

Characterisation of two methods involving  
double-stranded RNA for the functional genetic  
analysis of the murine embryonic stem cell

—

Antisense gene trapping & RNA interference

Timothy J. T. Chevassut

PhD Thesis



*To Ann,  
for her love and support,*

*&*

*to Emily, Ben and Hannah,  
who are my inspiration.*



---

## Abstract

Murine embryonic stem cells are derived from the inner cell mass of a pre-implantation mouse blastocyst. According to the conditions in which they are cultured, these cells exhibit dual defining characteristics of indefinite self-renewal and pluripotent differentiation. This thesis sought to explore two different methods of functional genetic analysis of embryonic stem cells, namely antisense gene trapping and RNA interference.

Antisense gene trapping in embryonic stem cells involves the sequential genomic integration of two vectors: the first to randomly trap active gene promoters and the second to induce functional antisense gene knockout. These putative “knockout” clones were screened for mutant phenotypes using a methylcellulose differentiation assay. However, for reasons of technical inefficiency and biological variability, such a screening strategy proved to be experimentally impractical. Instead, clones were selected according to the sequence identity of the trapped locus. Five such gene-trap clones were characterised (*Tsix*, *Catns*, *Rex2*, *Nasp*, and *Ctbp2*) and the analyses of these clones are presented.

RNA interference involves the targeted knockdown of a known gene via the action of a sequence-homologous double-stranded RNA molecule. In mammalian cells, gene knockdown can be achieved using short interfering RNA duplexes 19-21 base pairs in length. Using various mammalian cell lines, the nature of this interference process is characterised and shown to operate predominantly at the transcriptional DNA level, rather than the post-transcriptional mRNA level as exhibited by *Drosophila* cells. The use of RNA interference to study embryonic stem cell biology is explored including demonstration of knockdown of the maintenance methyltransferase gene, *Dnmt1*, that results in significant demethylation of genomic DNA.

Antisense gene trapping and RNA interference represent two complementary strategies for the functional genetic analysis of the mammalian cell. In both cases, the goal is to correlate genotype with phenotype, or vice versa, through the bi-allelic knockdown of a target gene. By contrast with one another, while gene trapping has the potential to define novel genes according to their function, RNA interference can only be used for known genes whose functions are predicted from their respective sequences. Intriguingly, both methodologies involve double-stranded RNA suggesting a commonality of underlying molecular mechanism: in the case of antisense gene trapping this is generated endogenously, while in the case of RNA interference it is introduced exogenously in the form of short duplexes. In this manner, double-stranded RNA represents a unifying thread to this thesis.

---

## Declaration

The work presented in this thesis is entirely that of the candidate and no one else. It contains no material previously submitted for any other degree or previously published except where stated. The thesis is entirely composed by the candidate.

Thesis presented for the degree of Doctor of Philosophy, University of Edinburgh.



20<sup>th</sup> October 2006

Dated

The John Hughes Bennett Laboratory  
Division of Molecular and Clinical Medicine  
University of Edinburgh

October 2006

---

## **Dedication**

I dedicate this thesis to my parents Anthony and Kathleen, for the love and support they have always shown me; to my wife Ann, for her encouragement and loving companionship; and to our three children, Emily, Benjamin, and Hannah, who are simply the most fantastic thing to have ever happened in our lives.

---

## Acknowledgements

I am indebted to countless individuals who have provided inspiration, advice and assistance during the period of the research and who have given me support during the write-up of this thesis:

- To Bing Lim, who took an enormous gamble in giving me the opportunity to study biology in his laboratory and who has been a guiding influence throughout; for all the crazy discussions we had over cups of Starbucks tea, and for his constant encouragement to always think differently.
- To Lesley Forrester, who similarly took a major gamble in granting me the chance to continue my research in Edinburgh; for her encouragement and thoughtful criticism of my work, and for her rightful judgement as my thesis supervisor that, sometimes, less is more.
- To Bernard Ramsahoye, a friend and mentor who is as good a scientist as I have known and who never declined the opportunity for discussing the mysteries of biology over a cup of coffee.
- To Professor John Ansell, who accepted me as a PhD candidate under his supervision, and to Marc Turner, who has provided helpful advice regarding juggling clinical and academic careers.
- To all the excellent colleagues of the John Hughes Bennett Laboratory, notably Melany Jackson, Anna Krassowska, Richard Axton, Ronnie Gallacher, Kay Samuel, Tim Fagge, and Olga Tura.
- To Iain Chambers, final-year undergraduate, and Anestis Tsakiridis, MSc Honours student, who kindly allowed me to act as supervisor of respective laboratory projects.
- To medical colleagues, who increasingly refrained from asking me if my PhD was written yet!
- To members of the Bing Lim laboratory, especially Adam Schmucker, Victor Li, and Xiaoyi Li, and members of the Divisions of Hematology / Oncology and Experimental Medicine at the HIM.
- To Stanley Cohen (Stanford) who kindly provided key genetic material; to En Li (Boston), and to Alexander Medvinsky and Austin Smith (Edinburgh), who all kindly provided valuable cell lines; and to Xinsheng Nan (Edinburgh) who kindly shared unpublished data used in this thesis.
- To Buzz Baum, for kind provision of experimental material and also for his advice, discussion, and friendship, and to Hannah Kuper, for her company I enjoyed on the M2 shuttle daily.

---

## **Background**

The work presented in this thesis was originally initiated at the Harvard Institutes of Medicine in Boston while employed as a research fellow in the Department of Hematology / Oncology at Beth Israel Deaconess Medical Center, Harvard Medical School. It was brought to fruition at the John Hughes Bennett Laboratory in the Division of Molecular and Clinical Medicine at the University of Edinburgh.

## **Funding**

I am indebted to the *Association for International Cancer Research*, without whose generous support this work would never have been possible, and also to the *Leukaemia Research Fund* for their support of the John Hughes Bennett Laboratory.

---

## Abbreviations

CCE	Embryonic stem cell line
cDNA	Complementary deoxyribonucleic acid
CMV	Cytomegalovirus
CHO	Chinese hamster ovary cell line
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
Dnmt	<i>De novo</i> methyltransferase
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
E14	Embryonic stem cell line
EB	Embryoid body
ES	Embryonic stem
GDID4	GDI Dissociation Inhibitor D4
GMEM	Glasgow's Modified Eagle Medium
GFP	Green fluorescent protein
HeLa	Human cervical carcinoma cell line
IMDM	Iscove's Modified Dulbecco's Medium
LB	Luria-Bertani
LIF	Leukaemia inhibitory factor
LTR	Long terminal repeat
MoMuLV	Moloney murine leukaemia virus
PCR	Polymerase chain reaction
PBS	Phosphate buffered saline
PEF	Primary embryonic fibroblast
PGK	Phosphoglycerate kinase
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
SAM	S-adenosyl-methionine
siRNA	Short interfering ribonucleic acid
siRNP	Small interfering ribonucleoprotein
SSEA	Stage-specific embryonic antigen
TdT	Terminal deoxynucleotidyl transferase
tTA	Tetracycline transactivator

---

**“Imagination is more important than knowledge”**

**Albert Einstein**

---

## Preface

Jacques Monod, in his essays on the natural philosophy of modern biology, claimed that evolution is without purpose and driven only by “chance and necessity”. The same might also be said of a PhD thesis. When the work for this thesis was commenced in 1999, the phenomenon of RNA interference had only recently been described and it was not until two years later that mammalian genes were first shown to be silenced using short interfering RNA duplexes. But the pace of scientific discovery is unrelenting and, since then, the use of double-stranded RNA to manipulate gene expression has almost become routine in the laboratory with genome-wide libraries of molecules now available for high-throughput cellular analysis using phenotypic screens. However, more often than not, RNA interference is being used merely as a convenient genetic tool. This thesis has attempted to answer some basic mechanistic questions about how double-stranded RNA silences genes in mammalian cells. Some of the findings presented here have been corroborated by published data elsewhere while others remain admittedly contentious. Either way, in reading this work, it should be borne in mind that the experimental results were generated over the period between 1999 and 2002 when RNA interference, as applied to mammalian systems, was still in its infancy.

The other major landmark in biology to occur during the period of completion of this thesis was the publication of the human (and mouse) genome, a resource that is now almost taken for granted but which represents the remarkable culmination of fifty years of modern molecular biology since the discovery of the double helix. As testament to this, the gene sequence of the first gene trap clone generated in the course of experimental work presented here was that of an X-linked gene of unknown function that was not pursued further. However, it was only in revisiting this sequence four years later, during the write-up of this thesis, that its identity was revealed as being *Tsix*, a key player in coordinating X chromosome inactivation.

Finally, there is no little irony in the fact that the two methods of gene silencing presented in this thesis, namely antisense gene trapping and RNA interference, share with one another a similar mechanism that involves the generation of double-stranded RNA. In antisense gene trapping, this is produced endogenously and envisioned to inhibit gene expression via an antisense mechanism, while in RNA interference, it is introduced exogenously in the form of short RNA duplexes and the mechanism is assumed to be one of homology-driven mRNA cleavage. However, this mechanistic overlap between the two methodologies was not obviously evident at the outset of the PhD and stands as testament to the endless capacity for biology to conjure up unforeseen surprises.

# Index

---

<b>CHAPTER 1 – INTRODUCTION</b>	<b>1</b>
<b>1.1 INTRODUCTION</b>	<b>2</b>
1.1.1 What is life?	2
1.1.2 The language of life	3
1.1.3 Syntax and semantics	3
1.1.4 What is differentiation?	4
1.1.5 Revealing gene function	5
<b>1.2 EMBRYONIC STEM CELLS</b>	<b>6</b>
1.2.1 Introduction	6
1.2.2 Maintenance of ES cell in an undifferentiated state	8
1.2.3 Differentiation of ES cell in culture	9
1.2.4 Haematopoietic development in the embryo	10
1.2.5 Haematopoietic differentiation of ES cells	11
<b>1.3 GENE TRAPPING</b>	<b>13</b>
1.3.1 Trapping genes in embryonic stem cells	13
1.3.2 Examples of genes identified by gene trapping	14
1.3.3 Strategies for screening of gene trap clones	15
1.3.4 High throughput screening approaches	15
1.3.5 Strategies for generating homozygous gene trap mutants	17

---

1.3.6	Functional homozygous knockouts by antisense gene trapping	18
1.3.7	Screening functional ES gene trap knockout clones for “differentiation genes”	19
1.4	<b>RNA INTERFERENCE</b>	20
1.4.1	Silencing genes by RNA interference	20
1.4.2	Short interfering RNAs (siRNAs)	21
1.4.3	Processing of double-stranded RNAs	22
1.4.4	Molecular mechanism of gene interference	23
1.4.5	Biochemistry of the RISC	24
1.4.6	Alternative theories of RNA interference	26
1.4.7	Transcriptional gene silencing	27
1.4.8	Role of the RNAi machinery in heterochromatin formation	28
1.4.9	RNAi-induced epigenetic changes in mammalian cells	29
1.4.10	Novel strategy variations designed to induce gene interference	29
1.4.11	Translation of RNAi technology <i>in vivo</i>	30
1.4.12	RNAi as applied to embryonic stem cells	31
1.4.13	Theories on the biological role of small RNA molecules	32
1.4.14	Future prospects for RNAi in biology and medicine	33
1.5	<b>AIMS OF THESIS</b>	34

---

<b>CHAPTER 2 – METHODS</b>	35
<b>2.1 CELL CULTURE</b>	36
2.1.1 Culture of embryonic stem cells	36
2.1.2 Production of primary embryonic fibroblasts	37
2.1.3 Methylcellulose differentiation of embryonic stem cells	37
2.1.4 Electroporation of embryonic stem cells	38
2.1.5 Antibiotic selection of embryonic stem cell clones	39
2.1.6 Assay for expression of $\beta$ -galactosidase	40
<b>2.2 MOLECULAR BIOLOGY</b>	41
2.2.1 Extraction of genomic DNA from cells	41
2.2.2 Extraction of total RNA from cells	41
2.2.3 PCR analysis for contained genomic sequence	42
2.2.4 RT-PCR analysis for transcribed sequence	43
2.2.5 Identification of gene trapped sequences by the method of 5' Rapid Amplification of cDNA Ends (5'RACE)	44
2.2.6 Sequence analysis of 5'RACE products	46
2.2.6 Identification of gene trapped loci by BLAST database analysis	46
2.2.7 Measurement of cytosine methylation using nearest neighbour analysis	46
2.2.8 Transformation of competent <i>E. coli</i>	47
2.2.8 Large-scale harvest of plasmid DNA	48

---

<b>2.3</b>	<b>RNAi METHODOLOGY</b>	49
2.3.1	Design and synthesis of short interfering RNA (siRNA) oligonucleotides	49
2.3.2	Synthesis of siRNAs using an <i>in vitro</i> transcription method	50
2.3.3	Method of Transfection of siRNAs into cells in culture	51
2.3.4	Sequences of siRNA molecules used in current experiments	52
2.3.5	Co-transfection of siRNAs with DNA or RNA gene reporters	53
2.3.6	Quantification of <i>Firefly / Renilla</i> luciferase expression by luminescence measurement	54
2.3.7	Flow cytometry analysis of siRNA-induced knockdown of endogenously expressed green fluorescent protein in embryonic stem cells	55
2.3.8	Analysis of telomerase activity using PCR/ELISA-based assay	55
2.3.9	Optimisation of siRNA cell transfection by various methods including use of HIV-derived <i>tat</i> transduction peptide molecules	56
<b>2.4</b>	<b>QUANTITATIVE METHODOLOGY</b>	
2.4.1	Statistical considerations	57
2.4.2	Statistical analysis: Results 2	58
2.4.3	Statistical analysis: Results 3	62
2.4.4	Statistical analysis: Results 4	66
2.4.5	Units of measurement	68

---

<b>CHAPTER 3 – RESULTS 1</b>	<b>69</b>
<b>ANTISENSE GENE TRAPPING IN EMBRYONIC STEM CELLS</b>	
3.1 Introduction	70
3.2 Overview of the vector constructs	72
3.3 Overview of the experimental strategy	74
3.4 Overview of the experimental design	75
3.5 Gene search vector (pGSV) traps active genes in ES cells	76
3.6 5' rapid amplification of cDNA ends (5'RACE) enables gene trap sequences to be identified	78
3.7 Gene trap clones C2, E5 and E6 exhibit uniform X-Gal staining suggestive of constitutive promoter activity in undifferentiated ES cells	82
3.8 Clones E3 and E12 exhibit discrete X-gal staining patterns that may reflect sub-cellular localisation of the respective gene trap fusion proteins	83
3.9 The gene trap genes, Rex2, Nasp, and Ctbp2, are <i>not</i> expressed in erythroid cells, as assessed by LacZ staining, during haematopoietic differentiation	84
3.10 Enforced" expression of the gene trapped gene via maintained G418 selection pressure during ES cell differentiation induced by LIF-withdrawal	85
3.11 B-galactosidase negative sub-clones are generated following introduction of the transactivator vector (pTX) into gene trap ES cell clones	87
3.12 Comparative differentiation of gene trap clones with antisense knockout sub-clones fails to identify any differences using the methylcellulose assay	90
3.13 Generation of Cre recombinase sub-clones leads to the reappearance of $\beta$ -galactosidase expression putatively via excision of the transactivator vector	91
3.14 Technical problems relating to the gene trap antisense knockout strategy for identification of differentiation genes in ES cells	92
3.15 Conclusion	93

---

<b>CHAPTER 4 – RESULTS 2</b>	103
<b>ESTABLISHING THE BIOLOGY OF RNA INTERFERENCE</b>	
4.1 Introduction	104
4.2 Sequence-specific short interfering RNA molecules suppress <i>Firefly</i> luciferase expression in a potent and dose-dependent manner	105
4.3 Interference of <i>Firefly</i> luciferase expression is dependent upon the presence of <i>both</i> sense and antisense RNA oligonucleotide strands	106
4.4 The interference effect is maximal when sense and antisense RNA strands are in equimolar proportions suggesting a stoichiometry of 1 to 1	107
4.5 The specificity of the interference effect is critically dependent upon sequence fidelity of the siRNA molecule with its target region	108
4.6 siRNA molecules that have sequence specificity for different regions within the target gene exhibit differential interference potency	109
4.7 Combinations of interfering siRNA molecules yield gene knockdown effects that approximate to the summation average of each individual siRNA	111
4.8 The combination of the sense strand of one siRNA with the non-homologous antisense strand of another siRNA fails to generate an interference effect	112
4.9 siRNA-mediated interference of co-transfected <i>Firefly</i> luciferase is apparent as soon as luminescence is detectable	113
4.10 siRNA-mediated knockdown of a co-transfected plasmid-borne gene exhibits first order competitive reaction kinetics	114
4.11 The magnitude of siRNA-mediated gene knockdown is diminished if siRNA transfection is delayed relative to transfection of the luciferase plasmid	115
4.12 Premature transfection of the siRNA up to two days prior to transfection of the luciferase plasmid results in robust gene knockdown	116
4.13 siRNAs can mediate suppression of a stably integrated <i>Firefly</i> luciferase gene	117

---

4.14	The interference-related gene knockdown effect persists for at least ten days post-siRNA transfection	118
4.15	RNA oligonucleotides in high dose provoke a non-specific transient enhancement of luciferase expression in CHO-LUC cells	119
4.16	siRNA molecules generated by <i>in vitro</i> transcription exhibit less gene knockdown potency than siRNA molecules generated by chemical synthesis	121
4.17	Using short basic peptides derived from the transduction domain of the HIV TAT protein to enhance liposomal delivery of siRNA molecules	122
4.18	Tat-derived transduction peptides appear to enhance liposomal-mediated delivery of siRNA molecules into cells	123
4.19	Conclusion	124
<b>CHAPTER 5 – RESULTS 3</b>		130
<b>DEFINING MOLECULAR ASPECTS OF RNA INTERFERENCE</b>		
5.1	Introduction	131
5.2	The antisense strand is pre-eminent over the sense strand in directing siRNA-mediated gene knockdown of transiently expressed <i>Firefly</i> luciferase	132
5.3	The importance of antisense-strand fidelity for proper siRNA function is recapitulated for knockdown of a stably integrated <i>Firefly</i> luciferase gene	133
5.4	Phosphorylation of the 5' end of the siRNA molecule does not influence knockdown of a stably integrated luciferase gene	134
5.5	<i>Firefly</i> -specific siRNAs fail to knock down luciferase expression following transient transfection of a <i>Firefly</i> luciferase mRNA reporter molecule	135
5.6	Evidence suggesting partial siRNA-induced knockdown of mRNA-encoded <i>Firefly</i> luciferase after a period of 24 hours post-transfection	137
5.7	Failure of <i>Firefly</i> -specific siRNAs to mediate knockdown of mRNA-encoded luciferase reporter confirmed likewise in CHO cells	138

---

5.8	<i>Firefly</i> -specific siRNAs knock down expression of a co-transfected mRNA luciferase reporter in <i>Drosophila</i> KC cells but not S2 cells	139
5.9	5'-phosphorylation of the siRNA molecule does not affect level of knockdown of <i>Firefly</i> luciferase expressed from a transiently transfected pGL3 plasmid	141
5.10	Overview of experimental strategy designed to assess the effect of 5'-terminal phosphorylation on siRNA-induced knockdown of transiently transfected mGL2 and pGL2 <i>Firefly</i> luciferase reporters in mammalian and <i>Drosophila</i> cells	143
5.11	Evidence suggesting that 5'-phosphorylation of the siRNA termini enhances the knockdown of mGL2, but not pGL2, luciferase in <i>Drosophila</i> S2 cells	144
5.12	5'-phosphorylation of a sense-mismatched heteroduplex siRNA is necessary for knockdown of pGL2 <i>Firefly</i> luciferase in <i>Drosophila</i> , but not mammalian, cells	147
5.13	Conclusion	150
<b>CHAPTER 6 – RESULTS 4</b>		157
<b>TARGETING ENDOGENOUS GENES BY RNA INTERFERENCE</b>		
6.1	General design of siRNA-mediated knockdown experiments against various gene targets in different cell lines	158
6.2	Transfection of <i>Firefly</i> -specific siRNAs result in dose-dependent knockdown of transiently expressed luciferase in murine ES cells	159
6.3	Sequence-specific siRNA molecules are able to suppress expression of a GFP transgene stably integrated by random genomic insertion in ES cells	160
6.4	Repeated knockdowns of GFP expression, via serial transfection of siRNA, result in diminished rather than enhanced knockdown effect	161
6.5	LacZ-specific siRNAs can suppress expression of a $\beta$ -galactosidase transgene targeted to the Oct3/4 locus in murine ES cells	162
6.6	siRNA-induced knockdown of $\beta$ -galactosidase expression in ES cells exhibits dose-dependency and clonal heritability	163

---

6.7	siRNAs directed against the telomerase reverse transcriptase gene (hTERT) fail to show significant functional knockdown of telomerase activity in human cancer cell lines	164
6.8	siRNA-induced knockdown in ES cells of the purine salvage pathway enzyme, Hprt, fails to generate surplus mutant clones by negative selection in 6-thioguanine	166
6.9	CHO cells can be efficiently killed by targeting the pyrimidine salvage pathway using siRNA-directed knockdown of thymidine kinase (Tk)	167
6.10	siRNA-induced knockdown in ES cells is not contingent upon the presence of either Dnmt1 or Dnmt3a/3b DNA methyltransferase activity	169
6.11	siRNAs directed against Dnmt1 result in a functional deficit of maintenance methyltransferase activity in ES cells	171
6.12	Conclusion	174
<b>CHAPTER 7 – DISCUSSION</b>		179
7.1	Introduction	180
7.2	Antisense gene trapping	180
7.3	Analysis of gene trap genes	184
7.4	RNA interference	188
7.5	Gene knockdown using siRNAs	189
7.6	Molecular biology of siRNAs	192
7.7	A model of siRNA action	195
7.8	Probing ES cells with siRNAs	198
7.9	Where is the siRNA site-of-action?	200
7.10	Conclusion	202

# Figures

<b>CHAPTER 3 – RESULTS 1</b>	<b>69</b>
Figure 3.1: Schematic illustration of classical gene trapping strategy using basic vector design.	71
Figure 3.2: Vectors used to derive functional knockouts via gene trap / antisense strategy.	73
Figure 3.3: Experimental strategy for deriving functional gene trap/antisense knockouts.	74
Figure 3.4: Diagrammatic overview of experimental design requiring three sequential steps.	75
Figure 3.5a: Schematic diagram illustrating method of rapid amplification of cDNA ends.	79
Figure 3.5b: Sequence, notation and location of the two sets of primer triplets used for 5'RACE.	79
Figure 3.6: Alignment of DNA sequence derived from 5'RACE analysis of clone C2 genomic DNA.	80
Figure 3.7: DNA sequencing of the 5'RACE products revealing gene trap genomic integration sites.	81
Figure 3.8: Photomicrographs revealing uniform X-gal staining of gene trap clones C2, E5, E6.	82
Figure 3.9: Photomicrographs revealing discrete X-gal staining patterns for clones E3 and E12.	83
Figure 3.10: X-gal staining of haematopoietic embryoid bodies derived from ES gene trap clones.	84
Figure 3.11: Gene trap clone E5 cells grown in the presence of G418 fail to differentiate.	86
Figure 3.12: Pie charts showing the loss of $\beta$ -galactosidase expression in "knockout" sub-clones.	88
Figure 3.13: Photomicrographs demonstrating $\beta$ -galactosidase negative "knockout" clones.	89
Figure 3.14: Cre excision of pTX transactivator vector leads to $\beta$ -galactosidase reappearance.	91
 <b>CHAPTER 4 – RESULTS 2</b>	 <b>103</b>
Figure 4.1: Sequence-specific siRNAs result in potent and specific gene interference.	105
Figure 4.2: Both sense and antisense RNA strands are required for gene interference.	106
Figure 4.3: Gene interference is maximal when sense and antisense strands are equimolar.	107
Figure 4.4: The 1:1 stoichiometry of RNA interference is apparent at all concentrations.	107
Figure 4.5: Gene knockdown via RNAi is critically dependent upon sequence fidelity.	108
Figure 4.6: Table and schematic diagram showing the site and sequence of various siRNAs.	109
Figure 4.7: Graph showing the differential effects of three siRNAs against pGL2 and pGL3.	110
Figure 4.8: Graph showing same experiment performed over a range of siRNA concentrations.	110
Figure 4.9: Graph showing effect of using combinations of siRNAs on knockdown of <i>Firefly</i> expression.	111

Figure 4.10: Graph showing effect of non-matched sense/antisense pairs on gene expression.	112
Figure 4.11: Graph showing gene knockdown effects of siRNA-B3 at various time-points.	113
Figure 4.12: Graph revealing the kinetics of siRNA-mediated gene knockdown (linear-log plot).	114
Figure 4.13: Graph revealing the kinetics of siRNA-mediated gene knockdown (log-log plot).	114
Figure 4.14: Graph showing how delayed siRNA transfection affects level of gene knockdown.	115
Figure 4.15: Graph showing luciferase gene knockdown up to 48 hours post siRNA transfection.	116
Figure 4.16: Graph showing knockdown of an integrated luciferase in CHO-LUC cells.	117
Figure 4.17: Graph showing the persistence of siRNA-induced luciferase knockdown.	118
Figure 4.18: Graph showing non-specific RNA-induced enhancement of luciferase activity.	120
Figure 4.19: Non-specific enhancement of luciferase activity occurs at 24 but not 48 hours.	120
Figure 4.20: <i>In vitro</i> transcribed siRNAs are less potent than chemically synthesised siRNAs.	121
Figure 4.21: <i>In vitro</i> transcribed siRNAs yield luciferase enhancement in CHO-LUC cells.	121
Figure 4.22: Tat peptides enhance liposomal delivery of siRNA as assessed by gene knockdown.	123
<b>CHAPTER 5 – RESULTS 3</b>	130
Figure 5.1: Graph showing <i>Firefly</i> luciferase knockdown using the B2/siRNA-B3 heteroduplex.	132
Figure 5.2: Graph showing luciferase knockdown using the siRNA-B2/B3 heteroduplex.	133
Figure 5.3: Graph showing effect of 5'-phosphorylation on siRNA-induced knockdown.	134
Figure 5.4: <i>Firefly</i> -specific siRNA-B2 knocks down pGL2 but not mGL2 luciferase.	136
Figure 5.5: <i>Firefly</i> siRNAs fail to knock down mGL2 luciferase reporter within 18 hours.	136
Figure 5.6: <i>Firefly</i> -specific siRNAs generate a modest and late knockdown of mGL2 luciferase.	137
Figure 5.7: <i>Firefly</i> -specific siRNAs fail to knockdown mGL2 luciferase in CHO cells at 24 hours.	138
Figure 5.8: <i>Firefly</i> siRNAs knock down mGL2 luciferase in KC but not S2 <i>Drosophila</i> cells.	140
Figure 5.9: 5'-terminal phosphorylation is irrelevant to siRNA function in CHO cells.	142
Figure 5.10: 5'-terminal phosphorylation is irrelevant to siRNA function in KC cells.	142
Figure 5.11: Schematic representation of experimental strategy to examine siRNA knockdown.	143
Figure 5.12: Effect of 5'-phosphorylation on luciferase knockdown using <i>Firefly</i> siRNA-B2.	145
Figure 5.13: Effect of 5'-phosphorylation on luciferase knockdown using <i>Firefly</i> siRNA-C.	145

Figure 5.14: Effect of 5'-phosphorylation on siRNA-induced <i>Firefly</i> knockdown in various cells.	146
Figure 5.15: Effect of 5'-phosphorylation on sense-mismatch heteroduplex siRNA function.	148
Figure 5.16: Effect of 5'-phosphorylation on antisense-mismatch heteroduplex siRNA function.	148
Figure 5.17: Effect of 5'-phosphorylation of heteroduplex siRNAs on <i>Firefly</i> luciferase knockdown.	149
<b>CHAPTER 6 – RESULTS 4</b>	157
Figure 6.1: Schematic illustration demonstrating the basic experimental strategy for studying siRNA-induced knockdown of various target genes in a number of different cell lines.	158
Figure 6.2: Graph showing siRNA-mediated knockdown of <i>Firefly</i> luciferase in murine ES cells.	159
Figure 6.3: Flow cytometric analysis showing siRNA-mediated knockdown of GFP in ES cells.	160
Figure 6.4: Repeated cycles of siRNA transfection result in diminished levels of GFP knockdown.	161
Figure 6.5: LacZ-specific siRNAs knock down expression of a stable $\beta$ -galactosidase targeted to the Oct3/4 locus in ES cells.	162
Figure 6.6: Graph showing siRNA-induced knockdown of $\beta$ -galactosidase expression in ES cells.	163
Figure 6.7: Graph showing effect of transfection of hTERT-specific siRNAs in MDA cells.	165
Figure 6.8: Hprt-specific siRNAs fail to generate surplus 6-thioguanine resistant ES clones compared with control siRNA.	166
Figure 6.9: Transfection of Tk-specific siRNA (sense + antisense) is toxic to CHO-LUC cells.	168
Figure 6.10: Transfection of Hprt-siRNA is also toxic in CHO cells but tenfold less than Tk-siRNA.	168
Figure 6.11: siRNA-induced knockdown is not contingent DNA methyltransferase activity.	170
Figure 6.12: siRNA-induced knockdown of Dnmt1 results in loss of genomic methylation in ES cells.	172
Figure 6.13: siRNA Dnmt1 knockdown leads to loss of cytosine methylation in wild-type ES cells.	173
<b>CHAPTER 7 – DISCUSSION</b>	179
Figure 7.1: Schematic diagram of the mechanism of action of siRNA in gene silencing in mammalian cells.	197

# **Chapter 1**

## **Introduction**

## 1.1 INTRODUCTION

### 1.1.1 What is life?

Molecular biology was born sixty years ago with the publication of Erwin Schrödinger's classic book, *"What Is Life?"*. In it, the Austrian physicist, and father of quantum wave mechanics, reflected upon the essence of living organisms, which hitherto had remained almost a complete mystery to science (Ref: Schrödinger E., *What is Life? The Physical Aspect of the Living Cell*, published 1944). Schrödinger argued that life might be thought of as the expression of biological information contained within a "hereditary code-script" whose structure would be that of an "aperiodic crystal". Through its internal arrangement of atoms, such a material would contain all the instructions necessary for making a living organism. Less than ten years later, the nature of this aperiodic crystal was elucidated by Watson and Crick, both strongly influenced by Schrödinger's writings, in their unravelling of the structure of deoxyribonucleic acid, or DNA (Ref: Watson J.D. and Crick F.H., 1953a). Their remarkable insight, based upon the crystal diffraction data of Wilkins and Franklin, revealed that this "life substance" comprised a non-repeating sequence of sugar bases contained within a linear molecular chain (Refs: Franklin R.E. and Gosling R.G., 1953; and Wilkins M.H. *et al.*, 1953). Moreover, the very structure of the DNA molecule embodied a clear mechanism for the copying of this coded sequence, something that had long been recognised as being a necessary property of life's genetic material (Ref: Watson J.D. and Crick F.H., 1953b).

Fifty years on from this seminal discovery, an international collaborative effort has succeeded in unravelling the entire DNA sequence of over three billion letters that comprises the human genome, the so-called "book of life" (Refs: Lander E.S. *et al.*, 2001; McPherson J.D. *et al.*, 2001; Venter J.C. *et al.*, 2001; International Human Genome Sequencing Consortium, 2004). This landmark achievement denotes the start of a new era in the exploration of human biology.

### 1.1.2 The language of life

DNA is comprised of four chemical bases, adenine (A), guanine (G), thymine (T), and cytosine (C), arranged in a non-repeating sequence. The discrete nature of this sequence of letters is akin to the digital nature of computing language and represents the fundamental hallmark feature of any information-containing system. The encoding of biological information in such a manner underpins the two defining antagonistic properties of Darwinian evolution, namely the heredity of phenotypic features via a high-fidelity copying process and the emergence of new phenotypic features via perturbation of the sequence through copying errors. This emergence of new features in an organism, and their subsequent inheritance in the progeny, gave rise to the notion of the gene as the basic unit of heredity. In eukaryotic organisms, two copies of each gene are inherited and segregated singly into the germ cells, then paired up at random via fertilisation allowing shuffling of hereditary units at each generation. The deciphering of the “genetic code” in the early 1960s heralded the defining of the gene as a molecular entity and gave birth to the notion of genotype as the molecular correlate of phenotype. The understanding of the genotypic basis of cell and organism phenotype remains the overarching goal of molecular biology today.

### 1.1.3 Syntax and semantics

The “Central Dogma” of biology states that the flow of information in a cell is unidirectional from the nucleic acid to the protein: the gene, defined by its sequence of nucleotides in the DNA, is transcribed into a mobile molecule, the messenger RNA, that becomes translated into the protein which is responsible for enacting the gene’s proscribed function. Using a figurative analogy, a computer program comprises a string of digital information that is transcribed into letters of the alphabet and translated into words on this page. To extend this analogy further, the code of a computer program must follow certain syntactical rules in order to generate the correct semantic output (Ref: Hofstadter D.R. *Godel, Escher, Bach*, published 1980). Such is the case also with the

genotype and the phenotype. The exhaustive identification of genes in the human genome represents a remarkable achievement, but the identification of the rules of grammar, the syntax of the genome, is far from complete. In a multicellular organism, this encompasses not only *which* gene is transcribed and translated but *when* and *where* this occurs. The “genetic syntax” responsible for coordinating the temporal and spatial expression of genes is necessarily embedded in the primary code of the DNA contained in the fertilised egg. The “semantic consequence” of this process is the coordinated interaction of myriad different proteins whose interacting functions determine cell behaviour and lead to development of an organism.

### 1.1.4 What is differentiation?

The totipotency of a fertilised egg refers to the remarkable feat whereby, given the right environment, it will develop into a multicellular organism composed of differentiated tissues. The demonstration, originally in amphibians (Ref: Gurdon J.C. *et al.*, 1958) and more recently in mammals (Ref: Campbell K.H. *et al.*, 1996), that an apparently normal animal can be produced following enucleation of the egg and the transfer of a donor nucleus derived from a somatic cell, reveals that the genome contained in committed cells is the same as that of the original egg. The further implication is that it is the intracellular environment of the egg that is able to “re-boot” the somatic cell derived genome and re-initiate the developmental program. This developmental program is characterised by a cascade of coordinated gene expression regulated spatially and temporally by feedback loops impinging on the genome. The dissection of this interplay between the genome and the microenvironment of the nucleus, which in turn is influenced by cytoplasmic and extracellular factors, represents a seemingly insurmountable challenge. However, while the complete description of differentiation in all its detail might not be fully realisable, its comprehension in broad principle should be. One experimental tool that has yielded countless insights into this differentiation process is the murine embryonic stem cell that provides an *in vitro* model of *in vivo* development.

### 1.1.5 Revealing gene function

While the cell is clearly more than the sum of all its genes, nonetheless the reductionist approach to elucidating gene function has served cell biology well over the past fifty years. This permits function to be ascribed to the gene in terms of the cellular phenotype yielded when the gene is deleted from the cell's repertoire. The bringing together of genotype and phenotype in this manner represents the overarching goal of molecular biology and has traditionally been achieved in one of two ways. The classical approach, or forward genetics, has been to isolate mutants defined in terms of a suitable biological screen and identify the causative gene lesion. The alternative approach, or reverse genetics, involves the isolation of the gene on the basis of its sequence or structure and its functional deletion, or mutagenisation, by appropriate means. These two approaches are exemplified in this thesis by two contrasting strategies: the forward genetic approach employs a novel variation of gene trapping to identify genes important to differentiation in murine ES cells; while the reverse genetic approach takes advantage of a biological process termed RNA interference to knockdown expression of pre-determined genes in murine ES cells as well as other cell lines. The nature of gene trapping and the biology of RNA interference are described in greater detail below, while the rationale for the use of these two approaches in the context of the current thesis is outlined thereafter.

## 1.2 EMBRYONIC STEM CELLS

### 1.2.1 Introduction

Embryonic stem (ES) cells are pluripotent cells derived from pre-implantation mouse embryos via dissection and culture of the inner cell mass of a developing blastocyst (Ref: Evans M.J. and Kaufman M.H., 1981). While maintained on embryonic fibroblasts and grown in the presence of Leukaemia Inhibitory Factor (LIF; also known as Differentiation Inhibiting Activity, or DIF), murine ES cells exhibit indefinite self-renewal (Refs: Williams R.L. *et al.* 1988; Smith A.G. *et al.*, 1988). However, if introduced into a host blastocyst, ES cells will contribute to all tissue lineages resulting in production of a chimaeric mouse (Ref: Bradley A. *et al.*, 1984), or using tetraploid blastocysts wholly ES-derived mice (Ref: Wang Z.Q. *et al.*, 1997). This includes contribution to the germ cells allowing passage of the ES genome into the progeny and testifying to totipotency of the ES cell. This is the basis whereby mouse models, carrying discrete genetic mutations, can be generated using appropriately manipulated ES cells. This includes both knockout animals by gene deletion through homologous recombination and over-expressing animals by introduction of a transgene under a powerful promoter.

While ES are best maintained in an undifferentiated state by culture on embryonic fibroblasts, it was demonstrated early on that recombinant LIF on its own is sufficient for culture of undifferentiated ES cells even in the absence of so-called feeders (Ref: Pease S. *et al.*, 1990) and moreover for the isolation of new ES cell lines (Ref: Nichols J. *et al.*, 1990). ES cells that are removed from the presence of LIF, or intimacy with embryonic fibroblasts, cease exhibiting self-renewal activity and instead undergo spontaneous differentiation into various lineages according to the culture conditions. Differentiation of ES cells *in vitro* has provided a powerful model system for addressing questions regarding lineage commitment that are less readily addressed *in vivo*. The most widely utilised approach has been to seed ES cells in methylcellulose and generate

clonal proliferations called embryoid bodies (EBs) that show evidence of mimicking early embryogenesis. This is reflected by the expression of genes indicative of primitive endoderm and mesoderm (Ref: Keller *et al.*, 1993). Moreover, after extended periods of culture, EBs can generate cells of haematopoietic (Ref: Wiles M.V. *et al.*, 1991), endothelial (Ref: Risau W. *et al.*, 1988), muscle (Ref: Robbins J. *et al.*, 1990) and neuronal (Ref: Bain G. *et al.*, 1995) lineages. Thus, ES cell differentiation *in vitro* appears to closely recapitulate that of embryo development *in vivo* highlighting the value of ES cells as a model for the study of lineage commitment (Ref: Keller G.M., 1995).

Thus, ES cells exhibit dual defining criteria: one is the ability to undergo indefinite self-renewal in an undifferentiated state while the other is the ability to undergo pluripotent differentiation when so induced. The POU transcription factor Oct4 appears to play an essential role in both the maintenance of this undifferentiated state and the initiation of normal differentiation as revealed by the observation that cells derived from the inner cell mass of Oct4-deficient embryos are not pluripotent (Ref: Nichols J. *et al.*, 1998). Indeed, it is the precise level of expression of the Oct4 locus that determines ES cell fate as revealed by the fact that both up and down-regulation of the gene, from normal self-renewal levels, induce divergent developmental programs (Ref: Niwa H. *et al.*, 2000). The role of LIF in this molecular switching process appears to be mediated through a receptor complex comprising the LIF receptor subunit (LIF-R) and the signal transducer gp130. However, the findings that embryos lacking LIF-R or gp130 can develop beyond gastrulation provide experimental evidence for existence of one or more alternative pathways governing the maintenance of pluripotent state (Ref: Dani C. *et al.*, 1998). This was verified by the identification of the homeodomain protein, Nanog, which directs propagation of undifferentiated ES cells and is not expressed in differentiated cells (Ref: Chambers I. *et al.*, 2003). Endogenous Nanog acts in parallel with LIF-induced activation of Stat3 to drive ES cell self-renewal via maintenance of Oct4 levels and establishes a central role for Nanog in the transcription factor hierarchy that defines ES cell identity (Ref: Mitsui K. *et al.*, 2003).

Recent interest in ES cells has been further fuelled by the isolation of cells from human blastocysts that exhibit features akin to those derived from mice (Ref: Thomson J.A. *et al.*, 1998). These human ES cells proliferate in an undifferentiated state while maintaining the capacity for pluripotent differentiation under appropriate conditions. More recently, the feasibility of isolating human ES cells from cloned blastocysts derived by somatic cell nuclear transfer has also been realised (Ref: Hwang W.S. *et al.*, 2004). However, despite these advances, knowledge of the broad principles regulating the pluripotent state, let alone precise understanding of the molecular cascade that regulates differentiation, remain surprisingly limited.

### **1.2.2 Maintenance of ES cells in an undifferentiated state**

As outlined above, Es cells can be maintained in an undifferentiated state by co-culture with mouse embryonic feeder cells (Ref: Evans M.J. and Kaufman M.H., 1981). LIF was subsequently identified as one of the feeder-cell-derived molecules that plays a pivotal role in this process (Ref: Smith *et al.*, 1992) and, in the presence of suitable batches of foetal calf serum, this is sufficient to replace the feeder cells altogether and maintain an undifferentiated ES cell state. More recently, BMP4 has also been shown to play a key role in ES cell growth (Ref: Ying Q.L. *et al.*, 2003) and in the presence of LIF can supplant the need for serum. Thus, it is now possible to grow ES cells with defined factors in the absence of both feeder cells and serum.

Molecular analyses have revealed that LIF functions through the gp130 activation of STAT3 (Ref: Niwa *et al.*, 1998), whereas BMP4 signalling appears to be mediated via Smad activation leading to induction of helix-loop-helix Id factors. In addition to STAT3 and Id, two other transcription factors, Oct3/4 (Ref: Niwa H. *et al.*, 2000) and nanog (Ref: Chambers I. *et al.*, 2003; Matsui K. *et al.*, 2003), have been shown to play pivotal roles in maintaining ES cells in an undifferentiated state.

### 1.2.3 Differentiation of ES cells in culture

ES cells will exhibit spontaneous differentiation when removed from the factors that would otherwise maintain them in an undifferentiated state. This differentiation will yield progeny identifiable as belonging to one of the three embryonic germ layers, mesoderm, ectoderm, and endoderm (Ref: Keller G., 1995). Interestingly, wild-type ES cells are unable to differentiate into trophoectoderm in culture reflecting the potential of their founder embryonic population, the inner cell mass.

Broadly speaking, there are two general approaches for studying ES cell differentiation. The first is to allow the ES cells to form embryoid bodies. These can be generated either by aggregation in medium, for instance suspending in hanging drops, (Ref: Doetschman T.C. *et al.*, 1985) or by clonal expansion, such as by seeding in semi-solid methylcellulose (Ref: Keller G., 1995). The alternative strategy is to culture the differentiating ES cells as a monolayer. This might be achieved either using stromal cells such as the OP9 cell line (Ref: Nakano T. *et al.*, 1994) or on extracellular matrix proteins (Ref: Nishikawa S.I. *et al.*, 1998). Both embryoid body formation and monolayer differentiation have their respective advantages and disadvantages. Embryoid bodies retain some semblance of three-dimensional structure that retains important cell-cell interactions important for certain developmental programs. Conversely, the complexity of embryoid body differentiation can complicate the interpretation of experiments where factors determining lineage commitment are being sought. The latter is sometimes better dissected out using two-dimensional monolayer growth despite the fact this loses some of the physiology of normal embryonic development. Indeed, it is often a combination of both methodologies that has been employed for analysing ES cell differentiation and in this regard the haematopoietic lineage represents probably the best characterised. Here, the processes determining ES cell lineage commitment appear to closely mimic normal haematopoietic development in the early embryo making *in vitro* ES cell differentiation an especially valuable research tool.

### 1.2.4 Haematopoietic development in the embryo

Haematopoiesis in the early embryo arises initially at two distinct sites. These are the extra-embryonic yolk sac and shortly thereafter an intra-embryonic region termed the aorto-gonadomesonephros, or AGM (Ref: Medvinsky A. and Dzierzak E., 1996). Haematopoietic development in these two sites appears to follow different genetic programs referred to respectively as primitive and definitive haematopoiesis. Primitive haematopoiesis, occurring in the yolk sac within distinct blood islands, gives rise to primitive erythrocytes that are large and nucleated and contain an embryonic form of haemoglobin. By contrast, definitive haematopoiesis, occurring within the AGM, gives rise to erythroid cells that undergo enucleation prior to entering the bloodstream, are smaller than those of primitive lineage and produce adult forms of haemoglobin.

The other major defining criterion that distinguishes primitive from definitive haematopoiesis relates to their potential for giving rise to so-called haematopoietic stem cells. Primitive haematopoiesis, occurring in the yolk sac, appears unable to give rise to haematopoietic stem cells while definitive haematopoiesis, occurring initially in the AGM, is able to give rise to such cells, which are defined by their ability for multi-lineage repopulation of the bone marrow. Of some relevance in this regard is the fact that the generation of haematopoietic stem cells from ES cells remains a major challenge to investigators in the field and, to date, efforts to identify such cells using ES cell culture have largely met with failure (Ref: Keller G., 2005). Only by forced over-expression of the HoxB4 gene in ES cells has it been possible to derive haematopoietic stem cells that are able to repopulate primary recipients with significant levels of engraftment (Ref: Kyba M. *et al.*, 2002). [A more recent paper showing bone marrow reconstitution of irradiated recipients using CD45+c-kit+ cells isolated from day 7-10 EBs is yet to be formally reproduced (Ref: Burt R.K. *et al.*, 2004)]. With the isolation of human ES cells, the derivation of potentially limitless haematopoietic stem cells from ES cell culture, for the purpose of therapeutic marrow rescue, represents a highly desirable goal.

### 1.2.5 Haematopoietic differentiation of ES cells

For the ES cell system to be a valid model of early haematopoietic development, and ultimately a source of haematopoietic stem cells, it has been necessary to demonstrate that *in vitro* ES cell differentiation closely recapitulates the genetic and transcriptional program of the embryo. Several different studies have demonstrated striking parallels between the ES cell model and the early embryo, providing important insights into the embryonic origins of the haematopoietic system (Ref: Keller G., 2005). Lineage commitment in ES cell culture can be followed in a number of ways: by monitoring gene expression patterns (Ref: Schmitt R. *et al.*, 1991); by noting the appearance of specific cell surface markers (Ref: Kabrun N. *et al.*, 1997); and by the development of clonable progenitor cells (Ref: Keller G. *et al.*, 1993). In optimised culture conditions following serum induction, ES cells will undergo haematopoietic differentiation with over 50% of cells expressing the haematopoietic/vascular receptor tyrosine kinase Flk-1 (Ref: Kabrun N. *et al.*, 1997) and up to 5% can represent a clonable haematopoietic progenitor (Ref: Keller G. *et al.*, 1993). Detailed analyses of the early stages of haematopoietic commitment have shown that both gene expression patterns and the kinetics of lineage development within EBs accurately reflect that found in the yolk sac (Ref: Robertson S.M. *et al.*, 2000). This is characterised by the transient appearance of the primitive erythroid lineage arising within EBs at around 4 days followed by the macrophage, definitive erythroid, and mast cell lineages occurring in the same temporal order as found in the yolk sac. Lymphoid progenitors and haematopoietic stem cells have not been identified among the progeny of the early stage EBs, suggesting that the initial stages of EB haematopoiesis represent the equivalent of yolk sac haematopoiesis.

The faithful recapitulation of haematopoietic development in the yolk sac provides good evidence that the ES/EB model is similar, if not identical, to that of the early embryo. Further support for this interpretation is provided by gene targeting studies that have helped to define the role of specific transcription factors. These include *Scl/tal-1* (Ref: Begley C.G. *et al.*, 1989), *Runx1* (Ref:

Wang S.W. and Speck N.A., 1992), and GATA-1 (Ref: Orkin S.H., 1992) in the establishment of the haematopoietic system. Each of these three factors have been shown to play key roles at certain stages of blood cell differentiation such that Scl/tal-1 is required for the development of all haematopoietic lineages, Runx1 for the definitive lineages but not primitive erythropoiesis, and GATA-1 for late-stage primitive and definitive erythroid maturation. All of these three defects have been accurately replicated *in vitro* using the ES cell differentiation model, once again validating the notion that embryoid bodies faithfully mimic embryonic haematopoiesis and in so doing provide an exceedingly versatile tool for functional gene analysis.

The single most useful by-product of the ES cell differentiation model is the fact that it provides insights to early developmental stages that are otherwise difficult to access in the embryo. One particular area of controversy that the model has shed light upon is the long-standing hypothesis that the haematopoietic and endothelial lineages develop from a common progenitor known as the haemangioblast. Formal demonstration that such a progenitor does indeed exist was provided by studies using the ES differentiation model (Ref: Choi K. *et al.*, 1998; Nishikawa S. *et al.*, 1998). Carefully timed analysis of very early EBs led to the identification of a progenitor known as the blast colony-forming cell (BL-CFC) that gives rise to blast colonies consisting of haematopoietic and vascular progenitors in methylcellulose cultures in the presence of vascular endothelial growth factor (Ref: Kennedy M. *et al.*, 1997). Subsequent studies have shown this progenitor expresses the receptor tyrosine kinase Flk-1, the transcription factor Runx1, and the mesoderm gene brachyury, suggesting it might represent a subpopulation of mesoderm undergoing commitment to the haematopoietic and vascular lineages. Recently, a similar progenitor has been demonstrated in the early mouse embryo confirming that BL-CFC is not an artefact of the ES cell model and that this entity is the best candidate for the illusive haemangioblast (Ref: Huber T.L. *et al.*, 2004).

## 1.3 GENE TRAPPING

### 1.3.1 Trapping genes in embryonic stem cells

Gene trapping in embryonic stem cells represents an efficient approach to the functional analysis of the murine genome (Refs: Joyner A.L., 1991; Evans M.J. *et al.*, 1997). In essence, gene trapping involves the random integration of a gene trap vector into genomic DNA. Where the site of insertion is within intronic DNA, the gene trap vector will be transcribed by the upstream promoter of the host gene. The resultant RNA transcript will then undergo nuclear processing whereupon the gene trap vector sequence is spliced into the mRNA via its own splice donor site and becomes contiguous with the gene's upstream exons (Refs: Skarnes W.C., 1993; Cecconi F. and Meyer B.I., 2000). The basic design of the gene trap vector comprises an antibiotic resistance gene in tandem with a reporter gene, thereby permitting selection of successful gene trap clones and the monitoring of the transcriptional activity of the trapped promoter in the cell (Ref: Skarnes W.C., 1992). Typically, the reporter gene is  $\beta$ -galactosidase and the selection gene is neomycin phosphotransferase (Ref: Friedrich G. and Soriano P., 1991) although the hygromycin phosphotransferase gene has also been successful utilised for gene trapping purposes (Ref: Natarajan D. and Boulter C.A., 1995). Gene trapping in ES cells represents an efficient approach to the functional analysis of the murine genome since it couples gene discovery with genomic mapping and expression pattern (Ref: Townley D.J. *et al.*, 1997). Moreover, since the integration event results in insertional mutagenesis at the site of the trapped gene locus, it is possible via blastocyst injection to derive chimaeric mice from which knockout animals can be bred (Ref: Auerbach W. *et al.*, 2000). This capacity to transmit the mutated gene trap locus into the germ line permits phenotype to be correlated with genotype and represents the most powerful facet of ES cell gene trapping technology (Refs: Brennan J. and Skarnes W.C., 1999; Cecconi F. and Gruss P., 2002; Chen W.V. and Soriano P., 2003).

### 1.3.2 Examples of genes identified by gene trapping

The functional characterisation of genes through gene trapping is exemplified by numerous examples: the secretory protein Sec8 required for paraxial mesoderm patterning (Ref: Friedrich G.A. *et al.*, 1997); a cardiac gene homologous to the human transcription factor TFEB and repressed by retinoic acid (Ref: McClive P. *et al.*, 1998); Bodenin, a novel gene expressed in restricted areas of the brain (Ref: Faisst A.M. and Gruss P., 1998); Aquarius, a mesodermal gene that is strongly induced by retinoic acid (Ref: Sam M. *et al.*, 1998); histone 3.3A gene, associated with sub-fertility, stunted growth, and neonatal lethality in transgenic mice (Ref: Couldrey C. *et al.*, 1999); Hzf and Hhl, two genes expressed during haematopoiesis and induced by culture on OP9 cells (Ref: Hidaka M. *et al.*, 2000); a novel gene expressed in foregut endoderm that serves as an marker of early liver development (Ref: Watt A.J. *et al.*, 2001); and Wnk1 kinase, deficiency of which in the heterozygous state is associated with significant fall in blood pressure (Ref: Zambrowicz B.P. *et al.*, 2003).

An analysis of the efficiency of the gene trap approach has been undertaken providing some quantitative measure of its success in yielding successful integration that will lead to a null mutation (Ref: Voss A.K. *et al.*, 1998). Finally, other modifications of the basic gene trap vector that incorporate a transmembrane peptide motif (Ref: Skarnes W.C., 1995) have been used to identify and functionally analyse cell surface proteins (Ref: Skarnes W.C., 2000) and secretory proteins (Ref: Mitchell K.J. *et al.*, 2001). Other gene trap vector modifications have included the incorporation of a thymidine kinase cassette for negative selection and identification of loci with potential haematopoietic function (Ref: Cannon J.P. *et al.*, 1999), development of a poly A trap that allows for capture of genes that are inactive in undifferentiated ES cells but become active during differentiation (Ref: Salminen M. *et al.*, 1998), and the use of a GFP reporter in a poly A trap retroviral vector that allows for reversible disruption and expression monitoring of genes in living cells (Ref: Ishida Y. and Leder P., 1999).

### 1.3.3 Strategies for screening of gene trap clones

In order to realise the undisputed potential of gene trapping for the identification and functional analysis of novel genes in ES cells, it is necessary to incorporate an effective screening strategy that will select those ES clones that harbour a gene trap integration within a gene of interest. Early gene trapping established that integration events occur predominantly within genes that lie adjacent to CpG islands (Ref: Macleod D. *et al.*, 1991) thereby allowing identification of genes active in ES cells. Subsequently, simple phenotypic screens were introduced that identified genes involved in early pattern formation (Ref: Hill D.P. and Wurst W., 1993). More sophisticated induction gene trap screens were developed to identify genes that were either induced or repressed by the presence of retinoic acid (Ref: Forrester L.M. *et al.*, 1996) that typically exhibited spatially restricted or tissue-specific patterns of expression and revealed a novel family of repeat sequences in the mouse genome responsive to retinoic acid (Ref: Sam M. *et al.*, 1996). Differentiation screens were developed to allow for the *in vitro* pre-selection of gene trapped ES clones that exhibit  $\beta$ -galactosidase expression in certain cell lineages during ES cell differentiation (Ref: Bonaldo P. *et al.*, 1998), thereby increasing the likelihood that blastocyst injection of the ES cells might yield a transgenic animal carrying a tissue-specific phenotype (Refs: Baker R.K. *et al.*, 1997; Stuhlmann H., 2003). Such an approach has generated a mouse strain, jumonji, that exhibits abnormal neural tube formation (Ref: Takaeuchi T., 1997). Finally, using a novel induction gene trap strategy to screen a library of sequence tagged gene trap clones (Ref: Wiles M.V. *et al.*, 2000) for homeoprotein gene targets, genes such as the bullous pemphigoid antigen 1, Bpag1, gene have been revealed and insertionally mutated (Ref: Mainguy G. *et al.*, 2000) by gene trapping.

### 1.3.4 High throughput screening approaches

The gene trap methodology lends itself to the efficient identification of genes whose functional disruption leads to a revealing biological phenotype (Ref: Joyner A.L. *et al.*, 1992). Increasingly,

the goal in gene trap technology is towards screening strategies that permit the high throughput identification of key genes (Ref: Zambrowicz B.P. and Friedrich G.A., 1998). Currently, there are of the order of approximately 35,000 genes identified by sequence and mapping of the human genome. However, expression data is available for only a relatively small proportion of these genes and functional data for still fewer again. Gene trap mutagenesis has established a public resource of over 8,000 ES cell lines, each carrying a potential loss-of-function insertionally-mutated gene ready for phenotypic analysis (Ref: Stanford W.L. *et al.*, 2001). Early large-scale screens simply looked for developmentally regulated genes from an extensive library of ES gene trap clones (Ref: Wurst W. *et al.*, 1995) or combined expression pattern data with DNA sequence information to provide a molecular characterisation of the trapped gene (Ref: Holzschu D. *et al.*, 1997). More sophisticated gene trap screens have been used for the large-scale disruption and identification of genes involved in the development of the mammalian nervous system (Ref: Stoykova A. *et al.*, 1998), of genes regulated during haematopoietic differentiation of ES cells (Ref: Muth K. *et al.*, 1998), and of genes destined for the endothelial lineage (Ref: Stanford W.L. *et al.*, 1998). Recently, with the completion of the sequencing of the mouse genome, it has become possible to undertake saturation mutagenesis of the mouse genome using gene trap vectors to insertionally mutate a major proportion of transcribed genes (Refs: Floss T. and Wurst W., 2002; Hansen J. *et al.*, 2003). Such a resource of gene trap ES cell clones, each clone carrying a sequence-tagged gene, has been made publicly available (Ref: Stryke D. *et al.*, 2003) and can be accessed via the website of the *International Gene Trap Consortium* ([www.igtc.ca](http://www.igtc.ca)). This library of cell lines has yielded a number of mutant mouse models, many of which have phenotypes that mimic the knockout animal. Similar technologies also lie behind at least one commercial venture (Omnibank™, Lexicon Inc.) whose stated goal is to perform saturation mutagenesis of mouse genome in order generate a library of transgenic mice for functional gene analyses (Ref: Zambrowicz B.P. *et al.*, 1998). Finally, by applying expression profiling of cDNA arrays obtained from gene trap mouse ES cell clones, it is possible to identify those genes whose absent expression suggests likely loss-of-function disruption (Ref: Matsuda E. *et al.*, 2004).

### 1.3.5 Strategies for generating homozygous gene trap mutants

The principal shortcoming of the gene trap methodology for elucidating gene function is that it can only lead to the loss, by random insertional mutagenesis, of one copy of the gene, thereby rendering the cell hemizygous for the mutation. In the presence of a normal functioning copy of the same gene on the homologous chromosome, this heterozygous state of affairs would typically be insufficient for yielding a discernible phenotype in ES cells, thus limiting the ability to perform recessive screens in cultured cells. As a consequence, the judgment of whether or not to derive chimaeric animals from the gene trap clone and proceed to breeding homozygous offspring must be based upon other data such as the expression pattern of the gene during *in vitro* ES cell differentiation or its inferred function as deduced from the DNA sequence (Ref: Van de Putte T. *et al.*, 2001). One strategy for circumventing this problem is based upon selection of gene trap clones for transgene homozygosity in order to create chromosome-specific loss of heterozygosity, thereby revealing recessive mutations through perturbation of the cellular phenotype (Ref: Lefebvre L. *et al.*, 2001). A more refined albeit similar approach has recently been described using ES cells that are deficient for the Bloom's syndrome protein (Blm) gene (Ref: Guo G. *et al.*, 2004). These cells exhibit a high rate of mitotic recombination that result in regions of loss of heterozygosity. By appropriate selection, it is possible to isolate gene trap ES cell clones that have undergone a chromosomal recombination event and duplication of the mutated gene, so converting the mutation from heterozygosity into homozygosity. This method, which involves a revertible gene trap retrovirus described previously (Ref: Ishida Y. and Leder P, 1999), has enabled the generation of a genome-wide library of mutant ES cell clones. By way of proof-of-principle of the methodology, this ES library has been successfully screened to isolate clones defective in DNA mismatch repair, revealing known and unknown genes in the process. In a similar fashion, using appropriate *in vitro* screens designed to detect discrete phenotype-based recessive genetic abnormalities, this novel approach will undoubtedly broaden the application of gene trapping by allowing phenotype and genotype to be more closely correlated at the cell level.

### 1.3.6 Functional homozygous knockouts by antisense gene trapping

A different approach that, like the one described above, aims to reveal recessive genetic mutations in diploid cells is termed random homozygous knockout (RHKO). This strategy, also referred to as antisense gene trapping, was originally used in NIH3T3 cells to reveal the tumour susceptibility gene, *tsg101* (Ref: Li L. and Cohen S.N. *et al.*, 1996). Antisense gene trapping permits the isolation of previously unknown genes encoding selectable recessive phenotypes through the random integration of a modified gene trap vector designed to yield an antisense RNA from a transactivated promoter. Thus, the trapped gene is homozygously disrupted, one copy through insertional mutagenesis and the other copy via functional antisense knockout. Removal of the transactivator, via Cre-lox recombination, permits the reversal of the antisense knockout, thereby reverting the cell back to a heterozygous state. As applied to NIH3T3 cells, the strength of this technique lay principally in the simplicity of the phenotypic screen, which selected cells exhibiting a transformed phenotype by growth in 0.5% agar and then tested their tumourigenicity by subcutaneous injection in nude mice. While claims that the human homologue of the *tsg101* gene was mutated in breast cancer (Ref; Li L. *et al.*, 1997) proved subsequently to be erroneous (Refs: Gayther S.A. *et al.*, 1997; Li L. *et al.* 1998), persuasive evidence has now accumulated to show that deficiency of this gene leads to genome instability (Ref: Xie W. *et al.*, 1998) possibly on account of its role in regulating the key cell-cycle genes *Mdm2* and *p53* (Ref: Li L. *et al.*, 2001). The latter appears to be borne out the fact that the targeted mouse knockout causes early embryonic lethality with accumulation of *p53* and defective cell proliferation (Ref: Ruland J. *et al.*, 2001) giving some confirmation that the original screening strategy was biologically valid.

This approach, of random homozygous knockout in NIH3T3 cells, has similarly been used to identify three other candidate tumour suppressor genes: the vasodilator stimulated phosphoprotein (*Vasp*) gene (Ref: Liu K. *et al.*, 1999), the calpain protease (*nCL-4*) gene (Ref: Liu K. *et al.*, 2000a), and a novel RNA-binding nuclear protein termed *Sam68* (Ref: Liu K. *et al.*, 2000b).

### **1.3.7 Screening functional ES gene trap knockout clones for “differentiation genes”**

Based upon its proven efficacy in identifying candidate tumour susceptibility genes in NIH3T3 cells, this strategy of random homozygous knockout (RHKO) outlined above was chosen here as a plausible experimental basis for deriving functional homozygous mutants in ES cells. The goal of the current work was to establish a proof-of-principle for this approach, hereafter referred to as antisense gene trapping, for the purpose of identifying recessive genes important for normal ES cell differentiation using an *in vitro* methylcellulose differentiation screen for mutant ES cell clones. The results are presented relating to the experimental efficiency of the methodology, the functional analysis of the trapped genes, and the technical shortcomings of the screening method chosen.

## 1.4 RNA INTERFERENCE

### 1.4.1 Silencing genes by RNA interference

RNA interference (RNAi) represents an evolutionarily ancient and highly conserved biological process that appears to be present in some form or other within most, if not all, eukaryotic organisms. As such, it represents an evolutionary relic of the earliest life forms on earth and might plausibly exist today as a distant echo of the so-called “RNA World”, when it is proposed that RNA may have been both the genetic material and biological catalyst for emerging life (Refs: Joyce G.F. and Orgel L.E., 1999; Gilbert W., 1986). When exposed to foreign genetic material (RNA or DNA), organisms mount a highly specific counter attack to silence the invading nucleic acid before the sequence is able to integrate into the host genome or subvert cellular processes. At the heart of this sequence-directed immunity is double-stranded RNA (dsRNA). The role of dsRNA in the interference process was first revealed in 1998 when it was shown that such molecules were able to mediate the potent and specific silencing of gene expression in the nematode worm, *Caenorhabditis elegans*, in a sequence directed manner (Refs: Fire A. *et al.*, 1998). To the considerable surprise of the investigators, this interference process, which was substantially more effective at gene silencing than either strand individually, could be produced throughout the body of the organism distant to the site of injection and could even be propagated to the progeny (Ref: Sharp P.A. and Zamore P.D., 2000).

With hindsight, it is clear that this interference phenomenon could explain a number of previously recognised but hitherto poorly explained processes (Ref: Guo S. and Kemphues K.J., 1995) whereby sense RNA had been found to be equally effective as antisense RNA in suppressing genes in *C. elegans* due to overlooked contamination of each preparation with dsRNA. Moreover, some years prior to this observation, plant biologists working with petunias had been surprised to find that the introduction of numerous copies of a gene that codes for deep purple flowers led, not

as expected, to an even darker purple hue, but rather to plants with white or patchy flowers (Refs: Napoli C. *et al.*, 1990; van der Krol A.R. *et al.*, 1990). Somehow the transgenes had silenced both themselves and the plants' own "purple flower" genes. Similarly, when plants were infected with an RNA virus that had been genetically engineered to contain fragments of a plant gene, the gene itself became silenced (Ref: Wassenegger M. *et al.*, 1994).

This phenomenon of "co-suppression" in plants later came to be termed post-transcriptional gene silencing (PTGS) and was shown to be functionally analogous to a similar process called "quelling" identified in the fungus *Neurospora crassa* which, when transformed with homologous sequences, underwent transient gene inactivation (Ref: Romano N. and Macino G., 1992). In nematodes, the usefulness of RNAi as a tool for functional gene analysis was enhanced still further by the discovery that genes could be efficiently silenced simply by soaking the worms in dsRNA (Ref: Tabara H. *et al.*, 1998) or providing food containing bacterially expressed dsRNA (Refs: Timmons L. and Fire A., 1998; Timmons L. *et al.*, 2001). The early realisation that RNAi is evolutionarily highly conserved, sharing a common underlying mechanism in many organisms, was soon followed by the demonstration of RNAi-induced gene silencing in *Drosophila* (Ref: Kennerdell J.R. and Carthew R.W., 1998) as well as several other organisms, some of which were otherwise unsuited to genetic analysis (Refs: Ngo H. *et al.*, 1998; Boshier J.M. and Labouesse M.).

#### **1.4.2 Short interfering RNAs (siRNAs)**

Early insight into the molecular mechanism responsible for the RNA interference came from the study of post-transcriptional gene silencing in plants which revealed that this process was accompanied by the production of short RNA molecules, some 20 to 25 nucleotides long, that matched the gene being silenced (Ref: Hamilton A.J. and Baulcombe D.C., 1999). These RNAs are much shorter than typical mRNAs and ribosomal RNAs, perhaps explaining why this fundamental process remained undiscovered until relatively recently.

At around the same time, efforts were made to recapitulate the biochemistry of the RNAi process *in vitro* using fruitfly extracts (Refs: Tuschl T. *et al.*, 1999; Zamore P.D. *et al.*, 2000). These studies demonstrated that the long dsRNAs used to trigger gene silencing in cells were diced up, by enzymes in the extract, into short RNA molecules that shared a common structure (Ref: Tuschl T., 2001). This structure was that of an RNA duplex comprised of two 21 nucleotide strands paired together in a staggered fashion so by leaving two unpaired nucleotides at each ends (Ref: Elbashir S.M. *et al.*, 2001a). This molecular species was dubbed a short interfering RNA (siRNA) and, in a landmark paper, it was demonstrated that these entities, when synthesised exogenously and introduced into mammalian cells, were able to mediate sequence specific interference of co-transfected target reporter genes (Ref: Elbashir S.M. *et al.*, 2001b). This use of synthetic siRNAs, 21-22 nucleotides in length, were found to bypass the well-documented so-called “interferon response”, which represents a non-specific anti-viral response triggered by longer dsRNA, leaving just the sequence-specific effect directed at the target gene (Ref: Caplen N.J. *et al.*, 2001). This discovery represents the first demonstration of an RNAi-induced phenotype in cultured mammalian cells and was soon extended to the knock down of endogenous genes also, thereby rendering the entire human genome theoretically amenable to this powerful new technology (Ref: Harborth J. *et al.*, 2001; Elbashir S.M. *et al.*, 2002).

### 1.4.3 Processing of double-stranded RNAs

The enzyme responsible for the cleavage of dsRNA into siRNAs is termed Dicer (Ref: Bernstein E. *et al.*, 2001). This is an RNase-III-type endonuclease that has a dsRNA-binding domain and a catalytic domain (Ref: Ketting R.F. *et al.* 2001). Processing of the dsRNA by Dicer yields RNA duplexes of about 21 nucleotides in length that have 5' phosphates and 2-nucleotide 3' overhangs (Ref: Elbashir S.M. *et al.*, 2001). Several organisms, such as *Drosophila* and *Arabidopsis*, contain more than one Dicer gene with each Dicer preferentially processing dsRNAs that come from a specific source, while in mammals and nematodes there appears to be only a single such enzyme.

In the mouse, disruption of the *Dicer1* gene by targeted mutation results in embryonic lethality and depletion of stem cells in *Dicer1* null embryos suggesting a functional link between RNA interference and early mammalian development (Ref: Bernstein E. *et al.*, 2003). In *C. elegans*, using a screen for RNAi deficient strains, two other key genetic elements of the RNAi process, the *rde-1* and *rde-4* loci, have been identified (Ref: Tabara H. *et al.*, 1999) and shown to mediate distinct and non-overlapping roles in the interference pathway (Ref: Parrish S. and Fire A., 2001).

#### 1.4.4 Molecular mechanism of gene interference

Following processing of the dsRNA into siRNA duplexes by Dicer, the siRNA molecules are incorporated into a complex of proteins termed the RNA-induced silencing complex (RISC) that appears to be responsible for mediating the actual interference process (Ref: Hammond S.M. *et al.*, 2000). This siRNA-containing effector complex, which has an estimated molecular mass of between 130 and 160 kDa in human cells (Ref: Martinez J. and Tuschl T., 2004), includes one key component that belongs to the Argonaute (Ago) protein family and appears to be responsible for direct binding to the RNA (Ref: Fagard M. *et al.* 2000). The assembly of the RISC is an ATP-dependent process that likely reflects the requirement for energy-driven unwinding of the siRNA that is required for sequence-specific base-pairing between the antisense strand of the duplex and the sense strand of the target mRNA (Ref: Nykanen A. *et al.*, 2001). The cleavage of the mRNA within the region of complementary base pairing with the antisense strand of the siRNA results in the post-transcriptional destruction of the gene with consequential knockdown of its expression (Ref: Martinez J. *et al.*, 2002). The RISC cleaves the target mRNA in the middle of the complementary region, ten nucleotides upstream of the nucleotide paired with the 5' end of the guide siRNA (Ref: Elbashir S.M. *et al.*, 2001c), after which the mRNA fragments are released and the RISC is able to mediate cleavage of further mRNA strands. In this manner, the siRNA-induced RISC-mediated interference process is highly efficient at bringing about the functional knock down of the target gene (Ref: Mello C.C. and Conte D., 2004).

### 1.4.5 Biochemistry of the RISC

A number of different proteins have been identified as essential for RNA interference or as components of the RISC. Of these it is the Dicer and Argonaute proteins that appear to play the major roles in assembly and function of the RISC. Dicers are approximately 200kDa proteins and generally contain ATPase/RNA helicase and PAZ domains, two catalytic RNase III domains, and a C-terminal dsRNA binding domain. Their primary role is to process precursor molecules into siRNAs and miRNAs, but the enzymes also function in downstream steps of RNAi. Vertebrates and *C. elegans* contain single Dicer genes, while *Drosophila* expresses at least two Dicers with specialised functions. Of the two *Drosophila* Dicers, Dcr-1 functions mainly in the processing of miRNA precursors, while Dcr-2 is required for RNAi (Ref: Lee Y.S. et al., 2004). Argonautes are 100kDa proteins with characteristic PAZ and PIWI domains and are the only proteins consistently found in all RISC and miRNP complexes. They fall into two families, namely Ago and Piwi. Of the Ago proteins, four are ubiquitously expressed in mammals and two in *Drosophila*, Ago1 functioning primarily in miRNA and Ago2 in RNAi pathways.

siRNA molecules are the effector molecules of RNA interference. They are short double-stranded RNAs of which only one strand, referred to as the guide strand, is incorporated into the RISC, while the other strand, a passenger strand, is discarded. The choice of which strand becomes chosen for incorporation appears to be determined by whose 5'-terminus resides at the thermodynamically less stable end of the duplex. In *Drosophila*, a heterodimer on one of its Dicers, Dcr-2, and another protein termed R2D2 senses the differential stability of the duplex ends and determines which strand will enter the RISC. The Dcr-2-R2D2-siRNA ternary complex is the best-characterised assembly in the RISC formation pathway that pre-emptly the unwinding of the duplex within a sub-compartment termed the RISC loading complex. The Dcr-2-R2D2 heterodimer is then displaced by Ago2 which then catalyses the cleavage of the target mRNA thereby carrying out the main functional role of the RISC.

That the Argonaute protein is indeed the catalytic engine of the RISC was shown by the demonstration that an active RISC can be reconstituted from a bacterially expressed human Ago2 and a single-stranded siRNA (Ref: Rivas F.V. *et al.*, 2005). Within the resulting Ago-siRNA complex, the 5'-phosphorylated guide strand end docks with the PIWI pocket and the opposite 3' end contacts the PAZ domain (Ref: Tomari Y. *et al.*, 2004). The elucidation of the crystal structure shows that the phosphorylated 5'-terminal nucleotide of the guide strand is anchored within a highly conserved basic pocket stabilising the binding of the siRNA. Evidence suggests that the 5'phosphate is required for efficient unwinding and release of the passenger strand than for subsequent steps of the RISC cycle.

Within the RISC, mRNA cleavage occurs between residues base paired to nucleotides 10 and 11 of the siRNA, and the cleavage itself does not require ATP. The guide siRNA remains associated with the complex, allowing it to carry out multiple rounds of RNA cleavage. The turnover of the enzyme is dependent on ATP, suggesting that release of the cleaved mRNA halves involves an RNA helicase. Two models of Argonaute function in the RISC have been proposed. These are the “two-state” model proposed by Tomari and Zamore (Ref: Tomari Y. and Zamore P.D., 2005) and the “fix-ends” model (Ref: Ma J.B. *et al.*, 2004). In essence, while the former requires recurrent disruption of the guide-PAZ contact, the latter does not. Further evaluation of these models is required to determine which one best describes the molecular machinery of the RISC complex.

