Application and development of methods towards the target identification of biologically-active small molecules

Rohitha Poornami SriRamaratnam

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2011

Abstract

Application and development of methods towards the target identification

of biologically-active small molecules

Rohitha Poornami SriRamaratnam

Small molecules have played an important role in defining the functions and identities of numerous proteins involved in fundamental biological processes as well as pathways involved in disease. Chemical genetics represents the formalization of this process into a defined field desiring to achieve the breadth and specificity of classical genetics. In order to gain full advantage of a small molecule's ability to perturb the cell for novel or desired phenotypes, a complete understanding of the molecule's mechanism of action must be achieved. Identification of the biological targets of a molecule represents the most direct approach to attaining this knowledge.

In our strategy to find novel mechanisms to target cancers with oncogenic *RAS* mutations, we have used small molecules to probe specific weaknesses of this cancerous network through synthetic lethal screening. One molecule identified in these screens, RSL3, attracted interest as a candidate for target identification studies because of its potent lethality and potentially unique mechanism of action. We used an affinity chromatography approach to directly isolate binding partners of RSL3 by modifying the molecules structure to incorporate various affinity tags. Through these experiments we ultimately identified a number of interesting candidate targets. Investigations validating these targets suggest that multi-targeted modulation of antioxidant and prostaglandin networks may be a mechanism for selectively killing cancers with oncogenic *RAS.*

The identification of biological targets of small molecules poses a difficult challenge to the field of forward chemical genetics. Thus, we attempted to optimize a unique method for target identification, the yeast three-hybrid system (Y3H), which detects small molecule-protein interactions through a transcriptional assay *in vivo*. We created a version of our Y3H system that incorporated a covalent anchor and compared it with the existing state-of-the-art, which uses a high affinity non-covalent anchor. Transcriptional assays indicated our new system was functional, but surprisingly could not improve upon the original Y3H system. These results highlight the complexities of manipulating ligand-receptor interactions *in vivo.*

Table of Contents

List of Figures and Tables

Figures

Tables

Acknowledgements

There are many people to whom I owe gratitude for their mentorship, friendship, help and love over the past five years so I will take this opportunity to briefly acknowledge them. First, I would like to thank my advisors, Brent and Virginia, for taking on a somewhat indecisive student trying to find her way when she first joined the Columbia PhD program. By conducting research in your labs you have instilled in me the same passion for the field of chemical biology that you both exhibit in your work and which I admire greatly. Thank you for giving me the opportunity to be a part of the research that you have established here at Columbia. I particularly want to acknowledge how grateful I am that you have always allowed me to choose my path during my graduate career – even at times when I'm sure you wished or thought that I should take another direction in my project, you allowed me to stubbornly pursue what I was most interested in and I hope that we will all agree that it was worth the risk in the end.

To my fellow lab colleagues, you are among my closest friends that I have made during my graduate career. I strongly believe that work environment makes a big impact in job satisfaction and having your support, both in terms of friendship and scientifically, has been essential to making it through the ups and downs of research. I particularly appreciate your support during recent times of hardship such as the preparation for this defense :)

I would like to specifically acknowledge the work and support of Lewis Brown and his staff at the Comparative Proteomics Center who have been my major collaborators performing the proteomic analysis for samples I have prepared in affinity chromatography studies with RSL3. Additionally, this thesis work also includes experiments, reagents and methods developed by my colleagues: Dr. Wan

Seok Yang, Dr. Scott Dixon, Dr. Rachid Skouta, Dr. Adam Wolpaw, Dr. Katie Lemberg, Dr. Laura Wingler, Kenichi Shimada, Chaoran Jing and Zhixing Chen.

I also thank my committee members, Dr. Ann McDermott and Dr. Ruben Gonzalez for serving on my graduate committee for the past five years as well as Dr. Kenneth Olive for serving on my defense committee. I greatly appreciate the input and feedback they have provided on my work.

To all the friends I have made here in New York, it has been a great experience meeting such a diverse group of people and going through grad school life in the big apple! Thanks for all the times we have spent exploring this city when we manage to find our way out of the lab. Special thanks to my roommate of five years, Burce, who has been a constant companion to me at home and at work, it has been great to have a friend working one floor above me to share day to day life with as well as one with expertise in protein biochemistry! I still hope to purify a protein one day with you.

To Raul, who has witnessed many important moments in my PhD, sometimes without even realizing it, I am glad that you now understand what I was going through after kindly proof reading almost all the pages of this manuscript. You have given me so much support and never grumbled that experiments in lab always seemed to come first in terms of our schedule.

To my parents, I dedicate this work to you. You have supported and cheered me from afar as well as put up with my monosyllabic responses to your questions about how research was going for many years. I always appreciate your concerns for me and my research work although I don't show it often enough.

To my parents.

Chapter 1: Introduction

1.1 Small molecules as tools for biological discovery

The functioning of the cell is made up of millions of diverse interactions between molecules of different sizes. When we purposefully introduce an exogenous molecule into the cell, understanding the consequences of this action involve identifying the corresponding molecular partners within the cell that interact with our molecule of interest. These molecular interactions are mainly comprised of 'binding events', and can take the form of reversible and irreversible covalent interactions as well as non-covalent interactions. In particular, this work is interested in the interactions of introduced ligands that are classified as small molecules (typically organic and under 2000 $gmol^{-1}$) with the endogenous cellular constituents.

The concept of a small molecule interacting with a particular component of the cell or biological system dates back to the work of Paul Erlich and others who established the theory of a 'receptor' - a single binding target of a small molecule that is responsible for the phenotypic consequences of the ligand's introduction. This concept originated from his work in histological staining where he observed that the organic dyes he was using reacted differentially with distinct bacteria and tissues [1]. His work in the staining of blood cells consolidated these ideas and provided him with the insight that this specificity was due to the unique interaction of a chemical with a distinct part of the cell. Although his model was initially met with a great deal of skepticism, he further developed his theory suggesting that a chemical could specifically target components in the cell not only for visualization, but also for therapeutic purposes by killing it (in particular, pathogenic microbial and tumor cells). He described these chemicals as 'magic bullets' and they are still the major pursuit of the drug industry today [2, 3].

Many small molecules exert phenotypic changes via protein-binding [4]. To date, proteins have played the most important role as the targets of molecules used as drugs [5]. Moreover, over the past century small molecules have been used as experimental probes to define the role and even the identity of proteins involved in fundamental biological processes such as cell cycle progression, cell proliferation and survival, DNA synthesis, protein translation and degradation, chromatin modification and microtubule polymerization to name a few (Table 1) [6-16]. These examples confirmed the idea that small molecules can be used to alter the function of gene products to gain biological understanding, providing similar power to biological research as the utilization of mutations in classical genetics.

Applying small molecules as tools to explore biology became formalized as a founding principle of the field of chemical genetics, in which compounds with novel or desired biological functions are systematically and actively sought, instead of serendipitously found [17]. Access to large libraries of small molecules (both synthetic and natural products) combined with the technology of high through-put screening (HTS) has enabled chemical genetic screens on the scale of those in genetics, except using molecules as perturbagens.

Small molecule	Protein target	Biological pathway/Mechanism of action
FK506 (Tacrolimus)	FK506 binding protein 12 (FKBP12)	As a small molecule-protein complex both FK506-FKBP12 and Cyclosporin A-cyclophilin inhibit calcineurin and ultimately IL2 secretion which blocks T-cell activation
Cyclosporin A	Cyclophilin	
Wortmannin	phosphatidylinositol 3-kinase (PI3K)	Blocks PI3K/Akt pathway. Inhibits cell growth, proliferation, differentiation, motility, survival
Rapamycin (Sirolimus)	FK506 binding protein 12 (FKBP12)	In complex with FKBP12 inhibits mTOR and thus blocks mTORC1 pathway, (downstream of PI3K/Akt), particularly inhibiting translation
Cycloheximide	Ribosome	Blocks translocation of the tRNA and thus inhibits translation
Lactacystin	Proteosome	Inhibits protein degradation
Trapoxin B	Histone deacetylases	Inhibits deacetylation of chromatin and affects gene expression
Doxorubicin & Daunorubicin	DNA/Topoisomerase II	Inhibits progression of topoisomerase II resulting in G2 arrest
Colchicine	Tubulin	Microtubule destabilizer, inhibits mitosis

Table 1.1: Small molecules that have been used to elucidate the function or identity of their protein binding partners

Although parallels between classical and chemical genetics are often made, small molecules and genetic mutations are not equivalent in both positive and negative ways. In particular, chemical genetics still lacks the selectivity and breadth of classical genetics. While methods have been developed to delete, overexpress or mutagenize every gene in the genome, small molecules that modulate the whole proteome have not been discovered. Achieving similar generality was one of the original goals laid down by founders of the field [17]. Instead, at least among molecules used as drugs, the current coverage is only 2% of the human proteome [18]. Moreover, roughly half of the proteins that are targeted by drugs tend to fall in the same four protein families [5, 19]. These proteins tend to have convenient hydrophobic pockets in which hydrophobic compounds can happily bind. Finding small molecule modulators for the rest of the 'undruggable' proteome is now a burgeoning area of research interest [18]. Additionally, improved methods to identify the targets of all the small molecule hits emerging from HTS could potentially reveal that a few more classes of targetable proteins exist already.

Very few small molecules bind and modulate only one target in the cell with the precision with which molecular genetics can selectively modify a gene [4, 20]. For the majority of them, the 'one drug, one target' idea does not hold true. Often it is the case that as more studies are directed at a molecule, their actual promiscuity is brought to light. Even imatinib (also known as Gleevec), a clinically used drug celebrated for its specificity against c-ABL kinase and the BCR-ABL kinase oncogene, has been shown to bind the receptor tyrosine kinases platelet-derived growth factor receptor (PDGF-R), c-kit and DDR1 and an oxidoreductase, NQO2 [21-24].

The quality of promiscuity is somewhat of a double-edged sword. While it may make it more difficult to use a small molecule as a probe in dissecting an area of biology, if a small molecule hits multiple targets involved in different pathways of a disease it can perform better as a drug. Network analysis studies indicate that multi-target rather than single-target drugs may be necessary to overcome

the phenotypic robustness of cell disease networks due to redundant pathways [25]. Additionally, modulating diverse targets may give the small molecule therapeutic potential against different diseases [20]. In the case of imatinib, the discovery of its additional targets has created the opportunity for its use in the treatment of cancers with increased c-kit activity [23]. This highlights the importance of fully identifying a small molecule's 'target profile' to enable complete understanding of all its potential biological applications as well as predict undesired off-target effects [20].

To their advantage, small molecules also possess greater flexibility in application. Compound concentration can easily be varied to study the effects of gradual modulation of a target. In classical genetics, the equivalent would be a conditional knockout or temperature sensitive allele of a gene, which is much more difficult to perform. Dose-dependent phenotypes can vary greatly as illustrated by comparing the chemical and genetic 'inhibition' of the yeast cyclin-dependent kinase, Cdc28 [26, 27]. Genetic approaches showed that temperature sensitive alleles of Cdc28 exhibited a G1 arrest phenotype, while treatment at low doses of a drug-sensitive Cdc28 with a chemical inhibitor resulted in arrest in the G2/M phase of the cell cycle. Interestingly, cell cycle-synchronized yeast treated with high concentrations of the inhibitor showed the same G1 arrest, illustrating different phenotypic response depending on the level of inhibition. This conditional effect was not reflected in the genetic approach [26]. Further, one has temporal control over the addition of a small molecule, and their action is reversible in some cases where it can be washed out to revert back to the original phenotype. Small molecules can be applied to a wide variety of cells or animals, while genetics screens are largely restricted to smaller organisms due to size constraints and rates of reproduction [4].

Further distinctions between chemical and classical genetics stem from the fact that studying the complete loss of a gene product is often not the same as studying its inhibition by a small molecule. Small molecules are able to modulate their targets with more nuances; inhibiting a particular domain

while the others remain functional, inhibiting enzyme activity or modulating a protein allosterically. One other advantage is that effects due to the inhibition of an individual target's function but not its absence from other protein complexes can be observed in isolation – a problem when a gene is knocked out, although this can be achieved with missense mutations or conditional alleles. Additionally, the effects of cellular compensation are greater in genetic experiments where the cell has more time to respond [26].

Finally, some small molecules that are identified in screens of disease models can become drugs or be used as a source of new targets for drug discovery endeavors [4]. Thus due to their complementary natures both classical and chemical genetic methods are valuable in elucidating the diverse functions of proteins.

Similar to genetics there are two directions of approach when probing a biological system. In forward chemical genetics, a screen is performed looking for compounds that produce a particular phenotype, the final result of a small molecule-target interaction. In contrast, reverse chemical genetics focuses on searching for compounds that bind to a particular target, in a way that is rationally predicted to affect its activity, in order to produce a desired phenotype in the cell.

At first glance, forward chemical genetics has many appealing aspects. The small molecule screen searches directly for the phenotype of interest, the small molecule-target interaction already occurs in a physiological, competitive cellular environment and it is performed as an unbiased, blind search more likely to yield novel targets and pathways. The main stumbling block of this approach however, results from the black box nature of the screen. For any hit molecule that is identified, its target in the cell is unknown [1] . The mechanism by which the molecule exerts its phenotype must be deciphered.

FK506, also known as the widely-used immunosuppressant Tacrolimus, is a shining example of the forward chemical genetic process. It is a natural product macrolide discovered in a screen for

compounds with immunosuppressant activity. Its molecular target was subsequently identified to be a peptidyl-prolyl isomerase that was then named FKBP (for FK506-binding protein) [28, 29]. The compound's immunosuppressant activity turned out not to be a result of inhibition of FKBP, but due to the inhibition of another target, the protein phosphatase calcineurin, by the FK506-FKBP complex. When calcineurin is activated, it dephosphorylates a transcription factor NF-AT (nuclear factor of activated T cells) enabling its entry into the nucleus and activating expression of genes involved in T-cell activation such as interleukin-2 (IL-2) [21, 30]. Due to its resulting immunosuppressive effects, it is now used as a drug to lower the risk of rejection during organ transplants.

In order to fully harness small molecules as tools for biological discovery, it is necessary to identify their targets within the cell. FK506 could have remained a useful immunosuppressant drug without our knowledge of its target, but the full elucidation of the biological pathways that involve its receptor, FKBP12, and their direct target, calcineurin, with its downstream effectors would have remained a mystery. Thus it is clear that the identification of biological targets of small molecules is an essential step in bringing the forward chemical genetic process to fruition.

Target identification (target ID) is the generic name given to the process of determining the binding partners of a small molecule in the cell. Unfortunately, target ID has proven notoriously difficult, gaining it the negative reputation of being a 'bottle-neck', 'the rate-limiting step' and 'the missing link' in forward chemical genetics [29, 30]. However, when it is performed successfully, major advances in our knowledge of the biology of the cell as well as disease pathogenesis can be gained. The following provides an overview of the most commonly used techniques for small molecule target identification. The first methods used to identify biological targets involved the direct isolation of the molecular target *in vitro* either through radiolabeling or affinity chromatography.

1.2 Direct methods of target identification

1.2a Radiolabelling

Radiolabelling involves the incorporation of a radioactive isotope within a small molecule for detection purposes. The first use of the radioligand assay was in the isolation of the *lac* repressor using a radioactive version of isopropyl thio galactopyranoside (IPTG). IPTG containing a radioactive isotope (^{14}C) was incubated with fractionated cell extracts and assayed for binding by equilibrium dialysis to isolate fractions that retained radioactivity within the dialysis bag and thus the molecule's binding partner [31]. Subsequently this technique was used in the target identification of many endogenous neurotransmitters and hormones such as the opioids and acetylcholine [32, 33].

Radiolabeling has also been successfully performed with a tritiated (^{3}H) form of the natural product, colchicine. Tritiated colchicine was incubated with tissue lysates, fractionated by size-exclusion chromatography and bound radioactivity was detected using a scintillation spectrometer. The researchers further noticed that binding activity was greatest in lysates derived from sources with high levels of microtubules. Combined with the knowledge that colchicine could cause the disassembly of these structures, they speculated that the binding protein they observed could be a subunit of microtubules, later identified as tubulin [34].

Advantages of this method include the incorporation of a relatively small modification into the molecule (compared with other target ID techniques discussed later) that is unlikely to disrupt binding to targets and gives the radioligand a very similar target profile to the original molecule. The technique is often used with molecules that bind covalently to their targets, but as illustrated in the example with colchicine, it can also be used with high affinity ligands. The main drawback is the lack of a method for enrichment of the target proteins that is addressed in the other method classically used for target ID, affinity chromatography.

1.2b Affinity chromatography ('pull down')

Affinity matrices

The first methods using affinity chromatography (commonly referred to as a 'pull down' experiment) involved immobilization of a small molecule on a matrix, commonly a resin of sepharose or agarose. This procedure was originally conceived for the isolation of enzymes based on 'catalytic specificity' (in other words, specific affinity to a ligand or inhibitor of the enzyme) rather than other physical properties of the protein. Through binding to the ligand or inhibitor, enzymes would be isolated from the lysate on a solid support, washed free of remaining non-binding lysate components and then released by denaturing conditions or excess competitive ligand [35] (Figure 1.1A). This purification method has since been extended to the isolation of proteins binding to compounds with unknown targets.

Trapoxin is a small molecule that was known for its interesting histone deacetylation activity and ability to cause cell cycle arrest in mammalian cells. This knowledge spurred the identification of the protein responsible for its activity, a mammalian histone deacetylase (HDAC), through the 20-step synthesis of an affinity matrix named 'K-trap' that incorporated the trapoxin structure along with a tether for attachment to the matrix. K-trap was subsequently used to purify trapoxin-binding proteins from nuclear fractions of bovine thymus tissue. Analysis of the enriched proteins by SDS-PAGE revealed a 46kD protein band that microsequencing demonstrated was related to a yeast homolog with transcriptional regulatory activity [8].

Current applications of affinity chromatography are similar with the difference that affinity-purified proteins are now typically identified using mass spectrometry (MS). Activated resins with amino, thiol, N-succimidyl ester and carboxylic acid functional groups are readily available commercially to facilitate tethering of small molecules.

Affinity tags

An alternative to affinity matrices is the affinity tag that utilizes a moiety with high affinity for a protein or antibody. Biotin is the most popular example as it possesses femtomolar affinity to various proteins such as avidin and streptavidin. Small molecules are connected to biotin via a linker and bound proteins are 'pulled down' through the use of resins coupled to streptavidin (Figure 1.1B). Streptavidin is typically favored over avidin due its lower levels of glycosylation making it less charged, which reduces non-specific binding with proteins in the lysate. Commercial unglycosylated versions of avidin such as Neutravidin also exist [5].

The strength of the biotin-streptavidin interaction prevents its disruption except under strongly denaturing conditions that are required to elute proteins from the matrix. This is an advantage in the sense that it allows for rigorous washing to remove non-specific protein binders. However, due to the inherent stickiness of many abundant proteins in cell lysates, this normally can never be fully achieved and thus non-specific binders remain a problem. Monomeric versions of avidin have lower affinity for biotin and thus excess biotin can be used for elution.

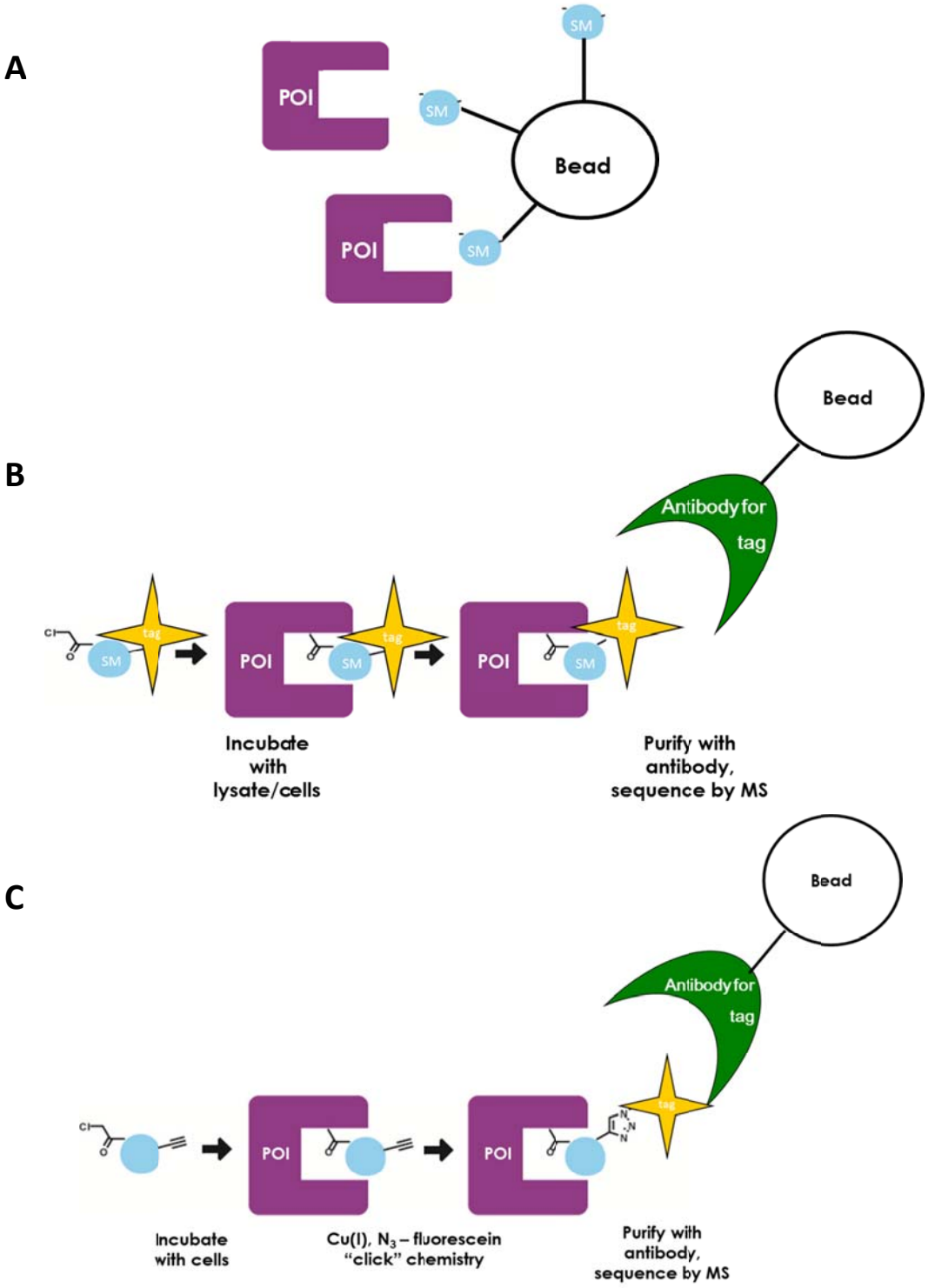


Figure 1.1: Three types of affinity chromatography experiment

(A) Affinity matrices: small molecules (blue, SM) are directly tethered to the bead (resin or matrix) and proteins of interest (POI) are specifically bound to the small molecule while the rest of the lysate is eluted.

(B) Affinity tags: A covalently binding small molecule is illustrated, but non-covalent molecules can also be used (blue, SM). A tag (yellow, star) is incorporated into the molecules structure via a tether. The small molecule-bound proteins are pulled out of the lysate using an antibody or protein bound to a bead (resin or matrix) with high affinity for the tag.

(B) Alkyne tags: A covalently binding small molecule must be used (blue, SM). An alkyne moiety is incorporated into the molecules structure. Compounds are incubated with cells or lysates and then click chemistry reactions are used to affix a tag (yellow, star) to all small-bound proteins. These are then pulled out of the lysate using an antibody or protein bound to a bead (resin or matrix) with high affinity for the tag.

In cell vs. *in vitro* treatment

A major advantage of affinity tags is that they can also be used in the treatment of cells instead of cell lysates. Incubating compounds in cells provide the benefits of determining a compound's target profile in the natural cellular environment where proteins are properly folded and compartmentalized. Cell lysis introduces bias against membrane proteins and other proteins with poor aqueous solubility. Proteins can also partially unfold *in vitro* and expose hydrophobic domains that results in non-specific binding and aggregate formation. Additionally, if a compound is metabolized into its active form in the cell, binding may not be detectable *in vitro* [20]. Regrettably, in cell treatment has generally only been applied with covalent molecules [36, 37] presumably because non-covalent small molecule-protein interactions cannot survive the lysing process [5].

Tether effect

A major hurdle in all affinity chromatography approaches is the need to derivatize the molecule with a long tether either for coupling to the matrix or for addition of a tag. Often, attachment of a tether affects the binding of the molecule to its targets, which is reflected in a loss in its biological activity or could result in an altered target profile. Extensive structure-activity relationship (SAR) studies must be performed to identify areas in the molecule essential for binding and areas that are dispensable and could be used as a site of attachment for an affinity tag or tether. This can be done through the testing of analogs that are commercially available or made synthetically in the laboratory.

Ideally a tether should be hydrophilic to minimize non-specific binding, for this reason polyethylene glycol (PEG) linkers are typically used [20]. One study on the optimal length of a linker for target ID purposes reported that longer linkers give better results [38]. Intuitively this makes sense as it allows the small molecule to assume different conformations in solution and also reduces the steric

effects of the resin. In the same study, linkers with a defined structure derived from a polyproline helix were used as a 'fishing rod' to extend the molecule away from the bead.

Whenever a molecule is modified with a tag, it is extremely important to determine whether its biological activity is retained. One disadvantage of affinity matrices is that the activity of the small molecule cannot be tested with the steric bulk of the bead attached. Instead, the molecule with the tether used to link it to the bead is assessed. In contrast, tagged small molecules can be fully characterized biologically.

Alkyne/azide tags

A solution to avoid the introduction of a tether was provided in the adaption of approaches developed for activity based proteomic profiling (ABPP). ABPP utilizes suicide inhibitors targeted towards specific classes of enzymes to profile their activities under specific conditions. In order to monitor binding events of their substrates they exploit a labeling strategy based on the popular copper (I)-catalyzed variant of the azide-alkyne Huisgen cycloaddition reaction, often dubbed 'click chemistry' [39, 40]. By incorporation of an alkyne into their suicide substrates, active enzymes can be tagged after binding by performing the click reaction with an azide-linked tag attached to rhodamine for visualization purposes or to a biotin for purification purposes [41] (Figure 1.1C). Use of the alkyne as their tag enables the minimal modification of only two additional carbons on a small molecule, which greatly increases the chances of retaining close to original activity. Furthermore, the later addition of an azide-linked tag provides modularity in its application. The technique has only been applied to covalent molecules like the suicide inhibitors used in ABPP, mainly due to the requirements of the click chemistry reaction [42, 43]. However, applications have been demonstrated with non-covalent molecules that were additionally modified with a photoaffinity label [44].

Photoaffinity labeling

As illustrated above, there are many advantages to target identification efforts involving covalent small molecules. Molecules that bind covalently to their targets typically have reactive electrophilic groups that are susceptible to attack from nucleophiles within their protein target. Engineering an electrophilic group into the molecule is not trivial unless good knowledge of the target structure is known *a priori*, which is not generally the case in phenotypic screens. It is difficult to predict the positioning of a suitable nucleophile close to the binding site and this is more likely to increase non-specific reactivity with other nucleophiles in the cellular milieu [5]. Instead, attachment of a photoaffinity label can be used to turn a molecule into a covalent binder.

A photoaffinity label consists of a photoreactive moiety that is inert under normal conditions but forms a highly reactive radical at specific wavelengths of UV irradiation. The radical then forms a covalent bond with nearby backbone or side chain carbons of target proteins. The beauty of this is that the molecule first binds to its target normally prior to covalent attachment via the photoreactive group. The most common photoaffinity labels used are benzophenones, aryl azides and diazirines, each of which possess different efficiencies of photo cross-linking, stability and additional steric bulk. The aryl azide is the smallest and easiest to incorporate into small molecules that often already contain a phenyl group.

Covalent bond formation by the radical generated upon photolysis is affected by the spatial orientation of the photoreactive group and distance relative to the small molecule. Positioning of the photoaffinity label in the binding pocket of the protein is crucial for the efficiency of the reaction. Overlapping with a residue that forms a contact with the small molecule would result in lower binding affinity but increase efficiency of labeling, while pointing out into solution would improve binding but result in poor covalent labeling – thus a balance needs to be achieved. One study observed a

discrepancy in labeling efficiencies achieved when using different enantiomers of a photoreactive unnatural amino acid, 4-benzoyl phenylalanine illustrating the dependence on spatial orientation [45]. Labeling efficiencies are also optimal in *in vitro* methods while poor within the cell where conditions for binding are more ideal.

Elution

The frequency with which 'non-specific' binding is mentioned in target ID literature indicates the magnitude of the problem. One way to tackle this is by performing more 'specific' elutions of the protein from the resin. Non-specific elutions are performed with high concentrations of sodium dodecyl sulphate (SDS), low pH and high heat. Specific elutions are most easily achieved using an excess of small molecule to compete off high affinity targets from the resin, thus selectively enriching the elution with specific targets while the background of matrix-binders and low affinity small molecule interactions remains on the matrix. In order to achieve this, the small molecule must possess high aqueous solubility. Alternatively, incorporation of a protease recognition sequence into the linker can facilitate cleavage using proteolysis, both TEV and HRV3C proteases have been used [46, 47]. The option of enzymatic elution is particularly attractive for the case of covalent molecules.

1.2c Strategies to improve the identification of 'real' targets

Performing a target identification experiment is often described as a 'fishing expedition' [30]. Disregarding the negative implications, analogies can be found in the concept in order to increase one's chance of success.

Use good bait

Success in target identification attempts has been observed to be proportional to the affinity of the small molecule and the abundance of the target [20, 48]. While no control can be gained over the latter, the choice of molecule can be carefully made. High potency in the biological activity of molecules

has been used to imply high affinity and selectivity towards a target [4]. This is clear empirically in the examples of specific natural product inhibitors used in the past (FK506 and rapamycin bind with low nanomolar affinity to FKBP). The importance of affinity is believed to be related to the molecule's kinetic half-life of binding that determines whether it will endure the rigors of the washing steps required to isolate specific over non-specific interactions [30]. While potency is often a good (or the only) indicator of affinity, it should be noted that it is not definitive. Small molecules can have high potency for a number of reasons including the modulation of multiple targets that synthetically produce a desired phenotype even though no particular target interaction is of high affinity or due to the low abundance of a target [47]. If the target is of high abundance, detection is greatly facilitated. Identification of low abundance targets require large amounts of cell or tissue lysate and increased fractionation. In the case of FK506 the situation was further ideal, as discovery was facilitated by the high abundance of its target FKBP [30].

Finding the right fish

A major problem in pull downs (affinity purification experiments) is sorting through the list of proteins (or bands on a gel) and identifying which of these are 'true', biologically relevant targets. As mentioned before, a number of proteins bind non-specifically to the matrix or the small molecule with low or no affinity. In the case of affinity matrices, the latter is reported to be the greater problem as abundant, sticky proteins are attracted to the localized hydrophobicity of the small molecules on the matrix [20]. One suggestion to address this problem has been to increase the ratio of fish to bait by using large quantities of protein lysate and decreasing the amount of small molecule loaded on to the resin [5]. Here, target proteins are forced to compete for the bait leading high affinity binders to emerge while low affinity binders are overpowered by interactions with the rest of the lysate. For example, mammalian cell extracts have even been spiked with *E. coli* lysate to achieve higher protein concentrations [30]. Additionally, non-specific matrix binders can be removed from the lysate by

incubation with resin-only controls that have been blocked with ethanolamine in what is described as ‘pre-clearing'.

Another strategy involves the use of good controls that enable elimination of non-specific interactions in post-experiment analysis. As the majority of background interactors are derived from low affinity binders to small molecules, similarly tagged ‘inactive’ analogs are often used to sort through specific and non-specific interactions. If an inactive analog is as structurally close as possible to the molecule of interest but with lower or no biological activity, in theory one will observe the same non-specific binders in pull downs performed with the inactive analog while the specific targets of the active molecule will not be present. An inactive enantiomer of a small molecule is ideal as was used in the target identification of *S* and *R*-roscovitine (*R*-roscovitine was the active molecule) [30, 49]. The caveat is that an inactive analog could be inactive for a number of reasons, whether because it binds some but not all the necessary targets required to give a particular phenotype, or because it has poor solubility or permeability [47]. Care needs to be taken to avoid ruling out potential targets incorrectly on this basis.

Another control is the use of ‘competitor’ pre-treatments. Analogous to the concept of specific elution described above, competitor controls refer to pre-treatment of samples with excess small unlabeled small molecule prior to treatment with tagged molecule. Upon incubation with the affinity matrix, the excess unlabeled small molecule selectively depletes targets from binding to the matrix by blocking them from interacting with tagged small molecule. In both cases, targets are identified in post-experiment comparative analysis by their absence in competitor and inactive- treated samples.

Additional sources of complexity in pull down results originate from proteins that interact with other proteins bound to the resin and come along for the ride. This can be decreased by higher stringency washes, though there is always the danger of removing relevant lower affinity interactions at the same time. Protein interaction databases also exist to aid deconvolution of proteomic results and

identify potential protein piggybackers such as IntAct and Mint [20]. Finally, using selective elutions as described above when possible aids to enrich specific targets.

1.2d Chemical proteomics

In target identification studies, chemical proteomics is a combination of the classical methods used in affinity chromatography with the power of modern mass spectrometry (MS). In recent years, advances in the MS field have provided great improvement in the capabilities of this technology, particularly in sensitivity, resolution, accuracy and speed. Thus in complex protein mixtures more proteins can be detected and quantified more accurately than ever before. Importantly, in recent years quantitative methods have also been developed using isotopic labeling as well as label-free techniques that enable comparative analysis of different sample treatments [50].

Typically, complex samples from affinity purification experiments are first either (a) fractionated by one or two dimensional gel electrophoresis followed by in-gel digestion and peptide extraction or (b) proteolytically digested into peptides and separated by one or two dimensions of high performance liquid chromatography (HPLC) prior to MS analysis [51]. The latter, analyzing complex protein mixtures using a combination of HPLC and MS, is referred to as 'shotgun proteomics'. A type of shotgun proteomics experiment that uses two orthogonal chromatographic separations (specifically strong cation exchange and reverse phase) is referred to as 'MudPIT' (multidimensional protein identification technology). While gel electrophoresis in the past was the predominant approach, gel-free purification has now become more popular due to its greater proteome coverage and higher resolution [52].

Peptides are generally identified by tandem MS (MS/MS), which involves acquiring the mass spectra of the precursor ion (parent peptide) that is then fragmented into multiple ions whose mass spectra are also recorded. The fragment mass spectra are used to determine the identity of the peptide (and ultimately the protein) using protein database analysis (e.g. Swiss Prot, US National Center for

Biotechnology Information or a search engine such as Mascot or Sequest) [20]. The strength of the identification is dependent on the quality of the mass spectra [5]. Quantitation is possible by recording the intensity of a peptide signal over time and calculating the area under the curve to give the extracted ion current (XIC), which is proportional to the abundance of the peptide. While different peptides within a given sample cannot be compared because of their unique ionization efficiencies and yields during sample preparation, the same peptide can be compared between different samples provided that the experimental conditions of the LC-MS run are kept constant [50].

Label-free quantitative methods

In quantitative label-free proteomic methods, different samples are run consecutively and in an identical manner. Identical peptides in different runs are then clustered according to their retention time on the column and their relative intensity indicates their relative abundance [53]. A known quantity of a foreign protein can be spiked in to determine absolute quantitation. This technique has the advantage that no labeling strategy is used and therefore sample preparation is less involved resulting in reduced costs, smaller sample requirements and greater applicability to a wider range of protein sources [52]. Although MS analysis conditions can be carefully controlled, the accuracy of quantitation is affected more by variability in sample preparation prior to analysis in some experimental setups [50].

Isotopic labeling quantitative methods

A number of different isotopic labeling methods exist that allow for the simultaneous analysis of two or more samples in a quantitative fashion.

Stable isotope labeling by amino acids in cell culture (SILAC) is a method that isotopically labels proteins through metabolic incorporation. Cells are grown in media containing either light or heavy stable isotope-labeled arginine (bearing six ^{13}C atoms) and lysine (bearing six ^{13}C atoms and two

^{15}N atoms) that are integrated into the proteome over several cell doublings and resultant protein turnover. Different sample treatments are prepared and then combined and analyzed by LC-MS/MS in one analysis. The same peptides from different samples will co-elute in pairs, but will be separated in mass spectra by the difference in mass of the labeled amino acid (Figure 1.2A). Thus quantitation of relative abundance of a peptide from a treated sample compared with a control sample can be calculated. Technical considerations include confirming complete labeling of proteins with isotopic amino acids and adapting cell media such that there are no unlabeled sources of arginine and lysine available to the cells (dialyzed serum is used as a result, which can affect growth of certain cell lines) [50].

In general, analyses using the SILAC technique possess an important advantage in that the cells from different treatments are mixed at an early stage and experimental variability from sample preparation involving lysis, fractionation, HPLC separation and MS analysis is identical, which greatly enhances the accuracy [54]. However, in the case of affinity purification experiments this benefit is reduced somewhat as cell lysis and affinity purification are performed separately and samples are combined at the washing and elution stage. Additionally, SILAC can only quantitatively compare two samples at a time. As twice the amount of peptides are being analyzed in one run, additional fractionation is generally performed by gel electrophoresis and different gel slices are analyzed separately by HPLC-MS/MS, which increases instrumental time requirements [55]. SILAC has been used recently to identify targets of a cancer cell-selective drug, piperlongumine, which is believed to inhibit proteins involved in the cellular response to oxidative stress, such as glutathione S-transferase pi 1 (GSTP1) and carbonyl reductase (CBR1) [56].

Isotope-code affinity tags (ICAT) is a technique analogous to SILAC except that protein samples are labeled with heavy or light isotopic tags post-treatment (Figure 1.2B). The ICAT tag is reactive with

cysteines (via iodoacetamide) on proteins. The current version uses nine ^{13}C isotopes with an acid-cleavable linker connected to biotin for avidin purification of labeled peptides after the tagging reaction and digestion are performed. After elution, heavy and light ICAT-tagged peptides are then pooled and analyzed together as described above [54]. The drawbacks of this technique are that accuracy is dependent on the efficiency of the labeling reaction and that an additional purification step of the ICAT-tagged peptides is involved, which adds more experimental variability to the samples. Additionally, only peptides containing cysteine can be quantified, which also reduces the number of peptides in the sample able to be used for protein identification and any protein lacking cysteine cannot be detected [50]. Oda et al. used the ICAT technique in combination with gene expression profiling and 2D DIGE (DIfference Gel Electrophoresis, discussed below) to identify malate dehydrogenase (MDH) as a target of E7070 (Indisulam), an anticancer agent that was in clinical trials at the time [57].

Isobaric tag for relative and absolute quantitation (iTRAQ) is another tagging strategy, similar to ICAT, except that the isobaric tag modification occurs at amino groups of peptides (via NHS ester). Theoretically by labeling amines all peptides should be tagged at their amino termini, which provides an advantage over ICAT. A series of eight isobaric tags (of the same weight) for labeling exist that fragment in tandem MS to give eight different types of reporter ions. Thus, peptides in each labeled sample all have the same mass and are only differentiated upon fragmentation by the ratios of the reporter ion tags of different mass [58]. Importantly, this advances the previous technology by enabling the quantitative analysis of multiple (up to eight) different protein samples simultaneously. Just as above, good resolution of peptides is necessary through further separation by gel electrophoresis to reduce overlap of reporter ions from different peptides. iTRAQ was used to determine the target profile of kinase inhibitors used as clinical drugs such as imatinib (Gleevec) from a subset of the human proteome comprised of kinases and purine-binding proteins. In addition to confirming known targets of imatinib

(ABL family kinases), they also identified the receptor kinase, DDR1, and a quinone oxidoreductase, NQO2, as imatinib binders, which they confirmed with biochemical assays [24].

A nice control with isotopic labeling experiments is generally performed by reversing the ‘heavy’ and ‘light’ tags in the treated and control samples in another replicate of the experiment. Differential ratios in peptide pairs or reporter ions should be inversed if the data from the previous run reflected a real interaction. Ratios that remain the same regardless of heavy or light labeling are likely due to inconsistency in the labeling or the experimental setup.

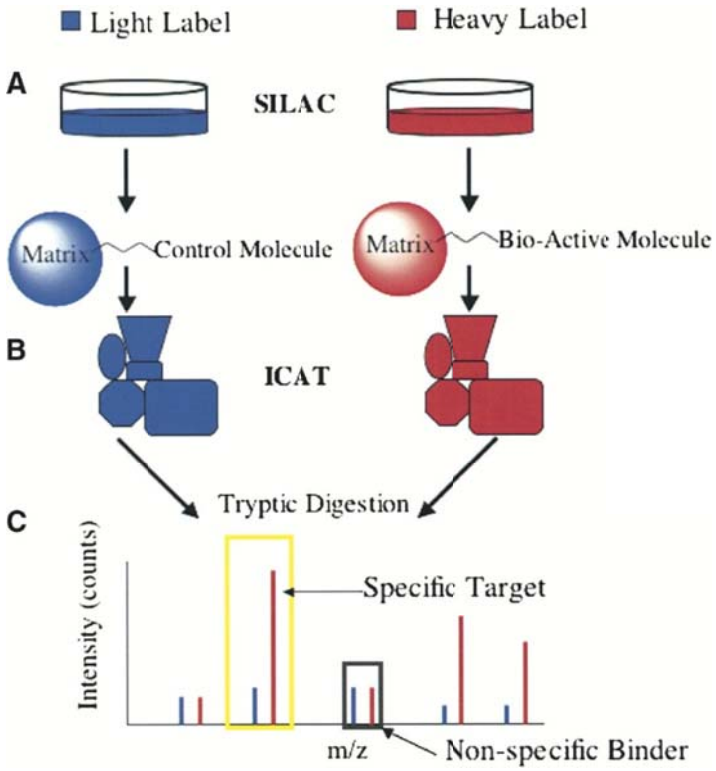


Figure 1.2: Quantitative proteomic techniques with isotopic labeling

(A) SILAC, stable isotope labeling with amino acids in cell culture, incorporates heavy amino acids into the proteome by supplementation of the media. Isotopically labeled cell lysates are then incubated with small molecule and control matrices.

(B) ICAT, istope coded affinity tags, are incorporated on proteins after elution from the matrix via a chemical reaction.

(C) The small molecule and control-treated samples are mixed together and analyzed by mass spectrometry. Peptides from both samples appear in pairs separated by the difference in mass of the isotopic tag. Proteins that are enriched in one sample show greater signal intensity. High ratios of heavy:light peptides indicate specific interactions with the small molecule. In a control experiment, the labeling direction is reversed and thus the ratio should also be inversed if it is a real interaction and not due to experimental artifact.

Figure was reproduced from [30] with permission.

1.2e Drug Affinity Response Target Stability (DARTS)

DARTS is a relatively new method for target identification based on the idea that proteins become more resistant to proteolysis upon ligand binding. The process negatively enriches candidate targets by depleting non-ligand stabilized proteins from the lysate. The major advantages of this approach are in its relative technical simplicity and that no derivitization of the small molecule is required - often a major hurdle in target ID approaches. In proof-of-principle experiments, DARTS was shown to enrich lysates with FKBP12 upon treatment with FK506 or rapamycin, but not upon treatment with wortmannin or vehicle controls. Additionally they identified the eukaryotic initiation factor, eIF4A as a potential target of reseveratrol a molecule popular for its suspected anti-aging effects [59].

Although this method is fairly accessible, some optimization is still necessary. Highly abundant proteins are the easiest to detect by eye in subsequent SDS PAGE analysis and thus lower abundance proteins may not be discerned. One solution is subcellular fractionation, or separation by 2D PAGE or DIGE, and a better alternative is coupling the technique with shotgun proteomics. For shotgun proteomics, dialysis of the samples to remove small digest leftovers (below a 10kDa cutoff) is necessary to avoid convolution with tryptic peptides generated later from stabilized targets. Protease choice has also been observed to affect degradation patterns and the inherent stability or sensitivity to proteolysis of certain proteins can complicate results [29]. This method can also be used for target validation of protein candidates as was demonstrated recently in a screen for inhibitors of ribosomal ternary complex accumulation. Resistance to proteolysis in a DARTS assay was used to show that purified recombinant heme-regulated inhibitor kinase (HRI) was the target of their active compound, a series of N,N'-diarylureas [60].

1.2f Yeast three-hybrid (Y3H)

The yeast three- hybrid system is a yeast transcriptional assay for the detection of small molecule- or RNA-protein interactions *in vivo* and is a derivative of the well-known yeast two-hybrid system (Y2H). It is comprised of three hybrid molecules, two of which are fusion proteins consisting of (a) the DNA binding domain (DBD) of a transcription factor and the receptor protein of an anchor moiety (typically a well-characterized high affinity ligand such as methotrexate or FK506) and (b) the transcriptional activation domain (AD) and a second ligand binding protein. The third hybrid molecule is a chemical heterodimer containing the anchor moiety connected via a linker to a small molecule of interest whose target is unknown [61] (Figure 1.3). Through the availability of genomic cDNA libraries, a library of cDNA-AD clones can be screened in this system to find fusion proteins that interact with the small molecule of interest. The interaction effectively reconstitutes the transcription factor, resulting in expression of a reporter gene (typically) essential for growth of the cell. The target protein can then be easily identified by sequencing of the cDNA-AD clone [62].

The screen possesses many of the same advantages and disadvantages of Y2H systems. First the assay utilizes yeast, a highly manipulatable and cost-effective organism that has already been adapted to high through-put screening formats due to advances in Y2H technology. Second, the cDNA-AD fusions are overexpressed, which facilitates detection of low abundance proteins that may be imperceptible to other techniques. Small molecule-protein interactions are also detected *in vivo* as opposed to *in vitro* allowing certain post-translational modifications to be maintained. This method also comes with caveats equivalent to those in Y2H as proteins must be expressed functionally as fusions that can translocate to the nucleus in order to be detected, and interactions that require complex formation also go unseen [61].

Since the first proof of principle experiments performed by Licitra and Liu, this technology has been used to profile targets of known cyclin-dependent kinase (CDK) inhibitors, roscovitine, Purvalanol B, and indenopyrazole [63, 64]. Recently, a version of the Y3H system incorporating a covalent anchor moiety based on SNAP-tag technology was used to profile dasatinib, Purvalanol B, erlotinib, atorvastatin and furosemide for known and unknown targets. Additionally, they identified sepiapterin reductase (SPR) as a target of sulfasalazine, an anti-inflammatory drug whose target was previously unknown [65].

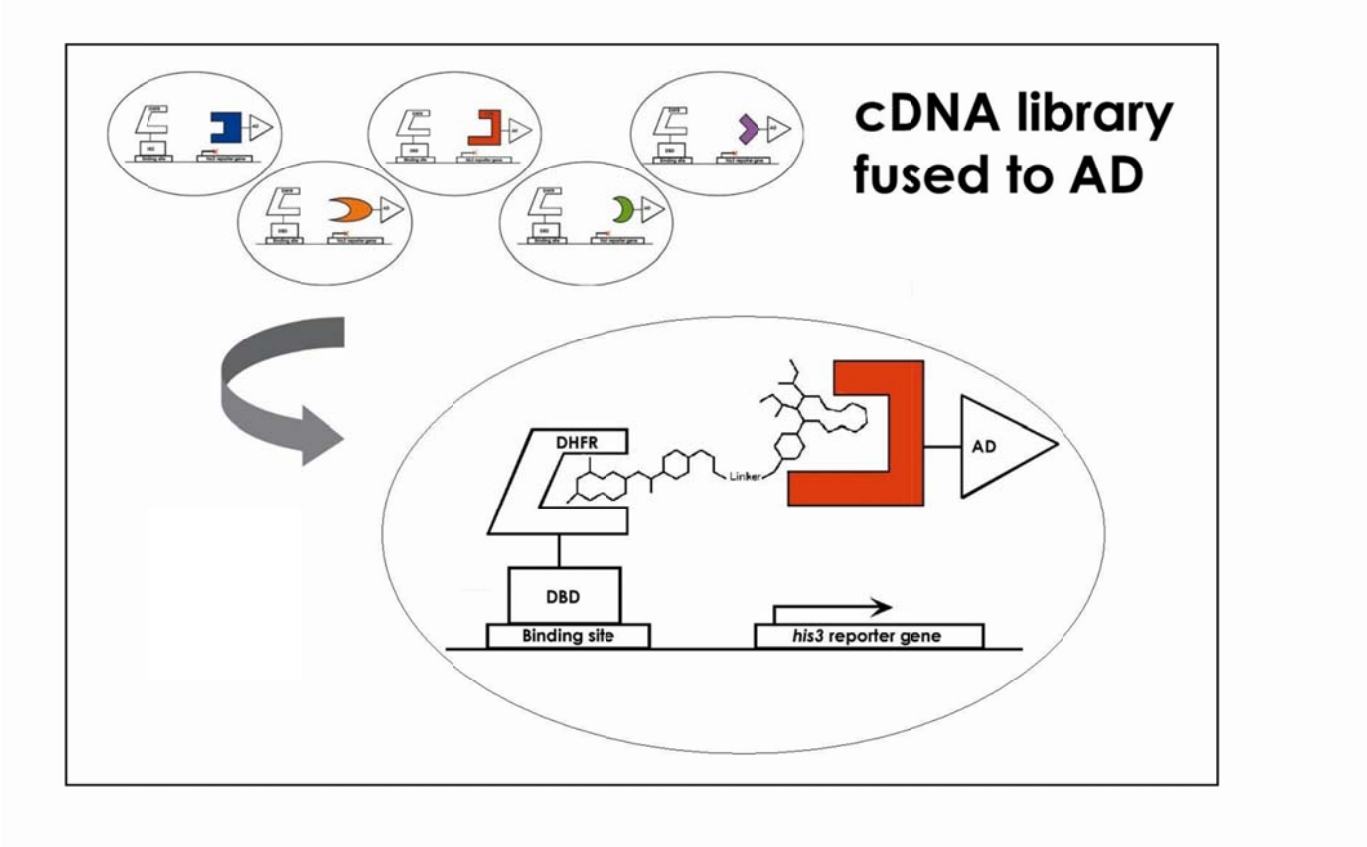


Figure 1.3: The yeast three-hybrid system as a method for target identification

Transcriptional activation of a reporter gene (*his3* shown as an example) is coupled to two small molecule-protein interactions, one between an anchor ligand (e.g. methotrexate) and a fusion protein of its receptor (dihydrofolate reductase, DHFR) and a DNA-binding domain (DBD). The other interaction is between a small molecule with unknown targets and a fusion protein expressed from a library of cDNA-activation domain (AD) clones. Interactions that result in transcriptional activation can be detected by growth or enzyme activity and then identified by sequencing.

1.2g Protein microarrays

Protein-binding assays can be adapted to identify targets of a small molecule though many of these techniques are difficult to perform in a high-throughput fashion [4]. One of these techniques, protein microarrays, is complementary to DNA microarrays in that it involves the immobilization of collections of proteins on a solid arrayed surface. Small molecule-protein interactions with a particular position on the array are detected by labeling of the small molecule with a fluorescent or radioactive tag [66, 67]. Technical complications that have restricted this technology include the difficulty in purification and immobilization of a large number of diverse, correctly folded, functional proteins that are required to truly make this a global assay (although promisingly, this has been done with membrane proteins such as G-protein coupled receptors [68]). Additionally, the proteins lack cofactors or complex partners found in a native environment that may be required for binding of certain small molecules.

A solution to this problem comes in the form of cellular microarrays that are arrays of plasmid cDNA and transfection reagent that are then plated with live mammalian cells (transfected cell microarray [69]) or a cell-free version utilizing mammalian lysate (nucleic acid programmable protein array, [70]) which produce the desired proteins *in vivo* or *in vitro*. Small molecule interactions are also detected as described above or through the use of antibodies to epitope-tagged proteins. A certain level of background in this array derives from endogenous targets expressed by the cells themselves. To date, these arrays have only been used in proof-of-principle experiments in the application of target ID.

1.2h Comparison of direct methods

When choosing a method for target identification experiments, the primary determining factor is whether one possesses the ability to modify the small molecule with a tag or tether, which is not trivial. In order to achieve this, extensive knowledge of the structure activity relationship (SAR) of the molecule is necessary. Typically a functional group on the molecule is required for connection with a

tether, however not all small molecules possess convenient functional groups or in certain cases these groups are essential for binding. Finding a site for modification may require exploration of many positions on the molecule for modification. Additionally, if the small molecule of interest is structurally complex, like many natural products, it may be difficult to synthesize, let alone modify with a tether. No generalized solution for this problem exists for normal small molecules found in compound libraries and thus this issue must be dealt with on a case by case basis. Unfortunately, the majority of direct target ID approaches involve the modification of the molecule in some way (except for DARTS), which emphasizes the importance of this problem. Indirect methods for target identification described in the next section provide an alternative solution for molecules that cannot be modified.

A distinction between various direct target ID methods is the setting in which the small molecule-protein interaction is assessed. The majority of methods probe these interactions *in vitro*. This has the advantage that proteins are in a near native state (provided lysis is done carefully), at normal abundance levels and in the presence of all cofactors and complex partners. Additionally all isoforms of a protein and post-translation modifications are also represented. The disadvantage is that *in vitro* methods generally require large amounts of cellular material that can be difficult and expensive to achieve with mammalian cell lines, although animal tissue or organs can be used as a cheaper alternative. Additionally the lysing process biases against proteins unstable in aqueous buffer, such as membrane proteins. With the exception of covalent small molecules that can be used in cells, the Y3H and cellular microarrays are direct methods that probe small molecule-protein interactions in a cellular setting, although as proteins are overexpressed they are not present at endogenous levels. As mentioned before, this can also facilitate detection of the interaction. Additionally, many indirect methods of target identification are also performed *in vivo*.

