2017

Phage display to identify functional resistance mutations to Rigosertib

Nedim Filipovic

Claremont McKenna College

Recommended Citation
http://scholarship.claremont.edu/cmc_theses/1475

This Open Access Senior Thesis is brought to you by Scholarship@Claremont. It has been accepted for inclusion in this collection by an authorized administrator. For more information, please contact scholarship@cuc.claremont.edu.
Phage display to identify functional resistance mutations to Rigosertib

A Thesis Presented

by

Nedim Filipovic

To the Keck Science Department

Of Claremont McKenna, Pitzer, and Scripps Colleges

In partial fulfillment of

The degree of Bachelor of Arts

Senior Thesis in Chemistry

December 5, 2016
Acknowledgments

Thank you to Dr. Aaron Leconte for his guidance, trust, and faith.

Thank you to Dr. Arlie Rinaldi for her time and support during a tumultuous semester.

A special thanks to Devin Cavero, Reid Dickerson, Sidd Mandava, and Sam Peterson for always having faith in me, even when I didn’t have any in myself, and for the laughs, late-night talks, and friendship.

Finally, I’d like to thank my family and friends who listened and supported me throughout the semester.
# Table of Contents

Abstract  iv  
Summary  1  
A. Phage Display  3  
A.1. Phage Biology  4  
A.1.1. Structure  5  
A.1.2. Cell Infection  8  
A.1.3. Phage Replication  9  
A.2. Display Proteins  10  
A.3. Phage Library Construction  11  
A.3.1. Phage Vectors  11  
A.3.1.a. Phage and Phagemid Vectors  13  
A.3.1.b. Vector Display Confirmation  16  
A.3.2. Library Construction  16  
A.3.2.a. Oligonucleotide-directed Mutagenesis  17  
A.3.2.b. *In vitro* DNA Recombination  20  
A.3.2.c. Other Methods  21  
A.3.3. Selection  22  
A.3.3.a. General Approach  22  
A.3.3.b. Methods  22  
A.3.3.b.1. Immobilization  22  
A.3.3.b.2. Solution Binding  24  
A.3.3.b.3. Selection Conditions  25  
B. High-Throughput Sequencing  25  
B.1. Short Read Illumina Sequencing  26  
B.2. Application and Results  28  
C. RAS Oncogene  30  
C.1. RAS Structure/Function  30  
C.2. Oncogenic RAS  31  
C.2.1. Cell Proliferation  33  
C.2.2. Cell Apoptosis  34  
C.2.3. Metabolism  35  
C.2.4. Cell Microenvironment  36  
C.2.5. Evasion of Immune Response  36  
C.2.6. Metastasis  37  
C.3. Cancer Therapy via RAS Inhibition  38  
C.4. RAS and the RAF-MEK-ERK Signaling Pathway  39  
C.4.1. RAF Structure/Function  39  
C.4.2. Overview of Signaling Pathway  41  
C.4.3. RAS Involvement in Pathway  42  
C.5. Past RAS Targeting Drug Therapies  42  
C.6. Rigosertib  43  
D. Proposed Experiments  45  
D.1. Significance  58  
References  59
Abstract

*In vitro* protein selection has had major impacts in the field of protein engineering. Traditional screens assay individual proteins for specific function. Selection, however, analyzes a pool of mutants and yields the best variants. Phage display, a successful selection technique, also provides a reliable link between variant phenotype and genotype. It can also be coupled with high throughput sequencing to map protein mutations; potentially highlighting vital mutations in variants. We propose to apply this technique to cancer therapy. RAF, a serine/threonine kinase, is critical for cell regulation in mammals. RAF can be activated by oncogenic RAS, found in over 30% of cancers, to drive cancer proliferation. Rigosertib, a benzyl styryl sulfone in phase III clinical trials for myelodysplastic syndrome (MDS), is an inhibitor of the RAS binding domain (RBD) in RAF. Phage display can be used to select RAF mutants for RAS binding affinity, in the presence of Rigosertib. High-throughput sequencing of these variants can provide a means of anticipating, and mapping resistance to this anti-cancer drug.
Summary

In vitro protein selection has had major impacts in the field of protein engineering and analysis. While traditional screens assay individual mutants or proteins of interest for specific function, selection analyzes a pool of mutants by subjecting them to a set of conditions that are designed to yield the best variants. Selection, however, is plagued by the difficulty of accurately identifying the genotypes of selected variants. Phage display is a successful selection technique that provides a reliable link between protein and genotype. Ff bacteriophage can display non-native proteins alongside its coat proteins. This is accomplished by cloning a protein of interest into the genome of Ff bacteriophage, which results in the display of that protein on the phage coat. Thus, a single phage links the genotype and phenotype of the protein that its genome includes. Furthermore, once cloned into a phage genome, a protein can be diversified to create a mutant library. This library of phage can be selected for specific function effectively and efficiently. Selected phage can then be characterized via high throughput sequencing to map vital protein mutations; highlighting mutations conserved across multiple variants that may be involved with the function of interest.

The coupling of selection and high throughput sequencing can have major implications for cancer therapy. RAS is a small-GTPase membrane bound protein, primarily involved with signal transmission in cells. Oncogenic mutations in RAS have found in over 30% of cancers. RAS oncogenes are involved in signal transduction pathways that affect cell apoptosis, proliferation, and division, and can initiate or perpetuate human carcinomas. RAF is a serine/threonine kinase directly downstream of RAS. RAF is critical for cell growth and regulation in mammals via its role in the cytoplasmic kinase cascade signaling pathway. Oncogenic RAS has been found to impact tumor growth through the RAF-mediated pathway. Because RAS is involved in a number of cellular pathways, drugs that specifically inhibit binding between RAS and RAF can affect the mutagenic effects of oncogenic RAS without adversely impacting other vital RAS function in the cell.

Rigosertib, a benzyl styryl sulfone in phase III clinical trials for myelodysplastic syndrome (MDS), is a small molecule inhibitor of the RAS binding domain of RAF (RBD). It is possible that RAF may develop a functional resistance mutation that reduces Rigosertib binding affinity, while retaining RAS binding affinity. By combining phage display selection of RAF variants, with high-throughput sequencing we can understand and map RAF resistance to Rigosertib. This information can be used to enhance current and future cancer therapeutics.

A library of RAF mutants will be successfully displayed on Ff bacteriophage. This library will then undergo a selection strategy that will procure variants with high binding affinity to RAS in the presence of Rigosertib. We can infer that these variants have developed resistance to the anti-cancer drug. Following selection, these variants will be characterized by high-throughput sequencing, and functional resistance mutations will be identified. Lastly, selected variants will be tested for functionality in
downstream signaling interactions; the development of resistance is not entirely effective if downstream function is impaired, and signal transduction cannot occur.

**Experimental Aim 1:** Create library of RAF mutants and confirm phage display by selecting for mutant binding affinity to RAS

**Experimental Aim 2:** Use phage display to identify RAF mutants that exhibit resistance to Rigosertib, while simultaneously retaining binding affinity to RAS

**Experimental Aim 3:** Test selected mutants for functionality within the RAF-MEK-ERK pathway
A. Phage Display

*In vitro* protein selection has gained ground in the field of protein engineering and analysis. Selection techniques are a better tool than screens for analyzing and specifying molecular function and efficiency. Screens assay individual mutants or proteins of interest for specific function, resulting in the identification of the best protein within the batch. The method of selection, on the other hand, is the analysis of a pool of variants, or mutants, under a set of selection conditions that yield the most efficient or functional variants. In the past, the difficulty associated with selection methods was grounded in identifying a realistic and efficient method with which to link genotype with the yielded proteins. Phage display has provided a method with which to select for specific function and efficiency, while simultaneously providing a reliable link between protein and genotype. Filamentous phage is used as a base for the display of peptides and proteins, while containing the desired deoxyribonucleic acid (DNA) sequence within its genome. The method was initially described by Smith in 1985, and has since been utilized to alter the function of polypeptides, and discover novel properties. It can even be used in conjunction with high throughput sequencing to map protein mutations, and identify amino acid mutations that are integral to novel protein function or increased efficiency. Phage display follows a simple scheme, as described in Figure 1. A protein of interest and a substrate are identified. The protein and substrate can represent protein-protein, protein-peptide, or protein-DNA interactions, among others. The gene encoding the protein of interest is then inserted into the bacteriophage genome, and expressed on the phage’s coat proteins. Selection (for
binding, affinity, catalysis, etc.) is completed via washing and elution of phages under various selection parameters.

