

DISSERTATION

DEVELOPMENT OF A PHAGE-BASED DIAGNOSTIC SENSOR FOR ACTIVE
TUBERCULOSIS

Submitted by

Ning Zhao

Department of Chemical and Biological Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2016

Doctoral Committee:

Advisor: John D. Fisk

John S. Spencer
Charles S. Henry
Christie Peebles

Copyright by Ning Zhao 2016

All Rights Reserved

ABSTRACT

DEVELOPMENT OF A PHAGE-BASED DIAGNOSTIC SENSOR FOR ACTIVE TUBERCULOSIS

Antibodies, the quintessential biological recognition molecules, are not ideal for many applications because of their large size, complex modifications, and thermal and chemical instability. Identifying alternative scaffolds that can be evolved into tight, specific binding molecules with improved physical properties is of increasing interest, particularly for biomedical applications in resource-limited environments. Hyperthermophilic organisms, such as *Sulfolobus solfataricus*, are an attractive source of highly stable proteins as starting points for alternative molecular recognition scaffolds. We describe the first application of phage display to identify binding proteins based on the *Sulfolobus solfataricus* protein Sso7d scaffold. Sso7d is a small (approximately 7 kDa, 63 amino acids), cysteine free DNA-binding protein with a melting temperature of nearly 100 °C. Tight binding Sso7d variants were selected for a diverse set of protein targets from a 10¹⁰ member library, demonstrating the versatility of the scaffold. These Sso7d variants are able to discriminate among closely-related human, bovine, and rabbit serum albumins. Equilibrium dissociation constants in the nanomolar to low micromolar range were measured via competitive ELISA. Importantly, the Sso7d variants continue to bind their targets in the absence of the phage context. Furthermore, phage-displayed Sso7d variants retain their binding affinity after exposure to temperatures up to 70 °C. Taken together, our results suggest that the Sso7d scaffold will be a complementary addition to the range of non-antibody scaffold proteins

that can be utilized in phage display. Variants of hyperthermostable binding proteins have potential applications in diagnostics and therapeutics for environments with extreme conditions of storage and deployment. One application for utilizing Sso7d evolved binding molecules is development of Tuberculosis (TB) diagnostic tests. TB is the leading cause of death from infectious disease worldwide. The low sensitivity, extended processing time and high expense of diagnostics are major challenges to the detection and treatment of TB. *Mycobacterium tuberculosis* ornithine transcarbamylase (*Mtb* OTC, Rv1656) has been identified in the urine of patients with active TB infection, making *Mtb* OTC a promising target for point-of-care diagnostics in resource-limited settings. We are motivated to engineer phage-based diagnostic systems that feature improved physical stability, cost of production and sensitivity relative to traditional antibody-based reagents. Specific binding proteins with low nanomolar affinities for *Mtb* OTC were selected from the naïve Sso7d phage library. Phage particles displaying Sso7d variants along with a monoclonal antibody (mAb) generated by hybridoma technology were utilized to generate a capture ELISA-based assay for *Mtb* OTC. The ELISA assay signal is linear over the target concentration range of 2.0-125.0 ng/mL with limit of detection 0.4 ng/mL (12 pM), which is comparable to commercial available antibody-based assays. Importantly, this assay maintains functionality at both neutral and basic pH in presence of salt and urea over the range of concentrations typical for human urine. Furthermore, towards our phage-based diagnostic test development goal, a test with a pair of phage displaying 2 different Sso7d variants was established with a limit of detection 4.5 ng/mL (130 pM). Stability of TB diagnostic test is improved at acidic conditions in presence of salt and urea in the typical concentration range of human urine, which may due to the replacement of mAb with phage particles. This result demonstrates that phage particles replacing antibodies in the diagnostic test has the potential to improve stability at harsh conditions.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. John D. (Nick) Fisk for the continuous support and mentor of my Ph. D study, without whom none of my accomplishment would have been possible. Dr. Fisk's strong motivation to solve challenging scientific problems and his immerse knowledge inspired me move forward on this scientific pathway. I also would like to express my gratitude to all my committee members, Dr. John S. Spencer, Dr. Charles S. henry and Dr. Christie Peebles, who gave me valuable advices and assistance to my Ph. D research. As a collaborator, Dr. Spencer gave critical support in protein purifications and monoclonal antibody preparations. Without his encouragement, I cannot be so confident about my scientific skills.

I am grateful to the generous help from the Fisk lab manager and a great scientist, Dr. Margaret A. Schmitt. She not only gave me valuable advices in science, but also in life. Meg, thank you for lending your ears when I was in a tough time.

I want to thank the laboratory of Dr. P. Shing Ho for assistance, expertise, and access to the ITC and DSC instrumentation, Dr. Charles S. Henry for generous supplying protein samples, and Dr. Andrew Bradbury for a generous gift of plasmids.

I would like to thank all past and present coworkers in the laboratory of Dr. Fisk, especially, Wil Biddle and David Schwark. This work cannot be accomplished so smoothly without your generous assistance. I cannot forget our beer time and those chicken wing nights.

Finally, I want to thank my family and friends who have encouraged me during these years. Especially, I would like to give my special thanks to my husband and my parents for their patient support behind me to complete this work.

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgements.....	iv
Chapter 1 Introduction.....	1
1.1 Diagnostics	2
1.2 Antibody-based affinity diagnostic tests.....	3
1.2.1 Antibodies introduction.....	3
1.2.2 Antibody-based diagnostic tests.....	5
1.2.3 Problems and challenges for the antibody-based diagnostic tests applying for resource-limited areas.....	7
1.3 Phage-based diagnostic tests.....	7
1.3.1 Filamentous bacteriophage introduction.....	7
1.3.2 Phage display.....	12
1.3.3 Antibody-based scaffolds for phage display.....	15
1.3.4 Non-antibody-based scaffolds for phage display.....	17
1.3.5 Phage as a component of diagnostic tests.....	19
1.4 Tuberculosis	21
References.....	31
Chapter 2 Phage display selection of tight specific binding variants from a hyperthermostable Sso7d scaffold protein library.....	40
2.1 Introduction.....	41
2.2 Results and Discussion.....	46

2.3 Conclusions.....	57
2.4 Materials and methods.....	58
References.....	80
Chapter 3 Selection of nanomolar affinity binding proteins for active Tuberculosis diagnosis from an Sso7d alternative scaffold phage display library.....	86
Chapter 4 A double-phage based diagnostic test established from an Sso7d alternative scaffold phage display library for active Tuberculosis diagnosis.....	87
Chapter 5 Single chain variable fragment (scFv) based on a monoclonal antibody and displayed on phage as a promising diagnostic test for active Tuberculosis.....	89
Appendix 1 Table S2-1. List of oligonucleotide primer sequences in chapter 2.....	90
Appendix 2 Table S2-2. DNA sequences of Sso7d display features in chapter 2.....	91
Appendix 3 4-parameter logistic equation.....	92
Appendix 4 Copyright agreement from FEBS J.	93

Chapter 1

Introduction

This thesis contains 5 chapters that describe the development of a platform for detecting biological molecules based on employing highly modified bacteriophage as replacements for antibodies. The first chapter provides a framework for understanding the relevance of the work described in subsequent chapters and includes a brief introduction to diagnostic tests, a more focused discussion of the use of antibodies in immunological tests and some drawbacks and alternatives to antibodies as molecular components employed in diagnostics. A particular focus of the thesis is the development of diagnostics for resource limited environments and a discussion of the particular challenges and requirements for point-of-care diagnostics in the developing work is presented. The governing idea behind all of the experiments described in the thesis is that bacteriophages have all of the desired qualities of detecting molecules for point-of-care diagnostics. A description of the filamentous bacteriophages, the use of filamentous bacteriophages to select binding molecules through phage display, and the use of bacteriophages in diagnostics is presented. The second chapter describes the development of an alternative scaffold based on a small protein Sso7d from the thermophilic organism *Sulfolobus solfataricus* for binding molecule evolution, the construction of a phage displayed library of Sso7d variants and characterization of selected Sso7d variants that bind a diverse set of target proteins. The third chapter describes the use of the Sso7d phage library to select binding molecules that recognize a protein marker of active tuberculosis infection, *Mycobacterium tuberculosis* ornithine transcarbamylase (*Mtb* OTC), and the development of a diagnostic test employing phages and antibodies. The fourth chapter describes the development of a diagnostic test that employs predominantly modified phages for the detection

of *Mtb* OTC. The final chapter describes the production of a monoclonal antibody for the *Mtb* OTC protein.

1.1 Diagnostics

Diagnosis, the critical first step in the treatment of disease, often relies on the detection of specific biochemical markers. Approximately 7 billion medical tests are performed annually in the US to diagnose as well as monitor the progression of diseases and treatments. Medical diagnostics are often characterized by the type of molecules or cells they interrogate or the specific technologies employed in making measurements. General chemical tests measure the compositions of bodily fluids (general protein levels in urine or Ca^{2+} levels in blood). Cytology tests monitor the types and properties of cells in a sample, most commonly blood tests count and size red and white blood cells, biopsies evaluate cellular morphology to detect cancerous cells. Microbiological screens are employed to detect disease causing bacteria and fungi by growing cultures from isolated sampled tissues. The most specific and sensitive biological tests measure the levels of individual proteins or nucleic acids that are indicative of particular disease states. Nucleic acid tests are extremely sensitive and specific due to amplification through polymerase chain reaction (PCR), and are used to detect infections, genetic markers of disease, and measure levels of viruses like human immunodeficiency virus (HIV). Nucleic acid tests generally provide information about the genetic background but not the specific biochemical state of an individual. Tests that measure the levels or functions of specific proteins are the most direct measurements of biochemical functioning of a patient. Proteins levels are typically measured in immunohistochemical assays.

Although new instruments and tests that rely on multiple different physical principles are constantly appearing, antibody-antigen interactions, the underlying biological recognition events

used in many of these systems, have not changed greatly in the past 40 years. Typically, antibodies are generated using whole animal systems and produced in mammalian cell cultures. The process of generating antibodies is time consuming, expensive, and in the end gives only limited amounts of material. Additional steps are necessary if further genetic modifications and optimizations are desired. Beyond these difficulties, the physical properties of whole antibodies are not ideal as the components of diagnostics. Antibodies are not physically stable; they have limited process ability in diagnostics manufacture and need to be stored at low temperatures. Even then, antibody-based reagents have limited shelf lives. These drawbacks are critical in the developing world where diagnostic laboratories and even basic infrastructure are lacking.

1.2 Antibody-based affinity diagnostic tests

1.2.1 Antibodies introduction

Antibodies, also called immunoglobulins, are antigen binding proteins that are secreted by B-cells. The basic structure of an antibody molecule is a “Y” shape, including two identical heavy (H) chains and two identical light (L) chains, which are connected by disulfide bonds (Figure 1-1). Both constant (C_H and C_L) and variable regions (V_H and V_L) exist in the heavy and light chains. The variable regions play a critical role in the binding of an antibody to an antigen, and the constant regions function as the supporting structure for the molecule. A comparison of the amino acid sequences of the V_L and V_H domains reveals that sequence variability is concentrated in three regions, named hypervariable regions, forming an antigen binding site. The hypervariable regions are also known as complementarity determining regions (CDRs) because the antigen binding site is complementary to the structure of an epitope which is the immunologically active region of an antigen. The remainder of the V_H and V_L regions exhibits far less variability than the CDRs and is termed relatively constant regions. These relatively constant regions form the β pleated sheet

structure of the V_H and V_L regions, and the six CDRs are located on the loops connecting the β strands. Although most antibodies are composed of two heavy and two light chains, both of which contribute to the antigen-binding sites, a heavy-chain antibody lacking light chains naturally exists in camelid¹ and shark², in which the antigen-binding site is formed by a single domain. The CDR3 region of the heavy-chain camelid antibodies, compared with the CDR3 region of the conventional V_H domain, possesses the capacity to form extended loops that can fit into cavities on antigens.

