ABSTRACT

BACON, KAITLYN BROOKE. Engineering Combinatorial Platforms for Design and Analysis of Biomolecular Recognition. (Under the direction of Drs. Balaji M. Rao and Stefano Menegatti).

The use of non-immunoglobulin proteins and peptides as alternative biomolecular recognition agents has gained increasing prominence to address the limitations of antibodies. Using combinatorial platforms, proteins and peptides with novel binding activity can be identified through iterative rounds of selection against a target protein. Despite the success of identifying binding ligands, the isolation and characterization of binders using combinatorial platforms relies heavily on the use of soluble, recombinant protein, which is cumbersome to obtain in an appropriate form. When considering cyclic peptide binders, their isolation has relied significantly on display platforms that utilize panning for selection, which reduces the efficiency of selecting high affinity binders. Here, we describe novel combinatorial platforms to address these limitations.

Generally, binding ligands have been isolated from combinatorial libraries by performing selections against a soluble, recombinantly expressed form of the target protein. Using this strategy, the efficiency of identifying binding ligands is limited as each target protein must be recombinantly expressed and subsequently purified. Many mammalian protein targets cannot be appropriately expressed in *E.coli* as they require specific post-translational modifications that can only be afforded by eukaryotic hosts. As a substitute, we detail the selection of combinatorial libraries against magnetic yeast cells expressing the target protein as a surface fusion. Target displaying cells co-expressing an iron oxide binding protein, SsoFe2, are magnetized after incubation with iron oxide nanoparticles. As an alternative, yeast displaying the target protein can be also magnetized by the non-specific adsorption of iron oxide nanopowder. Here, we demonstrate the use of magnetic yeast targets when selecting binding ligands from yeast and

mRNA-display libraries. Specifically, we used yeast-displayed targets to identify protein and peptide binders with affinity for the mitochondrial membrane protein TOM22. We confirmed the functionality of the identified nanobody and Sso7d binders for recognizing natively expressed TOM22 by testing their ability to recover mitochondria from cell lysate.

Biophysical characterization (i.e. binding affinity estimation) of mutants identified from combinatorial screens also depends on the use of soluble protein. Here, we detail a quantitative yeast-yeast two hybrid (qYY2H) platform that can be used to quantitatively estimate the binding affinity (K_D) of mutants identified from combinatorial screens. The qYY2H system relies on the interaction of yeast cells expressing the binding protein (prey) with magnetic yeast cells expressing the target (bait). The prey cells also co-express luciferase, which serves as a reporter to quantify the number of prey cells that bind the bait cells. Using measurements of multivalent yeast-yeast binding, we established a semi-empirical model for estimating the K_D of monovalent bait-prey interactions. Additionally, we used qYY2H to characterize the interactions of SMAD3 and the WW domains of YAP with binding partners isolated from a cDNA library to show how qYY2H can also be used to assess protein-protein interactions identified from cDNA library screens. Moreover, we engineered the qYY2H platform to quantitatively assess binding interactions that are facilitated by post-translationally modified bait.

Lastly, we describe a novel yeast display platform for the generation of chemically constrained peptide libraries that can subjected to Fluorescence Activated Cell Sorting (FACS). Specifically, we engineered the yeast surface display platform to express linear peptides that can be cyclized between their N-terminus and a C-terminal lysine residue using a disuccinimidyl glutarate crosslinker. Using this platform, we screened a library of cyclic heptapeptides to identify binders to lysozyme and interleukin-17 (IL-17) using magnetic selection followed by FACS.

Notably, we demonstrated that yeast displayed cyclic peptides can be used instead of chemically synthesized peptide to quantitatively estimate binding affinity. This platform will improve the efficiency of identifying cyclic peptide binders, especially in comparison to other combinatorial platforms that rely on panning for selection.

Engineering Combinatorial Platforms for Design and Analysis of Biomolecular Recognition

by
Kaitlyn Brooke Bacon

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Chemical Engineering

Raleigh, North Carolina
2021

APPROVED BY:

_____	_____
Dr. Balaji M. Rao Committee Co-Chair	Dr. Stefano Menegatti Committee Co-Chair
_____	_____
Dr. Ruben Carbonell	Dr. Albert Keung

DEDICATION

To my parents,

Thank you for teaching me the value of hard work.

None of this would have been possible with your love, encouragement, and support.

BIOGRAPHY

Kaitlyn Brooke Bacon was born on in Little Rock, Arkansas in 1993 to Colyn and Tammy Bacon. When she was four, her family moved to Raleigh, North Carolina, where she grew up, and cemented her love for Duke basketball. Growing up, Kaitlyn always had an interest in how science knowledge can be used to better public health as many genetic diseases run in her family. Beyond science, Kaitlyn also had an interest in learning about other cultures and languages. She was fortunate to participate in a government sponsored exchange program in high school and had the opportunity to live with a family in Seoul, South Korea for three months. Kaitlyn attended the North Carolina School of Science and Mathematics for high school where her interest in research began. While in high school, she participated as a research volunteer in a genetics lab at Duke University under the direction of Dr. Allison Ashley-Koch. For undergrad, Kaitlyn attended Cornell University and studied Chemical Engineering. During this time, she worked for four years in the lab of Dr. Susan Daniel studying the kinetics of Influenza and Parvovirus binding to receptor analogs embedded in supported lipid bilayers. She also participated in a six-month co-op experience at Momenta Pharmaceuticals in Cambridge, Massachusetts executing experiments to aid in the develop of a process to produce an antibody therapeutic. This experience laid the foundation for her interest in designing tools to improve bioprocessing.

After enduring the brutal cold for four years, Kaitlyn decided to move back home (to the warmth!) to attend North Carolina State University for her Ph.D. in Chemical & Biomolecular Engineering in 2015. She worked within the research groups of Dr. Balaji Rao and Dr. Stefano Menegatti using protein engineering techniques to design tools for biomolecular recognition applications, in particular ligands for bioseparations. Kaitlyn participated in an internship experience during the 2019 summer at the Duke Human Vaccine Institute. Her project focused on

designing and optimizing a small scale purification system for a mRNA vaccine. After graduation, Kaitlyn will begin a role as a technical development scientist at Taysha Gene Therapies to pursue her interest in downstream bioprocessing.

Kaitlyn's greatest passion outside of lab is to promote STEM education to young girls. Her interest in this area began in undergrad where she was an outreach volunteer for her school's Society of Women Engineers. Kaitlyn built upon this passion in graduate school by volunteering with the Young Women in Bio organization. Kaitlyn has served in multiple leadership roles within this organization and was the committee chair (2019 – present). This group designs free programs geared towards introducing girls in grades K-12 to possible areas of study and careers in all areas of STEM. Examples of events include tours of local biotechnology companies, activities with research labs at local universities, and science-based webinars. Kaitlyn also served as a Science Olympiad coach at Wiley elementary for two years during graduate school.

ACKNOWLEDGMENTS

To begin, I would like to thank my advisors Dr. Rao and Dr. Menegatti for their guidance, support, and encouragement that allowed me to grow as researcher. I am incredibly appreciative that Dr. Rao has allowed me to design my own research project to pursue my interest in bioprocessing. Thank you for allowing me to take time away from lab to grow professionally at my internship and conferences. I would like to acknowledge the members of my committee (Dr. Keung, Dr. Carbonell, and Dr. Freytes) for their invaluable insights. I would also like to thank Dr. Kelly for always having an open ear to listen and to give advice.

I am indebted to my undergraduates who helped me with many key experiments. This thesis wouldn't have been possible without their dedication. I can't wait to see what great researchers each of them will become in the future! Matt- Thank you for your patience when you joined the lab as we were both new researchers. Abby- Thank you for keeping our projects going while I was away at my internship. I am appreciative for your attention to detail when it comes to lab work. Nikki- I enjoyed watching you grow from an undergraduate to an independent researcher during your gap year. Your infectious personality has been missed in lab.

I am incredibly grateful for Dr. Carlos Cruz for his mentorship and continued friendship. Thank you for your patience when you were first teaching me about protein engineering. You showed me how hard work can pay off in research! I would also like to thank Dr. John Bowen for all your support in lab and always having my back. Our conversations, scientific and otherwise, got me through. Lab hasn't been the same without you. I would also like to thank my other lab members Dr. Adam Mischler and Jenna Meanor as well as Dr. Ashton Lavoie and Dr. John Schneible who helped me many times with equipment in the Menegatti Lab.

To Yunhu, thank you for being my best friend. I am going to miss our dinner dates, bubble tea outings, and Switch nights. One day we will beat Overcooked! Even though I will miss you once you move, I can't wait to see what great things are in store for you.

To James, thank you for being my rock, especially this past year, and for always believing in me, even when I don't. I am very appreciative how you always try to help me see the bigger picture. This has helped keep me sane, more than you know, throughout this crazy journey. I am forever grateful for our science conversations and how you challenge me to be a better researcher.

Most importantly, I would like to thank my parents whose love and support has fueled all of my accomplishments. Without your sacrifice of letting me leave home two years early to go to NCSSM and then Cornell, I may not be where I am today. I am incredibly grateful I chose to move closer to home for graduate school to spend more time with both of you. Thank you for all the meals, adventures, and laughter that helped me get through graduate school. Dad, thank you for always saying yes to my crazy ideas and always providing the means for me to explore my interests. I am incredibly grateful for all of our hiking adventures over the last few years. Thank you for teaching the value of hard work and how it can get you to where you want to be in life. Mommy, thank you for always being my number one cheerleader. I know how much you wanted to be here to see me defend my PhD, but I know you and grandpa will be watching me from above. I miss calling you when I leave lab every day to catch you up on my day. Your continued fight to live, especially in that last year, has taught me the value of perseverance. I am forever appreciative of the advances in science that allowed me an extra 22 years with you after your lung transplant. Hopefully, we aren't too far off from there being a gene therapy treatment for the disease that has taken too many of our family members.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Engineering Biomolecular Recognition Agents

1.1 Introduction

Biological systems rely significantly on interactions between biomolecules for a multitude of processes, including, but not limited to, cell structure, signal transduction, immune recognition, cell differentiation, and cell adhesion[1]. These interactions are based on biomolecular recognition, the process by which a biomolecule specifically recognizes and non-covalently binds to a molecular target[2]. Biomolecular recognition events include protein-protein[3], protein-nucleic acid[4–7], sugar-lectin[8], and RNA-ribosome[9] interactions. These events are driven by a combination of van der Waal forces, hydrogen bonds, hydrophobic interactions, and electrostatic interactions[6,7,10]. Examples include antibodies binding to pathogenic antigens to an illicit an immune response[11] or transcription factor proteins binding to DNA to modulate transcription[12].

1.2 Biomolecular recognition agents

Due to the high specificity of biomolecular recognition events, this phenomenon has been exploited to develop diagnostic and quantitative measurement tools[13–16] as well as for the capture of specific target molecules from complex mixtures[17–19]. In these applications, proteins typically serve as the molecular recognition reagent, or the affinity ligand. Antibodies are the most commonly used affinity reagent as binding molecules can be generated easily for almost any antigen of interest. Recently, several non-immunoglobulin protein scaffolds have emerged as alternative molecules that can be engineered to achieve biomolecular recognition[20–23]. Small peptides are also being explored as affinity ligands[24,25]. The success of non-antibody-based ligands has hinged on innovation in combinatorial library screening. Amino acid residues can be varied within protein and peptide scaffolds to generate large mutant libraries using techniques like phage display[26], yeast surface display[27], and mRNA display[28]. The libraries can then be screened against

a target protein of interest to identify particular variants of the starting scaffold that have affinity for the target protein.

1.2.1 Antibodies

Antibodies are the most commonly used biomolecular recognition agents for basic research, industrial, and clinical applications. Specifically, antibodies are large, Y shaped, multi-domain proteins (~150 kDa). While the structure of all antibodies is similar, a region at the tips of the Y differs between antibody variants, giving each antibody specificity for a different antigen[29]. This region comprises six hypervariable loops, called the complementarity determining regions (CDRs). Variations in these loops allow an antibody to bind a specific antigen.

Historically, antibodies against specific antigen targets have been generated by immunizing an animal, like a mouse or a rabbit, with the antigen of interest. Accordingly, the animal's immune system generates antibodies against the antigen which can then be harvested from serum[30]. While the generated antibodies are specific for the target, they typically are heterogenous in nature with each binding a different epitope; Thus, there is significant lot to lot variability, especially if the immunized animals exhibit a different immune response over time[31,32]. Antibodies produced in this manner are known as polyclonal antibodies. As an alternative, researchers have used hybridoma technology to generate mammalian cell lines that produce monoclonal antibodies (mAbs) that bind a single antigen epitope[33]. Specifically, researchers isolate the spleen of an animal after it is immunized with a desired antigen. The spleen is fused with a myeloma cell line, generating an immortalized B cell-myeloma hybridoma that can be grown continuously while co-producing antibodies. While hybridoma technology reduces variability, it is very expensive and time consuming to produce these cell lines[32].

Today, the governing production system for mAbs is recombinant expression. Antibodies are generally produced in mammalian systems as they require specific post-translational modifications that cannot be afforded by more simple prokaryotic hosts, like *E.coli*[34]. Generally, CHO and NS0 cell lines are used for recombinant mAb expression[35]. Genetic engineering tools are making it possible, however, to use other recombinant systems for expression. For example, glyco-engineered *Pichia pastoris* cell lines have been explored to produce mAbs with highly specific glycoforms by the company Glycofi[36]. Researchers are also engineering *E.coli* for antibody assemble. A major limitation of using bacterial cell lines for antibody production is that the disulfide bond that tethers an antibody's heavy and light chains is reduced in the cytoplasm where expression typically occurs in bacterial cells[34,37]. Using genetic engineering, protein expression can be altered to the periplasm to assemble full IgGs. However, the produced antibodies are aglycosylated[38].

1.2.2 Antibody fragments

While antibodies have been the most commonly used biomolecular recognition moiety, they suffer from other drawbacks beyond production limitations. In particular, the large size of antibodies results in low tumor penetration, excessive systemic accumulation, and slow clearance rates[39]. Additionally, the Fc portion on mAbs can cross-react with Fc-receptors on immune cells eliciting an off-target response[40]. As a result, antibody fragments are being explored to address these limitations. These moieties maintain the targeting capacity of antibodies by incorporating the variable loop structures responsible for antigen recognition[39]. Single chain antibody fragments (scFv) are a particular type of antibody fragment. Specifically, they are a fusion protein containing a flexible linker that connects the variable regions of an antibody's heavy (V_H) and light (V_L) chains[41]. Other examples include single domain antibodies that contain only the heavy chain.

Single domain antibodies are found in camelids (V_HH fragments) as well as sharks (V_{NAR} fragments)[42]. scFvs and single domain antibodies can be expressed using mammalian, yeast, and insect cells that secrete the antibody fragment into the culturing media, eliminating the need for cell lysis[41,43]. More often, scFvs and single domain antibodies are be produced in *E.coli*. However, most antibody fragments produced in *E.coli* must be expressed in the periplasm to maintain disulfide bridges; though, work is ongoing to improve cytoplasmic production[41,44–46].

When considering antibody fragments, camelid V_HH fragments are becoming more widely used than scFvs. Specifically, camelids do not suffer from solubility and aggregation issues typically associated with scFvs[47]. While the V_H and V_L chains of a scFv are connected via a linker, the proper orientation of these domains is dependent on hydrophobic interactions between each domain. The hydrophobic faces of each domain can dissociate and interact with other hydrophobic surfaces leading to aggregation[48]. Camelid V_HHs also differ from traditional antibodies when comparing the size of their CDR3 loop. The diversity in this loop contributes significantly to antibody specificity. The CDR3 loop in a camelid V_HH is generally comprised of more amino acid residues than human V_H domains, thereby increasing the potential interaction area with a target antigen[42,47,49]. Also, the CDR3 of a camelid V_HH exhibits a unique fingerlike conformation that can extend into antigen cavities while the CDR3 associated with human antibodies usually binds via a relatively flat interaction face[50]. Accordingly, these particular properties allow nanobodies access to binding epitopes that may not be accessible to traditional antibodies and scFvs[42,47,51,52].

1.2.3 Binding proteins derived from non-immunoglobulin scaffolds

Non-immunoglobulin scaffolds have emerged as another promising affinity ligand class to address the limitations of antibodies. Ligands developed from these alternative scaffolds are based on a naturally occurring protein template where specific amino acid residues are mutated to

produce entirely novel binding activity. Affinity ligands based on non-immunoglobulin scaffolds can be obtained by rational design, or more commonly by combinatorial protein engineering[53]. Utilizing these methods, binding proteins evolved from non-immunoglobulin scaffolds have been identified for an assortment of applications, including therapeutics[54], diagnostics[55], imaging[56], and bioseparations[57]. Numerous naturally occurring proteins have served as starting templates to engineer proteins with novel binding activity. Examples include affibodies[58], DARPins[59], adnectins[60], avimers[61], anticalins[62], fynomers[63], Kunitz domains[53,64], and knottins[65].

Several issues surround whether a protein template will make a successful scaffold. First, an ideal protein template must maintain its structural integrity after mutation. In comparison to antibodies, scaffold proteins should exhibit superior biophysical properties. Specifically, an ideal scaffold will be small (<20 kDa) and thermodynamically stable. It is also desirable for the scaffold to be comprised of a single domain that lacks disulfide bonds and glycosylation[66]. Consequently, any evolved mutants can be expressed microbially, which affords less expensive and more efficient expression in comparison to antibody production[21]. Most importantly, mutated scaffolds should result in ligands with affinities and specificities similar to antibodies.

Proteins derived from hyperthermophilic organisms are excellent candidates for alternative scaffolds as they exhibit many of the characteristics desired for an alternative scaffold. In particular, proteins derived from hyperthermophilic organisms exhibit high thermal stability (melting temperatures near 100°C) and proteolysis resistance, which result from these organisms' evolution to survive in environments prone to harsh temperatures and pH[67–70]. This stability allows hyperthermophilic proteins to maintain their structural integrity even after significant mutagenesis (~15% of residues)[20]. A particular hyperthermophilic protein that has been engineered extensively to generate novel affinity reagents is the Sso7d protein from *Sulfolobus solfataricus*. Gera et al.

mutagenized 10 surface exposed residues within Sso7d to generate a combinatorial protein library that was screened to identify mutants with antibody like affinity for targets including lysozyme, streptavidin, mIgG, and cIgY[20]. Sso7d has gained popularity as an alternative scaffold due to its small size (~7 kDa) and its efficient recombinant expression in *E.coli*[71]. Moreover, affinity ligands derived from Sso7d exhibit significant stability when exposed to extreme temperatures and pHs as well as chemical denaturation[20]. These combined characteristics promote the use of the Sso7d scaffold as an alternative to the traditional antibody framework when developing biomolecular recognition moieties.

1.2.4 Cyclic peptides

A demand for the further miniaturization of biomolecular recognition reagents has promoted the development of cyclic peptides as affinity ligands[72–75]. Affinity peptides generally contain between 5-30 amino acid residues, putting their size between non-immunoglobulin scaffolds and small molecules. When considering therapeutic, diagnostic, and imaging applications[76–79], their smaller size allows for improved tumor penetration[80,81] and cell membrane permeability[82]. However, it is important to note, that despite their small size, many peptide sequences cannot permeate tumors or cell membranes; permeability is dictated by multiple factors beyond peptide size[83]. Unlike antibody and non-IgG protein scaffold production, cyclic peptide synthesis does not require a biological host. Rather, cyclic peptides can be produced using chemical synthesis, which significantly reduces production costs[84].

Specific residues within a cyclic peptide can be randomized to evolve a peptide moiety with novel binding activity for a target of interest. Monocyclic peptides are most commonly used when generating affinity reagents due to the ease of engineering a single complementary loop[76]. More recently, multicyclic peptides, including bicycles[85] and tricycles[86–88], have been explored as

well. Cyclic peptides are more attractive than their linear counterparts as cyclic peptides typically exhibit higher affinity and specificity for their target due to increased conformational rigidity resulting from cyclization[89]. This rigidity decreases the entropic contribution to Gibbs Free Energy resulting in enhanced binding. Cyclic peptides also have enhanced stability over linear peptides as their decreased flexibility prevents them from taking on a confirmation that can fit in the active site of proteolytic endopeptidases[90]. Additionally, if head to tail cyclization is utilized, proteolytic exopeptidases cannot access the N or C termini of these peptides[91,92].

1.3 Engineering proteins and peptides for biomolecular recognition

The adoption of non-immunoglobulin scaffolds and cyclic peptides as alternative biomolecular reagents has depended on advances in protein engineering. Historically, two approaches have been used to engineer proteins and peptides with specific function, (1) rational design and (2) directed evolution. When using a rational design approach, protein structure and molecular modeling data are utilized to propose mutations that may lead to a specific phenotype[93]. After, individual point mutations are introduced to the protein or peptide template by site-specific mutagenesis followed by biophysical characterization of the new mutant.

Despite advances in computational modeling and an improved understanding of protein chemistry, it remains challenging and often is inefficient to engineer proteins using a rational design approach[94,95]. Instead, most researchers rely on directed evolution to engineer proteins to have a particular phenotype (Figure 1.1)[96]. Specifically, random mutagenesis[97] or DNA shuffling[98] techniques are used to generate combinatorial libraries of up to 10^{15} template variants[99]. After, the libraries are subjected to iterative rounds of selection and amplification of the isolated mutants. Additional mutagenesis can take place between selection rounds. This ultimately results in a homogenous population that exhibits the desired phenotype which can be evaluated to understand

which amino acid mutations gave rise to the desired function. This process is similar to Darwinian evolution as a selection pressure is applied over time to isolate the mutants from the library that exhibit the desired phenotype from the mutants lacking the desired phenotype[100]. When screening libraries to identify novel binding properties, an appropriate selection pressure would be a decrease in the concentration of target antigen over multiple selection rounds. Further, the stringency of the washing conditions could be manipulated to reduce non-specific binding of undesired mutants.

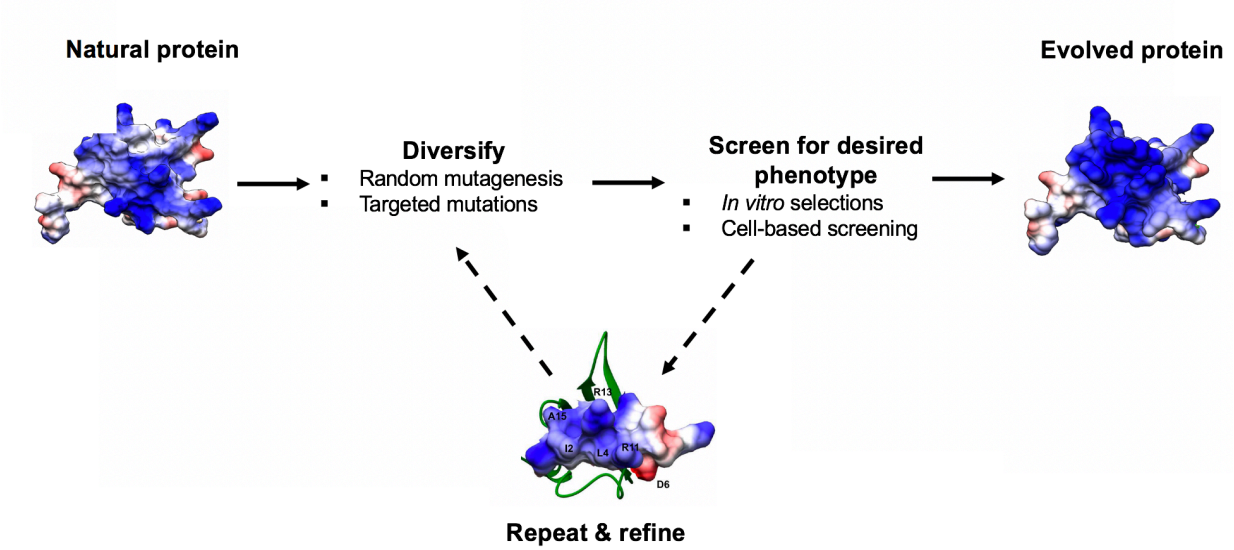


Figure 1.1. Proteins with novel activity can be engineered using directed evolution. A naturally occurring protein can be mutagenized to generate a library containing variants of the starting protein template. After, a selection pressure will be applied to the library. Mutants that exhibit the desired phenotype are isolated. Multiple rounds of diversification and screening are performed, ultimately resulting in a homogenous population of mutants that exhibit the desired phenotype.

The success of combinatorial library screening has hinged on the ability to genetically introduce library diversity. Genetic diversity is most commonly introduced by random mutagenesis where an amino acid within a protein sequence is randomly replaced by any of the other 19 amino acids. The template DNA can be amplified using a "error-prone" polymerase chain reaction (PCR)[97]. In these reactions, DNA polymerases are used that introduce errors at specific rates if metals or nucleotide analogs are added to the reaction[97,101,102]. If some structure-function

knowledge is known, a semi-rational approach can be used to generate a targeted library where selected residues within a protein template are mutagenized[103]. This approach relies on the use of degenerate codons during traditional oligonucleotide synthesis to generate targeted library diversity[104]. DNA shuffling is used less frequently to generate template diversity[98]. In this method, related genes are fragmented through the use of DNAse or sonication. Fragments with significant homology will anneal and can be extended by PCR allowing for related pieces of DNA to be randomly put together[105]. Researchers will often use a variety of these techniques when generating combinatorial libraries to evolve a protein or a peptide template[106].

1.3.1 Display technologies for combinatorial library expression

The critical challenge of combinatorial library screening is the identification of which mutants exhibit the desired phenotype. It is inefficient to recombinantly express, purify, and test each individual mutant in a combinatorial library. Instead, display technologies have been developed where each mutant within the library is barcoded. Biological organisms, like phage particles[107–109], bacteria[110,111], yeast[27,112], and mammalian cells[113–115], have been engineered to express individual mutants from a combinatorial library as fusions to surface proteins. This eliminates the need to express and purify each individual mutant for characterization. Each display organism carries DNA encoding the expressed mutant protein or peptide, linking phenotype to genotype, thereby barcoding each mutant[116]. For example, a yeast cell population can be generated where each individual cell expresses a different mutant protein from a combinatorial library. The library population can be incubated with an immobilized target protein. Any yeast cells that bind the target protein can be isolated. The DNA of the isolated cells can be extracted and sequenced to identify the mutants that have specificity for the target. Cell-free systems, like ribosome

display[117–119] and mRNA display[120–122], have also been developed to generate combinatorial libraries that take advantage of *in vitro* translation to couple genotype to phenotype.

1.3.1.1 Yeast surface display

Since its introduction over twenty years ago[27], the yeast surface display platform has been used to engineer proteins with novel binding, stability, and catalytic activity. A greater diversity of proteins can be properly displayed using yeast, in contrast to phage and bacterial display, as yeast are eukaryotic[123]. In particular, yeast possess machinery for glycosylation as well as the ability to express proteins within the oxidizing environment of the endoplasmic reticulum in the presence of molecular chaperones, thereby enhancing the expression of disulfide-bonded and natively glycosylated proteins, like antibody fragments[124–126]. These unique features have propelled the wide adoption of yeast surface display by the protein engineering community.

A variety of yeast strains and cell surface proteins have been used to display proteins of interest on the yeast cell surface[127–129]. Nevertheless, the most frequently used system depends on the *Saccharomyces cerevisiae* strain and its cell surface protein Aga2P (Figure 1.2)[27,112,130]. Specifically, proteins are expressed as fusions to Aga2P to afford display on the surface of *Saccharomyces cerevisiae*. Aag2P is a subunit of the yeast cell mating protein a-agglutinin. Aga2P is disulfide bonded to the cell wall associated-Aga1P subunit of α-agglutinin, thereby tethering any proteins fused to Aga2P to the yeast cell wall. To afford display, the protein-Aga2P fusion is genetically encoded into a single-copy plasmid vector that is transformed into a yeast strain, EBY100, that has the Aga1P gene stably integrated into the chromosome. Expression of Aga1P and Aga2P, and accordingly the displayed protein, is under the control of a galactose-inducible promoter.

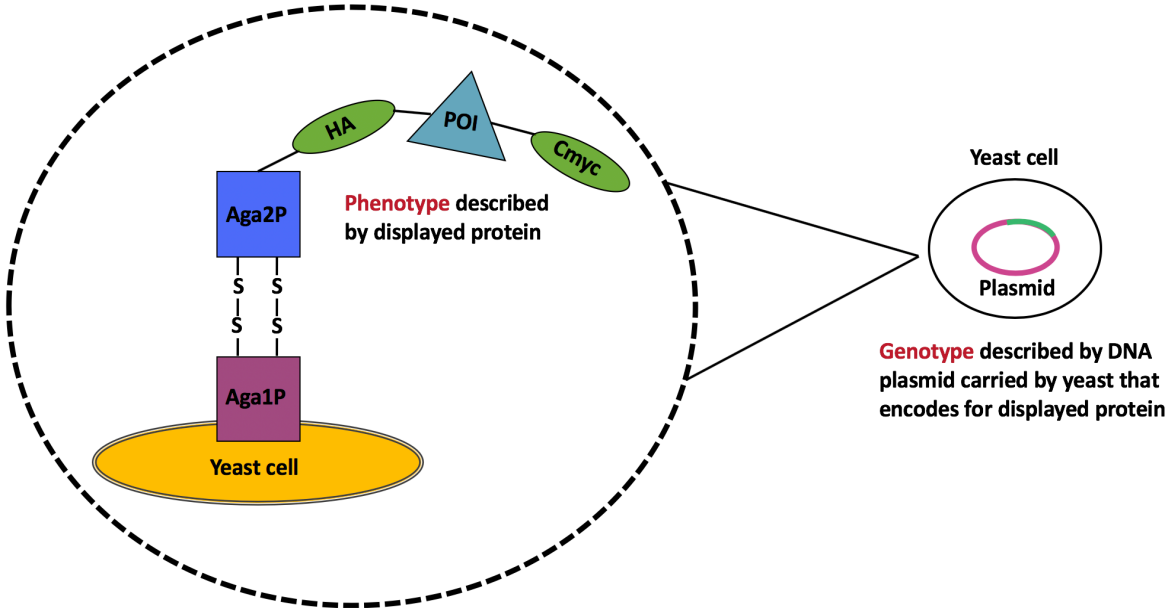


Figure 1.2. Yeast surface display of a protein. The protein of interest (POI) is genetically encoded as a fusion to Aga2P. The expression of this fusion is under the control of a galactose inducible promoter. After induction, Aga2P is linked to the Aga1P unit of the yeast mating protein a-agglutinin, thereby tethering the expressed protein of interest to the yeast surface. Epitope tags, like HA and cmyc, are genetically included in the fusion to quantify the expression of the displayed protein using immunofluorescence. Each yeast cell carries a plasmid that encodes the DNA for the expressed protein.

Yeast libraries with diversities up to 10^9 mutants can be produced using homologous recombination mediated plasmid gap repair[131]. After induction of protein expression, each cell displays ~50,000 copies of their encoded protein[112,132]. Epitope tags, like HA and c-myc tags, can be genetically encoded to flank the protein sequence. Accordingly, immunofluorescent detection of these tags can be used to quantify the expression of proteins on the yeast surface.

Mutants with novel binding affinity can be isolated from yeast display combinatorial libraries (Figure 1.3) using a combination of magnetic selection (MACS)[133] and fluorescence activated cell sorting (FACS)[134]. To perform a magnetic selection, the library cells are incubated with target functionalized magnetic beads. Any library cells that bind to the magnetic beads can be isolated using a magnet. In contrast, FACS involves the labeling of the library cell population

with a fluorescently labeled soluble target protein. Using a FACS analyzer, any library cells that bind the target can be isolated by detecting the target protein's fluorescent signal. Yeast surface display has gained popularity in comparison to phage display as yeast display libraries can be sorted using FACS. Phage libraries cannot be sorted via FACS due to their small size[134]. Selections based on FACS are desired as FACS allows for fine affinity discrimination between mutants[135]. Specifically, the higher affinity mutants from a library population can be isolated from the lower affinity mutants based on the level of target binding. This cannot be achieved when libraries are screened against an immobilized target protein, like when performing MACS, as all library cells that bind the target-decorated surface are collected irrespective of affinity.

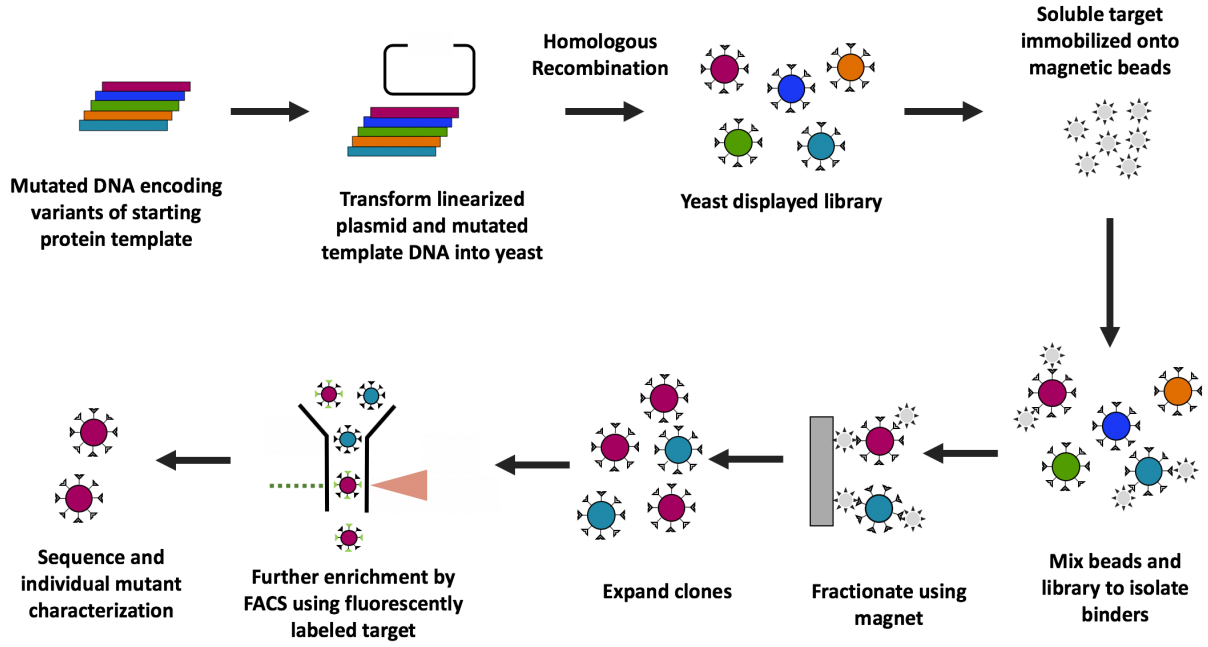


Figure 1.3. Generation and selection of a yeast displayed combinatorial library. A yeast displayed combinatorial library is created via homologous recombination mediated gap repair after transformation of linearized plasmid backbone DNA and a linear piece of DNA encoding the protein variant. When selecting variants for novel binding activity, a magnetic selection is typically performed first. Magnetic beads functionalized with soluble target protein are incubated with the yeast displayed library. Any library cells that bind the target-functionalized magnetic beads are isolated using a magnet. After expansion, this population is generally subjected to Fluorescence Activated Cell Sorting (FACS) where the library cells are incubated with a fluorescently labeled version of the target protein. The library cell population can be sorted into sub-populations based on the extent of fluorescent target labeling. Multiple rounds of FACS sorting can be performed where the selection stringency is increased for each round. After sorting, individual yeast clones can be sequenced for identification followed by biophysical characterization of the individual mutants.

1.3.1.2 mRNA display

Unlike yeast surface display, mRNA display is a completely *in vitro* system that does not rely on the transformation of cells for library generation[120]. Thus, libraries with diversities greater than 10^{12} can be produced using mRNA display in contrast to the diversities of libraries generated using yeast surface display (10^7-10^9)[136] or phage display libraries (10^9-10^{10})[137]. Because mRNA display libraries exhibit larger diversities, there is a greater probability of isolating rare sequences using mRNA display libraries in comparison to libraries generated via cell transformation. Unique chemical properties can be engineered into mRNA display libraries by using non-natural amino acids to expand the protein alphabet. The incorporation of non-canonical amino acids is more efficient using a cell free system, like mRNA display[138,139].

The genotype-phenotype linkage is unique for mRNA display; Each protein is covalently linked to its encoding mRNA[28]. Because the linkage between genotype and phenotype is through a covalent bond, a wider range of selection conditions can be used when screening mRNA display libraries in comparison to systems utilizing biological display hosts[140]. When screening libraries that are displayed using a biological host, all selection conditions must be compatible with maintaining the hosts' viability and ability to replicate. In contrast, the covalent linkage of mRNA and protein can be maintained even if harsh solvents, salts, or changes in temperature are applied. Thus, mRNA display affords the evolution of proteins with specific characteristics (enhanced stability, pH sensitivity, temperature sensitivity), which may be difficult when utilizing other systems[141].

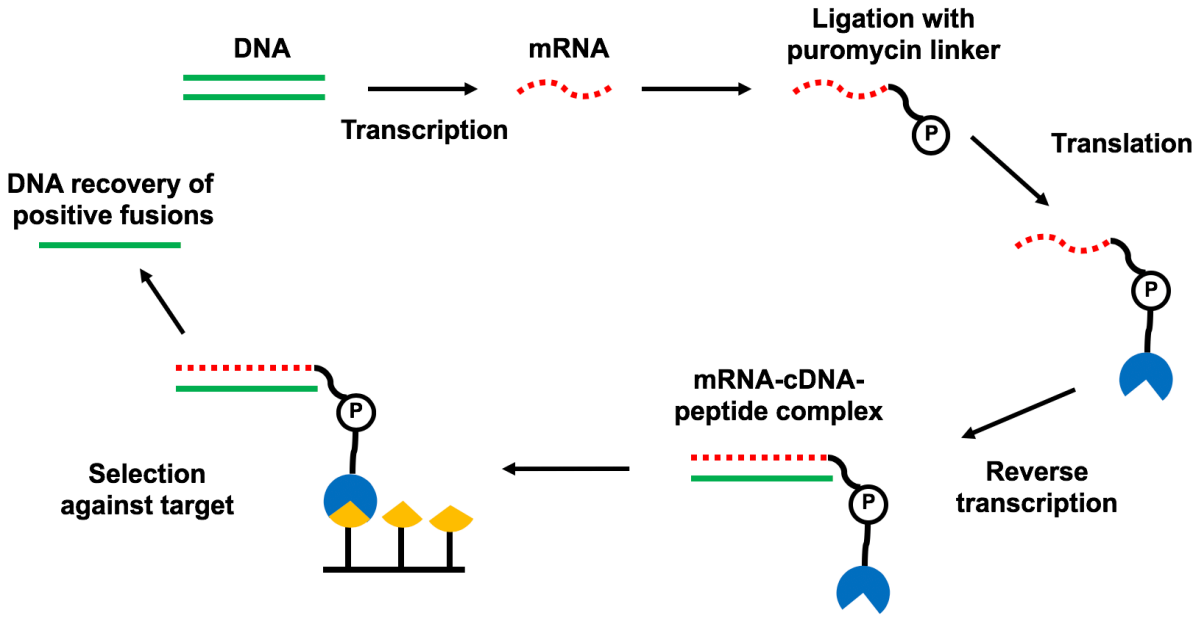


Figure 1.4. Generation and selection of a mRNA display combinatorial library. DNA encoding the library mutants is transcribed into mRNA. The mRNA is ligated with an oligonucleotide linker containing a puromycin residue. During *in vitro* translation, the encoding mRNA is linked to the translated peptide via the puromycin linker. To increase stability, the mRNA is reverse transcribed into cDNA. When identifying proteins with novel binding activity, the cDNA/mRNA-protein fusions can be selected against an immobilized form of the target protein. Subsequently, the positive fusions are recovered, and the linked cDNA can be sequenced for identification or used to generate a focused library for additional screening.

mRNA display libraries have been generated for a variety of starting templates, such as linear peptides[121,142], cyclic peptides[72,73,143–146], non-antibody scaffolds[147,148], single domain antibodies[149,150], and antibody fragments[151,152]. To construct a combinatorial library using mRNA display, a large DNA library is transcribed into mRNA (Figure 1.4). After, the 3' end of each encoding mRNA is ligated to a puromycin containing oligonucleotide prior to *in vitro* translation[153]. Puromycin is a translational inhibitor that mimics the aminoacyl end of tRNA. In mRNA display, the puromycin-ligated mRNA moieties are translated by the ribosome until the modified 3' end is reached. The ribosome stalls at the mRNA-DNA interface. Puromycin then enters the A site of the ribosome and stably attaches to the peptide chain, thereby forming a covalent link between the mRNA and the translated protein[154]. Following translation, the resulting mRNA-protein fusions are separated from the reaction components as well as any unlinked mRNA

or protein. To stabilize the nucleic acid region of the fusion, a DNA chain complementary to the linked mRNA is introduced using reverse transcription generating a cDNA/mRNA-protein fusion[155].

When screening mRNA display libraries to identify mutants with novel binding activity, a panning approach is generally used; FACS selections are typically not employed[72,73,147,148]. Specifically, the target protein is immobilized onto a surface, like a magnetic bead. The fusions are incubated with the immobilized target, and any fusions that interact with the target-decorated surface are separated from the fusions that do not bind (Figure 1.4). Subsequently, the positively bound fusions are eluted from the capture surface. The cDNA of the recovered fusions can be amplified by PCR for identification or used as a DNA template to generate a focused library for an additional round of selection.

The mRNA display platform is often used to generate libraries of peptides due to translation inefficiencies of larger proteins[140,149,156]. mRNA-display libraries of cyclic peptides can be generated by crosslinking amino acid residues contained within linear peptide-mRNA fusions. Head to tail peptide cyclization is not compatible with mRNA display because the c-terminus of the peptide is connected to its encoding mRNA via the puromycin linker[157]. Instead, cyclization can be achieved by incorporating amino acid residues with reactive side chains. For example, a linear peptide sequence can be cyclized if a disulfide bond forms between two cysteine side chains[158,159]. Beyond spontaneous cyclization, linear peptides can also be cyclized using a chemical crosslinker that reacts with side chains contained within the peptide sequence. Many chemical crosslinkers are only soluble in an organic solvent, which may be incompatible with display systems that rely on a biological host[143]. However, due to the stability of the mRNA-

peptide covalent linkage, crosslinking reactions requiring organic solvents can be compatible with mRNA-display.

A significant limitation of identifying protein or peptide binders from a mRNA display library is that FACS cannot be performed. It can take many rounds of panning (≥10) to isolate a homogenous, high affinity binding population from a mRNA display library due to the high diversity of these libraries as well as the inefficiencies of panning based selections[148,158]. Binding proteins can be readily identified using platforms, like yeast surface display, that are amenable to FACS. However, there are few examples of using yeast surface display to identify cyclic peptide binders[65,88,160–165]. It would be advantageous to develop a robust yeast surface display platform to generate libraries of cyclic peptides that can be screened via FACS to improve the efficiency of identifying cyclic peptide binders.

1.3.2 Limitations of recombinantly expressed target protein in combinatorial library screening

Traditionally, selections of combinatorial libraries are performed against a soluble version of the target protein that is expressed recombinantly. The most commonly used recombinant host for protein production is *E.coli* due to its fast growth, facile genetic manipulation, and scalability to produce high yields[166,167]. However, mammalian proteins produced in *E.coli* are often misfolded, aggregated, or expressed in an incorrect confirmation[168]. This can be contributed to the limitations of *E.coli* as a prokaryotic host, such as a lack of eukaryotic chaperone proteins, an absence of machinery to perform certain post-translational modifications, a shortage of sub-cellular organelles for protein maturation, or an inability to form complexes with stabilizing binding partners (specific cofactors or other small molecules)[169].

Most of these challenges can be overcome by using a eukaryotic host to express the target protein. When considering mammalian hosts, it still remains challenging and inefficient to recombinantly express proteins utilizing these hosts as many mammalian cell lines require more demanding culture conditions and are more difficult to genetically manipulate[170–172]. Protein yields may also be lower using eukaryotic hosts than those observed with *E.coli*[173,174].

1.3.2 Use of whole cell targets in combinatorial library screening

A significant bottleneck in producing recombinant protein for combinatorial selections is the purification of the expressed protein from the host cell material. To circumvent the need to purify soluble target protein, selections can take place against whole cells displaying the target protein of interest. When considering surface proteins, selections can be performed against cells that natively express the target protein[175–178]. This will promote the isolation of binders that recognize the target protein in its native confirmation. Selection are more likely to result in the isolation of functional binders if the target protein is expressed at a higher level than other surface proteins[175]. However, the native expression of many proteins is low in comparison to other surface proteins[179]. Instead, alternative eukaryotic cell lines, like HEK293 or yeast, can be used to express target proteins at a high level for whole cell selections[179–181].

Selections against whole cell targets generally use an adherent panning method. Briefly, the whole cell targets are incubated with a surface, like a culture dish, and the cells adhere to the surface. After, the library is incuabted with the capture surface. Non-binding library mutants can be removed via pipetting while positively bound library mutants can be eluted from the surface. Despite the success of adherent panning for particular targets, the panning process suffers from several limitations. For example, fixed-volume manual washing is inefficient and laborious. Nonspecifically or weakly bound library mutants are ineffectively removed[182]. Moreover, high

nonspecific binding to the immobilization surface and proteins used for blocking is frequent. The surface area available for target protein presentation is also limited when adherent cells are used[179].

As an alternative, an in solution panning method against whole cell targets has been proposed. When considering the CD146 and CD36 targets, researchers were only able to successfully isolate binders from a phage display library when panning against cells in solution rather than panning against adherent cells. The mutants isolated using the adherent panning approach were specific to the culture flask surface used to immobilize the cells as well as serum proteins that coated the flask surface[183,184]. It is likely that the expression level of the target protein was much lower than the concentration of the serum proteins on the capture surface; therefore, the library mutants were more likely to interact with the serum proteins. In contrast, if a library mutant has affinity for a surface or a non-specific, blocking protein, this mutant is less likely to be carried forward when using a suspension panning approach. More stringent washing can also be performed when target cells are in suspension (i.e. pipetting, vortexing, larger volumes, etc) to reduce the isolation of non-specific binders. If target protein expression levels are low, it is more desirable to have a greater surface area for protein display. Cells in suspension provide an increase in surface area as the entire cell surface is accessible for potential interaction[185].

The success of cell suspension panning when screening phage display libraries suggests that it would be advantageous to extend suspension based panning methods to libraries generated using other display systems, like yeast display and mRNA display[180,186]. It is important to note that phages bound to the target cells can be separated from nonbinding phages by differential centrifugation. When considering the suspension-based selection of a yeast displayed library against a whole cell target, like another yeast cell or a HEK293 cell, centrifugation cannot be easily used to separate target bound library yeast from nonbinding library yeast. During centrifugation,

all library yeast present would pellet, regardless of whether the library yeast was complexed with a target cell. In contrast, when considering libraries generated using mRNA-display, centrifugation does present an option for separating positively bound mRNA-protein hybrids from non-interacting fusions. However, carryover of unbound hybrids due to insufficient pipetting is consistently observed, resulting in strong contamination of the positive leads. Because mRNA-display libraries comprise a large number of copies, even a microliter carry over of non-binders can cause significant contamination. Thus, an alternative separation approach is needed to successfully extend suspension based panning to other display systems.

1.3.3 Biophysical characterization limitations for engineered proteins

Soluble protein target is also needed when characterizing the affinity of an identified binding protein. Traditional methods used to evaluate binding affinity, like isothermal calorimetry[187] or yeast surface titrations, require at least the binding protein or the target protein to be made solubly, if not both proteins. When estimating binding affinity using yeast surface titrations, yeast cells displaying the binding protein of interest are typically incubated with fluorescently labeled target protein[112]. Immunofluorescence detection can be used to estimate the fraction of cell surface fusions bound by the target protein as a function of the incubated target protein concentration. This data can be fit to a monovalent binding isotherm to estimate apparent binding affinity.

When considering whole cell targets, it may be difficult to recombinantly express the target protein of interest. Instead, each individual binding protein can be recombinantly expressed and purified. Interaction of the soluble binding protein can be confirmed using cells displaying the target of interest[188]. In some cases, binding affinity can be estimated when titrating whole cell targets with soluble binding protein[113,189]. While successful, this method limits the number of

putative binding proteins one can evaluate from a screened library population as each binding protein needs to be made recombinantly and purified. Accordingly, it would be beneficial to utilize a cell based assay that does not require soluble protein to confirm protein-protein interactions and more importantly to provide absolute binding affinity estimation of protein interaction pairs.

The most common cell based assay to evaluate the interaction of two proteins is the yeast two hybrid (Y2H) system[190]. Briefly, a protein of interest ("bait") and a putative binding partner ("prey") are fused to two different protein domains. The binding of the bait to the prey results in colocalization of the domains resulting in the stimulation of a readable output (i.e.: reporter gene activation). Variations of this assay can provide some quantitative information on binding affinities. For example, the affinities of different binding proteins can be ranked ordered for a single target[191,192]. However, no variation has been developed that can accurately estimate absolute binding affinities. A cell based assay that can be used for absolute binding affinity estimation will not only be applicable for characterizing engineered binding proteins but also naturally occurring protein-protein interactions.

1.4 Engineered affinity ligands for biologics purification

Unlike traditional protein purification, a unique phenomenon, known as the avidity effect, drives the purification of complex biologics, like cells, viruses, and organelles[193–195]. Often, a surface is immobilized with multiple copies of a capture ligand specific to a protein on the exterior of the biologic. The biologic, too, displays multiple copies of the targeted surface protein. When a heterogenous mixture containing the biologic is incubated with the surface, interactions between the capture surface and the biologic are driven by avidity; specifically, multiple copies of the target protein on the exterior of the biologic interact with multiple copies of the capture ligand. As a

result, the effective binding affinity between the biologic and the capture surface is a sum of the individual ligand-target interactions[196–199].

Accordingly, even ligands that exhibit low to moderate monovalent binding affinity can be used for the capture of biologics if the avidity effect is evoked as capture will be dictated by the combined strength of multiple ligand-target interactions. For example, a Sso7d mutant (K_D ~ 100 nM) was selected from a combinatorial library to bind red clover necrotic mosaic virus[200]. Subsequently, this mutant was immobilized onto magnetic beads enabling the efficient isolation of virus from a complex mixture. This example shows that biologics capture does not require ligands with antibody-like affinities (<10 nM), as avidity propels the interaction between the biologic and the capture surface. Consequently, other non-immunoglobulin ligand formats, like single domain antibodies and cyclic peptides, that are likely to exhibit moderate binding affinities in comparison to antibodies, can be explored to identify ligands for purification. It is likely that purification processes will begin to use non-antibody purification ligands more frequently as affinity matured antibodies may exhibit low stability and are relatively expensive to produce[24,73,195,201]. Thus, it would be valuable to establish a combinatorial screening platform for the rapid identification of ligands specific to proteins on the surface of biologics. Current combinatorial techniques are limited as it is difficult to properly express many biological surface proteins using *E.coli*.

1.5 Thesis Overview

This work focuses on the development of novel strategies to address the limitations of current techniques used to engineer biomolecular recognition agents. First, we demonstrate the functionality of using yeast cells as whole cell targets for screening both yeast surface display and mRNA display combinatorial libraries. We improve the efficiency of using whole cell targets by

developing two different methods to magnetize the target displaying yeast. Another chapter is dedicated to the development of a quantitative cell-cell based assay that can be used to estimate the absolute binding affinity between two surface displayed proteins. Lastly, we describe a novel yeast display platform for the efficient identification and characterization of cyclic peptide binders.

In **Chapter 2**, we establish a platform for the selection of yeast displayed libraries against yeast displayed targets. The separation of library cells that interact with the target displaying cells is dependent on the magnetization of the target cell population. The target yeast cells co-express an iron oxide binding protein which affords magnetization after incubation with iron oxide. This chapter demonstrates that binding proteins specific to the extracellular domain of mammalian membrane proteins can be isolated by screening a combinatorial library against yeast cells expressing these domains on their surface. Ultimately, this work details how yeast cells can be used as an alternative to mammalian whole cell targets when screening combinatorial libraries. This work also demonstrates how the avidity effect can be evoked to capture biologics from complex mixtures using moderate affinity binding proteins.

In **Chapter 3**, we develop a quantitative yeast-yeast two hybrid (qYY2H) assay that can be used to quantitatively estimate the binding affinity (K_D) between two proteins. In this system, each protein of interest is displayed as a yeast surface fusion. We have developed a semi-empirical framework for estimating the K_D of monovalent protein-protein interactions using measurements of multivalent yeast-yeast binding. This system was used to characterize the affinity of binding proteins identified from combinatorial library screens as well as protein-protein interactions identified from cDNA library screens. The system was also evolved to quantitatively characterize binding interactions mediated by post-translational modification.

In **Chapter 4**, we describe a platform for the selection of cyclic peptide libraries generated using mRNA display against yeast displayed targets. The target yeast cells in this work are non-specifically magnetized using iron oxide allowing for the efficient isolation of any mRNA-peptide fusions that bind the target yeast. Cyclic peptides specific to the mitochondrial surface protein TOM22 were identified. The specificity of the identified peptides was confirmed using the qYY2H system.

In **Chapter 5**, we describe the isolation and characterization of cyclic peptide binders from chemically constrained yeast display peptide libraries. Linear peptides displayed on the yeast cell surface are cyclized by disuccinimidyl glutarate-mediated crosslinking of amine groups within the peptide sequence. We show that absolute binding affinities can be estimated using yeast surface displayed cyclic peptides, eliminating the need to chemically synthesize selected peptides for biophysical characterization. The system was used to identify and characterize cyclic peptides specific to lysozyme as well as a more biologically relevant target, Interleukin-17 (IL-17). The identified peptides with specificity for IL-17 were evaluated *in vitro* and *in silico* for their ability to modulate IL-17's interaction with its receptor, which is implicated in a variety of autoimmune disorders.

1.6 References

(1) Bongrand, P. Biomolecular Recognition: The Current Challenge. In *Dynamic Force Spectroscopy and Biomolecular Recognition*; 2012; pp 1–50.

(2) Kahn, K.; Plaxco, K. W. Principles of Biomolecular Recognition. In *Recognition Receptors in Biosensors*; Springer New York, 2010; pp 3–45.

(3) Jones, S.; Thornton, J. M. Principles of Protein-Protein Interactions. *Proc. Natl. Acad. Sci.* **1996**, *93*, 13–20.

(4) Jolma, A.; Yan, J.; Whitington, T.; Toivonen, J.; Nitta, K. R.; Rastas, P.; Morgunova, E.; Enge, M.; Taipale, M.; Wei, G.; et al. DNA-Binding Specificities of Human Transcription Factors. *Cell* **2013**, *152* (1–2), 327–339.

(5) Hallikas, O.; Palin, K.; Sinjushina, N.; Rautiainen, R.; Partanen, J.; Ukkonen, E.; Taipale, J. Genome-Wide Prediction of Mammalian Enhancers Based on Analysis of Transcription-Factor Binding Affinity. *Cell* **2006**, *124* (1), 47–59.

(6) Luscombe, N. M. Amino Acid-Base Interactions: A Three-Dimensional Analysis of Protein-DNA Interactions at an Atomic Level. *Nucleic Acids Res.* **2001**, *29* (13), 2860–2874.

(7) Valuchova, S.; Fulnecek, J.; Petrov, A. P.; Tripsianes, K.; Riha, K. A Rapid Method for Detecting Protein-Nucleic Acid Interactions by Protein Induced Fluorescence Enhancement. *Sci. Rep.* **2016**, *6* (1), 1–10.

(8) Arnaud, J.; Audfray, A.; Imberty, A. Binding Sugars: From Natural Lectins to Synthetic Receptors and Engineered Neolectins. *Chem. Soc. Rev.* **2013**, *42* (11), 4798–4813.

(9) Savir, Y.; Tlusty, T. The Ribosome as an Optimal Decoder: A Lesson in Molecular Recognition. *Cell* **2013**, *153* (2), 471–479.

(10) Nilofer, C.; Sukhwal, A.; Mohanapriya, A.; Kangueane, P. Protein-Protein Interfaces Are VdW Dominant with Selective H-Bonds and (or) Electrostatics towards Broad Functional Specificity. *Bioinformation* **2017**, *13* (06), 164–173.

(11) Janeway Jr, C. A.; Travers, P.; Walport, M.; Shlomchik, M. J. The Interaction of the Antibody Molecule with Specific Antigen. In *Immunobiology: The Immune System in Health and Disease*; 2001; pp 122–127.

(12) Inukai, S.; Kock, K. H.; Bulyk, M. L. Transcription Factor–DNA Binding: Beyond Binding Site Motifs. *Curr. Opin. Genet. Dev.* **2017**, *43*, 110–119.

(13) Borrebaeck, C. A. K. Antibodies in Diagnostics - From Immunoassays to Protein Chips. *Immunol. Today* **2000**, *21* (8), 379–382.

(14) Al-Amri, S. S.; Hashem, A. M. Qualitative and Quantitative Determination of MERS-CoV S1-Specific Antibodies Using ELISA. In *Methods in Molecular Biology*; Humana Press Inc., 2020; Vol. 2099, pp 127–133.

(15) Chae, W.; Kim, P.; Kim, H.; Cheong, Y. C.; Kim, Y. S.; Kang, S. M.; Seong, B. L.

Hemagglutinin Quantitative ELISA-Based Potency Assay for Trivalent Seasonal Influenza Vaccine Using Group-Specific Universal Monoclonal Antibodies. *Sci. Rep.* **2019**, *9* (1), 1–12.

(16) Wang, L.; Degheidy, H.; Abbasi, F.; Mostowski, H.; Marti, G.; Bauer, S.; Hoffman, R. A.; Gaigalas, A. K. Quantitative Flow Cytometry Measurements in Antibodies Bound per Cell Based on a CD4 Reference. *Curr. Protoc. Cytom.* **2016**, *75* (1), 1.29.1-1.29.14.

(17) Andree, K. C.; Mentink, A.; Nguyen, A. T.; Goldsteen, P.; Van Dalum, G.; Broekmaat, J. J.; Van Rijn, C. J. M.; Terstappen, L. W. M. M. Tumor Cell Capture from Blood by Flowing across Antibody-Coated Surfaces. *Lab Chip* **2019**, *19* (6), 1006–1012.

(18) Hornig-Do, H.-T.; Günther, G.; Bust, M.; Lehnartz, P.; Bosio, A.; Wiesner, R. J. Isolation of Functional Pure Mitochondria by Superparamagnetic Microbeads. *Anal. Biochem.* **2009**, *389* (1), 1–5.

(19) Cupo, A.; Portillo, V. M. C.; Gelfand, P.; Yasmeen, A.; Klasse, P. J.; Moore, J. P. Optimizing the Production and Affinity Purification of HIV-1 Envelope Glycoprotein SOSIP Trimers from Transiently Transfected CHO Cells. *PLoS One* **2019**, *14* (4).

(20) Gera, N.; Hussain, M.; Wright, R. C.; Rao, B. M. Highly Stable Binding Proteins Derived from the Hyperthermophilic Sso7d Scaffold. *J. Mol. Biol.* **2011**, *409* (4), 601–616.

(21) Škrlec, K.; Štrukelj, B.; Berlec, A. Non-Immunoglobulin Scaffolds: A Focus on Their Targets. *Trends Biotechnol.* **2015**, *33* (7), 408–418.

(22) Chen, T. F.; De Picciotto, S.; Hackel, B. J.; Wittrup, K. D. Engineering Fibronectin-Based Binding Proteins by Yeast Surface Display. In *Methods in Enzymology*; Academic Press Inc., 2013; Vol. 523, pp 303–326.

(23) Lee, C. H.; Park, K. J.; Sung, E. S.; Kim, A.; Choi, J. Da; Kim, J. S.; Kim, S. H.; Kwon, M. H.; Kim, Y. S. Engineering of a Human Kringle Domain into Agonistic and Antagonistic Binding Proteins Functioning in Vitro and in Vivo. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (21), 9567–9571.

(24) Chu, W.; Prodromou, R.; Day, K. N.; Schneible, J. D.; Bacon, K. B.; Bowen, J. D.; Kilgore, R. E.; Catella, C. M.; Moore, B. D.; Mabe, M. D.; et al. Peptides and Pseudopeptide Ligands: A Powerful Toolbox for the Affinity Purification of Current and next-Generation Biotherapeutics. *J. Chromatogr. A* **2021**, *1635*, 461632.

(25) Wada, A. Development of Next-Generation Peptide Binders Using in Vitro Display Technologies and Their Potential Applications. *Front. Immunol.* **2013**, *4* (AUG).

(26) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science.* **1985**, *228* (4705), 1315–1317.

(27) Boder, E. T.; Wittrup, K. D. Yeast Surface Display for Screening Combinatorial Polypeptide Libraries. *Nat. Biotechnol.* **1997**, *15* (6), 553–557.

(28) 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), 12297–12302.

(29) Sela-Culang, I.; Kunik, V.; Ofran, Y. The Structural Basis of Antibody-Antigen Recognition. *Front. Immunol.* **2013**, *4* (OCT).

(30) Leenaars, M.; Hendriksen, C. F. M. Critical Steps in the Production of Polyclonal and Monoclonal Antibodies: Evaluation and Recommendations. *ILAR J.* **2005**, *46* (3), 269–279.

(31) Busby, M.; Xue, C.; Li, C.; Farjoun, Y.; Gienger, E.; Yofe, I.; Gladden, A.; Epstein, C. B.; Cornett, E. M.; Rothbart, S. B.; et al. Systematic Comparison of Monoclonal versus Polyclonal Antibodies for Mapping Histone Modifications by ChIP-Seq. *Epigenetics Chromatin* **2016**, *9* (1), 49.

(32) Lipman, N. S.; Jackson, L. R.; Trudel, L. J.; Weis-Garcia, F. Monoclonal Versus Polyclonal Antibodies: Distinguishing Characteristics, Applications, and Information Resources. *ILAR J.* **2005**, *46* (3), 258–268.

(33) Liu, J. K. H. The History of Monoclonal Antibody Development - Progress, Remaining Challenges and Future Innovations. *Ann. Med. Surg.* **2014**, *3* (4), 113–116.

(34) Demain, A. L.; Vaishnav, P. Production of Recombinant Proteins by Microbes and Higher Organisms. *Biotechnol. Adv.* **2009**, *27* (3), 297–306.

(35) Li, F.; Vijayasankaran, N.; Shen, A.; Kiss, R.; Amanullah, A. Cell Culture Processes for Monoclonal Antibody Production. *MAbs* **2010**, *2* (5), 466–479.

(36) Li, H.; Sethuraman, N.; Stadheim, T. A.; Zha, D.; Prinz, B.; Ballew, N.; Bobrowicz, P.; Choi, B. K.; Cook, W. J.; Cukan, M.; et al. Optimization of Humanized IgGs in Glycoengineered Pichia Pastoris. *Nat. Biotechnol.* **2006**, *24* (2), 210–215.

(37) de Marco, A. Strategies for Successful Recombinant Expression of Disulfide Bond-Dependent Proteins in Escherichia Coli. *Microb. Cell Fact.* **2009**, *8* (1), 26.

(38) Simmons, L. C.; Reilly, D.; Klimowski, L.; Raju, T. S.; Meng, G.; Sims, P.; Hong, K.; Shields, R. L.; Damico, L. A.; Rancatore, P.; et al. Expression of Full-Length Immunoglobulins in Escherichia Coli: Rapid and Efficient Production of Aglycosylated Antibodies. *J. Immunol. Methods* **2002**, *263* (1–2), 133–147.

(39) Richards, D. A. Exploring Alternative Antibody Scaffolds: Antibody Fragments and Antibody Mimics for Targeted Drug Delivery. *Drug Discov. Today Technol.* **2018**, *30*, 35–46.

(40) Deonarain, M.; Yahioglu, G.; Stamati, I.; Pomowski, A.; Clarke, J.; Edwards, B.; Diez-Posada, S.; Stewart, A. Small-Format Drug Conjugates: A Viable Alternative to ADCs for Solid Tumours? *Antibodies* **2018**, *7* (2), 16.

(41) Ahmad, Z. A.; Yeap, S. K.; Ali, A. M.; Ho, W. Y.; Alitheen, N. B. M.; Hamid, M. ScFv Antibody: Principles and Clinical Application. *Clin. Dev. Immunol.* **2012**, *2012*.

(42) Wesolowski, J.; Alzogaray, V.; Reyelt, J.; Unger, M.; Juarez, K.; Urrutia, M.; Cauerhff, A.; Danquah, W.; Rissiek, B.; Scheuplein, F.; et al. Single Domain Antibodies: Promising Experimental and Therapeutic Tools in Infection and Immunity. *Med. Microbiol. Immunol.* **2009**, *198* (3), 157–174.

(43) De Meyer, T.; Muyldermans, S.; Depicker, A. Nanobody-Based Products as Research and Diagnostic Tools. *Trends Biotechnol.* **2014**, *32* (5), 263–270.

(44) Järviluoma, A.; Strandin, T.; Lülf, S.; Bouchet, J.; Mäkelä, A. R.; Geyer, M.; Benichou, S.; Saksela, K. High-Affinity Target Binding Engineered via Fusion of a Single-Domain Antibody Fragment with a Ligand-Tailored SH3 Domain. *PLoS One* **2012**, *7* (7), e40331.

(45) Li, T.; Bourgeois, J.; Celli, S.; Glacial, F.; Le Sourd, A.; Mecheri, S.; Weksler, B.; Romero, I.; Couraud, P.; Rougeon, F.; et al. Cell-penetrating Anti-GFAP VHH and Corresponding Fluorescent Fusion Protein VHH-GFP Spontaneously Cross the Blood-brain Barrier and Specifically Recognize Astrocytes: Application to Brain Imaging. *FASEB J.* **2012**, *26* (10), 3969–3979.

(46) Vaks, L.; Benhar, I. Production of Stabilized ScFv Antibody Fragments in the E. Coli Bacterial Cytoplasm. *Methods Mol. Biol.* **2014**, *1060*, 171–184.

(47) Bannas, P.; Hambach, J.; Koch-Nolte, F. Nanobodies and Nanobody-Based Human Heavy Chain Antibodies as Antitumor Therapeutics. *Front. Immunol.* **2017**, *8* (NOV), 1603.

(48) Wörn, A.; Plückthun, A. Stability Engineering of Antibody Single-Chain Fv Fragments. *J. Mol. Biol.* **2001**, *305* (5), 989–1010.

(49) Muyldermans, S. Nanobodies: Natural Single-Domain Antibodies. *Annu. Rev. Biochem.* **2013**, *82* (1), 775–797.

(50) Harding, F. A.; Stickler, M. M.; Razo, J.; DuBridge, R. B. The Immunogenicity of Humanized and Fully Human Antibodies: Residual Immunogenicity Resides in the CDR Regions. *MAbs* **2010**, *2* (3), 256–265.

(51) Koch-Nolte, F.; Reyelt, J.; Schößow, B.; Schwarz, N.; Scheuplein, F.; Rothenburg, S.; Haag, F.; Alzogaray, V.; Cauerhff, A.; Goldbaum, F. A. Single Domain Antibodies from Llama Effectively and Specifically Block T Cell Ecto-ADP-ribosyltransferase ART2.2 in Vivo . *FASEB J.* **2007**, *21* (13), 3490–3498.

(52) Schmitz, K. R.; Bagchi, A.; Roovers, R. C.; Van Bergen En Henegouwen, P. M. P.; Ferguson, K. M. Structural Evaluation of EGFR Inhibition Mechanisms for Nanobodies/VHH Domains. *Structure* **2013**, *21* (7), 1214–1224.

(53) Hosse, R. J. A New Generation of Protein Display Scaffolds for Molecular Recognition. *Protein Sci.* **2006**, *15* (1), 14–27.

(54) Simeon, R.; Chen, Z. In Vitro-Engineered Non-Antibody Protein Therapeutics. *Protein Cell* **2018**, *9* (1), 3–14.

(55) Miller, E. A.; Traxlmayr, M. W.; Shen, J.; Sikes, H. D. Activity-Based Assessment of an Engineered Hyperthermophilic Protein as a Capture Agent in Paper-Based Diagnostic Tests. *Mol. Syst. Des. Eng.* **2016**, *1* (4), 377–381.

(56) Stern, L. A.; Case, B. A.; Hackel, B. J. Alternative Non-Antibody Protein Scaffolds for Molecular Imaging of Cancer. *Curr. Opin. Chem. Eng.* **2013**, *2* (4), 425–432.

(57) Fiedler, E.; Fiedler, M.; Proetzel, G.; Scheuermann, T.; Fiedler, U.; Rudolph, R. Affilin™

Molecules: Novel Ligands for Bioseparation. *Food Bioprod. Process.* **2006**, *84* (1 C), 3–8.

(58) Nygren, P. Å. Alternative Binding Proteins: Affibody Binding Proteins Developed from a Small Three-Helix Bundle Scaffold. *FEBS J.* **2008**, *275* (11), 2668–2676.

(59) Boersma, Y. L.; Plückthun, A. DARPins and Other Repeat Protein Scaffolds: Advances in Engineering and Applications. *Curr. Opin. Biotechnol.* **2011**, *22* (6), 849–857.

(60) Lipovsek, D. Adnectins: Engineered Target-Binding Protein Therapeutics. *Protein Eng. Des. Sel.* **2011**, *24* (1–2), 3–9.

(61) Silverman, J.; Lu, Q.; Bakker, A.; To, W.; Duguay, A.; Alba, B. M.; Smith, R.; Rivas, A.; Li, P.; Le, H.; et al. Multivalent Avimer Proteins Evolved by Exon Shuffling of a Family of Human Receptor Domains. *Nat. Biotechnol.* **2005**, *23* (12), 1556–1561.

(62) Skerra, A. Alternative Binding Proteins: Anticalins - Harnessing the Structural Plasticity of the Lipocalin Ligand Pocket to Engineer Novel Binding Activities. *FEBS J.* **2008**, *275* (11), 2677–2683.

(63) Grabulovski, D.; Kaspar, M.; Neri, D. A Novel, Non-Immunogenic Fyn SH3-Derived Binding Protein with Tumor Vascular Targeting Properties. *J. Biol. Chem.* **2007**, *282* (5), 3196–3204.

(64) Williams, A.; Baird, L. G. DX-88 and HAE: A Developmental Perspective. *Transfus. Apher. Sci.* **2003**, *29* (3), 255–258.

(65) Kintzing, J. R.; Cochran, J. R. Engineered Knottin Peptides as Diagnostics, Therapeutics, and Drug Delivery Vehicles. *Curr. Opin. Chem. Biol.* **2016**, *34*, 143–150.

(66) Yu, X.; Yang, Y.-P.; Dikici, E.; Deo, S. K.; Daunert, S. Beyond Antibodies as Binding Partners: The Role of Antibody Mimetics in Bioanalysis. *Annu. Rev. Anal. Chem.* **2017**, *10* (1), 293–320.

(67) Ladenstein, R.; Antranikian, G. Proteins from Hyperthermophiles: Stability and Enzymatic Catalysis Close to the Boiling Point of Water. *Adv. Biochem. Eng. Biotechnol.* **1998**, *61*, 37–85.

(68) Karshikoff, A.; Ladenstein, R. Ion Pairs and the Thermotolerance of Proteins from Hyperthermophiles: A “traffic Rule” for Hot Roads. *Trends Biochem. Sci.* **2001**, *26* (9), 550–557.

(69) Mao, Y. J.; Sheng, X. R.; Pan, X. M. The Effects of NaCl Concentration and PH on the Stability of Hyperthermophilic Protein Ssh10b. *BMC Biochem.* **2007**, *8* (1), 28.

(70) Daniel, R. M.; Cowan, D. A.; Morgan, H. W.; Curran, M. P. A Correlation between Protein Thermostability and Resistance to Proteolysis. *Biochem. J.* **1982**, *207* (3), 641–644.

(71) Knapp, S.; Karshikoff, A.; Berndt, K. D.; Christova, P.; Atanasov, B.; Ladenstein, R. Thermal Unfolding of the DNA-Binding Protein Sso7d from the Hyperthermophile Sulfolobus Solfataricus. *J. Mol. Biol.* **1996**, *264* (5), 1132–1144.

(72) Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. Design of Cyclic Peptides That

Bind Protein Surfaces with Antibody-like Affinity. *ACS Chem. Biol.* **2007**, *2* (9), 625–634.

(73) Menegatti, S.; Hussain, M.; Naik, A. D.; Carbonell, R. G.; Rao, B. M. MRNA Display Selection and Solid-Phase Synthesis of Fc-Binding Cyclic Peptide Affinity Ligands. *Biotechnol. Bioeng.* **2013**, *110* (3), 857–870.

(74) Wang, X. S.; Chen, P. H. C.; Hampton, J. T.; Tharp, J. M.; Reed, C. A.; Das, S. K.; Wang, D. S.; Hayatshahi, H. S.; Shen, Y.; Liu, J.; et al. A Genetically Encoded, Phage-Displayed Cyclic-Peptide Library. *Angew. Chemie - Int. Ed.* **2019**, *58* (44), 15904–15909.

(75) Hall, P. R.; Hjelle, B.; Njus, H.; Ye, C.; Bondu-Hawkins, V.; Brown, D. C.; Kilpatrick, K. A.; Larson, R. S. Phage Display Selection of Cyclic Peptides That Inhibit Andes Virus Infection. *J. Virol.* **2009**, *83* (17), 8965–8969.

(76) Liu, W.; Wu, C. A Mini-Review and Perspective on Multicyclic Peptide Mimics of Antibodies. *Chinese Chem. Lett.* **2018**, *29* (7), 1063–1066.

(77) Deyle, K.; Kong, X.-D.; Heinis, C. Phage Selection of Cyclic Peptides for Application in Research and Drug Development. *Acc. Chem. Res.* **2017**, *50* (8), 1866–1874.

(78) Choi, J. S.; Joo, S. H. Recent Trends in Cyclic Peptides as Therapeutic Agents and Biochemical Tools. *Biomol. Ther.* **2020**, *28* (1), 18–24.

(79) Mendive-Tapia, L.; Wang, J.; Vendrell, M. Fluorescent Cyclic Peptides for Cell Imaging. *Pept. Sci.* **2021**, *113* (1), e24181.

(80) Firer, M. A.; Gellerman, G. Targeted Drug Delivery for Cancer Therapy: The Other Side of Antibodies. *J. Hematol. Oncol.* **2012**, *5* (1), 70.

(81) Roveri, M.; Bernasconi, M.; Leroux, J. C.; Luciani, P. Peptides for Tumor-Specific Drug Targeting: State of the Art and Beyond. *J. Mater. Chem. B* **2017**, *5* (23), 4348–4364.

(82) Dougherty, P. G.; Sahni, A.; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **2019**, *119* (17), 10241–10287.

(83) Rhodes, C. A.; Pei, D. Bicyclic Peptides as Next-Generation Therapeutics. *Chem. - A Eur. J.* **2017**, *23* (52), 12690–12703.

(84) Zorzi, A.; Deyle, K.; Heinis, C. Cyclic Peptide Therapeutics: Past, Present and Future. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24–29.

(85) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-Encoded Combinatorial Chemical Libraries Based on Bicyclic Peptides. *Nat. Chem. Biol.* **2009**, *5* (7), 502–507.

(86) Yin, Y.; Fei, Q.; Liu, W.; Li, Z.; Suga, H.; Wu, C. Chemical and Ribosomal Synthesis of Topologically Controlled Bicyclic and Tricyclic Peptide Scaffolds Primed by Selenoether Formation. *Angew. Chemie Int. Ed.* **2019**, *58* (15), 4880–4885.

(87) Bashiruddin, N. K.; Nagano, M.; Suga, H. Synthesis of Fused Tricyclic Peptides Using a Reprogrammed Translation System and Chemical Modification. *Bioorg. Chem.* **2015**, *61*, 45–50.

(88) Hetrick, K. J.; Walker, M. C.; Van Der Donk, W. A. Development and Application of

Yeast and Phage Display of Diverse Lanthipeptides. *ACS Cent. Sci.* **2018**, *4* (4), 458–467.

(89) 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. *Biochemistry* **1998**, *37* (48), 16839–16845.

(90) Yudin, A. K. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem. Sci.* **2015**, *6* (1), 30–49.

(91) Werle, M.; Bernkop-Schnü rch, A. Strategies to Improve Plasma Half Life Time of Peptide and Protein Drugs Review Article. *Amino Acids* **2006**, *30*, 351–367.

(92) Li, X.; Wang, S.; Zhu, X.; Zhangsun, D.; Wu, Y.; Luo, S. Effects of Cyclization on Activity and Stability of α-Conotoxin TxIB. *Mar. Drugs* **2020**, *18* (4), 180.

(93) Steiner, K.; Schwab, H. Recent Advances in Rational Approaches for Enzyme Engineering. *Comput. Struct. Biotechnol. J.* **2012**, *2* (3), e201209010.

(94) Brustad, E. M.; Arnold, F. H. Optimizing Non-Natural Protein Function with Directed Evolution. *Curr. Opin. Chem. Biol.* **2011**, *15* (2), 201–210.

(95) Tobin, M. B.; Gustafsson, C.; Huisman, G. W. Directed Evolution: The "rational" Basis for "Irrational" Design. *Curr. Opin. Struct. Biol.* **2000**, *10* (4), 421–427.

(96) Cobb, R. E.; Chao, R.; Zhao, H. Directed Evolution: Past, Present, and Future. *AIChE J.* **2013**, *59* (5), 1432–1440.

(97) Cadwell, R. C.; Joyce, G. F. Randomization of Genes by PCR Mutagenesis. *Genome Res.* **1992**, *2* (1), 28–33.

(98) Stemmer, W. P. C. Rapid Evolution of a Protein in Vitro by DNA Shuffling. *Nature* **1994**, *370* (6488), 389–391.

(99) Galán, A.; Comor, L.; Horvatić, A.; Kuleš, J.; Guillemin, N.; Mrljak, V.; Bhide, M. Library-Based Display Technologies: Where Do We Stand? *Mol. Biosyst.* **2016**, *12* (8), 2342–2358.

(100) Cobb, R. E.; Sun, N.; Zhao, H. Directed Evolution as a Powerful Synthetic Biology Tool. *Methods* **2013**, *60* (1), 81–90.

(101) Zaccolo, M.; Williams, D. M.; Brown, D. M.; Gherardi, E. An Approach to Random Mutagenesis of DNA Using Mixtures of Triphosphate Derivatives of Nucleoside Analogues. *J. Mol. Biol.* **1996**, *255* (4), 589–603.

(102) Leung, D. W.; Chen, E.; Goeddel, D. V. A Method for Random Mutagenesis of a Defined DNA Segment Using a Modified Polymerase Chain Reaction. *Technique* **1989**, *1*, 11–15.

(103) Chica, R. A.; Doucet, N.; Pelletier, J. N. Semi-Rational Approaches to Engineering Enzyme Activity: Combining the Benefits of Directed Evolution and Rational Design. *Curr. Opin. Biotechnol.* **2005**, *16* (4), 378–384.

(104) Jacobs, T. M.; Yumerefendi, H.; Kuhlman, B.; Leaver-Fay, A. SwiftLib: Rapid Degenerate-Codon-Library Optimization through Dynamic Programming. *Nucleic Acids*

Res. **2015**, *43* (5), e34.

(105) Meyer, A. J.; Ellefson, J. W.; Ellington, A. D. Library Generation by Gene Shuffling. *Curr. Protoc. Mol. Biol.* **2013**, *105* (SUPPL.105), Unit.

(106) Hossain, G. S.; Shin, H. D.; Li, J.; Wang, M.; Du, G.; Liu, L.; Chen, J. Integrating Error-Prone PCR and DNA Shuffling as an Effective Molecular Evolution Strategy for the Production of α-Ketoglutaric Acid by l-Amino Acid Deaminase. *RSC Adv.* **2016**, *6* (52), 46149–46158.

(107) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science.* **1985**, *228* (4705), 1315–1317.

(108) Alfaleh, M. A.; Alsaab, H. O.; Mahmoud, A. B.; Alkayyal, A. A.; Jones, M. L.; Mahler, S. M.; Hashem, A. M. Phage Display Derived Monoclonal Antibodies: From Bench to Bedside. *Front. Immunol.* **2020**, *11*, 1986.

(109) Wu, C. H.; Liu, I. J.; Lu, R. M.; Wu, H. C. Advancement and Applications of Peptide Phage Display Technology in Biomedical Science. *J. Biomed. Sci.* **2016**, *23* (1), 1–14.

(110) Kenrick, S. A.; Daugherty, P. S. Bacterial Display Enables Efficient and Quantitative Peptide Affinity Maturation. *Protein Eng. Des. Sel.* **2010**, *23* (1), 9–17.

(111) Bessette, P. H.; Rice, J. J.; Daugherty, P. S. Rapid Isolation of High-Affinity Protein Binding Peptides Using Bacterial Display. *Protein Eng. Des. Sel.* **2004**, *17* (10), 731–739.

(112) Gera, N.; Hussain, M.; Rao, B. M. Protein Selection Using Yeast Surface Display. *Methods* **2013**, *60* (1), 15–26.

(113) Nguyen, A. W.; Le, K. C.; Maynard, J. A. Identification of High Affinity HER2 Binding Antibodies Using CHO Fab Surface Display. *Protein Eng. Des. Sel.* **2018**, *31* (3), 91.

(114) Beerli, R. R.; Bauer, M.; Buser, R. B.; Gwerder, M.; Muntwiler, S.; Maurer, P.; Saudan, P.; Bachmann, M. F. Isolation of Human Monoclonal Antibodies by Mammalian Cell Display. *Proc. Natl. Acad. Sci.* **2008**, *105* (38), 14336–14341.

(115) Ho, M.; Nagata, S.; Pastan, I. Isolation of Anti-CD22 Fv with High Affinity by Fv Display on Human Cells. *Proc. Natl. Acad. Sci.* **2006**, *103* (25), 9637–9642.

(116) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat. Rev. Genet.* **2015**, *16* (7), 379–394.

(117) Mattheakis, L. C.; Bhatt, R. R.; Dower, W. J. An in Vitro Polysome Display System for Identifying Ligands from Very Large Peptide Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (19), 9022–9026.