**Figure 1.** Schematic displaying a general phage display experiment.

**A.1. Phage Biology**

Ff filamentous *Escherichia coli* bacteriophage are most commonly used in phage display. These bacteriophages have proven ideal for phage display for a number of reasons. They are stable in a range of selection parameters. The filamentous phage genome is relatively small, about 6400 nucleotides, and tolerates insertions in a number of non-essential/non-coding regions. The phage does not kill the host cell, and thus
can replicate in high volumes.\textsuperscript{2,6} Lastly, the bacteriophage coat can be mutated or alerted while retaining the phage’s infectivity.\textsuperscript{2}

\textbf{A.1.1. Structure}

The structure and function of Filamentous bacteriophages has been widely studied. The Ff class of filamentous phage is made up of f1, fd, and M13, which are all f-pilus dependent viruses; they use the F conjugative pilus as a receptor.\textsuperscript{2,6} These bacteriophages encase a circular, single-stranded, DNA genome in a long protein capsid cylinder.\textsuperscript{2,6,7} The phage genome encodes for 11 proteins in total (Figure 2). Two proteins are involved with phage export, the entire capsid is made up of five coat proteins, and the other four proteins are minor coat proteins and are used for cell infection, and phage replication.\textsuperscript{2}
<table>
<thead>
<tr>
<th>General Function</th>
<th>Ff Bacteriophage Protein</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replication</td>
<td>pII / pX</td>
<td>pII initiates replication by nicking the RF and unwinding the + strand through a rolling circle mechanism. pX assists pII</td>
</tr>
<tr>
<td></td>
<td>pV</td>
<td>Inhibits RF conversion by binding single-stranded DNA + strands</td>
</tr>
<tr>
<td>Virion Structure</td>
<td>pVII</td>
<td>Starting point for phage assembly</td>
</tr>
<tr>
<td></td>
<td>pIX</td>
<td>Starting point for phage assembly</td>
</tr>
<tr>
<td></td>
<td>pVIII</td>
<td>Major coat protein that forms the capsid</td>
</tr>
<tr>
<td></td>
<td>pIII</td>
<td>Ending point for assembly, binds to F pilI, and serves as a display protein</td>
</tr>
<tr>
<td></td>
<td>pVI</td>
<td>Ending point for assembly, mediates release of assembled virions</td>
</tr>
<tr>
<td>Assembly / Export</td>
<td>pi / pXI</td>
<td>Both mediate virion assembly at the inner cell membrane</td>
</tr>
<tr>
<td></td>
<td>pIV</td>
<td>An exit channel through the cell membrane during assembly</td>
</tr>
</tbody>
</table>

**Figure 2.** Ff bacteriophage proteins and functions.

The gene VIII protein (pVIII), the major coat protein, is comprised of 50 amino acids and composes the protein cylinder that encapsulates the phages single-stranded DNA (ssDNA).6 pVIII molecules pack tightly, overlapping like fishescales,2 to form a right-handed helix. Phage particles have a fixed diameter of 6.5 nm, and their length is determined by their genome size.2 For example, the length of the 6400 nucleotide ssDNA Ff phage particle is 930 nm,6 while the length of a smaller variant of with 221 nucleotides is 50 nm.8 Thus, genome insertions result in longer phages. In order to contain these insertions, the phage elongates by producing more pVIII proteins.7 The
length of an insertion must, however, be considered since longer phages can be prone
to instability and breakage (Figure 3).

![Image: Gene insertions into the phage genome elongate the phage particle. The phage accomplishes this by producing and incorporating more pVIII capsid proteins in the major coat.]

**Figure 3.** Gene insertions into the phage genome elongate the phage particle. The phage accomplishes this by producing and incorporating more pVIII capsid proteins in the major coat.

One end of the phage contains about 5 copies of both the gene VII and gene IX proteins (pVII and pIX). pVII, 33 residues, and pIX, 32 residues, are located at the end of the phage where the genome's packaging signal is located. Phage assembly begins at this end of the phage, and both proteins are vital for phage formation.

The other end of the phage house 5 copies of both the gene III and gene VI proteins (pIII and pVI). pIII, 406 residues, is composed of three glycine-rich regions (Figure 4). Two of the regions are required for cell infection; one is required for binding to the cell membrane, and the other is necessary for the transportation of phage genome into the cytoplasm and insertion of phage proteins into the cell membrane. The third glycine-rich region is necessary for the formation of stable phage particles. This third pIII region, along with pVI, 112 residues, interacts with pVIII to form the second end of the phage particle. Both of these proteins are involved in the detachment of newly formed phage from host cell membranes.

Not much else is known.
about the function of pVI, except that it seems to form a complex with pIII,\textsuperscript{12,13} and that it degrades in the absence of pIII.\textsuperscript{11}

![Diagram of pIII structure and fusion points](image)

**Figure 4.** Structure of pIII and fusion points for display on phage. The CT domain (residues 257-406) is embedded in the phage particle, while the N domains (N1; residues 1-68, and N2; residues 87-217) are exposed. Glycine-rich linker regions connect the three domains (Gly1 and Gly2). Positions 198 and 249 are typical positions for protein of interest gene fusion. Image adapted from Ref. 2.

### A.1.2. Cell Infection

The f-pilus dependency of Ff bacteriophages, indicates that these phages typically infect bacteria cells that express and display F pili on their cell membranes.\textsuperscript{2,6} These pili exhibit normal assembly-retraction cycles,\textsuperscript{6} in which they continuously extend upon assembly and retract upon disassembly.\textsuperscript{2} Binding of the phage pIII to an F pilus retracts the pilus and brings the phage into close proximity of the cell surface.\textsuperscript{6}
Once the phage is close to the cell membrane, another domain within pIII binds to a co-receptor Tol protein (the F pilus is the receptor).\textsuperscript{2} Interaction of the pIII end of the phage with the host cell’s Tol proteins is vital for phage infection.\textsuperscript{14,15} Upon infection, the phage’s single stranded DNA genome is translocated into the cell’s cytoplasm, while the phage coat proteins are embedded within the cell membrane, where they await phage replication.\textsuperscript{2,6}

### A.1.3. Phage Replication

Phage replication can be separated into three distinct stages.

First, the phage’s single-stranded DNA is converted into a double-stranded molecule called replicative form (RF) via the host cells RNA and DNA polymerases and topoisomerase.\textsuperscript{2} The bacteria’s replicative enzymes and molecules replicate the original viral single-stranded DNA, + strand, to form a complimentary – strand, producing the RF. The RF is then transcribed into mRNA. Lastly, the cellular ribosomes then translate the mRNA into proteins. These proteins are transported to the cell membrane, where they are embedded for phage release.

The second stage of the replication process is initiated by the gene II protein (pII).\textsuperscript{2} The + strand of the RF is nicked by pII, producing a free 3’ end in the DNA strand. DNA polymerase then elongates the + strand DNA, beginning at the 3’ end and using the – strand as template. DNA polymerase is further assisted by pII, which unwinds the + strand through a rolling circle mechanism, thus, allowing DNA polymerase to fully replicate the DNA strand.\textsuperscript{2} Lastly, the new + strand is recircularized by pII, and converted into another RF molecule.
The third stage of replication is initiated by the gene V protein (pV). In the early stages of infection/replication the concentration of pV is low, and single-stranded DNA particles are converted into RF with little inhibition.\(^2\) As more RF is accumulated, and the concentration of expressed phage proteins increases, the presence of pV affects the replication cycle. pV is a binding protein which surrounds and binds newly synthesized single-stranded DNA + strands before they can be converted to RF. This is a vital step, as the double-stranded RF cannot be packaged into phage. These pV/single-strand complexes are then transported to the cell membrane,\(^2\) where minor and major coat capsid proteins have been embedded. The pV/single-strand complex is oriented such that the DNA begins to protrude from the cell. At this point, pV binding proteins release, and are replaced by the capsid proteins. As the final capsid proteins are assembled onto the new phage, the replication process of the phage terminates, and the newly synthesized phage particle is released from the host cell.

A.2. Display Proteins

Display on phage is achieved via fusion of the gene encoding the protein of interest to phage protein genes. The most commonly used coat proteins used for phage display are pVIII and pIII. pVIII, the protein that composes the phage cylinder, is used less frequently than pIII. Two factors contributing to this are: the protein can only display short sequences of (6-8) residues (longer sequences can prevent the packaging of phage molecules, likely due to interactions with pIV),\(^2\) and the protein will display on all copies of pVIII in a phage (this can cause avidity effects with multiple binding interactions and make it difficult to accurately characterize function of the protein of interest). There is a limitation to the use of pVIII; a requirement of short sequences
reduces the number of target peptides that can be analyzed. Recent work has extended
the applicability of pVIII phage display. pVIII proteins have been engineered to display
large polypeptides at numbers sufficient enough to produce an 100-fold improvement
in detection of fusion proteins. Other work has developed successful fusion of
proteins of interest to an artificial major coat protein (pVIII mimic) at the C-terminus
which expands the range of proteins that can be displayed on pVIII (as opposed to the
N-terminus, which is generally the location of insertion on pVIII).17

The common site of insertion for proteins of interest is the pIII gene. This gene
codes for the expression of five copies of pIII on the end of the virion. Unlike pVIII, pIII
can tolerate large insertions while retaining the phage’s infectivity. pIII is made up of
three domains (N1, N2, and CT), and two glycine-rich linker regions. The N domains are
located in the exposed N-terminal, and serve as contact points for interaction between
the phage and ligands. The CT domain is located in the C-terminal, which is buried in
the phage, and is linked to the assembly of the virus. The protein of interest can be
inserted into the pIII gene at the end of the N2 domain (residue 198) or to the glycine-
rich linker region that is near the CT domain (residue 249). Both fusion locations have
successfully displayed proteins on pIII in filamentous phage.18,19

A.3 Phage Library Construction

A.3.1 Phage Vectors

Once a display protein has been chosen as the fusion site for a protein of interest,
it must then be cloned into the gene and incorporated into the phage genome for
display. This is accomplished with a display vector. Some vectors are commercially
available, and are ready for use in various experiments. For other proteins, however,
that have not been displayed on phage, a vector must be constructed and its functionality must be tested. Two major factors in display vector construction are vector type, and valency.\textsuperscript{1,2} Display vectors are predominately of two forms: phage vectors and phagemid vectors (Figure 5).\textsuperscript{1,2,20} Valency refers to the number of foreign proteins that a phage displays.\textsuperscript{21} Multivalent, or polyvalent, phage are those that display more than one copy of the foreign protein of interest per phage particle. Monovalent phage predominately display one protein of interest per phage particle. Confirmation of display on phage is the final step in the construction of a vector. Vectors must successfully be able to facilitate display of the protein of interest, such that it is both displayed and functional.
Figure 5. Comparison of phage and phagemid vector cloning for pIII display. Target protein sequences are inserted before gene III. Both vector methods carry a phage replication origin (Ff) to permit phage genome replication. Phagemid vectors also carry a plasmid origin of replication, and an antibiotic resistance marker. Helper phage facilitates viral protein expression, but is packaging-deficient and does not compete with the phagemid genome during assembly. In this experiment, monovalent phage display is preferred as it reduces the avidity effects that plague polyvalent display. Image adapted from Ref. 2.

