4 AN SIRNA SCREEN OF 897 KINASES AND PHOSPHATASES

In Chapter 3 an assay for sensitivity to TRAIL-induced apoptosis was developed, and this assay was used to compare different methodologies for conducting an RNAi screen for genes involved in the TRAIL-induced apoptosis pathway. It was concluded that in this system a gene-by-gene siRNA screen is the most powerful approach. In this chapter these findings are applied to a screen of 897 genes classified as either protein kinases, protein

phosphatases or kinase/phosphatase associated. This screen serves both as a gene discovery experiment in its own right and as a pilot for possible larger screens. A rigorous confirmation strategy is employed to eliminate false positives. Attempts are made to define the point in the apoptotic pathway at which they act and the specificity of their regulation of apoptotic pathways is examined.

4.1 Introduction

4.1.1 Kinases and Phosphatases

The cell is a highly complex ordered system which depends on the correct regulation of the protein components of cellular pathways. The activity of proteins can be altered in many ways, but one of the most common is by the addition and removal of phosphate groups. Phosphate groups can alter the activity of proteins in several ways. Firstly, the process of phosphorylation has a large free energy. As such phosphorylation can radically alter the preferred conformational state of a protein. Secondly, the phosphate group carries two electrical changes and can accept three hydrogen bonds, which can alter the strength and specificity of the binding of the protein to other proteins and to substrates.

The phosphorylation state of proteins is controlled by kinases and phosphatases. Kinases catalyse the transfer of phosphate from a mono-nucleotide (usually ATP) to specific serine, threonine or tyrosine residues in the following, essentially irreversible, reaction:

$$XOH + ATP \xrightarrow{Kinase} XPO_4^{2-} + ADP$$

Kinases are the largest protein family in the genome and contain a characteristic 250 amino acid kinase domain. They are involved in many aspects of cellular regulation including control of the cell-cycle and mitogen-activated kinase signalling pathways.

The dephosphorylation of proteins is catalysed by protein phosphatases, which catalyse the removal of phosphate groups from protein in the following, again irreversible, reaction:

$$XPO_4^{2-} + H_2O \xrightarrow{Phosphatase} XOH + HPO_4^{2-}$$

Note that these two activities are not quite the opposite of each other, with the overall reaction of $ATP+H_2O \rightarrow ADP + HPO_4^{2-}$ effectively consuming one ATP molecule. There are fewer phosphatases in the genome than there are kinases, and the specificity of phosphatases is often regulated by phosphatase regulatory subunits.

While information in the apoptotic pathway is usually passed by proteolysis rather

than phosphorylation, many kinases have been implicated in regulation of apoptotic pathways in general and the TRAIL-induced apoptotic pathway in particular (see section 1.3.2.6). Indeed the Aza-Blanc screen concentrated mainly on kinases (Aza-Blanc et al. 2003). Repeating this should allow of an investigation of the reproducibility of results from RNAi screens. The involvement of kinases in the regulation of apoptotic pathways would also imply the involvement of phosphatases and their regulatory subunits.

4.1.2 The Library

The library used in this chapter is the Qiagen Kinase/Phosphatase library, which is a subset of the Qiagen Druggable Genome library v2. The library contains siRNAs targeting 691 kinases (there are 805 genes listed in Ensembl 46 associated with the GO term "Kinase activity") and 206 phosphatases and phosphatase associated proteins (159 genes associated with the GO term "Protein phosphatase activity" and 59 with "Protein phosphatase regulatory activity"). Each gene is targeted by 2 siRNAs designed using a neural network based algorithm (Huesken et al. 2005) and checked for specificity using a proprietary similarity based algorithm. The library is supplied on 24 96-well plates with one siRNA per well. Each plate contains 80 sample siRNAs and 2 non-targeting controls, a scrambled siRNA (QiaNeg) and an siRNA targeting GFP (siGFP), leaving 14 empty wells which can be used for controls.

4.1.3 Analysis of screening data

The aim of the analysis pipeline for siRNA screening data is to take raw assay outputs and produce a single score for each siRNA. The simplest analysis schemes calculate the number of standard deviations the average value of a well is from the mean value for the screen. A more comprehensive scheme has been proposed by Boutros *et al* and implemented in the R/Biocondutor package cellHTS. There are five steps from raw data to final score (Boutros, Brás & Huber 2006) :

1. **Production of a metric for each well.** Here that metric is survival calculated as described in Equation 6.
2. **Transformation of data.** A transformation such as a logarithmic transformation may be applied to the data.
3. **Plate normalisation.** As seen in Figure 3.13, data often shows plate to plate variation, and therefore data must be normalized before results from different plates can be directly compared.
4. **Standardisation.** Scores are standardized to produce a z score for each well. This

can be done using the mean and standard deviation, or using the median and median absolute deviation.

5. **Summary of replicates.** A function must be applied to produce one score from a number of screen replicates.

There are several different alternatives for plate normalization. Data maybe normalized using the values of the negative controls on each plate. A related method to this is the normalized percentage inhibition which rescales data so that the negative control has a value of 0 and the positive control a value of 1:

$$NPI(x) = \frac{x - \mu_{negatives}}{\mu_{positives} - \mu_{negatives}} \times 100\%$$

Equation 1

where: $\mu_{positives}$ is the geometric mean of the positives controls and $\mu_{negatives}$ is the geometric mean of the negative controls.

This method makes no assumption about the effect size of the samples, but is very sensitive to variance in the control wells. If there is an assumption that only a small number of siRNAs on each plate are "hits", then data maybe normalized to some measure of the central tendency of the sample data, such as the mean or median of samples on the plate. If the median is used, this method provides normalization that is robust with respect to small numbers of outliers. It does however rely on only a small number of siRNAs on each plate eliciting a strong phenotype. One situation in which this may not be the case is if the arraying of genes within the library is not random.

Different functions can also be applied to summarize the replicates. The most common are the mean, the minimum and the maximum. The use of the minimum is the most conservative summary. It asks the question: is the score of this siRNA high in every replicate? The maximum is the least conservative summary; it asks the question: is the score of this siRNA high in any of the replicates? The mean is between these two extremes. Other options include the calculation of p values using t-tests if there is sufficient data.

The cellHTS package provides a frame-work for carrying out each of these steps, calculates plate and screen level quality controls. Results are presented as a single R object and as a series of linked HTML reports.

4.2 Screen design, execution and data analysis

In order to screen the library of siRNAs for genes that affect the TRAIL-induced apoptosis pathway, aliquots of siRNAs from the library plates were transferred to new plates and transfected into HeLa cells. In order to reduce variability all cells used had been passaged exactly four times after defrosting. To each plate of siRNAs the following controls were added (Figure 4.1):

- **siNeg, QiaNeg, and siGFP**. Assay negative controls. These controls were designed not to target any gene which is expressed in human cells.
- **siCasp8, siBID, siSMAC**. Assay positive controls. These controls are siRNAs known to reduce the sensitivity of HeLa cells to TRAIL-induced apoptosis.
- **siKIF11**. RNAi positive control. siKIF11 induces growth arrest and cell death upon transfection. This allows that extent of the success in transfecting cells and inducing RNAi to be estimated.
- **No transfection**. Transfection negatives. Allows assessment of the effect of transfection on the assay
- **Blank**. Assay blanking wells. These wells contain no cells and allow for the assay to be zeroed.

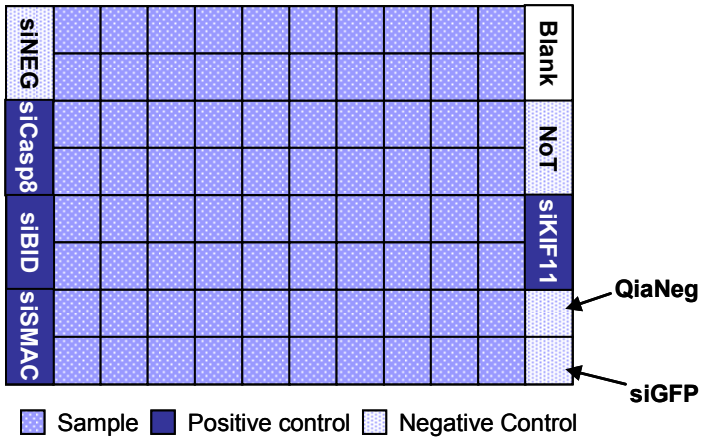

Figure 4.1 Plate layout used in screen

Plates were transfected in batches, and grown for 48 hours. Viability was assessed using alamarBlue. Cells were then treated with 0.25µg/ml TRAIL for 24 hours before viability was reassessed. Throughput was increased as the screen progressed, starting at six plates per day and reaching a maximum throughput of 24 plates per day.

The "quality" of the results from a screening plate can be judged by the dynamic range of the positive and negative controls, defined as the ratio of the geometric means of

the survival in siCasp8 transfected wells to the geometric mean survival of negative control transfected wells. A dynamic range of 2 means that the average survival in siCasp8 transfected wells is twice that in negative control wells. A quality control threshold for acceptable plates was arbitrarily set at a dynamic range of 2. The initial design required three replicates of the screen to be completed. After the completion of 2.75 replicates of the screen it was observed that the dynamic range of plates from the first of the replicates was lower than that found in plates from the other two replicates (Figure 4.2). The percentage of plates failing this quality control criterion for each replicate was 44%, 4% and 4% respectively. In all but two cases the quality of each plate was higher in replicates two and three than in replicate one. Including data from replicate one would therefore reduce the overall quality of data. On this basis it was decided to combine the three replicates to produce two "high-quality" replicates. Data from the first replicate were discarded, except in those cases where the dynamic range between siCasp8 and negative control transfected wells was less than 2 in one or other of the two remaining plates, and the plate from the first replicate had a higher dynamic range. In this case the plate with a dynamic range of less than 2 was discarded and replaced with the plate from the first replicate In this way the screen simulated a situation where the screen was performed in duplicate, and plates failing a quality control measure were repeated. Only 4% of plates in the resulting two replicates had dynamic ranges of less than 2.

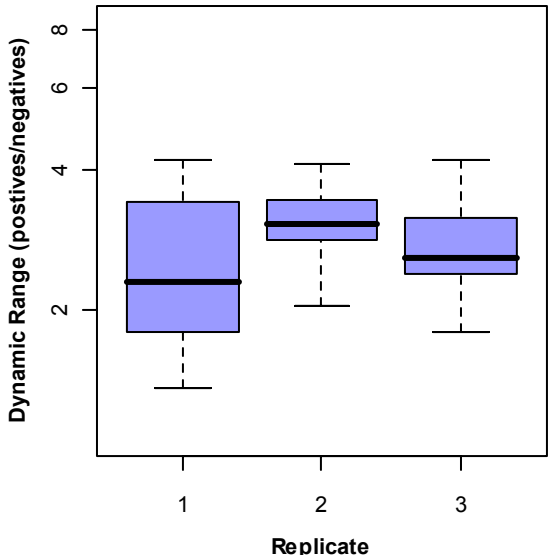

Figure 4.2 Boxplot of dynamic ranges of plates from different screen replicates
The dynamic range of results from each plate was calculated as the ratio of the geometric mean of survival in siCasp8 transfected wells to the geometic mean of survival in negative control transfected wells. Plates were grouped by the replicate of the screen from which they originated. n = 18, 24 and 24 for the first, second and third replicates respectively.

In order to investigate the effects of different plate normalization techniques, unnormalized data (Figure 4.3a) was compared to data normalized using the geometric mean of the plate negative controls (Figure 4.3b), the median of the survivals for the samples on each plate (Figure 4.3c) and using normalized percentage inhibition (Figure 4.3d). Examination of these plots indicates that of the different normalisation methods median normalization produces the most consistent results, with greater consistency in the position of the quartiles as well as the measures of centre. Data normalized using the geometric mean of the negative controls is little different from unnormalized data, while data normalized using NPI is more consistent, but has a few very obvious outliers.

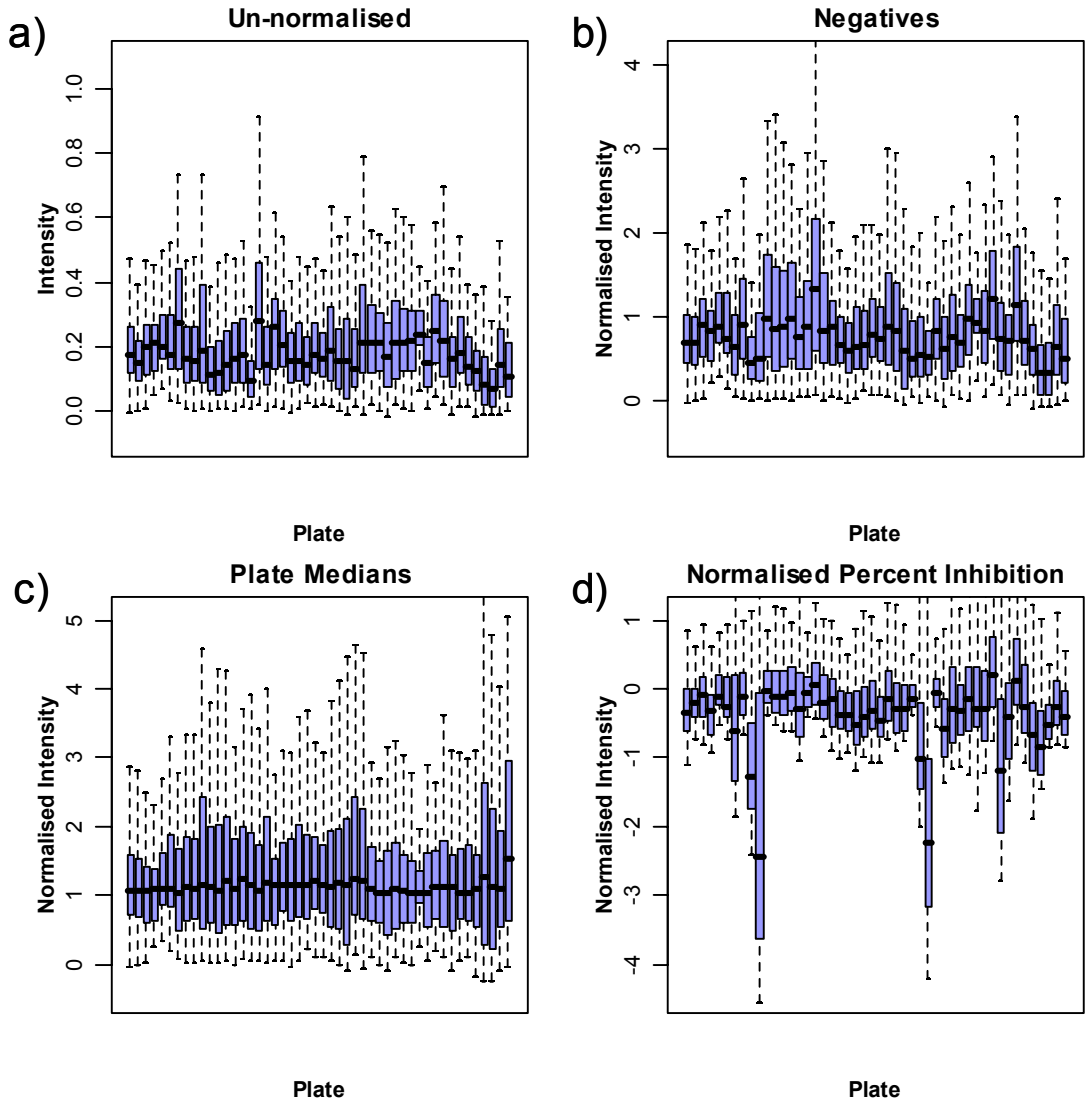

Figure 4.3 Normalization of data from screen of Kinases and phosphatases.
Boxplots showing values for each plate from a) un-normalised data, b) data normalized by dividing values by the geometric mean of negative controls, c) data normalized by dividing by the median of samples of each plate and d) data normalised using NPI as defined in Equation 1.