Finally, little is known about how the RISC processes miRNAs which are increasingly being revealed to play important roles in normal and abnormal cellular physiology. This, together with the further understanding of the other proteins comprising the RISC, remains a major challenge in elucidating the mechanism of post-transcriptional RNA interference (Ref: Filipowicz W. 2005).

### 1.4.6 Alternative theories of RNA interference

Despite the model outlined above, the complete dissection of the RNAi silencing pathway remains to be fully elucidated. Moreover, it is also becoming increasingly apparent that there exist distinct functional differences between various species in the exact mechanics of the interference process (Ref: Bass B.L., 2001). Thus, a key gene family involved in the silencing pathways in plants (Ref: Dalmay T. *et al.* 2000), fungi (Ref: Cogoni C. and Macino G.. 1999), and *C. elegans* (Ref: Sijen T. *et al.*, 2001) contains genes that encode putative cellular RNA-dependent RNA polymerases (RdRPs). Members of this family of proteins were identified in forward genetic screens and were found in plants to be required for transgene-induced silencing but not for viral genes, perhaps because viruses provide their own RNA polymerase. The mechanism whereby the RdRP mediates silencing appears to be via generation of a dsRNA intermediary upon recognition of foreign nucleic acid, which is then processed by Dicer leading to gene interference as per the classical route. In support of this mechanism is the observation of so-called transitive RNAi whereby, in *C. elegans*, dsRNA that targets one gene is able also to silence other splice variants of the gene, even if they do not share any sequence homology (Ref: Alder M.N. *et al.*, 2003). Such a mechanism has led some investigators to postulate that RNAi represents a process of “random degradative PCR” in which siRNA primers are able to convert mRNA into dsRNA that is then degraded by Dicer into new siRNAs thereby setting up an RNA-dependent polymerase chain reaction (Ref: Lipardi C. *et al.*, 2001). Such a mechanism would explain, through the inherent amplification that results from a small amount of “trigger” dsRNA, one of the most intriguing features of the interference process, namely its remarkable potency and ability to propagate throughout an organism (Ref: Nishikura K., 2001). However, while this theory has intrinsic appeal, it does not appear to be broadly applicable to all organisms and nor is RNA-dependent RNA polymerase activity an absolute requirement for successful RNAi-induced gene knockdown.

### 1.4.7 Transcriptional gene silencing

As described above, there are many common features between post-transcriptional gene silencing in plants and RNA interference in animals (Ref: Ketting R.F. and Plasterk R.H., 2000). However, by intricate experimental dissection of the gene silencing mechanism in *Arabidopsis*, it appears that the silencing effect in plants is mediated by two distinct molecular species of siRNA, one that is 21-22 nucleotides long and the other that is 24-26 nucleotides long (Ref: Hamilton A. *et al.*, 2002). While the former is associated with mRNA degradation, the latter correlates more closely with transcriptional gene repression operating via methylation of the genomic DNA (Ref: Timmons L., 2002). This consolidates a more permanent state of gene suppression that is induced both by transgene and viral RNAs (Refs: Mette M.F. *et al.*, 2000; Jones L. *et al.*, 2001). Recently these findings have been extended by the observations that siRNA-directed DNA methylation in plants is also linked to chromatin structure (Ref: Morel J.B. *et al.*, 2000) and histone modification (Refs: Zilberman D. *et al.*, 2003). The latter discovery tallies closely with the demonstration in fission yeast that heterochromatin formation at centromere boundaries is associated with siRNAs (Refs: Volpe T. *et al.*, 2002; Reinhart B.J. and Bartel D.P., 2002).

The notion that RNA interference might play a role in transcriptional regulation via chromatin modification is an attractive one (Ref: Allshire R., 2002) and suggests that gene silencing pathways might operate at more than one level. Such a conclusion is supported also in *Drosophila* where two types of transgene silencing were revealed in the case of the alcohol dehydrogenase unit (Ref: Pal-Bhadra M. *et al.*, 2002), one operating transcriptionally and is Polycomb dependent and the other operating post-transcriptionally with molecular features of classical RNAi. Moreover, recent data suggests that the RNAi machinery itself appears to play an indispensable role in the process that leads to heterochromatic silencing of genes in *Drosophila* cells (Ref: Pal-Bhadra M. *et al.*, 2004). Similarly, in *C. elegans*, using an elegant strategy whereby RNAi was used to identify genes required for RNAi, suppression of three *mes* genes (*mes-3*, *mes-4*, *mes-6*),

that are homologues of the *Drosophila* chromatin-binding Polycomb-group proteins, has identified their critical role in mediating gene silencing (Ref: Dudley N.R. *et al.*, 2002).

#### 1.4.8 Role of the RNAi machinery in heterochromatin formation

Thus, it is evident that components of the RNAi machinery can be involved in nuclear processes leading to heterochromatin formation and transcriptional gene silencing. Indeed, this process of nuclear RNAi appears to be a natural epigenetic gene regulation mechanism that is active in most eukaryotes and controls heritable changes in gene expression that are not caused by mutations. It also serves as a surveillance mechanism against foreign nucleic acids, is involved in the regulation of developmental genes, and contributes to accurate chromosome segregation during cell division.

The mechanism of RNA-mediated heterochromatin formation is perhaps best understood in *S. pombe*. Here, the initial step appears to require bidirectional transcriptional of the target loci to produce primary dsRNAs. From this, siRNAs are generated via the action of Dicer. These are incorporated into the RNA-induced initiator of transcriptional silencing (RITS) complex via the Argonaute protein, Ago1, thereby directing the RITS to complementary sites of the genome. It is unclear whether the RITS binds only to nascent target transcripts or pairs to corresponding DNA. Either way, binding of the complex enables the recruitment of chromatin-modifying proteins, including a histone H3 lysine 9 (H3K9) methyltransferase (Clr4). This appears to stabilise the binding of the RITS to the chromatin and suggests the likelihood of RITS/DNA binding in addition to RNA binding. Once stabilised, the RITS interacts with the RNA-directed RNA polymerase complex leading to the generation of secondary dsRNA using nascent transcripts as a templates (Ref: Motamedi M.R. *et al.*, 2004). Thus, a self-perpetuating loop is established that consolidates the heterochromatinisation of the gene locus. It also implies that this mode of gene silencing depends upon the primary transcription of the target locus to initiate the whole cascade of heterochromatin formation (Ref: Wassenegger M., 2005).

### 1.4.9 RNAi-induced epigenetic changes in mammalian cells

In the case of mammalian cells, recent evidence suggests that, here also, RNAi might play a key role in regulating gene expression at the transcriptional level. Using human tissue culture cells, it has been shown that promoter directed siRNAs inhibit transcription of an integrated proviral EF1A promoter-GFP reporter gene, and of endogenous EF1A (Ref: Morris K.V. *et al.*, 2004). Moreover, this silencing, which is dependent upon transport of the siRNA into the nucleus, was associated with DNA methylation of the targeted sequence. This evidence is further strengthened by another study that similarly demonstrated that synthetic siRNAs targeted to the CpG islands of an E-cadherin promoter are able to induce significant DNA methylation and histone H3 methylation both in the MCF-7 human cell line and in primary mammary epithelial cells (Ref: Kawasaki H. and Taira K., 2004). As a result, these siRNAs repressed expression of the E-cadherin gene at the transcriptional level. Likewise, vector-based siRNAs targeted to the erbB2 promoter also induced CpG methylation and gene down-regulation in MCF-7 cells. That this phenomenon was dependent upon the presence of a DNA methyltransferase was confirmed by the demonstration that disruption, by RNAi, of either the DNMT1 or DNMT3B genes leads to the abolishment of this siRNA-mediated DNA methylation. Both these papers point to the presence of a fundamental cellular process that exists in mammalian cells whereby the gene silencing effects of RNA interference can be consolidated at the level of genomic DNA.

### 1.4.10 Novel strategy variations designed to induce gene interference

The use of siRNAs to functionally dissect the mammalian genome has been based in the main upon the introduction of exogenously generated siRNAs, either synthesised chemically or by *in vitro* transcription (Ref: Donze O. and Picard D., 2002), or by transfection of PCR products from which sequence-specific siRNA molecules are transcribed intra-cellularly (Ref: Castanotto D. *et al.*, 2002). However, in view of the problems relating to delivery of the siRNAs into the cell and

the transience of the knockdown effect, other more novel strategies have been devised that can achieve stable suppression of gene expression. Such strategies typically rely upon the integration of a suitably designed construct which, when transcribed, yields a hairpin RNA molecule that functionally mimics the siRNA (Ref: Paddison P.J. *et al.*, 2002a). So-called short hairpin RNAs (shRNAs) undergo processing by Dicer into active siRNA molecules that are able to induce sequence-specific silencing in mammalian cells (Ref: Brummelkamp T.R. *et al.*, 2002). Typically, this shRNA construct is transcribed from an RNA polymerase III promoter (Ref: Paddison P.J. *et al.* 2002b) or alternatively from a U6 promoter that leaves four uridine 3' overhangs (Ref: Miyagishi M. and Taira K., 2002). A further recent elaboration of this methodology has resulted in the development of an inducible and reversible system for the conditional knockdown of a target gene (Ref: Gupta S. *et al.*, 2004).

#### **1.411 Translation of RNAi technology *in vivo***

The potential for applying RNAi to therapeutic ends by using shRNAs to direct knockdown of disease-related genes *in vivo* has been realised in a proof-of-principle experiment demonstrating the suppression of a reporter gene *in vivo* in mice and also the effective targeting of a sequence from the hepatitis C virus (Ref: McCaffrey A.P. *et al.*, 2002). The development of a lentivirus-based delivery system, for the efficient introduction of the DNA construct encoding the shRNA into cycling and non-cycling mammalian cells, represents an important further advance in achieving this goal of delivering therapeutically-directed shRNAs (Ref: Rubinson D.A. *et al.*, 2003). To this end, another proof-of-principle experiment has established that viral delivery of shRNA-expressing DNAs can markedly suppress both exogenous and endogenous genes *in vivo* in the mouse brain and liver (Ref: Xia H. *et al.*, 2002).

Other strategies for delivery of functional siRNA molecules have employed cationic liposomes (Ref: Sorensen D.R. *et al.*, 2003), facile high-pressure injection methods (Ref: Lewis D.L. *et al.*,

2002), and stabilising chemical modifications (Ref: Soutschek J. *et al.*, 2004). In addition to viral genes expressed by hepatitis C (Ref: Davidson B.L., 2003) and HIV-1 (Ref: Novina C.D. *et al.*, 2002), cancer genes have also been successfully targeted including the leukaemia fusion oncogenes BCR/ABL (Ref: Wilda M. *et al.*, 2002) and AML/ETO (Ref: Heidenreich O. *et al.*, 2003). These results testify to the considerable promise for the use of siRNA technology for therapeutic (Ref: Wall N.R. and Shi Y., 2003).

#### 1.4.12 RNAi as applied to embryonic stem cells

The first demonstration of RNA interference in murine cells used long double-stranded RNA molecules introduced into fertilised oocytes to induce knockdown of specific genes in pre-implantation embryos (Ref: Wianny F. and Zernicka-Goetz M., 2000). The effects of the interference were gene-specific and mimicked the phenotypes of the null mutants. This was followed by the demonstration of reporter gene knockdown, also in mouse oocytes, using expression vectors encoding short hairpin RNA (shRNAs) molecules (Ref: Svoboda P. *et al.*, 2001). The gene knockdown effects of long dsRNA molecules were subsequently revealed in mouse embryonal teratocarcinoma cell lines (Ref: Billy E. *et al.*, 2001) and in undifferentiated embryonic stem cells (Ref: Yang S. *et al.*, 2001). This gene knockdown could be rendered stable in embryonic stem cells using DNA vectors encoding short hairpin RNAs (Ref: Tang F.C. *et al.*, 2004). RNAi technology has found widespread application in dissecting the molecular processes leading to ES cell differentiation, as exemplified by the use of shRNAs to knockdown the Oct4 gene in murine cells (Ref: Velkey J.M. and O'Shea K.S., 2003) and siRNAs to do similarly in human cells (Ref: Hay D.C. *et al.*, 2004), in both cases inducing similar patterns of endoderm and trophoblast differentiation. Likewise, siRNAs have been used to key haematopoietic transcription factors in differentiating ES cells to influence lineage commitment (Ref: Zou G.M. *et al.*, 2003). Using the method of tetraploid aggregation to generate fully ES-derived embryos, ES cells carrying stably integrated vectors expressing gene-specific shRNAs have yielded mutant embryos

that mimick the genetic null phenotype (Ref: Kunath T. *et al.*, 2003). Thus, RNAi has rapidly established itself as an invaluable tool in the molecular understanding of differentiation both in ES cells and early embryonic development.

#### **1.4.13 Theories on the biological role of small RNA molecules**

The discovery that small RNA molecules exhibit sequence-directed properties has been one of the most remarkable breakthroughs in molecular biology in recent years (Ref: Zamore P.D., 2001). However, the physiological role of RNA interference remains largely conjecture. The fact that it is evolutionarily ancient and has been highly conserved throughout eukaryotes suggests that it serves indispensable cellular functions (Ref: Zamore P.D., 2002). The most attractive theory is that this mechanism of RNA silencing provides the cell's genome with its own immune system against invading transposons and viruses (Ref: Platerk R.H., 2002). Such a defence mechanism has dual attributes of specificity and potency necessary for the sequence-directed suppression of mobile genetic elements and silencing of foreign viral genes.

An alternative theory argues that RNAi is a major physiological regulator of gene expression. Such an argument has received further credence by the discovery of countless so-called microRNAs (Ref: Ambros V., 2004). These are small hairpin RNA molecules that regulate the expression of complementary messenger RNAs with roles in developmental timing, cell death, cell proliferation, haematopoiesis and patterning of the nervous system. Hundreds of microRNA genes have been found in diverse animals, and many of these are phylogenetically conserved suggesting that their regulatory impact is extremely pervasive. Indeed, bioinformatic analyses indicate that target genes in vertebrate species may number in the thousands. These target genes are regulated by the microRNAs through down-regulation at transcriptional, post-transcriptional and translational levels via RNA silencing pathways that are similar, if not identical, to those classical RNA interference (Ref: Sontheimer E.J., Carthew R.W., 2005).

#### 1.4.14 Future prospects for RNAi in biology and medicine

The realisation that small RNA molecules can carry out diverse physiological roles within the cell, including regulation of gene expression and interference of foreign genetic elements, represents a landmark discovery in molecular biology. Moreover, because of the ease with which a cell's genes can be knocked out using homologous double-stranded RNAs, RNA interference has become a convenient tool for the functional analysis of genes in many model organisms (Ref: Paddison P.J. and Hannon G.J., 2002). The identification of short interfering RNAs, as an intermediate in the interference process that can be used to knock down gene targets in mammalian cells, represents a timely advance in the post-genome era for the functional dissection of the mouse and human genomes (Ref: McManus M.T. and Sharp P.A., 2002). There is an emerging consensus on the necessary safeguards required for stringent experimental control when using RNAi as an experimental methodology (Ref: Lassus P. *et al.*, 2002), although these almost invariably affirm the specificity of the interference effect. The most powerful application of the RNAi technology has been when combined with high-throughput screening strategies to produce an exhaustive functional trawl of genes that yield a particular phenotype upon knockdown. Such functional genomic analyses have been performed on *C. elegans* (Ref: Fraser A.G. *et al.*, 2000) and *Drosophila* cells (Ref: Kiger A. *et al.*, 2003), and libraries of validated short hairpin RNAs have been constructed covering large portions of the mouse and human genomes (Ref: Paddison P.J. *et al.*, 2004).

Unsurprisingly, RNAi holds considerable promise in medicine too, as revealed by the proof-of-principle demonstration of knockdown of a variety of disease-related genes. The recent landmark publication describing a simple method for delivering siRNAs in a living mouse disease model (Ref: Soutschek J. *et al.*, 2004), represent a key step in realising this therapeutic potential.

## 1.5 AIMS OF THESIS

The central goal of this thesis is the examination of two different strategies for the functional analysis of genes in mammalian cells. These are *antisense gene trapping*, presented in results chapter one, and *RNA interference*, presented in results chapters two, three, and four. The individual aims of each of the chapters are as below:

- ❖ **Aim I:** To attain proof-of-principle for the use of antisense gene trapping in murine embryonic stem cells for the identification of genes relevant to differentiation.
  
- ❖ **Aim II:** To demonstrate that short interfering RNA molecules are able to knockdown reporter genes in mammalian cells and to explore the biology of this phenomenon.
  
- ❖ **Aim III:** To define the molecular determinants for effective interference and to identify the site of action of short interfering RNAs in mammalian cells.
  
- ❖ **Aim IV:** To investigate the feasibility of using short interfering RNAs to knockdown endogenous gene targets in various cell lines including murine embryonic stem cells.

# **Chapter 2**

# **Methods**

## 2.1 CELL CULTURE

### 2.1.1 Culture of embryonic stem cells

ES cells were cultured according to standard protocols (Refs: Smith A.G., 1991; Robertson E.J., 1987). ES cell culture medium was made up using Dulbecco's Modified Eagle's Medium (DMEM) containing 10% foetal calf serum and supplemented with non-essential amino acids (100 $\mu$ M), sodium glutamate (2mM), sodium pyruvate (1mM), 2-mercaptoethanol (1/10,000) and leukaemia inhibitory factor, LIF, (100U/ml). Alternatively, 10x stock concentration Glasgow's Modified Eagle's Medium (GMEM) was used at a 1 in 10 dilution in distilled autoclaved water that was appropriately buffered to a pH of 7.4 with sodium bicarbonate. The foetal calf serum was selected prior to purchase by careful batch testing using a limiting dilution clonal assay to ensure optimal ES cell growth conditions. Typically, medium was also supplemented with 1% penicillin/streptomycin (50U/ml and 50 $\mu$ g/ml respectively) to prevent bacterial contamination. All the above reagents were purchased from *Invitrogen Inc.* except serum, which was purchased from various commercial suppliers. The LIF was derived in-house using conditioned medium from a recombinant COS-7 cell line and assayed according to its ability to maintain ES cells in an undifferentiated state (Ref: Robertson E.J., 1987).

ES cells were cultured on gelatin coated tissue culture grade plastic flasks/plates and passaged when confluent. Passaging of ES cells involved the removal of serum-containing medium followed by washing in phosphate buffered saline (PBS) and finally treatment for one to two minutes with 0.1% trypsin (*Invitrogen Inc.*). The cells were then re-suspended in fresh medium and re-plated at an approximate 10x dilution, typically  $4 \times 10^4$  cells per  $\text{cm}^2$  (or  $10^6$  per T25 flask) using a haemocytometer to count cell density when necessary. Most experiments were performed using the stromal-independent ES cell line, E14, (Ref: Fisher J.P. *et al.*, 1989) which can be passaged in the absence of feeder cells. Early experiments were performed using the stromal-

dependent ES cell line, CCE, which requires passaging of cells on a bed of irradiated primary embryonic fibroblasts (PEFs) derived from day 12.5 mouse embryos.

### 2.1.2 Production of primary embryonic fibroblasts

Primary embryonic fibroblasts (PEFs) were generated from day 12.5 mouse embryos homozygous for a targeted knockout of the GDI D4 locus (Ref: Lelias J.M. *et al.*, 1993) on chromosome 12 (Ref: Adra C.N. *et al.*, 1994). These fibroblasts therefore carry two copies of the neomycin-resistance gene that is driven from a constitutive phosphoglycerate kinase (PGK) promoter (Ref: Guillemot J.C. *et al.*, 1996). The heads of the embryos were removed and the torsos were dissected mechanically and disaggregated using trypsin and collagenase. Embryonic fibroblasts were expanded for several passages until a population of homogenous fibroblasts were obtained (Ref: Robertson E.J., 1987). These were then harvested and irradiated using a standard dosage of gamma irradiation calibrated to yield viable but non-dividing fibroblasts.

### 2.1.3 Methylcellulose differentiation of embryonic stem cells

24g methylcellulose powder was dissolved in Iscove's Modified Dulbecco's Medium (IMDM) and autoclaved to generate a 2x stock of sterile methylcellulose suspension. This suspension was then submitted to a number of freeze-thaw cycles, with stirring, until the consistency was even and semi-fluid. A 1.2% methylcellulose differentiation mixture was then made up using the following recipe: 50% 2x methylcellulose suspension, 20% foetal calf serum, 10% ES cells suspended at 10,000 cells per ml in fresh medium (without LIF). The remaining 20% was comprised of DMEM supplemented with 1% monothioglycerol (10 $\mu$ l in 4ml stock), 1% penicillin/streptomycin (50U/ml and 50 $\mu$ g/ml respectively), 1% sodium glutamate (200mM), and 1% non-essential amino acids (10mM). Typically, growth factors were also added as follows: erythropoietin at 2U/ml, murine

interleukin-1 at 50U/ml, murine interleukin-3 at 500U/ml, stem cell factor at 10µg/ml, made up in a 1% cocktail mixture (Ref: Wiles M.V. *et al.*, 1991).

The methylcellulose-cell suspension was made up to a volume of 3mls at a concentration of 1000 cells/ml and this was split equally onto two 4cm non-tissue culture (bacterial) dishes which were placed together with a third dish, containing water for humidification, onto a 10cm plate. The plate was then incubated at 37°C. After a few days, clonally derived embryoid bodies emerged and underwent differentiation that was assessed visually under light microscopy (Ref: Keller G, 1995).

#### **2.1.4 Electroporation of embryonic stem cells**

Two ES cell lines were used, namely, CCE cells, which are cultured on irradiated fibroblasts, and E14 cells, which are a feeder-independent line. Otherwise, the protocol employed for electroporation of plasmid constructs was identical in both cases and outlined below:

Cells were grown to semi-confluence on a T75 flask giving approximately  $1-2 \times 10^7$  cells. They were then trypsinised using 0.25% trypsin (*Invitrogen Inc.*), ensuring single cell suspension with gentle pipetting, re-suspended in complete medium, and centrifuged for 5 minutes at 1,000rpm. The pellet was then re-suspended in PBS and the cell density estimated using a haemocytometer prior to re-centrifugation. Finally, the cells were once again re-suspended, this time at a concentration of  $10^7$  per ml in PBS or electroporation buffer (20mM Hepes pH7, 137mM NaCl, 5mM KCl, 0.7mM Na<sub>2</sub>HPO<sub>4</sub>, 6mM dextrose, 0.1mM β-mercaptoethanol, distilled water, filter-sterilised). An aliquot of 0.8ml was transferred to an Eppendorf tube and 10-50µg of plasmid DNA (either circularised or linearised) was added and gently mixed. The cells were allowed to stand for 10 minutes at room temperature before transfer to a 0.8ml volume electroporation cuvette with an electrode gap of 0.4cm. The electroporation apparatus used was a BioRad

GenePulser™ machine (*BioRad Inc.*) and the specific parameters used were voltage of 800V and capacitance of 3  $\mu$ F with a discharge time constant of approximately 4-5 milliseconds. Following electric discharge, the cells were immediately re-suspended in a suitable volume of complete medium and re-plated on four 10cm diameter gelatin-coated tissue-culture dishes at a concentration of approximately  $2 \times 10^6$  cells per dish.

### 2.1.5 Antibiotic selection of embryonic stem cell clones

Electroporated cells were allowed to recover by growth in fresh complete medium for approximately 24-36 hours. This was then replaced with medium containing the appropriate antibiotic selection agent at lethal concentration. This concentration was gauged for each cell line and sub-line by prior testing using a range of antibiotic dosages and confirmed by ensuring 100% lethality of parallel control plates (containing cells electroporated in the absence of plasmid).

Three antibiotic selective agents were used in the course of these experiments, namely neomycin and hygromycin, which are toxic to wild-type cells, and ganciclovir, which is only toxic to cells expressing the adenoviral thymidine kinase for which it is a suitable substrate. The toxic dosages of these three drugs are as listed below:

- Neomycin 300-400 $\mu$ g/ml
- Hygromycin 200-300 $\mu$ g/ml
- Ganciclovir 1-2 $\mu$ M

Fresh antibiotic-containing medium was initially replaced daily, while cell death was extensive, and subsequently on alternate days. Discrete colonies emerged at around day 8-10 after electroporation. These were left to grow to a size where they could be successfully picked using a fine long-tipped 200 $\mu$ l pipette tip under microscope guidance in a laminar-flow hood. Each

colony was picked directly into an individual well of a 96 well plate and disaggregated by brief incubation in a drop of trypsin before re-plating on a 48 well plate. Cells were then expanded under reduced antibiotic stringency, typically 25% of the selection concentration. Clones were split when confluent onto a number of wells of a 24 well plate for LacZ staining, RNA extraction, freezing, and further downstream experimentation.

### 2.1.6 Assay for expression of $\beta$ -galactosidase

$\beta$ -galactosidase expression was assessed by *in situ* staining of cells using the following protocol:

- 1) Cells were washed in PBS and fixed in freshly-prepared fixative solution (0.4ml 25% glutaraldehyde, 2.5ml 100mM EGTA, 0.1ml 1M magnesium chloride, 47ml 0.1M sodium phosphate at pH 7.3), or alternatively in 3% paraformaldehyde in PBS, for 15 minutes.
- 2) Cells were washed three times in wash buffer (0.1ml 1M magnesium chloride, 0.5ml 2% NP40, 0.5ml 1% deoxycholate, 48.9ml 0.1M sodium phosphate pH 7.3).
- 3) Cells were stained in X-gal stain (0.106g potassium ferrocyanide, 0.082g potassium ferricyanide, 2ml 25mg/ml X-gal reagent in dimethylformamide, 48ml 0.1M sodium phosphate pH 7.3) and incubated from 4 hours to overnight at room temperature or 37°C, as appropriate, alongside wild-type cells as negative controls.

$\beta$ -galactosidase staining reflects the transcriptional activity of the trapped promoter in driving expression of the  *$\beta$ geo* gene via the splice acceptor bridge. This can be performed on undifferentiated ES cells and also on differentiating embryoid bodies when staining will reflect evidence of lineage-restricted expression.

## 2.2 MOLECULAR BIOLOGY

### 2.2.1 Extraction of genomic DNA from cells

Cells were trypsinised, washed in ice cold PBS, then digested overnight in digestion buffer containing 100mM NaCl, 10mM TrisCl (pH8.0), 25mM EDTA (pH8.0), 0.5% (w/v) SDS and 0.1mg/ml proteinase K. Genomic DNA was then extracted from samples, using an equal volume of phenol, chloroform and isoamyl alcohol, and centrifuged at 1700 x g for 10 minutes. The upper aqueous layer was transferred to a fresh tube and mixed with ½ volume of 7.5M ammonium acetate (or 3M sodium acetate) and 2 volumes of 100% ethanol resulting in precipitation of the DNA. The mixture was then centrifuged at 1700 x g for 5 minutes and the fluid carefully decanted. Finally, the DNA precipitate was washed once with 70% ethanol and briefly air-dried before re-suspension in an appropriate volume of TE buffer or distilled water. The concentration of the DNA was calculated using a spectrophotometer to measure the optical density of an appropriately diluted aliquot at 260nm and multiplying the result by 50 to give a reading in mg/ml. The purity of the DNA solution was indicated by the ratio of OD readings at 260nm and 280nm with pure preparations of DNA having an  $OD_{260}/OD_{280}$  ratio of 1.8.

### 2.2.2 Extraction of total RNA from cells

RNA was extracted using the Total RNA Isolation Reagent™ (TRIzol™, purchased from *Invitrogen Inc.*) as per the manufacturer's instructions. Cells were lysed directly in the tissue culture dish by adding 1ml of TRIzol Reagent™ per 10cm<sup>2</sup> dish, and passing the cell lysate several times through a pipette. The homogenised mixture was then incubated for 5 minutes at room temperature to permit complete dissociation of nucleoprotein complexes. 0.2ml of chloroform was added per 1ml of TRIzol™ used and the mixture briefly vortexed and incubated for a further 2 minutes. The sample was then centrifuged at 12,000 x g for 15 minutes at 2 to 8°C

to achieve phase separation with a lower red, phenol-chloroform phase, an interphase layer, and a colourless upper aqueous phase. The RNA remains exclusively in the latter, which comprises about 60% of the total volume. This was transferred to a fresh tube and the RNA precipitated by mixing with isopropyl alcohol (0.5ml per 1ml TRIzol™) and incubating for 10 minutes at room temperature followed by centrifugation at 12,000 x g for 10 minutes at 2 to 8°C. Finally, the supernatant was removed and the RNA pellet washed with 75% ethanol, vortexed and re-centrifuged at 7,500 x g. The RNA was air-dried and dissolved in a suitable volume of RNase-free water (pre-treated with diethylpyrocarbonate, DEPC) and the concentration calculated by multiplying the optical density at 260nm by 40 (purity assessed by an OD<sub>260</sub>/OD<sub>280</sub> ratio of 1.6).

### 2.2.3 PCR analysis for contained genomic sequence

PCR analysis for contained genomic sequences was performed using *Taq* DNA polymerase and a PCR kit purchased from *Invitrogen Inc.* The manufacturer's recommended protocol was applied and optimised as appropriate. The following components were added to a sterile 0.5ml micro-centrifuge tube to a total volume of 50µl (or else a master mix prepared for multiple reactions):

<u>Component</u>	<u>Volume</u>	<u>Final Concentration</u>
10x PCR buffer	5µl	1x
10mM dNTP mixture	1µl	0.2mM each
50mM MgCl <sub>2</sub>	1.5µl	1.5mM
Primer mix (10µM each)	1µl	0.2µM each
Template DNA (1µg/µl)	1µl	20ng/µl
<i>Taq</i> DNA polymerase	0.5µl	2.5 units
Autoclaved distilled water	40µl	
<hr/>		
Total volume	50µl	

The mixture was incubated in a thermal cycler at 94°C for 2 minutes to completely denature the template. Where a “hot start” was performed, the *Taq* enzyme was added at this stage else a heat-activated enzyme (such as Platinum *Taq*<sup>™</sup>, *Invitrogen Inc.*) was used and added at the outset. PCR amplification was then achieved by performing 25-35 cycles, using a heated-lid (or oil emersion) to prevent evaporation, as below:

Denature	94°C for 30-60 seconds
Anneal	55-62°C for 30-60 seconds
Extend	72°C for 1 minute per kilobase of DNA

#### 2.2.4 RT-PCR analysis for transcribed sequence

Successful RT-PCR analysis for transcribed RNA sequences is dependent upon high purity RNA. The method involves an initial step of reverse transcription of the messenger RNA followed by amplification of the resultant cDNA using conventional PCR methods. The first strand cDNA synthesis reaction was carried out using a master mix comprising the following components:

<u>Component</u>	<u>Volume</u>	<u>Final Concentration</u>
RNA Guard <sup>™</sup> ( <i>Pharmacia Inc.</i> )	0.5µl	20 units
5x first strand buffer	4µl	1x
DTT (dithiothreitol at 0.1mM)	2µl	10mM
dNTPs at 25mM	0.8µl	1mM
Reverse transcriptase (MoMuLV)	1µl	200 units
Oligo dT (at 1µg/µl)	0.2µl	10ng/µl
Autoclaved distilled water	1.5µl	
<hr/>		
Total volume	10µl	

The RNA, extracted as described above, was diluted to a concentration of 1µg in 10µl H<sub>2</sub>O. Denaturation was achieved by incubation at 65°C for 5 minutes, then quenching on ice and spinning down briefly. 10µl of the reverse transcription master mix was then added, making a final volume of 20µl per tube, and incubated at 37°C for 1 hour. For each RT-PCR reaction, a control sample was prepared in parallel using the identical mix as above but lacking the reverse transcriptase enzyme (thereby distinguishing amplified cDNA sequences from those arising from contaminating genomic DNA). Finally, the samples were spun down and stored at -20°C or used directly in the second amplification step.

The amplification step, which includes the second strand cDNA synthesis, was performed as per the PCR protocol described above. Specifically, 1µl of the first strand reaction product was used in place of the template DNA and the mixture was denatured for 5 minutes at 95°C followed by 30 amplification cycles: denaturation at 94°C for 30seconds; annealing at 62°C for 60 seconds; and extension at 72°C for 1 minute.

Finally, the amplification product was analysed by agarose gel electrophoresis using a 1.5% gel and visualised by ethidium bromide staining alongside an appropriate molecular weight ladder.

### **2.2.5 Identification of gene trapped sequences by the method of 5' Rapid**

#### **Amplification of cDNA Ends (5'RACE)**

RNA was extracted from gene trap clones using the Total RNA Isolation Reagent™ (TRIzol™, purchased from *Invitrogen Inc.*) as per the manufacturer's instructions and described here in Section M.1.12. The identification of expressed gene trap sequences was achieved using the method of 5' rapid amplification of cDNA ends, or 5'RACE, using a kit purchased from *Invitrogen Inc.*, The 5'RACE System™, and applied as per the manufacturer's instructions. 5'RACE, or "anchored PCR", is a technique that facilitates the isolation and characterisation of

5'ends from low-copy messages (Ref: Frohman, 1994). The basic methodology is outlined in principle below and schematically overleaf (Figures M.2.3 and M.2.4):

- 1) First strand cDNA synthesis is primed from total RNA using a LacZ-specific antisense oligonucleotide, termed primer 1, using SuperScript™ II, an RNase H- derivative of the MuMoLV reverse transcriptase. This permits cDNA conversion of the gene trapped mRNA and maximises the potential for complete extension to the 5'-end of the message.
- 2) The first strand product, the original mRNA template is by treatment with the RNase Mix™ (mixture of RNaseH, which is specific for RNA:DNA heteroduplex molecules, and RNase T1).
- 3) The cDNA is purified by removal of unincorporated dNTPs, primer, and proteins using a SNAP column (GlassMax Spin Cartridge™).
- 4) A homopolymeric tail is then added to the 3'-end of the cDNA using TdT (terminal deoxynucleotidyl transferase) and dCTP to generate an oligo dC anchor sequence.
- 5) The dC-tailed cDNA is then amplified by PCR using a deoxyinosine-containing anchor primer (Abridged Universal Amplification Primer, AUAP™) and a nested LacZ-specific antisense oligonucleotide, termed primer 2, which anneals to a site upstream of primer 1.
- 6) An additional PCR amplification step is performed on a dilution of the product from the previous step, this time using a second nested LacZ-specific antisense oligonucleotide, termed primer 3.

### 2.2.6 Sequence analysis of 5'RACE products

Following amplification of upstream gene trapped sequences, as outlined above, 5'RACE products were either sequenced directly using the  $\beta$ -galR3 (or LacR3) primer or alternatively cloned into a TA cloning vector (pGEM-T Kit™, *Promega Inc.*) and sequenced bi-directionally using forward T7 and reverse M13 sequencing primers. The DNA was labelled as per the dideoxy method (Ref: Sanger, 1977) using the Big Dye Terminator Kit™ (*PE Systems Inc.*) and sequenced on an ABI Prism™ machine (*Applied Biosystems Inc.*).

T7 Forward: TAATACGACTCACTATAGGG

M13 Reverse: CAGGAAACAGCTATGAC

M13 Forward: GTAAAACGACGGCCAG

### 2.2.7 Identification of gene trapped loci by BLAST database analysis

DNA sequence data was manipulated using the DNASTar™ computer software package. DNA sequences were analysed using the nucleotide blast search facility available on-line from the National Center for Biotechnology Information, NCBI, at the following internet website: <http://www.ncbi.nlm.nih.gov/BLAST>. The success or otherwise of a gene trap integration event could then be determined by demonstrating correct fusion of the adenoviral splice acceptor site with the intronic sequence of a known, or unknown, gene.

### 2.2.8 Measurement of cytosine methylation using nearest neighbour analysis

Analysis of cytosine methylation at CpG dinucleotides within genomic DNA was kindly performed by B.H. Ramsahoye PhD (Ref: Ramsahoye 2002). Briefly, samples of genomic DNA were extracted, as described previously (Section M.1.10), from inducible ES cell clones

expressing either Dnmt3a or Mbd2b, and from an inducible CHO cell clone expressing Dnmt3a. For each inducible clone, cells were cultured in duplicate wells either with or without doxycycline (yielding with repression or induction of the corresponding gene respectively). 10µg of DNA was digested with 30 units of the restriction enzyme *FokI* (purchased from *New England Biolabs Inc.*), which cuts at the sequence GGATGN<sub>9,13</sub>, for 2 h at 37°C in a 150-µl reaction volume. The enzyme was heat-inactivated (60°C for 20 min) and the restriction fragments were precipitated in ethanol and resuspended in 20 µl of TE buffer (10 mM Tris/1 mM EDTA, pH 8). An aliquot of 2µg (4µl) of DNA was labelled in a 12-µl reaction volume with 20 µCi of <sup>32</sup>P dNTP (10 µCi/µl and 3000 Ci/mmol; 1 Ci = 37 GBq) using 2.5 units of Klenow (*Amersham Pharmacia Inc.*) and the manufacturer's fill-in buffer. After labelling for 15 min at 15°C, the reaction was stopped by addition of 2 µl of 0.2 M EDTA and unincorporated nucleotides were removed using Sephadex G50 spin columns (*Boehringer Mannheim Inc.*). The DNA was digested to deoxyribonucleotide 3'-monophosphates and separated by two-dimensional TLC using isobutyric acid-based solvent (66 parts isobutyric acid [Sigma] to 18 parts water to 3 parts 30% ammonia solution) in both dimensions of a 20 × 20 cm glass-backed cellulose TLC plate (*Merck Inc.*). Finally, the radioactive TLC plates were exposed onto photographic film allowing the relative proportions of methylated to unmethylated cytosine nucleotides to be quantified using Scion™ imaging software and the percentage CpG methylation calculated.

### 2.2.9 Transformation of competent *E. coli*

Competent bacteria were either purchased commercially (*Invitrogen Inc.*, Subcloning Efficiency DH5α Chemically Competent *E. coli*™) or generated as per established protocols using the calcium chloride method. Briefly, this involved the inoculation of a single colony of *E. coli* into 50ml of LB medium, overnight growth at 37°C, then transfer of 4mls culture into 400ml of LB medium and further growth until the optical density reached 0.375 at 590nm (representing early to mid-log phase growth). The culture was then divided into 8 pre-chilled 50ml tubes, left on ice for

5-10minutes, then centrifuged and re-suspended in 10mls ice-cold  $\text{CaCl}_2$  twice, and finally re-suspended in 2mls of ice-cold  $\text{CaCl}_2$  and aliquoted into 250 $\mu\text{l}$  tubes and frozen at minus 70°C.

Transformation of competent cells was performed for large-scale generations of plasmid DNA. Competent cells were thawed rapidly on ice, then inoculated with approximately 1ng of DNA, mixed gently, and incubated on ice for 30 minutes. The cells were then subjected to heat-shock for 20-45 seconds at 37-42°C as appropriate, placed on ice for 2 minutes, and finally diluted in 0.95ml of room temperature SOC or LB medium and shaken at 225rpm for 1 hour at 37°C. Various aliquots of a 1 in 10 dilution of this medium were then spread on LB plates containing 100 $\mu\text{g}/\text{ml}$  Ampicillin and incubated overnight at 37°C alongside negative (no plasmid DNA) and positive (pUC19 plasmid) control plates. Individual colonies were picked for further expansion.

#### **2.2.10 Large-scale harvest of plasmid DNA**

Bacterial cultures were performed under continuous antibiotic selection, initially in ~5mls LB medium for 6-8 hours and subsequently overnight in a volume of 400-600mls LB medium. This was then centrifuged to yield a pellet of bacteria from which the plasmid DNA was to be harvested. The method of plasmid extraction was based on a modification of the alkaline lysis procedure, followed by purification by binding in an anion-exchange resin column. The Qiagen Maxi Plasmid Kit™ (*Qiagen Inc.*) was used for this purpose according the manufacturer's recommended instructions. In general, this approach yielded relatively large amounts of pure plasmid DNA at a concentration of approximately 1 $\mu\text{g}/\mu\text{l}$  for use in subsequent experiments.

## 2.3 RNAi METHODOLOGY

### 2.3.1 Design and synthesis of short interfering RNA (siRNA) oligonucleotides

The siRNA oligonucleotides were designed according to published on-line recommendations available at <http://www.rockefeller.edu/labheads/tuschl/sirna.html> (Ref: Tuschl *et al.*, 2002, “The siRNA User Guide”). Subsequent experimental results lent additional insights into optimal design outlined in the corresponding results chapter. In brief, the targeted region is selected from a given cDNA sequence beginning 50 to 100 nucleotides downstream of the start codon. Sequences of the form AA(N19)TT are identified, or alternatively AA(N21) where the former is not present. The G/C content appears to be an important determinant of siRNA function with 50% being optimal. The sense RNA strand corresponds to the N19 sequence and the antisense to its complementary sequence. In either case, the 3' end carries a TT di-nucleotide pair. This creates a symmetric overhang that helps ensure that equal ratios of siRNPs are generated without affecting recognition of the target mRNA which is directed solely by the antisense strand. Finally, a blast search is performed using the NCBI database (Ref: <http://www.ncbi.nlm.nih.gov/BLAST/>) to ensure that the target region is uniquely represented within the genome of the cell of interest.

In addition to the above rules guiding optimal choice of siRNA sequence, a number of further insights were gained in the course of experimentation. These findings have since been borne out by other investigators and are now generally recommended. The most important finding was that there is a preference for at least one or more guanine-cytosine pairings at the proximal 5' end of the siRNA duplex and an adenine-thymine pairing at the distal 3' end of the siRNA duplex. It is plausible that this finding reflects the need for strong sense-antisense base-pairing at the 5' end, and conversely weaker base-pairing at the 3' end, of the siRNA duplex with implications regarding the molecular mechanism of interference. The other finding was for a general avoidance of palindromic sequences and G/C strings within the siRNA molecule.

siRNA oligonucleotides were chemically synthesised by a commercial custom siRNA service (*Xeragon inc.*) using a proprietary method involving protection of the 2'-OH group using 2'-O-Tri-isopropylsilyl-Oxy-Methyl to protect against RNA degradation. This so-called TOM-chemistry™ enables a high coupling efficiency of over 99.5% facilitating the automated synthesis of RNA oligonucleotides up to 100mers in length at high yield and purity.

### 2.3.2 Synthesis of siRNAs using an *in vitro* transcription method

The *in vitro* synthesis of siRNAs was achieved using the Silencer™ siRNA Construction Kit (*Ambion Inc.*) according to the manufacturer's recommendations (performed by Anestis Tsakiridis BSc, masters degree candidate, Edinburgh University, 2002). This is described in outline only here. The Silencer™ Kit employs an *in vitro* transcription method that uses T7 RNA polymerase to generate individual strands of the siRNA. Templates for the reactions are produced from two DNA oligonucleotides encoding the desired siRNA strands. These oligonucleotides are designed to include an 8 base sequence complementary to the 5' end of a T7 promoter primer. These are annealed to one another and a fill-in reaction, using Klenow fragment, generates a double-stranded template ready for use in the *in vitro* transcription reaction. After transcription, the reactions are combined to permit annealing of the two siRNA strands. These are then treated with DNase, to remove the template, and RNase, to polish the ends of the duplex, and finally, purified using a glass fibre column. The eluted siRNA duplexes are similar to those produced by chemical synthesis but have uridine dimers at both 3' ends rather than thymidine dimers.

### 2.3.3 Method of transfection of siRNAs into cells in culture

siRNA transfection experiments were typically performed in 24 well plate format. Cells were plated out the day before the experiment so as to have a confluence of around 50% at the time of transfection (around  $10^5$  cells per well depending on cell type). Fresh medium was added shortly before the experiment. The transfection reagent mix was made up comprising a 10 to 1 dilution of Oligofectamine™ in Opti-MEM™ I Reduced Serum Medium (*Invitrogen Inc.*). This was added to the siRNA oligonucleotide diluted to the desired concentration in 40µl Opti-MEM. The complete siRNA/Oligofectamine/Opti-MEM cocktail was allowed to stand for 15 minutes to allow complexes to form. 150µl of serum-free medium (minus antibiotics) was added to each well and the 50µl cocktail added making a final volume of 200µl. The range of quantities of siRNA used for each experiment typically varied in 10-fold log increments from 1pg up to 1µg per well. This equates to a molar amount of approximately 0.067fmoles up to 0.067nmoles or a molar concentration of approximately 335fM up to 335nM respectively. The cells were incubated at 37°C for 4 hours after which the medium was supplemented with 20µl of serum per well. siRNA-induced target gene knockdown was assayed at variable time intervals post-transfection, usually around 48 hours. All experiments were controlled using a secondary siRNA oligonucleotide whose sequence was alien to the cell genome, often a GFP or LacZ siRNA, and all experiments where possible were performed in parallel triplicates.

### 2.3.4 Sequences of siRNA molecules used in current experiments

The sequences of the various siRNAs used in gene knockdown experiments are listed below:

<i>Firefly</i> luciferase siRNA - A	G A A A G G C C C G G C G C C A U U C dT dT
<i>Firefly</i> luciferase siRNA - B2	C G U A C G C G G A A U A C U U C G A dT dT
<i>Firefly</i> luciferase siRNA - B3	C U U A C G C U G A G U A C U U C G A dT dT
<i>Firefly</i> luciferase siRNA - C	G C U A U G A A A C G A U A U G G G C dT dT
GFP (green fluorescent protein)	G U U C A G C G U G U C C G G C G A G dT dT
LacZ ( $\beta$ -galactosidase)	U G G C G C U U U G C C U G G U U U C dT dT
Dnmt1 ( <i>de novo</i> methyltransferase)	G G C U A C C U G G C U A A A G U C A dT dT
TK (murine thymidine kinase)	G U A U G C C A A A G A C A C G C G C dT dT
HPRT (murine Hprt gene)	G U G U U U A U U C C U C A U G G A C dT dT
TERT1 (human telomerase RT)	G C A C U U C C U C U A C U C C U C A dT dT
TERT2 (human telomerase RT)	C C A A G A A G U U C A U C U C C C U dT dT
ESG-1 (embryonal stem cell gene)	G A A C U U A U C G A G G U C U U C A dT dT
EHOX (murine EhoX gene)	U G A G G G A G A G A G U G U A C U G dT dT
AET (aml1-eto translocation)	C C U C G A A A U C G U A C U G A G A dT dT

NB: Not all the experimental data arising from use of the above siRNAs is presented here.

### 2.3.5 Co-transfection of siRNAs with DNA or RNA gene reporters

The prototypic siRNA knockdown experiment used here to study the RNA interference phenomenon employed a *Firefly* luciferase (*Photinus pyralis*) target and a *Renilla* luciferase (*Renilla reniformis*, also know as sea pansy) internal control. Experiments were generally performed by co-transfection of the relevant reporter constructs (purchased commercially from *Promega Inc.*). These included the following:

*Firefly* luciferase plasmids pGL2 and pGL3

*Firefly* mRNA reporter mGL2

*Renilla* luciferase plasmid pRL

Most experiments aimed to measure siRNA-induced *Firefly* luciferase knockdown (either plasmid pGL2/pGL3 or mRNA mGL2) and used the *Renilla* luciferase (pRL) as an internal control. In all cases, the reporter construct was simply added to the siRNA/Opti-MEM™ mix prior to adding to the Oligofectamine™ reagent mix, which is able to efficiently co-transfect both siRNA and luciferase reporter. Typically, the respective amounts of the three luciferase reporters used in the experiments were 2µg pGL2, 1µg pGL3, and 0.5µg mGL2 (together with 20ng of the pRL internal control), which yielded satisfactory luminescence for purpose of assaying expression.

### 2.3.6 Quantification of *Firefly* / *Renilla* luciferase expression by luminescence measurement

Luciferase siRNA knockdown experiments were assayed using the Dual-Luciferase Reporter™ (DLR) Assay System (*Promega Inc.*) according to the manufacturer's instructions. This system provides an efficient means of performing two reporter assays in parallel, namely the activities of *Firefly* (*Photinus pyralis*) and *Renilla* (*Renilla reniformis*, also know as sea pansy) luciferases.

These are measured sequentially from a single sample obtained by passive lysis of the adherent cells *in situ* in each well. Briefly, the *Firefly* luciferase reporter is measured first by adding the Luciferase Assay™ Reagent II to generate a luminescent signal after which the reaction is quenched and the *Renilla* luciferase reaction initiated simultaneously by adding Stop and Glo™ Reagent to the same sample. The luminescent signal was measured on a Turner Designs 20/20 luminometer system over a linear range of five orders of magnitude with negligible background endogenous activity.

### **2.3.7 Flow cytometry analysis of siRNA-induced knockdown of endogenously expressed green fluorescent protein in embryonic stem cells**

siRNA-induced knockdown of GFP in ES cells was performed using a constitutively-expressing cell line generated by random integration of a CMV-driven transgene (Ref: Gilchrist G.S. *et al.*, 2003). GFP expression levels were assessed at various time-points post-transfection using fluorescent microscopy. Formal quantitative measure of the GFP expression levels was obtained by FACS analysis of trypsinised cells suspended in a 0.1% solution of sodium azide in PBS. Appropriate gating parameters were set according to negative (wild-type ES cells transfected with GFP-specific siRNA) and positive (GFP-ES cells transfected with an irrelevant siRNA) controls.

### **2.3.8 Analysis of telomerase activity using PCR/ELISA-based assay**

Telomerase activity was assayed using the Telomerase PCR ELISA Kit™ (*Roche Inc.*) according to the manufacturer's instructions (assay performed by Anestis Tsakiridis BSc, masters degree candidate, Edinburgh University, 2002). This assay is based on the addition of telomeric repeats (TTAGGG) by telomerase to the 3' end of a biotin-labelled synthetic primer, P1-TS. The elongation products are then amplified by PCR using the primers P1-TS and P2, thus generating PCR products with the telomerase-specific hexanucleotide increments. An aliquot of the PCR

product is denatured and hybridised to a telomeric repeat-specific detection probe labelled with digoxigenin. The resulting product is then immobilised via the biotin labelled primer to a streptavidin-coated microtitre plate and detected using an antibody against digoxigenin (anti-DIG-POD) that is conjugated to peroxidase. Finally, the probe is visualised by virtue of peroxidase metabolising the TMB substrate (tetramethyl benzidine) to form a coloured reaction product whose absorption is measured at 450nm.

### **2.3.9 Optimisation of siRNA cell transfection by various methods including use of HIV-derived Tat transduction peptide molecules**

A number of transfection strategies were compared (by measurement of luciferase knockdown) for their ability to deliver siRNAs into cells, including the following: calcium phosphate precipitation, electroporation, a proprietary lipid vesicle formulation, and various commercial liposomal preparations (Lipofectamine 2000™ and Oligofectamine™ [Invitrogen Inc.] and Transmessenger Reagent™ [Qiagen Inc.]). Of these strategies, Oligofectamine™ was found to give the most consistent and impressive results (data not shown).

In collaboration with Xinsheng Nan PhD (Cystic Fibrosis Gene Therapy Group, Molecular Medicine Centre, Edinburgh University) we sought to enhance the effectiveness of Oligofectamine™ by combining it with a novel delivery agent, short peptide molecules derived from the HIV gene *tat*. This work was based on preliminary unpublished data that demonstrates the efficacy of Tat peptides in the delivery of large nucleic acid molecules, specifically reporter plasmids. The Tat peptide sequence comprises nine basic amino acids (RKKRRQRRR) derived from the transduction domain of the *tat* gene and is reminiscent of similar peptide molecules derived from other genes, including the *Drosophila* homeotic transcription factor Antennapedia, and the herpes simplex protein VP22.

Here, we sought to investigate whether Tat might serve as a lipid-sparing agent able to enhance Oligofectamine™-mediated delivery of luciferase-specific siRNAs, assessed by knockdown of *Firefly* luciferase in CHO-LUC cells. In order to exaggerate the magnitude of this enhancement, we used siRNAs homologous to the region 195-213 of GL2 that gave a maximum knockdown of approximately 80% in transient transfection experiments using the plasmid pGL2. The knockdown was performed, as per previous experiments, in 24 well format using 1µl siRNA per well and a reduced dose of 0.15µl Oligofectamine™, rather than 1µl, combined with 2µg of Tat peptide. The controls were Oligofectamine™ or Tat used as single agents at the same dosage.

## 2.4 QUANTITATIVE METHODOLOGY

### 2.4.1 Statistical considerations

Wherever possible, quantitative data have been subjected to appropriate statistical analyses. Where the data comprises paired outcome measurements for two distinct experimental starting conditions, the means of the two populations have been compared using the Student's paired t-test. Here, the test statistic is assumed to exhibit a Student's t distribution and the two populations are assumed to exhibit identical variances. The Student's t-test yields a p-value that represents the probability that the observed outcome might have occurred by chance alone. By convention, an arbitrary level of statistical significance of 5% has been used to conclude whether the measurement is statistically significant.

Where applicable, graphic representations of data contain error bars for each value. Where the value is the mean of two independent measurements the error bars represent the range and where the value represents the mean of more than two independent measurements the error bars represent the standard error. If a value was derived from a unique experimental measurement, for instance for experiments performed over a range of substrate concentrations where dosage effects apply, then error bars are not appropriate. On such occasions, the dosage effect has been used to support or refute statistically significant differences between the two populations of data by using the paired Student's t-test as described above.

The following web-based statistical program was used to perform Student t-test analyses and calculate p-values: <http://www.physics.csbsju.edu/stats/t-test.html>.

## 2.4.2 Statistical analyses: Results 2

All comparative data in this thesis has been represented graphically and, where appropriate, subjected to statistical analyses using paired Student's t-test as outlined above. Where the data-points are derived from single experiments, error bars have not been applied. Where the data-points are derived from duplicate experiments, error bars shown represent the range of those experiments. Where the data-points are derived from triplicate experiments, error bars shown represent the standard error of those experiments. All graphical representations of data have been performed using Microsoft Excel™ (*Microsoft Inc.*) and error bars have been calculated according to the software of this program.

The graphical figures appearing in the three results chapters relating to RNA interference, namely Chapters 4, 5 and 6, are listed here below by way of clarifying the actual data being compared and the whether the outcome is statistically significant as determined by the calculated p-value:

- Figure 4.1 Demonstration that knockdown of Firefly luciferase in ES and 3T3 cells using sequence-specific siRNA exhibits dose-dependent potency. No formal statistical analysis performed as data-points are derived from single experiments.
- Figure 4.2 Demonstration that double-stranded siRNAs yield strong gene interference while single-stranded sense or antisense RNAs are no different to control siRNAs. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed.
- Figure 4.3 Demonstration that siRNA-induced interference in 3T3 cells is maximal when sense and antisense RNA strands are in 1:1 equimolar ratio. No formal statistical analysis performed as data-points are derived from single experiments.

- Figure 4.4 Demonstration that both sense and antisense RNA strands are necessary for gene interference. Statistically significant difference between double-strands and single sense ( $p$ -value  $< 0.001$ ) or antisense ( $p$ -value  $< 0.001$ ) strands using paired Student's  $t$ -test to compare relevant sets of data.
- Figure 4.5 Demonstration of the importance of siRNA-target gene sequence fidelity for yielding strong gene interference by comparing matched with mismatched siRNAs. The result is corroborated by the reciprocal experiment in which the target gene sequence rather than the siRNA is altered (although no formal statistical analysis is possible as data-points are derived from single experiments).
- Figure 4.7 Comparison of three different siRNAs according to levels of interference of Firefly luciferase following transient co-transfection with a luciferase plasmid in 3T3 cells. Data-points are derived from duplicate experiments and error bars represent the range. Statistical analysis is not warranted.
- Figure 4.8 Comparison of three different siRNAs as for Figure 4.7 above but for a range of different siRNA doses. Data-points are derived from single experiments and hence error bars have not been applied.
- Figure 4.9 Comparison of levels of interference of Firefly luciferase in 3T3 cells using three different siRNAs either alone or in combination with one another. Data-points are derived from duplicate experiments and error bars represent the range.
- Figure 4.10 Comparison of levels of interference of Firefly luciferase in 3T3 cells using non-homologous sense and antisense pairs of RNA strands. Data-points are derived from duplicate experiments and error bars represent the range.

- Figure 4.11 Demonstration that there is a dose-dependent knockdown of Firefly luciferase throughout a time-course starting 4 hours post-transfection and up to 7 hours. Comparison of the highest dose of test siRNA with the control confirms that this is a statistically significant result using a paired Student's t-test ( $p=0.02$ ).
- Figure 4.12 Linear-log plot of Firefly luciferase interference observed for three different levels of siRNA for a range of plasmid doses. No error bars have been applied.
- Figure 4.13 Log-log plot of Firefly luciferase interference observed for three different levels of siRNA for a range of plasmid doses. No error bars have been applied.
- Figure 4.14 Demonstration of the effect of delayed siRNA transfection on the level of interference of Firefly luciferase observed for a range of siRNA doses. Data-points are derived from single experiments and error bars have not been applied.
- Figure 4.15 Demonstration of the effect of delayed plasmid transfection on the level of interference of Firefly luciferase observed for a range of siRNA doses. Data-points are derived from single experiments and error bars have not been applied.
- Figure 4.16: Demonstration that Firefly-specific siRNA cause interference of a stably expressed luciferase gene in CHO-LUC cells as compared with control siRNA. Data-points are derived from single experiments and hence error bars have not been applied. Comparison of the two sets of data (Firefly-specific siRNA versus control siRNA over the dose range 0.01 to 1000ng/well) using a paired Student's t-test reveals that this difference is statistically significant ( $p\text{-value} < 0.001$ ).
- Figure 4.17 Demonstration of the persistence of detectable luciferase interference observed in

CHO-LUC cells following transfection of Firefly-specific siRNAs at two different doses. Data-points are derived from single experiments and hence error bars have not been applied.

- Figure 4.18 Demonstration that transfection of high doses of RNA into CHO-LUC cells results in paradoxical enhancement of the Firefly luciferase activity that is dose-dependent. Statistically significant using paired Student's t-test (p-value < 0.001).
- Figure 4.19 Demonstration that the paradoxical enhancement of luciferase activity following transfection of siRNAs into CHO-LUC cells occurs irrespective of sequence specificity, is dose-dependent, and apparent at 24 but not 48 hours post-transfection. Data-points are derived from single experiments and hence error bars have not been applied.
- Figure 4.20 Comparison of chemically synthesised siRNAs with ones made using an *in vitro* transcription method revealing greater potency of the former in terms of level of knockdown of transient expressed luciferase in 3T3 cells. Data-points are derived from single experiments and hence error bars have not been applied.
- Figure 4.21 Demonstration that *in vitro* transcribed siRNAs, when transfected into CHO-LUC cells at high dose, cause enhancement of Firefly luciferase similar to chemically synthesised siRNAs. This is irrespective of sequence specificity. Data-points are derived from single experiments and hence error bars have not been applied.
- Figure 4.22 Analysis of the effect of adding the Tat peptide to the transfection mix on Firefly luciferase interference as a measure of transfection efficiency. The result is weakly statistically significant using a paired Student's t-test (p-value < 0.03).

### 2.4.3 Statistical analyses: Results 3

- Figure 5.1 Demonstration that homology of the siRNA antisense strand to the gene target is more important than that of the sense strand in a transient transfection model. Data-points are derived from single experiments (hence no error bars) but the deviation from the overall experimental mean is statistically significant ( $p < 0.001$ ).
- Figure 5.2 Demonstration that homology of the siRNA antisense strand to the gene target is more important than that of the sense strand in a stable transfection model. Data-points are derived from duplicate experiments to give range error bars and the deviation from the overall experimental mean is statistically significant ( $p < 0.001$ ).
- Figure 5.3 Demonstration that 5'-phosphorylation of the siRNA molecule does not alter the previous result showing that homology of the antisense strand to the gene target is more important than that of the sense strand in a stable transfection model. Data-points are derived from duplicate experiments to give range error bars and the deviation from the overall experimental mean is statistically significant ( $p < 0.001$ ).
- Figure 5.4 Demonstration that Firefly-specific siRNA cause interference of plasmid-encoded luciferase but not mRNA-encoded luciferase in a transient transfection model over a range of different doses. Data-points are derived from single experiments and hence error bars have not been applied but comparison of the two sets of data by a paired Student's t-test confirm the result is statistically significant ( $p < 0.001$ ).
- Figure 5.5 Demonstration that Firefly-specific siRNAs cause interference of plasmid-encoded luciferase but not mRNA-encoded luciferase in a transient transfection model in 3T3 cells for two out of three different siRNAs when assayed at 18

hours post-transfection. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed.

Figure 5.6 Demonstration that partial interference of mRNA-encoded luciferase by more potent Firefly-specific siRNAs might be detectable in a transient transfection model in 3T3 cells at the later time-point of 24 hours post-transfection. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed.

Figure 5.7 Demonstration that Firefly-specific siRNA cause interference of plasmid-encoded luciferase but not mRNA-encoded luciferase in a transient transfection model in CHO cells for two out of two different siRNAs when assayed at 24 hours post-transfection. Data-points are derived from duplicate experiments and error bars represent the range. A paired Student's t-test has been performed to show that the difference in observed interference of plasmid-encoded luciferase and mRNA-encoded luciferase by duplex siRNA is statistically significant ( $p < 0.001$ ).

Figure 5.8 Demonstration that Firefly-specific siRNAs cause interference of plasmid-encoded luciferase but not mRNA-encoded luciferase in a transient transfection model in *Drosophila* S2 cells and similarly for mammalian 3T3 and CHO cells. However, for *Drosophila* KC cells, Firefly-specific siRNAs cause interference of *both* plasmid-encoded luciferase and not mRNA-encoded luciferase. Data-points are derived from duplicate experiments and error bars represent the range. A paired Student's t-test has been performed to show that this difference in siRNA-induced interference of mRNA-encoded luciferase between *Drosophila* S2 and KC cells is statistically significant ( $p < 0.001$ ).

Figure 5.9 Demonstration that prior 5'-terminal phosphorylation of the Firefly-specific siRNAs does not affect the observed interference of luciferase in a transient transfection model in CHO cells. In addition, 5'-phosphorylated antisense RNA strands do not yield interference using this model assay. Data-points are derived from duplicate experiments and error bars represent the range.

Figure 5.10 Demonstration that prior 5'-terminal phosphorylation of the Firefly-specific siRNAs does not affect the observed interference of luciferase in a transient transfection model in *Drosophila* KC cells. In addition, 5'-phosphorylated antisense RNA strands do not yield interference using this model assay. Data-points are derived from duplicate experiments and error bars represent the range.

Figure 5.12 Comparison of luciferase interference observed in *Drosophila* KC and S2 cells and mammalian CHO and 3T3 cells using 5'-phosphorylated versus mock-phosphorylated siRNA-B2 in a transient transfection model. 5'-phosphorylation enhances the level of knockdown of mRNA-encoded luciferase in S2 cells. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed. However, the result is consistent with a similar finding in the parallel experiment using siRNA-C (Figure 5.12).

Figure 5.13 Comparison of luciferase interference observed in *Drosophila* KC and S2 cells and mammalian CHO and 3T3 cells using 5'-phosphorylated versus mock-phosphorylated siRNA-C in a transient transfection model. 5'-phosphorylation enhances the level of knockdown of mRNA-encoded luciferase in S2 cells. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed. However, the result is consistent with a similar finding in the parallel experiment using siRNA-B2 (Figure 5.11).

Figure 5.14      Compilation of all the data relating to 5'-phosphorylation of siRNAs and their ability to interfere with either mRNA or plasmid-encoded luciferase according to the four cell lines used and for two different siRNA molecules. Each data-point is derived from a single experiment and the error bars shown represent the overall standard error of the mean for all the experiments from that cell line.

Figure 5.15      Comparison of luciferase interference observed in *Drosophila* KC and S2 cells and mammalian CHO and 3T3 cells using 5'-phosphorylated versus mock-phosphorylated sense-mismatched heteroduplex siRNA in a transient transfection model. 5'-phosphorylation enhances the level of knockdown of mRNA-encoded luciferase, compared with plasmid-encoded luciferase, in KC and S2 cells. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed.

Figure 5.16      Comparison of luciferase interference observed in *Drosophila* KC and S2 cells and mammalian CHO and 3T3 cells using 5'-phosphorylated versus mock-phosphorylated antisense-mismatched heteroduplex siRNA in a transient transfection model. 5'-phosphorylation fails to enhance the level of knockdown of mRNA-encoded luciferase, compared with plasmid-encoded luciferase, in KC and S2 cells. Data-points are derived from single experiments and hence error bars have not been applied and statistical analysis not performed.

Figure 5.17      Compilation of all the data relating to 5'-phosphorylation of heteroduplex siRNAs (both sense and antisense mismatched) and their ability to interfere with mRNA or plasmid-encoded luciferase according to the four cell lines used. Each data-point is derived from a single experiment and the error bars shown represent the overall standard error of the mean for all the experiments from that cell line.

#### 2.4.4 Statistical analyses: Results 4

Figure 6.2 Demonstration that siRNA-induced interference in ES cells is maximal when sense and antisense RNA strands are in 1:1 equimolar ratio. No formal statistical analysis performed as data-points are derived from single experiments.

Figure 6.3 Flow-cytometric plots demonstrating the siRNA-induced knockdown of GFP expression in ES cells for three different incubation times. Comparison of the GFP versus control experiment using a paired Student's t-test confirms that this result is statistically significant ( $p=0.001$ ).

Figure 6.4 Demonstration that repeated siRNA transfections in ES cells result in diminished knockdowns of GFP expression as compared with the control. This result does not reach statistical significance using a paired Student's t-test ( $p=0.13$ ).

Figure 6.6 Demonstration of dose-dependent siRNA-induced knockdown of a stably expressed  $\beta$ -galactosidase transgene targeted to the Oct4 locus in murine ES cells. Each siRNA dose represents the result of a single experiment and hence error bars have not been applied and formal statistical analysis has not been performed.

Figure 6.7 Demonstration of the effect of transfection of hTERT-specific siRNAs into MDA435 tumour cells on levels of telomerase activity. Data-points are derived from single experiments and hence error bars have not been applied.

Figure 6.8 Comparison of the number of ES cell colonies generated resistant to 6-thioguanine following treatment with Hprt-specific and control siRNAs at two different doses. Data-points are derived from single experiments and hence error

bars have not been applied. However, cells treated with Hprt-specific siRNA fail to yield more colonies than the background rate (control siRNA-treated cells).

Figure 6.9 Qualitative photographic image demonstrating the effect of Tk-specific siRNA on CHO-LUC cells according to the colour change of the tissue culture medium and compared against control single-stranded RNAs and double-stranded siRNAs.

Figure 6.10 Qualitative photographic image demonstrating the effect of Tk and Hprt-specific siRNAs at two different doses on CHO cells according to the colour change of the tissue culture medium and compared against control double-stranded siRNA.

Figure 6.11 Demonstration of the effect of Firefly-specific siRNA on the level of luciferase activity of a transiently co-transfected Firefly plasmid in a number of different ES cell lines either deficient in DNA methyltransferase activity or rescued thereof. Error bars represent the range derived from duplicate experiments. While no formal statistical analysis has been performed, it is apparent that there is siRNA-induced interference in all these cell lines irrespective of methyltransferase status.

Figure 6.12 Demonstration of the effect of Dnmt1-specific siRNA on the level of cytosine CpG methylation of ES cells following four serial transfections. Data-points are derived from single experiments and hence error bars have not been applied. However, comparison with control cells reveals evidence of siRNA-induced knockdown resulting in a small reduction in global methylation (16.4%).

Figure 6.13 Qualitative representation of genomic CpG methylation: autoradiographs from nearest neighbour analyses showing levels of cytosine methylation following four serial transfections of ES cells with Dnmt1-specific or control siRNAs.

### 2.4.5 Units of measurement

In general, quantities of a given compound have been expressed in terms of moles. Where the compound is in solution, quantities have been expressed as molar concentrations. The exception to this rule is the concentration of siRNA molecules used in experiments of RNA interference. Here, amounts of siRNA have been expressed in terms of absolute mass rather than mole equivalents. This convention was adopted for two reasons. Firstly, to circumvent the fact that one mole of siRNA duplex has twice the molecular mass of one mole of single stranded RNA and that this situation might otherwise lead to confusion. And secondly, the practice of using absolute mass followed the usual convention at the time of writing with many landmark papers in the field using mass and moles interchangeably (Ref: Elbashir S.M. et al., 2001; see also “*The siRNA User Guide*” at <http://www.rockefeller.edu/labheads/tuschl/sirna.html>, Tuschl. T., 2002).

For a typical 21-nucleotide siRNA duplex, the approximate molecular weight is 14820. Therefore, siRNA mass can be simply converted into mole equivalents by dividing by molecular weight and hence to molar concentration by dividing by the volume of the transfection mix (typically 200 $\mu$ l per well of a 24-well plate). These calculations are shown below:

<u>Mass</u>	<u>Moles</u>	<u>Molar concentration</u>
1 pg siRNA	0.067 fmoles	335 fM
1 ng siRNA	0.067 pmoles	335 pM
1 $\mu$ g siRNA	0.067 nmoles	335 nM

# **Chapter 3**

## **Antisense gene trapping in embryonic stem cells**

### **Results 1**

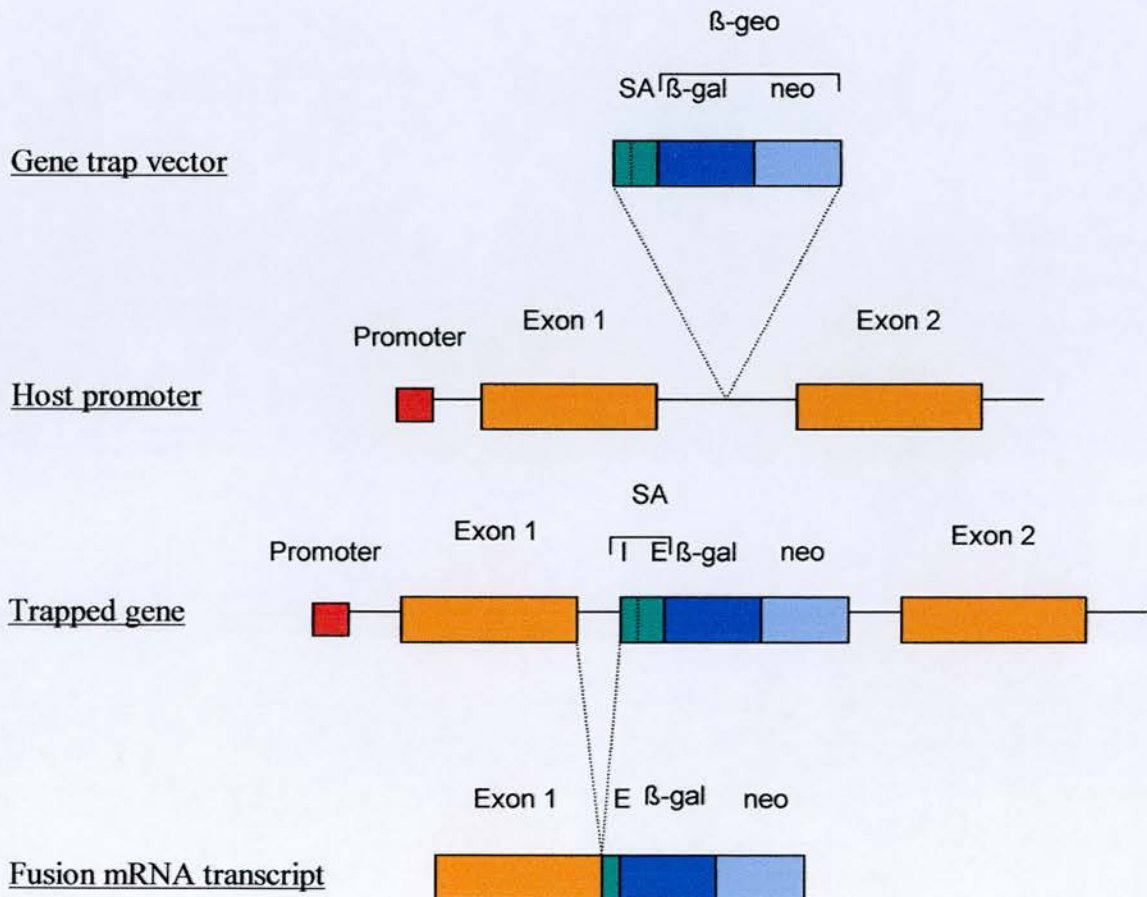
### 3.1 Introduction

Gene trap vectors, of which there are numerous designs, function by mutational insertion of a gene trapping construct at random sites within the genomic DNA of a given cell (Ref: Evans M.J. *et al.*, 1997). This construct typically includes a selectable marker gene (usually neomycin) together with a reporter gene (usually  $\beta$ -galactosidase), which both fall under the transcriptional control of the trapped upstream promoter, thereby facilitating the selection of clones in cell culture and the monitoring of gene expression within these cells (Ref: Friedrich G. and Soriano P., 1991). A splice acceptor site at the 5' end of the vector ensures the splicing of the RNA transcript to generate a gene/ $\beta$ -gal/neo fusion mRNA strand, thus achieving a functional disruption of the trapped gene and subsequent identification of the gene sequence using established PCR methods.

The gene trap approach has already demonstrated its considerable potential, particularly as applied to murine ES cells where it represents an efficient strategy for functional analysis of the mouse genome (Ref: Skarnes W.C. *et al.*, 1993; Forrester L.M. *et al.*, 1996). ES clones derived from gene trapping can be injected into early blastocysts to generate chimaeric animals, which can be bred to generate mice both heterozygous and homozygous for the mutated gene locus. In many cases, these animals might exhibit mutant phenotypes arising from functionally disrupted genes (Ref: Townley D.J. *et al.*, 1997).

