Development and characterization of affinity peptides using

mRNA display and dot blot method

by

Pankti Shah

A Thesis Presented in Partial Fulfillment
Of the Requirements for the Degree
Master of Science

Approved July 2014 by the
Graduate Supervisory Committee:

John Chaput, Chair
Sidney Hecht
Rebekka Wachter

ARIZONA STATE UNIVERSITY

August 2014

ABSTRACT

Protein affinity reagents have aptly gained profound importance as capture reagents and drugs in basic research, biotechnology, diagnostics and therapeutics. However, due to the cost, labor and time associated with production of antibodies focus has recently changed towards potential of peptides to act as protein affinity reagents. Affinity peptides are easy to work with, non-immunogenic, cost effective and amenable to scale up. Even though researchers have developed several affinity peptides, we are far from compiling library of peptides that encompasses entire human proteome. My thesis describes high throughput pipeline that can be used to develop and characterize affinity peptides that bind several discrete sites on target proteins.

Chapter 2 describes optimization of cell-free protein expression using commercially available translation systems and well-known leader sequences. Presence of internal ribosome entry site upstream of coding region allows maximal expression in HeLa cell lysate whereas translation enhancing elements are best suited for expression in rabbit reticulocyte lysate and wheat germ extract. Use of optimal vector and cell lysate combination ensures maximum protein expression of DNA libraries.

Chapter 3 describes mRNA display selection methodology for developing affinity peptides for target proteins using large diversity DNA libraries. I demonstrate that mild denaturant is not sufficient to increase selection pressure for up to three rounds of selection and increasing number of selection rounds increases probability of finding affinity peptides. These studies enhance fundamental understanding of mRNA display and pave the way for future optimizations to accelerate convergence of *in vitro* selections.

Chapter 4 describes a high throughput double membrane dot blot system to rapidly screen, identify and characterize affinity peptides obtained from selection output. I used dot blot to screen potential affinity peptides from large diversity of previously

uncharacterized mRNA display selection output. Further characterization of potential peptides allowed determination of several high affinity peptides from having Kd range 150–450 nM. Double membrane dot blot is automation amenable, easy and affordable solution for analyzing selection output and characterizing peptides without need for much instrumentation.

Together these projects serve as guideline for evolution of cost effective high throughput pipeline for identification and characterization of affinity peptides.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

List of Figures

Chapter 1

Introduction

Discovery of antibodies, nature's own protein affinity reagents, not only marked the beginning of field of immunology but also added whole new dimension to our understanding of protein-protein interactions. Realizing diagnostic and therapeutic applications of antibodies, researchers soon embarked upon the quest for cheaper, synthetic and simple alternatives to traditional antibodies. Understanding the structure of antibody has given rise to various protein affinity reagent platforms like monoclonal antibodies, recombinant antibodies, single chain antibodies, nucleic acid aptamers, peptides and synbodies. Hybridoma, recombinant DNA and systematic evolution of ligands by exponential enrichment (SELEX) technologies have been widely used to develop various types of affinity reagents. Each of these technologies aim to produce large diversity of affinity reagents by using permutation and combination of amino acids, a principle very similar to natural process of antibody production in primates. However, the time and cost of affinity reagents produced from each of these platforms varies widely. Moreover, rapidly growing field of proteomics has created an urgent need for technology that can develop variety of affinity reagents for array of target proteins at affordable cost in short time on a high throughput platform. SELEX technologies are highly preferred when developing a platform for high throughput protein affinity reagent production.

SELEX also known as *in vitro* selection, is a method of simulating natural evolution in laboratory to obtain biological molecules with specific function (1). It is based on the theory that functional biological molecules originated from large number of random sequences. As illustrated in Fig.1.1, every *in vitro* selection begins with synthesis of large DNA libraries with random sequences of fixed length. The library has constant 5' and 3' flanking regions for amplification of library by polymerase chain reaction (PCR). Depending on the number of randomized nucleotides a library of complexity 10^{10}-10^{15} can be generated. Having large diversity allows exploration of large sequence space thus increasing the probability of isolating desired functional biomolecules. The pool of library biomolecules is equilibrated with target protein followed by washing to select for desired ligand. Target protein bound sequences are PCR amplified and subjected to another

Double stranded DNA library

SELEX
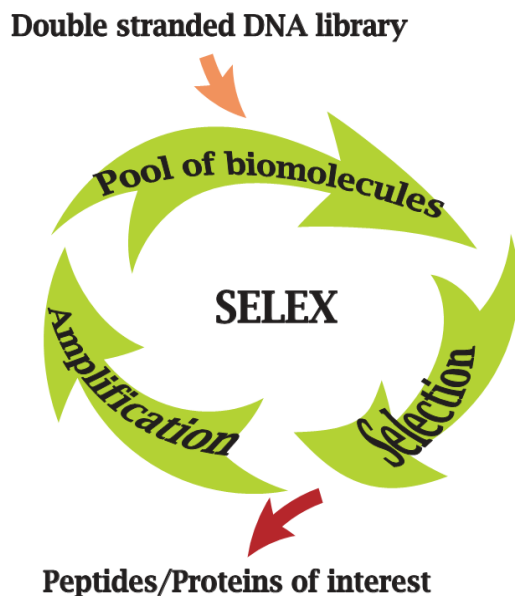
Peptides/Proteins of interest

Fig. 1.1 Schematic of SELEX technology Double stranded DNA library is used as starting point to obtain large diversity of biomolecules. Appropriate selection steps are performed in order to isolate biomolecules of interest. Output of selection is PCR amplified for another round of selection (39).

round of in vitro selection. Stringency of consecutive rounds can be increased to exclude weak binders. SELEX technology has been applied to identify DNA and RNA aptamers as well as peptides (2–5).Using peptides as affinity reagents is of great advantage due to their specificity, small size, ease of synthesis, ease of scaling up production, transportation and low cost involved. SELEX is fast, cost effective, less labor intensive and does not involve animal immunization. Variants of SELEX such as bacterial display, phage display, yeast display, cell surface display, ribosome display and mRNA display were developed to obtain desired peptides while retaining the linkage between genotype and phenotype during selections (6–11).

mRNA display is a method of choice for *in vitro* selections due to cell free nature of this system and possibility of using large diversity libraries for selections, to the order of 10^{15}. In most *in vitro* selections that involve at least one *in vivo* step, diversity of library is limited due to transformation efficiency of microorganism. Cell based *in vitro* selections introduce internal bias in selection process due to cellular bias during peptide expression or degradation. mRNA display uses *in vitro* translation in order to generate peptides/proteins. Chapter 2 describes optimization of template and commercially available cell lysate combinations to ensure maximum peptide/protein expression during selection. Furthermore, cell based *in vitro* selections do not permit use of most post translational modifications or unnatural amino acids. Alternatively, appropriate eukaryotic *in vitro* translation lysates can be used to translate variety of peptides/proteins with desirable post translational modifications and unnatural amino acids. As illustrated in Fig.1.2, mRNA display involves transcription of library followed by covalent linkage with puromycin linker. Puromycin is a structural mimic of tyrosyl-tRNA that allows covalent linkage between mRNA and corresponding peptide upon *in vitro* translation. Covalent linkage between genotype and phenotype in mRNA display, unlike ribosome display, allows application of stringent selection conditions to exclude majority of weak peptides to attain convergence earlier during selections. The small size of puromycin linker prevents selection of non-specific peptides interacting with linker. This is one of the major drawbacks of ribosome

3

display since large ribosomal complex consists of numerous proteins and ribosomal RNAs that interact with peptides in selection pool. mRNA display being completely *in vitro* technique allows use of *in vitro* mutagenesis or *in vitro* recombination at any step of selection. mRNA display allows fast, easy and cost effective selection of variety of functional peptides/proteins in few weeks. Chapter 3 provides detailed description and protocol of mRNA display.
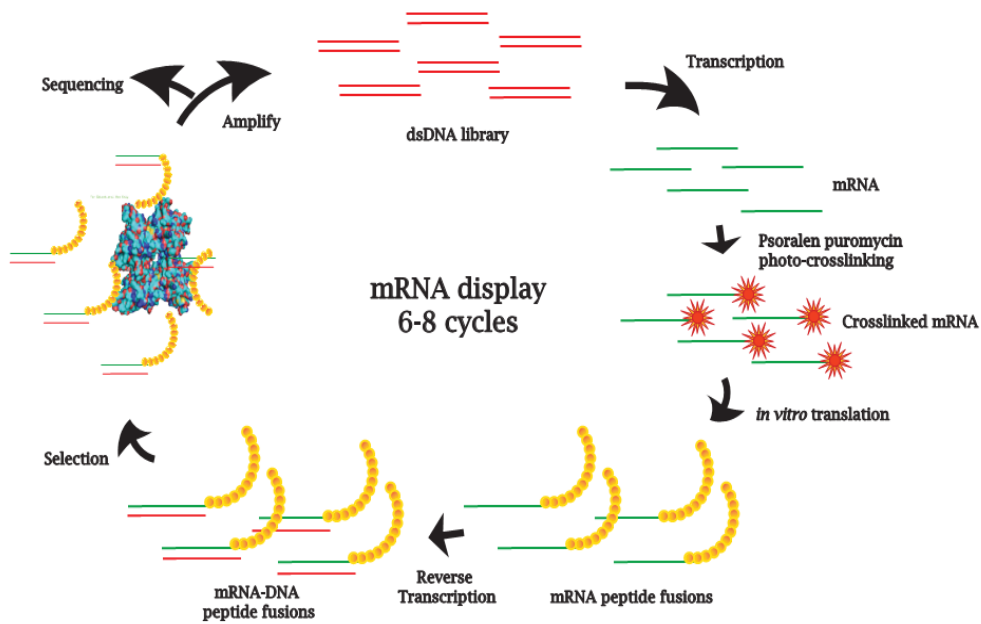

Fig. 1.2 mRNA display Double stranded DNA library is transcribed, crosslinked to psoralen - puromycin linker and translated using appropriate cell free translation lysate. Peptide – mRNA fusions are subjected to in vitro selection to obtain desirable functional peptides. Selection output is PCR amplified for next round of selection (40).

mRNA display is a powerful *in vitro* selection technique that allows isolation of desirable peptides/proteins from a large diversity library, as large as 10^{15}. However, starting selections with such vast complexity pool also increases possibilities of selecting nonspecific binders. Even after 10 rounds of selections it is not uncommon to find a few million peptides in the selection output. It can be very challenging to identify few high affinity peptides from a pool of millions of peptides if the downstream technology for characterizing peptides is not scalable. Most researchers perform next generation sequencing followed by computational analysis to identify affinity peptides. Computational algorithms that identify redundant amino acid sequences or motifs are valuable tools for analyzing large selection outputs. However, internal biases of selection technique, for instance bias introduced during PCR amplification of selection output, may add up over the rounds of selections thus misleading computational analysis towards false positives. Increasing use of selection strategies to obtain affinity peptides demands advanced experimental screening technologies that will help us identify potential high affinity peptides from large selection outputs. Most of the current techniques that characterize binding properties of peptides such as surface plasmon resonance (SPR), fluorescence resonance energy transfer (FRET), electrophoretic mobility shift assay (EMSA) are either time consuming, expensive, labor intensive or demand exceptional instrumentation and technical expertise dedicated to peptide characterization method. As a result most laboratories that perform selections are not able to identify and characterize valuable output that *in vitro* selections are generating.

