2013

Selection of a Non-Phosphorylated Peptide Inhibitor of BRCA1's (BRCT)2 Domain

Railey White
Virginia Commonwealth University

Selection of a Non-Phosphorylated Peptide Inhibitor of BRCA1's $(BRCT)_2$ Domain

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Chemical Biology)

at

Virginia Commonwealth University (2013)

By

ELIZABETH RAILEY WHITE

B.S. Chemistry (2007)

University of Kentucky, Lexington, KY

Advisor: **Matthew C. T. Hartman, Ph.D.**

Assistant Professor, Department Of Chemistry,

Virginia Commonwealth University,

Richmond, Virginia.

May, 2013

Acknowledgment

My road getting here, and surviving once I did, has fortunately been filled with supporters of many varieties, and without them I may not have arrived as sanely as I have, or perhaps not at all. I'll start with my family, who have encouraged me from day one, and convinced me that I can do anything I put my mind to, even if it meant going along with my plans as a five-year old to become either an archaeologist or a Ghostbuster. Without that sense of spirit, I would have given up long ago. I also owe a lot to Dr. David Atwood who allowed me to naively walk into his lab as a first semester freshman and who also casually mentioned one day the MD/PhD program as a means to sway me from the "dark side" of medicine. I have great appreciation for Lisa Blue and Trish Coakley who helped me grow up as a scientist, and for that I will always be grateful.

The four years I've spent working toward this degree would have been far more difficult if it hadn't been for the amazing lab mates I've been fortunate enough to work with. Overall we have had way more laughter and smiles than tears and frustration, and anyone who has received a PhD knows that is a blessing we should all hope to be bestowed. I've had the great fortune of working with Dr. Zhong Ma, Dr. Gajanan Dewkar, Dr. Tim Reed, Michael Dcona, Deboleena Mitra, Sara Ahadi, Daniela Selaya, David Hacker, Jonathan Sheldon, and Brittany Danzig. Each one of these individuals has helped me grow as a scientist, but more importantly made me laugh, often times when I needed it most. My fellow graduate students have not just remained lab mates, but have truly become my friends. In addition I must mention my Richmond family that

always know when I need a run, a hug, a drink, or a round of "The Game of Life." In particular that means you Albéric Rogman, Hope Bailey, Sarah Vunck, Greg Hawkins and Chris Faigle.

My committee has been extremely helpful throughout this journey, and they have each lent their own expertise to this work. I would like to thank Dr. Nicholas Farrell, Dr. Kris Valerie, Dr. Vladimir Sidorov, and Dr. David Williams. I appreciate the help offered by the Department of Chemistry and Massey Cancer Center. I also thank Altria and the Concern Foundation for financial support.

Obviously, I owe many thanks to Dr. Hartman whose ideas and mentorship have forever molded the way I will approach problems. I have to remind myself that most PhD candidates are not as lucky as I have been to have such a well matched mentor. It really has been a pleasure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

SELECTION OF A NON-PHOSPHORYLATED PEPTIDE INHIBITOR OF BRCA1'S (BRCT) DOMAIN

Elizabeth Railey White

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2013

Advisor: Matthew C. T. Hartman, PhD, Department of Chemistry

A growing body of literature suggests Breast Cancer-Associated Protein 1 (BRCA1) is important not only as a cause, but also as a target in the quest for cancer treatment. BRCA1 deficient cells treated with radiation as well as PARP inhibitors and other chemotherapeutics demonstrate a greater sensitivity than cells with wild type BRCA1. Inhibitors of BRCA1 would take advantage of this synthetic lethality and represent a significant advance in cancer treatment as well as an understanding of the biology of DNA repair. Despite significant study of BRCA1 protein and function, it is a large protein (220 KDa) that is still largely uncharacterized, but its N- and C-terminal domains have been described by significant structural data. The BRCT (BRCA1 C-Terminal) Domain is a phosphoprotein binding domain that is commonly mutated or lost in cancers and has a binding cleft seemingly very suitable for drug design. Small molecule screens have been conducted against this domain, but the resulting hits with moderate affinity have not been shown to induce BRCA1 deficient phenotypes. Phosphopeptides have also been studied as potential BRCA1 inhibitors, yet despite some having affinities in the mid-nanomolar range the

presence of a phosphate is not without its pharmacologic challenges. We generated an mRNA display library with 1.3×10^{13} cyclized peptides covalently attached to the mRNA that encoded them. Eight rounds of selection exposing the library to a GST-BRCT fusion resulted in selection of non-phosphorylated peptides that bind to a BRCT domain of BRCA1. The sequences resulting from the selection have common homologies and initial characterization has shown that these peptides may be the first viable non-phosphoserine containing inhibitors of BRCA1.

CHAPTER 1. INTRODUCTION

1.1 BRCA1 and Cancer

1.1.1 BRCA1 Discovery

The idea of a hereditary form of breast cancer was first put forth in a pedigree published in 1866,[1] but it would be more than 100 years later until the discovery that this syndrome had a genetic basis in the form of a gene now known as the Breast Cancer Associated Protein 1 (BRCA1). After the initial Broca report in 1866, only a few similar familial studies were published over the next 100 years.[2] It wasn't until the 1960's that Henry Lynch began collecting information from 120 families with hereditary breast and ovarian carcinomas creating the largest cohort of its kind in the world which came to be known as the Creighton families.[2] Through genetic linkage studies in families with early-onset breast and ovarian carcinomas, King and coworkers who named the protein were able to identify the location of the responsible BRCA1 gene to the long arm of chromosome 17.[3] Applying this finding to the large sample size of the Creighton families, the location of BRCA1 on the chromosome was identified and the gene was cloned in 1994.[4] Despite knowing the sequence of BRCA1 for nearly two decades, much of the structure of BRCA1 has yet to be fully elucidated, and the functions of this important gene are continuing to emerge.

1.1.2 A DNA Repair Protein with Many Functions

BRCA1 is a large protein, weighing in at 220 kDa.[5] Shuttling between the nucleus and cytoplasm, BRCA1 primarily acts as a scaffold protein forming many different complexes with other proteins to respond to many cellular functions and DNA damage. Many of BRCA1's 1863 amino acids have undetermined structure with the exception of the N-Terminal RING (Really Interesting New Gene) domain, C-terminal BRCT (BRCA1 C-Terminal) domain, and the coiled-coil PALB2 (partner and localizer of BRCA2) binding domain. The BRCA1 RING domain is constitutively bound to BARD1 as a heterodimer known as the core complex. This complex is known to mediate E3 ubiquitin ligation.[6] BRCT domains are found in many proteins that respond to DNA damage. Many, but not all BRCT domains are phosphoprotein binding modules.[7] The BRCT domain of BRCA1 is particularly important because of its many binding partners that interact at various times depending on their phosphorylation status. The three major BRCT binding partners are Abraxas (also known as CCDC98 and FAm175A), BACH1 (BRCA1-interacting protein C-terminal helicase 1, also known as FANJ and BRiP1), and CtIP (CtBP-interacting protein, also known as RBBP8).[8] The complexes that form with these three proteins, BRCA1 as well as other associated proteins are named the BRCA1-A (Abraxas), -B (BACH1), and -C (CtIP). Together these complexes are known to regulate control of the G2–M checkpoint, BRCA1 accumulation at damage-induced foci, DNA replication, S-phase progression, DNA resection, and G2–M checkpoint control.[9] The interaction of PALB2 and BRCA1 is known to be important to mediate homologous recombination (HR) repair of DNA, but this mechanism remains unclear.[9] Other evolutionarily conserved sequences throughout the rest of BRCA1's sequence imply that there are other important functions of this protein that have yet to be discovered. Yet even with the small percentage of BRCA1 function currently known it

has still been dubbed the "master regulator of genomic integrity."[9] It is, therefore, no surprise that deficiency of this important protein can lead to a predisposition to cancer.

1.1.3 The Role of BRCA1 in the Clinical Prognosis of Cancer

For a very long time, a careful family history was the only means of determining ones risk for cancer. In the 1990's genetic linkage studies were applied to risk assessment for high risk families.[2] Since then genetic counseling has progressed, and BRCA1 sequencing is now commonly applied for individuals at risk. In addition non-sequence based BRCA1 deficiency tests, such as the protein truncation test were adopted quickly after the gene was identified.[10] Most BRCA1 testing is now done with direct gene sequencing, but the results of these tests are providing more questions than answers with respect to risk assessment and management.[11] In addition to analysis of patient BRCA1 mutation carrier status, analysis of BRCA1 deficiency in individual tumors can have a significant impact on prognosis and treatment options.[12]

BRCA1 is a tumor suppressor gene (TSG), and its deficiency results in a whole host of cellular abnormalities. This is why inherited homozygous deficiency is embryonic lethal, and why as a TSG under the two-hit hypothesis carriers of a single deficient copy of the gene are pre-disposed to developing cancer.[13] However, studies of patients with BRCA1 deficient tumors have revealed that they typically respond to DNA-damaging cancer therapies at a higher rate than do patients with tumors expressing wild-type BRCA1.[14] In the past few years, researchers have begun discover the underlying mechanisms behind this observation, and have created new therapies to exploit these insights. The most prevalent example of this are inhibitors of the enzyme poly ADP ribose polymerase (PARP). PARP inhibitors (PARPi) prevent repair of

double-stranded DNA (dsDNA) breaks via the non-homologous end joining (NHEJ) pathway.[15] Because BRCA1 deficient cancer cells have aberrant homologous recombination repair (HRR) of dsDNA breaks, they therefore become more reliant on NHEJ to repair these breaks.[16] The toxicity of PARPi on BRCA1 deficient cells is an example of synthetic lethality, and PARPi are now being applied in the clinic as a much less toxic alternative to traditional cancer therapies in patients with BRCA1 deficiencies.[17]

1.1.4 Need for Inhibitors of BRCA1

From what has been learned from patients with BRCA1 deficient tumors, it is easy to see the implication that BRCA1 inhibitors could hold for cancer therapeutics. Application of less-toxic drugs such as PARPi with a targeted BRCA1 inhibitor could be applied to many different types of tumors regardless of BRCA1 status in order to mimic the BRCA1 deficiency that lends itself so nicely to synthetic lethality based treatments. Another potential application for BRCA1 inhibitors would be to prevent PARPi resistance. Although tumors with BRCA1 mutations respond well to treatment with PARPi, they can become resistant to the treatment.[18] Because tumors are heterogeneous and constantly evolving, it is no surprise that at least some of the cells express wild-type BRCA1, or acquire a mutation that results in reversion back to wild-type. Therefore combination of a PARPi or other therapeutics with a targeted BRCA1 inhibitor could help prevent this resistance and further improve cancer treatment in the high-risk BRCA1-positive populations. Additionally, despite the incredible amount of information that is known about BRCA1's functions, it is central to a large web of cellular interactions that is still far from being untangled. Site-specific, temporally controllable, and dosable inhibitors would be a

complementary and valuable tool in further understanding BRCA1 function leading eventually to better therapeutics not only for cancer but diseases that involve other BRCA1 related proteins such as Fanconi Anemia.[19]

1.2 Peptides and Drug Development

1.2.1 A Brief History of Drug Development

Surviving ancient texts from China, Egypt, India, Greece and Rome describe remedies for ailments of all sorts, but it was not until the 18th century that medicine progressed from herbal remedies to seminal observations that became the birth of modern medicine.[20] Jenner observed in 1796 that a patient exposed to cow pox was subsequently immune to small pox leading to development of the first vaccine. Preventative medicine has certainly been revolutionized by vaccination, but arguably drugs have made an even larger impact on medicine.[21] Advances in analytical chemistry lead to significant advances such as the isolation of morphine from opium in 1815.[22] However, at this time many of the foundations of chemical theory were still being formed, so pharmacology developed as a field in its own right.[21] By the 1930's natural product screening was the focus of drug development focused primarily on antibiotic discovery.[21]

In the 20th century, biological observations combined with advances in chemistry lead to an explosion of efficacious drugs. However, as of the late 1990's only about 500 molecular targets were successfully targeted with drugs, and among these targets the majority were either cell-membrane receptors, G-protein-coupled receptors or enzymes.[23] Despite advances in drug discovery techniques including combinatorial chemistry, and in silico based drug discovery, the number of new drugs reaching the market has dramatically diminished over the past 20 years to

the point where large companies develop only a single approved new chemical entity each year.[24] The slowed productivity of the drug discovery pipeline has led to more interest in non-traditional targets. Among the promising new targets for drug discovery are protein-protein interactions (PPIs), despite their previous label of 'un-druggable.'

1.2.2 A New Role for Peptides in Drug Discovery

Attempts to map the interactome have detected thousands of previously unknown PPIs and have only begun to brush the surface of this complicated network that governs the proteome.[25] The reason these interactions have been overlooked for drug design is they are often flat and featureless[26] and therefore not amenable to disruption by small molecules that have been so successfully applied to the classic biding pockets found in enzymes and many other proteins. More recent analysis of the surfaces of PPIs has shown that these surfaces are governed by more than hydrophobic interactions and often have longer binding grooves or a series of "hot spots" that combine to mediate the interaction.[26] Despite this new perspective, PPIs often remain out of reach for inhibition with traditional small molecules. This observation has led to attempts to inhibit these interactions instead through protein mimetics, such as peptides.[27] With their larger size comes the ability to span the larger surface areas associated with PPIs which can often lead to greater specificity than generally found with small molecule drugs. Despite the great potential of peptides to act as inhibitors to a wealth of new drug targets in the form of PPIs, the idea of such approaches has been met which much skepticism.

The criticism of peptides as drugs is due primarily to their inherent susceptibility to protease degradation and lack of inherent cellular permeability, which would make them unlikely

drug candidates in the traditional sense. However, there are a surprising number of peptides and peptide-derivatives found on the market today and many tools are being developed to overcome remaining challenges.[28] Many of these peptides mostly bind to extra-cellular targets, but with an ever increasing number of methods for intracellular peptide delivery (such as CPPs,[29] pHLIP,[30] lipid[31] and nano-particle[32] based delivery systems), peptides are becoming even more viable drugs to target PPIs. The issue of stability has been addressed in many simple ways including the synthesis of retro-inverso, N-methylated, and stapled peptides. In fact recently a stapled peptide entered human trials for the first time.[33] Although peptide drugs still have several obstacles to overcome, they are emerging as a very promising new class of drugs for a vast pool of untapped drug targets.

1.3 Current Development of BRCA1 Inhibitors

1.3.1 Introduction

The radio- and chemo-sensitive phenotype associated with BRCA1 deficiency points to inhibition of BRCA1 as a potential therapeutic strategy. However, most of BRCA1's functions are mediated by protein-protein interactions (PPIs).[9] Historically, achieving PPI inhibition has been challenging due to the fact that the contact surface of PPIs is often little more than a flat, large surface void of suitable binding pockets for small molecules.[34] However, an increased interest in PPIs, and the development of intermediate-sized therapeutic agents capable of binding to large surfaces, has made the inhibition of some PPIs a therapeutically attractive strategy.[35] The complete structure of BRCA1 (see Figure 1.1) has yet to be determined, but the crystal

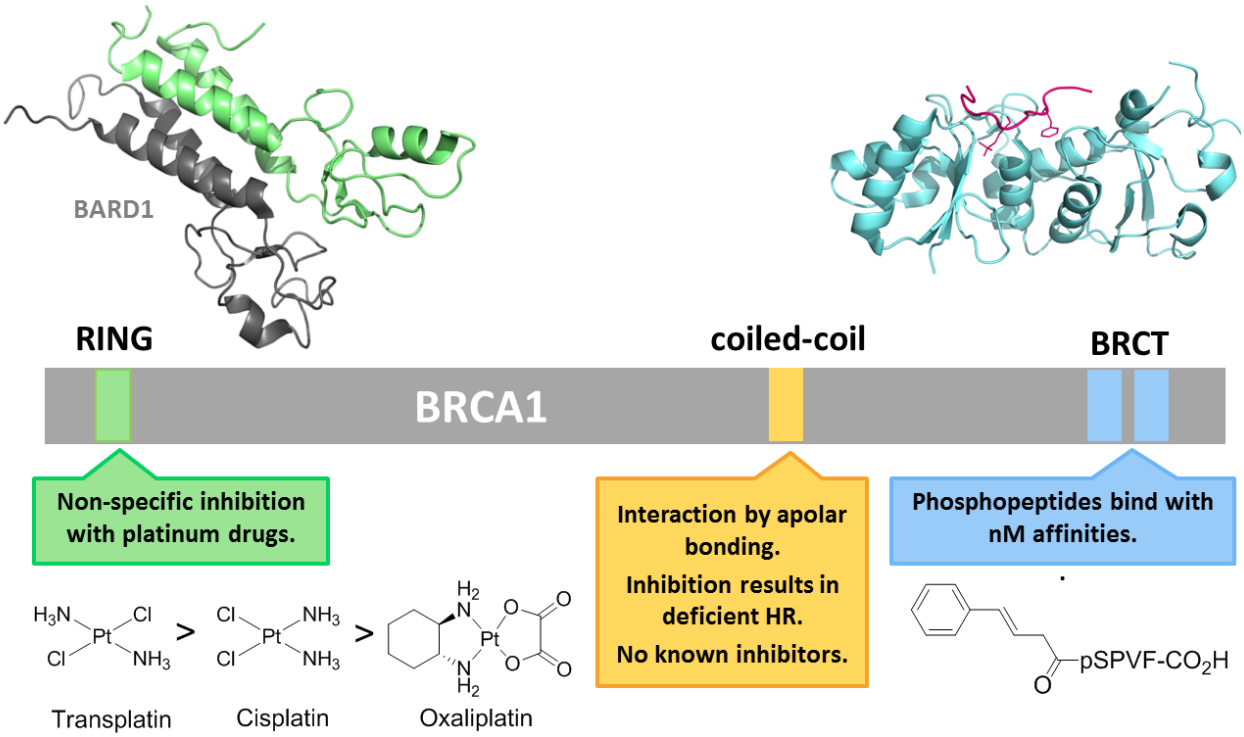


Figure 1.1 Potential BRCA1 Therapeutic Targets. BRCA1 with its RING, tandem BRCT, and overlapping SQ cluster and coiled-coil domains are indicated. Although the BRCT domain and, to a lesser extent, the RING domain, have been the focus of inhibitor design, others such as the coiled-coil domain may also be viable targets. The Zn^{2+}binding sites of the RING domain can be non-specifically inhibited by platinum compounds listed here in order of their affinity for the domain.[48] Extensive exploration of phosphopeptides that bind to the BRCT domain has resulted in the peptide shown which has a K_i of 40 nM.[47a] Structural representations are of the BRCA1 RING and BRCT domains co-crystallized with the RING domain of BARD1 and a BACH1 phosphopeptide, respectively, and were adapted from PDB entry codes 1JM7[49] and 1T29,[50] using the PyMOL Molecular Graphics System (Schrödinger, LLC).

structures of the N-terminal RING and C-terminal tandem BRCT domains are available to guide inhibitor development.[36]

1.3.2 N-terminal RING Domain

The BRCA1 RING domain is composed of a Zn^{2+} binding region of 8 Cys and His residues that form two separate Zn^{2+} binding sites and an adjacent coiled coil region.[37] This domain is known to interact primarily with BARD1, forming a heterodimer that possesses E3 ubiquitin ligase activity. Mutations that result in the loss of BRCA1 ubiquitin ligase activity, mainly due to the disruption of BARD1 binding,[38] render cells sensitive to ionizing radiation.[39] Until recently, the BRCA1-BARD1 complex was thought to be constitutive. However, it was recently demonstrated that when BRCA1-BARD1 binds to p53 in the nucleus, BARD1 dissociates, leading to the export of BRCA1 to the cytoplasm and concomitant sensitization of cells to DNA damage.[40] Therefore, inhibitors of BRCA1 and BARD1 interaction should lead to radio- and chemo-sensitization. The binding surface between BRCA1 and BARD1 is primarily composed of a 4-helical bundle, with two helices contributed by each protein. The interface is quite large (2,200 $Å^2$), and presents a formidable challenge for disruption. There is, however, some precedent for disruption of helical bundles. For example, the HIV protein gp41 assembles into a six helical bundle that is disrupted effectively with peptides, including the HIV drug Fuzeon.[41] Interestingly, most cancer-predisposing mutations in the BRCA1 RING domain occur not in the interface between BRCA1 and BARD1, but in the Zn^{2+} binding sites of the RING domain.[39b, 42] Platinum based anticancer drugs have previously been shown to preferentially bind to Zn^{2+} finger domains, replacing Zn^{2+} and thereby altering the protein tertiary structure.[43]

Recently, it was shown in vitro that platinum agents, shown in Figure 1.1, are able to bind to the RING domain and inhibit its E3 ubiquitin-ligase activity by ejecting Zn^{2+}.[42] The Zn-ligating residue H117 of BRCA1 was demonstrated to be the primary platinum binding residue.[44] Further work is required to develop specificity for BRCA1 prior to implementing this strategy in living cells.

1.3.3 C-terminal BRCT Domain

The BRCA1 tandem BRCT domain is a member of a family of BRCT motifs known to bind phosphorylated proteins involved in DNA repair as well as having other functions.[7a, 45] Sometimes these domains exist as a single motif, but often they are found in series, as is the case with BRCA1. Mutations in this domain are among the most common BRCA1 mutations in hereditary breast cancer.[36a] The BRCT domain has potential for inhibitor development due to its well-defined, relatively small binding cleft known to interact with proteins having pS-X-X-F motifs.[7a, 36a, 36c] Early work using SPOT peptide libraries identified the preferred binding sequence as phospho-Ser-aromatic β-branched/aromatic-Phe and confirmed that the phospho-Ser and Phe are the primary requirements for binding.[7, 46] The highest affinity peptide from this screen had an affinity of 162 nM.

Recently, a high-throughput assay based on fluorescence polarization to identify small molecules that bind to the BRCA1 BRCT domain was developed.[47] An initial screen of the NCI diversity database led to a single hit with an IC_{50} of 10 μM. Later, a dual fluorescence screen of 75,000 compounds identified 16 inhibitors with the lowest IC_{50} values in the single digit

micromolar range. However, some of these compounds have intrinsic fluorescence or act as fluorescence quenchers, suggesting, as the authors acknowledge, that they may be false positives.

Further optimization of peptide inhibitors by Natarajan and coworkers led to a tetrapeptide, whose structure is shown in Figure 1.1, with a 40 nM binding affinity.[48] Despite the challenges of drug delivery and cellular stability due to the phosphoserine, one report does exist showing that a phospopeptide is capable of inhibiting BRCA1 in cellular studies.[49] These studies required very high concentrations of drug (100 μM), and were examined only after very short time periods of drug exposure. With these limitations combined with the fact that the inhibitory results were small and far from clinically significant, it is unlikely that continuing to develop phosphate-based drugs for BRCA1 inhibition will be successful. Thus, therapeutically useful BRCA1 BRCT inhibitors remain elusive.

1.3.4 PALB2 and SQ Domains

At present, little is known about the BRCA1 structure outside of its two terminal domains. Much of the internal region of BRCA1 contains evolutionarily conserved sequences, but their function remains to be fully determined.[9] There is, however, the SQ cluster (amino acid residues 1,241–1,530) with a number of S-Q residues phosphorylated by ATM and ATR.[50] These regions constitute "non-druggable" BRCA1 targets except indirectly by inhibiting these kinases. Elucidating how the phosphorylation of this internal domain affects the activities of the N- and C-terminal domains is still an underexplored area, yet is critical to the design of BRCA1-based therapeutics targeting this region. In addition, the coiled-coil domain (amino acid residues 1,364–1,437) encompassed by the SQ cluster was shown to interact with PALB2 which, in turn,

associates with BRCA2.[51] Phosphorylation of S-Q residues within the SQ cluster was shown not to affect PALB2 binding.[51b] However, the disruption of this interaction by cancer patient-derived BRCA1 mutations lead to decreased HR and mitomycin C hypersensitivity,[51a] making it an interesting target from a therapeutic standpoint. To block the interaction of BRCA1 and PALB2, one approach could be the use of hydrocarbon-stapled peptides, which have been shown to disrupt protein-protein interactions involving helical interfaces.[52]

1.4 Peptide Selection with Libraries

1.4.1 Peptide Libraries: An Introduction

It is clear that BRCA1 is an important drug target, but it is a difficult target if viewed through a lens of traditional drug design. This is because its functions are mediated by protein-protein interactions, but somewhat non-traditional drugs such as peptides are ideally suited for the challenge. Much progress has been made in the rational design of peptide drugs, yet despite more advanced algorithms and increased computing power,[53] examples of high affinity peptide drugs from rational design are still rare.[54] Therefore, scientists have turned to powerful strategies for the creation of diverse peptide libraries using "molecular evolution" or "irrational design."[55] In addition to other advantage of peptides as drugs, small molecule screened as potential drugs are limited by the amount of time it takes to synthesize each unique library member, as well as the often laborious task of screening each member individually. Unlike small molecules, a wide variety of peptides can by synthesized relatively easily simply by changing the order of amino acid addition. Even the smallest peptide libraries far outreach the capacity of high-throughput screening techniques, and with the ability to easily generate a vast number of peptide members in

a library, it would be far too daunting a task to screen each library member individually.[56] Fortunately peptides have been very amenable to many different types of selections where an entire library is screened simultaneously against a single target thus dramatically reducing the time to identify lead sequences.

1.4.2. Immobilized Peptide Libraries (SPOT and OBOC)

Since the 1990s "spot synthesis" of peptides has emerged as a facile way to prepare and screen a large number of peptides against a desired target.[57] Unlike small molecule screening, a diverse library can easily be synthesized including incorporation of unnatural amino acids, and due to its immobilization onto solid support the sequence of each peptide is known by its location. Screening of a target protein against such a peptide array makes sequencing unnecessary for identification of lead peptides. The need to detect bound proteins is one disadvantage of this technique. Labeling of the protein with a fluorescent tag can lead to changes in protein properties and solubility, and secondary detection with antibodies can lead to false positives via non-specific interaction with peptides.[58] Additionally, even with advances in photo lithography a single micro-array contains only 768 members, which limits its power in the drug discovery process.[58]

Another method of generating a large peptide library on solid support is known as the "split and mix" method used in the "one bead one compound" (OBOC) approach developed in the late 1990s.[59] Instead of synthesizing peptides on known locations of a membrane support, this technique utilizes resin that is split apart before each amino acid addition then recombined and split again before the next addition. This results in the synthesis of a diverse library with a decreased number of chemical reactions.[60] This technique allows for the incorporation of

unnatural amino acids, which allows for the creation of a more "drug-like" library. Screening is often done via on bead binding with detection of bound proteins either by fluorescent probes fused to the protein or to an antibody.[61] Careful selection of resin also allows the application of flow cytometry for automated counting, which increases the speed of screening.[62] Fluorescence is proportional to the binding affinity, so leads are easily identified, but sequencing is reliant upon either Edman degradation or MS/MS sequencing, which can be challenging to perform. However, with the ability to screen 10^7 to 10^8 beads/compounds in only a few hours or less, it is easy to see the power of this technique.

1.4.3 Cellular Display Techniques: Phage and Cell Based Display

Although chemical synthesis of peptides has significantly improved since Emil Fisher's initial work in the 1930's,[63] chemical synthesis of peptides is still a highly inefficient process when compared to the power of ribosomal translation.[64] Libraries created through expression of peptide variants as surface displayed protein fusions have the dual advantages of easy "synthesis," and easy sequencing. Because each peptide library member is generated via translation, it is inherently linked to a cDNA sequence that can be PCR amplified and sequenced via routine methods allowing easy identification of library members that bind to the desired immobilized target. The library size for cell-based libraries is dependent upon transformation efficiency, and with recent advances can reach up to 10^{10} members in size,[65] but library sizes are more commonly around 10^8-10^9.[66]

Although cell based displays may provide a means to create larger libraries, they do not allow for efficient incorporation of unnatural amino acids. Although some labs have the ability

to incorporate one or two non-natural amino acids via amber and opal codon suppression as well as other orthogonal genetic incorporations,[67] these techniques are limited and not in wide use. Phage display has become the most widely used of the cell based peptide libraries due in part to its small particle size in selection, which allows for a smaller host surface to interfere via non-specific target binding compared to the size of a yeast or bacterial cell.[68] The close proximity of peptides displayed on a phage surface can lead to avidity effects where peptides bind cooperatively to generate a net high affinity. This is a disadvantage, as it complicates correlation of phage-displayed and synthesized peptide affinities.

1.4.4 In vitro Display Techniques: mRNA and Ribosome Display

Although the term display was originally used to describe the display of peptides on the surface of a host, it has come to be synonymous with library technologies that are genetically encoded. Each display technique shares the advantage that even a single peptide surviving a selection can have its sequence determined or amplified from its associated code. For bacteria, yeast and phage display this code or template of the peptide sequence is cDNA; however, as techniques to work with mRNA have improved, mRNA itself has become a viable purveyor of genotype directly, such that maintaining cDNA is not necessary to link genotype to phenotype. Using mRNA to determine peptide sequence requires reverse transcription followed by PCR amplification, but it has allowed peptide libraries to be created via cell free systems.

Ribosome Display. In Ribosome display the attachment of the mRNA template and translated peptide is mediated by the ribosome itself (See Figure 1.2). The mRNA is designed such that the code for the peptide library is followed by a spacer sequence and does not end in a stop codon

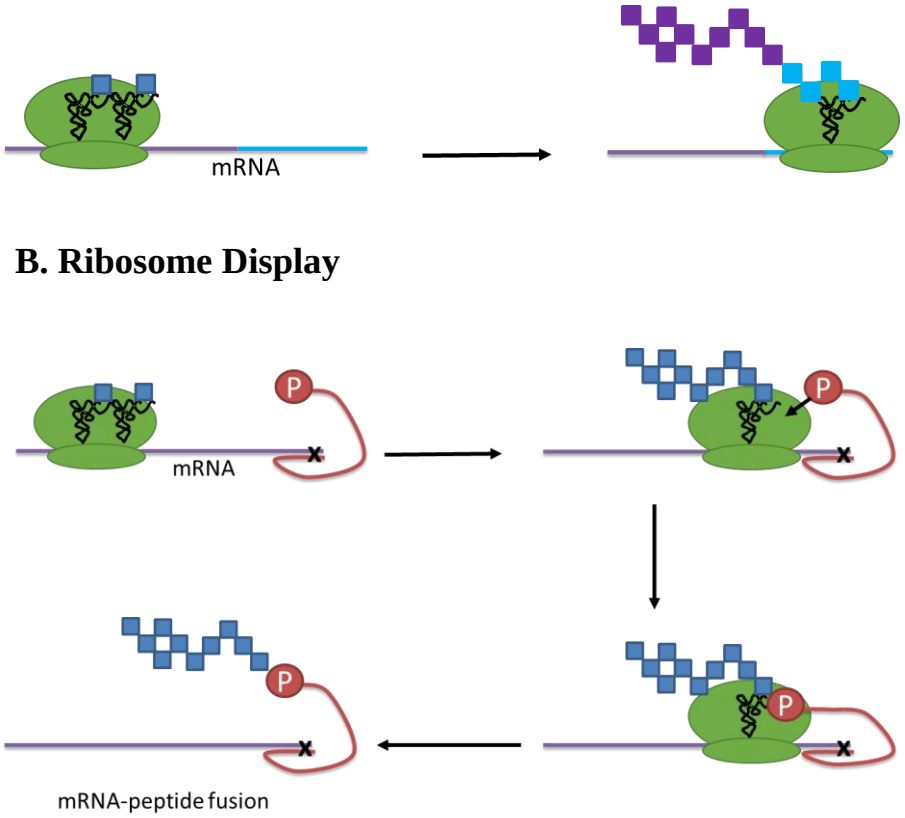


Figure 1.2 mRNA Display and Ribosome Display. mRNA display and ribosome display are two methods of generating ribosomaly transcribed peptide libraries *in vitro*. B) With ribosome display, the ribosome, mRNA and peptide remain non-covalently attached with a spacer filling the ribosomal tunnel, so that the library peptide can be "displayed" on the ribosome. A) In mRNA display, after the ribosome translates the mRNA template into peptide, it stalls at the double stranded region of the template allowing puromycin (P) to enter into the A site and form a covalent bond with the peptide forming a genetic link of genotype and phenotype.

which results in a persistent mRNA-ribosome-peptide complex. The spacer sequence is necessary so that the peptide chain can be long enough to exit the ribosomal tunnel and therefore is "displayed" during selection. The translation reaction is only 5-10 minutes, and after the reaction is quenched it must be kept at cold temperature (4 °C) throughout the selection.[69] One drawback of this technique stems from the fact that this complex is non-covalent in nature and therefore selection with ribosome display must be carefully conducted *in vitro* to preserve the fragile mRNA-ribosome-peptide complex.

mRNA display. The mRNA-peptide fusion formed during mRNA display overcomes fragile nature of ribosome display by creating a covalent linkage between the mRNA and peptide thus eliminating the need for the ribosome to persist in the complex (See Figure 1.2). This is accomplished with an mRNA lacking a stop codon which is replaced with an extended region that is complementary to a DNA oligo that acts as a linker for attachment of puromycin.[70] After annealing of the mRNA and DNA, UV crosslinking or splinted ligation forms a covalent attachment between them. Puromycin is an antibiotic that acts by inhibiting translation.[71] The slow kinetics of puromycin allow translation to continue at a normal rate without interruption; however, when the ribosome stalls due to lack of a stop signal at the juncture of double stranded mRNA/DNA the puromycin has time to enter the A site of the ribosome.[72] Once puromycin enters the A site, the ribosome catalyzes an amide bond between it and the C-terminal end of the peptide, thus creating a covalent linkage between the mRNA and the single peptide it encoded. This more stable structure allows a variety of selection conditions with no need to worry about mRNA and peptide dissociation.

Once this mRNA/peptide fusion has been generated, in reality there are two aptamer libraries: one of RNA and one of peptide. Because RNAs themselves can bind to a variety of

targets, the reverse transcription that would be necessary for eventual PCR is conducted prior to selection so that the RNA is more like a "featureless" negatively charged rod rather than a molecule with unique secondary structure which could interfere with selection. Because neither ribosome nor mRNA display are limited by transformation efficiency, libraries up to $>10^{15}$ have been generated.[73] In selection, diversity is very important, and a library of this diversity adds great power to these techniques.[74] Additionally being in a cell-free system allows more control and when conducted as a PURE translation system, UNAAs can easily be incorporated by substitution rather than cumbersome bio-orthogonal methods.[75]

1.4.5 Incorporation of Unnatural Amino Acids in Translation

The ability to incorporate UNAAs into proteins or peptides via the translational machinery offers many benefits.[76] For example, the ability to directly incorporate unnatural amino acids affords direct and simple access to functional groups generally only found in post-translationally modified proteins.[77] Similarly, the ability to site-specifically label proteins with unique functional groups not found in the standard proteinogenic AAs has enabled new means to control protein function inside cells.[78] For peptides, the incorporation of UNAAs can lead to enhanced stability and permeability, problems that have traditionally hindered the development of peptides as therapeutics.[79] For instance, peptides that contain even a single N-methyl amino acid can show enhanced bioavailability and protease stability.[80] A chief reason to pursue translational incorporation of UNAAs into peptides is that it in theory expands the chemical diversity of the already extremely diverse ($>10^{13}$-member) drug-like peptide libraries using

techniques like mRNA display.[71-72] This ability to create these libraries hinges on the development of methods to deliver UNAAs into the translation apparatus.

The first step for the introduction of UNAAs using in vitro translation is their ligation onto tRNAs. There are several strategies to achieve formation of non-natural aminoacyl-tRNAs. The original approach involved chemical attachment of the UNAA onto a dinucleotide followed by enzymatic ligation onto a truncated tRNA.[67a, 81] These chemically charged tRNAs could then be used with in vitro translation reactions.[82] This strategy has now been extended so that instead of chemically charging individual tRNAs, a whole family of orthogonal aminoacyl-tRNA synthetase/suppressor tRNA pairs that can be used to incorporate UNAAs site specifically in vivo.[83] An alternative strategy involves charging of a proteinogenic amino acid onto a tRNA, followed by converting it into an UNAA while attached to the tRNA. For example, reductive amination can convert a proteinogenic aminoacyl-tRNA into its N-methylated form. This approach can be used to synthesize peptides containing N-methylated backbones.[84] Finally, Suga, using an RNA catalyst, has developed a means to charge virtually any UNAA ester onto tRNAs. In their method, an artificial, flexible ribozyme, called flexizyme, recognizes the 3' end of the tRNA in conjunction with benzylic esters of amino acids and charges the amino acid to the tRNA.[85] Four different leaving groups have been developed that can be used to functionalize essentially any amino acid.[86] These leaving groups are recognized by three different flexizymes hypothetically allowing any amino acid to be mischarged onto a tRNA and therefore incorporated into translation products.[86] This general strategy has been applied for the incorporation of many UNAAs into peptides.[87]

Surprisingly, the wild-type aminoacyl-tRNA synthetases are able to charge a wide variety of unnatural amino acids onto tRNAs.[88] This ability suggests that such laborious engineering

approaches may not be necessary because incorporation of an UNAA would only require adding it to an in vitro translation reaction in place of its natural amino acid counterpart. While conceptually simple, standard cell extracts are highly contaminated with natural AAs, precluding this strategy. The reconstituted PURE (protein synthesis using recombinant elements) translation system[89] solves this problem because the natural amino acids can be withheld from translation. Thus the PURE system components can be tailored such that UNAAs are substituted for natural amino acids.

1.4.6 Unnatural Peptide Library Selections

Directed evolution and library selections have no doubt had significant impact on biology,[90] but there have been many fewer selections conducted with UNAAs present in the library. There have however, been several selections performed with UNAAs, that were successful in producing peptides that bind to various targets with drug-like modifications through UNAA incorporation.[91] One selection is of particular interest because it conducted as two parallel selections against thrombin, one with only natural amino acids and a second with an UNAA complement.[92] Although in this case, the UNAA peptides selected did not have higher affinities than the natural peptides selected (both had low nM affinities), they did in fact discover completely unique binding motifs that no longer bound to the target upon natural amino acid reversion. UNNA containing peptides such as these are more promising drug candidates from the outset because of their non-native structures.

1.5 Characterization of Binding Affinities

1.5.1 What is a Binding Affinity?

The term "binding affinity" is a general term used most commonly to refer to the equilibrium dissociation constant (K_d). For the simple reaction A + B → AB, a mathematical description of this constant would be the ratio of the concentration of A and B to the complex AB at equilibrium (Equation 1, concentrations are denoted in square brackets).

$$(1)\ K_d = \frac{[A][B]}{[AB]}$$

The Kinetic description of the dissociation constant is a ratio of measured on and off rates of binding (Equation 2, k_{off} = off-rate, k_{on} = on-rate).

$$(2)\ K_d = \frac{k_{off}}{k_{on}}$$

Thermodynamic analysis of a binding interaction is described in terms of Gibbs free energy (ΔG) (Equation 3, R = gas constant, T = temperature).

$$(3)\ K_d = \frac{1}{e^{-(\Delta G/RT)}}$$

The K_d of a molecular interaction can be measured through many different types of methods, many of which provide little description of the molecular details of how the molecules interact. However, there are other techniques that provide access to K_d through kinetic and thermodynamic measurements. These measurements provide additional molecular details that are important for guiding drug design. The strengths and weaknesses of several techniques for measuring the K_d are reviewed in the following sections.

1.5.2 Equilibrium Ultrafiltration Binding Assay

Many different methods have been developed to study the binding affinities of molecules.[93] Among the simplest of these techniques from a theoretical and application stand point is the radio-ligand spin assay.[94] As the name might imply, this technique requires radiolabeling of the ligand of interest. In the study of a peptide-protein interaction, this could easily be achieved via incorporation of ^{35}S-Met into the peptides' structure. A typical experiment, as shown in Figure 1.3, combines this radio-labeled peptide at constant concentration in multiple samples containing varying concentration of the protein of interest. After incubation to allow for the peptide to equilibrate between being free in solution and bound, the sample is placed into a spin filter with a membrane with specific molecular weight cut off such that the protein will remain in the upper portion of the spin filter and the unbound peptide will be evenly dispersed in both the sample remaining in the top portion of the spin filter and the flow through. Scintillation counting of the two separated solutions allows for mathematical determination of the fraction of radio-labeled peptide bound to the protein. Equation 4 holds when the volumes above and below the filter are identical after centrifugation. From a sigmoidal plot of fraction bound verses protein concentration, the K_d can be calculated as the inflection point of the graph where 50% of the protein binding sites are occupied.

$$(4)\ F_B = \frac{[\text{bound peptide}]}{[\text{total peptide}]} = \frac{[\text{top}-\text{bottom}]}{[\text{top}]}$$

These assays are technically challenging to perform because slight variations in the volume that passes through the membrane will skew the results. In addition the data obtained is not as

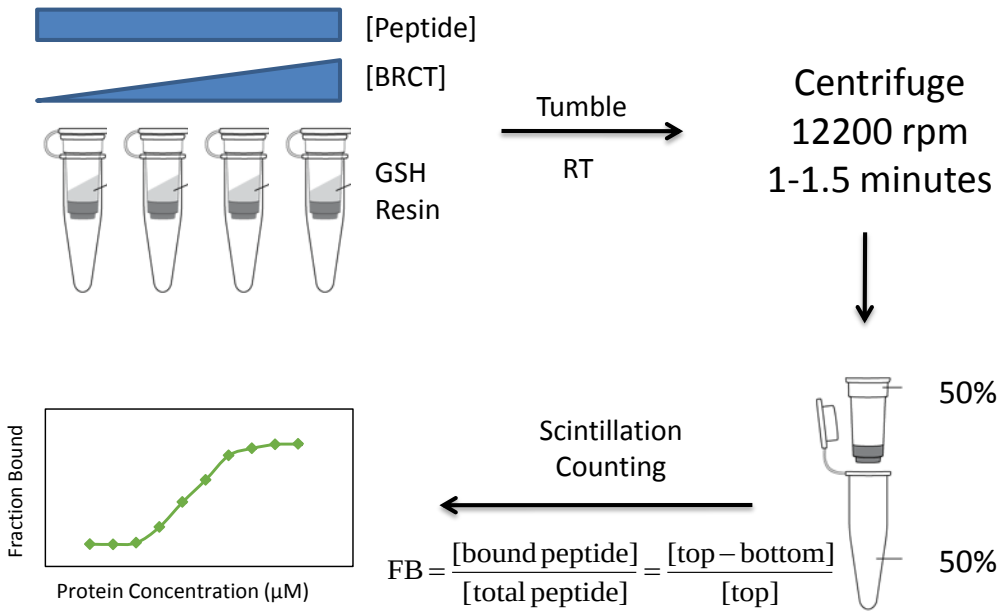


Figure 1.3 Equilibrium Ultrafiltration Binding Assay. Various concentrations of protein are added to a constant concentration of ^{35}S-labeled peptide. After incubation, the samples are centrifuged in 30,000 MWCO filters until the sample is divided, half filtered through the membrane, and half remaining in the top. A fraction bound (FB) can be calculated for each sample and then plotted to achieve a binding curve.

precise as many of the other techniques to determine binding affinity. Due to the sensitivity provided by monitoring binding with scintillation counting, only a small amount of radiolabelled peptide is need, such that peptides prepared on in vitro translation scale using ^{35}S methionine incorporation are more than sufficient.

Being the conceptually simplest techniques, as one might expect a radio-ligand spin assay does not require particularly expensive or sophisticated equipment. The primary hurdle to these assays is access to radio-labeled ligands. With radio-labeled spin assays reproducibility can be challenging. Although, this can be mathematically corrected in theory, in actual practice this is not a highly precise technique.

1.5.3 Isothermal Titration Calorimetry (ITC)

Rudimentary calorimetry qualitatively comparing the "heat of a breeding hen" and the "head of boiling water" dates back to the 17th century even before the invention of the first thermometer.[95] Since many biological interactions are accompanied by changes in heat, through the years calorimetry has become an increasingly useful tool. By the mid-20th century, calorimeter design had advanced significantly; however, it hasn't been until the past 20-30 years that advances have been significant enough to results in affordable, easy to use, stable, and sensitive enough to result in routine thermodynamic analysis of biological interactions.[95] Modern "microcalorimeters" require as little as 10 nanomoles of sample in a volume of as little as 200µL.

Isothermal Titration Calorimetry (ITC) is the modern form of Calorimetry used in laboratories to measure biological interactions. The instrument is composed of a reference chamber, containing water or a buffer, and sample chamber as shown in Figure 1.4. Both chambers are kept at a constant temperature (isothermal), and the sample chamber is fitted with a long syringe with a paddle-like end that rotates in order to stir the sample. Known quantities of the ligand are titrated into the sample chamber, and the heat of the interaction between macromolecule and ligand is indirectly measured after each injection. Because the instrument maintains a constant temperature, the measurement is the amount of power (microcalories per second) supplied to the reference or sample chambers in order to maintain a constant temperature. In the case of an exothermic reaction more power would be needed for the reference chamber, and for an endothermic reaction more power would need to be supplied to the sample chamber to maintain a constant temperature.

Figure 1.5 shows a typical ITC curve of an exothermic reaction. The area under each peak corresponds indirectly to the amount of heat change (ΔH) occurring as a result of each injection, with a return to baseline occurring in between each injection. As progressively more ligand is added, the free protein concentration decreases resulting in progressively smaller peak magnitude and eventual saturation. An ideal curve is sigmoidal. The inflection point gives the stoichiometry (N) of binding, and the binding constant (K) is most often determined from a "single site binding constant" model.[96] Thus ITC directly gives N, K and ΔH, from which ΔG and ΔS can be calculated given Equations 5 and 6.

$$(5)\ \Delta G = -RT \ln K$$

$$(6)\ \Delta G = \Delta H - T\Delta S$$

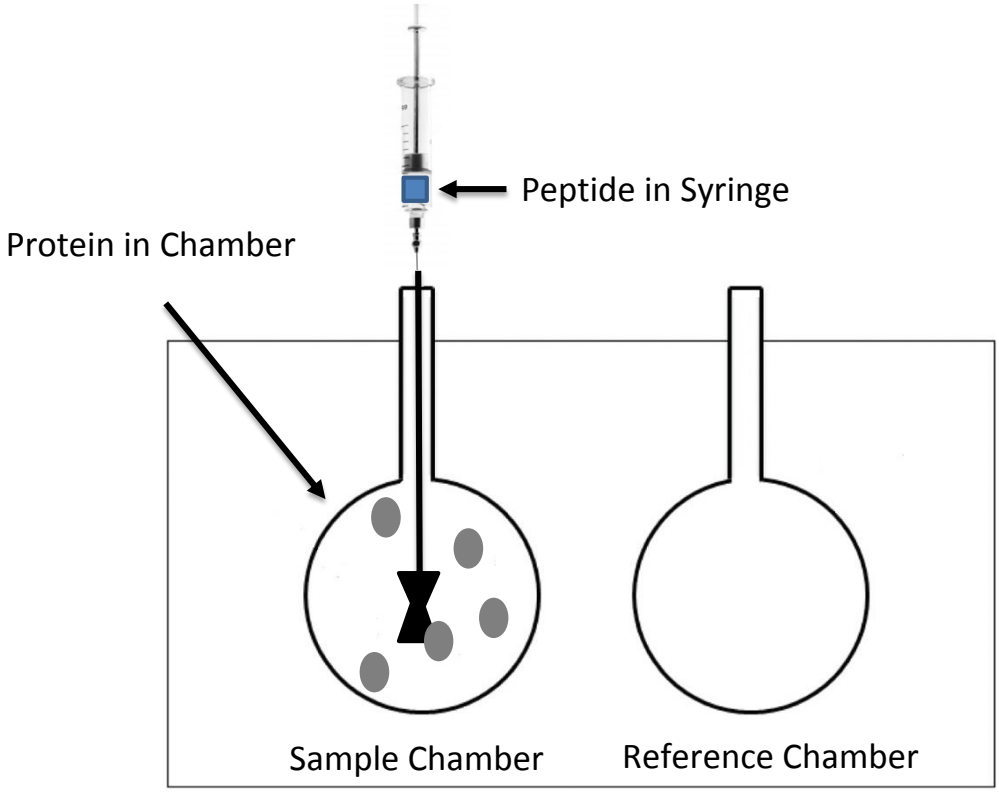


Figure 1.4 Setup of an Isothermal Titration Instrumentation. Consisting of two chambers kept at the same temperature, the reference chamber is filled with buffer or water. As peptide (or other ligand) is titrated into the protein in the cell, the current change needed to keep the chambers at a constant concentration is measured. As more peptide is added and the interaction saturates, a binding isotherm is obtained, as seen in Figure 1.5.

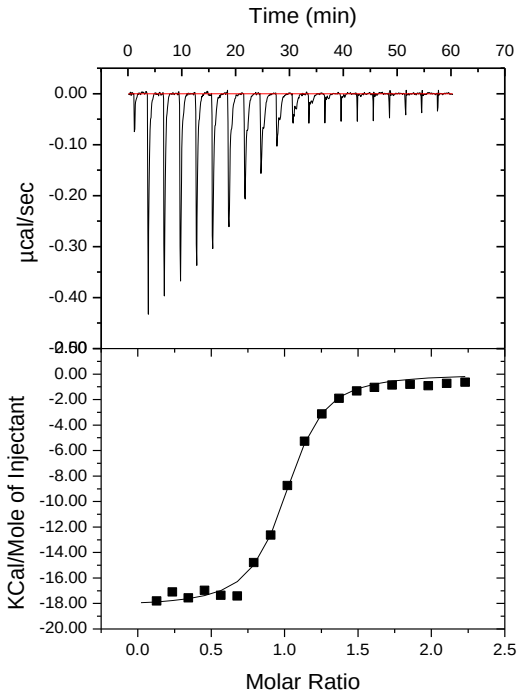


Figure 1.5 An Example ITC Isotherm. As aliquots of peptide are added over time the sample chamber containing protein, the heat released for the binding events is indirectly measured in µcal/sec. The "heat release" dissipates as binding sites are exhausted, and a binding curve is observed.

ITC is the only technique that gives the magnitude of the two thermodynamic values of ΔH and ΔS. This can be a very powerful tool in guiding drug design because thermodynamic properties are important for the elucidation of binding mechanisms. Despite this strength, ITC has its limitations. The biggest of these is governed by the constant c, which is expressed in equation 7,

$$(7)\ c = n[M]_T K$$

where n is the number of binding sites, $[M]_T$ is the total macromolecule (usually protein) concentration), and K is the equilibrium binding constant.[18] Sigmoidal curves generally are observed when c is between 10 and 100.[97] A c value that is too low will have a flat curve that gives little information about binding affinity or stoichiometry. However, a c value that is too high creates a steep, sigmoidal curve that cannot be accurately interpreted. Given that n and K are inherent to the binding interaction under study, the only means of manipulating the c value is to change the concentration of protein in the chamber. Protein solubility and aggregation typically limit the ability to compensate for low c values with high protein concentrations, while the sensitivity of ITC limits the ability to use a low protein concentration to compensate for high c values. If a fixed protein concentration of 20 μM is assumed in a reaction with 1:1 stoichimetry, ITC is typically limited to measurements of K_ds in the range of 2 μM to 200 nM.

ITC is the gold standard of binding affinity measurement because it allows for tag-free analysis of molecules and study of interactions in thermodynamic detail. This makes it a very powerful technique because a single experiment can determine all the major thermodynamic constants of an interaction; however, the high level of calorimeter needed to measure the small

heats of binding during a titration is quite expensive. After the initial purchase price, the cleaning and maintenance of these instruments is quite intensive, and can lead to inaccurate data if not maintained properly. Additionally even the new microcalorimeters require a fairly large amount of material, especially when compared to fluorescence polarization experiments that can be adapted to 384 well plates. When combined with an approximately 2 hour experiment time and the need to run a blank for every sample, it becomes clear to see that this technique is not adaptable to high-throughput screening.

1.5.4 Fluorescence Polarization (FP)

Also known as Fluorescence Anisotropy, this technique takes advantage of differing tumbling rates of molecules of varying mass in solution. In 1920 F. Weigert discovered that excitation of fluorescent dyes with polarized light resulted in emission of polarized light.[98] Additionally he observed that the degree of polarization observed in the emitted light was inversely proportional to the size of the dye measured. This is to say that smaller molecules that are tumbling faster in solution "scramble" the light so the emitted light is less polarized, and that larger molecules that tumble slower maintain more polarity in the emitted spectrum (see Figure 1.6). As complexes form, naturally the mass increases, which makes FP a suitable technique for studying binding interactions. In designing an experiment, ideally the fluorophore would be appended to the smaller of the two molecules under study because this would provide a larger difference in anisotropy upon binding; however, attaching a large fluorophore to the smaller molecule is more likely to alter the interaction of the two molecules under study. Polarization and Anisotropy measurements are described mathematically in equations 8 and 9,

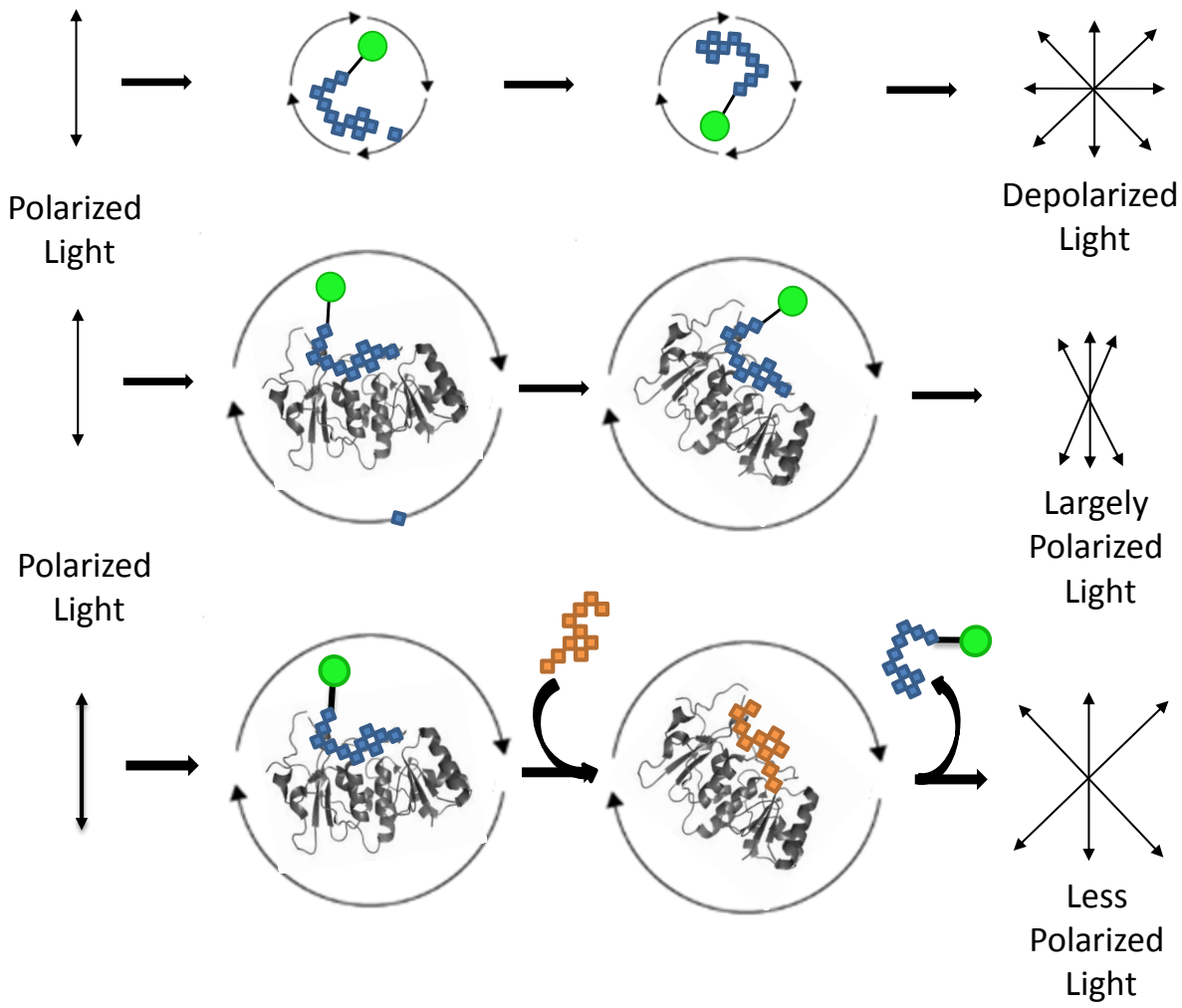


Figure 1.6 Fluorescence Polarization. FP analysis takes advantage of the fact that certain fluorophores emit polarized light when excited with polarized light. When smaller molecules, like peptides, are excited with polarized light their quick tumbling in solution scrambles the emitted polarized light. When such a labeled peptide binds to a protein, the rate of tumbling slows and more polarized light is observed. In a competition assay, a competing peptide (shown in orange) competes off the fluorescently labeled peptide resulting in the observation of less polarized (more scrambled) light emitted.

$$(8)\ \mathrm{P} = \frac{\mathrm{I_{VV}} - \mathrm{G}*\mathrm{I_{VH}}}{\mathrm{I_{VV}} + \mathrm{G}*\mathrm{I_{VH}}}$$

$$(9)\ \mathrm{r} = \frac{\mathrm{I_{VV}} - \mathrm{G}*\mathrm{I_{VH}}}{\mathrm{I_{VV}} + 2*\mathrm{G}*\mathrm{I_{VH}}} = \frac{2\mathrm{P}}{3-\mathrm{P}}$$

where P is polarization, r is anisotropy, and the intensity of fluorescent readings is represented as I with subscripts of V and H, where the first letter indicates vertical or horizontal polarization of the excitation wavelength, and the second indicates the polarization of the emission lens. The G factor is a corrective value specific to each instrument and fluorophore, and is a measurement of a sample containing only the fluorescent molecule with horizontally polarized excitation (See equation 10).

$$(10)\ \mathrm{G} = \frac{\mathrm{I_{HV}}}{\mathrm{I_{HH}}}$$

A typical FP experiment used for binding affinity measurement uses multiple samples at identical fluorphore concentrations, where each sample has a different known concentration of the biological molecule it is binding.

