1 INTRODUCTION

The systematic creation of mutations, study of resulting phenotypes and identification of the genes responsible, or genetic screening, has proved a very powerful way of studying gene function in model organisms. Yeast geneticists are known to wax lyrical about "the awesome power of yeast genetics". Genetic screens have, amongst other achievements, delineated the mechanism of cell-cycle control in yeast (Hartwell, Culotti & Reid 1970, Nurse, Thuriaux & Nasmyth 1976), identified the genes involved in embryonic development in flies (Nusslein-Volhard, Wieschaus 1980), and shown which genes are involved in programmed cell death in the worm (Ellis, Horvitz 1986). These studies, which go right to

the very heart of how cells function and how organisms are put together and work, earned those involved Nobel Prizes.

With the completion of the genome sequence of several organisms, including the human genome, and the prediction of the genes contained therein, one of the most important tasks for biologists today is understanding the function of each of the approximately 25,000 genes in the human genome and their roles in the disruption of normal cell operation in disease. Both forward genetic screens, where mutations causing a particular phenotype are associated with genes, and reverse genetic screens, where the phenotype of a defined mutation in a defined gene is studied, are powerful tools for adding functional annotation to the genome sequence. However, screening has thus far proved difficult in mammalian cell culture, due to the difficult and time-consuming nature of generating homozygous loss function mutations in a diploid, non-sexual system, and screens have mostly been restricted to gain-of-function, over-expression screens. Such over-expression screens can identify new gene function by expressing a gene in a situation where it would not normally be expressed, or by increasing its level leading to a change in the state of cellular pathways/networks and thereby revealing its role. This sort of screening is made more systematic by the availability of collections of clones representing full-length, sequence-verified open reading frames (ORFs), with exactly one clone representing each transcript in a genome, or a defined subset of a genome.

RNA interference (RNAi) is a relatively new method for "knocking-down" or reducing gene function, as opposed to complete loss of gene function produced by a gene "knock out" or deletion. First demonstrated by researchers in the worm *Caenorhabditis elegans* (Fire et al. 1998)and related to post-transcriptional gene silencing (PTGS) or co-suppression in plants (Napoli, Lemieux & Jorgensen 1990), this technique relies on the fact that the introduction of double-stranded RNA will trigger the degradation of complementary mRNA (Schwarz et al. 2002). Although it is true that RNAi results in a knock-down rather than a knock-out (a fact that can be an advantage in some cases) and that there are questions over the specificity of RNAi, possibly due to its overlap with the micro RNA pathway (see section 1.2.3), RNAi offers for the first time the possibility of genome-scale loss-of-function screening in mammalian cell systems.

RNAi screening in mammalian cell systems can be conducted in a number of ways, using a number of RNAi-inducing reagents, each of which has its strengths and weaknesses and each of which is suited to different tasks. Inducing RNAi with small interfering RNAs (siRNAs) is quick, efficient and reliable, but also transient and expensive. On the other hand,

bacterial clones carrying plasmids encoded short-hairpin RNAs (shRNAs) provide a limitless supply of reagent, and shRNAs can be used to generate stable or even conditional knock-down. The ability to recover and identify shRNAs at a later point also allows them to be used in pooled selections.

TNF-related apoptosis-inducing ligand (TRAIL) is a ligand that induces apoptosis in a subset of cancer cells, but not in normal cells (Walczak et al. 1999). However, the mechanism by which tumour cells are sensitive and normal cells are resistant is not fully understood. Understanding this mechanism allows us to increase the usefulness of TRAIL as an anti-cancer agent by being able to predict the sensitivity of cells, and also devise ways of sensitising insensitive tumour cells without sensitising normal cells. Furthermore, the understanding of a common weakness of diverse tumour cells over normal cells helps to understand the molecular basis of cancer itself, and may open the way for novel treatments based on this molecular Achilles' heel.

The TRAIL-induced apoptosis pathway presents a good candidate for genome-scale RNAi screening. Apoptosis is a simple and readily measurable phenotype which lends itself to screening by a number of different strategies. A number of genes involved are already known, which provides positive controls against which the performance of screens can be measured. Indeed, a small-scale screen for genes involved in the TRAIL-induced apoptosis pathway has already been successful in identifying new genes and new pathways involved (Aza-Blanc et al. 2003).

In this work, I will compare different methods for genome-scale screening of the TRAIL pathway, focusing on RNAi. Several large-scale screens for new genes in the TRAIL pathway will be described, along with confirmation of hits from these screens. Finally, effectiveness of the different screens and the usefulness of the results will be assessed.

1.1 Genetic Screening

Genetic screening is the practice of studying a particular biological process through the effects on that process of large numbers of genetic changes (mutations), and thereby identifying the genes involved. The term "mutation" can refer to the change in the organism or gene or the actual molecular change itself. At the organism level, mutations can be classified as hypomorphic (reduced gene function, of which a null mutation is most extreme example), hypermorphic (increased gene function) or neomorphic (generation of new functions for a gene). Of these, the most common generated by random mutagenesis are hypomorphic mutations. When considering screening two distinctions can be made: forward

and reverse genetic screens, and selective and non-selective screens.

Forward or traditional genetic screens start with a phenotype of interest and then generate more or less random mutations that change that process. The gene connected to mutations of interest must then be identified. The availability of complete genome sequences and the identification of all or at least many genes in these genomes allows the use of reverse genetics. A reverse genetic screen starts from the gene and then proceeds to assess the impact of mutating that gene.

The second distinction that can be drawn is between selective and non-selective screens. In a non-selective screen each mutant is assessed individually for its effects on the phenotype of interest. This can be very labour-intensive, particularly when mutants which give the phenotype of interest are expected to be rare. In a selective screen, a large number of mutants are exposed to a selective pressure such that mutants with the phenotype of interest are separated from the majority of uninteresting mutants or from wild-type cells or organisms.

1.1.1.1 Forward genetic screens in model organisms

Genetic screens and selections have been particularly successful when applied to model organisms. The vast array of screening paradigms and methodologies available in model organisms is outlined in the Nature Reviews Genetics series, "The Art and Design of Genetic Screens" (Casselton, Zolan 2002, Forsburg 2001, Grimm 2004, Jorgensen, Mango 2002, Kile, Hilton 2005, Patton, Zon 2001, Shuman, Silhavy 2003, St Johnston 2002).

The use of random mutagenesis by radiation or chemical mutagens in yeast and bacteria is greatly aided by the ability to grow these organisms as haploid clonal populations, as most often these methods lead to hypomorphic or null alleles which are likely to be recessive in heterozygous diploid organisms. Early screens in *Escherichia coli* delineated the repressor model of gene regulation in the *lac* operon (JACOB, MONOD 1961), and screens in bacteria continue to be useful for understanding the basic mechanisms of genetics; for example, screens for mutator genes defined the principles of DNA repair in bacteria.

The yeast is often seen as the screener's organism of choice for studying processes in higher organisms that are conserved in this organism. This is attributable to the ease of genetic manipulation, the fact that it can be grown as diploids or haploids, its ability to reproduce sexually or clonally, its very short generation time, and particularly its high recombination rate. Nevertheless, it is a eukaryote whose fundamental cellular mechanisms are remarkably similar to those of higher organisms, so that the knowledge gained has a

general applicability. The ability to generate and manipulate large numbers of mutants allows the isolation of rare mutations such as temperature-sensitive mutants, which are phenotypically normal at one temperature, but show a mutant phenotype at another. This allows the maintenance of homozygous mutants of genes that would otherwise be very deleterious, a strategy that was employed in the seminal screens for genes involved in the control of the cell-division cycle (Hartwell, Culotti & Reid 1970, Nurse, Thuriaux & Nasmyth 1976).

In a forward genetic screen, once a mutation has been isolated, the gene responsible must be identified. The high recombination rate and availability of a sexual cycle mean that mutant genes can be identified by recombination mapping. Alternatively, loss-of-function mutants can be identified by transforming cDNA libraries into cells and screening for rescue of the phenotype, a process known as cloning by complementation.

Although yeast is easy to manipulate, there are many processes that cannot be studied in yeast, particularly those related to the development of an organism and the interaction of cells. Here, worms, fish and flies are the most common and powerful models. Mutations in these models can be generated by radiation, chemical mutagenesis or insertional mutagenesis using mobile elements. Generally, these models require multi-generational breeding schemes to generate the homozygous mutants need for screening for the effects of recessive mutations. Such F_3 recessive screens have included the identification of recessive mutations involved in the embryonic development of *Drosophila melanogaster* (Nusslein-Volhard, Wieschaus 1980). Worms have an advantage here as they reproduce hermaphroditically, which can simplify breeding schemes, reducing number of generations necessary to produce homozygous mutants. This sort of screening was used to identify genes involved in apoptosis in *C. elegans* (Ellis, Horvitz 1986). *C. elegans* is particularly well suited to the study of the genetic basis of development because of its invariant cell lineage (Sulston, Horvitz 1977). The Zebrafish (*Danio rerio*) can be induced to grow as haploids for the first three days of life, or can be induced to reproduce parthenogenetically (Streisinger et al. 1981).

As in yeast and bacteria, once mutants have been isolated, the genes responsible must still be identified, a much slower process than in yeast owing to the longer generation times involved. Insertional mutagensis can help here as the insertion of an element into the genome tags the insertion site, sequences of which can be recovered by PCR and sequencing, and then mapped back on to the genome of a sequenced species.

Screening in the mouse is technically possible, but is slow and expensive since large numbers of homozygous mutant mice must be generated, and the genes responsible must be

mapped through recombination mapping, again requiring the crossing and examination of many mice. There are several methodologies that can speed up this process, such as the use of balancer chromosomes (Kile et al. 2003) or deletion strains (Rinchik, Carpenter & Selby 1990) for screen regions of the genome. An example of a genome-wide screening effort is the identification of genes involved in the innate immune system (Hoebe et al. 2003).

1.1.1.2 Screens vs. Selections

Screening as described above requires the examination of many individuals to identify mutants which match the phenotype in question. Selections can speed up the process by eliminating the need to examine every individual; they are particularly useful in isolating very rare events. Selections are common in yeast and bacteria where large numbers of mutants can be generated and maintained together and then a selective pressure applied such that only mutants of interest will be recovered. Selections are not however, only applicable to yeast and bacteria. Sydney Brenner used resistance to the drug alicarb to select for genes involved in the synapse function (Brenner 1974) in *C. elegans*.

The processes that can be studied by selections are different from those that can be studied using screens. Almost any phenotype can be studied using a screen, but selection restricts this. Firstly, only phenotypes that allow separation can be used. Secondly, in screens phenotypes can be measured in a quantitative way. However, in a selection genes are either isolated or not, and devising paradigms that allow for quantitative assessment of mutants is much more difficult. This is particularly important as the current paradigm for understanding cellular function shifts away from that of a pathway, where the presence of each component is required for the functioning of the pathway, to a network paradigm where each component has quantitative effect on the phenotypic output of the network.

1.1.1.3 Modifier screens

Modifier screens attempt to find mutations in second genes which alter the phenotype of initial mutants. Modifier screens can be divided into two catagories: screens for mutations which enhance the phenotype of the initial mutation (Enhancer screens), and screens for mutations which suppress the phenotype of the initial mutation (suppressor screens). Such screens are useful for identifying genes which are involved in the same pathways as the gene carrying the original mutation. Enhancer screens can also make recessive hypomorphic alleles dominant, as under normal circumstances a 50% reduction in gene activity seen with a heterozygous hypomorphic allele may be able to support a wild-type

phenotype, but cannot when paired with another mutation in the same pathway. This is particularly useful if a strong mutant in a gene is lethal. In this way the genes downstream of the Sevenless (sev) gene in flies were identified. Sev is a receptor tyrosine kinase involved in eye development. A weak hypomorph of sev can support eye development, but only just. Screening for genes which disrupted eye development in a weak sev background isolated Son of Sevenless (SOS), which is a Ras guanine exchange factor (RasGEF) that activates Ras by increasing the rate of exchange of the GDP bound to inactive Ras for GTP (Simon et al. 1991).

Suppressor screens work in the opposite way. Lackner *et al* screened worms that had an activating mutation in the *C. elegans* RAS homolog let-60 for mutations which suppressed the multi-vulval phenotype of this mutant and gave phenotypically normal worms (Lackner et al. 1994), identifying the MAP kinase homologue mpk-1 as being downstream of RAS.

1.1.1.4 Saturation and the limitations of screening

The Nussein-Volhard screen for genes essential for embryo development is an important example as it was an attempt at "saturation screening". In a saturating screen, the aim is to generate enough mutants that every gene in the genome has been mutated and so, in theory, identify all genes involved in a process. A screen is generally said to be reaching saturation when the same genes are identified multiple times. However, such screens are unlikely to ever identify all genes involved in a process for a number of reasons. Firstly, the chance of mutating a target is related to its size. Small targets are unlikely to be hit. For example, micro RNAs (miRNAs) such as lin-4 in the worm have active sequences of only 22nt, and although lin-4 is important for developmental timing, only two alleles have ever been identified. Secondly, a screen can only identify the first essential function of a gene. Imagine a screen for eye development. A gene involved in eye development will not be identified if it is essential for initial embryogenesis as no embryos will survive to show a defect in eye development. Finally, some genes are redundant, either by virtue of gene duplication or by the action of two pathways ultimately controlling the same aspect of phenotype.