Dot blot is widely used for detection and identification of biomolecules as quick cost effective alternative to chromatography and electrophoresis. In third world countries, dot blot has saved millions of lives by detecting sexually transmitted diseases, *Chlamydis trachomatis* infections and antibodies to tuberculosis and typhoid fever (12, 13). In Chapter 3, I propose a double membrane dot blot method for rapidly screening selection output to identify potential affinity peptides. This method can be used to further characterize binding properties of potential peptides to isolate high affinity binders. In order to analyze

peptides, they are *in vitro* expressed and purified from corresponding oligonucleotide sequence, which bypasses expensive and time consuming solid phase synthesis of peptides. The peptides are equilibrated with target protein to separate unbound free peptide from bound peptide-protein complex based on size. Dot blot method entails use of a very simple and inexpensive dot blot apparatus, which is assembled with regenerated cellulose and nylon membrane system. Selection output can be directly subjected to dot blot screen to identify potential affinity peptides, without need for prior sequencing, by equilibrating with appropriate concentration of target protein. Solution binding constant of identified potential peptides is determined by equilibration with wider range of concentration of target protein. Double membrane dot blot system is a rapid, high throughput and affordable solution to screen, identify and characterize selection output without need for exceptional instrumentation.
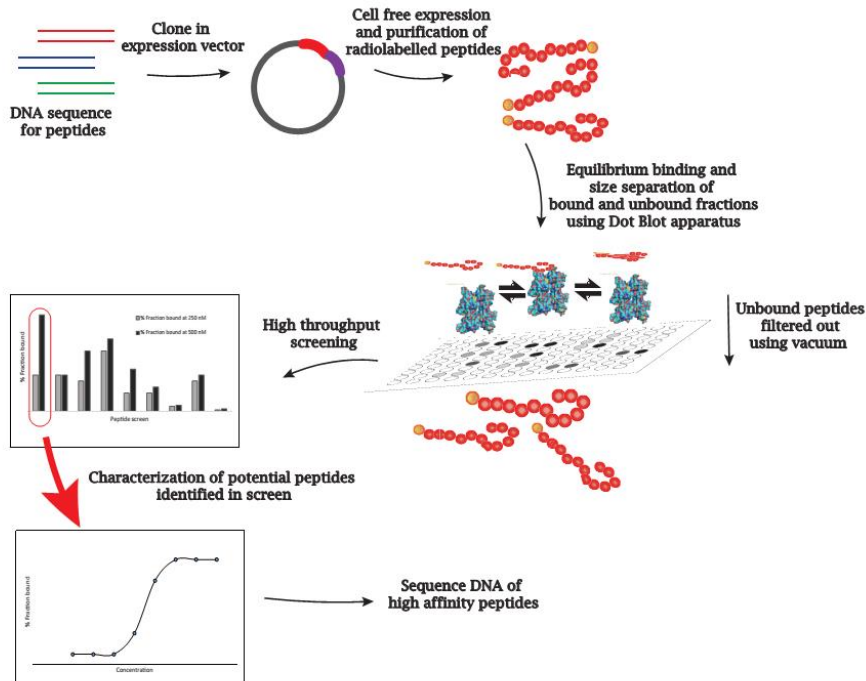
Fig. 1.3 Dot Blot method Radiolabelled peptides are expressed and
equilibrated with appropriate concentrations of target protein. Size based
separation of equilibrated reaction is used to screen peptides obtained from
selection output. Potential high affinity peptides can be further characterized by
equilibrating with wider range of target protein concentrations to determine Kd.
Sequencing DNA corresponding to high affinity peptides eliminates need for
next generation sequencing thus reducing cost of characterizing selection output.

My thesis proposes use of mRNA display technology to select for affinity peptides from
large sequence space using high complexity DNA libraries as described in Chapter 3. I
propose use of appropriate vector template and cell free lysate combination for in vitro
translation of peptides as illustrated in Chapter 2. Challenge of identifying high affinity
peptides from large diversity of in vitro selection output can be addressed using dot blot
method. Dot blot method allows efficient screening of selection peptides without need for

next generation sequencing. Potential high affinity peptides identified in screen can be further characterized using size based separation principle of dot blot method (Chapter 4). Thus I propose a time, labor and cost effective pipeline for generating high affinity peptides for diverse proteins in high throughput format.
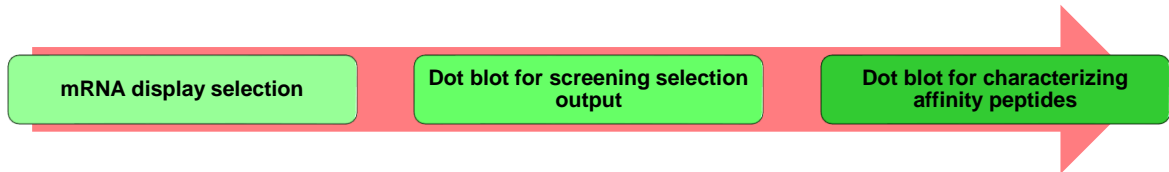
Fig. 1.4 Pipeline for isolating high affinity peptides Proposed pipeline illustrates three major components for isolating high affinity peptides; in vitro selection of desirable peptides using mRNA display, screening selection output to identify potential affinity peptides by dot blot method and characterization of high affinity peptides to determine Kd.

Chapter 2

Comparison of commercial cell-free translation systems for

optimal expression of DNA libraries

Contributions

The following chapter describes optimization of commercial cell free expression systems. The project was conceived by Dr. John Chaput. Reported experiments were performed by Pankti Shah,

Abstract

Cell-free protein synthesis is essential for many biotechnology and functional proteomics applications. However, not all protein expression vectors and translation systems are compatible. One of the most common reasons for low protein yields is reduced efficiency of translation initiation. Here, translation efficiency of three commonly used commercial cell-free translation systems is compared with an expression vector carrying three different translation enhancing elements (TEEs). Luciferase reporter vector was used in translation systems composed of wheat germ extract (WGE), HeLa cell lysate (HCL), and rabbit reticulocyte lysate (RRL) to compare encephalomyocarditis (EMCV) internal ribosome entry site (IRES) to the TEEs, which were derived from alfalfa mosaic virus (AMV) and tobacco mosaic virus (TMV). In addition, the role of poly A tail in enhancing protein expression as well as linear versus plasmid DNA templates were compared. The results of comparison of translation systems demonstrate that vectors with EMCV IRES are more efficient in HCL and vectors with TEEs are more efficient in WGE and RRL lysates. Besides, plasmid DNA increases translation 10-fold as compared to linear DNA and polyA tail failed to show any significant effect. We suggest a general guide for identifying vector and cell-free expression system compatibility. These findings will be used to optimally express DNA libraries to maximize unbiased representation of library diversity.

Introduction

Cell free translation systems have added new dimensions to the field of protein expression by coupling transcription and translation *in vitro*. Optimization of *in vitro* translation technology has received remarkable attention due to applications in creating protein arrays (10), expressing DNA libraries to generate pool of recombinant proteins (9), incorporation of unnatural amino acids (8, 12), high-throughput screening of peptide libraries (11, 17), structural and functional analysis of individual and membrane proteins (13–15) and understanding the mechanism of translation initiation (21). Continuous protein synthesis systems such as continuous-flow cell-free (CFCF) translation (24), microfluidic array devices (18) and protein producing gel (22) have facilitated protein expression in high throughput format with significantly reduced cost. In addition to these strategies, it is extremely valuable to optimize vectors for increased expression. Incorporation of critical sequence elements flanking the gene of interest greatly enhances gene expression. For instance, IRES is often incorporated into the 5' UTR of expression vectors to compensate for reduced cap-dependent initiation of translation *in vitro*. Optimal expression becomes crucial in case of mRNA display where expression of each variant of the DNA library is significant qualitatively and quantitatively. Commercial cell free translation systems were compared to ensure maximum diversity and equal representation of peptides during mRNA display selections. Comparison of EMCV, HCV (hepatitis C virus) and poliovirus IRES in HCL indicates that only the EMCV and HCV IRES increase protein yield when present upstream of a reporter gene (19). These results indicate that translation enhancing *cis* elements increase protein production *in vitro* by mechanisms that are lysate specific.

Cell free translation lysates from hyperthermophiles, hybridomas, Xenopus oocytes, insect, mammalian and human cells (7, 20), reconstituted translation systems from *E. coli* (23) *and Thermus thermophilus* (25) have been developed for varied applications in protein expression studies. Eukaryotic RRL, WGE and HCL based systems have been extensively

used for their unique characteristics each. Mammalian RRL allows cap-independent

translation but fails to glycosylate proteins (4). WGE produces high amounts of large

protein with negligible off target protein production but fails to confer co- and post –

translation modifications and often leads to premature termination (5). Human HCL allows

co- and post – translation modifications but is a relatively new system and has proven to

be sensitive to additives (3). The large variety of available translation systems has led to

a significant optimization of expression vectors; however, no direct comparison exists to

demonstrate which combinations are ideal starting points (1, 2, 6, 19). Promega

recommends presence of a Kozak sequence with an ATG start codon in an appropriate

context and a poly A tail for optimal expression in eukaryotic systems. Their observations

indicate that the addition of poly A tail compensates for absence of Kozak sequence in

RRL (1, 2, 6). *Langlais et. al*. demonstrated the significance of template sequence

optimization for protein production in *E. coli* and WGE based systems (16). Here we
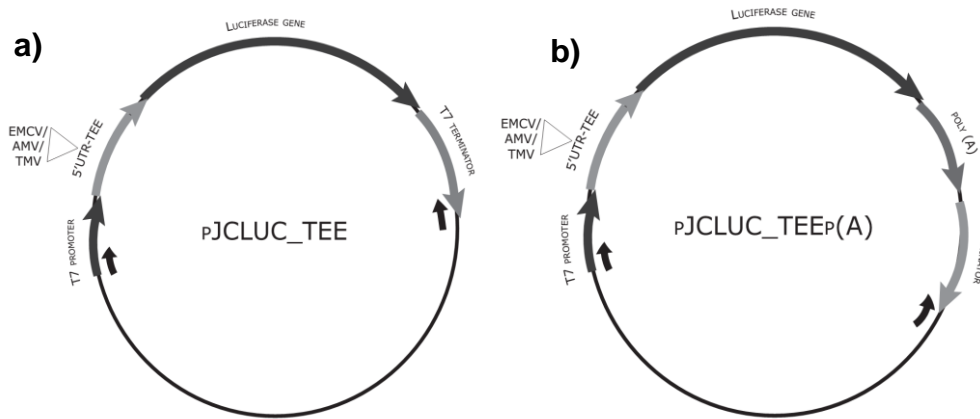
Fig. 2.1 Design of vectors to compare commercial cell free translation systems; a) pJCLuc_TEE and **b)** pJCLuc_TEEp(A) represent 6 vector constructs, 3 each. All of the vectors have T7promoter, either of unique 5'-UTR (EMCV, AMV or TMV), Firefly Luciferase reporter gene and T7 terminator as shown. pJCLuc_TEEp(A) has an additional poly A site between Luciferase gene and T7 terminator. Vectors have identical backbone sequence. Arrows indicate primer binding region for generating linear DNA template.

compare the potential of commonly used TEEs to enhance the functional luciferase protein yield across commercially available lysates. We followed recommended manufacturer's procedures in order to test the effect of three common 5' UTR sequences, EMCV, AMV, TMV in three common eukaryotic cell free translation systems, WGE, HCL and RRL. We also compared the effect of circular versus linear DNA template and the effect of adding a poly A tail downstream of a luciferase reporter gene.