A strong relationship between effect size and variance indicates non-normally distributed data. If there is a strong relationship between effect size and variance then this

may be reduced by applying a logarithmic transformation to the data. To investigate whether a log transformation would improve the interpretability of data here, the rank of the mean of the normalized values for each siRNA was plotted against the standard deviation between replicates (Figure 4.4). A clear relationship can be observed between the mean and the standard deviation in both non-transformed and log transformed data. In non-transformed data the standard deviation increases as the mean increases (Figure 4.4a). This is the expected behaviour for ratio data. Log transforming the data reverses the trend with smaller means having larger standard deviations than larger means (Figure 4.4b). The running median line, shown in red, should be flat if there is no relationship between the mean and the standard deviation. The deviation of the line from flat indicates the strength of the relationship. Comparing the running median line for non-transformed data and log-transformed data leads to the conclusion that the strength of the relationship is not reduced, and may even be increased by log transformation of the data; therefore data was used untransformed in further analyses. However, this means that data cannot be treated as normal and so the normal distribution cannot be used to determine p values representing the probability that any given siRNA affects the sensitivity of cells to TRAIL-induced cytotoxicity, when compared to the distribution of siRNA effects.

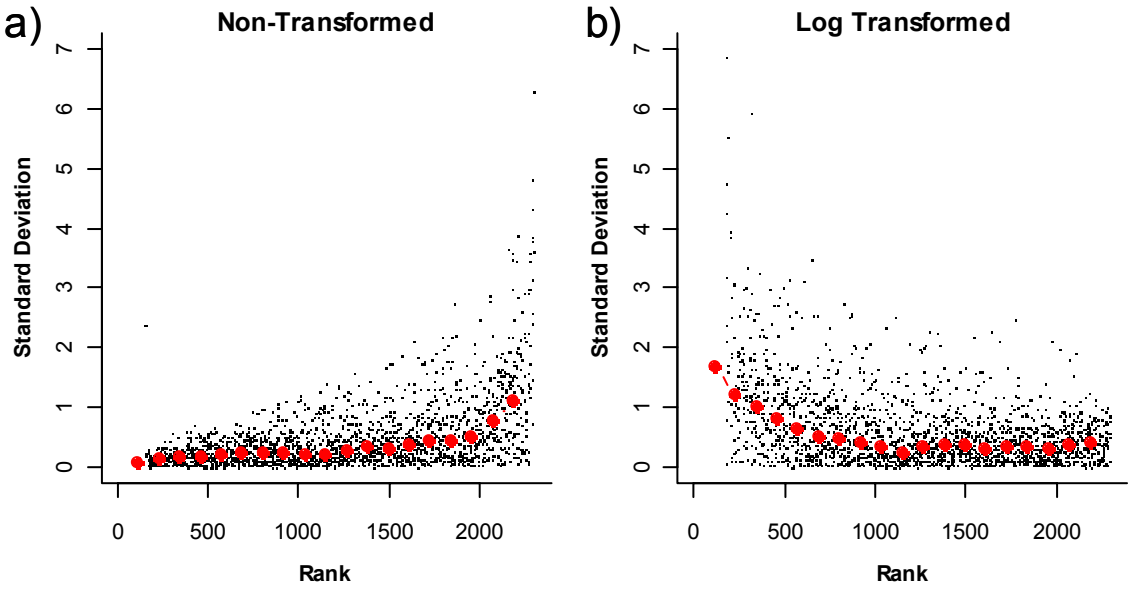

Figure 4.4 Relationship between standard deviation and rank of the mean for siRNAs.
For each siRNA the mean of the normalized data for each replicate was calculated. The rank of this mean was then plotted against the standard deviation between the replicates for a) Non-transformed data and b) Transformed data. The red line in each plot represents the running median standard deviation.

Given that only two replicates of the screen were conducted, it was decided to use the minimum of the replicates to summarise the data. This means that a high final score represents an siRNA which has large effect which repeated in both replicates.

Data was processed using the R/Bioconductor package cellHTS using a median

normalization without log transformation and the minimum as a summary function. The HTML report produced can be found on the CD accompanying this work or at web address http://www.sanger.ac.uk/HGP/Chr22/RNAi/TRAIL_Kinase/.

4.3 Screen quality and analysis of controls

The processed screening data was used to assess the quality of the screen (Figure 4.5). The Pearson's correlation co-efficient between the two repeats of the screen is 0.65 and spearman's $\rho = 0.66$. Both correspond to a p value of less than 2.2×10^{-16}, showing that this is a real correlation. The replicates are less well correlated at higher scores than at lower scores (Figure 4.5a), this is expected from the finding that variance increases with increasing mean (Figure 4.4a). The correlation between two siRNAs targeting the same gene is much weaker (Figure 4.5b), with a Pearson's correlation co-efficient of 0.2 and a spearman's ρ of 0.17. The weakness of this correlation suggests that while the effect of a particular siRNA on the sensitivity of cells to TRAIL-induced cytotoxicity is fairly reproducible, the effect of different siRNAs targeting the same gene is not. There are two possible explanations for this: the effectiveness of siRNA A and siRNA B at knocking down the targeted gene could vary, or alternatively siRNA A and siRNA B have a number of different off-target effects.

Figure 4.5c shows the distribution of plate dynamic ranges (ratio of the geometric mean of the siCasp8 transfected wells to the geometric mean of the negative control transfected wells). The majority of plates (83%) have a dynamic range between 2 and 4, and only two plates (4%) have a dynamic range of less than 2.

Figure 4.5d shows the distributions of scores for different well types and some of the same data is summarized in Table 4-1. The median score for all positive controls is greater than either the negative controls or samples. As expected from data in Chapter 3 the size of the difference between the negative control transfected wells and siBID or siSMAC transfected wells is smaller than the difference between negative controls transfected wells and siCasp8 transfected wells. The spread of results for siCasp8 is larger than that for other controls, possibly due to the relationship between mean and variance (Figure 4.4a). The Z' factors for comparing siCasp8, siBID and siSMAC to the negative controls are -0.35, -1.05 and -4.08. This is a large reduction on Z'-factor seen for siCasp8 in the previous chapter, although it is close to 0 when compared to the Z'-factors for siBID and siSMAC. In the case of siSMAC it is clear that even though the median score and survival (0.86 and 28.51% respectively) are higher than those for the negative controls (0.27 and 23% respectively) there is little chance of separating the individual values of negative control transfected wells and

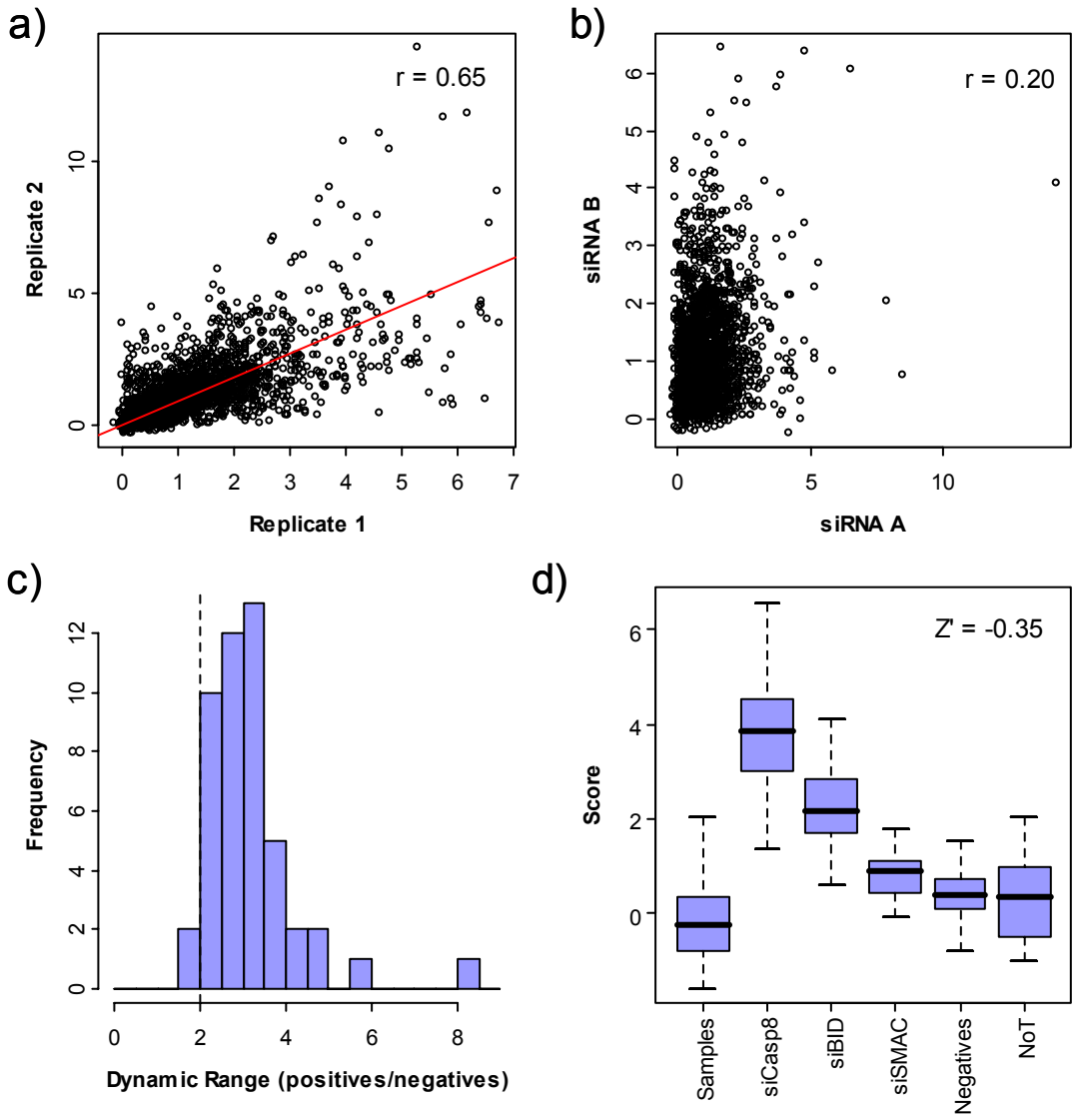

Figure 4.5 Assesment of screen quality and controls.
a) Normalized survival from replicate 1 plotted against normalised survival from replicate 2. Red line shows linear regression of replicate 2 on replicate 1. The Pearson's correlation co-efficient is shown in the top right corner. b) Plot showing normalised survival of the two siRNAs targeting the same gene. The Pearson's correlation co-efficient is shown in the top right corner. c) Histogram showing the distribution of plate dynamic range (see above for definition of plate dynamic range). Dashed line represents a dynamic range of 2. d) Box plot summarising the scores in different well types. Z'-factor between negative controls and siCasp8 is shown in top right corner. NoT = Untransfected

siSMAC transfected wells. This analysis of the positive controls shows that while the screen may be sensitive enough to pick up siRNAs with an effect as strong that of siCasp8, more subtle effects, such as those elicited by siSMAC are likely to be missed.

The scores and survivals of the non-transfected wells (0.32 and 24.9% respectively) are similar to those for the negative controls although the non-transfected wells have a higher spread of values (Figure 4.5d and Table 4-1). The median value of the samples is slightly reduced compared with the negative controls (-0.26 and 15.20% median score and median survival respectively), suggesting that either the library contains some siRNAs that actively increase sensitivity to TRAIL-induced cytotoxicity, or that one of the negative

Category	*Median Score*	*Median Survival*	*MAD Survival*
Samples	-0.26	15.20%	12.20%
siCasp8	3.788	64.00%	16.37%
siBID	2.2	48.32%	8.44%
siSMAC	0.86	28.51%	6.66%
Negatives	0.27	23%	8.47%
Untransfected	0.32	24.90%	16.18%

Table 4-1 Summary statistics of controls in Kinase and Phosphatase screen.
The median score, median survival and the median absolute deviation (MAD) of survival is shown

controls is triggering an off-target effect that is reducing the sensitivity to TRAIL-induced cytotoxicity.

4.4 Screen results

The results of the screen, in the form of single scores for each siRNA, are summarized in Figure 4.6. The distribution of scores has a long right hand tail and a foreshortened left hand tail (Figure 4.6a) compared to a normal distribution. The long right hand tail represents siRNAs that have reduced the sensitivity of cells to TRAIL. The foreshortened left hand tail is probably due to the lower limit of the measurement used i.e. it is not possible for less than 0% of cells to survive.

In order to further examine the distribution of the screening results, the rank of siRNA scores was plotted against the score for that siRNA (Figure 4.6b). The distribution of scores is continuous and the rate of increase in score with rank is constant for a large portion of the plot. The scores from this portion of the ranked list cover a large proportion of the observed range of scores. This implies that siRNA cannot be divided into two distinct classes: those that have an effect on sensitivity to TRAIL-induced cytotoxicity and those that do not have an effect, but that each siRNA has a more or less strong effect on the sensitivity of cells to TRAIL.

Figure 4.6c shows the spatial distribution of scores within the library in order to test for plate position effects. High and low scores are relatively evenly distributed between and within plates and there are not obvious signs of edge effects.

siRNAs were ranked by their score in the screen. A portion of the resulting table is shown in Table 4-2. The complete table is available as part of the screen report, on the included CD or online at http://www.sanger.ac.uk/HGP/Chr22/RNAi/TRAIL_Kinase/. The results of the screen were also summarized on a gene-by-gene basis. A portion of this summary is shown in Table 4-3.

