INVESTIGATION OF mRNA ABLATION STRATEGIES IN ES CELLS

SONIA ANITA LEE

Thesis presented for degree of Doctor of Philosophy
University of Edinburgh
2002

I declare that the work presented in this thesis is my own, except where otherwise stated

Sonia Lee

Acknowledgements

I would like to thank my supervisors Austin and Ian for their guidance and help throughout my PhD. I would also like to thank the members of our lab past and present whom I shall remember fondly for the laughs and friendships shared. Thanks also to Steve Sheardown for helpful discussions on my PhD and on life in Industry.

I would also like to thank my parents for their constant support and encouragment and for their faith in me. I also need to say thanks to Anna, Angela and Michael, the other Lee's who in their own ways have shown their support. Thanks Michael for the tapes, and Anna and Angela for your trips up to Edinburgh and taking me out.

I also want to thank my many friends, both the old and the new ones that I have made here. Thank you Jan for your phone calls and visits and Anita for my most beautiful God daughter Kimiko. Thanks to Adele for your friendship, T and T and Wednesday night dinners. Thanks also to Anna (Banana), Alison B and Alison W for your support, girly nights out and good chat. I would also like to thank the other CGR students, in particular Atif, Dean and Marios. Atif for his friendship and unique sense of humour, Dean for his chat and keeping me amused and Marios for all his computer technical advice which has been invaluable.

I also want to thank Patrick for being there, encouraging me and more importantly for just listening to me.

Finally I would also like to acknowledge Cicely Farnan, who was not only a neighbour but a dear friend who will be sadly missed.

CONTENTS

ABSTRACT

Embryonic stem (ES) cells are continuous cell lines derived from the inner cell mass of early mouse embryos and are capable of multilineage differentiation. ES cells can therefore be used to study events occurring during both the self renewal and differentiation processes in mice. However, many of the genes involved in ES cell differentiation or indeed self renewal have not been identified.

Loss of function analysis such as mRNA ablation strategies have been used as a way of determining the function of a gene. mRNA ablation strategies such as antisense RNA, ribozymes and more recently dsRNA have been used successfully to determine functions of various genes both *in vitro* and *in vivo*. Antisense RNA has been one of the predominant methods of ablation in previous studies but has shown a very variable rate of ablation. Ribozymes are a potentially more powerful ablation tool than antisense, as they are able to cleave their targets and be recycled for further use. Double stranded RNA mediates RNA interference (RNAi) which is the process of sequence specific degradation and silencing of genes in animals and plants. For all these ablation strategies the efficiency of gene silencing in ES cells is unknown.

The aim of the work presented in this thesis is to investigate mRNA ablation systems for assignment of gene function in ES cell self renewal and differentiation. Such a technology could be used to determine the potential role of candidate genes or to identify new genes in a functional library screen. My strategy was to express antisense RNA, ribozymes or double stranded hairpin RNA using a high efficiency episomal transduction and expression system. Antisense RNAs were targeted against the 5' and 3' UTRs of *LIFR*, *STAT3* and *Oct4* genes required for the self renewal of ES cells. However, stable expression of these antisense RNAs did not result in the expected differentiation phenotype when cells were maintained in the presence of LIF. Ribozyme ablation was also ineffective when targeting the coding region of the p300 gene target, using a ribozyme reported previously to be effective in embryonal carcinoma cells.

In contrast the operation of RNAi in ES cells was established. This was demonstrated by the specific suppression of luciferase expression from reporter plasmids co-transfected with 21 nt specific short interfering RNAs, siRNAs. The feasibility of an RNAi strategy by in situ production of dsRNA from stable electroporation of an episomally maintained vector harbouring long inverted repeats to the coding regions of eGFP, LIFR and SOCS3 targets was then investigated. dsRNA to eGFP appeared to produce sequence specific RNAi against a chromosomal eGFP gene randomly integrated in ES cells, where up to a 33% reduction of eGFP expression was observed. However, I found that the cells expressing reduced eGFP were dying as determined by staining with Annexin V and 7-AAD, so that reduced flourescence may be a secondary consequence. Cell death could be attributable to activation of an interferon response as it is known that long dsRNAs in mammalian cells are capable of inducing apoptosis. Evidence of ablation of the other targets was not observed, as there was neither phenotypic ablation of the LIFR target nor a reduction of SOCS3 mRNA. Compared to cells expressing eGFP dsRNA there was less cell death observed with SOCS3 dsRNA expression, which may be attributable to the higher transcript levels noted for eGFP dsRNA. The lack of dsRNA mediated ablation may be due to the choice of target site.

In conclusion the data presented here determines that of the various ablation methods studied, only 21 nt siRNAs are reliable and are a potential ablation tool to define or discover gene function in ES cells in the future.

Chapter 1: Introduction

1.1. Embryonic Stem Cells

1.1.1. Background

Embryonic stem cells (ES cells) are cell lines derived from the inner cell mass (ICM) of the early mouse embryo (Evans and Kaufman, 1981; Martin, 1981; Brook and Gardner, 1997). ES cells retain the pluripotent state of the cells that constitute the ICM, which is the key underlying feature that is essential to the generation of transgenic mouse lines and multilineage differentiation. *In vitro* ES cells are capable of multilineage differentiation, and are consequently used to study lineage commitment and differentiation of various cell types from ES cells, including neuronal precursors, neurons and cells from the haematopoietic lineage (Bain et al.1995; Okabe et al.1996; Li et al.1998; Wiles and Keller 1991; Kennedy et al.1997). Pluripotency of ES cells has been demonstrated *in vivo* by reintroducing ES cells into a mouse embryo by injection of ES cells into a blastocyst. ES cells incorporate into the ICM, and contribute to all tissues including the germ cells resulting in a chimeric mouse (Bradley et al.1984; Nichols et al.1990). Germline transmission has allowed the development of sophisticated ES cell gene targeting strategies to generate mice with defined genomic changes and has become standard practice in researching gene function. A current limitation however, is the time consuming and labour intensive nature of gene knockouts in mice by homologous recombination. As a direct result it would be of considerable use to determine if alternative messenger RNA (mRNA) ablation strategies can be employed to determine gene function more quickly yet accurately in ES cells. Current mRNA ablation strategies include antisense RNA,

ribozymes and double stranded RNA. The mechanisms involved as well as current evidence of their effectiveness as ablation strategies will now be discussed.

1.2. Ablation Strategies

1.2.1. Background

Gene "knockout" technology is the traditional method for determining the role of a gene in mice. Though an effective approach the major drawback is that this method is slow and labour intensive. Appropriate targeting constructs have to be generated followed by targeting of ES cells and isolation of correctly targeted clones before these can be injected into mouse blastocysts and bred to make mice homozygous for the desired mutation. With the recent rapid expansion of available gene sequences through such projects as the "genome project", quicker yet reliable strategies for discovering gene function are required. mRNA ablation strategies have been proposed. These ablation strategies have potential not only for discovering the function of a cloned gene of interest but for the identification of new genes by an *in vitro* screening approach. I have focused on three approaches: antisense RNA; ribozymes; and double stranded RNA (dsRNA).

1.2.2. Antisense RNA

1.2.2.1. Background

Antisense RNA is RNA that is complementary to a target mRNA. It exists naturally in both prokaryotes and eukaryotes and is usually transcribed from the same locus as sense RNAs but in the opposite direction (reviewed by Vanhee-

Brossollet and Vaquero 1998). Antisense RNA acts to control gene expression by hybridising specifically to the targeted mRNA through Watson-Crick base pair interactions.

The first demonstration of antisense RNA or DNA selectively preventing gene expression experimentally in eukaryotic cells was provided by Izant and Weintraub (1984). Expression of a thymidine kinase gene injected into mouse L cells (that are thymidine kinase negative) was significantly reduced when a plasmid directing synthesis of antisense thymidine kinase RNA was co-injected.

Although antisense RNA has been used for some time as a tool for silencing gene expression, the mechanisms involved are still not fully understood. It is thought that the duplex formation between antisense and sense RNAs interferes with one or more steps in the gene expression pathway including transcription, RNA processing, mRNA transport, translation and mRNA turnover (see figure 1.1.). However, whilst antisense RNA may interfere with transcription there is little evidence of this in the literature. Examples of antisense RNA interfering with some of the other steps of gene expression are described in the text below.

Antisense RNA has been shown to inhibit splicing of human β-globin pre-mRNA *in vitro* by Munroe (1988). Splicing reactions were performed in HeLa cell nuclear extracts. Antisense RNAs complementary to sequences between 80-140 nucleotides downstream of the globin 3' splice site inhibited splicing as efficiently as antisense RNAs complementary to the 5' splice site. Antisense RNA was shown to anneal directly to the pre-mRNA substrate by the formation of T1 RNase resistant duplexes. A reduction in splicing inhibition observed following

late addition of antisense RNA reflected a reduction in the formation of these duplexes. However, the mechanism by which an RNA duplex inhibits splicing remains to be established.

The next stage in the gene expression pathway is the export of mature mRNA transcripts out of the nucleus. Antisense induced inhibition of the transport of transcripts out of the nucleus has been demonstrated experimentally in transfected cells (Kim and Wold 1985). Antisense RNA against thymidine kinase (TK) reduced TK activity by 80%-90% in a generated TK positive L cell line (Kim and Wold 1985). Antisense TK was expressed as a fusion with dihydrofolate reductase (DHFR), which allowed higher levels of antisense expression by selecting cells resistant to methotrexate that overproduce DHFR. The observed reduction in TK activity was due to greater than 95% of the TK mRNA being localised to the nucleus. Antisense RNA in this instance is causing inhibition by hybridising to its target in the nucleus preventing subsequent transport to the cytoplasm. However, the mechanism of how antisense hybridisation prevents TK mRNA export was not determined by Kim and Wold (1985).

Alternatively, inhibition by antisense RNA can occur at the translational level, preventing initiation, elongation or termination. For example, translation initiation of exogenously added β-globin mRNA was specifically blocked in frog oocytes injected with antisense RNA to the 5' untranslated region (UTR) of β-globin cDNA (Melton et al.1985). This was determined by labelling newly synthesised proteins by incubating oocytes in ^{35}S methionine. Antisense RNAs that were specific to the 3' half of β-globin were unable to block protein synthesis in this instance. Repression of ribosomal protein L1 synthesis however, was only

observed when antisense RNAs were complementary to the 3'-terminal region, indicating that antisense mediated inhibition was interfering with translation termination (Wormington 1986). Again this was demonstrated by incubating antisense injected *Xenopus* oocytes in ^{35}S methionine.

In addition RNA degradation is also thought to assist antisense RNA mediated inhibition in some cases. For example RNA degradation by RNase H has been reported for antisense oligonucleotides. RNase H is a ubiquitous enzyme that specifically degrades the RNA strand of RNA-DNA hybrids. For example inhibition of ICAM-1 expression, an intercellular adhesion molecule, by an antisense oligonucleotide was found to be due in part to RNase H mediated cleavage of the target mRNA (Chiang et al.1991). Chiang et al.(1991) demonstrated that a 2'-O-Methyl analogue, which does not support RNase H, failed to inhibit ICAM-1 expression in both human umbilical vein endothelial cells and a human lung carcinoma cell line.

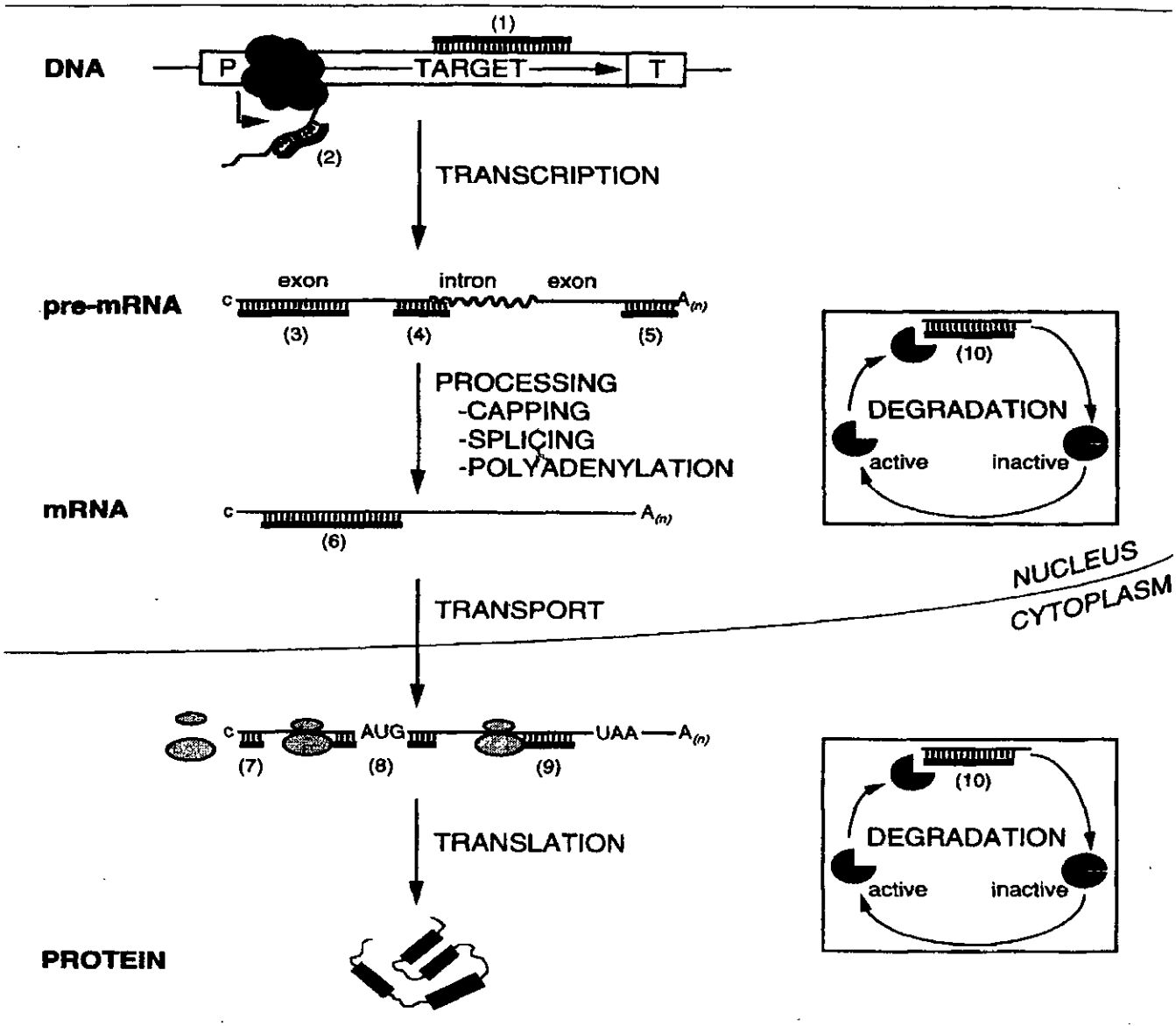


Figure 1.1. Potential modes of action of antisense in eukaryotic cells. Antisense RNA can interfere with one or more steps in eukaryotic gene expression, including: a) transcription by binding to one strand of duplex DNA at an open loop (1) or by hybridisation to nascent RNA (2); b) pre-mRNA processing by blocking 5' cap addition (3), splicing (4), or polyadenylation (5); c) transport of target mRNA from the nucleus to the cytoplasm (6); d) translation by interfering with ribosome binding (7), prevention of translation initiation (8), or truncation of translational elongation (9). All RNA-RNA hybrids formed during these steps may be degraded by double strand RNA specific RNase (10) Taken from Arndt and Rank 1997.

Antisense oligonucleotides and their analogues have been developed in parallel with antisense RNA studies and show similar potential for inhibiting gene expression in eukaryotic cells (reviewed in Colman 1990). Antisense oligonucleotides are short stretches of nucleotides that are complementary to a region of targeted mRNA, which can specifically suppress expression of that particular transcript. In some cases chemical modifications have been made to the phosphate-sugar backbone to render the oligonucleotides resistant to nuclease action i.e. substitution of one of the non-bridging oxygen atoms bound to phosphorus for a sulphur atom, generating phosphorothioate oligonucleotides. Phosphorothioate oligonucleotides act predominantly by an RNase H mechanism where the RNA strand of a DNA/RNA duplex is cleaved. The other antisense oligonucleotides act independently of RNase H and are generally thought to hybridise directly to the target mRNA and prevent translation (reviewed in Crooke 1992).

Morpholino oligos are DNA oligonucleotides that are made from morpholine subunits (C_4H_9NO) (Summerton and Weller 1997) which act by blocking translation. They are typically 25 mers in length and are complementary to regions within the 5' UTR or the first 25 bases 3' to the AUG translational start site. Morpholino oligos were initially developed to overcome the specificity problems of antisense oligonucleotides such as phosphorothioate linked DNA analogues which act primarily by an RNase H cleavage mechanism. Specificity problems arise because RNase H can cleave DNA/RNA duplexes as short as 5 or 6 base pairs in length (Monia et al.1993) which may result in non specific cleavage. Morpholino mediated ablation on the other hand is RNase H

independent. In addition targeting the 5' UTR, which is less conserved than coding regions and the fact that morpholinos are sensitive to several mismatched nucleotides reduces the problem of non specific ablation. Morpholinos have also been shown directly to be more efficient than antisense oligonucleotides (Summerton et al.1997). Summerton et al.(1997) demonstrated in HeLa cells that two morpholinos targeting globin were able to effectively inhibit translation at 300 nM, while phosphorothioate oligonucleotides were unable to produce significant inhibition of the target when using up to 3000 nM. However, high concentrations of morpholinos can result in non specific effects such as cell death (Braat et al.2001). To compensate for this control mRNA rescue experiments or the inclusion of two morpholinos complementary to different parts of the 5' UTR are required to investigate gene function. In addition the other major disadvantage of antisense oligonucleotides and morpholinos are that unlike antisense RNA they cannot be introduced into cells by stable expression from a vector making these ablation strategies less versatile than antisense RNA.

1.2.2.2. Methods for identifying optimal target sites

The efficiency of antisense ablation is highly dependent upon the region of the gene targeted. This is because mRNA not only possesses secondary structure but it is bound by cellular proteins. As a result the target site may be occluded, consequently reducing accessibility to the antisense reagent. There are methods available for selecting optimal hybridisation sites, although they are not trivial. The current methods available are RNase H mapping, combinatorial arrays and prediction of mRNA secondary structure by computational methods. RNase H

mapping of accessible sites within the target mRNA involves the hybridisation of a random oligonucleotide library and subsequent cleavage by RNase H, which cleaves the RNA strand of an RNA-DNA duplex. Fragments are then sequence analysed to identify accessible sites. This method has been shown to be successful, as oligonucleotides targeted to every accessible site found by RNase H mapping were shown to be active (Ho et al. 1998). Combinatorial arrays is a technique where all possible oligonucleotides of a chosen length targeting a specific mRNA are synthesised on a glass plate (Milner et al.1997). Accessible target sites of the targeted mRNA can be determined by radioactively labelling the target message and then hybridising it to the combinatorial array. Milner et al.(1997) found that the heteroduplexes identified by combinatorial arrays correlated well with sites identified by RNase H mapping of the rabbit β-globin target. Secondary structure analysis by computational methods usually predict secondary structures of RNA based upon the lowest free energy. However, computational methods are unable to account for interactions of the mRNA with possible RNA binding proteins present in the cell nor is it able to predict the secondary structures that form as the message is being transcribed. These comprehensive approaches for the identification of optimal hybridisation sites are however, often not adopted when a number of messages are being targeted primarily because of time constraints. Instead a number of antisense RNAs targeting different regions of the mRNA can be tested to determine the most effective site(s).

Optimal hybridisation sites for antisense RNA vary for different mRNAs. For example the optimal antisense RNA sequence for the β-globin message is specific to the 5' UTR (Melton 1985). Melton (1985) found that only antisense RNA, which included the 5' UTR was able to prevent translation of β-globin mRNA when injected into the cytoplasm of frog oocytes. Antisense transcripts complementary to the entire or 3' terminal coding region of L1 mRNA of *Xenopus* on the other hand were most effective at repressing translation (Wormington 1986). A review of ablation efficiencies obtained when targeting different regions within messages in transgenic animals is described by Sokol and Murray (1996). In general Sokol and Murray (1996) found that targeting the UTRs or coding sequence resulted in similar levels of either mRNA or protein reduction.

1.2.2.3. Target regions of antisense- specifically untranslated regions

Targeting accessible sites within the mRNA is not the only consideration when generating an effective antisense RNA. Non-specific effects due to cross-hybridisation of antisense RNA to a related gene transcript may also be a problem. However, specific antisense ablation is often difficult to demonstrate conclusively except when antisense RNAs target either the 5' or 3' untranslated regions (UTRs). In this instance transfection of the coding sequence for the targeted message should result in phenotypic rescue. How targeting either the 5' or 3' UTRs of the mRNA with antisense RNA is thought to interfere with gene expression is now described.

Antisense to 5' UTRs are thought to interfere with translation initiation. Translation initiation involves the association of the mRNA with the 43-S preinitiation complex. Several eukaryotic initiation factors (eIFs) are involved in the interaction of the mRNA with the 43-S preinitiation complex. The initial interaction involves the binding of the eukaryotic initiation factor-4F (eIF-4F) multiprotein complex to the cap at the 5' end of mRNA. The complex includes eIF-4E, which directly binds to cap structures and eIF-4A, which possesses RNA unwinding activity that is ATP dependent and requires eIF-4B. eIF-4F binding facilitates binding of the 43-S preinitiation complex which can then migrate along the mRNA until it reaches the initiation codon, where the 60-S ribosomal subunit associates to form the translating 80-S ribosome. This is the scanning model which predicts how the 43-S preinitiation complex associates with the initiation codon (Kozak 1978). Antisense RNA is thought to prevent appropriate binding of these molecules by hybridising to the mRNA and blocking access. Studies on the effect of mRNA/cDNA duplexes within the coding sequence of α and β-globin mRNAs have shown that positions downstream of the 80-S formation site at the AUG are not sensitive to inhibition by antisense cDNA during translation in rabbit reticulocyte lysates (Liebhaber et al.1984). This implies that the elongating ribosome is able to remove such blockages, an activity which may be critical for translation elongation. Hybridisation of cDNA fragments extending into the 5' UTR on the other hand were able to fully block translation. Translation is commonly inhibited by antisense molecules directed to the entire 5' UTR and a segment extending 20-30 nucleotides downstream of the AUG. For example additional analysis of a set of α or β-globin mRNA/cDNA hybrids which differ

from one another in the extent of the mRNA exposed at the 5' end revealed that either the first 18 or 26 nucleotides 3' to the initiation codon have to be exposed to allow translation respectively (Shakin and Liebhaber 1986). Translation of these mRNA/cDNA hybrids is associated with the complete removal of cDNA from the mRNA. However, this disruption of mRNA/cDNA duplex is blocked by inhibitors of translational initiation and elongation.

Antisense to the 3' UTR region of a gene is thought to ablate gene function by interfering with both polyadenylation and removal of the last intron of the pre-mRNA. Efficient polyadenylation of a message requires polyadenylation signals. In mammals polyadenylation signals usually consist of a hexanucleotide consensus element AAUAAA and GU-rich downstream elements, located 14-70 nucleotides downstream of the AAUAAA sequence (McDevitt et al.1984). The polyadenylation signal, AAUAAA is within the 3' UTR and so antisense directed to this region may disrupt the final processing of the transcript. In addition, there is evidence that both the AAUAAA and GU-rich downstream elements are coupled with splicing and are required for efficient splicing of the last exon (Berget 1995). Niwa and Berget (1991) have shown that in an *in vitro* splicing-polyadenylation system, mutations in polyadenylation signal not only eliminate polyadenylation but decrease splicing (the removal of the last intron). Thus targeting the 3' UTR with specific antisense RNA may inhibit the final production of mature cytoplasmic RNA by preventing both polyadenylation and splicing to remove the last intron.

1.2.2.4. Antisense ablation in ES cells

The premise that antisense technology is able to ablate gene function in ES cells is based upon a few indicative examples. For example, expression of antisense specific to the entire coding region of STAT3 resulted in a modest increase in the proportion of differentiated colonies grown in LIF, although there was no direct determination of mRNA or protein (Niwa et al.1998). Another example of antisense ablation in ES cells again with STAT3 as the target gene, was shown by Ernst et al.(1999). They used however, an 18bp phosphothioate antisense oligonucleotide directed to the 5' end of STAT3. They were able to show a ~66% decrease in the proportion of undifferentiated ES cells when grown in the presence of LIF. In addition the level of STAT3 protein was specifically decreased by more than 90% compared to controls. Antisense ablation in ES cells was also demonstrated with the inhibition of vav mRNA, a proto-oncogene expressed specifically in hematopoietic cells (Wulf et al.1993). Stable integration of a transgene expressing antisense vav transcripts in ES cell lines prevented ES cell differentiation into hematopoietic cells. The level of vav mRNA had also decreased in clones stably expressing antisense directed against vav. The experiment was done by co-transfecting the PGK driven vav antisense vector directed to two thirds of the 5' end of the vav cDNA with a PGKneo vector and selecting for G418 resistant clones. These examples are indicative that antisense RNA can be used as a method of ablating gene function in ES cells.

1.2.2.5. Antisense Ablation library Screens

An outcome of successfully establishing mRNA ablation in ES cells using antisense RNA is its potential use in a library screen to identify genes required for pluripotency or those that determine differentiation. Antisense RNA has been successfully used to identify novel genes in *Dictyostelium discoideum* (Spann et al.1996) and in mammalian cells (Levy-Strumpf and Kimchi 1998). Spann et al.(1996) transformed *Dictyostelium* cells with a cDNA library made from mRNA of vegetative and developing cells. 149 developmentally defective mutants were identified from an initial screen of 1426 transformants. Antisense cDNA fragment(s) were isolated and sequenced for 35 of these mutants and 6 found to be novel genes. Re-transformation of the corresponding antisense vectors produced the same phenotype and this was further confirmed by disruption of the corresponding genomic sequences. Levy-Strumpf and Kimchi (1998) identified genes involved in growth inhibition and apoptosis of HeLa cells. Cells were transfected with an antisense cDNA expression library derived from interferon γ treated cells. Antisense cDNA fragments that protected cells from interferon γ induced cell death were analysed. Of 7 cDNAs identified, 5 were novel. Full length sense cDNAs were then isolated and functionally characterised for four of the novel genes with overexpression studies which resulted in the death of HeLa cells. These examples demonstrate the ability of antisense to be used as a quick yet reliable screen in terms of phenocopying a disrupted gene's phenotype.

1.2.3. Ribozymes

1.2.3.1. Background

Ribozymes are RNA molecules with enzymatic activity that cleave themselves or other RNA molecules in a sequence specific manner. Like antisense reagents they hybridise to their target but their additional ability to cleave makes them more potent in principle, since this allows each molecule to inactivate multiple copies of the specific target.

The study of RNA processing led to the initial discovery that RNAs possess enzymatic properties (Kruger et al.1982; Guerrier-Takada et al.1983). Since then several types of ribozymes have been discovered; group I and II introns, the RNA subunit of RNase P, hairpin, hepatitis delta virus ribozymes and hammerhead ribozymes (Altman 1989; Cech 1990; Symons 1992; Feldstein et al.1989). The hammerhead ribozyme is one of the best characterised ribozymes. It is the smallest and has minimal target sequence requirements. It was discovered in virusoids, which have a circular single stranded RNA genome that is replicated by a rolling circle mechanism as a polycistronic message subsequently requiring cleavage into individual RNA transcripts (Forster and Symons 1987).

In vivo hammerhead ribozymes function in cis (intramolecularly), but have been engineered to work in trans (intermolecularly) *in vitro* (Uhlenbeck 1987; Haseloff and Gerlach 1988). This allows them to be designed to any target sequence making them potentially highly flexible tools for the inhibition of gene expression. The ribozyme sequence requirements consist of a catalytic domain and flanking sequences complementary to the target mRNA (helix I and III) (figure 1.2.). The optimal length of each of the flanking hybridising sequences are

9-12 nucleotides (Herschlag 1991) which are long enough to allow formation of a stable enzyme substrate complex, but short enough to allow a high rate of product release. The only sequence requirement of the target mRNA is a three base sequence, GUX, where cleavage occurs after the X base which is any base but G (Haseloff and Gerlach 1988). However, the efficiency of inhibition by ribozymes within the cell is affected by a number of factors. These include secondary and tertiary structure of both the target and the ribozyme and colocalization of the ribozyme with the sequences targeted. Futhermore efficient delivery methods combined with maintenance of high levels of expression of the ribozyme have to be considered.

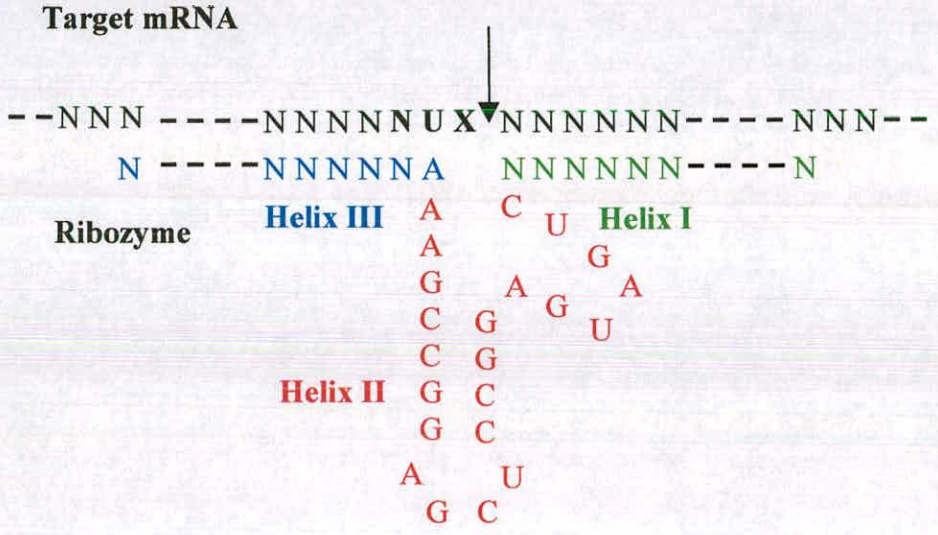


Figure 1.2. Structure of a Hammerhead ribozyme.
Helix I and III are the hybridising arms, while helix II is the catalytic domain required to catalyse cleavage of the mRNA target specified by an NUX sequence. N is any nucleotide and X is any nucleotide but G. The arrow denotes cleavage of the message.

1.2.3.2. Ribozyme expression vector

Ribozymes are to be introduced into ES cells through the use of a ribozyme expression vector. The efficiency of ribozyme ablation is dependent upon the choice of promoter to ensure both high expression levels and minimal inclusion of additional sequences to the ribozyme transcript, as these could mask the short hybridising regions. The promoter of choice is a modified tRNA gene which is transcribed by polymerase III (pol III). tRNA genes are highly expressed in most tissues and are small in size making their promoter ideal for expressing ribozyme sequences.

1.2.3.3. Ribozyme ablation in mammalian cells

A potential use of an effective ribozyme ablation strategy in ES cells would be in a library screen identifying novel genes required for either ES cell self renewal or differentiation. Although the majority of ribozyme mediated inhibition has been investigated in non mammalian cells such as *E.coli* and frog oocytes there is some evidence of successful ribozyme ablation in mammalian cells. For example specific ablation of α-lactalbumin expressed from a vector transfected into mouse C127I mammary cells was observed with five ribozymes targeted to either the coding or 3' untranslated region (L'Huillier et al.1992). Ribozyme mediated ablation was detected as both a decrease in mRNA and protein levels. Ablation was attributed directly to the catalytic activity of the ribozyme rather than to antisense effects of the flanking arms for one of the ribozymes, due to the fact that a catalytically inactive variant had no or very little effect. In another example, two ribozymes targeted to the hepatitis B virus X protein, a protein that

is required to establish infection, were able to decrease hepatitis B virus RNA in established liver cell lines (Weinberg et al.2000).

1.2.4. Double stranded RNA

1.2.4.1. Background

RNA interference (RNAi) is the process by which double stranded RNA induces the silencing of homologous genes post transcriptionally. The term RNAi was first introduced by Fire and colleagues in 1998 to describe the unexpected phenomenon that injection of both sense and antisense RNA into *C.elegans* silenced expression of unc-22 at least 10 fold more efficiently than either RNA strand alone. Previously however, RNAi had been observed when either antisense or control sense RNA to *par-1* gene in *C.elegans* produced an identical embryonic lethal phenotype (Guo and Kemphues 1995). This was due to RNA preparations being contaminated with small amounts of double stranded RNA (dsRNA). The biological significance of RNAi is to protect the genome against invasion by mobile genetic elements such as viruses and transposons, which produce dsRNA in the host cell or aberrant RNA when activated (Ketting et al.1999; Ratcliff et al.1999).

A characteristic feature of RNAi is its potency. Injection into *C.elegans* of only a few molecules of dsRNA per cell is sufficient to produce the null phenotype of unc-22 in 100% of worms (Fire et al.1998). This suggests that dsRNA either acts catalytically or undergoes an amplification step. RNAi ablation also possesses the characteristic in *C.elegans* that it is non cell autonomous, where

injection specifically into the gut of *C.elegans* caused silencing throughout the animal, which was then heritable to the F1 progeny (Fire et al.1998). RNAi suppresses gene expression by a post transcriptional mechanism, as dsRNA homologous to intron sequences is unable to produce RNAi (Montgomery et al.1998). Also dsRNA targeting one gene within an operon in *C.elegans* does not effect expression of a second within that operon, strongly indicating that RNAi occurs after transcription (Montgomery et al.1998).

Although originally discovered in *C.elegans*, RNAi has since been found in a wide variety of animals, including *Drosophila* (Kennerdell and Carthew 1998), trypanosomes (Ngo et al.1998), planaria (Sanchez-Alvarado and Newmark 1999), zebrafish (Wargelius et al.1999), mice (Wianny and Zernicka-Goetz 2000; Svoboda et al.2000) and post transcriptional gene silencing seen in plants (Waterhouse et al.1999) and *Neurospora* (Cogoni et al.1996). RNAi is most effective in *C.elegans* and *Drosophila*, and has become a valuable tool for reverse genetics in these organisms. In almost all cases, treatment of *C.elegans* with dsRNA phenocopies the genetic null mutant. RNAi has also been used successfully to study *Drosophila* embryology and gene function analysis in cultured *Drosophila* cells (Kennerdell and Carthew 1998; Clemens et al.2000). In addition dsRNA has been used specifically to interfere with gene function in mammals. Early mouse embryos display specific ablation. Ablation effects do not persist after implantation however, as the microinjections of dsRNA are diluted by an increase in cell number at this stage (Wianny and Zernicka-Goetz 2000; Svoboda et al.2000). Recently RNAi has been demonstrated in cultured

mammalian cells (Elbashir et al.2001b), including ES cells (Billy et al.2001; Paddison et al.2001; Yang et al.2001).

1.2.4.2. Proposed model of RNAi ablation

1.2.4.2.1. Biochemical investigation

A key insight into a possible mechanism for RNAi ablation was uncovered by Hamilton and Baulcombe (1999). They identified a small RNA species of ~25 nucleotides (nt) homologous to the gene being silenced that was only present in plants undergoing post transcriptional gene silencing. The importance of these small RNA species was evidenced by their subsequent discovery in cultured *Drosophila* cells induced with dsRNA (Hammond et al.2000), *Drosophila* embryo extracts carrying out RNAi *in vitro* (Zamore et al.2000) and in *Drosophila* embryos injected with dsRNA (Yang et al.2000). These small RNAs were thought to be the putative specificity determinants that act as guides to target specific mRNA, as the ~25 nt RNA species was found to be homologous to the target in *Drosophila* S2 cells (Hammond et al.2000). RNAi is due to specific mRNA degradation by nuclease digestion specified by the ~25 nt RNAs as a nuclease co-fractionating with ~25 nt RNAs was partially purified from *Drosophila* S2 cells transfected with dsRNA (Hammond et al.2000). The nuclease complex was found to have protein components and so was designated an RNA induced silencing complex (RISC). Zamore and colleagues (2000) using *Drosophila* embryo extracts found that both strands of the dsRNA were processed into 21-23 nt fragments and that this processing occurred independently of the targeted mRNA i.e. the dsRNA is cleaved first. The targeted mRNA is then

cleaved but only within the region of identity with the dsRNA and at 21-23 nt intervals.

Based on the 21-23 nt length of the processed dsRNA it was speculated that a RNase III enzyme was the nuclease responsible for RNAi (Bass 2000). RNase III is the only characterised nuclease known to cleave dsRNA at specific sites to generate dsRNA fragments of discrete sizes. A candidate protein named Dicer was isolated from *Drosophila* S2 cells in which RNAi had been initiated by transfection with dsRNA (Bernstein et al.2001). Dicer is a member of the RNase III family of nucleases and contains an amino terminal ATP dependent RNA helicase domain, two RNase III domains and one dsRNA binding domain at the carboxyl terminus. Dicer is evolutionarily conserved in flies, worms, plants, fungi and mammals (Jacobsen et al.1999; Filippov et al.2000; Matsuda et al.2000). At present only a similar role for the Dicer homologue in *C.elegans* has been determined (Ketting et al.2001).

Continued investigation of the RNAi mechanism using a *Drosophila in vitro* system (Tuschl et al.1999; Zamore et al.2000) mapped cleavage sites on the sense and antisense target RNA. The positioning of these cleavage sites revealed that both strands were predominantly cleaved 10 nt from the 5' end of the region covered by the dsRNA (Elbashir et al.2001a). Analysis of 21-23 nt fragments generated by processing of the target RNA revealed that the fragments possessed 5' phosphate and 3' hydroxyl termini with staggered 3' ends. This is another indication that dsRNAs are processed by an RNase III enzyme which generates similar reaction products (Nicholson 1999). The first demonstration of the ability of synthetic 21 and 22 nt RNAs to mediate target RNA cleavage was also shown

by Elbashir et al.(2001a). These 21 and 22 nt RNAs were subsequently named short interfering RNAs, or siRNAs. Synthetic siRNAs with 2 or 3 nt overhanging 3' ends were found to be more efficient than either blunt ended or dsRNAs with 4 nt overhangs and the cleavage site is located near the centre of the spanned region. They also provide some evidence suggesting that the orientation of the siRNA duplex with respect to proteins in the RISC complex, may determine which of the two complementary strands cleaves the mRNA.

1.2.4.2.2. Genetic investigation

In parallel to increasing biochemical data, genetic screens have identified a number of genes required for RNAi, predominantly in *C.elegans*. Genes *rde-1* and *rde-4* (where rde denotes "RNAi deficient") are essential for initiating RNAi as heterozygous worms exposed to dsRNA are able to transmit RNAi to homozygous mutant offspring (Tabara et al.1999). *rde-2* and *mut-7* (mutator-7) mutants are required directly for RNAi, as heterozygous worms injected with dsRNA are unable to transmit RNAi to homozygous offspring (Grishok et al.2000). The RDE-1 protein is homologous to the quelling defective, *qde-2* gene in *Neurospora* (Catalananotto et al.2000) and to *Argonaute*, *AGO-1* in *Arabidopsis* (Fagard et al.2000), indicating that there is conservation between species. The role of *rde-1* gene is unknown although its gene product is similar to the mammalian translation initiation factor, eIF2C (Tabara et al.1999). No homologs of *rde-4*, *rde-2* or *mut-7* have been identified in other species. However, the *mut-7* gene product is homologous to both RNase D of *E.coli* and to the

exonuclease domain of the human Werner syndrome protein, implicating its possible role as the nuclease degrading mRNA in *C.elegans* (Ketting et al.1999).

The incredible potency and systemic effects of RNAi in *C.elegans* has led to the suggestion that RNA dependent RNA polymerases (RdRPs) may play a role in amplification of the silencing effect. This was confirmed by the discovery of *qde-1* gene in *Neurospora* (Cogoni and Macino 1999), whose protein is a homologue of a previously characterised tomato RdRP (Schiebel et al.1993). Subsequently *Arabidopsis* and *C.elegans* homologs have been identified, *Sde-1* and *ego-1* respectively (Dalmay et al.2000; Smardon et al.2000).

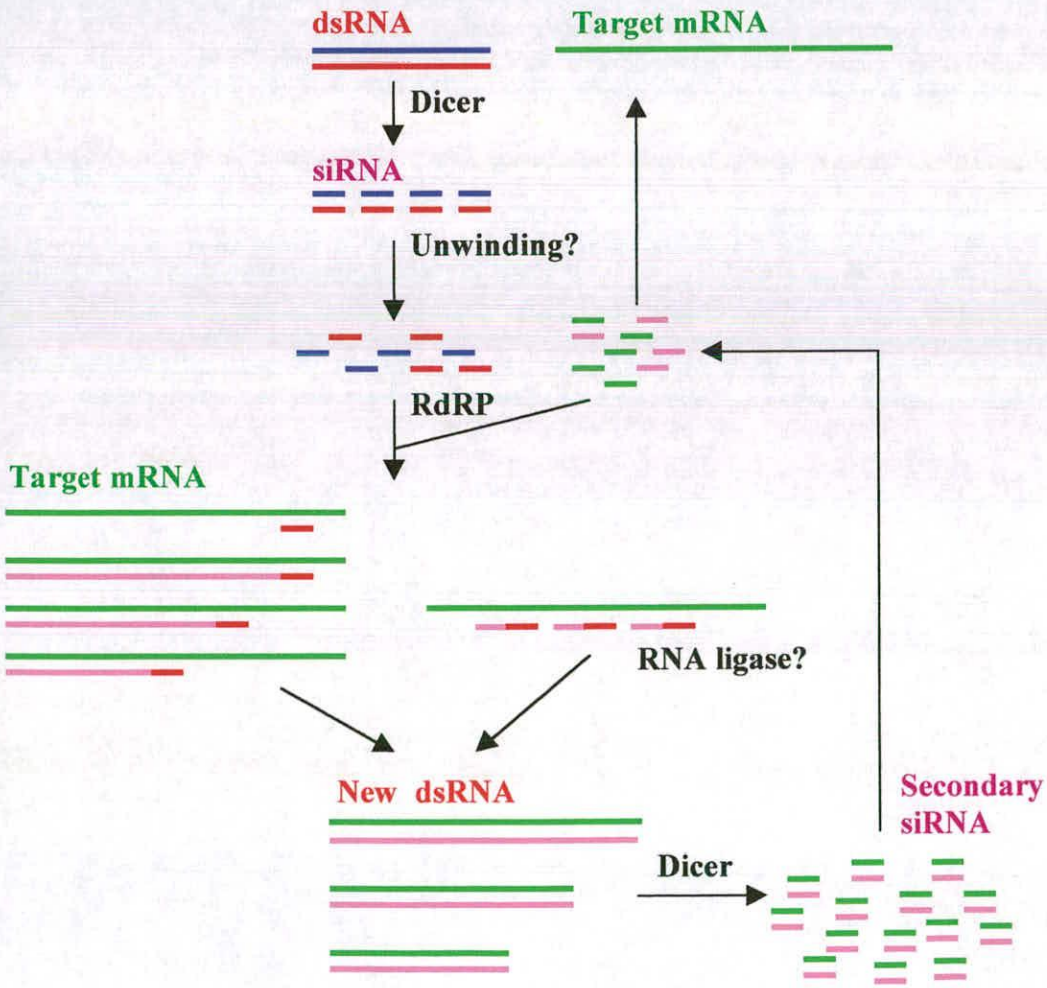
One explanation for the potency of RNAi is that dsRNAs provide templates for RdRP in the production of copy RNA. However, it has since been determined that the templates are predominantly the target mRNA and that the primers are the 21-23 nt fragments generated by cleavage of dsRNA (Lipardi et al.2001).

1.2.4.2.3. RNA dependent RNA polymerase involvement in RNAi potency

The presence of a 3'-hydroxyl group on the siRNAs (Elbashir et al.2001a) and its requirement for RNAi was an early indication that these siRNAs act as primers for an RdRP activity. The role of siRNAs as primers was determined by the incorporation of ^{32}P labelled GFP siRNAs into new full-length cognate dsRNA in a *Drosophila* embryo extract system (Lipardi et al.2001). Both single stranded (antisense RNA) and dsRNA served as templates for primer incorporation. Formation of full-length dsRNAs occurred rapidly reaching a

maximum level within 5-10 minutes. These newly formed dsRNAs were then themselves cleaved generating more siRNAs. Presumably these siRNAs act as additional primers for further dsRNA amplification. Full-length dsRNA appeared to be the dominant product formed from siRNA primed RdRP reaction. This may indicate that multiple partial length RdRP products are ligated on the full-length antisense strand, possibly by an RNA ligase. All these considerations are shown in the proposed model of RNAi as illustrated in figure1.3.

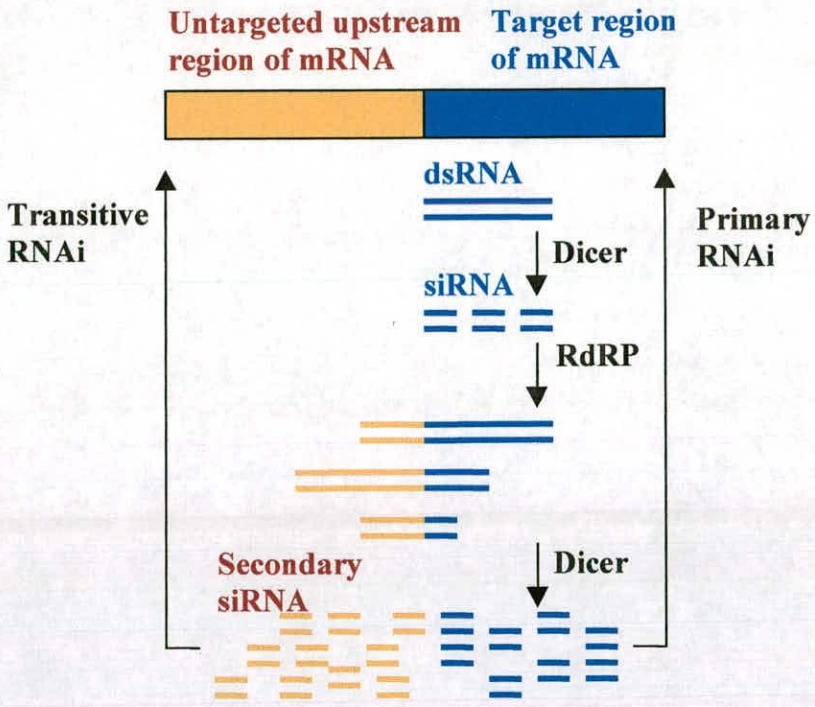
Figure 1.3. A Model for RNAi.
dsRNA is first cleaved into siRNAs by an RNase III related enzyme such as Dicer. The siRNAs act as RdRP primers upon target mRNA generating new dsRNA. This may occur by the alignment of one or many siRNA primers along the template strand and which are extended by RdRP. If many primers are used ligation by a putative RNA ligase is suggested as full length dsRNA appears to be the dominant product. New dsRNA are processed into secondary siRNAs which again act as primers to generate more dsRNA as well as mediate cleavage of the target mRNA.



1.2.4.2.4. Secondary siRNA and Transitive RNAi

The nature of RdRP amplification of the target mRNA predicts that sequences upstream of the target will also be amplified, producing a population of secondary siRNAs. The existence of these secondary siRNAs was determined in *C.elegans* by RNase protection experiments using ^{32}P labelled RNA probes specific to the siRNAs generated by the original dsRNA as well as regions both upstream and downstream of this region (Sijen et al.2001). As anticipated, synthesis of secondary siRNAs was limited to the region upstream of the trigger sequence, where abundance of the secondary siRNA molecules decreased as a function of distance from the primary trigger. Furthermore these secondary siRNAs were also able to induce secondary RNAi, a phenomenon termed transitive RNAi. This was demonstrated by targeting fused transgenes with a dsRNA specific to the downstream gene and observing a down regulation of expression of both. For example dsRNA specific for the lacZ region of a target mRNA encoding a nuclear GFP-lacZ fusion protein reduced nuclear localised GFP expression (Sijen et al.2001). In addition the lacZ specific dsRNA also reduced the expression of a separate mitochondrial localised GFP in a *C.elegans* cell line carrying both transgene constructs. Transitive RNAi by secondary siRNAs is outlined in figure 1.4.

Figure 1.4. Transitive RNAi by Secondary siRNA.
siRNAs used as RdRP primers extend upstream of the original targeted site, generating secondary siRNAs which are capable of promoting transitive RNAi.



1.2.4.3. Expression of dsRNA hairpin from a vector

Double stranded RNA can be introduced into cells by expression as a hairpin loop from an inverted repeat (IR) containing vector. The advantages of this are continuous, stable expression, allowing long term inhibition and the determination of the function of genes involved in the differentiation processes of cells. An inverted repeat transcript folds upon itself, hybridising to complementary regions so that the RNA is double stranded.

Stable silencing has been induced by directed expression of long dsRNA in *C.elegans* and *Drosophila* (Tavernarakis et al.2000; Kennerdell and Carthew 2000; Piccin et al.2001). *In vivo* promoter driven RNAi reproduces null phenotypes with similar efficiency to direct injection of dsRNA, which had

previously been the method of introducing dsRNA. For example, in *C.elegans* ablation of C37A2.5, a gene essential for progression past L2 larval stage, resulted in 82%-100% L2 stage arrested F1 progeny by dsRNA injection, compared to 49%-89% arrest observed with stable dsRNA expression (Tavernarakis et al.2000).