Experimental design

Molecular cloning and template preparation

pJCLuc plasmid with monocistronic Firefly Luciferase under the T7 promoter was received as a generous gift from Dr. Doudna's lab, University of California, Berkeley, USA. We used enzymes from New England Biolabs, Ipswich, MA, USA to modify pJCLuc plasmid by standard molecular cloning techniques. All the plasmids created from pJCLuc parent plasmid had T7 promoter, single unique translation enhancing element (TEE; EMCV, AMV or TMV), Luciferase reporter gene and T7 terminator. Some plasmids were further modified to insert an additional 62 nt. long poly A tail between Luciferase reporter and T7 terminator (fig. 2.1). We amplified all the plasmids in XL-1 Blue cells and purified plasmids using PureYield plasmid miniprep system (Promega, Madison, WI, USA). We quantified plasmid preparation by Nanodrop (Thermo Scientific, USA). We used above mentioned vectors as templates to generate linear DNA by polymerase chain reaction (PCR) using T730mer.F - 5' CAAGCTCATTAATACGACTCACTATAGGCC 3' and T7 terminator long.R - 5' GGTTATGCTAGTTATTGCTCAGCGG 3' primers. Linear DNA was purified (Quiagen, Valencia, CA, USA) and quantified by agarose gel electrophoresis using low mass DNA ladder (New England Biolabs, USA).

in vitro translation and Luciferase assay

We used 500 ng of circular plasmid DNA and 400 ng of linear PCR generated DNA as templates for cell free expression, unless otherwise stated. We followed recommended protocols from manufacturer (Promega; L1170 and L4140 and Thermo Scientific; 88882) to perform coupled *in vitro* transcription and translation reactions. T7 PCR enhancer was not included in the cell free expression reaction of linear template. We incubated all the reactions at 30 °C for 90 minutes. We performed luciferase assay with 10 times diluted *in vitro* translation reactions. Glomax 96-microplate luminometer (Promega, USA) was used to quantify the light signal from Firefly Luciferase with delay time of 2 seconds and

14

integration time of 10 seconds. Each data point in the figures represents three independent cell free expression reactions.

TEE Sequences

EMCV in pJCLuc vectors –

GGTTATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTC
TTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAA
TGTCGTGAAGGAAGCAGTTCCTCTGGAGGCTTCTTGAAGACAAACAACGTCTGTAGCGACC
CTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGT
GTATAAGATACACCCGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTG
TGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTCAACAAGGGGCGGAAGGATGCCCAGAA
GGTACCCCATTGTATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGT
CGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACA
CGATGAT

AMV in pJCLuc vectors – TTTTTTTATTTTTAATTTTCTTTCAAATACTTCCACCATGG

TMV in pJCLuc vectors – TTACAATTACTATTTACAATTACA

Results and Discussion

5'- UTR TEE, presence of poly A tail and mode of template generation i.e. circular plasmid versus PCR generated linear DNA template were tested for WGE, HCL and RRL based cell free translation. EMCV IRES, AMV and TMV TEEs were incorporated into the 5' UTR of the luciferase reporter gene to compare their potential to enhance cell free expression in different cell lysates. Each construct contains an identical T7 promoter, luciferase gene and T7 terminator site. The potential for a poly A tail to influence translation was studied by incorporating a 62-mer poly A tail downstream of the luciferase gene. Circular plasmid and PCR-generated, linear DNA templates were compared for each of the constructs (Fig. 2.1). Luciferase activity was measured for each construct in each of the three lysate systems using a standard luciferase assay read by a Glomax 96-microplate luminometer.

The template format has a significant impact on the amount of protein produced via cell-free translation. Luciferase activity from circular plasmid DNA templates (Fig. 2.2a) was at least 10-fold higher than luciferase activity from PCR generated linear DNA templates (Fig. 2.2b). Each of the templates contained identical information from T7 promoter to T7 terminator region. In certain situations, PCR fragments can be much easier to obtain but suffer significantly in terms of protein yields. Certain manufacturers recommend addition of a poly A tail in order to compensate for the absence of a Kozak sequence but we find it does not increase production in our system and long adenosine repeats are difficult to clone.

The selection of an appropriate translation enhancing element in the 5' UTR can significantly alter the outcome of a cell-free translation. TEEs, AMV and TMV are comparable in WGE. TMV in WGE produced 32423-fold more luciferase than EMCV. While in HCL, EMCV is the ideal choice with 234-fold more protein than either AMV or TMV. For RRL the results were all very similar. We found this odd and noticed that with

16

decreased amount of template there was a significant difference in ability to express protein (data not shown). However, when using the amount of plasmid recommended by the manufacturer there is no significant difference in the amount of luciferase produced. We would still recommend TMV because it performs better at lower concentrations of template and thus has a much better chance for higher production if diverse proteins are made or the template is not under saturating conditions.

Of all the constructs, EMCV IRES in HCL produced highest amount of luciferase protein. In contrast, we observed that EMCV IRES in WGE and RRL resulted in lowest protein yield, indicating that EMCV IRES has a lysate specific mechanism to enhance translation. Although commercially available HCL based cell free translation system is almost twice as

a) Plasmid

☐ p(A)62 ■ p(A)0

b) Linear

☐ p(A)62 ■ p(A)0
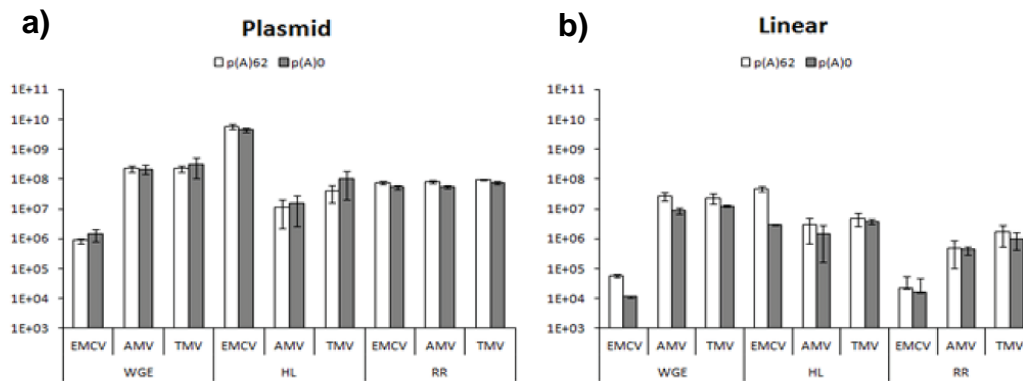
Fig. 2.2 Comparison of commercial cell free translation systems   Constructs containing different TEEs with poly A signal (white bars) and without poly A signal (gray bars) were used as template for cell free translation in three different cell lysates as indicated. Here, EMCV indicates pJCLuc_EMCV, AMV indicates pJCLuc_AMV and TMV indicates pJCLuc_TMV vectors. $p(A)_{62}$ indicates vector containing 62 bp long poly A signal downstream of Luciferase reporter and $p(A)_0$ indicates the gene devoid of a poly A signal followed by T7 terminator. **a)** Circular plasmid vector and **b)** Linear PCR generated DNA template.

expensive as RRL and WGE based systems, HCL produces up to 200-fold more protein when used with the appropriate *cis* element. This can be of great advantage when one wants to produce large amount of correctly folded human protein with associated post-translational modifications. Particularly, during expression of large DNA libraries where maximizing diversity is elementary for success of future applications. Appropriate template sequence choice during cell free translation can have dramatic effects on protein yield and our results represent ideal starting points that should yield significant material without exhaustive optimization.

Conclusion

Knowledge of using appropriate 5'- UTR TEE – lysate combination can save cost, effort and time of protein production significantly. Our results demonstrate that optimizing the template sequence for a given lysate can greatly increase protein production. We recommend use of EMCV IRES for coupled transcription and translation in HCL and AMV or TMV TEEs for WGE and RRL. These suggestions could serve as starting points for labs that are interested in producing protein from cell free lysates quickly and easily. These comparisons also serve as a benchmark for screening lysate - specific leader sequences.
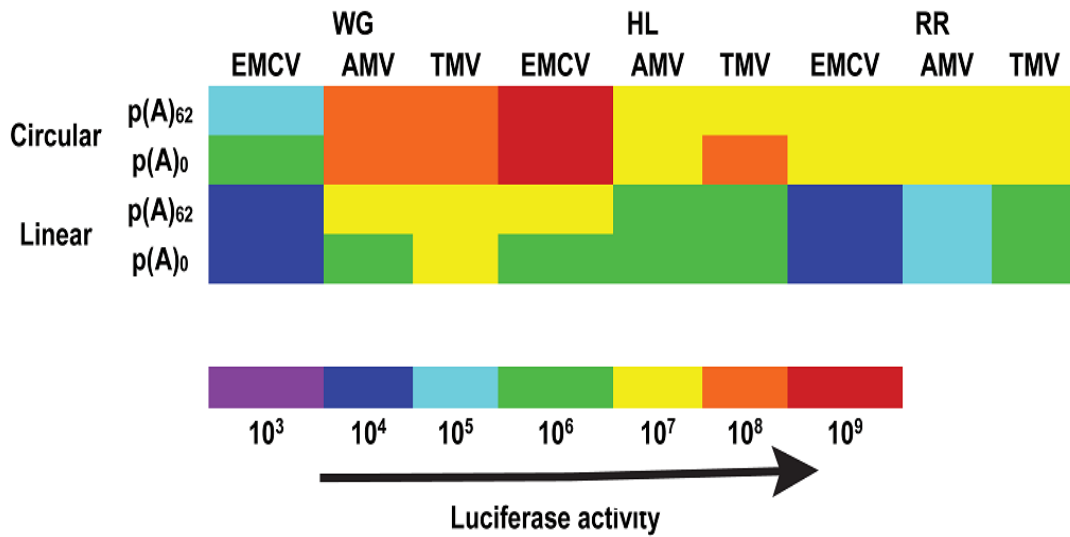
Fig. 2.3 Summary of comparison of commercial cell free translation systems Color-coded summary of results from comparison of EMCV IRES, AMV and TMV TEEs across commercially available WGE, HCL and RRL based cell free translation systems. The Luciferase activity of the construct goes on increasing 10-folds each time from Violet color to Red color as shown.

Chapter 3

Development of affinity peptides using mRNA display

Contributions

The following chapter describes the selection and characterization of affinity peptides using mRNA display technology. The project was conceived by Dr. John Chaput. Reported experiments were performed by Pankti Shah, Ian Shoemaker, Gokhan Demirkan and Columba Kim under the guidance of John Chaput, Mitch Magee and Joshua LaBaer. Pankti Shah performed mRNA display selections. Ian Shoemaker, Gokhan Demirkan and Columba Kim characterized selection output using surface plasmon resonance imaging.

Abstract

Growing knowledge of human proteome demands simultaneous development of protein affinity reagents to understand, diagnose and treat diseases. However, building novel target-specific affinity reagents is an extremely expensive, laborious and time-consuming process. Chaput lab has developed ligand interactions by nucleotide conjugates (LINC) technology that reconstructs two peptides, selected via mRNA display, which bind discrete sites on a protein surface in to a high affinity protein capture reagent linked at specific distance and orientation on a DNA scaffold. Virtues of LINC technology enable economic and rapid development of high affinity reagents against desirable targets at ease. We propose to automate the development of antibody like high affinity and high specificity bivalent affinity reagents against 96 human kinases in high throughput format. From our experience, convergence of mRNA display selection to obtain high affinity and high specificity peptides requires 6-8 rounds of selection. Our long term goal is to limit selection to a single round in order to facilitate automation. As a pilot project towards this goal, our aim was to reduce number of rounds of mRNA display selection by introducing mild denaturant thus increasing selection pressure. We performed three rounds of selection against AKT3 target protein under native and mild denaturing conditions. Surface plasmon resonance (SPRi) analysis of three rounds of AKT3 selections indicate that probability of encountering binders increases with increasing rounds of mRNA display selections. Addition of mild denaturant (0.75M GuHCl) however, is not sufficient to enrich binders during selections. Binders obtained from AKT3 selections will be transformed in to bivalent affinity reagents using LINC technology. Results from individual experiments and optimization strategies will enable reduction of mRNA display selections to single round in order to facilitate automation and develop binders at reduced cost, effort and time.