There are several ways around the problem of only being able to observe the first essential function of a gene. Enhancer screens can be important here: weak mutations can be isolated in genes where the mutation elicits the phenotype of interest because it is sensitised, but other phenotypes are not elicited because the mutation is not strong enough. Also important are conditional mutants as touched upon earlier in the context of temperature-

sensitive mutants in yeast. In other organisms mutations can be made conditional upon tissue type in so-called clonal screens. Here, meiotic recombination is induced after embryogenesis producing clones of homozygous mutant cells in otherwise heterozygous mutant individuals. This can be induced using X-rays or the Flp/FRT site-specific recombination system from the yeast 2μ plasmid. Flp/FRT can be heat inducible or made tissue specific by putting the Flp recombinase under the control of a tissue-specific promoter. Xu *et al* used this system to screen for genes which lead to uncontrolled proliferation in the imaginal discs and identified several candidate tumour suppressor genes (Xu et al. 1995).

1.1.1.5 Forward genetic screens in cell culture

Screening in tissue culture cells (the only sort of screening that can be performed in human systems) by random mutagenesis is hampered by the difficulty of generating the homozygotes necessary for screening for recessive mutants, and the lack of recombination makes identifying the causative gene difficult. Ting *et al* used a selection to isolate rare loss-of-function mutants that no longer activated NF-κB in response to TNF. Comparison of the mutant and parental lines showed that a candidate (RIP) was not expressed (Ting, Pimentel-Muinos & Seed 1996). More often, screening in cell culture involves over-expressing genes by introduction of cDNA libraries, a procedure known as expression cloning. cDNA libraries are introduced into cells. Clones showing the phenotype of interest are selected, the plasmids recovered and amplified in bacteria and retransformed into cells, enriching for the plasmid of interest. Alternatively, defined pools of cDNA clones are introduced into cells, which are then screened for the phenotype of interest. Positive pools are sub-divided and re-transfected. Multiple rounds of this can lead to pools small enough to allow the testing of single cDNA clones and the identification of the causative clone. Many mammalian genes have been isolated this way, including the transcription factor GATA1, where clones were tested for their ability to bind a certain DNA sequence (Tsai et al. 1989).

1.1.1.6 Reverse Genetics

The availability of genome sequences and gene predictions for most model organisms as well humans has opened the way for a move from forward genetics to reverse genetics. In reverse genetics rather than starting with the phenotype and working back to the gene via a mutation, the starting point is the gene, which is specially mutated and the effects studied. In yeast collections of strains exist with null mutation in more than 90% of ORFs. A similar project – the knock-out mouse project (KOMP) - aims to generate null mutations, tagged

with a suitable selectable marker in a large proportion of all mouse genes. Here, each allele is sequentially targeted with a disruption cassette. In other species collections of mutants in all genes are being assembled as they are generated. For example, the Bloomington stock centre now holds mutations in genes representing a large portion of genes in the *D. melanogaster* genome(Bellen et al. 2004). Genes of known sequence can also be targeted using RNA interference (see below). Collections of resources for over-expression also exist. For example the ORFeome collections in *Schizosaccharomyces pombe*, worms and humans aim to make available one expression clone representing each ORF in that organism (Matsuyama et al. 2006)

Reverse genetic screens involve taking a large or genome-wide collection of knock-out/overexpressing clones and screening them for the phenotype of interest. One interesting application is Synthetic Genomic Array (SGA) screens. Here yeast strains representing each of the yeast's non-essential genes are mated to systematically produce double mutants, which carry knock-outs in two genes. The study of these knock-outs allows for the examination of the interaction between the two genes and the discovery of synthetic phenotypes. A synthetic phenotype is one which is visible in a double mutant when both single mutants are normal. This has been used to systematically assay for synthetic effects of viability (Tong et al. 2001). Selections can also be used in reverse genetics. Here, a large pool of defined mutants is cultured together and a selection applied, with surviving clones isolated and the responsible gene determined, or the pool deconvolved. An application of this in yeast involved using a unique oligonucleotide sequence included with the deletion – a molecular barcode. Yeast were grown competitively for several generations in both rich and minimal media. DNA from the resulting mix of strains was isolated and a microarray was used to determine which barcodes were over- or under-represented, identifying genes which had an effect on fitness (Winzeler et al. 1999).

The advantage of genome-wide reverse genetic screening is that, in theory, every gene is tested and reaching saturation is not an issue. It should, in theory, also be efficient as only one mutation per gene is necessary, whereas in forward genetics, in order to ensure one mutation in most genes, far more mutants than genes must be generated.

1.2 RNA interference

The ability of exogenously introduced double-stranded RNA to induce a reduction in the activity of gene products – RNA interference (RNAi) – is one of a growing collection of small RNA mediated gene regulatory mechanisms, including the micro RNA pathway, the

piRNA pathway and the rasiRNA pathway. Initially observed when attempts to use sense RNA as a control for antisense-mediated knock-down resulted in a similar phenotype to that of antisense RNA in the *C. elegans* (Guo, Kemphues 1995), it was first rigorously described in the worm in 1998 by Fire and colleagues (Fire et al. 1998). RNAi mechanisms function in a wide range of organisms, including flies (Clemens et al. 2000), Zebrafish (Li et al. 2000), mouse embryos (Svoboda et al. 2000) and mammalian tissue culture (Elbashir et al. 2001). In plants, the silencing effect of double-stranded RNA is known as post-transcriptional gene silencing (PTGS) or co-suppression (Napoli, Lemieux & Jorgensen 1990). RNAi is triggered when small 21-22nt double-stranded RNAs (siRNAs) are incorporated into a multi-protein complex known as the RNA-Induced Silencing Complex (RISC) where they guide the cleavage of complementary mRNAs (Schwarz et al. 2002).

This approach has proved very useful in single-gene analysis, especially in mammalian cell culture, where other methods for reducing the function of particular genes are not as quick, easy or as effective. The use of RNAi as a therapeutic agent is also being rigorously investigated by many. However, the one very exciting prospect is the use of RNAi as a screening tool. Screening in this way is now well established in *C. elegans* and *Drosophila* and is becoming more common in mammalian cell culture. Genome-scale RNAi screening in mammalian cell culture is particularly exciting owing to the lack of other methods for generating such large numbers of hypomorphs in such a short time.

As with any technique, RNAi is not perfect. There are three main problems. Firstly, many of the methods generate only a transient knock-down of gene activity, although RNAi in the worm is heritable, at least for a number of generations, and knock-downs can be made heritable in other systems by expressing double-stranded RNA or hairpin constructs from transgenes. The second is that gene knock-down is not gene knock-out and the use of different RNAi-inducing reagents against different genes results in different levels of residual mRNA. This has the advantage that levels of essential genes can be reduced without being eliminated, and in some cases the construction of a series of knock-downs of different efficiencies can be useful (Hemann et al. 2003). However, it also means that a given RNAi-inducing agent may not knockdown the targeted mRNA sufficiently to elicit a phenotype. Finally, questions surround the specificity of RNAi knock-downs, particularly those induced using chemically synthesised siRNAs or hairpins in mammalian cell culture. Several reports have demonstrated unintended consequences of gene knock-down which can be dependent on the sequence of the siRNA/shRNA in question (Birmingham et al. 2006, Jackson et al. 2003, Lin et al. 2005, Lin et al. 2007) or independent of the sequence (Bridge et al. 2003, Kim

et al. 2004, Pebernard, Iggo 2004, Sledz et al. 2003).

1.2.1 Triggers of RNAi

There are a number of ways that RNAi can be triggered. In general, RNAi can be triggered by long dsRNA in non-mammalian model organisms and in mouse stem cells. In the worm, this dsRNA can be introduced by either injection, soaking worms in a solution of dsRNA, or feeding worms bacteria expressing the dsRNA (Fire et al. 1998, Maeda et al. 2001, Timmons, Fire 1998). A library of bacterial clones which can inducibly express dsRNA targeting each of the predicted genes in the *C. elegans* genome was constructed quickly after the discovery of RNAi in the worm (Fraser et al. 2000, Kamath et al. 2003). These clones can either be fed directly to worms or can be used to prepare dsRNA that can be used for soaking or micro-injection.

RNAi can also be induced in *D. melanogaster* by micro-injection into embryos (Kennerdell, Carthew 1998). However, a far more common use of RNAi in *Drosophila* genetics is knock-down of genes in cell culture. In some cases, dsRNA is taken up by the cells directly from the medium (Clemens et al. 2000), in others it is necessary to transfect the dsRNA (Lum et al. 2003). A number of large/genome-wide collections of dsRNA expressing clones are available for *Drosophila* genes (Reviewed in Echeverri, Perrimon 2006).

In mammalian cell culture, RNAi is induced by short dsRNAs known as siRNAs. These siRNA are 21bp long and have 2nt 3' overhangs, and either 5' terminal hydroxyl or phosphate groups. They can be introduced either by transfection of oligonucleotides (Elbashir et al. 2001)(Figure 1.1a) or by transcription of short hairpin structures, known as shRNAs, from plasmids introduced into the cell either by transfection or viral infection (Paddison et al. 2002). shRNAs are processed by the cellular RNA silencing machinery to produce siRNAs (see section 1.2.2). shRNAs come in two varieties: Type I shRNAs consist of the sense and anti-sense sequences separated by a short loop sequence(Paddison et al. 2002)(Figure 1.1b). Type II shRNAs are modelled on miRNAs and contain miRNA flanking and loop sequences (Silva et al. 2005)(Figure 1.1c). siRNAs can either be chemically synthesised or transcribed *in vitro* or *in vivo* (Zheng et al. 2004)(Figure 1.1d). Finally, a slightly different approach, known as endoribonuclease-prepared siRNA (esiRNA) involves transcribing long dsRNAs *in vitro* and using recombinant Dicer or RNaseIII to digest the long dsRNA into short siRNAs. This produces a pool of siRNAs, all targeting the same gene (Kittler et al. 2004)(Figure 1.1e).

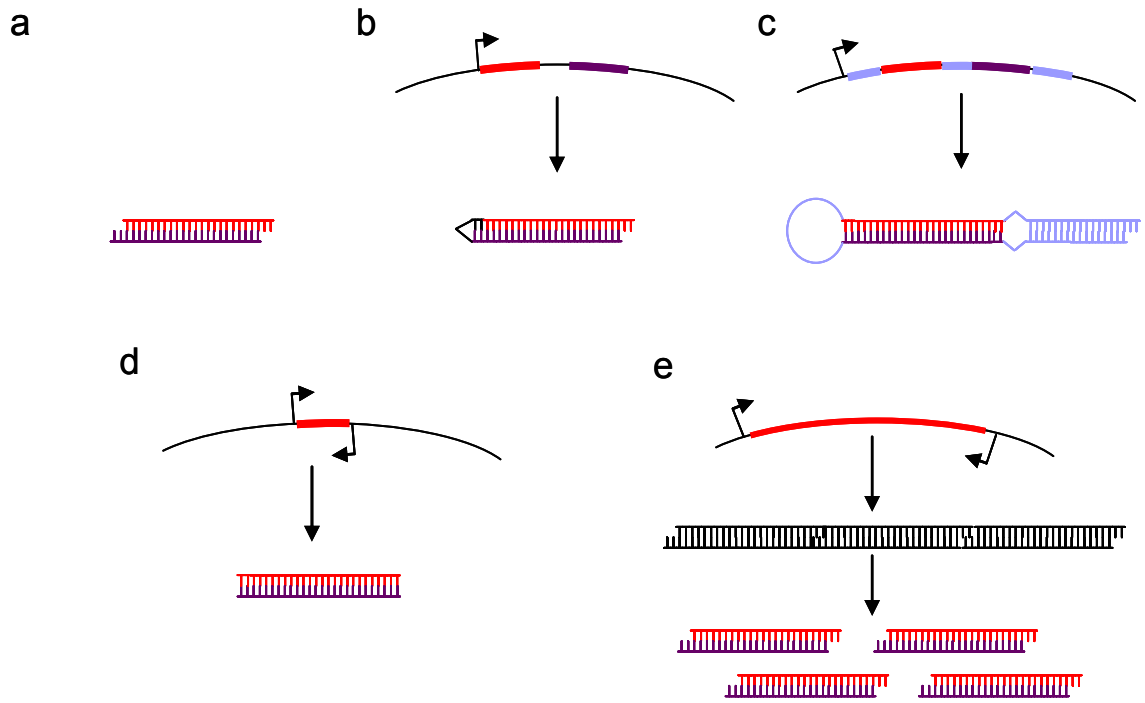

Figure 1.1 RNAi triggers used in mammalian cell culture.
Red – anti-sense strand (complementary to target), Purple – sense strand (identical to target), blue – miRNA sequence. a) Chemically synthesised siRNA, b) Type I shRNA, c) Type II shRNA, d) Transcribed siRNA, e) esiRNA.

There are a number of readily available libraries of RNAi-inducing reagents targeting human genes. Libraries of siRNAs targeting all or most mammalian genes are sold by several companies (including but not limited to Ambion, Dharmacon and Qiagen). These libraries contain between two and four siRNAs targeting each gene, with each siRNA either in a separate well (Ambion, Qiagen) or pooled according to gene (Dharmacon). The effect of pooling is disputed. Those that pool their siRNAs claim that it increases the potency compared with the average of the individual siRNAs contained within the pool and that, importantly, it also reduces off-target effects (see section 1.2.3) by diluting the off-target effect of each individual siRNA. However, those that do not pool claim that a pool is not as effective as the best siRNA in it and that one "dirty" siRNA can increase the off-target effects of the whole pool. It would, however, be important to ensure that in a siRNA pool each siRNA is incorporated into RISC with a similar efficiency to prevent competition.

At the time of starting this project, two large shRNA libraries were available (Berns et al. 2004, Silva et al. 2005). These libraries differ in several characteristics, summarised in Table 1-1. The most important differences are the type of hairpin (Type I vs. Type II), the inclusion of barcodes in the Silva library, and the fact that the Silva library is sequence verified (although the Berns library is being sequenced).