However, few strategies have been described for identifying which of these gene trap clones merit further analyses. This is because such clones, carrying single allele deletions, typically do not exhibit any perturbation of cellular phenotype. The method described here represents a novel variation on the conventional gene trapping approach that aims to permit *in vitro* screening of clones of interest according to whether or not they exhibit a mutant phenotype. This is achieved by employing a gene trap vector that possesses a powerful promoter inserted in reverse orientation upstream of the splice acceptor site, thus allowing the transcription of an antisense RNA strand via a secondarily transfected vector carrying a transactivating gene. This antisense RNA strand serves to functionally disrupt expression of the counterpart allele, thereby generating a clone that is effectively a functional “double knockout” for the locus. Such a strategy permits genotype and phenotype to be directly correlated with one another and has demonstrated its potential in NIH3T3 cells where it has been employed to identify two putative tumour susceptibility genes (Ref: Li L. *et al.* 1996; Liu K. *et al.* 1999).

**Figure 3.1: Schematic illustration of classical gene trapping strategy using basic vector design.**  
 The gene trap vector “traps” the host promoter via a splice acceptor site generating a fusion mRNA transcript (SA = Splice acceptor site; I = Intron; E = Exon;  $\beta$ -gal =  $\beta$ -galactosidase; neo = neomycin).



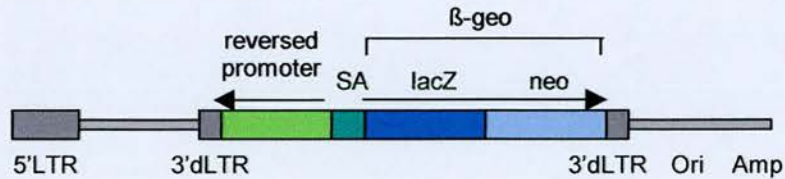
### 3.2 Overview of the vector constructs

Three vectors are relevant to this chapter, namely pGSV and pTX, both a kind gift from Stanley Cohen, Stanford University (Ref: Li L. & Cohen S., 1996), and pCRE (purchased commercially from *Invitrogen Inc.* as pBS185™). The essential design features of these vectors are outlined below (Figure 3.2).

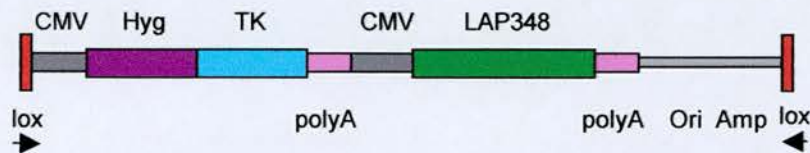
**Gene Search Vector** GSV is a pHHAM-derived retroviral vector designed to insert randomly within genomic DNA. It carries a  $\beta$ -*geo* reporter gene, which is transcribed only if insertion occurs within the intron of a gene that is being actively expressed. Splicing of the fusion transcript occurs via the upstream splice acceptor site resulting in juxtaposition of exonic and  $\beta$ -*geo* sequences and disruption of the gene. Therefore, neomycin-resistant clones (which should concomitantly produce  $\beta$ -galactosidase and so stain dark blue with X-gal) represent those cells in which there is a hemizygous deletion of the gene containing the search vector. The reversed promoter sequence acts to initiate transcription of an antisense RNA strand in the presence of transactivator LAP348, whose gene is contained in the vector TX.

**Transactivating Vector.** TX is derived from pHCMVLAP348. It contains a HyTK gene expression cassette ligated into a Hind III site upstream of the transactivator gene, LAP348; this is bracketed by *loxP* sites derived from pBS30<sup>4</sup>. LAP348 carries the operator-binding domain of the *E. coli lacI* repressor protein and the herpes simplex virus transactivation domain, VP16, and acts as a transcriptional activator in mammalian cells<sup>3</sup>. Activation of the reversed promoter in pLLGSV by LAP348 occurs in *trans* and results in transcription of a RNA strand that is antisense to the upstream portion of the trapped gene fusion transcript. Therefore, hygromycin-resistant clones represent those cells carrying successful integration of the TX vector and will include cells in which LAP348 expression has induced a functional disruption of both allelic copies of the trapped gene, rendering a functional homozygous gene knockout.

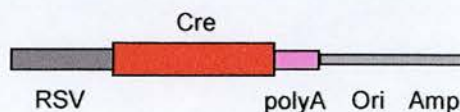
**Excision Vector.** Cre carries a gene encoding for a recombinase, Cre, which acts specifically on the *lox*-sites contained in TX. Therefore, integration of the excision vector into *hygro*-resistant cells results in deletion of the LAP348/HyTK segment located between these *lox* sites allowing selection of ganciclovir-resistant clones. Thus, the function of the trapped gene partner allele is restored resulting in return to a hemizygous state.

**Figure 3.2: Vectors used to derive functional knockouts via gene trap / antisense strategy.****A: Gene Search Vector (GSV)**

**Designations:** LTR, retroviral long terminal repeats; SA, splice acceptor site; reversed promoter, SV40 T antigen minimal early promoter linked to 14 tandemly repeated copies of *E. coli lac* operator sequence (arrow indicates direction of transcription);  $\beta$ -*geo*, reporter gene fusion of *E. coli lacZ* and *neo* genes (arrow indicates sense direction of fusion gene).

**B: Transactivating Vector (TX)**

**Designations:** CMV, cytomegalovirus promoter; HyTK, hygromycin resistance (*hyg*) and thymidine kinase (TK) fusion gene; LAP348, transactivator gene; poly(A), signals determining the site of mRNA polyadenylation; lox (arrowheads indicate the orientation of lox sites), recognition site for Cre recombinase encoded by segment depicted Cre figure.

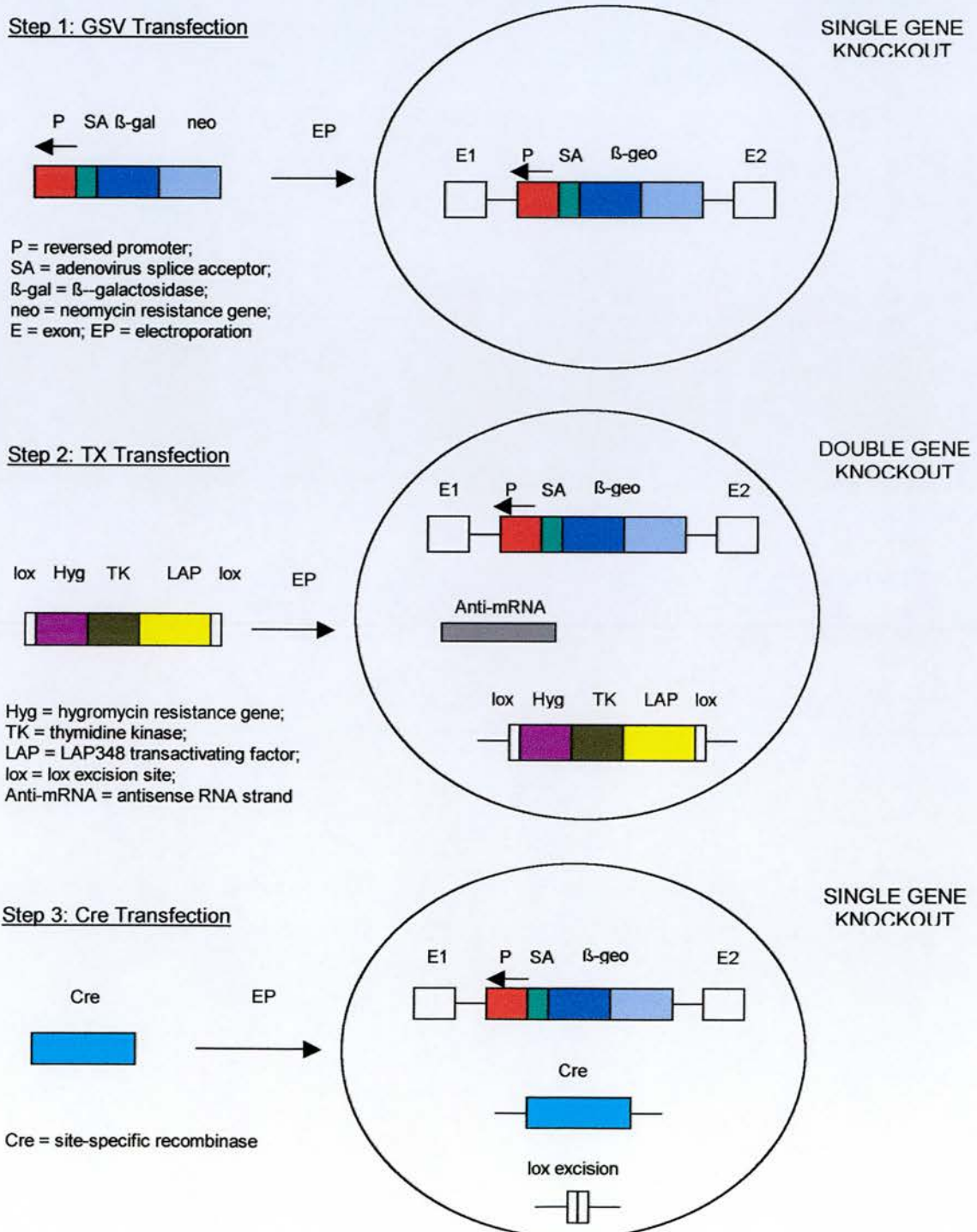
**C: Excision Vector (Cre)**

**Designations:** RSV, Rous sarcoma virus promoter; Cre, site-specific recombinase.

### 3.3 Overview of the experimental strategy

The experimental strategy comprises three steps: step one yields neomycin-resistant single-knockout gene trap clones; step two yields hygromycin-resistant functional double-knockout clones; and step three yields ganciclovir-resistant single-knockout revertant clones (Figure 3.3).

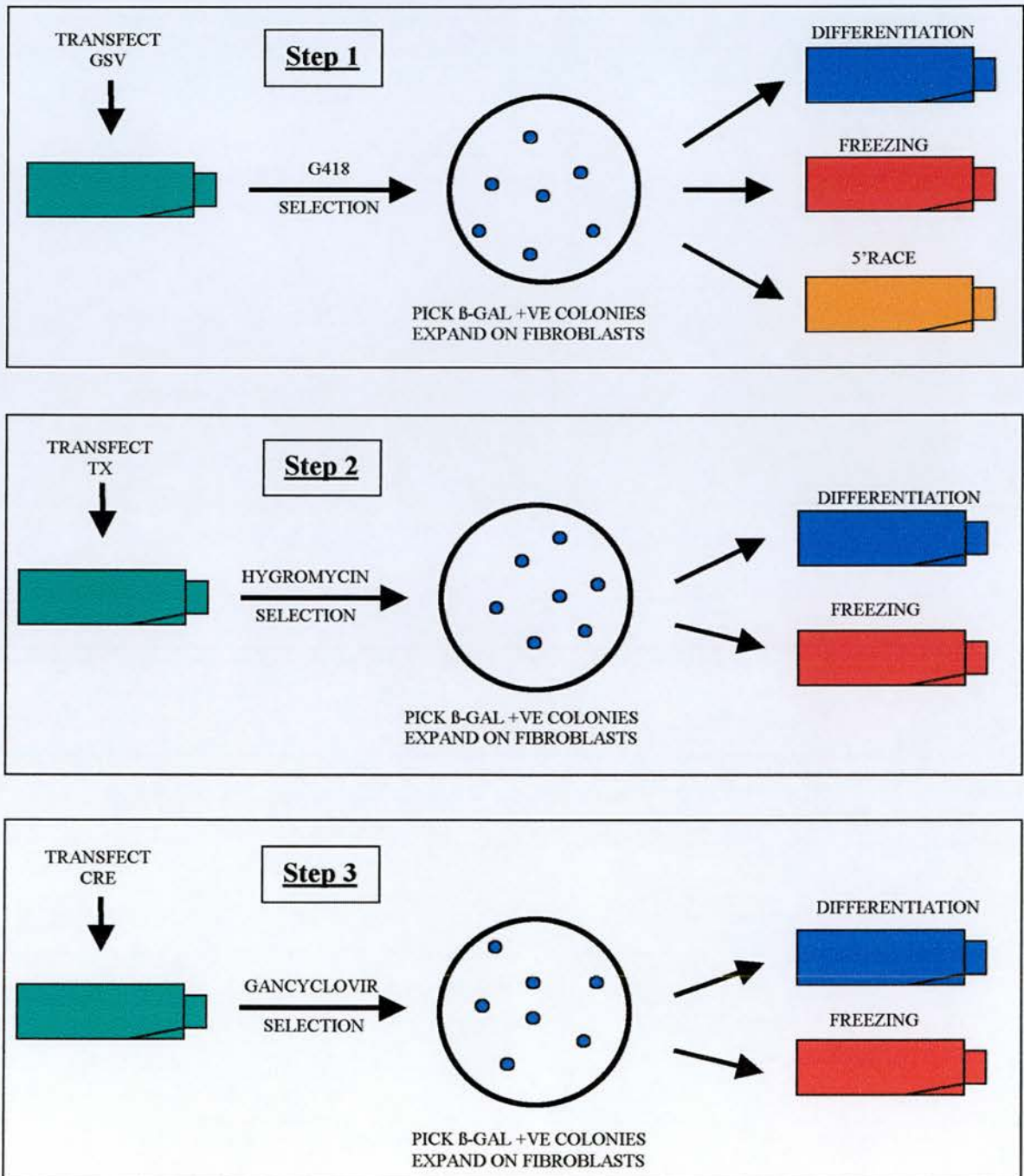
**Figure 3.3: Experimental strategy for deriving functional gene trap/antisense knockouts.**



### 3.4 Overview of the experimental design

The three steps, that together comprise the experimental strategy outlined above, were performed sequentially as illustrated below (Figure 3.4).

**Figure 3.4: Diagrammatic overview of experimental design requiring three sequential steps.**



### 3.5 Gene search vector (pGSV) traps active genes in ES cells

The gene search vector, pGSV, yielded approximately 10 neomycin-resistant colonies per  $10^7$  ES cells when electroporated with 50 $\mu$ g circularised (uncut) plasmid DNA. A total of 33 CCE-derived and 21 E14-derived gene trap clones were analysed.

The gene trap ES cell clones were assayed for *LacZ* expression and scored according to the pattern and intensity of staining which reflects the activity of the trapped promoter. Three transfection experiments were performed, the first two on CCE cells, termed C and D clones, and the third on E14 cells, termed E clones. 15 out of a total of 33 (45%) CCE clones were positive for  $\beta$ -galactosidase expression, at variable levels, while 17 out of 21 (81%) E14 clones were positive. The  $\beta$ -galactosidase positive clones are listed below together with the intensity and pattern of staining observed.

<u>Clone</u>	<u>Intensity of staining</u>	<u>Pattern of staining</u>
C1	++++	Heterogenous
C2	+++++	Uniform
C3	++++	Heterogenous
C4	0	Negative
C5	0	Negative
C6	0	Negative
C7	0	Negative
C8	0	Negative
C9	0	Negative
C10	0	Negative
C11	++	Heterogenous
C12	+++	Heterogenous
C13	+++	Uniform
C14	0	Negative
C15	0	Negative
C16	+++++	Uniform
D1	+	Scattered
D2	++++	Uniform
D3	0	Negative

<u>Clone</u>	<u>Intensity of staining</u>	<u>Pattern of staining</u>
D4	++++	Membrane
D5	+	Uniform
D6	0	Negative
D7	++++	Uniform
D8	0	Negative
D9	0	Negative
D10	+	Uniform
D11	+++	Uniform
D12	0	Negative
D13	+++	Uniform
D14	0	Negative
D15	0	Negative
D16	0	Negative
D17	0	Negative
E1	+++++	Uniform
E2	+++	Uniform
E3	+++	Uniform (discrete foci)
E4	0	Negative
E5	+++	Uniform
E6	+++	Uniform
E7	++	Uniform
E8	0	Negative
E9	+++	Uniform
E10	+	Uniform
E11	+++	Heterogenous
E12	+++	Heterogenous (membrane)
E13	+++	Uniform
E14	++++	Heterogenous
E15	++++	Heterogenous
E16	++++	Uniform
E17	++++	Uniform
E18	+	Scattered
E19	++	Heterogenous

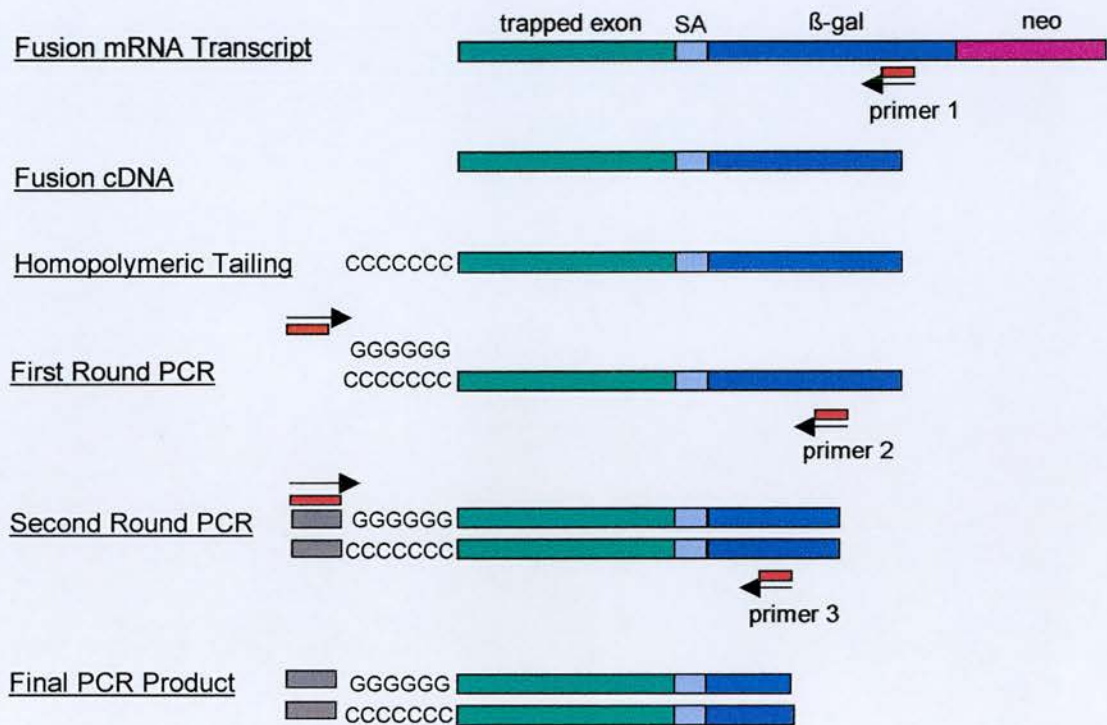
### 3.6 5' rapid amplification of cDNA ends (5'RACE) enables gene trap sequences to be identified

5'RACE was performed on total RNA extracted from all  $\beta$ -galactosidase positive clones listed above (Refs: Frohman M.A., 1994; Townley D.J. *et al.*, 1997). Out of a total of 32 clones, nine produced discrete PCR bands that could be directly sequenced without resort to cloning (27% success rate). All nine revealed gene trap integrations involving the 5' end of the  $\beta$ -galactosidase sequence and the splice acceptor region fused in frame with the upstream exon of the trapped gene. These are listed below:

<u>Clone</u>	<u>Trapped gene</u>	<u>Homology</u>
C2	Tsix (Xist antisense gene on X chromosome)	188bp
C16	X chromosome, clone RP23-436I3	123bp
D1	chromosome 4, clone RP23-282N17	180bp
D2	chromosome 2, clone RP23-216D20	105bp
D4	chromosome 19, clone RP23-117H14	310bp
E3	Catns (catenin src gene, p120)	146bp
E5	Rex2 (reduced expression gene)	145bp
E6	Nasp (somatic histone binding protein)	294bp
E12	Ctbp2 (C-terminal binding protein 2)	326bp

The 5'RACE methodology is shown schematically overleaf (Figure 3.5a) together with the sequences of the nested primers (Figure 3.5b). The cDNA sequence of the 5'RACE product derived from clone C2 is shown (Figure 3.6) for purpose of demonstrating homology alignment with the Tsix gene. Finally, the raw cDNA sequence analyses for the four gene trap clones, E3, E5, E6 and E12 is also included (Figure 3.7), for purpose of demonstrating the splice integration site of the GSV gene trap vector via the splice acceptor sequence. Note that all four integrations are shown in reversed orientation reflecting the 3' to 5' reading of the cDNA sequence from the nested  $\beta$ -galactosidase primer through the adenoviral splice acceptor and into the upstream portion of the endogenous gene. The actual site of insertion for each of these four gene trap integrations is given below the respective sequence analysis. For clones E3 and E5, the integration is within the 5' untranslated region of the gene upstream of the start codon, whereas for clones E6 and E12, the integration is juxtaposed to the gene's first coding exon downstream of the start codon.

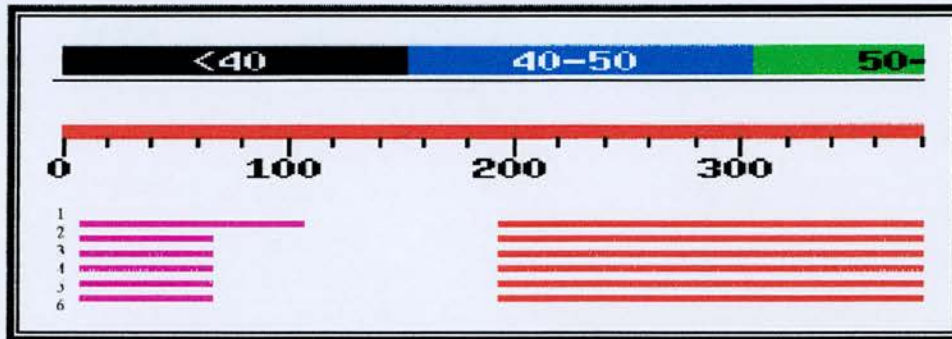
**Figure 3.5a: Schematic diagram illustrating method of rapid amplification of cDNA ends.** 5'RACE enables the amplification of gene sequences upstream of the gene trap insertion site. The method is based upon the following steps: reverse transcription of the upstream mRNA sequence, then homopolymeric tailing of the cDNA, and finally amplification through two cycles of PCR using nested downstream primers.



**Figure 3.5b: Sequence, notation and location of the two sets of primer triplets used for 5'RACE.** Two triplet sets of primers were used for amplification of gene trapped cDNA ends by the method of 5'RACE outlined in Figure 3.5a above. For each set, primer 1 was used for cDNA synthesis and primers 2 and 3 used with an anchor primer for the first round and nested round of PCR amplification respectively.

Primer 1:	$\beta$ -galR1	ATTCAGGCTGCGCAACTGTTGG	= bp 2023-2002
Primer 2:	$\beta$ -galR2	CTGCAAGGCGATTAAGTTGG	= bp 1943-1924
Primer 3:	$\beta$ -galR3	TAACGCCAGGGTTTTCCCAG	= bp 1922-1904
Primer 1:	LacR1	TAATGGGATAGGTTACG	= bp 2180-2164
Primer 2:	LacR2	AGTATCGGCCTCAGGAAGATCG	= bp 2107-2096
Primer 3:	LacR3	ATTCAGGCTGCGCAACTGTTGG	= bp 2023-2002

**Figure 3.6: Alignment of DNA sequence derived from 5'RACE analysis of clone C2 genomic DNA.** 5'RACE analysis of genomic DNA extracted from the gene trap clone C2 reveals a PCR product whose DNA sequence shows homology with the Tsix gene (testis-specific gene, on X chromosome). This gene is the antisense version of the Xist gene (X-inactive transcripts, also on X chromosome) with which the DNA sequence of the PCR product, as expected, also shares sequence homology. The 3' to 5' orientation of the PCR product, whose DNA sequence is read through in the antisense direction from the upstream neo gene, confirms that the site of the gene trap integration is within the Tsix gene and not the Xist gene.



- |    |                             |   |       |
|----|-----------------------------|---|-------|
| 1) | gi 25137577 gb AF541962.1   | Mus musculus Tsix antisense RNA ...     | 6e-96 |
| 2) | gi 13469702 gb AF138745.1   | AF138745 Mus musculus domesticus ...    | 6e-96 |
| 3) | gi 21425583 emb AJ421479.1  | Mus musculus X-inactivation cen...      | 6e-96 |
| 4) | gi 1673440 emb X99946.1     | MMTSXDNA M.musculus 94kb genomic se...  | 6e-96 |
| 5) | gi 25004494 emb AL669964.18 | Mouse DNA sequence from clone ...       | 6e-96 |
| 6) | gi 1150729 gb U41396.1      | MMMB414A3 Mus musculus X inactive sp... | 6e-96 |

#### Mus musculus Tsix antisense gene (Tsix)

```

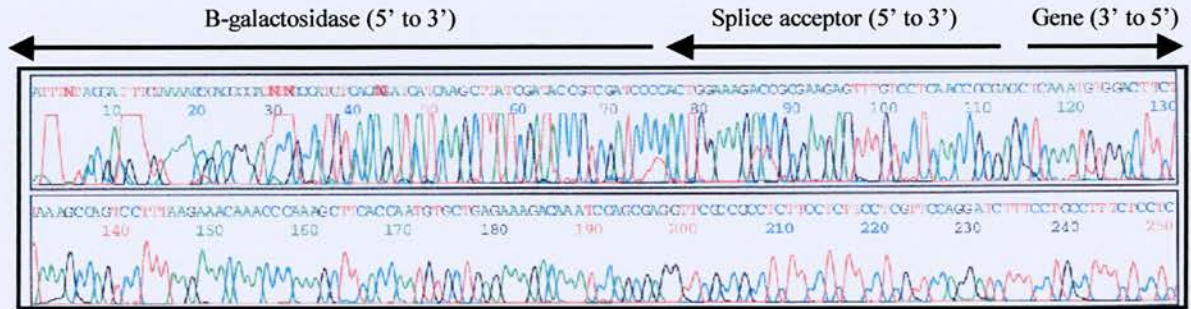
Query: 195  cgctccgctctggccacctgtctagctggctatcacgctcttctccatcacgcgcttgccgcagc 260
          |||
Sbjct: 250  cgctccgctctggccacctgtctagctggctatcacgctcttctccatcacgcgcttgccgcagc 185

Query: 261  cgctcccgggagtgatgagcaggactcaagatgctggatctcggaaagtgagaacagag 322
          |||
Sbjct: 184  cgctcccgggagtgatgagcaggactcaagatgctggatctcggaaagtgagaacagag 123

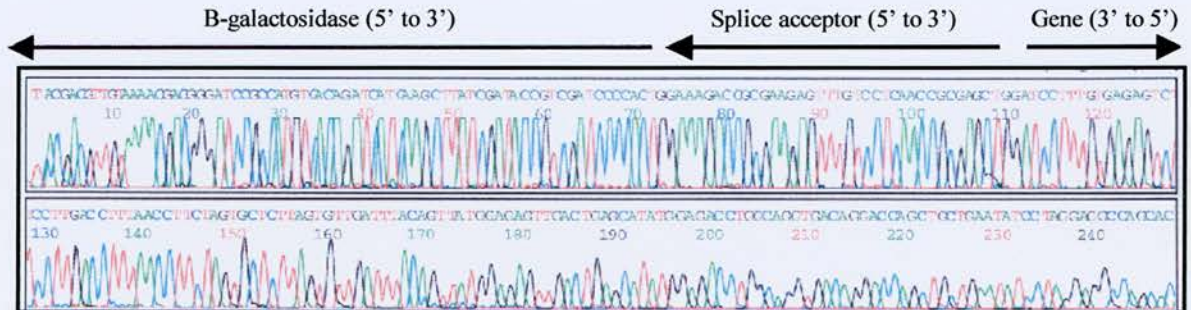
Query: 323  caagggcgccgagcgcaagtgttagcgcttcaggtacttttgggaccgagtgagcac 383
          |||
Sbjct: 122  caagggcgccgagcgcaagtgttagcgcttcaggtactt-ttgggaccgagtgagcac 63

```

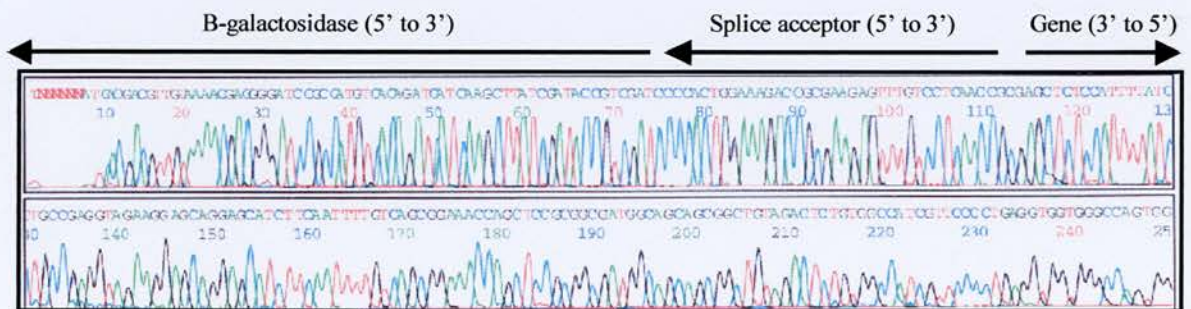
**Figure 3.7: 5'RACE cDNA sequence analysis reveals site of splice integration of the gene trap vector.** Sequence analysis of the 5'RACE cDNA products of four X-gal positive E14 clones reveals the splicing integration sites of the gene trap vector for the four clones E3, E5, E6 and E12 (respectively genes *Catns*, *Rex2*, *Nasp* and *Ctbp2*). Note that the splicing via the acceptor sequence is identical for each gene trap.



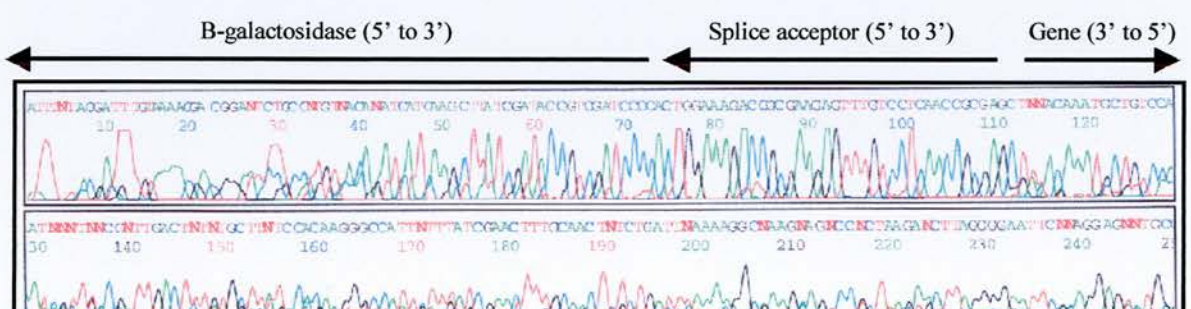
Clone E3 - *Catns* (catenin src gene, p120): Intron 1 insertion, 95bp upstream of start codon (5'UTR).



Clone E5 - *Rex2* (reduced expression gene): Intron 1 insertion, 209bp upstream of start codon (5'UTR).



Clone E6 - *Nasp* (somatic histone binding): Intron 2 insertion, 106bp downstream of start codon (exon 1).



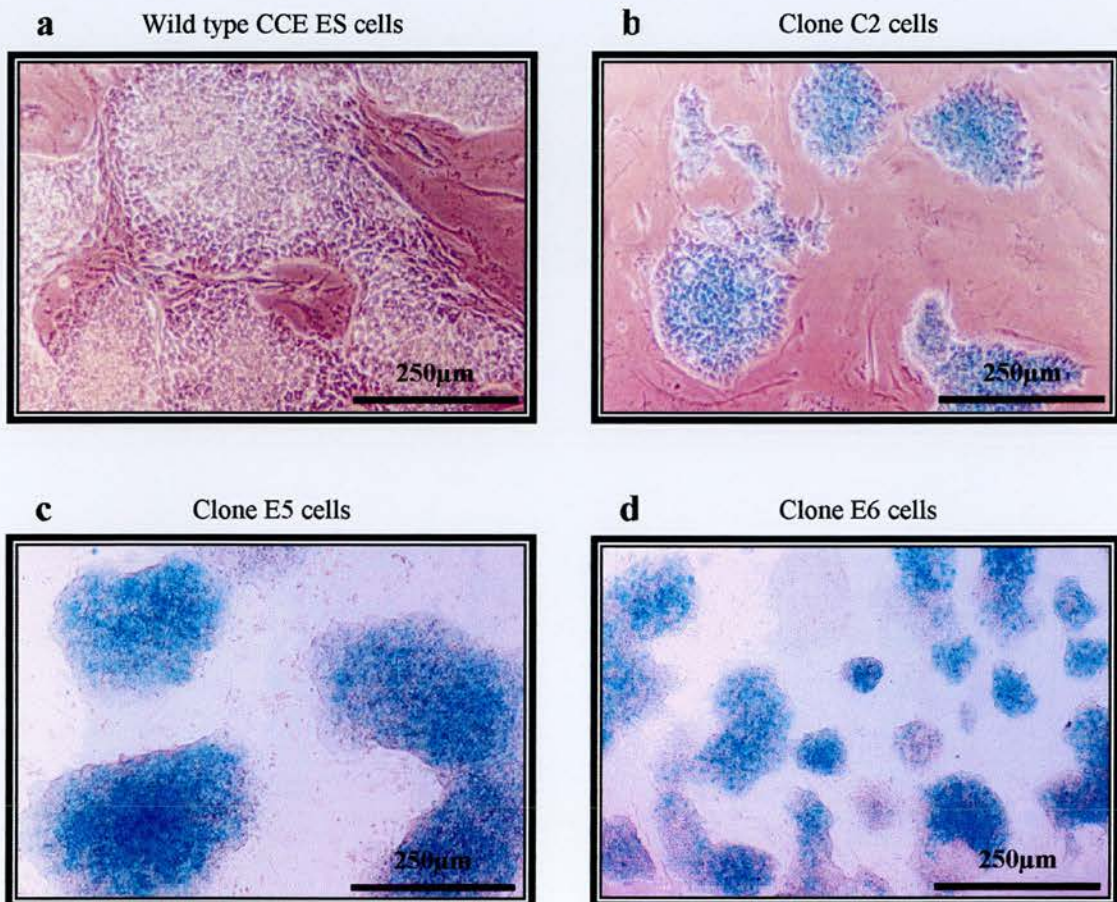
Clone E12 - *Ctbp2* (C-terminal binding prot.): Intron 2 insertion, 53bp downstream of start codon (exon 1).

### 3.7 Gene trap clones C2, E5 and E6 exhibit uniform X-Gal staining suggestive of constitutive promoter activity in undifferentiated ES cells

B-galactosidase expression was assessed by X-gal staining of undifferentiated ES cells. Of those clones listed above for which 5'RACE yielded the identity of the trapped gene, three exhibited a uniform expression pattern, namely C2, E5 and E6, shown below (Figure 3.8). This result suggests that the respective genes, Tsix, Rex2 and Nasp, are constitutively expressed in undifferentiated ES cells.

**Figure 3.8: Photomicrographs revealing uniform X-gal staining of gene trap clones C2, E5, E6.**

Photomicrographs demonstrating (a) control wild-type CCE ES cells (phase contrast), (b) clone C2 (phase contrast), (c) clone E5 (bright field), and (d) clone E6 (bright field). All three gene trap clones, C2, E5, E6 (respectively Tsix, Rex2, Nasp) exhibit uniform X-gal staining in undifferentiated ES cells.

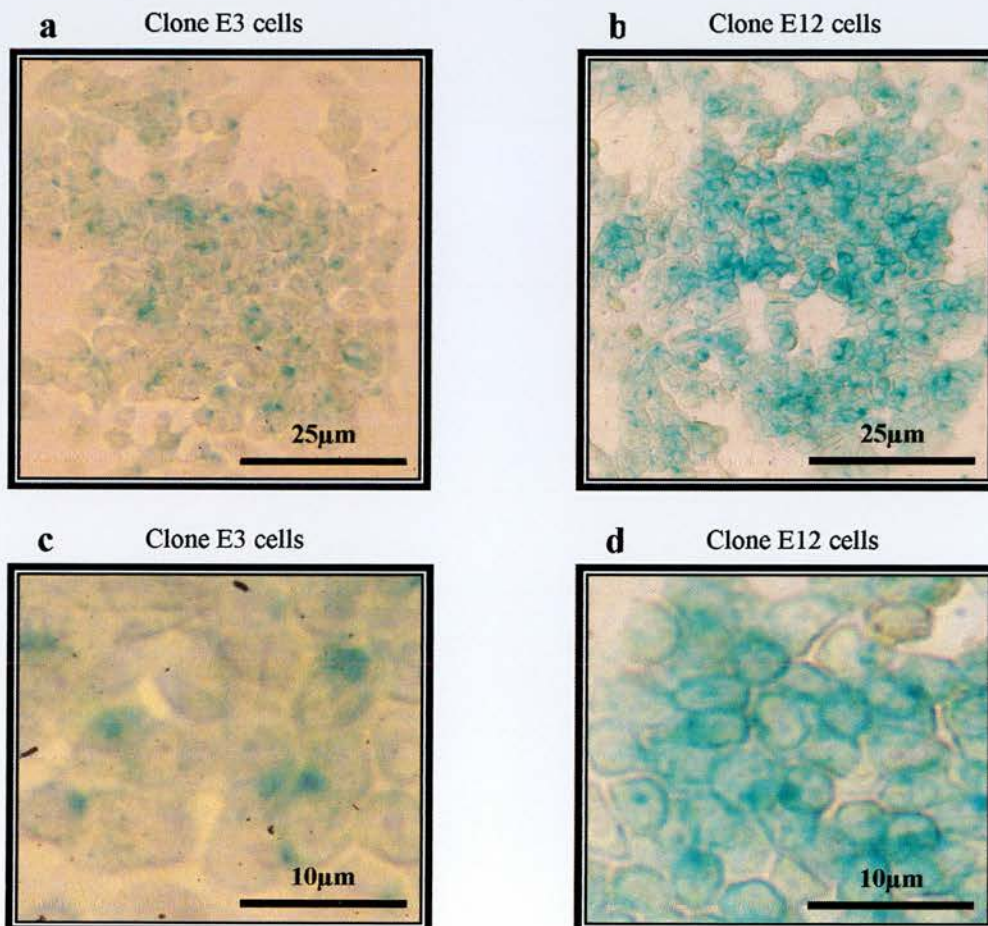


### 3.8 Clones E3 and E12 exhibit discrete X-gal staining patterns that may reflect sub-cellular localisation of the respective gene trap fusion proteins

Clones E3 and E12 both exhibit non-uniform discrete X-gal staining. For clone E3, in which the catenin-src gene is trapped, this is at discrete sites within each cell while for clone E12, in which the C-terminal binding protein 2 gene is trapped, this is localised most intensely in the perimembrane region of the cell (Figure 3.9). If the transcribed  $\beta$ -galactosidase retains the localisation signal of the trapped gene fused upstream via the splice acceptor site, then it is plausible that the staining pattern reflects the sub-cellular localisation of the trapped gene product with implications for gene function. Alternatively, this finding might represent staining artefact unrelated to the trapped gene. Both these possibilities are discussed in relation to known gene function later.

**Figure 3.9: Photomicrographs revealing discrete X-gal staining patterns for clones E3 and E12.**

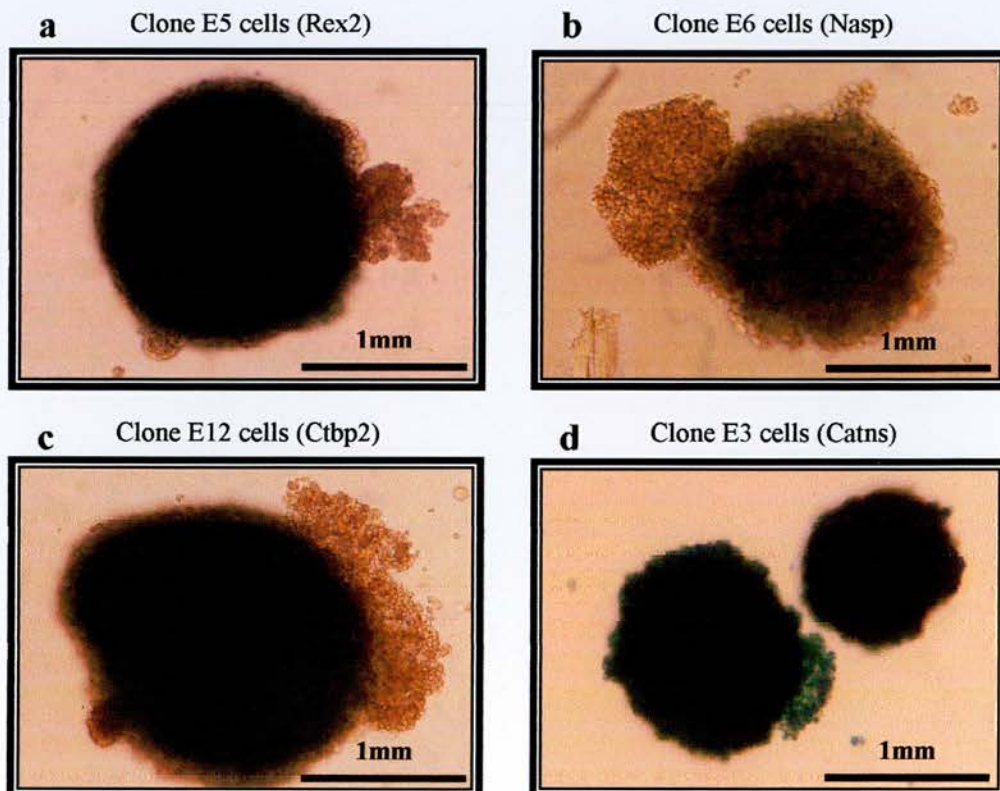
Photomicrographs revealing the  $\beta$ -galactosidase staining pattern for two gene trap clones: E3 (Cats gene) at x20 (a) and x50 (c) magnification; and E12 (Ctbp2) at x20 (b) and x50 (d) magnification. For E3 cells,  $\beta$ -galactosidase staining occurs in discrete peripheral foci, while for E12 cells,  $\beta$ -galactosidase staining is principally membranous, possibly reflecting localisation of respective fusion protein products.



### 3.9 The gene trap genes, *Rex2*, *Nasp*, and *Ctbp2*, are *not* expressed in erythroid cells, as assessed by LacZ staining, during haematopoietic differentiation

Ideally, the  $\beta$ -galactosidase staining should faithfully reflect the transcriptional activity of the trapped promoter and thereby that of the endogenous gene during ES cell differentiation. In the course of methylcellulose differentiation analysis of the three gene trap clones, E5, E6, and E12,  $\beta$ -galactosidase expression was *never* evident in haemoglobinised erythroid cells despite relatively strong staining in other haematopoietic and non-haematopoietic cells. By contrast, clone E3, in which the trapped gene is the ubiquitously-expressed *catenin-src*, exhibited  $\beta$ -galactosidase expression in *all* haematopoietic cells including those undergoing haemoglobinisation, thereby serving as a useful experimental control. Thus, these findings suggest that there is lineage-restricted expression of the three genes *Rex2*, *Nasp* and *Ctbp2* during haematopoiesis, but not of the gene *Catns* (Figure 3.10). Formal exclusion of the alternative explanation, namely that X-gal staining fails to reveal  $\beta$ -galactosidase expression in haemoglobinised cells for reasons such as changed pH, would require protein quantification by immunoblotting.

**Figure 3.10: X-gal staining of haematopoietic embryoid bodies derived from ES gene trap clones.** Haematopoietic embryoid bodies derived from methylcellulose differentiation of ES cell gene trap clones: (a) E5 (*Rex2*), (b) E6 (*Nasp*), and (c) E12 (*Ctbp2*) all fail to exhibit any X-gal staining of haemoglobinised cells as contrasted with (d) E3 (*Catns*) which does exhibit ubiquitous X-gal staining.



### 3.10 “Enforced” expression of the gene trapped gene via maintained G418 selection pressure during ES cell differentiation induced by LIF-withdrawal

This experiment was designed to explore the effect of functionally “enforced” expression of the trapped gene via maintenance of G418 antibiotic in the culture medium. The G418 prevents down-regulation of the promoter driving the neomycin resistance gene, which would otherwise result in cell death, and thereby results in concomitant constitutive expression of the homologous (normal) copy of the trapped gene.

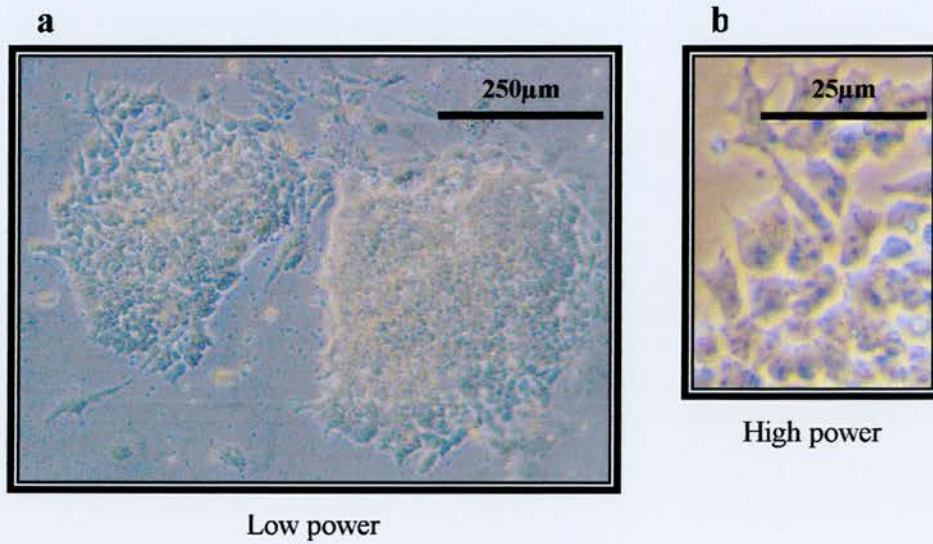
Experiments were performed on gene trap clones E3, E5, E6, E12 by plating individual cells at low density (100cells/ml) on paired tissue culture plates (1,000 cells per 10ml plate) that either contained G418 at a dose of 100µg/ml or were G418 free. The cells were cultured in medium lacking LIF and left for a period of one week to allow differentiation to occur.

For the three gene trap clones, E3, E6, E12 (respectively, trapped genes *Catns*, *Nasp*, *Ctbp2*), there were *no* qualitative morphological discrepancies in differentiation between the paired plates (G418 100µg/ml versus G418 free). However, for the gene trap clone E5, in which the gene trap insertion involved the gene *Rex2*, there was an obvious qualitative difference between the two plates. Here, the G418 100µg/ml plate contained relatively few cells that were almost exclusively undifferentiated while the G418 free plate contained ES cells undergoing normal differentiation (Figure 3.11). This observation suggests that *Rex2* expression, “enforced” by the presence of G418, is incompatible with normal ES cell differentiation and results in the persistence of an undifferentiated ES cell phenotype. This conclusion suggests that *Rex2* is phenotypically reminiscent of the unrelated gene *Rex1*. *Rex1*, which like *Rex2* is so-named because of its reduced expression in the presence of retinoic acid, is a well-recognised marker of undifferentiated ES cells that undergoes rapid down-regulation once differentiation is induced via withdrawal of LIF from the medium.

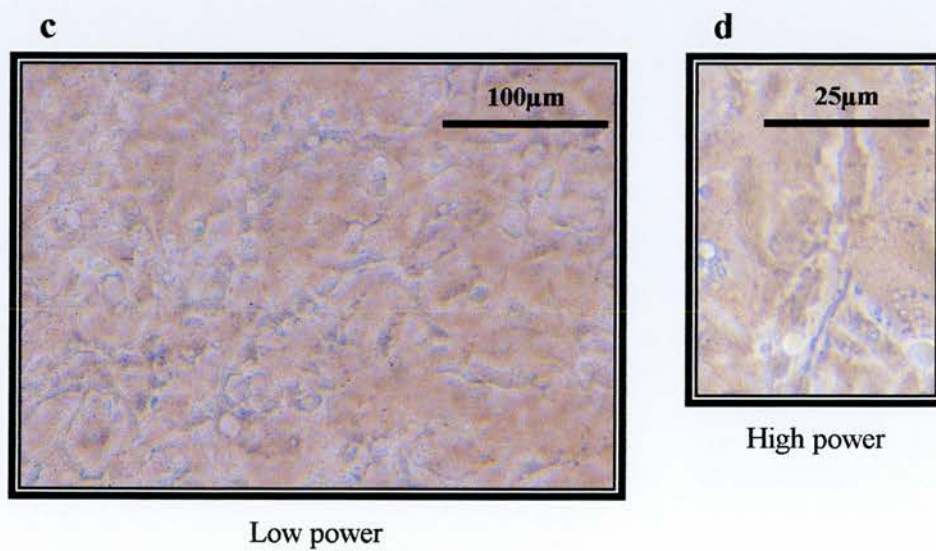
This observation raises the hypothesis that *Rex2* is an important gene regulating ES cell differentiation. However, the case is far from proven and the evidence presented here is circumstantial and not conclusive. Further experiments to test this hypothesis would include the introduction of the *Rex2* gene into ES cells under a constitutive or conditional promoter and the knockout of the gene using perhaps RNA interference, antisense approaches, or targeted disruption. The latter could be used for blastocyst injection to derive chimaeras and generate transgenic mice that could be bred to yield homozygous embryos and thereby true knockout ES cells.

**Figure 3.11: Gene trap clone E5 cells grown in the presence of G418 fail to differentiate.** ES cells of the E14 gene trap clone, E5, that carries the gene trap vector inserted into Rex2, were cultured at low density in the absence of LIF on paired G418 100 $\mu$ g/ml and G418 free plates. The photomicrographs reveal undifferentiated ES cells growing on the G418 100 $\mu$ g/ml plate at x 20 (a) and x50 (b), but differentiated ES cells growing on the G418 free plate at x 20 (c) and x50 (d).

Clone E5 ES cells grown without LIF in presence of G418



Clone E5 ES cells grown without LIF in absence of G418



### 3.11 B-galactosidase negative sub-clones are generated following introduction of the transactivator vector (pTX) into gene trap ES cell clones

Generation of gene trap sub-clones carrying the transactivator vector, pTX, was attempted initially on the four most strongly  $\beta$ -galactosidase positive CCE clones, C1, C2, C3 and C16. Each clone was grown to confluency on a T75, then trypsinised and divided equally into two cuvettes for electroporation, one with 50 $\mu$ g of pTX plasmid DNA and the other with 50 $\mu$ g of a control hygromycin vector, pPGK-hyg. Electroporated cells were plated onto duplicate plates and selected with hygromycin at a concentration of 300 $\mu$ g/ml until colonies emerged. Colonies were grown to day 12 when  $\beta$ -galactosidase expression was assessed by *in situ* staining using X-gal.

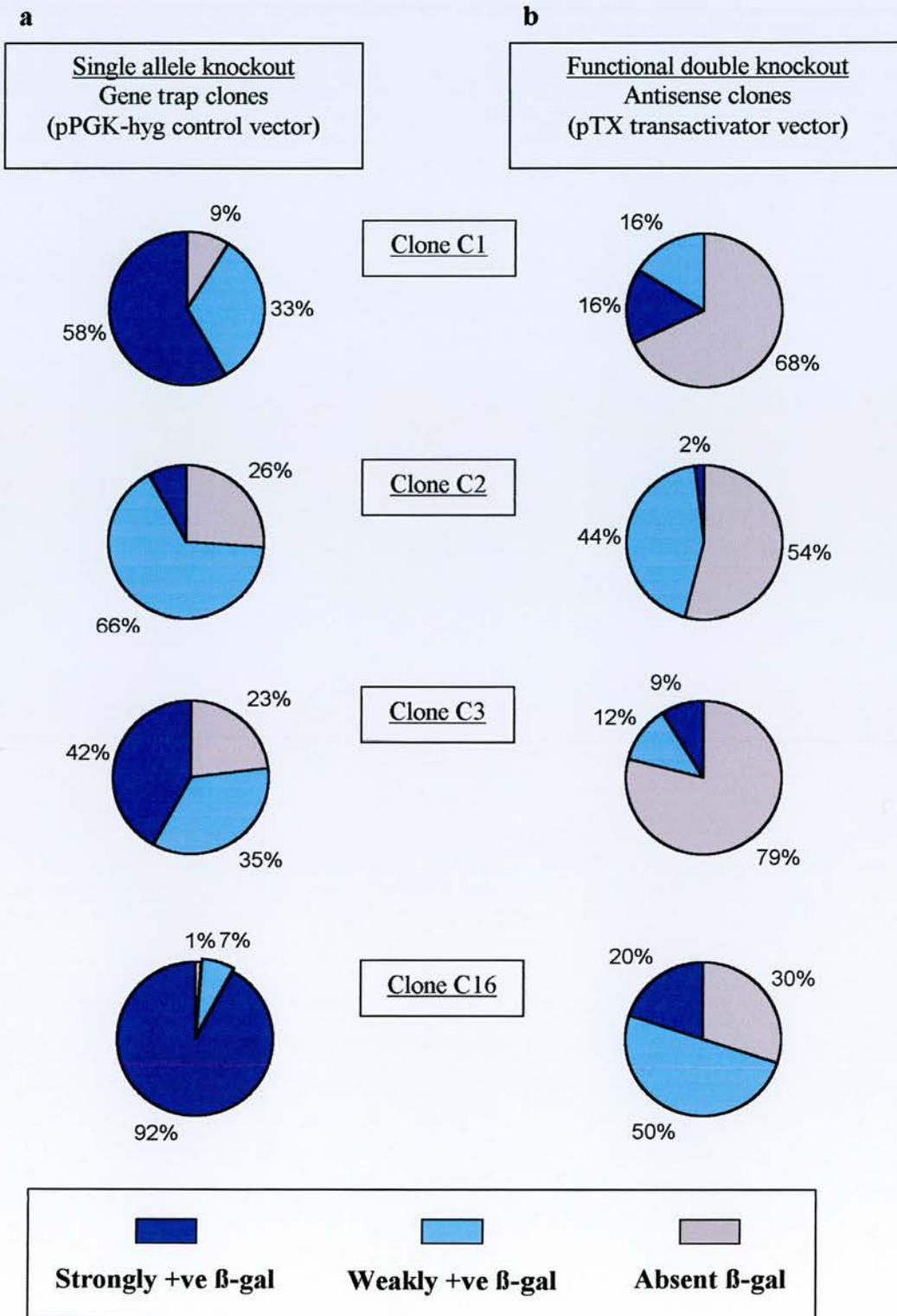
An independent arbiter, blinded to the identity of the plates, performed unbiased scoring of LacZ expression by crude subjective assessment of the intensity of colony staining into strong, weak and absent, thereby permitting comparison between the antisense and control plates. This revealed a higher proportion of LacZ negative and weakly staining colonies on the pTX plate than for the pPGK-hyg control plate (Figure 3.12). A total of 100 colonies were scored for each plate and the results analysed statistically using Student's t-test for paired data. This demonstrated that the observed reversal of  $\beta$ -galactosidase expression was statistically significant ( $p=0.003$ ).

Thus, there is evidence that antisense transcription, initiated from the reversed promoter upstream of the splice acceptor site, is able to knockdown expression of the  $\beta$ -galactosidase gene and, by inference, presumably that of the trapped gene locus likewise. However, this effect was not universal since some sub-clones remain LacZ positive implying that antisense knockdown has variable penetrance due to factors not determined. These CCE clones were not analysed further as attention was turned to E14-derived clones described below.

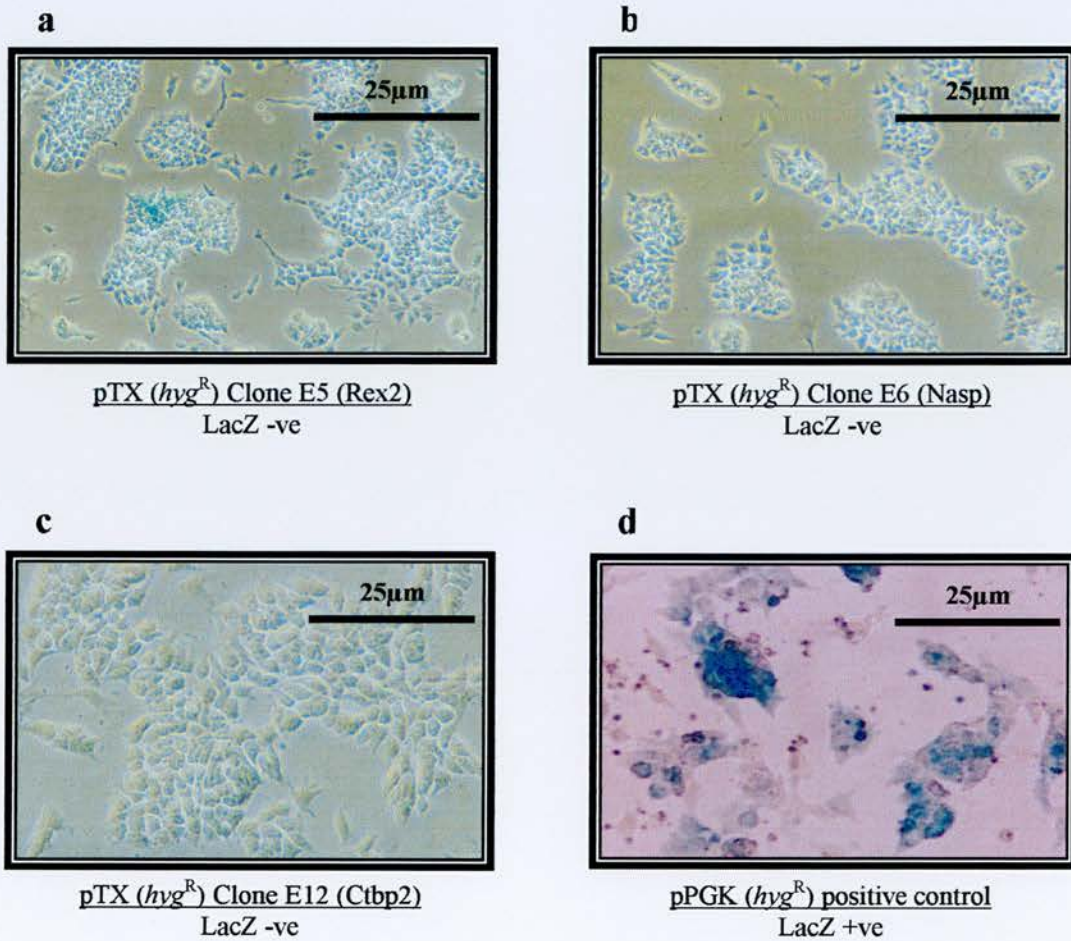
The four  $\beta$ -galactosidase-expressing E14 gene trap clones, E3, E5, E6 and E12, whose gene trap integration sites were characterised by 5'RACE, were expanded and electroporated with the pTX transactivator vector as described for the CCE clones above. On this occasion, hygromycin-resistant sub-clones were picked individually onto duplicate wells and one well was subjected to X-gal staining to assess  $\beta$ -galactosidase expression. Three of the four clones, E5 (Rex2), E6 (Nasp) and E12 (Ctbp2), yielded "antisense" sub-clones that were fully negative  $\beta$ -galactosidase expression (Figure 3.13). These were assumed to represent clones in which the trapped gene, together with the downstream  $\beta$ -geo fusion, had been functionally deleted by the antisense mechanism. They were individually expanded for further phenotypic analysis and for "antisense knockout reversal" through Cre-recombinase excision of the transactivator.

**Figure 3.12: Pie charts showing the loss of  $\beta$ -galactosidase expression in “knockout” sub-clones.**

Four gene trap CCE ES clones were used to generate hygromycin-resistant sub-clones by insertion of either a pPGK-hyg control vector (a) or the pTX transactivator vector (b). The sub-clones were scored for level of expression of  $\beta$ -galactosidase, according to the accompanying key, to permit comparison. Student's t-test for paired data was used to demonstrate that the observed reversal of  $\beta$ -galactosidase expression (pTX clones versus pPGK clones) was statistically significant ( $p=0.003$ ).



**Figure 3.13: Photomicrographs demonstrating  $\beta$ -galactosidase negative “knockout” clones.** Hygromycin-resistant sub-clones were generated from the E14 gene trap clones using the pTX transactivating vector and hygromycin selection. Putative “antisense knockout”  $\beta$ -galactosidase-negative sub-clones were identified for (a) E5 (Rex2), (b) E6 (Nasp), and (c) E12 (Ctbp2). By contrast, no  $\beta$ -galactosidase negative sub-clones were derived from the control experiments using a pPGK-hyg vector. A representative LacZ +ve example is shown for the E12 control experiment (d).



### 3.12 Comparative differentiation of gene trap clones with antisense knockout sub-clones fails to identify any differences using the methylcellulose assay

The identification of genes important for normal differentiation represents the principal goal of the current gene trap antisense knockout strategy. To this end, CCE-derived gene trap ES cell clones (carrying the pPGK control vector) were compared with their corresponding “antisense” sub-clone derivatives (carrying the pTX transactivator vector) using the methylcellulose assay to screen for differentiation mutants. In total, 15 CCE-derived clones that were positive for  $\beta$ -galactosidase expression (comprising seven C clones and eight D clones, as listed in section 3.5) were analysed. Of this total, *none* of the “antisense knockout” clones was found to exhibit reproducibly abnormal differentiation phenotypes when compared with their respective parental gene trap clones. In large part, this result reflects the fact that the methylcellulose assay is a qualitative assay that is not easily adapted for quantitative analysis. Moreover, it lacks the simplicity of classical phenotypic screens being comparatively labour-intensive to perform and therefore poorly suited to the aim of high-throughput screening of differentiation mutants. For these reasons, the methylcellulose screening strategy was abandoned and an alternative approach was sought.

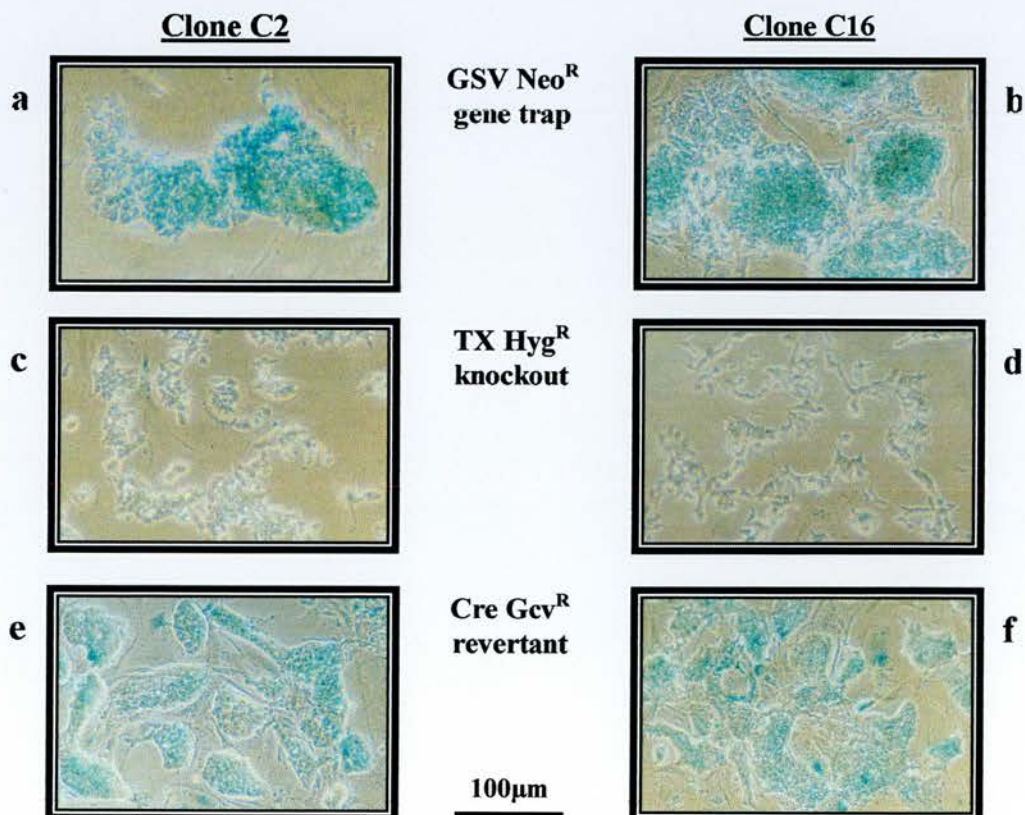
The alternative approach employed for selection of the E14-derived gene trap clones. Rather than performing methylcellulose differentiation on *all* clones as previously, this was performed instead on a restricted subset of  $\beta$ -galactosidase positive clones whose gene identity (revealed by 5'RACE and cDNA sequence analysis) was suggestive of a possible role in differentiation. Such an approach, which allowed a more detailed comparative analysis, was applied to the three clones E5 (Rex2), E6 (Nasp), and E12 (Ctbp2), each of which had yielded LacZ positive *Neo*<sup>R</sup> gene trap and LacZ negative *Hyg*<sup>R</sup> “antisense knockout” sub-clones. Methylcellulose differentiation assays were performed in triplicate for each clone pair and the plates were assessed at 10-12 days. This was done by scoring embryoid bodies qualitatively and quantitatively for recognisable features of lineage differentiation, namely, erythroid, myeloid/macrophage, mixed haematopoietic, cardiac, neuronal, and undifferentiated.

As for the CCE clones previously, this more refined differentiation analysis performed on the three E14 clones, E5, E6, and E12, again *failed* to reveal qualitative or quantitative abnormalities of differentiation when compared with their respective “antisense knockout” sub-clone derivatives (data not shown). If these clones are indeed true functional knockouts of the trapped genes Rex2, Nasp and Ctbp2, then this result suggests that none is required for normal *in vitro* differentiation. However, confirmation of successful antisense deletion by immunoblotting and further functional analyses of each of these three genes are warranted to corroborate this finding.

### 3.13 Generation of Cre recombinase sub-clones leads to the reappearance of $\beta$ -galactosidase expression putatively via excision of the transactivator vector

Reversal of antisense knockout was achieved by Cre-excision of the pTX transactivator vector in  $\beta$ -galactosidase negative clones via transfection of the pBS185™ vector. Sub-clones were selected using ganciclovir at a concentration of 2 $\mu$ M. Ganciclovir is metabolised into a toxic product by the thymidine kinase that is expressed from the HyTK cassette carried on the pTX vector in Hyg<sup>R</sup> antisense knockout clones. Cells in which the pBS185™ vector has successfully carried out Cre-excision of the pTX vector no longer express thymidine kinase and are therefore resistant to the effects of ganciclovir (like wild-type ES cells). Moreover, these hygromycin-sensitive / ganciclovir-resistant sub-clones no longer carry the transactivator gene and have therefore reverted to the original gene trap state. Note that formal demonstration of successful genomic excision of the lox-flanked pTX vector by Southern blot analysis was not performed but inferred from the reappearance of  $\beta$ -galactosidase expression indicated by positive X-gal staining. This is demonstrated for the two gene trap clones C2 and C16 and their derivative LacZ-ve/Hyg<sup>R</sup> knockout and LacZ+ve/Gcv<sup>R</sup> revertant sub-clones below (Figure 3.14)

**Figure 3.14: Cre excision of pTX transactivator vector leads to  $\beta$ -galactosidase reappearance.** X-gal staining of the CCE-derived C2 and C16 clones reveal  $\beta$ -galactosidase expression: positive Neo<sup>R</sup> gene trap (a,b), negative Hyg<sup>R</sup> antisense-knockout (c,d), and positive Gcv<sup>R</sup> Cre-excision (e,f).



### 3.14 Technical problems relating to the gene trap antisense knockout strategy for identification of differentiation genes in ES cells

The original aim of using antisense knockout of gene-trapped genes in ES cells was the identification of genes important to the differentiation process. However, two major impediments to this goal became apparent in the course of experimentation. These are listed below:

- 1) The identification of differentiation mutants amongst the “antisense knockout” clones, by comparative methylcellulose differentiation with the parental gene trap clones, proved excessively labour-intensive and inefficient as a screening strategy. Moreover, the methylcellulose assay lacked the qualitative and quantitative sensitivity necessary for adequate discrimination of subtle abnormalities of ES cell differentiation. For both these technical reasons, the methylcellulose assay was poorly suited to the purpose of reliable and reproducible high-throughput analysis of ES cell clones.
- 2) Both the gene trap clones and the “antisense knockout” sub-clones typically exhibited a broad degree of heterogeneity of  $\beta$ -galactosidase expression at the cellular level. This was true of all trapped genes not ubiquitously expressed and resulted in a mosaic population of X-gal negative and X-gal positive cells. Consequently, at the level of the individual cell, it is difficult to conclude that the absence of  $\beta$ -galactosidase expression is attributable to an antisense mechanism rather than the alternative explanation of spontaneous silencing of the trapped locus for stochastic reasons unrelated to the transactivator. Instead, knockdown of gene expression can only be inferred from an overall effect of  $\beta$ -galactosidase staining on the whole population of cells. In view of this technical difficulty, the formal proof of gene knockout using Western or Northern blot analysis was not pursued. Attempts to overcome the problem using redesigned gene trap and transactivating vectors that permitted inducible antisense knockout of the trapped locus were unsuccessful and data is not presented here.

On account of both these technical shortcomings of the overall experimental strategy, the primary goal of identifying novel genes involved in ES cell differentiation was not achievable. In terms of the secondary goal of demonstrating proof-or-principle of the antisense gene trap methodology to derive knockout ES cell clones, this was achieved only by measure of a surrogate reporter gene rather than the trapped gene itself. These conclusions are explored in greater detail in the discussion.

### 3.15 Conclusion

The primary objective of this chapter was the proof-of-principle demonstration that antisense gene trapping represents an experimentally legitimate and technically viable method for the identification of putative differentiation genes in murine ES cells. The secondary objective was the functional characterisation of these genes through analysis of the gene trap and antisense knockout clones.

#### **The gene search vector traps active genes in ES cells**

A total of 54 ES cell gene trap clones were generated, through genomic integration of the gene search vector, of which 32 demonstrated  $\beta$ -galactosidase expression. Amongst these, there was considerable variability of staining intensity, likely reflecting differences in the transcriptional activity of the trapped promoter. Similarly, there was considerable heterogeneity of staining patterns with some clones exhibiting diffuse uniform staining, likely reflecting constitutional promoter activity in all cells, and others exhibiting discrete patchy staining, likely reflecting stochastic loss of transcriptional activity in some cells.

#### **5'RACE analysis enables identification of gene trap sequences**

Identification of the trapped gene sequences, using the technique of 5'RACE, was possible in a total of nine out of 32  $\beta$ -galactosidase gene trap clones. Of these nine, five yielded known genes and four yielded unknown gene sequences. The five genes, whose functional characterisation is discussed later, are Xist, Catns, Rex2, Nasp, and Ctbp2.

#### **The transactivating vector yields $\beta$ -galactosidase negative gene trap sub-clones**

“Antisense knockout” sub-clones were generated from the four CCE-derived  $\beta$ -galactosidase positive gene trap clones, C1, C2 (Tsix), C3 and C16, via introduction of the transactivating vector. This resulted in overall reduction of  $\beta$ -galactosidase expression, by LacZ staining assay, compared with a control vector that lacked the transactivating gene. Similarly, for the four E14-derived  $\beta$ -galactosidase positive gene trap clones, E3 (Catns), E5 (Rex2), E6 (Nasp), and E12 (Ctbp2),  $\beta$ -galactosidase negative transactivator sub-clones were derived following integration of the transactivating vector. The transactivator recognises the reversed promoter sequence upstream of the splice donor site positioned at the 5' end of the gene trap construct, thereby generating an antisense RNA strand that includes upstream exon sequence. Thus, the observation of reduced  $\beta$ -galactosidase expression reflects some degree of loss of expression of the trapped gene, putatively via the mechanism of antisense inhibition. However, it is also possible for reduced expression of  $\beta$ -galactosidase to be explained in terms of promoter interference and this issue is discussed later.

**Deficiencies in the above methodology for inferring “gene knockout”**

The formal demonstration of reduced expression of the trapped gene in the transactivator sub-clones would require quantification of either the mRNA transcript or the protein level using northern or western blotting respectively. However, the high degree of clone-to-clone variability, as reflected by heterogeneity of  $\beta$ -galactosidase expression, meant that demonstration of reduced gene expression in a single clone might not be generalisable to all the other clones. Moreover, it is difficult to exclude the possibility that reduced mRNA or protein levels might simply reflect the stochastic loss of expression of the trapped gene rather than any antisense process. Hence, this direct evidence of antisense inhibition was not attained but rather the indirect evidence of reduced  $\beta$ -galactosidase expression as described above. This level of evidence sufficed in the original publication using antisense gene trapping in NIH3T3 cells (Ref: Li L. and Cohen S.N., 1996) where it was termed random homozygous knockout. Of course, reduced  $\beta$ -galactosidase expression can only reflect the events happening in *cis* at the site of the gene trap locus and not those in *trans* at the site of the homologous allelic copy of the gene. Therefore, in strict terms it is not possible to conclude that there has been actual bi-allelic antisense knockout (or knockdown) of the trapped gene or simply monoallelic loss of the reporter gene.