1.3 Indirect methods of target identification

1.3a Genetic modulator screens

Genetic screens in model organisms have been used to identify direct and related targets of small molecules. For example, the target of rapamycin in yeast was identified through the discovery of mutations in two alleles of rapamycin-resistant yeast strains encoding phosphoinositol kinase-related kinases, TOR1 and TOR2, of which mTOR (mammalian target of rapamycin) is a homolog [71, 72]. Yeast deletion libraries are also used as a method for identifying genes (potentially targets) that confer resistance or increased sensitivity to small molecules lethal in yeast [73-75]. Multicopy overexpression libraries are used analogously [76]. These screens are based on the concept that increasing or decreasing the abundance of a compound's target will modulate the cell's sensitivity to the compound.

A platform, dubbed as 'HIPHOP', combines the power of multiple genetic modulator assays [77]. Haploinsufficiency profiling (HIP) of lethal compounds is performed with a genome-wide library of heterozygous deletion mutants. Genes that sensitize yeast to a particular compound in this library represent potential targets of that compound. Homozygous profiling (HOP) is also performed with a library of homozygous deletion mutants of non-essential genes. Genes that sensitize yeast to a particular compound in this library represent genetic modifiers of that compound. Additionally a random genomic DNA transformant library is screened to give a multicopy suppressor profile (in the case that a molecule exerts a loss of function on its target, overexpression of the target should suppress the phenotype). Integration of these three assays was found to improve the correct prediction of small molecule targets. A number of known small molecule-protein interactions were confirmed using this platform and potential novel interactions with new bioactive compounds were also identified.

Genetic screens in model organisms possess the advantage that the phenotype of sensitivity or suppression of a strain to a compound is easily linked to the identity of the gene involved. Experiments

are performed *in vivo,* which can more accurately mimic the native environment of the original small molecule-protein interaction. In multicopy suppressor screens however, proteins are not expressed at endogenous levels, although this can facilitate the detection of interactions with low abundance targets. Shortcomings include the limitation of the approach to small organisms and the screens to fundamental processes connected to the fitness of yeast.

1.3b Profiling methods (guilt by association)

A number of profiling methods assimilating different experimental parameters have been used to predict a small molecule's mechanism of action through comparison with an existing database of small molecule profiles (colloquially termed 'guilt by association' [45]). Prominent among these is the use of gene expression profiling, which monitors global change in the transcription level of genes upon compound treatment.

Gene expression profiling

Early experiments using gene expression profiling in yeast treated with small molecules compared the expression signatures of FK506 and cyclosporin A-treated wild-type cells with cells containing a calcineurin-null mutant (the final target of these molecules) [78]. Their data showed that the expression profiles of wild-type yeast treated with the two drugs had high correlation. As the two drugs are known to exert similar downstream effects through calcineurin, these results implied that comparison of gene expression profiles could represent the correlation between two molecule's mechanisms of action. Treatment of calcineurin-null mutant cells with FK506 no longer showed the pattern observed in wild-type cells, implying a method for validation of potential small molecule targets [79]. Interestingly, when they treated the calcineurin-null mutant cells with FK506 at higher doses they observed a gene expression pattern that correlated with a signature from cells containing a mutant form of a transcription factor *GCN4*. Furthermore, treatment of mutant *GCN4* cells with FK506 showed no

further expression changes implying that *GCN4* could be a secondary target of FK506 at high concentration [78]. These initial experiments illustrated the relationships that could be drawn by analysis of gene expression signatures.

The Connectivity Map resource is a profiling system developed in recent years that aims to make "'connections between different small molecules, genes and disease". In their initial publication, Lamb *et al.* created a dataset of the transcriptional signatures of 164 compounds in human cells, primarily the breast epithelial cell line MCF7 [80]. 'Query signatures' can then be generated for a new compound treatment and compared to transcriptional profiles in the existing database for other molecules with signatures of high or low correlation (referred to as 'connectivity'). Using their system, they verified that they could identify HDAC inhibitors from signatures produced independently in other cell lines using those inhibitors. Additionally, they went on to identify heat shock protein 90 (HSP90) as a target of an uncharacterized compound, gedunin, through high connectivity with geldanamycin and its structural analogs, known HSP90 inhibitors, which they experimentally verified.

Cell Viability profiling

The National Cancer Institute's panel of sixty cancer cell lines has become a useful resource for deciphering molecular mechanisms of novel, lethal bioactive compounds. Toxicity profiles of small molecules against this panel were originally designed in the hope of finding drugs with tumor-specific lethality. In contrast, it soon became apparent that similarities in the sensitivity and resistance profiles of different small molecules were due mainly to related mechanisms of action. This realization was quickly made accessible to researchers in the form of the 'COMPARE' algorithm that enabled the generation of a ranked list of small molecule profiles in the database with similar profiles to a novel small molecule. Conversely, identification of molecules with 'COMPARE-negative' profiles indicated unique activity and potential new mechanisms of tumor lethality, also a desirable quality although

unhelpful in terms of mechanistic understanding [81]. One example of COMPARE's ability was the identification of the involvement of mitogen-activated ERK kinase (MEK) pathway inhibition in the toxicity of anthrax lethal factor through its high correlation with a known inhibitor of MEK, PD98059 [82].

Modulatory profiling

Modulatory profiling is a system developed in the Stockwell laboratory to characterize lethal molecules based on their profile of inhibition or sensitization with a panel of chemical modulators of known cell death processes [83]. Compounds with known mechanisms of action, acting through similar lethal mechanisms, are affected similarly by the same chemical modulators and thus have profiles that cluster together. Novel lethal molecules can be profiled in the system in order to classify their cell death mechanism as similar to a group of compounds with known mechanisms of action or they can be identified as possessing a unique cell death mechanism.

1.3c Additional mechanistic studies

Metabolite profiling

Although not yet advanced to the stage of a profiling database for small molecules, metabolite profiling can provide clues to a compound's mechanism of action through alterations in metabolite levels. It shares similarities to gene expression profiling in that global changes in metabolite levels in the cell are measured. Metabolites include all endogenous molecules under 1kDa within the cell and exclude biopolymers. The major application of metabolite profiling is in the area of biomarkers for drug efficacy and toxicity [84].

Difference Gel Electrophoresis (DIGE)

DIGE is an equivalent method to detect global changes in protein levels upon compound treatment. It is a gel-based technique that can be used to quantitatively compare protein levels in

different samples, although it is not as rigorous as the proteomic methods described above. A control and a treatment sample are labeled with different fluorescent dyes (e.g. Cy3 and Cy5), separated by 2D PAGE and scanned for fluorescence that is then analyzed computationally. Identical proteins from the different samples overlap with each other and, if they exist at the same level, a 'neutral' colour spot (yellow) is observed, whereas if a protein is present at higher or lower levels in either sample a red or green spot is observed depending on the dye-labeling. Differential protein spots are picked and sequenced by MS.

1.4 Getting around the problem

As target identification poses a major challenge for forward chemical genetics, some have pre-emptively tackled the hurdle of the future need to derivatize small molecule hits with a tag or tether by using libraries for screening that already have the tag incorporated into them. In a forward chemical genetic screen for small molecules exhibiting brain or eye developmental phenotypes in zebrafish embryos, a library of over 1500 tagged triazine molecules were tested. Each molecule had a triethylene glycol linker with a free amino group integrated into the structure. Hits that were identified in the screen were readily tethered to beads and used in affinity chromatography experiments to identify targets [85].

Although this approach is attractive, it comes with the caveat that the presence of tethers in the molecule may reduce the binding affinity of hits that would otherwise be active. Additionally, there is the possibility that the tether itself may be required for the biological activity of the molecule. To address these issues, smaller tags such as the azido/alkyne tags used in ABPP can be incorporated into the molecule. As the size of the tag relative to the molecule is decreased, the chances that it interferes with or is essential for binding are reduced. Evans *et al*. demonstrated this approach in a small library of

molecules containing both a spiroepoxide electrophilic group to induce covalent modification of the target and an alkyne tag to be used for subsequent affinity purification of hit molecules [86].

An alternative approach is to screen collections of small molecules with known biological activity (annotated libraries). Candidate hits that are identified most likely exert their phenotypes through their known mechanism of action. These libraries also possess the advantage that they are often enriched with active compounds in phenotypic screens [87].

1.5 Validation of targets

Following identification of candidate protein targets of a small molecule, potential targets are typically validated using a number of methods. The functional relevance of the potential target is assessed by modulating levels of the target in cell culture through RNAi and overexpression experiments. For molecules that induce a loss-of-function in their targets, RNAi knockdown will normally sensitize cells to the activity of the molecule while overexpression of the target will suppress its effect and *vice versa* for molecules that act through gain-of-function mechanisms. If the target is not an essential gene it can also be possible to test the compound's effect in mouse embryonic fibroblasts where the gene has been knocked out. Additionally co-localization of the small molecule and target in cells or lysates may be demonstrated through immunohistochemistry or by western blot using a tagged version of the molecule.

Ultimately it is necessary to establish direct binding of the small molecule with the target in isolation. Typically this is done *in vitro* using recombinant or purified proteins with an assay to determine the binding affinity of the molecule such as surface plasmon resonance (SPR), isothermal titration calorimetry (ITC) or fluorescence anisotropy. Additionally if the target is an enzyme, activity assays can be performed. Further validation demonstrates structural knowledge of the binding site of the small molecule in the protein. This can be achieved by examining the small molecule-protein

complex by X-ray crystallography or nuclear magnetic resonance (NMR). Binding-defective mutants can be created using site directed mutagenesis to abolish the small molecule-protein interaction and these can be expressed in cell culture to establish that no phenotype is observed in the presence of the defective mutant alone.

1.6 Applying and developing methods for target identification

In this thesis, a range of direct and indirect methods to define the targets and mechanism of a potential anti-cancer agent, RSL3, are explored. Additionally, optimization of an existing method for target identification, the yeast three-hybrid system is attempted. In Chapter 2, the development of an affinity reagent for direct target identification of RSL3 using affinity chromatography methods is discussed. In Chapter 3, development and characterization of a covalent version of the yeast three-hybrid system is described. In Chapter 4, a variety of mechanistic data gathered on the mechanism of action of RSL3 is summarized, including the identification of potential targets of RSL3 and their validation. Finally, in Chapter 5 future directions and perspectives are discussed.

1.7 References

1. Stockwell, B.R., *Chemical genetics: ligand-based discovery of gene function.* Nat Rev Genet, 2000. **1**(2): p. 116-25.
2. Schwartz, R.S., *Paul Ehrlich's magic bullets.* N Engl J Med, 2004. **350**(11): p. 1079-80.
3. Strebhardt, K. and A. Ullrich, *Paul Ehrlich's magic bullet concept: 100 years of progress.* Nat Rev Cancer, 2008. **8**(6): p. 473-80.
4. Stockwell, B.R., *Exploring biology with small organic molecules.* Nature, 2004. **432**(7019): p. 846-54.
5. Leslie, B.J. and P.J. Hergenrother, *Identification of the cellular targets of bioactive small organic molecules using affinity reagents.* Chem Soc Rev, 2008. **37**(7): p. 1347-60.
6. Bodley, A., L.F. Liu, M. Israel, R. Seshadri, Y. Koseki, F.C. Giuliani, S. Kirschenbaum, R. Silber, and M. Potmesil, *DNA topoisomerase II-mediated interaction of doxorubicin and daunorubicin congeners with DNA.* Cancer Res, 1989. **49**(21): p. 5969-78.
7. Weisenberg, R.C., G.G. Borisy, and E.W. Taylor, *The colchicine-binding protein of mammalian brain and its relation to microtubules.* Biochemistry, 1968. **7**(12): p. 4466-79.
8. Taunton, J., C.A. Hassig, and S.L. Schreiber, *A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p.* Science, 1996. **272**(5260): p. 408-11.
9. Hung, D.T., T.F. Jamison, and S.L. Schreiber, *Understanding and controlling the cell cycle with natural products.* Chem Biol, 1996. **3**(8): p. 623-39.
10. Yoshida, M., S. Horinouchi, and T. Beppu, *Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function.* Bioessays, 1995. **17**(5): p. 423-30.
11. Liu, J., J.D. Farmer, Jr., W.S. Lane, J. Friedman, I. Weissman, and S.L. Schreiber, *Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes.* Cell, 1991. **66**(4): p. 807-15.
12. Schreiber, S.L., *Chemistry and biology of the immunophilins and their immunosuppressive ligands.* Science, 1991. **251**(4991): p. 283-7.
13. Yano, H., S. Nakanishi, K. Kimura, N. Hanai, Y. Saitoh, Y. Fukui, Y. Nonomura, and Y. Matsuda, *Inhibition of histamine secretion by wortmannin through the blockade of phosphatidylinositol 3-kinase in RBL-2H3 cells.* J Biol Chem, 1993. **268**(34): p. 25846-56.
14. Sabatini, D.M., H. Erdjument-Bromage, M. Lui, P. Tempst, and S.H. Snyder, *RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs.* Cell, 1994. **78**(1): p. 35-43.
15. Fenteany, G., R.F. Standaert, W.S. Lane, S. Choi, E.J. Corey, and S.L. Schreiber, *Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin.* Science, 1995. **268**(5211): p. 726-31.
16. Kerridge, D., *The effect of actidione and other antifungal agents on nucleic acid and protein synthesis in Saccharomyces carlsbergensis.* J Gen Microbiol, 1958. **19**(3): p. 497-506.
17. Schreiber, S.L., *Chemical genetics resulting from a passion for synthetic organic chemistry.* Bioorg Med Chem, 1998. **6**(8): p. 1127-52.
18. Stockwell, B.R., *The Quest for the Cure: The Science and Stories Behind the Next Generation of Medicines.* 2011, New York: Columbia University Press. 284.
19. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nat Rev Drug Discov, 2006. **5**(12): p. 993-6.
20. Rix, U. and G. Superti-Furga, *Target profiling of small molecules by chemical proteomics.* Nat Chem Biol, 2009. **5**(9): p. 616-24.

21. Mayer, T.U. and A. Marx, *Five molecules we would take to a remote island.* Chem Biol. **17**(6): p. 556-60.
22. Druker, B.J., S. Tamura, E. Buchdunger, S. Ohno, G.M. Segal, S. Fanning, J. Zimmermann, and N.B. Lydon, *Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells.* Nat Med, 1996. **2**(5): p. 561-6.
23. Heinrich, M.C., D.J. Griffith, B.J. Druker, C.L. Wait, K.A. Ott, and A.J. Zigler, *Inhibition of c-kit receptor tyrosine kinase activity by STI 571, a selective tyrosine kinase inhibitor.* Blood, 2000. **96**(3): p. 925-32.
24. Bantscheff, M., D. Eberhard, Y. Abraham, S. Bastuck, M. Boesche, S. Hobson, T. Mathieson, J. Perrin, M. Raida, C. Rau, V. Reader, G. Sweetman, A. Bauer, T. Bouwmeester, C. Hopf, U. Kruse, G. Neubauer, N. Ramsden, J. Rick, B. Kuster, and G. Drewes, *Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors.* Nat Biotechnol, 2007. **25**(9): p. 1035-44.
25. Hopkins, A.L., *Network pharmacology: the next paradigm in drug discovery.* Nat Chem Biol, 2008. **4**(11): p. 682-90.
26. Knight, Z.A. and K.M. Shokat, *Chemical genetics: where genetics and pharmacology meet.* Cell, 2007. **128**(3): p. 425-30.
27. Bishop, A.C., J.A. Ubersax, D.T. Petsch, D.P. Matheos, N.S. Gray, J. Blethrow, E. Shimizu, J.Z. Tsien, P.G. Schultz, M.D. Rose, J.L. Wood, D.O. Morgan, and K.M. Shokat, *A chemical switch for inhibitor-sensitive alleles of any protein kinase.* Nature, 2000. **407**(6802): p. 395-401.
28. Harding, M.W., A. Galat, D.E. Uehling, and S.L. Schreiber, *A receptor for the immunosuppressant FK506 is a cis-trans peptidyl-prolyl isomerase.* Nature, 1989. **341**(6244): p. 758-60.
29. Lomenick, B., R.W. Olsen, and J. Huang, *Identification of direct protein targets of small molecules.* ACS Chem Biol, 2011. **6**(1): p. 34-46.
30. Burdine, L. and T. Kodadek, *Target identification in chemical genetics: the (often) missing link.* Chem Biol, 2004. **11**(5): p. 593-7.
31. Gilbert, W. and B. Muller-Hill, *Isolation of the lac repressor.* Proc Natl Acad Sci U S A, 1966. **56**(6): p. 1891-8.
32. Meunier, J.C., R. Sealock, R. Olsen, and J.P. Changeux, *Purification and properties of the cholinergic receptor protein from Electrophorus electricus electric tissue.* Eur J Biochem, 1974. **45**(2): p. 371-94.
33. Pert, C.B. and S.H. Snyder, *Opiate receptor: demonstration in nervous tissue.* Science, 1973. **179**(77): p. 1011-4.
34. Borisy, G.G. and E.W. Taylor, *The mechanism of action of colchicine. Colchicine binding to sea urchin eggs and the mitotic apparatus.* J Cell Biol, 1967. **34**(2): p. 535-48.
35. Lerman, L.S., *A Biochemically Specific Method for Enzyme Isolation.* Proc Natl Acad Sci U S A, 1953. **39**(4): p. 232-6.
36. Sin, N., L. Meng, M.Q. Wang, J.J. Wen, W.G. Bornmann, and C.M. Crews, *The anti-angiogenic agent fumagillin covalently binds and inhibits the methionine aminopeptidase, MetAP-2.* Proc Natl Acad Sci U S A, 1997. **94**(12): p. 6099-103.
37. Kwok, B.H., B. Koh, M.I. Ndubuisi, M. Elofsson, and C.M. Crews, *The anti-inflammatory natural product parthenolide from the medicinal herb Feverfew directly binds to and inhibits IkappaB kinase.* Chem Biol, 2001. **8**(8): p. 759-66.
38. Sato, S., Y. Kwon, S. Kamisuki, N. Srivastava, Q. Mao, Y. Kawazoe, and M. Uesugi, *Polyproline-rod approach to isolating protein targets of bioactive small molecules: isolation of a new target of indomethacin.* Journal of the American Chemical Society, 2007. **129**(4): p. 873-80.

39. Rostovtsev, V.V., L.G. Green, V.V. Fokin, and K.B. Sharpless, *A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes.* Angew Chem Int Ed Engl, 2002. **41**(14): p. 2596-9.
40. Kolb, H.C. and K.B. Sharpless, *The growing impact of click chemistry on drug discovery.* Drug Discov Today, 2003. **8**(24): p. 1128-37.
41. Speers, A.E. and B.F. Cravatt, *Profiling enzyme activities in vivo using click chemistry methods.* Chem Biol, 2004. **11**(4): p. 535-46.
42. Cohen, M.S., H. Hadjivassiliou, and J. Taunton, *A clickable inhibitor reveals context-dependent autoactivation of p90 RSK.* Nat Chem Biol, 2007. **3**(3): p. 156-60.
43. Sorensen, E.J., G.C. Adam, C.D. Vanderwal, and B.F. Cravatt, *(-)-FR182877 is a potent and selective inhibitor of carboxylesterase-1.* Angewandte Chemie-International Edition, 2003. **42**(44): p. 5480-5484.
44. MacKinnon, A.L., J.L. Garrison, R.S. Hegde, and J. Taunton, *Photo-leucine incorporation reveals the target of a cyclodepsipeptide inhibitor of cotranslational translocation.* Journal of the American Chemical Society, 2007. **129**(47): p. 14560-+.
45. Salisbury, C.M. and B.F. Cravatt, *Optimization of activity-based probes for proteomic profiling of histone deacetylase complexes.* Journal of the American Chemical Society, 2008. **130**(7): p. 2184-94.
46. Weerapana, E., A.E. Speers, and B.F. Cravatt, *Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in proteomes.* Nat Protoc, 2007. **2**(6): p. 1414-25.
47. Sato, S., A. Murata, T. Shirakawa, and M. Uesugi, *Biochemical target isolation for novices: affinity-based strategies.* Chem Biol, 2010. **17**(6): p. 616-23.
48. Walsh, D.P. and Y.T. Chang, *Chemical genetics.* Chem Rev, 2006. **106**(6): p. 2476-530.
49. Bach, S., M. Knockaert, J. Reinhardt, O. Lozach, S. Schmitt, B. Baratte, M. Koken, S.P. Coburn, L. Tang, T. Jiang, D.C. Liang, H. Galons, J.F. Dierick, L.A. Pinna, F. Meggio, F. Totzke, C. Schachtele, A.S. Lerman, A. Carnero, Y. Wan, N. Gray, and L. Meijer, *Roscovitine targets, protein kinases and pyridoxal kinase.* J Biol Chem, 2005. **280**(35): p. 31208-19.
50. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.
51. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, *Proteomics by mass spectrometry: approaches, advances, and applications.* Annu Rev Biomed Eng, 2009. **11**: p. 49-79.
52. Patel, V.J., K. Thalassinos, S.E. Slade, J.B. Connolly, A. Crombie, J.C. Murrell, and J.H. Scrivens, *A comparison of labeling and label-free mass spectrometry-based proteomics approaches.* J Proteome Res, 2009. **8**(7): p. 3752-9.
53. Silva, J.C., R. Denny, C. Dorschel, M.V. Gorenstein, G.Z. Li, K. Richardson, D. Wall, and S.J. Geromanos, *Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale.* Mol Cell Proteomics, 2006. **5**(4): p. 589-607.
54. Kline, K.G. and M.R. Sussman, *Protein quantitation using isotope-assisted mass spectrometry.* Annu Rev Biophys, 2010. **39**: p. 291-308.
55. Ong, S.E., M. Schenone, A.A. Margolin, X. Li, K. Do, M.K. Doud, D.R. Mani, L. Kuai, X. Wang, J.L. Wood, N.J. Tolliday, A.N. Koehler, L.A. Marcaurelle, T.R. Golub, R.J. Gould, S.L. Schreiber, and S.A. Carr, *Identifying the proteins to which small-molecule probes and drugs bind in cells.* Proc Natl Acad Sci U S A, 2009. **106**(12): p. 4617-22.
56. Raj, L., T. Ide, A.U. Gurkar, M. Foley, M. Schenone, X. Li, N.J. Tolliday, T.R. Golub, S.A. Carr, A.F. Shamji, A.M. Stern, A. Mandinova, S.L. Schreiber, and S.W. Lee, *Selective killing of cancer cells by a small molecule targeting the stress response to ROS.* Nature, 2011. **475**(7355): p. 231-4.

57. Oda, Y., T. Owa, T. Sato, B. Boucher, S. Daniels, H. Yamanaka, Y. Shinohara, A. Yokoi, J. Kuromitsu, and T. Nagasu, *Quantitative chemical proteomics for identifying candidate drug targets.* Anal Chem, 2003. **75**(9): p. 2159-65.
58. Ross, P.L., Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, and D.J. Pappin, *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
59. Lomenick, B., R. Hao, N. Jonai, R.M. Chin, M. Aghajan, S. Warburton, J. Wang, R.P. Wu, F. Gomez, J.A. Loo, J.A. Wohlschlegel, T.M. Vondriska, J. Pelletier, H.R. Herschman, J. Clardy, C.F. Clarke, and J. Huang, *Target identification using drug affinity responsive target stability (DARTS).* Proc Natl Acad Sci U S A, 2009. **106**(51): p. 21984-9.
60. Chen, T., D. Ozel, Y. Qiao, F. Harbinski, L. Chen, S. Denoyelle, X. He, N. Zvereva, J.G. Supko, M. Chorev, J.A. Halperin, and B.H. Aktas, *Chemical genetics identify eIF2alpha kinase heme-regulated inhibitor as an anticancer target.* Nat Chem Biol, 2011.
61. Kley, N., *Chemical dimerizers and three-hybrid systems: scanning the proteome for targets of organic small molecules.* Chem Biol, 2004. **11**(5): p. 599-608.
62. Lefurgy, S. and V. Cornish, *Finding Cinderella after the ball: a three-hybrid approach to drug target identification.* Chem Biol, 2004. **11**(2): p. 151-3.
63. Becker, F., K. Murthi, C. Smith, J. Come, N. Costa-Roldan, C. Kaufmann, U. Hanke, C. Degenhart, S. Baumann, W. Wallner, A. Huber, S. Dedier, S. Dill, D. Kinsman, M. Hediger, N. Bockovich, S. Meier-Ewert, A.F. Kluge, and N. Kley, *A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors.* Chem Biol, 2004. **11**(2): p. 211-23.
64. Licitra, E.J. and J.O. Liu, *A three-hybrid system for detecting small ligand-protein receptor interactions.* Proc Natl Acad Sci U S A, 1996. **93**(23): p. 12817-21.
65. Chidley, C., H. Haruki, M.G. Pedersen, E. Muller, and K. Johnsson, *A yeast-based screen reveals that sulfasalazine inhibits tetrahydrobiopterin biosynthesis.* Nat Chem Biol, 2011. **7**(6): p. 375-83.
66. MacBeath, G. and S.L. Schreiber, *Printing proteins as microarrays for high-throughput function determination.* Science, 2000. **289**(5485): p. 1760-3.
67. MacBeath, G., *Protein microarrays and proteomics.* Nat Genet, 2002. **32 Suppl**: p. 526-32.
68. Fang, Y., A.G. Frutos, and J. Lahiri, *Membrane protein microarrays.* Journal of the American Chemical Society, 2002. **124**(11): p. 2394-5.
69. Ziauddin, J. and D.M. Sabatini, *Microarrays of cells expressing defined cDNAs.* Nature, 2001. **411**(6833): p. 107-10.
70. Ramachandran, N., E. Hainsworth, B. Bhullar, S. Eisenstein, B. Rosen, A.Y. Lau, J.C. Walter, and J. LaBaer, *Self-assembling protein microarrays.* Science, 2004. **305**(5680): p. 86-90.
71. Cafferkey, R., P.R. Young, M.M. McLaughlin, D.J. Bergsma, Y. Koltin, G.M. Sathe, L. Faucette, W.K. Eng, R.K. Johnson, and G.P. Livi, *Dominant missense mutations in a novel yeast protein related to mammalian phosphatidylinositol 3-kinase and VPS34 abrogate rapamycin cytotoxicity.* Mol Cell Biol, 1993. **13**(10): p. 6012-23.
72. Kunz, J., R. Henriquez, U. Schneider, M. Deuter-Reinhard, N.R. Movva, and M.N. Hall, *Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression.* Cell, 1993. **73**(3): p. 585-96.
73. Zheng, X.S., T.F. Chan, and H.H. Zhou, *Genetic and genomic approaches to identify and study the targets of bioactive small molecules.* Chem Biol, 2004. **11**(5): p. 609-18.
74. Lum, P.Y., C.D. Armour, S.B. Stepaniants, G. Cavet, M.K. Wolf, J.S. Butler, J.C. Hinshaw, P. Garnier, G.D. Prestwich, A. Leonardson, P. Garrett-Engele, C.M. Rush, M. Bard, G. Schimmack, J.W. Phillips, C.J. Roberts, and D.D. Shoemaker, *Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes.* Cell, 2004. **116**(1): p. 121-37.

75. Parsons, A.B., A. Lopez, I.E. Givoni, D.E. Williams, C.A. Gray, J. Porter, G. Chua, R. Sopko, R.L. Brost, C.H. Ho, J. Wang, T. Ketela, C. Brenner, J.A. Brill, G.E. Fernandez, T.C. Lorenz, G.S. Payne, S. Ishihara, Y. Ohya, B. Andrews, T.R. Hughes, B.J. Frey, T.R. Graham, R.J. Andersen, and C. Boone, *Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast.* Cell, 2006. **126**(3): p. 611-25.
76. Luesch, H., T.Y. Wu, P. Ren, N.S. Gray, P.G. Schultz, and F. Supek, *A genome-wide overexpression screen in yeast for small-molecule target identification.* Chem Biol, 2005. **12**(1): p. 55-63.
77. Hoon, S., A.M. Smith, I.M. Wallace, S. Suresh, M. Miranda, E. Fung, M. Proctor, K.M. Shokat, C. Zhang, R.W. Davis, G. Giaever, R.P. StOnge, and C. Nislow, *An integrated platform of genomic assays reveals small-molecule bioactivities.* Nat Chem Biol, 2008. **4**(8): p. 498-506.
78. Marton, M.J., J.L. DeRisi, H.A. Bennett, V.R. Iyer, M.R. Meyer, C.J. Roberts, R. Stoughton, J. Burchard, D. Slade, H. Dai, D.E. Bassett, Jr., L.H. Hartwell, P.O. Brown, and S.H. Friend, *Drug target validation and identification of secondary drug target effects using DNA microarrays.* Nat Med, 1998. **4**(11): p. 1293-301.
79. Stockwell, B.R., *Frontiers in chemical genetics.* Trends Biotechnol, 2000. **18**(11): p. 449-55.
80. Lamb, J., E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, J. Lerner, J.P. Brunet, A. Subramanian, K.N. Ross, M. Reich, H. Hieronymus, G. Wei, S.A. Armstrong, S.J. Haggarty, P.A. Clemons, R. Wei, S.A. Carr, E.S. Lander, and T.R. Golub, *The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.* Science, 2006. **313**(5795): p. 1929-35.
81. Shoemaker, R.H., *The NCI60 human tumour cell line anticancer drug screen.* Nat Rev Cancer, 2006. **6**(10): p. 813-23.
82. Vande Woude, G.F., N.S. Duesbery, C.P. Webb, S.H. Leppla, V.M. Gordon, K.R. Klimpel, T.D. Copeland, N.G. Ahn, M.K. Oskarsson, K. Fukasawa, and K.D. Paull, *Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor.* Science, 1998. **280**(5364): p. 734-737.
83. Wolpaw, A.J., Shimada K., Skouta R., Welsch M.E., Akavia U., Pe'er, D., Shaik F., Bulisinski J.C., Stockwell, B.R., *Modulatory profiling identifies mechanisms of small molecule-induced death.* PNAS, 2011.
84. Keun, H.C. and T.J. Athersuch, *Application of metabonomics in drug development.* Pharmacogenomics, 2007. **8**(7): p. 731-41.
85. Khersonsky, S.M., D.W. Jung, T.W. Kang, D.P. Walsh, H.S. Moon, H. Jo, E.M. Jacobson, V. Shetty, T.A. Neubert, and Y.T. Chang, *Facilitated forward chemical genetics using a tagged triazine library and zebrafish embryo screening.* J Am Chem Soc, 2003. **125**(39): p. 11804-5.
86. Evans, M.J., A. Saghatelian, E.J. Sorensen, and B.F. Cravatt, *Target discovery in small-molecule cell-based screens by in situ proteome reactivity profiling.* Nat Biotechnol, 2005. **23**(10): p. 1303-7.
87. Root, D.E., S.P. Flaherty, B.P. Kelley, and B.R. Stockwell, *Biological mechanism profiling using an annotated compound library.* Chem Biol, 2003. **10**(9): p. 881-92.

Chapter 2: Application of target identification methods to a potential *RAS*-selective lethal anti-cancer agent, RSL3

2.1 Introduction

2.1a The RAS protein

The *RAS* genes, like other key oncogenes, were first identified by the presence of their tumorigenic forms in the Harvey and Kirsten rat sarcoma viruses in the 1960s [1-3]. Studies of these viruses revealed that their tumor-forming properties were a result of viral '*RAS*' (RAt Sarcoma) genes they encoded. In the late seventies and early eighties, these genes were subsequently detected to also be present in their normal form in rat and human cells as well as in their oncogenic forms in human cancers [4-8]. Four isoforms of the RAS protein exist encoded by three genes *HRAS*, *KRAS* (following the original Harvey and Kirsten v-*RAS* genes) and a third form known as *NRAS* [3].

The RAS protein is a GTPase which acts as a mediator between extracellular signals and extensive intracellular networks controlling cell growth, survival, differentiation and angiogenic processes. In its GTP-bound form it binds and activates downstream effector proteins such as RAF and PI3K, members of different pathways involved in processes under the control of RAS. Upon hydrolysis of GTP, the RAS protein's enzymatic activity effectively turns itself off by conversion into the GDP-bound state, thereby inactivating its ability to bind effectors. The oncogenic form of the protein present in cancer is the result of a single point-mutation commonly occurring at positions 12, 13 and 61 in the amino acid sequence which results in loss of GTPase activity. Loss of RAS enzyme activity results in it being in a permanently active state and thus it acts as a gain-of-function oncogene producing signals for pro-growth, survival and cell-invasiveness, important for the tumorigenic phenotype [9].

The *RAS* genes are mutated in 20-25% of human cancer and are particularly frequent in pancreatic, colon and lung carcinomas, three of the most prevalent killers [3, 10]. Thus, unsurprisingly, much effort has been made at targeting aberrant *RAS* signaling pathways. The main approaches employed thus far have aimed at blocking RAS localization to the membrane, utilized antisense oligonucleotides against *RAS* mRNA or targeted downstream kinases such as RAF, but issues such as lack of efficacy, specificity or difficulties in drug delivery as well as unexpected compensatory effects have impeded their success in clinical trials [11-13].

2.1b Synthetic lethal screening as an approach to find genotype-specific drugs

One approach to targeting the RAS pathway is through the use of synthetic lethal screening. The concept of synthetic lethality originated in classical genetics experiments in smaller organisms where certain combinations of two genetic mutations resulted in lethality, while either of the individual mutations alone was still viable [14]. Similarly, synthetic lethal screens are performed in chemical genetics using a small molecule as a substitute for one of the mutations. When this method is applied in conjunction with an oncogenic mutation, small molecules that are lethal only in an oncogene-specific genetic background and ideally innocuous in a wild-type background can be identified. Thus synthetic lethal screening offers an approach to discover small molecules whose toxicity specifically targets tumor cells, a crucial goal of cancer research.

Screening to identify synthetic lethal interactions with oncogenic RAS has revealed both genes and small molecules with mutant *RAS*-selective lethal properties. These screens typically utilize isogenic cell lines where one cell line possesses an oncogenic *RAS* mutation and the other has the wild-type allele. In 2009, three independent studies used RNAi-based synthetic lethal screens against mutant *KRAS* and identified lethal interactions with three kinases as their top hits: a mitotic polo-like kinase 1 (PLK1), serine/threonine kinase 33 (STK33) and the non-canonical IκB kinase (TBK1) [15-17].

Additionally, chemical synthetic lethal screening in oncogenic-*RAS* harboring cells has identified a sulfinyl cytidine derivative (SC-D), an indole-3-carbinol analog (onrasin-1) and more recently, lanperisone, a drug undergoing clinical testing as a muscle relaxant, as *RAS*-selective lethal compounds [18-20]. All compounds exhibited roughly four to six-fold selectivity in their lethality between the respective isogenic cell lines used in the studies with potencies ranging between sub to low micromolar.

2.1c Discovery of RSL3

In the Stockwell laboratory we have identified a number of mutant *RAS*-selective lethal (RSL) compounds in screens using a series of isogenic engineered cell lines created in the laboratory of Dr. Robert Weinberg. These cells were derived from primary human fibroblasts (BJ cells) and converted into tumorigenic cells by the introduction of vectors expressing human telomerase (hTERT), the large and small T antigen of the simian vacuolating virus (SV40 LT and ST) and an oncogenic allele of *HRAS* (*HRAS*^{G12V}) [21]. Three cell lines were created with a successively increasing number of oncogenic vectors. We named these cells lines BJ-TERT, BJ-TERT/LT/ST, and BJ-TERT/LT/ST/RAS^{V12} (hereafter referred to as **BJeH**, **BJeHLT** and **BJeLR**). Additionally a fourth cell line containing mutant *HRAS* was created BJ-TERT/p53DD/CDK4/cyclinD1/ST/RAS^{V12} (hereafter referred to as **DRD**) that substitutes the SV40 large T antigen for a dominant negative mutant of p53 and a constitutively active CDK4/cyclin D, which inactivates the RB protein [22]. The first two cell lines in this series BJeH and BJeHLT are immortalized cell lines that don't possess tumorigenic properties. The cell lines with mutant *HRAS*, BJeLR and DRD, both have the ability to form tumors in nude mice. BJeLR cells also rapidly proliferate while DRD cells have a longer doubling time. Identifying compounds that are lethal in both of these two cell lines with mutant *RAS* provides a control to eliminate compounds whose toxicity is targeted at rapidly dividing cells or is dependent on the presence of the large T antigen.

Using the four BJ cell series as a screening platform, the RSL compounds erastin, RSL3 and RSL5 were identified [23-25]. Erastin, RSL3 and RSL5 are not structurally related compounds, although their mechanisms of action possess similarities in that they all induce a non-apoptotic, oxidative, iron-dependent form of death, which will be discussed further in the following chapter. Erastin, a quinazolinone compound, is the most characterized of the three small molecules and its target, the voltage dependent anion channels (VDAC), have been identified [24]. RSL5 is also believed to act through the VDAC proteins and its lethal effects are rescued by RNAi directed at VDAC3, however the targets of RSL3 were still unknown [25].

RSL3 is a natural product derivative, with a tetrahydro-β-carboline backbone and exhibits the highest potency of all the RSL molecules with a low nanomolar EC50 in BJeLR and DRD cells. Further screening against the panel of sixty cancer cell lines by the National Cancer Institute (NCI60) identified RSL3 as displaying a novel pattern of killing when compared to its existing database of tested compounds. In order to understand its potentially novel mechanism of mutant-*RAS*-selective lethality, which could be a source of new therapeutic targets, we aimed to isolate the biologically-relevant protein interactions of RSL3.

2.1d Candidate for target identification

As a candidate for target identification, RSL3 had a number of desirable features. First, it possessed an interesting phenotype of mutant-*RAS*-selective lethality. Second it also had properties that increased chances of success in target identification experiments, particularly high potency as well as a putative covalent interaction with its target. The latter idea stemmed mainly from the reactive chloromethyl ketone functionality that preliminary structure activity relationship (SAR) studies indicated was essential for its lethality (Figure 2.1). Testing was performed using analogs of RSL3 that were purchased from commercial libraries as well as synthesized by an outside custom synthesis company.

These studies also provided information to direct possible sites for tether attachment to the RSL3 scaffold. Thus, in order to identify the biological targets of RSL3, we chose to pursue the classical affinity chromatography approach utilizing an affinity tag. In this chapter, the synthesis of an affinity reagent for target purification from the RSL3 scaffold is described. Crucial discoveries regarding the specific stereoselectivity of the active RSL3 molecule were made in this process that ultimately enabled the synthesis of a number of RSL3-probes with both alkyne and fluorescein tags for affinity purification. These probe molecules were subsequently used in pull down experiments for target identification using shotgun proteomic profiling.

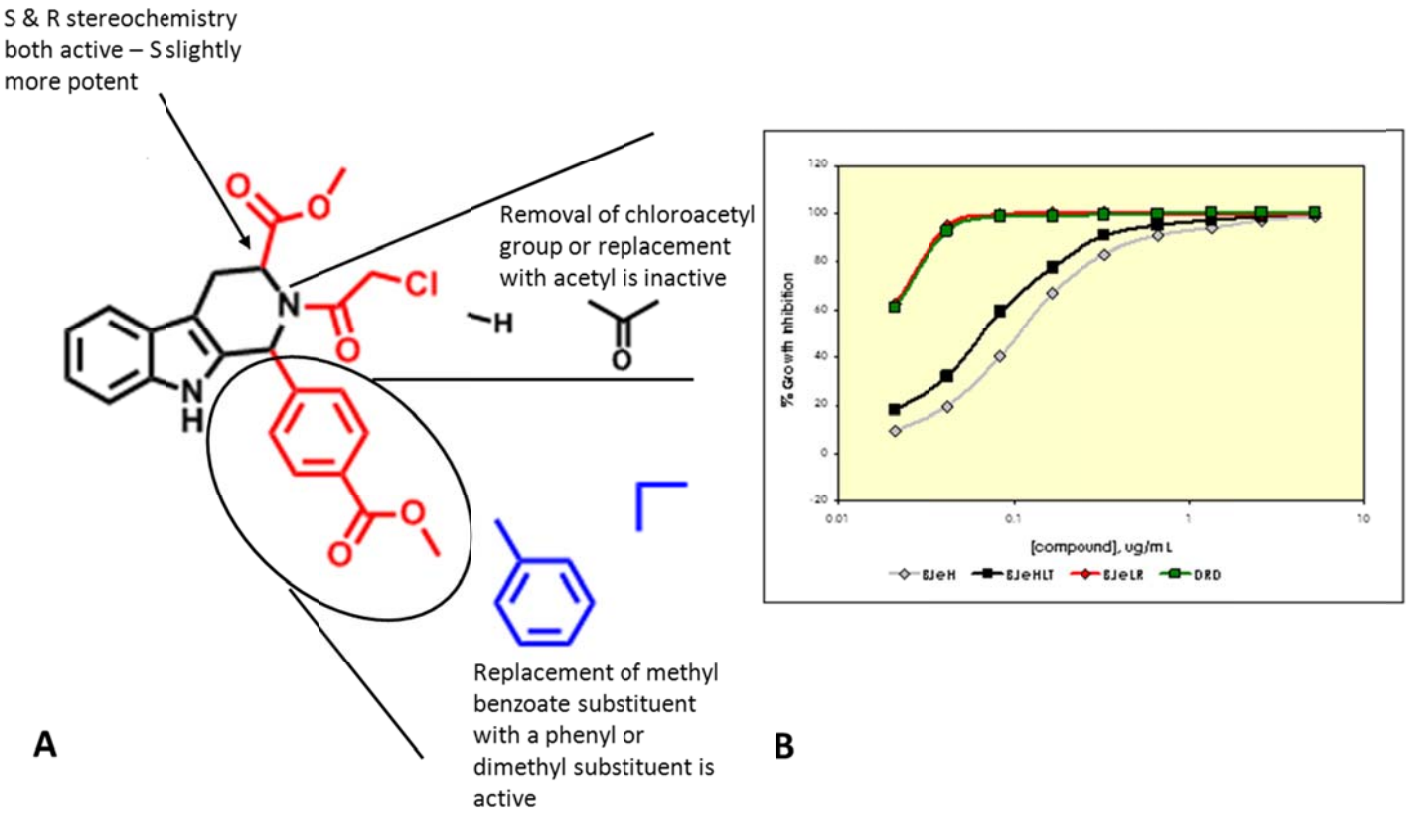


Figure 2.1: Preliminary structure activity relationship map of RSL3 and dose response curve

(A) Preliminary SAR of RSL3 revealed that the chloromethyl ketone functionality (chloroacetyl) group was essential for the activity of the molecule. Replacement with either an acetyl or complete removal to leave the 2° amine abolished all activity. Reduction of the methyl benzoate to a phenyl or dimethyl substituent retained activity with reduced potency. Both isomers with S & R stereochemistry at the 3 position were active; the S isomer had slightly greater potency. SAR data was obtained through testing analogs of RSL3 that were purchased from commercial libraries as well as synthesized by an outside custom synthesis company.

(B) Dose response curve of RSL3 in the four BJ cell series: BJeH (BJ-TERT), BJeHLT (BJ-TERT/LT/ST), BJeLR (BJ-TERT/LT/ST/RAS^{V12}) and DRD (BJ-TERT/p53DD/CDK4/cyclinD1/ST/RAS^{V12}). RSL3 displayed eight-fold selectivity between cell lines with oncogenic *RAS* and isogenic cell lines with the wild type allele.

2.2 Results

2.2a Making an affinity reagent from RSL3

Previous SAR data indicated a degree of structural flexibility in the phenyl ring located at the 1 position in the RSL3 scaffold, thus initial efforts were focused on modification at this location. A structural route based on a commercially available nitro phenyl analog was conceived to derivatize the molecule with a methylene linker connected to the ring with a biotin tag – commonly used in affinity purification. Upon synthesis and testing of this molecule (1), it became apparent that all lethal activity was lost (Figure 2.2). At the same time, an analog with a smaller modification on the phenyl ring comprising of a propargyl ester in place of the methyl ester was synthesized by an outside company (2). This modification incorporated a terminal alkyne in the molecule. The advantage of using alkyne groups is in their compactness, which makes them less likely to interfere sterically with the binding of the molecule. However, even this very similar analog, while possessing lethality albeit at a much lower potency, had no selectivity between cell lines with mutant *RAS* and cells with the wild-type allele. These data suggested that extension beyond the methyl ester of RSL3 would not be tolerated in terms of *RAS*-selective lethal activity.

Based on this study, we proposed to synthesize an alkyne version of RSL3 where the incorporation of the tag did not extend beyond the spatial limits of the parent molecule. Deletion of the methyl ester from the phenyl group and substitution of the phenyl group with a dimethyl still retained fairly potent and selective activity, so we felt that removing the ester should not impact negatively (Figure 2.1). Alkyne-RSL3 (3) was synthesized with an ethyne group connected to the phenyl at the para position (Figure 2.2). Confusingly, this again proved to be lethal but not selective. Our hypotheses about the possibility of modification at this site in the molecule appeared to be false, thus we chose to explore completely different positions in the compound.

5-propargyloxy RSL3 (4) was synthesized starting from 5-hydroxy tryptophan such that a propargyl group would extend from the indole in the beta-carboline scaffold (Figure 2.2). 3-ethyl ester RSL3 (5) was also synthesized easily to test the possibility of expansion beyond the methyl ester at the 3 position. Disappointingly, neither of these two analogs showed selective lethality in the four BJ series.

This stage proved to be a turning point in the project as an important positive control was performed. The original structure of RSL3 was resynthesized. Although resynthesis had been performed previously by an outside source, it had never been synthesized in our laboratory. The synthesis of RSL3 revealed that although the structure appeared to be chemically identical (by NMR and MS analysis) to that produced by the company, it did not exhibit lethality in a selective manner and behaved similarly to the previous affinity analogs synthesized.

These findings led us to question our pre-existing SAR related to RSL3. The fact that the two versions of the molecule on the surface appeared to be indistinguishable by standard structural characterization brought forward the idea that it would be worthwhile to synthesize all four stereoisomers. RSL3 contains two chiral centers, either in the 'R' or 'S' configuration, therefore four combinations of configurations can exist. Each stereoisomer was purified and characterized individually by ^{1}H and ^{13}C NMR, and the absolute stereo configuration was assigned using Nuclear Overhauser Effect spectroscopy (NOESY) (see Methods, NMR).

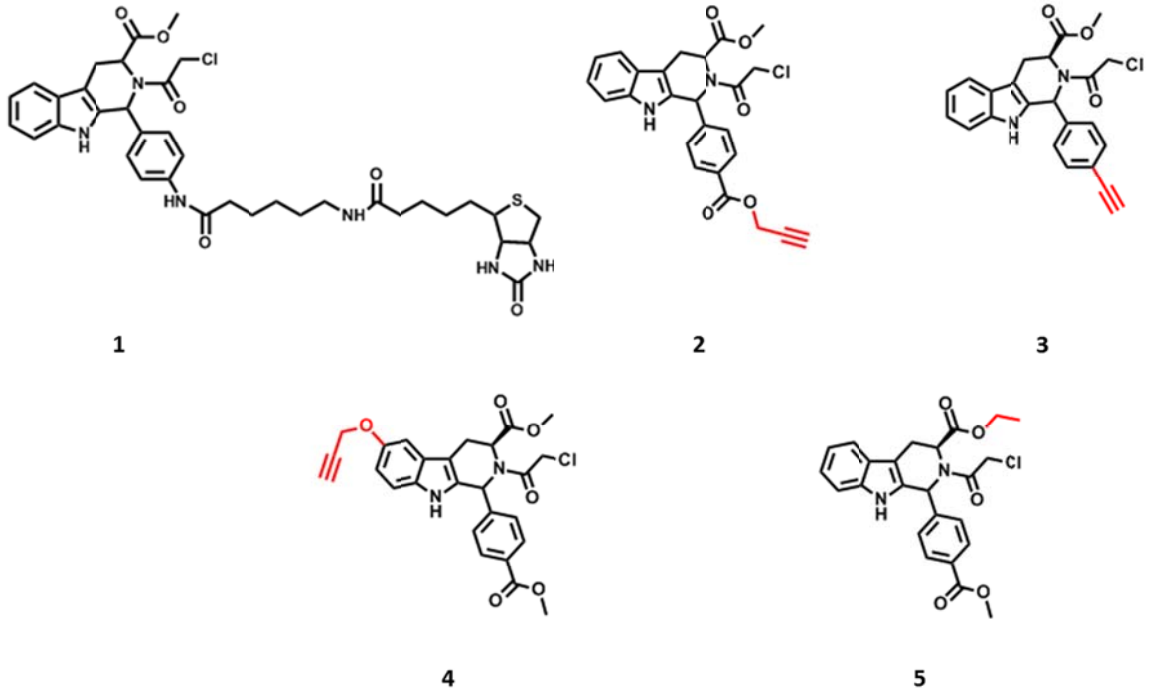


Figure 2.2: Structures of initial affinity tag and other analogs of RSL3

Biotin-RSL3 (1), a propargyl ester analog synthesized by an outside source (2), alkyne-RSL3 (3), 6-propargyloxy RSL3 (4) and 3-ethyl ester RSL3 (5) were either inactive or not *RAS*-selective lethal when tested in our RSL assay. Biotin-RSL3 (1) was synthesized from a diastereomeric mixture of the methyl 1-(4-nitrophenyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-3-carboxylate and the stereochemistry of (2) was indicated to be in the 'S' configuration at the 3 position, but there are doubts over the accuracy of the assignment provided by the outside company. Analogs (3), (4) and (5) were synthesized with 'S' configuration at the 3 position, starting from the L enantiomer of tryptophan.

Testing all four stereoisomers proved to be very illuminating. Only one of the four stereoisomers displayed potent and selective lethality in our RSL assay (Figure 2.3). The active stereoisomer in fact possessed 'R' configuration at the 3 position, a position that previous SAR based on the commercially synthesized analogs indicated was not a determinant in its activity. The three other stereoisomers showed the weaker but non-selective lethality characteristic of many of the affinity probes synthesized previously. This provided an explanation for the lack of selectivity in all previously modified versions of RSL3, as all of them had been synthesized with an 'S' configuration at the 3 position (using the (L) enantiomer of tryptophan) based on the data suggesting that this version was slightly more potent (Figure 2.1). The stereospecificity of RSL3's structure is intriguing as it implies a unique binding conformational requirement for the molecule's activity. Further, the non-selective stereoisomers provide excellent negative controls in affinity chromatography and other experiments.

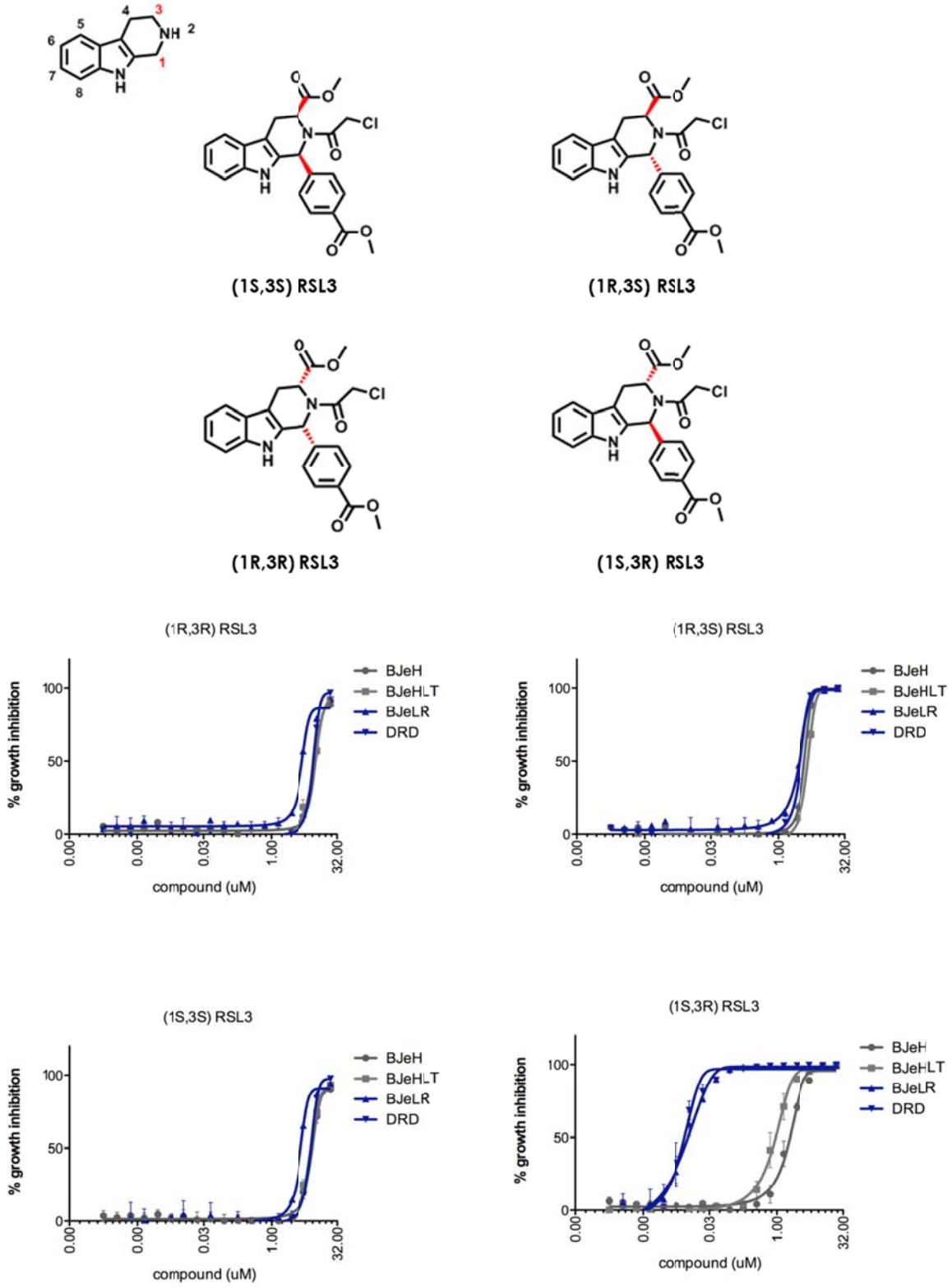


Figure 2.3: One stereoisomer of RSL3 uniquely possesses *RAS*-selective lethal activity

Only one of the four stereoisomers of RSL3 with 1S,3R configuration exhibits *RAS*-selective lethal activity while the remaining three isomers show no selectivity and lethality at a much higher dose. This result

explained why all previous analogs had been inactive or non-selective as they were synthesized with an S configuration at the 3 position. Data represents the mean of three replicates ±SD.