(118) Gersuk, G. M.; Corey, M. J.; Corey, E.; Stray, J. E.; Kawasaki, G. H.; Vessella, R. L. High-Affinity Peptide Ligands to Prostate-Specific Antigen Identified by Polysome Selection. *Biochem. Biophys. Res. Commun.* **1997**, *232* (2), 578–582.

(119) Lamla, T.; Erdmann, V. A. Searching Sequence Space for High-Affinity Binding Peptides Using Ribosome Display. *J. Mol. Biol.* **2003**, *329* (2), 381–388.

(120) 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), 12297–12302.

(121) Wilson, D. S.; Keefe, A. D.; Szostak, J. W. The Use of MRNA Display to Select High-Affinity Protein-Binding Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (7), 3750–3755.

(122) Barrick, J. E.; Takahashi, T. T.; Ren, J.; Xia, T.; Roberts, R. W. Large Libraries Reveal Diverse Solutions to an RNA Recognition Problem. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (22), 12374–12378.

(123) Sheehan, J.; Marasco, W. A. Phage and Yeast Display. In *Antibodies for Infectious Diseases*; American Society of Microbiology, 2015; Vol. 3, pp 105–127.

(124) Gai, S. A.; Wittrup, K. D. Yeast Surface Display for Protein Engineering and Characterization. *Curr. Opin. Struct. Biol.* **2007**, *17* (4), 467–473.

(125) Beal, D. M.; Bastow, E. L.; Staniforth, G. L.; Von Der Haar, T.; Freedman, R. B.; Tuite, M. F. Quantitative Analyses of the Yeast Oxidative Protein Folding Pathway in Vitro and in Vivo. *Antioxidants Redox Signal.* **2019**, *31* (4), 261–274.

(126) Boder, E. T.; Raeeszadeh-Sarmazdeh, M.; Price, J. V. Engineering Antibodies by Yeast Display. *Arch. Biochem. Biophys.* **2012**, *526* (2), 99–106.

(127) Kondo, A.; Ueda, M. Yeast Cell-Surface Display - Applications of Molecular Display. *Appl. Microbiol. Biotechnol.* **2004**, *64* (1), 28–40.

(128) Jacobs, P. P.; Ryckaert, S.; Geysens, S.; De Vusser, K.; Callewaert, N.; Contreras, R. Pichia Surface Display: Display of Proteins on the Surface of Glycoengineered Pichia Pastoris Strains. *Biotechnol. Lett.* **2008**, *30* (12), 2173–2181.

(129) Ren, R.; Jiang, Z.; Liu, M.; Tao, X.; Ma, Y.; Wei, D. Display of Adenoregulin with a Novel Pichia Pastoris Cell Surface Display System. *Mol. Biotechnol.* **2007**, *35* (2), 103–108.

(130) Chao, G.; Lau, W. L.; Hackel, B. J.; Sazinsky, S. L.; Lippow, S. M.; Wittrup, K. D. Isolating and Engineering Human Antibodies Using Yeast Surface Display. *Nat. Protoc.* **2006**, *1* (2), 755–768.

(131) Benatuil, L.; Perez, J. M.; Belk, J.; Hsieh, C.-M. An Improved Yeast Transformation Method for the Generation of Very Large Human Antibody Libraries. *Protein Eng. Des. Sel.* **2010**, *23* (4), 155–159.

(132) Kieke, M. C.; Shusta, E. V.; Boder, E. T.; Teyton, L.; Wittrup, K. D.; Kranz, D. M. Selection of Functional T Cell Receptor Mutants from a Yeast Surface-Display Library. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (10), 5651–5656.

(133) Ackerman, M.; Levary, D.; Tobon, G.; Hackel, B.; Orcutt, K. D.; Wittrup, K. D. Highly Avid Magnetic Bead Capture: An Efficient Selection Method for de Novo Protein Engineering Utilizing Yeast Surface Display. *Biotechnol. Prog.* **2009**, *25* (3), 774–783.

(134) VanAntwerp, J. J.; Wittrup, K. D. Fine Affinity Discrimination by Yeast Surface Display

and Flow Cytometry. *Biotechnol. Prog.* **2000**, *16* (1), 31–37.

(135) Bidlingmaier, S.; Su, Y.; Liu, B. Combining Phage and Yeast Cell Surface Antibody Display to Identify Novel Cell Type-Selective Internalizing Human Monoclonal Antibodies. *Methods Mol. Biol.* **2015**, *1319*, 51–63.

(136) Linciano, S.; Pluda, S.; Bacchin, A.; Angelini, A. Molecular Evolution of Peptides by Yeast Surface Display Technology. *Medchemcomm* **2019**, *10* (9), 1569–1580.

(137) Frei, J. C.; Lai, J. R. Protein and Antibody Engineering by Phage Display. In *Methods in Enzymology*; Academic Press Inc., 2016; Vol. 580, pp 45–87.

(138) Gao, W.; Cho, E.; Liu, Y.; Lu, Y. Advances and Challenges in Cell-Free Incorporation of Unnatural Amino Acids into Proteins. *Front. Pharmacol.* **2019**, *10* (611).

(139) Li, S.; Millward, S.; Roberts, R. In Vitro Selection of MRNA Display Libraries Containing an Unnatural Amino Acid. *J. Am. Chem. Soc.* **2002**, *124* (34), 9972–9973.

(140) Wang, H.; Liu, R. Advantages of MRNA Display Selections over Other Selection Techniques for Investigation of Protein-Protein Interactions. *Expert Rev. Proteomics* **2011**, *8* (3), 335–346.

(141) Newton, M. S.; Cabezas-Perusse, Y.; Tong, C. L.; Seelig, B. In Vitro Selection of Peptides and Proteins-Advantages of MRNA Display. *ACS Synth. Biol.* **2020**, *9* (2), 181–190.

(142) Baggio, R.; Burgstaller, P.; Hale, S. P.; Putney, A. R.; Lane, M.; Lipovsek, D.; Wright, M. C.; Roberts, R. W.; Liu, R.; Szostak, J. W.; et al. Identification of Epitope-like Consensus Motifs Using MRNA Display. *J. Mol. Recognit.* **2002**, *15* (3), 126–134.

(143) Millward, S. W.; Takahashi, T. T.; Roberts, R. W. A General Route for Post-Translational Cyclization of MRNA Display Libraries. *J. Am. Chem. Soc.* **2005**, *127* (41), 14142–14143.

(144) Hayashi, Y.; Morimoto, J.; Suga, H. In Vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors. *ACS Chem. Biol.* **2012**, *7* (3), 607–613.

(145) Morimoto, J.; Hayashi, Y.; Suga, H. Discovery of Macrocyclic Peptides Armed with a Mechanism-Based Warhead: Isoform-Selective Inhibition of Human Deacetylase SIRT2. *Angew. Chemie Int. Ed.* **2012**, *51* (14), 3423–3427.

(146) 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), 1562–1570.

(147) Hussain, M.; Gera, N.; Hill, A. B.; Rao, B. M. Scaffold Diversification Enhances Effectiveness of a Superlibrary of Hyperthermophilic Proteins. *ACS Synth. Biol.* **2013**, *2* (1), 6–13.

(148) Xu, L.; Aha, P.; Gu, K.; Kuimelis, R. G.; Kurz, M.; Lam, T.; Lim, A. C.; Liu, H.; Lohse, P. A.; Sun, L.; et al. Directed Evolution of High-Affinity Antibody Mimics Using MRNA

Display. *Chem. Biol.* **2002**, *9* (8), 933–942.

(149) Takahashi, K.; Sunohara, M.; Terai, T.; Kumachi, S.; Nemoto, N. Enhanced MRNA-Protein Fusion Efficiency of a Single-Domain Antibody by Selection of MRNA Display with Additional Random Sequences in the Terminal Translated Regions. *Biophys. Physicobiology* **2017**, *14* (0), 23–28.

(150) Doshi, R.; Chen, B. R.; Vibat, C. R. T.; Huang, N.; Lee, C. W.; Chang, G. In Vitro Nanobody Discovery for Integral Membrane Protein Targets. *Sci. Rep.* **2014**, *4* (1), 1–8.

(151) Nagumo, Y.; Fujiwara, K.; Horisawa, K.; Yanagawa, H.; Doi, N. PURE MRNA Display for in Vitro Selection of Single-Chain Antibodies. *J. Biochem.* **2016**, *159* (5), 519–526.

(152) Sumida, T.; Yanagawa, H.; Doi, N. In Vitro Selection of Fab Fragments by MRNA Display and Gene-Linking Emulsion PCR. *J. Nucleic Acids* **2012**, *2012*, 371379.

(153) 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.

(154) Seelig, B. MRNA Display for the Selection and Evolution of Enzymes from in Vitro-Translated Protein Libraries. *Nat. Protoc.* **2011**, *6* (4), 540–552.

(155) Lipovsek, D.; Plückthun, A. In-Vitro Protein Evolution by Ribosome Display and MRNA Display. *J. Immunol. Methods* **2004**, *290* (1–2), 51–67.

(156) Valencia, C. A.; Cotten, S. W.; Dong, B.; Liu, R. MRNA-Display-Based Selections for Proteins with Desired Functions: A Protease-Substrate Case Study. *Biotechnol. Prog.* **2008**, *24* (3), 561–569.

(157) Josephson, K.; Ricardo, A.; Szostak, J. W. MRNA Display: From Basic Principles to Macrocycle Drug Discovery. *Drug Discov. Today* **2014**, *19* (4), 388–399.

(158) Yamaguchi, J.; Naimuddin, M.; Biyani, M.; Sasaki, T.; Machida, M.; Kubo, T.; Funatsu, T.; Husimi, Y.; Nemoto, N. CDNA Display: A Novel Screening Method for Functional Disulfide-Rich Peptides by Solid-Phase Synthesis and Stabilization of MRNA-Protein Fusions. *Nucleic Acids Res.* **2009**, *37* (16), e108.

(159) Raffler, N. A.; Schneider-Mergener, J.; Famulok, M. A Novel Class of Small Functional Peptides That Bind and Inhibit Human α-Thrombin Isolated by MRNA Display. *Chem. Biol.* **2003**, *10* (1), 69–79.

(160) Van Rosmalen, M.; Janssen, B. M. G.; Hendrikse, N. M.; Van Der Linden, A. J.; Pieters, P. A.; Wanders, D.; De Greef, T. F. A.; Merkx, M. Affinity Maturation of a Cyclic Peptide Handle for Therapeutic Antibodies Using Deep Mutational Scanning. *J. Biol. Chem.* **2017**, *292* (4), 1477–1489.

(161) Ishii, J.; Yoshimoto, N.; Tatematsu, K.; Kuroda, S.; Ogino, C.; Fukuda, H.; Kondo, A. Cell Wall Trapping of Autocrine Peptides for Human G-Protein-Coupled Receptors on the Yeast Cell Surface. *PLoS One* **2012**, *7* (5), e37136.

(162) Jeong, M.-Y.; Rutter, J.; Chou, D. H.-C. Display of Single-Chain Insulin-like Peptides on

a Yeast Surface. *Biochemistry* **2019**, *58* (3), 182–188.

(163) Kimura, R. H.; Levin, A. M.; Cochran, F. V.; Cochran, J. R. Engineered Cystine Knot Peptides That Bind Avβ3, Avβ5, and A5β1 Integrins with Low-Nanomolar Affinity. *Proteins Struct. Funct. Bioinforma.* **2009**, *77* (2), 359–369.

(164) Kimura, R. H.; Jones, D. S.; Jiang, L.; Miao, Z.; Cheng, Z.; Cochran, J. R. Functional Mutation of Multiple Solvent-Exposed Loops in the Ecballium Elaterium Trypsin Inhibitor-II Cystine Knot Miniprotein. *PLoS One* **2011**, *6* (2), e16112.

(165) Silverman, A. P.; Levin, A. M.; Lahti, J. L.; Cochran, J. R. Engineered Cystine-Knot Peptides That Bind Avβ3 Integrin with Antibody-Like Affinities. *J. Mol. Biol.* **2009**, *385* (4), 1064–1075.

(166) Waegeman, H.; Soetaert, W. Increasing Recombinant Protein Production in Escherichia Coli through Metabolic and Genetic Engineering. *J. Ind. Microbiol. Biotechnol.* **2011**, *38* (12), 1891–1910.

(167) Tripathi, N. K.; Sathyaseelan, K.; Jana, A. M.; Rao, P. V. L. High Yield Production of Heterologous Proteins with Escherichia Coli. *Def. Sci. J.* **2009**, *59* (2), 137–146.

(168) Gasser, B.; Saloheimo, M.; Rinas, U.; Dragosits, M.; Rodríguez-Carmona, E.; Baumann, K.; Giuliani, M.; Parrilli, E.; Branduardi, P.; Lang, C.; et al. Protein Folding and Conformational Stress in Microbial Cells Producing Recombinant Proteins: A Host Comparative Overview. *Microb. Cell Fact.* **2008**, *7*, 11.

(169) Dyson, M. R.; Shadbolt, S. P.; Vincent, K. J.; Perera, R. L.; McCafferty, J. Production of Soluble Mammalian Proteins in Escherichia Coli: Identification of Protein Features That Correlate with Successful Expression. *BMC Biotechnol.* **2004**, *4* (1), 32.

(170) Ram, K.; Hatton, D.; Ahuja, S.; Bender, J.; Hunter, A.; Turner, R. Protein Production in Eukaryotic Cells. In *Topics in Medicinal Chemistry*; Springer Verlag, 2017; Vol. 21, pp 1–39.

(171) Hunter, M.; Yuan, P.; Vavilala, D.; Fox, M. Optimization of Protein Expression in Mammalian Cells. *Curr. Protoc. Protein Sci.* **2019**, *95* (1).

(172) Gray, D. Overview of Protein Expression by Mammalian Cells. *Curr. Protoc. Protein Sci.* **1997**, *10* (1), 591.

(173) Merlin, M.; Gecchele, E.; Capaldi, S.; Pezzotti, M.; Avesani, L. Comparative Evaluation of Recombinant Protein Production in Different Biofactories: The Green Perspective. *Biomed Res. Int.* **2014**, *2014*.

(174) Müller, D.; Bayer, K.; Mattanovich, D. Potentials and Limitations of Prokaryotic and Eukaryotic Expression Systems for Recombinant Protein Production – a Comparative View. *Microb. Cell Fact.* **2006**, *5* (Suppl 1), 1–2.

(175) Stern, L. A.; Schrack, I. A.; Johnson, S. M.; Deshpande, A.; Bennett, N. R.; Harasymiw, L. A.; Gardner, M. K.; Hackel, B. J. Geometry and Expression Enhance Enrichment of Functional Yeast-Displayed Ligands via Cell Panning. *Biotechnol. Bioeng.* **2016**, *113* (11), 2328–2341.

(176) Xin, X. W.; Shusta, E. V. The Use of ScFv-Displaying Yeast in Mammalian Cell Surface Selections. *J. Immunol. Methods* **2005**, *304* (1–2), 30–42.

(177) Sánchez-Martín, D.; Sørensen, M. D.; Lykkemark, S.; Sanz, L.; Kristensen, P.; Ruoslahti, E.; Álvarez-Vallina, L. Selection Strategies for Anti-Cancer Antibody Discovery: Searching off the Beaten Path. *Trends Biotechnol.* **2015**, *33* (5), 292.

(178) Popkov, M.; Rader, C.; Barbas, C. F. Isolation of Human Prostate Cancer Cell Reactive Antibodies Using Phage Display Technology. *J. Immunol. Methods* **2004**, *291* (1–2), 137–151.

(179) Alfaleh, M.; Jones, M.; Howard, C.; Mahler, S. Strategies for Selecting Membrane Protein-Specific Antibodies Using Phage Display with Cell-Based Panning. *Antibodies* **2017**, *6* (3), 10.

(180) Zhao, L.; Qu, L.; Zhou, J.; Sun, Z.; Zou, H.; Chen, Y.-Y.; Marks, J. D.; Zhou, Y. High Throughput Identification of Monoclonal Antibodies to Membrane Bound and Secreted Proteins Using Yeast and Phage Display. *PLoS One* **2014**, *9* (10), e111339.

(181) Yoon, H.; Song, J. M.; Ryu, C. J.; Kim, Y. G.; Lee, E. K.; Kang, S.; Kim, S. J. An Efficient Strategy for Cell-Based Antibody Library Selection Using an Integrated Vector System. *BMC Biotechnol.* **2012**, *12*, 62.

(182) Wang, J.; Liu, Y.; Teesalu, T.; Sugahara, K. N.; Kotamrajua, V. R.; Adams, J. D.; Ferguson, B. S.; Gong, Q.; Oh, S. S.; Csordas, A. T.; et al. Selection of Phage-Displayed Peptides on Live Adherent Cells in Microfluidic Channels. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (17), 6909–6914.

(183) Hoogenboom, H. R. Selection-Dominant and Nonaccessible Epitopes on Cell-Surface Receptors Revealed by Cell-Panning with a Large Phage Antibody Library. *Eur. J. Biochem.* **1999**, *260* (3), 774.

(184) Keller, T.; Kalt, R.; Raab, I.; Schachner, H.; Mayrhofer, C.; Kerjaschki, D.; Hantusch, B. Selection of ScFv Antibody Fragments Binding to Human Blood versus Lymphatic Endothelial Surface Antigens by Direct Cell Phage Display. *PLoS One* **2015**, *10* (5).

(185) Lown, P. S.; Hackel, B. J. Magnetic Bead-Immobilized Mammalian Cells Are Effective Targets to Enrich Ligand-Displaying Yeast. *ACS Comb. Sci.* **2020**, *22* (5), 274–284.

(186) Zhou, Y.; Zou, H.; Zhang, S.; Marks, J. D. Internalizing Cancer Antibodies from Phage Libraries Selected on Tumor Cells and Yeast-Displayed Tumor Antigens. *J. Mol. Biol.* **2010**, *404* (1), 88–99.

(187) Duff, M. R.; Jr.; Grubbs, J.; Howell, E. E. Isothermal Titration Calorimetry for Measuring Macromolecule-Ligand Affinity. *J. Vis. Exp.* **2011**, *55* (55), 2796.

(188) Jones, M. L.; Alfaleh, M. A.; Kumble, S.; Zhang, S.; Osborne, G. W.; Yeh, M.; Arora, N.; Hou, J. J. C.; Howard, C. B.; Chin, D. Y.; et al. Targeting Membrane Proteins for Antibody Discovery Using Phage Display. *Sci. Rep.* **2016**, *6*.

(189) Wang, J.; An, L.; Zhao, Y.; Zhang, C.; Li, S.; Ye, C.; Jing, S.; Hang, H. In Vitro Affinity Maturation of Antibody against Membrane-Bound GPCR Molecules. *Appl. Microbiol.*

Biotechnol. **2019**, *103* (18), 7703–7717.

(190) Brückner, A.; Polge, C.; Lentze, N.; Auerbach, D.; Schlattner, U. Yeast Two-Hybrid, a Powerful Tool for Systems Biology. *Int. J. Mol. Sci.* **2009**, *10* (6), 2763–2788.

(191) Younger, D.; Berger, S.; Baker, D.; Klavins, E. High-Throughput Characterization of Protein–Protein Interactions by Reprogramming Yeast Mating. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (46), 12166–12171.

(192) Cluet, D.; Amri, I.; Vergier, B.; Léault, J.; Audibert, A.; Grosjean, C.; Calabrési, D.; Spichty, M. A Quantitative Tri-Fluorescent Yeast Two-Hybrid System: From Flow Cytometry to in Cellula Affinities. *Mol. Cell. Proteomics* **2020**, *19* (4), 701–715.

(193) Vauquelin, G.; Charlton, S. J. Exploring Avidity: Understanding the Potential Gains in Functional Affinity and Target Residence Time of Bivalent and Heterobivalent Ligands. *Br. J. Pharmacol.* **2013**, *168* (8), 1771–1785.

(194) Mammen, M.; Choi, S. K.; Whitesides, G. M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chemie - Int. Ed.* **1998**, *37* (20), 2754–2794.

(195) Bacon, K.; Lavoie, A.; Rao, B. M.; Daniele, M.; Menegatti, S. Past, Present, and Future of Affinity-Based Cell Separation Technologies. *Acta Biomater.* **2020**, *112*, 29–51.

(196) Zhao, W.; Cui, C. H.; Bose, S.; Guo, D.; Shen, C.; Wong, W. P.; Halvorsen, K.; Farokhzad, O. C.; Teo, G. S. L.; Phillips, J. A.; et al. Bioinspired Multivalent DNA Network for Capture and Release of Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (48), 19626–19631.

(197) Sheng, W.; Chen, T.; Tan, W.; Fan, Z. H. Multivalent DNA Nanospheres for Enhanced Capture of Cancer Cells in Microfluidic Devices. *ACS Nano* **2013**, *7* (8), 7067–7076.

(198) Myung, J. H.; Gajjar, K. A.; Saric, J.; Eddington, D. T.; Hong, S. Dendrimer-Mediated Multivalent Binding for the Enhanced Capture of Tumor Cells. *Angew. Chemie - Int. Ed.* **2011**, *50* (49), 11769–11772.

(199) Xie, J.; Wang, J.; Chen, H.; Shen, W.; Sinko, P. J.; Dong, H.; Zhao, R.; Lu, Y.; Zhu, Y.; Jia, L. Multivalent Conjugation of Antibody to Dendrimers for the Enhanced Capture and Regulation on Colon Cancer Cells. *Sci. Rep.* **2015**, *5* (1), 1–10.

(200) Hussain, M.; Lockney, D.; Wang, R.; Gera, N.; Rao, B. M. Avidity-Mediated Virus Separation Using a Hyperthermophilic Affinity Ligand. *Biotechnol. Prog.* **2013**, *29* (1), 237–246.

(201) Rabia, L. A.; Desai, A. A.; Jhajj, H. S.; Tessier, P. M. Understanding and Overcoming Trade-Offs between Antibody Affinity, Specificity, Stability and Solubility. *Biochem. Eng. J.* **2018**, *137*, 365–374.

CHAPTER 2

Screening Yeast Display Libraries Against Magnetized Yeast Cell Targets Enables Efficient Isolation of Membrane Protein Binders

Adapted from Bacon K, Burroughs M, Blain A, Menegatti S, and Rao, B.M. (2019) Screening Yeast Display Libraries against Magnetized Yeast Cell Targets Enables Efficient Isolation of Membrane Protein Binders. *ACS Combinatorial Science 21(12)*: 817-832.

2.1 Introduction

Binding proteins that specifically target membrane proteins are important in several applications, including therapeutics, diagnostics, and the separation of biologics such as organelles and cells. Disease states are often defined by the overexpression of one or more specific membrane proteins[1,2]. Therefore, binding proteins with specificity to these targets are commonly used to direct drug therapies[3,4], to monitor disease advancement using imaging probes[5,6], or to predict disease progression[2]. Furthermore, advances in proteomics and genomics are enabling the identification of new clinically relevant membrane protein targets[7,8]. Binders to membrane proteins are also used to isolate intracellular organelles for basic research or to separate cell populations for transplant purposes[9–11].

Several high throughput combinatorial library screening platforms, such as phage display[12,13], yeast display[14], or mammalian display[15,16], have been used to identify ligands with unique or enhanced binding affinity for a target of interest. Nevertheless, isolation of binders to membrane protein targets can be challenging. Strategies for the isolation of binders to membrane proteins typically utilize a recombinantly expressed extracellular domain of a target protein. Recombinantly produced protein is often misfolded or adopts a conformation that is different from the native protein[17,18]. Misfolding is also compounded by storage and purification conditions that can promote aggregation[19]. Additionally, recombinant proteins are often fused with biological or chemical tags to assist in purification or selection. Binders with affinity for these tags rather than the protein may be isolated[14,20]. Consequently, use of recombinant protein targets may result in the isolation of binders that interact with epitopes that are inaccessible or not present when the membrane protein is natively expressed[21]. These challenges may be overcome by screening libraries against whole cell targets expressing the membrane protein[22]. However, here too, isolation

of specific binders may be difficult as the expression levels of the target protein may be low while other cell surface proteins are present at high density[23].

As an alternative to the aforementioned approaches, yeast surface display has emerged as an attractive strategy for expressing membrane protein targets in the context of structure-function studies as well as screening phage display libraries[24–30]. In yeast surface display, a protein of interest is expressed as an N- or C-terminal fusion to the Aga2 subunit of the yeast mating protein a-agglutinin[14,31]. The Aga2 subunit in turn is linked to the cell wall associated Aga1 subunit, tethering the protein of interest to the cell wall. Notably, Zhao et al. have used yeast surface display to express a variety of extracellular membrane domains[25]. Here, we describe the development and evaluation of a screening strategy wherein yeast displayed membrane protein targets are used for the isolation of binding proteins from a yeast display library.

The ease of expressing a membrane protein or a domain thereof as a cell surface fusion – in contrast to recombinant soluble expression – makes the use of yeast-displayed targets particularly attractive[32]. However, the critical challenge in screening yeast display libraries using yeast displayed targets is the separation of binder cells from the non-binders. To efficiently enable this separation, we investigated a strategy wherein an iron oxide binding protein (SsoFe2) is used for the magnetization of yeast cells displaying a target protein (Figure 2.1)[33]. Briefly, the incubation of yeast cells co-expressing a target protein and SsoFe2 with iron oxide nanopowder results in the magnetization of target displaying yeast. Subsequently, the magnetized target yeast cells are incubated with a yeast display library, and any putative binder cells that complex with the target-displaying yeast are separated using a magnet.

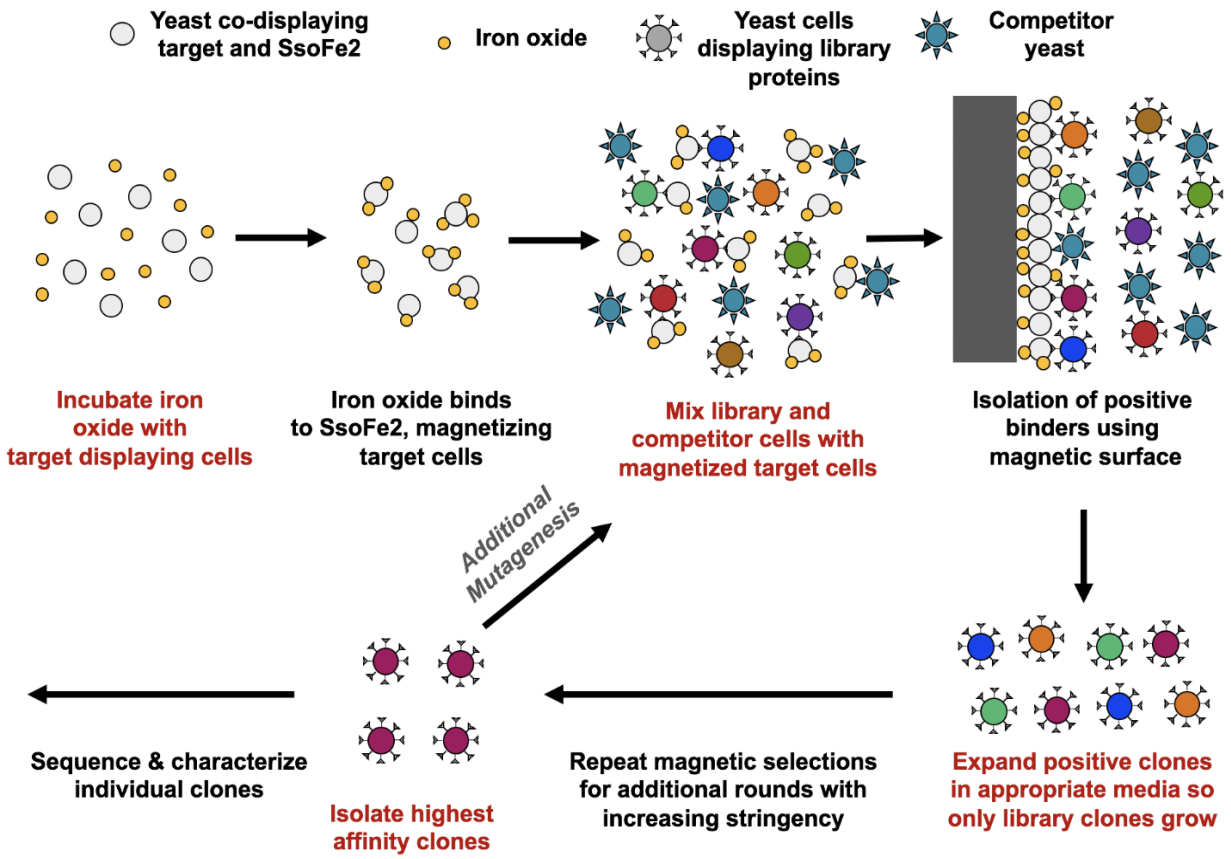


Figure 2.1. Overall strategy for screening a yeast displayed combinatorial library against yeast cells displaying a target protein. The target cells co-express the target protein along with SsoFe2, a protein with affinity for iron oxide. Incubation of the target cells with iron oxide nanoparticles enables magnetization of the target cells. After, the yeast display library is incubated with the magnetized target cells. Competitor yeast cells (cells displaying an irrelevant protein or no protein) are included to reduce non-specific binding. Any library cells that bind the target displaying cells can be expanded in appropriate media and screened further with increasing stringency to isolate higher affinity binders. Subsequently, individual clones can be sequenced and characterized.

We first demonstrated the feasibility of our strategy by performing quantitative studies focused on the enrichment of yeast cells displaying a model binder protein from a mixed population containing non-binder cells. Here, cells displaying the binder were recovered using magnetized yeast displaying the corresponding target protein. Subsequently, we evaluated our screening strategy for the isolation of binding proteins to two membrane protein targets of interest (TOM22 and c-Kit) from two yeast display libraries based on the Sso7d and nanobody scaffolds. Binding proteins for a wide spectrum of targets have been previously isolated from combinatorial

libraries derived from these scaffolds[34,35]. TOM22 is a membrane-associated mitochondrial protein; c-Kit is a cell surface receptor protein that is expressed on hematopoietic stem cells. Notably, antibody binders to the cytosolic domain of TOM22 are used in commercially available reagent kits for the isolation of mitochondria from cell lysates; these reagents serve as a benchmark when evaluating the binding proteins isolated using our strategy.

2.2 Results and Discussion

2.2.1 Co-expression of two proteins as yeast cell surface fusions

To obtain selective expansion of the binder cells that complex with the magnetized target cells in culture, the target-displaying yeast cells and the yeast displayed library cells must utilize yeast surface display plasmids with distinct nutritional selection markers. Towards that end, we sought to co-express the target protein and the iron-oxide binding protein, SsoFe2, as Aga2 fusions using a single plasmid. Two different approaches were investigated, involving the use of a T2A ribosomal skipping peptide from *Thosea asigna* (Figure 2.2A) or a bidirectional *GAL1/GAL10* promoter (Figure 2.2B). The T2A sequence has been successfully used to translate multiple proteins from a single mRNA transcript in *S. cerevisiae*[36,37]. The native yeast display construct contains the bidirectional *GAL1/GAL10* promoter[38].

The functionality of the two co-expression approaches was evaluated for the simultaneous yeast surface display of the Fc portion of human IgG (hFc) and SsoFe2 (Figure 2.2C). For each construct, the expression levels of hFc and SsoFe2 were evaluated by immunofluorescent detection of the c-myc and V5 epitope tags, respectively, using flow cytometry analysis. Both approaches resulted in the co-expression of hFc and SsoFe2 as cell surface fusions (Figures 2.2D-E). However, the cell surface expression level of SsoFe2 was higher when the T2A peptide was used, relative to the case when SsoFe2 expression was under control of the *GAL10* promoter using the bidirectional

promoter construct. Our results are consistent with previous studies by Rosowki et al[37]. The level of protein expression has been shown to be up to five-fold higher when the *GAL1* promoter was used instead of *GAL10*[39,40]. However, display levels are protein specific[41]; use of a stronger promoter may be desirable for proteins that are poorly displayed on the yeast cell surface relative to SsoFe2.

We chose to use the plasmid based on ribosomal skipping to achieve dual display in subsequent experiments because the cell surface display level was similar for the two co-expressed proteins during the initial evaluation of this construct. Note, however, the mechanism of ribosomal skipping and its effect on protein display must be considered when choosing a strategy for the co-display of a target protein and SsoFe2. 2A ribosomal skipping peptides share a conserved GDVEXNPGP sequence. Ribosomal skipping results in "cleavage" due to a failure to form a peptide bond between the C-terminal glycine and proline of this sequence during translation[42,43]. In the context of our plasmid construct, a major portion of the 2A peptide is fused to the C-terminus of the c-myc tag associated with the first translated protein (the target protein). It is important to note that the presence of these additional residues may affect the folding of certain proteins. More importantly, residues in the target protein upstream of the 2A sequence can affect the efficiency of ribosomal skipping. Finally, a proline residue is added to the prepro sequence associated with the second protein (SsoFe2) upon T2A cleavage; however, this did not affect surface display of SsoFe2.

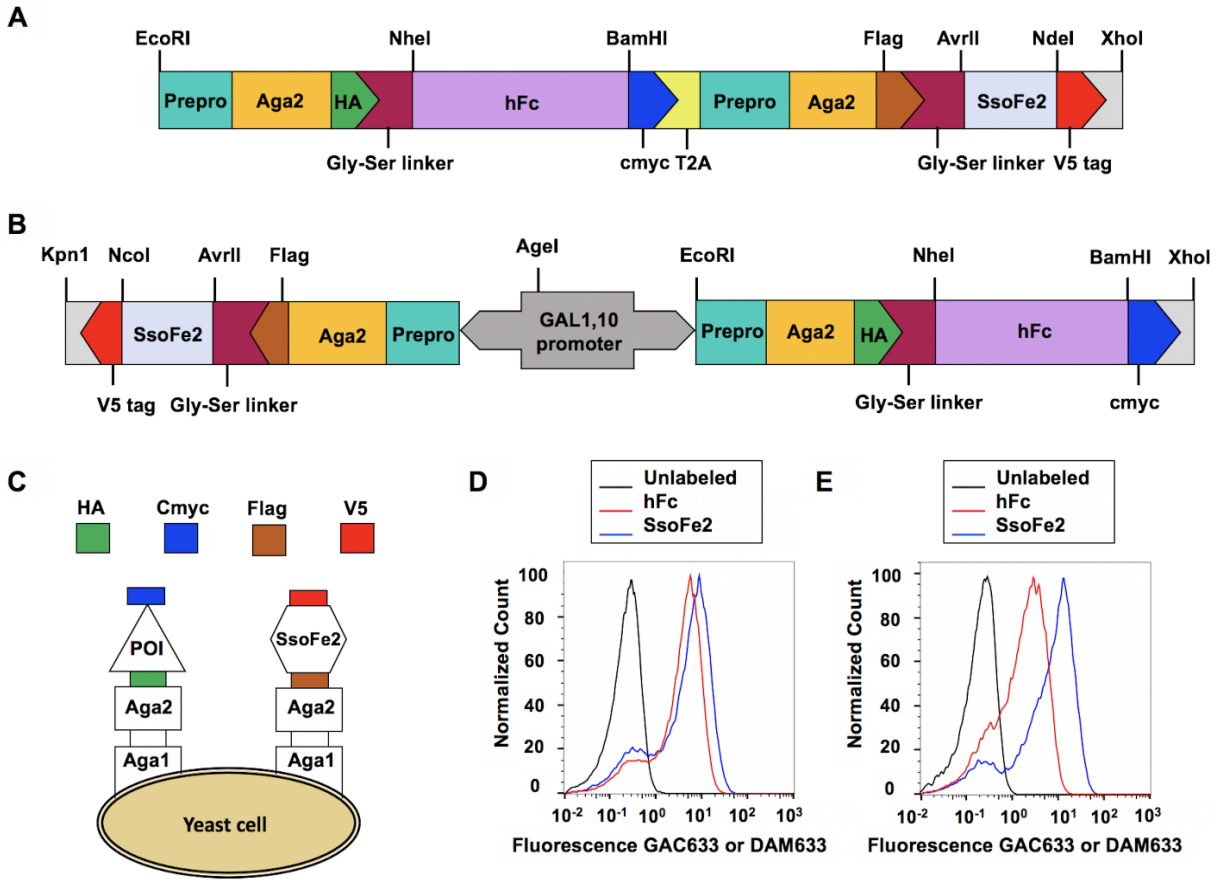


Figure 2.2. Constructs designed for the simultaneous display of two proteins on the surface of yeast as fusions to Aga2. In this example, the Fc portion of human IgG (hFc) and SsoFe2 are co-displayed as cell surface fusions. **(A)** Plasmid construct to achieve dual display of two proteins from a single mRNA transcript using a T2A ribosomal skipping peptide. **(B)** Plasmid construct to achieve dual display of two proteins using a bidirectional *GAL1/GAL10* promoter. In this example, hFc is under the direction of the *GAL1* promoter while SsoFe2 is under the direction of the *GAL10* promoter. **(C)** Visualization of protein co-expression as fusions to yeast cell surface including expression tags. **(D)** Representative flow cytometry analysis of hFc (blue) and SsoFe2 (red) expression by anti-cmyc and anti-V5 labeling, respectively, compared to unlabeled cells (black) using the ribosomal skipping construct. **(E)** Representative flow cytometry analysis of hFc (blue) and SsoFe2 (red) expression by anti-cmyc and anti-V5 labeling, respectively, compared to unlabeled cells (black) using the bidirectional promoter construct.

2.2.2 Yeast cells displaying a known binder protein can be enriched using magnetized yeast cells displaying the binder protein's interaction partner

We conducted enrichment studies on a known binder-target pair to quantitatively assess the likelihood of isolating binders from a combinatorial library using magnetized yeast displaying the target protein. Yeast cells co-expressing hFc and SsoFe2 as cell surface fusions were

magnetized by incubation with iron oxide powder (Figure 2.3A). We investigated the enrichment of cells expressing a previously identified Sso7d mutant (Sso7d-hFc) that binds hFc with modest binding affinity (K_D~ 400 nM)[44]. Cells displaying Sso7d-hFc or the cytosolic domain of TOM22, an irrelevant protein, were mixed in varying ratios (10-100,000x excess of the cells displaying TOM22). The Sso7d-hFc/TOM22 cell mixtures were incubated with magnetized yeast cells expressing hFc as the target. The TOM22 cells act as competitors to the Sso7d-hFc cells for interaction with the target hFc cells. Subsequently, the iron oxide and any complexed cells were separated with a magnet. The number of Sso7d-hFc and TOM22 cells recovered by the magnetic target cells was quantified by plating on selective agar plates and colony counting (Figure 2.4).

The Sso7d-hFc displaying cells were enriched for all population ratios tested (Figure 2.3B), and at least ~20% recovery of these cells was observed (Figure 2.3C). Notably over 100-fold enrichment at ~ 20% recovery was observed when TOM22-displaying cells were present at 10^5-fold excess. Based on the observed ~ 10-100-fold enrichment of binders in one round, one can reasonably assume that 4-5 rounds of screening with magnetized yeast cell targets would suffice, in general, to isolate specific binders from yeast display libraries. Additionally, a recovery of ~ 20% suggests that the library should be oversampled by at least five-fold, when screening using magnetized cell targets.

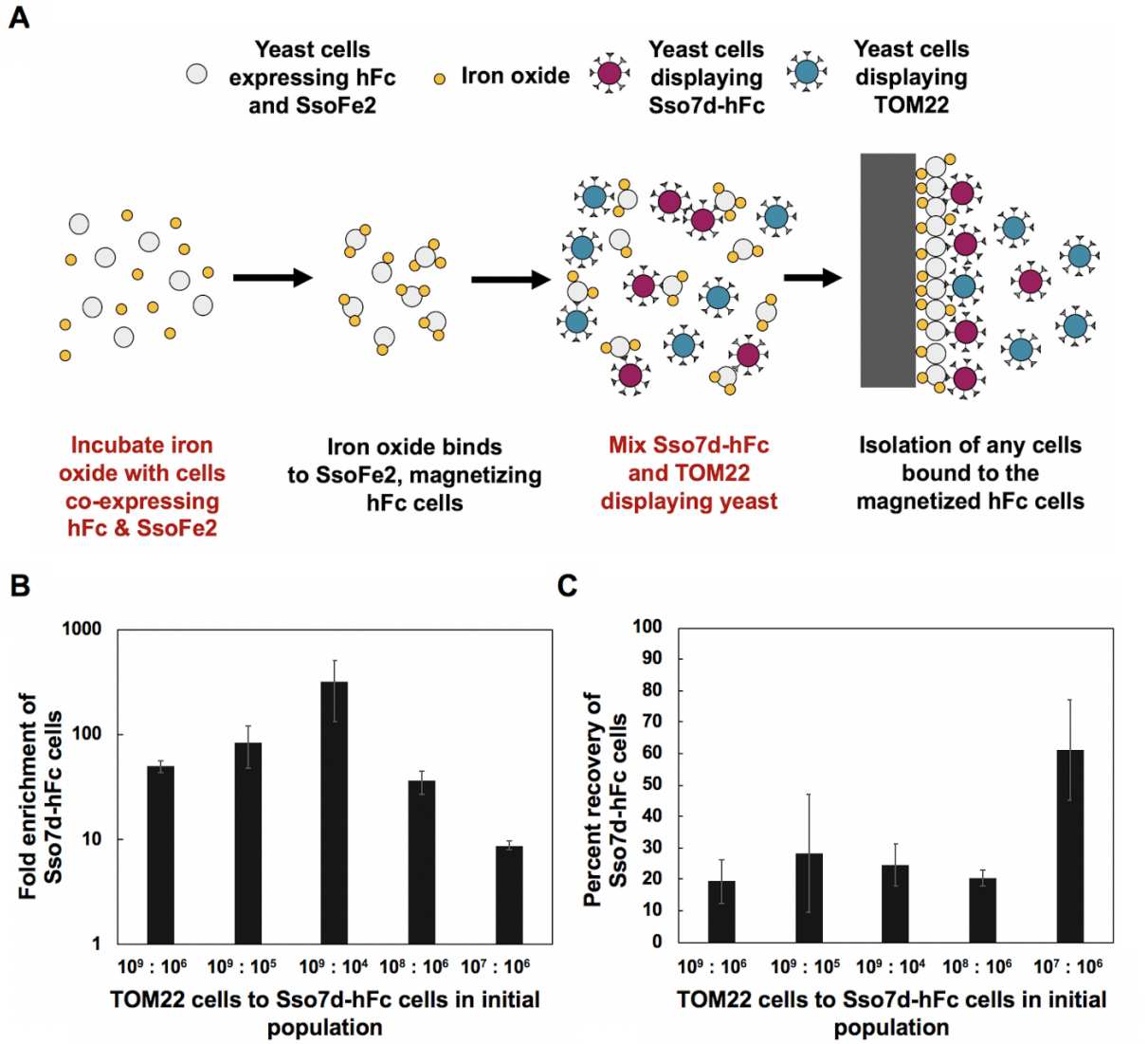


Figure 2.3. Specificity of magnetic yeast displaying target protein for yeast cells displaying binding partner. **(A)** Isolation of Sso7d-hFc displaying yeast using magnetic yeast co-expressing hFc and SsoFe2 in the presence of competitor yeast cells displaying TOM22. The hFc expressing cells are magnetized via iron oxide binding to SsoFe2 that is co-expressed on the cell surface. The number of TOM22 and Sso7d-hFc yeast cells was altered to vary the extent of competition from the TOM22 cells. Cells bound to the magnetic hFc cells were plated and quantitatively assessed to characterize the enrichment and recovery of the Sso7d-hFc population as well as the competitor TOM22 population. **(B)** Enrichment of Sso7d-hFc displaying yeast comparing the percentage of Sso7d-hFc cells in the recovered population to the percentage of Sso7d-hFc cells in the initial population. **(C)** Percentage of Sso7d-hFc displaying yeast in initial population that were recovered by the magnetic target cells. Error bars correspond to the standard error of the mean from three independent replicates.

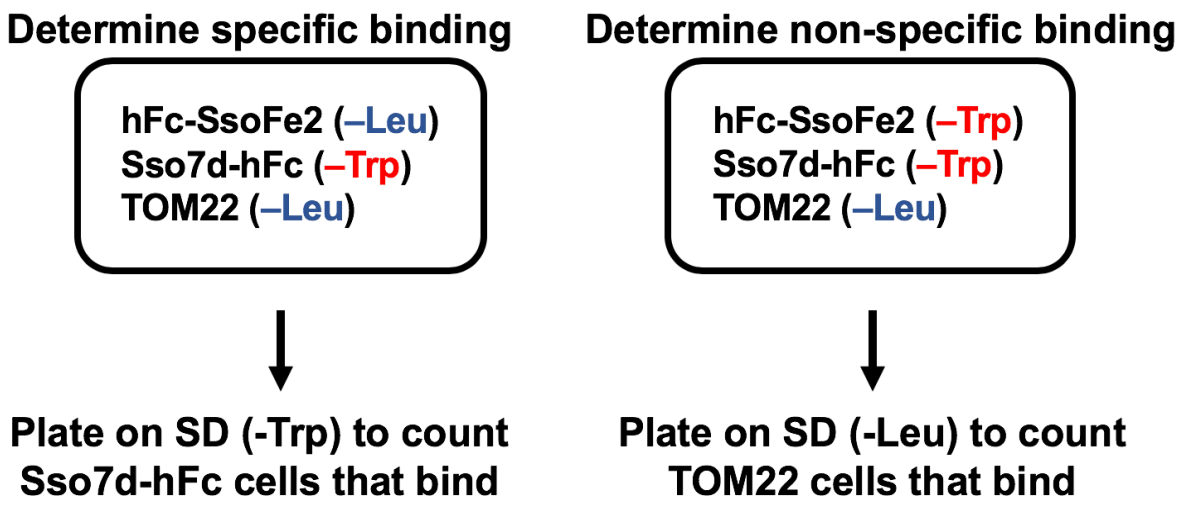


Figure 2.4. Plating method used to evaluate specificity of magnetic target yeast. Here, we considered the specific binding of Sso7d-hFc displaying yeast and the non-specific binding of TOM22 displaying yeast to magnetized yeast cells co-displaying the target protein (hFc) and SsoFe2.

It is instructive to compare our results with those reported by Stern et al. wherein they assessed the recovery of binder yeast upon panning against adherent cells expressing the target[45]. In that study, binder cells ($K_D \sim 2$ nM) were mixed with non-binders at 1000x excess and recovered at ~ 100-fold enrichment and ~ 35% recovery; these values are comparable with our results. Interestingly, however , yeast cells displaying a low affinity binder (K_D>600 nM) required multiple rounds of screening for enrichment[45]. In contrast, our approach could enrich a binder with K_D ~ 400 nM in a single round. Apart from the slightly higher binding affinity of Sso7d-hFc, the higher enrichment observed using our strategy may be attributed to there being a greater number of target cells present. We estimate that our screens contain at least 10x more target cells than when adherent cells on tissue culture plates are used as the target. The simplicity of increasing the number of magnetized target cells is a major advantage of our approach over classical cell panning using adherent cells as the scale up for large library screens is more efficient and even weak affinity binders can be isolated.

It is also important to note that yeast cells can be magnetized without the display of an iron oxide binding protein by simply incubating yeast cells with iron oxide nanopowder due to electrostatic interactions between the iron oxide and yeast cell surface proteins[33]. Therefore, non-specific adsorption of iron oxide to the yeast cell surface is an alternative strategy for obtaining magnetized yeast cell targets. However, use of SsoFe2 enables more robust cell magnetization. Wash steps and extended incubation in buffers containing high concentrations of carrier protein (e.g. bovine serum albumin (BSA)) can result in the loss of magnetization, presumably due to the dissociation of iron oxide particles from the yeast surface. The electrostatic interactions between the yeast cell surface proteins and the iron oxide are too weak to retain the iron oxide on the yeast surface when faced with competition from carrier protein or non-magnetized cells. Loss of magnetization has two deleterious consequences when screening combinatorial libraries – library binder cells associated with the de-magnetized target cells will be lost during the magnetic selection step, and non-specific re-binding of iron oxide to non-binder cells may result in their unwanted isolation. Our studies show that the loss of magnetization is significantly lower when SsoFe2 is present on the yeast cell surface (Figure 2.5). Therefore, yeast surface co-expression of SsoFe2 is the preferred approach for generating magnetized yeast cell targets.

Additionally, inclusion of excess yeast lacking the library's selective nutritional marker during library screening will minimize the isolation and expansion of library cells with no specificity for the target that could be isolated from non-specific adsorption of dissociated iron oxide particles. These yeast cells can display a non-relevant protein or be non-displaying EBY100 cells. By including excess non-relevant yeast cells, it is likely that any dissociated iron oxide will complex with the non-relevant yeast cells that are in excess rather than the library yeast cells. Finally, we noticed a decrease in efficiency of yeast magnetization – both SsoFe2-mediated and

non-specific – as the batch of iron oxide nanopowder aged. A potential explanation is that exposure to air causes changes in the properties of the iron oxide particles. Therefore, it helpful to test the efficiency of magnetization prior to screening libraries.

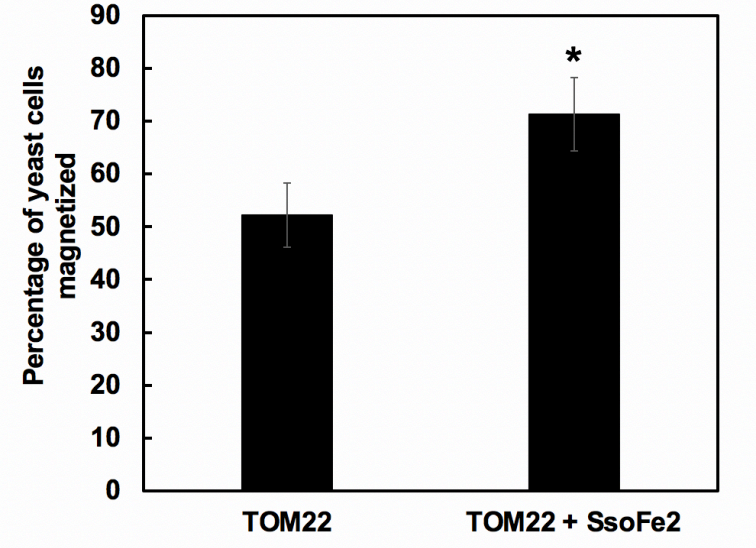


Figure 2.5. Comparison of non-specific and affinity based magnetization of yeast cells. The percentage of yeast cells initially magnetized that remain magnetized over the course of the wash, blocking, and incubation steps that would be associated with a library screen are described. Yeast cells expressing only TOM22 and yeast cells expressing both TOM22 and SsoFe2 were evaluated for their magnetization efficiency. Both cell populations were magnetized after incubation with iron oxide nanoparticles. SsoFe2 affords affinity-based magnetization while the cells not displaying SsoFe2 are magnetized through electrostatic interactions. Error bars represent the standard error of the mean for six repeats for cells displaying TOM22 only and three repeats for cells displaying TOM22, SsoFe2. * represents $p < .01$ for two tailed t-test in comparison to TOM22 only population.

2.2.3 Isolation of novel Sso7d binders to TOM22 and c-Kit using magnetic yeast expressing the target

We sought to investigate the use of magnetic yeast targets for the identification of novel binding proteins specific to the cytosolic domain of TOM22 and the extracellular domain of c-Kit. Towards that end, we screened a combinatorial library based on the Sso7d scaffold against magnetized yeast displaying the targets. Specifically, we used a combinatorial library designed by Cruz-Teran et al. wherein the yeast cells simultaneously express the Sso7d mutants as soluble protein as well as yeast cell surface fusions[46].

Yeast cells, co-expressing the target protein (TOM22 or c-Kit) and SsoFe2, as cell surface fusions, were magnetized by incubation with iron oxide powder as discussed earlier. The magnetized target cells were used to isolate binders from the library during four rounds of magnetic screening. Briefly, in each round of screening, magnetized cells were incubated with the library (Figure 2.1). Library cells that bound to the magnetized target cells were isolated using a magnet and expanded in selective yeast medium. The number of library and target cells was successively reduced for each round of screening. Additionally, excess EBY100 was also included to reduce the likelihood of isolating binders to native yeast surface proteins as well as to decrease the probability of isolating non-binders if iron oxide dissociated from the target cells.

The high avidity interaction between the binder and target displaying cells results in the isolation of even low affinity binders. Nevertheless, it is possible to bias selection towards high affinity binders by reducing the protein display levels on the library cells by DTT treatment, as previously described[47]. Accordingly, the Sso7d library cells were treated with DTT during the fourth round of screening to reduce the number of surface fusions on the Sso7d library cell population.

DNA isolated from the pool of binders obtained after the fourth round of screening was subjected to further random mutagenesis by error-prone PCR and used to construct a second yeast display library. This library was subjected to four rounds of screening as described earlier. Yeast displaying an irrelevant Sso7d mutant (Sso7d-hFc) was included along with EBY100 to increase the screening stringency and to minimize non-specific binding. For each target, plasmid DNA was isolated from ten individual clones recovered from the final screening round. DNA sequencing identified only two unique Sso7d mutants from the TOM22-binding population while only a single unique c-Kit binding Sso7d clone was identified (Figures 2.6A and 2.7A).

2.2.4 Characterization of yeast secreted Sso7d mutants

For each target, the pool of binders isolated after screening the mutagenized libraries was analyzed to confirm binding and specificity. Because a simultaneous secretion and display library was used, Sso7d binder protein could be secreted by the library yeast cells into the supernatant of the yeast display culture, without any additional sub-cloning steps. Individual clones were not characterized at this stage; rather, a mixture comprised of the final population clones was analyzed. Soluble protein, which contains a 6xHis tag, was purified from the supernatant using Ni-NTA agarose and concentrated as previously described[46].

Notably, the soluble protein obtained from the simultaneous secretion and yeast surface display system contains a portion of a F2A peptide at its C-terminus, which can be cleaved by treatment with TEV protease. The soluble protein obtained from the TOM22 library screens bound TOM22 displaying cells when the soluble protein was treated with TEV protease to cleave the F2A peptide. No binding was observed when the F2A peptide sequence was present (Figure 2.6B). The presence of the F2A peptide sequence at the secreted protein's C-terminus may impede the binder population from interacting with TOM22. Alternatively, the F2A peptide sequence may hinder binding of the anti-His antibody used for immunofluorescent detection. These results confirm that the pool of proteins isolated from the combinatorial library screens against yeast cells displaying TOM22 does in fact bind TOM22.

In contrast, Sso7d soluble protein derived from the c-Kit specific screens did not bind c-Kit displaying yeast cells at a protein concentration of 2.5 μM, even upon cleavage of the F2A peptide sequence by TEV protease (Figure 2.7B). These results suggest that the binding affinity of the isolated c-Kit Sso7d binder was significantly lower than that of the TOM22 Sso7d binders.

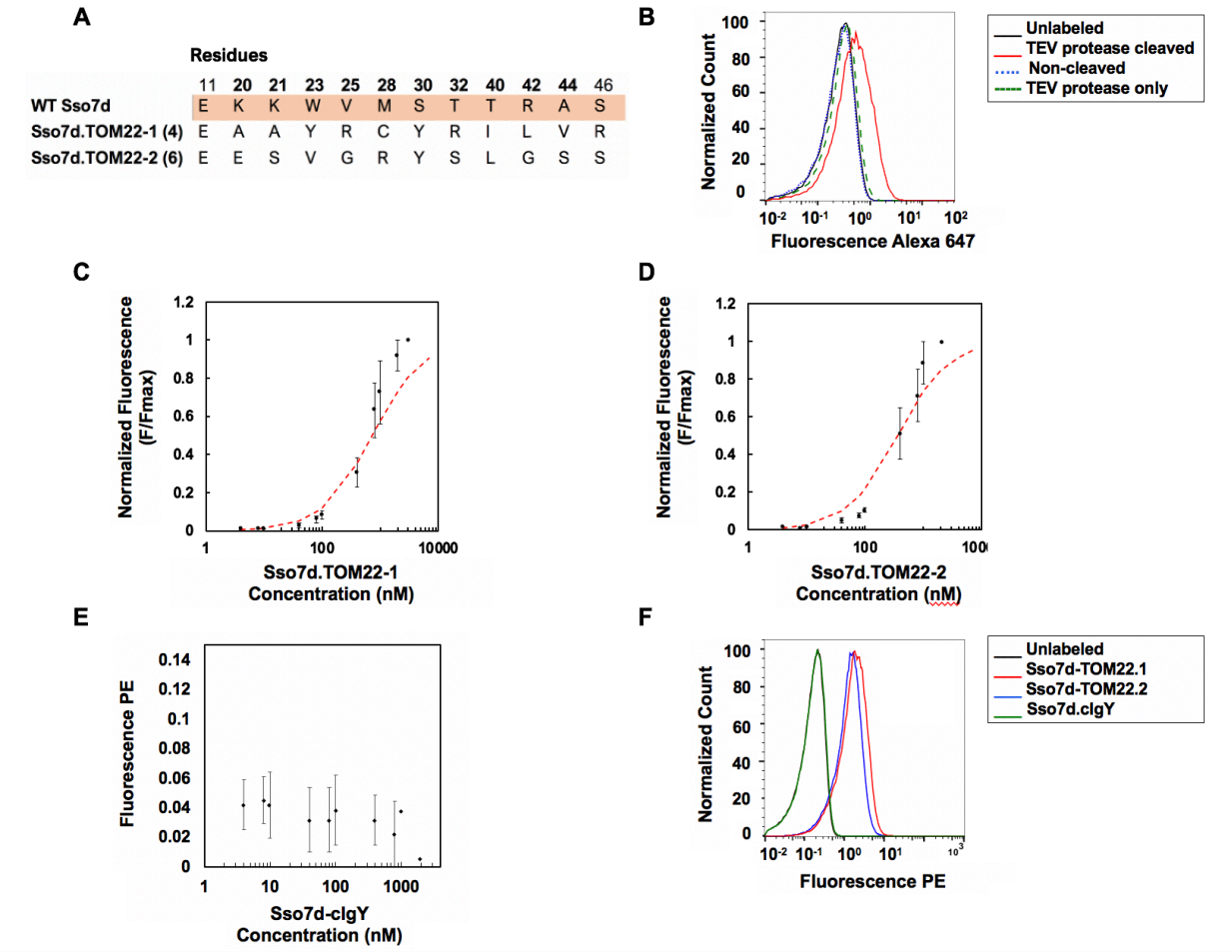


Figure 2.6. Characterization of isolated Sso7d mutants specific for TOM22 obtained from a selection using yeast-displayed targets. **(A)** Sequences of TOM22 binders isolated from a library of Sso7d mutants. Positions mutagenized in the original library are in bold font. The number in parentheses is the number of identical DNA sequences obtained. **(B)** Soluble protein from the pool of TOM22 mutants recovered from the final round of screening was purified from yeast culture supernatant. Flow cytometry analysis was completed for unlabeled cells (black) and yeast cells expressing TOM22 upon labeling with secreted binder protein cleaved with TEV protease (red), non-treated protein (blue), and TEV protease only (green). Soluble protein (2.5 μM) and TEV protease (0.67 μM) binding was detected using an Anti-His Alexa 647 antibody. Subsequently, individual mutant analysis was completed to estimate the apparent binding K_D of Sso7d.TOM22-1 **(C)** and Sso7d.TOM2-2 **(D)** to yeast cells displaying TOM22. A global fit was used to estimate the K_D of Sso7d.TOM22-1 as 732 nM (632-848 nM, 68% confidence interval). The K_D of Sso7d.TOM22-2 was estimated as 362 nM (310-423 nM, 68% confidence interval). Error bars correspond to the standard error of the mean from three independent replicates. **(E)** Background subtracted mean fluorescence values for Sso7d.cIgY binding to yeast cells displaying TOM22. Error bars correspond to the standard error of the mean from three independent replicates. Points with no error bars represent averages from two independent replicates. **(F)** Representative flow cytometry analysis of unlabeled cells (black) as well as yeast cells displaying TOM22 labeled with biotinylated Sso7d.cIgY (green), Sso7d.TOM22-1 (blue), and Sso7d.TOM22-2 (red) at a concentration of 1000 nM, followed by secondary labeling with SA-PE.

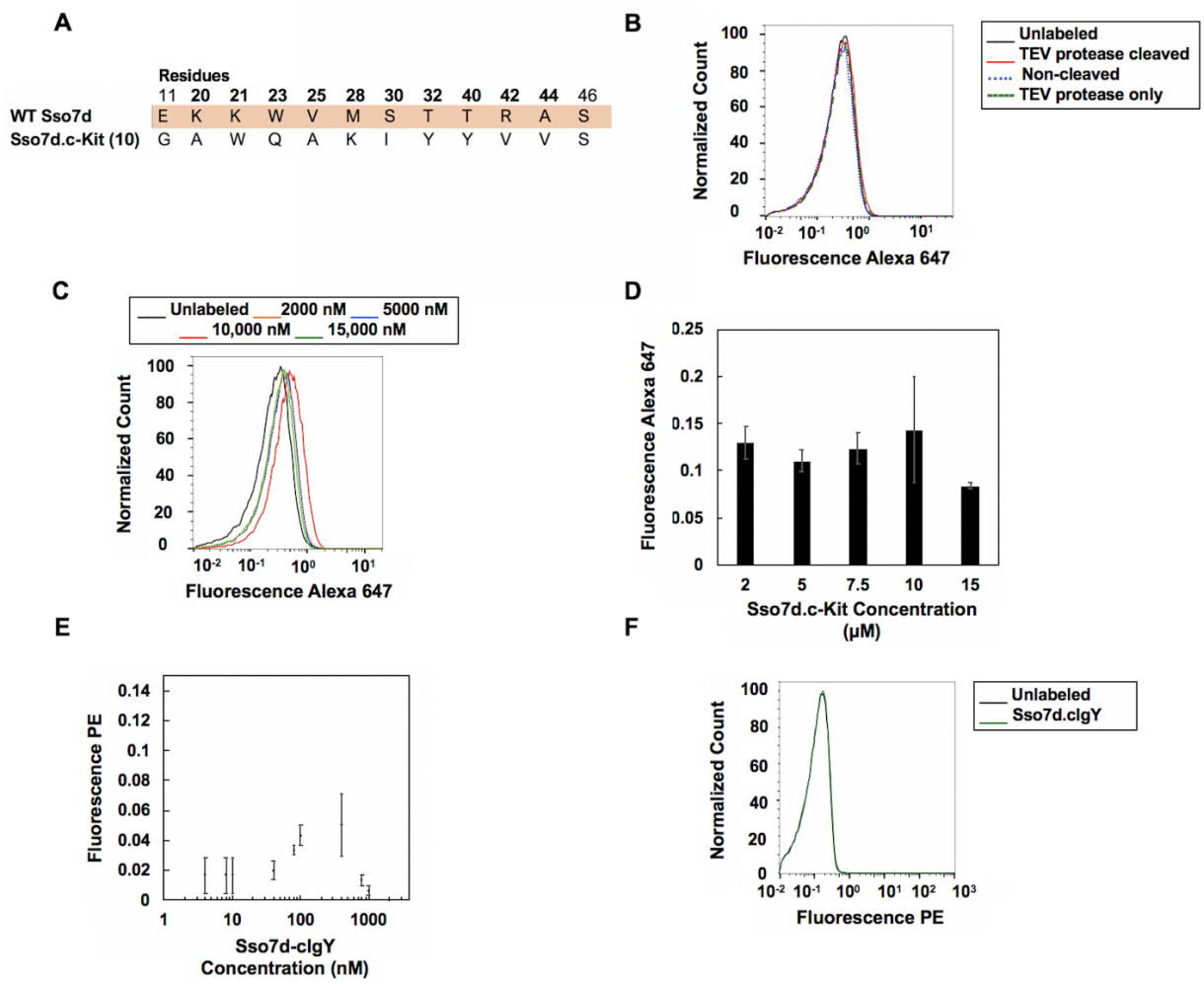


Figure 2.7. Characterization of isolated Sso7d mutants specific for c-Kit obtained from a selection using yeast-displayed targets. **(A)** Sequences of c-Kit binders isolated from a library of Sso7d mutants. Positions mutagenized in the original library are in bold font. The number in parentheses is the number of identical DNA sequences obtained. **(B)** Soluble protein from the c-Kit mutant population recovered from the final round of screening was purified from yeast culture supernatant. Flow cytometry analysis was completed for unlabeled cells (black) and yeast cells expressing c-Kit upon labeling with secreted binder protein cleaved with TEV protease (red), non-treated protein (blue), and TEV protease only (green). Soluble protein (2.5 μM) and TEV protease (0.67 μM) binding was detected using an Anti-His Alexa 647 antibody. **(C)** Representative flow cytometry analysis of unlabeled yeast cells and c-Kit displaying yeast cells labeled with varying concentrations of Sso7d.c-Kit and an Anti-His Alexa 647 antibody. **(D)** Background subtracted mean fluorescence values for Sso7d.c-Kit binding to c-Kit displaying yeast cells at various concentrations. Error bars correspond to standard error of the mean for three independent repeats. **(E)** Background subtracted mean fluorescence values for Sso7d.cIgY binding to yeast cells displaying c-Kit. Error bars correspond to the standard error of the mean from three independent replicates. Points with no error bars represent the average of two independent replicates. **(F)** Representative flow cytometry analysis of unlabeled cells (black) as well as yeast cells displaying c-Kit labeled with biotinylated Sso7d.cIgY (green) at a concentration of 2000 nM, followed by secondary labeling with SA-PE.

2.2.5 Characterization of binding affinity and specificity of recombinantly expressed Sso7d mutants

Each identified Sso7d binder for TOM22 and c-Kit was recombinantly produced in *E. coli*, and the purified protein was used to generate yeast surface titrations to estimate the apparent K_D of binding. Briefly, yeast cells displaying either TOM22 or c-Kit were incubated with varying concentrations of each mutant. The fraction of cell surface fusions bound by the mutant protein was quantified by immunofluorescent detection of biotinylated protein (for TOM22 binders) or a His tag on the recombinant protein (for the c-Kit binder). Note that the recombinantly produced Sso7d.c-Kit was not biotinylated to avoid potential modification of a lysine mutation introduced in the binding interface. The data was fit to a monovalent binding isotherm, and the K_D was estimated as described[31].

The apparent K_D of SSo7d.TOM22-1 and SSo7d.TOM22-2 was estimated as 732 nM and 362 nM respectively (Figures 2.6C-D). Similarly, the apparent K_D of Sso7d.c-Kit.1 was estimated to be greater than 2000 nM (Figures 2.7C-D). The exact K_D for Sso7d.c-Kit could not be determined due to low binding affinity; a complete yeast titration curve could not be generated. Nevertheless, weak, yet detectable levels of Sso7d.c-Kit binding to yeast displayed c-Kit was observed across concentrations ranging from 2,000 nM to 15,000 nM.

While the isolated binders have low to moderate affinity, they showed specific binding for their target. We evaluated the binding of an irrelevant Sso7d mutant, Sso7d.cIgY, to yeast displayed TOM22 and c-Kit to assess the binding specificity of the selected mutants. Sso7d.cIgY binds specifically to chicken IgY and has a reasonably similar composition of residues in its binding interface as the Sso7d TOM22 and c-Kit binding mutants[34]. Yeast surface titrations were conducted, as described earlier, for Sso7d.cIgY binding to TOM22 displaying cells (Figures 2.6E-

F). In contrast to the TOM22-binding Sso7d mutants, Sso7d.cIgY negligibly bound to yeast displayed TOM22 over the range of concentrations tested, and the data did not fit a monovalent binding isotherm. Taken together, these results confirm that Sso7d.TOM22-1 and -2 specifically bind the yeast displayed cytosolic domain of TOM22, and this binding is not a consequence of non-specific binding of the Sso7d scaffold to TOM22. Similarly, at a concentration of 2 μM, Sso7d.c-Kit – but not Sso7d.cIgY – exhibited binding to yeast displayed c-Kit (Figures 2.7C-2.7F). These results show that Sso7d.c-Kit binds to yeast displayed c-Kit specifically, though with weak binding affinity (K_D > 2 μM).

Despite the isolation of low to moderate affinity Sso7d binders, the convergence of the Sso7d libraries to one or two sequences when selected against yeast cells displaying the target domains suggests effectiveness of this screening strategy. The low binding affinity of Sso7d.c-Kit may represent a limitation of the Sso7d scaffold for generating binders to certain targets, as was previously observed for other target proteins[48]. The binding affinity of Sso7d.c-Kit can likely be improved through additional rounds of mutagenesis and screening. However, such affinity maturation was not pursued since the focus of our study was investigating the feasibility of screening yeast display libraries using magnetized yeast cell targets. Instead, we chose to extend our results by screening a combinatorial library based on a different scaffold protein.

2.2.6 Isolation of nanobodies specific to TOM22 & c-Kit

We investigated the screening of a synthetic combinatorial library based on a camelid single-domain antibody fragment (or nanobody) scaffold, as described in McMahon et al.[35], against magnetized yeast displaying a target of interest. In contrast to single chain antibody fragments, nanobodies have a single 15 kDa V_{HH} domain. The nanobody library was screened to obtain binders to TOM22 and c-Kit using magnetized yeast cell targets, as described earlier. The

selection stringency was increased in successive rounds of screening. Note, however, that the nanobody library cannot be treated with DTT to reduce the surface fusion copy number. This is because the nanobody mutants are covalently tethered to the cell wall by a synthetic amino acid chain designed to mimic low complexity yeast cell wall proteins, as opposed to being tethered by disulfide bonds to the cell wall via the interaction between Aga2 and Aga1, like the Sso7d library.

For each target, plasmid DNA was isolated from ten individual clones recovered during the final screening rounds. DNA sequencing identified 8 unique mutants from the TOM22 and c-Kit nanobody binding populations (Figures 2.8A-B). For each target, two unique mutants were recombinantly produced in *E.coli*, and the purified protein was biotinylated. The mutants chosen for characterization appeared twice in the sequenced populations. For each mutant, the apparent K_D of binding to its target was measured using yeast surface titrations in a similar manner as described for the Sso7d mutants. The apparent K_D of NB.TOM22-1 and NB.TOM22-2 (Figure 2.8C) was estimated as 272 and 1934 nM, respectively, while the apparent K_D of NB.c-Kit-1 and NB.c-Kit-2 (Figure 2.8D) was estimated as 93 and 274 nM, respectively. Except for NB.TOM22-2, the binding affinities of the isolated nanobody mutants were higher than the corresponding Sso7d mutants, underscoring differences between these scaffolds. Nevertheless, isolation of NB.TOM22-2 confirms that even binders with low affinity can be isolated when magnetized target cells are used.

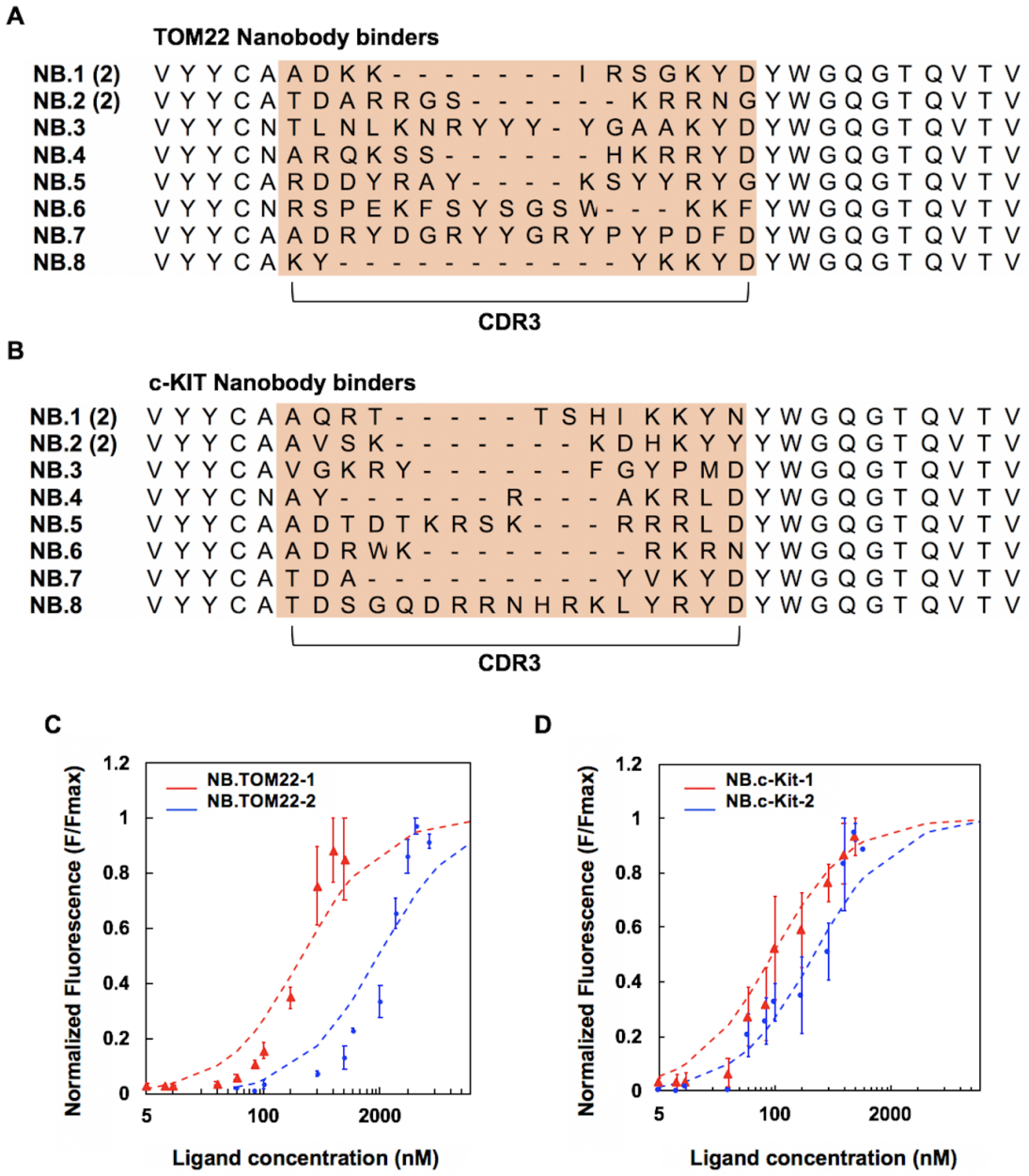


Figure 2.8. Characterization of isolated nanobody mutants specific for TOM22 or c-Kit obtained from a selection using yeast-displayed targets. Sequences of **(A)** TOM22 binders and **(B)** c-Kit binders isolated from a nanobody combinatorial library. CDR3 positions are highlighted in light brown. The number in parentheses is the number of identical DNA sequences obtained. **(C)** Yeast surface titrations to estimate the apparent K_D of NB.TOM22-1 (red) and NB.TOM22-2 (blue) for TOM22. A global fit was used to estimate the K_D for NB.TOM22-1 as 272 nM (238-312 nM, 68% confidence interval). The K_D for NB.TOM22-2 was estimated as 1934 nM (1693-2209 nM, 68% confidence interval). **(D)** Yeast surface titrations to estimate the apparent K_D of NB.c-Kit-1 (red) and NB.c-Kit-2 (blue) for c-Kit. The K_D for NB.c-Kit-1 and NB.c-Kit-2 was estimated as 93 nM (77-113 nM, 68% confidence interval) and 274 nM (230-327 nM, 68% confidence interval), respectively, using a global fit. Error bars correspond to the standard error of the mean from three or four independent replicates.

To further test the functionality of the isolated c-Kit binders, c-Kit-specific nanobodies were immobilized on magnetic beads and evaluated for their ability to deplete yeast cells expressing c-Kit from a heterogeneous population. However, the c-Kit displaying cells were only depleted around 10% after incubation with magnetic beads functionalized with NB.c-Kit-1 or NB.c-Kit-2 (Figures 2.9 and 2.10). It is likely that a higher affinity binder is needed to achieve significant depletion of c-Kit displaying cells. Therefore, we did not pursue further characterization of the c-Kit binding nanobodies in the context of isolating mammalian cells that naturally express c-Kit from a heterogenous mixture.

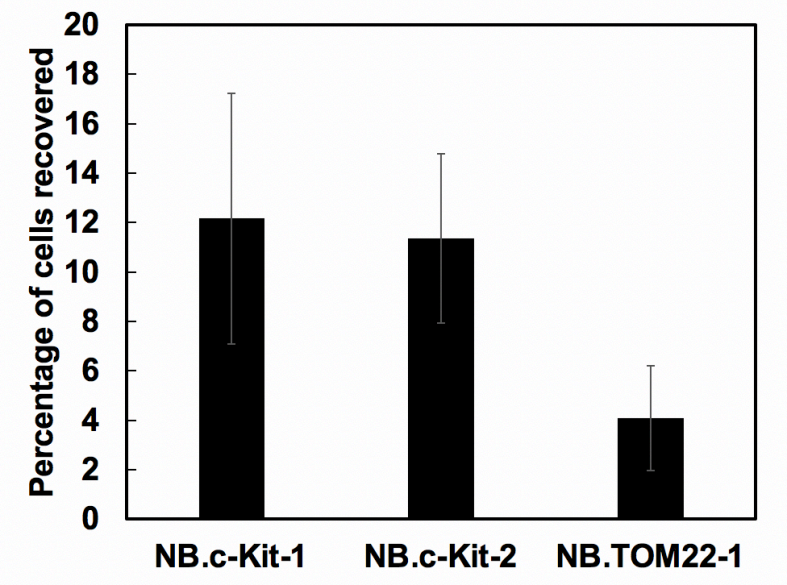


Figure 2.9. Specificity analysis of nanobodies recovered from selections against c-Kit displaying yeast using nanobody coated streptavidin beads. The percentage of c-Kit displaying cells recovered from a mixed population using magnetic streptavidin beads functionalized with binders NB.c-Kit-1, NB.c-Kit-2, and NB.TOM22-1 is described. The initial population contained $1x10^5$ c-Kit displaying yeast and $2x10^6$ non-displaying EBY100 yeast. The reduction in c-Kit displaying cells was quantified by comparing the c-myc expression between the unrecovered cell population and the initial population. These results suggest that higher affinity binders may be required to obtain more significant pull out of specific cell types for heterogenous populations. Error bars represent the standard error of the mean for three repeats.

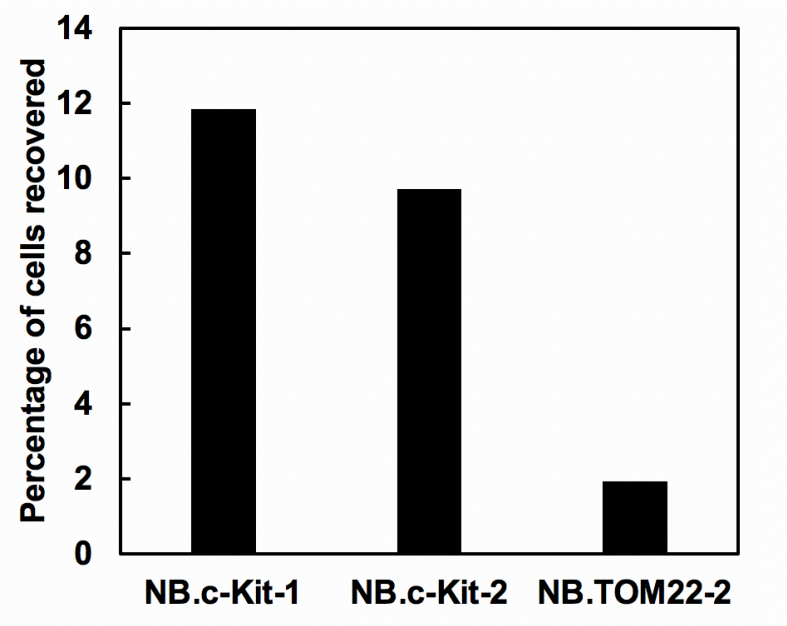


Figure 2.10. Specificity analysis of nanobodies recovered from selections against c-Kit displaying yeast using nanobody coated His beads. The percentage of c-Kit displaying cells recovered from a mixed population using magnetic His beads functionalized with binders NB.c-Kit-1, NB.c-Kit-2, and NB.TOM22-2 is described. The cell populations were only incubated with two sets of functionalized beads in comparison to the previous experiments that involved three bead incubations. Data from a single experiment is shown.

2.2.7 TOM22 binders can enrich mitochondria from cell lysates via recognition of natively expressed TOM22

To characterize the functionality of the TOM22 binders, we assessed the isolated TOM22 Sso7d and nanobody binders for their ability to recover mitochondria from a cell lysate. Each TOM22 binder protein was recombinantly produced, biotinylated, and immobilized onto streptavidin-functionalized 0.15 μm beads. Subsequently, lysates from HEK293-T cells were incubated with the binder-coated beads, and the beads were recovered. Immunoblotting was used to quantify the recovery of mitochondria pulled down by the binder-functionalized beads. Mitochondrial recovery was assessed by immunoblotting with an anti-TOM22 antibody (Figure 2.11A). Pulldown of a non-specific organelle, the endoplasmic reticulum, was assessed using an anti-calnexin antibody (Figure 2.11B); this serves as a proxy for the interaction of the binders with

non-target species. Commercially available micrometer-sized magnetic beads functionalized with an anti-TOM22 antibody are commonly used for the separation of mitochondria from cell lysates[49–52]. We carried out isolation of mitochondria from HEK293-T lysates using the commercially available antibody coated magnetic beads in parallel. Results from these beads served as a benchmark to evaluate the efficiency of the TOM22 binders identified in this study.

We observed that the mitochondrial recovery, as assessed by immunodetection of TOM22, was higher when each mutant was immobilized onto the streptavidin beads in comparison to the recovery using plain (unmodified) beads (Figure 2.11C); NB.TOM22-1 exhibited the greatest recovery among the identified ligands. A direct comparison of mitochondrial recovery between the beads coated with Sso7d or nanobody binders and the commercial product is not possible since the surface density of the anti-TOM22 antibody is unknown. However, we used the ratio of the TOM22 band intensity to that of the calnexin band intensity as a quantitative metric for specificity, to compare the commercial beads with the binder-coated streptavidin beads. Based on this metric, our mutants have specificity for mitochondria over the endoplasmic reticulum comparable to that of the commercial antibody beads (Figure 2.11D). The Sso7d and nanobody mutants were able to recover mitochondria from the heterogenous cell lysate with similar efficiency as the antibody coated beads. These results suggest the Sso7d and nanobody binders specific to TOM22, isolated from selections against a yeast displayed target, are functional in the context of the native mitochondrial protein.

