute to the promotion of cancer growth in adults.

When considered together the *in vitro* studies do not always provide a clear picture of the molecular mechanisms underlying a role for AP-2 proteins in breast cancer. In general, evidence exists in support of a tumour suppressor-like role for AP-2α. AP-2α can facilitate inhibition of proliferation by the transcriptional repression *EGFR*, in the case of skin epidermis, and the activation of *CDKN1A* transcription and repression of *BCL2* transcription by AP-2α has also been shown in various cells lines. The molecular mechanisms behind the association of high AP-2γ expression and increased breast tumour progression leading to poor patient prognosis remain elusive, as does the transcriptional program initiated by AP-2γ in the promotion of cell growth and inhibition of differentiation in breast tissues during development. AP-2γ expression is clearly associated with the regulation of *ERBB2* and ERα in breast cancer cell lines, however high levels of AP-2γ predict poor prognosis independently of these factors in patients with advanced primary invasive breast carcinomas, suggesting other molecular mechanisms may be involved.

In order to try to define the additional pathways activated by AP-2γ specifically, in breast cancer, the research presented in this thesis aimed to use Affymetrix GeneChip high-density oligonucleotide microarrays to define AP-2γ dependant changes in gene expression in breast epithelial cells.

6.1. A Role for AP-2γ in the control of proliferation in MCF-7 cells.

Although the primary aim of this study was to generate an AP-2γ dependant gene expression profile following RNAi mediated AP-2γ silencing in MCF-7 cells, it soon became apparent that depleting AP-2γ levels effected the growth of these cells.

Initial experiments used shRNA constructs to try to establish derivatives of the MCF-7 cells with stably reduced AP-2γ levels. Intriguingly, when carrying AP-2γ knockdown constructs MCF-7 cells produced reduced colony numbers after antibiotic selection and these colonies subsequently failed to thrive (Figure 3.7.). These results implied that lack of AP-2γ expression in these lines perturbs their normal growth. With the aim of generating material for gene expression profiling and to further analyse potential abnormalities in proliferation, transient transfection conditions for lipofection-mediated delivery of siRNA were carefully optimised with respect to off-target effects (see Figures 3.11 and 3.12 for details).

In early experiments, shown in Figure 5.1, a reduction in the number of cells was observed following transient AP-2γ silencing. These early effects on cell proliferation appeared to be independent of cell death, as neither the number of floating cells, nor the SubG1 population were increased in AP-2γ siRNA transfected cells. However, an increase in the number of cells with a G1 DNA content was observed and accompanied by a decrease in the proportion of cells in S phase. A more detailed cell cycle analysis, shown in Figure 5.2, was performed with the three independent AP-2γ targeting siRNAs compared to control conditions and confirmed a significant increase in G1 population ($p<0.0.1$) and a moderate but significant decrease in the proportion of cells in S phase ($p<0.0.1$) in MCF-7 cells depleted for AP-2γ. A significant reduction in the fraction of cells actively undergoing S phase was also seen in cells transfected with AP-2γ targeting siRNAs compared to the control populations. Together these data suggest that cells transfected with AP-2γ targeting siRNAs were undergoing a partial cell cycle arrest at the G1/S transition.

These results provide compelling evidence that AP-2γ expression is required in MCF-7 cells in order for them to proliferate and that siRNA mediated silencing of AP-2γ in these cells inhibits their growth by inducing cell cycle arrest at the G1/S transition.

6.2. AP-2γ dependant mechanisms for the control of proliferation in MCF-7 cells.

6.2.1. AP-2γ silencing in MCF-7 cells results in a cell cycle related gene expression profile.

Amongst the transcripts significantly regulated following AP-2γ silencing in MCF-7 cells were a subset of transcripts related to cell cycle. Using two different publicly available GO databases, the transcripts from the AP-2γ dataset were assigned GO or Panther biological process terms and then organised into related groups (Figure 4.9 and 4.10). In particular, both of these methods identified biological process terms highly significantly over-represented within the down-regulated genes in categories related to cell cycle (Tables 4.5 and 4.6).

A more detailed analysis of this subset of cell cycle regulated genes showed the majority of these are known to display cell cycle periodicity (Whitfield *et al.*, 2002), almost all showing their lowest expression at G1 or S phase, again consistent with a perturbation at the G1/S transition (Figure 5.2). It is reasonable to consider that many of these changes may have occurred as a consequence of the cell cycle arrest and therefore may not be direct targets of AP-2γ. Consistent with this hypothesis was that many of these genes are known to be transcriptionally regulated by E2F family members at the G1/S transition (Figure 4.11) and therefore AP-2γ may have a role in the regulation of genes upstream of this protein. The analysis of conserved TFBS within this data set supports this hypothesis. The most represented TFBS <2Kb upstream of transcripts within the AP-2γ dataset, were those for cell cycle related transcription factor families E2F and FOXO (Table 4.7).

An additional TFBS analysis, focusing on the AP-2 consensus-binding site upstream of the genes significantly regulated following AP-2γ silencing, was performed in order to discriminate between potential direct transcriptional targets for AP-2γ and indirect targets. No enrichment for the presence of the AP-2 consensus sequence was observed in promoter regions of genes within the AP-2γ data set compared to a similar region upstream of a set of 21787 genes (Table 4.8). The analysis highlighted how common the AP-2 consensus sequence is in upstream sequences of genes, and supports other studies which showed similar findings looking at a smaller set of promoters (Bajic *et al.*, 2004). Such computational genomics approaches to identify DNA regulatory elements should be interpreted with caution, especially when considering degenerate sequences that are common in GC-rich regions such as the AP-2 TFBS. It is likely, that constraints other than DNA sequence alone must confer the efficacy of these sites *in vivo*. In general the specific chromatin architecture that regulates the accessibility of TFBS within promoters or enhancers has been shown to be important (review in Li *et al.*, 2007). Also, interactions with other sequence-specific transcription factors binding within the vicinity of AP-2 binding sites may be important.

Shortly after establishing the findings of this study, Jin and colleagues addressed the later question for E2F1 using a more complex computational approach (Jin *et al.*, 2006). Of particular interest to this study was their analysis of E2F1 and AP-2 TFBS co-localisation. Using a small set of known E2F1 target promoters, the authors found AP-2 TFBS within the vicinity of 55% of functional E2F1 sites, compared to E2F1 sites in non-E2F target promoters. However the next most common association, E2F1 and NFAT TFBSs, was only enriched in 10% of functional E2F1 sites compared to the control promoters. In light of this second association, it seems debatable whether association of E2F1 and AP-2 TFBS in likely to have a functional consequence. Both E2F and AP-2 bind GC-rich TFBS. Functional E2F binding sites might be associated with GC-rich promoter regions where there would be an increased likelihood of chance occurrence of the AP-2 TFBS. Subsequent ChIP-chip experiments were performed on E2F target promoters in order to address the functional consequence of this association *in vivo*. ChIP-chip combines Chromatin IPs with microarray analysis (chip), where DNA enriched in the IPs is labelled and hybridised to microarrays containing tiled promoter sequences. The level of hybridisation over sequences enriched by a control

antibody can be directly related to occupancy of a factor at the promoter region to which it hybridises. Interestingly, it was shown that AP-2 ChIPs were enriched for E2F target promoters containing co-localisation of E2F1 and AP-2 TFBS. These results should be cautiously interpreted. Previous work in our lab has shown that the AP-2 antibody used in this ChIP has a high level of non-specific activity as determined using primers against a heterochromatic α-satellite region (Karsten Frederich, *Personal Communication*), and therefore perhaps it would have been useful to include similar controls in this study alongside known AP-2 target promoters to a provide positive control for binding. Little regard was given to the biological relevance of these findings and unfortunately the authors used the non discriminating AP-2 antibody. Although the potential association of AP-2 family members and regulation of E2F1 target genes is interesting, in light of AP-2 mediated regulation of p21 (see below), these findings remain to be more convincingly established.

6.2.2. AP-2γ silencing in MCF-7 cells results in an up-regulation *of the cyclin dependant kinase inhibitor p21.*

In parallel to the microarray experiments, changes in cell cycle regulators known to be important at the G1 transition were observed by Western blot following AP-2γ silencing in MCF-7 cells in order to delineate a mechanism for the cell cycle arrest (Figure 5.2). The most notable was an up-regulation of the CDKI, p21, observed for each AP-2γ targeting siRNA. As explained in Section 1.5, CDKIs like p21 exert their activity by inhibiting the kinase activity of Cyclin/CDK complexes. Indeed, a reduction in Cyclin/CDK activity was observed following AP-2γ silencing, marked by a reduction in phosphorylation of both Rb and p27. Changes in the levels of Cyclin and CDK proteins have also been shown to contribute to altered activity of Cyclin/CDK complexes. For example, the levels of cyclin D and CDK4/6 are transcriptionally activated by Myc in response to growth signalling (Reviewed by Pelengaris *et al.*, 2002). However, there was little or no change in expression levels of Cyclin D1, Cyclin D3 and CDK2, CDK4, and CDK6 or indeed of Myc itself following AP-2γ silencing in MCF-7 cells (Figure 5.3 and Figure 5.7), suggesting that changes in the overall levels of these proteins were not responsible for the observed altered activity of Cyclin/CDK complexes.

In addition to an up-regulation of p21 following AP-2γ silencing in MCF-7 cells, a moderate increase in the protein levels of p27 and cyclin E was observed. The increase in p27 levels was accompanied by a decrease in p27 phosphorylation at ser147. This suggests that the increase in p27 level could be largely due to a reduction in cyclinE/CDK2 activity which would, in normal cycling cells, promote p27 proteolysis (Sheaff *et al.*, 1997). The moderate increase in levels of Cyclin E2 however, can not be so easily explained, although this difference appears exaggerated due to the loading inaccuracies. However, cycle arrest occurred in spite of the moderate increase in cyclin E levels, suggesting that cyclin E up-regulation is not sufficient to override anti-proliferative mechanisms associated with AP-2γ silencing in MCF-7 cells.

In addition to the proteins described above an increase in both CBP and p300 protein levels were observed following AP-2γ silencing in MCF-7 cells (Figure 3.13). The mechanisms behind these observations clearly require further investigation; however the observed increases are likely be related to the involvement of these proteins at the G1/S transition. As well as their characterised role in the coactivation of positive and negative cell cycle regulating transcription factors including, p53, Myc, RB and E2F, CBP/p300 also play additional less well understood roles in the regulation of the cell cycle. A recent study has shown cells negative for p300 show defects at the G1/S transition and in particular, the HAT independent activity of p300 is required for orderly G1/S transition. In early G1, p300 specifically has been shown to associate with hypo-phosphorylated RB and inhibit Cyclin/CDK6 dependant RB phosphorylation (Iyer *et al.*, 2007). However, at the G1/S boundary p300 and possibly CBP have been suggested to associate with RB and augment RB phosphorylation (Iyer *et al.*, 2007). Furthermore both p300 and CBP are phosphorylated in a cell cycle dependent manner although the functional consequences are unclear (reviewed in Goodman & Smolik, 2000). Considering these observations, is it unclear how an increase in p300 or CBP following AP-2γ silencing might contribute to the observed cell cycle arrest in MCF-7 cells particularly in the context of p21 up-regulation.

6.2.2.1. Can p21 activation alone explain the cell cycle perturbation at the G1/S transition?

Overall evidence strongly supports that the cell cycle perturbation at the G1/S transition in AP-2γ silenced cells is largely due to the up-regulation in p21 levels. Indeed a decrease in the kinase activity of Cyclin/CDK complexes, a functional consequence of p21 induction, is marked by a reduction in phosphorylation of Rb and p27. Of the other cell cycle regulators studied, changes in cyclin E cannot explain a cell cycle arrest and p27 induction could largely be explained by posttranscriptional mechanisms caused by a cell cycle arrest. We cannot rule out the involvement of additional pathways regulated by AP-2γ silencing in MCF-7 cells that were not explored in this study or indeed that AP-2γ is normally required for the transcription of E2F target genes that were down-regulated in following AP-2γ silencing (discussed above).

In support of cell cycle arrest being caused by p21 induction alone, gene expression changes observed upon AP-2γ silencing in MCF-7 cells showed remarkable overlap with other studies observing gene expression changes following exogenous p21 expression in unrelated cell lines. In particular, in human fibrosarcoma cells approximately 45% of the genes down-regulated upon p21 induction (Chang *et al.*, 2000) were also down-regulated following AP-2γ silencing in MCF-7 cells in my experiments. Additionally, in human ovarian carcinoma cells approximately 47% of genes differentially regulated upon expression of exogenous p21 (Wu *et al.*, 2002) were also significantly regulated following AP-2γ silencing in MCF-7 cells. The details of these analyses are shown in Table 6.1. Clearly a number of E2F target genes (indicated by an underline in Table 6.1) are included in these overlapping gene lists. This is consistent with the idea that a major contributing factor leading to the down-regulation of E2F targets is the up-regulation of p21 following AP-2γ silencing in MCF-7, and not the direct regulation of E2F target genes by AP-2 family members (Jin *et al.*, 2006).

	Percentage Overlap	Overlapping Gene Symbols
Microarray derived gene expression changes following adenovirus mediated p21 expression in ovarian carcinoma cells. (Wu *et al.*, 2002)	56/117 = 47%	*CDC2, CCNB1, UBEC2, DKFZp762E1312, TTK, MKI67, TPX2, RACGAP1, DLG7, PRC1, TPX2, AURKA, TUBB2A, SMC2, TOP2A, HMGB2, UBE2T, BUB1, ANLN, WDHD1, KNTC1, RRM1, MCM7, TYMS, MCM4, ATAD2, AURKB, PBK, CEP55, ASPM, ATAD2, MCM2 ,CHEK1, MCM6, KIAA1794, HMGB2, C14ORF106, CDC6, CENPF, MYBL1, CDC20, KIF20A, PKB, CKS1B, PCNA, TRIP13, ARL6IP2, KIAA0101, SMC4, BUB1, CDKN3, RFC3, SMC4, BUB3, DTL, UHRF1, MAF.*
Microarray derived gene expression changes following inducible retroviral expression of p21 in fibrosarcoma cells. (Chang *et al.*, 2000)	Down 31/70 = 45% Up 1/46 = 2%	*TMPO, FOXM1, TKK, TYMS, AURKA, AURKB, PRC1, LAMB1, CENPA, CENPF, BUB1, KIF2C, KIFC1, CDC2, CKS1B, MYBL1, TRIP13, TOP2A, HMGB1, HMGB2, RRM1, RRM2, MCM7, MCM4, THOC4, PTBP1, DGL7, MAC30, HMMR, CCNA1, CCNB1* *INHBA.*

Table 6.1. Gene expression changes following AP-2γ silencing show overlap with those occurring following exogenous p21 expression. Significant gene expression changes following AP-2γ silencing in MCF-7 cells ($p<0.05$) were examined for overlapping genes within two studies exploring gene expression changes following exogenous p21 expression in unrelated cell lines. The percentage overlap is expressed along with the list of up (red) and down (green) regulated genes showing overlap. Genes were also annotated as known E2F targets (underlined) based on findings from the literature (Ren *et al.*, 2002; Polager *et al.*, 2002, Cam *et al.*, 2004).

In order to confirm whether the growth arrest caused by AP-2γ silencing in MCF-7 cells was dependant only on p21 induction, future experiments could examine the effect on cell cycle following co-transfection of AP-2γ and p21 siRNAs. If p21 induction following AP-2γ silencing in MCF-7 cells is the only mechanism by which cell cycle arrest is mediated, silencing of p21 should prevent this growth arrest. However, transient co-transfection of two different siRNAs can be technically difficult. Experiments in our lab trying to reduce AP-2α and AP-2γ levels in cell lines where they are both expressed have shown that when combined, one siRNA can be more effective that the other, despite each showing good knockdown when transfected individually. This suggests that the silencing machinery is more efficient at incorporating some

siRNA sequences than others and this is perhaps due to favourable structural characteristics.

6.2.2.2. Does p21 up-regulation occur at the level of transcription following AP-2γ silencing in MCF-7 cells?

Before exploring the potential role for AP-2γ in *CDKN1A* transcriptional repression, it was necessary to determine whether p21 was regulated at the level of transcription. It was perplexing that if this were the case, no significant gene expression change was observed on the microarray for the *CDKN1A* transcript following AP-2γ silencing in MCF-7 cells. However, a closer examination of normalised expression values for the *CDKN1A* probe set in Figure 5.3 (D), showed that the levels of *CDKN1A* mRNA in the control experiments were already high and, although a slight increase in *CDKN1A* gene expression was seen for the AP-2γ siRNA transfected conditions, this difference was minimal and not significant ($p<0.05$). I considered whether the discrepancies between p21 protein levels as detected by western (Figure 5.3) and those detected on the microarray might be due to differences in experimental set up. In particular, the gene expression profile was performed at 96 hours following siRNA transfection and the cells were approaching 100% confluency at the time of harvest while the western blot analysis was performed at 72 hours following transfection and on sub-confluent cells. p21 levels are known to be induced when cells reach high densities and contact inhibition occurs (Macleod *et al.*, 1995; Erhardt & Pittman, 1998) and perhaps in the microarray experiment differences in p21 mRNA level were less pronounced due to p21 already being above basal levels in the control (*reference*) samples. In order to test this hypothesis, p21 mRNA and protein levels were analysed in AP-2γ silenced MCF-7 cells, following a one in three passage at 48hrs, to ensure that cells were sub-confluent when harvested at 72 and 96 hours (Figure 5.5.). Here, elevated p21 mRNA and protein were observed in MCF-7 silenced cells, at both time points suggesting that p21 up-regulation does indeed occur at the level of transcription in MCF-7 cells. Furthermore, the level of induction of p21 mRNA shows a good correlation with the observed increase in p21 protein. Interestingly, p21 protein levels increased in the control extracts harvested at 96 hours compared to those harvested at 72 hours. This suggests that as cells increase in density on the plate contact inhibition contributes to p21 regulation in

MCF-7 cells and supports the explanation for the failure to detect this change in the microarray data.

Although the microarray gene expression changes were predominantly cell cycle related (Figure 4.9 and 4.10) and closer analysis of the cell cycle subset.clearly points to a cell cycle arrest at the G1/S transition (Figure 4.11), with hindsight it may have been more appropriate to have performed the experiments at a slightly earlier time point, or on sub-confluent cells. However, due to the high cost of gene expression profiling experiments it was felt that there was little value in revisiting these experiments in this cell line.

In summary, the change seen in cell cycle gene expression and proliferation in AP-2γ silenced MCF-7 cell can largely be explained by the up-regulation of the CKDI p21, although we have not ruled out the potential involvement of other pathways. This up-regulation occurs at the protein and mRNA level and therefore suggests p21 may be regulated at the level of transcription. This was of particular interest to this study as work by others had already described AP-2α as an important activator of transcription at the *CDKN1A* promoter (Zeng *et al.*, 1997; McPherson *et al.*, 2002). We hypothesised that AP-2γ could be acting in an opposing role in *CDKN1A* transcriptional regulation. This suggests that in normal cycling MCF-7 cells AP-2γ is absolutely required to repress *CDKN1A* transcription and upon AP-2γ silencing this repression is relieved, leading to p21 activation and growth arrest.

6.2.3. AP-2γ dependant transcriptional repression of CDKN1A – a mechanism for the control of proliferation by AP-2γ in MCF-7 cells.

Initial experiments suggested that exogenous AP-2γ was able to repress basal luciferase activity from the -2325 +8 *CDKN1A* reporter construct (Figure 5.8.B). In order to examine whether endogenous AP-2γ was associated with the *CDKN1A* promoter *in vivo*, AP-2γ occupancy was determined using ChIP. ChIP assays were conducted in cycling wt MCF-7 where p21 is maintained at basal levels and enrichment at four discrete *CDKN1A* promoter specific qPCR amplicons was explored (Figure 5.10 A&B). Importantly, enrichment for the *CDKN1A* promoter region amplified by the -21 / +44 primer set was observed in AP-2γ ChIP. This provides strong *in vivo* evidence that endogenous AP-2γ is recruited to the *CDKN1A* promoter region and implicates AP-2γ in the repression of transcription from this promoter. Interestingly, enrichment of the regions amplified by the other primer sets further upstream in the *CDKN1A* promoter was not observed in AP-2γ ChIP. This indicates that endogenous AP-2γ in not recruited to the region near the classical p53 binding site amplified by primer set -2290 / -2185 which has been suggested for AP-2 family member, AP-2α (Mcpherson *et al.*, 2002).

AP-2γ occupancy was also associated with p300 occupancy; the greatest enrichment was observed for the region amplified by the -21 / +44 primer set in p300 ChIP and not amplicons upstream. Although p300 is a known co-factor of AP-2γ, recruited through CITED adapter proteins, it has previously only been described in the context of AP-2 transcriptional activation and it is intriguing that p300 is associated with AP-2γ when it is mediating transcriptional repression. This suggests that the HAT activity of p300 must be altered or repressed. Strikingly the acH4 ChIP shows a lower enrichment for the -21 / +44 amplicon compared to the upstream amplified regions -2290 / -2185 and -935 / -864. This suggests that in cycling MCF-7 cells the -21 / +44 region associated with AP-2γ occupancy, is in a more condensed chromatin state than the regions upstream and perhaps less accessible to the general transcription machinery.

In order to determine whether AP-2γ occupancy of the *CDKN1A* promoter was in fact mediating transcriptional repression, changes in AP-2γ and p300 occupancy or H4 acetylation across the -21 / +44 region were examined following AP-2γ silencing in MCF-7 cells (Figure 5.11. A&B). An expected reduction in AP-2γ promoter occupancy was marked by a reduction in enrichment for -21 / +44 amplicon in the AP-2γ ChIP, following AP-2γ silencing in MCF-7 cells. Furthermore, a two fold increase in enrichment for the acH4 ChIP indicates that increased histone acetylation accompanies the observed decrease in AP-2γ occupancy. Increased histone acetylation correlates with activation of transcription from the *CDKN1A* promoter following AP-2γ silencing in MCF-7 cells, and suggests that an increased HAT activity is important for destabilising the chromatin in this region and allowing the access of the general transcription machinery to the DNA. Interestingly a two fold decrease in enrichment for the p300 ChIP indicates that decreased p300 occupancy accompanies AP-2γ silencing.

Together these data show that a reduction in AP-2γ occupancy at the *CDKN1A* promoter is associated with *CDKN1A* transcriptional activation and provide strong evidence for a role of AP-2γ in the repression of transcription form this promoter. However, many questions remain unanswered.