Introduction

With the advent of high throughput genomics, our abilities to study genes and gene expression have grown tremendously. Increasing information about entire human proteome demands development of new technologies and tools to work with repertoire of proteins quickly and easily at low cost. Protein affinity reagents are indispensible tools required for understanding, diagnosis and treatment of diseases caused by protein dysfunction. Future medicine needs individual specific affinity reagents to monitor individual health status and personalize diagnosis of diseases. Thus establishing need for technology that operates on thousands of proteins at a time and accelerates discovery of high affinity and high specificity protein capture reagents. In addition, it is desirable to have new class of affinity reagents that are small, simple, easy to produce on large scale, cost effective and easy to transfer in order to facilitate field of proteomics.

Hybridoma technology to develop monoclonal antibodies is oldest and reliable method of choice to produce high affinity reagents *in vivo*. Monoclonal antibodies are widely used in basic research, disease diagnosis, therapeutics and other consumer products due to high specificity and high affinity. Development of monoclonal antibody by classical hybridoma technology requires animal immunization, takes 5 to 10 months and costs tens of thousand dollars depending on method of characterization (9, 15). In addition, hybridomas produce antibodies without any encoding DNA sequence information making it difficult to manipulate these antibodies in lab. Phage display technology is based on *in vitro* selection of desirable peptides from library of peptides displayed on bacteriophage surface. This technology non-covalently couples phenotype i.e. peptides on phage surface, to its encoding genotype i.e. cDNA enclosed in corresponding phage. Selected phages have to be transfected in suitable bacterial host for amplification and next round of selection. This method involves single *in vivo* step of transfection, thus limiting the diversity of library by the transfection efficiency. Peptides from phage display selection are translated in bacteria hence codon usage, folding and post-translation modifications of

mammalian systems are not always feasible (7). Ribosome display is a completely *in vitro* protein selection method that relies on maintaining ribosome-mRNA-polypeptide ternary complex to couple phenotype to genotype. Ribosome is stalled during translation by removal of stop codon, low temperature and high cation concentration to facilitate peptide/protein display on ribosome surface. This completely *in vitro* selection technique allows use of synthetic and natural libraries with large diversity. However, selection conditions have to be carefully optimized for maintenance of ternary complex, which limits use of stringent selection strategies to enrich desirable peptides/proteins (2, 12). Apart from these, various other methods like colony screening, yeast display, cell surface display have also been used. Most of these technologies are cumbersome, time consuming, require large amount of target protein for selections and limit diversity of libraries due to *in vivo* steps.
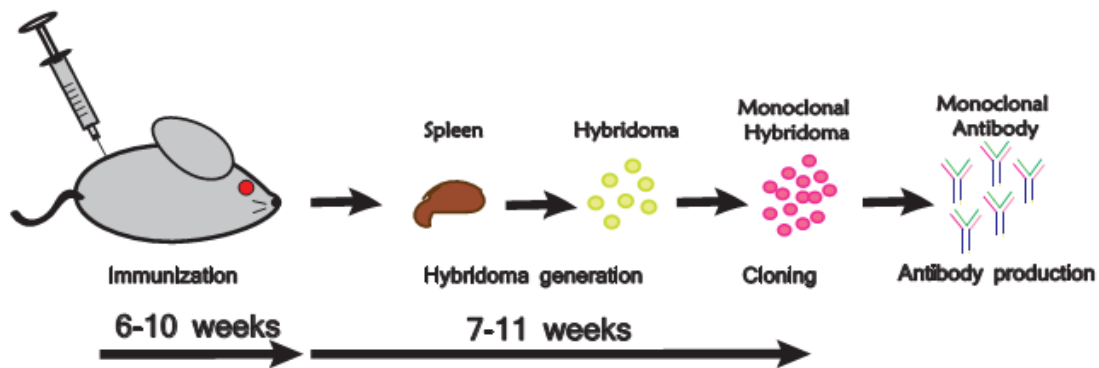

Fig. 3.1 Hybridoma Technology Antigen is injected in animal several times at interval of two weeks. Immunized spleen cells are extracted and fused to myeloma cells to create hybridoma. Hybridomas producing high affinity antibodies are cloned to acquire monoclonal antibodies.

mRNA display is extremely convenient completely *in vitro* selection technology which involves covalent linkage of phenotype i.e. peptide/protein to genotype i.e. its encoding mRNA via a psoralen puromycin linker (3). DNA library is *in vitro* transcribed using T7 RNA polymerase, pool of mRNA is purified and UV- crosslinked to psoralen puromycin linker. Puromycin, structural mimic of tyrosyl-tRNA, enters A site of ribosome towards the end of *in vitro* translation and covalently bonds to the translated peptide via natural peptidyl transferase activity of ribosome. Puromycin thus links mRNA to its encoding peptide. Different *in vitro* translation systems including, *E.coli* cell lysate, rabbit reticulocyte lysate, HeLa cell lysate, wheat germ lysate can be used for translation depending on purpose of selections. mRNA-peptide fusions are reverse transcribed to form stable RNA-DNA hybrid which also prevents selection of RNA aptamers. Desirable peptides can be selected by using wider range of conditions such as, pH, temperature, ionic strength, solvents, denaturants, due to profound stability of RNA-DNA-peptide fusions. Cell free nature of mRNA display reduces intrinsic biases introduced during transformation, cellular expression and easily aids construction of larger library sizes of the order of $10^{12} - 10^{14}$. Thus, mRNA display allows 10,000-fold greater sequence complexity than phage display, 10^6-fold greater complexity than yeast display and 10^9-fold greater sequence complexity as compared to colony screening methods (8). Higher diversity of libraries allows greater coverage of sequence space and better chances of finding high specificity high affinity peptides. mRNA display selections are performed at such low volumes that nM concentration of target protein is sufficient for multiple rounds of selections. Stringent selection strategy with freedom to use arbitrary selection conditions in addition to quick and easy amplification of selection output makes mRNA display easy and robust selection technology. mRNA display has been successfully used in past to develop proteins and peptides that bind target proteins and ligands with high specificity and high affinity (4, 11, 13).

For large protein library containing $10^{12} - 10^{14}$ unique sequences, distinguishing non-specific binders from specific binders can be challenging. Hence, it takes 6 - 8 rounds of mRNA display selections to achieve significant convergence of library. Each round of mRNA display typically takes 4 – 7 days depending on procedure of experiments. In most cases, it takes up to 2 months to obtain probable peptides/proteins for further characterization. Our goal is to reduce selection procedure to single round of mRNA display and facilitate automation for developing high affinity reagents from first rate binders using ligand interactions by nucleotide conjugates (LINC) technology. This technology uses DNA as a scaffold to hold two peptides (binding discrete epitopes of target protein) at specific distance and orientation to construct a high affinity bivalent protein capture reagent (5, 6, 10, 14). Our lab has previously used two low affinity Grb2 binding peptides that bind to different domains of Grb2 namely, SH2 and SH3 domain with $Kd = 0.5$ mM and $Kd = 5$ mM to create high affinity bivalent affinity reagent for Grb2. When two peptides were covalently linked on DNA scaffold, at specific distance and orientation, to recapitulate antibody like bivalent affinity reagent, they bound Grb2 with affinity of $Kd = 6.9 \pm 0.4$ nM, which is 5- to 10-fold stronger than commercially available Grb2 antibodies (5).

Reducing mRNA display to single round will enable quick and easy selection of first rate binders against various targets from a single parent pool of RNA-DNA-peptide fusions. These binders will be transformed in to high affinity bivalent reagents, called **nu**cleotide-**pro**tein apta**mers** (Nupromers), using LINC technology in automated high through put format. Automating high affinity protein capture reagent development will reduce time, labor and cost of production. As a pilot project towards our goal of reducing 6 – 8 rounds of mRNA display to single round of selection and amplification, we performed three rounds of mRNA display selections on AKT3 target protein, with and without 0.75 M guanidine hydrochloride (GuHCl) washes to obtain affinity binders. Surface plasmon resonance imaging (SPRi) analyses of small fraction of selection output suggest that

probability of finding binders increases with increasing rounds of selections however, adding denaturant at such low concentrations is not sufficient to enrich for higher affinity peptides. We selected human serine/threonine-protein kinase, AKT3 for its role as regulator of cell signaling. AKT3 is known to regulate cell proliferation, differentiation, apoptosis and tumorigenesis (1).

```
|-------T7  Promoter--------------|----|--------------TEE-------------------|Met|--------------N12  Random  Region----------------------|
5' TTC  TAA  TAC  GAC  TCA  CTA  TA  GGG  AGT  GGA  CGT  CAA  TAC  TTA  CGC  A  ATG  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS
3' AAG  AGG  ATG  CTG  AGT  GAT  AT  CCC  TCA  CCT  GCA  GTT  ATG  AAT  GCG  T  TAC  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS  NNS


    |--Thrombin Cleavage Site ---|------SbfI---|-------------|---X-Linker-----|
    CTG  GTG  CCT  AGA  GGA  TCT  CCT  GCA  GG  A  AGT  GGA  CAC  CGG  CTA  T  3'
    GAC  CAC  GGG  GCT  CCG  AGG  GGA  CGT  CC  T  TCA  CCT  GTG  GCC  GAT  A  5'
```
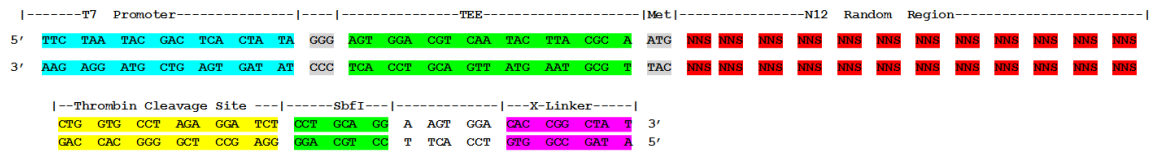
Fig. 3.2 DNA library used as starting pool for mRNA display selections DNA construct consists of T7 promoter (cyan), translation enhancing element (green), start codon (gray), random amino acid region (red), thrombin cleavage site (yellow), restriction enzyme cleavage site (green) and regionc omplementary to crosslinker (pink).

Experimental design

mRNA-Lib-For 5' TTCTAATACGACTCACTATAGGGAGTGGACGTCAATACTTACGC 3' and

mRNA-Lib-Rev2 5' ATAGCCGGTGTCCACTTCCTGCAGGAGATCCTCTAGGC 3' primers

were used for klenow and amplification of library.

mRNA-Lib-RT2 5' TTTTTTTTTTTTTTTTATCCACTTCCTGCAGGAGATCCTCTAGGC 3' primer

was used for reverse transcription. All oligo nucleotides were purchased from Integrated

DNA technologies (IDT, Coralville, IA). Quiaquick gel extraction kit was used to clean up

PCR (Quiagen, Valencia, CA). All enzymes were purchased from New England Biolabs

(Ipswich, MA) unless otherwise stated. DNA and RNA concentrations were determined by

absorbance readings at 260 nm using NanoDrop 1000 (Thermo Scientific, Waltham, MA).

5'-psoralen-TAG CCG GTG-(PEG9)2-A15-ACC-puromycin linker was used for psoralen-

DNA-puromycin crosslinking. Biotinylated AKT3 protein was a generous gift from Joshua

labaer lab. Commercial pJET cloning kit was purchased from Promega (Madison, WI).