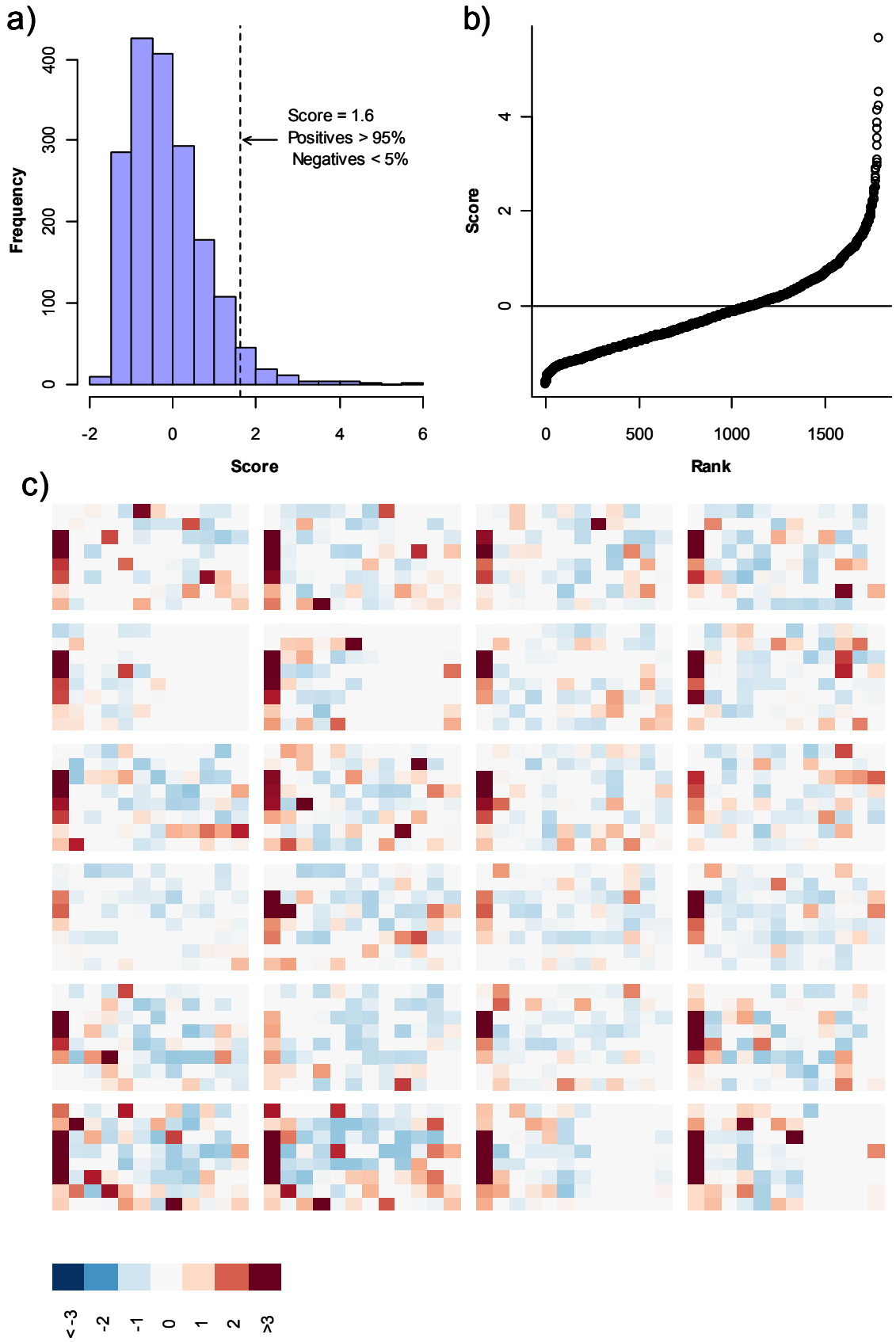

Figure 4.6 Results of siRNA screen of Kinases and Phosphatase
a) Histogram of scores from sample wells. Dashed line represents a score of 1.6. b) Rank of siRNA score plotted against score. c) Heat map of scores per plate. siRNAs with a highly positive score are shown in red, siRNAs with a highly negative score are shown in blue. Plates are arranged row wise

RefSeq ID	*Symbol*	*Description*	*Normalized Survival*		***score***
			Rep 1	Rep 2	
NM_030974	Sharpin	shank-interacting protein-like 1	5.3	14.2	**5.46**
NM_005541	INPP5D	inositol polyphosphate-5-phosphatase, 145kDa	4.8	4.9	**4.81**
NM_001396	DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	4.6	4.2	**4.21**
NM_018401	STK32B	serine/threonine kinase 32B	4.2	7.9	**4.10**
NM_198828	LOC375449	similar to microtubule associated testis specific serine/threonine protein kinase	4.0	4.3	**3.79**
NM_005399	PRKAB2	protein kinase, AMP-activated, beta 2 non-catalytic subunit	6.1	3.8	**3.65**
NM_014002	IKBKE	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase epsilon	3.7	4.1	**3.46**
XM_086287	PTPRV	protein tyrosine phosphatase, receptor type, V	3.6	3.8	**3.31**
NM_020791	TAOK1	TAO kinase 1	4.2	3.5	**3.28**
NM_007079	PTP4A3	protein tyrosine phosphatase type IVA, member 3	3.6	8.5	**3.24**
NM_006240	PPEF1	protein phosphatase, EF hand calcium-binding domain 1	3.5	5.2	**3.21**
NM_001570	IRAK2	interleukin-1 receptor-associated kinase 2	3.7	3.4	**3.08**
NM_002827	PTPN1	protein tyrosine phosphatase, non-receptor type 1	3.4	4.3	**3.04**
NM_001556	IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta	5.5	3.3	**2.98**
NM_016086	DUSP24	dual specificity phosphatase 24 (putative)	3.3	6.4	**2.88**
NM_173354	SNF1LK	SNF1-like kinase	3.2	3.7	**2.83**
NM_018638	ETNK1	ethanolamine kinase 1	3.1	6.4	**2.69**
AB033076	KIDINS220	likely homolog of rat kinase D-interacting substance of 220 kDa	3.0	3.9	**2.59**
NM_004577	PSPH	phosphoserine phosphatase	3.0	3.3	**2.57**
NM_022128	RBKS	ribokinase	3.0	3.0	**2.49**

Table 4-2 Top scoring siRNAs from an siRNA screen of Kinases and Phosphatases.
Table shows top 20 scoring siRNAs from the screen. The complete table is available on the included CD or online at http://www.sanger.ac.uk/HGP/Chr22/RNAi/TRAIL_Kinase

RefSeq ID	*Gene Symbol*	*Description*	Score siRNA A	Score siRNA B	Minimum Score	Maximum Score	Mean Score
NM_020836	KIAA1446	Brain-enriched guanylate kinase-associated protein	2.42	2.45	2.42	2.45	2.44
NM_000788	DCK	Deoxycytidine kinase	2.46	2.15	2.15	2.46	2.30
NM_030974	Sharpin	Shank-interacting protein-like 1	5.46	2.15	2.15	5.46	3.81
NM_018638	ETNK1	Ethanolamine kinase 1	1.81	2.69	1.81	2.69	2.25
NM_021132	PPP3CB	Protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform (calcineurin A beta)	1.82	1.65	1.65	1.82	1.74
NM_173354	SNF1LK	SNF1-like kinase	1.42	2.83	1.42	2.83	2.13
NM_018401	STK32B	Serine/threonine kinase 32B	4.10	1.35	1.35	4.10	2.72
NM_001896	CSNK2A2	Casein kinase 2, alpha prime polypeptide	1.32	1.81	1.32	1.81	1.57
NM_001556	IKBKB	Inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta	1.22	2.98	1.22	2.98	2.10
NM_183048	PRKCBP1	Protein kinase C binding protein 1	1.48	1.21	1.21	1.48	1.35
NM_033118	MYLK2	Myosin light chain kinase 2, skeletal muscle	1.21	1.45	1.21	1.45	1.33
NM_145203	CSNK1A1L	Casein kinase 1, alpha 1-like	1.44	1.18	1.18	1.44	1.31
NM_017823	DUSP23	Dual specificity phosphatase 23	1.66	1.17	1.17	1.66	1.42
NM_019892	INPP5E	Inositol polyphosphate-5-phosphatase, 72 kda	1.17	1.09	1.09	1.17	1.13
NM_007064	TRAD	Serine/threonine kinase with Dbl- and pleckstrin homology domains	1.07	2.21	1.07	2.21	1.64
NM_003558	PIP5K1B	Phosphatidylinositol-4-phosphate 5-kinase, type I, beta	1.07	1.94	1.07	1.94	1.50
NM_014683	ULK2	Unc-51-like kinase 2 (C. Elegans)	1.06	1.52	1.06	1.52	1.29
NM_004443	EPHB3	EPH receptor B3	1.04	1.09	1.04	1.09	1.07
NM_001798	CDK2	Cyclin-dependent kinase 2	1.10	0.99	0.99	1.10	1.04
NM_020639	RIPK4	Receptor-interacting serine-threonine kinase 4	0.94	1.58	0.94	1.58	1.26

Table 4-3 Extract from table summaring screen results on a per gene basis.
Genes are ranked on the basis of the minimum of the scores from the two siRNAs. Full table is available on included CD or online at http://www.sanger.ac.uk/HGP/Chr22/RNAi/TRAIL_Kinase/perGene.tab

4.4.1 Hit selection

The continuous distribution of scores makes distinction of "hit" genes from the negative genes difficult (and to a certain extent meaningless). However, in order to select genes for confirmation and follow up it is necessary to apply a cut off. A cut off was selected on the basis of the distributions of the siCasp8 positive control transfected wells and negative control transfected wells. In order to maximise the number of potential hits selected from this first pass screen, a cut-off was selected such that 95% of siCasp8 transfected wells score higher than the threshold. A cut off score of 1.6 selects 95% of siCasp8 transfected wells, but only 5% of negative control transfected wells (Figure 4.6a). 71 siRNAs targeting 66 genes have a score of greater than 1.6. In this way these siRNAs are identified as 'hit' siRNAs.

It is important to note that identifying an siRNA as a hit siRNA does not necessarily identify the gene it is targeted by as a hit gene. This is due to the possibility that the effect of the siRNA maybe due to off-target effects.

One possibility for selecting hit 'genes' is to look for genes where both siRNAs targeting the gene are hits – that is where both siRNAs score higher than 1.6. This method selects 5 genes as candidate hit genes for confirmation and follow-up. A second possibility for selecting hits for follow-up would be take genes targeted by a single scoring siRNA. Using a similar method for selecting a cut as was used above would make this a less conservative method for selecting hits. However, limitations on time and resource would make selecting all 66 genes targeted by a single siRNA scoring higher than 1.6 unfeasible. Therefore a smaller number of genes must be selected for confirmation and follow-up depending on the resources available. Here genes targeted by the top ten scoring siRNAs were selected for confirmation and follow-up as it was assumed that these were the genes most likely to confirm. Unfortunately, while this method does allow for a definitive answer for the involvement these ten genes, it offers no information on the other 55 genes targeted by the remaining 61 siRNAs: i.e. it is incorrect to say that they are not hit, simply that they are untested.

It is possible that where an siRNA scores below zero, transfection of the siRNAs maybe result in an increase in the sensitivity of cells to TRAIL-induced cytotoxicity. It might be interesting to follow-up some of these genes. However, the assay and the data analysis protocol was designed and optimised to find siRNAs which decrease the sensitivity of cells to TRAIL. Given limited resources it was decided that following up the hits that the screen

was designed to find was a more likely to yield interesting results, and so siRNAs that appear to increase sensitivity to TRIAL were not further investigated.

4.4.2 Analysis of genes previously associated with the TRAIL pathway.

The performance of siRNAs targeting genes previously associated with the TRAIL pathway provides a means of assessing the sensitivity of the screen. The results from the kinase and phosphatase screen conducted here for siRNAs targeting genes previously associated with the TRAIL pathway in the Aza-blanc *et al* screen (Aza-Blanc et al. 2003) and in other literature, are shown in Table 4-4. The Aza-blanc screen included siRNAs targeting 510 genes including 360 known and predicted kinases. The remaining genes were hand picked 'genes of interest'. The screen here screened siRNA targeted 691 kinases and 206 phosphatases. While the complete list of genes targeted by Aza-blanc *et al* is not available, it can be assumed that the majority of the 360 kinases screened in Aza-blanc overlap with the set screened here, while the remaining 150 do not. Of the 20 genes designated hits in the Aza-blanc screen, 14 were also targeted in the screen presented here. Only one siRNA targeting genes which reduced the sensitivity of cells to TRAIL-induced cytotoxicity in Aza-blanc *et al* or elsewhere in the literature has a score of greater than 1.6. One siRNA targeting ROS1 scored 1.7 and is ranked 65. The next highest scoring siRNA is an siRNA targeting GSK3α which scores 1.52 and is ranked 81. Thus using the first criteria for a hit – that both siRNAs targeting a gene must be about 1.6, none of the genes previously identified count as a hit. Under the second criteria for a candidate hit – that only one siRNA must score above 1.6, one gene counts as a hit, although none of the previously identified genes are within the 10 genes selected for confirmation based on this criteria. This gives the screen a nominal sensitivity of 3.5% as measured as the percentage of "true positives" selected from the sum of the true positives and the false negatives. However, the hit threshold was selected to ensure that 95% of genes with an effect as large as the effect of siCasp8 would be selected as hits. Since Caspase-8 was the gene that had the largest effect in Aza-blanc *et al* screen, other genes in this screen would be expected to have less effect and therefore a reduced probability of being selected as a hit. Further, many of the hits were not rigorously confirmed, with many of them only being targeted by one siRNA. Indeed, the only gene included in the overlap which was confirmed by multiple siRNAs was GSK3α. Nevertheless, as a group, the genes presented in Table 4-4 have a median score close to zero suggesting the results of this screen do not replicate the screen of Aza-blanc *et al.* siRNA against three of these genes

(ABL2, PRKCQ and PRKRIR) were shown to have an effect on TRAIL-induced cytotoxicity in Figure 3.13. This suggests that the screen presented here is not highly sensitive for the selection of genes previously associated with TRAIL-induced cytotoxicity. While it is concerning that many of the genes previously associated with TRAIL-induced cytotoxicity were not identified, this does not necessarily indicate that the novel genes targeted by siRNAs that did score highly in the screen are not involved in the process. That is, while this data shows that the screen might not be sensitive, it does not offer any data as to the accuracy of the screen.

Gene	*Survival*		*Score*		*Rank*	
	siRNA 1	*siRNA 2*	*siRNA A*	*siRNA B*	*siRNA 1*	*siRNA 2*
ABL2	45%	6%	1.22	-0.90	135	1412
BLK	2%	7%	-1.20	-0.95	1455	1655
GSK3α	20%	42%	0.1	1.52	594	81
GSK3β	16%	16%	-0.48	-0.10	1050	758
GUK1	19%	24%	0.12	0.48	578	376
IRAK1	19%	17%	0.16	-0.64	545	1188
MAPK1	16%	11%	0.95	0.77	765	929
MAPK10	22%	18%	0.05	0.05	625	630
PRKAA2	22%	17%	0.26	0.10	480	593
PRKCD	8%	3%	-1.06	-1.20	1532	1648
PRKCQ	11%	1%	-1.08	-1.24	1544	1688
PRKRIR	29%	30%	0.46	1.31	390	116
RAF1	29%	32%	0.10	1.04	593	186
ROS1	25%	38%	0.16	1.70	546	65
Median	19%	17%	0.11	0.08	594	694

Table 4-4 Results from this screen of genes previously associated with the TRAIL apoptosis pathway

4.5 Confirmation of hits

While the examination of the scores of siRNAs targeting genes previously associated with TRAIL-induced cytotoxicity showed the sensitivity of the screen may be low, the high scores of the positive controls suggests that the screen can identify siRNAs that do have an effect on the sensitivity of cells to TRAIL. There are several reasons not to take the results of the initial screen at face value. Firstly, while in order for a siRNA to have scored highly in the screen its effect must have repeated in both replicates, the results do not allow for a statistical assessment of the significance and reproducibility of the results. Further, effects elicited by an siRNA maybe due to knock-down of the target gene – implicating this gene in the TRAIL pathway, or effects may also be due to off target effects. In order to address these issues it is necessary to undertake a rigorous confirmation of hits before it is possible to declare that the genes targeted are involved in the TRAIL pathway with any confidence.