The immunoglobulins can be classified into five major classes or isotypes: IgA, IgD, IgE, IgG, and IgM based on the differences among the heavy chain constant regions. IgD, IgE, and IgG antibodies function as monomers, while IgA functions as a dimer and IgM functions as a pentamer. IgG is the most abundant antibody in human serum and the most common tool for immunology research. Antibodies can be further classified as polyclonal antibodies (pAb) and monoclonal antibodies (mAb). pAb can recognize multiple epitopes since it is a pool of multiple mAbs. Monoclonal antibodies are specific to a single epitope and are widely used in immunology assay applications.

Antibodies secreted from plasma cells continuously circulate in blood, recognizing or neutralizing antigens. The interaction between an antibody and an antigen is highly specific. The high specificity of antibodies led to the development of a variety of immunologic assays, including enzyme-linked immunosorbent assays (ELISA) and Western blots. These assays have been used to detect the presence of either antibody or antigen, which is important in disease diagnosis, mammal immune system response or other medical interests. mAbs have become the dominant form of new major biological pharmaceuticals with multibillion dollars of annual sales. Since the first therapeutic mAb was commercialized in 1986, forty-seven mAb drugs have been approved in the US or Europe for the treatment of various diseases³, such as Trastuzumab and Bevacizumab

for cancer therapy⁴, Rituximab for autoimmune disease treatment⁵, Alemtuzumab for transplant rejection prevention⁶.

1.2.2 Antibody-based diagnostic tests

Antibody-based diagnostic tests are usually designed based on the tight and specific interaction of the antibody and the biomarker to diagnose diseases. The most commonly used platform is an enzyme-linked immunosorbent assay (ELISA). Diagnostic tests usually employ some kind of complex body fluid, such as sputum, urine, or blood; the specific biomarker needs to be selected from the other components of the sample. Sandwich ELISA, which uses a pair of antibodies binding to different epitopes of the same biomarker, is the most commonly utilized platform. In a sandwich ELISA, an immobilized antibody captures the biomarker from the sample matrix, and then a second antibody binds to the captured biomarker. To develop a colorimetric or fluorescent signal, a third antibody conjugated with an enzyme, such as alkaline phosphatase (AP) or horseradish peroxidase (HRP), binds to the Fc region of the second antibody. Sandwich ELISA tests are often used in a qualitative (yes or no) mode to determine if a biomarker is present. Sandwich ELISA tests can also be quantitative, if a standard curve is made, the ELISA signal can be converted into a measure of the amount of analyte present.

Lateral flow assay (LFA) diagnostic tests are another very common design⁷. LFA tests are similar to the sandwich ELISAs in that a pair of antibodies with minimum binding competition to a same biomarker are needed. A detection antibody conjugated with an enzyme or colored particles⁷⁻¹⁰ binds the biomarker in a sample of some bodily fluid which is applied at one end of a paper strip. The sample and reagents flow through the paper strip towards the other end carrying the biomarker. A capture antibody is immobilized in a defined detection zone of the paper strip, captures the biomarker attached antibody/particle/enzyme conjugate as it flows across the

detection zone. The accumulation of the biomarker in the capture zone concentrates the enzyme-based signal generator or the accumulation of the particles causes a colorimetric band to show up when the biomarker is present. Unlike sandwich ELISA, LFAs only make qualitative diagnoses. The best-known lateral flow sensor is common pregnancy test⁷. Pregnancy tests detect the hormone human chorionic gonadotropin (hcg) produced during pregnancy in urine. Pregnancy test is a model for the desired performance of a point-of-care diagnostic test. The test can be performed without laboratory facilities and produces an easily readable colorimetric results within 10-20 min. The current generation of LFA test has high sensitivity and specificity due to integrating thin-layer chromatography, the identification of biomarker-specific antibodies and by labeling sensitive signaling component to LFA.¹¹

Dipstick assays are based on the immunoblotting principle.^{12,13} In dipstick assay, a paper or a membrane is usually used as an immobilization platform. However, it is different from LFA, not relying on lateral fluid flow through a membrane. In dipstick assay, a paper strip coated with a capture antibody is serially dipped into a multiple solution. Initially a blocking solution, then the diagnostic sample solution, then a solution containing a detection antibody conjugated with an enzyme or particles, and finally a substrate solution if needed. Multiple washes between each step of dipping are also needed. A blot will show the presence of biomarkers. The dipstick assay process is simple and easy, suitable for point-of-care diagnosis in resource-poor countries.

In addition to the above three antibody-based tests, many other combinations of antibodies and physical and chemical mechanisms of signal generation are applied as high sensitivity and specificity diagnostic tests. However, majority of those designs need laboratory facilities which are inaccessible in resource-limited areas.

1.2.3 Problems and challenges for the antibody-based diagnostic tests applying for resource-limited areas

A test for diagnosing diseases in resource-limited areas should be sensitive, specific, as well as affordable and suitable for point-of-care application. Point-of-care usage means the tests do not need laboratory facilities, ideally are laboratory-free, and do not need specially trained staff for the assay. Further, the physical components should be conveniently delivered to and stored in areas without refrigeration or places without reliable power. The ELISA, LFA and dipstick tests mentioned before are all excellent candidate diagnostic test designs for resource-limited areas applications that suffer due to some issues caused by antibodies. Antibodies usually are not stable to elevated temperature and need to be carefully stored at cold conditions, which constrains the delivery and storage of the test reagents in resource-limited areas. Additionally, the generation, production and purification processes of antibodies significantly extend the time for the assay development and increase the cost of the diagnostic tests. In undeveloped areas, like some areas in Africa, even \$1 per assay is too expensive. Cheaper is better and ideal situation is no cost. To solve these potential problems of antibodies, for a resource-limit areas diagnostic test development, other binding components need to be used, such as M13 phage particles.

1.3 Phage-based diagnostic tests

1.3.1 Filamentous bacteriophage introduction

The filamentous bacteriophages are a family of rod shaped viruses packaging a single-stranded DNA (ssDNA) genome capable of infecting a wide variety of gram-negative bacteria, including *Escherichia coli*, *Xanthomonas*, *Thermus*, *Pseudomonas*, *Salmonella* and *Vibrio* ¹⁴. The most well characterized filamentous phages are the Ff class. The Ff phages are so named because they enter the host bacterial cell via the tip of the F conjugative pilus on the surface of male (F⁺) *E. coli* cells.

The Ff phage family contains the phages M13, fl and fd, all of which show 98% homologous DNA sequences and nearly identical protein sequences.¹⁵⁻¹⁷

Filamentous phages are made and secreted from the host *E. coli* cells without cell killing or lysis. Therefore, the phages can accumulate to titers as high as 10^{11} to 10^{12} particles per mL along with propagation of the host *E. coli* cells. The length of engineered Ff phage depends on the length of packaged DNA. So far, the only known constraints on the length of packaged DNA results from shearing of assembled phages. After infection, the viral single-stranded DNA is converted to a double-stranded intermediate form during the replication process. Both the ssDNA and the dsDNA forms can be isolated alternatively from phage particles or infected cells. Ff phages found an early biotechnological application as a cloning vehicle to obtain ssDNA for sequencing. The Ff phage is tolerant to insertions of foreign genes into non-essential regions. A gene of interest can additionally be fused with one of the coat protein genes. The resulting phage particles display the encoded protein on the particle surface and containing the genes sequence packaged within the particles, providing a direct link between genotype and phenotype. The process of infection by phage, can be used to conveniently transfer genetic information between bacterial stains. The tolerance of foreign genes insertions, the direct linkage of phage phenotype and genotype, the possibility of manipulating both ssDNA and dsDNA, and the easy genome transfer between bacterial strains combine to make Ff phage, primarily M13, a convenient and versatile molecular biology tool.

1.3.1.1 Structure of the phage

The wild type Ff phages are approximately 6.5nm in diameter and 800 to 900nm in length.¹⁸ Five coat proteins, p3, p6, p7, p8, and p9, form a somewhat flexible cylinder encasing a single-stranded, covalently closed DNA phage genome of 6400 nucleotide (Figure 1-2). Approximately

2700 copies of the major coat protein, p8, form the cylindrical body of each wildtype M13 bacteriophage. The other coat proteins are termed minor coat proteins since only five copies of each are present in each phage particle. These minor proteins comprise the two ends of a phage particle: p3 and p6 form one end, while p7 and p9 form the other. The particle weight of wildtype M13 phage is approximately 16.3MD and 87% of the mass is contributed by coat proteins.¹⁸

The cylinder formed by major coat protein p8 comprises the largest component of the phage structure. The length of the cylinder depends on length of genome packaged in the particle. For wildtype M13 phage, approximately 2700 copies of p8 form the cylinder. Microphages can be constructed by packaging a 221 nucleotide single-stranded circle of DNA which contain only 95 copies of p8.¹⁹ A p8 monomer is a 50-residue protein and forms α -helix structure except in the N-terminal 5 residues. p8 proteins are arranged in an overlapping format with a symmetry defined as a 5-fold rotation axis and a 2-fold screw axis with a pitch of about 32Å.²⁰ The p8 monomer α -helices are uniformly parallel to the particle axis and tilted by an inclination angle in the range of 13° and 20° with respect to the filament axis as determined by polarized Raman spectroscopy.²¹

1.3.1.2 Phage genome and gene expressed proteins

Complete sequences of the genomes of Ff phage, F1, M13 and Fd are known and they encode nearly identical protein sequences.¹⁵⁻¹⁷ Each genome encodes 11 proteins involved in the phage life cycle (Table 1-1). The 11 protein-coding genes are grouped in the genome according to their function.¹⁸ g2, g5, and g10 encode proteins required for the replication of phage genome. g3, g6, g7, g8, and g9 encode the coat proteins. g1, g11, and g4 encode proteins involved in the membrane-associated assembly process. The phage genome is tightly packed containing overlapping genes except a short sequence, intergenic region, between g2 and g4. The intergenic region does not encode for any protein but contains the phage origins of DNA replication and a packaging signal

(PS). The intergenic region is tolerant to foreign gene insertion such as antibiotic resistance gene²². The phage origin of replication binds the phage DNA replication protein which initiates the phage genome replication process. The PS is a 78-nucleotide hairpin with function in phage assembly initiation.

All five phage coat proteins are synthesized in cytoplasm and inserted into cytoplasmic membrane of the host *E. coli* cell. The major protein, p8, is synthesized as a fusion with a 23 amino acid N-terminal signal peptide directing the insertion of the protein into cytoplasmic membrane. The signal peptide is removed by a specific peptidase after insertion, resulting in a 50-residue mature p8 with its N-terminal 20 amino acids located in the periplasm and its C-terminal 10 amino acids exposed in the cytoplasm.²³⁻²⁵ The minor protein p3 is also inserted in the membrane with the help of an N-terminal signal peptide that is removed after insertion, resulting in the C-terminal residues of mature p3 in the cytoplasm and the N-terminal residues exposed to the periplasm.²⁶ The other minor coat proteins, p6, p7, and p9, are synthesized without a signal peptide, and the mechanism of their insertion into the membrane is unknown.

The proteins involved in the assembly process, p1, p11, and p4, all reside in the cytoplasmic or outer membrane. Protein p4 is synthesized with a 21-residue N-terminal signal peptide and is suggested to be translocated to the periplasm via the Sec secretion system of the host cell.^{27,28} p1 and p11, are inserted into the cytoplasmic membrane by the SecA pathway.²⁸ Multiple copies of these assembly proteins cooperate to form a channel through the cytoplasmic and outer membrane allowing the phage particles to be released from the host cell without lysis.¹⁸ The Ff phage infection is not fatal and allows the host cells to produce phage particles and continue to grow. Proteins p2, p5, and p10 function in the replication process, and these 3 proteins all reside in the cytoplasm where the replication process occurs.

1.3.1.3 Life cycle of the phage

The Ff bacteriophage life cycle can be generalized in three major steps: infection, replication, and assembly.