The alkyne-RSL3 (**3**) described above was rapidly remade starting with the (R) enantiomer of tryptophan instead to give 1S3R (**6**) and 1R3R (**7**) alkyne-RSL3 probe 1 (Figure 2.4). Satisfyingly, the activity was as predicted. (1S,3R) alkyne-RSL3 was nearly as potent and selective as (1S,3R) RSL3, and (1R,3R) alkyne-RSL3 was only non-selectively lethal at higher doses. To confirm that the new probe was acting through a similar mechanism to 1S3R RSL3, the effect of some key chemical modulators of RSL3's lethality were also tested with the probe. As described in the next chapter, antioxidants (butylated hydroxyanisole), inhibition of MEK kinase (U0126) and necroptotic cell death (Necrostatin 1) are all known to rescue cells from RSL3's toxic effects. This rescue was also observed with the alkyne probe and thus these data supported the idea that their mechanisms of action were sufficiently similar to warrant use of 1S3R alkyne-RSL3 as an affinity reagent in the identification of biological targets of RSL3.

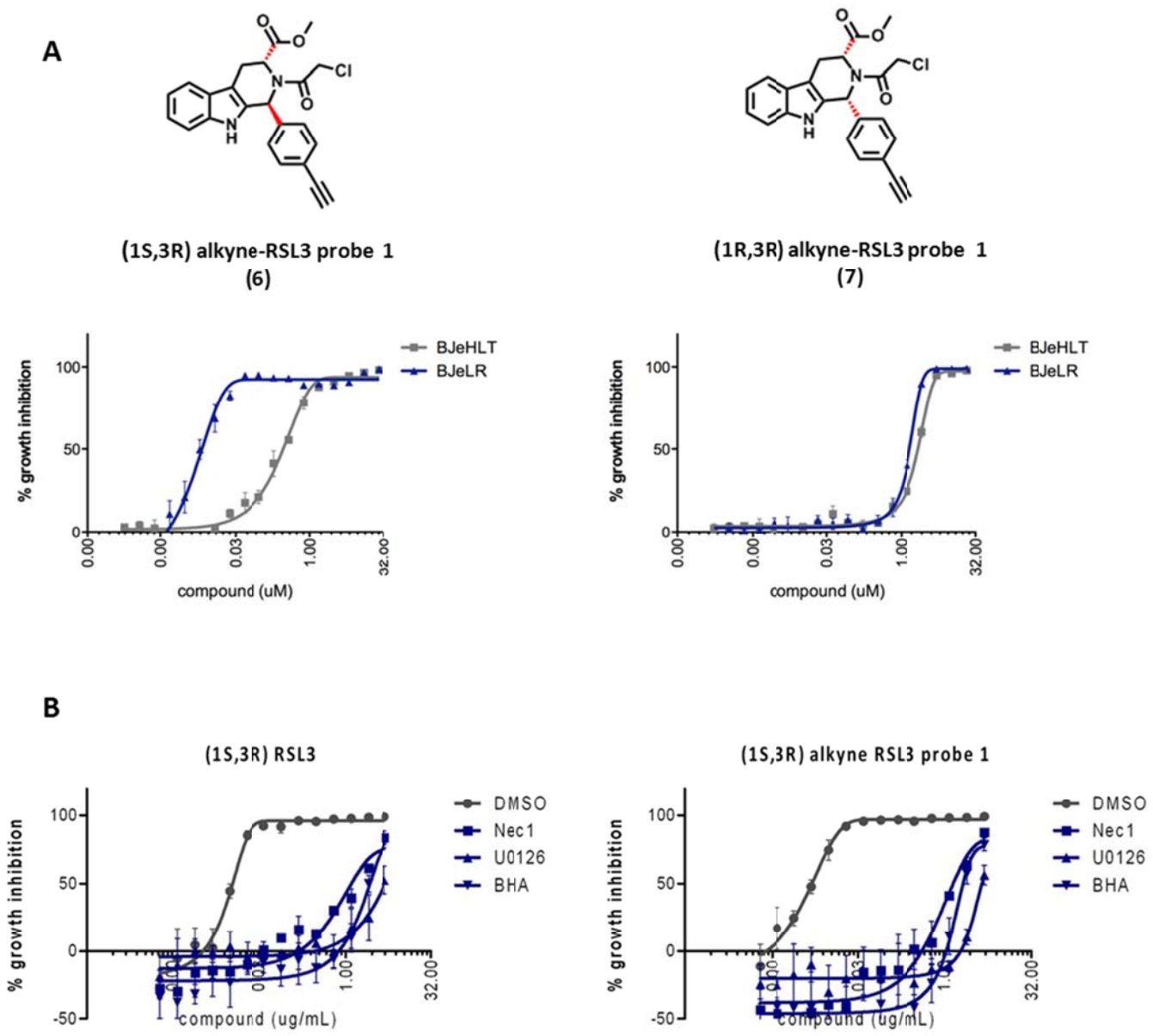


Figure 2.4: Structures and dose response of Alkyne-RSL3 probes 1 and inhibitor profile

(A) Structures and dose response of (1S,3R) and (1R,3R) alkyne-RSL3 probe 1 show that RSL3 scaffolds with (1S,3R) stereoconfiguration retained mutant *RAS*-selective lethality while stereoisomers in the other configurations did not. The (1R,3R) alkyne-RSL3 probe 1 can serve as a control affinity tag to distinguish non-specific interactions with the RSL3 scaffold from those that distinctly bind to the scaffold that exhibits selective lethality.

(B) Chemical inhibitors (BHA, butylated hydroxyanisole; Nec1, Necrostatin 1; U0126, MEK kinase inhibitor) known to suppress RSL3-induced death were used to characterize the probe's mechanism of action and confirm that its lethality is suppressed in a similar manner, implying comparable target profiles. Data represents the mean of three replicates ±SD.

2.2b Affinity pull downs with alkyne-RSL3

Click chemistry pull downs have been used in the Stockwell laboratory to identify the target of another tetrahydro-β-carboline discovered in a separate project involving mutant Huntington protein [26]. An alkyne-modified analog of the hit compound was first bound covalently to the target protein in cells via a similar chloromethylketone functionality. After cell lysis, the bio-orthogonal copper (I)-mediated Huisgen's azido alkyne cycloaddition (commonly referred to as 'click chemistry') was used to install a fluorescein azide tag onto the probe-labeled proteins [27, 28]. These were then purified with an anti-fluorescein antibody and analyzed by SDS PAGE. A band that was strongly labeled with the tagged probe was isolated and identified by MS.

Following this work, a similar strategy was employed with RSL3. Due to the presumed covalent interaction between RSL3 and its target, compound treatment was performed in cells to capture proteins in their native state and environment. The stereo selective feature of RSL3's activity provided an ideal 'inactive' analog control in the form of the 1R3R diastereomer. Excess RSL3 was also used to prepare competitor-treated samples (as described in Chapter 1.2c).

The alkyne tag was chosen in favor of the biotin tag because of the smaller modification it entailed resulting in good retention of potency in the probe molecule. The use of click chemistry as a method to detect covalent small molecule-protein interactions in cells has been largely pioneered by Cravatt and colleagues in their work on activity-based proteomic profiling (ABPP) [29-31]. In their system, a rhodamine azide tag is 'clicked' onto compounds following target binding and can be visualized by in gel fluorescence. For affinity purification they use a trifunctional tag that incorporates both a rhodamine and a biotin for use as an affinity tag.

Previous target ID work in our laboratory has found greater non-specific binding using NeutrAvidin beads than compared with antibodies against an exogenous tag such as fluorescein,

presumably due to the presence of endogenous biotinylated protein [32]. By using fluorescein as an affinity tag we hoped to reduce non-specific binding events between the antibody and the lysate. It also enabled selective elution of fluorescein-labeled proteins using excess carboxyfluorescein (this is only possible with biotin when using monomeric versions of avidin).

Samples of three different treatments were prepared, with either 1S3R alkyne RSL3 (referred to as 'probe') or 1R3R alkyne RSL3 ('inactive') as a well as a 'competitor' treated sample that was pretreated with untagged RSL3 in tenfold excess prior to treatment with 1S3R alkyne RSL3. After lysis, click chemistry reactions were performed on the treated lysates and established protocols were performed to remove excess click chemistry reagents from the lysate prior to incubation with blocked control beads to preclear the lysate and anti-fluorescein antibody coupled beads [33] (Figure 2.5A). Selective elution was performed with a solution of 5(6) carboxyfluorescein as utilized before in our laboratory and eluted proteins were precipitated via chloroform-methanol precipitation and resolubilized in detergent for MS analysis [32]. Instead of using gel analysis to identify candidate targets according to banding patterns, we chose to utilize the greater capabilities of label-free shotgun proteomics where multiple proteins from complex lysates can be identified quantitatively from individual samples [34, 35]. This was done in collaboration with Dr. Lewis Brown, Director of the Comparative Proteomics Center. By comparing the proteomes isolated in different pull down treatments, proteins that were selectively enriched in

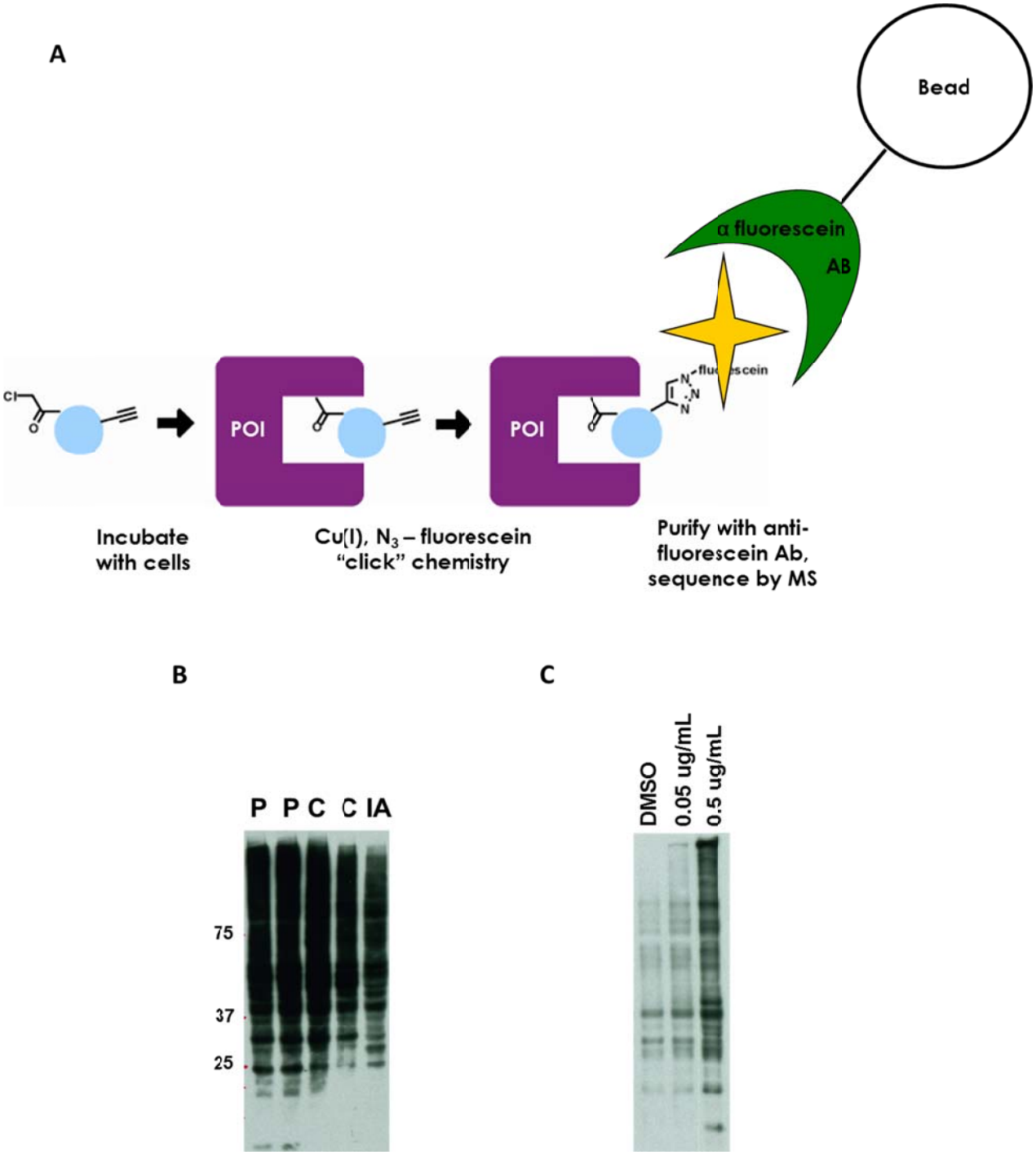


Figure 2.5: Affinity pull down strategy with alkyne-RSL3 probe1

(A) Overall schematic of pull down strategy. RSL3 (blue) is incubated with cells and binds to proteins of interest (POI). After cell lysis, copper (I)-catalyzed click chemistry reactions are performed with fluorescein azide that are then 'pulled down' using anti-fluorescein antibody coupled beads and then eluted with excess free carboxy-fluorescein.

(B) *In vitro* labeling of alkyne-RSL3 bound proteins with fluorescein azide shown by a western blot using anti-fluorescein antibodies. Click chemistry reactions were performed on lysates prepared from cells treated with the (1S,3R) alkyne-RSL3 probe 1 at 0.5ug/mL (**P**), pretreatment with 10X competitor (**C**) prior to treatment with probe at 0.5ug/mL, and 'inactive' (1R,3R) alkyne-RSL3 probe 1 (**IA**) at 0.5ug/mL.

(C) An example of background labeling observed during the click chemistry reaction. This was particularly noticeable when comparing click chemistry reactions performed on DMSO-treated cells with probe-treated at low concentrations (0.05ug/mL) vs. high concentrations (0.5ug/mL).

the probe samples over the competitor or inactive treated samples could be identified more effectively, particularly as lower abundance proteins can be obscured in gels. Additionally, the complexity and number of binding events most likely due to the reactivity of RSL3, meant that visualizing a differential between samples by eye would have been extremely difficult. Gels run with samples prepared at high concentrations of alkyne-RSL3 treatment had such a high intensity of fluorescein labeling that they appeared as black smears by Western blot and individual bands were hard to detect (Figure 2.5B).

During the optimization of the pull down protocol, background labeling was observed independent of the presence of the alkyne-RSL3 probe. Cells treated with DMSO only as a control and then subjected to click chemistry reaction conditions also showed labeling that was visible on SDS denaturing gels with either rhodamine azide (by in gel fluorescence) or fluorescein azide (by Western blot) despite not being treated with an alkyne-tagged molecule. The phenomenon of background labeling when performing click chemistry on cell lysates was reported in initial publications and had already posed a problem for its use in ABPP (ABPP described in Chapter 1.2b) [36]. However, in their original method the small molecule was tagged with an azide and their fluorescent tag possessed an alkyne. Their experiments concluded that the alkyne in addition with copper was the source of the background and they greatly reduced this issue by reversing the alkyne and azide moieties [30]. Since the fluorescent tag is present in excess in the reaction over the small molecule, having the azide on the tag instead was the likely reason.

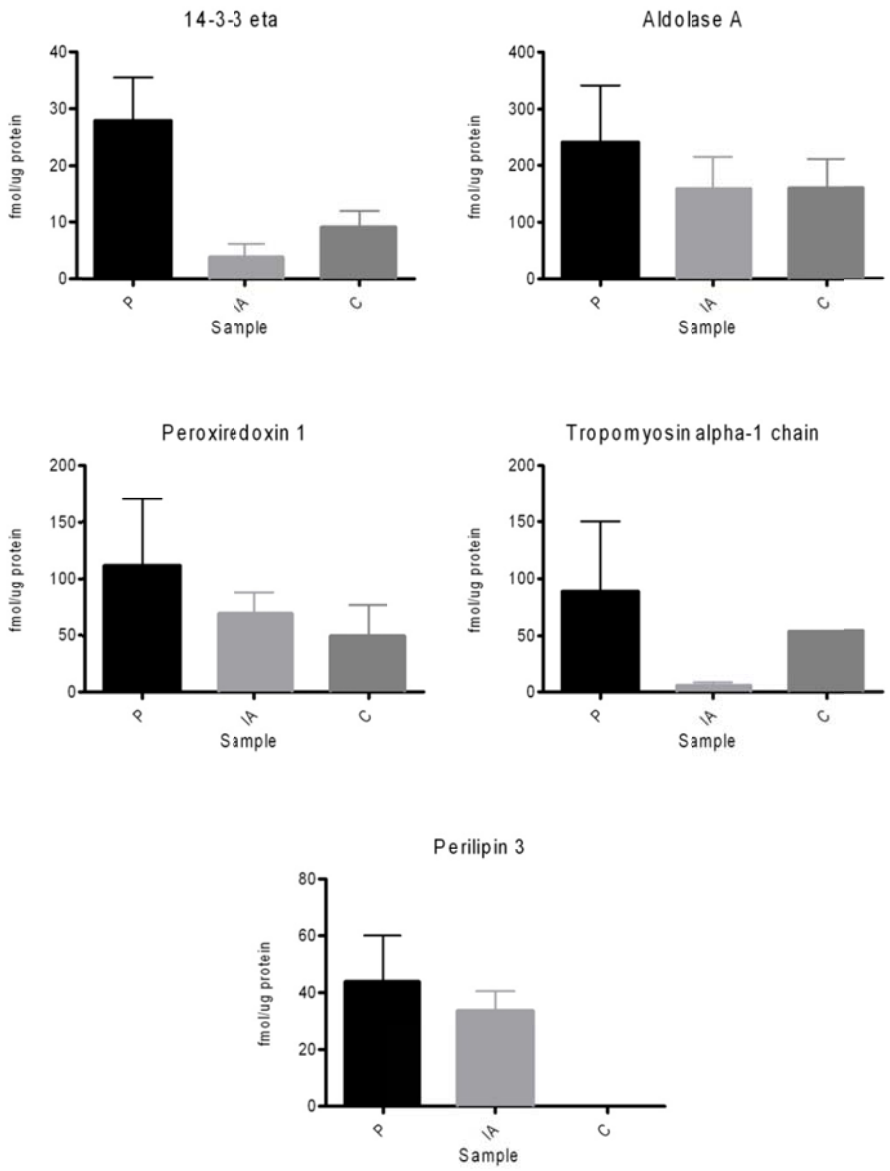
In our case, despite having the azide on the fluorescein tag, significant background labeling was still observed. At low concentrations of alkyne-RSL3 probe, labeling with fluorescein azide was barely distinguishable from lysates prepared from DMSO treated cells (Figure 2.5C). Further testing indicated that the labeling was dependent on the presence of the fluorescein azide tag and not on the other click chemistry reagents. If alkyne probes were treated at higher concentrations, more labeling could be

seen in proportion to this increase presumably due to actual labeling of probe-bound proteins. In ABPP, typical probe concentrations range between 5-20uM, however we were attempting to treat cells with low to mid nanomolar concentrations of alkyne-RSL3 due to its high potency. Additionally, the rhodamine tag used was generally in two- to ten-fold excess of the alkyne-small molecule. Thus, to ameliorate the situation, we increased the probe concentration to low micromolar level and decreased the fluorescein azide tag so the ratio of probe:tag was roughly 1:10. There is also a possibility that these differences could be due to the use of fluorescein instead of rhodamine as the fluorescent reporter.

Shotgun proteomic profiling of the samples and comparative analysis of the proteins eluted under different treatment conditions found a few proteins to be enriched in probe treated samples over inactive and competitor treated samples. However, the enrichment observed was not particularly high and the significance of the differences was somewhat poor due to high variability in the samples. Of the selected candidate targets, 14-3-3 eta, an adapter protein involved in signal transduction, showed the most significant enrichment in the probe samples as well as the greatest significance in the values for amount of protein detected (Figure 2.6 A, B). Peroxidredoxin 1 also stood out because of its role as an antioxidant enzyme that could be connected with RSL3's characteristic oxidative cell death phenotype.

Pulldown samples were prepared again to try to detect the presence of these proteins in elutions by Western blot. A band possibly corresponding to a degraded form of peroxidredoxin 1 was detected with the correct pattern of enrichment while 14-3-3 eta could not be detected (data not shown). Nonetheless, viral shRNA knockdown experiments were attempted with 14-3-3 eta to look at effects of decreasing protein levels on RSL3's potency (Figure 2.6 C). Unfortunately, no effect was observed and thus we decided not to pursue further investigation of these candidates.

A



B

Protein name	Ratio (P/IA)	Ratio (P/C)	PLGS score
ALDOA_HUMAN **Fructose-bisphosphate aldolase A**	1.5	1.5	1512
PRDX1_HUMAN **Peroxiredoxin-1**	1.6	2.2	1209
TPM1_HUMAN **Tropomyosin alpha-1 chain**	15.7	1.6	506
1433F_HUMAN **14-3-3 protein eta**	7.3	3.1	453
PLIN3_HUMAN **Perilipin-3 (Mannose-6-phosphate receptor-binding protein 1)**	1.3	N/A	218

C

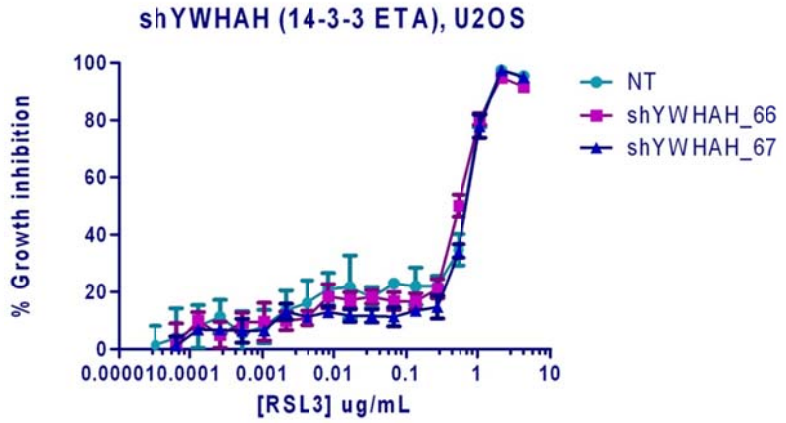


Figure 2.6: Results and follow-up of alkyne-RSL3 affinity pulldown

(A) Protein levels detected (in fmol/ug of total protein) of top hits in probe (P) treated cells vs. inactive (IA) and competitor (C) treated cells identified from alkyne-RSL3 affinity pulldown. Data represents the mean of three biological and three technical replicates ±SD.

(B) Fold enrichment of top hits in probe (P) treated cells vs. inactive (IA) and competitor (C) treated cells. Both tropomyosin alpha and 14-3-3 eta showed high enrichment in probe vs. inactive treated cells, however, data for 14-3-3 eta had greater significance.

(C) Viral shRNA knockdown of 14-3-3 eta (shYWHAH clone 66 and 67) in U20S cells (sensitive to RSL3) showed no sensitization or suppression effects on the compound's potency compared with cells infected with non-targeting (NT) shRNA controls. This experiment was performed by Dr. Wan Seok Yang. Data represents the mean of three replicates ±SD.

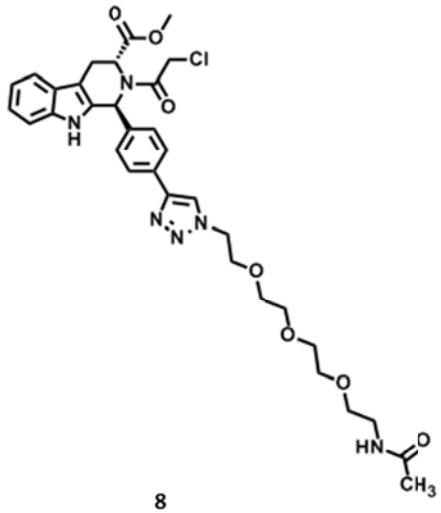
2.2c Affinity pull downs with fluorescein-RSL3

Given the doubts regarding the level of background labeling occurring in the experimental protocol and the poor level of enrichment found by proteomic analysis we decided to rethink our pull down strategy in order to improve confidence in the results obtained. Reducing the background labeling could require time-intensive optimization of the click chemistry reaction conditions so we planned to bypass this by attaching the affinity tag directly to the small molecule. This also had the advantage of greatly simplifying the sample preparation, but the greater size of the structural modification meant that biological activity could also be abolished. To test the possibility of attaching a tether to RSL3, 1S3R PEG RSL3 (8) was synthesized by click chemistry with the 1S3R alkyne RSL3 probe (6) and an acetylated polyethylene glycol (PEG) linker containing a terminal azide (Figure 2.7A). Although the EC50 of the molecule was reduced more than twenty fold compared with RSL3, an important window of *RAS*-selective lethality remained.

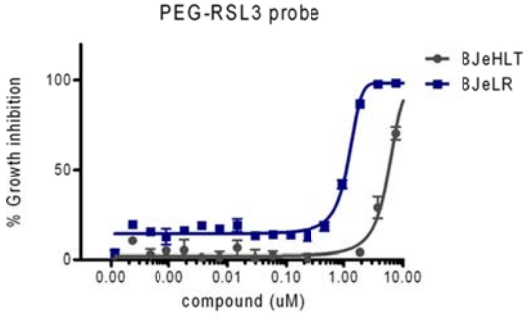
These results encouraged us to proceed with the synthesis of other tethered analogs of RSL3. We speculated that the triazole functionality could be impeding target binding by its close proximity to RSL3 and that if this connection could be extended further away or at a different position, it could result in improved binding perceived by higher potency. Thus, two new alkyne probes (9) and (10) were synthesized to enable attachment of affinity tags at alternative positions (Figure 2.7B). Alkyne-RSL3 probe 2 (9) was essentially the same structure as one of the original affinity tags (2), except remade in the (1S,3R) configuration. Alkyne-RSL3 probe 3 (10) replaced a methyl ester at the 3 position with a propargyl ester, an area that had been unexplored for tether attachment. Both molecules showed *RAS*-selective lethal activity and were used for the creation of affinity reagents connecting RSL3 directly to the fluorescein affinity moiety.

To prepare the fluorescein-RSL3 probes, a similar synthetic strategy as the one used to create (8) was again employed. A fluorescein connected via a PEG linker to an azide was first protected with isobutyryl groups to improve the cellular permeability of the fluorescein moiety as we planned to use these molecules for in-cell treatment. The protected fluorescein azide was then connected to the three alkyne-RSL3 probes (6, 9, 10) to give three fluorescein-RSL3 probes (11-13) (Figure 2.7C). While (11) and (12) showed similar activity to (8), which was modified at the same position, extension beyond the propargyl ester at position 3 in (13) severely reduced activity. As (12) showed the greatest potency and selectivity of the three fluorescein-RSL3 molecules, it was chosen for use as the optimal probe in the next set of affinity pull downs.

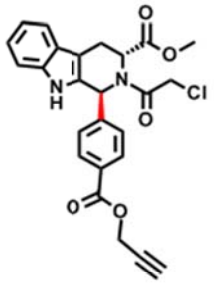
A



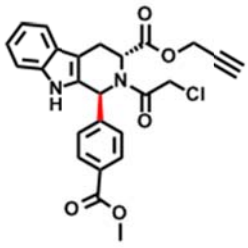
8


PEG-RSL3 probe
% Growth inhibition
100
50
0
0.00
0.00
0.01
0.10
1.00
10.00
compound (uM)
BJeHLT
BJeLR

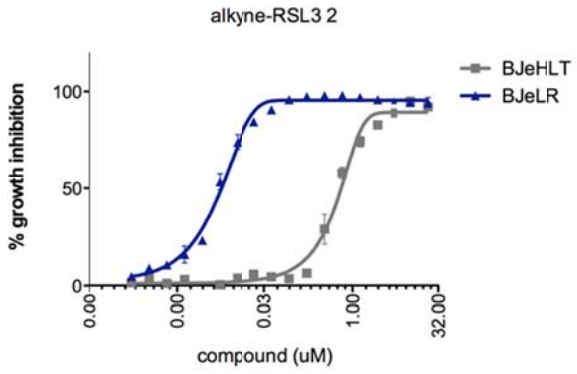
B

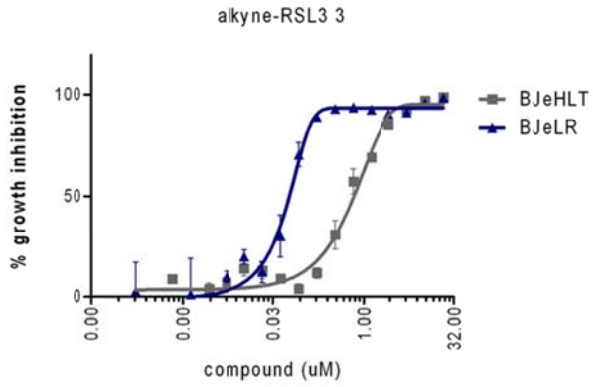


9



10


alkyne-RSL3 2
% growth inhibition
100
50
0
0.00
0.00
0.03
1.00
32.00
compound (uM)
BJeHLT
BJeLR


alkyne-RSL3 3
% growth inhibition
100
50
0
0.00
0.00
0.03
1.00
32.00
compound (uM)
BJeHLT
BJeLR

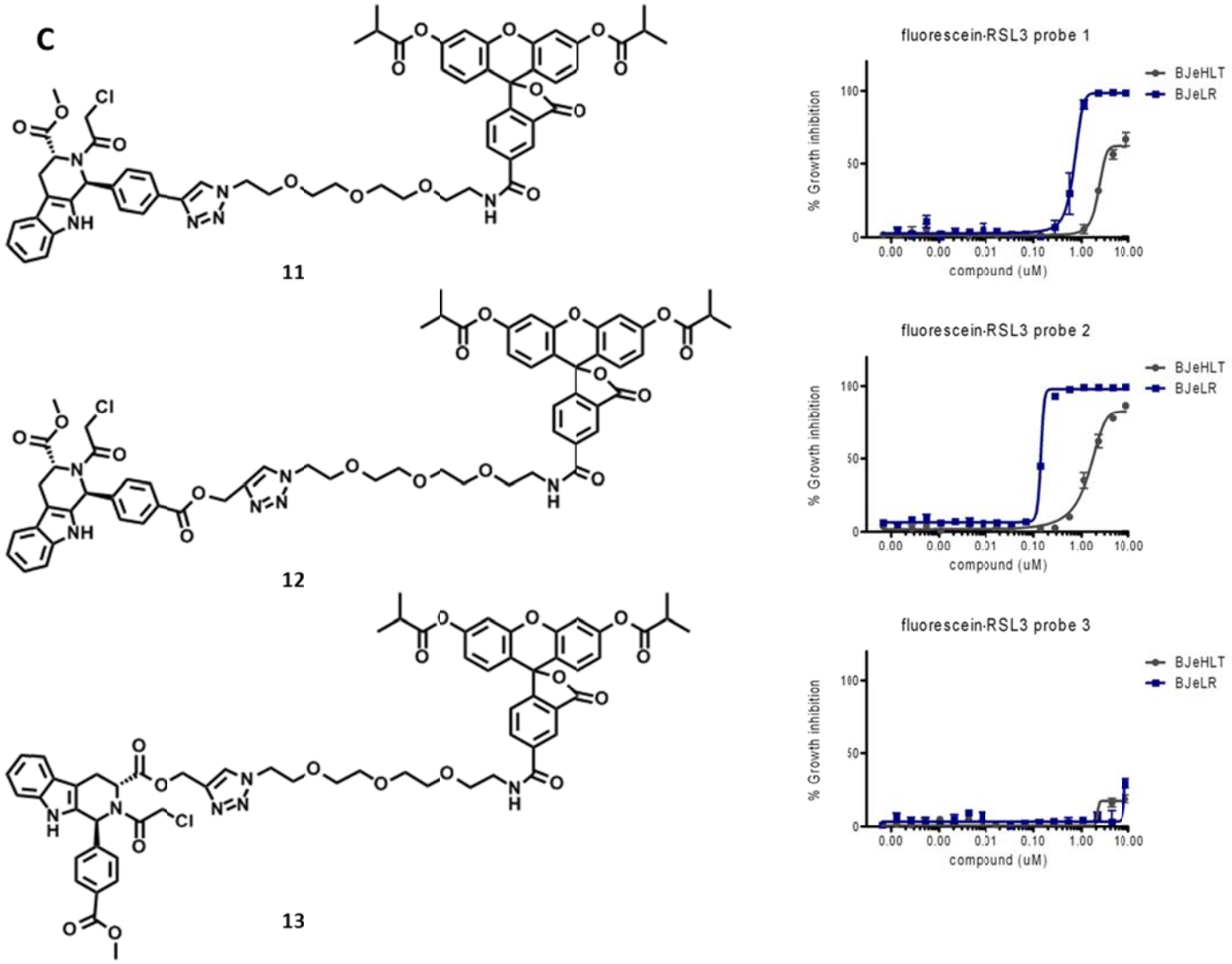


Figure 2.7: Structure and dose response of PEG RSL3, alkyne-RSL3 probes 9-10 and fluorescein-RSL3 probes 11-13

(A) A version of RSL3 with a tether was synthesized and tested to see whether a large modification on the RSL3 scaffold would still retain *RAS*-selective lethality. Although the potency was reduced greater than twenty fold in comparison with the original RSL3 scaffold, crucially a small window of *RAS*-selective lethality still remained implying that it could still be used to probe for interactions of RSL3.

(B) Alkyne-RSL3 probes 9 and 10 both showed *RAS*-selective lethality with potencies comparable to the parent RSL3 molecule, although 9 appeared the more potent of the two analogs.

(C)Fluorescein probes 11 and 12 showed *RAS*-selective lethality while 13 drastically appeared to lose all lethality. It indicates that position 1 on the RSL3 scaffold is the optimal position for modification thus tested while extension beyond the methyl ester at position 3 is not tolerated. Fluorescein probe 12 was used in further pull down experiments as it exhibited the greatest potency and selectivity.

Data represents the mean of three replicates ±SD.

Our new pull down strategy allowed us to solve the previous issues from background labeling due to click chemistry as evident when comparing fluorescein-probe treated lysates with vehicle-treated lysates by western blot - no labeled proteins were visible with vehicle-treatment (Figure 2.8A). Additional improvements were made by increasing the vigorousness of washing steps used during the affinity purification and the ionic strength of the buffers. The scale of lysate preparation was also increased six-fold as the amount of protein used in the first pull down experiment had been near minimal limits. Scaling up was also partially facilitated by the reduction in complexity of the protocol. In the final sample preparation, four different treatments were prepared. The first two treatments were with (1S,3R) fluorescein-RSL3 probe 12 (referred to as 'probe') or the corresponding diastereomer (1R,3R) fluorescein-RSL3 as an inactive negative control. Two sets of 'competitor' treated cells were made by pretreatment of cells with 0.1X and 1X equivalents (50nM and 500nM) of (1S, 3R) RSL3 prior to treatment with 1S3R fluorescein-RSL3 probe 12 (Figure 2.8B). The 'competitor' treated cells were not pretreated with an excess of parent molecule due to the difference in potency between the parent and probe molecule. By using two different competitor concentrations, we hoped to distinguish between higher affinity interactions with the lower dose of competitor and lower affinity interactions with the higher dose of competitor. The results of the fluorescein-RSL3 pull down experiments will be discussed in Chapter 4.

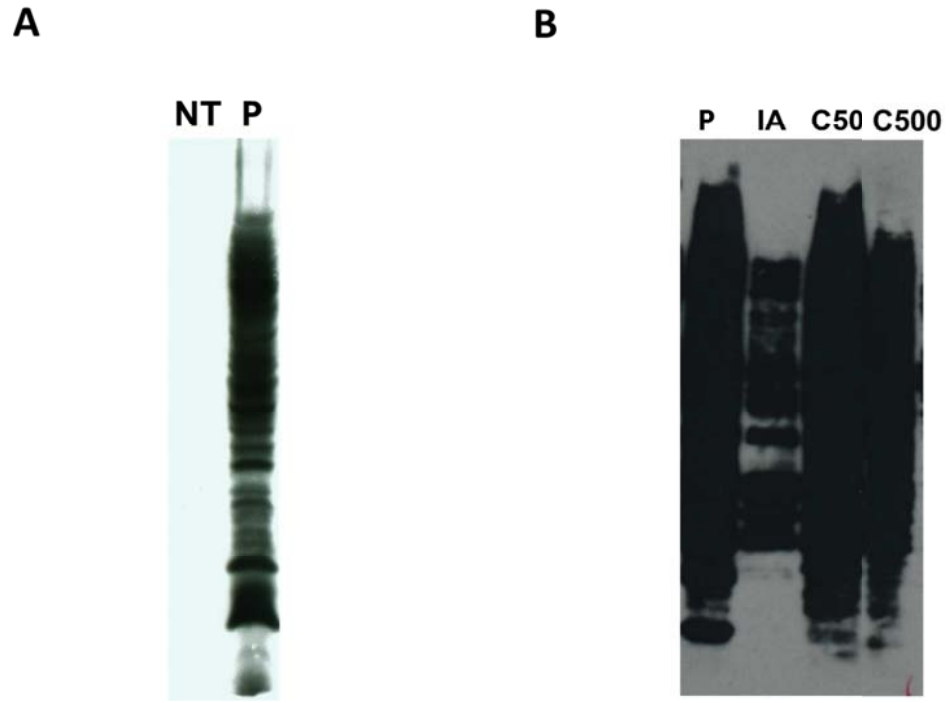


Figure 2.8: Lysates prepared from cells treated with fluorescein-RSL3 analogs

(A) Comparison of vehicle-treated (DMSO - **NT**) and fluorescein-RSL3 treated (**P**) cells (at 0.5uM) by anti-fluorescein western blot. The NT sample shows no non-specific interaction of the antibody with the lysate and all bands in the P lane represent fluorescein-RSL3 bound proteins.

(B) Four treatments were used in the fluorescein-RSL3 affinity pull down experiment, 0.5uM (1S,3R) fluorescein-RSL3 probe 12 (**P**) or the corresponding inactive diastereomer (1R,3R) fluorescein-RSL3 (**IA**). Two sets of 'competitor' treated cells were made by pretreatment of cells with 0.1X and 1X equivalents (50nM and 500nM) of (1S, 3R) RSL3 prior to treatment with **1**S3R fluorescein-RSL3 probe 12 (**C50** and **C500**).

2.2d Attenuating the reactivity of RSL3

The reactive chloromethyl ketone functional group on the RSL3 scaffold has been shown to be essential for the molecule's activity. It is likely that its high potency is due to the irreversible inhibition of its targets. The nature of its covalent interactions have proven beneficial for the identification of its binding partners, allowing for incubation in cells and easing the technical difficulties of maintaining proteins in a native state in pull down experiments. At the same time, it is also evident that the molecule reacts with a great number of proteins non-specifically, convoluting target identification efforts. Additionally, if RSL3 were to be considered as a scaffold for a drug candidate, concerns of toxicity risk would certainly be challenging. To improve the selectivity of RSL3's target profile by reducing non-specific reactivity with cellular nucleophiles we performed an SAR study exploring replacement of the electrophilic chloromethyl ketone group.

As the chloride ion is a good leaving group, our initial efforts focused on substituting it with groups of lower leaving group ability, which we approximated using the pKa value of the conjugate acid (pKa of HCl = -7). Analogs of RSL3 containing an azide (SRS1-44), an acetoxy (SRS1-91) and a fluoride (SRS1-99) were synthesized by Dr. Rachid Skouta and tested by the author to determine the effects on potency in the four BJ cell lines (Figure 2.9). Interestingly, while SRS1-99 containing a fluoride anion with the greatest leaving group ability of the three (pKa of HF = 3.2), SRS1-44 containing an azide group (pKa of HN_3 = 4.6) was the only compound to exhibit some lethality at high concentrations that appeared to be *RAS*-selective lethal (Figure 2.10B). The potency however had dropped approximately 2000-fold. Both SRS1-99 and SRS1-91 showed no activity at all at the highest concentration tested (data not shown). Thus the relationship between activity and leaving group ability of the electrophile in RSL3's case is not straightforward.

Next, we explored the effects of increasing the distance between the leaving group and the ketone. Analogs of RSL3 with a chloroethyl ketone (SRS7-09) and a chloropropyl ketone (SRS7-21) were synthesized. Again the activity was dramatically affected, SRS7-09 showed lethality at high doses with a loss in potency of approximately 1000-fold and appeared to have no *RAS* selective lethality (Figure 2.10C) while SRS7-21 was inactive (data not shown). The effects of the ketone's pi system clearly enhance the reactivity at the alpha carbon. Additionally, different types of electrophiles were also tested such as a vinyl sulfonamide (SRS7-22A) and substituted and unsubstituted acrylamides (SRS1-22, SRS2-42, SRS1-25) as well as electrophiles with increasing substitution on the alpha carbon (SRS6-62, SRS6-70). Only the vinyl sulfonamide (SRS7-22A) and the unsubstituted acrylamide (SRS1-22) had non-selective lethality at high doses (Figure 2.10D). The lack of selective lethality suggests that these analogs are no longer binding the targets of RSL3 and are instead killing via a non-specific mechanism.

Interestingly, two analogs that contained no leaving group (SRS7-37, SRS7-51) displayed weak but *RAS*-selective lethality (Figure 2.10 E-F). These analogs could show weak affinity to RSL3 targets, but lack potency due to the reversible nature of their interaction. Although this study did not identify any electrophile analogs with similar potency to RSL3, it does demonstrate that molecules with the RSL3 scaffold that cannot irreversibly bind targets still possess *RAS*-selective lethality. Ideally, we would like to find an analog where the intrinsic reactivity of the electrophile is attenuated to minimize non-specific reactions but the non-covalent affinity for the target is optimized [37].

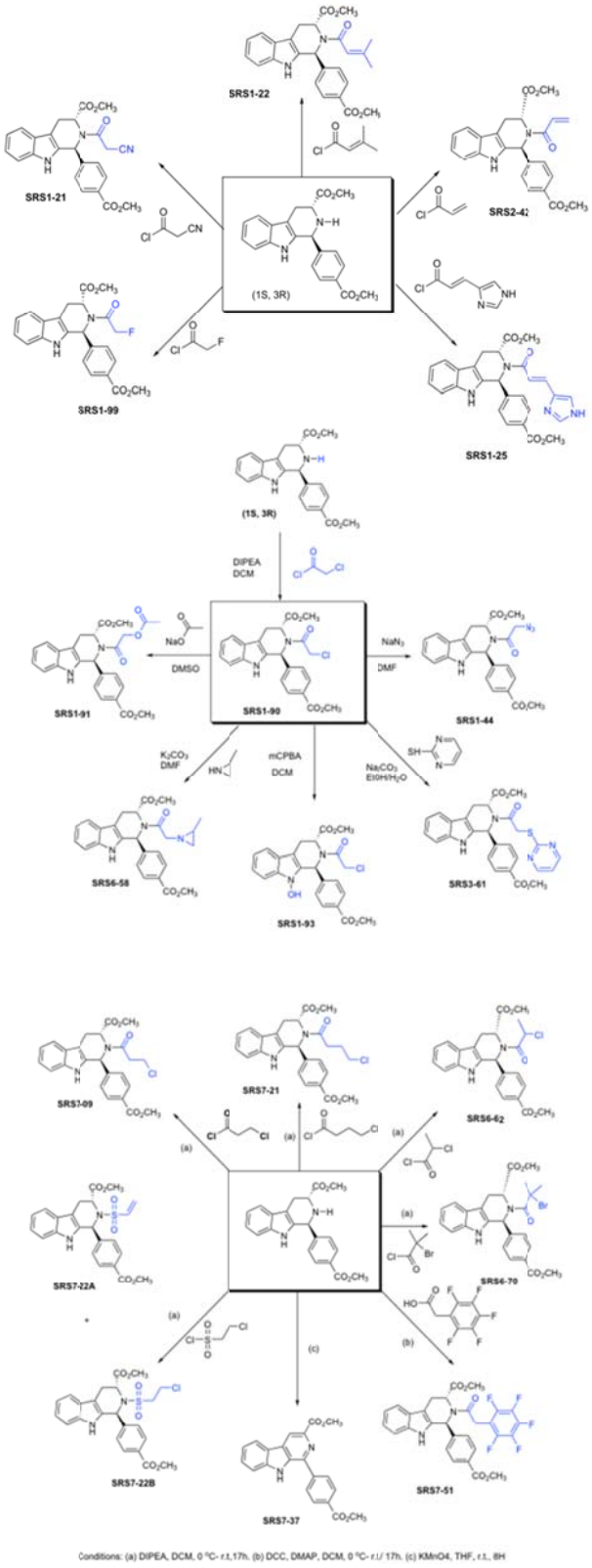


Figure 2.9: Electrophile analogs of RSL3

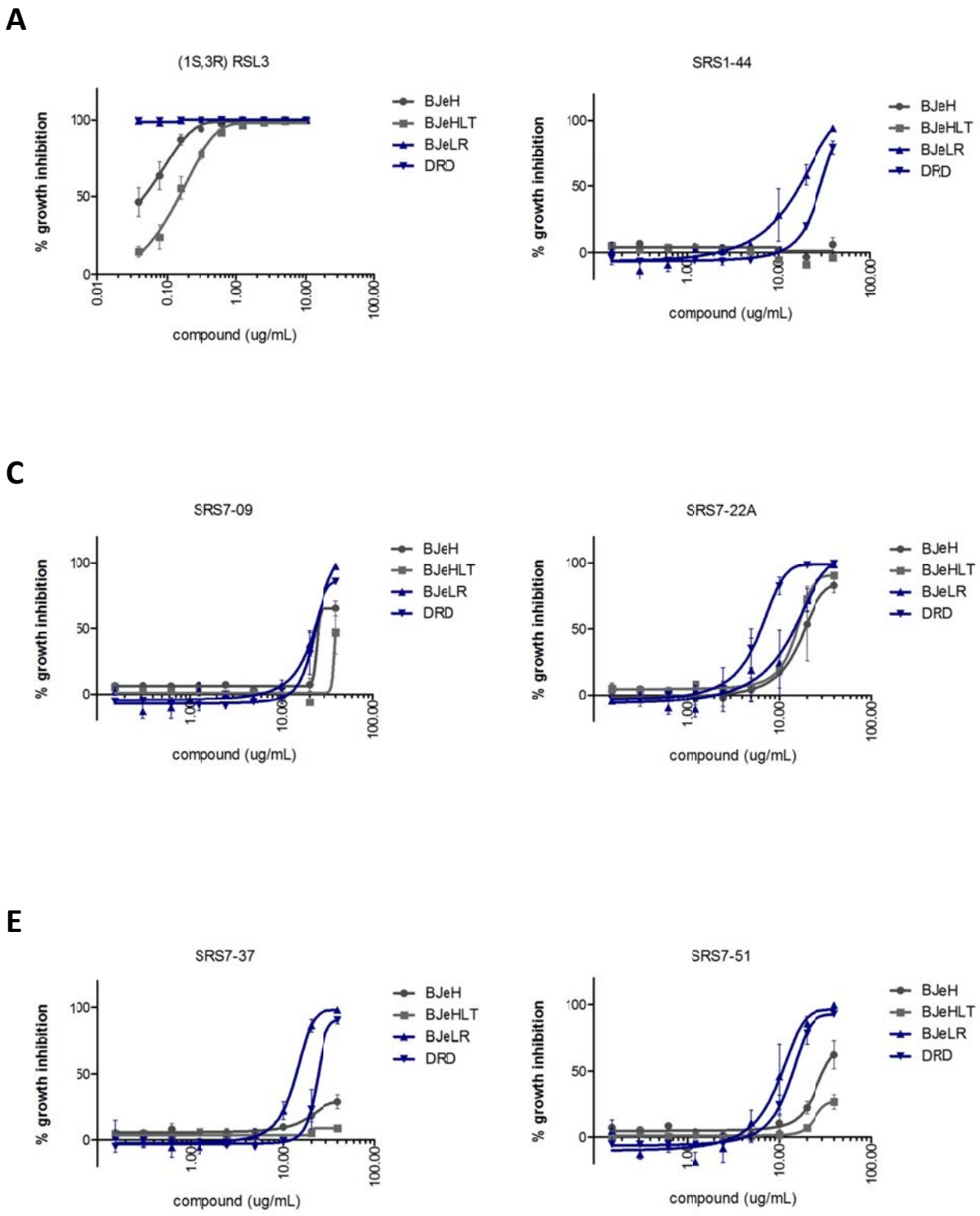


Figure 2.10: Dose response curves in the four BJ cell lines of RSL3 electrophile analogs

2.3 Discussion

In this work, different strategies for the target identification of RSL3 using affinity chromatography methods were implemented. These experiments first entailed the synthesis of an affinity reagent from the RSL3 scaffold. Before this could be achieved successfully, a vital discovery of the specific stereoconfiguration of the active RSL3 molecule was made. Further, the pure active diastereomer of RSL3 displays improved potency and selectivity in our *RAS*-selective lethal assays.

A number of affinity reagents for use in target identification were synthesized incorporating both alkyne and fluorescein tags. Pull down experiments were performed using the alkyne-tagged RSL3, which attaches a tag to the molecule after binding has occurred via a click reaction. Issues of background labeling resulting from click chemistry however hindered this method and results from proteomic experiments were not convincing. Thus the former strategy of directly connecting a tag to the small molecule was employed along with additional optimization. By simplifying the approach, background labeling was avoided and this enabled the experiment to be performed on a much larger scale, potentially facilitating detection of low abundance proteins. These results highlight the potential difficulties with using increasingly complex methods for target identification, which can add further variability into experiments.

Additionally, efforts were made to reduce the number of non-specific reactions of RSL3 with the proteome by attenuating its electrophile. These experiments revealed that replacement of the chloride was not trivial and that the relationship between the reactivity of the electrophile and biological activity of the molecule is complex.

2.4 Methods

2.4a Materials

All chemical nmaterials were purchased from Sigma Aldrich unless otherwise stated. Methyl 1-(4-nitrophenyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-3-carboxylate was obtained from Life Chemicals. The synthesis of biotin-6-aminohexanoic acid was previously published [32].

2.4b Synthesis

Biotin RSL3 (Compound 1)

Methyl 1-(4-nitrophenyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-3-carboxylate (100mg, 2.85e-4 mol) was dissolved in acetonitrile (CH_3CN, ~2mL) along with 120uL (8.54e-4 mol) of triethylamine and stirred under N_2. Boc anhydride (100 mg, 5.69 e-4 mol) and DMAP (7 mg, 5.69e-5 mol) were dissolved in 200uL of CH_3CN and then added dropwise to the reaction mixture, which changed from yellow to dark brown in color upon addition. The reaction was heated to 50°C and left to stir overnight. An additional 1-3 equivalents of boc anhydride were added until all starting material was observed to be consumed by TLC and MS. Upon reaction completion, CH_3CN was removed under reduced pressure and the mixture was redissolved in CH_2Cl_2. The reaction was quenched by addition of 2N NH_4Cl, the organic layer was removed, washed with brine and dried with Na_2SO_4. The resulting crude boc-protected intermediate was partially purified by silica chromatography using 2% MeOH in CH_2Cl_2 and taken forward to the next step. MS (m/z): $[M]^+$ = 451.84

<u>Reduction protocol</u>

The boc-protected product was dissolved in a small amount of CH_2Cl_2 and was added to 20-30mL ethanol. The reaction vessel was first evacuated and then refilled with Ar several times to replace the air in the vessel and a catalytic amount of 10% Pd/C catalyst was added. The reaction vessel was again evacuated and then refilled with H_2 (using a balloon attached to the flask with a three way joint). The

reaction was left to stir overnight. The next day full consumption of starting material was observed by TLC and the reduced product had an R_f value of 0.21-0.25 in 25% A (A = 10% methanol in EtOAc) in CH_2Cl_2. The crude product was purified by silica gel chromatography, eluting with a gradient between 10-25% A in CH_2Cl_2 to give 50mg of the aniline intermediate as determined by MS (2-tert-butyl 3-methyl 1-(4-aminophenyl)-3,4-dihydro-1H-pyrido[3,4-b]indole-2,3(9H)-dicarboxylate) with a 20% yield over 2 steps . MS (m/z): $[MH]^+$ = 422.3 This reaction was performed with the assistance of Dr. Matt Tremblay, a former graduate student in the Sames lab.