Fluorescence polarization, on the other hand, requires less specialized equipment since fluorescent readings are used in many types of research, and most commonly requires only an additional purchase of polarizing lenses. Although adaptable to multi-well plate format for high-throughput screening this usually requires more specialized instrumentation. If a multi-well plate reader is available, FP can use less than 50 uL samples, so its protein requirement can be fairly minimal. However low affinity interactions are not ideally suited for FP analysis, highly concentrated samples are needed leading to interference with anisotropy measurements, which can lead to sample scattering background through either protein aggregation of simply increased solution viscosity. This can be overcome with IC_{50} competition experiments, but this is not

always possible unless another tighter binding molecule is known. If this is the case, a label free technique such as SPR, ITC or a technique using radio-labels would be more appropriate.

1.5.5 Surface Plasmon Resonance (SPR)

As mentioned previously, SPR is one of the only techniques that continuously measures the kinetics of a binding event in real time. SPR is an optical technique that relies on activation of surface plasmons at the interface of a liquid and metal (usually gold). In SPR, the experiment is conducted with a metal chip where one side of the metal interacts with the optical system, and the other is a coated surface interacting with the liquid sample. The coating is usually carboxylated dextran, but other surfaces can be used to minimize surface effects due to non-specific interaction of the analyte with the coating. One binding partner is tethered to the dextran either directly by standard amine, thiol or aldehyde coupling or indirectly through interaction with an antibody coupled to the dextran surface. Light is reflected at a specific angle and wavelength from the side of the metal not interacting with the liquid. The refractive index is a function of mass bound to the surface, and is sensitive to changes in the mass due to binding of analyte, so the refractive angle of the incident light changes as more mass is bound to the chips surface. To perform a SPR experiment, as seen in Figure 1.7, an analyte is continuously flowed over the chip surface, and the changes in refractive index is recorded as a graph of response units (RU) versus time, where the RU reading is proportional to mass per surface area.[99] By repeating this experiment at various concentrations, the combined data can be fitted to various kinetic models to calculate the rate-constants which can then be used to calculate the K_d.[100] In the case

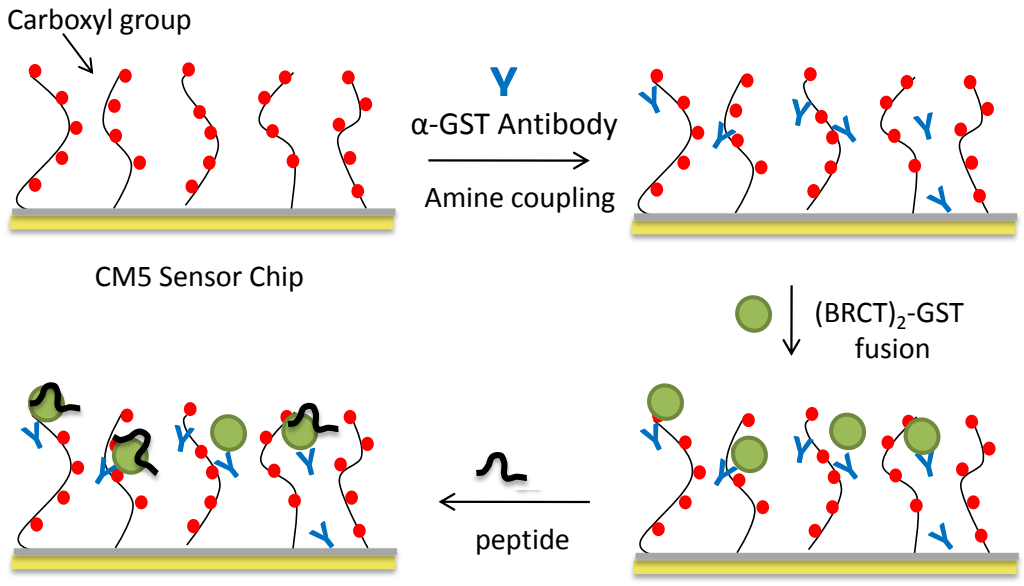


Figure 1.7 Surface Plasmon Resonance Experiment with α-GST Antibody. SPR experiments are based on a mass difference on the surface of a gold sensor chip. As binding occurs and the mass changes on the surface, the angle of reflected light from the bottom of the chip changes, and is observed. The most common chips are coated in dextran and therefore have carboxylic acids available for amine coupling of proteins to the surface. The protein of interest can be directly coupled to the chip surface, or secondarily via an antibody as shown here. Once the antibody is couple, the protein of interest immobilized on the surface via interaction with the antibody. Peptide (or other ligand) is flowed over the chip surface and the response is measure. On and off rates of binding are determined and used to find the binding affinity.

of incredibly fast binding kinetics when the on and off rates cannot be accurately determined, analysis of the binding at equilibrium at various concentrations is also possible.

Although technically SPR is a "label free" technique, one component of the binding pair, ideally the smaller of the two, is required to be attached to the surface of a chip. This can be problematic, because to optimize sensitivity the smaller molecule should be attached to the chip, and as with attaching a fluorescent tag, alteration of a smaller molecule has the potential to dramatically change the way it binds. Proteins are much easier to attach to the surface without altering the interacting interface; however, the additional mass increase from the binding of a small ligand is often insufficiently sensitive to acquire meaningful data. Because this format is measuring binding at a liquid/solid interface and not solution phase binding, the experimental results can complicated by non-specific surface effects that may be observed. In this case trying different chip surfaces may be required. In the end, SPR is a great technique for direct observation of an interaction's off and on rates, but it can be costly due to the necessary specialized equipment and maintenance, and is not amenable to a high-throughput format.

1.5.6 Summary of Current Techniques

The techniques discussed above are only some of the many that have been developed to study the interactions of biomolecules. One of the reasons that so many techniques have been developed is that there are many advantages and disadvantages to each technique. Each of these techniques is capable of measuring a K_d, but no single technique can produce all the information that is necessary to fully understand a binding interaction. A summary of each technique can be found in Table 1.1.

Technique	Advantages	Disadvantages
Equilibrium Ultrafiltration Binding Assay	• No special equipment needed • Only small quantities of ligand needed due to sensitivity of radiolabeling • Label-free technique • Solution phase analysis	• Not as precise as other techniques • Must have access to radiolabeled ligand
Isothermal Titration Calorimetry	• Allows determination of thermodynamic constants • Label-free technique • Solution phase analysis	• Narrow range of ligand affinity is suitable for this technique • Expensive instrumentation and maintenance
Fluorescence Polarization	• Adaptable to high throughput screening (with appropriate instrumentation • Solution phase analysis • Can be adapted to small volumes	• Can be difficult to analyze low affinity ligands • Requires addition of fluorescent tags, which can interfere with ligand binding • Special instrumentation needed for high-throughput analysis
Surface Plasmon Resonance	• Monitors on and off rates of interaction • Label-free technique	• Surface effects from monitoring reaction on a chip • Not amenable to high throughput screening • Expensive instrumentation and maintenance

Table 1.1 Summary of Binding Affinity Techniques. Listed in this table are key advantages and disadvantages of several techniques to determine binding affinities of a ligand.

CHAPTER 2. DRUG-LIKE PEPTIDES: UNNATURAL AMINO ACIDS IN P.U.R.E. TRANSLATION

E. Railey White, Timothy M. Reed, Zhong Ma, Matthew C.T. Hartman*

This chapter was published in part under the title "Replacing amino acids in translation: Expanding chemical diversity with non-natural variants."

***Ref: Methods,* 2012, ePub 27 March 2012**

Contributions:

ERW and Zhong Ma performed the translation experiments

2.1 Introduction

The standard set of amino acids has a wide variety of side chain redundancy. For example, Glu and Asp, Gln, and Asn, and many of the hydrophobic amino acids (Leu, Ile) have similar side chains. Therefore substitution of some of these with unnatural amino acids will expand the diversity of functional groups found in peptides. In the setting of a library selection, in addition to library size, this expanded group of available functional groups will also increase the diversity found in the library potentially leading to selection of higher affinity peptides as a result. Using the PURE translation system, substitution of natural with UNAAs is the simplest way to generate a large library containing many UNAAs.

The basis of this method is leaving out one or more natural amino acids from the translation reaction and replacing them with corresponding UNAA analogs that can be charged by the natural aminoacyl-tRNA synthetases (AARS) onto the tRNA corresponding to the absent natural amino acids (Figure 2.1). A previous assay to discover unnatural amino acids that are substrates for AARS found over 92 UNAAs that can be successfully charged by simple substitution in the PURE translation system.[88b] In this work we chose eight UNAAs (Figure 2.2) from this list for possible incorporation into our translation products, including two different possible tryptophan analogs. These amino acids are all commercially available in their Fmoc protected forms for solid phase peptide synthesis (SPPS) with the exception of canavanine. Additionally, the previous study found these UNAAs to be among the most efficiently incorporated using this substitution method, which makes them good candidates for use together in translation. After selecting the group of UNAAs to incorporate the next step is to test each of them individually to assure results similar to previous findings, and if needed optimize each UNAA individually in translation before combining them.

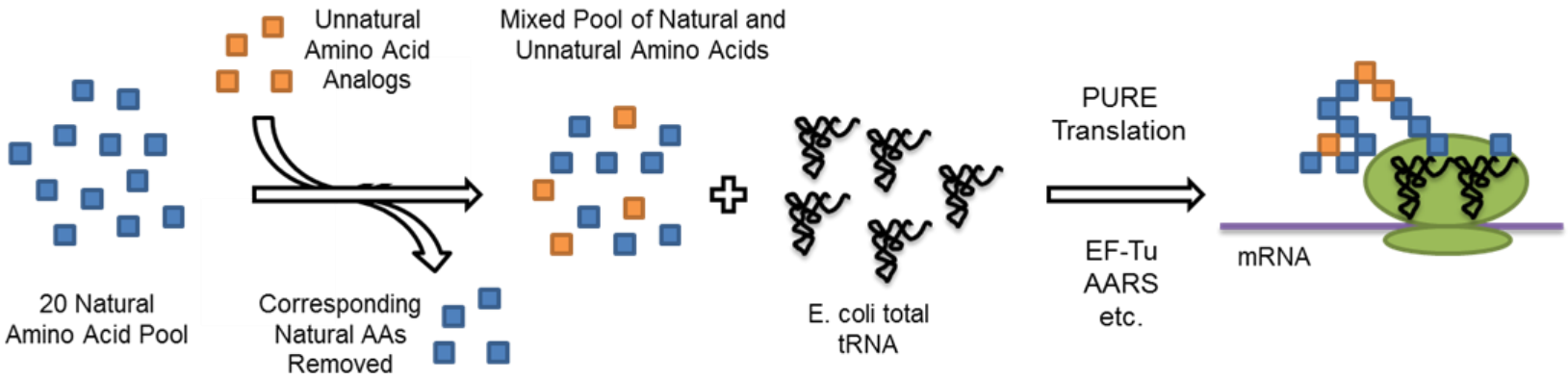
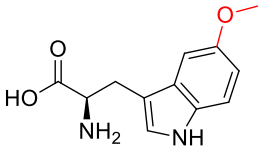
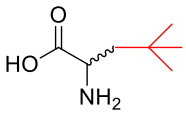


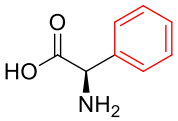
Figure 2.1 Incorporation of UNAA in PURE Translation by Substitution. Using the PURE translation system, natural amino acids can be selectively removed from translation allowing analogous UAAs to be incorporated in their place after being charged by the natural aminoacyl synthetase in the reaction.



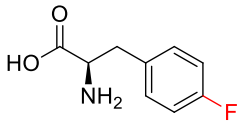
W_3 o-methyl tryptophan



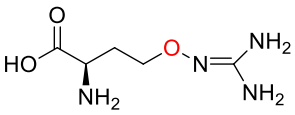
L_3 β-t-butyl alanine



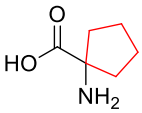
I_2 phenylglycine



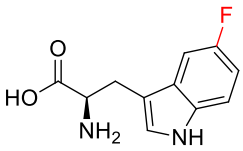
F_{26} 4-fluoro phenylalanine



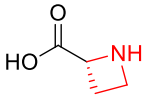
R_1 L-canavanine



V_3 1-amino cyclopentanoic acid



W_2 4-fluoro-tryptophan



P_1 L-azetidine-2-carboxylic acid

Figure 2.2 UNAA Analogs. These eight UNAAs were chosen to test for their suitability for use together in translation via substitution with their natural congeners. The letter in the abbreviation shown in bold indicates the amino acid for which each analog will be substituted.

2.2 Testing Individual Unnatural Amino Acids in PURE Translation

Each of the eight UNAAs chosen was added to a translation reaction using one of the four mRNA test templates listed in Table 2.1. Each translation was conducted on a 50 µL scale in duplicate with one reaction analyzed by MALDI-TOF-MS to monitor fidelity of translation and the other sample containing ^{35}S-methionine to monitor yield by scintillation counting. Each template encodes for a C-terminal FLAG and His_6 epitope tag that allows for purification of the peptides from translation. The C-terminal placement allows only for capture of full length peptides, so any truncations caused by inefficient incorporation of UNAAs are not observed. After capture of the peptides with either Ni-NTA or anti-FLAG antibody agarose, the translated peptides can be eluted and characterized via MALDI. The results of these initial translations are shown in Figure 2.3a and b. In each case of Figure 2.3 a MALDI analysis of a translation with all-natural amino reaction with a particular template is shown on the left, and a MALDI analysis of a translation reaction with an UNAA is shown on the right. Each of these is the initial test of this particular amino acid. The expected mass, and mass of the primarily observed peak is shown. The translation test of amino acids L_3, F_{26}, R_1 show high fidelity and yield; however, for the other three amino acids some optimization was required.

Troubleshooting yield and fidelity.

When UNAAs are not efficiently incorporated into a peptide during translation, the two types of errors that occur can generally be classified as either truncations or misincorporations. Truncation occurs when the UNAA is not an efficient translation substrate, and there are no competing AAs or AA-tRNAs. Truncations can only be directly visualized by MS if an N-terminal tag is used; however, in these experiments all mRNA templates encoded a C-terminal

Coding Region	Epitope Tag
MHFSW	DYKDDDDK
MTINR	DYKDDDDK
NLEPQ	DYKDDDDK
MVHM	HHHHHHM

Table 2.1 mRNA Test Templates. List of peptide-encoding sequences used to analyze analog translation. The full length mRNA sequences can be found in Ref 74.

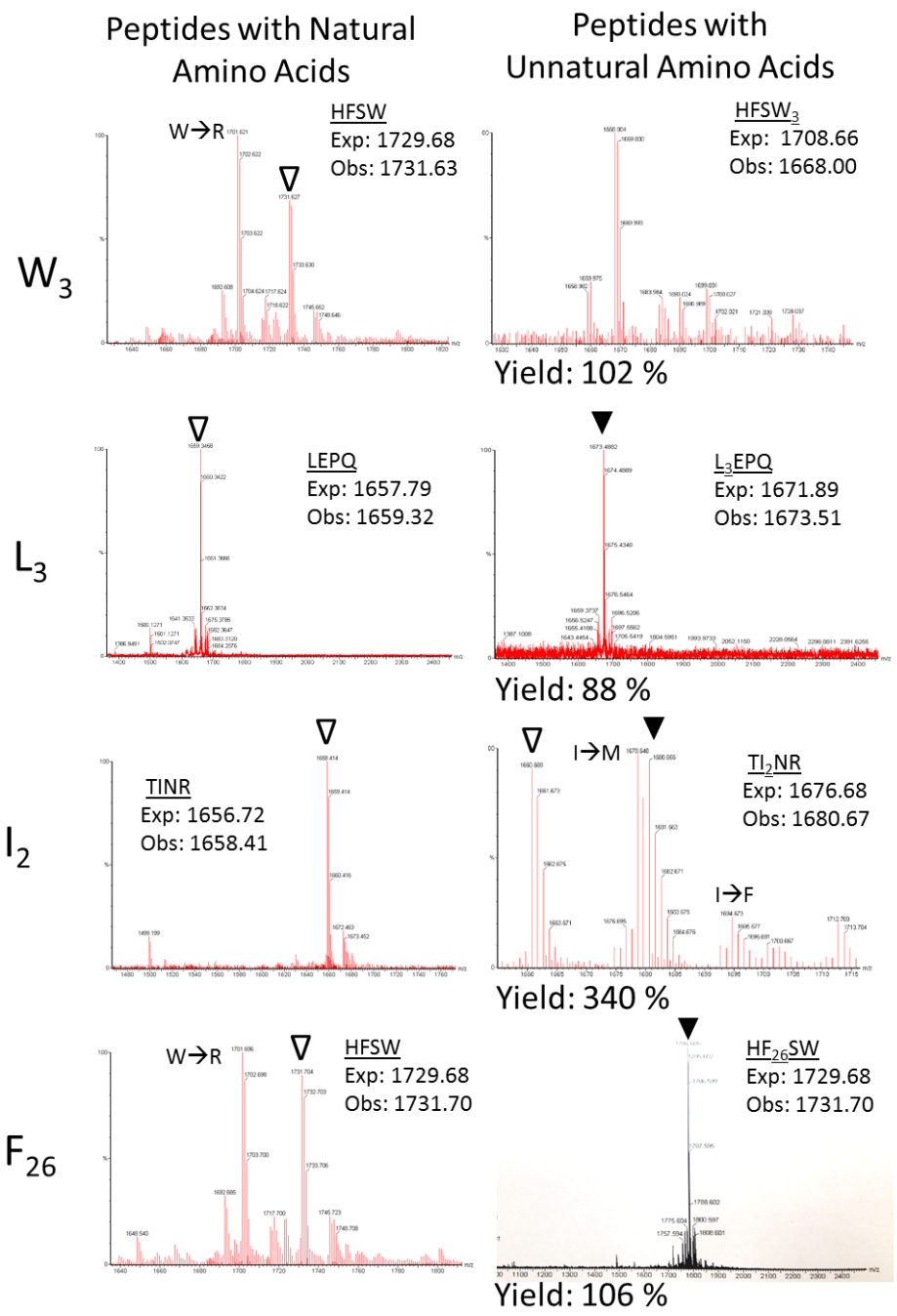


Figure 2.3a Initial MALDI Analysis of UNAA in Translation. The four amino acid analogues listed on the left were tested for their ability to incorporation into a translated test peptide via substitution for their corresponding natural amino acid, and analyzed by MALDI. The all-natural peptides are indicated by white triangle, and the UNAA peptides are indicated by a black triangle. The yields of the UNAA reaction compared to the all-natural is indicated. The expected and observed mass is shown on each MALDI spectrum.

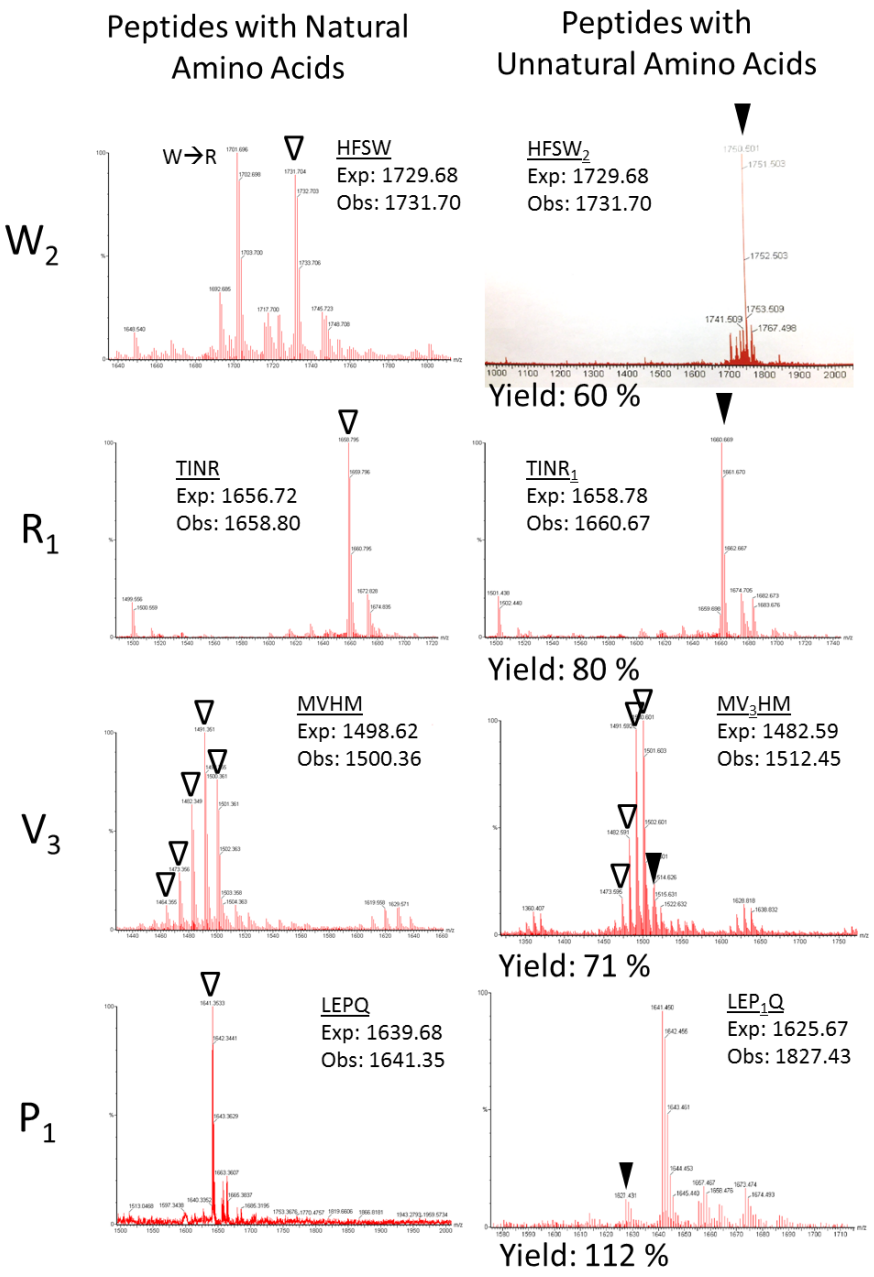


Figure 2.3b Initial MALDI Analysis of UNAA in Translation Continued. The four amino acid analogues listed on the left were tested for their ability to incorporation into a translated test peptide via substitution for their corresponding natural amino acid, and analyzed by MALDI. The all-natural peptides are indicated by white triangle, and the UNAA peptides are indicated by a black triangle. The yields of the UNAA reaction compared to the all-natural is indicated. The expected and observed mass is shown on each MALDI spectrum.

tag. A low translation yield with C-terminally tagged peptide is also typically indicative of premature truncation.

MALDI-TOF MS analysis provides information about the fidelity of the translation and provides evidence of any misincorporation of other amino acids in place of the desired UNAA. If the anticipated peptide mass with the UAA is not observed, use of a misincorporation mass table (Figure 2.4) can be useful to quickly determine the identity of the misincorporated amino acid. Misincoporations typically are observed when either (1) a near-cognate AA-tRNA is able to compete effectively at the ribosomal A-site with the UNAA-tRNA containing a cognate codon, or (2) when residual natural amino acid competes with the unnatural variant for the AARS.

Increasing the concentration of the UAA is typically the first strategy we use to improve fidelity and efficiency. Limiting factors to this strategy include the solubility of the UAA, as well as the possibility that very high concentrations may lead to competition with other AARS (if the UAA is charged onto tRNA by two different AARS). If increasing the concentration does not dramatically improve yield, we also typically try to increase the concentration of the appropriate AARS to enhance the rate of formation of the UAA-tRNA.

If the UAA is a relatively poor substrate for an AARS, even trace amounts of the contaminating natural AA can be a problem. There are several potential sources of contamination. A common source is the amino acids themselves. For example, we have found that Gln is contaminated with Glu and Asn is contaminated with Asp. It is often helpful in these situations to lower the concentration of the natural amino acid that contains the contaminant amino acid. Another contaminant source is the *E. coli* total tRNA which does contain aminoacylated-tRNA species. These residual AAtRNAs can be removed by deacylation at pH

	75	89	105	115	117	119	121	131	131	132	133	146	146	147	149	155	165	174	181	204
	G	A	S	P	V	T	C	I	L	N	D	K	Q	E	M	H	F	R	Y	W
75 G	0	-14	-30	-40	-42	-44	-46	-56	-56	-57	-58	-71	-71	-72	-74	-80	-90	-99	-106	-129
89 A	14	0	-16	-26	-28	-30	-32	-42	-42	-43	-44	-57	-57	-58	-60	-66	-76	-85	-92	-115
105 S	30	16	0	-10	-12	-14	-16	-26	-26	-27	-28	-41	-41	-42	-44	-50	-60	-69	-76	-99
115 P	40	26	10	0	-2	-4	-6	-16	-16	-17	-18	-31	-31	-32	-34	-40	-50	-59	-66	-89
117 V	42	28	12	2	0	-2	-4	-14	-14	-15	-16	-29	-29	-30	-32	-38	-48	-57	-64	-87
119 T	44	30	14	4	2	0	-2	-12	-12	-13	-14	-27	-27	-28	-30	-36	-46	-55	-62	-85
121 C	46	32	16	6	4	2	0	-10	-10	-11	-12	-25	-25	-26	-28	-34	-44	-53	-60	-83
131 I	56	42	26	16	14	12	10	0	0	-1	-2	-15	-15	-16	-18	-24	-34	-43	-50	-73
131 L	56	42	26	16	14	12	10	0	0	-1	-2	-15	-15	-16	-18	-24	-34	-43	-50	-73
132 N	57	43	27	17	15	13	11	1	1	0	-1	-14	-14	-15	-17	-23	-33	-42	-49	-72
133 D	58	44	28	18	16	14	12	2	2	1	0	-13	-13	-14	-16	-22	-32	-41	-48	-71
146 K	71	57	41	31	29	27	25	15	15	14	13	0	0	-1	-3	-9	-19	-28	-35	-58
146 Q	71	57	41	31	29	27	25	15	15	14	13	0	0	-1	-3	-9	-19	-28	-35	-58
147 E	72	58	42	32	30	28	26	16	16	15	14	1	1	0	-2	-8	-18	-27	-34	-57
149 M	74	60	44	34	32	30	28	18	18	17	16	3	3	2	0	-6	-16	-25	-32	-55
155 H	80	66	50	40	38	36	34	24	24	23	22	9	9	8	6	0	-10	-19	-26	-49
165 F	90	76	60	50	48	46	44	34	34	33	32	19	19	18	16	10	0	-9	-16	-39
174 R	99	85	69	59	57	55	53	43	43	42	41	28	28	27	25	19	9	0	-7	-30
181 Y	106	92	76	66	64	62	60	50	50	49	48	35	35	34	32	26	16	7	0	-23
204 W	129	115	99	89	87	85	83	73	73	72	71	58	58	57	55	49	39	30	23	0

Figure 2.4 Misincorporation Table. The values in the middle of the table correspond to changes in mass when changing from the amino acid on the top row to an amino acid on the left-side column.

8.8 followed by dialysis. Examples of how these strategies can be applied are shown below for the optimization of P1, V3, and W3

Optimization of P_1

Significant enhancement of P_1 incorporation was seen upon doubling the analog concentration from 200 µM to 400 µM (Figure 2.5). Supplementation of proline aminoacyl-tRNA Synthetase (PRS) was attempted to enhance yield of incorporation, but without success.

Optimization of V_3

One UNAA that we have shown can be incorporated with reasonable efficiency into peptides via translation is 1-aminocyclopentanoic acid (V3). To test its incorporation we used the template encoding $MVHMH_6M$. Monitoring incorporation by MALDI-TOF in templates containing a hexahistidine tag can be complicated by glutamine misincorporation. We presume that this misincorporation arises because the high proportion of His codons and low overall abundance of $tRNA^{His}$ leads to depletion of His-$tRNA^{His}$ during translation. The misincorporation results in serial -9 peaks in the MALDI spectrum corresponding to the number of glutamine misincorporations.

The initial translation experiment showed only a tiny amount of V_3 incorporation; the majority of the peptides contained valine. Because no valine is added to the translation reaction, the issue with V_3 incorporation is the presence of competing valine as a contaminant. Using deacylated tRNA significantly improved incorporation efficiency, suggesting that Val-$tRNA^{Val}$ present in the commercial tRNA mix was the culprit. To deacylate the total tRNA, it is treated at pH 8.8 followed by dialysis to remove any amino acids that had been deacylated. This deacylated tRNA led to a dramatic increase in the V3 peptide fidelity (Figure 2.6). Using this

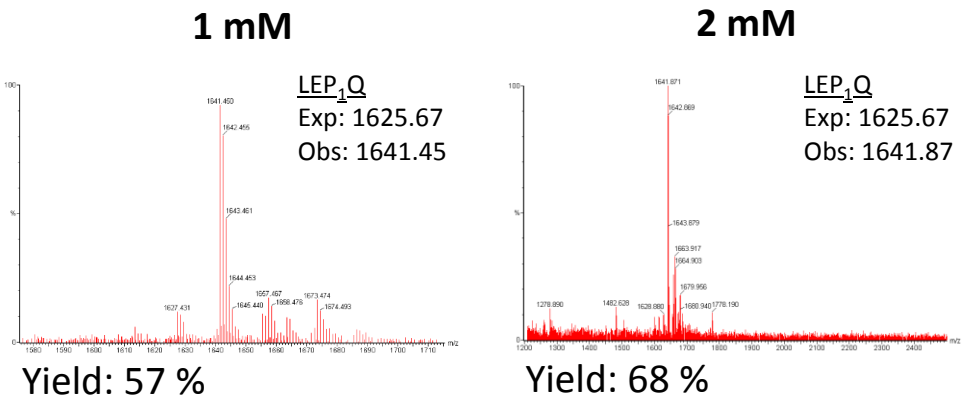


Figure 2.5 Titration of P_1 Analogue in Translation. In attempt to increase the fidelity of incorporation of the P1 analogue, its concentration was increased in the translation reaction. The expected and observed masses are shown on each MALDI spectrum, and the yield compared to an all-natural amino acid translation is shown below.

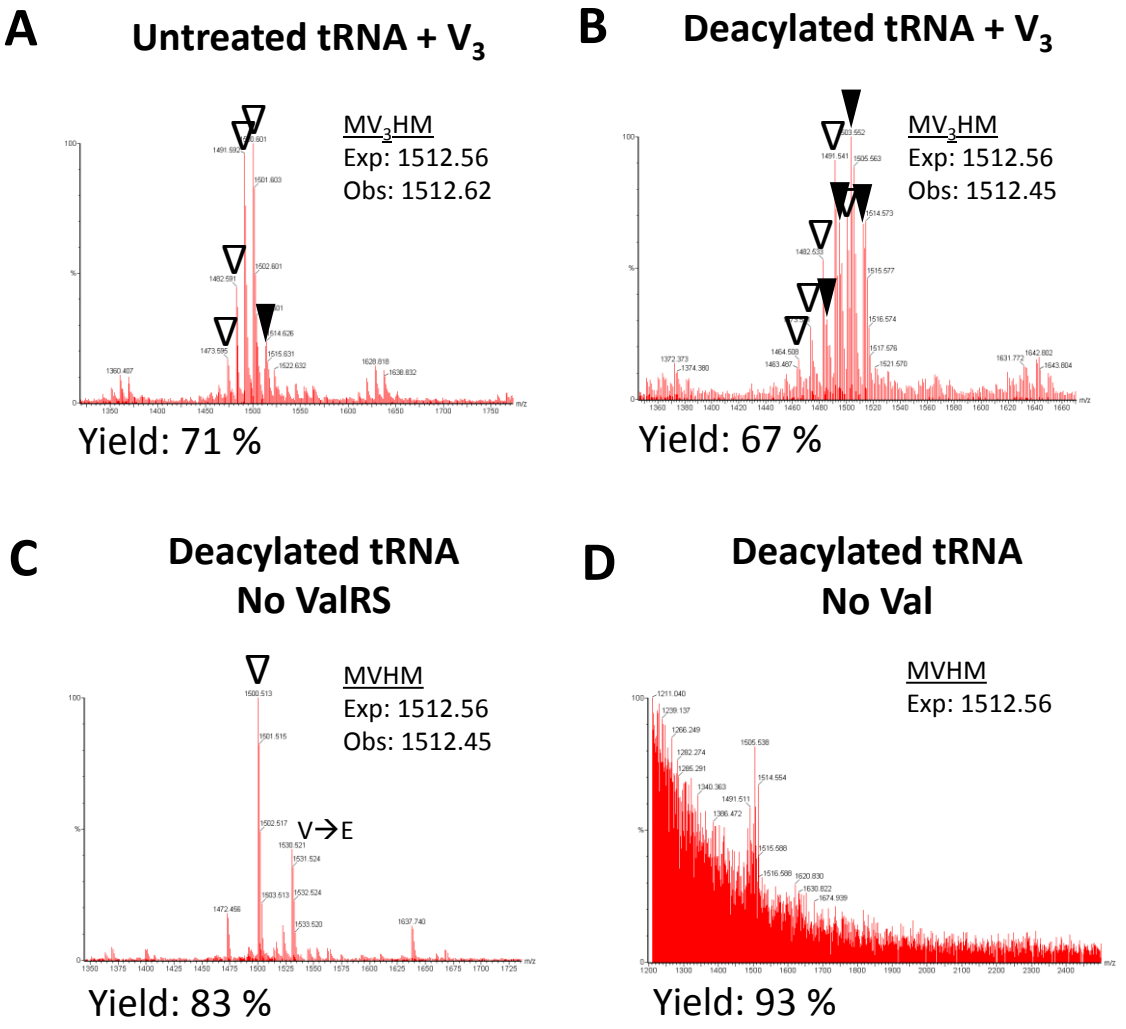


Figure 2.6 V_3 Incorporation with Deacylated tRNA. Fidelitiy of 1-amino-cylcopentanoic acid (V_3) incorporation in place of valine in a template encoding MVHMH$_6$M. A-B) Comparison of the effect on translation fidelity of V_3 incorporation with untreated and deacylated tRNA. C-D) Translation reactions were also conducted in the absence of either valine amino-acyl tRNA Synthetase (ValRS) or without valine. Yields of each reaction compared to an all-natural amino acid translation with all components are indicated. V3 incorporation is indicated by black arrows and valine with white arrows.

deacylated tRNA, we varied the concentration of V_3 and showed that 6.4 mM was the best concentration to use in translation (Figure 2.7). Use of deacylated tRNA did not lower the yield significantly of any of the peptide templates (Figure 2.8).

Optimization of W_3

A translation experiment was conducted along with the F_{26} and R_1 analogs that are also found in the HSFW template. Knowing that these two analogs incorporate very well and will likely be used in the final selection experiments W_3 is of little use to us if it cannot be incorporated along with the other analogs. In another experiment, the concentration of W_3 was significantly increased to 1.6 mM based on the available volume in the translation reaction, but despite this increase the expected mass was not observed at all (Figure 2.9). When a simple concentration increase failed to produce any W_3 incorporation, additional TrpRS was added to the translation, but this modification was also unsuccessful (Figure 2.10).

Optimization of I_2

The primary issue seen with the incorporation of the I_2 analog is the significant Ile incorporation that is seen in addition to the analog. Since no Ile was added to the translation reaction, it had to be present as a contaminant in one of the reaction components. When increasing the I_2 concentration up to 3.2 mM showed no improvement in out-competing the natural Ile for incorporation (Figure 2.11). Deacylated tRNA was tested as one means of contaminant removal (Figure 2.12). The source of Ile contamination does not appear to be from Ile-$tRNA^{Ile}$ because despite using treated tRNA, the misincorporations have the same pattern as untreated tRNA. Likely the Ile is a contaminant in another reaction component, perhaps a natural amino acid and is present in sufficient quantities to out compete I_2 for the most part. This type of

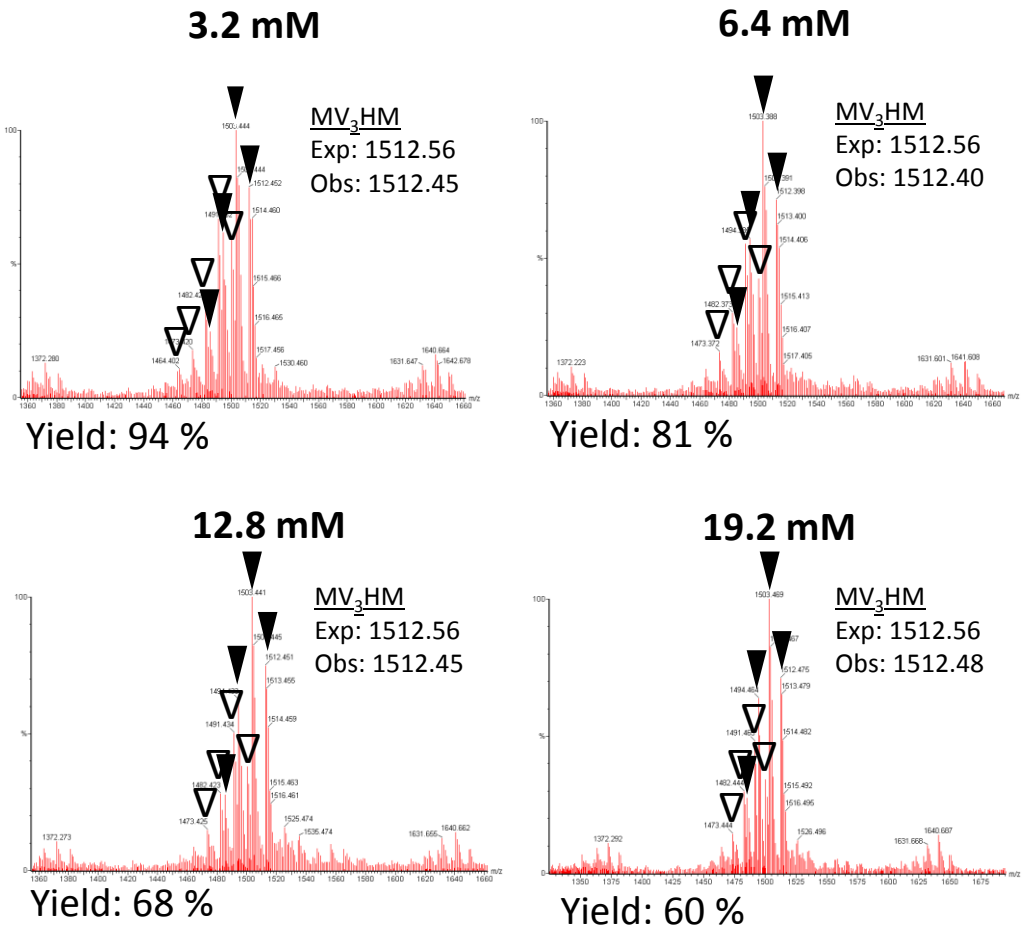


Figure 2.7 Titration of V_3 Concentration with Deacylated tRNA. In attempt to increase fidelity of V3 incorporation, translation reactions were conducted with tRNA as well as increasing V3 concentrations as indicated. The expected and observed masses are indicated, as well as the reaction yields. V3 incorporation is indicated by the black triangles and valine with white.

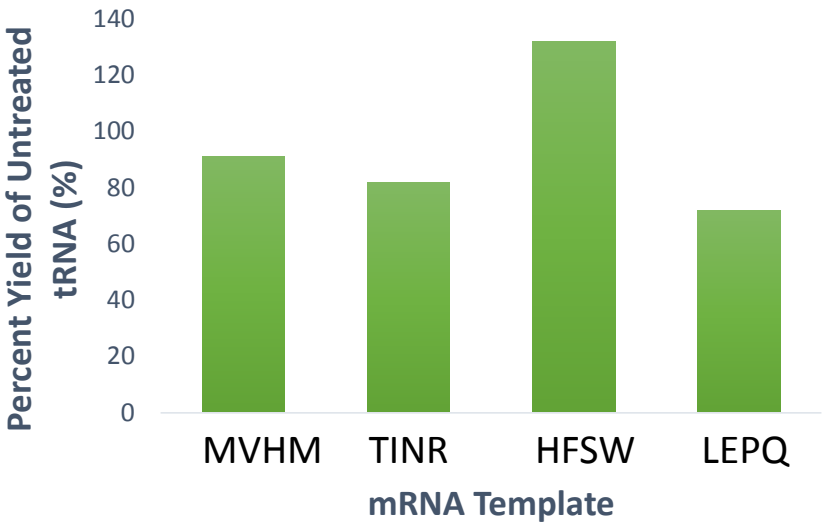


Figure 2.8 Effect of Deacylated tRNA on Peptide Yield. In order to asses a general effect of deacylated tRNA on peptide yield, translations were conducted with the test peptides. The graph shows the percent yield compared to translations with untreated tRNA.

1.6 mM

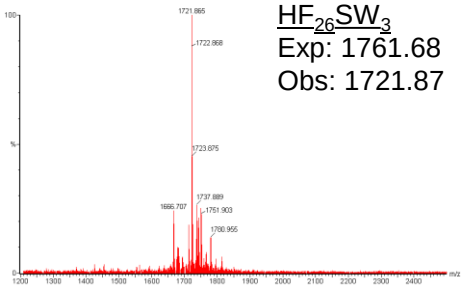


Figure 2.9 Increase in Concentration to Improve W_3 Fidelity. A 50 µL translation with an increased concentration of W3 in attempt to increase its incorporation. The peptide product was not observed.

12 µM TrpRS

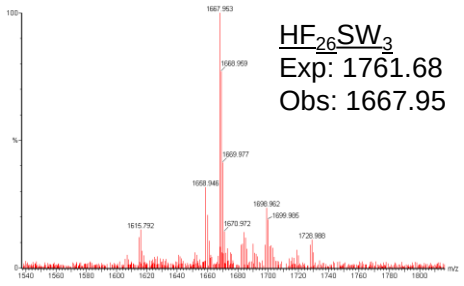


Figure 2.10 Increase in to Improve W_3 Fidelity. A 50 µL translation with an increased concentration of Tryptophan amino-acyl synthetase (TrpRS) in final attempt to observe its incorporation. The peptide product was unable to be identified.

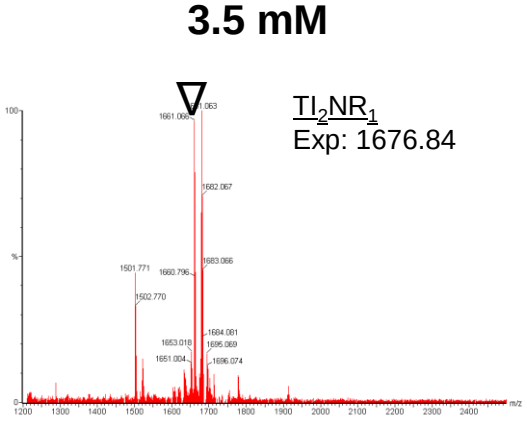


Figure 2.11 Incorporation of Analogue I_2 with a High Amino Acid Concentration. Despite the high concentration of the I2 analogue in the translation reaction, no peak with the expected mass was observed. However a peak corresponding to the all-natural peptide was seen and is marked with a triangle.

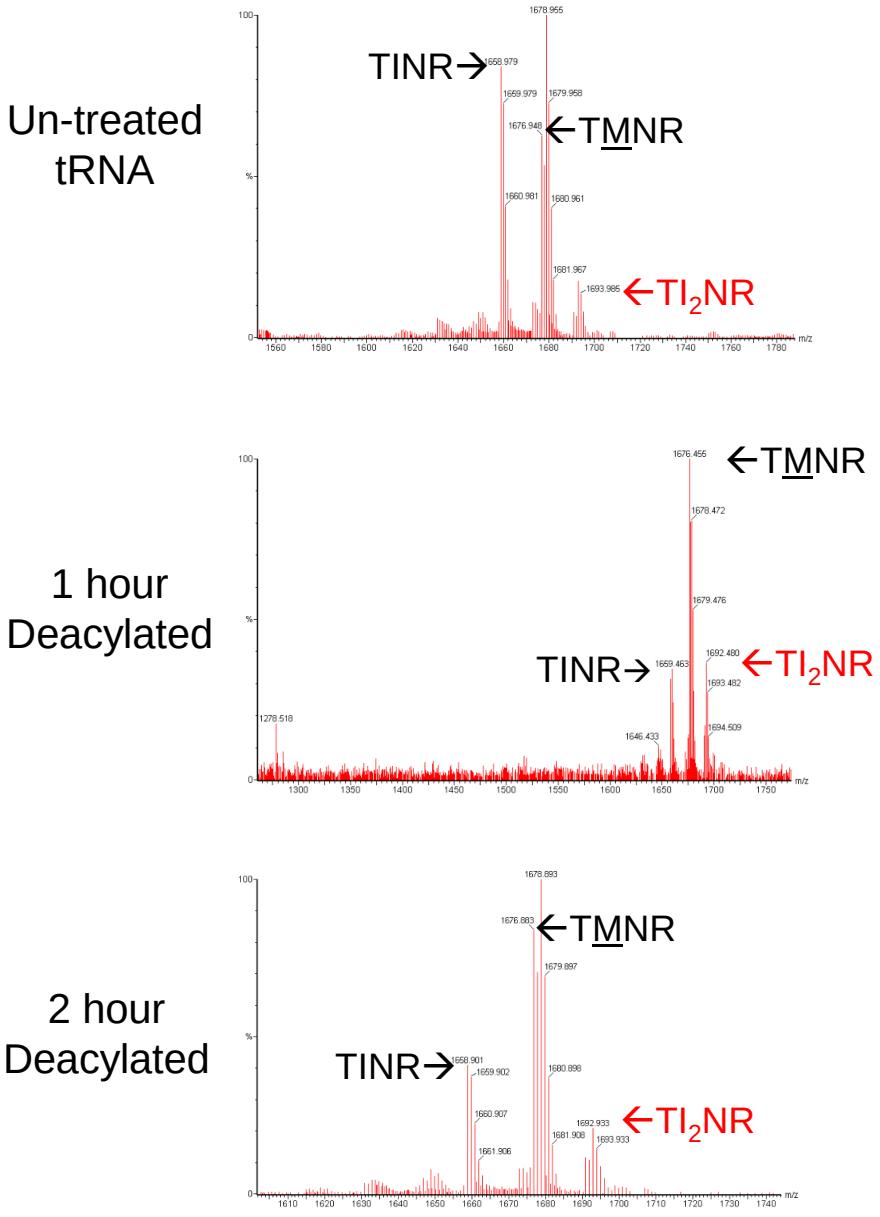


Figure 2.12 Analogue I_2 Deacylation Tests. To test the effect of tRNA deaclyation on the fideltity of I_2 incorporation, 50 µL translations were conducted with untreated, deacylated tRNA that had been treated with base for either 1 hour or 2 hours. Misincorporations are marked with an underline.

contaminant can be more difficult to remove, and it was decided not to include I_2 in the UNAA mix for selection.

2.3 Combination of UNAAs in PURE Translation

Efficient incorporation of the UAA into a single template does not guarantee that it will be efficiently incorporated into peptides when combined with other analogs. Before progressing to translation with the library templates, it was necessary to test all of the UNAAs with each of the templates. One reaction contained all the FLAG templates, and the other the single His_6 template, but both had all 6 desired UNAAs. The results from these tests (Figure 2.13) showed expected fidelity and yield based upon prior testing of individual amino acids, so testing with library templates was pursued. The final concentration of amino acids used in translation can be found in Table 2.2.

2.4 Testing UNAAs with Library Templates

After successful translation of the six final peptides in the same translation reaction a final assessment of the effect of UNAA incorporation on peptide yield was assessed using the actual template to be used in the selection (Figure 2.13). It is not possible to observe the peptides encoded by the randomized library because in theory they should each have different sequences. Thus fidelity is assumed to be the same as that observed in previous experiments. When the translation reactions with library mRNA templates were supplemented reaction (Figure 2.14).

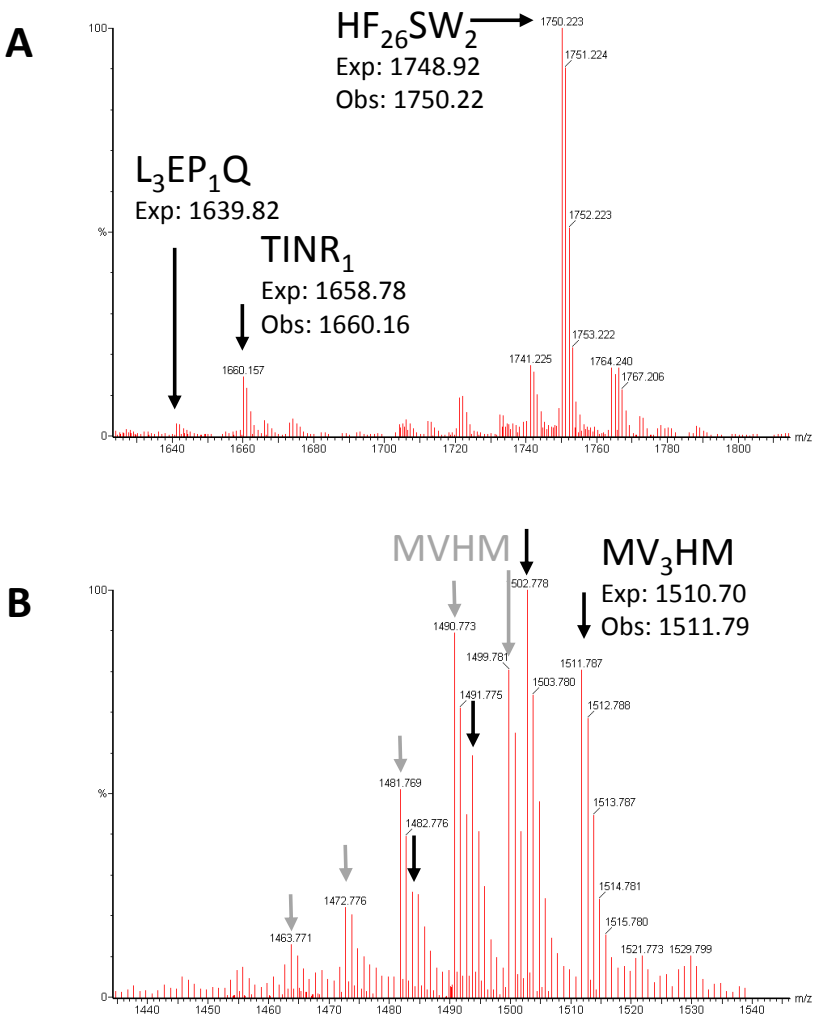


Figure 2.13 Translation with All Six UNAAs. For a final test of translation fidelity with all six UNAAs, two 50 µL translations were conducted with all six UNAAs present. A) The translation reaction containing the three FLAG-tagged templates. Peptide products containing UNAAs are marked with arrows, and the expected and observed masses are written. B) Another translation was conducted with the MVHM His_6-tagged template. V3 incorporation is marked with black arrows, and valine incorporation with gray arrows.

Table 2.2 Final Concentration of Unnatural Amino Acids and PRS. These are the only deviations to the *in vitro* translation reaction conditions described by Ma and Hartman (Ref 101).

Amino Acid/Synthetase	Symbol	Concentration in Translation Reaction
4-fluoro-DL-tryptophan	W_a	0.8 mM
L-canavanine	R_a	0.4 mM
L-azetidine-2-carboxylic acid	P_a	1.0 mM
4-fluoro-phenylalanine	F_a	1.6 mM
β-t-butyl-L-alanine	L_a	6.6 mM
1-aminocyclopentane-1-carboxylic acid	V_a	6.4 mM
Prolyl aminoacyl-tRNA Synthetase	PRS	0.31 µM

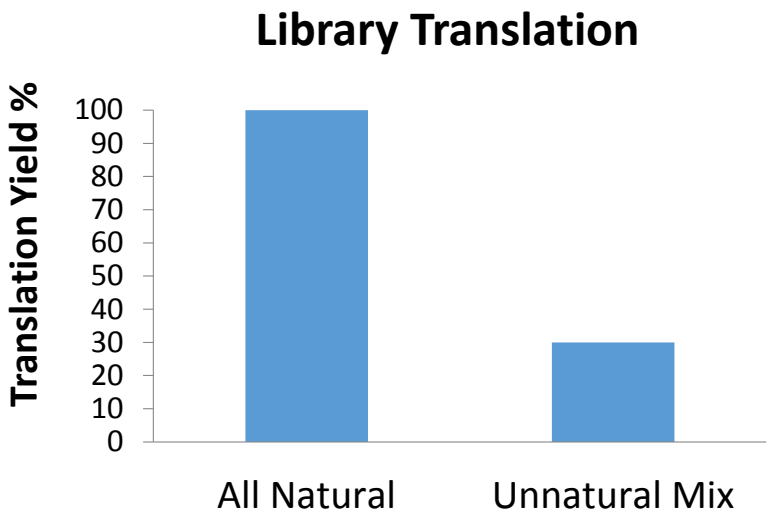


Figure 2.14 Yield of All 6 UNAAs with Library Templates. 250 µL translations were conducted with the CX12 library template that will be used for selection to compare the yield of incorporation of the six UNAAs to an all-natural amino acid translation.