1.3.1.3.1 Infection

Phage infection is a complex process requiring interaction of the particle with the F conjugative pilus of the host cell and the bacterial cytoplasmic membrane proteins TolQ, TolR, and TolA. Although the complete details of process are not fully understood, the main phage protein responsible for infection is p3. p3 contains three domains connected by flexible glycine rich linkers named N1, N2, and CT.²⁹ The N1 and N2 domains play an important role in the phage infection process, and removal of these 2 domains renders phage particles noninfective.^{30,31} The N2 domain binds to the F conjugative pilus at the initiation of infection.^{29,32} The phage genome is then translocated into the cytoplasm, facilitated by the N1 domain of p3 binding to domain 3 of the TolA (TolA-D3) protein.³³ The process by which the coat proteins are disassembled in the cytoplasmic membrane and the exact function for other coat proteins in the infection process are unclear.

1.3.1.3.2 Replication

Once the single-stranded phage genome denoted as (+) strand enters the cytoplasm, the complementary (-) strand is synthesized by bacterial enzymes, and the parental replicative double stranded form (RF) DNA is produced. The (-) strand serves as a template for transcription, and the synthesized messenger RNAs are translated into the 11 phage proteins. The (+) strand is replicated with the help of phage protein p2 and bacterial host enzymes. Early in infection newly synthesized (+) strands are converted to additional RF DNA with newly synthesized complementary (-) strands. Through this process, the pool of RF DNA is produced, leading to a buildup of the phage proteins.

However, the accumulation of viral ssDNA and phage proteins does not increase infinitely. Two regulation proteins, p5 and p10 serve to balance the production of viral ssDNA and the phage proteins. p5 regulates the synthesis of RF DNA by binding to the newly synthesized (+) strands (p5-DNA complex), preventing conversion of (+) strands to RF DNA. At a critical concentration of p5, RF DNA synthesis is terminated. p10 is also involved in the replication of phage DNA and appears to be an inhibitor of p2, although its exact role is unclear.³⁴

1.3.1.3.3 Assembly

Phage particle assembly involves the passage of the p5-DNA complex across the bacterial membranes with the combination removal of p5v and addition of p8 molecules. During assembly, p5 is removed from the p5-DNA complex, and the coat proteins are assembled around the ssDNA. The assembly process contains three steps: initiation, elongation, and termination. Assembly is initiated by the DNA packaging signal that interacts with p1, p7, p9, and the first few copies of p8.^{35,36} The p7/p9 end is the first part of the particle to be extruded from the assembly site, which is a channel between the cytoplasmic and the outer membranes formed by p1, p11, and p4. The elongation process continues by replacing p5 with additional copies of p8 around the ssDNA. Assembly is terminated near the end of the ssDNA by the addition of p3 and p6 to the end of the particle. Release of the phage particles from the host cell membrane requires the CT domain of p3.³⁷ If the CT domain of p3 is absent, the phage particle already protruding from the membrane will continue packaging multiple ssDNAs and will not be released from the cell membrane.

1.3.2 Phage display

Phage display is a biotechnology that primarily utilizes filamentous phage.³⁸ Phage display involves the generation of large combinatorial libraries of peptide, protein or antibody sequences displayed on the surface of a phage particles that encode the sequence of the displayed protein.

Phage display is not only widely used to select binding molecules like antibody fragments, peptides and other scaffolds but also has been used to map interacting surfaces or discover protease recognition sites.

Foreign proteins are presented on the surface by fusing the foreign protein gene to the gene for one of the phage coat proteins. The resulting phage particles display the encoded protein on the particle surface, providing a direct link between phenotype and genotype. The coat proteins p3 and p8 are most commonly used in phage display although display on each of the five coat proteins has been demonstrated.³⁹⁻⁴² Proteins displayed on the N-terminus of p3 or p8 stretch out from the phage particle surface, enabling access to binding partners. Large fusion proteins on p3 appear to be packaged into phage particles reasonably well, which is a distinct advantage for p3 display over display from the rest of the coat proteins. However, p3 display of large proteins may lower or even eliminate the infectivity of the phage particles.¹⁸ In order to overcome this limitation, monovalent p3 display is usually applied by utilizing phagemid/helper phage system which will be described later. A Maximum of five copies of a displayed molecule per phage particle can be possible employing p3 display.

While p3 display enables from one to five copies of a fusion protein to be displayed per particle, p8 has the potential to enable approximately 2700 fusion proteins to be displayed on the surface of each M13 phage particle. When the displayed peptide is smaller than ~8 amino acids, all copies of p8 can display the peptide of interest.⁴³ Larger peptides and proteins (>10 residues) are generally inefficiently displayed on multiple copies of p8. In particles where large fusions are made with p8, >99% of the p8 proteins tend to be wild type.⁴⁴⁻⁴⁶ The main reason behind the p8 display size limitations is the assembly process of the phage. One thought is that when a large protein is displayed on p8, the assembled phage particle cannot pass through the 7nm exit hole formed by

p1, p11, and p4.^{47,48} Additionally, protein size may interfere with the ability of p8 to pack properly for stable particle formation. Recently, efforts have been devoted to increase the level of large protein display on p8. The display level of human growth hormone on p8 was increased 100-fold by selecting the optimized p8 from 3 randomly mutated p8 libraries that span the N-terminal 30 amino acids of the wild type p8.⁴⁹

Two types of vectors are frequently used for phage display: phage and phagemid (Figure 1-3). Phage vectors typically employ a single copy of each coat protein and the nucleotide sequence of peptide to be displayed is inserted as a fusion to the coat protein. All copies of the coat protein used for display will be a fusion instead of wild type (Figure 1-3). This type of display is termed polyvalent display and is useful for selecting low affinity binders from a library. Phage particles produced in phage vectors are homogeneous both phenotypically and genotypically. However, the phage vectors are difficult to manipulate at the DNA level due to the large genome size and the overlapping DNA sequences of the 11 phage proteins.

Monovalent phage display (Figure 1-3), in which on average only one copy of the coat protein displays the inserted peptide or protein, is accomplished using a combination of a helper phage and a second type of vector, a phagemid. Phagemids are relatively small vectors that contain the phage origin of replication, the packaging signal (PS) and, usually, one of the phage coat proteins. A phagemid vector is easier to manipulate at the DNA level compared to the full phage genome due to its smaller size and simple DNA structure. A phagemid vector needs another “helper” vector, a helper phage, to supply the other proteins needed for phage production (Figure 1-3). A helper phage contains a defective PS and therefore when a phagemid is present, phage particles prefer to package the phagemid genes with an optimal PS, but a small portion (usually 1-10%) of the phage particles still packages the helper phage genome. Fusion p3 expressed in the phagemid and

wildtype p3 from the helper phage are randomly assembled into phage particles. As a result, phage particles produced in a phagemid/helper phage system are both phenotypically and genotypically heterogeneous. This type of display has two advantages: (i) since wild type p3 is always present on the particles, the infectivity will not be greatly affected; (ii) this type of display is better suited for selection of high affinity binding molecules because avidity effects are eliminated.³⁸

For phagemid display, a helper phage is essential for the replication and assembly of the phage particles. However, a helper phage can cause contamination because some of the phage particles still package the genome of the helper phage. Depending on a variety of factors related to the particular system, the number of phage particles packaging the helper phage genome can sometimes equal, or exceed, the number of those packaging the phagemid.⁵⁰ The Bradbury lab developed M13-based helper plasmids to eliminate the need for a helper phage.⁵⁰ These helper plasmids contain all the genes of the supplementary phage proteins but do not contain the packaging signal or phage origin of replication, meaning no ssDNA is produced. The helper plasmids perform the same function as a helper phage without their ssDNA being packaged. As a result, genetically pure phagemid particles are produced. Furthermore, by using helper plasmids with different versions of p3: full-length p3, truncated p3, or deleted p3, phage particles produced by a phagemid vector are either monovalent or polyvalent (Figure 1-4).

1.3.3 Antibody-based scaffolds for phage display

In molecular biology and medicine, antibodies have traditionally been used as specific binding molecules for diagnostics and therapeutics. Antibodies are large multidomain proteins, on the order of 150 kDa, which require dimerization and formation of key disulfide bonds to function properly. The antigen-binding elements of antibodies, complementarity determining regions (CDRs), consist of closely-packed loop structures connecting the β -strands in the immunoglobulin

fold. The majority of the amino acid diversity in antibodies is focused in the CDRs of the heavy and light chain variable domains. The remaining “scaffold” of the immunoglobulin domain contains only minor variability and serves to present the CDRs in a three dimensional context. In contrast to the 150 kDa size of antibodies, p3, the largest M13 coat protein, is only 42 kDa. Displaying the intact antibody on phage surface is challenging.

Two recombinant forms of antibodies with reduced size are commonly employed in phage display: the fragment antigen binding^{51,52} (Fab) and single-chain fragment variable^{53,54} (scFv) scaffolds (Figure 1-5). Both Fabs and scFvs represent considerable reductions in size relative to an entire antibody, 50 kDa and 25 kDa, respectively. The Fab format consists of two chains which are expressed independently, one chain as fusion to an M13 coat protein. The other chain must be assembled with the phage after expression. The scFv is an antibody fragment with a covalent linkage connecting the variable regions of the heavy chain and light chain. scFvs can be expressed as a fusion to one of the phage coat proteins from a single gene. Another interesting antibody fragment with an even smaller size is a single domain antibody⁵⁵ (sdAb) or nanobody with a molecular weight in the 12~15kDa range. sdAbs contain only one domain, an either variable domain or engineered constant domain of a typical antibody. One of the most widely used sdAb is VHH, which contains one heavy-chain variable domain, engineered from heavy-chain antibodies found in camelids.⁵⁶ VHH nanobodies have high affinity and specificity as well as stability, small size, high solubility and benefit of multiple re-formatting opportunities. For phage display, an sdAb gene is attached to one coat protein and expressed as a fusion.

Those antibody fragments with reduced size are demonstrated to have not only same antigen-binding properties as the intact antibodies, but also advantages in tumor penetration, drug delivery, plasma clearance, conjugation with imaging molecules as well as for phage display than the intact

antibody.^{57,58} Antibody fragments commonly service as library scaffolds for the evolution of binding molecules via phage display^{56,59,60}. However, the physical properties of antibody fragments are less than ideal for many applications, especially in point-of-care diagnostic applications with elevated temperature, extremely pH conditions, high ionic strength occurring. Additionally, Fabs and scFvs show a tendency toward aggregation, causing problem in many applications.⁶¹ A perfect scaffold that could be universally applied to all therapeutic and diagnostic applications does not exist. To overcome the potential limitations of the antibody fragments, a continuing search for non-antibody-based protein scaffolds that can be evolved into specific binding molecules is ongoing.

1.3.4 Non-antibody-based scaffolds for phage display

Although antibody-based scaffolds are the dominant class of engineered proteins for molecular recognition in therapeutic and diagnostic applications, antibody fragments have some limitations, such as large size, instability at high temperature or in harsh chemical conditions, dependence on intrachain disulfide bonds. Non-antibody-based scaffolds or alternative scaffolds are of increasing interest because *in vitro* selection techniques enable tight binding variants of alternative proteins to be selected that possess the improved properties over antibodies, such as small size, high stability, absence of cysteines, high yield bacterial expression and compatibility with multiple detection methods as well as maintain the high affinity and specificity. Multiple proteins have been identified to have the potential to be used as alternative scaffolds by phage display. Some alternative scaffolds, such as fibronectin domains (monobodies)⁶², mimic the loops structure of antibodies to form the antigen-binding sites. Other scaffolds, such as Designed ankyrin repeat proteins (DARPin)s⁶³, small helical proteins (affibodies)⁶⁴, form a flat binding surface instead of mimicking the CDR structure of antibodies.