A.3.1.a. Phage and Phagemid Vectors

Proteins displayed using vectors based directly within the Ff phage sequence – those incorporated directly into the coding sequence for either pIII or pVIII – are considered phage vectors. Phage vectors typically result in the display of the protein of interest on all copies of the coat protein, and, thus, are only used in polyvalent display. This phage vector display method has been used to analyze small peptides, and even functional enzymes. Polyvalent display, however, features some drawbacks. Polyvalent display prevents the identification of high-affinity clones. The use of polyvalent phage display to analyze the binding affinity of peptide libraries to a (beta)-endorphin antibody recovered sequences that were weakly binding relative to the original starting sequence. Similarly, attempts to use polyvalent phage display to select for high-affinity mutants of human growth hormone (hGH) from among a library of mutants with large binding affinity variations (over 20-fold) were unsuccessful. These studies found that polyvalent phage display results in multi-point attachment to ligands during selection, as a result of the number and proximity of displayed proteins of interest. This multi-point attachment, thus, makes weak-binding clones appear as if they have high affinity, an effect known as avidity (the strength of a receptor-ligand complex of N receptor-ligand interactions). Additionally,
phage vector display is complicated by the difficulty of cloning directly within the phage genome without disrupting important promoters, genes, or other genome structural components. For these reasons, monovalent phage display is preferred for experiments seeking to select mutants with the highest affinities.

Monovalent display was developed to account for, and avoid, potential avidity effects associated with polyvalent display. Phagemid vector display has been shown to produce monovalent display phages that can be used to sort between low- and high-affinity mutants. Phagemid vector display, unlike phage vector display, does not feature the introduction of the gene of interest directly into the phage genome. The displayed protein gene is, instead, fused to the pIII or pVIII genes in a plasmid under the control of a weak promoter. The plasmid typically contains: a plasmid origin, a phage replication origin (Ff origin), and an antibiotic resistance marker. Thus, this plasmid contains the sequence necessary for the production of the pIII-protein fusion and for packaging of single stranded vector into phage particles, but is incapable of encoding viral genes. Thus, the phagemid vector technique requires the infection of E. coli containing the phagemid with helper phage that is packaging-deficient, it has a compromised Ff origin. The helper phage introduces the genome necessary for the cell to express the viral proteins necessary for phage replication. Because the cells will express all of the wild-type proteins and the fusion protein encoded in the phagemid, the extruded phage particles will contain both proteins. Additionally, the deficient packaging of the helper phage genome ensures that most of the phage particles will contain the more stable phagemid genome, thus preserving the genotype-phenotype link. Typically, the wild-type proteins, including wild-type pIII, will be expressed in
excess relative to the fusion protein. This results in <10% of the phagemid particles displaying only one copy of the pIII-fusion protein, <1% displaying two copies, and the rest of the particles displaying solely wild-type pIII. The phage particles displaying only wild-type proteins will be eluted during selection, and, therefore, do not affect selection for binding affinity. Similarly, the small percentage of phage particles that display two copies of pIII-fusion protein is statistically insignificant upon multiple rounds of selection.

Not only does monovalent display, and, as an extension, phagemid vector construction mitigate any potential avidity effects, they also feature additional advantages.

The first advantage of monovalent display is the conservation of phage infectivity. The majority of pIII proteins on phagemid constructed phage are wild-type, thus these phage are able to infect host cells with nearly identical rates to wild-type phage. Polyvalent display of pIII proteins, on the other hand, results in display phage that are less infective than wild-type phage.

Another major advantage of phagemid display vectors is that they can be used to also freely express the protein of interest for biochemical analysis and sequencing. The addition of an amber stop codon (TAG) between the inserted gene and the pIII gene in the phagemid facilitates the controlled free expression of the protein of interest. Expression of vectors containing the amber stop codon within E. Coli strains that suppress the stop codon, amber-suppressors, allows for protein translation through the stop codon, resulting in pIII-fusion proteins. Vectors of interest (ones that have gone through a selection process, perhaps) can then be transformed into non-suppressor E.
*Escherichia coli* strains, resulting in the expression of the protein of interest (not fused to pIII).\textsuperscript{2,22} This method has been used to effectively display and express high levels of non-phage-linked Fabs and hGH without additional phage genome manipulation.\textsuperscript{19,27}

**A.3.1.b. Vector Display Confirmation**

Confirmation that a newly constructed phagemid vector, or phage vector, actually displays the protein of interest is vital. Phage enzyme-linked immunosorbent assays (ELISA) is a successful method for establishing the functional display of a protein on phage. In the case that display cannot be confirmed, modifications to the fusion gene, vector type, or display protein used must be considered.\textsuperscript{2}

**A.3.2. Library Construction**

Once a display vector has been constructed and confirmed for a protein, the protein can then be diversified. Protein diversification allows for the analysis of effects that various mutations have on protein folding, structure, and activity. Libraries with diversities of approximately $10^{10}$ are reasonably constructed,\textsuperscript{28,29} with some methods reaching even larger orders of magnitude. Diversification can aim for complete randomization, allowing for mutations among all amino acids and across the entire gene sequence. Other diversification methods can produce biased or incomplete mutant libraries, which can be composed of specific amino acid subsets or limited to specific locations within the gene sequence.

Regardless of which library construction method is used, the ability to create large libraries of mutant proteins allows for the analysis of, and insight into, the interactions in receptor-ligand complexes. Phage display is a valuable protein engineering tool as it allows for successful evaluation of a large library of protein
variants, in the range of $10^8$, as previously discussed. Traditional screening methods are unable to evaluate nearly as many variants, typically limited to 10-100, due to time and material constraints.

**A.3.2.a. Oligonucleotide-directed Mutagenesis**

Site-directed mutagenesis specifies mutation locations within a gene sequence, as opposed to completely random diversification. This method of mutagenesis features the replacement of specific codons in the gene sequence with degenerate codons that can encode for more than one amino acid. These degenerate codons are made up of mixtures of nucleotides at different positions in the codon’s three-nucleotide sequence. For example, the codon NNS is able to encode for all amino acids (where N = any nucleotide, and S = C or G). Because site-directed diversification requires the use of synthetic DNA, and synthetic DNA composition can be altered, facile control of diversity is achievable. For example, alteration of the design of the degenerate codons will dictate the focus and breadth of the library’s diversity. For complete diversification of a sequence, hard randomization, the use of degenerate codons that encode for all amino acids, is used. Alternatively, tailored randomization restricts the degenerate codons to those, which only encode for specific amino acid subsets. Lastly, soft randomization, a more restrictive model, limits degenerate codon variation and biases the mutations to a predetermined sequence.

Oligonucleotide-directed mutagenesis is a simple method for constructing site-directed mutant libraries as described in Figure 6. The method can be utilized to conduct hard randomization, tailored randomization, or soft randomization. The general scheme for oligonucleotide-directed mutagenesis remains follows the same
general protocol, regardless of the type of degenerate codon design.\textsuperscript{28} A template plasmid, containing the gene of interest, is first prepared for diversification. Transformation of the template plasmid into an \textit{E. Coli dut}⁻/\textit{ung}⁻ strain, like CJ236,\textsuperscript{31} purifies the plasmid into a single-stranded form that has a significant amount of thymine bases replaced by uracil.\textsuperscript{28} The synthetically designed oligonucleotide, containing the desired mutations, then anneals to bases preceding and following the desired mutation region. This oligonucleotide then serves as a primer for a DNA polymerase, which replicates a complementary strand to the template plasmid. A DNA ligase then ligates the replicated strand into a circular, double-stranded DNA. The resulting double-stranded DNA is not entirely complementary; the oligonucleotide region containing the mutation does not match the template. The introduction of the double-stranded DNA into an \textit{E. Coli ung}⁻ strain results in the preferential replication of the new mutagenic DNA strand.\textsuperscript{28,31,32} The higher the purity of a template, the more efficient the oligonucleotide mutagenesis is. Efficiency can range from >50\%, to >80\% depending on template purity.\textsuperscript{28}
Figure 6. Schematic displaying oligonucleotide-directed mutagenesis. a) A synthetic oligonucleotide is annealed to a single-stranded DNA template that has a significant amount of thymine bases replaced by uracil (ss-DNA). The oligonucleotide contains a degenerate codon, and is designed to encode mutations (*). b) A double-stranded DNA (CCC-dsDNA) is synthesized by polymerase and ligase. c) The CCC-dsDNA is then transformed into *E. coli* cells where it is replicated to either the mutant or wild-type sequence. Image adapted from Ref. 28.

A major limitation of oligonucleotide mutagenesis is the fact that it is far from 100% efficient, meaning that libraries constructed using this method will contain wild-type proteins. In cases where the wild-type protein exhibits high affinity for binding to the target, the wild-type proteins displayed on phage will be selected for alongside mutant proteins.28 Thus, it is imperative that mutant libraries constructed via oligonucleotide mutagenesis are purified of the wild-type protein. This can be accomplished by rendering the template inactive. An inactive template permits the construction of mutant libraries, retaining the functionality required for oligonucleotides to bind to the template strand, but does not allow for the expression of the wild-type protein encoded in the template. The incorporation of stop codons within the desired mutated region of the template sequence achieves this.28,33 Translation of
the stop codon incorporated template sequence results in truncated proteins. Because they are no longer translated as complete, pIII-fused proteins they cannot be displayed on phage. The inactive template still serves as the wild-type genome, yet can no longer interfere with selection.\textsuperscript{28} In general, templates that contain at least one stop codon per target region will likely be rendered inactive.