Although long dsRNAs can trigger RNAi effectively in some mammalian cells, this is often masked by non-specific effects, which are mediated by the dsRNA dependent protein kinase, PKR and RNase L. These enzymes form part of the interferon defence pathways that are present in mammals and are activated by long dsRNA (reviewed by Stark et al.1998). PKR is a serine/threonine specific protein kinase that acts by phosphorylating the α subunit of the eukaryotic initiation factor (eIF2). The role of eIF-2 is to form a complex with a methionine charged tRNA and to present it to the ribosomal machinery for translation initiation (Merrick 1992). Phosphorylation of eIF2 prevents the exchange of eIF2-GDP to the GTP bound form, catalysed by eIF2B. The resulting reduction in eIF2-GTP levels leads to an overall block of translation initiation. The other enzyme activated by dsRNA that produces a non-specific ablation effect is RNase L. dsRNA activates RNase L indirectly by first activating 2',5'-oligoadenylate synthetase which results in the increase of the synthesis of 2',5'-oligonucleotide required to activate RNase L. RNase L activation results in non-specific mRNA degradation (Floyd-Smith et al.1981).

Recently Tuschl and colleagues (Elbashir et al.2001b) demonstrated that RNAi can be provoked in numerous mammalian cell lines through the introduction of synthesised 21 nt siRNAs. These siRNAs avoid provoking a PKR

or RNase L response because dsRNA less than 35 bp are unable to activate these molecules (Manche et al.1992; Minks et al.1979). siRNAs are thought to be incorporated into RNAi by mimicking the products of the Dicer enzyme. Synthetic 21 nt siRNAs were able to specifically reduce expression of an exogenous luciferase gene expressed from a vector in NIH/3T3 mouse fibroblast cells, HeLa cells, 293 human embryonic kidney cells and COS-7 cells (Elbashir et al.2001b). The specific ablation observed for siRNAs contrasts with the nonspecific reduction of luciferase expression seen with longer dsRNAs of 50 bp and 500 bp in HeLa cells. siRNAs were also able to reduce protein levels of three out of four tested endogenous genes, lamin A/C, lamin B1 and nuclear mitotic apparatus protein, as visualised by antibody staining and Western blotting. So siRNAs provide an attractive alternative to ablate gene function in mammalian cells if expression of long dsRNA produces non-specific effects. However, the ablation efficiency of siRNAs still depends upon targeting an accessible site on the mRNA (Holen et al.2002).

1.3. A Polyoma based episomal replication system for efficient gene expression in ES cells

1.3.1. Background

The polyoma based episomal replication system allows maintenance of exogenous DNA sequences on episomal plasmids, that is as non-integrated circular DNA. It relies upon the replication cycle of the polyoma virus, a mouse DNA virus, in which the double stranded circular DNA is replicated as

unintegrated minichromosomes. The polyoma virus replicates and transforms its host by expression of four proteins, large tumor (T) antigen, middle T, small T and tiny T from its early region. These proteins are generated from alternative splicing of the same primary transcript. Only large T is essential for replication (Francke and Eckhart 1973), as middle T is required for cellular transformation (Treisman et al.1981), small T has only a minor role in both (Asselin et al.1984; Berger and Wintersberger 1986) and tiny T has an ill defined role (Riley et al.1997). Therefore the extrachromosomal maintenance of vectors constructed to express exogenous DNA requires only expression of large T protein, which can act in trans, by binding to and initiating replication from the polyoma origin of replication (ori).

A polyoma based episome (pMGD20neo, figure 1.5.) was originally constructed by Gassmann and colleagues (1995). pMGD20neo contains a polyoma early region modified to prevent expression of the transforming middle T, as well as a mutant enhancer region, PyF101 (Fujimura et al.1981), originally selected for the ability to replicate the polyoma genome in embryonal carcinoma cells. Electroporation of this plasmid in ES cells was reported to result in episomal replication in 15% of G418 resistant colonies.

It was further reported that a particular ES cell clone containing episomal copies of pMGD20neo, designated MG1.19, was able to efficiently support the replication of a second polyoma ori containing episome. Episomal maintenance resulted in stable transfection efficiencies approaching 1% which is about 100-1000 fold higher compared to non-episomal maintenance, and is termed supertransfection. Work in our laboratory has substantiated these findings (Niwa

et al.1998; I.Chambers unpublished). An ES cell line, E14/T has been established in which copies of the pMGD20neo vector are stably integrated. E14/T cells have indistinguishable self renewal and differentiation properties from parental ES cells. cDNA sequences for the different mRNA ablation strategies to be investigated will be expressed from a second polyoma ori vector by transfection of the E14/T polyoma large T expressing ES cell line.

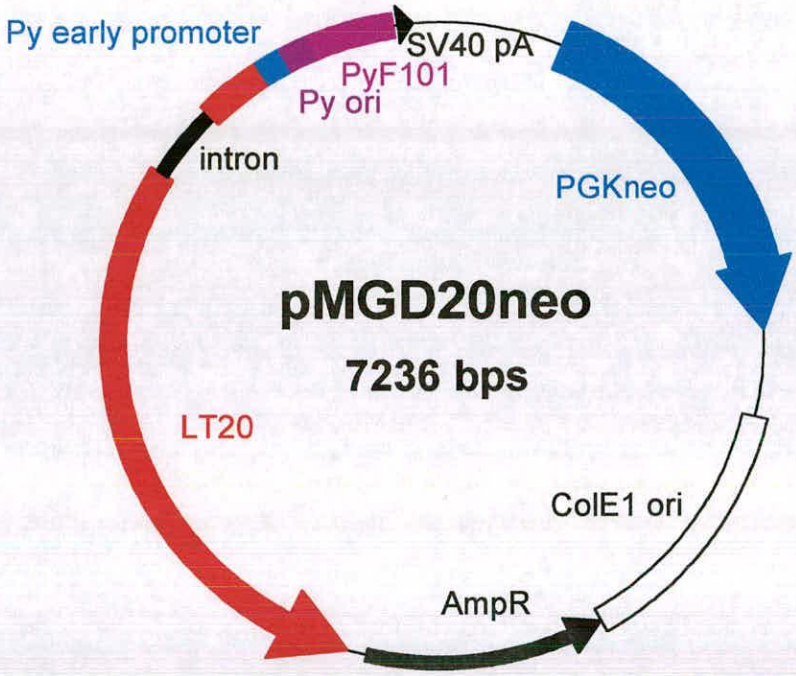
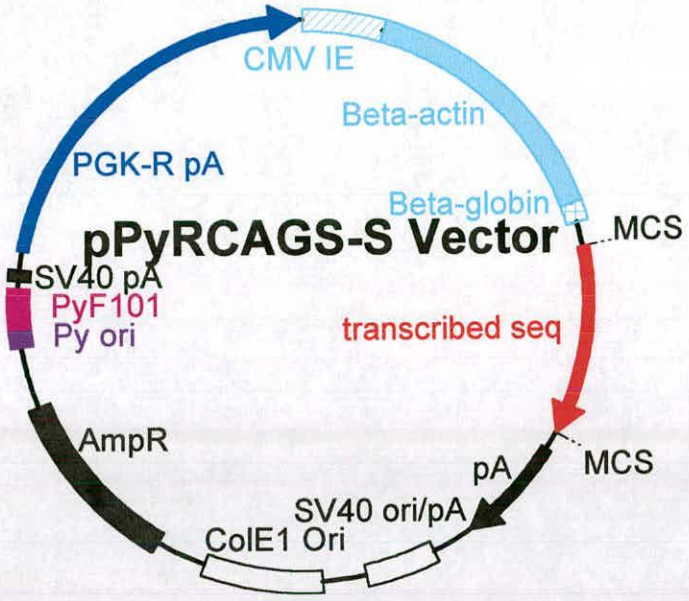


Figure 1.5. The primary plasmid, pMGD20neo.
pMGD20neo contains the mutant enhancer PyF101, the polyoma wild-type ori and modified polyoma early region (LT20) which is without the splice sites for processing of middle and small T mRNAs. The neomycin gene is expressed from the phosphoglycerate kinase (PGK) promoter, conferring resistance to G418. Details of construction of pMGD20neo are described by Gassmann et al. (1995).

The basic design of the polyoma ori episomal vector to be used for expression of ablation cDNAs under investigation is shown in figure 1.6. The vector contains the polyoma ori and the mutated enhancer region, PyF101. The mutated enhancer is required for propagation of the polyoma ori vector in mouse ES cells. The promoter used to drive expression is CAG (Niwa et al.1991) consisting of the human cytomegalovirus immediate early (hCMV-IE) enhancer, and the AG promoter. This is a strong transcriptional regulator that is ubiquitous in a wide variety of cell lines, including ES cells. The AG promoter consists of sequences from chicken β-actin including the promoter and first exon, which is untranslated, and most of the first intron, fused to the 3' portion of the second intron and part of the third exon from rabbit β-globin. The vector is made selectable in ES cells with the inclusion of a PGK driven drug resistance gene encoding either hygromycin phosphotransferase or puromycin N-acetyltransferase. An alternative to a separately driven drug resistance gene is to link transcription of a sequence of interest, cloned into the MCS by placing an internal ribosomal entry site (IRES) linked drug resistance gene between the 3' MCS and the polyadenylation site immediately downstream. Propagation in bacteria is made possible with an ampicillin resistance gene and ColE1 ori present in the vector backbone. While the presence of an SV40 origin of replication allows maintenance of the vector some transformed cell lines, such as C0S-7 cells.

Figure 1.6.
Plasmid map to illustrate all backbone components shown in subsequent vector maps and Cloning strategy diagrams.



Py ori - Polyoma origin of replication
derived from Polyoma virus, for episomal replication of the vector

PyF101 - mutated enhancer
Required for propagation of a polyoma ori vector in ES cells

CMVIE - Cytomegalovirus Immediate Early Enhancer

Beta-actin - Chicken beta-actin Promoter
including the first exon, which is non coding, and most of the first intron

Beta-globin - Rabbit beta-globin
Includes splice acceptor sequence

Sequence of interest to be expressed surrounded by multiple cloning sites

pA- polyadenylation site eiter Beta-globin pA or bovine growth hormone pA

SV40 ori and ColE1 Ori

AmpR - Ampicillin resistance gene, a bacterial selection marker

PGK-R pA - phosphoglycerate kinase 1 promoter driven expression of a resistance gene either hygromycin or puromycin, selectable in ES cells, with a polyadenylation site

1.4. Inducible system for expression of ablation sequences

1.4.1. Background

Inducible systems are often employed to tightly regulate gene expression and would be favourable for the regulation of transcription of the ablation cDNA sequences under investigation. This is important for gene products required for cell viability, for gene function either during differentiation or once cells have differentiated or genes involved in self renewal. Having an inducible system will allow ablation to be controlled in a temporal manner. Also as differentiated cells cannot be propagated an inducible system would allow a replica plate of the uninduced cell population to be maintained for DNA extraction.

1.4.2. Cre-loxP recombination System

The Cre-loxP recombination system was chosen for conditional activation of expression because it is well characterised and functions efficiently in mouse ES cells (Gu et al.1993) as well as in transgenic mice (Orban et al.1992). Moreover it can be used in an episomal context.

The Cre-loxP system is the recombination system from bacteriophage P1 and is so called because it consists of a site on the DNA where recombination occurs called loxP, and the Cre protein, which mediates the reaction. The loxP site is 34 bp in length and consists of two 13 bp inverted repeats separated by an 8 bp asymmetric spacer region. Cre is a member of the integrase family of recombinases, which recognise specific nucleotide sequences, the inverted repeats within the loxP site in the case of the Cre protein. Upon binding to the inverted repeats, Cre synapses with a second lox site on the same segment of DNA,

cleaving the DNA in the 8 bp spacer region initiating strand exchange between the recombining loxP sites. Cre mediated recombination results in either the excision or the inversion of intervening sequences depending on whether the loxP sites are in direct or opposite orientation to each other. If the former occurs a single loxP site is left at the site of excision. It is the excisional properties that are taken advantage of when generating an inducible expression system for sequences of interest. Inducible expression can be achieved by inserting a loxP flanked transcriptional terminating sequence upstream of the sequence to be expressed. The transcriptional terminating sequence prevents expression prior to induction with Cre.

The action of Cre can be made inducible in two ways, either under the control of a tissue specific or regulatable promoter or by fusion to an inducible ligand binding domain, which controls the activity of the protein. An example of the latter is the fusion to a steroid hormone receptor binding domain. In particular a mutated oestrogen binding domain that binds the synthetic oestrogen analogue tamoxifen or 4-hydroxy tamoxifen. The Cre-estrogen receptor binding domain (Cre-ER) fusion proteins are inactive in absence of tamoxifen as they are bound and sequestered in the cytosol by heat shock proteins such as heat shock protein 90 (Hsp90). However, addition of tamoxifen to cells displaces the heat shock proteins, allowing Cre to enter the nucleus and mediate site specific recombination (Mattioni et al.1994). Currently a Cre-ER fusion protein designated Cre-ER^{T2} has been developed which is ~10-fold more sensitive to 4-hydroxy tamoxifen than the previously characterised Cre-ER^{T} fusion protein (Indra et al.1999). 4-hydroxy tamoxifen concentrations as low as 10nM are effective at inducing expression,

while there is no or very little background expression prior to induction (Vallier et al.2001). Conditional gene expression using this Cre-ER fusion protein has been demonstrated in ES cells as well as differentiated cells both *in vitro* and *in vivo* (Vallier et al.2001).

1.5. Endogenous ablation targets in ES cells

1.5.1. Background

To investigate the feasibility of the mRNA ablation strategies to ablate gene function in ES cells an understanding of the genes and signalling pathways involved in ES cell self renewal and differentiation is required. Suitable ES cell gene targets whose functions are already established can then be targeted for ablation and provide a way of determining efficiencies of the different ablation strategies.

1.5.2. Cytokine dependence of ES cells

Historically, ES cells were initially co-cultured with mouse embryo fibroblasts (Evans and Kaufman 1981; Martin 1981), as removal from fibroblast cells caused differentiation of ES cells. Subsequently it was demonstrated that an isolated soluble glycoprotein, leukaemia inhibitory factor (LIF) (Gearing et al.1987) maintained ES cell pluripotency in the absence of fibroblasts (Smith et al.1988; Williams et al.1988; Smith and Hooper 1987). LIF is required for the propagation of self renewing ES cells in culture. In the absence of LIF, ES cells differentiate and pluripotency is not maintained.

LIF is a member of a family of related cytokines including oncostatin M (Rose et al.1994), ciliary neurotrophic factor (Conover et al.1993), cardiotrophin-1 (Pennica et al.1995), interleukin 6 (IL-6) and interleukin 11 (IL-11). LIF, Oncostatin M and cardiotrophin-1 are able to sustain self renewing ES cells as they bind to a heterodimeric complex consisting of two transmembrane glycoproteins, LIF receptor (LIFR) and gp130 (reviewed in Taga and Kishimoto 1997; Heinrich et al.1998). IL-6 and IL-11 are unable to promote self renewal, as the specific receptors are not expressed in ES cells. However, addition of a soluble form of the IL-6 receptor (sIL-6R) with IL-6 promotes self renewal (Nichols et al.1994; Yoshida et al.1994) by forming a hexameric receptor complex acting solely through a gp130 homodimer (Murakami et al.1993). This indicates that signals emanating from gp130 are sufficient for self renewal.

1.5.3. Activation of STAT3 signalling to maintain self renewal of ES cells

LIFR and gp130 are members of the cytokine receptor superfamily. These molecules possess no intrinsic tyrosine kinase activity (Bazan 1990) but this function is provided by members of a tyrosine kinase family called the Janus kinases (JAKs) (Stahl et al.1995). JAKs are constitutively associated with the receptor at proline rich conserved regions within the cytoplasmic domain (Darnell et al.1994). Ligand binding induces homo or hetero dimerisation of gp130 and other class I cytokine receptor family, bringing JAKs into close proximity, inducing their phosphorylation and activation. There are four Janus kinases, JAK 1, JAK 2, JAK 3 and Tyk 2, of which, all except JAK 3 have been shown to bind

gp130 (Stahl 1994). Activated JAKs phosphorylate tyrosine residues on the cytoplasmic domain of gp130, which form binding sites for Src homology 2 (SH2) domain containing molecules (Stahl et al.1995) including the signal transducer and activator of transcription (STAT) family members 1, 3 and 5 (Stahl et al.1995) and protein tyrosine phosphatase SHP-2 (Fuhrer et al.1995). These molecules are then activated by phosphorylation by JAKs. In addition to activation of STATs and SHP-2 other downstream signalling pathways are also activated including the mitogen activated protein kinase (MAPK) pathway (Yin and Yang 1994), the phosphatinositide-3' (PI-3') kinase pathway (Takahashi-Tezuka et al.1998), Src-like kinases Hck and Btk (Ernst et al.1994; Matsuda et al.1995) and insulin receptor substrate (IRS) family proteins (Argetsinger et al.1995).

STAT3 is the principle STAT factor activated by LIF or IL6 plus sIL-6R in ES cells (Hocke et al.1995). STAT proteins have a conserved amino terminus required for dimerization of activated STATs, a central DNA binding domain, an SH2 domain and transactivation domain at the C-terminus (Darnell 1997). Upon receptor binding, STAT3 is phosphorylated on a conserved tyrosine residue Y705, which is required for STAT3 dimerisation and translocation to the nucleus (Darnell et al.1994), where it binds specific DNA sequences, regulating transcription of target genes (Figure 1.7.).

STAT3 activation is directly required for ES cell self renewal. Inhibition of STAT3 activation not only blocks self renewal but promotes differentiation (Niwa et al.1998). This was demonstrated by mutation of STAT3 docking sites in gp130 and by using a dominant interfering mutant of STAT3, STAT3F, where the

tyrosine residue at amino acid position 705 is mutated to phenylalanine. High expression of STAT3F blocks activation of endogenous STAT3 possibly by competing for receptor docking sites in myeloid and M1 leukemia cells (Minami et al.1996; Nakajima et al.1996). ES cells supertransfected with a STAT3F expression vector resulted in predominantly differentiated colonies in the presence of LIF. These colonies resembled differentiated colonies formed in the absence of LIF. Expression of marker genes further confirmed STAT3F expressing ES cells had differentiated (Niwa et al.1998). The MAPK pathway is also activated downstream of gp130. However, ERK signalling unlike STAT3 signalling is not required for stem cell self renewal and may regulate differentiation of ES cells.

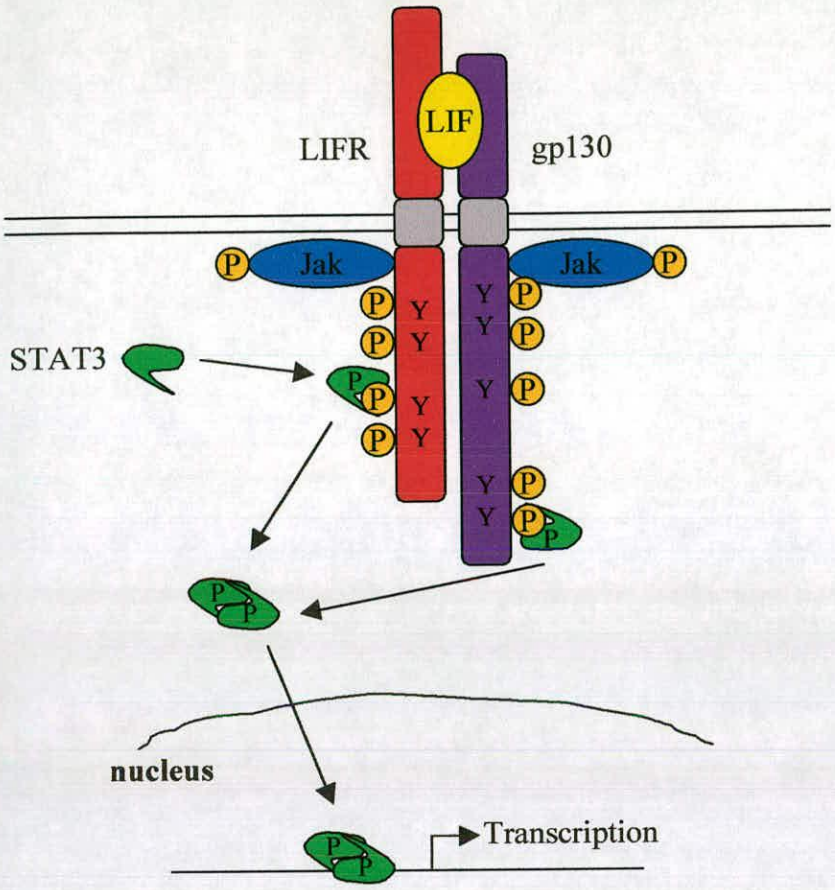


Figure 1.7. Figure showing STAT3 activation by LIF signalling

1.5.4. Negative regulator of STAT3 activation

- Suppressor of Cytokine Signalling 3 (SOCS3)

The duration and intensity of STAT signalling is tightly regulated by induction of STAT inducible STAT inhibitor molecules also referred to as suppressors of cytokine signalling (SOCS) proteins (Bousquet et al. 1999). The SOCS proteins consist of a family of eight members, SOCS1 through SOCS7 and the cytokine inducible SH2 containing protein (CIS) (reviewed in Krebs and Hilton 2000). Each protein contains two regions of homology, a central SH2 domain and a 40 amino acid C-terminal motif designated the "SOCS box". Each has different expression patterns or activity against different signalling pathways. SOCS3 alone is significantly stimulated in ES cells by LIF (I.Chambers, unpublished) and is induced by gp130 signalling by cytokine stimulation, which activates the JAK-STAT pathway inducing SOCS3 transcription (reviewed in Krebs and Hilton 2000). SOCS3 acts by binding to tyrosine Y118 present on gp130, blocking further activation (Nicholson et al.2000). Suppression of gp130 signalling by SOCS3 was impaired in 293T fibroblast cells transfected with a chimeric receptor containing a mutated gp130 intracellular domain, where the tyrosine residue at position Y118 was mutated to phenylalanine (Nicholson et al.2000).

1.5.5. Oct4 - Transcriptional regulator required for maintaining pluripotency of ES cells

Oct4 is a member of a group of related POU transcription factors containing a particular type of bipartite DNA binding domain (reviewed in Ruvkun and Finney 1991). The POU family was originally defined by the sequence homology of three mammalian transcription factors and one nematode regulatory protein, **P**it1/GHF1, **O**ct1, **O**ct2 and **U**nc86 (Herr et al.1988). Oct4 is expressed in oocytes (Rosner et al.1990) and the early embryo, where the maternal protein is present until the end of the two-cell stage (Palmieri et al.1994). Between the four and eight-cell stage zygotic Oct4 gene expression is activated, and Oct4 expression becomes confined to the inner cell mass (ICM) of the blastocyst and later becomes restricted to the germ cell lineage (Scholer et al.1990; Yeom et al.1996).

Oct4 expression is essential in maintaining the identity and pluripotency of the ICM of mouse blastocysts (Nichols et al.1998). Elimination of Oct4 expression results in differentiation along an extraembryonic trophoblast lineage in Oct4 deficient embryos and in ES cells (Nichols et al.1998; Niwa et al.2000). ES cells require a critical level of Oct4 to maintain pluripotency. Reduction of normal Oct4 levels by more than a half induces differentiation into the trophectoderm lineage, while less than a twofold increase results in the differentiation into endoderm and mesoderm (Niwa et al.2000). Oct4 is therefore a master regulator of pluripotency and can be used to unequivocally define undifferentiated ES cells.

1.6. Aims and Objectives

The aim of this thesis was to investigate the effectiveness of ablation strategies for the determination of gene function in ES cells. The three mRNA ablation strategies investigated were antisense RNA, ribozymes and double stranded RNA. These methods of ablation have all been used successfully in many organisms and *in vitro* systems, which suggested their potential in determining gene function in ES cells. To maximise ablation, ablation sequences were stably expressed from episomally replicated vectors which allowed continual expression. Inducible vectors were also designed to make the system more versatile. Effectiveness of the different ablation strategies in ES cells was evaluated by ablating defined targets, chosen to give either easily scorable phenotypes or molecular readouts. The development of an effective ablation strategy has potential not only for discovering the function of a cloned gene of interest but for the identification of new genes in a screen.

Chapter 2: MATERIALS AND METHODS

Unless otherwise stated, analytical grade chemicals were obtained from either Sigma or Fischer chemicals. Analytical grade agarose was supplied by Biowhittaker Molecular Applications. All bacterial media components were supplied by DIFCO laboratories. Synthetic oligonucleotides were synthesised by Oswel DNA Service (University of Southampton, U.K.). Radioisotopes were supplied by Amersham International plc (Little Chalfont,U.K.).

2.1 Molecular Biology Methods

General molecular biology techniques and the preparation of standard solutions were carried out according to Sambrook et al. (1989) unless otherwise stated.

2.1.1. General Cloning Techniques

2.1.1.1. Blunt ending of 5' and 3' overhangs

Cohesive ends were filled in with 1µl of 25µM dNTPs (Amersham Pharmacia) using 1µl of the Klenow fragment of DNA polymerase (2U/µl) (Roche) to a final volume of 100µl with dH_2O. The reaction was incubated for 5 minutes at room temperature and stopped by addition of 5mM EDTA.

2.1.1.2. Gel purification of DNA

2.1.1.2.1. Large fragments

Large restriction fragments of up to 10kb were purified by first running the digested fragments on 1% agarose gels. These gels, containing 0.5µg/ml of ethidium bromide, were cast and run in 1XTAE buffer (40mM Tris, 1mM EDTA,

pH adjusted to pH7.7 with acetic acid), which allows for better separation of larger fragments than TBE buffer. Single bands of interest were isolated from the gel with a scalpel under UV illumination. DNA was purified from the gel slice using QIAEX II gel purification kit (Qiagen) according to the manufacturers instructions.

2.1.1.2.2. Small fragments

The Oct4 5'UTR PCR product of 83bp was isolated for subcloning by resolving on a 4% NuSieve GTG agarose gel. The gel contained 0.5μg/ml of ethidium bromide and was cast and run in 0.5XTBE buffer (44mM Tris, 44mM Boric acid, 1mM EDTA). A PCR marker (Sigma) was loaded onto the gel as a size standard. The fragment was then isolated and purified as for larger fragments (see section 2.1.1.2.1.).

2.1.1.3. Ligations

Appropriate plasmid vectors were digested and gel purified, as were restriction fragments (section 2.1.1.2.). Ligations were set up at either a 1:1 or 1:3 vector:insert molarity ratio with 0.5μl of T4 DNA ligase (Roche) and 1μl of 10X ligase buffer (Roche) to a final volume of 10μl. Cohesive and blunt end ligations were incubated overnight at 4°C. Vector alone ligations including both T4 DNA ligase and ligase buffer were always included as a control for background vector re-ligation.

2.1.1.4. Preparation of Chemically Competent cells

A colony from a freshly streaked Luria Broth (LB) agar (1.5% w/v agar in LB-1% w/v tryptone (Difco), 0.5% w/v yeast extract (Difco) and 5mM NaCl) plate of DH10B cells was used to inoculate 10mls of LB and grown overnight at 37°C with shaking. 1ml of overnight culture was added to 100ml of LB and the culture grown to A_{550}=0.3-0.6 to ensure the cells were in log phase. The culture was then spun at 3,000 x g for 5 minutes at 4°C. The cell pellet was resuspended in 5ml of freezing solution (LB broth pH 6.1, 10% PEG, 5%DMSO, 10mM Magnesium chloride, 10mM Magnesium sulphate, 10% glycerol) at 4°C. Cells were then aliquoted at 500µl into pre-chilled eppendorf tubes and snap frozen in liquid nitrogen before storage at -80°C.

2.1.1.5. Transformation of Bacterial cells

Transformations were carried out to obtain either newly ligated vectors or to re-transform existing vectors from a diminished stock. Frozen competent bacteria were thawed on a mixture of ice and water. 100µl of cells were mixed with a pre-chilled mixture of 1µl of ligation reaction, 20µl of KCM (500mM KCl, 150mM $CaCl_2$, 250mM $MgCl_2$) and made up to 100µl with water. This mixture was incubated on ice for 20 minutes and then at room temperature for 10 minutes. To this 800µl of LB was added and the mixture incubated at 37°C with shaking for 1 hour. Aliquots were then plated onto LB agar plates (1.5% w/v agar in LB) with an appropriate drug selection e.g either 50µg/ml of ampicillin or 25µg/ml of zeocin and incubated at 37°C.

2.1.1.6. TOPO TA Cloning

TOPO TA Cloning was used to directly clone PCR fragments into pCR2.1 vector using the TOPO TA Cloning kit (Invitrogen) according to the manufacturers instructions.

2.1.1.7. Screening transformants

Recombinant bacterial colonies were screened for the presence of inserts by first preparing plasmid DNA (2.1.2.1.) and then cutting with appropriate restriction enzymes to produce informative digest bands when run on agarose gels.

2.1.1.8. In vitro Cre recombinase reaction

Cre recombinase was either from Invitrogen or was a gift from Dr. John Bishop. The Invitrogen enzyme was used according to the manufacturer's instructions with the exception that after inactivation at 65°C the reaction mix was phenol/chloroform extracted and ethanol precipitated. Reactions using Cre recombinase from Dr. John Bishop were performed as follows; 1μg of substrate DNA was incubated with 2μl 10x Buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol pH 7.9 at 25°C) and 2 μl of Cre recombinase to a final volume of 20μl for 15 minutes at 37°C.

2.1.2. Isolation of Nucleic Acids

2.1.2.1. Plasmid Preparation

Small scale plasmid preparation (miniprep) was performed using QIAprep miniprep plasmid isolation kit (Qiagen) using spin columns.

Large Scale plasmid preparations (Maxiprep) were prepared using Qiagen's Maxiprep plasmid purification kit.

2.1.2.2. Isolation of Total RNA

2.1.2.2.1. RNeasy minikit

Total RNA was isolated from ES cells using an RNeasy minikit (Qiagen) according to manufacturer's instructions. RNA concentration was then determined by measuring absorbance at 260nm and purity monitored by measuring the $A_{260/280}$ ratio.

2.1.2.3. Isolation of polyA$^+$ RNA

PolyA$^+$ RNA was purified from total RNA using Oligotex (Qiagen) according to the manufacturer's instructions.

2.1.3. Nucleic Acid transfer to membranes

2.1.3.1. Northern Blot Analysis

2.1.3.1.1. Formaldehyde gel electrophoresis for resolving RNA molecules

RNA was electrophoresed on a denaturing gel prepared by dissolving 2g of agarose in 174 ml of water (1%) before addition of 20 ml of 10x 3-N-Morpholino propaneSulfonic acid (MOPS) buffer and 10.2 ml of 37% formaldehyde. 10x MOPS buffer contains 200 mM MOPS, 50 mM sodium acetate and 10 mM EDTA, pH 7. RNA samples were prepared for loading onto a formaldehyde gel by adjusting the volume to 10 µl and adding 25 µl of sample buffer and 0.125 µl of ethidium bromide (10mg/ml). Sample buffer contains 50%v/v formamide, 2.2M formaldehyde, 1x MOPS, 5% glycerol and 0.125% bromophenol blue. Samples were heated to 70°C for 10 minutes and allowed to cool before loading. 6µg of 0.24-9.5kb RNA ladder (GibcoBRL-Life Technologies) was treated in the same way. Electrophoresis was performed at 80V for 4-5 hours.

2.1.3.1.2. Northern Blotting

Nucleic acids were transferred to Hybond-N nylon membrane (amersham pharmacia biotech) using a capillary transfer technique. This involves a dish containing 20x SSC with a sheet of 3MM paper pre-soaked in 20x SSC resting on a glass plate above the dish with the ends immersed in the SSC. The gel, washed in dH_2O, was inverted and placed on the 3MM paper onto which the pre-cut nylon membrane was placed. The top right hand corner of the filter was cut as an orientation aid. Ten sheets of 3MM paper cut to the same size as the gel and pre-

soaked in 2x SSC were placed on the filter. Finally, paper towels were stacked on the 3MM paper and a weight was placed on top, ensuring good contact throughout the assembly. The transfer was left overnight before disassembly, when the filter was washed in 2x SSC and baked at 80°C for 30 minutes.

2.1.4. Radiolabelling Probes

2.1.4.1. DNA Probes

20 ng of purified probe was made up to 12 µl with dH_2O and denatured by boiling for 5 minutes. When cooled, 4 µl of High Prime (Roche Molecular Biochemicals) and 4 µl of $\alpha^{32}P$ dCTP (10 mCi/ml; 3000Ci/mmol) were added and incubated at 37°C for 30 minutes. The reaction volume was then adjusted to 100 µl and unincorporated nucleotides were removed by centrifugation of the reaction through a G-50 column at 2000 x g for 2 minutes. Specific radioactivity of probes was determined by readings in ULTIMAGOLD using a liquid scintillation analyser (TRI-CARB®).

2.1.4.2. Oligonucleotide Probes

50 ng of synthesised oligonucleotide probe (Oswel, University of Southampton, UK.) was end labelled with 4 µl γATP ^{32}P (10 mCi/ml; 5000Ci/mmol) by including 1 µl of T4 polynucleotide kinase (Roche), 2 µl of 10x buffer (500 mM Tris-HCl, 100 mM $MgCl_2$, 1 mM EDTA, 50 mM dithiothreitol, 1 mM spermidine) and 12 µl of dH_2O in the reaction and incubated for 45-60 minutes at 37°C. The reaction volume was adjusted to 100 µl and unincorporated nucleotides

were removed by centrifugation of the reaction through a G-25 column at 2000 x g for 2 minutes. Specific radioactivity of probes was determined by readings in ULTIMAGOLD using a liquid scintillation analyser (TRI-CARB®). Probes with a specific radioactivity in the range of $1x\ 10^8$– $1x\ 10^9$ cpm per μg were used for hybridisation.

2.1.5. Hybridisation

2.1.5.1. Double stranded DNA Probes

All hybridisations with DNA probes were performed at 65°C in Techne hybridisation bottles in a Teche HB-1 oven. Filters were prehybridised for at least 30 minutes in 50 ml of Church and Gilbert hybridisation buffer (0.5 M Na_2HPO_4; 7% SDS; 1 mM EDTA). The double stranded radiolabelled probe was denatured at 100°C for 5 minutes, and added to the hybridisation buffer and incubated overnight. Following hybridisation, filters were washed for approximately 20 minutes sequentially in three wash solutions, as required, (A- 44 mM Na_2HPO_4, 5% SDS, 1 mM EDTA; B- 44 mM Na_2HPO_4, 1% SDS, 1 mM EDTA; C- 40 mM Na_2HPO_4 and 0.1% SDS). Filters were wrapped in Saran Wrap and exposed to either autoradiographic film at -70°C for an appropriate length of time until a signal appeared or a PhosphorImager screen for 1-2 days (FLA-3000 PhosphorImager).

2.1.5.2. Single stranded Oligonucleotide Probes

Hybridisation with an Oligonucleotide probe was performed as described for DNA probes except that the hybridisation temperature was 48°C. Filters were then washed twice in 6X SSC at room temperature for 10 minutes.

2.1.6. Stripping Membranes

All membranes were stripped by submerging into 1% of boiled SDS (100°C), and allowed to cool to room temperature. The efficiency of stripping was checked by autoradiography prior to re-probing.

2.1.7. DNA Sequencing

2.1.7.1. Automated Cycle Sequencing

Sequencing of DNAs was performed by the CGR sequencing facility.

2.1.7.2. Sequencing Primers

#	Description	Sequence
36	T7	AAT ACG ACT CAC TAT AG
37	T3	ATT AAC CCT CAC TTA AG
164	CAG cassette downstream (complementary to the 3rd exon of rabbit β-globin included in the CAG cassette)	TGC TGG TTG TTG TGC TGT
153	V9547 (complementary to rabbit β-globin polyadenylation signal)	ACC ACC TTC TGA TAG GCA
232	5'p300RIB	GGT TTT TTG TGT GGG CTT GG
233	3'p300RIB	GTT AGC TCA CTC ATT AGG CA

2.1.8. Polymerase Chain Reaction (PCR)

All PCRs were carried out on GeneAmp® PCR System 9700 (PE Applied Biosystems). Reactions were assembled on ice containing 100 ng DNA template, 300 ng of each primer, 10 mM dNTPS (Amersham Pharmacia), 5 µl of 10x reaction buffer (provides a final concentration of 1.5 mM $MgCl_2$) and 0.3 µl Taq Polymerase (1.5units; Qiagen) where the final volume is adjusted to 50µl with dH_20. PCRs were used to amplify the UTR sequences of LIFR, STAT3 and Oct4 from either vectors available in the lab with the appropriate cloned sequence or genomic DNA prepared from ES cells. The basic cycling parameters are 94°C 5 mins, annealing temperature 12 s, 72°C 30 s and 72°C 10 mins. PCR products were analysed by gel electrophoresis. Details of the source and cycling parameters used that were individually determined are described below. Subsequently these fragments were cloned using the TOPO TA Cloning kit (section 2.1.1.6.) except the STAT3 5' UTR where a single band was not obtained. In this case the fragment of the anticipated size was excised from a 2% agarose gel and purified (section 2.1.1.2.1.). 0.2 ng of the purified fragment was amplified by PCR and cloned into pCR2.1.

Gene Target		Source of DNA for PCR (Lab No. of vectors)	Accession Numbers	Cycling temperature (°C)	Number of Cycles
LIFR	5'UTR	Jo2 (311)	AF014933	45	20
	3'UTR	pDIARg M2 BS2.1 (301)	S81861	45	25
STAT3	5'UTR	E14TG2a ES cells	U08378	43	25
	3'UTR	STAT3 PRC/cmv (735)	U06922	41	20
Oct4	5'UTR	Oct4 pNLacF (215)	S58422S1	60	20
	3'UTR	pBluescript II SK-Oct4 (121)	X52437	45	25

2.1.8.1. Primers used in PCR

#	Description	Sequence
238	LIFR 5'UTR -5'	GCGGCCGCCGGAGCGTCTCGGCAGT
239	LIFR 5'UTR -3'	ACGCGTTGCCATTGTCCTTGCTGTT
240	LIFR 3'UTR -5'	GCGGCCGCCCAGGTCACCCTTTGTCA
241	LIFR 3'UTR -3'	ACGCGTACAGTAGGCACAGAAATCAG
253	STAT3 5'UTR-5'	GCGGCCGCCTGGAGGGGCTGTAATTCA
254	STAT3 5'UTR-3'	ACGCGTAGCCATCCTGCTGCAGTCAGGGGTCTCGAC TGTCTCCGGGGCTGAGG
245	STAT3 3'UTR-5'	GCGGCCGCGGAGCTGAAACCAGAA
246	STAT3 3'UTR-3'	ACGCGTAATTTAAAGAGGAACCTC
247	Oct4 5'UTR-5'	GCGGCCGCGAAACCGTCCCTAGGTGAGC
248	Oct4 5'UTR-3'	ACGCGTAGCCATGGGGAAGGTGGGCA
242	Oct4 3'UTR-5'	GCGGCCGCGGCACCAGCCCTCC
243	Oct4 3'UTR-3'	ACGCGTACTGTGTGTCCCAGTCTTTATTTAAGAACA A

2.1.9. RNAi Oligonucleotides

#	Description	Sequence
293	GL2F (*Photinus pyralis* luciferase)	CGUACGCGGAAUACUUCGATT
294	GL2R (*Photinus pyralis* luciferase)	UCGAAGUAUUCCGCGUACGTT
295	GL3F (*Photinus pyralis* luciferase)	CUUACGCUGAGUACUUCGATT
296	GL3R (*Photinus pyralis* luciferase)	UCGAAGUACUCAGCGUAAGTT
297	RLF (*Renilla reniformis* luciferase)	AAACAUGCAGAAAAUGCUGTT
298	RLR (*Renilla reniformis* luciferase)	CAGCAUUUUCUGCAUGUUUTT

2.2. Tissue Culture

2.2.1. ES Cell Culture

Methods for the routine culture of ES cells are based upon those described by Smith (1991). All ES cell culture was performed using sterile techniques within a laminar flow hood (GELAIRE®, Flow Laboratories). Cells were incubated in a humidified incubator (Heraeus Instruments) at 37°C with 7.5% CO_2. All solutions were tested for sterility and pre-warmed to 37°C prior to use. ES cells were viewed under a microscope (Olympus CK2).

2.2.1.1. Reagents

Media for growing ES cells is 1X GMEM containing 15% sodium bicarbonate, 0.1% MEM non essential amino acids, 4 mM glutamine, 2 mM sodium pyruvate, 0.1 mM 2-mercaptoethanol and 10% FCS. ES cells were maintained in an undifferentiated state by addition of LIF to the media at 100U/ml. LIF conditioned medium was prepared within the lab by transient expression of murine or human LIF expression plasmids in COS-7, as described by Smith (1991). IL6 together with soluble IL6/R (sIL-6R) also maintains ES cells in an undifferentiated state (Yoshida et al.1994). IL6 was prepared by transient transfection of a human IL-6 expression plasmid in COS-7 cells. sIL-6R conditioned medium was prepared by selecting in methotrexate, CHO/H27 cells transiently transfected with a sIL6-R expression construct linked to the dihydrofolate reductase (Dhfr) gene (Saito et al.1991).

2.2.1.2. Thawing ES cells

Frozen ES cell vials were taken directly from liquid nitrogen storage and thawed immediately in a 37°C water bath. The cells were transferred to a 20ml centrifugation tube to which 10ml of pre-warmed medium was added and then centrifuged at 250 x g for 3 minutes. The pellet was resuspended in 10ml of medium and transferred to a gelatinised 25 cm^2 flask. Medium was changed the next day to remove dead cells.

2.2.1.3. Passage and Expansion of ES cells

Cells were normally passaged every two days. The culture medium was aspirated and cells rinsed with PBS. 1ml of trypsin solution was added to cells grown in a 25 cm^2 flask (2mls for 75 cm^2 flask; 3mls for 175 cm^2 flask) and incubated at 37°C for 2-3 minutes until the cells detached from the plastic. 9mls of medium was added to neutralise 1ml of trypsin (18 mls or 17 mls for 2 or 3 mls of trypsin solution respectively) and the cell suspension centrifuged for 3 minutes at 250 x g. The cell pellet was resuspended in ES cell medium, pipetted vigorously to ensure a single cell suspension and transferred to freshly gelatinised flasks at a dilution of 1:10.

2.2.1.4. Freezing ES cells

Cells were trypsinised and pelleted as described in section 2.2.1.3. and counted in a haemocytometer by resuspending in an appropriate amount of medium. After centrifugation, the cell pellet was resuspended at a density of $5x10^6$ cells/ml in freezing solution (10% DMSO in cell culture medium) and 1ml was aliquoted into cryotubes (Nunc). Cryotubes were transferred to -80°C overnight and the next day removed to liquid nitrogen for long term storage.

2.2.1.5. Transfer of DNA to ES cells

2.2.1.5.1. Supertransfection of E14/T ES cells

E14/T cells were trypsinised (section 2.2.1.3.) washed twice with PBS and counted. Cell density was then adjusted to $6.25x10^6$/ml and 0.8 ml ($5x10^6$ cells) mixed with 20 μg of episomal vector in an electroporation cuvette and incubated

at room temperature for 3 minutes. The mixture was then pulsed using a Gene Pulser (Bio-rad) set at 0.2 kV and 960 μF, typically giving a time constant of 13-15. The cells were then immediately transferred to a centrifuge tube containing 24.2 ml of warmed ES cell medium and mixed thoroughly. Cells were plated at an appropriate density, usually $5x10^4$ and $2x10^5$ per 9cm gelatinised plate and incubated at 37°C overnight. Appropriate selection was applied 30 hrs after and subsequently replaced every 2 days.

2.2.1.5.2. Effectene

Transient transfection of ES cells was achieved by using a non-liposomal reagent, Effectene (Qiagen) following the manufacturers guide. $1x10^6$ E14/T ES cells were plated per well of a 6-well plate in 3 mls of medium containing 100U/ml LIF. For each transfection 20 μl of DNA (3 μg), 24 μl of enhancer and 430 μl of EC buffer were pipetted into an eppendorf and vortexed for 1 second. This mixture was left for 2 minutes at room temperature before transferring into an eppendorf containing 25μl of effectene reagent. Again the mixture was vortexed, this time for 10 seconds before leaving at room temperature for 9 minutes to allow DNA-Effectene complexes to form. The transfection complex was then added to 4 mls of pre-warmed medium containing 100U/ml LIF and transferred to the plated cells. Cells were incubated at 37°C for an appropriate length of time before analysis.

2.2.1.5.3. LIPOFECTAMINE 2000

Lipofectamine 2000 (LF2000) (Life Technologies-GibcoBRL) was used to transiently co-transfect COS-7 cells, E14/T and E14TG2a ES cells with firefly luciferase (pGL3-Control, Promega) and sea pansy luciferase (pRL-TK, Promega) reporter plasmids and siRNA. The annealing of siRNAs was achieved by incubating 20 μM single stranded RNAs in annealing buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) for 1 minute at 90°C followed by 1hr at 37°C.

A day before transfection cells were trypsinised and seeded into 24-well plates at $8x10^4$ for COS-7 cells or $2x10^5$ for ES cells. Cells were plated in 0.5 ml of GMEM medium containing 10% FCS, where ES cells were additionally supplemented with 100U/ml LIF. The medium was replaced on the day of transfection. For each well 1 μg pGL3-Control, 0.1 μg pRL-TK and 0.21 μg siRNA duplex were diluted into 50 μl of serum free GMEM. This mixture was then added to 3 μl of LF2000 diluted in 50μl of serum free GMEM that had been previously incubated at room temperature for 5 minutes. The final mixture was incubated at room temperature for 20 minutes before adding to cells. Cells were incubated for 20 hrs and assayed for luciferase expression with the Dual luciferase assay (Promega) as outlined by the manufacturer.

2.2.1.6. Induction of ES cell differentiation

2.2.1.6.1. Retinoic acid Induction

ES cells were seeded at $1x10^3$ or $5x10^3$ per well of a 6-well plate and allowed to grow for 1 day. Stock retinoic acid at 10^{-2} M (Aldrich) was diluted to 10^{-6} M or

10^{-7} M in ES cell medium supplemented with 100U/ml LIF. The culture medium was replaced with this medium and changed every 2 days for up to 6 days.

2.2.1.7. Induction of ES cells

2.2.1.7.1. SOCS3 Inductions

For SOCS3 inductions, ES cells were trypsinised and plated at $1x10^6$ density into 25 cm^2 flasks. The next day medium was replaced with GMEM containing 10% FCS without LIF for 1 hr. Cells were then washed and incubated in GMEM alone for 6 hrs. Medium was then replaced with GMEM alone or supplemented with 100U/ml LIF for the specified length of time, at which point cells were lysed for RNA analysis.

2.2.2. Staining of ES cells

2.2.2.1. Leishman Staining

ES cells were stained with Leishmans reagent (BDH) by removing the cell culture medium and replacing it with enough Leishmans reagent to cover the surface of the plate. The cells were left at room temperature for 5 minutes before diluting the stain with an equal volume of water, which again was left for 5 minutes before rinsing with water.

2.2.2.2. Alkaline Phosphatase Staining

Alkaline phosphatase staining was achieved by using the Leukocyte Alkaline Phosphatase kit (Sigma Diagnostics) following manufacturer's instructions. The

Alkaline Phosphatase assay kit works by reacting a sodium nitrite solution with a fast red violet Alkaline solution to form a diazonium salt solution. This in combination with a solution containing naphthol AS-BI phosphate stains sites of alkaline phosphatase activity red, the colouration is dependent upon the choice of diazonium salt.

2.3. Flow Cytometry

2.3.1. Preparation of cells for Flow Cytometry

Cells were trypsinised as described in section 2.2.1.3. The cells were counted in a haemocytometer, resuspended at $1x10^6$ cells/ml in ice cold PBS/10% FCS and analysed on FACSCalibur (Becton Dickinson).

2.3.2. Staining with Annexin V-PE and 7-AAD

Cells were prepared for staining as described in section 2.3.1. and stained with AnnexinV-PE and 7-AAD from the Annexin V-PE Apoptosis Detection Kit I according to the manufacturer's instructions (BD Biosciences).

Chapter 3:

Construction of an inducible episomal expression system in ES cells

3.1. Introduction

The feasibility of antisense, ribozyme and dsRNA as efficient mRNA ablation strategies in ES cells are investigated in subsequent results chapters. These sequences are to be introduced by continuous expression from an episomally maintained vector. This has the advantage that potential ablation effects will last longer compared to transfection of pre-synthesised ablation sequences and therefore produce clearer ablation phenotypes. This chapter focuses on developing an inducible episomal system with which to direct expression, giving a greater versatility to the expression system.

Generation of an inducible vector to express RNA ablation sequences is favourable as the knock down of gene expression is made regulatable. Having an inducible system will allow ablation to be controlled in a temporal manner. This is important for determining the function of genes in ES cells especially if the gene products are required for either cell viability or self renewal. An inducible system will allow cells transfected with vectors expressing ablation sequences to be propagated and maintained prior to induction so that ablation of genes required for self renewal will produce a differentiated phenotype, while genes required for cell viability will result in a reduction in colony number. An inducible system can also be potentially used to express ablation sequences in a screen to identify genes

involved in the self renewal of ES cells. If a successful ablation strategy is developed in ES cells, inducible expression will allow a replica plate of the uninduced cell population to be maintained for the isolation and characterisation of cDNA sequences responsible. Ablation of genes required for self renewal will cause ES cells to differentiate, which may limit the yield of recoverable episomal DNA. Generation of an inducible episomal expression system is therefore important. Additional applications of an inducible expression system is its possible use for temporal or spatial over or mis-expression of other interesting cDNA sequences.

The inducible system investigated in this chapter makes use of the Cre-loxP recombination system. This was done by generating an inducible episomally maintained vector by inserting an effective floxed transcription terminating sequence between the promoter and the cDNA to prevent its expression. Excision of the floxed sequence was then induced with Cre recombinase allowing the subsequent transcription and expression of the cDNA sequence downstream of the promoter.