In vitro transcription was performed using HEPES KOH Buffer pH 7.5 200 mM, freshly

prepared DTT 40 mM, bovine serum albumin 200 µg/ml, $MgCl_2$, nucleotide triphosphate

mix 7 mM each, spermidine 2 mM, rRNAsin 0.6 U/µl and T7 polymerase 0.2 U/µl and

template 50 nM – 200 nM final concentration. Reagents were mixed and incubated at 37

°C overnight.

In order to crosslink mRNA and psoralen-DNA-puromycin, HEPES KOH pH 7.5 20 mM,

KCl 100 mM, EDTA pH 8 1 mM, spermidine 1 mM, psoralen-DNA-puromycin linker 20

µM and mRNA template 5 µM final concentration were mixed in 1.5 ml black eppendorf.

Reaction mix was denatured at 65 °C for 5 min followed by annealing at room

temperature for 25 min. 50 µl aliquots were transferred to each well of 96-well plate and

irradiated at 366 nm for 15 min. Crosslinked material was collected in a single tube.

In vitro translation and reverse transcription reaction were performed as per

manufacturer's recommendations.

PCR cycle optimization was performed by visualizing PCR aliquots from 8-20 cycles at

interval of 2 cycles each on 2% agarose gel.

28

Peptides for SPRi analysis were purchased from Biomatik, Wilmington, DE.

Results and Discussion

A 20mer DNA library was designed with randomization at 12 amino acid positions. Library was klenow extended and PCR amplified for mRNA display selections. DNA library was constructed with 5' T7 promoter, to allow *in vitro* transcription by T7 RNA polymerase and a 22mer translation enhancing element (TEE) immediately upstream of peptide coding region to enhance translation in rabbit reticulocyte lysate. Enhancement of translation with this short 22mer sequence was comparable to tobacco mosaic virus derived TEE in rabbit reticulocyte lysate (data not shown). Coding region comprised of start codon followed by 12 random amino acids and a thrombin cleavage site will create ~10^{14} diverse peptides. Library complexity of approx. 10^{14} unique sequences would traverse sufficient sequence space to identify affinity binders against multiple motifs on various target proteins. We included an X linker region complementary to psoralen-puromycin crosslinker at the 3' end of the construct (fig. 3.2).

Run off transcription of DNA library produced pool of 107 bp mRNA molecules. Transcription reaction was DNase treated and purified using denaturing 10% Urea-polyacrylamide gel electrophoresis (PAGE) followed by electroelution and ethanol precipitation. Purified mRNA was annealed and photo-crosslinked to psoralen-DNA-puromycin crosslinker by irradiating at 366 nm for 15 min. Likewise, crosslinked mRNA was purified by denaturing 10% Urea-PAGE followed by electroelution and ethanol precipitation. Crosslinked mRNA was further translated *in vitro* using rabbit reticulocyte lysate for 1 hr at 30°C. Peptides were radiolabelled by addition of S^{35}-methionine in the translation reaction. Fusion formation was promoted by incubating translation reaction overnight with high concentration of KCl and $MgCl_2$ salts. Oligo-d(T) cellulose column was used to purify mRNA-peptide fusions. mRNA of the fusions was reverse transcribed with Superscript II to create stable RNA-DNA-peptide fusions. Reverse transcribing mRNA before selections protects degradation of mRNA by RNases and also prevents enrichment of RNA aptamers against target protein during selection.

Target protein, AKT3 was biotinylated for immobilization on streptavidin coated magnetic dynabeads (Life Technologies, Carlsbad, CA) before selections. In order to compare effect of mild denaturant on selections, one set of selections and washes were carried out in phosphate buffer saline containing 0.025% Tween (PBST) and 0.75 M GuHCl. We expected 0.75 M GuHCl to aid removal of unfolded weak interacting peptides while still maintaining major folds of peptides. RNA-DNA-peptide fusions were preselected for binding to streptavidin coated magnetic Dynabeads for 1 hr at room temperature to eliminate non-specific peptides binding to bead or eppendorf tube surfaces. Unbound fraction from pre-selection step was incubated with AKT3 coated magnetic dynabeads for 2 hrs at room temperature with rotation. After selection, supernatant was stored at -20 °C as unbound fraction and beads were washed 5 times with 1X PBST and suspended in nuclease free water. In order to prevent over amplification of bound fraction of library,

a)

RNA(107nts): 5'GGGAGUGGACGUCAAUACUUACGCAAUGNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSCUGGUGCCUAGAGGAUCUCCUGCAGGAAGUGGACACCGGCUAU 3'

Peptide: M X X X X X X X X X X X X L V P R G S P A G S G

b)

Crosslinked Fusion Schematic:

c)

Fig. 3.3 Transcription, crosslinking, translation and PCR cycle optimization during mRNA display a) 107 bp mRNA transcripts produced after run off transcription of dsDNA library and corresponding peptide sequence. **b)** Schematic of mRNA-psoralen puromycin crosslinked fusion. **c)** 2% agarose gel showing PCR cycle optimization, numbers in each lane correspond to number of cycles of PCR used for amplification of selection output.

number of PCR cycles required to amplify selected pool needs to be optimized. We used 20% of the bound beads as template for PCR cycle optimization (Fig. 3.3c). RNA-DNA-peptide fusions on the surface of rest of the dynabeads were then amplified using optimal number of PCR cycles. Amplified library was purified using Quiaquick PCR clean up kit, cloned in commercial pJET vector and sequenced to determine sequence of peptides that were selected as AKT3 binders. Part of this amplified library was used as starting pool for next round of mRNA display selections against AKT3. After three rounds of AKT3 selection with and without mild denaturant, we cloned and sequenced outputs from every round of selection.

Selection peptides were ordered in 96 well-format, sixteen from each round, with terminal cysteine containing GSC linker. Peptides were immobilized on gold surface using amine coupling chemistry for SPRi analysis. Plexera proteomic processor V3 was used to process array of peptides by flowing AKT3 target protein over chip surface while recording measurements for each peptide. After binding, running buffer alone was allowed to flow over the chip surface to measure off rates. SPRi experiments were performed in duplicates using a transferrin specific binder, B10 as control. Data collected from our experiments were extracted and analyzed using Scrubber software to calculate Rmax of each peptide. Selection peptides that showed Rmax greater than B10 were considered as binders. Rmax of binders was normalized with respect to B10. Our data indicates that probability of obtaining affinity binders increases slightly with more rounds of selection. Amongst the sequences we used for characterization, more binders were obtained from round three than from second or first round. Fig. 3.4b shows a plot of total number of binders obtained from each round of selection. As evident from graph, adding mild denaturant during selections was not sufficient to increase selection of binders. Binders with high, medium and low Rmax were found to be selected in every round of selection. Fig. 3.4c shows curves of three representative peptides from AKT3 selections. R3-8 is one of the highest affinity AKT3 binding peptide characterized so far.

Conclusion

Above results indicate that increasing number of rounds of mRNA display increases probability of encountering binders. Presence of mild denaturant (0.75 M GuHCl) is not sufficient to enrich for binders during mRNA display selections. Peptides obtained from selections can be used to generate bivalent affinity reagents, Nupromers for AKT3 using LINC technology. Results from our pilot experiments will be used to fine tune mRNA display selections to reduce cost, time and effort of obtaining high affinity high specificity binders for target proteins. These efforts will enable automation of LINC technology to develop bivalent affinity reagents in high throughput format.
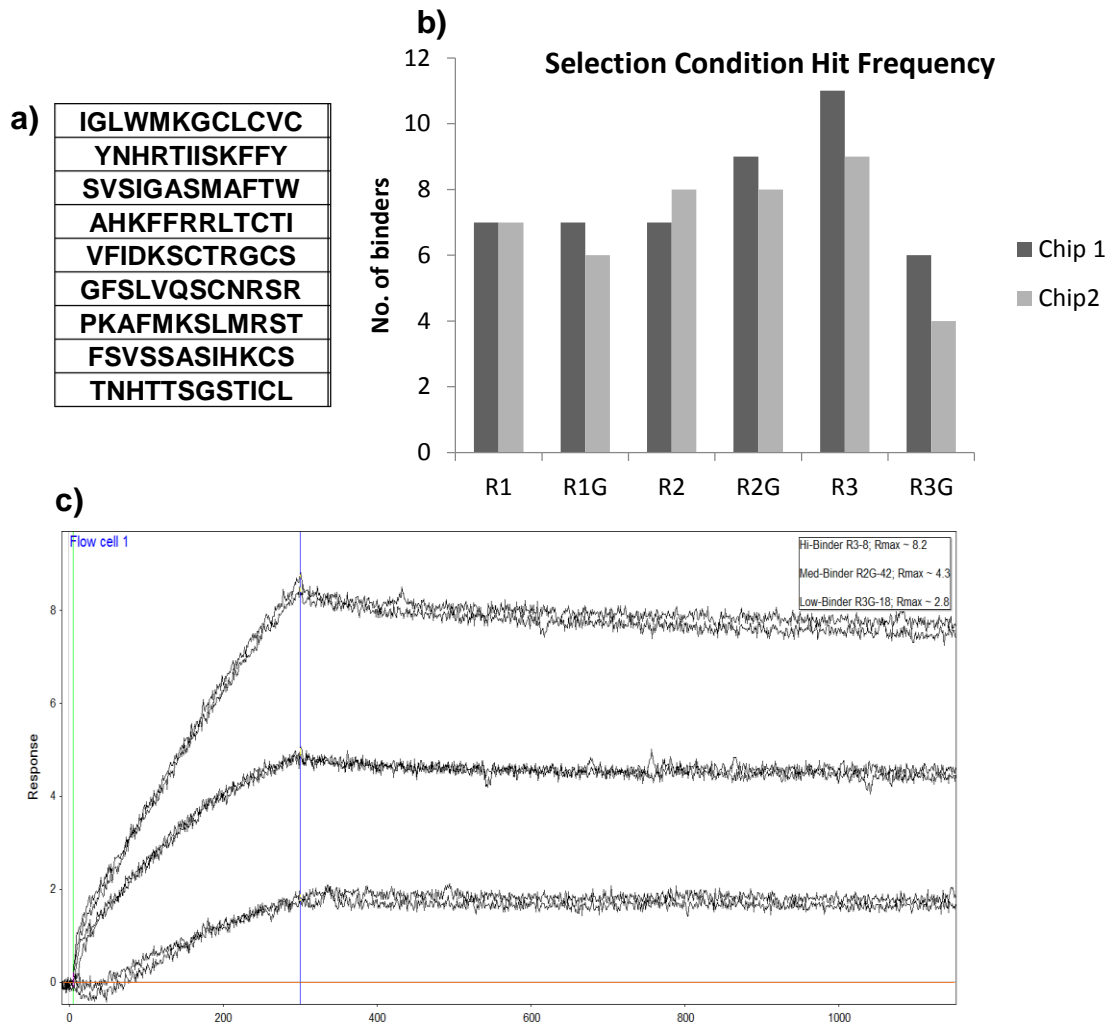
a)

IGLWMKGCLCVC
YNHRTIISKFFY
SVSIGASMAFTW
AHKFFRRLTCTI
VFIDKSCTRGCS
GFSLVQSCNRSR
PKAFMKSLMRST
FSVSSASIHKCS
TNHTTSGSTICL

b) Selection Condition Hit Frequency

c)

Hi-Binder R3-8; Rmax ~ 8.2
Med-Binder R2G-42; Rmax ~ 4.3
Low-Binder R3G-18; Rmax ~ 2.8

Fig. 3.4 Analysis and characterization of mRNA display selection output

a) Representative peptides from AKT3 selection. **b)** Plot of total number of binders obtained from each round. R is an abbreviation for round, number indicates number of rounds and G indicates presence of mild denaturant during selections. Chip 1 and Chip 2 are two replicates performed on different days. Total number of peptides screened for each round was 16. **c)** Representative SPRi plot of high, medium and low affinity peptides obtained from three rounds of selections against AKT3 via mRNA display.