**“Promoter interference” as an alternative explanation of loss of reporter gene expression**

Promoter interference represents an alternative explanation to antisense inhibition that could account for the observation of reduced  $\beta$ -galactosidase expression following introduction of the transactivating vector. Here, the transactivator initiates antisense transcription leading to the recruitment of RNA polymerases and other components of the transcriptional machinery that advance in reverse direction along the DNA strand. Inevitably, this would impede the generation of mRNA transcripts from sense direction transcription and moreover interfere with the upstream promoter by disrupting assembly of the transcriptional complex. Such a mechanism would result in reduced expression of the  $\beta$ -galactosidase gene trapped downstream of the promoter without resort to any antisense inhibition. Critically, this mode of gene/promoter interference would operate only in *cis* and not in *trans* and thus the knockout would be monoallelic and not biallelic. In other words, unlike the putative antisense inhibition mechanism, there would be no true knockout of the trapped gene.

In order to determine which of the two mechanisms described above, namely antisense or promoter interference, might explain the derivation of  $\beta$ -galactosidase negative transactivator sub-clones, it would be necessary to quantify the level of mRNA transcript (by northern blot analysis or by RT-PCR) or that of the translated protein (by immunoblotting). The reasons these experiments were not done are explored in the overview section of this conclusion.

**The Cre recombinase vector results in restoration of  $\beta$ -galactosidase expression**

The introduction of the Cre recombinase into  $\beta$ -galactosidase negative transactivator sub-clones of the gene trap clones C2 (Tsix) and C16 resulted in the restoration of  $\beta$ -galactosidase expression. This observation suggests that transactivator excision is accompanied by reversal of antisense inhibition (or promoter interference) of the trapped gene, thereby supporting the foregoing conclusion that loss of  $\beta$ -galactosidase expression is a reliable measure of expression of the trapped gene. Moreover, it concludes the primary objective of demonstration of the experimental legitimacy of the antisense gene trapping approach.

**Shortcomings of the methylcellulose differentiation assay as a screening strategy**

However, the use of antisense gene trapping as a technically viable means of identifying genes involved in the regulation of ES cell differentiation was contingent upon the use of a suitable screening strategy. Here, the methylcellulose differentiation assay was used to compare single knockout gene trap clones with functional antisense double knockout sub-clones to identify any phenotypic differences. The assay was tailored in particular to assess haematopoietic differentiation through the use of specific haematopoietic growth factors (erythropoietin, interleukin-1, interleukin-3, and stem cell factor). However, the identification of differentiation mutants in this manner proved excessively labour-intensive for the purpose of high throughput screening on account of the inherent inefficiency of the methylcellulose assay, which lacks the qualitative and quantitative sensitivity necessary for discrimination of subtle abnormalities. Out of a total of 15 CCE-derived  $\beta$ -galactosidase negative antisense knockout sub-clones, none exhibited any gross abnormality of differentiation as compared with the respective  $\beta$ -galactosidase positive parental gene trap clones. Similarly, using a more refined methodology, none of the four E14-derived  $\beta$ -galactosidase negative antisense knockout sub-clones, E3 (Catns), E5 (Rex2), E6 (Nasp), and E12 (Ctbp2), exhibited any significant abnormality of differentiation. Hence, the primary objective of demonstrating technical viability of the antisense gene trapping strategy to identify differentiation genes in ES cells was not achievable. However, a limited functional analysis of the genes trapped in clones C2, E3, E5, E6 and E12 was performed as outlined below.

**Functional analysis of the Tsix (antisense to Xist) gene trapped in clone C2**

The gene trapped in clone C2 was Tsix. This gene is antisense to the Xist gene and both Tsix and Xist reside on the murine X chromosome. The DNA sequence derived by 5'RACE analysis therefore exhibited homology with both these genes, but the corrected 5' to 3' orientation of the sequence confirmed the integration of the gene trap vector into the Tsix gene and not its antisense counterpart. Tsix is a 40-kb RNA originating 15kb downstream of the Xist which is known to regulate X chromosome dosage compensation by controlling X inactivation in collaboration with a

*cis*-acting element called the X inactivation centre (Ref: Lee J.T. *et al.*, 1999). At the onset of X inactivation, *Tsix* expression becomes monoallelic, is associated with the future active X and persists until *Xist* is turned off. *Tsix* is not found on the inactive X once cells enter the X inactivation pathway suggesting that it plays an important role in regulating the early steps of X inactivation. This is supported by the observation that targeted deletion of *Tsix* gene in female mouse ES cells results in non-random inactivation of the mutant X chromosome (Ref: Lee J.T. and Lu. N., 1999). Thus, in the mouse it appears that *Tsix* regulates *Xist* in *cis* and thereby determines which X chromosome is to become silenced (Ref: Sado T. *et al.*, 2002). By contrast, in whole mouse embryos, deletion of *Tsix* results in disruption of the normal paternal imprinting in the extra-embryonic tissues suggesting that *Tsix* regulates this process also (Ref: Lee J.T., 2000).

The C2 gene trap clone was derived from CCE cells that are a karyotypically male ES cell line with a normal male complement of single X and Y chromosomes. Consequently, it is self-evident that the integration of the gene search vector into the *Tsix* gene on the single X chromosome results in insertional mutagenesis of the only copy of *Tsix* in the cell. Thus, the single knockout clone represents a complete gene knockout rendering the antisense knockout sub-clone redundant in this instance. On the other hand, if the role of the *Tsix* gene is purely antisense inhibition of *Xist*, then arguably the insertion of the gene trap vector does little to diminish its functionality. Either way, neither the single knockout clone nor antisense knockout sub-clone exhibited any significant abnormality of ES cell differentiation.

Perhaps the most notable insight from this gene trap clone is rather counterintuitive and that is the fact the  $\beta$ -galactosidase gene is expressed at all. Since *Tsix* is an RNA gene whose putative function is mediated via antisense binding of the RNA molecule to *Xist*, there is no translation of the gene into protein (Ref: Ogawa Y. and Lee J.T., 2004). Hence, it is unclear why the  $\beta$ -galactosidase of the gene search vector, or for that matter the neomycin phosphotransferase that confers G418 resistance, *should* be expressed. The fact that it is suggests that the mRNA, either despite the gene trap integration or because of it, is indeed translated into protein resulting in ubiquitous expression of  $\beta$ -galactosidase in all cells. If this is attributable to upstream sequence elements within the mRNA transcript, then it suggests that *Tsix* RNA is able to translocate to ribosomes in the cytoplasm and is not solely contained within the nuclear compartment.

Future experiments would seek to address both these issues. Firstly, confirmation that gene trap insertion into *Tsix* does indeed result in loss of function of the gene and secondly functional studies to investigate how this leads to alteration in the ability of *Xist* to initiate X chromosome silencing.

**Functional analysis of the Catns (Catenin-src) gene trapped in clone E3**

The gene trapped in clone E3 was Catns, or catenin-src (also referred to as p120), that is located on the murine chromosome 2. This is a novel tyrosine kinase substrate with armadillo homology that is implicated in ligand-induced signalling via a number of growth factor receptors and also in cell transformation by p60v-src (Ref: Reynolds A.B. *et al.*, 1992). Catns associates with the cell-cell adhesion protein complex containing E-cadherin and its cytoplasmic co-factors alpha-catenin, beta-catenin, and plakoglobin, and like other components of the cadherin/catenin complex, defects in Catns have been implicated in tumourigenesis (Ref: Reynolds A.B. *et al.*, 1996). More recently, Catns exists in a membrane-associated cadherin-bound pool, a cytosolic pool that affects Rho GTPases, and a nuclear pool that is thought to associate with the methylation-relevant transcriptional repressor Kaiso (Ref: Yanagisawa M. *et al.*, 2004). The cytosolic Catns has been shown to associate with microtubules and this binding is inversely related to its ability to regulate Rho GTPases (Ref: Franz C.M. and Ridley A.K., 2004).

In the gene trap clone E3, the gene trap integration within the Catns gene occurred within the first intron, which is located in the 5' untranslated region approximately 95 base pairs upstream of the predicted start codon.  $\beta$ -galactosidase expression in undifferentiated ES cells was apparent within the majority of cells but the pattern of staining revealed discrete foci. Typically each cell possessed only a single focus of  $\beta$ -galactosidase staining that was located peripherally. It is tempting to speculate that this reflects localisation of the Catns protein within either the cytosolic or membrane-associated pool, but this possibility is tempered by the fact that the  $\beta$ -geo fusion protein cannot carry any upstream localisation sequence tag (although it could conceivably carry a downstream localisation tag). It is possible however that the fusion mRNA, rather than the protein, carries upstream sequence information to target the transcript to one portion of the cell thereby accounting for the discrete nature of the  $\beta$ -galactosidase expression. Introduction of the transactivator vector into the E3 gene trap clone failed to generate  $\beta$ -galactosidase negative sub-clones implying either that the antisense inhibition failed on this occasion, possibly due integration occurring within the first intron, or was insufficient to overcome the high level of expression of the gene. Methylcellulose differentiation of the E3 gene trap clone was unremarkable and  $\beta$ -galactosidase expression occurred ubiquitously without any evidence of cell lineage restriction.

Future experiments would seek to determine whether the  $\beta$ -galactosidase expression pattern observed in the gene trap clone truly reflects the intracellular localisation of the Catns protein using fluorescently labelled antibodies and to functionally characterise the gene using more established knockout technologies such as RNA interference, antisense or targeted deletion.

**Functional analysis of the Rex2 (retinoic acid repression) gene trapped in clone E5**

The gene trapped in clone E5 was Rex2, a gene that was originally identified using differential hybridisation on account of its reduced expression during retinoic acid-induced differentiation of F9 teratocarcinoma cells (Ref: Faria T.N. *et al.*, 1998). Rex2 has also been identified in a large-scale screen of mammalian proteins and their nuclear sub-compartments (Ref: Sutherland H.G. *et al.*, 2001). It is not clear, despite its name, that Rex2 bears any functional or evolutionary ties with Rex1, which is a zinc finger gene also regulated by retinoic acid and a commonly-used marker of undifferentiated ES cells.

In the gene trap clone E5, the gene trap integration within the Rex2 gene occurred within the first intron, which is located in the 5' untranslated region approximately 209 base pairs upstream of the predicted start codon.  $\beta$ -galactosidase expression was uniform throughout all undifferentiated ES cells but without any specific intracellular staining pattern. Upon ES cell differentiation in methylcellulose,  $\beta$ -galactosidase expression was apparent within the embryoid body but never within cells undergoing haemoglobinisation. Thus, it would appear that Rex2 is down regulated during erythroid differentiation. This conclusion is supported by the observation above that the trapped gene *Catns* is expressed in erythroid cells, thus serving as a convenient positive control (note that the alternative explanation, namely that haemoglobinised cells stain poorly with X-gal substrate, was not excluded). Apart from this finding, the methylcellulose differentiation of E5 ES cells was otherwise unremarkable both for the  $\beta$ -galactosidase gene trap clone itself and also for the derived  $\beta$ -galactosidase negative antisense knockout sub-clone.

When E5 clone ES cells were induced to undergo differentiation by seeding at low density on tissue culture plates and removal of LIF from the medium, an interesting observation was made. The presence of G418 at low dose (100 $\mu$ g/ml) resulted in failure of the normal fibroblastic differentiation seen on the G418-free plate and also in control plates that used other G418-resistant gene trap clones. Instead, the presence of low dose G418 caused the E5 ES cells to remain in small colonies of undifferentiated cells even despite the absence of LIF in the medium. Thus, it would appear that G418, either through enforced expression of the trapped Rex2 gene and the downstream neo that confers resistance or through antibiotic-induced death of those cells that switch off Rex2 and neo expression, results in persistence of the undifferentiated ES state in clone E5 cells.

The rationale for this experiment was based upon previously published work showing that ES cells carrying a transgenic neo gene, driven by the  $\alpha$ -cardiac myosin heavy chain promoter, produced a homogenous population of cardiomyocytes when differentiated in the presence of G418 due to

selection of cardiac lineage-restricted cells (Ref: Klug M.G. *et al.*, 1996). In the case of the E5 gene trap clone, where the neo is under the control of the Rex2 promoter, it appears that G418 selection restricts the ES cells to the undifferentiated state even in the absence of LIF.

Thus, Rex2 arguably represents a candidate stem cell gene in ES cells. This conclusion is consistent with published data showing that induction of differentiation in teratocarcinoma cells, using retinoic acid, results in down-regulation of Rex2 (Ref: Faria T.N. *et al.*, 1998) and leads to the testable hypothesis that enforced expression of Rex2 is able to maintain the pluripotent ES state in a LIF-independent fashion.

#### **Functional analysis of the Nasp (histone-binding protein) gene trapped in clone E6**

The gene trapped in clone E6 was Nasp, or nuclear autoantigenic sperm protein. This is a histone-binding protein that is a homologue of the N1/N2 gene expressed in oocytes of *Xenopus laevis* (Ref: Richardson R.T. *et al.*, 2000). It is expressed in both murine ES cells and somatic cells and is carried on the mouse chromosome 4 (Ref: Richardson R.T. *et al.*, 2001). Nasp has been shown to be complexed with H1 linker histone, suggesting a role as a histone transport protein, and evidence points to its involvement in regulating the cell cycle (Ref: Alekseev O.M. *et al.*, 2003).

In the gene trap clone E6, the gene trap integration within the Nasp gene occurred within the second intron, which is located approximately 106 base pairs downstream of the predicted start codon.  $\beta$ -galactosidase expression was evident throughout nearly all undifferentiated ES cells but without any specific intracellular staining pattern. Upon ES cell differentiation in methylcellulose,  $\beta$ -galactosidase expression was apparent within the embryoid body but never within cells undergoing haemoglobinisation. Thus, as for the Rex2 gene previously, it would appear that Nasp is down-regulated during erythroid differentiation (although again the alternative explanation that staining with X-gal substrate is unusually weak within haemoglobinised cells has not been excluded). Apart from this finding, the methylcellulose differentiation of E6 ES cells was otherwise unremarkable both for the  $\beta$ -galactosidase gene trap clone itself and also for the derived  $\beta$ -galactosidase negative antisense knockout sub-clone.

#### **Functional analysis of the Ctbp2 (C-terminal binding protein) gene trapped in clone E12**

The gene trapped in clone E12 was Ctbp2, or C-terminal binding protein 2. This gene, which resides on the mouse chromosome 21, was originally identified on the basis of its sequence similarity to the Ctbp1 gene whose protein product is known to associate with the C-terminus of the adenoviral oncoprotein E1A via a Pro-X-Asp-Leu-Ser motif (Ref: Katsanis N. and Fisher E.M., 1998). Unlike Ctbp1, which is expressed from embryo to adult, Ctbp2 is expressed mainly

during embryogenesis (Ref: Furusawa T. *et al.*, 1999) and appears to be critical for axial patterning leading to embryonic lethality in mutant null mice (Ref: Hildebrand J.D. and Soriano P., 2002). The characterisation of the CtBP2 protein has revealed that it functions as a potent co-repressor for a number of transcriptional factors including the Basic Kruppel-like factor (BKLF) that is highly expressed in erythroid cells (Ref: Turner J. and Crossley M., 1998), and that this action is mediated via binding to a minimal repression domain (Ref: Thio S.S. *et al.*, 2004). *Ctbp2* has also been shown to interact with the hematopoietic, zinc-finger protein FOG1 that is essential for the development of the erythroid and megakaryocytic lineages (Ref: Katz S.G. *et al.*, 2002).

In the gene trap clone E12, the gene trap integration within the *Ctbp2* gene occurred within the second intron, which is located approximately 53 base pairs downstream of the predicted start codon.  $\beta$ -galactosidase expression was apparent in nearly all undifferentiated ES cells and exhibited a specific staining pattern that was strongest in the peri-membrane region with occasional sub-membrane foci. However, given the fact that this protein is known to act within the nucleus as a transcriptional co-repressor, this unusual staining pattern is unlikely to indicate physiological localisation of the protein within the cell and perhaps more likely to represent staining artefact.

Upon ES cell differentiation in methylcellulose,  $\beta$ -galactosidase expression was apparent within the embryoid body but never within cells undergoing haemoglobinisation. Thus, as for both *Rex2* and *Nasp* genes previously, it would appear that *Ctbp2* is down-regulated during erythroid differentiation (although once again it is not possible to exclude the possibility that staining with X-gal substrate is inexplicably weak within haemoglobinised cells). Apart from this finding, the methylcellulose differentiation of E12 ES cells was otherwise unremarkable both for the  $\beta$ -galactosidase gene trap clone itself and also for the derived  $\beta$ -galactosidase negative antisense knockout sub-clone.

The apparent absence of expression in cells undergoing erythroid differentiation is in interesting contrast with a related gene called FOG1. This gene is highly expressed in erythroid cells and acts in concert with the transcriptional factor GATA1 to regulate erythropoiesis. In view of the known interaction between CtBP2 and FOG1 (Ref: Katz S.G. *et al.*, 2002), this reciprocal expression pattern in erythroid cells suggests a model in which the lifting of CtBP2-mediated repression of FOG1 is an important step during haematopoietic differentiation that leads to normal erythropoiesis.

### Overview

As stated previously, the primary objective of this chapter was the proof-of-principle demonstration that antisense gene trapping represents an experimentally legitimate and technically viable method for the identification of putative differentiation genes in murine ES cells with the secondary objective being the functional characterisation of these genes through analysis of the gene trap and antisense knockout clones. Gene trapping in ES cells represents a powerful form of random mutagenesis that has been used to reveal insights into the expression patterns of many genes and to generate transgenic mice with the potential for breeding knockout offspring. However, the inherent weakness of gene trapping lies in the fact that the gene-trap clones are only hemizygous knockouts and therefore unlikely to yield a mutant cellular phenotype. Antisense gene trapping represents a refinement of the established methodology designed to overcome this shortcoming by enabling generation of functional double-knockout ES cell clones through transactivation of a reversed promoter contained upstream of the gene-trap site. The RNA strand thus generated would then inhibit gene expression via antisense hybridisation with the mRNA transcript.

However, as explained previously, this *trans* inhibition is not the sole mechanism able to account for the resulting  $\beta$ -galactosidase negative transactivator sub-clones. The alternative explanation is that of promoter interference operating in *cis* via processes of mechanical hindrance of the transcriptional complex. In order to discern which of these explanations best accounts for the reversal of  $\beta$ -galactosidase expression, it would have been necessary to quantify either the level of mRNA transcript or alternatively that of the translated protein. For mRNA levels this would be by northern blot analysis or by RT-PCR, while for protein levels this would be by western blot analysis. In the case of antisense inhibition operating in *trans*, there should be complete biallelic knockout of gene expression resulting in minimal or absent levels of mRNA or protein. By contrast, in the case of promoter interference operating in *cis*, the knockout would only be monoallelic and therefore the mRNA and protein levels should only fall by approximately half compared with wild-type ES cells and not at all compared with the mother gene-trap clone that is already a hemizygous knockout.

In view of this alternative mechanistic explanation of the observed results, and the failure to discern experimentally between them, it is arguably the case that the primary objective of this chapter was not met. However, it should be borne in mind that the primary objective was not the proof-of-principle demonstration of the antisense gene-trap methodology but rather its ability to yield phenotypically mutant ES cell clones. This after all was the manner in which antisense gene trapping originally proved its worth in NIH3T3 cells as a screen for tumour susceptibility genes

(Ref: Li L. and Cohen S.N., 1996). Moreover, in this original demonstration of the methodology, those mutant 3T3 clones selected for anchorage independence by an agar assay did indeed show knockout of the trapped gene with loss of mRNA expression via northern blot analysis. In other words, there already exists a proof-of-principle of the antisense gene trapping methodology, albeit in 3T3 cells, and for this reason further experiments to prove the worth of the methodology in ES cells were not pursued. Rather, the difficulty encountered in applying this strategy to ES cells was that of failure to derive mutant clones due to a shortcoming of the methylcellulose differentiation assay. Indeed, it is the screening strategy, rather than the knockout methodology, that represents the major weakness of the whole forward genetics approach. This is particularly true in the case of identifying differentiation genes in ES cells where the screening strategy must exhibit attributes of discrimination, accuracy and reproducibility as well as being both high-throughput and non-labour intensive.

Despite the failure to fulfil the primary objective of this chapter and derive mutant ES cell clones, the secondary objective of functional characterisation of the gene-trap clones was arguably achieved. However, the failure of any of the  $\beta$ -galactosidase-negative transactivator sub-clones to yield aberrant differentiation meant abandonment of the antisense gene trapping methodology in favour of a new strategy for achieving gene knockout in ES cells, namely RNA interference.

# **Chapter 4**

## **Establishing the biology of RNA interference**

### **Results 2**

## 4.1 Introduction

RNA interference (RNAi) refers to the process whereby double-stranded RNA molecules are processed into short interfering RNAs (siRNAs) able to induce knockdown of expression of genes sharing complementary sequences. RNAi appears to represent an evolutionarily ancient cellular process whose primary role is probably that of defence against foreign transposable elements but may also have a secondary role in gene regulation. The principal aim of this chapter was to explore the biology of the interference effect using a simple and tractable experimental model, namely the knockdown of expression of a Firefly luciferase reporter plasmid using sequence-specific siRNAs that are transiently co-transfected with the plasmid into the cell. This aim comprises a number of goals:

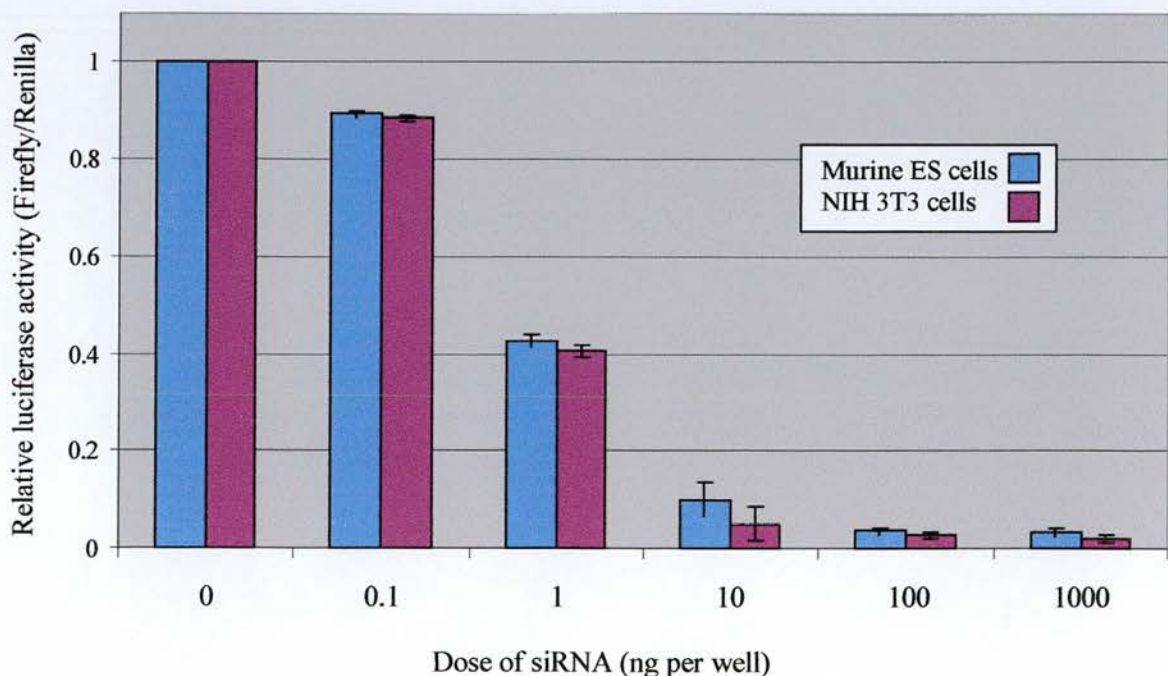
- ❖ To establish what is necessary and sufficient for induction of siRNA-induced gene knockdown
- ❖ To demonstrate the importance of sequence-specific fidelity in the interference process
- ❖ To investigate the potency and dose-dependency of siRNA-induced gene knockdown
- ❖ To establish the stoichiometric ratio of sense and antisense RNA strands for interference
- ❖ To investigate the effect of additional siRNAs on the potency of the interference effect
- ❖ To establish whether there is differential functionality between the sense and antisense strands
- ❖ To study the dynamics of the interference reaction by analysing the time course of knockdown
- ❖ To study the kinetics of the interference reaction with regard to competitive reversibility
- ❖ To study the magnitude and persistence of the interference knockdown effect
- ❖ To confirm that knockdown of a transient gene can be recapitulated against a stable gene
- ❖ To compare the efficacy of chemically synthesised siRNAs with *in vitro* transcribed siRNAs
- ❖ To explore methods of enhancing delivery of siRNAs molecules into cells

#### 4.2 Sequence-specific short interfering RNA molecules suppress *Firefly* luciferase expression in a potent and dose-dependent manner

Chemically synthesised siRNAs, sharing sequence homology to the *Firefly* luciferase gene (position 154-172 from start codon), result in a potent knockdown of gene expression using a transient transfection experimental system in both NIH 3T3 cells and murine ES cells. Here, the siRNAs were co-transfected together with the *Firefly* luciferase reporter plasmid (pGL3) and a *Renilla* luciferase internal control (pRL-CMV) into 3T3 cells and ES cells. This resulted in a highly specific and dose-dependent suppression of *Firefly* luciferase expression in both cell lines that was maximally potent at the highest concentration tested. The specificity of the interference was evident from the observation that there is *no* knockdown of expression of the *Renilla* internal control (except at very high doses of siRNAs where plasmid transfection itself is likely to be compromised). This experiment highlights the *potency* and *specificity* inherent to the RNAi process in 3T3 and ES cells, both of which exhibit remarkably similar knockdown profiles illustrated in the graph below (Figure 4.1).

**Figure 4.1: Sequence-specific siRNAs result in potent and specific gene interference.**

Graph demonstrating the potent and specific nature of the knockdown effect resulting from transient transfection of Firefly-specific siRNAs together with a Firefly luciferase target plasmid and *Renilla* luciferase internal control plasmid in NIH 3T3 and murine ES cells. Error bars represent the range derived from duplicate experiments.

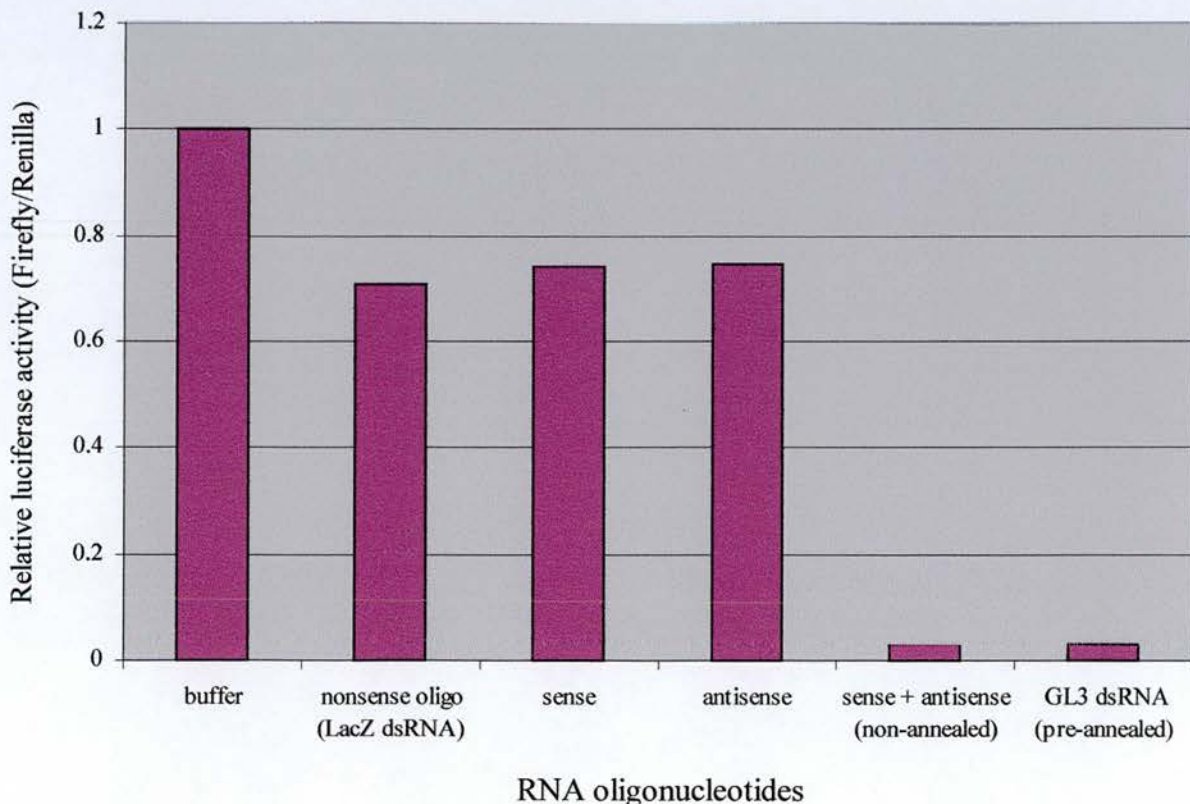


### 4.3 Interference of *Firefly* luciferase expression is dependent upon the presence of both sense and antisense RNA oligonucleotide strands

Transfection of either sense or antisense oligonucleotide strands *alone* is insufficient to induce knockdown of *Firefly* luciferase expression, compared with a non-specific siRNA control, in a repeat of the 3T3 transient transfection experiment above. However, if *both* sense and antisense strands are transfected in an equimolar ratio, then there is potent knockdown of *Firefly* expression irrespective of whether or not these have been pre-annealed by cooling slowly from 95°C. Hence, *both* sense and antisense RNA 21mer strands, sharing 19bp of homology to each other and the target gene, would appear to be necessary and sufficient for RNAi knockdown in mammalian cells as illustrated graphically below (Figure 4.2).

**Figure 4.2: Both sense and antisense RNA strands are required for gene interference.**

Sense and antisense RNA molecules, irrespective of a prior explicit annealing step, appear to be necessary and sufficient for knockdown of *Firefly* luciferase in this experimental system. Data-points are derived from single experiments and hence error bars have not been applied.

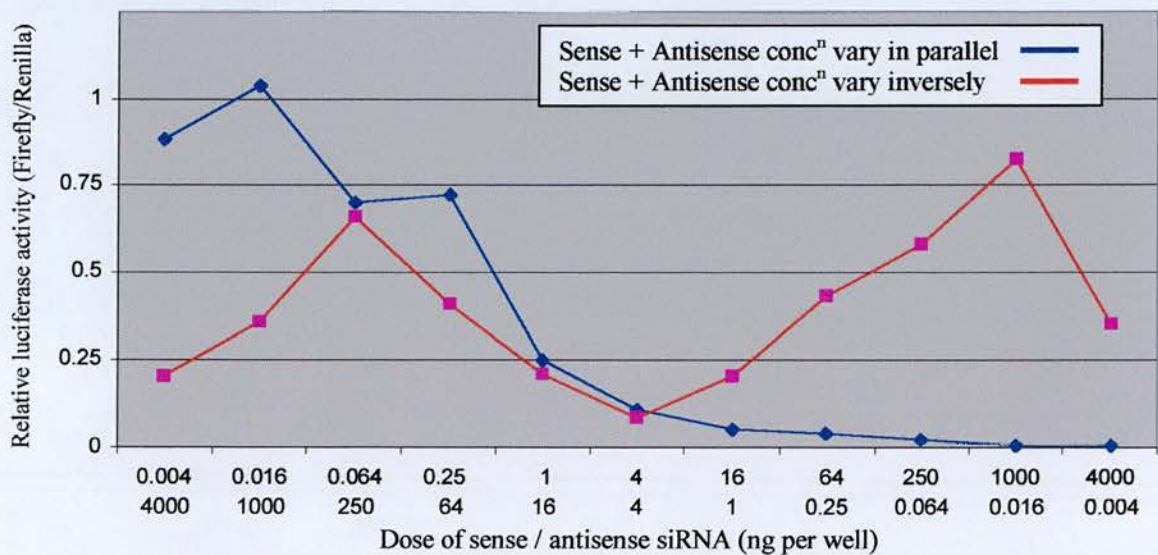


#### 4.4 The interference effect is maximal when sense and antisense RNA strands are in equimolar proportions suggesting a stoichiometry of 1 to 1

When the molar ratio of sense and antisense RNA strands are varied relative to one another, it is apparent that the interference effect is maximal when they are present in *equal* amounts (Figure 4.3). The result of this experiment, performed in *both* 3T3 and ES cells, implies that the interference process has an underlying stoichiometry of 1 to 1 (Figure 4.4). This observation strongly suggests that hybridisation of the two RNA molecules, to form a double-stranded RNA intermediary species, is a key step in the interference mechanism.

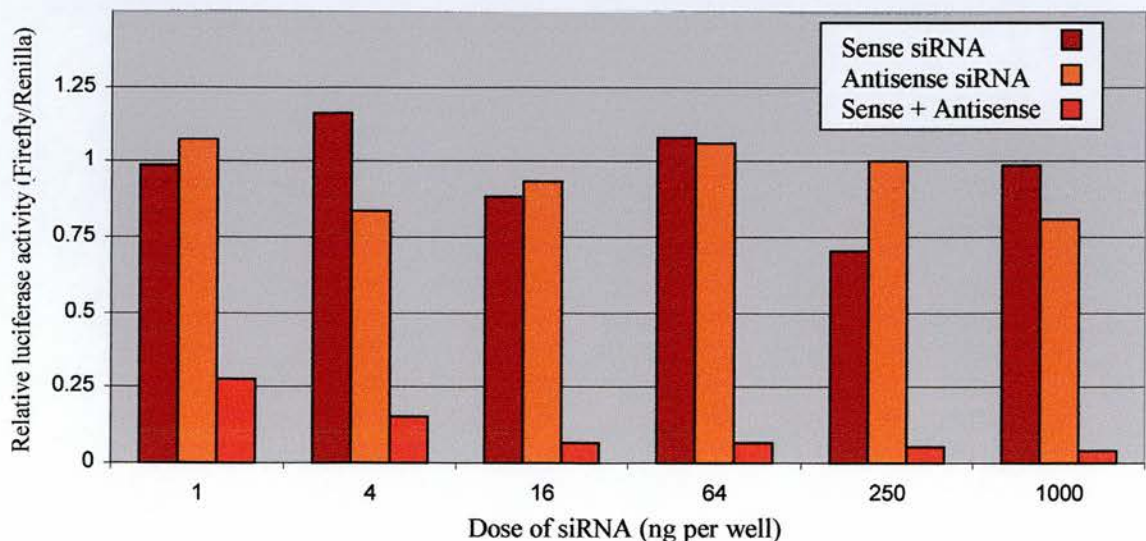
**Figure 4.3: Gene interference is maximal when sense and antisense strands are equimolar.**

[Note that the “interference” seen at either extreme might be attributable to the non-equal delivery of the Firefly and *Renilla* plasmids due compromised transfection on account of the high siRNA dose.] Data-points are derived from single experiments and hence error bars have not been applied.



**Figure 4.4: The 1:1 stoichiometry of RNA interference is apparent at all concentrations**

Double-stranded RNA yields more interference than sense ( $p < 0.001$ ) or antisense ( $p < 0.001$ ) alone.

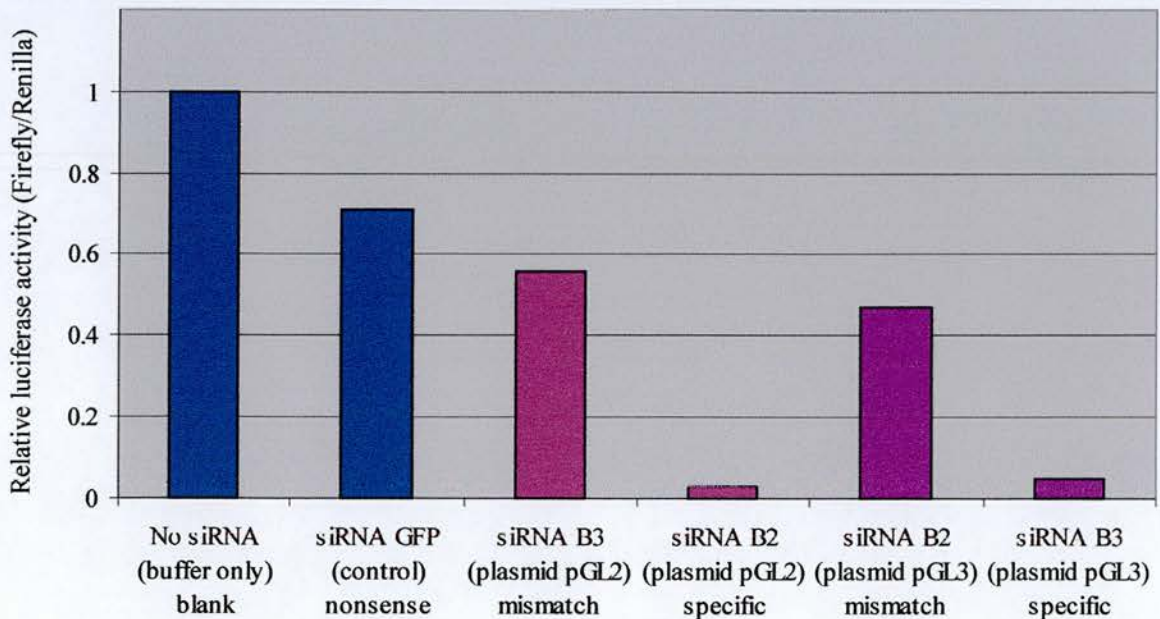


#### 4.5 The specificity of the interference effect is critically dependent upon sequence fidelity of the siRNA molecule with its target region

The RNAi knockdown effect is largely abrogated if the siRNA carries a three-nucleotide mismatch with the target gene sequence. Thus, if instead of the pGL3 *Firefly* plasmid the interference experiment is performed using the pGL2 plasmid, which carries discrepancies at three discrete sites within the region of homology, then the level of knockdown is much reduced. Conversely, if siRNA specific for pGL2 is co-transfected with the pGL3 plasmid, then likewise the level of knockdown is much reduced by comparison with co-transfected pGL2 plasmid. This reaffirms the specificity of the RNAi mechanism and highlights the importance of complete sequence fidelity within the region of homology as illustrated below (Figure 4.5). Moreover, these findings were borne out in *both* 3T3 cells and ES cells and thus appear to be intrinsic properties of the RNAi process and not dependent upon the cell line used.

**Figure 4.5: Gene knockdown via RNAi is critically dependent upon sequence fidelity.**

Potent knockdown of Firefly luciferase is observed in 3T3 cells when the siRNA used is specific for the target gene (siRNA-B2/plasmid-pGL2 or siRNA-B3/plasmid-pGL3). However, this is largely abrogated when the siRNA carries a three-nucleotide mismatch (siRNA-B3/plasmid-pGL2 or siRNA-B2/plasmid-pGL3). Data-points are derived from single experiments and hence error bars have not been applied.



<b>siRNA - B3 (pGL3)</b>	CUUACGCUGAGUACUUCGAdTdT
	*         *     *
<b>siRNA - B2 (pGL2)</b>	CGUACGCGGAUACUUCGAdTdT
<b>siRNA control (GFP)</b>	GUUCAGCGUGUCCGGCGAGdTdT

#### 4.6 siRNA molecules that have sequence specificity for different regions within the target gene exhibit differential interference potency

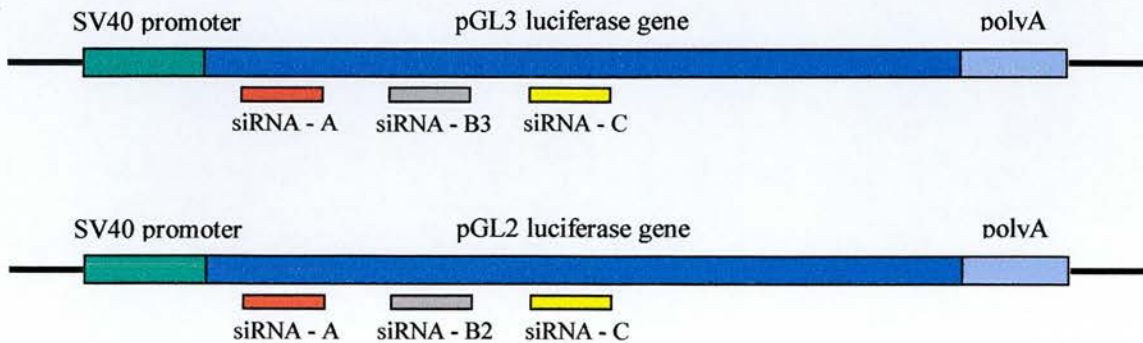
When the potency of different siRNAs are compared, each one sharing a different region of homology within the target gene sequence, there are clear differences in potency of the knockdown observed. Thus, for the pGL3 *Firefly* luciferase plasmid, siRNA sharing homology with the region 154-172 (siRNA-B3) gives 99% knockdown, whereas siRNA sharing homology with the upstream region 23-41 (siRNA-A) gives only 61% knockdown, and siRNA sharing homology with the downstream region 195-213 (siRNA-C) gives 89% knockdown (Figures 4.6, 4.7 and 4.8). Therefore, it is apparent that the potency of the interference effect is dependent upon a combination of both the sequence of the siRNA and also the site of homology within the target gene. These findings appear to be largely independent of cell line used being broadly similar for *both* 3T3 cells and ES cells.

**Figure 4.6: Table and schematic diagram showing the site and sequence of various siRNAs.**

Locations of the siRNA target sites within the Firefly gene: A) 26-46; B) 153-173; and C) 193-213.

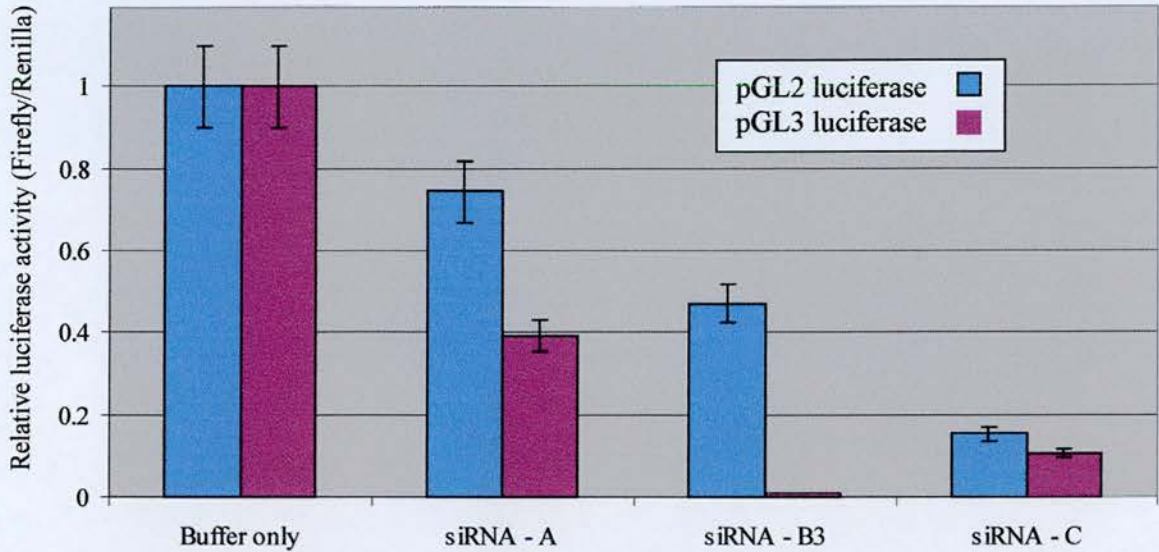
Note that siRNA-B3 is mismatched at three sites with pGL2 and vice versa for siRNA-B2 and pGL3.

NAME	POSITION	SEQUENCE	pGL2	pGL3
siRNA - A	26-46 bp	GAAAGGCCCGGCGCCAUCdTdT	Specific	Specific
siRNA -B3	153-173 bp	CUUACGCUGAGUACUUCGAdTdT	Mismatch	Specific
siRNA - B2	153-173 bp	CGUACGC <del>G</del> GAAUACUUCGAdTdT	Specific	Mismatch
siRNA - C	193-213 bp	GCUAUGAAACGAUAUGGGCdTdT	Specific	Specific



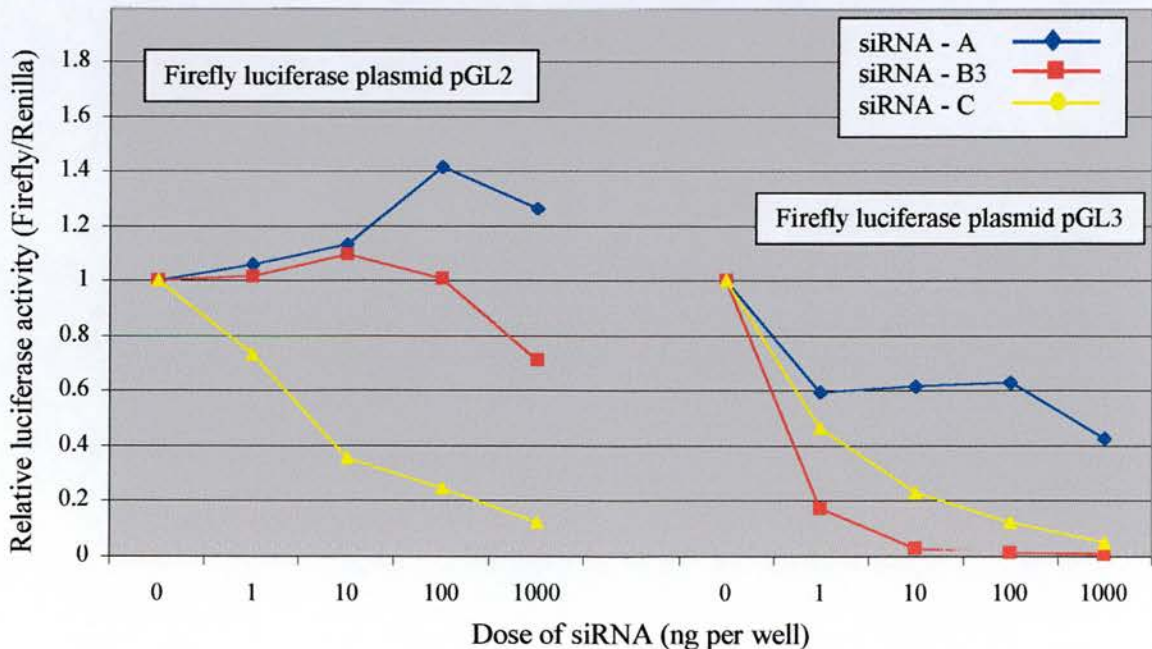
**Figure 4.7: Graph showing the differential effects of three siRNAs against pGL2 and pGL3.**

For the Firefly luciferase on plasmid pGL3, knockdown potency is greatest for siRNA B3 > C > A. For plasmid pGL2, knockdown is greatest for siRNA C > B3 > A. Note B3 is a mismatch for pGL2. Transfections performed in 3T3 cells. Error bars represent the range derived from duplicate experiments.



**Figure 4.8: Graph showing same experiment performed over a range of siRNA concentrations.**

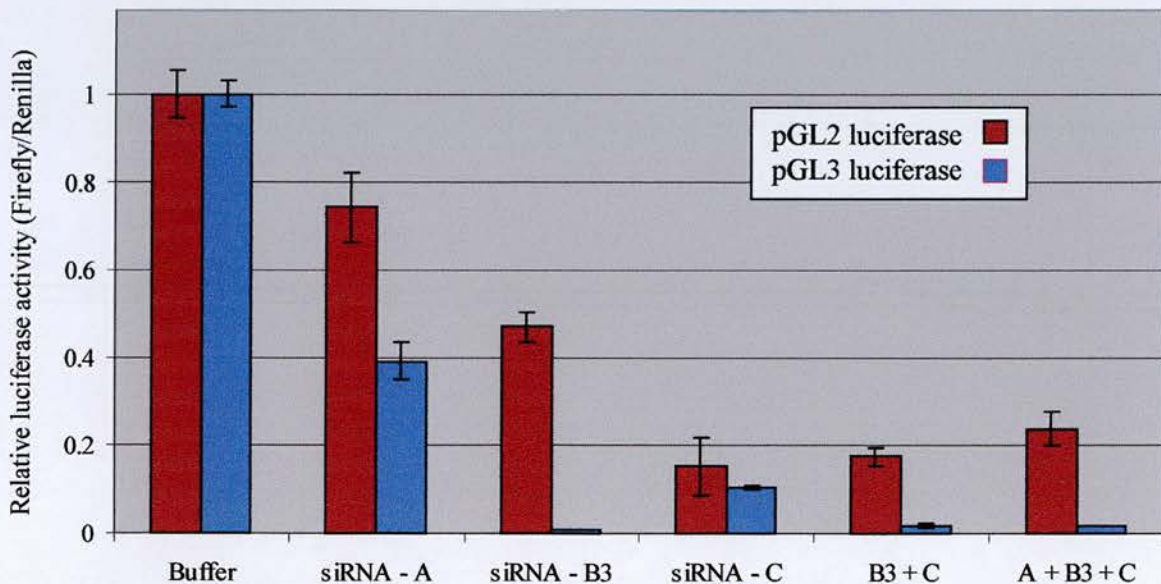
For the Firefly luciferase on plasmid pGL3, knockdown potency is greatest for siRNA B3 > C > A. For plasmid pGL2, knockdown is greatest for siRNA C > B3 > A. Note B3 is a mismatch for pGL2. Data-points are derived from single experiments and hence error bars have not been applied.



#### 4.7 Combinations of interfering siRNA molecules yield gene knockdown effects that approximate to the summation average of each individual siRNA

To determine the gene knockdown effects of using multiple siRNA molecules in combination, transient co-transfection experiments were performed using equimolar amounts of three siRNA species, A, B3, and C, to target *Firefly* luciferase in 3T3 cells. The results demonstrate that the observed level of knockdown of either the pGL2 or pGL3 *Firefly* gene is an approximate summation average of each siRNA used individually without any appreciable loss of the interference effect. This is illustrated in the graph below (Figure 4.9).

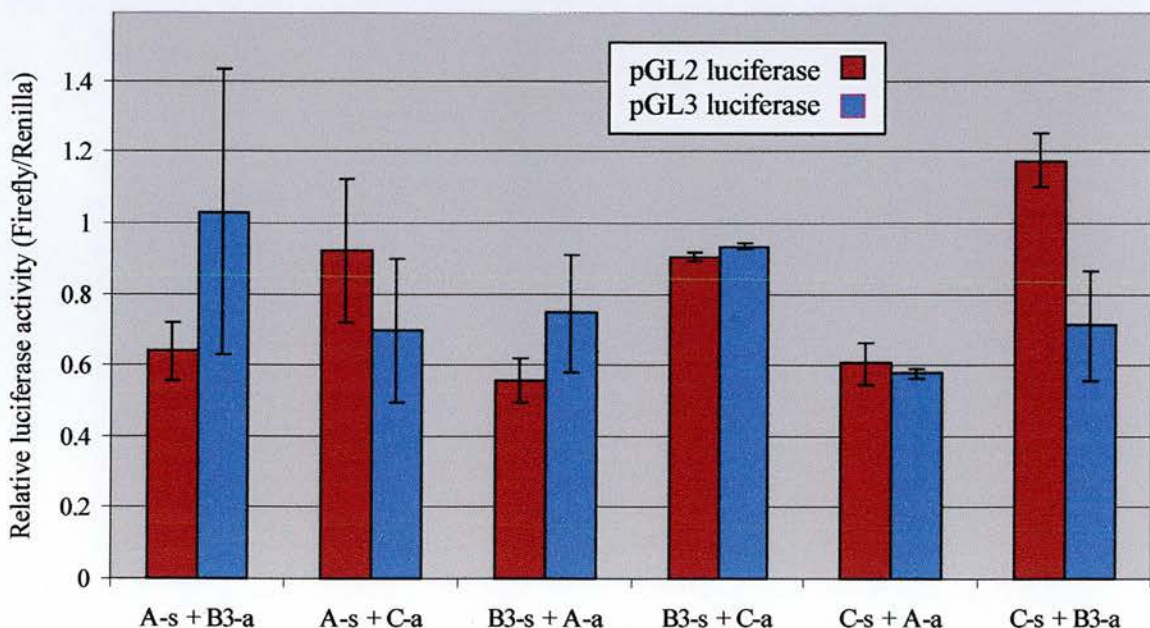
**Figure 4.9: Graph showing the effect of different combinations of siRNAs on Firefly knockdown.** Combinations of siRNAs yield a knockdown effect that is the summation average of each individual siRNA. Here, the total siRNA is maintained at 100 $\mu$ g/well for each experiment and thus the interference observed for two or three siRNAs is a weighted average of each used individually. [Note that B3 is a mismatch for pGL2]. Error bars represent the range derived from duplicate experiments.



#### 4.8 The combination of the sense strand of one siRNA with the non-homologous antisense strand of another siRNA fails to generate an interference effect

In order to explore the possibility that non-homologous pairs of sense and antisense RNA strands might be able to induce an interference response, incorrectly matched sense/antisense RNAs were co-transfected with either pGL2 or pGL3 plasmids in 3T3 cells and tested for their ability to knockdown luciferase expression. The results reveal that when the sense strand of one siRNA is used in conjunction with an equimolar amount of the antisense strand of a *different* siRNA, no discernible gene knockdown effect is observed (Figure 4.10). Thus, failure to generate a double-stranded RNA molecule due to non-homology between the sense and antisense strands appears to abrogate the interference phenomenon. This experiment was predicated on the possibility that sense and antisense RNA strands might serve as RNA primers for an endogenous RNA-dependent RNA polymerase resulting in secondary generation of double stranded RNA that could then mediate gene interference. A similar process, termed “transitive RNAi”, has been reported to occur in *C. elegans*. However, this experiment is a poor model of transitive RNAi since the non-homologous RNA strands fail to form siRNA duplexes and the only conclusion that might be drawn is that such duplex formation appears to be an obligatory step in the interference process.

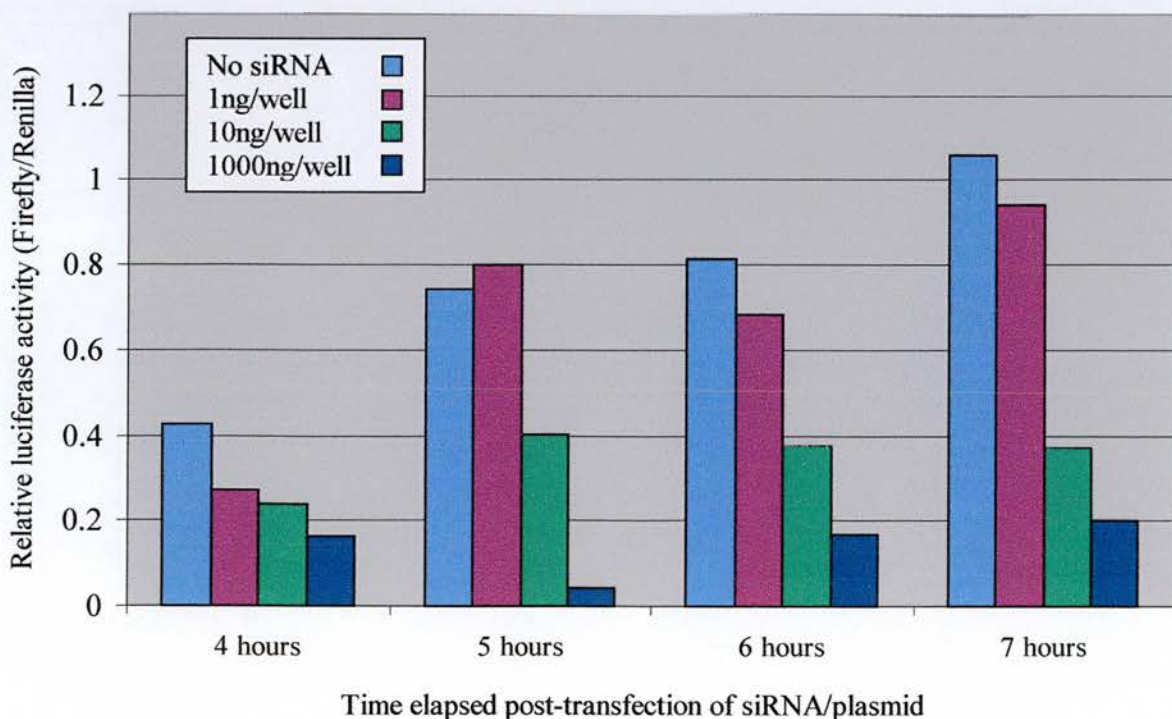
**Figure 4.10: Graph showing effect of non-matched sense/antisense pairs on gene expression.** Mismatched sense and antisense combinations of the siRNA molecules A, B3, or C, in order to generate a non-homologous pair, results in the complete abrogation of the gene interference effect. The notation in the graph uses “s” to denote sense strand and “a” to denote antisense strand. Note that the positive experimental control, using paired siRNA A, B3 and C, is shown in the previous graph. [Total siRNA used was 100µg/well for each experiment. Note that B3 is a mismatch for pGL2]. Error bars represent the range derived from duplicate experiments.



#### 4.9 siRNA-mediated interference of co-transfected *Firefly* luciferase is apparent as soon as luminescence is detectable

In order to explore the dynamics of the interference knockdown effect in 3T3 cells, luciferase activity was assayed at early time-points following plasmid/siRNA co-transfection. As previously, this experiment used the potent siRNA, B3 (homologous to region 154-172 of the *Firefly* luciferase gene), transfected at varying concentrations together with the pGL3 plasmid. The results demonstrate an unequivocal dose-dependent knockdown of *Firefly* luciferase as early as four hours post-transfection, which represents the earliest time-point for detection of luminescence (Figure 4.11). Thus, in transient co-transfection, there does not appear to be any discernable lag period for the induction of the siRNA-induced interference effect, which rather becomes evident commensurate with expression of the target luciferase. Using a paired Student's t-test to compare the knockdown induced by siRNA at 1000ng/well with that of the control for all four time-points, this result is shown to be statistically significant ( $p=0.02$ ).

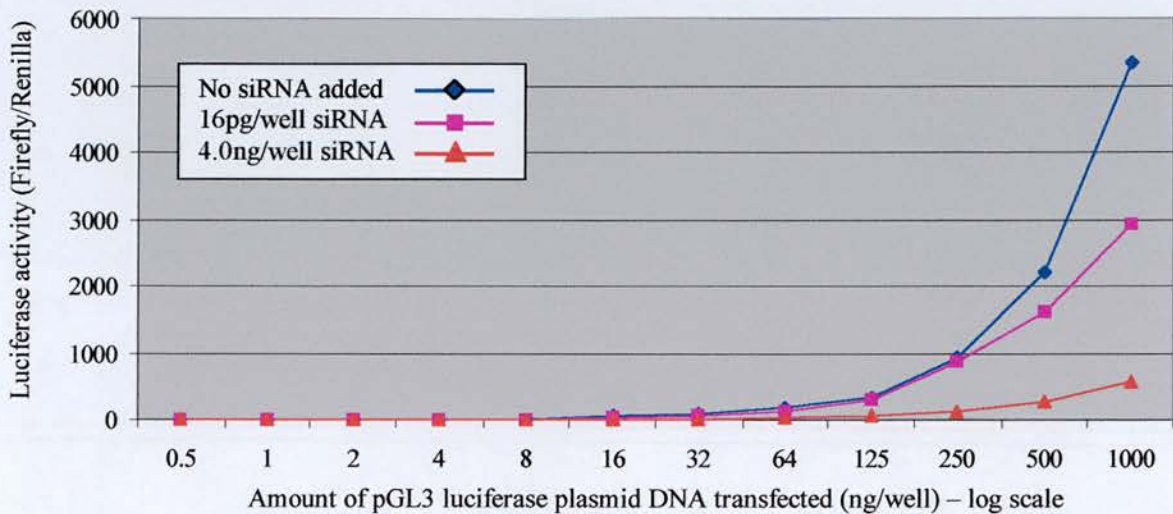
**Figure 4.11: Graph showing gene knockdown effects of siRNA-B3 at various time-points.** siRNA-B3, co-transfected together with pGL3, exhibits a dose-dependent knockdown of *Firefly* luciferase expression (compared with *Renilla* internal control), as soon as luminescence is detectable (four hours). This suggests that RNA interference acts commensurately with luciferase expression. Data-points are derived from single experiments and hence error bars have not been applied. There is a statistically significant difference between 1000ng/well siRNA-B3 and the control using a paired Student's t-test ( $p=0.02$ ) consistent with a dose-dependent knockdown throughout the experiment.



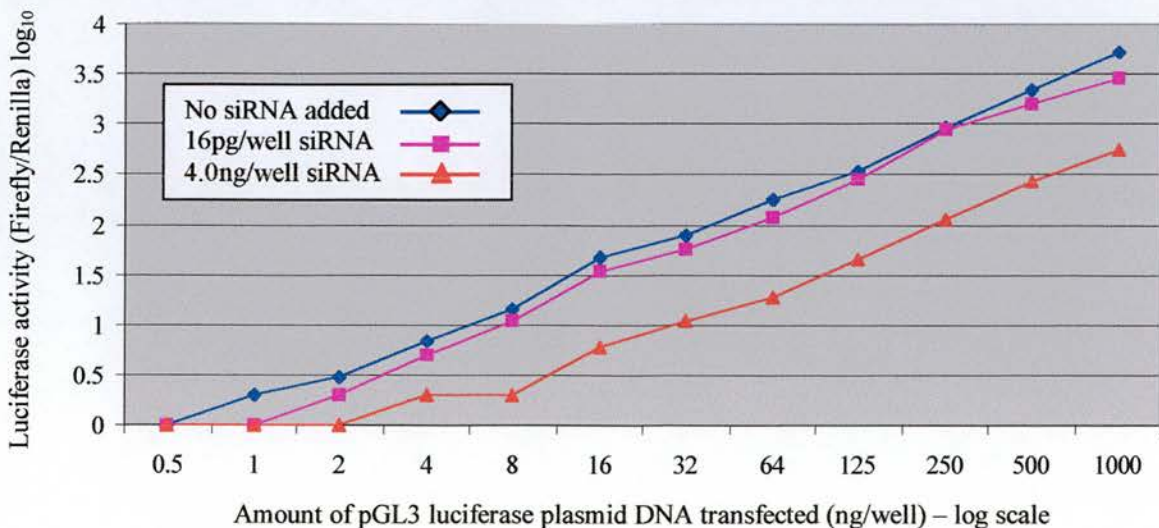
#### 4.10 siRNA-mediated knockdown of a co-transfected plasmid-borne gene exhibits first order competitive reaction kinetics

In order to explore the kinetics of the interference phenomenon, the level of knockdown of *Firefly* luciferase mediated using a fixed amount of siRNA co-transfected with *varying* concentrations of target plasmid was assayed. This reveals that the level of luciferase knockdown observed for a fixed amount of siRNA is proportional to the logarithm of the concentration of the target gene implying that siRNA-induced interference obeys first order reaction kinetics. Moreover, it can be overcome by using greater amounts of plasmid suggesting that interference is reversibly competitive. This is illustrated by the exponential curves of the knockdown profiles plotted on log-linear graphs (Figure 4.12) that become parallel when plotted on log-log graphs (Figure 4.13).

**Figure 4.12: Graph revealing the kinetics of siRNA-mediated gene knockdown (linear-log plot).**



**Figure 4.13: Graph revealing the kinetics of siRNA-mediated gene knockdown (log-log plot).**

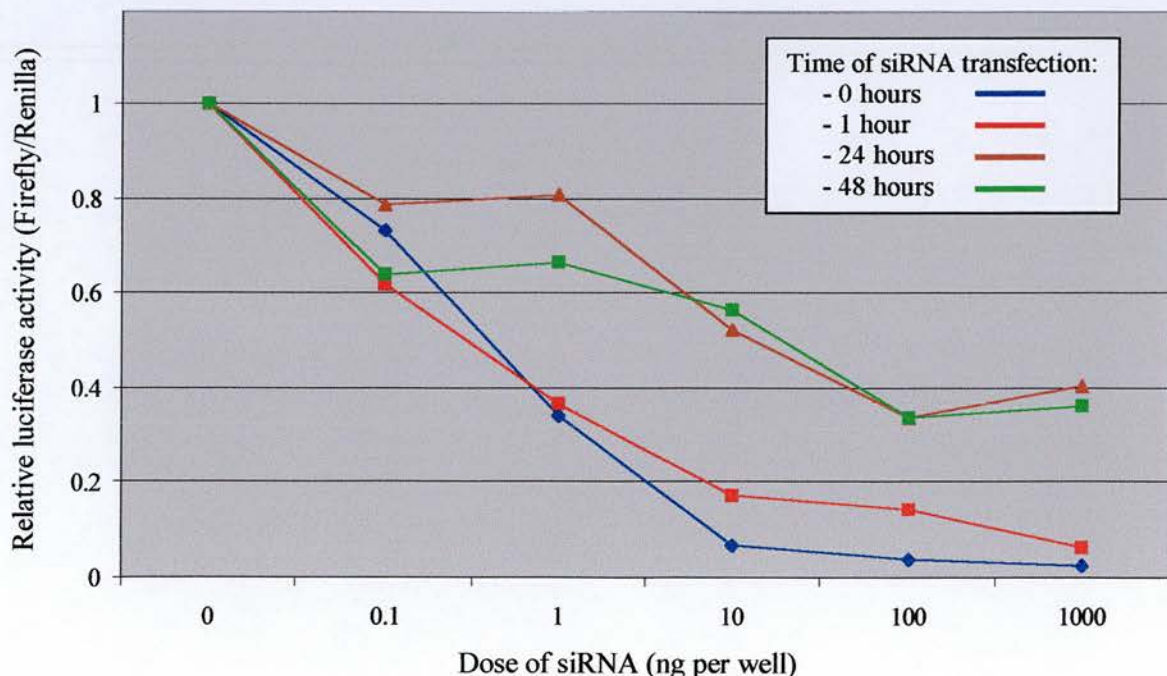


#### 4.11 The magnitude of siRNA-mediated gene knockdown is diminished if siRNA transfection is delayed relative to transfection of the luciferase plasmid

This experiment was designed to address the question of whether plasmid DNA, following transfection in 3T3 cells, might undergo changes that render it *less* susceptible to the effects of siRNAs. The pGL3 target plasmid and pRL internal control plasmid were transfected at time zero while siRNA-B3 (homologous to region 154-172) was transfected at time-points 0, +1, +24, and +48 hours later for concentrations covering five orders of magnitude (0.1 to 1000ng per well). The *Firefly/Renilla* luciferase activity was then assayed at +24, +25, +48 and +72 hours later (Figure 4.14). At the highest concentration of 1000ng per well siRNA, the results reveal that the level of *Firefly* luciferase knockdown to be 97% at 0 hours, at 95% at 1 hour, at 60% at 24 hours, and 64% at 48 hours, relative to the *Renilla* control. The facile explanation for this observation is simply the persistence of luciferase that is already produced prior to transfection of the siRNA (approximate half-life of *Firefly* luciferase is 6 hours; source Promega™ product literature). The alternative explanation, that the plasmid DNA undergoes qualitative changes that render it *less* susceptible to subsequent siRNA-induced gene knockdown, appears less plausible.

**Figure 4.14: Graph showing how delayed siRNA transfection affects level of gene knockdown.**

When transfection of the siRNA-B3 is delayed relative to that of the pGL3 plasmid carrying the target gene, the level of observed luciferase knockdown is diminished at all siRNA concentrations tested. Data-points are derived from single experiments and hence error bars have not been applied.

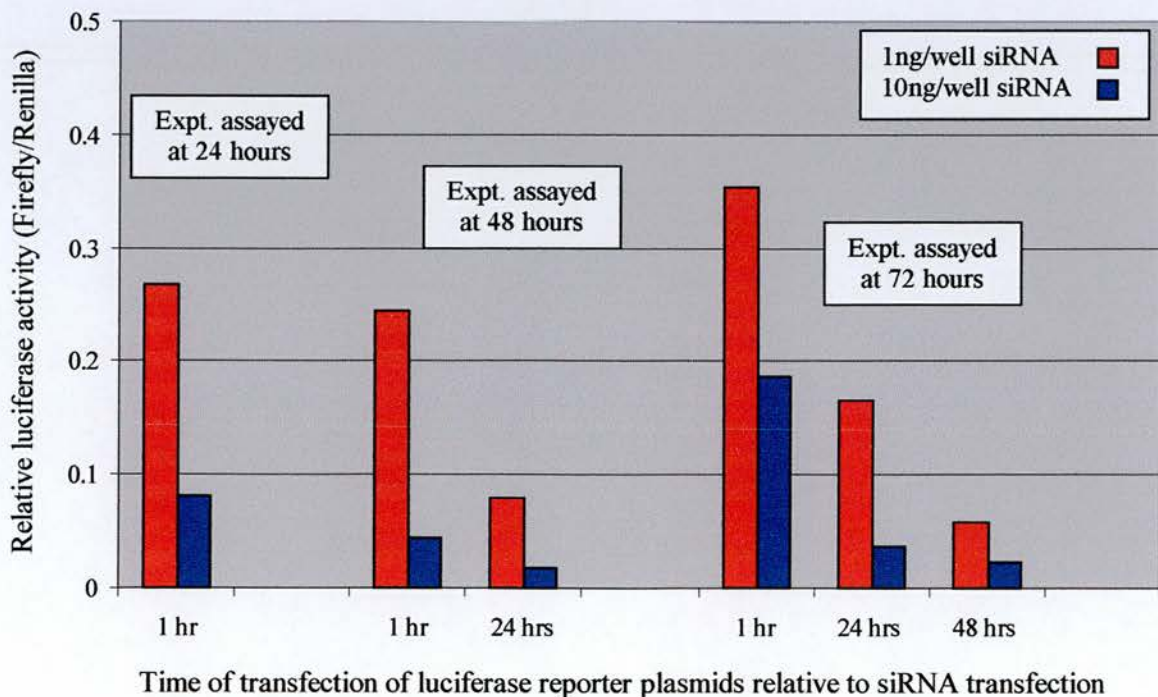


#### 4.12 Premature transfection of the siRNA up to two days prior to transfection of the luciferase plasmid results in robust gene knockdown

This experiment was designed to address the question of whether siRNAs can lead to gene knockdown in 3T3 cells if transfection of the luciferase plasmid is delayed by a period up to two days. siRNAs, at two different concentrations, were transfected at time zero, while the plasmids pGL3 and pRL, were transfected at times +1, +24, and +48 hours. Luciferase assays were performed at +24, +48, and +72 hours to generate a time course knockdown profile (Figure 4.15). The results reveal that the level of knockdown for a given dose of siRNA is approximately the same irrespective of whether the *Firefly* plasmid transfection is delayed by 24 or 48 hours relative to the siRNA. This implies that there is persistence within the cell of either the siRNA molecules themselves or their interference effects over at least this time period of 48 hours (Figure 4.15). Note that the data-points are from single experiments and hence error bars have not been applied and statistical analysis has not been performed. However, the dosage effect and time-dependency provides credibility to this conclusion.

**Figure 4.15: Graph showing luciferase gene knockdown up to 48 hours post siRNA transfection.**

When transfection of the siRNA-B3 precedes that of the pGL3 plasmid carrying the target gene by a period up to 48 hours, specific and dose-dependent knockdown of the Firefly luciferase is still observed. Note that the data-points are from single experiments and hence error bars have not been calculated.

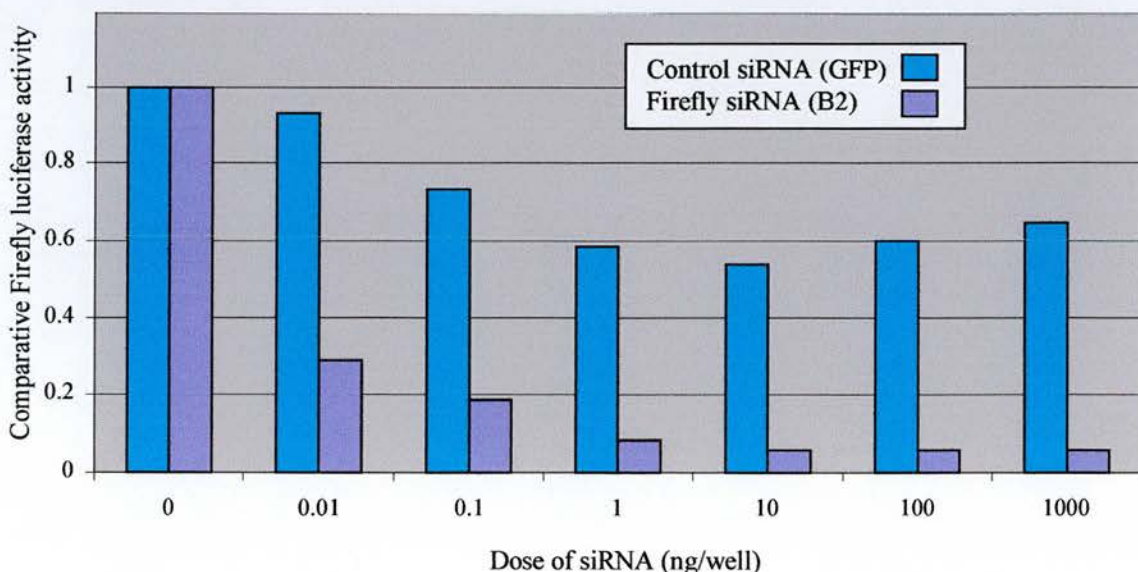


### 4.13 siRNAs can mediate suppression of a stably integrated *Firefly* luciferase gene

The siRNA-B2 that induces potent knockdown of the pGL2 *Firefly* luciferase (homologous to region 154-172) were tested for its ability to suppress expression of the same gene target *stably* integrated in a CHO cell line. The cell line used for this experiment is the CHO-LUC cell line (commercially available from *Clontech Inc.*) that has been generated from the CHO-AA8 tTA-expressing cell line by insertion of the GL2 *Firefly* luciferase gene under the TRE promoter (pRev-TRE-LUC). In the CHO-LUC cell-line, the GL2 luciferase is expressed constitutively (in the absence of doxycycline) and at high level making this a suitable cell line for studying the knockdown effects of GL2-specific siRNAs.