There are two possible methods by which genes could be selected for follow-upGenes targeted by 2 siRNAs which scored greater than 1.6 (Table 4-2), and genes targeted by the top 10 highest scoring siRNAs were selected for follow up (Table 4-3). Thus 14 genes were selected for rigorous confirmation. In order for a gene to be declared as confirmed, transfection of cells with at least two different siRNAs targeting the gene must significantly reduce the sensitivity of the cells to TRAIL-induced cytotoxicity. In order to rule out off-target effects siRNAs that significantly reduce the sensitivity of transfected cells to TRAIL-induced cytotoxicity must also reduce the mRNA level of targeted genes more efficiently than siRNAs that do not significantly reduce sensitivity.

In order to confirm the involvement of hit genes in the TRAIL-induced cytotoxicity pathway, multiple siRNAs targeting the selected genes were tested for their ability to alter sensitivity to the TRAIL ligand. Identical siRNAs to those used in the screen were resynthesised and used where both siRNAs targeting a gene scored over 1.6 (Figure 4.7a, first five genes) otherwise for each gene the siRNA which scored over 1.6 was resynthsised and used and in addition 2 novel siRNAs were used (Figure 4.7a, genes 6-14). In total 11 of the 37 siRNAs tested significantly increased the survival of transfected cells compared to negative control transfected cells on the same plate (using a Student's t-test on log transformed data with a 5% significance level). A total of 4 genes were targeted by two siRNAs that significantly increased the survival of cells treated with TRAIL ligand. The low confirmation rate may, in part, be attributable to the large amount of variation seen in the survival of cells transfected with negative control siRNA on one of the plates in this experiment (Figure 4.7a, siNeg (pl 1)).

siRNAs may fail to affect the sensitivity of transfected cells to TRAIL-induced cytotoxicity due to their inefficiency in reducing the mRNA level of the targeted gene or the lack of involvement of the gene targeted in the TRAIL-induced apoptosis pathway. To distinguish these two possibilities, the ability of the siRNAs used to reduce the mRNA levels of the targeted genes was measured using qRT-PCR (Figure 4.7b). Primers were designed to amplify from mRNA of targeted genes only, and tested for specificity and efficiency. Primers were successfully designed to amplify from 10 of the 14 targeted genes (see Appendix B). As a positive control, the ability of a well characterised siRNA targeting the gene Lamin A/C was also measured. Transfection of the siRNA targeting Lamin A/C reduced the mRNA level to 24% of the negative control level, demonstrating the effectiveness of the transfection and qRT-PCR process. 12 of the 26 siRNAs measured reduced the levels of the target siRNA by more than 70% (Figure 4.7b, dashed line).

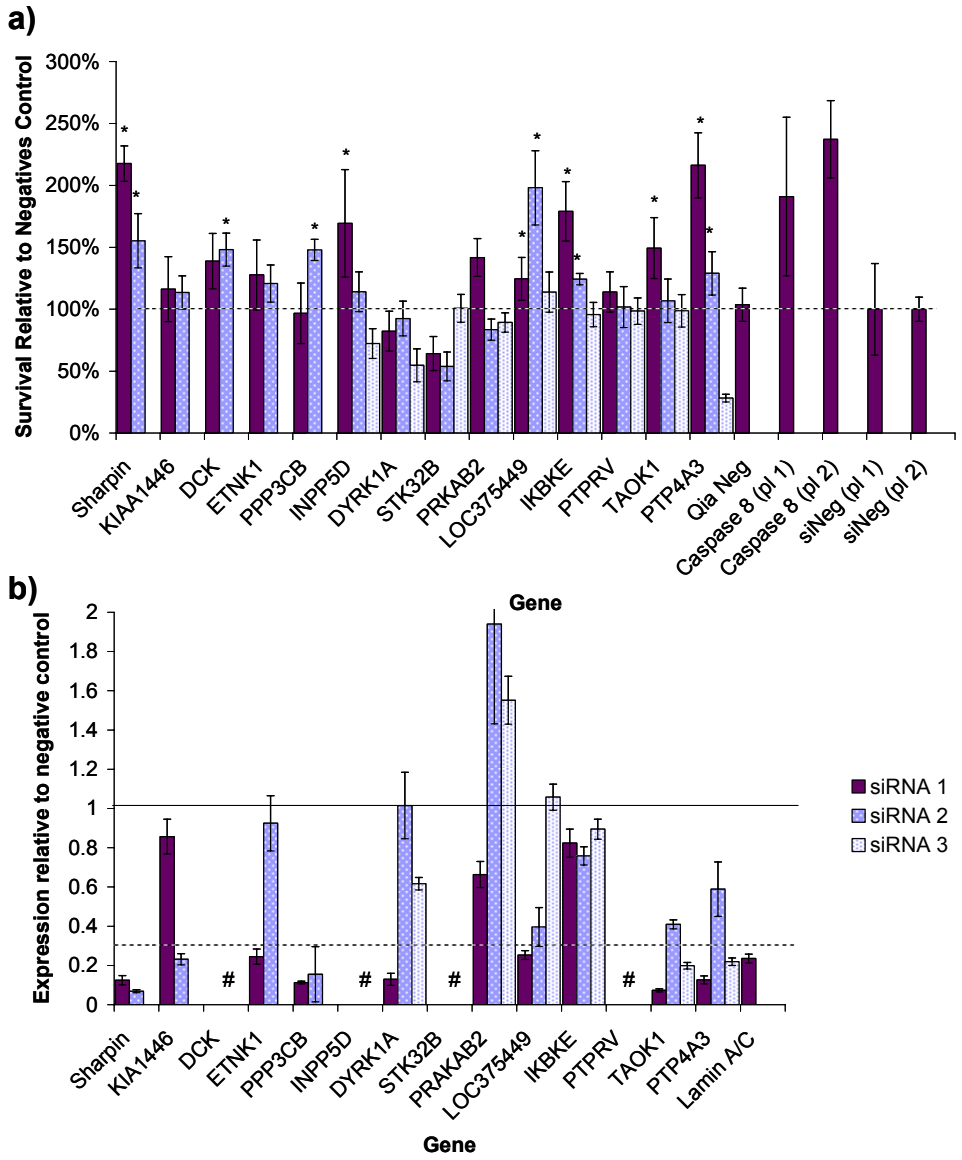

Figure 4.7 Confirmation of the effect of 14 genes from a screen of Kinases and Phosphatases.
a) Effect of siRNAs targeting 'hit' genes on sensitivity to TRAIL induced cytotoxicity. Cells were transfected with either two or three siRNAs targeting genes selected from the kinase and phosphatase screen. After 48 hours viability was measured using alamarBlue and cells were treated with 0.5µg/ml TRAIL. Viability was reassessed 24 hours later. Tests were carried out on two independent plates. Caspase 8 (pl 1) and Caspase 8 (pl 2) – positive controls from plates 1 and 2. siNeg (pl 1) and siNeg (pl 2) – negatives controls from plates 1 and 2. QiaNeg – Qiagen negative control found on screening plates. Data is shown as survival relative to negative control. Dashed line represents survival level of negative control. Error bars represent 1 standard deviation. n = 4. * = result significantly different from negative control using a Student's t-test on log transformed data (α = 0.05). b) Effect of siRNAs targeting 'hit' genes on mRNA levels of targeted genes. RNA was isolated from cells transfected with siRNAs targeting genes selected from the kinase and phosphatase screen or Lamin A/C as a positive control. cDNA was prepared by reverse transcription. SYBR green qPCR was carried out in triplicate using primers designed to amplify from mRNA of genes targeted, GAPDH and ACTB. Primers were designed and tested as described in Methods. Primers were successfully designed for 10 of the 14 genes tested. Expression levels are shown relative to negative control and were calculated using a variation of the Pfaffl method to allow normalization to multiple housekeeping genes using GAPDH and ACTB to normalize samples (Hellemans et al. 2007). # = genes for which no primers were successfully designed. Solid line represents 100% of negative control levels and dashed line represents 30% of control levels). Error bars represent 1 standard error of the mean.

Gene	*siRNA*	*TRAIL sensitivity*	*TRAIL Rank*	*> 70% KD*	*KD Rank*	*Conclusion*
Sharpin	1	+	1	+	2	**Confirmed Hit**
	2	+	2	+	1	
KIAA1446	1	-	1	-	2	**Unrepeatable**
	2	-	2	-	1	
DCK	1	-	2	N/A	N/A	**Unconfirmed**
	2	+	1	N/A	N/A	
ETNK1	1	-	2	-	1	**Unrepeatable**
	2	-	1	-	2	
PPP3CB	1	-	2	+	1	**Off-Target**
	2	+	1	+	2	
INPP5D	1	+	1	N/A	N/A	
	2	-	2	N/A	N/A	**Unconfirmed**
	3	-	3	N/A	N/A	
DYRK1A	1	-	2	+	1	
	2	-	1	-	2	**Unrepeatable**
	3	-	3	-	3	
STK32B	1	-	2	N/A	N/A	
	2	-	1	N/A	N/A	**Unrepeatable**
	3	-	3	N/A	N/A	
PRKAB2	1	+	1	-	1	
	2	-	3	-	2	**Unconfirmed**
	3	-	2	-	3	
IKBKE	1	+	1	-	2	
	2	+	2	-	1	**Confirmed Hit**
	3	-	3	-	3	
LOC375449	1	+	2	+	1	
	2	+	1	-	2	**Confirmed Hit**
	3	-	3	-	3	
PTPRV	1	-	1	N/A	N/A	
	2	-	2	N/A	N/A	**Unrepeatable**
	3	-	3	N/A	N/A	
TAOK1	1	+	1	+	1	
	2	-	2	-	3	**Unconfirmed**
	3	-	3	+	2	
PTP4A3	1	+	1	+	1	
	2	+	2	-	3	**Off-Target**
	3	-	3	+	2	

Table 4-5 Categorisation of genes selected from confirmation from kinase and phosphatase screen.
See text for definition of categories. N/A – not available. KD - Knockdown

Based on these results the genes selected for follow up were categorised into one of four categories by combining the effect of siRNAs on sensitivity to TRAIL-induced

cytotoxicity and the ability of the same siRNAs to reduce the mRNA levels of the targeted genes. Phenotypically active siRNAs are defined as siRNAs, transfection of which leads to a statistically significant reduction in TRAIL-induced cytotoxicity and the efficiency of an siRNA is defined as the degree to which it reduces the level of the intended target transcript.

- **Confirmed Hits** are genes targeted by at least two phenotypically active siRNAs. These two phenotypically active siRNAs must be more efficient than any siRNA tested that was not shown to be phenotypically active. That is, the inactivity of the non-phenotypically active siRNA can be explained by the lack of efficiency.
- **Unconfirmed Hits** are genes targeted by one phenotypically active siRNA. This phenotypically active siRNA must be more efficient than the siRNAs targeting the same gene which were not shown to be phenotypically active. This category also includes genes were the efficiency of siRNAs targeting it are not known.
- **Off-Targets** are genes targeted by both phenotypically active and non-phenotypically active siRNAs. At least one non-phenotypically active siRNAs must be more efficient than at least one phenotypically active siRNA. That is the activity of the phenotypically active siRNAs is at least in part due to off-target effects.
- **Unrepeatable** genes are genes where none of the siRNAs targeting the gene are phenotypically active. That is the original screen result is unrepeatable.

The categorisations of the genes selected for confirmation from the screen are shown in Table 4-5. Three genes are classed as Confirmed Hits (Sharpin, LOC375449 and IKBKE) and a further four genes are classed Unconfirmed Hits (DCK, INPP5D, PRKAB2 and TAOK1). Two genes are classed as confirmed Off-Targets (PPP3CB and PTP4A3). The remaining five genes were not targeted by any siRNAs that significantly altered the sensitivity of cells to TRAIL-induced cytotoxicity and so are classed as unrepeatable.

Of the top ten siRNAs selected for follow up from the ranked list of siRNAs, seven significantly reduced the sensitivity to TRAIL-induced cytotoxicity when retested (Figure 4.7a and Table 4-5), suggesting that the accuracy of this method for selecting 'hit' siRNAs is approximately 70%. However, only three of the genes targeted by these siRNAs were confirmed as hits by being targeted by two independent siRNAs that, when transfected, caused a significant reduction in sensitivity to TRAIL-induced cytotoxicity, and which reduce the level of mRNA more efficiently than siRNA which do not have an effect on TRAIL sensitivity. This suggests that the accuracy of this method for selecting 'hit' genes is 30%. This could possibly be raised to 60% if all the unconfirmed genes were confirmed. Of the five genes selected for being targeted by two siRNAs scoring greater than 1.6 in the screen,

only one gene confirmed, suggesting that the accuracy of this method for selecting 'hit' genes is approximately 20%, rising to 30% if the unconfirmed genes were also confirmed.

The screen has clearly not been as effective as might have been expected given the optimisation experiments, particularly with regard to finding genes previously associated with the pathway. It should be remembered however, that many of the previously associated genes are from the Aza-blanc screens. These hits were not confirmed with the appropriate rigor, and so it impossible to tell how many are genuinely involved in the pathway. The number of genes selected from the screen that confirmed is also lower than might have been expected. This may in part be due to the rigorous confirmation process applied. While the number of confirmed hits is small, the confidence that these hits are genuinely involved in the process is high

4.6 Experimental characterisation of confirmed hits

From rigorous follow up of the screen of siRNAs targeting 897 kinases, phosphatases and associated genes, three genes passed the confirmation process. These genes are targeted by two siRNAs that significantly reduce the sensitivity to cells to TRAIL-induced cytotoxicity and reduce the mRNA level of the target gene to a greater extent than any siRNA tested that does not significantly reduce the sensitivity of cells to TRAIL-induced cytotoxicity. It should be noted that in one case (IKBKE), although the phenotypically active siRNAs are more efficient than the non-active siRNAs, the difference in efficiency is small, while the difference in sensitivity is large. However, since the gene is targeted by two siRNAs that are phenotypically active and given that IKBKE is an activator of NF-κB which has been implicated in control of TRAIL sensitivity previously (1.3.2.6) it was decided that it was worth investigating this gene further. To investigate the significance of these genes to the TRAIL-induced apoptosis pathway a series of experiments were undertaken to characterise these them.