These experiments do not conclusively demonstrate that p300 and AP-2γ are bound together in the same complex, although an *in vivo* interaction has been shown previously, and requires the presence of CITED2 (Braganca *et al.*, 2003). Originally these ChIP experiments aimed to examine the occupancy of CITED2 alongside AP-2γ and p300. However problems with batch to batch variation in the commercial CITED2 antibody precluded such an analysis. The presence of CITED2 would support the coexistence of AP-2γ and p300 in the same complex. Although it is known that p300 is associated with other factors in this region, such as SP1 (Xiao *et al.*, 2000). Ultimately in order to confirm the co-occupancy of these factors sequential ChIP (re-ChIP) experiments could be attempted in the future.

6.2.3.1 Why does a reduction in p300 occupancy accompany *CDKN1A* transcriptional activation following AP-2γ silencing and how does this relate to histone acetylation?

The observed reduction in p300 occupancy that accompanies *CDKN1A* transcriptional activation following AP-2γ silencing appears to be in conflict with the increased levels of histone acetylation.

One explanation is that alternative HATs, recruited by activators such as SP1, are involved in *CDKN1A* transcriptional activation following AP-2γ silencing. However, it is likely that the p300 remaining, following AP-2γ silencing is associated with SP1 co-activation (Xiao *et al.*, 2000).

It is interesting that a very slight increase in p53 levels is observed following AP-2γ silencing (discussed in detail below). Future experiments should explore whether an increase in p53 occupancy in the p53 binding region (-2290 / -2185) of the *CDKN1A* promoter also occurs. An associated increase in p300, a known p53 coactivator, at this region would be expected to contribute to the observed increase histone acetylation at the -21 / +44 region, as the close proximity of the p53 binding region and core promoter region is thought to be responsible for p53 mediated transcriptional activation *in vivo* (Kaeser & Iggo, 2004).

Finally, the observed decrease in p300 enrichment following AP-2γ silencing could be explained by an increase in p300 HAT activity. In addition to the transcription factor targeted HAT activity by catalytically active p300, work by Black and colleagues has shown that autoacetylation mediated dissociation of p300 from chromatin, is required for ordered PIC assembly (Black *et al.*, 2006). This could lead to a more transient association of p300 with chromatin in the actively transcribing *CDKN1A* promoter and therefore the overall enrichment by p300 in ChIP might be reduced compared to that of the repressed promoter, where an inactive p300 (discussed below) more stably associated with chromatin.

6.2.3.1 How does AP-2γ repress transcription at the p21 promoter?

It is intriguing that AP-2γ and p300 promoter occupancy are observed when AP-2γ is mediating transcriptional repression of the *CDKN1A* promoter.

Indeed studies by others have shown p300 promoter occupancy in the context of transcriptional repression *in vivo*. The formation of a chromatin bound p53-p300-E6 complex was shown to require the presence of p300 in order to repressed *CDKN1A* transcription *in vivo* and was correlated with reduced acetylation of both histone H4 and p53 (Thomas & Chiang, 2005). This implies co-activating properties of p300 can be altered or counteracted by additional proteins.

SUMO modified forms of p300 have been shown to recruit HDAC6 *in vitro* and repress transcription (Girdwood *et al.*, 2003). Equally, conformational changes in p300 following its association with other transcription factors may be important in regulating its activity (Dial *et al.*, 2003). The basal catalytic activity of p300 is known to be stimulated by acetylation of several key lysine residues within an activation loop motif (Thompson *et al.*, 2004). It has been suggested that deacetlyases such as HDACs might associate with p300 to keep it in a catalytically inactive state, or counteract its catalytic activity (Pugh *et al.*, 2006). In support of this hypothesis, a recent study has shown that p300 can acetylate HDAC1 and inhibit its histone deacetylase activity leading to transcription activation (Qiu *et al.*, 2006). This suggests that HAT and HDAC complexes may occur concurrently on promoters and can regulate each other's activity.

Indeed, several lines of evidence support the association of HDAC activity with basal levels of p21 transcription. Small molecule inhibitors of HDACs including Suberoylanilide hydroxamic acid (SAHA) and trichostatin A (TSA) can induce p21 from the *CDKN1A* promoter in MCF7 cells. Using reporter assays, this activity has been shown to require a minimal region of the proximal p21 promoter containing the six SP1 sites (see Figure 5.14. for details) (Huang *et al.*, 2000; Varshochi *et al.*, 2005). Mutation of any of the SP1 sites reduced the activation of p21 transcription following treatment with HDAC inhibitors, suggesting that SP1 binding is important in activation of transcription in this proximal region. RNAi studies have also provided evidence for specific HDAC activity. For example, treatment with HDAC1, but not HDAC3 targeting siRNAs in MCF7 cells induced p21 levels (Varshochi *et al.*, 2005). A

separate study showed siRNA knockdown of HDAC2, but not HDAC1 enhanced p53 mediated induction of p21 and caused a cell cycle arrest in MCF7 cells (Harms & Chen, 2007).

Interestingly, comparisons can also be drawn between the observations described above for HDAC inhibitors and those observed in this study investigating the activity of a series of *CDKN1A* promoter reporter constructs following AP-2γ silencing in MCF-7 cells. Although the data from ChIP experiments show that AP-2γ is associated with the proximal region of *CDKN1A in vivo* (by enrichment of the -21 / +44 amplicon), additional luciferase reporter assays defined a minimal region from –74 to +20 relative to the start of transcription of the *CDKN1A* promoter as sufficient for repression by AP-2γ in the MCF-7 cells (Figures 5.11 & 5.14). As AP-2γ is associated with a region of the p21 promoter suggested to be repressed by HDACs, the AP-2γ mediated repression of transcription in this region could therefore involve the recruitment of HDACs or modulation of HDAC activity. Additionally, in a manner similar to the activation of p21 transcription following treatment with HDAC inhibitors, mutation of any of the 4 SP1 sites within the –74 to +20 promoter was also shown to reduce the activation of p21 transcription following AP-2γ silencing in MCF-7 cells (Figure 5.14). This suggests that the activity of the p21 promoter following AP-2γ silencing is largely mediated by SP1.

6.2.3.2 Does AP-2γ mediate its repression via interaction with AP-2 consensus binding sites in the p21 promoter?

The experiments presented in this study were insufficient to identify a specific binding sequence for AP-2 within the *CDKN1A* promoter. However a minimal +74 to +20 promoter region was shown to be sufficient for AP-2γ mediated repression. This does not include any consensus binding sites, as defined by the SCCNNVRGB consensus (McPherson & Weigel, 1999). Additionally sequences that contributed to AP-2α activation of *CDKNA1 in vitro* are upstream of this minimal region (Zeng *et al.*, 1997). Therefore, it is possible that AP-2γ might bind sequences alternative to the consensus within the +74 to +20 region. Future experiments could use electromobility shift assays (EMSA) to determine whether AP-2γ proteins can directly bind this DNA region *in vitro*. Alternatively, AP-2γ could facilitate its transcriptional repression through direct

interaction with other transcription factors that bind DNA within this region. A prime candidate would be SP1. Many other transcriptional regulators of p21 have been shown to exert their effects through the interaction with SP1 and not a direct interaction with the DNA in this region, these include Myc and HBx protein (reviewed in Gartel & Radhakrishnan, 2005). Indeed, evidence suggests that SP1 and AP-2 family members may interact to regulate transcription. AP2α and SP-1 have been shown to cooperate to activate the *KISS1* transcription *in vitro* (Mitchell *et al.*, 2006). A potential AP-2γ-SP1 interaction could be explored in MCF-7 cells via coIP or sequential ChIP.

6.2.3.3 Can Modulation of AP-2γ by SUMO proteins alter its transcriptional potential?

Previous work in our lab has shown that AP-2α, AP-2β and AP-2γ can interact with the SUMO conjugating enzyme UBC9. This interaction results in SUMO conjugation at the conserved lysine residue (lys 10) in the N-terminal region of AP-2, and was shown to reduce the ability of AP-2 to activate transcription when analysed in reporter gene assays (Eloranta *et al.*, 2002). However, until this study, the effect of SUMO modification of AP-2γ had not been explored in the context of a promoter that was a known target for AP-2γ mediated transcriptional repression. It order to examine whether SUMO modification could facilitate *CDKN1A* transcriptional repression by AP-2γ a series of mutant AP-2γ constructs were assessed for their ability to repress transcription from the –2325 to +8 *CDKN1A* promoter reporter Figure 5.15 (A&B). Interestingly, co-transfection of the *CDKN1A* promoter reporter with the K10R-AP-2γ or ΔN10-AP-2γ non-SUMO-modifiable mutant constructs did not lead to repression of luciferase activity, in contrast to the repression observed with wildtype AP-2γ. These data suggest that SUMO modification at lysine 10 of AP-2γ is important for mediating its transcriptional repression at the *CDKN1A* promoter. Constitutive SUMO-modified forms of AP-2γ reduced luciferase activity over and above that of wild-type AP-2γ when co-transfected with the *CDKN1A* promoter reporter. This shows that artificial conjugation of SUMO to the AP-2γ N-terminus can repress *CDKN1A* promoter activity in HepG2 cells.

Together these transient transfection studies indicate that SUMO modification of AP-2γ mediates its ability to repress *CDKN1A* promoter activity in HepG2 cells *in vitro* and highlight the potential for SUMO modification of AP-2γ to mediate transcriptional repression *in vivo*. However, the presence of SUMO modification of AP-2γ needs to be demonstrated at the *CDKN1A* promoter in MCF-7 cells. This could be explored by performing ChIP for SUMO at the *CDKN1A* proximal promoter region. Alternatively, the effect of siRNA-mediated silencing of UBC9 on p21 expression could be used to determine the involvement of SUMO modification in transcriptional repression of the *CDKN1A* promoter. If, in the absence of active UBC9, endogenous AP-2γ is no longer recruited to the *CDKN1A* promoter, this would be strong evidence implicating SUMO-modified AP-2γ in the repression of this gene *in vivo*.

SUMO modification of AP-2γ might facilitate the recruitment of HDAC repressor complexes, as has been proposed for other factors including p300 and Elk-1, and thus enhance the recruitment of AP-2γ associated complexes into repressive nuclear domains such as PML bodies (Girdwood, et al. 2003; Yang & Sharrocks *et al.*, 2004). Another unanswered question is how the SUMO modification of AP-2γ is regulated? In this context it is interesting that p300 SUMO modification has been reported to be inhibited by p21 protein. An interaction between p21 and the SUMO modification domain of p300 was hypothesised to prevent SUMO modification resulting in enhanced p300 activation *in vitro* (Snowden *et al.*, 2000).

6.2.3.4. Do changes in AP-2α or p53 contribute to p21 up-regulation following AP-2γ silencing in MCF-7 cells?

As both AP-2α and p53 had previously been shown to play a role in the transcriptional activation of *CDKN1A* (McPherson *et al.*, 2002; Wajapeyee & Somasundaram, 2003), changes in AP-2α or p53 protein levels were assessed following AP-2γ knock down in MCF-7 cells. No increase in AP-2α was observed following AP-2γ silencing; indeed AP-2α remained at barely detectable levels (Figure 5.5.A). However, there was a suggestion that p53 might increase slightly following AP-2γ silencing (Figure 5.5.A). Subsequent experiments, using additional AP-2γ targeting siRNAs, revealed a very slight increase in p53 levels accompanying AP-2γ knockdown (Figure 5.5.B). In order

to put these slight changes in p53 level in to context we compared them to the changes in p53 levels following DNA damage stress (Figure 5.6.A). This showed that the wild type p53 protein in our MCF-7 cells behaved predictably following DNA damage stress and rules out any p53 anomalies in our cells. The increase in p53 levels seen following AP-2γ silencing (Figure 5.5 (B) Lanes 3-9) was very slight compared to the stabilisation of p53 observed in cisplatin treated MCF-7 cells, even at the low 1μM cisplatin concentration (Figure 5.6 (A) Lane 2) and was inconsistent with the magnitude of p21 induction. Together these data suggested that the fairly small increase in p53 is not the major mechanism of p21 induction following AP-2γ silencing, although it may contribute to p21 activation.

Shortly after establishing these findings, work by Stabach and colleagues (Stabach *et al.*, 2006) showed a similar two fold p21 induction at the mRNA level 72 hours following AP-2γ knockdown in MCF-7 cells, in support of our findings. However, this AP-2γ knock down was accompanied by a substantial increase in p53 protein levels shown in Figure 6.1. The primary aim of this study by Stabach and colleagues was to follow up previous findings that suggested AP-2 family members could interact with p53 *in vitro*, and the suggestion that AP-2α and p53 could cooperate to activate *CDKN1A* transcription (McPherson *et al.*, 2002; Wajapeyee & Somasundaram, 2003). Indeed, Stabach and colleagues confirmed an interaction of exogenous AP-2α and p53 *in vivo* by co-IP (Stabach *et al.*, 2006). In contrast to previous findings by Wajapeyee & Somasundaram, Stabach and colleagues showed that AP-2α activation of a *CDKN1A* reporter construct was entirely dependant on the expression of p53. These discrepancies aside, Stabach and colleagues suggested using ChIP that occupancy by exogenous AP-2α near the -2250 p53 regulatory region only occurred in HCT116 wt cells and not in HCT116 p53 (-/-) derivatives, again supporting their hypothesis that p53 is required for AP-2α coactivation of p21. These results should be cautiously interpreted as previous work in our lab has shown that the antibody used in this ChIP has a high level of non-specific activity, and no control for non-specific activity was reported in this study (Stabach *et al.*, 2006). In addition to a role for AP-2α in the activation of *CDKN1A* transcription, Stabach and colleagues suggested that exogenous AP-2α expression was able to reduce p53 protein stability in cyclohexamide chase experiments. They hypothesised that AP-2γ might have a similar role in destabilising p53 and that the

observed increase in p53 protein levels following AP-2γ knock down in MCF-7 cells (shown in Figure 6.1) was due to stabilisation of p53. Closer examination of these data shows that these differences are also reflected in the Actin loading control blot and suggests that the majority of this difference could be accounted for by increased protein loading in the AP-2γ knockdown well.

In summary, a role for p53 in the regulation of *CDKN1A* transcription by AP-2 family members remains uncertain. What is clear is that AP-2α can activate *CDKN1A* transcription either by proximal promoter sites (Zeng *et al.*, 1997) or possibly via an interaction with p53 at upstream p53 sites (Stabach *et al.*, 2006) and this activation was shown to be reduced or abolished in the absence of p53 (Wajapeyee & Somasundaram, 2003; Stabach *et al.*, 2006). Perhaps in conflict with these findings exogenous AP-2α can also reduce the stability of p53 (Stabach *et al.*, 2006). The findings presented in this thesis show that *endogenous* AP-2γ can repress *CDKN1A* transcription by interacting with a proximal promoter region *in vivo* but not at the p53 binding sites in upstream regions. AP-2γ silencing in MCF-7 cell relieves this transcriptional repression and is accompanied by a very slight increase in p53 levels. The significance of this slight increase in p53 following AP-2γ silencing requires further analysis. Evidence suggests that the p53 binding region of the *CDKN1A* promoter accounts for a large component of the luciferase activity in MCF-7 cells following AP-2γ silencing (Figure 5.13.). Therefore, it is possible that p53 contributes to the induction of *CDKN1A* following AP-2γ silencing.

It is interesting to note that the only cell line in which we successfully derived stable AP-2γ silenced colonies was the p53-defective T47D cell line (Table 3.2). All of the other cell lines studied are wild type for p53 and exhibit a p21 induction following AP-2γ silencing. In future experiments it will be interesting to see if p21 is induced following transient AP-2γ silencing in T47D. It would be hypothesised that p21 induction would be absent or reduced due to defective p53 activity and thus explaining the derivation of stable AP-2γ silenced colonies in the T47D cell line. The lack of cell cycle arrest in T47D cells makes them a useful tool for future expression profiling experiments. Experiments could explore resultant gene expression profiles following

the silencing of AP-2α and AP-2γ both individually and together, in order to more clearly identify non-cell cycle related AP-2γ targets.

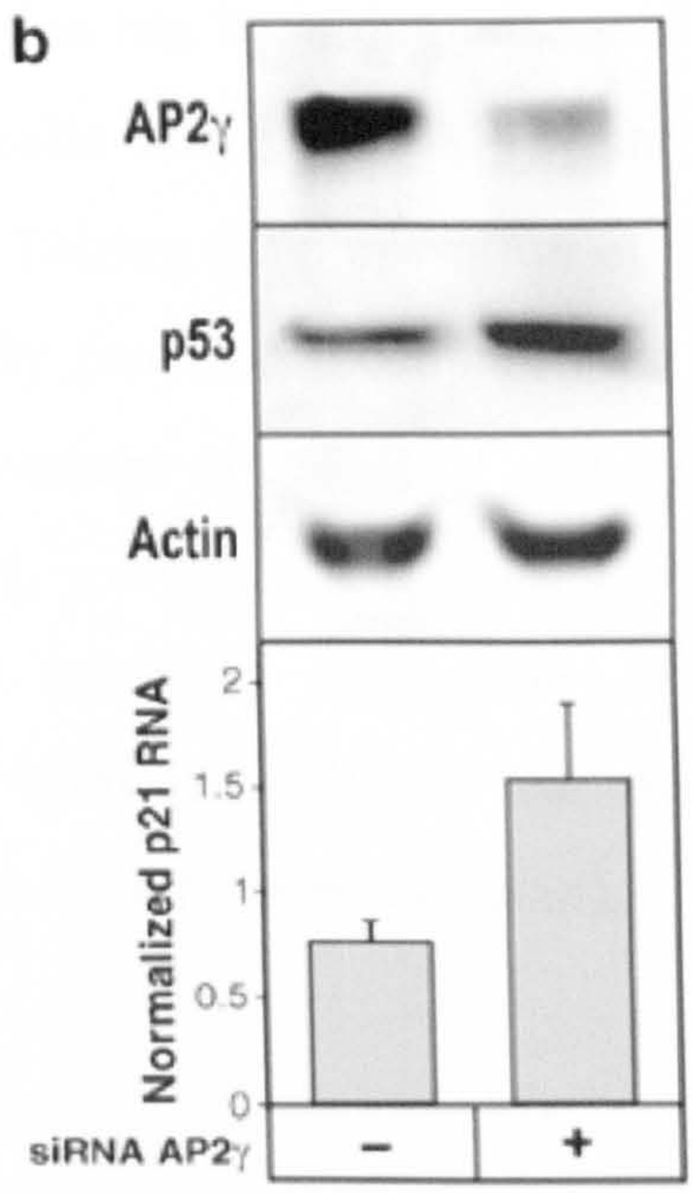


Figure 6.1. AP-2γ silencing in MCF-7 cells is accompanied by an up-regulation of p21 mRNA and p53 protein from Stabach *et al.*, 2006. A similar two fold p21 induction at the mRNA level 72 hours following AP-2γ knockdown in MCF-7 cells, in support of our findings. However, the AP-2γ knock down is also accompanied by a substantial increase p53 protein levels. Closer examination of these data shows that these differences are also reflected in the Actin loading control blot and suggests that the majority of this difference could be accounted for by increased protein loading in the AP-2γ knockdown well.

6.2.3.5 A Model for AP-2γ mediated repression of p21 in MCF-7 cells

Based on the findings in this study and the evidence discussed above a model for the repression of *CDKN1A* transcription by AP-2γ and the subsequent activation of *CDKN1A* transcription following AP-2γ silencing can be proposed, shown in Figure 6.2. There are clearly some gaps in this proposed model, the important one being the discovery of an AP-2γ associated HDAC. It would also be of interest to explore the role of AP-2α which is known to activate *CDKN1A* transcription. AP-2α and AP-2γ can heterodimerise *in vitro* and perhaps in a cell line in which they are both expressed the balance of these factors might be important in the propensity of p21 to be activated or repressed. Indeed, p21 induction has been demonstrated following AP-2γ silencing in ZR-75-1 cells and MCF10A cells where both AP-2γ and AP-2α are expressed at high levels (Figures 3.1 and 5.17). During the preparation of this thesis, preliminary work in the lab has shown that AP-2γ promoter occupancy is associated with HDAC1 and HDAC2, marked by the enrichment of the -21 / +44 *CDKN1A* promoter amplicon in HDAC1 and HDAC2 ChIPs. Studies are currently underway to confirm their involvement in the AP-2γ mediated repression of *CDKN1A* transcription by demonstrating that either or both HDACs are lost from the promoter after AP-2γ silencing (Karsten Friedrich, *personal communication*).

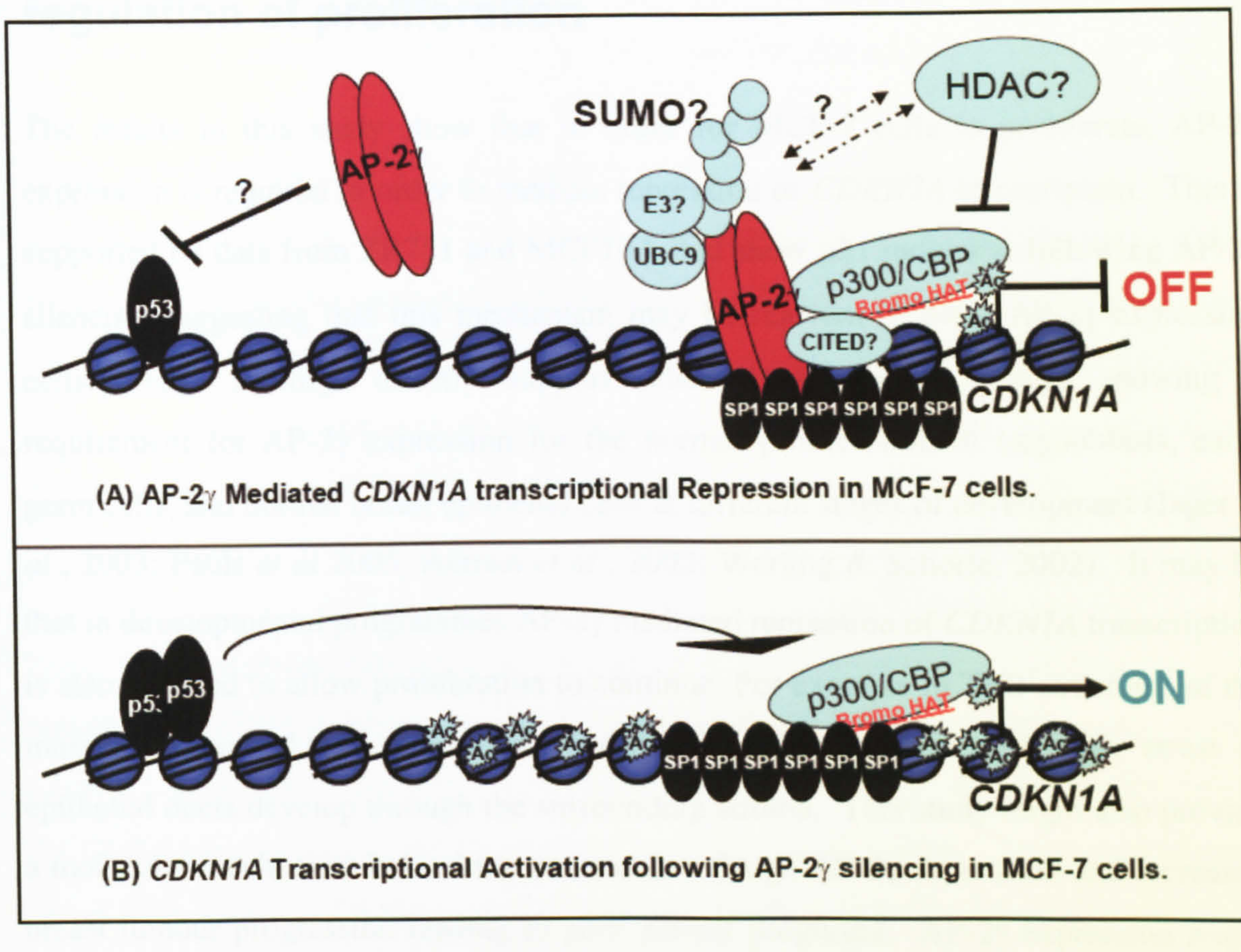


Figure 6.2. A model for AP-2γ mediated transcriptional repression at the *CDKN1A* promoter.