(Figure of SPRi was provided by Shoemaker I)

34

Chapter 4

Dot blot method for screening and characterization of affinity peptides

Contributions

The following chapter describes the screening and characterization of affinity peptides using double membrane dot blot system. The project was conceived by Dr. John Chaput. Reported experiments were performed by Pankti Shah, Andrew Larsen and Katie Fenton under the guidance of John Chaput. Pankti Shah expressed and purified peptides to screen and characterize peptides using dot blot method. Andrew Larsen designed oligonucleotides for cloning and performed dot blot. Katie Fenton made recombinant expression vectors using molecular cloning tools.

Abstract

Systematic evolution of ligands by exponential enrichment (SELEX) is widely accepted method of choice to meet the growing need for protein affinity reagents. Tremendous diversity of selection output however, demands careful characterization to be able to identify most appropriate affinity peptide. Here we demonstrate a rapid affordable and high throughput approach to screen and characterize binding kinetics of peptides with target protein. This method entails, *in vitro* translation and purification of labelled peptides followed by equilibrium binding with target protein and separation of bound peptide-protein complexes from unbound peptides. Size dependent separation of bound and unbound peptide using simple membrane system is fundamental to our dot blot method of screening and characterization. We validate our technology by screening and characterizing previously reported human α-thrombin mRNA display selection output. Rapid screening of uncharacterized selection output led to identification of several human α-thrombin binding peptides with low nanomolar solution binding affinity (Kd), range 150 – 450 nM. We offer economic and simple solution to analyze selection output and characterize peptides in high throughput format to accelerate discovery of protein capture reagents.

Introduction

Increasing demand for protein capture reagents over past decade has led to advancement of SELEX technology as a tool to identify protein affinity peptides (1–5). *in vitro* selection technologies allow cost effective exploration of large sequence space, starting with complex DNA libraries, to simulate evolution and identify affinity binders. However, much success and advancement of selection technologies has not directly translated into rapid affinity peptide development. Several factors slow down the discovery and development of affinity peptides despite of powerful selection platform. Starting with large diversity libraries of approx. 10^9 to 10^{13} complexity, even 10 rounds of selections result in millions of survivor sequences. Some actively selected low, medium and high affinity binders whereas some passively selected peptides. Next generation sequencing allows us to sequence voluminous selection output however, we fail to identify high affinity binders from millions of reads of potential binders. Although, computational tools help identification of convergence patterns over rounds of selections, analysis is limited by internal experimental bias that leads to over representation of certain amino acid sequences over others. Due to cost and time associated with peptide synthesis and instrumentation for measuring solution binding kinetics, it is not practical to characterize large number of potential peptides identified after next generation sequencing and computational analyses. Thus bottleneck for protein capture reagent development seems to be shifted from affinity peptide development to identification and characterization of potential peptides. Development of affordable, simple, and quick high throughput screening and characterization platform for peptides will solve major hurdles faced during protein capture reagent discovery.

To overcome the restraints of current methods, we developed double membrane dot blot method, an affordable solution to rapidly screen and characterize peptides for their affinity to diverse target proteins, using minimal instrumentation. Major advantage of our technology comes from user friendly and cost effective nature of steps involved. In our

37

system, coupled transcription and translation systems are employed to express labelled peptide from DNA constructs. Labelled peptides have affinity tag which make them amenable to purification via affinity column, thus producing large amounts of pure peptide easily in days as opposed to time, cost and effort associated with solid phase peptide synthesis. Labelled peptides are then equilibrated with target protein followed by partitioning using simple membrane partitioning system that separates bound and unbound peptide-protein complexes based on size. Our method is easily scalable to large number of peptides making it feasible to automation. Screening and characterization of sequences by dot blot method does not require large amounts of target protein or prior sequencing of selection output, which largely contribute to the cost of affinity reagent development. However, when combined with sequencing and computational analysis, our technology can greatly increase the efficiency and pace of affinity reagent development.

As illustrated in Fig. 1.3, DNA oligonucleotides, obtained from *in vitro* selection output, corresponding to peptide sequence is cloned in expression vector for expression and affinity purification of peptide. Expression vector consists of T7 promoter for efficient *in vitro* transcription, an appropriate translation enhancing element (TEE) for enhancing efficiency of cap independent translation *in vitro*, a C-terminal streptavidin binding protein (SBP) tag for affinity purification and a tobacco etch virus (TEV) protease cleavage site to specifically elute peptide bound to streptavidin affinity column. Recombinant plasmids are transcribed and translated using commercially available rabbit reticulocyte lysate coupled transcription and translation system in presence of S^{35}-labelled methionine to generate S^{35}-labelled peptide- SBP fusions. Peptides are affinity purified from crude lysate using streptavidin agarose column. Peptides are eluted by TEV cleavage or as SBP fusions with water. Purified peptides are equilibrated with target protein at fixed concentrations, 250 nM and 500 nM in present study, to screen and identify potential binders. In case fraction of peptide bound positively correlates with concentration of target protein, those particular peptides are further characterized to measure solution binding

affinity (Kd) by equilibration with various concentrations of target protein spanning 5 µM – 1 nM. Peptides of interest that exhibit high affinity to target protein (Kd in low nM range) are sequenced to determine amino acid composition.

Above mentioned screening and characterization is based on selective partitioning of bound peptide-protein complex from unbound peptides based on size which is accomplished by use of a double membrane dot blot system. Dot blot microfiltration apparatus consists of two 96 well plates and a vacuum base. Layers of different membranes can be sandwiched between two plates to achieve specific filtration goals. Combination of regenerated cellulose and nylon membranes have been used in past to measure binding affinity of DNA and RNA aptamers to protein targets (6–8). Many membranes such as regenerated cellulose, polyvinylidene fluoride and nylon bind protein and nucleic acids based on physiochemical properties like charge and hydrophobicity. Since our goal was to separate peptide-protein complexes based on size, we used 10K dialysis membrane made from regenerated cellulose. Small pore size allowed us to prevent filtration of peptide-protein complex combined with low power vacuum, approx. 40 torr, for slow and efficient filtration of free peptides from equilibrated solution.

Here we screen 24 peptides obtained from human α-thrombin mRNA display selection to identify potential affinity peptides using double membrane dot blot technique. Five potential high affinity peptides were further characterized to determine their binding affinity to human α-thrombin and was found to be in range of 150 – 450 nM.
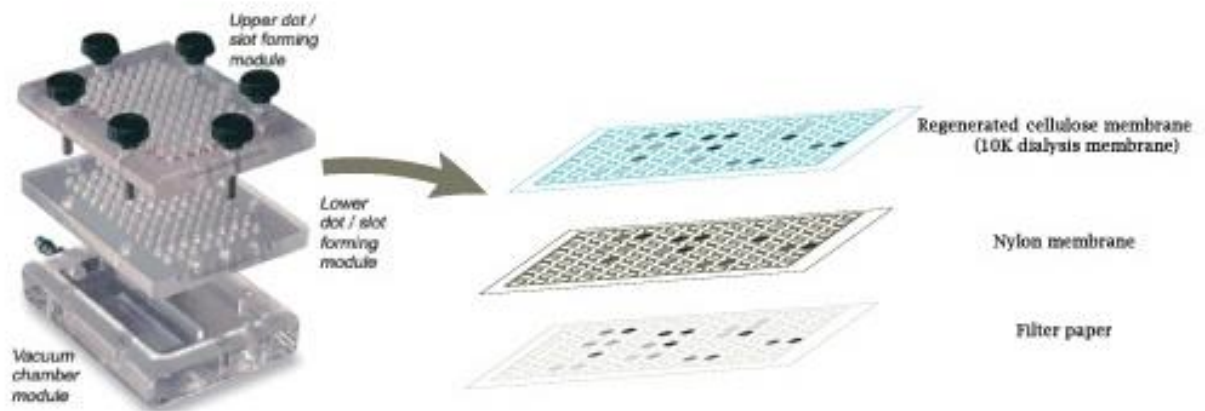
Fig. 4.1 Double membrane dot blot system Dot blot apparatus containing two membrane system for size separation of unbound peptides from bound peptide-protein complexes. Dot blot apparatus is assembled as shown with dialysis membrane directly facing top 96-well plate followed by nylon membrane and filter paper from top to bottom respectively.

(Figure of Dot Blot derived from Harvard Apparatus)

Experimental design

Construction and amplification of linear DNA constructs

DNA oligonucleotides encoding for various peptides were annealed and klenow extended to make double stranded peptide coding sequence (Integrated DNA Technologies Inc., Coralville,IA; New England Biolabs Inc., Ipswich, MA). These double stranded peptide encoding DNA constructs were PCR amplified and inserted in expression vector with T7 promoter and C-terminal Streptavidin binding protein tag (*Wilson et. al.,* 2001). Recombinant vectors were transformed in XL-1 Blue cells and transformants were grown to amplify pure recombinant plasmid. PureYield plasmid miniprep system (Promega, Madison, WI) was used to isolate plasmid from clones. Linear DNA with T7 promoter, translation enhancing element, peptide coding sequence followed by streptavidin binding protein tag were PCR amplified using T730 mer.F -

5'CAAGCTCATTAATACGACTCACTATAGGCC3' and T7 terminator long.R -

5'GGTTATGCTAGTTATTGCTCAGCGG3' primers.

In vitro expression of peptides

Peptides were *in vitro* transcribed and translated for 90 min at room temperature in rabbit reticulocyte system using linear DNA as template (Promega L1170, Madison, WI). SBP tagged peptides were affinity purified using Sterptavidin agarose column (Thermo Fisher 20347, Rockford, IL) (9). Peptides were eluted by ProTEV protease cleavage overnight at room temperature leaving streptavidin binding protein tag on Streptavidin agarose column (Promega, Madison, WI) (10).