Feature	Silva et al	Berns et al
Number of clones (2004/present)	18,882/81,500	23,742/23,742
Number of genes (2004/present)	13,456/28,500	7,914/7,914
Vector	pSHAG-MAGIC2	pRetroSuper
Vector introduction methods	Transient Transfection Stable Transfection Viral Infection	Transient Transfection Stable Transfection Viral Infection
Barcodes	Separate 60bp barcodes	Uses hairpin sequence
Promoter	U6	H1
Other vector features	MAGIC cloning system	
shRNA type	Type II	Type I
shRNA cassette design	Proprietary, rule based Based on miRNAs Three mismatches to any other gene	Based on 9 rules including: 19bp in length Start with C/G (to introduce strand bias) Share minimal homology with other targets 30%-70% CG
Redundancy	Variable: Some genes have only one hairpin, some have up to 9.	Three hairpins per gene
Pooling	None	Available as per gene, per plate and per library
Sequence verification	Yes	No

Table 1-1 Comparison of the shRNA library available at start of project

1.2.2 Mechanism

Biochemical and genetic analysis has outlined the mechanism through which double-stranded RNA silences gene products. A summary of the current model is presented in Figure 1.2.

RNA interference is triggered by short double-stranded RNAs (Elbashir, Lendeckel & Tuschl 2001) known as siRNAs. As well as being 19bp-22bp in length, siRNAs have several other features. Important other structural features include having characteristic 2nt 3' overhangs and phosphorylated 5' ends. They come from a range of sources including exogenous long double-stranded RNA (dsRNA) and short hairpin structures (shRNAs). Long dsRNAs are converted into siRNAs by Dicer endonucleases, which are multidomain ribonuclease III enzymes(Bernstein et al. 2001). shRNAs originate either from natural long primary transcripts known as pri-miRNAs, which are then processed to form the short-hairpin pre-miRNAs (Lee 2002), or from artificial constructs expressed from plasmid or viral vectors(Paddison et al. 2002). The first step in the processing of pre-miRNAs or constructs designed to mimic them is enacted by the Microprocessor complex, which contains the

RNase III enzyme Drosha and its dsRNA-binding protein (dsRBP) partner DGCR8 (known as Pasha in *Drosophila*) (Denli et al. 2004). Pre-miRNAs are formed in the nucleus and then exported into the cytoplasm by exportin-5, which also exports artificial shRNAs (Yi et al. 2003). Pre-miRNAs and artificial shRNAs are converted to siRNAs through the action of Dicer (Hutvagner et al. 2001, Paddison et al. 2002). In *Drosophila* miRNA processing is carried out by Dicer-1 and dsRNA processing by Dicer-2, suggesting distinct pathways for siRNAs and miRNAs (Lee et al. 2004 Apr 2). However, mammals and worms have only one Dicer protein.

The antisense strand, or guide strand, of the siRNA guides a multiprotein complex known as the RNA Induced Silencing Complex (RISC) to the target mRNA to which the guide strand is complementary. If the siRNA guide strand is completely complementary, RISC then cleaves the target at a position opposite the bond between the 10^{th} and 11^{th} nucleotides of the guide(Elbashir et al. 2001). The endonuclease at the heart of RISC has been shown to be Argonaute-2 (Liu et al. 2004, Rand et al. 2004). Argonaute proteins have various roles in small RNA-silencing pathways, but in humans only Argonaute-2 is capable of endonuclease activity ((Reviewed in Peters, Meister 2007).

Although purified Argonaute-2 can bind single-stranded RNA to reconstitute RISC activity, it is not able to load double-stranded siRNA (Rand et al. 2004). The loading of siRNA into RISC goes through two intermediates known as Complex B and the RISC Loading Complex (RLC, otherwise known as Complex A)(Tomari et al. 2004a). In *Drosophila,* the RLC contains at least Dicer-2 and the dsRBP R2D2 (RNA binding domains x 2 associated with Dicer-2). R2D2 binds to the phosphylated 5' strand at the most stable end of the siRNA duplex, simultaneously specifying which strand will be loaded into RISC (not the one it is bound to) and licensing the siRNA for entry into RISC (Tomari et al. 2004b). This provides a mechanism for the observed asymmetry, whereby the strand of the siRNA which has a lower thermodynamic stability at its 5' end is preferentially loaded into RISC (Schwarz et al. 2003). In mammals it has been suggested that the dsRBPs TRBP and PACT play this role. They contain multiple RNA-binding domains homologous to R2D2 and are found in complexes with Dicer and Argonaute-2(Chendrimada et al. 2005, Lee et al. 2006). However, the precise role of these dsRBPs in the RNAi pathway is unclear. Some have reported that they are required for siRNA assembly into RISC (Chendrimada et al. 2005, Gregory et al. 2005, Lee et al. 2006). Chendrimada *et al* demonstrated that knock-down of TRBP reduces miRNA processing, but that this was likely due to a destabilisation of Dicer (Chendrimada et al. 2005). In contrast, it has also been reported that both TRBP and PACT are dispensable

for siRNA mediated silencing, but are required for miRNA processing (Kin et al. 2007).

There are two models for the mechanism by which active RISC, bound to the guide strand of the siRNA, is formed from Dicer-TRBP-PACT, double-stranded siRNA and Argnaute-2. It was originally believed that the double-stranded siRNA was unwound by a helicase, with the guide strand retained, which was then progressively inserted into RISC. This was based on the finding that Argonaute was only found associated with single-stranded RNA and that RISC assembly was ATP dependent (Nykanen, Haley & Zamore 2001). A role in RISC formation has been shown for several helicases (Meister et al. 2005, Robb, Rana 2007, Tomari et al. 2004a).

However, recently a second model for the loading of siRNA into RISC has emerged. It has been demonstrated that in fact Argonaute-2 can be found associated with double-stranded siRNA and that it cleaves the passenger strand, facilitating its dissociation from RISC, and that this process is ATP independent (Matranga et al. 2005, Rand et al. 2005). It is not clear that these two models are mutually exclusive, since imperfectly matched miRNAs are not cleaved by Argonaute-2, and in humans, siRNAs and miRNAs associate with all four Argonaute proteins, but only Argonaute-2 has been demonstrated to have endonuclease activity.

In nematodes, plants and fungi, the primary siRNAs derived from long dsRNA trigger the production of secondary siRNAs outside the region covered by the original dsRNA, a process known as transitive RNAi. This process is dependent on the action of RNA dependent RNA polymerases (RdRPs). In *C. elegans* this activity is encoded by the RRF-1 gene (Sijen et al. 2001). These secondary siRNAs form a separate class of RNAs which are biochemically distinct, function through separate Argonaute proteins, and do not themselves trigger amplification (Pak, Fire 2007). RNAi in *Drosophila* and mammals is not dependent on RdRPs (Schwarz et al. 2002) and no evidence of transitive RNAi has been found for these species. In plants and some yeast (such as *Schizosaccharomyces pombe)*, siRNAs can act to silence transcription by triggering methylation of DNA. There is some evidence that siRNAs targeted at promoters, but not coding sequences, can direct methylation-dependent transcription silencing in mammals (Kawasaki, Taira 2004, Morris et al. 2004, Park et al. 2004). This has been shown to be dependent on Argonaute-1 (Kim et al. 2006).

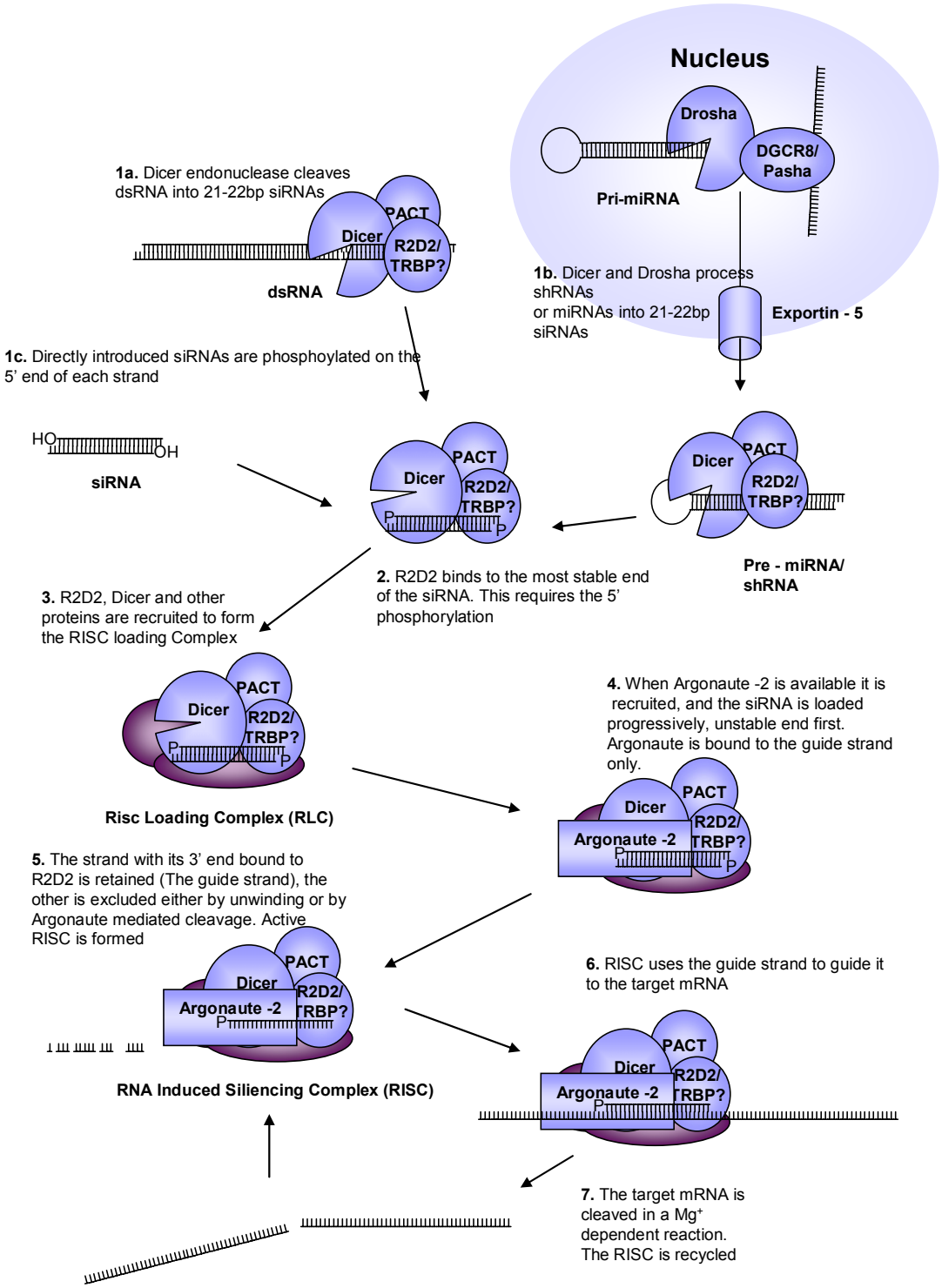

Figure 1.2 A possible model for siRNA mediated mRNA degredation

1.2.3 Design of potent RNAi inducing agents

Not all siRNAs are of equal potency in knocking down target mRNAs (Elbashir et al.

2001). The original guide-lines for the design of siRNAs specified few factors for increasing the sequence dependent potency of the siRNA. These included that the siRNAs should be 21bps in length, with 2nt overhangs, which should be UU or UG (or dTdT/dTdG) on both strands and the duplex should have a low GC content (Elbashir et al. 2001, Elbashir et al. 2002).

Subsequently much work has been carried out to determine what makes a potent siRNA. Some factors, such as that siRNAs should show a lower T_m at the 5' end of the guide strand were deduced from studies of the biochemistry of the RNAi pathway (Schwarz et al. 2003). Others have been deduced from analysis of the potency of a large number of randomly chosen siRNAs. Properties predicted include a lack of inverted repeat sequences, a lack of tracts of 9nt or more comprising entirely G or C and a variety position specific base preferences (Reynolds et al. 2004, Ui-Tei et al. 2004). Such determinants have been extensively reviewed (for example Patzel 2007)

The finds of these studies are drawn together in the many different algorithms for designing siRNAs. These can be based either rule based systems (Holen 2006) or use artificial intelligence methods such as neural networks (Huesken et al. 2005) or support vector machines (Jia et al. 2006). Many of these algorithms have been made available as design tools both by academic and commercial groups (Patzel 2007)

1.2.4 Specificity

One of the most important factors when considering the usefulness of RNA interference as a tool for large-scale screening, the confidence with which results can be regarded, is the specificity of gene knock-down. Originally, RNAi was reported to be very specific (Elbashir et al. 2001, Elbashir et al. 2001). However, it is now understood that there can be a range of unintended effects that accompany the knock-down of the targeted gene.

A discussion of the so called "off-target" effects of an siRNA or shRNA can be divided between those effects which are dependent on the sequence of the RNAi-inducing molecule and those effects which are independent of sequence. These can be distinguished by studying the effects of si/shRNAs targeting the same gene. Consider the hypothetical experiment presented in Figure 1.3. In Figure 1.3a, siRNAs targeting the same gene have the same effect, and that effect is different from the effect of siRNA targeting another gene. Note that in this situation, transcripts other than the one targeted may be affected by events downstream of silencing the targeted transcript. In Figure 1.3b, transcripts other than the targeted transcript are affected and the effects are different for different siRNAs targeting the

same gene – that is the effects are sequence dependent. Finally, in Figure 1.3c transcripts other than the targeted transcript are affected, but all siRNAs have the same effect, even though they are targeting separate genes, that is, they are sequence independent.