Fibronectin is a large protein essential in the formation of the extracellular matrix and cell-cell interactions and consists of many repeats of three types (I, II and III) domains. Fibronectin type III (FN3) is a small (94 amino acid residues, 10kDa) monomeric protein domain consisting seven β -sheets and has been used as a phage library scaffold by randomizing the loops connecting β sheets analogous to the CDR 1, 2, 3 structure in V_H of antibodies⁶². FN3 scaffold has been used for evolution of binding molecules against multiple protein targets in pM binding affinity.^{62,65,66}

DARPin are tightly packed repeat proteins.^{67,68} Each repeat module contains 33 amino acid residues and forms a structural unit consisting of a β -turn followed by two antiparallel α -helices. DARPins usually consist of four to six repeat modules and form a groove-like binding surface.⁶⁹ The molecular mass is about 14kDa for four-repeat DARPins. In each repeat module, six surface-exposed residues are potentially engaged in the interactions with antigens and randomized to construct a library. The rest residues are critical for maintaining the stable structure, and therefore kept conserved. DARPins have been display on phage via the cotranslational signal recognition particle (SRP) dependent system⁷⁰ instead of the traditional posttranslational Sec system. This is because DARPins fold very fast in cytoplasm. And the twin arginine transport (TAT) system⁷¹, a third transport system in *E. coli* which has been successfully used for super-folded green fluorescent protein (sfGFP) display on p3⁴⁰, was not functional for display of DARPins on p3⁷². From DARPins phage library, binding molecules have been selected against a wide targets, such as HER2⁷³, caspase-2⁷⁴, MAP kinase⁷⁵, aminoglycoside phosphotransferase⁷⁶, with affinity in the range of nM to pM.

Affibodies were originally engineered from a three-helix-bundle B domain in the immunoglobulin-binding region of *staphylococcus aureus* protein A. The B-domain was engineered, denoted as the Z-domain, by mutating a single site in helix 2 to improve chemical

stability, quench the binding affinity to Fab but retain the binding affinity to Fc region.⁷⁷ The affibody is composed with a 58-amino acid residue protein domain with small size (6kDa). The affibody has been used as a scaffold for the construction of combinatorial phage libraries and specific binding molecules have been identified from the phage libraries against a wide range of targets, such as insulin-like growth factor-1 receptor (IGF-IR)⁷⁸, HER2⁷⁹, Alzheimer amyloid β peptides⁸⁰, EGFR⁸¹, IgA⁸², TNF- α ⁸³, with affinity in the μ M to pM.

1.3.5 Phage as a component of diagnostic tests

Although antibodies are the dominantly utilized binding molecules in diagnostics, in recent years, phage particles have started to replace antibodies and be applied to the diagnosis of disease^{84,85}. Utilizing phage particles to establish a diagnostic test has significant advantages over antibodies: First, the growth and isolation of phage particles is much faster, easier and cheaper than the production of antibodies; Second, phages have increased stability towards pH, elevated temperatures and organic solvents compared to antibodies; Third, phage displaying binding molecules can be conveniently identified from phage-displayed libraries and panning process *in vitro*, the development period is much shorter compared to the generation of antibodies. Multiple antibody-based diagnostic test platforms have the potential to replacing antibodies with phage particles. Sandwich ELISA has been demonstrated to utilizing phage particles as binding components, such as a Staphylococcal enterotoxin B (SEB)⁸⁶ detection, a prostate cancer diagnostic test⁸⁷.

After the tight and specific binding molecules are selected from the phage library, the binding molecules are usually sequenced, cloned into a protein expression plasmid and purified from the cell culture. Although small binding molecules have advantages in biomedical imaging *in vivo* due to the easiness of conjugation with fluorophores and better tumor penetration properties, removal

of binding molecules from the phage platform can sometimes dramatically decrease the binding affinity, which is a universal problem for any kind of display technology. To overcome this issue, applying the whole phage particles to construct a diagnostic test *in vitro* not only reduces the risk of losing binding affinity to minimum, but also significantly reduces the cost and shortens the development time by avoiding the additional cloning and purification steps. Additionally, phage particles can service as an affinity matrix to extend the binding molecules from the immobilization surface and allow the binding molecules access to the biomarkers in solution. For example, in an ELISA assay milk powder or bovine serum albumin (BSA) are applied to passivate the surface and prevent non-specific binding of analyte. The size of BSA or albumin in milk powder is larger than some peptides or small protein binding molecules and can block the access of the binding molecules to the analyte in solution. By applying the whole phage particles in assay, phage particles function as a long linker to maintain the functional binding structures as well as allow the access to the analyte.

Multiple protein scaffolds have been displayed on phage surface with diverse structures, molecular sizes, physical and chemical properties. Construction of phage display libraries and identification of binding molecules against protein targets, small molecules, cell surface receptors via panning have become an important technology. The binding molecules selected through phage display technology compare favorably with the high binding affinity and specificity of traditional antibodies. Phage particles have the potential to replace antibodies to establish diagnostics and based on different diagnostic applications, appropriate phage library scaffolds could be selected without sacrificing the advantages of utilizing phage particles instead of antibodies.

1.4 Tuberculosis

Tuberculosis (TB) is an infectious disease caused by the bacillus *Mycobacterium tuberculosis* (*Mtb*), which typically affects the lungs (pulmonary TB), but can affect other organs or tissues as well (extrapulmonary TB). TB is spread from person to person via droplets forced into the air by pulmonary patients coughing.⁸⁸

The treatment of TB usually involves a six-month regimen of antibiotics. A combination of anti-TB medicines developed in 1940s and 1950s, such as isoniazid and rifampin, greatly reduces the mortality rate of TB patients. In 2009, the cure rate reached 87% at the global level, which is the highest cure rate in TB treatment history.⁸⁹ Although efficient treatment has existed for decades, TB was declared a global public health emergency by the World Health Organization (WHO) in 1993 when an estimated 7 ~ 8 million new TB cases and 1.3~1.6 million TB-related deaths were reported.⁸⁹ In 2010, still an estimated 8.5~9.2 million cases, among which an estimated 290,000 cases of multidrug-resistance TB (MDR-TB) existed, and 1.2~1.5 million deaths were reported.⁸⁹ In 2014, TB surpassing human immunodeficiency virus (HIV) became the leading cause of death from infectious disease worldwide for the first time and caused 1.2 million deaths.⁸⁸ Outbreaks of multiple drug resistance (MDR-TB) and TB-HIV-co-infection in recent years have contributed greatly to the high mortality rate and have made diagnosis and treatment of TB more complex.

A great challenge to reducing the morbidity and mortality of TB is the lack of sensitive and reliable diagnostics that are affordable and suitable for point-of-care application. The primary TB diagnostic method at point-of-care in the under developed areas is sputum smear microscopy developed 100 years ago. Sputum smear microscopy lacks sensitivity especially for cases where co-infected with HIV is involved. In routine clinical practice, the sensitivity of sputum smear microscopy varies between 35% and 80% in HIV-negative patients, but drops to as low as 20%

among HIV-infected patients.⁹⁰ A much more sensitive TB diagnostic is culture incubation, regarded as gold standard for TB diagnosis by the WHO. Culture incubation can diagnose TB as well as screen for drug resistant TB. However, culture incubation requires a laboratory, highly trained staff and an efficient transport system to ensure viable sputum samples arrive in a timely manner. Additionally, the culture process takes weeks to obtain results. In high TB burden areas, such as South-East Asia, Western Pacific Regions, and most of Africa, laboratory facilities are usually inaccessible, which constrains the application of culture diagnostics at point-of-care in resource-limited areas. Since 2010, the WHO started to provide a nucleic acid amplification based diagnostic Xpert MTB/RIF which can diagnose TB and rifampin resistance within 2 h with high sensitivity and specificity.⁹¹ However, in one study 12 false-positive rifampin resistance cases out of 22 tested TB cases were identified using the Xpert MTB/RIF test.⁹² And Xpert MTB/RIF was found to identify DNA from nonviable, degraded bacilli from patients recovered from previous TB infections also leading to false positive results.⁹³ Additionally the cost of Xpert/MTB/RIF is often prohibitively high considering the cost of the equipment (\$17,000) and each test cartridge (\$9.98).

None of the above described diagnostic methods satisfies the “affordable, effective and suitable for point-of-care” criteria for an ideal TB diagnostic test. And the above described methods are constrained in diagnosing extrapulmonary TB, which requires that samples of the suspected organ or tissue be collected by biopsy. Biopsy needs highly trained staff and laboratory facilities. Additionally even for pulmonary TB, since the main method of diagnosis involves identifying presence of *Mtb* in sputum samples, sputum samples sometimes are challenging to obtain for example in children and are risky for health workers when appropriate protections are unavailable.

Because of the above issues, development of novel affordable, effective and suitable for point-of-care (ideally laboratory-free) TB diagnostics is highly encouraged by the WHO.

Pathogen-derived biomarkers or host response based diagnostics triggered wide interest for TB diagnosis. For pathogen-derived biomarkers or host response based diagnostics, urine or blood serves as diagnostic samples. Urine and blood are much safer to handle compared with sputum. Additionally, those diagnostics have the potential to diagnose pulmonary TB as well as extra pulmonary TB using one sample assay. However, the pathogen biology and interaction between host and pathogen are only partially understood causing frustration in attempts to exploit the host response in antibody detection diagnostic assays.^{94,95} In 2011, the WHO recommended against using host response based tests for diagnosing pulmonary and extrapulmonary TB.⁸⁹ A major effort has been devoted to identifying pathogen-related biomarkers which could be translated into a specific and sensitive diagnostic tests.

Keeping in mind the criteria of an ideal TB diagnostic, non-invasive samples like urine triggered the greatest interest because urine is more convenient to obtain, safer to handle and has a simpler composition than other body fluids, such as sputum and blood. The process of obtaining urine is painless to patients and special training for health workers is not needed, all of which make urine based diagnostics very suitable for applications in local health centers of under-developed countries.

Multiple groups have been devoted to finding pathogen-related biomarkers in urine of TB patients. Among all the identified biomarker candidates, lipoarabinomannan (LAM), which is a structurally important component of the outer cell wall of *Mtb*, is the most frequently targeted molecule. Several clinical studies have been applied to screen people with suspected TB by using anti-LAM ELISA tests. Tessema reported a sensitivity of 74.0% and specificity of 86.9% in

Ethiopia.⁹⁶ Boehme reported a sensitivity of 80.3% and specificity of 99% in Tanzania.⁹⁷ However, another 4 studies show consistently poor overall test performance with sensitivities ranging from 13% to 51%.⁹⁸⁻¹⁰¹ These results suggest little consistency for the detection of urine LAM in patients with suspected TB. Determine™ TB-LAM (Alere Inc., USA), a commercial lateral flow based strip TB diagnostic by detecting LAM, was identified to have good sensitivity in HIV-co-infected patients with the most advanced immunodeficiency¹⁰² and therefore is promising to be used as an “add-on” test instead of replacing the primary test sputum smear microscopy. Although there is progress for applying LAM detection in TB diagnosis, the LAM-based diagnostic is still far from ideal. Identification and validation of other pathogen-derived markers in urine is a top priority in the development of new TB diagnostics.

The Campos-Neto laboratory found that *Mtb* ornithine transcarbamylase (*Mtb* OTC, Rv1656) in urine could work as a possible molecular marker of active pulmonary TB.¹⁰³ They first discovered *Mtb* OTC in urine of patients with pulmonary TB by mass spectrometric survey. Western blot analysis showed this protein was also present in *Mtb* crude cell extracts as well as in the culture supernatant, which suggests the protein in urine probably originates from *Mtb* and is not degraded by the host metabolism. The protein was detected by the IgG antibodies from active TB patients but not by those from uninfected healthy subjects. These results suggest *Mtb* OTC as a promising marker for active TB diagnosis. The *Mtb* OTC was chosen as the diagnostic target of our diagnostic tests. To our knowledge, no available diagnostic tests have been developed towards this target.