(A) Repression. AP-2γ is recruited to the promoter by an undefined TFBS or via an interaction with SP1. In this repressed state AP-2γ is associated with p300 which is probably recruited via CITED proteins. AP-2γ may also recruit an undefined HDAC, via SUMO modification of its conserved N-terminal lysine residues, which then counteracts or inactivates the HAT activity of p300. In this state chromatin remains inaccessible to other activating factors, which is marked by a lowered level of histone acetylation.

(A) Activation. Following AP-2γ silencing in MCF-7 cells, AP-2γ is depleted from the *CDKN1A* promoter. In turn the associated HDAC activity is lost. The remaining p300, in association with SP1 and perhaps p53, is now active at its associated HAT activity leads to increased Histone acetylation, the resulting chromatin remodelling increases the accessibility to the general transcription machinery and transcription is activated.

6.3 What is the functional significance of repression of the *CDKN1A* transcription by AP-2γ and the associated regulation of proliferation

The results in this study show that in order for MCF-7 cells to proliferate, AP-2γ expression is required in order to mediate repression of *CDKN1A* transcription. This is supported by data from ZR751 and MCF10A that show p21 induction following AP-2γ silencing suggesting that this mechanism may be relevant in other AP-2γ expressing cells. These findings directly support other developmental studies showing a requirement for AP-2γ expression for the normal proliferation of trophoblasts, early germ cells, and normal breast epithelial cells at different stages of development (Jager *et al.*, 2003; Pauls *et al* 2005; Auman *et al.*, 2002; Werling & Schorle, 2002). It may be that in developmental programmes AP-2γ mediated repression of *CDKN1A* transcription is also required to allow proliferation to continue. For example in TEB structures of the mammary gland AP-2γ expression in cap cells might prevent cell cycle arrest as epithelial ducts develop through the surrounding stroma. This study might also provide a molecular mechanism behind the association of high AP-2γ expression and increased breast tumour progression leading to poor patient prognosis. AP-2γ expression might represent increased insensitivity to anti-growth signals. Repression of *CDKN1A* transcription by AP-2γ might lead to failure of cells to arrest via *CDKN1A* up-regulation in response to anti-proliferative cues, leading to unchecked cell proliferation.

Equally, repression of *CDKN1A* transcription by AP-2γ might also represent a challenge to growth arrest in response to DNA damage stress. It is interesting that following Cisplatin treatment of MCF-7 cells in this study, AP-2γ levels were slightly reduced (Figure 5.6. B). This suggests the presence of a mechanism to down-regulate AP-2γ in response to DNA damage stress.

High AP-2γ expression has been linked to poor patient prognosis in patients positive for ERα expression. In these patients AP-2γ expression represents a challenge to anti-oestrogen therapy as AP-2γ (a ligand activated target gene of ERα (Orso *et al.*, 2004)) would be expected to be down-regulated by anti-oestrogens. Indeed, poor patient

prognosis in this subset of ERα positive tumours might be due to failed ERα regulation of *TFAPC* transcription. The prevailing levels of AP-2γ in these tumours might lead to failure of cells to arrest via *CDKN1A* induction following anti-oestrogen treatment contributing to enhanced proliferation and poor patient prognosis.

Work presented in this thesis shows that ERα can indeed mediate some of its proliferative effects in MCF-7 cells via AP-2γ dependant transcriptional regulation of *CDKN1A*. ICI182,780 anti-oestrogen treatment in MCF-7 cells was accompanied by a decrease in AP-2γ occupancy across the -21 / +44 region of *CDKN1A* promoter (Figure 5.12. A). This change was consistent with the reduction in AP-2γ protein levels observed following ICI182,780 treatment and was accompanied by an up-regulation of p21 mRNA and protein levels and a reduction in RB phosphorylation (Figure 5.12. B).

These results were interesting in the light of findings by Varshochi and colleagues. The authors used ChIP to show that ICI182,780 treatment of MCF-7 cells could induce *CDKN1A* transcription through the release of ERα and HDAC1 association with SP1 at the *CDKN1A* proximal promoter *in vivo* (Varshochi *et al.*, 2005). These data suggest that a dual level of transcriptional control occurs following ICI182,780 treatment of MCF-7 cells. Firstly, the antagonism of ligand activated ERα transcription of *TFAP2C* leads to a reduction in AP-2γ levels and a concurrent reduction in AP-2γ occupancy of the *CDKN1A* proximal promoter leading to *CDKN1A* up-regulation. Secondly, the ICI182,780 mediated disruption of an ERα-SP1-HDAC1 complex at the *CDKN1A* proximal promoter leads to release of HDAC1 repression and *CDKN1A* up-regulation.

Interestingly, work in our lab exploring cell line models of anti-oestrogen resistance, has shown that, although wt MCF-7 cells down-regulate AP-2γ in response to anti-oestrogen treatment, Tamoxifen-resistant derivatives express normal levels of AP-2γ. As these cells now cycle in the presence of Tamoxifen, it is interesting to speculate that part of the adaptation process has been the re-expression of AP-2γ allowing the cells to effectively repress p21 expression and re-enter the cell cycle. In order to determine whether AP-2γ expression in these cell lines might represent a general mechanism of anti-oestrogen resistance, work is currently underway to explore whether AP-2γ is down-regulated in response to acute levels of anti-oestrogens in other ERα positive cell

lines (ZR75-1, T47D) and whether their resistant derivatives also re-express AP-2γ to allow their re-entry into the cell cycle via repression of *CDKN1A*.

As well as being associated with increased proliferation in developmental studies, enhanced apoptosis was also observed in cells with forced over-expression of AP-2γ (Jager *et al.*, 2003; Meredith, 2006). No significant increase in apoptosis was observed following AP-2γ silencing in MCF-7 cells. Developmental data suggest that in order to sustain AP-2γ expression in cells, normal apoptotic pathways must be bypassed or silenced. Sustained expression of AP-2γ in the immortal MCF-7 cell line might therefore represent insensitivity to any potential pro-apoptotic effects of AP-2γ. Therefore, induction of apoptosis would not be an expected consequence of AP-2γ silencing in MCF-7 cells and these experiments cannot rule out an involvement for AP-2γ in the normal regulation of apoptosis. Interestingly, ~~that~~ treatment of MCF-7 cells with cisplatin results in an up-regulation of AP-2α and moderate down-regulation of AP-2γ (Figure 5.6. B). An up-regulation of AP-2α is consistent with findings from others who showed that AP-2α expression is required for transcriptional repression of *BCL2* during chemotherapy mediated apoptosis in colon carcinoma cells (Wajapeyee *et al.*, 2006), and suggests a similar mechanism in MCF-7 cells. The moderate AP-2γ down-regulation following cisplatin treatment in MCF-7 cells might also have consequences for apoptosis, although if this were the case it would be in conflict with data relating to forced over-expression of AP-2γ to enhance apoptosis. It would be interesting to explore whether the response to chemotherapy-induced apoptosis is altered in cells silenced for AP-2γ compared to their wild type counterparts.

6.4 Other Potential Targets for AP-2γ regulation

As well as revealing a role AP-2γ in the regulation of the cell cycle (discussed above), the transcription profiling experiments presented in this thesis have also identified changes in the abundance of additional transcripts following AP-2γ silencing in MCF-7 cells. A number of these changes were validated using qPCR, and where, possible western blots were used to test if these translated into changes at the protein level (Section 4.9).

6.4.1. SPDEF (SAM Pointed Domain containing Ets transcription Factor)

A significant down-regulation of the *SPDEF* transcript following AP-2γ silencing in MCF-7 cells was verified using qPCR (Figure 4.12). Although, further investigations are required to validate this change by western blot, it is interesting that several AP-2 TFBS were found in the –500 to +50 bp relative to the start of *SPDEF* transcription, implicating AP-2γ in the activation of transcription from this promoter (Appendix 2). High levels of *SPDEF* transcript have been detected in a large proportion of human breast cancers, where *SPDEF* expression has been related to invasive potential (Feldman *et al.*, 2003). In particular, SPDEF has been shown to enhance migration and invasion in normal and tumourigenic breast epithelial cell lines (Gunawardane *et al.*, 2005). Interestingly, *SQSTM1* (Sequestosome 1), a previously described transcriptional target of SPDEF was also identified within the gene expression changes following AP-2γ silencing in MCF-7 cells (244804_at: Fold Change = 1.74, p Value = 0.00492). Unfortunately, very little is known about additional transcriptional targets of SPDEF and it was not possible to identify further potential SPDEF transcriptional effects within the AP-2γ microarray data. These observations lead to the speculation that AP-2γ expression might be important for the regulation of invasion and migration of cells during normal breast development and carcinogenesis.

6.4.2. TIMP2 (Tissue Inhibitor of MetalloProteinases 2)

In support of a role for AP-2γ in the regulation of invasion is the qPCR verification of the significant up-regulation of the *TIMP2* transcript following AP-2γ silencing in MCF-7 cells (Figure 4.12). Again, further investigations are required to validate this change at the functional level and indeed any involvement for AP-2γ in *TIMP2* regulation. TIMPs have been shown to irreversibly inhibit the activity of matrix metalloproteinases (MMPs). The activity of MMPs is associated with many types of cancer and is thought to be important for the degradation of connective tissue contributing to invasion by facilitating tumour cell migration (reviewed in Overall & Kleifeld, 2006). Therefore the down-regulation of *TIMP2* in AP-2γ expressing cells

suggests that these cells might be more capable of MMP mediated degradation of connective tissue required for invasion.

6.4.5. TGLN (Transgelin)

The up-regulation of the *TGLN* transcript following AP-2γ silencing in MCF-7 cells was verified by qPCR and western blot (Figure 4.12 and Figure 4.14). Again, further investigations are required to validate this change at the functional level and indeed an for AP-2γ in *TGLN* regulation. Transgelin is a 22kDa actin binding protein that was first discovered as an early marker that was lost during the onset of cellular transformation (Shapland *et al.*, 1993) Indeed, Transgelin expression is frequently down-regulated in colon and breast tumour tissue (Shields *et al.*, 2002). A recent study has shown that Transgelin is able to repress MMP-9 gene expression, via modulation of ERK signalling in colon carcinoma cells (Nair *et al.*, 2006), although no change was observed for the MMP-9 transcript following AP-2γ silencing in MCF-7 cell in this study.

6.4.3. Kiaa1324

A significant down-regulation of the *KIAA1324* transcript following AP-2γ silencing in MCF-7 cells was verified by qPCR and western blot (Figure 4.12 and Figure 4.14). This protein has no known function, but was selected for further analysis in this study as it was the most statistically significant down-regulated gene on the microarray, following AP-2γ silencing in MCF-7 cells (Figure 4.8) and was represented by multiple probes sets. Subsequent immunofluorescence revealed kiaa1324 was localised to the membrane in MCF-7 and was regulated by oestrogens, supporting observations by others (Deng *et al.*, 2005). The AP-2 TFBS analysis presented in this study (Appendix 2) revealed several AP-2 sites within the region –500 to +50 bp relative to the start of *Kiaa1324* transcription. During the preparation of this thesis, enrichment of a -180 / -129 *KIAA1324* promoter specific amplicon was observed in AP-2γ antibody ChIP elutes from MCF-7 cells, implicating AP-2γ in the direct regulation of *KIAA1324* transcription (Karsten Friedrich, *personal communication*). Kiaa1324 protein was also detected in other breast carcinoma cell lines where AP-2γ was expressed, including ZR751 and

T47D, however Kiaa1324 protein was absent from the normal MCF10A cell line suggesting mechanisms other than AP-2 control its expression in this cell line (data not shown). AP-2γ silencing in these cell lines should be explored with respect to Kiaa1324 expression in order to see if AP-2γ can also regulate *KIAA1324* transcription in these additional tumour cell lines. Future work should also examine the relationship between AP-2γ expression and Kiaa1324 via immunohistochemistry in human breast tumours and, in order to identify a functional role for Kiaa1324, the effect of siRNA mediated Kiaa1324 silencing in MCF-7 cells could be explored with respect to tumourigenic character such as cell proliferation, apoptosis and migration.

6.4.4. IGFBP3 (Insulin-like Growth Factor Binding Protein 3)

The up-regulation of the *IGFBP3* transcript following AP-2γ silencing in MCF-7 cells was verified by qPCR and western blot (Figure 4.12 and Figure 4.14). Interestingly, a two fold increase in the relative abundance of *IGFPB3* mRNA was also observed following AP-2γ silencing in ZR75-1 cells (Figure 4.15), suggesting that the mechanisms of *IGFBP3* up-regulation following AP-2γ silencing are similar in MCF-7 and ZR75-1 cells. Several AP-2 TFBS were found in the –500 to +50 bp relative to the start of *IGFPB3* transcription (Appendix 2), implicating AP-2γ in the repression of transcription from this promoter. Future ChIP experiments will be important in verifying such an interaction but it is interesting that others have identified IGFBP3 as a potential target for AP-2α activation (Wajapeyee *et al.*, 2005). The functional significance of IGFBP3 up-regulation following AP-2γ silencing in these AP-2 expressing cell lines is unclear. IGFBP3 expression can both mediate anti-proliferative signalling and apoptosis in breast cancer cell lines (reviewed in Cohen, 2006). Studies examining forced expression of IGFP3 in mouse models of prostate cancer have showed that IGFBP3 can act in two distinct modes to mediate a dramatic suppression of tumour development (Silha *et al.*, 2006). Firstly, this can be indirectly mediated by IGFBP3 through complex formation with IGFs, preventing IGFs from activating IGF receptors and inhibiting cell survival signalling. Alternatively, IGFBP3 suppression of tumour development can be mediated directly through IGF-independent mechanism (reviewed in Cohen, 2006). Evidence exists for IGF-independent mechanisms in MCF-7 cells.

Recombinant IGFBP3 has been shown to stimulate phosphorylation of the exogenously expressed cell surface receptor TGF-RI, although the initiating mechanism remains unclear, this was accompanied by an increase in Smad transcription factor phosphorylation (Fanayan *et al.*, 2002). Although not explored with reference to IGFBP3, TGF-β signalling can lead to growth inhibition via Smad mediated activation of target genes such as *CDKN1A*. Work presented in this thesis, demonstrated that a mutation in the region of the *CDKN1A* promoter previously shown to be critical for reporter activity after TGFβ signalling (Datto *et al.*, 1995) did not effect *CDKN1A* reporter activity following AP-2γ silencing (mut#3, Figure 5.14), suggesting that increased TGFβ signalling is not involved in the up-regulation of *CDKN1A*, following AP-2γ silencing. These data suggest that if IGFBP3 up-regulation following AP-2γ silencing is exerting any anti-proliferative effects in MCF-7 cells, these are not mediated through any of previously described mechanisms. IGFBP3 has also been shown to activate apoptosis in breast cancer cell lines, however this mechanism is less well understood (Reviewed in Cohen, 2006). Although, no significant increase in apoptosis was observed following AP-2γ silencing in MCF-7 cells, it maybe that the normal mechanisms of IGFBP3-mediated apoptosis are not functional in this cell line. It is interesting that proteins suggested to activate apoptosis, both IGFBP3 and the related protein IGFBP5, are potential targets for down-regulation by AP-2γ in breast cancer cell lines. Although more work is required to establish the functional relevance of these observations this suggests that in some contexts AP-2γ can repress apoptotic mechanisms.

6.5 Concluding Remarks

Presented within this thesis is a large scale microarray analysis conducted in order to identify potential AP-2γ regulated genes. As well as a large proportion of cell cycle regulated genes down-regulated in response to AP-2γ silencing, additional transcripts involved in the regulation of apoptosis, invasion and developmental pathways were also identified. The functional relevance of these AP-2γ dependant regulated targets requires further investigation, although they implicate AP-2γ expression in regulation of pathways additional to proliferation.

Also in this study we have shown that in order for MCF-7 cells to proliferate, AP-2γ expression is required to mediate repression of *CDKN1A* transcription. This provides the first evidence for a molecular mechanism behind the requirement for AP-2γ expression in the normal proliferation required in developing tissues. Importantly AP-2γ mediated repression of *CDKN1A* transcription may also be important for the control of proliferation in beast cancers. Evidence suggests that breast cancer cell lines can provide good models in which to identify molecular events that are likely to be important in human breast cancers (Zhu *et al.*, 2006). Therefore, the correlation of high AP-2γ expression and poor prognosis in patients with breast cancer might be due to failure of cells to arrest via *CDKN1A* up-regulation in response to anti-proliferative cues, leading to unchecked cell proliferation.

REFERENCES

Affymetrix (url). Available from: http://www.affymetrix.com/.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-402.

Anttila, M. A., Kellokoski, J. K., Moisio, K. I., Mitchell, P. J., Saarikoski, S., Syrjanen, K. and Kosma, V. M. (2000). Expression of transcription factor AP-2alpha predicts survival in epithelial ovarian cancer. *Br J Cancer* **82**, 1974-83.

Applied Biosystems (url). Standard Curve Method for Relative Quantification, Available from: http://www.appliedbiosystems.com/.

Aqeilan, R. I., Palamarchuk, A., Weigel, R. J., Herrero, J. J., Pekarsky, Y. and Croce, C. M. (2004). Physical and functional interactions between the Wwox tumour suppressor protein and the AP-2gamma transcription factor. *Cancer Res* **64**, 8256-61.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T. et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-9.

Auman, H. J., Nottoli, T., Lakiza, O., Winger, Q., Donaldson, S. and Williams, T. (2002). Transcription factor AP-2gamma is essential in the extra-embryonic lineages for early postimplantation development. *Development* **129**, 2733-47.

Bajic, V. B., Choudhary, V. and Hock, C. K. (2004). Content analysis of the core promoter region of human genes. *In Silico Biol* 4, 109-25.

Bamforth, S. D., Braganca, J., Eloranta, J. J., Murdoch, J. N., Marques, F. I., Kranc, K. R., Farza, H., Henderson, D. J., Hurst, H. C. and Bhattacharya, S. (2001). Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking Cited2, a new Tfap2 co-activator. *Nat Genet* **29**, 469-74.

Bar-Eli, M. (2001). Gene regulation in melanoma progression by the AP-2 transcription factor. *Pigment Cell Res* **14**, 78-85.

Bartek, J., Iggo, R., Gannon, J. and Lane, D. P. (1990). Genetic and immunochemical analysis of mutant p53 in human breast cancer cell lines. *Oncogene* **5**, 893-9.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-97.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. et al. (2004). The Pfam protein families database. *Nucleic Acids Res* **32**, D138-41.

Bates, N. P. and Hurst, H. C. (1997). An intron 1 enhancer element mediates oestrogen-induced suppression of ERBB2 expression. *Oncogene* **15**, 473-81.

Batsche, E., Muchardt, C., Behrens, J., Hurst, H. C. and Cremisi, C. (1998). RB and c-Myc activate expression of the E-cadherin gene in epithelial cells through interaction with transcription factor AP-2. *Mol Cell Biol* **18**, 3647-58.

Bauer, R., McGuffin, M. E., Mattox, W. and Tainsky, M. A. (1998). Cloning and characterization of the Drosophila homologue of the AP-2 transcription factor. *Oncogene* **17**, 1911-22.

Beattie, J., Allan, G. J., Lochrie, J. D. and Flint, D. J. (2006). Insulin-like growth factor-binding protein-5 (IGFBP-5): a critical member of the IGF axis. *Biochem J* **395**, 1-19.

Beckmann, M. W., Niederacher, D., Schnurch, H. G., Gusterson, B. A. and Bender, H. G. (1997). Multistep carcinogenesis of breast cancer and tumour heterogeneity. *J Mol Med* **75**, 429-39.

Beissbarth, T. and Speed, T. P. (2004). GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20**, 1464-5.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Roy. Statist. Soc. Ser. B* **57**, 289-300.

Berger, A. J., Davis, D. W., Tellez, C., Prieto, V. G., Gershenwald, J. E., Johnson, M. M., Rimm, D. L. and Bar-Eli, M. (2005). Automated quantitative analysis of activator protein-2alpha subcellular expression in melanoma tissue microarrays correlates with survival prediction. *Cancer Res* **65**, 11185-92.

Bernstein, B. E., Meissner, A. and Lander, E. S. (2007). The mammalian epigenome. *Cell* **128**, 669-81.

Black, J. C., Choi, J. E., Lombardo, S. R. and Carey, M. (2006). A mechanism for coordinating chromatin modification and preinitiation complex assembly. *Mol Cell* **23**, 809-18.

Bosher, J. M., Totty, N. F., Hsuan, J. J., Williams, T. and Hurst, H. C. (1996). A family of AP-2 proteins regulates c-erbB-2 expression in mammary carcinoma. *Oncogene* **13**, 1701-7.

Bosher, J. M., Williams, T. and Hurst, H. C. (1995). The developmentally regulated transcription factor AP-2 is involved in c-erbB-2 overexpression in human mammary carcinoma. *Proc Natl Acad Sci U S A* **92**, 744-7.

Braganca, J., Eloranta, J. J., Bamforth, S. D., Ibbitt, J. C., Hurst, H. C. and Bhattacharya, S. (2003). Physical and functional interactions among AP-2 transcription factors, p300/CREB-binding protein, and CITED2. *J Biol Chem* **278**, 16021-9.