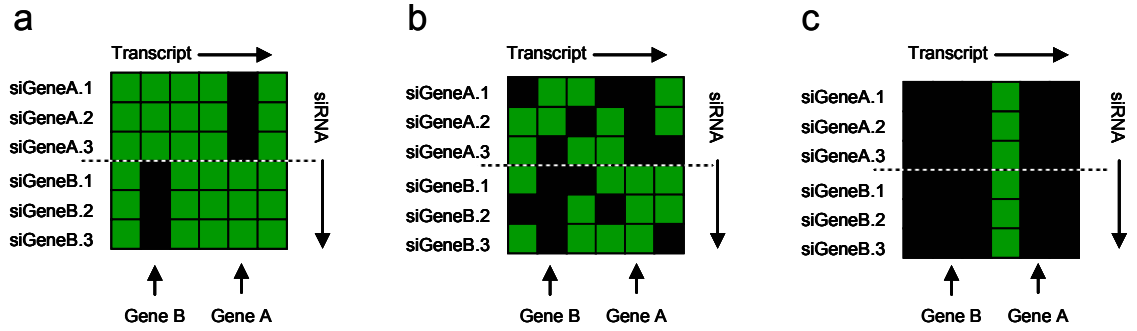

Figure 1.3 Sequence specific and sequence independent effects of siRNAs.
A hypothetical microarray experiment. Six siRNAs, targeting two genes are transfected into six populations of cells. mRNA is harvested and anaylsed on microarrays. Black cells represent genes that under-expressed compared with green cells a) On-target effects. Each siRNA regulates the expression of the same transcript as the other siRNAs targeting that transcript. b) Sequence specific off-target effects. Expression of transcripts other than the transcript targeted are changed and different siRNAs targeting the same gene have differing effects. c) Sequence independent off-target effects. Expression of transcripts other than the transcript targeted are changed and siRNAs targeting different genes have the same effect. After (Jackson, Linsley 2004).

1.2.4.1Sequence specific off-target effects

Original reports of RNAi in mammalian cells using siRNAs found high levels of specificity, with as little as a one base mismatch resulting in a drastic reduction in knock-down efficiency (Elbashir et al. 2001). However, it has been shown that the RNAi machinery is mostly tolerant to single-base mismatches, although the degree of tolerance is dependent on the position and identity of the mismatched bases (Du et al. 2005).

Initially, microarray experiments surveying the effect of siRNA transfection on a transcriptome level supported the idea that RNAi was specific (Chi et al. 2003, Semizarov et al. 2003), finding either that knock-down affected only the targeted gene (Chi et al. 2003), or that off-targets could be eliminated by reducing the concentration of siRNA or by more restrictive design (Semizarov et al. 2003). However, another report in the same year found large numbers of transcripts were affected by transfection of siRNAs and that only a small number were in common between different siRNAs targeting the same gene(Jackson et al. 2003). It was found that off-target effects fell into one of two groups. One group consisted of transcripts that shared a region of homology to the central 14 bases of the siRNAs. The second group consisted of transcripts that had much shorter regions of homology, as little as 7 bases, to the 5' end of one strand of the siRNA.

This was supported by a larger, more systematic study into the relationship between

siRNA sequence and the targets silenced. Birmingham *et al* created a large database of off-target effects elicited by a number of siRNAs targeting a number of genes. They found that overall sequence identity below 1 or 2 mismatched nucleotides was not correlated with knock-down. However the off-target effects were associated with the presence of one or more 6 or 7 base "seed" sequences homologous to nucleotides 2-7 or 2-8 of the siRNA guide strand in the 3' UTR of transcripts (Figure 1.4). The more of these seed sequences present, the greater the probability that the gene would be affected. However, not all transcripts containing seed sequences in their 3' UTR were affected, indicating that other determinants are important (Birmingham et al. 2006).

Lim *et al*, also found that 7nt complementary in the 5' seed region of the siRNA was sufficient to cause off-target knock-down of transcripts. They conducted a screen for regulators of the HIF-1 pathway. They found that two of their top three hit siRNA shared the same 7nt sequence in the seed region and that this was complementary to two sequences in the 3' UTR of HIF-1α. They also found that the context of the 7nt match was important (Lin et al. 2005).

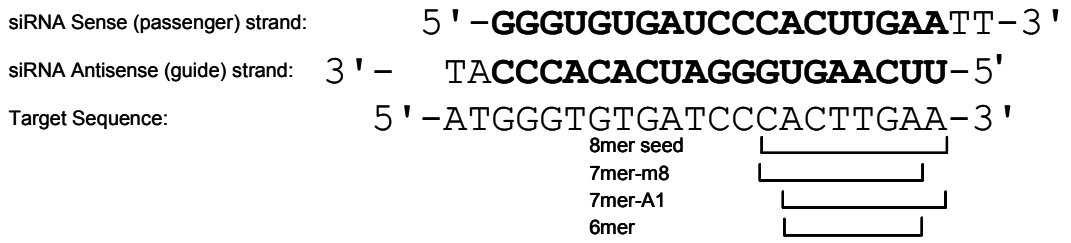

Figure 1.4 Different types of seed sequence as defined by (Lewis, Burge & Bartel 2005)

The importance of 6nt-7nt seed sequences in the 5' region of the siRNA guide strand is reminiscent of the 6nt-7nt seed sequence that has been found to be important in the specificity of micro RNAs (Brennecke et al. 2005, Doench, Sharp 2004). Several lines of evidence support the hypothesis that at least some off-target effects are due to siRNAs acting as miRNAs. Firstly, siRNAs, like miRNAs, may regulate genes with which they are mismatched at the transcript level or protein level (Bagga et al. 2005, Wu, Fan & Belasco 2006). There are several reports of siRNAs affecting protein levels of genes but not affecting transcript levels to the same extent. (Alemán, Doench & Sharp 2007, Saxena, Jonsson & Dutta 2003) . So-called GU wobble base pairing between the siRNA and the mRNA has been shown to reduce mRNA silencing, but not protein silencing (Alemán, Doench & Sharp 2007, Saxena, Jonsson & Dutta 2003) . Secondly, it has been shown that degradation of off-target mRNA is less well correlated with cleavage at the canonical siRNA cleavage site, and

may be independent of Ago2, which is the endonuclease involved in RNAi mediated silencing (Alemán, Doench & Sharp 2007) . miRNAs associate with all four human Argonaute proteins, but only Ago2 is capable of specific cleavage of the target mRNA. miRNA associated mRNA degradation is thought to occur at least partly via the recruitment of decapping and de-adenylation proteins causing a destabilisation of the message (Behm-Ansmant et al. 2006, Wu, Fan & Belasco 2006)

The 3' UTRs of genes targeted by miRNAs often contain multiple seed sequences for multiple miRNAs. It has been shown that the optimal spacing for these seeds is 13nt-26nt. Seeds that are closer than 13nt actively interfere with one another, and any cooperative effect from having multiple seeds in a UTR is dramatically reduced for seeds spaced further than 26nt away from one another (Grimson et al. 2007, Sætrom et al. 2007) . Sætrom *et al* also examined the location of matches to the seed sequences of the siRNAs used in Birmingham *et al* in the 3'UTRs of genes which were shown to be regulated by off-target effects. They found that there was an under-representation of matches to the seed sequences of miRNAs known to be expressed in the cell type used by Birmingham *et al* within 13nt of the matches to the seed sequences of siRNAs used in the study. They also found an over-representation of matches to the seeds of these siRNAs within 26nt of seed sequences for co-expressed miRNAs in the 3'UTRs of transcripts shown to be regulated by these siRNAs. This suggests that the effect of siRNAs is dependent on co-targeting of transcripts by miRNAs and as such would mean that off-target effects would be cell-type specific to a certain extent.

There is now convincing evidence that siRNAs can affect the expression of un-targeted genes at both the mRNA and protein level. It is possible that these effects are mediated in two ways. Firstly, the RNAi machinery would seem to be less sensitive to single or double mismatches in otherwise perfectly matching targets. Secondly, some of these effects are due to siRNAs acting in the miRNA pathway. As miRNA target-site prediction improves, it might be possible to predict off-target effects (Grimson et al. 2007), but this is not currently feasible. It has also been reported that chemical modification of siRNAs can increase their specificity. Addition of 2' O-methyl groups to bases of the passenger strand reduces its incorporation into RISC, therefore reducing any off-target effects due to this strand. Addition of 2' O-methyl groups to base 2 of the guide strand can also reduce, but not eliminate, off-target effects, from this strand. However, off-targeted genes with perfect matches to the seed region are less affected.(Jackson et al. 2006). The mechanism of this effect is not known. The authors of the original report speculate that the addition of the modification to bases in the seed region reduces the stability of pairings between the siRNA

and the target mRNA in this region, meaning longer matches are required for proper pairing. Alternatively, they also suggest that possibly the modification renders RISC incapable of cleaving non-perfectly matched targets (Jackson et al. 2006). However, if off-target effects are due to miRNA like effects, as is suggested by multiple lines of evidence explored above, then this second explanation would seem unlikely as degradation of miRNA targets is probably not due to cleavage by RISC (Alemán, Doench & Sharp 2007) . In order to control for possible off-target effects it is important that all RNAi experiments are confirmed by at least 2 independent siRNAs/shRNAs. This is particularly important in screening applications, as hits are likely to be enriched in siRNAs/shRNAs that have off-target effects on genes involved in the process under study.

One interesting possibility is that this knowledge of off-target effects could be used to extract more information from screening datasets. One group has reported that several hits identified in a screen could be ascribed to off-target effects. They took the seed sequences of siRNAs which were causing off-target effects and used them to search for matches to a database of 3' UTRs. In this way they were able to find genes that could be regulated by the off-target effects of these siRNAs. In this way they identified a gene which was regulated by several of siRNAs causing off-target hits and which was important in the process they were studying, but which had been missed by their primary screen. (Lin et al. 2007).

1.2.4.2 Sequence-independent off-target effects

It was initially thought that RNAi in mammalian cells would not prove possible because double-stranded RNA triggered a general anti-viral response termed the interferon response in such cells. This response includes a general shut-down of translation and non-specific degradation of mRNA. It was originally thought that double-stranded sequences of a length shorter than 30 bp could escape this response (Manche et al. 1992). However, there are now several reports of induction of interferon-response-related genes in a non-specific manner after the introduction of RNAi-inducing molecules.

One study used microarrays to demonstrate that up to 50 interferon-response related genes were up regulated when siRNAs targeting several genes were transfected into human cells. Mouse cells with a knock-out in the gene responsible for recognising dsRNA and activating the interferon response, PKR, did not exhibit this response (Sledz et al. 2003). The mechanism of this induction remains unknown, although it is known that blunt-ended siRNAs can trigger the interferon repsonse via an interaction with RIG-1 and that siRNA

purity can affect interferon induction (Marques et al. 2006). The response has also been shown to be cell-type and transfection-reagent specific (Yoo et al. 2006). Others have found that shRNAs expressed from plasmid vectors using the U6 or H1 promoters can induce an interferon response (Bridge et al. 2003) and that in the case of the U6 promoter this was due to the presence of an AA motif in the promoter sequence (Pebernard, Iggo 2004). Yet another possible trigger of the interferon response is the triphosphate found on the end of siRNAs produced by *in vitro* transcription with T7 RNA polymerase (Kim et al. 2004). Although in none of these cases do the authors examine the effects of the interferon response on the viability of the transfected/infected cells, it can be implied that there is no massive effect on viability at the time points the investigators examined. Most of the reports involve measuring the transcript levels of genes involved in the response either by microarray or by real-time PCR, therefore, viable cells most be present to allow the isolation of RNA. This is despite in some cases a report induction of interferon response markers of up to 500 fold (Bridge et al. 2003).

It has been found that certain GU-rich sequence motifs in siRNAs can trigger an inflammatory immune response in some cell types, possibly through the activation of the Tol receptor pathway (Judge et al. 2005). Although this is technically a sequence specific reaction, it is included here as it is a reaction to the siRNA itself rather than a consequence of an off-target RNAi effect.

A separate way in which siRNAs/shRNAs can trigger a non-specific response is by the overloading of the endogenous RNAi/miRNA machinery. In theory, one could imagine that saturating the RNAi/miRNA machinery with siRNA or shRNA might lead to a reduction in the activity of endogenous miRNA. Thus far there have been no reports of this for siRNAs. However, shRNAs have been reported to cause toxicity in the livers of mice after virus-mediated delivery. This effect was accompanied by a reduction in miRNA levels and miRNA mediated silencing. This competition could be relieved by the over expression of exportin-5, a protein believed to be involved in the transport of pre-miRNAs and shRNAs from the nucleus into the cytoplasm (Grimm et al. 2006).

Although it is possible to eliminate some of these general responses using careful siRNA/shRNA and vector design, this still cannot guarantee a complete lack of response. All of the effects reported here are concentration dependent, so it is important to use the lowest concentration of siRNA/shRNA that elicits the intended response. Further, it also emphasises the importance of comparing the effect of siRNA silencing with non-targeting controls as well as non-transfected controls. Ideally, induction of the interferon response

could be controlled for by measuring the induction of interferon response associated genes, such as OAS1 by quantitative reverse-transcription PCR (qRT-PCR).

1.2.5 RNAi Screening

1.2.5.1Screening Paradigms

There are a multitude of ways in which an RNAi screen can be performed depending on the organism, the process being studied, and the library being used. Several different paradigms are summarised in Figure 1.5.