As for the transient transfection experiment, the GL2-specific siRNA-B2 results in specific knockdown of the stably integrated luciferase gene as compared with a non-specific siRNA control (Figure 4.16). This achieves 93% knockdown at the highest concentration tested and is highly statistically significant by Student's t-test ( $p < 0.001$ ). Even ignoring the fact that siRNA delivery to CHO-LUC cells is likely to be sub-maximal, this level of gene knockdown approaches the level seen for the pGL2 plasmid-encoded luciferase in the 3T3 transient transfection experiment. Moreover, neither sense nor antisense RNA strands alone, nor the three-nucleotide mismatched siRNA-B3, produce any significant knockdown whatsoever. This result confirms that the siRNA-induced interference of stably integrated endogenous genes is qualitatively similar, in terms of potency and specificity, to that of transiently expressed plasmid-borne epigenes.

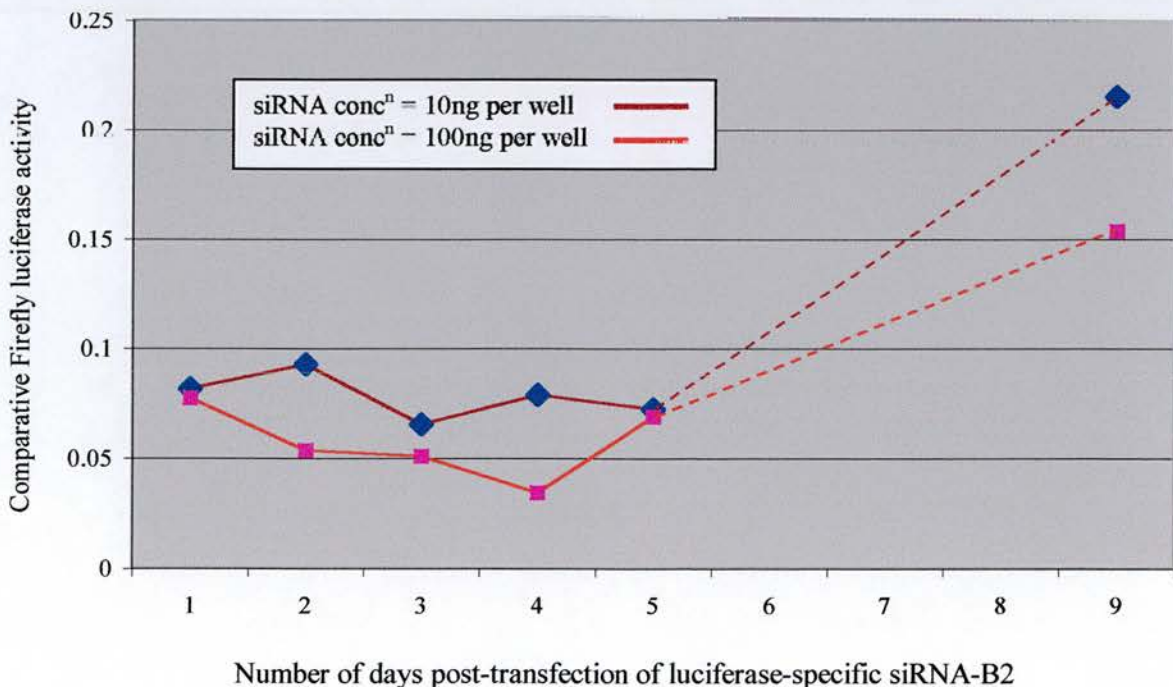
**Figure 4.16: Graph showing siRNA knockdown of an integrated *Firefly* luciferase in CHO-LUC cells.** siRNA-B2 transfection yields potent and specific knockdown of a stably integrated transgene. Data-points are derived from single experiments and hence error bars have not been applied. Comparison of the two sets of data using the paired Student's t-test confirms that this result is statistically significant ( $p < 0.001$ ).



#### 4.14 The interference-related gene knockdown effect persists for at least ten days post-siRNA transfection

Using CHO-LUC cells, it was possible to determine the longevity of the interference effect. Cells transfected with siRNA-B2 (homologous to region 154-172) resulted in over 90% knockdown compared with a non-specific control siRNA until day 6 and around 80% until day 10 (Figure 4.17). This demonstrates the remarkable persistence of the interference phenomenon. Moreover, it should be emphasised that the cells were serially diluted through passaging approximately 1 in 10 every third day. The doubling time of CHO-LUC cells is 16-20 hours and hence it is expected that they would undergo at least ten cell divisions over this period (equivalent to an expansion of  $2^{10}$ , or approximately 1,000 fold). It follows, therefore, that siRNA-induced gene knockdown can persist through ten cell divisions despite dilution of the cytosolic siRNA concentration at each cell doubling. This conclusion ignores the possibility that there exists either an amplification process or an inherited epigenetic process that underpins the propagation of the knockdown effect from cell generation to cell generation.

**Figure 4.17: Graph showing the persistence of siRNA-induced luciferase knockdown.** CHO-LUC cells carrying a stably integrated Firefly luciferase were transfected with siRNA-B2 at two different doses and the level of luciferase knockdown measured thereafter up to ten days. The knockdown is measured comparative to cells transfected with a non-specific control siRNA. Data-points are derived from single experiments and hence error bars have not been applied.



#### 4.15 RNA oligonucleotides in high dose provoke a non-specific transient enhancement of luciferase expression in CHO-LUC cells

In the course of experiments designed to study siRNA-induced knockdown of the GL2 luciferase gene in CHO-LUC cells, an unexpected observation was made. This was the paradoxical *enhancement* of luciferase expression under some circumstances. These included the transfection of *all* RNA oligonucleotides irrespective of whether they were single or double stranded and irrespective of any shared luciferase sequence homology (Figure 4.18). The enhancement was *only* apparent at concentrations in excess of 1µg per well, and was not apparent at lower concentrations (Figure 4.19). Finally, the enhancement was only observed over a time window of 24 to 48 hours, causing up to a 10-fold increase in *Firefly* luciferase activity, after which its magnitude waned and true interference-mediated gene knockdown became apparent (Figure 4.19).

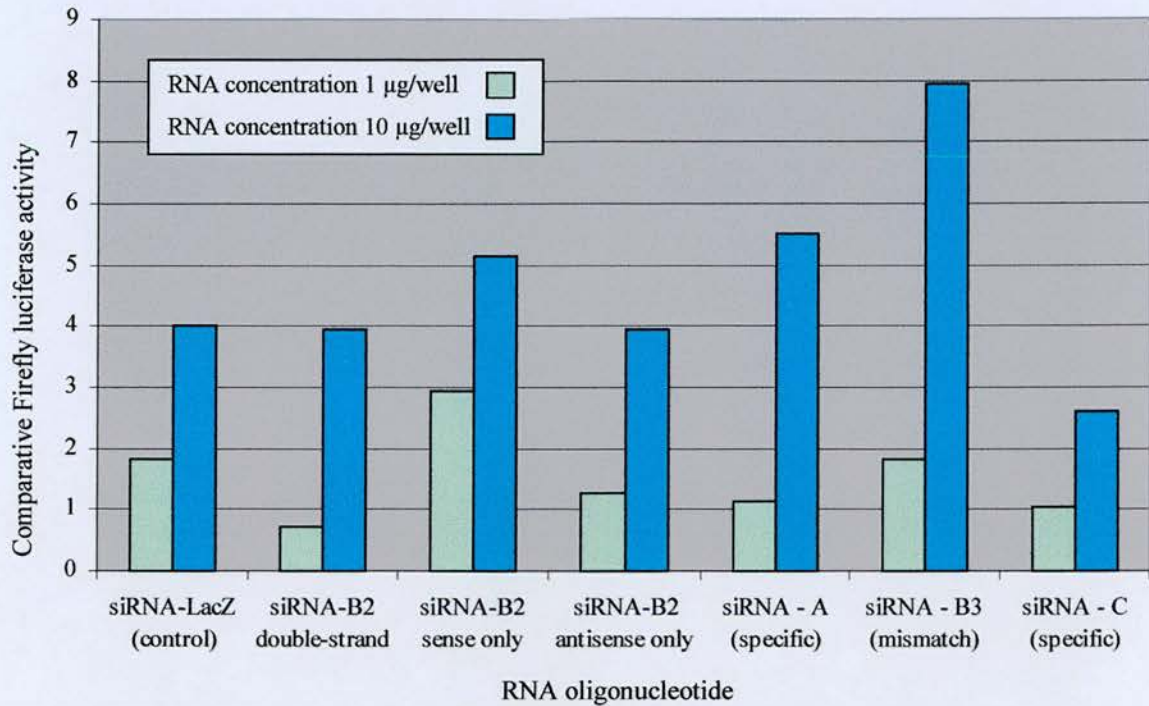
Thus, this phenomenon is qualitatively distinct from the RNAi effect. Firstly, it arises in response to challenge with *all* RNAs, rather than just double-stranded RNA as for genuine interference; secondly, it is completely non-specific, rather than the sequence-specific homology demanded of siRNA-mediated gene knockdown; thirdly, it results in enhancement rather than knockdown; fourthly, it is not apparent at lower concentrations where siRNAs retain their gene-interfering effect; and finally, it results in an effect that is relatively short-lived compared with true RNAi.

These features are more consistent with the well characterised “interferon response”. This view is strengthened by the observation that at high doses of RNA, where significant luciferase enhancement is seen to occur, this is accompanied by widespread cellular death (data not shown), presumably via apoptosis. However, the interferon response is typically associated with translational arrest rather than enhanced gene expression as seen here. A possible explanation of this apparent paradox is the release of oxidative radical species during cellular apoptosis that are capable of accelerating the luciferase reaction used to assay *Firefly* expression. In this case, the apparent gene “enhancement” would simply reflect an artefact arising as a consequence of assaying luciferase activity in apoptotic cells, rather than an actual increase in the expression of the gene. However, this remains untested speculation.

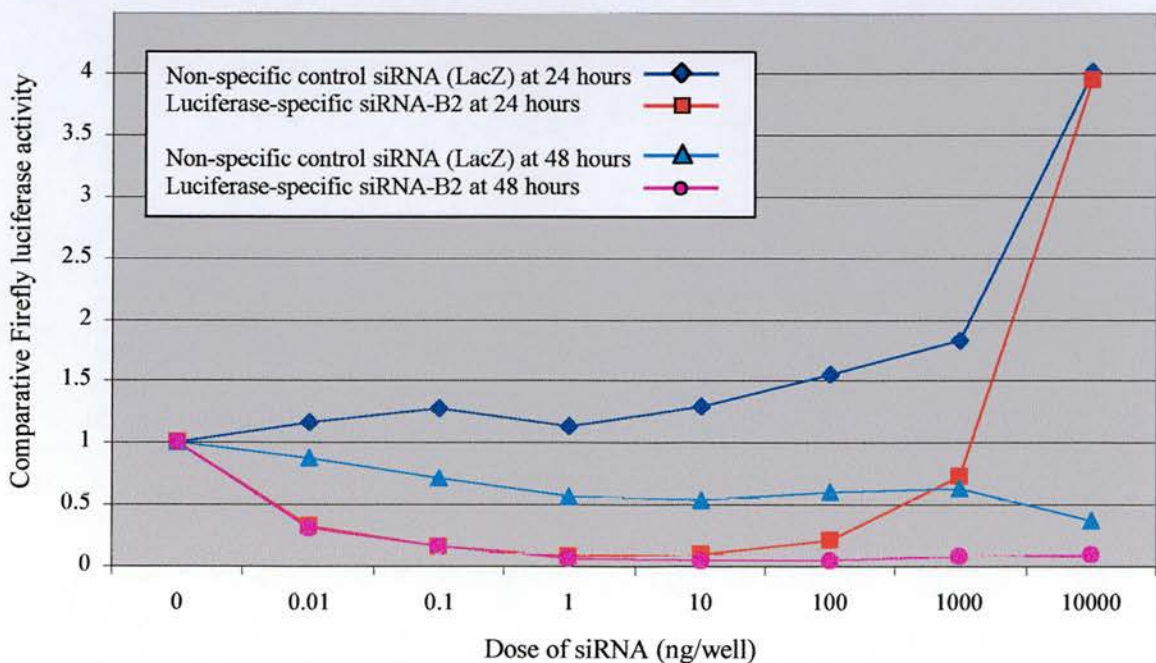
Note that the “enhancement” phenomenon might quite easily mask any specific siRNA-related knockdown effect with a superimposition of both occurring at higher siRNA concentrations. For this reason, all experiments involving siRNA-induced knockdown were performed at concentrations well below the enhancement-trigger threshold (i.e.: 100ng siRNA per well or less) and were assayed at 48 rather than 24 hours post-transfection.

**Figure 4.18: Graph showing non-specific RNA-induced enhancement of luciferase activity.**

CHO-LUC cells transfected with RNA oligonucleotides, whether single or double-stranded, induce enhanced luciferase activity, irrespective of sequence specificity, at concentrations above 1  $\mu\text{g}/\text{well}$ . Knockdown is measured comparative to control cells that had not undergone siRNA transfection. Data-points are derived from single experiments and hence error bars have not been applied. The difference between the two siRNA doses is statistically significant using a paired Student's t-test ( $p < 0.001$ ).

**Figure 4.19: Non-specific enhancement of luciferase activity occurs at 24 but not 48 hours.**

CHO-LUC cells were transfected with either non-specific (LacZ) or luciferase-specific (B2) siRNA. Increased luciferase activity is recorded at 24 hours post-transfection and at doses above 100ng/well. However, by 48 hours post-transfection, and at doses below 100ng/well, RNA interference prevails. The knockdown is measured comparative to control cells that had not undergone siRNA transfection. Data-points are derived from single experiments and hence error bars have not been applied.

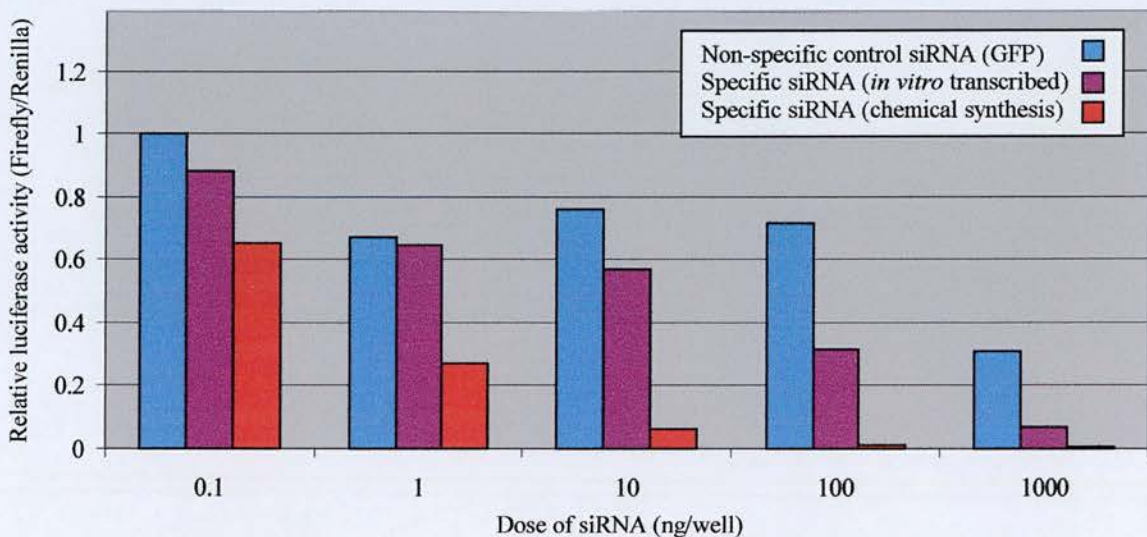


#### 4.16 siRNA molecules generated by *in vitro* transcription exhibit less gene knockdown potency than siRNA molecules generated by chemical synthesis

siRNA molecules generated by *in vitro* transcription, using the Silencer™ siRNA Construction Kit (Ambion Inc.), were compared with identical sequence chemically synthesised siRNAs (Xeragon Inc.). This revealed that the *in vitro* transcribed siRNAs were consistently less potent and more erratic in their ability to knockdown transiently transfected luciferase in 3T3 cells than chemically synthesised siRNAs (Figure 4.20). In CHO-LUC cells, when used at high doses over 1mg/well, enhanced luciferase expression was found similar to that described in Section 4.15 (Figure 4.21).

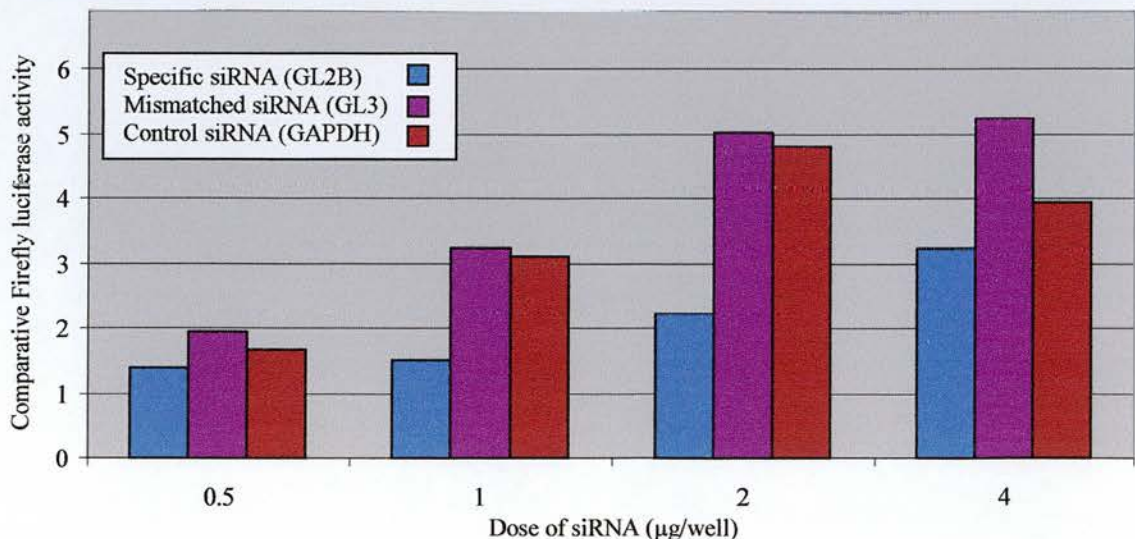
##### **Figure 4.20: *In vitro* transcribed siRNAs are less potent than chemically synthesised siRNAs.**

Knockdown of transient luciferase in 3T3 cells using *in vitro* transcribed siRNAs is up to two orders of magnitude lower than chemically synthesised siRNAs. Data-points are derived from single experiments.



##### **Figure 4.21: *In vitro* transcribed siRNAs yield luciferase enhancement in CHO-LUC cells.**

High doses of *in vitro* transcribed siRNAs result in enhanced luciferase activity in CHO-LUC at 24 hours similar to chemically synthesised siRNAs. Data-points are derived from single experiments.



#### 4.17 Using short basic peptides derived from the transduction domain of the HIV TAT protein to enhance liposomal delivery of siRNA molecules

The improved delivery of siRNA molecules into cells represents one of the major hurdles in bringing the RNAi technology to *in vivo* applications. The latter includes the generation of models of disease in mice and the treatment of disease in humans. The strategy chosen to enhance siRNA delivery was based upon preliminary data generated by X. Nan and colleagues (Molecular Medicine Centre, Edinburgh University) who have used peptide fragments derived from the human immunodeficiency virus transcriptional activator TAT protein, in combination with cationic liposomes, to facilitate the transfer of plasmid-borne reporter genes into cells (Ref: Hyndman L. *et al.*, 2004). These so-called transduction peptides, of which HIV TAT is the best studied, were originally discovered by the serendipitous observation that some proteins are able to enter cells in an apparently passive manner (Ref: Becker-Hapak M. *et al.*, 2001). Such proteins, which include the *Drosophila* homeotic transcription factor Antennapedia (Antp) and the herpes simplex protein VP22, as well as HIV TAT protein, all contain short embedded basic peptide sequences of around 9-16 amino-acids in length that are able to mediate the transduction of the whole protein across the plasma membrane (Ref: Wadia J.S. and Dowdy S.F., 2002). Structurally dissimilar to one another, their common feature appears to be the presence of numerous positively charged lysine and arginine residues, as exemplified by the TAT sequence of nine amino acids (RKKRRQRRR).

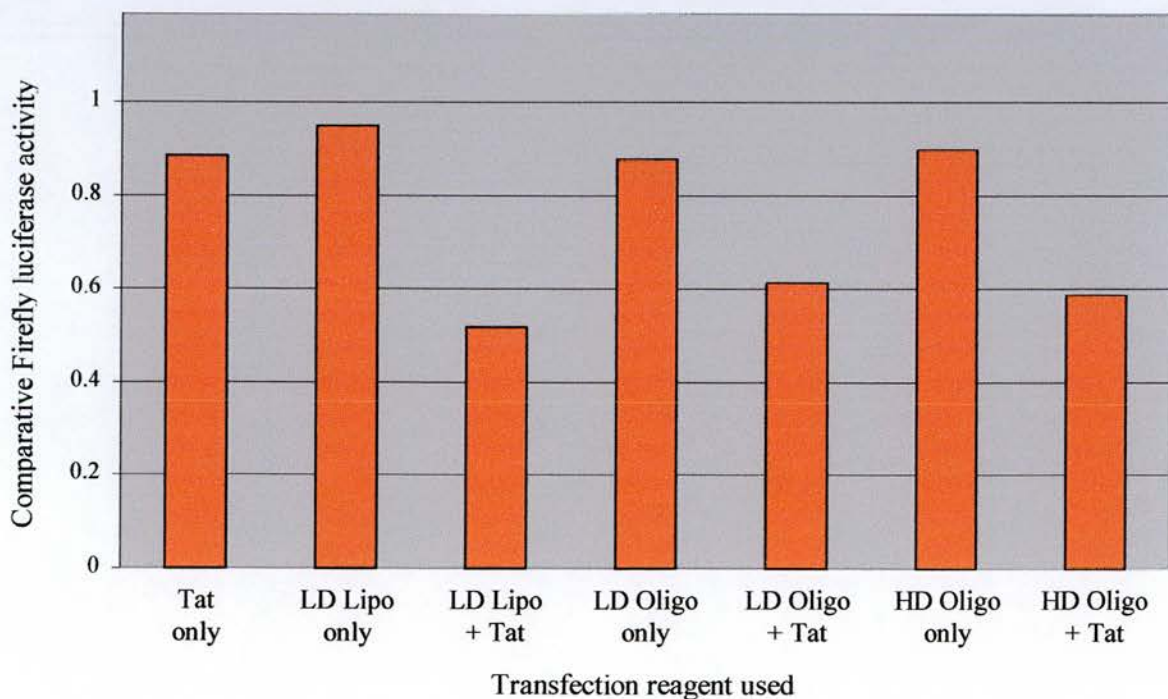
The transduction process, which remains poorly defined, appears to operate via a mechanism that is independent of receptors, transporters, or endocytosis, occurring at 4°C and in the presence of cellular transport process inhibitors (Ref: Schwarze S.R. and Dowdy S.F., 2000). The potency of this mechanism to deliver functional proteins to cells *in vivo* is well illustrated in mice by the demonstration that intraperitoneal injection of a TAT-coupled  $\beta$ -galactosidase reporter results in rapid uptake of the protein throughout all tissues of the animal (Ref: Schwarze S.R. *et al.*, 1999). These results have been extended to therapeutic ends in a mouse model of terminal peritoneal carcinomatosis successfully treated by TAT-mediated delivery of a transducible p53-activating peptide (Ref: Snyder E.L. *et al.*, 2004). In a similar fashion, TAT has also been shown to be able to deliver large DNA molecules into cells (Ref: Snyder E.L., and Dowdy S.F., 2001) and in this evidence, together with the preliminary data kindly provided by X. Nan and colleagues (not shown), that provided the experimental rationale for exploring its use for delivery of siRNAs into cells.

#### 4.18 Tat-derived transduction peptides appear to enhance liposomal-mediated delivery of siRNA molecules into cells

Based on preliminary data showing how a short ten amino-acid peptide sequence derived from the HIV *tat* gene is able to enhance the liposomal-mediated transfection of reporter plasmids (personal communication, X. Nan), experiments were designed to explore this approach for delivery of interfering siRNAs in CHO-LUC cells. siRNA-C, which has homology to region 195-213 of the GL2 luciferase gene, was chosen for this purpose at a dose of 100ng per well that would permit greatest sensitivity of detection of enhanced gene knockdown effect. Luciferase knockdown was compared for low dose liposome (0.15 $\mu$ l Lipofectin™ or Oligofectamine™), with and without 2 $\mu$ g Tat added, and also high dose liposome (1 $\mu$ l Oligofectamine™ only), with and without 2 $\mu$ g Tat added. The results revealed that Tat enhances the transfection efficiency of liposomal the reagents, Lipofectin™ and Oligofectamine™ (Figure 4.22). However, this finding exhibited a variable consistency and requires further methodological refinement. Comparison of Tat+ve and Tat-ve transfection mixes using a paired Student's t-test reveals statistical significance ( $p=0.03$ ).

**Figure 4.22: Tat peptides enhance liposomal delivery of siRNA as assessed by gene knockdown.**

The ability of Tat (2 $\mu$ g/well) to improve delivery of siRNA-C (100ng/well) by either Lipofectin™ (Lipo) at low dose (LD=0.15 $\mu$ g/well) or Oligofectamine™ (Oligo) at low dose (LD=0.15 $\mu$ g/well) and high dose (HD=1 $\mu$ g/well) was tested by assessing knockdown of luciferase in CHO-LUC cells. The knockdown is measured comparative to control cells that had not undergone siRNA transfection. Data-points are derived from single experiments and hence error bars have not been applied. The finding of enhanced knockdown when using Tat is weakly statistically significant using a paired Student's t-test to compare transfection mixes +Tat versus -Tat ( $p=0.03$ ).



## 4.19 Conclusion

The primary objective of this chapter was to characterise the nature of the RNA interference effect in mammalian cells using an experimental system that yielded robust and reproducible results. This system employed siRNA molecules to knockdown a *Firefly* luciferase reporter gene transiently expressed in NIH3T3 cells. A secondary objective was the demonstration that siRNA-mediated knockdown operates in a similar fashion against a stably integrated *Firefly* luciferase in a transgenic CHO cell line.

### **Double-stranded siRNA molecules cause dose-dependent knockdown of target epigenes**

Co-transfection of the *Firefly* luciferase reporter plasmid together with sequence-specific siRNA oligonucleotides in 3T3 cells resulted in the potent and dose-dependent knockdown of luciferase expression. This knockdown effect was dependent upon the presence of *both* sense and antisense RNA strands, although explicit annealing of these strands at high temperature prior to transfection appeared to be unnecessary. The complementary strands were designed to form an RNA duplex, via hybridisation over 19 base pairs with 3' dinucleotide thymidine overhangs, and this was shown to be the physiologically relevant species by the 1:1 stoichiometry of the interference effect. Neither sense nor antisense alone yielded any significant effect.

### **siRNA-induced knockdown requires sequence fidelity with the target gene**

The sequence specificity of the interference effect was demonstrated by the fact that three discrete sequence discrepancies between the siRNA and its target largely abrogated any significant knockdown of the luciferase target. Moreover, sequence homology of the siRNA is not sufficient for effective interference since different siRNA molecules yield considerable variability of knockdown potency. Hence, for the three siRNAs used here to knockdown transiently expressed *Firefly* luciferase, siRNA-B (153-173 bp from start codon) was more potent than siRNA-C (193-213 bp from start codon), which was more potent than siRNA-A (26-46 bp from start codon).

### **The level of gene knockdown is determined by key siRNA design criteria**

Based upon the above findings, together with those from siRNA knockdown experiments in subsequent chapters, it is possible to derive a number of broad conclusions regarding successful siRNA design. These are listed below:

- Site of siRNA sequence homology approximately 150 base pairs from start codon
- Ratio of G/C nucleotides to A/U nucleotides close to unity
- Absence of long runs of G/C base pairings

- Absence of contained palindromic sequences
- Presence of one or more G/C nucleotides at the 5' end of the sense strand
- Presence of one or more A/U nucleotides at the 5' end of the antisense strand

In particular, high G/C content within the siRNA, especially where this generates a palindromic sequence (such as GGCGCC in siRNA-A), which would lead to erroneous sense-sense and antisense-antisense strand pairings, appears to significantly lessen the interference effect. Conversely, the presence of one or more G/C nucleotides at the 5' end of the sense strand together with one or more A/U nucleotides at the 5' end of the antisense strand appears to significantly enhance the interference effect. We can speculate that these generally applicable features of effective siRNA design, which have also been validated in other comparative analyses of siRNA potency (Ref: Holen T. *et al.*, 2002), reflect the importance of suitable annealing/melting properties of the RNA duplex which appear to be different at the two opposite ends of the molecule. It is possible that this reflects aspects of the basic interference mechanism. This requires incorporation of the siRNA molecule into the RISC machinery, unwinding of the duplex siRNA into single RNA strands, and hybridisation of the relevant strand to its target sequence.

#### **The use of siRNA molecules in combination provides useful mechanistic insights**

When siRNA molecules were used in combination with one another, it was shown that the overall knockdown is approximately equivalent to the average of each siRNA used individually. In other words, the interference effect of two or more siRNAs appears to be additive and not multiplicative such that the knockdown potency of a 50:50 mix of weak and strong siRNA is midway between each siRNA used alone. Based upon this assessment, one could envisage that destruction of the target mRNA by two siRNAs might result in cleavage at two different sites, although clearly a single cleavage site on its own will suffice in preventing mRNA translation.

An alternative model of RNA interference, termed “transitive RNAi”, argues that guide siRNAs can serve as primers for RNA-dependent RNA polymerisation of the mRNA template. In this way, a pair of non-overlapping siRNAs could produce a long double-stranded RNA molecule from which new siRNAs are secondarily generated through the action of Dicer. In order to test this possibility in mammalian cells, mismatched non-complementary pairs comprising the sense strand of one siRNA and the antisense strand of another siRNA, were tested and shown to be incapable of any significant knockdown compared with the non-specific control. However, while this is an interesting negative result, it does not rule out transitive RNAi in mammalian cells since the experiment is arguably based upon the wrong model. This is because the RNA strand pairs fail to form duplexes that previous experiments have shown to be necessary for RNA interference.

**Analysis of the temporal and dosage aspects of siRNA-induced gene knockdown**

Analysis of temporal and dosage aspects of the siRNA-induced knockdown effect was performed in order to provide insight into the dynamic and kinetic nature of the interference process. The results of these experiments, and their interpretations, are as below.

Firstly, the time-course of the knockdown effect is commensurate with that of *Firefly* expression such that there is a dose-dependent reduction of luciferase activity even at the earliest moment this is detectable (four hours post-transfection of siRNA/plasmid mix). Thus, in a transient transfection setting, there appears to be no significant lag period, between gene expression and gene interference, required for the initiation or amplification of the siRNA-induced knockdown effect. Rather, the level of observed interference is directly related to the amount of siRNA that is co-transfected with the reporter plasmid.

Secondly, for a fixed amount of siRNA, the effect of increasing the amount of co-transfected plasmid results in an exponential-shaped “knockdown curve” on a linear-log plot that becomes a straight line on a log-log plot that is indicative of first order reaction kinetics. Different amounts of siRNA result in parallel straight lines on the log-log plot consistent with reversible competitive behaviour. In other words, the level of knockdown observed following siRNA/plasmid co-transfection in 3T3 cells is determined solely by the amount of siRNA transfected and can be overcome by adding excess plasmid.

Thirdly, if the transfection of the siRNA is delayed relative to transfection of the plasmid by an interval of 24 to 48 hours, then the level of observed knockdown is reduced by approximately 30-50% over a range of siRNA doses by comparison with that observed for co-transfection. In other words, if the plasmid is transfected a significant while before the siRNA, then the ability of the siRNA to knockdown the target gene is diminished. However, these findings most easily explained in terms of the persistence of luciferase protein translated prior to transfection of the siRNA, thereby limiting interpretation of this result.

Fourthly, in a reversal of the previous experiment, if the transfection of the siRNA precedes transfection of the plasmid by an interval of 24 to 48 hours, then the level of observed knockdown for a given dose of siRNAs is approximately the same as if the siRNA and plasmid were co-transfected. In other words, there is persistence of the siRNA molecules themselves, or their interference effect, within the cell over this time period without significant loss of the potency of the observed knockdown. Thus, RNA interference exhibits intracellular molecular stability over at least 48 hours even in the absence of the gene target.

**siRNA molecules are able to induce knockdown of a stable transgene in CHO cells**

The demonstration of siRNA-induced knockdown of a stably integrated *Firefly* luciferase transgene, using a CHO cell line termed CHO-LUC, confirmed that RNAi operates against an endogenous gene just as for an exogenous gene. The knockdown effect exhibits the same hallmark features of dose-dependent potency and sequence-directed specificity as identified in the transient co-transfection experiments previously, thereby providing the experimental basis whereby siRNA molecules might be used as tools for analysing the function of native genes.

The siRNA-induced knockdown of the *Firefly* luciferase in CHO-LUC cells showed evidence of persistence up to ten days post-transfection. This result demonstrates that the interference effect is able to persist through multiple rounds of cell division. Given that cellular division in this cell line occurs every 16-20 hours and the cells were passaged regularly throughout the experiment, then a conservative approximation for the persistence of the interference effect is ten cell divisions. This is equivalent to a population expansion of  $2^{10}$ , or approximately 1,000 fold. Thus, siRNA-induced gene knockdown can persist despite a halving of the cytosolic concentration of the siRNA at each cell doubling, equivalent to a 1,000-fold overall dilution. This demonstration is most simply explained in terms of the remarkable potency of the siRNA molecules in mediating knockdown of their target gene. However, it is also compatible with two alternative possibilities, namely either the presence of an amplification process or an inherited epigenetic process, both of which would lead to the propagation of the knockdown effect from cell generation to cell generation.

**Excess siRNA induces paradoxical luciferase enhancement and apoptosis of CHO-LUC cells**

In the course of experiments aimed at knocking down expression of stable *Firefly* luciferase in CHO-LUC cells, it was shown that RNA oligonucleotides, under certain circumstances, provoke a paradoxical *enhancement* of luciferase activity, on occasion up to ten-fold. This was observed with *all* RNA oligonucleotides transfected irrespective of whether they were single or double stranded and irrespective of any shared luciferase sequence homology. The enhancement was *only* apparent at the highest doses of RNA in excess of 1 µg per well and not below. Finally, the enhancement was transient, occurring over 24 to 48 hour time window after which its magnitude waned and true interference-mediated gene knockdown became apparent.

Thus, this phenomenon is qualitatively distinct from that of RNA interference. Firstly, it arises in response to *all* RNA oligonucleotides, rather than just double-stranded RNA; secondly, it is completely non-specific, rather than the sequence-specific; thirdly, it results in enhancement rather than knockdown; fourthly, it is not apparent at lower concentrations where siRNAs retain their physiological activity; and finally, it results in an effect that is relatively short-lived.

These features are more consistent with the well characterised “interferon response” that is typical of mammalian cells. Such a view is strengthened additionally by the observation that at high doses of RNA, where significant luciferase enhancement is seen to occur, there is widespread cellular apoptosis (data not shown). However, the interferon response is typically associated with translational arrest rather than enhanced gene expression as seen here. A plausible explanation of this apparent paradox might be the fact that apoptotic death is accompanied by release of oxidative radicals that are capable of accelerating the luciferase reaction used to assay *Firefly* expression. If so, then the apparent gene “enhancement” simply reflects an artefactual result, caused by assaying luciferase activity in apoptotic cells, rather than a true increase in gene expression. However, this explanation remains speculative and further characterisation of the effect would require proper gene expression analysis, such as Southern or Northern blot assays, detection of interferon levels and apoptosis markers, and testing against other reporter gene targets such as GFP.

#### **Chemically synthesised siRNAs are more potent than *in vitro* transcribed siRNAs**

Two technical aspects of siRNA experimentation were addressed in the course of this chapter. The first was a comparative assessment of chemically synthesised siRNAs with molecules having identical nucleotide sequence but generated by *in vitro* transcription. This revealed that the *in vitro* transcribed siRNAs were consistently *less* potent and *more* erratic in their action than their chemically synthesised counterparts, whether used to knockdown transiently expressed luciferase in 3T3 cells or stably integrated luciferase in CHO-LUC cells. Naturally, this finding does not preclude technical shortcomings arising during the synthesis of the *in vitro* transcribed molecules, which were prepared using a commercial kit. However, based upon this result, chemically synthesised siRNAs were used for all subsequent knockdown experiments.

#### **Liposomal delivery of siRNAs into cells can be improved using “transduction peptides”**

The second technical aspect addressed in this chapter pertained to the delivery of siRNA molecules into cells. The reason for exploring this issue is the fact that efficient siRNA delivery represents one of the major hurdles in the translation of RNAi technology to *in vivo* applications. The work was conducted in collaboration with X. Nan and colleagues who have demonstrated that transduction peptides derived from the HIV-1 Tat protein, in combination with cationic liposomes, are able to facilitate the transfer of plasmid-borne reporter genes into cells (Ref: Hyndman L. *et al.*, 2004). This nine amino-acid peptide was tested here for its ability to enhance liposomal delivery of siRNAs using *Firefly*-specific molecules to knockdown luciferase expression in CHO-LUC cells. The siRNA chosen for this purpose was of intermediate potency and used at low dose with knockdown assayed after 48 hours to avoid any spurious enhancement artefact.

The findings of the Tat-liposome experiment revealed that Tat peptide can significantly enhance siRNA delivery by liposomes, as measured by luciferase knockdown, albeit with some experimental variability. Because of this, the method was not used in other knockdown experiments and remains to be more fully characterised. This would require refinement of the methodology and doses. As for the mechanism of enhanced transfection, evidence from plasmid transfection studies suggests that the liposomal reagent acts as a bridging molecule between the oligonucleotide and the transduction peptide (Refs: Torchilin V.P. *et al.*, 2001; personal communication X. Nan). However, this explanation remains speculative and the results presented here in relation to siRNA transfection will require future work to clarify.

# **Chapter 5**

## **Defining molecular aspects of RNA interference**

### **Results 3**

## 5.1 Introduction

The molecular mechanism whereby siRNA molecules can induce knockdown of expression of genes remains to be fully elucidated. Current evidence suggests that RNA interference operates predominantly at a post-transcriptional level by sequence-directed destruction of the mRNA transcript. However, other non-conflicting experimental data supports the conclusion that siRNAs operate also at a transcriptional level that might induce local epigenetic modifications of the gene. This possibility was addressed in this chapter using an elaboration of the transient transfection model developed previously to dissect out aspects of siRNA action. Specifically, experiments were designed to try to answer the following questions:

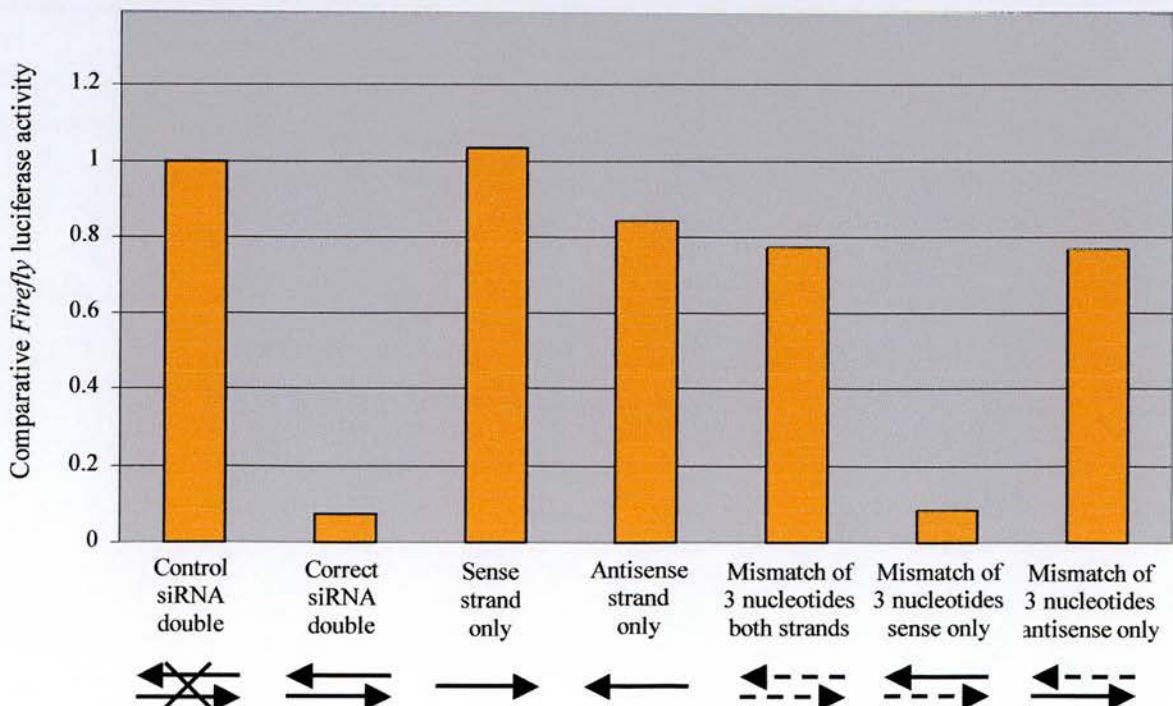
- ❖ Is one RNA strand pre-eminent over the other in directing interference?
- ❖ To what degree are discrete nucleotide mismatches tolerated within either strand?
- ❖ Is siRNA-induced knockdown of a plasmid reporter recapitulated using an mRNA reporter?
- ❖ What is the effect of 5'-terminal phosphorylation of the siRNA molecule on the interference?
- ❖ Are the foregoing findings consistent for stable gene targets as for transiently expressed ones?
- ❖ Are the features of RNA interference in mammalian cells recapitulated in *Drosophila* cells?
- ❖ Are there qualitative inter-species differences (mammalian/*Drosophila*) with regard to RNAi?
- ❖ Are there qualitative intra-species differences between various cell lines with regard to RNAi?
- ❖ Is there evidence that siRNAs act at a transcriptional level as well as post-transcriptionally?
- ❖ If so, is it possible to experimentally dissect one mechanism of siRNA action from the other?

## 5.2 The antisense strand is pre-eminent over the sense strand in directing siRNA-mediated gene knockdown of transiently expressed *Firefly* luciferase

The siRNAs B2 and B3 differ from one another by three distinct nucleotides. Thus, by mixing the sense of one with the antisense of the other, it is possible to generate “chimaeric” heteroduplex siRNAs from B2 and B3 (termed either siRNA-B2/3 or B3/2) that are internally mismatched and therefore impaired in their ability to undergo complementary base-pairing. This situation permits insight into *which* RNA strand is responsible for directing gene knockdown specificity.

The potency of these heteroduplex siRNAs to yield *Firefly* luciferase gene knockdown was tested experimentally using the transient transfection model in 3T3 cells. This reveals that if the sense strand is mismatched but the antisense strand is correct, then there is potent gene knockdown comparable to that of the fully matched siRNA at 94%. If, conversely, the antisense strand is mismatched but the sense strand is correct, then there is minimal gene knockdown at 22%. This represents a statistically significant experimental difference (mean = 0.88, SD = 0.12,  $p < 0.001$ ). Thus, it appears that antisense homology is paramount while sense strand mismatch is largely tolerated (Figure 5.1). This result was corroborated similarly for ES cells (data not shown).

**Figure 5.1: Graph showing *Firefly* luciferase knockdown using the siRNA-B2/B3 heteroduplex.** The knockdown of transiently expressed *Firefly* luciferase in 3T3 cells using 100ng/well heteroduplex siRNAs reveals that antisense-strand fidelity is paramount while sense-strand mismatch is tolerated. Data-points are derived from single experiments and hence error bars have not been applied. This result is statistically significant (mean = 0.88, SD = 0.12,  $p < 0.001$ ).

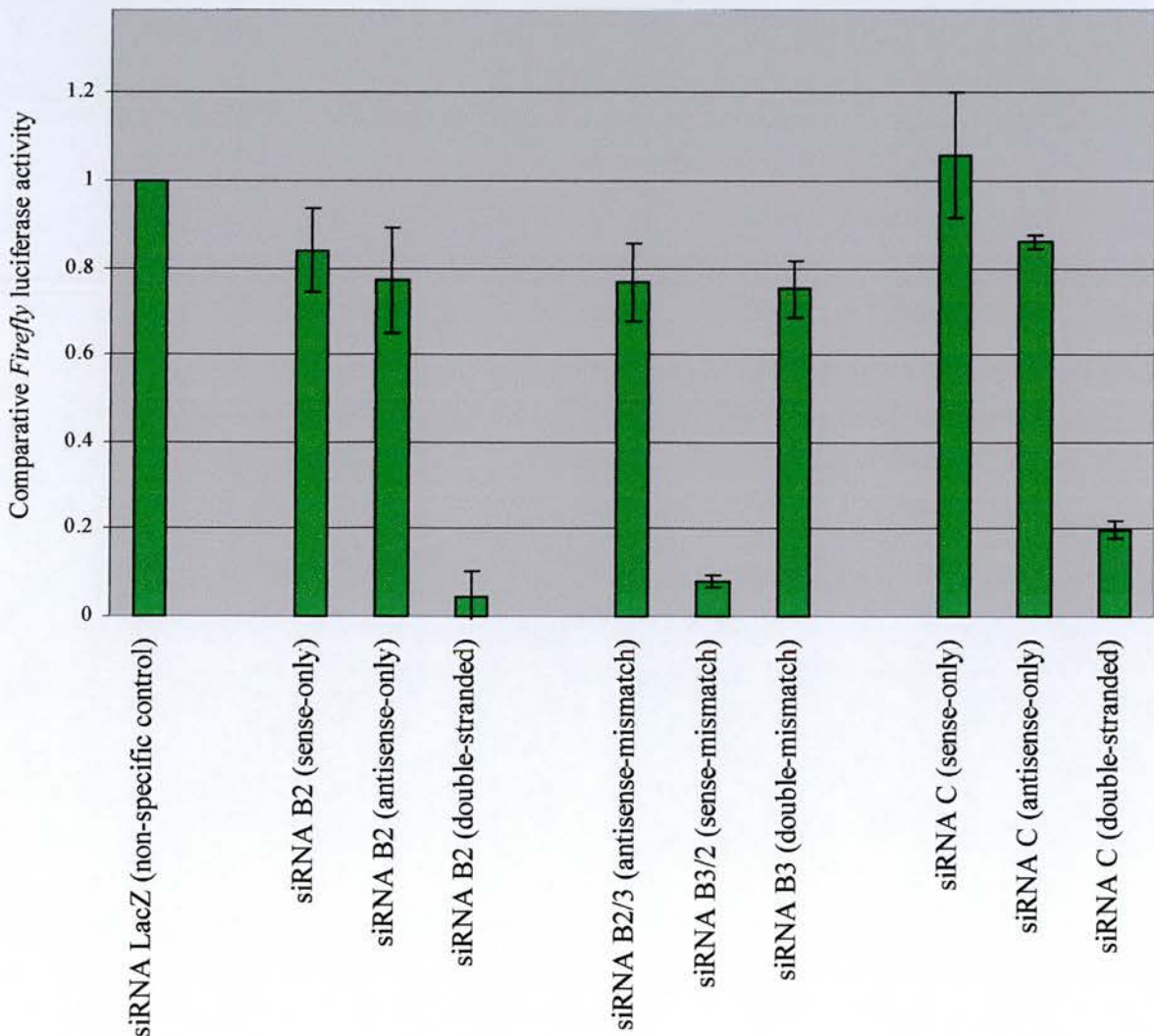


### 5.3 The importance of antisense-strand fidelity for proper siRNA function is recapitulated for knockdown of a stably integrated *Firefly* luciferase gene

The above experiment was performed in a similar fashion using the CHO-LUC cell line that carries a stably integrated *Firefly* luciferase (GL2) gene. Heteroduplex siRNAs, comprising either the sense strand of B2 and antisense strand of B3 (siRNA-B2/3) or vice versa (siRNA-B3/2), were tested for their ability to knockdown expression of the *Firefly* luciferase in CHO-LUC cells alongside a variety of controls (Figure 5.2). As for the transient transfection experiments in 3T3 cells previously, only those heteroduplex siRNAs that retained antisense strand fidelity yielded gene knockdown, even if the sense strand carried a three-nucleotide mismatch. Once again this points to the pre-eminence of the antisense strand in directing the interference process and is a statistically significant finding ( $p < 0.001$ ).

#### **Figure 5.2: Graph showing luciferase knockdown using the siRNA-B2/B3 heteroduplex.**

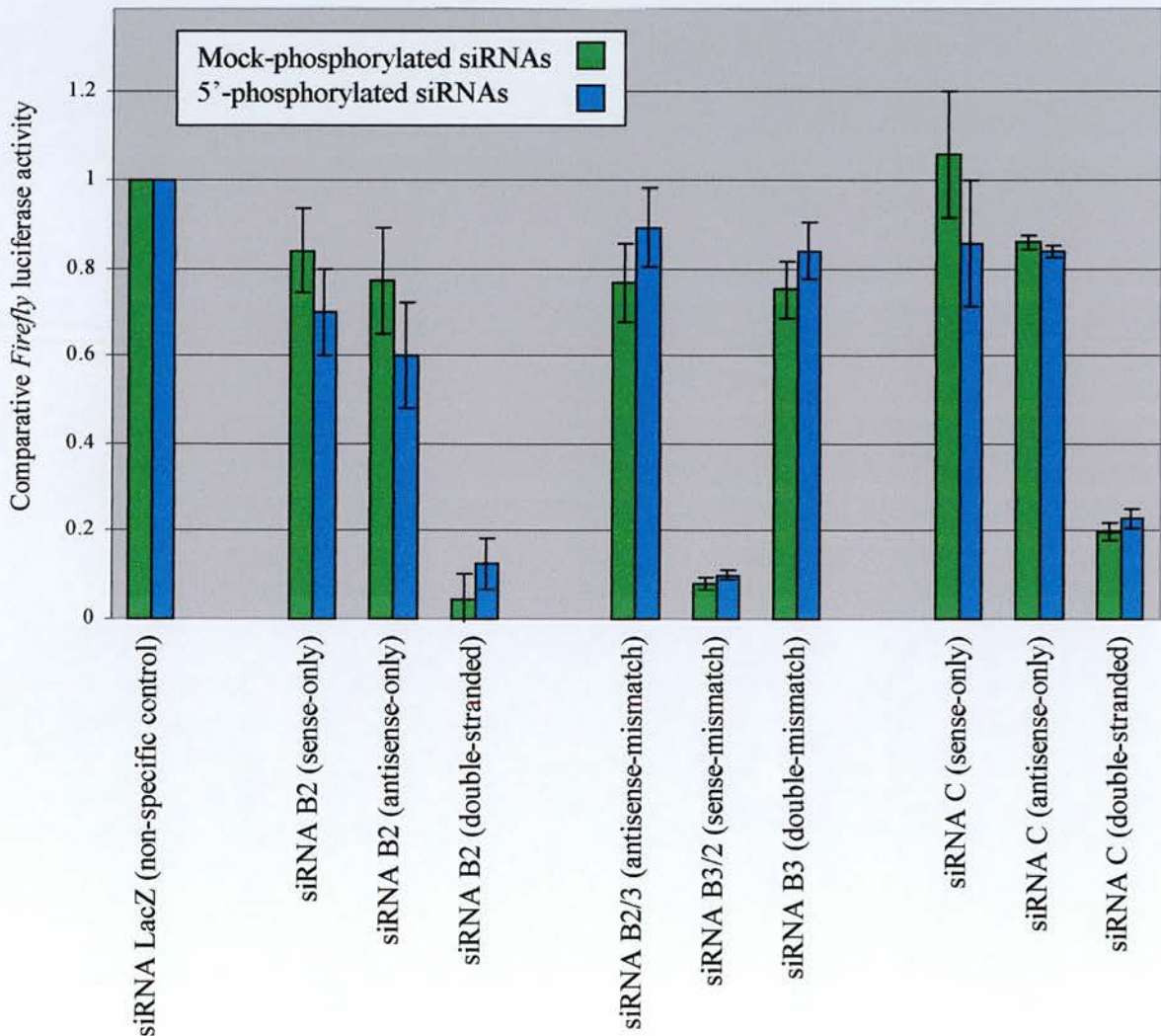
Knockdown of a stably integrated *Firefly* luciferase in CHO cells, using 100ng/well heteroduplex siRNAs, reveals that antisense-strand fidelity is paramount while sense-strand mismatch is tolerated. Error bars represent the range derived from duplicate experiments. The result is statistically significant ( $p < 0.001$ ).



#### 5.4 Phosphorylation of the 5' end of the siRNA molecule does not influence knockdown of a stably integrated luciferase gene

In order to address the question of whether or not phosphorylation at the 5' end of the siRNA molecule might influence the RNA interference process, the previous experiment was performed in duplicate using both mock-phosphorylated and 5'-phosphorylated siRNAs (Figure 5.3). This analysis holds relevance for experiments presented subsequently comparing the interference potential of mock and 5'-phosphorylated siRNAs against transiently expressed *Firefly* luciferase. The results reveal that there is *no* significant difference between the two siRNAs and, thus, exogenous 5'-phosphorylation of the siRNA molecule appears to be unnecessary for the knockdown of a stably integrated *Firefly* luciferase gene in CHO-LUC cells.

**Figure 5.3: Graph showing effect of 5'-phosphorylation on siRNA-induced knockdown.** 5'-phosphorylated siRNAs (100ng/well) show no significant difference in their ability to knockdown a stably integrated *Firefly* luciferase in CHO-LUC cells compared with mock-phosphorylated controls. Error bars represent the range derived from duplicate experiments.



### 5.5 *Firefly*-specific siRNAs fail to knock down luciferase expression following transient transfection of a *Firefly* luciferase mRNA reporter molecule

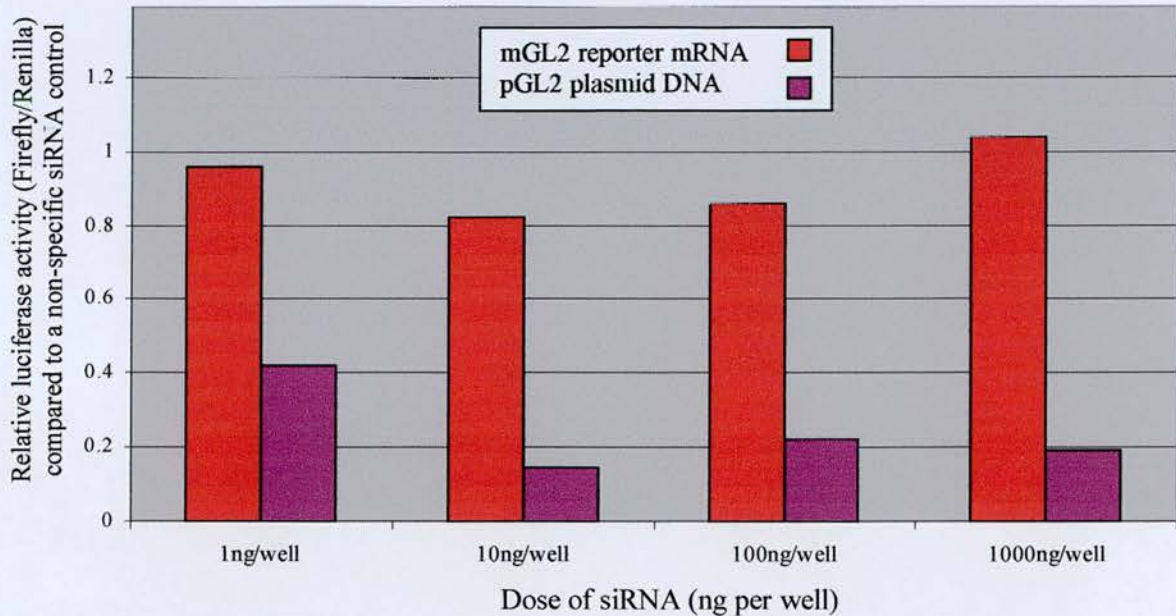
The paradigm view of RNA interference argues in favour of a post-transcriptional mechanism in which targeted binding of the siRNA to the messenger RNA transcript leads to cleavage of the mRNA via an RNA-induced silencing complex (RISC). In order to challenge this hypothesis, the ability of *Firefly*-specific siRNAs to direct knockdown of a messenger RNA luciferase reporter molecule was compared with knockdown of a DNA-encoded luciferase plasmid. The *Firefly* gene sequence was identical for both the mRNA, denoted mGL2, and the plasmid, denoted pGL2, (both products purchased from *Promega Inc.*). The mGL2 reporter is an uncapped mRNA molecule containing a 30-base polyA tail that produces functional luciferase when translated, while the pGL2 plasmid is a circular DNA molecule carrying a luciferase gene sandwiched between SV40 promoter and enhancer regions that control transcription.

All experiments comparing mGL2 *Firefly* knockdown with pGL2 *Firefly* knockdown were performed, as for previous transient transfection experiments, using liposomal-mediated co-transfection of the relevant siRNA together with either the mGL2 (mRNA) or the pGL2 (DNA plasmid) luciferase target. Typically, experiments were internally controlled using the pRL *Renilla* luciferase. Luciferase-specific siRNA molecules were compared with non-specific control siRNAs (typically LacZ or GFP siRNAs) using maximal concentrations of 100ng/well, which is below the threshold at which double-stranded RNA exhibits non-specific influences. The experiments were performed on NIH 3T3 or CHO-AA8 cells and were assayed for luciferase expression using a luminometer within 6-18 hours of transfection (except where otherwise stated).

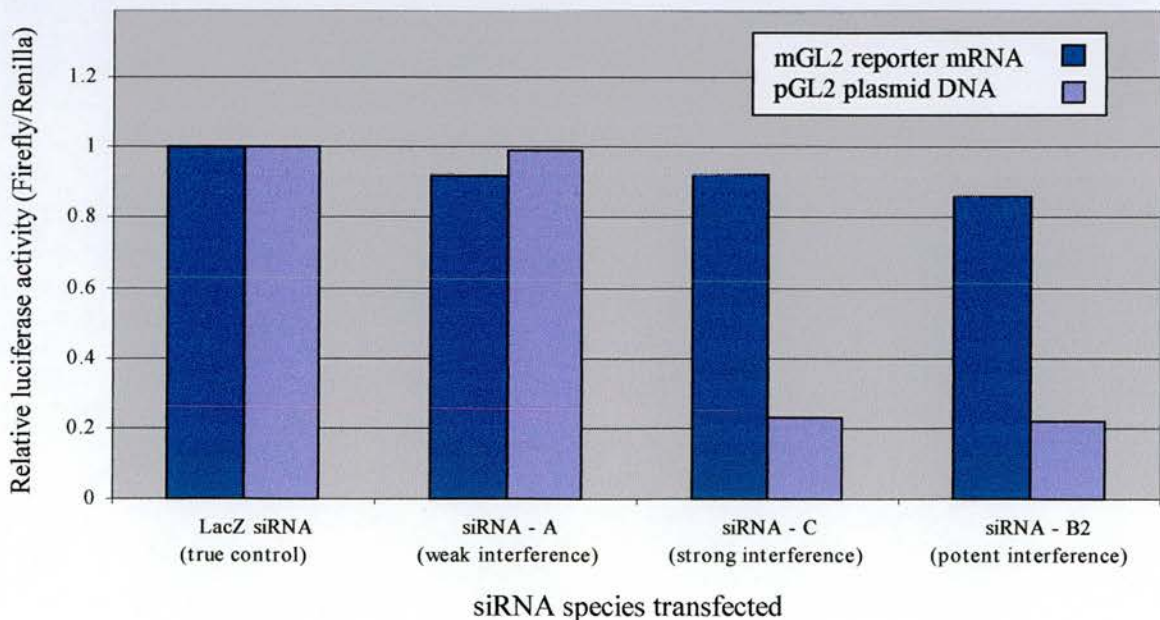
In the initial experiment to compare knockdown of mGL2 luciferase expression with pGL2 luciferase expression using *Firefly*-specific siRNA, the latter was used at a variety of concentrations against a non-specific siRNA control. The findings demonstrate that when luciferase expression is analysed within 18 hours post-transfection, there is *no* significant *Firefly* knockdown when it is encoded as an mRNA reporter while there is robust *Firefly* knockdown when it is encoded as a plasmid reporter. The results, which were internally controlled using *Renilla* expression, are illustrated graphically overleaf (Figure 5.4). Here, the change in luciferase expression resulting from the *Firefly*-specific siRNA at each dose has been calculated relative to the change in expression seen for the non-specific control siRNA at the same dose. This lack of efficacy of the B2 siRNA to knockdown the mGL2 target was also borne out for two other *Firefly*-specific siRNAs, A and C, that have previously been shown to knockdown pGL2 (Figure 5.5). This finding is statistically significant using Student's t-test for paired data ( $p=0.002$ ).

**Figure 5.4: Firefly-specific siRNA-B2 knocks down pGL2 but not mGL2 luciferase.**

The *Firefly*-specific siRNA, B2, is able to knock down expression of luciferase when encoded on a DNA plasmid, pGL2, but not an mRNA reporter, mGL2, when introduced into 3T3 cells by transient transfection. This was apparent for all siRNA concentrations tested with luciferase expression being assayed at 18 hours post-transfection. The quantities used were either 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, together with 100ng/well pRL *Renilla* internal control. Data-points are derived from single experiments and hence error bars have not been applied. This finding is statistically significant using a paired Student's t-test ( $p=0.002$ ).

**Figure 5.5: Firefly siRNAs fail to knock down mGL2 luciferase reporter within 18 hours.**

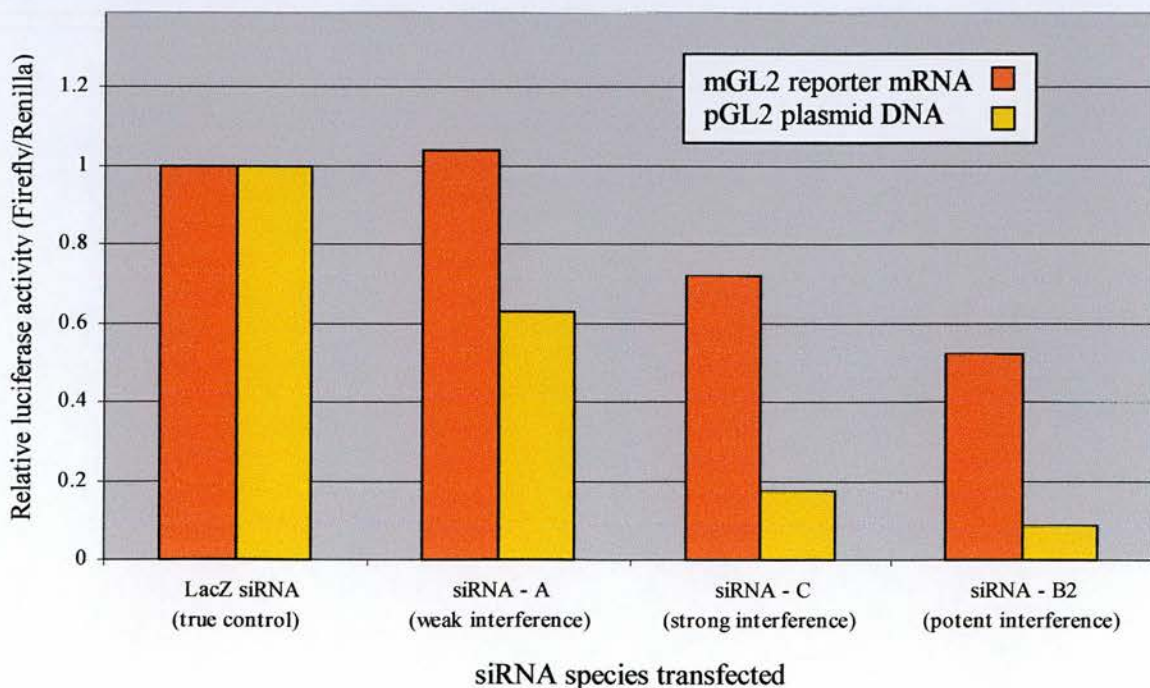
*Firefly*-specific siRNAs generate knockdown of transiently transfected luciferase when it is encoded on a plasmid, pGL2, but fail to generate any significant knockdown when it is encoded by a messenger RNA, mGL2, reporter molecule. This experiment was performed on 3T3 cells using 100ng/well siRNA, either 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL *Renilla* internal control. Luciferase levels were assayed at 18 hours post-transfection. Data-points are derived from single experiments and hence error bars have not been applied.



### 5.6 Evidence suggesting partial siRNA-induced knockdown of mRNA-encoded *Firefly* luciferase after a period of 24 hours post-transfection

When the previous experiment comparing three *Firefly*-specific siRNA molecules (A, C and B2) was assayed for luciferase expression after 24-36 hours post-transfection (rather than 12-18 hours), then a modest knockdown of the mGL2 *Firefly* luciferase is detectable. However, this is quantitatively much smaller than the knockdown of the pGL2 *Firefly* luciferase (Figure 5.6). If correct, then it suggests that the action of the siRNA at the level of the mRNA is a delayed one that is unable to account for the earlier knockdown of the pGL2 luciferase. However, this finding represents the only occasion that siRNA-induced mGL2 knockdown was observed and as such might simply represent experimental variation rather than genuine result. Moreover, the data-points are derived from single experiments and hence error bars and statistical analysis have not been applied. As previously, this experiment was performed on 3T3 cells in 24-well plates using 100ng/well siRNA, 0.5 $\mu$ g mGL2/well or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL *Renilla* control.

**Figure 5.6: *Firefly*-specific siRNAs generate a modest and late knockdown of mGL2 luciferase.** *Firefly*-specific siRNAs generate a modest knockdown of the mGL2 luciferase *only* if luciferase activity is assayed after 24 hours post-transfection. This represents a late effect of the siRNA acting at the level of the mRNA and is unable to account for the much larger knockdown of the pGL2 luciferase detectable within 12 hours post-transfection. As previously, the experiment was performed on 3T3 cells using 100ng/well siRNA, either 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL *Renilla* internal control. Data-points are derived from single experiments and hence error bars have not been applied.

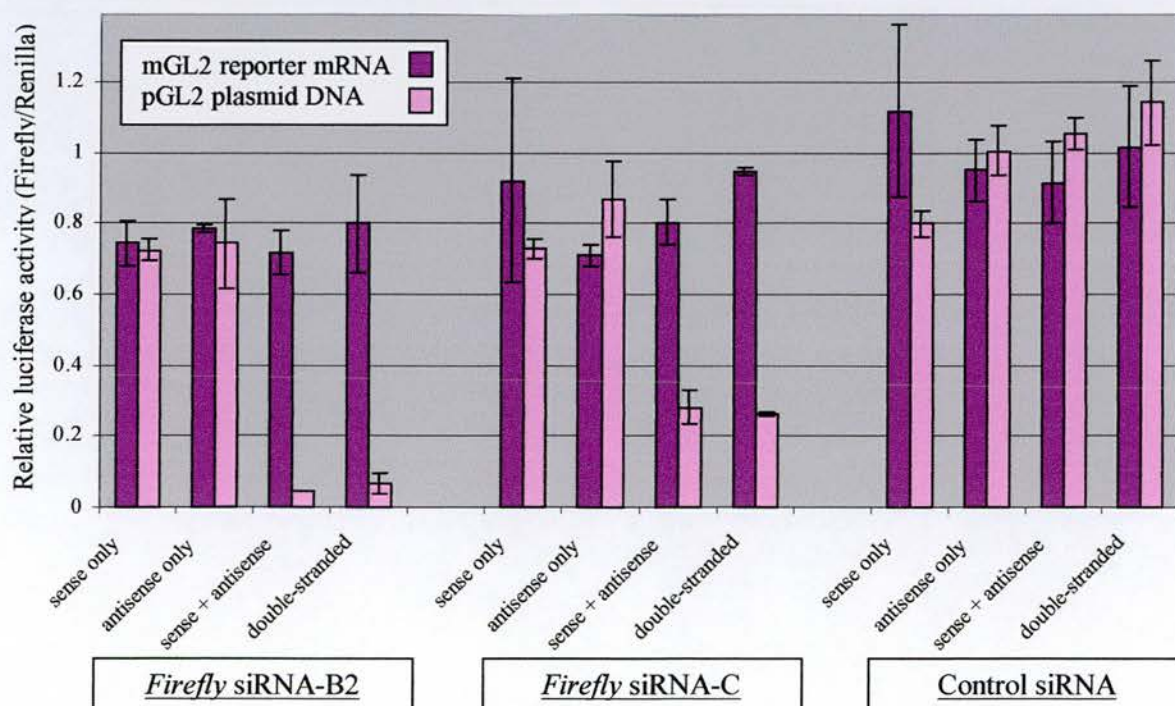


### 5.7 Failure of *Firefly*-specific siRNAs to mediate knockdown of mRNA-encoded luciferase reporter confirmed likewise in CHO cells

Based upon the previous findings in 3T3 cells, revealing an unexpected inability of *Firefly*-specific siRNA to mediate knockdown of an mRNA-encoded luciferase, experiments were undertaken to extend this result to another mammalian cell line, specifically CHO cells. Here, CHO-AA8 cells were co-transfected using either mGL2 reporter or pGL2 plasmid together with the relevant siRNA (B2, C, or non-specific control), and pRL internal control. For each of the three siRNAs, the following combinations were compared: sense alone, antisense alone, sense/antisense equimolar mixture, and double-stranded annealed (Figure 5.7). The results illustrate that for the two *Firefly*-specific siRNAs, B2 and C, but not the control, there is unambiguous knockdown of the pGL2 luciferase for both the sense/antisense mixture and the double-stranded annealed. However, there is no demonstrable knockdown of mGL2 luciferase in the parallel experiment suggesting that siRNA-induced knockdown operates in CHO cells against the DNA-encoded, but not the mRNA-encoded, *Firefly* reporter. This result is statistically significant using Student's t-test for paired data ( $p=0.001$ ).

#### **Figure 5.7: *Firefly*-specific siRNAs fail to knockdown mGL2 luciferase in CHO cells at 24 hours.**

In CHO-AA8 cells, *Firefly*-specific siRNAs yield knockdown of pGL2 luciferase but not mGL2 luciferase (controlled by non-specific siRNAs). This experiment was performed using 100ng/well siRNA, either 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL and luciferase levels were analysed at 24 hours. Error bars represent the range derived from duplicate experiments. This finding is statistically significant using a paired Student's t-test ( $p=0.001$ ).



### 5.8 *Firefly*-specific siRNAs knock down expression of a co-transfected mRNA luciferase reporter in *Drosophila* KC cells but not S2 cells

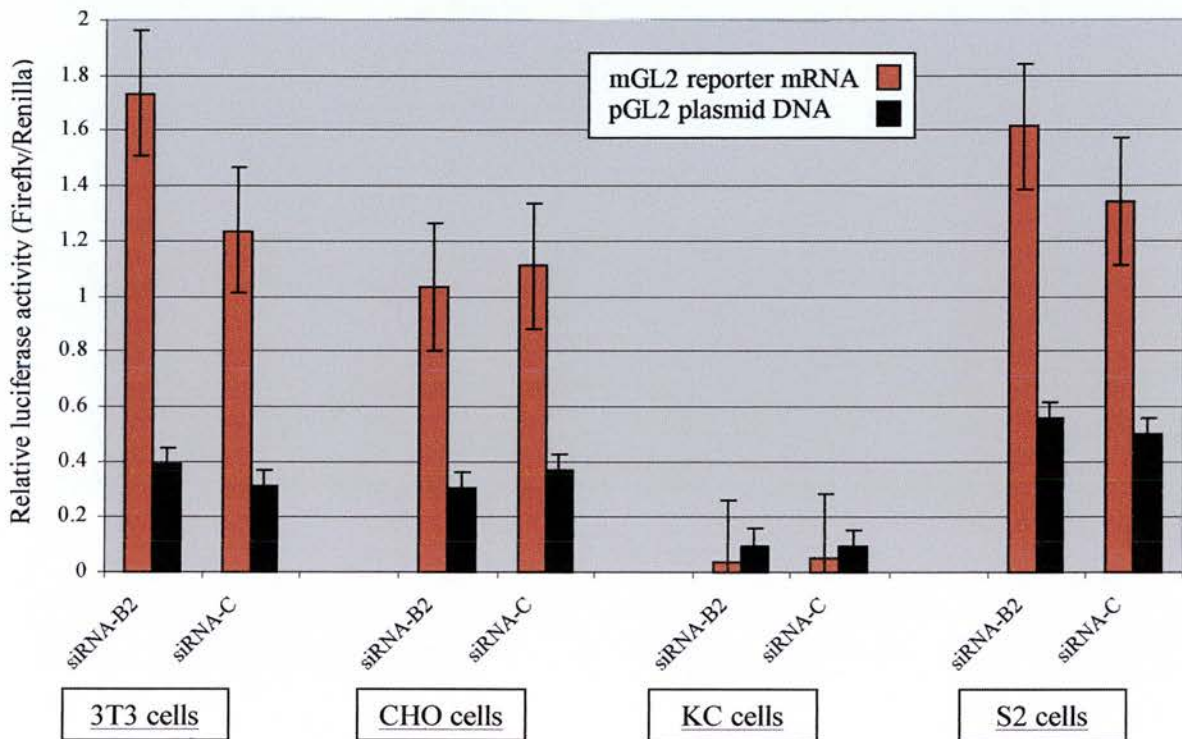
The foregoing results, showing failure of *Firefly*-specific siRNAs to knock down expression of the mGL2 mRNA luciferase reporter, argues against the paradigm view that siRNA-induced interference operates at a post-transcriptional mRNA level in mammalian cells. In order to address this issue, it was necessary to demonstrate that the *Firefly*-specific siRNAs are able to induce knockdown of the mGL2 reporter in a different cell system where a post-transcriptional site-of-action is well established. The *Drosophila* cell lines, KC and S2, were chosen for this purpose. These are immortal cell lines originally derived from embryonic haemocytes that have been well characterised and were kindly provided by B. Baum. KC (Kc<sub>167</sub>) cells are small and round with a high doubling rate and poor plastic adherence while S2 (S2R<sup>+</sup>) cells are large and flat with a lower doubling rate and strong plastic adherence. Both cell lines exhibit robust RNA interference and have been used in a functional RNAi-based genomic analysis of cell morphology using double-stranded RNA mediated knockdown (Ref: Kiger A. *et al.* 2003).