\textbf{A.3.2.b. \textit{In vitro} DNA Recombination}

Another robust library construction method is the use of \textit{in vitro} DNA recombination in directed evolution. Directed evolution is an \textit{in vitro} simulation of natural Darwinian evolution. A protein of interest is run through iterative rounds of mutation, selection, and screening. Each round is akin to a generation, or another representative period of time, as in Darwinian evolution. This method allows for the production of mutants with specific properties.\textsuperscript{34,35}

Stemmer reported the first use of \textit{In vitro} DNA recombination as a method for evolution.\textsuperscript{36,37} DNA shuffling, as Stemmer called it, incorporates a low rate of point mutations within recombination of related sequences of the template.\textsuperscript{34} Other \textit{In vitro} DNA recombination methods have followed Stemmer’s DNA shuffling. These include random priming recombination,\textsuperscript{38} and Staggered extension process (StEP).\textsuperscript{39}

\textit{In vitro} DNA recombination has proven successful in a number of applications. Stemmer initially utilized DNA shuffling to improve drug resistance.\textsuperscript{36,37} Others have used the method to improve and alter various protein characteristics and functions. The thermostability of enzymes has been generated and improved by recombination random mutagenesis and selection.\textsuperscript{40–42} Additionally, the folding and solubility for single-chain antibody fragments,\textsuperscript{43} and the green fluorescent protein were improved.\textsuperscript{44}
While this library construction method has proven successful in the optimization and analysis of protein function, it does feature drawbacks. One of its major limitations is a result of the method’s logic. Because *In vitro* DNA recombination requires that the “best,” or most optimized, sequence be selected and then further mutated, it does not allow for the discovery of other successful mutations. Those mutants can only be discovered in separate recombination experiments; in which they are found to be the “best” mutant. This is a seemingly small limitation, yet can result in the loss of vital, or optimal, mutant discoveries.

**A.3.2.c. Other Methods**

Other methods for diversification of a mutant library have been described. These include but are not limited to: the oligo cassette method, recursive polymerase chain reaction (PCR), oligo-splinted assembly, error-prone PCR, and the use of PCR with randomized primers. Other methods for phage display library construction involve collections of genes, as opposed alterations to single genes (cDNA libraries, and antibody V-region libraries, for example). Choosing a method for a specific experiment depends on a few factors. First, the type of mutagenesis must be considered. Random mutagenesis can lead to novel function discoveries, while site-directed mutagenesis allows for the analysis of the impact that specific positions, or mutations have on function. Second, the target region for diversification must be defined and considered. If the mutagenesis of an entire gene is desired, it will require a different method to mutagenesis that aims to contain mutations within a defined region. Lastly, the efficiency of a method must be weighed against its accuracy, faster methods may not necessarily have more accurate results, and vice-versa.
A.3.3. Selection

A.3.3.a. General Approach

Once a library of mutant proteins has been constructed and displayed on phage, the library population must be sorted by binding affinity. This stage in phage display is called selection, and results in the optimization of protein binding to a specific target. Generally, selection will allow for the separation of tight binders, or “fit” binders, from non-binders or weakly binding variants. Consecutive rounds of successful selection yields the emergence of a single clone, or set of clones, that have relatively optimal binding affinities.\textsuperscript{3,20}

A.3.3.b. Methods

Selection methods all follow a general idea, the separation of “good” binders, those that retain or have developed high binding affinity to the target relative to a standard like wild-type affinity, from non-functional or weak binders. The type of selection method, however, depends on the mechanism of binding that is being analyzed, or optimized. In general, selection can be separated into two methods: immobilization of the displayed protein, or binding of the protein within solution.\textsuperscript{20} Additionally, selection conditions must be stringent enough to allow for the retention of fit mutants and the extraction of other variants from the mutant pool. These conditions must also, however, allow for the recovery of fit variants in order to enable further selection rounds and final analysis.\textsuperscript{20} Therefore, a selection strategy must also facilitate pairing with an efficient and effective elution strategy.

A.3.3.b.1. Immobilization
The phage target can be immobilized onto a solid support, thus rendering the pIII-fusion protein : target interaction immobile. The general mechanism follows as such: the target (protein, molecule, ligand, etc.) is covalently attached to the solid support, then the phage library, in the presence of buffer, is introduced into the system inducing interaction, and finally the mutant variants that do not bind are washed away and those that do are eluted for further analysis.\textsuperscript{20} The solid state target can select for fit variants in one of two ways. Variants can be selected for binding affinity to a substrate covalently attached to the solid support. Thus, only variants with a relatively high binding affinity will remain in the system after subsequent washing. In the case that catalysis optimization is the goal, another setup can be utilized. All variants can be attached to the solid support, where they are introduced to a substrate, and then catalysis can result in cleavage of fit variants from the solid support. Upon which washing of the system would elute fit variants, while less-fit variants remain attached to the support.\textsuperscript{5}

Elution of directly immobilized phage typically follows one of two general schemes. In the situation where selection conditions are designed to immobilize phage based on binding affinity to the solid support, elution aims to dissociate the phage from the binding target, or to dissociate the target, with the attached phage, from the stable support itself.\textsuperscript{20} Phage dissociation from target can be induced by alteration of pH levels to extremes or addition of denaturants; phage infectivity can be retained after exposure to extreme conditions.\textsuperscript{2} Alternatively, binding targets can be covalently attached to solid supports via specific tags, or engineered cleavage sites (cloning specific amino acid chains, or binding sites into a protein is an example of engineering). Attachment to solid
support can be cleaved at these sites upon the incorporation of specific molecules or conditions. For example, thrombin can be added to a solution to cleave an engineered thrombin site on a binding target protein.\textsuperscript{20}

If selection aims to optimize the catalytic function of a protein, as previously described, the elution conditions focus on collection of catalytically active variants. For catalytic processes that involve cleavage, selection and subsequent elution can be accomplished by inhibiting cleaving function during the selection stage, and then inducing cleavage as an elution strategy. Pedersen et al. were able to optimize the enzyme staphylococcal nuclease (SNase) for the catalysis of DNA cleavage by utilizing the immobilization method, and eluting variants that had successfully cleaved from the solid support.\textsuperscript{5} The activity of the nuclease was inhibited by control of Ca\textsuperscript{2+} levels; SNase does not cleave its substrate in the absence of Ca\textsuperscript{2+}. Following selection, elution was accomplished through the introduction of Ca\textsuperscript{2+}, which detached fit variants from the solid support by inducing cleavage.

### A.3.3.b.2. Solution Binding

Phage library variants can also be selected via binding in a solution and subsequent capture by a solid support (i.e. an affinity matrix). This method can reduce the potential avidity effects of polyvalent phage, because interaction with a target in solution does not have the same intramolecular effects as binding to targets on a solid support.\textsuperscript{20} On a solid support, as previously discussed, a polyvalent phage will initially bind to a single target, but will undergo multiple binding events as a result of target proximity the multiple displayed proteins present on a polyvalent phage particle. Solution binding can still facilitate multiple binding events, however, the likelihood of
these subsequent binding events is not increased to the degree that binding on a
surface does. This is because each target molecule is independent, and binding does not
inherently bring the phage closer to other targets.

Elution strategies for solution bound phage are similar to those for phage
selected with the immobilization strategy. Generally, elution strategies aim to break the
link between the phage and the binding target, or the affinity matrix. The introduction
of basic or acidic conditions, at levels that the phage can tolerate (which are typically
extreme for most other molecules or organisms), is a typical technique for
accomplishing this. For example, hydrochloric acid (HCl) and triethylamine are typical
reagents.²

A.3.3.b.3. Selection Conditions

The first round of selection, typically, segregates functional variants from non-
functional variants.²⁰ This initial step is focused on producing a resultant library that is
diverse, but contains only functional variants. Further rounds of selection aim to obtain
the variants with the highest levels of function. Thus, the stringency of selection
conditions will increase to remove low affinity variants. Methods for increasing
stringency include longer washing times, increased number of washings, the addition of
denaturants,⁵² increased temperatures or the addition of competitive binders or
ligands.⁵³ The increased stringency results in the selection of more and more optimal
variants, with the final selection aimed at isolating the “best” variant from a specific
library.

B. High-throughput Sequencing.
While phage display is an effective and efficient method for selecting protein variants with specific functions and affinities, it does not address the issues associated with the analysis of the selected variants. Traditionally, mutational analysis was the primary method with which mutant variants of a protein were examined. Traditional scanning of mutants does have its drawbacks.\(^4\) Originally, mutants were cloned, expressed, and purified in preparation for assays and further analysis. Phage display, a method which directly and accurately links phenotype to genotype, reduces the need for purification of individual protein variants.\(^4\) The requirements of Sanger sequencing, however, still restricted the number of total variants that could be analyzed to a few thousand. Recent work in high-throughput sequencing, such as combinatorial scanning,\(^5\) and short-read Illumina sequencing,\(^4\) has substantially expanded the volume of variants that can be assessed and sequenced in a reasonable period of time.\(^5\)–\(^7\) For example, combinatorial “quantitative saturation (QS) scanning” enabled effective and efficient assessment of the interaction between human growth hormone (hGH) and its receptor (hGHR).\(^5\) High-throughput sequencing, coupled with phage display selection can be used to map resistance of protein mutants to antibiotics or anticancer drugs.

**B.1. Short Read Illumina sequencing**

The principle behind Illumina’s high-throughput sequencing method is that a DNA template strand is broken into smaller fragments, and then the bases of those fragments are identified as each section of the entire sequence is re-synthesized as described in Figure 7. The DNA to be sequenced is extracted and then fragmented into a collection of short segments. These segments are then simultaneously sequenced via
Illumina’s sequencing reactions. The resulting sequences, or reads, are then reassembled in comparison to a reference genome. In the case that a reference genome is unavailable, or does not exist, the segments can be reassembled via de novo sequencing methods.\textsuperscript{58,59}

**Figure 7.** Illumina high throughput sequencing scheme. 1) Multiple DNA sequences are extracted from selected phage. 2) Sequences are fragmented. 3) Each set of fragments is labelled with a tag. 4) Fragments are simultaneously sequenced. 5) Fragments separated by tag. 6) Each set of sequencing reads is compared to a reference sequence.