2.5 Discussion

In search of a bioavailable and high affinity peptide, unnatural amino acids were substituted for some of the 20 natural amino acids in translation. Based on our previous work describing amino acid analogs that can be incorporated by substitution in ribosomal translation,[101] a group of six unnatural amino acids were optimized for use together with the other 14 natural amino acids in translation. Translational efficiency was measured by ^{35}S-Met incorporation, and translational fidelity was determined by MALDI-TOF analysis. Amino acid analogs W_2, F_{26}, R_1 and L_3 were incorporated with high efficiency in place of their natural congeners; however, additional modifications were required for incorporation of the W_3, I_2 and V_3 and P_1 analogs. Attempts to increase amino acid concentration and add the corresponding AARS or use deacylated tRNA were unsuccessful in achieving high fidelity incorporation of the W_3 and I_2 analogs. However, treatment of total *E. coli* tRNA with mildly basic conditions to remove residual Val-tRNA^{Val} along with addition of higher V_3 concentrations was sufficient to remove the majority of valine contamination in translated peptides allowing successful incorporation of the V_3 analog.[102] Attempts to improve the yield of P_1 by increasing its concentration in the reaction were successful; however, improving yield by supplementing the translation reaction with additional proline aminoacyl-tRNA Synthetase (PRS) was unsuccessful.

Because W_3 was unable to be successfully optimized and was therefore replaced with the efficiently incorporated W_2 analog, while I_2 was dropped altogether.

2.6 Experimental

General Reagents

Putrescine, spermidine, potassium chloride (KCl), ammonium chloride, magnesium acetate tetrahydrate ($Mg(OAc)_2$), calcium chloride ($CaCl_2$), potassium hydroxide (KOH), nucleoside 5′-diphosphate kinase from bovine liver, D,L-dithiothreitol (DTT), myokinase from rabbit muscle, adenosine 5′-triphosphate disodium salt, guanosine 5′-triphosphate sodium salt hydrate, ANTI-FLAG M2-Agarose from mouse, trifluoroacetic acid spectrophotometric grade and a-cyano-4-hydroxy-cinnamic acid were purchased from Sigma–Aldrich. Potassium acetate (KOAc), water Optima LC/MS grade and acetonitrile Optima LC/MS grade were purchased from Fisher. Potassium phosphate dibasic (K_2HPO_4) was purchased from Caledon. Creatine kinase and Escherichia coli total tRNA were purchased from Roche Applied Science. Creatine phosphate potassium salt was purchased from Merck/EMD. (6R,S)-5,10-formyl-5,6,7,8-tetrahydrofolic acid (methyl tertahydrofolate) was purchased from Schircks Laboratory. All natural L-amino acids were purchased from Fluka in their highest purity form. 35S-Met (Specific Activity: >1000 Ci (37.0 TBq)/mmol) was purchased from Perkin-Elmer. 1-Aminocyclopentanoic acid was purchased from Chem-Impex International. Zip Tip C-18 columns were purchased from Millipore. Ni-NTA agarose was purchased from Qiagen.

Amino acids. Unnatural amino acid sources: W_a: 4-fluoro-DL-tryptophan (Sigma), R_a: L-canavanine (Sigma), P_a: L-azetidine-2-carboxylic acid (ChemImpex), F_a: 4-fluoro-L-phenylalanine (ChemImpex), L_a: beta-*t*-butyl-L-alanine (ChemImpex), and V_a: 1-aminocyclopentane-1-carboxylic acid (ChemImpex). All natural amino acids were purchased from Fluka in their highest purity form. All amino acids were dissolved in H_2O at a concentration of 10 mM or at maximal solubility and KOH was added to a final pH of 7.0-7.5

followed by sterile filtration. Isotopically labeled ^{35}S-Met (1000 Ci/mmol) was from Perkin Elmer.

Translation Factors, Enzymes and Ribosomes. All purified enzymes were stored at -80 °C in enzyme storage buffer (50mM HEPES-KOH pH 7.6, 100mM KCl, 10mM $MgCl_2$, 7mM BME, 30% glycerol) with the exception of MetRS which was kept in enzyme storage buffer with 50% glycerol at -20 °C. Ribosomes were prepared as described.[103]

Instrumentation

MALDI-MS experiments were performed on a Micromass MALDI-R MALDI-TOF Mass Spectrometer.

Preparation of mRNAs for Translation

mRNAs were prepared using T7 in vitro transcription according to described protocols.[88a] The DNA templates were created using two different methods. The first method involved ligating synthetic DNA duplexes with sticky ends into a pET12b vector, followed by PCR using the primers complementary to the vector encoded T7 promoter and terminator. Alternatively, a single synthetic oligonucleotide complementary to the T7 promoter, Epsilon enhancer, Shine-Dalgarno, and coding region was used as a template for runoff T7-mediated in vitro transcription.

Preparation of Amino Acid Stocks for Translation

Each UAA was dissolved to a final concentration of 10 mM, the pH was adjusted to 7.0–7.5 with 1 M KOH, filtered through 0.22 μm syringe filter, and stored at -20 °C

In vitro Translation

Each translation reaction was carried out as previously described.[88a] Each reaction (50 μL) contained putresciene (8 mM), spermidine (1 mM), potassium phosphate (5 mM), potassium

chloride (95 mM), ammonium chloride (5 mM), magnesium acetate (5 mM), calcium chloride (0.5 mM), dithiothreitol (1 mM), inorganic pyrophosphatase (1 µg/mL), creatine kinase (4 µg/mL), nucleotide diphosphate kinase (1.1 µg/mL), (6R,S)-5,10-formyl-5,6,7,8-tetrahydrofolic acid (30 lM), myokinase (3 µg/mL), creatine phosphate (20 mM), ATP (2 mM), GTP (2 mM), *E. coli* total tRNA (2.4 mg/mL), IF-1 (1 µg), IF-2 (2 µg), IF-3 (0.75 µg), EF-G (1 µg), EF-TS (1 µg), EF-Tu (2.24 µg), RF-1 (0.5 µg), RR-F (0.5 µg), RF-3 (0.5 µg), ribosomes (0.5 µM), 35S-Methionine (0.4 µM), methionine (10 µM), 19 AA (200 µM), MetRS (0.1 µM), LeuRS (0.3 µM), GluRS (0.6 µM), ProRS (0.2 µM), GlnRS (1.0 µM), HisRS (1.0 µM), PheRS A294G (2.5 µM), TrpRS (1.5 µM), SerRS (0.2 µM), IleRS, (0.2 µM) ThrRS (0.4 µM), AsnRS (0.6 µM), AspRS (0.6 µM), TyrRS (0.5 µM), LysRS (0.5 µM), ArgRS (0.4 µM), ValRS (0.2 µM), AlaRS (0.2 µM), CysRS (0.5 µM), GlyRS (0.6 µM), MTF (0.2 µM) and mRNA template (1.14 µM).

The translations were initiated by addition of the appropriate mRNAs. For initial testing of UAAs, 19 natural amino acids were included (200 µM each) with only one UAA. After incubation of the translation for 1 h at 37 °C, the reactions were quenched with 150 µL PBS (if using FLAG tag) or 150 µL TBS with 5 mM BME (if using His-tag). Forty microliters of Ni-NTA resin or 10 µL ANTI-FLAG M2 agarose was added to a 500 µL centrifugal filter along with the quenched translation reaction, and the mixture was tumbled at room temperature. After 1 h, the resin was washed three times with 500 µL TBS and eluted with 1% trifluoroacetic acid (TFA) (50 µL). For reactions labeled with 35S-Met, the yield was determined by scintillation counting of half (25 µL) of the elution. To examine the fidelity of the UAA incorporation, the non-radiolabeled reactions were purified and concentrated by Zip-Tip C18 chromatography. The Zip-tips were first wetted with acetonitrile, followed by 1:1 acetonitrile, then with 0.1% TFA. Then the peptide was loaded onto the tip by pipetting up and down 15 times in the peptide

solution. The tip was washed three times with 0.1% TFA, and then eluted with 5 µL CHCA matrix (a-cyano-4-hydroxycinnamic acid in 1:1 MeCN:0.1% TFA). An aliquot (1 µL) of the resulting suspension was spotted on a MALDI plate and analyzed.

tRNA Deacylation

E. coli total tRNA (100 mg/mL) dissolved in 1 M Tris–HCl (pH 8.8) and was incubated at 37 °C for 2 h, followed by dialysis overnight at 37 °C against 50 mM Tris–HCl (pH 8.8). The tRNA was precipitated by first adding 0.1 volume of a solution of KOAc (3.0 M, pH 5.5) and 3 volumes of ethanol. The pellet was washed twice with 70% ethanol, and allowed to air-dry at room temperature. The tRNA was resuspended in ddH_2O and the concentration was adjusted to 100 mg/mL (1.6A260/µL). The tRNA was aliquoted and stored at -80 °C.

2.7 Summary

Of the eight UNAA tested, six were found to work well together in translation. The F_{26}, W_2, L_3 and R_1 analogs were successful in initial testing. P_1 and V_3 required increased amino acid concentration to be successfully incorporated, with V_3 also requiring the use of deacylated tRNA. W_2 and I_2 on the other hand were not able to be incorporated with high fidelity despite modifications to the translation reactions. Fidelity of these UNAA incorporations was maintained when they were combined, and overall yield was 30% of the all-natural yield with library templates.

Future Directions. These 6 UNAAs will be used together in translation for mRNA display selection against the $(BRCT)_2$ domain of BRCA1.

CHAPTER 3. STRENGTH IN NUMBERS: mRNA-DISPLAY LIBRARY SELECTION AGAINST BRCA1-$(BRCT)_2$

Contributions:

Zhong Ma and ERW performed experiments leading to the optimization of selection as well as the actual selection experiments.

Melissa B. Huie, was responsible for cloning of GST-$(BRCT)_2$ fusion protein.

3.1 Introduction

Perhaps the best studied $(BRCT)_2$ domain protein is BRCA1. This protein is involved in various DNA repair pathways, and its C-terminal tandem BRCT domain[36c, 45, 104] is known to bind selectively to phosphoprotein partners. The well-defined and relatively small binding cleft of the $(BRCT)_2$ domain, and the fact that the many mutations that occur in this domain lead to chemotherapeutic and radiation sensitization by disruption of DNA repair, make it a promising therapeutic target.[105] In addition, BRCA1 is involved in at least three different protein complexes mediated by the $(BRCT)_2$ domain.[106] Its involvement in each of these dynamic complexes is dependent on the cell cycle and on the extent of DNA damage. Inhibitors of this domain would therefore allow dosable and temporal control of BRCA1 complex formation.

High throughput in vitro small molecule screens have uncovered molecules that bind to BRCA1 in the 5-10 micromolar range,[47a] yet their activity in cell culture and target selectivity have not been established. Currently, the best binders of this domain are phosphoserine-containing peptides. Several phosphopeptide library screens as well as comparisons of endogenous protein binders have determined that the preferred binding sequence for this domain is phosphoserine (pS)-X-X-F.[104, 107] Libraries lacking phosphoserine were shown not to bind.[104, 107] Natarajan and coworkers have used rational design to optimize binding based on pS-X-X-F and have recently found a pS-containing modified tetrapeptide with a K_d of 40 nM.[108] A recent report showed this phosphopeptide can abrogate BRCA1 function, but the effect was minimal and required high concentrations of peptide making it far from a useful therapeutic or tool.[49] Yet, attempts to replace phosphoserine with phosphomimetic groups have led to dramatically weakened binding affinity.[48b, 109]

We reasoned that by using a peptide library of sufficient size and functional diversity, a non-pS containing peptide could be found that binds to the $(BRCT)_2$ domain of BRCA1. There are many methodologies that have been developed for creation of peptide libraries, but mRNA display[110] provides advantages over other techniques because of the potential for incorporation of non-canonical amino acids[111] as well as the ability to prepare libraries with up to 10^{13} unique peptides,[110b, 111c] 3-6 orders of magnitude more diverse than is possible with on-bead synthesis or phage display. With this increased diversity may come an improved chance of finding higher affinity peptides.[112]

In choosing these UNAAs we have purposefully left out phosphoserine which has been regarded in the literature as a requirement for binding to the BRCA1-$(BRCT)_2$ domain. Additionally, attempts to incorporate phosphomimetic analogs into BRCA1-$(BRCT)_2$ binding peptides have not found great success and in many cases nearly abolish peptide binding. We hope to uncover an alternative means of binding to this domain that is not reliant upon phosphoserine. Such a peptide would be more amenable for use as a drug owing to increased stability and cell permeability.

3.2 Selection Preparation and Optimization

Library Design. The DNA library contained a 12 amino acid random region encoding the peptide sequence $MCX_{12}GSGSL_aGH_6R_aL_a$, with the random, X, amino acids designated by the codon NNB (B = G, T or C) (Figure 3.1A). Usage of this codon vs. the standard NNS/NNK served to decrease the number of stop codons present in the random region (1/48 vs. 1/32 for NNS/K) while also increasing the likelihood of a second Cys (2/48 vs. 1/32) such that

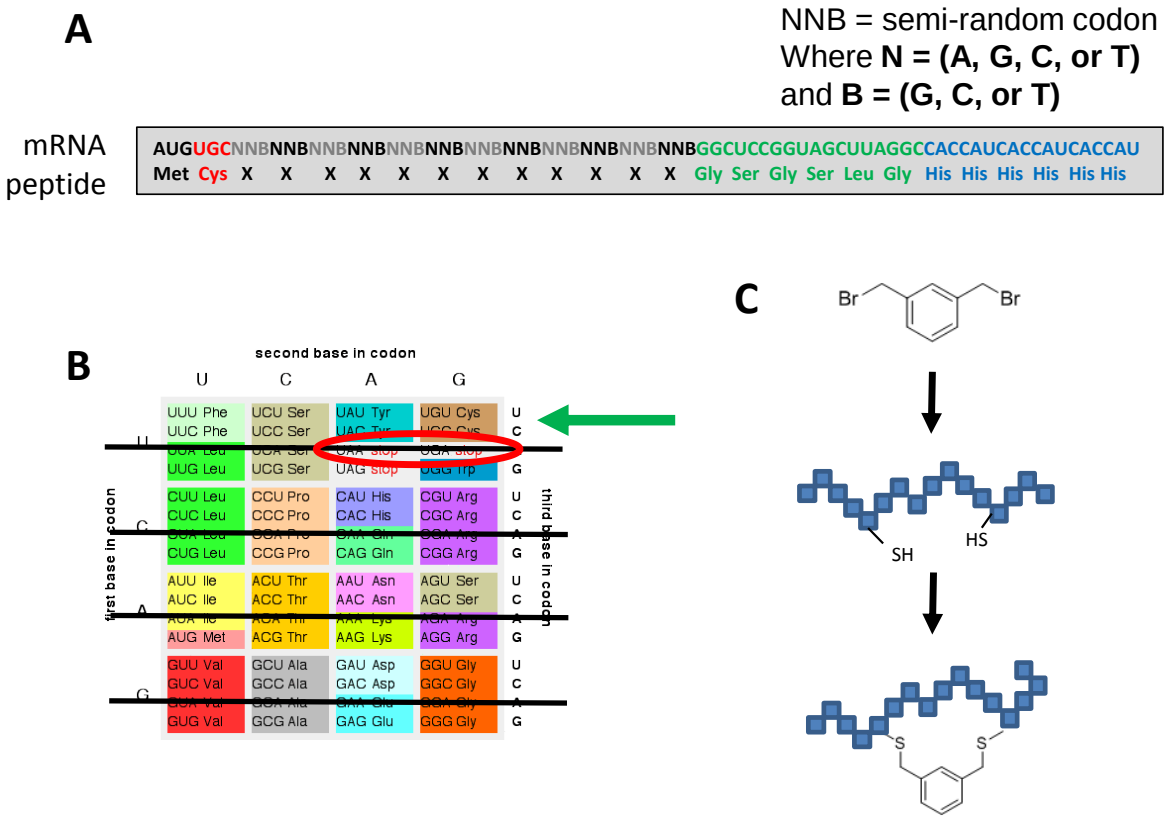


Figure 3.1 Library Design. Peptide library design for selection against the $(BRCT)_2$ domain of BRCA1. A) the mRNA sequence encoding the peptide library including an N-terminal fix cysteine (red) followed by a 12 amino acid random region (black and gray), a Gly-Ser linker (green) and a His_6 tag (blue). B) A codon table displaying remaining available codons with use of the NNB codon where B = G, C, or T, thus eliminating two of three potential stop codons and enriching the presence of cysteine in the library. C) Cyclization of library members with an additional cysteine in the random region will be accomplished via a cyclization reaction with dibromoxylene.

approximately 40% of the initial library should contain at least one additional Cys (Figure 3.1B). The addition of a second cysteine was desired to enable cyclization with α-α'-dibromo-*m*-xylene (Figure 3.1C).[111c, 113] Cyclization is beneficial because it can enhance affinity, cell permeability, and stability.[114]

Expression of GST-$(BRCT)_2$ fusion (01-85)

The BRCA1 $(BRCT)_2$ had previously been cloned into pGEXx4T-1 vector for expression as an N-terminal glutathione S-transferase (GST) fusion. Using a BL21 Codon Plus (DE3) RIL strain, the protein was expressed via ZYM-5052 auto induction overnight at 18 °C.[115] The protein was purified with tandem Ni-NTA agarose (Qiagen) and Glutatione agarose (Thermo) affinity resin and gave 3 mg/L culture. Resultant protein was > 95 % pure by SDS-PAGE analysis (See Figure 3.2), and was stored in enzyme storage buffer at -80 until needed.

Capture and Release of GST-$(BRCT)_2$ with Magnetic GSH Beads In order to perform the capture and elution of the library it is necessary to optimize the binding and release of the protein from its immobilized support. An initial test of bead binding and elution (Figure 3.3A) showed an acceptable level of protein fusion binding to the bead; however a significant amount of protein remained on the beads after four rounds of exposure to 120 µL elution buffer (125 mM Tris pH 8.25, 50 mM GSH, 500 mM NaCl, 1 % Triton X-100) for 10 minutes each. To improve the elution, we tested four buffers noted for their enhanced stringency: 1) 250 mM Tris pH 9.0, 100 mM GSH, 500 mM NaCl, 1 % Triton X-100, 2) 125 mM Tris pH 7.5, 500 mM NaCl, 0.1 % SDS, 5 mM DTT, 3) 100 mM Glycine pH 2.85, 500 mM NaCl, 0.1 % Triton X-100, 4) Boiling in TBS (50 mM Tris pH 8.0, 300 mM NaCl) for 5 minutes.

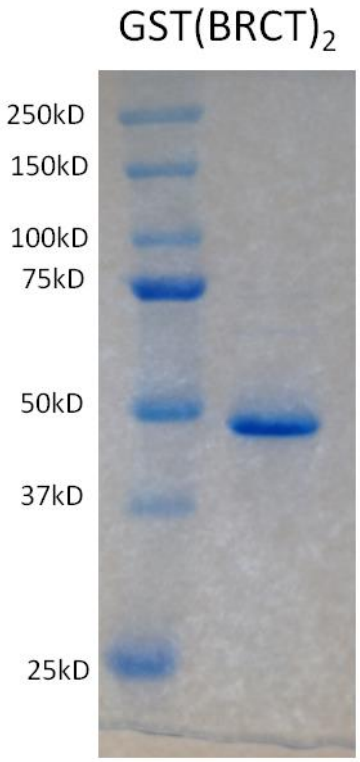


Figure 3.2. SDS-PAGE Analysis of Purified Fusion Proteins. Analysis of purified BRCT fusions by SDS-PAGE and staining with Coomassie Blue shows proteins estimated to be >95% pure. The expected mass of GST-BRCT is 51.3 kDa, and TR-BRCT is 42.9 kDa.

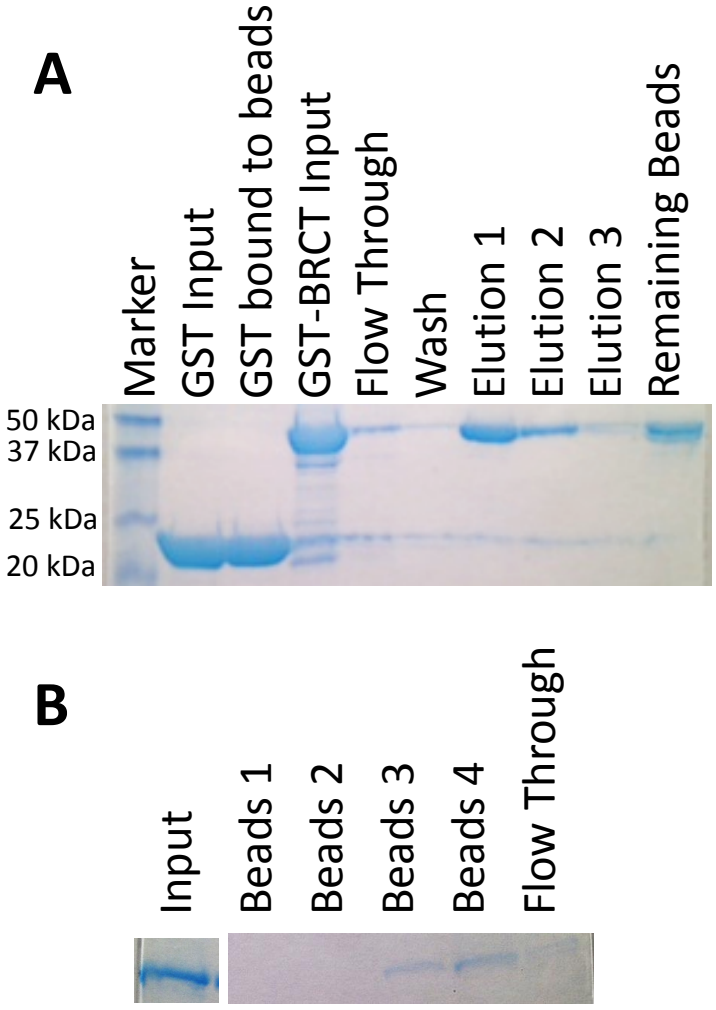


Figure 3.3 Elution of GST-$(BRCT)_2$ Fusion from Glutathione Beads. Only peptides bound to immobilized GST-$(BRCT)_2$ that is successfully eluted will be carried on in the next round of selection so it is important that as much of the fusion be eluted as possible after selection. A) An initial test of the protein fusion binding an elution shows significant protein fusion remaining on the beads (see lane on far right). B) Four elution buffers (see text for details) were tested to find a more stringent buffer capable of complete elution of the protein fusion.

After one round of elution with 200 µL for 10 minutes, buffers 1 and 2 left no detectable protein bound to the bead (Figure 3.3B). Buffer 1 was chosen for use in the library selection.

3.3 mRNA Display Library Selection Against BRCA1-$(BRCT)_2$

The general scheme of the selection process is shown in Figure 3.4. The mRNA peptide fusion library was prepared in the standard way.[111c, 116] Briefly, mRNA was photocrosslinked onto a puromycin-DNA linker. After translation in the presence of unnatural amino acids, a library of mRNAs covalently linked to the peptides they encoded was formed. The mRNA-peptide fusions were purified via Oligo-dT cellulose and cyclized on the resin, followed by reverse transcription and Ni-NTA purification. The translation was performed on a large (10 mL) scale for round 1 leading to the creation of 1.3×10^{13} mRNA peptide fusions, each theoretically unique. The $(BRCT)_2$ domain of BRCA1 was prepared as N-terminal GST-fusion. This fusion protein was immobilized onto glutathione (GSH) magnetic beads, and the mRNA peptide fusions were allowed to bind. Bound mRNA-peptide fusions were sequestered using a magnet, followed by several washes. Elution of the GST fusion from the beads was achieved via competition with excess GSH. PCR of the mRNA-peptide fusions amplified the recovered fusions. The resulting DNAs were in vitro transcribed and the process repeated iteratively. The percentage of the eluted ^{35}S-Met-containing mRNA-peptide fusions vs. the total input was calculated in each round. In the first six rounds, very little enrichment was seen, but after the initial spike in round seven, the beginning of a plateau was seen in round 8 (Figure 3.5).

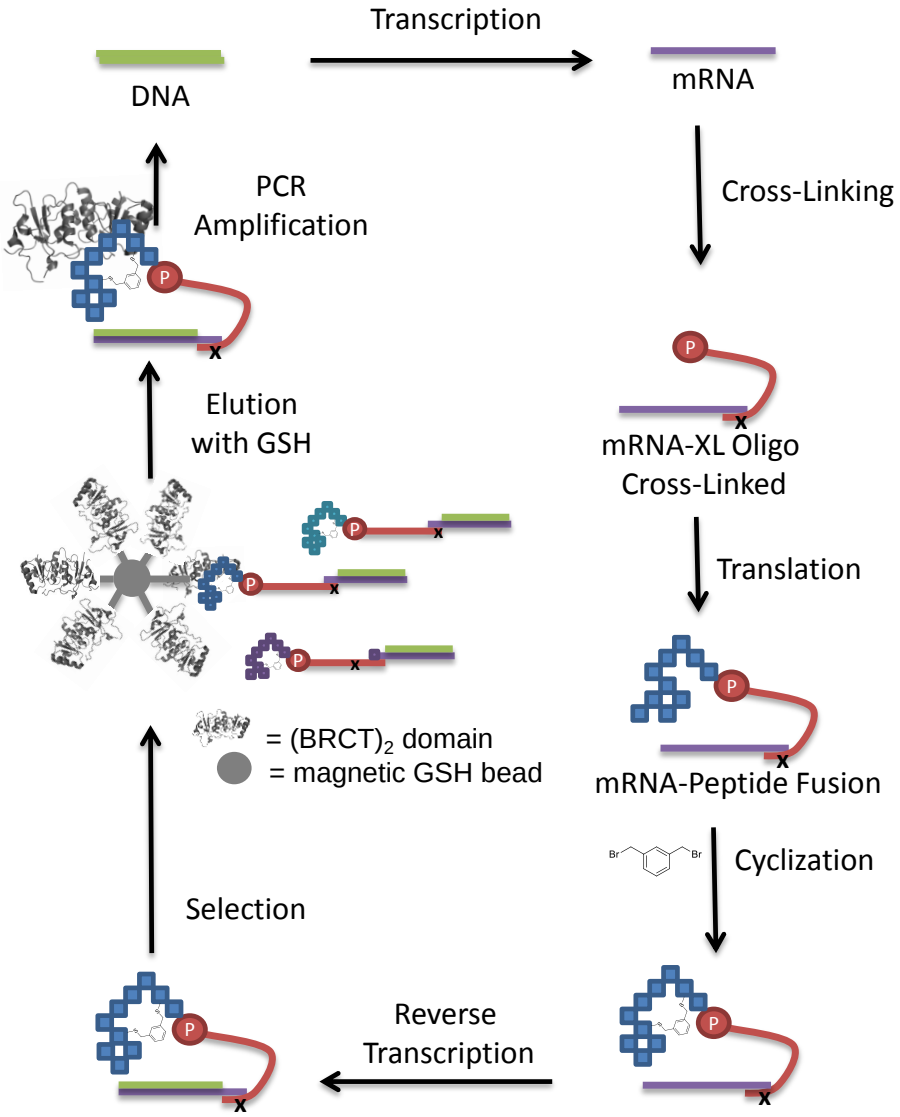


Figure 3.4 In vitro Selection Scheme. The DNA library encodes a twelve amino acids random region with an N-terminal cysteine. After peptide fusion formation, peptides with a second cysteine are cyclized with dibromoxylene. Purified mRNA-peptide fusions undergo reverse transcription and his-tagged purification before being selected for binding to the GST-$(BRCT)_2$ fusion immobilized on magnetic resin. Unbound peptides are washed away and bound peptides are eluted, PCR amplified, and carried through another round of selection. Structural representation of the BRCA1 $(BRCT)_2$ domain was adapted from PDB entry code 1T29 using the PyMOL molecular graphics system (Schrodinger, LLC).

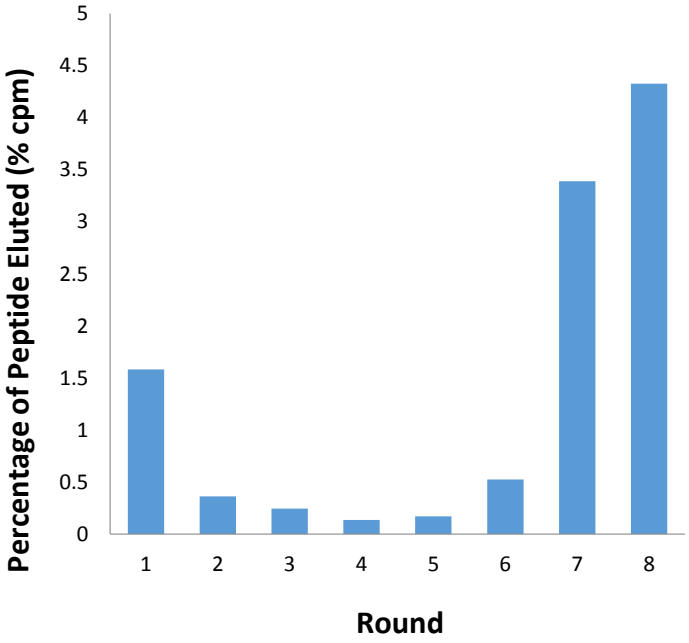


Figure 3.5 Selection Progress. The percentage of ^{35}S-Met labeled peptide eluted with GSH relative to the total ^{35}S-Met input from each round of selection was calculated and is shown for each of eight rounds.

3.4 Sequencing

After the plateau beginning in the eighth round, the cDNAs corresponding to the selected peptides were sequenced following cloning. As shown in Figure 3.6, of the 85 sequences found, 7 sequences appeared more than once, and these duplicate sequences comprised more than 80 % of the total sequencing hits. Alignment of the top seven sequences is shown in Figure 3.7.

3.5 Discussion

Unnatural amino acids appeared many times in the random region of the selected sequences, but the variety of different amino acids was limited to primarily F_{26} and R_1. Analogs L_3 and P_1 were present in families 8.3/8.4 and 8.7 respectively, while V_3 and W_2 were limited to single appearances in the "other" sequences. The most notable motif is the recurring D/E-X-X-F_a sequence that is found in 80 of the 85 sequences. This motif is remarkably similar to the pS-X-X-F motif known to be found in all known BRCA1 $(BRCT)_2$-domain binding proteins,[106, 117] where the Phe is replaced with 4-fluoro-Phe and the pSer is replaced with either aspartic or glutamic acid, both being known, but weak pS-mimetics.[118] These results were somewhat surprising, because prior attempts to mutate pS to E led to abrogated BRCA1 binding.[109] Previously oriented SPOT library screens have shown that peptides containing β-branched and aromatic AAs in the X-X positions are preferred.[107] This preference is also mirrored in many of our sequences.

3.6 Experimental

Protein Expression. The plasmid construct containing the BRCA1 $(BRCT)_2$ domain (amino acids 1646-1859)[119] was digested with restriction enzymes NdeI and XhoI. PCR amplification

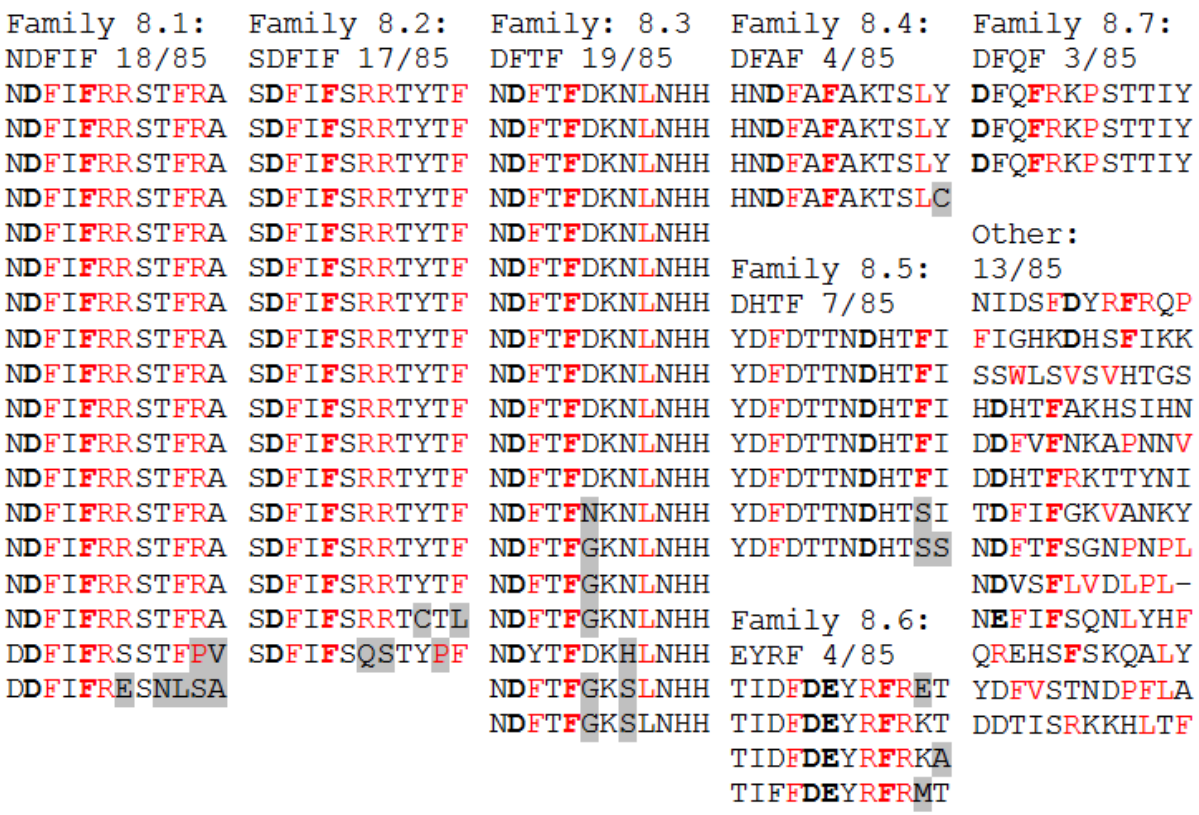


Figure 3.6 Selection Sequencing Results. After round eight, cDNA from the library pool was cloned and sequenced giving the 85 sequences shown, and similar sequences were arranged into 7 families. Unnatural amino acids are designed in red.

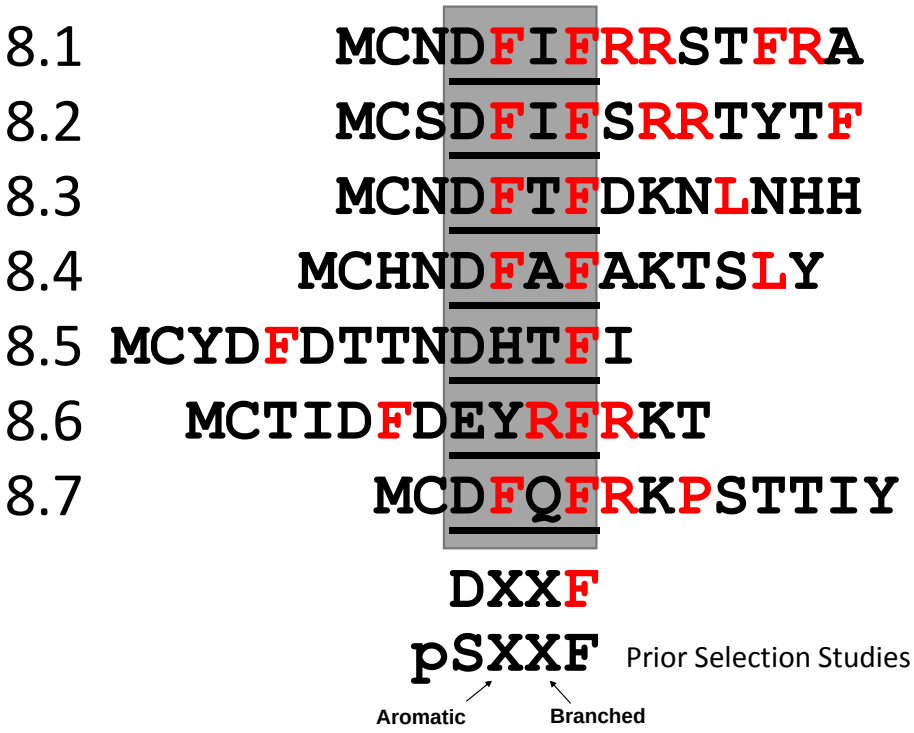


Figure 3.7 Family Sequence Alignment. The top seven families are aligned with respect to the D/E-X-X-F motif found in each family (emphasized with a gray rectangle). Un-nautral amino acids are indicated in red.

7

with primers BRCT FWD (5'-CGGGATCCGTCAACAAAAGAATGTCCATGGTGGTGTC-3') and REV (5'-CCGCTCGAGTCAGTGGTGGTGGTGGTGGTGGGGGATCTGGGGTATC-3') were used to amplify the insert creating a BamHI restriction site at the N-terminus and adding a hexahistidine tag to the C-terminus. The insert was then ligated into a pGEX-4T-1 vector to form a C-terminal Glutathione S-Transferase (GST) fusion which was then verified by sequencing. After transforming into *E. coli* strain BL21(DE3)-RIL cells, protein expression was induced with auto-inducing ZYM-5052 media[120] with overnight incubation at 18 °C. Additionally a modified pET32a vector previously described[121] was used to prepare an N-terminal thioredoxin fusion with C-terminal His-tag (TR-BRCT). The pGEX-4T-1 vector was digested with BamHI and XhoI and ligated into the modified pET32a. After transformation into BL21 (DE3) strain of *E. coli*, cells were expressed using auto-inducing media overnight at 18 °C.[115] The recombinant protein was purified via tandem affinity purification, first with binding to Ni-NTA agarose (QIAGEN) while tumbling at 4 °C for 60 minutes. After washing twice (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 5mM BME), six elutions (1.5 ml each of 50 mM NaH2PO4, 300 mM NaCl, 250 mM imidazole, 5 mM BME) were combined an dialyzed overnight into 1X PBS. After dialysis, protein was combined with glutatione-agarose resin (Thermo) while tumbling for 1.5 hours at 4 °C. The resin was washed twice with 1X PBS and the GST-fusion protein was eluted with 6 x 1 mL fractions (50 mM Tris pH 8.0, 10 mM glutathione). The resulting yield was 3 mg/L and fusion proteins were >95% pure as estimated by SDS–PAGE analysis. Protein was stored at -80 °C after dialysis in enzyme storage buffer (50 mM HEPES-KOH pH 7.6. 100 mM KCl, 10 mM MgCl2, 7 mM BME, 30 % glycerol).

Test of GST-$(BRCT)_2$ Immobilization onto Magnetic Beads. Purified GST-$(BRCT)_2$ fusion protein was diluted 1:5 with binding buffer (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 4 mM

MgCl2, 0.25 % Triton X-100, 0.1 mg/ml BSA) and incubated with glutathione magnetic beads (Thermo) with tumbling for 1 hour at 4 °C. Any unbound protein was taken as the flow through and beads were washed with twice with binding buffer. Bound protein was eluted with elution buffer indicated, and fractions of each step were analysed with a 10 % SDS-PAGE gel.

Preparation of mRNA Fusions. The in vitro selection protocol has previously been outlined in detail[116] and is described here in brief. The sequence of the cDNA encoding the mRNA display library was: 5'-CTAGCTACCTATAGCCGGTGGTGATGGTGATGGTGGCCTAAGCTACCG GAGCCVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNGCACATTTAGCTGT CCTCCTTACTAAAGTTAACCCTATAGTGAGTCGTATTA-3'. The library was ordered from the HHMI/Keck synthesis facility and was purified by Urea-PAGE. Following PCR amplification of library templates, sequences were verified with TOPO cloning and sequencing. An in vitro transcription reaction was followed by Urea-PAGE purification and electroelution to obtain purified mRNAs to be used for selection. The library mRNA was photochemically crosslinked to a puromycin linker after annealing followed by exposure to UV irradiation, and was then used in translation to produce peptide-mRNA fusions. A 10 mL scale translation was used to generate peptides for the first round of selection. After translation, fusion formation was maximized through the addition of KCl/$Mg(OAc)_2$ as described.[122] The mRNA peptide fusions were bound to Oligo(dT) resin via the poly-A sequence contained in the puromycin linker. Peptides were cyclized on the resin using dibromoxylene, and eluted with water. The cyclization reaction was conducted as described.[111c] This was followed by reverse transcription of the mRNA with a primer of the sequence 5'-TTTTTTTTTTTTTTT CTAGCTACCTATAGCCGGT GGTGATGGTGATGGTGGCCTAAGC-3', and Ni-NTA purification of the full-length peptide fusions containing a C-terminal hexahistidine tag. The library was then dialyzed into selection

buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $MgCl_2$, 0.25% Triton X-100) and used in the first round of selection. The yield of peptide fusions after all purification steps was 22 pmol, equivalent to 1.3 x 10^{13} peptides.

Selection. Prior to the selection, beads containing GST and GST-BRCT were prepared. GST beads: 200 µL of magnetic glutathione beads (Pierce) were washed three times with 1 mL GSH beads wash buffer (125 mM Tris-HCl pH 8, 150 mM NaCl). 1 mL of 10 µM GST in GSH beads wash buffer was added to the beads and tumbled for 1 hour at 4°C. The supernatant was then discarded and the beads were washed 2x with 1 mL GSH beads wash buffer and 1 x 1 mL selection buffer. GST-BRCT beads: a separate 400 µL of bead suspension was washed 3x with GSH beads wash buffer. 1mL of 10 µM GST-$(BRCT)_2$ was added and tumbled 1 hour at 4°C. GST-$(BRCT)_2$ bound beads were washed 2x with GST beads wash buffer and 1x with selection buffer. The previously purified peptide-mRNA fusions were dissolved in 1200 µL selection buffer. Pre-clear: fusions were added to GST bound beads and tumbled 1 hour at 4°C. Selection: the supernatant containing the library was transferred from the GST bound beads to the GST-$(BRCT)_2$ bound beads. With an additional three washes, the total volume transferred to the GST-$(BRCT)_2$ beads was 1300 µL. 13 µL of 10 mg/mL BSA was also added and the tube was tumbled for 1 hour at 4°C. The beads were washed 3x with 1 mL selection buffer. Peptide-fusions were eluted along with the bound protein by the addition of 6 x 100 µL freshly prepared GSH elution buffer (250 mM Tris-HCl pH 9, 500 mM NaCl, 100mM L-glutathione reduced (Sigma), 1% Triton X-100). Each addition of elution buffer was allowed to incubate for 5 minutes. Portions of selection input, flow through, washes, re-suspended beads and elution fractions were quantified by scintillation counting of ^{35}S-Met and used to monitor selection enrichment. Elution fractions were combined and dialyzed overnight at 4°C into 0.1% Triton X-

100 prior to PCR amplification with Library FWD primer: 5'-TAATACGACTCACTAT AGGGTTAACTTTAGTAAGGAGGACAGC-3', and Library REV 5'-CTAGCTACCTATAG CCGGTGGTGATGGTGATGGTGGCC-3' primers. The amplified cDNA was then transcribed and used for the subsequent round of selection. The scale of the second round of selection was a 1 mL translation, followed by rounds 3-8 that were started with 500 µL translation reactions. Additionally, pre-clearing of the library bound to GST beads was only performed in rounds 1 and 2.

cDNA Sequencing. To analyze the results of the selection, cDNAs were amplified by PCR and cloned into the pCR-TOPO vector (TOPO TA Cloning Kit, Invitrogen). Unnatural amino acids were assigned on the basis of the tRNA/AARS pairs responsible for their incorporation into peptides. For DNA and peptide sequence alignments and homology ranking, Jalview Version 2 was used.[123]

3.7 Summary

An mRNA display selection with 1.3×10^{13} peptide members was selected against a recombinant GST fusion of the BRCA $(BRCT)_2$ domain. After 8 rounds, the surviving peptides' sequences showed 7 sequences occurring multiple times, making up more than 80 % of determined sequences. No evidence of cyclization is apparent from the sequences, with an UNAA presence biased toward F_{26} and R_1.

Future Directions. With peptide sequences in hand, it is necessary to rank the peptides by affinity as well as to determine how tightly they bind. Where these peptides bind on the BRCT

domain is also important for these peptides to be applied as inhibitors, and will need to be investigated.

CHAPTER 4. LIFE AFTER SEQUENCING: RANKING THE TOP PEPTIDE HITS

Contributions:

ERW and David E. Hacker are responsible for peptide synthesis and purification.

ERW conducted all of the biacore, ITC and FP experiments.

David E. Hacker is responsible for the MS/MS experiments of cyclized peptides.

Chapter 4: Life after Sequencing: Ranking the Top Peptide Hits

4.1 Introduction

After 8 rounds of mRNA-display selection against the $(BRCT)_2$ domain of BRCA1, a group of 85 sequences were determined. Nearly 80% of these sequences occurred more than once and these 7 sequences were grouped together into families 8.1-8.7, ordered by frequency. The D/E-X-X-F motif found in each of these sequences suggests these peptides may bind to the same binding pockets as the pS-X-X-F peptides; however, before we can begin to investigate such questions it is necessary to first confirm that these peptides bind to the $(BRCT)_2$ domain. There are many ways to determine a binding affinity, but before investing too much into determining an exact affinity, a quick ranking of these peptides via a radiation spin assay was pursued.

4.2 Equilibrium Ultrafiltration Binding Assay

Although there still isn't an efficient means of incorporating the arginine analog canavanine via solid phase peptide synthesis, synthesis of these peptides should be accessible via *in vitro* translation. The scale of these reactions is small, but due to the presence of an ^{35}S-methionine it is sufficient for radiation spin assays to determine an approximate K_d of each peptide. This assay, as shown in Figure 4.1, involves the combination of a constant concentration of radiolabeled peptide with varying concentrations of GST-$(BRCT)_2$ protein. Each sample will then be centrifuged in a 30,000 MWCO spin filter such that approximately half of the solution remains in the top portion of the filter and half has been filtered through. By calculating the fraction of peptide bound in each sample, a binding curve can be established

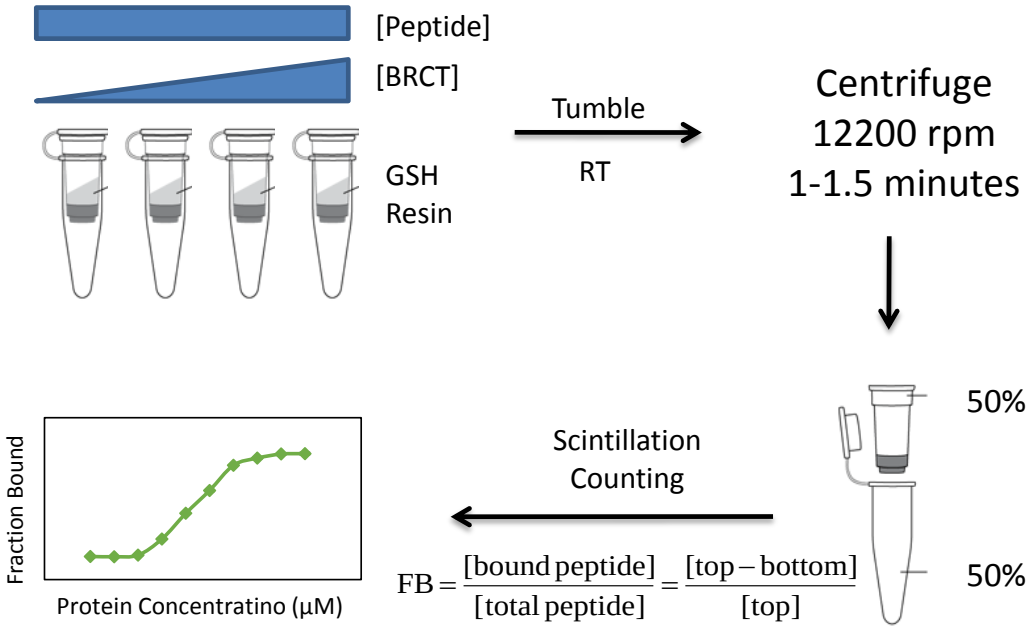


Figure 4.1 Equilibrium Ultrafiltration Binding Assay. Various concentration of GST-$(BRCT)_2$ fusion protein is added to concentration of ^{35}S-labeled peptide. After incubation, the samples are centrifuged in 30,000 MWCO filters so that the sample is divided, half filtered through the membrane, and half remaining in the top. A fraction bound can ben calculated in each sample to achieve a binding curve.

effectively ranking these 7 peptide families.

A cDNA representative of each peptide family was selected to generate the mRNA template used for selection. A second peptide was chosen from family 2 that appeared to have a second cysteine, which was called family 8.2c. All sequences chosen had a full length His-tag. The glycerol stocks from sequencing were used to generate purified plasmid which was used in PCR amplification of the cDNA insert (Figure 4.2). *In vitro* transcription was used to generate the mRNA template used for *in vitro* transcription.

The resulting mRNA was used in 250-500 µL scale translation reactions. The large scale reaction included ^{35}S-methionine, but a 50 µL reaction was conducted in parallel for characterization by MALDI. Many of the translation reactions suffered from poor yield for unknown reason, and even those reactions with sufficient yield to conduct the spin assays were not able to be characterized by MALDI. Typically we can detect as little as 1 pmol after zip-tipping; however, not a single peptide produced from these mRNA templates was able to be characterized even in reactions with yields greater than 1 pmol. Despite the lack of characterization, attempts at the radio-labeled spin assay were conducted with no observed binding of any peptide. At this point it was decided to try to synthesize peptides with solid phase peptide synthesis and characterized their binding to the $(BRCT)_2$ domain via surface plasmon resonance (SPR).

4.3 Surface Plasmon Resonance

Being able to synthesize peptides via *in vitro* translation should have been a rapid method of synthesis and characterization; however efficiency was not its only advantage. As of this

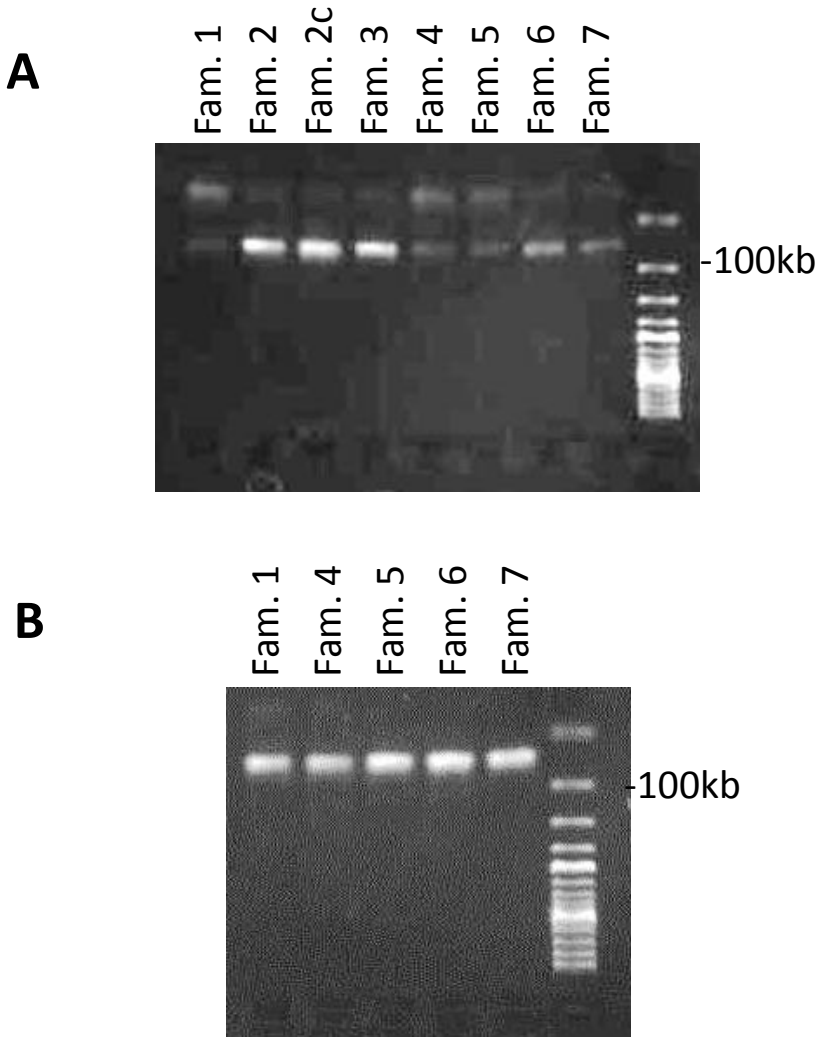


Figure 4.2 PCR Amplificaton of Library cDNA. One cDNA was selected for each of the over represented families from the selection. After 24 rounds of PCR amplification, the product was analyzed on a 1 % agarose gel. A) PCR amplification with primers from selection. B) PCR product using shorter primers.

writing, canavanine in an appropriately protected form for solid phase peptide synthesis (SPPS) was not commercially available. Only one published report of the synthesis of a similar analog Fmoc-canavanine(Mtr)-OH exists, and attempts to repeat the reported procedure have been unsuccessful.[124] Although progress is being made toward alternative routes of synthesis, it has remained a formidable challenge owing largely to the lability of the O-N bond in procedures typically used for Fmoc-arginine(Pbf)-OH as well as the lack of nucleophilicity of the guanidine side chain of canavanine. With synthesis of peptides via *in vitro* translation an unsuccessful endeavor, the decision was made to work toward synthesis via SPPS, with the possibility of including arginine in the place of canavanine if necessary.

At this time it was hoped that we would soon be able to synthesize a suitably protected canavanine, thus the peptide characterization began with those peptides without canavanine. The first peptide chosen was peptide 8.3. Optimistic thinking that cell studies may quickly follow the peptide ranking, a nuclear localization sequence (NLS) was added to the C-terminus of the peptide making the full peptide sequence: MCNDF$_{26}$TF$_{26}$DKNL$_{3}$NHHGS**PKKKRKV**, where the underlined portion is the sequence of peptide 8.3 and the bold indicates the NLS sequence.

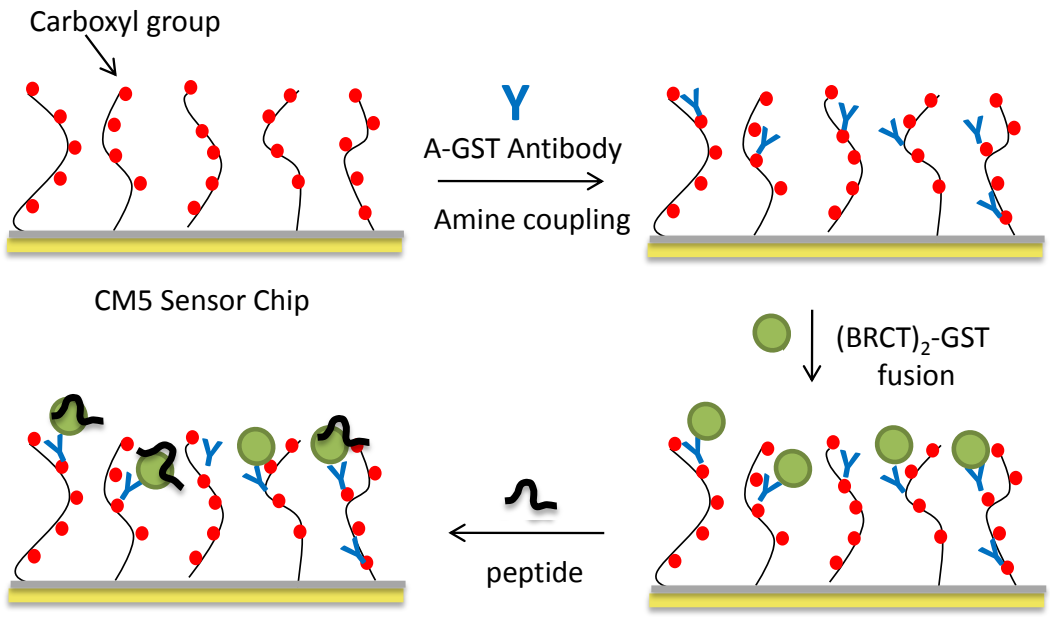
It was decided that surface plasmon resonance (SPR) would be a potentially quick means of ranking the affinity of the peptides and would also give us kinetic information about both the on and off rates of each peptide. SPR is one of the few techniques that allows measurement of the kinetics of binding in real time, which is very powerful in characterizing a binding affinity because both k_{on} and k_{off} can be easily determined. However, because this measurement takes place on the surface of a chip surface binding artifacts can obscure the solution phase kinetics.