Bioconductor (url). The Bioconductor Project. Available from: http://www.bioconductor.org/

Braganca, J., Swingler, T., Marques, F. I., Jones, T., Eloranta, J. J., Hurst, H. C., Shioda, T. and Bhattacharya, S. (2002). Human CREB-binding protein/p300-interacting transactivator with ED-rich tail (CITED) 4, a new member of the CITED family, functions as a co-activator for transcription factor AP-2. *J Biol Chem* **277**, 8559-65.

Brenneisen, P., Blaudschun, R., Gille, J., Schneider, L., Hinrichs, R., Wlaschek, M., Eming, S. and Scharffetter-Kochanek, K. (2003). Essential role of an activator protein-2 (AP-2)/specificity protein 1 (Sp1) cluster in the UVB-mediated induction of the human vascular endothelial growth factor in HaCaT keratinocytes. *Biochem J* **369**, 341-9.

Brewer, S., Feng, W., Huang, J., Sullivan, S. and Williams, T. (2004). Wnt1-Cre-mediated deletion of AP-2alpha causes multiple neural crest-related defects. *Dev Biol* **267**, 135-52.

Brewer, S., Jiang, X., Donaldson, S., Williams, T. and Sucov, H. M. (2002). Requirement for AP-2alpha in cardiac outflow tract morphogenesis. *Mech Dev* **110**, 139-49.

Brewer, S. and Williams, T. (2004). Loss of AP-2alpha impacts multiple aspects of ventral body wall development and closure. *Dev Biol* **267**, 399-417.

Bridge, A. J., Pebernard, S., Ducraux, A., Nicoulaz, A. L. and Iggo, R. (2003). Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* **34**, 263-4.

Brummelkamp, T. R., Bernards, R. and Agami, R. (2002). Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* **2**, 243-7.

Brummelkamp, T. R., Bernards, R. and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550-3.

Burdette, J. E., Jeruss, J. S., Kurley, S. J., Lee, E. J. and Woodruff, T. K. (2005). Activin A mediates growth inhibition and cell cycle arrest through Smads in human breast cancer cells. *Cancer Res* **65**, 7968-75.

Caldon, C. E., Daly, R. J., Sutherland, R. L. and Musgrove, E. A. (2006). Cell cycle control in breast cancer cells. *J Cell Biochem* **97**, 261-74.

Cam, H., Balciunaite, E., Blais, A., Spektor, A., Scarpulla, R. C., Young, R., Kluger, Y. and Dynlacht, B. D. (2004). A common set of gene regulatory networks links metabolism and growth inhibition. *Mol Cell* **16**, 399-411.

Cerda, S. R., Mustafi, R., Little, H., Cohen, G., Khare, S., Moore, C., Majumder, P. and Bissonnette, M. (2006). Protein kinase C delta inhibits Caco-2 cell proliferation by selective changes in cell cycle and cell death regulators. *Oncogene* **25**, 3123-38.

Chan, H. M. and La Thangue, N. B. (2001). p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J Cell Sci* **114**, 2363-73.

Chang, B. D., Watanabe, K., Broude, E. V., Fang, J., Poole, J. C., Kalinichenko, T. V. and Roninson, I. B. (2000). Effects of p21Waf1/Cip1/Sdi1 on cellular gene expression: implications for carcinogenesis, senescence, and age-related diseases. *Proc Natl Acad Sci U S A* **97**, 4291-6.

Chazaud, C., Oulad-Abdelghani, M., Bouillet, P., Decimo, D., Chambon, P. and Dolle, P. (1996). AP-2.2, a novel gene related to AP-2, is expressed in the forebrain, limbs and face during mouse embryogenesis. *Mech Dev* **54**, 83-94.

Chen, C. Z., Li, L., Lodish, H. F. and Bartel, D. P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83-6.

Chen, T. T., Wu, R. L., Castro-Munozledo, F. and Sun, T. T. (1997). Regulation of K3 keratin gene transcription by Sp1 and AP-2 in differentiating rabbit corneal epithelial cells. *Mol Cell Biol* **17**, 3056-64.

Chi, J. T., Chang, H. Y., Wang, N. N., Chang, D. S., Dunphy, N. and Brown, P. O. (2003). Genomewide view of gene silencing by small interfering RNAs. *Proc Natl Acad Sci U S A* **100**, 6343-6.

Cleator, S., Heller, W. and Coombes, R. C. (2007). Triple-negative breast cancer: therapeutic options. *Lancet Oncol* **8**, 235-44.

Cohen, P. (2006). Insulin-like growth factor binding protein-3: insulin-like growth factor independence comes of age. *Endocrinology* **147**, 2109-11.

Coleman, M. L., Marshall, C. J. and Olson, M. F. (2003). Ras promotes p21(Waf1/Cip1) protein stability via a cyclin D1-imposed block in proteasome-mediated degradation. *Embo J* **22**, 2036-46.

Dallas, P. B., Gottardo, N. G., Firth, M. J., Beesley, A. H., Hoffmann, K., Terry, P. A., Freitas, J. R., Boag, J. M., Cummings, A. J. and Kees, U. R. (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? *BMC Genomics* **6**, 59.

Datto, M. B., Yu, Y. and Wang, X. F. (1995). Functional analysis of the transforming growth factor beta responsive elements in the WAF1/Cip1/p21 promoter. *J Biol Chem* **270**, 28623-8.

Debernardi, S., Bassini, A., Jones, L. K., Chaplin, T., Linder, B., de Bruijn, D. R., Meese, E. and Young, B. D. (2002). The MLL fusion partner AF10 binds GAS41, a protein that interacts with the human SWI/SNF complex. *Blood* **99**, 275-81.

deConinck, E. C., McPherson, L. A. and Weigel, R. J. (1995). Transcriptional regulation of estrogen receptor in breast carcinomas. *Mol Cell Biol* **15**, 2191-6.

Delacroix, L., Begon, D., Chatel, G., Jackers, P. and Winkler, R. (2005). Distal ERBB2 promoter fragment displays specific transcriptional and nuclear binding activities in ERBB2 overexpressing breast cancer cells. *DNA Cell Biol* **24**, 582-94.

Deng, L., Broaddus, R. R., McCampbell, A., Shipley, G. L., Loose, D. S., Stancel, G. M., Pickar, J. H. and Davies, P. J. (2005). Identification of a novel estrogen-regulated gene, EIG121, induced by hormone replacement therapy and differentially expressed in type I and type II endometrial cancer. *Clin Cancer Res* **11**, 8258-64.

Dial, R., Sun, Z. Y. and Freedman, S. J. (2003). Three conformational states of the p300 CH1 domain define its functional properties. *Biochemistry* **42**, 9937-45.

Ding, X., Fan, C., Zhou, J., Zhong, Y., Liu, R., Ren, K., Hu, X., Luo, C., Xiao, S., Wang, Y. et al. (2006). GAS41 interacts with transcription factor AP-2beta and stimulates AP-2beta-mediated transactivation. *Nucleic Acids Res* **34**, 2570-8.

Doench, J. G., Petersen, C. P. and Sharp, P. A. (2003). siRNAs can function as miRNAs. *Genes Dev* **17**, 438-42.

Douglas, D. B., Akiyama, Y., Carraway, H., Belinsky, S. A., Esteller, M., Gabrielson, E., Weitzman, S., Williams, T., Herman, J. G. and Baylin, S. B. (2004). Hypermethylation of a small CpGuanine-rich region correlates with loss of activator protein-2alpha expression during progression of breast cancer. *Cancer Res* **64**, 1611-20.

Downward, J. (2004). RNA interference-based functional genomics in cancer research--an introduction. *Oncogene* **23**, 8334-5.

Duan, C. and Clemmons, D. R. (1995). Transcription factor AP-2 regulates human insulin-like growth factor binding protein-5 gene expression. *J Biol Chem* **270**, 24844-51.

Dunn, S. M., Keough, R. A., Rogers, G. E. and Powell, B. C. (1998). Regulation of a hair follicle keratin intermediate filament gene promoter. *J Cell Sci* **111 (Pt 23)**, 3487-96.

Dykxhoorn, D. M., Novina, C. D. and Sharp, P. A. (2003). Killing the messenger: short RNAs that silence gene expression. *Nat Rev Mol Cell Biol* **4**, 457-67.

Eckert, D., Buhl, S., Weber, S., Jager, R. and Schorle, H. (2005). The AP-2 family of transcription factors. *Genome Biol* **6**, 246.

Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**, 14863-8.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-8.

el-Deiry, W. S., Tokino, T., Velculescu, V. E., Levy, D. B., Parsons, R., Trent, J. M., Lin, D., Mercer, W. E., Kinzler, K. W. and Vogelstein, B. (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817-25.

Eloranta, J. J. and Hurst, H. C. (2002). Transcription factor AP-2 interacts with the SUMO-conjugating enzyme UBC9 and is sumolated in vivo. *J Biol Chem* **277**, 30798-804.

Erclik, M. S. and Mitchell, J. (2005). Activation of the insulin-like growth factor binding protein-5 promoter by parathyroid hormone in osteosarcoma cells requires activation of an activated protein-2 element. *J Mol Endocrinol* **34**, 713-22.

Erhardt, J. A. and Pittman, R. N. (1998). p21WAF1 induces permanent growth arrest and enhances differentiation, but does not alter apoptosis in PC12 cells. *Oncogene* **16**, 443-51.

Fanayan, S., Firth, S. M. and Baxter, R. C. (2002). Signaling through the Smad pathway by insulin-like growth factor-binding protein-3 in breast cancer cells. Relationship to transforming growth factor-beta 1 signaling. *J Biol Chem* **277**, 7255-61.

Fei, P. and El-Deiry, W. S. (2003). P53 and radiation responses. *Oncogene* **22**, 5774-83.

Feldman, R. J., Sementchenko, V. I., Gayed, M., Fraig, M. M. and Watson, D. K. (2003). Pdef expression in human breast cancer is correlated with invasive potential and altered gene expression. *Cancer Res* **63**, 4626-31.

Feng, W. and Williams, T. (2003). Cloning and characterization of the mouse AP-2 epsilon gene: a novel family member expressed in the developing olfactory bulb. *Mol Cell Neurosci* **24**, 460-75.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-11.

Friedrichs, N., Jager, R., Paggen, E., Rudlowski, C., Merkelbach-Bruse, S., Schorle, H. and Buettner, R. (2005). Distinct spatial expression patterns of AP-2alpha and AP-2gamma in non-neoplastic human breast and breast cancer. *Mod Pathol* **18**, 431-8.

Fukuda, A., Nakadai, T., Shimada, M., Tsukui, T., Matsumoto, M., Nogi, Y., Meisterernst, M. and Hisatake, K. (2004). Transcriptional coactivator PC4 stimulates promoter escape and facilitates transcriptional synergy by GAL4-VP16. *Mol Cell Biol* **24**, 6525-35.

Garcia, M. A., Campillos, M., Marina, A., Valdivieso, F. and Vazquez, J. (1999). Transcription factor AP-2 activity is modulated by protein kinase A-mediated phosphorylation. *FEBS Lett* **444**, 27-31.

Gartel, A. L. and Radhakrishnan, S. K. (2005). Lost in transcription: p21 repression, mechanisms, and consequences. *Cancer Res* **65**, 3980-5.

Gartel, A. L. and Tyner, A. L. (1999). Transcriptional regulation of the p21((WAF1/CIP1)) gene. *Exp Cell Res* **246**, 280-9.

Gaubatz, S., Imhof, A., Dosch, R., Werner, O., Mitchell, P., Buettner, R. and Eilers, M. (1995). Transcriptional activation by Myc is under negative control by the transcription factor AP-2. *Embo J* **14**, 1508-19.

Gautier, L., Cope, L., Bolstad, B. M. and Irizarry, R. A. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-15.

Gee, J. M., Robertson, J. F., Ellis, I. O., Nicholson, R. I. and Hurst, H. C. (1999). Immunohistochemical analysis reveals a tumour suppressor-like role for the transcription factor AP-2 in invasive breast cancer. *J Pathol* **189**, 514-20.

Genome Browser (url). University of California, Satna Clara. Available from: http://genome.ucsc.edu/.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80.

Gille, J., Swerlick, R. A. and Caughman, S. W. (1997). Transforming growth factor-alpha-induced transcriptional activation of the vascular permeability factor (VPF/VEGF) gene requires AP-2-dependent DNA binding and transactivation. *Embo J* **16**, 750-9.

Girdwood, D., Bumpass, D., Vaughan, O. A., Thain, A., Anderson, L. A., Snowden, A. W., Garcia-Wilson, E., Perkins, N. D. and Hay, R. T. (2003). P300 transcriptional repression is mediated by SUMO modification. *Mol Cell* **11**, 1043-54.

Goodman, R. H. and Smolik, S. (2000). CBP/p300 in cell growth, transformation, and development. *Genes Dev* **14**, 1553-77.

GOstat (url). Beissbarth, T, Available from: http://gostat.wehi.edu.au/.

Grewal, S. I. and Jia, S. (2007). Heterochromatin revisited. *Nat Rev Genet* **8**, 35-46.

Gruvberger, S., Ringner, M., Chen, Y., Panavally, S., Saal, L. H., Borg, A., Ferno, M., Peterson, C. and Meltzer, P. S. (2001). Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* **61**, 5979-84.

Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L. and Amati, B. (2006). Myc-binding-site recognition in the human genome is determined by chromatin context. *Nat Cell Biol* **8**, 764-70.

Gunawardane, R. N., Sgroi, D. C., Wrobel, C. N., Koh, E., Daley, G. Q. and Brugge, J. S. (2005). Novel role for PDEF in epithelial cell migration and invasion. *Cancer Res* **65**, 11572-80.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* **100**, 57-70.

Harborth, J., Elbashir, S. M., Bechert, K., Tuschl, T. and Weber, K. (2001). Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J Cell Sci* **114**, 4557-65.

Harms, K. L. and Chen, X. (2005). The C terminus of p53 family proteins is a cell fate determinant. *Mol Cell Biol* **25**, 2014-30.

Harms, K. L. and Chen, X. (2007). Histone deacetylase 2 modulates p53 transcriptional activities through regulation of p53-DNA binding activity. *Cancer Res* **67**, 3145-52.

Hay, R. T. (2005). SUMO: a history of modification. *Mol Cell* **18**, 1-12.

Hennighausen, L. and Robinson, G. W. (2001). Signaling pathways in mammary gland development. *Dev Cell* **1**, 467-75.

Hennighausen, L. and Robinson, G. W. (2005). Information networks in the mammary gland. *Nat Rev Mol Cell Biol* **6**, 715-25.

Hilger-Eversheim, K., Moser, M., Schorle, H. and Buettner, R. (2000). Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. *Gene* **260**, 1-12.

Hoei-Hansen, C. E., Nielsen, J. E., Almstrup, K., Sonne, S. B., Graem, N., Skakkebaek, N. E., Leffers, H. and Meyts, E. R. (2004). Transcription factor AP-2gamma is a developmentally regulated marker of testicular carcinoma in situ and germ cell tumors. *Clin Cancer Res* **10**, 8521-30.

Hoei-Hansen, C. E., Nielsen, J. E., Almstrup, K., Sonne, S. B., Graem, N., Skakkebaek, N. E., Leffers, H. and Rajpert-De Meyts, E. (2004). Transcription factor AP-2gamma is a developmentally regulated marker of testicular carcinoma in situ and germ cell tumors. *Clin Cancer Res* **10**, 8521-30.

Hoei-Hansen, C. E., Rajpert-De Meyts, E., Carlsen, E., Almstrup, K., Leffers, H. and Skakkebaek, N. E. (2005). A subfertile patient diagnosed with testicular carcinoma in situ by immunocytological staining for AP-2gamma in semen samples: case report. *Hum Reprod* **20**, 579-82.

Hollywood, D. P. and Hurst, H. C. (1993). A novel transcription factor, OB2-1, is required for overexpression of the proto-oncogene c-erbB-2 in mammary tumour lines. *Embo J* **12**, 2369-75.

Holzschuh, J., Barrallo-Gimeno, A., Ettl, A. K., Durr, K., Knapik, E. W. and Driever, W. (2003). Noradrenergic neurons in the zebrafish hindbrain are induced by retinoic acid and require tfap2a for expression of the neurotransmitter phenotype. *Development* **130**, 5741-54.

Huang, L., Sowa, Y., Sakai, T. and Pardee, A. B. (2000). Activation of the p21WAF1/CIP1 promoter independent of p53 by the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) through the Sp1 sites. *Oncogene* **19**, 5712-9.

Huang, S., Jean, D., Luca, M., Tainsky, M. A. and Bar-Eli, M. (1998). Loss of AP-2 results in downregulation of c-KIT and enhancement of melanoma tumorigenicity and metastasis. *Embo J* **17**, 4358-69.

Huang, Z., Xu, H. and Sandell, L. (2004). Negative regulation of chondrocyte differentiation by transcription factor AP-2alpha. *J Bone Miner Res* **19**, 245-55.

Huppi, K., Martin, S. E. and Caplen, N. J. (2005). Defining and assaying RNAi in mammalian cells. *Mol Cell* **17**, 1-10.

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**, 395-8.

Ihaka, R. a. G., R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**, 299-314.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-64.

Ishida, S., Huang, E., Zuzan, H., Spang, R., Leone, G., West, M. and Nevins, J. R. (2001). Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. *Mol Cell Biol* **21**, 4684-99.

Ito, T., Nomura, S., Okada, M., Katsumata, Y., Kikkawa, F., Rogi, T., Tsujimoto, M. and Mizutani, S. (2002). Ap-2 and Ikaros regulate transcription of human placental leucine aminopeptidase/oxytocinase gene. *Biochem Biophys Res Commun* **290**, 1048-53.

Iyer, N. G., Xian, J., Chin, S. F., Bannister, A. J., Daigo, Y., Aparicio, S., Kouzarides, T. and Caldas, C. (2007). p300 is required for orderly G1/S transition in human cancer cells. *Oncogene* **26**, 21-9.

Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G. and Linsley, P. S. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**, 635-7.

Jager, R., Friedrichs, N., Heim, I., Buttner, R. and Schorle, H. (2005). Dual role of AP-2gamma in ErbB-2-induced mammary tumorigenesis. *Breast Cancer Res Treat* **90**, 273-80.

Jager, R., Werling, U., Rimpf, S., Jacob, A. and Schorle, H. (2003). Transcription factor AP-2gamma stimulates proliferation and apoptosis and impairs differentiation in a transgenic model. *Mol Cancer Res* **1**, 921-9.

Jean, D., Gershenwald, J. E., Huang, S., Luca, M., Hudson, M. J., Tainsky, M. A. and Bar-Eli, M. (1998). Loss of AP-2 results in up-regulation of MCAM/MUC18 and an increase in tumor growth and metastasis of human melanoma cells. *J Biol Chem* **273**, 16501-8.

Jiang, G., Yang, F., Sanchez, C. and Ehrlich, M. (2004). Histone modification in constitutive heterochromatin versus unexpressed euchromatin in human cells. *J Cell Biochem* **93**, 286-300.

Jin, V. X., Rabinovich, A., Squazzo, S. L., Green, R. and Farnham, P. J. (2006). A computational genomics approach to identify cis-regulatory modules from chromatin immunoprecipitation microarray data--a case study using E2F1. *Genome Res* **16**, 1585-95.

Jin, Y., Norquay, L. D., Yang, X., Gregoire, S. and Cattini, P. A. (2004). Binding of AP-2 and ETS-domain family members is associated with enhancer activity in the hypersensitive site III region of the human growth hormone/chorionic somatomammotropin locus. *Mol Endocrinol* **18**, 574-87.

Johnston, R. J. and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* **426**, 845-9.

Kaeser, M. D. and Iggo, R. D. (2002). Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity in vivo. *Proc Natl Acad Sci U S A* **99**, 95-100.

Kaeser, M. D. and Iggo, R. D. (2004). Promoter-specific p53-dependent histone acetylation following DNA damage. *Oncogene* **23**, 4007-13.

Kalkhoven, E. (2004). CBP and p300: HATs for different occasions. *Biochem Pharmacol* **68**, 1145-55.

Kannan, P., Buettner, R., Pratt, D. R. and Tainsky, M. A. (1991). Identification of a retinoic acid-inducible endogenous retroviral transcript in the human teratocarcinoma-derived cell line PA-1. *J Virol* **65**, 6343-8.

Kannan, P. and Tainsky, M. A. (1999). Coactivator PC4 mediates AP-2 transcriptional activity and suppresses ras-induced transformation dependent on AP-2 transcriptional interference. *Mol Cell Biol* **19**, 899-908.

Kannan, P., Yu, Y., Wankhade, S. and Tainsky, M. A. (1999). PolyADP-ribose polymerase is a coactivator for AP-2-mediated transcriptional activation. *Nucleic Acids Res* **27**, 866-74.

Karjalainen, J. M., Kellokoski, J. K., Eskelinen, M. J., Alhava, E. M. and Kosma, V. M. (1998). Downregulation of transcription factor AP-2 predicts poor survival in stage I cutaneous malignant melanoma. *J Clin Oncol* **16**, 3584-91.

Kaufman, C. K., Sinha, S., Bolotin, D., Fan, J. and Fuchs, E. (2002). Dissection of a complex enhancer element: maintenance of keratinocyte specificity but loss of differentiation specificity. *Mol Cell Biol* **22**, 4293-308.

Kawasaki, H., Taira, K. and Morris, K. V. (2005). siRNA Induced Transcriptional Gene Silencing in Mammalian Cells. *Cell Cycle* **4**.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res* **12**, 996-1006.

Kim, J. H. and Rho, H. M. (2002). Activation of the human transforming growth factor alpha (TGF-alpha) gene by the hepatitis B viral X protein (HBx) through AP-2 sites. *Mol Cell Biochem* **231**, 155-61.

Knight, R. D., Javidan, Y., Nelson, S., Zhang, T. and Schilling, T. (2004). Skeletal and pigment cell defects in the lockjaw mutant reveal multiple roles for zebrafish tfap2a in neural crest development. *Dev Dyn* **229**, 87-98.

Knight, R. D., Javidan, Y., Zhang, T., Nelson, S. and Schilling, T. F. (2005). AP2-dependent signals from the ectoderm regulate craniofacial development in the zebrafish embryo. *Development* **132**, 3127-38.

Knight, R. D., Nair, S., Nelson, S. S., Afshar, A., Javidan, Y., Geisler, R., Rauch, G. J. and Schilling, T. F. (2003). lockjaw encodes a zebrafish tfap2a required for early neural crest development. *Development* **130**, 5755-68.

Koster, M. I., Kim, S., Huang, J., Williams, T. and Roop, D. R. (2006). TAp63alpha induces AP-2gamma as an early event in epidermal morphogenesis. *Dev Biol* **289**, 253-61.

Koster, M. I., Kim, S., Mills, A. A., DeMayo, F. J. and Roop, D. R. (2004). p63 is the molecular switch for initiation of an epithelial stratification program. *Genes Dev* **18**, 126-31.