Reduction protocol

The aniline intermediate was dissolved (25mg, 5.94e-5 mol) in anhydrous $CHCl_3$ (~1mL) and stirred under N_2 in a dry conical vial. Biotin-6-aminohexanoic acid (23 mg, 6.53e-5 mol) was dissolved along with EDC and DMAP in anhydrous DMF (~1mL). This solution was then added dropwise to the amine solution and left to stir at 50°C overnight. Upon reaction completion, CH_2Cl_2 was added and the aqueous layer was extracted several times. The combined organic layers were washed successively with 1N HCl, H_2O, saturated aqueous $NaHCO_3$ and brine, and then dried with Na_2SO_4. The coupled biotin-RSL3 intermediate was purified by silica chromatography using 7% MeOH in CH_2Cl_2 with an 11% yield. MS (m/z): $[MH]^+$ = 761.6

Boc-deprotection protocol

The coupled intermediate (3 mg, 7.1 umol) was dissolved in a solution of 30% TFA/CH_2Cl_2 and stirred for an hour at room temperature. The solution turned a reddish pink color. The solvents and TFA were removed in vacuo, redissolved in CH_2Cl_2 and the crude product was extracted with 5% aqueous K_2CO_3 solution. The red color disappeared and the organic phase resumed a yellow color. The organic phase

was washed twice with H_2O, dried with Na_2SO_4, concentrated and the deprotected intermediate was taken directly to the next step. LRMS (m/z): $[MH]^+$ = 661.5

Chloroacetylation protocol

The deprotected intermediate (~1mg) was dissolved in anhydrous $CHCl_3$ and 2-3 equivalents of triethylamine were added. To this was added 0.5 equivalents of chloroacetyl chloride diluted in a solution of anhydrous $CHCl_3$. Consecutive half equivalents were added and the reaction was followed by MS until all the deprotected intermediate was chloroacetylated. Attempts to characterize the final product by NMR were made but the low amount of material prevented full characterization. Subsequent negative biological data led to a decision not to pursue this analog further. MS (m/z): $[MH]^+$ = 737.5

The synthesis of alkyne-RSL3 using the pictet-spengler reaction (Compound 3)

The hydrochloride salt of L-tryptophan methyl ester (0.4 g, 1.84mmol) was "free-based" by dissolving it in CH_2Cl_2 and then adding 5% aqueous K_2CO_3. The organic layer was extracted, dried with Na_2SO_4 and concentrated in vacuo. To a dry vial was added L-tryptophan methyl ester along with molecular sieves dissolved in dry CH_2Cl_2. 4-ethynylbenzaldehyde (0.2 g, 1.54mmol) and 0.1 eq of TFA (11.4uL, 0.154mmol) were added to the reaction mixture at 0°C and the solution was then warmed to room temperature and stirred overnight. The following day, 3 equivalents of TFA (342uL, 4.62mmol) were added to the solution. The reaction was allowed to stir for an additional 3-8 hours then quenched with saturated aqueous $NaHCO_3$. The organic phase was separated, washed with brine and dried with Na_2SO_4 and then concentrated to give the crude product. The compound was purified twice by dry-loaded silica gel chromatography in 20:80 EtOAc:Hexane with 1% NEt_3 in order to separate the two diastereomers ('cis' Rf = 0.29, 'trans' Rf = 0.13). As a result of significant losses during purification, the total final yield was ~5%. MS (m/z): $[MH]^+$ = 331.5

The following protocol was performed on both diastereomers separately, a representative example using the cis diastereomer is described. The 'cis' para ethynyl RSL3 intermediate (12mg, 37.2umol) was dissolved in anhydrous CH_2Cl_2 and 2-3 equivalents of triethylamine (10uL, 74.4umol) were added. To this was added 0.5 equivalents of chloroacetyl chloride (1.5uL, 18.6umol) diluted in a solution of dry CH_2Cl_2 (30uL)dropwise at 0°C. Consecutive half equivalents were added and the reaction was followed by MS and TLC until all the starting material was chloroacetylated. The reaction was quenched by addition of saturated $NaHCO_3$ and the organic layer was extracted, washed with brine and dried with Na_2SO_4 and then concentrated. The desired compound was purified by dry-loaded silica gel chromatography in 20:80 EtOAc:Hexane with close to quantitative yield.

1H NMR (500 MHz, CDCL3) δ 7.88 (s, 1H), 7.65 (d, J = 7.9 Hz, 2H), 7.44 (d, J = 7.5 Hz, 4H), 7.36 (d, J = 8.1 Hz, 2H), 7.31 – 7.22 (m, 5H), 7.19 (t, J = 7.4 Hz, 2H), 6.96 (s, 2H), 4.98 (s, 2H), 4.42 (s, 2H), 4.27 (d, J = 13.1 Hz, 2H), 3.70 (d, J = 16.1 Hz, 2H), 3.23 (dd, J = 15.8, 6.7 Hz, 2H), 3.12 (s, 5H). MS (m/z): $[MH]^+$ = 407.7

Assignment of NMR – D1 and D2 were assigned cis and trans using NOESY spectra, for the cis diastereomer a cross peak is visible between the O and B protons, which is not visible in the trans diastereomer.

Certain protons around the chloroacetylated N were significantly downshifted from their original chemical shifts in the unchloroacetylated intermediate. HSQC was used to assign their new locations, the O proton is now found in near 7 ppm (downshifted in the proton spectrum but a regular C13 shift), and the B proton in the 4-5 range. The two alpha protons are diastereotopic and appear as a doublet of doublets.

5-Propargyloxy RSL3 (Compound 4)

L-methyl ester 5-hydroxytryptophan (0.3g, 1.1mmol) was dissolved in water and stirred along with ~7 equivalents of $NaHCO_3$ (0.65g, 7.8mmol) in order to basify the solution to pH 8-9. BOC-ON (1.37g,

5.54mmol) was dissolved in anhydrous dioxane and added dropwise to the reaction at 0°C. The reaction was monitored by MS and additional equivalents of BOC-ON were added until reaction completion. The solvent was removed in vacuo. The boc-protected intermediate was purified by silica gel column chromatography in 1% EtOH/$CHCl_3$ with a 26% yield (Rf = 0.19 in 1% EtOH/$CHCl_3$). HRMS (m/z): $[MH]^+$ = 372.1677

The boc-protected intermediate (30mg, 90umol) was dissolved in anhydrous CH_2Cl_2 with a small amount of acetone and ~7 equivalents of K_2CO_3 was added (87mg, 0.628mmol)) and stirred rapidly under N_2. To this propargyl bromide was added (7uL, 81umol) and the extent of propargylation was monitored by MS. Additional equivalents of propargyl bromide were added until reaction completion was observed. The reaction mixture was filtered and the propargylated intermediate was purified by silica gel column chromatography in 2% EtOH/$CHCl_3$. The propargylated intermediate was deprotected by stirring in 30% TFA/CH_2Cl_2 for an hour. The solvents and TFA were removed in vacuo, redissolved in CH_2Cl_2 and the crude product was washed with 5% aqueous K_2CO_3 solution. Subsequently the organic phase was washed twice with ddH_2O, dried with Na_2SO_4 and concentrated to give a 61% yield over two steps. LRMS (m/z): $[MH]^+$ = 374.2

To a dry vial was added the propargylated 5-hydroxy L-tryptophan methyl ester (15mg, 55umol) along activated with molecular sieves in anhydrous CH_2Cl_2. Methyl 4-formyl benzoate (13mg, 81umol) and TFA (0.3uL, 4umol) were added to the reaction mixture at 0°C and the solution was then warmed to room temperature and stirred overnight. The following day, more TFA (9uL, 120umol) was added to the solution. The reaction was allowed to stir for an additional 3-8 hours then quenched with saturated aqueous $NaHCO_3$, the organic phase was separated, washed with brine and dried with Na_2SO_4 and then concentrated to give ~10mg of crude product. The compound was then chloroacetylated as described above. The final product was purified by silica gel chromatography in 50:50 EtOAc:Hexane. A small

amount of pure compound (~1mg) was used for preliminary biological testing. Rf = 0.15 in 1% EtOH/$CHCl_3$. LRMS (m/z): $[MH]^+$ = 495.6

1H NMR (400 MHz, CDCL3) δ 7.98 (d, J = 8.1 Hz, 2H), 7.51 – 7.37 (m, 3H), 7.09 (dd, J = 7.8, 5.5 Hz, 2H), 6.84 (dd, J = 8.7, 2.2 Hz, 1H), 5.26 (s, 1H), 4.72 (dd, J = 10.8, 2.1 Hz, 2H), 3.98 – 3.91 (m, 1H), 3.88 (d, J = 6.3 Hz, 3H), 3.81 (d, J = 8.1 Hz, 3H), 3.27 – 3.09 (m, 1H), 3.06 – 2.89 (m, 1H). LRMS (m/z): $[MH]^+$ = 495.6

1-Ethyl ester RSL3(Compound 5)

1-Ethyl ester RSL3 was prepared in an analogous procedure to that used in the preparation of alkyne RSL3, except an L-tryptophan ethyl ester and methyl 4-formyl-benzoate were used in place of L-tryptophan methyl ester and 4-ethynyl benzaldehyde in the pictet-spengler reaction. The two diastereomers were separated by silica gel chromatography in 40:60 EtOAc/Hexane. The first diastereomer, D1 has an Rf value of 0.14 and D2 an Rf value of 0.06 in 20:80 EtOAc:Hexane. Chloroacetylation was carried out as described above.

1H NMR (500 MHz, D2O) δ 7.90 (d, J = 7.5 Hz, 2H), 7.84 (s, 1H), 7.62 (d, J = 7.7 Hz, 1H), 7.37 (d, J = 6.0 Hz, 2H), 7.31 (d, J = 8.0 Hz, 1H), 7.23 (t, J = 7.5 Hz, 1H), 7.18 (t, J = 7.4 Hz, 1H), 6.96 (s, 1H), 4.96 (s, 1H), 4.36 (d, J = 12.1 Hz, 1H), 4.22 (d, J = 12.5 Hz, 1H), 3.91 (d, J = 12.3 Hz, 3H), 3.72 (d, J = 16.8 Hz, 2H), 3.23 (dd, J = 15.8, 7.0 Hz, 1H), 3.13 (s, 1H).

Acryloyl RSL3

Acryloyl RSL3 was prepared in an analogous procedure to that used in the preparation of alkyne RSL3, except that methyl 4-formyl-benzoate was used in place 4-ethynyl benzaldehyde. The intermediate was purified by dry loaded silica gel chromatography in 20:80 EtOAc/Hexanes. Instead of chloroacetyl chloride, acryloyl chloride was used to acryloylate the intermediate in an identical procedure used for chloroacetylation to give the acrylamide version of RSL3.

1H NMR (400 MHz, CDCL3) δ 8.32 (s, 3H), 7.90 (d, J = 7.2 Hz, 6H), 7.63 (d, J = 7.6 Hz, 3H), 7.35 (d, J = 7.8 Hz, 9H), 7.22 (dtd, J = 19.4, 7.2, 1.1 Hz, 6H), 7.10 (s, 2H), 6.80 – 6.62 (m, 3H), 6.34 (d, J = 16.5 Hz, 3H), 5.82 (dd, J = 10.8, 1.3 Hz, 3H), 5.08 (s, 2H), 3.89 (s, 9H), 3.79 – 3.62 (m, 3H), 3.11 (dd, J = 15.7, 6.7 Hz, 4H), 3.02 (s, 7H).

Cyclized RSL3

Following a reported protocol [38], glycolic acid (200mg, 2.63mmol) was added to an oven-dried flask and dried under vacuum for an hour. The solid was then dissolved in anhydrous THF under N_2 and DMAP (64mg, 0.526mmol) and 1.01 equivalents of anhydrous NEt_3 (0.370mL, 2.66mmol) were added to the solution at 0°C. Following five minutes of stirring at 0°C, 1.05 equivalents of TMSCl (0.352mL, 2.76mmol) was added dropwise, yielding a white slurry (NEt_3HCl salt forms). This suspension was stirred for an additional hour at 0°C, and then mixed for a second hour at RT. A second equivalent of NEt_3 was added at room temperature followed by 1.05 equivalents of TBDPSCl (0.707mL, 2.76mmol), and this was stirred for 8 hours at 25°C (or overnight). Upon completion, the reaction was quenched by the addition of an aqueous solution of HOAc (~ 30% w/v) and stirred for 1 hour at room temperature. Water was added, and this solution was extracted with Et_2O three times. The combined organic layers were then washed with water twice, dried with Na_2SO_4 and concentrated in vacuo. The resultant light yellow oil was purified by silica gel chromatography in 50:50 Et_2O:Hexane, Rf = 0.18 with a yield of 78%.

The protected glycolic acid (200mg, 0.637mmol) was concentrated in an oven-dried flask and dried under vacuum for an hour. The starting material was dissolve in anhydrous CH_2Cl_2 and thionyl chloride (83mg, 0.7mmol) was added along with a drop of anhydrous DMF as a catalyst. After stirring for 2 hours under N_2 the reaction mixture was concentrated and redissolved in anhydrous CH_2Cl_2 and re-concentrated. This process was repeated twice to remove excess $SOCl_2$. The resultant oil was dissolved in anhydrous CH_2Cl_2 and NEt_3 (78uL, 0.700mmol) and RSL3 intermediate (prepared as described in

protocol for acryloyl RSL3) (100mg, 0.274mmol). The reaction mixture was allowed to stir overnight under N_2. The The solvent was removed in vacuo and crude siloprotected intermediate was purified by silica gel chromatography by a gradient elution from 100% hexane to 50:50 ether:Hexane followed by elution with 20:80 EtOAc:Hexane (Rf = 0.27 and 0.15) in 66% yield.

The purified siloprotected RSL3 intermediate (50mg, 75.7umol) was dissolved in anhydrous CH_2Cl_2 in an oven-dried flask under N_2. To this TBAF in THF (83ul of 1.0M solution, 83.3umol) was added and allowed to stir for five hours. The reaction was quenched with saturated NH_4Cl and extracted three times into CH_2Cl_2. The combined organic layers were washed with brine, dried with Na_2SO_4, filtered and concentrated. The final product was purified by a gradient elution of EtOAc: Hexane (Rf = 0.43 in 80:20 EtOAc: Hexane) with a 17% yield.

1H NMR (500 MHz, D2O) δ 7.99 (d, J = 8.2 Hz, 3H), 7.56 (d, J = 7.8 Hz, 1H), 7.35 (d, J = 8.3 Hz, 3H), 7.27 (d, J = 7.4 Hz, 1H), 7.20 (t, J = 7.5 Hz, 1H), 7.04 (s, 1H), 5.03 – 4.81 (m, 2H), 4.48 (dd, J = 12.2, 4.3 Hz, 1H), 3.91 (s, 3H), 3.54 (dd, J = 15.5, 4.2 Hz, 1H), 3.28 – 3.08 (m, 1H).

Synthesis of four diastereomers of RSL3

A generalized protocol for the Pictet-Spengler and chloroacetylation reactions is provided; procedure is slightly different to that used to prepare the para alkyne RSL3.

To a suspension of 1.2 equivalents of the corresponding S or R tryptophan methyl ester hydrochloride in CH_2Cl_2 was added 1.3 equivalents of NEt_3 at room temperature. The mixture was stirred for 1 hour and then filtered. The filtrate was concentrated down to give the required product, tryptophan methyl ester as a clear oil, which was dried under vacuum for 10 minutes. Tryptophan methyl ester, along with activated molecular sieves, was dissolved in anhydrous dichloromethane. Methyl 4-formylbenzoate (1 equivalent) and 0.1 equivalent of TFA were added to the reaction mixture and the solution was refluxed for an hour. After an hour, 3 equivalents of TFA were added to the solution and the reaction was

allowed to stir under reflux overnight. The reaction mixure was cooled to room temperature and quenched with 30% NaOH. Tthe organic phase was separated, washed with brine and dried with Na_2SO_4 and then concentrated to give the crude product. The compound was purified by dry-loaded silica gel chromatography in a gradient elution from 100% hexane with 1% NEt3 to 20:80 EtOAc:Hexane with 1% NEt_3 to give the two separate diastereomers. Mixed fractions were repurified again with the same protocol. The four diastereomers were obtained in the following yields: 1R3S RSL3 intermediate 18% and 1S3S RSL3 intermediate 34%, 1R3R RSL3 intermediate 41% and 1S3R RSL3 intermediate 16%.

The corresponding RSL3 intermediate (one of the stereoisomers) was dissolved in anhydrous CH_2Cl_2 and 1.1 equivalents of sodium bicarbonate was added. To this was added 0.5 equivalents of chloroacetyl chloride dropwise at 0°C. Consecutive half equivalents of chloroacetyl chloride were added and the reaction was followed by MS and TLC until all the starting material was completely consumed. The reaction was filtered and the filtrate was extracted, washed with brine, dried with Na_2SO_4 and concentrated. The compound was purified by silica gel chromatography in 20:80 EtOAc:Hexane. A similar protocol was used to form the chloroacetamide moiety for the other three diastereomers. The four diastereomers were obtained in the following yields: 1R3S RSL3 (49%), 1S3S RSL3 (55%), and 1R3R RSL3 (42%), 1S3R RSL3 (79%).

1S3S RSL3

1H NMR (400 MHz, CDCL3) δ 10.94 (s, 1H), 8.05 (s, 1H), 7.98 (d, J = 8.2 Hz, 2H), 7.66 (d, J = 7.7 Hz, 1H), 7.41 (d, J = 8.0 Hz, 1H), 7.36 (d, J = 8.0 Hz, 2H), 7.24 – 7.16 (m, 1H), 7.13 (dd, J = 10.9, 3.9 Hz, 2H), 7.03 (s, 1H), 5.42 (d, J = 6.5 Hz, 1H), 5.00 (d, J = 13.8 Hz, 1H), 4.61 (d, J = 13.8 Hz, 1H), 3.92 (s, 3H), 3.68 (d, J = 15.9 Hz, 2H), 3.28 (dd, J = 15.5, 6.4 Hz, 1H), 3.03 (s, 3H). 13C NMR (75 MHz, DMF) δ 171.11, 168.52, 167.84, 167.58, 166.84, 165.61, 163.05, 162.66, 162.27, 157.85, 145.68, 144.09, 137.74, 136.34, 130.11, 129.84, 129.53, 127.46, 127.01, 123.73, 118.74, 111.01, 107.66, 79.56, 56.18, 55.36, 55.07, 53.79.

1R3R RSL3

1H NMR (400 MHz, CDCL3) δ 10.93 (s, 1H), 7.98 (d, J = 8.2 Hz, 2H), 7.65 (d, J = 7.7 Hz, 1H), 7.40 (d, J = 8.0 Hz, 1H), 7.35 (d, J = 8.0 Hz, 2H), 7.25 – 7.16 (m, 1H), 7.11 (d, J = 7.1 Hz, 1H), 7.02 (s, 1H), 5.41 (d, J = 6.5 Hz, 1H), 5.00 (d, J = 13.8 Hz, 1H), 4.61 (d, J = 13.7 Hz, 1H), 3.67 (d, J = 15.9 Hz, 1H), 3.27 (dd, J = 15.5, 6.5 Hz, 1H), 3.02 (s, 3H). 13C NMR (75 MHz, CDCl3) δ 170.74, 167.25, 144.89, 136.95, 136.00, 135.98, 132.86, 132.60, 130.49, 130.47, 129.86, 129.66, 129.47, 128.32, 128.27, 126.84, 122.91, 120.10, 118.99, 111.63, 108.19, 65.53, 27.21, 22.12, 19.68. HRMS (m/z): [MH]+ calculated for $C_{23}H_{21}ClN_2O_5$, 440.1139; found 440.1146.

1R3S RSL3

1H NMR (300 MHz, CDCL3) δ 10.99 (s, 1H), 8.10 (d, J = 8.2 Hz, 2H), 7.86 (d, J = 8.2 Hz, 2H), 7.70 (d, J = 7.4 Hz, 1H), 7.49 (d, J = 7.8 Hz, 1H), 7.34 – 7.04 (m, 2H), 6.52 (s, 1H), 5.56 (s, 1H), 4.84 (d, J = 13.6 Hz, 1H), 4.51 (d, J = 13.6 Hz, 1H), 4.04 (s, 3H), 3.83 (d, J = 15.6 Hz, 1H), 3.65 (dd, J = 15.6, 4.6 Hz, 1H). 13C NMR (75 MHz, DMF) δ 172.08, 168.79, 166.86, 138.03, 134.27, 130.08, 129.81, 127.36, 126.99, 122.21, 119.72, 118.54, 112.02, 105.65, 57.75, 57.17, 52.59, 52.09, 43.39, 23.92.

1S3R RSL3

1H NMR (500 MHz, D2O) δ 8.19 (s, 1H), 8.10 (s, 2H), 7.87 (s, 2H), 7.69 (d, J = 6.9 Hz, 1H), 7.48 (d, J = 6.9 Hz, 1H), 7.22 (dd, J = 19.1, 7.2 Hz, 2H), 6.52 (s, 1H), 5.56 (s, 1H), 4.83 (d, J = 12.8 Hz, 1H), 4.51 (d, J = 12.3 Hz, 1H), 4.04 (s, 3H), 3.90 – 3.71 (m, 4H), 3.65 (d, J = 13.4 Hz, 1H). 13C NMR (75 MHz, DMF) δ 171.93, 171.27, 168.63, 167.63, 166.71, 166.34, 137.86, 137.30, 134.12, 133.31, 129.91, 129.63, 127.19, 126.82, 125.25, 122.03, 119.54, 118.37, 111.85, 110.35, 57.58, 57.00, 52.42, 51.93, 43.23, 23.75. HRMS (m/z): [MH]+ calculated for $C_{23}H_{21}ClN_2O_5$, 440.1139; found 440.1143.

1S3R & 1R3R alkyne- RSL3 probe 1 (Compound 6 & 7)

Synthesis of 1S3R and 1R3R para RSL3 analogs was performed analogously to the four diastereomers of RSL3, only 4-ethynyl benzaldehyde was used instead of methyl 4-formylbenzoate. The structure was identical to that synthesized in para alkyne-RSL3, however instead of using the L-tryptophan methyl ester hydrochloride, the R enantiomer was used. 1S3R and 1R3R alkyne RSL3 intermediate were synthesized with 36% and 59% yield respectively. The chloroacetylation described above was altered to remove the extraction procedure and the resulting yields of the final products became near quantitative.

1R3R

1H NMR (400 MHz, CDCL3) δ 10.90 (s, 1H), 7.64 (d, J = 7.7 Hz, 1H), 7.49 (d, J = 8.0 Hz, 2H), 7.40 (d, J = 8.0 Hz, 1H), 7.24 – 7.14 (m, 3H), 7.13 – 7.05 (m, 1H), 6.97 (s, 1H), 5.39 (d, J = 6.4 Hz, 1H), 4.97 (d, J = 13.6 Hz, 1H), 4.59 (d, J = 13.7 Hz, 1H), 4.17 (s, 1H), 3.65 (d, J = 16.0 Hz, 1H), 3.25 (dd, J = 15.6, 6.6 Hz, 1H), 3.06 (s, 3H). 13C NMR (75 MHz, DMF) δ 171.13, 168.46, 167.76, 166.10, 163.05, 162.66, 162.27, 145.67, 141.36, 137.70, 137.37, 132.55, 132.10, 130.43, 129.90, 129.48, 127.02, 122.25, 118.74, 107.62, 103.73, 83.79, 62.22, 53.70. HRMS (m/z): [MH]+ calculated for $C_{23}H_{19}ClN_2O_3$, 406.1084; found 406.1097.

1S3R

1H NMR (300 MHz, Aceton) δ 10.88 (s, 2H), 7.70 (t, J = 9.3 Hz, 6H), 7.60 (d, J = 8.1 Hz, 4H), 7.48 (d, J = 7.6 Hz, 2H), 7.30 – 7.06 (m, 4H), 6.46 (s, 2H), 5.47 (s, 2H), 4.78 (d, J = 13.5 Hz, 2H), 4.47 (d, J = 13.6 Hz, 2H), 3.96 (s, 2H), 3.76 (s, 7H), 3.59 (dd, J = 15.7, 4.8 Hz, 2H). 13C NMR (75 MHz, CDCl3) δ 168.69, 138.01, 132.60, 127.53, 127.04, 122.12, 119.66, 118.47, 111.98, 83.91, 79.54, 57.72, 52.44, 43.30, 23.82. HRMS (m/z): [MH]+ calculated for $C_{23}H_{19}ClN_2O_3$, 406.1084; found 406.1068.

1S3R PEG RSL3 (Compound 8)

11-Azido-3,6,9-trioxaundecan-1-amine (45uL, 0.23mmol) and anhydrous NEt_3 (42uL, 0.30mmol) was added to an oven-dried flask containing anhydrous CH_2Cl_2 and acetyl chloride (17uL, 0.24mmol) was added drop wise and allowed to stir for an hour. The resulting intermediate was concentrated and taken forward to the next step where the acetylated-PEG-azide (25mg, 0.12mmol) and 1S3R RSL3-alkyne probe 1 (30mg, 74umol) as described above were dissolved in DMF and sodium ascorbate (7.3 mg, 37umol) and copper sulphate pentahydrate (0.9mg, 3.7umol) dissolved in water was added to give a 50:50 solution. The reaction was run for an hour at 100°C in a microwave reactor. The compound was purified by silica gel chromatography using a gradient elution from 100% Hexanes to 100% EtOAc (Rf = 0.24 in EtOAc) with a yield of 23%.

1H NMR (500 MHz, D2O) δ 10.79 (s, 2H), 8.42 (s, 3H), 7.86 (s, 6H), 7.64 (s, 6H), 7.55 (s, 3H), 7.35 (s, 3H), 7.17 – 6.93 (m, 7H), 6.37 (s, 3H), 5.34 (s, 3H), 4.66 (d, J = 13.8 Hz, 10H), 4.34 (d, J = 13.4 Hz, 3H), 4.16 (s, 5H), 3.99 (s, 7H), 3.61 (d, J = 19.8 Hz, 47H), 3.46 (d, J = 13.9 Hz, 4H), 2.02 (s, 8H).

1S3R fluorescein-RSL3 probe 1(Compound 11)

Based on published protocol [39], 11-Azido-3,6,9-trioxaundecan-1-amine (63uL, 0.3mmol) was added to a solution of 5(6)-carboxyfluorescein N-succinimidyl ester (50mg, 0.1mmol) dissolved in anhydrous DMF (0.5mL). The reaction was stirred under N_2 overnight and then concentrated in vacuo. The crude product was then dissolved in dry pyridine (0.5mL) and isobutyric anhydride (~1mL, 1mmol) was added and allowed to stir at room temperature for 2 hours. Methanol (0.25mL) was added to the mixture and incubated for 5 minutes to quench unreacted isobutyric anhydride. The solvent was removed in vacuo and the crude product was purified by silica gel chromatography in 5:95 MeOH: CH_2Cl_2 (Rf = 0.28) to give the protected fluorescein azide with 73% yield over two steps.

1H NMR (500 MHz, D2O) δ 8.43 (s, 1H), 8.22 (d, J = 7.7 Hz, 1H), 8.11 (d, J = 6.0 Hz, 1H), 7.62 (s, 1H), 7.28 (d, J = 7.2 Hz, 1H), 7.10 (d, J = 6.5 Hz, 2H), 6.82 (d, J = 5.0 Hz, 4H), 3.72 (t, J = 5.9 Hz, 7H), 3.68 – 3.64 (m, 2H), 3.61 (dd, J = 13.2, 5.1 Hz, 2H), 3.56 – 3.47 (m, 2H), 3.41 – 3.33 (m, 1H), 3.32 – 3.24 (m, 1H), 2.82 (dt, J = 13.9, 7.0 Hz, 2H), 1.33 (d, J = 7.0 Hz, 12H). 13C NMR (126 MHz, CDCl3) δ 174.95, 168.29, 165.55, 155.19, 153.38, 152.54, 151.58, 151.48, 141.45, 137.03, 134.79, 129.44, 128.95, 128.78, 128.14, 126.53, 125.47, 124.49, 123.51, 122.66, 117.89, 117.84, 115.72, 115.67, 110.47, 110.42, 81.95, 81.85, 50.65, 50.59, 40.21, 40.13, 34.19, 18.83.

The protected fluorescein azide (30mg, 42umol) and 1S3R RSL3-alkyne probe 1 (20mg, 50umol), as described above, were dissolved in THF and sodium ascorbate (4.2 mg, 21umol) and copper sulphate pentahydrate (0.5mg, 2.1umol) dissolved in water was added. The reaction was stirred overnight at 40°C. The final product was purified using TLC preparatory plates using 5:95 MeOH: CH_2Cl_2 in 79% yield.

1H NMR (400 MHz, DMF) δ 10.83 (s, 2H), 8.62 (s, 1H), 8.58 (s, 1H), 8.51 (s, 1H), 8.45 – 8.35 (m, 3H), 8.32 (d, J = 8.4 Hz, 1H), 8.15 (d, J = 7.9 Hz, 1H), 7.92 (s, 1H), 7.83 (s, 4H), 7.62 (d, J = 7.7 Hz, 4H), 7.53 (d, J = 7.8 Hz, 3H), 7.35 – 7.26 (m, 6H), 7.12 – 7.05 (m, 2H), 7.03 (d, J = 5.5 Hz, 9H), 6.35 (s, 2H), 5.32 (s, 2H), 4.68 (s, 1H), 4.67 – 4.56 (m, 5H), 4.33 (d, J = 13.5 Hz, 2H), 3.94 (dt, J = 14.5, 5.3 Hz, 4H), 3.69 (d, J = 5.4 Hz, 3H), 3.61 (s, 21H), 3.57 – 3.52 (m, 4H), 3.52 – 3.39 (m, 7H), 2.89 – 2.82 (m, 3H), 1.32 (d, J = 7.0 Hz, 28H).

13C NMR (101 MHz, DMF) δ 174.44, 168.10, 165.25, 156.42, 153.03, 151.67, 151.53, 146.57, 139.29, 137.40, 134.89, 129.09, 127.86, 127.22, 126.56, 126.06, 125.63, 124.58, 124.06, 121.48, 119.06, 118.41, 117.90, 117.31, 116.49, 116.24, 111.41, 110.35, 78.52, 70.28, 69.32, 57.24, 57.11, 51.77, 49.88, 47.85, 42.87, 41.06, 39.95, 18.25, 15.67.

1S3R & 1R3R alkyne- RSL3 probe 2 (Compound 9)

Propargyl 4-formyl benzoate was prepared by stirring 4-formyl benzoic acid (500mg, 3.33mmol) and K_2CO_3 (507mg, 3.67mmol) in anhydrous DMF for 30min at 70°C. The reaction mixture was then cooled

to 0°C and propargyl bromide (1.15mL, 3.67mmol) was added dropwise. The mixture was again heated overnight until it became a clear yellow color. The reaction was filtered, concentrated and purified by silica gel column chromatography in 10:90 EtOAc:Hexane to give propargyl 4-formyl benzoate in 80% yield.

The pictet-spengler reaction was performed with (D) tryptophan methyl ester hydrochloride (638mg, 2.92mmol) and propargyl 4-formyl benzoate (500mg, 2.66mmol) as described above. The 1-propargyl ester RSL3 intermediate was purified by silica gel chromatography by gradient elution from 0:100 to 50:50 EtOAc:Hexane with 1% NEt_3 (Rf = 0.53 (1R3R), 0.35 (1S3R))

The 1-propargyl ester intermediates were chloroacetylated as described above and purified by gradient elution from 0:100 to 50:50 EtOAc:Hexane to give 1S3R and 1R3R RSL3-alkyne probe 2 in near quantitative yield.

1R3R: ^{1}H NMR (400 MHz, DMF) δ 10.90 (s, 1H), 8.02 (d, J = 7.9 Hz, 2H), 7.66 (d, J = 7.6 Hz, 1H), 7.43 (d, J = 7.5 Hz, 3H), 7.21 (d, J = 7.1 Hz, 1H), 7.13 (t, J = 7.3 Hz, 1H), 7.04 (s, 1H), 5.42 (s, 1H), 5.06 (s, 2H), 4.97 (d, J = 13.0 Hz, 1H), 4.61 (d, J = 13.0 Hz, 1H), 3.79 – 3.65 (m, 1H), 3.55 (s, 1H), 3.29 (dd, J = 15.8, 6.5 Hz, 1H). ^{13}C NMR (101 MHz, DME) δ 171.13, 167.90, 165.69, 146.12, 137.78, 130.30, 129.71, 129.53, 127.01, 122.51, 119.62, 118.82, 112.00, 107.71, 78.90, 77.46, 53.89, 53.01, 52.67, 52.02, 43.76, 42.03, 22.09.

1S3R: ^{1}H NMR (400 MHz, DMF) δ 10.83 (s, 1H), 7.99 (d, J = 7.9 Hz, 2H), 7.75 (d, J = 7.9 Hz, 2H), 7.57 (d, J = 7.6 Hz, 1H), 7.35 (d, J = 7.9 Hz, 1H), 7.09 (dt, J = 21.2, 7.2 Hz, 3H), 6.40 (s, 1H), 5.43 (s, 1H), 5.01 (d, J = 1.9 Hz, 2H), 4.69 (d, J = 13.6 Hz, 1H), 4.38 (d, J = 13.5 Hz, 1H), 3.70 (d, J = 15.3 Hz, 1H), 3.52 (d, J = 11.8 Hz, 1H). ^{13}C NMR (101 MHz, DMF) δ 178.32, 171.54, 168.29, 165.14, 137.54, 133.66, 129.72, 128.65, 126.97, 126.47, 121.71, 119.21, 118.03, 111.51, 78.35, 76.36, 57.25, 52.28, 52.07, 42.83, 23.41.

1S3R alkyne-RSL3 probe 3 (Compound 10)

N-Boc protected (D) tryptophan (500mg, 1.64mmol), EDC (470mg, 2.46mmol) and DMAP (40mg, 0.34mmol) were dried under vacuum for an hour. In another flask propargyl alcohol (117uL, 1.97mmol) and a few drops of anhydrous NEt_3 were dissolved in anhydrous DMF. The tryptophan mixture was also dissolved in anhydrous DMF, added dropwise to the propargyl alcohol mixture and allowed to stir overnight. The propargyl ester (D) N-boc tryptophan was purified by silica gel column chromatography in 10:90 EtOAc:Hexanes. After characterization, the compound was dissolved in 30% TFA/ CH_2Cl_2 and allowed to stir overnight and then concentrated without further purification to give the unprotected propargyl ester (D) tryptophan with 75% yield.

The pictet-spengler reaction was performed with propargyl ester (D) tryptophan (300mg, 1.23mmol), methyl 4-formyl benzoate (324mg, 1.97mmol) and activated molecular sieves as described above. The 3-propargyl ester RSL3 intermediate was purified by column chromatography in 30:70 EtOAc:Hexane with 1% NEt_3 to give 1S3R and 1R3R 3-propargyl ester intermediate with yields of 13% and 23% respectively.

The 3-propargyl ester intermediates were chloroacetylated as described above and purified by gradient elution from 0:100 to 50:50 EtOAc:Hexane to give 1S3R and 1R3R RSL3-alkyne probe 3 in near quantitative yield.

1R3R: 1H NMR (400 MHz, CDCl3) δ 8.39 (s, 1H), 7.86 (d, J = 7.8 Hz, 2H), 7.63 (d, J = 7.7 Hz, 1H), 7.35 (d, J = 7.9 Hz, 1H), 7.33 – 7.28 (m, 2H), 7.22 (dtd, J = 20.6, 7.3, 1.0 Hz, 2H), 6.97 (s, 1H), 5.00 (d, J = 5.7 Hz, 1H), 4.40 (d, J = 12.7 Hz, 1H), 4.32 – 4.15 (m, 2H), 3.89 (s, 3H), 3.71 (d, J = 15.9 Hz, 1H), 3.52 (d, J = 15.3 Hz, 1H), 3.26 (ddd, J = 15.9, 6.9, 1.4 Hz, 1H), 2.35 (t, J = 2.3 Hz, 1H). 13C NMR (101 MHz, CDCl3) δ 167.24, 166.71, 143.76, 136.58, 129.91, 129.54, 129.42, 128.65, 126.22, 122.79, 119.95, 118.65, 111.28, 107.54, 76.38, 75.64, 53.73, 52.77, 52.29, 51.88, 42.03, 21.68.

1S3R: 1H NMR (400 MHz, DMF) δ 11.27 (s, 4H), 8.04 (t, J = 13.5 Hz, 8H), 7.85 (d, J = 7.0 Hz, 6H), 7.73 (d, J = 7.7 Hz, 4H), 7.46 (d, J = 7.3 Hz, 4H), 7.35 – 7.11 (m, 8H), 6.41 (s, 3H), 5.83 (s, 3H), 5.05 (d, J = 13.9 Hz, 3H), 4.89 (t, J = 11.8 Hz, 8H), 4.73 (d, J = 13.9 Hz, 3H), 4.06 (d, J = 19.5 Hz, 13H), 3.92 (d, J = 15.0 Hz, 3H), 3.77 (d, J = 12.9 Hz, 3H), 3.69 (s, 5H). 13C NMR (101 MHz, DMF) δ 168.31, 166.45, 137.66, 133.85, 129.68, 126.96, 126.63, 121.79, 119.28, 118.17, 111.59, 77.81, 76.66, 57.40, 52.94, 51.67, 42.91, 23.50.

1S3R fluorescein-RSL3 probe 3 (Compound 13)

Protected fluorescein azide (30mg, 42umol) and 1S3R RSL3-alkyne probe 3 (23mg, 50umol), as described above, were dissolved in THF and sodium ascorbate (4.2 mg, 21umol) and copper sulphate pentahydrate (0.6 mg, 2.1umol) dissolved in water was added. The reaction was stirred overnight at 40°C. The final product was purified using TLC preparatory plates run in 5:95 MeOH: CH_2Cl_2 in 89% yield.

1H NMR (400 MHz, DMF) δ 10.85 (s, 1H), 8.64 (s, 1H), 8.59 (s, 1H), 8.52 (s, 1H), 8.39 (d, J = 7.6 Hz, 13H), 8.32 (d, J = 8.3 Hz, 1H), 8.15 (d, J = 7.5 Hz, 1H), 7.92 (d, J = 8.2 Hz, 2H), 7.68 (s, 2H), 7.60 (s, 1H), 7.52 (dd, J = 16.9, 7.9 Hz, 2H), 7.30 (s, 3H), 7.08 (d, J = 10.3 Hz, 1H), 7.04 (d, J = 2.0 Hz, 113H), 6.34 (s, 1H), 5.46 (s, 1H), 5.16 (d, J = 4.0 Hz, 2H), 4.67 (d, J = 13.3 Hz, 1H), 4.47 – 4.25 (m, 3H), 3.88 (s, 2H), 3.85 – 3.76 (m, 2H), 3.73 (d, J = 5.6 Hz, 1H), 3.69 – 3.56 (m, 173H), 3.53 (s, 6H), 2.90 – 2.84 (m, 2H), 1.33 (d, J = 6.9 Hz, 12H). MS (m/z): [MH]+ calculated for $C_{62}H_{61}ClN_6O_{16}$, 1181.63; found 1181.6.

1S3R & 1R3R fluorescein-RSL3 probe 2 (Compound 12)

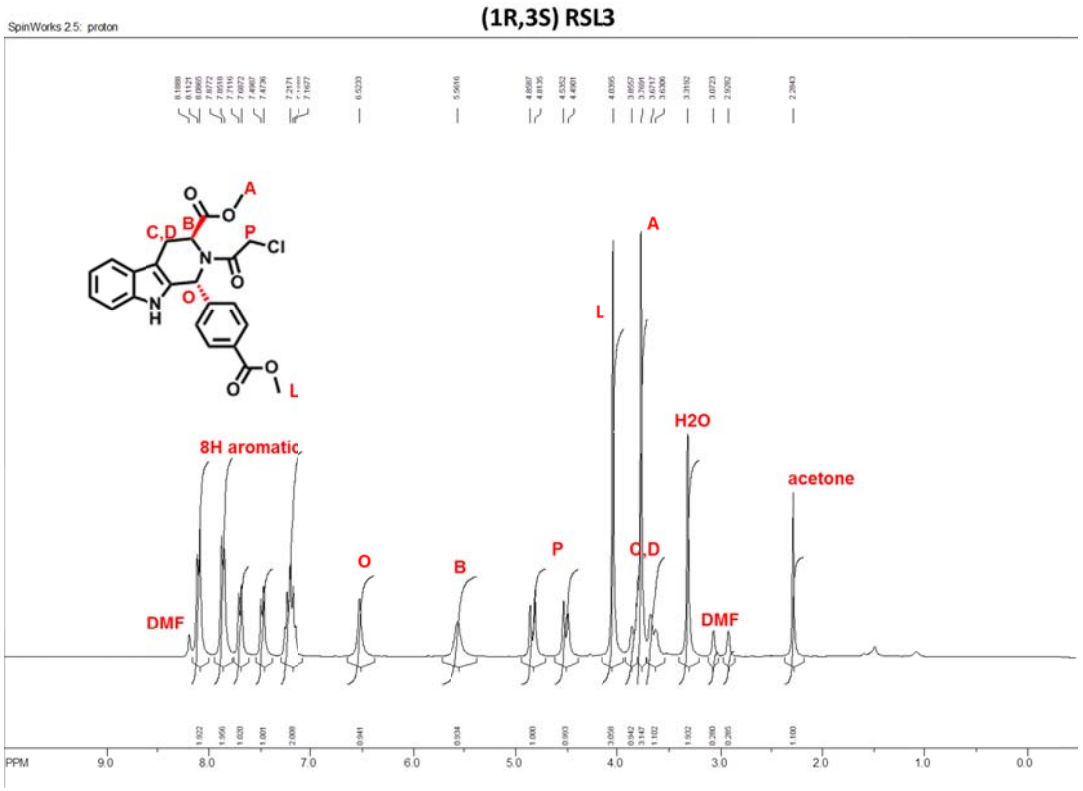
Protected fluorescein azide (20mg, 28umol) and either 1S3R or 1R3R RSL3-alkyne probe 2 (15mg, 34umol) as described above, were dissolved in THF and sodium ascorbate (2.8 mg, 14umol) and copper sulphate pentahydrate (0.4 mg, 1.4umol) dissolved in water was added. The reaction was stirred overnight at 40°C. The final product was purified using TLC preparatory plates run in 5:95 MeOH: CH_2Cl_2 in 83% yield.

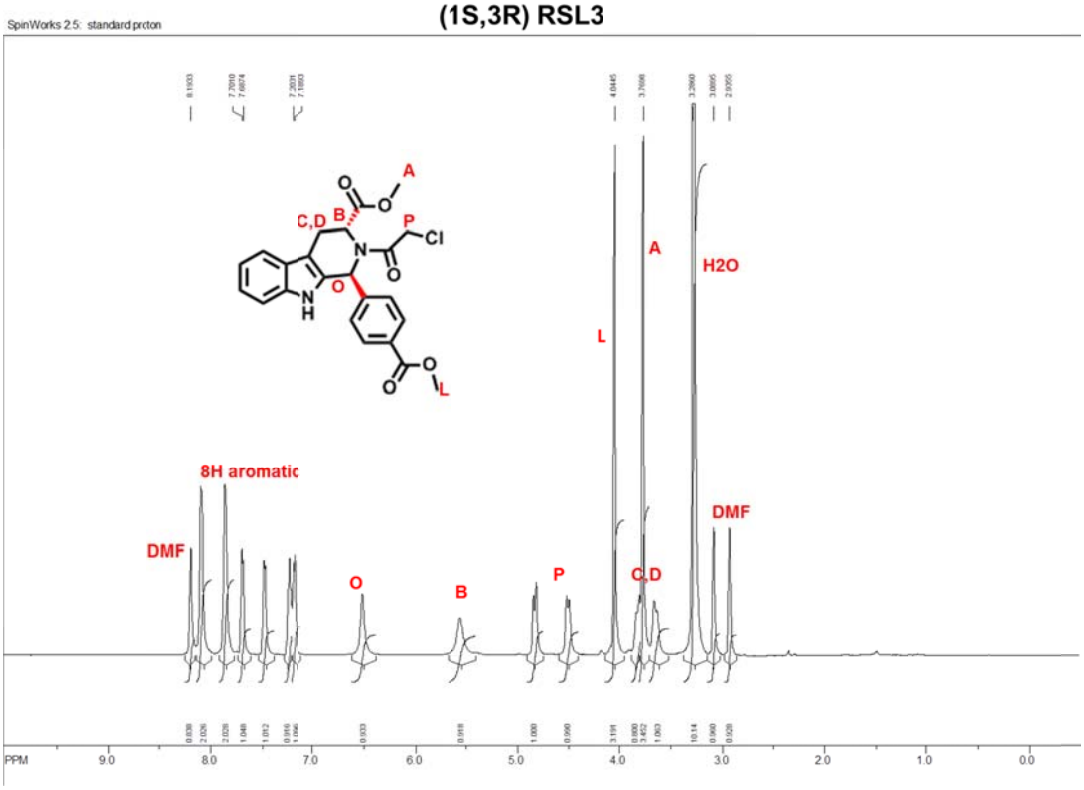
1S3R fluorescein-RSL3: 1H NMR (400 MHz, DMF) δ 10.81 (s, 1H), 8.63 (s, 1H), 8.59 (s, 1H), 8.51 (s, 1H), 8.38 (d, J = 6.6 Hz, 1H), 8.32 (d, J = 8.0 Hz, 1H), 8.17 (d, J = 9.5 Hz, 1H), 7.92 (d, J = 5.6 Hz, 3H), 7.68 (d, J = 7.5 Hz, 2H), 7.53 (dd, J = 7.9, 4.1 Hz, 2H), 7.30 (d, J = 2.0 Hz, 3H), 7.12 – 6.94 (m, 6H), 6.33 (s, 1H), 5.44 (d, J = 5.1 Hz, 3H), 4.67 (d, J = 13.5 Hz, 1H), 4.59 (dt, J = 10.9, 5.4 Hz, 2H), 4.35 (d, J = 13.3 Hz, 1H), 3.91 (dt, J = 14.9, 5.4 Hz, 2H), 3.71 (t, J = 5.6 Hz, 2H), 3.68 – 3.53 (m, 12H), 3.52 – 3.39 (m, 4H), 2.89 (dd, J = 14.0, 7.0 Hz, 2H), 1.32 (d, J = 7.0 Hz, 12H). 13C NMR (101 MHz, DMF) δ 174.56, 171.70, 168.12, 154.4, 129.95, 128.28, 124.64, 123.78, 120.02, 114.84, 110.42, 72.18, 69.94, 68.68, 52.17, 50.82, 46.81, 43.35, 41.34, 32.59, 22.35,18.51. HRMS (m/z): [MH]+ calculated for $C_{62}H_{61}ClN_6O_{16}$, 1181.3911; found 1182.4039.

1R3R fluorescein-RSL3: 1H NMR (300 MHz, DMF) δ 8.68 (d, J = 10.0 Hz, 1H), 8.36 (s, 3H), 8.17 (dd, J = 8.0, 1.5 Hz, 3H), 8.10 (dd, J = 8.0, 1.3 Hz, 2H), 7.96 (d, J = 8.0 Hz, 2H), 7.89 – 7.80 (m, 2H), 7.77 (s, 2H), 7.68 (s, 2H), 7.66 – 7.55 (m, 1H), 7.41 – 7.29 (m, 2H), 7.23 – 7.14 (m, 2H), 7.12 (d, J = 5.8 Hz, 1H), 7.07 – 7.03 (m, 1H), 6.81 (t, J = 2.9 Hz, 1H), 6.79 (d, J = 1.2 Hz, 1H), 6.77 (dd, J = 7.5, 2.3 Hz, 1H), 5.40 (d, J = 20.6 Hz, 2H), 4.98 (s, 1H), 4.61 – 4.50 (m, 1H), 4.48 – 4.30 (m, 1H), 4.22 (dd, J = 12.6, 3.6 Hz, 1H), 3.94 – 3.85 (m, 1H), 3.84 – 3.75 (m, 1H), 3.70 (d, J = 15.7 Hz, 1H), 3.65 – 3.53 (m, 7H), 3.53 – 3.36 (m, 4H), 3.24 (dd, J = 15.9, 6.8 Hz, 1H), 3.04 (s, 2H), 2.92 – 2.71 (m, 2H), 1.32 (t, J = 10.6 Hz, 12H). 13C NMR (101 MHz, DMF) δ 174.8, 173.06, 171.70, 168.03, 154.45, 143.61, 137.84, 135.35, 130.81, 129.19, 124.64, 120.02, 114.84, 110.42, 69.94, 52.17, 48.71, 46.81, 41.34, 25.04, 19.21. HRMS (m/z): [MH]+ calculated for $C_{62}H_{61}ClN_6O_{16}$, 1181.3911; found 1181.3926.