Amine coupling of the GST-$(BRCT)_2$ to the chip surface was tried at first, but this technique does not control for orientation of the protein on the chip surface. Direct coupling of a protein makes the expected response from peptide binding to the protein more difficult to estimate. Due to the multiple protein orientations, peptides will not be able to bind to every protein. When peptide binds to the protein on the surface, the amount of peptide bound can be correlated to the response units (RU) measured by the detector with the following equation.

$$R_{max} = R_{ligand}(\text{stoichiometry})\left(\frac{MW_{ligand}}{MW_{protein\ fusion}}\right) \qquad [4.1]$$

where R_{max} is the the predicted total response of an assay in RU; R_{ligand} is the response of the bound protein in RU; the stoichiometry is equal to 1; MW_{ligand} is the molecular wieght of the peptide, 2939 amu; and $WM_{protein\ fusion}$ is the molecular weight of the GST-$(BRCT)_2$ protein bound to the surface 51.2 kDa. Although directly coupling protein to the surface may underestmate the maximum observable response from peptide binding, a more important disadvantage of this technique is the inability to regenerate the bound peptide. Peptides are only stable for a certain period of time, and it is advantageous to use fresh fusion peptide for each analysis.

By switching to a configuration, shown in Figure 4.3, where an anti-GST antibody was first coupled to the chip surface, the RU could accurately be predicted, and the fusion protein could be freshly added for each experiment leading to more accurate analysis. From initial experiments it was found that the RU resulting from peptide biding was orders of magnitude greater than the expected RU. This likely indicates that the peptide is non-specifically binding to the chip surface. The surface of the chip is a negatively charged dextran, so it is not surprising that the positively charged NLS sequence would bind to the chip surface. This was confirmed when a series of samples were run at various salt concentrations (Figure 4.4), where salt



Figrue 4.3 SPR Experiment with α-GST Antibody. First a α-GST antibody is covalently coupled to the dextran coated chip using standard amine coupling chemistry. The $(BRCT)_2$-GST fusion is then immobilized on the chip via interaction with the antibody. Various concentrations of peptide are then flowed over the chip and the resulting response is recorded.

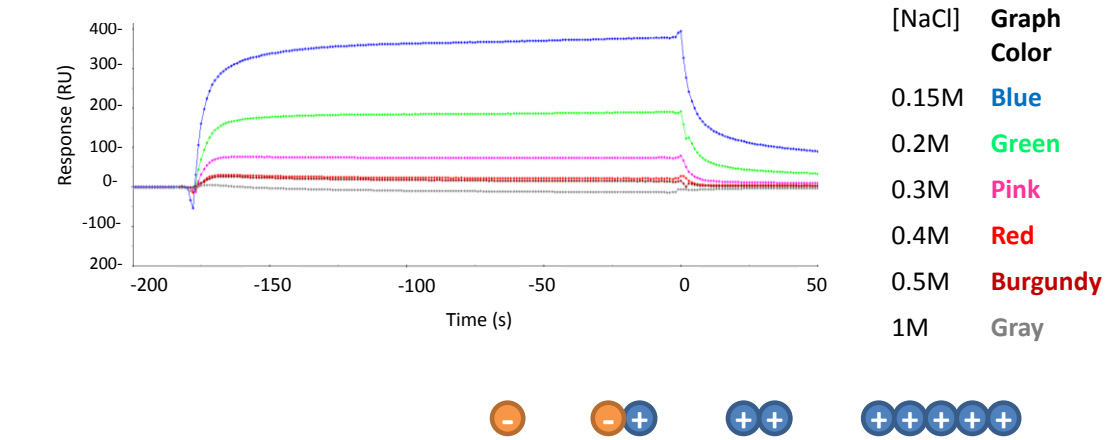


Figure 4.4 SPR with Nuclear Localization Sequence. The above trace shows the response units recorded for peptide samples in buffer containing different salt concentrations binding to the immobilized GST-$(BRCT)_2$. Below is the sequence of the peptide used in this study with the charges indicated.

concentrations as high as 0.5 M NaCl were able to able to interrupt the non-specific interactions bringing the RU down to the expected range. Although, there are other chip surfaces, such as an amine based chip that could also eliminate the non-specific interaction, it may have been just as simple to re-synthesize the peptide without the NLS which was likely causing the problem. It was while making this decision, the instrument developed technical difficulties not readily remedied, and the decision was made to pursue alternative means of peptide characterization.

4.4 Isothermal Titration Calorimetry

Because Isothermal Titration Calorimetry (ITC) is a solution based technique, the presence of a NLS in theory should not interfere with characterization of peptide binding. The NLS may alter the affinity of the peptide, but the same surface effects seen with SPR should not be an issue. There was a significant learning curve when using ITC that can be mostly summed up into a single concept: the ITC is a sensitive instrument and must be very clean to get accurate results. This however, is often easier said than done, and in the case of our peptides this learning curve was complicated due to the low affinity of the peptides. With ITC, there is a range of affinities that are optimal for analysis, which can be defined by the following equation

$$c = nK_a[M]_t \qquad [4.2]$$

where c is the Wiseman c parameter, n is the molar ratio, K_a is the association constant, and $[M]_t$ is the total protein concentration.[97]

The optimal range for any ITC analysis is a c value between 10 and 100. At high c values, the titration curve becomes too steep, and at low c values the curve becomes a shallow almost linear slope. Both of these scenarios make accurate data analysis difficult, if not impossible. As can be seen in the equation above, c is dependent up on n (reaction

stoichiometry) and K_a (the association constant) which are inherent to the interaction under study and therefore cannot be altered. This leaves the concentration of protein, $[M]_t$, as the only parameter that can be manipulated to produce an experiment with an optimal c value. For high affinity inhibitors, the limitation of protein dilution to lower the c value is dependent upon the amount of heat than can be measure in dilute solutions. For low affinity interactions, achieving high protein concentration is limited by protein solubility. The GST-$(BRCT)_2$ fusion protein that was used for selection worked very well at the low concentrations needed for immobilization to magnetic beads, or to the surface of an SPR chip, but for ITC it was necessary to have a protein concentration as high as possible. Solubility of the GST-$(BRCT)_2$ fusion became a significant problem during attempts to achieve the high concentrations needed for ITC analysis. Precipitation was observed during the concentrating process as well as during the actual ITC analysis. One means of explanation for this observation is because GST exists as a dimer, and the BRCA1 $(BRCT)_2$ domain that displays weak homodimerization. At low concentrations, solubility was not a significant problem, but as the concentration increased for ITC analysis, perhaps these two dimerizing proteins formed a polymer resulting in significant precipitation. Despite these hurdles, reasonably successful ITC analysis was achieved for 4 peptides (Figure 4.5). The only peptide with a discernable sigmoidal curve is peptide 8.3 D4pS. The plots that do not appear sigmoidal have high error as they fall below the recommended c value for ITC analysis.

4.5 MS/MS Analysis of Peptide Cyclization

While initial SPR and ITC studies were being conducted with peptide 3, DEH began investigating what product was formed in the cyclization reaction with these peptides that did not

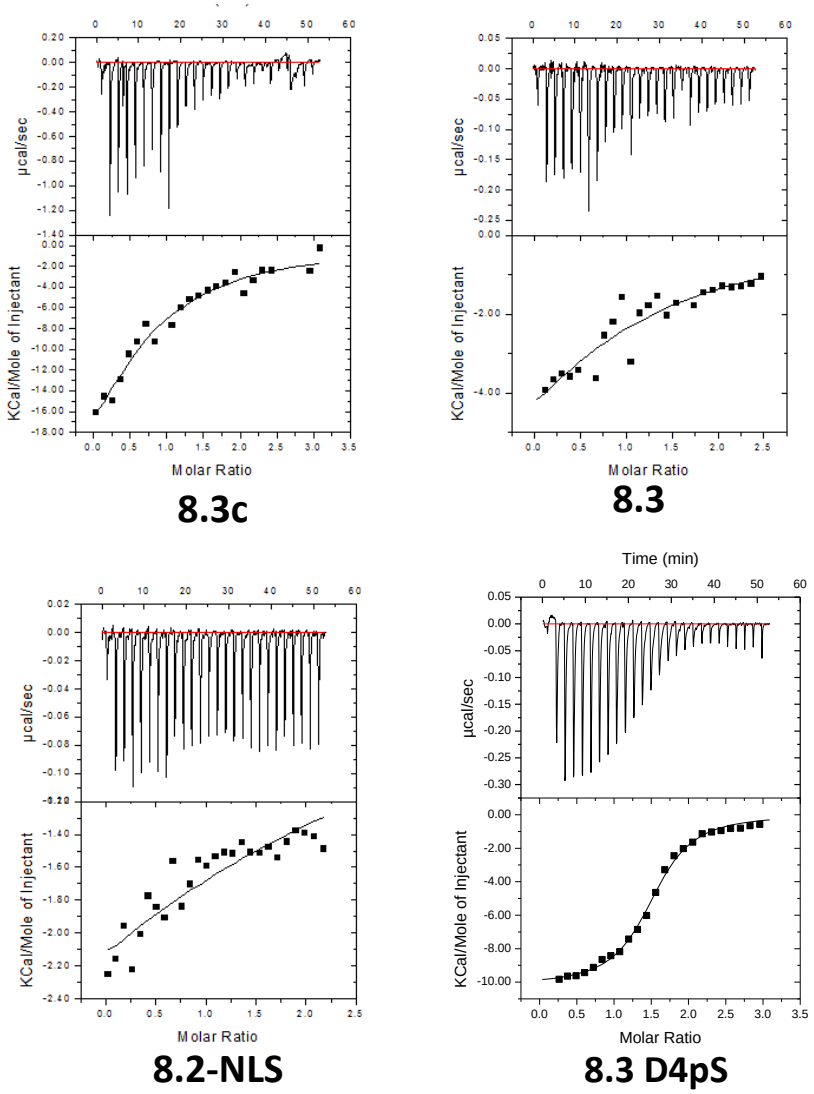


Figure 4.5 Initial ITC Data with NLS Peptides. Shown here are the initial ITC curves obtain for four peptides containing nuclear localization sequences (NLS).

have a second cysteine. It was a bit surprising that none of the sequences contained a second cysteine, and it was unclear the fate of mono-Cys peptides. A peptide of the sequence MCNDF_{26}TF_{26}DKNL_3NHHGSPKKKRKV was subject to cyclization and analyzed by ESI-MS/MS. The results of this analysis indicate cyclization is limited to an interaction between the peptide's N-terminal Met and Cys, forming a cyclic sulfonium group (Figure 4.6).

4.6 Fluorescence Polarization

After running into technical difficulties with the radiolabeled spin assay, SPR and ITC analysis of peptides, it was determined that a fluorescence polarization (FP) assay might be the best alternative to rank the peptides from selection. Like ITC, FP has some limitations with regard to the range of binding affinities it can accurately measure. ITC is limited by its c value; however, FP is often not suitable to measure low affinity proteins due to the high protein concentrations needed to produce a full binding curve. In a typical assay, a fluorescently labelled peptide is kept at a constant concentration with increasing concentration of protein. At the high concentrations needed to examine low binding affinities, protein solubility and aggregation begin to interfere with analysis due to non-specific light scattering that exists in samples containing aggregates or with higher viscosity. This rationale combined with the simplicity of synthesizing a single fluorescently-labelled inhibitor instead of labelling each peptide under study made a competition study an ideal means to rank the peptides. Although this will not directly determine a binding affinity in terms of K_d, it will allow for determination of an IC50 with respect to the known inhibitor. In other words, a competitive FP study will determine the half maximal concentration of our competing peptides that can successfully prevent the know inhibitor from binding.

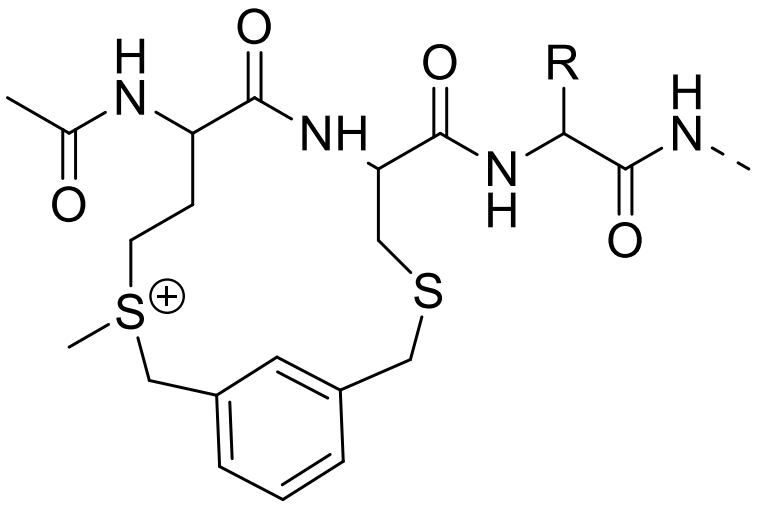


Figure 4.6 Proposed Cyclization Structure. This structure was proposed based on MS/MS data of library peptide 8.3 cyclized with dibromoxylene. In the absence of a second cysteine, it is thought that cyclization is occurring between the N-terminal Met-Cys producing the sulfonium ion shown.

FAM-β-A-pSPTF-NH_2 is a known BRCA1 $(BRCT)_2$ binding peptide that has been used previously in FP competition assays.[108] All competitor peptides from the library were generated via SPPS and purified by HPLC, with verification by MALDI (See Appendix I). As of this writing, protected canavanine suitable for SPPS has still not been successfully synthesized, so each of the peptides characterized by FP has arginine where canavanine was present in the peptide sequence. The major difference between the two molecules is their side chain pK_as (Arg pKa = 12.5, canavanine pKa = 7.0).[125]

Initial fluorescence polarization competition assays were conducted with representative peptides in their full length, un-cyclized form. All assays utilized a thioredoxin (TR) fusion protein, TR-$(BRCT)_2$, which has superior solubility to the GST-(BRCT)2 fusion. SDS-PAGE analysis of this proteins purity can be found in Figure 4.7. All seven peptides were tested in their linear forms, and five of them were tested as cyclization products as solubility and stability allowed. The curves resulting from this analysis can be found in Appendix II. The IC_{50} values determined for all twelve peptides can be seen in Table 4.1. Linear peptide 8.6 had the highest binding affinity with an IC_{50} of 10.5 μM, binding twice as strongly as peptides 8.1 and 8.5, and more than five-fold more strongly than any other peptide tested. The effect of cyclization had a negative impact on affinity except in the case of peptide 8.1 where there was a neutral effect on binding.

4.7 Discussion

In the selection sequences, no second cysteines were found in the random region, but MS/MS analysis of a cyclized peptide revealed the presence of dibromoxylene cyclization between the N-terminal Met-Cys. It is possible that the reactive sulfonium ion formed in

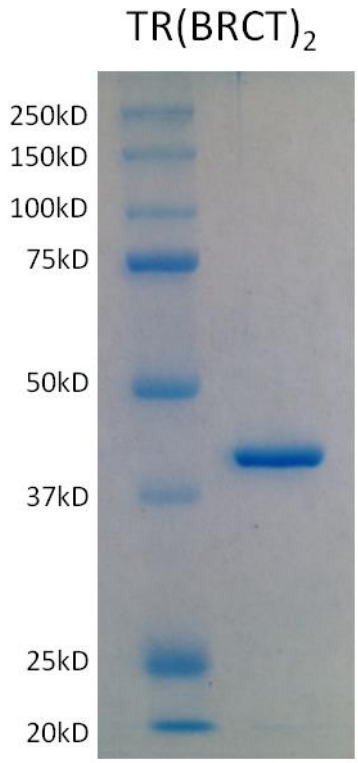


Figure 4.7 SDS-PAGE Analysis of Purified Thioredoxin-$(BRCT)_2$ Fusion Protein. Analysis of the purified BRCT fusions by SDS-PAGE and staining with Coomassie Blue shows proteins estimated to be >95% pure. The expected mass of TR-$(BRCT)_2$ is 42.9 kDa.

Peptide	Sequence	IC_{50} (μM)
8.6	MCTIDFDEYRFRKT	10 ± 1
8.6c	**MC**TIDFDEYRFRKT	24 ± 2
8.1c	**MC**NDFIFRRSTFRA	25 ± 2
8.1	MCNDFIFRRSTFRA	24 ± 2
8.5	MCYDFDTTNDHTFI	26 ± 2
8.2	MCSDFIFSRRTYTF	49 ± 5
8.4	MCHNDFAFAKTSLY	56 ± 4
8.7	MCDFQFRKPSTTIY	56 ± 6
8.3	MCNDFTFDKNLNHH	99 ± 8
8.7c	**MC**DFQFRKPSTTIY	122 ± 9
8.4c	**MC**HNDFAFAKTSLY	125 ± 12
8.3c	**MC**NDFTFDKNLNHH	142 ± 15

Table 4.1 Ranking of Peptides from Selection. The linear form of all seven peptides and the cyclized form of five, were tested in a competition assay with FAM-β-A-pSPTF, a peptide known to bind the $(BRCT)_2$ domain.

cyclization could result in a covalent linkage upon binding to the protein if the right functional groups were near; however, our Biacore data did not support a covalent mechanism.

The seven peptide families in both linear and cyclized forms were eventually ranked by their binding affinity by FP after technical hurdles were uncovered with other techniques. Comparison of the binding affinity of linear and cyclized peptides with FP showed that cyclization had a negative effect on binding affinity in all cases except peptide 8.1 where the affect was neutral. Since 40% of the library should have had a second cysteine somewhere in the random region, and none of these were selected seems to suggest that the $(BRCT)_2$ domain prefers a more extended binding motif. The question of whether the entire length of these peptides is required to bind remains unanswered, but suggests that the full length may be necessary, or that the presence of cyclization with a second cysteine actually produced peptides in unfavourable binding confirmations.

Although each of the techniques employed here are suitable for ranking the peptides from selection based on their approximate affinities, technical challenges were a significant hurdle in this endeavor. The radiation spin assay could have allowed us to easily rank the peptides as well as to examine the effects of UNAA compared to the all-natural analogs. The frequently poor yields obtained from these reactions as well as the inability to characterize these peptides by MALDI-TOF led us to pursue techniques using peptides made with SPPS. Initially peptide 8.3 (which was chosen due to its lack of canavanine which was not able to be incorporated in the solid phase) was examined by SPR in both its linear and cyclized state with the addition of an NLS. In the end, the positive charge of the NLS led to significant non-specific interaction with the negatively charged SPR chip. By the time this was confirmed, technical difficulties with the instrumentation led us to pursue other avenues of analysis.

Testing these same peptides with ITC led to less than optimal results partially due to technical problems with instrument maintenance and GST-$(BRCT)_2$ solubility, but also because these peptides have low enough affinity, perhaps affected by the presence of the NLS, that they are outside the optimal range of analysis for ITC. It was then decided that all 7 peptides should be synthesized on the solid phase without an NLS or any other modification with arginine in the place of canavanine. These peptides in their linear and cyclized forms were ranked by IC_{50} through a competitive FP assay against a known inhibitor FAM-β-A-pSPTF. The FP curves collected can be found in the Appendix II. The highest affinity peptide was fond to be the linear form of peptide 8.6. This peptide will be the subject of further investigation into how this peptide binds to the $(BRCT)_2$ domain.

Unfortunately, not being able incorporate the canavanine that was present in selection leaves many questions unanswered. Linear peptide 8.6 had the highest affinity with an IC_{50} of 10.5 μM. Would this have been lower if canavanine were in place of arginine? Inclusion of arginine means that this residue has a positive charge where none was present during selection, as well as a carbon where an oxygen would be which could alter potential hydrogen bonding interactions. Other arginine analogs such as citrulline could be used to approximate the effect of canavanine's neutral charge, but the only current means to investigate peptides with canavanine is to synthesize them with *in vitro* translation.

4.8 Experimental

Generation of peptides with *in vitro* translation. Glyerol stocks from selection were grown up overnight in 1ml LB containing kanamycin. The next day the cultures were removed from the shaker and cells were collected via centrifugation. The plasmids were purified with Qiagen

miniprep. PCR amplification was conducted with the same primers used for selection with serial dilutions of the template plasmid with 24 rounds of amplification. The primers used for amplification of families, 1, 4, 5, 6, and 7 had the following sequences: FWD TAATACGACTC ACTATAGGGTTAACTTTAGTAAG, REV CTAGCTACCTATAGCCGGTGGTG. A 500 uL *in vitro* transcription reaction was set up for each PCR product as previously described.[103] The resulting mRNAs were purified by urea-SDS-PAGE and electroelution followed by ethanol precipitation. 500 uL *scale in vitro* translation reactions were conducted with each mRNA template as previously described.[103]

Radiation Spin Assay. A constant concentration of ^{35}S-methionine containing peptide from *in vitro* translation reactions was incubated with different concentrations of GST-$(BRCT)_2$ fusion protein in 200 μL selection buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $MgCl_2$, 0.25% Triton X-100) while tumbling for 5, 20, 60 and 120 minutes at 4°C in a YM-30 filter. After incubation samples were spun at 12200 rpm for 75 seconds until approximately half of the volume had been filtered into the bottom collection chamber. The radioactivity in the top and bottom chamber was determined by removing 75 μL from each. From this the total peptide ("top") and free peptide ("bottom") concentrations were used to calculate the fraction of peptide bound from the formula f_a = [bound peptide]/[total peptide], where [bound peptide] is the concentration of peptide bound to the GST-$(BRCT)_2$ fusion protein and [total peptide] is the total peptide concentration, by substituting top-bottom for [bound peptide] and top for [total peptide].

Peptide synthesis. The peptides were synthesized using a Liberty Automated Microwave Peptide Synthesizer (CEM). The peptides were synthesized on Fmoc-PAL-PEG-PS Resin

(Applied Biosystems) using N-α-Fmoc-protected amino acids (CEM or AAPPTEC) or unnatural N-α-Fmoc protected amino acids (Chem Impex). After each coupling step a capping step was performed using 20% acetic anhydride (Fisher Certified ACS). The peptides were cleaved from the resin using trifluoracetic acid (TFA) (Chem Impex)/TIS (Sigma)/DODT (Sigma)/water (92.5:2.5:2.5:2.5) with incubation at room temperature for 3 hours, and the resin was filtered off. The filtrate containing the crude peptides was precipitated with cold ether, and collected by centrifugation. The supernatant was discarded and the peptide was dissolved in CH_3CN (Fisher HPLC Grade) and water with 10% acetic acid (Fisher certified ACS PLUS) (1:1) followed by freezing and lyophilization. The peptides were then purified by reverse phase HPLC using a Shimadzu Prominence system with a Vydac (218TP C18 5μ) column with 0.1% TFA in water (A) and CH_3CN (B) as the mobile phase with monitoring at 215nm or 264nm. A typical gradient was 10-100% B over 30 minutes, but was adjusted for each peptide. MS data was collected using a Micromass MALDI-R spectrometer.

MALDI-TOF Analysis. HPLC fractions were prepared for analysis by 1:1 dilution in a 1:0.99:0.01 CH_3CN:H_2O:TFA solution containing 10 mg/mL α-cyano-4-hydroxy-cinnamic acid (CHCA). After spotting on the sample plate, samples were allowed to co-crystallize by slow evaporation at rt. Samples resulting from translation reactions were desalted and concentrated with ZipTipC_{18} Pipette Tips (Millipore) according to the manufacturer's protocol.

Peptide Cyclization with α-α'-dibromo-*m*-xylene: A 50 mL oven-dried flask was charged with water (19.34 mL) and acetonitrile (5.28 mL) and was deoxygenated by bubbling Argon for 10 min. Then 200 mM ammonium bicarbonate (3.5 mL, pH 8.4), tris-carboxyethylphosphine (TCEP) (2.0 mg, 7.0 μmol) in 3.5 mL water and peptide (10 mg, 3.5 μmol) were added and the reaction was kept under argon. After 30 min, α-α'-dibromo-*m*-xylene linker (10.2 mg, 38.6

μmol) was added as a solution in 3.5 mL of acetonitrile. The reaction was incubated at rt and monitored by MALDI-TOF. After 2 h, β-mercaptoethanol (BME) (13.7 mg, 176 μmol) or dithiothreitol (DTT) (27.1 mg, 176 μmol) was added to quench unreacted linker, which serendipitously led to formation of peptide-linker-BME (or DTT) adducts upon concentration during lyophilization. Omission of the quench step resulted in the desired α-α'-dibromo-*m*-xylene cyclized peptide product. The reaction was then frozen and lyophilized. The resulting white powder was dissolved in 25% acetonitrile and purified by reverse phase semi-preparative HPLC under the following conditions. Column: Vydac 218TP52210 22 x 100 mm: Flow rate: 10 mL/min: Solvents: A=water/0.1% TFA, B=acetonitrile/0.1% TFA): Gradient: 10 min at 10% B, 30 min at 10-55% B. Injection occurred at 5 min. 4.8 mg (1.63 μmol, 47%) of pure, unquenched product was recovered.

Determination of Protein and Peptide Concentrations. All protein and peptide concentrations were determined using their UV absorbance at 280 nm according to the method of Gill and von Hippel [126]. Extinction coefficients: GST-$(BRCT)_2$ = 76810 $M^{-1}cm^{-1}$; TR-$(BRCT)_2$ = 50070 $M^{-1}cm^{-1}$; FAM = 75800 $M^{-1}cm^{-1}$; Peptides = 1280 $M^{-1}cm^{-1}$ (each contains a single Tyr).

Surface Plasmon Resonance. All experiments were performed at 25 °C on a Biacore 2000 instrument (GE Healthcare). Immobilization of either GST-$(BRCT)_2$ fusion or α-GST antibody was achieved via amine coupling to a CM5 chip (GE Healthcare) per the manufacturer's instructions (Amine coupling kit, GST capture kit, GE Healthcare). Peptides were diluted in HBS-EP running buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% Surfactant P20) with the adjusted NaCl concentrations as indicated.

Isothermal Titration Calorimetry. ITC experiments were performed on an ITC200 calorimeter

(Microcal Inc.) in 300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4. Experiments were carried out by titration of 200-400 µM peptide in the syringe into 20-30 µM TR-BRCT protein in the sample cell. The cell was thermostated at 25 °C and the syringe was stirred at 400 rpm. Each titration consisted of 20 injections of 2 µl except for the first addition which was only 0.5 µL. This first data point was deleted prior to data analysis. Control experiments consisted of titration of peptide into buffer alone to determine the heat of dilution which was subtracted from the data collected from the peptide into protein titration. The resulting data was fitted using a one-set binding-site model analysis using Origin 7 software (Microcal Inc.) to obtain binding stoichiometry (N), association constant (Ka), change in enthalpy (ΔH). ΔG and ΔS were calculated with the following equations: $\Delta G = -RT*\ln(K_a)$ and $\Delta S = (\Delta H - \Delta G)/T$.

Cloning and Protein Expression. To prepare an N-terminal thioredoxin fusion with C-terminal His-tag (TR-BRCT). The pGEX-4T-1 vector containing the BRCA1 $(BRCT)_2$ domain (amino acids 1646-1859)[119] previously described (Chapter 3) was digested with BamHI and XhoI and ligated into the modified pET32a.[127] After transformation into Rosetta 2(DE3) strain of *E. coli*, cells were induced using IPTG and expressed overnight at 18 °C. The recombinant protein was purified on FPLC using nickel affinity chromatography with Chelating Sepharose Fast Flow resin (GE Healthcare) and elution with a gradient of imidazole (50 mM phosphate pH 8.0, 300 mM NaCl, 25 mM imidazole wash, 250 mM imidazole elute) followed by chromatographic isolation by size exclusion chromatography with HiLoad 26/60 Superdex 75 prep grade resin (GE healthcare) in Buffer A (300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4). The resulting fusion proteins had a yield of 20 mg/L and were >95% pure as estimated by SDS–PAGE analysis. Proteins were stored at -80 °C after dialysis in 1L enzyme storage buffer (50 mM HEPES-KOH pH 7.6. 100 mM KCl, 10 mM MgCl2, 7 mM BME, 30 % glycerol).

Determination of IC_{50} with Fluorescence Polarization. The indicated amounts of competing peptide were added to a total of 150 µL containing 2.5 µM TR-$(BRCT)_2$ and 20 nM FAM-β-Ala-pSPTF in 300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4. The samples were incubated at room temperature for 2 hours. The samples were then analyzed in a 100 µL quartz fluorescence cuvette using a Cary Eclipse Fluorescence Spectrophotometer (Varian Instruments) with polarizing lens attachment. The G-factor was experimentally determined with sample containing 20 nM FAM-β-Ala-pSPTF in the same buffer. Anisotropy values were automatically calculated by the accompanying Advance Reads software. After plotting anisotropy versus peptide concentration, an IC_{50} was determined as the peptide concentration at which 50% of the FAM-β-Ala-pSPTF was bound by fitting to the four parameter logistic equation using SigmaPlot 11.

4.9 Summary

On the road to ranking the top 7 hits from the library selection many technical hurdles were encountered. After attempting radiation spin assay, SPR, and ITC, finally FP was used to rank the peptides in a competition assay with a phopsphoserine containing peptide known to bind to the BRCA1 $(BRCT)_2$ domain.

Future Directions: Now that an initial ranking has been achieved, there are many more questions remaining with regard to the highest affinity peptide: linear peptide 8.6. Whether or not the full length of the peptide is necessary will be investigated as well as mutational analysis of individual amino acids. The presence of an E-X-X-F motif of this peptide suggests it may

bind in the same pockets as the pS-X-X-F motif, and the importance of these amino acids can be partially addressed through alanine mutations.

CHAPTER 5. WHERE'S THE PHOSPHATE?: INVESTIGATION OF PEPTIDE 8.6 BINDING MECHANISM THROUGH MUTATIONAL ANALYSIS

5.1 Introduction

The ultimate goal for this project was to discover a non-phosphorylated peptide that binds to the $(BRCT)_2$ domain of BRCA1 in the same location as proteins known to interact with this domain. The presence of our D/E-X-X-F motif in each peptide is an encouraging indication that the selected peptides are in fact competitive inhibitors due to the analogous pS-X-X-F motif that is currently thought to be necessary for proteins binding to this domain. Presence of a similar motif does not guarantee binding and needs further testing to support this hypothesis. The focus of this chapter is mutational studies with the aim of understanding how the highest affinity peptide 8.6 is binding to the $(BRCT)_2$ domain, as well as attempts to improve up on the affinity of this peptide.

5.2 Examination of Peptide 8.6 Truncation

With linear peptide 8.6 being the highest affinity peptide from selection it was decided to focus all further efforts on this peptide. A series of truncated peptides were made and their affinities were compared to linear peptide 8.6 via the same fluorescence polarization assay. The results from these studies can be found in Table 5.1. It was first investigated to what extent the putative core amino acid, DEYRF and FDEYRF could bind to the $(BRCT)_2$ domain, but these peptides had little to no binding to the protein. This provides evidence that amino acids outside the core tetrapeptide provide significant binding interactions, which is in contrast to previous phosphoserine-containing peptide binders.[109] To determine the contribution of both termini of the peptide we made a few truncation mutants. Further truncation mutants showed that removal of the C-terminal RKT led to a small 2-fold increase in IC_{50} while N-terminal truncation led to a

Peptide	Sequence	IC_{50} (μM)
8.6	MCTIDFDEYRFRKT	10.5 ± 1.0
8.6 7-11	DEYRF	>500
8.6 6-11	FDEYRF	>500
8.6 1-11	MCTIDFDEYRF	23.8 ± 1.9
8.6 7-14	DEYRFRKT	274 ± 59
-MC	TIDFDEYRFRKT	25.5 ± 5.4

Table 5.1 Truncation Analysis of Peptide 8.6. Peptide 8.6 was incrementally truncated to investigate whether the full length of the peptide was necessary to achieve its highest binding affinity. These peptides were tested via an FP competition assay and the IC_{50}s resulting are shown here. 4-F-phenylalanine is indicated by red lettering.

25-fold increase. Therefore the N-terminal 6 amino acids are more important for binding than the C-terminal "RKT." Another peptide was made to investigate the presence of the N-terminal fixed

Met-Cys. As part of the non-random region, this was not optimized from the library other than its position relative to the D/E-X-X-F motif. Removal of these amino acids resulted in a two-fold decrease in affinity which is the same effect as "blocking" these amino acids with cyclization. This was yet another indication that the full length is necessary to achieve the highest affinity.

5.3 Mutational Analysis of Peptide 8.6

For results of FP analysis, see Table 5.2. To investigate whether the EYRF motif is binding in the same manner as the pS-X-X-F motif in other known peptides we mutated the Phe to Ala (F11A) and changed the Glu to Ala (E8A). Both mutants showed a dramatic increase in the IC_{50} value, highlighting their importance for binding. It is interesting that the individual amino acids in this motif are necessary yet not sufficient for binding.

If the elimination of the glutamic acid had such a dramatic decrease in binding, then what effect would the incorporation of a phosphoserine have? In previous studies the mutation of a phosphoserine to the phosphomimetic amino acids glutamic and aspartic acid has led to a several order of magnitude decrease in binding if not complete elimination of binding.[128] Because of these results we were surprised that 8.6 E8pS peptide had only a three-fold increase in binding affinity. Because of the small increase, it was hypothesized that the presence of phosphoserine perhaps changed the mechanism of binding. The truncation studies showed that the N-terminal MCTID was important for peptide binding, and this same truncated peptide with an E→pS

Peptide	Sequence	IC_{50} (μM)
8.6	MCTIDFDEYRFRKT	10.5 ± 1.0
8.6 F11A	MCTIDFDEYRARKT	356 ± 35
8.6 E8pS	MCTIDFD$_p$SYRFRKT	3.45 ± 0.34
8.6 E8A	MCTIDFDAYRFRKT	> 300
8.6 E8pS 8-14	$_p$SYRFRKT	2.73 ± 0.09
8.6 All Natural	MCTIDFDEYRFRKT	4.54 ± 0.22

Table 5.2 Mutational Analysis of Peptide 8.6. In attempt to investigate whether peptide 8.6 binds in the same manner as native (BRCT)2 domain binders, peptides with key mutations were analyzed via an FP competition assay. The results are shown here. 4-F-phenylalanine is indicated by red lettering.

substitution had an even lower IC_{50} than the full length 8.6 E8pS peptide. Could the presence of the phosphate change the way this peptide binds? Additionally an all-natural (AN) version of Peptide 8.6 was tested to probe the effect of the 4-fluoro-Phe. Unfortunately canavanine was unable to be incorporated in these peptides, so its importance is currently unable to be examined. Some previous investigations of other hydrophobic amino acid substitutions for phenylalalanine in the pS-X-X-F motif indicate that this hydrophibic binding pocket is quite selective and no analogs were found to be superior to phenylalanine.[128b] Although none of the phenylalanine analogs tested was as small of a modification as a single fluorine, the FP data from AN 8.6 indicates elimination of the two fluorines resulted in better binding.

5.4 Hybridization of Peptides 8.1 and 8.6

Given that the random region is only twelve amino acids in length, one can't help but to wonder what might have been found if the random region had been longer. The study of truncated peptide 8.6 seems to indicate than the extended linear confirmation is necessary for high affinity binding. Not every peptide has the same placement of the D/E-X-X-F domain, so is it possible to generate a higher affinity peptide than any found in selection by hybridizing our highest affinity peptides through mixing and matching those amino acids outside the core domain?

Peptides 8.1 and 8.6 are the two highest affinity peptides from the selection and these peptides have significantly different placement of the E/D-X-X-F domain. Combining the long N-terminal portion and core of peptide 8.6 that has shown to be important with the long C-terminal portion of peptide 8.1 did not result in a peptide with higher affinity than peptide 8.6

(See Table 5.3). Although a disappointing result, it also seems likely that the length of this peptide (18 amino acids) may create additional contacts, but also increases the entropic cost paid in binding. Entropic cost in mind, another hybrid peptide was created where the shorter N-terminus of peptide 8.1 was combined with the core and shorter C-terminus of peptide 8.6. Since the N-terminus of peptide 8.6 was shown to be critical it was surprising that this short peptide was only weaker by a factor of 2-3. This may suggest that the N-terminal MCN of peptide 8.1 makes important contacts with the $(BRCT)_2$ domain. In addition to the D/E-X-X-F motif, the majority of the peptide families also have a threonine somewhere in the C-terminal sequence. Could the hybridization of the N-terminal portion of peptides 8.1 and 8.6 result in a higher affinity peptide? The N-terminal RKT was added to the end of hybrid peptide 2, but this actually resulted in a slight decrease in binding affinity rather than an increase.

Mixing and matching of peptide pieces outside the core domain did not prove successful in producing higher affinity peptides, but questions remained about the effect of amino acids in the core domain. Instead of substituting any individual core domain for another, these were investigated incrementally. Comparisons were again made between peptides 8.6 and 8.1 because they were the highest affinity peptides, but peptide 8.1 also contains the core domain that occurred the largest number of times in selection since the DFIF motif found in 8.1 is also the core domain of 8.2. First, an E8D mutation was made of peptide 8.6 to compare the two phosphoserine analogs, but this mutation was not beneficial indicating preference of the glutamic acid. The difference between the core domains primarily lies in the amino acids selected for the "X-X" portion of the core domain. A substitution of the FI from peptide 8.1 for the YR in peptide 8.6 also had a minimally negative effect providing further evidence that the core domains are not completely interchangeable. Another interesting observation in the sequence of peptide

Peptide	Sequence	IC_{50} (μM)
8.6 All Natural	MCTIDFDEYRFRKT	4.5 ± 0.2
8.1, 8.6 Hybrid	MCTIDFDEYRF**RRSTFRA**	18.6 ± 0.8
8.6 E8D	MCTIDFD**D**YRFRKT	9.6 ± 0.5
8.6 Y9F, R10I	MCTIDFDE**FI**FRKT	7.0 ± 0.4
8.6 D7N	MCTIDF**N**EYRFRKT	5.6 ± 0.1
8.1, 8.6 Hybrid 2	MCNEYRFRKT	12.7 ± 2.0
8.1, 8.6 Hybrid 3	MCNEYRFRKTFRA	14.5 ± 2.3

Table 5.3 Analysis of Hybridized Peptides. Six peptides were synthesized that were hybridized sequences of peptide 8.1 and 8.6. These peptides were subjected to an FP competition assay and the resulting IC_{50} are shown in this table.

8.1 is the presence of an asparagine -1 from the DFIF. This D7N mutation was tested and was found to be a nearly neutral mutation. Although the D7N mutation may not have made an improvement of peptide affinity, but it did show an improvement in solubility, which may be useful in later studies.

5.5 Examination of Highest Affinity Peptides by ITC

Ranking peptides via their IC_{50} values proved useful for comparison between peptides; however, to compare these peptides to other known binders determination of a K_d is necessary. Another point to consider is that it is well known that IC_{50} values approaching the K_d of a fluorescent ligand (in our case 2.3 μM)[108] are obscured by the fluorescent ligand's intrinsic affinity,[129] so comparisons between ligands with IC_{50} values approaching the K_d of the fluorescent ligand are not reliably quantitative.

Further characterization of peptides (8.6, AN, 8.6 E8pS, and 8.6 8-11 8EpS) by ITC was conducted (Figure 5.1). The relative ranking of the binding affinity of the three peptides is roughly the same as that determined by FP (E8pS > AN > 8.6). The relative entropic and enthalpic contributions to binding were also determined (Table 5.4, Figure 5.2). The binding of all three peptides is driven by large, negative ΔH values. Surprisingly, the glutamic acid-containing peptides from our libraries have more negative ΔH values than E8pS. One potential rationale for this increase in ΔH is that the glutamic acid containing peptides form additional contacts (presumably at the N-terminus) with the protein which result in the more favorable value. This pattern holds true except for the E8pS 8-14 peptide that actually has a larger negative ΔH than the full length E8pS peptide.

The binding of all four peptides have negative ΔS values. The entropic term of binding is

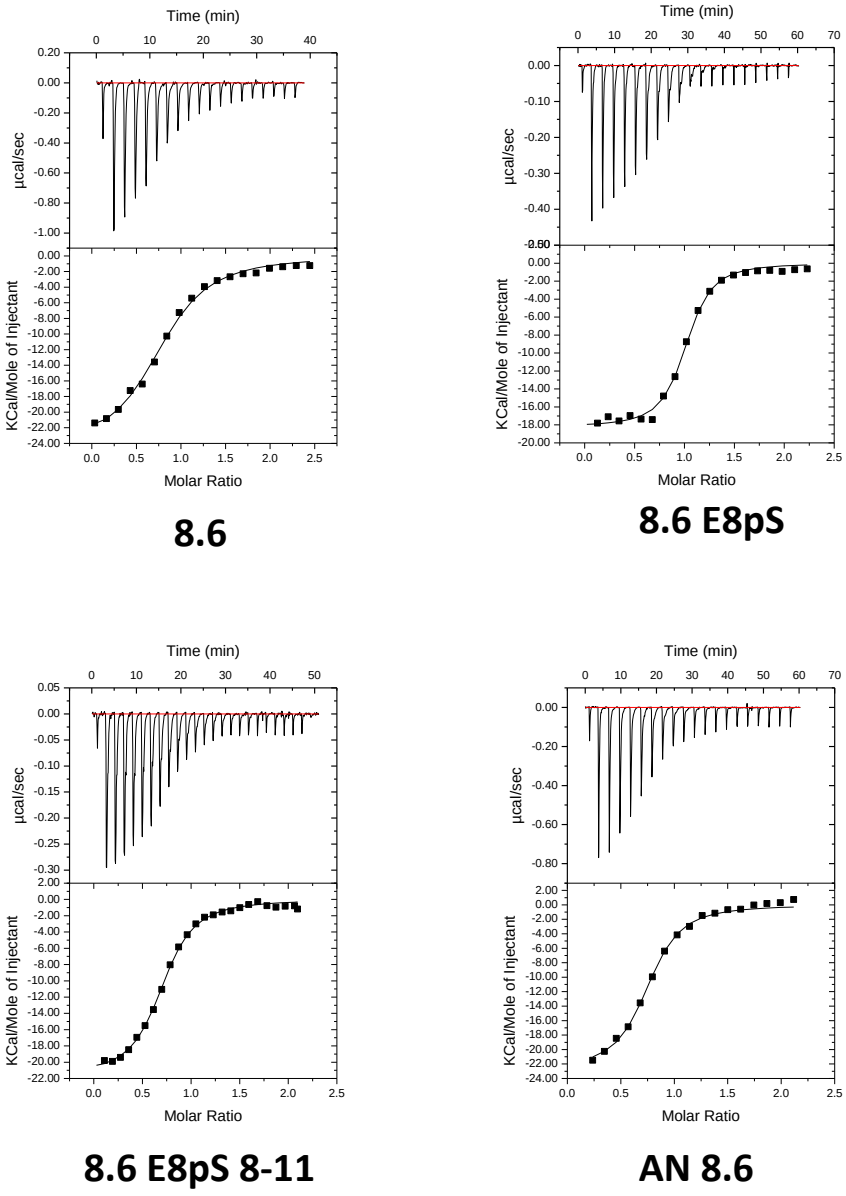


Figure 5.1 ITC Analysis of Highest Affinity Peptides. The curves shown above are the ITC curves collected for four of the highest affinity peptides uncovered from selection and mutation of peptide 8.6.

Peptide	N (Sites)	K (M^{-1})	ΔH (kcal/mol)	ΔG (kcal/mol)	ΔS (cal/mol/deg)	ITC: 1/K (μM)
8.6	0.80 ± 0.01	$2.71 \cdot 10^5 \pm 2.9 \cdot 10^4$	-24.6 ± 0.6	-7.40 ± 0.06	-57.7 ± 2.0	3.69 ± 0.11
8.6 E8pS	0.98 ± 0.01	$2.44 \cdot 10^6 \pm 3.3 \cdot 10^5$	-18.6 ± 0.3	-8.70 ± 0.08	-33.0 ± 1.0	0.41 ± 0.14
8.6 E8pS 8-11	0.705 ± 0.007	$1.14 \cdot 10^6 \pm 8.6 \cdot 10^4$	-21.6 ± 0.3	-8.25 ± 0.04	-44.8 ± 1.1	0.877 ± 0.066
8.6 All Natural	0.73 ± 0.01	$9.95 \cdot 10^5 \pm 1.3 \cdot 10^4$	-22.7 ± 0.5	-8.17 ± 0.01	-48.7 ± 1.7	1.01 ± 0.01

Table 5.4 Thermodynamic Values from ITC Analysis. The four highest affinity peptides excluding the hybrid mutants were analyzed by ITC. The values for N, K, ΔH and ΔS were obtained from analysis of the binding curves. The values for ΔG and the binding affinity (1/K) were calculated.

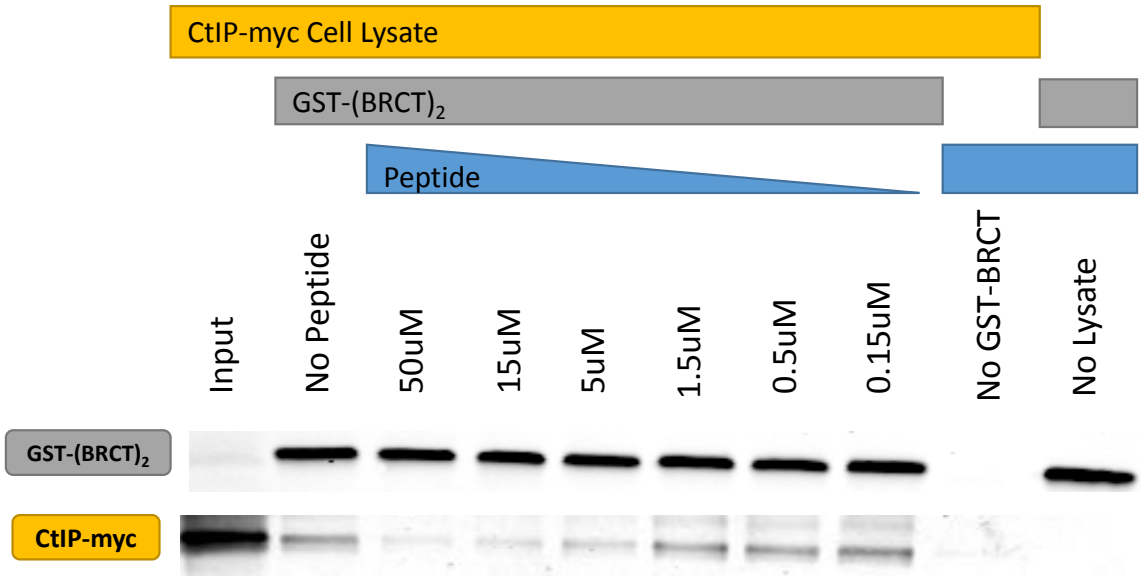


Figure 5.2 Inhibition of $(BRCT)_2$ and CtIP in Interaction in Cell Lysate. A pull down study was conducted in which CtIP-myc was over expressed in cell lysate and bound to immobilized GST-(BRCT)2 on agarose. This interaction was inhibited with increasing concentrations of AN peptide 8.6.

7

correspondingly less favorable for the peptides containing Glu vs. pSer except in the E8pS 8-14 peptide. Before testing of the truncated pSer peptide, it was hypothesized that the Glu peptides are gaining additional N-terminal contacts not found in E8pS. These additional contacts would be thought to increase the entropic cost of binding; however, this model does not fit given the increased entropy of binding found in the shorter E8pS 8-14 peptide.

Finally, it is interesting that the replacement of two fluorines with two hydrogens (8.6 to AN) results in a significant increase in ΔS. This is the opposite of what is typically observed with fluorine substitutions, since the more hydrophobic fluorinated molecules typically have more favorable entropy due to desolvation of their increased hydrophobic surface.[130] Although this could be explained by different solution structure preferences of the two peptides, we think it is more likely that the increase in entropy results from more flexibility in the binding of Phe vs. the slightly more sterically demanding 4-F-Phe. We are currently pursuing crystallographic studies to investigate these hypotheses.

5.6 Inhibition of CtIP-BRCA1 Interaction in Cell Lysate

In order to assess the ability of the highest affinity non-phosphorylated peptide (AN) to inhibit protein-protein interactions in cell lysates, a known BRCA1 $(BRCT)_2$ binding partner phospho-CtIP-myc[131] was overexpressed in 293T cells and captured with purified GST-$(BRCT)_2$ fusion in the presence of varying concentrations of the peptide. Proteins were separated by SDS-PAGE and analysed via western blotting (Figure 5.3). A dose-dependent inhibition of the CtIP-BRCT complex was observed, resulting in near complete abrogation of this interaction at 50 μM.

5.7 Discussion

Mutational Analysis of Peptide 8.6. Truncation of peptide 8.6 down to its core DEYRF domain resulted in nearly complete abolishment of binding. No truncation mutations were found that maintained an IC_{50} as low as the full length peptide, but deletions of the N-terminal and C-terminal regions outside of the core domain showed that overall the N-terminal region of the peptide plays a large role in the interaction. The pSer and Phe of native peptides are generally regarded as being necessary and sufficient for binding to the $(BRCT)_2$ domain. Therefore it stands to reason that if our D/E-X-X-F peptides are biding in the same pockets, elimination of the key E and F via alanine mutation might also abrogate the interaction.

Despite the inability of the core domain to bind alone, these alanine mutations had a significant impact on binding affinity indicating they are playing a very important role. It is interesting that the E8pS mutant did not have a more dramatic increase in binding affinity. We also attempted to generate a higher affinity peptide by recombining the highest affinity sequences from selection. Although none of these attempts generated a higher affinity peptide, it is interesting to note that the point mutations had much less of an effect on binding that any attempts to recombine whole sections of the sequences.

ITC analysis showed that AN 8.6 had a K_d of 1 μM making this the tightest non-phosphorylated $(BRCT)_2$ domain binder discovered to date. The thermodynamic information from the ITC analysis raised more questions than it answered about the potential binding mechanism of these peptides. Hopefully these questions can be answered with the crystallographic studies that are underway.

Comparison to current inhibitors. Our peptide may not be the tightest-binding molecules discovered to date that bind to the $(BRCT)_2$ domain, but there is much more to drug design than affinity. A molecule without a phosphate is arguably much more cell permeable and stable, which is a significant advance in the quest for an inhibitor of BRCA1.[79] Small molecule screens have resulted in peptides in the mid-micromolar range, yet despite fewer drug delivery issues with small molecules they were unable to effectively inhibit BRCA1 function.[132] With small molecules seemingly unable to do the job, many phosphopeptides have been tested that are higher affinity that the previous small molecules, and have been shown to inhibit $(BRCT)_2$ binding proteins in cell lysate. Translation of these effects into cells has been much more of a challenge. Even the highest affinity peptide required high (100 μM) concentration and had minimal inhibitory function. The trade-off in the history of BRCA1 $(BRCT)_2$ inhibition seems to have been presence of a phosphate or an inhibitor. Our highest affinity peptide is not close to the 40 nM affinity achieved with phosphopeptides, but how tight does the affinity need to be?

Comparison to Native $(BRCT)_2$ binding partners. In the development of inhibitors it is usually the goal to develop a molecule that will bind to the desired site with the tightest affinity possible. Affinity is important, but so is cell permeability and stability. A lower affinity inhibitor may be more affective in the end depending on the system. There are several known proteins that bind to the $(BRCT)_2$ domain, and of these the proteins BACH1 and CtIP currently have the highest affinity (see Table 5.5). Several studies have been conducted reporting the binding affinity via ITC for these interactions. For BACH1 reports have ranged from 0.9 to 0.17 μM and for CtIP reports are a bit higher ranging from 3.7 to 1.32 μM. Thus our non-phosphorylated pepeptide is in the same range as the natural protein binders.

Table 5.5 Peptides Known to Bind to the $(BRCT)_2$ Domain. This table is a list of peptides known to interact with a reasonably high affinity with the $(BRCT)_2$ domain of BRCA1.

Peptide	Sequence	Affinity (μM)	Reference
p-ACC1	DSPPQpSPTFPEAGH	5.2	[133]
1	Ac-pSPTF-$CONH_2$	2.51	[128b]
CtIP	PTRVpSPVFGAT	1.32-3.7	[134]
7	Ac-pSPVF-$CONH_2$	1.62	[128b]
AN Peptide 8.6	MCTIDFDEYRFRKT	1.0	
BACH1	ISRSTpSPTFNKQ	0.17-0.9	[134a, 135]
18	Ac-pSPTF-COOH	0.19	[128b]
19	Ac-pSPVF-COOH	0.29	[128b]
15	*pSPVF-COOH	0.04	48a

* N-terminal constrained by a 3-carbon linker with a phenyl ring.[134, 135c] These affinities are quite similar to the peptides discovered with this selection. Although a higher affinity inhibitor is desirable, and will be a future direction for this work, given that the affinities of BACH1 and CtIP for this domain are in the low micromolar and high nanomolar range, our peptide will likely have a significant functional improvement with advancements in its current affinity.

Functional Analysis of Peptide 8.6. Perhaps the most important result from these studies, is the initial demonstration that AN 8.6 can function as an inhibitor of BRCA1 $(BRCT)_2$ interactions in cell lysates in a dose dependent manner. Although this study does address protease resistance of the peptide due to the presence of protease inhibitors in the assay, it does lend support to the specificity of the inhibitor. If non-specific binding were to be significant, much larger quantities of the peptide would be needed to inhibit the $(BRCT)_2$-CtIP interaction. As it stands here the concentration needed to significantly inhibit the interaction approximates the K_d measured by ITC.

It is unlikely that we would have been able to find peptide 8.6 if we had made a library of 8 randomized AAs since the N-terminus through AA F11 are required for binding. However, the extreme diversity of our peptide libraries has allowed us to find tight-binding non-phosphorylated inhibitors when other library approaches failed. Given enough diversity, binders to challenging targets can be uncovered, including those previously requiring phosphoserine.

5.8 Experimental

Peptide synthesis. The peptides were synthesized using a Liberty Automated Microwave Peptide Synthesizer (CEM). The peptides were synthesized on Fmoc-PAL-PEG-PS Resin (Applied Biosystems) using N-α-Fmoc-protected amino acids (CEM or AAPPTEC) or unnatural N-α-Fmoc protected amino acids (Chem Impex). After each coupling step a capping step was performed using 20% acetic anhydride (Fisher Certified ACS). The peptides were cleaved from the resin using trifluoracetic acid (TFA) (Chem Impex)/TIS (Sigma)/DODT (Sigma)/water (92.5:2.5:2.5:2.5) with incubation at room temperature for 3 hours, and the resin was filtered off. The filtrate containing the crude peptides was precipitated with cold ether and collected by centrifugation. The supernatant was discarded and the peptide was dissolved in CH_3CN (Fisher HPLC Grade) and water with 10% acetic acid (Fisher certified ACS PLUS) (1:1) followed by freezing and lyophilization. The peptides were then purified by reverse phase HPLC using a Shimadzu Prominence system with a Vydac (218TP C18 5μ) column with 0.1% TFA in water (A) and CH_3CN (B) as the mobile phase with monitoring at 215nm or 264nm. A typical gradient was 10-100% B over 30 minutes, but was adjusted for each peptide. MS data was collected using a Micromass MALDI-R spectrometer.

Determination of Protein and Peptide Concentrations. All protein and peptide concentrations were determined using their UV absorbance at 280 nm according to the method of Gill and von Hippel [126]. Extinction coefficients: GST-$(BRCT)_2$ = 76810 $M^{-1}cm^{-1}$; TR-$(BRCT)_2$ = 50070 $M^{-1}cm^{-1}$; FAM = 75800 $M^{-1}cm^{-1}$; Peptides = 1280 $M^{-1}cm^{-1}$ (for peptides containing a single Tyr); 4-F-Phe = 500 $M^{-1}cm^{-1}$ (for peptides with no Tyr, but contain at least one 4-F-Phe).