3.2. Experimental Strategy

3.2.1. Design of an Inducible vector

The inducible vector contains a polyoma origin of replication and the CAG promoter to drive expression (described in section 1.3.1.). The test sequence to be expressed from the vector is the enhanced form of the green fluorescent protein (eGFP). eGFP is used because its expression is easily detected and can

therefore be used to investigate whether the transcriptional termination sequence is working efficiently. Subsequently eGFP could be replaced with appropriate sequences of interest. A floxed transcriptional terminating sequence is inserted between the promoter and eGFP, consisting of a synthetic polyadenylation (SPA) signal followed by a C2MAZ sequence. A functional polyadenylation signal is an essential component of the RNA polymerase II termination signal (Proudfoot and Brownlee 1976) while sequences located downstream of the polyadenylation site have also been implicated in the termination process (McDevitt et al.1984). The SPA C2MAZ transcription terminating sequence was chosen because it efficiently prevents transcription *in vitro* (Eggermont and Proudfoot 1993). It was found that the C2MAZ sequence potentiates polyadenylation at the upstream SPA site and that the effect is additive as inclusion of both sequences terminates transcription more efficiently than either sequence alone (Eggermont and Proudfoot 1993). Eggermont and Proudfoot (1993) assayed for transcription termination by constructing an expression vector containing two HIV-1 long terminal repeats (LTRs) in tandem each driving the expression of an α2 globin gene. Transcription of each α2 globin gene from this vector resulted in transcriptional interference of the LTR promoters when transfected into HeLa cells, which was alleviated by insertion of the SPA and C2MAZ sequences between the globin genes (Eggermont and Proudfoot 1993). Initially the SPA and C2MAZ sequences were characterised individually, as summarised below.

The SPA signal is based upon the highly efficient polyadenylation signal of the rabbit β-globin gene and contains the AATAAA sequence, a GT-rich sequence and a T-rich sequence (Levitt et al.1989). The SPA signal differs from

the rabbit β-globin polyadenylation signal in that the GT and T rich sequences normally separated by two AA nucleotides are fused. In addition sequences between AATAAA and the GT/T-rich sequence differ slightly, although the correct spacing is maintained. Inclusion of these three elements was found to be required for efficient polyadenylation (Levitt et al.1989). The SPA signal is better than natural polyadenylation sites that may not have either a GT- or T-rich downstream element or that possess only one. For example when the SPA signal was placed 3' to the polyadenylation site of the human α2-globin gene, the SPA site was used exclusively, indicating that SPA is a more efficient polyadenylation site than α-globin, which does not possess a T-rich downstream sequence (Levitt et al.1989).

The C2MAZ sequence was discovered between two closely spaced human complement genes, C2 and Factor B, located in the MHC class III locus by Ashfield et al.(1991). These genes are expressed simultaneously in the adult liver, but are very close together, 421 bp from the C2 polyadenylation site to the Factor B cap site (Wu et al.1987). This suggested that there may be efficient termination of the C2 transcripts as read-through of a polymerase into the promoter of the downstream gene, Factor B might have been expected to result in a down-regulation of that gene. This is known as transcriptional interference or promoter occlusion and has been observed *in vitro* and *in vivo* (Eggermont and Proudfoot 1993; Cullen et al.1984). In addition to the polyadenylation signal of the C2 gene a functionally defined sequence was found downstream, 160 bp in length required for efficient termination of the C2 transcript. Subsequently this sequence was found to bind the zinc finger protein MAZ, which was shown to halt or slow the

progression of the transcribing polymerase by bending the DNA to which it is bound (Ashfield et al.1994). MAZ dependent termination occurs in the presence of an upstream polyadenylation site, so the SPA site was utilised in the transcriptional terminating sequence. Hence the combined transcriptional terminating sequence is referred to as SPA C2MAZ.

The pPyHCAGeGFP vector was initially constructed in two steps. In the first step eGFP from the commercially available vector, peGFP-N1 was cloned as a SalI/NotI fragment into pCAGGS-ori (lab number 577) digested with XhoI/NotI to generate an eGFP expression vector driven by the CAG promoter (see appendix 1). The second step involved subcloning the CAG eGFP expression cassette from the newly cloned pCAGGSeGFP plasmid (lab number 444) as a SalI/SfiI fragment into the backbone of pPHCAGGS to generate a CAG eGFP expression vector also containing a polyoma origin of replication and a hygromycin resistance cassette (see appendix 2). The completed pPyHCAGeGFP vector is shown in figure 3.1.

The transcriptional terminating sequence was cloned into pBluescript loxP (lab number 885) between two loxP sites using BamHI restricted fragments (see appendix 3). The loxP sites are in the same direction as each other and the terminating sequence is inserted in the opposite orientation with respect to the loxP sites. This is important because the loxP sequence contains an ATG in the 5' to 3' direction which can potentially act as an alternative start codon to that of the gene to be expressed. Therefore the loxP sequences have to be inserted in a 3' to 5' orientation with respect to the direction of expression from the CAG promoter.

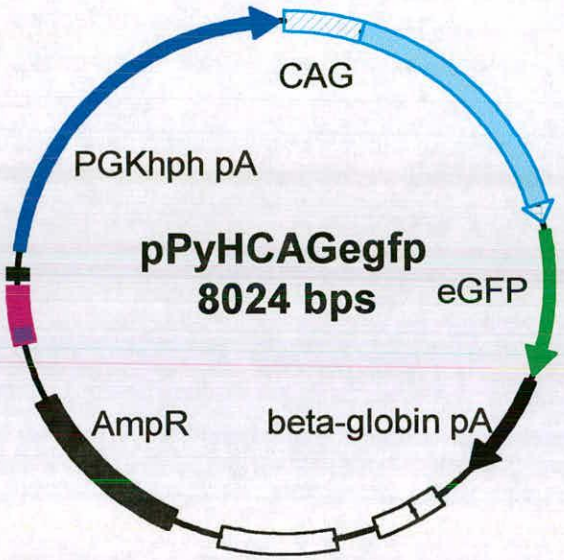


Figure 3.1. pPyHCAGeGFP Vector

The CAG promoter contains an intron derived from a splice donor site from the chicken β-actin gene and a splice acceptor site from the rabbit β-globin gene (see figure 3.2.). In order to effectively prevent eGFP expression the floxed SPA C2MAZ sequence should be cloned between the CAG promoter and the eGFP gene. There is therefore an opportunity to determine whether placement of the termination cassette in an intronic or an exonic location would differentially affect expression of downstream sequences. Plasmids in which the SPA C2MAZ was placed in the CAG intron or downstream of the intron were therefore generated (Appendices 4 and 5). This was done by cloning the floxed SPA C2MAZ sequence as an AvrII fragment into either the intron or the exon of the CAG promoter present in vector pPyHCAGeGFP (Lab number 719) linearised with XbaI or SmaI respectively. The terminating sequence is upstream of eGFP in both cases.

3.3. Results

3.3.1. Testing transcriptional termination by SPA C2MAZ in intronic and exonic locations

The inducible eGFP expressing vector, pFlox-C2MAZ-Exon was electroporated into E14/T, the polyoma large T protein expressing cell line. Cells were grown in LIF (100 units/ml) and G418 (200 μg/ml) and selected in hygromycin (200 μg/ml) for 7 days until stable transfectants were obtained. At this hygromycin concentration control untransfected E14/T cells were killed after 7 days. This is in good agreement with the minimum effective concentration of

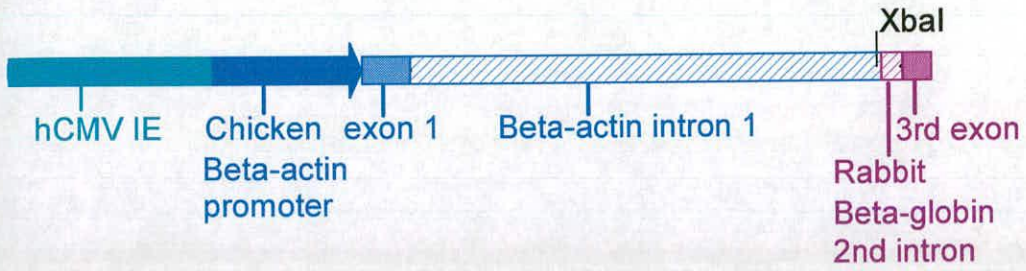


Figure 3.2. The CAG cassette construct.
The CAG cassette consists of the human cytomegalovirus immediate early (hCMVIE) enhancer, and the AG promoter. The previously characterised AG promoter (Miyazaki et al.1989) was fused to the CMV IE enhancer sequence, to generate a strong transcriptional regulator that is ubiquitous in a wide variety of cell lines, including ES cells. The AG promoter consists of sequences from chicken β-actin including the promoter and first exon, which is untranslated, and most of the first intron. This is fused to the 3' portion of the second intron and part of the third exon from rabbit β-globin.

160 μg/ml required to kill untransfected cells after ~10 days. $5x10^4$ and $2x10^5$ cells were plated on three 9cm plates at each density. Colonies were viewed under bright field and fluorescence and images were taken. No fluorescent colonies were detected for ES cells maintaining the pFlox-C2MAZ-Exon vector (figure 3.3.a). Identical results were obtained from a duplicate experiment where there were approximately 20 colonies growing on plates plated at $5x10^4$ density and ~70 colonies growing on plates plated at $2x10^5$. This indicates that the transcriptional terminating sequence is effective and is able to prevent expression when inserted in an exonic location. As the pFlox-C2MAZ-Exon vector proved successful in preventing expression of a cloned gene, this vector was used in a subsequent analysis to determine if the floxed C2MAZ sequence could be efficiently removed with Cre. This was done by electroporating E14/T cells maintaining the pFlox-C2MAZ-Exon vector with a Cre expression vector, pPyCAGCreIP (lab number 759). Cells were selected in LIF (100units/ml), hygromycin (200 μg/ml) and puromycin (2 μg/ml) for 4 days and examined by fluorescence microscopy (figure 3.3.b). Expression of GFP indicates that the floxed C2MAZ sequence has been removed by Cre expression. Approximately one in five colonies now expresses eGFP, where 20 were the total number of colonies growing on plates plated at $5x10^4$ density. Therefore an inducible expression system has been developed, activatable with the expression of Cre protein.

The other eGFP inducible expression vector generated was pFlox-C2MAZ-Intron, where the floxed transcriptional terminating sequence was cloned within the intron of the CAG cassette. This vector was electroporated into E14/T cells and selected in LIF, G418 and hygromycin in an identical experiment to that

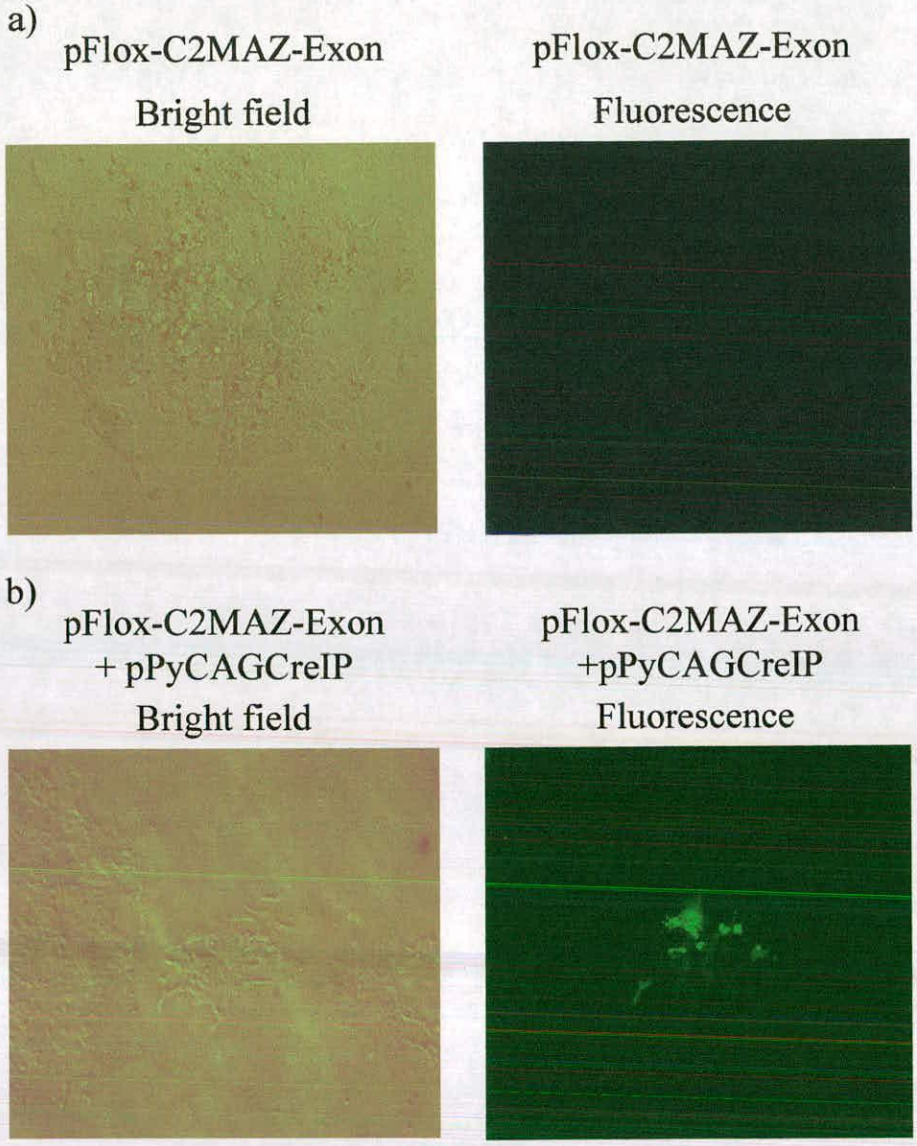


Figure 3.3. Determining the effectiveness of the floxed transcriptional terminating sequence C2MAZ in preventing expression of the downstream gene-eGFP, when cloned within the exon of the CAG cassette and its subsequent removal with Cre expression.

E14/T ES cells were electroporated with pFlox-C2MAZ-Exon, where the floxed C2MAZ sequence is cloned within the exon of the CAG cassette. Cells were selected in LIF, G418 and hygromycin (200 μg/ml) for 7 days. a) Colonies were viewed under bright field (left panel); with the corresponding image under fluorescence (right panel). b) ES cells stably maintaining pFlox-C2MAZ-Exon were then electroporated with a Cre expression vector, pPyCAGCreIP. Cells were selected in puromycin (2 μg/ml) for 4 days and viewed under bright field (left panel) and fluorescence (right panel).

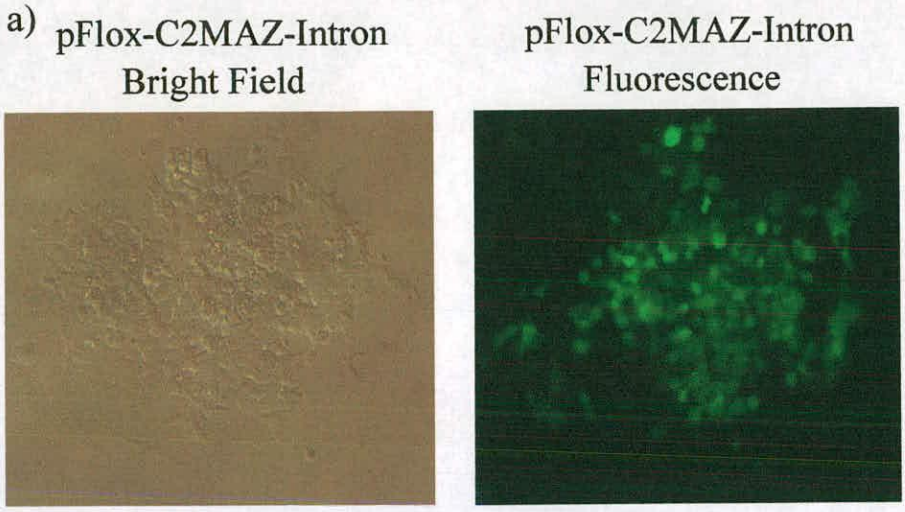


Figure 3.4. Determining the effectiveness of the transcriptional terminating sequence C2MAZ in preventing expression of the downstream gene-eGFP, when cloned within the intron of the CAG cassette.

E14/T ES cells were electroporated with pFlox-C2MAZ-Intron. Cells were selected in LIF, G418 and hygromycin (200 μg/ml) for 7 days. A representative image of a colony observed is viewed under bright field (left panel); with the corresponding image under fluorescence (right panel).

described for pFlox-C2MAZ-Exon vector. However, most of the colonies stably maintaining the pFlox-C2MAZ-Intron vector were green (figure 3.4.), where the transfection efficiency is similar to that of pFlox-C2MAZ-Exon vector. A lack of eGFP expression from the pFlox-C2MAZ-Exon vector indicates that the transcriptional terminating sequence itself is effective and that the difference in the ability to prevent expression is due to the insertion location of SPA C2MAZ.

3.4. Discussion

In this chapter I have investigated the development of an inducible expression system based upon the Cre-loxP recombination system. This was used to generate an inducible episomal expression system from which RNA ablation sequences to be investigated in subsequent chapters could be expressed.

Inducibility relied upon cloning a previously determined transcriptional terminating sequence, SPA C2MAZ (Ashfield et al.1991) into a polyoma ori containing vector between the promoter and the gene to be expressed, eGFP. The floxed sequence was cloned into either the intron or exon of the CAG cassette. Only when positioned within the exon could SPA C2MAZ prevent eGFP expression. The reason for the ineffectiveness of the transcriptional terminating sequence when cloned into the intron is likely to be that the SPA signal is not utilised in this configuration. This notion fits with current ideas about how modules within genes are recognised (Berget et al.1995). Exon definition proposes that splicing of pre-mRNAs occurs by defining exons rather than introns.

Internal exons are defined by biochemical interactions between the splicing machinery recognizing a 3' splice acceptor site and the downstream 5' splice donor site. The first and last exons however are defined slightly differently. In the former case the 7-methyl guanosine cap is recognised instead of a splice acceptor site, in the latter a polyadenylation signal is recognised instead of a splice donor site. The 5' cap and nuclear proteins that bind the cap structure are essential along with the downstream splice donor site for definition of the first exon and removal of the first intron *in vitro* (Izaurralde et al.1994). Definition of the last exon involves recognition of the splice acceptor site of the last intron and the polyadenylation signal. The consequences of this model for the plasmids I have tested are shown in figure 3.5.a) and b), which are diagrams of the proposed splicing events that are presumed to occur when the SPA C2MAZ sequence is cloned into either the intron or exon of the CAG cassette respectively. The diagrams depict two exons and an intron, where the first exon is defined by the 5' cap and the 5' splice donor site and the last exon is defined by the 3' splice acceptor site and a polyadenylation signal in both instances (figure 3.5.a) and b)). However, the source of the polyadenylation signal used depends upon whether the SPA C2MAZ sequence is cloned upstream or downstream of 3' splice acceptor site. According to this model, when cloned into the intron, the polyadenylation signal downstream of eGFP is used as it defines the last exon and eGFP is expressed. However, when cloned into the exon, the polyadenylation signal from the SPA C2MAZ sequence is the first recognised polyadenylation signal so that eGFP is not included in the mature message.

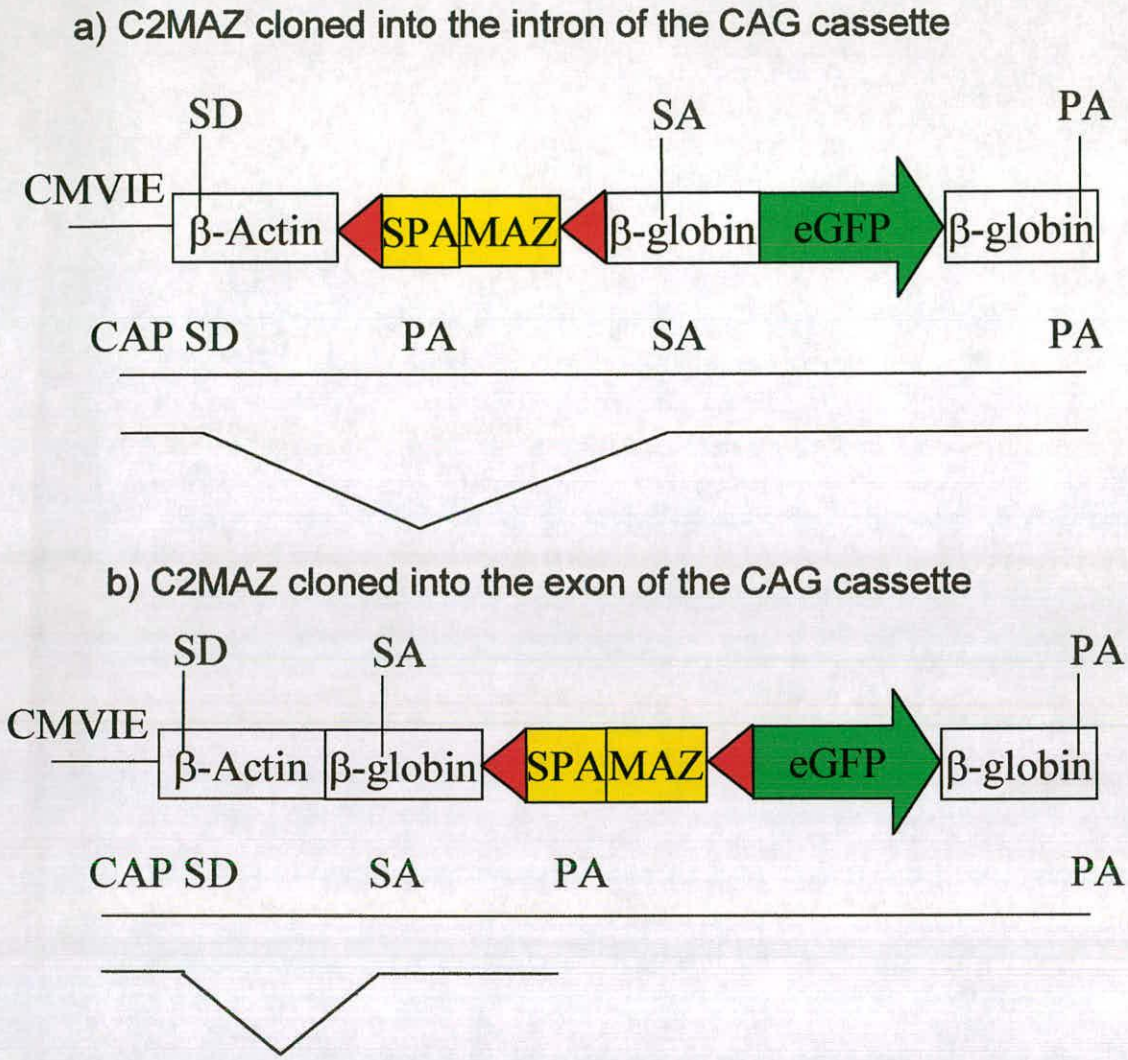


Figure 3.5. Diagram of splicing event and polyadenylation signal used when the C2MAZ transcriptional terminating sequence is cloned within the intron or exon of the CAG cassette.

The CMVIE, β-actin promoter and 5' part of intron I and the first β-globin sequence shown form the CAG cassette. There is a splice donor (SD) site and a splice acceptor (SA) site within the β-actin and β-globin sequences of the cassette respectively. Downstream is eGFP. The C2MAZ sequence is cloned either within the a) intron or the b) exon as defined by the SD and SA site within the promoter. Cloning C2MAZ into the intron results in the sequence being spliced out and polyadenylation (PA) occurring at the PA site downstream of eGFP. However, cloning within the exon results in the PA site present in the C2MAZ sequence being used and causing efficient transcriptional termination upstream of the eGFP sequence.

Evidence in support of the exon definition model and the recognition and definition of the last exon has been the demonstration that polyadenylation is coupled to splicing. *In vitro* splicing and polyadenylation reactions reveal that polyadenylation efficiency is significantly reduced when the terminal splice acceptor is mutated (Cooke et al.1999). A synthetic pre-mRNA substrate containing an adenovirus splicing cassette and the simian virus 40 late polyadenylation signal was used and found that a point mutation in the splice acceptor site, which completely eliminates splicing also reduces polyadenylation by a third compared to wild type substrates (Cooke et al.1999). This study also demonstrated mutating polyadenylation elements such as the hexanucleotide consensus element AAUAAA, which is absolutely required for polyadenylation, also decreased splicing confirming that splicing and polyadenylation are coupled. Requirement of an upstream splice acceptor site for efficient *in vitro* polyadenylation was initially identified by Berget and colleagues. The chimeric precursor RNA used for processing in HeLa nuclear extracts consisted of two exons and an intron from the major late transcription unit of human adenovirus fused to a polyadenylation cassette from the late transcription unit of SV40 (Niwa et al.1990). Mutant precursor RNAs lacking the splice acceptor resulted in less polyadenylated transcript compared to the wild type RNA.

The floxed episome was made inducible by the direct expression of Cre protein from a stably selected vector. This is an adequate system to induce expression. However, the percentage of cells expressing Cre could be increased by generating an ES cell line with an integrated regulatable Cre construct, which would result in more efficient induction. Such regulation is routinely achieved by

fusing Cre to a mutated hormone binding domain of the human estrogen receptor (called Cre-ER), which is then activated by anti-oestrogen drugs such as 4-hydroxy tamoxifen. Originally a ligand dependent recombinase Cre-ER^T was constructed where the glycine residue at position 521 was mutated to an arginine residue making it inducible with 4-hydroxy tamoxifen (Feil et al.1996). However, Cre-ER^{T2} has since been generated, which has a triple mutation, mutated at positions 400 from a glycine to a valine, 543 from a methionine to an alanine and position 544 from a leucine to an alanine (Feil et al.1997), which make the fusion protein ~10-fold more sensitive to 4-hydroxy tamoxifen (Indra et al.1999). Savatier's group (Vallier et al.2001) have demonstrated conditional gene expression using the ligand dependent recombinase Cre-ER^{T2} in ES cells as well as differentiated cells both *in vitro* and *in vivo*. ES cell lines were generated stably expressing Cre-ER^{T2} and an integrated reporter transgene, conditionally expressing human alkaline phosphatase (hAP). Efficient recombination of a loxP flanked cassette containing a selection marker as well as three tandemly repeated transcriptional termination sequences was detected as hAP expression following 24 hr treatment with 4-hydoxy tamoxifen at concentrations as low as 10nM (Vallier et al.2001). A few ES cell clones that displayed strong hAP activation after addition of 4-hydroxy tamoxifen also displayed no or very little background expression prior to induction, making this system highly regulatable for conditional gene expression in ES cells. So future work could be the generation of a 4-hydroxy tamoxifen regulatable Cre-ER^{T2} fusion protein expressing ES cell line.

In conclusion a tightly regulated inducible system for the expression of sequences from episomally maintained vectors has been successfully developed in this chapter.

Chapter 4: Results

Investigation of Antisense RNA as a method to ablate gene function in ES cells

4.1. Introduction

Antisense RNA has potential to inhibit gene function by hybridising specifically to complementary mRNA target sequences and preventing protein expression. The advantage of antisense RNA as an ablation strategy is that it is rapid and relatively more easy than the traditional transgenic technology available in mice to uncover gene function. Recent findings have indicated that antisense RNA can be used to ablate gene function in ES cells (Niwa et al.1998; Ernst et al.1999; Wulf et al.1993). In this chapter I investigate the ability of antisense RNA to produce an ablation phenotype in ES cells.

To assess this three genes *LIFR*, *STAT3* and *Oct4*, whose functions and ablation phenotypes are established in ES cells (see chapter 1 section 1.5.2., 1.5.3. and 1.5.5.) were chosen as antisense targets. As all the genes chosen are required for a self renewal phenotype, ablation induced by antisense should be revealed by the differentiation of ES cells while grown under self renewing conditions. Antisense sequences targeting both the 5' and 3' UTRs were stably expressed from vectors that took advantage of a polyoma virus based expression system that maintains the vector episomally (Gassmann et al.1995). A possible application of a successful antisense RNA strategy would be its use in a library screen for genes

that are important in either ES cell self renewal or differentiation as assessed by function.

4.2. Background

4.2.1. Targeting the UTRs of the antisense target genes

Both 5' and 3' UTRs are to be investigated as effective antisense ablation targets in this thesis. An advantage of targeting UTRs are that they allow for rescue of an ablation phenotype by expression of coding sequence for the targeted gene. This would conclusively demonstrate that an observed phenotype was specific to the gene targeted and not due to cross-hybridisation to a related gene.

Antisense RNAs specific to the UTRs for each target gene investigated are defined as follows. 5' UTR specific antisense begin from the transcription start site and extend to three bases 3' of the ATG translational start site, included to maximise ablation efficiency (see chapter 1 section 1.2.2.4.). The rationale for including a maximum of 6 nts of coding sequence, including the ATG site in the 5' UTR directed antisense is based on evidence that oligonucleotides as short as 7 nts are able to inhibit gene expression (Wagner et al.1996). Wagner et al.(1996) demonstrated selective and potent inhibition of exogenous SV40 large T antigen expression, expressed from a vector co-injected with 7 nt oligonucleotides into African green monkey kidney cells. Inclusion of 7 nts complementary to the coding sequence for 5' UTR specific antisense might therefore not allow the potential rescue of an observed ablation phenotype. Where there are two transcription start sites identified i.e. for Oct4 (Okazawa et al.1991), then the

antisense extends from the first start site, and so encompasses the smaller transcript.

Antisense sequences to the 3'UTR begin after the stop codon and extend to the stretch of A residues in the poly A tail. The poly A tail for both the LIFR and Oct4 mRNA have been identified, as they possess the hexanucleotide polyadenylation signal, AAUAAA. Although the STAT3 sequence cloned by Zhong et al.(1994a) from a mouse thymus cDNA library possesses a stretch of poly As downstream of the stop codon, the hexanucleotide polyadenylation signal is absent.

The LIFR exists in two forms, soluble and transmembrane, of which there are two LIFR transcripts for each, which differ in size. The sizes of the transcripts for the transmembrane form are 11 and 4.4 kb and those for the soluble form are 3 and 2.8 kb. The 5'UTR specific antisense to LIFR is directed to the 11, 4.4 and 3kb transcripts. It does not include the 2.8 kb transcript, which uses an alternative exon 1a promoter, and is expressed predominantly in the liver (Chambers et al.1997). The 3' UTR specific antisense to LIFR is directed to one of the transmembrane forms of the message, the smaller 4.4 kb, as the 3'UTR sequence of the 4.4 kb transcript is also present in the 11 kb 3' UTR. The transmembrane forms of the LIFR are targeted because it is this form of the LIFR that delivers the self renewal signal. The soluble form of the LIFR is unable to deliver a self renewal signal and is therefore not targeted. In fact soluble LIFR has been shown to inhibit the ability of LIF to induce terminal macrophage differentiation of the M1 myeloid leukemic cell line (Layton et al.1992). The possible role that the

soluble LIFR has to play in LIF signalling may be to prevent any systemic effects of locally produced LIF.

Although the lengths of antisense used for each target are different, which may contribute to differences in potency, the constant are that the complete UTR sequences are targeted for the target genes. Exceptions are the 11kb form of the LIFR and the STAT3 3'UTR.

Table 1. Table of sizes of 5' and 3' UTRs, including coding sequence and the predicted size of the mRNA of the target genes

Target genes	Length of 5'UTR (bp)	Length of coding sequence (bp)	Length of 3'UTR (bp)	Size of mRNA (bp) (including 230bp poly A tail)
LIFR	279	3278	575	4362
STAT3	235	2312	472*	4.8 or 5kb**
Oct4	69	974	218	1491

* As determined by Zhong et al.1994 a) b), accession number U06922.

**However, the 3' UTR is longer, as the size of the STAT3 mRNA is about 5kb or 4.8kb, as determined by Northern blot analysis by Zhong et al.(1994a) and Akira et al.(1994) respectively.

4.3. Experimental Strategy

4.3.1. Vector Design

Generation of antisense vectors to 5' and 3' UTR sequences

The vector into which antisense sequences have been cloned is pPyHCAG-S (poly2) vector (Lab plasmid number 403). The construct includes a polyoma origin of replication, a hygromycin resistance gene and a CAG promoter (see chapter 1, section 1.3.1). The vector also contains a stuffer fragment downstream of the promoter with convenient cloning sites.

Two forms of the antisense vectors for each target gene were made, one that constitutively expresses antisense and one in which the antisense sequence is cloned but is not expressed due to the presence of a Stop sequence. The Stop sequence consists of a premature polyadenylation site and termination sequence, cloned upstream of the antisense sequence, downstream of the CAG promoter and was originally characterised in section 3.3.1. (see Appendix 6a). An inducible vector gives the option for more detailed analyses of phenotypic effects, as a self renewing population transfected with inducible antisense vectors can be maintained before transiently expressing Cre and causing differentiation. The inducible antisense expression vectors were constructed by cloning the antisense sequences into the vector containing the Stop sequence, pPyHCAGS-flox C2MAZ-S(poly2). Derivation of the constitutive expressing form was then generated by excising the loxP flanked Stop sequence by an *in vitro* cre recombinase reaction (see Appendix 6b) and 6c). The loxP flanked Stop sequence was cloned into the exon of the CAG promoter of pPyHCAG-S (poly2) which efficiently prevented transcription of downstream sequences as analysed in

chapter 3 section 3.3.1. A unique XhoI site present in the multiple cloning site downstream of the CAG promoter, was cut to produce a linearised vector that was then subsequently blunt ended by a Klenow reaction. The loxP flanked Stop sequence digested with AvrII was also blunt ended by using the Klenow fragment of DNA polymerase and the fragment ligated into the linearised vector. Correctly ligated vectors, with the Stop sequence cloned in a coding orientation were then identified by appropriate restriction digests. See appendix 6a) for the cloning strategy.

5' and 3' UTR sequences were generated as PCR fragments with 5' NotI and 3' MluI restriction sites, for cloning into unique MluI and NotI sites downstream of the CAG promoter and Stop sequence (see Appendix 6 b). The UTR specific PCR fragments are inserted in an antisense orientation with respect to the promoter. PCR primers were designed to amplify the 5' or 3' UTR sequence as defined in section 4.2.1. UTR sequences were amplified from plasmids available in the lab, except STAT3 5'UTR for which the cloned vector sequence was unavailable. Genomic DNA from E14TG2a cells was therefore used as a source to amplify STAT3 5'UTR. The primer sequences used to generate the UTRs of targeted genes with 5' NotI and 3' MluI ends are shown in chapter 2 section 2.1.8.1. PCR reaction conditions were optimised to amplify only the band of interest for the direct cloning into the vector pCR2.1 (TOPO TA Cloning Kit-with pCR2.1-TOPO vector. Invitrogen). Recombinant white colonies were picked for the identification of correctly ligated vectors. Correct vector constructs identified by restriction digests were verified by sequencing. For vectors that showed no mutations of the amplified UTR sequence, a large scale

plasmid preparation was made from which the inserts were removed by NotI/MluI digestion for ligation into pPyH-flox C2MAZ-CAG-S (poly2) (vector number 877). Correctly ligated vectors were identified and used in subsequent supertransfection experiments (see 4.3.2.) along with the constitutive antisense expressing vectors where the Stop sequence was excised by *in vitro* cre recombinase reactions. A strategy diagram of the cloning of 5' and 3' UTR PCR fragments into appropriate vectors is shown in figure 4.1.

4.3.2. Cell Culture

The experimental strategy was to supertransfect into 5 $x10^6$ E14/T ES cells 20 μg of a control vector which does not contain any antisense sequences, pPyH-flox C2 MAZ-CAG-S (poly2) (the parental vector from which the antisense vectors are derived), and the constitutive and non expressing 5' or 3' UTR specific antisense vectors for each target gene. Cells were plated at 5 $x10^4$ and 2 $x10^5$ densities on 9cm diameter gelatin coated tissue culture plates to allow easy identification of the different morphologies of ES cell colonies. Transfectants were selected in hygromycin (400μg/ml) and G418 (200μg/ml) for 7 days, and stained with Leishman's reagent which reveals both undifferentiated and differentiated cells to determine total colony counts. The phenotypic screen when targeting STAT3 and Oct4 was to grow transfected cells in either 0 or 100units/ml of LIF. Phenotypic ablation of these targets was assayed by the differentiation of cells supertransfected with an antisense expressing vector when grown in LIF. For the LIFR ablation experiments, transfected cells were additionally grown in IL6/sIL6R. Again the phenotypic readout of LIFR ablation by LIFR antisense

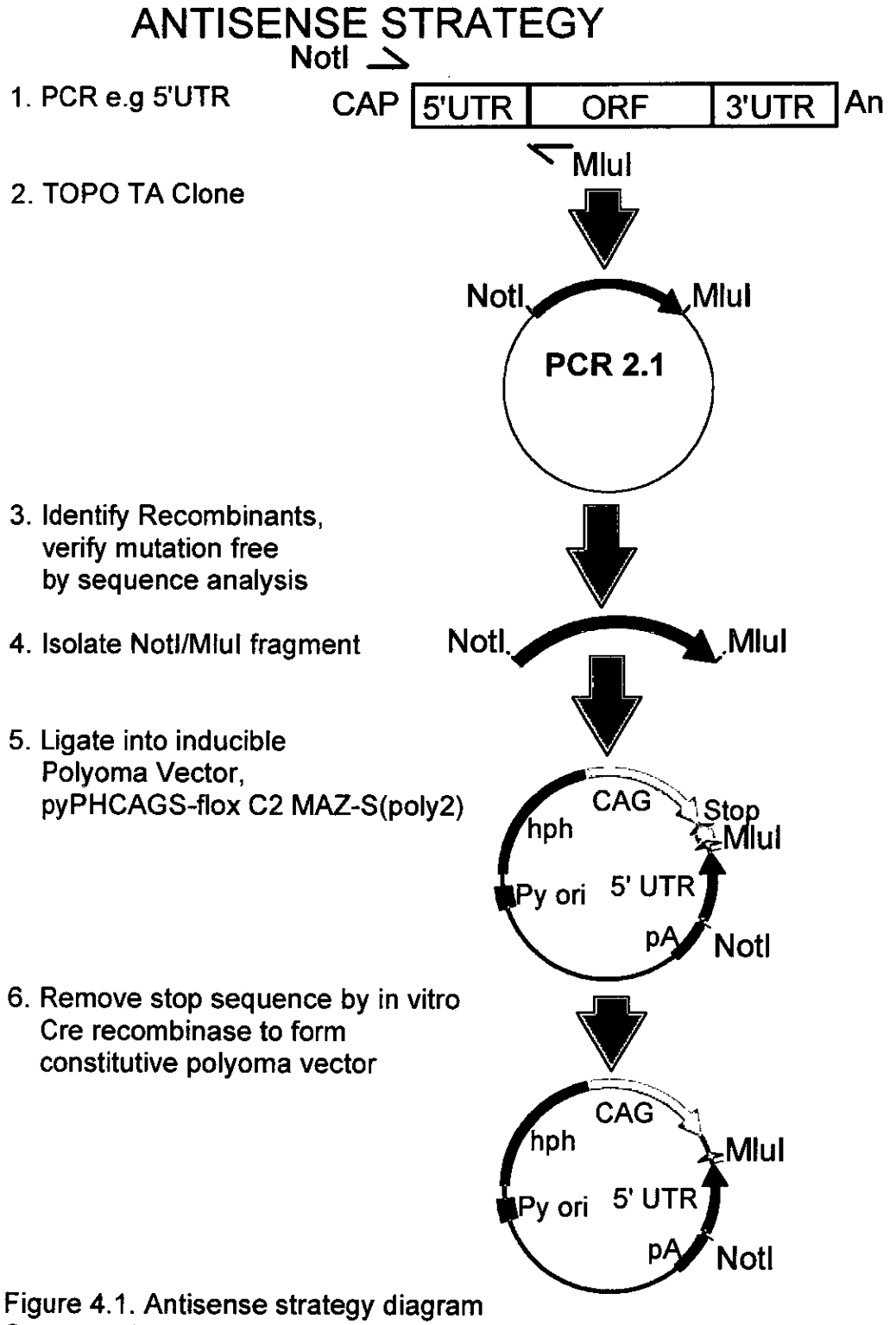


Figure 4.1. Antisense strategy diagram
Strategy diagram to outline the generation of 5' or 3' UTR expressing vectors from initial PCR fragments.

expression vectors will be the differentiation of ES cells grown in LIF. Targeting the LIFR has the advantage that cells can be maintained in an undifferentiated state by growing in a combination of IL6/sIL6R (Yoshida et al. 1994), in which case the signal for stem cell self renewal occurs via the formation of gp130 homodimers i.e. independent of the LIFR.

Additional plates were also set up for each electroporated vector, so that total RNA of the pooled population could be extracted and used subsequently in Northern analyses to determine the level of antisense expression and whether expression of the targeted mRNA is reduced.

4.4. Results

4.4.1. Phenotypic effect of antisense to target genes

To determine the phenotypic effect of UTR specific antisense targeting LIFR, STAT3 or Oct4, stably selected colonies were stained with Leishman's reagent. Colonies were scored based on one of three morphologies; those which were self renewing only; those which had a self renewing core with differentiated extremities; or those that were totally differentiated. These results are shown as a percentage of cells with a purely differentiated morphology, for each of the three vectors, pPyH-flox C2MAZ-CAG-S (poly2) (parental vector, containing no antisense sequences), pPyH-flox C2MAZ-CAG-5'/3' UTR (antisense containing vector, but not expressed) and pPyHCAG-5'/3' UTR (constitutive antisense expressing vector), for each of the cytokine conditions used. The results shown in figure 4.2. are the mean data from two independent transfections.

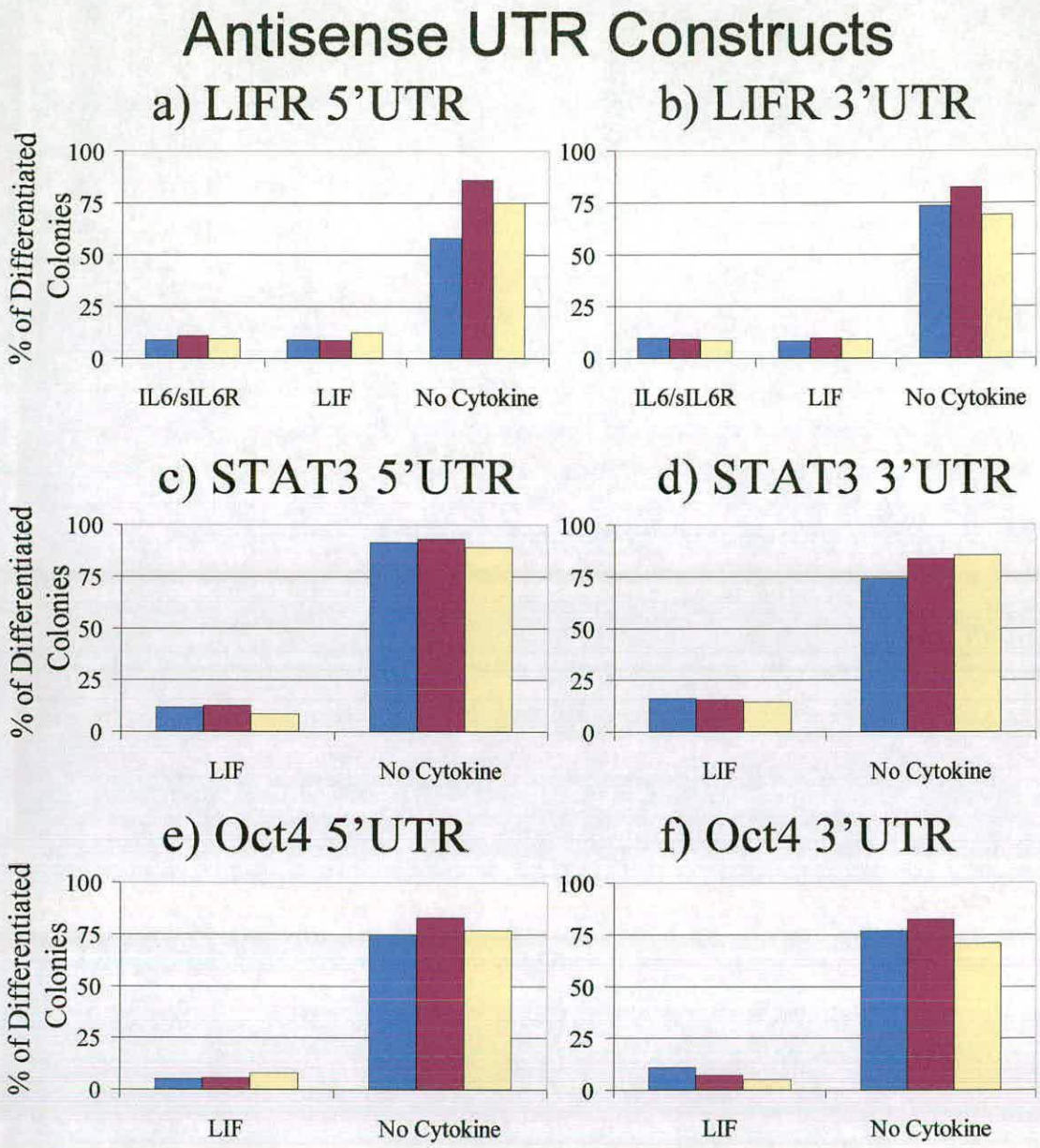


Figure 4.2. A Lack of phenotypic effect with Antisense directed to UTRs of target genes

Effect of Antisense on ES Cell differentiation in the presence or absence of a cytokine, LIF for all targets, and additionally IL6/sIL6R for LIFR targets. E14/T ES Cells were supertransfected with 20μg of DNA of; pPyH-flox C2 MAZ-CAG-S(poly2) (purple bar), pPyH-flox C2 MAZ-CAG-5'/3' UTR (mauve bar) or pPyHCAG 5'/3' UTR (yellow bar). Colonies were fixed and stained with Leishman's reagent after 7 days of selection at 400μg/ml of hygromycin. The number of stem cell colonies and differentiated colonies were scored based on morphology observed with Leishman stain, and the percentage of those cells which had differentiated are shown. Data are the mean for two independent transfections.

Figure 4.2. shows a lack of phenotypic ablation with antisense directed towards the UTRs of all the genes targeted. Growing supertransfected ES cells in the absence of cytokine results in the expected differentiation of cells, where the percentages of fully differentiated colonies is 58% and greater. Supertransfected cells grown in either LIF or IL6/sIL6R showed similar degrees of self renewal irrespective of whether the supertransfecting DNA was empty vector, floxed Stop or the antisense expression vector. This result was confirmed using an appropriate paired t statistical test to analyse the data verifying that there is no significant difference at the 5% level. Therefore there is no evidence of these antisense plasmids inducing an effect of sufficient strength to be visible in this phenotypic screen.

4.4.2. Northern hybridisation analyses to determine transcript levels of endogenous target genes and those from supertransfected vectors

The lack of phenotype may be due to either low or no expression of the antisense transcript. To determine if this was the case, antisense transcript levels were analysed. Total RNA was prepared from pooled populations of cells supertransfected with vectors pPyH-flox C2MAZ-CAG-S (poly2), pPyH-flox C2MAZ-CAG 5'/3' UTR or pPyHCAG-5'/3' UTR for each gene targeted. RNA (10μg) was fractionated on a denaturing agarose gel and probed for 5' or 3' UTR antisense transcripts with a UTR specific probe. Only those cells supertransfected with the constitutively expressing antisense vector i.e. pPyHCAG-5'/3' UTR, produced an antisense transcript of a correctly predicted size (figures 4.3., 4.4. and

4.5.). Transcript sizes are predicted based upon transcription beginning from the β-actin transcription start site present in the CAG promoter and ending at the β-globin polyadenylation site downstream of the cloned antisense fragment. The fact that antisense transcripts are only detected with the constitutive antisense vector demonstrates that the stop sequence present in the pPyH-flox C2 MAZ-CAG-5'/3' UTR vectors is working efficiently and prevents transcription of downstream sequences, as determined in the initial characterisation of the inducible vector (chapter 3).