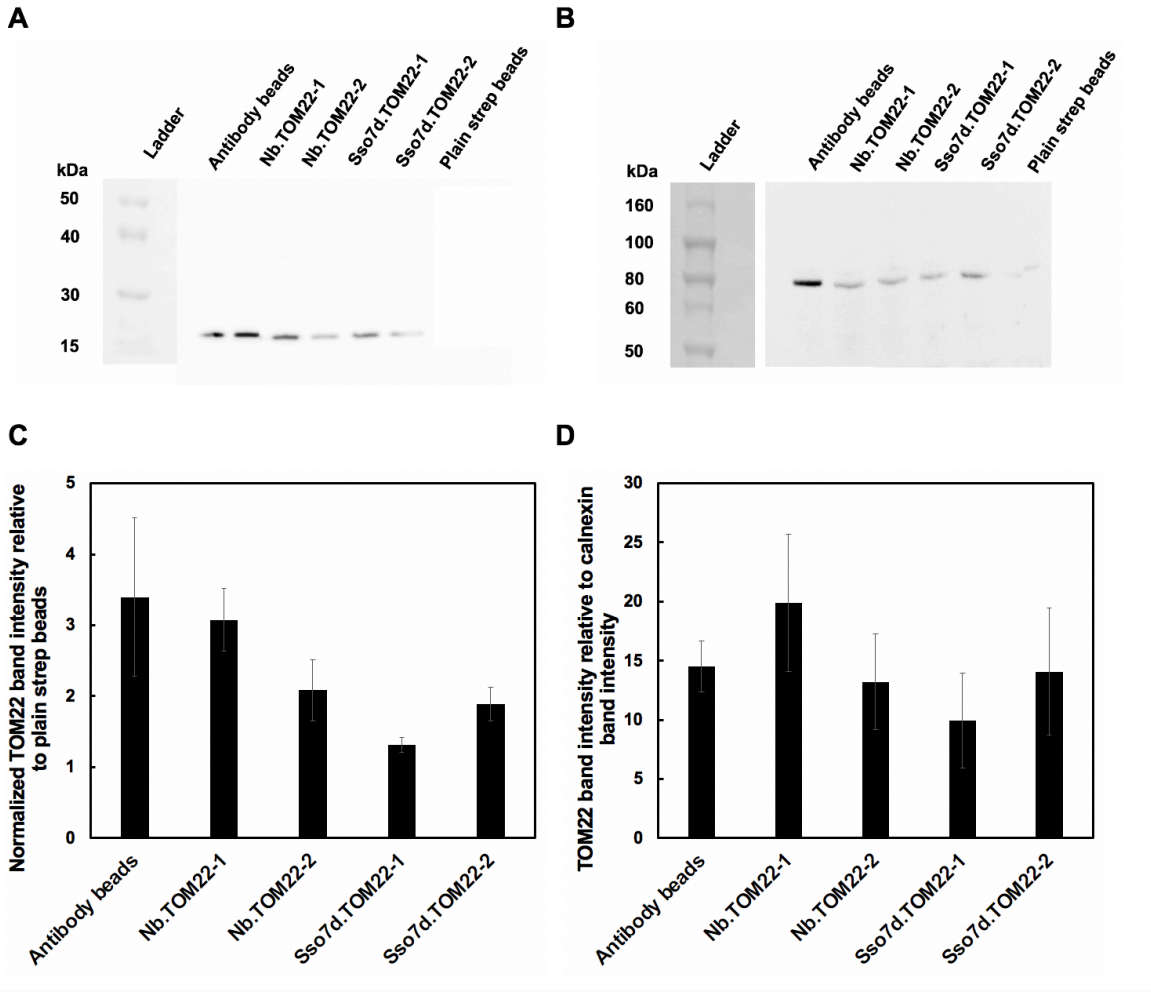


Figure 2.11. Enrichment of mitochondria from cell lysate using nanobody and Sso7d binders specific to TOM22. The binder proteins were biotinylated and functionalized onto streptavidin beads (0.15 µm). The mitochondrial recovery of the TOM22 nanobody and Sso7d binders was compared to the recovery observed using non-functionalized streptavidin beads as well as commercially available micro-sized magnetic beads functionalized with an anti-TOM22 antibody. **(A)** Mitochondrial recovery was quantified using anti-TOM22-HRP immunoblotting. **(B)** Endoplasmic reticulum recovery was quantified using anti-Calnexin-HRP immunoblotting. An image of the ladder was obtained in a separate image than the blot. Blot images were inverted for display. The ladder and experimental condition sections of the blot were spliced together after alignment. No other manipulations of the image occurred. These blots are representative of three replicates. Mitochondria and endoplasmic reticulum recovery was quantified by band intensity analysis. **(C)** The mitochondrial recovery by beads functionalized with a nanobody, Sso7d, or antibody specific for TOM22 was compared to the mitochondrial recovery of non-functionalized streptavidin beads to evaluate the enrichment of each binding protein against background binding. **(D)** The specificity of the binder proteins was evaluated by comparing the mitochondrial recovery to the endoplasmic reticulum recovery for each bead type considered. Error bars represent the standard error of the mean for at least three independent replicates.

Increases in yield and specificity may be possible if the pull-down protocol is further optimized. However, we did not optimize the binding and wash conditions. Furthermore, our pull-

down protocol with the streptavidin coated beads used centrifugation to isolate the beads as opposed to magnetic separation, which was used to recover the commercial beads. It is also noteworthy that NB.TOM22-2 showed similar or greater mitochondria recovery in comparison to the Sso7d mutants, despite its weak affinity ~2000 nM. A potential explanation is that protein immobilization may differentially affect the binding activity of the various mutants. Alternatively, these differences may be attributed to increased non-specific binding of proteins present in cell lysates to the Sso7d mutants. While the specificity of the ligands for mitochondrial capture over endoplasmic reticulum was explored, there are a wide variety of other proteins and organelles found in the cell lysate that were not considered when evaluating specificity that could be outcompeting the mitochondria for binding to the functionalized beads. Therefore, inclusion of target-depleted cell homogenates during later rounds of library screening may be beneficial for isolation of binders with greater specificity for TOM22. Nevertheless, our results confirm that our isolated Sso7d and nanobody mutants bind natively expressed TOM22 on the surface of a mitochondria.

2.3 Conclusions

We have shown that the isolation of binding proteins from yeast display libraries using yeast displayed targets is feasible. A critical aspect of our strategy is affinity-based magnetization of target-displaying cells, mediated by interactions between iron oxide nanoparticles and a co-expressed iron oxide binding protein; this facilitates facile, magnetic separation of library cells that bind the target. Binding proteins discovered when selecting against yeast displayed targets were functional in the context of a natively expressed membrane protein, as seen with the TOM22 binders depleting cell lysates of mitochondria. We expect that this strategy will enable efficient isolation of binders to membrane protein targets.

2.4 Materials and Methods

2.4.1 Plasmids and yeast culture

The *Sacchromyces cerevisiae* strain EBY100 was used in conjunction with the pCTCON vector containing a *TRP* selectable marker, or a pCT302 vector variant with a *LEU* selectable marker[53]. Plasmid DNA was transformed into chemically competent EBY100 using the Frozen-EZ yeast transformation Kit II (Zymo Research).

For cells with the pCTCON vector, Trp-deficient SDCAA and SGCAA medium was used for culturing cells and for inducing cell surface protein expression, respectively, as previously described[31]. Similarly, Leu-deficient SDSCAA (-Leu) and SGSCAA (-Leu) media was used for cells with the pCT302-based vector. Leu-deficient media has similar composition to SDCAA and SGCAA media except it contains synthetic dropout mix (1.62 g/L) lacking leucine instead of casamino acids. Yeast cells were cultured in SDCAA or SDSCAA medium, as appropriate, at 30°C with shaking at 250 RPM. Protein expression was induced by transferring the yeast cells into SGCAA or SGSCAA medium at an OD_{600} of 1 and cultured overnight at 20 °C with shaking at 250 rpm. EBY100 without plasmid was grown in YPD medium (10.0 g/L yeast extract, 20.0 g/L peptone, and 20.0 g/L dextrose) at 30 °C with shaking at 250 rpm.

The nanobody library was grown in TRP deficient NB.SDCAA and NB.SGCAA medium (3.8 g -TRP drop-out media supplement (US Biological), 6.7 g Yeast nitrogen base (HiMedia), 20 g dextrose for growth or galactose for protein induction, pH 6). The nanobody library cells were cultured in NB.SDCAA medium at 30 °C with shaking at 250 RPM. Protein expression was induced by transferring the yeast cells into NB.SGCAA medium at an OD_{600} of 1 and cultured overnight at 20 °C with shaking at 250 rpm.

2.4.2 Construction of co-expression plasmids

pCTCON-SsoFe2-T2A-TOM22 was constructed by amplifying gene block 1 with Pf1 and Pr1 and introduced to the pCTCON vector between the EcoRI and XhoI sites. The DNA corresponding to the Fc portion of human IgG (residues 100-330 of immunoglobulin heavy constant gamma 1) was amplified from gene block 2 using primers Pf2 and Pr2 and inserted into pCTCON-T2A-SsoFe2-TOM22 between the NheI and BamHI sites creating pCTCON-T2A-SsoFe2-hFc.

To construct pCT302-SSoFe2-T2A-TOM22, the pCT302 vector was cut between NheI and XhoI. Gene block 1 was amplified by PCR using primers Pf1 and Pr1 and introduced into pCT302 using homologous recombination. Briefly, 200 ng of cut plasmid and 250 ng of insert were transformed into EBY100 cells prepared with the Frozen-EZ yeast transformation kit II. The plasmid and insert were concentrated using ethanol precipitation such that the volume of DNA used in the transformation step was less than 5 μL in total. The extracellular domain of c-Kit (residues 2-83) was amplified from gene block 3 using primers Pf3 and Pr3. pCT302-SsoFe2-T2A-c-Kit was constructed by inserting the amplified c-Kit DNA between the NheI and BamHI sites in pCT302-SsoFe2-T2A-TOM22. pcT302-SsoFe2-T2A-hFc was constructed in a similar manner by amplifying the hFc DNA from gene block 2 with primers Pf2 and Pr2 and inserting between the NheI and BamHI sites of pCT302-SsoFe2-T2A-TOM22.

To generate the vector where simultaneous display is mediated by the bidirectional *GAL1/GAL10* promoter, gene block 4 was amplified using primers Pf4 and Pr4 and introduced to the pCTCON vector between the EcoRI and XhoI sites to eliminate an AgeI site from the prepro sequence used in previous constructs to generate pCTCON-modified-prepro. Gene block 5 was amplified using primers Pf5 and Pr5 and inserted into the pCTCON-modified-prepro plasmid

between the KpnI and AgeI sites creating pCTCON-SsoFe2-Bi-TOM22. DNA corresponding to hFc was amplified using primers Pf2 and Pr2 and inserted into pCTCON-SsoFe2-Bi via the NheI and BamHI sites to generate pCTCON-SsoFe2-Bi-hFc.

Sso7d-hFc was inserted between the NheI and BamHI sites of pCTCON and pCT302 to create pCTCON-Sso7d-hFc and pCT302-Sso7d-hFc[44]. The Sso7d-hFc DNA was amplified using primers Pf6 and Pr6 from previously constructed plasmids[44]. The TOM22 mitochondrial domain was cloned into pCT302 by amplifying gene block 1 using primers Pf7 and Pr7 and inserting the amplified, digested DNA between the NheI and the BamHI sites to create pCT302-TOM22. The extracellular domain of c-Kit was similarly cloned into pCT302 using gene block 3 and primers Pf3 and Pr3 to generate pCT302-c-Kit.

All gene fragments were purchased from Integrated DNA technologies. Primer oligonucleotides were purchased from IDT or Eton Biosciences. Gene fragment and primer sequences can be found in Tables 2.1 and 2.2. All PCR reactions were performed in 50 μL reactions with high-fidelity Phusion polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. All restriction enzymes were purchased from New England Biolabs. Restriction digestions were performed in 50 μL with a 5 times excess of each restriction enzyme for 4 hours at 37°C. Digested plasmid backbones were treated with Antarctic Phosphatase purchased from New England Biolabs for 1 hour at 37°C. Digested products and PCR amplicons were purified using the 9K Series Gel and PCR extraction kit from Biobasic. Ligations of the digested plasmid backbones and PCR products occurred overnight at 16°C using T4 DNA ligase (Promega) prior to transformation into chemically competent Novablue *E. coli* cells. The Novablue cells were made chemically competent using the Mix & Go! E. coli transformation kit and buffer

(Zymo Research). Overnight *E. coli* cultures were harvested for their plasmid using the GeneJET plasmid miniprep kit (Thermo Fisher Scientific).

Table 2.1. List of gene fragments described in Chapter 2.

Gene Block 1	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGACAGGAACTGACAACTATA TGCGAGCAAATCCCCTCACCAACTTTAGAATCGACGCCGTACTCTT TGTCAACGACTACTATTTTGGCCAACGGGAAGGCAATGCAAGGAG TTTTTGAATATTACAAATCAGTAACGTTTGTCAGTAATTGCGGTTC TCACCCCTCAACAACTAGCAAAGGCAGCCCCATAAACACACAGTA TGTTTTTTATCCGTACGACGTTCCAGACTACGCTGGTGGTGGTGGT TCCGGTGGTGGTGGTTCTGGTGGTGGTGGTTCAGCTAGCGCTGCCG CCGTCGCTGCTGCCGGTGCAGGGGAACCCCAGTCCCCGGACGAAT TGCTCCCGAAAGGCGACGCGGAGAAGCCTGAGGAGGAGCTGGAG GAGGACGACGATGAGGAGCTAGATGAGACCCTGTCGGAGAGACT ATGGGGCCTGACGGAGATGTTTCCGGAGAGGGTCCGGTCCGCGGC CGGAGCCACTTTTGATCTTTCCCTCTTTGTGGCTCAGAAAATGTAC AGGTTTTCCAGGGCAGGATCCGAACAAAAGCTTATCTCCGAAGAA GACTTGGAGGGAAGGGGATCTTTGCTAACGTGCGGGGATGTCGAA GAAAACCCTGGCCCTATGAAGGTGCTGATTGTTCTATTGGCTATAT TTGCTGCCTTGCCCCTGGCCCTGGCACAGCCAGTCATAAGTACAAC AGTGGGCAGTGCTGCGGAAGGTTCACTAGATAAAAGACAGGAACT TACTACAATTTGTGAACAAATACCTAGTCCCACGTTGGAGTCAACA CCGTACTCCCTGTCTACTACAACAATACTTGCGAACGGTAAGGCCA TGCAAGGAGTCTTTGAATACTACAAGTCAGTTACCTTCGTATCTAA CTGCGGGAGTCACCCGAGCACCACATCCAAGGGGAGTCCTATAAA CACTCAGTATGTATTCGACTATAAGGATGACGATGACAAGGGCGG CGGAGGCTCCGGTGGAGGCGGAAGCGGCGGTGGAGGTAGTCCTA GGGCGACCGTGAAATTTAAATATAAAGGCGAAGAAAAACAGGTG GATATTAGCAAAATTAGACGGGTGCGTCGCAAGGGCAAATGCATT AGGTTTTACTATGATCTGGGCGGCGGCAAATATGGCAGGGGCATA GTGAGCGAAAAAGATGCGCCGAAAGAACTGCTGCAGATGCTGGA AAAACAGAAAAAACATATGGGTAAACCCATACCCAATCCCCTGCT GGGTCTTGATAGTACGTAATAG

Table 2.1 (continued).

Gene Block 2	CCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCA CCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAAC CCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCG TGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACT GGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGC GGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCA CCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCA AGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCT CCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGC CCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCT GCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGG AGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCC GTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCG TGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG TGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCT CCCTGTCTCCGGGTAAA

Table 2.1 (continued).

Gene Block 3	CAACCATCTGTGAGTCCAGGGGAACCGTCTCCACCATCCATCCATC CAGGAAAATCAGACTTAATAGTCCGCGTGGGCGACGAGATTAGGC TGTTATGCACTGATCCGGGCTTTGTCAAATGGACTTTTGAGATCCT GGATGAAACGAATGAGAATAAGCAGAATGAATGGATCACGGAAA AGGCAGAAGCCACCAACACCGGCAAATACACGTGCACCAACAAA CACGGCTTAAGCAATTCCATTTATGTGTTTGTTAGAGATCCTGCCA AGCTTTTCCTTGTTGACCGCTCCTTGTATGGGAAAGAAGACAACGA CACGCTGGTCCGCTGTCCTCTCACAGACCCAGAAGTGACCAATTAT TCCCTCAAGGGGTGCCAGGGGAAGCCTCTTCCCAAGGACTTGAGG TTTATTCCTGACCCCAAGGCGGGCATCATGATCAAAAGTGTGAAA CGCGCCTACCATCGGCTCTGTCTGCATTGTTCTGTGGACCAGGAGG GCAAGTCAGTGCTGTCGGAAAAATTCATCCTGAAAGTGAGGCCAG CCTTCAAAGCTGTGCCTGTTGTGTCTGTGTCCAAAGCAAGCTATCT TCTTAGGGAAGGGGAAGAATTCACAGTGACGTGCACAATAAAAGA TGTGTCTAGTTCTGTGTACTCAACGTGGAAAAGAGAAAACAGTCA GACTAAACTACAGGAGAAATATAATAGCTGGCATCACGGTGACTT CAATTATGAACGTCAGGCAACGTTGACTATCAGTTCAGCGAGAGT TAATGATTCTGGAGTGTTCATGTGTTATGCCAATAATACTTTTGGA TCAGCAAATGTCACAACAACCTTGGAAGTAGTAGATAAAGGATTC ATTAATATCTTCCCCATGATAAACACTACAGTATTTGTAAACGATG GAGAAAATGTAGATTTGATTGTTGAATATGAAGCATTCCCCAAAC CTGAACACCAGCAGTGGATCTATATGAACAGAACCTTCACTGATA AATGGGAAGATTATCCCAAGTCTGAGAATGAAAGTAATATCAGAT ACGTAAGTGAACTTCATCTAACGAGATTAAAAGGCACCGAAGGAG GCACTTACACATTCCTAGTGTCCAATTCTGACGTCAATGCTGCCAT AGCATTTAATGTTTATGTGAATACAAAACCAGAAATCCTGACTTAC GACAGGCTCGTGAATGGCATGCTCCAATGTGTGGCAGCAGGATTC CCAGAGCCCACAATAGATTGGTATTTTGTCCAGGAACTGAGCAG AGATGCTCTGCTTCTGTACTGCCAGTGGATGTGCAGACACTAAACT CATCTGGGCCACCGTTTGGAAAGCTAGTGGTTCAGAGTTCTATAGA TTCTAGTGCATTCAAGCACAATGGCACGGTTGAATGTAAGGCTTAC AACGATGTGGGCAAGACTTCTGCCTATTTTAACTTTGCATTTAAAG GTAACAACAAAGAGCAAATCCATCCCCACACCCTGTTCACTCCT

Table 2.1 (continued).

Gene Block 4	ATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTTGCCATT GGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCCGCTGCA GAAGGCTCTTTGGACAAGAGACAGGAACTGACAACTATATGCGAG CAAATCCCCTCACCAACTTTAGAATCGACGCCGTACTCTTTGTCAA CGACTACTATTTTGGCCAACGGGAAGGCAATGCAAGGAGTTTTTG AATATTACAAATCAGTAACGTTTGTCAGTAATTGCGGTTCTCACCC CTCAACAACTAGCAAAGGCAGCCCCATAAACACACAGTATGTTTT TCCCGGGTATCCGTACGACGTTCCAGACTACGCTGGTGGTGGTGGT TCCGGTGGTGGTGGTTCTGGTGGTGGTGGTTCAGCTAGCGCTGCCG CCGTCGCTGCTGCCGGTGCAGGGGAACCCCAGTCCCCGGACGAAT TGCTCCCGAAAGGCGACGCGGAGAAGCCTGAGGAGGAGCTGGAG GAGGACGACGATGAGGAGCTAGATGAGACCCTGTCGGAGAGACT ATGGGGCCTGACGGAGATGTTTCCGGAGAGGGTCCGGTCCGCGGC CGGAGCCACTTTTGATCTTTCCCTCTTTGTGGCTCAGAAAATGTAC AGGTTTTCCAGGGCAGGATCCGAACAAAAGCTTATCTCCGAAGAA GACTTGTAATGACTCGAG
Gene Block 5	CTATTACGTACTATCAAGACCCAGCAGGGGATTGGGTATGGGTTT ACCCCATGGTTTTTTCTGTTTTTCCAGCATCTGCAGCAGTTCTTTCG GCGCATCTTTTTCGCTCACTATGCCCCTGCCATATTTGCCGCCGCCC AGATCATAGTAAAACCTAATGCATTTGCCCTTGCGACGCACCCGTC TAATTTTGCTAATATCCACCTGTTTTTCTTCGCCTTTATATTTAAAT TTCACGGTCGCCCTAGGACTACCTCCACCGCCGCTTCCGCCTCCAC CGGAGCCTCCGCCGCCCTTGTCATCGTCATCCTTATAGTCGAATAC ATACTGAGTGTTTATAGGACTCCCCTTGGATGTGGTGCTCGGGTGA CTCCCGCAGTTAGATACGAAGGTAACTGACTTGTAGTATTCAAAG ACTCCTTGCATGGCCTTACCGTTCGCAAGTATTGTTGTAGTAGACA GGGAGTACGGTGTTGACTCCAACGTGGGACTAGGTATTTGTTCAC AAATTGTAGTAAGTTCCTGTCTTTTATCTAGTGAACCTTCCGCAGC ACTGCCCACTGTTGTACTTATGACTGGCTGTGCCAGGGCCAGGGGC AAGGCAGCAAATATAGCCAATAGAACAATCAGCACCTTCATAATT CCTTGGAATTTTCAAAAATTCTTACTTTTTTTTTTGGATGGACGCAA AGAAGTTTAATAATCATATTACATGGCATTACCACCATATACATAT CCATATACATATCCATATCTAATCTTACTTATATGTTGTGGAAATG TAAAGAGCCCCATTATCTTAGCCTAAAAAAACCTTCTCTTTGGAAC TTTCAGTAATACGCTTAACTGCTCATTGCTATATTGAAGTACGGAT TAGAAGCCGCCGAGCGGGTGACAGCCCTCCGAAGGAAGACTCTCC TCCGTGCGTCCTCGTCTTC

Table 2.2. List of oligonucleotide primers described in Chapter 2.

Primer	Sequence
Pf1	TCTTATTCAAATGTAATAAAAGATCGAATTCATGAAGGTTTTGATTGTC
Pr1	TACATCTACACTGTTGTTATCAGATCTCGAGCTATTACGTACTATCAAGA CCC
Pf2	GTTGACGCTAGCCCCAAATCTTGTGACAAAACTCAC
Pr2	GCACTTGGATCCTTTACCCGGAGACAGGGAGA
Pf3	GTTGACGCTAGCCAACCATCTGTGAGTCCAGGG
Pr3	GCACTTGGATCCAGGAGTGAACAGGGTGTGGG
Pf4	GCTACGGAATTCATGAAGGTTTTGATTGTCTTGTTGGC
Pr4	GCTACGCTCGAGTCATTACAAGTCTTCTTCGGAGATAA
Pf5	GCTATCGGTACCCTATTACGTACTATCAAGACCCAG
Pr5	GTACTGACCGGTGAAGACGAGGACGCACGGA
Pf6	GAGTCAGCTAGCATGGCGACCGTGAAATTTAAATAT
Pr6	GAGTCAGGATCCTTTTTTCTGTTTTTCCAGCATCTG
Pf7	GTTCTCGCTAGCGCTGCCGCCGTCGCTG
Pr7	GCACTTGGATCCTGCCCTGGAAAACCTGTACATTT
Pf8	CAGAAGGCTCTTTGGACAAGAGACACCACCACCACCACCACGCTAGCAT GGCGACCGTGAAATTTAAATAT
Pr8	GGAGATAAGCTTTTGTTCCCCTTGGAAGTACAGGTTCTCGGATCCTTTTT TCTGTTTTTCCAGCATCTG
Pf9	GATGCACATATGATGGCGACCGTGAAATTT
Pr9	CCGCCGCTCGAGTTTTTTCTGTTTTTCCAGCATC
Pf10	GTCTCGCATATGGCACAAGTTCAGCTTGTAGAGT
Pr10	GTCACTCTCGAGCGATGATACAGTTACTTGGGTAC

2.4.3 Expression analysis of dual display constructs

The expression level of the co-expressed proteins for both dual display constructs was estimated using flow cytometry analysis using pCTCON-T2A-SsoFe2-hFc and pCTCON-SsoFe2-Bi-hFc. Briefly, $5x10^6$ cells were labeled with a 1:100 dilution of chicken anti-c-myc antibody or mouse anti-V5 antibody (Thermo Fisher Scientific) for 15 minutes at room temperature. Subsequently, cells were washed and secondary labeling was carried out using a 1:250 dilution of goat-anti-chicken 633 or donkey anti-mouse 633 (Immunoreagents, Raleigh, NC) for 10 minutes on ice. All labeling was conducted in 50 μL of 0.1% PBS BSA (PBSA). Cells were analyzed using a Miltenyi Biotec MACsQuant VYB cytometer.

2.4.4 Comparison of non-specific and affinity based magnetization

$1x10^7$ pCT302-TOM22 yeast were magnetized in 2 mL of 50 mM sodium acetate, 50 mM NaCl pH 5 buffer by incubating 100 µL of a 4 mg/mL iron oxide solution for 30 minutes at room temperature. $1x10^7$ pCT302-SsoFe2-T2A-TOM22 yeast cells were similarly magnetized, but using 50 mM Tris-HCl, 300 mM NaCl pH 7.4 (TN buffer). Two separate samples were magnetized per cell type. Cells complexed with iron oxide were removed from the unbound cells using a magnet and were washed with the appropriate magnetization buffer. After a single wash, one sample was serially diluted and plated onto SD(-Leu) plates to quantify the initial number of cells magnetized. The other sample was carried forward and washed two further times to simulate the wash and incubation steps that would occur during a typical library selection. After, this sample was incubated with 1% TNBSA pH 7.4. (TN buffer + 1% BSA) for one hour at room temperature. Next, this sample was washed 2X with 0.1% TNBSA pH 7.4 (TN buffer + 0.1% BSA) and incubated for 2 hours at room temperature in this buffer. This sample was finally washed three more times with 0.1% TNBSA, resuspended in 1 mL 0.1% TNBSA, serially diluted, and plated onto SD (-Leu) plates to quantify the cells that remained bound to the iron oxide.

2.4.5 Magnetic pull downs of a known binder-target pair

$5x10^7$ pCTCON-T2A-SsoFe2-hFc yeast cells and $5x10^7$ cells pCT302-T2A-SsoFe2-hFc yeast cells were spun down in separate tubes and resuspended in 2 mL of 1% BSA TN buffer (50 mM Tris-HCl, 300 mM NaCl pH 7.4; TNBSA). To magnetize the cells, 100 µL of a 4 mg/mL iron oxide (II, III) (Thermo Fisher Scientific) solution in water was added to the cells and incubated for 30 minutes. The non-magnetized cells were removed using a magnet. OD_{600} of the solution prior to the addition of iron oxide (OD_i) and after the removal of yeast bound to iron oxide (OD_f) was

measured using 100 μL of the sample. The number of cells bound to the iron oxide was calculated as (OD_i-OD_f) and used for normalization between replicates, as described later. The cells complexed with the iron oxide were washed three times with 1% TNBSA and blocked with 2 mL of 1% TNBSA for 1 hour. After blocking, the cells complexed with the iron oxide were washed two times with 0.1% TNBSA.

Varying amounts of pCT302-TOM22 yeast cells and pCTCON-Sso7d-hFc yeast cells were incubated with the magnetized target cells (10^7:10^6, 10^8:10^6, 10^9:10^6, 10^9:10^5, and 10^9:10^4 pcT302-TOM22 cells: pCTCON-Sso7d-hFc cells) for two hours at room temperature in 2 mL of 0.1% TNBSA with rotation. Two separate cell aliquots were generated for each ratio. One aliquot was incubated with magnetic target cells that contained the pCTCON-T2A-SsoFE2-hFc plasmid (*TRP* selectable marker) while the other aliquot was incubated with magnet target cells that contained the pCT302-T2A-SsoFe2-hFc plasmid (*LEU* selectable marker). At the end of the incubation period, the supernatants containing the unbound cells were removed after placing the mixtures on a magnet, and the cells complexed with the iron oxide were washed three times with 0.1% TNBSA. The iron oxide and any bound cells were resuspended in 1 mL of 0.1% TNBSA to generate a final sample mixture.

100 μL of 1/10, 1/100, and 1/1000 dilutions of the final sample mixture were plated. Samples that used magnetic target cells containing the pCTCON-T2A-SsoFe2-hFc plasmid were plated on SD (-Leu) plates to quantify non-specific, TOM22 yeast binding to the target hFc yeast cells. Samples containing magnetic target cells based on the pCT302-T2A-SSoFe2-hFc plasmid were plated on SD (-Trp) plates to quantify specific, Sso7d-hFc yeast binding to the target hFc yeast cells (Figure 2.4). Colony counting was used to determine the number of Sso7d-hFc and TOM22 yeast recovered by the magnetic target cells. Ultimately, these values were used to

quantify enrichment and recovery of the Sso7d-hFc binder populations. There were slight differences in the number of cells magnetized for each replicate. The number of captured cells was scaled by the number of magnetized cells for each replicate.

Additionally, $1x10^6$ pCTCON-Sso7d-hFc and pCT302-TOM22 cells were spun down and resuspended in 1 mL of 0.1% TNBSA. 100 μL of a 1/1000 dilution of these cells were plated on -Trp and -Leu plates, respectively, and counted to quantify the number of these cells initially incubated with the magnetized target cells. Enrichment is defined as the percentage of Sso7d-hFc cells in the final population compared to the percentage of Sso7d-hFc cells in the initial population. Recovery is defined as the percentage of Sso7d-hFc cells from the initial population that were isolated in the population isolated by the magnetic target cells.

2.4.6 Sso7d library screening using targets displayed on magnetic yeast

Multiple rounds of magnetic sorting were carried out to isolate binders to TOM22 and c-Kit. A simultaneous secretion and display library of Sso7d mutants (-Trp) was used (diversity $2x10^8$)[46]. The pCT-NT-F2A vector and a synthetic oligonucleotide containing NNK codons at randomized positions described by Gera et al. [34] were used to construct the library, using protocols previously described by Cruz et al[46]. The library was screened against pCT302-T2A-SsoFe2-TOM22 (-Leu) or pCT302-T2A-SsoFe2-cKit (-Leu).

In the first round, $5x10^8$ cells expressing the target and SsoFe2 were resuspended in 9 mL of 1% TNBSA. The cells were magnetized by adding 1 mL of an iron oxide solution (4 mg/mL) followed by a 30 minute incubation at room temp. Based on the OD_{600} of the solution prior to the addition of the iron oxide and after the removal of yeast bound to the iron oxide, $2.5x10^8$ magnetized cells were aliquoted for use in the magnetic selections. The magnetized cells were washed 3X with 1% PBSA buffer and blocked with 1% PBSA for 1 hour at room temp in a 10 mL

volume. After, the cells were washed 3X with 0.1% PBSA prior to the addition of $2x10^9$ library cells and $2x10^9$ EBY100 cells in 10 mL of 0.1% PBSA. The selection took place for 2 hours at room temperature. Library cells bound to the magnetized target cells were separated from the unbound cells using a magnet. The iron oxide was washed 3 times with 0.1% PBSA prior to expansion in 20 mL of SDCAA (-TRP) media.

A similar process was carried out in the subsequent magnetic selection rounds. However, there was a reduction in number of magnetized target cells and library cells for these rounds. In the second round, $2.5x10^7$ magnetic target cells were incubated with $1x10^7$ library cells and $1x10^9$ EBY100 cells. The target-bound library cells were expanded in 5 mL of SDCAA (-Trp) media. In the third round, $2.5x10^6$ magnetic target cells were incubated with $1x10^5$ library cells and $1x10^9$ EBY100 cells. In these rounds, magnetization of the target cells took place in a total volume of 2 mL by incubating $5x10^7$ or $5x10^6$ cells in 1% PBSA, respectively for rounds 2 and 3. For round 2, 100 μL of the 4 mg/mL iron oxide solution was added while 10 μL of the iron oxide solution was incubated for round 3. After magnetization, the cells were appropriately aliquoted to carry forward the correct number of magnetized target cells.

In the fourth round, the library cells were treated with DTT to reduce display levels by ~80%, similar to Stern et al[47]. The DTT concentration resulting in a ~80% reduction of c-myc expression was empirically determined for each library using the following procedure. $5x10^6$ cells from the third round of the library screening were pelleted and washed twice with 10 mM Tris pH 7.5 buffer (1.24 g/L Tris-HCl, 0.26 g/L Tris base). Solutions of varying DTT concentration (1, 5, 7.5 10, 15, or 20 mM) were created using the 10 mM Tris pH 7.5 buffer. The yeast cells were resuspended in 20 μL of the different DTT concentration solutions at 30°C for 20 minutes with shaking. Subsequently, the cells were pelleted and washed twice with 0.1% PBSA. To quantify

ligand expression, the treated yeast cells were labeled with 50 μL of a 1:100 dilution of chicken-anti-c-myc antibody for 15 minutes at room temperature and washed once with 0.1% PBSA. Secondary labeling was performed using 50 μL of a 1:250 dilution of goat-anti-chicken 633 on ice. The mean fluorescence of 50,000 events from each sample was compared to an untreated control to determine the reduction in the display level. It was determined that 5 mM and 7.5 mM of DTT was required to reduce the display levels of the putative TOM22 and c-Kit binders by ~80%, respectively. During the fourth round of selection, $1x10^5$ DTT treated library cells and $1x10^9$ EBY100 were incubated with $1x10^6$ magnetized target cells. The target cells were magnetized as previously described for round 3 and appropriately aliquoted.

2.4.7 Construction of a second Sso7d yeast display library by random mutagenesis using error-prone PCR

DNA associated with the mutants recovered from the fourth round of screening was isolated using a Zymoprep yeast plasmid miniprep II kit (Zymo Research). For both the TOM22 and c-Kit screens, the Sso7d sequences from the isolated DNA were amplified by error prone PCR using nucleotide analogs with primers Pf6 and Pr6, as detailed by Gera et al[44]. The amplified error-prone PCR DNA was combined and amplified further using primers Pf8 and Pr8 in six identical 50 μL reactions. The PCR products were combined and purified using phenol: chloroform extraction. The DNA was concentrated using ethanol precipitation. Briefly, 1/10 volume of potassium acetate was added to the extracted DNA along with 10 μL of linear acrylamide followed by 2 volumes of ice-cold ethanol. The mixture was incubated at -20 °C overnight and centrifuged at 15,000 g for 10 minutes followed by removal of the supernatant. The pellet was washed once with 70% ethanol followed by a 100% ethanol wash and allowed to dry prior to resuspension in water. In parallel, 40 μg of pCT-NT-F2A-Sso7dhFc from Cruz-Teran et al. was digested with NheI

and BamHI restriction enzymes and concentrated by phenol:chloroform extraction and ethanol precipitation[46]. A yeast display library was generated by transforming the PCR product and the digested vector into electrocompetent yeast using the lithium acetate method previously described[54]. Two electroporation reactions were performed using a Bio-Rad Gene Pulser system (Bio-Rad) where 12 μg of PCR product and 4 μg of digested vector was added to 400 μL of electrocompetent *S. cerevisiae strain* EBY100. The electroporation was carried out at 2500 V, 25 μF, and 250 Ω. As a control, an identical electroporation reaction containing only the digested vector was also performed. The library diversity was quantified by plating serial dilutions of the transformation reaction onto SDCAA plates and estimated as $6x10^7$ for the TOM22-binder library and $5x10^7$ for the c-Kit-binder library.

2.4.8 Additional screening of mutagenized Sso7d library

Four additional rounds of magnetic selection were carried out using the libraries constructed by error prone PCR. Yeast cells dually displaying SsoFe2 and TOM22 or SsoFe2 and c-Kit were used as targets. For rounds 1-3, $1x10^9$ EBY100 and $1x10^9$ pCT302-Sso7d-hFc cells were incubated along with the library cells. All selections took place in 0.1% PBSA. For round 1, $6x10^8$ and $5x10^8$ library cells were incubated for the TOM22 and c-Kit screens, respectively, in a volume of 10 mL. $3x10^8$ cells with pCT302-TOM22-T2A-SsoFe2 or pCT302-c-Kit-T2A-SsoFe2 were magnetized. All magnetization and incubation steps were carried out as previously described. In round two, $1x10^7$ library cells were incubated with $2.5x10^7$ magnetized target cells while in round three $1x10^6$ library cells were incubated with $2.5x10^6$ magnetized target cells in a volume of 2 mL total. In the fourth round, the library cells were treated with 7.5 mM DTT as previously described to reduce the surface display level. $1x10^5$ DTT treated library cells were incubated with $1x10^6$ magnetized target cells in a volume of 2 mL. A Zymoprep yeast plasmid miniprep II kit

(Zymo Research) was used to extract the plasmid DNA of the clones recovered in the final populations. After, the DNA was transformed into electrocompetent Novablue cells and plated. 10 colonies were sequenced per target.

2.4.9 Nanobody library screening against targets displayed on magnetic yeast

The synthetic nanobody yeast display library used was gifted by the Kruse lab at Harvard University (diversity $5x10^8$) [35]. This library was screened against pCT302-T2A-SsoFe2-TOM22(-Leu) or pCT302-T2A-SsoFe2-c-Kit (-Leu) yeast cells for multiple rounds.

During the first screening round, a negative selection against pCT302-T2A-SsoFe2-hFC cells was performed prior to carrying out a positive selection against cells expressing either TOM22 or c-Kit. $2.5x10^8$ magnetized cells were used in both the negative and positive selection steps. Magnetization took place in 11.5 mL of 1% PBSA for $5x10^8$ cells using 2.6 mL of a 4 mg/mL iron oxide solution. The magnetized cells were appropriately aliquoted to carry forward the correct number of magnetized cells. After, the magnetized cells were blocked in 10 mL of 1% PBSA for one hour prior to their use and subsequently washed. Next, $5x10^9$ nanobody library cells in 10 mL of 0.1% PBSA were incubated with the hFc displaying magnetized cells for one hour at room temperature. After, the selection tube was placed on a magnet. Any cells that did not bind to the hFc cells were removed and incubated with the target displaying magnetized cells for 1 hour at room temperature. Finally, the iron oxide was washed four times with 0.1% PBSA prior to expansion in 20 mL of NB.SDCAA (-TRP) media.

A similar process was performed for subsequent rounds of magnetic selection with a reduction in the number of magnetized target cells and library cells. Negative and positive selection as previously described were carried out for each round. In the second round, $2.5x10^7$ magnetic cells were incubated with $5x10^7$ library cells and $1x10^9$ EBY100 cells in 2 mL of 1% PBSA +

0.05% Tween-20. The target-bound library cells were expanded in 5 mL of NB.SDCAA (-TRP media). In the third round, $1x10^6$ magnetic target cells were incubated with $5x10^5$ library cells, $1x10^9$ EBY100 cells, and $1x10^9$ pCT302-Sso7d-hFc cells in 2 mL of 1% PBS BSA + 0.05% Tween-20. No negative selection took place for the third round. In rounds two and three, the magnetization of the hFc displaying cells as well as the target cells took place in 2 mL of 1% PBSA using 200 μL of a 4 mg/mL iron oxide solution and $5x10^7$ cells. The magnetized cells were appropriately aliquoted to ensure the correct number of magnetized cells were carried forward. For both rounds, the iron oxide was washed five times with 1% PBS BSA + 0.05% tween with 10 seconds of vortexing in between washes. A Zymoprep yeast plasmid miniprep II kit was used to recover the plasmid DNA of the clones isolated in the final populations. The DNA was transformed into electrocompetent Novablue cells and plated. 10 colonies were sequenced per target.

2.4.10 Specificity analysis of secreted Sso7d binder populations

For the Sso7d screens, soluble protein representing the entire pool of binders isolated from the final selection was obtained by inducing the isolated cell populations for 72 hours in SGCAA medium 30°C. The secreted protein was purified from the supernatant using Ni-NTA agarose resin (Qiagen) as previously described[46]. The purified protein was concentrated using a Vivaspin 6, 3 kDa MWCO concentrator (Sartorius), and the protein concentration was estimated using a BCA assay.

Binding of the secreted pool to its respective target was analyzed using flow cytometry. The secreted TOM22 and c-Kit binders were incubated with pCT302-TOM22 and pCT302-c-Kit yeast cells, respectively. The secreted protein was treated overnight at 4°C with TEV protease to cleave the c-myc tag and F2A sequence from the protein. 1 μg of TEV protease was incubated per 25 μg of Sso7d protein. 2.5 μM of TEV protease-treated ligand was incubated with $5x10^6$ target

cells for 1 hour at room temperature. The cells were washed 1X with 0.1% PBSA prior to the incubation. The TEV protease and the secreted protein both contain 6xHis tags. Therefore, target cells were also incubated with TEV protease alone to assess non-specific binding of the protease. The TOM22 and c-Kit cells were labeled with 0.046 μM of TEV protease in 50 μL of 0.1% PBSA, respectively. The target cells were also labeled with 2.5 μM of untreated, secreted binder protein. After these incubations, the cells were washed once with 0.1% PBSA and incubated with 50 μL of a 1:100 dilution of anti-penta-His Alexa 647 antibody (Qiagen) for 10 minutes on ice. A secondary only control was also performed. After a final wash, the cells were analyzed using flow cytometry. The estimated molecular weight of the Sso7d mutant population is 13 kDa.

The TEV protease was expressed in Rosetta™ *E. coli* cells using the plasmid pRK793 (Addgene, Plasmid # 8827)[55]. Cells expressing the TEV protease were grown using the following protocol. A 5 mL overnight culture was grown in LB media and ampicillin and added to 1 liter of LB media that also contained 1X M9 minimal media, 1X 5052, 1 mM $MgSO_4$, and 1X ampicillin. 50X M9 minimal media contains 33.9 g/L Na_2HPO_4, 15 g/L KH_2PO_4, 5 g/L NH_4Cl, 2.5 g/L NaCl while 50X 5052 is 25% glycerol, 2.5% glucose, 10% alpha-D-lactose. The cells were grown at 37°C until an OD_{600} of 1 was reached. After, the cells were grown at 18 °C and harvested 24 hours later. The protein was purified using a BioLogic LP FPLC system (Bio-Rad) by metal affinity chromatography. The cells were collected and lysed via sonication in 35 mL of Buffer A-IMAC (50 mM Tris, 300 mM NaCl pH 8), loaded onto a 5 mL Bio-Rad Mini-Proafinity IMAC column (Bio-Rad), washed with 40 mL of Buffer A-IMAC, and eluted over a 40 mL linear gradient of Buffer B-IMAC (50 mM Tris, 300 mM NaCl, 500 mM imidazole pH 8). Pure fractions were dialyzed into 20 mM HEPES, 150 mM NaCl pH 7.8.

2.4.11 Protein purification of Sso7d & nanobody mutants

Sso7d.TOM22-1, Sso7d.TOM22-2, Sso7d.c-Kit, and Sso7d.cIgY were amplified by PCR from the isolated library plasmids with forward primer Pf9 and reverse primer Pr9 and inserted into pet22B™ between the NdeI and XhoI restriction sites. NB.TOM22-1, NB.TOM22-2, NB.c-Kit-1, and NB.c-Kit-2 were amplified from isolated library plasmids by PCR with forward primer Pf10 and reverse primer Pr10 and similarly cloned into pet22B™. Positive clones were transformed into Rosetta™ *E. coli* cells for expression. Expression was carried out in 2XYT media (10 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl) plus 1X ampicillin. A 5 mL overnight culture was used to inoculate a 1 L culture of 2XYT. The cells were induced using 0.5 mM IPTG when the OD_{600} reached between 0.6 and 0.8. Expression took place overnight at 20°C. The Sso7d proteins were purified by cation exchange and metal affinity chromatography while the nanobody proteins were only purified by metal affinity chromatography.

Induced Sso7d cell culture was collected and lysed using a French press into 35 mL of Buffer A-cat (50 mM Tris-HCl, 50 mM NaCl, pH 8), loaded onto a 5 mL Bio-Rad High S column (Bio-Rad), washed with 40 mL of Buffer A-cat, and eluted with a 40 mL linear gradient of Buffer B-cat (50 mM Tris-HCl, 1 M NaCl, pH 8) from 0-50% B. The fractions were analyzed by SDS-Page and fractions containing the protein of interest were pooled and loaded onto a 5 mL Mini-profinity IMAC column (Bio-Rad). Metal affinity chromatography was performed as previously described for the TEV protease. Because the nanobody proteins were only purified using IMAC, these proteins were initially lysed into 35 mL of Buffer-A-IMAC using a French press. Pure fractions were dialyzed into PBS pH 7.4. All protein purification steps were completed at room temperature.

Each protein, expect Sso7d.c-Kit, was biotinylated using a 5:1 molar excess of Ez-Link™ Sulfo-NHS-LC-Biotin (Thermo-Fisher Scientific) overnight at 4°C. Excess reagent was inactivated through dialysis against 50 mM Tris-HCl, 300 mM NaCl pH 7.5. The purified protein was concentrated using a Vivaspin 6, 3 kDa MWCO concentrator (Sartorius). Protein concentrations were estimated by BCA assay.

2.4.12 Estimation of K_D

K_D values of the isolated TOM22 Sso7d, TOM22 nanobody, and c-Kit nanobody mutants were estimated using yeast surface titrations. Depending on the binder type, pCT302-TOM22 or pCT302-c-Kit yeast cells were labeled with varying concentrations of the biotinylated Sso7d or nanobody mutants followed by secondary labeling with a 1:250 dilution of streptavidin R-phycoerythrin conjugate (SA-PE; Thermo Fisher Scientific). Flow cytometry analysis was completed as previously described[31]. All labeling was performed in 50 µL of 0.1% PBSA. At low labeling concentrations (≤10 nM), yeast displaying the target of interest were combined with un-induced yeast cells for ease of obtaining cell pellets by centrifugation. Under these conditions, the volume of the labeling reaction was sufficient to ensure that the Sso7d binder mutants were present in at least 10-fold excess over the cell surface targets. Samples were incubated at room temperature for 1 hour. The K_D of the binding interaction between the binder protein and its target was estimated using the following relationship:

$$F = \frac{F_{Max}[L]_0}{K_D + [L]_0} \quad \text{(Equation 1)}$$

Where F is background subtracted mean observed fluorescence, $[L]_O$ is the concentration of Sso7d mutant used for labeling the cells, and F_{Max} is the background subtracted fluorescence intensity when surface saturation is attained. Fluorescence signal was observed to decrease at

high values of $[L]_O$ (hook effect), as previously observed in other uses of yeast surface titrations[34]; these values were not included in K_D calculations. For each replicate, F_{max} was set to the highest value of fluorescence observed. A global non-linear least squares regression was used to fit the data to Equation 1 as previously described[31]. A single K_D value was used as the fitted parameter across all repeats.

A complete yeast titration was not generated for Sso7d.c-Kit due to its low affinity. $2x10^6$ pCT302-c-Kit yeast cells were incubated with 2, 5, 7.5, 10, and 15 μM of Sso7d.c-Kit. Sso7d.c-Kit was not biotinylated as one of its mutated residues was a lysine. Binding of Sso7d.c-Kit to the c-Kit displaying cells was detected using an anti-penta-His Alexa 647 antibody and flow cytometry analysis as previously described. Similar yeast surface titration analysis was also used to evaluate binding of Sso7d.cIgY to yeast cells displaying TOM22 and c-Kit. Concentrations of Sso7d.cIgY tested ranged from 4 nM to 2 μM. Because Sso7d.cIgY was biotinylated, binding was detected using SA-PE.

2.4.13 Enrichment of c-kit displaying yeast using magnetic streptavidin beads functionalized with c-kit nanobody

The c-Kit nanobodies were tested for their ability to isolate yeast cells displaying c-Kit from a heterogeneous population. Briefly, 25 μL of biotin binder Dynabeads (Thermo Fisher Scientific) were washed two times with 0.1% PBSA (PBS pH 7.4, 0.1% BSA). The beads were functionalized with either 1 μM NB.c-Kit-1, NB.c-Kit-2, or NB.TOM22-1 diluted in 100 μL of 0.1% PBSA overnight at 4°C. The beads were washed 2X with 0.1% PBSA the next day. $1x10^5$ c-Kit displaying yeast were mixed with $2x10^6$ non-displaying EBY100 yeast in 1 mL of 0.1% PBSA and added to the functionalized binder beads for 1 hour at room temperature. After one hour, the tube was placed on a magnet and the supernatant was removed. The supernatant was transferred

to a new set of functionalized binder beads and a second incubation was carried out. In all, the cell mixture was incubated with three different sets of beads all functionalized with the same binding protein.

After the third incubation, the supernatant containing the cells that did not bind to the functionalized beads was centrifuged, washed 1X with 0.1% PBSA, and then labeled with a 1:100 dilution of chicken anti-c-myc antibody for 15 minutes at room temperature. Subsequently, the cells were washed and secondary labeling was carried out with a 1:250 dilution of goat-anti-chicken 488 for 10 minutes on ice. All labeling was conducted in 50 μL of 0.1% PBSA. Anti-c-myc labeling was also performed on a cell population that did not undergo incubation with the binder proteins that contained $1x10^5$ c-Kit displaying yeast and $2x10^6$ EBY100 yeast. The EBY100 cells do not express c-myc, so any c-myc expression is attributed to cells expressing c-Kit. If c-Kit expressing cells were pulled out by the functionalized magnetic beads then the percentage of c-myc expressing cells would be greater in the initial population when compared to the population of cells that were not pulled out by the beads. Thus, the percent decrease in cells expressing c-myc between the final and initial populations was quantified to assess the depletion of c-Kit displaying yeast cells by the binder functionalized beads.

The c-Kit specific nanobodies were also immobilized onto His dynabeads (Thermo Fisher Scientific) following a similar procedure (2 μL His-Tag Isolation Dynabeads, 2.3 μM nanobody protein) to understand if immobilization onto streptavidin beads was impeding the nanobody binding. NB.TOM22-2 was used as the negative control and the incubations were carried out cold. The cell populations were only incubated twice with the functionalized beads, instead of three incubations. The depleted population was labeled with an antibody against the c-myc tag and

analyzed using flow cytometry as previously described to evaluate the depletion of c-Kit displaying cells from the population.

2.4.14 Immunoblotting analysis of mitochondria isolation

2 μL of streptavidin Hi-Sur Mag 0.15 μm beads (Ocean Nanotech) were washed two times with 0.1% PBSA and incubated overnight with 80 nM of Sso7d.TOM22-1, Sso7d.TOM22-2, NB.TOM22-1, or NB.TOM22-2 diluted in 100 μL using 0.1% PBSA. Thereafter, the beads were washed two times with 0.1% PBSA and blocked with 1 mL of 1% PBSA for two hours. Subsequently, the beads were washed two times with 0.1% PBSA, resuspended in 1 mL of 0.1% PBSA, and incubated with the cell lysate. As a control, 2 μL of plain streptavidin magnetic beads (i.e. without binder functionalization) were also incubated overnight in 100 μL of 0.1% PBSA and also subjected to similar washing and blocking prior to incubation with the cell lysate. As a benchmark control, 25 μL of anti-TOM22 magnetic microbeads (Miltenyi Biotec) were incubated with the cell lysate per the manufacturers recommended protocols[49]. For each experimental condition, two bead samples were generated and incubated separately with cell lysate to allow separate analysis of mitochondrial and endoplasmic reticulum binding when immunoblotting.

To generate the cell lysate, HEK-293T cells were grown to 70-90% confluency in Dulbecco's modified Eagle medium containing 10% fetal bovine serum (Thermo Fisher Scientific), washed with PBS, and lifted from the plate using Gibco TrypLe Select (Thermo Fisher Scientific). The cells were washed twice with ice cold PBS and resuspended in ice cold PBS containing cOmplete, mini protease inhibitor cocktail (Roche) at a concentration of $1x10^7$ cells/mL. 1 mL of cells was sheared through a 26 G needle at a time on ice about 20 times. A small aliquot of lysed cells was taken to ensure that ~80% of the cells were lysed.

1 mL of crude lysate was mixed with 8 mL of 0.1% PBSA and incubated with either the Sso7d immobilized magnetic beads, nanobody immobilized magnetic beads, or plain magnetic beads. For the commercially available TOM22 microbeads, 1 mL of crude lysate was mixed with 8.975 mL of 0.5% PBSA with 2 mM EDTA pH 7.2 (PEB buffer), per the manufacturer's recommendations. Incubations took place for 1 hour at 4 °C with rotation. After, the streptavidin Hi-Sur bead samples were centrifuged at 13,000 g for 1 minute and washed 3x with 0.1% PBSA. The anti-TOM22 microbead samples were recovered using the manufacturer's recommendations, supplied columns, and supplied magnet (Miltenyi Biotec). The column was placed into the magnet and washed with 3X with 1 mL of PEB buffer. The sample was then loaded onto the column adding 3.3 mL at a time and washed 3X with 3.3 mL of PEB buffer. Subsequently, the column was removed from the magnet and 1.5 mL of PEB buffer was pushed through the column using the provided syringe to release the beads into a collection tube.

After these initial washes, all samples, including the binder functionalized streptavidin beads and the anti-TOM22 microbeads, were pelleted twice at 13,000 g for 1 minute and washed twice with 0.32 M sucrose, 1 mM EDTA, and 10 mM Tris-HCl. The samples were pelleted again and resuspended in 9.75 µL of PBS and 3.75 µL of NuPAGE LDS sample buffer (Thermo Fisher Scientific) prior to boiling for 15 minutes at 98°C. The samples were cooled at room temperature for 10 minutes and then 1.5 µL of 50 mM TCEP was added. After a 10-minute incubation, the samples were loaded onto a 4-12% Bis Tris NuPAGE gel (Thermo Fisher Scientific) along with 10 µL of Novex™ Sharp Pre-stained Protein Standard (Thermo Fisher Scientific).

Two gels were run with each having lanes corresponding to the following experimental conditions: ladder, commercial anti-TOM22 micro beads, NB.TOM22-1, NB.TOM22-2, Sso7d.TOM22-1, Sso7d.TOM22-2, and plain streptavidin magnetic beads. Gels were blotted onto

a PDVF membrane using the iBlot system's program 3 for 6.5 minutes. After blotting, the membranes were blocked in 5% milk TBST (50 mM Tris HCl, 150 mM NaCl, 0.05% Tween 20 pH 7.4) for three hours. One membrane was incubated with a 1:200 dilution of mouse anti-TOM22 HRP (Santa Cruz Biotechnology; Dallas, Texas; Catalog number sc-58308 HRP; Lot number F2018) while the other membrane was incubated with a 1:200 dilution of mouse anti-calnexin HRP (Santa Cruz Biotechnology; Catalog number sc-23954 HRP; Lot number E0218) both in 15 mL of 2% milk TBST overnight. The next morning the membranes were flipped and allowed to incubate for three hours before being washed 3 times with PBS for 10 minutes each. Detection was carried out using the SuperSignal™ West Femto maximum sensitivity substrate (Thermo Fisher Scientific) according to the manufacturer's protocol. The West-Femto substrate was left on the membranes for 1 minute prior to imaging using a Syngene Gbox with an exposure time of 2 milliseconds and the West Femto protocol. Thereafter, the gel was illuminated with white light to image the ladder only.

For the images presented in this text, the blot images were inverted. The ladder images were aligned with the blot images, the regions of interest were clipped, and irrelevant lanes were removed. No other manipulation of the images occurred.

2.5 References

(1) Hirsch, F. R.; Varella-Garcia, M.; Cappuzzo, F. Predictive Value of EGFR and HER2 Overexpression in Advanced Non-Small-Cell Lung Cancer. *Oncogene* **2009**, *28* (S1), S32–S37.

(2) Várady, G.; Cserepes, J.; Németh, A.; Szabó, E.; Sarkadi, B. Cell Surface Membrane Proteins as Personalized Biomarkers: Where We Stand and Where We Are Headed. *Biomark. Med.* **2013**, *7* (5), 803–819.

(3) Firer, M. A.; Gellerman, G. Targeted Drug Delivery for Cancer Therapy: The Other Side of Antibodies. *J. Hematol. Oncol.* **2012**, *5* (1), 70.

(4) Diamantis, N.; Banerji, U. Antibody-Drug Conjugates--an Emerging Class of Cancer Treatment. *Br. J. Cancer* **2016**, *114* (4), 362–367.

(5) Boonstra, M. C.; de Geus, S. W. L.; Prevoo, H. A. J. M.; Hawinkels, L. J. A. C.; van de Velde, C. J. H.; Kuppen, P. J. K.; Vahrmeijer, A. L.; Sier, C. F. M. Selecting Targets for Tumor Imaging: An Overview of Cancer-Associated Membrane Proteins. *Biomark. Cancer* **2016**, *8*, 119–133.

(6) Case, B. A.; Kruziki, M. A.; Johnson, S. M.; Hackel, B. J. Engineered Charge Redistribution of Gp2 Proteins through Guided Diversity for Improved PET Imaging of Epidermal Growth Factor Receptor. *Bioconjug. Chem.* **2018**, *29* (5), 1646–1658.

(7) Tainsky, M. A. Genomic and Proteomic Biomarkers for Cancer: A Multitude of Opportunities. *Biochim. Biophys. Acta* **2009**, *1796* (2), 176–193.

(8) Boschetti, E.; D'Amato, A.; Candiano, G.; Righetti, P. G. Protein Biomarkers for Early Detection of Diseases: The Decisive Contribution of Combinatorial Peptide Ligand Libraries. *J. Proteomics* **2017,** *188*, 1-14.

(9) Satori, C. P.; Kostal, V.; Arriaga, E. A. Review on Recent Advances in the Analysis of Isolated Organelles. *Anal. Chim. Acta* **2012**, *753*, 8–18.

(10) Spohn, G.; Wiercinska, E.; Karpova, D.; Bunos, M.; Hümmer, C.; Wingenfeld, E.; Sorg, N.; Poppe, C.; Huppert, V.; Stuth, J.; et al. Automated CD34+ Cell Isolation of Peripheral Blood Stem Cell Apheresis Product. *Cytotherapy* **2015**, *17* (10), 1465–1471.

(11) Sun, C.; Hsieh, Y.-P.; Ma, S.; Geng, S.; Cao, Z.; Li, L.; Lu, C. Immunomagnetic Separation of Tumor Initiating Cells by Screening Two Surface Markers. *Sci. Rep.* **2017**, *7*, 40632.

(12) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science* **1985**, *228* (4705), 1315–1317.

(13) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. Phage Antibodies: Filamentous Phage Displaying Antibody Variable Domains. *Nature* **1990**, *348* (6301), 552–554.

(14) Boder, E. T.; Wittrup, K. D. Yeast Surface Display for Screening Combinatorial Polypeptide Libraries. *Nat. Biotechnol.* **1997**, *15* (6), 553–557.

(15) Beerli, R. R.; Bauer, M.; Buser, R. B.; Gwerder, M.; Muntwiler, S.; Maurer, P.; Saudan, P.;

Bachmann, M. F. Isolation of Human Monoclonal Antibodies by Mammalian Cell Display. *Proc. Natl. Acad. Sci.* **2008**, *105* (38), 14336–14341.

(16) Ho, M.; Nagata, S.; Pastan, I. Isolation of Anti-CD22 Fv with High Affinity by Fv Display on Human Cells. *Proc. Natl. Acad. Sci.* **2006**, *103* (25), 9637–9642.

(17) Yoshino, T.; Shimojo, A.; Maeda, Y.; Matsunaga, T. Inducible Expression of Transmembrane Proteins on Bacterial Magnetic Particles in Magnetospirillum Magneticum AMB-1. *Appl. Environ. Microbiol.* **2010**, *76* (4), 1152–1157.

(18) Baneyx, F.; Mujacic, M. Recombinant Protein Folding and Misfolding in Escherichia Coli. *Nat. Biotechnol.* **2004**, *22* (11), 1399–1408.

(19) Chae, P. S.; Rasmussen, S. G. F.; Rana, R. R.; Gotfryd, K.; Chandra, R.; Goren, M. A.; Kruse, A. C.; Nurva, S.; Loland, C. J.; Pierre, Y.; et al. Maltose-Neopentyl Glycol (MNG) Amphiphiles for Solubilization, Stabilization and Crystallization of Membrane Proteins. *Nat. Methods* **2010**, *7* (12), 1003–1008.

(20) Ackerman, M.; Levary, D.; Tobon, G.; Hackel, B.; Orcutt, K. D.; Wittrup, K. D. Highly Avid Magnetic Bead Capture: An Efficient Selection Method for de Novo Protein Engineering Utilizing Yeast Surface Display. *Biotechnol. Prog.* **2009**, *25* (3), 774–783.

(21) Lipes, B. D.; Chen, Y.-H.; Ma, H.; Staats, H. F.; Kenan, D. J.; Gunn, M. D. An Entirely Cell-Based System to Generate Single-Chain Antibodies against Cell Surface Receptors. *J. Mol. Biol.* **2008**, *379* (2), 261–272.

(22) Yang, Z.; Wan, Y.; Tao, P.; Qiang, M.; Dong, X.; Lin, C.-W.; Yang, G.; Zheng, T.; Lerner, R. A. A Cell-Cell Interaction Format for Selection of High-Affinity Antibodies to Membrane Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (30), 14971–14978.

(23) Stark, Y.; Venet, S.; Schmid, A. Whole Cell Panning with Phage Display. In *Methods in Molecular Biology*, Synthetic Antibodies ed.; Tiller T., Eds.; Humana Press: New York, 2017; Vol. 1575, pp 67–91.

(24) Zhou, Y.; Zou, H.; Zhang, S.; Marks, J. D. Internalizing Cancer Antibodies from Phage Libraries Selected on Tumor Cells and Yeast-Displayed Tumor Antigens. *J. Mol. Biol.* **2010**, *404* (1), 88–99.

(25) Zhao, L.; Qu, L.; Zhou, J.; Sun, Z.; Zou, H.; Chen, Y.-Y.; Marks, J. D.; Zhou, Y. High Throughput Identification of Monoclonal Antibodies to Membrane Bound and Secreted Proteins Using Yeast and Phage Display. *PLoS One* **2014**, *9* (10), e111339.

(26) Aggen, D. H.; Chervin, A. S.; Insaidoo, F. K.; Piepenbrink, K. H.; Baker, B. M.; Kranz, D. M. Identification and Engineering of Human Variable Regions That Allow Expression of Stable Single-Chain T Cell Receptors. *Protein Eng. Des. Sel.* **2011**, *24* (4), 361–372.

(27) Levy, R.; Forsyth, C. M.; LaPorte, S. L.; Geren, I. N.; Smith, L. A.; Marks, J. D. Fine and Domain-Level Epitope Mapping of Botulinum Neurotoxin Type A Neutralizing Antibodies by Yeast Surface Display. *J. Mol. Biol.* **2007**, *365* (1), 196–210.

(28) Piatesi, A.; Howland, S. W.; Rakestraw, J. A.; Renner, C.; Robson, N.; Cebon, J.; Maraskovsky, E.; Ritter, G.; Old, L.; Wittrup, K. D. Directed Evolution for Improved

Secretion of Cancer–Testis Antigen NY-ESO-1 from Yeast. *Protein Expr. Purif.* **2006**, *48* (2), 232–242.

(29) Johns, T. G.; Adams, T. E.; Cochran, J. R.; Hall, N. E.; Hoyne, P. A.; Olsen, M. J.; Kim, Y.-S.; Rothacker, J.; Nice, E. C.; Walker, F.; et al. Identification of the Epitope for the Epidermal Growth Factor Receptor-Specific Monoclonal Antibody 806 Reveals That It Preferentially Recognizes an Untethered Form of the Receptor. *J. Biol. Chem.* **2004**, *279* (29), 30375–30384.

(30) Cochran, J. R.; Kim, Y.-S.; Olsen, M. J.; Bhandari, R.; Dane Wittrup, K. Domain-Level Antibody Epitope Mapping through Yeast Surface Display of Epidermal Growth Factor Receptor Fragments. *J. Immunol. Methods* **2004**, *287* (1-2), 147–158.

(31) Gera, N.; Hussain, M.; Rao, B. M. Protein Selection Using Yeast Surface Display. *Methods* **2013**, *60* (1), 15–26.

(32) Eckart, M. R.; Bussineau, C. M. Quality and Authenticity of Heterologous Proteins Synthesized in Yeast. *Curr. Opin. Biotechnol.* **1996**, *7* (5), 525–530.

(33) Cruz-Teran, C. A.; Bacon, K.; McArthur, N.; Rao, B. M. An Engineered Sso7d Variant Enables Efficient Magnetization of Yeast Cells. *ACS Comb. Sci.* **2018**, *20* (10), 579–584.

(34) Gera, N.; Hussain, M.; Wright, R. C.; Rao, B. M. Highly Stable Binding Proteins Derived from the Hyperthermophilic Sso7d Scaffold. *J. Mol. Biol.* **2011**, *409* (4), 601–616.

(35) McMahon, C.; Baier, A. S.; Pascolutti, R.; Wegrecki, M.; Zheng, S.; Ong, J. X.; Erlandson, S. C.; Hilger, D.; Rasmussen, S. G. F.; Ring, A. M.; et al. Yeast Surface Display Platform for Rapid Discovery of Conformationally Selective Nanobodies. *Nat. Struct. Mol. Biol.* **2018**, *25* (3), 289–296.

(36) Grzeschik, J.; Hinz, S. C.; Könning, D.; Pirzer, T.; Becker, S.; Zielonka, S.; Kolmar, H. A Simplified Procedure for Antibody Engineering by Yeast Surface Display: Coupling Display Levels and Target Binding by Ribosomal Skipping. *Biotechnol. J.* **2017**, *12* (2), 1600454.

(37) Rosowski, S.; Becker, S.; Toleikis, L.; Valldorf, B.; Grzeschik, J.; Demir, D.; Willenbücher, I.; Gaa, R.; Kolmar, H.; Zielonka, S.; et al. A Novel One-Step Approach for the Construction of Yeast Surface Display Fab Antibody Libraries. *Microb. Cell Fact.* **2018**, *17* (1), 3.

(38) Chao, G.; Lau, W. L.; Hackel, B. J.; Sazinsky, S. L.; Lippow, S. M.; Wittrup, K. D. Isolating and Engineering Human Antibodies Using Yeast Surface Display. *Nat. Protoc.* **2006**, *1* (2), 755–768.

(39) Yocum, R. R.; Hanley, S.; West, R.; Ptashne, M.; Ptashne, M. Use of LacZ Fusions to Delimit Regulatory Elements of the Inducible Divergent GAL1-GAL10 Promoter in Saccharomyces Cerevisiae. *Mol. Cell. Biol.* **1984**, *4* (10), 1985–1998.

(40) Da Silva, N. A.; Bailey, J. E. Influence of Plasmid Origin and Promoter Strength in Fermentations of Recombinant Yeast. *Biotechnol. Bioeng.* **1991**, *37* (4), 318–324.

(41) Zhao, L.; Qu, L.; Zhou, J.; Sun, Z.; Zou, H.; Chen, Y.-Y.; Marks, J. D.; Zhou, Y. High Throughput Identification of Monoclonal Antibodies to Membrane Bound and Secreted

Proteins Using Yeast and Phage Display. *PLoS One* **2014**, *9* (10), e111339.

(42) Ryan, M. D.; Mehrotra, A.; Gani, D.; Donnelly, M. L. L.; Hughes, L. E.; Luke, G.; Li, X. Analysis of the Aphthovirus 2A/2B Polyprotein 'Cleavage' Mechanism Indicates Not a Proteolytic Reaction, but a Novel Translational Effect: A Putative Ribosomal 'Skip.' *J. Gen. Virol.* **2001**, *82* (5), 1013–1025.

(43) ten Dam, E.; Donnelly, M. L. L.; Gani, D.; Luke, G.; Ryan, M. D.; Mendoza, H.; Hughes, L. E. The 'Cleavage' Activities of Foot-and-Mouth Disease Virus 2A Site-Directed Mutants and Naturally Occurring '2A-like' Sequences. *J. Gen. Virol.* **2001**, *82* (5), 1027–1041.

(44) Gera, N.; Hill, A. B.; White, D. P.; Carbonell, R. G.; Rao, B. M. Design of PH Sensitive Binding Proteins from the Hyperthermophilic Sso7d Scaffold. *PLoS One* **2012**, *7* (11), e48928.

(45) Stern, L. A.; Schrack, I. A.; Johnson, S. M.; Deshpande, A.; Bennett, N. R.; Harasymiw, L. A.; Gardner, M. K.; Hackel, B. J. Geometry and Expression Enhance Enrichment of Functional Yeast-Displayed Ligands via Cell Panning. *Biotechnol. Bioeng.* **2016**, *113* (11), 2328–2341.

(46) Cruz-Teran, C. A.; Tiruthani, K.; Mischler, A.; Rao, B. M. Inefficient Ribosomal Skipping Enables Simultaneous Secretion and Display of Proteins in *Saccharomyces Cerevisiae*. *ACS Synth. Biol.* **2017**, *6* (11), 2096–2107.

(47) Stern, L. A.; Csizmar, C. M.; Woldring, D. R.; Wagner, C. R.; Hackel, B. J. Titratable Avidity Reduction Enhances Affinity Discrimination in Mammalian Cellular Selections of Yeast-Displayed Ligands. *ACS Comb. Sci.* **2017**, *19* (5), 315–323.

(48) Hussain, M.; Gera, N.; Hill, A. B.; Rao, B. M. Scaffold Diversification Enhances Effectiveness of a Superlibrary of Hyperthermophilic Proteins. *ACS Synth. Biol.* **2013**, *2* (1), 6–13.

(49) Hornig-Do, H.-T.; Günther, G.; Bust, M.; Lehnartz, P.; Bosio, A.; Wiesner, R. J. Isolation of Functional Pure Mitochondria by Superparamagnetic Microbeads. *Anal. Biochem.* **2009**, *389* (1), 1–5.

(50) Devall, M.; Burrage, J.; Caswell, R.; Johnson, M.; Troakes, C.; Al-Sarraj, S.; Jeffries, A. R.; Mill, J.; Lunnon, K. A Comparison of Mitochondrial DNA Isolation Methods in Frozen Post-Mortem Human Brain Tissue—Applications for Studies of Mitochondrial Genetics in Brain Disorders. *Biotechniques* **2015**, *59* (4), 241-247.

(51) Macheiner, T.; Fengler, V. H. I.; Agreiter, M.; Eisenberg, T.; Madeo, F.; Kolb, D.; Huppertz, B.; Ackbar, R.; Sargsyan, K. Magnetomitotransfer: An Efficient Way for Direct Mitochondria Transfer into Cultured Human Cells. *Sci. Rep.* **2016**, *6* (1), 35571.

(52) Sarkar, P.; Randall, S. M.; Muddiman, D. C.; Rao, B. M. Targeted Proteomics of the Secretory Pathway Reveals the Secretome of Mouse Embryonic Fibroblasts and Human Embryonic Stem Cells. *Mol. Cell. Proteomics* **2012**, *11* (12), 1829–1839.

(53) Carlin, K. B.; Cruz-Teran, C. A.; Kumar, J. P.; Gomes, C.; Rao, B. M. Combinatorial Pairwise Assembly Efficiently Generates High Affinity Binders and Enables a "Mix-and-Read" Detection Scheme. *ACS Synth. Biol.* **2016**, *5* (12), 1348–1354.

(54) Benatuil, L.; Perez, J. M.; Belk, J.; Hsieh, C.-M. An Improved Yeast Transformation Method for the Generation of Very Large Human Antibody Libraries. *Protein Eng. Des. Sel.* **2010**, *23* (4), 155–159.

(55) Kapust, R. B.; Tözsér, J.; Fox, J. D.; Anderson, D. E.; Cherry, S.; Copeland, T. D.; Waugh, D. S. Tobacco Etch Virus Protease: Mechanism of Autolysis and Rational Design of Stable Mutants with Wild-Type Catalytic Proficiency. *Protein Eng. Des. Sel.* **2001**, *14* (12), 993–1000.

CHAPTER 3

Quantitative Yeast-Yeast Two Hybrid for Discovery and Binding Affinity Estimation of Protein-Protein Interactions

Adapted from Bacon K, Blain A, Bowen J, Burroughs M, McArthrur N, Menegatti S, and Rao, B.M. (2021) Quantitative yeast-yeast two hybrid for discovery and binding affinity estimation of protein-protein interactions. *ACS Synthetic Biology*. Accepted for publication February 15, 2021.

3.1 Introduction

Quantitative binding affinity estimation of protein-protein interactions (PPIs) is important for system-level analysis of intracellular biochemical signaling pathways[1–6], as well as characterization of binding proteins isolated from combinatorial libraries using screening platforms such as phage display[7,8] and yeast surface display[9,10]. Yeast two hybrid (Y2H) assays are a powerful means to discover putative protein interaction partners ("prey") for specific proteins of interest ("bait") in a high throughput manner[11,12]. Modifications to Y2H have enabled the elucidation of interactions that rely on post-translational modifications by tethering a modifying enzyme to the bait protein[13,14], as well as semi-quantitative characterization of relative binding affinities using yeast surface displayed bait proteins[15]. More recently, yeast cell surface expression of bait and prey proteins as fusions to yeast mating proteins has also been used to identify PPIs in a high throughput fashion; the quantitative relationship between binding affinities and efficiencies of yeast mating can be used to assess the strength of bait-prey binding[16]. However, despite these advances, obtaining quantitative estimates of binding affinity – specifically equilibrium dissociation constant (K_D) values – is generally not feasible using the aforementioned approaches.

Here, we describe a quantitative yeast-yeast two hybrid (qYY2H) system that enables the discovery of protein interaction pairs, as well as efficient and quantitative estimation of interaction binding affinities (K_D). The requirement for recombinant, soluble protein is often a significant hurdle for high throughput binding affinity estimation[17,18]. In qYY2H, both the bait and prey proteins are expressed as cell surface fusions using the widely implemented Aga2-based yeast display platform[9,10], which bypasses the need for recombinant protein (Figure 3.1). Bait cells co-express the iron oxide binding protein, SsoFe2, as a cell surface fusion, enabling magnetization of bait yeast and separation of bait-prey complexes[19,20]. An engineered luciferase reporter,

NanoLuc[21,22], is expressed on the surface of prey yeast that can be used to quantify the number of prey cells complexed with the bait cells, and therefore, the strength of yeast-yeast binding interactions that are mediated by multivalent associations between bait and prey proteins expressed on each yeast population. Importantly, we describe a semi-empirical framework that enables quantitative estimation of the K_D of monovalent interactions between the bait and prey proteins using measurements of yeast-yeast binding interactions. Further, we show that qYY2H can be used to efficiently identify and quantitatively characterize putative interaction partners of bait proteins from a yeast displayed cDNA library. Finally, quantitative assessment of binding affinities is particularly challenging using conventional methods when one or both proteins under investigation are post-translationally modified. We have modified qYY2H to address this challenge by expressing and sequestering the modifying enzyme in the endoplasmic reticulum (ER), which can act on the expressed bait protein prior to surface display. Using this approach, we show that qYY2H can be used to quantitatively characterize binding interactions mediated by phosphorylated bait.

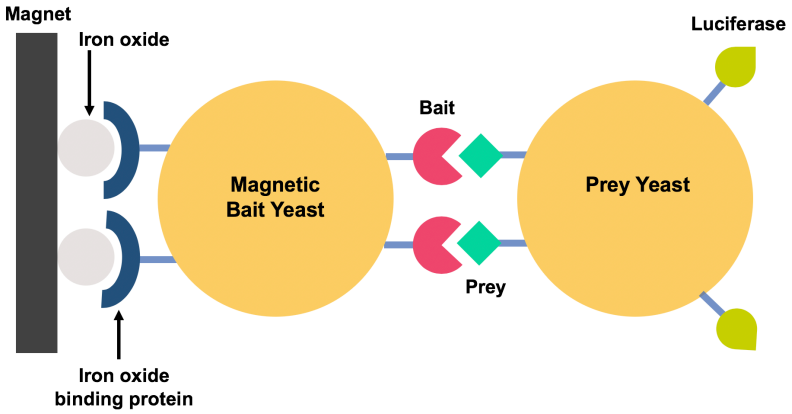


Figure 3.1. Quantitative yeast-yeast two hybrid (qYY2H). Yeast cells co-expressing the bait protein and SsoFe2 are magnetized by incubation with iron oxide particles. Magnetic separation is used to isolate yeast cells expressing prey proteins that interact with the bait cells. Prey yeast also display a luciferase reporter. The number of prey cells captured can be accurately quantified using a luminescence assay, in conjunction with a standard curve.

3.2 Results and Discussion

3.2.1 qYY2H enables relative binding affinity discrimination between bait-prey interactions

In qYY2H, the interaction of magnetized bait yeast with prey cells results in the formation of bait-prey, yeast-yeast complexes that can be separated using a magnet (Figure 3.2). Subsequently, the number of prey cells captured can be accurately quantified using a luminescence assay. We hypothesized that the number of prey cells captured can be used to quantitatively discriminate between the relative strengths of bait-prey binding interactions. To test this hypothesis, we investigated, in the context of qYY2H, the interaction of two previously characterized binding proteins (prey) based on the Sso7d protein scaffold – Sso7d.hFc (K_D = 450 nM) and Sso7d.ev.hFc (K_D = 5280 nM) – with the Fc portion of human immunoglobulin G (hFc; bait). The number of Sso7d.hFc prey cells captured was ~3-fold higher than the number of Sso7d.ev.hFc prey cells recovered (Figure 3.3A). Further, the number of cells captured displaying the lower affinity Sso7d.ev.hFc was ~ 4-fold higher than background (number of cells displaying an irrelevant prey protein). These results show that qYY2H can be used to discriminate between the relative strengths of bait-prey interactions. It is important to note that the high sensitivity of the NanoLuc reporter enzyme is critical for discrimination of relative binding strengths in qYY2H. Relative binding strengths of bait-prey interactions could not be assessed when glucose oxidase (GOx) was used as a reporter (Figure 3.4). Additionally, non-specific binding between yeast cells is common. Therefore, inclusion of excess competitor, non-displaying EBY100 yeast and a non-ionic detergent, such as Tween-20, is important (Figure 3.2).

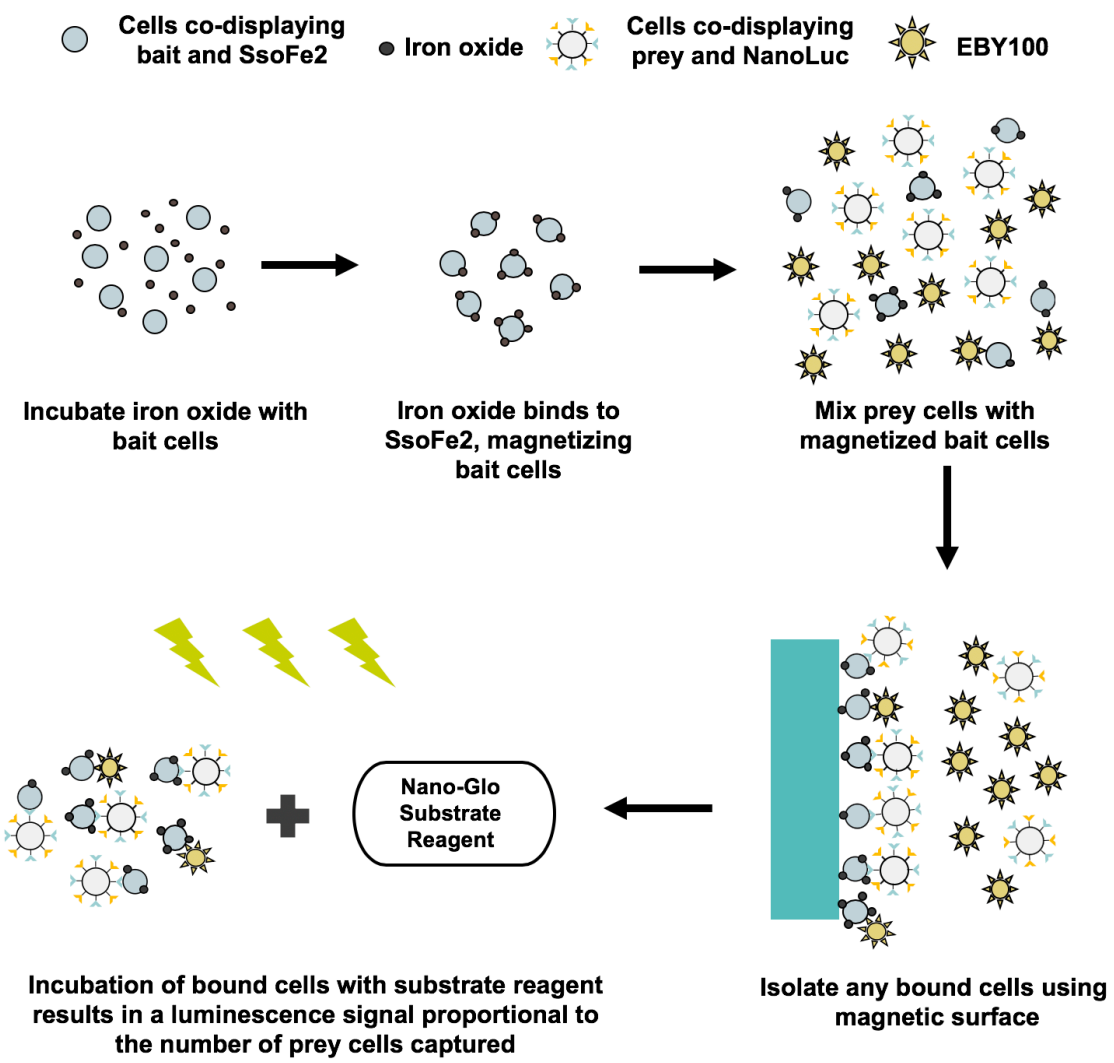


Figure 3.2. Detailed strategy for quantitative yeast-yeast two hybrid assay. The steps include the magnetization of the bait cells after incubation with iron oxide, incubation of the bait and prey cells, isolation of any prey cells bound by the magnetic bait cells using a magnet, and quantification of prey cell capture using the luminescence read out after addition of the Nano-Glo reagent.