Determination of IC_{50} with Fluorescence Polarization. The indicated amounts of competing peptide were added to a total of 150 µL containing 2.5 µM TR-$(BRCT)_2$ and 20 nM FAM-β-Ala-pSPTF in 300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4. The samples were incubated at room temperature for 2 hours. The samples were then analyzed in a 100 µL quartz fluorescence cuvette using a Cary Eclipse Fluorescence Spectrophotometer (Varian Instruments) with polarizing lens attachment. The G-factor was experimentally determined with sample containing 20 nM FAM-β-Ala-pSPTF in the same buffer. Anisotropy values were automatically calculated by the accompanying Advance Reads software. After plotting anisotropy versus peptide concentration, an IC50 was determined as the peptide concentration at which 50% of the FAM-β-Ala-pSPTF was bound by fitting to the four parameter logistic equation using SigmaPlot 11.

Isothermal Titration Calorimetry. ITC experiments were performed on an ITC-200 calorimeter (Microcal Inc.) in 300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4. Experiments were carried out by titration of 200-400 µM peptide in the syringe into 20-30 µM TR-BRCT protein in the sample cell. The cell was thermostated at 25 °C and the syringe was stirred at 400 rpm. Each titration consisted of 20 injections of 2 µl except for the first addition which was only 0.5 µL. This first data point was deleted prior to data analysis. Control experiments consisted of titration of peptide into buffer alone to determine the heat of dilution

which was subtracted from the data collected from the peptide into protein titration. The resulting data was fitted using a one-set binding-site model analysis using Origin 7 software (Microcal Inc.) to obtain binding stoichiometry (N), association constant (Ka), change in enthalpy (ΔH). ΔG and ΔS were calculated with the following equations: $\Delta G = -RT*\ln(Ka)$ and $\Delta S = (\Delta H - \Delta G)/T$.

Preparation of Cell Lysates. HEK293T cells were cultured at 37 °C in DMEM (Gibco) medium supplemented with 10% FBS and 1% pen-strep. The CtIP-myc expressing pcDNA3-5X vector was as described. [117] Five million 293T cells were transfected with 5 µg plasmid using an Amaxa Nucleofector II (Lonza) according to the manufacturer's protocol (program A-023). 5-days post-transfection, total cell lysates were prepared in 1.5 mL RIPA buffer (10mM Tris pH 7, 1% Triton-X, 0.1% SDS, 1 mM EDTA, 0.5 mM EGTA, 140 mM NaCl, 0.1% sodium deoxycholate) containing 1:100 dilutions of Phosphatase Inhibitor Cocktails 2 and 3 (Sigma) and 1:100 dilution of Protease Inhibitor Cocktail for use with mammalian cell and tissue extracts (Sigma). Cell lysates were used immediately for pull-down experiments.

Pulldown-Western. 40 µL of the commercial Glutathione magnetic bead suspension (Pierce) were washed three times with 200 µL beads wash buffer (125 mM Tris-HCl pH 8, 150 mM NaCl). To the washed beads, 200 µL of 250 nM GST-$(BRCT)_2$ fusion protein was added and incubated while tumbling at 4 °C overnight. Beads were then washed three times with 300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4, and 100 µL of peptide solution in the same buffer was added. The samples were incubated at 4 °C for 1.5 hours. To these samples 100 µL of cell lysate was added and incubation continued for another hour. The magnetic beads were washed three times with RIPA with inhibitors, and boiled for 10 minutes in 1x Laemmli buffer (BioRad) followed by separation on 10% SDS-PAGE in running buffer (25 mM Tris, 192 mM

glycine, 0.1% SDS) and Western blotting.

Western Blot Analysis. After separation by SDS-PAGE, proteins were transferred to a polyvinyl difluoride membrane (PVDF, 0.45 mm, Immobilon-FL, Millipore) by wet transfer at 0.6A for 2 h with transfer buffer (25 mM Tris, 192 mM glycine). Subsequently, the PVDF membranes were blocked in blocking buffer (0.1X PBS and 0.1 % casein, BioRad) for 1 hour at room temperature with gentle agitation. The primary antibodies (mouse anti-myc, Cell Signaling; rabbit anti-GST, Cell Signaling) were diluted 1:1000 in blocking buffer and incubated overnight at 4 °C with gentle agitation. The blot was washed 3 x 15 min at room temperature with PBS-T (1X PBS, 0.2% Tween). Bound antibodies were labeled using secondary antibodies (Invitrogen goat anti-mouse AlexaFluor 680, Rockland goat anti-rabbit Dylight 800) in a concentration of 1:2000 in blocking buffer. After incubation for 1 hour at room temperature, three further washing steps of 15 min each at room temperature in PBS-T followed. The protein was imaged and quantified using the Odyssey Infrared Imaging System and application software version 3.0 (Li-Cor Biosciences).

5.9 Summary

None of the truncated version of peptide 8.6 that were tested had a binding affinity lower than that of the full length peptide 8.6. This supports the hypothesis that none of the selected library members had a second cysteine in their sequences due to a preference for an extended linear conformation. Mutational analysis supports the theory that the D/E-X-X-F motif is important for binding and perhaps binds in the same site as the known pS-X-X-F motif. Attempts to hybridize peptides 8.6 and 8.1 did not result in discovery of a higher affinity peptide, but did uncover multiple nearly neutral mutations. In the end, the all-natural analog of peptide

8.6 (AN 8.6) proved to be the highest affinity peptide found with a K_d of 1 μM. Initial in vitro inhibition of the $(BRCT)_2$ domain CtIP in cell lysate appears to shows that AN 8.6 is the first good, non-phosphorylated inhibitor of BRCA1.

Future Directions: Cell studies with all-natural peptide 8.6 (AN 8.6) are needed to test its potential as an inhibitor of BRCA1. Because this peptide uses natural amino acids it could be over expressed in a cell as an alternative to delivery with a cell penetrating peptide tag or other delivery means. Although, initial in vitro studies show that AN 8.6 can inhibit protein-protein interactions of the $(BRCT)_2$ domain, additional studies are needed to support the binding of this peptide in the same binding cleft. Binding studies with protein mutations would be useful in demonstrating this. A crystal structure of an 8.6 analog and the $(BRCT)_2$ domain is also being pursued and will hopefully provide definitive evidence of how the peptide is binding to the protein surface. This study has also served as a proof of principle for using large mRNA-display libraries for discovery of peptides that bind to a BRCT domain that is traditionally thought only to bind phosphopeptides. This same technique could be applied to the many other BRCT domains as well as other targets to select inhibitors of other protein-protein interactions.

CHAPTER 6. OVERCOMING CURRENT LIMITATIONS:

EXPANSION OF DRaCALA TO PEPTIDES

6.1 Introduction

The most difficult challenge faced in finding a non-phosphorylated peptide inhibitor of the BRCA1 $(BRCT)_2$ domain was actually characterizing the peptides that were discovered in selection. This experience highlighted the shortcomings of many of the standard binding affinity analysis techniques we use. Although many of these techniques have great power, the more complicated a technique is the more prone to technical difficulties it becomes. Work by the Lee lab produced a technique that has been named the differential radial capillary action of ligand assay (DRaCALA**)** which provides a very simple means of characterizing the binding affinity of DNA to protein based simply on their differential diffusion on nitrocellulose.[136] Protein tends to bind non-specifically to nitrocellulose, and the high negative charge of DNA results in its diffusion across the membrane. In the experiments conducted by the Lee lab, radiolabelled DNA in a constant concentration is mixed in samples with varying concentrations of an interacting protein. After spotting only five microliters of each sample onto nitrocellulose, the amount of bound DNA can be ascertained from the diffusion pattern. Any DNA bound to the protein will be observed as a darker inner circle, while unbound DNA is seen as a diffuse outer ring, as seen in Figure 6.1. The fraction of DNA bound can be calculated be the equation

$$F_B = \frac{V_{inner} - \left[A_{inner} \times \frac{(V_{outer} - V_{inner})}{(A_{outer} - A_{outer})}\right]}{V_{outer}} \qquad [6.1]$$

where F_B is the fraction bound, V_{inner} is the volume of the inner circle, V_{outer} is volume of the outer circle, A_{inner} is the area of the inner circle, and A_{outer} is the area of the outer circle.

The simplicity of this technique has quite an appeal. Not only does it not require expensive equipment that requires training and maintenance, but it also uses very small quantities of material for each assay. The authors posit that this technique might be applied to

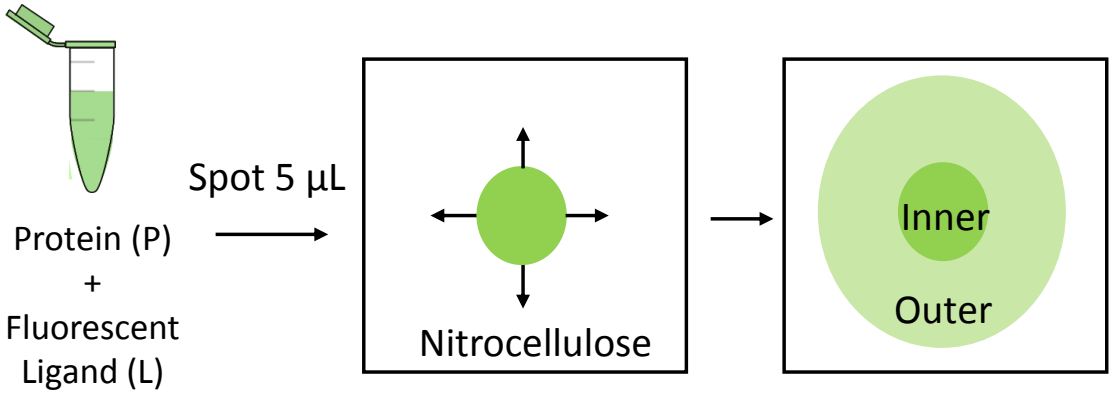


Figure 6.1 DRaCALA Experiment. In the proposed DRaCALA experiments, protein and fluorescent ligand are incubated together in a sample of small volume. 5 µL of this sample is spotted onto a dry, untreated sheet of nitrocellulose membrane. As the sample diffuses from the center spot, the protein interacts strongly with the nitrocellulose and remains in the center. An ideal ligand for this interaction would diffuse uniformly with the solvent away from the center. If the ligand binds to the protein an inner and outer circle will appear, as shown, upon visualization with a fluorescent scanner. From analysis of this pattern a fraction bound can be calculated with Equation 6.1.

small molecules, but with its many advantages, we wondered if it could be applied to peptides. Nitrocellulose binds proteins which are chemically no different than peptides, except of course for their size. It was hypothesized that there was likely some set of peptides that would have sufficient diffusion on nitrocellulose such that a DRaCALA assay could be used to study their binding affinity with proteins.

6.2 Diffusion of Peptides on Nitrocellulose

A set of test peptides was established to test for their ability to diffuse on nitrocellulose. These peptides were designed to test for different characteristics, including charge and length (Table 6.1). Instead of using radiolabelled peptides, each peptide was fluorescently labelled with fluorescein as a means of detection. Each set of peptides was dissolved at various concentrations in PBS and spotted onto nitrocellulose. After drying, each was scanned at the appropriate wavelength to examine the diffusion pattern (Figure 6.2). It was thought that perhaps the longer peptides would not readily diffuse on the nitrocellulose, but it was found that the approximately neutral peptides all displayed a central ring upon spotting which indicates they did not freely diffuse on nitrocellulose. It was however, found, that the small and negatively charged pSPTF peptide did readily diffuse on nitrocellulose at all concentrations making it perhaps suitable for binding affinity analysis with this technique.

6.3 Model Systems for Comparison to Current Techniques

After observing that BRCA1 peptide, FAM-β-A-pSPTF, could in fact diffuse readily across the nitrocellulose even at the highest concentration of 1 μM, it was tested in a binding

Peptide	Label	Sequence	Length (AAs)	Overall Charge
Beta-Amyloid	Rhodamine Green	DAEFRHDSGYEVHHQKLVFFA EDVGSNKGAIIGLMVGGVV	40	0
18mer	FITC	TYSCHFGPLTWVCKPQGG	18	2+
Angiotensin II	FAM	DRVIHPF	7	1+
BRCA1 peptide	FAM	B-A-pSPTF	5	2-

Table 6.1 Peptides Chosen to Test for Diffusion on Nitrocellulose. Peptides of differing lengths were tested for their ability to diffuse onto nitrocellulose with detection via the N-terminal fluorescent tags listed. There sequences are shown and labeled as acidic residues (blue), basic residues (red) and neutral uncharged residues (green).

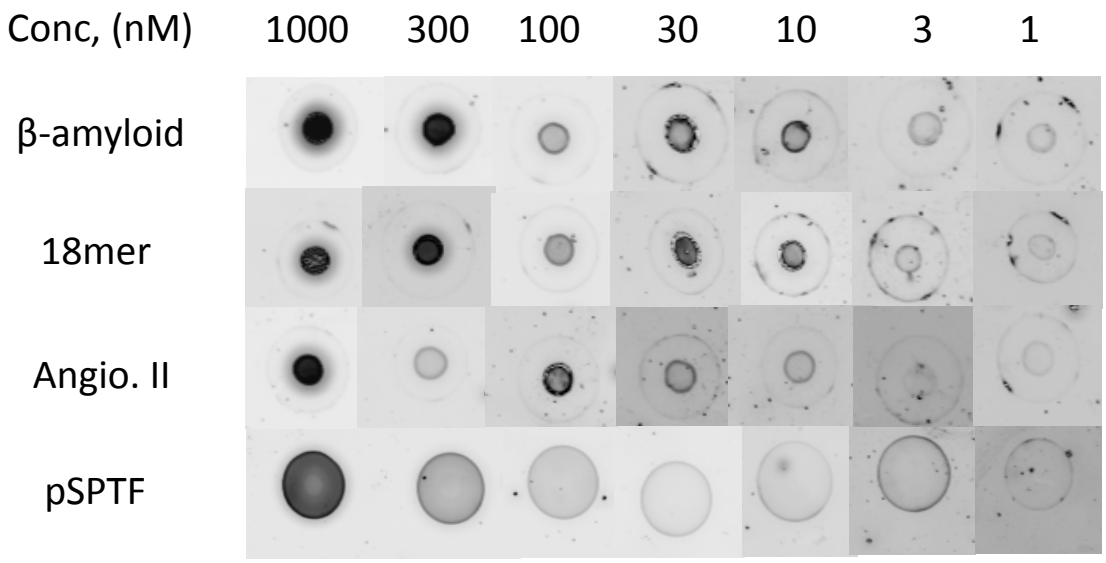


Figure 6.2 Test of Peptide Diffusion on Nitrocellulose. The fluorescently labeled peptides indicated (whose sequences can be found in Table 6.1) were spotted at the concentration indicated onto dry, untreated nitrocellulose and examined on a fluorescence scanner. Each spot is the result of 5 µL of sample.

study with the thioredoxin (TR) fusion of the $(BRCT)_2$ domain. Solutions with 20nM of FAM-β-A-pSPTF and concentrations of TR-$(BRCT)_2$ ranging from 100-0.05 μM were incubated for two hours in buffer. The samples were spotted in six replicates due to irregularities in the samples shape and background interference in some samples. These samples were analyzed on a fluorescence scanner, and the fraction bound for each circle was calculated. The scanned sheet of nitrocellulose and the resulting binding curve from analysis is shown in Figure 6.3. The K_d determined was 4.8 ± 0.1 μM.

6.4 Discussion

Four peptides were spotted at concentrations ranging from 1000-1 nM, and were spotted on nitrocellulose. After drying their diffusion patters were examined with a fluorescence scanner. It is clearly shown in Figure 6.3 that the three longest peptides did not diffuse readily across the nitrocellulose even at the lowest concentrations test. In each spot there is clearly visible both an inner and an outer ring. The fourth and shortest peptide however, shows a single diffuse circle the same size as the outer ring for all the other peptides. This was present at each concentration tested indicating this peptide is clearly diffusing even at the highest concentrations tested. This makes the short, dense negatively charged FAM-β-A-pSPTF peptide potentially suitable for binding affinity analysis via DRaCALA.

DRaCALA is an attractive alternative to FP and ITC in the measurement of the binding affinity of the FAM-β-A-pSPTF peptide. First, the small volume of the DRaCALA assay conserves significant amounts of material. Second, and more importantly, the BRCA1 protein tends to aggregate at higher concentrations. This aggregation leads to light scattering and prevents measurement of the upper bounds of the K_d by FP.

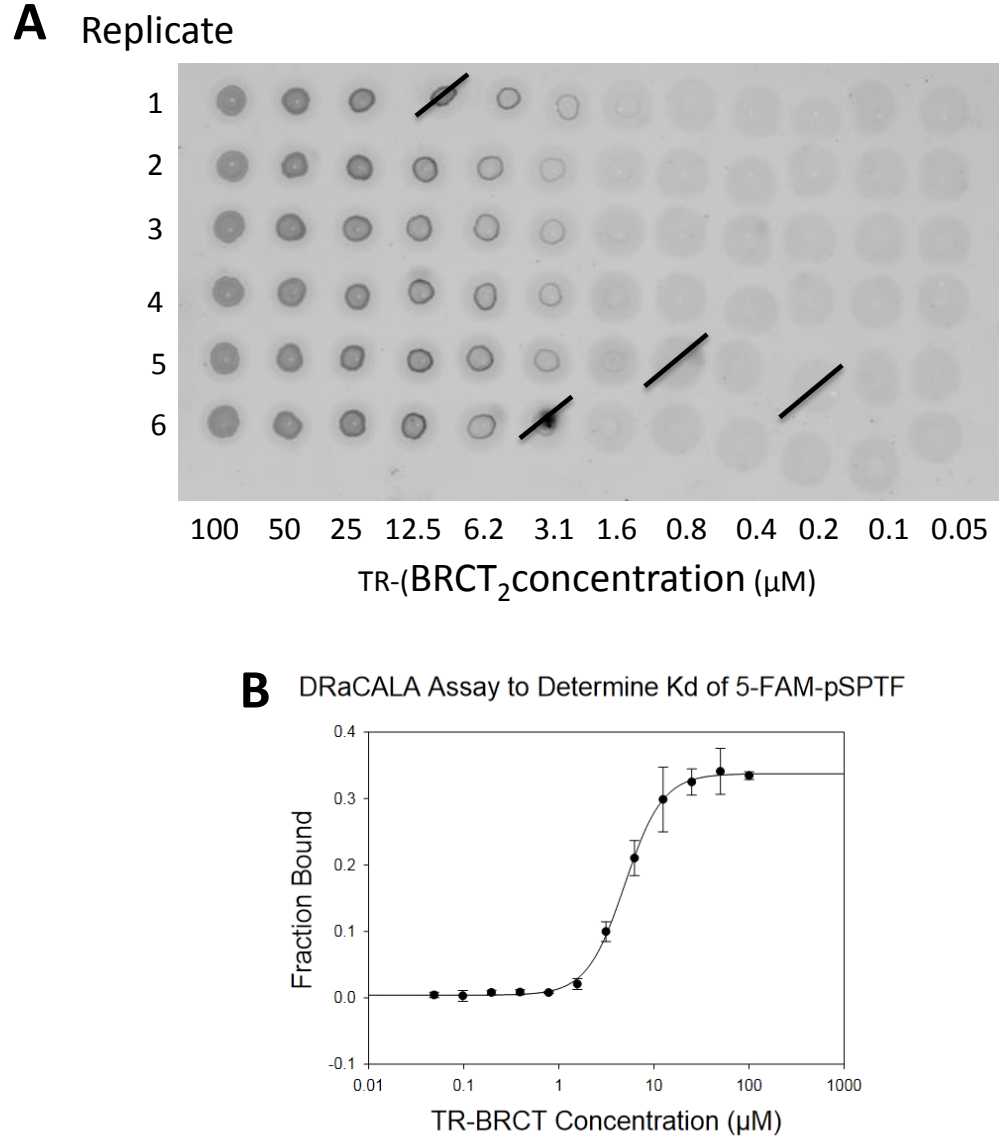


Figure 6.3 DRaCALA of 5-FAM-pSPTF. Analysis of a peptide known to bind to the BRCA1 (BRCT)2 domain, 5-FAM-pSPTF, was conducted by incubating the peptide at 20 nM and the TR-(BRCT)2 fusion protein for two hours. 5 µL of each sample was spotted in six replicates on nitrocellulose A) The resulting spots analyzed by a fluorescent scanner. B) The binding curve resulting from analysis of the spots shown in A.

β-A-pSPTF labeled with FITC has previously been reported to bind to the $(BRCT)_2$ domain of BRCA1 with an affinity of 2.28 ± 0.09 μM.[48a] Testing the affinity of FAM-β-A-pSPTF to TR-$(BRCT)_2$ with DRaCALA produced a binding affinity of 4.8 ± 0.1 μM, which is quite close to the previously reported value, suggesting that in this case DRaCALA is able to measure accurately the binding constant. The difference of these values may owe more to the difference in attached fluorophore or the difference in protein used for the study. It is of course difficult to draw general conclusions about this technique from a single study, however, this bodes well for future uses of this technique to measure the binding affinity of short, negatively charged petpides.

6.5 Experimental

Peptide synthesis. The fluorescently labelled β-amyloid, angiotensin II and 18 mer peptides were purchased from Anaspec, and no further purification was performed. The β-A-pSPTF peptide was synthesized using a Liberty Automated Microwave Peptide Synthesizer (CEM). The peptides were synthesized on Fmoc-PAL-PEG-PS Resin (Applied Biosystems) using N-α-Fmoc-protected amino acids (CEM or AAPPTEC) or unnatural N-α-Fmoc protected amino acids (Chem Impex). After each coupling step a capping step was performed using 20% acetic anhydride (Fisher Certified ACS), but no capping was performed on the final N-terminal amino acid to leave the amine free for labelling. The peptide was cleaved from the resin using trifluoracetic acid (TFA) (Chem Impex)/TIS (Sigma)/DODT (Sigma)/water (92.5:2.5:2.5:2.5) with incubation at room temperature for 3 hours, and the resin was filtered off. The filtrate containing the crude peptide was precipitated with cold ether, and collected by centrifugation. The supernatant was discarded and the peptide was dissolved in CH_3CN (Fisher HPLC Grade)

and water with 10% acetic acid (Fisher certified ACS PLUS) (1:1) followed by freezing and lyophilization. The peptide was then purified by reverse phase HPLC using a Shimadzu Prominence system with a Vydac (218TP C18 5μ) column with 0.1% TFA in water (A) and CH_3CN (B) as the mobile phase with monitoring at 443 nm.

Spotting Peptides on Nitrocellulose. All peptides were dissolved in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4) and diluted to the indicated concentrations. 5 μL was spotted of each was spotted onto dry-untreated nitrocellulose (Bio-Rad Trans Blot Transfer Medium) at least 2 cm apart. Spots were allowed to dry for 20 minutes and scanned on a Typhoon 9410 Variable Mode Imager (Amersham Biosciences) with an excitation wavelength of 488 nm. Resulting images were analysed with ImageQuant 5.1 software.

DRaCALA. Protein was mixed with 20 nM radiolabeled nucleotide in buffer (300 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4) and allowed to incubate for 2 hours at room temperature. These mixtures were pipetted (5 μL) onto dry untreated nitrocellulose (BioRad Trans Blot Transfer Medium) at least 1 cm apart in six replicates and allowed to dry completely before scanning on a Typhoon 9410 Variable Mode Imager (Amersham Biosciences) with an excitation wavelength of 488 nm. Resulting images were analysed with ImageQuant 5.1 software.

6.6 Summary

Four fluorescently labeled peptides were tested for their suitability for use as the ligand in a DRaCALA assay. This was measured simply by spotting 5 μL of serially diluted peptide solutions onto nitrocellulose, and examining the pattern of diffusion on a fluorescence scanner. Only the shortest peptide, which also had the densest negative charge, freely diffused onto the

nitrocellulose. This peptide was used in a DRaCALA assay to determine its binding affinity to the $(BRCT)_2$ domain of BRCA1. A value of 4.9 ± 0.1 μM was obtained, which is quite close to a previous literature report of 2.28 ± 0.09 μM.[48a] This technique will not be generalizable to all peptides, but has shown here to work in special cases.

Future Directions. Although this technique may not be applicable to all peptides, it did help us confirm the affinity of a peptide that was difficult to determine via other methods. This might be a viable option for these special cases, but is not likely suited for hydrophobic and perhaps overall neutral peptides. Preliminary data (not shown) indicates that larger negatively charged peptides diffuse readily on nitrocellulose, and that positively charged peptides like poly-arginine diffuse well on polyethyleneimine (PEI) membranes. Additional study is required to confirm these conclusions.

Overall Summary and Conclusions

Inhibitors of the BRCA1 C-Terminal $(BRCT)_2$ domain have been sought after, yet many of the traditional routes to drug discovery have failed to produce viable inhibitors. Even successful attempts to design peptides that bind to the $(BRCT)_2$ domain have been unable to produce a promising drug candidate due to reliance on the incorporation of a phosphoserine. Peptide libraries have previously proven themselves to be a powerful tools in selecting peptides that bind to protein surfaces, and an important aspect of this has been their large diversity which adds great power to these techniques. Can this technique be used to overcome the phosphate "requirement?" We proposed that the large diversity ($>10^{13}$ members) of an mRNA display library that does not contain phospho-amino acids could be applied to the selection of a non-phsophorylated BRCA1 inhibitor.

In addition to the large number of library members, unnatural amino acids (UNAAs) were also used in an attempt to increase the chemical diversity of the library. This was accomplished via substation of some of the 20 naturally occurring amino acids with unnatural analogues. This is possible because these libraries are constructed within a PURE translation system that allows for control of the in vitro translation reaction components. By simply leaving out a natural amino acid, we can simply substitute a similar UNAAs that will be incorporated by the translation machinery instead. Six UNAAs were found that work well together and were used in translation of the peptide library.

The library peptides consisted of a fixed Met-Cys followed by a 12-amino acid random region encoded by an NNB codon (where N = A, C, T, or G and B = C, T, or G). This decreases the number of stop codons, as well as enriches the number of cysteines found in the random

region so that 40% of the library members should have a second cysteine. This is important as the library will be subject to cyclization with dibromoxylene that will act as a covalent linker selectively between two cysteines. This will hopefully result in peptides with a more rigid peptide scaffolds that may lead to tighter peptide binding due to decreased entropic cost in binding.

The selection was conducted over eight rounds against GST-$(BRCT)_2$ fusion protein immobilized on magnetic glutathione beads. In all but the first round, a pre-clearing step was conducted against immobilized GST. The selection enrichment was monitored via the ^{35}S-Met incorporated in the peptides and calculated for each round. In round 7 a spike in enrichment was observed followed by the beginning of a plateau in round 8. The surviving peptides were sequenced and in the resulting sequences, seven peptides recurred more than once and comprised more than 80 % of the sequences. These sequences were assigned numbers 8.1 to 8.7 based on their frequency in the sequencing results. Each sequence contains a D/E-X-X-F motif, which is analogous to the pS-X-X-F motif known to bind the BRCA1 $(BRCT)_2$ domain where the D and E are known phosphoserine mimetics. This was a promising result indicating the peptide likely binds in known binding groove. UNAAs that were selected were primarily the Phe analog (4-fluoro-phe) and the Arg analog (canavanine). It was interesting however, than none of the peptide sequences contained a second cysteine in the random region. MS/MS analysis showed that the dibromoxylene is instead forming a sulfonium ion linkage with the fixed methionine.

An initial ranking of the 7 peptide families in both their linear and cyclized (treated with dibromoxylene) forms was achieved with a competitive fluorescence polarization (FP) study with a peptide (FITC-β-A-pSPTF) known to inhibit the $(BRCT)_2$ domain. However, these tests are not completely indicative of what was selected because one of the UNAAS, canavanine, was

unable to be incorporated in solid phase peptide synthesis, so arginine was substituted instead. Upon this initial ranking it was clear that linear peptide 8.6 (with the sequence MCTIDF$_{26}$DEYRF$_{26}$KRT) had the highest affinity with peptides 8.1 and 8.5 just behind with approximately two-fold lower affinity.

To further investigate how this non-phosphorylated peptide binds to the (BRCT)$_2$ domain, a series of truncated mutants were tested in the same FP assay. The core domains (DEYRF$_{26}$ and F$_{26}$DEYRF$_{26}$) were not sufficient to observe binding to the domain. C-Terminal (MCTIDF$_{26}$DERYF$_{26}$) and N-terminal (DEYRF$_{26}$RKT) deletions revealed that the contribution of the N-terminal amino acids are much more significant that the C-terminal amino acids due to over an order of magnitude decrease compared to the mere two-fold decrease in affinity observed with the C-terminal deletion. Additionally deletion of the fixed MC at the N-terminus showed an approximately two-fold decrease affinity. This seems to indicate that the full length peptide is necessary to achieve the highest binding affinity.

Further mutational studies were conducted to investigate the importance of individual amino acids in the sequence of peptide 8.6. First, mutations of the E8 and F11 to A both nearly abolished binding altogether indicating that these amino acids are necessary, but not sufficient for binding. The contrary mutation, replacing E8 with a pS, surprisingly had only a three-fold increase in binding affinity when previous pS to E mutations resulting in a multiple order of magnitude decrease in affinity. Because of this observation, we wondered if the N-terminal amino acids would still be as important in the presence of the pS containing peptide. This N-terminal truncated pS peptide had a very slight increase in affinity leading to the conclusion that the N-terminal amino acids are no longer necessary when the pS is present in the E-X-X-F motif. Additionally, replacement of the two F$_{26}$ amino acids with F resulted in a two-fold increase in

binding affinity. The question of what role canavanine would play in these peptides is still open, but for now the all-natural (AN 8.6) will be used in future study.

One interesting variation between the 7 peptides found in selection is where the D/E-X-X-F was positioned in the random region. The two highest affinity hits 8.6 and 8.1 had very different position of this core motif. With the core boing closer to the C-terminus in peptide 8.6 and closer to the N-terminus in peptide 8.1, it was hypothesized that perhaps combining these sequences into longer peptides could result in a peptide that covers more of the protein surface and has a higher affinity. However, after making three different hybrid combinations, none of these peptides had an affinity higher than peptide 8.6. Mutations of individual amino acids of peptide 8.6 to amino acids found in the same position around the core domain of peptide 8.1 were also investigated. Although neither mutation of E8D, or Y9F and R10I achieved a peptide with a higher affinity than AN 8.6, they were less than a two-fold decrease in affinity. The mutation of D7N was nearly a neutral mutation.

Isothermal titration calorimetry (ITC) was conducted on peptides 8.6, AN 8.6, 8.6 E8pS and 8.6 E8pS 8-14. The affinities determined from ITC were found to be very similar to those determined by FP with AN 8.6 having a K_d of 1 μM. No differences in binding mechanism could be determined, but each binding interaction was enthalpically driven.

Although these peptides seem to inhibit the $(BRCT)_2$ domain *in vitro*, AN 8.6 was tested for its ability to inhibit binding interactions in cell lysate. In a pull-down western experiment, binding of CtIP-myc in cell lysate was able to be inhibited from binding to immobilized GST-$(BRCT)_2$ fusion in a dose-dependent manner. At concentrations of 50 μM in solution, this protein-protein interaction (PPI) was almost completely abrogated.

This work has shown a means of developing a non-phosphorylated peptide that can inhibit a PPI that has been previously thought to be 'un-druggable' due to the "requirement" of a phosphate. It has shown to be able to inhibit its target PPI in cell lysate, but there is still much work left to be done. With a 1 μM affinity, it is unclear as to whether or not it will be a viable inhibitor in cells. Preliminary data of cellular overexpression is promising, but can sufficient quantities of AN 8.6 be delivered exogenously? Various cell-penetrating techniques can be applied, but such an effort would also be aided by an increased affinity peptide. From this point, perhaps rational design would be a good place to start. Preliminary crystal structure data has given us some information about how this peptide binds, but can this be peptide be optimized from this structural information?

Another possibility is conducting another selection. It was found that a linear conformation was preferred, and the current selection was limited to a 12 amino acid random region. Would a longer peptide be beneficial? Throughout this work, it has been assumed that this peptide is in a floppy, linear conformation in solution, and remains linear upon binding, but what is its actual conformation in solution? If there is helical propensity, is it possible that altering the sequence to change this propensity could increase the affinity? What about another selection with a semi-random region based on the crystal structure data? Is optimization even necessary to begin to use this peptide as a tool? With an all-natural amino acid sequence, it could be over-expressed in cells to study BRCA1 inhibition. This work has also opened up the possibility of performing selections on other $(BRCT)_2$ domains. Each of these domains are known to interact with phosphoproteins, and in theory could be targeted with the same mRNA display library approach.

There are many directions that this work could continue, and in the grand tradition of scientific investigation it has opened more questions that it has answered. However, it did answer some important questions. We were able to use the power in numbers to find a non-phosphate containing inhibitor of BRCA1's $(BRCT)_2$ domain, and demonstrate that this inhibitor could be used successfully in cell lysate to disrupt a PPI. As PPIs become increasingly viewed as viable drug targets, studies like this one may help pave the way for a new method of drug design.

References

1. Broca, P., *Traité des tumeurs. Tome premier, Des Tumeurs en général*. P. Asselin: Paris, 1866.

2. Sinilnikova, O. M.; Mazoyer, S.; Bonnardel, C.; Lynch, H. T.; Narod, S. A.; Lenoir, G. M., BRCA1 and BRCA2 mutations in breast and ovarian cancer syndrome: reflection on the Creighton University historical series of high risk families. *Fam Cancer* **2006,** *5* (1), 15-20.

3. Hall, J. M.; Lee, M. K.; Newman, B.; Morrow, J. E.; Anderson, L. A.; Huey, B.; King, M. C., Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **1990,** *250* (4988), 1684-9.

4. Neuhausen, S. L.; Swensen, J.; Miki, Y.; Liu, Q.; Tavtigian, S.; Shattuck-Eidens, D.; Kamb, A.; Hobbs, M. R.; Gingrich, J.; Shizuya, H.; et al., A P1-based physical map of the region from D17S776 to D17S78 containing the breast cancer susceptibility gene BRCA1. *Hum Mol Genet* **1994,** *3* (11), 1919-26.

5. Chen, Y.; Farmer, A. A.; Chen, C. F.; Jones, D. C.; Chen, P. L.; Lee, W. H., BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. *Cancer Res* **1996,** *56* (14), 3168-72.

6. Wu, L. C.; Wang, Z. W.; Tsan, J. T.; Spillman, M. A.; Phung, A.; Xu, X. L.; Yang, M. C.; Hwang, L. Y.; Bowcock, A. M.; Baer, R., Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nat Genet* **1996,** *14* (4), 430-40.

7. (a) Manke, I. A.; Lowery, D. M.; Nguyen, A.; Yaffe, M. B., BRCT repeats as phosphopeptide-binding modules involved in protein targeting. *Science* **2003,** *302* (5645), 636-9; (b) Yu, X.; Chini, C. C.; He, M.; Mer, G.; Chen, J., The BRCT domain is a phospho-protein binding domain. *Science* **2003,** *302* (5645), 639-42.

8. (a) Cantor, S. B.; Bell, D. W.; Ganesan, S.; Kass, E. M.; Drapkin, R.; Grossman, S.; Wahrer, D. C.; Sgroi, D. C.; Lane, W. S.; Haber, D. A.; Livingston, D. M., BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* **2001,** *105* (1), 149-60; (b) Yu, X.; Wu, L. C.; Bowcock, A. M.; Aronheim, A.; Baer, R., The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *J Biol Chem* **1998,** *273* (39), 25388-92; (c) Wang, B.; Matsuoka, S.; Ballif, B. A.; Zhang, D.; Smogorzewska, A.; Gygi, S. P.; Elledge, S. J., Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* **2007,** *316* (5828), 1194-8; (d) Kim, H.; Huang, J.; Chen, J., CCDC98 is a BRCA1-BRCT domain-binding protein involved in the DNA damage response. *Nature structural & molecular biology* **2007,** *14* (8), 710-5; (e) Liu, Z.; Wu, J.; Yu, X., CCDC98 targets BRCA1 to DNA damage sites. *Nature structural & molecular biology* **2007,** *14* (8), 716-20.

9. Huen, M. S.; Sy, S. M.; Chen, J., BRCA1 and its toolbox for the maintenance of genome integrity. *Nat Rev Mol Cell Biol* **2010,** *11* (2), 138-48.

10. Hogervorst, F. B.; Cornelis, R. S.; Bout, M.; van Vliet, M.; Oosterwijk, J. C.; Olmer, R.; Bakker, B.; Klijn, J. G.; Vasen, H. F.; Meijers-Heijboer, H.; et al., Rapid detection of BRCA1 mutations by the protein truncation test. *Nat Genet* **1995,** *10* (2), 208-12.

11. Eitan, R.; Michaelson-Cohen, R.; Levavi, H.; Beller, U., The counseling and management of young healthy BRCA mutation carriers. *Int J Gynecol Cancer* **2009,** *19* (7), 1156-9.

12. Price, M.; Monteiro, A. N., Fine tuning chemotherapy to match BRCA1 status. *Biochem Pharmacol* **2010,** *80* (5), 647-53.

13. Smith, S. A.; Easton, D. F.; Evans, D. G.; Ponder, B. A., Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nat Genet* **1992,** *2* (2), 128-31.

14. Rigakos, G.; Razis, E., BRCAness: finding the Achilles heel in ovarian cancer. *The oncologist* **2012,** *17* (7), 956-62.

15. Wang, M.; Wu, W.; Wu, W.; Rosidi, B.; Zhang, L.; Wang, H.; Iliakis, G., PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res* **2006,** *34* (21), 6170-82.

16. Mizuarai, S.; Kotani, H., Synthetic lethal interactions for the development of cancer therapeutics: biological and methodological advancements. *Hum Genet* **2010,** *128* (6), 567-75.

17. Sandhu, S. K.; Yap, T. A.; de Bono, J. S., The emerging role of poly(ADP-Ribose) polymerase inhibitors in cancer treatment. *Curr Drug Targets* **2011,** *12* (14), 2034-44.

18. Jaspers, J. E.; Kersbergen, A.; Boon, U.; Sol, W.; van Deemter, L.; Zander, S. A.; Drost, R.; Wientjens, E.; Ji, J.; Aly, A.; Doroshow, J. H.; Cranston, A.; Martin, N. M.; Lau, A.; O'Connor, M. J.; Ganesan, S.; Borst, P.; Jonkers, J.; Rottenberg, S., Loss of 53BP1 causes PARP inhibitor resistance in Brca1-mutated mouse mammary tumors. *Cancer Discov* **2013,** *3* (1), 68-81.

19. (a) Litman, R.; Gupta, R.; Brosh, R. M., Jr.; Cantor, S. B., BRCA-FA pathway as a target for anti-tumor drugs. *Anticancer Agents Med Chem* **2008,** *8* (4), 426-30; (b) Xie, J.; Litman, R.; Wang, S.; Peng, M.; Guillemette, S.; Rooney, T.; Cantor, S. B., Targeting the FANCJ-BRCA1 interaction promotes a switch from recombination to poleta-dependent bypass. *Oncogene* **2010,** *29* (17), 2499-508.

20. Ng, R., *Drugs from Discovery to Approval*. Second ed.; Wiley-Blackwell: 2009.

21. Drews, J., Drug Discovery: A Historical Perspective. *Science* **2000,** *287* (5460), 1960-1964.

22. Sertuerner, F. W., Ueber eins der fürchterlichsten Gifte der Pflanzenwelt, als ein Nachtrag zu seiner Abhandlung über die Mekonsäure und das Morphium; mit Bemerkungen, den aciden Extractivstoff des Opiums und seine Verbindungen betreffend. *Annalen der Physik* **1817,** *57* (10), 183-202.

23. Drews, J.; Ryser, S., Classic drug targets. *Nat Biotech* **1997,** *15* (13), 1350-1350.

24. Bunnage, M. E., Getting pharmaceutical R&D back on target. *Nat Chem Biol* **2011,** *7* (6), 335-9.

25. (a) Rual, J. F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G. F.; Gibbons, F. D.; Dreze, M.; Ayivi-Guedehoussou, N.; Klitgord, N.; Simon, C.; Boxem, M.; Milstein, S.; Rosenberg, J.; Goldberg, D. S.; Zhang, L. V.; Wong, S. L.; Franklin, G.; Li, S.; Albala, J. S.; Lim, J.; Fraughton, C.; Llamosas, E.; Cevik, S.; Bex, C.; Lamesch, P.; Sikorski, R. S.; Vandenhaute, J.; Zoghbi, H. Y.; Smolyar, A.; Bosak, S.; Sequerra, R.; Doucette-Stamm, L.; Cusick, M. E.; Hill, D. E.; Roth, F. P.; Vidal, M., Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **2005,** *437* (7062), 1173-8; (b) Stelzl, U.; Worm, U.; Lalowski, M.; Haenig, C.; Brembeck, F. H.; Goehler, H.; Stroedicke, M.; Zenkner, M.; Schoenherr, A.; Koeppen, S.; Timm, J.; Mintzlaff, S.; Abraham, C.; Bock, N.; Kietzmann, S.; Goedde, A.; Toksöz, E.; Droege, A.; Krobitsch, S.; Korn, B.; Birchmeier, W.; Lehrach, H.; Wanker, E. E., A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome. *Cell* **2005,** *122* (6), 957-968.

26. Blundell, T. L.; Sibanda, B. L.; Montalvao, R. W.; Brewerton, S.; Chelliah, V.; Worth, C. L.; Harmer, N. J.; Davies, O.; Burke, D., Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. *Philos Trans R Soc Lond B Biol Sci* **2006,** *361* (1467), 413-23.

27. Fletcher, S.; Hamilton, A. D., Protein surface recognition and proteomimetics: mimics of protein surface structure and function. *Curr Opin Chem Biol* **2005,** *9* (6), 632-8.

28. Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatisky, M., Synthetic therapeutic peptides: science and market. *Drug Discov Today* **2010,** *15* (1-2), 40-56.

29. Milletti, F., Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov Today* **2012,** *17* (15-16), 850-60.

30. Yao, L.; Daniels, J.; Wijesinghe, D.; Andreev, O. A.; Reshetnyak, Y. K., pHLIP(R)-mediated delivery of PEGylated liposomes to cancer cells. *J Control Release* **2013,** *167* (3), 228-37.

31. Arias, J. L.; Clares, B.; Morales, M. E.; Gallardo, V.; Ruiz, M. A., Lipid-based drug delivery systems for cancer treatment. *Curr Drug Targets* **2011,** *12* (8), 1151-65.

32. Du, J.; Jin, J.; Yan, M.; Lu, Y., Synthetic nanocarriers for intracellular protein delivery. *Curr Drug Metab* **2012,** *13* (1), 82-92.

33. Grigoryev, Y., Stapled peptide to enter human testing, but affinity questions remain. *Nat Med* **2013,** *19* (2), 120-120.

34. Raimundo, B. C.; Oslob, J. D.; Braisted, A. C.; Hyde, J.; McDowell, R. S.; Randal, M.; Waal, N. D.; Wilkinson, J.; Yu, C. H.; Arkin, M. R., Integrating fragment assembly and biophysical methods in the chemical advancement of small-molecule antagonists of IL-2: an approach for inhibiting protein-protein interactions. *J Med Chem* **2004,** *47* (12), 3111-30.

35. Verdine, G. L.; Walensky, L. D., The challenge of drugging undruggable targets in cancer: lessons learned from targeting BCL-2 family members. *Clin Can Res* **2007,** *13* (24), 7264-70.

36. (a) Shiozaki, E. N.; Gu, L.; Yan, N.; Shi, Y., Structure of the BRCT Repeats of BRCA1 Bound to a BACH1 Phosphopeptide: Implications for Signaling. *Mol Cell* **2004,** *14* (3), 405-412; (b) Clapperton, J. A.; Manke, I. A.; Lowery, D. M.; Ho, T.; Haire, L. F.; Yaffe, M. B.; Smerdon, S. J., Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. *Nat Struct Mol Biol* **2004,** *11* (6), 512-8; (c) Williams, R. S.; Lee, M. S.; Hau, D. D.; Glover, J. N. M., Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1. *Nat Struct Mol Biol* **2004,** *11* (6), 519-525; (d) Brzovic, P. S.; Rajagopal, P.; Hoyt, D. W.; King, M.-C.; Klevit, R. E., Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat Struct Mol Biol* **2001,** *8* (10), 833-837.

37. Brzovic, P. S.; Rajagopal, P.; Hoyt, D. W.; King, M. C.; Klevit, R. E., Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat Struct Biol* **2001,** *8* (10), 833-7.

38. Brzovic, P. S.; Keeffe, J. R.; Nishikawa, H.; Miyamoto, K.; Fox, D., 3rd; Fukuda, M.; Ohta, T.; Klevit, R., Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proc Nat Acad Sci USA* **2003,** *100* (10), 5646-51.

39. (a) Scully, R.; Ganesan, S.; Vlasakova, K.; Chen, J.; Socolovsky, M.; Livingston, D. M., Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol Cell* **1999,** *4* (6), 1093-9; (b) Ruffner, H.; Joazeiro, C. A. P.; Hemmati, D.; Hunter, T.; Verma, I. M., Cancer-predisposing mutations within the RING domain of BRCA1: Loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. *ProcNat Acad Sci* **2001,** *98* (9), 5134-5139.

40. Jiang, J.; Yang, E. S.; Jiang, G.; Nowsheen, S.; Wang, H.; Wang, T.; Wang, Y.; Billheimer, D.; Chakravarthy, A. B.; Brown, M.; Haffty, B.; Xia, F., p53-Dependent BRCA1 Nuclear Export Controls Cellular Susceptibility to DNA Damage. *Can Res* **2011,** *71* (16), 5546-5557.

41. Wild, C.; Dubay, J. W.; Greenwell, T.; Baird, T., Jr.; Oas, T. G.; McDanal, C.; Hunter, E.; Matthews, T., Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex. *Proc Nat Acad Sci USA* **1994,** *91* (26), 12676-80.

42. Brzovic, P. S.; Meza, J. E.; King, M.-C.; Klevit, R. E., BRCA1 RING Domain Cancer-predisposing Mutations. *J Biol Chem* **2001,** *276* (44), 41399-41406.

43. Almaraz, E.; de Paula, Q. A.; Liu, Q.; Reibenspies, J. H.; Darensbourg, M. Y.; Farrell, N. P., Thiolate bridging and metal exchange in adducts of a zinc finger model and Pt(II) complexes: biomimetic studies of protein/Pt/DNA interactions. *J Am Chem Soc* **2008,** *130* (19), 6272-80.

44. Atipairin, A.; Canyuk, B.; Ratanaphan, A., Cisplatin Affects the Conformation of Apo Form, not Holo Form, of BRCA1 RING Finger Domain and Confers Thermal Stability. *Chem Biodiversity* **2010,** *7* (8), 1949-1967.

45. Yu, X.; Chini, C. C. S.; He, M.; Mer, G.; Chen, J., The BRCT Domain Is a Phospho-Protein Binding Domain. *Science* **2003,** *302* (5645), 639-642.

46. Rodriguez, M.; Yu, X.; Chen, J.; Songyang, Z., Phosphopeptide Binding Specificities of BRCA1 COOH-terminal (BRCT) Domains. *J Biol Chem* **2003,** *278* (52), 52914-52918.

47. (a) Lokesh, G. L.; Rachamallu, A.; Kumar, G. D. K.; Natarajan, A., High-throughput fluorescence polarization assay to identify small molecule inhibitors of BRCT domains of breast cancer gene 1. *Analytical Biochemistry* **2006,** *352* (1), 135-141; (b) Simeonov, A.; Yasgar, A.; Jadhav, A.; Lokesh, G. L.; Klumpp, C.; Michael, S.; Austin, C. P.; Natarajan, A.; Inglese, J., Dual-fluorophore quantitative high-throughput screen for inhibitors of BRCT phosphoprotein interaction. *Anal Biochem* **2008,** *375* (1), 60-70.

48. (a) Yuan, Z.; Kumar, E. A.; Campbell, S. J.; Palermo, N. Y.; Kizhake, S.; Mark Glover, J. N.; Natarajan, A., Exploiting the P-1 pocket of BRCT domains toward a structure guided inhibitor design. *ACS medicinal chemistry letters* **2011,** *2* (10), 764-767; (b) Yuan, Z.; Kumar, E. A.; Kizhake, S.; Natarajan, A., Structure-activity relationship studies to probe the phosphoprotein binding site on the carboxy terminal domains of the breast cancer susceptibility gene 1. *Journal of Medicinal Chemistry* **2011,** *54* (12), 4264-4268; (c) Lokesh, G. L.; Muralidhara, B. K.; Negi, S. S.; Natarajan, A., Thermodynamics of Phosphopeptide Tethering to BRCT:â€‰ The Structural Minima for Inhibitor Design. *J Am Chem Soc* **2007,** *129* (35), 10658-10659.

49. Pessetto, Z. Y.; Yan, Y.; Bessho, T.; Natarajan, A., Inhibition of BRCT(BRCA1)-phosphoprotein interaction enhances the cytotoxic effect of olaparib in breast cancer cells: a proof of concept study for synthetic lethal therapeutic option. *Breast Cancer Res Treat* **2012,** *134* (2), 511-7.

50. (a) Gatei, M.; Scott, S. P.; Filippovitch, I.; Soronika, N.; Lavin, M. F.; Weber, B.; Khanna, K. K., Role for ATM in DNA damage-induced phosphorylation of BRCA1. *Cancer Res* **2000,** *60* (12), 3299-304; (b) Gatei, M.; Zhou, B. B.; Hobson, K.; Scott, S.; Young, D.; Khanna, K. K., Ataxia telangiectasia mutated (ATM) kinase and ATM and Rad3 related kinase mediate phosphorylation of Brca1 at distinct and overlapping sites. In vivo assessment using phospho-specific antibodies. *J Biol Chem* **2001,** *276* (20), 17276-80; (c) Xu, B.; O'Donnell, A. H.; Kim, S. T.; Kastan, M. B., Phosphorylation of serine 1387 in Brca1 is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation. *Cancer Res* **2002,** *62* (16), 4588-91; (d) Cortez, D.; Wang, Y.; Qin, J.; Elledge, S. J., Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. *Science* **1999,** *286* (5442), 1162-6.

51. (a) Sy, S. M. H.; Huen, M. S. Y.; Chen, J., PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc Nat Acad Sci* **2009,** *106* (17), 7155-7160; (b) Zhang, F.; Ma, J.; Wu, J.; Ye, L.; Cai, H.; Xia, B.; Yu, X., PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Current biology : CB* **2009,** *19* (6), 524-9; (c) Xia, B.; Sheng, Q.; Nakanishi, K.; Ohashi, A.; Wu, J.; Christ, N.; Liu, X.; Jasin, M.; Couch, F. J.; Livingston, D. M., Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* **2006,** *22* (6), 719-29.

52. Kritzer, J. A., Stapled peptides: Magic bullets in nature's arsenal. *Nat Chem Biol* **2010,** *6* (8), 566-567.

53. Mavromoustakos, T.; Durdagi, S.; Koukoulitsa, C.; Simcic, M.; Papadopoulos, M. G.; Hodoscek, M.; Grdadolnik, S. G., Strategies in the rational drug design. *Curr Med Chem* **2011,** *18* (17), 2517-30.

54. Vanhee, P.; van der Sloot, A. M.; Verschueren, E.; Serrano, L.; Rousseau, F.; Schymkowitz, J., Computational design of peptide ligands. *Trends Biotechnol* **2011,** *29* (5), 231-9.

55. (a) Smith, G. P., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985,** *228* (4705), 1315-7; (b) Parmley, S. F.; Smith, G. P., Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **1988,** *73* (2), 305-18; (c) Ullman, C. G.; Frigotto, L.; Cooley, R. N., In vitro methods for peptide display and their applications. *Brief Funct Genomics* **2011,** *10* (3), 125-34.

56. (a) Krumpe, L. R.; Mori, T., Potential of phage-displayed peptide library technology to identify functional targeting peptides. *Expert Opin Drug Discov* **2007,** *2* (4), 525; (b) Hoogenboom, H. R., Selecting and screening recombinant antibody libraries. *Nat Biotechnol* **2005,** *23* (9), 1105-16; (c) Sidhu, S. S., Phage display in pharmaceutical biotechnology. *Curr Opin Biotechnol* **2000,** *11* (6), 610-6; (d) Sergeeva, A.; Kolonin, M. G.; Molldrem, J. J.; Pasqualini, R.; Arap, W., Display technologies: application for the discovery of drug and gene delivery agents. *Adv Drug Deliv Rev* **2006,** *58* (15), 1622-54.

57. Frank, R., The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods* **2002,** *267* (1), 13-26.

58. Katz, C.; Levy-Beladev, L.; Rotem-Bamberger, S.; Rito, T.; Rudiger, S. G.; Friedler, A., Studying protein-protein interactions using peptide arrays. *Chem Soc Rev* **2011,** *40* (5), 2131-45.

59. Lam, K. S.; Lebl, M., Synthesis of a one-bead one-compound combinatorial peptide library. *Methods Mol Biol* **1998,** *87*, 1-6.

60. Lebl, M.; Krchnak, V.; Sepetov, N. F.; Seligmann, B.; Strop, P.; Felder, S.; Lam, K. S., One-bead-one-structure combinatorial libraries. *Biopolymers* **1995,** *37* (3), 177-98.

61. Lam, K. S., Enzyme-linked colorimetric screening of a one-bead one-compound combinatorial library. *Methods Mol Biol* **1998,** *87*, 7-12.

62. Lorthioir, O.; Carr, R. A.; Congreve, M. S.; Geysen, M. H.; Kay, C.; Marshall, P.; McKeown, S. C.; Parr, N. J.; Scicinski, J. J.; Watson, S. P., Single bead characterization using analytical constructs: application to quality control of libraries. *Anal Chem* **2001,** *73* (5), 963-70.

63. Kent, S. B., Chemical synthesis of peptides and proteins. *Annu Rev Biochem* **1988,** *57*, 957-89.

64. Mitchell, A. R., Studies in solid-phase peptide synthesis: a personal perspective. *Biopolym* **2008,** *90* (3), 215-33.

65. Benatuil, L.; Perez, J. M.; Belk, J.; Hsieh, C.-M., An improved yeast transformation method for the generation of very large human antibody libraries. 2010; Vol. 23, pp 155-159.

66. Mersich, C.; Jungbauer, A., Generation of bioactive peptides by biological libraries. *J Chromatogr B* **2008,** *861* (2), 160-170.

67. (a) Robertson, S. A., Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C., Schultz, P.G. , The use of 5'-phospho-2'-deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of tRNA. *Nucleic Acids Res.* **1989,** *17* (23), 9649-9660; (b) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R., Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *J Amer Chem Soc* **1989,** *111* (20), 8013-8014; (c) Rodriguez Ea Fau - Lester, H. A.; Lester Ha Fau - Dougherty, D. A.; Dougherty, D. A., In vivo incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. **2006,** (0027-8424 (Print)); (d) Kohrer, C.; Sullivan, E. L.; RajBhandary, U. L., Complete set of orthogonal 21st aminoacyl-tRNA synthetase-amber, ochre and opal suppressor tRNA pairs: concomitant suppression of three different termination codons in an mRNA in mammalian cells. *Nucleic Acids Res* **2004,** *32* (21), 6200-11.

68. Smith, G. P.; Petrenko, V. A., Phage Display. *Chem Rev* **1997,** *97* (2), 391-410.

69. Zahnd, C.; Amstutz, P.; Pluckthun, A., Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. *Nat Meth* **2007,** *4* (3), 269-79.

70. Kurz, M.; Gu, K.; Al-Gawari, A.; Lohse, P. A., cDNA - protein fusions: covalent protein - gene conjugates for the in vitro selection of peptides and proteins. *Chembiochem* **2001,** *2* (9), 666-72.

71. Roberts, R. W.; Szostak, J. W., RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A* **1997,** *94* (23), 12297-302.