The most obvious paradigms, and the ones most widely applicable, are the gene-by-gene approaches, where the screen consists of a large number of experiments, usually carried out in a multi-well plate, where the output is either some sort of bulk measurement, usually fluorescent or luminescence based, read by a plate reader, or recorded down a microscope. Bulk outputs provide the quickest, easiest, and possibly most quantitative outputs, but microscopy can provide a higher density of information including the recording of multiple phenotypes. Micoscopy also allows for the examination of a wider range of phenotypes including unexpected ones. However, the quantitative recording of phenotypic information from a large number of RNAi knock-downs requires either a large number of man-hours, or highly specialised automated microscopy platforms and image-analysis pipelines.

One way in which the throughput of high-content screening can be increased in cell culture screens is by using RNAi arrays. Here, the RNAi-inducing agents are mixed with transfection reagent and spotted on to known locations on a glass slide, along with some sort of transfection marker. Cells are then grown as a monolayer on the surface of the slide. The slide is scanned using a slide-imaging platform. Transfected cells are marked and their phenotype recorded, the gene knocked-down being determined by the position on the slide.

Pooling RNAi-inducing agents and performing selections rather than screens has the advantage of greatly increasing the throughput of a study, allowing the entire genome to be examined in 100-1000 fold fewer individual experiments than gene-by-gene approaches. A collection of RNAi-inducing agents targeting a number of genes is introduced into cells as a pool and a selection is applied such that cells showing the phenotype of interest are separated from the majority of cells. The phenotype-causing RNAi-inducing agent is then identified. Here, the use of shRNAs is particularly applicable, as the knock-down of interest can be determined by examining the hairpins present in the selected clones. This can be done by the identification of either hairpins sequences themselves or the identification of associated

barcodes. Identification can either be by sequencing or hybridisation to a microarray of hairpin/barcode sequences. Unless the selection employed has a very low escape rate and some method of ensuring only one shRNA is introduced into each cell, it is the enrichment of hairpins that is important rather than just the presence of hairpins. Here, barcode arrays probably have the edge unless a method of sequencing many hairpins is employed. One possibility would be to extract hairpin/barcode sequences and ligate several together to form a concatermer, such that a single sequencing reaction could cover several hairpins/barcodes. Although such avselection based screen would greatly increase the throughput of a study, it would restrict the processes that can be studied to those that allow selection, such as survival assays or assays where outcomes can be separated by fluorescence activated cell sorting (FACS). Also, the assay must allow few negative clones to escape.

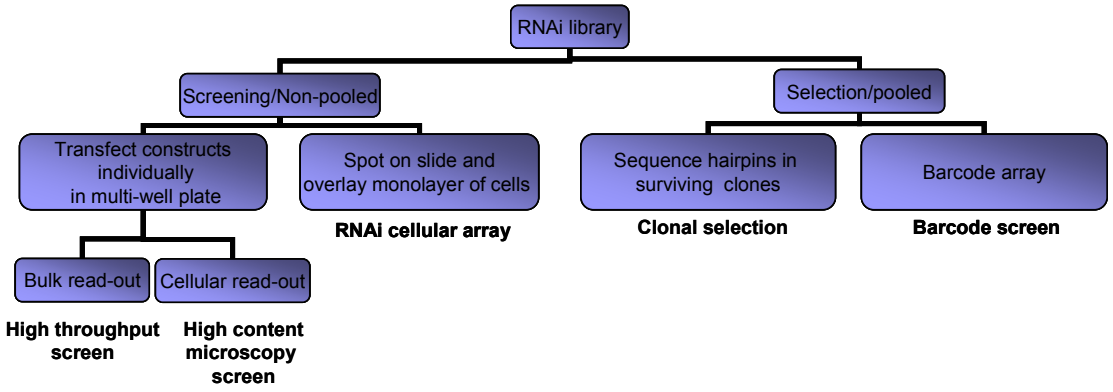

Figure 1.5 Paradigms for genome-scale RNAi screening

1.2.5.2 Screening in model organisms

The first RNAi screens in model organisms, indeed the first RNAi screens in any organism, were screens of all predicted genes on particular chromosomes of the worm *C. elegans* (Fraser et al. 2000, Gonczy et al. 2000). Gönczy *et al* micro-injected dsRNAs in pairs, covering most of the genes on chromosome III, and screened for defects in embryogenesis using time-lapse microscopy. Pairs of dsRNAs that elicited a phenotype were tested separately. In this way they identified 133 genes necessary for a range of cellular processes in embryogenesis. In contrast, Fraser *et al* preferred to assay for a wider range of phenotypes. They fed worms on bacteria expressing dsRNA homologous to predicted genes on chromosome I. They assayed the offspring of these worms for obvious embryonic and post-embryonic phenotypes. In this way, they assigned functions to 13.9% of the genes they examined, noting that knock-down of conserved genes was more likely to elicit a phenotype than knock-down of non-conversed genes.

Soon after, the first genome-wide survey of gene function in the worm was published

(Kamath et al. 2003). The screen was carried out in a similar fashion to Fraser *et al.* Worms were fed bacteria expressing dsRNA corresponding to 86% of the worm genome. Information on sterility, embryonic or larval lethality, slow growth and a wide range of post-embryonic phenotypes was recorded from the offspring of fed worms. Phenotypes were recorded for 10% of genes, including 77.2% of genes whose mutation is known to cause a non-viable phenotype and 43% of genes whose mutation is known to cause a post-embryonic phenotype.

Many screens for more focused phenotypes have now been conducted, including screens for genes involved in fat regulation (Ashrafi et al. 2003), mutator genes (Pothof et al. 2003), genes involved in poly-glutamine aggregation (Nollen et al. 2004), genes involved in the RNAi pathway (Kim et al. 2005), genes involved in axon guidance (Schmitz, Kinge & Hutter 2007), genes involved in the non-sense mediated mRNA decay pathway(Longman et al. 2007) and many more.

C. elegans presents an excellent model for studying the organisation and development of multicellular organisms and the malfunction of such in disease. However, direct study of biochemical processes underlying cell biology is often best undertaken in cell-culture systems where the output is measurement of some surrogate for the biochemical state of the cell. Knock-down of a gene may have no obvious effect on the organism as a whole, and yet may profoundly change the state of the cell. Among the model organisms, *D. melanogaster* is the system most widely used. The first RNAi screen in the *Drosophila* cell culture involved the knock-down of 1,000 randomly selected genes taken from a cDNA library. The effect of these knock-downs on the phagocytotic abilities of S2 cells was examined, identifying a receptor for *E. coli* cells (Ramet et al. 2002). A more general survey of the function of the same number of genes was carried out by Kiger *et al* who assayed for defects in "cell morphology" using automated flourescence microscopy (Kiger et al. 2003). Lum *et al* were the first to use a transcriptional reporter to measure the output of a signalling pathway, a paradigm that has proved popular. They surveyed the effect of knock-down of 40% of predicted *Drosophila* genes on transcription from a hedgehog pathway reporter, identifying two new components of the pathway (Lum et al. 2003).

The first genome-wide RNAi screen in *Drosophila* was for genes affecting growth and viability (Boutros et al. 2004). The GenomeRNAi database of RNAi screening results in *Drosophila* (http://rnai.dkfz.de) lists 21 genome-wide screens in Drosophila cell culture. Many of these screens focus on finding genes involved in various signalling pathways, infection/phagocytosis or the cell cycle: phenotypes it would harder to study at the

organismal level.

A recent screen with relevance to the work here is a screen of the *Drosophila* genome for genes involved in DNA damage induced apoptosis. This screen identified genes connected with a wide range of cellular processes including basic metabolism. The conservation of the function of several of these genes was also demonstrated in mammalian cells (Yi et al. 2007).

1.2.5.3 Screening in mammalian cell culture.

Screening in mammalian cell culture has mainly focused on screens in human cells. Screens have been conducted on a number of scales and using a range of paradigms. One of the earliest screens was a screen of kinases for genes that modulated the TRAIL-induced apoptosis pathway (see section 1.2.5). This screen employed a library of around 500 siRNAs targeting the kinases along with a few candidate genes, each siRNA targeting a gene, which were transfected in duplicate. One duplicate was then treated with the apoptosis-inducing ligand TRAIL, and the sensitivity of the cells to TRAIL was measured by comparing the viability of treated and untreated cells 24 hours later. Several genes were identified in this way, including several genes previously unassociated with the pathway and one gene of no previous known function. However, only a very small number of genes were tested with more than one siRNA, which makes it difficult to assess the accuracy of the screen (Aza-Blanc et al. 2003).

The first shRNA-mediated screens were published back to back. Paddison *et al* sought to validate their shRNA library and took a clone-by-clone approach, co-transfecting about 7,000 shRNAs with a construct designed to measure proteasome activity and a transfection marker. In this way they identified genes involved in proteasome function, including aboutt 50% of the shRNAs expected to alter proteasome function (Paddison et al. 2004b). In contrast, Berns *et al* used a pooled selection to identified genes that allowed escape from a p53-induced growth arrest. They transduced cells with pools of viruses containing shRNAs targeting 96 genes. A p53-mediated arrest was then induced, escaping colonies were identified, and their shRNA content determined by sequencing. In this way they tested more than 7,000 genes and identified one known and five new modulators of p53-mediated growth arrest. All their hits were confirmed by multiple shRNAs. (Berns et al. 2004).

One common way in which RNAi has been implemented as a screening tool, while keeping costs and labour down, has been to test sets of candidate genes. For example, in one study, 257 growth regulated genes were tested for their ability to differentially affect the

viability of two cell lines (Machida et al. 2006), revealing that 25% of the genes tested affected viability in one but not both cell lines. Using a small set of candidates allows for a more detailed examination of the phenotype studied. Simpson *et al* used 97 siRNAs targeting 37 putative membrane trafficking genes in a high-content microscopy screen for genes involved in secretion (Simpson et al. 2007). Candidate gene sets can be small, such as in a screen for components of the mammalian retromer which screened 30 nexins (Wassmer et al. 2007) or large, such as a screen for modulators of Glutamatergic and GABAergic synapse development using esiRNA pools targeting 160 genes up or down regulated during rodent synapse development (Paradis et al. 2007). When using candidate gene sets the expectation is that the chance of generating hits is increased.

Although candidate gene studies have been successful in confirming the involvement of genes in a variety of processes, the speed of RNAi promises the ability to identify previously unsuspected genes. Many screens focus on a particular subset of the genome, such as the kinases. shRNAs targeting the kinases have been used in a selective screen for genes involved in RAS signalling. In this screen a cell line which is arrested upon Ras induction was transduced with pools of shRNAs, each targeting 96 genes. shRNAs in colonies which escaped arrest upon RAS induction were determined. In this way two known and one novel RAS-regulated kinase was identified (Nicke et al. 2005). Screens of subsets of the genome have also been carried out using a gene-by-gene approach, such as a screen targeting 650 kinases and 222 phosphatases for genes which affect the sensitivity of cells to various apoptosis inducers (MacKeigan, Murphy & Blenis 2005). High-content automated microscopy has also been used with targeted sets of shRNAs. Moffat *et al* screened a set of 1,028 genes including kinases, phosphatases, tumour suppressors and DNA-binding/modifying genes for effects on mitotic progression using a histone modification specific antibody (Moffat et al. 2006)

Pooled selections aid the throughput of larger, unbiased, genome-scale screens. Screens using both a clonal selection and barcoding paradigm have been performed (see section 1.2.5.1, Figure 1.5). A pooled, clonal selection of 8,500 genes revealed a new gene involved in all-trans retinoic acid (ATRA)-induced cell growth arrest. Here, cells were infected with a library of 43,800 shRNAs targeting 8,500 genes and then grown on ATRA containing soft agar media. Colonies which grew despite the presence of ATRA were divided into those that expressed a GFP marker contained on the shRNA vector and those that did not, allowing the removal of colonies that grew owing to spontaneous mutation. Hairpins from the GFP positive colonies were identified. Positive hairpins were re-transfected

individually to confirm their effect. However, no new hairpins were synthesised and so the hits are based on results from only one hairpin per gene (Hattori et al. 2007). Another pooled clonal selection involved examining the ability of pools of shRNA encoding viruses targeting a total of 4,000 genes to transform primary human cells. This screen identified two known and one novel gene, the knock-down of which transformed cells with a similar efficiency to over-expression of oncogenic RAS (Kolfschoten et al. 2005). Here the result was confirmed with multiple shRNAs targeting the same gene. shRNA libraries have also been used to carry out barcode screens. Westbrook *et al* used a barcoding strategy to look for genes, knock-down of which would induce anchorage independent growth in large-T and telomerase-immortalised primary cells. This identified a transcriptional repressor, REST, as a tumour repressor. Results were confirmed using multiple hairpins, and a sequencing of colonies based approach which closely matched the barcode array results, validating the use of barcode arrays, where previously only proof-of-principle experiments had been published (Westbrook et al. 2005).

Although pooled strategies are very attractive when dealing with such large sets of reagents, they do limit the phenotypes that can be studied. The use of one well one gene approaches increases the phenotypes that can be studied. One screen introduced shRNAs along with a GFP marker into cells and used time-lapse microscopy used to track the motility of transfected cells. In this way CUTL1 was identified as a target of TGFβ signalling and as being involved in invasion and metastasis (Michl et al. 2005).

siRNA libraries have been employed in large genome-scale screens in mammalian cells. Examples include screens for genes that are required for tumour-cell survival (Morgan-Lappe et al. 2007) and genes that prevent differentiation of mesenchymal stem cells into osteoclasts (Zhao, Ding 2007). siRNAs libraries have also been employed for essentially genome-wide screens targeting nearly all genes in the human genome. One such screen tested the effect of gene knock-down on sensitivity of cells to the cancer drug paclitaxel (Whitehurst et al. 2007). Such designs allow the gathering of phenotypic information on a quantitative rather than qualitative basis.