Dot blot

Minifold I 96 well dot blot apparatus from Whatman was used to quantify fraction of bound peptide (GE Healthcare Life Sciences, Piscataway, NJ). In order to assemble dot blot apparatus, a filter paper, two nylon membranes and a 10K dialysis membrane were placed between sample well plate and filter support plate from bottom to top respectively

(Thermo fisher, Rockford, IL; EMD Millipore Corporation, Billerica, MA). All membranes were pre-wetted in PBS at 4 °C for at least 10 min before assembly. Peptides were equilibrated with human α-thrombin at various concentrations for 1 hr at 4 °C in PBS buffer with 0.025% Tween-20 and 0.3% BSA and then loaded on to assembled dot blot apparatus. Vacuum was applied to facilitate filtration of unbound peptide. Membranes were exposed to phosphor imager overnight in order to quantify peptide retained on membranes. Image quant software was used to quantify dot blots and R software was used to analyze results (GE Healthcare Lifesciences, Piscataway, NJ).
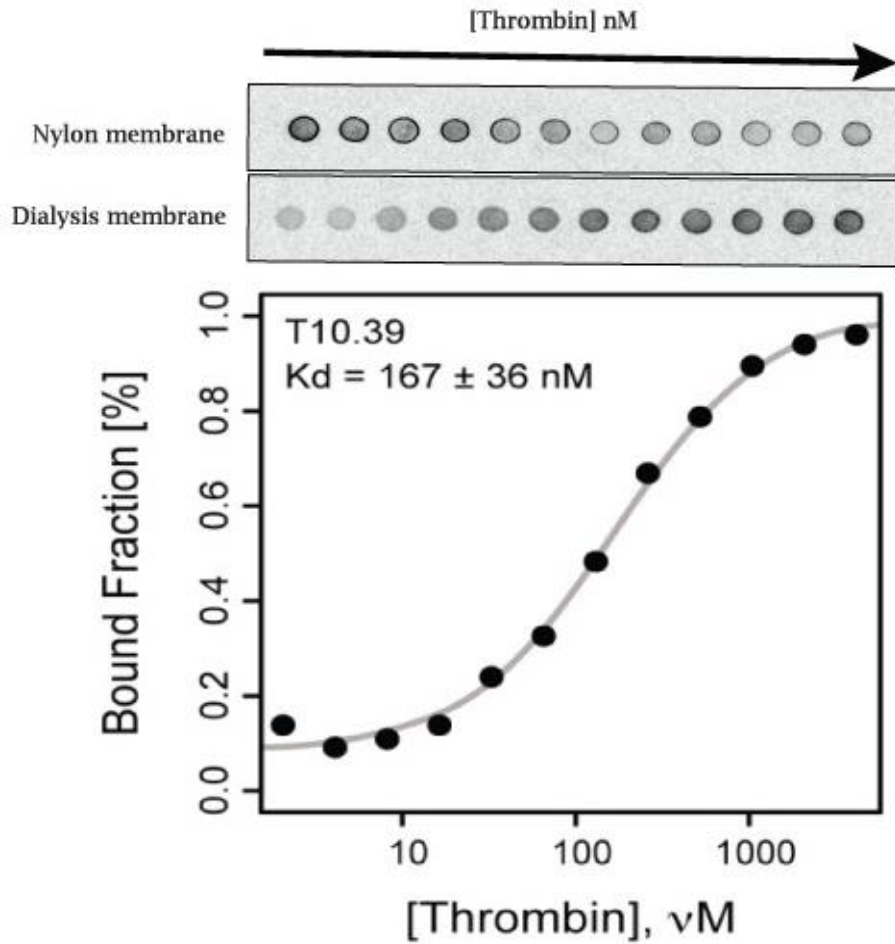
Fig. 4.2 Validation of dot blot double membrane system for determination of Kd of T10.39, a well characterized human α-thrombin binding peptide; a) phosphor image of dialysis membrane (top) and nylon membrane (bottom) b) Equilibrium dissociation plot of T10.39. Kd is the average of at least three independent replicates with standard deviations (19).

Results and Discussion

To validate our rationale of measuring peptide binding kinetics by size based separation, regenerated cellulose and nylon membrane were used to separate unbound peptide from bound peptide-protein complexes. We obtained solution binding affinity constant (Kd) for T10.39, a known thrombin binding peptide isolated from human α-thrombin mRNA display selection. Our results depict that T10.39 binds thrombin with Kd of 167 ± 36 nM which is consistent with reported Kd of 166 nM (11). T10.39 and T10.11 were the only two peptides characterized from human α-thrombin mRNA display selection output. Due to time and cost associated with peptide characterization most selection outputs remain uncharacterized thus bypassing many high affinity binders that might bind to discrete sites on target protein.

Above mentioned mRNA display selection was performed on starting library of 10^{11} different sequences. 45 potential binders were reported after 10 rounds of selection and only two peptides were characterized reporting a Kd of 166nM and 520 nM for human α-thrombin. To further validate power of our technology, we selected 24 peptides from set of 45 peptides reported in human α-thrombin mRNA display selection output to perform a parallel screen (11). DNA oligonucleotides encoding 24 peptides were cloned in expression vector and expressed as S^{35}-labelled peptides to perform equilibrium binding analysis as mentioned in Fig. 1.3. Three peptides, T10.35, T10.46 and T10.57, failed to show sufficient expression for dot blot analysis and hence were discarded. 21 other peptides were equilibrated with 250 nM and 500 nM human α-thrombin for 1 hr at 4°C with rotation. Equilibrated samples were spotted in dot blot apparatus assembled with regenerated cellulose and nylon membrane in presence of slow vacuum. In order to account for non-specific binding of peptide to membranes, we set up samples devoid of human α-thrombin as control. Analysis of our screen indicates that five peptides, T10.06, T10.13, T10.25, T10.30 and T10.37, show increased binding with increasing human α-thrombin concentration. We noticed some of the peptides with high hydrophobicity showed

up as weak binders, peptide solubility may also affect binding to target protein. Otherwise insoluble peptides might become soluble during mRNA display due to association with nucleic acids thus rendering them to selections. Also, insoluble peptides forming large aggregates may lead to high background during dot blot analysis.

We further characterized five potential thrombin binding peptides obtained from our screen using range of human α-thrombin concentrations, 5 μM - 1 nM. Peptides were incubated with various concentration of human α-thrombin for 1 hr at 4°C with rotation. Separation of bound peptide-protein complexes from unbound peptide resulted in a binding curve indicating Kd of 167 $\pm$ 36 nM for T10.39, 159 $\pm$ 41 nM for T10.13, 363 $\pm$ 104 nM for T10.30 and 313 $\pm$ 78 nM for T10.37. T10.39 was used as standard throughout our experiments and gave a Kd of 167 $\pm$ 36 nM consistent with literature. DPGR motif present in T10.39 is also present in T10.06, T10.25 and T10.30 peptides which bind human α-thrombin with Kd of 200 – 500 nM. This indicates importance of DPGR motif in human α-thrombin binding. Peptides T10.13 and T10.37 lacking DPGR element were also found to bind human α-thrombin with Kd of 159 nM and 313 nM, suggesting that they might bind to a different site on human α-thrombin protein surface. Discovery of peptides that bind to discrete locations on protein surface paves way to construction of bivalent affinity reagents, NuPromers, using LINC technology. Linking two affinity peptides on a flexible surface at appropriate distance and orientation has been reported to increase binding affinity up to 1000 fold as compared affinity of individual peptides by themselves (12, 13).

Peptides are excellent therapeutics owing to small size, ease of *in vivo* delivery and capability to bind specific surfaces with high affinity. Affinity peptides overcome the limitations of antibodies as protein capture reagents due to simple structure, non-immunogenicity, affordability, ease and time of synthesis as well as convenience of tailoring them to our needs. With increasing applications of peptides as therapeutics (14),

drugs (15), disease diagnostics, carriers for drugs (16), activity based probes for enzyme profiling and cell imaging (17, 18), there is increasing demand for high affinity peptides. We seek to fill this gap by proposing double membrane dot blot method for screening and characterization of peptides.
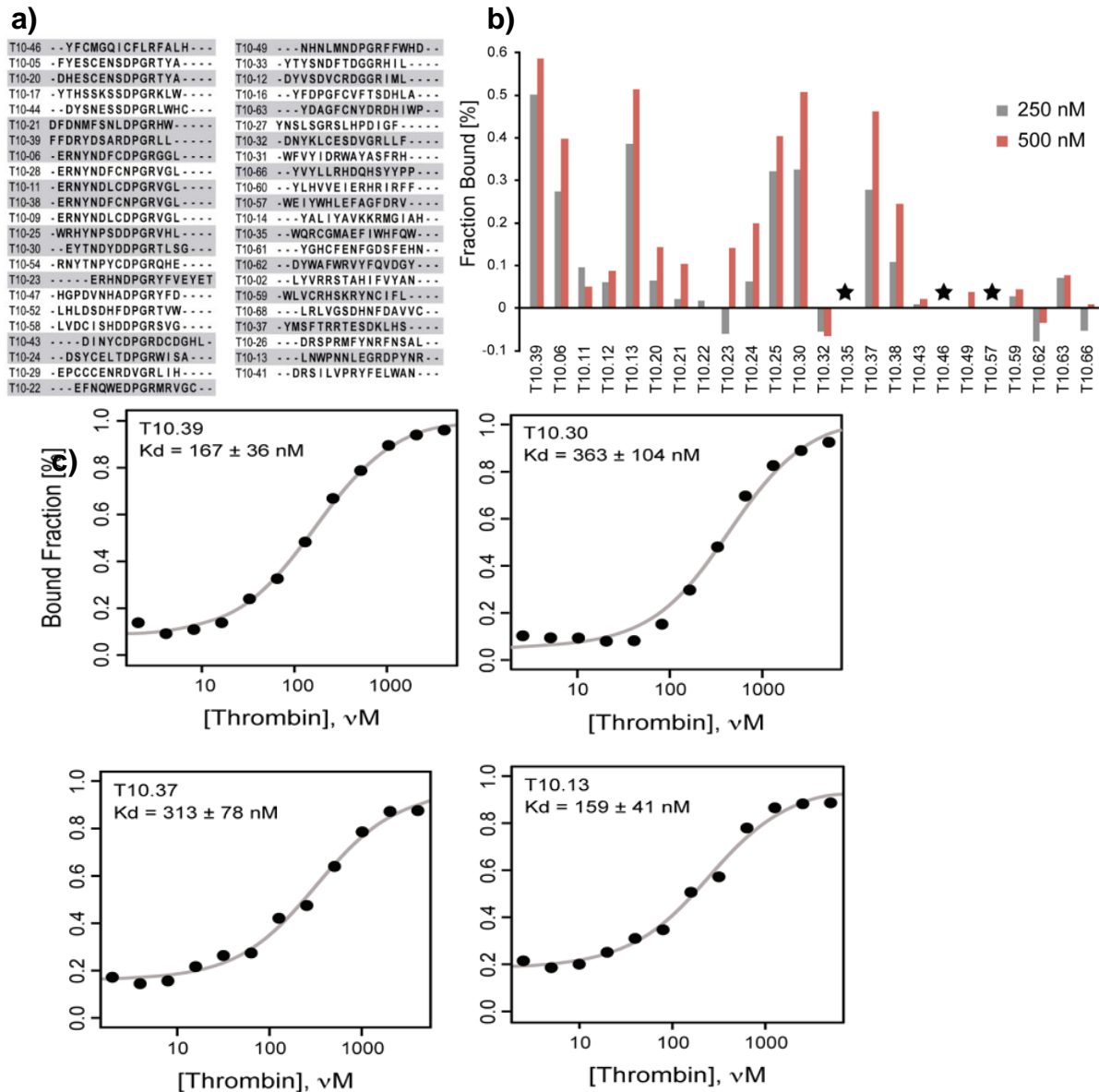
Fig. 4.3 Screening, identification and characterization of affinity peptides a) amino acid sequence of forty five peptides obtained from mRNA display selection (11). Amino acid sequences with grey highlight indicate peptides chosen for present study. **b)** double membrane based dot blot screen of twenty four peptides for affinity to human α-thrombin. Relative fraction of peptide bound to thrombin at 250 nM (grey bars) and 500 nM (red bars) were comparred to identify potential affinity peptides. Stars indicate peptides that had negligible expression insufficient for dot blot analysis. **c)** Equilibrium dissociation plots of four thrombin-binding peptides. Kd values represent atleast three independent replicates with standard deviations (19).

Conclusion

Coupling *in vitro* selections with double membrane dot blot approach provides an advantage of exploring larger sequence space, rapidly and cost effectively without much investment in expensive instruments. A single screen of dot blot can be multiplexed and completed in 3 days, including peptide expression and purification. Thus establishing double membrane dot blot as ideal method for peptide screening and characterization.

BIBLIOGRAPHY

1. **Ellington, A.D., and J.W. Szostak**. 1990. In vitro selection of RNA molecules that bind specific ligands. Nature *346*:818–822.

2. **Stoltenburg, R., C. Reinemann, and B. Strehlitz**. 2007. SELEX-A (r)evolutionary method to generate high-affinity nucleic acid ligands. Biomol. Eng. *24*:381–403.

3. **Yu, H., B. Jiang, and J.C. Chaput**. 2011. Aptamers can discriminate alkaline proteins with high specificity. Chembiochem *12*:2659–2666.