As antisense RNA is being expressed in cell culture the levels of endogenous mRNA were then examined for Oct4 and STAT3 to see if a corresponding ablation could be detected. The levels of LIFR transcript are too low in ES cells to determine readily by Northern blotting and so were not analysed. The appropriate Northerns were subsequently stripped and probed for either endogenous Oct4 or STAT3 mRNA levels with a coding sequence specific probe. Oct4 mRNA was readily detected as a 1.4kb transcript, however the levels had not decreased with respect to the Oct4 directed antisense (figure 4.5.). Analysis of STAT3 mRNA levels using total RNA yielded no detectable message. However, STAT3 mRNA had been determined previously by Zhong et al.(1994a) and Akira et al.(1994) in mouse liver, spleen and kidney using poly(A)$^+$ RNA derived from these tissues. Therefore 3μg of poly(A)$^+$ RNA from the STAT3 3'UTR ablation experiment was fractionated on a gel and probed for STAT3 mRNA. The STAT3 mRNA was now detectable as seen in figure 4.4. as a transcript of approximately 5kb as observed by Akira et al.(1994) and Zhong et al.(1994a). However, ablation of the message was not detected with respect to

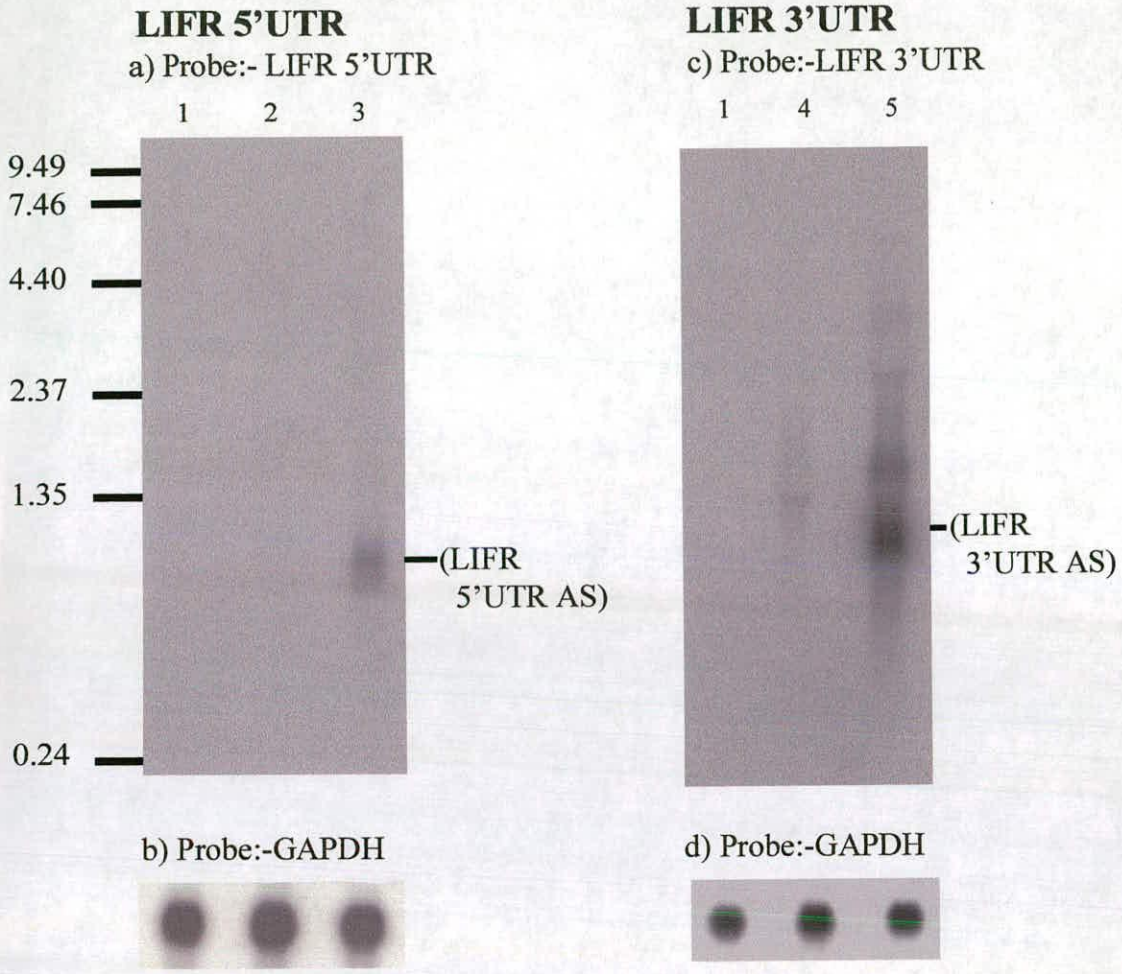


Figure 4.3. Relative antisense transcript levels of LIFR 5' and 3'UTR

E14/T ES cells were electroporated with 20μg of the indicated plasmids:-
1) pPyH-flox C2 MAZ-CAG- S(poly2); 2) pPyH-flox C2 MAZ-CAG-LIFR 5'UTR; 3) pPyHCAG-LIFR 5'UTR; 4) pPyH-flox C2 MAZ-CAG-LIFR 3'UTR and 5) pPyHCAG-LIFR 3'UTR.
Total RNA was prepared after 7 days of selection in 400μg/ml hygromycin, G418 and LIF. The LIFR 5'UTR northern was probed with a probe specific to a) The LIFR 5'UTR and b) GAPDH. The LIFR3'UTR northern was probed with a probe specific to c) The LIFR 3'UTR and d) GAPDH. Brackets indicate the size of the LIFR 5' and 3' antisense transcripts 888bp and 1184bp respectively (inclusive of a polyA tail). Blot a) and c) were exposed for 1 day on a phosphorimager screen; b) and d) for 1 day on autoradiographic film.

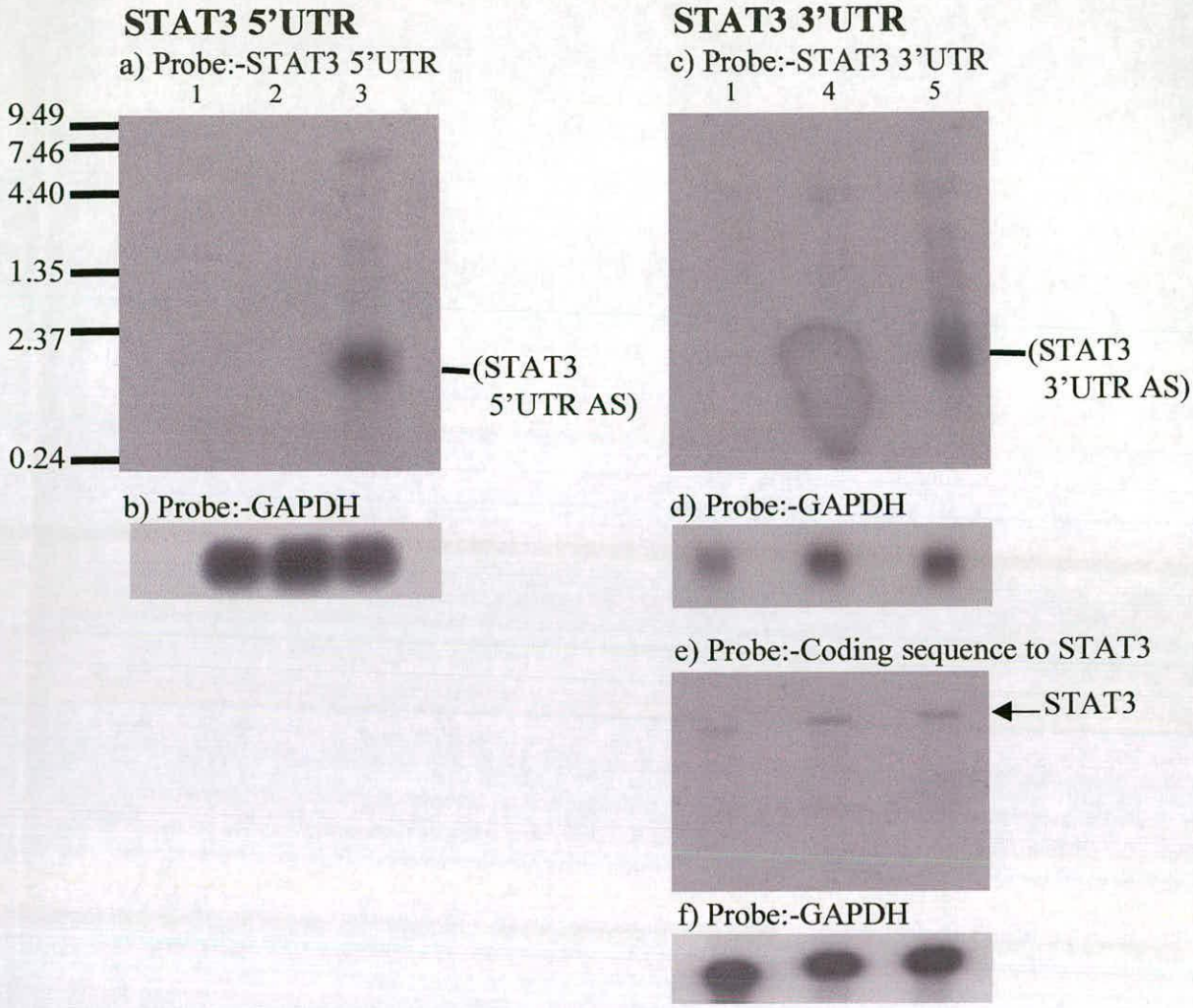


Figure 4.4. Relative levels of endogenous STAT3 transcript compared to UTR specific Antisense transcript.

E14/T ES cells were electroporated with 20μg of the indicated plasmids:- 1) pPyH-flox C2 MAZ-CAG-S(poly2); 2) pPyH-flox C2 MAZ-CAG-STAT3 5'UTR; 3) pPyHCAG-STAT3 5'UTR; 4) pPyH-flox C2 MAZ-CAG-STAT3 3'UTR and 5) pPyHCAG-STAT3 3'UTR.

Total RNA was prepared after 7 days of selection in 400μg/ml hygromycin, G418 and LIF. The STAT3 5'UTR northern was probed with a probe specific to a) The STAT3 5'UTR and b) GAPDH. The STAT3 3'UTR northern was probed with probe specific to c) The STAT3 3'UTR and d) GAPDH.

PolyA RNA was prepared from total RNA from the STAT3 3'UTR antisense ablation experiment and the northern was probed with a probe specific to e) the Coding sequence to STAT3 and f) GAPDH. An arrowhead indicates the size of full length STAT3 mRNA ~4.8-5 kb. Brackets indicate the size of the STAT3 5' and 3' antisense transcripts 873bp and 1081bp respectively (inclusive of a polyA tail). Blots a), c) and e) were exposed for 1 day on a phosphorimager screen; b) and d) for 4 days on autoradiographic film; f) for 1 day on autoradiographic film.

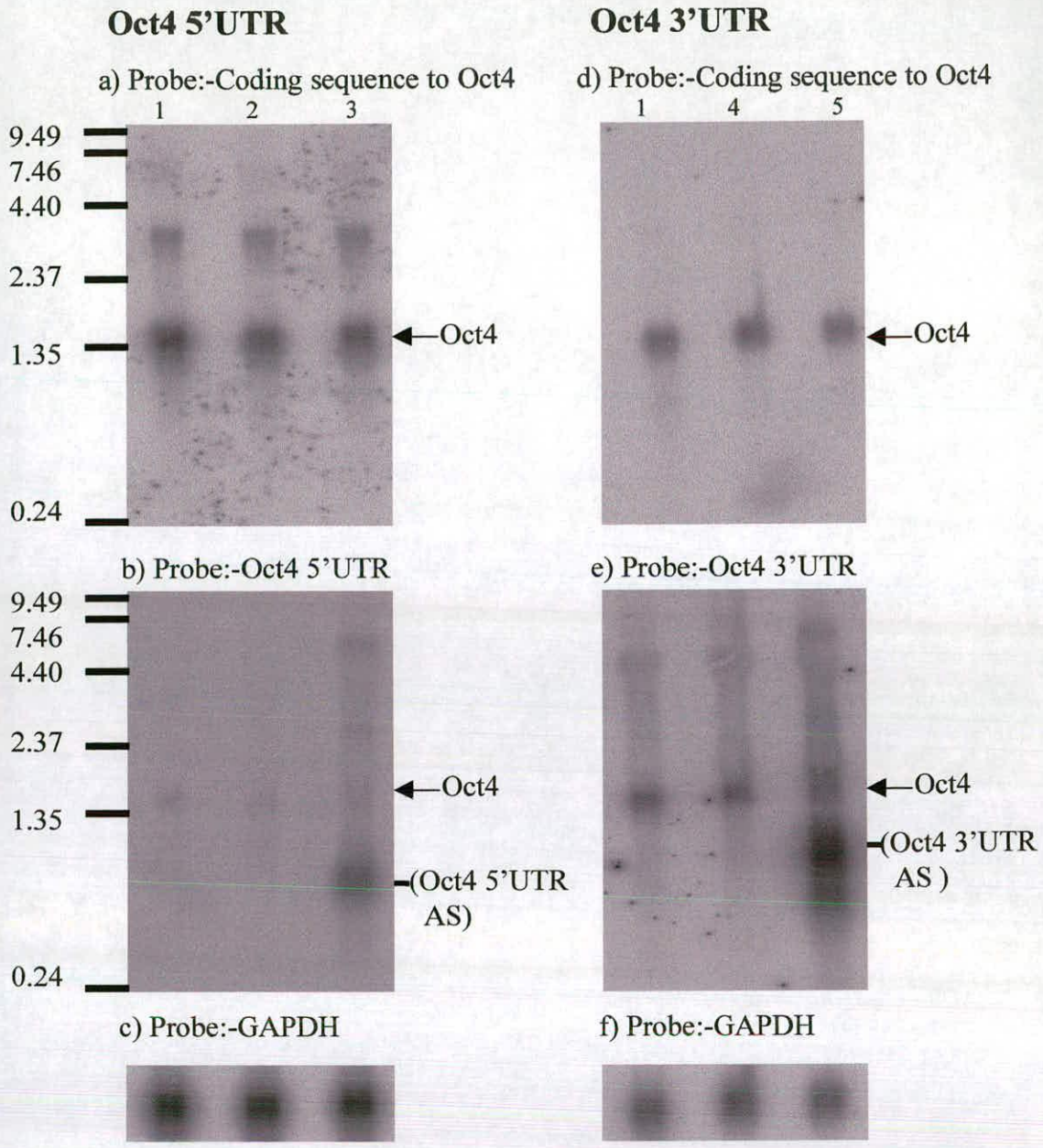


Figure 4.5. Relative levels of endogenous Oct4 transcript compared to UTR specific Antisense transcript.

E14/T ES cells were electroporated with 20μg of the indicated plasmids:-
1) pPyH-flox C2 MAZ-CAG-S(poly2); 2) pPyH-flox C2 MAZ-CAG-Oct4 5'UTR; 3) pPyHCAG-Oct4 5'UTR; 4) pPyH-flox C2 MAZ-CAG-Oct4 3'UTR and 5) pPyHCAG-Oct4 3'UTR.
Total RNA was prepared after 7 days of selection in 400μg/ml hygromycin, G418 and LIF. The Oct4 5'UTR northern was probed with a probe specific to
a) The Coding sequence to Oct4; b) The Oct4 5'UTR; and c) GAPDH.
The Oct4 3'UTR northern was probed with probe specific to d) The Coding sequence to Oct4; e) The Oct4 3'UTR; and f) GAPDH. An arrowhead indicates the size of full length Oct4 mRNA ~1.4kb. While the bracket indicates the size of the Oct4 5' and 3' antisense transcripts 678bp and 827bp respectively (inclusive of a polyA tail). Blots a), b), d) and e) were exposed for 1 day; c) and f) for 2 days on a phosphorimager screen.

expression of STAT3 3'UTR specific antisense. So although the antisense transcripts are being expressed the level of expression may be insufficient to ablate the targeted message. However, both the STAT3 and LIFR antisense transcripts are well in excess of their endogenous targets as the STAT3 mRNA is only detectable as a faint band while the LIFR message is undetectable by Northern hybridisation analysis. The Oct4 antisense transcripts are also in excess of the Oct4 mRNA (Figures 4.5.b) and e)) which are Northern blots probed with either Oct4 5' or 3' UTR specific probes detecting both the specific antisense transcript and endogenous Oct4. Therefore, the lack of a phenotypic effect with UTR specific antisense is not due to the steady state levels of antisense in the cell being below those of the target transcripts.

4.5. Discussion

The data presented here addresses the ability of antisense RNA to ablate gene function of known targets in ES cells and produce an expected ablation phenotype, based upon previous evidence that antisense is able to cause phenotypic ablation in ES cells (Niwa et al.1998; Ernst et al.1999; Wulf et al.1993). My results have however, indicated that there is a lack of an expected phenotype with the use of antisense to both the 5' and 3' UTRs of the target genes LIFR, STAT3 and Oct4. The lack of ablation is not due to a lack of antisense RNA being transcribed. In fact in all cases the levels of antisense RNA are in excess of the target they are directed to ablate. The lack of ablation may be the

result of the choice of sites targeted which may be poor ablation sites either in terms of possible inaccessibility due to mRNA secondary structure or in terms of the UTRs being ineffective target sites.

This study focused upon antisense specific to UTRs, which have been predicted and demonstrated to be good targets for antisense ablation (Pepin et al.1992; Moxham et al.1993). There are however, some examples where UTRs are ineffective targets. For example Munir et al.(1990) were unable to show a reduction in the activity of the HPRT target gene with either a 544 bp or a 1390 bp antisense directed to 5' UTR, first coding exon and intron of which 136 bp or 963 bp is specific to the 5'UTR respectively. This was demonstrated by generating transgenic mice expressing these antisense sequences under the control of the metallothionein I promoter. Not all of the mice generated produced detectable quantities of antisense RNA by Northern analysis, but of those that did none showed a reduction in HPRT activity. This corresponded to a lack of reduction of endogenous HPRT mRNA in all but one transgenic line which expressed the longer antisense construct in the brain and showed a corresponding reduction of the HPRT mRNA in this tissue. However, there are many reports of UTR specific antisense being successful. For example a 50-70% reduction in type II glucocorticoid receptor mRNA in the hypothalmus and cerebral cortex was observed with a 1815bp antisense directed to the 3' UTR of rat type II glucocorticoid receptor cDNA (Pepin et al.1992). Transgenic mice of the type II glucocorticoid receptor were created, made by microinjecting into mouse oocytes the antisense cDNA fragment cloned downstream of the human neurofilament L gene promoter, to restrict expression of the antisense transgene to the brain

(although expression was also seen in the heart, kidney and adrenal tissues). Transgenic mice expressing a 5' UTR and start codon specific antisense RNA to $G\alpha_{i2}$ efficiently reduced $G\alpha_{i2}$ expression in liver and adipose tissues when expressed by the phosphoenolpyruvate carboxykinase promoter (Moxham et al.1993). The antisense used was only 39 bp long and was sufficient to cause almost complete ablation. Therefore UTRs are not necessarily ineffective target sites and whether protein expression can be effectively inhibited by targeting either the 5' or 3' UTR of a message has to be determined for the individual mRNA. The effectiveness of an antisense mediated ablation may however depend upon the length of the antisense RNA used. Although constructs utilised to date have varied in length, the general trend in antisense experiments has been the use of longer antisense constructs ~1kb (see reviews Sokol and Murray 1996; Erickson 1999). The requirement of longer antisense to ablate gene function may be a reason why the antisense RNAs used in my experiments were unsuccessful. The antisense used were relatively short, ranging from 69-575 bp.

The failure of UTR specific antisense to ablate gene function may also be due to the presence of mRNA secondary structures preventing access of antisense RNA to these regions. However, although accessible target sites are required for efficient antisense hybridisation it has been found that some secondary structure is required for heteroduplex formation between oligonucleotides and mRNA (Mir and Southern 1999). Mir and Southern (1999) have shown that efficient duplex formation between short oligonucleotides and RNA requires both single and double stranded regions of the target sequence. They propose that for duplex formation to occur a small number of bases that are accessible in single stranded

regions of the target are required for initial hybridisation between two complementary strands, in a step known as nucleation. Duplex formation is then propagated from this site, which may be facilitated by double stranded helically configured regions. Of the four regions of the tRNA ^{phe} target analysed that produced high yields of heteroduplexes, all contained double stranded stems (Mir and Southern 1999).

Methods of identifying the most suitable sites on the mRNA to be targeted will therefore be highly desirable, as the efficiency of antisense ablation is largely dependent upon these chosen target sites. A couple of approaches have been developed to identify potentially accessible sites. One involves a computer programme to predict secondary structures of antisense RNA molecules (Patzel and Sczakiel 1998). The other approach involves mapping the mRNA transcript for accessible sites with oligonucleotide libraries (Ho et al.1998). For the first approach antisense RNA with favourable secondary structures against an HIV-1 target were predicted by computer. This was based upon the selection of parameters from effective hepatitis B virus directed antisense sequences found previously, i.e. a high number of terminal unpaired nucleotides. Effectiveness of the chosen antisense RNA sequences to inhibit HIV-1 replication in 293T mammalian cells was confirmed, where a reduction of up to 96% of HIV-1 replication was observed. The second approach which is commonly employed to identify effective antisense sequence is the use of random libraries of antisense oligodeoxyribonucleotides (ODNs). An example of this technique was investigated by Ho et al.(1998), to identify sites on their target message, the angiotensin type-1 receptor, that are most accessible to hybridization with

antisense ODNs. Their method involved probing the entire coding sequence of the targeted mRNA with four semirandom libraries composed of three randomized 2'-deoxyribonucleotides in the centre, the fourth position fixed with one of the four deoxyribonucleotides, flanked by three and four randomized 2'-methoxyribonucleotides at either side. Once annealed, accessible sites were determined by cleavage of the RNA by addition of RNase H and then sequencing of the cleaved fragments by primer extension. The active antisense ODNs identified in this way were tested *in vitro* and *in vivo* and found to efficiently inhibit receptor expression. Angiotensin receptor levels were reduced by at least 50% compared to a negative control in cell culture, with a 70-80% maximum reduction seen in some cases. When injected into the brains of rats, ODNs inhibited receptor levels by 65%. These two methods show that there is a link between the accessibility of the target mRNA to antisense RNA and the subsequent ability of the antisense to mediate ablation. However, the situation *in vivo* may be more complex as proteins may occlude potential hybridisation sites, which cannot be accounted for by *in vitro* selection methods. Although such methods can be used to find potential sites for efficient antisense hybridization with regard to the targets used in this thesis, the work is not trivial and ultimately the predicted accessible sites still have to be tested individually. Determining the most efficient sites in this way will therefore be contradictory to the aim of this thesis which was to develop a rapid and efficient ablation method. Had antisense RNA ablation been successfully established in ES cells the ablation strategy could potentially be used to identify genes required for either pluripotency or differentiation. Such screens could be set up by constructing antisense cDNA

libraries, where the source of cDNAs are derived from either self renewing or differentiating ES cells. A potential screen would be for colonies that differentiate in the presence of LIF or those that self renew in the absence of LIF respectively. Library expression from the polyoma based episomal vector, investigated for target gene ablation studies, would have allowed cDNAs of interest to be recovered as episomes from Hirt extracts (see section 1.3.).

It may be more profitable to target coding sequences of the gene to be ablated as longer antisense RNA can be generated, which as described earlier in the discussion may be better at ablation. UTRs may still be effective as antisense targets of a gene, as the number analysed in this study is too small to rule out their potential as a target for antisense ablation. This study has however been unable to ablate the gene function of three targeted genes in ES cells using antisense RNA directed to the UTRs. Therefore it seems unlikely that this could be developed as a general approach for testing gene function in ES cells.

Chapter 5: Results

Investigation of the ability of ribozymes to ablate gene function in ES cells

5.1. Introduction

Ribozymes are an established ablation technology (reviewed in Sokol and Murray 1996). Here the ability of ribozymes to ablate gene function in ES cells was determined using a previously described ribozyme sequence shown to generate a defined phenotype under set conditions in EC cells (Kawasaki et al.1998).

5.1.1. Ribozyme Target - p300

Kawasaki and colleagues (1998) defined a critical role for p300 in mediating retinoic acid (RA) induced differentiation of embryonal carcinoma F9 cells. p300 is a transcriptional coactivator of 300 kD. It is closely related to CBP (CREB-binding protein) with 57% amino acid identity. p300/CBP possess three conserved cysteine/histidine rich regions and a bromodomain interposed between two of them. These domains serve as binding sites for sequence specific transcription factors and other components regulating gene expression. Originally p300 was identified as a consequence of its ability to interact with the adenovirus E1A oncoprotein. Subsequently, a variety of different DNA binding transcription factors and coactivators have been shown to rely on p300 to function *in vivo*. For example members of the Jun family, STAT3 and ligand dependent nuclear

receptors such as retinoic acid, oestrogen and progesterone receptors are but a few examples (reviewed by Shikama et al.1997). p300 is intimately involved in growth control, transformation and cell differentiation. This is evident by the interaction of p300 with the tumor suppressor p53 (Lill et al.1997)and the viral oncoprotein E1A (Arany et al.1995). p53 interacts with the carboxy terminal region of p300 leading to activation of the p53 responsive promoters, thus attributing p300 growth suppressor activities with enhancement of p53 mediated transcription (Lill et al.1997). E1A directly binds to the third cysteine/histidine rich domain of p300 and is important for cell transformation (Smits et al.1996). In addition, p300 also has a histone acetyltransferase activity (Bannister and Kouzarides 1996; Ogryzko et al. 1996), suggesting that p300 may facilitate transcription of p300 responsive genes by regulating the chromatin environment.

Further insights into the developmental importance of p300 have been gained from gene knockout studies (Yao et al.1998). The homozygous p300 knockout is embryonic lethal, with the lethality occurring between E9 and E11.5. The cause of death appears to be multifactorial, as knockout mice have defects in neural tube closure, cell proliferation and cardiac development. p300 heterozygotes were also underrepresented assuming normal Mendelian inheritance. p300/CBP double heterozygotes were embryonic lethal, demonstrating an absolute requirement of p300 and CBP for normal animal development (Yao et al.1998). Moreover mouse development is sensitive to the overall gene dosage of p300 and CBP. As these proteins interact with numerous transcription factors involved in many signalling pathways it could be that a

reduction in the combined level of these proteins is sufficient to result in embryonic lethality.

5.1.2. Rationale of Ribozyme screen

The ability of ribozymes to ablate ES cell gene function was assessed by reproducing an experiment set up by Kawasaki and colleagues (1998) targeting p300 mRNA in EC cells with a p300 specific ribozyme. Their aim was to investigate the role of p300 during retinoic acid induction in EC cells. In their experiment Kawasaki et al.(1998) established stable F9 cell lines expressing a p300 directed ribozyme specific to part of the N-terminal region of p300. This targeted region of the p300 gene had been characterised in previous experiments and shown to be a successful ribozyme target (Kawasaki et al.1996). The ribozyme was expressed from a vector driven by the human $tRNA^{Val}$ promoter, which is suitable for expressing ribozymes because of its small size and high rate of transcription (section 1.2.3.2.). F9 cells were transfected with the ribozyme expression vector using the LipofectAMINE reagent (GIBCO-BRL). The level of p300 mRNA was shown to be reduced compared to wild type F9 cells as determined by competitive RT-PCR. Nuclear hormone receptors including those for retinoic acid (RA) are known to bind p300, so p300 deficient F9 cells were examined to determine if they were still able to respond to RA. However, these cells failed to undergo RA induced differentiation as determined by morphology and a lack of downregulation of SSEA-1, a stem cell specific surface antigen (Solter and Knowles, 1978). In addition synthesis of either collagen IV or laminin B1, which are extracellular matrix proteins, used as markers of differentiation was

not observed. However, cells that contained a CBP specific ribozyme sequence or those that contained the p300 ribozyme and a vector expressing a non-cleavable form of p300 mRNA (i.e. where the GUC cleavage site was mutated to GCC) were able to differentiate. Firstly this demonstrates that only p300 and not CBP is required for RA induced differentiation. Secondly that the p300 ribozyme is specific as the stable introduction of the non cleavable form of p300 into cells expressing p300 ribozyme effectively rescued p300 mRNA expression. p300 is therefore required for RA induced differentiation of F9 cells. The same p300 specific ribozyme sequence was used in a similar RA induced differentiation screen in ES cells to determine the feasibility of ribozyme mediated ablation in ES cells.

5.2. Experimental Strategy

5.2.1. Vector Design

The ribozyme sequence was cloned into a polyoma origin containing vector for episomal expression. The vector from which the p300 ribozyme sequence was cloned was p300 ribozyme in TRS-Neo obtained from Dr. Kawasaki. It contains the $tRNA^{Val}$ promoter driving expression of a p300 ribozyme sequence followed by a polymerase III termination sequence (Kawasaki et al.1996). The ribozyme sequence is complementary to nucleotides 1392-1410 of the p300 message (numbering according to Ecker et al.1994 where ATG is nt 1200) and includes the catalytic loop (figure 5.1.a) and b). This whole fragment was excised as a HindIII/ SalI 166bp fragment and cloned into

Figure 5.1. Sequences of p300 mRNA targeted, specific ribozyme and oligonucleotide probe for ribozyme detection

a) p300 sequence targeted.
Position of target site within p300 mRNA is indicated above the target.

b) Secondary structure of p300 ribozyme with its target.
The arrow defines the site of cleavage. Nucleotides in red denote those to which the oligonucleotide shown in c) are complementary to.

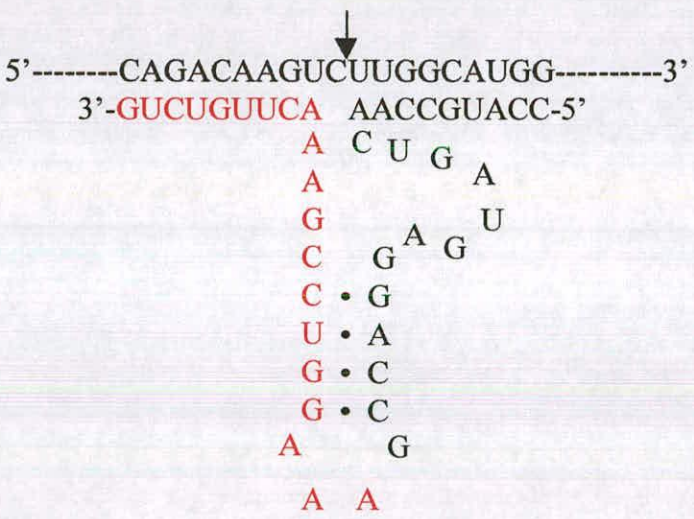
c) Sequence of oligonucleotide probe to detect p300 ribozyme.

a)

1392 1410

Target sequence: . **5'-CAGACAAGUCUUGGCAUGG-3'**

b)



c)

5'-CAGACAAGTTTCGGACCTTT-3'

pPHC2MAZ-CAGeGFP (Lab plasmid number 477) cut with HindIII and SalI to remove the CAG-eGFP and β-globin polyadenylation site of the vector (see Appendix 7.).

The p300 ribozyme expression unit was cloned into this polyoma ori containing vector with a C2 MAZ transcriptional stop sequence downstream of the hygromycin resistance gene. This was to circumvent any possible promoter occlusion of the $tRNA^{Val}$ promoter by the upstream hygromycin phosphotransferase transcription unit. Promoter occlusion has been observed in various eukaryotic systems and is due to inefficient transcriptional termination of an upstream gene, which can interfere with transcription from a downstream promoter (Proudfoot 1986). However, Greger et al.(1998) have demonstrated that insertion of the same transcriptional termination element used here, SPA C2MAZ, consisting of a strong polyadenylation site and the C2 MAZ pause element between two HIV-1 promoters integrated into the genome of HeLa cells, alleviated interference. The p300 ribozyme sequence was verified by sequencing.

5.2.2. Cell Culture

E14/T ES cells were electroporated with either 20μg of pPHC2MAZ-p300R (Lab plasmid number 879) or 20 μg of pPHCAGGSBstXSv40ori, the control vector (Lab plasmid number 470). Cells were selected 30 hrs after supertransfection in 100 units/ml of LIF, 200 μg/ml of G418 and 200 μg/ml of hygromycin for 9 days and both total population and clones of each electroporation were established. Transfected cells were selected in G418 to maintain cells expressing polyoma large T and in hygromycin to select for the

electroporated vectors. A clone maintaining either the ribozyme or the control vectors were then treated with RA. Kawasaki et al. (1998) used 3 x 10^{-7} M of RA to induce differentiation of F9 cells. 1 x 10^{-6} and 1 x 10^{-7} M of RA are typically used to induce differentiation of ES cells (Smith 1991) and as these molarities are similar to that used by Kawasaki et al. (1998) both of these concentrations were used in subsequent RA induced differentiation assays. 1 x 10^3 or 5 x 10^3 cells of the stably maintained clone were plated into wells of a 6-well plate. p300 ribozyme and control vector expressing cells were grown in LIF (100 units/ml), G418 (200 μg/ml), hygromycin (200 μg/ml) and RA at 1 x 10^{-6} or 1 x 10^{-7} M. This was to determine if p300 ribozyme expressing cells are able to maintain a self renewal phenotype in the presence of RA. The same cells were also grown without RA as a control to show that these cell lines maintain a self renewal phenotype in the absence of RA. Cells were plated in triplicate for the same clone and selective media was replaced every two days. After 6 days of selection morphologies were observed by staining cells plated at 1 x 10^3 with alkaline phosphatase, a stem cell marker, and cells plated at 5 x 10^3 with Leishman stain, which stains all cells (see figures 5.2. and 5.3. and 5.4. and 5.5. respectively).

In addition to the phenotypic screen set up to observe p300 ablation, total RNA was also extracted from both the total and clonal population of control and p300 ribozyme expressing cells for Northern analysis. This was so that the expression levels of both p300 ribozyme and endogenous p300 mRNA could be determined.

5.3. Results

5.3.1. RA induced phenotypic screen of p300 ablation

Stable ES cell lines expressing either the p300 directed ribozyme or the control vector without ribozyme sequences were established with a similar transfection efficiency of 0.5%. It was thought that ablation of p300 could prevent ES cell differentiation based on the findings by Kawasaki et al.(1998), who showed that F9 cells expressing the same p300 specific ribozyme, blocked RA induced differentiation. As seen from both alkaline phosphatase staining and Leishman staining of a clonal p300 ribozyme expressing cell line, there is no prevention of RA induced differentiation. Figures 5.2. and 5.3. are colonies stained with alkaline phosphatase. When grown without RA, but in the presence of LIF and hygromycin selection (to maintain the vector) the clonal p300 ribozyme expressing cells maintain a self renewal phenotype. However, when the same cells were additionally grown in the presence of RA (10^{-6} or 10^{-7} M) and LIF, the RA was sufficient to cause the cells to differentiate. This was also true for cells containing the control vector as there were no cells (or very few) that stained positive for alkaline phosphatase. However, there were still cells growing in RA in the same experiment evident by staining with Leishman's reagent (Figures 5.4. and 5.5.). As the controls demonstrate that these cells are able to self renew in the absence of RA, failure of self renewal in the presence of RA may be due to lack of or poor expression of the p300 ribozyme. To determine this and to see if any ablation of p300 mRNA can be detected, RNAs were prepared and analysed by filter hybridisation.

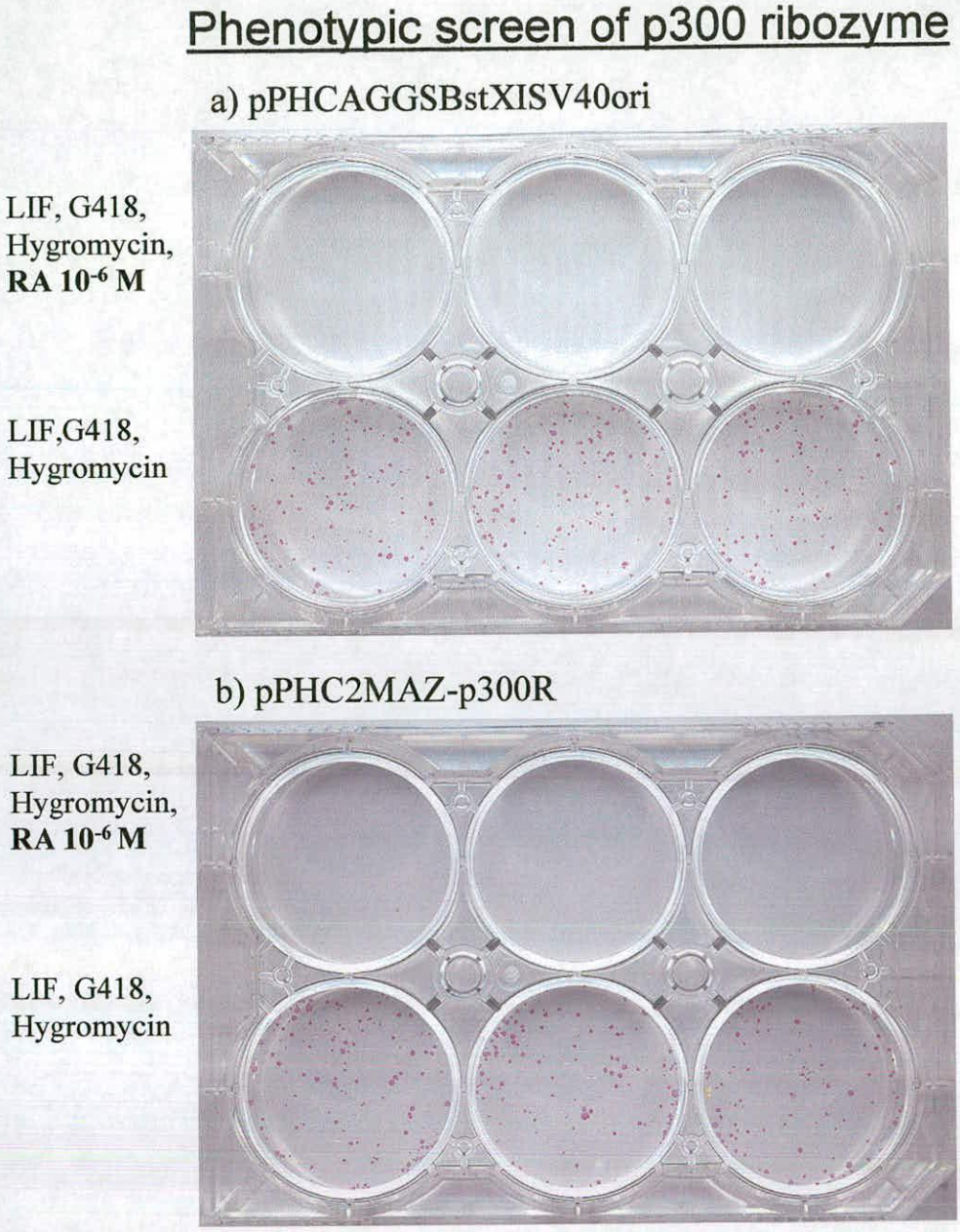


Figure 5.2. Effect of p300 directed ribozyme on ES cell RA induced differentiation.

A stably maintained clone derived from ES cells electroporated with either a control vector a) pPHCAGGSBstXISV40ori; or b) pPHC2MAZ-p300R were plated at 1 x 10^3 density per well of a 6-well plate in triplicate. Cells were grown in LIF (100 units /ml), G418 (200 μg/ml), hygromycin (200 μg/ml) alone or with RA at 10^{-6} M. Cells were grown for 6 days before staining with alkaline phosphatase, a marker of undifferentiated cells.

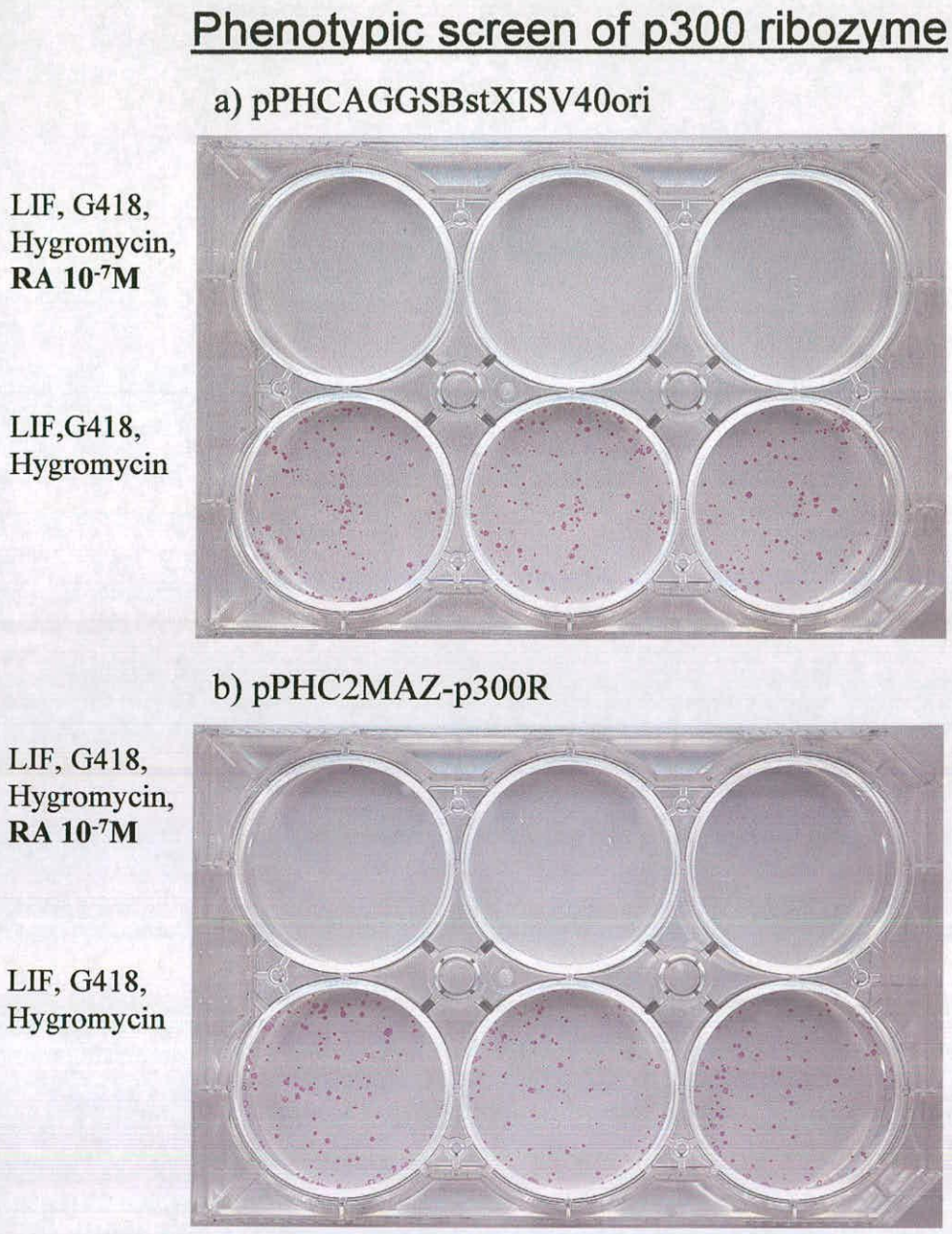


Figure 5.3. Effect of p300 directed ribozyme on ES cell RA induced differentiation.

A stably maintained clone derived from ES cells electroporated with either a control vector a) pPHCAGGSBstXISV40ori; or b) pPHC2MAZ-p300R were plated at 1 x 10^3 density per well of a 6-well plate in triplicate. Cells were grown in LIF (100 units /ml), G418 (200 μg/ml), hygromycin (200 μg/ml) alone or with RA at 10^{-7} M. Cells were grown for 6 days before staining with alkaline phosphatase, a marker of undifferentiated cells.

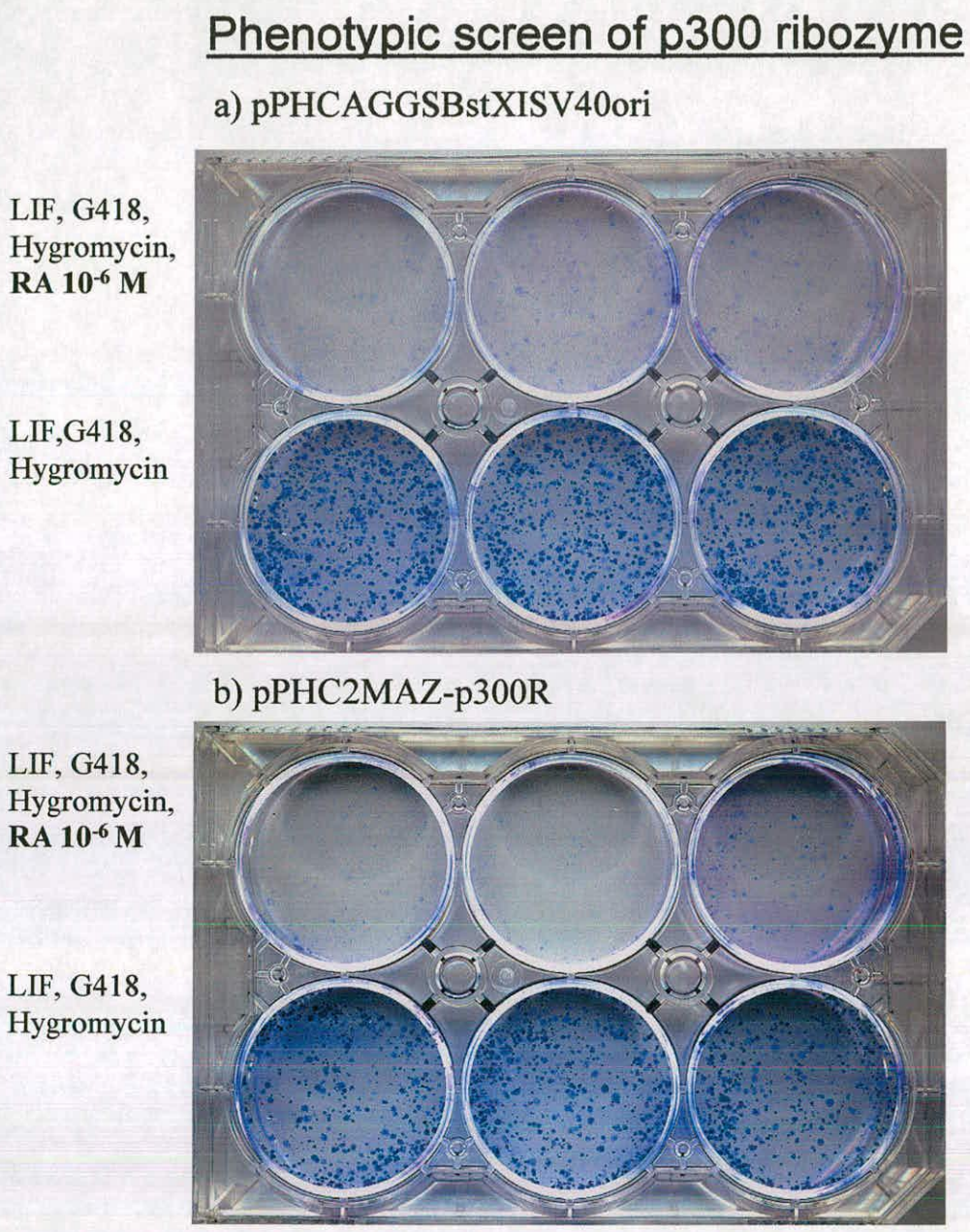


Figure 5.4. Effect of p300 directed ribozyme on ES cell RA induced differentiation.

A stably maintained clone derived from ES cells electroporated with either a control vector a) pPHCAGGSBstXISV40ori; or b) pPHC2MAZ-p300R were plated at 5 x 10^3 density per well of a 6-well plate in triplicate. Cells were grown in LIF (100 units /ml), G418 (200 μg/ml), hygromycin (200 μg/ml) alone or with RA at 10^{-6} M. Cells were grown for 6 days before staining with Leishmans stain, staining all cell morphologies.

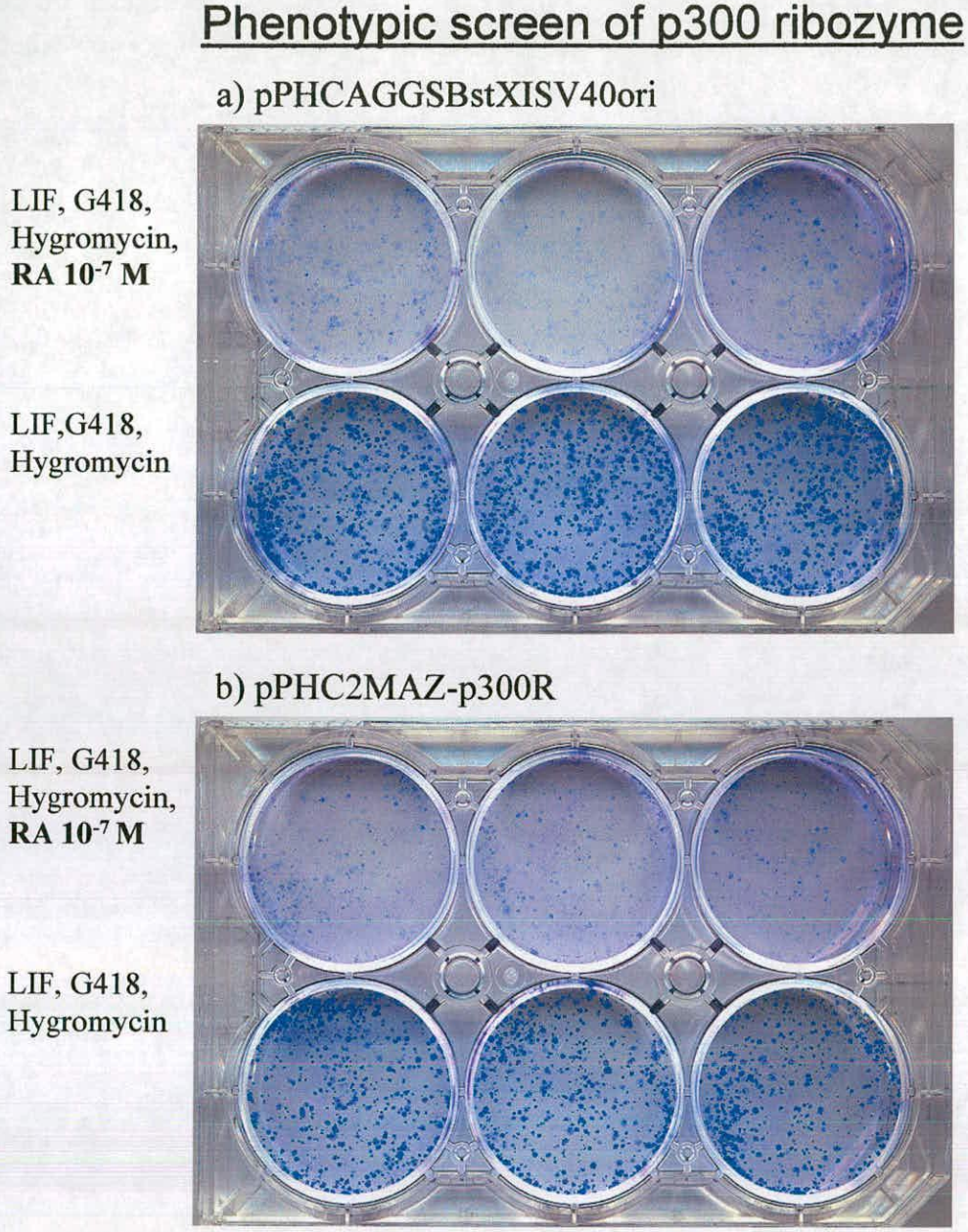


Figure 5.5. Effect of p300 directed ribozyme on ES cell RA induced differentiation.

A stably maintained clone derived from ES cells electroporated with either a control vector a) pPHCAGGSBstXISV40ori; or b) pPHC2MAZ-p300R were plated at 5 x 10^3 density per well of a 6-well plate in triplicate. Cells were grown in LIF (100 units /ml), G418 (200 μg/ml), hygromycin (200 μg/ml) alone or with RA at 10^{-7} M. Cells were grown for 6 days before staining with Leishmans stain, staining all cell morphologies.