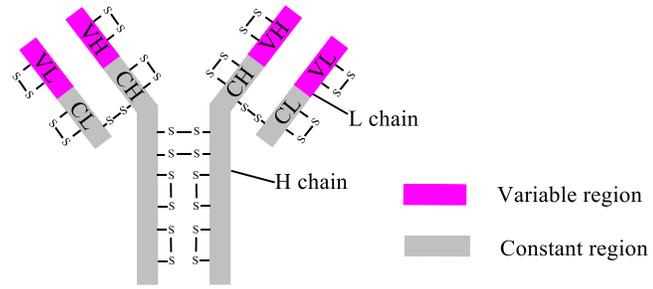


Figure 1-1. Basic structure of an antibody molecule

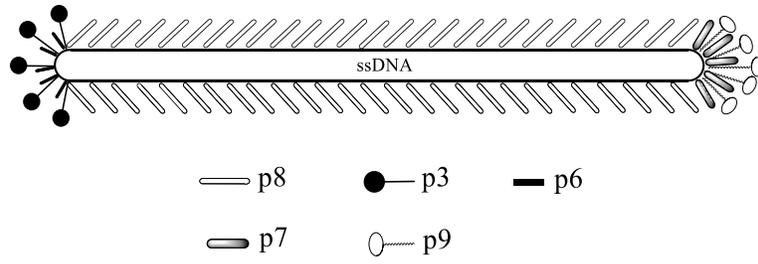
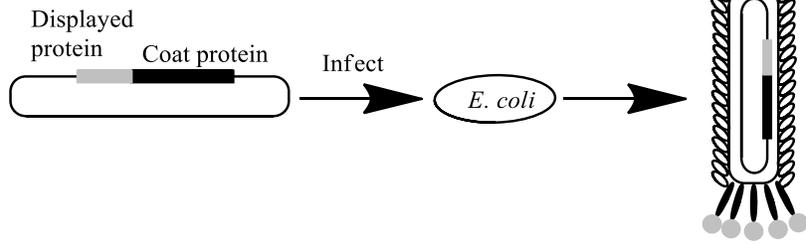


Figure 1-2. Diagram of an Ff bacteriophage particle

Phage vector: polyvalent display



Phagemid vector: monovalent display

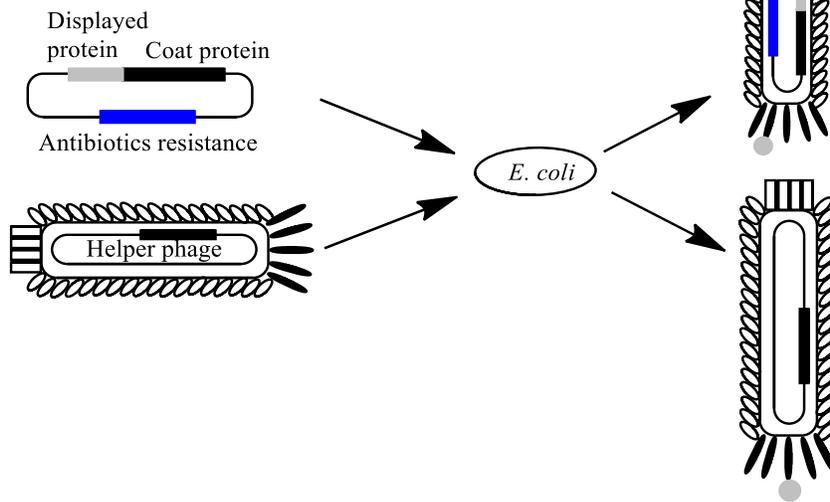


Figure 1-3. Types of phage display vectors³⁸

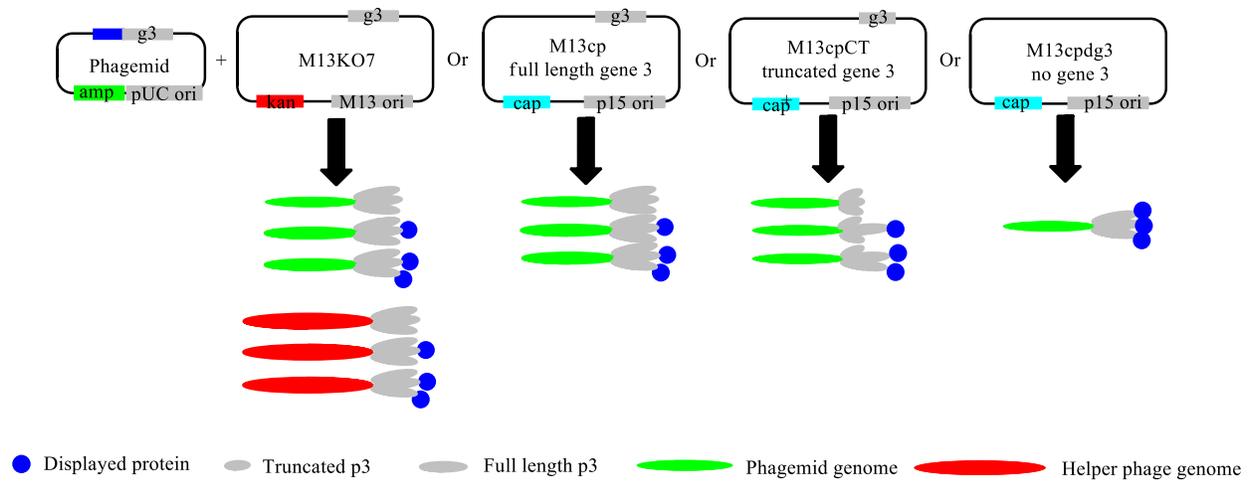


Figure 1-4. The phage particle genotypes and phenotypes when using the standard phagemid and helper phage (M13KO7) as well as the three helper plasmids (M13cp, M13cpCT, M13cpdg3).⁵⁰

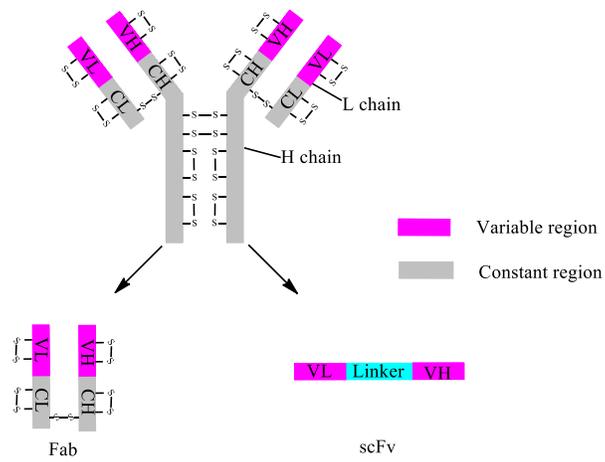


Figure 1-5. Structure of a full length antibody molecule and two shortened versions of antibody termed Fab and scFv.

Table 1-1 Proteins involved in the f1 phage life cycle and their function as well as molecular weight¹⁸

Protein	Function	No. of amino acids^Δ	Protein MW^Δ
p2	DNA replication	410	46,137
p10	DNA replication	111	12,672
p5	Binding ssDNA	87	9,682
p8	Major coat protein	50*	5,235*
p3	Minor coat protein	406*	42,522*
p6	Minor coat protein	112	12,342
p7	Minor coat protein	33	3,599
p9	Minor coat protein	32	3,650
p1	Assembly	348	39,502
p4	Assembly	405*	43,476*
p11	Assembly	108	12,424

^ΔThe initiating methionine is included in the proteins.

*The number of amino acids and the molecular weight for p3, p8 and p4 are for the mature proteins, which do not contain an N-terminal signal sequence.

References

1. Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E.B., Bendahman, N. & Hammers, R. (1993). Naturally occurring antibodies devoid of light chains. *Nature* **363**, 446–448
2. Roux, K.H., Greenberg, A.S., Greene, L., Strelets, L., Avila, D., McKinney, E.C. & Flajnik, M.F. (1998). Structural analysis of the nurse shark (new) antigen receptor (NAR): molecular convergence of NAR and unusual mammalian immunoglobulins. *Proceedings of the National Academy of Sciences* **95**, 11804–11809
3. Ecker, D.M., Jones, S.D. & Levine, H.L. (2015). The therapeutic monoclonal antibody market. *MAbs* **7**, 9–14
4. Oldham, R.K. (1983). Monoclonal antibodies in cancer therapy. *Journal of Clinical Oncology* **1**, 582–590
5. Kattah, A.G. & Fervenza, F.C. (2012). Rituximab: emerging treatment strategies of immune-mediated glomerular disease. *Expert review of clinical immunology* **8**, 413–421
6. Cahoon Jr, W., Ensor, C. & Shullo, M. (2012). Alemtuzumab for cytolytic induction of immunosuppression in heart transplant recipients. *Progress in Transplantation* **22**, 344–350
7. Leuvering, J.H., Thal, P., Waart, M. van der & Schuurs, A. (1980). Sol particle immunoassay (SPIA). *J. Immunoassay* **1**, 77–91
8. Van Dam, G.J., Wichers, J.H., Ferreira, T.M.F., Ghati, D., van Amerongen, A. & Deelder, A.M. (2004). Diagnosis of schistosomiasis by reagent strip test for detection of circulating cathodic antigen. *J. Clin. Microbiol.* **42**, 5458–61
9. Fernández-Sánchez, C., McNeil, C.J., Rawson, K., Nilsson, O., Leung, H.Y. & Gnanapragasam, V. (2005). One-step immunostrip test for the simultaneous detection of free and total prostate specific antigen in serum. *J. Immunol. Methods* **307**, 1–12
10. Wang, S., Quan, Y., Lee, N. & Kennedy, I.R. (2006). Rapid determination of fumonisin B1 in food samples by enzyme-linked immunosorbent assay and colloidal gold immunoassay. *J. Agric. Food Chem.* **54**, 2491–2495
11. Posthuma-Trumpie, G.A., Korf, J. & van Amerongen, A. (2009). Lateral flow (immuno) assay: its strengths, weaknesses, opportunities and threats. A literature survey. *Analytical and bioanalytical chemistry* **393**, 569–582
12. Henderson, K. & Stewart, J. (2001). A dipstick immunoassay to rapidly measure serum oestrone sulfate concentrations in horses. *Reproduction, Fertility and Development* **12**, 183–189

13. Snowden, K. & Hommel, M. (1991). Antigen detection immunoassay using dipsticks and colloidal dyes. *J. Immunol. Methods* **140**, 57–65
14. Rodi, D.J., Mandava, S. & Makowski, L. (2005). Filamentous bacteriophage structure and biology. *Phage display in biotechnology and drug discovery* **3**, 1–61
15. Beck, E. & Zink, B. (1981). Nucleotide sequence and genome organisation of filamentous bacteriophages f1 and fd. *Gene* **16**, 35–58
16. Hill, D.F. & Petersen, G.B. (1982). Nucleotide sequence of bacteriophage f1 DNA. *Journal of virology* **44**, 32–46
17. Van Wezenbeek, P.M., Hulsebos, T.J. & Schoenmakers, J.G. (1980). Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd. *Gene* **11**, 129–148
18. Webster, R. (2001). Filamentous phage biology. *Phage Barbas CF, Burton DR, Scott JK and Silverman (eds) Phage Display: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor pp 1–1*
19. Specthrie, L., Bullitt, E., Horiuchi, K., Model, P., Russel, M. & Makowski, L. (1992). Construction of a microphage variant of filamentous bacteriophage. *Journal of molecular biology* **228**, 720–724
20. Caspar, D.L. & Makowski, L. (1981). The symmetries of filamentous phage particles. *J. Mol. Biol.* **145**, 611–7
21. Overman, S.A., Tsuboi, M. & Thomas Jr, G.J. (1996). Subunit Orientation in the Filamentous VirusFf (fd, f1, M13). *Journal of molecular biology* **259**, 331–336
22. Zacher, A.N., Stock, C.A., Golden, J.W. & Smith, G.P. (1980). A new filamentous phage cloning vector: fd-tet. *Gene* **9**, 127–140
23. Williams, K.A., Farrow, N.A., Deber, C.M. & Kay, L.E. (1996). Structure and dynamics of bacteriophage IKe major coat protein in MPG micelles by solution NMR. *Biochemistry* **35**, 5145–57
24. McDonnell, P., Shon, K., Kim, Y. & Opella, S. (1993). fd Coat protein structure in membrane environments. *Journal of molecular biology* **233**, 447–463
25. Papavoine, C.H., Remerowski, M.L., Horstink, L.M., Konings, R.N., Hilbers, C.W. & van de Ven, F.J. (1997). Backbone dynamics of the major coat protein of bacteriophage M13 in detergent micelles by ¹⁵N nuclear magnetic resonance relaxation measurements using the model-free approach and reduced spectral density mapping. *Biochemistry* **36**, 4015–4026