72. Nemoto, N.; Miyamoto-Sato, E.; Husimi, Y.; Yanagawa, H., In vitro virus: Bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett* **1997,** *414* (2), 405-408.

73. Mersich, C.; Jungbauer, A., Generation of bioactive peptides by biological libraries. *J Chromatogr B* **2008,** *861* (2), 160-70.

74. Gold, L., mRNA display: diversity matters during in vitro selection. *Proc Natl Acad Sci U S A* **2001,** *98* (9), 4825-6.

75. Hartman, M. C.; Josephson, K.; Lin, C. W.; Szostak, J. W., An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PloS one* **2007,** *2* (10), e972.

76. Liu, C. C.; Schultz, P. G., Adding New Chemistries to the Genetic Code. *Annu Rev Biochem* **2010,** *79* (1), 413-444.

77. (a) Nguyen, D. P.; Garcia Alai, M. M.; Virdee, S.; Chin, J. W., Genetically directing varepsilon-N, N-dimethyl-L-lysine in recombinant histones. *Chem Biol* **2010,** *17* (10), 1072-6; (b) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W., Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat Chem Biol* **2008,** *4* (4), 232-4; (c) Groff, D.; Chen, P. R.; Peters, F.

B.; Schultz, P. G., A genetically encoded epsilon-N-methyl lysine in mammalian cells. *Chembiochem* **2010,** *11* (8), 1066-8; (d) Liu, C. C.; Mack, A. V.; Tsao, M. L.; Mills, J. H.; Lee, H. S.; Choe, H.; Farzan, M.; Schultz, P. G.; Smider, V. V., Protein evolution with an expanded genetic code. *Proc Natl Acad Sci U S A* **2008,** *105* (46), 17688-93; (e) Park, H. S.; Hohn, M. J.; Umehara, T.; Guo, L. T.; Osborne, E. M.; Benner, J.; Noren, C. J.; Rinehart, J.; Soll, D., Expanding the genetic code of Escherichia coli with phosphoserine. *Science (New York, N.Y.)* **2011,** *333* (6046), 1151-4.

78. (a) Hao, Z.; Hong, S.; Chen, X.; Chen, P. R., Introducing bioorthogonal functionalities into proteins in living cells. *Acc Chem Res* **2011,** *44* (9), 742-51; (b) Best, M. D., Click chemistry and bioorthogonal reactions: unprecedented selectivity in the labeling of biological molecules. *Biochem* **2009,** *48* (28), 6571-84.

79. Adessi, C.; Soto, C., Converting a peptide into a drug: strategies to improve stability and bioavailability. *Curr Med Chem* **2002,** *9* (9), 963-78.

80. (a) Biron, E.; Chatterjee, J.; Ovadia, O.; Langenegger, D.; Brueggen, J.; Hoyer, D.; Schmid, H. A.; Jelinek, R.; Gilon, C.; Hoffman, A.; Kessler, H., Improving oral bioavailability of peptides by multiple N-methylation: somatostatin analogues. *Angew Chem Int Ed* **2008,** *47* (14), 2595-9; (b) Fiacco, S. V.; Roberts, R. W., N-Methyl scanning mutagenesis generates protease-resistant G protein ligands with improved affinity and selectivity. *Chembiochem* **2008,** *9* (14), 2200-3.

81. Heckler, T. G.; Chang, L. H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M., T4 RNA ligase mediated preparation of novel "chemically misacylated" tRNAPheS. *Biochem* **1984,** *23* (7), 1468-73.

82. (a) Forster, A. C.; Tan, Z.; Nalam, M. N.; Lin, H.; Qu, H.; Cornish, V. W.; Blacklow, S. C., Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc Nat Acad Sci USA* **2003,** *100* (11), 6353-7; (b) Frankel, A.; Millward, S. W.; Roberts, R. W., Encodamers: unnatural peptide oligomers encoded in RNA. *Chem Biol* **2003,** *10* (11), 1043-50.

83. (a) Young, T. S.; Schultz, P. G., Beyond the canonical 20 amino acids: expanding the genetic lexicon. *J Biol Chem* **2010,** *285* (15), 11039-11044; (b) Wang, Q.; Parrish, A. R.; Wang, L., Expanding the genetic code for biological studies. *Chem Biol* **2009,** *16* (3), 323-6.

84. (a) Merryman, C.; Green, R., Transformation of aminoacyl tRNAs for the in vitro selection of "drug-like" molecules. *Chem Biol* **2004,** *11* (4), 575-582; (b) Subtelny, A. O.; Hartman, M. C. T.; Szostak, J. W., Ribosomal Synthesis of N-methyl peptides. *J Am Chem Soc* **2008,** *130* (19), 6131-6 .

85. Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H., A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat Meth* **2006,** *3* (5), 357-9.

86. Passioura, T.; Suga, H., Flexizyme-Mediated Genetic Reprogramming As a Tool for Noncanonical Peptide Synthesis and Drug Discovery. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2013**.

87. (a) Ohuchi M, M. H., Suga H., The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Curr Opin Chem Biol* **2007,** *11* (5), 537-542; (b) Nakajima, E.; Goto, Y.; Sako, Y.; Murakami, H.; Suga, H., Ribosomal synthesis of peptides with C-terminal lactams, thiolactones, and alkylamides. *Chembiochem* **2009,** *10* (7), 1186-1192; (c) Goto, Y.; Iwasaki, K.; Torikai, K.; Murakami, H.; Suga, H., Ribosomal synthesis of dehydrobutyrine- and methyllanthionine-containing peptides. *Chem Commun* **2010,** *23*, 3419-3421; (d) Kawakami, T.; Murakami, H.; Suga, H., Ribosomal synthesis of polypeptoids and peptoid-peptide hybrids. *J Am Chem Soc* **2008,** *130* (50), 16861-16863.

88. (a) Hartman, M. C.; Josephson, K.; Lin, C. W.; Szostak, J. W., An expanded set of amino cid analogs for the ribosomal translation of unnatural peptides. *PLoS ONE* **2007,** *2* (10), e972; (b) Hartman, M. C.; Josephson, K.; Szostak, J. W., Enzymatic aminoacylation of tRNA with unnatural amino acids. *Proc Nat Acad Sci USA* **2006,** *103* (12), 4356-61.

89. Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T., Cell-free translation reconstituted with purified components. *Nat Biotechnol* **2001,** *19* (8), 751-5.

90. Matsuura, T.; Yomo, T., In vitro evolution of proteins. *J Biosci Bioeng* **2006,** *101* (6), 449-56.

91. (a) Reid, P. C.; Goto, Y.; Katoh, T.; Suga, H., Charging of tRNAs using ribozymes and selection of cyclic peptides containing thioethers. *Methods Mol Biol* **2012,** *805*, 335-48; (b) Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H., Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem Biol* **2011,** *18* (12), 1562-70; (c) Day, J. W.; Kim, C. H.; Smider, V. V.; Schultz, P. G., Identification of metal ion binding peptides containing unnatural amino acids by phage display. *Bioorg Med Chem Lett* **2013,** *23* (9), 2598-600.

92. Schlippe, Y. V.; Hartman, M. C.; Josephson, K.; Szostak, J. W., In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J Am Chem Soc* **2012,** *134* (25), 10469-77.

93. Vuignier, K.; Schappler, J.; Veuthey, J. L.; Carrupt, P. A.; Martel, S., Drug-protein binding: a critical review of analytical tools. *Anal Bioanal Chem* **2010,** *398* (1), 53-66.

94. Jenison, R. D.; Gill, S. C.; Pardi, A.; Polisky, B., High-resolution molecular discrimination by RNA. 1994; Vol. 263, pp 1425-1429.

95. Perozzo, R.; Folkers, G.; Scapozza, L., Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *J Recept Signal Transduct Res* **2004,** *24* (1-2), 1-52.

96. Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L. N., Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **1989,** *179* (1), 131-7.

97. Turnbull, W. B.; Daranas, A. H., On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J Am Chem Soc* **2003,** *125* (48), 14859-66.

98. Jameson, D. M.; Ross, J. A., Fluorescence Polarization/Anisotropy in Diagnostics and Imaging. *Chem Rev* **2010,** *110* (5), 2685-2708.

99. Szabo, A.; Stolz, L.; Granzow, R., Surface plasmon resonance and its use in biomolecular interaction analysis (BIA). *Curr Opin Struct Biol* **1995,** *5* (5), 699-705.

100. Jason-Moller, L.; Murphy, M.; Bruno, J., Overview of Biacore systems and their applications. *Curr Protoc Protein Sci* **2006,** *Chapter 19*, Unit 19 13.

101. Hartman, M. C. T.; Josephson, K.; Lin, C.-W.; Szostak, J. W., An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS ONE* **2007,** *2* (10), e972.

102. (a) White, E. R.; Reed, T. M.; Ma, Z.; Hartman, M. C. T., Replacing amino acids in translation: Expanding chemical diversity with non-natural variants. *Methods* **2013,** *60* (1), 70-74; (b) Traboni, C.; Cortese, R.; Salvatore, F., Selective 32P-labelling of individual species in a total tRNA population. *Nucleic Acids Res* **1980,** *8* (22), 5223-5232.

103. Ma, Z.; Hartman, M. C., In vitro selection of unnatural cyclic peptide libraries via mRNA display. *Meth Mol Biol* **2012,** *805*, 367-90.

104. Manke, I. A.; Lowery, D. M.; Nguyen, A.; Yaffe, M. B., BRCT repeats as phosphopeptide-binding modules involved in protein targeting. *Science* **2003,** *302* (5645), 636-639.

105. (a) Farmer, H.; McCabe, N.; Lord, C. J.; Tutt, A. N. J.; Johnson, D. A.; Richardson, T. B.; Santarosa, M.; Dillon, K. J.; Hickson, I.; Knights, C.; Martin, N. M. B.; Jackson, S. P.; Smith, G. C. M.; Ashworth, A., Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **2005,** *434* (7035), 917-921; (b) Konstantinopoulos, P. A.; Spentzos, D.; Karlan, B. Y.; Taniguchi, T.; Fountzilas, E.; Francoeur, N.; Levine, D. A.; Cannistra, S. A., Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J Clin Oncol* **2010,** *28* (22), 3555-3561.

106. Huen, M. S. Y.; Sy, S. M. H.; Chen, J., BRCA1 and its toolbox for the maintenance of genome integrity. *Nat Rev Mol Cell Biol* **2010,** *11* (2), 138-148.

107. Rodriguez, M.; Yu, X.; Chen, J.; Songyang, Z., Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains. *J Biol Chem* **2003,** *278* (52), 52914-52918.

108. Yuan, Z.; Kumar, E. A.; Campbell, S. J.; Palermo, N. Y.; Kizhake, S.; Glover, J. N. M.; Natarajan, A., Exploiting the P-1 pocket of BRCT domains toward a structure guided inhibitor design. *ACS Med Chem Lett* **2011,** *2* (10), 764-767.

109. Lokesh, G. L.; Muralidhara, B. K.; Negi, S. S.; Natarajan, A., Thermodynamics of phosphopeptide tethering to BRCT: the structural minima for inhibitor design. *J Am Chem Soc* **2007,** *129* (35), 10658-10659.

110. (a) Nemoto, N.; Miyamoto-Sato, E.; Husimi, Y.; Yanagawa, H., In vitro virus: bonding of mRNA bearing puromycin at the 3-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett* **1997,** *414* (2), 405-408; (b) Roberts, R. W.; Szostak, J. W.,

RNA-peptide fusions for the in vitro selection of peptides and proteins. *ProcNat Acad Sci* **1997,** *94* (23), 12297-12302.

111. (a) Forster, A. C.; Tan, Z.; Nalam, M. N. L.; Lin, H.; Qu, H.; Cornish, V. W.; Blacklow, S. C., Programming peptidomimetic syntheses by translating genetic codes designed de novo. *ProcNat Acad Sci* **2003,** *100* (11), 6353-6357; (b) Frankel, A.; Millward, S. W.; Roberts, R. W., Encodamers: unnatural peptide oligomers encoded in RNA. *Chem Biol* **2003,** *10* (11), 1043-1050; (c) Guillen Schlippe, Y. V.; Hartman, M. C. T.; Josephson, K.; Szostak, J. W., In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J Am Chem Soc* **2012,** *134* (25), 10469-10477; (d) Hipolito, C. J.; Suga, H., Ribosomal production and in vitro selection of natural product-like peptidomimetics: the FIT and RaPID systems. *Curr Opin Chem Biol* **2012,** *16* (1-2), 196-203; (e) Josephson, K.; Hartman, M. C. T.; Szostak, J. W., Ribosomal synthesis of unnatural peptides. *J Am Chem Soc* **2005,** *127* (33), 11727-35.

112. (a) Gold, L., mRNA display: diversity matters during in vitro selection. *Proc Nat Acad Sci* **2001,** *98* (9), 4825-4826; (b) Zhang, J.; Williams, B. A. R.; Nilsson, M. T.; Chaput, J. C., The evolvability of lead peptides from small library screens. *Chem Commun* **2011,** *46* (41), 7778-80.

113. (a) Dewkar, G. K.; Carneiro, P. B.; Hartman, M. C. T., Synthesis of novel peptide linkers: simultaneous cyclization and labeling. *Org Lett* **2009,** *11* (20), 4708-11; (b) Timmerman, P.; Beld, J.; Puijk, W. C.; Meloen, R. H., Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem* **2005,** *6* (5), 821-4.

114. (a) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. G., Lowering the entropic barrier for binding conformationally flexible inhibitors to enzymes. *Biochem* **1998,** *37* (48), 16839-16845; (b) Kwon, Y.-U.; Kodadek, T., Quantitative comparison of the relative cell permeability of cyclic and linear peptides. *Chem Biol* **2007,** *14* (6), 671-677; (c) Piserchio, A.; Salinas, G. D.; Li, T.; Marshall, J.; Spaller, M. R.; Mierke, D. F., Targeting specific PDZ domains of PSD-95: structural basis for enhanced affinity and enzymatic stability of a cyclic peptide. *Chem Biol* **2004,** *11* (4), 469-473.

115. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein Express Purif* **2005,** *41* (1), 207-34.

116. Ma, Z.; Hartman, M. C., In vitro selection of unnatural cyclic peptide libraries via mRNA display. *Meth Mol Biol* **2012,** *805*, 367-390.

117. Dever, S. M.; White, E. R.; Hartman, M. C. T.; Valerie, K., BRCA1-directed, enhanced and aberrant homologous recombination: mechanism and potential treatment strategies. *Cell Cycle* **2012,** *11* (4), 687-694.

118. (a) Coster, G.; Hayouka, Z.; Argaman, L.; Strauss, C.; Friedler, A.; Brandeis, M.; Goldberg, M., The DNA damage response mediator MDC1 directly interacts with the anaphase-promoting complex/cyclosome. *J Biol Chem* **2007,** *282* (44), 32053-32064; (b) Gautel, M., Zuffardi, O, Freiburg, A, and Labeit, S, Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? *EMBO J.* **1995,** *14* (9), 1952-1960; (c) Maciejewski, P. M.; Peterson, F. C.; Anderson, P. J.; Brooks, C. L., Mutation of serine 90 to glutamic acid mimics phosphorylation of bovine prolactin. *J Biol Chem* **1995,** *270*

(46), 27661-27665; (d) Tyson, D. R.; Swarthout, J. T.; Jefcoat, S. C.; Partridge, N. C., PTH Induction of transcriptional activity of the cAMP response element-binding protein requires the serine 129 site and glycogen synthase kinase-3 activity, but not casein kinase II sites. *Endocrinol* **2002,** *143* (2), 674-682.

119. Xu, B.; O'Donnell, A. H.; Kim, S. T.; Kastan, M. B., Phosphorylation of serine 1387 in Brca1 is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation. *Can Res* **2002,** *62* (16), 4588-91.

120. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein ExpressPurif* **2005,** *41* (1), 207-34.

121. Cai, M.; Williams, D. C.; Wang, G.; Lee, B. R.; Peterkofsky, A.; Clore, G. M., Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIAGlucose and the cytoplasmic domain of the glucose transporter IICBGlucose of the Escherichia coli glucose phosphotransferase system. *J Biol Chem* **2003,** *278* (27), 25191-25206.

122. Josephson, K.; Hartman, M. C.; Szostak, J. W., Ribosomal synthesis of unnatural peptides. *J Am Chem Soc* **2005,** *127* (33), 11727-35.

123. Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J., Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009,** *25* (9), 1189-1191.

124. Pajpanova, T.; Stoev, S.; Golovinsky, E.; Krauß, G. J.; Miersch, J., Canavanine derivatives useful in peptide synthesis. *Amino Acids* **1997,** *12* (2), 191-204.

125. Boyar, A.; Marsh, R. E., l-Canavanine, a paradigm for the structures of substituted guanidines. *J Am Chem Soc* **1982,** *104* (7), 1995-1998.

126. Gill, S. C.; von Hippel, P. H., Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **1989,** *182* (2), 319-326.

127. Cai M Fau - Williams, D. C., Jr.; Williams Dc Jr Fau - Wang, G.; Wang G Fau - Lee, B. R.; Lee Br Fau - Peterkofsky, A.; Peterkofsky A Fau - Clore, G. M.; Clore, G. M. *J Biol Chem* **2003,** *278* (27), 25191-206.

128. (a) Lokesh, G. L.; Muralidhara, B. K.; Negi, S. S.; Natarajan, A., Thermodynamics of phosphopeptide tethering to BRCT: the structural minima for inhibitor design. *J Am Chem Soc* **2007,** *129* (35), 10658-9; (b) Yuan, Z.; Kumar, E. A.; Kizhake, S.; Natarajan, A., Structure-activity relationship studies to probe the phosphoprotein binding site on the carboxy terminal domains of the breast cancer susceptibility gene 1. *J Med Chem* **2011,** *54* (12), 4264-8.

129. Xinyi, H., Fluorescence polarization competition assay: the range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand. *J Biomol Screen* **2003,** *8* (1), 34-38.

130. Biffinger, J. C.; Kim, H. W.; DiMagno, S. G., The polar hydrophobicity of fluorinated compounds. *ChemBioChem* **2004,** *5* (5), 622-627.

131. Yu, X.; Fu, S.; Lai, M.; Baer, R.; Chen, J., BRCA1 ubiquitinates its phosphorylation-dependent binding partner CtIP. *Genes Dev* **2006,** *20* (13), 1721-1726.

132. Lokesh, G. L.; Rachamallu, A.; Kumar, G. D.; Natarajan, A., High-throughput fluorescence polarization assay to identify small molecule inhibitors of BRCT domains of breast cancer gene 1. *Anal Biochem* **2006,** *352* (1), 135-41.

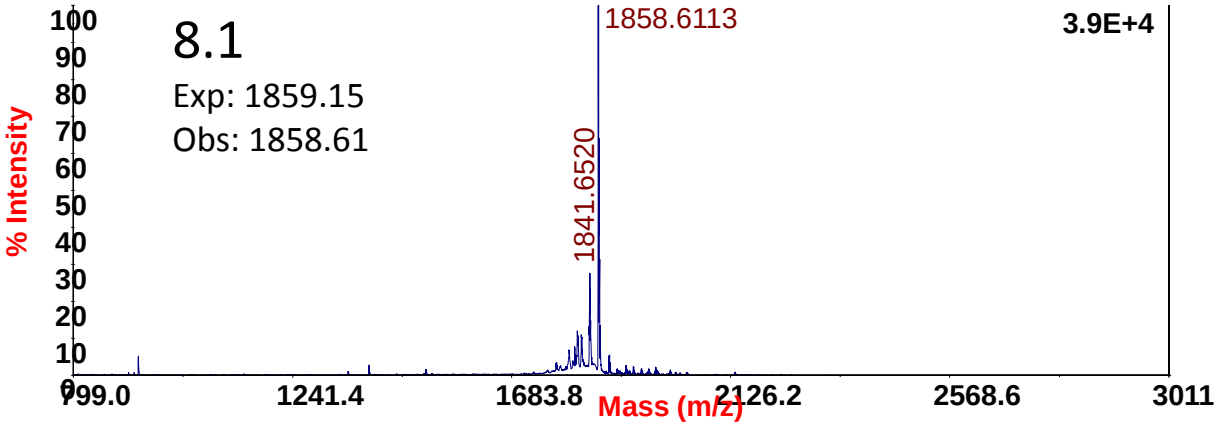
133. Shen, Y.; Tong, L., Structural evidence for direct interactions between the BRCT domains of human BRCA1 and a phospho-peptide from human ACC1. *Biochem* **2008,** *47* (21), 5767-73.

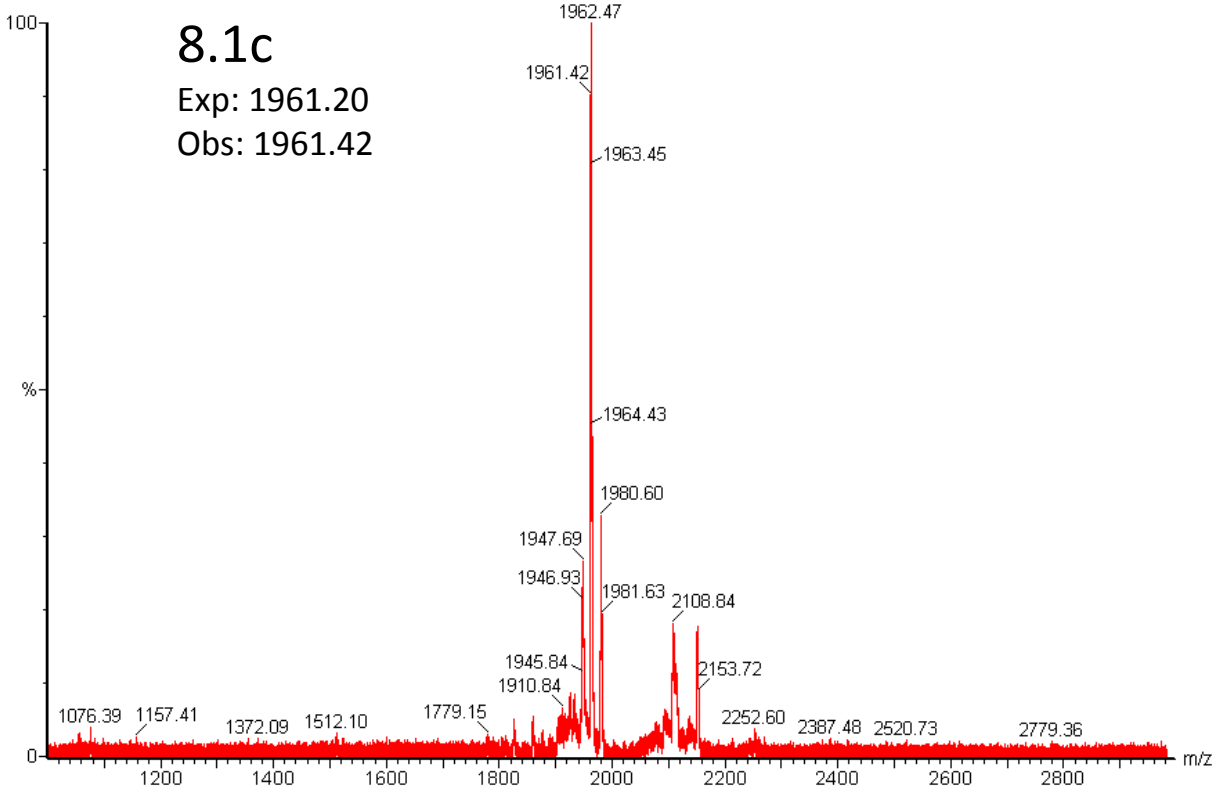
134. (a) Drikos, I.; Nounesis, G.; Vorgias, C. E., Characterization of cancer-linked BRCA1-BRCT missense variants and their interaction with phosphoprotein targets. *Proteins* **2009,** *77* (2), 464-76; (b) Varma, A. K.; Brown, R. S.; Birrane, G.; Ladias, J. A., Structural basis for cell cycle checkpoint control by the BRCA1-CtIP complex. *Biochem* **2005,** *44* (33), 10941-6.

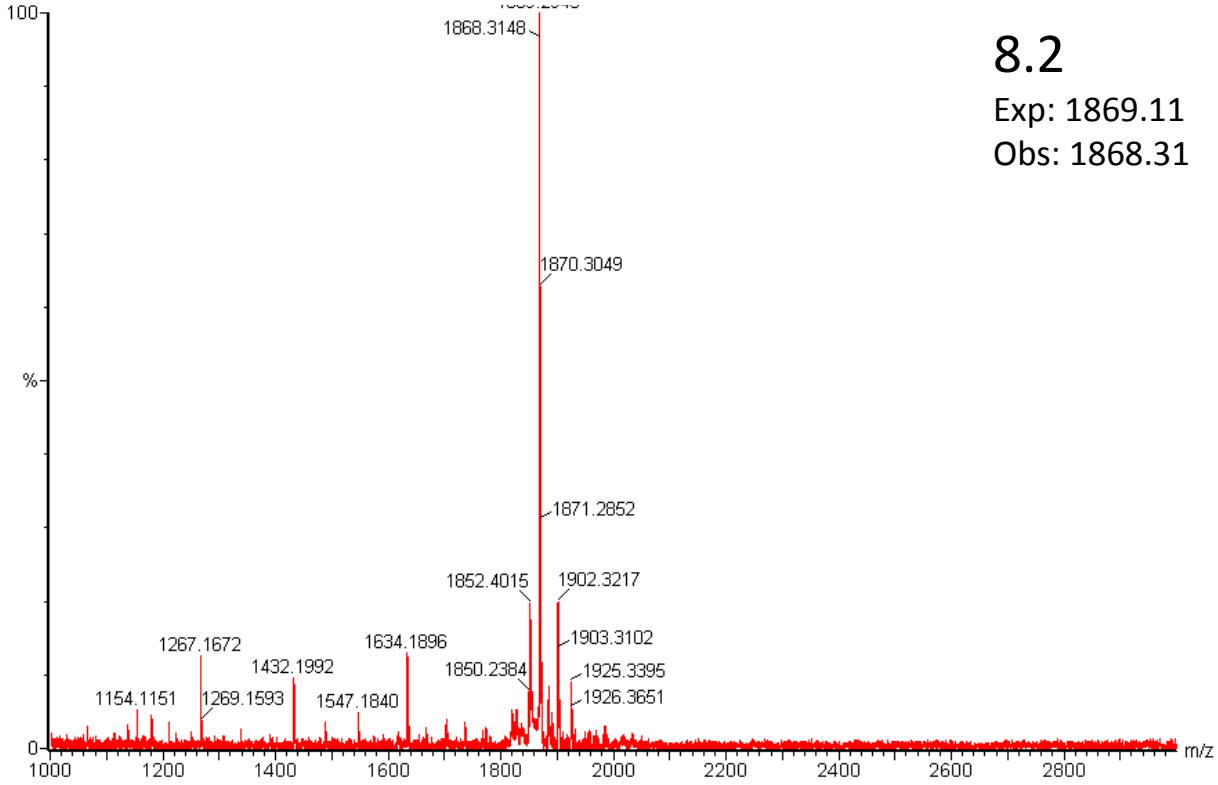
135. (a) Nomine, Y.; Botuyan, M. V.; Bajzer, Z.; Owen, W. G.; Caride, A. J.; Wasielewski, E.; Mer, G., Kinetic analysis of interaction of BRCA1 tandem breast cancer c-terminal domains with phosphorylated peptides reveals two binding conformations. *Biochem* **2008,** *47* (37), 9866-79; (b) Tischkowitz, M.; Hamel, N.; Carvalho, M. A.; Birrane, G.; Soni, A.; van Beers, E. H.; Joosse, S. A.; Wong, N.; Novak, D.; Quenneville, L. A.; Grist, S. A.; Nederlof, P. M.; Goldgar, D. E.; Tavtigian, S. V.; Monteiro, A. N.; Ladias, J. A.; Foulkes, W. D., Pathogenicity of the BRCA1 missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach. *Eur J Hum Genet* **2008,** *16* (7), 820-32; (c) Shiozaki, E. N.; Gu, L.; Yan, N.; Shi, Y., Structure of the BRCT repeats of BRCA1 bound to a BACH1 phosphopeptide: implications for signaling. *Molecular cell* **2004,** *14* (3), 405-12.

136. (a) Donaldson, G. P.; Roelofs, K. G.; Luo, Y.; Sintim, H. O.; Lee, V. T., A rapid assay for affinity and kinetics of molecular interactions with nucleic acids. *Nucleic Acids Res* **2012,** *40* (7), e48; (b) Roelofs, K. G.; Wang, J.; Sintim, H. O.; Lee, V. T., Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc Nat Acad Sci USA* **2011,** *108* (37), 15528-33.

Appendix I MALDI-TOF Analysis of Peptides



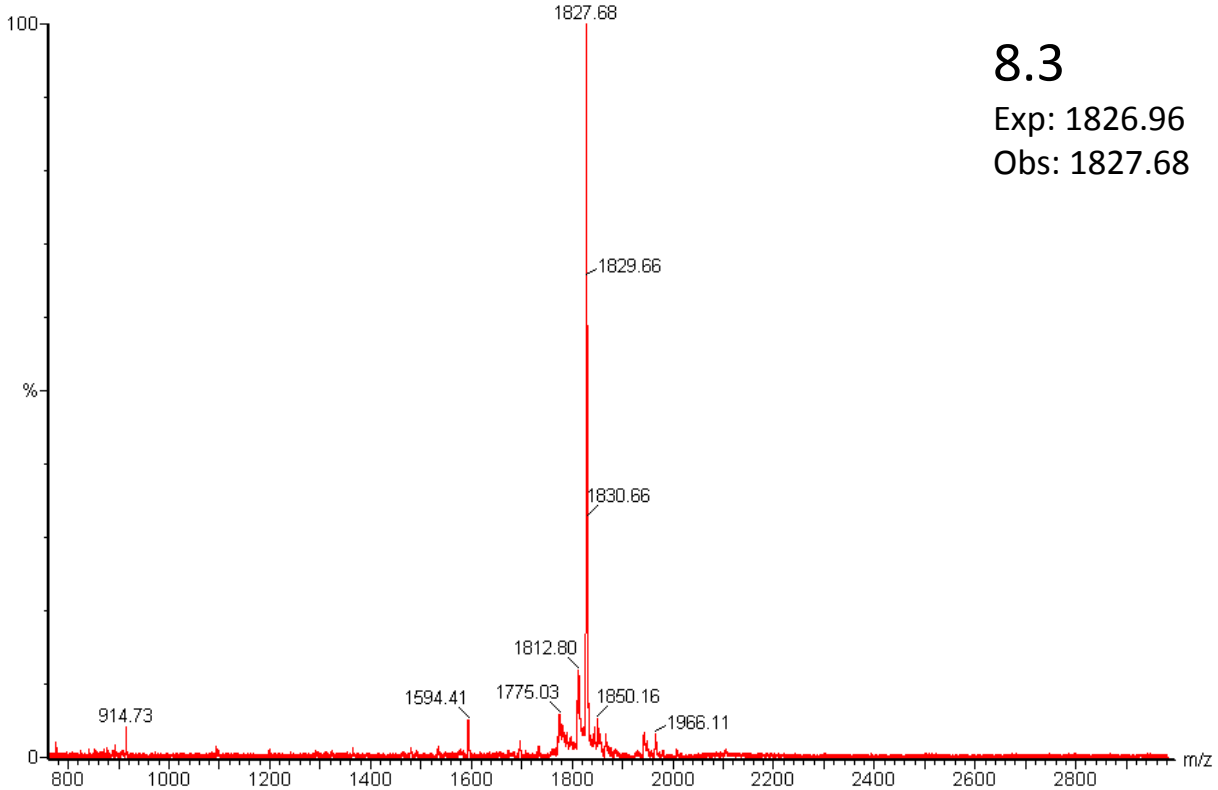




8.2

Exp: 1869.11

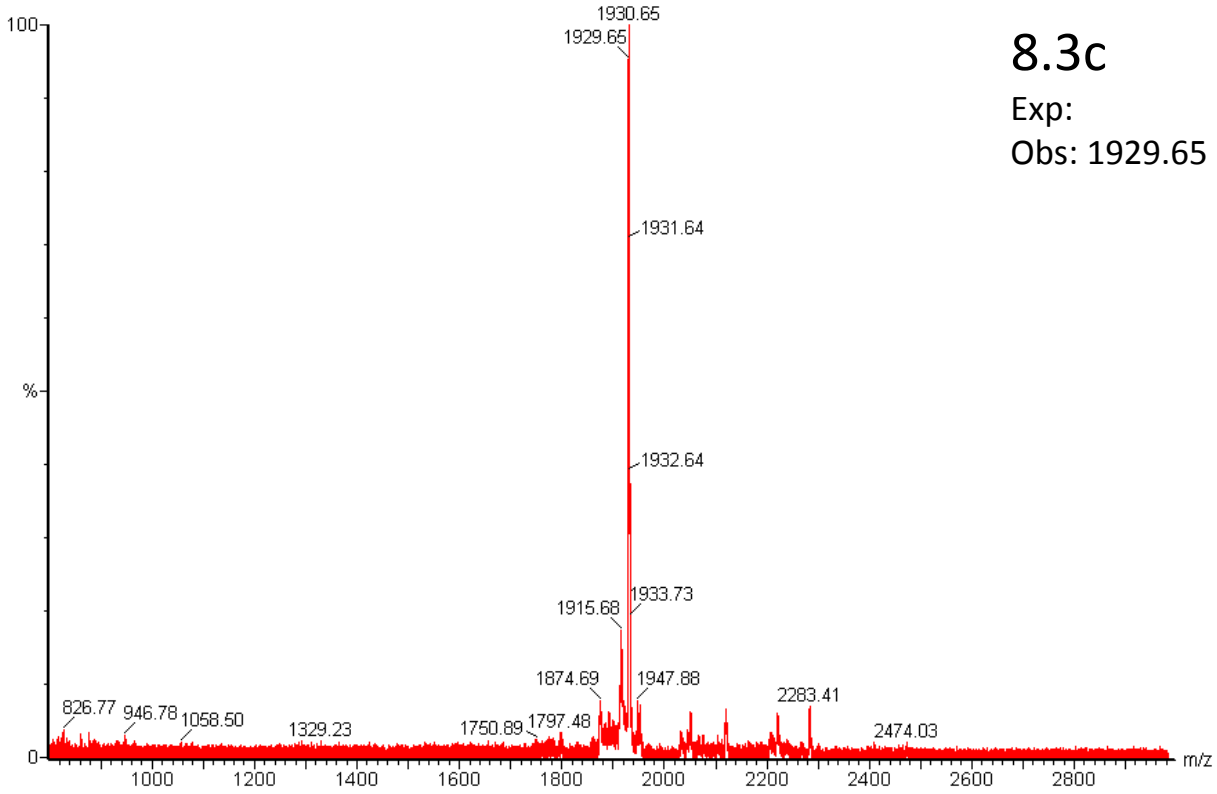
Obs: 1868.31



8.3

Exp: 1826.96

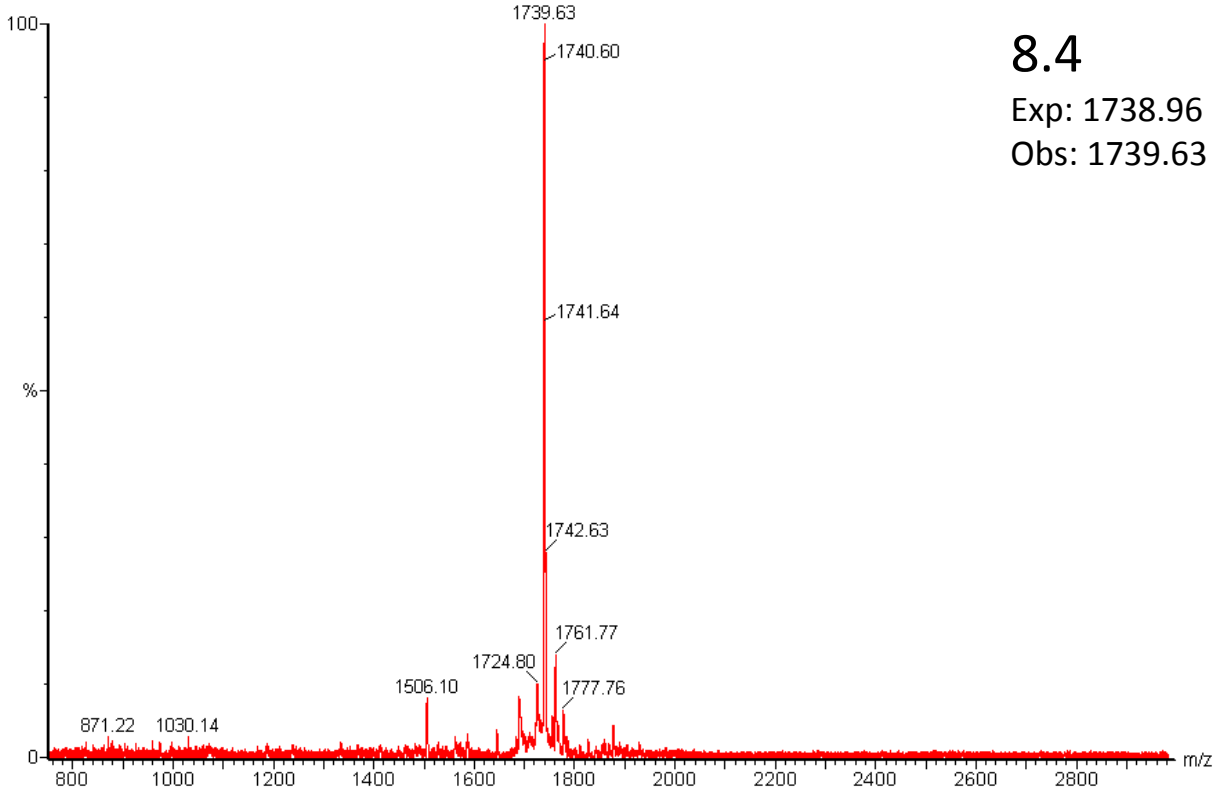
Obs: 1827.68



8.3c

Exp:

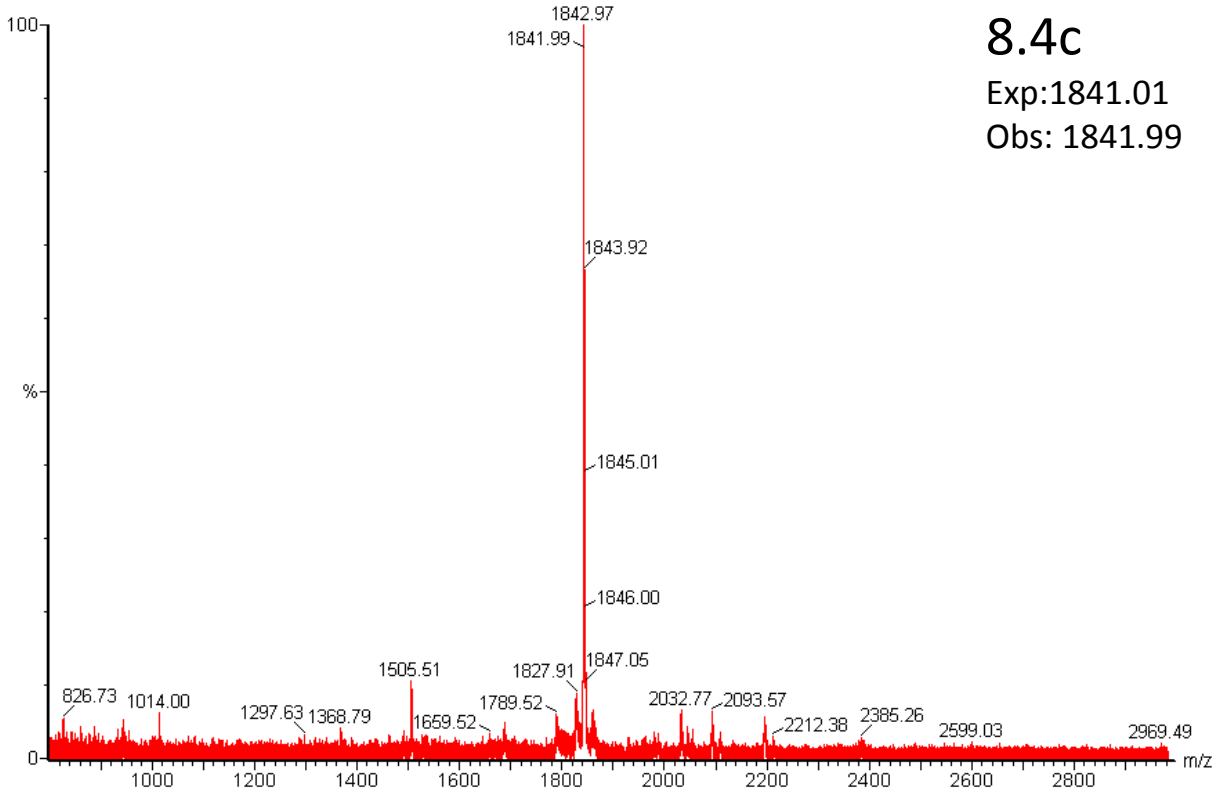
Obs: 1929.65



8.4

Exp: 1738.96

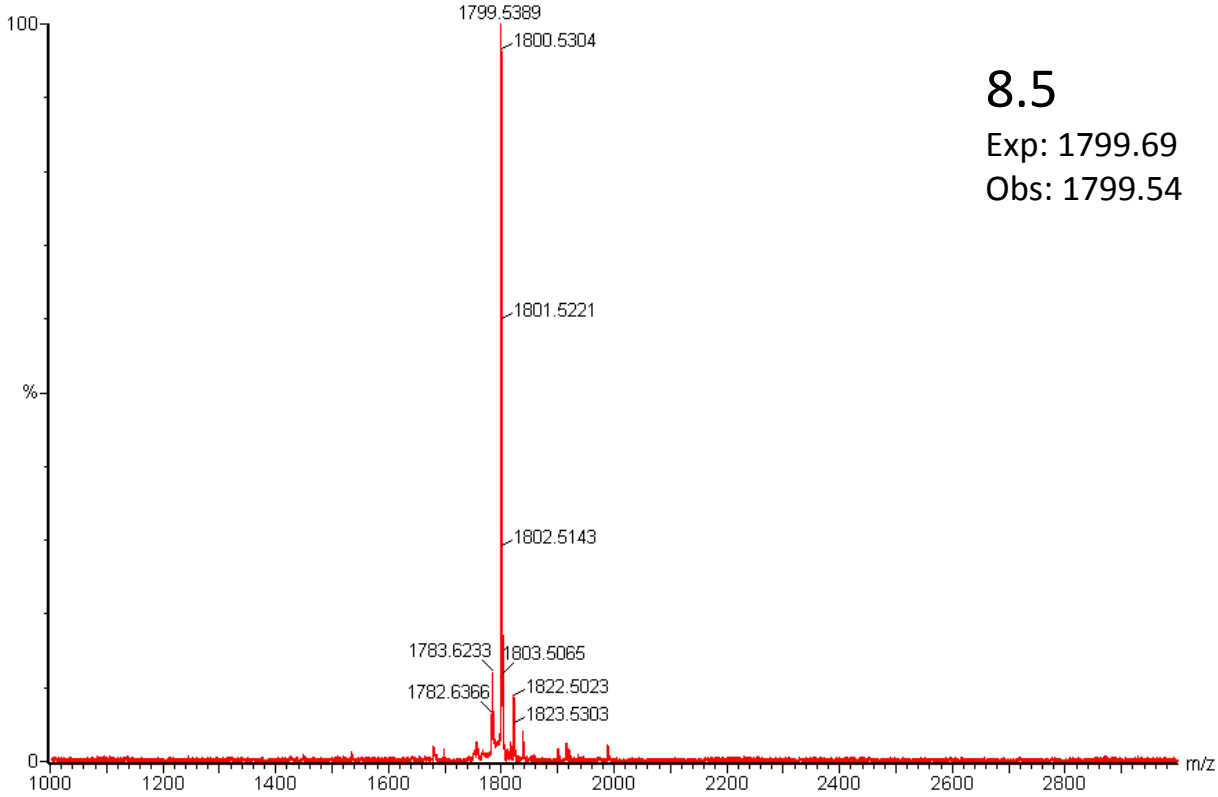
Obs: 1739.63



8.4c

Exp:1841.01

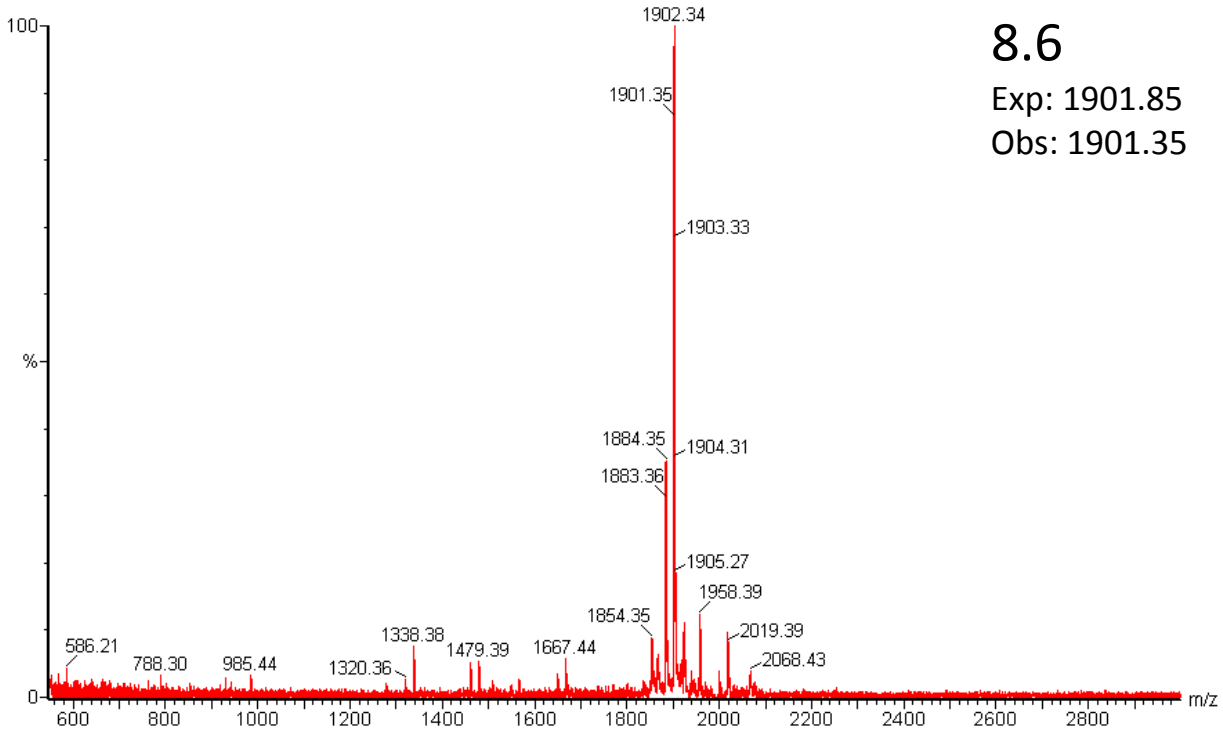
Obs: 1841.99



8.5

Exp: 1799.69

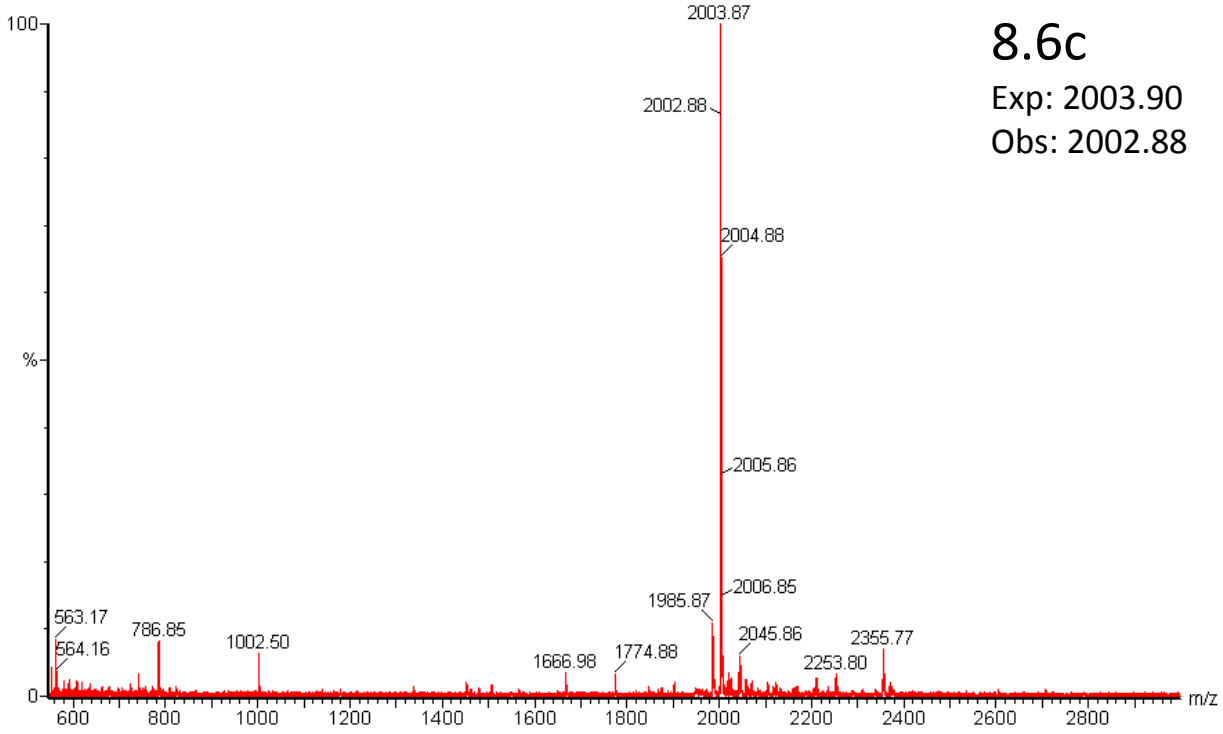
Obs: 1799.54



8.6

Exp: 1901.85

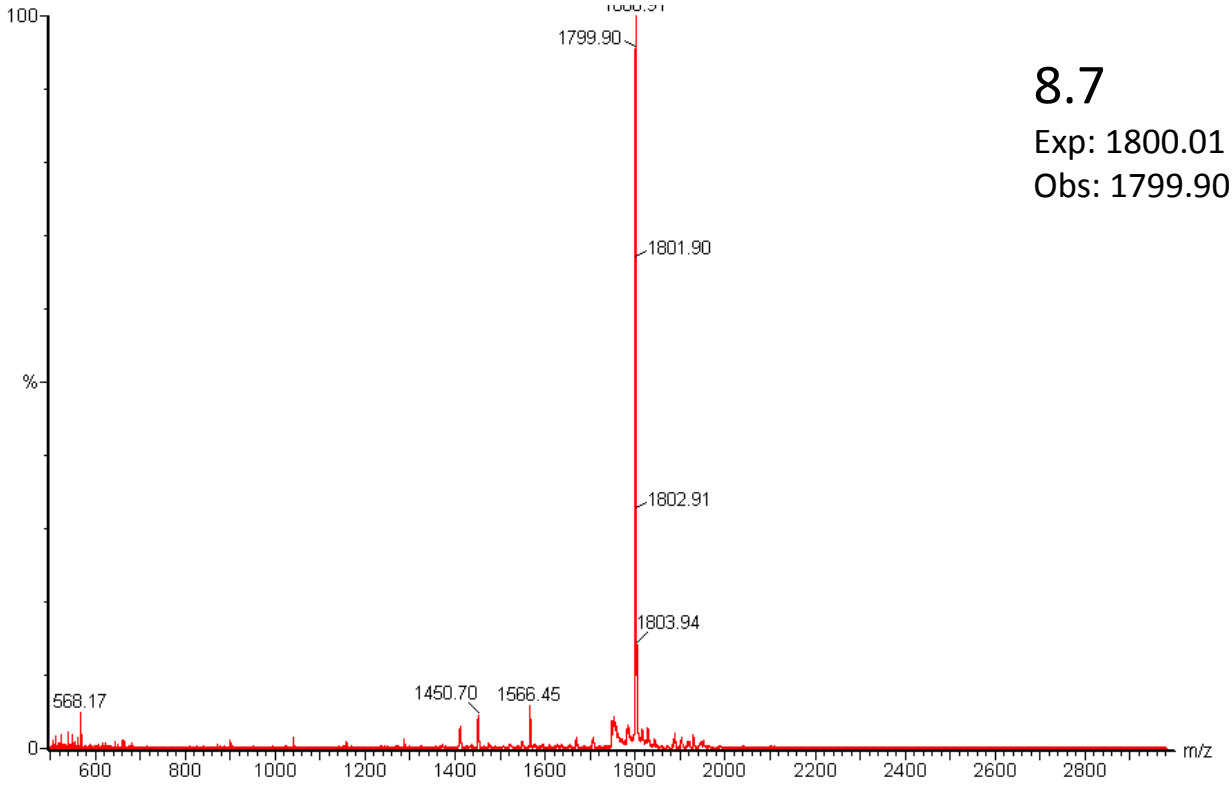
Obs: 1901.35



8.6c

Exp: 2003.90

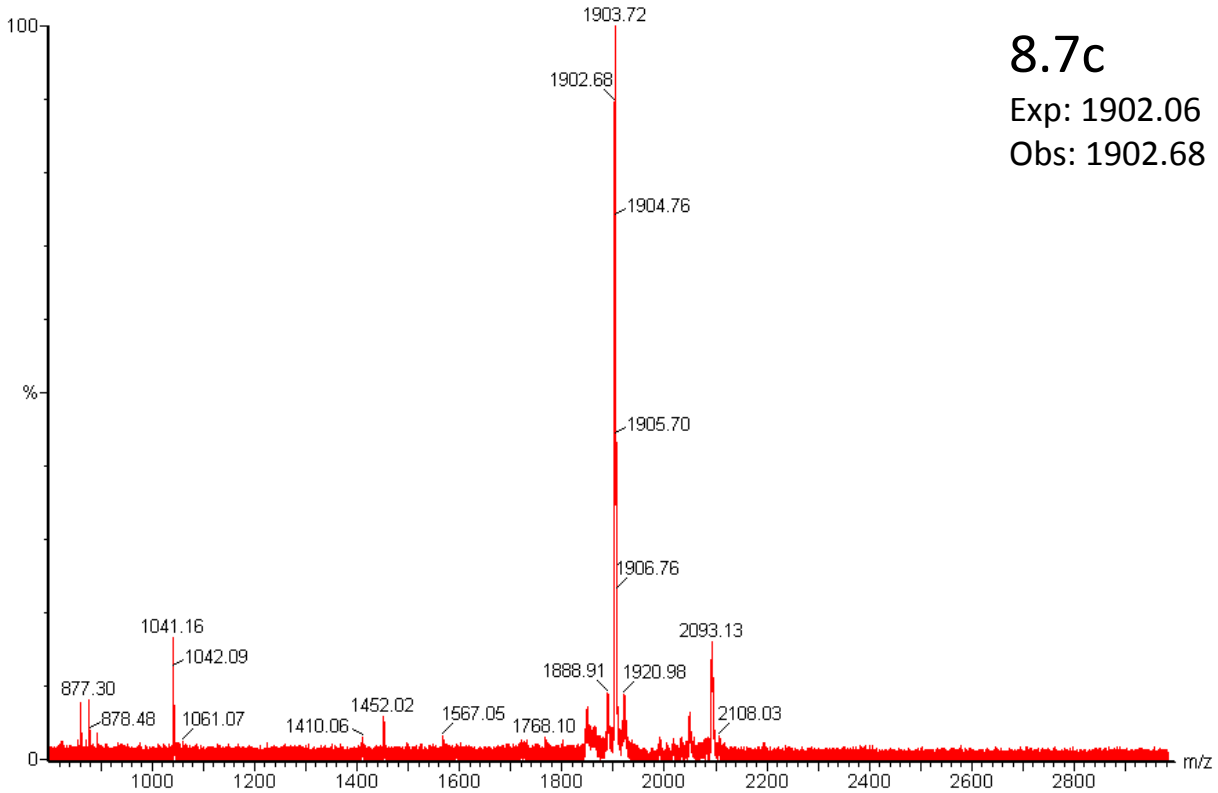
Obs: 2002.88



8.7

Exp: 1800.01

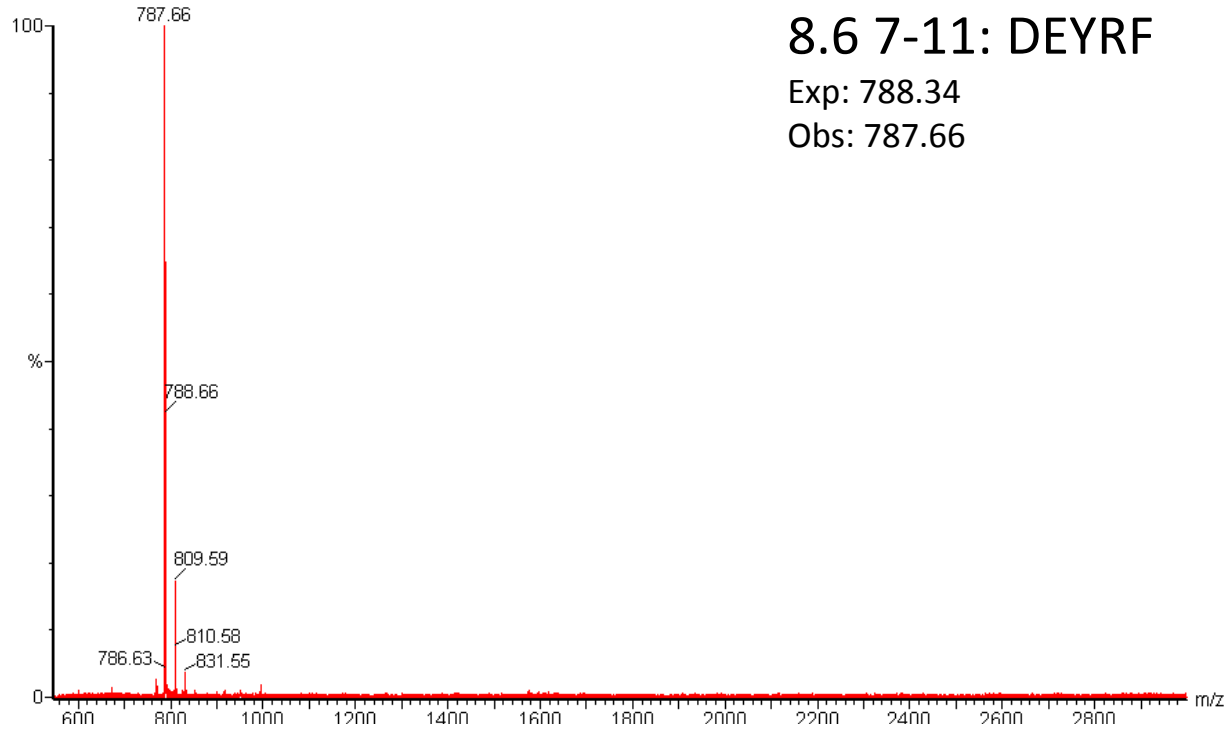
Obs: 1799.90



8.7c

Exp: 1902.06

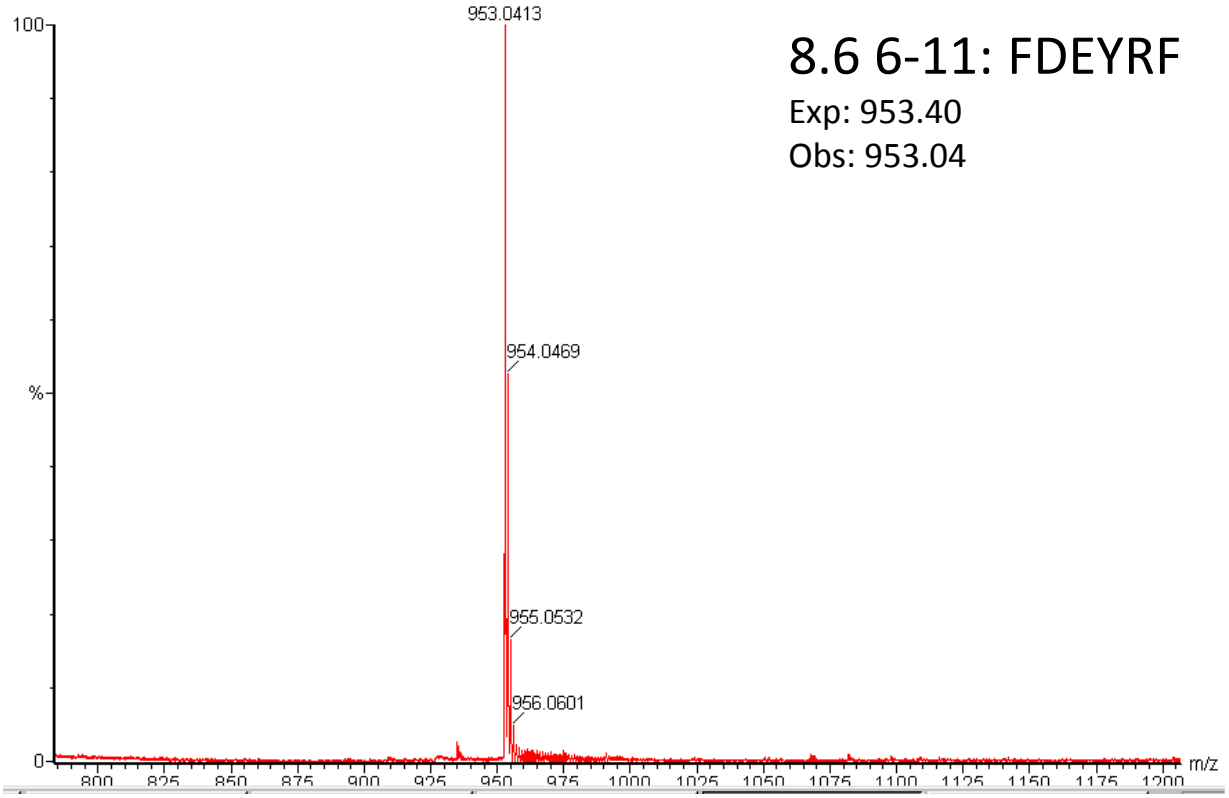
Obs: 1902.68



8.6 7-11: DEYRF

Exp: 788.34

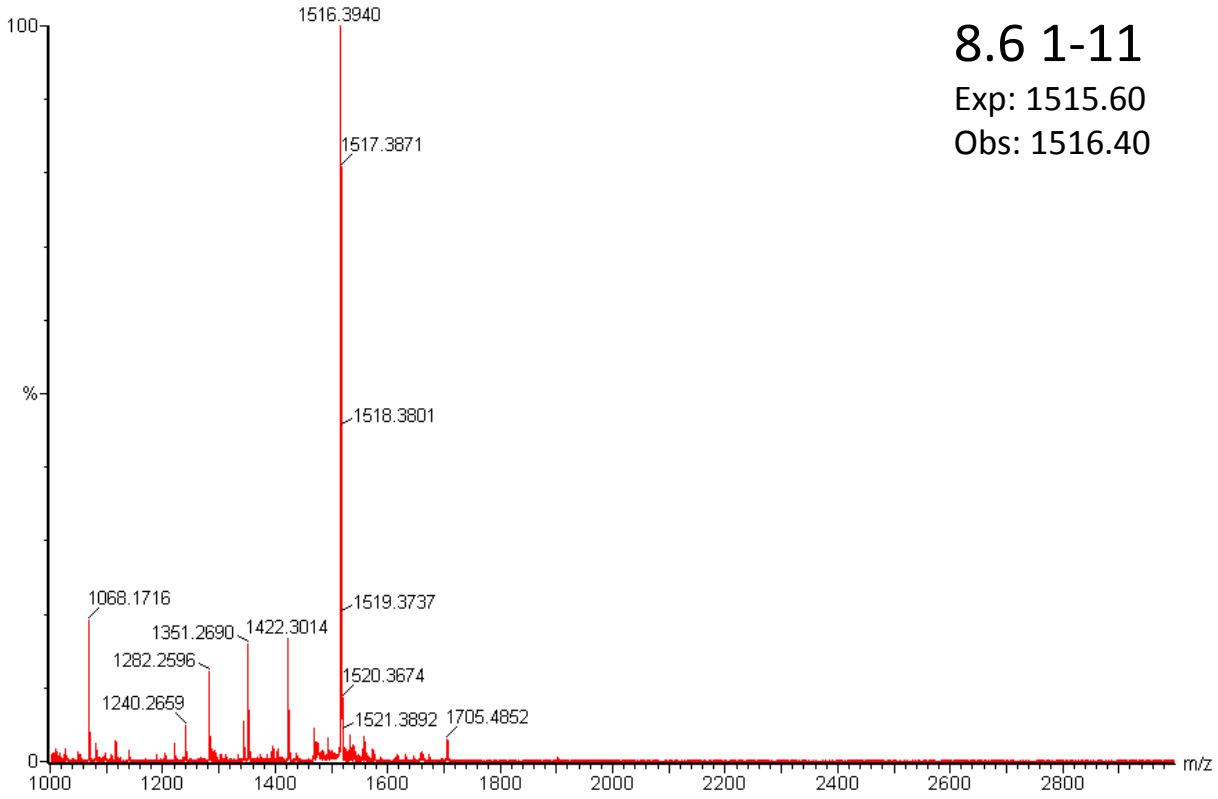
Obs: 787.66



8.6 6-11: FDEYRF

Exp: 953.40

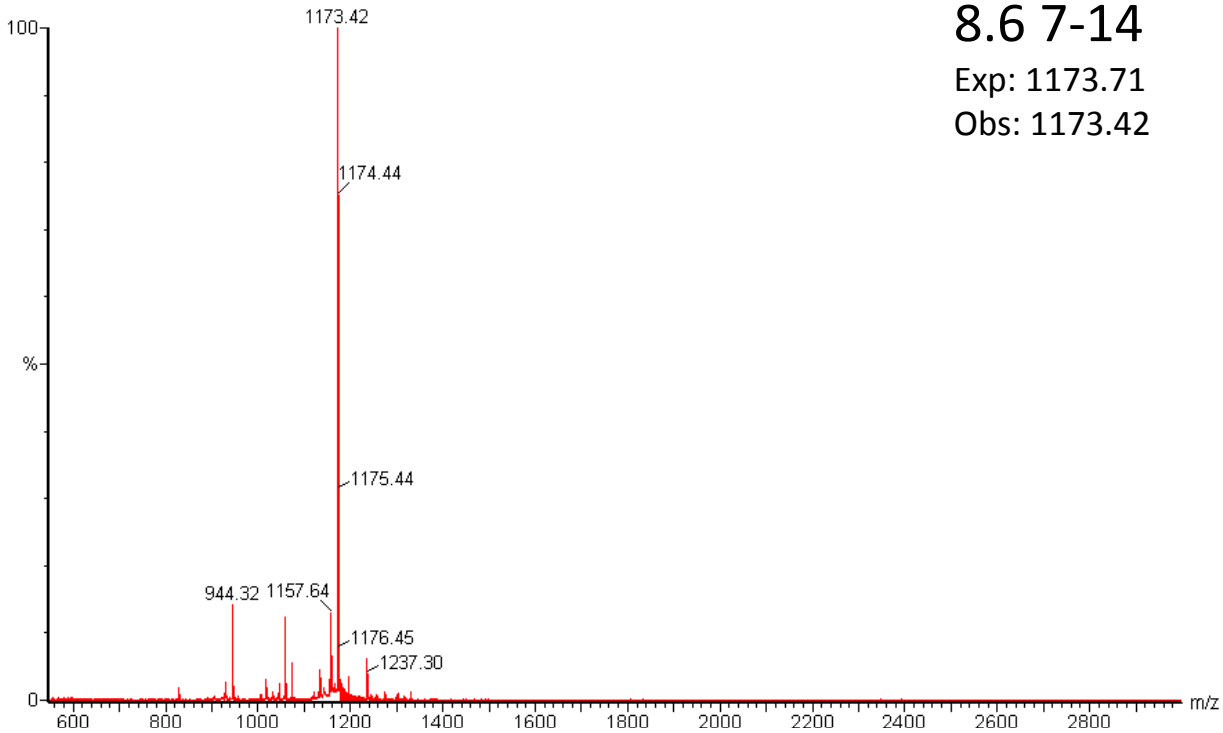
Obs: 953.04



8.6 1-11

Exp: 1515.60

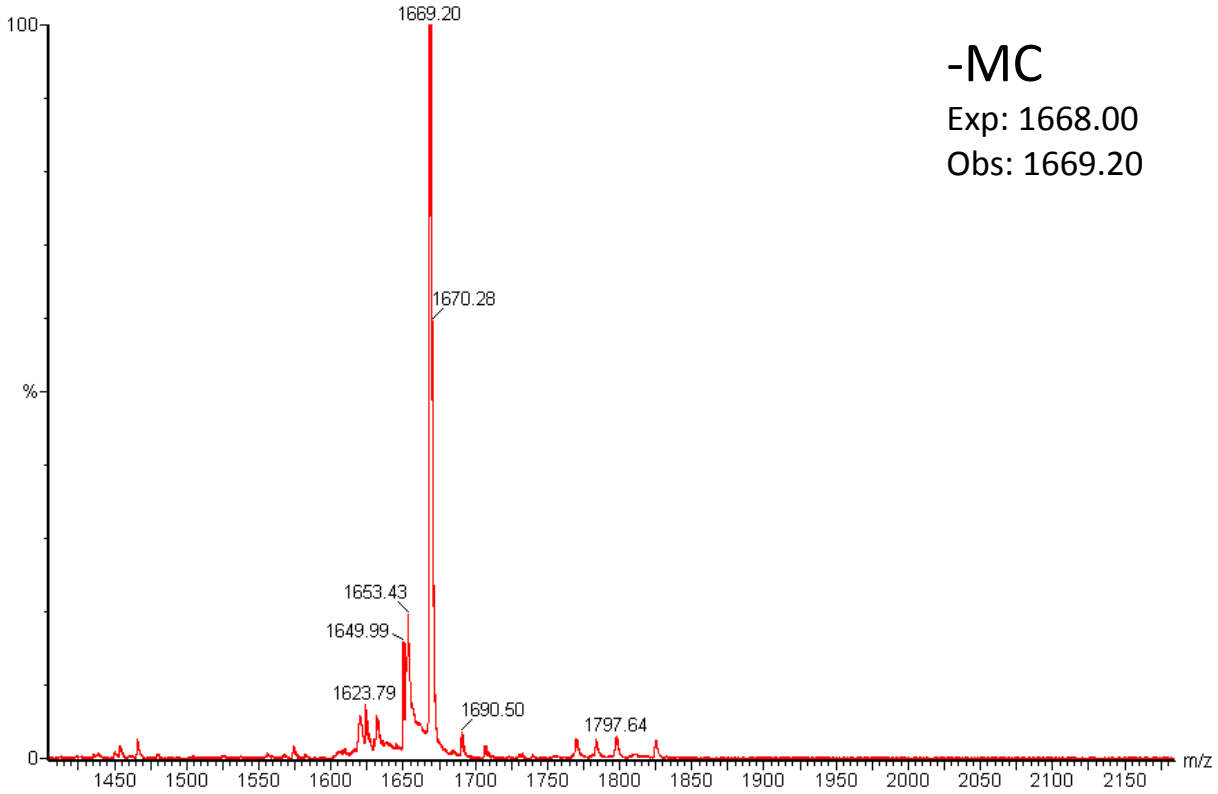
Obs: 1516.40



8.6 7-14

Exp: 1173.71

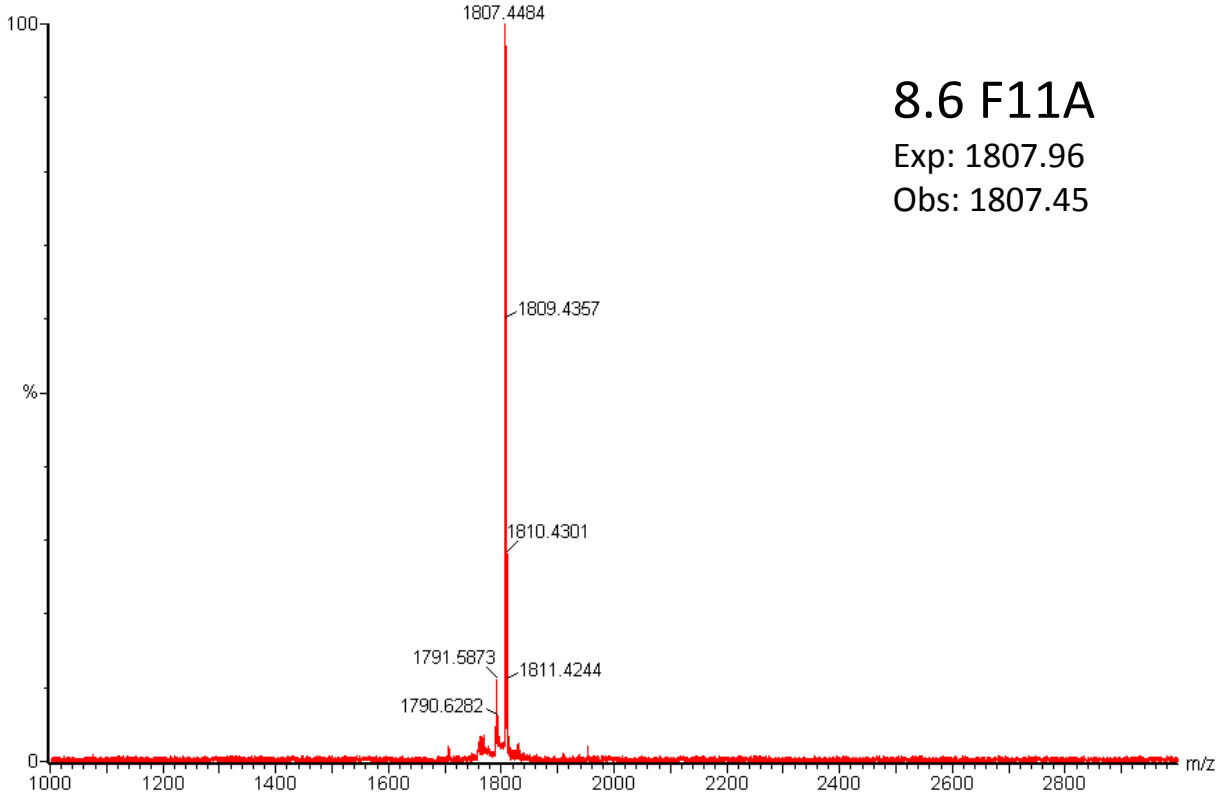
Obs: 1173.42



-MC

Exp: 1668.00

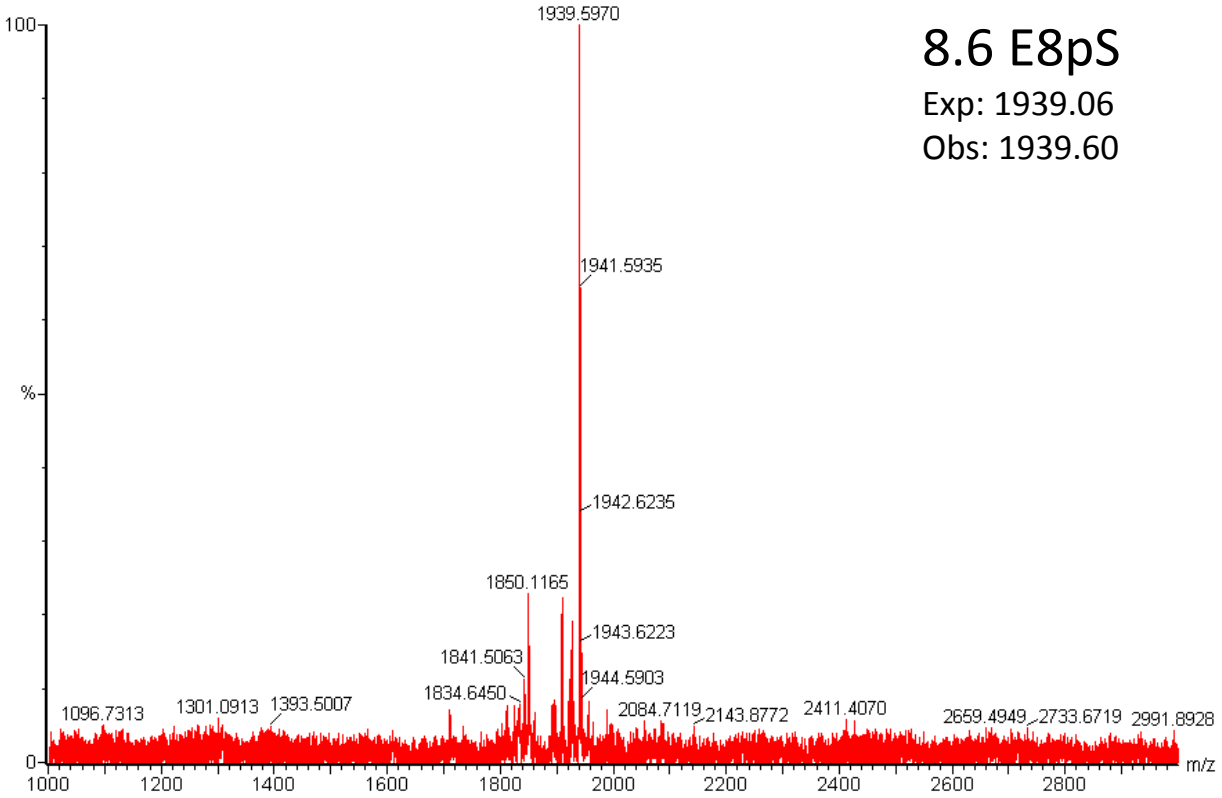
Obs: 1669.20

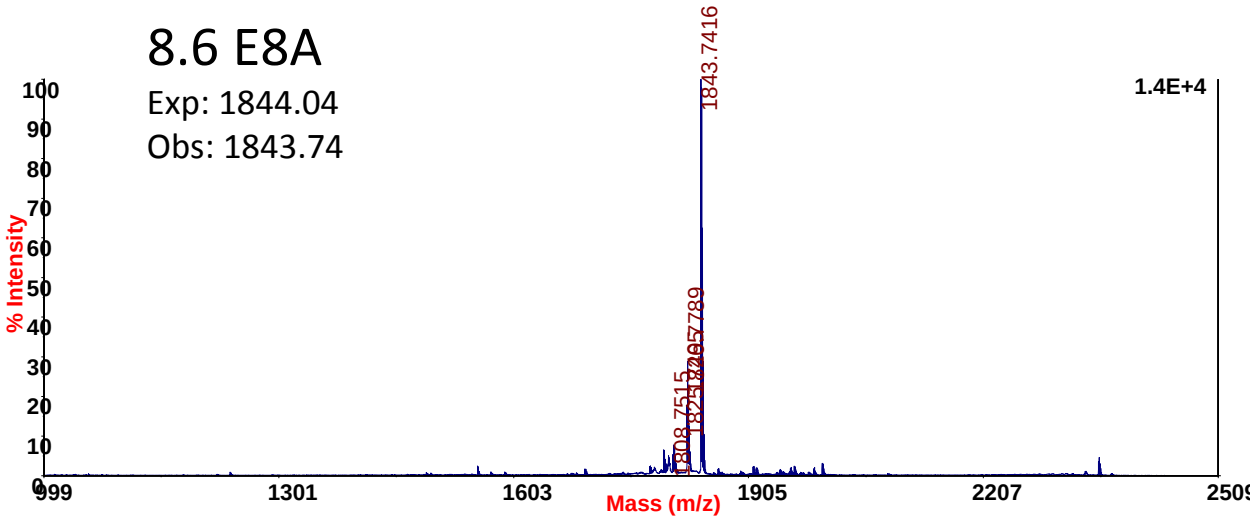


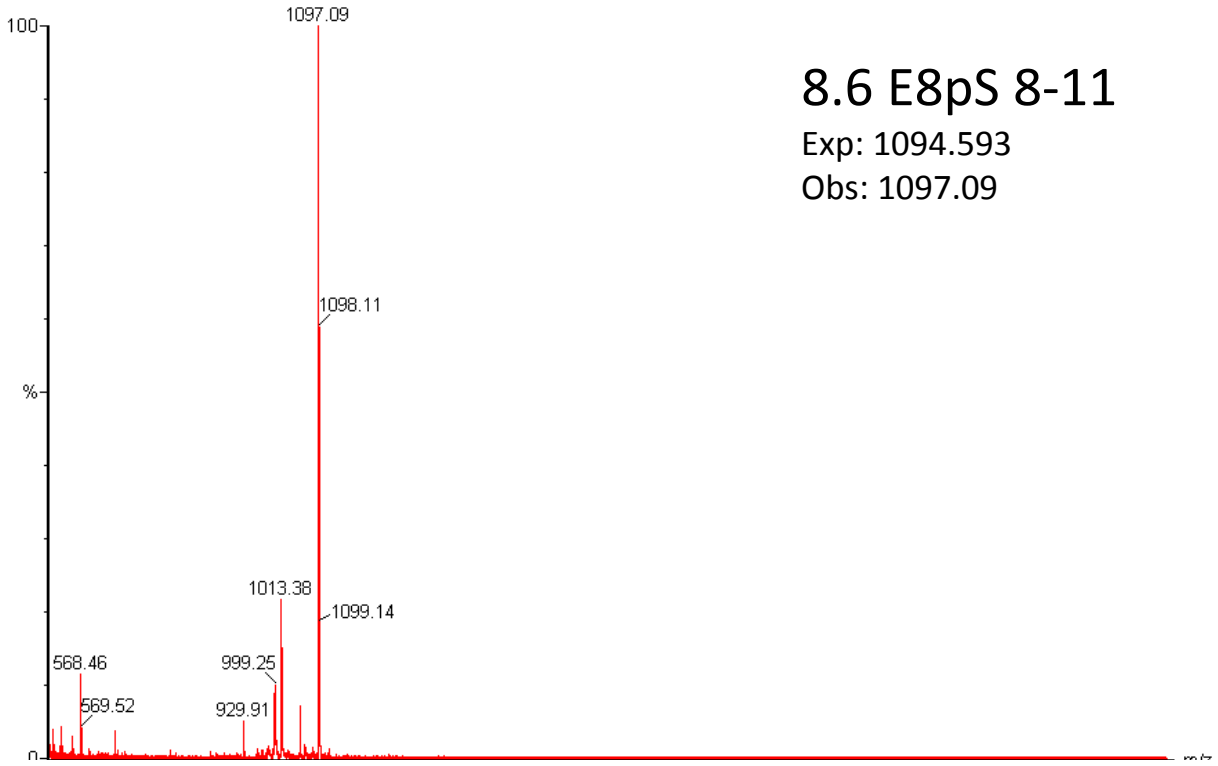
8.6 F11A

Exp: 1807.96

Obs: 1807.45


8.6 E8pS
Exp: 1939.06
Obs: 1939.60
1939.5970
1941.5935
1942.6235
1850.1165
1943.6223
1841.5063
1944.5903
1834.6450
1096.7313
1301.0913
1393.5007
2084.7119
2143.8772
2411.4070
2659.4949
2733.6719
2991.8928
100
%
0
m/z
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800

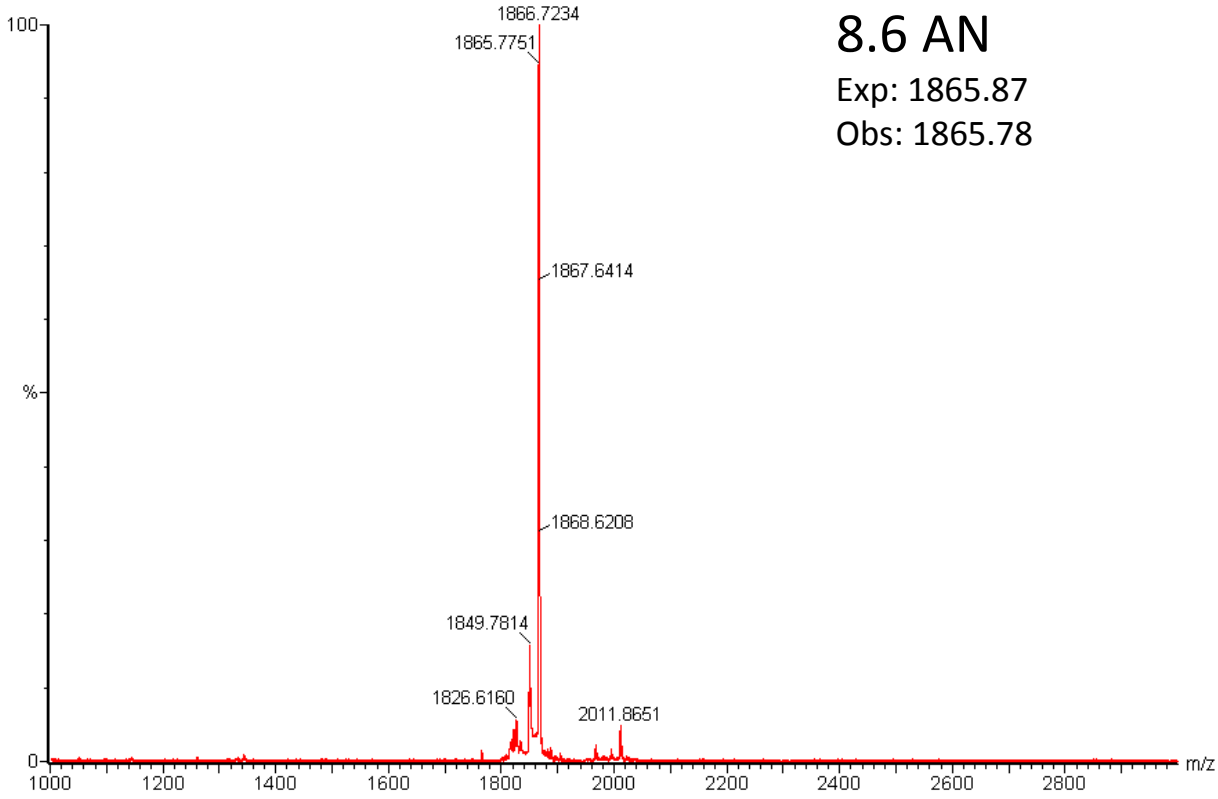

8.6 E8A
Exp: 1844.04
Obs: 1843.74
1843.7416
1.4E+4
% Intensity
Mass (m/z)
100
90
80
70
60
50
40
30
20
10
0
999
1301
1603
1905
2207



8.6 E8pS 8-11

Exp: 1094.593

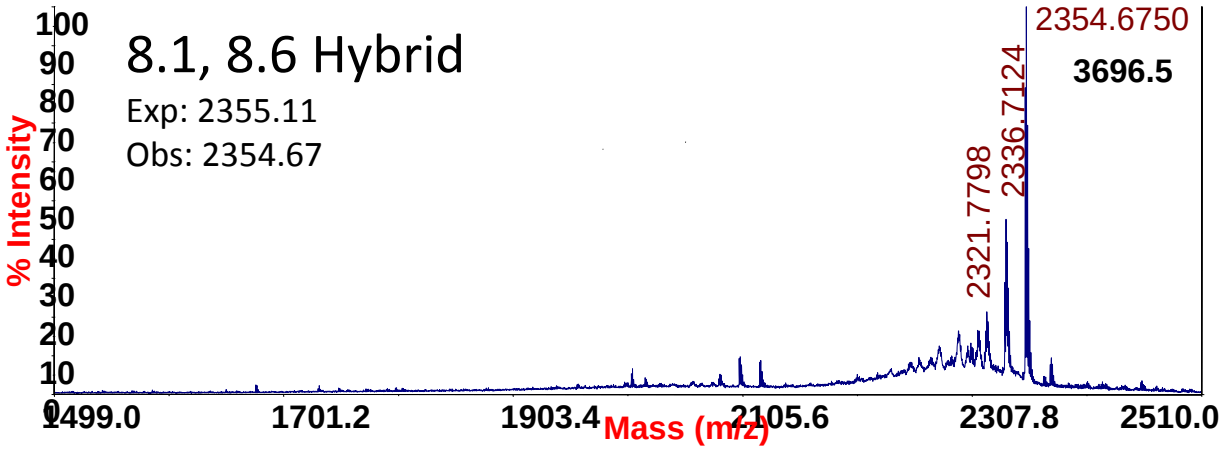
Obs: 1097.09

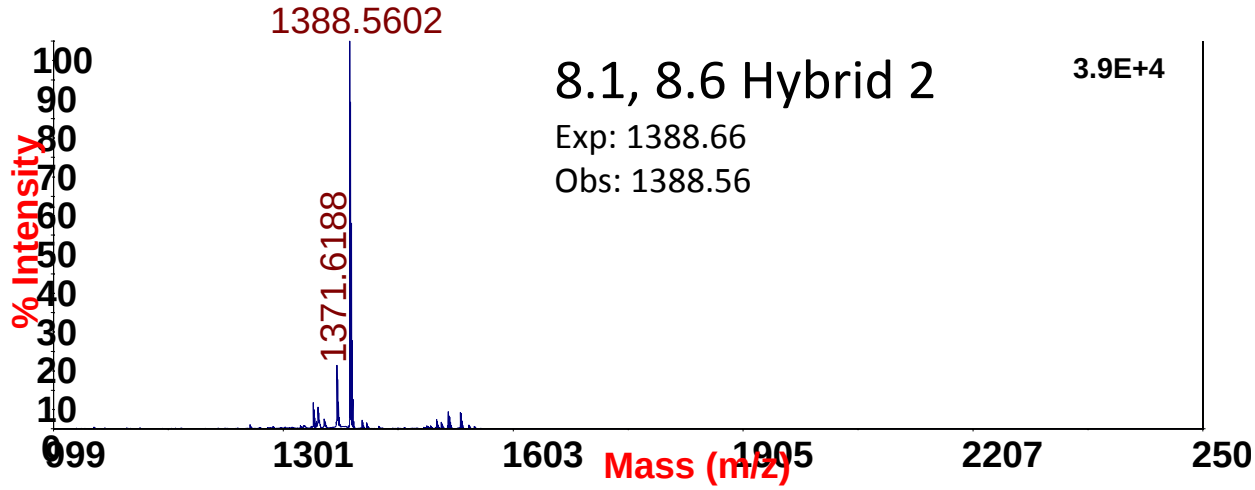


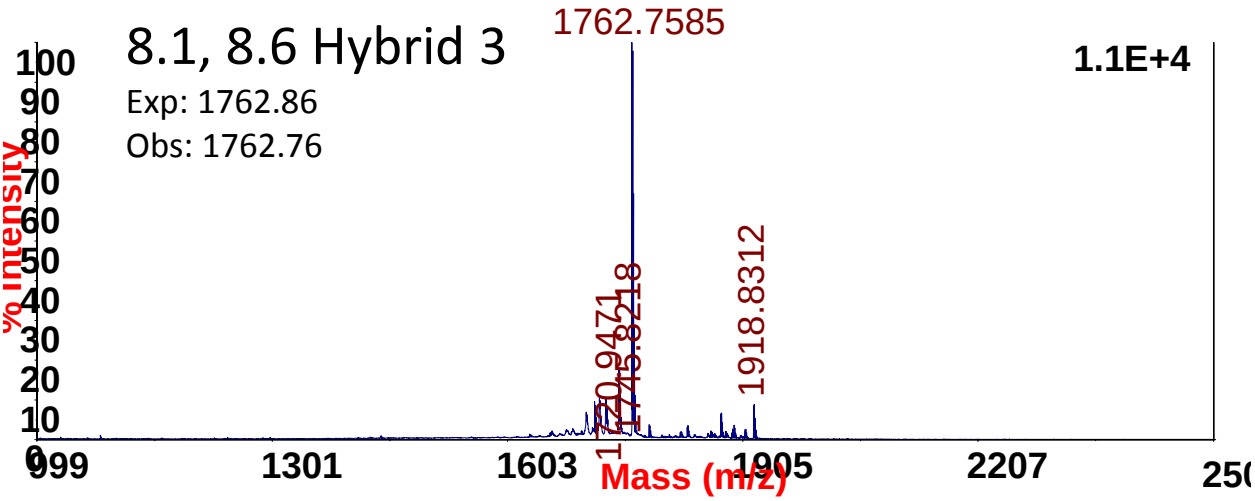
8.6 AN

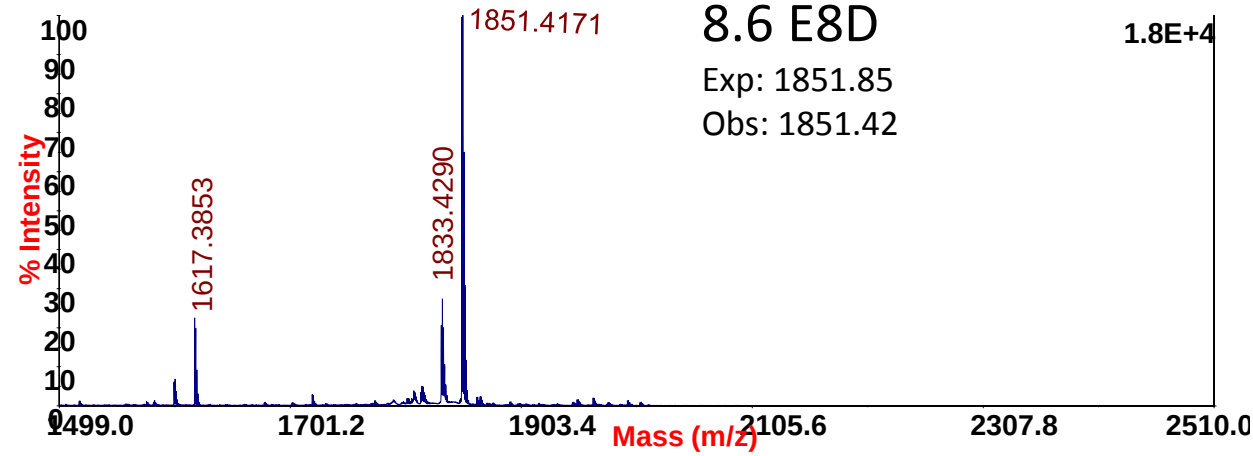
Exp: 1865.87

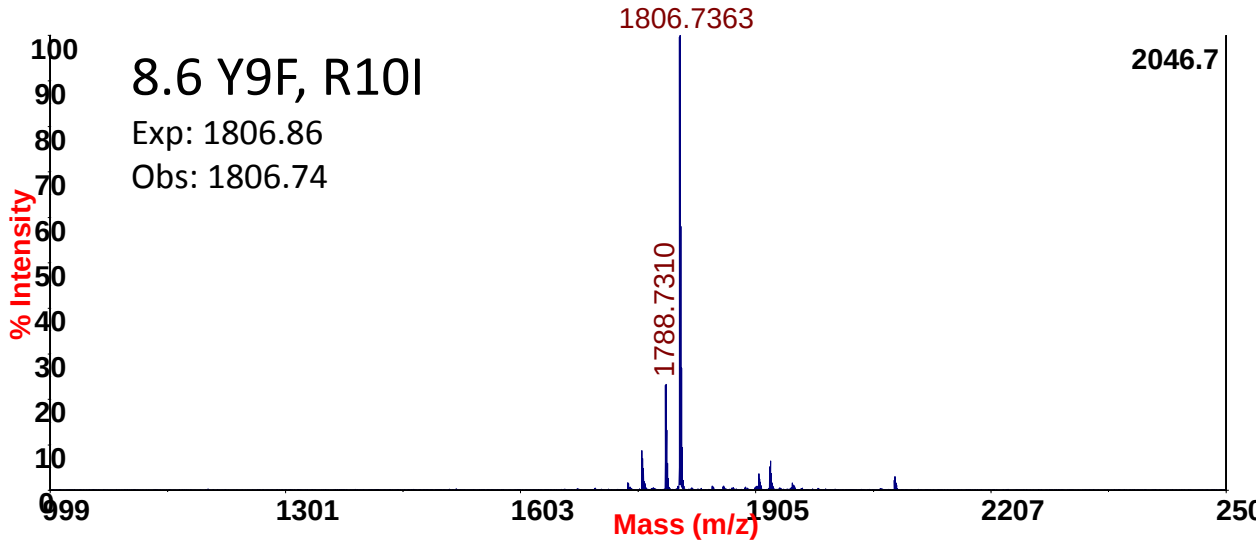
Obs: 1865.78

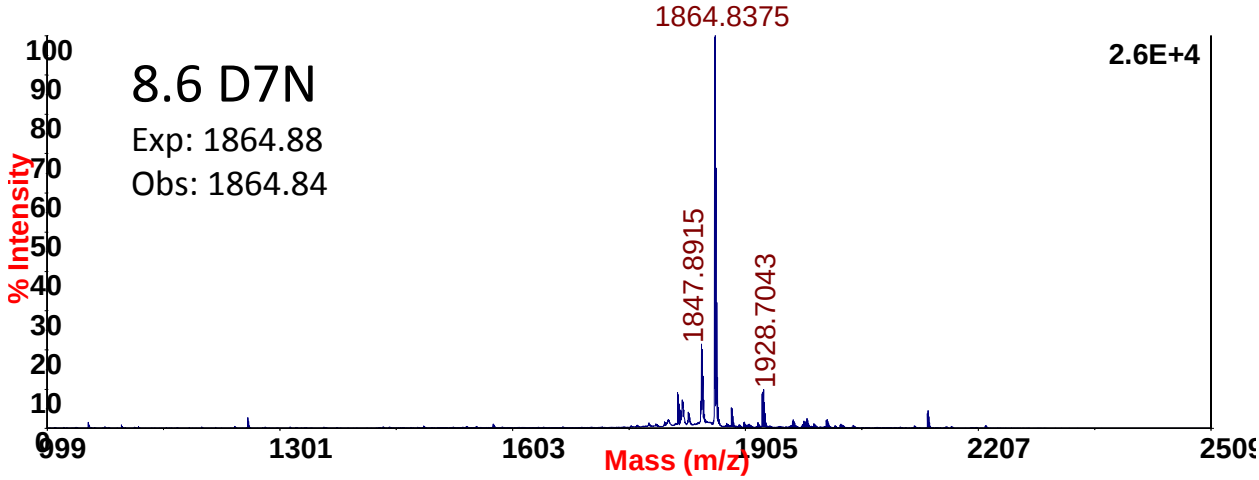

8.1, 8.6 Hybrid
Exp: 2355.11
Obs: 2354.67
2354.6750
3696.5
2336.7124
2321.7798
% Intensity
Mass (m/z)
1499.0
1701.2
1903.4
2105.6
2307.8
2510.0

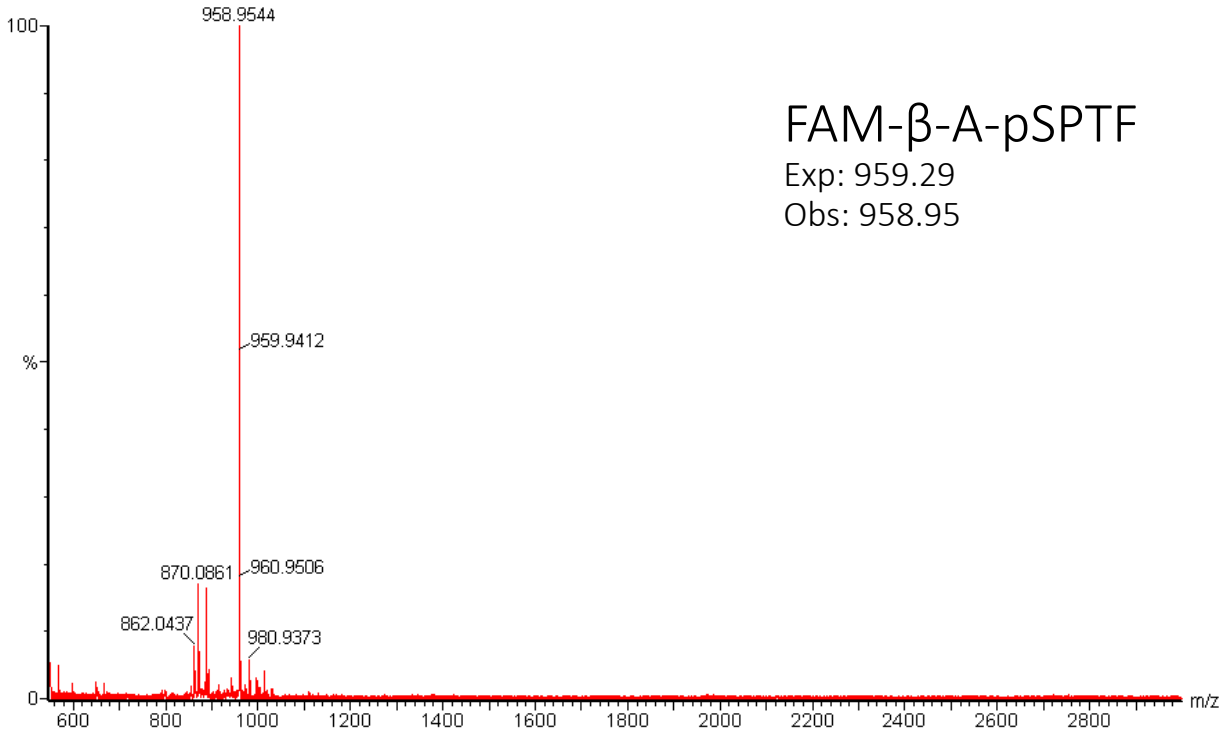

1388.5602
8.1, 8.6 Hybrid 2
Exp: 1388.66
Obs: 1388.56
3.9E+4
1371.6188
% Intensity
Mass (m/z)
999
1301
1603
1905
2207


8.1, 8.6 Hybrid 3
Exp: 1762.86
Obs: 1762.76
1762.7585
1.1E+4
1918.8312
% Intensity
Mass (m/z)
999
1301
1603
1905
2207


8.6 E8D
Exp: 1851.85
Obs: 1851.42
1851.4171
1833.4290
1617.3853
% Intensity
Mass (m/z)
1.8E+4


8.6 Y9F, R10I
Exp: 1806.86
Obs: 1806.74
1806.7363
1788.7310
% Intensity
Mass (m/z)
2046.7


8.6 D7N
Exp: 1864.88
Obs: 1864.84
1864.8375
1847.8915
1928.7043
% Intensity
Mass (m/z)
2.6E+4

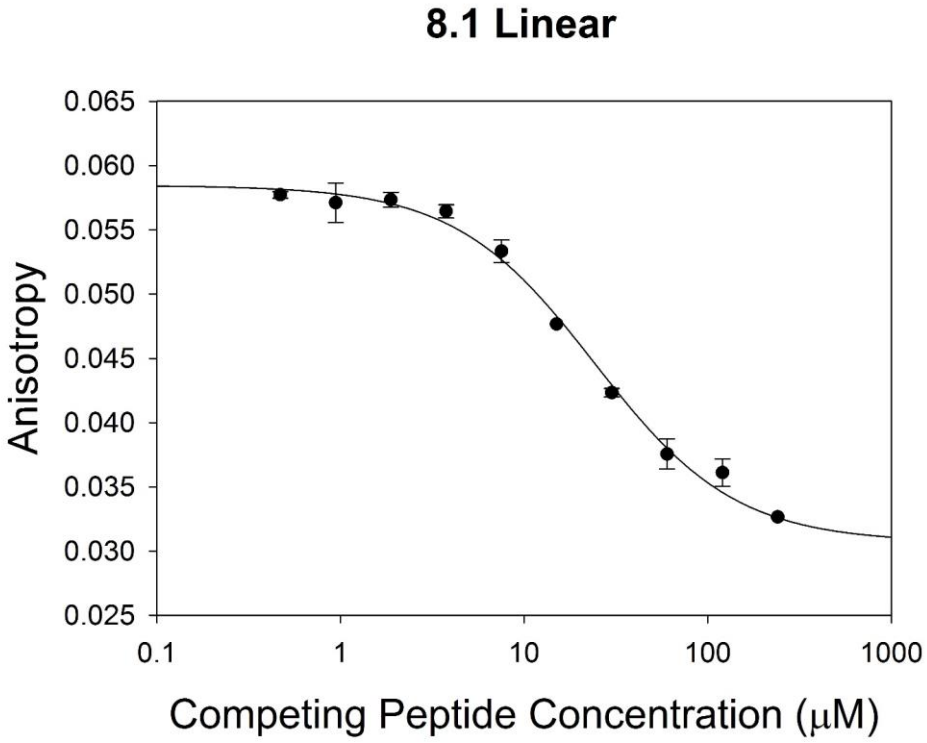


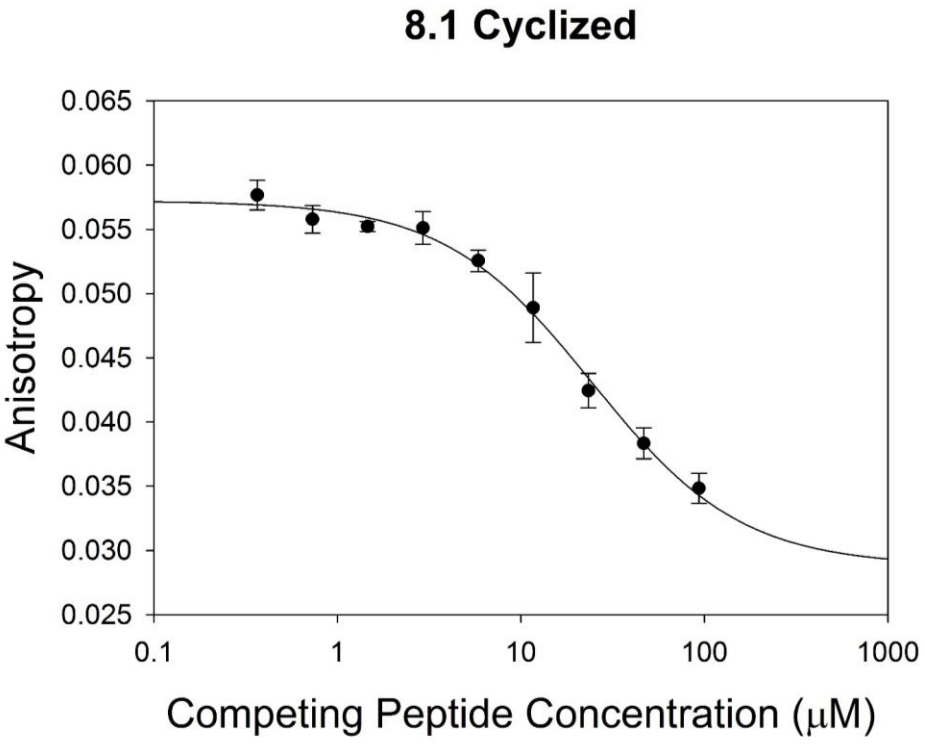
FAM-β-A-pSPTF

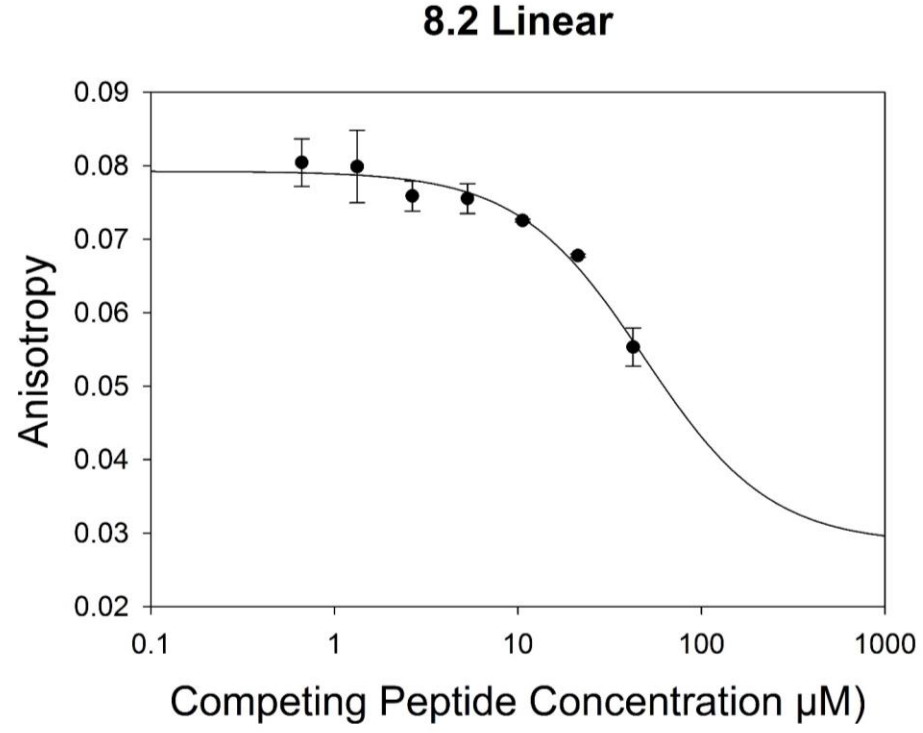
Exp: 959.29

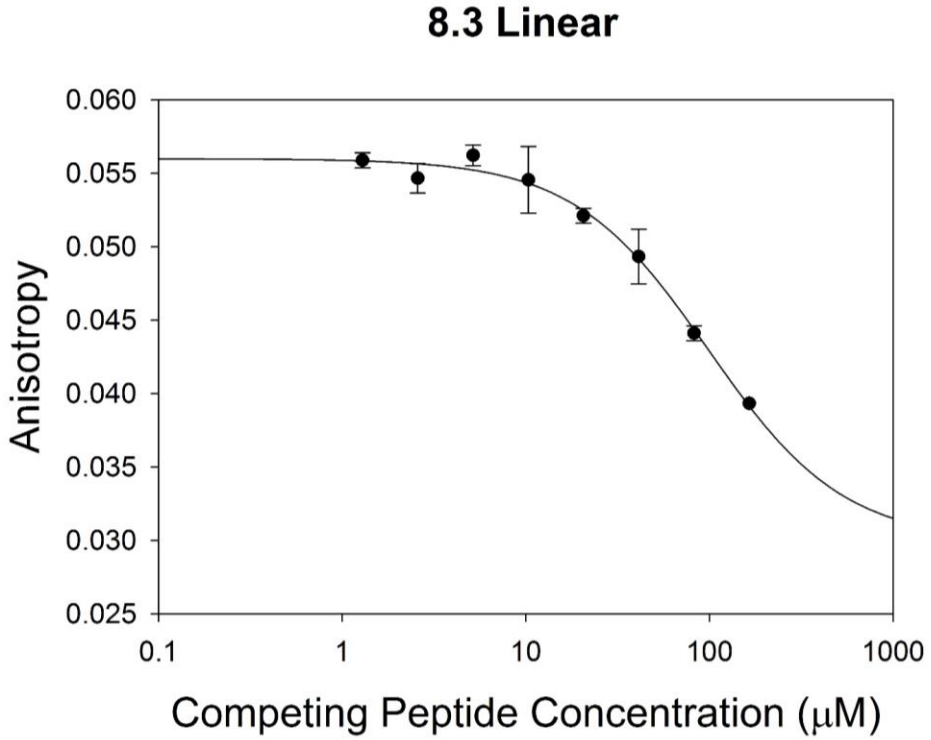
Obs: 958.95

Appendix II Fluorescence Polarization Binding Curves

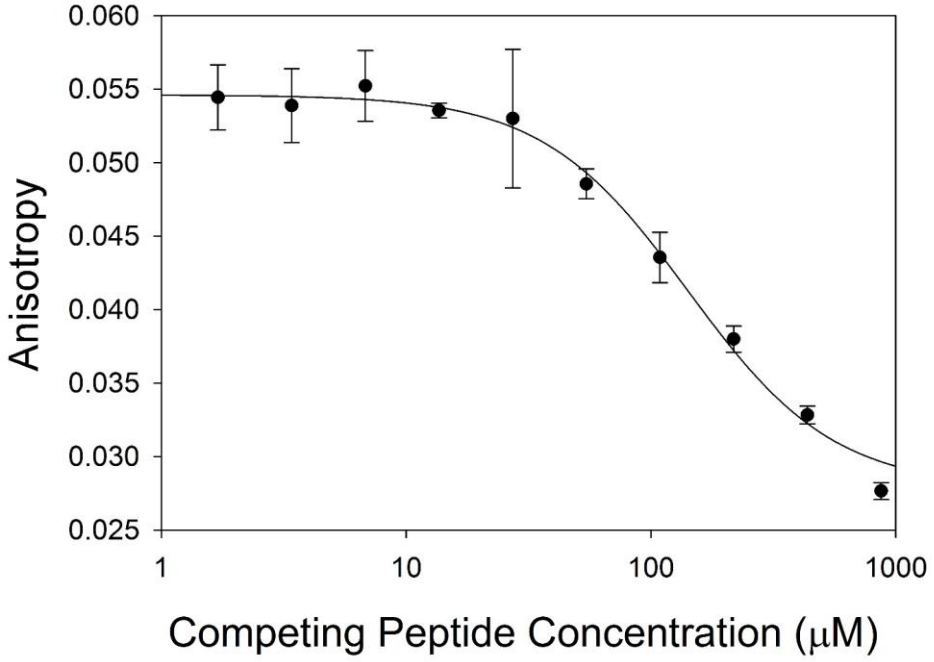




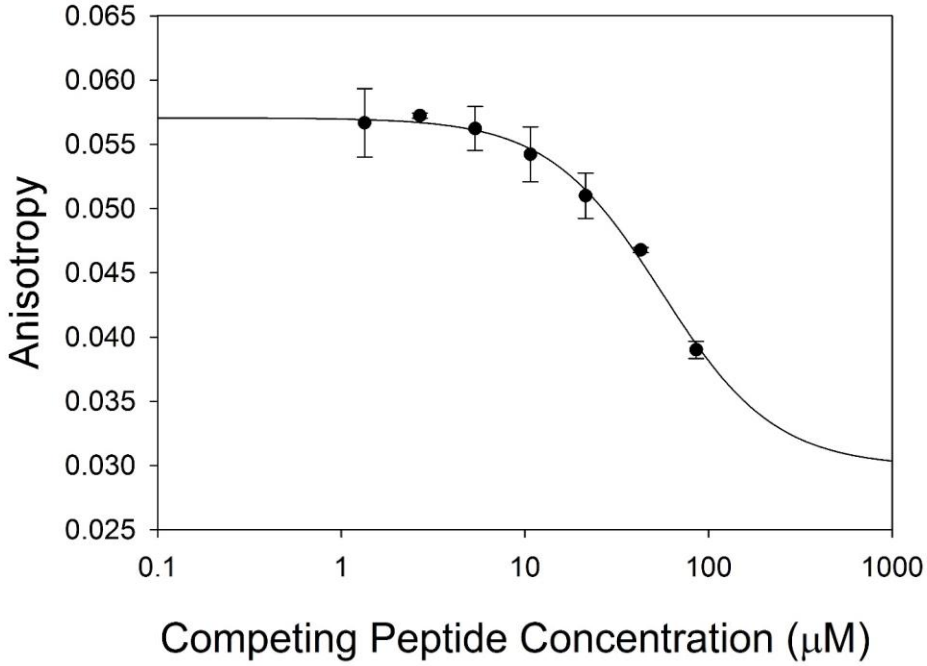

8.2 Linear
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
Anisotropy
0.1
1
10
100
1000
Competing Peptide Concentration µM)