4.6.1 Effect of knock-down of confirmed hits on TRAIL ligand dependent activation of caspases.

As noted in earlier chapters, the assay implemented in the screening for and confirmation of, genes involved in the TRAIL pathway has measured the effect of siRNAs on TRAIL-induced cytotoxicity rather than TRAIL-induced apoptosis. In order to establish that hit genes affect TRAIL-induced apoptosis, the effect of siRNA mediated knock-down of hit genes on the TRAIL-ligand dependent activation of several caspases was investigated. As well as confirming the involvement of these genes in TRAIL-induced apoptosis, in

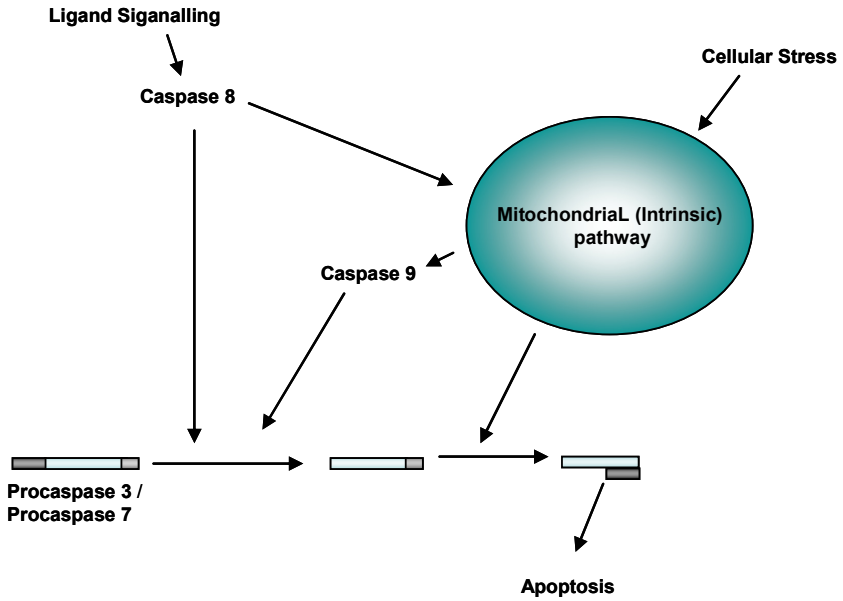

Figure 4.8 Order of caspases in the caspase cascade

addition to TRAIL-induced cytotoxicity, these experiments will indicate the point in the apoptotic pathway at which the genes act. Most genes involved in TRAIL-induced apoptosis would be expected to affect the activation of the executioner caspases, Caspase-3 and Caspase-7 by the TRAIL ligand. Genes involved in the regulation of the intrinsic, mitochondrial pathway, or the connection between the extrinsic and intrinsic pathways would be expected to affect the induction of Caspase-9 as well as caspases -3 and -7. Genes involved in the regulation of the extrinsic pathway would be expected to affect Caspase-8 in addition to caspases -9,-3 and -7 (Figure 4.8). The effect of the knock-down of hit genes on the induction of these caspases was determined using luminescent caspase assays to compare caspase activity in TRAIL treated vs. mock treated cells (Figure 4.9).

As expected TRAIL treatment of cells increases the activity of Caspase-8 more than three fold (from 1,026 to 3,418, Figure 4.9a). The levels of Caspase-8 activity in TRAIL treated, siCasp8 transfected cells are 45% of those in negative control transfected cells (Figure 4.9a). Reduction of Caspase-8 levels in TRAIL-treated cells is significant in cells transfected with both siRNAs against IKBKE (siIKBKE.1 and siIKBKE.2) and one of the two siRNAs against each of Sharpin and LOC37449. Levels are reduced to 60% and 73% of the negative control respectively for cells transfected with siIKBKE.1 and siIKBKE.2 and to 36% and 38% for cells transfected with siRNA 1 targeting Sharpin (siSharpin.1) and siRNA 2 targeting LOC375449 (siLOC375449.2) respectively. Transfection of siRNA 2 against Sharpin (siSharpin.2) and siRNA 1 against LOC375449 (siLOC375449.1) did not significantly reduce the activity of Caspase 8 in TRAIL treated cells compared to the negative control.

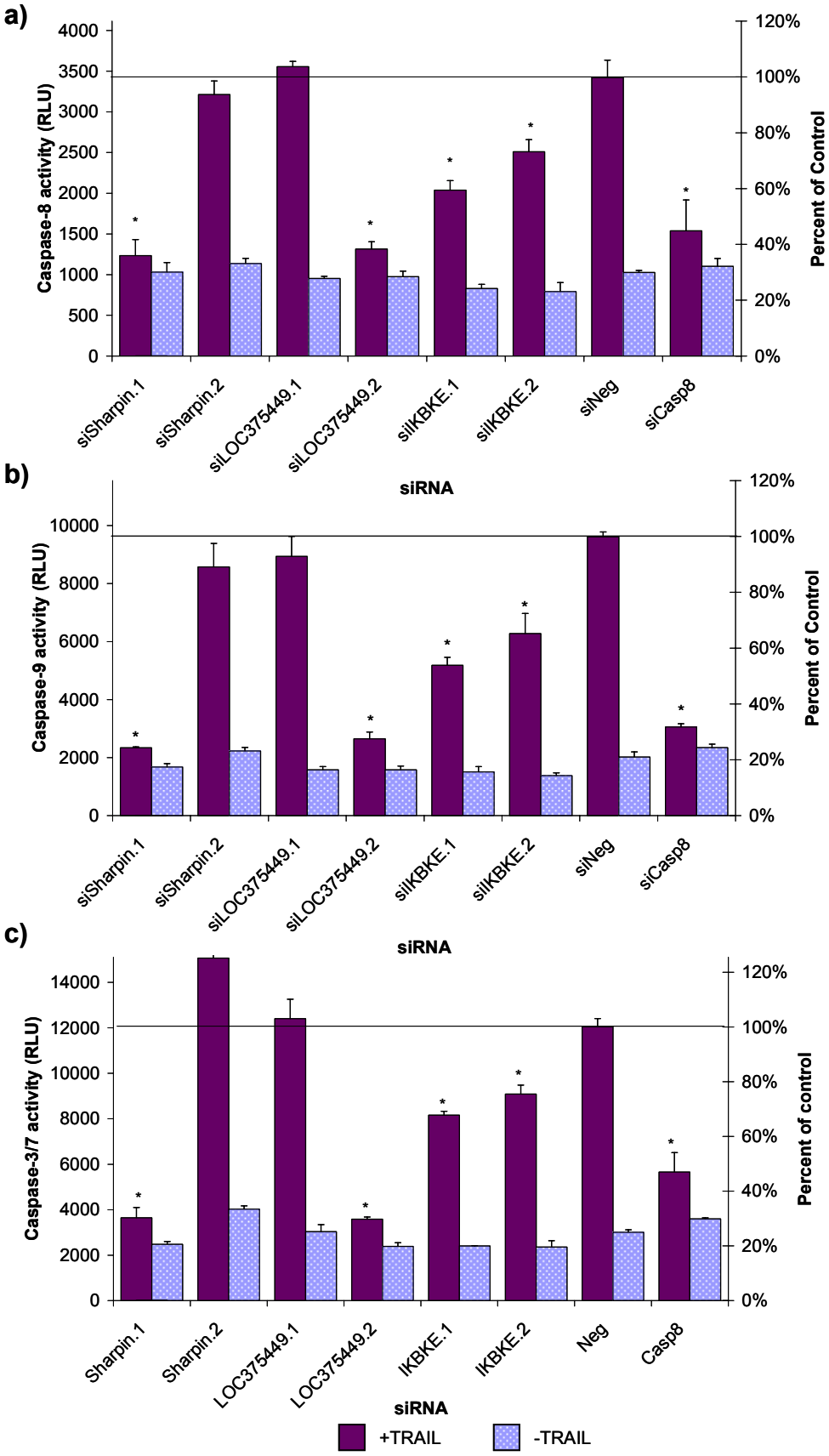

Figure 4.9 Effect of siRNAs targeting confirmed hits on TRAIL-induced caspase activation
Cells were transfected with siRNAs targeting confirmed hits. 48 hours later cells were treated with either 0.5µg/ml TRAIL or media for six hours. Caspase-8 (a), Caspase-9 (b) or Caspase-3/7 (c) luminescent assay reagent was added to cells and incubated for 1 hour before luminescence was determined. Error bars represent 1 standard deviation. n = 3. * = significantly different from negative control using a Student's t-test (Bonferroni corrected $\alpha = 0.05$). Horizontal line represents negative control level. RLU = relative luminescent units

These results are almost identical to the results obtained by measuring the effect of transfecting these siRNAs on the activity of Caspase-9 (Figure 4.9b) and Caspases 3/7 (Figure 4.9c). The correlation coefficient between the effects on Caspase-8 and the effects on Caspase-9 is 0.99 ($r^2 = 0.98$) and the correlation coefficient between effects on Caspase-8 and Caspases 3/7 is 0.95 ($r^2 = 0.90$). The effect of knockdown on the level of Caspase-9 activity is generally stronger: transfection of siCasp8 reduces the activity of Caspase-9 in TRAIL treated cells to 31% of activity in negative control transfected cells, compared to reducing the activity of Caspase-8 it self to 45% of control levels.

It is unclear why only one siRNA targeting Sharpin and LOC375449 significantly reduce the levels of caspase activity (Figure 4.9), when both siRNAs targeting both these genes reduce the level of cytotoxicity induced by TRAIL treatment (Figure 4.7). However, those siRNAs which do not significantly reduce the levels of caspase activity are the siRNAs which reduced the TRAIL-induced cytotoxicity least of the two siRNAs targeting each gene. There is a strong correlation between the effect of siRNAs targeting hit genes on TRAIL-induced cytotoxicity and their effects on the level of Caspase-8 activity ($r^2 = 0.80$).

It can be concluded that there is strong evidence that the genes identified as confirmed hits from the screen of kinases, phosphatases and associated genes affect TRAIL-induced apoptosis as well as TRAIL-induced cytotoxicity. Further there is evidence that all three genes act to regulate the pathway at or above the level of Caspase-8 activation.

4.6.2 Effect of knock-down of confirmed hits on the sensitivity of HeLa cells to a selection of apoptosis inducing conditions

Genes regulating the sensitivity of cells to TRAIL-induced apoptosis may do so at several levels. They may be specifically involved in the regulation of the TRAIL-induced apoptosis pathway. In this case knock-down of these genes should affect the sensitivity of cells to TRAIL-induced apoptosis, but not the sensitivity of cells to apoptosis induced by other conditions. Alternatively genes maybe involved in regulating the sensitivity of cells to ligand-induced apoptosis. In this case knock-down of these genes should affect the sensitivity of cells to both TRAIL-induced apoptosis and also the sensitivity of cells to other apoptosis inducing ligands, such as FAS ligand. Finally genes may be involved in the regulation of the general sensitivity of cells to apoptosis. In this case knock-down of the genes should affect the sensitivity of cells to TRAIL-induced apoptosis, apoptosis induced by other ligands, and sensitivity to apoptosis induced by non-ligand apoptosis inducers such as hydrogen peroxide or UV radiation.

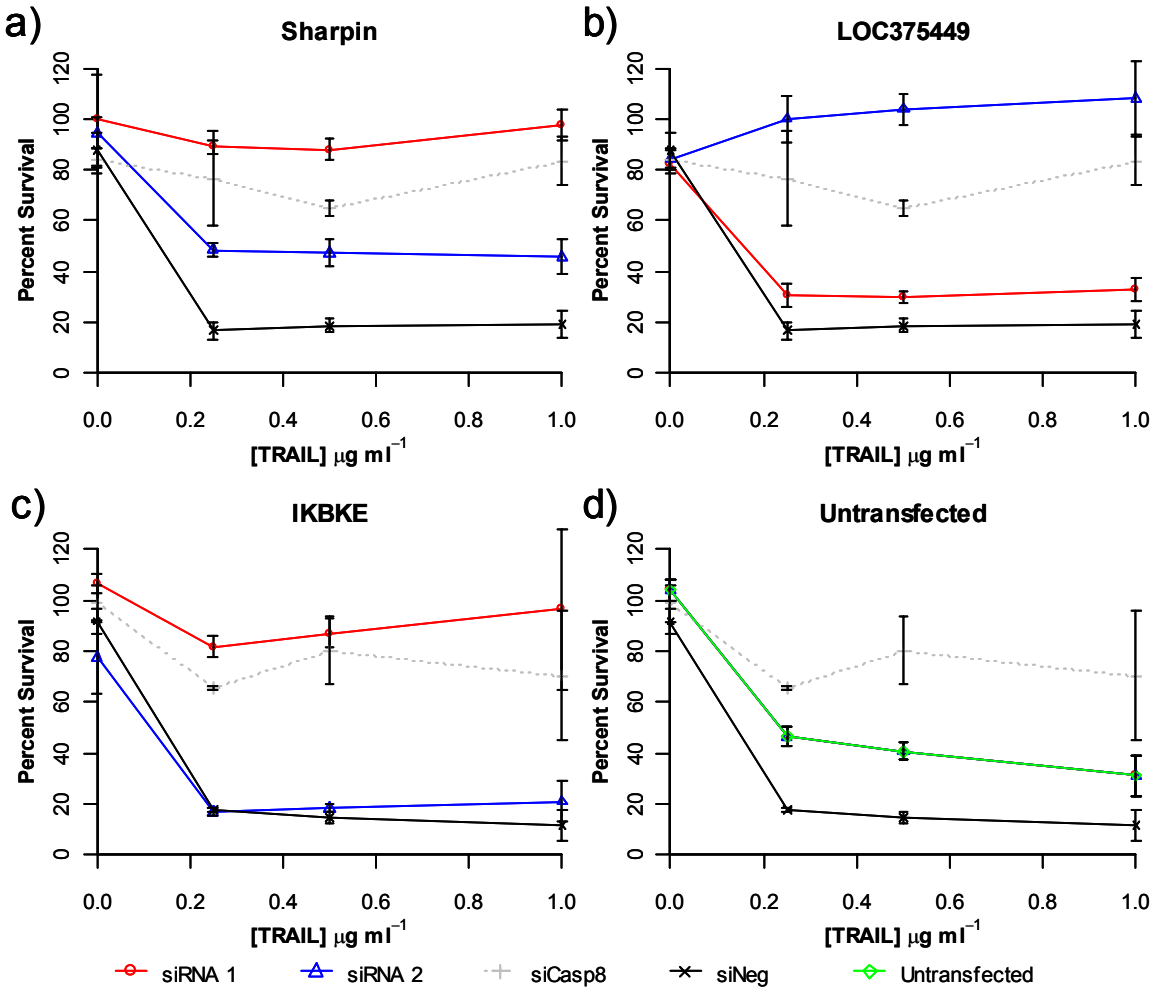

Figure 4.10 Knockdown of confirmed hit genes affects sensitivity of cells to a range of concentrations of TRAIL ligand.
Cells were transfected with siRNAs targeting a) Sharpin, b) LOC375449, c) IKBKE, siCasp8 or siNEG or d) mock transfected. Viability was measured using alamarBlue 48 hours later. Cells were treated with either 0.25µg/ml, 0.5µg/ml, 1 µg/ml TRAIL ligand or serum free media for 24 hours and the viability reassessed. Error bars represent 1 standard deviation, n = 3.

In order to further confirm the effect of hit siRNAs on TRAIL-induced cytotoxicity, and also to provide a comparison for the effect of other apoptosis inducers, the effects of knockdown of confirmed hit genes on the sensitivity of cells to TRAIL-induced apoptosis at a range of TRAIL concentrations was determined (Figure 4.10). The effects of the transfection of siRNAs targeting Sharpin and LOC375449 reflect the effects seen in the confirmation experiments. siSharpin.1 and siSharpin.2 increased the number of cells which survived treatment with 1µg/ml TRAIL from 19% to 98% and 46% respectively, while 83% of cells transfected with siCasp8 survived (Figure 4.10a). Similarly transfection with siLOC375449.1 and siLOC735449.2 increased survival of cells treated with 1µg/ml TRAIL from 19% to 33% and 106% respectively (Figure 4.10b). The concentration of TRAIL had little effect on survival in both cases. Knock-down of IKBKE by siIKBKE.1 increased the survival of cells treated with 1µg/ml TRAIL from 11% to 97% compared with siCasp8,

which increased survival to 71%. The concentration of TRAIL had little effect on the level of cytotoxicity induced by TRAIL. However, knockdown IKBKE by siIKBKE.2 only increased survival of cells treated with 0.5μg/ml TRAIL from 14% to 18%, although the increase was marginally more substantial when cells were treated with 1μg/ml where siIKBKE.2 increased survival from 11% to 21% (Figure 4.10c). While at 7% the standard deviation of the mean survival of siIKBKE.2 transfected cells treated with 1μg/ml TRAIL is within the range seen for other siRNAs in this experiment, the standard deviation as a proportion of the mean, that is the coefficient of variance, is much higher for siIKBKE.2 than for data from other siRNAs. This pattern is in line with previous findings with siIKBKE.1 having a larger effect on both TRAIL-induced cytotoxicity and TRAIL-induced caspase activation than siIKBKE.2 (Figure 4.7a and Figure 4.9), but here the pattern is much more pronounced. This presents a difficulty. Transfection of siIKBKE.2 produced a significant change in sensitivity to TRAIL-induced cytotoxicity in confirmation experiments, and also reduced the level of TRAIL-induced caspase activity, yet here, doesn't produce a significant change in TRAIL-induced cytotoxicity, although it does cause a non-significant reduction in sensitivity. Therefore the weight of evidence is in favour of this gene being involved in TRAIL-induced apoptosis. Similar to findings in the assay development experiments, but unlike findings from the screen, untransfected cells are more resistant to TRAIL-induced cytotoxicity than negative control transfected cells.