26. Davis, N.G., Boeke, J.D. & Model, P. (1985). Fine structure of a membrane anchor domain. *J. Mol. Biol.* **181**, 111–21
27. Russel, M. & Ka'zmiarczyk, B. (1993). Analysis of the structure and subcellular location of filamentous phage pIV. *Journal of bacteriology* **175**, 3998–4007
28. Rapoza, M.P. & Webster, R.E. (1993). The filamentous bacteriophage assembly proteins require the bacterial SecA protein for correct localization to the membrane. *J. Bacteriol.* **175**, 1856–9
29. Stengele, I., Bross, P., Garces, X., Giray, J. & Rasched, I. (1990). Dissection of functional domains in phage fd adsorption protein: discrimination between attachment and penetration sites. *Journal of molecular biology* **212**, 143–149
30. Gray, C.W., Brown, R. & Marvin, D. (1981). Adsorption complex of filamentous fd virus. *Journal of molecular biology* **146**, 621–627
31. Armstrong, J., N Perham, R. & Walker, J.E. (1981). Domain structure of bacteriophage fd adsorption protein. *FEBS letters* **135**, 167–172
32. Deng, L.W., Malik, P. & Perham, R.N. (1999). Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pilus of Escherichia coli. *Virology* **253**, 271–7
33. Holliger, P., Riechmann, L. & Williams, R.L. (1999). Crystal structure of the two N-terminal domains of g3p from filamentous phage fd at 1.9 Å: evidence for conformational lability. *Journal of molecular biology* **288**, 649–657
34. Fulford, W. & Model, P. (1984). Gene X of bacteriophage f1 is required for phage DNA synthesis. Mutagenesis of in-frame overlapping genes. *J. Mol. Biol.* **178**, 137–53
35. Russel, M. & Model, P. (1989). Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. *J. Virol.* **63**, 3284–95
36. Lopez, J. & Webster, R.E. (1983). Morphogenesis of filamentous bacteriophage f1: orientation of extrusion and production of polyphage. *Virology* **127**, 177–193
37. Rakonjac, J., Feng, J. & Model, P. (1999). Filamentous phage are released from the bacterial membrane by a two-step mechanism involving a short C-terminal fragment of pIII. *Journal of molecular biology* **289**, 1253–1265
38. Hoess, R.H. (2001). Protein design and phage display. *Chem. Rev.* **101**, 3205–18

39. Hufton, S.E., Moerkerk, P.T., Meulemans, E.V., de Bruïne, A., Arends, J.W. & Hoogenboom, H.R. (1999). Phage display of cDNA repertoires: the pVI display system and its applications for the selection of immunogenic ligands. *J. Immunol. Methods* **231**, 39–51
40. Velappan, N., Fisher, H.E., Pesavento, E., Chasteen, L., D'Angelo, S., Kiss, C., Longmire, M., Pavlik, P. & Bradbury, A.R. (2010). A comprehensive analysis of filamentous phage display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins. *Nucleic Acids Res.* **38**, e22–e22
41. Løset, G.Å., Bogen, B. & Sandlie, I. (2011). Expanding the versatility of phage display I: efficient display of peptide-tags on protein VII of the filamentous phage. *PLoS One* **6**, e14702
42. Løset, G.Å., Roos, N., Bogen, B. & Sandlie, I. (2011). Expanding the versatility of phage display II: improved affinity selection of folded domains on protein VII and IX of the filamentous phage. *PLoS One* **6**, e17433
43. Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. (1995). Modifying filamentous phage capsid: limits in the size of the major capsid protein. *J. Mol. Biol.* **248**, 835–44
44. Kretzschmar, T. & Geiser, M. (1995). Evaluation of antibodies fused to minor coat protein III and major coat protein VIII of bacteriophage M13. *Gene* **155**, 61–5
45. Greenwood, J., Willis, A.E. & Perham, R.N. (1991). Multiple display of foreign peptides on a filamentous bacteriophage: peptides from Plasmodium falciparum circumsporozoite protein as antigens. *Journal of molecular biology* **220**, 821–827
46. Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J. & Lerner, R.A. (1991). Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proceedings of the National Academy of Sciences* **88**, 4363–4366
47. Marciano, D.K., Russel, M. & Simon, S.M. (1999). An aqueous channel for filamentous phage export. *Science (80)*. **284**, 1516–1519
48. Linderoth, N.A., Simon, M.N. & Russel, M. (1997). The filamentous phage pIV multimer visualized by scanning transmission electron microscopy. *Science (80-)*. **278**, 1635–1638
49. Sidhu, S.S., Weiss, G.A. & Wells, J.A. (2000). High copy display of large proteins on phage for functional selections. *Journal of molecular biology* **296**, 487–495
50. Chasteen, L., Ayris, J., Pavlik, P. & Bradbury, A. (2006). Eliminating helper phage from phage display. *Nucleic Acids Res.* **34**, e145–e145
51. Better, M., Chang, C.P., Robinson, R.R. & Horwitz, A.H. (1988). Escherichia coli secretion of an active chimeric antibody fragment. *Science (80)*. **240**, 1041–1043

52. Horwitz, A.H., Chang, C.P., Better, M., Hellstrom, K.E. & Robinson, R.R. (1988). Secretion of functional antibody and Fab fragment from yeast cells. *Proceedings of the National Academy of Sciences* **85**, 8678–8682
53. Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.-M., Lee, T., Pope, S.H., Riordan, G.S. & Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science* (80). **242**, 423–426
54. Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.-S., Novotny, J., Margolies, M.N., Ridge, R.J., Brucoleri, R.E., Haber, E. & Crea, R. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **85**, 5879–5883
55. Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T. & Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature* **341**, 544–546
56. Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R. & Muyldermans, S. (1997). Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* **414**, 521–6
57. Colcher, D., Pavlinkova, G., Beresford, G., Booth, B.J. & Batra, S.K. (1999). Single-chain antibodies in pancreatic cancer. *Ann. N. Y. Acad. Sci.* **880**, 263–80
58. Colcher, D., Bird, R., Roselli, M., Hardman, K.D., Johnson, S., Pope, S., Dodd, S.W., Pantoliano, M.W., Milenic, D.E. & Schlom, J. (1990). In vivo tumor targeting of a recombinant single-chain antigen-binding protein. *Journal of the National Cancer Institute* **82**, 1191–1197
59. Clackson, T., Hoogenboom, H.R., Griffiths, A.D. & Winter, G. (1991). Making antibody fragments using phage display libraries.
60. Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P. & Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res.* **19**, 4133–4137
61. Ewert, S., Honegger, A. & Plückthun, A. (2004). Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. *Methods* **34**, 184–199
62. Koide, A., Bailey, C.W., Huang, X. & Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. *Journal of molecular biology* **284**, 1141–1151

63. Binz, H.K., Stumpp, M.T., Forrer, P., Amstutz, P. & Plückthun, A. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *Journal of molecular biology* **332**, 489–503
64. Nord, K., Gunneriusson, E., Ringdahl, J., Ståhl, S., Uhlén, M. & Nygren, P.-Å. (1997). Binding proteins selected from combinatorial libraries of an α -helical bacterial receptor domain. *Nature biotechnology* **15**, 772–777
65. Wojcik, J., Hantschel, O., Grebien, F., Kaupe, I., Bennett, K.L., Barkinge, J., Jones, R.B., Koide, A., Superti-Furga, G. & Koide, S. (2010). A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nature structural & molecular biology* **17**, 519–527
66. Hackel, B.J., Kapila, A. & Wittrup, K.D. (2008). Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling. *Journal of molecular biology* **381**, 1238–1252
67. Bork, P. (1993). Hundreds of ankyrin-like repeats in functionally diverse proteins: Mobile modules that cross phyla horizontally? *Proteins: Structure, Function, and Bioinformatics* **17**, 363–374
68. Li, J., Mahajan, A. & Tsai, M.-D. (2006). Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry* **45**, 15168–15178
69. Plückthun, A. (2015). Designed ankyrin repeat proteins (DARPs): binding proteins for research, diagnostics, and therapy. *Annual review of pharmacology and toxicology* **55**, 489–511
70. Steiner, D., Forrer, P., Stumpp, M.T. & Plückthun, A. (2006). Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat. Biotechnol.* **24**, 823–31
71. Paschke, M. & Höhne, W. (2005). A twin-arginine translocation (Tat)-mediated phage display system. *Gene* **350**, 79–88
72. Nangola, S., Minard, P. & Tayapiwatana, C. (2010). Appraisal of translocation pathways for displaying ankyrin repeat protein on phage particles. *Protein expression and purification* **74**, 156–161
73. Zahnd, C., Wyler, E., Schwenk, J.M., Steiner, D., Lawrence, M.C., McKern, N.M., Pecorari, F., Ward, C.W., Joos, T.O. & Plückthun, A. (2007). A designed ankyrin repeat protein evolved to picomolar affinity to Her2. *Journal of molecular biology* **369**, 1015–1028
74. Schweizer, A., Roschitzki-Voser, H., Amstutz, P., Briand, C., Gulotti-Georgieva, M., Prenosil, E., Binz, H.K., Capitani, G., Baici, A., Plückthun, A. & others (2007). Inhibition

of caspase-2 by a designed ankyrin repeat protein: specificity, structure, and inhibition mechanism. *Structure* **15**, 625–636

75. Amstutz, P., Koch, H., Binz, H.K., Deuber, S.A. & Plückthun, A. (2006). Rapid selection of specific MAP kinase-binders from designed ankyrin repeat protein libraries. *Protein Eng. Des. Sel.* **19**, 219–29
76. Amstutz, P., Binz, H.K., Parizek, P., Stumpp, M.T., Kohl, A., Grütter, M.G., Forrer, P. & Plückthun, A. (2005). Intracellular kinase inhibitors selected from combinatorial libraries of designed ankyrin repeat proteins. *Journal of Biological Chemistry* **280**, 24715–24722
77. Nilsson, B., Moks, T., Jansson, B., Abrahmsén, L., Elmblad, A., Holmgren, E., Henrichson, C., Jones, T.A. & Uhlén, M. (1987). A synthetic IgG-binding domain based on staphylococcal protein A. *Protein engineering* **1**, 107–113
78. Li, J., Lundberg, E., Vernet, E., Larsson, B., Höidén-Guthenberg, I. & Gräslund, T. (2010). Selection of affibody molecules to the ligand-binding site of the insulin-like growth factor-1 receptor. *Biotechnology and applied biochemistry* **55**, 99–109
79. Orlova, A., Magnusson, M., Eriksson, T.L., Nilsson, M., Larsson, B., Höidén-Guthenberg, I., Widström, C., Carlsson, J., Tolmachev, V., Ståhl, S. & others (2006). Tumor imaging using a picomolar affinity HER2 binding affibody molecule. *Cancer research* **66**, 4339–4348
80. Grönwall, C., Jonsson, A., Lindström, S., Gunneriusson, E., Ståhl, S. & Herne, N. (2007). Selection and characterization of Affibody ligands binding to Alzheimer amyloid β peptides. *Journal of biotechnology* **128**, 162–183
81. Friedman, M., Orlova, A., Johansson, E., Eriksson, T.L., Höidén-Guthenberg, I., Tolmachev, V., Nilsson, F.Y. & Ståhl, S. (2008). Directed evolution to low nanomolar affinity of a tumor-targeting epidermal growth factor receptor-binding affibody molecule. *Journal of molecular biology* **376**, 1388–1402
82. Rönmark, J., Grönlund, H., Uhlén, M. & Nygren, P.-Å. (2002). Human immunoglobulin A (IgA)-specific ligands from combinatorial engineering of protein A. *European Journal of Biochemistry* **269**, 2647–2655
83. Jonsson, A., Wällberg, H., Herne, N., Ståhl, S. & Frejd, F.Y. (2009). Generation of tumour-necrosis-factor- α -specific affibody molecules capable of blocking receptor binding in vitro. *Biotechnology and applied biochemistry* **54**, 93–103
84. Penner, G.A.W. & M, R. (2008). The promise of phage display: customized affinity and specificity. *Analytical Chemistry* **80**, 3082–3089