The experiment performed was the same as for the 3T3 and CHO cells previously, namely the direct comparison of *Firefly*-specific siRNAs (B2 and C) to knockdown pGL2 and mGL2 luciferase reporters using transient co-transfection. There are two clear findings:

- 1) In *Drosophila* KC cells, *Firefly*-specific siRNAs are able to knockdown *both* pGL2 and mGL2 luciferase suggesting that the site of action is post-transcriptional ( $p < 0.001$ ).
- 2) However, in *Drosophila* S2 cells, as for mammalian 3T3 and CHO cells, *Firefly*-specific siRNAs are able to knockdown pGL2 but *not* mGL2, luciferase ( $p < 0.001$ ).

Thus, there are qualitative differences not only between *Drosophila* and mammalian cells but also between KC and S2 cells. The levels of statistical significance shown are calculated using Student's t-test for paired data. The results are shown graphically overleaf (Figure 5.8).

**Figure 5.8: Firefly siRNAs knockdown mGL2 luciferase in KC but not S2 *Drosophila* cells.** Firefly-specific siRNAs (B2 and C) knockdown both pGL2 and mGL2 in *Drosophila* KC cells but only pGL2 and not mGL2 in *Drosophila* S2 cells (similar to the pattern of knockdown for mammalian 3T3 and CHO cells). The experiment was performed using 100ng/well siRNA, either 0.5µg/well mGL2 or 2.0µg/well pGL2, and 100ng/well pRL. Luciferase levels were analysed at 24 hours. Error bars represent the range derived from duplicate experiments. This qualitative difference between KC and S2 cells is statistically significant using a paired Student's t-test ( $p < 0.001$ ).



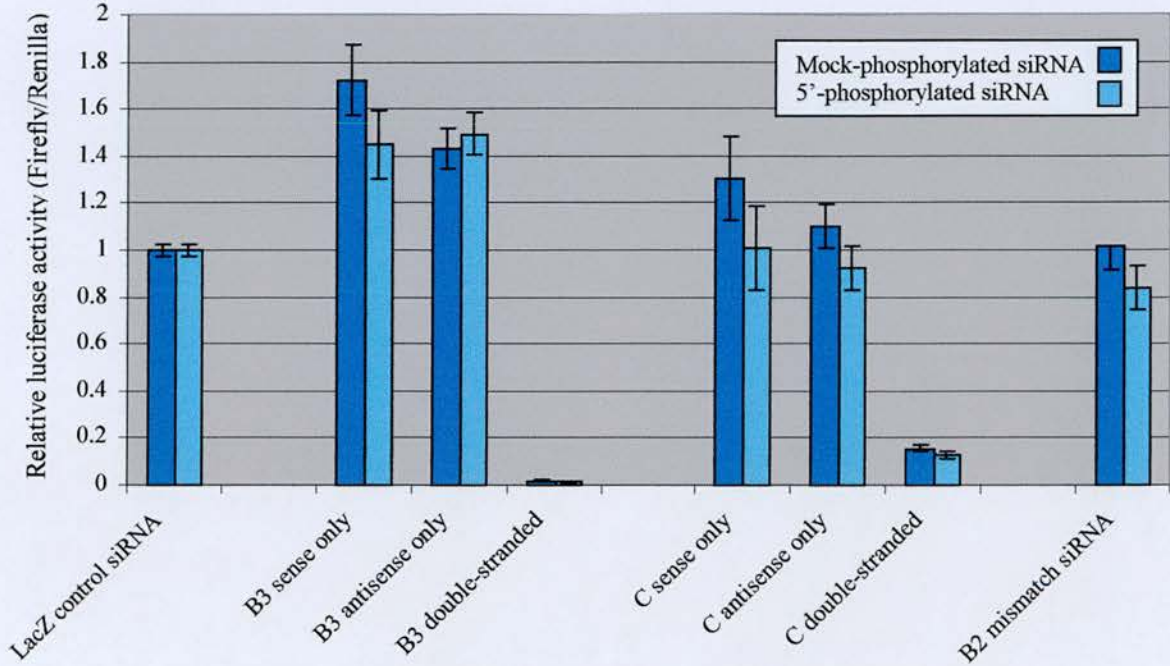
### 5.9 5'-phosphorylation of the siRNA molecule does not affect level of knockdown of *Firefly* luciferase expressed from a transiently transfected pGL3 plasmid

The influence of phosphorylation of the 5' terminus of the siRNA molecule on the level of target gene knockdown was investigated in mammalian (CHO) and *Drosophila* (KC) cells. siRNA molecules were either phosphorylated using the enzyme T4 polynucleotide kinase, which catalyses the transfer of the  $\gamma$ -phosphate from ATP to the 5'-terminus of polynucleotides bearing a 5'-hydroxyl group, or else mock phosphorylated (no enzyme added). The luciferase plasmid pGL3 was used rather than pGL2 on account of its overall greater level of expression in cells. Two *Firefly*-specific siRNAs (B3 and C) were tested for the knockdown potential (sense alone versus antisense alone versus double-stranded) and these were compared with the three-nucleotide mismatch siRNA (B2) and a non-specific control siRNA (LacZ). The experiment was performed identically in both CHO and KC cells using co-transfection of the relevant siRNA together with the pGL3 *Firefly* (target) and pRL *Renilla* (internal control) plasmids. Knockdown was assayed at 36 hours post-transfection in the conventional manner.

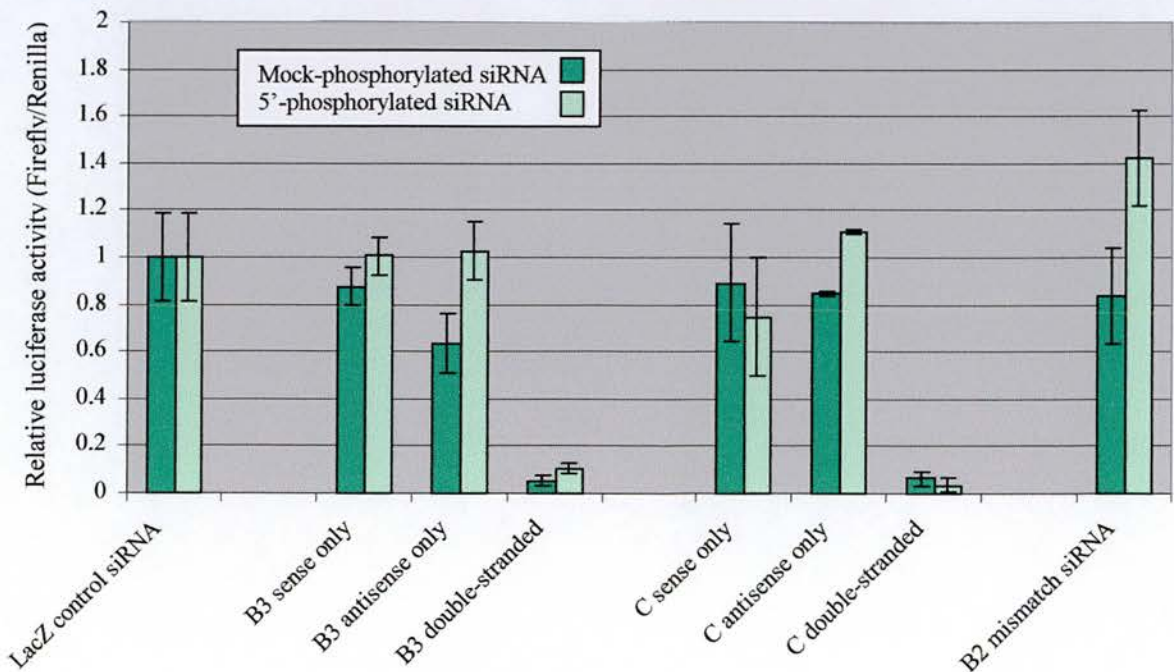
The results confirm the comparable magnitude of siRNA-induced knockdown in KC cells as for CHO cells and also its exquisite sensitivity to contained sequence mismatch. More pertinently, the results demonstrate that phosphorylation of the 5'-terminus of the oligonucleotide appears to have no effect on the potency of the siRNA-induced knockdown, nor on the knockdown inactivity of the single-stranded RNAs. Note however that no test to confirm successful 5'-phosphorylation by T4 polynucleotide kinase was performed. These results are presented for CHO cells and KC cells overleaf (Figures 5.9 and 5.10).

**Figure 5.9: 5'-terminal phosphorylation is irrelevant to siRNA function in CHO cells.**

Phosphorylation of the 5'-terminus of the siRNA molecule has no influence on the function of the siRNA in CHO cells. Both mock and 5'-phosphorylated *Firefly*-specific siRNA molecules yield pGL3 luciferase knockdown while control, mismatch and single-stranded forms are all ineffective. The experiment used co-transfection of 100ng/well siRNA (5'-phosphorylated, or mock), 1.0µg/well pGL3 and 100ng/well pRL, and was analysed at 36 hours post-transfection. Error bars represent the range derived from duplicate experiments.

**Figure 5.10: 5'-terminal phosphorylation is irrelevant to siRNA function in KC cells.**

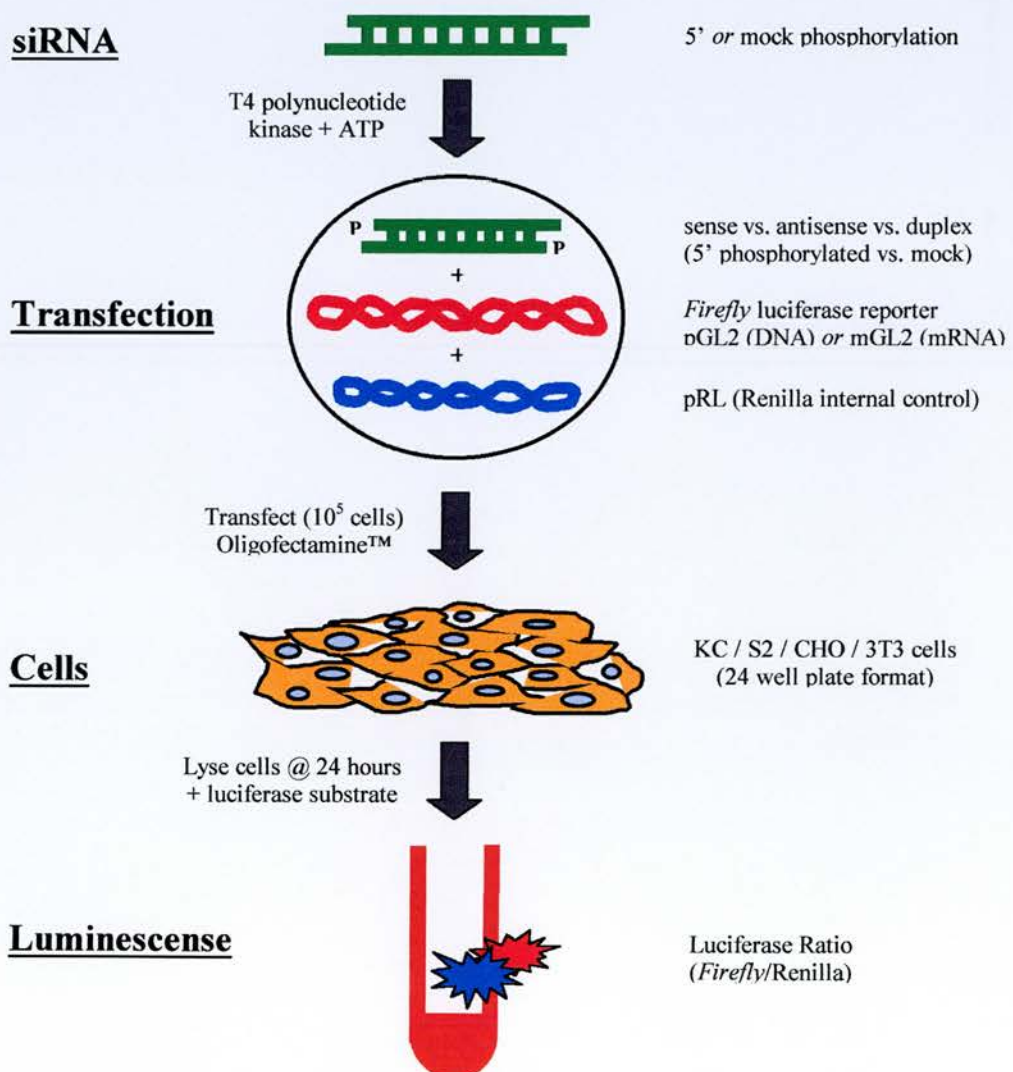
Phosphorylation of the 5'-terminus of the siRNA molecule has no influence on the function of the siRNA in KC cells. Both mock and 5'-phosphorylated *Firefly*-specific siRNA molecules yield pGL3 luciferase knockdown while control, mismatch and single-stranded forms are all ineffective. The experiment used co-transfection of 100ng/well siRNA (5'-phosphorylated, or mock), 1.0µg/well pGL3 and 100ng/well pRL, and was analysed at 36 hours post-transfection. Error bars represent the range derived from duplicate experiments.



### 5.10 Overview of experimental strategy designed to assess the effect of 5'-terminal phosphorylation on siRNA-induced knockdown of transiently transfected mGL2 and pGL2 *Firefly* luciferase reporters in mammalian and *Drosophila* cells

The previous experiment reveals that 5'-terminal phosphorylation of the siRNA molecule apparently has no bearing on knockdown of a plasmid-borne *Firefly* luciferase target, pGL3, in either mammalian CHO cells or *Drosophila* KC cells. To test the hypothesis that this is also true of knockdown of the mRNA luciferase reporter, mGL2, in those cells that fail to show knockdown with non-phosphorylated siRNAs (i.e.: CHO, 3T3 and S2 cells), the following experiment was performed represented schematically below (Figure 5.11).

**Figure 5.11: Schematic representation of experimental strategy to examine siRNA knockdown.** Experimental strategy to explore the effect of 5'-terminal siRNA phosphorylation on knockdown of a target mGL2 or pGL2 *Firefly* luciferase in mammalian (CHO & 3T3) and *Drosophila* (KC & S2) cells.



### 5.11 Evidence suggesting that 5'-phosphorylation of the siRNA termini enhances the knockdown of mGL2, but not pGL2, luciferase in *Drosophila* S2 cells

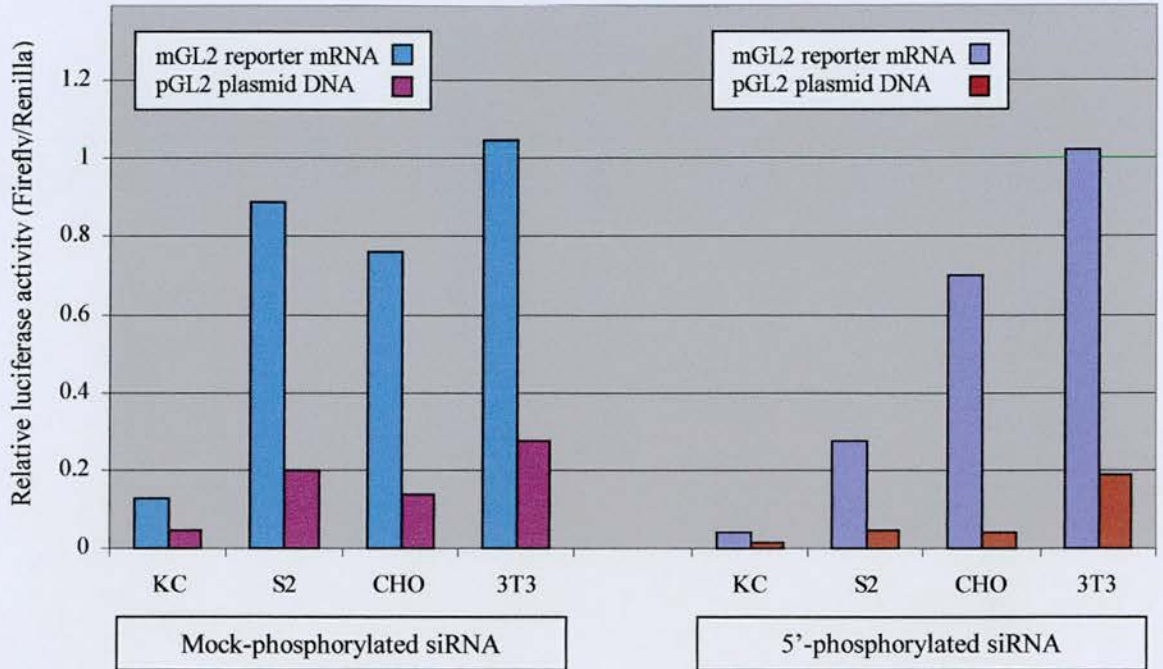
The results of the foregoing experiments reveal that *Firefly*-specific siRNAs are unable to knockdown luciferase expression in *Drosophila* S2 when the gene target is borne on an mRNA molecule despite being able to knockdown the same gene target when it is borne on a plasmid molecule, while in *Drosophila* KC cells, by contrast, *Firefly*-specific siRNAs are able to knockdown *both* mRNA and plasmid-borne luciferase targets. Unexpectedly, when the same experiment is performed using 5'-phosphorylated siRNAs in *Drosophila* S2 cells, there is significant knockdown of the mRNA-borne luciferase suggesting that phosphorylation of the 5'-termini of the siRNA molecule has “rescued” its interference activity. Thus, in S2 cells, siRNA-induced knockdown of an mRNA-borne target is 5'-phosphate dependent while knockdown of a plasmid-borne target is 5'-phosphate independent. This is qualitatively different from KC cells, where knockdown of *both* mRNA and plasmid-borne targets operates in the absence of 5'-phosphorylation. In the case of mammalian CHO and 3T3 cells, there remains failure of knockdown of the mRNA-borne target, despite knockdown of the plasmid-borne target, irrespective of 5'-phosphorylation of the siRNA.

The validity of this result, showing that knockdown of mRNA-encoded luciferase in S2 cells is enhanced by 5'-phosphorylation of the siRNA molecule, is strengthened by the fact that it is consistent for two different *Firefly*-specific siRNA molecules, namely siRNA-B2 (89% to 28%) and siRNA-C (81% to 50%). However, it should be noted that data-points are derived from single experiments and hence error bars and statistical analysis have not been applied. Consequently, these results might simply reflect experimental variation rather than genuine finding. The results are shown overleaf: the upper graph (Figure 5.12) shows the effect of 5'-phosphorylation of siRNA-B2 on mGL2 versus pGL2 knockdown in KC, S2, CHO and 3T3 cells, while the lower graph (Figure 5.13) does likewise for siRNA-C.

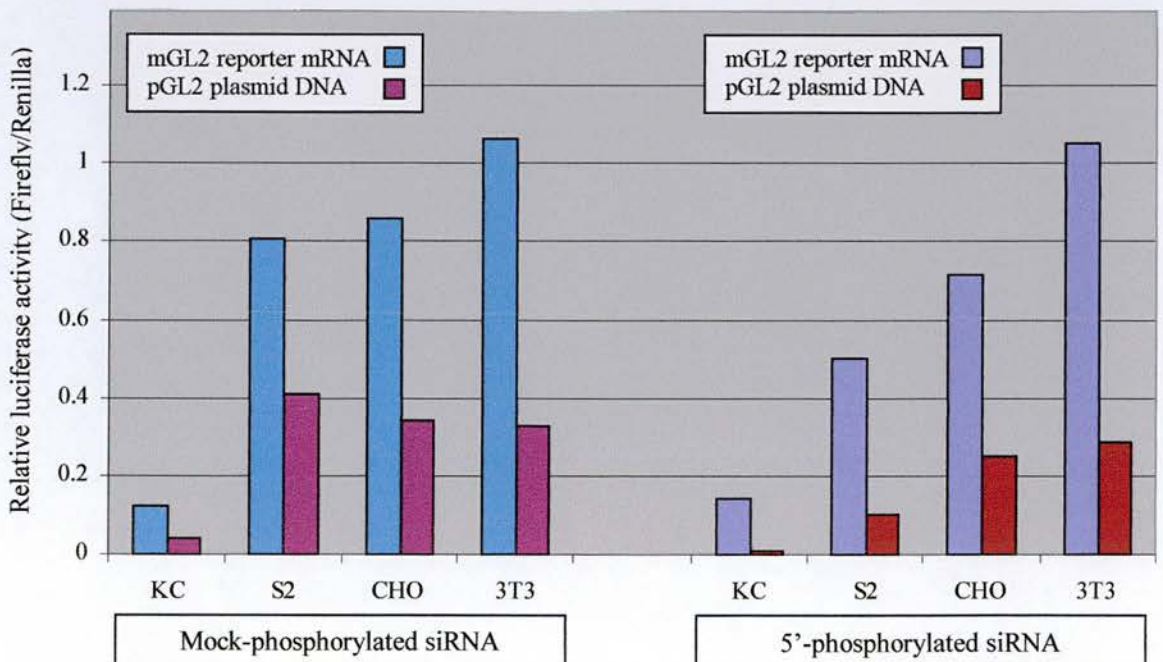
The set of eight graphs on the subsequent page (Figure 5.14) presents the complete set of experimental data. Here, each cell line (KC, S2, CHO and 3T3) has been subjected to luciferase knockdown experiments using a variety of siRNAs (*Firefly*-specific B2 and C, both single-stranded and double-stranded, three-nucleotide mismatched B3, and control LacZ) that has been either mock or 5'-phosphorylated. The results affirm the findings outlined above that 5'-phosphorylation of the double-stranded *Firefly*-specific siRNAs, B2 and C appears to “rescue” the knockdown effect of the mRNA luciferase reporter, mGL2, in *Drosophila* S2 cells (but not mammalian CHO or 3T3 cells).

**Figure 5.12: Effect of 5'-phosphorylation on luciferase knockdown using *Firefly* siRNA-B2.**

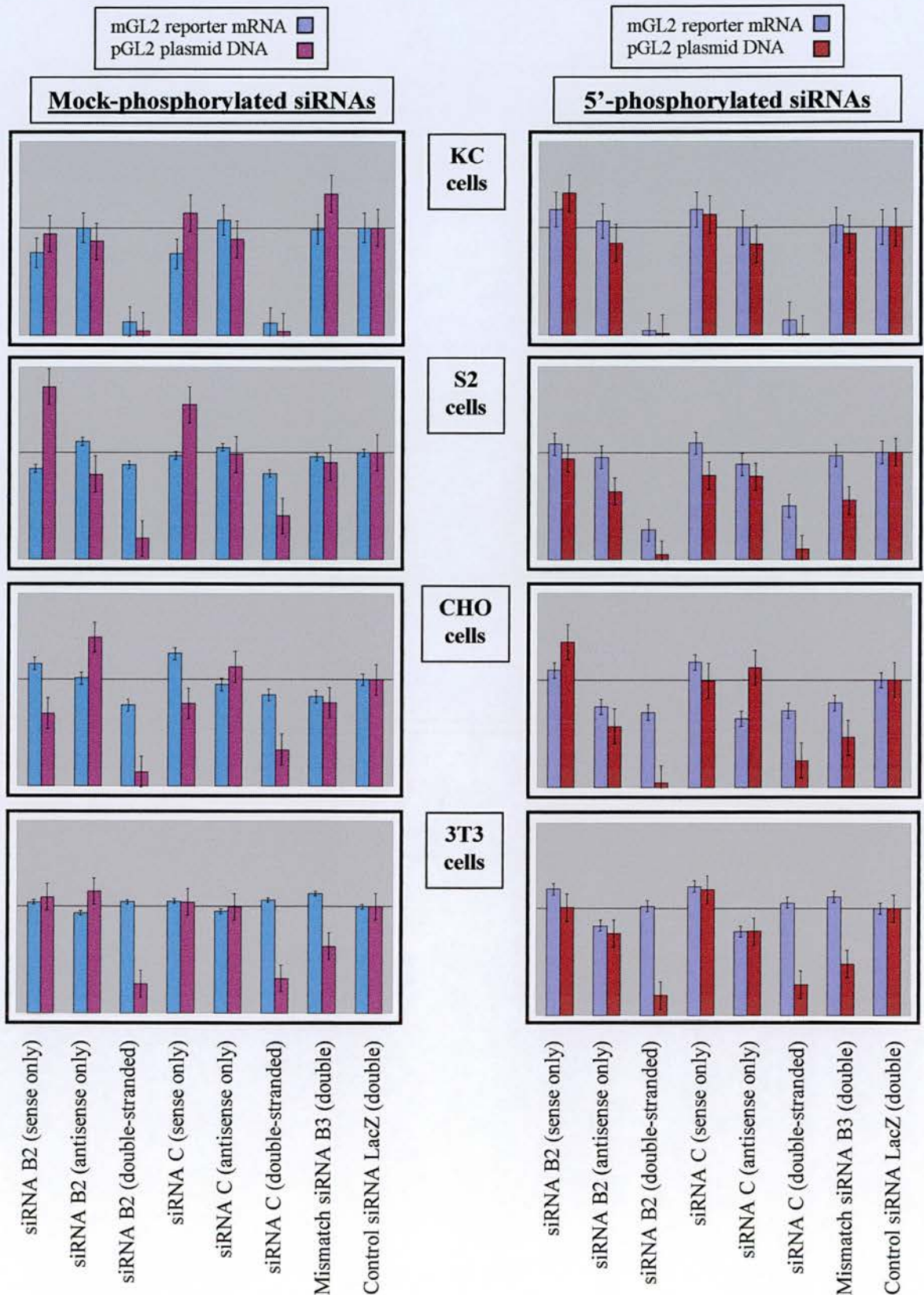
Luciferase knockdown was assessed for both mGL2 and pGL2 using either mock or 5'-phosphorylated siRNA-B2 in *Drosophila* KC and S2 cells and mammalian CHO and 3T3 cells. 5'-phosphorylation "rescues" the siRNA knockdown effect with mGL2 in S2 cells. This experiment was performed using 100ng/well siRNA, 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL internal control.

**Figure 5.13: Effect of 5'-phosphorylation on luciferase knockdown using *Firefly* siRNA-C.**

Luciferase knockdown was assessed for both mGL2 and pGL2 using either mock or 5'-phosphorylated siRNA-C in *Drosophila* KC and S2 cells and mammalian CHO and 3T3 cells. 5'-phosphorylation "rescues" the siRNA knockdown effect with mGL2 in S2 cells. This experiment was performed using 100ng/well siRNA, 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL internal control.



**Figure 5.14: Effect of 5'-phosphorylation on siRNA-induced *Firefly* knockdown in various cells.** Comparison of *Firefly* luciferase knockdown using either mock or 5'-phosphorylated siRNAs in KC, S2, CHO and 3T3 cells. For each graph, two *Firefly*-specific siRNAs, B2 and C, were compared (sense vs. antisense vs. double-stranded) alongside B3 mismatch siRNA and LacZ control siRNA. This experiment was performed using 100ng/well siRNA, 0.5µg/well mGL2 or 2.0µg/well pGL2, and 100ng/well pRL.



### 5.12 5'-phosphorylation of a sense-mismatched heteroduplex siRNA is necessary for knockdown of pGL2 *Firefly* luciferase in *Drosophila*, but not mammalian, cells

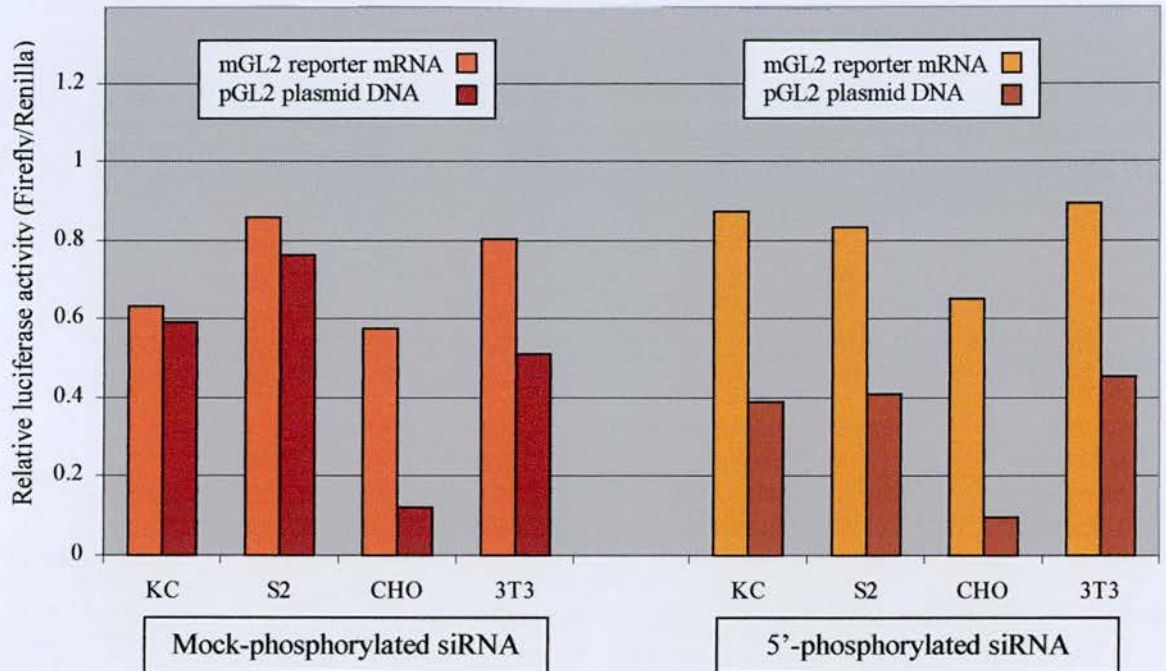
Previous results in mammalian CHO cells have demonstrated that there are profound differences in the ability of a mismatched heteroduplex siRNA to knockdown a target luciferase. Thus, if the heteroduplex is comprised of a matching sense-strand but three-nucleotide mismatched antisense strand, then there is complete abrogation of the knockdown effect. Conversely, if the heteroduplex is comprised of a matching antisense strand but three-nucleotide mismatched sense strand, then the level of knockdown is similar to that measured for the fully matched siRNA duplex. In other words, sequence fidelity of the antisense strand is paramount in determining the effectiveness of siRNA-induced target gene knockdown while minor infidelity of the sense strand is relatively well tolerated. This finding was demonstrated both for transiently transfected pGL3 luciferase target in CHO-AA8 cells (Section 5.1) and also for stably integrated GL2 luciferase gene target in CHO-LUC cells (Section 5.2). Moreover, the phenomenon was apparently independent of 5' terminal phosphorylation of the siRNA (Section 5.3).

When the sense-mismatch heteroduplex was tested against the mGL2 mRNA luciferase reporter, using transient co-transfection in CHO and 3T3 cells, it was not possible to demonstrate any significant knockdown. This demonstration was consistent with previous findings revealing the ineffectiveness of siRNAs against mRNA targets in mammalian cells and contrasts with a clear knockdown of the pGL2 in the parallel experiment. Of greater significance was the demonstration in *Drosophila* KC and S2 cells that the sense-mismatch heteroduplex siRNA fails to result in knockdown of luciferase for *both* pGL2 *and* mGL2. In other words, there is a qualitative difference between *Drosophila* and mammalian cells in regard to their ability to generate knockdown of pGL2 luciferase in response to a sense-mismatch heteroduplex siRNA.

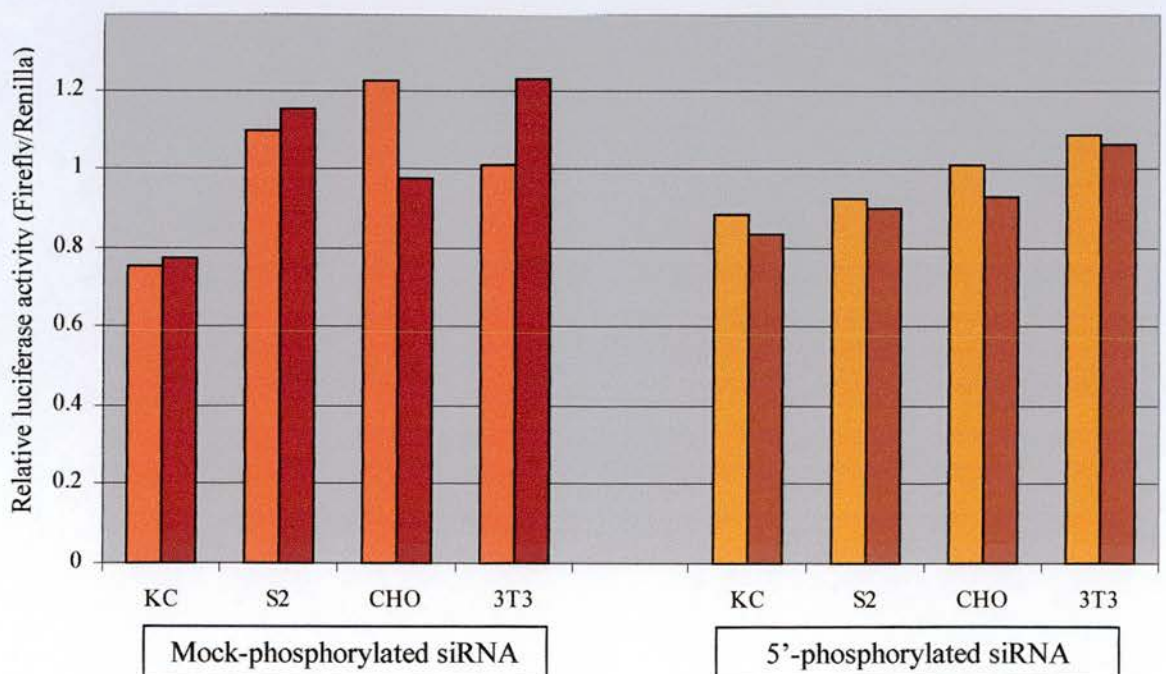
However, if the same sense-mismatch heteroduplex siRNA is 5'-terminally phosphorylated, then there is enhanced luciferase knockdown in *Drosophila* cells, but *only* for the pGL2 luciferase target and not the mGL2 luciferase target. Thus, pGL2 luciferase knockdown using sense-mismatch heteroduplex siRNA is largely 5'-phosphorylation-independent in mammalian CHO and 3T3 cells but largely 5'-phosphorylation-dependent in *Drosophila* KC and S2 cells. However, it should be noted that data-points are derived from single experiments and hence error bars and statistical analysis have not been applied. Consequently, these results might simply reflect experimental variation rather than genuine finding. The results are shown below according to the siRNA heteroduplex (Figures 5.15 and 5.16) and thereafter are compiled according to each cell line used (Figure 5.17).

**Figure 5.15: Effect of 5'-phosphorylation on sense-mismatch heteroduplex siRNA function.**

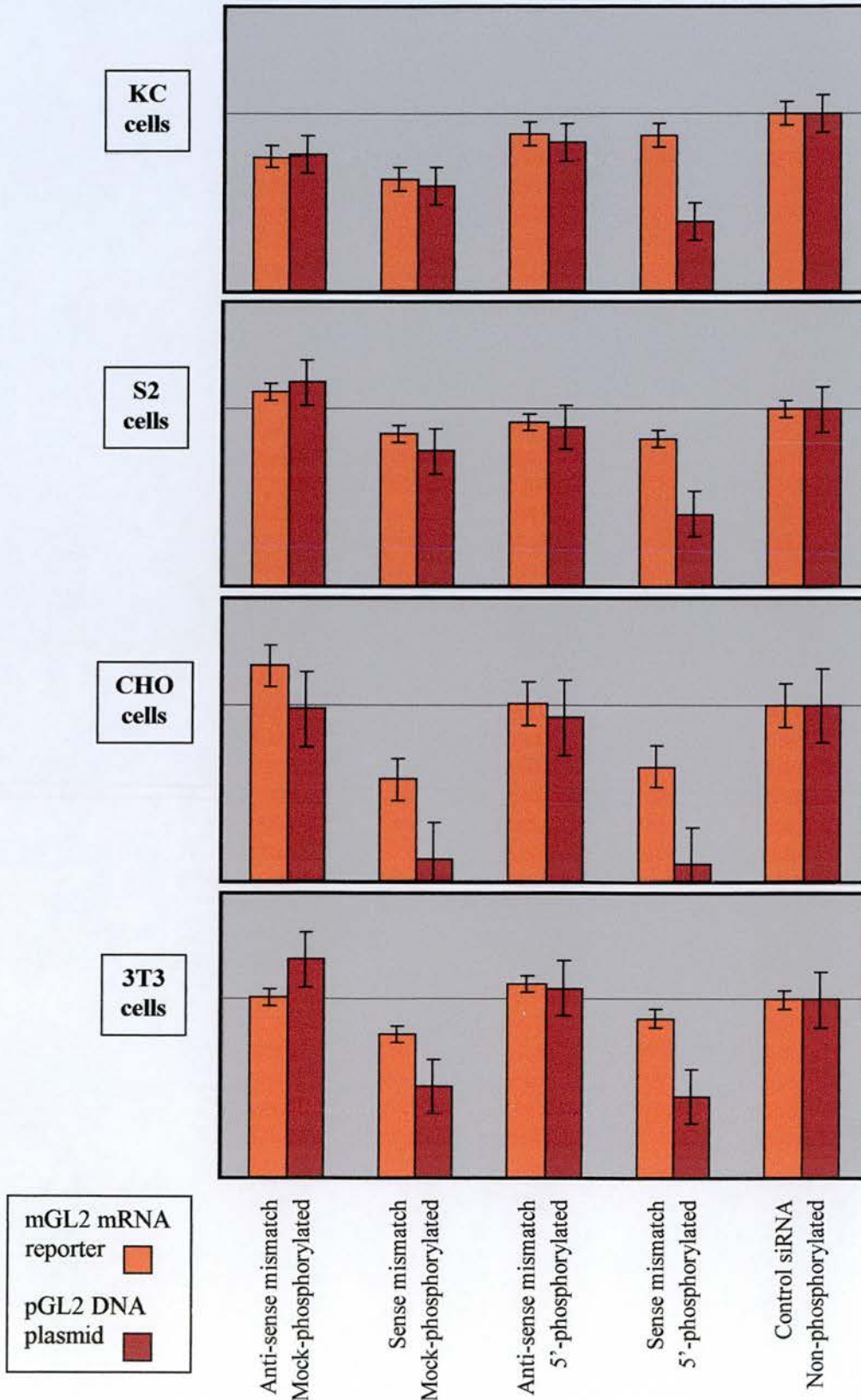
*Firefly* knockdown using the B3/B2 heteroduplex siRNA (sense-mismatch/antisense-correct) occurs in CHO, and to a lesser degree 3T3, mammalian cells but not in KC or S2 *Drosophila* cells unless there is 5'-terminal phosphorylation of the siRNA. This experiment was performed using 100ng/well siRNA, 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL internal control.

**Figure 5.16: Effect of 5'-phosphorylation on antisense-mismatch heteroduplex siRNA function.**

*Firefly* knockdown using the B2/B3 heteroduplex siRNA (antisense-mismatch/sense-correct) is not demonstrable in either mammalian or *Drosophila* cells irrespective of siRNA 5'-terminal phosphorylation (i.e.: there is a total intolerance of antisense-strand mismatch). This experiment was performed using 100ng/well siRNA, 0.5 $\mu$ g/well mGL2 or 2.0 $\mu$ g/well pGL2, and 100ng/well pRL internal control.



**Figure 5.17: Effect of 5'-phosphorylation of heteroduplex siRNAs on *Firefly* luciferase knockdown.** Same data as previous figures presented according to the cell line in which the experiment was performed. The results demonstrate that luciferase knockdown occurs in mammalian CHO and 3T3 cells using sense-mismatch, but not antisense-mismatch, heteroduplex siRNAs. However, in *Drosophila* KC or S2 cells there is no luciferase knockdown using sense-mismatch heteroduplex siRNAs unless it is first 5'-phosphorylated.



### 5.13 Conclusion

The primary objective of this chapter was the experimental dissection of the nature of RNA interference in mammalian cells. This dissection was twofold: firstly, to define the key molecular determinants for effective siRNA-induced knockdown; and secondly, to establish the site of action of the siRNA molecule.

#### **Antisense strand fidelity is obligatory for siRNA action**

In order to determine which strand is pre-eminent in directing effective target knockdown, “chimaeric” heteroduplex siRNA molecules were generated comprising three-nucleotide mismatched sense or antisense strands. When the mismatch is within the antisense strand, there is almost complete abrogation of the knockdown potency of the siRNA against transiently transfected luciferase in 3T3 cells. Conversely, when the mismatch is within the sense strand, then the siRNA retains similar potency to a fully matched siRNA. In other words, siRNA action is intolerant of mismatch in the antisense strand while remarkably tolerant of the corresponding mismatch in the sense strand. Thus, fidelity of the antisense strand of the siRNA molecule is paramount in determining effective siRNA function. The same outcome was observed when the experiment was performed in CHO-LUC cells using siRNA to knockdown a stably integrated luciferase transgene. The clear mechanistic implication of this finding is that it is the antisense strand that directs recognition of the target sequence probably through complementary base pairing with the mRNA or, possibly, with the non-coding DNA.

#### **5'-phosphorylation of the siRNA is irrelevant in mammalian cells**

In the foregoing experiment using CHO cells, it was shown that phosphorylation of the 5'-ends of the siRNA molecule prior to transfection (using polynucleotide kinase) had no influence on the observed pattern or potency of the knockdown. Thus, either the presence of a terminal 5'-phosphate is physiologically irrelevant for effective siRNA action or alternatively the cell contains its own endogenous 5'- nucleotide kinase activity able to perform this function post-transfection. This finding has relevance to results derived from subsequent experiments.

#### **siRNA knockdown of mRNA luciferase is weak or absent in 3T3 cells**

The paradigm view of RNA interference argues in favour of a post-transcriptional mechanism in which targeted binding of the siRNA to the mRNA transcript leads to cleavage of the latter via an RNA-induced silencing complex (RISC). A direct prediction from this theory is that active siRNAs should be able to knockdown a target gene irrespective of whether that gene is encoded by a DNA molecule or an mRNA molecule. However, when this prediction was tested in 3T3

cells, *Firefly*-specific siRNAs that could induce knock down of a plasmid-borne luciferase gene failed to do likewise if the same luciferase gene was encoded on an mRNA molecule. This was the case for two different *Firefly*-specific siRNAs and over all concentrations tested. Only if the luciferase activity is assayed after 24 hours post-transfection was there any detectable knockdown of mRNA-encoded luciferase but this was consistently much less than the knockdown of the plasmid-borne luciferase. Thus, in 3T3 cells, siRNA-induced knockdown of mRNA luciferase appears to be qualitatively distinct from that of DNA luciferase being at best rather weak and delayed in its effect.

#### **siRNAs show knockdown of DNA luciferase but not mRNA luciferase in CHO cells**

The same result was demonstrated likewise in CHO cells except that here it proved impossible to detect *any* significant knockdown of the mRNA luciferase despite unequivocal knockdown of the DNA luciferase for two different siRNAs tested. This result suggests that the siRNA-induced knockdown measured in all the transient transfection experiments operates at the level of the DNA and not the mRNA. Such an interpretation of the results is clearly at odds with the prevailing view that siRNA operates primarily at a post-transcriptional level. However, it is possible that alternative explanations could account for these results including experimental variation or artefact, problems with the time-course of mRNA transfection and expression, and the non-physiological nature of the experimental approach.

#### **Reasons for failure of siRNA-induced knockdown of mRNA luciferase in mammalian cells**

It is possible to contrive alternative explanations for these unexpected findings. Firstly, if the mRNA was rapidly translated into luciferase prior to incorporation of the siRNA into the silencing complex, then clearly no knockdown would be detectable. However, luciferase activity was apparent for at least 36 hours suggesting that mRNA translation must itself take place over several hours allowing time for the siRNAs to exert their effect. Secondly, the siRNAs might operate in a cellular compartment, such as the nucleus, that the exogenously introduced mRNA (but not the plasmid DNA) is unable to access. However, this would be contrary to the established view that interference takes place in the cytoplasm. Finally, the mRNA molecule used here, which lacks a 5'-cap but possesses a 30-nucleotide polyA tail, has not been generated via the normal physiological route and may therefore bypass the RNA interference machinery.

#### **KC and S2 cells are *Drosophila* haemocyte lines known to exhibit robust RNA interference**

In order to determine if these findings were specific to mammalian cells, and in so doing exclude alternative explanations suggested above, it was necessary to demonstrate that transiently transfected mRNA luciferase could be successfully knocked down using *Firefly* specific siRNAs

in non-mammalian cells instead. Two well-characterised *Drosophila* cell lines (kindly provided by B. Baum) were chosen for this purpose, both derived by immortalisation of embryonic haemocytes. KC (Kc<sub>167</sub>) cells are small and round with a high doubling rate and poor plastic adherence while S2 (S2R<sup>+</sup>) cells are large and flat with a lower doubling rate and strong plastic adherence. Both cell lines exhibit robust RNA interference and have been used in a functional RNAi-based genomic analysis of cell morphology using double-stranded RNA mediated knockdown (Ref: Kiger A. *et al.* 2003).

#### **siRNAs show knockdown of mRNA luciferase in KC cells but not S2 cells**

The findings reveal an unexpected but clear qualitative difference between KC and S2 cells *Drosophila* cells. Thus, in KC cells, *Firefly*-specific siRNAs are able to knockdown *both* DNA luciferase and RNA luciferase in-keeping with the notion that the interference of gene expression is operating post-transcriptionally at the level of the mRNA. However, in S2 cells, *Firefly*-specific siRNAs are only able to knockdown DNA luciferase and *not* RNA luciferase, which is the pattern of knockdown seen in the mammalian cell lines, 3T3 and CHO, previously. Moreover, these findings were corroborated for two different siRNA molecules that recognise different target sequences within the luciferase gene. The demonstration that *Firefly*-specific siRNAs do indeed knockdown RNA luciferase in KC cells confirms that the experimental design is robust and that their failure to do so in mammalian cells is not attributable to trivial technical reasons.

#### **5'-phosphorylated antisense siRNAs alone do not exhibit knockdown activity**

Some evidence has suggested that single-stranded antisense siRNAs, particularly when 5'-phosphorylated, are sufficient on their own to induce target gene knockdown, at least in HeLa cells (Ref: Martinez J. *et al.*, 2002). However, this was not a finding in the experiments of the previous chapter performed in 3T3 cells. In order to address this issue, sense alone, antisense alone, and double-stranded versions of two different *Firefly* siRNAs were tested in both mammalian CHO and *Drosophila* KC cells alongside a non-specific control and a three-nucleotide mismatched control. The results reveal, both for unphosphorylated and 5'-phosphorylated siRNAs, that *only* double-stranded siRNAs are able to yield significant luciferase knockdown. By contrast, single-stranded antisense (and sense) siRNAs on their own, whether phosphorylated or not, are unable to produce knockdown in either CHO or KC cells, thus confirming earlier conclusions that *both* homologous strands, irrespective of prior 5'-phosphorylation, are necessary for interference of a transiently expressed plasmid-borne gene. In addition, again confirming previous findings, discrete mismatches at three disparate sites within the target sequence largely abrogate this effect.

**5'-phosphorylation partially rescues siRNA action against mRNA luciferase in S2 cells**

In light of the above result using a DNA luciferase target, the effect of 5'-phosphorylation of siRNAs on knockdown of RNA luciferase was similarly assessed in *Drosophila* and mammalian cell lines. This revealed in S2 cells that, whereas previously the *Firefly* siRNA was unable to induce knockdown of the RNA luciferase, the addition of a 5'-phosphate partially rescues the knockdown effect of the siRNA. In other words, unlike in KC cells where siRNA knockdown of RNA luciferase occurs irrespective of 5'-phosphorylation, in S2 cells the siRNA requires a 5'-phosphate for effective knockdown of RNA luciferase. Thus, prior 5'-phosphorylation of the siRNA molecule appears to rescue the knockdown effect against an RNA luciferase target in S2 cells resulting in an outcome similar to KC cells. However, this result is also subject to more prosaic explanations including experimental variation and technical error. If on the other hand it is a genuine finding, then it suggests that KC cells contain a high level of endogenous 5' nucleotide kinase activity relative to S2 cells, a prediction that might be formally tested.

**5'-phosphorylated siRNAs fail to knockdown mRNA luciferase in mammalian cells**

Two additional salient negative findings from the foregoing experiment are also relevant. Firstly, unlike for S2 cells, 5'-phosphorylation fails to rescue siRNA knockdown function against RNA luciferase in either CHO or 3T3 cells, revealing an important qualitative difference between *Drosophila* and mammalian cells. Secondly, there is no demonstrable knockdown of RNA luciferase using antisense siRNAs alone, whether 5'-phosphorylated or not, in any of the cell lines tested, confirming again that both sense and antisense strands are necessary for interference against RNA, as well as DNA, gene targets.

**5'-phosphorylation is able to rescue the knockdown activity of (antisense-fidelitous)****heteroduplex siRNAs in *Drosophila* cells against DNA luciferase but not mRNA luciferase**

As presented previously, mismatched heteroduplex siRNAs are able to knockdown both plasmid-borne luciferase and endogenous luciferase in CHO and 3T3 cells but *only* if the antisense strand retains complete sequence fidelity with the target. In other words, siRNA knockdown in mammalian cells appears to be relatively tolerant of sense strand mismatches. However, these results are not reproducible in *Drosophila* cells. Instead, the same heteroduplex siRNAs that are able to knockdown DNA luciferase in mammalian cells are unable to do likewise in KC and S2 cells. However, by 5'-phosphorylating these heteroduplex siRNAs, their knockdown action is partially restored in both KC and S2 cells, but only against DNA luciferase and *not* against RNA luciferase. Thus, the knockdown action of (antisense fidelitous) heteroduplex siRNAs against DNA luciferase appears to be 5'-phosphate-independent in mammalian cells but 5'-phosphate-dependent in *Drosophila* cells. In addition, there is no demonstrable knockdown effect against

RNA luciferase in either mammalian or *Drosophila* cells irrespective of 5'-phosphorylation status. Once again, it is possible that more prosaic explanations of experimental variation and technical error might account for these results.

### **Summary of findings**

This chapter sought to define the key molecular determinants of effective siRNA-induced knockdown and in so doing shed light on the site of action of the siRNA molecule. Some of the findings in the two mammalian cell lines used were unexpected and, in view of this, efforts were made to corroborate these findings by repeating the same experiments in two *Drosophila* cell lines. The summary of all these results are listed below:

- siRNA knockdown activity in mammalian cells, as revealed by the use of heteroduplex molecules, requires antisense strand fidelity but is tolerant of sense strand mismatch
- siRNAs that yield robust knockdown of a transiently transfected plasmid-borne DNA luciferase in 3T3 or CHO mammalian cells fail to knockdown an mRNA-borne luciferase
- siRNAs produce knockdown of *both* DNA and mRNA luciferase in KC *Drosophila* cells but only DNA and not mRNA luciferase in S2 *Drosophila* cells
- 5'-phosphorylation of the siRNA molecule appears to partially rescue its knockdown activity against mRNA luciferase in S2 *Drosophila* cells
- 5'-phosphorylation of the siRNA molecule *fails* to rescue its knockdown activity against mRNA luciferase in 3T3 or CHO mammalian cells
- Heteroduplex siRNAs that yield robust knockdown of DNA luciferase in mammalian cells fail to produce knockdown of either DNA or mRNA luciferase in *Drosophila* cells
- 5'-phosphorylation of the heteroduplex siRNA is able to partially rescue its knockdown activity in KC and S2 *Drosophila* cells against DNA luciferase but *not* mRNA luciferase
- Antisense-strand molecules alone, whether 5'-phosphorylated or not, fail to exhibit any knockdown activity against DNA or mRNA luciferase in either mammalian or *Drosophila* cells

### **Statistical considerations**

This chapter contains a large amount of data that has been analysed using various quantitative methods to determine levels of statistical significance. However, statistical significance alone cannot prove that an observation is indeed experimentally genuine or biologically meaningful. This is especially true where large numbers of data-points are being correlated with outcomes. Moreover, while most these data-points were derived from replicated experiments, some were derived from non-replicated experiments for which standard errors cannot be calculated. Where this occurred, the validity of the results was considered according to whether the effect showed dosage consistency or occurred similarly for two different siRNAs. However, inevitably, there is the risk that results are over-interpreted leading to incorrect conclusions. For this reason, the findings listed above are provisional and intended to serve as hypotheses that might be tested by future experiments.

### **Future experiments**

The most consistently reproducible finding from this chapter was the demonstration that Firefly-specific siRNAs, when co-transfected with the target gene into CHO cells, are unable to induce knockdown of mRNA-encoded luciferase. It was borne out similarly in 3T3 cells, except perhaps for a weak delayed effect, and also *Drosophila* S2 cells. However, in KC cells, the siRNAs *do* induce knockdown of the mRNA-encoded luciferase with similar potency as for DNA-encoded luciferase. This observation served to strengthen the validity of the experimental model and confirm that the lack of knockdown seen in CHO, 3T3 and S2 cells is indeed genuine. In order to explore this finding further, future experiments are proposed below:

- The key experiment to establish the precise site of action of siRNA interference in these cells is the nuclear run-on transcription assay. This would allow a transcriptional level of interference to be distinguished from a post-transcriptional/translational level via incorporation of radio-labelled nucleotides into nascent mRNA strands in isolated nuclei and subsequent autoradiography. In the former case, mRNA transcripts would cease to be produced upon siRNA action while in the latter, they would continue to be produced and then destroyed by siRNA-induced strand cleavage.
  
- A second strategy would be to quantify the mRNA using immunoblotting or even microarray analysis. However, neither of these approaches possesses the ability to measure nascent transcripts and hence they would fail to distinguish between mRNA transcriptional arrest and mRNA destruction.

- Finally, detection of mRNA cleavage products, by immunoblotting for size or sequencing of 5'RACE products, would reveal whether the mRNA was cleaved within the siRNA target site, yielding proximal and distal fragments, or alternatively interrupted during transcription, when only a proximal fragment would result.

Of these experiments, the nuclear run-on experiment would be the most important in determining whether siRNA knockdown in these cells operates predominantly at the transcriptional DNA level or post-transcriptional mRNA level. It is conceivable that gene expression might be interrupted at any point along this route, from accessibility of the DNA to the transcriptional machinery to generation of an mRNA molecule and its translocation out of the nucleus. Finally, the possibility that the non-physiological nature of the experimental model used in this chapter, which uses the transient transfection of mRNA or DNA molecules into cells, might have resulted in artefactual results should be considered. Undoubtedly, this strategy produces an artificial cellular situation raising the possibility that this alone accounts for the unexpected outcome that co-transfected siRNAs knockdown DNA but not mRNA genes. This is explored further in the forthcoming discussion.

# **Chapter 6**

## **Targeting endogenous genes by RNA interference**

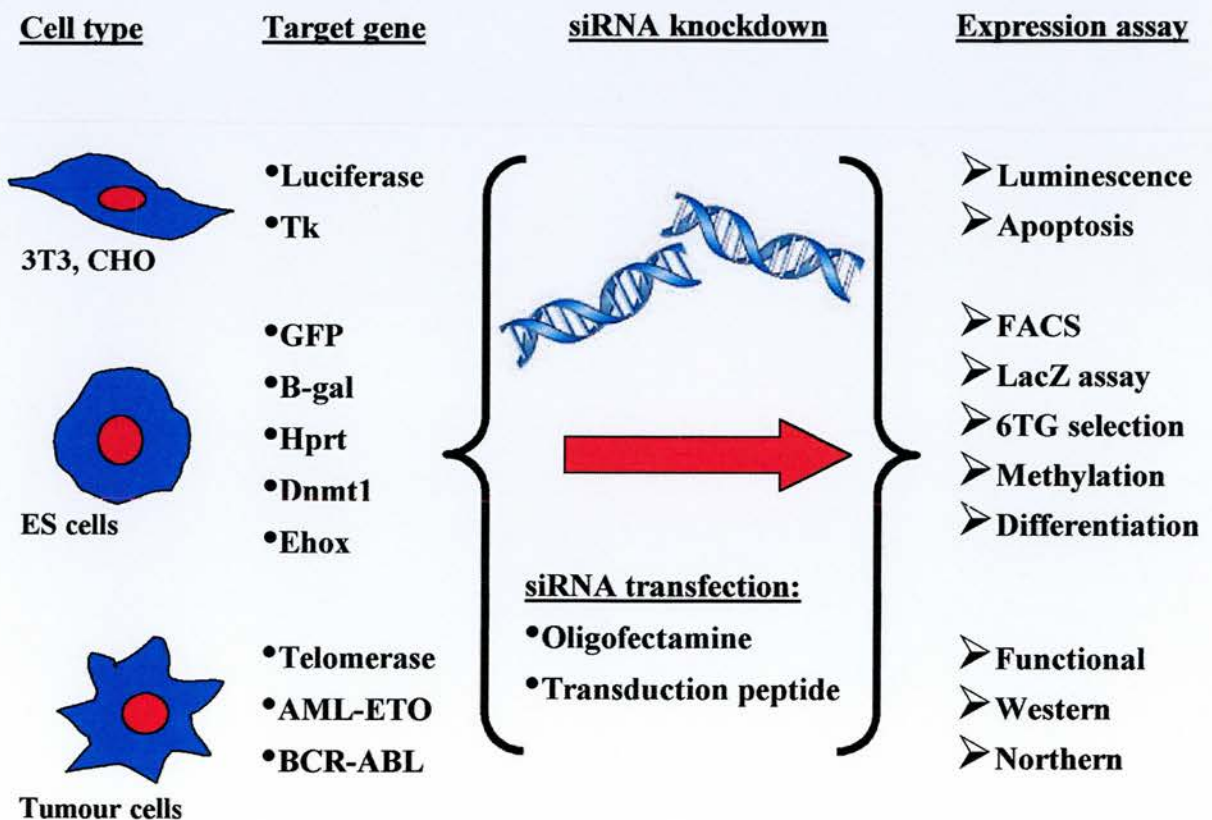
### **Results 4**

## 6.1 General design of siRNA-mediated knockdown experiments against various gene targets in different cell lines

The principal goal of this chapter was to demonstrate that siRNA knockdown can operate in murine ES cells by targeting a number of different gene targets: transiently expressed *Firefly* luciferase; randomly integrated GFP transgene;  $\beta$ -galactosidase gene targeted to an endogenous locus (Oct3/4); and the endogenous DNA methyltransferase gene (Dnmt1). In addition, results are presented from siRNA-mediated knockdown of the telomerase reverse transcriptase gene (hTERT) in two human breast cancer cell lines (MDA and T47D) and of the pyrimidine salvage pathway thymidine kinase gene (Tk) in CHO cells. Finally, the pertinent negative results from an experiment designed to induce permanent silencing of the hypoxanthine phosphoribosyl transferase gene (Hprt) by siRNA targeting and 6-thioguanine selection of ES cells are presented.

The basic design of all the siRNA knockdown experiments was essentially identical with gene expression assayed either qualitatively, quantitatively, or functionally as appropriate (Figure 6.1).

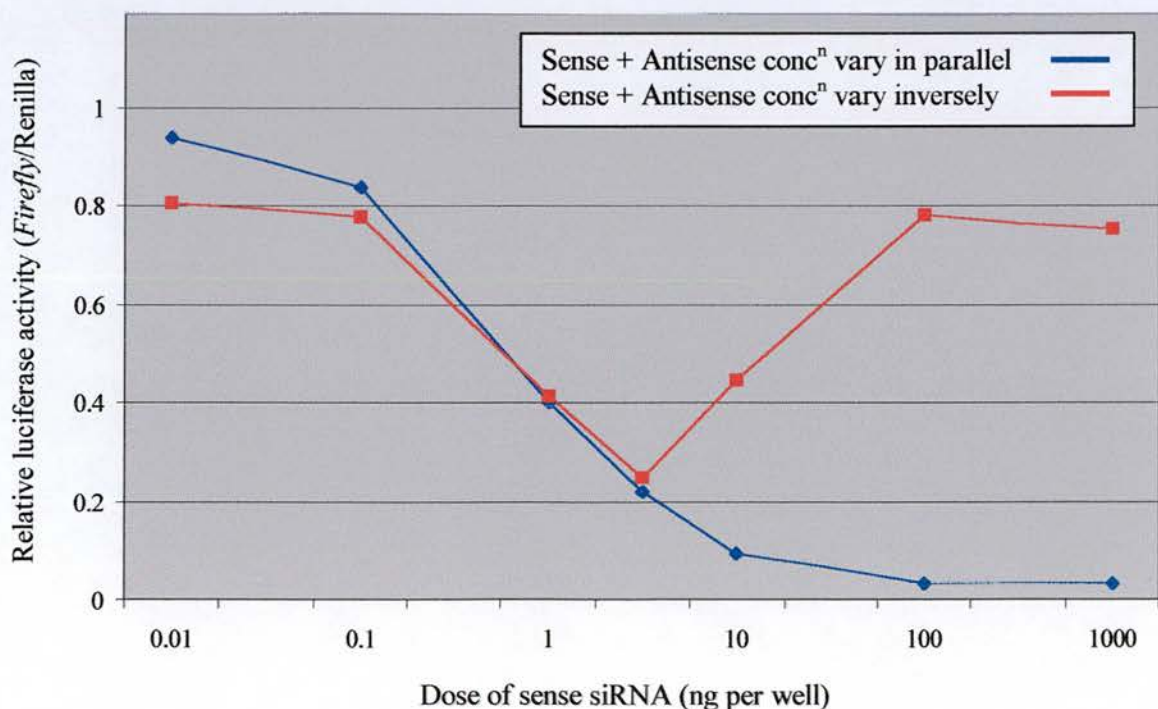
**Figure 6.1: Schematic illustration demonstrating the basic experimental strategy for studying siRNA-induced knockdown of various target genes in a number of different cell lines.**



## 6.2 Transfection of *Firefly*-specific siRNAs result in dose-dependent knockdown of transiently expressed luciferase in murine ES cells

Results presented in previous chapters have established that siRNA molecules bearing sequence homology with a target reporter luciferase are able to mediate knockdown of gene expression in a potent and dose-dependent fashion in mammalian CHO and 3T3 cell lines. Co-transfection of *Firefly*-specific siRNAs with pGL3 luciferase plasmid in murine ES cells recapitulates these findings and confirms that RNA interference operates in a similar fashion in ES cells as for CHO and 3T3 cells. The results reveal that the step-wise increase of both sense and antisense siRNA strands in the transfection mix leads to a dose-dependent knockdown of luciferase expression. By comparison, variance of sense and antisense siRNA strands in inverse amounts results in a “V”-shaped curve in which maximum knockdown occurs when sense and antisense strands are present in equimolar amounts (Figure 6.2). This confirms that the stoichiometry of sense and antisense strands in the interference process is 1:1 consistent with the view that the active intermediary species is double-stranded RNA. Finally, the demonstration that RNA interference operates in ES cells opens up the possibility of using siRNAs to functionally analyse differentiation genes.

**Figure 6.2: Graph showing siRNA-mediated knockdown of *Firefly* luciferase in murine ES cells.** The knockdown of transiently expressed *Firefly* luciferase (from pGL3 plasmid) in murine ES cells using specific siRNA molecules exhibits dose-dependent potency, sequence specificity and 1:1 stoichiometry. The amount of antisense strand goes up in parallel or down inversely with the sense strand dose shown.

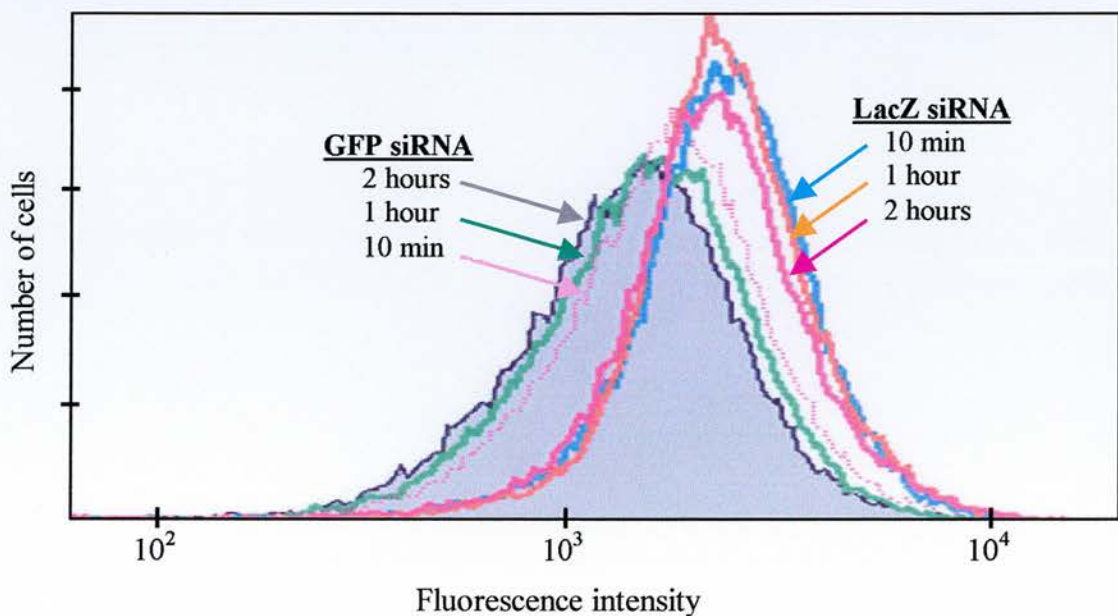


### 6.3 Sequence-specific siRNA molecules are able to suppress expression of a GFP transgene stably integrated by random genomic insertion in ES cells

Sequence-specific siRNAs were tested for their ability to knockdown expression of enhanced GFP in a genetically engineered ES cell line kindly donated by A. Medvinsky (Ref: Gilchrist D.S. *et al.*, 2003). Here, the GFP gene is expressed from a powerful constitutive promoter and carried on a transgene stably integrated by random genomic insertion. siRNA molecules sharing sequence homology with region 125 to 143 of the EGFP gene were transfected using Oligofectamine™ using a modified protocol in which cells were exposed for variable periods to the transfection reagent containing siRNA at 100ng/ml. ES cells were then assessed for GFP expression by flow cytometric analysis of fluorescence at 48 hours post-transfection and compared with ES cells transfected with control (LacZ-specific) siRNAs. The results reveal a modest knockdown of fluorescence level that is maximal (34%) for the two-hour siRNA incubation experiment (Figure 6.3). Comparison of the two sets of data using a paired Student's t-test confirms statistical significance ( $p=0.001$ ). The limited potency of the knockdown effect may reflect poor efficacy of the siRNA molecule or inefficiency of the transfection process. In addition, demonstration of EGFP knockdown is made more difficult given its prolonged intracellular stability with a half-life of approximately 12-24 hours (personal communication A. Medvinsky).

**Figure 6.3: Flow cytometric analysis showing siRNA-mediated knockdown of GFP in ES cells.**

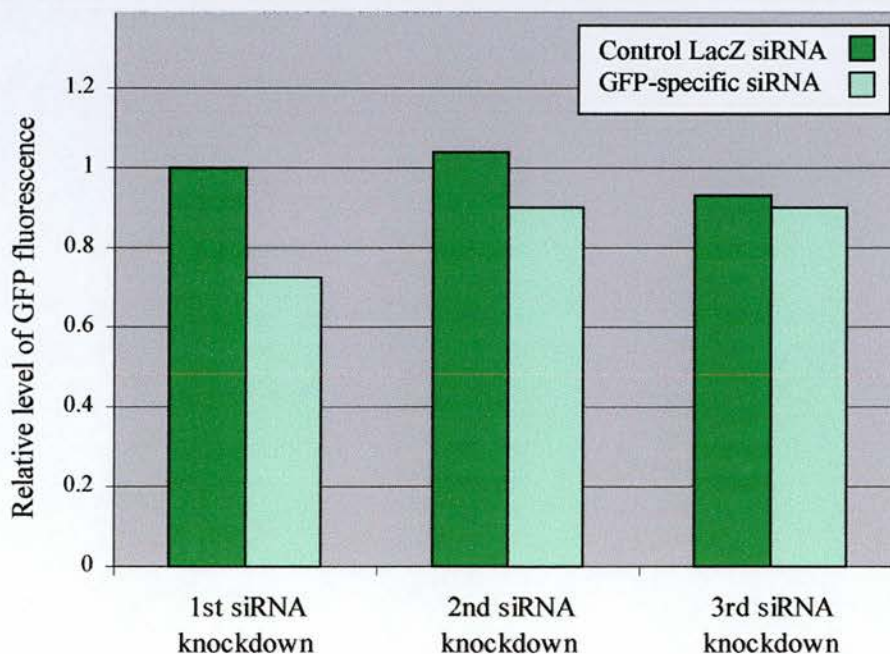
siRNAs sharing sequence homology with EGFP were tested against control LacZ siRNAs for their ability to knockdown expression of GFP in an ES cell line carrying a stably integrated GFP transgene. The two keys indicate the lengths of time the GFP and LacZ siRNAs were incubated with the ES cells. Fluorescence was assessed by flow cytometry at 48 hours for three different siRNA incubation times. Analysis of the data using a paired Student's t-test demonstrates statistical significance ( $p=0.001$ ).



#### 6.4 Repeated knockdowns of GFP expression, via serial transfection of siRNA, result in diminished rather than enhanced knockdown effect

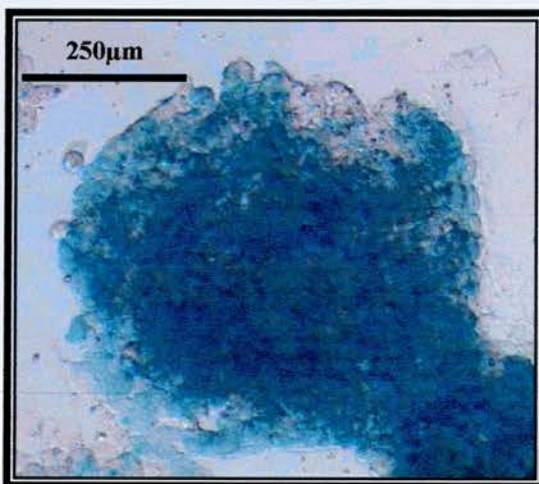
In order to test whether repeated knockdowns of GFP expression would result in an enhanced cumulative effect, serial transfections of GFP-specific siRNA were performed at three-day intervals on the same ES cells. At each passage, an aliquot of cells were analysed by flow cytometry while the remainder were re-suspended directly into transfection mix for one hour prior to re-plating on gelatin-coated wells. Contrary to expectation, this experiment revealed that the magnitude of the knockdown was diminished at each passage from 28% to 14% to 4% relative to the control ES cells that were transfected with non-specific (LacZ) siRNA (Figure 6.4). It is possible that this apparent fall-off in magnitude of GFP knockdown at each transfection cycle reflects a progressive depletion of the molecular components of the RNAi machinery. However, the finding fails to reach statistical significance using Student's t-test for paired data ( $p=0.13$ ) and might be more easily explained in terms of experimental variation. Nonetheless, it is included here as an observation that is worthy of further investigation.

**Figure 6.4: Repeated cycles of siRNA transfection result in diminished levels of GFP knockdown.** Repeated cycles of transfection of ES cells with GFP-specific siRNAs result in diminished, rather than enhanced, levels of GFP knockdown relative to the baseline cells transfected with control LacZ siRNAs. Data-points are derived from single experiments and hence error bars have not been applied. The result is not statistically significant using a paired Student's t-test ( $p=0.13$ ).



### 6.5 LacZ-specific siRNAs can suppress expression of a $\beta$ -galactosidase transgene targeted to the Oct3/4 locus in murine ES cells

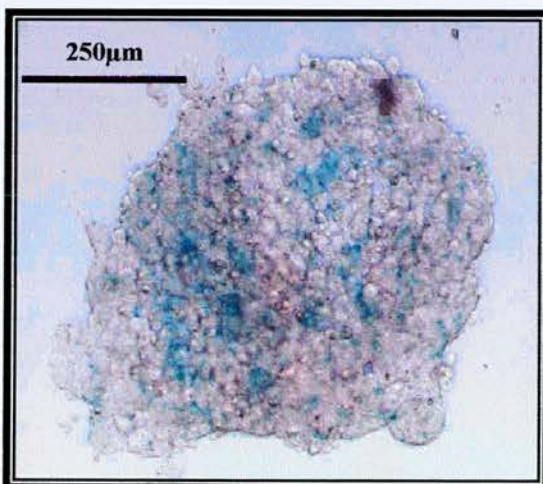
LacZ-specific siRNAs were tested for their ability to knockdown a  $\beta$ -galactosidase transgene that had been targeted to the endogenous Oct3/4 locus in ES cells kindly provided by A. Smith (Ref: Dani C. *et al.*, 1998). This ES cell line, termed IOUD2, permits the testing of siRNAs against a reporter gene integrated within a physiological gene locus rather than a random integration locus as for the GFP cell line used in the previous experiments. The photomicrographs below illustrate the unequivocal knockdown of  $\beta$ -galactosidase expression within discrete ES cell colonies (Figure 6.5). Here, ES cells have been seeded at low density following siRNA transfection to allow colonies to form. The level of LacZ knockdown was then quantified after five days using X-gal staining. This analysis reveals the persistence of  $\beta$ -galactosidase knockdown as the ES colonies grow indicating clonal inheritance of the interference effect through repeated cell divisions.