In order to determine the involvement of confirmed hit genes in regulation of apoptosis induced by other ligands, the effect of knockdown of these genes on the sensitivity of cells to FAS ligand-induced apoptosis was determined. (Figure 4.11). An average of 19% of negative control transfected cells survived treatment with 100ng/ml recombinant FAS ligand plus™ (FAS ligand fused to a FLAG epitope, referred to as FAS ligand from here on). An average of 76% of siCasp8 transfected cell survived treatment with 100ng/ml FAS ligand (Figure 4.11). Knockdown of Sharpin by siSharpin.1 but not siSharpin.2 reduces the sensitivity of cells to FAS ligand at all concentrations tested, with 43% of cells surviving treatment with 100ng/ml of FAS ligand, compared with 20% of negative control transfected cells ($p = 0.009$, calculated using Student's t-test on log transformed data) and 26% of siSharpin.2 transfected cells (Figure 4.11a). Knockdown of LOC375449 by both siRNAs targeting this gene reduced the sensitivity to cells to FAS ligand at all concentrations tested. Transfection with siRNAs siLOC375449.1 and siLOC375449.2 increased survival of cells treated with 100ng/ml of FAS ligand to 40% and 57% respectively compared with 20% of negative control transfected cells ($p = 0.011$ and 0.004 respectively, Figure 4.11b). As with

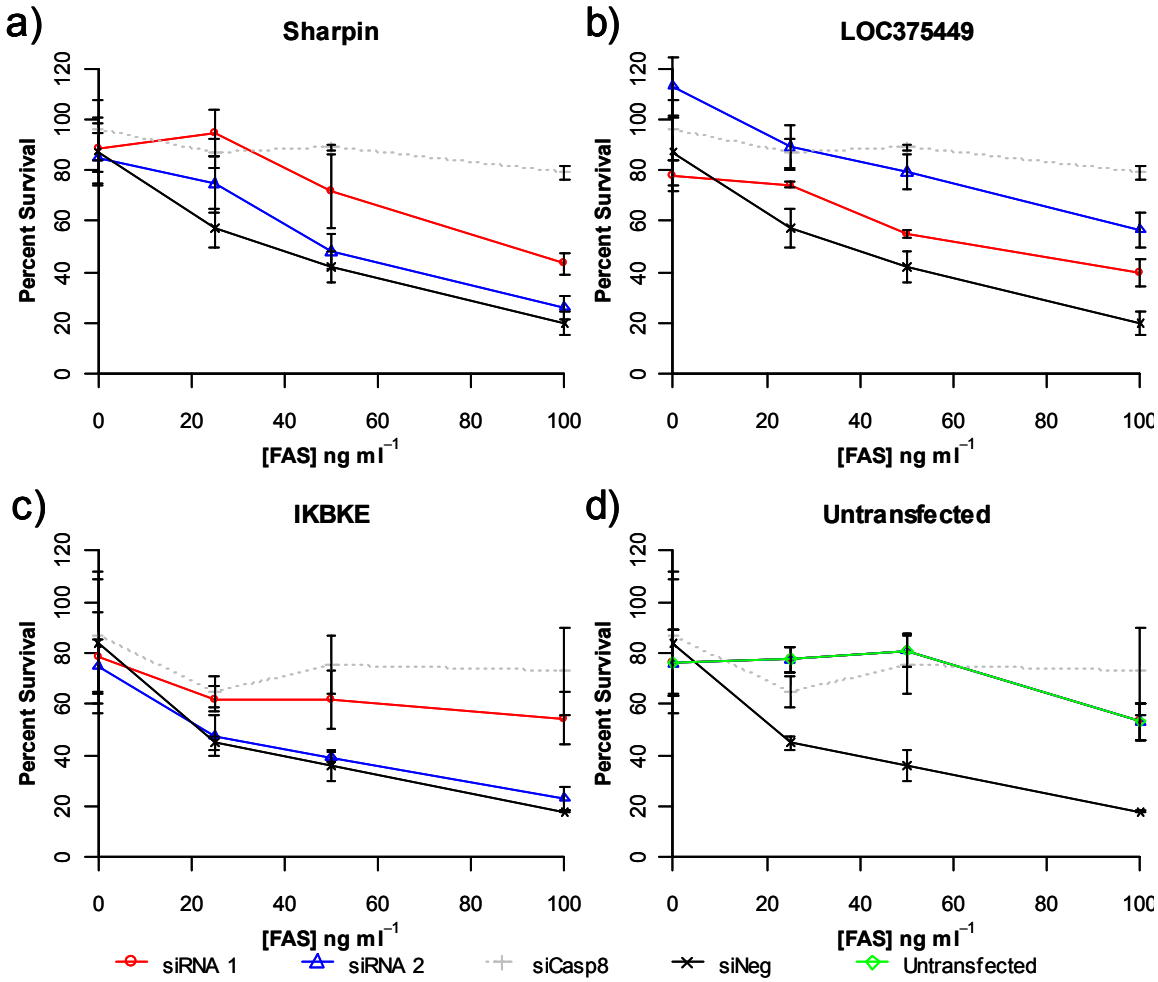

Figure 4.11 Knockdown of confirmed hit genes affects sensitivity of cells to FAS ligand induced apoptosis.
Cells were transfected with siRNAs targeting a) Sharpin, b) LOC375449, c) IKBKE, siCasp8 or siNEG or d) mock transfected. After 48 hours viability was assessed using alamarBlue. Cells were then treated with the concentration of FAS ligand Plus™ indicated in serum free media or serum free media only for 24 hours and viability was reassessed. Error bars represent 1 standard deviation, n = 3

the effect of knock-down of IKBKE on TRAIL sensitivity, knockdown of IKBKE by siIKBKE.1, but not siIKBKE.2 reduced sensitivity of cells to FAS ligand induced apoptosis. As with sensitivity to TRAIL-induced cytotoxicity, the transfection processes itself increases the sensitivity of cells to FAS (Figure 4.11d).

Knockdown of all three confirmed hit genes affects the sensitivity of cells to apoptosis inducing ligands other than TRAIL, although the effects are not as strong. siRNAs that have a small effect on the sensitivity of cells to TRAIL-induced apoptosis, do not have a significant effect on FAS ligand induced death. This is similar to the effect of increased cFLIP levels. Over-expression of cFLIP has a stronger inhibitory effect on TRAIL-induced apoptosis than it does on FAS-induced apoptosis (Irmler et al. 1997).

The role of the confirmed hit genes in the regulation of cell death induced by physiological conditions was examined by determining the sensitivity of cells to H_2O_2-(which

increases oxidative stress) and UV radiation- (which leads to DNA damage) induced cell death (Figure 4.12 and Figure 4.13). High concentrations of H_2O_2 killed all cells, irrespective of siRNA transfection. At lower concentrations significantly more siNeg transfected cells died than siCasp3 transfected cells (average survival after treatment with 100μM H_2O_2 was 36% and 86% respectively). Knock-down of Sharpin had no effect at either 200μM or 100μM H_2O_2, however knockdown of Sharpin (by siSharpin.1 only) did significantly increase the survival of cells treated with 50μM H_2O_2, from 81% to 122%, with a p-value of 0.010 (Figure 4.12a). Knock-down of LOC375449 had an effect on sensitivity of cells only at 100μM H_2O_2, however the effect of neither siRNA is significant at the 5% level (survival of 70% and 73% for siLOC375449.1 and siLOC375449.2 transfected cells compared with 43% for siNeg transfected cells, p-value = 0.72 and 0.62 respectively, Figure 4.12b). siIKBKE.2 had no effect on the sensitivity of cells to H_2O_2. However siIKBKE.1 had a large, significant effect, increasing the survival of cells treated with 100μM from 30% to 77% (Figure 4.12c, p value = 0.003).

It is known that high enough concentrations of H_2O_2 induce a necrotic cell death rather than an apoptotic cell death. This could explain the lack of effect of knock down of IKBKE and Capase-3 at higher concentrations, when a large effect is observed at lower concentrations (Nosseri, Coppola & Ghibelli 1994).

Of all six siRNAs, only siIKBKE.1 has a large and significant effect on the sensitivity of cells to UV radiation-induced cell death (Figure 4.13c). siIKBKE transfection increased the survival of cells treated with 100 J/ml^2 UV radiation from 31% to 85% (p-value = 0.0002). siSharpin.1 had a smaller effect on the sensitivity of cells to UV radiation, increasing the survival of cells exposed to 100 J/ml^2 UV radiation from 28% to 50%, (p-value = 0.02, Figure 4.13a). It is worth noting that the positive control siCasp3 had a similar effect, increasing survival from 28% to 51%. Unfortunately the interpretation of data on the sensitivities of cells to UV radiation is hampered by low viability in cells supposedly not exposed to UV radiation, particularly the negative control. One possible explanation for this is that a defect in the experimental protocol allowed the cells which were not to be exposed to UV to be exposed, through insufficient shielding. This would account for the observed increased survival of supposedly unexposed cells transfected with siRNAs which increase the survival of cells exposed to UV radiation. That is, if they were accidentally exposed to some level of UV radiation, the transfection of an siRNA regulating apoptotic responses would protect them, while transfection of the non-targeting siNeg would not.

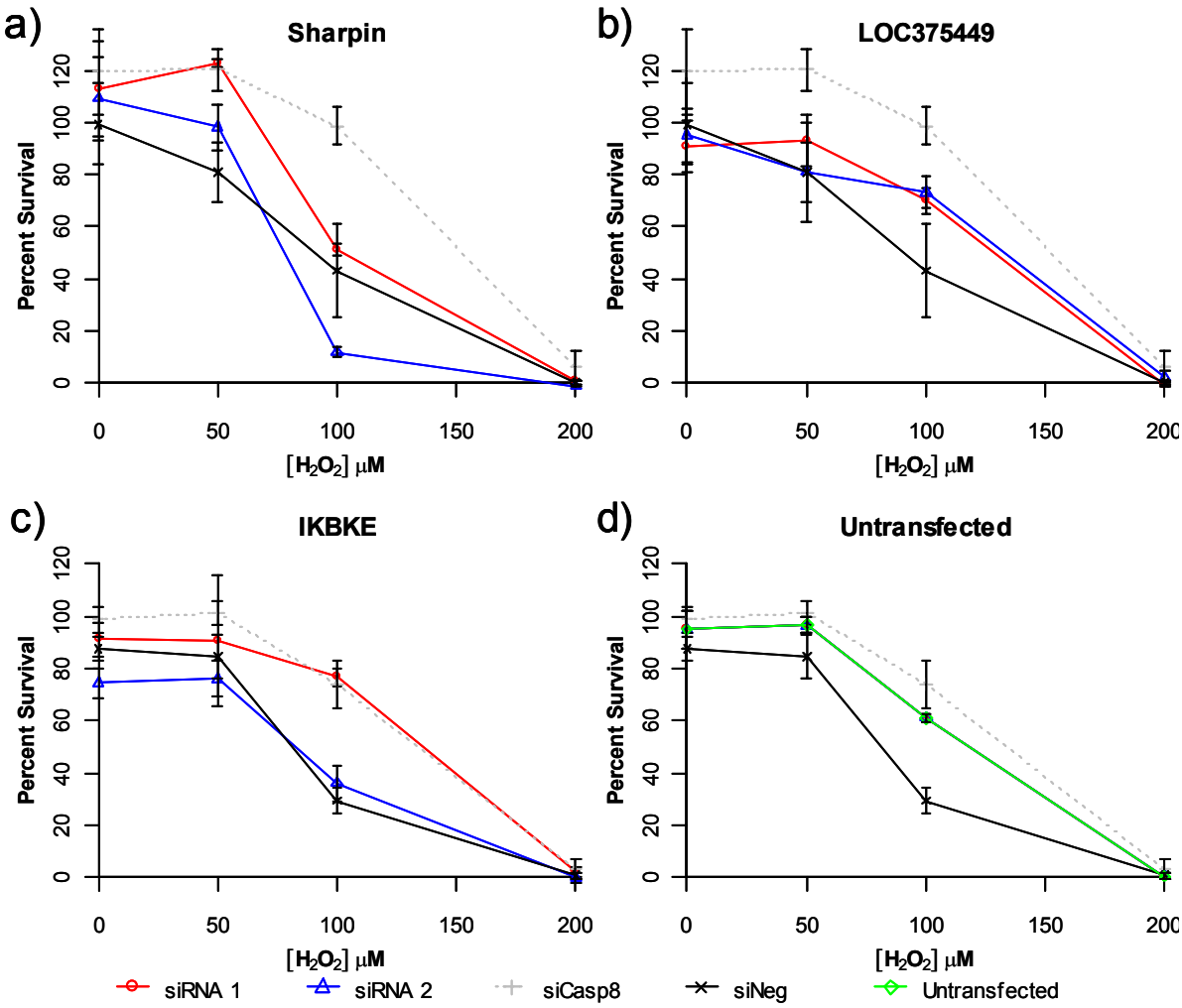

Figure 4.12 Effect of knockdown of confirmed hit genes on sensitivity to H_2O_2 induced cytotoxicity. Cells were transfected with siRNA targeting a) Sharpin, b) LOC375449, c)IKBKE, an siRNA targeting Capase-3 (siCasp3) as a positive control or siNeg or d)mock transfected. After 24 hours the viability of cells was measured using alamarBlue. Cells were then incubated in serum free media containing the concentration of H_2O_2 indicated for 24 hours and viability was reassessed. Error bars represent 1 standard deviation, n = 3.

The results of the biological investigations undertaken are summarised in Table 4-6. The strength of the effects of transfection with each siRNA in each of the assays correlated with strength of the effects in the original confirmation experiment (Figure 4.7a). All of the siRNAs that significantly reduced the activity of Caspases -3 and -7 in TRAIL treated cells, also significantly reduced the activity of Caspases -8 and -9, suggesting these genes regulate the apoptotic pathway at the level of Caspase-8 activation. Consistent with this knockdown of LOC37449 reduces the sensitivity of cells to FAS ligand induced cell death as well as TRAIL-induced cell death, but not to H_2O_2 induced or UV induced cell death. Knockdown Sharpin also protects against both TRAIL and FAS ligand induced cell death. It may also protect to some extent against H_2O_2 induced cell death and UV induced cell death, but this effect is much weaker. In contrast knockdown of IKBKE by the siRNA that gave the strongest protection against TRAIL-induced cell death also had a strong protective effect

against FAS, H_2O_2 and UV induced cell death.