Interestingly, quantitative discrimination of relative binding strengths was also observed when yeast cells displaying the lysozyme-binding proteins Sso7d.BVL.Lys (K_D = 1.7 nM) and Sso7d.NTL.Lys (K_D = 1300 nM) were used as prey, and biotinylated bait protein (lysozyme) was immobilized on streptavidin coated magnetic beads (Figure 3.3B). Recovery after magnetic separation was ~ 10-fold greater for cells displaying the higher affinity Sso7d.BVL.Lys. Our results are strikingly different from those previously reported by Ackerman et al [23]. In that study,

the recovery of cells displaying lysozyme-binding proteins was not affected by the displayed protein's binding affinity for lysozyme, which was similarly immobilized on streptavidin-coated magnetic beads. A likely explanation for this discrepancy is the lower ratio of cells to magnetic beads used in our study. Taken together, our results show that the efficiency of magnetically isolating prey cells – using magnetized yeast cells displaying the bait or magnetic beads coated with bait – has a quantitative dependence on the binding affinity of the bait-prey interaction. Notably, this observation is contrary to the generally accepted view that magnetic selections, which are employed during combinatorial screening of yeast display libraries, do not discriminate between high and low affinity binders.

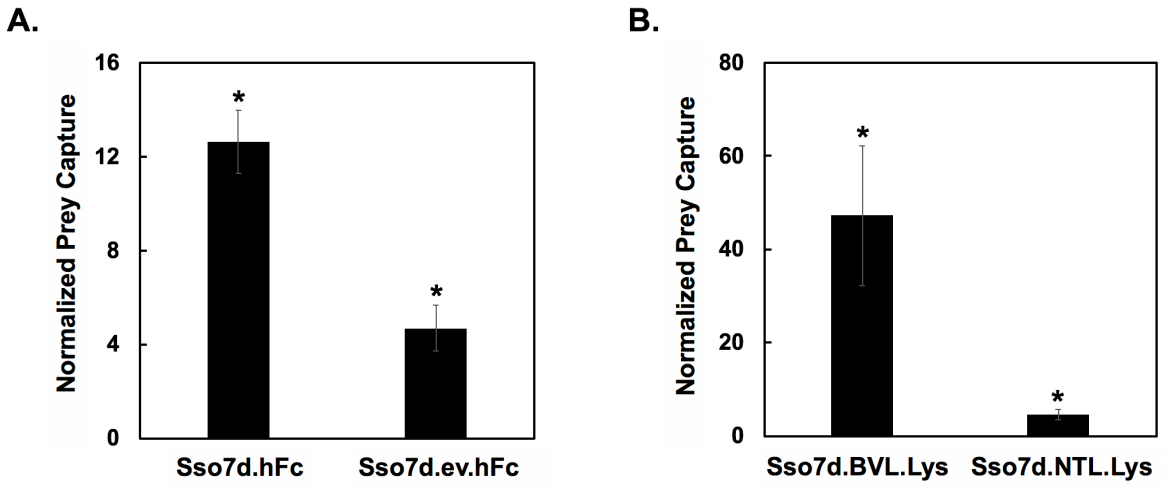


Figure 3.3. qYY2H enables discrimination of relative binding affinities associated with bait-prey interactions through display of a NanoLuc reporter. **(A)** Recovery of prey cells by magnetic yeast displaying the bait protein (hFc). The prey cells displayed either Sso7d.hFc, Sso7d.ev.hFc, or an irrelevant protein, NanoLuc. Recovery of the cells displaying the hFc binding proteins was normalized by the capture of cells displaying NanoLuc. * represents $p < 0.05$ for a two tailed, paired t-test in comparison to the capture of NanoLuc cells. **(B)** Recovery of prey cells by magnetic beads immobilized with the bait protein (lysozyme). The prey cells displayed either Sso7d.BVL.Lys, Sso7d.NTL.Lys, or an irrelevant protein, Sso7d.hFc. Recovery of the cells displaying the lysozyme binding proteins was normalized by the capture of cells displaying Sso7d.hFc. * represents $p < 0.05$ for a two tailed, paired t-test in comparison to the capture of Sso7d.hFc cells. Error bars represent the standard error of the mean for three replicates throughout.

3.2.2 Enzyme sensitivity is important to quantitatively assess bait-prey interactions

While the qYY2H assay ultimately relies on the use of luciferase as a reporter, other reporter enzymes were evaluated for their sensitivity to distinguish the binding interactions between prey and bait yeast. Another enzyme considered was glucose oxidase (GOx), which uses a colorimetric assay for detection. First, we incubated yeast cells co-expressing prey proteins and GOx with magnetic beads functionalized with the soluble bait protein. Any prey cells bound to the magnetic beads were isolated and incubated with the GOx assay reagent. The assay reagent contains dextrose, 3,3′,5,5′-Tetramethylbenzidine (TMB), and horseradish peroxidase (HRP). In the presence of GOx, glucose reacts with oxygen to form glucono delta-lactone and hydrogen peroxide. HRP reduces the generated hydrogen peroxide to water using TMB as a hydrogen donor, resulting in the formation of a diimine that causes the solution to turn blue. The reaction is halted by the addition of sulfuric acid, resulting in the solution turning yellow. The intensity of the color change can be read using a spectrophotometer. The intensity of the color change is directly proportional to the concentration of GOx present in the solution and, accordingly, the number of prey cells captured.

Three previously identified prey, derived from the Sso7d scaffold, with affinity for lysozyme were considered, Sso7d.BVL.Lys (K_D=1.7 nM), Sso7d.CTL.Lys (K_D=250 nM), and Sso7d.NTL.Lys (K_D=1300 nM) as well as an irrelevant Sso7d mutant, Sso7d.hFc[24,25]. The GOx signal was higher for each of the lysozyme mutants in comparison to the negative control, suggesting that the bait functionalized magnetic beads were capturing cells expressing the lysozyme binding proteins at a higher level than cells expressing the irrelevant Sso7d mutant (Figure 3.4A). However, the difference in signal between the lysozyme binding prey was not significantly different, though the mutants vary in affinity. These results suggest that GOx as a

reporter is not sensitive enough to distinguish the affinities between different mutants. GOx can be used to distinguish if two proteins interact, but it cannot be used to rank order affinities between different prey proteins for a single bait protein. The sensitivity of the assay may increase if the TMB substrate concentration is higher. However, the solubility of the TMB substrate was reached in these assays[26].

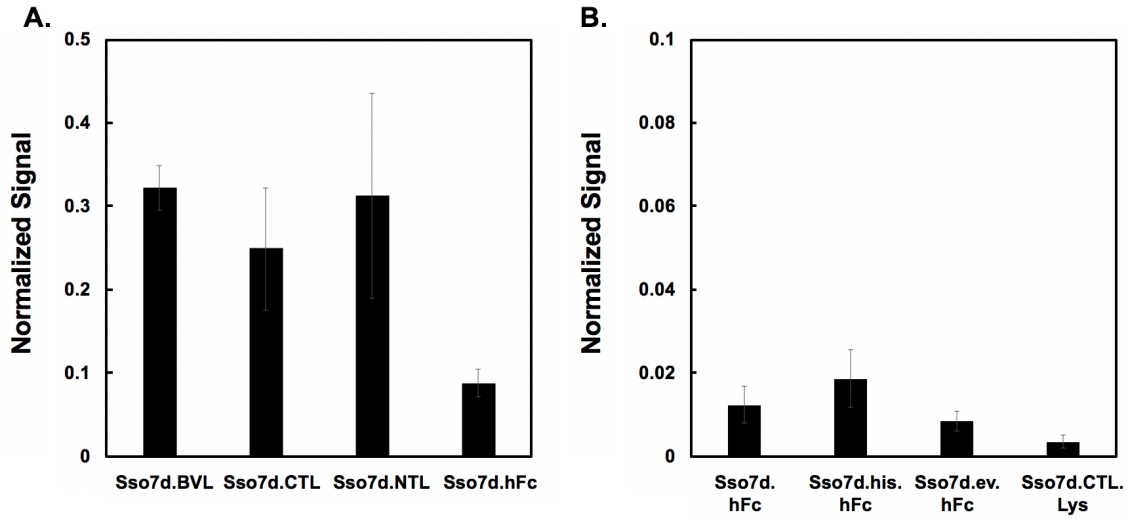


Figure 3.4. Evaluation of GOx as a reporter enzyme for the quantitative assessment of bait-prey interactions as a function of prey cell recovery. Prey proteins were expressed as yeast surface fusions along with GOx. **(A)** Normalized GOx signal proportional to recovery of prey cells by magnetic beads functionalized with lysozyme. The prey cells displayed either Sso7d.BVL.Lys, Sso7d.CTL.Lys, Sso7d.NTL.Lys, or an irrelevant protein, Sso7d.hFc. **(B)** Normalized GOx signal proportional to recover of prey cells by magnetic yeast displaying the bait protein (hFc). The prey cells displayed either Ssso7d.hFc, Sso7d.his.hFc, Ssso7d.ev.hFc, or an irrelevant protein, Sso7d.CTL.Lys. GOx signals were normalized throughout by the mean fluorescence associated with prey expression (detected using an anti-c-myc antibody) and by a GOx activity control. Error bars represent the standard error of the mean for three repeats throughout.

We also evaluated the use of GOx as a reporter when the bait protein is expressed on the surface of magnetic yeast cells (Figure 3.4B)[19]. For these assays, three previously identified prey, derived from the Sso7d scaffold, with affinity for the Fc portion of human IgG (hFc), Sso7d.hFc (K_D=450 nM), Sso7d.his.hFc (K_D=1647 nM), and Sso7d.ev.hFc (K_D=5280 nM), were co-expressed with GOx[25]. An irrelevant Sso7d mutant, Sso7d.CTL.Lys, was also considered[24]. Here, the GOx assays performed on the captured prey cells barely generated a signal above background.

It is important to note that the prey cells were not removed from the magnetic bait cells prior to performing the GOx assay. Color formation depends on the interaction of hydrogen peroxide with HRP which could be negatively affected if the generated hydrogen peroxide decomposes in the presence of the iron oxide used for magnetization instead of interacting with HRP[27–29]. Iron oxide has also been shown to promiscuously bind to many proteins in solution[20,30–33]. If HRP were to associate with iron oxide, thereby blocking the substrate binding site, then the assay's productivity would be negatively affected.

Subsequently, HRP was explored in a similar manner for its use as a reporter enzyme due to the limitations observed with GOx. HRP activity was detected using 1-Step Ultra TMB Elisa Substrate solution. However, it is likely the substrate solution is lysing any yeast cells present due to its low pH (3.35-3.75). After lysis, any peroxidases from inside the cell are released, compounding the signal readout beyond the signal associated with surface displayed HRP. We tested this hypothesis by incubating non-displaying EBY100 cells with the substrate solution. We were able to obtain a signal even though these cells do not express surface displayed HRP, suggesting that the cells are lysed after the substrate solution is added. Due to the limitations of GOx and HRP, we chose to use luciferase as the reporter when performing qYY2H assays.

3.2.3 A semi-empirical framework enables estimation of K_D using qYY2H

Formation of bait-prey, yeast-yeast complexes in qYY2H is driven by high avidity interactions between surface displayed bait and prey proteins. The number of bait-prey, yeast-yeast complexes isolated by magnetic separation, as quantified via the luminescence signal, was related to the concentration of prey cells used, through a parameter similar to the K_D term used when fitting a monovalent binding isotherm to yeast surface titrations with soluble prey protein; this

parameter was denoted as $K_{D,\,MV}$ to indicate the multivalent interaction between the bait and prey yeast:

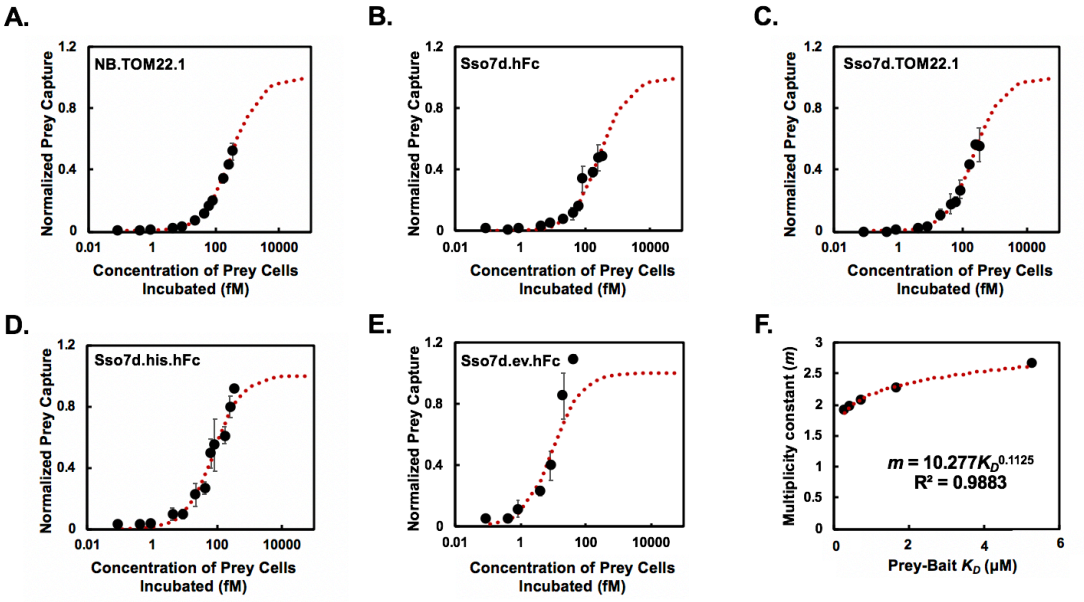
$$[\text{Complex}] = \frac{[\text{Max}][\text{Prey}]}{\text{K}_{\text{D,MV}}+[\text{Prey}]} \text{ (Equation 1)}$$

Here [*Complex*] is the molar concentration of bait-prey, yeast-yeast complexes, [*Prey*] is the initial molar concentration of prey yeast used in the experiment, and [*Max*] is the maximum molar concentration of the prey cells captured for a given number of magnetized bait yeast. We generated titration curves for seven bait-prey pairs where varying concentrations of prey cells (10^5 – $4x10^8$ cells) were incubated with 10^7 magnetized bait cells in the presence of 10^9 non-displaying, EBY100 yeast. The number of prey cells captured was quantified using a luminescence assay. The bait-prey pairs comprise binders to hFc and TOM22 based on the Sso7d or nanobody scaffolds and span a range of binding affinities (K_D ~ 200 nM to K_D ~ 5 μM). Binding affinities were previously estimated via yeast surface titration with soluble protein. $K_{D,MV}$ and [*Max*] were estimated for each bait-prey pair (Figure 3.5A-E and Figure 3.6) by non-linear regression using Equation 1. It is important to note that saturation was not observed in several titration curves, particularly for higher affinity bait-prey pairs; increasing the concentration of incubated prey yeast further to saturate the binding of the bait yeast is not feasible in these cases.

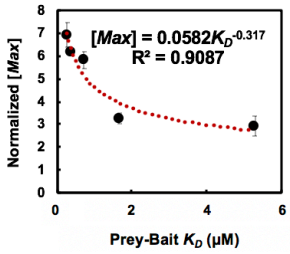
We hypothesized that a quantitative relationship exists between the binding affinity of the monovalent bait-prey interaction (K_D) and $K_{D,MV}$, similar to a previously outlined relationship that describes the multivalent interaction between viral surface proteins and cell surface receptors[34]:

$$\text{K}_{\text{D,MV}} = \text{K}_{\text{D}}{}^{\text{m}} \text{ (Equation 2)}$$

where m is a multiplicity constant that accounts for contributions of multiple monovalent bait-prey interactions occurring between bait and prey yeast. We used fitted-values of $K_{D,MV}$ and previously estimated values of K_D to calculate m for five randomly chosen bait-prey pairs (Figure 3.5H). We observed that m differed with K_D, and this variation could be described using a power law relationship (Figure 3.5F), with m monotonically increasing with K_D. We also observed a power law relationship between [*Max*] and previously estimated values of K_D (Figure 3.5G). It is important to note that the parameter $K_{D,MV}$ is not directly analogous to the monovalent K_D parameter. Indeed, unlike monovalent interactions where K_D decreases with increasing strength of monovalent binding, we observe that $K_{D,\,MV}$ is lower for weaker bait-prey binding interactions (i.e. higher K_D of monovalent bait-prey interaction) (Figure 3.5H). We also observe that prey with weaker affinity of monovalent interaction with the bait (i.e. higher K_D) exhibited smaller [*Max*] values than prey with stronger monovalent affinity for the bait (Figure 3.5G). Nevertheless, fewer bait-prey yeast-yeast complexes are isolated for bait-prey interactions with lower affinity of monovalent interaction (i.e. higher K_D), as intuitively expected. Thus, the strength of the multivalent bait-prey yeast-yeast interaction is dependent on both $K_{D,\,MV}$ and [*Max*]; these parameters allow us to link the K_D of monovalent bait-prey interaction to the number of bait-prey yeast-yeast complexes isolated. Notably, Equation 2 and the relationship developed between m and K_D (Figure 3.5F) can be used to estimate K_D values using experimentally determined values of $K_{D,MV}$ obtained from yeast-yeast titrations. Using this approach, K_D values were calculated for two bait-prey interactions randomly designated as "unknowns". The obtained K_D values were in reasonable agreement with K_D estimates from yeast surface titrations using soluble prey protein (Figure 3.5I). These results show that qYY2H can be used to obtain quantitative estimates of binding affinity (K_D) using yeast-displayed bait and prey proteins.



G.



H.

	$K_{D,MV}$ (fM)	YST K_D (nM)	m
NB.TOM22.1	318 (303.0 – 333.9)	272 (238 – 312)	1.90
Sso7d.hFc	273.6 (237.1 – 317.2)	400 (366 – 437)	1.96
Sso7d.TOM22.1	221.0 (164.1 – 305.4)	732 (632 – 848)	2.06
Sso7d.his.hFc	75.7 (58.6 – 97.9)	1670 (1337 – 3000)	2.27
Sso7d.ev.hFc	9.8 (5.2 – 18.4)	5280 (4500 – 6700)	2.65

I.

	$K_{D,MV}$ (fM)	YST K_D (nM)	qYY2H K_D (nM)	m
Sso7d.TOM22.2	684.3 (620.0 – 760.1)	362 (310 – 423)	209 (193 – 225)	1.82
NB.TOM22.2	49.5 (38.2– 64.3)	1934 (1693 – 2209)	1711 (1363 – 2155)	2.31

Figure 3.5. A semi-empirical framework enables estimation of the equilibrium dissociation constant of monovalent bait-prey binding (K_D) using qYY2H. Titration curves describing the capture of prey cells as a function of the concentration of prey cells incubated were generated for a given concentration of magnetic bait cells. For each repeat, the concentration of prey cells captured is normalized by the estimated maximum concentration of prey cells captured, [*Max*]. Titration curves are shown for **(A)** NB.TOM22.1, **(B)** Sso7d.hFc, **(C)** Sso7d.TOM22.1, **(D)** Sso7d.his.hFc, and **(E)** Sso7d.ev.hFc. **(F)** A global, non-linear regression was used to estimate $K_{D,MV}$. Associated multiplicity constants, m, were calculated using the fitted $K_{D,MV}$ values and experimental K_D values. A plot of m vs. K_D is shown and fits a power law model. K_D values for these bait-prey interactions were previously estimated from soluble yeast surface titration (YST K_D). **(G)** A plot of normalized [*Max*] values vs. K_D fits a power law model. [*Max*] values for each repeat were normalized by the average number of prey cells captured across each concentration of prey cells incubated. **(H)** Fitted values of $K_{D,MV}$ and calculated values of m based on previous estimates of K_D (YST K_D) are shown. **(I)** The K_D of Sso7d.TOM22.2 and NB.TOM22.2 is 209 nM (193 – 225 nM, 68% confidence interval) and 1711 nM (1363 – 2155 nM), respectively, as estimated using the multivalent binding model (qYY2H K_D), compared to previously estimated K_D values of 362 nM (310 – 423 nM) and 1934 nM (1693 – 2209 nM) using soluble yeast surface titrations (YST K_D). Values in parentheses represent the bounds of a 68% confidence interval. Error bars represent the standard error of the mean for three repeats for all data shown.

The strength of the multivalent interaction between bait and prey yeast will likely depend on the cell surface density of bait and prey proteins. In our studies, a robust correlation between *m* and K_D was obtained (Figure 3.5F) even though the expression level of bait proteins on the yeast cell surface varied ~3 fold across our studies (Fig. 3.7). Therefore, the relationship between *m* and K_D described in Figure 3.5F – generated using bait-prey pairs with known K_D – may be used to obtain reasonable estimates of K_D for other bait-prey interactions, despite some variation in cell surface display levels. To further investigate the effect of bait surface density, we conducted titrations with three hFc binding prey proteins displayed on the yeast surface using hFc, immobilized on magnetic beads, as bait (Figure 3.8). The surface density of hFc is significantly higher on magnetic beads (>$2x10^5$ based on the manufacturer's capacity for IgG binding) compared to level of hFc expressed on the yeast surface (~$5x10^4$)[9]. We observed that $K_{D,MV}$ is significantly lower for yeast-bead interactions relative to yeast-yeast interactions (Figure 3.9), indicating that the multiplicity constant *m* is greater when the surface density of the bait protein is higher. This is consistent with previous studies which show that *m* for virus-cell interactions is higher for cases where the expression level of virus-binding receptors on the cell surface is greater[34]. Finally, it is important to note that the relationship between K_D and *m* in Figure 3.5F was established using a set of bait-prey pairs with K_D in the range 272-5280 nM. We anticipate that expansion of the "standard curve" in Figure 3.5F with additional bait-prey pairs that have K_D values outside this range will help refine fitted model parameter values, improve K_D estimates, and identify potential limitations of the proposed model.

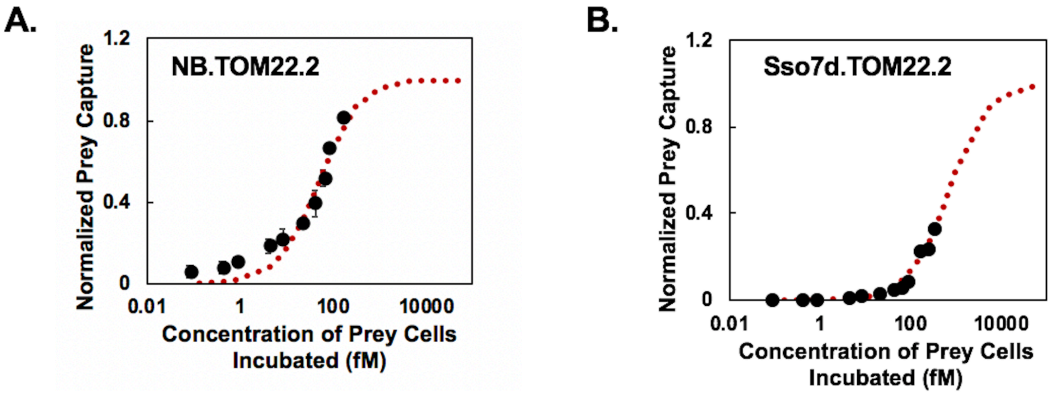


Figure 3.6. Binding of yeast expressing NB.TOM22.2 or Sso7d.TOM22.2 to TOM22 expressing yeast as assessed by qYY2H. Titration curves are presented describing the binding of prey cells expressing **(A)** NB.TOM22.2 or **(B)** Sso7d.TOM22.2 as a function of the concentration of prey cells incubated were generated for a given concentration of magnetic bait cells expressing TOM22. A global fit was used to predict $K_{D,MV}$ using the proposed non-linear, multivalent binding model as represented by the red line. Error bars represent the standard error of the mean for 3 or 4 repeats.

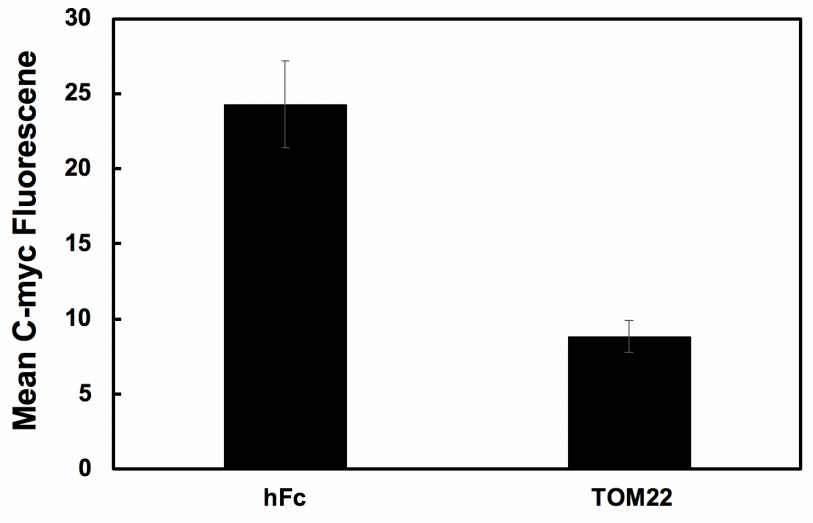
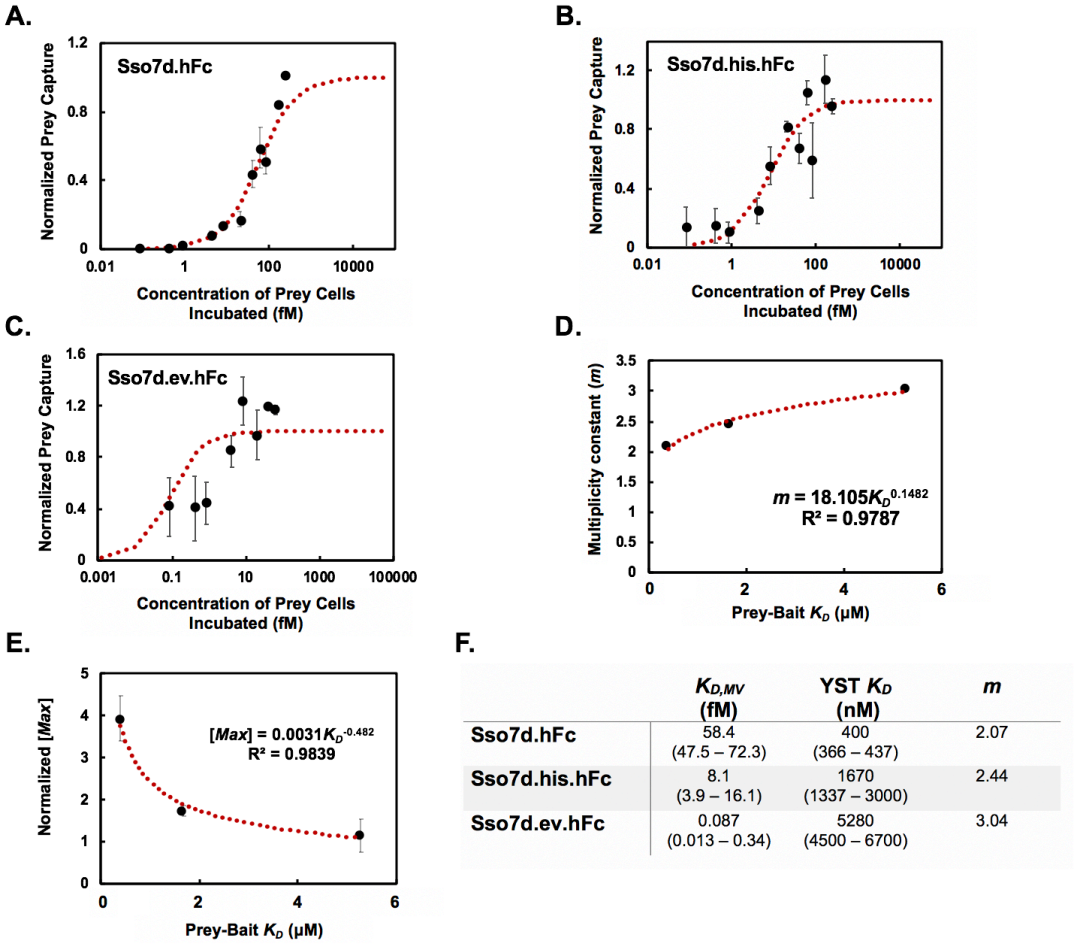


Figure 3.7. Comparison of bait expression levels between cells co-expressing either hFc and SsoFe2 or TOM22 and SsoFe2. The expression of the bait protein was quantified by labeling with an anti-c-myc antibody followed by detection with a goat-anti-chicken secondary antibody and flow cytometry analysis. The expression of the hFc bait on the surface of yeast is ~2.7 fold higher than the expression of the TOM22 bait on the yeast surface. Error bars represent the standard error of the mean for eight independent replicates.



F.

	$K_{D,MV}$ (fM)	YST K_D (nM)	m
Sso7d.hFc	58.4 (47.5 – 72.3)	400 (366 – 437)	2.07
Sso7d.his.hFc	8.1 (3.9 – 16.1)	1670 (1337 – 3000)	2.44
Sso7d.ev.hFc	0.087 (0.013 – 0.34)	5280 (4500 – 6700)	3.04

Figure 3.8. A semi-empirical framework enables the development of a relationship between monovalent K_D and $K_{D,MV}$ when the bait protein is immobilized on magnetic beads. Titration curves describing the binding of prey cells as a function of the concentration of prey cells incubated were generated for a given concentration of magnetic beads immobilized with the bait protein (biotinylated hFc). For each repeat, the concentration of prey cells captured is normalized by the estimated maximum concentration of prey cells captured, [*Max*]. Titration curves are shown for **(A)** Sso7d.hFc, **(B)** Sso7d.his.hFc, and **(C)** Sso7d.ev.hFc. **(D)** A global non-linear regression was used to estimate $K_{D,MV}$ as well as the multiplicity constant m. A plot of m vs. K_D is shown and fits a power law model. K_D values for these bait-prey interactions had been previously estimated from soluble yeast surface titration (YST K_D). **(E)** A plot of normalized [*Max*] values vs. K_D fits a power law model. [*Max*] values for each repeat were normalized by the average number of prey cells captured across each concentration of prey cells incubated. **(F)** Experimentally determined $K_{D,MV}$ and calculated values of m based on previous estimates of K_D (YST K_D) are shown . Values in parentheses represent the bounds of a 68% confidence interval. Error bars represent the standard error of the mean for three repeats for all data shown.

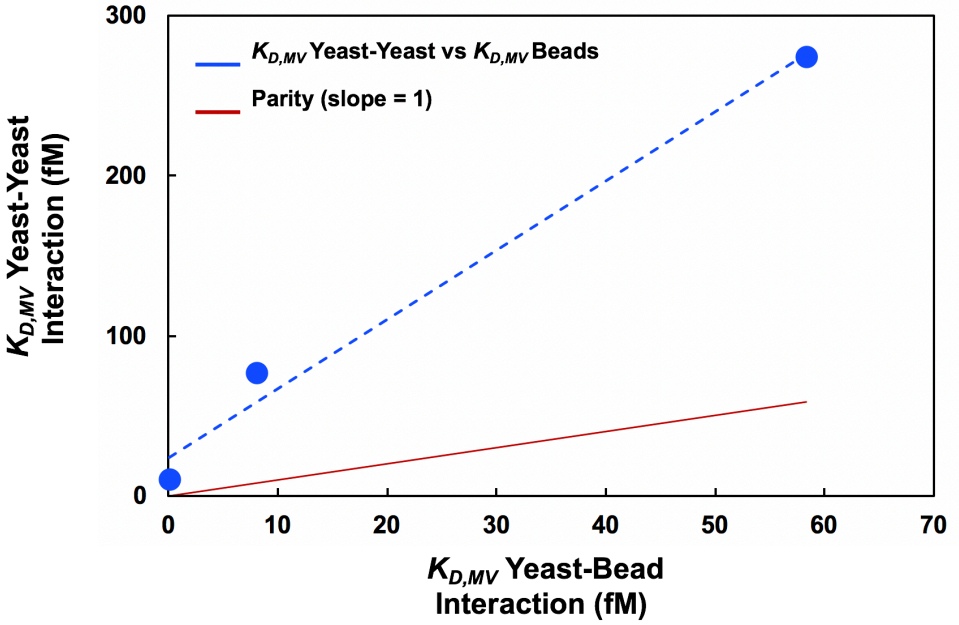


Figure 3.9. Comparison of $K_{D,MV}$ for yeast-yeast interactions and $K_{D,MV}$ for yeast-bead interactions. The interaction of prey cells displaying Sso7d.hFc, Sso7d.his.hFc, or Sso7d.ev.hFc with either magnetic cells expressing the hFc bait or magnetic beads functionalized with soluble hFc were considered (blue). Included in the figure is a parity line (red) that details the trend if the $K_{D,MV}$ for the yeast-yeast interaction was the same as the $K_{D,MV}$ for yeast-bead interaction. Accordingly, the $K_{D,MV}$ is significantly lower for yeast-bead interaction relative to yeast-yeast interaction.

3.2.4 qYY2H enables screening of cDNA libraries to identify putative PPIs and estimation of their binding affinities

Y2H is routinely utilized to identify putative interaction partners for a given bait protein by employing a cDNA-encoded protein library[35–37]. To assess if qYY2H can be used for discovery of PPIs, we screened a cDNA library to identify proteins that interact with SMAD3, a signaling molecule that transduces TGF-beta signals from the cell surface to the nucleus by acting as a transcriptional regulator for various genes[38]. Briefly, we constructed a yeast displayed cDNA library that expresses prey proteins as well as NanoLuc as cell surface fusions. The bait protein, SMAD3, was co-expressed with the iron oxide binding protein SsoFe2 to enable magnetization of bait yeast. Selections were conducted by incubating the yeast cells displaying prey proteins from the cDNA library with magnetized bait cells, followed by magnetic separation of bait-prey

complexes, and selective expansion of positively bound prey yeast. After two rounds of selection, plasmid DNA was isolated from individual prey yeast clones. DNA sequencing of thirty clones identified nine unique sequences corresponding to putative SMAD3-binding proteins (Table A.1). Note that we also recovered prey cells displaying a Sso7d mutant protein that arise from incomplete digestion of the plasmid vector used as the starting backbone to construct the cDNA library. A likely explanation for recovery of these cells is association between the Sso7d prey mutant and the SsoFe2 protein used for magnetizing bait cells. It is important to note that a 100X excess of non-displaying, EBY100 yeast cells, which are not expanded in selective yeast medium, are included in our screens to prevent non-specific stickiness between bait and prey yeast. The inadvertently included Sso7d-displaying prey cells further act as competitor for the cDNA clones, thereby increasing the likelihood that the putative cDNA prey isolated do not bind non-specifically to the bait protein.

We compared the interaction strength between SMAD3 and a subset of the identified putative binding partners, with that of the SMAD3-SARA binding interaction. SARA binds unphosphorylated SMAD3[39–41] and thus acts as a positive control. Specifically, we quantified the number of prey cells captured by magnetized yeast cells displaying SMAD3, relative to the number of cells captured displaying the SMAD binding domain of SARA (Figure 3.10A). Cells displaying a portion of the cohesin loading factor NIPBL protein were captured at the highest rate by the SMAD3 bait cells, followed by cells displaying fragments of bromodomain containing protein 7 (BRD7), gametogenetin binding protein 2 (GGNBP2), and endothelin-1 (END1). In comparison to cells displaying SARA, cells expressing each of these prey were captured at a higher level, suggesting that the binding strength of these protein fragments for SMAD3 is at least comparable to that of SARA. On the other hand, fewer cells displaying portions of the mitochondrial calcium

uniporter (MCU) protein and the dead box polypeptide 18 (DDX18) protein were recovered than cells displaying SARA, suggesting that these proteins may bind SMAD3 with lower affinity than SARA. Notably, BRD7 has been previously shown to interact with SMAD3[42]. Similarly, other RNA helicase dead box proteins, similar to DDX18, have been shown to complex with SMAD3[39,43]. However, two of the identified putative binding partners – MCU and endothelin 1 – are localized to the mitochondria and extracellular space, respectively[44,45]. Therefore, these binding interactions may not be biologically relevant. Additionally, we generated complete titration curves for yeast cells expressing the identified portions of BRD7 and GGNBP2 to estimate their binding affinities for SMAD3 using the multivalent binding model discussed earlier (Figure 3.10B-C). The K_D values associated with monovalent bait-prey binding were estimated as 365 nM (309 – 453 nM, 68% confidence interval) and 1165 nM (964 – 1408 nM, 68% confidence interval) for the BRD7 and GGNBP2 fragments respectively (Figure 3.10D).

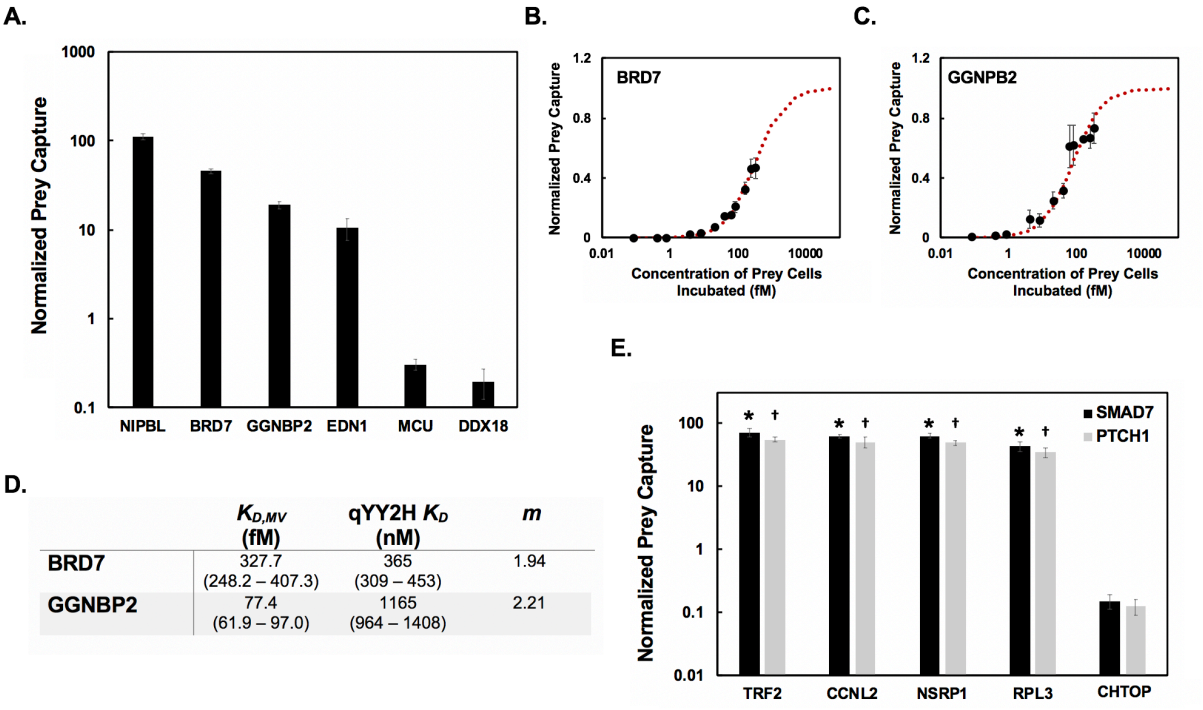
To further investigate the use of qYY2H in the context of identifying putative PPIs, we screened the cDNA library to identify interaction partners for the tandem WW domains of the YAP protein. The WW domain is a frequently occurring protein-interaction domain that mediates PPIs in signaling pathways by recognizing proline-rich motifs[46–49]. Besides proline-rich motifs, WW domains have also been shown to interact with phosphoserine-proline or phosphothreonine-proline motifs[50], LPxY motifs[51], and polyprolines flanked by arginine residues or interrupted by leucine residues[52–54]. Using similar methods as described for SMAD3, we identified sequences corresponding to 10 unique proteins that putatively interact with the WW domains of YAP after analyzing 30 clones by DNA sequencing (Table A.2). While we did not identify sequences containing the most common PPXY motif, many of the prey identified contained other WW binding motifs. For example, four of the isolated sequences contained a previously identified

arginine motif[52]. Five sequences contained no previously identified WW binding motifs, consistent with previous cDNA screens performed to identify WW-binding prey, where at least 20% of the isolated prey did not contain a known binding motif [55].

We further analyzed five of the isolated WW-domain binding sequences using qYY2H (Figure 3.10E) and compared the strength of their binding interaction with the bait to that of known WW domain binding peptides, SMAD7 and PTCH1. SMAD7 and PTCH1 both contain a PPPY motif, and each binds the WW domains of YAP with K_D = 8 μM and 24.7 μM, respectively[56,57]. Isolated prey containing fragments of the TATA-box binding protein associated factor 2 (TRF2), cyclin L2 (CCNL2), and nuclear speckle splicing regulatory protein 1 (NSRP1) contain an arginine-based binding motif. In contrast, some of the characterized prey, such as sequences from ribosomal protein L3 (RPL3) and chromatin target of PRMT1 protein (CHTOP), do not contain any known WW binding motifs. For the five prey considered, the number of prey cells captured by the WW domain bait cells was greater than the capture of the control cells expressing either SMAD7 or PTCH1, except for the CHTOP prey (Figure 3.10E); this suggests that the identified CHTOP sequence likely binds the WW domains of YAP with weaker affinity than the SMAD7 and PTCH1 controls, whereas all other sequences likely bind with similar or higher affinity than SMAD7 and PTCH1.

Collectively, these results show that qYY2H can be utilized for first discovering putative PPIs by screening cDNA libraries followed by quantitative assessment of the relative binding strengths associated with the identified putative interactions. Importantly, we unveiled a putative binding interaction between the WW domains of YAP and a fragment of CHTOP, with an affinity likely weaker than the YAP-PTCH1 interaction (24.7 μM). Despite the suggested weak affinity of CHTOP for the WW domains of YAP (<24.7 μM), the number of CHTOP-displaying cells

captured using the qYY2H assay is well above the detection limit of the assay, as assessed by the lowest number of luciferase-displaying cells quantified using a standard curve in this study. Thus, even weak affinity, putative binding interactions can be identified and quantitatively assessed with qYY2H.



	$K_{D,MV}$ (fM)	qYY2H K_D (nM)	m
BRD7	327.7 (248.2 – 407.3)	365 (309 – 453)	1.94
GGNBP2	77.4 (61.9 – 97.0)	1165 (964 – 1408)	2.21

Figure 3.10. Use of qYY2H to characterize bait-prey pairs identified from a cDNA library screen. For prey isolated from a cDNA library screen, cells co-displaying these prey protein sequences and NanoLuc were incubated with magnetic cells expressing the bait. The luminescence signal of the captured prey cells was used to numerate prey cell capture. **(A)** Prey cell recovery by magnetic bait cells expressing SMAD3 for prey protein sequences identified in cDNA library selections against SMAD3 using qYY2H. Cell capture data for each cDNA prey was normalized by the recovery of cells displaying the SMAD binding domain of SARA. Bait-prey, yeast-yeast titration curves were generated for two of the identified prey, BRD7 **(B)** and GGNBP2 **(C)**. **(D)** A global, non-linear regression was used to estimate $K_{D,MV}$. Using the previously established semi-empirical model, K_D values (qYY2H K_D) describing the binding strength of monovalent bait-prey interactions were estimated as well as associated multiplicity constants, m. Values in parentheses represent the bounds of a 68% confidence interval. **(E)** Prey cell recovery by magnetic bait cells expressing the WW domains of YAP using qYY2H. The prey protein sequences considered were identified in a cDNA library screen against the WW domains of YAP. Cell capture data for each cDNA prey was normalized by the recovery of cells displaying either SMAD7 (black) or PTCH1 (grey), known WW binding peptides that contain a PPPY motif. * represents $p < 0.05$ for a two tailed, paired t-test in comparison to the capture of SMAD7 cells while † represents $p < 0.05$ for a two tailed, paired t-test in comparison to the capture of PTCH1 cells. Error bars represent the standard error of the mean for three repeats for all data shown.

3.2.5 qYY2H enables evaluation of binding interactions mediated by post-translational modifications

Many processes that regulate biological function rely on protein domains that specifically bind post-translationally modified targets to relay signals[58]. For instance, the SH2 and PTB domains bind targets modified by phosphorylation[59]. Previous studies have demonstrated the identification of binding interactions dependent on post-translational modifications by linking the modifying enzyme to the bait protein when using the Y2H system[13,14] or through co-expression of the modifying enzyme in a bacterial two hybrid system[60]. We investigated if qYY2H can be extended to quantitatively assess binding interactions mediated by post-translational modifications by utilizing enzymatically modified bait proteins displayed on the yeast surface as previously described[61]. In this approach, the modifying enzyme and an Aga2-substrate fusion are co-expressed with ER retention tags to increase their ER residence times. The enzyme acts on the substrate within the ER prior to surface display of the modified Aga2-substrate fusion. Specifically, we evaluated the binding interactions mediated by a phosphopeptide (F1163 – F1183, containing pY1173) derived from epidermal growth factor receptor (EGFR). We used the Abelson tyrosine kinase, which can phosphorylate Y1173 in EGFR, as our modifying enzyme[61,62]. To afford expression of an enzymatically modified bait on the surface of magnetic yeast cells, we refashioned the yeast surface display plasmid to encode the modifying enzyme downstream of the Gal10 promoter (Figure 3.11). The Aga2 fusions containing the substrate peptide and the iron-oxide binding protein SsoFe2 are encoded downstream of a Gal1 promoter on the same plasmid. The Aga2-substrate peptide and the Aga2-SsoFe2 fusions are translated as separate proteins through the inclusion of a T2A ribosomal skipping peptide as we have previously described[19].

Flow cytometry analysis using an anti-phosphotyrosine antibody showed that the EGFR peptide containing Y1173 is phosphorylated only in the presence of Abelson kinase (Figure 3.12A).

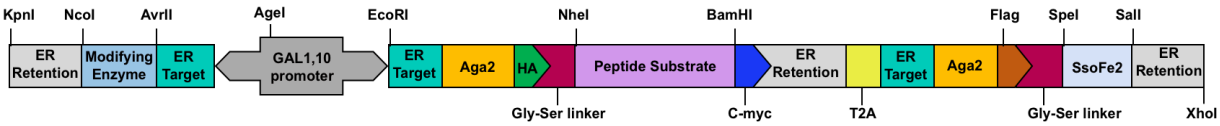


Figure 3.11. Plasmid design for co-expression of an enzyme modified bait peptide and SsoFe2, an iron oxide binding protein, as fusions to Aga2 for display on the yeast surface. The DNA encoding the peptide substrate and SsoFe2 is under the control of the Gal1 promoter and is separated by a T2A ribosomal skipping peptide, which affords separate translation of the fusion proteins. A bait-modifying enzyme is encoded downstream of the Gal10 promoter. Each protein is trafficked to the ER by an included N-terminal targeting sequence as well being retained in the ER for a longer than average residence time due to an included C-terminal ER retention sequence. This increases the probability that the modifying enzyme will have time to interact with the peptide substrate. Ultimately, the modified peptide substrates along with SsoFe2 are trafficked to the yeast cell surface for display.

To assess the use of qYY2H for studying phosphorylation-dependent binding, we first evaluated the interaction between the SH2 domains of the adapter protein APS and pY1173. Previous cDNA library screens identified that the SH2 domains of adapter protein APS bind to EGFR pY1173[63]. APS was shown to not interact with the non-phosphorylated EGFR peptide. Prey cells expressing the SH2 domains of APS and NanoLuc were incubated with magnetized bait cells expressing the phosphorylated EGFR peptide pY1173, obtained via enzymatic modification by Abelson kinase within the ER, or the non-phosphorylated peptide as a control. Subsequently, bait-prey complexes were separated using a magnet and the number of prey cells captured was quantified. A fraction of cells from the population that co-expresses Abelson kinase and the substrate peptide will display non-phosphorylated peptides. Hence, the cell capture data for the bait population co-expressing Abelson kinase was normalized by the fraction of cells expressing peptide fusions that are also phosphorylated as evaluated by immunofluorescent detection (Figure 3.12A-B). A greater number of APS displaying prey cells was captured when the phosphorylated

peptide was used as bait (Figure 3.12C), suggesting the qYY2H can be used to identify interactions that rely on post-translational modifications.

We further investigated if qYY2H could be used to quantitatively discriminate between proteins that bind a post-translationally modified target with varying interaction strengths. SH2 domains bind phosphotyrosine residues promiscuously via a consensus sequence pYXXP motif[64–67]. The tandem SH2 domains of phospholipase C–γ1 (PLCγ1), denoted tSH2, bind to phosphorylated EGFR. tSH2, as well as its C-terminal SH2 domain, denoted cSH2, have been evaluated as biosensors for live cell imaging of EGFR phosphorylation at Y992[68–70]. However, both tSH2 and cSH2 show promiscuous binding to other phosphorylated residues in EGFR[68,71]. To overcome the promiscuity of tSH2 and cSH2, Tiruthani et al. identified two proteins with enhanced specificity for pY992, mSH2 and SPY992, which were isolated by screening combinatorial libraries generated by mutagenesis of cSH2 and a Sso7d protein scaffold, respectively[68]. SPY992 and mSH2 exhibit high specificity of binding to pY992 over other phosphorylation sites in EGFR, relative to tSH2 and cSH2. We used qYY2H to investigate binding of tSH2, cSH2, mSH2, and SPY992 to pY1173. Consistent with lower binding of mSH2 and SPY992 to sites other than pY992, fewer prey cells expressing mSH2 and SPY992 were captured by bait expressing pY1173 than cells expressing tSH2 and cSH2 as prey (Figure 3.12D). These results show that qYY2H can be used to quantitatively compare binding interactions that are dependent on post-translational modification of the bait protein.

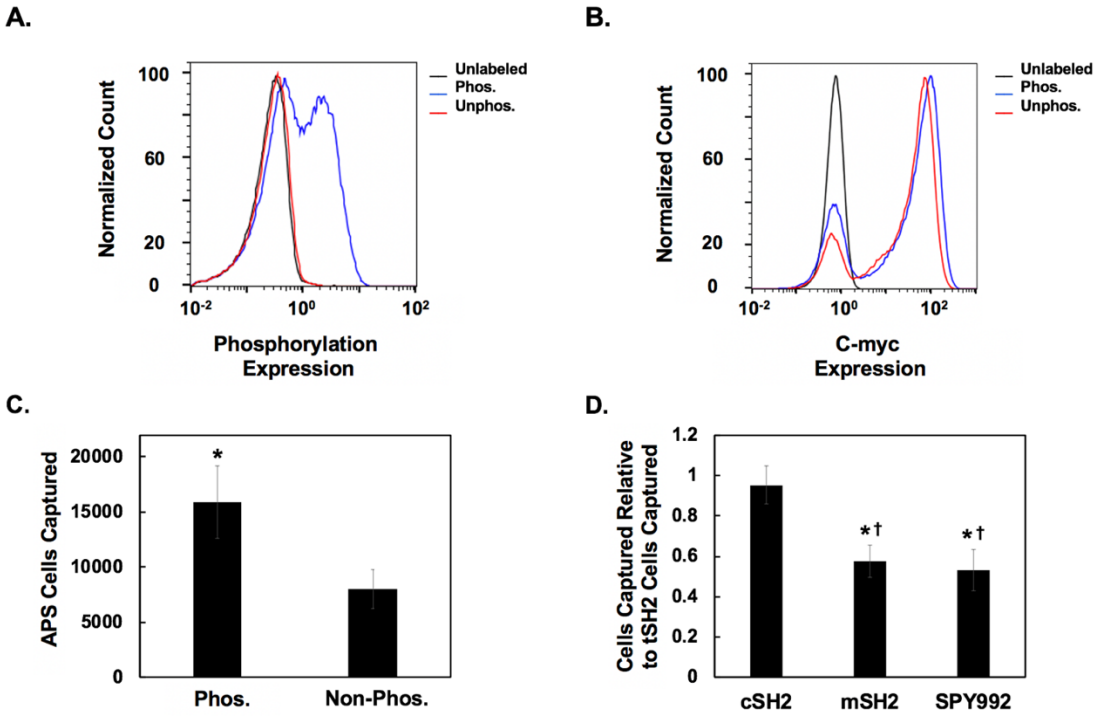


Figure 3.12. Use of qYY2H system to evaluate protein-protein interactions that depend on a post-translationally modified binding domain. **(A)** Yeast-cells expressing a phosphorylated EGFR peptide (pY1173) were used as bait. The EGFR peptide was enzymatically modified within the ER by Abelson tyrosine kinase prior to surface display. Flow cytometry analysis was used to evaluate the fraction of the bait population expressing the phosphorylated peptide above background. Yeast cells containing the modifying kinase (blue) as well as non-enzymatically modified yeast cells (red) were labeled with an anti-phosphotyrosine Alexa Fluor 647 antibody followed by flow cytometry detection. The fluorescence of unlabeled cells (black) is also considered. **(B)** Peptides may be displayed prior to modification by the kinase. Hence, a fraction of cells from the population that co-expresses Abelson kinase and the substrate peptide will display non-phosphorylated peptides. To quantify total peptide display, regardless of phosphorylation state, the cell population containing the modifying kinase (blue) as well as the non-enzymatically modified yeast (red) were labeled with an anti-c-myc antibody followed by detection using a goat-anti-chicken 488 antibody and flow cytometry analysis. The fluorescence of unlabeled cells (black) is also considered. Representative flow cytometry plots from three independent repeats are shown. **(C)** Yeast cells co-displaying the SH2 domains of adaptor protein APS and NanoLuc were incubated with magnetic bait cells expressing either an EGFR peptide phosphorylated at Y1173 or an unphosphorylated EGFR peptide. The cell capture values associated with the phosphorylated bait were normalized by the fraction of bait cells expressing c-myc fusions that were also phosphorylated. A greater number of APS prey cells were recovered by the phosphorylated bait cells than the non-phosphorylated bait cells. * represents $p < 0.05$ for a two-tailed, paired t-test in comparison to the non-phosphorylated bait. Error bars represent the standard error of the mean for seven repeats. **(D)** Yeast cells co-expressing tSH2, cSH2, mSH2, or SPY992 and NanoLuc were incubated with magnetic bait yeast expressing the pY1173 EGFR peptide. Cell capture values for each prey were first normalized by the fraction of bait cells expressing c-myc fusions that were also phosphorylated. After, the normalized cell capture values for each repeat were compared to the normalized capture of tSH2 cells by the phosphorylated bait cells. * represents $p< 0.05$ for a two-tailed, paired t-test in comparison to the capture of tSH2 cells while † represents $p< 0.05$ for a two-tailed, paired t-test in comparison to the capture of cSH2 cells. Error bars represent the standard error of the mean for five repeats.

3.3 Conclusions

In conclusion, we have developed a quantitative yeast-yeast two-hybrid system, denoted qYY2H, wherein the bait and prey proteins are expressed as yeast cell surface fusions. qYY2H can be used not only to screen cDNA libraries to identify PPIs, but also to quantitatively assess the strength of bait-prey binding interactions. Notably, PPIs dependent on post-translational modification of the bait can be analyzed using qYY2H. Due to the multivalent nature of the interaction between yeast displayed bait and prey, qYY2H can be used to investigate binding interactions with low binding affinities. In this study we have identified binding interactions with K_D ~ at least 25 μM, where the signal was well above the limit of detection of qYY2H; therefore, we expect that interactions with weaker affinities may be assessed. Inclusion of a large excess of competitor yeast (e.g. EBY100) mitigates non-specific stickiness of bait and prey yeast cells. Further, we have shown that the strength of the multivalent interaction between bait and prey yeast bears a quantitative relationship with the binding affinity of the monovalent bait-prey interaction (K_D). We have described a mathematical framework to exploit this relationship for quantitative estimates of monovalent bait-prey interaction K_D. We anticipate that qYY2H will be a powerful tool for quantitative analysis of protein-protein interactions. In particular, we expect that qYY2H will be very useful for efficient characterization of binding proteins or peptides isolated from combinatorial libraries[72]. Finally, the quantitative framework discussed herein for yeast-yeast binding interactions will be useful for designing combinatorial screens for isolating binders from yeast display libraries using whole cell targets[73] or when the target protein is expressed as a yeast cell surface fusion[19].

3.4 Materials and Methods

3.4.1 Yeast culturing and plasmids

Saccharomyces cerevisiae strain EBY100 cells containing the pCTCON plasmid were cultured using SDCAA (-Trp) media while cell surface protein expression was induced using SGCAA (-Trp) medium, as previously described[9]. Cells containing the pCT302 plasmid were cultured and induced with SDSCAA (-Leu) and SGSCAA (-Leu) media, which is similar in composition to SDCAA and SGCAA media except it contains a synthetic drop out mix (1.62 g/L; US Biological Life Sciences) lacking leucine instead of casamino acids. Cells were grown overnight in SDCAA or SDSCAA media at 30°C with shaking at 250 RPM. Protein expression was induced at 20°C with shaking at 250 rpm after transferring the yeast cells into SGCAA or SGSCAA medium at an OD of 1. EBY100 cells without plasmid were grown in YPD medium (10.0 g/L yeast extract, 20.0 g/L peptone, and 20.0 g/L dextrose) at 30°C with shaking at 250 rpm. Plasmid DNA was transformed using chemically competent EBY100 generated using the Frozen-EZ Yeast Transformation Kit II (Zymo Research).

3.4.2 Plasmid construction for dual display of prey and reporter enzyme

Plasmids were constructed to afford the co-expression of a prey protein and a reporter enzyme as yeast cell surface fusions. Previously described plasmids that confer dual protein expression via a ribosomal skipping T2A peptide were modified[19]. Specifically, the pCT302-SsoFe2-T2A-TOM22 plasmid was altered by digesting between the AvrII and NdeI sites. DNA encoding specific reporter enzymes, including glucose oxidase (GOx), horseradish peroxidase (HRP), and NanoLuc, were inserted between these sites to generate different dual enzyme display plasmids, pCT302-T2A-GOx, pCT302-T2A-HRP, and pCT302-T2A-NanoLuc, respectively. The GOx gene was amplified using primers Pf1 and Pr1 from genomic DNA extracted from *A.niger*

spores as described in Cruz et al[74]. DNA encoding horseradish peroxidase (Q1 – S308, assuming no signal peptide) was amplified from gene block 1 using primers Pf2 and Pr2[75]. Similarly, DNA encoding the NanoLuc protein was amplified using gene block 2 and primers Pf3 and Pr3[21]. Previously identified binding proteins evolved from the Sso7d scaffold were inserted into each of these plasmids via the NheI and BamHI sites. Sso7d.hFc, Sso7d.his.hFc, and Sso7d.ev.hFc were amplified using primers Pf4 and Pr4 from plasmids described in Gera et al[25]. Similarly, DNA encoding TOM22.Sso7d.1 and TOM22.Sso7d.2 were amplified using primers Pf4 and Pr4 from plasmids described in Bacon et al[19]. DNA encoding SSo7d.NTL.Lys and Sso7d.BVL.Lys was amplified from plasmids described by Carlin et al. with primers Pf4 and Pr4[24] while Sso7d.Lys.CTL was amplified with primers Pf5 and Pr4[24]. DNA encoding TOM22.NB.1 and TOM22.NB.2 was amplified with primers Pf6 and Pr6.

A plasmid, pcTCON-NanoLuc, was also constructed that solely expresses NanoLuc as an Aga2-yeast surface fusion. This plasmid was created by amplifying gene block 2 with primers Pf7 and Pr7 and inserting this product between the NheI and BamHI cut sites in pCTCON.

Double stranded gene fragments were purchased from Integrated DNA technologies (IDT). Primer oligonucleotides were purchased from IDT or Eton Biosciences. Gene fragment and primer sequences can be found in Tables 3.1 and 3.2. PCR reactions using Phusion Polymerase (Thermo Fisher Scientific) were performed in a 50 μL volume following the manufacturer's protocols. Restriction enzymes were purchased from New England Biolabs. Restriction digestions of plasmids and PCR products were performed for 4 hours at 37°C in 50 μL reactions with a 5X excess of each restriction enzyme. Digested plasmids were treated for 1 hour at 37°C with Antarctic phosphatase (New England Biolabs). Subsequently, digested products were purified using the 9K series gel and PCR extraction kit (BioBasic). Ligations of digested plasmids and PCR

product inserts were performed overnight using T4 DNA ligase (Promega) at 16°C. Ligations were transformed into electrocompetent Novablue *E.coli* cells. Each electroporation was performed at 1.6 KV, 25 μF, and 200 Ω. Plasmid was harvested from overnight *E.coli* cultures using the GeneJET plasmid miniprep kit (Thermo Fisher Scientific).

Table 3.1. List of gene fragments described in Chapter 3.

Gene Block 1	CAACTTACCCCTACCTTCTACGACAATTCATGTCCTAACGTCTCAA ACATAGTACGGGACACTATTGTCAATGAGTTACGATCGGACCCTA GAATCGCCGCGAGCATCCTTCGTCTTCACTTCCACGACTGCTTTGT TAACGGTTGTGACGCATCGATCTTGTTAGACAACACAACATCATT TCGAACAGAGAAAGATGCGTTTGGGAACGCAAACTCGGCGCGCG GATTTCCTGTGATTGACAGAATGAAGGCGGCCGTGGAGAGTGCAT GCCCAAGAACTGTATCCTGCGCAGATCTGCTTACCATTGCAGCTC AACAATCTGTTACTTTGGCAGGAGGTCCCTCTTGGAGGGTTCCTTT GGGACGTCGAGACAGCCTACAAGCATTTTTAGATCTCGCGAATGC GAATCTTCCAGCTCCATTCTTCACACTTCCACAACTTAAGGATTCT TTTAGAAATGTTGGTTTAAACCGTTCTTCTGATCTCGTTGCCCTCA GCGGTGGTCACACATTTGGTAAAAATCAATGTCGATTTATTATGG ACAGATTATACAACTTCAGCAACACCGGGTTACCCGACCCTACCC TCAACACTACTTACCTTCAAACTCTTCGTGGACTATGTCCCCTTAA TGGCAACCTAAGCGCCTTGGTCGATTTCGATCTGCGTACGCCAAC AATTTTCGATAACAAATACTATGTGAATCTTGAAGAGCAAAAAGG TCTCATCCAGAGTGATCAAGAGTTGTTCTCTAGCCCCAATGCCAC TGACACAATCCCACTGGTGAGATCATTTGCTAATAGCACACAAAC ATTCTTCAACGCGTTTGTGGAGGCCATGGATAGGATGGGAAACAT TACACCTCTTACAGGAACTCAAGGACAAATCAGGTTGAACTGTAG GGTGGTGAACTCCAACTCT

Table 3.1 (Continued).

Gene Block 2	ATGGTTTTCACCTTAGAGGATTTCGTAGGCGACTGGAGACAGACG GCAGGCTATAACTTGGACCAAGTGCTGGAACAGGGCGGTGTCAG CTCCTTATTCCAGAATCTGGGGGTATCCGTTACTCCTATACAGAG AATAGTTCTTTCCGGAGAAAACGGGCTAAAGATAGATATTCATGT CATTATTCCTTACGAAGGATTGTCCGGCGACCAAATGGGTCAAAT CGAAAAGATTTTCAAGGTGGTTTACCCGGTAGACGATCATCACTT CAAGGTGATTTTGCATTACGGTACATTAGTTATTGACGGGGTAAC GCCAAACATGATAGACTACTTTGGCAGACCCTACGAGGGAATTGC AGTTTTCGATGGTAAAAAGATCACGGTTACGGGTACTCTGTGGAA CGGGAATAAAATTATAGATGAGCGTTTAATCAATCCAGATGGATC ACTACTTTTTAGAGTGACAATTAACGGTGTCACTGGCTGGAGGTT ATGCGAAAGAATACTTGCG
Gene Block 3	TTCGAAATTCCCGACGATGTGCCGCTGCCAGCCGGTTGGGAGATG GCCAAAACTTCCTCAGGCCAGCGTTACTTCTTAAACCACATAGAT CAGACAACTACCTGGCAGGATCCGCGGAAAGCGATGTTAAGTCA AATGAATGTGACGGCACCTACATCCCCGCCAGTCCAGCAGAATAT GATGAACTCTGCTAGTGGTCCTCTTCCTGATGGGTGGGAACAGGC GATGACGCAAGACGGCGAAATTTATTACATAAACCACAAGAACA AAACAACCTCGTGGCTTGATCCACGGCTGGACCCACGCTTTGCGA TGAATCAG

Table 3.1 (Continued).

Gene Block 4	ATGTCTTCAATCCTTCCGTTTACTCCGCCGATCGTCAAGAGGCTAT TAGGCTGGAAAAAGGGCGAACAAAATGGGCAGGAAGAAAAGTG GTGTGAAAAAGCCGTCAAATCACTGGTGAAGAAACTGAAAAAAA CTGGCCAGCTAGATGAGTTGGAAAAGGCGATTACGACCCAGAAT GTAAACACTAAGTGCATCACCATTCCTAGATCTCTGGACGGCCGT CTACAGGTTTCTCACAGAAAGGGGCTACCGCATGTGATCTATTGT AGGCTTTGGCGTTGGCCTGATCTGCATTCCCACCATGAATTAAGG GCTATGGAGCTGTGTGAATTTGCATTCAACATGAAGAAGGATGAG GTGTGCGTTAACCCCTATCATTATCAAAGGGTCGAGACCCCCGTC CTGCCCCCCGTTCTGGTGCCACGTCACACTGAGATACCAGCGGAG TTCCCCCCGCTTGACGACTATTCCCACTCTATCCCAGAGAATACG AATTTTCCGGCTGGTATTGAGCCTCAAAGCAACATCCCAGAGACG CCGCCGCCTGGCTACCTGTCTGAAGATGGAGAGACTAGCGACCAC CAGATGAACCATAGTATGGACGCTGGCTCCCCCAATCTAAGCCCC AATCCAATGAGTCCAGCGCACAACAACCTGGATTTGCAGCCGGTC ACTTATTGTGAACCAGCATTTTGGTGTTCAATAAGCTACTATGAA CTTAACCAAAGGGTTGGAGAGACTTTCCATGCTTCCCAGCCATCC ATGACGGTTGATGGATTCACCGATCCCTCTAATTCAGAGAGGTTC TGTTTGGGGCTATTGTCCAACGTTAACCGTAACGCTGCGGTTGAG CTGACTCGTAGGCACATAGGCAGGGGCGTAAGACTTTATTATATT GGTGGGGAGGTGTTTGCTGAGTGTCTGTCTGACAGCGCAATTTTC GTCCAGAGCCCGAATTGCAATCAAAGGTATGGATGGCATCCAGC GACGGTCTGCAAGATCCCCCCTGGGTGTAACTTAAAGATCTTTAA TAATCAAGAATTTGCTGCACTTCTTGCTCAAAGTGTGAACCAAGG GTTTGAGGCAGTGTATCAATTAACCAGGATGTGTACGATCCGTAT GAGCTTCGTAAAGGGGTGGGGAGCAGAATACCGTAGGCAGACCG TTACCTCCACACCTTGTTGGATAGAGCTGCACTTAAACGGCCCTCT GCAATGGCTGGATAAAGTTTTAACACAAATGGGTTCACCTTCAAT TAGATGCTCTTCAGTTAGC

Table 3.1 (Continued).

Gene Block 5	GGTGGTTCAGCTAGCAGATATAGTCCACCTCCGCCGTACTCTTCTC ATAGTGGATCCGAACAAAAG
Gene Block 6	GGTGGTTCAGCTAGCGAATTGGAATCCCCGCCCCCGCCGTATTCC CGTTACCCTATGGACGGATCCGAACAAAAG
Gene Block 7	GCGCCCGACTCTCCTGATAATGATCTTCGTGCTGGACAGTTCGGC ATAAGTGCCAGGAAGCCATTTACGACCCTGGGCGAAGTCGCGCCC GTATGGGTTCCCGATTCTCAGGCTCCGAACTGTATGAAGTGCGAG GCGAGATTTACGTTTACGAAACGTAGACACCACTGC
Gene Block 8	GAAACTGACCCAGAAGCAGAACCGGAATTAGAGCTATCTGACTA TCCTTGGTTTCACGGGACATTGTCTAGGGTTAAGGCGGCCCAGTT GGTGCTGGCAGGTGGTCCTCGTAACCATGGTTTATTCGTTATCCGT CAAAGTGAAACTAGGCCAGGGGAGTACGTCCTAACTTTTAATTTC CAGGGGAAAGCGAAACACTTGCGTCTTTCTCTAAATGGTCACGGA CAGTGCCATGTCCAACACCTTTGGTTTCAATCTGTTCTTGACATGT TAAGGCACTTTCATACACATCCTATACCTCTGGAGTCCGGTGGTA GCGCCGATATTACACTTCGTAGCTACGTGCGTGCCCAGGACCCTC CTCCCGAACCAGGGCCAACC

Table 3.1 (Continued).

Gene Block 9	GGTACCCTATTATAATTCATCGTGTTCAAACCATGGAAACATGGT TTCGAACGCCTGGTGAATCTCAGCGAAGGAGGGTCTATCTGATGG ATTCCACTGCCAGCACGCACGCATTAATTCGTATACTTTTTCGGGA CAGCCCTCAGGACGTTCCATACGATAATCTTTTTCCAGTAGCTCGT ACACTTGAGACAGGTCTATCCCGGGATAAGGAGACATTCCGTAGG TGGCAATCTCCCATAGCAGCACTCCAAATGCCCAGACGTCTGACT TAATGCTGAACTTATTATAAGCTAGTGATTCCGGCGCTGTCCATTT AATGGGAAATTTTGCACCAGCATGGGCGGTATAGGTATCTCCAGT CATAAGCCTGGAAAGTCCGAAATCAGCCACCTTGACAAGGTGATT CTCACCAACTAAGCAGTTTCTGGCCGCTAAGTCTCTGTGTATGAA ATTTTTCTTTTCCAGATATTCCATTGCACTGCTTATCTGGGTAGCC ATGTAAAGTAAAACCACGGCATTGACTTCCTGACGATTACACTCT CTAAGATAGTCAAGCAAGTTGCCATATGTCATAAACTCTGTAATA ATATAAAACGGGGGCTCCCTAGTACATACCCCCAACAGCTGAACT AGATTCGGATGTTTAATTTCCTTCATAACAGCCGCTTCTTTCAAGA ACTCTTCAACTTCCATTGTATCTTCTTTCAACGTCTTTACGGCAAC CGTTAGGGAGTACTTCTTCCATACTCCTTCGTATACTTCGCCGTAC TGACCCCCTCCTAGTTTGTGTTTCATTGTAATGTCTGTACGTTCCA TCTCCCAGGCTAGTACGCTCGCGATTACGGAGAAGATACTAAAAC AACGTAGAAGCTGCATCCTAGGAATTCCTTGTTGAATTCGAATTT TCAAAAATTCTTACTTTTTTTTTGGATGGACGCAAAGAAGTTTAAT AATCATATTACATGGCATTACCACCATATACATATCCATATACAT ATCCATATCTAATCTTACTTATATGTTGTGGAAATGTAAAGAGCC CCATTATCTTAGCCTAAAAAAACCTTCTCTTTGGAACTTTCAGTAA TACGCTTAACTGCTCATTGCTATATTGAAGTACGGATTAGAAGCC GCCGAGCGGGTGACAGCCCTCCGAAGGAAGACTCTCCTCCGTGCG TCCTCGTCTTCACCGGT

Table 3.1 (Continued).

Gene Block 10	GAATTCATGCAATTGTTGAGATGTTTCTCTATCTTCTCTGTTATCG CTTCTGTTTTGGCTCAGGAACTGACAACTATATGCGAGCAAATCC CCTCACCAACTTTAGAATCGACGCCGTACTCTTTGTCAACGACTA CTATTTTGGCCAACGGGAAGGCAATGCAAGGAGTTTTTGAATATT ACAAATCAGTAACGTTTGTCAGTAATTGCGGTTCTCACCCCTCAA CAACTAGCAAAGGCAGCCCCATAAACACACAGTATGTTTTTTATC CGTACGACGTTCCAGACTACGCTGGTGGTGGTGGTTCCGGTGGTG GTGGTTCTGGTGGTGGTGGTTCAGCTAGCTTCAAGGGTTCTACTG CTGAAAACGCTGAATACTTGAGAGTTGCTCCACAATCTTCTGAAT TTGGATCCGAACAAAAGCTTATCTCCGAAGAAGACTTGTTCGAAC ACGACGAATTGGAGGGAAGGGGATCTTTGCTAACGTGCGGGGAT GTCGAAGAAAACCCTGGCCCTATGCAGTTGTTAAGATGCTTCTCT ATCTTTTCCGTTATCGCTTCTGTTTTAGCGCAGGAACTTACTACAA TTTGTGAACAAATACCTAGTCCCACGTTGGAGTCAACACCGTACT CCCTGTCTACTACAACAATACTTGCGAACGGTAAGGCCATGCAAG GAGTCTTTGAATACTACAAGTCAGTTACCTTCGTATCTAACTGCG GGAGTCACCCGAGCACCACATCCAAGGGGAGTCCTATAAACACT CAGTATGTATTCGACTATAAGGATGACGATGACAAGGGCGGCGG AGGCTCCGGTGGAGGCGGAAGCGGCGGTGGAGGTAGTACTAGTG CGACCGTGAAATTTAAATATAAAGGCGAAGAAAAACAGGTGGAT ATTAGCAAAATTAGACGGGTGCGTCGCAAGGGCAAATGCATTAG GTTTTACTATGATCTGGGCGGCGGCAAATATGGCAGGGGCATAGT GAGCGAAAAAGATGCGCCGAAAGAACTGCTGCAGATGCTGGAAA AACAGAAAAAAGTCGACTTTGAGCATGATGAGTTGTAATAGCTCG AG

Table 3.2. List of oligonucleotide primers described in Chapter 3.

Primer	Sequence
Pf1	GATTACCCTAGGATGCAGACTCTCCTTGTGAGCTCGCTT
Pr1	GAGCACCATATGCTGCATGGAAGCATAATCTTCCAAGATAG
Pf2	TCACATCCTAGGCAACTTACCCCTACCTTCTAC
Pr2	GACCGTCATATGAGAGTTGGAGTTCACCACCC
Pf3	TCGGCACCTAGGATGGTTTTCACCTTAGAGGATTTC
Pr3	TCGTCACATATGCGCAAGTATTCTTTCGCATAACC
Pf4	GAAGTCGCTAGCATGGCGACCGTGAAATTTAAATATAAAG
Pr4	GCACTTGGATCCTTTTTTCTGTTTTTCCAGCATCTGCAG
Pf5	GAAGTCGCTAGCATGGCGACCGTGAAATTAAAATAT
Pf6	GTCTCGGCTAGCGCACAAGTTCAGCTTGTAGAGT
Pr6	GTCACTGGATCCCGATGATACAGTTACTTGGGTAC
Pf7	TCGGCAGCTAGCATGGTTTTCACCTTAGAGGATTTC
Pr7	TCGTCAGGATCCCGCAAGTATTCTTTCGCATAACC
Pf8	TCCGGTGGTGGTGGTTCTGGTGGTGGTGGTTCAGCTAGCGCAGAG TGGCCATTACGGCC
Pr8	CTCCAAGTCTTCTTCGGAGATAAGCTTTTGTTCGGATCCGAGGCC GAGGCGGCC
Pf9	GGTGGTGGTGGTTCTGGTGGTGGTGGTTCAGCTAGCTTCGAAAT TCCCGACGATGTGC
Pr9	CAAGTCTTCTTCGGAGATAAGCTTTTGTTCGGATCCCTGATTCA TCGCAAAGCGTGG
Pf10	AGCACTGCTAGCATGTCTTCAATCCTTCCGTTTACTC
Pr10	ACGTGCGGATCCGCTAACTGAAGAGCATCTAATTGAA
Pf11	GGTGGTGGTGGTTCTGGTGGTGGTGGTTCAGCTAGCAGA

Table 3.2 (Continued).

Pr11	CAAGTCTTCTTCGGAGATAAGCTTTTGTTCGGATCCACTATG
Pf12	GGTGGTGGTGGTTCTGGTGGTGGTGGTTCAGCTAGCGAA
Pr12	CAAGTCTTCTTCGGAGATAAGCTTTTGTTCGGATCCGTCCAT
Pf13	GTTGGAGCTAGCGCGCCCGACTCTCCTGAT
Pr13	GTTGGAGGATCCGCAGTGGTGTCTACGTTTCG
Pf14	CGTCCGGCTAGCGAAACTGACCCAGAAGCAGAAC
Pr14	GGTCGTGGATCCGGTTGGCCCTGGTTCGGG
Pf15	GGTCGGGCTAGCGGACTCAGATCTAGTGGCAGC
Pr15	GGTCGTGGATCCTGCCCCATAATCGGGTTCAGC
Pf16	GGTTACGCTAGCCGCCTTTCAGAGCCTGTTCCA
Pf17	GCTATCGGTACCCTATTATAATTCATCGTGTTCAAACC
Pr17	GTACTGACCGGTGAAGACGAGGACGCACGGAGG
Pf18	CTGACGGAATTCATGCAATTGTTGAGATGT
Pr18	CTGATCCTCGAGCTATTACAACTCATC

3.4.3 Magnetization of bait expressing cells

Yeast cells co-expressing the bait and SsoFe2 were magnetized and blocked as previously described[19]. Briefly, $5x10^8$ cells co-expressing the bait protein and SsoFe2 were incubated with 2.6 mL of an iron oxide solution (4 mg/mL iron oxide in water) in a total volume of 13 mL using 1% PBSA (PBS pH 7.4, 1% BSA) for 20 minutes at room temperature. Non-magnetized cells were removed using a magnet. OD_{600} of the cells prior to the addition of the iron oxide (OD_i) and after the removal of yeast bound to iron oxide (OD_f) was measured using 100 µL of the sample. The number of cells bound to the iron oxide was calculated as (OD_i-OD_f). The magnetized cells were washed 3X with 1% PBSA and then aliquoted for use in each assay. The number of magnetized

cells aliquoted varied with the particular assay being performed. The aliquoted magnetized bait cells were then blocked for 1 hour using 2 mL of 1% PBSA and then washed two more times with 0.1% PBSA (PBS pH 7.4, 0.1% BSA).

3.4.4 Luciferase-based binding quantification assays using bait-expressing magnetic yeast

Briefly, $1x10^7$ magnetic bait cells were incubated with varying concentrations of yeast cells co-expressing the prey and NanoLuc in addition to $1x10^9$ EBY100 cells in 2 mL of 0.1% PBSAT (PBS pH 7.4, 0.1% BSA, 0.05% tween-20) for 1 hour. When comparing the binding of mutants for a signal prey concentration, only $1x10^7$ prey cells were incubated. When generating titration curves to estimate binding affinity values, the number of prey cells incubated ranged from $1x10^5$ - $4x10^8$ cells. After, any cells not bound to the magnetic bait cells were removed using a magnet. The magnetic bait cells and any bound prey cells were washed 3X with 0.1% PBSAT and then resuspended in 100 µL of PBS. The Nano-Glo Luciferase Assay system (Promega) was used to detect binding of prey cells to the immobilized bait proteins. 100 µL of the Nano-Glo reagent was added to the magnetic bead solution. The reaction proceeded for 3 minutes before placing the tube onto the magnet and plating 100 µL of the reaction in duplicate onto a 96 white-well plate with a clear bottom. The luminescence was read using a Tecan Infinite 200 plate reader using an integration time of 1000 ms, settle time of 0 ms, and no attenuation.

A calibration curve was made for each prey construct to develop a relationship between luminescence signal and the number of prey cells present. Luminescence assays were carried out as previously described to determine the points in the calibration curve. The curve was generated using $1x10^3$ - $1x10^6$ prey cells resuspended in 100 µL of PBS. The luminescence of just PBS and the Nano-Glo reagent was also measured as a blank. The calibration curves were generated by plotting background subtracted luminescence signals vs number of cells and fitting a linear

regression. Accordingly, the generated calibration curve was used to predict the exact number of prey cells captured by the magnetic bait cells by relating the luminescence signal produced by the captured prey cells.

3.4.5 Luciferase-based binding quantification assays using bait-functionalized magnetic beads

25 µL of Dynabeads Biotin Binder Beads (Thermo Fisher Scientific) were functionalized overnight with biotinylated, soluble bait protein (10.8 µM biotinylated lysozyme or 0.1 µM biotinylated IgG) using 0.1% PBSA. The following morning the beads were washed 3X with 0.1% PBSA followed by blocking in 1 mL of 1% PBSA for two hours. After, the beads were incubated with varying concentrations of yeast cells co-expressing the prey protein and NanoLuc along with $1x10^9$ EBY100 cells in 2 mL of 0.1% PBSAT for 1 hour at room temperature. After, any cells not bound to the bait-functionalized magnetic beads were removed using a magnet. The beads were washed 3X with 0.1% PBSAT and then resuspended in 100 µL of PBS. The Nano-Glo Luciferase Assay was performed as previously described to detect binding of the prey cells to the bait-functionalized magnetic beads. Subsequently, a luminescence calibration curve was used to quantify the number of prey cells captured.

The lysozyme protein used as the bait was biotinylated by incubating a 3:1 molar excess of Ez-Link^{TM}Sulfo-NHS-LC-Biotin (Thermo-Fisher Scientific) for two hours at 4°C. Excess reagent was removed through dialysis against 50 mM Tris-HCl, 300 mM NaCl pH 7.5. A BCA assay was used to estimate the protein concentration. Biotinylated IgG was purchased from Immunoreagents.