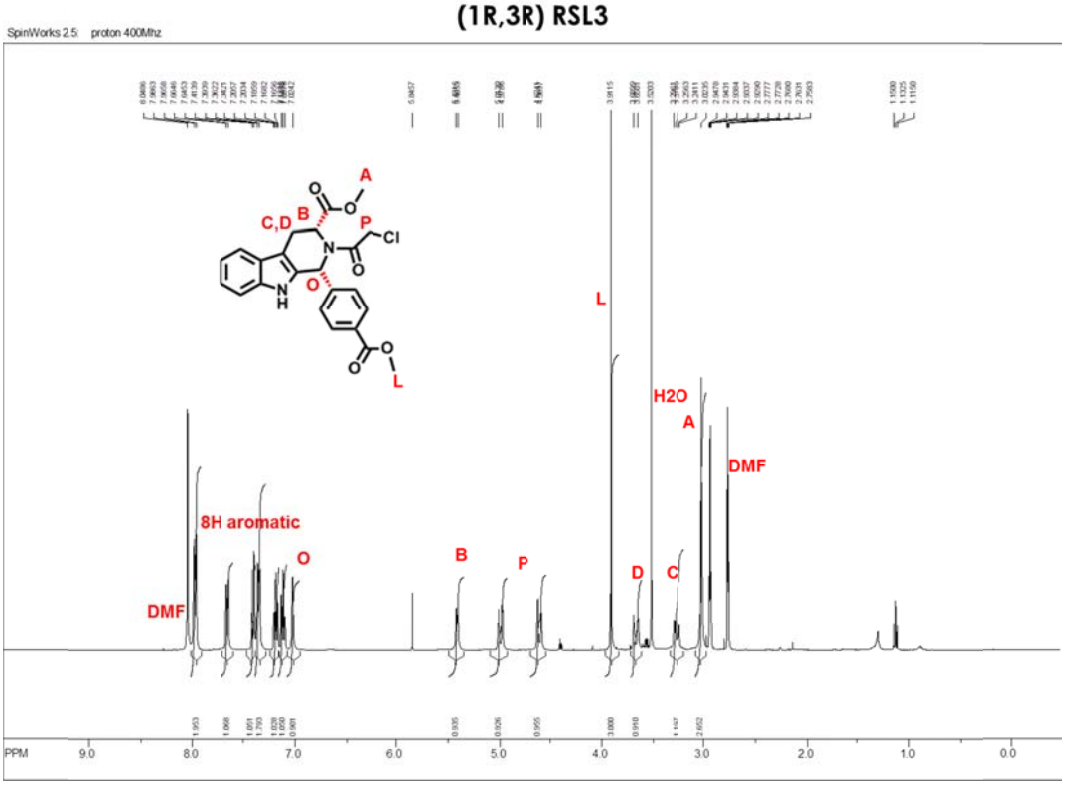
2.4c NMR spectra

Four diastereomers of RSL3

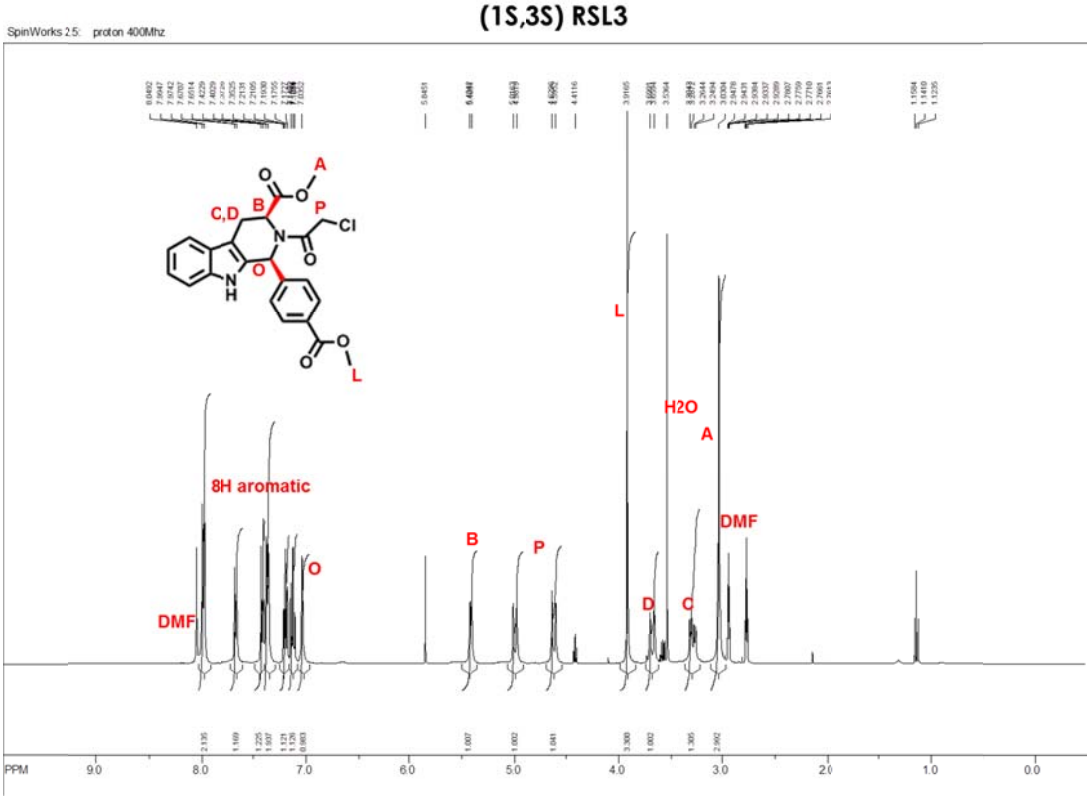




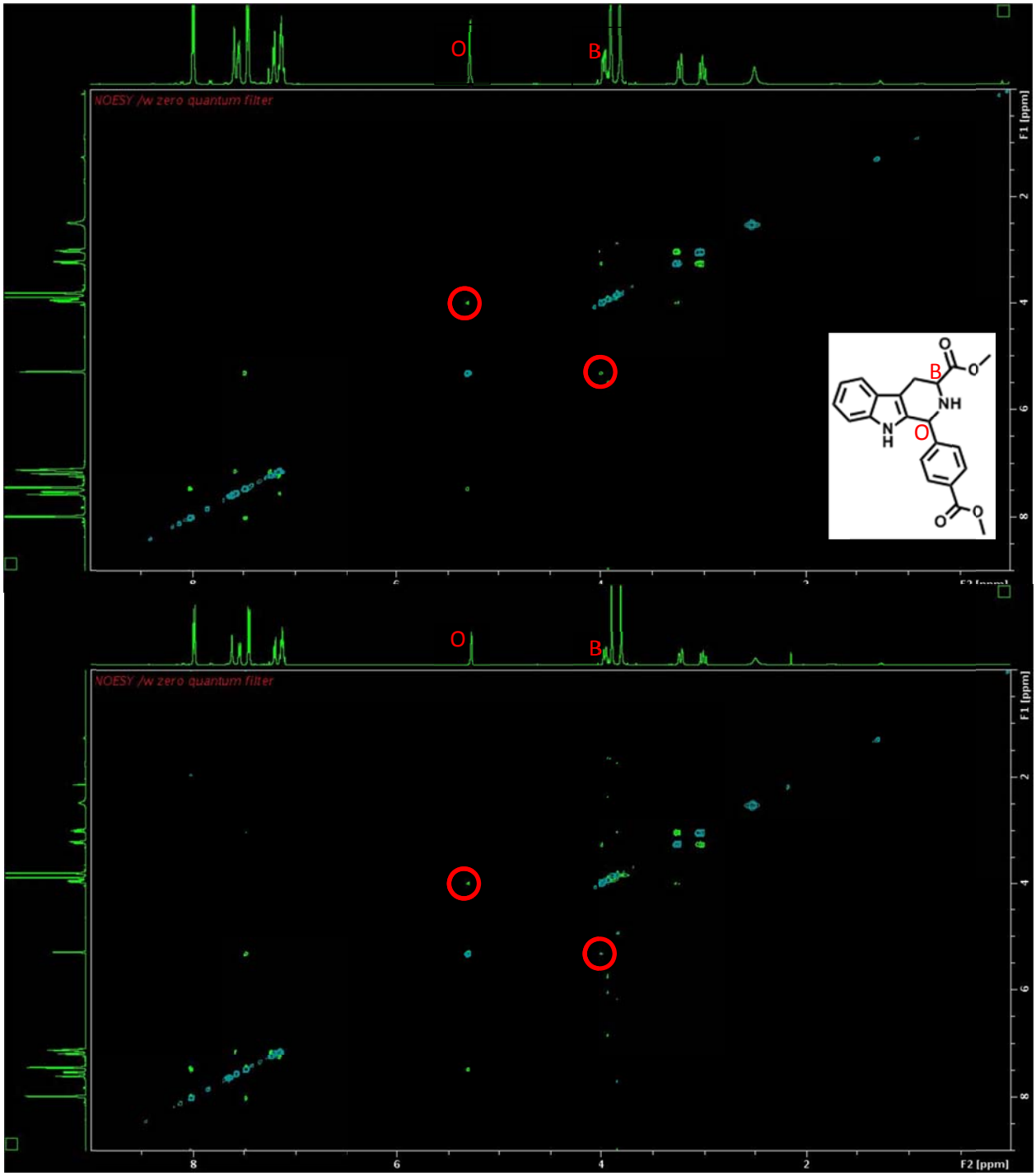
(1R,3R) RSL3



(1S,3S) RSL3

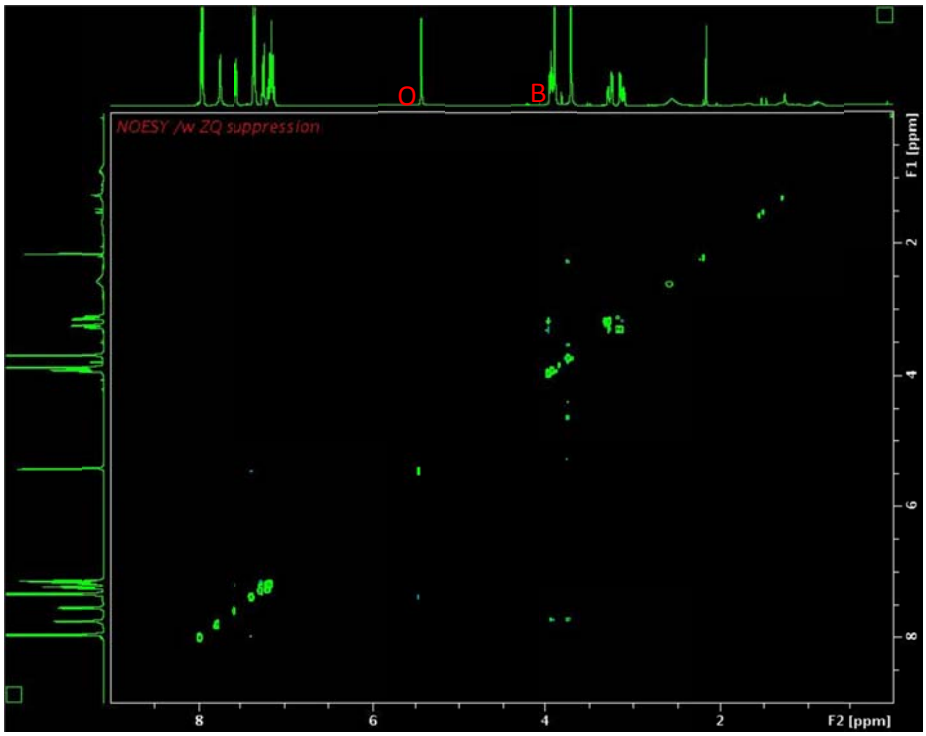


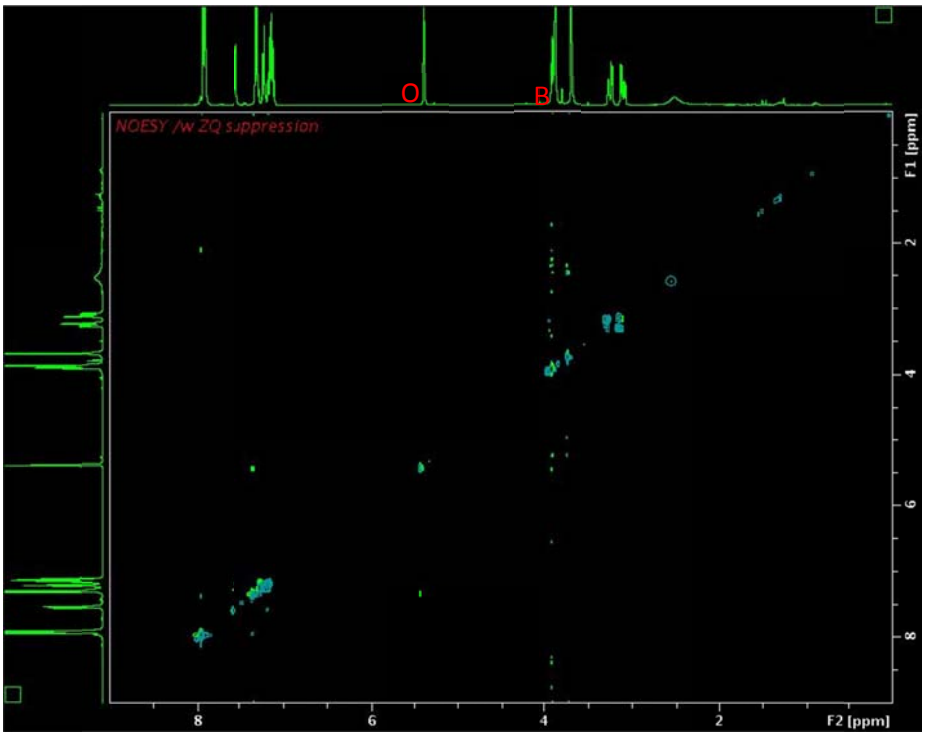
NOESY spectra of (1R,3R) and (1S,3S) RSL3 intermediates



In RSL3's structure two protons exist at either chiral center, described as 'O' and 'B' in the structures depicted above. In two of the diastereomers, a NOESY cross peak is observed between the probes that is absent in the other two diastereomers. Based on these data, we assigned the above two spectra to the 'cis' configuration of the molecule, (1S,3S) and (1R,3R) and the spectra on the next page.

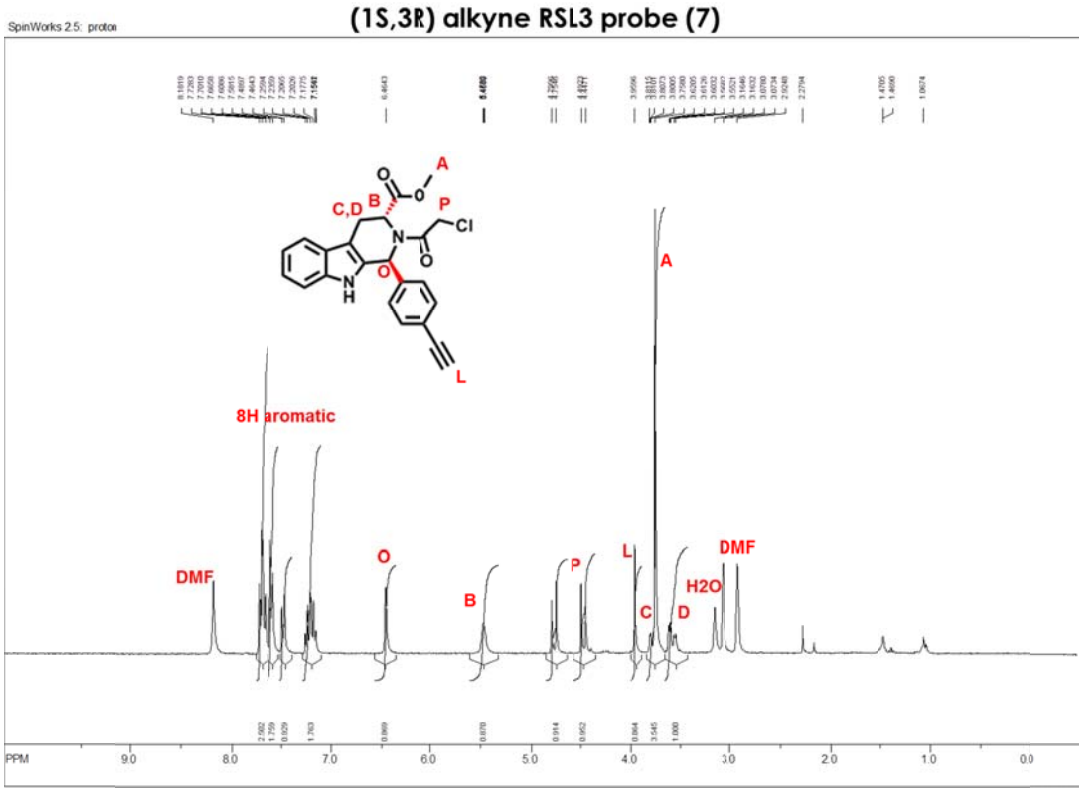
NOESY spectra of (1R,3S) and (1S,3R) RSL3 intermediates



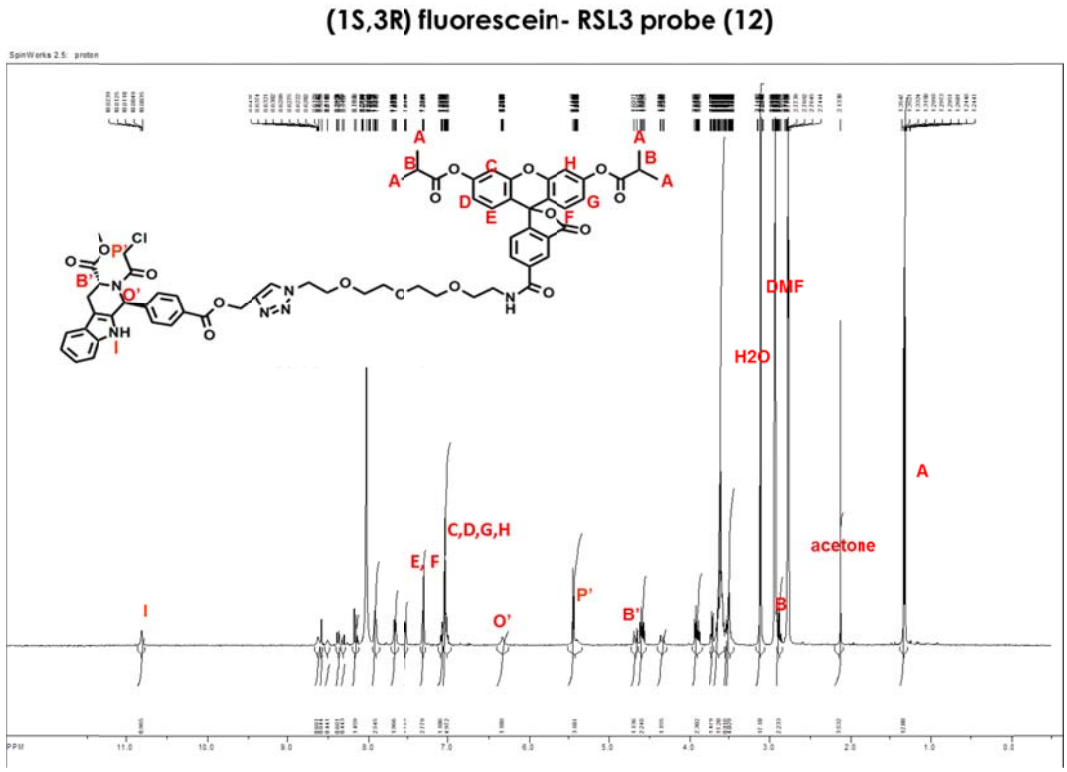


In the ‘trans’ diastereomers, no NOESY cross peak is observed between protons ‘O’ and ‘B’.

1S3R RSL3-alkyne probe 1(Compound 6)



1S3R fluorescein probe 2 (Compound 12)



2.4d Testing compounds in four BJ cell lines

Cells from four BJ cell lines were seeded the night before the experiment in 384-well format at 1000 cells/well in 32 ul of media. Compounds were dissolved in anhydrous DMSO to a stock concentration of 10mg/mL. Stocks were diluted into media to a concentration of 50ug/mL and further diluted to 10ug/mL upon transfer to the assay plate. Assay plates were incubated for 24 hours and viability was assessed using alamar blue. Viability measurements are described in Chapter 4.

2.4e Mechanism studies of RSL3 probes

BJeLR cells were seeded in 384-well format at 2000 cells/well in 32 ul of media with or without the addition of a set concentration of inhibitor. Compounds were dissolved in anhydrous DMSO to a stock concentration of 10mg/mL. Stocks were diluted into media to a concentration of 50ug/mL and further diluted to 10ug/mL upon transfer to the assay plate. Assay plates were incubated for 24 hours and viability was assessed using alamar blue. Viability measurements are described in Chapter 4. Necrostatin 1 and BHA (Butylated hydroxyanisole) were purchased from Sigma and U0126 (MEK inhibitor) was purchased from LC laboratories. Necrostatin 1 and U0126 (MEK inhibitor) were used at a concentration of 5ug/mL and BHA was used at 25ug/mL.

2.4f Pulldown sample preparation with alkyne probe

Lysate preparation

Cells were seeded into T175 flasks the day before with four million cells in order to be confluent on the day of treatment. The following day three flasks were first washed with PBS three times to remove serum proteins from the media and then treated with either 1.23uM 1S3R RSL3-alkyne probe (hereafter referred to as 'probe') or 1.23uM 1R3R RSL3-alkyne probe (hereafter referred to as 'inactive') in serum-free media (DMEM) followed by incubation for an hour. For the competitor-treated samples, three flasks of cells were treated with 12.3uM (10X) of 1S3R RSL3 (hereafter referred to as 'competitor') for 20

min prior to treatment with 1.23uM of probe for an hour. (The probe and inactive samples were treated with the equivalent amount of DMSO for the time period of the competitor pretreatment) After treatment, the cells were washed three times with PBS, trypsinized and pelleted at 1000 rpm. Following trypsinization all following steps were performed at 4°C. The pelleted cells were again washed in PBS and pelleted twice to remove serum proteins added during the trypsinization. Cells were resuspended in a 10mM potassium phosphate buffer (PB) for 3-5 min on ice and then lysed by 30-40 rapid strokes with a dounce homogenizer. The resulting lysate was centrifuged at 12,000 g for 15 min to pellet out insoluble materials and the supernatant was removed. Protein concentration was determined using the Bradford assay (Biorad) and the lysate was diluted to 2mg/mL protein concentration using PB buffer.

Click chemistry copper-catalyzed cycloaddition reactions

This protocol is adapted from that published by Cravatt lab and via personal communication [33]. Click chemistry was performed on the samples by the addition of 18.75 uM fluorescein-azide, 1 mM TCEP (tris(2-carboxyethyl)phosphine), 100 uM TBTA(tris[(1-benzyl-1H -1,2,3-triazol-4-yl)methyl]amine) and 1 mM $CuSO_4$ for 1 hour at room temperature. Aliquots (15 ul) from each reaction were removed for analysis and mixed with 5x SDS loading buffer and run out by SDS-PAGE and transferred to a nitrocellulose membrane using an Invitrogen iBlot system. Membranes were probed with an anti-fluorescein/ Oregon green HRP conjugate rabbit IgG antibody (Invitrogen, A21253) overnight at 4°C, and developed with SuperSignal West Pico Substrate (Pierce, 34080). The remaining sample was pelleted at 6500g at 4°C and the supernatant was run through a Zeba Desalt Spin Column (Pierce, 89892 prepared according to manufacturer's directions) to remove excess cycloaddition reagents from protein that did not precipitate. The pelleted protein was resuspended in cold methanol, incubated on a rotator at 4°C for 10min and then pelleted again (this was repeated twice - also to remove excess cycloaddition reagents). The pellet was resuspended in 1.2% SDS/PBS. The lysate is then denatured to allow

accessibility to the fluorescein tag on the protein by heating at 80-90°C and then diluted 5 fold to 0.2% SDS (at this point supernatant from before is added back in).

Pre-clearing and pulldown

Anti-fluorescein beads were prepared by coupling NHS-activated Sepharose 4 Fast Flow matrix (GE Healthcare, 17-0906-01) to mouse IgG anti-FITC (Jackson Labs, 200-002-037) according to the manufacturer's instructions. For use in pre-clearing, "negative" beads were prepared that had the NHS-active site blocked with ethanolamine. The bead preparations were first tested for their ability to affinity-purify a positive control protein prior to use (FITC-tagged goat anti-rabbit IgG (Jackson labs, 111-095-003)).

Lysate was first incubated with packed negative beads (~60uL packed beads) for an hour at 4˚C on a rotator to remove non-specific interactions with the resin alone. Next, the lysate was then incubated with anti-fluorescein beads (~120uL) overnight at 4˚C on a rotator. The supernatant was removed the next morning and the remaining beads were washed twice in 0.2% SDS in PBS and once in PBS. The beads were then eluted with 120uL of 0.1mM 5(6) carboxyfluorescein in PBS for an hour at 4˚C on a rotator to specifically elute fluorescein-tagged proteins. This was followed by a denaturing elution with 2% SDS and boiling to remove remaining bound proteins for comparison. The fluorescein-eluted proteins were then precipitated using standard chloroform-methanol precipitation and the final pellet was redissolved in Rapigest buffer (Waters Corp, 186001861 -prepared according to manufacturer's directions) and 1.2uL of DTT followed by vortexing, sonification and boiling for 5 min each. The protein concentration was determined by the Bradford assay and samples were snap-frozen in liquid nitrogen and stored at -80°C until MS analysis.

Chromatography, mass spectrometry and data analysis

Prior to MS analysis, cysteines in the samples were alkylated with iodoacetamide. Proteins were digested with trypsin (6 ng/µL Promega Corp, #V511A in 50mM NH_4HCO_3) and 50 fmol of a digest of yeast alcohol dehydrogenase was added as an internal detection control (also can be used for absolute quantification). Three chromatograms were recorded for each of nine biological replicates (three probe, three inactive, three competitor), yielding 27 chromatograms. Prior to analytical separation on a NanoAcquity UPLC (Waters Corp.), peptides were trapped on a Symmetry C18 Trap column, 5 µm particles, 180 µm × 20 mm (Waters Corp.), for 1 min at 15 µL/minute in 1% solvent B (0.3% formic acid in acetonitrile)/99% solvent A (0.3% formic acid, aqueous). Peptides were analyzed in a 120-min chromatogram on a 75 µm ID x 10 cm reverse phase 1.7 µm particle diameter bridged ethyl hybrid (BEH) C18 column at a flow rate of 300 nL/minute. For the analytical separation, Solvent B was increased in a 90-min linear gradient between 3 and 40%, and postgradient cycled to 95% B for 7 min, followed by postrun equilibration at 3% B. Spectra were recorded in V-positive mode with a Synapt quadrupole-time-of-flight mass spectrometer (Waters Corp). Source settings included extraction cone at 3.5 V, sampling cone at 24 V, and source temperature 80 °C. Collision energy was held at 6 V for low energy scan and ramped from 15 to 35 V for the high energy scan with a collision gas flow (Ar) of 1.5 mL/minute. Alternate 0.6 s scans at low and high energy were recorded for the range between 100 and 1990 *m*/*z*. A reference sprayer was operated at 300nL/minute to produce a lockmass spectrum with Glu-1-Fibrinopeptide B (*m*/*z* 785.8426) every 30 s.

Spectra were analyzed with ProteinLynx Global Server (Vers. 2.4, RC7) (Waters Corp.) and searched against a database of human protein sequences (reviewed canonical sequences with isoforms) from UniProt release 15.5 (July 7, 2009). The database also contained sequences for yeast alcohol dehydrogenase, porcine trypsin, bovine serum proteins (BSA, serotransferrin, fibrinogen alpha chain, fibrinogen beta chain and fibrinogen gamma-B). This database was comprised of 34602 protein

sequences (20,005,831 amino acid residues). Data mining and statistical analysis was performed with the Elucidator Protein Expression Data Analysis System, Version 3.3 (3.3.0.1.SP3_CRE52.21) (Rosetta Biosoftware). Proteins were identified with a PeptideTeller predicted error (false positive rate) of 1% and a calculated decoy error rate of 0.2%. Ratio P-values for differential expression were calculated by the Elucidator program using the xdev parameter [40, 41]. The P-values calculated are not from an analysis of variance, but instead derive from an application of an error model developed for large scale microarray data as adapted for proteomics within the Elucidator program. All of the above was performed by Dr. Lewis Brown and staff of the Comparative Proteomics Center.

2.4g Pulldown sample preparation with fluorescein probe

Lysate preparation

Cells were seeded into T225 flasks beforehand with ~1.5 million cells in order to be confluent on the day of treatment. Three days later, twenty flasks were first washed with PBS three times to remove serum proteins from the media and then treated with either 0.5uM 1S3R RSL3-fluorescein probe (hereafter referred to as 'probe') or 0.5uM 1R3R RSL3-fluorescein probe (hereafter referred to as 'inactive') in serum-free media (DMEM) followed by incubation for two hours. For the competitor-treated samples, twenty flasks of cells were treated with either 50nM or 500nM (0.1X or 1X) of 1S3R RSL3 (hereafter referred to as 'competitor 50' or 'competitor 500') for 20 min prior to treatment with 0.5uM of probe for two hours. (The probe and inactive samples were treated with the equivalent amount of DMSO for the time period of the competitor pretreatment) After treatment, the cells were washed once with PBS, trypsinized and pelleted at 1000 rpm. Following trypsinization all following steps were performed at 4°C. The pelleted cells were again washed in PBS twice to remove serum proteins added during the trypsinization. Cells were resuspended in non-denaturing lysis buffer for 20 min (50mM HEPES, 40mM NaCl, 2mM EDTA, 0.5% Triton-X, 1.5mM Na_3VO_4, 50mM NaF, 10mM Na-pyrophosphate, 10mM Naβ-

glycerophosphate, Roche protease inhibitor tablet). The resulting lysate was centrifuged at 12,000 g for 15 min to pellet out insoluble materials and the supernatant was removed. Protein concentration was determined using the Bradford assay (Biorad). Dr. Hemant Varma, Dr. Gisun Park, Dr. Andras Bauer, Dr. Adam Wolpaw, Dr. Wan Seok Yang, Dr. Scott Dixon and Dr. Reka Letso kindly aided in the trypsinization and collection of over 240 flasks of cells for this experiment.

Pre-clearing and pulldown

Protocol and buffer recipes gratefully adapted from protocol published by MacKinnon and Taunton [42]. Anti-fluorescein beads and negative beads were prepared as described above. Lysate was diluted two-fold with 'affinity purification' buffer (50mM HEPES, 100mM NaCl, 1% Triton X-100, pH 7.4) and first incubated with negative beads (~0.8mL of packed beads per treatment sample) for an hour at 4˚C on a rotator to remove non-specific interactions with the resin alone. Next, the lysate was then incubated with anti-fluorescein beads (~1.6mL) overnight at 4˚C on a rotator. The supernatant was removed the next morning and the remaining beads were washed twice in 'affinity purification' buffer (this involved resuspension of the beads and rotating at 4°C for 10 min followed by centrifugation at 10,000rpm). The beads were then washed similarly in 'wash' buffer (50mM HEPES, 500mM NaCl, 1% Triton X-100, pH 7.4) twice. The beads were then eluted with 750uL of saturated 5(6)-carboxyfluorescein in PBS (137mM NaCl, 10mM Na2HPO4, 2.7mM KCl) for an hour at 4˚C on a rotator to specifically elute fluorescein-tagged proteins, this was repeated twice more. Finally a denaturing elution was performed with 2% SDS and boiling to remove remaining bound proteins for comparison. The fluorescein elutions were combined and precipitated using standard chloroform-methanol precipitation and the final pellet was redissolved in Rapigest buffer (Waters Corp, 186001861 - prepared according to manufacturer's directions) and 1.2uL of DTT followed by vortexing, sonification and boiling for 5 min each. The protein concentration was determined by the Bradford assay and samples were snap-frozen in liquid nitrogen

and stored at -80°C until MS analysis. Three biological replicates were prepared for each treatment condition.

Chromatography, mass spectrometry and data analysis

Prior to MS analysis, cysteines in the samples were alkylated with iodoacetamide. Proteins were digested with trypsin (6 ng/µL Promega Corp, #V511A in 50mM NH_4HCO_3) and 50 fmol of a digest of yeast alcohol dehydrogenase was added as an internal detection control. Three chromatograms were recorded for each of twelve biological replicates (three probe, three inactive, three competitor at 50 nM, and three competitor 500 nM), yielding 36 chromatograms (35 analyzed in the study). Prior to analytical separation on a NanoAcquity UPLC (Waters Corp.), peptides were trapped on a Symmetry C18 Trap column, 5 µm particles, 180 µm × 20 mm (Waters Corp.), for 2 min at 10 µL/minute in 1% solvent B (0.3% formic acid in acetonitrile)/99% solvent A (0.3% formic acid, aqueous). Peptides were analyzed in a 120-min chromatogram on a 75 µm ID x 10 cm reverse phase 1.7 µm particle diameter bridged ethyl hybrid (BEH) C18 column at a flow rate of 300 nL/minute. For the analytical separation, Solvent B was increased in a 90-min linear gradient between 5 and 40%, and postgradient cycled to 95% B for 7 min, followed by postrun equilibration at 5% B. Spectra were recorded in resolution positive mode with a Synapt G2 quadrupole-time-of-flight mass spectrometer (Waters Corp). Source settings included extraction cone at 2 V, sampling cone at 30 V, and source temperature 80 °C. Collision energy was held at 4 V for low energy scan and ramped from 15 to 45 V for the high energy scan with a collision gas flow (Ar) of 2 mL/minute. Alternate 0.6 s scans at low and high energy were recorded for the range between 50 and 2000 *m*/*z*. A reference sprayer was operated at 300nL/minute to produce a lockmass spectrum with Glu-1-Fibrinopeptide B (*m*/*z* 785.8426) every 30 s.

Spectra were analyzed with ProteinLynx Global Server (PLGS) (Vers. 2.4, RC7) (Waters Corp.) and searched against a database of human protein sequences (reviewed canonical sequences with isoforms)

from UniProt release **2011_04** (**April 5, 2011**). The database also contained sequences for yeast alcohol dehydrogenase, porcine trypsin, bovine serum proteins (BSA, serotransferrin, fibrinogen alpha chain, fibrinogen beta chain and fibrinogen gamma-B). This database was comprised of 35,353 protein sequences (20,652,583 amino acid residues). Data mining and statistical analysis was performed with the Elucidator Protein Expression Data Analysis System, Version 3.3 (3.3.0.1.SP3_CRE52.21) (Rosetta Biosoftware). Proteins included in the study were identified with a PLGS predicted error (false positive rate) of 4% a minimum PLGS Protein Score of 75 and were also represented by two or more unique identified peptides. Ratio P-values for differential expression were calculated as described above. All of the above was performed by Dr. Lewis Brown and staff of the Comparative Proteomics Center.

2.5 References

1. Harvey, J.J., *An Unidentified Virus Which Causes the Rapid Production of Tumours in Mice.* Nature, 1964. **204**: p. 1104-5.
2. Kirsten, W.H. and L.A. Mayer, *Morphologic responses to a murine erythroblastosis virus.* J Natl Cancer Inst, 1967. **39**(2): p. 311-35.
3. Malumbres, M. and M. Barbacid, *RAS oncogenes: the first 30 years.* Nat Rev Cancer, 2003. **3**(6): p. 459-65.
4. DeFeo, D., M.A. Gonda, H.A. Young, E.H. Chang, D.R. Lowy, E.M. Scolnick, and R.W. Ellis, *Analysis of two divergent rat genomic clones homologous to the transforming gene of Harvey murine sarcoma virus.* Proc Natl Acad Sci U S A, 1981. **78**(6): p. 3328-32.
5. Ellis, R.W., D. DeFeo, M.E. Furth, and E.M. Scolnick, *Mouse cells contain two distinct ras gene mRNA species that can be translated into a p21 onc protein.* Mol Cell Biol, 1982. **2**(11): p. 1339-45.
6. Der, C.J., T.G. Krontiris, and G.M. Cooper, *Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses.* Proc Natl Acad Sci U S A, 1982. **79**(11): p. 3637-40.
7. Parada, L.F., C.J. Tabin, C. Shih, and R.A. Weinberg, *Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene.* Nature, 1982. **297**(5866): p. 474-8.
8. Santos, E., S.R. Tronick, S.A. Aaronson, S. Pulciani, and M. Barbacid, *T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes.* Nature, 1982. **298**(5872): p. 343-7.
9. Weinberg, R.A., *Biology of Cancer*. 1st Edition ed. 2006: Taylor & Francis, Inc. .
10. Downward, J., *Targeting RAS signalling pathways in cancer therapy.* Nat Rev Cancer, 2003. **3**(1): p. 11-22.
11. Singh, S.B. and R.B. Lingham, *Current progress on farnesyl protein transferase inhibitors.* Curr Opin Drug Discov Devel, 2002. **5**(2): p. 225-44.
12. Mukhopadhyay, T., M. Tainsky, A.C. Cavender, and J.A. Roth, *Specific inhibition of K-ras expression and tumorigenicity of lung cancer cells by antisense RNA.* Cancer Res, 1991. **51**(6): p. 1744-8.
13. Rodriguez-Viciana, P., O. Tetsu, K. Oda, J. Okada, K. Rauen, and F. McCormick, *Cancer targets in the Ras pathway.* Cold Spring Harb Symp Quant Biol, 2005. **70**: p. 461-7.
14. Chan, D.A. and A.J. Giaccia, *Harnessing synthetic lethal interactions in anticancer drug discovery.* Nat Rev Drug Discov, 2011. **10**(5): p. 351-64.
15. Scholl, C., S. Frohling, I.F. Dunn, A.C. Schinzel, D.A. Barbie, S.Y. Kim, S.J. Silver, P. Tamayo, R.C. Wadlow, S. Ramaswamy, K. Dohner, L. Bullinger, P. Sandy, J.S. Boehm, D.E. Root, T. Jacks, W.C. Hahn, and D.G. Gilliland, *Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells.* Cell, 2009. **137**(5): p. 821-34.
16. Barbie, D.A., P. Tamayo, J.S. Boehm, S.Y. Kim, S.E. Moody, I.F. Dunn, A.C. Schinzel, P. Sandy, E. Meylan, C. Scholl, S. Frohling, E.M. Chan, M.L. Sos, K. Michel, C. Mermel, S.J. Silver, B.A. Weir, J.H. Reiling, Q. Sheng, P.B. Gupta, R.C. Wadlow, H. Le, S. Hoersch, B.S. Wittner, S. Ramaswamy, D.M. Livingston, D.M. Sabatini, M. Meyerson, R.K. Thomas, E.S. Lander, J.P. Mesirov, D.E. Root, D.G. Gilliland, T. Jacks, and W.C. Hahn, *Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1.* Nature, 2009. **462**(7269): p. 108-12.
17. Luo, J., M.J. Emanuele, D. Li, C.J. Creighton, M.R. Schlabach, T.F. Westbrook, K.K. Wong, and S.J. Elledge, *A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene.* Cell, 2009. **137**(5): p. 835-48.

18. Vogelstein, B., C.J. Torrance, V. Agrawal, and K.W. Kinzler, *Use of isogenic human cancer cells for high-throughput screening and drug discovery.* Nature Biotechnology, 2001. **19**(10): p. 940-945.
19. Guo, W., S. Wu, J. Liu, and B. Fang, *Identification of a small molecule with synthetic lethality for K-ras and protein kinase C iota.* Cancer Res, 2008. **68**(18): p. 7403-8.
20. Shaw AT, W.M., Magendantz M, Ouyang C, Dowdle J, Subramanian A, Lewis TA, Maglathin RL, Tolliday N, Jacks T., *Selective killing of K-ras mutant cancer cells by small molecule inducers of oxidative stress.* Proc Natl Acad Sci U S A, 2011. **108**(21).
21. Hahn, W.C., C.M. Counter, A.S. Lundberg, R.L. Beijersbergen, M.W. Brooks, and R.A. Weinberg, *Creation of human tumour cells with defined genetic elements.* Nature, 1999. **400**(6743): p. 464-8.
22. Hahn, W.C., S.K. Dessain, M.W. Brooks, J.E. King, B. Elenbaas, D.M. Sabatini, J.A. DeCaprio, and R.A. Weinberg, *Enumeration of the simian virus 40 early region elements necessary for human cell transformation.* Mol Cell Biol, 2002. **22**(7): p. 2111-23.
23. Dolma, S., S.L. Lessnick, W.C. Hahn, and B.R. Stockwell, *Identification of genotype-selective antitumor agents using synthetic lethal chemical screening in engineered human tumor cells.* Cancer Cell, 2003. **3**(3): p. 285-96.
24. Yagoda, N., M. von Rechenberg, E. Zaganjor, A.J. Bauer, W.S. Yang, D.J. Fridman, A.J. Wolpaw, I. Smukste, J.M. Peltier, J.J. Boniface, R. Smith, S.L. Lessnick, S. Sahasrabudhe, and B.R. Stockwell, *RAS-RAF-MEK-dependent oxidative cell death involving voltage-dependent anion channels.* Nature, 2007. **447**(7146): p. 864-8.
25. Yang, W.S. and B.R. Stockwell, *Synthetic Lethal Screening Identifies Compounds Activating Iron-Dependent, Nonapoptotic Cell Death in Oncogenic-RAS-Harboring Cancer Cells.* Chem Biol, 2008. **15**(3): p. 234-45.
26. Hoffstrom, B.G., A. Kaplan, R. Letso, R.S. Schmid, G.J. Turmel, D.C. Lo, and B.R. Stockwell, *Inhibitors of protein disulfide isomerase suppress apoptosis induced by misfolded proteins.* Nat Chem Biol. **6**(12): p. 900-6.
27. Rostovtsev, V.V., L.G. Green, V.V. Fokin, and K.B. Sharpless, *A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes.* Angew Chem Int Ed Engl, 2002. **41**(14): p. 2596-9.
28. Wang, Q., T.R. Chan, R. Hilgraf, V.V. Fokin, K.B. Sharpless, and M.G. Finn, *Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition.* J Am Chem Soc, 2003. **125**(11): p. 3192-3.
29. Adam, G.C., B.F. Cravatt, and E.J. Sorensen, *Profiling the specific reactivity of the proteome with non-directed activity-based probes.* Chem Biol, 2001. **8**(1): p. 81-95.
30. Speers, A.E. and B.F. Cravatt, *Profiling enzyme activities in vivo using click chemistry methods.* Chem Biol, 2004. **11**(4): p. 535-46.
31. Evans, M.J., A. Saghatelian, E.J. Sorensen, and B.F. Cravatt, *Target discovery in small-molecule cell-based screens by in situ proteome reactivity profiling.* Nat Biotechnol, 2005. **23**(10): p. 1303-7.
32. Smukste, I., O. Bhalala, M. Persico, and B.R. Stockwell, *Using small molecules to overcome drug resistance induced by a viral oncogene.* Cancer Cell, 2006. **9**(2): p. 133-46.
33. Weerapana, E., A.E. Speers, and B.F. Cravatt, *Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in proteomes.* Nat Protoc, 2007. **2**(6): p. 1414-25.
34. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.
35. Silva, J.C., R. Denny, C. Dorschel, M.V. Gorenstein, G.Z. Li, K. Richardson, D. Wall, and S.J. Geromanos, *Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale.* Mol Cell Proteomics, 2006. **5**(4): p. 589-607.

36. Speers, A.E., G.C. Adam, and B.F. Cravatt, *Activity-based protein profiling in vivo using a copper(i)-catalyzed azide-alkyne [3 + 2] cycloaddition.* J Am Chem Soc, 2003. **125**(16): p. 4686-7.
37. Singh, J., R.C. Petter, T.A. Baillie, and A. Whitty, *The resurgence of covalent drugs.* Nat Rev Drug Discov, 2011. **10**(4): p. 307-17.
38. Nicolaou, K.C., D.Y.K. Chen, X.H. Huang, T.T. Ling, M. Bella, and S.A. Snyder, *Chemistry and biology of diazonamide A: First total synthesis and confirmation of the true structure (vol 126, pg 12888, 2004).* Journal of the American Chemical Society, 2004. **126**(46): p. 15316-15316.
39. Gallagher, S.S., J.E. Sable, M.P. Sheetz, and V.W. Cornish, *An in vivo covalent TMP-tag based on proximity-induced reactivity.* ACS Chem Biol, 2009. **4**(7): p. 547-56.
40. Stoughton, R., H.Y. Dai, M. Meyer, S. Stepaniants, and M. Ziman, *Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays.* Nucleic Acids Research, 2002. **30**(16).
41. Weng, L., H. Dai, Y. Zhan, Y. He, S.B. Stepaniants, and D.E. Bassett, *Rosetta error model for gene expression analysis.* Bioinformatics, 2006. **22**(9): p. 1111-21.
42. MacKinnon A., T.J., *Target Identification by Diazirine Photo-Cross-Linking and Click Chemistry.* Current Protocols in Chemical Biology, 2009. **1**(1): p. 55-73.

Chapter 3: Comparing the transcription readout of the yeast three-hybrid system with covalent and non-covalent ligand-receptor pairs

3.1 Introduction

The yeast three-hybrid (Y3H) system is unique among methods used to detect small molecule-protein interactions in that it directly assays drug-receptor binding in an *in vivo* setting. Surprisingly however, despite impressive applications for target identification [1, 2] and enzyme evolution [3], the Y3H system has been relatively underutilized since its initial development.

The Y3H system is the child of two parent technologies, the yeast two-hybrid (Y2H) system and chemical dimerizers. The Y2H system, first introduced by Fields and Song, is a commonly used method for profiling protein-protein interactions [4]. Interaction of two proteins is tied to the transcription of a reporter gene through the use of two hybrid protein chimeras, one to the DNA-binding domain (DBD) of a transcriptional activator and the other to its activation domain (AD). When the two chimeras interact, the transcription factor is effectively reconstituted to drive the expression of a detectable reporter gene, most commonly via an enzymatic (*lac* z) or growth assay (*LEU2, HIS3*).

Chemical dimerizer technology originated from the use of small molecules, particularly the high affinity immunophilin ligands FK506, rapamycin and cyclosporin A, to crosslink hybrid proteins that contained their natural ligand binding domains (FKBP12, mTOR/FRAP or cyclophilin, respectively). Thus, addition of bivalent ligands (connected via a linker) could be used to dimerize other proteins that are fused to the respective binding partners of the ligands [5]. Chemical dimerizers (often referred to as chemical inducers of dimerization, CIDs) have been used to artificially induce events controlled by protein-protein interactions such as gene expression, subcellular localization and membrane recruitment and consequent activation of intracellular signaling [6-9].

Licitra and Liu united these two technologies in their first description of the yeast three-hybrid system, where the DBD and AD domains used in the Y2H system are fused to two receptor proteins that are connected via a dimerizing ligand [10]. Importantly, they were the first to recognize its capabilities as a screening platform to identify *de novo* small molecule-protein interactions in yeast. To demonstrate this concept they showed that a Y3H system containing a hybrid protein of a DBD-glucocorticoid receptor (LexA-GR) could be used along with a heterodimeric molecule comprised of the GR-ligand dexamethasone (Dex) and FK506 to identify the hybrid target of FK506, FKBP12, from a Jurkat cDNA library of cDNA-AD clones.

Other Y3H systems have been developed since then incorporating different ligand-receptor pairs and transcription factors or using different organisms. In the system used in the Cornish laboratory, an 'anchor' comprised of the ligand-receptor pair methotrexate (Mtx) and dihydrofolate reductase (DHFR) has been used to effectively activate transcription in conjunction with a Dex-GR pair [11]. Meldrum and colleagues have modified this system with a GAL4 transcription factor instead [12]. Use of the Mtx-DHFR pair has also been adapted to create a bacterial and a mammalian three-hybrid system system [13, 14]. Additional ligand-receptor pairs have been introduced such as those based on the estrone-estrogen receptor and biotin-streptavidin interactions [15, 16]. A covalent ligand-receptor pair was first demonstrated using a technology developed for the covalent labeling of fusion proteins by small molecules *in vivo*. Now known as 'SNAP-tag', the system is based on O^6-alkylguanine-DNA alkyltransferase (hAGT) that irreversibly reacts with O^6-benzylguanine derivatives (BG). [17]

A major advantage of the Y3H system, similar to other genetic screens, is that the transcriptional phenotype of a small-molecule protein interaction can easily be connected with the genotype of the target by sequencing the cDNA-AD clone of candidate hits that emerge from a screen. As the interaction occurs *in vivo*, detection of proteins that are unstable in cell lysates or that require post-

translational modifications conserved in yeast is improved [5]. The Y3H system is also highly useful for facilitating the detection of low-abundance targets since the protein (in the form of the cDNA-AD clone) is overexpressed. Furthermore, once a system has been created and optimized, it can be applied to profile targets of a range of small molecules connected to an anchor ligand with a number of different cDNA libraries created from various tissues and organisms with relative ease [5]. Thus, this method affords many desirable properties over affinity chromatography, which is largely performed in cell lysates, requires careful detection of enriched proteins by MS and low abundance targets are easily obscured.

Inherent caveats of yeast-based transcription screens as described in Chapter 1.2f still remain. Moreover, just as the affinity of a small molecule to its protein target poses a limitation on affinity chromatography, it has also been a restrictive factor in the transcriptional readout of the Y3H system and is perhaps a reason for its limited application. Even in the early days of the Y3H it was evident that the transcription readout of the system depended heavily on the receptor-small molecule interactions involved. In their proof of principle experiment, Licitra and Liu were required to use mutant versions of the GR protein with higher affinity for dexamethasone in the nanomolar range in order to detect Y3H interactions with Dex-FK506 CIDs [10]. These data suggested that only small molecule-protein interactions of high affinity could be isolated with this method. The relationship between ligand-receptor affinity and transcriptional readout has subsequently been shown to be complex and not necessarily predictable based on known biochemical affinities *a priori* [18]. The Cornish laboratory's version of the Y3H replaced the use of Dex-GR as an anchor with Mtx, an irreversible inhibitor of DHFR with picomolar affinity, to improve the sensitivity of the system [11]. The applicability of the system has been demonstrated by the use of Mtx as an anchor moiety in cDNA screens to identify novel protein targets (Shepard, Rezwan, Cornish; unpublished results) [1].

The transcriptional readout also reflects the affinity of the interaction at the variable end of the chemical dimerizer. For a given anchor molecule and its receptor, there is a K_d threshold for detection of small molecule-protein interactions in the Y3H system. In the Mtx-based system our laboratory has characterized this cut-off to be ca. 50nM [19]. We hypothesized that development of a covalent ligand for DHFR as an anchor in our Y3H system could improve the sensitivity of assay. In this chapter, assays to determine whether a covalent interaction would enhance the transcriptional readout of our system are described. An improvement could potentially lower the affinity threshold and enable the detection of weaker interactions in target identification experiments. Additionally, attempts to use the Y3H system to identify protein targets of RSL3 are also performed.

3.2 Results

3.2a Comparing the transcription readout of the yeast three-hybrid system with covalent and non-covalent ligand-receptor pairs

To test the abilities of a covalent yeast three-hybrid system we took advantage of an engineered eDHFR protein that has been used with acrylamide analogs of trimethoprim (TMP), another substrate of DHFR, for in vivo imaging applications [20]. The eDHFR variant is designed with a cysteine residue positioned close to the binding site of TMP such that a covalent bond specifically forms with the acrylamide moiety in the TMP tag. TMP itself has been previously characterized for use in the Y3H system as an orthogonal CID with greater specificity for the E. coli form of DHFR, however, transcriptional activation in this system was not as effective as with Mtx [21]. The reason for this is possibly the lower affinity of TMP for DHFR (K_i = 1.3nM) compared with Mtx. Comparing a CID incorporating the acrylamide version of TMP (A-TMP) with the original Mtx and TMP CIDs could provide

an explanation for this disparity by creating a version of TMP that effectively acts as a covalent inhibitor of DHFR.

In order to achieve this a Y3H system was constructed with a chimera of a LexA DNA binding domain with the cysteine mutant of eDHFR (LexA-eDHFR:L28C) and a B42 activation domain-GR chimera and *LacZ* reporter gene transcriptionally controlled by four consecutive *LexA* operators to facilitate direct comparison with earlier systems (Figure 3.1B).

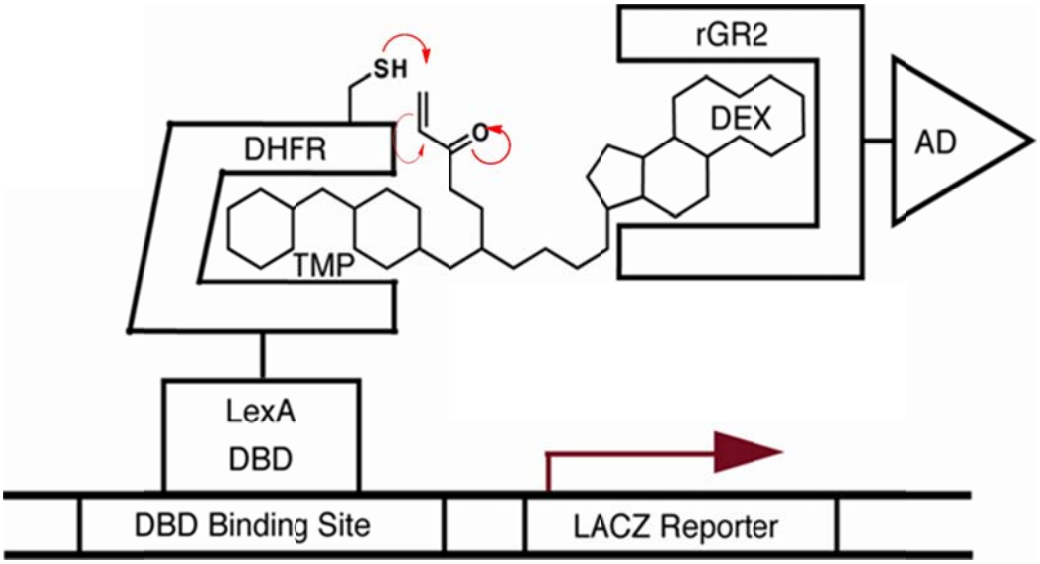


Figure 3.1: A covalent yeast three-hybrid system using acrylamide-trimethoprim-dexamethasone (A TMP-Dex)

In the covalent Y3H system, a protein chimera of the LexA DNA-binding domain (DBD) fused to a L28C variant of dihydrofolate reductase (DHFR) reacts covalently with an acrylamide moiety on the A TMP-Dex heterodimer. Upon interaction of Dex with its receptor, the glucocorticoid receptor (GR) fused to a transcriptional activation domain (AD), the transcriptional activator is reconstituted resulting in transcription of the reporter gene (*Lac Z*).

The A TMP-Dex CID was designed according to previously used syntheses in our laboratory and incorporated a ten carbon methylene linker, found to be optimal for the Y3H in earlier characterization studies [18](Figure 3.2A). The ability of the covalent CID was tested in the new Y3H strain alongside Mtx-Dex and TMP-Dex using a standard transcriptional activation assay in liquid culture of β-galactosidase activity using ortho-nitrophenyl-β-galactoside (ONPG) as a substrate (Figure 3.2B). Yeast strains were grown in synthetic complete liquid media lacking histidine, tryptophan and uracil with 2% galactose and raffinose, respectively, in the absence of glucose. All three CIDs were assayed at 1, 5 or 10uM to observe β-galactosidase activity at a range of concentrations.

Interestingly, while comparable, Mtx-Dex still showed greater fold transcriptional activation (29-fold) over background (no CID) than A TMP-Dex (25-fold) at the highest concentration (10uM) of CID tested (Figure 3.2B). The disparity was greater at 5uM concentration. TMP Dex displayed the weakest levels of transcriptional activation in the system with 12-fold activation over background only observed at 10uM. This suggests that a covalent interaction between DHFR and its substrate does not provide better display of the chemical heterodimer than the very tight non-covalent interaction between Mtx and DHFR.

At the same time we also evaluated the three CIDs in the original wild-type eDHFR Y3H system [11, 22]. Here, transcriptional activation of A TMP-Dex was no longer comparable to Mtx Dex and was instead almost identical to TMP Dex (Figure 3.2C). This provides further evidence that the strength of the activation of A TMP Dex was due to the covalent interaction with eDHFR:L28C in our new system. Due to strain variance, direct comparison between the wild-type and L28C variant systems is not possible.

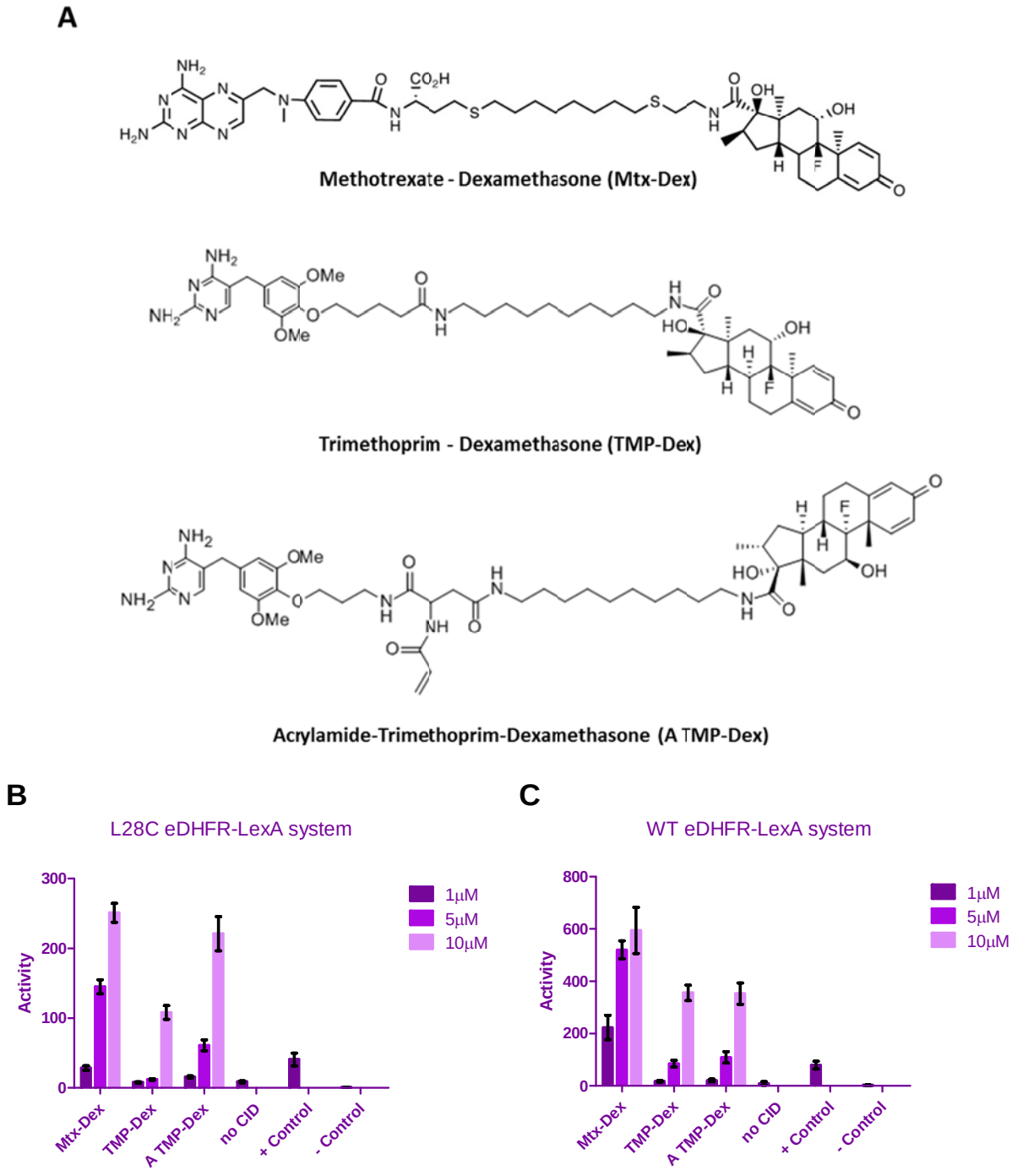


Figure 3.2: *Lac Z* transcription assays comparing the abilities of three Dex heterodimers in L28C variant and wild-type Y3H system

(A) The structures of methotrexate-dexamethasone (Mtx-Dex), trimethoprim-dexamethasone (TMP-Dex) and acrylamide trimethoprim-dexamethasone (A TMP-Dex) used in this study.