5.3.2. Northerns to detect p300 ribozyme transcript levels and endogenous p300 mRNA

Total RNA was extracted from both pooled populations and a clone of ES cells supertransfected with either pPHCAGGSBstXISV40ori or pPHC2MAZ-p300R. RNA from the same p300 ribozyme expressing ES cell clone that was used for the phenotypic screen was analysed. Northern analysis was performed to detect either p300 ribozyme transcript levels or endogenous p300 mRNA. 10 μg of total RNA was loaded onto a 2% formaldehyde gel and probed for p300 ribozyme with an oligonucleotide probe complementary to the 3' flanking arm of the ribozyme and part of the catalytic domain (see figure 5.1. showing the p300 target, ribozyme and oligonucleotide probe sequence specific to the given ribozyme). As seen from figure 5.6.a) ES cells maintaining the ribozyme expression vector, pPHC2MAZ-p300R express the p300 ribozyme transcript, while those cells maintaining the control vector do not. Interestingly the pooled population of ribozyme expressing cells express more ribozyme than the clone of the ribozyme expression vector. This may indicate a variation of expression levels between individual clones. However, RNA from the p300 ribozyme expressing clone was prepared at a later time post transfection than RNA from the population. It is known that the episomal copy number decreases with time (I.Chambers, unpublished) so that a decrease in p300 ribozyme RNA might be expected.

For detection of p300 mRNA levels the Northern was stripped and subsequently re-probed with a 5018 bp NdeI cDNA fragment of the human p300 gene from vector p300HA (lab plasmid number 880). Figure 5.6.b) shows that

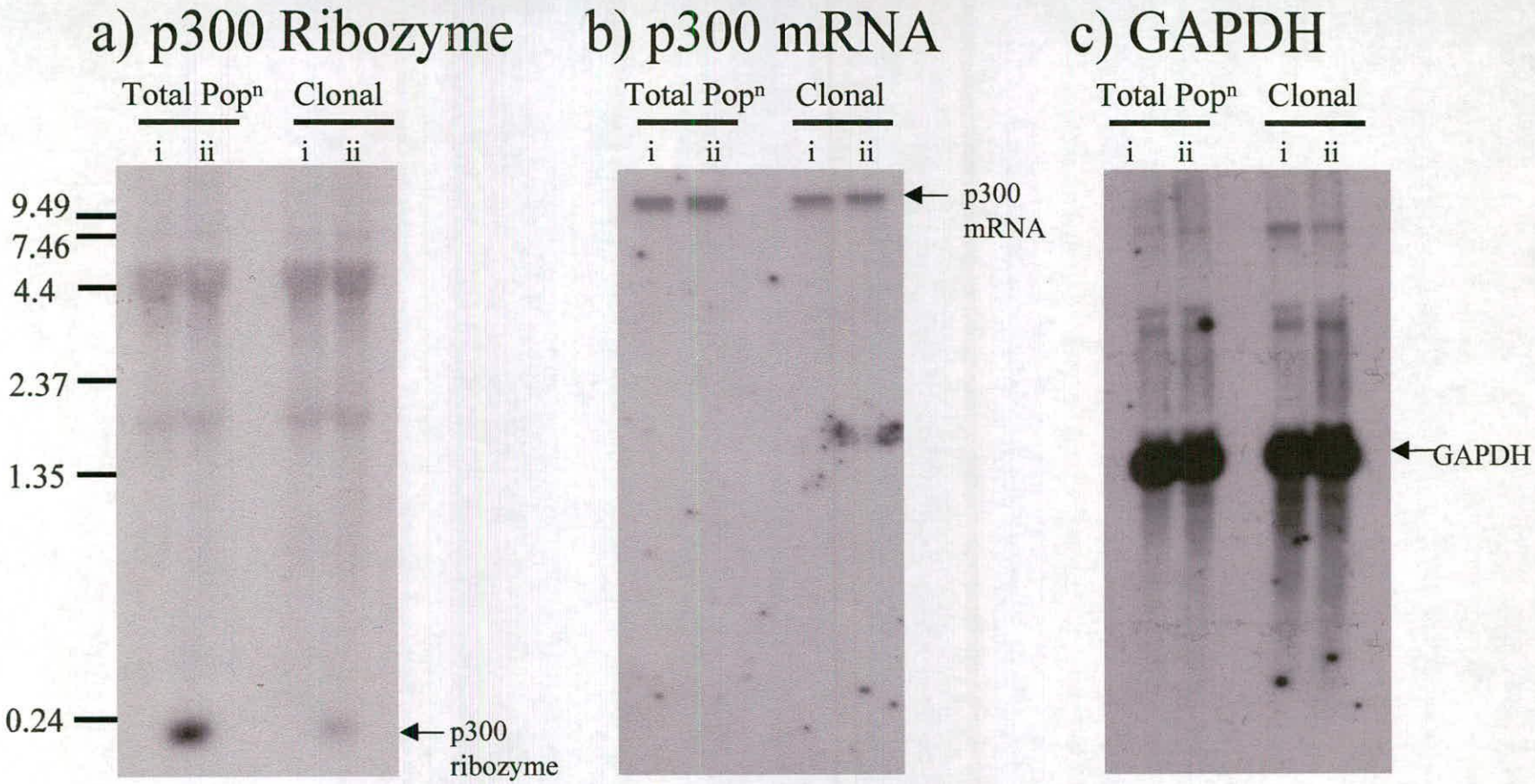


Figure 5.6. Expression of p300 mRNA in comparison to p300 Ribozyme levels.
E14/T ES Cells were electroporated with i) pPHCAGGSBSTXSV40Ori and ii) pPHC2MAZ-p300R and both total and clonal populations were isolated. Total RNA was prepared after 2 weeks of selection and growth in 200μg/ml hygromycin and 100 units/ml of LIF and probed for:- a) p300 Ribozyme ; b) p300 endogeous mRNA ; or c) GAPDH.

there is no obvious decrease in p300 mRNA level for either the total population or the clone of cells expressing p300 ribozyme. Subsequent probing of the same blot with a GAPDH probe shows that loading is almost equivalent in all lanes (figure 5.6.c). These data establish that p300 ribozyme is being expressed in cells selected for this expression vector, however, a decrease of p300 mRNA levels is not observed. Although the clonal p300 ribozyme expressing cell line was used in the phenotypic screen, for which there was less ribozyme expression compared to the pooled population, p300 mRNA ablation was not seen for either of the p300 ribozyme expressing cell lines.

5.4. Discussion

There is both a lack of an ablation phenotype and any mRNA reduction of p300 when targeted with a p300 specific ribozyme. One possible reason for the lack of p300 ablation may be that the phenotypic screen used to measure p300 ablation is inappropriate in ES cells. p300 protein may not be required for RA induced differentiation of ES cells as it is in EC cells. There have been no direct studies of the effect of p300 ablation on the ability of ES cells to differentiate. However, retinoic acid does induce both ES and EC cells to differentiate and into similar cell types. There is evidence that ES cells exposed to RA (10^{-6} M) in the presence of LIF resemble F9 EC cells grown in RA (10^{-6} M) after 5 days as determined by changes in expression of extraembryonic endoderm markers such

as laminin, type IV collagen and tissue plasminogen activator (Mummery et al.1990).

Another reason for the lack of p300 ablation may be that although the ribozyme is being transcribed as seen in the Northern, figure 5.6.a), the levels may be insufficient. In the p300 ribozyme mediated ablation experiments performed by Kawasaki et al.(1998) the level of ribozyme expressed was obviously sufficient to cause phenotypic ablation. However, it is difficult to compare the level of p300 ribozyme expressed in F9 cells versus ES cells, especially as the level of p300 ribozyme produced in F9 cells was not quantitated. It may be useful to repeat the transfections performed by Kawasaki et al.(1998) in F9 cells, so that the level of ribozyme transcription can be directly compared between their experiment and mine. This would help to determine whether a lack of p300 ablation in ES cells is due to insufficient ribozyme expression. The level of p300 ribozyme required to inhibit luciferase activity in HeLa cells co-transfected with a vector expressing a p300 luciferase fusion gene however, has been quantitated (Kawasaki et al.1996). This p300 ribozyme targets a region further upstream from the ribozyme used in the phenotypic screen. It was found that equimolar amounts of this ribozyme and target vector reduced luciferase activity by 50% with almost complete ablation observed with a 4:1 ratio. I have looked at p300 ribozyme expression levels using an oligonucleotide probe and endogenous p300 mRNA using a much longer DNA probe. However, it is not possible to accurately compare the results since the probes are of different specific activities and the hybridisation kinetics of the two probes differ widely. All that can be determined is that this level may not be sufficient to ablate p300 message or produce a predicted ablation phenotype.

Another possible explanation for the lack of ablation may be that although the ribozyme forms properly in EC cells it may not be doing so in ES cells as they possess different cellular environments. There may be additional proteins present in ES cells that bind non-specifically and either prevent correct ribozyme folding or occlude the ribozyme binding site. The secondary structure of the p300 ribozyme was not determined in my experiments, although it could be done by extracting RNA from stably selected cells electroporated with the ribozyme expression vector, pPHC2MAZ-p300R, and the RNA subjected to denaturing and then annealing conditions to allow any secondary structures to form. The folded nature of the RNA could then be analysed on a non denaturing gel by hybridising to a radiolabelled oligonucleotide probe. Direct comparison of migration of the same sample digested with RNase A and RNase T_1 will determine whether the ribozyme possesses any secondary structure. However, this would only demonstrate that the transcribed ribozyme is capable of forming a secondary structure, but not that it had formed an active structure *in vivo*. In addition to the ribozyme potentially misfolding into an inactive structure because of additional non specific proteins binding, the p300 target site could also be occluded by binding of non specific proteins.

Although ablation of p300 mRNA was unsuccessful in terms of a lack of an expected phenotype, the screen and ribozyme used were predefined making this a quick test for the ability of ribozymes to ablate gene function in ES cells. Therefore I conclude that to determine if ribozyme mediated ablation could be a potential strategy to ablate gene function in ES cells more ribozymes will need to be investigated, preferably targeting a gene of known function in ES cells.

Chapter 6: Results

Investigation of the feasibility of RNAi as a method to ablate gene function in ES cells

6.1. Introduction

RNA interference or RNAi as it has been shortened to, is a new and emerging tool to ablate gene function. At the start of my investigation of RNA ablation strategies to silence gene function in ES cells RNAi had been demonstrated to operate in *C.elegans*, *Drosophila* and plants, but not in mammalian cells. In these organisms RNAi was established as being highly specific, functioning by specifically degrading targeted mRNA, promoted by dsRNA. However, the biochemical mechanisms underlying RNAi were still unknown. Subsequently there has been a large amount of work and interest generated in this area due to its potential use as a new and more efficient method of ablation. In this chapter I investigate the feasibility of RNAi ablation in ES cells by either directly transfecting pre-synthesised oligonucleotide 21nt siRNAs or by expressing dsRNA as a hairpin loop from an inverted repeat (IR) containing vector against targeted mRNA.

Transfecting siRNAs directly into ES cells was used to establish whether RNAi occurred in ES cells. RNAi ablation was then investigated by stable expression of dsRNA taking advantage of the episomal expression system. As stated previously transfection efficiencies for the episomal vector system are high with stable and long lasting expression of cloned cDNA sequences. This allows

the investigation of genes that are required for differentiation, where ablation effects have to be propagated for longer. dsRNAs were stably expressed from episomally maintained vectors by cloning the mRNA sequence to be targeted as an inverted repeat. This method of introducing dsRNA was originally developed by Tavernarakis et al.(2000) in *C.elegans* and subsequently shown to work in *Drosophila* (Kennerdell and Carthew 2000; Piccin et al.2001). Usually a spacer (consisting of an unrelated sequence) is included between the two cloned inverted repeat sequences.

The ability of RNAi to cause specific gene ablation in ES cells was assessed by targeting the coding sequences of *LIFR*, *eGFP* and *SOCS3*. LIFR was chosen because it is an endogenous target for which there is a clear ablation phenotype, the differentiation of ES cells in the presence of LIF (I.Chambers-Personal Communication). In addition LIFR ablation experiments possess an internal control, as self renewal can be maintained independently of the LIFR by signalling via IL6/sIL6. eGFP was chosen, because as a visible target its ablation is readily detectable by eye and can be quantitated by flow cytometry. SOCS3 was chosen because it is not continuously expressed and can be induced by stimulation of gp130, after 6-7 hrs of LIF withdrawal, with peak mRNA levels occurring approximately 45 minutes post-stimulation. Coding sequences of these genes are targeted rather than the UTRs, which were targeted previously in the antisense experiments. The reasons for not using the UTRs is that that there is evidence that for RNAi experiments, longer dsRNA are disproportionally able to ablate gene function, at least in *Drosophila* (Yang et al.2000). Injections of dsRNAs varying in size from 80bp to 928 bp resulted in a linear relationship between dsRNA

length and the log value of inhibition for dsRNAs in a single embryo assay expressing a luciferase target from an injected plasmid. The regions targeted are the entire coding sequence of eGFP, the 5'region of SOCS3 (471 bp out of 678bp) and the 5' region of LIFR (1071 bp out of 3279 bp). All include the ATG translational start site.

Determining whether RNAi is a feasible technology to ablate gene function in ES cells is important because it has been demonstrated to possess great potential as a quick yet reliable tool to determine gene function in organisms such as *C.elegans* and *Drosophila*. If RNAi can be developed as such a tool in ES cells it could be used to study the role of genes in either self renewal or differentiation to gain a better understanding of these processes.

6.2. Background

6.2.1. Expression of dsRNA as a hairpin loop – inclusion of a spacer sequence

As already described part of the investigation of this chapter will be to study the potency of RNAi ablation from continuous *in vitro* expression of dsRNA hairpin loops via transfection of inverted repeat containing vectors. An unrelated spacer sequence can be cloned between the IRs, the advantages for which are three fold. Firstly the sequence can be that of a selectable marker, thus allowing easy isolation of correctly ligated vectors at the cloning stage. Secondly the sequence also acts as a spacer between the IRs which should facilitate the transcript forming a stable hairpin loop structure, where this sequence forms the

loop. The third and most important reason for inclusion of a spacer sequence is that it allows replication of the IR in *E.coli*. This is because it has been well documented that long DNA palindromes are unstable in *E.coli* (Leach 1996), where there is an upper limit of ~150 bps total length for a palindromic DNA sequence that can be cloned in wild type *E.coli* (Warren and Green 1985). There are one of two fates associated with trying to clone IRs into *E.coli*; either clones containing the palindrome are not recovered or a low frequency of clones with either complete or partial deletion are propagated to circumvent palindromic inviability (Hagan and Warren 1983). These effects are the direct result of hairpin structures forming, as these structures have been shown to stop the progression of the replication fork *in vitro* (LaDuca et al.1983). However, both inviability and instability of a palindrome can be reduced with the inclusion of a central spacer between the repeats (Warren and Green 1985). A possible way in which all or part of the palindromic sequence is removed is explained by strand slippage. This is where palindromes are fully or partially deleted by recombination between complementary sequences found either downstream or within the palindrome respectively (Leach 1996). Recombination can occur as a result of the stalling of DNA replication by the hairpin secondary structure.

Tavernarakis et al.(2000) performed RNAi ablation experiments in *C.elegans* using IRs without a spacer. The cloning efficiency was found to be low in the *E.coli*, DH5α strain, only a few per hundred screened, but a 1 in 20 efficiency was observed when the *E.coli* SURE strain (Stratagene) was transformed. Kennerdell and Carthew (2000) included a 5bp spacer sequence, which was sufficient to allow the stable maintenance the IR vectors even in *E.coli*,

DH5α. Piccin et al.(2001) included a much larger unrelated DNA sequence, a 330 bp spacer from part of the coding region of the green fluorescent protein. They noted that inclusion of the spacer greatly increased cloning efficiency. The spacer at 330 bps was about one third of one repeat length implying that relatively long spacers can be used without interfering with the ability of dsRNA hairpins to mediate RNAi. The zeocin resistance cassette to be used as the spacer in my subsequent RNAi experiments is 568 bp which ranges from being about the size of one repeat length to about four times its length. This is larger than the spacer used by Piccin et al.(2001) and may interfere with the correct formation of the dsRNA hairpin structure. However, there have not been any experiments that have determined if the size of spacer used affects RNA interference. After the completion of the work presented in this thesis, Paddison et al.(2002) used a similar GFP hairpin plasmid specific to the first 500 coding base pairs with a zeocin resistance gene between, and were able to show down regulation of GFP expression in P19 cells. This demonstrates that long spacer sequences of approximately equal length to the hybridising arms of a hairpin can be included between inverted repeats to form effective dsRNA hairpin structures.

6.3. Experimental strategy

6.3.1. Vector design

The vector into which IR sequences were cloned was pPyCAGIP (Lab plasmid number 564). It contains a polyoma origin of replication and the CAG promoter upstream of a stuffer fragment with unique BstXI restriction sites for

cloning of desired sequences, followed by an IRES puromycin resistance cassette. IRES linked selection was chosen for the direct selection of gene expression to give higher levels of expression compared to selection from a separate drug resistance marker.

The spacer sequence cloned between the IRs was a zeocin selection cassette with a bacterial EM7 promoter driving its expression to facilitate isolation of correctly ligated IR vector constructs. The cloning strategy involved a four way ligation of the vector backbone, the target sequence which forms the IR, and the zeocin selection cassette. This was achieved by cutting the cloning vector, pPyCAGIP with BstXI and the eGFP and LIFR target sequence with BstXI/EcoRV or BstXI/SmaI for SOCS3 and the zeocin selection cassette with PvuII/EcoRV. Constructs where the zeocin cassette had ligated in either a sense or antisense orientation with respect to the promoter were isolated and both were used in ES cell supertransfections (See appendix 8 for the cloning strategy).

6.4. Results

6.4.1. Establishing RNAi in ES cells

- Ablation of exogenous luciferase target in ES cells using 21nt siRNAs

To establish RNAi ablation in ES cells 21 nt siRNA duplexes were used to suppress luciferase expression from transfected luciferase reporter plasmids. RNAi using 21 nt siRNAs by Elbashir et al.(2001b) was first demonstrated in cultured mammalian cells including COS-7, NIH/3T3 mouse fibroblasts, HeLa S3

and 293 human embryonic kidney cells. The same assay used by Elbashir et al.(2001b) was used to establish whether RNAi operated in ES cells. Two ES cell lines E14/T cells which have the polyoma large T protein expressed from a randomly integrated transgene and E14TG2a were both investigated. COS-7 cells were included as a control to compare the efficiency of ablation and confirm reproducibility.

The targeted reporter genes were those coding for sea pansy (*Renilla reniformis*, RL) and two sequence variants of firefly (*Photinus pyralis*, GL2 and GL3) luciferases. The luciferase reporter genes were expressed from plasmids pGL2-Control, pGL3-Control and pRL-TK (Promega). Elbashir et al.(2001b) synthesized five siRNA duplexes targeting GL2, GL3 and RL luciferase as well as a control inverted GL2 duplex and a GL2 duplex with ribo-uridine 3' overhangs, instead of the 2'-deoxythymidine 3' overhangs present in all other siRNA duplexes. Thymidine overhangs were chosen because they reduce costs of synthesizing the duplexes and may increase nuclease resistance.

In my experiments I chose to make duplexes to GL2, GL3 and RL luciferase and to target only pGL3-Control and pRL-TK vectors (Promega). The regions of the luciferase gene reporters targeted were those used by Elbashir et al.(2001b) and the method of annealing the oligos was the same. Cells were co-transfected with the luciferase reporter plasmids and the siRNAs using Lipofectamine 2000 reagent (Life Technologies), and luciferase expression analysed 20hrs after with the Dual luciferase assay (Promega). The assay was carried out in three separate experiments for each ES cell line in conjunction with COS cells and the results are shown in figure 6.1.

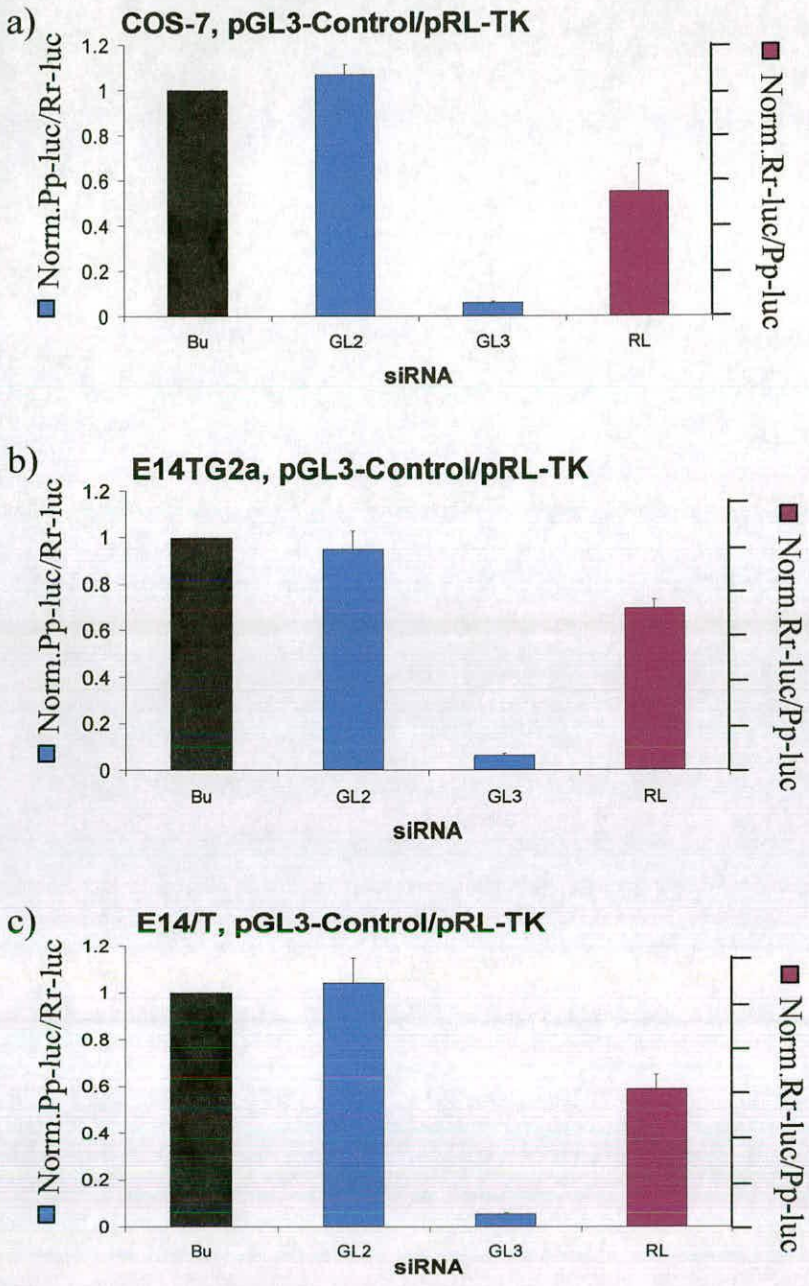


Figure 6.1. RNA Interference by siRNA duplexes.
Ratios of target to control luciferase were normalized to a buffer control (Bu, black bars); Purple bars show ratios of Photinus pyralis (Pp-luc) GL3 luciferase to Renilla reniformis (Rr-luc) luciferase (left axis), mauve bars show RL to GL3 ratios (right axis). The cell lines used for each experiment are designated as a) COS-7 cells; b) E14TG2a; c) E14/T ES cells. The data were averaged from three independent experiments + s.d.

In COS cells specific inhibition of luciferases was observed and to a similar level as demonstrated by Elbashir et al.(2001b). Transfection of GL3 specific siRNA alone reduced *Photinus pyralis* luciferase (Pp-luc) expression from the pGL3-Control plasmid, whilst GL2 specific siRNA is without effect, giving similar levels of luciferase expression as the buffer control (that is pGL3-Control and pRL-TK transfected with water alone). This result was confirmed using an appropriate statistical test, a single analysis of variance test to analyse the data verifying that the result is significant at the 0.1% level. The ratio of Pp-luc when normalised to the transfection control *Renilla reniformis* luciferase (Rr-luc) is below 0.2 (where 1 unit is arbitrarily given as the maximum value when buffer alone is transfected). This value is similar to that seen by Elbashir et al.(2001b) for COS cells. The ratio of Rr-luc reduction normalised to Pp-luc as a transfection control was not as great however, 0.5 compared to a 0.4 ratio observed by Elbashir et al.(2001b).

In both ES cell lines investigated the levels of Pp-luc are also reduced to the same amount as observed with COS cells. However, Rr-luc reduction was not as great in ES cells where the ratio of luciferase expression were 0.6 to 0.7 compared to 0.5 seen with COS cells. These results indicate that RNAi does operate in ES cells.

6.4.2. LIFR as a target for RNAi Ablation

6.4.2.1. Strategy for LIFR ablation experiments - A functional readout

RNAi has been established in ES cells using 21 nt siRNAs. However, the usefulness of this approach to investigate gene function phenotypically is limited as ablation effects are not long lasting. The focus from this point was therefore the investigation of a stable expression system for continuous expression of dsRNA as a double stranded hairpin loop.

LIFR was chosen as a target for ablation because it is an endogenous target for which ablation can be screened phenotypically, as its requirement for ES cell self renewal is well characterised. Ablation of LIFR can also be internally controlled for as the self renewal phenotype can be maintained by growing ES cells in IL6 and sIL6R, which signals self renewal independently of LIF and the LIFR. Also levels of LIFR mRNA are very low, detectable by RNAse protection which may increase the likelihood of LIFR ablation. The strategy was to ablate gene function by continuous *in vitro* expression of a LIFR specific dsRNA hairpin from an episomally maintained LIFR inverted repeat containing vector. The 5' region of the coding sequence of the LIFR including the ATG site, 1162 bp in total was targeted.

E14/T ES cells were supertransfected with pPyCAGIP, the control parental vector and two forms of the LIFR inverted repeat containing vectors, pPyCAG-S LIFR IRzeo-IP and pPyCAG-AS LIFR IRzeo-IP, and grown in LIF or IL6/sIL6R. The prefix S or AS before the gene name denotes either a sense or antisense orientation of the zeocin cassette with respect to the direction of the CAG promoter. The experiment was plated at $1x10^5$ and $5x10^5$ on 9cm diameter

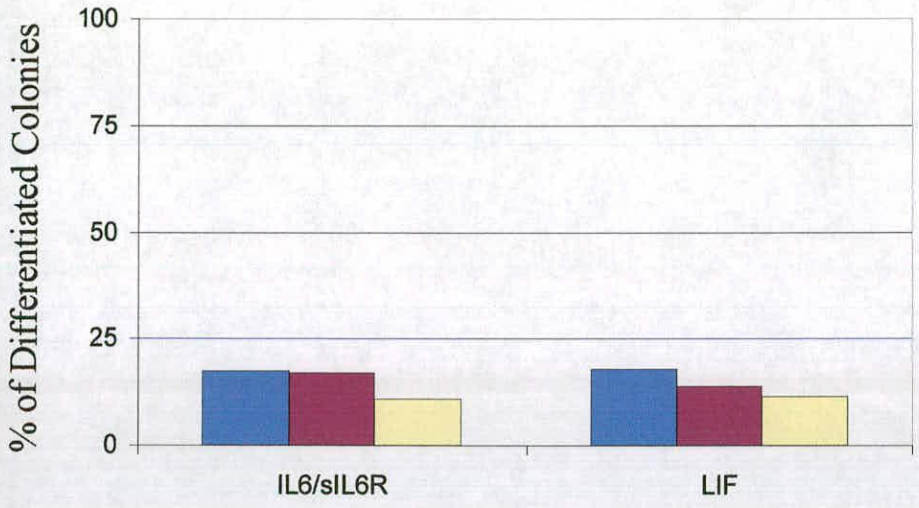


Figure 6.2. A Lack of phenotypic effect with LIFR specific dsRNA
Effect of *in vitro* transcribed LIFR specific dsRNA upon ES cell differentiation grown in IL6/sIL6R or LIF. E14/T ES cells were supertransfected with 20μg of DNA of; pPyCAGIP (blue bar), pPyCAG-S LIFR IRzeo-IP (mauve bar) or pPyCAG-AS LIFR IRzeo-IP (yellow bar). Colonies were fixed and stained with Leishman's reagent after 8 days of selection at 1.4μg/ml of puromycin. The number of differentiated colonies were scored based on morphology observed with Leishman stain, and are shown as a percentage of the total number of colonies scored.

plates and selected in puromycin (1.4μg/ml) for 8 days before staining with Leishmans stain. At 1.4μg/ml puromycin concentration control untransfected E14/T cells were killed. The number of stem cell colonies propagated in either cytokine, LIF or IL6/sIL-6R were similar at this concentration of puromycin for LIFR IR transfected cells. However, there was no increase observed in the numberof differentiated cells when grown in LIF compared to those grown in IL6/sIL6R (Figure 6.2.). This result was confirmed using an appropriate paired t statistical test to analyse the data verifying that there is no significant difference at the 5% level. Therefore this particular LIFR dsRNA targeting the 5' region of the LIFR coding sequence was unable to produce a specific ablation phenotype.

6.4.3. eGFP as a Target for RNAi Ablation

6.4.3.1. FACS Analysis Data

eGFP was chosen as a target because it has been previously ablated from a transgene in both *C.elegans* in the adult (Parrish et al.2000) and in the mouse at the preimplantation stage (Wianny and Zernicka-Goetz, 2000). Fluorescence also provides a sensitive and accurate means of quantitating eGFP ablation. Ablation of eGFP was examined by electroporating eGFP inverted repeat expression vectors into a constitutively expressing eGFP ES cell line, GFPzeo-E14/T. The cell line was made by Dr. Ian Chambers with technical assistance from Morag Robertson and Douglas Colby. The construct pCAGGsEGFPIZ was randomly integrated into E14/T cells. The clone was shown to express eGFP as well as remaining supertransfectable. The GFPzeo-E14/T cell line was selected in

20μg/ml of zeocin to maintain high eGFP expression before electroporation with inverted repeat containing vectors at which point zeocin selection was removed.

eGFP inverted repeat vectors cloned as pPyCAG-S eGFP IRzeo-IP or pPyCAG-AS eGFP IRzeo-IP as well as the parental control vector, pPyCAGIP were electroporated individually into GFP-zeo-E14/T ES cells. The cells were then selected in 3μg/ml of puromycin for 3 days. At this time cells had been grown for 5 days post electroporation and were analysed by FACS. This was to determine whether those cells that expressed the eGFP inverted repeat reduced eGFP expression from the random CAG-eGFP IRES zeocin transgene. Figure 6.3. shows the results for this FACS analysis. A population of cells with reduced fluorescence (indicated as M2), of the order of two magnitudes lower than the level of eGFP expression observed normally is seen. This population is confined to only those cells transfected with eGFP IR vectors. The percentages of the cells expressing reduced eGFP is 33.3% for pPyCAG-S eGFP IRzeo-IP and 39.6% for pPyCAGIP-AS eGFP IRzeo-IP vector (percentages are based upon the total number of gated cells). The percentage of cells expressing less eGFP for pPyCAGIP vector is 6.7%, which is comparable to the background levels observed for the untransfected GFPzeo-E14/T cell line control.

These initial results suggest that RNAi may be induced in ES cells by stable expression of a long hairpin dsRNA against an integrated eGFP transgene. A separate eGFP ablation experiment was therefore carried out to confirm reproducibility. An additional control was included in the form of LIFR inverted repeat vectors. The concentration of puromycin used for selection was 1.5μg/ml, as the transfection efficiency was very low for the eGFP inverted repeat vectors

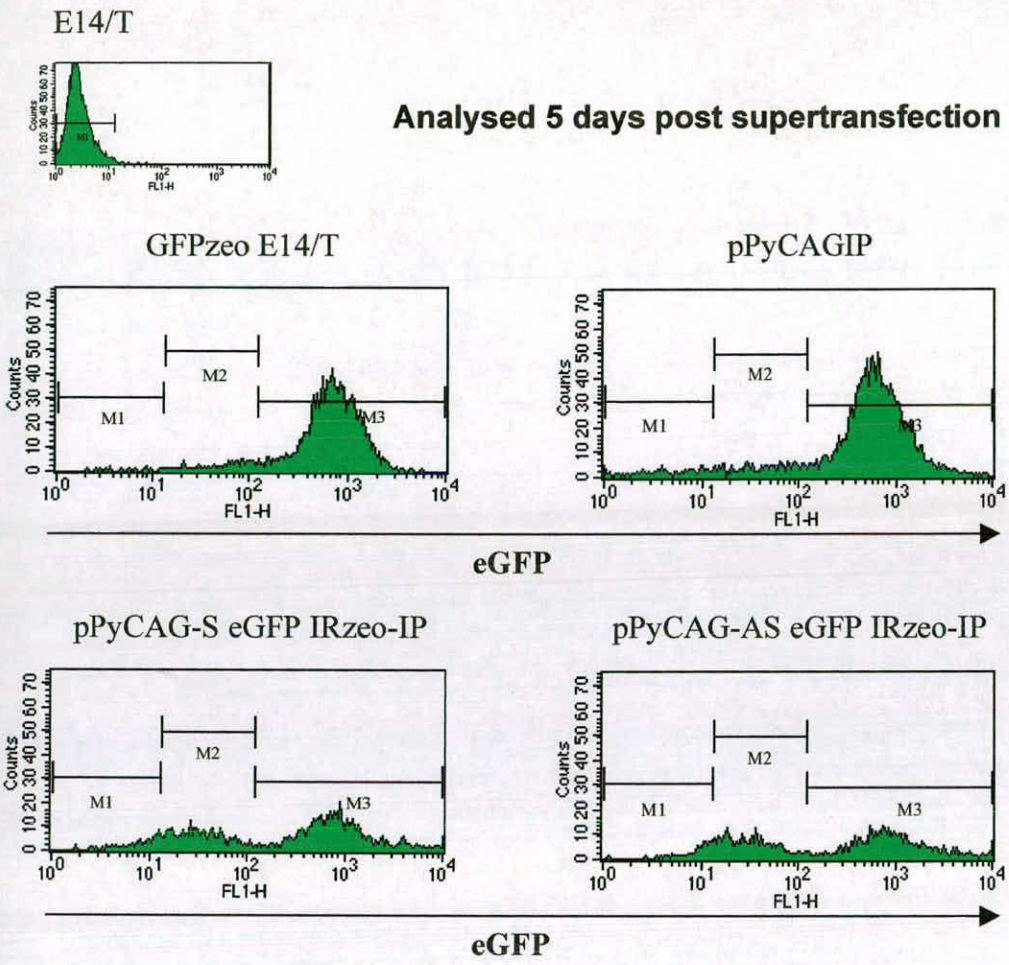


Figure 6.3. Sequence specific inhibition of eGFP fluorescence from a GFPzeo-E14/T ES cell line by expression of an eGFP double stranded hairpin RNA.

FACS analysis of eGFP expression from the stable GFP-zeo-E14/T ES cell line supertransfected with the control vector, pPyCAGIP or two of the eGFP IR expressing vectors, pPyCAG-S eGFP IRzeo-IP and pPyCAG-AS eGFP IRzeo-IP. The untransfected E14/T ES cell line was analysed as a negative control for no eGFP expression, while the non transfected GFP-zeo-E14/T cell line acted as a positive control. Cells were analysed 5 days post supertransfection, selected in 3μg/ml of puromycin. M1 indicates the gating of eGFP negative cells, M2 the gating of cells with reduced fluorescence and M3 the gating of cells expressing high eGFP. The data are the results from one experiment.

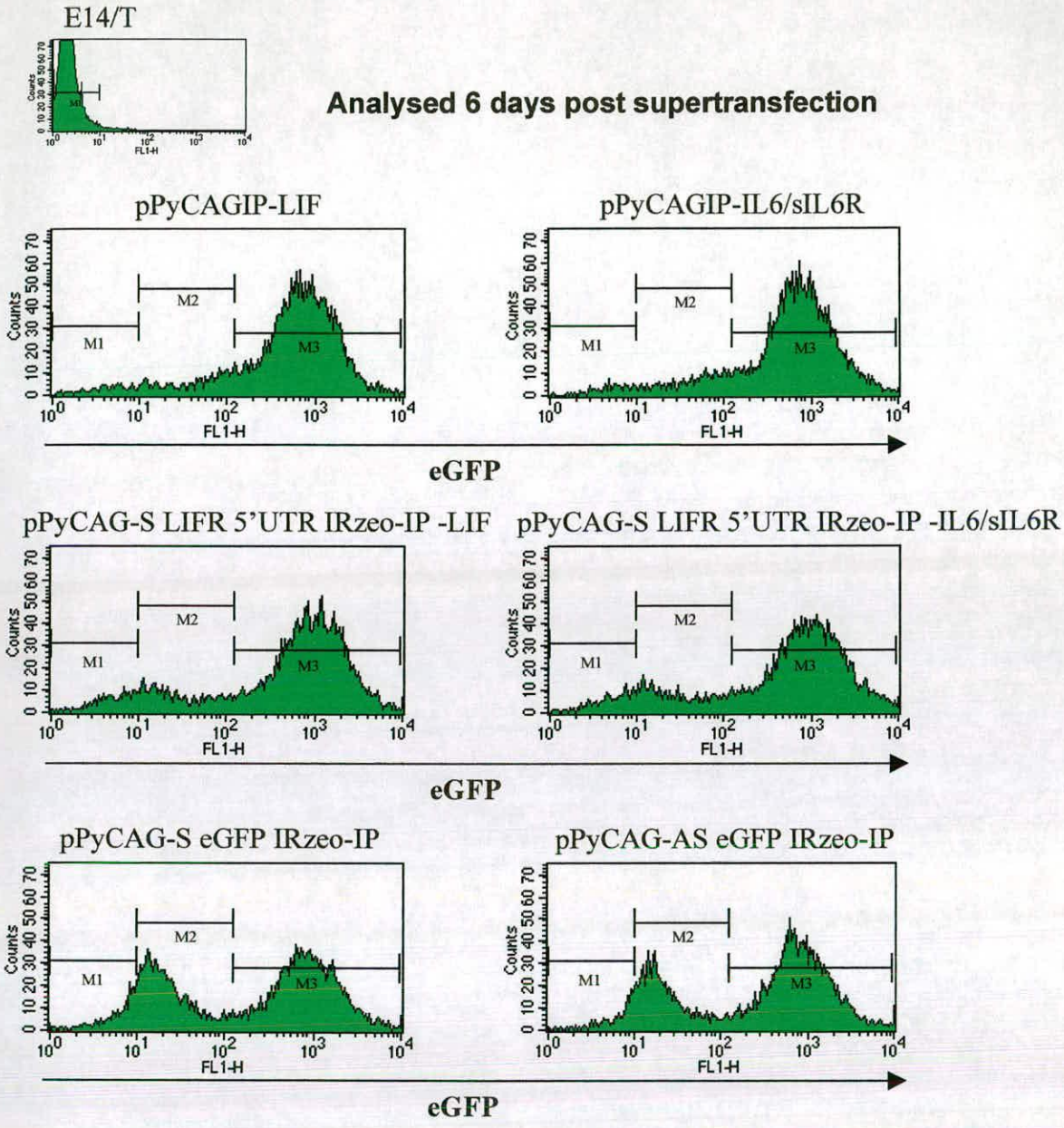


Figure 6.4. Sequence specific inhibition of eGFP expression from a GFP-zeo-E14/T ES cell line by transcription of an eGFP double stranded hairpin RNA.

FACS analysis of eGFP expression from the stable GFP-zeo-E14/T ES cell line supertransfected with the control vector, pPyCAGIP and pPyCAG-S LIFR 5'UTR IRzeo-IP or two of the eGFP IR expressing vectors, pPyCAG-S eGFP IRzeo-IP and pPyCAG-AS eGFP IRzeo-IP. The untransfected E14/T ES cell line was analysed as a negative control for no eGFP expression. Cells were analysed 6 days post supertransfection, selected in 1.5μg/ml of puromycin. pPyCAGIP and the LIFR 5'UTR IR vector was additionally grown in IL6/sIL6R, as well as LIF as for all other vectors. M1 indicates the gating of eGFP negative cells, M2 the gating of cells with reduced fluorescence and M3 the gating of cells expressing high eGFP.

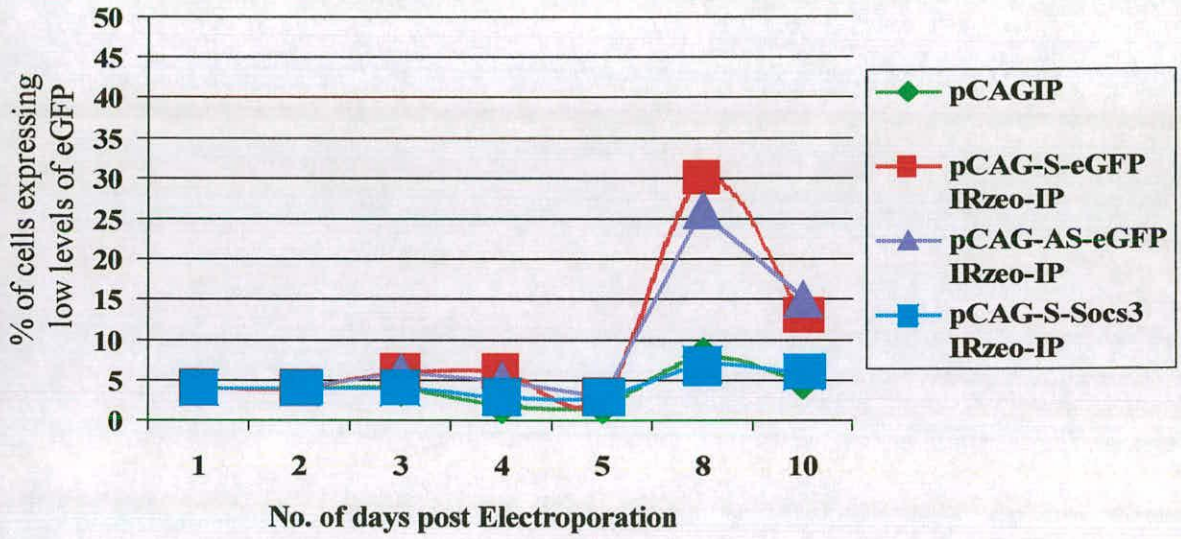
compared to pPyCAGIP (discussed in 6.4.4.2.). Cells were selected for 5 days, a total of 6 days post electroporation. Figure 6.4. shows the results of the FACS analyses. Cells transfected with the eGFP inverted repeat vectors possess two distinct peaks of eGFP. The low eGFP expressing population at 10^{-1} comprises about a third of those gated.

To determine when reduced eGFP is first detected and how long the effects last, eGFP expression was analysed over 1-10 days post electroporation. ES cells were electroporated individually with pPyCAGIP, pPyCAG-S eGFP IRzeo-IP, pPyCAG-AS eGFP IRzeo-IP and pPyCAG-S SOCS3 IRzeo-IP and analysed by FACS at 1, 2, 3, 4, 5, 8 and 10 days post electroporation. As seen from figure 6.5.a) the low expressing eGFP population emerges at day 8 comprising 30.3% and 26.4% of the total number of gated cells for pPyCAG-S eGFP IRzeo-IP and pPyCAG-AS eGFP IRzeo-IP respectively. However, by day 10 fewer cells expressing less eGFP are observed, where the population has decreased to 12.6% and 14.7%. The experiment was repeated looking more specifically at time points just before and after 8 days post electroporation when a reduced eGFP population was seen (figure 6.5.b). At 6 days the population of low expressing eGFP cells for both eGFP IR vectors is ~11% compared to control vectors (~2-3%) which increases to a maximum of ~42% at 7 days. By day 8 this population begins to decrease to 21% and continues to fall to 11-13% at day 10. So the loss of eGFP appears to be transitory, occurring at a specific time after the transfection of an eGFP specific inverted repeat vector. A possible explanation for the transitory effect may be that the level of dsRNA expression differs from day 6 to day 10. This is a possibility as it is known that the copy number of episomes

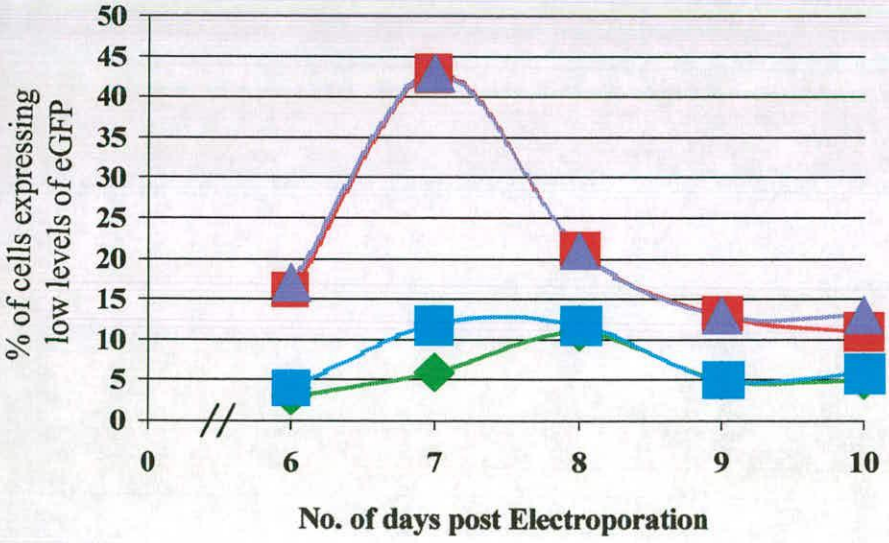
Figure 6.5. Time course for specific ablation of eGFP from a constitutively expressing eGFP ES cell line by in vitro dsRNA expression from inverted repeat vectors.

A constitutively expressing eGFP ES cell line, GFPzeo-E14/T was electroporated with the following vectors:- pPyCAGIP, pPyCAG-S eGFP IRzeo-IP, pPyCAG-AS eGFP IRzeo-IP and pPyCAG-S SOCS3 IRzeoIP and selected in LIF, G418 and 1.5μg/ml of puromycin. The cells were then analysed by FACS for eGFP expression after a) 1, 2, 3, 4, 5, 8 and 10 days; or b) 6, 7, 8, 9 and 10 post electroporation. The data presented are each from a single experiment and are shown as those cells expressing less eGFP as a percentage of gated cells.

a) Analysis after 1, 2, 3, 4, 5, 8 and 10 days post electroporation



b) Analysis after 6, 7, 8, 9 and 10 days post electroporation



decreases from day 10 to day 35 of selection (I.Chambers, unpublished). A reduction of eGFP expression might therefore only be seen when there is highest episomal copy number. However, reduced fluorescence may not be immediately apparent if GFP has a long half-life.

6.4.3.2. Staining cells with Annexin V conjugated PE and 7-AAD

It was apparent from the forward and side scatter plots of the FACS analysis data for all the eGFP ablation experiments that there were two populations of cells within the gated region analysed whenever a reduction of eGFP was observed. The low eGFP expressing population was entirely within the smaller of the two populations as determined by the lower forward scatter. The small size of these cells may be indicative of the cells either dying or undergoing apoptosis, which is a specific mechanism of induced cell death.

Annexin V-PE staining was used as an indication of whether the reduced eGFP expressing cells are apoptotic. eGFP ablation experiments were repeated and cells stained with Annexin V-PE (Annexin V-PE apoptosis detection kit, BD PharMingen). Annexin V is a phospholipid binding protein with a high affinity for phosphatidylserine (PS). PS is a membrane phospholipid that is translocated from the inner to the outer leaflet of the plasma membrane during the early stages of apoptosis when the membrane begins to loose its integrity (Martin et al.1995). The Annexin V used in this experiment is conjugated to the fluorochrome phycoerythrin (PE) so that it can be detected by flow cytometry simultaneously with the vital dye 7-Amino-actinomycin (7-AAD) (which is usually included to detect the loss of membrane integrity) and eGFP expression. Inclusion of 7-AAD

distinguishes early apoptotic cells from late apoptotic, as early apoptotic cells are Annexin V-PE positive and 7-AAD negative, while late apoptotic are positive for both.

E14/T ES cells were supertransfected as before. Cells were plated at 1 x 10^6 per 9cm plate for analysis by flow cytometry at 5-8 days post electroporation and selected with 1.5 μg/ml of puromycin. At 5 and 6 days post electroporation the low expressing eGFP population was not observed. By day 8 this population was seen for cells expressing the eGFP dsRNA hairpin (see figure 6.6.). As there are three variables to compare they were analysed on separate plots i.e. eGFP against Annexin V-PE and Annexin V-PE against 7-AAD. R6 is the gating of supposed live cells as determined by the size distribution from the forward and side scatter plots. The reduced eGFP population (designated as R9) stain highly with Annexin V-PE and at an intermediate level with 7-AAD (figure 6.6.). This is in comparison to equivalent plots obtained for cells supertransfected with pPyCAG-S SOCS3 IRzeo-IP (figure 6.7.). This staining profile indicates that the low eGFP population of cells are dying, most likely via apoptosis.