High-throughput sequencing of an entire library of DNA (genomes extracted from a phage display library, for example) can be similarly sequenced. The above method is simply run in parallel for each sequence of interest. This is accomplished by
attaching a specific known barcode sequence to each DNA fragment originating from a specific original sequence. The fragments are all pooled and sequenced simultaneously, and the barcode sequences are used to distinguish fragments. Each set of fragments is then reassembled, based on a reference genome or via de novo sequencing. This sequencing technique can produce large amounts of sequencing data, with small, measurable margins of error.\textsuperscript{4,50,59}

\textbf{B.2. Application and Results}

Fowler \textit{et al.} used high-throughput sequencing to successfully investigate the binding of the human Yes-associated protein 65 (hYAP65) WW domain to its cognate peptide ligand for upwards of 600,000 hYAP65 mutants.\textsuperscript{4} The hYAP65 WW domain was displayed on the surface of T7 bacteriophage, and mutant variants were selected for binding affinity to the cognate peptide, attached to a solid support via beads. The WW domain library varied, or mutated, 33 residues within the domain, and consisted of 99 varied bases. The input variant library, and the libraries after three and six rounds of selection were sequenced via short-read Illumina sequencing. The sequence data was then used to map and compare the three libraries, highlighting the relative abundances of mutations at specific locations through six rounds of selection (Figure 8).
Figure 8. Sample high throughput sequencing maps. Maps display the frequency of reads of a specific amino acid mutation at a specific position in the target protein for the variant library after zero (initial library), three, and six three rounds of selection, respectively. Image adapted from Ref. 4.
Through the use of high-throughput sequencing, coupled with phage display, Fowler et al. were able make conclusions about mutations at specific positions within the sequence that improved protein function. More importantly, the mapping of sequences obtained via high-throughput sequencing could be used to identify positions within a gene that are vital for protein function, structure, or folding.¹

C. RAS Oncogene

C.1. RAS Structure/Function

Four distinct, yet very homologous, RAS proteins are encoded for in the human genome. The HRAS, NRAS, and KRAS genes encode HRAS, NRAS, KRAS4A, and KRAS4B.⁶⁰ RAS proteins are small-GTPase membrane bound proteins, ~21 kDa, that are primarily associated with signal transmission in cells by pairing cell surface receptors with intracellular signaling pathways.¹⁴⁷ They cycle between active, GTP-bound, and inactive, GDP-bound, states; acting as ON/OFF switches for cell growth, differentiation and growth.⁶¹ Guanine nucleotide exchange factors (GEFs), which can facilitate the exchange of GDP for GTP on GTPase proteins, activate RAS by promoting the GTP-bound state.⁶⁰,⁶² On the other hand, GTPase-activating proteins (GAPs), which accelerate the hydrolysis of GTP into GDP, promote the formation of the inactive, GDP-bound, state.⁶⁰,⁶² This GAP mediated inactivation of RAS is the main target of somatic mutations found in oncogenic RAS variants.⁶⁰ Mutations in RAS at amino acid positions G12, G13, or Q61 interfere with the ability of GAP to inactivate GTP-bound RAS. In the presence of any of these mutations, activated RAS will persist within the cell where it will signal through various pathways facilitating the transformation of mammalian cells.⁶⁰,⁶²
RAS carries signals to these pathways by binding proteins, or other ligands, at the effector domain. Binding at this domain, spanning residues 32-40, is essential for signal transduction. It has been shown that mutations in, or near, this region impair interaction between RAS and its binding partners. RAS has specific binding interactions at the RAS effector domain with a number of RAS effectors, such as: phosphatidylinositol 3-kinase (PI3K), p120 RAS GAP, and RIN1.

C.2. Oncogenic RAS

RAS genes have been found to have oncogenic function via point mutations (at amino acid positions 12, 13 and 61, for example), and that these mutations occur in 30% of human cancers. Thus, it has been found that RAS proteins are the most commonly mutated oncogene in human cancer. For this reason, inhibition of RAS oncogenic activation has major implications in cancer therapy.

The oncogenic activity of the four RAS isoforms (HRAS, NRAS, KRAS4A, and KRAS4B) has not been completely determined, nor has the issue of whether some isoforms exhibit unique activity been resolved. On one hand, the isoforms follow a non-random distribution within the range of known cancer types. Mutations in HRAS are frequently detected in skin, head, and neck tumors. KRAS mutations are associated with colorectal tumors, and lung and pancreatic carcinomas. Lastly, NRAS mutations are typically found haematopoietic and lymphoid tissue tumors. On the other hand, RAS isoforms can substitute other isoforms within specific tissues, where they perform the same function as the oncogenic RAS isoform typically associated with that specific tissue. Aberrations from this trend have been discovered; especially when RAS
isoforms are analyzed *in vivo*. Thus, it has been difficult to determine the exact mechanisms and interactions that oncogenic isoforms have within the cell cycle.

Independent of RAS isoform variations, oncogenic RAS has phenotypic effects across a wide range of typical cell functions. In fact, the contribution of oncogenic RAS to human malignancies is typically underestimated by the frequency of RAS mutations found in cancer types.\(^{60}\) Mutations in RAS have a wide impact on cancer function, and the growth and spread of cancer. Oncogenic RAS can affect the promotion of cell proliferation, the suppression of cell apoptosis, metabolism, the cell microenvironment, immune response evasion, and metastasis (Figure 9). All of these are vital mechanisms, functions, or components of cells. Alterations in any of them can induce cancer, or tumor growth.
### C.2.1. Cell Proliferation

A common characteristic attributed to carcinoma is the continual expansion of cancer cells; that is, the growth of cell population via uncontrolled, or atypical, cell division. Cell proliferation depends on the interference of appropriate cell response to mitogenic, and anti-mitogenic, cell signals.\(^{60}\) Since RAS is directly involved in the transmission of mitogenic signals within a cell, oncogenic RAS is able to promote cell proliferation via a number of mechanisms. The expression of oncogenic HRAS has been discovered to push a GO phase-arrested cell, a cell that is in a non-dividing equilibrium state, toward cell division without stimulus from cell growth factors.\(^{72}\)

---

**Figure 9.** Cancer-related impacts of oncogenic RAS. Table adapted from Ref. 60.
Additionally, RAS oncogenes can facilitate cell proliferation by signaling for the upregulation of various transcription factors, such as: FOS, activating transcription factor 2 (ATF2), JUN (leucine zipper protein, and ELK1.\textsuperscript{60} The main mechanism in which oncogenic RAS mediated upregulation of transcription factors leads to cell transformation is the expression of cyclin D1, the G1 cyclin; a regulator of cell cycle progression and a transcription co-regulator. This mechanism is validated by the fact that mice lacking cyclin-D1 have been found to exhibit resistance to carcinomas and tumors that are promoted by oncogenic HRAS.\textsuperscript{73,74}

There are other ways in which oncogenic RAS can affect cell proliferation via the upregulation of pathways that lead to cell division (or other cellular pathways that affect such pathways) or the downregulation of factors involved with the inhibition of cell growth. Overall, due to the massive role that RAS plays in the regulation of cell population it is not surprising that RAS oncogenes have been found in a wide range of cancers.

\textbf{C.2.2. Cell Apoptosis}

The elimination of, or interference with, signaling pathways that induce cell apoptosis is a major characteristic of cancer, and the spread of cancer cells. Apoptosis is a programmed process of cell death that can be initiated by extracellular or intracellular factors. Typically, apoptosis is a response to high levels of cellular stress, and acts as a defense mechanism against malignancies.\textsuperscript{60}

Oncogenic RAS erodes apoptotic pathways via the upregulation of anti-apoptotic factors and downregulation of pro-apoptotic factors. Chin \textit{et al.} discovered that the RAS oncogene not only arrests cellular apoptosis pathways, but is also important for tumor
maintenance. The downregulation of oncogenic HRAS$^{G12V}$ lead to tumor cell apoptosis, and subsequent regression of melanomas. Similarly, the removal of oncogenic KRAS$^{G12D}$ also caused tumor cell apoptosis, accompanied by the regression of lesions and malignant tumors.

C.2.3. Metabolism

As RAS oncogenes induce high levels of cell proliferation, accompanied by the erosion of apoptotic pathways, the need for cellular building blocks increases to accommodate the increased rates of growth. In order to proliferate, a cell must accomplish two tasks. It must generate sufficient energy and produce biomolecules at a rate that meets proliferation demands. Thus, cancer is dependent on metabolic pathways as they provide the components needed for the construction of new cells, and provide mechanisms for energy production.

In order to meet these needs of proliferation, cancerous cells will shift from mitochondrial oxidative phosphorylation (OXPHOS) to aerobic glycolysis. Glycolysis in typical mammalian cells is inhibited by the presence of oxygen, which is used by the mitochondria to oxidize pyruvate to carbon dioxide and water. Aerobic glycolysis instead, converts pyruvate to lactate. This mechanism allows the cell to recover the NAD$^+$ required for the cycle repeat. This self-sustaining cycle can produce more ATP, and at a faster rate, than OXPHOS, while also providing metabolic intermediates that can be used for the production of biomolecules required for cell proliferation.

Oncogenic RAS facilitates aerobic glycolysis through numerous pathways. First, oncogenic RAS stimulates the shift from OXPHOS to glycolysis in cells via the upregulation of hypoxia-inducible factor 1α. Additionally, oncogenic RAS improves a
cells ability to transport the glucose necessary for glycolysis by upregulating the glucose transporter GLUT1. Lastly, oncogenic RAS affects the synthesis of biomolecules by increasing the levels of enzymes associated with the processing of glycolytic intermediates. Hence, RAS oncogenes play a critical role in the cancer-characteristic metabolic shift that cells undergo.