8.3 Linear
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
Anisotropy
0.1
1
10
100
1000
Competing Peptide Concentration (µM)

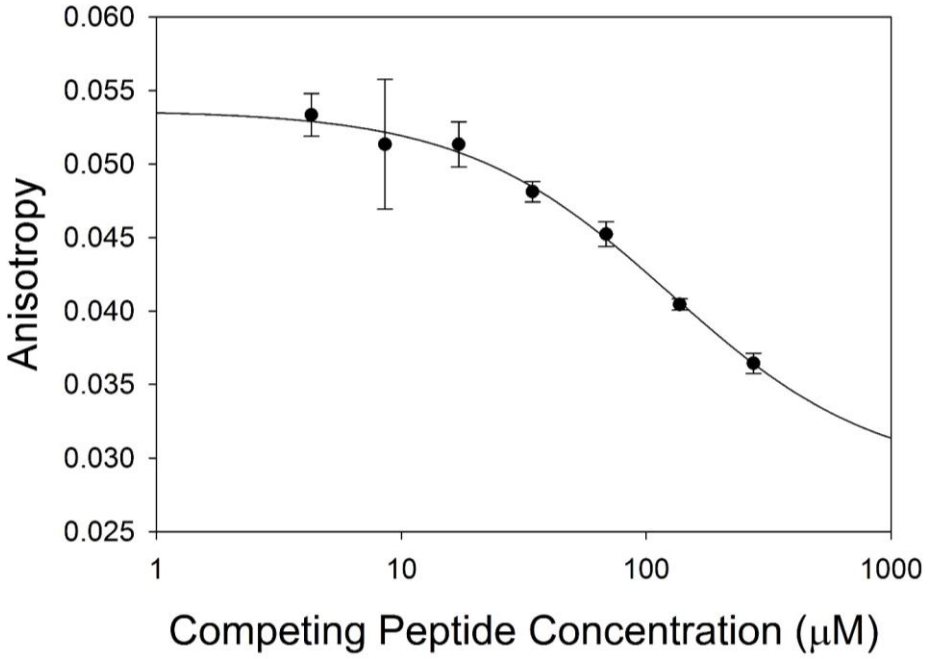
8.3 Cyclized



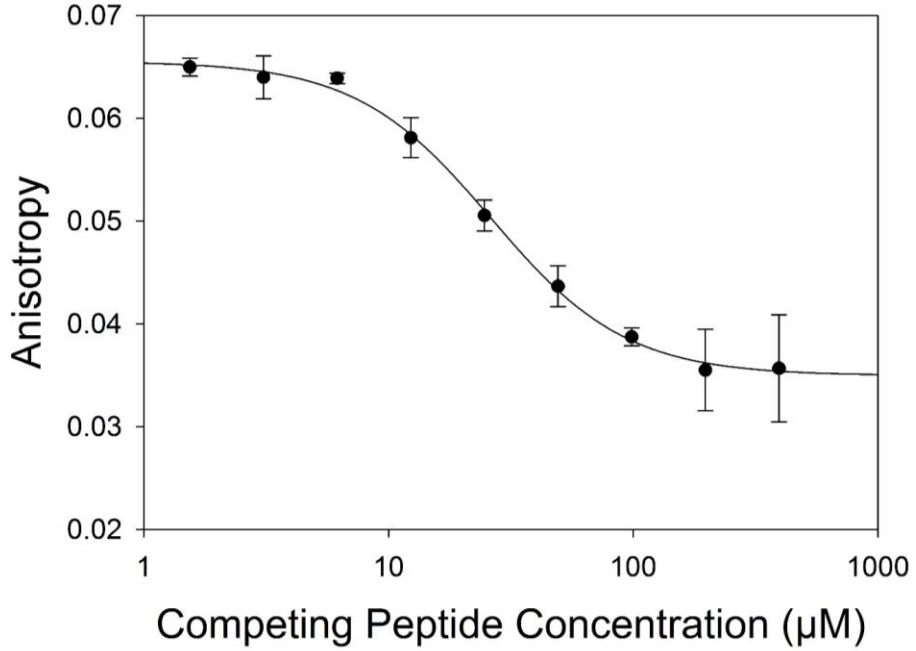
8.4 Linear

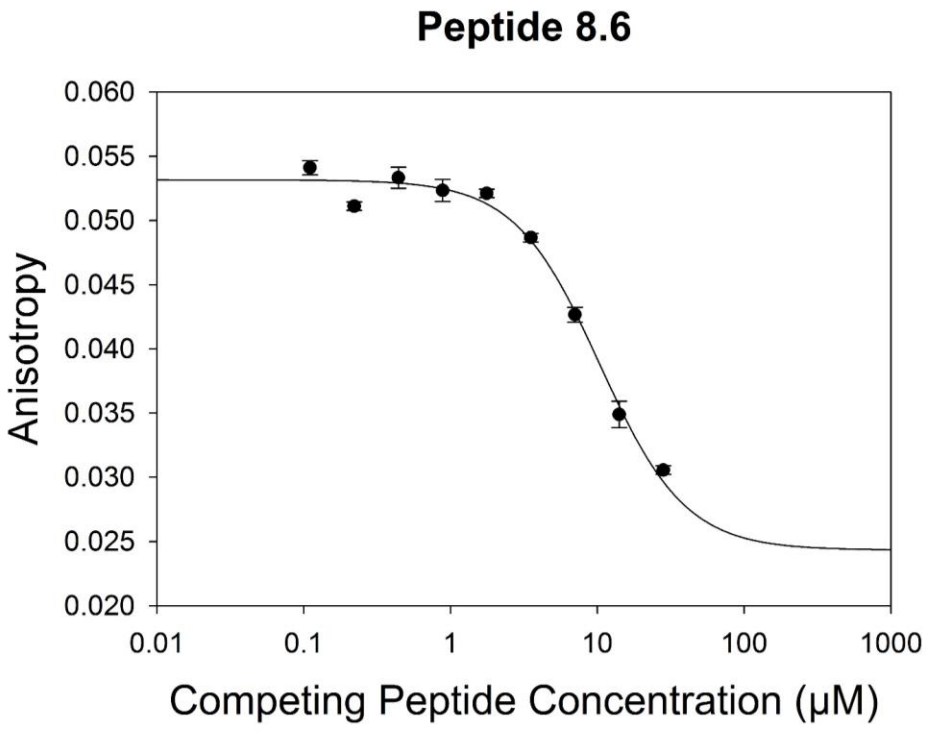


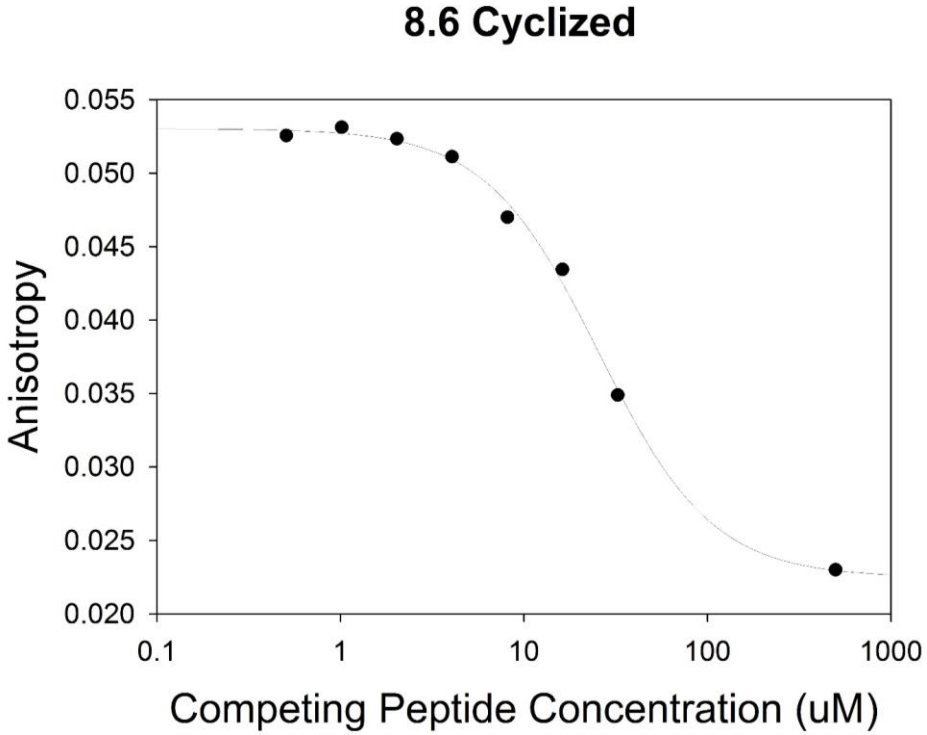
8.4 Cyclized

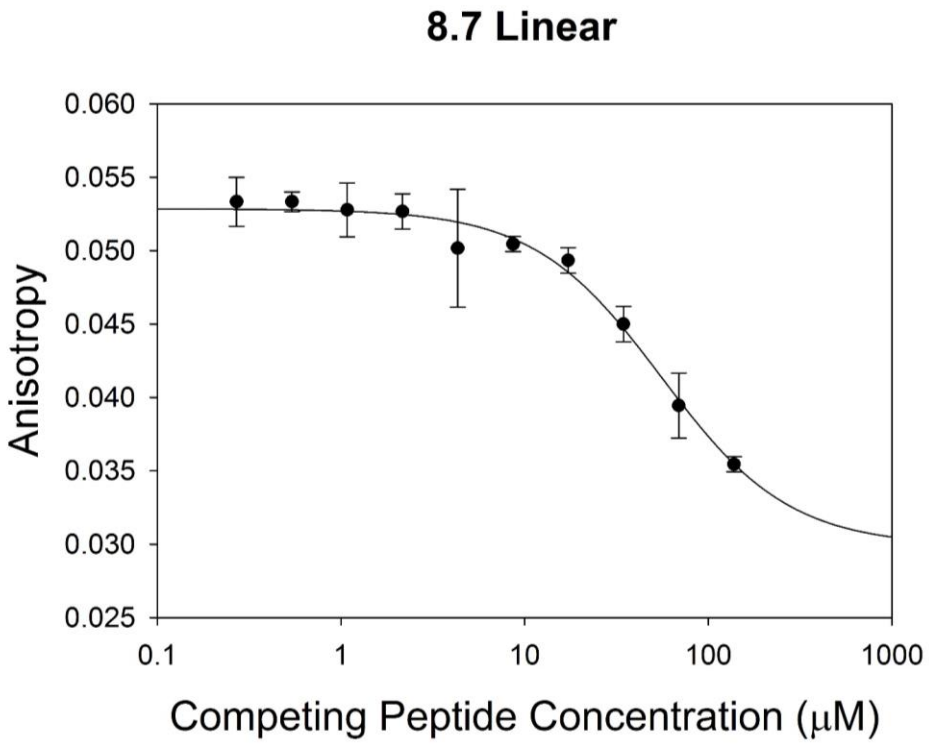


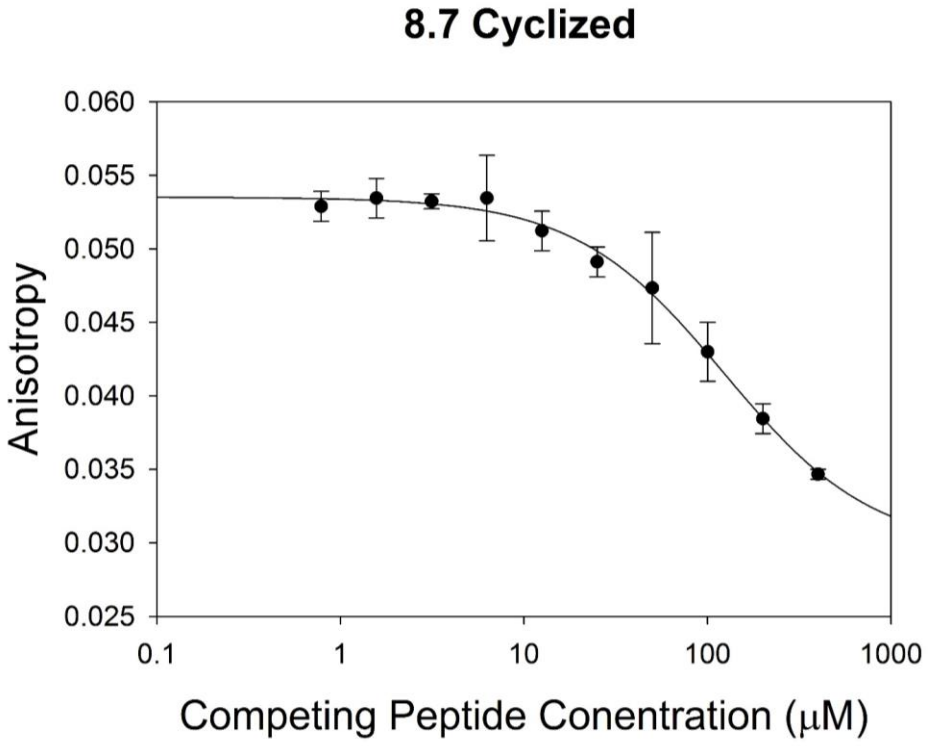
8.5 Linear

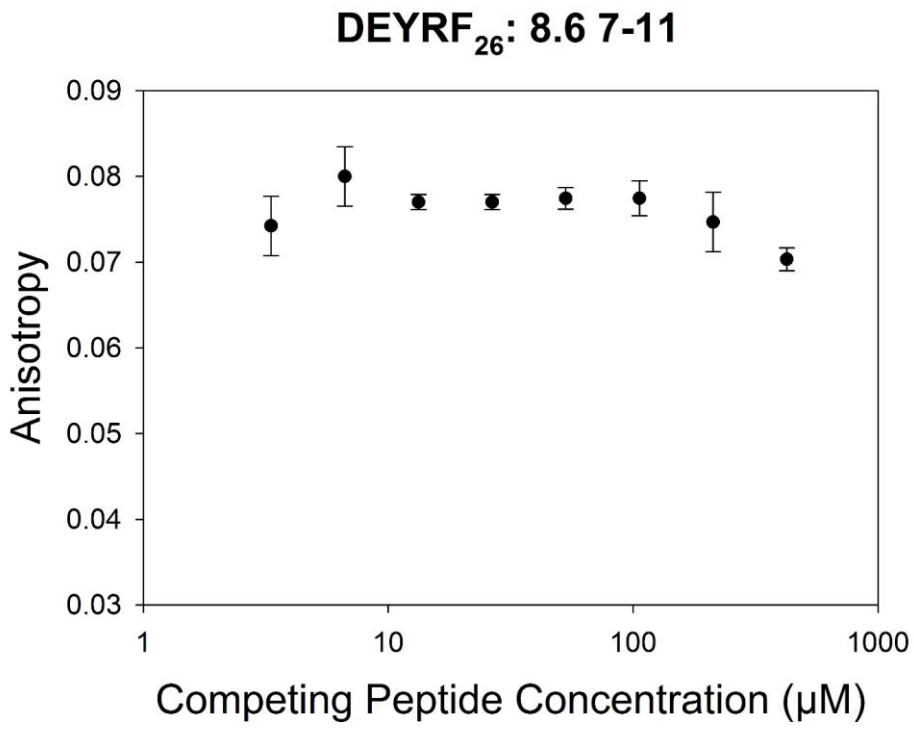


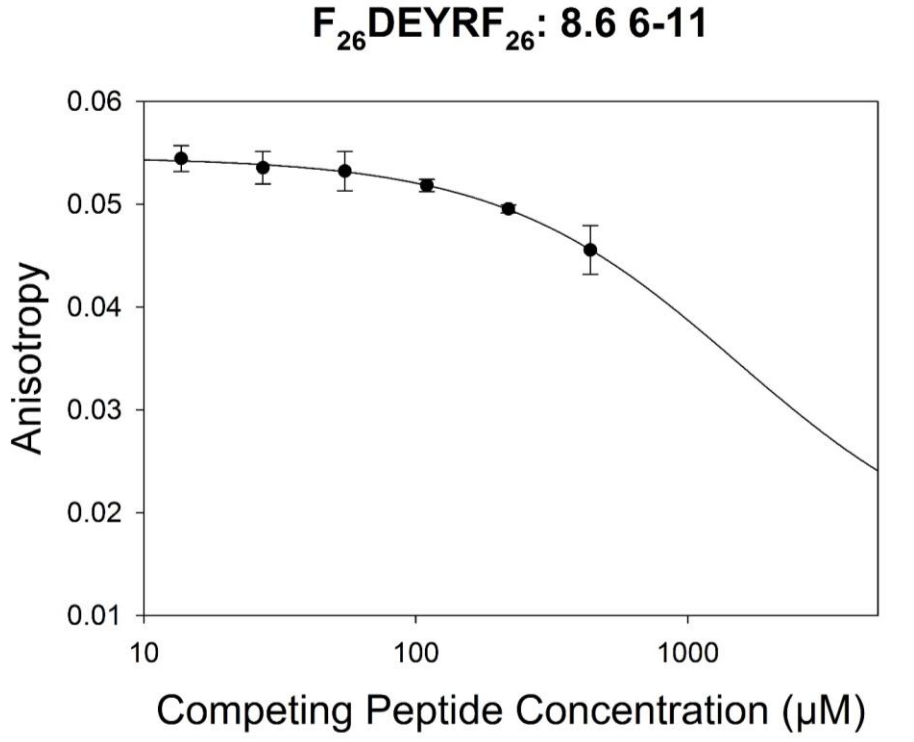

Peptide 8.6
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
Anisotropy
0.01
0.1
1
10
100
1000
Competing Peptide Concentration (μM)


8.6 Cyclized
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
Anisotropy
0.1
1
10
100
1000
Competing Peptide Concentration (uM)

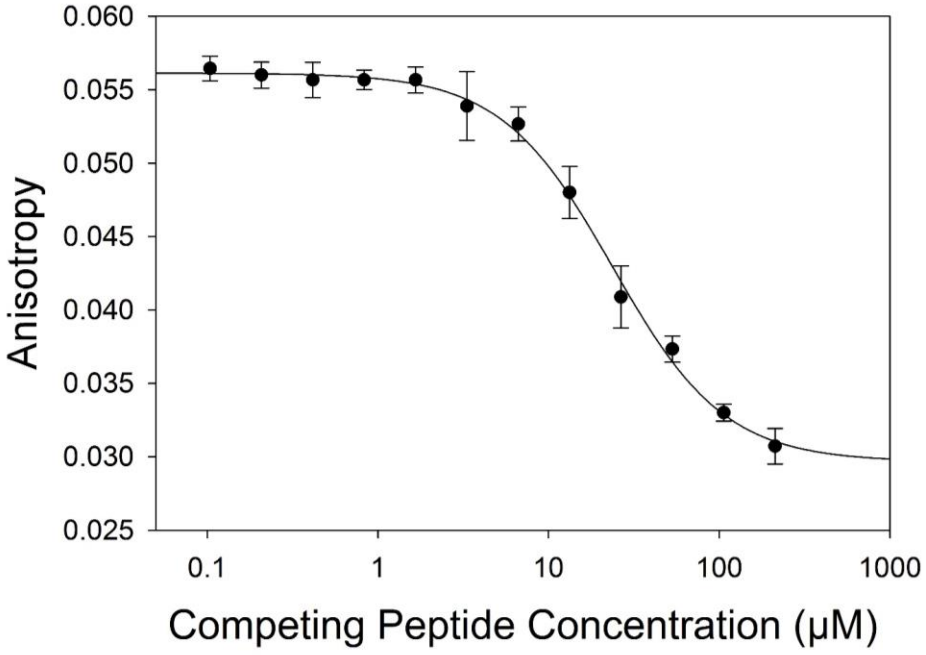

8.7 Linear
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
Anisotropy
0.1
1
10
100
1000
Competing Peptide Concentration (μM)


8.7 Cyclized
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
Anisotropy
0.1
1
10
100
1000
Competing Peptide Conentration (μM)

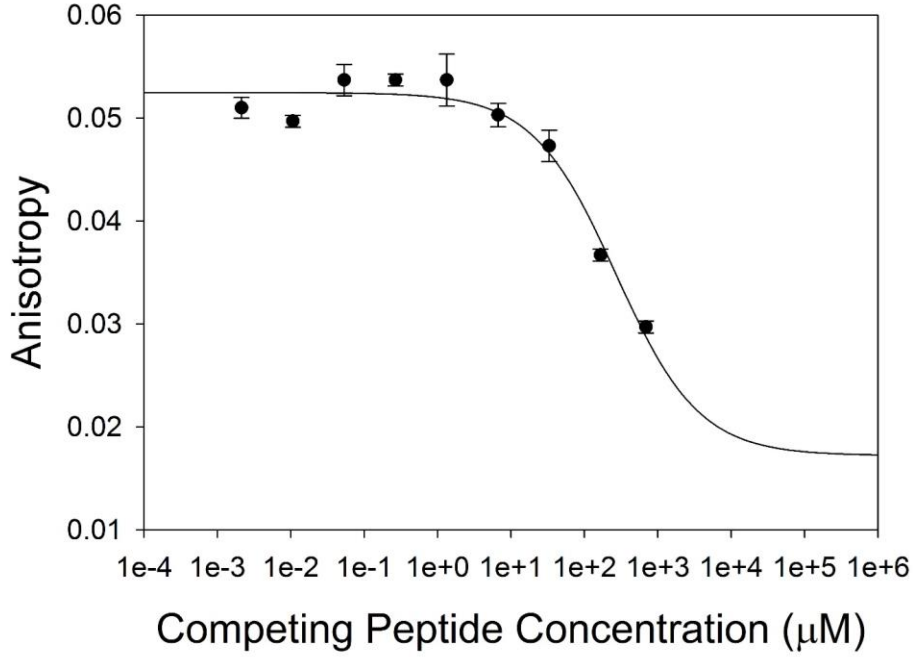

DEYRF26: 8.6 7-11
Anisotropy
0.09
0.08
0.07
0.06
0.05
0.04
0.03
1
10
100
1000
Competing Peptide Concentration (μM)


F26DEYRF26: 8.6 6-11
Anisotropy
0.06
0.05
0.04
0.03
0.02
0.01
10
100
1000
Competing Peptide Concentration (μM)

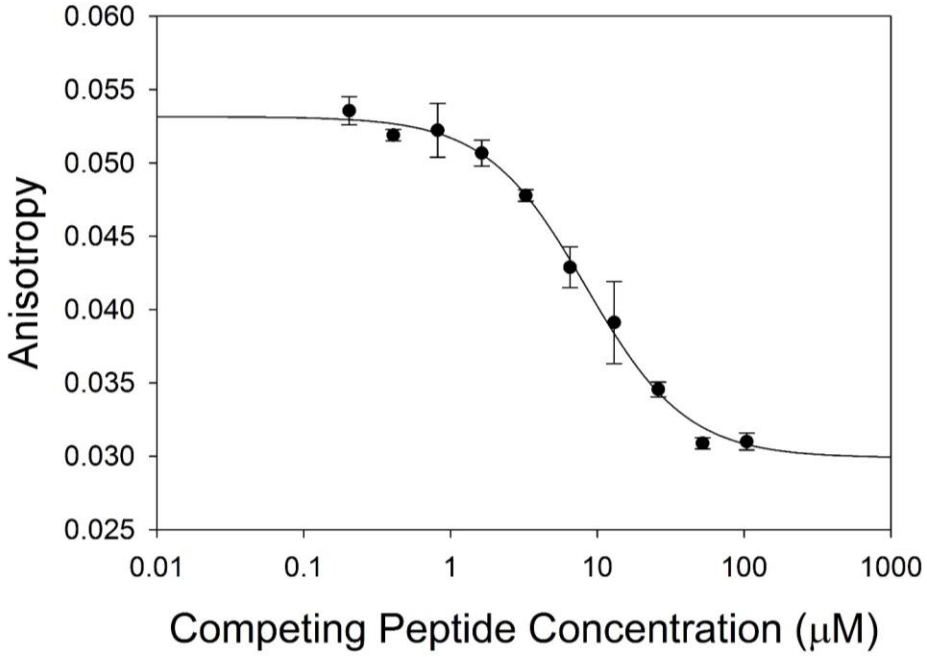
MCTIDF$_{26}$DEYRF$_{26}$: 8.6 1-11



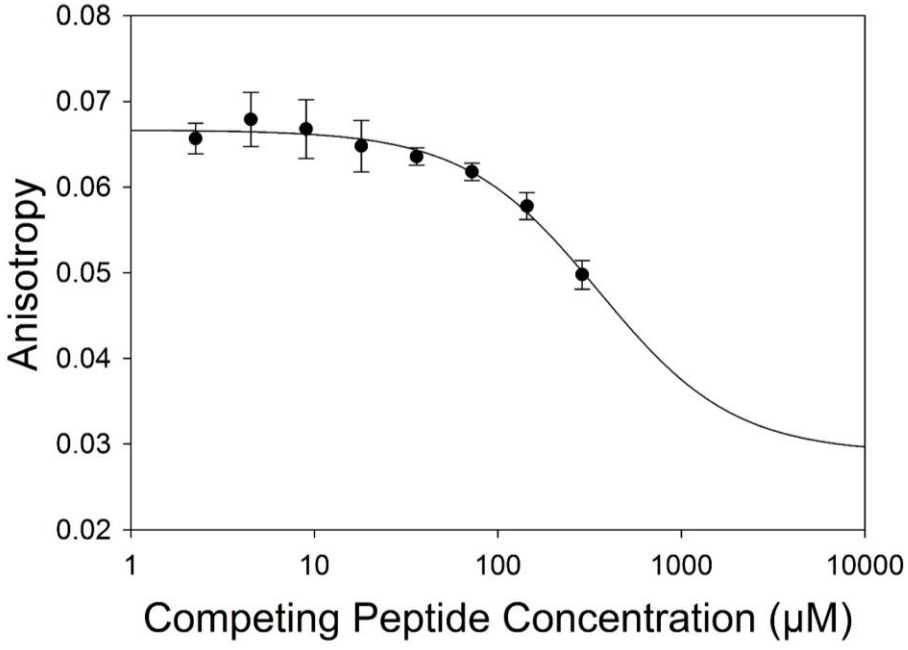
DEYRF$_{26}$RKT: 8.6 7-14

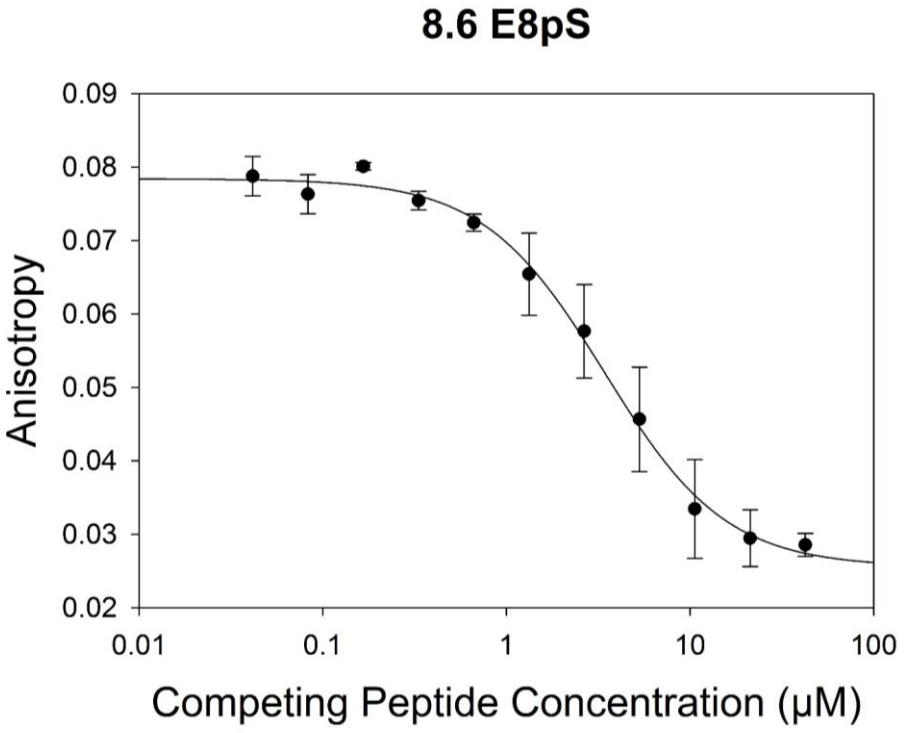


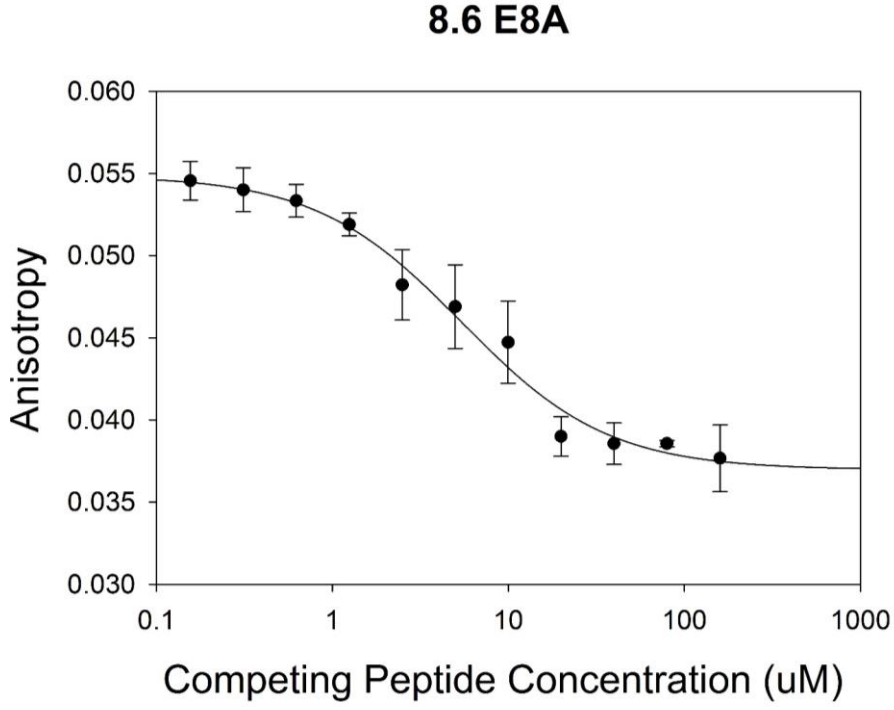
-MC: 8.6 3-14

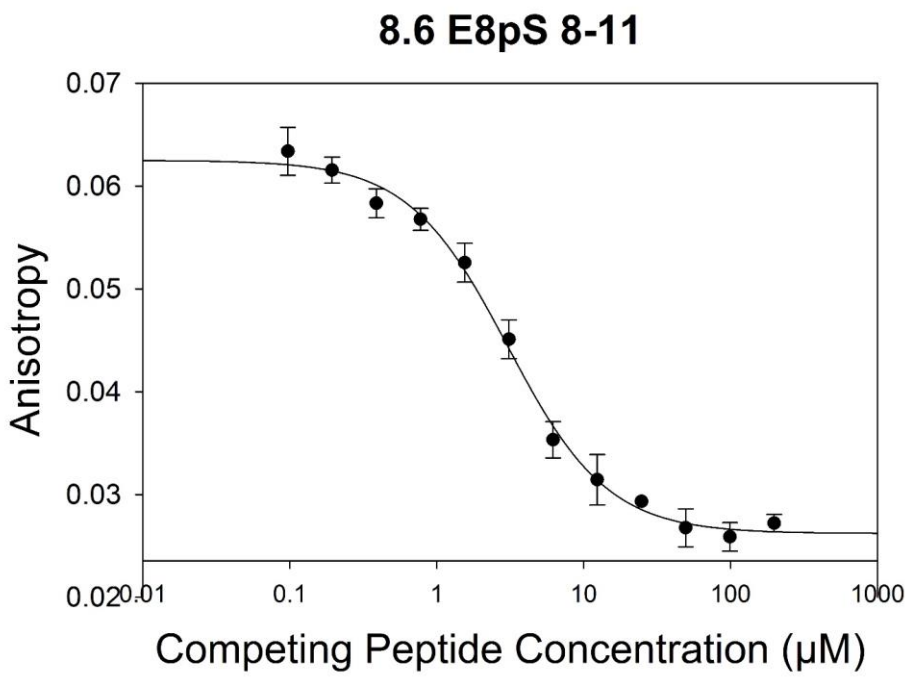


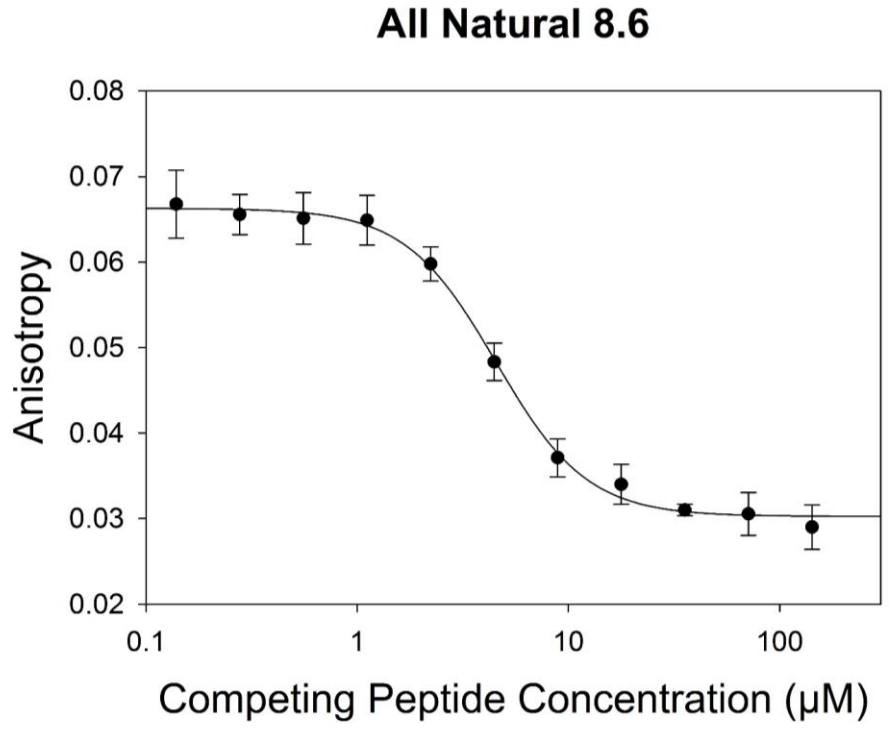
8.6 F11A

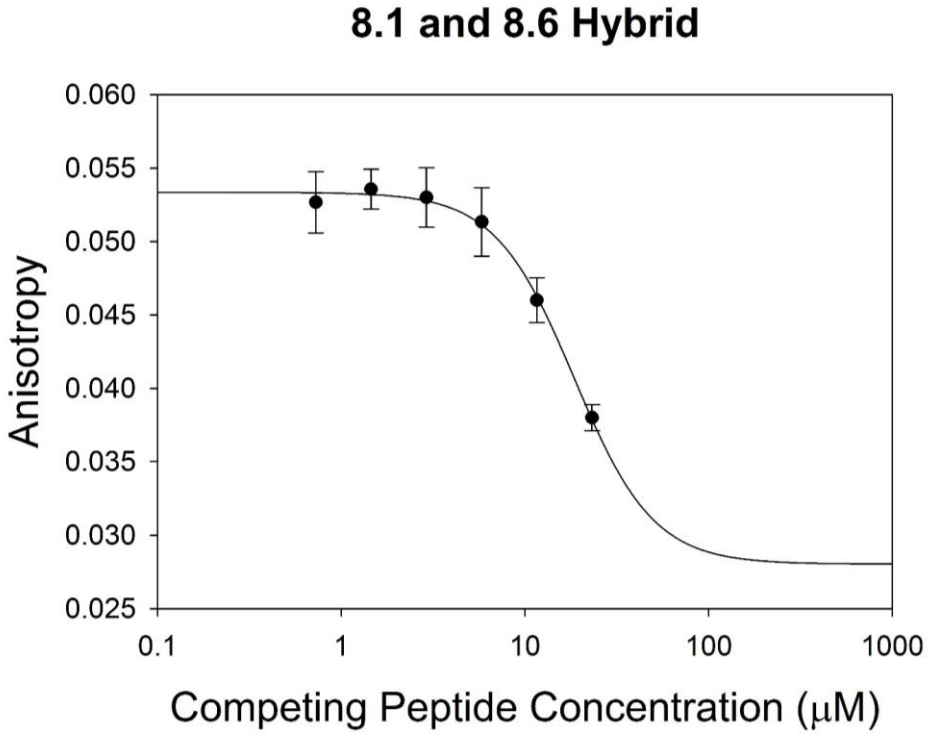


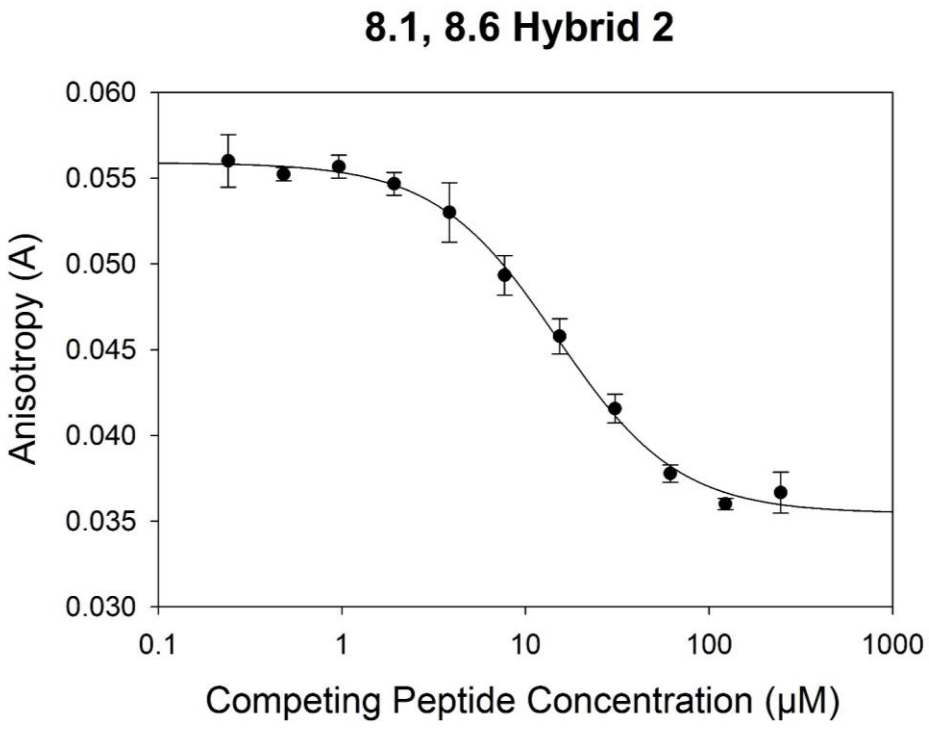

8.6 E8pS
Anisotropy
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.1
1
10
100
Competing Peptide Concentration (μM)

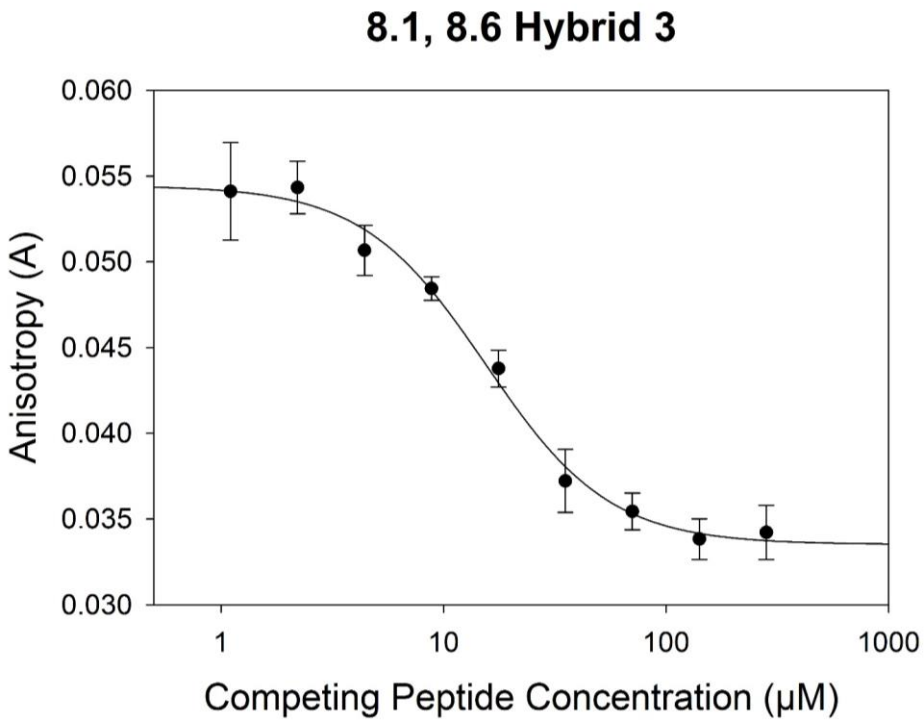

8.6 E8A
Anisotropy
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.1
1
10
100
1000
Competing Peptide Concentration (uM)

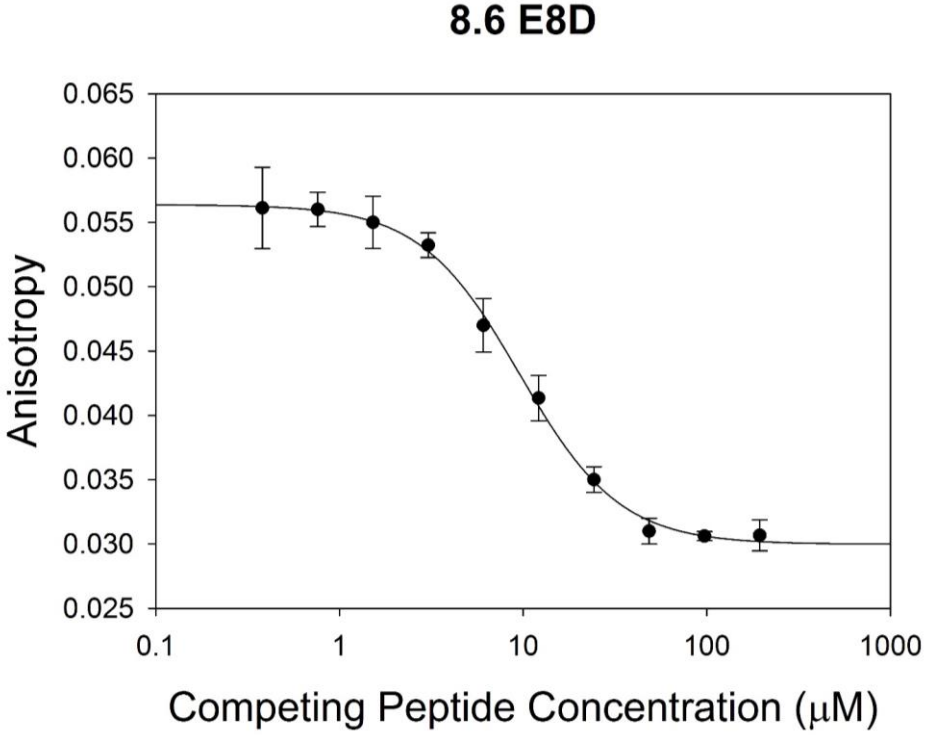

8.6 E8pS 8-11
Anisotropy
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.1
1
10
100
1000
Competing Peptide Concentration (μM)


All Natural 8.6
Anisotropy
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.1
1
10
100
Competing Peptide Concentration (μM)

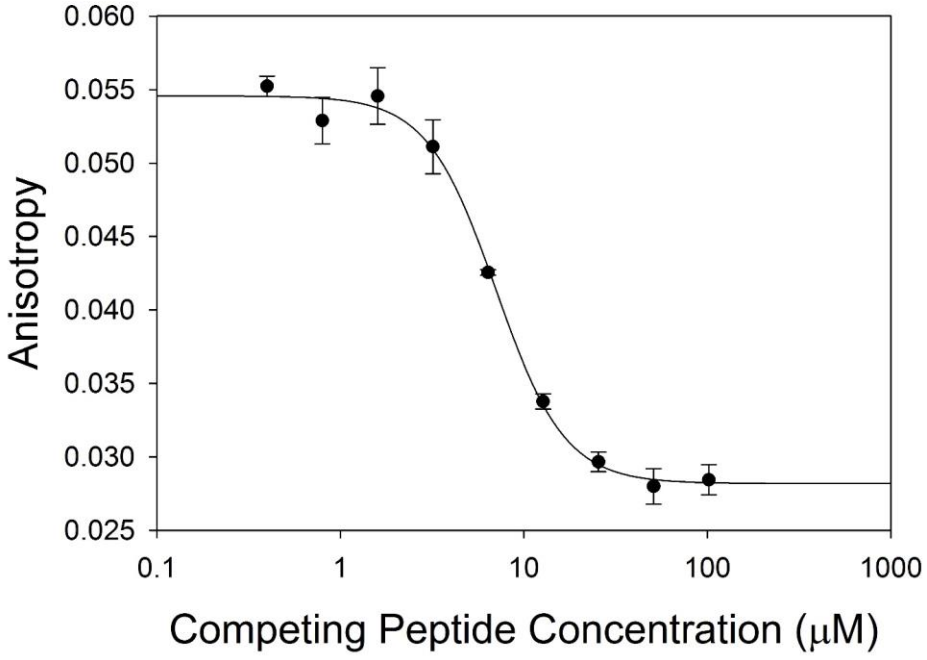

8.1 and 8.6 Hybrid
Anisotropy
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.1
1
10
100
1000
Competing Peptide Concentration (μM)


8.1, 8.6 Hybrid 2
Anisotropy (A)
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.1
1
10
100
1000
Competing Peptide Concentration (μM)

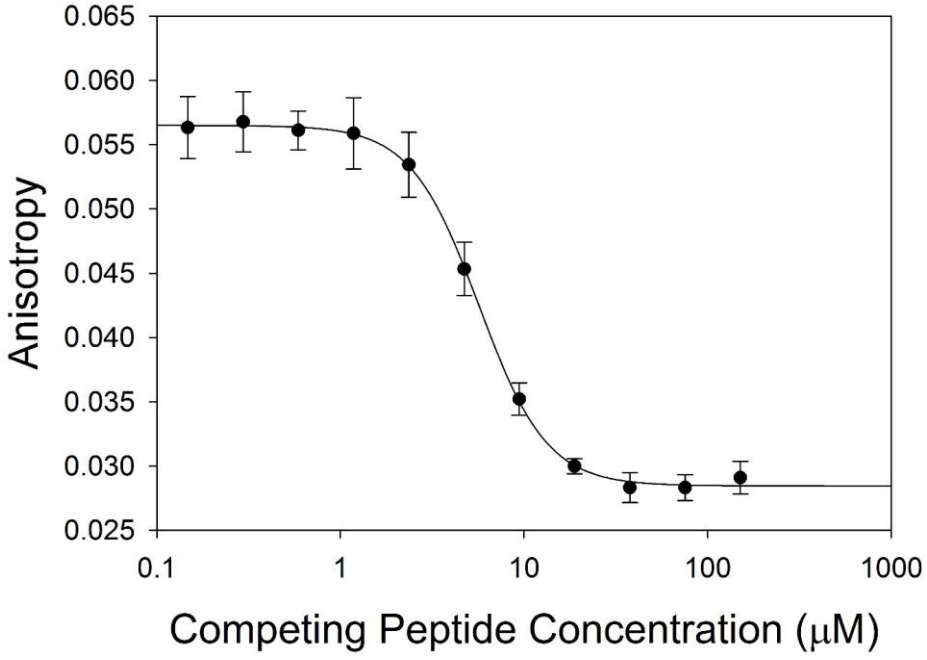

8.1, 8.6 Hybrid 3
0.060
0.055
0.050
0.045
0.040
0.035
0.030
Anisotropy (A)
1
10
100
1000
Competing Peptide Concentration (μM)


8.6 E8D
0.065
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
Anisotropy
0.1
1
10
100
1000
Competing Peptide Concentration (μM)

8.6 Y9F, R10I



8.6 D7N



Vita

Elizabeth Railey White was born on June 4, 1985 in Winston-Salem, North Carolina and is an American citizen. She graduated from Madison Central High School in Richmond, KY in 2003 with a Commonwealth Diploma. She received her Bachelor of Science in Chemistry from the University of Kentucky, Lexington, Kentucky in 2007. During her four years of undergraduate study she worked as a research assistant in the laboratory of Dr. David Atwood.

EDUCATION

M.D./Ph.D., anticipated May 2015
Virginia Commonwealth University School of Medicine, Richmond VA
(in the tradition of the Medical College of Virginia)

B.S. in Chemistry, summa cum laude, May 2007
The University of Kentucky, Lexington, KY

PUBLICATIONS

Dever, S. M., White, E. R., Hartman, M. C. T., and Valerie, K. (2012) BRCA1-directed, enhanced and aberrant homologous recombination: mechanism and potential treatment strategies, Cell Cycle 11, 687-694.

White, E. R., Reed, T. M., Ma, Z., and Hartman, M. C. T. (2012) Replacing amino acids in translation: expanding chemical diversity with non-natural variants, Methods Epub March 27.

CONFERENCE PRESENTATIONS

E. Railey White, Zhong Ma, David. E. Hacker, Melissa B. Huie, Jason M. Beckta, David C. Williams, Kristoffer Valerie, and Matthew C. T. Hartman (April, 2013). Selection of non-phophorylated peptide inhibitors of BRCA1. Oral presentation at the *Virginia Commonwealth University 3rd Annual Chemical Biology Symposium*, Richmond, VA.

E. Railey White, Zhong Ma, David. E. Hacker, Melissa B. Huie, Jason M. Beckta, David C. Williams, Kristoffer Valerie, and Matthew C. T. Hartman (April, 2013). Selection of non-phophorylated peptide inhibitors of BRCA1. Poster presentation at the *American Association of Cancer Research (AACR) Annual Meeting*, Washington D.C.

E. Railey White, Zhong Ma, David E. Hacker, and Matthew C.T. Hartman (July, 2012). Selection of Unnatural Peptide Inhibitors of the BRCT Domain of BRCA1. Poster presentation MD/PhD National Student Conference, Keystone, CO.

E. Railey White, Zhong Ma, David E. Hacker, and Matthew C.T. Hartman (October, 2011). Selection of Unnatural Peptide Inhibitors of the BRCT Domain of BRCA1. South East Regional Meeting of the American Chemical Society (SERMACS), Richmond, VA.

Zhong Ma, E. Railey White, David E. Hacker, Matthew C.T. Hartman (2010, August). Developing Unnatural Cyclic Peptide Inhibitors of BRCA1. Poster presentation at the *Chemical Insights into Biological Processes Symposium*, Frederick, MD.

RESEARCH EXPERIENCE

▪ Graduate Research Assistant August 2009-May 2013
Department of Chemistry, Chemical Biology Program, Virginia Commonwealth University, Advisor: Dr. Matthew Hartman

▪ Undergraduate Research Assistant September 2003-May 2007
Department of Chemistry, University of Kentucky, Advisor: Dr. David Atwood

TEACHING EXPERIENCE

▪ Physical Chemistry I Teaching Assistant Spring 2013
Department of Chemistry, Virginia Commonwealth University

▪ General Chemistry Teaching Assistant Spring 2006
Department of Chemistry, University of Kentucky

SERVICE

▪ MD/PhD Admissions Committee Fall 2009-Present
Student Interviewer, Fall 2009-Present
Admissions File Reviews Fall 2009-Spring 2010

▪ Missions of Mercy (MOM) Spring 2009
Medical Student Volunteer for screening and counseling at dental clinics that serve the more than 700 patients in a single weekend

▪ Fan Free Clinic Fall 2007
Medical Student Volunteer in evening clinic

HONORS

Phi Beta Kappa Society (inducted 2007)
Willard Riggs Meredith Award for Outstanding Undergraduate Chemistry Student (2007)