(B) The transcriptional activation of a *Lac Z* reporter gene was compared between methotrexate-dexamethasone (Mtx-Dex), trimethoprim-dexamethasone (TMP-Dex) and acrylamide trimethoprim-dexamethasone (A TMP-Dex) in the Y3H system comprising the L28C variant of dihydrofolate reductase (DHFR) fused to the Lex A DNA-binding domain at three different concentrations.

(C)The assay described in (B) was repeated in the original Y3H system with a wild-type variant of DHFR.

Data represent the mean of three replicates ±SD.

Next we assayed the new A TMP CID in a mock selection assay that more closely tests the ability of a yeast hybrid system in a setting analogous to which it will be used. Y3H screens are typically performed using growth selections to isolate cDNA-AD fusion proteins that interact with a small molecule of interest and activate transcription of an auxotrophic marker (such as *LEU2* and *HIS3*). Thus a robust Y3H system requires a ligand-receptor pair that enables the enrichment of 'positive' interactions from a background of 'null' interactions or false positives (non-specific interactions). As the CID presence is essential for these interactions, it enables powerful counter screens to eliminate false positives from potential hits by performing the selection in the absence of the CID.

We utilized an experimental framework developed in our laboratory to evaluate the enrichment capabilities of yeast hybrid growth selections [23]. This framework is comprised of a colorimetric assay to monitor the enrichment of 'positive' vs. 'null' interactions in a mock selection experiment. Using a different marker gene in each strain, the positive and null strains enzymatically produce a different color in the presence of different artificial reporter substrates. By plating cells at various time points in a selection experiment and assaying colonies with these chemical substrates, enrichment of positive strains can be monitored by determining the percentage of positive and null strains using this color indicator.

Strains representing positive Y3H interactions were constructed with a B42-GR chimera on a plasmid that also constitutively expressed β-glucuronidase (*gusA*) that turns blue in the presence of the compound X-Gluc. Corresponding null strains contained a B42 construct with no GR on a plasmid that constitutively expressed β-galactosidase (*lacZ*) that turns purple in the presence of the compound Magenta-Gal. Both positive and null strains possessed LexA-eDHFR:L28C chimeras and a *LEU2* reporter gene under the control of six *Lex A* operators integrated into the chromosome, which has been shown to stabilize the transcription readout [24].

The positive strains were diluted in a range of successively larger library sizes composed of the null strain to measure the enrichment in increasingly difficult selection conditions. Strains were grown in the absence or presence of either 5uM A TMP-Dex CID or Mtx-Dex CID in synthetic complete liquid media lacking histidine, tryptophan and leucine with 2% galactose and raffinose respectively. Enrichment was measured after 3 days of selection.

Preliminary results from mock selections indicate that, while comparable, enrichment observed in the presence of Mtx-Dex was higher than with A TMP-Dex (Figure 3.3). Intriguingly this relationship appeared to be drastically reversed in libraries with complexity of 10^7, where A TMP-Dex showed significantly greater enrichment than Mtx-Dex. This phenomenon needs to be further investigated. Finally, beyond Day 3 an overall drop in enrichment was observed while general patterns between the different CIDs remained the same (data not shown).

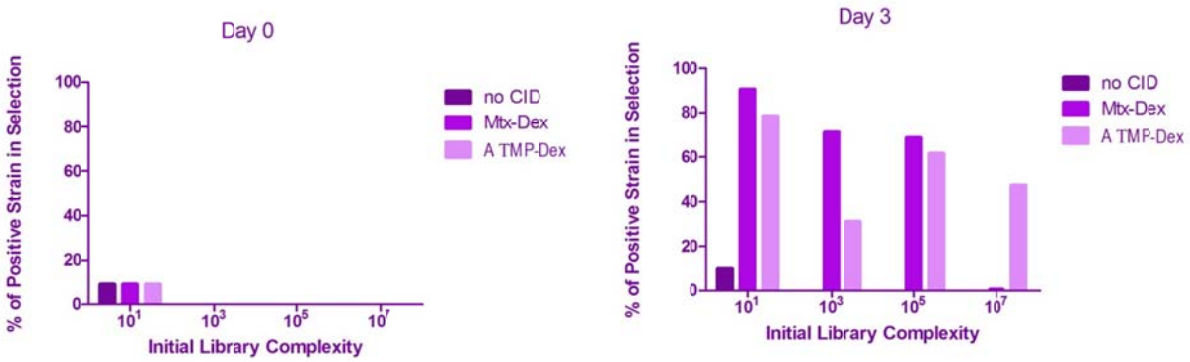


Figure 3.3 Mock selections comparing the enrichment abilities of the non-covalent and covalent heterodimers in the yeast three-hybrid system

Enrichment capabilities of methotrexate-dexamethasone (Mtx-Dex) and acrylamide trimethoprim-dexamethasone (A TMP-Dex) were compared in the L28C eDHFR variant of the Y3H system by calculating the percentage of positive Y3H interactions detected in a mock selection over three days with increasing library complexity (10^1-10^7). Preliminary results showed that Mtx-Dex again performed better than A TMP-Dex, although the relationship appeared to be reversed unusually at higher library complexity.

3.2b Efforts to use the yeast three-hybrid assay for the target identification of RSL3

The Y3H system has been successfully applied to profile targets of small molecules with known and unknown mechanisms of action (as described in Chapter 1.2f). Thus, we wanted to apply this promising technology to identify targets of RSL3. In order to do this we needed to prepare a CID containing RSL3 and Mtx.

Using SAR knowledge gained in the synthesis of other affinity reagents from the RSL3 scaffold (discussed in Chapter 2) along with established syntheses of Mtx in the Cornish laboratory, we designed a synthetic strategy to create an RSL3-Mtx CID (Figure 3.4A). A t-butyl protected Mtx was coupled with a heterobifunctional PEG molecule containing a terminal amine and azide (1). Compound (1) was then coupled to the (1S,3R) version of alkyne-RSL3 probe 2 (Chapter 2.2c) via a copper-catalyzed 1,3-dipolar cycloaddition reaction (click chemistry) to give (2) a t-butyl protected version of the RSL3-Mtx CID. The next step required deprotection of the t-butyl group to give the active form of Mtx as compound (3). During this reaction we observed a large amount of side products that we speculated could be due to decomposition of the RSL3 structure under the highly acidic conditions required. Reversing the last two steps of the synthesis to avoid this did achieve the final product, but the molecule proved to be unstable during characterization. Note that this molecule was never fully characterized and thus results from testing were inconclusive.

Furthemore, testing of the intermediate compound (2) in our *RAS*-selective lethal assay showed that the molecule had no activity (Figure 3.4B) (Note that the intermediate molecule was only characterized by ^{1}H NMR). The Mtx was located at a distance analogous to that of the protected fluorescein synthesized as an affinity reagent in Chapter 2, thus it was surprising that a molecule with Mtx at this position showed no activity. We speculated that binding of Mtx to the mammalian DHFR may be mislocalizing the RSL3-Mtx CID although crystal structures have shown that the free carboxylic acid in the structure is important for binding. Nonetheless, we attempted to co-treat cells with an

excess of Mtx and the RSL3-Mtx CID to see whether the activity was recovered. These experiments also exhibited no lethality of the RSL3-Mtx CID, illustrating the complexity of creating active heterodimeric ligands. Due to these unpromising results, we decided to recontinue pursuit of target identification via affinity chromatography methods.

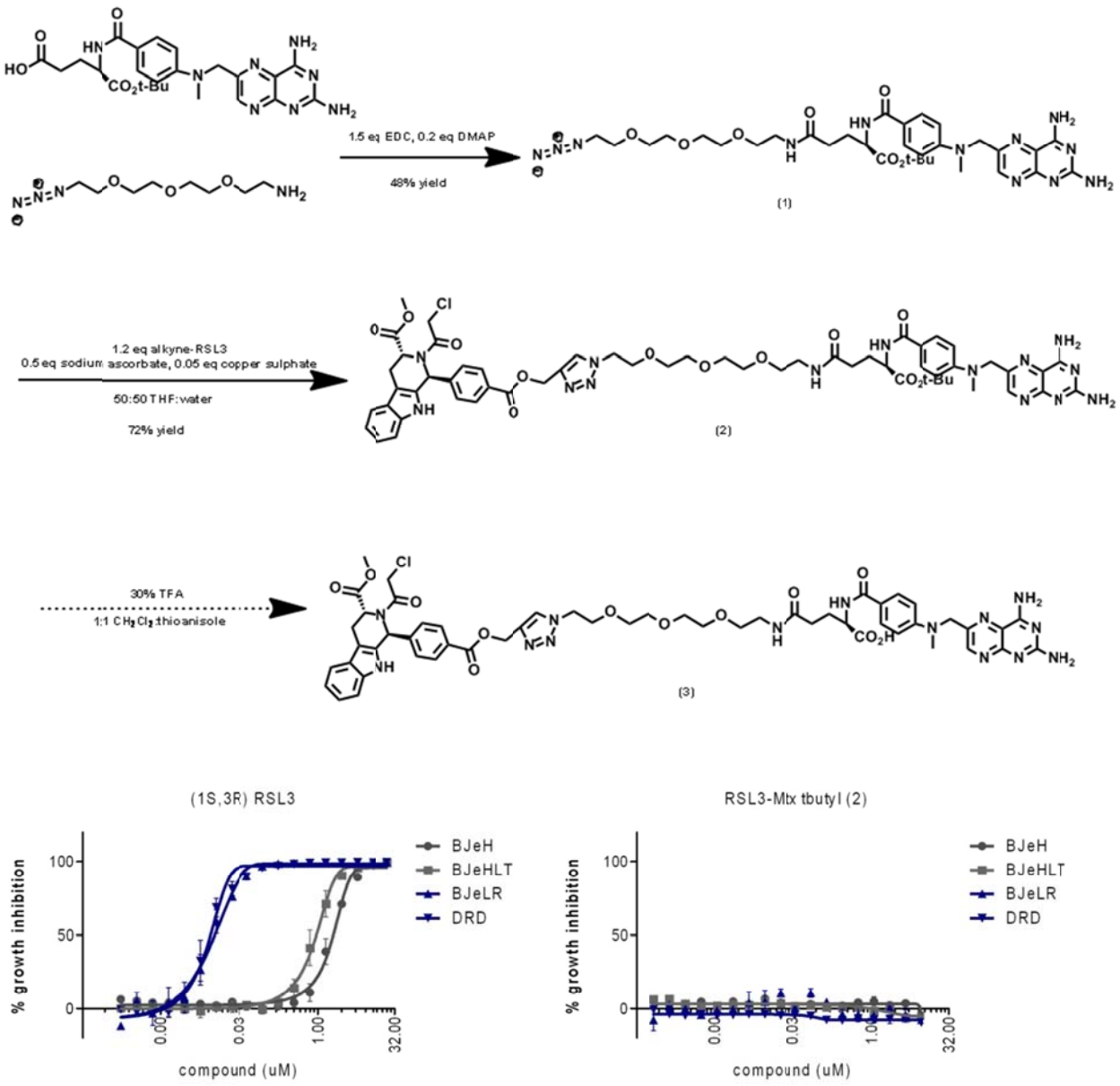


Figure 3.4 Synthetic route for a RSL3-Mtx chemical inducer of dimerization and dose response

(A) A t-butyl protected Mtx was coupled with a heterobifunctional PEG molecule containing a terminal amine and azide to give (2). Compound (2) was then coupled to the (1S,3R) version of alkyne-RSL3 probe 2 (Chapter 2.2c) to give (3) a t-butyl protected version of the RSL3-Mtx CID. Deprotection under highly acidic conditions in the last step resulted in a number of side products.

(B) Dose response curve comparing (1S,3R) RSL3 with compound (2), a t-butyl protected form of the RSL3-Mtx CID.

3.3 Discussion

In this study we have shown that a covalent ligand of DHFR can be used in the Y3H system to give a robust transcriptional readout. The A TMP-eDHFR:L28C interaction provides a third ligand-receptor pair, adding to the previously used Mtx and TMP. The additional covalent ligand-receptor pair can also be utilized as a method for validation of candidate target proteins following a Y3H screen through affinity chromatography studies. As demonstrated in another Y3H screening platform, a GST-tagged receptor protein such as eDHFR:L28C could be coupled to beads and loaded with an A TMP ligand connected to the small molecule of interest for immunoprecipitation-like experiments in lysates or with purified candidate target proteins [2].

Ultimately, we observed that the covalent A TMP-Dex CID does not provide greater transcriptional activation than the original Mtx-Dex. Measurements have been performed to characterize the affinity and off-rate of Mtx and TMP with wild-type DHFR, but it is not clear how these would correlate with derivatized version of these compounds (in the form of the CID) with the eDHFR:L28C variant. Data exist that suggest that derivatized TMP analogs have affinity in the same order of magnitude for both the wild-type and L28C variant of eDHFR (Aaron Hoskins, personal communication). Future experiments characterizing the K_d of A-TMP Dex and Mtx-Dex with the L28C variant of eDHFR will provide better insight into their differences. Additonally, while the acrylamide moiety has been demonstrated to react very specifically with eDHFR:L28C in mammalian cells, this has not been fully characterized in yeast. Non-specific reactivity of the A TMP-Dex CID could reduce cellular concentrations.

Finally, the ability of A TMP Dex to enrich positive strains from libraries of null strains has been preliminarily explored. Initial results indicate although in general its enrichment abilities were not

greater than that of Mtx Dex, it has the potential to show enrichment in libraries of large complexity size, which is an interesting prospect that requires future investigation.

In order to directly determine whether a covalent ligand-receptor interaction as an anchor improves the sensitivity of the system, the limits of its ability to detect small molecule-protein interactions of lower affinity needs to be assessed. The same assays described here can be applied to a series of AD-target protein chimeras with varying affinities to determine the K_d threshold for the new system [19].

3.4 Methods

Synthesis of A-TMP Dex and TMP-Dex were performed by Chaoran Jing.

3.4a Construction of strains

Strains for use in these experiments were constructed using the *S. cerevisiae* strain EGY48 (MATa trp1 his3 ura3 6LexAop-LEU2 GAL+) as a parent strain that contains a *LEU2* reporter under the control of 6 *lexA* operators and is auxotrophic for histidine, uracil, and tryptophan. Strains for the Lac Z assay were constructed by transforming vectors into EGY48 containing LexA-eDHFR:L28C (p2363 prepared by Vanessa Mondol) and B42-rGR2 (pBC398E) under control of $P_{gal1}$and a third plasmid containing the *lacZ* gene under the control of 4 *lexA* operators (pMW112). Yeast electroporation was carried out using a Bio-Rad Gene Pulser Xcell.

Strains for the mock selection experiments were constructed through integration of *LexA-eDHFR:L28C* into the EGY48 chromosome. This was done by performing site-directed mutagenesis using Strategene's QuikChange Mutagenesis Kit on an integration plasmid (pLW2666) containing a wild-type *pGAL-LexA-DHFR* construct to create the cysteine variant. Primers for mutagenesis: 5' - GGA ACC TGC CTG CCG ATT GCG CAT GGT TTA AAC GCA AC - 3' and 5' - GTT GCG TTT AAA CCA TGC GCA ATC GGC AGG CAG GTT GC - 3'. DNA sequencing was confirmed using Genewiz. The *pGAL-LexA-DHFR:L28C* fragment was excised by

digestion with ApaII and isolated by gel purification. The fragment was transformed into EGY48 and insertion into the chromosome was verified by genomic DNA purification using a YeaStar Genomic DNA kit (Zymo research) and sequenced using Genewiz. Plasmids containing *B42-GR gusA* constructs (positive strain) and *B42-lacZ* constructs (null strain) were transformed into the new EGY48 strain. Plasmids were prepared by Laura Wingler as described in [24](pLW2577 + pLW2578).

3.4b Lac Z assay

The yeast strains expressing the three-hybrid constructs with either wild-type or mutant versions of DHFR were assayed for β-galactosidase activity in liquid cultures. Each of the yeast strains was inoculated from patches into SC media containing 2% glucose and lacking histidine, uracil, and tryptophan. Incubation at 30 °C, shaking at 250 rpm, was allowed to proceed until saturation overnight. For the liquid assays, 5 uL of the saturated yeast cultures was used to inoculate 95 uL of SC media containing 2% galactose and 2% raffinose and lacking histidine, uracil, and tryptophan. Cells were treated with 1, 5 or 10uM of CID or the equivalent amount of vehicle. The cultures were allowed to grow for 3 days at 30 °C, shaking at 250 rpm. The cells were harvested by centrifugation, and the pellets were subsequently resuspended in 100 uL of distilled water. Next, the cultures were transferred to a flat-bottomed 96-well plate (Fisher) for reading at OD_{600}. The cultures were centrifuged, and the pellets were resuspended in 100 uL of the Y-Per protein extraction reagent (Pierce). Lysis was allowed to proceed for 30 min. Then, 8.5 uL of a 10 mg/mL ONPG (ortho-nitrophenyl-β-galactoside) solution was added to the extracts and allowed to incubate for 10 min at 37 °C. The β-galactosidase activity reaction was stopped with 110 uL of 1 M sodium carbonate. The extracts were centrifuged, and the supernatant was transferred to a flat-bottomed 96-well plate, where the OD_{420} was measured. The following equation was used to calculate β-galactosidase units: β-galactosidase = 1000 X [OD_{420}/(OD_{600} X time (in min) X volume assayed (in mL))].

3.4c Mock selections

One mL of SC media containing 2% glucose and lacking histidine and tryptophan was inoculated from patches of the positive or null strain. These starter cultures were used to inoculate overnight cultures (10-50 mL) of the same media except containing 2% galactose and 2% raffinose to pre-induce the yeast three-hybrid construct expression. Cells were harvested, washed 3X with sterile water, and resuspended in the SC media containing 2% galactose and 2% raffinose and lacking histidine, tryptophan and leucine for selection. The OD_{600} of the each strain was determined and cells were mixed to give 10:1, 10^3:1, 10^5:1, and 10^7:1 ratios of the null:positive strains. For yeast three-hybrid selections, Mtx-Dex and A-TMP-Dex were added to a final concentration of 5uM. Each selection had a final volume of 3 mL and a calculated initial OD_{600} of 1. Selections were shaken at 30°C. On days 0, 3, 6, and 9, cells were plated on plates containing SC media with 2% glucose and lacking histidine and tryptophan and after 2 days of growth, plates were assayed using the overlay assay described below. After the plates developed, the number of red and blue colonies on each plate was counted.

3.4d Magenta-Gal/X-Gluc overlay assay.

Potassium phosphate buffer (300 mL; 0.5 M, pH 7.0), 20 mL of DMF, 3.3 mL of 10% SDS, and 3.3 g low-melting agarose were combined in an Erlenmeyer flask. The solution was microwaved until the SDS and agarose went into solution. The flask was cooled in a 65°C water bath. β-mercaptoethanol (165 uL), X-Gluc (165 mg dissolved in 1 mL DMF), and Magenta-Gal (50 mg dissolved in 1 mL DMF) were added. After gentle mixing, a pipette was used to carefully cover each plate with approximately 10 mL of the agarose solution. The color typically developed sufficiently overnight for plates to be counted the next day.

3.4e Synthesis of RSL3-Mtx tbutyl (2)

Mtx tbutyl (98mg, 0.19mmol), EDC (55mg, 0.288mmol) and DMAP (5mg, 38umol) were dried under vacuum for an hour. In another flask 11-azido-3,6,9-trioxaundecan-1-amine (42uL, 0.21mmol) and a few

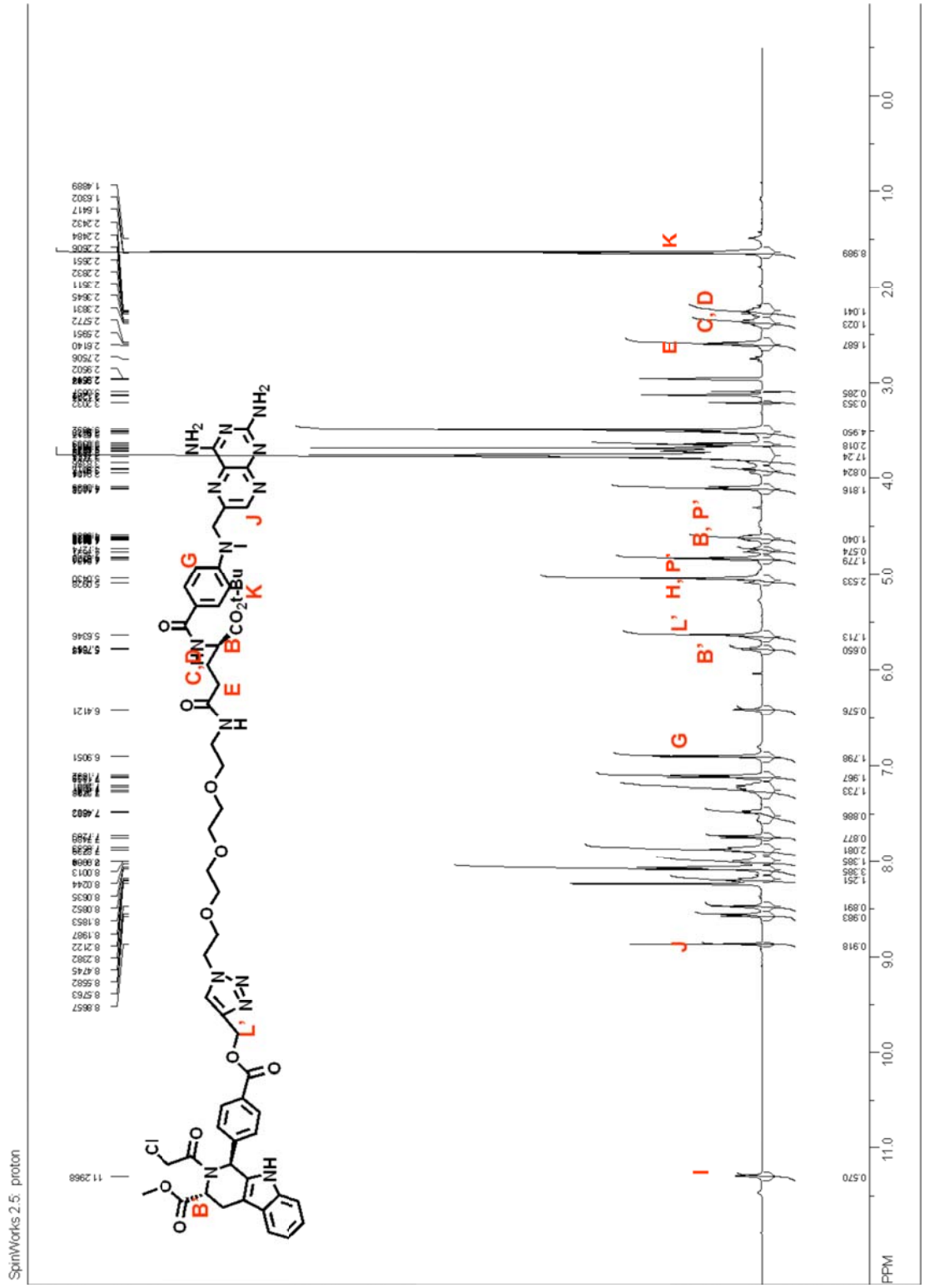
drops of anhydrous NEt3 were dissolved in anhydrous DMF. The tryptophan mixture was also dissolved in anhydrous DMF, added dropwise to the amine mixture and allowed to stir overnight. The Mtx tbutyl azide (1) was purified by silica gel column chromatography in 5:95 MeOH:CH_2Cl_2 to give the product in 48% yield. MS, m/z 711.6 (MH+) calculated 710.4

1H NMR (400 MHz, CDCl3) δ 8.61 (s, 1H), 7.71 (d, J = 8.4 Hz, 2H), 7.24 (d, J = 7.1 Hz, 1H), 6.82 (s, 1H), 6.71 (d, J = 8.6 Hz, 3H), 5.80 (s, 2H), 4.70 (s, 2H), 4.63 (dd, J = 13.6, 6.5 Hz, 1H), 3.74 – 3.61 (m, 8H), 3.59 (s, 2H), 3.55 – 3.46 (m, 2H), 3.45 – 3.41 (m, 1H), 3.40 (tt, J = 18.7, 9.5 Hz, 4H), 3.16 (s, 3H), 2.37 (ddd, J = 21.2, 14.7, 7.6 Hz, 2H), 2.19 (ddd, J = 20.8, 14.0, 6.2 Hz, 3H), 1.56 – 1.37 (m, 9H).

Compound (1) (10mg, 414umol) and 1S3R RSL3-alkyne probe 2 (8mg, 17umol) were dissolved in THF and sodium ascorbate (1.4 mg, 70umol) and copper sulphate pentahydrate (0.2 mg, 0.7umol) dissolved in water were added. The reaction was stirred overnight at 40°C. The final product was purified using TLC preparatory plates run in 5:95 MeOH: CH_2Cl_2 in 72% yield.

1H NMR (400 MHz, DMF) δ 11.30 (s, 1H), 8.87 (s, 1H), 8.57 (d, J = 7.2 Hz, 1H), 8.19 (d, J = 5.4 Hz, 1H), 8.07 (d, J = 8.7 Hz, 3H), 8.00 (s, 1H), 7.86 (d, J = 7.9 Hz, 2H), 7.74 (d, J = 7.7 Hz, 1H), 7.48 (d, J = 7.6 Hz, 1H), 7.37 – 7.16 (m, 2H), 7.11 (d, J = 8.8 Hz, 2H), 6.90 (s, 2H), 6.41 (s, 1H), 5.78 (s, 1H), 5.63 (s, 2H), 5.07 (d, J = 19.9 Hz, 3H), 4.83 (t, J = 5.1 Hz, 2H), 4.74 (d, J = 13.9 Hz, 1H), 4.69 – 4.56 (m, 1H), 4.10 (t, J = 5.1 Hz, 2H), 3.98 – 3.87 (m, 1H), 3.76 (d, J = 5.9 Hz, 12H), 3.71 (dd, J = 5.8, 3.7 Hz, 3H), 3.68 (s, 3H), 3.65 (t, J = 5.9 Hz, 2H), 3.57 – 3.50 (m, 2H), 2.60 (t, J = 7.4 Hz, 2H), 2.44 – 2.32 (m, 1H), 2.31 – 2.18 (m, 1H), 1.64 (d, J = 4.6 Hz, 9H). MS, m/z 1175.6 (MH+) calculated 1174.5

1S3R RSL3-Mtx tbutyl



3.4f Testing in four BJ cell lines

This assay was performed as described in Chapter 2.4d.

3.4g Strains and plasmids

Strain/Plasmid		Source/Ref
EGY48	(*MATα trp1 his3 ura3 6LexAop-LEU2 GAL+*)	R. Brent/[25]
pBC398E	(*pGAL1-B42-(GSG)$_2$-rGR2 2μ TRP1* pUC ori kan^R)	B. Carter/[18]
pLW2578	(pBC398E with *pADH-gusA* inserted)	L.Wingler/24
pMW112	(8LexAop-lacZ 2 μ URA3 pBR ori kan^R)	R. Brent/[26]
pLW2577	(*pGAL1-B42 pADH-lacZ 2μ TRP1* pUC ori amp^R)	L.Wingler/24
pLW2666	(HO-polylinker-KanMX4-HO with h*pGAL-LexA-eDHFR-tADH-HIS3)*	L.Wingler/24

3.5 References

1. Becker, F., K. Murthi, C. Smith, J. Come, N. Costa-Roldan, C. Kaufmann, U. Hanke, C. Degenhart, S. Baumann, W. Wallner, A. Huber, S. Dedier, S. Dill, D. Kinsman, M. Hediger, N. Bockovich, S. Meier-Ewert, A.F. Kluge, and N. Kley, *A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors.* Chem Biol, 2004. **11**(2): p. 211-23.
2. Chidley, C., H. Haruki, M.G. Pedersen, E. Muller, and K. Johnsson, *A yeast-based screen reveals that sulfasalazine inhibits tetrahydrobiopterin biosynthesis.* Nat Chem Biol, 2011. **7**(6): p. 375-83.
3. Peralta-Yahya, P., B.T. Carter, H. Lin, H. Tao, and V.W. Cornish, *High-throughput selection for cellulase catalysts using chemical complementation.* J Am Chem Soc, 2008. **130**(51): p. 17446-52.
4. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-6.
5. Kley, N., *Chemical dimerizers and three-hybrid systems: scanning the proteome for targets of organic small molecules.* Chem Biol, 2004. **11**(5): p. 599-608.
6. Pollock, R. and T. Clackson, *Dimerizer-regulated gene expression.* Curr Opin Biotechnol, 2002. **13**(5): p. 459-67.
7. Spencer, D.M., T.J. Wandless, S.L. Schreiber, and G.R. Crabtree, *Controlling signal transduction with synthetic ligands.* Science, 1993. **262**(5136): p. 1019-24.
8. Belshaw, P.J., S.N. Ho, G.R. Crabtree, and S.L. Schreiber, *Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins.* Proc Natl Acad Sci U S A, 1996. **93**(10): p. 4604-7.
9. Ho, S.N., S.R. Biggar, D.M. Spencer, S.L. Schreiber, and G.R. Crabtree, *Dimeric ligands define a role for transcriptional activation domains in reinitiation.* Nature, 1996. **382**(6594): p. 822-6.
10. Licitra, E.J. and J.O. Liu, *A three-hybrid system for detecting small ligand-protein receptor interactions.* Proc Natl Acad Sci U S A, 1996. **93**(23): p. 12817-21.
11. Cornish, V.W., H.N. Lin, W.M. Abida, and R.T. Sauer, *Dexamethasone-methotrexate: An efficient chemical inducer of protein dimerization in vivo.* Journal of the American Chemical Society, 2000. **122**(17): p. 4247-4248.
12. Henthorn, D.C., A.A. Jaxa-Chamiec, and E. Meldrum, *A GAL4-based yeast three-hybrid system for the identification of small molecule-target protein interactions.* Biochem Pharmacol, 2002. **63**(9): p. 1619-28.
13. Althoff, E.A. and V.W. Cornish, *A bacterial small-molecule three-hybrid system.* Angew Chem Int Ed Engl, 2002. **41**(13): p. 2327-30.
14. Caligiuri, M., L. Molz, Q. Liu, F. Kaplan, J.P. Xu, J.Z. Majeti, R. Ramos-Kelsey, K. Murthi, S. Lievens, J. Tavernier, and N. Kley, *MASPIT: three-hybrid trap for quantitative proteome fingerprinting of small molecule-protein interactions in mammalian cells.* Chem Biol, 2006. **13**(7): p. 711-22.
15. Hussey, S.L., S.S. Muddana, and B.R. Peterson, *Synthesis of a beta-estradiol-biotin chimera that potently heterodimerizes estrogen receptor and streptavidin proteins in a yeast three-hybrid system.* J Am Chem Soc, 2003. **125**(13): p. 3692-3.
16. Muddana, S.S. and B.R. Peterson, *Facile synthesis of cids: biotinylated estrone oximes efficiently heterodimerize estrogen receptor and streptavidin proteins in yeast three hybrid systems.* Org Lett, 2004. **6**(9): p. 1409-12.
17. Gendreizig, S., M. Kindermann, and K. Johnsson, *Induced protein dimerization in vivo through covalent labeling.* J Am Chem Soc, 2003. **125**(49): p. 14970-1.
18. Abida, W.M., B.T. Carter, E.A. Althoff, H. Lin, and V.W. Cornish, *Receptor-dependence of the transcription read-out in a small-molecule three-hybrid system.* Chembiochem, 2002. **3**(9): p. 887-95.

19. de Felipe, K.S., B.T. Carter, E.A. Althoff, and V.W. Cornish, *Correlation between ligand-receptor affinity and the transcription readout in a yeast three-hybrid system.* Biochemistry, 2004. **43**(32): p. 10353-63.
20. Gallagher, S.S., J.E. Sable, M.P. Sheetz, and V.W. Cornish, *An in vivo covalent TMP-tag based on proximity-induced reactivity.* ACS Chem Biol, 2009. **4**(7): p. 547-56.
21. Gallagher, S.S., L.W. Miller, and V.W. Cornish, *An orthogonal dexamethasone-trimethoprim yeast three-hybrid system.* Anal Biochem, 2007. **363**(1): p. 160-2.
22. Baker, K., D. Sengupta, G. Salazar-Jimenez, and V.W. Cornish, *An optimized dexamethasone-methotrexate yeast 3-hybrid system for high-throughput screening of small molecule-protein interactions.* Anal Biochem, 2003. **315**(1): p. 134-7.
23. Wingler, L.M. and V.W. Cornish, *A library approach for the discovery of customized yeast three-hybrid counter selections.* Chembiochem, 2011. **12**(5): p. 715-7.
24. Wingler, L.M., Cornish V.W, *An Experimental Framework for the Characterization and Optimization of Yeast Hybrid Selections.* In preparation, 2011.
25. Estojak, J., R. Brent, and E.A. Golemis, *Correlation of two-hybrid affinity data with in vitro measurements.* Mol Cell Biol, 1995. **15**(10): p. 5820-9.
26. Watson, M.A., R. Buckholz, and M.P. Weiner, *Vectors encoding alternative antibiotic resistance for use in the yeast two-hybrid system.* Biotechniques, 1996. **21**(2): p. 255-259.

Chapter 4: Characterization of the mechanism of action of RSL3

4.1 Introduction

RSL3 is an interesting molecule from many angles. Originally it was identified for its ability to preferentially kill tumor cells with oncogenic *RAS*. Further investigation of its mechanism of action revealed that it didn't function through an apoptotic form of cell death and instead induced a form of cell death that was oxidative and required free cellular iron. These conclusions were inferred from data showing that RSL3's toxicity could not be rescued by inhibitors of caspases, classical effectors of apoptosis, and did not result in cleavage of poly (ADP-ribose) polymerase-1 (PARP-1), a substrate of activated caspase 3 and 7 [1]. A counter screen of bioactive small molecules led to the discovery that cell death could be rescued with iron chelators (deferoxamine mesylate) and a range of antioxidants (e.g. Vitamin E, glutathione). The role of iron was further confirmed by genetic evidence showing the effect of iron-related genes on RSL3-induced cell death as well as oncogenic-*RAS* signaling on iron-related genes (Dixon, Stockwell; unpublished results) [1]. Generation of reactive oxygen species (ROS) were also detected in cells soon after RSL3 treatment. Additionally, among the kinase inhibitors tested, an inhibitor of MEK1 & 2 kinases (U0126) suppressed RSL3's lethality while inhibitors of kinases in other downstream signaling pathways were ineffective. These results implied that the primary RAS pathway required for death was the RAS-RAF-MEK pathway.

Many of the compounds identified in our *RAS*–selective lethal (RSL) screens possess similar properties; Erastin was the first molecule to be discovered and was shown to act by a ROS-dependent mechanism through its targets, the VDAC proteins [2]. Later, further RSL agents were identified and also investigated, including RSL3. RSL5 in particular, also acts through the VDAC proteins and its lethal effects are rescued by RNAi directed at *VDAC3*. RSL3's mechanism of action however, does not involve the VDAC proteins, and is the primary subject of this chapter.

Mutant KRAS is the most prevalent isoform of RAS found in cancer, with mutations in NRAS and HRAS comprising less than 20% [3]. The tumorigenic forms of the BJ series contain mutant HRAS, thus the effectiveness of RSL3 in cell lines with the other isoforms of mutant RAS was tested. RSL3 showed lethal effects in cell lines possessing mutant KRAS (e.g. Calu-1) and NRAS (e.g. HT1080) suggesting a general mechanism against all isoforms of oncogenic *RAS*.

Death induced by RSL3 is also noticeably rapid, with significant cell death occurring by eight hours. As discussed in Chapter 2, the molecule exhibits high potency in certain cell lines (ca. 10-20nM). This feature, along with its unique pattern of killing in the National Cancer Institute's panel of sixty human cancer cell lines (NCI60) resulted in its selection for further testing in mouse models by their Developmental Therapeutics Program (described in Chapter 1.3b). Analysis using the NCI COMPARE algorithm of RSL3's lethal profile in the NCI60 showed poor correlation with other compounds in their database, highlighting a potentially novel mechanism of action. These preliminary studies showed convincingly that RSL3 had desirable medicinal properties and acted in a unique and novel way, warranting significant investigation into its mechanism of action in order to fully understand its therapeutic potential.

4.1a Polypharmacology

As tools to understand biological function, small molecules are most useful when they specifically modulate a particular protein, because only then can they be used with the same clarity as gene modulation in genetic experiments. However, for small molecules that are used as probes for drug discovery efforts, there are diverging opinions in the field about what is most desirable. The existing concept of finding 'targeted' drugs implies 'magic bullet' specificity against a disease-causing gene product to maximize efficacy against the target and reduce unwanted off-target side effects. Efforts have been made to distinguish compounds that act through single targets from those with multiple

target ('side') effects [4]. Others have argued, however, that drugs designed against specific targets do not necessarily translate into good clinical response as the modulation of multiple targets can be required to overcome complicated, redundant disease networks [5-7]. Additionally single-target drugs are vulnerable to the development of drug resistance in cells, particularly in cancer, which is aided by genomic instability [8]. Many effective clinically-used drugs work through multiple targets simultaneously [9]. Thus, a new paradigm in drug discovery is the use of drugs with rational polypharmacology [5]. Among the many challenges this paradigm faces, balancing improved clinical efficacy with toxicity is one of the greatest concerns. In applying this concept to identify mechanisms to selectively kill tumors with oncogenic *RAS*, one must accept that it may require the perturbation of multiple targets. Thus small molecules can be used as tools to decipher potential multi-target mechanisms.

In this chapter the lethal mechanism of RSL3 is investigated through the use of affinity chromatography to identify potential direct targets of this molecule with the eventual goal of understanding a potential therapeutic approach to selectively target oncogenic *RAS* in cancer.

4.2 Results

4.2a RSL3 potentially interacts with GPx4, MCM7, PTGES2 and SMG9

An affinity chromatography experiment using a fluorescein-tagged version of RSL3 was performed and submitted for label-free shotgun proteomic analysis as described in Chapter 2.2c. Analysis of the proteomics samples prepared in the fluorescein-RSL3 pull down identified a number of candidate target proteins that were enriched in the probe-treated samples when compared with samples treated with the inactive diastereomer of fluorescein-RSL3 or 500 namomolar (nM) (1X) of the parent RSL3 as a competitor. Enrichment of proteins in the probe-treated sample compared to the 50nM competitor treated sample was weaker in general and was therefore not used to filter candidate

targets. The hits were ranked according to the ratios between probe and inactive and probe and competitor (500nM) and discriminated on the basis of the strength of the identification (the number of individual peptides that matched that protein), as well as the consistency of the enrichment pattern among the different peptides identified for a specific protein. Based on these criteria, four proteins that showed the desired pattern of enrichment were selected for further preliminary validation (Table 4.1, Figure 4.1). Other proteins ranked below the top four did not show as great enrichment in probe-treated samples, the significance of the protein intensity measured was poorer and often the pattern of enrichment did not correlate with the treatments.

Protein name	Ratio (P/IA)	Ratio (P/C 500 nM)	Ratio (P/C 50 nM)
GPx4_HUMAN **Phospholipid hydroperoxide glutathione peroxidase 4**	5.35	4.45	1.53
SMG9_HUMAN **Nonsense mediated mRNA decay factor**	4.04	3.01	1.35
PTGES2_HUMAN **Prostaglandin E synthase 2**	2.60	3.26	1.85
MCM7_HUMAN **Minichromosome maintenance complex component 7**	4.55	4.56	1.67

Table 4.1: Top candidate targets from fluorescein-RSL3 pull down proteomic analysis.
Comparing fluorescein-RSL3 probe (**P**)-treated and inactive fluorescein-RSL3 (**IA**) probe-treated cells, greatest fold enrichment was observed of glutathione peroxidase 4 (GPx4), followed by minichromosome maintenance complex component 7 (MCM7). Similarly**,** MCM7 and GPx4 showed the greatest fold enrichment when comparing fluorescein-RSL3 probe (**P**)-treated and competitor (**C500**)-treated cells.

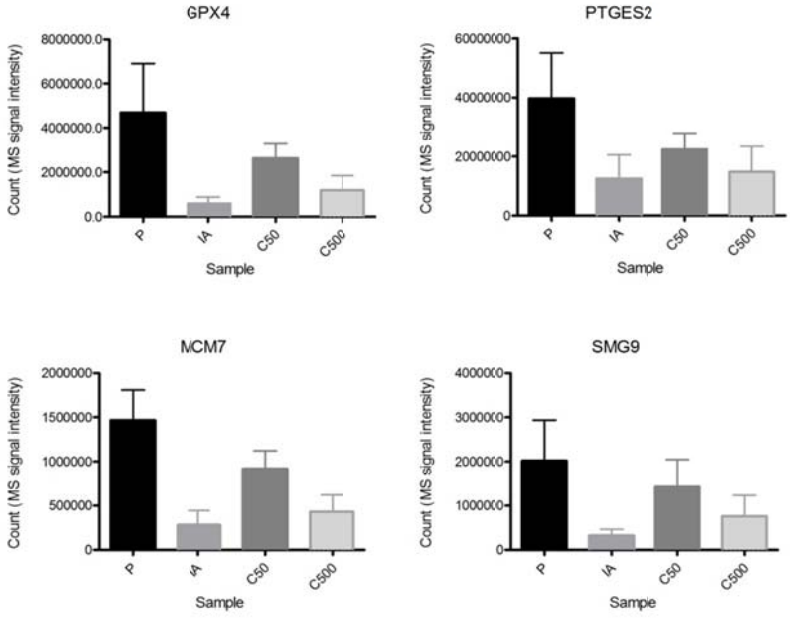


Figure 4.1: Comparison of candidate target protein levels of top hits found in 'probe', 'inactive' and 'competitor'-treated samples.

Protein levels are indicated by the intensity of the mass spectrometry signal for that particular protein. A pattern where protein levels are highest in the probe-treated sample (**P**), and consecutively lower in the competitor-treated samples at 50nM and 500nM (**C50** & **C500**) was used to distinguish between potential 'real' targets and non-specific targets. Additionally these targets were present at low levels in the inactive probe-treated samples (**IA**) Data represent the mean of three biological and three technical replicates ±SD.

4.2b The role of GPX4 in RSL3-induced cell death

Glutathione peroxidase 4 (GPx4) is a member of a family of eight glutathione peroxidases found in mammals [10]. GPxs function as a part of the cellular antioxidant defense network to reduce hydrogen peroxide, organic hydroperoxides, and lipid peroxides through the use of glutathione. GPx4 is unique amongst its family members in four ways: (1) it is the only GPx isoform that can reduce lipid peroxides (2) it doesn't essentially require glutathione and can use protein thiols as a reducing agent (3) it exists as a monomer while the others are tetramers and (4) it is essential for life as knockout mice result in embryonic lethality at day 8. It is also one of twenty-five known selenoproteins in the human genome, containing a selenocysteine in its active site that is essential for its catalytic activity [10]. Interestingly, two other selenoproteins were also among the top twenty candidates showing enrichment in the probe-treated samples. The presence of many selenocysteine and cysteine containing proteins in this set may be indicative of the reactive nature of RSL3 due to its electrophilic group.

4.2c Lipid peroxidation is involved in RSL3-induced death

Upon first inspection of the candidate targets, GPx4 immediately stood out because of its key role in the antioxidant system of the cell and in particular, for being the only glutathione peroxidase that reduces lipid peroxides. Investigation of the phenomenon of glutathione depletion in erastin-induced death in our laboratory had already focused on reports in the literature of the connection between glutathione depletion and activation of lipoxygenases [11, 12]. In mammalian cells, lipoxygenase (LOX) enzymes dioxygenate arachidonic acid, a fatty acid involved in cellular signaling and the inflammatory response, to form the reactive lipid peroxide hydroperoxyeicosatetraenoic acid (HPETE), a precursor in leukotriene biosynthesis. LOX activation requires oxidation by basal levels of lipid peroxides, and thus the activation and the products of these enzymes are reduced by GPx4 activity. Additionally, studies in GPx4-deficient MEFs have shown that loss of GPx4 leads to unregulated activity of lipoxygenases that

unleashes massive lipid peroxidation that results in cell death [12]. Cyclooxygenases (COX), which convert arachidonic acid to PGH_2, a precursor of other eicosanoids, are also regulated by GPx4.

Numerous inhibitors of both the LOX and COX enzymes exist and some of these were tested for their effects on RSL3-induced cell death (Figure 4.2). Co-treatment of all LOX inhibitors tested showed significant suppression of death by RSL3 while a broad-range COX inhibitor, indomethacin, showed slight sensitization. ROS was already identified to play a major role in RSL3's mechanism of action. These data implicate the source of ROS as lipid peroxides formed by lipoxygenases. Preliminary data using lipid peroxide-sensitive dyes in flow cytometry assays show formation of lipid peroxides in cells upon RSL3 treatment. LOX and COX inhibitor profiling was performed by Dr. Wan Seok Yang.

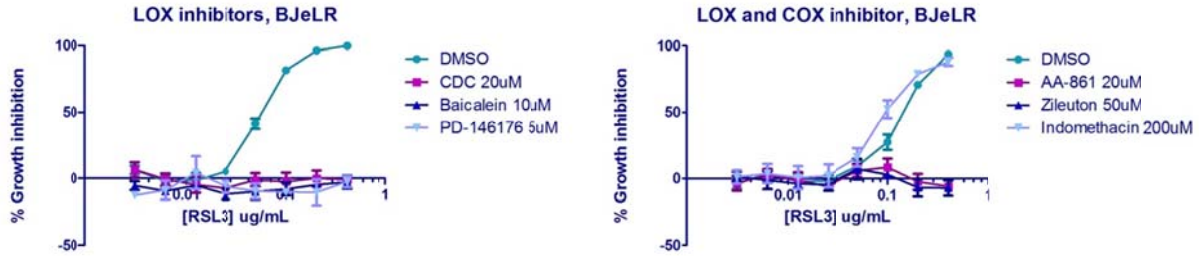


Figure 4.2: Lipoxygenase inhibitors rescue cells from RSL3-induced death

(A) A range of lipoxygenase (LOX) and one cyclooxygenase (COX) inhibitor were co-treated with RSL3 and effects on growth inhibition were measured. All lipoxygenase inhibitors (CDC, Baicalein, PD-146176, AA-861 and Zileuton) were observed to rescue cells from RSL3 treatment, while the cyclooxygenase inhibitor (indomethacin) had no effect. Thus lipid peroxides produced by lipoxygenases, which are regulated by GPx4, could be involved in RSL3-induced death. Data represent the mean of three replicates ±SD.

4.2d RNAi knockdown of *GPx4* sensitizes cells to RSL3-induced death

Modulating the expression of candidate targets via RNAi knockdown and cDNA overexpression and observing the effects on cell sensitivity to the small molecule is a standard way to initially verify small molecule-protein interactions. To test the effect of depleting *GPx4* cellular target levels on RSL3 sensitivity we performed viral shRNA knockdowns of *GPx4*. Quantitative PCR was performed to check the efficacy of the knockdown (Figure 4.3). RNAi silencing of *GPx4* showed sensitization to RSL3 treatment (Figure 4.3). Sensitization by knockdown of *GPx4* suggests that RSL3 may be targeting the GPx4 protein through a loss-of-function mechanism. If RSL3 causes death through the functional loss of this protein, depletion of GPx4 alone might be expected to show a decrease in cell viability.

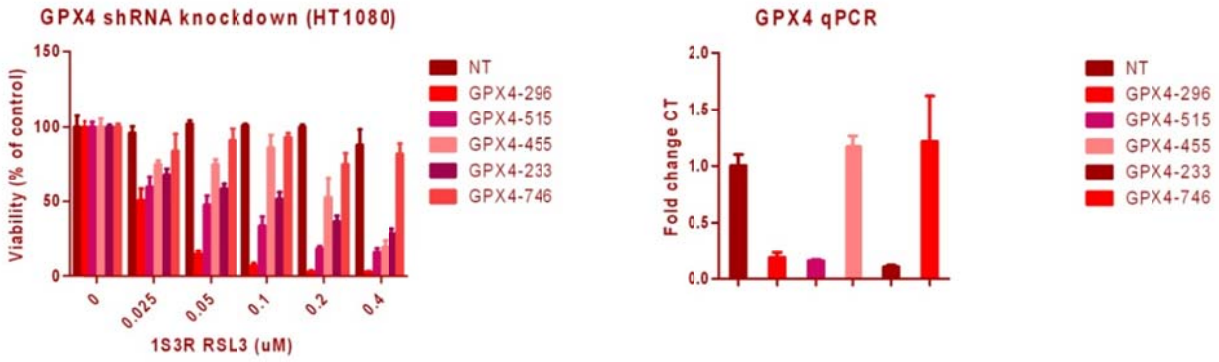


Figure 4.3: RNAi knockdown of *GPx4* sensitizes cells to RSL3-induced death
(A) Viral shRNA knockdowns of *GPx4* alongside non-targeting controls (NT) were performed in HT1080 cells and their effects on RSL3 sensitivity were assessed. Knockdown of *GPx4* resulted in sensitization to RSL3 treatment. Data represent the mean of three replicates ±SD.

4.2e General sensitization effects of *GPx4* knockdown

GPx4 is an important antioxidant enzyme in the cell and we were curious whether depletion of this protein would have effects on other oxidative *RAS*-selective lethal compounds or compounds known to kill through other mechanisms. To examine the generality of the sensitization effects of knockdown of *GPx4* we treated viral shRNA infected cells with a range of compounds with similar *RAS*-selective lethal mechanisms as well as other lethal mechanisms. Knockdown of GPx4 showed a pattern of general sensitization to *RAS*-selective lethal compounds tested (Figure 4.4). Lipid peroxides have already been implicated in erastin-induced cell death and modulatory profiling of ML162 has shown that its profile clusters strongly with RSL3's profile, suggesting similarity in their mechanisms of action (Shimada, Stockwell; unpublished results). This compound also contains a chloromethylketone functionality although its general scaffold is different. (ML162 was identified in *RAS*-selective lethal screening performed in collaboration with the Broad institute [13]). Another more recently identified RSL molecule, CIL56, has been shown to induce lipid peroxides rapidly (Shimada, Stockwell; unpublished results). These data collectively suggests that lipid peroxides are involved in the non-apoptotic mechanism of death induced by these compounds, perhaps through perturbation of GPx4.

Although not as prominent, slight sensitization effects are observed with compounds of other lethal mechanisms (on average a 40% decrease in viability when compared with cells infected with non-targeting virus) particularly trichostatin A (TSA), a histone deacetylase inhibitor. One report has observed that trichostatin A's inhibition of histone deacetylase 2 (HDAC2) causes instability of Nrf2, a transcription factor that induces expression of multiple-antioxidant genes [14]. Due to its role in antioxidant defense, GPx4 depletion could result in sensitization to any compound that kills partly through the generation of ROS or by affecting the antioxidant network.

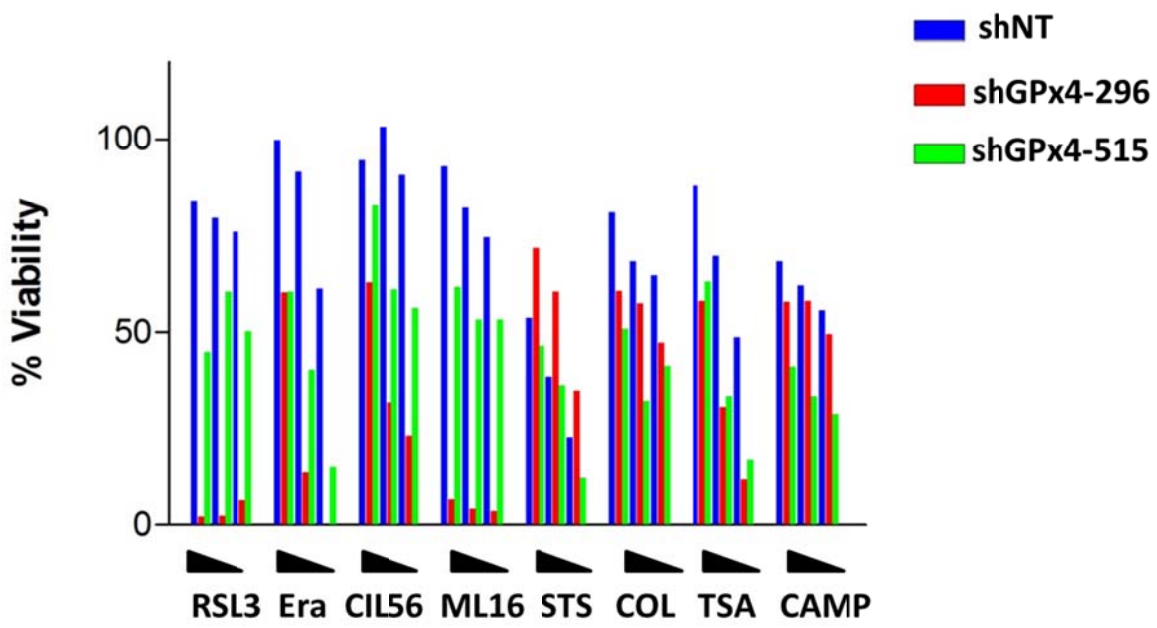


Figure 4.4: Sensitization by knockdown of *GPx4* to other lethal compounds
HT1080 cells infected with shRNA targeting *GPX4* and non-targeting controls (NT) were treated with a range of lethal compounds (**RSL3**, **Era** = erastin, **CIL56**, **ML16**2, **STS** = a kinase inhibitor, staurosporine, **COL** = a microtubule destabilizer, colchicine, **TSA** = an HDAC inhibitor, trichostatin A, **CAMP** = a topoisomerase I inhibitor) at decreasing doses (black triangle). Compounds were treated for 48 hours. Knockdown of GPx4 showed a pattern of general sensitization to *RAS*-selective lethal compounds tested. Data represent the mean of three replicates.