**C.2.4. Cell Microenvironment**

As the number of cells in a tumor growth increases, the competition for oxygen and nutrients among cells also increases. Angiogenesis is the growth of new blood vessels from pre-existing blood vessels. Angiogenesis within a tumor allows for the delivery of adequate supplies of oxygen, among other nutrients, to sustain tumor growth. Oncogenic RAS reaches beyond the confines of the cell membrane to impact the extracellular microenvironment, and induce angiogenesis. It accomplishes this by upregulating angiogenic growth factors and by changing the composition of the extracellular matrix, the system of biological molecules that provide structural and signaling support to cells. For example, oncogenic RAS targets vascular endothelial growth factor A (VEGFA), a promoter of endothelial cell growth and the formation of blood vessels. Additionally, matrix metalloproteinase 2 (MMP2), MMP9 and urokinase-type plasminogen activator (uPA) are upregulated by RAS oncogenes, resulting in the removal of physical constraints to angiogenesis. The role that oncogenic RAS plays in the promotion of angiogenesis indicates that RAS does not only increase tumor growth intracellular, but can also impact the extracellular environment.

**C.2.5. Evasion of Immune Response**
Humans have mechanisms within their immune systems that can identify and defend against cancerous tumor growth. These immune responses are vital components of the body’s defense against cancer. Tumors that can subvert these response, are able to grow unimpeded by antitumor immunity pathways, which can lead to metastasis. RAS oncogenes can provide cells with mechanisms with which to evade immune responses within the body.60

Oncogenic RAS has been found to suppress the effectiveness of the immune response via two mechanisms.60 First, RAS oncogenes can reduce the expression of antigens on cell surfaces, thereby aiding in the cell’s ability to evade detection of lymphocytes.81 Second, oncogenic activity of RAS can actually overcome immune response, as opposed to evading them. For example, while patients with some forms of cancers induced by oncogenic RAS possess T-cells that are specific for antigens formed by these RAS variants, these lymphocytes are rendered angenic and cannot actually bind the cancer cells.60

C.2.6. Metastasis

Metastasis is the biological process in which tumor cells spread from the organ of origin to surrounding organs and tissues. Cancers in this stage are difficult to treat due to the mobile nature of the tumors, and the sheer number of locations in which tumors are and can be located. Many tumors that have developed metastatic properties contain oncogenic RAS variants.60

Metastasis is initiated by local tumor cell invasion. In order for a tumor to enter this stage, individual tumor cells must develop the ability to detach from the tumor.60 Oncogenic RAS facilitates cancer cell acquisition of this property by affecting
interactions between cells and the extracellular matrix, and by promoting cell
phenotypes that aid in migration.\textsuperscript{60} RAS oncogenes are able to reduce cellular adhesion
to the ECM by decreasing the presence of integrins that promote cell linkage to the
ECM.\textsuperscript{62} Additionally, cells must develop a front-rear asymmetry in order to migrate.\textsuperscript{60}
RAS oncogenic activity is able to induce the structural shifts necessary for a cell to
adopt front-rear asymmetry.\textsuperscript{60}

In order for a cancer to continue through metastasis, cancer cells must be able to
detach from the original tumor and enter the bloodstream. Oncogenic activity of RAS
has been shown to degrade the ECM, and impact the integrity of the basement
membrane in order to promote angiogenesis for cell proliferation. The degradation of
the basement membrane, which is vital for the containment of cancer cells, and the
increase of blood vessels near the tumor aid in a cell's progression through the
metastatic process.\textsuperscript{60} Thus, oncogenic RAS is able to further facilitate the migration of
cancer cells.

C.3. Cancer Therapy via RAS Inhibition

RAS has been shown to facilitate tumor growth through numerous pathways,
and mechanisms. Oncogenic activation of RAS has been commonly found in the
progression of human carcinomas via these pathways.\textsuperscript{62} Therefore, the inhibition of the
activation or function of RAS oncogenes has been a major focus in cancer therapy.

There have been attempts to completely inhibit the function of RAS.\textsuperscript{62,83} RAS
function, however, is involved in many cellular processes. For that reason, complete
inhibition or alteration of its function is not ideal. The focus has shifted to cancer
therapeutics that aim to reduce the effects of oncogenic RAS without completely abolishing all of its function.62

C.4. RAS and the RAF-MEK-ERK Signaling Pathway

C.4.1. RAF Structure/Function

RAF, a serine/threonine kinase, is critical for cell growth and regulation in mammals.84 RAF is directly downstream of RAS in signal transduction pathways that affect cell apoptosis, proliferation, and division.85 Furthermore, oncogenic RAS variants effectively bind, and activate RAF, resulting in cell processes that promote tumor growth. RAS oncogenes have been shown to induce tumor-related proliferation in cells by up-regulating cyclin D1 and down-regulating a cyclin-dependent kinase inhibitor, p27, through RAF and its signaling pathway.86,87 Additionally, oncogenic-RAS mediated signaling through RAF results in abnormal apoptotic activity. RAF activation results in both pro-apoptotic and pro-survival signaling. The balance of these two signaling effects determines cell action; oncogenic RAS pushes the balance toward pro-survival signaling, and tumor growth.60,88

Vertebrate genomes encode three RAF genes: A-raf, B-raf, and c-raf-1.84 All three genes feature three highly conserved regions (CR1, CR2, and CR3) within species-dependent conserved variable sequences.84 The C1 region contains both an 81 amino acid sequence (residues 51-131) binding active site, and a zinc-finger component.84 The binding active site is called the RAS binding domain (RBD), and provides a pocket that the RAS effector domain uses to bind RAF (Figure 10). The zinc-finger of the CR1 region has been found to be involved with RAS binding. (Bruder J.T. 1992) The CR2 region contains many serine and threonine residues that participate in the regulation of RAF
via phosphorylation/dephosphorylation mechanisms. Together, CR1 and CR2 make up the regulatory part of RAF proteins, while CR3 contains the catalytic binding domain for downstream activation.\textsuperscript{84} RAF mutants with extensions to the amino terminal of the protein, CR2 region disrupting mutations, alteration of the zinc-finger structure, and even complete deletion of the CR1-CR2 encoding region,\textsuperscript{89} are still constitutively active, and have oncogenic characteristics.\textsuperscript{84}

\textbf{Figure 10.} RAF-RBD contact points with the RAS effector domain. Image adapted from Ref. 89.
C.4.2. Overview of Signaling Pathway

RAF is the initial protein involved in the cytoplasmic kinase cascade, a signal transduction pathway that affects cell proliferation and differentiation (Figure 11).\textsuperscript{84} The mechanism of this kinase cascade is relatively straightforward. RAF activates the MAP/ERK protein kinase (MEK), the RAF protein’s only known substrate.\textsuperscript{84} In turn, MEK activates the mitogen-activated protein kinase (MAPK, or ERK, extracellular-signal-regulated kinase) which transmits the initial signal to various proto-oncogenic transcription factors and other protein kinases.\textsuperscript{90} Cell signal receptors are connected to the RAF-MEK-ERK pathway via RAS.\textsuperscript{84}

\textbf{Figure 11.} The RAF-MEK-ERK kinase cascade signaling pathway. Image adapted from Ref. 62.
C.4.3. RAS Involvement in Pathway

The RAF-MEK-ERK kinase cascade is activated by the binding of RAF by RAS. It has been shown that the RBD of RAF is crucial for interactions between RAS and RAF.\textsuperscript{91,92} RAS signaling relies on a highly conserved mechanism\textsuperscript{62} that features RAS protein switch I and switch II regions and the RBD of RAS effector proteins, which includes over 100 other mammalian proteins in addition to RAF.\textsuperscript{67,91,93,94} Upon binding, RAS recruits RAF to the plasma membrane.\textsuperscript{84,95} At the plasma membrane, in the presence of co-factors, RAF initiates the cytoplasmic kinase cascade.

Oncogenic activity in this pathway has been linked to malignant transformation of cells, cancer proliferation, and the evasion of cell apoptosis.\textsuperscript{95} The RAF-MEK-ERK pathway participates in the expression of cyclin D1.\textsuperscript{60} Thus, oncogenic RAS can yield overexpression of cyclin D1, which is associated with uncontrolled cell growth and cancer progression, via this pathway. Additionally, the RAF pathway contributes to the suppression of cell-apoptosis induced by oncogenic RAS. RAS-activated RAF downregulates prostate apoptosis response 4 (PAR4), a protein that inhibits transcription and acts to promote apoptosis,\textsuperscript{96} and upregulates anti-apoptotic factors, such as BCL-2.\textsuperscript{97} Furthermore, RAS-activated RAF has been demonstrated to phosphorylate BCL-2-associated agonist of cell death (BAD), which results in the inactivation of the BCL-2 family anti-apoptotic factors.\textsuperscript{98}

C.5. Past RAS Targeting Drug Therapies

RAS’ role in eukaryotic kinase pathways positions it as a viable drug target for inhibition of uncontrollable cell growth, and RAS hyperactivation. Prevention of productive interactions between RAS and its downstream effectors is, in theory, an
effective inhibition method.\textsuperscript{61} This could be achieved by limiting the amount of RAS in the GTP-bound state, reducing the overall number of RAS proteins in a cell, or directly blocking protein-protein interactions between RAS and its binding partners.\textsuperscript{61,67} While these seem like reasonable inhibition mechanisms, there is an issue with direct RAS inhibition. Inhibitors that cannot distinguish between wild-type and mutant RAS proteins would likely result in high levels of toxicity.\textsuperscript{61} Therefore, inhibition mechanisms that target the binding domains of specific RAS binding partners, such as RAF in the case of this experiment, can provide effective RAS inhibition, while reducing resultant toxicity. This is possible because RAS affects multiple signaling pathways.\textsuperscript{62} By targeting a specific RAS/RAF effector interaction, drugs that inhibit binding can affect the mutagenic effects of oncogenic RAS without adversely impacting other vital RAS function in the cell.