Koster, M. I. and Roop, D. R. (2004). The role of p63 in development and differentiation of the epidermis. *J Dermatol Sci* **34**, 3-9.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.

Landis, G. N., Abdueva, D., Skvortsov, D., Yang, J., Rabin, B. E., Carrick, J., Tavare, S. and Tower, J. (2004). Similar gene expression patterns characterize aging and oxidative stress in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **101**, 7663-8.

Leask, A., Byrne, C. and Fuchs, E. (1991). Transcription factor AP2 and its role in epidermal-specific gene expression. *Proc Natl Acad Sci U S A* **88**, 7948-52.

Levy, D., Adamovich, Y., Reuven, N. and Shaul, Y. (2007). The Yes-associated protein 1 stabilizes p73 by preventing Itch-mediated ubiquitination of p73. *Cell Death Differ* **14**, 743-51.

Li, B., Carey, M. and Workman, J. L. (2007). The role of chromatin during transcription. *Cell* **128**, 707-19.

Li, W. and Cornell, R. A. (2007). Redundant activities of Tfap2a and Tfap2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Dev Biol* **304**, 338-54.

Liu, G., Loraine, A. E., Shigeta, R., Cline, M., Cheng, J., Valmeekam, V., Sun, S., Kulp, D. and Siani-Rose, M. A. (2003). NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res* **31**, 82-6.

Luo, T., Matsuo-Takasaki, M., Thomas, M. L., Weeks, D. L. and Sargent, T. D. (2002). Transcription factor AP-2 is an essential and direct regulator of epidermal development in Xenopus. *Dev Biol* **245**, 136-44.

Luscher, B., Mitchell, P. J., Williams, T. and Tjian, R. (1989). Regulation of transcription factor AP-2 by the morphogen retinoic acid and by second messengers. *Genes Dev* **3**, 1507-17.

Macleod, K. F., Sherry, N., Hannon, G., Beach, D., Tokino, T., Kinzler, K., Vogelstein, B. and Jacks, T. (1995). p53-dependent and independent expression of p21 during cell growth, differentiation, and DNA damage. *Genes Dev* **9**, 935-44.

Magnaldo, T., Bernerd, F., Freedberg, I. M., Ohtsuki, M. and Blumenberg, M. (1993). Transcriptional regulators of expression of K#16, the disease-associated keratin. *DNA Cell Biol* **12**, 911-23.

Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A. E., Kel-Margoulis, O. V. et al. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**, 374-8.

Maytin, E. V., Lin, J. C., Krishnamurthy, R., Batchvarova, N., Ron, D., Mitchell, P. J. and Habener, J. F. (1999). Keratin 10 gene expression during differentiation of mouse epidermis requires transcription factors C/EBP and AP-2. *Dev Biol* **216**, 164-81.

McPherson, L. A., Loktev, A. V. and Weigel, R. J. (2002). Tumor suppressor activity of AP2alpha mediated through a direct interaction with p53. *J Biol Chem* **277**, 45028-33.

McPherson, L. A. and Weigel, R. J. (1999). AP2alpha and AP2gamma: a comparison of binding site specificity and trans-activation of the estrogen receptor promoter and single site promoter constructs. *Nucleic Acids Res* **27**, 4040-9.

Meredith, J. (2006). Activation and Repression of Gene Expression by The AP-2 Family of Transcription Factors, vol. PhD Thesis (ed.: University of London.

Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabkin, S., Guo, N., Muruganujan, A., Doremieux, O., Campbell, M. J. et al. (2005). The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* **33**, D284-8.

Minks, M. A., West, D. K., Benvin, S. and Baglioni, C. (1979). Structural requirements of double-stranded RNA for the activation of 2',5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. *J Biol Chem* **254**, 10180-3.

Mitchell, D. C., Abdelrahim, M., Weng, J., Stafford, L. J., Safe, S., Bar-Eli, M. and Liu, M. (2006). Regulation of KiSS-1 metastasis suppressor gene expression in breast cancer cells by direct interaction of transcription factors activator protein-2alpha and specificity protein-1. *J Biol Chem* **281**, 51-8.

Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W. and Tjian, R. (1991). Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes Dev* **5**, 105-19.

Mitchell, P. J., Wang, C. and Tjian, R. (1987). Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. *Cell* **50**, 847-61.

Miyashita, T. and Reed, J. C. (1995). Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell* **80**, 293-9.

Mohibullah, N., Donner, A., Ippolito, J. A. and Williams, T. (1999). SELEX and missing phosphate contact analyses reveal flexibility within the AP-2[alpha] protein: DNA binding complex. *Nucleic Acids Res* **27**, 2760-9.

Monge, I. and Mitchell, P. J. (1998). DAP-2, the Drosophila homolog of transcription factor AP-2. *Mech Dev* **76**, 191-5.

Moser, M., Imhof, A., Pscherer, A., Bauer, R., Amselgruber, W., Sinowatz, F., Hofstadter, F., Schule, R. and Buettner, R. (1995). Cloning and characterization of a second AP-2 transcription factor: AP-2 beta. *Development* **121**, 2779-88.

Moser, M., Pscherer, A., Roth, C., Becker, J., Mucher, G., Zerres, K., Dixkens, C., Weis, J., Guay-Woodford, L., Buettner, R. et al. (1997). Enhanced apoptotic cell death of renal epithelial cells in mice lacking transcription factor AP-2beta. *Genes Dev* **11**, 1938-48.

Moser, M., Ruschoff, J. and Buettner, R. (1997). Comparative analysis of AP-2 alpha and AP-2 beta gene expression during murine embryogenesis. *Dev Dyn* **208**, 115-24.

Naef, F., Hacker, C. R., Patil, N. and Magnasco, M. (2002). Empirical characterization of the expression ratio noise structure in high-density oligonucleotide arrays. *Genome Biol* **3**, RESEARCH0018.

Nair, R. R., Solway, J. and Boyd, D. D. (2006). Expression cloning identifies transgelin (SM22) as a novel repressor of 92-kDa type IV collagenase (MMP-9) expression. *J Biol Chem* **281**, 26424-36.

Nelson, J. D., Denisenko, O. and Bomsztyk, K. (2006). Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat Protoc* **1**, 179-85.

Newman, S. P., Bates, N. P., Vernimmen, D., Parker, M. G. and Hurst, H. C. (2000). Cofactor competition between the ligand-bound oestrogen receptor and an intron 1 enhancer leads to oestrogen repression of ERBB2 expression in breast cancer. *Oncogene* **19**, 490-7.

Nottoli, T., Hagopian-Donaldson, S., Zhang, J., Perkins, A. and Williams, T. (1998). AP-2-null cells disrupt morphogenesis of the eye, face, and limbs in chimeric mice. *Proc Natl Acad Sci U S A* **95**, 13714-9.

Nyormoi, O. and Bar-Eli, M. (2003). Transcriptional regulation of metastasis-related genes in human melanoma. *Clin Exp Metastasis* **20**, 251-63.

Nyormoi, O., Wang, Z., Doan, D., Ruiz, M., McConkey, D. and Bar-Eli, M. (2001). Transcription factor AP-2alpha is preferentially cleaved by caspase 6 and degraded by proteasome during tumor necrosis factor alpha-induced apoptosis in breast cancer cells. *Mol Cell Biol* **21**, 4856-67.

Orso, F., Cottone, E., Hasleton, M. D., Ibbitt, J. C., Sismondi, P., Hurst, H. C. and De Bortoli, M. (2004). Activator protein-2gamma (AP-2gamma) expression is specifically induced by oestrogens through binding of the oestrogen receptor to a canonical element within the 5'-untranslated region. *Biochem J* **377**, 429-38.

Overall, C. M. and Kleifeld, O. (2006). Tumour microenvironment - opinion: validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. *Nat Rev Cancer* **6**, 227-39.

Paddison, P. J., Caudy, A. A. and Hannon, G. J. (2002). Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* **99**, 1443-8.

PANTHER Classification System (url). Available from: http://www.pantherdb.org/.

Park, K. and Kim, K. H. (1993). The site of cAMP action in the insulin induction of gene expression of acetyl-CoA carboxylase is AP-2. *J Biol Chem* **268**, 17811-9.

Pauls, K., Jager, R., Weber, S., Wardelmann, E., Koch, A., Buttner, R. and Schorle, H. (2005). Transcription factor AP-2gamma, a novel marker of gonocytes and seminomatous germ cell tumors. *Int J Cancer.*

Pear, W. S., Nolan, G. P., Scott, M. L. and Baltimore, D. (1993). Production of high-titer helper-free retroviruses by transient transfection. *Proc Natl Acad Sci U S A* **90**, 8392-6.

Pelengaris, S., Khan, M. and Evan, G. (2002). c-MYC: more than just a matter of life and death. *Nat Rev Cancer* **2**, 764-76.

Pellikainen, J., Kataja, V., Ropponen, K., Kellokoski, J., Pietilainen, T., Bohm, J., Eskelinen, M. and Kosma, V. M. (2002). Reduced nuclear expression of transcription factor AP-2 associates with aggressive breast cancer. *Clin Cancer Res* **8**, 3487-95.

Perissi, V., Menini, N., Cottone, E., Capello, D., Sacco, M., Montaldo, F. and De Bortoli, M. (2000). AP-2 transcription factors in the regulation of ERBB2 gene transcription by oestrogen. *Oncogene* **19**, 280-8.

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A. et al. (2000). Molecular portraits of human breast tumours. *Nature* **406**, 747-52.

Persengiev, S. P., Zhu, X. and Green, M. R. (2004). Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). *Rna* **10**, 12-8.

Pfisterer, P., Ehlermann, J., Hegen, M. and Schorle, H. (2002). A subtractive gene expression screen suggests a role of transcription factor AP-2 alpha in control of proliferation and differentiation. *J Biol Chem* **277**, 6637-44.

Polager, S., Kalma, Y., Berkovich, E. and Ginsberg, D. (2002). E2Fs up-regulate expression of genes involved in DNA replication, DNA repair and mitosis. *Oncogene* **21**, 437-46.

Pugh, B. F. (2006). HATs off to PIC assembly. *Mol Cell* **23**, 776-7.

Qiu, Y., Zhao, Y., Becker, M., John, S., Parekh, B. S., Huang, S., Hendarwanto, A., Martinez, E. D., Chen, Y., Lu, H. et al. (2006). HDAC1 acetylation is linked to progressive modulation of steroid receptor-induced gene transcription. *Mol Cell* **22**, 669-79.

Ragione, F. D., Russo, G. L., Oliva, A., Mercurio, C., Mastropietro, S., Pietra, V. D. and Zappia, V. (1996). Biochemical characterization of p16INK4- and p18-containing complexes in human cell lines. *J Biol Chem* **271**, 15942-9.

Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R. A. and Dynlacht, B. D. (2002). E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev* **16**, 245-56.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W. S. and Khvorova, A. (2004). Rational siRNA design for RNA interference. *Nat Biotechnol* **22**, 326-30.

Richardson, B. D., Langland, R. A., Bachurski, C. J., Richards, R. G., Kessler, C. A., Cheng, Y. H. and Handwerger, S. (2000). Activator protein-2 regulates human placental lactogen gene expression. *Mol Cell Endocrinol* **160**, 183-92.

Robinson, G. W. (2004). Identification of signaling pathways in early mammary gland development by mouse genetics. *Breast Cancer Res* **6**, 105-8.

Ropponen, K. M., Kellokoski, J. K., Pirinen, R. T., Moisio, K. I., Eskelinen, M. J., Alhava, E. M. and Kosma, V. M. (2001). Expression of transcription factor AP-2 in colorectal adenomas and adenocarcinomas; comparison of immunohistochemistry and in situ hybridisation. *J Clin Pathol* **54**, 533-8.

Ruiz, M., Pettaway, C., Song, R., Stoeltzing, O., Ellis, L. and Bar-Eli, M. (2004). Activator protein 2alpha inhibits tumorigenicity and represses vascular endothelial growth factor transcription in prostate cancer cells. *Cancer Res* **64**, 631-8.

Samejima, K. and Earnshaw, W. C. (2005). Trashing the genome: the role of nucleases during apoptosis. *Nat Rev Mol Cell Biol* **6**, 677-88.

Samuel, C. E. (2001). Antiviral actions of interferons. *Clin Microbiol Rev* **14**, 778-809, table of contents.

Satoda, M., Zhao, F., Diaz, G. A., Burn, J., Goodship, J., Davidson, H. R., Pierpont, M. E. and Gelb, B. D. (2000). Mutations in TFAP2B cause Char syndrome, a familial form of patent ductus arteriosus. *Nat Genet* **25**, 42-6.

Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-70.

Schorle, H., Meier, P., Buchert, M., Jaenisch, R. and Mitchell, P. J. (1996). Transcription factor AP-2 essential for cranial closure and craniofacial development. *Nature* **381**, 235-8.

Schramke, V. and Allshire, R. (2003). Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. *Science* **301**, 1069-74.

Schreiber, V., Dantzer, F., Ame, J. C. and de Murcia, G. (2006). Poly(ADP-ribose): novel functions for an old molecule. *Nat Rev Mol Cell Biol* **7**, 517-28.

Schuur, E. R., McPherson, L. A., Yang, G. P. and Weigel, R. J. (2001). Genomic structure of the promoters of the human estrogen receptor-alpha gene demonstrate changes in chromatin structure induced by AP2gamma. *J Biol Chem* **276**, 15519-26.

Schwartz, B., Melnikova, V. O., Tellez, C., Mourad-Zeidan, A., Blehm, K., Zhao, Y. J., McCarty, M., Adam, L. and Bar-Eli, M. (2007). Loss of AP-2alpha results in deregulation of E-cadherin and MMP-9 and an increase in tumorigenicity of colon cancer cells in vivo. *Oncogene*.

Semizarov, D., Frost, L., Sarthy, A., Kroeger, P., Halbert, D. N. and Fesik, S. W. (2003). Specificity of short interfering RNA determined through gene expression signatures. *Proc Natl Acad Sci U S A* **100**, 6347-52.

Shan, B., Xu, J., Zhuo, Y., Morris, C. A. and Morris, G. F. (2003). Induction of p53-dependent activation of the human proliferating cell nuclear antigen gene in chromatin by ionizing radiation. *J Biol Chem* **278**, 44009-17.

Shapland, C., Hsuan, J. J., Totty, N. F. and Lawson, D. (1993). Purification and properties of transgelin: a transformation and shape change sensitive actin-gelling protein. *J Cell Biol* **121**, 1065-73.

Sheaff, R. J., Groudine, M., Gordon, M., Roberts, J. M. and Clurman, B. E. (1997). Cyclin E-CDK2 is a regulator of p27Kip1. *Genes Dev* **11**, 1464-78.

Shen, H., Wilke, T., Ashique, A. M., Narvey, M., Zerucha, T., Savino, E., Williams, T. and Richman, J. M. (1997). Chicken transcription factor AP-2: cloning, expression and its role in outgrowth of facial prominences and limb buds. *Dev Biol* **188**, 248-66.

Sherr, C. J. and Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev* **13**, 1501-12.

Shi, D. and Kellems, R. E. (1998). Transcription factor AP-2gamma regulates murine adenosine deaminase gene expression during placental development. *J Biol Chem* **273**, 27331-8.

Shi, H., Hoffman, B. E. and Lis, J. T. (1999). RNA aptamers as effective protein antagonists in a multicellular organism. *Proc Natl Acad Sci U S A* **96**, 10033-8.

Shields, J. M., Rogers-Graham, K. and Der, C. J. (2002). Loss of transgelin in breast and colon tumors and in RIE-1 cells by Ras deregulation of gene expression through Raf-independent pathways. *J Biol Chem* **277**, 9790-9.

Siegel, P. M., Ryan, E. D., Cardiff, R. D. and Muller, W. J. (1999). Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. *Embo J* **18**, 2149-64.

Silha, J. V., Sheppard, P. C., Mishra, S., Gui, Y., Schwartz, J., Dodd, J. G. and Murphy, L. J. (2006). Insulin-like growth factor (IGF) binding protein-3 attenuates prostate tumor growth by IGF-dependent and IGF-independent mechanisms. *Endocrinology* **147**, 2112-21.

Skinner, A. and Hurst, H. C. (1993). Transcriptional regulation of the c-erbB-3 gene in human breast carcinoma cell lines. *Oncogene* **8**, 3393-401.

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. et al. (1989). Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* **244**, 707-12.

Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. and Williams, B. R. (2003). Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* **5**, 834-9.

Sledz, C. A. and Williams, B. R. (2004). RNA interference and double-stranded-RNA-activated pathways. *Biochem Soc Trans* **32**, 952-6.

Snowden, A. W., Anderson, L. A., Webster, G. A. and Perkins, N. D. (2000). A novel transcriptional repression domain mediates p21(WAF1/CIP1) induction of p300 transactivation. *Mol Cell Biol* **20**, 2676-86.

Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S. et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **98**, 10869-74.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S. et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* **100**, 8418-23.

Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B. et al. (2006). Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* **98**, 262-72.

Stabach, P. R., Thiyagarajan, M. M., Woodfield, G. W. and Weigel, R. J. (2006). AP2alpha alters the transcriptional activity and stability of p53. *Oncogene* **25**, 2148-59.

Sternlicht, M. D. (2006). Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. *Breast Cancer Res* **8**, 201.

Sumigama, S., Ito, T., Kajiyama, H., Shibata, K., Tamakoshi, K., Kikkawa, F., Williams, T., Tainsky, M. A., Nomura, S. and Mizutani, S. (2004). Suppression of invasion and peritoneal carcinomatosis of ovarian cancer cells by overexpression of AP-2alpha. *Oncogene* **23**, 5496-504.

Talbot, D., Loring, J. and Schorle, H. (1999). Spatiotemporal expression pattern of keratins in skin of AP-2alpha-deficient mice. *J Invest Dermatol* **113**, 816-20.

Tellez, C. and Bar-Eli, M. (2003). Role and regulation of the thrombin receptor (PAR-1) in human melanoma. *Oncogene* **22**, 3130-7.

Tellez, C., McCarty, M., Ruiz, M. and Bar-Eli, M. (2003). Loss of activator protein-2alpha results in overexpression of protease-activated receptor-1 and correlates with the malignant phenotype of human melanoma. *J Biol Chem* **278**, 46632-42.

The Gene Ontology Project (url). Available from: http://www.geneontology.org/.

Thomas, M. C. and Chiang, C. M. (2005). E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation. *Mol Cell* **17**, 251-64.

Thompson, P. R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R. G., Wong, J. et al. (2004). Regulation of the p300 HAT domain via a novel activation loop. *Nat Struct Mol Biol* **11**, 308-15.

Transfac (url). Transfac Matrix Database (v.7.0). Available from: http://www.gene-regulation.com/pub/databases.html#transfac.

Turner, B. C., Zhang, J., Gumbs, A. A., Maher, M. G., Kaplan, L., Carter, D., Glazer, P. M., Hurst, H. C., Haffty, B. G. and Williams, T. (1998). Expression of AP-2 transcription factors in human breast cancer correlates with the regulation of multiple growth factor signalling pathways. *Cancer Res* **58**, 5466-72.

Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* **13**, 3191-7.

Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R. and Saigo, K. (2004). Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res* **32**, 936-48.

van de Rijn, M., Perou, C. M., Tibshirani, R., Haas, P., Kallioniemi, O., Kononen, J., Torhorst, J., Sauter, G., Zuber, M., Kochli, O. R. et al. (2002). Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. *Am J Pathol* **161**, 1991-6.

van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J. et al. (2002). A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* **347**, 1999-2009.

Varshochi, R., Halim, F., Sunters, A., Alao, J. P., Madureira, P. A., Hart, S. M., Ali, S., Vigushin, D. M., Coombes, R. C. and Lam, E. W. (2005). ICI182,780 induces p21Waf1 gene transcription through releasing histone deacetylase 1 and estrogen receptor alpha from Sp1 sites to induce cell cycle arrest in MCF-7 breast cancer cell line. *J Biol Chem* **280**, 3185-96.

Vernimmen, D., Begon, D., Salvador, C., Gofflot, S., Grooteclaes, M. and Winkler, R. (2003). Identification of HTF (HER2 transcription factor) as an AP-2 (activator protein-2) transcription factor and contribution of the HTF binding site to ERBB2 gene overexpression. *Biochem J* **370**, 323-9.

Wajapeyee, N., Britto, R., Ravishankar, H. M. and Somasundaram, K. (2006). Apoptosis induction by activator protein 2alpha involves transcriptional repression of Bcl-2. *J Biol Chem* **281**, 16207-19.

Wajapeyee, N., Raut, C. G. and Somasundaram, K. (2005). Activator protein 2alpha status determines the chemosensitivity of cancer cells: implications in cancer chemotherapy. *Cancer Res* **65**, 8628-34.

Wajapeyee, N. and Somasundaram, K. (2003). Cell cycle arrest and apoptosis induction by activator protein 2alpha (AP-2alpha) and the role of p53 and p21WAF1/CIP1 in AP-2alpha-mediated growth inhibition. *J Biol Chem* **278**, 52093-101.

Wang, D., Shin, T. H. and Kudlow, J. E. (1997). Transcription factor AP-2 controls transcription of the human transforming growth factor-alpha gene. *J Biol Chem* **272**, 14244-50.

Wang, X., Bolotin, D., Chu, D. H., Polak, L., Williams, T. and Fuchs, E. (2006). AP-2alpha: a regulator of EGF receptor signaling and proliferation in skin epidermis. *J Cell Biol* **172**, 409-21.

Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M. E., Yu, J. et al. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **365**, 671-9.

Wankhade, S., Yu, Y., Weinberg, J., Tainsky, M. A. and Kannan, P. (2000). Characterization of the activation domains of AP-2 family transcription factors. *J Biol Chem* **275**, 29701-8.

Wanner, R., Zhang, J., Dorbic, T., Mischke, D., Henz, B. M., Wittig, B. and Rosenbach, T. (1997). The promoter of the HaCaT keratinocyte differentiation-related gene keratin 4 contains a functional AP-2 binding site. *Arch Dermatol Res* **289**, 705-8.

Werling, U. and Schorle, H. (2002). Conditional inactivation of transcription factor AP-2gamma by using the Cre/loxP recombination system. *Genesis* **32**, 127-9.

Werling, U. and Schorle, H. (2002). Transcription factor gene AP-2 gamma essential for early murine development. *Mol Cell Biol* **22**, 3149-56.

Whitfield, M. L., Sherlock, G., Saldanha, A. J., Murray, J. I., Ball, C. A., Alexander, K. E., Matesė, J. C., Perou, C. M., Hurt, M. M., Brown, P. O. et al. (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell* **13**, 1977-2000.