6.4.3.3. Northern Analysis to look for ablation of eGFP mRNA

As low eGFP cells are dying, possibly by apoptosis, the ablation observed may not be specific. Therefore to establish whether ablation is occurring via an RNAi mechanism, evidence of mRNA cleavage into predicted product sizes was investigated. Previously it had been shown by Zamore et al.(2000) that RNAi acts by specifically cleaving target mRNA at regions of homology with the introduced dsRNA. However, recent evidence by both Sijen et al.(2001) and Lipardi et

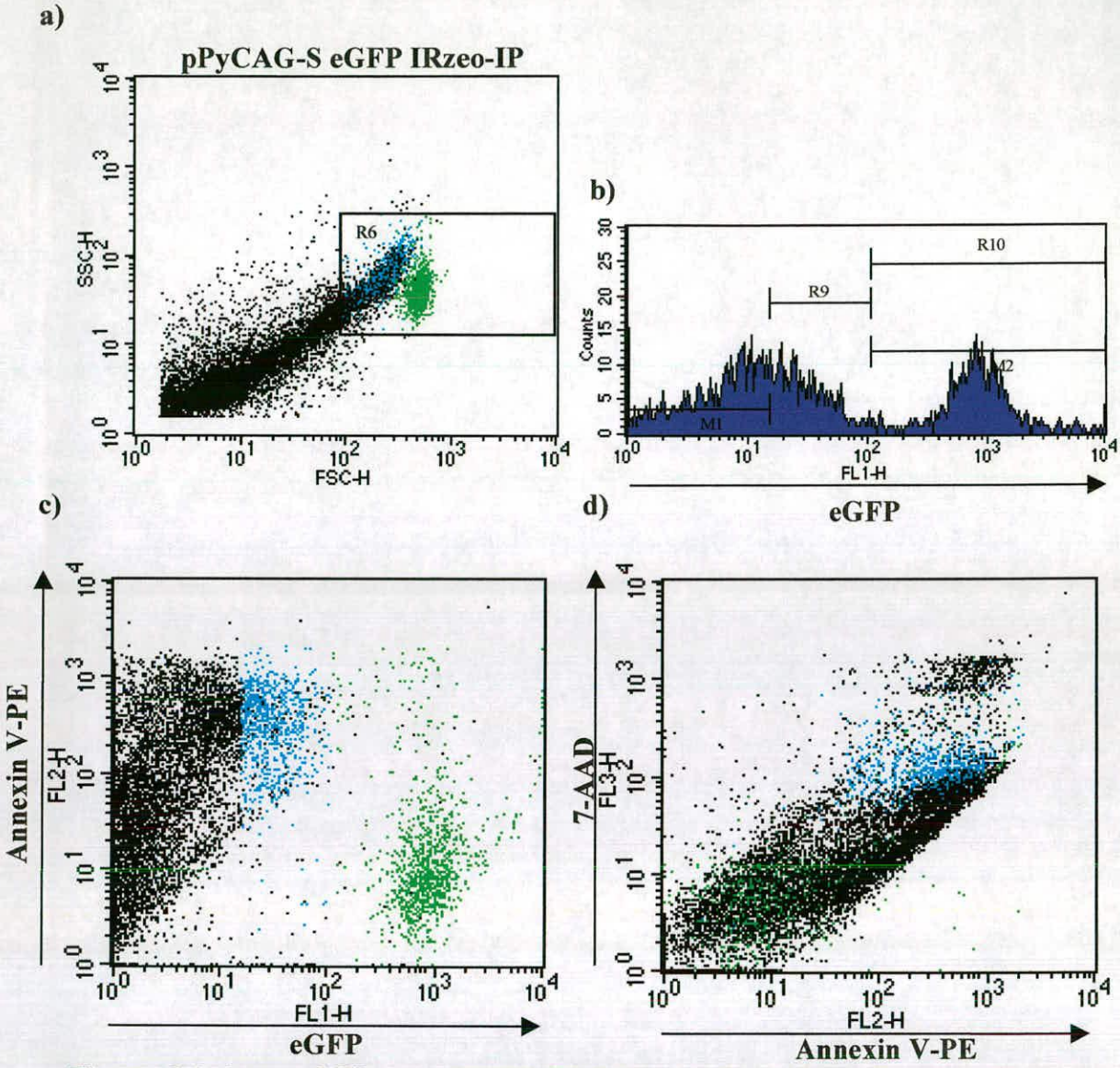


Figure 6.6. Low eGFP expressing ES cells generated by constitutive transcription of eGFP dsRNA hairpin from IR vectors are apoptotic, as determined by Annexin V-PE and 7-AAD staining.

The stable eGFP expressing cell line, GFP-zeo-E14/T supertransfected with pPyCAG-S eGFP IRzeo-IP were stained with Annexin V-PE and 7-AAD to determine apoptotic cells. Cells were analysed 8 days post electroporation, selected in 1.5μg/ml of puromycin. Cells determined as live cells from the forward/side scatter plot (a) were gated as R6, and analysed for eGFP expression (b). Low eGFP expressing cells are gated as R9 and high eGFP expressing as R10. The "live" low eGFP expressing cells, R6 and R9, are coloured blue, while the "live" high eGFP expressing cells, R6 and R10 are coloured green. These are then subsequently analysed on c) eGFP Vs AnnexinV-PE plot; and d) Annexin V-PE Vs 7-AAD plot.

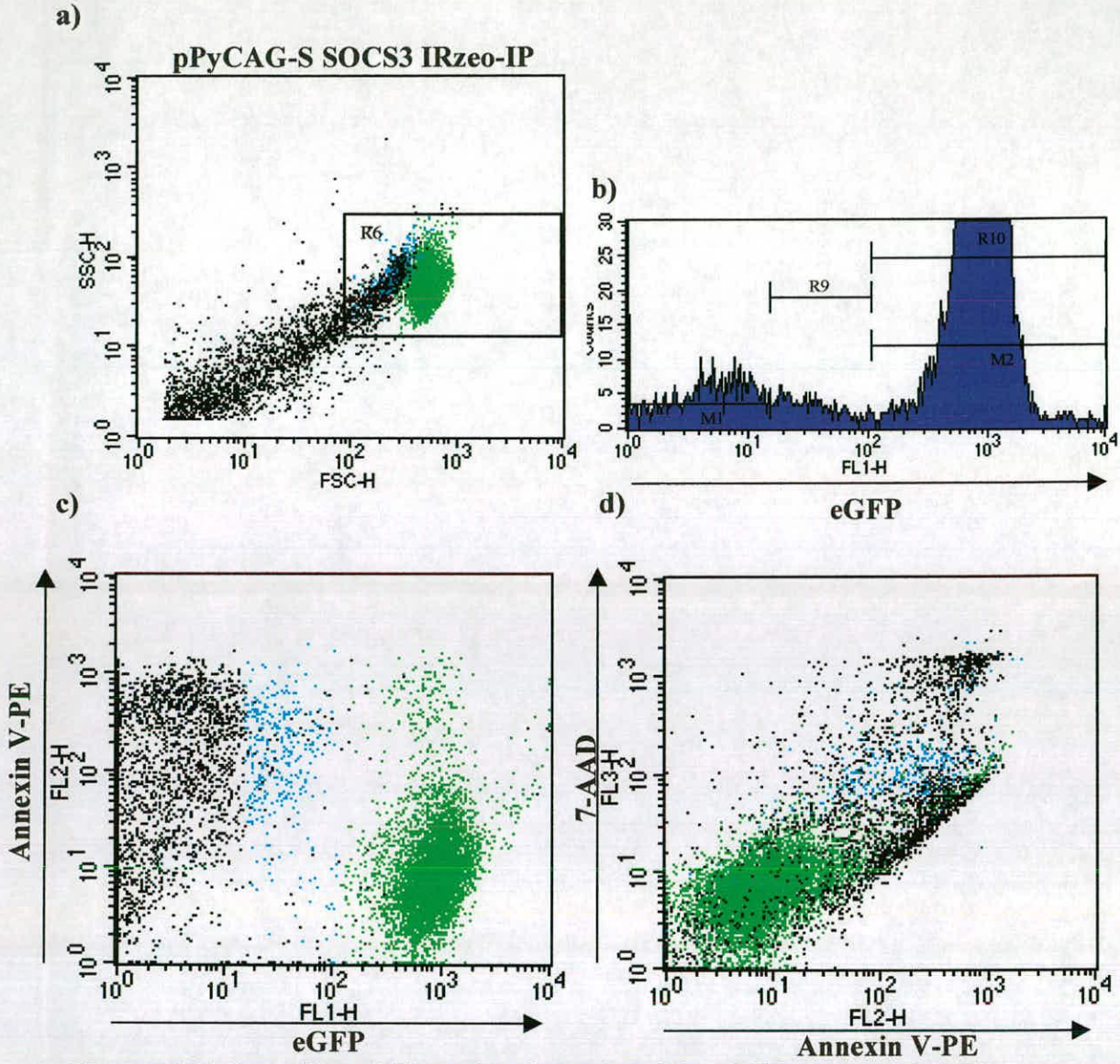


Figure 6.7. Specific eGFP reduction of stable eGFP expressing ES cells as demonstrated with a control vector, pPCAG-S SOCS3 IRzeo-IP determined by Annexin V-PE and 7-AAD staining.

The stable eGFP expressing cell line, GFP-zeo-E14/T supertransfected with pPyCAG-S SOCS3 IRzeo-IP were stained with Annexin V-PE and 7-AAD to determine apoptotic cells. Cells were analysed 8 days post electroporation, selected in 1.5μg/ml of puromycin. Cells determined as live cells from the forward/side scatter plot (a) were gated as R6, and analysed for eGFP expression (b). Low eGFP expressing cells are gated as R9 and high eGFP expressing as R10. The "live" low eGFP expressing cells, R6 and R9, are coloured blue, while the "live" high eGFP expressing cells, R6 and R10 are coloured green. These are then subsequently analysed on c) eGFP Vs AnnexinV-PE plot; and d) Annexin V-PE Vs 7-AAD plot.

al.(2001) in different systems, either *C.elegans* or a cell free extract of *Drosophila* embryos respectively, have demonstrated the existence of transitive RNAi. That is a secondary RNAi effect, where regions upstream of the original target site for dsRNA are ablated. This is due to an RNA dependent RNA polymerase (RdRP) amplification step primed by the siRNA, which results in elongation of regions upstream of the target and the subsequent generation of secondary siRNAs (see chapter 1 section 1.2.4.2.4.).

Although, transitive RNAi effects may be a factor in the eGFP ablation experiments, the main cleavage products of both target eGFP mRNA and the eGFP hairpin itself will be those expected from the initial hairpin sequences introduced (figure 6.8. and 6.9.). So to try to correlate the observed down regulation of eGFP expression seen with vectors expressing eGFP specific inverted repeats, RNA was extracted at 6 days post electroporation from the same experiment analysed in figure 6.4. At this time point for this experiment a population of cells expressing less eGFP is clearly seen. In addition, RNA was also prepared from a separate experiment at 13 days post electroporation at which time the lower eGFP expressing population had disappeared, as a negative control. These RNA samples were run on the same formaldehyde gel and probed sequentially (where the blot is stripped between probes) for eGFP mRNA with a probe specific to the β globin polyadenylation sequence and for equal loading by probing with GAPDH. The results are shown in figure 6.10. The expected sizes of the largest and smallest cleavage products of eGFP mRNA generated from the introduced eGFP hairpin have been determined as 2072bp and 1353bp. This is based upon the eGFP IR sequence, which is complementary to the entire coding

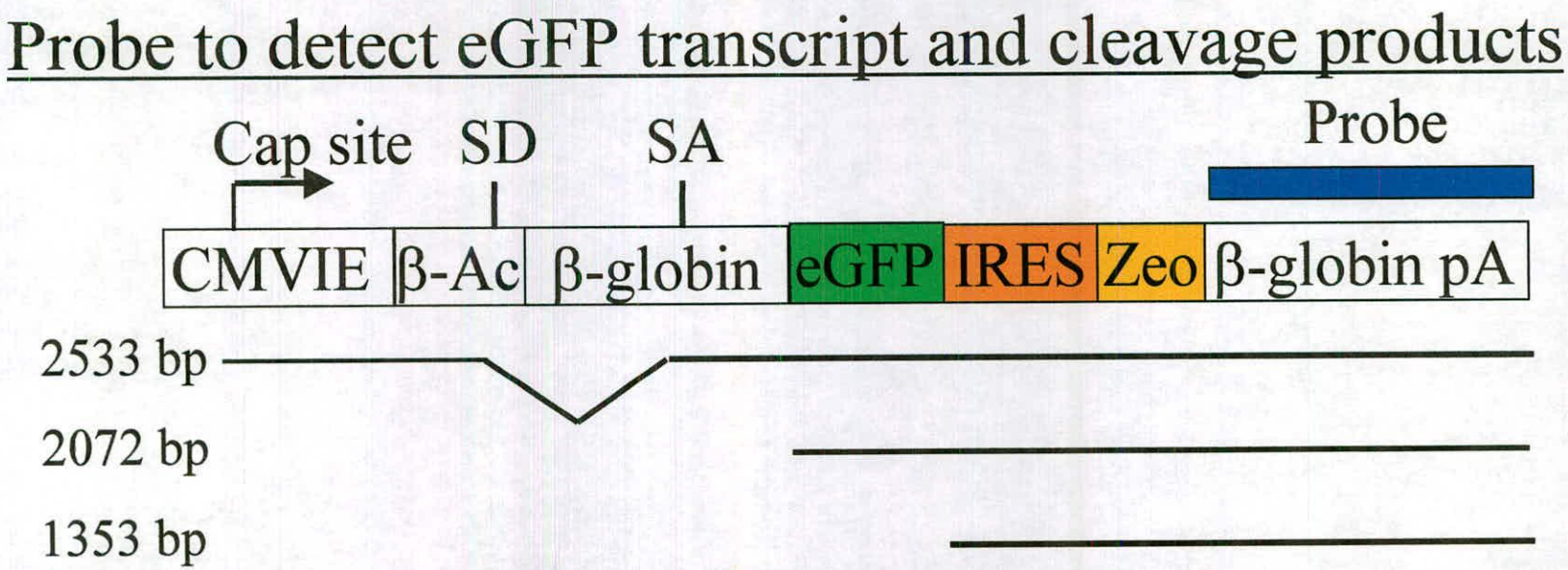


Figure 6.8. Diagram to show predicted sizes of the eGFP transcript from a randomly integrated CAG-eGFP IRES zeocin transgene and its primary cleavage products, generated through RNAi.

The CAG promoter contains a splice site, which is shown in the transcript below and is taken into account when calculating the size of the full length transcript. The two other transcripts show the sizes of the primary cleavage products formed, as the eGFP inverted repeat vector used to ablate eGFP is directed to the whole of the eGFP coding sequence. The probe used is a 459bp β-globin polyadenylation sequence specific probe.

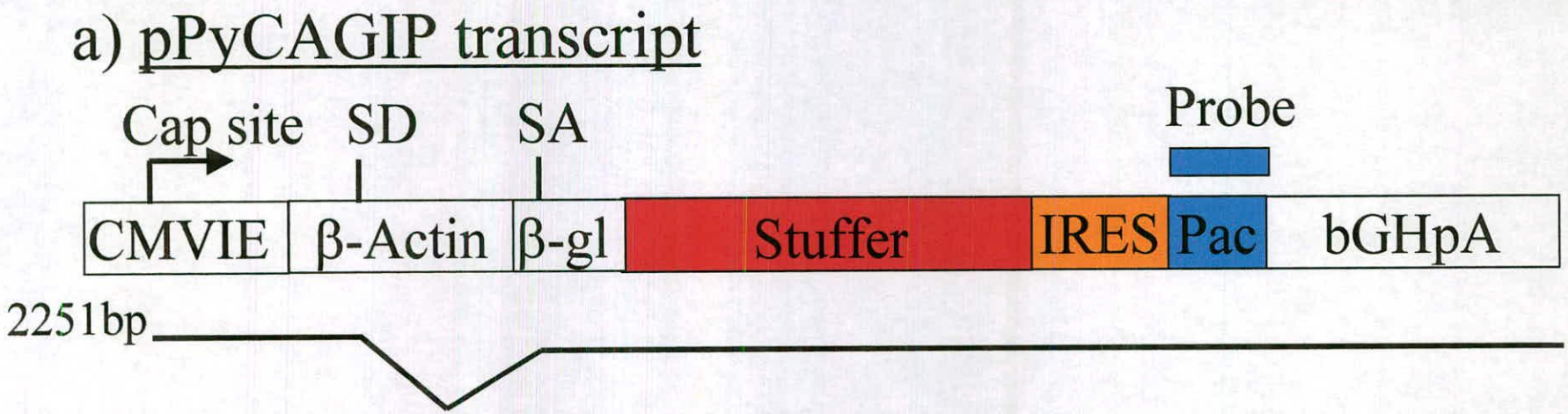
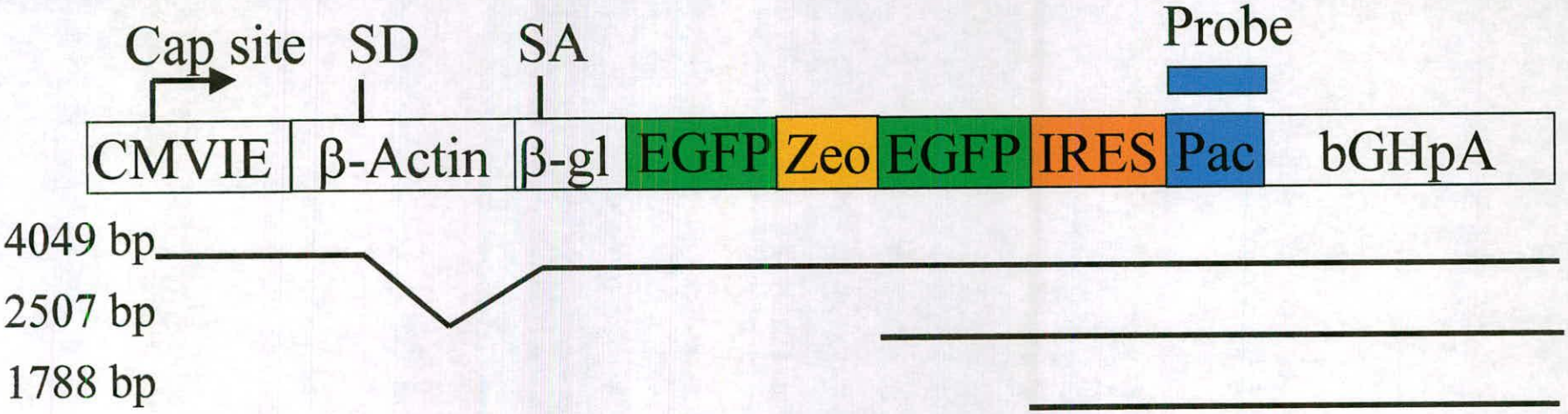


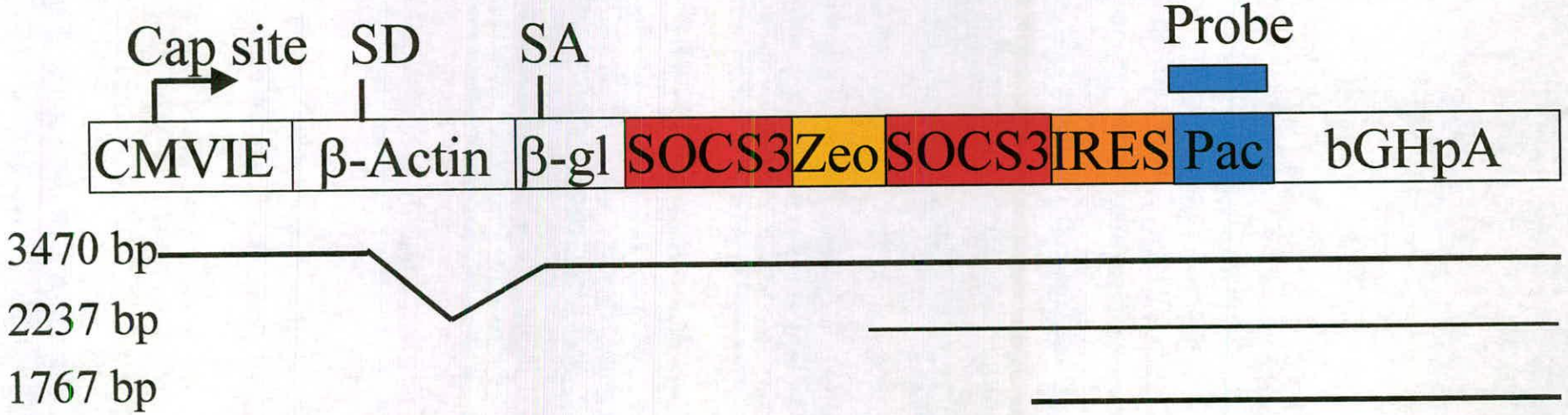
Figure 6.9. Diagram to show predicted sizes of the inverted repeat transcripts and their primary cleavage products, generated through RNAi.

The CAG promoter contains a splice site, which is shown in the full length transcripts and is taken into account when calculating the size. Transcript sizes also include the poly A tail ~230bp. Below the full length transcript of the inverted repeats only are shown the extremes in sizes of two primary cleavage products:- a) pPyCAGIP transcripts; b) pPyCAG-S eGFP IRzeo-IP or pPyCAG-AS eGFP IRzeo IP transcripts; and c) pPyCAG-S SOCS3 IRzeo-IP or pPyCAG-AS SOCS3 IRzeo-IP transcripts. The eGFP inverted repeat vector used to ablate eGFP is directed to the whole of the eGFP coding sequence, while that to SOCS3 is directed to 471bp of the 5'region. The probe used is a 668bp puromycin specific probe.

b) eGFP IR Transcript



c) SOCS3 IR Transcript



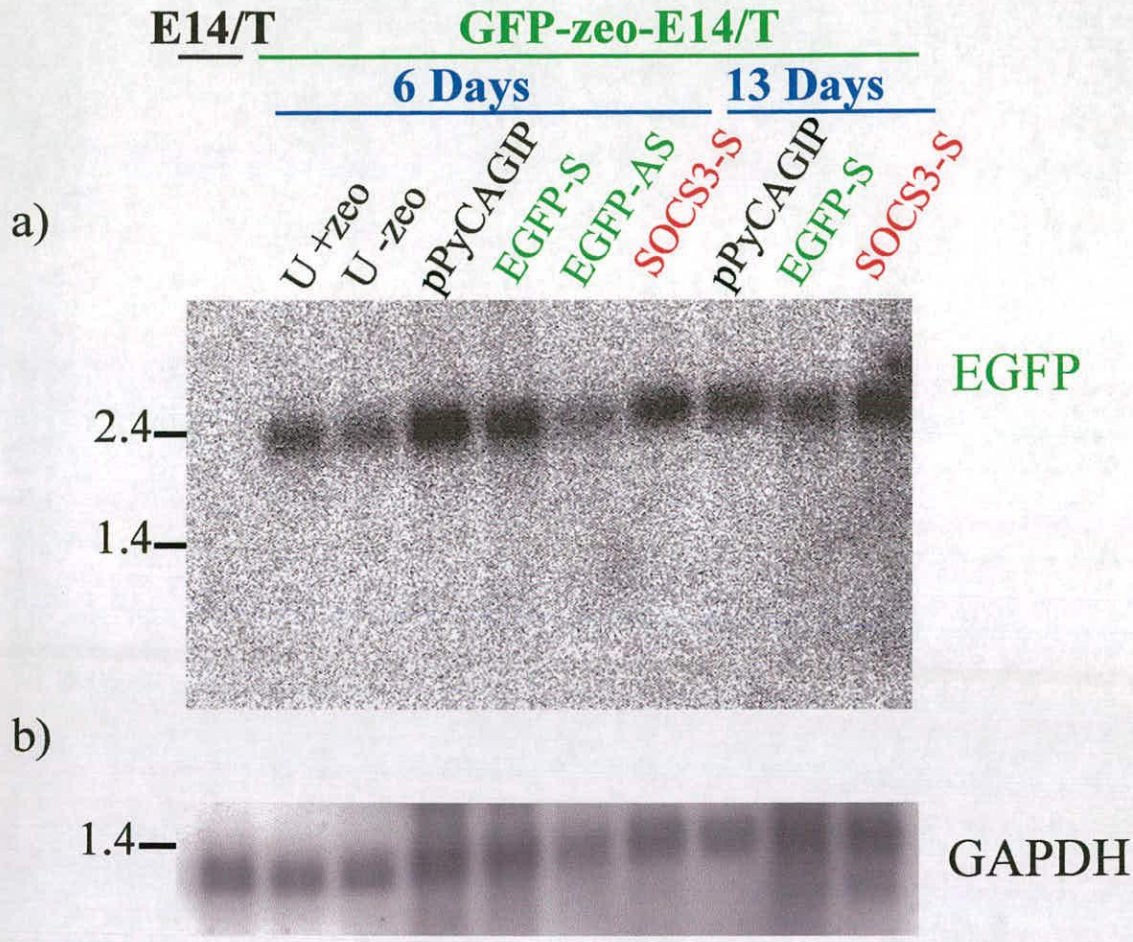


Figure 6.10. Northerns to detect ablation of eGFP from a randomly integrated CAG-eGFP IRES zeocin transgene.

GFP-zeo-E14/T ES cells were supertransfected with the following inverted repeat containing vectors which were either eGFP specific i.e. pPyCAG-S eGFP IRzeo-IP (EGFP-S) and pPyCAG-AS eGFP IRzeo-IP (EGFP-AS) or a non specific inverted repeat control i.e. pPyCAG-S SOCS3 IRzeo-IP (SOCS3-S) and pPyCAG-AS SOCS3 IRzeo-IP (SOCS3-AS). A non inverted repeat containing vector, pPyCAGIP was also included as a control. Total cell populations were selected for either 6 or 13 days as indicated with 1.5μg/ml of puromycin. 10μg of total RNA was loaded onto each lane along with untransfected E14/T ES cells and untransfected GFP-zeo-E14/T cells selected with (U +zeo) or without zeocin (U -zeo) (20μg/ml), where all other supertransfections are not selected in zeocin. The blot was then sequentially probed for a) eGFP with a 459bp β-globin specific probe; and b) equal loading with a GAPDH probe. Blot a) was exposed for 1 day on a phosphorimager screen; b) for 2 days.

sequence of eGFP, cleaving eGFP at its 5' or 3' most end (see figure 6.8.). These define the maximum and minimum sizes of expected eGFP mRNA cleavage products, however, products between these sizes will also be observed.

As seen from the eGFP mRNA probed blot, figure 6.10. a), the predicted eGFP full length transcript is detected at about 2533bp. However, no eGFPmRNA cleavage products were detected. The level of eGFP transcript between the lanes was then compared. This was done taking into account the GAPDH levels observed for each, so that a true level of eGFP message was noted, taking into account variations in loading. However, no specific ablation of eGFP mRNA was apparent for either of the eGFP inverted repeat vectors compared to the control vectors. It therefore seems unlikely that the reduced eGFP expression observed for cells transfected with an eGFP specific dsRNA is occurring via an RNAi mechanism.

6.4.3.4. Transfection efficiencies of eGFP inverted repeat vectors in a non-eGFP vs. an eGFP expressing cell line

All of my experiments to ablate integrated copy(ies) of eGFP with either of the eGFP inverted repeat vectors have consistently given more cells expressing less eGFP than any of the control vectors, including other vectors expressing a non specific inverted repeat. A possible reason for this may be that these cells are undergoing cell death due to a specific response of eGFP dsRNA to its target. To determine whether this is true the transfection efficiency of vectors was compared when electroporated into two different cell lines either expressing the *eGFP* target gene or not. The two different cell lines used were either GFPzeo-E14/T or E14/T ES cells from which the GFP expressing cell line was derived. These cell lines

were electroporated on the same day with the following vectors;- pPyCAGIP, pPyCAG-S eGFP IRzeo-IP and pPyCAG-AS eGFP IRzeo-IP. The cells were plated at three different densities into a 9cm dish, $5x10^4$, $2x10^5$ and $5x10^5$ in duplicate and then selected 30hours after transfection with LIF, G418 and 1.5µg/ml of puromycin. The results are shown below in table 2.

Table 2. Transfection efficiencies of eGFP inverted repeat vectors in a non-eGFP vs. an eGFP expressing cell line.

Vector	Plating Density	Cell Line					
		E14/T			GFPzeo-E14/T		
		Colony Number	Transfection Efficiency (%)		Colony Number	Transfection Efficiency (%)	
pPyCAGIP	$5x10^4$	256	0.512	Av.=	398	0.796	Av.=
		257	0.514	0.513	393	0.786	0.791
pPyCAG-S eGFP IRzeo-IP	$2x10^5$	31	0.016	Av.=	49	0.025	Av.=
		19	0.01	0.013	48	0.024	0.025
	$5x10^5$	124	0.025	Av.=	200	0.04	Av.=
		114	0.023	0.024	229	0.046	0.043
pPyCAG-AS eGFP IRzeo-IP	$2x10^5$	17	0.009	Av.=	39	0.02	Av.=
		12	0.006	0.008	34	0.017	0.019
	$5x10^5$	72	0.014	Av.=	157	0.031	Av.=
		59	0.012	0.013	132	0.026	0.029

The transfection efficiency for all vectors electroporated into the GFPzeo-E14/T cell line, i.e. containing an eGFP target, were about twice that of E14/T. However, as the control vector pPyCAGIP also has a higher transfection efficiency when electroporated into GFPzeo-E14/T cells, this indicates that this cell line is inherently more amenable for transfection. So taking this into account

and instead comparing the ratios of transfection efficiency (TxE) obtained for the higher plating density:-

1.E14/T Cell line

	pPyCAGIP		pPyCAG-SeGFP IRzeo-IP		pPyCAG-ASeGFP IRzeo-IP
TxE	0.513		0.024		0.013
Ratio	1	:	0.047	:	0.025

2. GFPzeo-E14/T

	pPyCAGIP		pPyCAG-SeGFP IRzeo-IP		pPyCAG-ASeGFP IRzeo-IP
TxE	0.791		0.043		0.029
Ratio	1	:	0.054	:	0.037

Comparing the ratios of transfection efficiency where the transfection efficiency of pPyCAGIP is given the value of 1, the transfection efficiencies of the eGFP specific inverted repeat vectors are about the same whether the eGFP target is present or not. So it would appear that the low transfection efficiency of eGFP IR vectors is not due to a specific response of eGFP dsRNA to its target and that the greater number of dying cells seen is not due to a specific RNAi response.

6.4.4. SOCS3 as a target for RNAi Ablation

6.4.4.1. Rationale and Strategy for SOCS3 ablation experiments

The rationale for choosing SOCS3 as a target for ablation is that SOCS3 mRNA is an endogenous target that can be induced with the addition of LIF (I.Chambers-Personal Communication). This is achieved by withdrawing LIF for 6-7 hrs before reapplication. The ability to induce SOCS3 mRNA was thought to

increase the likelihood of detecting ablation. In many model RNAi ablation experiments the targets are exogenous, which allows dsRNA directed against the target to be added to the system first and become processed before addition of the target. For example Tuschl et al.(1999) developed a *Drosophila* embryo lysate *in vitro* system and found that preincubation of a luciferase dsRNA specific to its target, in the lysate before addition of the target mRNA potentiates its RNAi activity. This allows time for the long dsRNA to be cleaved into 21-23nt siRNAs, which are the mediators of RNAi that direct the cleavage of the target mRNA. SOCS3 is a good target, because it is endogenous yet can be depleted from cells allowing time for the dsRNA hairpin to be transcribed and processed into active siRNAs, before its subsequent induction.

The strategy for SOCS3 ablation experiments was to supertransfect individually into E14/T ES cells both the SOCS3 inverted repeat containing vectors pPyCAG-S SOCS3 IRzeo-IP and pPyCAG-AS SOCS3 IRzeo-IP as well as two control vectors, one without an inverted repeat, pPyCAGIP and one with an irrelevant inverted repeat sequence i.e. pPyCAG-S eGFP IRzeo-IP. Stably maintained populations from these supertransfections were selected with 1.5μg/ml of puromycin for 18-24 days, as the transfection efficiency for each vector was different, and expanded until there were enough cells for the subsequent induction experiments. Induction experiments involved incubation in GMEM with 10% FCS without LIF for 1 hr to remove any residual LIF before replacing the media with GMEM alone for 6 hrs. Induction of SOCS3 mRNA was initiated by applying GMEM containing LIF, for a certain amount of time. The time required for maximal induction had already been analysed previously in the lab

(I.Chambers, unpublished). SOCS3 mRNA is maximally induced after addition of LIF for 45 minutes, but then is down regulated after 3 hours. A number of different time points were included to observe SOCS3 mRNA being induced over time and also to determine whether the message is cleaved. The time points examined at were 0, 10, 25, 45, 90 and 180 minutes. RNA for Northern analysis was extracted at these time points for subsequent analysis. RNA samples were loaded on two separate gels, cells stably maintaining pPyCAGIP or pPyCAG-S SOCS3 IRzeo-IP on one gel and pPyCAG-AS SOCS3 IRzeo-IP or pPyCAG-S eGFP IRzeo-IP on another gel. The Northerns were sequentially probed for SOCS3 endogenous mRNA (figure 6.11.a), IR transcript levels using a puromycin probe (figure 6.11.b) and equal loading using a probe directed against GAPDH (figure 6.11.c).

As can be seen SOCS3 mRNA levels are maximally induced at about 45-90 minutes and decrease after 3 hours for all the electroporated maintained cell lines. However, no cleavage products were observed. The endogenous SOCS3 transcript is 2.8kb, and predicted primary cleavage products are expected to be between 2.3kb and 1.7kb.

To determine whether transcripts from the introduced vector were being expressed and cleaved, the Northerns were stripped and probed with a puromycin probe. The predicted transcript sizes and their cleavage products are seen in figure 6.11. The full length transcripts for both SOCS3 and eGFP inverted repeat vectors were detected at ~3.5 kb and ~4 kb respectively. In addition to the smearing seen there is a distinct band at ~1.7 kb which is the predicted size of a primary cleavage product expected from the cleavage of these inverted repeat transcripts (figure

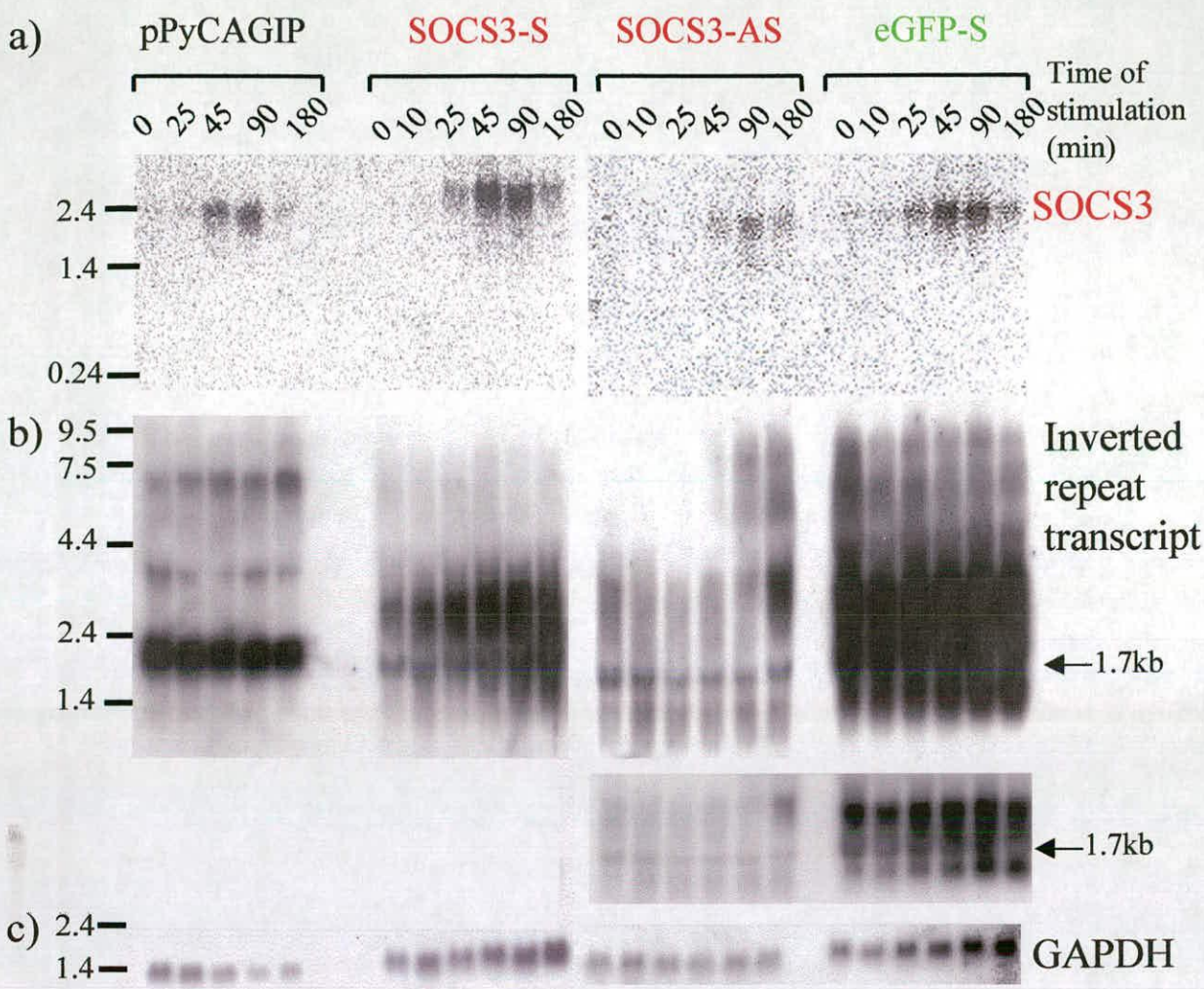


Figure 6.11. Analysis of SOCS3 mRNA and inverted repeat transcripts from SOCS3 induction experiments induced with LIF.

E14/T ES cell lines supertransfected with either two SOCS3 inverted repeat containing vectors, pPyCAG-S SOCS3 IRzeo-IP (SOCS3-S) and pPyCAG-AS SOCS3 IRzeo-IP (SOCS3-AS) or a non specific inverted repeat containing vector pPyCAG-S eGFP IRzeo-IP or a non inverted repeat containing vector pPyCAGIP were established prior to induction. Cell lines were maintained from pooled populations for 18-24 days by selecting with 1.5μg/ml of puromycin. The cell lines were then induced with LIF for 0, 10, 25, 45, 90 and 180 minutes and subsequently lysed and RNA loaded (10μg) onto a formaldehyde gel. Blots were then probed for a) endogenous SOCS3 mRNA with a 559bp SOCS3 probe; b) levels of inverted repeat transcripts with a 668bp puromycin probe; and c) equal loading with a GAPDH probe. The expected size of SOCS3 mRNA is 2.8kb, its predicted primary cleavage products are between 1.7 and 2.3kb. SOCS3 inverted repeat transcripts are 3470bp and primary cleavage products are between 2237bp and 1767bp. The eGFP inverted repeat transcript is 4049bp, and its predicted primary cleavage products are 2507bp-1788bp. The size of pPyCAGIP's main transcript is 2251bp, while due to the presence of two additional polyadenylation sites downstream of the main one, additional transcripts are of 2837-2866bp and 5432-5461bp are detected. Blot a) was exposed for 1 day on a phosphorimager screen; b) the first blot for 4 days, the second blot for 7 or 4 days on autoradiographic film; c) for 2 days on autoradiographic film.

6.9.b) c). Thus demonstrating that the inverted repeats are being transcribed once electroporated into cells and that there is some cleavage taking place. The bands detected for pPyCAG-AS SOCS3 IRzeo-IP plasmid are fainter than the corresponding SOCS3 IR vector where the zeocin is cloned in the reverse orientation as the loading of total RNA was lower than for the other cell lines as seen with the GAPDH probe (figure 6.11.c). The control vector pPyCAGIP, which does not contain an inverted repeat sequence correspondingly does not show any cleavage products. The major transcript at 2.25kb is detected as well as two minor transcripts at 2.8kb and 5.4kb. The minor transcripts are due to termination at two additional SV40 polyadenylation sites found in the vector downstream of the major polyadenylation signal. The bands observed for the control vector are distinct with no smearing, a further indication that the transcript is not being cleaved.

The transcript level from the eGFP inverted repeat vector appears to be higher than those for the SOCS3 inverted repeat vectors. This implies that the reduced eGFP expression observed previously with the eGFP IR plasmid may be due to this difference in transcript levels. It is known that dsRNA can trigger an interferon response in mammalian cells leading to non specific mRNA degradation and a global shut down of protein synthesis (Stark et al.1998). This would lead to cell death, which is consistent with the results of my eGFP ablation experiments (see discussion).

6.4.4.2. Transfection Efficiences of Inverted repeat containing vectors vs. non Inverted repeat vectors

It was noticed that the transfection efficiencies of the inverted repeat vectors compared to the control, pPyCAGIP vector were consistently lower. Transfection efficiencies are tabulated below, including those of the LIFR, eGFP and SOCS3 specific inverted repeat vectors which were used in the previously described eGFP ablation experiments. See Table 3.

Table 3. Transfection efficiencies of inverted repeat containing vectors vs. non inverted repeat vectors

Experiment	Vector	Plating density	No.of Colonies per plate	Transfection Efficiency (%)
1. 3μg/ml puromycin selection; fluorescence analysed at 3 days of selection. Equivalent to 5 days post electroporation	pPyCAGIP	$5x10^4$	499	0.998
	pPyCAG-SeGFP IRzeo-IP	$2x10^5$	104	0.052
	pPyCAG-ASeGFP IRzeo-IP	$2x10^5$	88	0.044
2. 1.5μg/ml of puromycin selection; fluorescence analysed at 5 days of selection. Equivalent to 6 days post electroporation	pPyCAGIP (grown in LIF)	$5x10^4$	804	1.608
	pPyCAGIP (grown in IL6/sIL6R)	$5x10^4$	1027	2.054
	pPyCAGLIFR 5'UTR IRzeo-IP (grown in LIF)	$5x10^4$	125	0.25
	pPyCAGLIFR 5'UTR IRzeo-IP (grown in IL6/sIL6R)	$5x10^4$	304	0.608
	pPyCAG-SSOCS3 IRzeoIP	$5x10^4$	62	0.124

	pPyCAG-SeGFP IRzeo-IP	$5x10^4$	2	0.004
	pPyCAG-ASeGFP IRzeo-IP	$5x10^4$	10	0.02
3. 1.5μg/ml of puromycin selection; fluorescence analysed at 0-4, 7 and 9 days of selection. Equivalent to 1-5, 8 and 10 days post electroporation	pPyCAGIP	$5x10^5$	*	
	pPyCAG-SSOCS3 IRzeoIP	$5x10^5$	900	0.18
	pPyCAG-SeGFP IRzeo-IP	$5x10^5$	446	0.089
	pPyCAG-ASeGFP IRzeo-IP	$5x10^5$	272	0.054
4. 1.5μg/ml of puromycin selection; fluorescence analysed at 4-8 days of selection. Equivalent to 6-10 days post electroporation	pPyCAGIP	$5x10^5$	*	
	pPyCAG-SSOCS3 IRzeo-IP	$5x10^5$	2920	0.584
	pPyCAG-SeGFP IRzeo-IP	$5x10^5$	737	0.147
	pPyCAG-ASeGFP IRzeo-IP	$5x10^5$	860	0.172

* Too many to count at this plating density.

Results to note from comparing transfection efficiencies are that the eGFP inverted repeat vectors consistently give much lower transfection efficiencies than either pPyCAGIP or other inverted repeat containing vectors. The transfection efficiency is normally about 20 fold lower than pPyCAGIP (but up to 80-fold lower has been seen), and 2-13 fold lower than the other inverted repeat containing vectors e.g. SOCS3 and LIFR5'UTR. This implies a general phenomenon of a decrease of transfection efficiency with inverted repeat

containing vectors, where the transfection efficiency for the eGFP inverted repeat vectors are the lowest observed. This suggests that expression of inverted repeat transcripts in the cell is either toxic or detrimental for cell growth, possibly by inducing an interferon response, or that either the replication of such vectors or transcription of inverted repeats is less efficient, as has been observed of the replication of direct repeat sequences in bacteria (LaDuca et al.1983). The lower transfection efficiency of the eGFP inverted repeat vectors are consistent with more cell death seen with transfection of these vectors which may be attributable to the higher transcript levels for these vectors (see section 6.4.3.2. and 6.4.4.1.).

6.4.4.3. Transfection efficiencies of supertransfected SOCS3 inverted repeat vectors vs. integration of these vectors

As transfection efficiencies for inverted repeat containing vectors are so low compared to vectors without the inverted repeat, and the values are similar to that observed when the vector has integrated, a comparison of the two situations was examined. E14/T ES cells were supertransfected as before with 20μg of pPyCAGIP, pPyCAG-S SOCS3 IRzeo-IP, pPyCAG-AS SOCS3 IRzeo-IP and pPyCAG-S eGFP IRzeo-IP and selected in 1.5μg/ml of puromycin for 6 days before staining with Leishmans. E14TG2a ES cells were electroporated with the same vectors linearised by digestion with ScaI (100μg). Cells were selected with 1.5μg/ml of puromycin and stained after 9 days. The transfection efficiencies observed are shown in table 4.

Table 4.Transfection efficiencies of SOCS3 inverted repeat vectors vs. integration of these vectors.

Vector	Plating Density	Cell Line					
		E14/T			E14TG2a		
		Colony Number per plate	Transfection Efficiency (%)		Colony Number per plate	Transfection Efficiency (%)	
pPyCAGIP	$5x10^4$	334	0.668	Av.= 0.645	-	-	Av.= 0.049
	$2x10^5$	1245	0.623		100	0.05	
	$5x10^5$	-	-		236	0.047	
pPyCAG-S SOCS3 IRzeo-IP	$2x10^5$	242	0.121	Av.= 0.161	-	-	0.048
	$5x10^5$	1008	0.202		242	0.048	
pPyCAG-AS SOCS3 IRzeo-IP	$2x10^5$	552	0.276	Av.= 0.298	-	-	0.037
	$5x10^5$	1601	0.320		187	0.037	
pPyCAG-S eGFP IRzeo-IP	$5x10^5$	392	0.078	Av.= 0.100	144	0.028	0.028
	$1x10^6$	1224	0.122		-	-	

As can be seen from the results above, the frequency of isolating stable integrants is reduced compared to supertransfection, at three to eight fold lower. This may indicate that the low transfection efficiency observed with the IR containing vectors is not because they cannot be propagated episomally. However, to demonstrate this conclusively copy numbers of IR containing vectors will need to be scored.

6.5. Discussion

6.5.1. RNAi ablation in ES cells using siRNA

Results from the siRNA mediated ablation experiments performed in this Thesis clearly show that RNAi exists in ES cells. This has subsequently been shown to be a general phenomenon of ES cells (Yang et al.2001; Billy et al.2001; Paddison et al.2002). Yang et al.(2001) demonstrated RNAi by targeting eGFP by either transiently transfecting a plasmid with an eGFP IR directed sequence or by directly transfecting eGFP specific dsRNA transcribed *in vitro*. Billy et al.(2001) again targeted eGFP and were able to ablate 60% of GFP fluorescence when electroporating ES cells, with a GFP plasmid and a ~700bp long GFP specific dsRNA. Paddison et al.(2002) were able to show specific suppression of firefly and Renilla luciferase by transfecting ES cells with a 500 bp dsRNA corresponding to the first 500 nts of each luciferase, where inhibition was 80-90% complete.

6.5.2. The feasibility of stable expression of long hairpin dsRNA for gene ablation in ES cells

Phenotypic ablation of LIFR mRNA target was not observed when targeting 1071 bps of 5' region of the coding sequence, including the ATG translational start site. The level of transcript needed to ablate LIFR mRNA is probably not an issue here, as the level of LIFR mRNA in the cell is very low, as it is detectable by RNase protection only. However, to understand more clearly what was happening at a biochemical level eGFP and SOCS3 mRNA were targeted, as these mRNAs are easily detected by Northern analysis.

The first thing to note of the eGFP RNAi experiments is the appearance of a population of cells expressing less eGFP when a stable GFP expressing ES cell clone is supertransfected with an eGFP specific IR vector. This effect is specific as it is only observed with the eGFP IR vectors and not with IR vectors specific to SOCS3 or LIFR 5' UTR. The results suggest that the reduction of eGFP expression maybe due to an RNAi effect. Similar results have been observed by Yang et al.(2001), where they saw specific inhibition of eGFP expression from ES cell clones with a single integrated copy of the *eGFP* gene. However, they introduced dsRNA by transfecting 3 x 10^5 ES cells with 3μg of *in vitro* transcribed dsRNA specific to 547 bp beginning from the ATG start codon of eGFP. Similar to my data they only observed a new population of cells with reduced fluorescence when dsRNA to eGFP was transfected, where their control was dsRNA to lacZ. The greatest decrease in fluorescence of over 70% was observed at 48 hrs after transfection which diminished steadily until at 124 hrs ablation was no longer seen. The authors claim that the transitory population was probably due to the dilution of dsRNA per cell as ES cells replicated. As dsRNA was *in vitro* transcribed cells could not be selected so would contain less and less dsRNA. In my experiments cells expressing the dsRNA IRES puromycin transcript were selected for. The transitory nature of eGFP reduction in my experiments may instead be due to a possible decrease in the copy number of the episome resulting in a reduction of the inverted repeat transcript level. This could be determined by examining both the copy number and transcript levels at time points before and after a reduction in eGFP expression is seen. However, there are potential complications with the IRES linked selection strategy employed. This is because

once the long dsRNA hairpin has been processed the cleaved products are no longer selectable with an appropriate drug. However, I have been able to detect cleavage products of the inverted repeat transcripts, implying that the cleaved products are present in the cell and are not selected against immediately.

Although it appears that the reduction in eGFP could be due to an RNAi effect based upon both my data and published work (Yang et al.2001), I have found that those cells expressing less eGFP are dying. This may be by apoptosis as determined by staining with Annexin V-PE and 7-AAD (section 6.4.3.2). The first indication of this came initially from observing that the population with reduced fluorescence almost entirely had a lower forward scatter, and so were smaller than the higher fluorescent cells, as observed in forward/side scatter plots from flow cytometric analysis (see figure 6.12.). The difference in size indicated that they may be dying, as they are still quite large, but are beginning to loose their membrane integrity. This can be seen on the scatter plots as these cells are part of a population that joins with the large amount of debris that is present. Staining with Annexin V-PE and 7-AAD showed that the cells were either apoptotic or necrotic because they stained highly with Annexin V and at an intermediate level with 7-AAD, where Annexin V positive and 7-AAD negative cells are defined as being early apoptotic and Annexin V positive 7-AAD positive cells are in late apoptosis or are already dead. Staining with Annexin V is an indication that the cells are apoptotic, however additional methods of identification such as DNA fragmentation assays should be performed to confirm this.