\textbf{C.6. Rigosertib}

Rigosertib, a benzyl styryl sulfone, is a small molecule inhibitor of the RBD interactions between RAF and RAS. The drug is currently in phase III clinical trials for myelodysplastic syndrome (MDS), a type of cancer. Rigosertib binds to the RBDs of a number of RAS effectors, and limits RAS' ability to bind to them.\textsuperscript{67} This in turn reduces RAF activity, and inhibits the RAF-MEK-ERK pathway (Figure 12). Rigosertib has been shown to inhibit tumor growth \textit{in vitro} and limit growth \textit{in vivo}.\textsuperscript{67,99}
Rigosertib specifically binds to six proteins: three RAF variants (A-RAF, B-RAF, and C-RAF), FUBP3, Hsp27, and Hsp73. In the RAF protein interactions, Rigosertib bound similar residues to those which RAS has been found to bind (all of which are conserved in the RAF family of proteins). Based on these studies, Athuluri-Divakar et al. determined that Rigosertib, in fact, acts as a RAS mimetic and utilizes a similar binding mechanism for interactions with RAF.

Inhibition of oncogenic RAS activity is a reasonable and effective target for cancer therapy. Rigosertib does exactly this. By acting as a RAS mimetic, Rigosertib
inhibits binding between RAF and RAS, thus abolishing function in the RAF-MEK-ERK kinase cascade. This drug has promising implications for the future of therapeutic intervention in RAS oncogenic activity, and cancer therapy in general. RAF, however, may develop mutations that reduce the binding affinity of Rigosertib. In the case that these mutations also reduce binding function between RAS and RAF, they may not have oncogenic effects; while Rigosertib will not inhibit RAS and RAF interactions by binding RAF, the binding interaction between RAS and RAF will still be inhibited by the mutation. Some mutations may, however, have oncogenic effects. RAF may develop a mutation that negatively impacts RAF binding by Rigosertib, yet retains the binding function between RAS and RAF. In this case, RAF will have developed functional resistance to Rigosertib, and oncogenic RAS could signal through the RAF-MEK-ERK pathway uninhibited by Rigosertib.

By combining phage display selection of RAF variants, with high-throughput sequencing of variants selected for binding affinity to RAS in the presence of Rigosertib, we can provide a method for understanding and mapping RAF resistance to the anti-cancer drug, Rigosertib. This information can be used to alter the structure of Rigosertib or its mechanism of binding to anticipate resistance, or for the development of other therapeutics.

D. Proposed Experiments

Our experiment proposes to pair the phage display of a library of RAF variants with high-throughput sequencing of variants selected for functional resistance mutations to Rigosertib to shed light on potential mechanisms for Rigosertib resistance in the human organism. An overview of our experiment is described in Figure 13.
**Figure 13.** Overview of proposed experiment.

Aim 1: Create library of RAF mutants and use phage display to select for mutant binding affinity to RAS

Aim 1 will result in a phage displayed library of mutant RAF variants that have been validated for binding affinity to RAS. A library of RAF variants displayed on phage will be created, and selected for binding to RAS as described in Figure 14. Previously characterized effects of specific RAF mutations will also be used to assess successful library creation. Block et al. have characterized specific amino acid positions found in the RBD of RAF (residues 51-131) that appear to be critical for RAS/RAF binding. This
data will be compared to the mutant RAF library as an additional validation of successful library creation, and accurate functionality of the RAF variants.

**Figure 14.** Schematic displaying Aim 1 selection. The phage displayed RAF variant library is selected for RAS binding function via biotinylated RAS, which binds to streptavidin beads. The selected variants are then sequenced and compared to the input library.

First, RAF must be displayed on phage. The gene encoding the RAF protein will be fused to the filamentous phage M13 gene encoding pIII in a phagemid vector, and transformed into E. Coli cells with the addition of helper phage (M13K07, as previously described). We have decided to fuse RAF to pIII with the phagemid vector method because it produces phages that exhibit monovalent display. Monovalency is desirable as it reduces any avidity effects that would result from the display of multiple copies of RAF on a single phage. Additionally, since phagemid vector construction results in a larger quantity of wild-type pIII displayed on phage, phage infectivity will be conserved. Lastly, this method will facilitate biochemical analysis and sequencing of selected phage; free expression of RAF can be conducted with phagemid vectors. Prior to protein diversification, successful display on phage will be confirmed with a phage enzyme-linked immunosorbant assay (ELISA), a successful and commonly used assay for display confirmation. (xxref Russel M. Lowman H. B. 2004)

Upon confirmation of display on phage, a library of RAF variants will be created. The library will be constructed by random mutagenesis within the RBD of wild-type
RAF. Error-prone PCR, a variation of mutagenic PCR, can be used to introduce non-biased, random mutations within a desired segment of a target sequence.\textsuperscript{100} For the purposes of this experiment, error-prone PCR provides distinct advantages to oligonucleotide site-directed mutagenesis, and \textit{In vitro} DNA recombination. We are aiming to uncover functional mutations within the RAF RBD, and thus desire random mutagenesis across the entire target region. Oligonucleotide site-directed mutagenesis specifies mutation locations in the target sequence, and cannot efficiently produce a library of RAF variants with random RBD mutations. \textit{In vitro} DNA recombination is prone to the loss of functional mutations as a result of the method's concept. Furthermore, error-prone PCR features a number of additional advantages. Primarily, it is a simple and efficient method. It closely follows a well understood PCR mechanism, with slight variations, and can produce target levels of mutagenesis relatively quickly.\textsuperscript{100} Furthermore, error-prone PCR can be controlled for specific rates of mutagenesis.\textsuperscript{101}

Error-prone PCR, as a method of random mutagenesis, takes advantage of the relatively high error rate of \textit{Taq} DNA polymerase. In a typical PCR reaction the error rate is around $10^{-3}$ errors per nucleotide.\textsuperscript{102} The relatively low-fidelity of Taq polymerase can be further exacerbated by varying the reaction conditions, such as increasing the concentration of Mg\textsuperscript{2+} or Mn\textsuperscript{2+} of the reaction mixture.\textsuperscript{101} Cadwell et al. utilized such manipulation of reaction conditions to reach an error rate of about $7 \times 10^{3}$ errors per nucleotide with minimal sequence bias.\textsuperscript{102} Additional parameters may be altered in order to increase or decrease this error rate to produce the desired rate of mutagenesis of the target sequence.\textsuperscript{100}
For the purposes of this experiment, the aim is to produce a library of mutants that contain an average of one mutation per variant. This rate of mutagenesis will allow for the accurate attribution of any potential novel function to a specific amino acid change. While variants with more mutations might exhibit novel function, it would be difficult to link novel function to the mutations that caused them. As the RBD, an 80-residue region, is composed of more than 200 base-pairs of DNA the error-rate that Cadwell et al. managed to reach, a frequency of 0.66%±0.13% errors per position, will be sufficient to guarantee at least one mutation per variant.100

To diversify the RAF gene fused to pIII, error-prone PCR will follow a standard PCR protocol with vital modifications, as previously described.102 Some of the modifications that are engineered to increase the mutation rate are: an increased MgCl₂ concentration to facilitate the pairing of noncomplementary bases, addition of MnCl₂ which increases the polymerases error rate, an increase in dCTP and dTTP concentrations, and an increase in quantity of Taq polymerase units.

We will confirm that the error-prone PCR method successfully diversified the RAF gene. The quantity and molecular weight of the product will be confirmed via an agarose gel embedded with ethidium bromide.100 Lastly, we will confirm achievement of the desired mutation rate by sequencing a sample of the PCR product. Successful implementation of this protocol will result in a phage displayed library of RAF mutants that includes no more than 10⁸ mutants.48

Next, the newly synthesized library must be selected for binding to RAS. Selection will be achieved by a solution binding technique; active RAS binding RAF mutants will be collected upon successful binding to wild-type RAS. Solution binding
was chosen over alternative selection techniques, such as immobilization, because it decreases the likelihood that multiple binding events occur. This will mitigate collection of variants that appear to have high binding affinity, but in reality only appear to do so because of avidity effects.

The wild-type RAS, the substrate in this case, will be prepared for selection via biotinylation, the process of covalently linking a biotin molecule to a protein, or other molecule. This process typically does not impede or affect protein function due to the relatively small size of Biotin. In our experiment, wild-type RAS will be enzymatically biotinylated. Chemical biotinylation, another major biotinylation method, risks protein inactivation and low biotinylation yield, while enzymatic biotinylation conserves protein function and leads to high yields. The biotin carboxy carrier protein (BCCP), a 13 amino acid peptide, will be fused to wild-type RAS. This sequence contains a lysine that is specifically biotinylated by BirA, a biotin ligase found in E. Coli. The BCCP-tagged RAS will be transformed into E. Coli and biotinylated as previously described. Successful biotinylation will be verified by ELISA, with absorption at 405 nm directly proportional to the amount of biotinylated products.

Upon successful biotinylation of wild-type RAS, the phage library will be incubated in solution with the biotinylated RAS. Streptavidin-coated magnetic beads will be added to the solution. Only the RAF variants that retain binding functionality should bind the biotinylated RAS, and thus only those variants will be attached to the streptavidin beads. The bound phages will then be removed from the solution by magnet; the streptavidin-coated beads are magnetic. Any background phage will be removed by buffer washes, while bound phage will remain bound to the streptavidin
beads. The selected phage will then be eluted from the beads by 100mM triethylamine, a reagent that is not toxic to Ff bacteriophage.

Selection will be monitored by incubation of eluted phage samples in *E. Coli* cells, as previously described. The cells will be plated, incubated, and then colonies will be counted. Phage concentration (cfu/ml) will be determined. A minimum of $10^4$ cfu/ml is expected, as this is the number of phage usually required for growth in a culture.