Whitfield, M. L., Sherlock, G., Saldanha, A. J., Murray, J. I., Ball, C. A., Alexander, K. E., Matese, J. C., Perou, C. M., Hurt, M. M., Brown, P. O. et al. (url). Supplemental Data, Available from: http://genome-www.stanford.edu/Human-CellCycle/Hela/.

Williams, T., Admon, A., Luscher, B. and Tjian, R. (1988). Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. *Genes Dev* **2**, 1557-69.

Williamson, J. A., Bosher, J. M., Skinner, A., Sheer, D., Williams, T. and Hurst, H. C. (1996). Chromosomal mapping of the human and mouse homologues of two new members of the AP-2 family of transcription factors. *Genomics* **35**, 262-4.

Wilson, C. L. and Miller, C. J. (2005). Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. *Bioinformatics* **21**, 3683-5.

Winger, Q., Huang, J., Auman, H. J., Lewandoski, M. and Williams, T. (2006). Analysis of transcription factor AP-2 expression and function during mouse preimplantation development. *Biol Reprod* **75**, 324-33.

Winning, R. S., Shea, L. J., Marcus, S. J. and Sargent, T. D. (1991). Developmental regulation of transcription factor AP-2 during Xenopus laevis embryogenesis. *Nucleic Acids Res* **19**, 3709-14.

Wooster, R. and Weber, B. L. (2003). Breast and ovarian cancer. *N Engl J Med* **348**, 2339-47.

Wu, Q., Kirschmeier, P., Hockenberry, T., Yang, T. Y., Brassard, D. L., Wang, L., McClanahan, T., Black, S., Rizzi, G., Musco, M. L. et al. (2002). Transcriptional regulation during p21WAF1/CIP1-induced apoptosis in human ovarian cancer cells. *J Biol Chem* **277**, 36329-37.

Xiao, H., Hasegawa, T. and Isobe, K. (2000). p300 collaborates with Sp1 and Sp3 in p21(waf1/cip1) promoter activation induced by histone deacetylase inhibitor. *J Biol Chem* **275**, 1371-6.

Yagi, R., Chen, L. F., Shigesada, K., Murakami, Y. and Ito, Y. (1999). A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *Embo J* **18**, 2551-62.

Yang, S. H. and Sharrocks, A. D. (2004). SUMO promotes HDAC-mediated transcriptional repression. *Mol Cell* **13**, 611-7.

Zeng, Y. X., Somasundaram, K. and el-Deiry, W. S. (1997). AP2 inhibits cancer cell growth and activates p21WAF1/CIP1 expression. *Nat Genet* **15**, 78-82.

Zhang, J., Brewer, S., Huang, J. and Williams, T. (2003). Overexpression of transcription factor AP-2alpha suppresses mammary gland growth and morphogenesis. *Dev Biol* **256**, 127-45.

Zhang, J., Hagopian-Donaldson, S., Serbedzija, G., Elsemore, J., Plehn-Dujowich, D., McMahon, A. P., Flavell, R. A. and Williams, T. (1996). Neural tube, skeletal and body wall defects in mice lacking transcription factor AP-2. *Nature* **381**, 238-41.

Zhao, C., Yasui, K., Lee, C. J., Kurioka, H., Hosokawa, Y., Oka, T. and Inazawa, J. (2003). Elevated expression levels of NCOA3, TOP1, and TFAP2C in breast tumors as predictors of poor prognosis. *Cancer* **98**, 18-23.

Zhao, F., Satoda, M., Licht, J. D., Hayashizaki, Y. and Gelb, B. D. (2001). Cloning and characterization of a novel mouse AP-2 transcription factor, AP-2delta, with unique DNA binding and transactivation properties. *J Biol Chem* **276**, 40755-60.

Zhao, F., Weismann, C. G., Satoda, M., Pierpont, M. E., Sweeney, E., Thompson, E. M. and Gelb, B. D. (2001). Novel TFAP2B mutations that cause Char syndrome provide a genotype-phenotype correlation. *Am J Hum Genet* **69**, 695-703.

Zhu, Y., Wang, A., Liu, M. C., Zwart, A., Lee, R. Y., Gallagher, A., Wang, Y., Miller, W. R., Dixon, J. M. and Clarke, R. (2006). Estrogen receptor alpha positive breast tumors and breast cancer cell lines share similarities in their transcriptome data structures. *Int J Oncol* **29**, 1581-9.

Zhuang, L., Lee, C. S., Scolyer, R. A., McCarthy, S. W., Zhang, X. D., Thompson, J. F. and Hersey, P. (2007). Mcl-1, Bcl-XL and Stat3 expression are associated with progression of melanoma whereas Bcl-2, AP-2 and MITF levels decrease during progression of melanoma. *Mod Pathol* **20**, 416-26.

APPENDIX 1

Appendix 1 displays the 578 probe sets that changed their expression significantly (FDR corrected, p value <0.05) between the *reference* (non silencing control siRNA and cells treated with transfection reagent only) and *test* (AP-2γ targeting siRNAs: AP-2γ siRNA1, AP-2γ siRNA2 and AP-2γ siRNA3) groups. The table displays the Affymetrix Probe Set ID, Gene Symbol, Gene Name, Fold Change (*test* vs. *reference*) and the associated statistical significance.

Probe Set ID	Gene Symbol	Gene Name	Fold Change	p value
232306_at	CDH26	cadherin-like 26	2.07	8.43E-06
224925_at	PREX1	phosphatidylinositol trisphosphate-dependent RAC exchanger 1	-1.89	8.43E-06
221874_at	KIAA1324	KIAA1324	-2.57	8.43E-06
235408_x_at	ZNF117	Zinc finger protein 117 (HPF9)	2.01	1.09E-05
225207_at	PDK4	pyruvate dehydrogenase kinase, isozyme 4	1.96	1.09E-05
228176_at	EDG3	endothelial differentiation, sphingolipid G-protein-coupled receptor, 3	-2.49	1.09E-05
204244_s_at	DBF4	DBF4 homolog (S. cerevisiae)	-2.08	1.09E-05
203571_s_at	C10orf116	chromosome 10 open reading frame 116	-2.53	1.09E-05
210511_s_at	INHBA	inhibin, beta A (activin A, activin AB alpha polypeptide)	1.97	1.56E-05
201458_s_at	BUB3	BUB3 budding uninhibited by benzimidazoles 3 homolog	-1.77	1.56E-05
203060_s_at	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	2.18	1.92E-05
210868_s_at	ELOVL6	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	-1.72	2.62E-05
243010_at	MSI2	musashi homolog 2 (Drosophila)	-2.01	3.15E-05
220414_at	CALML5	calmodulin-like 5	-1.90	3.15E-05
209754_s_at	TMPO	thymopoietin	-1.78	3.76E-05
222557_at	STMN3	stathmin-like 3	-1.97	3.76E-05
220192_x_at	SPDEF	SAM pointed domain containing ets transcription factor	-2.24	3.76E-05
204070_at	RARRES3	retinoic acid receptor responder (tazarotene induced) 3	-2.06	3.76E-05
203304_at	BAMBI	BMP and activin membrane-bound inhibitor	-2.10	4.41E-05
201896_s_at	PSRC1	proline/serine-rich coiled-coil 1	-1.78	5.29E-05
227140_at	INHBA	inhibin, beta A (activin A, activin AB alpha polypeptide)	3.06	5.36E-05
212148_at	PBX1	Pre-B-cell leukemia transcription factor 1	-1.79	5.36E-05
227850_x_at	CDC42EP5	CDC42 effector protein (Rho GTPase binding) 5	-2.16	5.36E-05
239007_at	ZNF616	zinc finger protein 616	1.87	5.38E-05
203058_s_at	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	1.90	5.38E-05
212143_s_at	IGFBP3	insulin-like growth factor binding protein 3	2.30	5.38E-05
213599_at	OIP5	Opa interacting protein 5	-1.72	5.38E-05
205240_at	GPSM2	G-protein signalling modulator 2 (AGS3-like, C. elegans)	-1.78	5.38E-05
210095_s_at	IGFBP3	insulin-like growth factor binding protein 3	2.39	5.90E-05
204319_s_at	RGS10	regulator of G-protein signalling 10	-1.75	5.90E-05
210297_s_at	MSMB	microseminoprotein, beta-	-2.11	6.39E-05
209291_at	ID4	inhibitor of DNA binding 4	2.28	6.57E-05
201830_s_at	NET1	neuroepithelial cell transforming gene 1	-1.85	6.57E-05
207076_s_at	ASS1	argininosuccinate synthetase 1	-2.30	6.57E-05
227099_s_at	LOC387763	hypothetical LOC387763	2.88	7.45E-05
226248_s_at	KIAA1324	KIAA1324	-2.55	7.62E-05
224734_at	HMGB1	high-mobility group box 1	-1.95	7.62E-05
242245_at	SYDE2	Synapse defective 1, Rho GTPase, homolog 2 (C. elegans)	-2.25	7.98E-05
213906_at	MYBL1	v-myb myeloblastosis viral oncogene homolog (avian)-like 1	-2.00	7.98E-05
209301_at	CA2	carbonic anhydrase II	-2.53	7.98E-05
204602_at	DKK1	dickkopf homolog 1 (Xenopus laevis)	-1.85	9.42E-05
217744_s_at	PERP	PERP, TP53 apoptosis effector	1.81	0.000106
201829_at	NET1	neuroepithelial cell transforming gene 1	-1.79	0.000106
219918_s_at	ASPM	asp (abnormal spindle)-like, microcephaly associated	-2.10	0.000106

205380_at	PDZK1	PDZ domain containing 1	-2.59	0.000109
204709_s_at	KIF23	kinesin family member 23	-1.84	0.000113
229305_at	MLF1IP	MLF1 interacting protein	-1.79	0.000125
218542_at	CEP55	centrosomal protein 55kDa	-2.03	0.000125
202870_s_at	CDC20	CDC20 cell division cycle 20 homolog (S. cerevisiae)	-2.04	0.000125
228729_at	CCNB1	cyclin B1	-2.10	0.000143
212314_at	KIAA0746	KIAA0746 protein	1.83	0.000158
213253_at	SMC2L1	SMC2 structural maintenance of chromosomes 2-like 1 (yeast)	-1.87	0.000158
204886_at	PLK4	polo-like kinase 4 (Drosophila)	-1.66	0.000158
221258_s_at	KIF18A	kinesin family member 18A /// kinesin family member 18A	-1.87	0.000158
228273_at	---	Transcribed locus	-2.38	0.000158
204033_at	TRIP13	thyroid hormone receptor interactor 13	-1.86	0.000163
228597_at	C21orf45	chromosome 21 open reading frame 45	-1.85	0.000166
209684_at	RIN2	Ras and Rab interactor 2	1.87	0.000167
211080_s_at	NEK2	NIMA (never in mitosis gene a)-related kinase 2	-1.73	0.00017
211708_s_at	SCD	stearoyl-CoA desaturase (delta-9-desaturase)	-1.72	0.000178
218755_at	KIF20A	kinesin family member 20A	-2.26	0.000187
224909_s_at	PREX1	phosphatidylinositol 3,4,5-trisphosphate-dependent RAC exchanger 1	-1.98	0.000193
209343_at	EFHD1	EF-hand domain family, member D1	-1.79	0.000195
242138_at	DLX1	distal-less homeobox 1	-1.72	0.000195
202183_s_at	KIF22	kinesin family member 22	-1.86	0.000198
240633_at	DOK7	docking protein 7	-1.68	0.000225
235564_at	ZNF117	Zinc finger protein 117 (HPF9)	2.01	0.000228
201292_at	TOP2A	topoisomerase (DNA) II alpha 170kDa	-2.28	0.000228
201456_s_at	BUB3	BUB3 budding uninhibited by benzimidazoles 3	-1.90	0.000228
204446_s_at	ALOX5	arachidonate 5-lipoxygenase	1.70	0.000238
240572_s_at	LOC374443	CLR pseudogene	-1.81	0.000245
219004_s_at	C21orf45	chromosome 21 open reading frame 45	-1.68	0.000246
203554_x_at	PTTG1	pituitary tumor-transforming 1	-2.02	0.000254
219540_at	ZNF267	zinc finger protein 267	1.70	0.000275
213711_at	KRT81	keratin 81	-4.02	0.000275
219294_at	CENPQ	centromere protein Q	-1.99	0.000275
212958_x_at	PAM	peptidylglycine alpha-amidating monooxygenase	1.66	0.000276
236641_at	KIF14	kinesin family member 14	-1.86	0.000276
225710_at	---	CDNA FLJ34013 fis, clone FCBBF2002111	-1.68	0.000276
235775_at	TMTC2	transmembrane and tetratricopeptide repeat containing 2	2.02	0.000303
1554614_a_at	PTBP2	polypyrimidine tract binding protein 2	-1.63	0.000312
207695_s_at	IGSF1	immunoglobulin superfamily, member 1	1.73	0.000314
202766_s_at	FBN1	fibrillin 1	1.71	0.000314
237435_at	---	Transcribed locus	-1.86	0.00032
213664_at	SLC1A1	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1	-1.77	0.000339
206204_at	GRB14	growth factor receptor-bound protein 14	1.90	0.00034
227452_at	LOC146795	Hypothetical protein LOC146795	1.77	0.000354
204141_at	TUBB2A	tubulin, beta 2A	2.23	0.000358
218683_at	PTBP2	polypyrimidine tract binding protein 2	-1.64	0.000382
216568_x_at	KRT8	Keratin 8	-1.66	0.000383
1555758_a_at	CDKN3	cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)	-2.23	0.000392
213154_s_at	BICD2	bicaudal D homolog 2 (Drosophila)	1.68	0.000408
221726_at	RPL22	ribosomal protein L22	-1.67	0.000424
203343_at	UGDH	UDP-glucose dehydrogenase	-1.90	0.000452
225687_at	FAM83D	family with sequence similarity 83, member D	-1.87	0.000455
243349_at	KIAA1324	KIAA1324	-2.08	0.000502
204641_at	NEK2	NIMA (never in mitosis gene a)-related kinase 2	-2.14	0.000505
209714_s_at	CDKN3	cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)	-2.25	0.000505
231855_at	KIAA1524	KIAA1524	-1.63	0.000516
201897_s_at	CKS1B	CDC28 protein kinase regulatory subunit 1B	-2.00	0.000532
224719_s_at	C12orf57	chromosome 12 open reading frame 57	-1.90	0.000532
202565_s_at	SVIL	supervillin	1.75	0.000573
206364_at	KIF14	kinesin family member 14	-1.94	0.000574

230165_at	SGOL2	shugoshin-like 2 (S. pombe)	-1.69	0.000594
219306_at	KIF15	kinesin family member 15	-1.63	0.000616
209715_at	CBX5	chromobox homolog 5 (HP1 alpha homolog, Drosophila)	-1.87	0.000638
219529_at	CLIC3	chloride intracellular channel 3	-1.82	0.000742
219032_x_at	OPN3	opsin 3 (encephalopsin, panopsin)	1.96	0.000751
207430_s_at	MSMB	microseminoprotein, beta-	-2.17	0.00077
206115_at	EGR3	early growth response 3	-1.91	0.000777
211812_s_at	B3GALNT1	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	-1.72	0.000777
229964_at	C9orf152	chromosome 9 open reading frame 152	-1.63	0.000792
203276_at	LMNB1	lamin B1	-2.12	0.000814
223374_s_at	B3GALNT1	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	-1.90	0.000946
229309_at	---	Beta-1 adrenergic receptor mRNA, 3' UTR	1.90	0.00101
232155_at	KIAA1618	KIAA1618	-1.59	0.00101
227165_at	C13orf3	chromosome 13 open reading frame 3	-1.64	0.00103
204058_at	ME1	Malic enzyme 1, NADP(+)-dependent, cytosolic	1.74	0.00112
209581_at	HRASLS3	HRAS-like suppressor 3	-1.77	0.00112
214710_s_at	CCNB1	cyclin B1	-2.10	0.00112
211513_s_at	OGFR	opioid growth factor receptor	1.64	0.00113
205110_s_at	FGF13	fibroblast growth factor 13	1.78	0.00115
202705_at	CCNB2	cyclin B2	-1.93	0.00117
228574_at	TMTC2	transmembrane and tetratricopeptide repeat containing 2	1.77	0.0012
212281_s_at	TMEM97	transmembrane protein 97	-1.74	0.0012
228323_at	CASC5	cancer susceptibility candidate 5	-1.63	0.00122
213682_at	NUP50	nucleoporin 50kDa	-1.67	0.00123
222039_at	SLC35E1	solute carrier family 35, member E1	-1.73	0.00126
226936_at	C6orf173	chromosome 6 open reading frame 173	-1.80	0.00126
211162_x_at	SCD	stearoyl-CoA desaturase (delta-9-desaturase)	-1.65	0.00131
231894_at	SARS	Seryl-tRNA synthetase	-1.74	0.00131
232278_s_at	DEPDC1	DEP domain containing 1	-1.67	0.00131
225297_at	CCDC5	coiled-coil domain containing 5 (spindle associated)	-1.71	0.00131
211379_x_at	B3GALNT1	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	-1.81	0.00131
203856_at	VRK1	vaccinia related kinase 1	-1.72	0.00147
226980_at	DEPDC1B	DEP domain containing 1B	-1.72	0.00147
224944_at	TMPO	thymopoietin	-1.73	0.00148
203418_at	CCNA2	cyclin A2	-1.95	0.00149
209366_x_at	CYB5A	cytochrome b5 type A (microsomal)	-1.64	0.00151
210052_s_at	TPX2	TPX2, microtubule-associated, homolog (Xenopus laevis)	-1.82	0.00152
235456_at	---	CDNA clone IMAGE:4819084	1.69	0.00156
204128_s_at	RFC3	replication factor C (activator 1) 3, 38kDa	-1.95	0.00156
201663_s_at	SMC4	structural maintenance of chromosomes 4	-1.80	0.00159
210983_s_at	MCM7	MCM7 minichromosome maintenance deficient 7	-1.73	0.00164
208712_at	CCND1	cyclin D1	-1.77	0.00164
213931_at	ID2	inhibitor of DNA binding 2	1.80	0.00167
202336_s_at	PAM	peptidylglycine alpha-amidating monooxygenase	1.69	0.00169
204727_at	WDHD1	WD repeat and HMG-box DNA binding protein 1	-1.66	0.00169
202489_s_at	FXYD3	FXYD domain containing ion transport regulator 3	-2.12	0.00171
204235_s_at	GULP1	GULP, engulfment adaptor PTB domain containing 1	1.73	0.00177
209172_s_at	CENPF	centromere protein F, 350/400ka (mitosin)	-1.72	0.00177
204962_s_at	CENPA	centromere protein A	-1.89	0.00177
219555_s_at	CENPN	centromere protein N	-1.57	0.00188
209373_at	MALL	mal, T-cell differentiation protein-like	2.31	0.00192
238756_at	GAS2L3	Growth arrest-specific 2 like 3	-1.62	0.00193
213624_at	SMPDL3A	sphingomyelin phosphodiesterase, acid-like 3A	-1.58	0.00194
201525_at	APOD	apolipoprotein D	-2.06	0.00194
204237_at	GULP1	GULP, engulfment adaptor PTB domain containing 1	1.68	0.00195
222699_s_at	PLEKHF2	pleckstrin homology domain containing, family F (with FYVE domain) member 2	-1.59	0.00198
237160_at	CCDC83	coiled-coil domain containing 83	-1.73	0.00199
207843_x_at	CYB5A	cytochrome b5 type A (microsomal)	-1.65	0.00206
201368_at	ZFP36L2	zinc finger protein 36, C3H type-like 2	-1.66	0.00222
209408_at	KIF2C	kinesin family member 2C	-1.58	0.00226

201540_at	FHL1	four and a half LIM domains 1	2.10	0.00235
209774_x_at	CXCL2	chemokine (C-X-C motif) ligand 2	2.47	0.00236
235113_at	PPIL5	peptidylprolyl isomerase (cyclophilin)-like 5	-1.62	0.0024
203764_at	DLG7	discs, large homolog 7 (Drosophila)	-2.16	0.00245
208988_at	FBXL11	F-box and leucine-rich repeat protein 11	1.67	0.00249
226140_s_at	OTUD1	OTU domain containing 1	1.65	0.00257
228286_at	FLJ40869	hypothetical protein FLJ40869	-1.67	0.00257
205003_at	DOCK4	dedicator of cytokinesis 4	1.89	0.00283
201664_at	SMC4	structural maintenance of chromosomes 4	-1.96	0.00283
204822_at	TTK	TTK protein kinase	-1.88	0.0029
208808_s_at	HMGB2	high-mobility group box 2	-1.93	0.0029
1552619_a_at	ANLN	anillin, actin binding protein	-1.71	0.0029
203967_at	CDC6	CDC6 cell division cycle 6 homolog (S. cerevisiae)	-1.67	0.00296
223839_s_at	---	PRO1933	-1.70	0.00305
204127_at	RFC3	replication factor C (activator 1) 3, 38kDa	-1.77	0.00323
210299_s_at	FHL1	four and a half LIM domains 1	2.07	0.00328
204444_at	KIF11	kinesin family member 11	-2.17	0.00329
218640_s_at	PLEKHF2	pleckstrin homology domain containing, family F (with FYVE domain) member 2	-1.70	0.00337
202954_at	PAK3	p21 (CDKN1A)-activated kinase 3	-1.77	0.00344
215189_at	KRT86	keratin 86	-2.12	0.00344
202625_at	LYN	v-yes-1 Yamaguchi sarcoma viral related oncogene homolog	1.62	0.0036
1554696_s_at	TYMS	thymidylate synthetase	-1.68	0.0036
203210_s_at	RFC5	replication factor C (activator 1).5, 36.5kDa	-1.61	0.0036
218723_s_at	RGC32	response gene to complement 32	1.75	0.00378
224866_at	MLSTD2	male sterility domain containing 2	-1.54	0.00398
235425_at	SGOL2	shugoshin-like 2 (S. pombe)	-1.54	0.00399
228281_at	FLJ25416	hypothetical protein FLJ25416	-1.73	0.00401
227350_at	HELLS	Helicase, lymphoid-specific	-1.71	0.00402
223381_at	CDCA1	cell division cycle associated 1	-2.00	0.00413
203217_s_at	ST3GAL5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	1.55	0.00422
222608_s_at	ANLN	anillin, actin binding protein	-1.77	0.00423
203640_at	MBNL2	muscleblind-like 2 (Drosophila)	-1.56	0.00429
205909_at	POLE2	polymerase (DNA directed), epsilon 2 (p59 subunit)	-1.65	0.00433
213088_s_at	DNAJC9	DnaJ (Hsp40) homolog, subfamily C, member 9	-1.62	0.00433
235419_at	ERRFI1	ERBB receptor feedback inhibitor 1	1.62	0.00439
205287_s_at	TFAP2C	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	-2.09	0.00439
229551_x_at	ZNF367	zinc finger protein 367	-1.95	0.0044
205234_at	SLC16A4	solute carrier family 16, member 4 (monocarboxylic acid transporter 5)	-2.23	0.00454
228242_at	---	---	-1.60	0.00462
216237_s_at	MCM5	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	-1.69	0.00487
1554768_a_at	MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	-2.24	0.00487
212112_s_at	STX12	syntaxin 12	1.74	0.00488
222077_s_at	RACGAP1	Rac GTPase activating protein 1	-1.71	0.00489
209104_s_at	NOLA2	nucleolar protein family A, member 2 (H/ACA small nucleolar RNPs)	-1.55	0.00489
244804_at	SQSTM1	Sequestosome 1	1.74	0.00492
210028_s_at	ORC3L	origin recognition complex, subunit 3-like (yeast)	-1.53	0.00504
218726_at	DKFZp762E1312	hypothetical protein DKFZp762E1312	-1.51	0.00508
228640_at	---	CDNA clone IMAGE:4800096	1.63	0.00509
213008_at	KIAA1794	KIAA1794	-1.57	0.00509
212185_x_at	MT2A	metallothionein 2A	-1.96	0.00511
203362_s_at	MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	-2.07	0.00511
203887_s_at	THBD	thrombomodulin	1.68	0.00518
217996_at	PHLDA1	pleckstrin homology-like domain, family A, member 1	1.81	0.00531
213092_x_at	DNAJC9	DnaJ (Hsp40) homolog, subfamily C, member 9	-1.69	0.00531
230360_at	GLDN	gliomedin	1.75	0.00541
202859_x_at	IL8	interleukin 8	1.87	0.00544
222848_at	CENPK	centromere protein K	-1.88	0.00545
205339_at	STIL	SCL/TAL1 interrupting locus	-1.60	0.00552