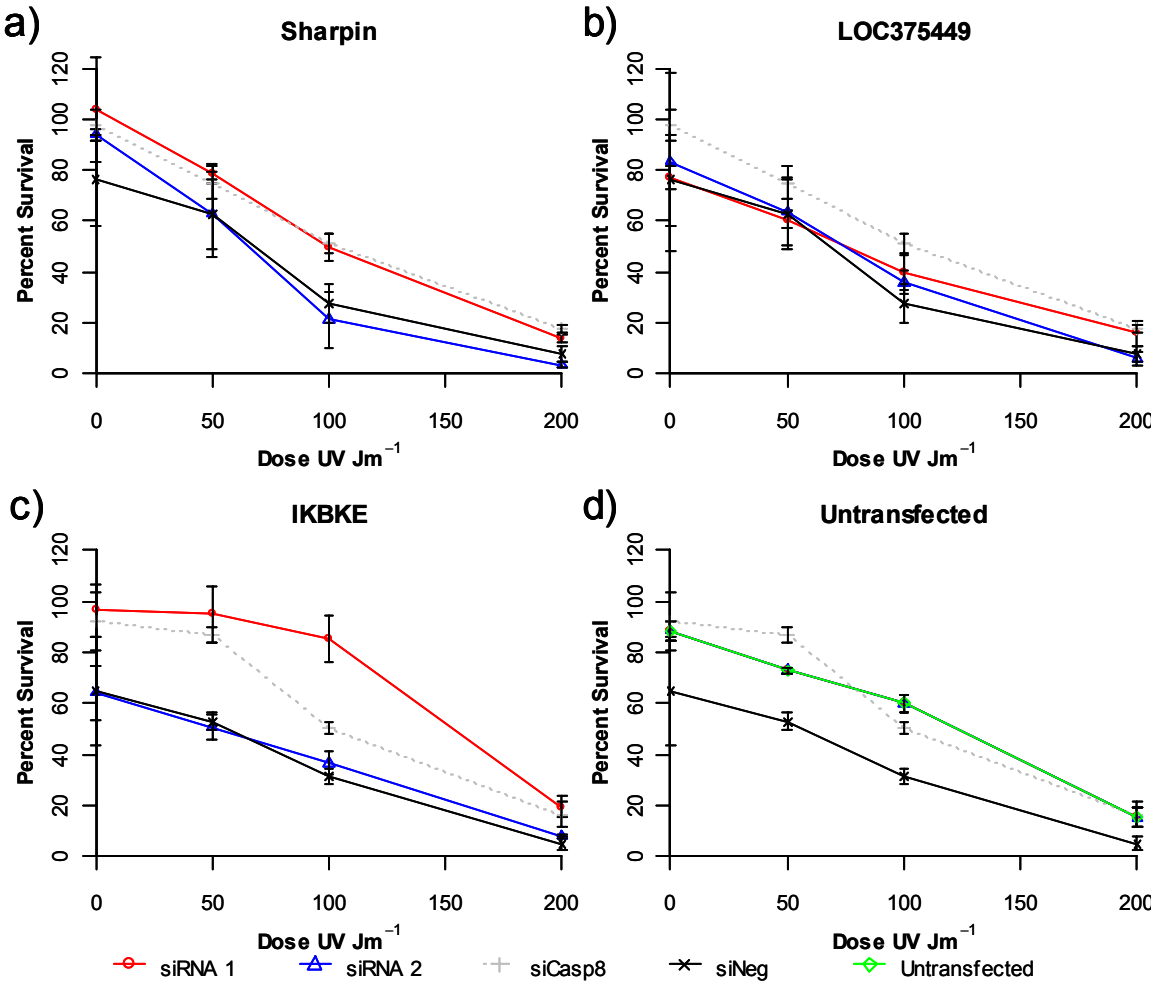

Figure 4.13 Effects of knockdown of confirmed hit genes on sensitivity to UV radiation induced cell death.
Cells were transfected with siRNAs targeting a) Sharpin, b) LOC375449, c) IKBKE, Caspase -3 (siCasp 3), siNEG or d) mock transfected. 48 hours later viability of cells was assessed using alamarBlue. Cells were exposed to doses of UV radiation indicated. Cells were incubated for 24 hours in serum free media and viability was reassessed. Error bars represent 1 standard deviation. n = 3.

siRNA	Caspase Cascade			Apoptosis Inducers			
	Caspase 8	Caspase 9	Caspase 3/7	TRAIL	FAS	H_2O_2	UV
siSharpin.1	+	+	+	+	+	-/+	-/+
siSharpin.2	nsd	nsd	nsd	+	nsd	nsd	nsd
siLOC375449.1	nsd	nsd	nsd	+	+	nsd	nsd
siLOC375449.2	+	+	+	+	+	nsd	nsd
siIKBKE.1	+	+	+	+	+	+	+
siIKBKE.2	+	+	+	nsd	nsd	nsd	nsd

Table 4-6 Summary of experimental characterisation of confirmed hits from screen of Kinase and Phosphatase screen.
Nsd – No significant difference under any of the tested conditions.

4.7 Discussion and conclusions

A screen of genes annotated as kinases, phosphatases or kinase/phosphatase associated was carried out (Figure 4.5 and Figure 4.6). Fourteen genes were selected for confirmation by virtue of being targeted by two siRNAs scoring highly in the screen or being targeted by one siRNA scoring very highly (Table 4-2 and Table 4-3). Of these genes three could be designated confirmed hits, while four were designated as unconfirmed hits, two as off-target hits and five as unreproducible (Table 4-5 and Figure 4.7).

The three confirmed hits were further investigated by experiments to measure the effect of knock-down of these genes on TRAIL ligand induced activation of several caspases (Figure 4.9) and their ability to affect sensitivities to apoptosis inducers other than TRAIL (Figure 4.10,Figure 4.11,Figure 4.12 and Figure 4.13).

4.7.1 The Screen

Different normalization and data transformation strategies were explored. Median normalization produced the most consistent results and was thus selected as the normalization method for data analysis (Figure 4.3). Normalisation methods based on values of control wells are sensitive to the effects of small numbers of outliers in these wells, particularly when only two replicates of each well are present on each plate.

It was observed that there was a strong relationship between mean normalized survival and variance between screen replicates in mean/standard deviation plots of the screen results (Figure 4.4a). This relationship was also observed in scatter-plots of one replicate of the screen against the other and in the spread of scores for siCasp8 transfected wells compared to negative control transfected wells (Figure 4.5d). While such relationships can complicate analysis, log transforming data did not remove this relationship and may have made it stronger (Figure 4.4b).

Three positive controls were included in each plate. These controls had previously been observed to affect the sensitivity of cells to TRAIL to different degrees, with siCasp8 having the strongest effect, followed by siBID and then siSMAC. Transfection of each of these controls reduced the sensitivity of cells to TRAIL-induced cytotoxicity as expected (Figure 4.5d), although there was a reduction in the separation between the positive controls and negative controls compared to the assay development experiments (as measured by Z'-factor). This could represent the extra variability that is inevitable with an increase in throughput, or the reduced concentration of TRAIL used in the screen compared to earlier

experiments (0.25μg/ml here compared with 1μg/ml previously) despite the fact that previous experiments showed little reduction in cytotoxicity when concentration of TRAIL was reduced.

Comparison of the inter-replicate correlation and the inter-siRNA correlation reveals that there is a much higher degree of reproducibility between replicates of the same siRNA than there is between different siRNAs targeting the same gene (Figure 4.5a and Figure 4.5b). This demonstrates that while the screen could be regarded as accurate at selecting highly active siRNAs, the interpretation of this in terms of selecting hit genes is more difficult. This was also seen in the confirmation process where 70% of siRNAs from the top ten scoring siRNAs in the screen reproducibly reduced the sensitivity of cells to TRAIL-induced cytotoxicity, but only 30% of genes targeted by siRNAs in the top ten scoring siRNAs were confirmed. There are two possible explanations for this. Either the siRNAs not eliciting a phenotype do not sufficiently knock down the target mRNA, or alternatively the siRNAs that are eliciting a phenotype are doing so through the knock-down of an off-target mRNA not targeted by the second siRNA. Two lines of evidence suggest the former explanation applies in more cases than the latter. Firstly, in five cases, when attempts were made to confirm genes targeted by only a single siRNA by obtaining further siRNAs, these siRNAs did not knock-down the mRNA of the targeted gene as efficiently as the siRNA originally selected from the screen. In two of these cases the original siRNA elicited a significant phenotype while the additional, less efficient siRNAs did not. These are two of the four "unconfirmed hits", with the other two being genes for which siRNA efficiency was not measured. Effects of siRNAs targeting these genes could still turn out to be due to off-target effects. However, this interpretation is also suggested by the finding that only one of the 28 siRNAs targeting genes previously associated with the TRAIL-induced apoptosis pathway scored over the hit threshold score of 1.6. It must, however, be remembered that the threshold was set to capture 95% of siRNAs that had an effect of the same strength as knock-down of Caspase-8. This same cut off would have only selected 70% of wells with siRNAs targeting BID and only 2% of wells with siRNAs target SMAC/DIABLO, even though transfection with this siRNA had a clear (albeit small) effect on the sensitivity of cells to TRAIL-induced cytotoxicity. Further the majority of genes identified in the Aza-Blanc screen were not rigorously confirmed. In the majority of cases only a single siRNA targeting the gene was used, the efficiency of the knockdown elicited was not measured and no statistics were employed to determine if differences were significant and so it is possible that a number of these are false positives (Aza-Blanc et al. 2003).

The low correlation between the effects of an siRNA on TRAIL-induced cytotoxicity and the effect of other siRNAs targeting the same gene would suggest that the score of both siRNAs targeting the gene must be taken into account when selecting genes for follow up. Despite this, selecting genes for follow up on the basis of the top scoring siRNAs proved more efficient than selecting genes for follow up on the basis of two siRNAs scoring higher than a pre-selected threshold based on the distribution of the control siRNA scores, since the former method led to the confirmation of 3 genes from 10 as genuine hits, an accuracy of 30%, while the latter method lead to the confirmation of 1 from 5 genes, an accuracy of 20%. It is to be noted that these differences do not reach statistical significance, indeed 30% is similar to 20% given the sample sizes. However, no gene was confirmed which was selected by virtue of being targeted by two siRNAs scoring over 1.6 in the screen that was not also selected by being targeted by one siRNA in the top ten siRNAs. The small sample size also points to another problem with selecting genes in this way – the small number of genes selected. While a higher threshold may have helped to select hit genes, the small number of genes with two siRNAs scoring higher than 1.6 suggests that this strategy would select very few genes.

Assessing the sensitivity and accuracy of the screen is difficult. 70% of siRNAs from the top ten siRNAs reconfirmed, this means that the accuracy of the screen at selecting siRNAs which have an effect is 70%, but only for siRNAs in the top ten, undoubtedly the accuracy will fall further down the ranked list of siRNAs. Further, only 30% of genes targeted by these siRNAs were confirmed (although with 3 genes "unconfirmed" this could rise to 60%), suggesting that the accuracy of the screen for selecting hit genes is much lower, even for this portion of the ranked list of siRNAs. In terms of sensitivity, 95% of siCasp8 transfected wells scored more than 1.6 (by design), but only 70% of siBID transfected wells and 2% siSMAC transfected wells scored more than this threshold. Further, only 3.5% of siRNAs targeting genes previously associated with TRAIL-induced apoptosis scored more than this threshold, although it is possible that some of these are false positives (see above). Genes were also selected for confirmation by virtue of being targeted by an siRNA in the top ten highest scoring siRNAs in the screen. The lowest scoring of these had a score of 3.24. 66% of wells transfected with siCasp8, 16% of wells transfected with siBID and none of the wells transfected with siSMAC scored higher than this threshold. None of the wells transfected with negative controls siRNAs scored higher than this threshold. However, it should be noted that this criteria was not designed to definitively separate all hits from all none hits, but rather to select high confidence candidates for confirmation.

These calculations rely on classifying genes as either hits or non-hits, assuming that the distribution of siRNA scores is a mixture of two distributions: the distribution of siRNAs that do affect TRAIL-induced cytotoxicity and the distribution of those that do not. In a pathway model of cellular signalling, if a gene in the pathway is present then the pathway is intact and signals. In the absence of this component the pathway is incomplete and does not signal. As such, genes are either involved or not in the process. Thus the position of an siRNA score in the distribution of all siRNA scores is purely determined by the efficiency of the knockdown, and the technical variation in the experiment – differences in transfection efficiency, variation in the measurement of the phenotype etc. In a network model of cellular signalling, each component has a quantitative effect on the output signal of the network. Thus there is a relationship between the knock-down efficiency and network output. If network output falls below a certain level then the apoptotic program is not initiated in response to the TRAIL ligand. Crucially however, this relationship is different for different genes. As such, the distribution of siRNA scores depends on two distributions: the variations due to technical variation in the system (transfection efficiency, assay variation, variation in the state of other network components etc.) and also variation due to the different quantitative effects on network output. In this case it is not only difficult to divide genes into two categories: those that have an effect on TRAIL-induced cytotoxicity, and those that do not, it is also somewhat meaningless to do so. The distribution of siRNA scores seen in the screen is continuous over a large portion of the total range of scores (Figure 4.6), suggesting a continuous distribution of effects. However this could reflect the distribution of siRNA mediated knock-down efficiencies and technical variation and disentangling the contribution of these different sources of variation in siRNA score is difficult given only two siRNAs targeting each gene and only two replicates for each siRNA.

The failure to identify genes previously associated with TRAIL-induced apoptosis, along with the small number of genes for which the two siRNAs calls for a critical evaluation of the design of the screen. In particular, the decision only to include two replicates of the screen, leads to the necessity of using the conservative minimum replicate summary. It would have been useful to have repeated the third replicate of the screen and then compared the results from a mean of three replicate design to those from a minimum of two replicates design.

Although previously identified genes were not identified in the screen presented here, the screen has been successful in identifying siRNAs that have a large and reproducible affect the sensitivity of cell to TRAIL-induced cytotoxicity with a low false positive rate. The

involvement of three novel genes targeted by these siRNAs was confirmed. As such the screening system presented here has proved useful for identifying new genes involved in the TRAIL-induced apoptosis pathway.

4.7.2 The Hits

Three confirmed hits were identified from the screen. Each of these genes is targeted by two siRNAs that significantly reduced the sensitivity of the transfected cells to TRAIL-induced cytotoxicity and also reduced the target mRNA more efficiently than siRNAs targeting the same gene that did not significantly reduce the sensitivity of transfected cells to TRAIL-induced cytotoxicity (Figure 4.7 and Table 4-5).

These hits were further investigated for their effect on TRAIL ligand-induced caspase activation. Both siRNAs targeting IKBKE significant reduced the activity of Caspases -8,-9 and -3/7 in TRAIL treated cells compared to siNeg transfected cells. Only one of the two siRNAs targeting each of Sharpin and LOC375449 significantly reduced the activity of the caspases, although those that did reduced the activity to a level comparable with cells not exposed to TRAIL ligand. Further, the siRNA that had the largest effect in the confirmation experiments correlated with the siRNA that that had the largest effect in the caspase activity experiments.