3.4.6 GOx-based binding quantification assays using bait-functionalized magnetic beads

Magnetic beads were functionalized with lysozyme by incubating 25 µL of Dynabeads Biotin Binder Beads (Thermo-fisher Scientific) with 10.8 µM biotinylated lysozyme overnight at 4°C using 0.1% PBSA. The beads were washed 2 times with 1% PBSA the next day and blocked with 1 mL of 1% PBSA for 3 hours at 4°C. Next, the magnetic beads were washed two times with 0.1% PBSA followed by incubation with $1x10^8$ cells co-expressing the prey protein and GOx along with $1x10^9$ EBY100 cells in 2 mL of 0.1% PBSA for 2 hours. The prey proteins considered were Sso7d.BVL.Lys, Sso7d.CTL.Lys, and Sso7d.NTL.Lys as well as an irrelevant protein, Sso7d.hFc. After, the incubations were placed onto a magnet to isolate any cells bound to the magnetic beads. The magnetic beads and any associated prey cells were then washed 3X with PBS. The beads were resuspended in 100 µL of PBS prior to adding 100 µL of the 2X TMB solution (0.1 mg/mL TMB, 0.4 M dextrose, 1.5 µg/mL HRP diluted in PBS). This reaction took place in the dark for 1 hour at room temperature before it was quenched with 50 µL of 2 M sulfuric acid. The reaction tubes were placed onto a magnet, and 100 µL aliquots were transferred onto a 96 well plate in duplicate. Each well's absorbance was measured at 450 nm using a plate reader.

The cell surface expression level of the co-displayed prey and GOx proteins was estimated using immunofluorescence detection. The expression of the prey protein was detected using a chicken anti-c-myc antibody (Thermo Fisher Scientific) while the expression of the GOx protein was detected using a mouse anti-V5 antibody (Thermo Fisher Scientific). To begin, $5x10^6$ cells were labeled with a 1:100 dilution of the primary antibodies for 15 minutes at room temperature. After, a secondary labeling was performed using a 1:250 dilution of goat-anti-chicken 488 or a donkey anti-mouse 633 antibody (Immunoreagents) for 10 minutes on ice. All labelings were

conducted in 50 µL of 0.1% PBSA. Washes took place between each labeling. The binding of the antibodies was detected using a Miltenyi Biotec MACsQuant VYB cytometer.

An activity control was also carried out for each prey construct. Briefly, $1x10^6$ cells co-expressing the prey and GOx were washed two times with PBS and ultimately resuspend in 100 µL of PBS prior to the addition of 100 µL of the 2X TMB solution. The control reactions were allowed to proceed for 5 minutes prior to being quenched with 50 µL of sulfuric acid. The quenched reactions were centrifuged. 100 µL aliquots of the supernatant plated in duplicate were read at 450 nm using a plate reader as previously described.

The background absorbance associated with an equal volume mixture of PBS and the TMB solution at 450 nm was subtracted from each raw, experimental absorbance value. For each prey, the background subtracted absorbance value associated with bait-prey binding was normalized by first dividing by the prey cells' mean fluorescence of c-myc expression and then dividing by the background subtracted absorbance of that particular prey's activity control. These normalizations were performed to account for differences in prey and GOx expression levels, respectively.

3.4.7 GOx-based binding quantification assays using bait expressing magnetic yeast

Yeast cells co-expressing hFc and SsoFe2 (pCT302-SsoFe2-T2A-hFc) were used as bait as previously described[19]. The bait cells were magnetized and blocked as previously described. However, tris based buffers (50 mM Tris, 300 mM NaCl pH 7.4) were used instead of PBS based buffers; This includes 1% TBSA (50 mM Tris, 300 mM NaCl, 1% BSA pH 7.4) and 0.1% TBSA (50 mM Tris, 300 mM NaCl, 0.1% BSA pH 7.4). $1x10^7$ magnetized target cells were incubated with $5x10^8$ yeast cells co-expressing the prey protein and GOx along with $1x10^9$ EBY100 for 2 hours at room temperature in 2 mL of 0.1% TBSA (50 mM Tris, 300 mM NaCl, 0.1% BSA pH 7.4). The incubations were then placed onto a magnet, and unbound cells were removed. The

magnetic bait cells and any bound prey cells were washed 3X with 0.1% PBSA prior to being resuspended in 100 μL of PBS. 100 μL of the 2X TMB solution was subsequently added. The reaction proceeded for 1 hour in the dark prior to quenching with 50 μL of sulfuric acid. The reactions were placed onto a magnet and a similar process, as previously described for the magnetic beads, was used to measure the absorbance of the reactions. Activity control assays and flow cytometry analysis of the prey cells were completed as previously described for normalization.

3.4.8 Semi-empirical framework to estimate K_D using qYY2H

Using the qYY2H system, titration curves were generated from bait-prey binding assays using luminescence quantification as described. $1x10^7$ magnetic bait cells were used for each assay while the number of the number of prey cells varied from $1x10^5$ cells - $4x10^8$ cells. Also included were $1x10^9$ EBY100 cells. The Nano-Glo assay described in the main text was used to detect binding of the prey cells to the magnetic bait cells. A luminescence standard curve was used to predict the number of prey cells captured as described in the main text. Three or four repeats were completed for each bait-prey pair. Prey considered with affinity for hFc included Sso7d.hFc, Sso7d.his.hFc, and Sso7d.ev.hFc[25]. Prey considered with affinity for TOM22 included Sso7d.TOM22.1, Sso7d.TOM22.2, NB.TOM22.1, and NB.TOM2.2[19]. Bait cells expressing hFc and SsoFe2 (pCT302-SsoFe2-T2A-hFc) as well as TOM22 and SsoFe2 (pCT302-SsoFe2-T2A-TOM22) were previously generated[19].

The number of prey cells incubated and subsequently captured were converted to a concentration using the following relationship:

$$[\text{Cell}] = \frac{\text{N}}{\text{N}_{\text{AV}}\text{V}} \text{ (Equation 3)}$$

Where [*Cell*] is the concentration of cells in moles, *N* is the number of cells, N_{AV} is Avogadro's number, and *V* is the volume the assay took place in (i.e.: 2 mL).

Curves were generated comparing the concentration of prey cells captured vs the concentration of prey cells initially incubated. Using Equation 1, the concentration of bait-prey, yeast-yeast complexes isolated by magnetic separation was related to the concentration of prey cells initially incubated through a $K_{D,MV}$ term that is similar, but not directly analogous, to the K_D parameter used when fitting a monovalent binding isotherm to yeast surface titrations that rely on soluble prey protein.

Typically, all data points in the titration curve were used when fitting the monovalent isotherm. However, some data points in the titration curves for individual repeats were not included when fitting the non-linear regression to estimate *[Max]* and $K_{D,MV}$ values due to non-specific binding or observation of the Hook effect. When evaluating non-specific binding, the slope between adjacent data points was calculated and plotted against the concentration of prey cells incubated. When generating this plot, the larger of the incubated prey cell concentrations, among the data points, was used as the x coordinate. For a given replicate, the first four slope points were ignored. When the slope reached ~0 (generally <0.005), we assumed that binding was reaching saturation. Sometimes multiple instances of a slope ~0 would occur for a particular plot. In these instances, we chose to not include data that occurred for concentrations higher than the second instance of slope ~0 to eliminate bias if the first minimum slope occurred due to error. For some prey, the slope increased, after reaching a minimum, therefore this data was not included when fitting the monovalent binding isotherm, as we assumed this increase in binding was due to non-specific binding. Data was also not included if the number of cells captured was significantly greater than the number of cells captured at the neighboring concentrations of incubated prey cells.

For yeast-yeast binding, Figures A.1 – A.7 in Appendix A describe the data used when fitting the monovalent binding isotherm to the generated titration curves for each repeat as well as the slope plots used to determine what data to include. For each bait-prey interaction, a global fit was used to estimate [*Max*] values for each repeat and a single $K_{D,MV}$ across all repeats by minimizing the sum of the squared error.

Using Equation 2, it was assumed that a quantitative relationship exists between the binding affinity of the monovalent bait-prey interaction (K_D) and $K_{D,MV.}$ For each bait-prey considered, estimates for K_D had been previously predicted by performing yeast surface titrations using soluble protein. We used the fitted values of $K_{D,MV}$ and the previously estimated K_D values to calculate *m* for five of the bait-prey pairs. We used a random number generator to choose which interaction prey to consider (TOM22.NB.1, Sso7d.hFc, Sso7d.TOM22.1, Sso7d.his.hFc, Sso7d.ev.hFc). After, we generated a plot of comparing *m* and K_D for the interaction pairs considered. This data fit a power law relationship where

$$\mathrm{m} = 10.277\mathrm{K_D}^{0.1125} \text{ (Equation 4)}$$

This relationship can be inserted into Equation 2 to predict $K_{D,\,MV}$ as a function of K_D as follows

$$\mathrm{K_{D,MV}} = (\mathrm{K_D})^{10.277*\mathrm{K_D}^{0.1125}} \text{ (Equation 5)}$$

We used estimates of $K_{D,MV}$ for the other considered prey, Sso7d.TOM22.2 and NB.TOM22.2, to predict values of K_D using Equation 5 to evaluate the accuracy of our developed semi-empirical model.

A relationship between [*Max*] and K_D was also established for the five bait-prey pairs randomly chosen above. For each repeat, the [*Max*] value estimated when applying the monovalent

binding isotherm to the titration curve data was normalized by dividing by the average concentration of prey cells captured for the data included when applying the non-linear regression model. If a particular piece of data was not included when fitting the monovalent binding isotherm, then it was not used to calculate this cell capture average. For a given interaction pair, normalized [*Max*] values were averaged across the repeats. A plot was generated of the averaged, normalized [*Max*] values vs K_D and a power law regression was applied. K_D values were previously estimated for these bait-prey pairs using yeast surface titrations of soluble protein.

A similar procedure was used to generate titration curves when soluble bait protein was immobilized onto magnetic beads. We only considered human IgG (hFc) as a bait when evaluating this system using the NanoLuc reporter, but it could be extended to other soluble bait proteins. First, 25 μL of Dynabeads Biotin Binder Beads were washed two times with 0.1% PBSA. After, the beads were incubated overnight with 100 μL of 0.1 μM biotinylated human IgG (Immunoreagents) diluted in 0.1% PBSA based on the manufacturer's recommendations. The next day the beads were washed 3X with 0.1% PBSA followed by blocking in 1 mL of 1% PBSA for two hours. After, the beads were washed an additional two times in 0.1% PBSA prior to incubation with the prey cells and $1x10^9$ EBY100 in 2 mL of 0.1% PBSAT. The number of prey cells incubated varied ranging from $1x10^5$ cells - $4x10^8$ cells. After separation of unbound cells and washing (3X with 0.1% PBSAT), the Nano-Glo assay was used to detect binding of the prey cells to the magnetic beads as described in the main text. Standard curves generated using known quantities of prey cells were used to estimate the number of prey cells captured. Three repeats were repeated for each prey. The monovalent binding isotherm as described in Equation 1 was fit to this data, in a similar manner as previously described, to estimate [*Max*] values for each repeat and a single $K_{D,MV}$ across all repeats for each prey. Previously described methods were used to determine

which points to include when applying the fit (Figures A.8-A.10). Plots of m vs K_D and normalized [Max] vs K_D were generated and fit using a power law regression as previously described (Figure 3.8).

3.4.9 Construction of a yeast displayed cDNA library with a luciferase reporter

The Clontech Mate & Plate Library- Universal Mouse (Normalized) cDNA library was obtained from Takara Bio Inc. The library was grown in SDSCAA (-Leu) media at 30°C. Plasmids were extracted from these cells using a Zymoprep Yeast Plasmid Miniprep Kit II (Zymo Research). Twenty four plasmid extractions using $9x10^6$ cells were performed. The prey DNA was amplified using primers Pf8 and Pr8 in 24 identical 50 µL PCR reactions to add DNA homologous with the pCTCON vector to the ends of the amplified prey DNA. PCR products were combined and purified using phenol:chloroform extraction. DNA was concentrated using ethanol precipitation where a 1/10 volume of potassium acetate was added to the extracted DNA along with 10 µL of linear acrylamide followed by 2 volumes of ice-cold ethanol. The mixture was incubated at -20°C overnight followed by centrifugation at 15,000 g for 10 minutes the next day. After, the pellet was washed once with 70% ethanol followed by a 100% ethanol wash and allowed to dry prior to resuspension in water. PCR amplification and the DNA purification/concentration steps were repeated to obtain enough DNA for library construction.

In parallel, the NanoLuc protein was cloned into a dual display T2A construct that uses pCTCON as its backbone to generate pCTCON-T2A-NanoLuc. A similar procedure was used as described for the pCT302-T2A-NanoLuc construct. However, the pCTCON-T2A-SSoFe2-hFc construct as previously described in Bacon et al. was digested[19]. 40 µg of pCTCON-T2A-NanoLuc was digested with NheI and BamHI and concentrated by phenol:chloroform extraction and ethanol precipitation.

The yeast-displayed cDNA library was generated by completing four electroporation reactions using a Bio-Rad Gene Pulser system (Bio-Rad) where 12 μg of amplified cDNA PCR product and 4 μg of digested vector was added to 400 μL of electrocompetent EBY100 cells. The electroporation settings used were 2.5 KV, 25 μF, and 250 Ω. An additional electroporation using only the digested vector was also performed. The number of transformed cells was estimated as $1x10^7$ as determined by plating serial dilutions of the combined transformation reactions onto SDCAA plates. However, the manufacturer of the cDNA library used for prey DNA amplification listed the diversity of their library as $6.64x10^6$.

3.4.10 Plasmid construction for dual display of bait protein and SsoFe2 as yeast surface fusions

The cDNA library screens were carried out against magnetic yeast co-displaying the bait protein and SsoFe2. To generate these cells, DNA encoding the bait proteins was cloned into the pCT302-SsoFe2-T2A-TOM22 plasmid between the NheI and BamHI sites as previously described[19]. This plasmid confers the concurrent display of a target protein of interest and SsoFe2 as Aga2-yeast surface fusions using a T2A ribosomal skipping peptide. The bait proteins cloned into this construct include the WW domain of YAP (F164 – Q270) and SMAD3 by amplifying from gene block 3 using primers Pf9 and Pr9 and gene block 4 using Pf10 and Pr10, respectively. The WW domains of YAP were cloned using homologous recombination while SMAD3 was cloned via restriction digest using the NheI and BamHI sites. For homologous recombination, 200 ng of cut plasmid and 250 ng of insert were transformed into competent EBY100 cells generated using the Frozen-EZ Yeast Transformation Kit II. The plasmid and insert were concentrated using ethanol precipitation to ensure the total volume of DNA used in the transformation was less than 5 μL.

3.4.11 Screening a yeast-displayed cDNA library with a luciferase reporter to discover putative PPIs

Two rounds of magnetic sorting were carried out to isolate putative binders to SMAD3 and the WW domains of YAP. In these sorts, the bait protein was displayed on the surface of magnetic yeast cells. For the first screening round, $5x10^7$ magnetic bait cells were prepared as previously described followed by incubation with $1x10^8$ induced cDNA, NanoLuc library cells and $1x10^9$ EBY100 cells in 2 mL of 0.1% PBSAT for 2 hours at room temperature. Any library cells bound to the magnetic bait cells were isolated using a magnet followed by washing with 0.1% PBSAT (5X) prior to expansion in 20 mL of SDCAA (-TRP) media.

In the second magnetic screening round, $5x10^6$ magnetic bait cells were incubated with $1x10^7$ library cells from the first screening round and $1x10^9$ EBY100 cells. The library cells bound to the magnetic bait cells were expanded in 5 mL of SDCAA (-TRP) media after washing. DNA was recovered from 30 individual clones using a Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research) and sequenced. Additional details can be found in the supplemental methods.

3.4.12 Cloning to perform qYY2H assays on identified cDNA bait-prey pairs

Some clones obtained from the cDNA screens were initially out of frame with the luciferase DNA due to the length of their poly A tail. For these constructs, the protein encoded by the cDNA was expressed correctly on the yeast surface, but luciferase was not displayed. Cloning was carried in a similar fashion as previously described using primers that changed the length of each clone's poly A tail. DNA encoding the in frame clones was inserted between the NheI and BamHI sites of pCTCON-T2A-NanoLuc.

All positive control peptides and proteins were cloned into the pCT302-T2A-NanoLuc construct via the NheI and BamHI sites as previously described. PTCH was amplified using gene

block 5 with primers Pf11 and Pr11 while SMAD7 was amplified using gene block 6 with primers Pf12 and Pr12. The inserts were incorporated into the plasmid backbone using homologous recombination as previously described. The SMAD binding domain of SARA (A664 – C720) was amplified from gene block 7 using primers Pf13 and Pr13 and inserted into the plasmid backbone using restriction digest cloning via the NheI and BamHI sites.

3.4.13 qYY2H assays to analyze identified cDNA bait-prey pairs

A similar procedure as the other qYY2H assays was performed where $1x10^7$ magnetic bait cells (SMAD3 or WW domains of YAP) were incubated with $1x10^7$ prey cells co-expressing NanoLuc for 1 hour at room temperature along with $1x10^9$ EBY100. Complete titration curves were generated for prey cells expressing BRD7 and GGNBP2 binding to magnetic bait cells expressing SMAD3 as previously described. The individual titration curves from each repeat are described in Figures A.11-A.12 along with details on which points were included when fitting the monovalent binding isotherm to predict $K_{D,MV}$. Equation 5 was used to estimate K_D values for these prey.

3.4.14 Cloning for analysis of bait-prey pairs that depend on post-translational modifications

All prey proteins were cloned into pCTCON-T2A-NanoLuc construct via the NheI and BamHI sites as previously described. The SH2 domains of APS (E400 – T526) were amplified using gene block 8 with primers Pf14 and Pr14[63]. tSH2 was amplified using primers Pf15 and Pr15 while cSH2 and mSH2 were amplified using primers Pf16 and Pr15. SPY992 was amplified using primers Pf4 and Pr4. The DNA encoding tSH2, cSH2, mSH2, and SPY992 was amplified from plasmids described in Tiruthani et al [68].

The plasmid conferring expression of an EGFR peptide (F1163 – F1183, assuming no signal peptide) phosphorylated at the 1173 tyrosine was cloned in multiple steps. First, gene block 9 was amplified with primers Pf17 and Pr17. This DNA piece was inserted into the pCTCON plasmid via the AgeI and KpnI sites to generate pCTCON-AblTK. This part of the plasmid affords expression of the Abelson tyrosine kinase (AblTK) in the endoplasmic reticulum under the direction of the Gal10 promoter. Next, gene block 10 was amplified with primers Pf18 and Pr18 and inserted into the pCTCON-AblTK plasmid via the EcoRI and XhoI sites to generate pCTCON-AblTK-EGFR-T2A. Gene block 10 encodes the EGFR peptide substrate to be modified by AblTK as an Aga2 fusion as well as an Aga2-SsoFe2 fusion both under the direction of the Gal1 promoter. The DNA encoding the EGFR peptide substrate and SsoFe2 fusions are separated by a T2A ribosomal skipping peptide sequence that affords separate products of these two fusions after translation. N-terminal ER targeting tags are included for all expressed proteins as well as C-terminal ER retention tags to increase the residence time of the proteins within the ER. The final plasmid construct is described in Figure 3.11.

In a similar manner, pCTCON-EGFR-T2A was generated by only inserting gene block 10 between the EcoRI and XhoI sites of pCTCON. This plasmid encodes a non-phosphorylated version of the described EGFR peptide as well as SsoFe2 as Aga2-yeast surface fusions.

3.4.15 qYY2H assays for binding analysis of bait-prey pairs that depend on post-translational modifications

Yeast cells containing the pCTCON-AblTK-EGFR-T2A plasmid were induced for protein expression for at least 24 hours at 20°C. These bait cells were magnetized as previously described. A similar procedure as the other qYY2H assays was performed where $1x10^7$ magnetic bait cells

were incubated with $1x10^7$ prey cells co-expressing NanoLuc for 1 hour at room temperature along with $1x10^9$ EBY100.

A fraction of cells from the population that co-expresses Abelson kinase and the substrate peptide (pCTCON-AblTK-EGFR-T2A) will display non-phosphorylated peptides. Accordingly, flow cytometry analysis was used to estimate the percentage of peptide displaying cells that are express phosphorylated versions of the peptide. Peptide display, irrespective of phosphorylation state, was analyzed by labeling with an anti-c-myc antibody. Display of the phosphorylated peptide was analyzed by labeling with an anti-phosphotyrosine antibody. Briefly, $2x10^6$ cells were labeled with a 1:100 dilution of chicken anti-c-myc antibody (Thermo Fisher Scientific) or a 1:50 dilution of Alexa Fluor 647 anti-phosphotyrosine (Biolegend) for 15 minutes at room temperature. Subsequently, cells were washed and secondary labeling was carried out for the cells labeled with the anti-c-myc antibody using a 1:250 dilution of goat-anti-chicken 488 (Immunoreagents) for 10 minutes on ice. All labeling were conducted in 50 μL of 0.1% PBSA, and washes with 0.1% PBSA took place after each labeling. Cells were analyzed using a Miltenyi Biotec MACsQuant VYB cytometer. The cells were not double labeled for c-myc and phosphorylation expression as it was observed that the binding of the anti-phosphotyrosine antibody decreased during double labeling due to steric hindrance. Similar analysis was completed for cells containing the pCTCON-EGFR-T2A plasmid (no modifying enzyme) to quantify their c-myc expression. For the pCTCON-AblTK-EGFR-T2A cells, 75.2% of the total cell population expressed a peptide while only 37.6% of the total cell population expressed a phosphorylated version of the peptide (Figures 3.12A-B). Therefore, only around ~50% of the peptide-displaying pCTCON-AblTK-EGFR-T2A cells are actually displaying a phosphorylated version of the peptide.

Prey cell recovery by the phosphorylated bait cell population was normalized by the percentage of peptide-displaying cells expressing a phosphorylated version of the peptide. This normalization took place as it is likely a higher number of prey cells could have been recovered if all cells within the bait population displayed a phosphorylated version of the EGFR peptide.

3.5 References

(1) Mehla, J.; Caufield, J. H.; Sakhawalkar, N.; Uetz, P. A Comparison of Two-Hybrid Approaches for Detecting Protein-Protein Interactions. *Methods Enzymol.* **2017**, *586*, 333–358.

(2) Keskin, O.; Tuncbag, N.; Gursoy, A. Predicting Protein–Protein Interactions from the Molecular to the Proteome Level. *Chem. Rev.* **2016**, *116* (8), 4884–4909.

(3) Barabási, A. L.; Oltvai, Z. N. Network Biology: Understanding the Cell's Functional Organization. *Nat. Rev. Genet.* **2004**, *5* (2), 101–113.

(4) Moal, I. H.; Agius, R.; Bates, P. A. Protein-Protein Binding Affinity Prediction on a Diverse Set of Structures. *Bioinformatics* **2011**, *27* (21), 3002–3009.

(5) Buntru, A.; Trepte, P.; Klockmeier, K.; Schnoegl, S.; Wanker, E. E. Current Approaches toward Quantitative Mapping of the Interactome. *Front. Genet.* **2016**, *7* (MAY), 74.

(6) Perkins, J. R.; Diboun, I.; Dessailly, B. H.; Lees, J. G.; Orengo, C. Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. *Structure* **2010**, *18* (10), 1233–1243.

(7) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science.* **1985**, *228* (4705), 1315–1317.

(8) Scott, J. K.; Smith, G. P. Searching for Peptide Ligands with an Epitope Library. *Science.* **1990**, *249* (4967), 386–390.

(9) Gera, N.; Hussain, M.; Rao, B. M. Protein Selection Using Yeast Surface Display. *Methods* **2013**, *60* (1), 15–26.

(10) Boder, E. T.; Wittrup, K. D. Yeast Surface Display for Screening Combinatorial Polypeptide Libraries. *Nat. Biotechnol.* **1997**, *15* (6), 553–557.

(11) Stynen, B.; Tournu, H.; Tavernier, J.; Van Dijck, P. Diversity in Genetic in Vivo Methods for Protein-Protein Interaction Studies: From the Yeast Two-Hybrid System to the Mammalian Split-Luciferase System. *Microbiol. Mol. Biol. Rev.* **2012**, *76* (2), 331–382.

(12) Fields, S.; Song, O. A Novel Genetic System to Detect Protein–Protein Interactions. *Nature* **1989**, *340* (6230), 245–246.

(13) Guo, D.; Hazbun, T. R.; Xu, X. J.; Ng, S. L.; Fields, S.; Kuo, M. H. A Tethered Catalysis, Two-Hybrid System to Identify Protein-Protein Interactions Requiring Post-Translational Modifications. *Nat. Biotechnol.* **2004**, *22* (7), 888–892.

(14) Spektor, T. M.; Rice, J. C. Identification and Characterization of Posttranslational Modification-Specific Binding Proteins in Vivo by Mammalian Tethered Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (35), 14808–14813.

(15) Hu, X.; Kang, S.; Chen, X.; Shoemaker, C. B.; Jin, M. M. Yeast Surface Two-Hybrid for Quantitative in Vivo Detection of Protein-Protein Interactions via the Secretory Pathway. *J. Biol. Chem.* **2009**, *284* (24), 16369–16376.

(16) Younger, D.; Berger, S.; Baker, D.; Klavins, E. High-Throughput Characterization of Protein–Protein Interactions by Reprogramming Yeast Mating. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (46), 12166–12171.

(17) Miura, K. An Overview of Current Methods to Confirm Protein- Protein Interactions. *Protein Pept. Lett.* **2018**, *25* (8), 728–733.

(18) Kastritis, P. L.; Bonvin, A. M. J. J. On the Binding Affinity of Macromolecular Interactions: Daring to Ask Why Proteins Interact. *Journal of the Royal Society Interface*. Royal Society February 6, 2013.

(19) Bacon, K.; Burroughs, M.; Blain, A.; Menegatti, S.; Rao, B. M. Screening Yeast Display Libraries against Magnetized Yeast Cell Targets Enables Efficient Isolation of Membrane Protein Binders. *ACS Comb. Sci.* **2019**, *21* (12), 817–832.

(20) Cruz-Teran, C. A.; Bacon, K.; McArthur, N.; Rao, B. M. An Engineered Sso7d Variant Enables Efficient Magnetization of Yeast Cells. *ACS Comb. Sci.* **2018**, *20* (10), 579–584.

(21) Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; et al. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem. Biol.* **2012**, *7* (11), 1848–1857.

(22) England, C. G.; Ehlerding, E. B.; Cai, W. NanoLuc: A Small Luciferase Is Brightening Up the Field of Bioluminescence. *Bioconjug. Chem.* **2016**, *27* (5), 1175–1187.

(23) Ackerman, M.; Levary, D.; Tobon, G.; Hackel, B.; Orcutt, K. D.; Wittrup, K. D. Highly Avid Magnetic Bead Capture: An Efficient Selection Method for de Novo Protein Engineering Utilizing Yeast Surface Display. *Biotechnol. Prog.* **2009**, *25* (3), 774–783.

(24) Carlin, K. B.; Cruz-Teran, C. A.; Kumar, J. P.; Gomes, C.; Rao, B. M. Combinatorial Pairwise Assembly Efficiently Generates High Affinity Binders and Enables a "Mix-and-Read" Detection Scheme. *ACS Synth. Biol.* **2016**, *5* (12), 1348–1354.

(25) Gera, N.; Hill, A. B.; White, D. P.; Carbonell, R. G.; Rao, B. M. Design of PH Sensitive Binding Proteins from the Hyperthermophilic Sso7d Scaffold. *PLoS One* **2012**, *7* (11), e48928.

(26) Frey, A.; Meckelein, B.; Externest, D.; Schmidt, M. A. A Stable and Highly Sensitive 3,3′,5,5′-Tetramethylbenzidine-Based Substrate Reagent for Enzyme-Linked Immunosorbent Assays. *J. Immunol. Methods* **2000**, *233* (1–2), 47–56.

(27) Zhang, S.; Zhao, X.; Niu, H.; Shi, Y.; Cai, Y.; Jiang, G. Superparamagnetic Fe3O4 Nanoparticles as Catalysts for the Catalytic Oxidation of Phenolic and Aniline Compounds. *J. Hazard. Mater.* **2009**, *167* (1–3), 560–566.

(28) Sun, S. P.; Lemley, A. T. P-Nitrophenol Degradation by a Heterogeneous Fenton-like Reaction on Nano-Magnetite: Process Optimization, Kinetics, and Degradation Pathways. *J. Mol. Catal. A Chem.* **2011**, *349* (1–2), 71–79.

(29) Pereira, M. C.; Oliveira, L. C. A.; Murad, E. Iron Oxide Catalysts: Fenton and Fentonlike Reactions – a Review. *Clay Miner.* **2012**, *47* (3), 285–302.

(30) Mahdizadeh, F.; Karimi, A.; Ranjbarian, L. Immobilization of Glucose Oxidase on Synthesized Superparamagnetic Fe 3 O 4 Nanoparticles ; Application for Water Deox-Ygenation. *Int. J. Sci. Eng. Res.* **2012**, *3* (5), 3–8.

(31) Mohamed, S. A.; Al-Harbi, M. H.; Almulaiky, Y. Q.; Ibrahim, I. H.; El-Shishtawy, R. M. Immobilization of Horseradish Peroxidase on Fe3O4 Magnetic Nanoparticles. *Electron. J. Biotechnol.* **2017**, *27*, 84–90.

(32) Yang, Q.; Liang, J.; Han, H. Probing the Interaction of Magnetic Iron Oxide Nanoparticles with Bovine Serum Albumin by Spectroscopic Techniques. *J. Phys. Chem. B* **2009**, *113* (30), 10454–10458.

(33) Mahmoudi, M.; Shokrgozar, M. A.; Sardari, S.; Moghadam, M. K.; Vali, H.; Laurent, S.; Stroeve, P. Irreversible Changes in Protein Conformation Due to Interaction with Superparamagnetic Iron Oxide Nanoparticles. *Nanoscale* **2011**, *3* (3), 1127–1138.

(34) Xiong, X.; Coombs, P. J.; Martin, S. R.; Liu, J.; Xiao, H.; McCauley, J. W.; Locher, K.; Walker, P. A.; Collins, P. J.; Kawaoka, Y.; et al. Receptor Binding by a Ferret-Transmissible H5 Avian Influenza Virus. *Nature* **2013**, *497* (7449), 392–396.

(35) Hannon, G. J.; Demetrick, D.; Beach, D. Isolation of the Rb-Related P130 through Its Interaction with CDK2 and Cyclins. *Genes Dev.* **1993**, *7* (12 A), 2378–2391.

(36) Durfee, T.; Becherer, K.; Chen, P. L.; Yeh, S. H.; Yang, Y.; Kilburn, A. E.; Lee, W. H.; Elledge, S. J. The Retinoblastoma Protein Associates with the Protein Phosphatase Type 1 Catalytic Subunit. *Genes Dev.* **1993**, *7* (4), 555–569.

(37) Zheng, Y.; Tan, X.; Pyczek, J.; Nolte, J.; Pantakani, D. V. K.; Engel, W. Generation and Characterization of Yeast Two-Hybrid CDNA Libraries Derived from Two Distinct Mouse Pluripotent Cell Types. *Mol. Biotechnol.* **2013**, *54* (2), 228–237.

(38) Zi, Z.; Chapnick, D. A.; Liu, X. Dynamics of TGF-β/Smad Signaling. *FEBS Lett.* **2012**, *586* (14), 1921–1928.

(39) Brown, K. A.; Ham, A.-J. L.; Clark, C. N.; Meller, N.; Law, B. K.; Chytil, A.; Cheng, N.; Pietenpol, J. A.; Moses, H. L. Identification of Novel Smad2 and Smad3 Associated Proteins in Response to TGF-Beta1. *J. Cell. Biochem.* **2008**, *105* (2), 596–611.

(40) Tsukazaki, T.; Chiang, T. A.; Davison, A. F.; Attisano, L.; Wrana, J. L. SARA, a FYVE Domain Protein That Recruits Smad2 to the TGFβ Receptor. *Cell* **1998**, *95* (6), 779–791.

(41) Qin, B. Y.; Lam, S. S.; Correia, J. J.; Lin, K. Smad3 Allostery Links TGF-Beta Receptor Kinase Activation to Transcriptional Control. *Genes Dev.* **2002**, *16* (15), 1950–1963.

(42) Liu, T.; Zhao, M.; Liu, J.; He, Z.; Zhang, Y.; You, H.; Huang, J.; Lin, X.; Feng, X.-H. Tumor Suppressor Bromodomain-Containing Protein 7 Cooperates with Smads to Promote Transforming Growth Factor-β Responses. *Oncogene* **2017**, *36* (3), 362–372.

(43) Warner, D. R.; Bhattacherjee, V.; Yin, X.; Singh, S.; Mukhopadhyay, P.; Pisano, M. M.; Greene, R. M. Functional Interaction between Smad, CREB Binding Protein, and P68 RNA Helicase. *Biochem. Biophys. Res. Commun.* **2004**, *324* (1), 70–76.

(44) Baughman, J. M.; Perocchi, F.; Girgis, H. S.; Plovanich, M.; Belcher-Timme, C. A.; Sancak, Y.; Bao, X. R.; Strittmatter, L.; Goldberger, O.; Bogorad, R. L.; et al. Integrative Genomics Identifies MCU as an Essential Component of the Mitochondrial Calcium Uniporter. *Nature* **2011**, *476* (7360), 341–345.

(45) Davenport, A. P.; Hyndman, K. A.; Dhaun, N.; Southan, C.; Kohan, D. E.; Pollock, J. S.; Pollock, D. M.; Webb, D. J.; Maguire, J. J. Endothelin. *Pharmacol. Rev.* **2016**, *68* (2), 357–418.

(46) Chen, H. I.; Sudol, M. The WW Domain of Yes-Associated Protein Binds a Proline-Rich Ligand That Differs from the Consensus Established for Src Homology 3-Binding Modules. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (17), 7819–7823.

(47) Chen, H. I.; Einbond, A.; Kwak, S. J.; Linn, H.; Koepf, E.; Peterson, S.; Kelly, J. W.; Sudol, M. Characterization of the WW Domain of Human Yes-Associated Protein and Its Polyproline-Containing Ligands. *J. Biol. Chem.* **1997**, *272* (27), 17070–17077.

(48) Staub, O.; Dho, S.; Henry, P. C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D. WW Domains of Nedd4 Bind to the Proline-Rich PY Motifs in the Epithelial Na+ Channel Deleted in Liddle's Syndrome. *EMBO J.* **1996**, *15* (10), 2371–2380.

(49) Rentschler, S.; Linn, H.; Deininger, K.; Bedford, M. T.; Espanel, X.; Sudol, M. The WW Domain of Dystrophin Requires EF-Hands Region to Interact with β-Dystroglycan. *Biol. Chem.* **1999**, *380* (4), 431–442.

(50) Lu, P. J.; Zhou, X. Z.; Shen, M.; Lu, K. P. Function of WW Domains as Phosphoserine- or Phosphothreonine-Binding Modules. *Science.* **1999**, *283* (5406), 1325–1328.

(51) Bruce, M. C.; Kanelis, V.; Fouladkou, F.; Debonneville, A.; Staub, O.; Rotin, D. Regulation of Nedd4-2 Self-Ubiquitination and Stability by a PY Motif Located within Its HECT-Domain. *Biochem. J.* **2008**, *415* (1), 155–163.

(52) Bedford, M. T.; Sarbassova, D.; Xu, J.; Leder, P.; Yaffe, M. B. A Novel Pro-Arg Motif Recognized by WW Domains. *J. Biol. Chem.* **2000**, *275* (14), 10359–10369.

(53) Ermekova, K. S.; Zambrano, N.; Linn, H.; Minopoli, G.; Gertler, F.; Russo, T.; Sudol, M. The WW Domain of Neural Protein FE65 Interacts with Proline-Rich Motifs in Mena, the Mammalian Homolog of *Drosophila* Enabled. *J. Biol. Chem.* **1997**, *272* (52), 32869–32877.

(54) Bedford, M. T.; Chan, D. C.; Leder, P. FBP WW Domains and the Abl SH3 Domain Bind to a Specific Class of Proline-Rich Ligands. *EMBO J.* **1997**, *16* (9), 2376–2383.

(55) Ingham, R. J.; Colwill, K.; Howard, C.; Dettwiler, S.; Lim, C. S. H.; Yu, J.; Hersi, K.; Raaijmakers, J.; Gish, G.; Mbamalu, G.; et al. WW Domains Provide a Platform for the Assembly of Multiprotein Networks. *Mol. Cell. Biol.* **2005**, *25* (16), 7092–7106.

(56) Aragón, E.; Goerner, N.; Xi, Q.; Gomes, T.; Gao, S.; Massagué, J.; Macias, M. J. Structural Basis for the Versatile Interactions of Smad7 with Regulator WW Domains in TGF-β Pathways. *Structure* **2012**, *20* (10), 1726–1736.

(57) Iglesias-Bexiga, M.; Castillo, F.; Cobos, E. S.; Oka, T.; Sudol, M.; Luque, I. WW Domains of the Yes-Kinase-Associated-Protein (YAP) Transcriptional Regulator Behave as

Independent Units with Different Binding Preferences for PPxY Motif-Containing Ligands. *PLoS One* **2015**, *10* (1), e0113828.

(58) Duan, G.; Walther, D. The Roles of Post-Translational Modifications in the Context of Protein Interaction Networks. *PLoS Comput. Biol.* **2015**, *11* (2).

(59) Wagner, M. J.; Stacey, M. M.; Liu, B. A.; Pawson, T. Molecular Mechanisms of SH2- and PTB-Domain-Containing Proteins in Receptor Tyrosine Kinase Signaling. *Cold Spring Harb. Perspect. Biol.* **2013**, *5* (12).

(60) Erce, M. A.; Low, J. K. K.; Hart-Smith, G.; Wilkins, M. R. A Conditional Two-Hybrid (C2H) System for the Detection of Protein-Protein Interactions That Are Mediated by Post-Translational Modification. *Proteomics* **2013**, *13* (7), 1059–1064.

(61) Yi, L.; Gebhard, M. C.; Li, Q.; Taft, J. M.; Georgiou, G.; Iverson, B. L. Engineering of TEV Protease Variants by Yeast ER Sequestration Screening (YESS) of Combinatorial Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (18), 7229–7234.

(62) Tanos, B.; Pendergast, A. M. Abl Tyrosine Kinase Regulates Endocytosis of the Epidermal Growth Factor Receptor. *J. Biol. Chem.* **2006**, *281* (43), 32714–32723.

(63) Bidlingmaier, S.; Liu, B. Construction and Application of a Yeast Surface-Displayed Human CDNA Library to Identify Post-Translational Modification-Dependent Protein-Protein Interactions. *Mol. Cell. Proteomics* **2006**, *5* (3), 533–540.

(64) Zhou, S.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J.; et al. SH2 Domains Recognize Specific Phosphopeptide Sequences. *Cell* **1993**, *72* (5), 767–778.

(65) Machida, K.; Thompson, C. M.; Dierck, K.; Jablonowski, K.; Kärkkäinen, S.; Liu, B.; Zhang, H.; Nash, P. D.; Newman, D. K.; Nollau, P.; et al. High-Throughput Phosphotyrosine Profiling Using SH2 Domains. *Mol. Cell* **2007**, *26* (6), 899–915.

(66) Nollau, P.; Mayer, B. J. Profiling the Global Tyrosine Phosphorylation State by Src Homology 2 Domain Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (24), 13531–13536.

(67) Machida, K.; Mayer, B. J.; Nollau, P. Profiling the Global Tyrosine Phosphorylation State. *Mol. Cell. Proteomics* **2003**, *2* (4), 215–233.

(68) Tiruthani, K.; Mischler, A.; Ahmed, S.; Mahinthakumar, J.; Haugh, J. M.; Rao, B. M. Design and Evaluation of Engineered Protein Biosensors for Live-Cell Imaging of EGFR Phosphorylation. *Sci. Signal.* **2019**, *12* (584), eaap7584.

(69) Stauffer, T. P.; Meyer, T. Compartmentalized IgE Receptor-Mediated Signal Transduction in Living Cells. *J. Cell Biol.* **1997**, *139* (6), 1447–1454.

(70) Haugh, J. M.; Meyer, T. Active EGF Receptors Have Limited Access to Ptdlns(4,5)P2 in Endosomes: Implications for Phospholipase C and PI 3-Kinase Signaling. *J. Cell Sci.* **2002**, *115* (2), 303–310.

(71) Rotin, D.; Margolis, B.; Mohammadi, M.; Daly, R. J.; Daum, G.; Li, N.; Fischer, E. H.; Burgess, W. H.; Ullrich, A.; Schlessinger, J. SH2 Domains Prevent Tyrosine

Dephosphorylation of the EGF Receptor: Identification of Tyr992 as the High-Affinity Binding Site for SH2 Domains of Phospholipase C Gamma. *EMBO J.* **1992**, *11* (2), 559–567.

(72) Bacon, K.; Bowen, J.; Reese, H.; Rao, B. M.; Menegatti, S. Use of Target-Displaying Magnetized Yeast in Screening MRNA-Display Peptide Libraries to Identify Ligands. *ACS Comb. Sci.* **2020**, *22* (12), 738–744.

(73) Lown, P. S.; Hackel, B. J. Magnetic Bead-Immobilized Mammalian Cells Are Effective Targets to Enrich Ligand-Displaying Yeast. *ACS Comb. Sci.* **2020**, *22* (5), 274–284.

(74) Cruz-Teran, C. A.; Tiruthani, K.; Mischler, A.; Rao, B. M. Inefficient Ribosomal Skipping Enables Simultaneous Secretion and Display of Proteins in Saccharomyces Cerevisiae. *ACS Synth. Biol.* **2017**, *6* (11), 2096–2107.

(75) Lipovšek, D.; Antipov, E.; Armstrong, K. A.; Olsen, M. J.; Klibanov, A. M.; Tidor, B.; Wittrup, K. D. Selection of Horseradish Peroxidase Variants with Enhanced Enantioselectivity by Yeast Surface Display. *Chem. Biol.* **2007**, *14* (10), 1176–1185.

CHAPTER 4

Use of Target-Displaying Magnetized Yeast in Screening mRNA-Display Peptide Libraries to Identify Ligands

Adapted from Bacon K, Bowen J, Reese H, B.M Rao, and Menegatti S. (2020) Use of Target-Displaying Magnetized Yeast in Screening mRNA Display Peptide Libraries to Identify Ligands. *ACS Combinatorial Science 22(12)*: 738-744.

4.1 Introduction

Cell-free combinatorial libraries, such as ribosomal- and mRNA-display, are attractive tools in ligand discovery owing to their chemical diversity and ease of post-translational modifications[1]. In the mRNA-display platform, a peptide sequence is connected to its coding mRNA via a puromycin linkage[2]. The mRNA is reverse transcribed to form mRNA-cDNA-peptide fusions that can be used in library screening to identify peptide affinity ligands. Following library selection, the DNA linked to the peptide fusions that bind the target is amplified and sequenced. This technology is widely utilized in the discovery of short peptides[3], which are of great interest for use as tags or modulators of protein-protein interactions[4]; short peptides are also amenable to post-translational modifications, such as cyclization, which typically imparts higher target binding affinity and proteolysis resistance[5]. In prior work, our group developed a method to generate mRNA-display libraries of cyclic peptides via head-to-side chain chemical crosslinking of peptide-mRNA fusions adsorbed on a solid phase[6].

The selection of mRNA-display peptide libraries is commonly performed against a target protein conjugated onto synthetic magnetic nanoparticles[6]. Generally, recombinant expression and subsequent purification is required to obtain a soluble form of the target protein to be immobilized. As an alternative, the target protein can be expressed as a yeast cell surface fusion using the yeast surface display platform.[7] The use of yeast-displayed targets is particularly impactful for proteins that require a eukaryotic host for appropriate expression[8]. Yeast-displayed targets have been employed to identify binding proteins specific to membrane protein targets from both phage- and yeast-display libraries[7,9]. However, their use in screening mRNA-display libraries has not been previously explored.

The use of yeast-displayed targets for screening mRNA-display libraries poses a challenge: how to separate the selected fusions from the unbound portion of the library. As the fusions are soluble, the use of centrifugation for separating yeast-bound fusions represents an option; however, removal of the unbound fusions is tedious. Unbound fusions from even small amounts of residual liquid may lead to a significant fraction of false positives. This highlights the need for improved isolation techniques capable of reducing carryover and increasing selection stringency.

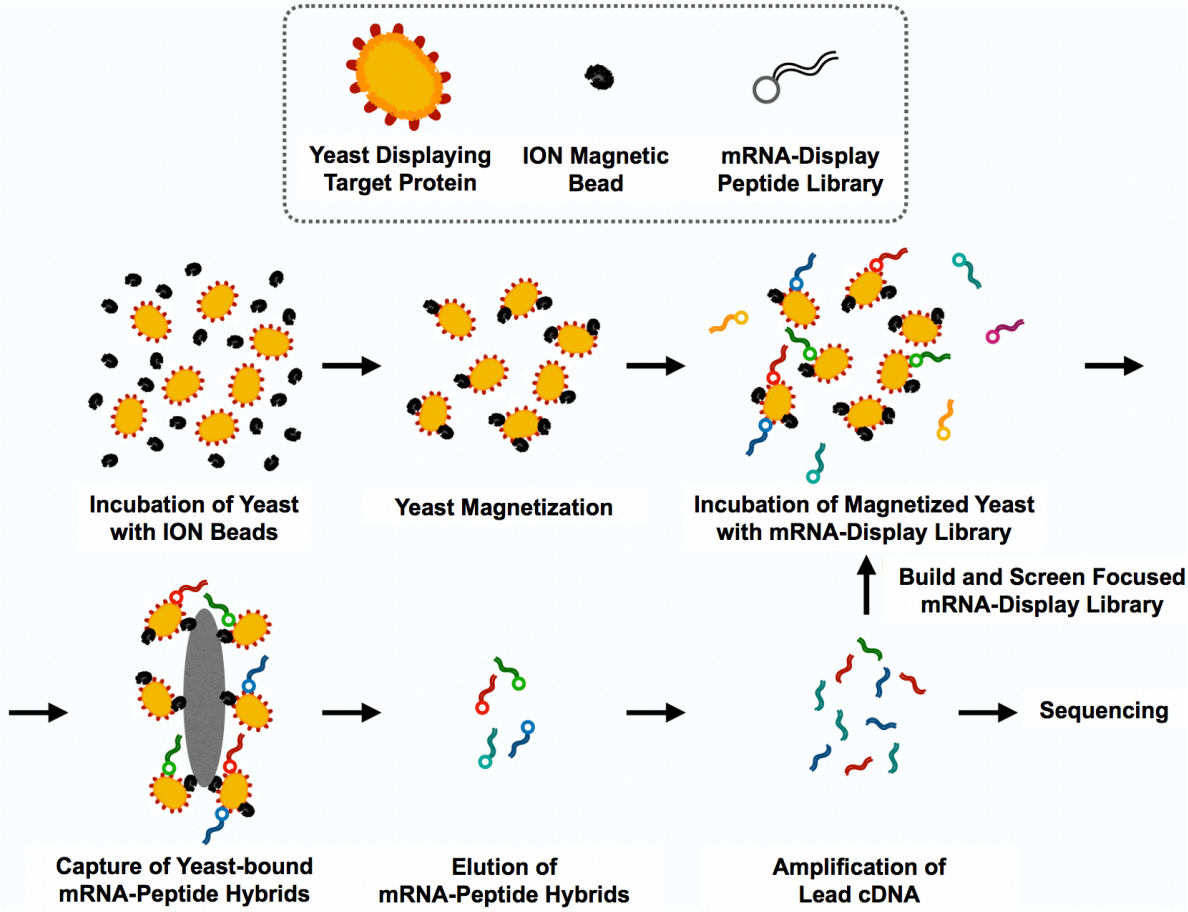


Figure 4.1. Selection of a mRNA-display peptide library against non-specifically magnetized target yeast cells. Yeast cells expressing WW-YAP or TOM22 were incubated with iron oxide magnetic nanoparticles (ION magnetic beads). Magnetic yeast cells were separated from non-magnetized yeast cells using a magnet. The mRNA-display-peptide library was then incubated with the magnetized target cells. Any mRNA-display peptide fusions bound to the target cells were isolated using a magnet. After, the positively bound mRNA-display peptide fusions were eluted from the target cells. DNA linked to the eluted clones was amplified, and a new library was synthesized using this DNA as the template for the next round of screening. Additional rounds of selection against the target cells were performed. The stringency of elution increased for each selection round.

To address this challenge, this work presents the use of magnetized yeast cells as protein-display particles to increase the efficiency of isolating lead mRNA-peptide fusions (Figure 4.1). Yeast are magnetized by adsorbing iron oxide nanoparticles (IONs) onto anionic moieties that constellate on the cell wall in a pH-controlled environment[10]. Yeast expressing proteins with varying isoelectric points can be magnetized by adjusting the pH of the buffer in which the IONs are adsorbed[11]. Following adsorption, the IONs are blocked with albumin to prevent non-specific adsorption of mRNA-peptide hybrids and other species in solution. The selection of mRNA-display libraries generally takes place at a pH of 7.4. If yeast cells are magnetized at a different pH, IONs can dissociate when the magnetic yeast cells are incubated in the selection buffer. mRNA-peptide hybrids associated with a demagnetized yeast cell are subsequently lost. Nonetheless, more than 50% of non-specifically magnetized cells maintain their magnetized state over the course of a selection[7]. As an alternative, affinity-based magnetization methods can be used, which result in lower losses due to target cell de-magnetization (~30% loss)[7]. The size of mRNA-display libraries (~10^{12} hybrids, corresponding to ~1000 copies per peptide in case of 7-mers) and the number of target proteins expressed on yeast (~50,000 per cell[12,13]) ensure ample contact between the library and the target proteins. Also, because there are multiple copies of each peptide in the library, it is likely that any peptide with affinity for the target will be isolated even if some fusions are lost due to ION desorption.

In this work, a combinatorial library of cyclic peptides generated using mRNA-display was screened against magnetized target yeast to identify peptides with affinity for mitochondrial surface protein TOM22. TOM22 is a mitochondrial membrane protein involved in the recognition and translocation of mitochondrial pre-proteins produced in the cytosol[14]. Peptides with affinity for TOM22 have the potential be used for targeted delivery of therapeutic payloads to treat patients

with mitochondrial deficiencies[15]. TOM22 was expressed on the surface of EBY100 yeast using the yeast surface display platform (surface expression levels detailed in Figure 4.2)[16]. The cells were magnetized with IONs and utilized as targets during subsequent selections of mRNA-display peptide libraries.

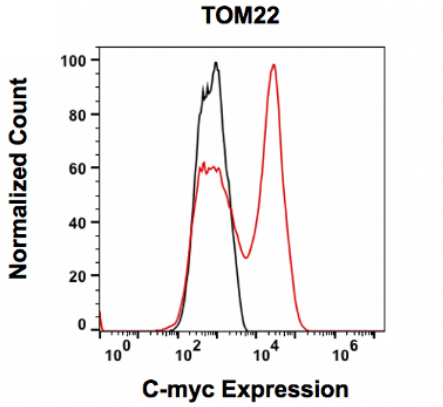


Figure 4.2. Expression of TOM22 on the surface of yeast. Expression levels were quantified through the immunofluorescent detection of a fused c-myc tag. In the plasmid used for yeast display, a c-myc tag is included that is c-terminal to surface expressed WW-YAP and TOM22. To analyze expression, yeast cells displaying WW-YAP or TOM22 were labeled with an anti-c-myc antibody. Goat-anti-chicken-488 was used to detect binding of the anti-c-myc antibody via flow cytometry.

4.2 Results and Discussion

4.2.1 Screening a mRNA-display library of cyclic peptides against magnetic yeast displaying target protein

A mRNA-display library of cyclic peptides was constructed using a previously developed method of head-to-side chain cyclization that employs a glutarate linker to tether amine groups within methionine and lysine residues flanking the variable peptide segment (NNN degeneracy)[6,17]. Prior to selecting against target yeast cells, the library was negatively screened against magnetized yeast displaying a non-target protein (*i.e.*, Fc portion of human IgG, hFc) to remove fusions bound to non-specific, yeast surface proteins. Throughout the library selection process, the stringency of the elution conditions was increased to promote isolation of high-affinity

binders. Initially, the mRNA-peptide fusions were eluted from the target cells using 0.15 M potassium hydroxide; in later rounds, the yeast-bound peptides were washed with a mild acidic buffer (50 mM NaCl in 50 mM sodium acetate at pH 5) prior to alkaline elution to remove weakly bound peptides and increase the probability of isolating true affinity binders. Following elution, the cDNA linked to the isolated mRNA-peptide fusions was amplified to generate a focused sub-library for subsequent screening rounds. We noted that, during the generation of the sub-libraries, heptamer and decamer sequences emerged. While potentially present as minor contaminants in the original library, these longer peptides may have outcompeted the original pentamers; the longer peptides likely benefit from a higher enthalpic contribution to their binding free energy due to the presence of more residues that can contact the target protein. Accordingly, the longer peptides were identified as lead fusions after multiple rounds of selection.

4.2.2 Isolation of cyclic peptides specific to TOM22

After five rounds of selection of the mRNA-display cyclic peptide library against magnetic yeast displaying TOM22, ten individual colonies were sequenced, and seven complete sequences were returned (Figure 4.3A). The majority of these sequences exhibit an overall positive charge, which likely promotes interaction with negatively charged TOM22 (isoelectric point ~4)[18]. Only two of the isolated sequences exhibit either a neutral or negative overall charge, TOM22.4 and TOM22.7 respectively. The anionic character of TOM22 is instrumental for protein translocation across mitochondria, as translocating proteins contain positively charged pre-sequences. Previous studies identified a TOM22-binding peptide, KTGALLLQ[19], which is cationic and rich in hydrophobic residues, like the sequences identified in our selection. Another TOM22-binding sequence identified in the same study, LCTKVPEL, contains 1 cationic and 1 anionic residue, and a balance of hydrophobic and hydrophilic amino acids, similar to cyclo[M-PELNRAI-K]

identified in this work. Based on the diversity of these previously identified TOM22-binding peptides, it is not surprising that we also isolated a diverse set of peptides from our screens.

4.2.3 Binding specificity evaluation of isolated cyclic peptides for TOM22 using a luciferase based binding assay

We resolved to study the TOM22 binding of cyclo[M-PELNRAI-K]. The chemical diversity of this peptide, which contains positively (R) and negatively charged (E), polar (N), and hydrophobic (A, I, L, and P) residues, is conducive to TOM22-binding by true affinity. Briefly, yeast cells co-expressing TOM22 and luciferase were incubated with magnetic beads functionalized with cyclo[M-PELNRAI-K] in the presence of non-displaying EBY100 yeast cells, which were included to reduce non-specific binding. Yeast cells bound to the peptide-functionalized beads were isolated using a magnet and subjected to a luminescence-based assay, wherein the luminescence signal is proportional to the number of TOM22-displaying cells captured by the magnetic beads. A study by Bacon et al. describes how this assay can be used to quantitatively rank ligands with different affinities for a target protein[20]; the correlation between the number of TOM22-displaying cells and luminescence signal is reported in a calibration curve detailed in Figure 4.4.

We also considered the binding of cells co-displaying a non-target protein (c-Kit) and luciferase to magnetic beads functionalized with peptide cyclo[M-PELNRAI-K]. We observed that $TOM22^{+}$ yeast cells were captured by the peptide-functionalized beads at a statistically (~4.5-fold) higher level than $TOM22^{-}$/c-KIT^{+} cells (Figure 4.3B), suggesting that cyclo[M-PELNRAI-K] is selective for TOM22. Based on the capture levels of the $TOM22^{+}$ cells by the magnetic peptide beads, in comparison to other protein interactions studied previously using this luciferase assay, it is likely that the binding strength of the TOM22:cyclo[M-PELNRAI-K] complex is rather

moderate[20]. Nonetheless, the described luminescence-based assay was capable of detecting mild interactions owing to the multi-point binding mechanism (avidity) between cells and the peptide-functionalized beads as well as the inherent sensitivity of the luciferase reporter. Due to its low affinity, we did not evaluate this peptide in the context of targeted delivery to mitochondria.

A)

	M	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	K
TOM22.1		S	V	H	L	R						
TOM22.2		G	V	H	R	N						
TOM22.3		R	G	Q	A	R	A	Y				
TOM22.4		P	E	L	N	R	A	I				
TOM22.5		R	C	S	H	A	L	G	A	R	E	
TOM22.6		T	V	K	D	R	F	C	R	M	F	
TOM22.7		L	Q	A	E	G	G	D	M	W	M	

B)

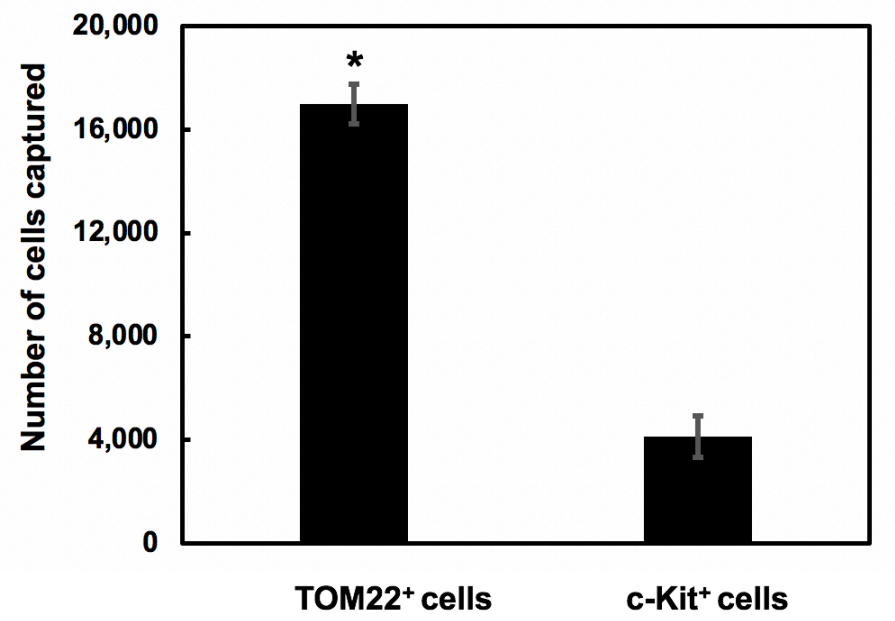


Figure 4.3. Identified TOM22-binding cyclic peptides and specificity characterization. **(A)** Amino acid sequences of peptides isolated from the selection of a mRNA-display cyclic peptide library against yeast-displayed TOM22. **(B)** Recovery of yeast cells displaying TOM22 or c-Kit by magnetic beads functionalized with cyclo[M-PELNRAI-K]. Yeast cells expressing TOM22 or c-Kit also displayed an engineered luciferase protein as a yeast surface fusion. Cell capture was quantified as a function of the luminescence signal produced by the recovered population using a generated luminescence standard curve. * represents $p < 0.05$ for a two tailed, paired t-test in comparison to the recovery of c-Kit displaying yeast cells.

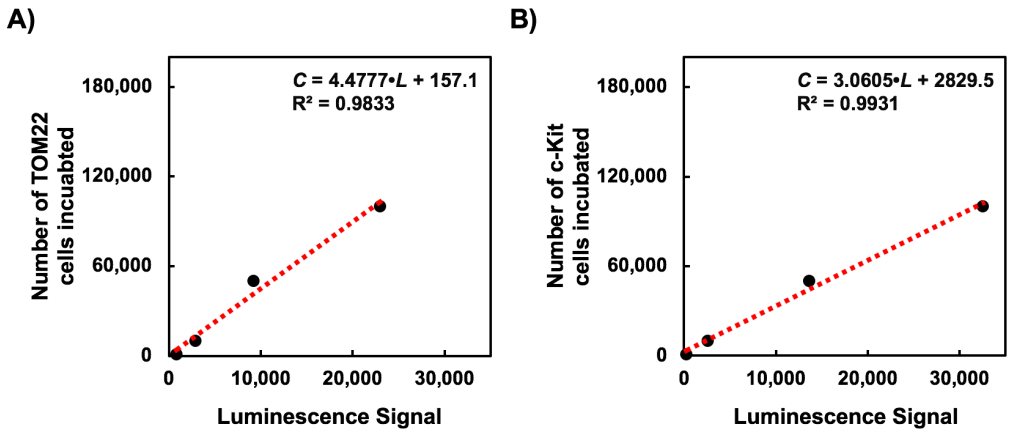


Figure 4.4. Calibration curves used to correlate luminescence signal to the number of cells incubated. Curves were generated for cells expressing TOM22 **(A)** or c-Kit **(B)**. These cells also co-expressed a luciferase protein. For each repeat, a new calibration curve specific to each cell type was generated. A linear regression was fit to the data to relate the number of cells present (*C*) to the produced luminescence signal (*L*). This relationship was used to estimate the number of TOM22 or c-Kit cells captured by the peptide functionalized magnetic beads based on the luminescence signal produced by the isolated cells.

Here, we have demonstrated the isolation of cyclic peptide affinity ligands for TOM22 by combinatorial screening of mRNA display libraries using magnetic yeast displaying the target proteins. This method was also used to identify cyclic peptide affinity ligands for a different protein target, the WW domains of YAP (data not shown). Values of protein-binding affinity in the μM range, as those measured in this work, are typical for linear and cyclic peptides of 6-to-10 amino acids in length[21]. It is however possible to enhance the affinity of peptide ligands selected combinatorially to reach a low-μM/high-nM range by adjusting the sequence length of the protein-binding segment to increase the enthalpic component of the binding energy[21]. Care should be taken in adjusting the protein-binding segment of cyclic peptides to prevent excessive flexibility of the peptide backbone, which would unfavorably affect the entropic component of the binding energy. The library screening process can also be revisited to improve the binding affinity of isolated mutants by increasing the stringency of the selections. In this work, the selection conditions were maintained throughout the successive screening rounds, and only the elution conditions were

adjusted; alternatively, competitive conditions can be adopted when incubating the library with the target cells by adding a mixture of soluble competitors, or by adjusting the composition, concentration, and pH of the selection buffer. Additional wash steps can also be performed to eliminate weak-affinity binders. Further, while held constant in this work, the number of protein-displaying cells could be reduced through the successive screening rounds to bias the isolation of higher affinity mutants. Finally, a larger number of selection rounds, beyond the four or five employed here, up to ten, can be performed to attain strong sequence homology among the identified peptides[22]; nonetheless, this work has shown that specific binding peptides can be identified using fewer rounds of screening.

4.3 Conclusions

Collectively, our results show that the use of magnetic yeast displaying a target protein is an effective tool to increase the throughput of library screening. Our approach limits the need for recombinant soluble protein and provides an alternative approach for screening mRNA display libraries. The use of yeast-displayed targets will be most effective for proteins that do not require complex post-translational modifications for their function, since glycosylation patterns may vary between yeast and other higher order eukaryotes.

4.4 Materials and Methods

4.4.1 Plasmids and yeast culture

The pCTCON and pCT302 vectors containing TRP or LEU selectable markers, respectively, were used in conjugation with *Saccharomyces cerevisiae* strain EBY100. Specifically, the Frozen-EZ yeast transformation Kit II (Zymo Research) was used to transform plasmid DNA into chemically competent EBY100. Trp-deficient SDCAA and SGCAA media was used, respectively, for culturing and inducing cells harboring the pCTCON plasmid[1]. Similarly, cells harboring

pCT302 plasmids were grown and induced with Leu-deficient SDSCAA (-Leu) and SGSCAA (-Leu) media.[2] Both Trp-deficient and Leu-deficient media have similar compositions. However, the Leu-deficient media does not contain casamino acids, but rather uses synthetic dropout mix (1.62 g/L; US Biological Life Sciences) lacking leucine. When culturing, yeast cells were grown in SDCAA or SDSCAA media at 30°C with shaking at 250 RPM. To induce protein expression, yeast cells were transferred into SGCAA or SGSCAA medium at an OD_{600} of 1 and incubated overnight at 20 °C with shaking at 250 rpm.

4.4.2 Plasmid construction for display of proteins as yeast surface fusions

Plasmids were constructed that afford the expression of the proteins used for positive and negative selections of the mRNA display libraries as yeast surface fusions. These proteins were encoded as fusions to Aga2, a yeast cell mating protein. pCTCON-TOM22, affording the expression of TOM22, was constructed by amplifying gene block 1 with Pf1 and Pr1. Similarly, pCTCON-hFc, affording the expression of hFc (Fc portion of human IgG), was constructed by amplifying gene block 2 with primers Pf2 and Pr2. DNA amplified from gene blocks 1 and 2 were inserted into pCTCON between the NheI and BamHI cut sites.

Double stranded gene fragments were purchased from Integrated DNA technologies (IDT). Primer oligonucleotides were bought from IDT or Eton Biosciences. Gene fragment and primer sequences can be found in Table 4.1-4.2, respectively. Phusion Polymerase (Thermo Fisher Scientific) was used for PCR reactions that took place in a 50 μL volume reaction following the manufacturer's protocols. Restriction enzyme digests of plasmid backbones and PCR products were executed at 37°C for 2 hours in a 50 μL volume using a 5-times excess of each appropriate restriction enzyme. Digested plasmid backbones were incubated with Antarctic phosphatase (New England Biolabs) for 1 hour at 37°C. Digested plasmids and PCR products were purified using a

9K series gel and PCR extraction kit (BioBasic). Overnight ligations using T4 DNA ligase (Promega) were performed with the digested plasmid backbones and inserts. Ligations were transformed into chemically competent Novablue *E.coli* cells. The cells were made chemically competent using Mix&Go! *E.coli* transformation buffers (Zymo Research). The GeneJET™ plasmid miniprep kit (Thermo Fisher Scientific) was used to harvest plasmid from overnight *E.coli* cultures.

Table 4.1. List of gene fragments described in Chapter 4.

Gene Block 1	GCTGCCGCCGTCGCTGCTGCCGGTGCAGGGGAACCCCAGTCCCCG GACGAATTGCTCCCGAAAGGCGACGCGGAGAAGCCTGAGGAGGA GCTGGAGGAGGACGACGATGAGGAGCTAGATGAGACCCTGTCGG AGAGACTATGGGGCCTGACGGAGATGTTTCCGGAGAGGGTCCGGT CCGCGGCCGGAGCCACTTTTGATCTTTCCCTCTTTGTGGCTCAGAA AATGTACAGGTTTTCCAGGGCA
Gene Block 2	CCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCA CCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAAC CCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCG TGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACT GGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGC GGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCA CCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCA AGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCT CCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGC CCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCT GCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGG AGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCC GTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCG TGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG TGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCT CCCTGTCTCCGGGTAAA

Table 4.1 (Continued).

Gene Block 3	CAACCATCTGTGAGTCCAGGGGAACCGTCTCCACCATCCATCCATC CAGGAAAATCAGACTTAATAGTCCGCGTGGGCGACGAGATTAGGC TGTTATGCACTGATCCGGGCTTTGTCAAATGGACTTTTGAGATCCT GGATGAAACGAATGAGAATAAGCAGAATGAATGGATCACGGAAA AGGCAGAAGCCACCAACACCGGCAAATACACGTGCACCAACAAA CACGGCTTAAGCAATTCCATTTATGTGTTTGTTAGAGATCCTGCCA AGCTTTTCCTTGTTGACCGCTCCTTGTATGGGAAAGAAGACAACGA CACGCTGGTCCGCTGTCCTCTCACAGACCCAGAAGTGACCAATTAT TCCCTCAAGGGGTGCCAGGGGAAGCCTCTTCCCAAGGACTTGAGG TTTATTCCTGACCCCAAGGCGGGCATCATGATCAAAAGTGTGAAA CGCGCCTACCATCGGCTCTGTCTGCATTGTTCTGTGGACCAGGAGG GCAAGTCAGTGCTGTCGGAAAAATTCATCCTGAAAGTGAGGCCAG CCTTCAAAGCTGTGCCTGTTGTGTCTGTGTCCAAAGCAAGCTATCT TCTTAGGGAAGGGGAAGAATTCACAGTGACGTGCACAATAAAAGA TGTGTCTAGTTCTGTGTACTCAACGTGGAAAAGAGAAAACAGTCA GACTAAACTACAGGAGAAATATAATAGCTGGCATCACGGTGACTT CAATTATGAACGTCAGGCAACGTTGACTATCAGTTCAGCGAGAGT TAATGATTCTGGAGTGTTCATGTGTTATGCCAATAATACTTTTGGA TCAGCAAATGTCACAACAACCTTGGAAGTAGTAGATAAAGGATTC ATTAATATCTTCCCCATGATAAACACTACAGTATTTGTAAACGATG GAGAAAATGTAGATTTGATTGTTGAATATGAAGCATTCCCCAAAC CTGAACACCAGCAGTGGATCTATATGAACAGAACCTTCACTGATA AATGGGAAGATTATCCCAAGTCTGAGAATGAAAGTAATATCAGAT ACGTAAGTGAACTTCATCTAACGAGATTAAAAGGCACCGAAGGAG GCACTTACACATTCCTAGTGTCCAATTCTGACGTCAATGCTGCCAT AGCATTTAATGTTTATGTGAATACAAAACCAGAAATCCTGACTTAC GACAGGCTCGTGAATGGCATGCTCCAATGTGTGGCAGCAGGATTC CCAGAGCCCACAATAGATTGGTATTTTTGTCCAGGAACTGAGCAG AGATGCTCTGCTTCTGTACTGCCAGTGGATGTGCAGACACTAAACT CATCTGGGCCACCGTTTGGAAAGCTAGTGGTTCAGAGTTCTATAGA TTCTAGTGCATTCAAGCACAATGGCACGGTTGAATGTAAGGCTTAC AACGATGTGGGCAAGACTTCTGCCTATTTTAACTTTGCATTTAAAG GTAACAACAAAGAGCAAATCCATCCCCACACCCTGTTCACTCCT

Table 4.2. List of oligonucleotide primers described in Chapter 4.

	Sequence
Pf1	GTTCTCGCTAGCGCTGCCGCCGTCGCTG
Pr1	GCACTTGGATCCTGCCCTGGAAAACCTGTACATTT
Pf2	GTTGACGCTAGCCCCAAATCTTGTGACAAAACTCACA
Pr2	GCACTTGGATCCTTTACCCGGAGACAGGGAGA
Pf3	GCAAATTTCTAATACGACTCACTATAGGGACAATTACTATTTACAATTAC
Pr3	ATAGCCGGTGCCAGATCCAGACATTCCCATATGGTGATGGT
Pf4	GTTGACGCTAGCCAACCATCTGTGAGTCCAGGG
Pr4	GCACTTGGATCCAGGAGTGAACAGGGTGTGGG

4.4.3 Surface expression analysis

Cell surface expression levels of TOM22 as an Aga2 fusion was estimated using immunofluorescence detection. A c-myc tag is encoded as a c-terminal fusion to the displayed proteins. Accordingly, the expression of each protein fusion was detected using a chicken anti-c-myc antibody (Thermo Fisher Scientific). To begin, $5x10^6$ cells were labeled with a 1:100 dilution of the primary antibody for 15 minutes at room temperature. After, a secondary labeling was performed using a 1:250 dilution of goat-anti-chicken 488 (Immunoreagents) for 10 minutes on ice. All labelings were conducted in 50 μL of 0.1% PBSA. Washes took place between each labeling. The binding of the antibodies was detected using a Miltenyi Biotec MACsQuant VYB cytometer.

4.4.4 Construction of the cyclic peptide mRNA-display libraries

mRNA display libraries of randomized peptides were constructed using previously described protocols for guidance[3,4]. Oligonucleotides 1-3 (Table 4.3) encoding either a pentapeptide ($MX_1X_2X_3X_4X_5K$), heptapeptide ($MX_1X_2X_3X_4X_5X_6X_7K$), or decapeptide ($MX_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}K$) library were PCR-amplified using primers Pf3 and Pr3 (Table 4.2) X represents any of the 20 amino acids. All oligonucleotides and primers were purchased from

IDT. These primers add a 5' consensus sequence containing a T7 RNA polymerase promoter, a TMV translation enhancer, and a sequence coding the FLAG epitope tag. Also added is a 3' consensus sequence including a 6xHis tag followed by a sequence that affords conjugation of a puromycin linker. 25 PCR reactions were performed in a volume of 50 μL that contained 1 U of Phusion HF DNA polymerase (Thermo Fisher Scientific), 1X Phusion Buffer, 0.2 mM deoxynucleotide triphosphate (dNTPs) (Thermo Fisher Scientific), 0.1 μM of the forward and reverse primers, 1 M betaine, 3% dimethyl sulfoxide (DMSO), and 20 ng of the template DNA. The PCR was performed using the following conditions: Initial denaturation at 98°C for 2 minutes, followed by 30 cycles of denaturation at 98°C for 1 minute, annealing at 66°C for 1 minute, extension at 72°C for 15 seconds, and a final extension at 72°C for 10 minutes.

Table 4.3. Oligonucleotides encoding randomized peptides described in Chapter 4.

	Sequence
Oligo 1	5'-GGA CAA TTA CTA TTT ACA ATT ACA ATG NNN NNN NNN NNN NNN AAA GGC AGC GGC TCC GGT CAT CAC CAC CAT CAC CAT ATG GGA ATG-3'
Oligo 2	5'-GGA CAA TTA CTA TTT ACA ATT ACA ATG NNN NNN NNN NNN NNN NNN NNN AAA GGC AGC GGC TCC GGT CAT CAC CAC CAT CAC CAT ATG GGA ATG-3'
Oligo 3	5'-GGA CAA TTA CTA TTT ACA ATT ACA ATG NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN AAA GGC AGC GGC TCC GGT CAT CAC CAC CAT CAC CAT ATG GGA ATG-3'

After, EDTA was added to the pooled PCR reactions at a final concentration of 5.5 mM. The amplified DNA was purified using phenol: chloroform: isoamyl alcohol extraction followed by ethanol precipitation. Briefly, the DNA was extracted 2X with 1 volume of phenol: chloroform: isoamyl alcohol followed by an extraction with 1 volume of chloroform. After, 0.1 volumes of 3M potassium acetate, 2 volumes of ethanol, and 10 μL of linear acrylamide (Invitrogen) were added to the extracted DNA prior to overnight incubation at -20°C. The following day, the precipitated

DNA was pelleted and washed 1X with 75% ethanol followed by a wash with 100% ethanol. After, the pellet was air dried prior to resuspension in 0.1% DEPC water.

The purified DNA was then used in an *in vitro* transcription reaction. The 300 μL reaction contained 12 μg of DNA, 5 mM ribonucleotide triphosphate (rNTPs) (Promega), 19 mM $MgCl_2$, 1X transcription buffer, 45 μL of T7 RNA polymerase (Thermo Fisher Scientific), and 0.1% DEPC water to reach the final volume. The reaction was incubated for 8 hours at 37°C. EDTA was added to a final concentration of 50 mM. After, the mRNA was purified by acidic phenol chloroform extraction (2X, 1 volume) followed by an extraction with chloroform (1X, 1 volume). The extracted mRNA was further purified using a Nap 5 column (GE Healthcare Life Sciences) by eluting in 0.1% DEPC water. Subsequently, the template DNA was digested in an 850 μL reaction for 4 hours at 37°C with 42.5 μL of RNase-free Turbo DNASE (1 U/μL, Thermo Fisher Scientific) and 1X DNase buffer. After, EDTA was added to a final concentration of 20 mM. The reaction mixture was purified using acidic phenol chloroform extraction followed by ethanol precipitation as previously described. This ethanol precipitation used lithium chloride instead of potassium acetate.

A puromycin linker ([psoralen-$(ATAGCCGGTG)_2$-OMe-dA_{15}-C9C9-Acc-puromycin]; Keck Oligo Synthesis Lab, Yale University) was conjugated to the purified mRNA. The conjugation reaction contained 200 μg of purified mRNA, 20 mM HEPES, 100 mM KCl, and the puromycin linker at 2.5 times the total molar concentration of mRNA. The total reaction volume was 250 μL. Conjugation took place in a thermocycler using the following conditions: 85°C for 8 minutes, then 60 cycles with a 1°C decrease in each cycle from 85°C to 25°C, followed by 25°C for 25 minutes, and a 4°C hold. Next, the puromycin linker was crosslinked to the mRNA using ultraviolet light (360 nm) for 20 minutes. The crosslinked mRNA was precipitated overnight using

lithium acetate and ethanol. Prior to *in vitro* translation, the precipitated crosslinked mRNA was recovered via centrifugation and washed as previously described.

After, a 500 µL translation reaction was set up containing the crosslinked mRNA, 340 µL of rabbit reticulocyte lysate (Invitrogen), 50 µM methionine, 1X buffer without methionine, and nuclease free water to reach the final volume. The reaction mixture was incubated at 30°C for 1.5 hours prior to the addition of 1 M MgCl2 and 1 M KCl to final concentrations of 76 mM and 880 mM, respectively. After, this mixture was held at room temperature for 1 hour and then stored overnight at -20°C.

The puromycin linker contains a poly(dA) sequence allowing for purification of mRNA-puromycin and mRNA-puromycin-peptide fusions using oligo dT beads. The translation mix was added to 1 mL of magnetic oligo-dT beads (New England BioLabs) that had been washed prior 3X with 0.1% DEPC water and 3X with 10 mM Tris pH 7.5, 1 mM EDTA, and 0.05% SDS. The translation mix and washed beads were incubated in 9 mL of fresh binding buffer (0.5 M NaCl, 10 mM Tris pH 7.5, 1 mM EDTA, 0.05% SDS, and 1 mM DTT) for 2 hours at 4°C. Next, the beads were washed 3x10 minute with wash buffer (0.2 M NaCl, 10 mM Tris pH 7.5, 1 mM EDTA, 0.05% SDS, 1 mM DTT) followed by washes 3x10 minutes with crosslinking buffer (0.2 M NaCl, 1 mM EDTA, 0.05% SDS, 1 mM DTT). Subsequently, cyclization was performed by resuspending the beads in 800 µL of crosslinking buffer and adding 50 µL of a 3 mg/mL solution of a disuccinimidyl glutarate crosslinker (Thermo Fisher Scientific) dissolved in dimethylformamide. The cyclization reaction took place for two hours at 4°C. After, the beads were washed 3x10 minute with crosslinking buffer. The cyclization reaction was repeated a second time. Lastly, the beads were washed 3x10 minute with wash buffer and then eluted in 600 µL of 0.1% DEPC water with 1 mM dTT overnight at 4°C.

The cyclized, oligo(dT) purified product underwent a reverse transcription reaction the following day. Initially, 25 μL of 100 μM reverse transcription primer (5'- TTT TTT TTT TNN CCA GAT CCA GAC ATT CCC AT-3') was incubated with the oligo(dT) purified product for 15 minutes at room temperature. Next, 200 μL of 5X first strand buffer, 50 μL of 10 mM dNTPs, 100 μL of 0.1 M DTT, and 20 μL of 0.1% DEPC water was added to the reaction mixture. The reaction mixture was then incubated at 42°C for 2 minutes prior to the addition of 5 μL of Superscript™ II reverse transcriptase (Invitrogen). Subsequently, the reaction was held for 50 minutes at 42°C. After, EDTA was added to the reaction at a final concentration of 6 mM. The reaction mixture was then passed through a Nap-10 column (GE Healthcare Life Sciences) using 20 mM Tris pH 7.5, 150 mM NaCl, 0.05% Tween 20 as the equilibration and elution buffers.

Magnetic Ni-NTA agarose beads (Thermo Fisher Scientific, Pierce) were used to isolate the mRNA-cDNA-peptide fusions by taking advantage of the 6xHis tag on the translated peptide. 1 mL of beads was washed with binding buffer (50 mM Sodium Phosphate pH 8, 300 mM NaCl). After, the Nap-10 column eluate was incubated with the beads at room temperature for 1 hour. The beads were washed 3X with binding buffer and eluted with 500 μL of 50 mM sodium phosphate pH 8, 300 mM NaCl, 300 mM imidazole. Subsequently, the eluted mRNA-cDNA-peptide fusions were desalted using a Nap-5 column by eluting with 0.1% DEPC water. Library diversity was estimated using a A_{260} measurement.

4.4.5 Magnetization of yeast cells

$5x10^7$ yeast cells expressing each selection protein were incubated in 2 mL of their respective magnetization buffer with 0.4 mg iron oxide (4 mg/mL in water) for 30 minutes at room temperature. Yeast cells displaying TOM22 were magnetized in 50 mM sodium acetate pH 5, 50 mM NaCl buffer. Yeast cells displaying hFc were magnetized in 50 mM Tris-HCl pH 7.4, 300

mM NaCl buffer. Different buffers were used as it appeared iron oxide binding was dependent on the isoelectric point of the yeast displayed target protein. Any cells bound to the iron oxide were isolated from the unbound cells using a magnet. The OD_{600} of the solution prior to the addition of the iron oxide (OD_I) and after the removal of the yeast bound to the iron oxide (OD_f) was measured. Yeast recovery was calculated as $(OD_I - OD_f)/OD_f$. ~ $2x10^7$ cells were magnetized and used in each screen.

After the initial magnetization, the yeast-iron oxide conjugates were washed three times to remove any unbound or weakly bound cells using the appropriate magnetization buffer. Subsequently, the yeast-iron oxide conjugates were incubated with PBS pH 7.4, 1% BSA, 0.1% salmon sperm DNA (1% PBSASD) for 1 hour at room temperature to block any unbound iron oxide. After blocking, the yeast-iron oxide conjugates were washed two more times with 1% PBSASD to remove any yeast cells that may have dissociated from the iron oxide. After these steps, the magnetic yeast were used as targets in the selection of mRNA display libraries.