227180_at	ELOVL7	ELOVL family member 7, elongation of long chain fatty acids (yeast)	-1.49	0.00552
202095_s_at	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	-1.54	0.00552
230964_at	FREM2	FRAS1 related extracellular matrix protein 2	-1.76	0.0057
204092_s_at	AURKA	aurora kinase A	-1.85	0.00572
201041_s_at	DUSP1	dual specificity phosphatase 1	1.54	0.00601
204146_at	RAD51AP1	RAD51 associated protein 1	-1.92	0.00601
218990_s_at	SPRR3	small proline-rich protein 3	2.88	0.0061
203209_at	RFC5	replication factor C (activator 1) 5, 36.5kDa	-1.66	0.00611
207165_at	HMMR	hyaluronan-mediated motility receptor (RHAMM)	-1.87	0.00617
225834_at	LOC652689	hypothetical protein LOC652689	-2.11	0.00625
204472_at	GEM	GTP binding protein overexpressed in skeletal muscle	2.60	0.00629
201739_at	SGK	serum/glucocorticoid regulated kinase	1.81	0.00635
203022_at	RNASEH2A	ribonuclease H2, subunit A	-1.59	0.00635
219148_at	PBK	PDZ binding kinase	-2.43	0.00635
242560_at	FANCD2	Fanconi anemia, complementation group D2	-1.61	0.00635
216598_s_at	CCL2	chemokine (C-C motif) ligand 2	1.50	0.00637
211623_s_at	FBL	fibrillarin	-1.54	0.00642
218875_s_at	FBXO5	F-box protein 5	-1.64	0.00643
201539_s_at	FHL1	four and a half LIM domains 1	1.56	0.00645
204204_at	SLC31A2	solute carrier family 31 (copper transporters), member 2	1.56	0.00646
205286_at	TFAP2C	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	-2.39	0.00646
224788_at	ARF6	ADP-ribosylation factor 6	-1.52	0.00646
209464_at	AURKB	aurora kinase B	-1.52	0.00684
202413_s_at	USP1	ubiquitin specific peptidase 1	-1.52	0.00691
226456_at	C16orf75	chromosome 16 open reading frame 75	-1.57	0.00692
218883_s_at	MLF1IP	MLF1 interacting protein	-1.89	0.00695
201563_at	SORD	sorbitol dehydrogenase	-1.55	0.00702
227196_at	RHPN2	rhophilin, Rho GTPase binding protein 2	-2.03	0.00702
213226_at	CCNA2	cyclin A2	-1.89	0.00702
228033_at	E2F7	E2F transcription factor 7	-1.49	0.00705
218662_s_at	HCAP-G	chromosome condensation protein G	-1.66	0.00721
212021_s_at	MKI67	antigen identified by monoclonal antibody Ki-67	-1.61	0.00723
223229_at	UBE2T	ubiquitin-conjugating enzyme E2T (putative)	-1.70	0.00724
204240_s_at	SMC2	structural maintenance of chromosomes 2	-1.77	0.00736
205694_at	TYRP1	tyrosinase-related protein 1	1.55	0.0074
205016_at	TGFA	transforming growth factor, alpha	1.78	0.0074
236224_at	RIT1	Ras-like without CAAX 1	1.51	0.00751
205394_at	CHEK1	CHK1 checkpoint homolog (S. pombe)	-1.63	0.00752
204256_at	ELOVL6	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	-1.58	0.00761
218039_at	NUSAP1	nucleolar and spindle associated protein 1	-1.89	0.00772
204112_s_at	HNMT	histamine N-methyltransferase	-1.70	0.00772
205807_s_at	TUFT1	tuftelin 1	1.90	0.00775
242338_at	TMEM64	transmembrane protein 64	-1.55	0.00775
218009_s_at	PRC1	protein regulator of cytokinesis 1	-1.97	0.00775
205393_s_at	CHEK1	CHK1 checkpoint homolog (S. pombe)	-1.73	0.00781
204510_at	CDC7	CDC7 cell division cycle 7 (S. cerevisiae)	-1.55	0.00786
206461_x_at	MT1H	metallothionein 1H	-1.68	0.00793
211959_at	IGFBP5	insulin-like growth factor binding protein 5	2.18	0.00817
219471_at	C13orf18	chromosome 13 open reading frame 18	1.56	0.00817
202567_at	SNRPD3	small nuclear ribonucleoprotein D3 polypeptide 18kDa	-1.51	0.00817
203214_x_at	CDC2	cell division cycle 2, G1 to S and G2 to M	-2.03	0.00817
232238_at	ASPM	asp (abnormal spindle)-like, microcephaly associated	-1.63	0.00817
214505_s_at	FHL1	four and a half LIM domains 1	1.51	0.00826
218807_at	VAV3	vav 3 oncogene	-1.70	0.00826
205053_at	PRIM1	primase, polypeptide 1, 49kDa	-1.60	0.00826
220651_s_at	MCM10	MCM10 minichromosome maintenance deficient 10	-1.57	0.00826
204542_at	ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	-1.64	0.00827
1559060_a_at	KIAA1961	KIAA1961 gene	1.77	0.00836
219978_s_at	NUSAP1	nucleolar and spindle associated protein 1	-1.66	0.00836
222118_at	CENPN	centromere protein N	-1.55	0.00839

229256_at	PGM2L1	phosphoglucomutase 2-like 1	1.89	0.0143
226621_at	---	---	1.77	0.0143
201504_s_at	TSN	translin	-1.51	0.0143
203432_at	TMPO	thymopoietin	-1.48	0.0143
214042_s_at	RPL22	ribosomal protein L22	-1.52	0.0143
231981_at	PRLR	Prolactin receptor	-1.74	0.0143
218768_at	NUP107	nucleoporin 107kDa	-1.49	0.0143
235178_x_at	ESCO2	establishment of cohesion 1 homolog 2 (S. cerevisiae)	-1.63	0.0143
1553984_s_at	DTYMK	deoxythymidylate kinase (thymidylate kinase)	-1.45	0.0143
219000_s_at	DCC1	defective in sister chromatid cohesion homolog 1	-1.53	0.0143
226790_at	MORN2	MORN repeat containing 2	2.17	0.0146
241652_x_at	LIN7A	Lin-7 homolog A (C. elegans)	1.48	0.0146
201202_at	PCNA	proliferating cell nuclear antigen	-1.55	0.0146
218294_s_at	NUP50	nucleoporin 50kDa	-1.57	0.0146
204105_s_at	NRCAM	neuronal cell adhesion molecule	-1.50	0.0146
206205_at	MPHOSPH9	M-phase phosphoprotein 9	-1.51	0.0146
231121_at	HPS3	Hermansky-Pudlak syndrome 3	-1.49	0.0146
236223_s_at	RIT1	Ras-like without CAAX 1	1.51	0.0147
204825_at	MELK	maternal embryonic leucine zipper kinase	-1.55	0.0147
205018_s_at	MBNL2	muscleblind-like 2 (Drosophila)	-1.49	0.0147
201324_at	EMP1	epithelial membrane protein 1	-1.51	0.0147
221667_s_at	HSPB8	heat shock 22kDa protein 8	1.74	0.0148
204162_at	KNTC2	kinetochore associated 2	-1.94	0.0151
226189_at	ITGB8	integrin, beta 8	2.37	0.0152
224764_at	ARHGAP21	Rho GTPase activating protein 21	1.49	0.0152
224989_at	---	Full-length cDNA clone CS0DE011YO20 of Placenta	-1.53	0.0153
224830_at	NUDT21	nucleoside diphosphate linked moiety X-type motif 21	-1.47	0.0154
220840_s_at	C1orf112	chromosome 1 open reading frame 112	-1.53	0.0155
218806_s_at	VAV3	vav 3 oncogene	-1.62	0.0159
212022_s_at	MKI67	antigen identified by monoclonal antibody Ki-67	-1.59	0.0159
232654_s_at	UGT1A6	UDP glucuronosyltransferase 1 family, polypeptide A6	1.48	0.0163
209270_at	LAMB3	laminin, beta 3	1.53	0.0163
225706_at	GLCCI1	glucocorticoid induced transcript 1	1.58	0.0163
218117_at	RBX1	ring-box 1	-1.48	0.0165
218349_s_at	ZWILCH	Zwilch, kinetochore associated, homolog (Drosophila)	-1.56	0.0169
202267_at	LAMC2	laminin, gamma 2	1.44	0.0173
219493_at	SHCBP1	SHC SH2-domain binding protein 1	-1.78	0.0173
209687_at	CXCL12	chemokine (C-X-C motif) ligand 12	-1.47	0.0173
228468_at	MASTL	microtubule associated serine/threonine kinase-like	-1.51	0.0175
1562836_at	---	CDNA FLJ11653 fis, clone HEMBA1004538	-1.45	0.0176
1554168_a_at	SH3KBP1	SH3-domain kinase binding protein 1	1.45	0.0178
212141_at	MCM4	MCM4 minichromosome maintenance deficient 4	-1.50	0.0178
223556_at	HELLS	helicase, lymphoid-specific	-1.61	0.0178
223758_s_at	GTF2H2	general transcription factor IIH, polypeptide 2, 44kDa	-1.49	0.0178
209676_at	TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	1.63	0.0179
206316_s_at	KNTC1	kinetochore associated 1	-1.46	0.0179
211712_s_at	ANXA9	annexin A9 /// annexin A9	-1.52	0.0179
205590_at	RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated)	1.84	0.0181
217678_at	SLC7A11	solute carrier family 7, (cationic amino acid transporter, y+ system) member 11	1.54	0.0184
203404_at	ARMCX2	armadillo repeat containing, X-linked 2	1.69	0.0185
209187_at	DR1	down-regulator of transcription 1, TBP-binding (negative cofactor 2)	-1.52	0.0185
222740_at	ATAD2	ATPase family, AAA domain containing 2	-1.64	0.0186
200934_at	DEK	DEK oncogene (DNA binding)	-1.44	0.0187
225489_at	TMEM18	transmembrane protein 18	-1.45	0.0188
237968_at	ARL6IP2	ADP-ribosylation factor-like 6 interacting protein 2	-1.48	0.0189
227003_at	RAB28	RAB28, member RAS oncogene family	-1.50	0.0193
202903_at	LSM5	LSM5 homolog, U6 small nuclear RNA associated	-1.53	0.0197
221563_at	DUSP10	dual specificity phosphatase 10	1.49	0.0198
222680_s_at	DTL	denticleless homolog (Drosophila)	-1.47	0.0198
1554314_at	C6orf141	chromosome 6 open reading frame 141	1.52	0.0201

220083_x_at	UCHL5	ubiquitin carboxyl-terminal hydrolase L5	-1.54	0.0201
200831_s_at	SCD	stearoyl-CoA desaturase (delta-9-desaturase)	-1.60	0.0201
203176_s_at	TFAM	transcription factor A, mitochondrial	-1.46	0.0206
223700_at	MND1	meiotic nuclear divisions 1 homolog (S. cerevisiae)	-1.71	0.0206
218585_s_at	DTL	denticleless homolog (Drosophila)	-1.45	0.0206
1557765_at	LOC643401	hypothetical protein LOC643401	1.49	0.0207
227525_at	GLCCI1	glucocorticoid induced transcript 1	1.57	0.0207
242836_at	ATP1B3	ATPase, Na+/K+ transporting, beta 3 polypeptide	1.61	0.0207
202219_at	SLC6A8	solute carrier family 6, member 8	1.55	0.0208
227998_at	S100A16	S100 calcium binding protein A16	-1.72	0.0212
225700_at	GLCCI1	glucocorticoid induced transcript 1	1.43	0.0214
228401_at	ATAD2	ATPase family, AAA domain containing 2	-1.44	0.0215
215913_s_at	GULP1	GULP, engulfment adaptor PTB domain containing 1	1.49	0.0216
226320_at	THOC4	THO complex 4	-1.61	0.0216
202107_s_at	MCM2	MCM2 minichromosome maintenance deficient 2,	-1.64	0.0216
220060_s_at	C12orf48	chromosome 12 open reading frame 48	-1.52	0.0216
215780_s_at	SET	SET translocation (myeloid leukemia-associated)	-1.46	0.0217
204768_s_at	FEN1	flap structure-specific endonuclease 1	-1.50	0.0217
203665_at	HMOX1	heme oxygenase (decycling) 1	2.01	0.0219
209084_s_at	RAB28	RAB28, member RAS oncogene family	-1.51	0.0219
228834_at	TOB1	transducer of ERBB2, 1	-1.45	0.0224
209337_at	PSIP1	PC4 and SFRS1 interacting protein 1	-1.55	0.0227
214507_s_at	EXOSC2	exosome component 2	-1.45	0.023
221970_s_at	NOL11	nucleolar protein 11	-1.56	0.0232
222872_x_at	OBFC2A	oligonucleotide/oligosaccharide-binding fold containing 2A	1.59	0.0235
200832_s_at	SCD	stearoyl-CoA desaturase (delta-9-desaturase)	-1.56	0.0236
238736_at	REV3L	REV3-like, catalytic subunit of DNA polymerase zeta	1.43	0.0237
222921_s_at	HEY2	hairy/enhancer-of-split related with YRPW motif 2	-1.48	0.0237
229744_at	SSFA2	Sperm specific antigen 2	-1.56	0.0242
209921_at	SLC7A11	solute carrier family 7, (cationic amino acid transporter, y+ system) member 11	1.47	0.0243
200628_s_at	WARS	tryptophanyl-tRNA synthetase	-1.56	0.0245
230261_at	ST8SIA4	ST8 α-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	1.49	0.0246
217788_s_at	GALNT2	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 2 (GalNAc-T2)	1.45	0.0246
211421_s_at	RET	Ret proto-oncogene (multiple endocrine neoplasia and medullary thyroid carcinoma 1, Hirschsprung disease)	-1.62	0.0246
206632_s_at	APOBEC3B	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	-1.53	0.0246
201295_s_at	WSB	WD repeat and SOCS box-containing 1	1.86	0.0247
1558014_s_at	MLSTD2	male sterility domain containing 2	-1.47	0.0248
205034_at	CCNE2	cyclin E2	-1.60	0.0249
209344_at	TPM4	tropomyosin 4	1.95	0.0251
202503_s_at	KIAA0101	KIAA0101	-1.58	0.0251
224731_at	HMGB1	high-mobility group box 1	-1.58	0.0251
228692_at	---	CDNA FLJ13569 fis, clone PLACE1008369	-1.52	0.0251
230748_at	SLC16A6	solute carrier family 16, member 6	-1.60	0.0253
225429_at	---	---	-1.41	0.0253
202436_s_at	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	1.62	0.0256
235004_at	RBM24	RNA binding motif protein 24	-1.42	0.0256
210735_s_at	CA12	carbonic anhydrase XII	-1.42	0.0258
204767_s_at	FEN1	flap structure-specific endonuclease 1	-1.49	0.026
202149_at	NEDD9	neural precursor cell expressed, developmentally down-regulated 9	1.48	0.0261
204798_at	MYB	v-myb myeloblastosis viral oncogene homolog (avian)	-1.68	0.0263
218195_at	C6orf211	chromosome 6 open reading frame 211	-1.50	0.0267
238890_at	---	---	-1.50	0.0268
201975_at	RSN	Reed-Steinberg cell-expressed intermediate filament-associated protein	1.52	0.0269
224221_s_at	VAV3	vav 3 oncogene	-1.82	0.0271
209932_s_at	DUT	dUTP pyrophosphatase	-1.55	0.0271
209642_at	BUB1	budding uninhibited by benzimidazoles 1	-1.64	0.0271
203998_s_at	SYT1	synaptotagmin I	1.49	0.0274
201477_s_at	RRM1	ribonucleotide reductase M1 polypeptide	-1.51	0.0276
203200_s_at	MTRR	methyltetrahydrofolate methyltransferase reductase	-1.40	0.0283

231579_s_at	TIMP2	TIMP metallopeptidase inhibitor 2	1.63	0.0284
212111_at	STX12	syntaxin 12	1.46	0.0284
235976_at	SLITRK6	SLIT and NTRK-like family, member 6	1.53	0.0287
238002_at	GOLPH4	golgi phosphoprotein 4	-1.42	0.0287
217989_at	DHRS8	dehydrogenase/reductase (SDR family) member 8	-1.46	0.0287
202901_x_at	CTSS	cathepsin S	-1.69	0.0287
201012_at	ANXA1	annexin A1	2.45	0.0288
222037_at	MCM4	MCM4 minichromosome maintenance deficient 4	-1.65	0.0288
206500_s_at	C14orf106	chromosome 14 open reading frame 106	-1.43	0.0288
218498_s_at	ERO1L	ERO1-like (S. cerevisiae)	1.55	0.0289
208886_at	H1F0	H1 histone family, member 0	-1.44	0.0289
235716_at	---	Transcribed locus	1.49	0.029
208827_at	PSMB6	proteasome (prosome, macropain) subunit, beta type, 6	-1.40	0.029
237460_x_at	LOC283551	hypothetical protein LOC283551	-1.62	0.029
205083_at	AOX1	aldehyde oxidase 1	-1.40	0.0291
218493_at	C16orf33	chromosome 16 open reading frame 33	-1.63	0.0292
206658_at	UPK3B	uroplakin 3B	1.54	0.0298
214455_at	HIST1H2BC	histone 1, H2bc	1.70	0.0298
203231_s_at	ATXN1	ataxin 1	1.48	0.0298
1552609_s_at	IL28A	interleukin 28A (interferon, lambda 2)	-1.53	0.0298
1566096_x_at	ARHGEF12	Rho guanine nucleotide exchange factor (GEF) 12	1.39	0.0302
213007_at	KIAA1794	KIAA1794	-1.63	0.0302
235698_at	ZFP90	zinc finger protein 90 homolog (mouse)	1.52	0.0304
222646_s_at	ERO1L	ERO1-like (S. cerevisiae)	1.54	0.0306
219588_s_at	LUZP5	leucine zipper protein 5	-1.47	0.031
243296_at	PBEF1	Pre-B-cell colony enhancing factor 1	1.59	0.0311
208596_s_at	UGT1A10	UDP glucuronosyltransferase 1 family, polypeptide A10	1.46	0.0312
207928_s_at	GLRA3	glycine receptor, alpha 3	1.82	0.0312
205822_s_at	HMGCS1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1	-1.43	0.0312
207628_s_at	WBSCR22	Williams Beuren syndrome chromosome region 22	-1.43	0.0318
231772_x_at	CENPH	centromere protein H	-1.52	0.0321
213293_s_at	TRIM22	tripartite motif-containing 22	-1.67	0.0322
201930_at	MCM6	MCM6 minichromosome maintenance deficient 6	-1.58	0.0323
204238_s_at	C6orf108	chromosome 6 open reading frame 108	-1.39	0.0328
210592_s_at	SAT1	spermidine/spermine N1-acetyltransferase 1	1.52	0.0331
1554408_a_at	TK1	thymidine kinase 1, soluble	-1.59	0.0331
202551_s_at	CRIM1	cysteine rich transmembrane BMP regulator 1 (chordin-like)	1.39	0.0336
202338_at	TK1	thymidine kinase 1, soluble	-1.44	0.034
1557366_at	KIAA0565	KIAA0565 gene product	1.44	0.0343
224579_at	---	CDNA FLJ32108 fis, clone OCBBF2001492	-1.39	0.0345
214606_at	TSPAN2	tetraspanin 2	1.50	0.0346
230333_at	SAT	Spermidine/spermine N1-acetyltransferase	1.41	0.0347
238825_at	ACRC	acidic repeat containing	1.44	0.0351
208175_s_at	DMP1	dentin matrix acidic phosphoprotein	-1.39	0.0351
1560622_at	TPARL	TPA regulated locus	1.41	0.0358
211596_s_at	LRIG1	leucine-rich repeats and immunoglobulin-like domains 1	-1.39	0.036
211429_s_at	SERPINA1	serpin peptidase inhibitor, clade A member 1	1.71	0.0364
209861_s_at	METAP2	methionyl aminopeptidase 2	-1.39	0.0367
223315_at	NTN4	netrin 4	1.67	0.0368
200972_at	TSPAN3	tetraspanin 3	1.42	0.0373
219888_at	SPAG4	sperm associated antigen 4	1.49	0.0376
204222_s_at	GLIPR1	GLI pathogenesis-related 1 (glioma)	2.26	0.0376
205176_s_at	ITGB3BP	integrin beta 3 binding protein (beta3-endonexin)	-1.42	0.0377
204083_s_at	TPM2	tropomyosin 2 (beta)	1.97	0.0384
207996_s_at	C18orf1	chromosome 18 open reading frame 1	1.41	0.0388
218979_at	RMI1	RMI1, RecQ mediated genome instability 1	-1.39	0.0395
203650_at	PROCR	protein C receptor, endothelial (EPCR)	-1.38	0.0396
236193_at	HIST1H2BC	histone 1, H2bc	1.42	0.0403
210085_s_at	ANXA9	annexin A9	-1.40	0.0408
225239_at	---	CDNA FLJ26120 fis, clone SYN00419	1.66	0.0409
229553_at	PGM2L1	phosphoglucomutase 2-like 1	1.63	0.0411
201667_at	GJA1	gap junction protein, alpha 1, 43kDa (connexin 43)	1.97	0.0413
226077_at	FLJ31951	hypothetical protein FLJ31951	1.47	0.0413