4.2f RNAi knockdown of *GPx4* does not phenocopy RSL3 treatment

The effect of viral shRNA knockdown of *GPx4* on cell viability was examined to see whether protein depletion mimicked the possible inhibition of GPx4 by RSL3. We observed no effect of knockdown of this candidate on cell viability by viral shRNA infection of HT1080 cells or by treatment of the four BJ cell series with *GPx4* siRNA (Figure 4.5). The implication is that loss of GPx4 is not sufficient to induce RSL3's mechanism of lethality and that RSL3 may act through multiple targets simultaneously. Thus knockdown of multiple targets could be required to phenocopy RSL3. Alternatively, GPx4 may be downstream of an actual target of RSL3 in a relationship that enhances its toxicity, and it was detected through its interactions with this target. Protein interaction databases were searched to explore the possibility of direct interactions between top candidate targets or the existence of a common interactor however this did not reveal any connections. The siRNA experiment with GPx4 and quantitative PCR were performed by Dr. Wan Seok Yang and Darnelle Delva.

A

B

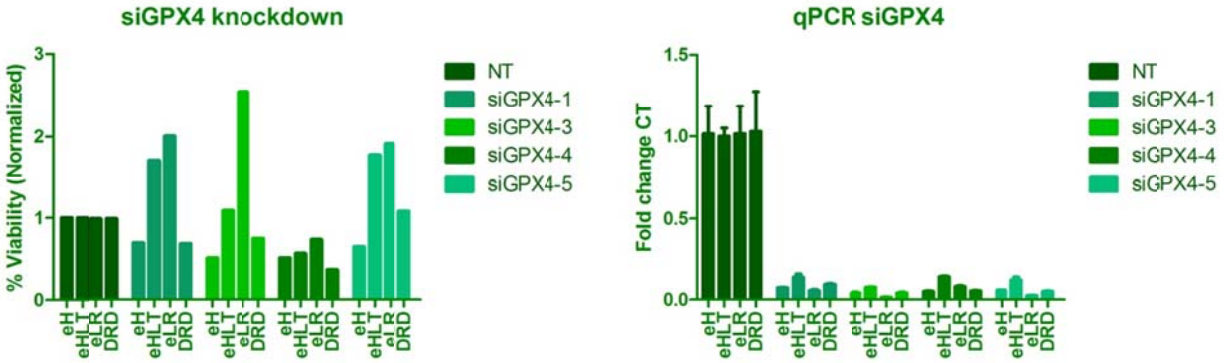


Figure 4.5: RNAi knockdown of *GPx4* does not cause loss of viability

(A) HT1080 cells were infected with viral shRNA clones targeting *GPx4* and non-targeting controls (NT) and viability was measured six days later using trypan blue exclusion.

(B) The four cell lines in the BJ series were transfected with siRNA targeting *GPx4* and non-targeting controls (NT) and effects on their viability were determined by trypan blue exclusion. Knockdown with the siRNA clones was verified by quantitative PCR (data represent the mean of three replicates ±SD).

4.2g The role of PTGES2 in RSL3-induced cell death

Inhibition of GPx4 alone did not appear to be sufficient to explain RSL3's mechanism of lethality, as evident in the experiments where knockdown alone of this target showed no effect on cell viability. RSL3 was shown to enrich multiple candidate targets that were identified through affinity purification experiments. Thus, we considered the possibility that multiple targets need to be simultaneously perturbed to induce the *RAS*-selective lethal death observed with RSL3.

On the surface, the top candidates did not seem directly connected, however GPx4 and prostaglandin E synthase 2 (PTGES2) function on close pathways involved in the metabolism of arachidonic acid described above. PTGES2 is a membrane-associated protein and one of three isozymes in the eicosanoid biosynthetic pathway that converts the unstable intermediate prostaglandin PGH_2 to PGE_2. It functions in conjunction with the cyclooxygenase enzymes (COX) that perform the initial step converting arachidonic acid into PGH_2. PTGES2's product, PGE_2, is a key mediator of the inflammatory response through secretion and binding to the E-prostanoid receptors [15].

The role of PTGES2 in cancer is mainly attributed to the effects of its product, PGE_2, a protumoral eicosanoid that has been found in addition to its pro-inflammatory effects, to enhance tumor cell proliferation, invasiveness, angiogenesis and to inhibit apoptosis – all hallmarks of cancer [16]. The importance of PGE_2, particularly in colorectal cancer, has been deduced from its correlation with increased expression of its upstream enzyme (COX-2) and levels of PGE_2 themselves [17, 18]. It is believed to exert these effects through downstream E prostanoid receptor signaling networks whose targets include PI3K/Akt , Ras-MAPK, vascular endothelial growth factor (VEGF) and Bcl-2 pathways [19]. Sequestration of PGE_2 by a specific monoclonal antibody in a xenograft model of lung cancer in mice retarded tumor growth [20].

Given the evidence of the connection particularly with colorectal cancer, which is known to have a high prevalence of *RAS* mutations (~ 50%), we decided to investigate the involvement of PTGES2 in RSL3's mechanism [21, 22].

4.2h RNAi knockdown of *PTGES2* sensitizes cells to RSL3-induced death

To test the effect of depleting *PTGES2* cellular target levels on RSL3 sensitivity we performed viral shRNA knockdowns of *PTGES2*. Surprisingly, RNAi silencing of *PTGES2* also showed sensitization to RSL3 treatment (Figure 4.6). Sensitization by knockdown of *PTGES2* suggests that RSL3 may be targeting both GPx4 and PTGES2 in order to induce cell death. If RSL3 causes death through the functional loss of both of these proteins, it could explain why depletion of GPx4 alone did not show a decrease in cell viability.

4.2i PTGES2 is present in fluorescein RSL3-treated cells prepared for pull downs

New pull down samples with the fluorescein-RSL3 probes were prepared on a smaller scale to verify the presence of GPx4 and PTGES2 in elutions from probe, inactive and competitor treated cells by western blot (Figure 4.7). An enrichment pattern of PTGES2 could be observed in elutions analogous to that seen in the proteomic results. These data provide a confirmation that RSL3 indeed interacts with PTGES2. Experiments to detect the presence of GPx4 and other candidate targets in pull down elutions have been hampered by lack of effective antibodies, however efforts are ongoing.

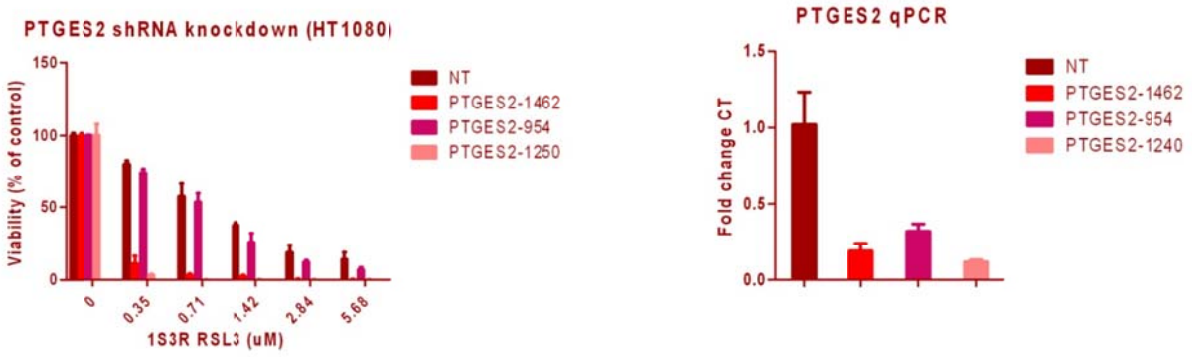


Figure 4.6: RNAi knockdown of *PTGES2* sensitizes cells to RSL3-induced death
(A) Viral shRNA knockdowns of *PTGES2* alongside non-targeting controls (NT) were performed in HT1080 cells and their effects on RSL3 sensitivity were assessed. Knockdown of *PTGES2* resulted in sensitization to RSL3 treatment. Data represent the mean of three replicates ±SD.

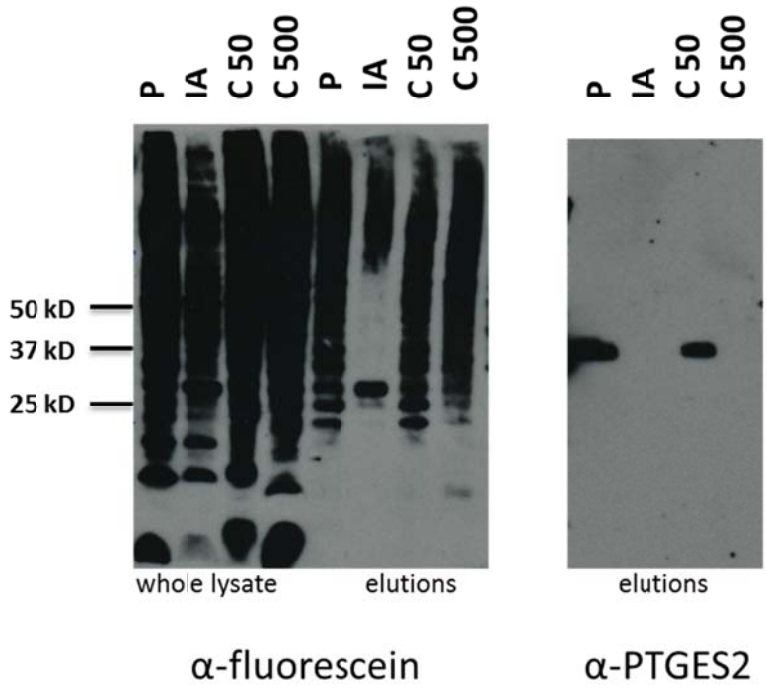


Figure 4.7: PTGES2 is enriched in elutions of pull downs from probe-treated samples

Western blot showing (1S,3R) fluorescein-RSL3 probe (**P**), (1R,3R) fluorescein-RSL3 probe (**IA**), and competitor (50nM = **C50**, 500nM = **C500**)- treated samples detected by α-fluorescein (left) or α-PTGES2 (right). Western-blot detection with α-PTGES2 (right) shows the presence of PTGE2 in elutions of **P** and **C50** treated samples and its absence in **IA** and **C500**. The same elutions were probed with α-fluorescein (left, right four lanes), there does not appear to be proteins in the corresponding position in the gel of IA-treated sample, however these samples are not as concentrated as the original whole lysate (left, left four lanes). Samples were prepared as described in Chapter 2, on a reduced scale.

4.2j General sensitization effects of *PTGES2* knockdown

The generality of the sensitization effects of knockdown of *PTGES2* were also examined by treating viral sh*PTGES2* infected cells with the same panel of compounds with similar *RAS*-selective lethal mechanisms as well as other lethal mechanisms. The sensitization effect of PTGES2 knockdown was observed to be relatively unique and strong in RSL3 (and ML162)-induced death (Figure 4.8). Effects on other RSL compounds, such as erastin were not as great and only slight with CIL56. Interestingly, cell sensitivity to camptothecin also showed a reasonable effect. These data demonstrate a more unique sensitization with knockdown of *PTGES2* in RSL3-induced death, which suggests it may be a distinguishing feature of RSL3's mechanism in comparison with other RSL compounds we have identified.

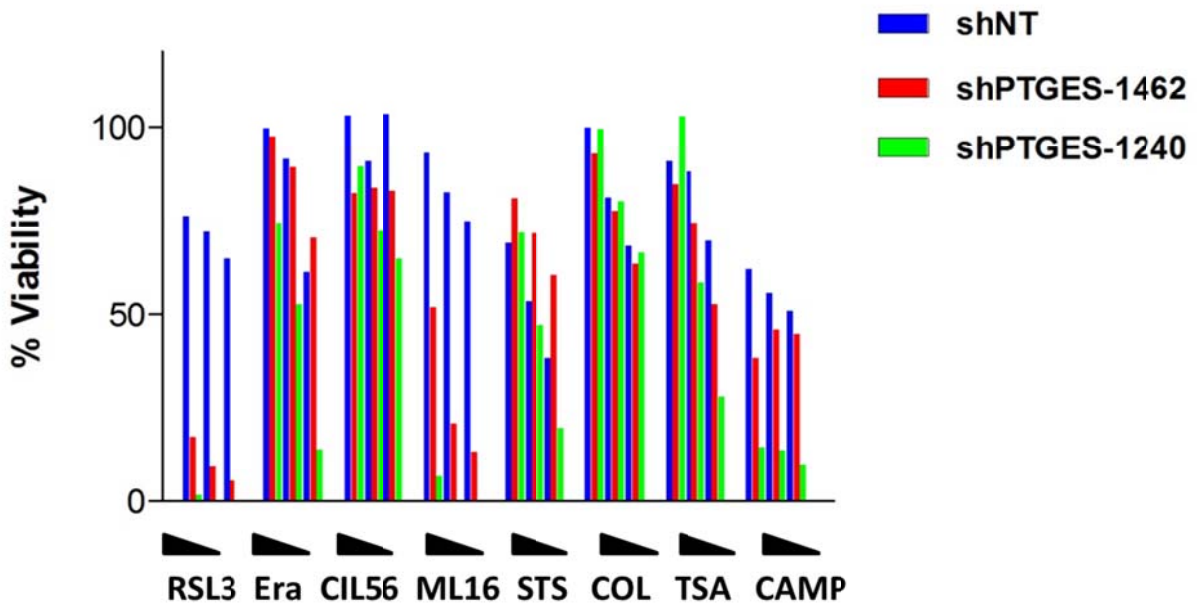


Figure 4.8: Sensitization by knockdown of *PTGES2* to other lethal compounds
HT1080 cells infected with shRNA targeting *PTGES2* and non-targeting controls (NT) were treated with a range of lethal compounds (**RSL3**, **Era** = erastin, **CIL56**, **ML162**, **STS** = a kinase inhibitor, staurosporine, **COL** = a microtubule destabilizer, colchicine, **TSA** = an HDAC inhibitor, trichostatin A, **CAMP** = a topoisomerase I inhibitor) at decreasing doses (black triangle). Compounds were treated for 48 hours. Knockdown of *PTGES2* showed a unique pattern of sensitization to RSL3 and a related analog ML162. Sensitization to camptothecin was also noticeable. Data represent the mean of three replicates.

4.2k RNAi knockdown of *PTGES2* does not phenocopy RSL3 treatment

The effect of viral shRNA knockdown of *PTGES2* on cell viability was also examined to see whether protein depletion could phenocopy RSL3 treatment in cells. As expected from our hypothesis of a multi-target mechanism, we observed little or no effect of knockdown of *PTGES2* on cell viability by viral shRNA infection of HT1080 cells (Figure 4.9). Knockdown of both or even additional targets could be required to phenocopy RSL3's mechanism, emphasizing the power of small molecules in modulating multiple targets concurrently. Knockdown of both *GPX4* and *PTGES2* simultaneously was attempted by viral infection with multiple shRNA clones and no effect on viability was observed, however the knockdown of both of these genes was not confirmed by qPCR or western blot and thus these results are inconclusive at this stage.

4.2l Overexpression of *PTGES2* does not rescue cells from RSL3-induced death

Next we looked at the effects of overexpression of the candidate targets on cell sensitivity to RSL3. Overexpression of a cDNA construct of *PTGES2* was performed successfully (as confirmed by western blot), however no suppression effects of RSL3-induced death were observed (Figure 4.10). The lack of effect was puzzling as the sensitization effect of these knockdowns had been very strong. One possible explanation is that PTGES2 is acting as part of a complex. Depletion of one protein from the complex would inhibit its function, but overexpression of one component would not enhance the complex's activity. Another explanation is that PTGES2 modulates the death in a way that can only enhance it, but it is not ultimately causal in death and therefore cannot prevent it. Additional experiments to overexpress more than one protein at the same time could be performed. Transfection of cDNA encoding two GFP-tagged transcript variants and one DDK-tagged variant of *GPx4* was attempted, but no overexpression was observed by microscopy or through western-blot analysis. Technical issues need to be overcome in order to investigate this further.

Figure 4.9: RNAi knockdown of *PTGES2* does not cause loss of viability

(A) HT1080 cells were infected with viral shRNA clones targeting the *PTGES2* and non-targeting controls (NT) and viability was measured six days later using trypan blue exclusion.

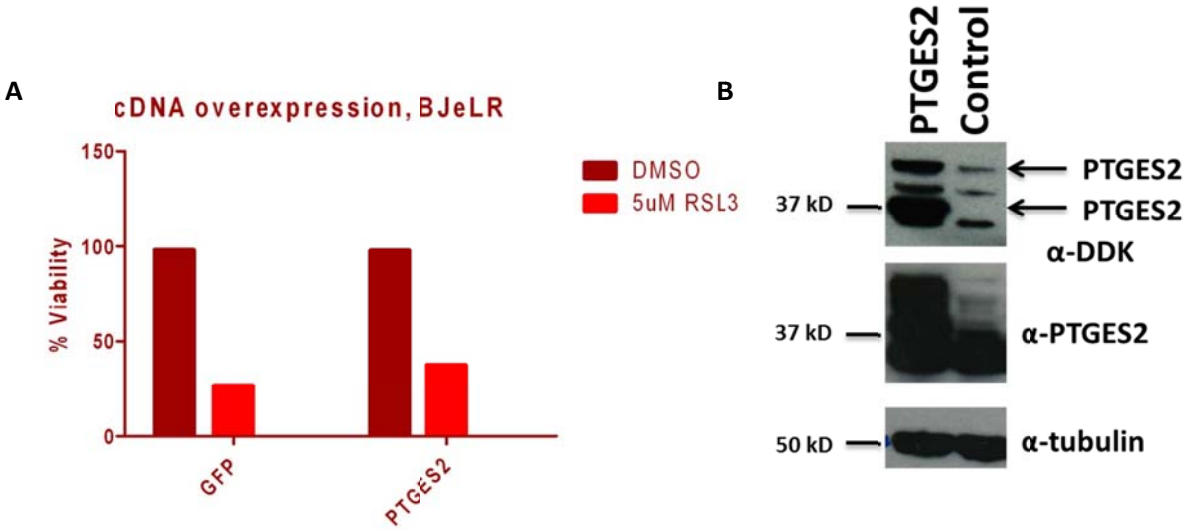


Figure 4.10: Overexpression of *PTGES2* does not rescue cells from RSL3-induced cell death

(A) BJeLR cells were transfected with a DDK-tagged cDNA construct of *PTGES2.* No significant suppression effects were observed on RSL3 treatment (5uM). Cells were transfected for 36 hours prior to compound treatment. Viability was measured 24 hours after treatment with RSL3 or vehicle using trypan blue exclusion.

(B) Western blot verification of cDNA overexpression.

4.2m RSL3's mechanism of action

RSL3's lethality is sensitized by depletion of both GPx4 and PTGES2, targets that were identified through affinity purification experiments. In general, GPx4 depletion is observed to sensitize to *RAS*-selective lethal compounds implicating a role for lipid peroxides as a species of ROS important in their death phenotype. Additionally, inhibitors of lipoxygenases directly regulated by GPx4 are able to rescue RSL3-induced death suggesting an excess of lipid peroxides are responsible for cell death, an effect which has been observed in other studies. Evidence that PTGES2 directly binds to RSL3 is further confirmed by its presence in elutions from pull down samples prepared with the fluorescein-RSL3 probe. The product of PTGES2, PGE_2, has a firmly established connection with colorectal cancer resulting in multiple clinical efforts to inhibit its upstream enzyme, COX [16]. We hypothesize that both targets must be simultaneously perturbed to induce the *RAS*-selective lethal death observed with RSL3 and depletion of any target alone cannot reproduce RSL3's mechanism of action. Intuitively this makes sense. A complex cancerous network designed to evade the normal regulation of the cellular machinery requires a complex solution to kill it.

By modulating the arachidonic acid metabolic pathway, RSL3 could concurrently allow activation of LOX and COX and induce formation of lipid peroxides through inhibition of GPx4, at the same time blocking PGE_2 production through inhibition of PTGES2 (Figure 4.11). COX activation increases prostaglandin synthesis, however PTGES2 inactivation results in decreased prostaglandin synthesis, particularly the pro-tumoral prostaglandin PGE_2. Thus both PGE_2 reduction and lipid peroxidation appear to be required for RSL3's mechanism of action. Reduction of GPx4 levels by RNAi induces lipid peroxidation, but still allows COX activation, which increases PGE_2 production. Reduction of PTGES2 levels by RNAi would be expected to lower PGE_2 production but have no effect on lipid peroxide formation. Only RSL3 could theoretically induce both effects simultaneously. Death induced by lipid peroxides could explain why LOX inhibitors were effective at rescuing cell death caused by RSL3 but the

COX inhibitor, indomethacin, showed no suppression effect. Inhibition of COX would be redundant if RSL3 is already blocking PGE_2 production and instead sensitization might be observed, perhaps at higher doses. This does not explain why overexpression of PTGES2 showed no suppression effects, this aspect will need to be investigated further. One study has shown that the growth of certain tumors are sensitive to the precise levels of PGE_2. In colorectal adenomas, low concentrations of PGE_2 resulted in stimulation of growth while high concentrations were inhibitory [23]. Overexpression experiments may not capture the dynamic range of this effect. The proposed model also ignores potential contributions from other RSL3 targets that may further add to the complexity of the mechanism. Using RSL3 as a probe, we have potentially identified novel mechanisms to target oncogenic *RAS* dependence through the loss of function of GPx4 and PTGES2.

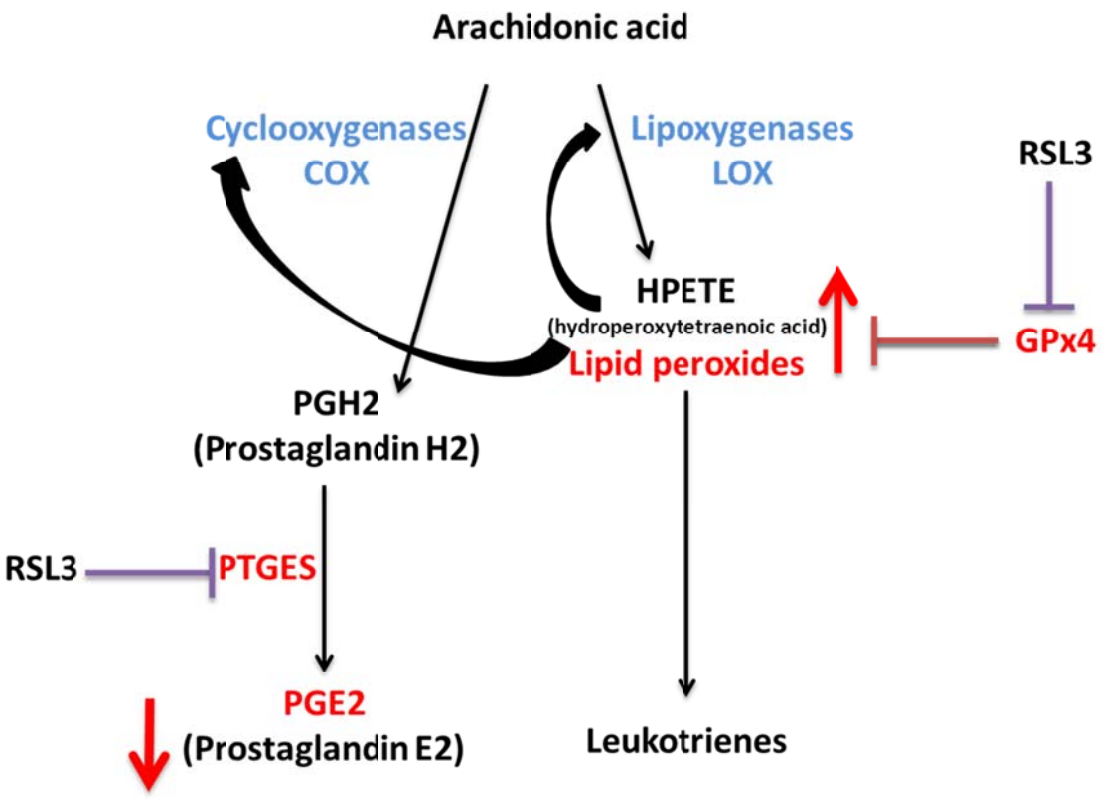


Figure 4.11: A potential model for RSL3's mechanism of action involving the arachidonic acid pathway
GPx4 normally functions to reduce lipid peroxides that are required basally to activate LOX and COX. Upon depletion or inhibition of GPx4, LOX and COX activity are reduced and the resulting lipid peroxide increase creates damage in the cell. In combination with PGE_2 reduction, which decreases pro-growth and invasiveness signaling, these effects are lethal.

4.3 Discussion

The roles of PTGES2 and GPx4 in cancer and cell death

As PGE_2 is a major downstream product of the COX enzymes, the use of a class of general COX inhibitors, nonsteroidal anti-inflammatory drugs (NSAIDS), has been of interest as chemopreventative therapeutics in cancer due to their negative regulation of PGE_2 production [24]. Regular use of NSAIDS has been shown to reduce the incidence of colorectal cancer [19]. However, adverse side effects from prolonged use are believed to be due to inhibition of the constitutively expressed COX-1, while antitumor and anti-inflammatory effects are derived from inhibition of the inducible COX-2. To resolve this problem selective COX-2 inhibitors were developed and approved by the FDA as a chemopreventative treatment of colorectal polyp formation in familial adenomatous polyposis [25]. Unfortunately, prolonged use of these inhibitors was found to have complications and increase cardiovascular risk [26]. A further point of note is that mutant K-ras and mutant β-catenin have been shown to increase COX-2 expression, suggesting a mechanistic link between COX activity and oncogenic transformation [27]. Measuring levels of COX-2 and PGE_2 in our engineered tumor cells with oncogenic *RAS* are important points to investigate in the future.

It has been proposed that drugs that specifically reduce levels of PGE_2 or inhibit PGE_2 downstream signaling pathways could be useful in chemoprevention of cancer [19, 23] . Efforts have been made to target PTGES1 (also called mPGES-1, which is considered to be the main prostaglandin E synthase), an isozyme of PTGES2, as alternatives to COX enzymes and consequently inhibitors of this protein have been developed and described [28, 29]. Doubts however have been raised about the potential side effects of the diversion of its substrate PGH_2 towards other eicosanoid synthesis [30, 31]. Ultimately studies have shown the importance of PTGES along with COX2 in cancer [32, 33]. Inhibition of PTGES2 may provide an alternative route to reducing levels of PGE_2.

Interestingly, GPx4's role in the regulation of COX and LOX was used in a study to reduce tumor growth or invasiveness through the inhibition of PGE_2 synthesis. Through specific overexpression of GPx4 in two tumorigenic mouse models they were able to inactivate COX, suppress PGE_2 production and either block or lower the malignancy of the tumor depending on its tumorigenicity [34]. Curiously, by potentially inhibiting GPx4, RSL3 could be doing the opposite but would still block PGE_2 synthesis through inhibition of PTGES2.

GPx4's various isoforms have been shown to play specific roles in the cell. The cytosolic form has been shown to regulate the production of eicosanoids through LOX and COX activation while the mitochondrial form has been shown to inhibit the intrinsic pathway of apoptosis by reducing cardiolipin hydroperoxide and thereby suppressing release of cytochrome c [35, 36]. The role of GPx4 has also been of interest in neuronal cell death [37]. It will be interesting to see whether the cytosolic isoform of GPx4 plays the major role in RSL3-induced cell death and whether the mitochondrial form is involved in this cell death mechanism.

To identify new therapeutic targets selective towards oncogenic *RAS*, we embarked on a project to identify the biological targets of RSL3, a molecule identified in the Stockwell laboratory with potent *RAS*-selective lethality that appeared to act through a novel mechanism. Through the synthesis of affinity reagents of RSL3 we discovered the stereochemistry of the active RSL3 isomer, which enabled the creation of many tagged RSL3 probes. Although initial affinity chromatography experiments were unfruitful, we ultimately found a method that identified a number of promising candidate targets. Two of these candidates, GPX4 and PTGES2, appear connected with RSL3's mechanism of action through multiple pieces of evidence. RNAi knockdown of GPX4 and PTGES2 shows strong sensitization to RSL3 treatment, suggesting that RSL3 modulates them through a loss-of-function mechanism. In agreement with this, RSL3 is known to induce lipid peroxidation that is normally regulated by GPx4, and inhibitors of

lipoxygenases (a main source of lipid peroxides) can suppress RSL3's death phenotype. Additionally, the presence of PTGES2 in samples prepared with RSL3 by affinity chromatography was detected in a pattern that resembled its enrichment in the proteomic analysis. Thus, these data suggest that a multi-targeting strategy disabling the cellular antioxidant network through GPx4 and reducing prostaglandin E2 synthesis through PTGES2 could be useful mechanisms to target cancers with oncogenic *RAS.*

4.4 Methods

4.4a Cell culture

HT-1080, 293-T and U2OS cells were obtained from American Type Culture Collection. HT-1080 cells were grown in DMEM High-Glucose media (Gibco) supplemented with 10% fetal bovine serum (FBS) and 1% non-essential amino acids (Gibco). 293-T cells were grown in DMEM High-Glucose supplemented with 10% FBS. When used for transfections to generate virus, 293-T cells were seeded in the above media lacking antibiotics. 293-T viral collection media contained 30% HyClone FBS. BJ/TERT/LT/ST/RASV12 (BJeLR) and related isogenic cell lines (eH, eHLT, DRD) were obtained from William Hahn (Harvard University). BJ series cells were grown in DMEM plus 20% M199 and 15% heat-inactivated FBS. All media were supplemented with penicillin and streptomycin (Gibco). All cell lines were grown in humidified tissue culture incubators (Thermo Scientific) at 37°C with 5% CO_2.

4.4b Viability measurements

Viability was assessed using one of two viability dyes: Alamar blue fluorescence or Trypan Blue dye exclusion. Alamar blue (Invitrogen) is a cell-permeable resazurin-based dye that is reduced by metabolically active cells to rezorufin, a dye with red fluorescence. Fluorescence was measured 6-16 hours after alamar blue addition to cells using a Victor3 platereader with excitation filter at 530 nm and emission filter centered on 590 nm. Trypan Blue is a viability dye that is excluded from live cells but permeates dead cells. Trypan Blue negative and positive cells were counted using an automated cell counter (ViCell, Beckman-Coulter).

4.4c Lentiviral shRNA production

Virus production followed the protocol of The RNAi Consortium. For production of virus, 293-T cells were seeded in 6-well dishes in antibiotic free media (280,000 cells/well). The next day, cells were transfected using FuGENE with shRNA encoding plasmid (450 ng), viral packaging plasmid (p-Delta8.9,

400 ng), and viral envelope plasmid (p-VSV-G, 45 ng). After 24 hours, media was removed and replaced with 1.6 ml Viral Collection Media. A total of 3 collections of viral supernatant in VCM per shRNA were made over 36 hours and pooled. Pooled supernatant was centrifuged (2,250 rpm, 2 min), aliquoted and stored at -80°C.

4.4d Lentiviral shRNA Infection

For lentiviral infections in 384-well format, 400 cells were seeded per well the night before the infection. On the day of the infection, media was replaced with 38 ul media + polybrene (8 ug/ml) and 2 ul of the viral supernatant per well. Plates were centrifuged at 2,250 rpm at 37°C for 90 minutes in a tabletop Sorvall centrifuge. Plates were incubated for 48 hours, the media was replaced with media containing 1.5 ug/ml puromycin and incubated for 24 hours. Media was subsequently changed again for 40 ul of RSL3 or compound-containing media at the appropriate concentration. Plates were incubated for 24 hours and viability was assessed using alamar blue. For lentiviral infections in 6-well format, 35,000 cells were seeded per well in 2mL of media the night before the infection. Cells were infected in the presence of 8 ug/ml polybrene in 1.85mL with 150 ul viral supernatant. Puromycin selection, compound treatment were the same as above, but scaled for 6-well format. Viability was assessed by trypan-blue exclusion. Where possible, viability was also assessed by light microscopy.

4.4e Gene expression by RT-q-PCR

Cells from 6-well format infection were harvested following puromycin selection and RNA was purified using the QIAshredder and QIAgen RNAeasy extraction kits according to the manufacturer's instructions. 2 ug total RNA per sample were subsequently used in a reverse transcription reaction using the TaqMan RT Kit priming with Random Hexamers (Applied Biosystems). Primers for q-PCR were designed with either Primer Express. Quantitative PCR was performed on triplicate samples in 96-well format using Power SYBR Green Master Mix (Applied Biosystems) on an Applied Biosystems 7300 Cycler set to

absolute quantification. The change in expression of a gene between experimental and control conditions was computed using the DeltaDelta Ct method with *ACTIN* as an internal reference gene.

4.4f Inhibitor profiling experiments

RAS-selective lethal molecules (RSL3, erastin, CIL56 and ML162) were synthesized in our laboratory or at the Broad Institue. All other compounds used in the panel of lethal small molecules were purchased from Sigma.

4.4g cDNA overexpression

Cells were seeded in 6-well dishes at a density of 250,000 cells/well the day before transfection. The next day cells were transfected with 1ug of cDNA construct with 2uL of Lipofectamine 2000 in 500uL of OptiMEM that had been pre-mixed together for 20 min before drop-wise addition to the cells. After 6 hours the media was replaced with fresh media to lower toxicity from the lipofectamine reagent. All experiments were performed with a GFP cDNA construct alongside as a positive control and transfection efficiency was examined by microscopy or through western blot analysis (see methods Chapter 2.4f). Antibodies for DDK tag (Origene, TA100014), PTGES2 (Cayman chemical, 160145) and tubulin (Santa Cruz,sc-32293) were used at 1:2000 dilution and detected using a goat α mouse HRP conjugate secondary antibody (Santa Cruz, sc-2030) at 1:5000 dilution and enhanced chemiluminiscence.

4.5 References

1. Yang, W.S. and B.R. Stockwell, *Synthetic Lethal Screening Identifies Compounds Activating Iron-Dependent, Nonapoptotic Cell Death in Oncogenic-RAS-Harboring Cancer Cells.* Chem Biol, 2008. **15**(3): p. 234-45.
2. Yagoda, N., M. von Rechenberg, E. Zaganjor, A.J. Bauer, W.S. Yang, D.J. Fridman, A.J. Wolpaw, I. Smukste, J.M. Peltier, J.J. Boniface, R. Smith, S.L. Lessnick, S. Sahasrabudhe, and B.R. Stockwell, *RAS-RAF-MEK-dependent oxidative cell death involving voltage-dependent anion channels.* Nature, 2007. **447**(7146): p. 864-8.
3. Downward, J., *Targeting RAS signalling pathways in cancer therapy.* Nat Rev Cancer, 2003. **3**(1): p. 11-22.
4. Klekota, J., E. Brauner, F.P. Roth, and S.L. Schreiber, *Using high-throughput screening data to discriminate compounds with single-target effects from those with side effects.* Journal of Chemical Information and Modeling, 2006. **46**(4): p. 1549-1562.
5. Hopkins, A.L., *Network pharmacology: the next paradigm in drug discovery.* Nat Chem Biol, 2008. **4**(11): p. 682-90.
6. Couzin, J., *Cancer drugs. Smart weapons prove tough to design.* Science, 2002. **298**(5593): p. 522-5.
7. Frantz, S., *Drug discovery: playing dirty.* Nature, 2005. **437**(7061): p. 942-3.
8. Huang, P., D. Trachootham, and J. Alexandre, *Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach?* Nature Reviews Drug Discovery, 2009. **8**(7): p. 579-591.
9. Csermely, P., V. Agoston, and S. Pongor, *The efficiency of multi-target drugs: the network approach might help drug design.* Trends Pharmacol Sci, 2005. **26**(4): p. 178-82.
10. Savaskan, N.E., C. Ufer, H. Kuhn, and A. Borchert, *Molecular biology of glutathione peroxidase 4: from genomic structure to developmental expression and neural function.* Biol Chem, 2007. **388**(10): p. 1007-17.
11. Li, Y., P. Maher, and D. Schubert, *A role for 12-lipoxygenase in nerve cell death caused by glutathione depletion.* Neuron, 1997. **19**(2): p. 453-63.
12. Seiler, A., M. Schneider, H. Forster, S. Roth, E.K. Wirth, C. Culmsee, N. Plesnila, E. Kremmer, O. Radmark, W. Wurst, G.W. Bornkamm, U. Schweizer, and M. Conrad, *Glutathione peroxidase 4 senses and translates oxidative stress into 12/15-lipoxygenase dependent- and AIF-mediated cell death.* Cell Metab, 2008. **8**(3): p. 237-48.
13. Bittker, J.A., M. Weiwer, K. Shimada, W.S. Yang, L. MacPherson, S. Dandapani, B. Munoz, M. Palmer, B.R. Stockwell, and S.L. Schreiber, *Screen for RAS-Selective Lethal Compounds and VDAC Ligands.* 2010.
14. Mercado, N., R. Thimmulappa, C.M. Thomas, P.S. Fenwick, K.K. Chana, L.E. Donnelly, S. Biswal, K. Ito, and P.J. Barnes, *Decreased histone deacetylase 2 impairs Nrf2 activation by oxidative stress.* Biochem Biophys Res Commun, 2011. **406**(2): p. 292-8.
15. Park, J.Y., M.H. Pillinger, and S.B. Abramson, *Prostaglandin E2 synthesis and secretion: the role of PGE2 synthases.* Clin Immunol, 2006. **119**(3): p. 229-40.
16. Greenhough, A., H.J. Smartt, A.E. Moore, H.R. Roberts, A.C. Williams, C. Paraskeva, and A. Kaidi, *The COX-2/PGE2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour microenvironment.* Carcinogenesis, 2009. **30**(3): p. 377-86.
17. Elder, D.J., J.A. Baker, N.A. Banu, M. Moorghen, and C. Paraskeva, *Human colorectal adenomas demonstrate a size-dependent increase in epithelial cyclooxygenase-2 expression.* J Pathol, 2002. **198**(4): p. 428-34.

18. Pugh, S. and G.A. Thomas, *Patients with adenomatous polyps and carcinomas have increased colonic mucosal prostaglandin E2.* Gut, 1994. **35**(5): p. 675-8.
19. Wang, D. and R.N. Dubois, *Prostaglandins and cancer.* Gut, 2006. **55**(1): p. 115-22.
20. Dubinett, S.M., M. Stolina, S. Sharma, Y. Lin, M. Dohadwala, B. Gardner, J. Luo, L. Zhu, M. Kronenberg, P.W. Miller, J. Portanova, and J.C. Lee, *Specific inhibition of cyclooxygenase 2 restores antitumor reactivity by altering the balance of IL-10 and IL-12 synthesis.* Journal of Immunology, 2000. **164**(1): p. 361-370.
21. Bos, J.L., E.R. Fearon, S.R. Hamilton, M. Verlaan-de Vries, J.H. van Boom, A.J. van der Eb, and B. Vogelstein, *Prevalence of ras gene mutations in human colorectal cancers.* Nature, 1987. **327**(6120): p. 293-7.
22. Forrester, K., C. Almoguera, K.Y. Han, W.E. Grizzle, and M. Perucho, *Detection of High-Incidence of K-Ras Oncogenes during Human-Colon Tumorigenesis.* Nature, 1987. **327**(6120): p. 298-303.
23. Chell, S., A. Kaidi, A.C. Williams, and C. Paraskeva, *Mediators of PGE2 synthesis and signalling downstream of COX-2 represent potential targets for the prevention/treatment of colorectal cancer.* Biochim Biophys Acta, 2006. **1766**(1): p. 104-19.
24. Smalley, W.E. and R.N. DuBois, *Colorectal cancer and nonsteroidal anti-inflammatory drugs.* Adv Pharmacol, 1997. **39**: p. 1-20.
25. Steinbach, G., P.M. Lynch, R.K.S. Phillips, M.H. Wallace, E. Hawk, G.B. Gordon, N. Wakabayashi, B. Saunders, Y. Shen, T. Fujimura, L.K. Su, B. Levin, L. Godio, S. Patterson, M.A. Rodriguez-Bigas, S.L. Jester, K.L. King, M. Schumacher, J. Abbruzzese, R.N. DuBois, W.N. Hittelman, S. Zimmerman, J.W. Sherman, and G. Kelloff, *The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis.* New England Journal of Medicine, 2000. **342**(26): p. 1946-1952.
26. Solomon, S.D., J.J. McMurray, M.A. Pfeffer, J. Wittes, R. Fowler, P. Finn, W.F. Anderson, A. Zauber, E. Hawk, and M. Bertagnolli, *Cardiovascular risk associated with celecoxib in a clinical trial for colorectal adenoma prevention.* N Engl J Med, 2005. **352**(11): p. 1071-80.
27. Araki, Y., S. Okamura, S.P. Hussain, M. Nagashima, P. He, M. Shiseki, K. Miura, and C.C. Harris, *Regulation of cyclooxygenase-2 expression by the Wnt and ras pathways.* Cancer Res, 2003. **63**(3): p. 728-34.
28. Gosselin, F., S. Lau, C. Nadeau, T. Trinh, P.D. O'Shea, and I.W. Davies, *A Practical Synthesis of m-Prostaglandin E Synthase-1 Inhibitor MK-7285.* Journal of Organic Chemistry, 2009. **74**(20): p. 7790-7797.
29. Xu, D.G., S.E. Rowland, P. Clark, A. Giroux, B. Cote, S. Guiral, M. Salem, Y. Ducharme, R.W. Friesen, N. Methot, J. Mancini, L. Audoly, and D. Riendeau, *MF63 [2-(6-chloro-1H-phenanthro[9,10-d]imidazol-2-yl)-isophthalonitrile], a selective microsomal prostaglandin E synthase-1 inhibitor, relieves pyresis and pain in preclinical models of inflammation.* Journal of Pharmacology and Experimental Therapeutics, 2008. **326**(3): p. 754-763.
30. FitzGerald, G.A., M. Wang, W.L. Song, and Y. Cheng, *Microsomal prostaglandin E synthase-1 inhibition in cardiovascular inflammatory disease.* Journal of Internal Medicine, 2008. **263**(5): p. 500-505.
31. Scholich, K. and G. Geisslinger, *Is mPGES-1 a promising target for pain therapy?* Trends Pharmacol Sci, 2006. **27**(8): p. 399-401.
32. Nakagawa, H., J.J. Lee, M. Natsuizaka, S. Ohashi, G.S. Wong, M. Takaoka, C.Z. Michaylira, D. Budo, J.W. Tobias, M. Kanai, Y. Shirakawa, Y. Naomoto, A.J.P. Klein-Szanto, and V.H. Haase, *Hypoxia activates the cyclooxygenase-2-prostaglandin E synthase axis.* Carcinogenesis, 2010. **31**(3): p. 427-434.

33. Oshima, H., H. Itadani, H. Kotani, M.M. Taketo, and M. Oshima, *Induction of prostaglandin E2 pathway promotes gastric hamartoma development with suppression of bone morphogenetic protein signaling.* Cancer Res, 2009. **69**(7): p. 2729-33.
34. Heirman, I., D. Ginneberge, R. Brigelius-Flohe, N. Hendrickx, P. Agostinis, P. Brouckaert, P. Rottiers, and J. Grooten, *Blocking tumor cell eicosanoid synthesis by GP x 4 impedes tumor growth and malignancy.* Free Radic Biol Med, 2006. **40**(2): p. 285-94.
35. Imai, H. and Y. Nakagawa, *Biological significance of phospholipid hydroperoxide glutathione peroxidase (PHGPx, GPx4) in mammalian cells.* Free Radic Biol Med, 2003. **34**(2): p. 145-69.
36. Nomura, K., H. Imai, T. Koumura, M. Arai, and Y. Nakagawa, *Mitochondrial phospholipid hydroperoxide glutathione peroxidase suppresses apoptosis mediated by a mitochondrial death pathway.* J Biol Chem, 1999. **274**(41): p. 29294-302.
37. Savaskan, N.E., A. Borchert, A.U. Brauer, and H. Kuhn, *Role for glutathione peroxidase-4 in brain development and neuronal apoptosis: specific induction of enzyme expression in reactive astrocytes following brain injury.* Free Radic Biol Med, 2007. **43**(2): p. 191-201.

Chapter 5: Conclusions and Future Directions

This thesis presents the application of a number of methods to characterize the mechanism of action of a small molecule, RSL3, a potential anti-cancer agent selectively lethal in cells possessing oncogenic activation of the *RAS* pathway. Work performed to identify the direct targets of RSL3 led to the discovery of two potential targets, glutathione peroxidase 4 (GPx4) and prostaglandin E synthase 2 (PTGES2). Investigation into the role of GPx4 and PTGES2 in RSL3's lethal mechanism revealed the importance of lipid peroxidation and the potential involvement of the lipid metabolite prostaglandin E2 (PGE_2). Work performed on a new covalent version of a yeast three-hybrid (Y3H) system was aimed at determining whether it could provide an improvement over the existing state-of-the-art.

5.1 Identifying the biological targets of RSL3

5.1a Summary

To identify therapeutic targets for the treatment of cancer cells expressing oncogenic *RAS* pathway mutations, synthetic lethal screening in a series of isogenic cell lines was performed. Compounds that exhibited selective lethality to cells containing a mutant *HRAS* allele compared to an otherwise isogenic cell line with a wild-type *HRAS* allele were identified, one of them being a molecule that was called RSL3 [1]. This compound was of interest due to its potent lethality and studies that indicated a novel mechanism of action, thus we desired to characterize it further. The main approach taken was to directly identify the relevant binding partners of RSL3 in the cell through a method known as affinity chromatography. In order to achieve this, first an analog of RSL3 that could act as an affinity reagent for use in target identification studies was synthesized. The synthesis of multiple RSL3 'probes' revealed that only one stereoisomer of the molecule possessed *RAS*-selective lethality. The use of a fluorescein-tagged probe in an affinity chromatography experiment combined with shotgun proteomics

identified a number of promising candidate targets of RSL3. The top candidate, GPx4, was first investigated revealing that its role in reduction of lipid peroxides and inactivation of lipoxygenases (LOX) is likely connected to RSL3's mechanism of action. Loss of function of GPx4, induced by RSL3, would result in an excess of lipid peroxides that could contribute to the cell death. Indeed, we have observed production of lipid peroxides upon RSL3 treatment and inhibitors of the LOX enzymes are able to suppress RSL3-induced death. Further, the role of another potential target, prostaglandin E synthase 2 (PTGES2), was also implicated in RSL3-induced death. A direct enrichment of PTGES2 through affinity chromatography with the RSL3 probe was detected by Western blot. Inhibition of PTGES2 by RSL3 likely results in reduction of its product, the protumoral eicosanoid PGE_2, contributing to the death phenotype by removing activation of signaling pathways that promote growth and invasiveness. We hypothesize that dependency on these pathways could further weaken the cell to a lethal mechanism, such as an increase in lipid peroxides. Thus, we have determined that RSL3 could act as a multi-targeted drug causing the generation of lipid peroxides and inhibiting production of PGE_2, pathways that could be therapeutically targeted in the treatment of oncogenic *RAS.*

5.1b Significance

In order to find therapies that target cancer cells, it is important to identify their unique characteristics that can be exploited to cause their preferential death over normal cells. Studies of RSL3's mechanism of action identify two of these qualities that could be targeted, potentially in concert, elevated levels of ROS and PGE_2. A large amount of evidence has demonstrated that elevated levels of ROS are present in many types of cancer cells that are necessary for the promotion of cell proliferation and differentiation [2]. Consequently, in order to maintain redox homeostasis and prevent oxidative damage, cancer cells often adapt by enhancing their antioxidant networks; enzymes consisting of superoxide dismutases, peroxiredoxins, glutaredoxins, thioredoxins, catalase and glutathione peroxidases. Targeting of antioxidant enzymes by small molecules such as β-phenylethyl isothiocyanate

and piperlongumine have been shown to exhibit selective killing of cancer cells by exploiting their dependence on these networks [3, 4]. Uniquely, RSL3 appears to target not only this characteristic dependence, but also another pathway implicated in cancer, PGE_2 signaling. Due to the evidence that PGE_2 is involved in tumorigenesis and sustainment of tumor growth and invasiveness, a great deal of investigation of the use of COX-2 inhibitors to inhibit PGE_2 production has been conducted. Another method to inhibit PGE_2 production could be through the inhibition of PTGES, the enzyme that directly produces PGE_2. By targeting both PTGES2 and GPX4 simultaneously, RSL3 could be acting through multiple mechanisms of tumor-selective lethality. This agrees with an emerging shift in drug discovery where the use of multi-targeted drugs has been identified as necessary to overcome redundant cellular networks in disease such as cancer and to prevent the development of tumor resistance [5]. Additionally, it highlights the fact that many existing clinically-used drugs work through multiple targets simultaneously [6, 7].

5.1c Future directions

The main future directions for this project involve further defining the role of GPx4 and PTGES2 in RSL3's mechanism of action, particularly PTGES2. To address this, PGE_2 levels and their relationship to PTGES2 loss-of-function by sh*PTGES2* knockdown or RSL3 inhibition could be measured through the use of immunoassays. Additionally the effects on cell sensitivity to RSL3 upon addition of key metabolites in the arachidonic acid pathway, such as PGE_2, arachidonic acid and hydroperoxy eicosatetraenoic acid (HPETE) could also be explored. Addition of PGE_2 might be expected to rescue from RSL3's lethality while arachidonic acid and HPETE would be expected to sensitize by feeding into the generation of lipid peroxides. Inhibitors of an isozyme of PTGES2, PTGES1, exist and they could be used to compare the effects of inhibition of other prostaglandin E synthases on RSL3-induced death.

To further characterize GPx4's involvement in RSL3's mechanism, technical hurdles need to be addressed in order to gather data on the effects of overexpression of GPx4 and to detect the presence of GPx4 in proteins enriched by affinity chromatography with fluorescein-RSL3 as a confirmation of the proteomic results. The role of lipid peroxide generation and its relation to RSL3's activity can be defined through the use of flow cytometry assays with dyes that specifically detect lipid peroxides [8]. Treatment with the active stereoisomer of (1S,3R) RSL3 could be compared with the other stereisomers to observe effects specific to the active molecule for example. Additional functions, such as the role of the mitochondrial form of GPx4 in reducing cardiolipins can also be assessed by flow cytometry [9]. Lipoxygenase activity can also be specifically measured by monitoring levels of its product, HPETE, through cellular extraction and high performance liquid chromatography. Furthermore, the relevance of these events can be assessed *in vivo* through the use of mouse xenograft models. Tumors treated with RSL3 could be sectioned and examined with lipid peroxide-staining dyes.

The next major step would involve characterizing the connections of these pathways with oncogenic *RAS* signaling. All of the above methods could be employed to compare key events, such as lipid peroxide generation or changes in PGE_2 levels, upon treatment of RSL3 across the four BJ cell lines. Use of these cell lines provide a convenient test for identifying characteristics of RSL3-induced death specific to mutant *RAS.* Evidence already exists suggesting that mutant *RAS* in combination with mutant β-catenin increases COX-2 expression [10]. Endogenous levels of GPx4 and PTGES2 as well as COX-2 could be measured to see if cell sensitivity is related to their differential expression.

Following full confirmation of their role in RSL3's mechanism, direct binding of RSL3 with GPx4 and PTGES2 *in vitro* needs to be demonstrated. As illustrated in Chapter 1.5, a number of binding assays exist, such as surface plasmon resonance (SPR), isothermal titration calorimetry (ITC) or fluorescence anisotropy. Additionally, as GPx4 and PTGES2 are both enzymes, assays can performed to measure the

inhibition of their catalytic activity by (1S,3R) RSL3, using the other diastereomers as a control. Assays measuring GPx activity by a coupled reaction with glutathione reductase are well established and catalytic activity of PTGES2 has also been measured [11]. Crystal structures exist of both enzymes (for GPx4, only the U46C mutant), thus there is also the possibility of solving a co-crystal structure of the RSL3-bound complex to elucidate the binding site and gain insight into the mode of inhibition [12, 13]. Purification of GPx4 however, has been shown to be difficult and requires expression at low levels under anaerobic conditions to avoid oxidative polymerization of the enzyme [14].

Further, mouse embryonic fibroblasts that are conditionally deficient in GPx4 exist and could be used to compare RSL3's function in a *GPx4* null background [8]. Additionally they could be useful for the expression of binding site-defective mutants.

Finally, it will be important to investigate the potential roles of the other RSL3 targets identified in our affinity chromatography studies. As two targets already appear to be involved in RSL3's mechanism, it is possible that further targets could also contribute to the phenotype. Preliminary data show that knockdown of sh*MCM7* also shows sensitization to RSL3 treatment, although its major role in DNA replication appears to be unlikely to cause effects on the time scale of RSL3's lethality (see Appendix A.8)[15]. Co-treatment of cells with DNA synthesis inhibitors and RSL3 could determine whether this is a possibility as well as FACS analysis to observe effects on the cell cycle. Validation of SMG9, involved in nonsense-mediated mRNA decay, has been impeded by the lack of suitable shRNA clones, however siRNA and overexpression studies could be performed instead.