The effect of knock-down of the confirmed hits on sensitivity to other apoptosis inducers was investigated. siRNAs targeting both Sharpin and LOC375449 affected the sensitivity of cells to both TRAIL and FAS ligand induced cell death, but did not have a statistically significant effect on non-ligand inducers of apoptosis , expect at one concentration for one siRNA (Figure 4.10a,b, Figure 4.11a,b ,Figure 4.12a,b and Figure 4.13a,b). This is in line with the finding from caspase activity assays that both genes act to reduce the level of Caspase-8 activation. However, it must be remembered that a lack of significance does not necessarily imply that there is no actual difference. Indeed, small differences between wells transfected with siRNAs targeting these genes did show a small effect at one concentration of H_2O_2 tested (50μM for Sharpin and 100μM for LOC375449). Further, in the UV experiments, the positive control had a small effect, similar to that of one of the Sharpin siRNAs, and interpretation of this experiment was complicated by problems with the viability of negative control transfected, untreated cells. Therefore, all that can be concluded is that there is no strong evidence for the involvement of these genes in the induction of apoptosis in response to these conditions.

Knock-down of IKBKE by one, but not both of the siRNAs targeting this gene had

a large and significant effect on apoptosis induction by both TRAIL, FAS and the non-ligand inducers tested (Figure 4.10c, Figure 4.11c ,Figure 4.12c and Figure 4.13c). It is unclear why the effects were seen by only one siRNA, when the siRNAs had similar effects in the caspase assays (Figure 4.9), particularly since both siRNAs targeting Sharpin and LOC375449 had effects on ligand induced cytotoxicity, but differing effects on ligand induced caspase activity.

4.7.2.1 Sharpin

Sharpin was originally identified as a binding partner for the post-synaptic density protein Shank (Lim et al. 2001). Protein localisation studies in rat neurons showed that the protein localised in a punctuate pattern near synapses and co-localised with shank, although fractionation experiments relieved that significant portion of the protein was localised in the cytosol (Lim et al. 2001). As well as binding Shank, Sharpin is also able to homodimerise through its N terminal domain (Lim et al. 2001). Lim *et al* hypothesised that Sharpin acts as a scaffold protein. The interaction between Sharpin and Shank is mediated through the C-terminal domain, which contains a RanBP type zinc finger domain (Pfam, http://pfam.sanger.ac.uk) and is homologous to the N terminal domain of the Protein Kinase C binding ubiquitin ligase RBCK1 (Lim et al. 2001).

Strangely considering its inclusion in this targeted siRNA set, Sharpin contains neither kinase nor phosphatase domains. Its inclusion here in the kinase, phosphatase and kinase/phosphatase associated gene set could be due it similarity to a kinase binding protein. Alternatively Pfam reports a "kinase-like" domain at the N terminal end, although BLAST does not reveal similarity to any kinases in this domain.

A recent report has demonstrated that a mutation in the mouse Sharpin gene leads to a dermatitis like phenotype, which is accompanied by a multi-organ inflammatory repsonse and various immune cell and cytokine defects (Seymour et al. 2007).

The GNF expression atlas (http://symaltas.gnf.org) contains expression profiles for a large number of transcripts in a range of tissue types and in number of organisms. The GNF atlas profile for Sharpin reveals high mRNA expression in testis associated tissues and in the heart. This correlates with the finds of Lim *et al* who found expression in a wide range of tissues, but strong expression in the heart and testis. However, protein expression did not correlate with mRNA expression with strong protein expression in the brain, lung and spleen, and more modest expression in heart, liver, muscle, kidney and testis tissue (Lim et al. 2001).

Although no interaction has been reported between protein kinase C (PKC) and

Sharpin, the homology of Sharpin to RBCK1 in its PKC biding domain is interesting as PKC has been implicated in the regulation of TRAIL-induced apoptosis through control of FADD recruitment to the DISC and therefore Caspase-8 activation (Harper et al. 2003). If Sharpin were to interact with, and regulate, PKC then this would fit with the finding that Sharpin regulates apoptosis at or above the level of Caspase-8 activation (Figure 4.9) and could be a clue to its involvement. Although in the absence of experimental evidence this remains purely speculative but could serve as a hypothesis for further investigation.

There was a small effect of the stronger of the two siRNAs targeting Sharpin on UV- and H_2O_2-induced cell death. It has previously been reported that death receptor knock-outs show a subtle deficiency in radiation induced apoptosis (Finnberg et al. 2005), suggesting that the integrity of the extrinsic apoptosis pathway has an effect on activity of the intrinsic pathway.

4.7.2.2 LOC375449 (MAST4)

LOC375449 has recently been renamed Microtubule Associated Serine Threonine kinase 4 (MAST4) due to its similarity to a family of MAST kinases (MAST1, MAST2, MAST3 and MASTL) (Sun et al. 2006). These kinases, MAST4 included, have a conserved domain structure, with a conserved domain of unknown function at the N terminal end, followed by a kinase domain, a kinase C terminal domain and a PDZ domain (Pfam, http://pfam.sanger.ac.uk). PDZ domains are involved in protein/protein interactions and in MAST1-3 are involved in the binding of these kinases to the tumour suppressor PTEN (Valiente et al. 2005). However it was reported that this is not the case for MAST4. MAST1 and MAST2 are also involved in regulation of TNFα mediated activation of NF-κB, through their phosphorylation of TRAF6 and its subsequent ubiquitination and degradation by the proteasome (Xiong et al. 2004).

The GNF atlas reports a generally even expression across tissue types, with perhaps a slight increase in expression in immune cells. Oncomine (http://www.oncomine.org) is a database containing the results from microarray transcription profiles for cancer cells. Oncomine reports 11 independent studies where MAST4 has a significantly increased expression in Estrogen Receptor positive breast cancer samples compared to Estrogen Receptor negative breast cancer samples. The Estrogen Receptor is a ligand activated transcription factor that has long been associated with breast cancer and several breast cancer treatments such as tamoxifen function by targeting this receptor (reveiwed in Ali, Coombes 2000). A study of the effects of transfection of known oncogenes into primary human

mammary epithelial cells revealed that MAST4 was strongly up-regulated when cells were transfected with activated H-Ras (Bild et al. 2006).

The COSMIC database contains details of somatic mutations in cancer cells lines (http://www.sanger.ac.uk/genetics/CGP/cosmic/). Currently the database focuses on results from systematic resequencing of kinases from a large range of cancer cell lines and primary cancer samples. The database lists six mutations in the MAST4 gene, 3 missense substitutions, 2 nonsense substitutions and one complex substitution resulting in a missense and a nonsense substitution. None of these mutations are in the kinase domain, or any other Pfam domain, as would be expected for an activating mutation. Two of these mutations (E784K and E2276*) were found in an Estrogen Receptor positive primary breast cancer sample shown to have a mutator phenotype (Stephens et al. 2005). Analysis of the spectrum of mutations found suggested that while some of the mutations found in this sample probably contributed to the oncogenic phenotype, the majority were "passenger mutations" and did not contribute to the phenotype.

Since MAST4 is a member of a protein family, it is worth considering the possibility that the siRNAs designed to target MAST4 are also targeting related genes from the same family. Pairwise alignment of the sequence of the MAST4 transcript targeted by the siRNAs used in this study to each of the other 4 members of the MAST family showed that MAST4 shares 39.9%, 49.8%, 40.91%, and 34.3% sequence similarity to MAST1, MAST2, MAST3 and MASTL respectively. When deciding if it is possible that MAST4 siRNAs target other members of the MAST family the alignment of the siRNA to the sequence is of more importance than the global alignment of the two targeted sequences. The best alignment between each of the MAST4 siRNAs is shown in Table 4-7. Thus it seems unlikely that these siRNAs are targeting these transcripts. This is not unexpected since siRNA design includes a specificity check for sequences that match this siRNA with a high sequence identity.

siRNA	*MAST1*	*MAST2*	*MAST3*	*MASTL*
siLOC375449.1	12/21	13/21	14/21	12/21
siLOC375449.2	14/21	15/21	14/21	12/21

Table 4-7 – Best matchs of MAST4 siRNAs against other members of the MAST family
Sequences of MAST4 familiy members were retrieved from the ReqSeq database and aligned to the sequences of MAST4 siRNAs using a smith waterman algorithm.

Here it has been shown that MAST4 is involved in the regulation of sensitivity of cells to ligand mediated apoptosis at or above the level of Caspase-8 activation. Expression experiments seem to suggest that it is downstream of RAS (Bild et al. 2006), which is of interest as activation of the RAS pathway is known to be involved in the sensitisation of cells

to TRAIL-mediated apoptosis (see 1.3.2.6). It has also been shown to be overexpressed and mutated in estrogen receptor positive breast cancer samples.

4.7.2.3 IKBKE

The IKBKE gene codes for the Inhibitor of kappa B Kinase ε protein (IKKε, also known as IKK-i). Upon stimulation IKKε phosphorylates Inhibitor of kappa B alpha (IκBα) targeting it for degradation and subsequently leading to activation of NF-κB (Shimada et al. 1999). IKKε has also been shown to regulate the constitutive activity levels of NF-κB by direct phosphorylation of both the cRel and RelA NF-κB subunits (Adli, Baldwin 2006, Harris et al. 2006). IKKε is also known to directly regulate the interferon anti-viral pathways by the phosphorylation and activation of the IRF-3 and IRF-7 (Fitzgerald et al. 2003, Sharma et al. 2003). Both the NF-κB and innate antiviral responses are known to be involved in the regulation of TRAIL sensitivity (see 1.3.2and 1.3.4). IKBKE amplification has recently been shown to be important in breast cancer oncogenesis (Boehm et al. 2007).

In addition to a kinase domain the IKKε protein, like Sharpin, contains a ubiquitin like domain (Pfam, http://pfam.sanger.ac.uk). This domain has been shown to be required for the regulation of antiviral response genes (Ikeda et al. 2007). Other IKK proteins also contain ubiquitin-like domains and the IKKβ ubiquitin-like domain is required for its NF-κB activating activity (Suzuki, Nakabayashi & Takahashi 2001).

The GNF expression atlas reports that IKBKE is expressed at a low level in most tissues, but expression is particularly strong in immune related cell types, particularly T cell lineages and dendritic cells. The strongest expression is found in CD8+ T cells, which is interesting given the role of TRAIL in preventing the secondary expansion of "Helpless" CD8+ T cells (Janssen et al. 2005).

Knock-down of IKBKE by siIKBKE.1 reduces sensitivity to not only TRAIL-induced apoptosis, but also to FAS-, H_2O_2- and UV--induced cell death. Caspase assays showed that IKBKE acts at or above the level of Caspase-8 activation. The effect on H_2O_2- and UV-induced cell death could be mediated by the same action of Caspase 8 on the intrinsic death pathway mentioned above in connection to the effect of Sharpin knockdown on UV- and H_2O_2-induced cell death. However the effect of IKBKE knockdown is much stronger than the effect of Sharpin knockdown. If IKBKE is involved both the NF-κB activation, which it self has multiple effects on apoptosis regulatory proteins, and the interferon response, it is possible that IKBKE is involved in regulation of apoptotic pathways at multiple points.

4.7.3 Conclusions

In the previous chapter an assay for sensitivity to TRAIL-induced cytotoxicity was developed and used to compare different methods for performing RNAi-mediated screens for genes which affect the TRAIL-induced cytotoxicity pathway. Here the findings from that work were applied to a screen of 897 kinases, phosphatases and kinase/phosphatase associated genes. The aims of this study were two fold. Firstly the screen acts as a gene discovery exercise in it own right. Secondly the relatively small screen acts as a pilot for larger screens.

siRNAs targeting 897 kinases, phosphatases and kinase/phosphatase associated genes were transfected into HeLa cells and the sensitivity of cells to TRAIL determined. Each gene was targeted by two siRNAs and 2.75 replicates of the screen were performed, which was later condensed into 2 "high-quality" replicates. The survival of cells in each well was determined. Data was normalised using the plate median and scored by standardising the normalised survival and selecting the minimum of the two replicates as the score for each siRNA.

The correlation between the two replicates ($r = 0.65$) was much stronger than the correlation between the scores of different siRNAs targeting the same gene ($r=0.20$). This was reflected in the finding that seven of the top ten siRNAs retested for confirmation showed a significant effect on TRAIL-induced cytotoxicity, but only in three cases did additional siRNAs targeting these genes also significantly reduce the sensitivity of cell to TRAIL. The difference in the proportion of either genes or siRNAs selected on the basis of both siRNAs targeting a gene scoring higher than a threshold is not statistically different. However, selecting genes using the single high-scoring siRNA method, gave a higher number of candidates, and thus a higher number of eventual hits. The two-hit method is more restrictive, and risks missing strong hits where one siRNA has failed due to poor siRNA design, or for technical reasons. Selecting genes for follow up on the basis of two siRNAs targeting the gene scoring higher than a threshold determined using the positive controls did not successfully identify any additional genes which could be confirmed as having a role in TRAIL-induced cytotoxicity, and thus fails to increase the rate of confirmation despite the increased restrictiveness of the candidate selection. More candidates could be selected by reducing the threshold, yet this would almost certainly reduce the confirmation rate. Thus the screen is accurate at selecting siRNAs with an effect on the phenotype, but less accurate at selecting genes that are involved. Additionally, only one siRNA targeting genes previously implicated in TRAIL-induced apoptosis scored higher than this threshold (in addition to the

positive controls).

In both identifying previously associated genes and identifying novel genes (as opposed to siRNAs) the screen was clearly less successful than might have been hoped. However, while the screen clearly has not identified all genes in the set which are involved in the pathway, several high confidence genes have been isolated.

Taken together, these results suggest that the screening system described here can be successfully used to identify siRNAs which affect sensitivity to TRAIL-induced cytotoxicity. Rigorous follow up of these siRNAs can allow for the identification of novel genes involved in the pathway. They demonstrate that a follow up strategy based on following the top siRNA hits is at least as good as and possibly better than following genes that are hit by two weaker siRNAs. The screen presented is less successful at defining the absolute involvement of any one gene in TRAIL-induced apoptosis. These results emphasize the importance of careful confirmation of hits from siRNA screening experiments.

A scale up of this experiment to a larger gene set could be expected to produce a similar standard of results. Although it is unlikely that all or genes involved in the process would be isolated from a genome-scale screen, novel targets that would not have been otherwise isolated should be identified.

Three genes with a novel involvement in the TRAIL-induced apoptosis pathway have been identified. Their involvement in TRAIL-induced apoptosis as well as TRAIL-induced cytotoxicity was demonstrated using luminescent caspase activity assays, which also showed that these genes acted at the level of Caspase-8 activation. Study of the effects of knock-down of these genes on other inducers of apoptosis demonstrated that Sharpin and MAST4 are involved in ligand induced apoptosis while IKBKE is involved more generally in regulation of apoptosis, possibly acting at multiple points in the pathway. Little is known about both Sharpin and MAST4, although mutation of Sharpin has been implicated in a spontaneous inflammation phenotype in mice(Seymour et al. 2007) and MAST4 is over-expressed in estrogen receptor positive cancers and in response to RAS activation(Bild et al. 2006), and is mutated in some breast cancers samples (Stephens et al. 2005)□. IKBKE is involved in regulating the constitutive levels of NF-κB(Adli, Baldwin 2006, Harris et al. 2006), a transcription factor with multiple effects on TRAIL sensitivity and also in the regulation of anti-viral responses, which is interesting as virally infected cells are sensitive to TRAIL-induced apoptosis.