4. **Xu, D., and H. Shi**. 2009. Composite RNA aptamers as functional mimics of proteins. Nucleic Acids Res. *37*:e71.

5. **Tuerk, C., and L. Gold**. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*:505–510.

6. **Wittrup, K.D.** 2001. Protein engineering by cell-surface display. Curr. Opin. Biotechnol. *12*:395–399.

7. **Boder, E.T., and K.D. Wittrup**. 1997. Yeast surface display for screening combinatorial polypeptide libraries. Nat. Biotechnol. *15*:553–557.

8. **Hanes, J., and A. Plückthun**. 1997. In vitro selection and evolution of functional proteins by using ribosome display. Proc. Natl. Acad. Sci. U. S. A. *94*:4937–4942.

9. **Roberts, R.W., and J.W. Szostak**. 1997. RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc. Natl. Acad. Sci. U. S. A. *94*:12297–12302.

10. **Clackson, T., H.R. Hoogenboom, A.D. Griffiths, and G. Winter**. 1991. Making antibody fragments using phage display libraries. Nature *352*:624–628.

11. **Molek, P., B. Strukelj, and T. Bratkovic**. 2011. Peptide phage display as a tool for drug discovery: targeting membrane receptors. Molecules *16*:857–887.

12. **Sperhacke, R.D., F.C.Q. Mello, A. Zaha, A. Kritski, and M.L.R. Rossetti**. 2004. Detection of Mycobacterium tuberculosis by a polymerase chain reaction colorimetric dot-blot assay. Int. J. Tuberc. Lung Dis. *8*:312–317.

13. **Bishop, P.N., A.B. Tullo, R. Killough, and S.J. Richmond**. 1991. An immune dot-blot test for the diagnosis of ocular infection with Chlamydia trachomatis. Eye (Lond). *5 (Pt 3)*:305–308.

14. **He, M., O. Stoevesandt, and M.J. Taussig**. 2008. In situ synthesis of protein arrays. Curr. Opin. Biotechnol. *19*:4–9.

15. **Goshima, N., Y. Kawamura, A. Fukumoto, A. Miura, R. Honma, R. Satoh, A. Wakamatsu, J. Yamamoto, et al.** 2008. Human protein factory for converting the transcriptome into an in vitro – expressed proteome. *5*.

16. **Forster, A.C., Z. Tan, M.N.L. Nalam, H. Lin, H. Qu, V.W. Cornish, and S.C. Blacklow**. 2003. Programming peptidomimetic syntheses by translating genetic codes designed de novo. Proc. Natl. Acad. Sci. U. S. A. *100*:6353–6357.

17. **Josephson, K., M.C.T. Hartman, and J.W. Szostak**. 2005. Ribosomal synthesis of unnatural peptides. J. Am. Chem. Soc. *127*:11727–11735.

18. **Mattheakis, L.C., R.R. Bhatt, and W.J. Dower**. 1994. An in vitro polysome display system for identifying ligands from very large peptide libraries. Proc. Natl. Acad. Sci. U. S. A. *91*:9022–9026.

19. **Hospital, M.G.** 1997. RNA-peptide fusions for the in vitro selection of peptides. *94*:12297–12302.

20. **Katzen, F., T.C. Peterson, and W. Kudlicki**. 2009. Membrane protein expression: no cells required. Trends Biotechnol. *27*:455–460.

21. **Kaiser, L., J. Graveland-Bikker, D. Steuerwald, M. Vanberghem, K. Herlihy, and S. Zhang**. 2008. Efficient cell-free production of olfactory receptors: detergent optimization, structure, and ligand binding analyses. Proc. Natl. Acad. Sci. U. S. A. *105*:15726–15731.

22. **Junge, F., L.M. Luh, D. Proverbio, B. Schäfer, R. Abele, M. Beyermann, V. Dötsch, and F. Bernhard**. 2010. Modulation of G-protein coupled receptor sample quality by modified cell-free expression protocols: a case study of the human endothelin A receptor. J. Struct. Biol. *172*:94–106.

23. **Mikamo-Satoh, E., A. Takagi, H. Tanaka, T. Matsumoto, T. Nishihara, and T. Kawai**. 2009. Profiling of gene-dependent translational progress in cell-free protein synthesis by real-space imaging. Anal. Biochem. *394*:275–280.

24. **Spirin, a S., V.I. Baranov, L. a Ryabova, S.Y. Ovodov, and Y.B. Alakhov**. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. Science *242*:1162–1164.

25. **Mei, Q., C.K. Fredrickson, A. Simon, R. Khnouf, and Z.H. Fan**. 2007. Cell-free protein synthesis in microfluidic array devices. Biotechnol. Prog. *23*:1305–1311.

26. **Park, N., S.H. Um, H. Funabashi, J. Xu, and D. Luo**. 2009. A cell-free protein-producing gel. Nat. Mater. *8*:432–437.

27. **Mikami, S., T. Kobayashi, M. Masutani, S. Yokoyama, and H. Imataka**. 2008. A human cell-derived in vitro coupled transcription/translation system optimized for production of recombinant proteins. Protein Expr. Purif. *62*:190–198.

28. **Endoh, T., T. Kanai, Y.T. Sato, D. V Liu, K. Yoshikawa, H. Atomi, and T. Imanaka**. 2006. Cell-free protein synthesis at high temperatures using the lysate of a hyperthermophile. J. Biotechnol. *126*:186–195.

29. **Mikami, S., T. Kobayashi, S. Yokoyama, and H. Imataka**. 2006. A hybridoma-based in vitro translation system that efficiently synthesizes glycoproteins. J. Biotechnol. *127*:65–78.

30. **Shimizu, Y., a Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, and T. Ueda**. 2001. Cell-free translation reconstituted with purified components. Nat. Biotechnol. *19*:751–755.

31. **Zhou, Y., H. Asahara, E. a Gaucher, and S. Chong**. 2012. Reconstitution of translation from Thermus thermophilus reveals a minimal set of components sufficient for protein synthesis at high temperatures and functional conservation of modern and ancient translation components. Nucleic Acids Res. *40*:7932–7945.

32. **Corporation, P.** Rabbit Reticulocyte Lysate System.

33. **Corporation, P.** Wheat Germ Extract Technical Manual TM230.

34. **Contents, K., H. Lysate, A. Proteins, R. Mix, N. Water, and C. Color**. 1-Step Human Coupled IVT Kit – DNA. *0747*:1–6.

35. **Corporation, P.** 2008. Protein expression selecting the cell-free protein expression system that meets your experimental goals choosing a cell-free expression system rn. 6–9.

36. **Brouette, B.C., N. Betz, G. Kobs, and P. Corporation**. Optimized Gene Expression with the T7 Sample System.

37. **Betz, B.N., and D. Ph**. Characterization of T N T ® T7 Quick for PCR DNA. 19–22.

38. **Langlais, C., B. Guilleaume, N. Wermke, T. Scheuermann, L. Ebert, J. LaBaer, and B. Korn**. 2007. A systematic approach for testing expression of human full-length proteins in cell-free expression systems. BMC Biotechnol. *7*:64.

39. **Larsen, A.C., A. Gillig, P. Shah, S.P. Sau, K.E. Fenton, and J.C. Chaput**. 2014. A General Approach for Characterizing In Vitro Selected Peptides with Protein Binding Affinity. Anal. Chem.

40. 1975. © 1975 Nature Publishing Group.

41. **Uhlen, M.** 2009. Mapping the Human Proteome Using Antibodies. 2006–2007.

42. **Rhyner, C., M. Weichel, S. Flückiger, S. Hemmann, T. Kleber-Janke, and R. Crameri**. 2004. Cloning allergens via phage display. Methods *32*:212–218.

43. **Zahnd, C., P. Amstutz, and A. Plückthun**. 2007. Ribosome display : selecting and evolving proteins in vitro that specifically bind to a target. *4*:269–279.

44. **Takahashi, T.T., R.J. Austin, and R.W. Roberts**. 2003. mRNA display: Ligand discovery, interaction analysis and beyond. Trends Biochem. Sci. *28*:159–165.

45. **Zhang, J., B. a R. Williams, M.T. Nilsson, and J.C. Chaput**. 2010. The evolvability of lead peptides from small library screens. Chem. Commun. (Camb). *46*:7778–7780.

46. **Li, S., and R.W. Roberts**. 2003. A novel strategy for in vitro selection of peptide-drug conjugates. Chem. Biol. *10*:233–239.

47. **Xu, L., P. Aha, K. Gu, R.G. Kuimelis, M. Kurz, T. Lam, A.C. Lim, H. Liu, et al.** 2002. Directed evolution of high-affinity antibody mimics using mRNA display. Chem. Biol. *9*:933–942.

48. **Williams, B.A.R., C.W. Diehnelt, P. Belcher, M. Greving, N.W. Woodbury, S.A. Johnston, and J.C. Chaput**. 2009. Creating protein affinity reagents by combining peptide ligands on synthetic DNA scaffolds. J. Am. Chem. Soc. *131*:17233–17241.

49. **Manuscript, A., and H. Crosslinker**. 2011. NIH Public Access. 1–29.

50. **Liu, R., B. Jiang, H. Yu, and J.C. Chaput**. 2011. Generating DNA synbodies from previously discovered peptides. Chembiochem *12*:1813–1817.

51. **Datta, S.R., A. Brunet, M.E. Greenberg, S.R. Datta, A. Brunet, and M.E. Greenberg**. 1999. Cellular survival : a play in three Akts Cellular survival : a play in three Akts. 2905–2927.

52. **Wong, I., and T.M. Lohman**. 1993. A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. Proc. Natl. Acad. Sci. U. S. A. *90*:5428–5432.

53. **Keefe, A.D., D.S. Wilson, B. Seelig, and J.W. Szostak**. 2001. One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. Protein Expr. Purif. *23*:440–446.

54. **Kapust, R.B., and D.S. Waugh**. 2000. Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr. Purif. *19*:312–318.

55. **Raffler, N.A., J. Schneider-Mergener, and M. Famulok**. 2003. A novel class of small functional peptides that bind and inhibit human α-thrombin isolated by mRNA display. Chem. Biol. *10*:69–79.

56. **Williams, B.A.R., K. Lund, Y. Liu, H. Yan, and J.C. Chaput**. 2007. Self-assembled peptide nanoarrays: an approach to studying protein-protein interactions. Angew. Chem. Int. Ed. Engl. *46*:3051–3054.

57. **Petsalaki, E., A. Stark, E. García-Urdiales, and R.B. Russell**. 2009. Accurate prediction of peptide binding sites on protein surfaces. PLoS Comput. Biol. *5*.

58. **McGregor, D.P.** 2008. Discovering and improving novel peptide therapeutics. Curr. Opin. Pharmacol. *8*:616–619.

59. **Svensen, N., J.G.A. Walton, and M. Bradley**. 2012. Peptides for cell-selective drug delivery. Trends Pharmacol. Sci. *33*:186–192.

60. **Heal, W.P., T.H.T. Dang, and E.W. Tate**. 2011. Activity-based probes: discovering new biology and new drug targets. Chem. Soc. Rev. *40*:246–257.

61. **Reubi, J.C., and H.R. Maecke**. 2008. Peptide-based probes for cancer imaging. J. Nucl. Med. *49*:1735–1738.

62. **Esposito, C.L., S. Catuogno, V. de Franciscis, and L. Cerchia**. 2011. New insight into clinical development of nucleic acid aptamers. Discov. Med. *11*:487–496.

63. **Roberts, R.W.** 1999. Totally in vitro protein selection using mRNA-protein fusions and ribosome display. Curr. Opin. Chem. Biol. *3*:268–273.