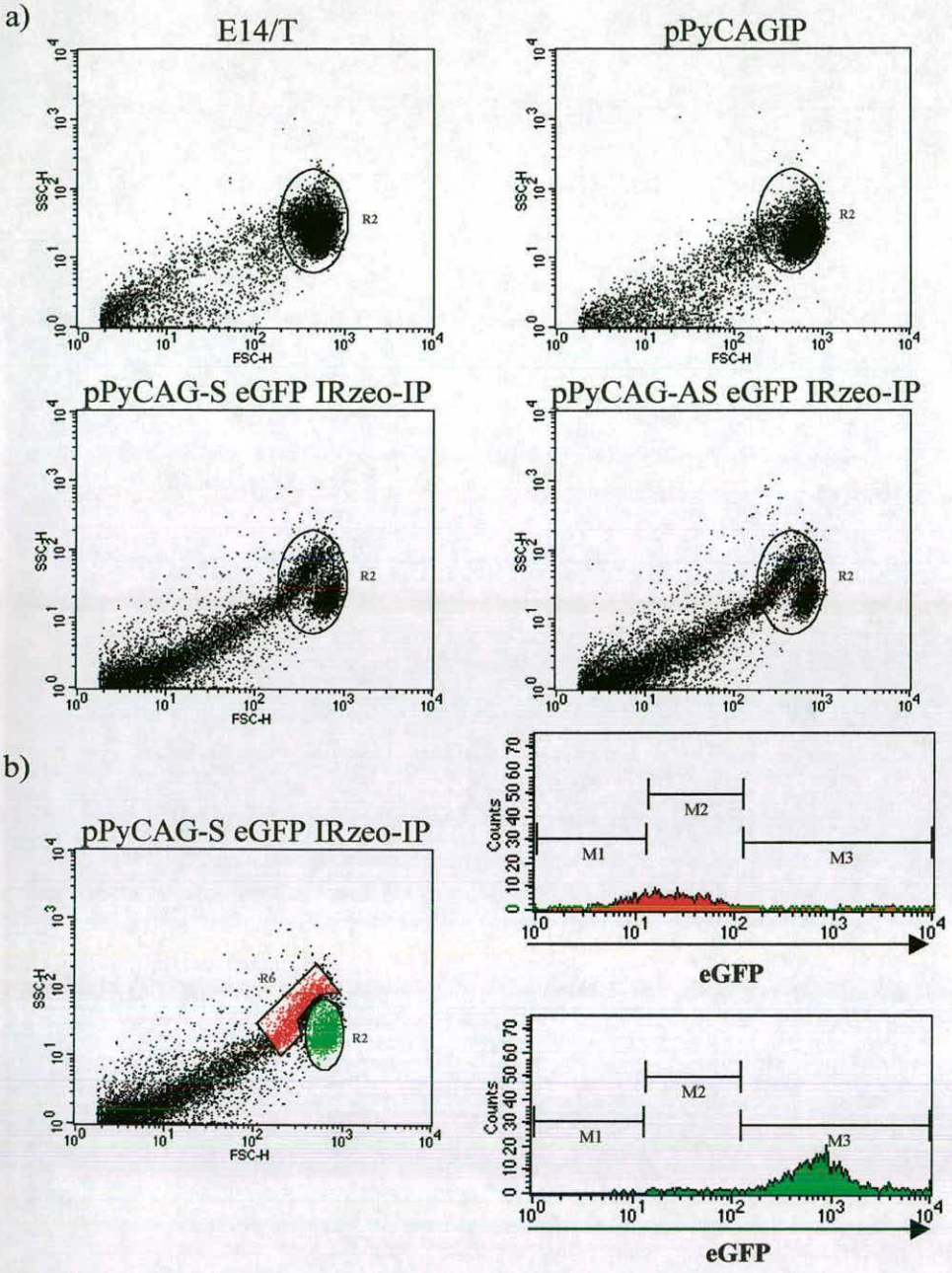


Figure 6.12. Cell morphology of GFP-zeo-E14/T ES cells supertransfected with IR versus non IR containing vectors.

a)Forward and Side scatter plots of the stable eGFP expressing cell line, GFP-zeo-E14/T supertransfected with a control non IR containing vector, pPyCAGIP, or two eGFP IR containing vectors, pPyCAG-S eGFP IRzeo-IP and pPyCAG-AS eGFP IRzeo-IP. E14/T ES cells are the parental cells from which the GFP expressing cell line is derived. Cells were analysed 5 days post supertransfection, selected in 3μg/ml of puromycin. b)Analysis of eGFP expression from the two populations seen with pPyCAG-S eGFP IRzeo-IP transfected cells denoted as R6 and R2. Where R6 are those cells expressing reduced eGFP and R2 are those cells expressing high eGFP.

As a population of cells expressing reduced eGFP was seen, the most conclusive way to prove that it is due to RNAi was to show that the target mRNA and dsRNA are being cleaved. Northerns were examined for cleavage of RNA extracted from cells which showed two distinct populations in terms of fluorescence intensity (figure 6.4.) as well as RNA from cells where the low expressing eGFP population had disappeared. Ablation of the eGFP mRNA was not seen nor was cleavage of the message. This was also true of another mRNA target, SOCS3 targeted in separate SOCS3 ablation experiments. Ablation of the eGFP message may not have been detected as the percentage of cells expressing less eGFP was ~30% at the time point looked at (6 days post electroporation), a decrease which may be undetectable by Northern analysis. Alternatively, there may be no appreciable ablation for either mRNA target, especially as the reduced eGFP expression is probably a consequence of the cells dying. However, cleavage of the inverted repeat transcripts for eGFP and SOCS3 were detected, observed as smears rather than distinct bands. This is expected as there will be a variety of different sized cleavage products generated as is the nature of dsRNA processing into siRNAs. Although this shows that the inverted repeat transcripts are being processed, siRNAs of 21-23nt may not be produced or not in sufficient quantity, which may be another reason why ablation is not seen.

The transfection efficiencies of cells supertransfected with inverted repeat containing vectors specific to LIFR, eGFP and SOCS3 were much lower compared to non inverted repeat containing vectors. This could be due to expression of dsRNA from inverted repeat sequences being either toxic or detrimental to cell growth as there is evidence that accumulation of small amounts

of dsRNA in mammalian cells results in the interferon response and the subsequent onset of apoptosis (reviewed in Tan and Katze 1999). The interferon response is mediated by two main enzymes: the double stranded RNA protein kinase, PKR (Manche et al.1992), and 2',5'-oligoadenylate synthetase (Minks et al.1979). Both are activated by long dsRNA and interferons (IFNs), typically required to mediate an antiviral response, as viruses replicate by producing dsRNA. Alternatively, the low transfection efficiencies of inverted repeat containing vectors may be due to either poor replication of such vectors or poor transcription of inverted repeat sequences. This has been observed of the replication of direct repeat sequences in bacteria (LaDuca et al.1983). These possibilities will be discussed further here.

In the first instance there is evidence in the literature that dsRNA induces the interferon response causing apoptosis via PKR, a serine/threonine specific protein kinase. Presently PKR has not been identified in ES cells, however it has been cloned in the mouse (Tanaka and Samuel 1994). dsRNA activates the inactive form of PKR by either causing a conformational change in the PKR molecule so that the catalytic domain becomes accessible or by promoting kinase dimerization which could lead to PKR activation through a transphosphorylation mechanism (Wu and Kaufman 1997). Once activated, PKR acts by phosphorylating the α subunit of the eukaryotic initiation factor (eIF-2). The role of eIF-2 is to form a complex with a methionine charged tRNA and present it to the ribosomal machinery for translation initiation (Merrick 1992). Phosphorylation of this protein on Ser^{51} is thought to result in tighter binding of eIF-2 to eIF-2B, which is the translation initiation exchange factor required to

recycle eIF-2-GDP to eIF-2-GTP. During the process of initiation codon recognition, eIF-2 is associated to GTP, which is subsequently hydrolysed to GDP after each round of initiation. Sequestering eIF2B reduces the rate of eIF-2 nucleotide exchange, thereby inhibiting protein synthesis and viral replication. The other pathway activated by dsRNA and interferons is the 2',5'-oligoadenylate synthetase (OAS) / RNase L pathway. OAS is activated by binding to dsRNA. Once activated, OAS coverts ATP into unusual short 2',5'-oligoadenylates (2-5A). 2-5A are able to bind to and activate RNase L, which cleaves single stranded RNA 3' of UpA, UpU and UpG sequences (Floyd-Smith et al.1981), leading to degradation of cellular RNA.

It is becoming increasingly clear that PKR triggers apoptosis. The first indication for such a role came from a study where HeLa cells expressing a catalytically inactive form of PKR did not rapidly undergo apoptosis compared to cells overexpressing the wild type PKR (Lee and Esteban 1994). Conversely overexpression of wild type human PKR, but not the catalytically inactive form in a transient DNA transfection system was sufficient to initiate an apoptotic response in COS-1 cells (Srivastava et al.1998). In addition there is evidence that embryo fibroblasts from PKR knock out mice develop resistance to apoptosis induced by dsRNA (Der et al.1997). These findings strongly demonstrate the existence of a PKR dependent apoptotic pathway in many cell lines, although its existence in ES cells is at present unknown. The apoptotic pathway is thought to be mediated in part through phosphorylation of eIF2α. As already demonstrated by Srivastava et al.(1998) with wild type PKR, overexpression of a serine 51 phosphorylated eIF-2 variant was sufficient to trigger apoptosis in COS-1 cells. A

possible way in which translational arrest induces apoptosis may be by suppression of translation of mRNAs encoding anti-apoptotic functions. In addition the apoptotic pathway is also known to be triggered by PKR through the FADD-caspase 8 pathway (Fas-associated death domain) (Balachandran et al.1998).

RNase L has also been shown to activate apoptosis. Castelli et al.(1997) have shown that inducible expression of the human RNase L gene results in the loss of NIH3T3 cell viability, which was confirmed as apoptotic cell death by in situ DNA fragmentation detection. DNA fragmentation quantitates the number of cells dying via DNA cleavage, which is a characteristic of apoptosis. A low rate of spontaneous apoptosis was observed prior to induction at 0-2.2%, which then increased to 8.8-13.5% of the total cells after induction of RNase L. Transient transfection of poliovirus infected HeLa cells with a dominant negative form of RNase L prevented virus induced apoptosis, maintaining cell viability at 80% compared to only 22% when HeLa cells were transfected with a control vector (Castelli et al.1998). More direct evidence of the role RNase L plays in apoptosis was determined from mice with a targeted disruption of the gene, preventing the synthesis of a functional ribonuclease (Zhou et al.1997). In addition to the antiviral effect of interferon α being impaired in RNase L $^{-/-}$ mice, these mice had unusually enlarged thymuses containing excess numbers of thymocytes, which was the direct result of the suppression of apoptosis. This was determined by measuring levels of apoptosis by in situ assays for DNA fragmentation, where there was a 47% decrease in the number of apoptotic cells in the thymuses of RNase L $^{-/-}$ mice compared with wild type.

In the second instance replication or transcription of inverted repeat containing vectors may be less efficient compared to non IR containing vectors (see section 6.4.4.2.). This phenomenon is well characterised in bacteria (for review see Leach 1996). Cloning IRs into *E.coli* results in either inviability or instability, where clones containing the construct are either not recovered or are recovered but at a low frequency with partial deletions of the inverted repeat sequence. However, inclusion of a spacer sequence between the inverted repeats is sufficient to allow *E.coli* cells to resume the replication of complete sequences (Warren and Green 1985). Warren and Green (1985) found that including a spacer sequence of at least 50 bp was sufficient to restore viability, while a longer spacer of ~150 bp was required for stability. This was determined by ligating a range of restriction fragments into the centres of inverted repeat sequences of ~950 bp. Cloning of the inverted repeat sequences specific to LIFR, eGFP and SOCS3 included a 568 bp zeocin selection cassette as a spacer sequence. The length of this spacer could be sufficient in reducing both the inviability and instability of replicating inverted repeat sequences in ES cells. However, whether the same replication problems are found in mammalian cells as in bacteria is not known. The possible replication or transcription problems of inverted repeat vectors may also be linked to the poor transfection efficiencies that have been observed for these vectors. Possible replication problems could be determined in ES cells by supertransfection of inverted repeat containing vectors into E14/T and preparing DNA at various days of selection for digestion with DpnI as well as maintaining an uncut control. Southern analysis of these DNAs can then be performed to

determine copy number as any unreplicated DNA will be eliminated by DpnI digestion.

6.6. Summary

I conclude that RNAi operates in ES cells using 21 nt siRNAs against a luciferase target. Reduced expression of an exogenous target, eGFP was also observed using an eGFP specific long hairpin dsRNA. However, this does not appear to be due to RNAi as cells with reduced fluorescence are dying as determined by Annexin V-PE staining. Cell death or apoptosis is most probably due to an interferon response activated by long dsRNA. The episomal expression system used efficiently expresses dsRNA hairpin loop structures, as determined by probing for transcripts from a number of different inverted repeat vectors. Expression was highest from the eGFP inverted repeat vector, which may account for the reduced eGFP expression specifically observed with this vector. Phenotypic ablation of an endogenous target of ES cells, the LIFR was not observed and a reduction of targeted messages of either eGFP or SOCS3 was not detected. This was despite efficient transcription of inverted repeat transcripts and cleavage of the dsRNA transcripts.

Chapter 7: CONCLUSIONS

The aim of this thesis was to develop a rapid and reliable ablation expression system in ES cells. The three RNA ablation strategies investigated were antisense RNA, ribozymes and dsRNA. Ablation using these various gene silencing strategies was investigated principally by looking for a phenotypic effect and a corresponding reduction in the targeted mRNA. Antisense RNA was directed against three genes that are involved in the self renewal of ES cells, ribozyme against p300, and dsRNA against two endogenous genes, LIFR and SOCS3, and a transgene eGFP. Genes required for ES cell self renewal were expected to produce a differentiated phenotype when propagated in the presence of LIF, while p300 ablation in F9 EC cells had been previously shown to result in the maintenance of self renewing colonies when induced to differentiate with retinoic acid (Kawasaki et al.1998). Phenotypic ablation of targeted genes was not observed using any of the ablation strategies. A corresponding reduction of the targeted mRNAs was also not detected. To determine whether this was due to insufficient expression of the RNA ablation transcripts, transcript levels were determined. For all ablation strategies transcripts were expressed at appreciable levels implying that the lack of phenotypic ablation is probably not due to inefficient expression from the episomal vector system.

A probable explanation for the lack of ablation observed for either antisense or ribozyme is the likelihood of secondary structure formation for both antisense RNA and ribozyme transcripts. As the antisense RNA investigated are quite long, 69-575 bp it is possible that either during or after these sequences are transcribed they form secondary structures, which could reduce or prevent

hybridisation to the complementary mRNA target, which may also possess secondary structure. Transcribed ribozyme sequences on the other hand have to be able to be able to form their characteristic functional secondary structure consisting of a looped catalytic domain and two binding arms that are complementary to the mRNA to be targeted. If the transcribed ribozyme sequence does not form this structure the ribozyme will be non functional. Even though the p300 ribozyme has been previously shown to work in EC cells, the intracellular conditions and the non specific proteins present may be different compared to ES cells. In addition, for both ablation strategies the accessibility of the targeted message is important, where hybridisation sites could be occluded by either secondary structure or the binding of non-specific proteins.

The antisense ablation strategy presented in this thesis specifically targeted the UTRs of genes. However, it seems unlikely that UTR sequences in general are poor ablation targets as there are examples in the literature of both effective (Pepin et al.1992; Moxham et al.1993) as well as ineffective ablation (Munir et al.1990) when targeting these sequences.

An alternative explanation for the lack of ablation observed for the ribozyme strategy is that the role of p300 in ES cells remains undetermined so the phenotypic screen set up may have been inappropriate. Also there may be redundancy with CBP, as p300 and CBP are highly homologous (Arany et al.1994) and bind to similar target proteins. Although only p300 and not CBP was found to be required for retinoic acid induced differentiation of F9 EC cells (Kawasaki et al.1998) this may not be the situation in ES cells.

Although phenotypic ablation was not detected for any of the mRNA ablation strategies, reduction of eGFP expression from an integrated transgene was observed following transfection with an eGFP specific inverted repeat expression vector. A similar reduction of eGFP expression was reported for undifferentiated ES cells by in situ production of dsRNA from the transient transfection of a plasmid with an eGFP specific inverted repeat (Yang et al.2001). However, the FACS profile and Annexin V staining data of cells with reduced fluorescence presented in this thesis suggests that the reduction of eGFP expression may not be due to RNAi. Annexin V is an apoptotic marker which demonstrated that the cells were dying probably by an apoptotic mechanism. So the reduced eGFP expression seen is most likely a secondary consequence of the cells dying. This data has not been presented previously. Published papers that have shown RNAi induced ablation in ES cells of either an eGFP target (Yang et al.2001; Billy et al.2001) or luciferase targets (Paddison et al.2002) have not commented upon whether cells are dying or if there is more cell death observed for dsRNA transfected cells.

A second result to note from the transfection of inverted repeat containing vectors as the method of introducing dsRNA into ES cells was that the transfection efficiency of inverted repeat containing vectors compared to non inverted repeat control vectors was always reduced. Although reduced transfection efficiencies were not dependent upon the presence of a target. A further indication that the reduction of eGFP expression observed may not be due to RNAi. Transfection efficiencies were reduced by ~20 fold for eGFP specific inverted repeat vectors and by 2-13 fold for other inverted repeat containing

vectors. This suggested that expression of inverted repeat sequences are either toxic or detrimental for cell growth, which may be linked to the increase in cell death observed for eGFP inverted repeat vectors. Alternatively the replication of inverted repeat vectors or the transcription of inverted repeat sequences may be less efficient as has been observed for the replication of direct repeat sequences in bacteria (LaDuca et al.1983). Lower stable transfection efficiencies observed might also be the result of integration of these vectors. However, the frequency of isolating stable integrants was 3-8 fold reduced compared to supertransfection which may indicate that the inverted repeat vectors have not integrated. To demonstrate this conclusively copy numbers of IR containing vectors will need to be scored.

Although inverted repeat sequences may not form dsRNA once transcribed this is probably not the reason for the lack of ablation observed when investigating dsRNA as an ablation strategy in this thesis. This is unlikely because cleavage products were detected from the transcribed inverted repeat sequences (figure 6.10.), indicating that a dsRNA hairpin loop structure has probably formed. Ablation is most probably not detected because significant expression of long dsRNA results in cell death (section 6.4.3.2.). Cell death was specific to cells transfected with an eGFP inverted repeat vector, which is probably due to the higher expression level of this transcript compared to other tested inverted repeat containing vectors (figure 6.10.). The observed cell death is consistent however, with the emerging evidence that long dsRNAs are capable of inducing apoptosis in mammalian cells (reviewed by Tan and Katze 1999). Evidence that dsRNA triggered apoptosis is mediated by PKR is demonstrated by genetic studies with

PKR knockout mice, whose embryo fibroblasts are resistant to apoptosis induced by dsRNA (Der et al.1997). In addition PKR is already known to be activated as part of the interferon response. At present PKR remains unidentified in ES cells however, it has been cloned in the mouse (Tanaka and Samuel 1994). If indeed it is expressed in ES cells PKR could be reducing eGFP expression in two ways, by inhibiting protein synthesis and by actively inducing apoptosis.

A possible route therefore for initiating RNAi in ES cells could be the additional transfection of dsRNA to mouse PKR which could prevent a non specific response. However, a simpler strategy would be the use of siRNAs rather than long dsRNA. siRNAs have recently been used successfully to initiate effective RNA interference in mammalian cells including ES cells as demonstrated in both data presented in this thesis and in published data (Yang et al.2001; Billy et al.2001; Paddison et al.2002). siRNAs are more effective than dsRNA at inducing a specific RNAi response because dsRNA of 49 nts trigger an interferon response that is sequence non-specific (Elbashir et al.2001b). This is demonstrated by the activation of both PKR and 2',5'-oligoadenylate synthetase by dsRNA longer than 35 nts (Minks et al.1979). Maximal activation measured in terms of phosphorylation for PKR and the synthesis of 2',5'-oligoadenylate synthetase by dsRNA was observed with 80 nts and 65 nts respectively (Minks et al.1979). Manche et al.(1992) found similar results for PKR where 23 and 34 bp dsRNAs resulted in little activation of the enzyme. The efficiency of activation increased however with increasing dsRNA length, reaching a maximum at ~85 bp.

The usefulness of siRNAs is now being realised with the development of vector based expression systems for the expression of siRNAs in cells (Miyagishi

and Taira 2002; Lee et al.2002; Paul et al.2002). siRNA expression is driven by U6, which is a polymerase III promoter that does not contain internal promoter sequences and has a short termination signal of about four uridines. This results in 4 nt 3' overhangs which make the siRNA more potent than corresponding blunt ended siRNAs (Elbashir et al.2001a). siRNAs have been generated by either expressing sense and antisense sequences from separate U6 promoters on the same or two separate plasmids or by including a short tetraloop sequence between the sense and antisense strands so that a single transcript is generated. A possible way therefore of producing a randomised siRNA library for a functional library screen could be to transcribe random 19 nt sequences between U6 promoters placed in opposing orientations. The randomised vector library could then be transfected into ES cells and the appropriate resistance marker present on the vector selected for (figure 7.1). Both the sense and antisense strands of the random library would be transcribed allowing them to hybridise and form siRNAs. The effect of these siRNAs could then be assayed in an appropriate functional screen. For example by growing transfected cells in media without LIF to screen for genes involved in the differentiation of ES cell. To identify the sequence of the effective siRNA(s) individual self renewing colonies can be picked and grown from which the stably maintained episomes are isolated and sequenced. However, the potential strategy for screening novel genes by expressing a randomised siRNA library is not trivial. The efficiency of siRNA ablation must first be determined using such a U6 promoter driven siRNA expression system. Another consideration is that individual siRNA sequences may target a number of mRNAs if the sequence is complementary to a conserved

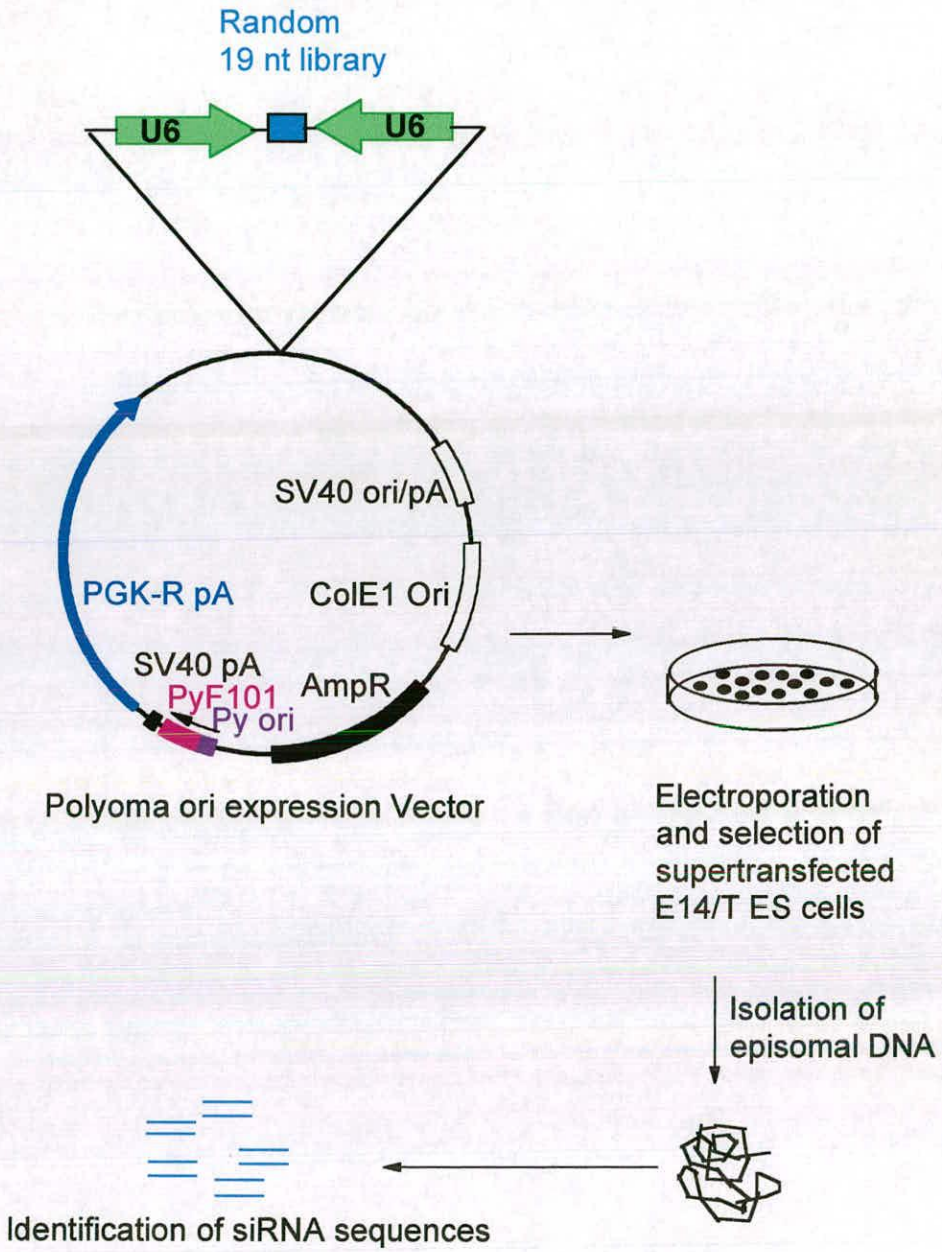
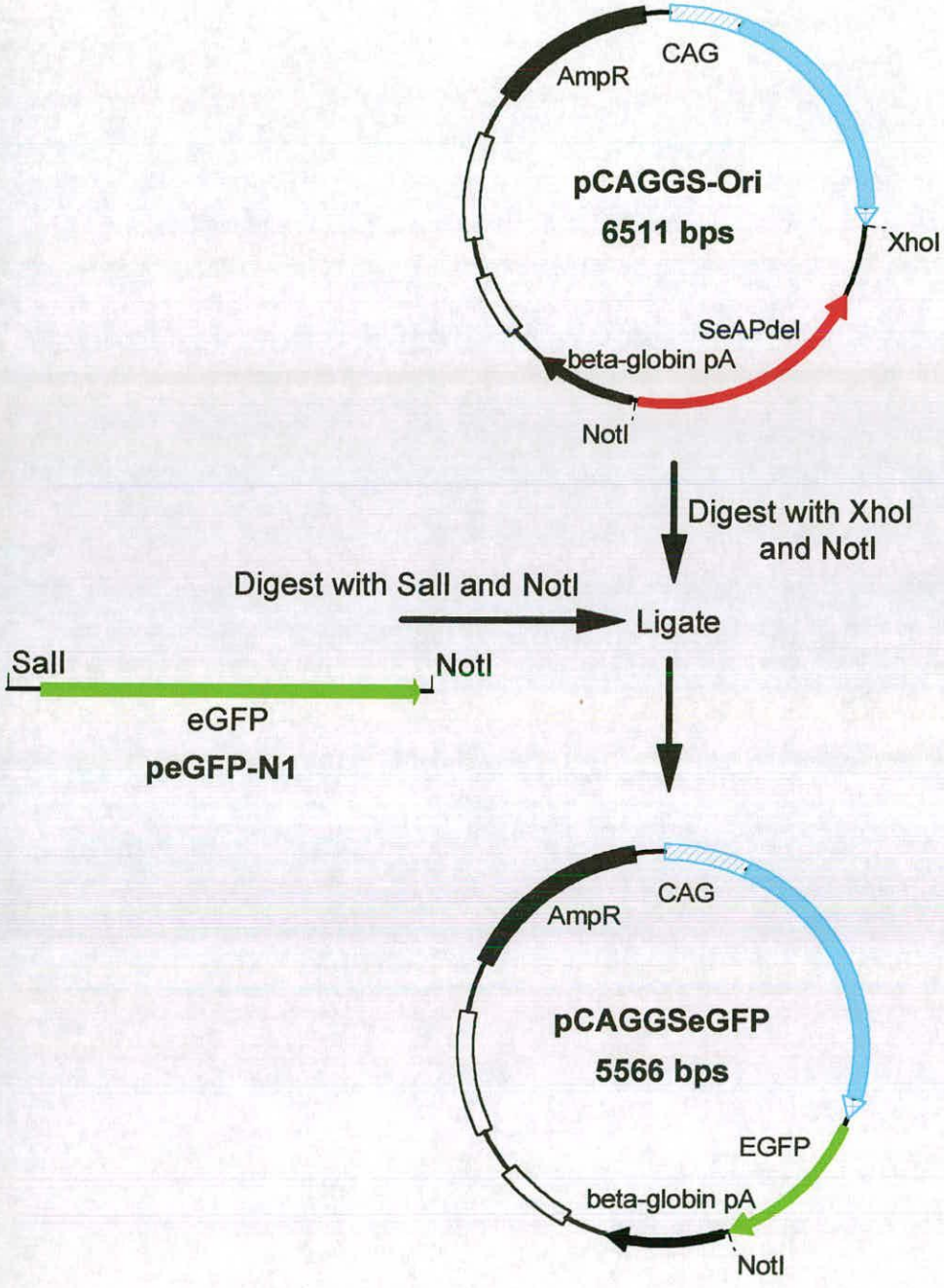


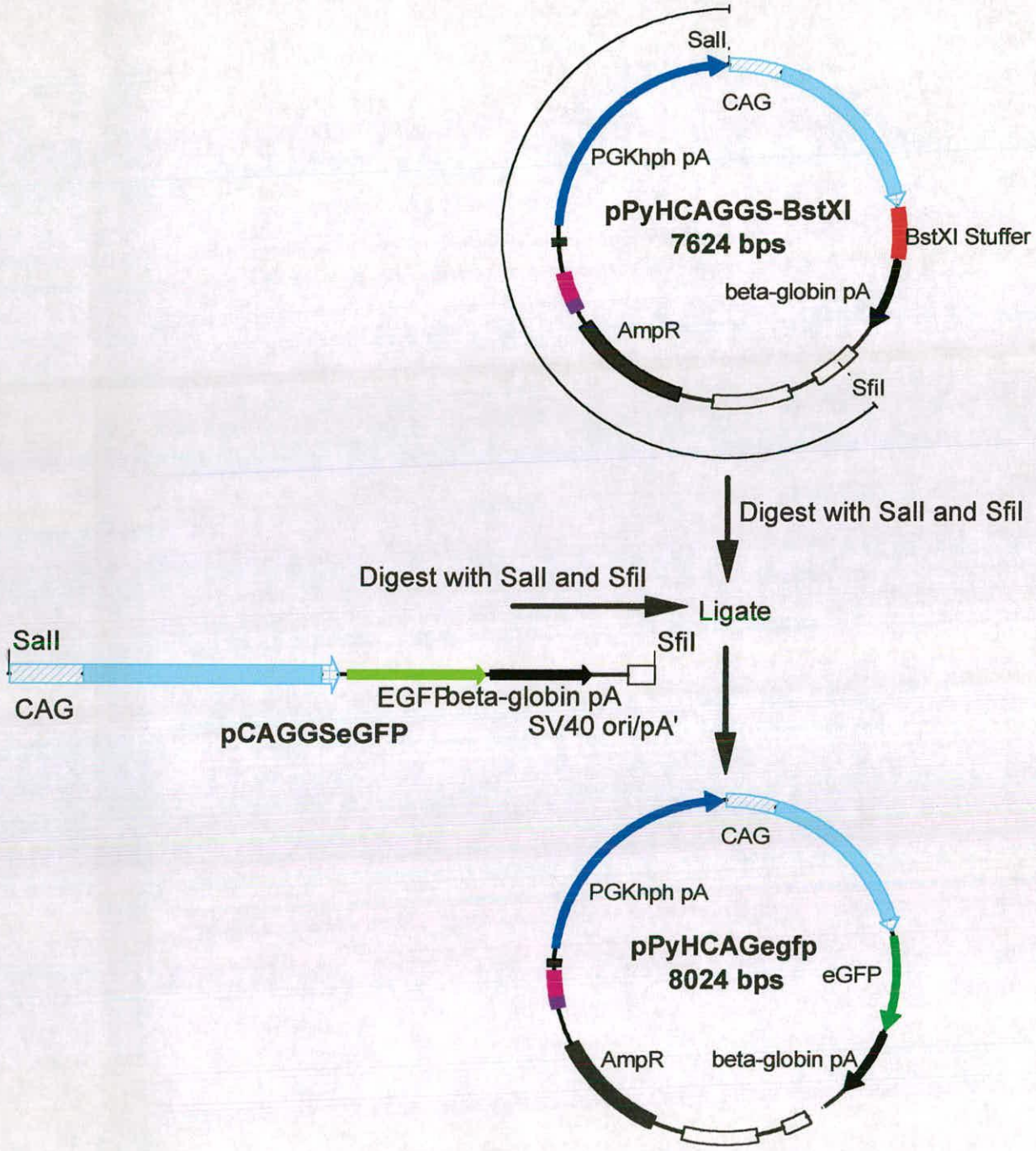
Figure 7.1. Strategy for novel gene function in ES cells using a randomised siRNA library generated by opposing U6 promoters.

domain. This may make identification of the ablated targets responsible for the phenotype difficult. Identification of the targeted mRNA(s) could be further complicated by the generation of secondary siRNAs complementary to regions upstream of the original trigger siRNA by an RdRP reaction and the induction of transitive RNAi. Although an RdRP homolog has yet to be identified in the mammalian or *Drosophila* genome an RdRP-like activity has been demonstrated in *Drosophila* embryo extracts (Lipardi et al.2001) which does not rule out the possibility that mammalian cells may also be able to generate secondary siRNAs. Therefore the ablation efficiency of siRNA as well as secondary ablation effects must first be investigated before the potential of siRNAs as a tool to uncover novel gene function in ES cells as well as other mammalian cells can be exploited.

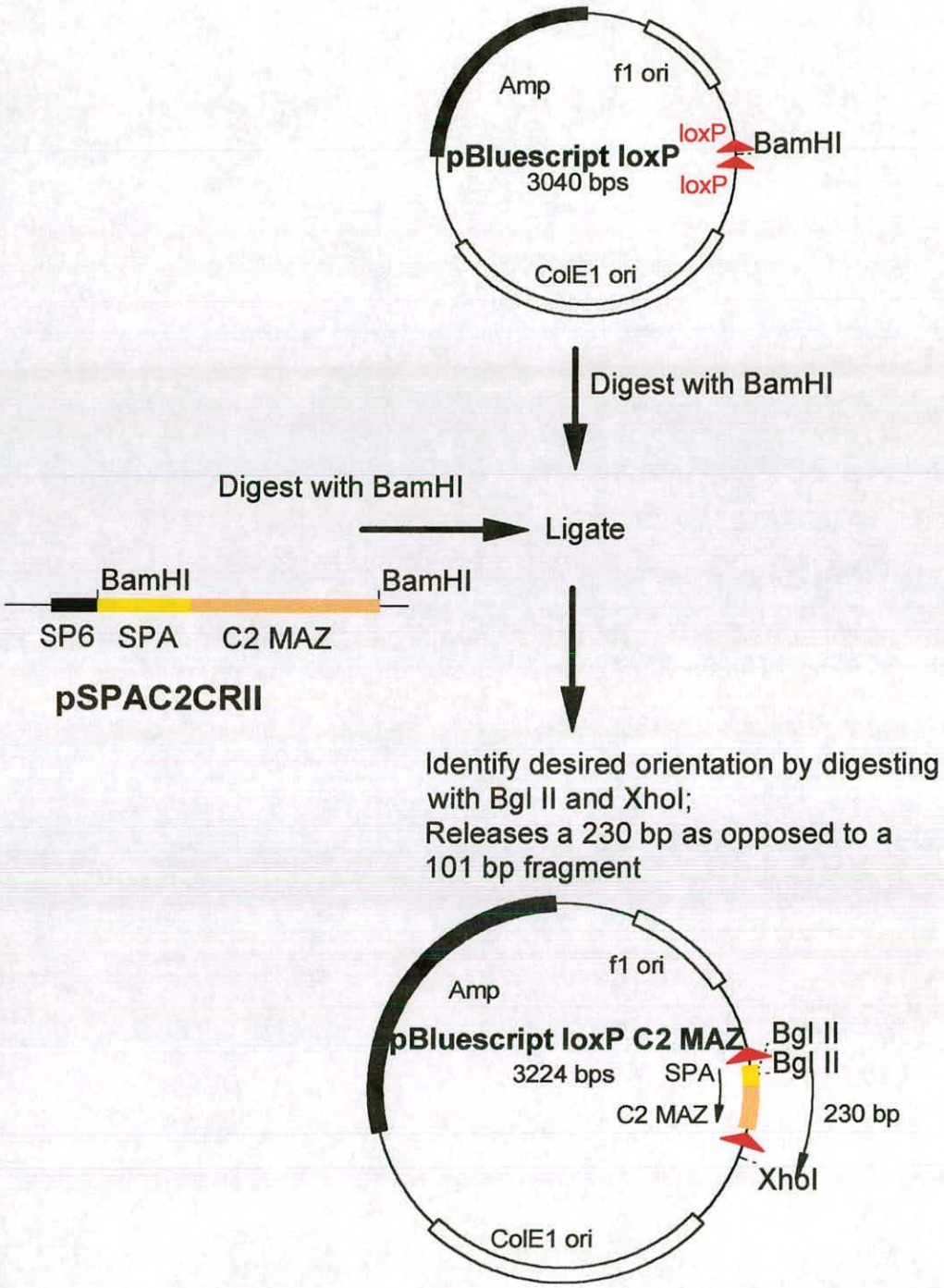
Appendix 1
Cloning strategy for pCAGGSeGFP



Appendix 2
Cloning strategy for pPyHCAGeGFP

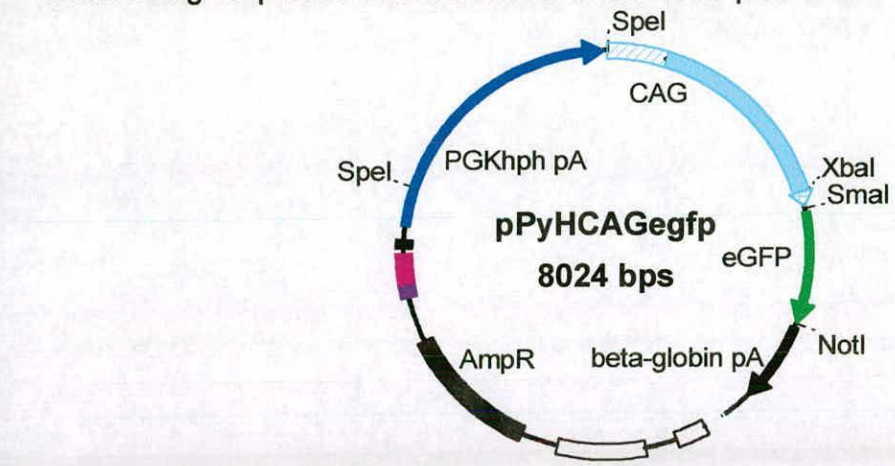


Appendix 3:
Cloning Strategy for pBluescript loxP C2 MAZ



Appendix 4

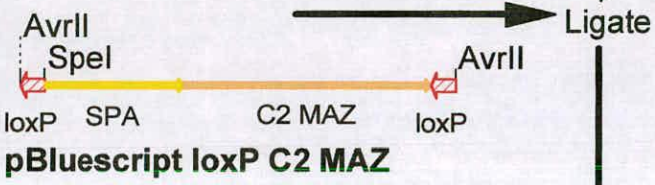
Cloning strategy for pFlox-C2MAZ-Intron. Cloning the transcriptional terminating sequence into the intron of the CAG promoter.



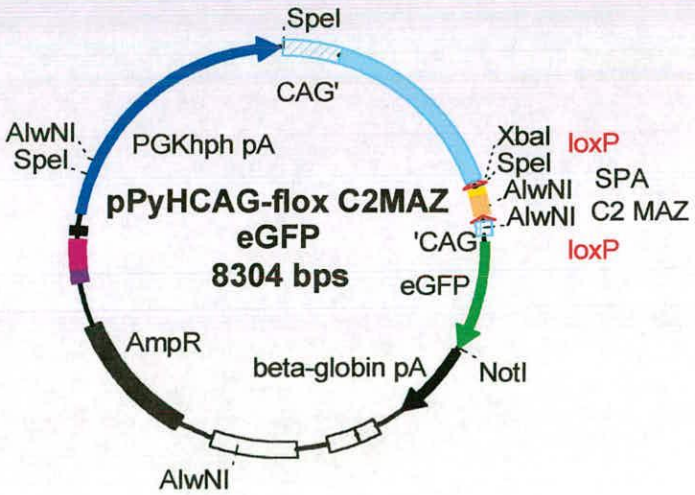
Digest with XbaI

Digest with AvrII

Ligate

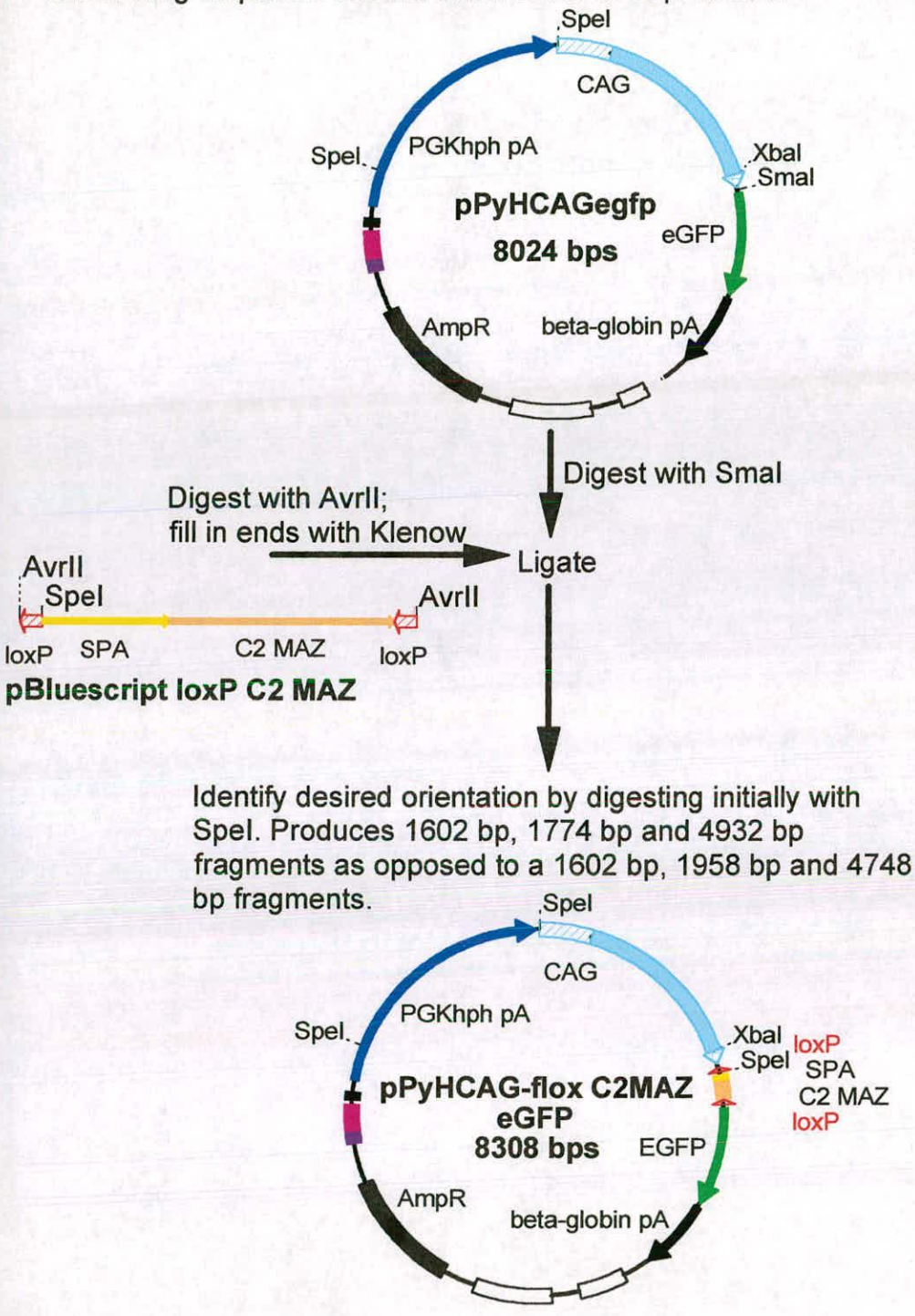


Identify desired orientation by digesting initially with SpeI.Produces 1602 bp, 1653 bp and 5049 bp fragments as opposed to a 1602 bp, 1861 bp and 4841 bp fragments.



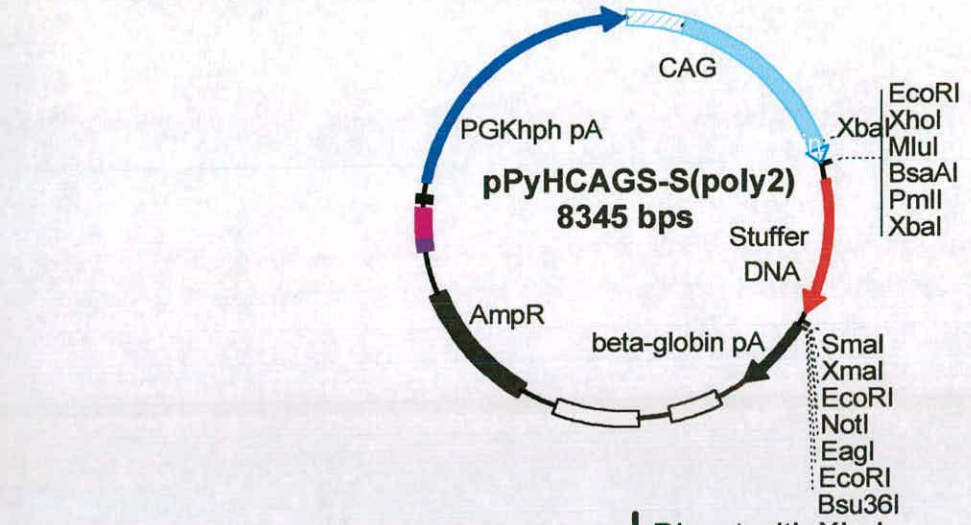
Appendix 5

Cloning strategy for pFlox-C2MAZ-Exon. Cloning the transcriptional terminating sequence into the exon of the CAG promoter.



Appendix 6:

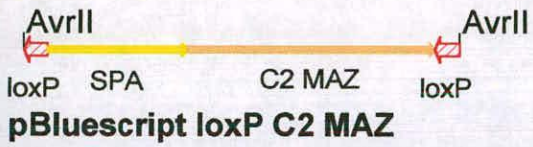
a) Cloning strategy for pPyHCAGS-flox C2MAZ-S(poly2);
b) and subsequent pPyHCAG-S flox C2MAZ 5'/3' UTR vectors;
c) and pPyHCAG-S 5'/3' UTR vectors.