However, this can present a problem for "hit" selection, particularly since such large quantities of data are involved. In this case a very stringent statistical hit selection was applied. The screen was carried out in sextuplicate, with three replicates treated and three untreated. This allowed the application of a student's t-test to test the significance of the difference between the treated and untreated replicates. After the application of a correction for multiple testing, genes that gave a significant difference and also were in the top 2.5

centile-rank for absolute size of the treated/untreated ratio were selected as hits. This still left 87 hits, only 6 of which were selected for confirmation and follow up. Another genome-wide screen employed high-content microscopy to examine the effects of gene knockdown on cell-cycle progression. Information on a number of phenotypic parameters was recorded and the screen was carried out in duplicate. A total of 1,152 genes were reported to reproducibly alter at least one of the phenotypes studied. Although 24 of these genes were carefully confirmed with the use of multiple siRNAs, the authors mainly tried to avoid selecting genes for follow up by employing systems-level analysis. They found that their hits clustered into 8 phenotypic groups, and that they could combine the information from the screen with information from the literature or from published interaction experiments to build phase specific networks. (Mukherji et al. 2006).

Several groups have published protocols and proof-of-principle experiments for cell-array based RNAi screens, but no novel screens have yet been published using this technique (Bailey et al. 2006, Erfle et al. 2007, Silva et al. 2004).

Selecting genes for follow up can be approached in several ways. In many of the pooled shRNA screens, the number of hits seems to be small enough that all novel hits can be confirmed. For screens which produce a larger number of hits, genes are often selected on the basis of biological interest (e.g. Aza-Blanc et al. 2003, Whitehurst et al. 2007, Zhao, Ding 2007). Although this may yield biologically interesting results, it seems a shame given the ability of RNAi screening to identify genes in an unbiased manner.

Given the controversy over the specificity of RNAi (see section 1.2.3), it is perhaps surprising that many screens are published in which hits are not confirmed by multiple siRNAs targeted against the same gene (Hattori et al. 2007, Zhao, Ding 2007). At best, large screens tend only to verify a the few genes that they select for follow up (Mukherji et al. 2006, Westbrook et al. 2005, Whitehurst et al. 2007). The danger of not confiming hits with multiple siRNAs is shown by the rigorous confirmation process undertaken by Morgan-Lappe *et al* in their screen for genes which reduce the viability of tumors cells. Of 48 genes selected as hits only 23 were confirmed on retesting. Of these, only 10 confirmed with multiple siRNAs in multiple assays and of these only 3 had phenotypes that correlated with the level of knock-down (Morgan-Lappe et al. 2007). Others have found that their screens have returned only hits that are due to off-target effects (Lin et al. 2005, Lin et al. 2007).

RNAi screening has yielded new insights in many areas. There are multiple ways that screening can be carried out, depending on the system under study, using a range of reagents. Whichever method is chosen, careful strategies are necessary to confirm hits and to select

genes for follow-up studies.

1.3 TRAIL-induced apoptosis

Tumour Necrosis Factor (TNF)-Related Apoptosis-inducing Ligand (TRAIL) was originally identified using searches of EST databases for homology to other members of the TNF family by two groups independently and shown to have apoptosis-inducing activity (Pitti et al. 1996, Wiley et al. 1995). The TRAIL ligand is a type II membrane-bound protein, although it can also be expressed as a soluble ligand, and is trimeric. It is though to have a number of roles in the immune system (see 1.3.4). The ligand is widely expressed in almost all tissues (Wiley et al. 1995), as are its receptors (Pan et al. 1997b, Walczak et al. 1997). This suggested that the regulation of its apoptosis-inducing activity must be different to that for the FAS system, where expression of both ligand and receptor is tightly controlled. TRAIL is interesting as it induces apoptosis in between a half and two-thirds of tumour cells but not normal cells (Ehrhardt et al. 2003, Walczak et al. 1999, Zhang et al. 1999).

The mechanism by which the sensitivity of cells to TRAIL is regulated is of interest for two reasons. Firstly, by understanding the mechanism of sensitivity, we can predict which cell types – tumour and normal – will be sensitive. This will aid decisions as to the effectiveness of TRAIL as an anti-cancer agent. It will also help to identify combination treatments that increase the effectiveness of TRAIL. Secondly, understanding the regulation of TRAIL sensitivity gives an insight into why so many tumour cells are sensitive when the molecular basis of their transformation is so diverse.

Below is reviewed the current state of knowledge on the mechanism of TRAIL-induced apoptosis and the regulation of sensitivity to TRAIL. A brief overview of the physiological function of TRAIL and the prospects for its use as an anti-cancer treatment is also given.

1.3.1 Mechanism of TRAIL-induced apoptosis

The mechanism of apoptosis induction by the TRAIL ligand is similar to that of other TNF family apoptosis-inducing ligands, particularly that of the FASL/Apo1L/CD95L ligand (Reviewed: Ashkenazi, Dixit 1999) with ligand-bound receptor triggering apoptosis through both a caspase cascade and via the mitochondrial associated intrinsic pathway. The pathway is summarised in Figure 1.6.

The TRAIL ligand can bind four membrane-bound TNF family receptors – DR4 (also known as Death Receptor 4, TRAIL-R1, and TNFRSF10A), DR5 (Death Receptor 5,

TRAIL-R2 or TNFRSF10B), DcR1 (Decoy Receptor 2, TRAIL-R3, TNFRSF10C, TRID or LIT) and DcR2 (Decoy Receptor 2, TRAIL-R4, TNFRSF10D, TRUNDD) (Degli-Esposti et al. 1997a, Degli-Esposti et al. 1997b, Pan et al. 1997a, Pan et al. 1997b, Walczak et al. 1997). TRAIL can also bind the soluble protein Osteoprotegerin (OPG) (Emery et al. 1998). Of these 5 receptors, only DR4 and DR5 contain complete death domains, and it is the binding of the TRAIL ligand to these receptors that induces apoptosis (Pan et al. 1997b, Walczak et al. 1997).

The binding of TRAIL to either DR4 or DR5 causes recruitment of the FADD adaptor protein via interaction of the death domains of the receptors and those of the FADD protein. Recruitment of FADD in turn leads to the recruitment of the cysteine-protease Caspase-8 (also known as FLICE or MACH)(Kuang et al. 2000) through interactions between the Death Effector Domains (DED) of FADD and Caspase-8. Recruitment of Caspase-8 results in the processing from the long inactive form, termed procaspase-8, to the active form, which involves two cleavage events, separating the long and short active domains from one another and from the pro-domain. The complex of Death receptor, FADD and Caspase-8 is termed the DISC (Death Inducing Signalling Complex).

The activated DISC can induce apoptosis by both the extrinsic and intrinsic pathways (Suliman et al. 2001). The extrinsic pathway involves the cleavage and activation of the executioner caspases such as Caspase-3 and Caspase-7. The inactive form of procaspase-3 is cleaved by caspase-8 to give p24 and p12 caspase-3 subunits. The p24 subunit is further self-processed to give the p19/p17 active subunit. (Martin et al. 1996).

The intrinsic pathway is activated by the cleavage of Bid by Caspase-8, which leads to mitochondrial outer membrane permeabilization (MOMP) and release of cytochrome c and DIABLO (also known as Smac), a process for which the protein BAX is required (Deng, Lin & Wu 2002, Suliman et al. 2001). Cytochrome c associates with caspase-9 and APAF-1 to form the apoptosome. The apoptosome activates caspase-9, which in turn activates procaspase-3. However, Caspase-9 activation has been found to be disposable for TRAIL-induced apoptosis (Deng, Lin & Wu 2002), and therefore is not considered part of the TRAIL-induced apoptosis mechanism. DIABLO functions to promote the self-processing of p24 Caspase-3 to the active p20/p17. This is achieved by the binding of DIABLO to XIAP, a member of the IAP (inhibitor of apoptosis protein) family, which would otherwise bind Caspase-3 and inhibit its self-activating activity (Deng, Lin & Wu 2002). Thus blockage of the intrinsic pathway can delay or prevent TRAIL-induced apoptosis, possibly depending on the level of XIAP (Deng, Lin & Wu 2002, Suliman et al. 2001).

1.3.2 Regulation of TRAIL sensitivity

Since it was demonstrated that most transformed, but not normal, cells are sensitive to TRAIL-induced apoptosis, much research has addressed the question of how sensitivity to TRAL is regulated. This intensified when it was found that up to 50% of tumour cells are resistant to TRAIL (Ehrhardt et al. 2003, Zhang et al. 1999). The TRAIL ligand is widely expressed in many normal tissues, suggesting that sensitivity is expressed at a level other than ligand expression. Interest has focused on five areas: Regulation of the TRAIL receptors, the so-called decoy receptors, the inhibitory molecule c-FLIP, the expression of IAP proteins, and the effects of other signalling pathways.

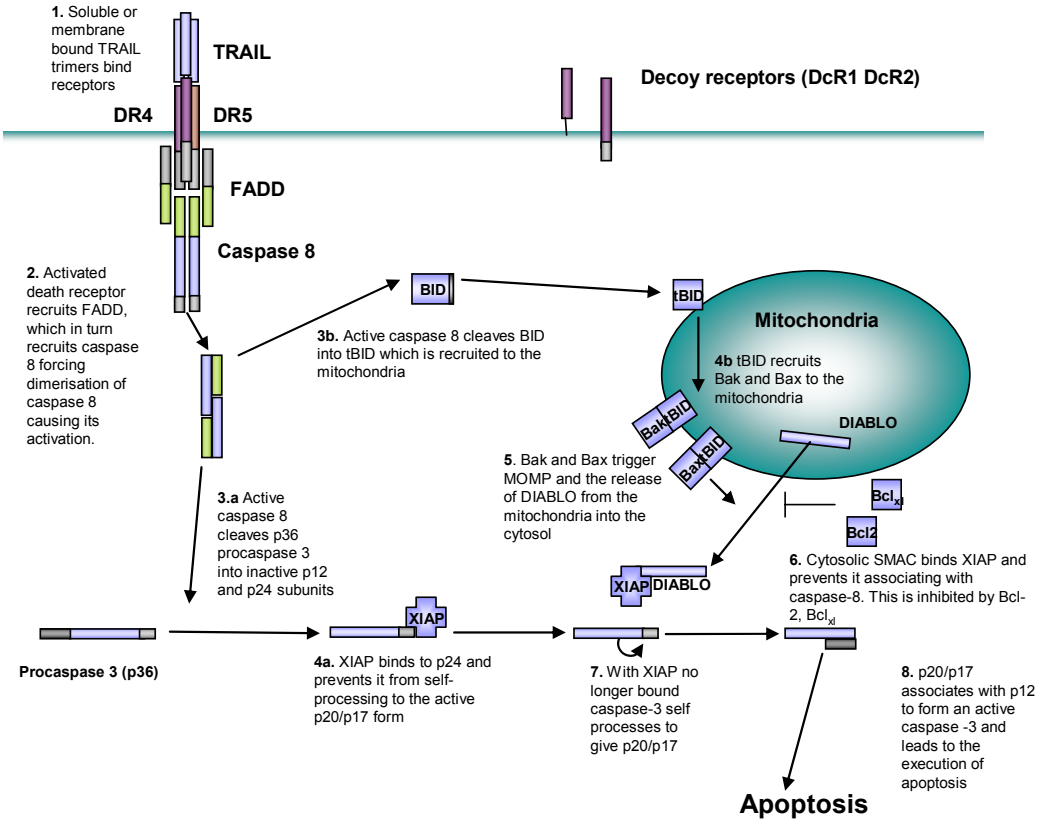

Figure 1.6 A simplified schematic of events leading to TRAIL-induced Apoptosis

1.3.2.1 Regulation of the TRAIL receptors

The original TRAIL receptor cloning papers found that both receptors were widely expressed in a range of normal tissue at the RNA level (Pan et al. 1997b, Walczak et al. 1997). However, studies have looked at the relationship between the levels of the two receptors and the sensitivity of cells to TRAIL. Kim *et al* found that levels of DR4, but not

DR5, correlated with the level of apoptosis induced by TRAIL in a panel of sensitive and insensitive tumour cell lines, but that this could not completely explain sensitivity in all cases(Kim et al. 2000). The same study also found that resistance to TRAIL in one case could be explained by a mutation in the DR4 gene. Resistance in tumour cells can also be explained by deletion of the DR4 gene (Zhang et al. 1999).

This stands in opposition to the finding that in lung carcinoma cells a mutant TRAIL selective for DR4 was less able to induce apoptosis than a DR5-specific TRAIL ligand mutant. This finding was replicated in several tumor cell lines leading the authors to conclude that DR5 was more important in the induction of apoptosis than DR4 (Kelley et al. 2005). However, this study had several problems. Firstly, the DR4-specific ligand had reduced affinity for DR5, but not for the decoy receptors (see below), but the DR5 specific ligand had reduced affinity for all the TRAIL receptors except DR5. Secondly, all the cell lines used in the study expressed DR5 at a higher level than DR4. Indeed, it has been demonstrated, using receptor-sensitive TRAIL mutants, that some sensitive cells signal purely through DR4 (MacFarlane et al. 2005) and so it is likely that the relative importance of the two receptors is cell-type specific.

Expression of DR5 can be regulated by p53 (Takimoto, El-Deiry 2000) and indeed resistant tumor cells can be sensitised to TRAIL by cellular stressors such as γ-irradiation, or anti-cancer drugs such as etopotside, 5-Fluorouracil and proteasome inhibitors in both p53-dependent and p53-independent manners (Anan et al. 2006, Chinnaiyan et al. 2000, Frese et al. 2003, Ganten et al. 2004, Ganten et al. 2005, Gibson et al. 2000). Sensitisation by these agents results in increased expression of either DR5 alone or both DR5 and DR4. However, one report has demonstrated that this increase in receptor expression is dispensable for the sensitisation (Ganten et al. 2004). These treatments do not sensitise normal cells, which generally already express TRAIL receptors, both at an mRNA and a protein level (Ganten et al. 2005, Pan et al. 1997b, Walczak et al. 1997, Wu, Ogawa & Kakehi 2004).