Successful selection of RAS binding RAF variants will yield between $10^4$ and $10^9$ cfu/ml.

As a control for this selection, the phage library variants will also be selected for in the absence of biotinylated RAS. None of the phage variants should be attached to the streptavidin beads in this trial, if this is not the case then the phage library must be further analyzed to find the cause of binding, or simply re-constructed.

Following selection, successful binding will be experimentally validated. The bound phages will then be sequenced, and compared to the input RAF variant library to control for library biases and other potential artifacts. Finally, mutations that retained RAS-binding affinity will be identified and compared to recent studies.

For validation of binding after selection, we will monitor a sample of 5-10 of the selected RAF variants for binding affinity to activated RAS, the GTP-bound state. Selected RAF mutants will be individually incubated in HeLa cells, treated with DMSO, with glutathione S-transferase (GST) fused RAS$^{G12D}$ bound to glutathione-agarose beads. The level of binding for each RAF variant will then be determined by immunoblot analysis. This technique was chosen as it has previously been used to successfully examine RAF binding to activated RAS. GST tagging of proteins is useful as it can be used to induce immobilization by glutathione, which will allow us to immobilize
functionally binding RAF variants. HeLa cells are a desirable medium for this experiment as they are a human cell line. The decision to construct the phage library with phagemid vectors will expedite this experiment; phagemid vectors can be used to express RAF variants that are not linked to phage via amber (TAG) stop codon incorporation, as previously described.

Once the selected RAF variant population has been validated for binding to RAS, it will be sequenced with high-throughput short-read Illumina sequencing techniques, as previously described. The technique requires the isolation of phage library DNA by phenol-chloroform extraction and precipitation by ethanol. The DNA is then amplified, purified, and sequenced using a Genome Analyzer IIx, an Illumina product. The input library will also be sequenced. The populations will then be compared to account for any artifacts or library biases.

Lastly, mutants that are validated for binding affinity to RAS will then be compared to prior studies. Block et al. found that mutations at specific locations within the RBD of RAF reduced RAF/RAS binding affinity (Figure 15). Validation of the library of RAF variants will depend on the mutations found in variants that were selected for RAS binding. This method will serve as a secondary validation. If any of the selected variants contain mutations that have been found to reduce binding affinity, then there are three plausible explanations. First, the less likely explanation, the findings in previous studies may be incorrect, such that the mutations that have been found to impede RAF/RAS binding actually do not do so. Second, the much more likely explanation, it is possible that the creation of the phage library either didn’t successfully incorporate mutations into the RAF sequence, or that the RAF variants were not
successfully displayed on the bacteriophage. Third, it is also possible that these mutant variants contain more than one mutation, and that the additional mutation(s) restored function in RAF in the presence of the function inhibiting mutation. In either case, these findings would require troubleshooting and additional analysis; anywhere from re-sequencing the selected variants to re-constructing the phage library.

<table>
<thead>
<tr>
<th>Amino Acid Position in RAF-RBD</th>
<th>Mutations that reduced function</th>
</tr>
</thead>
<tbody>
<tr>
<td>59</td>
<td>R59A</td>
</tr>
<tr>
<td>64</td>
<td>N64A</td>
</tr>
<tr>
<td>65</td>
<td>K65A, K65E, K65M</td>
</tr>
<tr>
<td>66</td>
<td>Q66A</td>
</tr>
<tr>
<td>67</td>
<td>R67A</td>
</tr>
<tr>
<td>68</td>
<td>T68A</td>
</tr>
<tr>
<td>69</td>
<td>V69A</td>
</tr>
<tr>
<td>84</td>
<td>K84A</td>
</tr>
<tr>
<td>89</td>
<td>R89L</td>
</tr>
</tbody>
</table>

**Figure 15.** RAF mutations with known negative impacts on RAS binding affinity. This will be used as an additional validation of the library of RAF mutants. Variants that bind RAS should not contain any of these mutations. Image adapted from Ref. 91.

**Aim 2:** Use phage display to identify RAF mutants that exhibit resistance to Rigosertib while, simultaneously, retaining binding affinity to RAS

The goal of incorporating Rigosertib as a selection condition is to evaluate whether the selected RAF variants have developed the novel function of resistance to
RAS-mimetic cancer therapeutics, primarily Rigosertib. Theoretically, RAF variants, that have been validated for binding to RAS, that are able to retain binding affinity for RAS in the presence of Rigosertib will have developed resistance to the cancer drug. These variants can then be characterized to determine which mutations developed this resistance in the variant. Identification of functional resistance mutations to Rigosertib can provide vital insight for future cancer therapy.

Once the phage library of RAF variants has been validated, it will be used to select for RAF/RAS binding function retaining RAF mutants with developed resistance to Rigosertib as described in Figure 16. Following selection and purification of these mutants, they will be characterized by high-throughput sequencing and mapped to display mutational tolerance and function conservation in the RBD.

**Figure 16.** Schematic displaying Aim 2 selection strategy. The phage displayed RAF variant library undergoes two rounds of selection for RAS binding: in the absence and presence of Rigosertib, the RAS mimetic inhibitor. Both sets of selected mutants are sequenced, and compared to the input library.

The same phage library, which has already been subjected to two selection trials (no RAS in the selection conditions, and then with the addition of biotinylated RAS), will undergo another selection trial. The library must now be selected for RAS binding in the
presence of Rigosertib, the RAS mimetic. The selection conditions will likely remain similar to the two other trials, with the exception of the presence of the inhibitory drug. Due to the mimetic nature of Rigosertib, it is likely that RAF mutations that reduce RAF/RAS binding affinity will also interfere with the binding of Rigosertib. Variants for which this is the case will be washed away during selection, as they will not be able to bind the biotinylated-RAS regardless of their resistance to Rigosertib. This is to be expected. In fact, these variants are of little interest outside of the role they may play in validating the mechanism with which Rigosertib binds to RAF. The variants that are selected, via immobilization by streptavidin beads, will feature mutations that inhibit Rigosertib binding while retaining RAF/RAS binding function.

Multiple rounds of selection are expected to produce a small subset (50-100 variants) of the original phage library that exhibit high affinity for Rigosertib resistance and RAS binding. At this point, short-read Illumina sequencing will be used to sequence these Rigosertib-resistant RAF variants, as previously described.\(^4\) This sequence data will be compared to the data of the pre-selection input library, and the phage selected in the absence of Rigosertib to control for library biases and artifacts. The sequence data will also be used to map the prevailing mutations that conferred desired function with a sequence logo, as previously described (Figure 17).\(^{108}\) These results will highlight vital mutations for resistance to Rigosertib, thus providing critical insight into potential mechanisms of resistance that may arise to the use of Rigosertib as a cancer therapy treatment.
Figure 17. Sample sequence logo. DNA sequences are aligned and used to create a sequence logo that highlights conserved nucleotides. This technique will be used to compare amino acid sequences of selected RAF variants. Image adapted from Ref. 108.

_Aim 3: Test selected mutants for functionality within the RAF-MEK-ERK pathway_

The goal of Aim 3 is to determine the downstream functionality of RAF variants that have gained RAS/RAF functional resistance to Rigosertib. While these variants
have displayed the ability to resist binding by Rigosertib, and can still be bound by RAS, they will not have an oncogenic effect if they cannot be activated by RAS, or if they cannot continue the signal transduction pathway through MEK and ERK. Therefore, selected variants must be analyzed for functional binding and activation of the other components of the cytoplasmic kinase cascade.

Selected RAF variants will be tested for functional signaling activity in the RAF-MEK-ERK pathway. The MEK-ERK pathway activation assay and subsequent protein analysis that will be used have been previously described, and conditions are shown in Table X. HeLa cells will be incubated with DMSO (negative control), or Rigosertib and treated with epidermal growth factor (EGF), which activates RAS. The cell lysates will then be treated with each of the Rigosertib-resistant RAF variants, following variant expression. Co-immunoprecipitation assays with RAF-specific antibodies and western blot analyses will be used to determine the level of RAF heterodimerization (a vital function of wild-type RAF upon RAS activation), the level of phosphorylated-ERK relative to total ERK, and the level of phosphorylated-MEK relative to total MEK. Functional signaling RAF variants will readily form heterodimers, and will induce downstream phosphorylation with little to no change relative to wild-type levels.

This method was chosen over others, such as the MAP/ERK Kinase Assay, because of its similarity to the techniques that we will use to validate binding affinity in Aim 1. We will not need to obtain many new reagents, or a new cell line. Additionally, this technique has successfully assessed RAF signaling function in previous studies.

This experiment will identify and characterize mutants that successfully activate downstream kinases, thus completing signal transduction. These mutants will
represent fully functional mechanisms for RAF resistance to Rigosertib, and potentially other RAS mimetic drugs.

D.1. Significance

Successful implementation of this experiment will yield vital information and insight into the use of Rigosertib, and other RAS mimetic therapeutics, as a cancer therapy treatment. Pharmacologists and doctors may be able to anticipate resistance to Rigosertib in vivo and administer treatments that can counter or prevent the development of RAF resistance. Rigosertib could, for example, be coupled with another therapeutic cancer drug to anticipate resistance and prevent a second wave of oncogenic RAF mutants. Furthermore, the experiment may uncover mutants that, in addition to resisting Rigosertib, successfully activate downstream kinases, thus completing signal transduction. These mutants will be identified and characterized, thus improving understanding of the binding mechanism of RAF and RAS, and their interactions with downstream kinases. The development of additional drug therapies will be aided by this characterization of oncogenic activity of RAF and RAS in the cytoplasmic kinase cascade signaling pathway.
References


(58) Whole-Genome Sequencing.


Nuclear magnetic resonance and molecular dynamics studies on the interactions of the ras-binding domain of raf-1 with wild-type and mutant ras proteins. J. Mol. Biol. 286, 219–232.