4.4.6 mRNA-display screening to identify affinity peptides using magnetic yeast target

A cyclic peptide library with five randomized amino acid positions was used to screen for peptide binders to TOM22. Five rounds of screening were performed. During each round, a negative selection was initially performed against yeast cells expressing hFc. The mRNA-cDNA-peptide fusions were incubated with magnetic yeast cells displaying hFc in 1 mL of PBS pH 7.4, 0.1% BSA, 0.1% salmon sperm DNA for 1 hour at room temperature. Subsequently, the incubation mixture was placed on a magnet and the supernatant containing unbound mRNA-cDNA-peptide fusions was removed and incubated with magnetic yeast cells displaying either WW-YAP or TOM22. After a 1-hour incubation, the selection mixture was placed on a magnet and the

supernatant was removed to isolate any mRNA-cDNA-peptide fusions positively bound to the target displaying cells.

The selection stringency was increased for each round of screening. For round one, peptide-yeast-iron oxide conjugates were not washed after separation using a magnet. The bound peptide fusions were eluted twice in 200 μL of 0.15 M potassium hydroxide for 1 hour. The eluted fractions were combined, and the cDNA associated with the eluted fusions was used as a template to generate the library for the next screening round. For round two, the elution steps were similar, but the bound peptides fusions were washed 3 times with gentle pipetting using 0.1% BSA PBS (0.1% PBSA) prior to elution. For round three, the same wash steps were carried out. For elution, peptide bound cells were first incubated in 200 μL of a weak, low pH buffer (50 mM sodium acetate, 50 mM NaCl pH 5) for 1 hour at room temperature, followed by 2X incubation in 0.15 M potassium hydroxide for 1 hour at room temperature. The library was remade for the next screening round with cDNA obtained after elution using potassium hydroxide to bias the isolation of higher affinity binders. For rounds four and five, the same washing and elution steps were carried out as in round three, but an additional hour of elution in the weak, low pH 5 buffer was performed. All eluates were neutralized using 5N HCl.

4.4.7 Library reconstruction for subsequent screening rounds

For each elution condition, the mRNA-cDNA-peptide fusions eluted in each incubation were pooled together. For rounds three through five, only the fusions eluted using the harsher potassium hydroxide buffer were pooled and precipitated for DNA amplification. cDNA was precipitated using ethanol precipitation and linear acrylamide as previously described. The precipitated cDNA was used to generate a new mRNA display library for the subsequent screening round. The precipitated cDNA was amplified using primers Pf3 and Pr3. A 50 μL PCR reaction

was performed containing 1 U Phusion HF DNA polymerase, 1X HF Phusion buffer, 0.2 mM dNTPS, 0.2 μM of each primer, 1 M betaine, 3% DMSO, and the precipitated cDNA template. The PCR was performed using the following conditions: Initial denaturation at 98°C for 2 minutes, followed by 30 cycles of denaturation at 98°C for 1 minute, annealing at 66°C for 1 minute, extension at 72°C for 15 seconds, and a final extension at 72°C for 10 minutes. The PCR product was ethanol precipitated using linear acrylamide as previously described and used as the template for additional PCRs. A visible product band was not observed on a gel until the second round of PCR. Additional PCR reactions were performed to obtain enough DNA for library generation. Similar steps as described were carried out using this amplified DNA to generate mRNA-cDNA-peptide fusions for the next round of screening.

After the final round of screening, the fusions eluted using potassium hydroxide were precipitated and amplified using similar conditions. The amplified DNA was inserted into the pJet2.1 vector via blunt-end ligation using the CloneJET kit (Thermo Fisher Scientific). The ligation mixture was transformed into chemically competent Novablue *E. coli* cells. DNA from 10 individual colonies was extracted using the GeneJET™ plasmid miniprep kit and sent for sequencing. Some isolated clones contained premature stop codons and were not considered.

4.4.8 Specificity analysis of a TOM22 binding peptide using a luminescence binding assay

The specificity of cyclo[M-PELNRAI-K] for TOM22 was evaluated using a luminescence based binding assay that takes advantage of a luciferase reporter to quantify recovery[40]. We used plasmids that afford the co-expression of a target protein and NanoLuc, an engineered luciferase. Previously described plasmid pCT302-T2A-NanoLuc was digested between its NheI and BamHI sites.[40] Gene block 1 encoding the DNA for TOM22 was amplified using primers Pf1 and Pr1, digested, and inserted into the digested plasmid backbone to create pCT302-TOM22-T2A-

NanoLuc. Similarly, Gene block 3 encoding the DNA for the extracellular domain of c-Kit was amplified using primers Pf4 and Pr4, digested, and inserted into the digested plasmid backbone to construct pCT302-cKit-T2A-NanoLuc.

Peptide cyclo[M-PELNRAI-K] was synthesized in its cyclic form by the UNC High-Throughput Peptide Synthesis and Array facility. The synthesized peptide included a C-terminal biotin and was cleaved from resin. The peptide was purified by RP18 HPLC (purity 95%) and its mass was confirmed by MALDI mass spec.

cyclo[M-PELNRAI-K] was immobilized onto the surface of magnetic streptavidin beads. Briefly, 25 μL of magnetic biotin binder Dyanbeads^{TM} (Invitrogen) were washed 2X with 0.1% PBSA. After, the beads were incubated with 11 μM of the TOM22 binding peptide overnight in a total volume of 100 μL using 0.1% PBSA at 4°C. The next morning the beads were washed 3X with 0.1% PBSA followed by an hour incubation in PBS pH 7.4, 1% BSA for blocking. The beads were then washed 2X with 0.1% PBSA, 0.05% tween-20 (0.1% PBSAT). Subsequently, the functionalized beads were incubated with either $1x10^7$ cells co-expressing TOM22 and NanoLuc or c-Kit and NanoLuc in a total volume of 2 mL using 0.1% PBSAT for 1 hour at room temperature. Also included were $1x10^9$ EBY100 cells to reduce non-specific binding. When the incubation was complete, the beads were washed 3X with 0.1% PBSAT and then resuspended in 100 μL of PBS.

The binding of the cells to the functionalized magnetic beads was detected using the Nano-Glo Luciferase Assay system (Promega). 100 μL of the reconstituted reagent was incubated with the magnetic beads. The reaction proceeded for 3 minutes. After, the tube was placed on a magnet and 100 μL of the supernatant was plated onto a 96 white-well plate with a clear bottom (Corning)

in duplicate. A Tecan Infinite 200 Plate Reader was used to read the luminescence signal with the following settings: integration time of 1000 ms, settle time of 0 ms, and no attenuation.

Calibration curves were made for the TOM22 and c-Kit reporter cells to develop a relationship between luminescence signal and number of reporter cells incubated. A known number of reporter cells ranging from $1x10^5$ - $1x10^7$ were resuspended in 100 μL of PBS. Luminescence assays were carried out as previously described for each concentration of cells to determine the points in the calibration curve. The luminescence of an equal volume of PBS and the Nano-Glo Luciferase Assay Reagent was also considered as a blank. Background subtracted luminescence signals were plotted against the number of reporter cells incubated to generate a standard curve. A linear regression was also fit to the data points. For each pull down assay, the fitted linear regression was used to estimate the number of reporter cells captured by the peptide functionalized beads using the luminescence signal produced by the captured cells.

4.5 References

(1) Galán, A.; Comor, L.; Horvatić, A.; Kuleš, J.; Guillemin, N.; Mrljak, V.; Bhide, M. Library-Based Display Technologies: Where Do We Stand? *Mol. Biosyst.* **2016**, *12* (8), 2342–2358.

(2) 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), 12297–12302.

(3) Wang, H.; Liu, R. Advantages of MRNA Display Selections over Other Selection Techniques for Investigation of Protein-Protein Interactions. *Expert Rev. Proteomics* **2011**, *8*, 335–346.

(4) Wójcik, P.; Berlicki, Ł. Peptide-Based Inhibitors of Protein-Protein Interactions. *Bioorganic Med. Chem. Lett.* **2016**, *26* (3), 707–713.

(5) Davies, J. S. The Cyclization of Peptides and Depsipeptides. *J. Pept. Sci.* **2003**, *9* (8), 471–501.

(6) Menegatti, S.; Hussain, M.; Naik, A. D.; Carbonell, R. G.; Rao, B. M. MRNA Display Selection and Solid-Phase Synthesis of Fc-Binding Cyclic Peptide Affinity Ligands. *Biotechnol. Bioeng.* **2013**, *110* (3), 857–870.

(7) Bacon, K.; Burroughs, M.; Blain, A.; Menegatti, S.; Rao, B. M. Screening Yeast Display Libraries against Magnetized Yeast Cell Targets Enables Efficient Isolation of Membrane Protein Binders. *ACS Comb. Sci.* **2019**, *21* (12), 817–832.

(8) Vieira Gomes, A.; Souza Carmo, T.; Silva Carvalho, L.; Mendonça Bahia, F.; Parachin, N. Comparison of Yeasts as Hosts for Recombinant Protein Production. *Microorganisms* **2018**, *6* (2), 38.

(9) Zhao, L.; Qu, L.; Zhou, J.; Sun, Z.; Zou, H.; Chen, Y.-Y.; Marks, J. D.; Zhou, Y. High Throughput Identification of Monoclonal Antibodies to Membrane Bound and Secreted Proteins Using Yeast and Phage Display. *PLoS One* **2014**, *9* (10), e111339.

(10) Safarik, I.; Maderova, Z.; Pospiskova, K.; Baldikova, E.; Horska, K.; Safarikova, M. Magnetically Responsive Yeast Cells: Methods of Preparation and Applications. *Yeast* **2015**, *32* (1), 227–237.

(11) Schwegmann, H.; Feitz, A. J.; Frimmel, F. H. Influence of the Zeta Potential on the Sorption and Toxicity of Iron Oxide Nanoparticles on S. Cerevisiae and E. Coli. *J. Colloid Interface Sci.* **2010**, *347* (1), 43–48.

(12) Gera, N.; Hussain, M.; Rao, B. M. Protein Selection Using Yeast Surface Display. *Methods* **2013**, *60* (1), 15–26.

(13) Kieke, M. C.; Shusta, E. V.; Boder, E. T.; Teyton, L.; Wittrup, K. D.; Kranz, D. M. Selection of Functional T Cell Receptor Mutants from a Yeast Surface-Display Library. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (10), 5651–5656.

(14) Wilpe, S. Van; Ryan, M. T.; Hill, K.; Maarse, A. C.; Meisinger, C.; Brix, J.; Dekker, P. J. T.; Moczko, M.; Wagner, R.; Meijer, M.; et al. Tom22 Is a Multifunctional Organizer of the Mitochondrial Preprotein Translocase. *Nature* **1999**, *401* (6752), 485–490.

(15) Kang, Y. C.; Son, M.; Kang, S.; Im, S.; Piao, Y.; Lim, K. S.; Song, M.; Park, K.; Kim, Y.; Pak, Y. K. Cell-Penetrating Artificial Mitochondria- Targeting Peptide-Conjugated Metallothionein 1A Alleviates Mitochondrial Damage in Parkinson's Disease Models. *Exp. Mol. Med.* **2018**, *50* (8), 105.

(16) Boder, E. T.; Wittrup, K. D. Yeast Surface Display for Screening Combinatorial Polypeptide Libraries. *Nat. Biotechnol.* **1997**, *15* (6), 553–557.

(17) Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. Design of Cyclic Peptides That Bind Protein Surfaces with Antibody-like Affinity. *ACS Chem. Biol.* **2007**, *2* (9), 625–634.

(18) Yano, M.; Hoogenraad, N.; Terada, K.; Mori, M. Identification and Functional Analysis of Human Tom22 for Protein Import into Mitochondria. *Mol. Cell. Biol.* **2000**, *20* (19), 7205–7213.

(19) Bellot, G.; Cartron, P.-F.; Er, E.; Oliver, L.; Juin, P.; Armstrong, L. C.; Bornstein, P.; Mihara, K.; Manon, S.; Vallette, F. M. TOM22, a Core Component of the Mitochondria Outer Membrane Protein Translocation Pore, Is a Mitochondrial Receptor for the Proapoptotic Protein Bax. *Cell Death Differ.* **2007**, *14* (4), 785–794.

(20) Bacon, K.; Blain, A.; Bowen, J.; Burroughs, M.; McArthur, N.; Menegatti, S.; Rao, B. M. Quantitative Yeast-Yeast Two Hybrid for Discovery and Binding Affinity Estimation of Protein-Protein Interactions. *bioRxiv* **2020**, *ID# 247874*.

(21) Tavassoli, A. SICLOPPS Cyclic Peptide Libraries in Drug Discovery. *Curr. Opin. Chem. Biol.* **2017**, *38*, 30–35.

(22) 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.

CHAPTER 5

Isolation of Chemically Cyclized Peptide Binders Using Yeast Surface Display

Adapted from Bacon K, Blain A, Burroughs M, McArthrur N, B.M Rao, and Menegatti S. (2020) Isolation of Chemically Cyclized Peptide Binders Using Yeast Surface Display. *ACS Combinatorial Science 22(10)*: 519-532.

5.1 Introduction

Cyclic peptides have emerged as a promising class of affinity ligands for use in basic research as well as diagnostic and therapeutic applications[1–4] owing to their favorable properties such as higher binding affinity[5] and stability relative to linear peptides[6,7]. In particular, they are ideal for inhibiting protein-protein interactions as their small size and modularity allows them to interact efficiently with viable binding sites on protein surfaces[8]. Other attractive features of cyclic peptides include scalable and affordable synthesis[9], metabolic stability[10], and biocompatibility[11]. Additionally, they can be easily modified with labels (*e.g.*, fluorescence or radioactive probes)[12–14] or conjugated to other biomolecules[15] to endow further biochemical functionalities.

Cyclic peptides are commonly isolated from combinatorial libraries that are either chemically synthesized or genetically encoded. Among chemical synthesis-based approaches, one-bead-one-peptide libraries are frequently employed[16]; these libraries are a collection of beads, each displaying multiple copies of a unique peptide variant that can be sequenced by Edman degradation or mass spectrometry for identification[17,18]. As an alternative, genetically-encoded peptide libraries have gained widespread application as high-affinity candidates can be isolated using convergent steps of library selection (*i.e.*, directed evolution). Commonly used methods for generating genetically-encoded peptide libraries include mRNA display[19–21], ribosomal display[22–24], phage-display[25–28], and bacterial display[29,30]. Multiple strategies have been developed for the cyclization and stabilization of displayed peptides, both chemical[31] and enzymatic[32–34], although disulfide bond formation remains the most widely employed[35–37]. Recently, there has been interest in designing cyclic peptides that target intracellular proteins for therapeutic and research applications. Cyclic peptide sequences specific to intracellular targets can be attached to cell penetrating peptide sequences for intracellular delivery[15,38–40]. However, disulfide-containing cyclic peptides are

unlikely to maintain a stable cyclic structure *in vivo* due to the reducing environment of the cytosol[41,42]. Accordingly, a general cyclization method that affords stable bond formation is desired that can be applied to identify cyclic peptides for both intracellular and extracellular application. Alternative routes to stable peptide cyclization, involving the formation of thioether and amide bonds via chemical crosslinking, have been demonstrated with mRNA display[42–45] and phage-display libraries[31,46]. The inclusion of noncanonical amino acids has also been explored to expand the types of chemical crosslinking chemistries available[47–51].

Chemically synthesized and genetically encoded peptide libraries are typically screened using a "panning" method, wherein the library is incubated with a surface that is functionalized with the target protein. Any clones that bind the surface are isolated. However, panning-based selections do not provide an efficient means to discriminate high-affinity binders from low affinity binders[52]. Further, panning methods typically yield pools of peptide sequences whose biochemical characterization (*e.g.*, binding affinity) requires prior chemical synthesis. To overcome these limitations, we describe the use of yeast surface display for the construction of chemically cyclized peptide libraries that can be subjected to fluorescence activated cell sorting (FACS) to efficiently isolate the highest affinity library clones. While frequently employed, phage display libraries are not amenable to sorting via traditional FACS due to the small size of phage particles[53]. Only a few instances are reported of using yeast display to express cyclic peptides[34,54–56], most of which are limited to disulfide-constrained peptides, like knottins[57–60]. However, to our knowledge, chemical crosslinkers have not been commonly used for the cyclization of yeast-displayed peptides, despite their widespread use for cyclizing mRNA, phage, and bacterial display peptide libraries.

In our approach, a linear peptide precursor is displayed on the yeast cell surface as an N-terminal fusion to the Aga2p subunit of the yeast mating protein α-agglutinin. The Aga2p subunit

forms a disulfide bond with the Aga1p subunit of α-agglutinin, thereby tethering the linear peptide to the yeast cell surface. Peptide cyclization is mediated by disuccinimidyl glutarate (DSG) crosslinking between the peptide's N-terminal amine and an ε-amine of an included lysine residue (Figure 5.1). Amino acid residues between a N-terminal methionine and a C-terminal lysine are randomized to generate a combinatorial library. It is important to note that a prepro secretory signal sequence precedes the linear peptide precursor[61–63]. A portion of this sequence directs the nascent polypeptide through the secretion pathway involving the endoplasmic reticulum and the Golgi apparatus[64–68]. Before leaving the Golgi, the secretory sequence is cleaved between its C-terminal lysine and arginine residues[69,70], resulting in an arginine residue at the N-terminus of the displayed peptide. Thus, the surface displayed peptides are crosslinked between the α-amine of the N-terminal arginine and the ε-amine of the lysine residue located at the C-terminus of the peptide sequence. Other components of the displayed peptide fusion include a HA epitope tag to characterize the expression of the fusion; no lysine residues are present in this tag. Whilst Aga2p includes three lysine residues, DSG-mediated crosslinking between these residues and other amine groups within the linear precursor peptide is unlikely due to the large spatial separation between these moieties.

To evaluate the use of yeast surface display for isolating cyclic peptide binders, we first screened a combinatorial library of chemically crosslinked cyclic peptides against a model target protein, lysozyme, using magnetic selection and FACS. We further evaluated if yeast-displayed cyclic peptides can be used to efficiently assess binding selectivity and to obtain quantitative estimates for equilibrium binding dissociation constants (K_D), thereby avoiding the need to chemically synthesize isolated peptides for such evaluation. Finally, we investigated if our

approach could be used to isolate and characterize cyclic peptides that bind human interleukin-17 (IL-17).

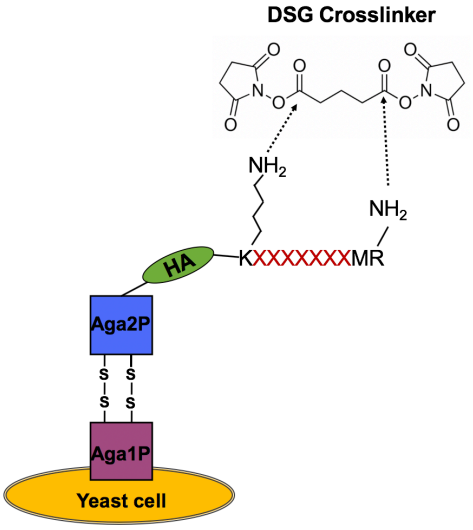


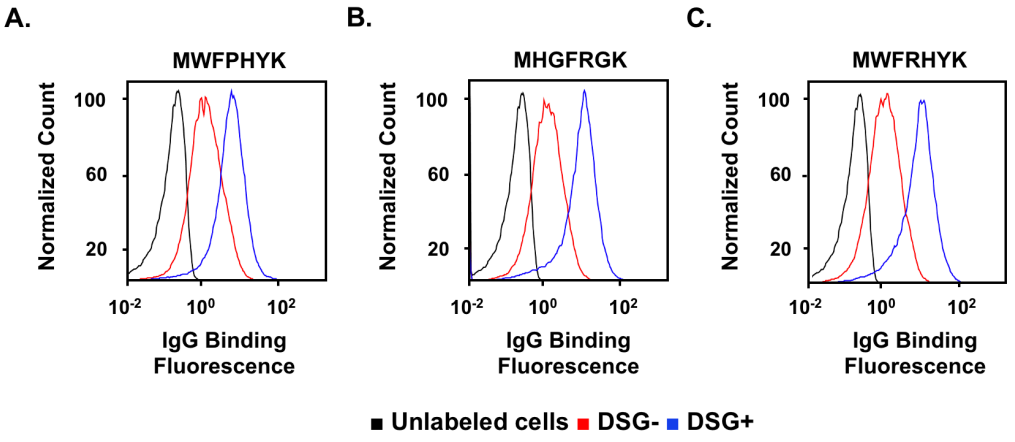
Figure 5.1. Yeast surface display of chemically crosslinked cyclic peptides. A linear peptide sequence is displayed as an N-terminal fusion to Aga2p, which is tethered to the yeast cell surface through a disulfide bond with Aga1p. The displayed peptides are crosslinked between the α-amine of an arginine residue and the ε-amine of a lysine residue, located at the C-terminus of the peptide's variable region, using a disuccinimidyl glutarate (DSG) crosslinker.

5.2 Results and Discussion

5.2.1 Confirmation of linear peptide cyclization on the yeast surface

In our approach, the cyclization of a yeast-displayed linear peptide is mediated by a DSG crosslinker. While previously utilized to cyclize peptide fusions in solution[42,44], to our knowledge, DSG has not been commonly used to cyclize peptide sequences expressed as cell surface fusions. Accordingly, we utilized a protein-binding assay to indirectly assess DSG-mediated cyclization of yeast-displayed linear peptide precursors. It is generally accepted that cyclic peptides exhibit higher target-binding affinity compared to their linear counterparts[71]. We hypothesized that the target protein would bind the DSG-treated cells more efficiently if DSG appropriately cyclized the linear displayed peptides as crosslinked peptide sequences would exhibit a higher binding affinity for the target protein. Consequently, we considered yeast displaying known IgG-binding cyclic

peptide sequences; the peptides cyclo-[*DSG*-MWFPHY-*K*], cyclo-[*DSG*-MHGFRG-*K*], and cyclo-[*DSG*-MWFRHY-*K*] had been selected in prior work to bind IgG[44]. We compared the binding of IgG to yeast cells displaying these peptide sequences with or without DSG treatment. To promote complete cyclization of all linear displayed peptides, DSG was incubated 550-fold over the assumed total amount of peptides displayed on the yeast surface (~50,000 per cell[72,73]) in two consecutive crosslinking reactions. Subsequently, DSG-treated and untreated cells were incubated with biotinylated IgG, followed by immunofluorescent detection using Streptavidin R-phycoerythrin (SA-PE) (Figure 5.2). In comparison to the non-treated cells, the binding of IgG to DSG-treated cells was significantly higher, as evaluated by the mean fluorescence signal. This difference in binding is likely attributed to DSG appropriately crosslinking the linear displayed peptides resulting in higher affinity for IgG. It is important to note that we observed detectable binding of IgG at high labeling concentrations (2 μM and 5 μM *vs.* 450 nM for studies in Figure 5.2) to DSG-treated EBY100 yeast not displaying the peptides (Figure 5.3). However, this non-specific binding was significantly lower than the binding of IgG to DSG-treated cells displaying the IgG-binding peptides (Figure 5.2). In addition, other proteins (*i.e.*, BSA, lysozyme) bound to a negligible extent to DSG-treated EBY100 yeast (Figure 5.3).



D.

Mutant	Mean Fluorescence Unlabeled Cells	Mean IgG Binding Fluorescence DSG-	Mean IgG Binding Fluorescence DSG+
MWFPHYK	0.17 (0.003)	1.34 (0.12)	4.91 (0.48)*
MHGFRGK	0.21 (0.006)	1.28 (0.021)	8.13 (0.31)*
MWFRHYK	0.18 (0.003)	1.30 (0.043)	8.30 (0.44)*

Figure 5.2. Evaluation of disuccinimidyl glutarate-mediated crosslinking of linear peptides displayed as yeast surface fusions. The binding of IgG to yeast cells displaying a linear peptide sequence that were subjected to DSG crosslinking (blue) or not subjected to DSG crosslinking (red) is compared to an unlabeled control (no IgG, black). Three cyclic peptide sequences with known binding to IgG were tested: **(A)** cyclo-[*DSG*-MWFPHY-*K*], **(B)** cyclo-[*DSG*-MHGFRG-*K*], and **(C)** cyclo-[*DSG*-MWFRHY-*K*]. Representative data from three independent experiments is shown. **(D)** Mean fluorescence values describing IgG binding to yeast cells displaying a linear peptide sequence that were treated or not treated with DSG are reported. The mean fluorescence of the unlabeled cell (no IgG) control is also shown. The mean fluorescence values are not background subtracted. The reported mean fluorescence values represent the average of three independent repeats while the values in parentheses represent the standard error of the mean for the repeats. * represents statistical significance ($p<0.05$) for IgG binding compared to the non-DSG treated cells using a two-tailed, paired t-test.

While these results suggest that DSG can be used to cyclize yeast-displayed linear peptides containing six residues between the N and C terminal linking residues, it may be desirable to generate cyclic peptides of different sizes. Accordingly, we explored the impact of peptide length on cyclization through *in silico* studies (Figure 5.4) using the IgG binding peptide sequence RMWFPHYK as a model cyclic peptide template. An arginine residue was included at the N-terminus, as all yeast-displayed peptides contain this residue. Variations of this peptide were designed using the molecular editor Avogadro[74,75] and equilibrated using atomistic molecular dynamics (MD) simulations. Amino acids were eliminated in the N-to-C terminal direction along

the peptide sequence to reduce the peptide length, or a glycine residue was added to the N-terminus to elongate the peptide sequence. The resultant free energy of each peptide structure cyclized by DSG was estimated via MD simulations using the OPLS all atom force field[76–78]. Figure 5.4G suggests that the peptide free energy normalized against the peptide length does not significantly vary with the number of amino acids for the peptide sequences considered herein. It is important to note that peptides containing different amino acid residues will likely cyclize with different free energies. Nevertheless, these studies suggest DSG-mediated cyclization may be reliably used for generating cyclic peptides of varying lengths.

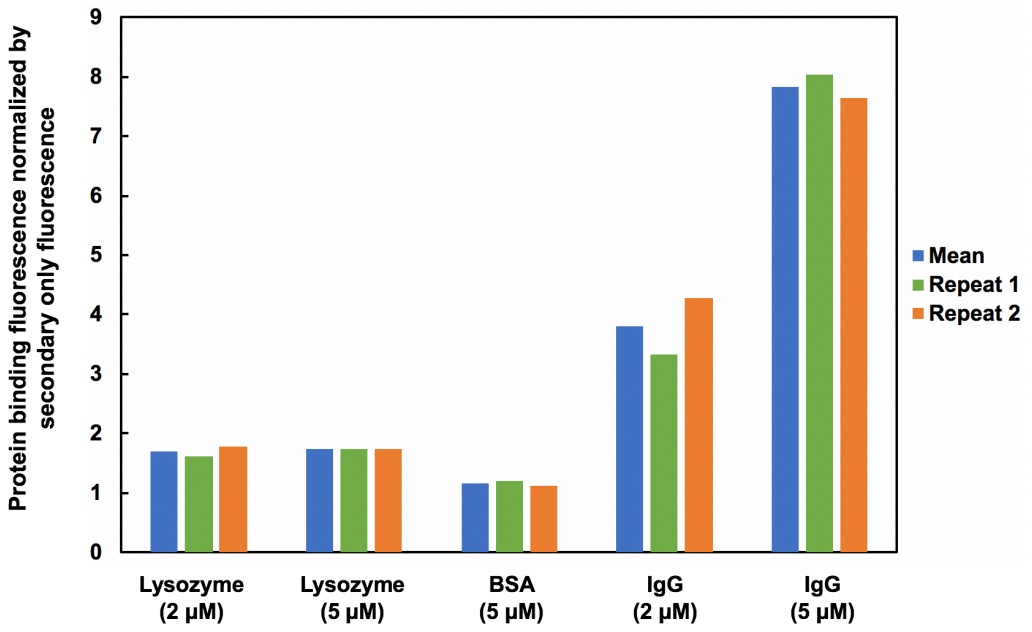


Figure 5.3. Detection of different biotinylated proteins (lysozyme, bovine serum albumin (BSA), and human IgG) interacting with disuccinimidyl glutarate-treated, non-displaying EBY100 at varying concentrations. The binding of each biotinylated protein was detected by flow cytometry after SA-PE labeling. Each mean fluorescence value associated with protein binding was normalized by dividing the mean fluorescence of SA-PE binding to DSG-treated EBY100 cells. The normalized mean fluorescence values of protein binding were averaged across two independent repeats (blue). Error bars represent the standard deviation of the two independent repeats. The mean fluorescence of protein binding for each repeat is also shown (green and orange).

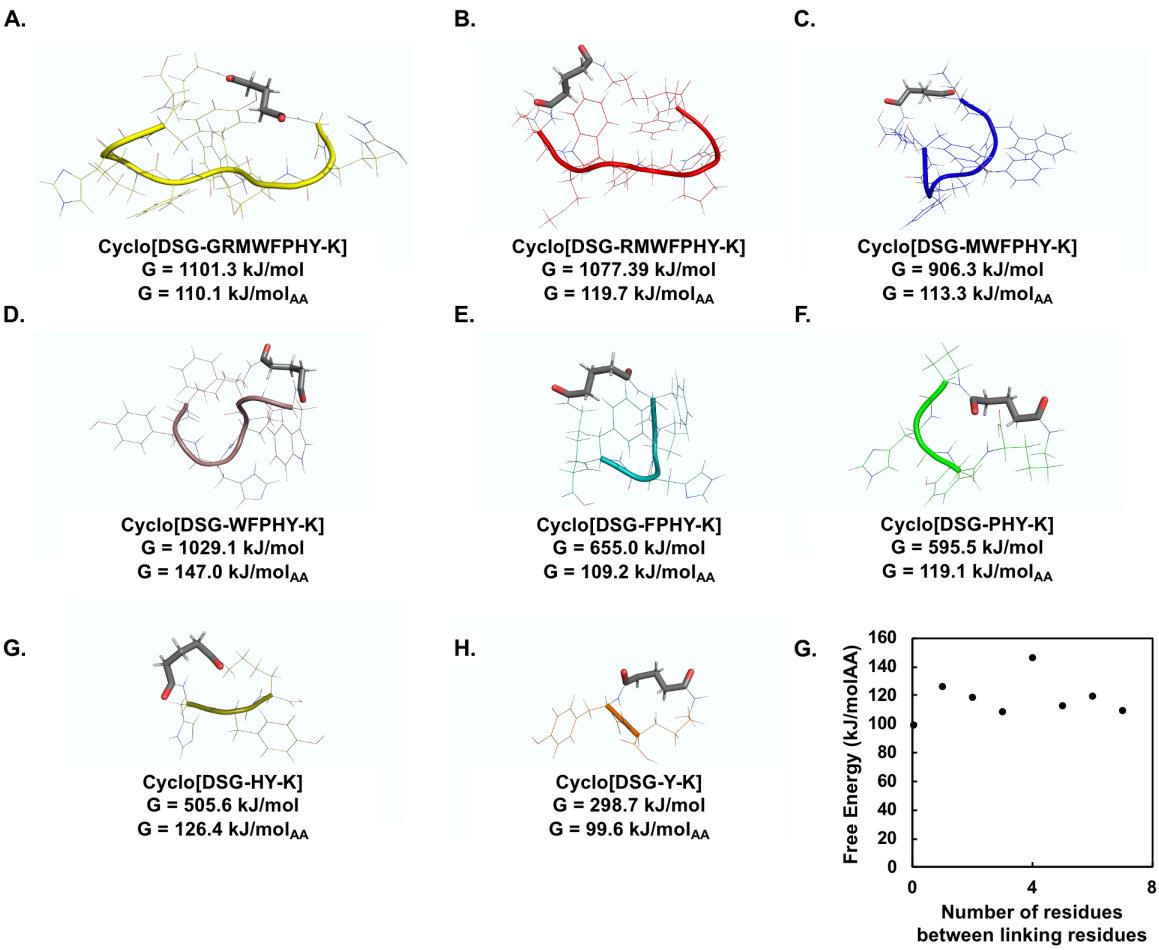


Figure 5.4. *In silico* structure of peptide sequences of varying lengths cyclized using disuccinimidyl glutarate as predicted by molecular dynamics. Each peptide in **A-H** is based on the starting sequence RMWFPH. The peptides are crosslinked by DSG between their N-terminal residue and a C-terminal lysine (denoted as linking residues). **(G)** The peptide free energy per mole of amino acids are compared for the different sized peptides. The peptide free energy does not significantly differ between the different sized cyclic peptides, suggesting DSG cyclization is feasible for each of the peptide lengths considered in this work based on the starting sequence RMWFPHY.

5.2.2 Isolation of lysozyme-binding cyclic peptides using yeast surface display

To identify novel cyclic peptide binders, we generated a yeast-display combinatorial library of linear peptide precursors with a diversity of ~ $5x10^7$ variants as estimated by plating of the transformed yeast. The library variants were encoded as $RMX_1X_2X_3X_4X_5X_6X_7K$, where X is a randomized amino acid residue. It is worth noting that the methionine that precedes the variable residue region is not required. We chose to include a methionine residue near the N-terminus since it is present in the IgG-binding peptide sequences used to confirm cyclization of surface displayed

linear peptides. Yeast-displayed linear peptides were cyclized using DSG as described earlier. We initially screened the library against a model target protein, lysozyme, using one round of magnetic selection against lysozyme-coated magnetic beads, followed by two rounds of FACS. Interestingly, we observed that simultaneous labeling of yeast cells with both lysozyme and an anti-HA antibody resulted in a decrease of lysozyme binding, relative to the case when the anti-HA antibody was not present (Figure 5.5). The binding of lysozyme to the yeast-displayed cyclic peptides is likely affected by steric hindrance if the anti-HA antibody is bound. Therefore, unlike typical protocols used for the FACS selection of yeast display libraries, the cyclic peptide library was solely labeled with the target protein for sorting.

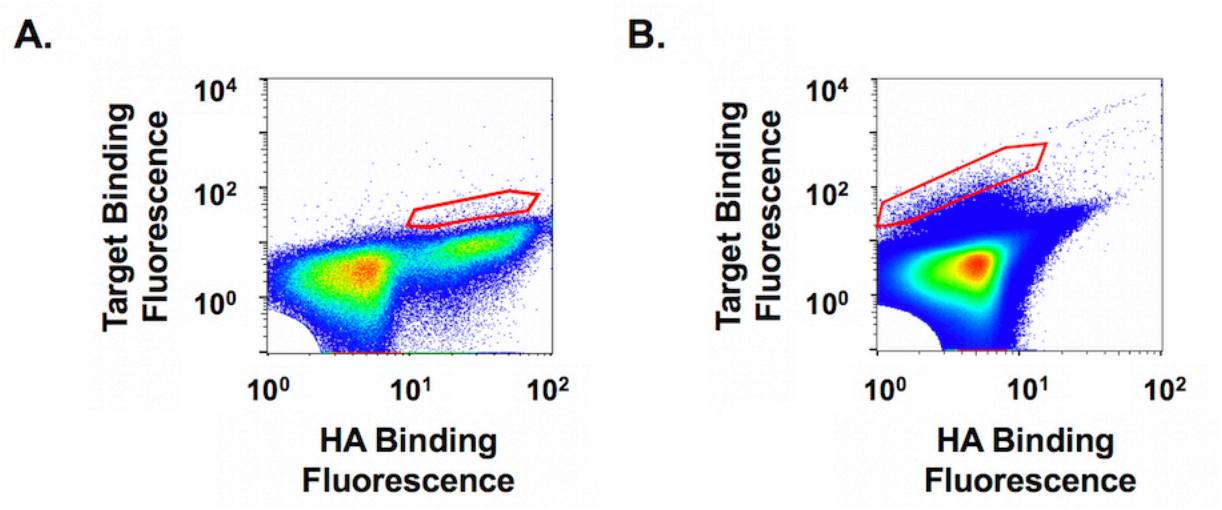


Figure 5.5. Comparison of single and double labeled populations when sorting a yeast displayed cyclic peptide library via fluorescence activated cell sorting. Population **(A)** was double labeled for target binding and HA expression while population **(B)** was single labeled for target binding only when performing fluorescence activated cell sorting.

After FACS, plasmid DNA was isolated from 13 different colonies and sequenced (Table 5.1). Sequence homology, especially the presence of arginine (R) at position X_7, and the enrichment in cationic (R and K), aromatic (W, F, and Y), and hydrogen-bonding (Q and S) residues is apparent throughout the selected sequences. The cationic character of the sequences is

surprising given that lysozyme is a basic protein (pI ~ 11[79]). It is noted that one of the sequences (Lys.7) contains multiple lysine residues within the randomized residue segment. Lysine residues within the randomized segment introduce additional amine groups that are available for crosslinking via DSG. Consequently, we cannot predict the exact cyclization of peptides containing multiple lysine residues. The surface of yeast cells expressing these peptide sequences will likely feature a heterogenous population of peptides with varying crosslinking patterns, making characterization using yeast surface display difficult. Therefore, we chose to not evaluate peptide Lys.7 further. Similarly, the peptide sequence Lys.6 was abandoned due to the presence of arginine in five out of the seven randomized residues positions.

Table 5.1. Lysozyme binding peptides isolated from a yeast-displayed library of cyclic heptapeptides. Mutagenized positions are denoted with an X. Numbers in parentheses represent the number of times a sequence occurred.

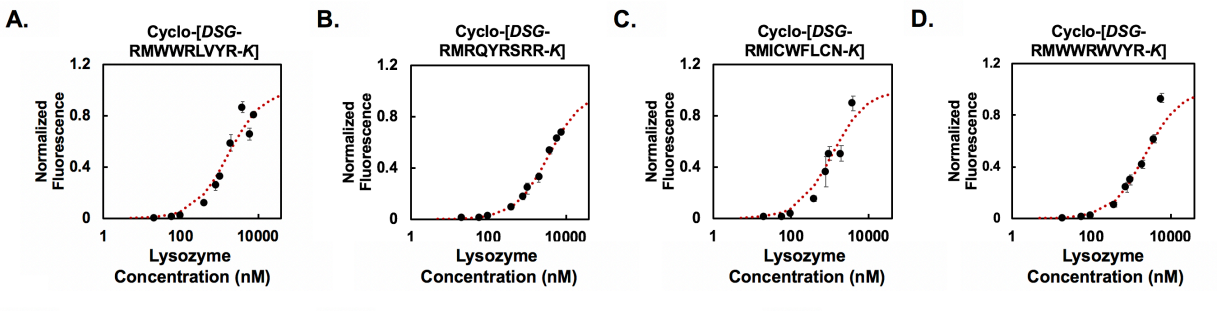
	M	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**	**X_6**	**X_7**	**K**
LYS.1 (4)		W	W	R	W	V	Y	R	
LYS.2 (3)		I	C	W	F	L	C	N	
LYS.3 (2)		R	Q	Y	R	S	R	R	
LYS.4		W	W	R	L	V	Y	R	
LYS.5		R	R	Q	G	L	R	R	
LYS.6		R	R	R	R	P	V	R	
LYS.7		K	R	F	K	T	R	R	

5.2.3 Binding affinity analysis of lysozyme cyclic peptides

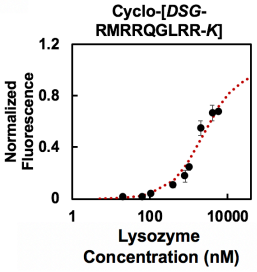
Lysozyme binding isotherms were generated for yeast cells displaying cyclo-[*DSG*-RMWWRWVYR-*K*], cyclo-[*DSG*-RMICWFLCN-*K*], cyclo-[*DSG*-RMRQYRSRR-*K*], cyclo-[*DSG*-RMWWRLVYR-*K*], and cyclo-[*DSG*-RMRRQGLRR-*K*]. Yeast cells displaying these cyclic peptides were incubated with varying concentrations of biotinylated lysozyme (20 nM - 8 μM). The fraction of cell surface fusions bound by biotinylated lysozyme was quantified via

immunofluorescent detection of SA-PE. The resulting data were fit using a monovalent binding isotherm (Figures 5.6A-E) to estimate the apparent binding affinity ($K_{D,Yeast}$) of the yeast-displayed cyclic peptides for lysozyme (Figure 5.6F). It is important to note that the Hook effect was observed at high concentrations of lysozyme (> 8 μM). This phenomenon, where the fluorescence corresponding to target binding decreases at high target concentrations, has been observed previously when performing yeast surface titrations to estimate binding affinity[80]. Consequently, fluorescence values at concentrations greater than 8 μM were not included when calculating the apparent binding affinity ($K_{D, Yeast}$).

Cyclo-[*DSG*-RMWWRLVYR-*K*] exhibited the highest affinity for lysozyme as estimated via yeast surface titration. Accordingly, this cyclic peptide was chemically synthesized and characterized via isothermal titration calorimetry (ITC). The binding affinity of cyclo-[*DSG*-MWWRLVYR-*K*] for lysozyme, as measured by ITC (2996 nM +/-1410 nM, Figure 5.7), is of a similar magnitude (low micromolar) as the binding affinity predicted via yeast surface titration (862 – 1641 nM, 68% confidence interval) (Figure 5.6G). The minor difference in the predicted binding affinity between the two methods may be attributed to non-affinity binding interactions (*e.g.*, lysozyme-lysozyme, peptide-peptide, and lysozyme-peptide binding) that can occur during ITC titrations due to high concentrations of titrant (500 μM of peptide) and titrand (50 μM of lysozyme), thereby resulting in a lower predicted binding affinity between the titrant and titrand when performing ITC[81]. Alternatively, the binding affinity of the chemically synthesized peptide may be slightly different from the yeast-displayed peptide due to the absence of the N-terminal arginine residue in the former. The N-terminal arginine residue present in the yeast-displayed peptide was not included in the chemically synthesized peptide, as arginine residues contain bulky side chains that can negatively impact cyclization efficiency[82].



E.



F.

Mutant	$K_{D,Yeast}$ (nM)	68% CI (nM)
Cyclo-[*DSG*-RMWWRLVYR-*K*]	1186	862 – 1641
Cyclo-[*DSG*-RMRQYRSRR-*K*]	1728	1375 - 2175
Cyclo-[*DSG*-RMICWFLCN-*K*]	2332	1958 - 2791
Cyclo-[*DSG*-RMWWRWVYR-*K*]	2391	2058 - 2784
Cyclo-[*DSG*-RMRRQGLRR-*K*]	3662	3227 - 4176

G.

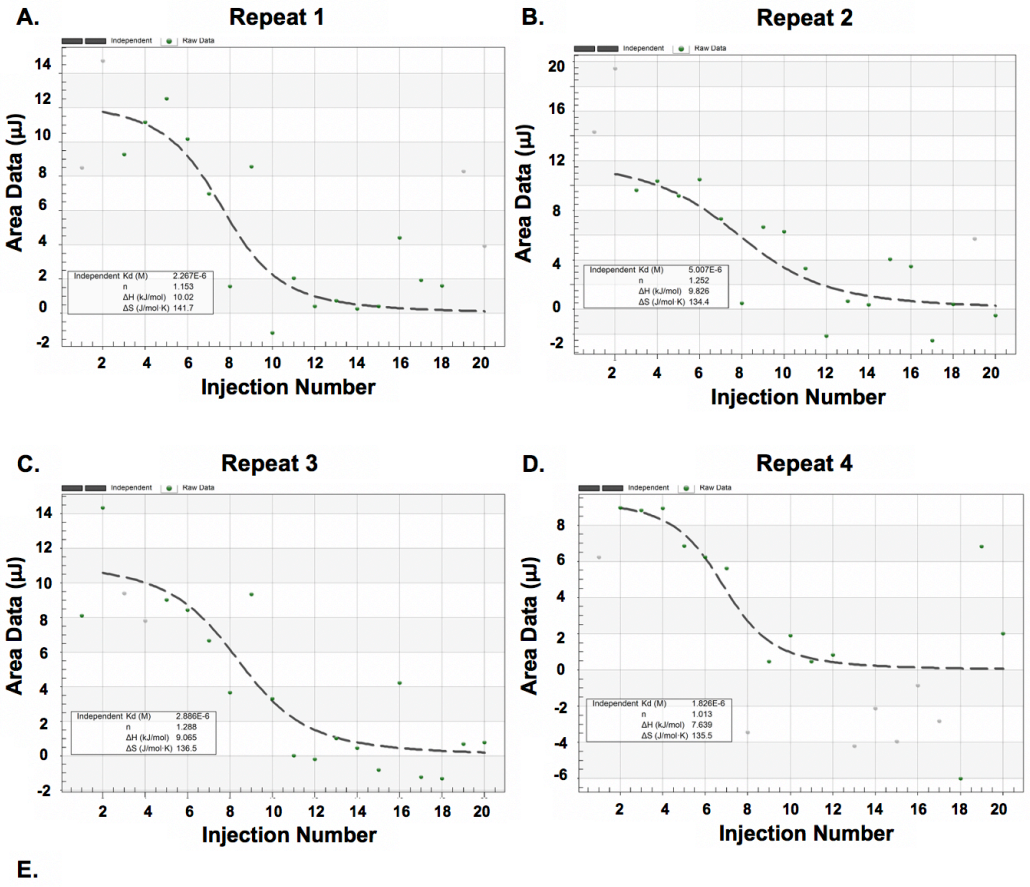
Mutant	$K_{D,Yeast}$ (nM)	$K_{D,ITC}$ (nM)
Cyclo-[*DSG*-MWWRLVYR-*K*]	1186 (862 – 1641)	2996 (1586 – 4406)

Figure 5.6. Estimation of apparent K_D describing the affinity of lysozyme for yeast-displayed cyclic peptides. The cyclic peptides considered include **(A)** cyclo-[*DSG*-RMWWRLVYR-*K*], **(B)** cyclo-[*DSG*-RMRQYRSRR-*K*], **(C)** cyclo-[*DSG*-RMICWFLCN-*K*], **(D)** cyclo-[*DSG*-RMWWRWVYR-*K*], and **(E)** cyclo-[*DSG*-RMRRQGLRR-*K*]. A normalized mean fluorescence describing lysozyme binding to the yeast-displayed cyclic peptides as a function of the concentration of lysozyme incubated is reported. Error bars represent the standard error of the mean for three independent experiments. **(F)** A global, non-linear regression was used to estimate $K_{D,Yeast}$ describing the apparent affinity between the yeast-displayed cyclic peptides and lysozyme as well as the corresponding 68% confidence interval (CI). The reported mean fluorescence values associated with lysozyme binding were first background subtracted and then normalized relative to the maximum fluorescence values fit by the non-linear regression **(A-E)**. The mean background fluorescence values originate from the signals produced after incubation of the cells with SA-PE alone. **(G)** The binding affinity ($K_{D,ITC}$) of soluble cyclo-[*DSG*-MWWRLVYR-*K*] for lysozyme was also estimated using isothermal calorimetry and compared to $K_{D,Yeast}$. In parentheses, a 68% confidence interval is provided for $K_{D,Yeast}$ while a range representing the standard deviation of the mean for four independent experiments is described for $K_{D,ITC}$.

Overall, these results suggest that yeast surface titrations can be used to reasonably estimate the binding affinity of cyclic peptides, similar to proteins[72]. Quantitative assessment of binding affinity using yeast-displayed cyclic peptides is particularly attractive as this method can be high throughput and circumvents the need to chemically synthesize peptides.

The low micromolar affinity of the lysozyme binding peptides was anticipated based on the affinity of previously identified cyclic peptides that are similar in size[83]. Other groups, using display-based technologies, have identified cyclic peptides with low nanomolar affinities[84]; these

cyclic peptides, however, feature a larger ring structure (up to 14 randomized amino acid positions), and therefore benefit from a higher enthalpic contribution to the binding free energy. In other instances, the peptide macrocycle is constrained via multifunctional crosslinkers to increase both the enthalpic and entropic contributions of binding free energy[85]. While larger ringed structures likely exhibit increased binding affinity due to enthalpic gains, the efficiency of yeast transformation limits library size, and, consequently, the fraction of the theoretical diversity that can be captured in a yeast library; combinatorial libraries generated using yeast surface display typically do not contain more than 10^{10} variants[86]. This limitation was highlighted by Sieber et al[87]. who showed that a heptapeptide library using NNK codon degeneracy with a library size of 10^8 will only contain ~5.6×10^7 unique sequences, corresponding to ~6.3% of the theoretical diversity for a peptide this size. Further, chemical synthesis of cyclic peptides becomes more challenging as the number of amino acid residues increases and/or additional macrocycles are included. Due to these aforementioned considerations, we chose to focus our analysis on screening heptapeptide libraries. We anticipate monocyclic peptides of this size will be useful for applications where low micromolar affinities are desired, like affinity chromatography.



E.

	Kd (nM)	N	ΔH (kJ/mol)	ΔS (J/mol·K)
Cyclo-[*DSG*-MWWRLVYR-*K*]	2996 +/- 1410	1.18 +/- 0.35	9.14 +/- 22.33	137.02

Figure 5.7. Isothermal titration calorimetry profiles describing the binding of cyclo-[*DSG*-MWWRLVYR-*K*] to lysozyme. The integrated heats of binding obtained from the raw data after subtracting the heat of dilution are shown in **(A-D)** for each repeat. The dashed line represents the best curve fit to the experimental data using the Independent model. Associated fitted and calculated parameters are described in **(E)**. These values as well as the standard deviations presented were averaged across the four repeats.

5.2.4 Selectivity of cyclic peptides for lysozyme and effect of peptide cyclization on binding affinity

To assess the binding selectivity of the lysozyme-targeting peptides, we incubated yeast cells displaying the isolated cyclic peptides with varying concentrations (0.1 - 2 µM) of lysozyme or bovine serum albumin (BSA). Lysozyme binding was significantly higher than BSA binding to the yeast cells displaying each cyclic peptide (Figures 5.8A-E), demonstrating selectivity of the

cyclic peptides for lysozyme over BSA. Even at higher concentrations (5 μM), lysozyme bound at a higher level than BSA to the yeast-displayed cyclic peptides (Figure 5.9). BSA was adopted as a model protein to test binding selectivity owing to its anionic character (pI ~ 4.7)[88] and because it was included in the buffers used during library screening.

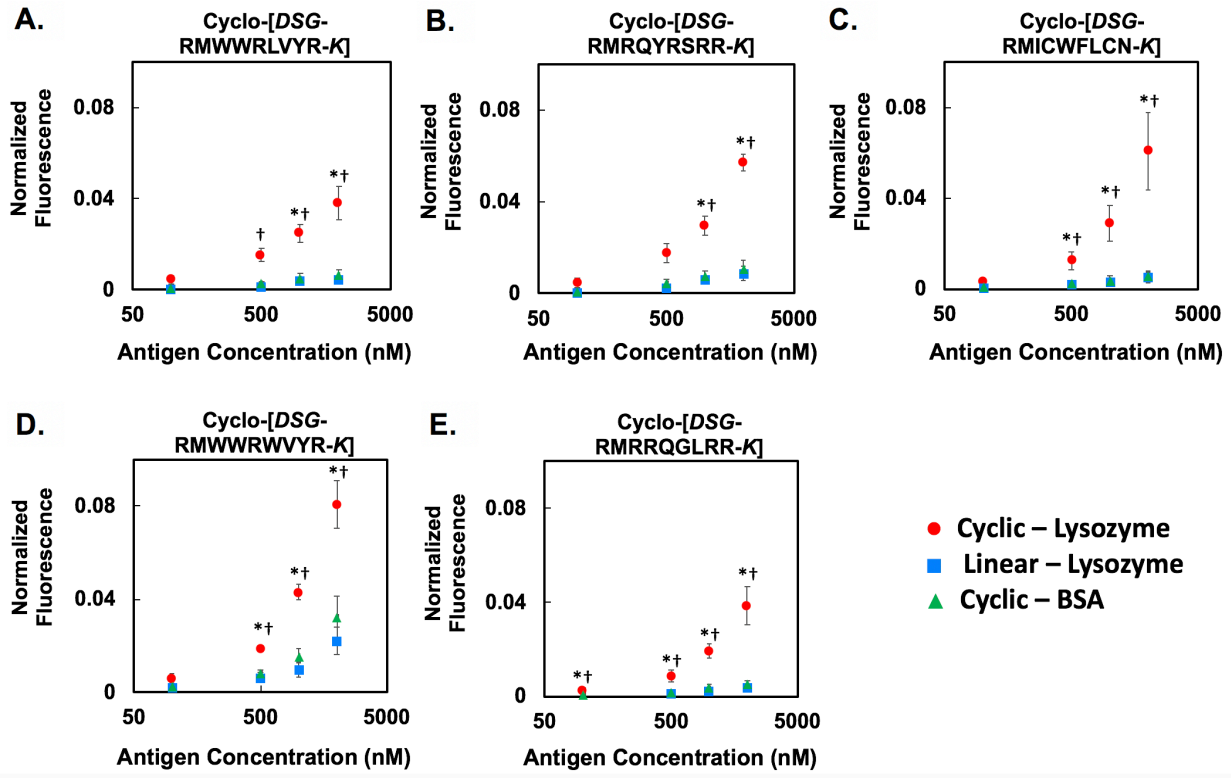


Figure 5.8. Specificity analysis of lysozyme binding cyclic peptides. The normalized mean fluorescence associated with lysozyme binding to yeast cells displaying cyclic peptides (red), **(A)** cyclo-[*DSG*-RMWWRLVYR-*K*], **(B)** cyclo-[*DSG*-RMRQYRSRR-*K*], **(C)** cyclo-[*DSG*-RMICWFLCN-*K*], **(D)** cyclo-[*DSG*-RMWWRWVYR-*K*], and **(E)** cyclo-[*DSG*-RMRRQGLRR-*K*], was compared to the normalized mean fluorescence associated with lysozyme binding to yeast cells displaying their linear counterparts (blue). Additionally, for a given yeast-displayed cyclic peptide, the normalized mean fluorescence associated with lysozyme binding (red) was compared to the normalized mean fluorescence associated with BSA binding (green). Binding of each biotinylated protein was detected by flow cytometry after SA-PE labeling. Error bars represent the standard error of the mean for three independent experiments. For a given lysozyme concentration, * represents statistically significant lysozyme binding ($p < 0.05$) for a yeast-displayed cyclic peptide in comparison to its linear counterpart. For a given antigen concentration, † represents statistically significant lysozyme binding ($p < 0.05$) in comparison to the binding of BSA for a given yeast-displayed cyclic peptide. A two-tailed, paired t-test was performed to evaluate statistical significance. All mean fluorescence values describing protein binding were normalized by first subtracting the mean background fluorescence and then dividing by the background subtracted mean fluorescence associated with HA expression. The mean background fluorescence values are derived from the signals produced after incubation of the secondary detection agents (SA-PE or DAR488) alone.

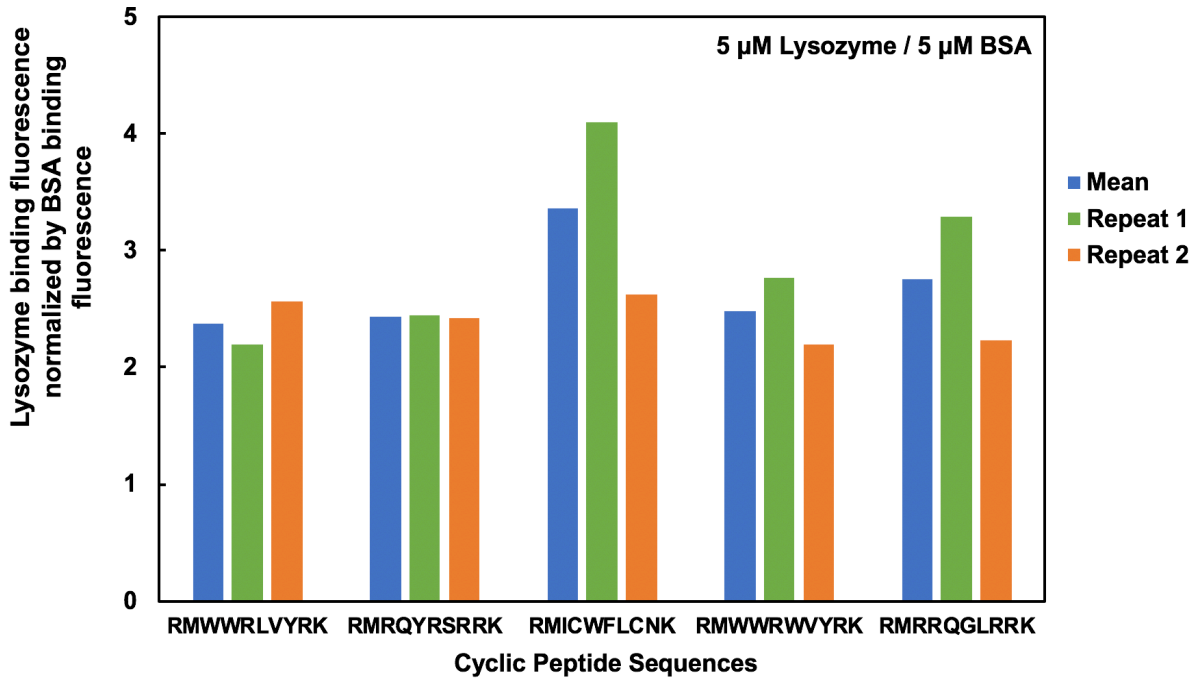


Figure 5.9. Selectivity analysis of cyclic peptides for lysozyme compared to bovine serum albumin. Yeast cells displaying each lysozyme binding cyclic peptide were labeled with either 5 µM biotinylated lysozyme or 5 µM biotinylated bovine serum albumin (BSA). Binding of each protein was detected by flow cytometry after SA-PE labeling. For each mutant, the background subtracted mean fluorescence value of lysozyme binding was normalized by dividing the background subtracted mean fluorescence value of BSA binding. The mean background fluorescence originates from the signals produced after incubating the cyclic peptide displaying cells with SA-PE alone. The normalized mean fluorescence values were averaged across two independent repeats (blue). Error bars represent the standard deviation of the two independent repeats. The normalized mean fluorescence for each repeat is also shown (green and orange).

We also tested other proteins to further characterize the binding selectivity of the isolated lysozyme binding cyclic peptides. At a concentration of 2 and 5 µM, green fluorescent protein (GFP, pI ~ 5)[89] was found to bind to a similar extent to cells expressing the lysozyme binding cyclic peptides and DSG-treated, non-displaying EBY100 cells, suggesting that the lysozyme binding cyclic peptides have no significant affinity for GFP (Figure 5.10). This is noteworthy due to the anionic nature of GFP and the enrichment of basic residues in the selected peptides. As discussed earlier (Figure 5.3), detectable binding of human IgG to DSG-treated, non-displaying EBY100 cells was observed at 2 µM and 5 µM IgG, whereas lysozyme bound negligibly at these

concentrations. IgG also interacts with cells displaying the lysozyme-targeting cyclic peptides, producing a signal above background, when labeling at 2 μM and 5 μM IgG (Figure 5.11). However, at a concentration of 5 μM, the cyclic peptides are selective for binding to lysozyme over IgG as the background subtracted signal of lysozyme binding to yeast cells displaying each cyclic peptide was higher than the corresponding signal produced after IgG labeling (Figure 5.11B). This selectivity is notable given that IgG is known to have an increased propensity for non-specific binding due to ionic and hydrophobic interactions[90].

Nevertheless, the detectable interaction of the identified lysozyme binding cyclic peptides with IgG (Figure 5.11), a protein that the library was not selected against, highlights an important limitation of using monocyclic peptides as affinity ligands. Cyclic peptides are likely to be less specific than binders derived from protein scaffolds, such as antibodies, due to their smaller size and simple structure. Unlike antibodies, which typically form multiple points of contact with a target protein, thereby providing greater specifcity[83], monocyclic peptides are likely to form only a single continuous point of contact. Consequently, monocyclic peptides are less likely to distinguish between proteins that have similar regions of amino acids. Therefore, negative selections are critical when screening peptide combinatorial libraries to ensure desired peptide selectivity against particular proteins. Note that the selected peptides, despite their strong cationic and amphiphilic character, showed no or minimal binding to BSA, the only protein we performed a negative selection against. Collectively, these results show that with appropriate negative selection steps, selective cyclic peptide binders can be isolated from a yeast display library of chemically modified peptides.

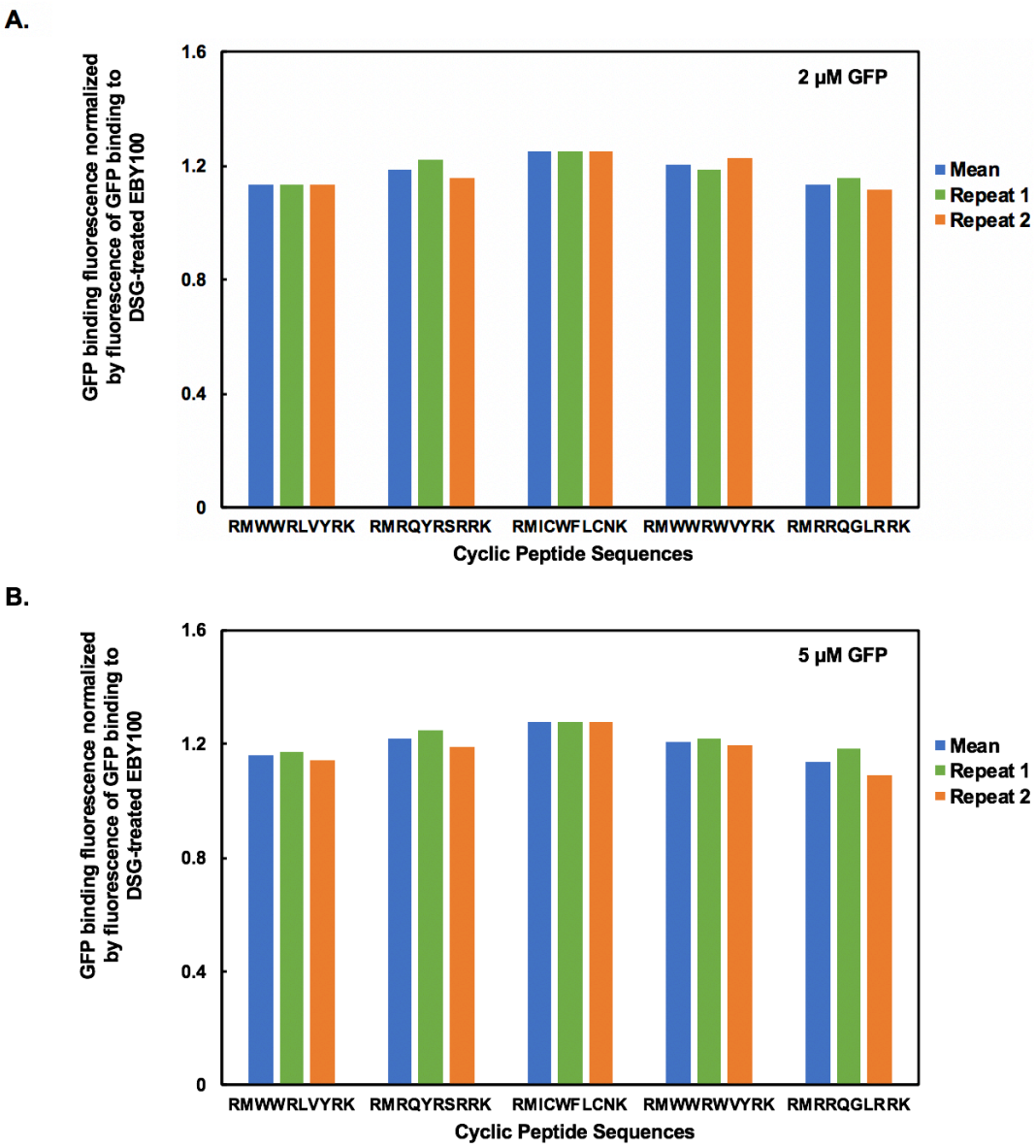


Figure 5.10. Interaction of a non-specific protein, green fluorescent protein, with yeast cells displaying the lysozyme binding cyclic peptides. Yeast cells displaying each lysozyme binding cyclic peptide as well as DSG-treated, non-displaying EBY100 cells were labeled with 2 μM **(A)** or 5 μM green fluorescent protein (GFP) **(B)**. The binding of GFP to each yeast population was analyzed using flow cytometry. The mean fluorescence associated with GFP binding to the cyclic peptide displaying yeast cells was normalized by dividing by the mean fluorescence associated with GFP binding to DSG-treated EBY100 cells. The normalized mean fluorescence values were averaged across two independent repeats (blue). Error bars represent the standard deviation of the two independent repeats. The normalized mean fluorescence for each repeat is also shown (green and orange).

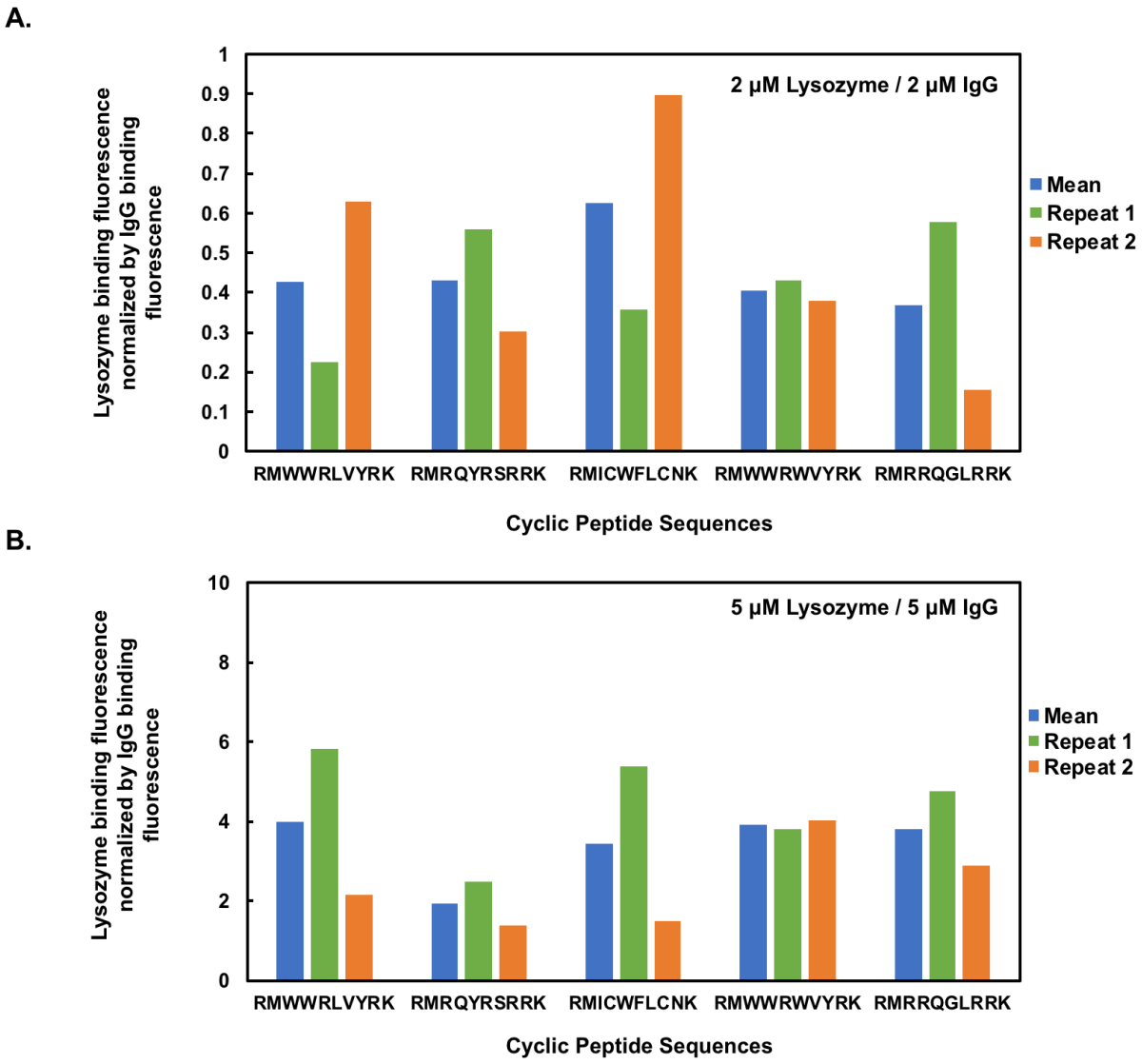


Figure 5.11. Selectivity analysis of cyclic peptides for lysozyme compared to human IgG. Yeast cells displaying each lysozyme binding cyclic peptide were labeled with biotinylated lysozyme or biotinylated IgG at a concentration of 2 μM **(A)** or 5 μM **(B)**. DSG-treated, non-displaying EBY100 cells were also labeled in a similar manner. Binding of each protein was detected by flow cytometry following SA-PE labeling. For each mutant, the background subtracted mean fluorescence value associated with lysozyme binding was normalized by dividing the background subtracted mean fluorescence value associated with IgG binding. For each respective protein and concentration tested, the mean background fluorescence is derived from the signals produced after incubating DSG-treated EBY100 cells with that respective concentration of protein. The normalized mean fluorescence values were averaged across two independent repeats (blue). Error bars represent the standard deviation of the two independent repeats. The normalized mean fluorescence for each repeat is also shown (green and orange).

We also evaluated the effect of peptide cyclization on lysozyme binding activity by comparing the extent of lysozyme binding (0.1 - 2 μM) to yeast cells expressing either the linear or cyclized form of each peptide. Lysozyme binding was significantly higher to the yeast cells

displaying the DSG-treated, cyclic form of each peptide (Figures 5.8A-E) for concentrations as high as 2 μM. This indicates that DSG-mediated cyclization critically confers binding activity to the selected sequences, and that the biorecognition activity of cyclized peptides cannot be inferred from their linear counterparts alone.

5.2.5 Discovery of IL-17-binding cyclic peptides

To demonstrate broader applicability, we screened the previously described cyclic heptapeptide library against Interleukin-17A (IL-17), a proinflammatory cytokine. The interaction of IL-17 with its receptor, IL-17RA, results in the stimulation of inflammatory signaling. Several autoimmune diseases, like plaque psoriasis and psoriatic arthritis, are characterized by excessive levels of IL-17[91]. Cyclic peptides specific to IL-17 can potentially be used as labeling agents[92] or inhibitors of the IL-17:IL-17RA interaction[93].

Cyclic peptide binders specific to IL-17 were isolated from the yeast display heptapeptide library using magnetic selection and FACS. The magnetic selection was performed using magnetized yeast cells displaying IL-17 in lieu of target-functionalized magnetic beads[94]. The use of a yeast-displayed target circumvents the need for purified target protein. The population resulting from FACS was subjected to Illumina Next Generation Sequencing (NGS). The isolated sequences (Table 5.2) are enriched in hydrophobic and basic residues, especially tryptophan (W) and arginine (R). The latter was not anticipated given the cationic character of IL-17 (sequence-based pI of 8.82).

We also performed NGS on the populations obtained after each round of screening against IL-17. As expected, the majority of the sequenced clones contained 7 randomized amino acid positions. There did not appear to be any bias for particular amino acid residues in the sequences

obtained after MACS (Figure 5.12A). However, the screened populations became enriched in hydrophobic and basic residues after FACS (Figures 5.12B-C).

Table 5.2. Sequences of IL-17 binders isolated from a yeast-displayed library of cyclic heptapeptides. Mutagenized positions are denoted with an X. Numbers in parentheses represent the number of times a sequence appeared during Illumina Next Generation Sequencing.

	M	X_1	X_2	X_3	X_4	X_5	X_6	X_7	K
IL17.1 (4191)		R	W	L	R	G	R	R	
IL17.2 (3446)		L	F	H	W	P	F	P	
IL7.3 (2914)		S	Y	F	W	R	W	L	
IL17.4 (2906)		A	W	E	W	L	W	K	
IL17.5 (2742)		L	F	F	D	W	L	L	
IL17.6 (2324)		R	V	P	W	L	W	L	
IL17.7 (2099)		W	R	N	F	A	W	R	
IL17.8 (2030)		F	P	W	V	Q	W	R	
IL17.9 (1895)		W	L	N	D	L	W	R	
IL17.10 (1891)		F	L	W	H	F	R	T	

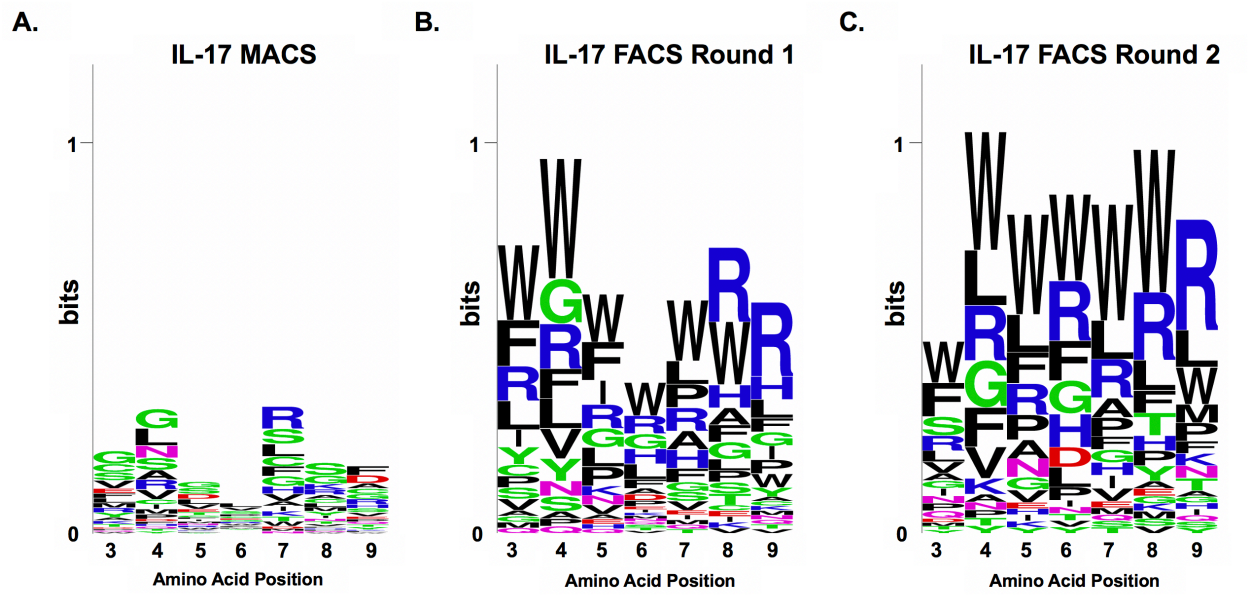


Figure 5.12. Evalution of peptide sequences returned after screening a yeast-displayed cyclic peptide library against IL-17. The distribution of amino acid redues at each position in peptide RMXXXXXXXK is described for the mutants isolated from the **(A)** MACS, **(B)** FACS Round 1, and **(C)** FACS Round 2 selections. The top 50 hits obtained from Ilumina Next Generation Sequencing were analyzed. Each sequence logo was generated using WebLogo.

5.2.6 *In silico* analysis of IL-17:cyclic peptide complexes

Given the strong amphiphilic nature of the selected peptides, we used the hydropathy index (GRAVY) to select specific peptides for *in silico* analysis from the NGS results[95]. We chose peptides that were likely to be water soluble and exhibit low-nonspecific binding resulting from potential hydrophobic interactions (GRAVY < -0.5). Ultimately, seventeen selected peptides (Table 5.3) were chosen to be docked against IL-17 (PDB ID: 4HSA)[96]. It is important to note that the simulated peptide sequences lacked a N-terminal arginine residue, as peptides lacking this residue would likely be easier to chemically synthesize. In preparation for the docking studies, the crystal structure of each peptide was generated via atomistic MD simulation, and the structure of IL-17 was prepared using Schrödinger's Protein Preparation Wizard[97,98]. After, the peptide structures were docked against IL-17 using HADDOCK[99–101]. The IL-17-binding linear peptide IHVTIPADLWDWINK[93] and cyclo-[*DSG*-MGSGGGGSG-*K*] were also docked against IL-17 as positive and negative controls, respectively. Subsequently, representative IL-17:peptide complexes selected from the top docking clusters were further refined via MD simulations and analyzed using the MM/GBSA method to estimate the free energy of binding (ΔG_b) and calculate the corresponding values of affinity ($K_{D,in\ silico}$) (Table 5.4)[102,103]. Twelve of the tested cyclic peptides were predicted to form stable complexes with IL-17, featuring $K_{D,in\ silico}$ ~ 23 nM – 1 **μ**M (Figure 5.13 and Figure 5.15A). Each of these peptides docked at the interface between IL-17 and IL-17RA, suggesting that they may have the potential to modulate this interaction. As anticipated, the positive control peptide was predicted to bind IL-17 with a high affinity, $K_{D,in\ silico}$ ~ 49 nM. Similarly, the negative control cyclic peptide returned a $K_{D,in\ silico}$ of 80 μM.

Table 5.3. Sequences of IL-17 binding cyclic peptides analyzed in a molecular docking study with the IL-17 interface. Mutagenized positions are denoted with an X. Numbers in parentheses represent the number of times a sequence appeared during Illumina Next Generation Sequencing.

	M	X_1	X_2	X_3	X_4	X_5	X_6	X_7	K
IL17.1 (4191)		R	W	L	R	G	R	R	
IL17.2 (1197)		F	G	L	L	H	R	G	
IL17.3 (1082)		D	L	W	R	W	R	R	
IL17.4 (747)		S	F	R	F	W	R	L	
IL17.5 (498)		S	F	F	D	I	W	R	
IL17.6 (428)		I	G	Q	W	W	R	R	
IL17.7 (318)		R	F	W	P	F	R	V	
IL17.8 (307)		F	L	F	R	L	G	R	
IL17.9 (285)		E	F	C	L	F	R	C	
IL17.10 (275)		Y	R	F	H	R	H	G	
IL17.11 (251)		F	R	I	F	L	P	R	
IL17.12 (229)		C	L	L	W	C	R	R	
IL17.13 (228)		F	K	G	Y	Y	R	W	
IL17.14 (219)		N	R	L	K	F	W	F	
IL17.15 (145)		K	H	R	P	F	R	Y	
IL17.16 (141)		F	G	W	R	M	R	L	
IL17.17 (119)		P	F	G	H	W	R	R	

It is important to note that the *in silico* method used to evaluate the isolated cyclic peptide sequences is more rigorous than traditional docking simulations[104]. Molecular docking alone, in fact, is not sufficient to predict the binding energies with sufficient accuracy to reproduce experimental values, since the scoring functions used to rank the docked complexes provide a rather approximative evaluation of the binding energy[105,106]. In this work, we used scoring functions to select top hits that were subsequently refined via MD simulations. Upon stabilization of the simulation trajectories, the IL-17:peptide binding energies were calculated using the MM/GBSA method, which has been shown to provide results that compare well with experimental measurements[102,105,107–109]. We selected six peptide sequences predicted to bind IL-17 with *in silico* affinities reasonably comparable to the positive control peptide for further *in vitro* evaluation.

Table 5.4. Estimation of binding energy (ΔG_B) and calculated affinity (K_D) of the IL-17:peptide complexes refined using molecular dynamics simulations. IHVTIPADLWDWINK and cyclo-[*DSG*-MGSGGGGSG-K] were used as positive and negative controls, respectively.

Sequence	Site 1 ΔG_B (kcal/mol)	Site 1 K_D (nM)	Site 2 ΔG_B (kcal/mol)	Site 2 K_D (nM)
IHVTIPADLWDWINK	-9.97	49		
Cyclo-[*DSG*-MGSGGGGSG-*K*]	-5.57	81000		
Cyclo-[*DSG*-MRWLRGRR-*K*]	-10.14	36	-9.59	91
Cyclo-[*DSG*-MIGQWWRR-*K*]	-9.26	160	-9.93	51
Cyclo-[*DSG*-MNRLKFWF-*K*]	-9.97	48	-9.34	140
Cyclo-[*DSG*-MSFFDIWR-*K*]	-8.93	280	-9.88	56
Cyclo-[*DSG*-MYRFHRHG-*K*]	-9.67	79	-9.83	61
Cyclo-[*DSG*-MFGLLHRG-*K*]	-8.20	950	-9.67	80
Cyclo-[*DSG*-MDLWRWRR-*K*]	-9.68	78	-8.57	510
Cyclo-[*DSG*-MEFCLFRC-*K*]	-9.29	150		
Cyclo-[*DSG*-MFGWRMRL-*K*]	-8.67	430	-9.02	240
Cyclo-[*DSG*-MFKGYYRW-*K*]	-9.02	240	-9.00	250
Cyclo-[*DSG*-MFRIFLPR-*K*]	-9.63	86	-8.18	1000
Cyclo-[*DSG*-MRFWPFRV-*K*]	-8.96	270	-8.74	390

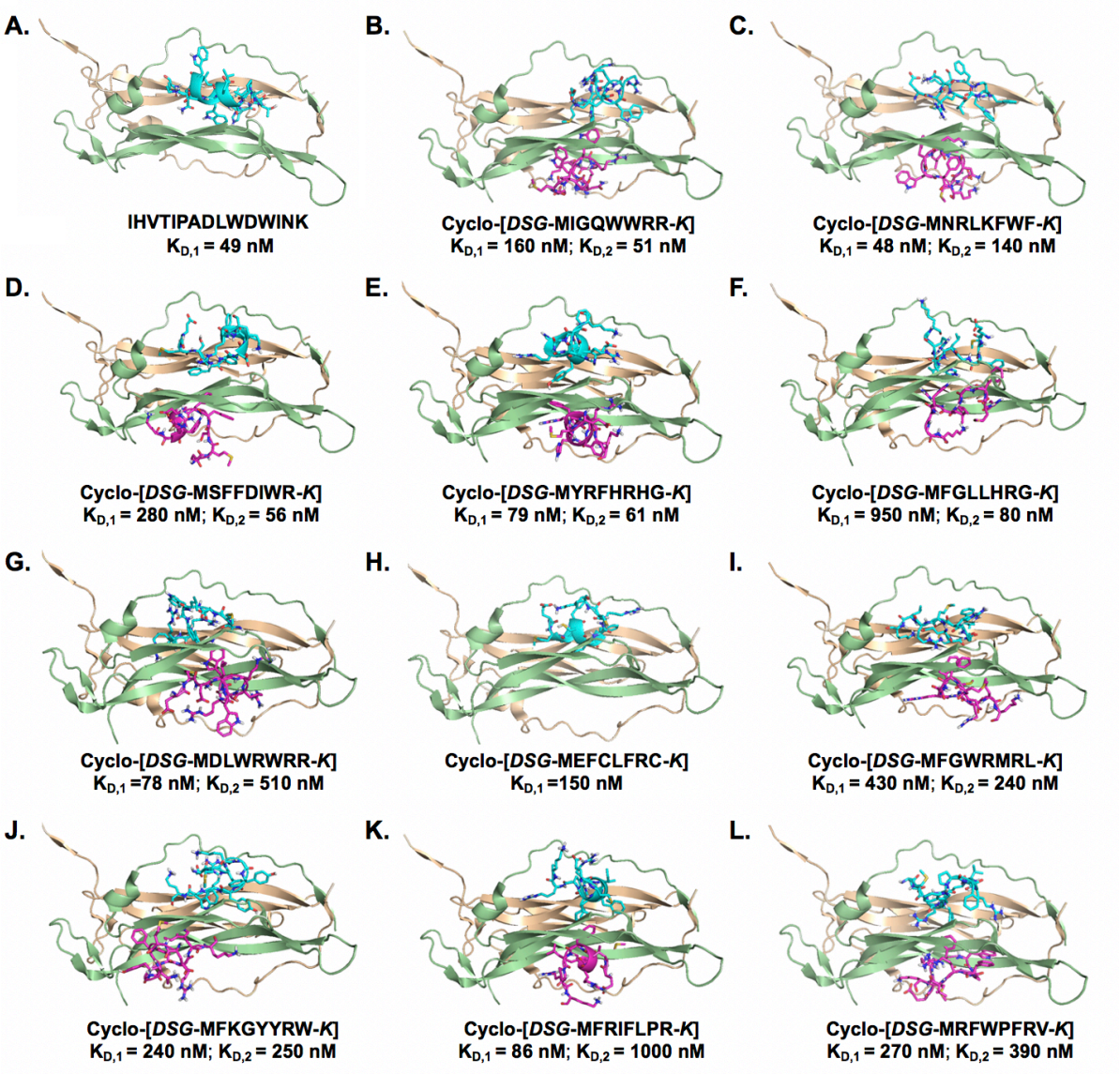


Figure 5.13. *In silico* IL-17:cyclic peptide complexes obtained by molecular docking and refined by molecular dynamic simulations. The monomers of IL-17 are denoted in green and brown while the respective cyclic peptide is modeled in cyan or magenta. The positive control, linear peptide, IHVTIPADLWDWINK, is modeled in **(A)**.