85. Arugula, M.A. & Simonian, A. (2014). Novel trends in affinity biosensors: current challenges and perspectives. *Measurement Science and Technology* **25**, 032001
86. Goldman, E.R., Pazirandeh, M.P., Mauro, J.M., King, K.D., Frey, J.C. & Anderson, G.P. (2000). Phage-displayed peptides as biosensor reagents. *Journal of Molecular Recognition* **13**, 382–387
87. Lang, Q., Wang, F., Yin, L., Liu, M., Petrenko, V.A. & Liu, A. (2014). Specific probe selection from landscape phage display library and its application in enzyme-linked immunosorbent assay of free prostate-specific antigen. *Analytical chemistry* **86**, 2767–2774
88. WHO (2015). Global tuberculosis report 2015. *Geneva: World Health Organization*
89. WHO (2011). Global tuberculosis report 2011. *Geneva: World Health Organization*
90. Davies, P.D.O. & Pai, M. (2008). The diagnosis and misdiagnosis of tuberculosis. *Int. J. Tuberc. Lung Dis.* **12**, 1226–34
91. Boehme, C.C., Nabeta, P., Hillemann, D., Nicol, M.P., Shenai, S., Krapp, F., Allen, J., Tahirli, R., Blakemore, R., Rustomjee, R. & others (2010). Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine* **363**, 1005–1015
92. Ocheretina, O., Byrt, E., Mabou, M.-M., Royal-Mardi, G., Merveille, Y.-M., Rouzier, V., Fitzgerald, D.W. & Pape, J.W. (2016). False-positive Rifampin Resistant Results with Xpert MTB/RIF Version 4 Assay in clinical samples with a low bacterial load. *Diagn. Microbiol. Infect. Dis.*
93. Theron, G., Venter, R., Calligaro, G., Smith, L., Limberis, J., Meldau, R., Chanda, D., Esmail, A., Peter, J. & Dheda, K. (2016). Xpert MTB/RIF Results in Patients With Previous Tuberculosis: Can We Distinguish True From False Positive Results? *Clinical Infectious Diseases* **62**, 995–1001
94. Young, D.B., Perkins, M.D., Duncan, K. & Barry, C.E. (2008). Confronting the scientific obstacles to global control of tuberculosis. *J. Clin. Invest.* **118**, 1255–1265
95. Steingart, K.R., Flores, L.L., Dendukuri, N., Schiller, I., Laal, S., Ramsay, A., Hopewell, P.C. & Pai, M. (2011). Commercial serological tests for the diagnosis of active pulmonary and extrapulmonary tuberculosis: an updated systematic review and meta-analysis. *PLoS Med* **8**, e1001062
96. Tessema, T.A., Bjune, G., Assefa, G., Svenson, S., Hamasur, B. & Bjorvatn, B. (2002). Clinical and radiological features in relation to urinary excretion of lipoarabinomannan in Ethiopian tuberculosis patients. *Scand. J. Infect. Dis.* **34**, 167–71
97. Boehme, C., Molokova, E., Minja, F., Geis, S., Loscher, T., Maboko, L., Koulchin, V. & Hoelscher, M. (2005). Detection of mycobacterial lipoarabinomannan with an antigen-

- capture ELISA in unprocessed urine of Tanzanian patients with suspected tuberculosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **99**, 893–900
98. Daley, P., Michael, J., Hmar, P., Latha, A., Chordia, P., Mathai, D., John, K. & Pai, M. (2009). Blinded evaluation of commercial urinary lipoarabinomannan for active tuberculosis: a pilot study. *The international journal of tuberculosis and lung disease: the official journal of the International Union against Tuberculosis and Lung Disease* **13**, 989
 99. Dheda, K., Davids, V., Lenders, L., Roberts, T., Meldau, R., Ling, D., Brunet, L., van Zyl Smit, R., Peter, J., Green, C. & others (2010). Clinical utility of a commercial LAM-ELISA assay for TB diagnosis in HIV-infected patients using urine and sputum samples. *PLoS One* **5**, e9848
 100. Mutetwa, R., Boehme, C., Dimairo, M., Bandason, T., Munyati, S.S., Mangwanya, D., Mungofa, S., Butterworth, A.E., Mason, P.R. & Corbett, E.L. (2009). Diagnostic accuracy of commercial urinary lipoarabinomannan detection in African tuberculosis suspects and patients. *Int. J. Tuberc. Lung Dis.* **13**, 1253–9
 101. Reither, K., Saathoff, E., Jung, J., Minja, L.T., Kroidl, I., Saad, E., Huggett, J.F., Ntinginya, E.N., Maganga, L., Maboko, L. & others (2009). Low sensitivity of a urine LAM-ELISA in the diagnosis of pulmonary tuberculosis. *BMC infectious diseases* **9**, 1
 102. Lawn, S.D., Dheda, K., Kerkhoff, A.D., Peter, J.G., Dorman, S., Boehme, C.C. & Nicol, M.P. (2013). Determine TB-LAM lateral flow urine antigen assay for HIV-associated tuberculosis: recommendations on the design and reporting of clinical studies. *BMC infectious diseases* **13**, 407
 103. Napolitano, D.R., Pollock, N., Kashino, S.S., Rodrigues, V. & Campos-Neto, A. (2008). Identification of Mycobacterium tuberculosis ornithine carboamyltransferase in urine as a possible molecular marker of active pulmonary tuberculosis. *Clinical and Vaccine Immunology* **15**, 638–643

Chapter 2

Phage display selection of tight specific binding variants from a hyperthermostable Sso7d scaffold protein library*

Antibodies, the quintessential biological recognition molecules, are not ideal for many applications because of their large size, complex modifications, and thermal and chemical instability. Identifying alternative scaffolds that can be evolved into tight, specific binding molecules with improved physical properties is of increasing interest, particularly for biomedical applications in resource-limited environments. Hyperthermophilic organisms, such as *Sulfolobus solfataricus*, are an attractive source of highly stable proteins as starting points for alternative molecular recognition scaffolds. We describe the first application of phage display to identify binding proteins based on the *Sulfolobus solfataricus* protein Sso7d scaffold. Sso7d is a small (approximately 7 kDa, 63 amino acids), cysteine free DNA-binding protein with a melting temperature of nearly 100 °C. Tight binding Sso7d variants were selected for a diverse set of protein targets from a 10¹⁰ member library, demonstrating the versatility of the scaffold. These Sso7d variants are able to discriminate among closely-related human, bovine, and rabbit serum albumins. Equilibrium dissociation constants in the nanomolar to low micromolar range were measured via competitive ELISA. Importantly, the Sso7d variants continue to bind their targets in the absence of the phage context. Furthermore, phage-displayed Sso7d variants retain their binding affinity after exposure to temperatures up to 70 °C. Taken together, our results suggest that the

**This chapter was adapted from a publication on FEBS Journal 2016 Vol 238:1351~1367. The agreement of reprinting and reusing this publication in Ning Zhao's PhD dissertation is attached in Appendix 4.*

Sso7d scaffold will be a complementary addition to the range of non-antibody scaffold proteins that can be utilized in phage display. Variants of hyperthermostable binding proteins have potential applications in diagnostics and therapeutics for environments with extreme conditions of storage and deployment.

2.1 Introduction

Tight, specific binding molecules are important components of medical diagnostics and therapeutics. Phage display developed as an *in vitro* method to select and evolve binding molecules through the creation and screening of large libraries of peptides or proteins displayed on the surface of phage particles. Binding molecules are selected by panning a library against an immobilized target and retaining those particles that bind^{1,2}. The bound particles are eluted, amplified through propagation in bacteria, and rescreened. Several rounds of binding and amplification allow the selection of tight binding variants from billions of starting sequences. Beyond phage display, the nanoscale size, simple life cycles, and speed and ease with which they can be prepared, manipulated, and characterized render phages attractive and versatile systems for myriad biotechnological applications.

The filamentous bacteriophage M13 is most commonly employed for phage display, and M13 systems are at the forefront of research in nanomaterials³⁻⁷ and nanomedicine⁸⁻¹¹. Fusion proteins have been displayed on every one of the five M13 coat proteins¹²⁻¹⁵, and platforms utilizing combinations of several, but not all, coat proteins have been developed^{14,16-18}. Modified phage displaying simultaneous fusions to two coat proteins have been used build nanowires¹⁶, construct components of batteries¹⁹, produce components for biosensors⁵, and act as tissue engineering scaffolds²⁰ and cell-targeted therapeutic delivery vehicles^{8,21}. The ability to select functional

elements and produce large quantities of decorated phage particles is central to the biotechnological uses of M13.

Generating highly decorated phage is challenging because the small size and close packing of phage coat proteins as well as the phage assembly process all place constraints on the display of fusion proteins²². M13 particles consist of a single-stranded DNA genome encased in approximately 2700 copies of a major coat protein (p8) and five copies each of four minor coat proteins: p3 and p6 on one end and p7 and p9 on the other. Protein fusions are most commonly displayed on the minor coat protein p3, the largest coat protein whose sequence of 406 amino acids forms three globular domains. p3 is employed for monovalent display of large proteins for selection, but attempts to display proteins on all 5 copies of p3 often lead to decreased phage yields. p7, p8 and p9 are all 50 or fewer amino acids in length, and packing interactions leave limited space for displayed fusions. p6 is rarely used for phage display. Particle production involves passing the assembling phage through a membrane-spanning pore complex that further limits the size of fusions on the phage surface^{23,24}. Many of the advanced applications of phage particles as components of materials, as imaging probes and therapeutic delivery vehicles would benefit from the ability to display multiple fusions in a controlled and high valent manner. One approach to overcome the bias against the display of large proteins on phage is to develop functional proteins of reduced size.

In molecular biology and medicine, antibodies have traditionally been used as specific binding molecules for diagnostics and therapeutics. Antibodies are large multidomain proteins, on the order of 150 kDa, which require dimerization and formation of key disulfide bonds to function properly. The antigen-binding elements of antibodies, complementarity determining regions (CDRs), consist of closely-packed loop structures connecting the β -strands in the immunoglobulin

fold. The majority of the amino acid diversity in antibodies is focused in the CDRs of the heavy and light chain variable domains. The remaining “scaffold” of the immunoglobulin domain contains only minor variability and serves to present the CDRs in a three dimensional context. In contrast to the 150 kDa size of antibodies, p3, the largest M13 coat protein, is only 42 kDa. Two recombinant forms of antibodies with reduced size are commonly employed in phage display: the fragment antigen binding (Fab) and single-chain fragment variable (scFv) scaffolds. Both Fabs and scFvs represent considerable reductions in size relative to an entire antibody, 50 kDa and 25 kDa, respectively. The Fab format consists of two chains which are expressed independently, with one as fusion to an M13 coat protein. The other chain must assemble with the phage after expression. scFvs, while expressed as a fusion to one of the phage coat proteins from a single gene, exhibit a tendency to dimerize, which can complicate selections ^{25,26}.