5.2 Comparing the transcription readout of the yeast three-hybrid system with covalent and non-covalent ligand-receptor pairs

5.2a Summary

The yeast three hybrid (Y3H) system is a transcriptional assay for the detection of small molecule or RNA-protein interactions *in vivo*. It is comprised of three hybrid molecules, two of which are fusion proteins consisting of (a) the DNA binding domain (DBD) of a transcription factor and the receptor protein of an anchor moiety and (b) the transcriptional activation domain (AD) and a second ligand binding protein. The third hybrid molecule is a chemical heterodimer containing the anchor moiety connected via a linker to a small molecule of interest whose target is unknown. The use of the Y3H system for target identification of small molecules has been clearly demonstrated, however it has still not been widely adopted [16, 17]. One reason is concern about the sensitivity of the assay to detect small molecule-protein interactions of greater than nanomolar affinity. We aimed to determine whether a covalent interaction between the anchor ligand and the DNA-binding domain-receptor fusion protein could improve transcriptional readout of the system, and thereby potentially improve sensitivity. In order to achieve this we created a version of the Y3H using an L28C mutant of *E. coli* dihydrofolate reductase (DHFR) that has been designed previously to form a covalent bond with a substrate of DHFR, trimethoprim (TMP), containing an acrylamide moiety [18]. The covalently interacting acrylamide-TMP Dex (A-TMP Dex) chemical inducer of dimerization (CID) was compared alongside the original methotrexate-dexamethasone (Mtx-Dex) CID developed in the Cornish laboratory as well as a TMP-Dex CID to directly compare the difference between non-covalent and covalent heterodimers. We found that while the covalent A-TMP Dex produced transcriptional activation greater than two fold higher than the non-covalent TMP-Dex, the Mtx-Dex CID still exhibited slightly higher fold activation in a β-galactosidase activity assay. Preliminary characterization of enrichment abilities of the

Mtx-Dex CID also showed comparable but greater enrichment than A-TMP Dex in a mock selection assay.

5.2b Significance

The use of A-TMP Dex represents the first use of a covalent ligand with DHFR in our Y3H system. In addressing the question of whether a covalent ligand-receptor pair could provide an improved transcriptional readout over the original molecule, Mtx, we have illustrated the strength of the non-covalent interaction between Mtx and DHFR in our existing system. Mtx has affinity for DHFR in the picomolar range with an off-rate of $10^{-4}s^{-1}$ [19]. Measurement of the K_d of both Mtx-Dex and A-TMP Dex with the L28C variant of eDHFR will allow for better mechanistic understanding of their differences.

5.2c Future directions

Although we have compared the transcriptional readout of the covalent Y3H system with the original Mtx-Dex system, experiments to assess the ability of the new system to detect interactions of increasingly weaker affinity would provide a more accurate evaluation of the system in a target identification scenario. A series of AD-target protein chimeras with varying affinities could be used to determine the K_d threshold for the new system in terms of transcriptional readout and enrichment ability [20].

In summary, through the use of affinity chromatography studies we have discovered new targets for a novel anti-cancer agent, RSL3, that are involved in antioxidant cellular networks and prostaglandin synthesis with a common connection in lipid metabolism. Follow-up studies have revealed that targeting of GPx4 by RSL3 could be responsible for the generation of ROS that we observe in RSL3-induced death, particularly lipid peroxides. Further, direct interaction of RSL3 with PTGES2 could result in reduction in PGE_2 signaling, a known pathway that is activated in cancer. These findings

suggest that these networks could be therapeutically targeted for the selective treatment of oncogenic *RAS.*

5.3 References

1. Yang, W.S. and B.R. Stockwell, *Synthetic Lethal Screening Identifies Compounds Activating Iron-Dependent, Nonapoptotic Cell Death in Oncogenic-RAS-Harboring Cancer Cells.* Chem Biol, 2008. **15**(3): p. 234-45.
2. Huang, P., D. Trachootham, and J. Alexandre, *Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach?* Nature Reviews Drug Discovery, 2009. **8**(7): p. 579-591.
3. Huang, P., D. Trachootham, Y. Zhou, H. Zhang, Y. Demizu, Z. Chen, H. Pelicano, P.J. Chiao, G. Achanta, R.B. Arlinghaus, and J.S. Liu, *Selective killing of oncogenically transformed cells through a ROS-mediated mechanism by beta-phenylethyl isothiocyanate.* Cancer Cell, 2006. **10**(3): p. 241-252.
4. Raj, L., T. Ide, A.U. Gurkar, M. Foley, M. Schenone, X. Li, N.J. Tolliday, T.R. Golub, S.A. Carr, A.F. Shamji, A.M. Stern, A. Mandinova, S.L. Schreiber, and S.W. Lee, *Selective killing of cancer cells by a small molecule targeting the stress response to ROS.* Nature, 2011. **475**(7355): p. 231-4.
5. Hopkins, A.L., *Network pharmacology: the next paradigm in drug discovery.* Nat Chem Biol, 2008. **4**(11): p. 682-90.
6. Csermely, P., V. Agoston, and S. Pongor, *The efficiency of multi-target drugs: the network approach might help drug design.* Trends Pharmacol Sci, 2005. **26**(4): p. 178-82.
7. Ehsanian, R., C. Van Waes, and S.M. Feller, *Beyond DNA binding - a review of the potential mechanisms mediating quinacrine's therapeutic activities in parasitic infections, inflammation, and cancers.* Cell Commun Signal, 2011. **9**: p. 13.
8. Seiler, A., M. Schneider, H. Forster, S. Roth, E.K. Wirth, C. Culmsee, N. Plesnila, E. Kremmer, O. Radmark, W. Wurst, G.W. Bornkamm, U. Schweizer, and M. Conrad, *Glutathione peroxidase 4 senses and translates oxidative stress into 12/15-lipoxygenase dependent- and AIF-mediated cell death.* Cell Metab, 2008. **8**(3): p. 237-48.
9. Ferlini, C. and G. Scambia, *Assay for apoptosis using the mitochondrial probes, Rhodamine123 and 10-N-nonyl acridine orange.* Nature Protocols, 2007. **2**(12): p. 3111-3114.
10. Araki, Y., S. Okamura, S.P. Hussain, M. Nagashima, P. He, M. Shiseki, K. Miura, and C.C. Harris, *Regulation of cyclooxygenase-2 expression by the Wnt and ras pathways.* Cancer Res, 2003. **63**(3): p. 728-34.
11. Yamada, T. and F. Takusagawa, *PGH2 degradation pathway catalyzed by GSH-heme complex bound microsomal prostaglandin E2 synthase type 2: the first example of a dual-function enzyme.* Biochemistry, 2007. **46**(28): p. 8414-24.
12. Scheerer, P., A. Borchert, N. Krauss, H. Wessner, C. Gerth, W. Hohne, and H. Kuhn, *Structural basis for catalytic activity and enzyme polymerization of phospholipid hydroperoxide glutathione peroxidase-4 (GPx4).* Biochemistry, 2007. **46**(31): p. 9041-9.
13. Yamada, T., J. Komoto, K. Watanabe, Y. Ohmiya, and F. Takusagawa, *Crystal structure and possible catalytic mechanism of microsomal prostaglandin E synthase type 2 (mPGES-2).* J Mol Biol, 2005. **348**(5): p. 1163-76.
14. Savaskan, N.E., C. Ufer, H. Kuhn, and A. Borchert, *Molecular biology of glutathione peroxidase 4: from genomic structure to developmental expression and neural function.* Biol Chem, 2007. **388**(10): p. 1007-17.

15. Costa, A. and S. Onesti, *The MCM complex: (just) a replicative helicase?* Biochem Soc Trans, 2008. **36**(Pt 1): p. 136-40.
16. Becker, F., K. Murthi, C. Smith, J. Come, N. Costa-Roldan, C. Kaufmann, U. Hanke, C. Degenhart, S. Baumann, W. Wallner, A. Huber, S. Dedier, S. Dill, D. Kinsman, M. Hediger, N. Bockovich, S. Meier-Ewert, A.F. Kluge, and N. Kley, *A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors.* Chem Biol, 2004. **11**(2): p. 211-23.
17. Chidley, C., H. Haruki, M.G. Pedersen, E. Muller, and K. Johnsson, *A yeast-based screen reveals that sulfasalazine inhibits tetrahydrobiopterin biosynthesis.* Nat Chem Biol, 2011. **7**(6): p. 375-83.
18. Gallagher, S.S., J.E. Sable, M.P. Sheetz, and V.W. Cornish, *An in vivo covalent TMP-tag based on proximity-induced reactivity.* ACS Chem Biol, 2009. **4**(7): p. 547-56.
19. Appleman, J.R., E.E. Howell, J. Kraut, M. Kuhl, and R.L. Blakley, *Role of aspartate 27 in the binding of methotrexate to dihydrofolate reductase from Escherichia coli.* J Biol Chem, 1988. **263**(19): p. 9187-98.
20. de Felipe, K.S., B.T. Carter, E.A. Althoff, and V.W. Cornish, *Correlation between ligand-receptor affinity and the transcription readout in a yeast three-hybrid system.* Biochemistry, 2004. **43**(32): p. 10353-63.

Appendix A: Further mechanistic studies of RSL3

A range of mechanistic and profiling studies were performed to gain insight into RSL3's mechanism of action. These studies probed many aspects of the cell, including protein, mRNA and metabolite levels, and cellular morphology.

A.1 RSL3 induces morphological disruption of the nuclei and mitochondria

In order to observe morphological changes which might give us insight into important cellular locales or events involved in RSL3-induced death, samples were prepared for analysis by transmission electron microscopy (TEM). Cells treated for different time periods were analyzed to monitor corresponding cellular alterations occurring until death took place. We observed as early as two hours after RSL3 treatment distinct populations of cells either with disrupted nuclei, internally disrupted mitochondria or normal organelles (Figure A.1). Those cells with disrupted organelles however, didn't necessarily possess both disrupted nuclei and mitochondria. This pattern was still observed at four and six hours. After eight hours, extensive cell death had occurred and any live cells remaining were shrunken. One thing of note was that although mitochondria largely retained their tubular morphology, there were sections within the mitochondria where prominent gaps were observed between areas of defined cristae. These studies provide clues that potential key events in RSL3's mechanism of action are occurring in both the mitochondria and nucleus. Furthermore, none of the classic morphological features of established cell death types such as apoptosis (chromatin condensation), necrosis (cytoplasmic swelling, disruption of plasma membrane, swollen mitochondria) or autophagy (double-membrane enclosed compartments) were observed which corroborates with previous studies indicating RSL3 functions through a novel form of cell death. TEM sample preparation was performed by the author, TEM sections were cut by Kristy R. Brown in the Department of Pathology and Cell Biology and examined and analyzed by Dr. Kathryn Lemberg.

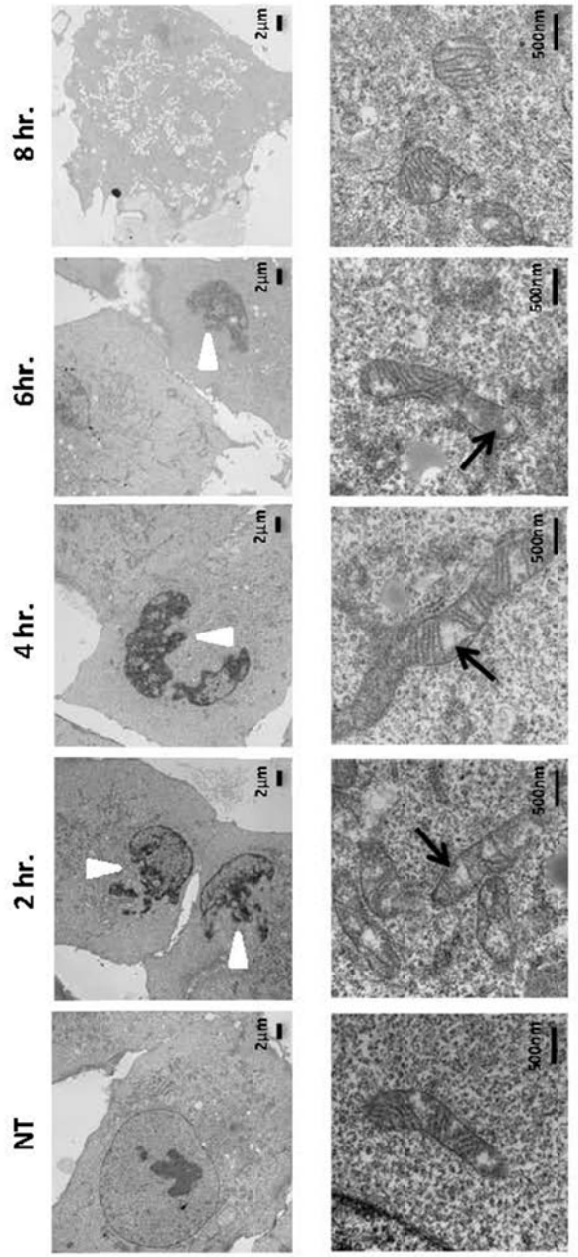


Figure A.1: RSL3 induces morphological disruption of the nuclei and mitochondria

BJeLR cells were treated with RSL3 for 2, 4, 6 and 8 hour time points at 0.5ug/mL or vehicle (DMSO) equivalent for 8 hours. Compared with vehicle-treated cells (NT), RSL3-treated cells showed populations with distinctly disrupted nuclei (white triangles) and mitochondria which appeared to possess disrupted cristae (black arrows) while overall retaining tubular morphology.

A.2 Difference gel electrophoresis analysis

To examine changes in global protein expression in response to RSL3 treatment, either through resulting protein up or down regulation or degradation, we performed two-dimensional difference gel electrophoresis (DIGE) in collaboration with the Comparative Proteomics Center in the Biology department. Two sets of DMSO- and RSL3-treated samples were prepared, labeled with Cy3 and Cy5 dyes respectively and separated using 2D PAGE. In this analysis, spots marking proteins which are present in greater or lesser quantities in the sample appear red or green, while proteins whose level is unchanged have an intermediate color of yellow, indicating equal amounts of red and green protein from either sample (Figure A.2).

This study did not reveal changes in any particular protein level over three-fold. Separate data showing that cyclohexamide, a protein synthesis inhibitor, could not rescue cells from RSL3 treatment supports the idea that protein expression is not required for cell death [1]. Of thirty spots selected for excision from the gel and subsequent identification by MS, six spots corresponded to bovine protein contamination and seven spots produced weak mass spectra such that they were unidentifiable. Of the remaining spots, stathmin 1 and vimentin were the most upregulated with an average 2.3 and 2.5 fold increase in RSL3 treated samples (Table 4.1). Glutathione S-transferase omega 1 (GSTO1) was observed to be down-regulated 1.6 fold. Both the changes in stathmin and GSTO1 were verified by western blot, but because of the small effects, they have not been pursued further. DIGE sample preparation and follow-up studies were performed by the author. DIGE analysis was performed by Dr. Lewis Brown, Director of the Comparative Proteomics Center.

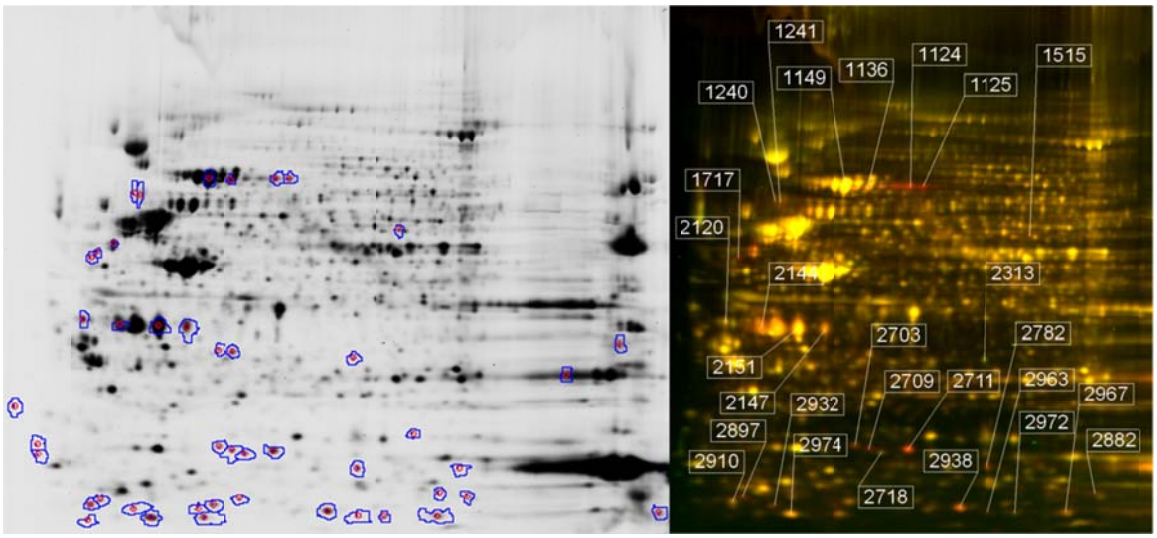


Figure A.2: Example of a 2D difference gel electrophoresis experiment
Cells were treated with 1ug/mL RSL3 or equivalent vehicle control, labeled with Cy3 and Cy5 dyes respectively and separated using 2D PAGE. Spots (numbered) marking proteins which are present in greater or lesser quantities in the sample appear red or green, while proteins whose level is unchanged have an intermediate color of yellow, indicating equal amounts of red and green protein from either sample.

Table A.1: Proteins observed with highest fold change upon RSL3 treatment in two DIGE replicates

Protein name	Fold-change Replicate 1	Fold-change Replicate 2	UniProtKB ID
heat shock 70kDa protein 1B	1.5	1.5	HSP71_HUMAN
heat shock 70kDa protein 8 isoform 1	1.1	1.1	HSP7C_HUMAN
alpha isoform of regulatory subunit A, protein phosphatase 2	1.8	2.2	2AAA_HUMAN
alpha isoform of regulatory subunit A, protein phosphatase 2	1.3	1.4	2AAA_HUMAN
minor mix: pyruvate kinase, muscle isoform M2 and heterogeneous nuclear ribonucleoprotein H1(5031753)	1.5	1.5	KPYM_HUMAN
proliferating cell nuclear antigen	1.0	1.0	PCNA_HUMAN
vimentin	1.8	2.6	VIME_HUMAN
vimentin	2.9	2.7	VIME_HUMAN
glutathione-S-transferase omega 1	-1.6	-1.6	GSTO1_HUMAN
stathmin 1	2.4	1.5	STMN1_HUMAN
stathmin 1	3.1	1.4	STMN1_HUMAN
stathmin 1	3.9	1.5	STMN1_HUMAN

Proteins shown more than once were detected in different spots.

A.3 RSL3-induced death is caspase and Bax/Bak-independent

Modulatory profiling is a system developed in the Stockwell laboratory to characterize lethal molecules based on their profile of inhibition or sensitization with a panel of chemical modulators of known cell death processes as described in Chapter 1.3b [2]. Compounds acting through similar lethal mechanisms have profiles which cluster together and thus novel small molecules can either be classified in a group of compounds with known mechanisms of action or identified as possessing a unique cell death mechanism.

RSL3 and another *RAS*-selective lethal compounds erastin were also included in the profiling system and were found to cluster separately from characterized lethal compounds such as topoisomerase inhibitors, microtubule destabilizers and alkylating agents. These findings were consistent with previous data indicating the novelty of RSL3's mechanism. Compounds in this unique cluster which included RSL3 were found to induce a caspase-independent form of cell death as determined by a caspase activity assay using a fluorogenic reporter (Figure A.3). These data provide further confirmation that RSL3 is working through a non-apoptotic mechanism. To further exclude cell death types that are caspase-independent yet dependent on Bax and Bak-mediated mitochondrial outer membrane permeabilization, RSL3 was tested in $Bax^{-/-}Bak^{-/-}$ double knockout mouse embryonic fibroblasts (MEFs). Rather than suppressing death, cells were instead slightly sensitized to RSL3. These experiments were performed by Dr. Adam Wolpaw and Kenichi Shimada.

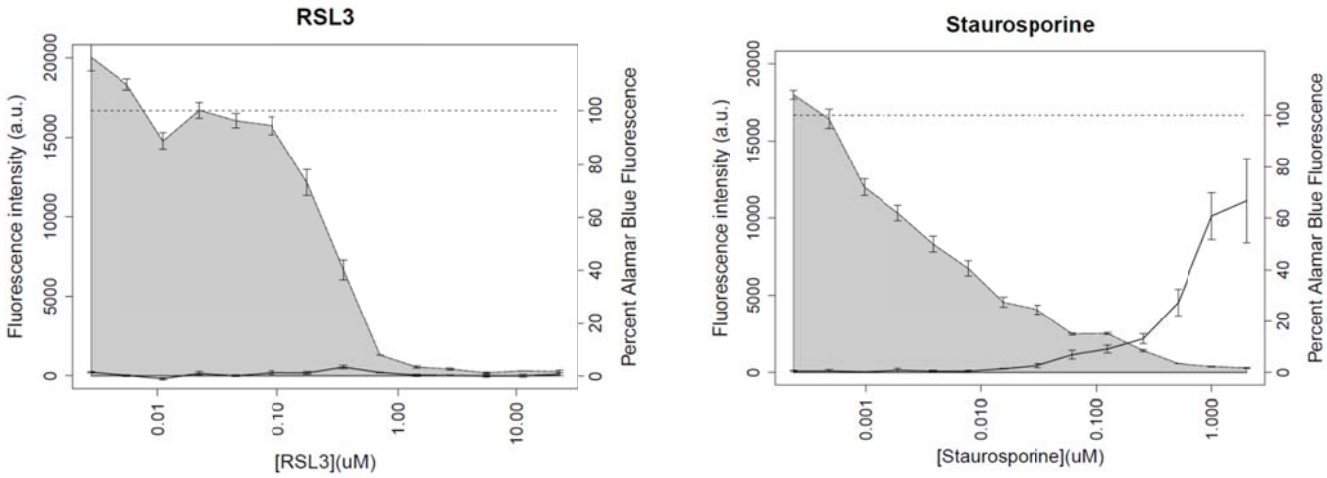


Figure A.3: Caspase activation is not observed in RSL3-induced death

Fluorescence intensity of a fluorogenic caspase substrate indicating caspase activation is plotted on the left *y* axis with percent alamar blue fluorescence indicating cell viability (shaded grey) on the right axis. Staurosporine, a known inducer of apoptosis and caspase activation, is used as a positive control and caspase activation is observed at lethal concentrations. No caspase activation is observed, however at lethal concentrations of RSL3.

A.4 RSL3 reduces tumor size in a mouse xenograft model from BJeLR cells

Activity in mouse models is a standard and important assessment of a small molecule's ability to perform *in vivo*. Small molecules identified in chemical genetic screens aimed at finding new therapeutics in disease models must be proven to act in an organismal environment to verify that their mechanism is not an artifact of cell culture. Further, if the molecule is to serve as a drug lead, a number of pharmacokinetic parameters need to be met to warrant further drug discovery investment.

To evaluate RSL3's *in vivo* efficacy we tested RSL3's effect on mouse xenografts formed from BJeLR cells. RSL3 was able to decrease the size of tumor formation in this model in both a prevention and therapeutic study (Figure A.4). In the prevention study athymic nude mice were injected with BJeLR cells subcutaneously and then with 100mg/kg (1S,3R) RSL3 to the same site one day later. In the therapeutic study mice were allowed to form tumors for 1 week and then injected with 100mg/kg (1S,3R) RSL3 intratumorally twice a week for two weeks. Tumor size was measured four weeks after the cell injection. Reduction in tumor volume was greater in the prevention study than the therapeutic study, but decrease in tumor size was observed in both. These results are promising considering that the RSL3 scaffold is almost the same as the original hit identified in our screen and contains a reactive chloromethylketone functionality. Further optimization of the structure through medicinal chemistry to improve solubility and potency as well as attenuate the electophilicity of the molecule could produce improved results *in vivo*. Mouse studies were performed by Dr. Wan Seok Yang.

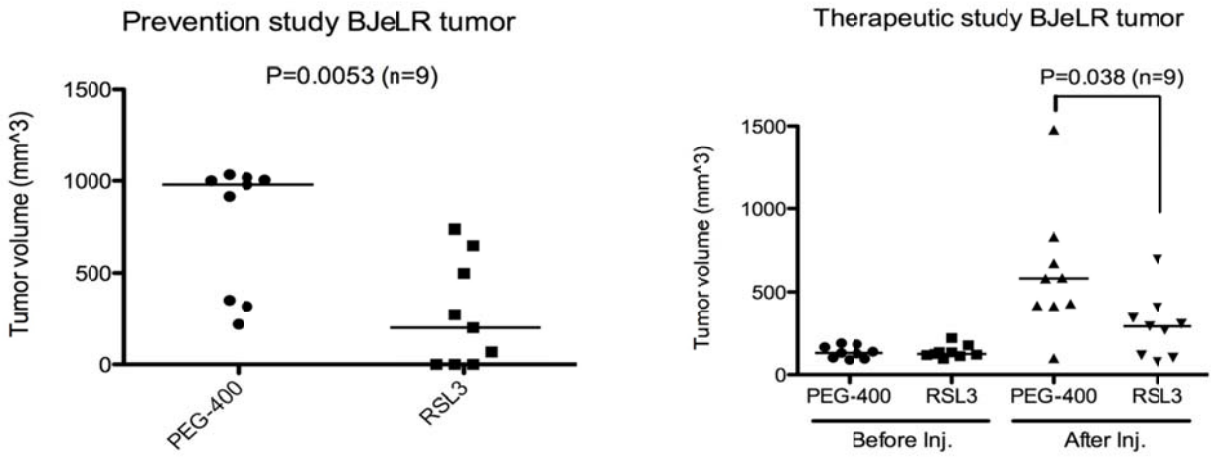


Figure A.4: RSL3 reduced tumor size in a mouse xenograft model

RSL3 was shown to decrease the size of tumor formation in a mouse xenograft model of BJeLR cells in both a prevention study (left) and a therapeutic study (right). PEG-400 was used as a vehicle and small molecule was administered intratumorally.

A.5 HIP/HOP profiling

RSL3 was submitted for profiling in a platform of genetic modulator screens in yeast composed of a whole-library genome of heterozygous deletion mutants and a library of homozygous deletion mutants for non-essential genes (described in Chapter 1.3a). Yeast from either library were treated with RSL3 in a pooled format and mutants with increased sensitivity to RSL3 compared with vehicle control were identified. A greater level of sensitization was observed in the homozygous deletion library whose genes indicate mutants that modulate RSL3-induced death but are not direct targets. One gene from this library, inositol heptakisphosphate kinase (KCS1), which produces high energy inositol pyrophosphates, was sensitized ca. eleven-fold. Among genes sensitized in the heterolozygous deletion library, which indicates potential direct targets, the yeast frataxin homolog showed a two-fold change in sensitivity. This was noticeable as one of the few genes to show any increase in sensitization and due to its involvement in the assembly of iron-sulfur clusters as iron is already known to play a role in RSL3's mechanism. Silencing of the human gene in our laboratory caused a degree of cell sensitivity to erastin which also acts in an iron-dependent manner (Dixon, Stockwell; unpublished results). No further validation of these results has been performed. HIP/HOP profiling was performed by the Nislow laboratory at the University of Toronto.

A.6 Pyrimidine metabolites levels change upon RSL3 treatment

Changes in metabolite levels are an additional source of clues about a molecule's effects on cellular metabolic networks. In our laboratory, metabolite profiling has revealed that glutathione depletion plays a major role in erastin-induced cell death (Yang, Stockwell; unpublished results). RSL3 treatment was shown to affect levels of pyrimidine metabolites, particularly the RNA nucleoside uridine was upregulated five-fold in comparison with vehicle-treated controls and its corresponding nucleobase, uracil, was downregulated similarly (Figure A.5). Another RNA nucleoside, cytidine, and a DNA nucleoside, thymidine, were also upregulated. Orotate, also down-regulated, is a pyrimidine carboxylic

acid. We speculate that some process which increases or requires the synthesis of pyrimidine nucleosides may be involved in RSL3-induced death. Samples for metabolite analysis were prepared by Dr. Wan Seok Yang and analyzed by Dr. Clary Clish, Director of the Metabolite Profiling Platform at the Broad.

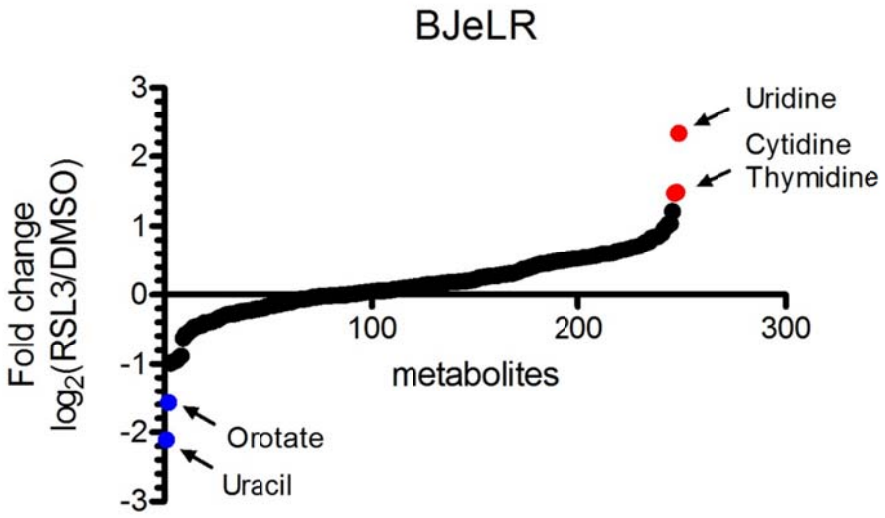


Figure A.5: Pyrimidine metabolite levels change upon RSL3 treatment

Among metabolite levels increased upon RSL3 treatment, RNA nucleosides uridine and cytidine were increased most significantly with uracil showing a corresponding decrease. Levels of thymidine, a DNA nucleoside were also increased.

A.7 Gene expression profiling

Gene expression profiling is a useful hypothesis generator when analyzing a small molecule's mechanism of action. RSL3 was submitted for gene expression analysis to the Connectivity Map (CMAP) profiling database previously described (Chapter 1.3b). While looking at changes in the expression levels of genes upon RSL3 treatment does not provide any direct evidence about targets (unless they are strongly related to gene expression), comparing RSL3 gene expression data against a database to find other small molecules with similar expression profiles could provide information about overlapping targets if those molecules are well characterized. Additionally, looking at enrichment of gene categories in up and down regulated genes can give indications of downstream effector pathways involved in the death.

RSL3 did not show high connectivity with other molecules in the CMAP database besides known *RAS*-selective lethal compounds such as erastin and RSL5. Compounds that did cluster with RSL3 caused oxidative stress and protein misfolding. Of the top one hundred up and down regulated genes, gene categories involving prostaglandin metabolic processes, glucuronosyl transferases, tyrosine/serine/threonine MAP kinases and regulation of cell redox homeostasis were enriched upon RSL3 treatment. RSL3 is known to induce ROS, so enrichment of genes regulating redox homeostasis and clustering with other compounds causing oxidative stress is expected. RSL3 is also known to be dependent on RAS-RAF-MEK signaling which perhaps ties in with upregulation of MAP kinases. Glucuronosyl transferases carry out glucuronidation reactions which are generally used to help export xenobiotics from the cells. These enzymes transfer glucoronic acid in the form of UDP-glucoronic acid to small hydrophobic molecules which increase their solubility and aids secretion from the cell. Up-regulation of these genes may signify a cellular response to remove RSL3 from the cell. A connection could perhaps exist between increased levels of these genes and an increase in uridine metabolites.

Microarray data was generated and provided by Dr.Todd Golub and colleagues at the Broad institute and was analyzed by Dr. Adam Wolpaw [2].

A.8 Preliminary investigation of other candidate targets of RSL3 from proteomic data

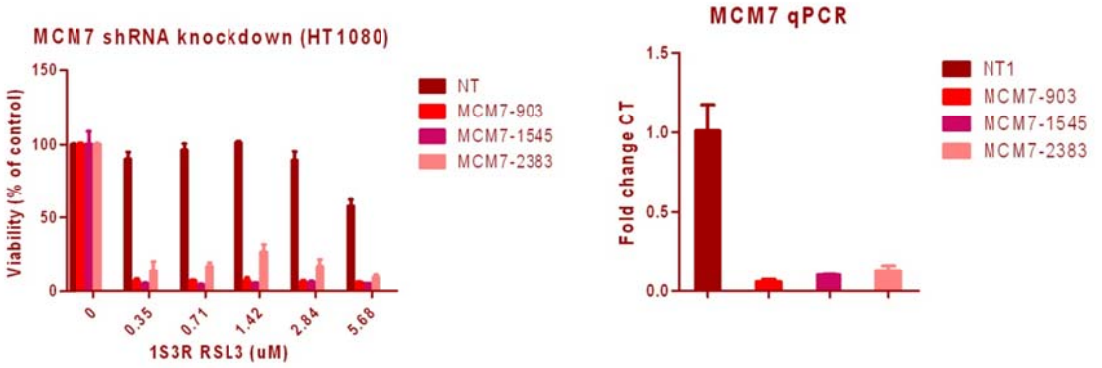
The other top candidates that were shown to interact with RSL3 in the proteomic results both have fundamental roles within the cell and these were also preliminarily investigated using RNAi. SMG9 is a protein that is part of a surveillance complex in translation which senses mRNAs containing premature translation termination codons and marks them for degradation in a process called nonsense-mediated mRNA decay. Other members of the complex include SMG-1, a kinase which phosphorylates another member of complex, Upf, indicating that a premature termination codon has been recognized. Intriguingly, SMG-8, which is involved in the recruitment of SMG-1 to the complex, was also among the top ten candidates in the proteomic results. Comparison of the two sequences shows only modest homology, thus it is unlikely that this is due to an identification error in the MS analysis [3, 4].

Minichromosome maintenance complex component 7 (MCM7) is a part of a complex of proteins (MCM2-7) which are involved in the initiation of DNA replication as a component of the pre-replication complex. Certain complex members possess helicase activity and thus it is believed to be the replicative helicase. These complexes form during G1 phase and only dissociate upon phosphorylation by cyclin-dependent kinases acting to ensure that each origin of replication is used only once per cell cycle. MCM7 and other MCM proteins are present in excess of the required amount needed for their role in DNA replication, thus they are suspected to possess other functions in the cell [5, 6].

To investigate the role of these proteins we attempted to deplete cellular levels of MCM7 and SMG9 by performing viral shRNA knockdowns of these genes and observing the effects on cell sensitivity to RSL3. Quantitative PCR was performed to check the efficacy of the knockdowns (Figure A.6A). RNAi silencing of *MCM7* showed sensitization to RSL3 treatment similar to that observed with *GPx4* and

PTGES2. Viral knockdown of the fourth hit, *SMG9*, did not show any effects on RSL3 potency (Figure A.6B). Quantitative PCR however, showed that the knockdown of the shRNA clones used was poor. RNAi validation of this hit will need to be performed with siRNA or another method. Sensitization by knockdown of *MCM7* suggests that RSL3 may be targeting this protein as well through a loss-of-function mechanism.

A



B

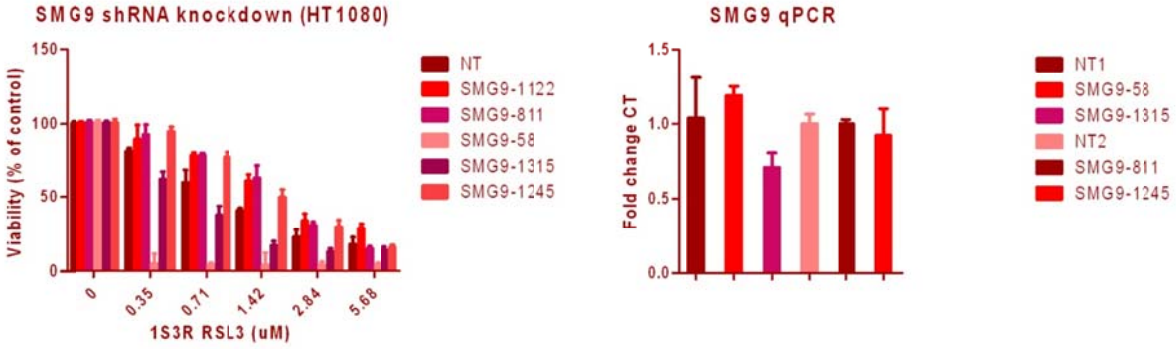


Figure A.6: RNAi knockdown of *SMG9* and *MCM7* sensitizes cells to RSL3-induced death
Viral shRNA knockdowns of (A) *MCM7* and (B) *SMG9* alongside non-targeting controls (NT) were performed in HT1080 cells and their effects on RSL3 sensitivity were assessed. Knockdown of *MCM7* resulted in sensitization to RSL3 treatment while *SMG9* knockdown showed no effect, except for one clone. Verification by quantitative PCR (**qPCR**) showed poor knockdown by *SMG9* shRNAs. Data represent the mean of three replicates ±SD.

RNAi knockdown of MCM7 does not phenocopy RSL3 treatment

The effect of viral shRNA knockdown of *MCM7* on cell viability was examined to see whether protein depletion mimicked the possible loss-of-function of this protein by RSL3. We observed no effect of knockdown of *MCM7* on cell viability by viral shRNA infection of HT1080 cells (Figure A.7). In later experiments we did note a loss in viability with infection of sh*MCM7* clones over a longer time period of infection presumable due to interference with replication and the cell cycle. However, the knockdown of *MCM7* did not cause an effect analogous to the rapid death induced by RSL3.

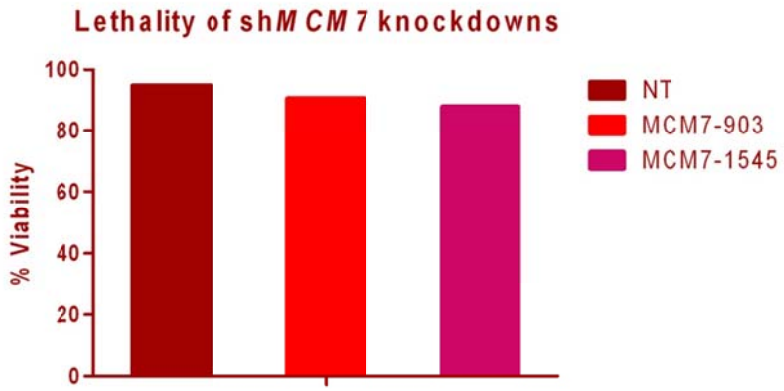


Figure A.7: RNAi knockdown of *MCM7* does not cause loss of viability

HT1080 cells were infected with viral shRNA clones targeting *MCM7* and non-targeting controls (NT) and viability was measured six days later using trypan blue exclusion.

A.9 Discussion

The mechanism of RSL3 was investigated using a number of techniques to look at changes in morphology, protein levels, mRNA levels and metabolite levels upon RSL3 treatment. Additionally RSL3 was profiled in a genetic modulator platform (HIP/HOP), a gene expression profiling data base (CMAP) and in a modulatory profiling system to characterize its form of cell death. RSL3 was also shown to be active *in vivo* in a mouse xenograft model study. Through these studies we gained a few clues to aid us in our target validation studies.

The roles of MCM7 and SMG9 in RSL3-induced death were investigated by RNAi and knockdown of *MCM7* also resulted in sensitization to RSL3 treatment. The connection of MCM7 with GPx4 and PTGES2 or its role in RSL3's mechanism is more obscure. Its primary role in DNA replication initiation is an important one, however prevention of DNA synthesis appears unlikely to explain the rapid induction of death induced by RSL3. There is a possibility that halting cell cycle progress sensitizes the cells to the lethal mechanisms of other targets.

Evidence from TEM studies shows disruption of the nuclei and it will be interesting to know if this is due to involvement of MCM7, or from damage due to lipid peroxides or some other mechanism. The increase in RNA nucleosides observed in metabolic profiling could also be due to the involvement of SMG9 and nonsense-mediated mRNA decay although their role in RSL3'S mechanism remains to be investigated. Up-regulation of prostaglandin metabolic processes observed in gene expression profiling may be a result of modulation by RSL3 of PTGES2 and COX through GPx4.

MCM7 has been reported to play a role in cancer. The genes oncogenic properties were first recognized when it was noticed to be overexpressed or amplified in a number of malignant cancers and was thus used as a biomarker [7-9]. Further studies in prostate cancer cells and mouse models showed

that MCM7 overexpression actively promotes cell proliferation and invasiveness, although it is not believed to act as an initiator in tumorigenesis [8, 10, 11].

A number of proteins have been found to interact with MCM7, and depending on their interaction, might either promote replication (and thus proliferation) or induce growth arrest. The retinoblastoma protein (Rb) was first shown to interact with MCM7 in Y2H screening [12]. A recent study has shown that the ability of TGF-β1 to block entry into the S-phase of the cell cycle in late G1 phase is due to blocking the dissociation of Rb and MCM7, thus preventing activation of the pre-replication complexes. The suggestion is that this interaction could be how Rb controls the S phase check point [13]. The androgen receptor and integrin-linked kinase 1 have also been shown to interact with MCM7 and either positively or negatively affect MCM replication activity [10]. Additionally, MCM7 possesses an intronic miRNA cluster which has been shown to target multiple tumor suppressor genes such as pTEN, p21, and BIM [14, 15]. Whether the effects of RSL3 could abrogate the ability of this cluster appears unlikely. It is also unclear whether MCM7's transforming properties are primarily due to expression of its oncogenic miRNA cluster or to the function of the protein.

Further evidence of the role of MCM7 in RSL3-induced cell death must be gained in order to understand its potential involvement.

A.10 Methods

Cell culture

HT-1080 and 293-T cells were obtained from American Type Culture Collection. HT-1080 cells were grown in DMEM High-Glucose media (Gibco) supplemented with 10% fetal bovine serum (FBS) and 1% non-essential amino acids (Gibco). 293-T cells were grown in DMEM High-Glucose supplemented with 10% FBS. When used for transfections to generate virus, 293-T cells were seeded in the above media lacking antibiotics. 293-T viral collection media contained 30% HyClone FBS. BJ/TERT/LT/ST/RASV12 (BJeLR) and related isogenic cell lines (eH, eHLT, DRD) were obtained from William Hahn (Harvard University). BJ series cells were grown in DMEM plus 20% M199 and 15% heat-inactivated FBS. All media were supplemented with penicillin and streptomycin (Gibco). All cell lines were grown in humidified tissue culture incubators (Thermo Scientific) at 37°C with 5% CO_2.

Viability measurements

Viability was assessed using one of two viability dyes: Alamar blue fluorescence or Trypan Blue dye exclusion. Alamar blue (Invitrogen) is a cell-permeable resazurin-based dye that is reduced by metabolically active cells to rezorufin, a dye with red fluorescence. Fluorescence was measured 6-16 hours after alamar blue addition to cells using a Victor3 platereader with excitation filter at 530 nm and emission filter centered on 590 nm. Trypan Blue is a viability dye that is excluded from live cells but permeates dead cells. Trypan Blue negative and positive cells were counted using an automated cell counter (ViCell, Beckman-Coulter).

A.3c Transmission Electron Microscopy

BJeLR cells were seeded (400,000 cells/dish) in 35mm tissue culture dishes and grown overnight. The next day cells were treated with DMSO for 8 h and RSL3 (diastereomeric version) at 0.5ug/mL for 2, 4, 6 and 8 h. Cells were fixed with 2.5% glutaraldehyde in 0.1 M Sorenson's buffer (0.1 M H_2PO_4, 0.1 M HPO_4 (pH 7.2)) for at least 1 h and then stored at 4°C until transportation. The rest of the sample preparation was performed by Kristy R. Brown in the Department of Pathology and Cell Biology. Cells were treated

with 1% OsO4 in 0.1 M Sorenson's buffer for 1 h, followed by Enblock staining used 1% tannic acid. After dehydration through an ethanol series, cells were embedded in Lx-112 (Ladd Research Industries) and Embed-812 (EMS). Thin sections were cut on an MT-7000 ultramicrotome and stained with 1% uranyl acetate and 0.4% lead citrate. Sections were examined and analyzed by Dr. Katie Lemberg using a Jeol JEM-1200 EXII electron microscope. Pictures were taken on an ORCA-HR digital 52 camera (Hamamatsu) at 5,000-50,000-fold magnification, and measurements were made using the AMT Image Capture Engine.

Two-dimensional difference gel electrophoresis.

Sample preparation: 10cm dishes were seeded with 4 million BJeLR cells the day before and treated the following day with either 1ug/mL RSL3 (diastereomeric version) or the equivalent amount of DMSO for 7hr. Treatment conditions were then removed and cells were washed in cold PBS and lysed in 50 mM Tris, pH 7.5, 150 mM NaCl, 0.3% SDS in the presence of Serine/Threonine Phosphatase Inhibitors and Protease Inhibitors. Cells were then treated with 50ul DNAse/RNAse for 30min. Protein was precipitated using chloroform/methanol precipitation and the resulting pellet was incubated in 200uL of reaction buffer overnight (7 M Urea, 2 M Thiourea, 4% CHAPS). The next day the lysate was vortexed and homogenized, 10uL of DTT was added and the mixture was incubated for an hour followed by vortexing and homogenization again. Undissolved material was removed by centrifugation at 16,000g and the final supernatant was removed. Protein concentration was determined using the Bradford Assay and aliquots were snap frozen in liquid nitrogen and stored at -80°C. Covalent Labeling with Cyanine Dyes, 2D Gel Electrophoresis, Confocal Laser Imaging of Gels, Image Analysis, Spot Picking and Enzymatic Digestion and Protein Identification by Peptide Mass Fingerprinting were all performed by Dr. Lewis Brown as described in Cardinale CJ et. al. (ref 175, Nidhi's thesis)

Mouse studies

Athymic nude mice were injected with BJeLR cells subcutaneously. In the prevention study mice were injected with 100mg/kg (1S,3R) RSL3 to the same site one day later. In the therapeutic study mice were allowed to form tumors for 1 week and then injected with 100mg/kg (1S,3R) RSL3 intratumorally twice a week for two weeks. The tumor size was measured using a caliper four weeks after the cell injection. Mouse studies were performed by Dr. Wan Seok Yang.

HIPHOP profiling

'HIPHOP' profiling was performed by the Nislow lab at the University of Toronto. Libraries of yeast were treated with RSL3 at 125ug/mL and 150ug/mL.

Gene expression analysis

Microarray data were downloaded from the Broad's Connectivity Map website (http://www.broadinstitute.org/cmap/). For the gene expression experiments, MCF7 and PC3 cells were treated with compound for 6 hours with 10uM of RSL3 before lysis and mRNA collection. A more detailed description of the experimental protocols is available on the website and in their publication about the project (insert CMAP reference). Analysis of micro array data was performed and described by Dr. Adam Wolpaw [2].

Lentiviral shRNA production

Virus production followed the protocol of The RNAi Consortium. For production of virus, 293-T cells were seeded in 6-well dishes in antibiotic free media (280,000 cells/well). The next day, cells were transfected using FuGENE with shRNA encoding plasmid (450 ng), viral packaging plasmid (p-Delta8.9, 400 ng), and viral envelope plasmid (p-VSV-G, 45 ng). After 24 hours, media was removed and replaced with 1.6 ml Viral Collection Media. A total of 3 collections of viral supernatant in VCM per shRNA were made over 36 hours and pooled. Pooled supernatant was centrifuged (2,250 rpm, 2 min), aliquoted and stored at -80°C.

Lentiviral shRNA Infection

For lentiviral infections in 384-well format, 400 cells were seeded per well the night before the infection. On the day of the infection, media was replaced with 38 ul media + polybrene (8 ug/ml) and 2 ul of the viral supernatant per well. Plates were centrifuged at 2,250 rpm at 37°C for 90 minutes in a tabletop Sorvall centrifuge. Plates were incubated for 48 hours, the media was replaced with media containing 1.5 ug/ml puromycin and incubated for 24 hours. Media was subsequently changed again for 40 ul of RSL3 or compound-containing media at the appropriate concentration. Plates were incubated for 24 hours and viability was assessed using alamar blue. For lentiviral infections in 6-well format, 35,000 cells were seeded per well in 2mL of media the night before the infection. Cells were infected in the presence of 8 ug/ml polybrene in 1.85mL with 150 ul viral supernatant. Puromycin selection, compound treatment were the same as above, but scaled for 6-well format. Viability was assessed by trypan-blue exclusion. Where possible, viability was also assessed by light microscopy.

Gene expression by RT-q-PCR

Cells from 6-well format infection were harvested following puromycin selection and RNA was purified using the QIAshredder and QIAgen RNAeasy extraction kits according to the manufacturer's instructions. 2 ug total RNA per sample were subsequently used in a reverse transcription reaction using the TaqMan RT Kit priming with Random Hexamers (Applied Biosystems). Primers for q-PCR were designed with either Primer Express. Quantitative PCR was performed on triplicate samples in 96-well format using Power SYBR Green Master Mix (Applied Biosystems) on an Applied Biosystems 7300 Cycler set to absolute quantification. The change in expression of a gene between experimental and control conditions was computed using the DeltaDelta Ct method with *ACTIN* as an internal reference gene.

A.11 References

1. Yang, W.S. and B.R. Stockwell, *Synthetic Lethal Screening Identifies Compounds Activating Iron-Dependent, Nonapoptotic Cell Death in Oncogenic-RAS-Harboring Cancer Cells.* Chem Biol, 2008. **15**(3): p. 234-45.
2. Wolpaw, A.J., Shimada K, Skouta R, Welsch M.e, Akavia U D, Pe'er D, Shaik F, Bulinski J.C, Stockwell, B.R, *Modulatory profiling identifies mechanisms of small molecule-induced death.* PNAS, 2011.
3. Yamashita, A., N. Izumi, I. Kashima, T. Ohnishi, B. Saari, Y. Katsuhata, R. Muramatsu, T. Morita, A. Iwamatsu, T. Hachiya, R. Kurata, H. Hirano, P. Anderson, and S. Ohno, *SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay.* Genes Dev, 2009. **23**(9): p. 1091-105.
4. Fernandez, I.S., A. Yamashita, E. Arias-Palomo, Y. Bamba, R.A. Bartolome, M.A. Canales, J. Teixido, S. Ohno, and O. Llorca, *Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex.* Nucleic Acids Res, 2011. **39**(1): p. 347-58.
5. Costa, A. and S. Onesti, *The MCM complex: (just) a replicative helicase?* Biochem Soc Trans, 2008. **36**(Pt 1): p. 136-40.
6. Hubbi, M.E., W. Luo, J.H. Baek, and G.L. Semenza, *MCM proteins are negative regulators of hypoxia-inducible factor 1.* Mol Cell, 2011. **42**(5): p. 700-12.
7. Brake, T., J.P. Connor, D.G. Petereit, and P.F. Lambert, *Comparative analysis of cervical cancer in women and in a human papillomavirus-transgenic mouse model: identification of minichromosome maintenance protein 7 as an informative biomarker for human cervical cancer.* Cancer Res, 2003. **63**(23): p. 8173-80.
8. Ren, B., G. Yu, G.C. Tseng, K. Cieply, T. Gavel, J. Nelson, G. Michalopoulos, Y.P. Yu, and J.H. Luo, *MCM7 amplification and overexpression are associated with prostate cancer progression.* Oncogene, 2006. **25**(7): p. 1090-8.
9. Li, S.S., W.C. Xue, U.S. Khoo, H.Y. Ngan, K.Y. Chan, I.Y. Tam, P.M. Chiu, P.P. Ip, K.F. Tam, and A.N. Cheung, *Replicative MCM7 protein as a proliferation marker in endometrial carcinoma: a tissue microarray and clinicopathological analysis.* Histopathology, 2005. **46**(3): p. 307-13.
10. Luo, J.H., *Oncogenic activity of MCM7 transforming cluster.* World J Clin Oncol, 2011. **2**(2): p. 120-4.
11. Honeycutt, K.A., Z. Chen, M.I. Koster, M. Miers, J. Nuchtern, J. Hicks, D.R. Roop, and J.M. Shohet, *Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis.* Oncogene, 2006. **25**(29): p. 4027-32.
12. Sterner, J.M., S. Dew-Knight, C. Musahl, S. Kornbluth, and J.M. Horowitz, *Negative regulation of DNA replication by the retinoblastoma protein is mediated by its association with MCM7.* Mol Cell Biol, 1998. **18**(5): p. 2748-57.
13. Mukherjee, P., S.L. Winter, and M.G. Alexandrow, *Cell cycle arrest by transforming growth factor beta1 near G1/S is mediated by acute abrogation of prereplication complex activation involving an Rb-MCM interaction.* Mol Cell Biol, 2010. **30**(3): p. 845-56.
14. Poliseno, L., L. Salmena, L. Riccardi, A. Fornari, M.S. Song, R.M. Hobbs, P. Sportoletti, S. Varmeh, A. Egia, G. Fedele, L. Rameh, M. Loda, and P.P. Pandolfi, *Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation.* Sci Signal, 2010. **3**(117): p. ra29.
15. Kan, T., F. Sato, T. Ito, N. Matsumura, S. David, Y. Cheng, R. Agarwal, B.C. Paun, Z. Jin, A.V. Olaru, F.M. Selaru, J.P. Hamilton, J. Yang, J.M. Abraham, Y. Mori, and S.J. Meltzer, *The miR-106b-25 polycistron, activated by genomic amplification, functions as an oncogene by suppressing p21 and Bim.* Gastroenterology, 2009. **136**(5): p. 1689-700.