5.2.7 Binding affinity and selectivity of selected cyclic peptides for IL-17

The binding of IL-17 to yeast cells displaying peptides cyclo-[*DSG*-RMRWLRGRR-*K*], cyclo-[*DSG*-RMIGQWWRR-*K*], cyclo-[*DSG*-RMNRLKFWF-*K*], cyclo-[*DSG*-RMSFFDIWR-*K*], cyclo-[*DSG*-RMYRFHRHG-*K*], and cyclo-[*DSG*-RMFGLLHRG-*K*] was evaluated at a fixed IL-17 concentration (500 nM). No significant difference in binding was observed between the

various mutants (Figure 5.14). Hence, cyclo-[*DSG*-RMRWLRGRR-*K*] was chosen for subsequent evaluation, as this sequence had the highest predicted *in silico* affinity and was the most abundant mutant in the population after FACS, as identified by NGS. Note that the *in silico* results do not provide an absolute binding affinity estimation, but rather suggest, through binding affinity rank ordering, that cyclo-[*DSG*-MRWLRGRR-*K*] is likely the highest affinity mutant from those simulated. Also, the abundance of a particular clone in a sorted population does not necessarily correlate with the clone's affinity in that population[110,111].

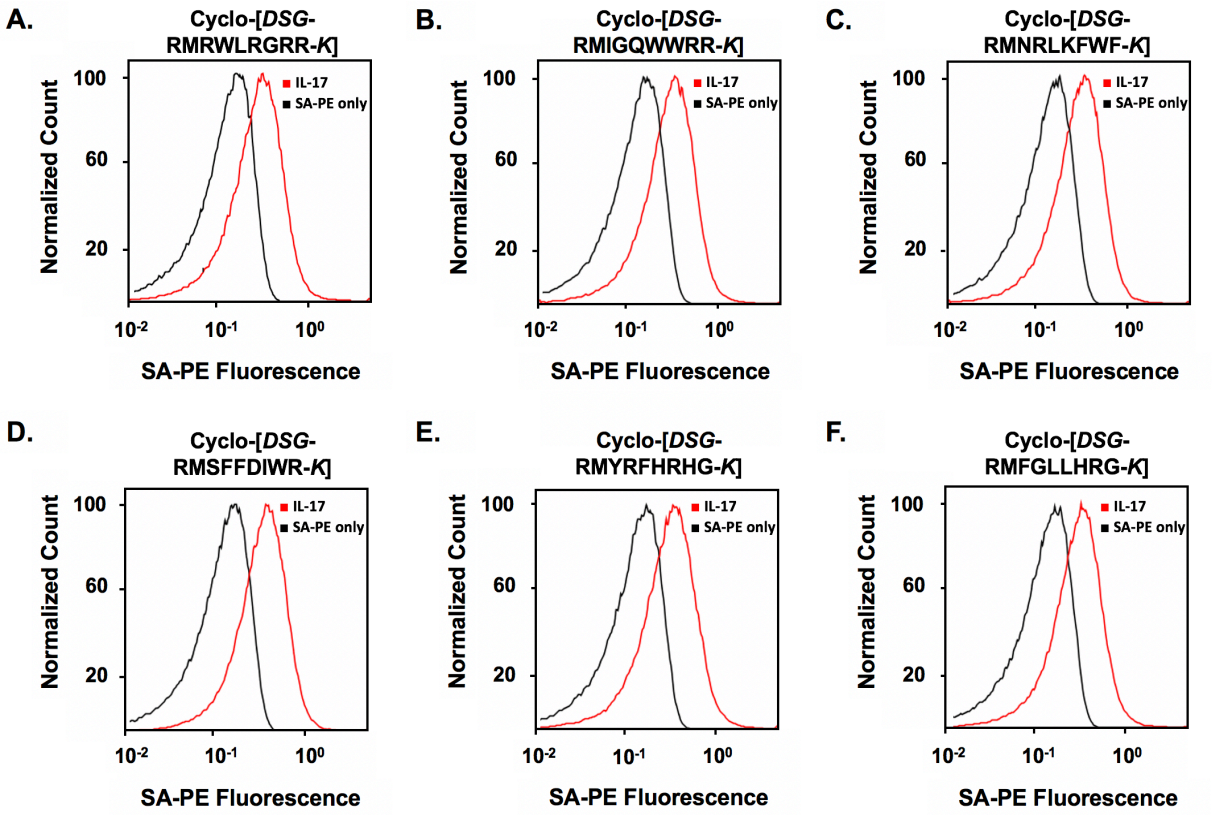


Figure 5.14. Binding of IL-17 to yeast-displayed cyclic peptides. The interaction of IL-17 (500 nM) to the yeast-displayed cyclic peptides was detected using SA-PE (red). The binding of SA-PE alone was also considered (black). Cyclic peptides considered include: **(A)** cyclo-[*DSG*-RMRWLRGRR-*K*], **(B)** cyclo-[*DSG*-RMIGQWWRR-*K*], **(C)** cyclo-[*DSG*-RMNRLKFWF-*K*], **(D)** cyclo-[*DSG*-RMSFFDIWR-*K*], **(E)** cyclo-[*DSG*-RMYRFHRHG-*K*], and **(F)** cyclo-[*DSG*-RMFGLLHRG-*K*].

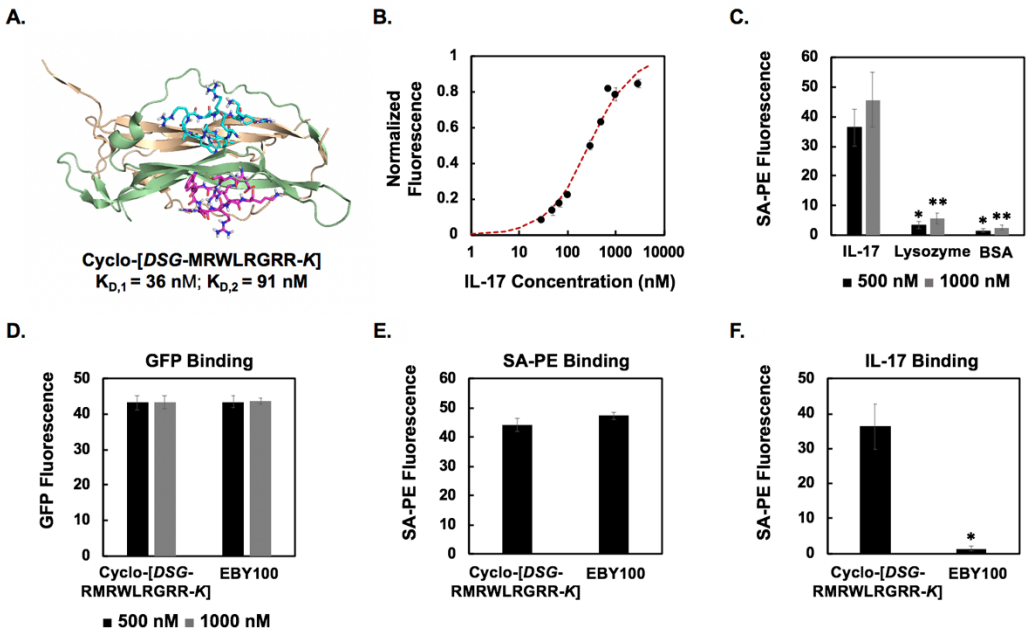


Figure 5.15. Affinity and specificity analysis of cyclo-[*DSG*-RMRWLRGRR-*K*] for IL-17. **(A)** *In silico* IL-17:cyclo-[*DSG*-MRWLRGRR-*K*] complex obtained by molecular docking and refined by molecular dynamic simulation. The monomers of IL-17 are denoted in green and brown while cyclo-[*DSG*-MRWLRGRR-*K*] is modeled in cyan or magenta. **(B)** Estimation of apparent K_D describing the affinity between IL-17 and yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*]. Normalized mean fluorescence values describing the binding of biotinylated IL-17 to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] as a function of the concentration of IL-17 incubated are reported. Error bars represent the standard error of the mean for three independent experiments. A global, non-linear regression was used to estimate the apparent K_D as 279 nM (243 – 320 nM, 68% confidence interval). The reported mean fluorescence values associated with IL-17 binding were first background subtracted and then normalized relative to the maximum fluorescence values estimated by the non-linear regression for each repeat. The mean background fluorescence is associated with the signals produced after incubation of the cyclic peptide displaying cells with SA-PE alone. **(C)** Background subtracted mean fluorescence of biotinylated IL-17, biotinylated lysozyme, and biotinylated BSA binding to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] at 500 and 1000 nM. Binding of these biotinylated proteins was detected by flow cytometry after SA-PE labeling. * denotes statistical significance in comparison to IL-17 binding at 500 nM for $p<0.05$ while ** denotes statistical significance in comparison to IL-17 binding at 1000 nM for $p<0.05$. The mean background fluorescence is derived from the signals produced after the incubation of the cyclic peptide displaying cells with SA-PE alone. **(D)** Mean fluorescence of non-biotinylated GFP (500 and 1000 nM) binding to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] and DSG treated, non-displaying EBY100 yeast as detected using flow cytometry. **(E)** Mean fluorescence of SA-PE (500 nM) binding to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] and DSG treated, non-displaying EBY100 yeast as detected using flow cytometry. **(F)** Background subtracted mean fluorescence of IL-17 binding to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] and DSG-treated, non-displaying EBY100 yeast. Binding of biotinylated IL-17 was detected by flow cytometry after SA-PE labeling. * represents statistical significance ($p< 0.05$) in comparison to the binding of IL-17 to yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*]. Here, the mean background fluorescence stems from the signals produced after incubation of each respective cell population with SA-PE alone. Error bars represent the standard error of the mean for three individual replicates throughout. Statistical significance was evaluated using a two-tailed, paired t-test throughout. Mean fluorescence values were scaled by a factor of 100 throughout.

The apparent affinity of cyclo-[*DSG*-RMRWLRGRR-*K*] for IL-17 was evaluated using the yeast surface titration method established earlier (Figure 5.15B). The apparent binding affinity of

cyclo-[*DSG*-RMRWLRGRR-*K*] was estimated as 279 nM (243 – 320 nM, 68% confidence interval). It is worth noting that the soluble IL-17 protein used in this assay is a homodimer, which results in a multi-site interaction with the yeast-displayed peptides. The true binding affinity of cyclo-[*DSG*-RMRWLRGRR-*K*] for IL-17, as estimated when both the cyclic peptide and the IL-17 are in solution, is likely lower.

To confirm that DSG-mediated cyclization is needed to elicit binding of this identified sequence, we simulated the binding of the linear version of cyclo-[*DSG*-RMRWLRGRR-*K*] to IL-17 *in silico* using the previously described method (Figure 5.16). The linear form of this peptide, MRWLRGRR, did not significantly bind to IL-17 during the simulation. This provides further confidence that DSG cyclization contributes to the binding specificity of cyclo-[*DSG*-RMRWLRGRR-*K*] for IL-17, rather than simply the particular amino acids in the sequence.

The binding selectivity of cyclo-[*DSG*-RMRWLRGRR-*K*] was evaluated by comparing the binding of IL-17 and multiple putative, non-specific proteins to yeast cells displaying this peptide. Our results indicate that IL-17 bound to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] at a significantly higher level ($p<0.05$) than lysozyme (pI ~ 11)[79] and BSA (pI ~ 4.7)[88] (Figure 5.15C). While the poor binding of lysozyme can be imputed to electrostatic repulsion from the cationic cyclic peptide, it should be noted that IL-17 readily bound despite its mild cationic character (sequence-based pI of 8.82), indicating true binding affinity. The binding selectivity of the peptide against BSA, while engineered through the inclusion of BSA in the buffers utilized for library screening, is nonetheless notable, given the potential of BSA binding solely due to Coulomb interaction. Finally, green fluorescent protein (GFP) (pI ~ 5)[89] and phycoerythrin (pI ~ 4.2)[112] conjugated streptavidin (pI ~ 5-6)[113], otherwise known as SA-PE, were found to bind similarly to cells expressing cyclo-[*DSG*-RMRWLRGRR-*K*] and non-displaying EBY100 yeast

(Figures 5.15D-E), suggesting that cyclo-[*DSG*-RMRWLRGRR-*K*] has no affinity for these proteins.

Lastly, we confirmed that IL-17 exhibits no interactions with non-specific proteins displayed on the surface of DSG-treated yeast. The large excess of DSG employed during the cyclization of linear displayed peptides is likely to promiscuously modify other yeast surface proteins. The possibility of IL-17 interacting with DSG-modified surface proteins cannot be excluded *a priori*. To evaluate this, we compared the binding of IL-17 to DSG-treated, null EBY100 cells *vs.* DSG-treated, peptide-displaying yeast cells (Figure 5.15F). The lack of binding to the null cells corroborates the claim that yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] specifically interacts with IL-17 and provides confidence that off-target DSG modifications did not influence the identification of peptide binders during the library screening process to discover IL-17 affinity ligands.

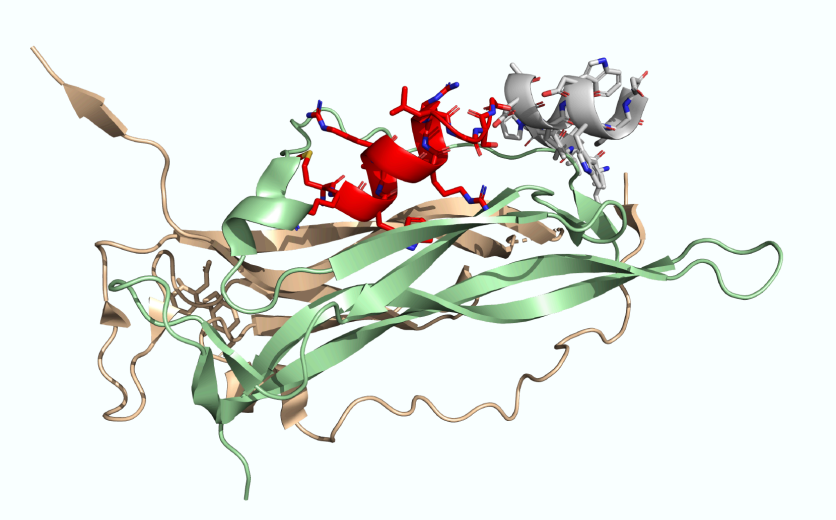


Figure 5.16. *In silico* IL-17:MRWLRGRRK complex obtained by molecular docking and refined by molecular dynamic simulations. MRWLRGRRK is simulated in its linear form. MRWLRGRRK does not appreciably bind to IL-17 and can be seen flying off of the IL-17 structure (transition from red to grey molecule) during the simulation. The monomers of IL-17 are denoted in green and brown while the respective linear peptide is modeled in red or grey.

5.2.8 Evaluation of cyclo-[DSG-MRWLRGRR-K] as a modulator of the IL-17:receptor interaction

While the selectivity of cyclo-[*DSG*-RMRWLRGRR-*K*] for IL-17 supports the use of this peptide as an affinity ligand, we also explored the use of cyclo-[*DSG*-MRWLRGRR-*K*] to modulate the interaction between IL-17 and its receptor. Peptides and other small molecules have been used to modulate protein-protein interactions, typically as inhibitors[8,114]. Specific to the IL-17:receptor interface, linear peptide IHVTIPADLWDWINK was previously developed as an inhibitor of this interaction for potential use as an anti-inflammatory agent[93].

Initially, we evaluated whether cyclo-[*DSG*-MRWLRGRR-*K*] modulates the interaction between IL-17 and IL-17RA using *in silico* simulations. We first used MD-refined, docking simulations to model the interaction of cyclo-[*DSG*-MRWLRGRR-*K*] with IL-17. Subsequently, the binding of IL-17RA to IL-17 was superimposed on the simulated IL-17:peptide complex (Figure 5.17A). The results show that cyclo-[*DSG*-MRWLRGRR-*K*] likely binds to IL-17 in a similar region as IL-17RA, thereby suggesting that the peptide could potentially inhibit the interaction between IL-17 and IL-17RA.

To evaluate the effect of cyclo-[*DSG*-MRWLRGRR-*K*] on IL-17:IL-17RA binding, we used a two-step approach. First, yeast cells displaying IL-17 were incubated with soluble IL-17RA (20 nM or 50 nM) in the presence of varying concentrations of soluble cyclo-[*DSG*-MRWLRGRR-*K*]. The binding of IL-17RA to yeast-displayed IL-17 was detected by flow cytometry following biotinylated protein A and SA-PE labeling; note that the IL-17RA used in this work is a Fc fusion. For concentrations less than 30 μM of cyclo-[*DSG*-MRWLRGRR-*K*], the binding of IL-17RA to yeast-displayed IL-17 was not significantly altered, indicating no inhibition of the IL-17:IL-17RA

interaction (Figures 5.17B-C). For high peptide concentrations, a slight increase in IL-17RA binding was observed. However, this increase was not statistically significant ($p>0.05$).

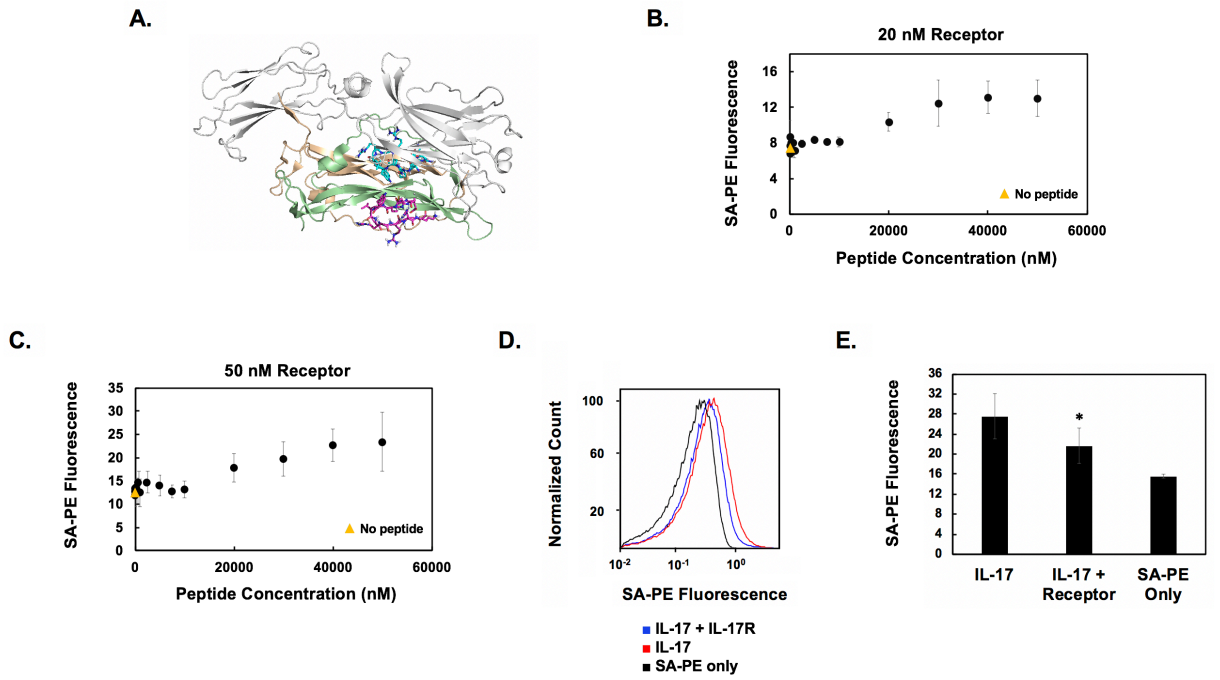


Figure 5.17. Evaluation of cyclo-[*DSG*-MRWLRGRR-*K*] as a potential modulator of IL-17's interaction with its receptor, IL-17RA. **(A)** *In silico* modeling of cyclo-[*DSG*-MRWLRGRR-*K*] (cyan or magenta) interacting with IL-17 (monomers denoted in green and brown). IL-17RA was not present during the simulation. Rather, the interaction of IL-17RA with IL-17 was superimposed after modeling and is visualized in grey. This visualization suggests cyclo-[*DSG*-MRWLRGRR-*K*] interacts with IL-17 in the same region as IL-17RA. To evaluate the *in silico* results, a two-step *in vitro* approach was explored. **(B-C)** First, yeast cells displaying IL-17 were incubated with soluble cyclo-[*DSG*-MRWLRGRR-*K*] and IL-17RA. The binding of IL-17RA to the IL-17 displaying cells was detected using flow cytometry after labeling with biotinylated protein A and SA-PE. The concentration of peptide was varied from 0 nM to 50 μM while the concentration of receptor was held at either 20 nM **(B)** or 50 nM **(C)**. Background subtracted mean fluorescence values describing IL-17RA binding to yeast-displayed IL-17 in the presence and absence of the peptide are reported. The mean background fluorescence is derived from the signals produced after the incubation of the cyclic peptide displaying cells with biotinylated protein A and SA-PE alone. The binding of IL-17RA when no peptide was present is denoted with a yellow triangle for comparison. No significant inhibition of the interaction between IL-17 and IL-17RA is observed likely due to the peptide not having sufficient affinity in its monovalent form to compete with IL-17RA for binding to IL-17. **(D-E)** In the second approach, yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] were incubated with biotinylated IL-17 (100 nM) in the presence (blue) or absence (red) of IL-17RA (500 nM) followed by flow cytometry analysis using SA-PE detection. Yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] were also incubated with only SA-PE (black). A representative histogram of three repeats is shown **(D)**. The mean fluorescence values describing the binding of biotinylated IL-17 to the yeast-displayed peptides in the presence and absence of IL-17RA are reported **(E)**. The binding of IL-17 to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] is statistically lower when IL-17RA is present (* denotes $p<0.001$ for a two tailed, two sample unequal variance t-test). Here, the mean fluorescence values are not background subtracted. In this approach, it is probable that the yeast-displayed peptides can partially inhibit the interaction between IL-17 and its receptor as the peptide decorated yeast surface provides sufficient avidity. Error bars represent the standard error of the mean for three individual replicates throughout. The mean fluorescence values were multiplied by a factor of 100 throughout.

When considering the results of this assay, it is important to note that the binding affinity of IL-17 for IL-17RA is ~ 2.8 nM[115], whereas yeast-displayed cyclo-[*DSG*-MRWLRGRR-*K*] binds IL-17 with an apparent K_D ~ 279 nM, as determined by yeast titration. In this first *in vitro* inhibition assay, the binding affinity of the soluble peptide for IL-17 is likely significantly lower than the apparent affinity predicted via yeast-surface titration (K_D ~ 279 nM), as the interaction between IL-17 and the soluble peptide is monovalent in nature. In contrast to the yeast-displayed peptides, the soluble peptide cannot take advantage of avidity to increase its apparent binding affinity for IL-17. Moreover, the strength of interaction between the yeast-displayed IL-17 and the dimeric IL-17RA Fc fusion is increased due to avidity. Therefore, it is not surprising that the monomeric cyclic peptide does not inhibit the IL-17:IL-17RA interaction in this first approach, as the binding affinity of soluble cyclo-[*DSG*-MRWLRGRR-*K*] is not sufficient to compete with IL-17RA for binding to IL-17.

For the second approach, yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] were incubated with soluble, biotinylated IL-17 (100 nM) in either the presence or absence of IL-17RA (500 nM). The binding of biotinylated IL-17 to the cells was measured by flow cytometry using SA-PE for detection (Figure 5.17D). The fluorescence signal corresponding to the binding of IL-17 was lower ($p<0.001$) in presence of IL-17RA (Figure 5.17E), suggesting that cyclo-[*DSG*-RMRWLRGRR-*K*] and IL-17RA compete for binding to IL-17. It is also important to note that yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] do not interact with IL-17RA (Figure 5.18). The outcome of these two *in vitro* approaches vary due to differences in the avidity of the IL-17:peptide interaction. In the second approach, the interaction strength between the peptide-decorated yeast surface and the IL-17 homodimer is increased, due to avidity, as the IL-17 homodimer can interact with multiple copies of the yeast surface expressed peptides (Figures

5.17D-E). This is in contrast to the first approach where the interaction between the soluble peptide and IL-17 is monovalent in nature (Figures 5.17B-C). Consistent with this explanation, Liu et al. observed increased inhibition of the IL-17:IL-17RA interaction when using a dimeric form of their linear peptide specific to IL-17[93]. Taken together with *in silico* simulations, our results show that cyclo-[*DSG*-RMRWLRGRR-*K*] may inhibit the interaction between IL-17 and IL-17RA by competing for binding to IL-17 when the peptide is expressed in a form with sufficient avidity.

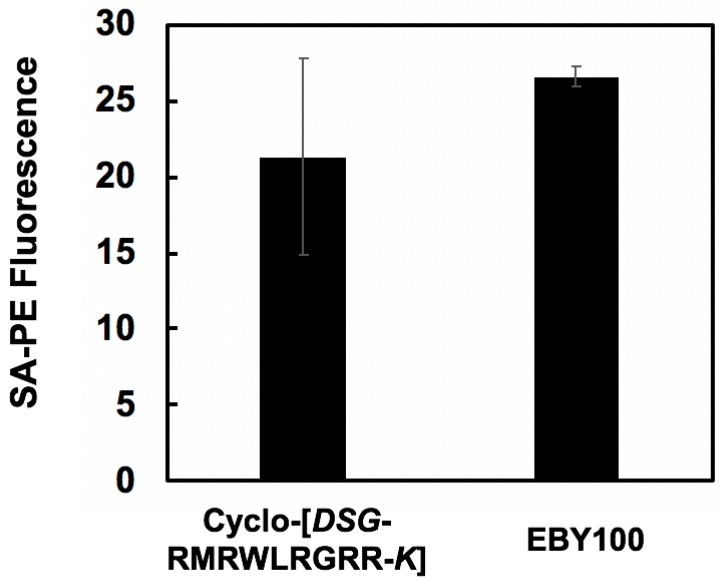


Figure 5.18. Evaluation of Cyclo-[*DSG*-RMRWLRGRR-*K*]'s interaction with IL-17RA. Yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] and DSG-treated EBY100 yeast were incubated with 500 nM of IL-17RA followed by flow cytometry detection after labeling with biotinylated protein A and SA-PE. The background subtracted mean fluorescence of IL-17RA binding to each cell type is reported. The mean background fluorescence stems from the signals produced after incubating each cell type with biotinylated protein A and SA-PE alone. Error bars represent the standard error of the mean for three individual repeats. Cyclo-[*DSG*-RMRWLRGRR-*K*] does not interact with IL-17RA.

5.3 Conclusions

Collectively, our results demonstrate the successful development of a method for the efficient identification and characterization of cyclic peptide ligands with selective protein-binding activity. Of note is the integration of FACS, which accelerates the library screening process. Due to the

limitations of monocyclic peptides, binding selectivity must be engineered into the peptides during combinatorial library screening by performing negative selections against specific proteins. The use of yeast-displayed cyclic peptides allows on-the-fly selectivity and affinity evaluation of potential peptide ligands. We envision this method will enable the identification of cyclic peptides to be used as affinity ligands for protein purification[44] and biosensor[116] applications that require moderate affinity biorecognition moieties as well as inhibitors of moderate affinity protein-protein interactions.

5.4 Materials and Methods

5.4.1 Plasmids and yeast culture

Cells containing a pCTCON-based plasmid were cultured using TRP-deficient SDCAA media, whereas SGCAA medium was used to induce cell surface protein expression for these cells, as described previously[72]. Leu-deficient SDSCAA (-Leu) and SGSCAA (-Leu) media was similarly used for cells containing a pCT302-based plasmid[117]. Leu-deficient media is similar in composition to SDCAA and SGCAA media except it does not contain casamino acids, but rather synthetic drop out mix (1.62 g/L) lacking leucine (US Biological Life Sciences). Yeast cells were cultured in the appropriate growth media, SDCAA or SDSCAA, at 30°C with shaking at 250 RPM. To induce protein expression, yeast cells were transferred to SGCAA or SGSCAA medium at an OD_{600} of 1 and cultured overnight at 20 °C with shaking at 250 rpm. EBY100 yeast without plasmid was grown in YPD medium (10.0 g/L yeast extract, 20.0 g/L peptone, and 20.0 g/L dextrose) at 30°C with shaking at 250 rpm. Plasmid DNA was transformed into chemically competent *Saccharomyces cerevisiae* strain EBY100 using the Frozen-EZ Yeast Transformation Kit II (Zymo Research).

5.4.2 Construction of plasmids for linear peptide yeast surface display

pCTCON-Nterm was constructed by amplifying gene block 1 with Pf1 and Pr1 and introduced into pCTCON via the EcoRI and XhoI sites. This particular gene block displays a random linear peptide, between the NheI and BamHI restriction sites. To display specific linear peptide sequences on the yeast surface, pCTCON-Nterm was cut with NheI and BamHI, and DNA encoding a linear peptide sequence was introduced via homologous recombination. The linear peptide sequence DNA was purchased with the necessary homology arms as single stranded oligonucleotides and amplified using Pf2 and Pr2. These homology arms eliminate the NheI and BamHI sites present in pCTCON-Nterm (Figure 5.19A). The linear sequence of three previously identified DSG-cyclized peptides[44], with affinity for IgG, were displayed on the yeast surface using oligonucleotide DNA blocks that encode the sequences MWFPHYK, MHGFRGK, and MWFRHYK, respectively.

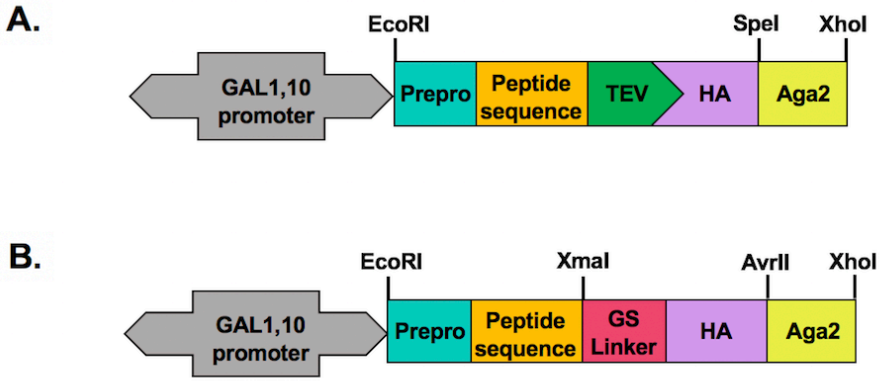


Figure 5.19. Plasmid design for the expression of a linear peptide sequence as a N-terminal Aga2p fusion. Two different plasmid designs were used when expressing the **(A)** IgG binding peptides and **(B)** variants from the heptapeptide yeast display library.

All double stranded gene block fragments were purchased from Integrated DNA Technologies (IDT). All primer oligonucleotides and oligonucleotide DNA fragments were

purchased from IDT or Eton Biosciences. Gene block fragments, oligonucleotide primers, and oligonucleotide DNA fragments can be found in Tables 5.5 – 5.7, respectively. PCR reactions were performed at a 50 μL scale using Phusion Polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. The amplified DNA was purified using the 9K Series Gel and PCR extraction kit from BioBasic. Restriction enzymes were bought from New England Biolabs. Plasmids and PCR products were digested in a 50 μL reaction volume with a 5X excess of each restriction enzyme for 4 hours at 37°C. Digested plasmid backbones were subsequently incubated with Antarctic Phosphatase purchased from New England Biolabs for 1 hour at 37°C. After, digested plasmids and PCR products were purified using the BioBasic 9K Series Gel and PCR extraction kit. Ligations combining digested plasmid backbone and inserts were carried out overnight at 16°C using T4 DNA Ligase (Promega) prior to transformation into electrocompetent Novablue *E. coli* cells. Transformations were carried out at 1.6 KV, 25 μF, and 200 Ω. The GeneJET plasmid miniprep kit (Thermo Fisher Scientific) was used to harvest DNA plasmids from overnight *E.coli* cultures.

When cloning via homologous recombination, 200 ng of cut plasmid and 250 ng of insert were transformed into chemically competent EBY100 cells prepared with the Frozen-EZ Yeast Transformation Kit II. Here, the digested plasmid was not treated with Antarctic Phosphatase. The plasmid and insert were concentrated using ethanol precipitation such that the volume of DNA used in the transformation step was less than 5 μL in total. Briefly, a 1/10 volume of potassium acetate and 10 μL of linear acrylamide was added to the DNA to be concentrated followed by 2 volumes of ice-cold ethanol. The mixture was stored at -20 °C overnight. The next day, the DNA was centrifuged at 15,000 g for 10 minutes followed by removal of the supernatant. The DNA pellet was washed once with 70% ethanol followed by a wash with 100% ethanol and dried prior

to resuspension in water. Recombined plasmid was extracted from the transformed yeast cells using the Zymoprep™ Yeast Plasmid Miniprep II kit following the manufacturer's recommendations. The recovered plasmids were transformed into electrocompetent Novablue *E. coli* cells, and the DNA was recovered from overnight *E. coli* cultures using the GeneJET plasmid miniprep kit.

Table 5.5. List of gene block fragments described in Chapter 5.

Gene Block 1	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAGCTAGCGGACAATTACTA TTTACAATTACAATGAGTCATAGAGCAGGTCTGGAGAAGGGCAGC GGCGGCAGCGGATCCGAGAACCTGTACTTCCAAGGGTATCCGTAC GACGTTCCAGACTACGCTACTAGTCAGGAACTGACAACTATATGC GAGCAAATCCCCTCACCAACTTTAGAATCGACGCCGTACTCTTTGT CAACGACTACTATTTTGGCCAACGGGAAGGCAATGCAAGGAGTTT TTGAATATTACAAATCAGTAACGTTTGTCAGTAATTGCGGTTCTCA CCCCTCAACAACTAGCAAAGGCAGCCCCATAAACACACAGTATGT TTTTTAATGACTCGAG
Gene Block 2	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGNNKNNKNNKNNKNN KNNKNNKAAGCCCGGGGGTGGTGGTGGTTCTGGAGGAGGCTCTGG TG
Gene Block 3	ATGACGCCTGGGAAAACGAGCTTGGTCTCACTACTATTGCTTCTGA GCCTTGAGGCGATAGTTAAGGCGGGTATCACAATACCTAGAAACC CAGGCTGCCCGAACAGTGAAGACAAAAACTTCCCGAGAACTGTGA TGGTCAACCTAAACATACACAATAGGAATACAAACACTAATCCAA AGAGGAGCAGTGATTACTACAATAGATCAACCTCTCCTTGGAACT TACATCGTAATGAGGACCCTGAGAGGTATCCTTCAGTCATCTGGG AGGCTAAGTGTAGACATTTGGGCTGCATCAACGCTGACGGTAACG TCGATTATCACATGAACTCCGTGCCTATTCAACAGGAAATTCTGGT ATTGAGGCGTGAGCCTCCTCATTGTCCCAATAGTTTCAGGCTTGAG AAGATCCTTGTTTCAGTTGGCTGCACTTGCGTCACTCCAATCGTAC ACCACGTTGCA

Table 5.5 (Continued).

Gene Block 4	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGCGTTGGTTAAGGGGA AGAAGGAAACCCGGGGGTGGTGGTGGTTCT
Gene Block 5	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGATAGGCCAATGGTGG AGAAGAAAGCCCGGGGGTGGTGGTGGTTCT
Gene Block 6	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGAACCGTCTTAAATTT TGGTTTAAACCCGGGGGTGGTGGTGGTTCT
Gene Block 7	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGTCCTTCTTCGACATAT GGCGTAAACCCGGGGGTGGTGGTGGTTCT
Gene Block 8	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGTACCGTTTCCATAGG CATGGAAAGCCCGGGGGTGGTGGTGGTTCT
Gene block 9	GAATTCATGAAGGTTTTGATTGTCTTGTTGGCTATCTTCGCTGCTTT GCCATTGGCCTTAGCTCAACCAGTAATTTCTACTACCGTCGGTTCC GCTGCAGAAGGCTCTTTGGACAAGAGAATGTTTGGCTTGTTGCACC GTGGCAAGCCCGGGGGTGGTGGTGGTTCT

Table 5.6. List of oligonucleotide primers described in Chapter 5.

Primer	Sequence
Pf1	CTGGACGAATTCATGAAGGTTTTGATTGTCTT
Pr1	CTGGACCTCGAGTCATTAAAAAACATACTGT
Pf2	GCAGAAGGCTCTTTGGACAA
Pr2	ATACCCTTGGAAGTACAGGTT
Pf3	GTATTACTTCTTATTCAAATGTAATAAAAGATCGAATTCATGAAGGTTTTG ATTGTCTTGTTG
Pr3	CACCAGAGCCTCCTCCAGA
Pf4	CATAGCGCTAGCATGACGCCTGGGAAAACGAG
Pr4	CACTGCGGATCCTGCAACGTGGTGTACGATTGG
Pf5	GAATTCATGAAGGTTTTGATTGTCT
Pr5	AGCGTAGTCTGGAACGTCGT
Pf6	GCACGAGAATTCATGAAGGTTTTGATTGTC
Pr6	GCACGAAGAACCACCACCACCCC

Table 5.7. Oligonucleotide DNA fragments described in Chapter 5.

Oligo	Sequence
Oligo 1	GCAGAAGGCTCTTTGGACAAGAGAATGTGGTTTCCGCATTACAAGGAGAACCTGTACTTCCAAGGGTAT
Oligo 2	GCAGAAGGCTCTTTGGACAAGAGAATGCACGGTTTTAGAGGCAAGGAGAACCTGTACTTCCAAGGGTAT
Oligo 3	GCAGAAGGCTCTTTGGACAAGAGAATGTGGTTTAGGCACTACAAGGAGAACCTGTACTTCCAAGGGTAT

5.4.3 Cyclization of linear peptides expressed as yeast surface fusions

$1x10^8$ yeast cells expressing a linear peptide sequence were washed 3X with crosslinking buffer (0.2 M NaCl, 1 mM EDTA). The cells were then resuspended in 2 mL of crosslinking buffer and 10 μL of DSG crosslinker (Thermo Fisher Scientific) dissolved in N,N'-dimethylformamide (DMF) at 1.5 mg/mL. The crosslinking reaction was performed for 2 hours at 4°C with rotation. After, the cells were washed 3X with crosslinking buffer, and the crosslinking reaction was repeated. Finally, the cells were washed three times with 50 mM Tris, 300 mM NaCl pH 7.4 and stored in this buffer overnight prior to subsequent use.

5.4.4 Binding analysis to confirm cyclization of yeast-displayed linear peptides

The interaction of IgG with DSG-treated or non-treated yeast cells expressing the linear form of select IgG binding cyclic peptides was characterized via flow cytometry. $5x10^6$ cells were labeled with 450 nM biotinylated IgG (Immunoreagents) for 15 minutes at room temperature. After, the cells were washed and labeled with a 1:250 dilution of SA-PE (Thermo Fisher Scientific) for 12 minutes on ice. Cells were analyzed using a Miltenyi Biotec MACsQuant VYB cytometer.

5.4.5 Generation of a cyclic heptapeptide yeast display library

A yeast display combinatorial library of linear peptides was generated by first amplifying gene block 2 with primers Pf3 and Pr3, that contain consensus sequences with pCT-NT-F2A-Sso7dhFc[118], in eight identical 50 μL reactions. Gene block 2 encodes a linear peptide sequence that contains seven randomized amino acid positions flanked between a N-terminal methionine and a C-terminal lysine. The peptide sequence is expressed as a N-terminal fusion to Aga2p (Figure 5.19B). A HA tag follows the peptide sequence to confirm expression of the peptide on the yeast surface via immunofluorescence. The amplified product from gene block 2 was purified using

phenol: chloroform extraction followed by DNA concentration using ethanol precipitation as previously described. Concurrently, 40 μg of a previously engineered plasmid, pCT-NT-F2A-Sso7dhFc[118], was linearized using EcoRI and XmaI. The digested backbone was purified and concentrated using phenol: chloroform extraction and ethanol precipitation. The yeast library was generated using a previously described lithium acetate transformation protocol where the amplified product from gene block 2 and the digested backbone vector were transformed into EBY100 yeast via electroporation[86]. Three electroporations were performed using a Bio-Rad Gene Pulser system where 12 μg of PCR product and 4 μg of digested vector were combined with 400 μL of electrocompetent EBY100. The electroporation settings used were 2.5 KV, 25 μF, and 250 Ω. A vector only electroporation was performed as a control. The library diversity was determined by plating serial dilutions of the combined transformation reactions onto SDCAA plates and estimated as $5x10^7$.

5.4.6 Screening a cyclic heptapeptide yeast display library against lysozyme

The cyclic heptapeptide yeast display library was screened against lysozyme in one round of magnetic selection and two rounds of fluorescence activated cell sorting (FACS). To perform the magnetic selection, 25 μL of Biotin Binder Dynabeads (Thermo Fisher Scientific) were pre-coated with 10.8 μM biotinylated lysozyme overnight at 4°C in 100 μL of 0.1% BSA PBS, pH 7.4 (0.1% PBSA). The biotin binder beads were washed 3X with 0.1% PBSA prior to functionalization. Lysozyme (Thermo Fisher Scientific) was previously biotinylated by incubating a 3:1 molar excess of Ez-Link™Sulfo-NHS-LC-Biotin (Thermo Fisher Scientific) for two hours at 4°C (resulted in ~2 biotins per molecule). Excess reagent was removed by performing dialysis against 50 mM Tris-HCl, 300 mM NaCl pH 7.4. Protein concentrations were estimated using a BCA assay.

For selection, the library was induced and cyclized as described in the main methods section. The following day, a negative selection was performed where $5x10^8$ cyclized library cells were incubated with 25 µL of non-functionalized Biotin Binder Dynabeads at room temp for 1 hour. Unbound cells were recovered, and an additional negative selection was carried out using the recovered cells and a new set of non-functionalized beads. Any unbound cells were subsequently removed and used in a positive selection against the pre-coated lysozyme beads for 1 hour at room temperature. Cells that did not bind the lysozyme-functionalized magnetic beads were removed using a magnet. After, the lysozyme-functionalized magnetic beads and any bound library cells were washed 3X with 0.1% PBSA, resuspended in SDCAA media, and incubated for 48 hours at 30°C to allow for expansion of the positively bound cell population. All selections took place in 4 mL of 0.1% PBSA. All beads were washed 3X with 0.1% PBSA prior to incubation with the library cells.

The pool of cells obtained from the magnetic selection were cyclized, and FACS was performed on this population using a MoFlo Cell Sorter (Beckman Coulter) as previously described[72]. For FACS, $1x10^7$ cells were labeled with biotinylated lysozyme on ice for 1 hour. After, the cells were labeled with SA-PE (1:250 dilution) on ice in the dark for 12 minutes. All labelings took place in 100 µL of 0.1% PBSA. Washes with 0.1% PBSA were performed after each incubation. The cells were not double labeled for lysozyme binding and HA expression as the binding of the anti-HA antibody appeared to hinder the peptide's ability to interact with lysozyme. Two rounds of FACS were performed where the library cells were labeled with 1 µM of lysozyme for each round. The population obtained after sorting was expanded in SDCAA media. After the second round of FACS and expansion, the isolated cells were plated on SDCAA plates (20 g/L dextrose, 5 g/L casamino acids, 6.7 g/L yeast nitrogen base, 182 g/L sorbitol, 5.4

g/L Na_2HPO_4, 8.6 g/L $NaH_2PO_4 \cdot H_2O$, 12 g/L Agar), and 13 clones were selected for sequencing. Plasmids were extracted from yeast cultures of the individual clones using a Zymoprep™ Yeast Plasmid Miniprep II kit. For each clone, the extracted plasmid DNA was transformed into electrocompetent Novablue *E.coli* cells. An overnight *E.coli* culture was grown for each clone followed by DNA extraction and sequencing.

5.4.7 Binding affinity estimation using yeast surface titrations

The apparent binding affinity (K_D) between the isolated cyclic peptides and their target protein (lysozyme or IL-17) was estimated using a yeast surface titration method. Briefly, yeast-displayed linear peptide sequences were cyclized using the technique described in the main methods section. Subsequently, the yeast cells displaying a cyclic peptide of interest were labeled with varying concentrations of the target antigen, either biotinylated lysozyme (20 nM – 8 μM) or biotinylated IL-17 (30 nM – 3 μM; R&D Systems), for 1 hour at room temperature. This was followed by secondary labeling with SA-PE (1:250 dilution) for 12 minutes on ice. The cells were then analyzed by flow cytometry as previously described[72]. All labeling took place in 50 μL of 0.1% PBSA. Washes with 0.1% PBSA were performed after each incubation. A SA-PE only labeling was also performed using the cyclic peptide displaying yeast cells to quantify the mean background fluorescence. The K_D describing the binding affinity between the yeast-displayed peptides and a target protein was estimated using the following relationship:

$$F = \frac{F_{Max}[L]_O}{K_D + [L]_O} \quad \text{(Eq. 1)}$$

Where F is the mean background subtracted fluorescence, $[L]_0$ is the concentration of antigen used to label the cells, and F_{Max} is the background subtracted fluorescence intensity when surface saturation is attained. The fluorescence binding data was fit to Eq.1 using a global, non-linear least

squares regression across three repeats as previously described[72]. A single K_D value and a unique F_{Max} for each repeat were used as the fitted parameters. At high antigen concentrations, the fluorescence signal decreased due to the Hook effect[80]. We ignored these points when applying the non-linear regression. Error bars represent the standard error of the mean for three independent experiments. The titration curves reported in Figure 5.6 and Figure 5.15 are generated using normalized fluorescence values that describe the background subtracted mean fluorescence of protein binding (lysozyme or IL-17) to the cyclic peptide displaying yeast cells relative to the maximum fluorescence for that respective repeat as estimated from the non-linear regression fit.

5.4.8 Binding affinity estimation using isothermal calorimetry

The affinity of cyclo-[DSG-MWWRLVYR-K] for lysozyme was also estimated using a nano-ITC calorimeter (TA Instruments). A soluble form of this peptide was synthesized by the UNC High-Throughput Peptide Synthesis and Array Facility[44]. The chemically synthesized, soluble peptide lacks the N-terminal arginine residue that is present in the yeast-displayed peptides. 500 μM of the peptide dissolved in water was injected into 50 μM of lysozyme dissolved in water. 20 injections of 2.5 μL were carried out 250 seconds apart at 25°C. The reference cell contained water. We performed this titration in quadruplicate. A blank titration was also performed where water was injected into 50 μM of lysozyme using the same conditions. The blank titration was only performed once. We used software provided by TA instruments to obtain the integrated heats of binding from the raw data after subtracting the heat of dilution associated with the blank titration. The experimental data was fit using the Independent model to estimate and calculate associated parameters, like K_D, N, ΔH, and ΔS, for each repeat. Ultimately, we averaged these values across the repeats as well as the associated standard deviations.

5.4.9 Selectivity characterization of lysozyme-binding cyclic peptides

The selectivity of the lysozyme-binding cyclic peptides was evaluated using two different approaches. First, the binding of biotinylated lysozyme to yeast cells displaying the cyclic form of each peptide was compared to the binding of biotinylated lysozyme to yeast cells displaying the linear counterpart. Briefly, $2x10^6$ cells displaying either the linear or cyclic form of each peptide were incubated with 100, 500, 1000, or 2000 nM of biotinylated lysozyme for 1 hour at room temperature followed by secondary labeling with SA-PE (1:250 dilution) for 12 minutes on ice and immunofluorescence detection using flow cytometry. Using a similar approach, we also compared the binding of biotinylated lysozyme and biotinylated BSA (100, 500, 1000, and 2000 nM) to yeast cells displaying cyclic peptides of interest. BSA was biotinylated in a similar manner as previously described for lysozyme (~2 biotins per molecule). The peptide expression level of each cell population was analyzed by labeling with a 1:100 dilution of an anti-HA antibody (Thermo Fisher Scientific) for 1 hour at room temperature followed by labeling with a 1:250 dilution of a donkey anti-rabbit 488 antibody (DAR488; Immunoreagents) for 12 minutes on ice and immunofluorescence detection using flow cytometry. For each peptide sequence, expression and secondary only (SA-PE and DAR488) labelings were performed for both the linear and cyclic populations. All labelings took place in 50 μL using 0.1% PBSA. Washes with 0.1% PBSA were performed after each incubation. The presented values describing the mean fluorescence of biotinylated lysozyme and biotinylated BSA binding were background subtracted. Here, the mean background fluorescence originates from the signals produced after incubating each cell population with SA-PE alone. Similarly, the mean fluorescence values associated with HA expression were background subtracted. For these samples, the mean background fluorescence was derived from the signals produced after the incubation of each cell population with the

DAR488 detection antibody alone. For each experimental sample, the background subtracted mean fluorescence values for lysozyme and BSA binding were further divided by the background subtracted mean fluorescence values associated with HA expression for that particular cell population. Error bars represent the standard error of the mean for three independent experiments.

Similar immunofluorescence experiments were performed to detect the binding of biotinylated BSA (5 μM), GFP (2 and 5 μM), and biotinylated IgG (2 and 5 μM) to yeast cells displaying each identified lysozyme binding cyclic peptide as well as DSG-treated, non-displaying EBY100 cells. Non-displaying EBY100 yeast were subjected to the DSG crosslinking procedure in a similar manner as the yeast cells displaying the linear peptide sequences. Cells were also labeled with SA-PE alone as a negative control. Biotinylated Human IgG was purchased from Immunoreagents. The mean fluorescence values derived from the signals produced after labeling with the biotinylated proteins were background subtracted using one of the following approaches: 1) by subtracting the mean fluorescence originating from the signals produced after labeling with SA-PE alone (comparison of biotinylated lysozyme vs biotinylated BSA) or 2) by subtracting the mean fluorescence stemming from the signals produced when labeling the DSG-treated EBY100 cells with the same respective protein and concentration (comparison of biotinylated lysozyme vs biotinylated IgG). The GFP protein used was not biotinylated. Its binding was detected using the 488 laser within the flow cytometer. Accordingly, the mean fluorescence values associated with the signals produced after incubation of the cells with GFP were not background subtracted.

We recombinantly expressed and purified the GFP protein used in this analysis. Specifically, GFP was expressed in Rosetta *E.coli* cells using the plasmid pet22B. DNA for GFP was inserted between the NdeI and XhoI cut sites of pet22B. Expression was carried out in 2XYT media (10 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl) plus 1X ampicillin. A 5 mL overnight

culture was used to inoculate a 1 L culture of 2XYT. GFP expression was induced using 0.5 mM isopropyl β-D-1-thiogalactopyr-anoside (IPTG) when the OD_{600} reached between 0.6 and 0.8. Expression was performed overnight at 20°C. The GFP protein was purified using hydrophobic interaction chromatography. Briefly, the induced cells were collected and lysed using sonication in 35 mL of Buffer A (50 mM sodium phosphate, 1.5 M ammonium sulfate, pH 7.0). After, the supernatant collected from lysis was loaded onto a 5 mL HiTrap Phenyl FF (High Sub) column (GE Healthcare Life Sciences), washed with 40 mL of Buffer A, and eluted over a 40 mL linear gradient of Buffer B (50 mM sodium phosphate, pH 7.0). Pure fractions were dialyzed into PBS 7.4.

5.4.10 Screening a cyclic heptapeptide yeast display library against IL-17

Cyclic peptides with affinity for IL-17 were discovered by performing magnetic selections against yeast-displayed IL-17 followed by FACS selections against soluble, biotinylated IL-17. Initially, IL-17 was co-displayed on the yeast surface with SsoFe2, an iron oxide binding protein, using the pCT302-T2A-SSoFe2 plasmid as previously described[94]. To generate pCT302-T2A-SSoFe2-IL17 (-Leu), DNA corresponding to human IL-17 (Met1-Ala155, gene block 3) was amplified using primers Pf4 and Pr4 followed by insertion between the NheI and BamHI sites of the pCT302-T2A-SSoFe2 plasmid. For the magnetic selection, a negative selection against magnetic yeast cells expressing TOM22[94] (-Leu) and a positive selection against magnetic yeast cells expressing IL-17 was performed. To begin, $7x10^8$ cells co-expressing IL-17 and SsoFe2 were resuspended in 10 mL of 1% PBS BSA, pH 7.4 (1% PBSA). Cells co-expressing TOM22 and SsoFe2 were also treated in a similar manner. The cells were magnetized by adding 2 mL of an iron oxide solution (4 mg/mL in water) followed by a 20-minute, rotating incubation at room temperature. Any unbound cells were removed. The magnetized cells were then washed 3X with

1% PBSA. Based on the OD_{600} of the solution prior to the addition of the iron oxide and after the removal of the yeast bound to the iron oxide, $1x10^8$ magnetized cells were aliquoted for the magnetic selections. The aliquoted cells were blocked with 1% PBSA for 1 hour at room temperature (10 mL). Subsequently, the TOM22 magnetized cells were washed 3X with 0.1% PBSA prior to the addition of $5x10^8$ DSG-cyclized heptapeptide library cells and $1x10^9$ EBY100 cells in 10 mL of 0.1% PBSA. After an hour incubation, library cells that did not bind the magnetized TOM22 cells were recovered. The IL-17 magnetized cells were washed three times with 0.1% PBSA after blocking and then incubated for 1 hour with the cells recovered from the negative selection at room temperature. Library cells bound to the magnetized IL-17 cells were isolated using a magnet. Magnetic IL-17 cells and any complexed library cells were washed 5 times with 0.1% PBSA prior to expansion in 20 mL of SDCAA (-TRP) media.

Using the library cells isolated from the magnetic selection, FACS selections were performed against soluble IL-17 in a similar manner as the FACS selections against lysozyme. Briefly, $1x10^7$ cyclized, library cells were labeled with biotinylated IL-17 (R&D systems) for 1 hour on ice followed by secondary labeling using SA-PE (1:250 dilution) for 12 minutes on ice. The cells were labeled with 1 μM and 500 nM of biotinylated IL-17 for the first and second rounds of FACS, respectively. All labelings took place in 100 μL of 0.1% PBSA. Washes with 0.1% PBSA were performed after each incubation. DNA was extracted from the yeast cells isolated from the final IL-17 FACS round using a Zymoprep™ Yeast Plasmid Miniprep II kit following the manufacturer's recommendations. DNA encoding the isolated cyclic peptide sequences was amplified from the extracted plasmid DNA using primers Pf5 and Pr5. Multiple rounds of PCR were performed to obtain the concentration of DNA required for sequencing. After the first round of PCR, amplified DNA was purified by phenol: chloroform extraction and concentrated using

ethanol precipitation, as previously described. For subsequent PCR rounds, the amplified DNA was purified using the BioBasic 9K Series Gel and PCR extraction kit. Partial Illumina adapters were added to the amplified DNA by Genewiz. The DNA was sequenced using Genewiz's Amplicon Ez protocol. Genewiz provided bioinformatic details about the recovered sequences, including how many times a particular sequence occurred in the sequenced population. In total, 102,896 reads were obtained after performing Illumina sequencing on the final FACS population.

5.4.11 *In silico* evaluation of IL-17:cyclic peptide complexes

The crystal structure of IL-17 (PDB ID: 4HSA)[96] was initially prepared using Schrödinger's Protein Preparation Wizard[97,98] to identify and correct missing atoms and/or side chains, remove salt ions and other small molecules, add explicit hydrogens, assign tautomeric states with EPIK, optimize the hydrogen-bonding networks, and minimize the protein energy using the OPLS force field[76–78]. The adjusted structure was subjected to a "druggability" study using SiteMap to identify putative binding sites capable of accommodating the cyclic heptapeptides[119–121]. The SiteMap algorithm by Schrödinger identifies binding pockets by locating spherical "site points" onto the protein surface; these points are clustered based on *(i)* their ability to form favorable protein-ligand interactions, *(ii)* solvent exposure, and *(iii)* hydrophobic/philic character. The regions on the protein surface that possess a sufficient number of site points and volume are scored using an S-score (likelihood of the protein's surface to be a binding pocket) and D-score (measure of the pocket's "druggability"). Sites with S-score and D-score values greater than or equal to 0.7 and 0.9, respectively, were selected as putative sites for ligand binding. In this work, given the topology of IL-17, the SiteMap analysis was performed with the '*detect shallow binding sites*' option selected, which adjusts amino acid atomic van der Waals radii to be more accommodating for peptide binding when locating potential binding pockets.

An ensemble, containing 17 cyclic peptides identified from the library screen, was designed using the molecular editor Avogadro[74,75], and equilibrated via atomistic molecular dynamics (MD) simulations in the GROMACS package[122–124] using the OPLS all-atom force field[76–78] and periodic boundary conditions[125–127]. Every peptide was placed in a simulation box with periodic boundary conditions containing 800 water molecules (TIP3P water model)[128]. The system was minimized by running 10,000 steps of the steepest gradient descent, heated to 300 K in an NVT ensemble for 250 ps (1 fs time steps), and equilibrated to 1 atm by running a 500-ps NPT simulation (2 fs time steps). The production run for every peptide was performed in the NPT ensemble at 300 K and 1 atm using the Nosé-Hoover thermostat[129–131] and Parrinello-Rahman barostat, respectively[132,133]. The leap-frog algorithm was used to integrate the equations of motion. All covalent bonds were constrained by means of the LINCS algorithm[134]. The short-range electrostatic and Lennard-Jones interactions were calculated within a cutoff of 1.0 nm and 1.2 nm, respectively, whereas the particle-mesh Ewald method was utilized to treat the long-range electrostatic interactions[135–137]. The atomic coordinates were saved every 2 ps, and the non-bonded interaction pair-list was updated every 5 fs using a cutoff of 1.4 nm.

The resulting IL-17 specific peptide structures were then docked *in silico* against the selected putative binding sites on IL-17 using the docking software HADDOCK (High Ambiguity Driven Protein-Protein Docking) (v.2.1)[99–101]. The residues within IL-17's binding sites were defined as "active", whereas the residues surrounding the binding sites were defined as "passive". All variable amino acid positions on the peptide ligands were also denoted as 'active'. A GSG tripeptide segment located on the C-terminus of the peptide was defined as not involved in the interaction to account for the directionality of binding. When displayed on the surface of yeast, in fact, the peptide is connected to the Aga2p display system through its C-terminus. Docking in

HADDOCK proceeded through a 3-stage protocol: *(i)* rigid, *(ii)* semi-flexible, and *(iii)* water-refined fully flexible docking. A total of 1000, 200, and 200 structures were calculated at each stage, respectively. Final structures were clustered using the cutoff of Cα RMSD < 7.5 Å as calculated in ProFit (Martin, A.C.R, http://www.bioinf.org.uk/software/profit/). The peptides in the identified clusters were ranked using FireDock[138,139] and Xscore[140,141]. The selected binding poses were finally refined via 100-ns atomistic MD simulations using the GROMACS simulation package. The IL-17:peptide complexes were embedded in a cubic box of 9.7 nm side and periodic boundary and solvated with 30,000 TIP3P water molecules. The MD simulations were performed at 300 K and 1 atm using the Amber99SB force field. The MM/GBSA method was used for processing the refined IL-17:peptide complexes and to estimate the free energy of binding (ΔG_B)[102,103]. The *in silico* predicted K_D values were calculated directly from the estimated ΔG_B values assuming $\Delta G_B = RT\ln(K_d)$. The binding of IL-17RA to IL-17 was superimposed over the refined IL-17:peptide complexes to visualize if the peptides bind to IL-17 in a similar region as IL-17RA. Similar methods were also used to simulate the binding of the linear version of cyclo-[*DSG*-MRWLRGRR-*K*] to IL-17.

When evaluating the effect of peptide length on cyclization efficiency (Figure 5.4), similar methods were used to simulate the peptide structures based off the parent sequence RMWFPHYK. These peptides were designed using the molecular editor Avogadro followed by equilibration using atomistic MD simulations within the GROMACS package as well as the OPLS all-atom force field and all other previously described parameters and algorithms.

5.4.12 Cloning of IL-17 cyclic peptide sequences for yeast display

Gene blocks 4 – 9 encoding the linear peptide sequences for cyclo-[*DSG*-RMRWLRGRR-*K*], cyclo-[*DSG*-RMIGQWWRR-*K*], cyclo-[*DSG*-RMNRLKFWF-*K*], cyclo-[*DSG*-

RMSFFDIWR-*K*], cyclo-[*DSG*-RMYRFHRHG-*K*], and cyclo-[*DSG*-RMFGLLHRG-*K*], respectively, were amplified using primers Pf6 and Pr6. These gene blocks were digested and inserted between the EcoRI and XmaI cut sites of the pCT-NT-F2A-Sso7dhFc vector[118].

5.4.13 Affinity and binding selectivity characterization of the IL-17 cyclic peptides

The interaction of the yeast-displayed IL-17 binding cyclic peptides with IL-17 and other putative, non-specific proteins was evaluated using flow cytometry. Induced yeast cells displaying linear peptide sequences were cyclized using the DSG crosslinking protocol as previously described. Non-displaying EBY100 yeast cells were also subjected to the DSG crosslinking protocol. The binding of biotinylated IL-17 at a concentration of 500 nM was studied for yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*], cyclo-[*DSG*-RMIGQWWRR-*K*], cyclo-[*DSG*-RMNRLKFWF-*K*], cyclo-[*DSG*-RMSFFDIWR-*K*], cyclo-[*DSG*-RMYRFHRHG-*K*], and cyclo-[*DSG*-RMFGLLHRG-*K*] as well as for DSG-treated EBY100 cells. The binding of biotinylated lysozyme (500 & 1000 nM), biotinylated BSA (500 & 1000 nM), GFP (500 & 1000 nM), and SA-PE (500 nM) to yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] was evaluated using immunofluorescence detection. Binding of GFP and SA-PE at the described concentrations was also evaluated for the DSG-treated EBY100 cells. All primary incubations using biotinylated protein took place for 30 minutes at room temperature. The binding of biotinylated protein was detected by flow cytometry following SA-PE labeling (1:250 dilution, 12 minutes on ice in the dark). When evaluating the binding of GFP and SA-PE alone, no secondary detection agent was used as these proteins, by themselves, can provide a fluorescent signal that, when excited, can be detected via flow cytometry. When cells were only labeled with GFP and SA-PE, incubations took place in the dark for 30 minutes at room temperature. All labelings took place in a volume of 50 µL using 0.1% PBSA. Washes with 0.1% PBSA were performed after each incubation.

Background subtracted mean fluorescence values describing the binding of biotinylated IL-17, biotinylated lysozyme, and biotinylated BSA are reported. Here, the mean background fluorescence originates from the signals produced after incubating the cells with SA-PE alone. The background subtracted mean fluorescence values were then multiplied by a factor of 100. The mean fluorescence values describing the signals produced by GFP and SA-PE binding to the cells were not background subtracted and only multiplied by a factor of 100. Error bars represent the standard error of the mean for three individual replicates.

5.4.14 *In vitro* evaluation of cyclo-[DSG-MRWLRGRR-K] as an inhibitor

Two approaches were used to study if cyclo-[*DSG*-MRWLRGRR-*K*] modulates the interaction between IL-17 and its receptor. In approach one, yeast cells displaying IL-17 were incubated with soluble cyclo-[*DSG*-MRWLRGRR-*K*] and soluble IL-17RA (R&D Systems). The UNC High-Throughput Peptide Synthesis and Array facility synthesized a soluble form of cyclo-[*DSG*-MRWLRGRR-*K*][44]. This peptide lacked the N-terminal arginine residue present in the yeast-displayed peptide sequence. To confer expression of IL-17 as a yeast surface fusion, gene block 3 was amplified using primers Pf4 and Pr4 and inserted between the NheI and BamHI sites of pCTCON to generate pCTCON-IL17[142]. $2x10^6$ yeast cells displaying IL-17 were incuabted overnight at 4°C with varying concentrations of soluble cyclo-[*DSG*-MRWLRGRR-*K*] (20 nM – 50 μM) and either 20 or 50 nM of IL-17RA. All incubations took place in 100 μL of 0.1% PBSA. A control was also performed for each concentration of IL-17RA where no peptide was present. The next morning the cells were incubated with a 1:100 dilution of biotinylated protein A (Thermo Fisher Scientific) for 30 minutes at room temperature (50 μL labeling in 0.1% PBSA). Thereafter, the cells were labeled with a 1:250 dilution of SA-PE for 12 minutes on ice (50 μL labeling in 0.1% PBSA) and subsequently analyzed by flow cytometry. A wash with 0.1% PBSA took place

between all incubations. A negative control was also performed where cells were incubated with biotinylated protein A followed by SA-PE in a similar fashion. The reported values describing the mean fluorescence of IL-17RA binding to the IL-17 displaying cells were background subtracted. Here, the mean background fluorescence stems from the signals produced after incubating the cell population with biotinylated protein A (1:100 dilution) and SA-PE (1:250 dilution). The background subtracted mean fluorescence values were then multiplied by a factor 100.

The second approach differed by incubating yeast-displayed cyclo-[*DSG*-RMRWLRGRR-*K*] with soluble IL-17 and soluble IL-17RA. Specifically, $2x10^6$ yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] were incubated with 100 nM of biotinylated IL-17 in the presence or absence of 500 nM IL-17RA overnight at 4°C in a total volume of 100 μL using 0.1% PBSA. The following day, the cells were washed and incubated with a 1:250 dilution of SA-PE for 12 minutes on ice (50 μL labeling in 0.1% PBSA). SA-PE binding was detected using flow cytometry. A secondary only labeling using SA-PE alone was also performed. The reported mean fluorescence values stemming from the signals produced by IL-17 binding to the peptide displaying cells were not background subtracted. A two-tailed, two sample unequal variance t-test was performed using individual data points from the flow cytometry analysis (n=390,763) comparing the mean fluorescence of IL-17 binding for the populations where IL-17RA was present or absent. Data points used in the t-test were not background subtracted. The binding of IL-17 to the yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] statistically differed when the receptor was present ($p<0.001$).

Additionally, a control experiment was performed where yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] were incubated with IL-17RA. $2x10^6$ yeast cells displaying cyclo-[*DSG*-RMRWLRGRR-*K*] were incuabted with 500 nM of IL-17RA overnight at 4ºC in a total

volume of 100 μL using 0.1% PBSA followed by labeling with biotinylated protein A (1:100 dilution) for 30 minutes at room temperature and SA-PE labeling for 12 minutes on ice in the dark (1:250 dilution). Non-displaying EBY100 yeast subjected to the DSG crosslinking procedure were also labeled in a similar manner. All protein A and SA-PE labelings took place in 50 μL using 0.1% PBSA. Washes took place between each incubation with 0.1% PBSA. Binding was detected using flow cytometry. A negative control was also performed where each cell population was incubated with biotinylated protein A followed by SA-PE in a similar fashion. The reported values describing the mean fluorescence of IL-17RA binding to each cell population were background subtracted. Here, the mean background fluorescence stems from the signals produced after incubating each cell population with biotinylated protein A and SA-PE. The background subtracted mean fluorescence values were then multiplied by a factor 100.

5.5 References

(1) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class. *Nature Reviews Drug Discovery*. July 2008, pp 608–624.

(2) Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **2011**, *54* (7), 1961–2004.

(3) Mallinson, J.; Collins, I. Macrocycles in New Drug Discovery. *Future Medicinal Chemistry*. July 2012, pp 1409–1438.

(4) Choi, J. S.; Joo, S. H. Recent Trends in Cyclic Peptides as Therapeutic Agents and Biochemical Tools. *Biomolecules and Therapeutics*. Korean Society of Applied Pharmacology 2020, pp 18–24.

(5) Horswill, A. R.; Benkovic, S. J. Cyclic Peptides, a Chemical Genetics Tool for Biologists. *Cell Cycle*. Taylor and Francis Inc. 2005, pp 552–555.

(6) Bird, G. H.; Madani, N.; Perry, A. F.; Princiotto, A. M.; Supko, J. G.; He, X.; Gavathiotis, E.; Sodroski, J. G.; Walensky, L. D. Hydrocarbon Double-Stapling Remedies the Proteolytic Instability of a Lengthy Peptide Therapeutic. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (32), 14093–14098.

(7) Szewczuk, Z.; Gibbs, B. F.; Yue, S. Y.; Purisima, E. O.; Konishi, Y. Conformationally Restricted Thrombin Inhibitors Resistant to Proteolytic Digestion. *Biochemistry* **1992**, *31* (38), 9132–9140.

(8) Nevola, L.; Giralt, E. Modulating Protein-Protein Interactions: The Potential of Peptides. *Chem. Commun.* **2015**, *51* (16), 3302–3315.

(9) Jaradat, D. M. M. Thirteen Decades of Peptide Synthesis: Key Developments in Solid Phase Peptide Synthesis and Amide Bond Formation Utilized in Peptide Ligation. *Amino Acids*. Springer-Verlag Wien January 1, 2018, pp 39–68.

(10) Qian, Z.; Rhodes, C. A.; McCroskey, L. C.; Wen, J.; Appiah-Kubi, G.; Wang, D. J.; Guttridge, D. C.; Pei, D. Enhancing the Cell Permeability and Metabolic Stability of Peptidyl Drugs by Reversible Bicyclization. *Angew. Chemie - Int. Ed.* **2017**, *56* (6), 1525–1529.

(11) Zorzi, A.; Deyle, K.; Heinis, C. Cyclic Peptide Therapeutics: Past, Present and Future. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24–29.

(12) Zheng, Y.; Ji, S.; Czerwinski, A.; Valenzuela, F.; Pennington, M.; Liu, S. FITC-Conjugated Cyclic RGD Peptides as Fluorescent Probes for Staining Integrin Avβ3/Avβ5 in Tumor Tissues. *Bioconjug. Chem.* **2014**, *25* (11), 1925–1941.

(13) Subiros-Funosas, R.; Mendive-Tapia, L.; Sot, J.; Pound, J. D.; Barth, N.; Varela, Y.; Goñi, F. M.; Paterson, M.; Gregory, C. D.; Albericio, F.; et al. A Trp-BODIPY Cyclic Peptide for Fluorescence Labelling of Apoptotic Bodies. *Chem. Commun.* **2017**, *53* (5), 945–948.

(14) Beer, A. J.; Haubner, R.; Goebel, M.; Luderschmidt, S.; Spilker, M. E.; Wester, H.-J.; Weber, W. A.; Schwaiger, M. Biodistribution and Pharmacokinetics of the Alphavbeta3-Selective Tracer 18F-Galacto-RGD in Cancer Patients. *J. Nucl. Med.* **2005**, *46* (8), 1333–1341.

(15) Lian, W.; Jiang, B.; Qian, Z.; Pei, D. Cell-Permeable Bicyclic Peptide Inhibitors against Intracellular Proteins. *J. Am. Chem. Soc.* **2014**, *136* (28), 9830–9833.

(16) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierskit, W. M.; Knappt, R. J. A New Type of Synthetic Peptide Library for Identifying Ligand-Binding Activity. *Nature* **1991**, *354* (6348), 82–84.

(17) Qian, Z.; Upadhyaya, P.; Pei, D. Synthesis and Screening of One-Bead-One-Compound Cyclic Peptide Libraries. *Methods Mol. Biol.* **2015**, *1248*, 39–53.

(18) Liu, R.; Marik, J.; Lam, K. S. A Novel Peptide-Based Encoding System for "One-Bead One-Compound" Peptidomimetic and Small Molecule Combinatorial Libraries. *J. Am. Chem. Soc.* **2002**, *124* (26), 7678–7680.

(19) 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), 12297–12302.

(20) Wilson, D. S.; Keefe, A. D.; Szostak, J. W. The Use of MRNA Display to Select High-Affinity Protein-Binding Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (7), 3750–3755.

(21) Barrick, J. E.; Takahashi, T. T.; Ren, J.; Xia, T.; Roberts, R. W. Large Libraries Reveal Diverse Solutions to an RNA Recognition Problem. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (22), 12374–12378.

(22) Mattheakis, L. C.; Bhatt, R. R.; Dower, W. J. An in Vitro Polysome Display System for Identifying Ligands from Very Large Peptide Libraries. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (19), 9022–9026.

(23) Gersuk, G. M.; Corey, M. J.; Corey, E.; Stray, J. E.; Kawasaki, G. H.; Vessella, R. L. High-Affinity Peptide Ligands to Prostate-Specific Antigen Identified by Polysome Selection. *Biochem. Biophys. Res. Commun.* **1997**, *232* (2), 578–582.

(24) Lamla, T.; Erdmann, V. A. Searching Sequence Space for High-Affinity Binding Peptides Using Ribosome Display. *J. Mol. Biol.* **2003**, *329* (2), 381–388.

(25) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science.* **1985**, *228* (4705), 1315–1317.

(26) Gram, H.; Schmitz, R.; Zuber, J. F.; Baumann, G. Identification of Phosphopeptide Ligands for the Src-Homology 2 (SH2) Domain of Grb2 by Phage Display. *Eur. J. Biochem.* **1997**, *246* (3), 633–637.

(27) Böttger, V.; Böttger, A.; Howard, S. F.; Picksley, S. M.; Chène, P.; Garcia-Echeverria, C.; Hochkeppel, H. K.; Lane, D. P. Identification of Novel Mdm2 Binding Peptides by Phage Display. *Oncogene* **1996**, *13* (10), 2141–2147.

(28) Zwick, M. B.; Shen, J.; Scott, J. K. Phage-Displayed Peptide Libraries. *Curr. Opin.*

Biotechnol. **1998**, *9* (4), 427–436.

(29) Kenrick, S. A.; Daugherty, P. S. Bacterial Display Enables Efficient and Quantitative Peptide Affinity Maturation. *Protein Eng. Des. Sel.* **2010**, *23* (1), 9.

(30) PH, B.; JJ, R.; PS, D. Rapid Isolation of High-Affinity Protein Binding Peptides Using Bacterial Display. *Protein Eng. Des. Sel.* **2004**, *17* (10).

(31) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-Encoded Combinatorial Chemical Libraries Based on Bicyclic Peptides. *Nat. Chem. Biol.* **2009**, *5* (7), 502–507.

(32) Bosma, T.; Kuipers, A.; Bulten, E.; de Vries, L.; Rink, R.; Moll, G. N. Bacterial Display and Screening of Posttranslationally Thioether-Stabilized Peptides. *Appl. Environ. Microbiol.* **2011**, *77* (19), 6794–6801.

(33) Urban, J. H.; Moosmeier, M. A.; Aumüller, T.; Thein, M.; Bosma, T.; Rink, R.; Groth, K.; Zulley, M.; Siegers, K.; Tissot, K.; et al. Phage Display and Selection of Lanthipeptides on the Carboxy-Terminus of the Gene-3 Minor Coat Protein. *Nat. Commun.* **2017**, *8* (1), 1–10.

(34) Hetrick, K. J.; Walker, M. C.; Van Der Donk, W. A. Development and Application of Yeast and Phage Display of Diverse Lanthipeptides. *ACS Cent. Sci.* **2018**, *4* (4), 458–467.

(35) McLafferty, M. A.; Kent, R. B.; Ladner, R. C.; Markland, W. M13 Bacteriophage Displaying Disulfide-Constrained Microproteins. *Gene* **1993**, *128* (1), 29–36.

(36) O'Neil, K. T.; Hoess, R. H.; Jackson, S. A.; Ramachandran, N. S.; Mousa, S. A.; DeGrado, W. F. Identification of Novel Peptide Antagonists for GPIIb/IIIa from a Conformationally Constrained Phage Peptide Library. *Proteins Struct. Funct. Genet.* **1992**, *14* (4), 509–515.

(37) Lupold, S. E.; Rodriguez, R. Disulfide-Constrained Peptides That Bind to the Extracellular Portion of the Prostate-Specific Membrane Antigen. *Mol. Cancer Ther.* **2004**, *3* (5), 597–603.

(38) Qian, Z.; Larochelle, J. R.; Jiang, B.; Lian, W.; Hard, R. L.; Selner, N. G.; Luechapanichkul, R.; Barrios, A. M.; Pei, D. Early Endosomal Escape of a Cyclic Cell-Penetrating Peptide Allows Effective Cytosolic Cargo Delivery. *Biochemistry* **2014**, *53* (24), 4034–4046.

(39) Qian, Z.; Dougherty, P. G.; Pei, D. Targeting Intracellular Protein–Protein Interactions with Cell-Permeable Cyclic Peptides. *Current Opinion in Chemical Biology*. Elsevier Ltd June 1, 2017, pp 80–86.

(40) Qian, Z.; Liu, T.; Liu, Y. Y.; Briesewitz, R.; Barrios, A. M.; Jhiang, S. M.; Pei, D. Efficient Delivery of Cyclic Peptides into Mammalian Cells with Short Sequence Motifs. *ACS Chem. Biol.* **2013**, *8* (2), 423–431.

(41) Gilbert, H. F. Thiol/Disulfide Exchange Equilibria and Disulfidebond Stability. *Methods in Enzymology*. January 1, 1995, pp 8–28.

(42) Millward, S. W.; Takahashi, T. T.; Roberts, R. W. A General Route for Post-Translational Cyclization of MRNA Display Libraries. *J. Am. Chem. Soc.* **2005**, *127* (41), 14142–14143.

(43) Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. Design of Cyclic Peptides That Bind Protein Surfaces with Antibody-like Affinity. *ACS Chem. Biol.* **2007**, *2* (9), 625–634.

(44) Menegatti, S.; Hussain, M.; Naik, A. D.; Carbonell, R. G.; Rao, B. M. MRNA Display Selection and Solid-Phase Synthesis of Fc-Binding Cyclic Peptide Affinity Ligands. *Biotechnol. Bioeng.* **2013**, *110* (3), 857–870.

(45) Howell, S. M.; Fiacco, S. V.; Takahashi, T. T.; Jalali-Yazdi, F.; Millward, S. W.; Hu, B.; Wang, P.; Roberts, R. W. Serum Stable Natural Peptides Designed by MRNA Display. *Sci. Rep.* **2014**, *4* (1), 1–5.

(46) Diderich, P.; Bertoldo, D.; Dessen, P.; Khan, M. M.; Pizzitola, I.; Held, W.; Huelsken, J.; Heinis, C. Phage Selection of Chemically Stabilized α-Helical Peptide Ligands. *ACS Chem. Biol.* **2016**, *11* (5), 1422–1427.

(47) Navaratna, T.; Atangcho, L.; Mahajan, M.; Subramanian, V.; Case, M.; Min, A.; Tresnak, D.; Thurber, G. M. Directed Evolution Using Stabilized Bacterial Peptide Display. *J. Am. Chem. Soc.* **2020**, *142* (4), 1882–1894.

(48) Hacker, D. E.; Hoinka, J.; Iqbal, E. S.; Przytycka, T. M.; Hartman, M. C. T. Highly Constrained Bicyclic Scaffolds for the Discovery of Protease-Stable Peptides via MRNA Display. *ACS Chem. Biol.* **2017**, *12* (3), 795–804.

(49) Hayashi, Y.; Morimoto, J.; Suga, H. In Vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors. *ACS Chem. Biol.* **2012**, *7* (3), 607–613.

(50) Kawamura, A.; Münzel, M.; Kojima, T.; Yapp, C.; Bhushan, B.; Goto, Y.; Tumber, A.; Katoh, T.; King, O. N. F.; Passioura, T.; et al. Highly Selective Inhibition of Histone Demethylases by de Novo Macrocyclic Peptides. *Nat. Commun.* **2017**, *8* (1), 1–10.

(51) Owens, A. E.; Iannuzzelli, J. A.; Gu, Y.; Fasan, R. MOrPH-PhD: An Integrated Phage Display Platform for the Discovery of Functional Genetically Encoded Peptide Macrocycles. *ACS Cent. Sci.* **2020**, *6* (3), 368–381.

(52) Bidlingmaier, S.; Su, Y.; Liu, B. Combining Phage and Yeast Cell Surface Antibody Display to Identify Novel Cell Type-Selective Internalizing Human Monoclonal Antibodies. *Methods Mol. Biol.* **2015**, *1319*, 51–63.