One explanation for these discrepancies could be the localisation of the receptors. Stimulation of natural killer cells (NK cells) and $CD8^+$ T cells leads to movement of TRAIL receptor molecules from the cytoplasm of resting NK and $CD8^+$ T cells to the surface of stimulated cells. However, these cells are still TRAIL resistant (Mirandola et al. 2004). Zhang *et al* have shown that although a panel of melanoma cells lines, expressing all the TRAIL-receptors at both the mRNA and protein level, have a wide variability in the sensitivity to TRAIL, there was a correlation between the level of surface expression of the receptors and the degree of apoptosis induction in sensitive cell lines. Interestingly, the localisation of DR4

and DR5 is controlled in different ways, with expression of DR4, but not DR5, on the cell surface being dependent on the signal recognition particle (SRP) (Ren et al. 2004).

It would therefore seem that, since both sensitive and insensitive cells express the TRAIL receptors on their surface, this does not provide a conclusive explanation of why transformed cells, but not normal, cells are sensitive to TRAIL. However, modulation of TRAIL receptor surface expression does appear to provide a mechanism for the resistance of some tumour cell lines.

1.3.2.2 The TRAIL decoy receptors

Of the four receptors that the TRAIL ligand binds only two are capable of transmitting an apoptosis-inducing signal. This has lead to the remaining receptors being termed "decoy receptors" based on the hypothesis that their function is to titrate away TRAIL from the death-inducing receptors. DcR1 and DcR2 both share homology to the two apoptosis-inducing receptors in their extra-cellular domains. However, DcR1 completely lacks an intracellular domain and is anchored to the cell membrane via a phosphatidyl-inositol glycine (GPI) anchor (Degli-Esposti et al. 1997b, Pan et al. 1997a), and DcR2 has a truncated, inactive death-domain (Degli-Esposti et al. 1997a). The final "decoy" receptor is OPG, which is soluble and contains no death domain (Emery et al. 1998).

Overexpression of any one of the non-apoptosis-inducing ligands can protect sensitive cells from the cytotoxic effects of TRAIL (Degli-Esposti et al. 1997a, Degli-Esposti et al. 1997b, Emery et al. 1998, Pan et al. 1997a). Although this was originally assumed to be due to a titration of the TRAIL ligand away from the death-inducing receptors, other mechanisms have been proposed. Mérino *et al* have suggested that DcR1 and DcR2 function in different ways. They have demonstrated that due to its GPI attachment to the membrane, DcR1 is mainly found within lipid rafts and that in cells over-expressing this receptor TRAIL ligand is also mainly found in lipid rafts, while DR4 and DR5 are excluded. When they over-expressed DcR2 however, they found that it was recruited to the DISC upon TRAIL treatment and that DR4 was excluded. They found that Caspase-8 was recruited to this DR5/DcR2 DISC, but not activated (Merino et al. 2006). Some groups have found that binding of the TRAIL ligand to the DcR2 receptor can trigger the NF-κB pathway, which could lead to an increase in the expression of anti-apoptotic factors (Degli-Esposti et al. 1997a). This could suggest a wider role for DcR2 in TRAIL-mediated pathways.

However, these and other studies into the functions of the non-death-inducing ligands have relied upon overexpression, and so may not necessarily reflect the physiological

role of these receptors. Although some reports have claimed that normal cells express higher levels of the decoy receptors than transformed cells (Pan et al. 1997a), others have not found a correlation between sensitivity and decoy receptor expression either at the mRNA level(Kim et al. 2000) or expression on the cell surface(Zhang et al. 1999). Furthermore, several of the drug treatments that sensitise resistant tumour cells to TRAIL also increase expression on the surface of the decoy receptors (Ganten et al. 2004, Ganten et al. 2005).

Since studies showing the protective effect of decoy receptors rely on over-expression systems and no correlation has been found between sensitivity and expression of these receptors, the regulatory role of decoy receptors in a physiological situation remains unclear.

1.3.2.3 Bcl -2 and Bcl-XL

Bcl-2 and Bcl-XL are general anti-apoptotic factors which inhibit the action of Bax in inducing MOMP (Figure 1.1). Several overexpression studies have shown little effect of Bcl-2 and Bcl-XL on TRAIL-induced apoptosis (Gazitt, Shaughnessy & Montgomery 1999, Walczak et al. 2000), nor does BCL-2 levels correlate with TRAIL sensitivity, although they were isolated in an expression screen for genes involved in TRAIL-induced apoptosis (Burns, El-Deiry 2001).

1.3.2.4 The cFLIP protein

Cellular FLICE Inhibitory Protein (cFLIP, also known as CFLAR, Usurpin, Casper or FLAME) was first cloned using its homology to Caspase-8(Rasper et al. 1998). cFLIP is expressed as at least three different isoforms: $cFLIP_s$, $cFLIP_R$ and $cFLIP_L$ (Golks et al. 2005, Irmler et al. 1997, Rasper et al. 1998). The $cFLIP_S$ and $cFLIP_R$ proteins contain two DED domains, homologous to those in both Caspase-8 and FADD, and short N-terminal domains. In addition to the DED domains, $cFLIP_L$ contains a catalytically inactive caspase-like domain (Golks et al. 2005, Irmler et al. 1997).

Overexpression of cFLIP protects cells from the apoptosis-inducing ligand FAS by associating with Caspase-8 and FADD and preventing the assembly/activity of the DISC (Golks et al. 2005, Irmler et al. 1997, Rasper et al. 1998). Furthermore, siRNA mediated knock-down of cFLIP sensitises otherwise resistant NK cells to TRAIL-induced apoptosis (Mirandola et al. 2004) and mouse embryonic fibroblasts from $cFLIP^{-/-}$ mice have an increased sensitivity to FAS. However, there have been some reports that $cFLIP_L$ might actually activate Caspase-8 in some situations (Micheau et al. 2002).

Expression of cFLIP is widespread in most normal tissues, although expression was not detected in colon, placenta and testis samples (Rasper et al. 1998). High levels of cFLIP have been found to correlate with resistance to TRAIL-induced apoptosis in tumour cells. One study that found that 5 out 6 TRAIL-resistant transformed cell lines expressed cFLIP mRNA, but only 1 out of 5 sensitive lines expressed FLIP mRNA (Kim et al. 2000). However, a separate study found that few melanoma cells lines expressed cFLIP protein, irrespective of their TRAIL sensitivity, and that there was no correlation between protein expression levels and sensitivity (Zhang et al. 1999).

Sensitisation of cells to TRAIL via treatment with 5-FU involves a reduction both in cFLIP protein levels and its recruitment to the DISC (Ganten et al. 2004). In contrast, siRNA mediated knock-down of cFLIP leads to further sensitisation of cells treated with proteasome inhibitors, arguing that these compounds are acting though a different mechanism (Ganten et al. 2005).

Given that cFLIP has been shown to inhibit apoptosis and that its knock-down can sensitise cells to TRAIL-induced apoptosis, it would appear that cFLIP is an important regulator of sensitivity to apoptosis. Although not all studies have shown a correlation between cFLIP levels and sensitivity to TRAIL, this demonstrates that this is not the only mechanism by which cells can become resistant to TRAIL, not that it is a physiologically unimportant one. However, cFLIP is a general inhibitor of apoptosis and its effects are not restricted to TRAIL-induced apoptosis and so cannot entirely explain the sensitivity of some cells, but not others, to TRAIL-induced apoptosis when other ligands, such as FAS do not follow this pattern.

1.3.2.5 Inhibitor of Apoptosis Proteins (IAPs)

IAPs are characterised by the presence of BIR domains (baculoviral IAP repeat) and RING zinc-finger domain. XIAP, cIAP1 and cIAP2 bind directly to the executioner caspases and inhibit the processing of the pro-caspase form to the active form (Deveraux et al. 1997, Roy et al. 1997). Over-expression of these can lead to a general resistance to apoptosis(Deveraux et al. 1997). The RING zinc-finger domain is also involved in the ligation of ubiquitin to caspases therefore leading to their degradation (Suzuki, Nakabayashi & Takahashi 2001).

The IAP Survivin has been associated with resistance to TRAIL in renal cell carcinoma (Griffith et al. 2002) and overexpression of Survivin in sensitive, Survivin-negative cells can lead to resistance. Survivin is predominantly expressed during the G2/M phase of

the cell cycle and has been demonstrated to have a role in the control of the cell cycle. It co-localises with several microtubule associated factors, and knock-out mice show a failure of cytokinesis (Li et al. 1998, Uren et al. 2000). It has been suggested that Survivin may form a link between cell-cycle control and cell death with its expression during G2/M preventing a default cell death activation (Li et al. 1998). The role of BIR domain containing proteins in cell cycle regulation is conserved in a group of BIR containing proteins in the worm and in yeast. Worm embryos deficient in the protein BIR-1 are unable to compete cytokinesis and the null phenotype is indistinguishable from that of Aurora kinase null embryos (Fraser et al. 1999, Speliotes et al. 2000).

1.3.2.6 Regulation of TRAIL sensitivity by other signalling pathways

It is known that several other signalling pathways are connected to the TRAIL pathway. The NF-κB, RAS/MAPK/ERK, MYC, PKC and ATK/PKB pathways have all been implicated in the regulation of TRAIL sensitivity. Several of these pathways have also been shown to be regulated by the TRAIL pathway, suggesting a complex, interconnected regulatory network whose output is the decision to live or die.

The ligation of the TRAIL ligand to any of its receptors, with the exception of DcR1 and OPG, can lead to the activation of the NF-κB transcription factor in a process that involves the recruitment of the death domain kinase RIP to the DISC (Degli-Esposti et al. 1997a, Lin et al. 1999). Treatment of cells with proteasome inhibitors leads to a reduction in the activity levels of NF-κB. However, whether or not this is the cause of proteasome inhibitor-mediated sensitisation to TRAIL is controversial (Ganten et al. 2005, Ravi et al. 2001). NF-κB activity can also sensitise cells to TRAIL. This is accompanied by an up regulation of TRAIL receptor expression. The balance of pro- and anti-apoptotic signals is thought to be regulated by the ratio of c-Rel and RelA in the NF-κB dimer (Ravi et al. 2001)

Most studies into the sensitivity of cells to TRAIL have been conducted using either sensitive or resistant transformed cells lines, and as such we know more about the resistance of certain transformed cells to TRAIL than we do about the sensitivity of transformed cells compared with normal cells. Several groups have approached this problem using a system that involves creation of transformed cell lines from normal cells using defined genetic changes (Nesterov et al. 2004, Wang et al. 2004/5).

Immortalisation of HEK or fibroblast cells with the early region of the SV40 virus and a constitutively active telomerase does not sensitise these cells to TRAIL. However, the addition of an oncogenic mutant RAS transforms to these immortalised cells and sensitises

them to TRAIL-induced apoptosis. The surface expression of DR5, and thus the recruitment of Caspase-8 to the DISC, is also increased, suggesting a possible mechanism (Nesterov et al. 2004). A constitutively active MEK could substitute for oncogenic RAS in the sensitization, showing that the MAPK/ERK pathway was important down-stream of RAS. ERK2 signalling is required for the PG490-mediated sensitisation of resistant tumour cells (but not normal cells) to TRAIL (Frese et al. 2003). At the same time it was shown that over-expression of MYC could also sensitise immortalised fibroblast cells to TRAIL-induced apoptosis *in vitro* and *in vivo*. This also leads to the up regulation of DR5 expression on the cell surface (Wang et al. 2004/5). It has also been shown that stabilisation of MYC by knock-down of GSK3β or FBW7 leads to sensitisation (Rottmann et al. 2005).

It has been suggested that RAS sensitises cells to TRAIL by stabilising the MYC protein. This is supported by the fact that siRNA mediated knock-down of MYC removes the sensitizing effect of oncogenic RAS (Wang et al. 2005). Against the model that sensitisation is induced by RAS-mediated stabilisation of MYC and therefore an increase in DR5 surface expression is the fact that ERK2-mediated sensitisation of a cell to TRAIL does not involve an increase in DR5 surface expression (Frese et al. 2003). Indeed, as discussed above, some normal, insensitive, cell types have been shown to express DR5 on the surface. However, both RAS and MYC function through other mechanisms. Overexpression of both MYC and oncogenic RAS leads to a slight increase in levels of Caspase-8, FADD and BID in addition to DR5 (Nesterov et al. 2004, Wang et al. 2004/5). MYC also directly represses the transcription of cFLIP (Ricci et al. 2004), while RAS activation inhibits the translation of cFLIP via a RalA/cdc25 pathway (Panner et al. 2006).

Although a correlation has been shown between MYC levels and TRAIL sensitivity in TRAIL-sensitive cell lines, not all sensitive cells showed a high expression of MYC, and not all lines expressing a high level of MYC are sensitive(Ricci et al. 2004). Indeed, not all transformed cells have increased levels of MYC. This can be partly explained by the fact that the effects of MYC on DR5 levels are indirect (Wang et al. 2004/5). This means that it is possible that other effectors can work downstream of MYC.