200081_s_at	RPS6	ribosomal protein S6 /// ribosomal protein S6	-1.39	0.0413
206346_at	PRLR	prolactin receptor	-1.53	0.0413
1553810_a_at	KIAA1524	KIAA1524	-1.58	0.0413
202435_s_at	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	1.66	0.0414
201294_s_at	WSB1	WD repeat and SOCS box-containing 1	1.45	0.0416
214056_at	MCL1	Myeloid cell leukemia sequence 1 (BCL2-related)	-1.42	0.0416
220329_s_at	C6orf96	chromosome 6 open reading frame 96	-1.48	0.0416
211506_s_at	IL8	interleukin 8	1.98	0.0418
226430_at	LOC253981	hypothetical protein LOC253981	-1.42	0.0418
231202_at	ALDH1L2	aldehyde dehydrogenase 1 family, member L2	1.40	0.042
1561951_at	SLC5A12	solute carrier family 5, member 12	1.42	0.0423
228825_at	LTB4DH	leukotriene B4 12-hydroxydehydrogenase	1.39	0.0423
203721_s_at	UTP18	UTP18, small subunit (SSU) processome component	-1.40	0.0423
211814_s_at	CCNE2	cyclin E2	-1.64	0.0423
1554224_at	---	CDNA clone IMAGE:4794631	1.45	0.0425
209377_s_at	HMGN3	high mobility group nucleosomal binding domain 3	-1.48	0.0425
200679_x_at	HMGB1	high-mobility group box 1	-1.47	0.0426
217025_s_at	DBN1	drebrin 1	1.45	0.0427
219363_s_at	MTERFD1	MTERF domain containing 1	-1.36	0.0427
218162_at	OLFML3	olfactomedin-like 3	1.56	0.0429
244094_at	C9orf93	Chromosome 9 open reading frame 93	-1.38	0.0429
201656_at	ITGA6	integrin, alpha 6	1.39	0.0432
1552306_at	ALG10	asparagine-linked glycosylation 10 homolog	-1.56	0.0436
225189_s_at	RAPH1	Ras association and pleckstrin homology domains 1	-1.56	0.0438
240383_at	UBE2D3	ubiquitin-conjugating enzyme E2D 3	1.43	0.044
210015_s_at	MAP2	microtubule-associated protein 2	1.43	0.044
244447_at	KLF10	Kruppel-like factor 10	1.42	0.044
201117_s_at	CPE	carboxypeptidase E	1.60	0.044
1557915_s_at	GSTO1	glutathione S-transferase omega 1	-1.40	0.0447
206363_at	MAF	v-maf musculoaponeurotic fibrosarcoma oncogene	1.37	0.0451
208840_s_at	G3BP2	RasGTPase activating protein SH3domain-binding protein 2	1.41	0.0451
201476_s_at	RRM1	ribonucleotide reductase M1 polypeptide	-1.47	0.0451
214113_s_at	RBM8A	RNA binding motif protein 8A	-1.40	0.0451
1567107_s_at	TPM4	tropomyosin 4	1.50	0.0464
223551_at	PKIB	protein kinase (cAMP-dependent, catalytic) inhibitor beta	-1.47	0.0464
221521_s_at	GINS2	GINS complex subunit 2 (Psf2 homolog)	-1.48	0.0464
208721_s_at	ANAPC5	anaphase promoting complex subunit 5	-1.44	0.0464
223156_at	MRPS23	mitochondrial ribosomal protein S23	-1.38	0.0466
208669_s_at	CRI1	CREBBP/EP300 inhibitor 1	-1.61	0.0475
202619_s_at	PLOD2	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	1.45	0.0482
201614_s_at	RUVBL1	RuvB-like 1 (E. coli)	-1.42	0.0485
204326_x_at	MT1X	metallothionein 1X	-1.49	0.0485
225540_at	MAP2	microtubule-associated protein 2	1.87	0.0486
225957_at	LOC153222	adult retina protein	1.52	0.0486
239845_at	---	Transcribed locus	1.37	0.0486
207181_s_at	CASP7	caspase 7, apoptosis-related cysteine peptidase	1.49	0.0487
203545_at	ALG8	asparagine-linked glycosylation 8 homolog	-1.43	0.0487
208828_at	POLE3	polymerase (DNA directed), epsilon 3 (p17 subunit)	-1.37	0.0493
210561_s_at	WSB1	WD repeat and SOCS box-containing 1	1.38	0.0494
233085_s_at	OBFC2A	oligonucleotide/oligosaccharide-binding fold containing 2A	1.51	0.0497
218907_s_at	LRRC61	leucine rich repeat containing 61	1.45	0.0497
209773_s_at	RRM2	ribonucleotide reductase M2 polypeptide	-1.89	0.0497
218802_at	CCDC109B	coiled-coil domain containing 109B	-1.38	0.0497
229128_s_at	---	Transcribed locus	-1.37	0.0497

APPENDIX 2

An AP-2 consensus-binding site analysis conducted using the "HumanGenome9999" human promoter sequence data (9999bp up stream sequences of 21787 genes represented in the ENSEMBL database) imported into the Genespring package. A subset of 192 genes out of the 254 probe sets that changed their expression significantly (FDR corrected, p value <0.01) following AP-2γ silencing in MCF-7 cells were successfully mapped to the ENSEMBL promoter sequences. The presence of the AP-2 consensus binding sequence (SCCNNVRGB) was mapped in a –500 to +50 bp upstream region. The table shows the location relative to the start of transcription and the potential AP-2 TFBS sequence.

Gene Symbol	Location	Sequence
ALOX5	91	GCCTGGAGC
ALOX5	98	GCCGGGAGC
ALOX5	193	GCCCGAGGC
ALOX5	216	CCCAGGAGC
ALOX5	359	GCCTTGGGC
ALOX5	374	GCCTACAGG
ALOX5	381	CCCCAGGGC
ALOX5	394	CCCCGCAGC
ANLN	137	CCCCGCGGC
ASK	303	GCCGGCAGT
ASK	370	GCCTTCAGC
ASK	417	GCCCCGGGC
ASPM	97	CCCTGAGGG
ASPM	98	CCCCTGAGG
ASPM	269	CCCGGAGGT
ASPM	270	GCCCGGAGG
ASPM	427	GCCCCGGGT
ASPM	471	CCCTTCAGC
ASS	6	GCCCCCGGC
ASS	63	CCCGGGGGC
ASS	64	GCCCGGGGG
ASS	216	CCCGCCGGC
ASS	255	CCCGCGGGC
ASS	256	GCCCGCGGG
ASS	263	GCCTCCAGC
ASS	306	CCCCTGAGT
ASS	350	GCCAGGAGG
ASS	357	CCCCAGGGC
AURKA	195	GCCACAGGT
AURKA	211	CCCTGGGGC
AURKB	326	CCCGGGAGG
B3GALT3	65	CCCAGAGGG
B3GALT3	231	GCCTCGAGG
B3GALT3	243	CCCTCCAGG
BENE	53	GCCCAGGGG
BENE	278	CCCAAGAGT
BENE	343	CCCAGGGGC
BENE	408	CCCATGAGG
ELOVL6	282	GCCGCGGGG
ELOVL6	361	GCCCGGAGT
ELOVL6	424	CCCAGAGGT
ELOVL7	27	GCCTTCAGT
ELOVL7	198	GCCTTCAGT
ERRFI1	46	CCCGGCGGC
ERRFI1	173	CCCTGCAGC
ERRFI1	298	GCCGCGGGC
ERRFI1	338	GCCCTCAGT
ERRFI1	383	GCCCGGAGC
ERRFI1	423	GCCAGCGGC
ERRFI1	450	CCCTCCAGG
ERRFI1	484	GCCGCAAGC
FANCD2	111	GCCTTCAGC
FANCD2	306	CCCAGAGGC
FBN1	47	GCCCCGGGC
FBN1	94	CCCGGCGGC
FBN1	125	GCCCGGGGG
FBN1	168	GCCTGAAGT
FBN1	208	GCCTCGGGG
FBN1	355	GCCGGGGGC
FBN1	358	GCCGCCGGG
FBN1	390	CCCGGGAGG
FBN1	449	GCCCGGGGC
FBXO5	231	GCCCCCGGC
FHL1	292	CCCCGCAGC
FHL1	361	GCCCCCGGC
FHL1	370	GCCCGAGGC
FHL1	407	CCCGCAGGC
FHL1	427	GCCTGCAGT
FLJ10719	473	GCCTCAAGG
FLJ11029	210	CCCTCCGGC
FLJ40869	-2	GCCTAAGGC
FOLR1	42	GCCTTAGGC
GEM	42	GCCGTCGGG
GEM	72	CCCGGCGGG
GEM	150	CCCTCCGGC
GEM	237	GCCACGGGT
GEM	298	CCCGACAGC

BICD2	84	GCCGGAGGG
BICD2	101	CCCAGCGGG
BICD2	104	GCCCCCAGC
BICD2	167	GCCGGGGGC
BICD2	206	CCCGAGGGC
BICD2	207	GCCCGAGGG
BICD2	237	CCCCTGGGC
BICD2	246	CCCCACAGC
BICD2	400	GCCGCGGGG
BICD2	467	GCCAGAGGC
BICD2	477	GCCGCCGGC
BIRC5	18	CCCCGCGGC
BIRC5	63	GCCGCGGGG
BIRC5	71	CCCAGAAGG
C10orf116	60	CCCTGGGGC
C10orf116	61	GCCCTGGGG
C10orf116	120	CCCCGGAGC
C10orf116	256	GCCTCCAGC
C10orf116	343	CCCTGGAGC
C10orf116	357	GCCCTGAGC
C10orf3	158	CCCAGAAGC
C10orf3	233	CCCGTAGGC
C12orf57	-47	GCCAACAGC
C13orf18	131	GCCTGCGGC
C13orf18	269	GCCGCGGGT
C16orf75	136	CCCGGCGGT
C20orf129	157	GCCAGGAGC
C20orf129	404	CCCGAGGGT
C21orf45	83	CCCTTCAGG
C4orf34	9	GCCGGCGGC
C4orf34	89	GCCCGCGGG
C4orf34	201	GCCGCAGGC
C6orf139	114	CCCTCCAGC
CA2	133	CCCCCGAGC
CA2	144	CCCCCGGGC
CA2	216	GCCCCGAGC
CA2	268	GCCGAGAGG
CA2	282	CCCGCCGGG
CA2	291	CCCCGGAGC
CA2	299	CCCCGGAGC
CA2	307	CCCCGGAGC
CA2	337	GCCGAGGGG
CA2	344	GCCGTGGGC
CA2	390	GCCCAGAGC
CA2	416	CCCAGGAGC
CA2	473	GCCGGGAGG
CALML5	0	CCCCCGGGC
CALML5	1	GCCCCCGGG
CALML5	101	CCCTGCGGC
CALML5	213	CCCCAGGGC
CALML5	396	CCCTTGGGC
CALML5	489	CCCTCCAGT
CBX5	271	GCCAGGAGG
CCDC5	23	GCCAACGGC
GEM	324	CCCCCCAGT
GEM	347	CCCGTGGGC
GEM	384	GCCTGGAGG
GFPT2	68	CCCAGAAGC
GFPT2	411	CCCTGGGGG
GFPT2	412	GCCCTGGGG
GPSM2	35	GCCCGAGGC
GPSM2	68	GCCTTGGGG
GPSM2	114	GCCGGGGGC
GPSM2	180	CCCCTAGGT
GPSM2	240	GCCTGGAGC
HCAP-G	61	GCCCAAAGC
HCAP-G	194	CCCTCCGGT
HCAP-G	208	GCCAGCAGC
HCAP-G	315	GCCACAAGG
HCAP-G	322	CCCTTCGGC
HCAP-G	431	CCCAGCAGC
HMGB2	474	CCCAGGAGC
HRASLS3	-49	CCCATGGGT
HRASLS3	184	GCCTGCAGG
ID2	128	CCCGCCGGG
ID2	290	CCCGCGGGG
ID2	291	CCCCGCGGG
ID2	358	CCCGCCAGC
IGFBP3	-11	GCCACGAGC
IGFBP3	155	CCCCGCAGC
IGFBP3	196	CCCCCGAGC
IGFBP3	309	GCCTGGAGT
IGFBP3	346	GCCAGGAGT
IGFBP3	403	CCCGAGAGC
IGFBP3	421	GCCGCAGGG
INHBA	232	GCCAGAGGT
KIAA1324	208	CCCCGGGGG
KIAA1324	217	CCCCTCAGT
KIAA1324	480	CCCCTCAGT
KIAA1618	12	CCCAGCAGG
KIF14	268	GCCTGAAGT
KIF15	387	CCCGACAGC
KIF15	442	GCCTGGGGG
KIF18A	149	GCCGCGAGC
KIF18A	282	CCCACCAGC
KIF18A	325	CCCGCCGGT
KIF18A	374	GCCCGCAGC
KIF23	141	CCCTCCAGC
KIF2C	130	CCCCAAGGC
KIF2C	131	CCCCCAAGG
KIF2C	211	CCCAGAAGG
KIF2C	281	CCCCGCAGT
KLIP1	313	GCCTTAGGT
KRT81	81	GCCCTGAGG
KRT81	262	CCCCAGAGC
KRT86	411	CCCCGCAGG
KRT86	500	CCCCCAAGT
LMNB1	-27	CCCGGAAGT

CCDC5	83	GCCAGGAGT
CCDC83	66	GCCTGGGGG
CCDC83	225	CCCTGCAGG
CCDC83	245	GCCCAGGGC
CCDC83	277	GCCGGCGGG
CCNA2	216	GCCGCGAGC
CCNB1	158	CCCGAGAGC
CCNB2	380	CCCGGAGGG
CCND1	140	GCCTCAGGG
CCND1	159	GCCGCAGGG
CCND1	448	CCCGTGGGT
CDC20	106	CCCGGAAGG
CDC20	340	CCCTGAAGC
CDC20	434	GCCTTGGGC
CDC20	473	GCCCGCGGC
CDC42EP5	410	CCCTGAAGG
CDC6	114	GCCCAAGGC
CDC6	418	CCCATCAGG
CDC7	493	GCCAAGGGC
CDCA1	1	GCCCACAGC
CDCA5	29	CCCGCAGGG
CDCA5	134	CCCCTGGGC
CDCA5	215	CCCTTCAGC
CDCA5	236	CCCCAGGGC
CDCA5	432	GCCCCCGGC
CDCA5	465	GCCTCCGGC
CDH26	2	GCCAAAGGT
CDH26	328	CCCCAGAGC
CDKN3	65	CCCAAAGGC
CDKN3	144	CCCACAGGG
CDKN3	245	CCCCCAAGC
CKS1B	86	CCCATAAGG
CKS1B	87	GCCTCGAGC
CKS1B	143	CCCATAAGG
CLIC3	61	CCCCTCGGC
CLIC3	149	CCCCGGGGC
CLIC3	165	CCCTCAGGC
CLIC3	178	CCCTGAGGT
CLIC3	406	GCCATAGGG
CXCL2	63	GCCACAGGC
CXCL2	206	CCCAGAGGC
CXCL2	207	CCCCAGAGG
CXCL2	446	CCCCCCGGT
CYB5	44	CCCGACGGG
CYB5	113	GCCCCCAGT
CYB5	132	CCCCTGAGC
CYB5	393	CCCGCGGGG
DEPDC1B	9	GCCAACGGT
DEPDC1B	44	CCCGCGGGC
DEPDC1B	45	GCCCGCGGG
DEPDC1B	52	GCCCGCAGC
DEPDC1B	125	GCCGCGGGT
DEPDC1B	158	GCCCCGGGC
DEPDC1B	228	GCCCCCGGC

LMNB1	8	CCCTCCAGC
LMNB1	29	GCCTCCAGG
MCM10	63	GCCTAAAGT
MCM10	83	CCCATGGGG
MCM5	63	GCCGCGAGG
MCM5	159	CCCGGAAGG
MCM5	201	GCCAGAAGT
MCM5	397	CCCAAGAGG
MCM7	29	GCCCTGGGC
MCM7	132	CCCGTCAGC
MCM7	207	CCCGACGGG
MCM7	264	CCCCCAAGT
MKI67	45	CCCGCGGGC
MKI67	46	CCCCGCGGG
MKI67	116	GCCACCGGG
MKI67	371	CCCGCGAGC
MSI2	71	GCCGGGGGG
MSI2	245	GCCGTGGGG
MSI2	279	CCCGCGGGG
MSI2	280	GCCCGCGGG
MSI2	301	CCCGGAGGC
MSI2	302	GCCCGGAGG
MSI2	364	CCCCGCAGC
NEK2	96	GCCTAGGGT
NEK2	152	CCCACAGGT
NEK2	217	GCCGAGAGG
NEK2	278	GCCTGAGGG
NEK2	455	CCCGGGAGT
NET1	128	CCCGGGAGC
NET1	200	GCCTTAGGG
NOLA2	126	GCCTTAAGG
NOLA2	352	CCCGGGAGG
NOLA2	493	GCCTGAGGT
NUP50	79	CCCCAAAGC
NUP50	145	CCCCTCAGC
NUP50	232	CCCTGCAGC
NUP50	261	CCCGAGGGT
NUP50	367	CCCTCCGGG
NUSAP1	283	GCCCCAAGC
NUSAP1	380	CCCCGGGGC
NUSAP1	381	CCCCCGGGG
NUSAP1	382	CCCCCCGGG
NUSAP1	489	CCCTGCGGG
NUSAP1	497	GCCCCCAGC
OGFR	494	CCCTCCGGG
OIP5	90	GCCAAAGGG
OIP5	429	CCCTCAGGT
OIP5	486	CCCCGAGGC
OPN3	28	CCCCTCAGT
OPN3	93	GCCACGGGG
OPN3	111	GCCGCGGGG
OPN3	139	GCCGGGAGC
OPN3	168	CCCGCGGGG
OPN3	169	GCCCGCGGG

DEPDC1B	365	GCCCAAAGC
DKFZp762E1312	8	CCCTGCAGT
DKFZp762E1312	137	GCCGGAAGG
DKFZp762E1312	359	CCCAGCGGC
DKK1	21	CCCTGCAGT
DKK1	74	GCCCGCGGT
DKK1	97	CCCAAGGGG
DKK1	244	CCCGAAGGT
DKK1	245	CCCCGAAGG
DKK1	383	CCCCTCGGC
DLG7	157	CCCTCGGGC
DLG7	158	CCCCTCGGG
DLG7	271	GCCTAAGGG
DOK7	-42	CCCACCGGT
DOK7	-23	GCCTGAGGT
DOK7	25	GCCAAGGGT
DOK7	117	GCCGGGAGT
DOK7	285	CCCAGGGGG
DOK7	286	GCCCAGGGG
DOK7	298	GCCCTCAGT
DOK7	327	CCCCCAAGC
DOK7	352	GCCACCGGG
DOK7	498	CCCTGGGGC
DOK7	499	CCCCTGGGG
DUSP1	37	CCCAGAGGC
DUSP1	38	CCCCAGAGG
DUSP1	210	GCCGCGAGC
DUSP1	214	GCCAAGGGG
DUSP1	245	GCCAAGAGC
DUSP1	282	GCCTGGAGC
DUSP1	323	GCCTGGAGC
DUSP1	330	CCCGGAGGT
DUSP1	428	GCCGAGGGG
DUSP1	475	GCCCTGGGC
EDG3	24	CCCGGGAGC
EDG3	56	GCCGTGGGT
EDG3	430	GCCGCCAGC
EDG3	496	CCCCTCAGG
EFHD1	74	CCCACCAGG
EFHD1	141	GCCGCGGGG
EFHD1	148	GCCCGGAGC
EFHD1	171	CCCGGGGGG
EFHD1	227	CCCCGGAGC
EFHD1	314	GCCGCGAGC
EGR3	271	CCCTTCGGC
EGR3	363	GCCGCGAGC
ELOVL6	84	GCCAAGGGT
ELOVL6	126	CCCCCGGGC
ELOVL6	127	CCCCCCGGG
ELOVL6	233	CCCACGAGG
ELOVL6	282	GCCGCGGGG

OPN3	182	GCCGCGAGG
OPN3	257	CCCCCAGGC
OPN3	258	GCCCCCAGG
OPN3	271	CCCTGCGGC
ORC3L	14	GCCGCGAGG
ORC3L	21	CCCTGCAGC
ORC3L	76	CCCGGAGGG
ORC3L	77	CCCCGGAGG
ORC3L	274	GCCCGCAGC
ORC3L	296	CCCTCCAGC
PDZGEF2	79	CCCTCCAGC
PDZGEF2	163	CCCGCCAGT
PDZGEF2	499	CCCACGGGT
PDZGEF2	500	CCCCACGGG
PDZK1	122	GCCCCCAGC
PERP	14	CCCCGCGGC
PERP	43	GCCCGGAGC
PERP	100	CCCGGCAGG
PERP	175	CCCTCGGGC
PIR51	91	GCCATCGGC
PLEKHF2	60	GCCCGGAGT
PLEKHF2	161	CCCTAGAGC
PLEKHF2	175	GCCAAAGGC
PLEKHF2	207	GCCCGCGGC
PLK4	149	CCCCGAAGT
PLK4	270	GCCATAAGT
PMSCL1	60	GCCCAGGGC
PMSCL1	74	CCCCCAAGT
PMSCL1	194	CCCTCAGGC
PMSCL1	472	GCCTCGGGG
POLE2	-47	CCCAGGAGT
POLE2	201	CCCGGCAGC
POLE2	212	GCCCTCAGT
POLE2	348	CCCAACAGC
POLE2	455	CCCTCAAGT
PPIL5	91	GCCATCGGG
PPIL5	153	CCCAACGGC
PPIL5	211	CCCCAAGGC
PPIL5	212	GCCCCAAGG
PPIL5	425	GCCGGGAGC
PPIL5	454	CCCCAAGGC
PPIL5	455	CCCCCAAGG
PRC1	-9	GCCCAGGGT
PRC1	105	CCCGAGGGG
PRC1	106	GCCCGAGGG
PRC1	254	CCCTTGAGG
PRC1	390	CCCGGCAGC
PREX	70	CCCGCCGGC
PREX	100	GCCGGGAGC
PREX	121	GCCCCCAGT
PREX	141	GCCCCCGGC