(53) VanAntwerp, J. J.; Wittrup, K. D. Fine Affinity Discrimination by Yeast Surface Display and Flow Cytometry. *Biotechnol. Prog.* **2000**, *16* (1), 31–37.

(54) Van Rosmalen, M.; Janssen, B. M. G.; Hendrikse, N. M.; Van Der Linden, A. J.; Pieters, P. A.; Wanders, D.; De Greef, T. F. A.; Merkx, M. Affinity Maturation of a Cyclic Peptide Handle for Therapeutic Antibodies Using Deep Mutational Scanning. *J. Biol. Chem.* **2017**, *292* (4), 1477–1489.

(55) Ishii, J.; Yoshimoto, N.; Tatematsu, K.; Kuroda, S.; Ogino, C.; Fukuda, H.; Kondo, A. Cell Wall Trapping of Autocrine Peptides for Human G-Protein-Coupled Receptors on the Yeast Cell Surface. *PLoS One* **2012**, *7* (5), e37136.

(56) Jeong, M.-Y.; Rutter, J.; Chou, D. H.-C. Display of Single-Chain Insulin-like Peptides on a Yeast Surface. *Biochemistry* **2019**, *58* (3), 182–188.

(57) Kintzing, J. R.; Cochran, J. R. Engineered Knottin Peptides as Diagnostics, Therapeutics, and Drug Delivery Vehicles. *Current Opinion in Chemical Biology*. Elsevier Ltd October

1, 2016, pp 143–150.

(58) Kimura, R. H.; Levin, A. M.; Cochran, F. V.; Cochran, J. R. Engineered Cystine Knot Peptides That Bind Avβ3, Avβ5, and A5β1 Integrins with Low-Nanomolar Affinity. *Proteins Struct. Funct. Bioinforma.* **2009**, *77* (2), 359–369.

(59) Kimura, R. H.; Jones, D. S.; Jiang, L.; Miao, Z.; Cheng, Z.; Cochran, J. R. Functional Mutation of Multiple Solvent-Exposed Loops in the Ecballium Elaterium Trypsin Inhibitor-II Cystine Knot Miniprotein. *PLoS One* **2011**, *6* (2), e16112.

(60) Silverman, A. P.; Levin, A. M.; Lahti, J. L.; Cochran, J. R. Engineered Cystine-Knot Peptides That Bind Avβ3 Integrin with Antibody-Like Affinities. *J. Mol. Biol.* **2009**, *385* (4), 1064–1075.

(61) Kjærulff, S.; Jensen, M. R. Comparison of Different Signal Peptides for Secretion of Heterologous Proteins in Fission Yeast. *Biochem. Biophys. Res. Commun.* **2005**, *336* (3), 974–982.

(62) Lee, M. A.; Cheong, K. H.; Shields, D.; Park, S. D.; Hong, S. H. Intracellular Trafficking and Metabolic Turnover of Yeast Prepro-α-Factor-SRIF Precursors in GH3 Cells. *Exp. Mol. Med.* **2002**, *34* (4), 285–293.

(63) Oka, C.; Tanaka, M.; Muraki, M.; Harata, K.; Suzuki, K.; Jigami, Y. Human Lysozyme Secretion Increased by Alpha-Factor Pro-Sequence in Pichi. *Biosci. Biotechnol. Biochem.* **1999**, *63* (11), 1977–1983.

(64) Brake, A. J.; Merryweather, J. P.; Coit, D. G.; Heberlein, U. A.; Masiarz, F. R.; Mullenbach, G. T.; Urdea, M. S.; Valenzuela, P.; Barr, P. J. α-Factor-Directed Synthesis and Secretion of Mature Foreign Proteins in Saccharomyces Cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81* (15 I), 4642–4646.

(65) Otte, S.; Barlowe, C. Sorting Signals Can Direct Receptor-Mediated Export of Soluble Proteins into COPII Vesicles. *Nat. Cell Biol.* **2004**, *6* (12), 1189–1194.

(66) Rakestraw, J. A.; Sazinsky, S. L.; Piatesi, A.; Antipov, E.; Wittrup, K. D. Directed Evolution of a Secretory Leader for the Improved Expression of Heterologous Proteins and Full-Length Antibodies in Saccharomyces Cerevisiae. *Biotechnol. Bioeng.* **2009**, *103* (6), 1192–1201.

(67) Caplan, S.; Green, R.; Rocco, J.; Kurjan, J. Glyosylation and Structure of the Yeast MFα1 α-Factor Precursor Is Important for Efficient Transport through the Secretory Pathway. *J. Bacteriol.* **1991**, *173* (2), 627–635.

(68) Chaudhuri, B.; Steube, K.; Stephan, C. The Pro-Region of the Yeast Prepro-Alpha-Factor Is Essential for Membrane Translocation of Human Insulin-like Growth Factor 1 in Vivo. *Eur. J. Biochem.* **1992**, *206* (3), 793–800.

(69) Fuller, R. S.; Sterne, R. E.; Thorner, J. Enzymes Required for Yeast Prohormone Processing. *Annu. Rev. Physiol.* **1988**, *50* (1), 345–362.

(70) Mizuno, K.; Nakamura, T.; Ohshima, T.; Tanaka, S.; Matsuo, H. Characterization of KEX2-Encoded Endopeptidase from Yeast Saccharomyces Cerevisiae. *Biochem. Biophys. Res.*

Commun. **1989**, *159* (1), 305–311.

(71) 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. *Biochemistry* **1998**, *37* (48), 16839–16845.

(72) Gera, N.; Hussain, M.; Rao, B. M. Protein Selection Using Yeast Surface Display. *Methods* **2013**, *60* (1), 15–26.

(73) Kieke, M. C.; Shusta, E. V.; Boder, E. T.; Teyton, L.; Wittrup, K. D.; Kranz, D. M. Selection of Functional T Cell Receptor Mutants from a Yeast Surface-Display Library. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (10), 5651–5656.

(74) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeerschd, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4* (8), 17.

(75) Avogadro: An Open-Source Molecular Builder and Visualization Tool. Version 1.2.0 Http://Avogadro.Cc/.

(76) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110* (6), 1657–1666.

(77) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the Role of the Crystal Environment in Determining Protein Side-Chain Conformations. *J. Mol. Biol.* **2002**, *320* (3), 597–608.

(78) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies Using Molecular Dynamics Free Energy Perturbation and the Opls Force Field. *J. Chem. Theory Comput.* **2010**, *6* (5), 1509–1519.

(79) Wetter, L. R.; Deutsch, H. F. Immunological Studies on Egg White Proteins. IV. Immunochemical And. *J. Biol. Chem.* **1951**, *192* (1), 237–242.

(80) Gera, N.; Hussain, M.; Wright, R. C.; Rao, B. M. Highly Stable Binding Proteins Derived from the Hyperthermophilic Sso7d Scaffold. *J. Mol. Biol.* **2011**, *409* (4), 601–616.

(81) Winiewska, M.; Bugajska, E.; Poznański, J. ITC-Derived Binding Affinity May Be Biased Due to Titrant (Nano)-Aggregation. Binding of Halogenated Benzotriazoles to the Catalytic Domain of Human Protein Kinase CK2. *PLoS One* **2017**, *12* (3), e0173260.

(82) Thakkar, A.; Trinh, T. B.; Pei, D. Global Analysis of Peptide Cyclization Efficiency. *ACS Comb. Sci.* **2013**, *15* (2), 120–129.

(83) Tavassoli, A. SICLOPPS Cyclic Peptide Libraries in Drug Discovery. *Curr. Opin. Chem. Biol.* **2017**, *38*, 30–35.

(84) McAllister, T. E.; Yeh, T. L.; Abboud, M. I.; Leung, I. K. H.; Hookway, E. S.; King, O. N. F.; Bhushan, B.; Williams, S. T.; Hopkinson, R. J.; Münzel, M.; et al. Non-Competitive Cyclic Peptides for Targeting Enzyme-Substrate Complexes. *Chem. Sci.* **2018**, *9* (20), 4569–4578.

(85) Kale, S. S.; Villequey, C.; Kong, X. D.; Zorzi, A.; Deyle, K.; Heinis, C. Cyclization of Peptides with Two Chemical Bridges Affords Large Scaffold Diversities. *Nat. Chem.* **2018**, *10* (7), 715–723.

(86) Benatuil, L.; Perez, J. M.; Belk, J.; Hsieh, C.-M. An Improved Yeast Transformation Method for the Generation of Very Large Human Antibody Libraries. *Protein Eng. Des. Sel.* **2010**, *23* (4), 155–159.

(87) Sieber, T.; Hare, E.; Hofmann, H.; Trepel, M. Biomathematical Description of Synthetic Peptide Libraries. *PLoS One* **2015**, *10* (6).

(88) Li, R.; Wu, Z.; Wangb, Y.; Ding, L.; Wang, Y. Role of PH-Induced Structural Change in Protein Aggregation in Foam Fractionation of Bovine Serum Albumin. *Biotechnol. Reports* **2016**, *9*, 46–52.

(89) Richards, D. P.; Stathakis, C.; Polakowski, R.; Ahmadzadeh, H.; Dovichi, N. J. Labeling Effects on the Isoelectric Point of Green Fluorescent Protein. *J. Chromatogr. A* **1999**, *853* (1–2), 21–25.

(90) Buchwalow, I.; Samoilova, V.; Boecker, W.; Tiemann, M. Non-Specific Binding of Antibodies in Immunohistochemistry: Fallacies and Facts. *Sci. Rep.* **2011**, *1*.

(91) McGeachy, M. J.; Cua, D. J.; Gaffen, S. L. The IL-17 Family of Cytokines in Health and Disease. *Immunity* **2019**, *50* (4), 892–906.

(92) Werner, J. L.; Gessner, M. A.; Lilly, L. M.; Nelson, M. P.; Metz, A. E.; Horn, D.; Dunaway, C. W.; Deshane, J.; Chaplin, D. D.; Weaver, C. T.; et al. Neutrophils Produce Interleukin 17A (IL-17A) in a Dectin-1- and IL-23-Dependent Manner during Invasive Fungal Infection. *Infect. Immun.* **2011**, *79* (10), 3966–3977.

(93) Liu, S.; Desharnais, J.; Sahasrabudhe, P. V.; Jin, P.; Li, W.; Oates, B. D.; Shanker, S.; Banker, M. E.; Chrunyk, B. A.; Song, X.; et al. Inhibiting Complex IL-17A and IL-17RA Interactions with a Linear Peptide. *Sci. Rep.* **2016**, *6*.

(94) Bacon, K.; Burroughs, M.; Blain, A.; Menegatti, S.; Rao, B. M. Screening Yeast Display Libraries against Magnetized Yeast Cell Targets Enables Efficient Isolation of Membrane Protein Binders. *ACS Comb. Sci.* **2019**, *21* (12), 817–832.

(95) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* **1982**, *157* (1), 105–132.

(96) Liu, S.; Song, X.; Chrunyk, B. A.; Shanker, S.; Hoth, L. R.; Marr, E. S.; Griffor, M. C. Crystal Structures of Interleukin 17A and Its Complex with IL-17 Receptor A. *Nat. Commun.* **2013**, *4*, 1888.

(97) Schrödinger Release 2020-1: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2020; Impact, Schrödinger, LLC, New York, NY, 2020; Prime, Schrödinger, LLC, New York, NY, 2020.

(98) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput. Aided. Mol. Des.* **2013**, *27* (3), 221–234.

(99) Dominguez, C.; Boelens, R.; Bonvin, A. M. J. J. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731–1737.

(100) De Vries, S. J.; Van Dijk, M.; Bonvin, A. M. J. J. The HADDOCK Web Server for Data-Driven Biomolecular Docking. *Nat. Protoc.* **2010**, *5* (5), 883–897.

(101) Van Zundert, G. C. P.; Rodrigues, J. P. G. L. M.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; Van Dijk, M.; De Vries, S. J.; Bonvin, A. M. J. J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428* (4), 720–725.

(102) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51* (1), 69–82.

(103) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* **2015**, *10* (5), 449–461.

(104) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. the Accuracy of Ranking Poses Generated from Docking. *J. Comput. Chem.* **2011**, *32* (5), 866–877.

(105) Zhang, X.; Perez-Sanchez, H.; C. Lightstone, F. A Comprehensive Docking and MM/GBSA Rescoring Study of Ligand Recognition upon Binding Antithrombin. *Curr. Top. Med. Chem.* **2017**, *17* (14), 1631–1639.

(106) Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M. Fast and Accurate Predictions of Binding Free Energies Using MM-PBSA and MM-GBSA. *J. Comput. Chem.* **2010**, *31* (4), 797–810.

(107) Xu, L.; Sun, H.; Li, Y.; Wang, J.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. the Impact of Force Fields and Ligand Charge Models. *J. Phys. Chem. B* **2013**, *117* (28), 8408–8421.

(108) Sun, H.; Li, Y.; Tian, S.; Xu, L.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 4. Accuracies of MM/PBSA and MM/GBSA Methodologies Evaluated by Various Simulation Protocols Using PDBbind Data Set. *Phys. Chem. Chem. Phys.* **2014**, *16* (31), 16719–16729.

(109) Weng, G.; Wang, E.; Chen, F.; Sun, H.; Wang, Z.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 9. Prediction Reliability of Binding Affinities and Binding Poses for Protein-Peptide Complexes. *Phys. Chem. Chem. Phys.* **2019**, *21* (19), 10135–10145.

(110) Derda, R.; Tang, S. K. Y.; Li, S. C.; Ng, S.; Matochko, W.; Jafari, M. R. Diversity of Phage-Displayed Libraries of Peptides during Panning and Amplification. *Molecules*. Multidisciplinary Digital Publishing Institute (MDPI) February 2011, pp 1776–1803.

(111) Abdeen, S. J.; Swett, R. J.; Feig, A. L. Peptide Inhibitors Targeting Clostridium Difficile Toxins A and B. *ACS Chem. Biol.* **2010**, *5* (12), 1097–1103.

(112) Orta-Ramirez, A.; Merrill, J. E.; Smith, D. M. PH Affects the Thermal Inactivation Parameters of R-Phycoerythrin from Porphyra Yezoensis. *J. Food Sci.* **2000**, *65* (6), 1046–1050.

(113) Michael Green, N. Avidin and Streptavidin. *Methods Enzymol.* **1990**, *184* (C), 51–67.

(114) Helmer, D.; Schmitz, K. Peptides and Peptide Analogs to Inhibit Protein-Protein Interactions. In *Advances in Experimental Medicine and Biology*; Springer New York LLC, 2016; Vol. 917, pp 147–183.

(115) Ely, L. K.; Fischer, S.; Garcia, K. C. Structural Basis of Receptor Sharing by Interleukin 17 Cytokines. *Nat. Immunol.* **2009**, *10* (12), 1245–1251.

(116) Tiruthani, K.; Mischler, A.; Ahmed, S.; Mahinthakumar, J.; Haugh, J. M.; Rao, B. M. Design and Evaluation of Engineered Protein Biosensors for Live-Cell Imaging of EGFR Phosphorylation. *Sci. Signal.* **2019**, *12* (584), eaap7584.

(117) Carlin, K. B.; Cruz-Teran, C. A.; Kumar, J. P.; Gomes, C.; Rao, B. M. Combinatorial Pairwise Assembly Efficiently Generates High Affinity Binders and Enables a "Mix-and-Read" Detection Scheme. *ACS Synth. Biol.* **2016**, *5* (12), 1348–1354.

(118) Cruz-Teran, C. A.; Tiruthani, K.; Mischler, A.; Rao, B. M. Inefficient Ribosomal Skipping Enables Simultaneous Secretion and Display of Proteins in *Saccharomyces Cerevisiae*. *ACS Synth. Biol.* **2017**, acssynbio.7b00144.

(119) Halgren, T. New Method for Fast and Accurate Binding-Site Identification and Analysis. *Chem. Biol. Drug Des.* **2007**, *69* (2), 146–148.

(120) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49* (2), 377–389.

(121) Schrödinger Release 2020-1: SiteMap, Schrödinger, LLC, New York, NY, 2020.

(122) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.

(123) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. Long-Timescale Molecular Dynamics Simulations of Protein Structure and Function. *Curr. Opin. Struct. Biol.* **2009**, *19* (2), 120–127.

(124) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1–3), 43–56.

(125) Nguyen, P. H.; Mu, Y.; Stock, G. Structure and Energy Landscape of a Photoswitchable Peptide: A Replica Exchange Molecular Dynamics Study. *Proteins Struct. Funct. Bioinforma.* **2005**, *60* (3), 485–494.

(126) Nguyen, P. H.; Gorbunov, R. D.; Stock, G. Photoinduced Conformational Dynamics of a Photoswitchable Peptide: A Nonequilibrium Molecular Dynamics Simulation Study. *Biophys. J.* **2006**, *91* (4), 1224–1234.

(127) Nguyen, P. H.; Stock, G. Nonequilibrium Molecular Dynamics Simulation of a Photoswitchable Peptide. *Chem. Phys.* **2006**, *323* (1), 36–44.

(128) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(129) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52* (2), 255–268.

(130) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695–1697.

(131) Fu, J.; Yang, H.; Wang, J. Computational Design of the Helical Hairpin Structure of Membrane-Active Antibacterial Peptides Based on RSV Glycoprotein Epitope Scaffold. *Comput. Biol. Chem.* **2018**, *73*, 200–205.

(132) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(133) Yu, H.; Lin, Y. S. Toward Structure Prediction of Cyclic Peptides. *Phys. Chem. Chem. Phys.* **2015**, *17* (6), 4210–4219.

(134) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

(135) Cheatham, T. E. I.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *J. Am. Chem. Soc.* **1995**, *117* (14), 4193–4194.

(136) Quimbar, P.; Malik, U.; Sommerhoff, C. P.; Kaas, Q.; Chan, L. Y.; Huang, Y. H.; Grundhuber, M.; Dunse, K.; Craik, D. J.; Anderson, M. A.; et al. High-Affinity Cyclic Peptide Matriptase Inhibitors. *J. Biol. Chem.* **2013**, *288* (19), 13885–13896.

(137) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.

(138) Mashiach, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H. J. FireDock: A Web Server for Fast Interaction Refinement in Molecular Docking. *Nucleic Acids Res.* **2008**, *36* (Web Server issue), W229-32.

(139) Andrusier, N.; Nussinov, R.; Wolfson, H. J. FireDock: Fast Interaction Refinement in Molecular Docking. *Proteins Struct. Funct. Genet.* **2007**, *69* (1), 139–159.

(140) Wang, R.; Lai, L.; Wang, S. Further Development and Validation of Empirical Scoring Functions for Structure-Based Binding Affinity Prediction. *J. Comput. Aided. Mol. Des.* **2002**, *16* (1), 11–26.

(141) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46* (12), 2287–2303.

(142) Gera, N.; Hill, A. B.; White, D. P.; Carbonell, R. G.; Rao, B. M. Design of PH Sensitive

Binding Proteins from the Hyperthermophilic Sso7d Scaffold. *PLoS One* **2012**, *7* (11), e48928.

CHAPTER 6

Conclusions and Future Perspectives

6.1 Conclusions

In this work, we have extended the capabilities of the yeast surface display platform to enable efficient identification and characterization of affinity ligands. The use of affinity ligands evolved from non-antibody scaffold templates has increased as directed evolution and combinatorial library screening methods have progressed. Through iterative rounds of selection, proteins or peptides with novel binding affinity can be isolated from combinatorial libraries. Despite the success of identifying proteins with novel binding activity using established combinatorial techniques, limitations still exist when considering target protein expression and screening combinatorial libraries to identify cyclic peptide binders. To address these limitations, we have integrated the use of yeast-displayed targets into the selection of combinatorial libraries as well as the characterization of isolated binders and have additionally demonstrated a novel yeast display platform for the selection of cyclic peptides using fluorescence activated cell sorting (FACS).

In **Chapter 2**, we introduced the use of yeast-displayed targets as an alternative format for target protein expression when screening combinatorial libraries. The efficient isolation of library mutants that interact with the target-displaying cells is dependent on the magnetization of the cells displaying the target. The target-displaying cells are magnetized after incubation with iron oxide nanoparticles, which bind to a co-expressed SsoFe2 protein that was evolved to have affinity for iron oxide. In this chapter, we screened yeast display combinatorial libraries against magnetic yeast cells expressing the extracellular domain of a membrane protein target. Using this platform, we isolated binders specific to the cytosolic domain of the mitochondrial membrane protein TOM22 (K_D ~ 272 – 1934 nM) and the extracellular domain of the c-Kit receptor (K_D ~ 93 to K_D > 2000 nM). We also showed that the identified TOM22 binding proteins could be used to deplete

mitochondria from a cell lysate when immobilized onto magnetic beads, demonstrating how moderate affinity binding proteins can be used for biologics capture by taking advantage of the avidity effect. Moreover, the successful enrichment of mitochondria using the identified TOM22 binding proteins provides support that binding proteins isolated from selections against magnetic yeast targets are functional in the context of a natively expressed membrane protein. We expect the use of magnetic yeast cell targets will improve the efficiency of identifying binding proteins from combinatorial libraries, in particular for protein targets, like those from eukaryotes, that may not be appropriately expressed by *E.coli*.

In **Chapter 3,** we demonstrated a yeast-yeast two hybrid platform (qYY2H) that can be used to quantitatively assess the binding affinity (K_D) between two proteins without the need for soluble protein. This system relies on the interaction of yeast cells co-expressing one of the proteins (prey) and luciferase with magnetic yeast cells displaying the other protein (bait). The number of prey cells isolated by the magnetic bait cells can be quantified using a luminescence based assay. Due to the multivalent nature of yeast-yeast binding, we have been able to detect interactions as low as $K_D \sim 25$ μM using the qYY2H platform, suggesting that qYY2H can be used to assess even weak affinity interactions. We have also engineered the platform to study protein-protein interactions that rely on a post-translationally modified bait through the co-expression of a modifying enzyme. After quantifying the interactions of multiple bait and prey using this system, we found that the strength of the multivalent interaction between bait and prey yeast is quantitatively related to the binding affinity of the monovalent bait-prey interaction (K_D). Accordingly, we exploited this relationship to develop a mathematical framework to quantitatively estimate the binding affinity (K_D) of monovalent bait-prey interactions using measurements of yeast-yeast binding. In particular, we believe qYY2H will improve the efficiency of characterizing

binders isolated from combinatorial library screens as well as novel protein-protein interactions identified from cDNA library screens as the need for soluble protein is eliminated.

In **Chapter 4**, we extended the use of yeast-displayed targets to isolate binders from a mRNA display library of cyclic peptides. While the yeast target cells could be separated from the soluble mRNA-peptide fusions using centrifugation, magnetization helps improve the efficiency of removing unbound fusions. Here, the yeast targets are magnetized via non-specific surface adsorption of iron oxide nanoparticles rather than relying on an engineered protein with affinity for iron oxide. Non-specific magnetization represents a simpler approach to magnetize yeast cells as it does not require co-expression of SsoFe2. However, it has several potential disadvantages; Namely, a potential loss of positively bound library mutants if the iron oxide dissociates from the target cells over time as well as the occlusion of the displayed target proteins by the iron oxide particles as the particles can bind anywhere on the cell surface. Despite these disadvantages, we isolated cyclic peptide binders specific to TOM22, as confirmed using the qYY2H assay. The binding affinity of the isolated peptide was likely in the low micromolar range, which is not unexpected for cyclic peptides containing ~6-10 amino acids. We recommend using more stringent selection conditions as well as more rounds of selection if greater affinity mutants are desired.

In **Chapter 5**, we developed a novel yeast display platform for the isolation of cyclic peptide binders from chemically crosslinked peptide libraries that relies on FACS for selection. Here, we cyclized linear peptides displayed on the yeast surface using disuccinimidyl glutarate that crosslinks amine groups within the peptide sequence. In comparison to traditional display platforms that rely on panning for selection, the use of a yeast display platform for library generation enables the integration of FACS to improve the efficiency of isolating high affinity binders. Moreover, we have demonstrated how yeast-displayed cyclic peptides can be used to

estimate binding affinities and evaluate specificity, which has traditionally relied on chemically synthesized peptide. Using this approach, we isolated and characterized cyclic peptide binders specific to lysozyme (K_D ~1.2 – 3.7 μM) and interleukin-17 (IL-17; apparent K_D ~ 300 nM). We also provided an experimental and *in silico* framework for evaluating the potential use of cyclic peptide binders as inhibitors of protein-protein interactions, specifically in regard to the interaction of IL-17 and its receptor. Ultimately, we believe this platform will accelerate the identification and characterization of cyclic peptide binders for applications that require moderate affinity biorecognition agents, like protein purification or biosensors.

6.2 Future Directions

Nonimmunoglobulin proteins and peptides evolved to bind specific protein targets have the potential to significantly advance therapeutics, diagnostics, and basic research. However, the efficiency of established combinatorial screening techniques used to identify functional binders is being challenged by the desire to isolate ligands to protein targets that are difficult to recombinantly express, like mammalian proteins, as well as the use of unique scaffold templates, like cyclic peptides. Here, we have demonstrated baseline platforms that to begin to address these limitations. However, there are multiple ways that our described strategies can be evolved to further advance the functionality of combinatorial platforms.

In particular, several details of utilizing magnetic yeast targets for combinatorial selections, as described in **Chapters 2** and **4**, need to be evaluated further. For instance, we have not fully explored how selection conditions influence the isolation of binders with certain affinities when performing magnetic yeast panning. Further investigation is needed into what is the optimal number of magnetic target cells to utilize in each screening round. Also, the number of library cells or mRNA-peptide fusions incubated should be evaluated further. When considering yeast-yeast

screening, the qYY2H data suggests that the recovery of target bound-library cells depends on the affinity of the monovalent interaction between the displayed surface proteins. Thus, when considering yeast-yeast screening, it is likely that the concentration of library cells, and more specifically individual library mutants, can affect enrichment based on the affinity of the expressed mutants. Accordingly, a larger number of library cells may be required to isolate weaker affinity mutants if desired. In contrast, the interactions between magnetic yeast targets and mRNA-peptide fusions are monovalent in nature. Thus, the best selection conditions for yeast-yeast screens may differ from yeast, mRNA-peptide fusion screens. The wash conditions utilized can also significantly influence the affinity of the isolated clones. Wash conditions to explore include: the length of the wash step, presence of competing protein, presence of detergents, buffer pH, buffer salt concentration, and the of use vortexing. It is likely that some of these wash conditions will be better than others at diminishing the interactions of non-specific and weak-affinity binders. Lastly, we have not fully explored how many rounds of screening are required to achieve significant enrichment when screening combinatorial libraries utilizing magnetic yeast targets. Each of these features could be explored using a mixed population of known binders with varying affinity for a target protein. For each selection condition, the enrichment of each binder by the magnetic yeast can be determined via Next Generation Sequencing to quantify the effects of each condition tested.

A natural extension of the SsoFe2 protein is to use SsoFe2 to magnetize mammalian cells. While yeast are eukaryotic, yeast lack some post-translational machinery of higher order eukaryotes[1,2]. Thus, some proteins that require certain post-translational modifications may best expressed by a mammalian host. To generate magnetized mammalian cells, SsoFe2 can be expressed as a surface fusion to HEK293 cells that also express the target protein on their surface. When considering cell magnetization in general, other strategies could be explored to generate

magnetic target cells including the expression of peptide sequences with affinity for iron oxide[3] or the production of intracellular magnetic particles through iron sequestration[4–6].

The mathematical model developed in **Chapter 3** using the qYY2H platform has the potential to guide the selection conditions used when screening yeast display libraries against magnetic yeast targets. However, several features of the qYY2H platform need to be evaluated further. For example, the current mathematical model was developed only using 5 bait-prey pairs. Additional bait-prey pairs should be evaluated to refine the model. When developing the model, we only used two different bait, TOM22 and hFc (Fc portion of human IgG); To eliminate any bias, the model should be expanded using other bait proteins. Additionally, binders evolved from scaffold proteins, other than Sso7d and nanobody, should be considered. The model should also be expanded to include lower and higher affinity binders than those considered. We only evaluated bait-prey interactions ranging from ~400 – 5000 nM. Moreover, the detection limit of the qYY2H assay has not been determined. We have only tested bait-prey pairs with affinities as low as 25 μM. The detection limit could be evaluated using binders to the WW domains of YAP. WW-specific peptides bind with a wide range of affinities, as low as 320 μM[7,8]. It is also possible that once a certain affinity is reached, we may not be able to distinguish between high affinity binders due to the limitations of our assay (i.e.: the number of cells we can reasonably incubate). Ultimately, if the qYY2H model is better defined, we may be able to use the developed mathematical framework to predict the conditions (i.e.: how many magnetic target and library cells to incubate, incubation volume, etc) needed to isolate specific affinity binders from yeast display combinatorial libraries when performing magnetic yeast selections.

A major advantage of the qYY2H platform in comparison to other yeast two hybrid assays is that the qYY2H system can be used to efficiently study protein-protein interactions that rely on

a post-translationally modified bait. However, limitations exist for the strategy we utilized to express enzymatically modified bait proteins on the surface of magnetic yeast cells. When considering a phosphorylation modification, we observed that ~50% of the cells expressing the bait protein actually expressed a phosphorylated version of the bait. Thus, the bait cell population is heterogenous, expressing both modified and unmodified bait protein, which influences the interaction between the bait and prey cell populations. We have attempted to account for this limitation by scaling the number of prey cells captured by the percentage of bait cells actually expressing a modified version of the bait as it is likely that a greater number of prey cells would be captured if all bait cells expressed a modified version of the bait. If the majority of the bait cells expressed a modified version of the bait, we would not have to perform this scaling, and we would be able to more confidently quantify bait-prey cell interactions that rely on a post-translational modification.

The number of bait cells expressing a modified version of the bait protein can be improved by increasing the amount of modifying enzyme present in the endoplasmic reticulum. In our system, the bait protein is under the control of the Gal1 promoter while the enzyme that affords the modification is under the control of the Gal10 promoter. Protein expression under the Gal1 promoter has been shown to be up to five times greater than when expression is coordinated by the Gal10 promoter[9,10]. To gain greater control over enzyme expression, the expression of the bait protein and the modifying enzyme should be uncoupled and rely on different induction additives. In particular, the enzyme should likely be placed under the control of a promoter that is stronger than the promoter used for bait expression to ensure an excess of enzyme. Other promoters to explore include a DDI2[11] or β-estradiol promoter[12]. Additionally, stronger endoplasmic reticulum

retention sequences could be explored to increase the residence time of enzyme and bait protein within the endoplasmic reticulum prior to surface display of the bait protein[13].

Our yeast display platform for cyclic peptide identification, as described in **Chapter 5**, can be manipulated to utilize other crosslinking strategies allowing for the platform to potentially be used to identify multi-ringed peptides. For example, we could use two individual crosslinkers with different reactive chemistries to generate bicyclic peptides. Specifically, we could explore using a thiol-reactive 1,3-dibromopropane (DBP) linker to stably crosslink cysteine residues in conjunction with a disuccinimidyl gultarate (DSG) linker that crosslinks amine groups[14–16]. A linear peptide sequence, like MXXXXXK-GSGSGS-CXXXXXC, can be expressed as a yeast surface fusion. After, yeast cells displaying the described linear peptide sequence can be incubated with DSG followed by DBP. The DSG crosslinker will cyclize the linear peptide sequence between its N-terminus and the lysine residue while the DBP crosslinker will cyclize the sequence between the included cysteine residues. This will result in the display of a bicyclic peptide on the yeast surface. Optimization will be required to ensure complete modification by both crosslinkers. Other crosslinkers can be explored, include thiol reactive tris-(bromomethyl)benzene[17] as well as bis-electrophile linkers, like di(bromomethyl)pyridine, that can crosslink amine and thiol groups[18].

The described yeast platform for cyclic peptide display could also be extended to identify stimuli-responsive cyclic peptides if stimuli-responsive crosslinkers with appropriate reactivity chemistry (thiol or amine) are incorporated. The use of a stimuli-responsive linker allows the design of peptides that can reversibly bind target proteins. For example, peptides can be selected that bind a target protein when the crosslinker is in a particular conformation and lose binding after a stimulus is applied that changes the conformation of the crosslinker. A variety of materials exist that incorporate particular amino acid sequences or chemical linkers that respond to different

environmental triggers, including pH[19,20], light[21–24], and temperature[25,26]. Specifically, peptides incorporating crosslinkers sensitive to light and temperature could be used for the purification of biologics that require gentle processing conditions to maintain viability and functionality.

6.3 References

(1) Wildt, S.; Gerngross, T. U. The Humanization of N-Glycosylation Pathways in Yeast. *Nat. Rev. Microbiol.* **2005**, *3* (2), 119–128.

(2) Gemmill, T. R.; Trimble, R. B. Overview of N- and O-Linked Oligosaccharide Structures Found in Various Yeast Species. *Biochim. Biophys. Acta - Gen. Subj.* **1999**, *1426* (2), 227–237.

(3) Schwaminger, S. P.; Blank-Shim, S. A.; Scheifele, I.; Pipich, V.; Fraga-García, P.; Berensmeier, S. Design of Interactions Between Nanomaterials and Proteins: A Highly Affine Peptide Tag to Bare Iron Oxide Nanoparticles for Magnetic Protein Separation. *Biotechnol. J.* **2019**, *14* (3), 1800055.

(4) Zurkiya, O.; Chan, A. W. S.; Hu, X. MagA Is Sufficient for Producing Magnetic Nanoparticles in Mammalian Cells, Making It an MRI Reporter. *Magn. Reson. Med.* **2008**, *59* (6), 1225–1231.

(5) Nishida, K.; Silver, P. A. Induction of Biogenic Magnetization and Redox Control by a Component of the Target of Rapamycin Complex 1 Signaling Pathway. *PLoS Biol.* **2012**, *10* (2), e1001269.

(6) Liu, X.; Lopez, P. A.; Giessen, T. W.; Giles, M.; Way, J. C.; Silver, P. A. Engineering Genetically-Encoded Mineralization and Magnetism via Directed Evolution. *Sci. Rep.* **2016**, *6* (1), 1–10.

(7) Iglesias-Bexiga, M.; Castillo, F.; Cobos, E. S.; Oka, T.; Sudol, M.; Luque, I. WW Domains of the Yes-Kinase-Associated-Protein (YAP) Transcriptional Regulator Behave as Independent Units with Different Binding Preferences for PPxY Motif-Containing Ligands. *PLoS One* **2015**, *10* (1), e0113828.

(8) McDonald, C. B.; McIntosh, S. K. N.; Mikles, D. C.; Bhat, V.; Deegan, B. J.; Seldeen, K. L.; Saeed, A. M.; Buffa, L.; Sudol, M.; Nawaz, Z.; et al. Biophysical Analysis of Binding of WW Domains of the YAP2 Transcriptional Regulator to PPXY Motifs within WBP1 and WBP2 Adaptors. *Biochemistry* **2011**, *50* (44), 9616–9627.

(9) Yocum, R. R.; Hanley, S.; West, R.; Ptashne, M.; Ptashne, M. Use of LacZ Fusions to Delimit Regulatory Elements of the Inducible Divergent GAL1-GAL10 Promoter in Saccharomyces Cerevisiae. *Mol. Cell. Biol.* **1984**, *4* (10), 1985–1998.

(10) Da Silva, N. A.; Bailey, J. E. Influence of Plasmid Origin and Promoter Strength in Fermentations of Recombinant Yeast. *Biotechnol. Bioeng.* **1991**, *37* (4), 318–324.

(11) Wang, Y.; Zhang, K.; Li, H.; Xu, X.; Xue, H.; Wang, P.; Fu, Y. V. Fine-Tuning the Expression of Target Genes Using a DDI2 Promoter Gene Switch in Budding Yeast. *Sci. Rep.* **2019**, *9* (1).

(12) Denard, C. A.; Paresi, C.; Yaghi, R.; McGinnis, N.; Bennett, Z.; Yi, L.; Georgiou, G.; Iverson, B. L. YESS 2.0, a Tunable Platform for Enzyme Evolution, Yields Highly Active TEV Protease Variants. *ACS Synth. Biol.* **2021**, *10* (1), 63–71.

(13) Mei, M.; Zhai, C.; Li, X.; Zhou, Y.; Peng, W.; Ma, L.; Wang, Q.; Iverson, B. L.; Zhang, G.;

Yi, L. Characterization of Aromatic Residue-Controlled Protein Retention in the Endoplasmic Reticulum of Saccharomyces Cerevisiae. *J. Biol. Chem.* **2017**, *292* (50), 20707–20719.

(14) Byrne, M. P.; Stites, W. E. Chemically Crosslinked Protein Dimers: Stability and Denaturation Effects. *Protein Sci.* **1995**, *4* (12), 2545–2558.

(15) Witkowski, A.; Joshi, A. K.; Rangan, V. S.; Falick, A. M.; Witkowska, H. E.; Smith, S. Dibromopropanone Cross-Linking of the Phosphopantetheine and Active-Site Cysteine Thiols of the Animal Fatty Acid Synthase Can Occur Both Inter- and Intrasubunit. Reevaluation of the Side-by-Side, Antiparallel Subunit Model. *J. Biol. Chem.* **1999**, *274* (17), 11557–11563.

(16) Millward, S. W.; Takahashi, T. T.; Roberts, R. W. A General Route for Post-Translational Cyclization of MRNA Display Libraries. *J. Am. Chem. Soc.* **2005**, *127* (41), 14142–14143.

(17) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-Encoded Combinatorial Chemical Libraries Based on Bicyclic Peptides. *Nat. Chem. Biol.* **2009**, *5* (7), 502–507.

(18) Kale, S. S.; Bergeron-Brlek, M.; Wu, Y.; Kumar, M. G.; Pham, M. V.; Bortoli, J.; Vesin, J.; Kong, X. D.; Franco Machado, J.; Deyle, K.; et al. Thiol-to-Amine Cyclization Reaction Enables Screening of Large Libraries of Macrocyclic Compounds and the Generation of Sub-Kilodalton Ligands. *Sci. Adv.* **2019**, *5* (8), eaaw2851.

(19) Matsumoto, K.; Kawamura, A.; Miyata, T. Conformationally Regulated Molecular Binding and Release of Molecularly Imprinted Polypeptide Hydrogels That Undergo Helix-Coil Transition. *Macromolecules* **2017**, *50* (5), 2136–2144.

(20) Raza, F.; Zhu, Y.; Chen, L.; You, X.; Zhang, J.; Khan, A.; Khan, M. W.; Hasnat, M.; Zafar, H.; Wu, J.; et al. Paclitaxel-Loaded PH Responsive Hydrogel Based on Self-Assembled Peptides for Tumor Targeting. *Biomater. Sci.* **2019**, *7* (5), 2023–2036.

(21) Day, K.; Schneible, J. D.; Young, A. T.; Pozdin, V. A.; Van Den Driessche, G.; Gaffney, L. A.; Prodromou, R.; Freytes, D. O.; Fourches, D.; Daniele, M.; et al. Photoinduced Reconfiguration to Control the Protein-Binding Affinity of Azobenzene-Cyclized Peptides. *J. Mater. Chem. B* **2020**, *8* (33), 7413–7427.

(22) Beharry, A. a.; Sadovski, O.; Woolley, G. A. Azobenzene Photoswitching without Ultraviolet Light. *J. Am. Chem. Soc.* **2011**, *133* (49), 19684–19687.

(23) Samanta, S.; Beharry, A. A.; Sadovski, O.; McCormick, T. M.; Babalhavaeji, A.; Tropepe, V.; Woolley, G. A. Photoswitching Azo Compounds in Vivo with Red Light. *J. Am. Chem. Soc.* **2013**, *135* (26), 9777–9784.

(24) Wiedbrauk, S.; Dube, H. Hemithioindigo—an Emerging Photoswitch. *Tetrahedron Lett.* **2015**, *56* (29), 4266–4274.

(25) Bian, S.; Cai, H.; Cui, Y.; He, M.; Cao, W.; Chen, X.; Sun, Y.; Liang, J.; Fan, Y.; Zhang, X. Temperature and Ion Dual Responsive Biphenyl-Dipeptide Supramolecular Hydrogels as Extracellular Matrix Mimic-Scaffolds for Cell Culture Applications. *J. Mater. Chem. B* **2017**, *5* (20), 3667–3674.

(26) Hassouneh, W.; Christensen, T.; Chilkoti, A. Elastin-like Polypeptides as a Purification Tag for Recombinant Proteins. *Curr. Protoc. Protein Sci.* **2010**, *Chapter 6* (SUPPL. 61).

APPENDIX

Appendix A: Supplementary data for chapter 3

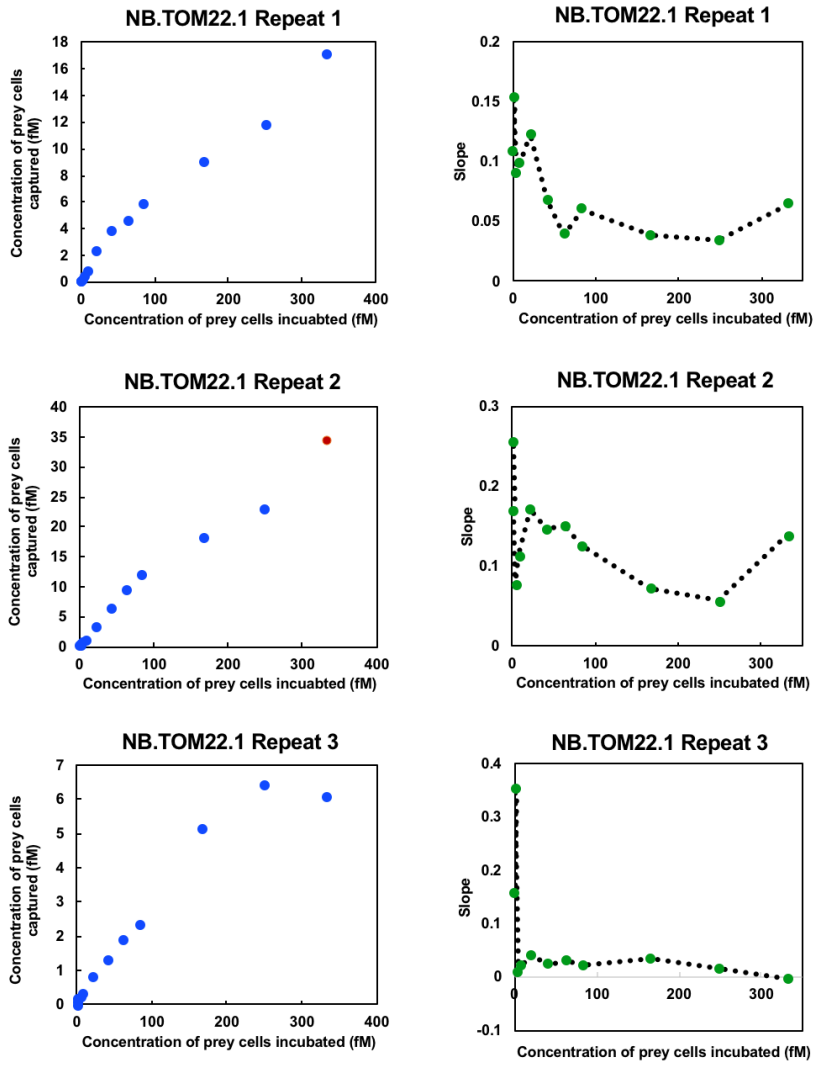


Figure A.1. qYY2H titration curve for NB.TOM22.1. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, TOM22. The number of magnetic bait cells displaying TOM22 was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

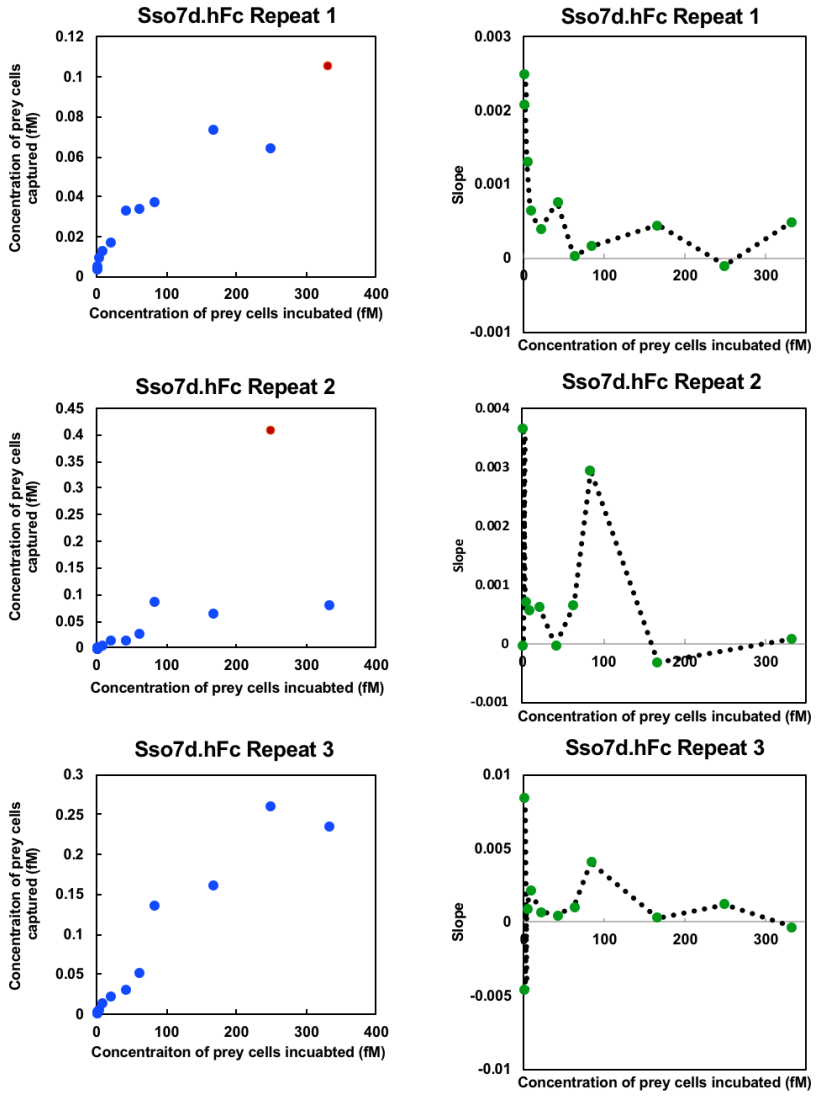


Figure A.2. qYY2H titration curve for Sso7d.hFc. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, hFc. The number of magnetic bait cells displaying hFc was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

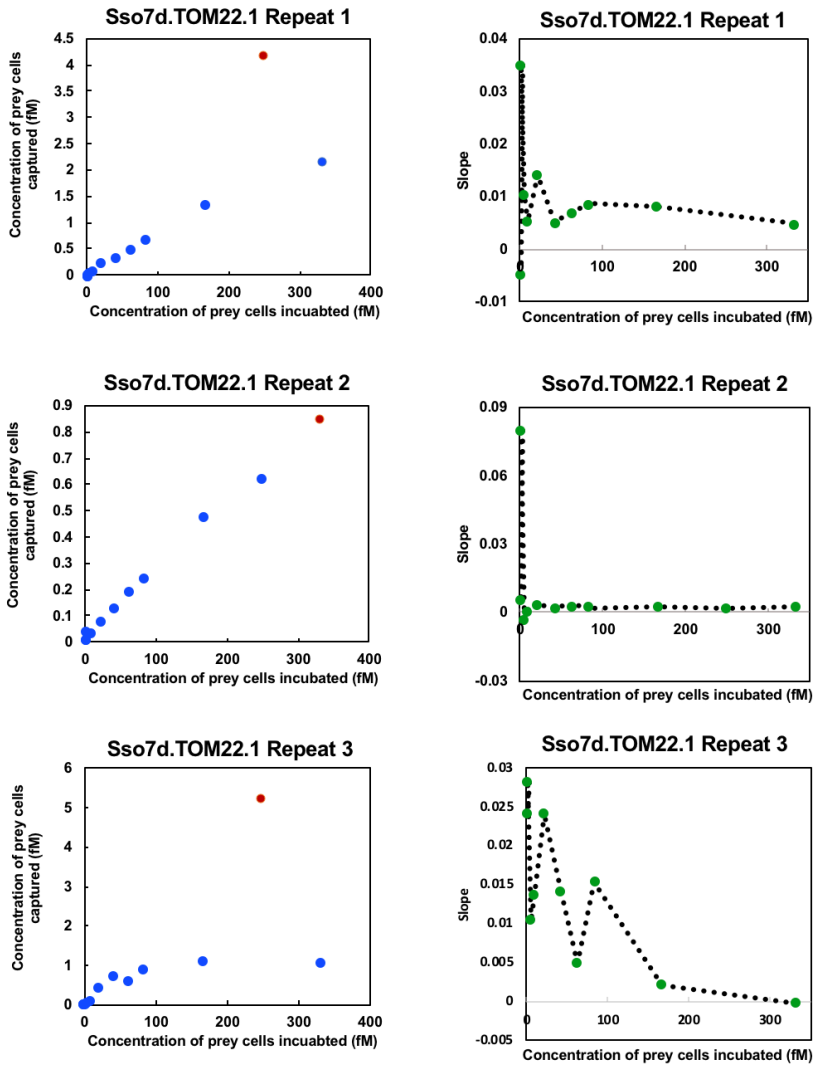


Figure A.3. qYY2H titration curve for Sso7d.TOM22.1. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, TOM22. The number of magnetic bait cells displaying TOM22 was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

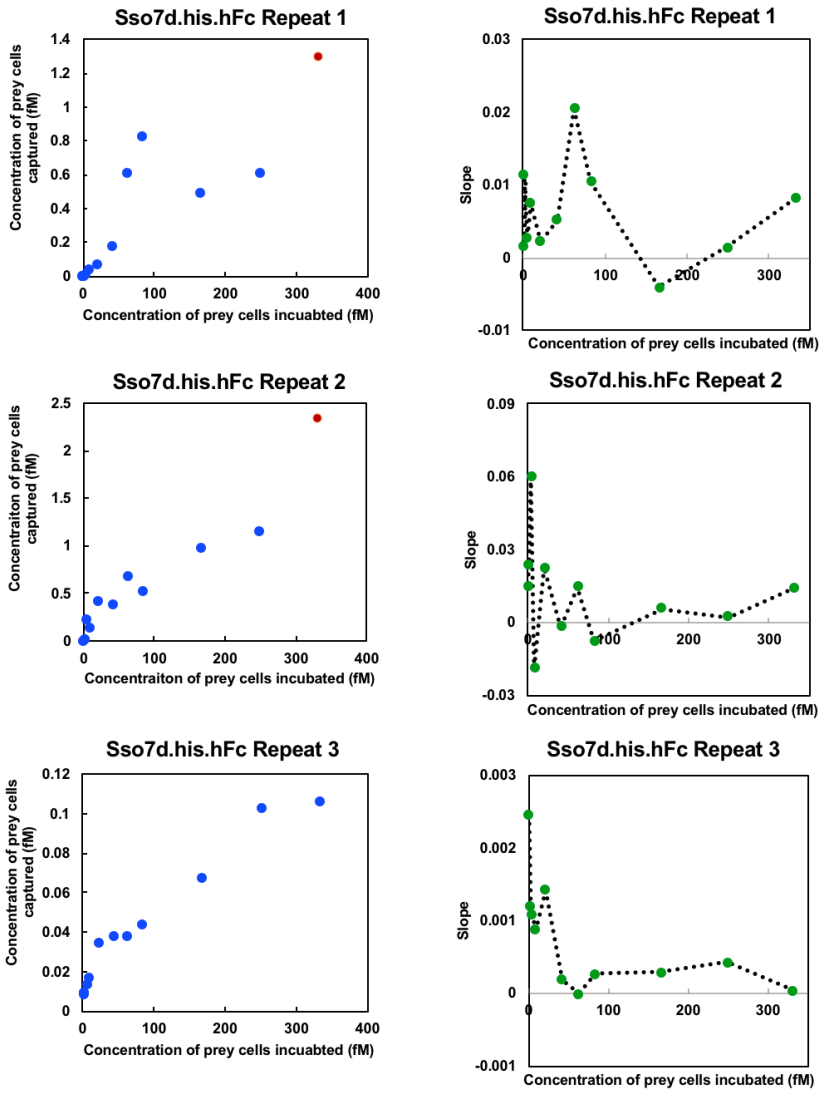


Figure A.4. qYY2H titration curve for Sso7d.his.hFc. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, hFc. The number of magnetic bait cells displaying hFc was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

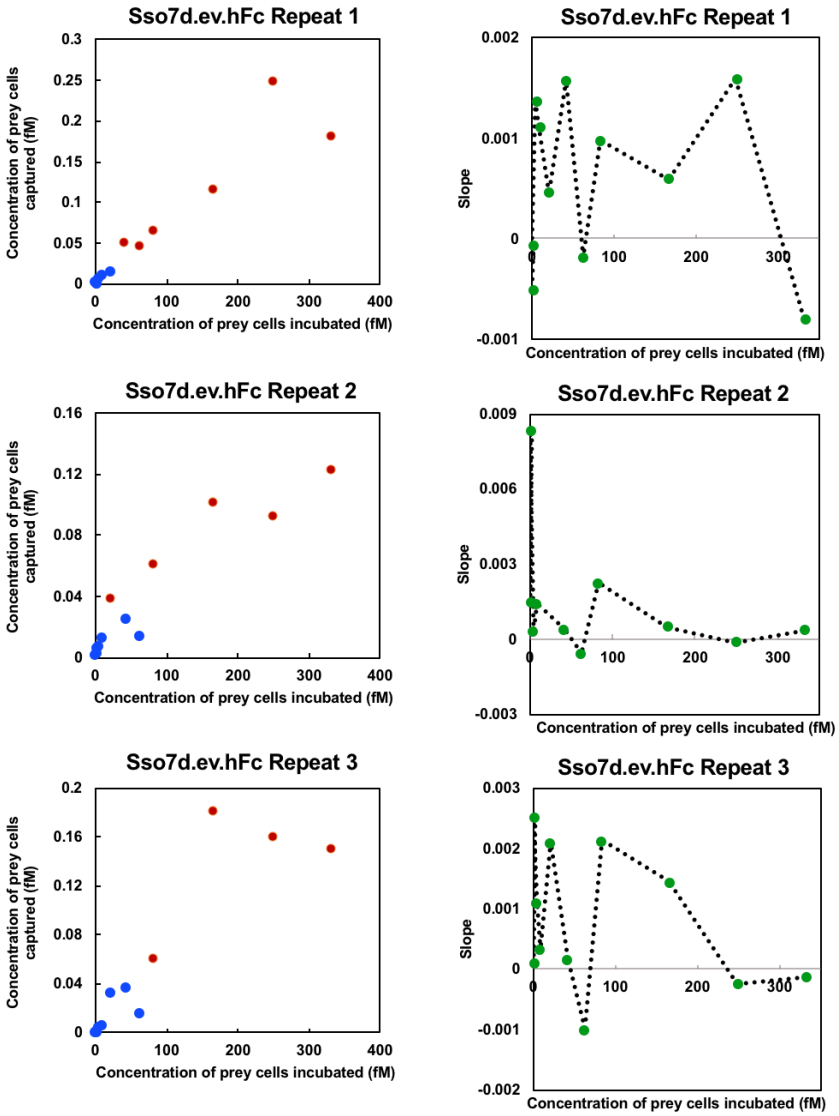


Figure A.5. qYY2H titration curve for Sso7d.ev.hFc. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, hFc. The number of magnetic bait cells displaying hFc was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

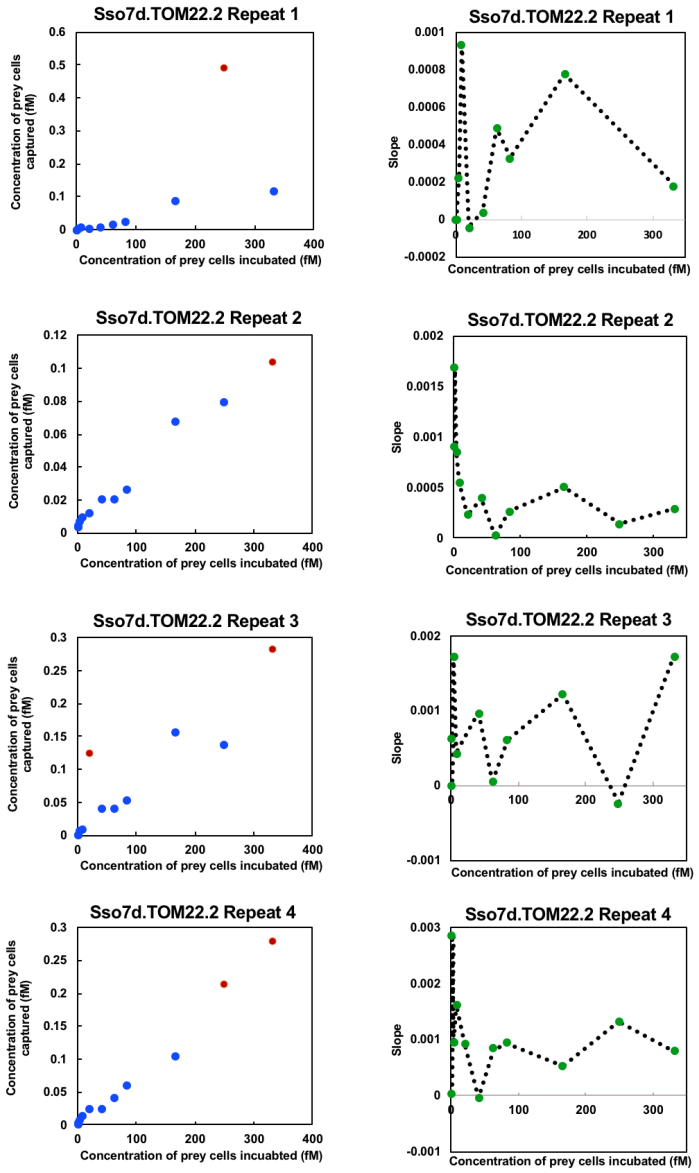


Figure A.6. qYY2H titration curve for Sso7d.TOM22.2. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, TOM22. The number of magnetic bait cells displaying TOM22 was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

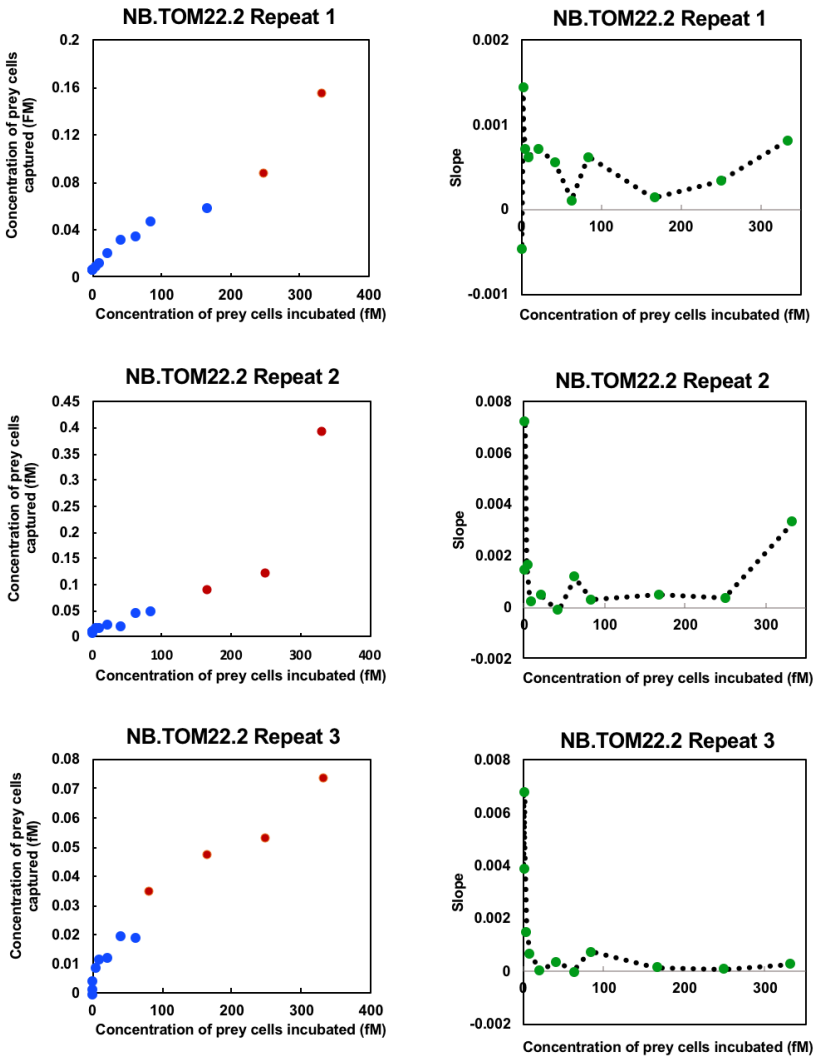


Figure A.7. qYY2H titration curve for NB.TOM22.2. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, TOM22. The number of magnetic bait cells displaying TOM22 was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

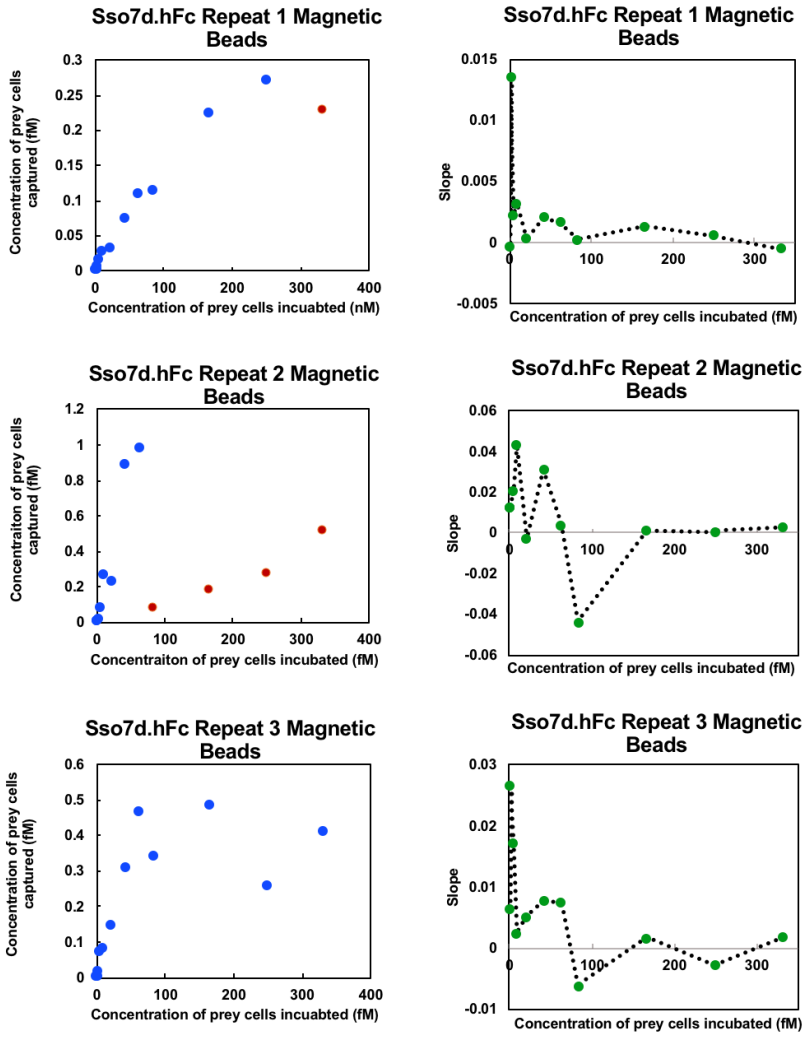


Figure A.8. Titration curve for Sso7d.hFc utilizing a bait immobilized onto magnetic beads. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic beads functionalized with the bait, hFc. The number of magnetic beads functionalized with hFc was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

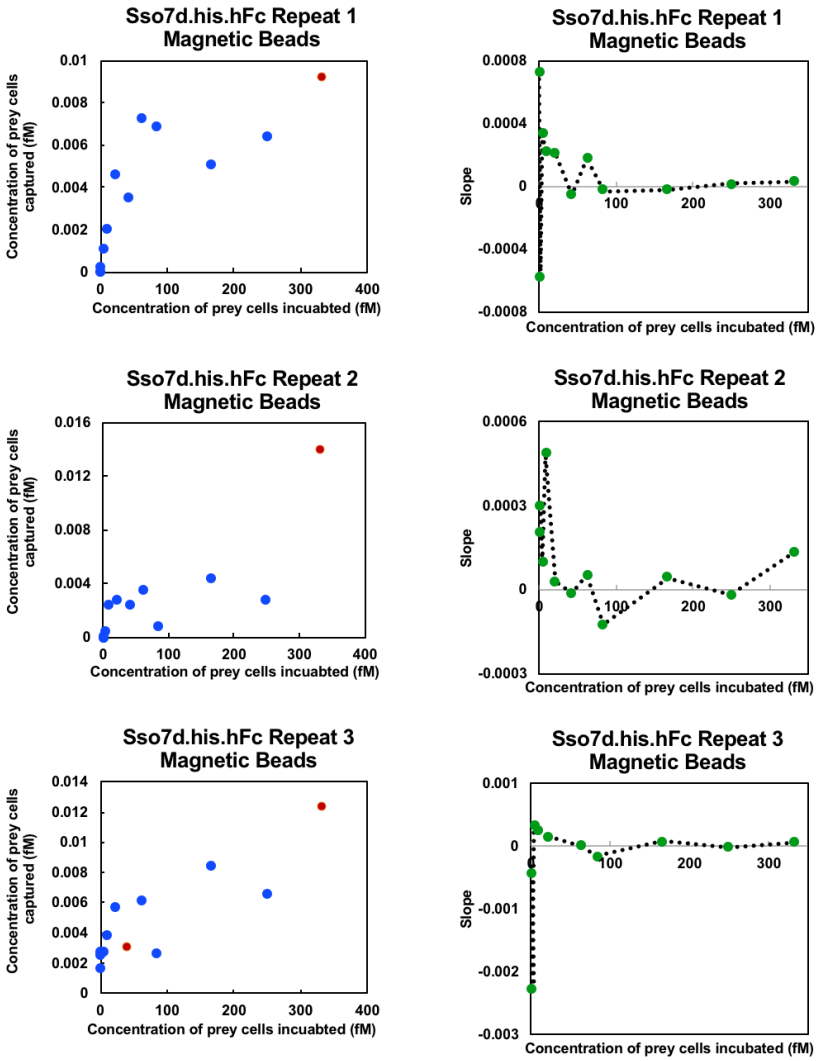


Figure A.9. Titration curve for Sso7d.his.hFc utilizing a bait immobilized onto magnetic beads. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic beads functionalized with the bait, hFc. The number of magnetic beads functionalized with hFc was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

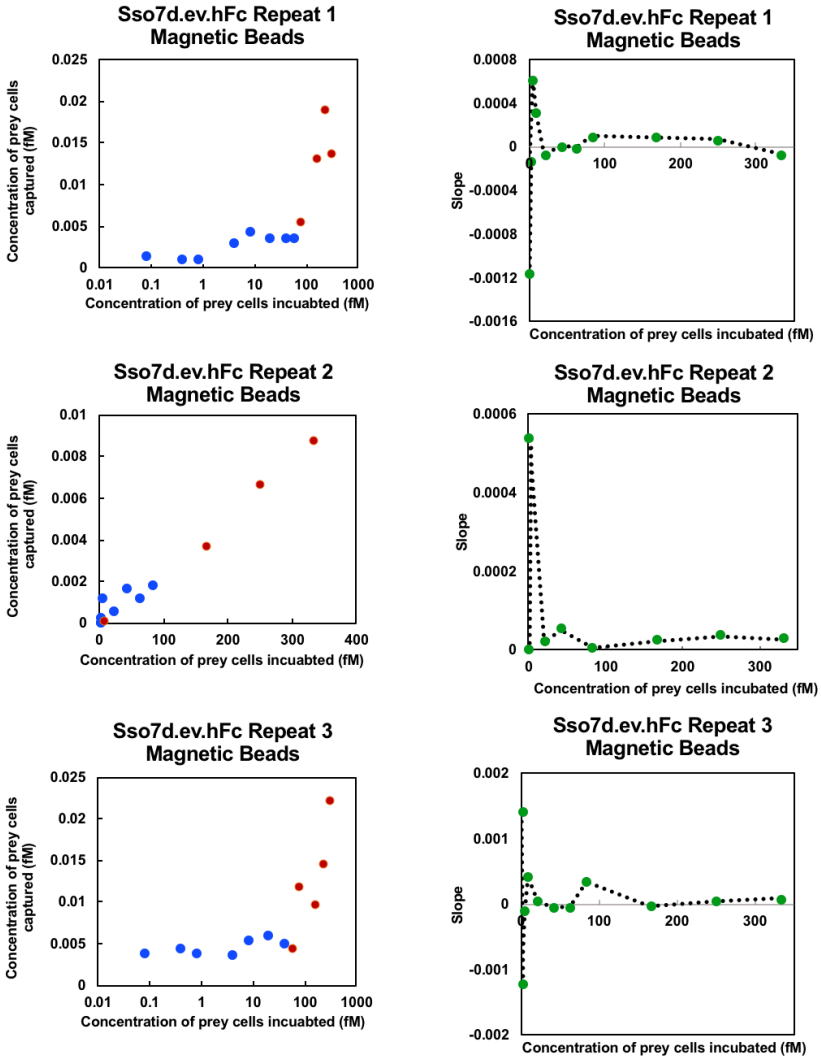


Figure A.10. Titration curve for Sso7d.ev.hFc utilizing a bait immobilized onto magnetic beads. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic beads functionalized with the bait, hFc. The number of magnetic beads functionalized with hFc was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

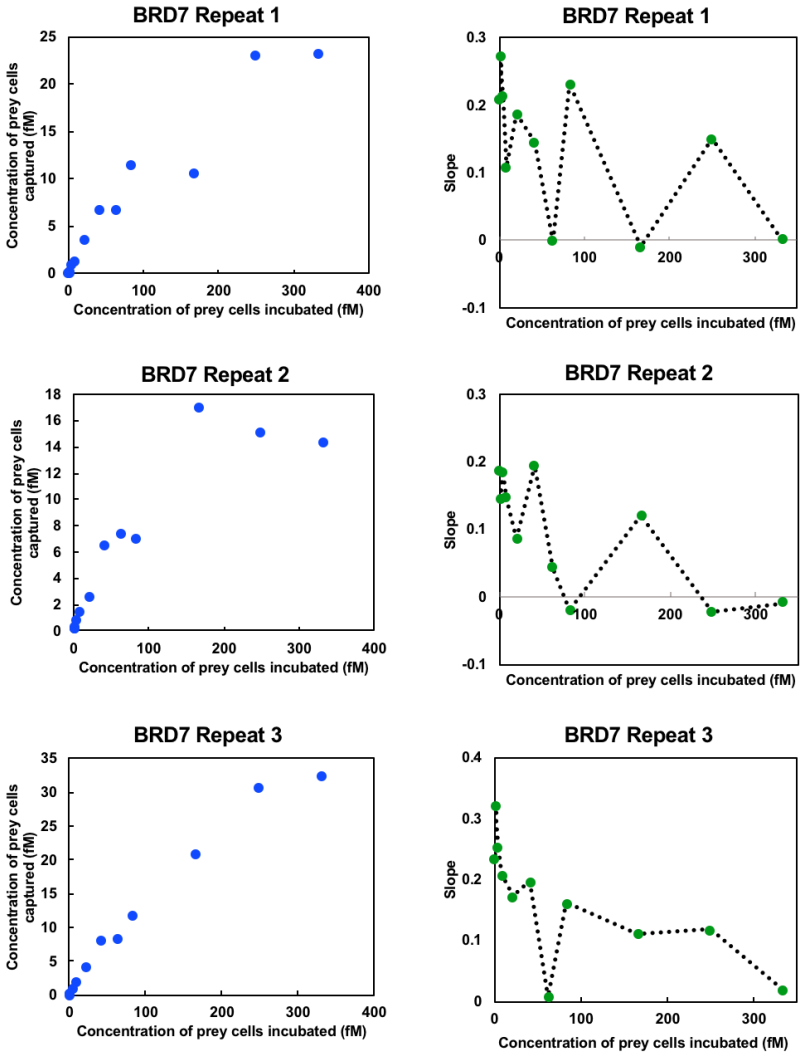


Figure A.11. qYY2H titration curve for BRD7. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, SMAD3. The number of magnetic bait cells displaying SMAD3 was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

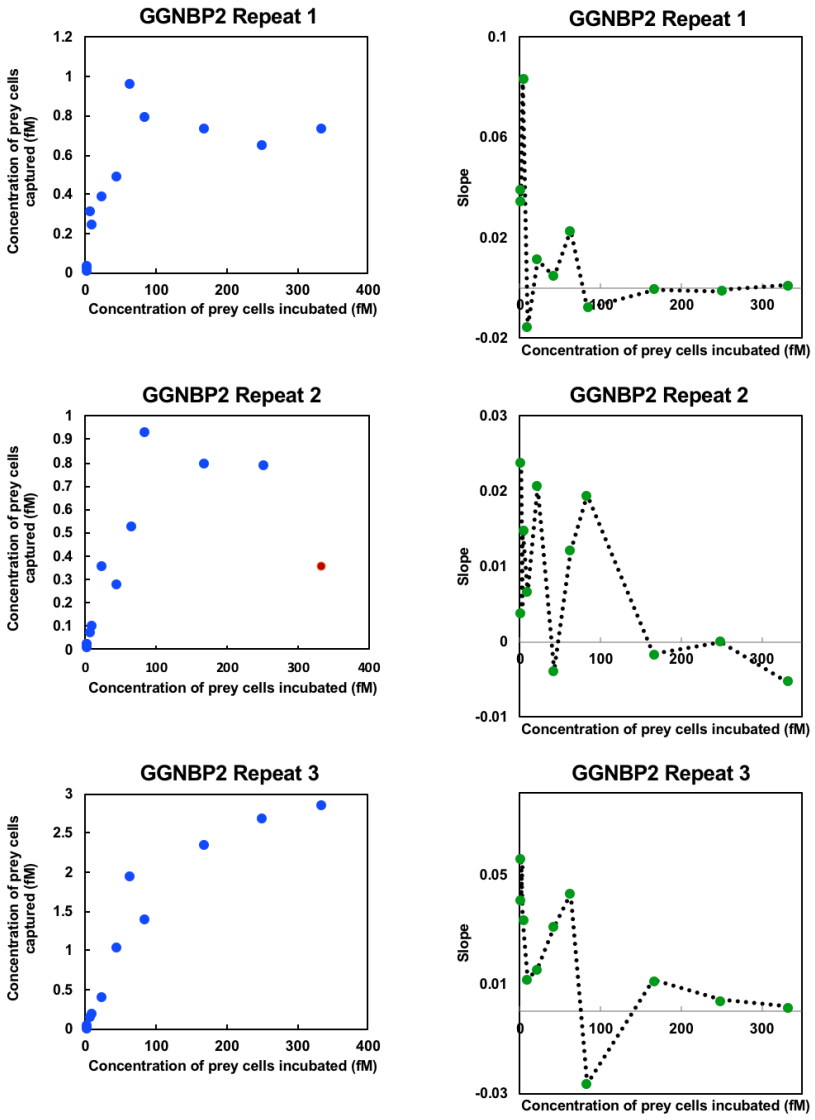


Figure A.12. qYY2H titration curve for GGNBP2. The left plot compares the concentration of prey cells captured as a function of the concentration of prey cells incubated with magnetic yeast cells expressing the bait, SMAD3. The number of magnetic bait cells displaying SMAD3 was held constant. A plot of the slope between adjacent titration curve data points was also generated (right). The slope plot was used to determine which points to include when applying the monovalent binding isotherm to the titration curve data. Typically, points were ignored after saturation was met (slope ~0). Non-specific binding of the prey cells can occur after saturation is met. For the titration curve plots, blue points were included when applying the monovalent binding isotherm while red points were discarded.

Table A.1. Putative interaction partners for SMAD3 isolated from a yeast-displayed cDNA library screened against magnetic yeast cells displaying SMAD3. Prey cells displaying the protein encoded by the uncut vector used for library construction were isolated due to nonspecific binding.

Protein	*Abbreviation*	*Occurrence*	*Protein Subunit Amino Acid Sequence*
NIPBL cohesin loading factor	NIPBL	1	ETPKQKSEGRPETPKQKGDGRPETPKQKSE GRPETPKQKGEGRPETPKHRHENRRDSGKP STEKKPDVSKHKQDIKSDSPRLKSERAEALK QRPDGRWESLRRDHDSKQKSDDRGESERH RGDQSRVRRPETLRSSSRNDHSTKSDAGKT EK
Bromodomain containing protein 7	BRD7	1	GGPVGHGQEAQEAQVGPPLLRGVRGEALK LVLKVGGSEVTELSTGSSGHDSSLFEDRSDH DKHKDRKRKKRKKGEKQAPGEEKGRKRRR VKEDKKKRDRDRAENEVDRDLQCHVPIRL DLPPEKPLTSSLAKQEEVEQTPLQEALNQL MRQLQRKDPSAFFSFPVTDFIAPGYSMIIKH PMDFSTMKEKIKNNDYQSIEELKDNFKLMC TNAMIYNKPETIYYKAAKKLLHSGMKILSQ ERIQSLKQSIDFMSDLQKTRKQKERTDACQS GEDSGCWQREREDSGDAGKTEKK
Gametogenetin binding protein 2	GGNBP2	1	EQLCEEFSEEERVRELKQEKKRQKRKNRK
Endothelin 1	EDN1	1	NTPERVVPYGLGGSSRSKRSLKDLLPNKAT DQAVRCQCAHQKDKKCWKNRK
Mitochondrial calcium uniporter	MCU	1	RHHVLSQLHHPRCWKNRKKD
Dead box polypeptide 18	DDX18	1	RHKPVPTSKWRCWKNRK
Pre-mRNA processing factor 40A	PRPF40A	1	AEWHYGRGKKKSKKRRHKSDSPESDAGKT EK
TATA-Box binding protein associated factor 2	TAF2	1	KKKKHKHKHKHKRKHDSKDKDREPFAFSS PASGRPIQMLEKQKK
Nucleotide binding protein 1	NUBP1	1	RVGRPRVQTPGQWGTGGTIATRCWKNRK
Uncut Vector		21	

Table A.2. Putative interaction partners for the WW domains of YAP isolated from a yeast-displayed cDNA library screened against magnetic yeast displaying the WW domains of YAP. Some of the isolated prey contain previously identified binding motifs for the WW domain. Prey cells displaying the protein encoded by the uncut vector used for library construction were isolated due to nonspecific binding.

Protein	Abbreviation	Occurrence	Binding Motif	Protein Subunit Amino Acid Sequence
Tata-Box binding protein associated factor 2	TRF2	2	RP	HKHKHKHKHKHDSKDKDR EPFAFSSPASGRPIQMLEKQ KK
Cyclin L2	CCNL2	1	PPR	AEGHYGRGGGSKSQSRSRS RSDSPPRQVHRGAPYKGSE VRGSRKSKDCKYLTQKPHK SRSRSSSRSRSRSRERTDNA GKTEK
Nuclear speckle splicing regulatory protein 1	NSRP1	1	RPP, RP	AEWPLRPERSSSEERGLGT KHHSRGSQSRGHEHQDRQS RDQESCHKDRSHREEKSSH RHREASHKDHYWKKHEQE DKLKGREQEERQDREGKRE KYSSREQERDRQRNDHDR YSEKEKKRKEKEEHTKARR ERCEDSGKHREREKPEGHG QSSERHRDRRESSPRSRPKD DLDQERSSKARNTEKDKGE QGKPSHSETSLATKHRLAE ERPEKGSEQERPPEAVSKFA KRSNEETVMSARDRYLAR QMARINAKTYIEKEDD
Ribosome protein L3	RPL3	1	None	VFAEHISDECKRRFYKNWH KSKKKAFTKYCKKWQDDT GKKQLEKDFNSMKKYCQV IRIIAHTQMLEKQKK
Chromatin target of PRMT1	CHTOP	1	None	SHHIYKTTLRKHEKKSDRC WKNRK
Chromosome 12 RP23-33124		1	RP	RKCGRPAHSWKNRK
Chromosome 7, RP24-233E11		1	RP	EKKKKKKKPAHTHKHTITH ISVYSNSATLLTRIN
Guanine nucleotide binding protein-like 2	GNL2	1	None	KEEKKTSAEDSDAAPTKKA RKWDAQMEEEPSNKTQRM LTCKERGRAARQQQSKKV GVRYYETHNLKNRNRNKK KTSDSEGQKHRRNKFRQKQ

Table A.2 (Continued).

ADP-ribosylation factor like GTPAse 9	*ARL9*	*2*	*None*	KQGEGEREGEGRGEKEVC GTEKQEGERDRERRREREG GRSENQPRERERREEKEEC GPGNQREWERQGEGGREK EECGPKNQGEWERQGEGG REKEECGPETKENGKDKGK EEEKKKSVDLKTKENGKD KGKEEGKKKSVDLKTKEN EKDKGKEEEKKKSVDLKT KENGKDKGKEEEKKKSVD PKTKENEKDKGKEEEKKKS VDLKTKENGKDKGKEEEK KKSVDLKTKENGKDKEKE EKKRKSADLKTKETGKEKG RREKK
Transcription elongation regulator 1	TCERG1	1	PXY	REKKNKIMQAKEDFKKMM EEAKFNPRATFSEFAAKRA KDSRFKAIEKMKDREALFN EFVAAARKKEKEDSKTRGE KIKSDFFELLSNHHLDSQSR WSKVKDKVESDPRYKAVD SSSMREDLFKQYIEKIAKNL DSEKEKELERQARIEASLRE RREREVQKARSEQTKEIDRE REQHKREEAIQNFKALLSD MVRSSDVSWSDTRRTLRK DHRWESGSLLEREEKEKLF NEHIEALTKKKKKKAASAS RGWASIRDPSSSSCR
Undigested vector		14		
Small cDNA with no similarity		4		