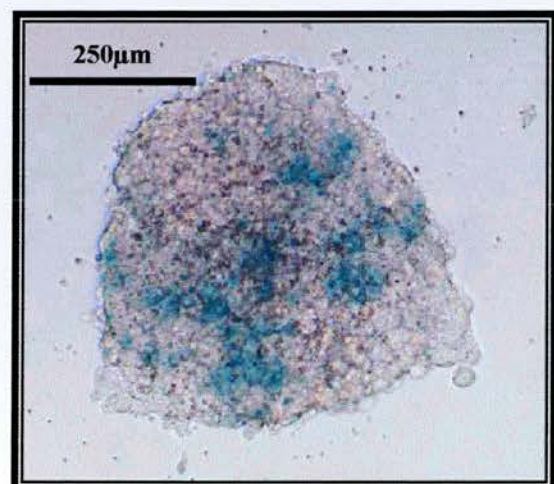
Control (GFP) siRNA at 4ug/well

#### **Figure 6.5: LacZ-specific siRNAs knock down expression of a stable $\beta$ -galactosidase targeted to the Oct3/4 locus in ES cells.**

ES cells were transfected with LacZ-specific siRNA and then re-plated at low dilution. At five days, colonies were assessed for level of expression of  $\beta$ -galactosidase using X-gal staining. This revealed a dose-dependent knockdown of  $\beta$ -galactosidase expression compared with control (GFP) siRNA. The level of knockdown was highly variable and colonies exhibited considerable heterogeneity of  $\beta$ -galactosidase expression (see Figure 6.6).



LacZ siRNA at 1ug/well

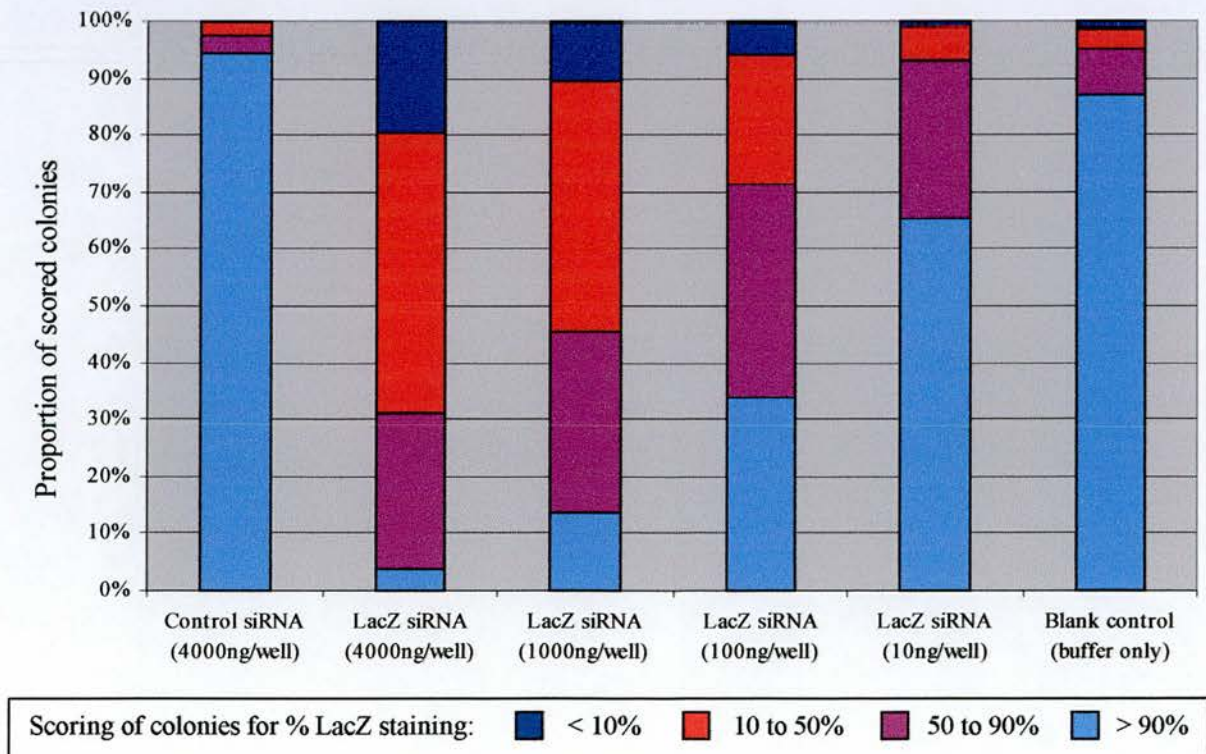


LacZ siRNA at 4ug/well

### 6.6 siRNA-induced knockdown of $\beta$ -galactosidase expression in ES cells exhibits dose-dependency and clonal heritability

B-galactosidase-expressing IOUD2 cells were transfected in suspension using either LacZ-specific siRNAs or control (GFP) siRNAs, at varying doses, and seeded on 10cm plates at a concentration of  $10^3$  cells per plate. Colonies were allowed to emerge over the following five days after when the plates were stained for  $\beta$ -galactosidase activity using the X-gal assay. Colonies were scored under light microscopy for the intensity of blue staining quantified according to the approximate proportion of positive cells per colony. Each experiment was unique and hence statistical analysis has not been performed. However, the dose-dependent nature of the effect supports the conclusion that siRNA-induced knockdown of  $\beta$ -galactosidase expression is genuine. Indeed, at the highest siRNA dosage tested, a significant proportion of colonies showed negligible blue staining five days post-transfection. This confirms that the RNAi effect can be propagated from parental cell to daughter cell through several rounds of cell division. However, even at this dosage there remain many colonies that do not exhibit any  $\beta$ -galactosidase knockdown suggesting either failure of transfection of the siRNA into the original cell or alternatively incomplete knockdown or even resistance of the cell to the interference effect.

**Figure 6.6: Graph showing siRNA-induced knockdown of  $\beta$ -galactosidase expression in ES cells.** LacZ-specific siRNAs are able to knockdown expression of a  $\beta$ -galactosidase transgene in IOUD2 ES cells in a dose-dependent fashion. Each experiment was unique and hence error bars have not been applied and statistical analysis has not been performed.



### **6.7 siRNAs directed against the telomerase reverse transcriptase gene (hTERT) fail to show significant functional knockdown of telomerase activity in human cancer cell lines**

Sequence-specific siRNAs were tested for their ability to knockdown expression of the reverse transcriptase subunit of the telomerase enzyme (hTERT) in human breast cancer cell lines. hTERT is an integral part of the telomerase complex required for the retention of telomeric sequences present at the termini of the chromosomes. This involves the catalytic restoration of hexanucleotide (TTAGGG) repeats that are otherwise eroded during DNA replication. Evidence suggests that these are required for chromosomal stability in cells undergoing continual self-renewal including both stem cells and cancer cells, both of which exhibit high levels of telomerase activity (Ref: Cech T.R., 2004). In this experiment, which was part a collaborative project with A. Tsakiridis (Honours student, Edinburgh University), hTERT-specific siRNAs were tested for their ability to knockdown telomerase activity as measured by a functional PCR-based ELISA telomerase assay (Telomerase PCR ELISA™, Roche Inc.).

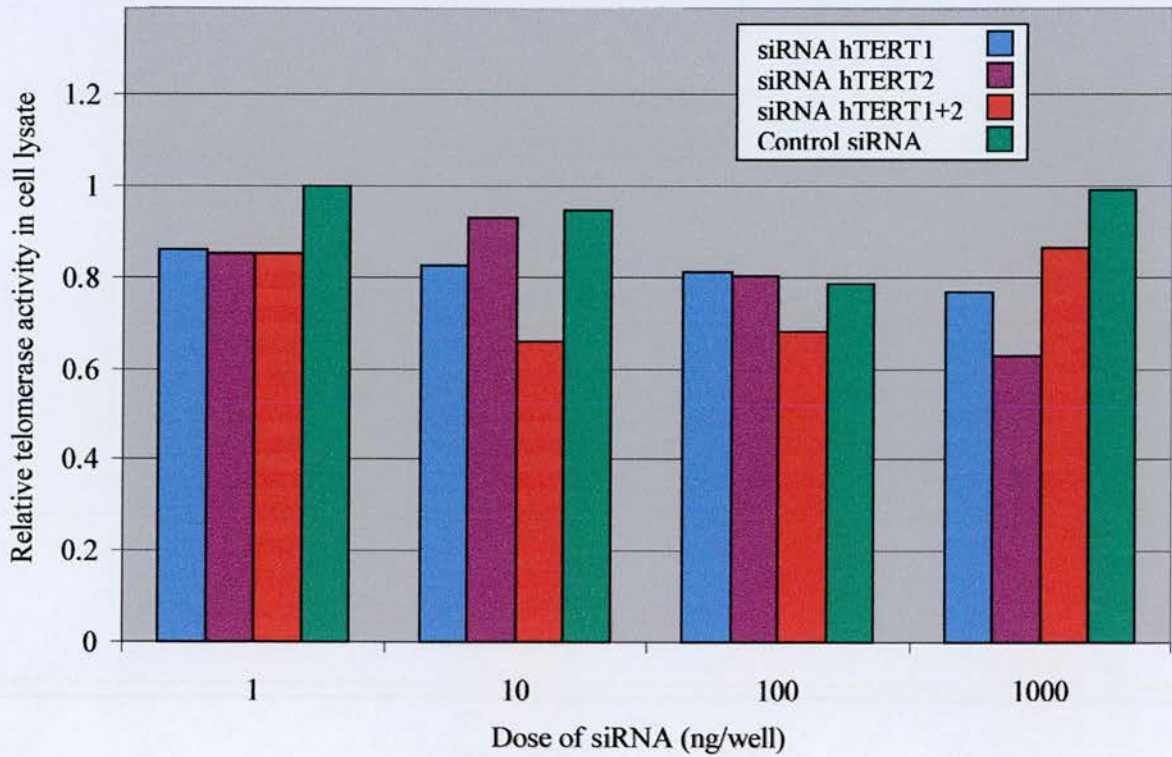
Two human breast cancer cell lines were used, namely MDA435 and T47D. These were serially transfected four times, as per the suspension method described previously, using two different siRNAs homologous to distinct regions of the hTERT gene and a non-specific control siRNA. The cells were then lysed and telomerase activity assayed using the TRAP method (Telomeric Repeat Amplification Protocol) and quantified via a sensitive photometric enzyme immunoassay. However, the results fail to demonstrate a consistent and dose-dependent knockdown of telomerase activity. Only at the highest level of 1µg siRNA was there any evidence of loss of telomerase activity (77% for the upstream hTERT siRNA, 63% for the downstream hTERT siRNA, 86% for both hTERT siRNAs combined, and 99% for the control siRNA) although these results are derived from single experiments and hence statistical analysis is not possible

Besides technical failure of the experiment itself, or of the siRNA molecules used, other reasons for the absence of interference effect on telomerase activity include the following:

- 1) hTERT may not be the rate-limiting step determining telomerase activity within a cell
- 2) hTERT protein may persist causing sustained telomerase activity
- 3) the PCR-based assay may be poorly sensitive to low levels of activity.

Finally, any functional assay of telomerase activity will inevitably under-measure the true level of gene knockdown compared with an RNA-based assay such as northern blot analysis or RT-PCR.

**Figure 6.7: Graph showing effect of transfection of hTERT-specific siRNAs in MDA435 cells.** siRNAs specific for the hTERT gene fail to demonstrate knockdown of telomerase activity when transfected into MDA435 breast cancer cells. Data-points are derived from single experiments and hence error bars have not been applied. The apparent lack of telomerase knockdown might be attributable to a number of problems including technical and biological issues.



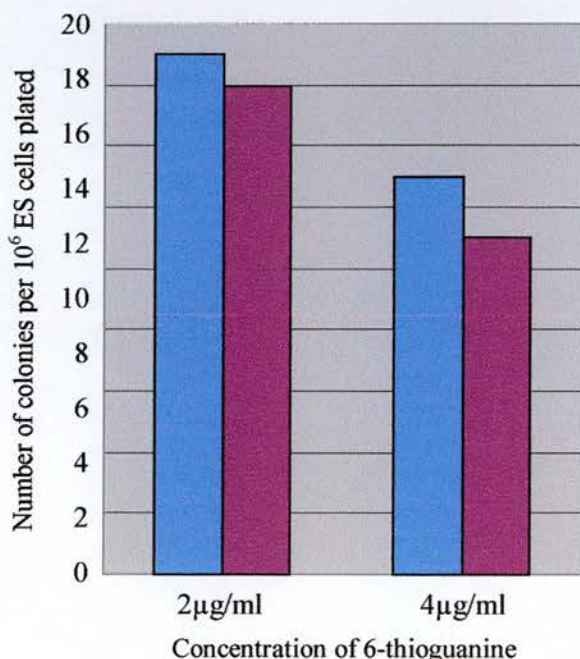
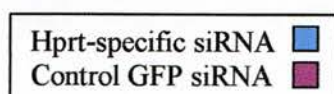
### 6.8 siRNA-induced knockdown in ES cells of the purine salvage pathway enzyme, Hprt, fails to generate surplus mutant clones by negative selection in 6-thioguanine

Hypoxanthine phosphoribosyl transferase (Hprt), which resides on the X chromosome and hence is present in a single copy in male ES cells, is a key enzyme involved in purine synthesis via the salvage pathway. ES cells that are mutant for the Hprt locus, whether arising spontaneously or via a genotoxic insult, can be isolated using the purine analogue 6-thioguanine. This is converted by Hprt into a cytotoxic metabolite that kills wild-type cells and thereby allows negative selection of Hprt deficient cells. In order to test the hypothesis that Hprt-specific siRNAs might be able to produce permanent silencing of the Hprt locus, for instance via methylation or some other epigenetic change at the gene locus, ES cells were selected with 6-thioguanine post-knockdown to see if this yielded surplus mutant clones above the background rate.

A total of  $10^6$  ES cells were transfected with siRNA specific for Hprt gene or a non-specific control siRNA. The cells were then re-plated on a total of six 10cm dishes and allowed to recover for 36 hours. Selection was then commenced in medium containing 6-thioguanine at two different doses,  $2\mu\text{g/ml}$  or  $4\mu\text{g/ml}$ . Colonies resistant to 6-thioguanine emerged at 8-10 days and these were counted and compared (Figure 6.8). The results demonstrate that cells treated with Hprt-siRNA fail to generate fewer 6-thioguanine-resistant clones, as hypothesised above, than the spontaneous background rate for cells treated with the control siRNA. A parallel experiment to derive thymidine kinase deficient ES cells, via siRNA-induced knockdown of the Tk locus followed by negative selection in 5-bromodeoxyuridine, yielded a similar result (data not shown).

**Figure 6.8: Hprt-specific siRNAs fail to generate surplus 6-thioguanine resistant ES clones compared with control siRNA.**

$10^6$  ES cells were transfected with either Hprt siRNA or control GFP siRNAs and then re-plated on 10cm dishes. The graph shows the number of ES colonies emerging after one-week selection with 6-thioguanine at toxic concentrations of  $2\mu\text{g/ml}$  and  $4\mu\text{g/ml}$ . Data points are derived from single experiments and hence error bars have not been applied. Cells treated with Hprt-siRNA fail to generate fewer 6-thioguanine-resistant clones than the spontaneous background rate (control-siRNA treated cells).



### 6.9 CHO cells can be efficiently killed by targeting the pyrimidine salvage pathway using siRNA-directed knockdown of thymidine kinase (Tk)

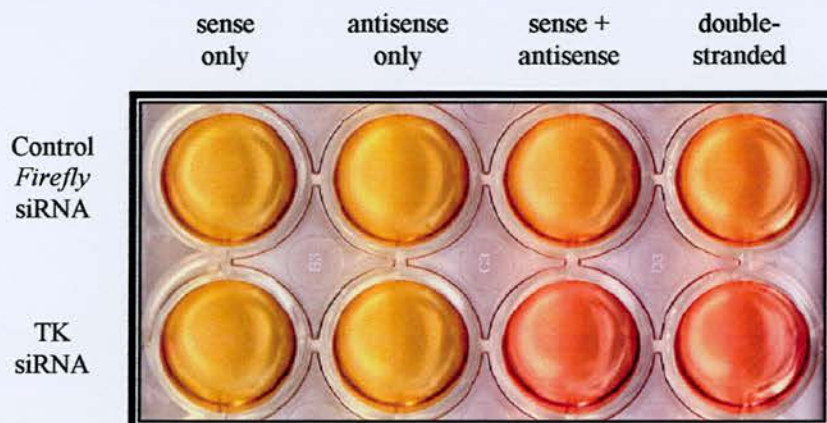
When CHO cells were transfected with siRNAs specific for the pyrimidine salvage pathway enzyme, thymidine kinase (Tk), widespread cell death was observed. This occurred even at very low siRNA doses of 10ng/well (which represents only 1% of the concentration used in most other siRNA knockdown experiments) and was never observed using a variety of control siRNAs (GFP, LacZ and *Firefly*). At 10ng/well, Tk-specific siRNA induced death of over 99% of cells (estimated according to the ratio of surviving colonies to cells plated < 0.01) while at higher concentrations of 100ng/well and 1µg/well there was absolute cell death (no surviving colonies). This result was qualitatively reproducible on a number of occasions and occurred in the absence of any selective toxic agent. By comparison, there was no significant toxic effect on murine ES cells and only a minor toxic effect on 3T3 cells suggesting that the salvage pathway is functionally redundant in these cells relative to CHO cells.

The above observation was exploited further to address the question of whether cells surviving Tk-siRNA knockdown were either spontaneously arising Tk mutants or otherwise resistant to the effects of RNA interference. CHO-LUC cells were selected using Tk-specific siRNA at 10ng/well and surviving cells were pooled and expanded. These were then subjected to repeat transfection with Tk-specific siRNA and shown to be equally susceptible to cell death as the original cells, thereby confirming that they were not mutant at the targeted Tk locus. Similarly, surviving CHO-LUC cells were subjected to transfection with *Firefly*-specific siRNA-B2 and shown to exhibit an expected knockdown of luciferase expression, thereby excluding any intrinsic resistance to the RNA interference process.

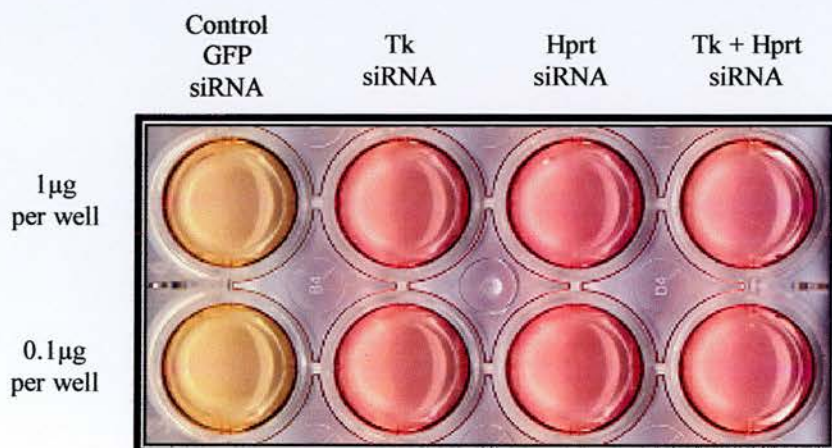
The observation of CHO cell toxicity arising from transfection with low dose Tk-specific siRNAs suggests a specific mechanism such as functional depletion of thymidine kinase activity leading to cellular apoptosis. If this is the case however, it is not clear why CHO cells should be peculiarly susceptible to targeting of the Tk gene within the pyrimidine salvage pathway given the redundancy of this enzyme to nucleotide synthesis. Nonetheless, it is of interest perhaps that the drug, 5-bromodeoxyuridine, acts as a chemosensitising agent via its action on thymidine kinase.

In a similar fashion, but to a tenfold lesser degree, Hprt-specific siRNA were also found to be toxic to CHO cells (data not shown). The cellular toxicity of both the Tk and Hprt siRNAs are demonstrated by the photographs overleaf showing the relevant wells of a 24-well plate in which the colour of the medium is a crude measure of metabolic cell activity (Figures 6.9 and 6.10).

**Figure 6.9: Transfection of Tk-specific siRNA (sense + antisense) is toxic to CHO-LUC cells.** Transfection of 10ng/well of Tk-specific siRNA is toxic to CHO cells only when both sense and antisense are used (either mixed in 1:1 ratio or pre-annealed double-stranded) while sense alone, antisense alone, and control siRNA are innocuous to the cells. This finding is evident here from failure of 1ml of medium to change colour from pink to yellow after 24 hours incubation indicating reduced metabolic activity. That this is a Tk-specific phenomenon is supported by the observation that the control *Firefly* siRNA, which yields potent knockdown of the stably integrated GL2 luciferase in CHO-LUC cells, is not associated with any significant cell death at equivalent doses.



**Figure 6.10: Transfection of Hprt-siRNA is also toxic in CHO cells but tenfold less than Tk-siRNA.** Hprt-specific siRNA was shown to induce cell death following transfection in CHO cells but to a tenfold lesser degree than Tk-specific siRNA. The toxicity of both siRNAs is indicated here by failure of 1ml of medium to change from pink to yellow after 24 hours incubation reflecting reduced metabolic activity. The non-specific control siRNA used in this experiment has specificity for the GFP gene.



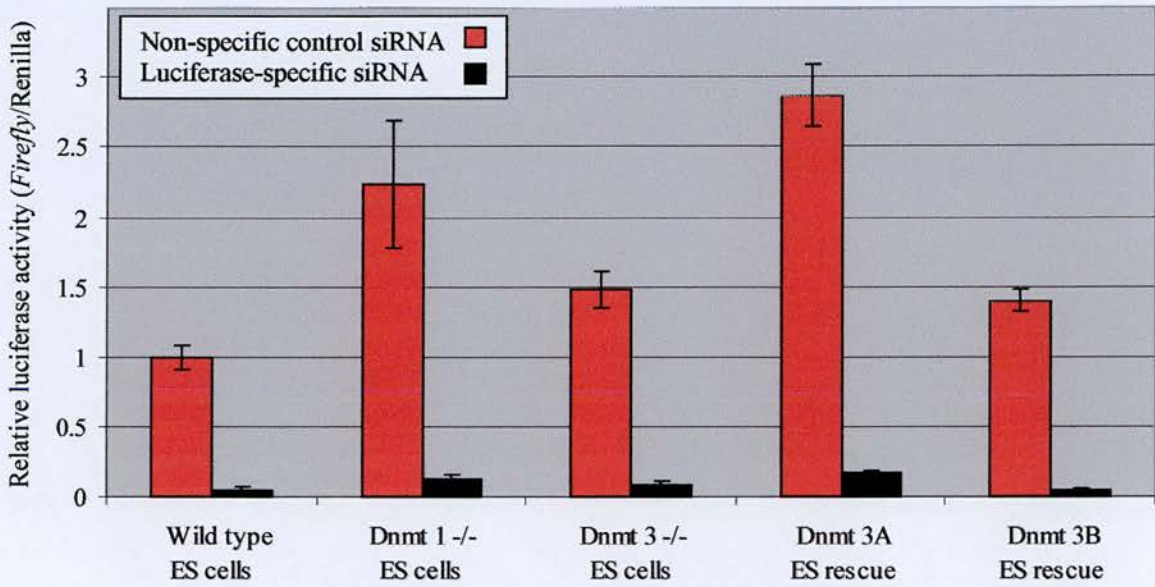
### 6.10 siRNA-induced knockdown in ES cells is not contingent upon the presence of either Dnmt1 or Dnmt3a/3b DNA methyltransferase activity

RNA interference in plants occurs initially via a post-transcriptional mechanism followed by transcriptional silencing mediated via DNA methylation and epigenetic change at the site of the targeted gene locus (Ref: Baulcombe D., 2004). Similarly, in mammalian cells, recent evidence has suggested that siRNAs can induce DNA methylation and epigenetic change likewise here also (Refs: Kawasaki H. and Taira K., 2004; Morris K. *et al*, 2004). In view of this, the relevance of DNA methyltransferase activity to the RNA interference process was tested. siRNA-B3, homologous to region 154-172 of the GL3 gene, was tested for its ability to knock down luciferase expression from a transiently transfected pGL3 plasmid in ES cells defective in either maintenance methyltransferase activity (Dnmt1<sup>-/-</sup>) or *de novo* methyltransferase activity (Dnmt3a<sup>-/-</sup>3b<sup>-/-</sup>) using cells kindly provided by B. Ramsahoye (Refs: Li E. *et al.*, 1992; Xie S. *et al.*, 1999; Okano M. *et al.*, 1999).

The results demonstrate that there is no significant compromise of knockdown efficacy in either Dnmt1<sup>-/-</sup> or Dnmt3a<sup>-/-</sup>3b<sup>-/-</sup> ES cells. This indicates that the presence of methyltransferase activity is *not* necessary for suppression of a transiently expressed epigene (Figure 6.11). Moreover, since Dnmt1<sup>-/-</sup> and Dnmt3a<sup>-/-</sup>3b<sup>-/-</sup> ES cells both have hypomethylated genomes (approximately 20-30% CpG methylation in Dnmt1<sup>-/-</sup> cells and <1% in late-passage Dnmt3a<sup>-/-</sup>3b<sup>-/-</sup> cells compared with >70% in wild-type ES cells, Refs: Li E. *et al.*, 1992; Jackson M. *et al*, 2004), then this result also demonstrates that RNA interference is not contingent upon a normally methylated genome. The control experiments using Dnmt3a<sup>-/-</sup>3b<sup>-/-</sup> ES cell lines rescued for either Dnmt3a or Dnmt3b also show normal siRNA-induced luciferase knockdown.

This finding, namely that methyltransferase activity is not necessary for siRNA-induced knockdown of an epigene, contrasts with the claim mentioned previously that siRNAs can bring about methylation of the gene promoter (Refs: Kawasaki H. and Taira K., 2004; Morris K. *et al*, 2004). While the two results are not wholly mutually exclusive, it does suggest that the DNA methylation is not the primary mechanism of siRNA-induced gene silencing. This conclusion is supported by the result of other investigators who have employed a similar strategy of using cells deficient in *de novo* methyltransferase activity and shown that double-stranded RNAs can induce transcriptional repression with appropriate chromatin changes independently of DNA methylation (Ref Ting A.H. *et al.*, 2005).

**Figure 6.11: siRNA-induced knockdown is not contingent upon DNA methyltransferase activity.** There is no significant difference in the potency of siRNA-induced knockdown of *Firefly* luciferase transiently transfected into ES cells deficient in maintenance methyltransferase activity (Dnmt1) or *de novo* methyltransferase activity (Dnmt3A/3B) compared with wild-type or 3A/3B rescue ES lines. Error bars represent the range derived from duplicate experiments.



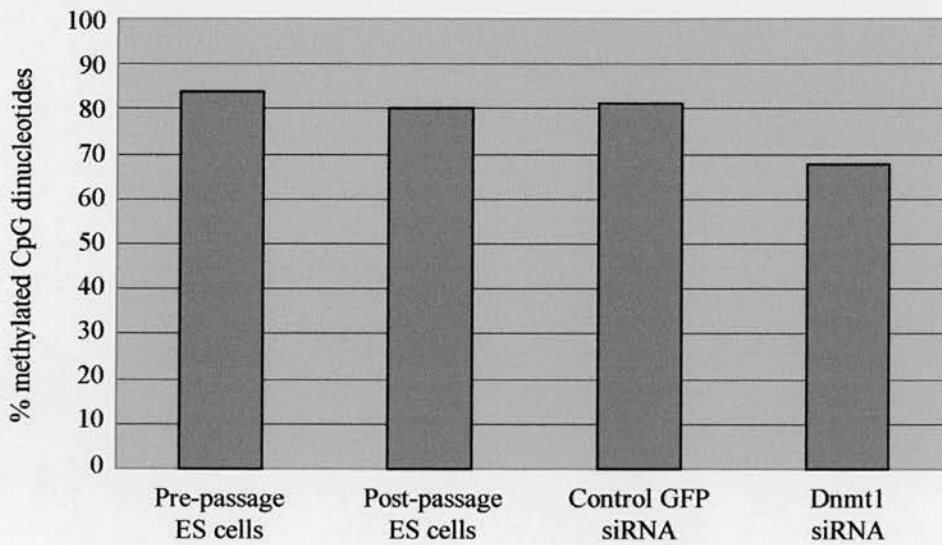
### 6.11 siRNAs directed against Dnmt1 result in a functional deficit of maintenance methyltransferase activity in ES cells

Wild-type ES cells exhibit approximately 70-80% methylation of CpG sites compared with approximately 20-30% for targeted knockout *Dnmt1*<sup>-/-</sup> ES cells (Ref: Li. E. *et al.*, 1992; Lei H. *et al.*, 1996). Dnmt1-specific siRNAs were tested for their ability to knockdown expression of maintenance DNA methyltransferase activity in wild-type ES cells resulting in an overall loss of methylated CpGs within the genome. ES cells were transfected using Dnmt1-specific siRNAs and seeded at low density on 10cm plates (10<sup>5</sup> cells per plate). The cells were cultured for four days to allow several rounds of cell division to occur with passive loss of methylated CpGs that are not maintained in the absence Dnmt1. This process was repeated serially four times over two weeks alongside a control experiment using a non-specific siRNA. Genomic DNA was then harvested for subsequent measurement of global CpG methylation level using the nearest neighbour analysis kindly performed by B. Ramsahoye.

The results are presented as raw data in the form of autoradiograph images (Figure 6.13). Quantification of CpG methylation (ratio of <sup>m</sup>CpG to total CpG) was performed by densitometry using the Scion Image™ software (available from [www.scioncorp.com](http://www.scioncorp.com)). This reveals that the proportion of methylated CpG dinucleotides is lower for the Dnmt1 knockdown (67.8%) than for the control knockdown (81.1%) and the non-treated pre-passaged (83.7%) and post-passaged ES cells (80.0%). Each methylation analysis was derived from a single experiment and hence error bars cannot be applied. However, the three control experiments yield a mean CpG methylation of 81.6% with standard deviation of 1.9% suggesting that Dnmt1 knockdown is indeed statistically significant and can produce partial genomic DNA demethylation (Figure 6.12).

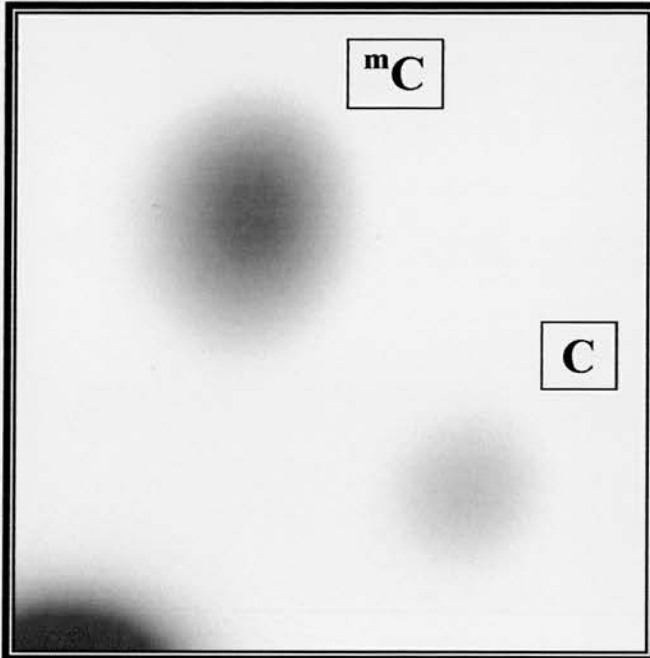
**Figure 6.12: siRNA-induced knockdown of Dnmt1 results in loss of genomic methylation in ES cells.**

Nearest neighbour quantification of genomic CpG methylation levels in wild-type ES cells as follows: pre-passaged (83.7%), post-passaged (80.0%), control GFP siRNA knockdown (81.1%), and Dnmt1-specific knockdown (67.8%). This represents an overall loss of genomic CpG methylation attributable to Dnmt1 siRNA knockdown of 16.4%. Data-points are derived from single experiments and hence error bars have not been applied. The difference in CpG methylation is statistically significant (mean 81.6, SD 1.9  $p < 0.001$ ).



**Figure 6.13: siRNA Dnmt1 knockdown leads to loss of cytosine methylation in wild-type ES cells.** Dnmt1 was knocked down by serial siRNA transfection of wild-type ES cells. Genomic DNA was harvested and assayed for cytosine methylation using the nearest neighbour method. Autoradiograph densitometry studies shown below were analysed using Scion Imaging™ software and reveal a reduction of methylated CpG dinucleotides of 16.4% relative to control siRNA knockdown cells.

#### Knockdown using non-specific control siRNAs



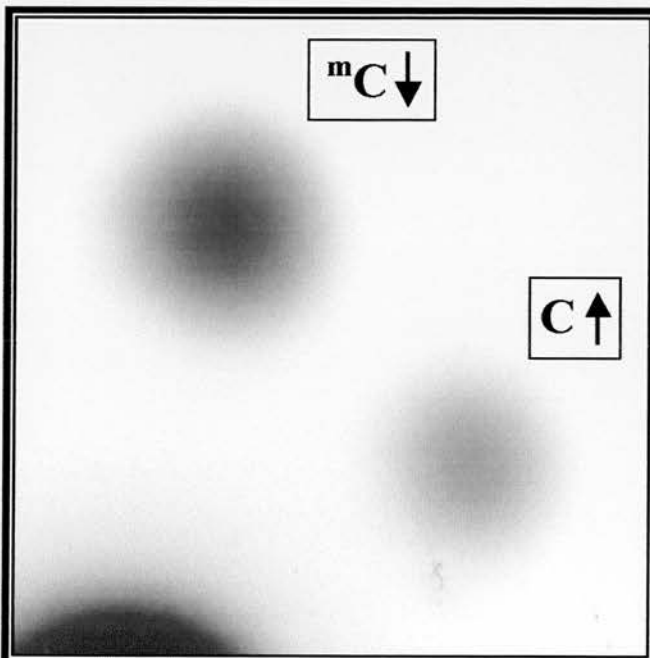
#### Cytosine Methylation

Percentage methylation  
at CpG dinucleotides:

**Methylated CpG:**  
81.1%

**Unmethylated CpG:**  
18.9%

#### Knockdown of Dnmt1 using specific siRNAs



#### Cytosine Methylation

Percentage methylation  
at CpG dinucleotides:

**Methylated CpG:**  
67.8%

**Unmethylated CpG:**  
32.2%

## 6.12 Conclusion

The primary objective of this chapter was to demonstrate that siRNA molecules are able to generate gene interference in murine embryonic stem cells by targeting various reporter genes expressed either transiently or stably in the cell. The secondary objective was to extrapolate these findings to various endogenous genes in ES cells, and other cell lines, with the goal of yielding an informative knockdown phenotype. Finally, results are presented from experiments designed to investigate a putative link between siRNA-induced gene interference and induction of epigenetic silencing of that gene in the cell.

### **siRNAs induce knockdown of a transiently expressed *Firefly* epigene in ES cells**

The initial experiment sought to recapitulate the findings of earlier experiments in 3T3 and CHO cells, namely the demonstration of knockdown of a transiently expressed *Firefly* luciferase gene by sequence-specific siRNAs. This confirmed, as expected, that siRNA molecules co-transfected with the *Firefly* plasmid result in the dose-dependent knockdown of luciferase activity with a sense to antisense strand stoichiometry of 1 to 1. Thus, RNAi in ES cells, as for other mammalian cells, requires equimolar amounts of matching oligonucleotide strands which base pair to yield a double-stranded RNA molecule that shares obligatory sequence homology with the target gene.

### **siRNAs induce knockdown of a stable randomly integrated GFP gene in ES cells**

Subsequent experiments in ES cells sought to extrapolate these findings from the setting of a transiently expressed target gene to that of a stably expressed one. A transgenic ES cell line carrying a randomly integrated enhanced green fluorescent protein (EGFP) reporter gene was chosen for this purpose. The results reveal that GFP-specific siRNA yields a modest but definite knockdown of fluorescence activity compared with non-specific control siRNA. The limited potency of this knockdown effect, compared with that of transiently expressed luciferase, can be explained in terms of the considerable stability of green fluorescent protein in the cell.

In an attempt to produce a more substantial GFP knockdown, through serial transfections of siRNA into the same cells, it was found that each round of transfection yielded a diminished knockdown effect compared with the previous round. In other words, the cells appear perversely *less* susceptible to the gene interference effect of the siRNA following prior exposure to the same siRNA, possibly due to exhaustion of the molecular constituents of the RNAi machinery. However, this result was not amenable to statistical analysis on account of non-replicated experimental data and the finding requires verification against other endogenous gene targets. Further exploration of this phenomenon was not pursued here.

**siRNAs induce knockdown of a  $\beta$ -galactosidase gene targeted to the Oct3/4 locus in ES cells**

The next logical experiment, building on the previous result, was the demonstration that siRNA molecules can mediate the knockdown a reporter gene located within a physiological gene locus. A  $\beta$ -galactosidase expressing ES cell line, IOUD2, in which the transgene construct has been targeted to the Oct3/4 locus, was used for this purpose. Expression of the Oct3/4 gene is critical for maintaining an undifferentiated ES state. In the IOUD2 cell line, which by necessity is hemizygous for the native Oct3/4 gene, expression from a single allele appears to suffice for this purpose. The knockdown of the  $\beta$ -galactosidase gene in these ES cells, using LacZ-specific siRNAs, resulted in the dose-dependent reduction of expression of the reporter gene by X-gal staining analysis, while non-specific siRNA had no significant effect. Of note is the fact that this knockdown was *not* associated with any observable ES cell differentiation implying that silencing of the  $\beta$ -galactosidase gene targeted to the Oct3/4 locus did not result in the concomitant silencing of the homologous Oct3/4 locus.

**siRNA-induced  $\beta$ -galactosidase knockdown is clonally inherited through many cell divisions**

By seeding the IOUD2 ES cells at low density following transfection of LacZ-specific siRNAs, it was possible to observe the staining pattern of discrete colonies emerging after five days, each one typically derived from a single ES cell. This revealed considerable heterogeneity of staining from mostly negative colonies at high siRNA doses to mostly mosaic colonies at low siRNA doses. Thus, it appears that the RNAi effect is clonally propagated from parent cell to daughter cell through several rounds of cell division and the permanence of this effect is determined by siRNA dosage. At all siRNA doses, however, there were some colonies that remained completely positive for  $\beta$ -galactosidase. This implies either that the transfection process is incomplete or that the interference process is not universally effective; reasons for the latter might include those of siRNA dosage, RNAi resistance, or stochastic knockdown failure.

**hTERT-specific siRNAs fail to knockdown telomerase activity in human cancer cells**

A number of endogenous genes were chosen as candidate targets for siRNA-induced knockdown. The first of these was the human telomerase reverse transcriptase (hTERT) gene that is highly expressed in both stem cells and transformed cells (Ref: Weinberg R.A., 1998), and which appears to play a key role in the process of tumourigenesis (Ref: Hahn W.C. *et al.*, 1999a). A specific telomere terminal transferase activity was originally identified in *Tetrahymena* extracts (Ref: Greider C.W. and Blackburn E.H., 1985). This so-called telomerase activity was subsequently shown to be required for maintaining the integrity of the telomeric structures that reside at the ends of chromosome (Ref: Blackburn E.H., 1991). It comprises an RNA subunit (TER) and a reverse transcriptase subunit (TERT), with the latter providing the rate-limiting catalytic activity

(Ref: Poole J.C. *et al.*, 2001). The selective activation of the hTERT gene in the vast majority of human tumours has highlighted this gene as a promising target for cancer therapy (Ref: Hahn W.C. *et al.*, 1999b). To date, a variety of diverse strategies for inhibiting telomerase have been proposed (Ref: Kelland L.R., 2000). However, most of these suffer from problems of potency, specificity or delivery.

RNA interference, using sequence-specific siRNAs, represents an obvious means for achieving targeted knockdown of telomerase activity in the cell. This proposal was tested here using two different siRNA molecules, sharing sequence homology with the hTERT gene. siRNA-induced knockdown was examined in two immortalised cell lines, derived from human breast cancer, using a functional assay of telomerase activity. However, each siRNA on its own failed to demonstrate consistent and dose-dependent knockdown of telomerase activity although both siRNAs combined did possibly yield a small but non-statistically significant knockdown effect. Apart from reasons of technical failure, the explanation for the lack of success in this experiment is not clearly apparent. However, persistence of the pre-translated hTERT protein following siRNA transfection may be partly to blame. In order to test such a possibility, quantification of hTERT mRNA, by northern blot analysis or quantitative RT-PCR, rather than a functional assay of telomerase activity would be required.

#### **Hprt-specific siRNAs fail to induce permanent silencing the Hprt locus in ES cells**

While cells are able to synthesise nucleotides *de novo*, these are generally acquired exogenously by the cell from the environment. However, when purines and pyrimidines are in short supply, so-called salvage pathways become a major source of nucleotides. The enzyme hypoxanthine phosphoribosyltransferase (Hprt), which is carried on the X chromosome and therefore exists as a single copy in male ES cells, catalyses a key step in the salvage of purines. ES cells that are mutant for the Hprt locus, whether arising spontaneously or via a genotoxic insult, can be isolated using the purine analogue 6-thioguanine. This is converted by Hprt into a cytotoxic metabolite that kills wild-type cells and thereby allows negative selection of Hprt deficient cells.

In order to test the hypothesis that Hprt-specific siRNAs might occasionally produce permanent silencing of the Hprt locus, for instance via methylation or epigenetic modification, ES cells were selected with 6-thioguanine post-knockdown. However, this strategy failed to generate 6-thioguanine resistant clones in excess of the background rate using a control siRNA. Thus, Hprt-specific siRNAs targeting the coding region of the Hprt gene are apparently *unable* to produce a permanent silencing of the gene locus. Naturally, this assumes that there was no technical failing of the experiment.

The relevance of this result is placed in context by two recent papers that have demonstrated that gene knockdown can indeed lead to permanent gene locus silencing via CpG methylation (Refs: Kawasaki H. and Taira K., 2004; Morris K.V. *et al.*, 2004) associated with histone methylation (Ref: Kawasaki H. and Taira K., 2004) and histone deacetylation (Ref: Morris K.V. *et al.*, 2004). This notable difference between the current Hprt knockdown experiment and these papers above is that while the former used siRNAs directed against the coding sequence of the gene, the latter used promoter-directed siRNAs. However, the comparison is nonetheless noteworthy.

#### **Tk-specific siRNAs fail to induce permanent silencing the Tk locus in ES cells**

The salvage of pyrimidines involves the key enzyme thymidine kinase whose activity varies with the cell cycle being particularly important for the replication of DNA shortly prior to cell division. In an analogous fashion to previously, mutant cells lacking the enzyme can be selected using the nucleotide analogue in 5-bromodeoxyuridine, which is converted by thymidine kinase into a toxic compound leading to death of normal cells. Attempts to generate thymidine kinase deficient ES cells via siRNA-induced knockdown of the gene followed by 5-bromodeoxyuridine selection failed to produce resistant colonies in excess of the spontaneous background rate, thereby supporting the previous conclusion that siRNAs directed at the coding region of a gene are *unable* to produce permanent silencing of that locus.

#### **siRNA-induced knockdown of key salvage pathway genes in CHO cells causes apoptosis**

The targeted knockdown in CHO cells of the Tk and Hprt nucleotide salvage pathway genes, using Tk-specific and Hprt-specific siRNAs, unexpectedly resulted in widespread cellular apoptosis compared with non-specific control siRNAs. Crude measurement of the relative toxicities of these siRNAs, using semi-quantitative methods of counting surviving colonies, showed that Tk siRNA had 100x and Hprt siRNA had 10x the toxicity of control GFP siRNA on CHO cells. The possibility of non-specific toxicity arising as a by-product of the interference process itself was excluded by the fact that *Firefly*-specific siRNAs were not associated with any significant death of CHO-LUC cells. Assuming then that this toxicity reflects a gene specific effect, then it would appear both pyrimidine and purine nucleotide salvage pathways are critical for CHO cell survival, perhaps because of its high rate of cell division, although it is not clear why this is not the case for ES and 3T3 cells likewise.

#### **siRNA-induced epigenetic knockdown is not dependent upon methyltransferase activity**

It is well recognised in plants that post-transcriptional gene silencing is able to lead to the methylation of the targeted gene leading to epigenetic silencing of the locus (Ref: Baulcombe D., 2004). Similarly, and as alluded to previously, it has also been claimed in mammalian cells that

siRNAs directed at gene promoter sequences are able to initiate cytosine methylation within and around the targeted region (Refs: Kawasaki H. and Taira K., 2004; Morris K.V. *et al.*, 2004). However, it is not clear that this process is relevant for gene knockdown by coding sequence directed siRNAs. In order to test this, siRNA-induced gene knockdown was assessed in cell lines kindly provided by B. Ramsahoye (Ref: Jackson M. *et al.*, 2004) that were deficient either in maintenance methyltransferase activity ( $Dnmt1^{-/-}$ ) or *de novo* methyltransferase activity ( $Dnmt3a^{-/-}3b^{-/-}$ ). The results reveal that there is no significant compromise of knockdown efficacy in either  $Dnmt1^{-/-}$  or  $Dnmt3a^{-/-}3b^{-/-}$  ES cells, thereby confirming that methyltransferase activity is *not* necessary for suppression of a transiently expressed epigene. However, the possibility that siRNA-induced knockdown of endogenous genes is compromised in methyltransferase-deficient cells was not tested.

#### **Dnmt1-specific siRNAs result in significant genomic demethylation in wild-type ES cells**

It is known that murine ES cells carrying a targeted knockout of the Dnmt1 maintenance methyltransferase are deficient in genomic methylation with only 20-30% of CpG dinucleotides methylated compared with 70-80% in wild-type ES cells (Ref: Li E. *et al.*, 1992). This was the rationale for testing the demethylating potential of Dnmt1-specific siRNAs in wild-type ES cells. Using serial knockdowns to induce a cumulative effect, genomic DNA showed an overall reduction of cytosine methylation of approximately 16% compared with control cells. Thus, it is possible using this strategy to induce partial genomic demethylation similar to that seen with nucleotide analogues such as azadeoxycytidine. It is predicted that the combined knockdown of both maintenance and *de novo* methyltransferase activity, by targeting all three Dnmt genes active in mammalian cells, would have a much greater aggregate demethylating effect.

# **Chapter 7**

# **Discussion**

## 7.1 Introduction

The overarching goal of this thesis was the exploration of two novel strategies for elucidating gene function in mammalian cells. The first strategy, termed antisense gene trapping, was applied using a classical forward genetics approach to screen for differentiation genes in murine embryonic stem cells, while the second strategy, RNA interference, was applied using a reverse genetics approach to manipulate discrete genes both in ES cells and other cell lines. Arguably, while quite distinct from one another, both strategies share in common a key dependence upon double-stranded RNA to mediate silencing of a sequence-homologous gene. In antisense gene trapping, the double-stranded RNA is generated endogenously by hybridisation of the messenger RNA with an antisense RNA strand (produced through the action of the transactivator on the reversed gene trap promoter). By contrast, in RNA interference, the double-stranded RNA is introduced exogenously in the form of small RNA duplexes termed short interfering RNAs. The main results from the foregoing results chapters in regard to these two methodologies are discussed below.

## 7.2 Antisense gene trapping

The primary objective of chapter three, termed “antisense gene trapping in embryonic stem cells”, was the demonstration that such a methodology might permit the identification of putative differentiation genes in murine ES cells. The secondary objective was the functional characterisation of these genes through analysis of the gene trap and antisense knockout clones.

It should be stated that the formal proof-of-principle demonstration of the antisense gene trapping methodology was not explicitly sought at the outset of this work. The reason for this was simply the fact that antisense gene trapping is already established as a viable experimental strategy for achieving the functional deletion of random genes. The evidence for this is derived from a number of publications (Refs: Li L. and Cohen S.N., 1996; Xie W. *et al.*, 1998; Liu K. *et al.*,

1999; Liu K. *et al.*, 2000a; Liu K. *et al.*, 2000b) that have used antisense gene trapping to induce homozygous functional knockout of allelic loci in NIH3T3 cells for the purpose of identifying tumour susceptibility genes. Moreover, the primary objective as stated above was to use the methodology of antisense gene-trap to yield phenotypically mutant ES cell clones using a suitable differentiation assay. Only if and when such mutant ES cell clones were isolated was it intended to confirm knockout of the trapped gene in order to show that the phenotype was attributable to the mutation. Further confirmation could then be sought by restoring the ES clone to the wild type state through Cre excision of the transactivator.

The major failure of the methodology therefore lay not in the shortcomings of antisense gene trapping to knockout genes but rather those of methylcellulose differentiation as a screening strategy for identifying mutant ES cell clones. This failed to identify *any* mutant clones either through deficiencies of the screening method itself or simply due to too few “knockout” clones being analysed. Indeed, herein lies one of the key problems of forward genetics, namely the dependency upon an optimal mutant screen. Such a screen should exhibit attributes of discrimination, accuracy and reproducibility as well as being both high-throughput and non-labour intensive. These are not features of methylcellulose differentiation of ES cells.

Despite this being the case, it did prove possible to isolate a number of  $\beta$ -galactosidase positive gene trap clones that clearly became  $\beta$ -galactosidase negative upon introduction of the transactivator and then positive again upon introduction of the Cre recombinase. However, it is not clear that these results are necessarily attributable to antisense knockout of the trapped gene locus though binding with mRNA transcript. An equally plausible explanation for these results is that of promoter/transcriptional interference due to the movement of reverse-direction RNA polymerase moving counter to forward-moving RNA polymerase. Such a situation would result in loss of expression of the trapped gene, and therefore its  $\beta$ -galactosidase marker, due to interference with gene transcription or promoter initiation. The major difference between these

two alternative explanations for the observed loss of expression of  $\beta$ -galactosidase lies in the fact that while antisense knockout operates in *trans* to functionally delete the trapped gene, promoter interference operates in *cis* and will not therefore lead to gene knockout.

In order to discern experimentally between antisense knockout and promoter interference, it would be necessary to quantify either the level of mRNA transcript or alternatively that of the translated protein. For mRNA levels this would be by northern blot analysis or by RT-PCR, while for protein levels this would be by western blot analysis. In the case of antisense inhibition operating in *trans*, there should be complete biallelic knockout of gene expression resulting in minimal or absent levels of mRNA or protein. By contrast, in the case of promoter interference operating in *cis*, the knockout would only be monoallelic and therefore the mRNA and protein levels should only fall by approximately half compared with wild-type ES cells and not at all compared with the mother gene-trap clone that is already a hemizygous knockout. It is worth noting perhaps that in the original use of the antisense gene trapping technology (Ref: Li L. and Cohen S.N., 1996) the mutant 3T3 clones selected for anchorage independence by an agar assay did indeed show knockout of the trapped gene with loss of mRNA expression via northern blot analysis. In other words, antisense gene trapping appears to have been the correct explanation at least in this particular case using 3T3 cells.

As stated already, the feasibility of using antisense gene trapping to identify genes important for differentiation of ES cells was undermined by the inefficiency of the chosen phenotype screen. The methylcellulose differentiation assay, while well characterised for the analysis of haematopoietic and other lineage differentiation of ES cells, is largely a qualitative assay that is good for identification of gross abnormalities of differentiation but poor at picking up subtle ones requiring precise quantitative analysis. Moreover, the methylcellulose assay is highly labour intensive making it less practical from an experimental perspective as a screening strategy for mutant clones, especially in the context of comparing gene trap clones with antisense knockout

sub-clones that each need to be performed in triplicate for purposes of statistical validation. By contrast, in the original application of the antisense gene trapping approach, for identification of putative tumour susceptibility genes in NIH3T3 cells, the screening strategy was growth in simple agar allowing for the emergence of knockout clones that had lost anchorage dependence.

What alternative screening strategies might be applicable to the task of identifying mutant ES cell clones? In essence, there are two possible approaches. Either to use a phenotypic screen, of which methylcellulose despite its drawbacks is arguably the best characterised, or else to use a genotypic screen. This would seek to show altered expression levels, for instance by RT-PCR, of key genes known to be important for ES cell pluripotency, of which Oct3/4 is perhaps the best such candidate. Alternatively, genes known to up-regulated by induction of differentiation could be analysed perhaps by micro-array analysis of clones. However, of the two approaches, the phenotypic screen is undoubtedly more powerful and one that allows selection of ES cell clones exhibiting a newly acquired phenotypic trait. Unfortunately, for the purpose of the current study, such a screen remained an elusive goal.

In an attempt to circumvent the lack of an effective screen, the original strategy was modified thus. Instead of subjecting all clones to methylcellulose differentiation, only those clones for which the trapped gene yielded to 5'RACE analysis and were deemed suitable candidates for further study were scrutinised. This avoided needless assays on "dead-end" clones and at the same time allowed for a more thorough assessment of differentiation. Moreover, it took advantage of the rapid completion and annotation of the murine genome that was occurring throughout the project. However, as previously, this approach also failed to identify differentiation mutants. It did nonetheless allow for a more refined phenotypic analysis of a number of gene trap clones. The major findings of interest arising from these clones are summarised below.

### 7.3 Analysis of gene trap genes

Using the modified approach described above, whereby genes-of-interest were selected according to 5'RACE identification of the trapped gene sequence, a total of five gene trap clones were selected for closer functional analysis. These were *Tsix*, *Catns*, *Rex2*, *Nasp*, and *Ctbp2*. It should be noted that the possibility of multiple insertions of the gene search vector in any of these clones was not explored. It is known that gene trap vectors are prone to integrate at multiple genomic sites leading to problems of interpretation of phenotypic outcomes. While such a situation can only be formally ruled out by Southern blotting of genomic DNA, the finding of a single 5'RACE band on gel electrophoresis would appear to argue against multiple insertions in the case of these particular five clones listed above. However, multiple insertions of the vector do represent a genuine possibility that could potentially affect the interpretation of phenotypic findings described here.

All five of the clones exhibited  $\beta$ -galactosidase positivity, and four out of five yielded derivative  $\beta$ -galactosidase negative knockout sub-clones; the *Catns* gene trap clone failed to do so for reasons not apparent. The most significant findings arising from analysis of these clones are listed below:

- (1) The gene trap insertion into the *Tsix* gene results in insertional mutagenesis of the *only* copy of this gene in male ES cells. It is not clear, however, that *Tsix* performs any specific function in male cells that are not required to undergo X chromosome inactivation. *Tsix* is a 40-kb RNA originating 15kb downstream of the *Xist* which is known to regulate X chromosome dosage compensation by controlling X inactivation in collaboration with a *cis*-acting element called the X inactivation centre (Ref: Lee J.T. *et al.*, 1999). At the onset of X inactivation, *Tsix* expression becomes monoallelic, is associated with the future active X and persists until *Xist* is turned off. *Tsix* is not found on the inactive X once cells enter the X inactivation pathway suggesting that it plays an important role in regulating the early steps of X

inactivation. This is supported by the observation that targeted deletion of *Tsix* gene in female mouse ES cells results in non-random inactivation of the mutant X chromosome (Ref: Lee J.T. and Lu. N., 1999). Thus, in the mouse it appears that *Tsix* regulates *Xist* in *cis* and thereby determines which X chromosome is to become silenced (Ref: Sado T. *et al.*, 2002).

Perhaps the most interesting insight from this gene trap clone comes from the  $\beta$ -galactosidase expression. Since *Tsix* is an RNA gene whose putative function is mediated via antisense binding of the RNA molecule to *Xist*, there is no translation of the gene into protein (Ref: Ogawa Y. and Lee J.T., 2004). Hence, it is unclear why the  $\beta$ -galactosidase of the gene search vector, or for that matter the neomycin phosphotransferase that confers G418 resistance, should be expressed at all. The fact that it is suggests that the mRNA, either despite the gene trap integration or because of it, is indeed translated into protein resulting in uniform and ubiquitous expression of  $\beta$ -galactosidase in all cells. If this is attributable to upstream sequence elements within the mRNA transcript, then it suggests that *Tsix* RNA is able to translocate to ribosomes in the cytoplasm and is not solely contained within the nuclear compartment. However, this conclusion remains speculative.

- (2) The three genes, *Rex2*, *Nasp*, and *Ctbp2*, all exhibited specific loss of  $\beta$ -galactosidase expression within erythroid cells during haematopoietic differentiation of embryoid bodies. It is not clear whether this is simply an artefact of the assay, possibly related to failure of the X-gal substrate to act properly in haemoglobinised cells, or reflection of true lineage restricted expression of these genes. In support of the latter is the fact that *Catn5* is expressed in erythroid cells and serves here as a convenient positive control to the experiment. In addition, in the case of *Ctbp2*, it is known that this is a negative regulator of the *Gata1*-associated transcription factor *Fog1* that is highly expressed in erythroid cells. Thus, down-regulation of *Ctbp2*, as reflected by the absence of  $\beta$ -galactosidase activity in haemoglobinised cells, may allow up-regulation of *Fog1* leading to normal erythropoiesis.

C-terminal binding protein 2 resides on the mouse chromosome 21 and was originally identified on the basis of its sequence similarity to the Ctbp1 gene whose protein product is known to associate with the C-terminus of the adenoviral oncoprotein E1A via a Pro-X-Asp-Leu-Ser motif (Ref: Katsanis N. and Fisher E.M., 1998). Unlike Ctbp1, which is expressed from embryo to adult, Ctbp2 is expressed mainly during embryogenesis (Ref: Furusawa T. *et al.*, 1999) and appears to be critical for axial patterning leading to embryonic lethality in mutant null mice (Ref: Hildebrand J.D. and Soriano P., 2002). The characterisation of the CtBP2 protein has revealed that it functions as a potent co-repressor for a number of transcriptional factors including the Basic Kruppel-like factor (BKLF) that is highly expressed in erythroid cells (Ref: Turner J. and Crossley M., 1998), and that this action is mediated via binding to a minimal repression domain (Ref: Thio S.S. *et al.*, 2004). CtBP2 has also been shown to interact with the hematopoietic, zinc-finger protein FOG1 that is essential for the development of the erythroid and megakaryocytic lineages (Ref: Katz S.G. *et al.*, 2002). Interestingly, FOG1 is highly expressed in erythroid cells and acts in concert with the transcriptional factor GATA1 to regulate erythropoiesis. Given the known interaction between CtBP2 and FOG1 (Ref: Katz S.G. *et al.*, 2002), this reciprocal expression pattern in erythroid cells is perhaps not surprising and suggests that the lifting of CtBP2-mediated repression of FOG1 is an important step in the initiation of normal erythropoiesis.

- (3) Expression of the Rex2 gene, during LIF-withdrawal differentiation, appears to result in persistence of the undifferentiated ES state. This conclusion is inferred from the fact that maintenance of G418 selection pressure on the gene trap cells, thereby “enforcing” expression of the Rex2 locus carrying the gene trap vector, prevents normal ES differentiation. Instead of fibroblasts emerging, the presence of low dose G418 causes the ES cells to remain in small colonies of undifferentiated cells even despite the absence of LIF in the medium. Thus, it appears that G418, either through enforced expression of the trapped Rex2 gene and the downstream neo that confers resistance or through antibiotic-induced death of those cells that

switch off Rex2 and neo expression, results in persistence of the undifferentiated ES state. The rationale for this experiment was based upon previously published work showing that ES cells carrying a transgenic neo gene, driven by the  $\alpha$ -cardiac myosin heavy chain promoter, produced a homogenous population of cardiomyocytes when differentiated in the presence of G418 due to selection of cardiac lineage-restricted cells (Ref: Klug M.G. *et al.*, 1996).

Thus, Rex2 arguably represents a candidate stem cell gene in ES cells. This conclusion is consistent with published data showing that induction of differentiation in F9 teratocarcinoma cells, using retinoic acid, results in down-regulation of Rex2 (Ref: Faria T.N. *et al.*, 1998). Indeed, this was the basis by which the gene was originally identified using differential hybridisation. Rex2 has also been identified in a large-scale screen of mammalian proteins residing in nuclear sub-compartments (Ref: Sutherland H.G. *et al.*, 2001). Note that despite its name, Rex2 does not appear to bear any functional or evolutionary ties with Rex1, which is a zinc finger gene also regulated by retinoic acid and a commonly-used marker of undifferentiated ES cells.

Thus, Rex2 appears to represent a candidate differentiation gene whose expression correlates with an undifferentiated state in ES cells. Such a claim is obviously contentious but consistent with the data presented. Alternatively, it is possible that the findings are artefactual and arise due to the artificial nature of the experimental conditions in which ES cells are induced to differentiate in the presence of G418 that might otherwise limit cell proliferation. Nonetheless, the hypothesis that Rex2 is indeed an important gene whose expression helps to maintain ES cells in an undifferentiated state is a testable one. For instance, it predicts that the targeted knockout of Rex2 is not compatible with ES cells or their derivation from the inner cell mass of the embryo. The converse prediction that down-regulation of Rex2 might lead to initiation of differentiation might be similarly tested using RNA interference to knockdown expression of the Rex2 gene.

## 7.4 RNA interference

The use of antisense gene trapping to achieve functional gene knockout was originally described in 1996. This pre-dates the discovery of RNA interference in 1998 and the first application of short interfering RNAs in mammalian cells in 2001. With hindsight, it is intriguing to speculate, as suggested previously, that both processes share a common molecular mechanism that requires the presence of double-stranded RNAs. In the case of antisense gene trapping, these are generated endogenously within the cell, while in the case of RNA interference, they are introduced exogenously into the cell.

Methods of genetic analysis are necessarily contingent upon the timing of biological discoveries and this is reflected in this thesis by the eventual abandonment of antisense gene trapping in favour of the emerging technology of RNA interference for the purpose of knocking down gene expression. The major rationale for switching methodologies was the realisation that antisense gene trapping was fatally undermined by the lack of an effective screening strategy and that the goal of gene knockdown could be more easily achieved using RNA interference. One additional factor that cannot be overlooked in this decision process was the rapid completion and annotation of both the human and mouse genomes that is necessary for the sequence data required for designing short interfering RNAs. However, as emphasised previously, it is important not to overlook the fact that while both antisense gene trapping and RNA interference aim to discover gene function through gene knockdown, they are conceptually very different approaches with the former being phenotype-driven and the latter genotype-driven. As such, this change of direction of the thesis was driven as much by experimental pragmatism as it was by the original aims of the research to identify key differentiation genes. In addition, given the infancy of RNA interference at the time of embarking upon this work, particularly in relation to its applicability to ES cells, it was recognised that preliminary groundwork in other cell lines would be necessary to establish key aspects of this new methodology.

## 7.5 Gene knockdown using siRNAs

The experimental results presented in this thesis provide useful insights into the nature of the RNA interference mechanism and shed light on some of the molecular determinants of siRNA-induced gene knockdown in mammalian cells. For instance, the requirement for equimolar amounts of complementary sense and antisense oligonucleotide RNA strands testifies to the fact that gene interference is mediated via a double-stranded RNA molecule, although an explicit annealing step prior to siRNA transfection does not appear to be obligatory (see figures 4.1, 4.2, 4.3 and 4.4). The potency of gene knockdown appears to be partly dependent upon the actual sequence of the siRNA and also upon the site of binding within the target gene, and this exhibits dose-dependency with sigmoid-shaped response curve (see figures 4.5, 4.6, 4.7 and 4.8). It is worth commenting that while these findings are now well established, at the time of embarking on this work most aspects of siRNA function in mammalian cells was simply inferred from earlier work using long dsRNA molecules in non-mammalian organisms.

While key design features for effective siRNA action were not fully explored, some tentative conclusions can be drawn from the results of these experiments presented in this thesis. Many, if not all, of these findings have been corroborated by other published analyses (Refs: Tomari Y. and Zamore P.D., 2005) and also by commercially available resources (see Dharmacon Inc. and Xeragon Inc. company websites). They include the avoidance of high G/C content or runs of G/C within the siRNA sequence and also the avoidance of palindromic sequences that might allow same strand base pairing or intra-strand folding. The presence of a pair of G/C bases at the 5' end of the siRNA and a pair of A/U bases at its 3' end appears to enhance the interference effect and confirm the functional non-symmetry of the siRNA molecule (Ref: Tomari Y. *et al.*, 2004). Moreover, this non-symmetry suggests differential roles of the sense and antisense RNA strands. For instance, it supports the notion that unzipping of the double-stranded RNA molecule (required for hybridisation of the antisense strand with the target gene) occurs in a 3' to 5' direction.

Another pertinent observation was the fact that combinations of siRNAs, each targeting different sites within the same gene, do not enhance the knockdown effect achieved by using a single siRNA (see figure 4.9). This finding is perhaps somewhat intuitive if indeed the interference effect occurs through siRNA-directed cleavage of the mRNA molecule when a single site would clearly suffice in rendering the mRNA functionally useless.

The demonstration that non-complementary sense and antisense strands, that cannot form a double-stranded RNA molecule, are unable to produce gene knockdown was used as evidence to argue against so-called “transitive RNAi” operating in mammalian cells (see figure 4.10). The basis for this conclusion assumes that each of the non-complementary RNA strands might act as guide siRNAs able to prime RNA-dependent RNA polymerisation of the mRNA template. In so doing, they would generate a double-stranded RNA molecule that could then initiate gene interference. Such a mechanism appears to be physiologically relevant in *C. elegans* but not in *Drosophila* or mammalian systems. However, in *C. elegans*, transitive RNAi requires the presence of a pair of double-stranded RNAs (Ref: Alder M.N. *et al.*, 2003) to occur rather than a pair of single-stranded RNAs as used here. Hence, this experiment, and any conclusions derived from it, is potentially flawed as it is arguably based upon the wrong model of transitive RNAi.

Analysis of the dynamic aspects of the siRNA-induced interference revealed that gene knockdown occurs commensurately with expression of the co-transfected reporter such that there is a discernible dose-dependent suppression at every time-point measured from four hours onwards (see figure 4.11). Notably, there is no lag period of the interference effect relative to reporter gene translation implying that the siRNAs act at least as rapidly as the epigene is expressed. This argues against the requirement for induction of the necessary effectors of RNA interference and their accumulation in the cell. Rather, it suggests that the interference machinery is already primed within the cell, possibly for the purpose of handling endogenously produced miRNAs, and is simply hijacked by the transfected siRNAs. Such a conclusion also comes from the

demonstration that siRNA-induced knockdown is overcome in a competitive manner by increasing amounts of co-transfected epigene (see figures 4.12 and 4.13). Indeed, in this regard, RNA interference might be considered akin to a chemical reaction with a dosage-response curve that is reminiscent of classical first order kinetics. Moreover, pre-transfection of the siRNA, up to 48 hours prior to the plasmid, leads to gene knockdowns of the expected magnitude (see figures 4.14 and 4.15), thereby indicating that the siRNA molecules themselves, whether free or bound to the RNA silencing machinery, possess considerable cellular stability.

The demonstration that siRNAs can knockdown a stably integrated reporter gene in an apparently similar fashion to a transiently expressed one suggests that the molecular mechanisms are closely analogous to one another if not identical (see figure 4.16) with persistence of many days (see figure 4.17). The unexpected finding of paradoxical “enhancement” of luciferase activity when the cells were challenged with high doses of RNA was shown to have temporal and dosage profiles that were markedly different to that of true RNA interference (see figures 4.18 and 4.19). While this phenomenon was not fully characterised, it is tempting to speculate that it represents an artefactual measurement of luciferase activity arising due to widespread cellular apoptosis possibly triggered by the non-specific interferon response.

Attempts to improve the liposomal delivery of siRNA molecules into cells were based upon the use of short peptides derived from the transduction domain of the HIV-1 TAT protein. The results suggest a small but statistically significant enhancement of transfection efficiency, as judged by the level of reporter gene knockdown, but the method requires further characterisation and optimisation (see figure 4.22). The other technical aspect of knockdown efficiency that was explored in this study was the use of *in vitro* transcribed siRNAs. In our hands, these failed to show either the consistency or potency of their chemically synthesised counterparts (see figures 4.20 and 4.21) and hence were not adopted in subsequent experiments.

## 7.6 Molecular biology of siRNAs

The siRNA/target-gene co-transfection method used in this thesis provided a versatile strategy for investigating mechanistic aspects of the RNAi process facilitating a piecemeal dissection of the interference effect in mammalian cells. This shed light on some of the molecular determinants of effective siRNA design and yielded provocative insights into the level at which the siRNA molecule acts to suppress gene expression. Using chimaeric siRNA duplexes, it was shown that sequence fidelity of the antisense strand is paramount while fidelity of the sense strand is less critical (see figures 5.1, 5.2 and 5.3). Thus, antisense mismatch at three discrete nucleotide sites within the siRNA molecule abrogates the interference effect while the same mismatch in the sense strand is well tolerated (Refs: Amarzguioui M. *et al.*, 2003, Schwarz D.S. *et al.*, 2003, Holen T. *et al.*, 2005).

This important finding suggests that the role of the sense strand is principally to serve as a chaperone for its complementary antisense partner by annealing with it to produce an RNA duplex that is less susceptible to the action of endogenous RNases. Subsequent unravelling of the duplex would then allow the antisense strand to recognise and base pair with the target mRNA leading to sequence-directed cleavage of the latter and post-transcriptional knockdown of the gene. It is also possible that this base pairing occurs with the nascent transcribed mRNA strand, while still attached to the DNA and prior to RNA processing, thereby inhibiting gene expression at a pre-transcriptional level. A third conceivable possibility would involve the antisense strand recognising and base pairing with the non-coding strand of the DNA, leading in some manner to inhibition of gene expression at the transcriptional level. The theoretical possibility that the sense strand might play an analogous role by base pairing with the coding strand of the DNA seems to be inconsistent with the observation that sequence fidelity of the sense strand is not an absolute requirement for successful gene knockdown to occur.

In mammalian cells, prior 5'-terminal phosphorylation of the siRNA molecule did not have any discernible effect on the knockdown of a co-transfected plasmid-borne gene (see figure 5.3). This implies either that 5'-phosphorylation is physiologically irrelevant for effective siRNA action or alternatively that mammalian cells contain their own endogenous 5'-nucleotide kinase activity able to perform this function post-transfection. There was also no evidence in these experiments to support the view that 5'-phosphorylated antisense strands on their own (and not duplexed with a sense strand partner) also possess gene knockdown potential (see figures 5.2 and 5.3) as suggested by other investigators (Ref: Martinez J. *et al.*, 2002).

A modification of the siRNA/target-gene co-transfection method, using an mRNA reporter molecule in place of the plasmid-borne reporter, was employed to challenge the paradigm-held view that interference acts at a post-transcriptional level. This revealed the unexpected result that siRNAs that are effective at knocking down a DNA-encoded gene are unable to do likewise for an RNA-encoded gene, in both mammalian CHO and arguably 3T3 cells (see figures 5.4, 5.5, 5.6 and 5.7). The simplest interpretation of this finding is that the siRNA-induced interference can be bypassed using mRNA that is pre-transcribed. The converse explanation is that siRNA action is dependent in some fashion upon the presence of the DNA. Either way, this evidence argues that in the mammalian cell lines tested, the site-of-action of the siRNA is not at a post-transcriptional level but rather at a transcriptional or peri-transcriptional level instead.

However, the non-physiological nature of the previous experiment raises the possibility that the outcome described is simply an artefact that does not reflect the true situation operating within the cell. For instance, the exogenous RNA reporter molecule may behave unlike the normally transcribed endogenous mRNA molecule for reasons of spatial, temporal, or chemical differences between the two. In order to address this possibility, it was necessary to validate the experimental strategy of siRNA/mRNA co-transfection in mammalian cells by showing that this could indeed yield gene knockdown in non-mammalian cells.

Validation of the siRNA/mRNA co-transfection experimental system was achieved by using two *Drosophila* cell lines in place of the mammalian ones used above. This revealed that *Drosophila* KC cells do indeed produce a potent and unequivocal knockdown of the RNA-encoded reporter gene by co-transfected siRNAs that is qualitatively identical to the result obtained using the DNA-encoded gene (see figure 5.8). Strangely, the same is not also true of *Drosophila* S2 cells which, like CHO and 3T3 cells, exhibit siRNA-induced knockdown of the DNA-encoded but not RNA-encoded reporter gene (see figure 5.8). Only by using 5'-phosphorylated siRNAs can S2 cells be made to behave similarly to KC cells and produce knockdown of *both* DNA and RNA-encoded reporter genes, suggesting that this is an important pre-requisite for siRNA action. If so, then it points to a higher level of endogenous 5' nucleotide kinase activity in KC cells relative to S2 cells (see figures 5.12, 5.13 and 5.14). Thus, siRNAs in *Drosophila* cells do appear to act at a post-transcriptional level, but only if they are 5'-phosphorylated.

By contrast with the above result in *Drosophila* cells, 5'-phosphorylation of the siRNA molecule fails to make any significant difference in mammalian cells (see figures 5.12, 5.13 and 5.14). Here, both 5'-phosphorylated and non-phosphorylated siRNAs behave identically to one another in terms of their ability to knockdown DNA-encoded but not RNA-encoded reporter genes. Thus, there is a clear qualitative difference between the cells derived from the two species.

A further important qualitative difference between cells of the two species is apparent with the use of mismatched heteroduplex siRNAs described previously (see figures 5.15, 5.16 and 5.17). Here, in mammalian CHO and 3T3 cells, mismatch in the sense strand of the siRNA duplex is well tolerated so long as there is fidelity of the antisense strand. By contrast, in *Drosophila* KC and S2 cells, these same heteroduplex siRNAs fail to produce any significant gene knockdown. However, if they are first modified by 5'-phosphorylation of their terminal ends then their knockdown action is restored in both KC and S2 cells, but only against the DNA-encoded and not RNA-encoded reporter gene.