Digest with XhoI;
fill in ends with dNTPs
using Klenow fragment
of DNA polymerase

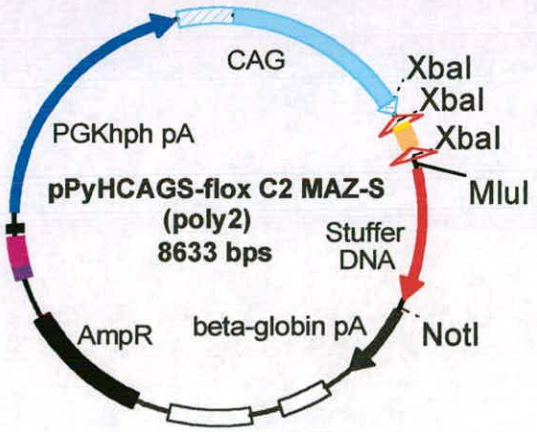
Digest with AvrII;
fill in ends with Klenow

Ligate

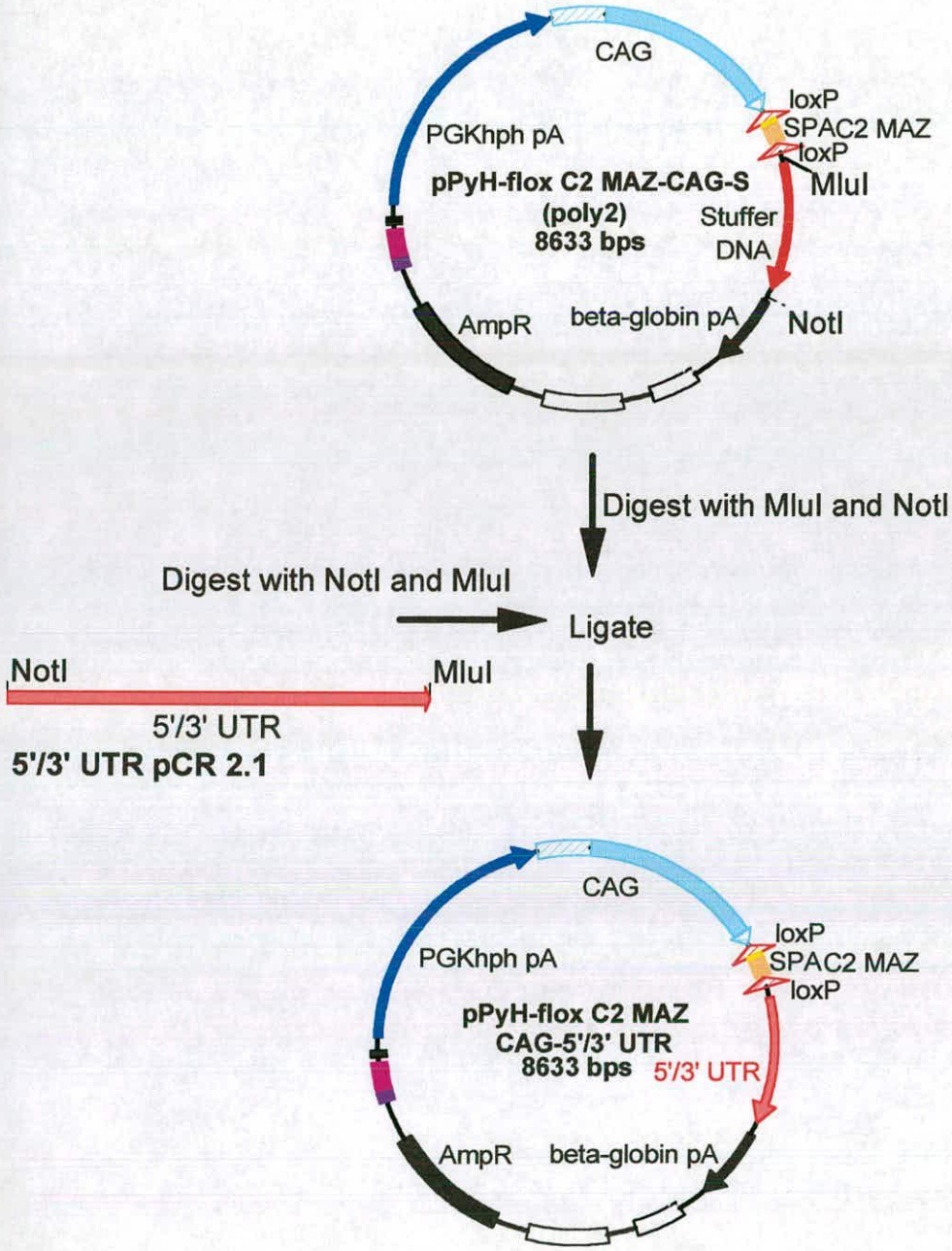


pBluescript loxP C2 MAZ

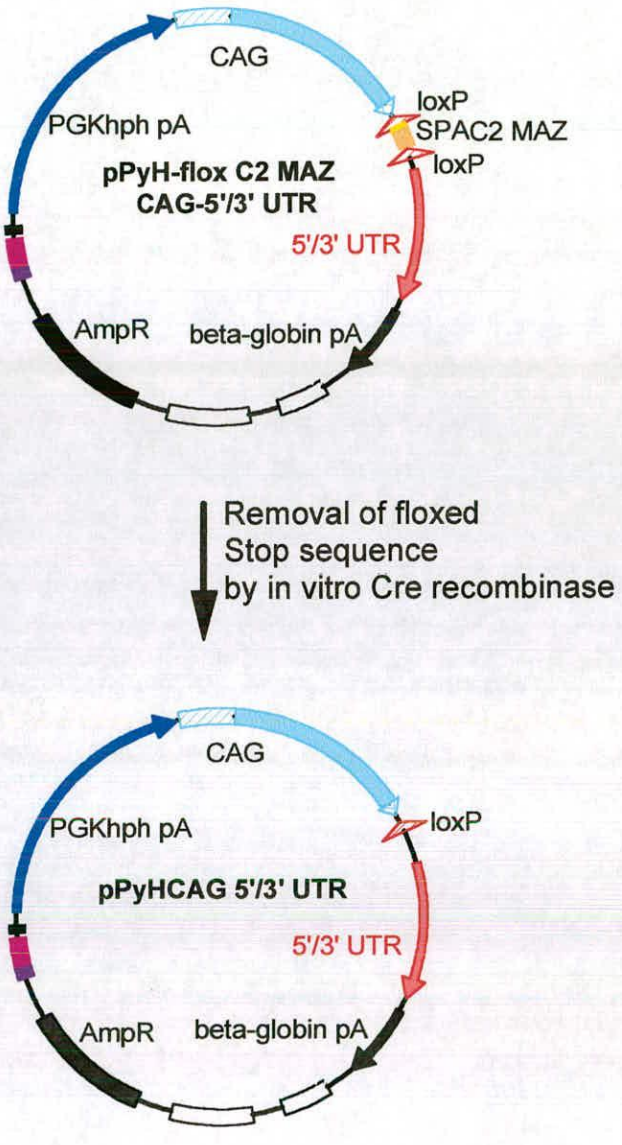
Identify desired orientation by digesting with XbaI.
Releases 148 bp and 260 bp fragments as opposed
to a 344 bp and 64 bp fragments



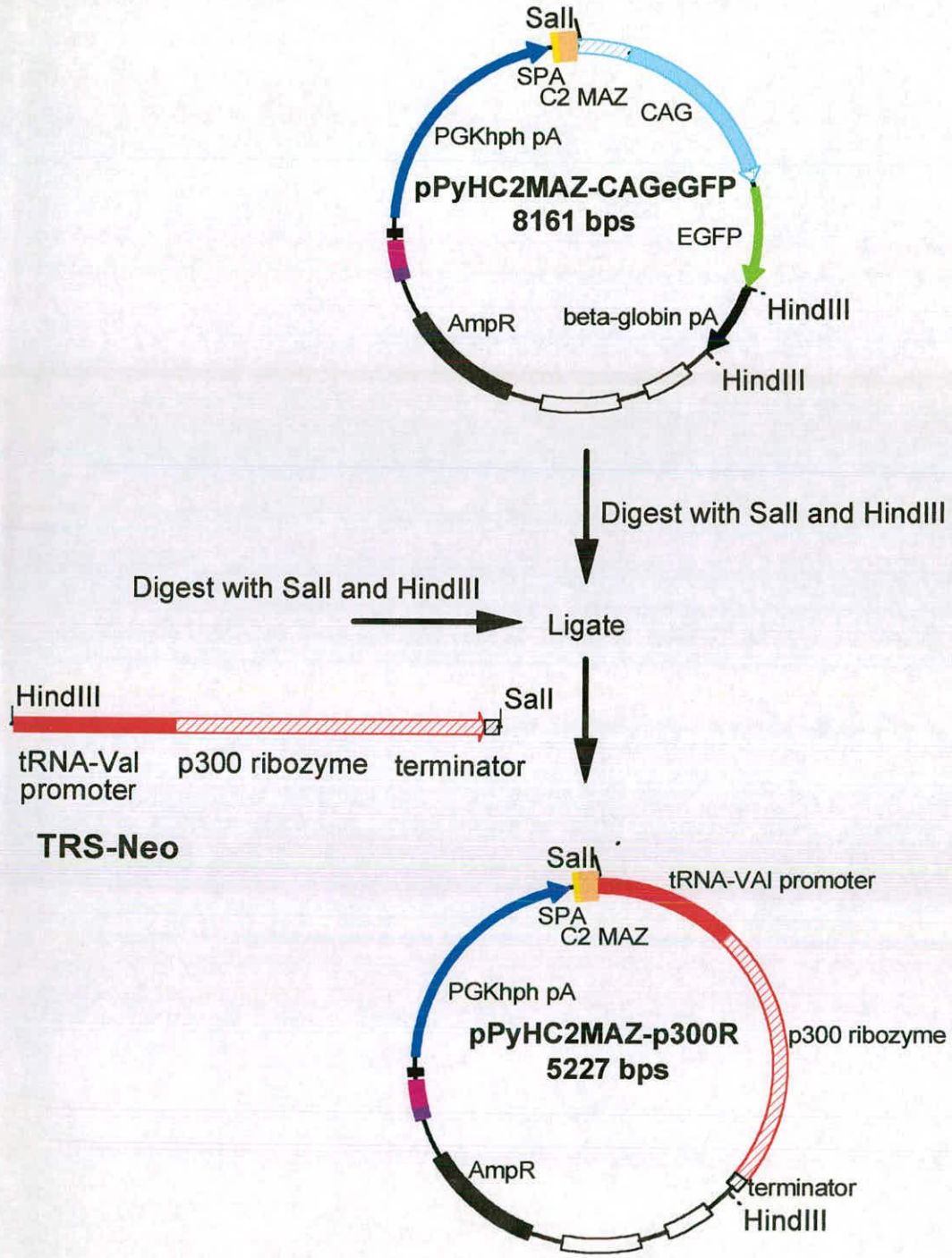
Appendix 6:
b) Cloning Strategy for pPyHCAG-S flox C2MAZ 5'/3' UTR vectors



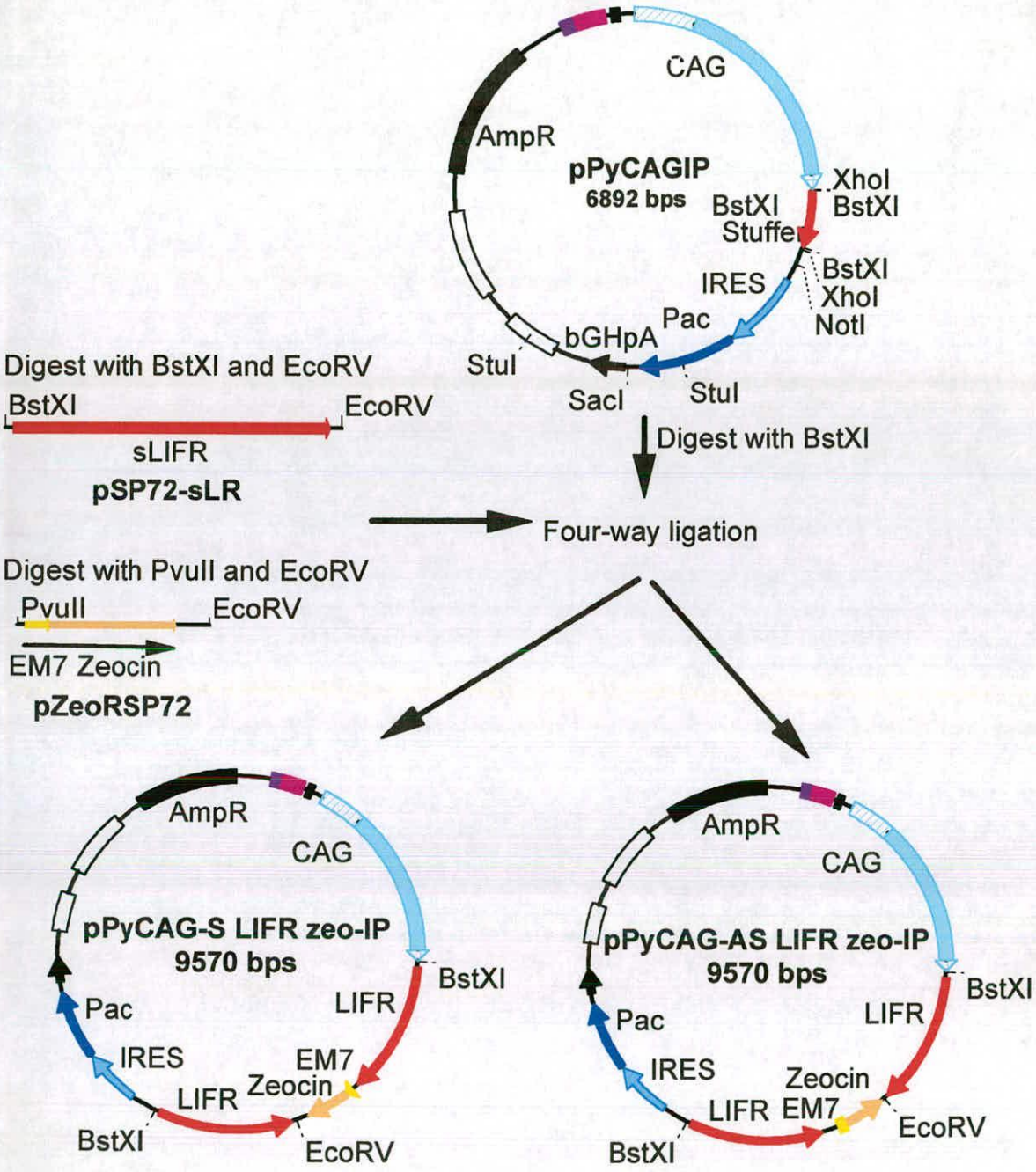
Appendix 6:
c) Cloning Strategy for pPyHCAG 5'/3' UTR vectors



Appendix 7:
Cloning strategy for pPyHC2MAZ-p300R



Appendix 8:
Cloning Strategy for pPyCAG-S/AS LIFR zeo-IP



List of Figures

List of Tables

List of Abbreviations

bp	base pairs
CAG	cytomegalovirus immediate early nhancer, actin promoter, globin intron
DHFR	dihydrofolate reductase
DMSO	dimethylsulphoxide
DsRNA	double stranded RNA
EDTA	ethylenediaminetetracetic acid
EGFP	enhanced green fluorescent protein
EIF	elongation initiation factor
ERK	extracellular regulated kinase
ES cell	embryonic stem cell
ICM	inner cell mass
IL-6	interleukin 6
IR	inverted repeat
IRES	internal ribosomal entry site
JAK	Janus kinase
Kb	kilobase
LIF	Leukaemia inhibitory factor
MAPK	mitogen activated protein kinase
MOPS	3-N-Morpholino propanesulfonic acid
Nt	nucleotide
ODN	oligodeoxyribonucleotides
PE	phycoerythrin
PGK	phosphoglycerate kinase
Pp-luc	*Photinus pyralis* luciferase
PS	phosphatidylserine
RA	retinoic acid
RdRP	RNA dependent RNA polymerase
RISC	RNA induced silencing complex
RNAi	RNA interference
Rr-luc	*Renilla reniformis* luciferase
SDS	sodium dodecyl sulphate
SH2	src homology 2
sIL-6R	soluble interleukin 6 receptor
siRNA	short interfering RNA
SOCS3	suppressor of cytokine signalling
SSC	standard saline citrate
STAT3	signal transducer and activator of transcription
TK	thymidine kinase
UTR	untranslated region

REFERENCES

Akira, S., Nishio, Y., Inoue, M., Wang, X. J., Wei, S., Matsusaka, T., Yoshida, K., Sudo, T., Naruto, M., and Kishimoto, T. (1994). Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway. *Cell* **77,** 63-71.

Altman, S. (1989). Ribonuclease P: an enzyme with a catalytic RNA subunit. *Adv Enzymol Relat Areas Mol Biol* **62,** 1-36.

Arany, Z., Newsome, D., Oldread, E., Livingston, D. M., and Eckner, R. (1995). A family of transcriptional adaptor proteins targeted by the E1A oncoprotein. *Nature* **374,** 81-4.

Arany, Z., Sellers, W. R., Livingston, D. M., and Eckner, R. (1994). E1A-associated p300 and CREB-associated CBP belong to a conserved family of coactivators. *Cell* **77,** 799-800.

Argetsinger, L. S., Hsu, G. W., Myers, M. G., Jr., Billestrup, N., White, M. F., and Carter-Su, C. (1995). Growth hormone, interferon-gamma, and leukemia inhibitory factor promoted tyrosyl phosphorylation of insulin receptor substrate-1. *J Biol Chem* **270,** 14685-92.

Arndt, G. M., and Rank, G. H. (1997). Colocalization of antisense RNAs and ribozymes with their target mRNAs. *Genome* **40,** 785-97.

Ashfield, R., Enriquez-Harris, P., and Proudfoot, N. J. (1991). Transcriptional termination between the closely linked human complement genes C2 and factor B: common termination factor for C2 and c-myc? *Embo J* **10,** 4197-207.

Ashfield, R., Patel, A. J., Bossone, S. A., Brown, H., Campbell, R. D., Marcu, K. B., and Proudfoot, N. J. (1994). MAZ-dependent termination between closely spaced human complement genes. *Embo J* **13,** 5656-67.

Asselin, C., Gelinas, C., Branton, P. E., and Bastin, M. (1984). Polyoma middle T antigen requires cooperation from another gene to express the malignant phenotype in vivo. *Mol Cell Biol* **4,** 755-60.

Bain, G., Kitchens, D., Yao, M., Huettner, J. E., and Gottlieb, D. I. (1995). Embryonic stem cells express neuronal properties in vitro. *Dev Biol* **168,** 342-57.

Balachandran, S., Kim, C. N., Yeh, W. C., Mak, T. W., Bhalla, K., and Barber, G. N. (1998). Activation of the dsRNA-dependent protein kinase, PKR, induces apoptosis through FADD-mediated death signaling. *Embo J* **17,** 6888-902.

Bannister, A. J., and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature* **384,** 641-3.

Bass, B. L. (2000). Double-stranded RNA as a template for gene silencing. *Cell* **101,** 235-8.

Bazan, J. F. (1990). Structural design and molecular evolution of a cytokine receptor superfamily. *Proc Natl Acad Sci U S A* **87,** 6934-8.

Bedinger, P., Munn, M., and Alberts, B. M. (1989). Sequence-specific pausing during in vitro DNA replication on double-stranded DNA templates. *J Biol Chem* **264,** 16880-6.

Berger, H., and Wintersberger, E. (1986). Polyomavirus small T antigen enhances replication of viral genomes in 3T6 mouse fibroblasts. *J Virol* **60,** 768-70.

Berget, S. M. (1995). Exon recognition in vertebrate splicing. *J Biol Chem* **270,** 2411-4.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363-6.

Billy, E., Brondani, V., Zhang, H., Muller, U., and Filipowicz, W. (2001). Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc Natl Acad Sci U S A* **98,** 14428-33.

Bousquet, C., Susini, C., and Melmed, S. (1999). Inhibitory roles for SHP-1 and SOCS-3 following pituitary proopiomelanocortin induction by leukemia inhibitory factor. *J Clin Invest* **104,** 1277-85.

Braat, A. K., van de Water, S., Korving, J., and Zivkovic, D. (2001). A zebrafish vasa morphant abolishes vasa protein but does not affect the establishment of the germline. *Genesis* **30,** 183-5.

Bradley, A., Evans, M., Kaufman, M. H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309,** 255-6.

Brook, F. A., and Gardner, R. L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. *Proc Natl Acad Sci U S A* **94,** 5709-12.

Castelli, J., Wood, K. A., and Youle, R. J. (1998). The 2-5A system in viral infection and apoptosis. *Biomed Pharmacother* **52,** 386-90.

Castelli, J. C., Hassel, B. A., Wood, K. A., Li, X. L., Amemiya, K., Dalakas, M. C., Torrence, P. F., and Youle, R. J. (1997). A study of the interferon

antiviral mechanism: apoptosis activation by the 2-5A system. *J Exp Med* **186,** 967-72.

Catalanotto, C., Azzalin, G., Macino, G., and Cogoni, C. (2000). Gene silencing in worms and fungi. *Nature* **404,** 245.

Cech, T. R. (1990). Self-splicing of group I introns. *Annu Rev Biochem* **59,** 543-68.

Chambers, I., Cozens, A., Broadbent, J., Robertson, M., Lee, M., Li, M., and Smith, A. (1997). Structure of the mouse leukaemia inhibitory factor receptor gene: regulated expression of mRNA encoding a soluble receptor isoform from an alternative 5' untranslated region. *Biochem J* **328,** 879-88.

Chiang, M. Y., Chan, H., Zounes, M. A., Freier, S. M., Lima, W. F., and Bennett, C. F. (1991). Antisense oligonucleotides inhibit intercellular adhesion molecule 1 expression by two distinct mechanisms. *J Biol Chem* **266,** 18162-71.

Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A., and Dixon, J. E. (2000). Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. *Proc Natl Acad Sci U S A* **97,** 6499-503.

Cogoni, C., Irelan, J. T., Schumacher, M., Schmidhauser, T. J., Selker, E. U., and Macino, G. (1996). Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *Embo J* **15,** 3153-63.

Cogoni, C., and Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. *Nature* **399,** 166-9.

Colman, A. (1990). Antisense strategies in cell and developmental biology. *J Cell Sci* **97,** 399-409.

Conover, J. C., Ip, N. Y., Poueymirou, W. T., Bates, B., Goldfarb, M. P., DeChiara, T. M., and Yancopoulos, G. D. (1993). Ciliary neurotrophic factor maintains the pluripotentiality of embryonic stem cells. *Development* **119,** 559-65.

Cooke, C., Hans, H., and Alwine, J. C. (1999). Utilization of splicing elements and polyadenylation signal elements in the coupling of polyadenylation and last-intron removal. *Mol Cell Biol* **19,** 4971-9.

Crooke, S. T. (1992). Therapeutic applications of oligonucleotides. *Annu Rev Pharmacol Toxicol* **32,** 329-76.

Cullen, B. R., Lomedico, P. T., and Ju, G. (1984). Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis. *Nature* **307,** 241-5.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D. C. (2000). An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101,** 543-53.

Darnell, J. E., Jr. (1997). STATs and gene regulation. *Science* **277,** 1630-5.

Darnell, J. E., Jr., Kerr, I. M., and Stark, G. R. (1994). Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264,** 1415-21.

Der, S. D., Yang, Y. L., Weissmann, C., and Williams, B. R. (1997). A double-stranded RNA-activated protein kinase-dependent pathway mediating stress-induced apoptosis. *Proc Natl Acad Sci U S A* **94,** 3279-83.

Eckner, R., Ewen, M. E., Newsome, D., Gerdes, M., DeCaprio, J. A., Lawrence, J. B., and Livingston, D. M. (1994). Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev* **8,** 869-84.

Eggermont, J., and Proudfoot, N. J. (1993). Poly(A) signals and transcriptional pause sites combine to prevent interference between RNA polymerase II promoters. *Embo J* **12,** 2539-48.

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

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001a). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15,** 188-200.

Erickson, R. P. (1999). Antisense transgenics in animals. *Methods* **18,** 304-10.

Ernst, M., Gearing, D. P., and Dunn, A. R. (1994). Functional and biochemical association of Hck with the LIF/IL-6 receptor signal transducing subunit gp130 in embryonic stem cells. *Embo J* **13,** 1574-84.

Ernst, M., Novak, U., Nicholson, S. E., Layton, J. E., and Dunn, A. R. (1999). The carboxyl-terminal domains of gp130-related cytokine receptors are necessary for suppressing embryonic stem cell differentiation. Involvement of STAT3. *J Biol Chem* **274,** 9729-37.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of

pluripotential cells from mouse embryos. *Nature* **292,** 154-6.

Fagard, M., Boutet, S., Morel, J. B., Bellini, C., and Vaucheret, H. (2000). 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* **97,** 11650-4.

Feil, R., Brocard, J., Mascrez, B., LeMeur, M., Metzger, D., and Chambon, P. (1996). Ligand-activated site-specific recombination in mice. *Proc Natl Acad Sci U S A* **93,** 10887-90.

Feil, R., Wagner, J., Metzger, D., and Chambon, P. (1997). Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem Biophys Res Commun* **237,** 752-7.

Feldstein, P. A., Buzayan, J. M., and Bruening, G. (1989). Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. *Gene* **82,** 53-61.

Filippov, V., Solovyev, V., Filippova, M., and Gill, S. S. (2000). A novel type of RNase III family proteins in eukaryotes. *Gene* **245,** 213-21.

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

Floyd-Smith, G., Slattery, E., and Lengyel, P. (1981). Interferon action: RNA cleavage pattern of a (2'-5')oligoadenylate--dependent endonuclease. *Science* **212,** 1030-2.

Forster, A. C., and Symons, R. H. (1987). Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell* **49,** 211-20.

Francke, B., and Eckhart, W. (1973). Polyoma gene function required for viral DNA synthesis. *Virology* **55,** 127-35.

Fuhrer, D. K., Feng, G. S., and Yang, Y. C. (1995). Syp associates with gp130 and Janus kinase 2 in response to interleukin-11 in 3T3-L1 mouse preadipocytes. *J Biol Chem* **270,** 24826-30.

Fujimura, F. K., Deininger, P. L., Friedmann, T., and Linney, E. (1981). Mutation near the polyoma DNA replication origin permits productive infection of F9 embryonal carcinoma cells. *Cell* **23,** 809-14.

Gassmann, M., Donoho, G., and Berg, P. (1995). Maintenance of an extrachromosomal plasmid vector in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* **92,** 1292-6.

Gearing, D. P., Gough, N. M., King, J. A., Hilton, D. J., Nicola, N. A.,

Simpson, R. J., Nice, E. C., Kelso, A., and Metcalf, D. (1987). Molecular cloning and expression of cDNA encoding a murine myeloid leukaemia inhibitory factor (LIF). *Embo J* **6,** 3995-4002.

Greger, I. H., Demarchi, F., Giacca, M., and Proudfoot, N. J. (1998). Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res* **26,** 1294-301.

Grishok, A., Tabara, H., and Mello, C. C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. *Science* **287,** 2494-7.

Gu, H., Zou, Y. R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73,** 1155-64.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35,** 849-57.

Guo, S., and Kemphues, K. J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81,** 611-20.

Hagan, C. E., and Warren, G. J. (1983). Viability of palindromic DNA is restored by deletions occurring at low but variable frequency in plasmids of Escherichia coli. *Gene* **24,** 317-26.

Hamilton, A. J., and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286,** 950-2.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404,** 293-6.

Haseloff, J., and Gerlach, W. L. (1988). Simple RNA enzymes with new and highly specific endoribonuclease activities. *Nature* **334,** 585-91.

Heinrich, P. C., Behrmann, I., Muller-Newen, G., Schaper, F., and Graeve, L. (1998). Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem J* **334,** 297-314.

Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D., Sharp, P. A., Ingraham, H. A., Rosenfeld, M. G., Finney, M., Ruvkun, G., and et al. (1988). The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. *Genes Dev* **2,** 1513-6.

Herschlag, D. (1991). Implications of ribozyme kinetics for targeting the

cleavage of specific RNA molecules in vivo: more isn't always better. *Proc Natl Acad Sci U S A* **88,** 6921-5.

Ho, S. P., Bao, Y., Lesher, T., Malhotra, R., Ma, L. Y., Fluharty, S. J., and Sakai, R. R. (1998). Mapping of RNA accessible sites for antisense experiments with oligonucleotide libraries. *Nat Biotechnol* **16,** 59-63.

Hocke, G. M., Cui, M. Z., and Fey, G. H. (1995). The LIF response element of the alpha 2 macroglobulin gene confers LIF-induced transcriptional activation in embryonal stem cells. *Cytokine* **7,** 491-502.

Holen, T., Amarzguioui, M., Wiiger, M. T., Babaie, E., and Prydz, H. (2002). Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res* **30,** 1757-66.

Indra, A. K., Warot, X., Brocard, J., Bornert, J. M., Xiao, J. H., Chambon, P., and Metzger, D. (1999). Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res* **27,** 4324-7.

Izant, J. G., and Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* **36,** 1007-15.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78,** 657-68.

Jacobsen, S. E., Running, M. P., and Meyerowitz, E. M. (1999). Disruption of an RNA helicase/RNAse III gene in Arabidopsis causes unregulated cell division in floral meristems. *Development* **126,** 5231-43.

Kawasaki, H., Eckner, R., Yao, T. P., Taira, K., Chiu, R., Livingston, D. M., and Yokoyama, K. K. (1998). Distinct roles of the co-activators p300 and CBP in retinoic-acid-induced F9-cell differentiation. *Nature* **393,** 284-9.

Kawasaki, H., Ohkawa, J., Tanishige, N., Yoshinari, K., Murata, T., Yokoyama, K. K., and Taira, K. (1996). Selection of the best target site for ribozyme-mediated cleavage within a fusion gene for adenovirus E1A-associated 300 kDa protein (p300) and luciferase. *Nucleic Acids Res* **24,** 3010-6.

Kennedy, M., Firpo, M., Choi, K., Wall, C., Robertson, S., Kabrun, N., and Keller, G. (1997). A common precursor for primitive erythropoiesis and definitive haematopoiesis. *Nature* **386,** 488-93.

Kennerdell, J. R., and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in

the wingless pathway. *Cell* **95,** 1017-26.

Kennerdell, J. R., and Carthew, R. W. (2000). Heritable gene silencing in Drosophila using double-stranded RNA. *Nat Biotechnol* **18,** 896-8.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev* **15,** 2654-9.

Ketting, R. F., Haverkamp, T. H., van Luenen, H. G., and Plasterk, R. H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99,** 133-41.

Kim, S. K., and Wold, B. J. (1985). Stable reduction of thymidine kinase activity in cells expressing high levels of anti-sense RNA. *Cell* **42,** 129-38.

Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15,** 1109-23.

Krebs, D. L., and Hilton, D. J. (2000). SOCS: physiological suppressors of cytokine signaling. *J Cell Sci* **113,** 2813-9.

LaDuca, R. J., Fay, P. J., Chuang, C., McHenry, C. S., and Bambara, R. A. (1983). Site-specific pausing of deoxyribonucleic acid synthesis catalyzed by four forms of Escherichia coli DNA polymerase III. *Biochemistry* **22,** 5177-88.

Layton, M. J., Cross, B. A., Metcalf, D., Ward, L. D., Simpson, R. J., and Nicola, N. A. (1992). A major binding protein for leukemia inhibitory factor in normal mouse serum: identification as a soluble form of the cellular receptor. *Proc Natl Acad Sci U S A* **89,** 8616-20.

Leach, D. R. (1996). Cloning and characterization of DNAs with palindromic sequences. *Genet Eng (N Y)* **18,** 1-11.

Lee, N. S., Dohjima, T., Bauer, G., Li, H., Li, M. J., Ehsani, A., Salvaterra, P., and Rossi, J. (2002). Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* **20,** 500-5.

Lee, S. B., and Esteban, M. (1994). The interferon-induced double-stranded RNA-activated protein kinase induces apoptosis. *Virology* **199,** 491-6.

Levitt, N., Briggs, D., Gil, A., and Proudfoot, N. J. (1989). Definition of an efficient synthetic poly(A) site. *Genes Dev* **3,** 1019-25.

Levy-Strumpf, N., and Kimchi, A. (1998). Death associated proteins (DAPs): from gene identification to the analysis of their apoptotic and tumor

suppressive functions. *Oncogene* **17,** 3331-40.

L'Huillier, P. J., Davis, S. R., and Bellamy, A. R. (1992). Cytoplasmic delivery of ribozymes leads to efficient reduction in alpha-lactalbumin mRNA levels in C127I mouse cells. *Embo J* **11,** 4411-8.

Li, M., Pevny, L., Lovell-Badge, R., and Smith, A. (1998). Generation of purified neural precursors from embryonic stem cells by lineage selection. *Curr Biol* **8,** 971-4.

Liebhaber, S. A., Cash, F. E., and Shakin, S. H. (1984). Translationally associated helix-destabilizing activity in rabbit reticulocyte lysate. *J Biol Chem* **259,** 15597-602.

Lill, N. L., Grossman, S. R., Ginsberg, D., DeCaprio, J., and Livingston, D. M. (1997). Binding and modulation of p53 by p300/CBP coactivators. *Nature* **387,** 823-7.

Lipardi, C., Wei, Q., and Paterson, B. M. (2001). RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. *Cell* **107,** 297-307.

Manche, L., Green, S. R., Schmedt, C., and Mathews, M. B. (1992). Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol Cell Biol* **12,** 5238-48.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78,** 7634-8.

Martin, S. J., Reutelingsperger, C. P., McGahon, A. J., Rader, J. A., van Schie, R. C., LaFace, D. M., and Green, D. R. (1995). Early redistribution of plasma membrane phosphatidylserine is a general feature of apoptosis regardless of the initiating stimulus: inhibition by overexpression of Bcl-2 and Abl. *J Exp Med* **182,** 1545-56.

Matsuda, S., Ichigotani, Y., Okuda, T., Irimura, T., Nakatsugawa, S., and Hamaguchi, M. (2000). Molecular cloning and characterization of a novel human gene (HERNA) which encodes a putative RNA-helicase. *Biochim Biophys Acta* **1490,** 163-9.

Matsuda, T., Takahashi-Tezuka, M., Fukada, T., Okuyama, Y., Fujitani, Y., Tsukada, S., Mano, H., Hirai, H., Witte, O. N., and Hirano, T. (1995). Association and activation of Btk and Tec tyrosine kinases by gp130, a signal transducer of the interleukin-6 family of cytokines. *Blood* **85,** 627-33.

Mattioni, T., Louvion, J. F., and Picard, D. (1994). Regulation of protein activities by fusion to steroid binding domains. *Methods Cell Biol* **43 Pt A,** 335-52.

McDevitt, M. A., Imperiale, M. J., Ali, H., and Nevins, J. R. (1984). Requirement of a downstream sequence for generation of a poly(A) addition site. *Cell* **37,** 993-9.

Melton, D. A. (1985). Injected anti-sense RNAs specifically block messenger RNA translation in vivo. *Proc Natl Acad Sci U S A* **82,** 144-8.

Merrick, W. C. (1992). Mechanism and regulation of eukaryotic protein synthesis. *Microbiol Rev* **56,** 291-315.

Miki, N., Hatano, M., Wakita, K., Imoto, S., Nishikawa, S., and Tokuhisa, T. (1992). Role of I-A molecules in early stages of B cell maturation. *J Immunol* **149,** 801-7.

Milner, N., Mir, K. U., and Southern, E. M. (1997). Selecting effective antisense reagents on combinatorial oligonucleotide arrays. *Nat Biotechnol* **15,** 537-41.

Minami, M., Inoue, M., Wei, S., Takeda, K., Matsumoto, M., Kishimoto, T., and Akira, S. (1996). STAT3 activation is a critical step in gp130-mediated terminal differentiation and growth arrest of a myeloid cell line. *Proc Natl Acad Sci U S A* **93,** 3963-6.

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

Mir, K. U., and Southern, E. M. (1999). Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat Biotechnol* **17,** 788-92.

Miyagishi, M., and Taira, K. (2002). U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat Biotechnol* **20,** 497-500.

Monia, B. P., Lesnik, E. A., Gonzalez, C., Lima, W. F., McGee, D., Guinosso, C. J., Kawasaki, A. M., Cook, P. D., and Freier, S. M. (1993). Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *J Biol Chem* **268,** 14514-22.

Montgomery, M. K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **95,** 15502-7.

Moxham, C. M., Hod, Y., and Malbon, C. C. (1993). Induction of G alpha i2-specific antisense RNA in vivo inhibits neonatal growth. *Science* **260,** 991-5.

Mummery, C. L., Feyen, A., Freund, E., and Shen, S. (1990). Characteristics

of embryonic stem cell differentiation: a comparison with two embryonal carcinoma cell lines. *Cell Differ Dev* **30,** 195-206.

Munir, M., Rossiter B.,and Caskey C. (1990). Antisense RNA production in transgenic mice. *Somatic Cell Mol. Genet.* **16,** 383-94.

Munroe, S. H. (1988). Antisense RNA inhibits splicing of pre-mRNA in vitro. *Embo J* **7,** 2523-32.

Murakami, M., Hibi, M., Nakagawa, N., Nakagawa, T., Yasukawa, K., Yamanishi, K., Taga, T., and Kishimoto, T. (1993). IL-6-induced homodimerization of gp130 and associated activation of a tyrosine kinase. *Science* **260,** 1808-10.

Nakajima, K., Yamanaka, Y., Nakae, K., Kojima, H., Ichiba, M., Kiuchi, N., Kitaoka, T., Fukada, T., Hibi, M., and Hirano, T. (1996). A central role for Stat3 in IL-6-induced regulation of growth and differentiation in M1 leukemia cells. *Embo J* **15,** 3651-8.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. *Proc Natl Acad Sci U S A* **95,** 14687-92.

Nichols, J., Evans, E. P., and Smith, A. G. (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development* **110,** 1341-8.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95,** 379-91.

Nicholson, A. W. (1999). Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol Rev* **23,** 371-90.

Nicholson, S. E., De Souza, D., Fabri, L. J., Corbin, J., Willson, T. A., Zhang, J. G., Silva, A., Asimakis, M., Farley, A., Nash, A. D., Metcalf, D., Hilton, D. J., Nicola, N. A., and Baca, M. (2000). Suppressor of cytokine signaling-3 preferentially binds to the SHP-2-binding site on the shared cytokine receptor subunit gp130. *Proc Natl Acad Sci U S A* **97,** 6493-8.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* **12,** 2048-60.

Niwa, H., Miyazaki, J., and Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24,** 372-6.

Niwa, H., Yamamura, K., and Miyazaki, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108,** 193-9.

Niwa, M., and Berget, S. M. (1991). Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev* **5,** 2086-95.

Niwa, M., Rose, S. D., and Berget, S. M. (1990). In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev* **4,** 1552-9.

Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87,** 953-9.

Okabe, S., Forsberg-Nilsson, K., Spiro, A. C., Segal, M., and McKay, R. D. (1996). Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. *Mech Dev* **59,** 89-102.

Okazawa, H., Okamoto, K., Ishino, F., Ishino-Kaneko, T., Takeda, S., Toyoda, Y., Muramatsu, M., and Hamada, H. (1991). The oct3 gene, a gene for an embryonic transcription factor, is controlled by a retinoic acid repressible enhancer. *Embo J* **10,** 2997-3005.

Orban, P. C., Chui, D., and Marth, J. D. (1992). Tissue- and site-specific DNA recombination in transgenic mice. *Proc Natl Acad Sci U S A* **89,** 6861-5.

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

Palmieri, S. L., Peter, W., Hess, H., and Scholer, H. R. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Dev Biol* **166,** 259-67.

Parrish, S., Fleenor, J., Xu, S., Mello, C., and Fire, A. (2000). Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. *Mol Cell* **6,** 1077-87.

Patzel, V., and Sczakiel, G. (1998). Theoretical design of antisense RNA structures substantially improves annealing kinetics and efficacy in human cells. *Nat Biotechnol* **16,** 64-8.

Paul, C. P., Good, P. D., Winer, I., and Engelke, D. R. (2002). Effective expression of small interfering RNA in human cells. *Nat Biotechnol* **20,** 505-8.

Pennica, D., Shaw, K. J., Swanson, T. A., Moore, M. W., Shelton, D. L., Zioncheck, K. A., Rosenthal, A., Taga, T., Paoni, N. F., and Wood, W. I. (1995). Cardiotrophin-1. Biological activities and binding to the leukemia inhibitory factor receptor/gp130 signaling complex. *J Biol Chem* **270,** 10915-22.

Pepin, M. C., Pothier, F., and Barden, N. (1992). Impaired type II glucocorticoid-receptor function in mice bearing antisense RNA transgene. *Nature* **355,** 725-8.

Piccin, A., Salameh, A., Benna, C., Sandrelli, F., Mazzotta, G., Zordan, M., Rosato, E., Kyriacou, C. P., and Costa, R. (2001). Efficient and heritable functional knock-out of an adult phenotype in Drosophila using a GAL4-driven hairpin RNA incorporating a heterologous spacer. *Nucleic Acids Res* **29,** E55-5.

Proudfoot, N. J. (1986). Transcriptional interference and termination between duplicated alpha-globin gene constructs suggests a novel mechanism for gene regulation. *Nature* **322,** 562-5.

Proudfoot, N. J., and Brownlee, G. G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **263,** 211-4.

Ratcliff, F. G., MacFarlane, S. A., and Baulcombe, D. C. (1999). Gene silencing without DNA. rna-mediated cross-protection between viruses. *Plant Cell* **11,** 1207-16.

Riley, M. I., Yoo, W., Mda, N. Y., and Folk, W. R. (1997). Tiny T antigen: an autonomous polyomavirus T antigen amino-terminal domain. *J Virol* **71,** 6068-74.

Rose, T. M., Weiford, D. M., Gunderson, N. L., and Bruce, A. G. (1994). Oncostatin M (OSM) inhibits the differentiation of pluripotent embryonic stem cells in vitro. *Cytokine* **6,** 48-54.

Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W., and Staudt, L. M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345,** 686-92.

Ruvkun, G., and Finney, M. (1991). Regulation of transcription and cell identity by POU domain proteins. *Cell* **64,** 475-8.

Saito, T., Yasukawa, K., Suzuki, H., Futatsugi, K., Fukunaga, T., Yokomizo, C., Koishihara, Y., Fukui, H., Ohsugi, Y., Yawata, H., and et al. (1991). Preparation of soluble murine IL-6 receptor and anti-murine IL-6 receptor antibodies. *J Immunol* **147,** 168-73.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: a Laboratory Manual." Cold Spring Harbor NY.

Sanchez Alvarado, A., and Newmark, P. A. (1999). Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc Natl Acad Sci U S A* **96,** 5049-54.

Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., and Sanger, H. L. (1993). RNA-directed RNA polymerase from tomato leaves. I. Purification and physical properties. *J Biol Chem* **268,** 11851-7.

Scholer, H. R., Dressler, G. R., Balling, R., Rohdewohld, H., and Gruss, P. (1990). Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *Embo J* **9,** 2185-95.

Shakin, S. H., and Liebhaber, S. A. (1986). Destabilization of messenger RNA/complementary DNA duplexes by the elongating 80 S ribosome. *J Biol Chem* **261,** 16018-25.

Shi, W., Inoue, M., Minami, M., Takeda, K., Matsumoto, M., Matsuda, Y., Kishimoto, T., and Akira, S. (1996). The genomic structure and chromosomal localization of the mouse STAT3 gene. *Int Immunol* **8,** 1205-11.

Shikama, N., Lyon, J., and La Thangue, N. (1997). The p300/CBP family:integrating signals with transcription factors and chromatin. *Trends in Cell Biology* **7,** 230-236.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107,** 465-76.

Smardon, A., Spoerke, J. M., Stacey, S. C., Klein, M. E., Mackin, N., and Maine, E. M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. *Curr Biol* **10,** 169-78.

Smith, A. G. (1991). Culture and differentiation of Embryonic stem cells. *J Tiss Cult Meth* **13,** 89-94.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336,** 688-90.

Smith, A. G., and Hooper, M. L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev Biol* **121,** 1-9.

Smith, L., Andersen, K. B., Hovgaard, L., and Jaroszewski, J. W. (2000). Rational selection of antisense oligonucleotide sequences. *Eur J Pharm Sci* **11,** 191-8.

Smits, P. H., de Wit, L., van der Eb, A. J., and Zantema, A. (1996). The

adenovirus E1A-associated 300 kDa adaptor protein counteracts the inhibition of the collagenase promoter by E1A and represses transformation. *Oncogene* **12,** 1529-35.

Sokol, D. L., and Murray, J. D. (1996). Antisense and ribozyme constructs in transgenic animals. *Transgenic Res* **5,** 363-71.

Solter, D., and Knowles, B. B. (1978). Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *Proc Natl Acad Sci U S A* **75,** 5565-9.

Spann, T. P., Brock, D. A., Lindsey, D. F., Wood, S. A., and Gomer, R. H. (1996). Mutagenesis and gene identification in Dictyostelium by shotgun antisense. *Proc Natl Acad Sci U S A* **93,** 5003-7.

Srivastava, S. P., Kumar, K. U., and Kaufman, R. J. (1998). Phosphorylation of eukaryotic translation initiation factor 2 mediates apoptosis in response to activation of the double-stranded RNA-dependent protein kinase. *J Biol Chem* **273,** 2416-23.

Stahl, N., Boulton, T. G., Farruggella, T., Ip, N. Y., Davis, S., Witthuhn, B. A., Quelle, F. W., Silvennoinen, O., Barbieri, G., Pellegrini, S., and et al. (1994). Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science* **263,** 92-5.

Stahl, N., Farruggella, T. J., Boulton, T. G., Zhong, Z., Darnell, J. E., Jr., and Yancopoulos, G. D. (1995). Choice of STATs and other substrates specified by modular tyrosine-based motifs in cytokine receptors. *Science* **267,** 1349-53.

Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H., and Schreiber, R. D. (1998). How cells respond to interferons. *Annu Rev Biochem* **67,** 227-64.

Summerton, J., Stein, D., Huang, S. B., Matthews, P., Weller, D., and Partridge, M. (1997). Morphowlino and phosphorothioate antisense oligomers compared in cell-free and in-cell systems. *Antisense Nucleic Acid Drug Dev* **7,** 63-70.

Summerton, J., and Weller, D. (1997). Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev* **7,** 187-95.

Svoboda, P., Stein, P., Hayashi, H., and Schultz, R. M. (2000). Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127,** 4147-56.

Symons, R. H. (1992). Small catalytic RNAs. *Annu Rev Biochem* **61,** 641-71.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons,

L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell* **99,** 123-32.

Taga, T., and Kishimoto, T. (1997). Gp130 and the interleukin-6 family of cytokines. *Annu Rev Immunol* **15,** 797-819.

Takahashi-Tezuka, M., Yoshida, Y., Fukada, T., Ohtani, T., Yamanaka, Y., Nishida, K., Nakajima, K., Hibi, M., and Hirano, T. (1998). Gab1 acts as an adapter molecule linking the cytokine receptor gp130 to ERK mitogen-activated protein kinase. *Mol Cell Biol* **18,** 4109-17.

Tan, S. L., and Katze, M. G. (1999). The emerging role of the interferon-induced PKR protein kinase as an apoptotic effector: a new face of death? *J Interferon Cytokine Res* **19,** 543-54.

Tanaka, H., and Samuel, C. E. (1994). Mechanism of interferon action: structure of the mouse PKR gene encoding the interferon-inducible RNA-dependent protein kinase. *Proc Natl Acad Sci U S A* **91,** 7995-9.

Tavernarakis, N., Wang, S. L., Dorovkov, M., Ryazanov, A., and Driscoll, M. (2000). Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat Genet* **24,** 180-3.

Treisman, R., Cowie, A., Favaloro, J., Jat, P., and Kamen, R. (1981). The structures of the spliced mRNAs encoding polyoma virus early region proteins. *J Mol Appl Genet* **1,** 83-92.

Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* **13,** 3191-7.

Uhlenbeck, O. C. (1987). A small catalytic oligoribonucleotide. *Nature* **328,** 596-600.

Vallier, L., Mancip, J., Markossian, S., Lukaszewicz, A., Dehay, C., Metzger, D., Chambon, P., Samarut, J., and Savatier, P. (2001). An efficient system for conditional gene expression in embryonic stem cells and in their in vitro and in vivo differentiated derivatives. *Proc Natl Acad Sci U S A* **98,** 2467-72.

Vanhee-Brossollet, C., and Vaquero, C. (1998). Do natural antisense transcripts make sense in eukaryotes? *Gene* **211,** 1-9.

Wagner, R. W., Matteucci, M. D., Grant, D., Huang, T., and Froehler, B. C. (1996). Potent and selective inhibition of gene expression by an antisense heptanucleotide. *Nat Biotechnol* **14,** 840-4.

Wargelius, A., Ellingsen, S., and Fjose, A. (1999). Double-stranded RNA induces specific developmental defects in zebrafish embryos. *Biochem Biophys Res Commun* **263,** 156-61.

Warren, G. J., and Green, R. L. (1985). Comparison of physical and genetic properties of palindromic DNA sequences. *J Bacteriol* **161,** 1103-11.

Waterhouse, P. M., Smith, N. A., and Wang, M. B. (1999). Virus resistance and gene silencing: killing the messenger. *Trends Plant Sci* **4,** 452-457.

Weinberg, M., Passman, M., Kew, M., and Arbuthnot, P. (2000). Hammerhead ribozyme-mediated inhibition of hepatitis B virus X gene expression in cultured cells. *J Hepatol* **33,** 142-51.

Wianny, F., and Zernicka-Goetz, M. (2000). Specific interference with gene function by double-stranded RNA in early mouse development. *Nat Cell Biol* **2,** 70-5.

Wiles, M. V., and Keller, G. (1991). Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. *Development* **111,** 259-67.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336,** 684-7.

Wormington, W. M. (1986). Stable repression of ribosomal protein L1 synthesis in Xenopus oocytes by microinjection of antisense RNA. *Proc Natl Acad Sci U S A* **83,** 8639-43.

Wu, L. C., Morley, B. J., and Campbell, R. D. (1987). Cell-specific expression of the human complement protein factor B gene: evidence for the role of two distinct 5'-flanking elements. *Cell* **48,** 331-42.

Wu, S., and Kaufman, R. J. (1997). A model for the double-stranded RNA (dsRNA)-dependent dimerization and activation of the dsRNA-activated protein kinase PKR. *J Biol Chem* **272,** 1291-6.

Wulf, G. M., Adra, C. N., and Lim, B. (1993). Inhibition of hematopoietic development from embryonic stem cells by antisense vav RNA. *Embo J* **12,** 5065-74.

Yang, D., Lu, H., and Erickson, J. W. (2000). Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in Drosophila embryos. *Curr Biol* **10,** 1191-200.

Yang, S., Tutton, S., Pierce, E., and Yoon, K. (2001). Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol Cell Biol* **21,** 7807-16.

Yao, T. P., Oh, S. P., Fuchs, M., Zhou, N. D., Ch'ng, L. E., Newsome, D., Bronson, R. T., Li, E., Livingston, D. M., and Eckner, R. (1998). Gene dosage-dependent embryonic development and proliferation defects in

mice lacking the transcriptional integrator p300. *Cell* **93,** 361-72.

Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H. R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122,** 881-94.

Yin, T., and Yang, Y. C. (1994). Mitogen-activated protein kinases and ribosomal S6 protein kinases are involved in signaling pathways shared by interleukin-11, interleukin-6, leukemia inhibitory factor, and oncostatin M in mouse 3T3-L1 cells. *J Biol Chem* **269,** 3731-8.

Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T., and Kishimoto, T. (1994). Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. *Mech Dev* **45,** 163-71.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101,** 25-33.

Zhong, Z., Wen, Z., and Darnell, J. E., Jr. (1994a). Stat3 and Stat4: members of the family of signal transducers and activators of transcription. *Proc Natl Acad Sci U S A* **91,** 4806-10.

Zhong, Z., Wen, Z., and Darnell, J. E., Jr. (1994b). Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science* **264,** 95-8.

Zhou, A., Paranjape, J., Brown, T. L., Nie, H., Naik, S., Dong, B., Chang, A., Trapp, B., Fairchild, R., Colmenares, C., and Silverman, R. H. (1997). Interferon action and apoptosis are defective in mice devoid of 2',5'-oligoadenylate-dependent RNase L. *Embo J* **16,** 6355-63.