Other signalling molecules that have been shown to affect the sensitivity of cells to TRAIL include Protein Kinase C (PKC) and the Akt/PKB pathway. PKC has been shown to regulate FADD recruitment to the DISC independently of receptor surface expression or FADD phosphorylation (Harper et al. 2003). Akt levels in transformed cells correlate with TRAIL sensitivity and cells expressing high levels of Akt do not cleave BID upon TRAIL treatment. This finding has been experimentally replicated by using dominant negative or

constitutively active Akt to regulate TRAIL sensitivity. (Chen et al. 2001, Thakkar et al. 2001)

Sensitivity of cells to TRAIL can be regulated at the level of receptor surface expression, decoy receptors and the expression of inhibitory proteins, such as cFILP. However, perhaps with the exception of cFLIP, these proteins do not seem to be able to mark the difference between resistant, normal cells and sensitive transformed cells. The fact that the resistance of some transformed cells can be explained by modulation of these factors suggests that the resistance of this subset has a different mechanism to the general resistance of normal cells. That is, transformation sensitises cells, some of which can escape by a secondary, acquired mechanism, such as losing expression of the TRAIL receptor. This is supported by the observation that many treatments which sensitise resistant transformed cells to TRAIL do not sensitise normal cells to TRAIL. Although cFLIP is widely expressed in normal cells, and high levels in tumour cells correlates with resistance, cFLIP inhibits apoptosis induced by any Caspase-8 dependent stimulus, and cannot explain, for instance, the sensitivity of some cells to FAS that are resistant to TRAIL. Many pathways involved in transformation have been implicated in TRAIL sensitivity, including NF-κB, RAS and MYC. However, the effects of these appear to be cell-type specific, and function through regulation of the same factors already discussed, although other, unknown mechanisms may exist. As such, they cannot provide an explanation as to why transformed cells generally are sensitive as changes in these pathways are not common to all transformed cell types. This leaves two options: 1) All pathways that lead to transformation regulate TRAIL sensitivity via independent mechanisms or 2) Some factor, common to all transformed cell types, remains to be found. This makes the regulation of TRAIL sensitivity an interesting and important problem to study.

1.3.3 Genes involved in TRAIL-induced apoptosis

The collation of information from the literature on the mechanism of TRAIL-induced apoptosis, work on determinants of the sensitivity of cell to TRAIL-induced apoptosis, and genes identified in the Aza-Blanc screen allows the production of a list of genes which, when knocked-down, should block TRAIL-induced apoptosis in sensitive cells. This list is presented in Table 1-2. The genes, which have been selected from the Aza-blanc screen, but are not reported elsewhere in the literature, are shown in a separate column since the lack of rigourous confirmation of hits in this study means the results must be treated with caution.

Confirmed positive controls	*Controls from* (Aza-Blanc et al. 2003)
TNFRSF10A (DR4)	BLK
TNFRSF10B (DR5)	PKM2 like
FADD	GSK3A
Caspase-8	CCDC139 (FLJ32312)
Caspase 3	FBXO11 (FLJ12673)
Bax	ROS1
Bid	ABL2
DIABLO (SMAC)	PRKRIR (DAP4)
SRP72	MAPK10 (JNK3)
MYC	TCF4
GSK3β	VPS16
	GUK1
	PRKCQ
	PRKAA2
	WDFY4 (FLJ00156)
	PRKCD
	IRAK1
	DVL2

Table 1-2 Genes involved required for TRIAL induced apoptosis
Table shows genes reported to be involved in TRAIL mediated apoptosis, which when knocked-down could be expected to protect cells from TRAIL-induced apoptosis. Genes from the general literature are separated from genes reported in Aza-Blac *et al*, owing to the lack of confirmation of genes from the screen. Names presented are HUGO names for genes, names in brackets are the names originally reported by Aza-blanc *et al*, or other commonly used names for the gene.

1.3.4 The physiological role of TRAIL

Although the cytotoxic effect of TRAIL against transformed cell lines has been intensively studied, less is known about the physiological role of TRAIL. Homozygous TRAIL knockout (TRAIL $^{-/-}$) mice do not show any grossly abnormal phenotype. They have normal tissue architecture, lymphoid cell homeostasis and bone density (Sedger et al. 2002). TRAIL does appear to function in both the innate and adaptive immune system, and has some role in tumour surveillance and anti-viral responses. There is also possibly conflicting data regarding TRAIL's involvement in auto-immune reactions and the maintenance of immune privileged sites.

Many cell types in the innate immune system up-regulate the expression of TRAIL upon stimulation, including monocytes, dendritic cells (DCs) and natural killer (NK) cells (Ehrlich et al. 2003, Sato et al. 2001, Takeda et al. 2001). There is also evidence that these cells exert their cytotoxic effect, especially against tumour cells, via TRAIL (Kayagaki et al. 1999, Kemp, Elzey & Griffith 2003, Liu et al. 2001, Sato et al. 2001, Takeda et al. 2001). This is particularly important for NK cells as this is their main function in the body. Indeed, liver NK cells express TRAIL constitutively (Takeda et al. 2001).

Although T and B cells can also express TRAIL upon stimulation and cytotoxic T

cells have been shown to exert their killing effect through TRAIL (Ehrlich et al. 2003, Janssen et al. 2005, Kayagaki et al. 1999), TRAIL also seems to have a role in the regulation of cells from the adaptive immune system. Helpless $CD8^+$ T cells, which do not require priming by $CD4^+$ T cells, do not undergo an expansion upon encountering their stimulating antigen a second time. This secondary expansion is suppressed by TRAIL and blocking TRAIL allows a second expansion to take place (Janssen et al. 2005).

The involvement of TRAIL in the cytotoxicity of immune cells against tumour cells indicates a role for TRAIL tumour surveillance mechanisms. $TRAIL^{-/-}$ mice do not have an increased incidence of tumours at an early age (Sedger et al. 2002), but they do have an increased risk of lymphomas if aged for a much longer period (Zerafa et al. 2005). Further, a lack of TRAIL in $p53^{+/-}$ mice leads to an increase in lymphomas and carcinomas. $TRAIL^{-/-}$ mice also show increased growth and metastasis of introduced tumours (Cretney et al. 2002, Sedger et al. 2002, Takeda et al. 2001). In one case it was demonstrated that this effect was due to the NK cells rather than cells of the adaptive immune system (Sedger et al. 2002). Consistent with this, the growth and metastasis of mammary tumours in a Her2/neu background was unaffected despite sensitivity to TRAIL. These tumours were MHC class I expressing and therefore would not have be subject to NK cell killing.

The issue of TRAIL's involvement with tumorigenesis is somewhat complicated by the finding that in certain situations TRAIL can promote tumour growth and metastasis. One study found that 50% of freshly isolated leukemia cell lines were resistant to TRAIL and that in a subset of these TRAIL actually reduced apoptosis and even promoted proliferation (Ehrhardt et al. 2003). *In vivo* work has also shown that TRAIL-resistant xenograft tumours in SCID mice are induced to metastasise by TRAIL treatment (Trauzold et al. 2006). Both groups found that this was mediated through TRAIL's ability to activate the NF-κB pathway, a pathway that can itself be pro- or anti-apoptotic (see 1.3.2.6). Thus it seems that TRAIL can induce a "proliferate or die" signal depending on the apoptotic sensitivity of the cell.

As well as its role in tumour surveillance, TRAIL also has an anti-viral role. Depleting TRAIL in mice that were infected with EMCV increased viral load (Sato et al. 2001) and is equivalent to the effect of depleting NK cells. It has also been shown that HIV-infected T-cells and macrophages are sensitive to TRAIL-induced apoptosis and that treatment with TRAIL reduces production of HIV mRNA and protein (Lum et al. 2001). Again, however, TRAIL seems to have effects in both directions, as $TRAIL^{-/-}$ mice have an enhanced resistance to MCMV infection (Diehl et al. 2004). This was shown to be due to an increased level of IL-2 and interferon-γ, suggesting a role for TRAIL in negative regulation of immune

cells.

Speculation on a possible role for TRAIL in the regulation of cells of the adaptive immune system led to a study of the possible effects of TRAIL on auto-immune disorders. Similar to tumorigenesis, it was found that the rate of spontaneous auto-immune disease was similar in TRAIL$^{-/-}$ mice to wild-type mice, but deficient mice were more sensitive to induced auto-immune disorders, including collagen-induced arthritis and streptozotocin-induced diabetes(Lamhamedi-Cherradi et al. 2003). Systematic administration of TRAIL can reduce the effects of experimental autoimmune encephalomyelitis (EAE), an experimentally induced model of multiple sclerosis in which immune cells infiltrate the central nervous system (CNS) leading to the destruction of the myelin around axons (Hilliard et al. 2001). Interestingly, direct blockage of TRAIL in the brain by injection of a modified TRAIL blocking, soluble TRAIL receptor leads to a reduction in the severity of the disease (Aktas et al. 2005), suggesting a duel role for TRAIL: firstly in regulating auto-immune T-cells and secondly in action of auto-immune T-cells on their targets.

The regulation of auto-immune T-cells in the brain suggests a role for TRAIL in the maintenance of so called immune privileged sites. The brain has very low numbers of T-cells and is one of a number of sites where cells of the immune system are excluded, presumably because the effects of inflammation would be very damaging. The brain does contain a small number of CNS-derived dendritic cells (CNS-DCs). However, unlike normal DCs, CNS-DC do not act to stimulate naive T-cells. Further, they actively block the proliferation of activated T-cells, a function that is dependent on TRAIL (Suter et al. 2003). TRAIL has also been implicated in the establishment of immune privilege at the interface between mother and placenta (Phillips et al. 1999).

This evidence suggests that TRAIL has a multitude of roles, both in the regulation and action of cells of the immune system. In addition to its role in tumour surveillance, it also has roles in anti-viral responses, auto-immune reactions and the establishment of immune privilege. However, the role of TRAIL is often complicated and contradictory. In some cases this is due to its ability to stimulate proliferation via the NF-κB pathway, in others, to the fact that it is both an effector and a regulator of the immune system.

1.3.5 Clinical prospects for TRAIL

There were initially very high hopes for TRAIL as an anti-cancer drug. Its ability to induce apoptosis in tumour cells, but not normal cells, gave it an obvious advantage over other apoptosis-inducing ligands such as FAS and TNFα, which lead severe effects on

normal tissue. Anti-Fas antibodies cause massive apoptosis of mouse liver cells *in vivo* (Ogasawara et al. 1993). Systemic TNFα, administration leads to a sepsis-like syndrome (Lejeune et al. 2006) and this has limited its use to limb salvage by regional limb perfusion in soft-tissue sarcoma treatment.

Pre-clinical studies showed promising remission of xenograph tumours in nude mice without toxicity (Ashkenazi et al. 1999, Sedger et al. 2002). However, this was challenged by the finding that normal human hepatocytes, not hepatocytes from mice or non-human primates were sensitive to recombinant TRAIL (Jo et al. 2000). It was noted that this study used a his-tagged version of the recombinant TRAIL and also that the native TRAIL dimer contained a zinc ion, while the his.TRAIL had not been optimised for physiological zinc content. Two groups later showed that the toxicity observed could be due to the tag, and that un-tagged recombinant TRAIL with the correct zinc content did not induce this apoptosis in normal cells. (Lawrence et al. 2001, Qin et al. 2001). Another group has confirmed this finding and also shown that hepatocytes are not sensitised to TRAIL on proteasome inhibition (Ganten et al. 2005).

Several TRAIL-based treatments are now in clinical trails (Reviewed in Duiker et al. 2006). Phase I clinical trials with a recombinant TRAIL have been completed and Phase II trials have recently been initiated. As well as TRAIL itself, several trials have been initiated using antibodies which target the TRAIL receptors. HGS-ETR1 targets DR4 and there are positive preliminary results for two phase I and three phase II trials using HGS-ET1 as a treatment for non-Hodgkin's lymphoma, colorectal cancer and non-small-cell lung cancer. HGS-ETR2 targets DR5. Preliminary results from two phase I trials show that although high doses of the antibody can lead to adverse effects including renal failure, possibly due to liver malfunction, lower doses show minimal toxicity. Phase I trials have also been initiated with a second DR5 targeting antibody known as HGS-TR2J.

There is hope TRAIL's ability to synergise with other treatments, such as those that activate the p53 pathway or proteasome inhibitors, may lead to more effective treatments for tumours refractory to either of these treatments alone. So far, the only trials initiated are phase I trials for HGS-ETR1 combined with gemcitabine and cisplatin, and with paclitaxel and carboplatin. Initial results are apparently promising.

1.4 Aims

High-throughput genetic screens represent an important route to add functional annotation to genes identified within the genome. In mammalian cell culture, gene levels can

be increased by expressing open-reading frames from constitutive promoters on introduced plasmid vectors. RNAi presents a novel way to reduce gene function on a large scale. These techniques present us with the tools to necessary to carry out genome-scale screens in mammalian cell culture.

The TRAIL apoptosis system is an interesting and medically important pathway which is amenable to study using such genome-scale gene-perturbation studies. Previous experiments have shown that the pathway can be studied using over-expression and RNAi-mediated gene-knockdown. A body of knowledge already exists, which allows the assessment of the success of screens. However, our understanding of the regulation of the pathway is incomplete, and this provides an opportunity for new discoveries to be made.

Screens can be carried out in a number of ways, using a range of different reagents. This work focuses on materials that are widely available, without recourse to specialist equipment or custom reagents. Although this may mean that better reagents and paradigms could have been constructed, ultimately the usefulness of these techniques as everyday tools for examining gene function relies on the use of standard, widely available resources.

Thus the aims of this thesis are:

1. To assess different methods for carrying out genome-scale RNAi screens using an assay for TRAIL-induced apoptosis.
2. To carry out screens to identify new genes in the TRAIL-induced apoptosis pathway and rigorously confirm hits for their reproducibility and specificity.
3. To assess the success and usefulness of the screens and the methodologies with which they were carried out for elucidating gene function.