Thus, at risk of over-interpreting this data, the knockdown action of (antisense-fidelitous/sense-mismatch) heteroduplex siRNAs against DNA-encoded genes appears to be 5'-phosphate independent in mammalian cells but 5'-phosphate dependent in *Drosophila* cells, with no demonstrable knockdown against RNA-encoded genes in either mammalian or *Drosophila* cells irrespective of 5'-phosphorylation status. These findings are consistent with other published evidence (Refs: Martinez J. and Tuschl T., 2004; Schwarz D.S. et al., 2002; Nykanen A. *et al.*, 2001). However, further studies are warranted to confirm that these results are experimentally robust and applicable equally to other siRNA molecules and cell lines.

### 7.7 A model of siRNA action

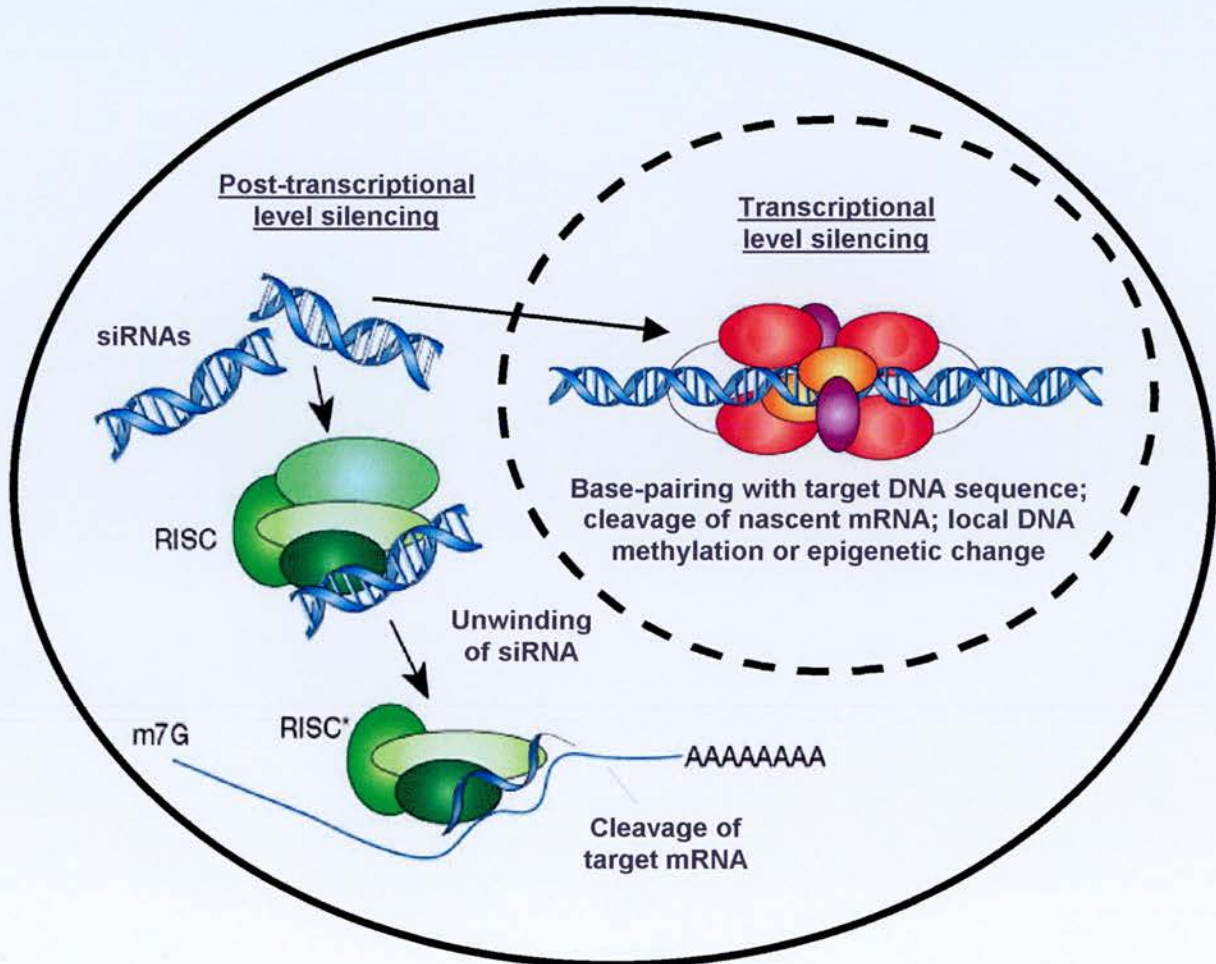
How is it possible to interpret these inter-species differences in terms of a mechanistic understanding of the mode of action of siRNA molecules to induce gene interference in cells? One interpretation argues that siRNA molecules can affect gene expression at both the DNA transcriptional level and at the mRNA post-transcriptional level, and that in any given cell one or other of these mechanisms predominates. Thus, in mammalian CHO and 3T3 cells, transcriptional interference prevails while in *Drosophila* KC cells, post-transcriptional interference prevails; *Drosophila* S2 cells behave similarly to KC cells but only if 5'-phosphorylated siRNA molecules are used. Finally, mammalian and *Drosophila* cells behave differently to one another in regard to heteroduplex siRNA molecules that have sense-strand mismatch and antisense-strand fidelity. These produce transcriptional interference but not post-transcriptional interference in both mammalian and *Drosophila* cells, but in the latter only when 5'-phosphorylated.

In summary, post-transcriptional level interference appears to be the predominant mechanism of siRNA-induced gene silencing in *Drosophila* KC cells. By contrast, in *Drosophila* S2 cells post-transcriptional level interference occurs but only if the molecule is 5'-phosphorylated. This suggests that post-transcriptional level interference requires 5'-phosphorylation of the siRNA

molecule and that this occurs endogenously in KC cells but not S2 cells. The situation in mammalian cells is qualitatively different. Here, transcriptional level interference appears to be the predominant mechanism of siRNA-induced gene silencing, and this is irrespective of siRNA 5'-phosphorylation. This transcriptional level interference appears to be remarkably tolerant of sense-strand mismatch so long as the antisense strand of the heteroduplex molecule retains sequence fidelity. The same heteroduplex siRNAs produce gene silencing in *Drosophila* cells only if 5'-phosphorylated when transcriptional level, and *not* post-transcriptional level, gene silencing is seen, suggesting that this pattern is perhaps a hallmark feature of transcriptional level knockdown.

There is a risk in constructing this model of over-interpretation of the data and hence the model remains purely speculative at this stage pending further experiments. However, the demonstration of siRNA-induced knockdown of an mRNA luciferase reporter molecule in *Drosophila* KC cells not only validates the experimental methodology but reveals the absence of knockdown in CHO cells to be apparently genuine. If so, then the simplest interpretation is that siRNA knockdown in mammalian cells is not solely post-transcriptional as previously assumed but also transcriptional. A schematic model illustrating this proposed mechanistic dichotomy of siRNA-induced gene silencing at post-transcriptional and transcriptional levels is shown overleaf (Figure 7.1).

**Figure 7.1: Schematic model of the mechanism of action of siRNAs in gene silencing in eukaryotic cells.** siRNA molecules that are introduced exogenously into mammalian cells appear to be able to mediate gene silencing at two different levels. The first is at a post-transcriptional level via cleavage of the target mRNA molecule within the site of binding. The other is at a transcriptional level via mechanisms that are less well defined and may include base-pairing with the DNA, cleavage of the nascent mRNA strand, and induction of epigenetic changes such as CpG methylation and histone modification that cause heterochromatinisation.



Post-transcriptional level gene silencing:

- 5'-phosphorylation obligatory?
- *Intolerant* of sense strand mismatch?
- Predominant in *Drosophila* KC cells and in S2 cells if 5'-phosphorylated

Transcriptional level gene silencing:

- 5'-phosphorylation *not* obligatory?
- *Tolerant* of sense strand mismatch?
- Predominant in mammalian cells and in S2 cells if non-phosphorylated

## 7.8 Probing ES cells with siRNAs

The original intention of the experimental work presented above relating to the biology of RNA interference in mammalian cells was to explore the application of this technology for the purpose of analysing gene function specifically in embryonic stem cells. As outlined previously, the knockdown of discrete genes using sequence specific siRNAs is a versatile tool for probing gene function since it permits a tight correlation between phenotype and genotype. Given appropriate methods of phenotypic analysis in the case of embryonic stem cells, siRNA technology represents a powerful method for dissecting out the function of differentiation genes that play key roles during early development. The demonstration that siRNAs operate to induce gene knockdown in embryonic stem cells, as they do in other more tractable mammalian cell lines, was the final goal of this thesis.

A stepwise approach was taken with murine ES cells: firstly, the demonstration that siRNAs can induce knockdown of a transiently expressed luciferase epigene in a manner that was qualitatively identical to other mammalian cell lines (see figure 6.2); secondly, that siRNAs can induce knockdown a stably expressed GFP transgene introduced by random genomic integration (see figure 6.3); and thirdly, that siRNAs can induce knockdown of stably expressed  $\beta$ -galactosidase transgene that had been targeted to the physiologically active Oct3/4 locus (see figure 6.5). This latter experiment also served to confirm that the siRNA-induced knockdown is clonally inherited through several cell divisions, that knockdown is specific to the gene targeted and not the allelic locus, and that failure of knockdown arises stochastically for reasons either technical or biological (see figure 6.6).

Thus, it is clear that RNA interference does indeed operate in murine ES cells and this conclusion is consistent with other evidence published contemporaneously with this work (Refs: Yang S. *et al.*, 2001; Zou G.M. *et al.*, 2003; Hay D.C. *et al.*, 2004; and Tang F.C. *et al.*, 2004).

In light of the earlier results suggesting that siRNAs act at a transcriptional level in mammalian cells, experiments were devised to investigate the possibility that siRNA-induced gene knockdown might lead to permanent epigenetic silencing of the targeted locus. These experiments aimed to derive mutant “knockout” clones using negative selection of ES cells transfected with siRNAs specific for the purine and pyrimidine salvage pathway genes, Hprt and Tk respectively. In neither case were such clones isolated (in excess of the spontaneous background rate) indicating failure of siRNA-induced permanent silencing of the targeted gene locus (see figure 6.8). Importantly, these experiments used coding sequence siRNAs and not promoter sequence siRNAs used by other investigators to induce epigenetic modifications such as DNA methylation (Refs: Morris K. V. et al., 2004; Kawasaki H. and Taira K., 2004).

When the foregoing experiment was performed in CHO cells, there was widespread cellular apoptosis caused by Tk-specific siRNAs (see figure 6.9) and to a lesser extent Hprt-specific siRNAs (see figure 6.10). This finding validated the biological activity of both these siRNAs, compared with non-specific control siRNAs, and suggests critical dependency of CHO cells on salvage pathway synthesis of nucleotides that is not the case for ES cells. If this result is also borne out in tumour cell lines, then targeting of the Tk and Hprt genes using sequence-specific siRNAs represents a feasible anti-cancer strategy that would target malignant cells but spare normal cells. Such was the rationale likewise for targeted knockdown of telomerase activity in human breast cancer cell lines using hTERT-specific siRNAs (see figure 6.7), although for reasons that were not fully explored this experiment proved unsuccessful.

With regard to ES cells, three endogenous genes were targeted using siRNAs. Two of these were putative stem cell genes expressed in undifferentiated ES cells, namely Ebox (Ref: Jackson M. *et al.*, 2002) and Esg1 (Ref: Tanaka T.S. *et al.*, 2002). However, for both these genes, sequence-specific siRNAs failed to induce any discernible differentiation of ES cells grown in LIF-containing medium (data not shown) and hence further experiments were not undertaken.

The third endogenous gene chosen for siRNA-induced gene knockdown was the maintenance methyltransferase gene Dnmt1, which is responsible for maintaining normal levels of genomic methylation in undifferentiated ES cells. Here, serial transfection of ES cells using sequence-specific siRNAs resulted in a small but significant loss (approximately 16.4%;  $p < 0.001$ ) of methylated cytosines residing within CpG dinucleotides (see figure 6.12). Thus, siRNA-induced knockdown of Dnmt1 can induce modest genomic demethylation of ES cells even in the face of continued expression of the *de novo* methyltransferase genes Dnmt3a and Dnmt3b. The use of a combination of siRNAs that target all three of the DNA methyltransferase genes expressed in ES cells might be expected to produce a more potent demethylating effect. Demethylating agents such as azadeoxycytidine and zebularine have been shown to induce “reprogramming” of differentiated cells and reversion of transformed cells suggesting that siRNA-induced knockdown of Dnmt genes, commonly over-expressed in tumour cells, represents a further potential anti-cancer application of the RNAi technology (Refs: Gowher H. and Jeltsch A., 2004; Esteller M. 2005; Laird P.W., 2005).

### **7.9 Where is the siRNA site-of-action?**

The experiments presented in this thesis sought to cast light on how siRNA molecules induce gene silencing in mammalian cells. In particular, they were designed to determine the site-of-action of the siRNA within the molecular chain of events that lead to gene expression. In this regard, do the siRNA molecules act a post-transcriptional level by targeting the mRNA or at a transcriptional or peri-transcriptional level involving the DNA possibly in association with the nascent RNA strand. Clearly, the experimental results presented fail to resolve this issue but do nonetheless raise some intriguing questions of their own. In particular, they uncover qualitative differences between mammalian and *Drosophila* cells as well as intra-species differences between the cell lines. Most notably, this relates to the discrepant way in which mammalian CHO cells and *Drosophila* KC cells respond to co-transfection of the siRNA molecules together with the target mRNA reporter

molecule: while in the KC cells the siRNAs silence the reporter, in CHO cells there is no discernable reporter knockdown. This major qualitative difference points to fundamental mechanistic discrepancies in the manner of action of siRNAs in mammalian cells as compared with *Drosophila* cells. Obviously, it remains possible that these results have arisen on account of the artificial nature of the experiment, for instance how the two species' cells handle transfected mRNAs, but the fact that there are differences also between the two *Drosophila* lines would perhaps argue against this.

How might this question of siRNA site-of-action be resolved? The most important experiment that was not performed here is the nuclear run-on assay. This would identify the precise site of action of the siRNA by distinguishing a transcriptional level of interference from a post-transcriptional/translational level based upon incorporation of radio-labelled nucleotides into nascent mRNA strands in isolated nuclei and subsequent autoradiography. In the former case, mRNA transcripts would cease to be produced upon siRNA action while in the latter, they would continue to be produced and then destroyed by siRNA-induced strand cleavage. Thus, the nuclear run-on assay represents a key future experiment that might be applied to the siRNA/plasmid co-transfection knockdown assay.

Another obvious strategy for discerning the site of action of the siRNA-induced interference might be to knockdown endogenous genes using siRNAs that target splice junction sites that bridge introns. Such siRNAs, it would be predicted, should only induce gene knockdown if their site-of-action occurs after splicing of the hnRNA into mRNA. In other words, it would confirm a post-transcriptional site-of-action. An alternative approach would be to use siRNAs to target non-coding promoter regions that are not transcribed into mRNA and therefore not relevant to post-transcriptional gene knockdown. Indeed, such an experiment has been performed and shown to lead to promoter methylation associated with gene silencing (Refs: Morris K. V. et al., 2004; Kawasaki H. and Taira K., 2004).

However, the phenomenon of siRNA-induced promoter methylation in mammalian cells remains somewhat contentious. The debate has been fuelled lately by the partial retraction of one of the papers originally claiming to show this phenomenon (Ref: Kawasaki H. and Taira K., 2006). Interestingly, one final experiment presented in this thesis might have some bearing on this matter. This is the demonstration that siRNAs are able to induce knockdown of a luciferase reporter gene even in cells lacking DNA methyltransferase activity through targeted deletion of both the maintenance methyltransferase Dnmt1 or the *de novo* methyltransferases Dnmt3a and Dnmt3b (see figure 6.11).

Thus, based on these results, siRNA-induced knockdown in mammalian cells does not appear to be contingent upon methyltransferase activity. A similar conclusion has been borne out by other investigators who have targeted promoter regions of endogenous genes using the same methyltransferase deficient ES cells as used here (Ref: Ting A.H. et al., 2005).

## 7.10 Conclusion

This thesis aimed to explore two novel strategies, antisense gene trapping and RNA interference, for elucidation of gene function in mammalian cells and in particular sought to apply them to biological analysis of the murine embryonic stem cell. The antisense gene trapping strategy was applied using a classical forward genetic approach in an attempt to identify putative differentiation genes in ES cells. By contrast, the RNA interference strategy was applied using a reverse genetic approach using sequence-specific short interfering RNAs to manipulate discrete genes both in ES and other mammalian cells. Thus, antisense gene trapping is essentially a phenotype-based strategy while RNA interference is essentially a genotype-based strategy. The former is inherently a more powerful approach for genetic analysis but is wholly dependent upon the effectiveness of the phenotypic screen. The latter, by contrast, requires prior gene sequence knowledge and as such is a timely new technology in the post-genomic era.

The functional dissection of the mammalian genome represents the overarching goal of contemporary molecular biology. Antisense gene trapping and RNA interference represent two complementary methodologies, each with its own strengths and weaknesses relative to the other. At the heart of both these processes is double-stranded RNA suggesting that there might exist a commonality in their respective molecular mechanisms. Such a possibility, namely that RNA antisense and RNA interference represent different facets of the same cellular process, represents an intriguing hypothesis worthy of future research.

# Bibliography

- Adra CN, Kobayashi H, Rowley JD, Lim B.**  
*Assignment of the human GDID4 gene, a GDP/GTP-exchange regulator, to chromosome 12p12.3.*  
Genomics. 1994 Nov 1;24(1):188-90.
- Alder MN, Dames S, Gaudet J, Mango SE.**  
*Gene silencing in Caenorhabditis elegans by transitive RNA interference.*  
RNA. 2003 Jan;9(1):25-32.
- Alekseev OM, Bencic DC, Richardson RT, Widgren EE, O'Rand MG.**  
*Overexpression of the Linker histone-binding protein tNASP affects progression through the cell cycle.*  
J Biol Chem. 2003 Mar 7;278(10):8846-52. Epub 2002 Dec 30.
- Allshire R.**  
*Molecular biology. RNAi and heterochromatin--a hushed-up affair.*  
Science. 2002 Sep 13;297(5588):1818-9. Epub 2002 Aug 22.
- Amarzguioui M, Holen T, Babaie E, Prydz H.**  
*Tolerance for mutations and chemical modifications in a siRNA.*  
Nucleic Acids Res. 2003 Jan 15;31(2):589-95.
- Ambros V.**  
*The functions of animal microRNAs.*  
Nature. 2004 Sep 16;431(7006):350-5. Review.
- Ansell J.D., Micklem H.S.**  
*"Handbook of Experimental Immunology", Volume 2: Cellular Immunology*  
Editors: Weir DM, chapter 56  
Blackwell Scientific Publications, 1986
- Auerbach W, Dunmore J, Fairchild-Huntress V, Fang Q, Auerbach A, Huszar D, Joyner A.**  
*Establishment and chimera analysis of 129/SvEv- and C57BL/6-derived mouse embryonic stem cell lines.*  
Biotechniques. 2000 Nov;29(5):1024-8, 1030, 1032.
- Baker RK, Haendel MA, Swanson BJ, Shambaugh JC, Micales BK, Lyons GE.**  
*In vitro preselection of gene-trapped embryonic stem cell clones for characterizing novel developmentally regulated genes in the mouse.*  
Dev Biol. 1997 May 15;185(2):201-14.
- Bass BL.**  
*RNA interference. The short answer.*  
Nature. 2001 May 24;411(6836):428-9.
- Baulcombe D.**  
*RNA silencing in plants.*  
Nature. 2004 Sep 16;431(7006):356-63.
- Becker-Hapak M, McAllister SS, Dowdy SF.**  
*TAT-mediated protein transduction into mammalian cells.*  
Methods. 2001 Jul;24(3):247-56. Review.

**Begley CG, Aplan PD, Denning SM, Haynes BF, Waldmann TA, Kirsch IR.**

*The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif.*

Proc Natl Acad Sci U S A. 1989 Dec;86(24):10128-32.

**Bernstein E, Caudy AA, Hammond SM, Hannon GJ.**

*Role for a bidentate ribonuclease in the initiation step of RNA interference.*

Nature. 2001 Jan 18;409(6818):363-6.

**Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ.**

*Dicer is essential for mouse development.*

Nat Genet. 2003 Nov;35(3):215-7. Epub 2003 Oct 05.

**Billy E, Brondani V, Zhang H, Muller U, Filipowicz W.**

*Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines.*

Proc Natl Acad Sci U S A. 2001 Dec 4;98(25):14428-33. Epub 2001 Nov 27.

**Blackburn EH.**

*The end of the (DNA) line.*

Nat Struct Biol. 2000 Oct;7(10):847-50. Review.

**Blackburn EH.**

*Structure and function of telomeres.*

Nature. 1991 Apr 18;350(6319):569-73. Review.

**Bonaldo P, Chowdhury K, Stoykova A, Torres M, Gruss P.**

*Efficient gene trap screening for novel developmental genes using IRES beta geo vector and in vitro preselection.*

Exp Cell Res. 1998 Oct 10;244(1):125-36.

**Bradley A, Evans M, Kaufman MH, Robertson E.**

*Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines.*

Nature. 1984 May 17-23;309(5965):255-6.

**Brennan J, Skarnes WC.**

*Gene trapping in mouse embryonic stem cells.*

Methods Mol Biol. 1999;97:123-38. Review.

**Brummelkamp TR, Bernards R, Agami R.**

*A system for stable expression of short interfering RNAs in mammalian cells.*

Science. 2002 Apr 19;296(5567):550-3. Epub 2002 Mar 21.

**Burt RK, Verda L, Kim DA, Oyama Y, Luo K, Link C.**

*Embryonic stem cells as an alternate marrow donor source: engraftment without graft-versus-host disease.*

J Exp Med. 2004 Apr 5;199(7):895-904.

**Campbell KH, McWhir J, Ritchie WA, Wilmut I.**

*Sheep cloned by nuclear transfer from a cultured cell line.*

Nature. 1996 Mar 7;380(6569):64-6.

**Cannon JP, Colicos SM, Belmont JW.**

*Gene trap screening using negative selection: identification of two tandem, differentially expressed loci with potential hematopoietic function.*  
Dev Genet. 1999;25(1):49-63.

**Caplen NJ, Parrish S, Imani F, Fire A, Morgan RA.**

*Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems.*  
Proc Natl Acad Sci U S A. 2001 Aug 14;98(17):9742-7. Epub 2001 Jul 31.

**Caplen NJ, Fleenor J, Fire A, Morgan RA.**

*dsRNA-mediated gene silencing in cultured Drosophila cells: a tissue culture model for the analysis of RNA interference.*  
Gene. 2000 Jul 11;252(1-2):95-105.

**Castanotto D, Li H, Rossi JJ.**

*Functional siRNA expression from transfected PCR products.*  
RNA. 2002 Nov;8(11):1454-60.

**Caudy AA, Hannon GJ.**

*Induction and biochemical purification of RNA-induced silencing complex from Drosophila S2 cells.*  
Methods Mol Biol. 2004;265:59-72.

**Cecconi F, Gruss P.**

*From ES cells to mice: the gene trap approach.*  
Methods Mol Biol. 2002;185:335-46. Review.

**Cecconi F, Meyer BI.**

*Gene trap: a way to identify novel genes and unravel their biological function.*  
FEBS Lett. 2000 Aug 25;480(1):63-71. Review.

**Cech TR.**

*Beginning to understand the end of the chromosome.*  
Cell. 2004 Jan 23;116(2):273-9. Review.

**Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A.**

*Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.*  
Cell. 2003 May 30;113(5):643-55.

**Chen WV, Soriano P.**

*Gene trap mutagenesis in embryonic stem cells.*  
Methods Enzymol. 2003;365:367-86.

**Choi K, Kennedy M, Kazarov A, Papadimitriou JC, Keller G.**

*A common precursor for hematopoietic and endothelial cells.*  
Development. 1998 Feb;125(4):725-32.

**Cogoni C, Macino G.**

*Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase.*  
Nature. 1999 May 13;399(6732):166-9.

**Couldrey C, Carlton MB, Nolan PM, Colledge WH, Evans MJ.**

*A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice.*  
Hum Mol Genet. 1999 Dec;8(13):2489-95.

**Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC.**

*An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus.*  
Cell. 2000 May 26;101(5):543-53.

**Dani C, Chambers I, Johnstone S, Robertson M, Ebrahimi B, Saito M, Taga T, Li M, Burdon T, Nichols J, Smith A.**

*Paracrine induction of stem cell renewal by LIF-deficient cells: a new ES cell regulatory pathway.*  
Dev Biol. 1998 Nov 1;203(1):149-62.

**Davidson BL.**

*Hepatic diseases--hitting the target with inhibitory RNAs.*  
N Engl J Med. 2003 Dec 11;349(24):2357-9.

**Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R.**

*The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium.*  
J Embryol Exp Morphol. 1985 Jun;87:27-45.

**Donze O, Picard D.**

*RNA interference in mammalian cells using siRNAs synthesized with T7 RNA polymerase.*  
Nucleic Acids Res. 2002 May 15;30(10):e46.

**Elbashir SM, Harborth J, Weber K, Tuschl T.**

*Analysis of gene function in somatic mammalian cells using small interfering RNAs.*  
Methods. 2002 Feb;26(2):199-213.

**Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T.**

*Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate.*  
EMBO J. 2001c Dec 3;20(23):6877-88.

**Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T.**

*Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells.*  
Nature. 2001b May 24;411(6836):494-8.

**Elbashir SM, Lendeckel W, Tuschl T.**

*RNA interference is mediated by 21- and 22-nucleotide RNAs.*  
Genes Dev. 2001a Jan 15;15(2):188-200.

**Esteller M.**

*DNA methylation and cancer therapy: new developments and expectations.*  
Curr Opin Oncol. 2005 Jan;17(1):55-60. Review.

**Evans MJ, Carlton MB, Russ AP.**

*Gene trapping and functional genomics.*  
Trends Genet. 1997 Sep;13(9):370-4. Review.

**Evans MJ, Kaufman MH.**

*Establishment in culture of pluripotential cells from mouse embryos.*  
Nature. 1981 Jul 9;292(5819):154-6.

**Fagard M, Boutet S, Morel JB, Bellini C, Vaucheret H.**

*AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals.*

Proc Natl Acad Sci U S A. 2000 Oct 10;97(21):11650-4.

**Faisst AM, Gruss P.**

*Bodenin: a novel murine gene expressed in restricted areas of the brain.*

Dev Dyn. 1998 Jun;212(2):293-303.

**Faria TN, LaRosa GJ, Wilen E, Liao J, Gudas LJ.**

*Characterization of genes which exhibit reduced expression during the retinoic acid-induced differentiation of F9 teratocarcinoma cells: involvement of cyclin D3 in RA-mediated growth arrest.*

Mol Cell Endocrinol. 1998 Aug 25;143(1-2):155-66.

**Filipowicz W.**

*RNAi: the nuts and bolts of the RISC machine.*

Cell. 2005 Jul 15;122(1):17-20. Review.

**Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC.**

*Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*.*

Nature. 1998 Feb 19;391(6669):806-11.

**Fisher JP, Hope SA, Hooper ML.**

*Factors influencing the differentiation of embryonal carcinoma and embryo-derived stem cells.*

Exp Cell Res. 1989 Jun;182(2):403-14.

**Floss T, Wurst W.**

*Functional genomics by gene-trapping in embryonic stem cells.*

Methods Mol Biol. 2002;185:347-79. Review.

**Forrester LM, Nagy A, Sam M, Watt A, Stevenson L, Bernstein A, Joyner AL, Wurst W.**

*An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid in vitro.*

Proc Natl Acad Sci U S A. 1996 Feb 20;93(4):1677-82.

**Franklin RE, Gosling RG.**

*Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate.*

Nature. 1953 Jul 25;172(4369):156-7.

**Franz CM, Ridley AJ.**

*p120 catenin associates with microtubules: inverse relationship between microtubule binding and Rho GTPase regulation.*

J Biol Chem. 2004 Feb 20;279(8):6588-94. Epub 2003 Dec 02.

**Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, Ahringer J.**

*Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference.*

Nature. 2000 Nov 16;408(6810):325-30.

**Friedrich G, Soriano P.**

*Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice.*

Genes Dev. 1991 Sep;5(9):1513-23.

**Friedrich GA, Hildebrand JD, Soriano P.**

*The secretory protein Sec8 is required for paraxial mesoderm formation in the mouse.*

Dev Biol. 1997 Dec 15;192(2):364-74.

**Frohman MA.**

*On beyond classic RACE (rapid amplification of cDNA ends).*

PCR Methods Appl. 1994 Aug;4(1):S40-58. Review.

**Gayther SA, Barski P, Batley SJ, Li L, de Foy KA, Cohen SN, Ponder BA, Caldas C.**

*Aberrant splicing of the TSG101 and FHIT genes occurs frequently in multiple malignancies and in normal tissues and mimics alterations previously described in tumours.*

Oncogene. 1997 Oct 23;15(17):2119-26.

**Gilbert W.**

*The RNA World*

Nature 319, 618 (1986).

**Gilchrist DS, Ure J, Hook L, Medvinsky A.**

*Labeling of hematopoietic stem and progenitor cells in novel activatable EGFP reporter mice.*

Genesis. 2003 Jul;36(3):168-76.

**Gough NM, Williams RL, Hilton DJ, Pease S, Willson TA, Stahl J, Gearing DP, Nicola NA, Metcalf D.**

*LIF: a molecule with divergent actions on myeloid leukaemic cells and embryonic stem cells.*

Reprod Fertil Dev. 1989;1(4):281-8. Review.

**Gowher H, Jeltsch A.**

*Mechanism of inhibition of DNA methyltransferases by cytidine analogs in cancer therapy.*

Cancer Biol Ther. 2004 Nov;3(11):1062-8. Epub 2004 Nov 12. Review.

**Greider CW, Blackburn EH.**

*Identification of a specific telomere terminal transferase activity in Tetrahymena extracts.*

Cell. 1985 Dec;43(2 Pt 1):405-13.

**Guillemot JC, Kruskal B, Adra CN, Zhu S, Ko J, Burch P, Nocka K, Seetoo K, Simons E, Lim B.**

*Targeted disruption of guanosine diphosphate-dissociation inhibitor for Rho-related proteins, GDID4: normal hematopoietic differentiation but subtle defect in superoxide production by macrophages derived from in vitro embryonal stem cell differentiation.*

Blood. 1996 Oct 1;88(7):2722-31.

**Guo S, Kempthues KJ.**

*par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed.*

Cell. 1995 May 19;81(4):611-20.

**Guo G, Wang W, Bradley A.**

*Mismatch repair genes identified using genetic screens in Blm-deficient embryonic stem cells.*  
Nature. 2004 Jun 24;429(6994):891-5.

**Gupta S, Schoer RA, Egan JE, Hannon GJ, Mittal V.**

*Inducible, reversible, and stable RNA interference in mammalian cells.*  
Proc Natl Acad Sci U S A. 2004 Feb 17;101(7):1927-32. Epub 2004 Feb 04.

**Gurdon JB, Elsdale TR, Fischberg M.**

*Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei.*  
Nature. 1958 Jul 5;182(4627):64-5.

**Hahn WC, Stewart SA, Brooks MW, York SG, Eaton E, Kurachi A, Beijersbergen RL, Knoll JH, Meyerson M, Weinberg RA.**

*Inhibition of telomerase limits the growth of human cancer cells.*  
Nat Med. 1999 Oct;5(10):1164-70.

**Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA.**

*Creation of human tumour cells with defined genetic elements.*  
Nature. 1999 Jul 29;400(6743):464-8.

**Hamada M, Ohtsuka T, Kawaida R, Koizumi M, Morita K, Furukawa H, Imanishi T, Miyagishi M, Taira K.**

*Effects on RNA interference in gene expression (RNAi) in cultured mammalian cells of mismatches and the introduction of chemical modifications at the 3'-ends of siRNAs.*  
Antisense Nucleic Acid Drug Dev. 2002 Oct;12(5):301-9.

**Hamilton AJ, Baulcombe DC.**

*A species of small antisense RNA in posttranscriptional gene silencing in plants.*  
Science. 1999 Oct 29;286(5441):950-2.

**Hamilton A, Voinnet O, Chappell L, Baulcombe D.**

*Two classes of short interfering RNA in RNA silencing.*  
EMBO J. 2002 Sep 2;21(17):4671-9.

**Hammond SM, Bernstein E, Beach D, Hannon GJ.**

*An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells.*  
Nature. 2000 Mar 16;404(6775):293-6.

**Hansen J, Floss T, Van Sloun P, Fuchtbauer EM, Vauti F, Arnold HH, Schnutgen F, Wurst W, von Melchner H, Ruiz P.**

*A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome.*  
Proc Natl Acad Sci U S A. 2003 Aug 19;100(17):9918-22. Epub 2003 Aug 06.

**Harborth J, Elbashir SM, Bechert K, Tuschl T, Weber K.**

*Identification of essential genes in cultured mammalian cells using small interfering RNAs.*  
J Cell Sci. 2001 Dec;114(Pt 24):4557-65.

**Hay DC, Sutherland L, Clark J, Burdon T.**

*Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells.*  
Stem Cells. 2004;22(2):225-35.

**Heidenreich O, Krauter J, Riehle H, Hadwiger P, John M, Heil G, Vornlocher HP, Nordheim A.**

*AML1/MTG8 oncogene suppression by small interfering RNAs supports myeloid differentiation of t(8;21)-positive leukemic cells.*

Blood. 2003 Apr 15;101(8):3157-63. Epub 2002 Dec 12.

**Hidaka M, Caruana G, Stanford WL, Sam M, Correll PH, Bernstein A.**

*Gene trapping of two novel genes, Hzf and Hhl, expressed in hematopoietic cells.*

Mech Dev. 2000 Jan;90(1):3-15.

**Hildebrand JD, Soriano P.**

*Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development.*

Mol Cell Biol. 2002 Aug;22(15):5296-307.

**Hill DP, Wurst W.**

*Screening for novel pattern formation genes using gene trap approaches.*

Methods Enzymol. 1993;225:664-81.

**Hofstadter D.R.**

*Gödel, Escher, Bach: an Eternal Golden Braid*

*A metaphorical fugue on minds and machines in the spirit of Lewis Carroll*

Published by Penguin Books 1980; reprinted 2000.

**Holen T, Amarzguoui M, Wiiger MT, Babaie E, Prydz H.**

*Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor*

Nucleic Acids Res. 2002 Apr 15;30(8):1757-66.

**Holen T, Moe SE, Sorbo JG, Meza TJ, Ottersen OP, Klungland A.**

*Tolerated wobble mutations in siRNAs decrease specificity, but can enhance activity in vivo.*

Nucleic Acids Res. 2005 Aug 19;33(15):4704-10.

**Holzschu D, Lapierre L, Neubaum D, Mark WH.**

*A molecular strategy designed for the rapid screening of gene traps based on sequence identity and gene expression pattern in adult mice.*

Transgenic Res. 1997 Jan;6(1):97-106.

**Huber TL, Kouskoff V, Fehling HJ, Palis J, Keller G.**

*Haemangioblast commitment is initiated in the primitive streak of the mouse embryo.*

Nature. 2004 Dec 2;432(7017):625-30.

**Hyndman L, Lemoine JL, Huang L, Porteous DJ, Boyd AC, Nan X.**

*HIV-1 Tat protein transduction domain peptide facilitates gene transfer in combination with cationic liposomes.*

J Control Release. 2004 Oct 19;99(3):435-44.

**International Human Genome Sequencing Consortium.**

*Finishing the euchromatic sequence of the human genome.*

Nature. 2004 Oct 21;431(7011):931-45.

**Ishida Y, Leder P.**

*RET: a poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells.*

Nucleic Acids Res. 1999 Dec 15;27(24):e35.

**Jackson M, Baird JW, Cambray N, Ansell JD, Forrester LM, Graham GJ.**

*Cloning and characterization of Ehox, a novel homeobox gene essential for embryonic stem cell differentiation.*

J Biol Chem. 2002 Oct 11;277(41):38683-92. Epub 2002 Jun 26.

**Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B.**

*Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells.*

Mol Cell Biol. 2004 Oct;24(20):8862-71.

**Jones L, Ratcliff F, Baulcombe DC.**

*RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance.*

Curr Biol. 2001 May 15;11(10):747-57.

**Joyce G.F., Orgel L.E.**

*"The RNA World", 49-77.*

Editors Gestland, R.F., Cech T.R. and Atkins J.F.

Cold Spring Harbor Laboratory Press, New York, 1999

**Joyner AL, Auerbach A, Skarnes WC.**

*The gene trap approach in embryonic stem cells: the potential for genetic screens in mice.*

Ciba Found Symp. 1992;165:277-88; discussion 288-97. Review.

**Joyner AL.**

*Gene targeting and gene trap screens using embryonic stem cells: new approaches to mammalian development.*

Bioessays. 1991 Dec;13(12):649-56. Review.

**Kabrun N, Buhring HJ, Choi K, Ullrich A, Risau W, Keller G.**

*Flk-1 expression defines a population of early embryonic hematopoietic precursors.*

Development. 1997 May;124(10):2039-48.

**Katsanis N, Fisher EM.**

*A novel C-terminal binding protein (CTBP2) is closely related to CTBP1, an adenovirus E1A-binding protein, and maps to human chromosome 21q21.3.*

Genomics. 1998 Jan 15;47(2):294-9.

**Katz SG, Cantor AB, Orkin SH.**

*Interaction between FOG-1 and the corepressor C-terminal binding protein is dispensable for normal erythropoiesis in vivo.*

Mol Cell Biol. 2002 May;22(9):3121-8.

**Kawasaki H, Taira K.**

*Induction of DNA methylation and gene silencing by short interfering RNAs in human cells.*

Nature. 2004 Sep 9;431(7005):211-7. Epub 2004 Aug 15.

**Kawasaki H, Taira K.**

*Retraction: Induction of DNA methylation and gene silencing by short interfering RNAs in human cells.*

Nature. 2006 Jun 29;441:1176.

**Kelland LR.**

*Telomerase inhibitors: targeting the vulnerable end of cancer?*

Anticancer Drugs. 2000 Aug;11(7):503-13. Review.

**Keller G, Kennedy M, Papayannopoulou T, Wiles MV.**

*Hematopoietic commitment during embryonic stem cell differentiation in culture.*

Mol Cell Biol. 1993 Jan;13(1):473-86.

**Keller GM.**

*In vitro differentiation of embryonic stem cells.*

Curr Opin Cell Biol. 1995 Dec;7(6):862-9. Review.

**Keller G.**

*Embryonic stem cell differentiation: emergence of a new era in biology and medicine.*

Genes Dev. 2005 May 15;19(10):1129-55. Review.

**Kennedy M, Firpo M, Choi K, Wall C, Robertson S, Kabrun N, Keller G.**

*A common precursor for primitive erythropoiesis and definitive haematopoiesis.*

Nature. 1997 Apr 3;386(6624):488-93.

**Kennerdell JR, Carthew RW.**

*Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway.*

Cell. 1998 Dec 23;95(7):1017-26.

**Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH.**

*Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*.*

Genes Dev. 2001 Oct 15;15(20):2654-9.

**Kiger A, Baum B, Jones S, Jones M, Coulson A, Echeverri C, Perrimon N.**

*A functional genomic analysis of cell morphology using RNA interference.*

J Biol. 2003;2(4):27. Epub 2003 Oct 1.

**Klug MG, Soonpaa MH, Koh GY, Field LJ.**

*Genetically selected cardiomyocytes from differentiating embryonic stem cells form stable intracardiac grafts.*

J Clin Invest. 1996 Jul 1;98(1):216-24.

**Kyba M, Perlingeiro RC, Daley GQ.**

*HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors.*

Cell. 2002 Apr 5;109(1):29-37.

**Laird PW.**

*Cancer epigenetics.*

Hum Mol Genet. 2005 Apr 15;14 Spec No 1:R65-76. Review.

**Lander ES et al.**

**International Human Genome Sequencing Consortium.**

*Initial sequencing and analysis of the human genome.*

Nature. 2001 Feb 15;409(6822):860-921.

**Lassus P, Rodriguez J, Lazebnik Y.**

*Confirming specificity of RNAi in mammalian cells.*

Sci STKE. 2002 Aug 27;2002(147):PL13. Print 2002 Aug 27.

**Lee JT.**

*Disruption of imprinted X inactivation by parent-of-origin effects at Tsix.*

Cell. 2000 Sep 29;103(1):17-27.

**Lee JT, Lu N.**

*Targeted mutagenesis of Tsix leads to nonrandom X inactivation.*

Cell. 1999 Oct 1;99(1):47-57.

**Lee JT, Davidow LS, Warshawsky D.**

*Tsix, a gene antisense to Xist at the X-inactivation centre.*

Nat Genet. 1999 Apr;21(4):400-4.

**Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW.**

*Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways.*

Cell. 2004 Apr 2;117(1):69-81.

**Lefebvre L, Dionne N, Karaskova J, Squire JA, Nagy A.**

*Selection for transgene homozygosity in embryonic stem cells results in extensive loss of heterozygosity.*

Nat Genet. 2001 Mar;27(3):257-8.

**Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E.**

*De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells.*

Development. 1996 Oct;122(10):3195-205.

**Lelias JM, Adra CN, Wulf GM, Guillemot JC, Khagad M, Caput D, Lim B.**

*cDNA cloning of a human mRNA preferentially expressed in hematopoietic cells and with homology to a GDP-dissociation inhibitor for the rho GTP-binding proteins.*

Proc Natl Acad Sci U S A. 1993 Feb 15;90(4):1479-83.

**Lewis DL, Hagstrom JE, Loomis AG, Wolff JA, Herweijer H.**

*Efficient delivery of siRNA for inhibition of gene expression in postnatal mice.*

Nat Genet. 2002 Sep;32(1):107-8.

**Li E, Bestor TH, Jaenisch R.**

*Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.*

Cell. 1992 Jun 12;69(6):915-26.

**Li L, Liao J, Ruland J, Mak TW, Cohen SN.**

*A TSG101/MDM2 regulatory loop modulates MDM2 degradation and MDM2/p53 feedback control.*

Proc Natl Acad Sci U S A. 2001 Feb 13;98(4):1619-24.

**Li L, Francke U, Cohen SN.**

*Retraction. The TSG101 tumor susceptibility gene is located in chromosome 11 band p15 and is mutated in human breast cancer.*  
Cell. 1998 May 15;93(4):following 660.

**Li L, Li X, Francke U, Cohen SN.**

*The TSG101 tumor susceptibility gene is located in chromosome 11 band p15 and is mutated in human breast cancer.*  
Cell. 1997 Jan 10;88(1):143-54.

**Li L, Cohen SN.**

*Tsg101: a novel tumor susceptibility gene isolated by controlled homozygous functional knockout of allelic loci in mammalian cells.*  
Cell. 1996 May 3;85(3):319-29.

**Lipardi C, Wei Q, Paterson BM.**

*RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs.*  
Cell. 2001 Nov 2;107(3):297-307.

**Liu K, Li L, Cohen SN.**

*Antisense RNA-mediated deficiency of the calpain protease, nCL-4, in NIH3T3 cells is associated with neoplastic transformation and tumorigenesis.*  
J Biol Chem. 2000a Oct 6;275(40):31093-8.

**Liu K, Li L, Nisson PE, Gruber C, Jessee J, Cohen SN.**

*Reversible tumorigenesis induced by deficiency of vasodilator-stimulated phosphoprotein.*  
Mol Cell Biol. 1999 May;19(5):3696-703.

**Liu K, Li L, Nisson PE, Gruber C, Jessee J, Cohen SN.**

*Neoplastic transformation and tumorigenesis associated with sam68 protein deficiency in cultured murine fibroblasts.*  
J Biol Chem. 2000b Dec 22;275(51):40195-201.

**Macleod D, Lovell-Badge R, Jones S, Jackson I.**

*A promoter trap in embryonic stem (ES) cells selects for integration of DNA into CpG islands.*  
Nucleic Acids Res. 1991 Jan 11;19(1):17-23.

**Mainguy G, Montesinos ML, Lesaffre B, Zevnik B, Karasawa M, Kothary R, Wurst W, Prochiantz A, Volovitch M.**

*An induction gene trap for identifying a homeoprotein-regulated locus.*  
Nat Biotechnol. 2000 Jul;18(7):746-9.

**Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T.**

*Single-stranded antisense siRNAs guide target RNA cleavage in RNAi.*  
Cell. 2002 Sep 6;110(5):563-74.

**Martinez J, Tuschl T.**

*RISC is a 5' phosphomonoester-producing RNA endonuclease.*  
Genes Dev. 2004 May 1;18(9):975-80. Epub 2004 Apr 22.

- Matsuda E, Shigeoka T, Iida R, Yamanaka S, Kawaichi M, Ishida Y.**  
*Expression profiling with arrays of randomly disrupted genes in mouse embryonic stem cells leads to in vivo functional analysis.*  
Proc Natl Acad Sci U S A. 2004 Mar 23;101(12):4170-4. Epub 2004 Mar 09.
- McCaffrey AP, Meuse L, Pham TT, Conklin DS, Hannon GJ, Kay MA.**  
*RNA interference in adult mice.*  
Nature. 2002 Jul 4;418(6893):38-9.
- McClive P, Pall G, Newton K, Lee M, Mullins J, Forrester L.**  
*Gene trap integrations expressed in the developing heart: insertion site affects splicing of the PT1-ATG vector.*  
Dev Dyn. 1998 Jun;212(2):267-76.
- McManus MT, Sharp PA.**  
*Gene silencing in mammals by small interfering RNAs.*  
Nat Rev Genet. 2002 Oct;3(10):737-47. Review.
- McPherson JD et al.**  
**International Human Genome Mapping Consortium.**  
*A physical map of the human genome.*  
Nature. 2001 Feb 15;409(6822):934-41.
- Medvinsky A, Dzierzak E.**  
*Development of the hematopoietic stem cell: can we describe it?*  
Blood. 1999 Nov 15;94(10):3613-4.
- Meister G, Tuschl T.**  
*Mechanisms of gene silencing by double-stranded RNA.*  
Nature. 2004 Sep 16;431(7006):343-9.
- Mello CC, Conte D Jr.**  
*Revealing the world of RNA interference.*  
Nature. 2004 Sep 16;431(7006):338-42.
- Mette MF, Aufsatz W, van der Winden J, Matzke MA, Matzke AJ.**  
*Transcriptional silencing and promoter methylation triggered by double-stranded RNA.*  
EMBO J. 2000 Oct 2;19(19):5194-201.
- Mitchell K, Pinson K, Kelly O, Brennan J, Zupicich J, Scherz P, Leighton P, Goodrich LV, Lu X, Avery B, Tate P, Dill K, Pangilinan E, Wakenight P, Tessier-Lavigne M, Skarnes WC.**  
*Functional analysis of secreted and transmembrane proteins critical to mouse development.*  
Nat Genet. 2001 Jul;28(3):241-9.
- Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S.**  
*The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.*  
Cell. 2003 May 30;113(5):631-42.

**Miyagishi M, Taira K.**

*U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells.*

Nat Biotechnol. 2002 May;20(5):497-500.

**Motamedi MR, Verdel A, Colmenares SU, Gerber SA, Gygi SP, Moazed D.**

*Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs.*

Cell. 2004 Dec 17;119(6):789-802.

**Morel JB, Mourrain P, Beclin C, Vaucheret H.**

*DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in Arabidopsis.*

Curr Biol. 2000 Dec 14-28;10(24):1591-4.

**Morris KV, Chan SW, Jacobsen SE, Looney DJ.**

*Small interfering RNA-induced transcriptional gene silencing in human cells.*

Science. 2004 Aug 27;305(5688):1289-92. Epub 2004 Aug 05.

**Muth K, Bruyns R, Thorey IS, von Melchner H.**

*Disruption of genes regulated during hematopoietic differentiation of mouse embryonic stem cells.*

Dev Dyn. 1998 Jun;212(2):277-83.

**Nakano T, Kodama H, Honjo T.**

*Generation of lymphohematopoietic cells from embryonic stem cells in culture.*

Science. 1994 Aug 19;265(5175):1098-101.

**Napoli C, Lemieux C, Jorgensen R.**

*Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans.*

Plant Cell. 1990 Apr;2(4):279-289.

**Natarajan D, Boulter CA.**

*A lacZ-hygromycin fusion gene and its use in a gene trap vector for marking embryonic stem cells.*

Nucleic Acids Res. 1995 Oct 11;23(19):4003-4. No abstract available.

**Nichols J, Zevnik B, Anastassiadis, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith AG.**

*Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.*

Cell. 1998 Oct 30;95(3):379-91.

**Nichols J, Evans EP, Smith AG.**

*Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity.*

Development. 1990 Dec;110(4):1341-8.

**Nishikawa SI, Nishikawa S, Hirashima M, Matsuyoshi N, Kodama H.**

*Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages.*

Development. 1998 May;125(9):1747-57.

**Nishikura K.**

*A short primer on RNAi: RNA-directed RNA polymerase acts as a key catalyst.*  
Cell. 2001 Nov 16;107(4):415-8. Review.

**Niwa H, Burdon T, Chambers I, Smith A.**

*Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3.*  
Genes Dev. 1998 Jul 1;12(13):2048-60.

**Niwa H, Miyazaki J, Smith AG.**

*Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.*  
Nat Genet. 2000 Apr;24(4):372-6.

**Novina CD, Murray MF, Dykxhoorn DM, Beresford PJ, Riess J, Lee SK, Collman RG, Lieberman J, Shankar P, Sharp PA.**

*siRNA-directed inhibition of HIV-1 infection.*  
Nat Med. 2002 Jul;8(7):681-6.

**Nykanen A, Haley B, Zamore PD.**

*ATP requirements and small interfering RNA structure in the RNA interference pathway.*  
Cell. 2001 Nov 2;107(3):309-21.

**Okano M, Bell DW, Haber DA, Li E.**

*DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.*  
Cell. 1999 Oct 29;99(3):247-57.

**Okano M, Xie S, Li E.**

*Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.*  
Nat Genet. 1998 Jul;19(3):219-20.

**Ogawa Y, Lee JT.**

*Antisense regulation in X inactivation and autosomal imprinting.*  
Cytogenet Genome Res. 2002;99(1-4):59-65. Review.

**Orkin SH.**

*GATA-binding transcription factors in hematopoietic cells.*  
Blood. 1992 Aug 1;80(3):575-81. Review.

**Paddison PJ, Silva JM, Conklin DS, Schlabach M, Li M, Aruleba S, Balija V, O'Shaughnessy A, Gnoj L, Scobie K, Chang K, Westbrook T, Cleary M, Sachidanandam R, McCombie WR, Elledge SJ, Hannon GJ.**

*A resource for large-scale RNA-interference-based screens in mammals.*  
Nature. 2004 Mar 25;428(6981):427-31.

**Paddison PJ, Hannon GJ.**

*RNA interference: the new somatic cell genetics?*  
Cancer Cell. 2002c Jul;2(1):17-23. Review.

**Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS.**

*Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells.*  
Genes Dev. 2002b Apr 15;16(8):948-58.

**Paddison PJ, Caudy AA, Hannon GJ.**

*Stable suppression of gene expression by RNAi in mammalian cells.*

Proc Natl Acad Sci U S A. 2002a Feb 5;99(3):1443-8. Epub 2002 Jan 29.

**Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA, Elgin SC.**

*Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery.*

Science. 2004 Jan 30;303(5658):669-72.

**Pal-Bhadra M, Bhadra U, Birchler JA.**

*RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila.*

Mol Cell. 2002 Feb;9(2):315-27.

**Parrish S, Fire A.**

*Distinct roles for RDE-1 and RDE-4 during RNA interference in Caenorhabditis elegans.*

RNA. 2001 Oct;7(10):1397-402.

**Parrish S, Fleenor J, Xu S, Mello C, Fire A.**

*Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference.*

Mol Cell. 2000 Nov;6(5):1077-87.

**Pease S, Braghetta P, Gearing D, Grail D, Williams RL.**

*Isolation of embryonic stem (ES) cells in media supplemented with recombinant leukemia inhibitory factor (LIF).*

Dev Biol. 1990 Oct;141(2):344-52.

**Plasterk RH.**

*RNA silencing: the genome's immune system.*

Science. 2002 May 17;296(5571):1263-5.

**Poole JC, Andrews LG, Tollefsbol TO.**

*Activity, function, and gene regulation of the catalytic subunit of telomerase (hTERT).*

Gene. 2001 May 16;269(1-2):1-12. Review.

**Ramsahoye BH.**

*Nearest-neighbor analysis.*

Methods Mol Biol. 2002;200:9-15. No abstract available.

**Reinhart BJ, Bartel DP.**

*Small RNAs correspond to centromere heterochromatic repeats.*

Science. 2002 Sep 13;297(5588):1831. Epub 2002 Aug 22.

**Reynolds AB, Herbert L, Cleveland JL, Berg ST, Gaut JR.**

*p120, a novel substrate of protein tyrosine kinase receptors and of p60v-src, is related to cadherin-binding factors beta-catenin, plakoglobin and armadillo.*

Oncogene. 1992 Dec;7(12):2439-45.

**Reynolds AB, Jenkins NA, Gilbert DJ, Copeland NG, Shapiro DN, Wu J, Daniel JM.**

*The gene encoding p120cas, a novel catenin, localizes on human chromosome 11q11 (CTNND) and mouse chromosome 2 (Catns).*

Genomics. 1996 Jan 1;31(1):127-9.

**Richardson RT, Bencic DC, O'Rand MG.**

*Comparison of mouse and human NASP genes and expression in human transformed and tumor cell lines.*

Gene. 2001 Aug 22;274(1-2):67-75.

**Richardson RT, Batova IN, Widgren E, Zheng LX, Whitfield M, Marzluff WF, O'Rand MG.**  
*Characterization of the histone H1-binding protein, NASP, as a cell cycle-regulated somatic protein.*

J Biol Chem. 2000 Sep 29;275(39):30378-86.

**Rivas FV, Tolia NH, Song JJ, Aragon JP, Liu J, Hannon GJ, Joshua-Tor L.**

*Purified Argonaute2 and an siRNA form recombinant human RISC.*

Nat Struct Mol Biol. 2005 Apr;12(4):340-9.

**Robertson E.J.**

*"Teratocarcinoma and Embryo-derived Stem Cells: A Practical Approach"*

IRL, Oxford, 1997.

**Robertson SM, Kennedy M, Shannon JM, Keller G.**

*A transitional stage in the commitment of mesoderm to hematopoiesis requiring the transcription factor SCL/tal-1.*

Development. 2000 Jun;127(11):2447-59.

**Romano N, Macino G.**

*Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences.*

Mol Microbiol. 1992 Nov;6(22):3343-53.

**Rubinson DA, Dillon CP, Kwiatkowski AV, Sievers C, Yang L, Kopinja J, Rooney DL, Ihrig MM, McManus MT, Gertler FB, Scott ML, Van Parijs L.**

*A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference.*

Nat Genet. 2003 Mar;33(3):401-6.

**Ruland J, Sirard C, Elia A, MacPherson D, Wakeham A, Li L, de la Pompa J, Cohen SN, Mak TW.**

*p53 accumulation, defective cell proliferation, and early embryonic lethality in mice lacking tsg101.*

Proc Natl Acad Sci U S A. 2001 Feb 13;98(4):1859-64.

**Sado T, Li E, Sasaki H.**

*Effect of TSIX disruption on XIST expression in male ES cells.*

Cytogenet Genome Res. 2002;99(1-4):115-8.

**Salminen M, Meyer BI, Gruss P.**

*Efficient poly A trap approach allows the capture of genes specifically active in differentiated embryonic stem cells and in mouse embryos.*

Dev Dyn. 1998 Jun;212(2):326-33.

**Sam M, Wurst W, Kluppel M, Jin O, Heng H, Bernstein A.**

*Aquarius, a novel gene isolated by gene trapping with an RNA-dependent RNA polymerase motif.*

Dev Dyn. 1998 Jun;212(2):304-17.

**Sam M, Wurst W, Forrester L, Vauti F, Heng H, Bernstein A.**

*A novel family of repeat sequences in the mouse genome responsive to retinoic acid.*  
Mamm Genome. 1996 Oct;7(10):741-8.

**Sanger F, Nicklen S, Coulson AR.**

*DNA sequencing with chain-terminating inhibitors.*  
Proc Natl Acad Sci U S A. 1977 Dec;74(12):5463-7.

**Schmitt RM, Bruyins E, Snodgrass HR.**

*Hematopoietic development of embryonic stem cells in vitro: cytokine and receptor gene expression.*  
Genes Dev. 1991 May;5(5):728-40.

**Schrödinger E.**

*What is Life? The Physical Aspect of the Living Cell.*  
*With Mind and Matter and Autobiographical Sketches*  
Published Cambridge University Press, 1944.

**Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD.**

*Asymmetry in the assembly of the RNAi enzyme complex.*  
Cell. 2003 Oct 17;115(2):199-208.

**Schwarz DS, Hutvagner G, Haley B, Zamore PD.**

*Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways.*  
Mol Cell. 2002 Sep;10(3):537-48.

**Schwarze SR, Dowdy SF.**

*In vivo protein transduction: intracellular delivery of biologically active proteins, compounds and DNA.*  
Trends Pharmacol Sci. 2000 Feb;21(2):45-8.

**Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF.**

*In vivo protein transduction: delivery of a biologically active protein into the mouse.*  
Science. 1999 Sep 3;285(5433):1569-72.

**Sharp PA, Zamore PD.**

*Molecular biology. RNA interference.*  
Science. 2000 Mar 31;287(5462):2431-3.

**Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A.**

*On the role of RNA amplification in dsRNA-triggered gene silencing.*  
Cell. 2001 Nov 16;107(4):465-76.

**Skarnes WC.**

*Gene trapping methods for the identification and functional analysis of cell surface proteins in mice.*  
Methods Enzymol. 2000;328:592-615.

**Skarnes WC, Moss JE, Hurtley SM, Beddington RS.**

*Capturing genes encoding membrane and secreted proteins important for mouse development.*  
Proc Natl Acad Sci U S A. 1995 Jul 3;92(14):6592-6.

**Skarnes WC.**

*The identification of new genes: gene trapping in transgenic mice.*  
Curr Opin Biotechnol. 1993 Dec;4(6):684-9. Review.

**Skarnes WC, Auerbach BA, Joyner AL.**

*A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice.*  
Genes Dev. 1992 Jun;6(6):903-18.

**Smith, A.G.**

*Culture and differentiation of embryonic stem cells.*  
J. Tissue Culture Methods 13, 89-94 (1991)

**Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D.**

*Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.*  
Nature. 1988 Dec 15;336(6200):688-90.

**Snyder EL, Meade BR, Saenz CC, Dowdy SF.**

*Treatment of Terminal Peritoneal Carcinomatosis by a Transducible p53-Activating Peptide.*  
PLoS Biol. 2004 Feb;2(2):E36. Epub 2004 Feb 17.

**Snyder EL, Dowdy SF.**

*Protein/peptide transduction domains: potential to deliver large DNA molecules into cells.*  
Curr Opin Mol Ther. 2001 Apr;3(2):147-52. Review.

**Sontheimer EJ, Carthew RW.**

*Silence from within: endogenous siRNAs and miRNAs.*  
Cell. 2005 Jul 15;122(1):9-12. Review.

**Sorensen DR, Leirdal M, Sioud M.**

*Gene silencing by systemic delivery of synthetic siRNAs in adult mice.*  
J Mol Biol. 2003 Apr 4;327(4):761-6.

**Soutschek J, Akinc A, Bramlage B, Charisse K, Constien R, Donoghue M, Elbashir S, Geick A, Hadwiger P, Harborth J, John M, Kesavan V, Lavine G, Pandey RK, Racie T, Rajeev KG, Rohl I, Toudjarska I, Wang G, Wuschko S, Bumcrot D, Kotliansky V, Limmer S, Manoharan M, Vornlocher HP.**

*Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs.*  
Nature. 2004 Nov 11;432(7014):173-8.

**Stanford WL, Cohn JB, Cordes SP.**

*Gene-trap mutagenesis: past, present and beyond.*  
Nat Rev Genet. 2001 Oct;2(10):756-68. Review.

**Stanford WL, Caruana G, Vallis KA, Inamdar M, Hidaka M, Bautch VL, Bernstein A.**

*Expression trapping: identification of novel genes expressed in hematopoietic and endothelial lineages by gene trapping in ES cells.*  
Blood. 1998 Dec 15;92(12):4622-31.

**Stoykova A, Chowdhury K, Bonaldo P, Torres M, Gruss P.**

*Gene trap expression and mutational analysis for genes involved in the development of the mammalian nervous system.*  
Dev Dyn. 1998 Jun;212(2):198-213.

**Stryke D, Kawamoto M, Huang CC, Johns SJ, King LA, Harper CA, Meng EC, Lee RE, Yee A, L'Italien L, Chuang PT, Young SG, Skarnes WC, Babbitt PC, Ferrin TE.**

*BayGenomics: a resource of insertional mutations in mouse embryonic stem cells.*  
Nucleic Acids Res. 2003 Jan 1;31(1):278-81.

**Stuhlmann H.**

*Gene trap vector screen for developmental genes in differentiating ES cells.*  
Methods Enzymol. 2003;365:386-406. No abstract available.

**Sutherland H, Mumford G, Newton K, Ford L, Farrall, Dellaire G, Caceres J, Bickmore WA**

*Large-scale identification of mammalian proteins localized to nuclear sub-compartments.*  
Hum Mol Genet. 2001 Sep 1;10(18):1995-2011.

**Svoboda P, Stein P, Schultz RM.**

*RNAi in mouse oocytes and preimplantation embryos: effectiveness of hairpin dsRNA.*  
Biochem Biophys Res Commun. 2001 Oct 12;287(5):1099-104.

**Tabara H, Grishok A, Mello CC.**

*RNAi in C. elegans: soaking in the genome sequence.*  
Science. 1998 Oct 16;282(5388):430-1.

**Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC.**

*The rde-1 gene, RNA interference, and transposon silencing in C. elegans.*  
Cell. 1999 Oct 15;99(2):123-32.

**Takeuchi T.**

*A gene trap approach to identify genes that control development.*  
Dev Growth Differ. 1997 Apr;39(2):127-34. Review.

**Tanaka TS, Kunath T, Kimber WL, Jaradat SA, Stagg CA, Usuda M, Yokota T, Niwa H, Rossant J, Ko MS.**

*Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity.*  
Genome Res. 2002 Dec;12(12):1921-8.

**Tang FC, Meng GL, Yang HB, Li CJ, Shi Y, Ding MX, Shang KG, Zhang B, Xue YF.**

*Stable suppression of gene expression in murine embryonic stem cells by RNAi directed from DNA vector-based short hairpin RNA.*  
Stem Cells. 2004;22(1):93-9.

**Thio SS, Bonventre JV, Hsu SI.**

*The CtBP2 co-repressor is regulated by NADH-dependent dimerization and possesses a novel N-terminal repression domain.*  
Nucleic Acids Res. 2004 Mar 22;32(5):1836-47. Print 2004.

**Thomson JA, Itskovitz-Eldor J, Shapiro S, Waknitz MA, Swiergiel J, Marshall V, Jones JM.**

*Embryonic stem cell lines derived from human blastocysts.*  
Science. 1998 Nov 6;282(5391):1145-7.

**Timmons L.**

*The long and short of siRNAs.*  
Mol Cell. 2002 Sep;10(3):435-7. Review.

**Timmons L, Court DL, Fire A.**

*Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans.*  
Gene. 2001 Jan 24;263(1-2):103-12.

**Timmons L, Fire A.**

*Specific interference by ingested dsRNA.*  
Nature. 1998 Oct 29;395(6705):854.

**Ting AH, Schuebel KE, Herman JG, Baylin SB.**

*Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation.*  
Nat Genet. 2005 Aug;37(8):906-10.

**Tomari Y, Du T, Haley B, Schwarz DS, Bennett R, Cook HA, Koppetsch BS, Theurkauf WE, Zamore PD.**

*RISC assembly defects in the Drosophila RNAi mutant armitage.*  
Cell. 2004 Mar 19;116(6):831-41.

**Tomari Y, Matranga C, Haley B, Martinez N, Zamore PD.**

*A protein sensor for siRNA asymmetry.*  
Science. 2004 Nov 19;306(5700):1377-80.

**Tomari Y, Zamore PD.**

*Perspective: machines for RNAi.*  
Genes Dev. 2005 Mar 1;19(5):517-29. Review.

**Torchilin VP, Rammohan R, Weissig V, Levchenko TS.**

*TAT peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors.*  
Proc Natl Acad Sci U S A. 2001; 98, 8786-91.

**Townley DJ, Avery BJ, Rosen B, Skarnes WC.**

*Rapid sequence analysis of gene trap integrations to generate a resource of insertional mutations in mice.*  
Genome Res. 1997 Mar;7(3):293-8.

**Turner J, Crossley M.**

*Cloning and characterization of mCtBP2, a co-repressor that associates with basic Kruppel-like factor and other mammalian transcriptional regulators.*  
EMBO J. 1998 Sep 1;17(17):5129-40.

**Tuschl T.**

*RNA interference and small interfering RNAs.*  
Chembiochem. 2001 Apr 2;2(4):239-45. Review.

**Tuschl T.**

*The siRNA User Guide*  
<http://www.rockefeller.edu/labheads/tuschl/sirna.html>

**Tuschl T, Zamore PD, Lehmann R, Bartel DP, Sharp PA.**

*Targeted mRNA degradation by double-stranded RNA in vitro.*  
Genes Dev. 1999 Dec 15;13(24):3191-7.

**Ui-Tei K, Naito Y, Takahashi F, Haraguchi T, Ohki-Hamazaki H, Juni A, Ueda R, Saigo K.**  
*Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference.*

Nucleic Acids Res. 2004 Feb 09;32(3):936-48. Print 2004.

**Vallis KA, Chen Z, Stanford WL, Yu M, Hill RP, Bernstein A.**

*Identification of radiation-responsive genes in vitro using a gene trap strategy predicts for modulation of expression by radiation in vivo.*

Radiat Res. 2002 Jan;157(1):8-18.

**van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR.**

*Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression.*

Plant Cell. 1990 Apr;2(4):291-9.

**van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR.**

*Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression.*

Plant Cell. 1990 Apr;2(4):291-9.

**Van de Putte T, Zwijsen A, Lonnoy O, Rybin V, Cozijnsen M, Francis A, Baekelandt V, Kozak CA, Zerial M, Huylebroeck D.**

*Mice with a homozygous gene trap vector insertion in mgcRacGAP die during pre-implantation development.*

Mech Dev. 2001 Apr;102(1-2):33-44.

**van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR.**

*Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression.*

Plant Cell. 1990 Apr;2(4):291-9.

**Velkey JM, O'Shea KS.**

*Oct4 RNA interference induces trophectoderm differentiation in mouse embryonic stem cells.*

Genesis. 2003 Sep;37(1):18-24.

**Venter JC et al.**

*The sequence of the human genome.*

Science. 2001 Feb 16;291(5507):1304-51.

**Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA.**

*Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi.*

Science. 2002 Sep 13;297(5588):1833-7. Epub 2002 Aug 22.

**Voss AK, Thomas T, Gruss P.**

*Efficiency assessment of the gene trap approach.*

Dev Dyn. 1998 Jun;212(2):171-80.

**Wadia JS, Dowdy SF.**

*Protein transduction technology.*

Curr Opin Biotechnol. 2002 Feb;13(1):52-6. Review.

**Wall NR, Shi Y.**

*Small RNA: can RNA interference be exploited for therapy?*  
Lancet. 2003 Oct 25;362(9393):1401-3. Review.

**Wang SW, Speck NA.**

*Purification of core-binding factor, a protein that binds the conserved core site in murine leukemia virus enhancers.*  
Mol Cell Biol. 1992 Jan;12(1):89-102.

**Wang ZQ, Kiefer F, Urbanek P, Wagner EF.**

*Generation of completely embryonic stem cell-derived mutant mice using tetraploid blastocyst injection.*  
Mech Dev. 1997 Mar;62(2):137-45.

**Wassenegger M.**

*The role of the RNAi machinery in heterochromatin formation.*  
Cell. 2005 Jul 15;122(1):13-6. Review.

**Wassenegger M, Heimes S, Riedel L, Sanger HL.**

*RNA-directed de novo methylation of genomic sequences in plants.*  
Cell. 1994 Feb 11;76(3):567-76.

**Watson JD, Crick FH.**

*Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.*  
Nature. 1953a Apr 25;171(4356):737-8.

**Watson JD, Crick FH.**

*Genetical implications of the structure of deoxyribonucleic acid.*  
Nature. 1953b May 30;171(4361):964-7.

**Watt AJ, Jones EA, Ure JM, Peddie D, Wilson DI, Forrester LM.**

*A gene trap integration provides an early in situ marker for hepatic specification of the foregut endoderm.*  
Mech Dev. 2001 Feb;100(2):205-15.

**Weinberg RA.**

*Telomeres. Bumps on the road to immortality.*  
Nature. 1998 Nov 5;396(6706):23-4.

**Wianny F, Zernicka-Goetz M.**

*Specific interference with gene function by double-stranded RNA in early mouse development.*  
Nat Cell Biol. 2000 Feb;2(2):70-5.

**Wilda M, Fuchs U, Wossmann W, Borkhardt A.**

*Killing of leukemic cells with a BCR/ABL fusion gene by RNA interference (RNAi).*  
Oncogene. 2002 Aug 22;21(37):5716-24.

**Wiles MV, Vauti F, Otte J, Fuchtbauer EM, Ruiz P, Fuchtbauer A, Arnold HH, Lehrach H, Metz T, von Melchner H, Wurst W.**

*Establishment of a gene-trap sequence tag library to generate mutant mice from embryonic stem cells.*  
Nat Genet. 2000 Jan;24(1):13-4.

**Wiles MV, Keller G.**

*Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture.*  
Development. 1991 Feb;111(2):259-67.

**Wilkins MH, Stokes AR, Wilson HR.**

*Molecular structure of deoxyribose nucleic acids.*  
Nature. 1953 Apr 25;171(4356):738-40.

**Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM.**

*Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.*  
Nature. 1988 Dec 15;336(6200):684-7.

**Worby CA, Simonson-Leff N, Dixon JE.**

*RNA interference of gene expression (RNAi) in cultured Drosophila cells.*  
Sci STKE. 2001 Aug 14;2001(95):PL1. Print 2001 Aug 14.

**Wurst W, Rossant J, Prideaux V, Kownacka M, Joyner A, Hill DP, Guillemot F, Gasca S, Cado D, Auerbach A.**

*A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice.*  
Genetics. 1995 Feb;139(2):889-99.

**Xia H, Mao Q, Paulson HL, Davidson BL.**

*siRNA-mediated gene silencing in vitro and in vivo.*  
Nat Biotechnol. 2002 Oct;20(10):1006-10.

**Xie S, Wang Z, Okano M, Nogami M, Li Y, He WW, Okumura K, Li E.**

*Cloning, expression and chromosome locations of the human DNMT3 gene family.*  
Gene. 1999 Aug 5;236(1):87-95.

**Xie W, Li L, Cohen SN.**

*Cell cycle-dependent subcellular localization of the TSG101 protein and mitotic and nuclear abnormalities associated with TSG101 deficiency.*  
Proc Natl Acad Sci U S A. 1998 Feb 17;95(4):1595-600.

**Yanagisawa M, Kaverina IN, Wang A, Fujita Y, Reynolds AB, Anastasiadis PZ.**

*A novel interaction between kinesin and p120 modulates p120 localization and function.*  
J Biol Chem. 2004 Mar 5;279(10):9512-21. Epub 2003 Dec 04.

**Yang S, Tutton S, Pierce E, Yoon K.**

*Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells.*  
Mol Cell Biol. 2001 Nov;21(22):7807-16.

**Ying QL, Nichols J, Chambers I, Smith A.**

*BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3.*  
Cell. 2003 Oct 31;115(3):281-92.

**Zambrowicz BP, Abuin A, Ramirez-Solis R, Richter LJ, Piggott J, BeltrandelRio H, Buxton E, Edwards J, Finch R, Friddle C, Gupta A, Hansen G, Hu Y, Huang W, Jaing C, Key B Jr, Kipp P, Kohlhauff B, Ma Z, Markesich D, Payne R, Potter D, Qian N, Shaw J, Schrick J, Shi Z, Sparks MJ, Van Sligtenhorst I, Vogel P, Walke W, Xu N, Zhu Q, Person C, Sands AT.** *Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention.* Proc Natl Acad Sci U S A. 2003 Nov 25;100(24):14109-14. Epub 2003 Nov 10.

**Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C, Sands AT.** *Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells.* Nature. 1998 Apr 9;392(6676):608-11.

**Zambrowicz BP, Friedrich GA.** *Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse.* Int J Dev Biol. 1998;42(7):1025-36. Review.

**Zamore PD.** *Ancient pathways programmed by small RNAs.* Science. 2002 May 17;296(5571):1265-9.

**Zamore PD.** *RNA interference: listening to the sound of silence.* Nat Struct Biol. 2001 Sep;8(9):746-50. Review.

**Zamore PD, Tuschl T, Sharp PA, Bartel DP.** *RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals.* Cell. 2000 Mar 31;101(1):25-33.

**Zeng Y, Cullen BR.** *RNA interference in human cells is restricted to the cytoplasm.* RNA. 2002 Jul;8(7):855-60.

**Zou GM, Wu W, Chen J, Rowley JD.** *Duplexes of 21-nucleotide RNAs mediate RNA interference in differentiated mouse ES cells.* Biol Cell. 2003;95(6):365-71.

**Zilberman D, Cao X, Jacobsen SE.** *ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation.* Science. 2003 Jan 31;299(5607):716-9.