Although antibody fragments remain the most common framework for the evolution of binding molecules via phage display, the physical properties of antibodies are less than ideal for many applications and other protein scaffolds are continuously being developed ^{27,28}. Additionally, a recent review of commercial antibodies has suggested that many antibody reagents are of highly variable quality and often poorly characterized.²⁹ Some non-antibody-based alternative scaffolds such as fibronectin domains (monobodies) ³⁰ and single immunoglobulin domains (nanobodies) ³¹ mimic the CDRs of antibodies. These scaffolds are on the order of 10 kDa, less than half the size of a scFv. Other scaffolds such as Ankyrin repeat proteins (DARpins) ³² and small helical proteins (affibodies) ³³ display diversity in structural contexts that do not mimic the CDRs of antibodies. While larger than fibronectin domains and monobodies, 12-18 kDa DARpins and affibodies are still smaller than scFvs. Binding interfaces presented on non-antibody-based scaffolds have been

utilized as reagents for protein purification and detection and have potential as diagnostic imaging and therapeutic agents^{34,35}.

Three properties are particularly important when considering a protein as a possible scaffold. Small size, on the order of 60 amino acids or fewer, is desirable for production and for ease of modification. Protein stability, preferably in the absence of disulfide bonds, improves formation and maintenance of the desired structure in diverse environments. The tolerance of a protein toward mutation at several simultaneous positions is related to protein stability, and stability of the “parent” protein correlates with the ability to select functional variants via directed evolution³⁶. For greatest utility, binding variants should retain their function in the absence of the phage context. Facile expression in bacterial culture is desirable for the production of large quantities of a selected binding protein in an industrial setting. Further, scaffold domains of fewer than 60 amino acids have the potential to be synthesized chemically, rather than biologically, which expands the possibilities of diversification through chemical modification, including employment of non-natural amino acids³⁷.

The continued identification and development of small, stable binding scaffolds is important because increasing the diversity of three-dimensional architectures available as library templates should allow the selection of binding molecules with improved physical properties for diverse targets of different shapes. Because the space of target molecular structures is vast, scaffolds of different shapes are needed to identify appropriate binding interfaces. Some architectures will naturally be more compatible than others. The increased stability and decreased size of alternative scaffolds comes at the potential price of generality of interaction. The use of multiple scaffolds with different folds in a single selection experiment has the potential to allow the direct selection of specific binders for multiple epitopes on a single target³⁸.

A natural, but largely unexplored, source for small, highly stable protein scaffolds are the genomes of hyperthermophilic organisms. Because of the severe conditions under which these organisms live, their proteins have evolved to withstand greatly elevated temperatures and often extremes of pH and salt. Many proteins or protein domains from these organisms satisfy all of the desirable criteria for an alternative binding scaffold. High thermostability is especially attractive for biomedical applications in resource-limited environments.

Sso7d is a small 7 kDa, 63 amino acid, cysteine free, DNA-binding protein with a melting temperature of nearly 100 °C from *Sulfolobus solfataricus*. Sso7d is representative of an abundant class of 6-8 kDa, putative histone-like molecules isolated from thermophilic *Sulfolobus* species³⁹. Sso7d and the closely related Sac7d from *Sulfolobus acidocaldarius* possess a nearly identical three-dimensional structure and share 85% sequence identity and have been studied extensively as model thermophilic and DNA binding proteins^{40,41}. Their three-dimensional structures have been characterized both in solution and in the solid state and consist of an incomplete five-stranded β -barrel closed off by a C-terminal α -helix (Figure 2-1)^{42,43}. This fold, labeled the OB-fold, is found in diverse binding molecules that recognize oligonucleotides, oligosaccharides, proteins, and metal ions⁴⁴.

With an eye toward expanding the toolkit for high level, multivalent display of evolved binding molecules on phage and a particular interest in thermostable scaffolds for molecular recognition, we demonstrate the efficient display of Sso7d on M13 as a p3 fusion and describe the generation of a large phage display library of Sso7d variants from which tight binding proteins for a variety of targets were isolated. Sso7d and Sac7d have been used as scaffolds in yeast and ribosome display libraries⁴⁵⁻⁴⁸, but to our knowledge, this report is the first demonstration of successful phage display of Sso7d.

2.2 Results and Discussion

2.2.1 Sso7d phagemid constructs

The gene for an Sso7d variant selected via yeast display to bind streptavidin (Sso7d-strep) was assembled by sequential thermodynamically balanced inside out synthesis (seqTBIO) ^{47,49,50} and inserted into phagemid vectors designed to evaluate superfolder green fluorescent protein (sfGFP) display on M13 phage. The vectors (generous gifts from Andrew Bradbury) contain cassettes encoding TAT, SRP or Sec leader peptides that direct a sfGFP-p3 fusion protein to the inner membrane via alternative secretion systems ⁵¹. DNA encoding a hexahistidine tag is located between the genes for the phage displayed protein and full length p3 as part of a single reading frame. The Sso7d gene was cloned into each of the phagemids in place of the sfGFP gene to produce TAT-Sso7d-strep, SRP-Sso7d-strep, and Sec-Sso7d-strep. In the conversion of the vectors from sfGFP to Sso7d-strep displaying phagemids, a suppressible stop codon between the hexahistidine tag and the gene for p3 was removed. Successful cloning was monitored by loss of fluorescence in transformed cells and verified by sequencing.

2.2.2 Translocation signal dependent display of Sso7d

Successful insertion of the Sso7d-p3 fusion protein into the *E. coli* inner membrane is critical for effective phage display, as all phage coat proteins contain transmembrane segments prior to particle assembly ⁵². Protein transport in *E. coli* can accommodate either folded or unfolded proteins and be accomplished either co- or post-translationally. The two phage proteins most commonly used for phage display, p3 and p8, have peptide leader sequences that direct translocation across the cytoplasmic membrane post-translationally while the coat proteins are unfolded via the Sec pathway ⁵³. The pelB leader sequence in Np3-Sec triggers translocation via the Sec pathway and is commonly used for phage display vectors. The Signal Recognition Particle

pathway (SRP) is similar to the Sec pathway in that proteins are translocated in an unfolded state, but the pathway operates co-translationally⁵⁴. SRP leader sequences have been used to improve the phage display of ankyrin-based scaffolds for which the Sec leader was ineffective⁵⁵. Proteins that are stably folded or that require the reducing environment of the cytoplasm for proper folding can be transported in a native form via the Twin Arginine Transport (TAT) pathway. Phage display of functional, fluorescent sfGFP was achieved using a TAT leader peptide^{51,56}.

Sso7d is small, highly stable and disulfide-free. The most efficient system for translocation of this protein in *E. coli* is not obvious *a priori*. Bacteria containing the Sec-Sso7d-strep, SRP-Sso7d-strep or TAT-Sso7d-strep phagemid were infected with M13KO7 helper phage. Similar yields of particles (on the order of 10^{12} /mL of bacterial culture) were achieved for each of the three phagemids. ELISA with an anti-hexahistidine tag antibody was utilized to evaluate the extent of incorporation of the Sso7d-strep-p3 fusion protein into phages as a function of leader peptide, independent of the streptavidin-binding function of the Sso7d-strep variant. Immobilization of $2-3 \times 10^{10}$ Sec-Sso7d-strep and SRP-Sso7d-strep phages gave rise to strong ELISA signals, greater than 20-fold over background, suggesting successful incorporation of the Sso7d-p3 fusion (Figure 2-2). ELISA signals for both M13KO7 (no Sso7d variant displayed) and TAT-Sso7d-strep were not above that observed for blocked wells. Preliminary experiments suggested that the phage displayed context of Sso7d-strep may have compromised its ability to bind streptavidin. The same phages did not show any signal over background when evaluated with a streptavidin-HRP conjugate in ELISAs.

2.2.3 Library design and construction

Based on the levels of display facilitated by the three leader peptides, Sec-Sso7d-strep was selected as the template for the construction of a library of Sso7d variants. Analysis of the binding

faces of OB-fold proteins and the previously-reported yeast display Sso7d and ribosome display Sac7d libraries suggested 9 positions for randomization^{39,45,47,48}. The 9 amino acids are located on the exposed surfaces of the 3-stranded β -sheet that makes up the DNA binding interface of Sso7d (Figure 2-1b). The library was constructed using the degenerate codon NNC which encodes 15 of 20 amino acids (excluding Met, Lys, Glu, Gln, and Trp) and does not introduce any stop codons. The theoretical size of the library is $16^9 = 6.9 \times 10^{10}$ DNA sequences encoding $15^9 = 3.9 \times 10^{10}$ protein sequences.

In an effort to reduce background of unmodified template in the library preparation, a starting template was constructed in which 7 of 10 codons between positions 22 and 31 in the Sso7d scaffold were replaced by the rare arginine codons AGA or AGG (Sec-Sso7d-rArg, Table 2-1). The rare Arg codons are expected to reduce the translation of the starting template and place a metabolic burden on cells⁵⁷. For ease of identification of library members, a unique XbaI restriction site was also introduced into Sec-Sso7d-rArg. Both the XbaI restriction site and the rare Arg codons are removed by the mutagenic primer during library construction. The Sec-Sso7d-rArg library template displays poorly on phage based on anti-His tag ELISA (Figure 2-2). 10^{13} purified Sec-Sso7d-rArg phage particles gave rise to an ELISA signal that was approximately 2.5-fold over background. For comparison, strong ELISA signals approximately 30-fold over background were observed for 10^{10} phage particles prepared from Sec-Sso7d-strep without the rare arginine codons. Two of seven arginine mutations, V22R and I29R, place positively charged residues into the hydrophobic core of Sso7d, possibly lowering protein stability or inducing misfolding. Increased metabolic burden, reduced stability and lower levels of correctly folded protein may all contribute to the observed decrease in the level of display of the starting rare arginine template.

The phage displayed Sso7d library was constructed using Kunkel mutagenesis with a single mutagenic primer. Transformation of the reaction into SS320 cells resulted in 1.3×10^{11} unique transformants. The mutation efficiency was evaluated by analyzing XbaI digests of PCR products from 95 isolated colonies. The amplified fragment from a library member is not cut by XbaI, while the unmodified rare Arg template is digested. 12 of 95 colonies were identified as library members. DNA sequencing confirmed that each of those variants were unique library members. Considering mutagenic efficiency, the library contained 1.6×10^{10} individual variants. To further evaluate the diversity of the library, a total of 56 randomly selected library members were sequenced. The diversity of the library was evaluated using the statistical methodology of Rodi, Soares and Makowski, which describes a measure of library diversity that is based on approximating peptide probabilities in the library using observed residue frequencies and comparing the observed distribution to a theoretical ideal where each peptide sequence is encoded by an equal number of nucleotide sequences.^{58,59} This method has the advantage of calculating a measure of diversity that can be compared across libraries produced using different encoding strategies. The diversity in the first six (of nine total) randomized positions exhibit an essentially unbiased distribution of amino acids. Some bias toward the amino acids that were present in the starting template was apparent in the last three diversified positions. Even with this bias, our library ($d_e = 0.096$) includes approximately 1.5x the functional diversity of the commercially available 7 and 12 amino acid peptide phage display libraries ($d_e = 0.079$ and $d_e = 0.043$ for New England BioLabs PhD 7 and 12 libraries)⁵⁹.

The yeast display Sso7d and ribosome display Sac7d libraries largely under sampled their theoretical diversity. The ribosome display libraries were constructed using NNS codons at either 10 or 14 positions. The theoretical sizes of these libraries were 10^{15} or 10^{21} DNA sequences of

which approximately 10^{12} independent variants were obtained^{45,48}. The yeast display libraries varied 10 positions using NNN codons (DNA sequence diversity of approximately 10^{18}) and contained approximately 10^7 surface displayed sequences⁴⁷. Effective directed evolution experiments generally strike a balance between theoretical diversity and adequate coverage of that diversity⁶⁰. Variants with high affinity for the intended targets were identified using both the yeast and ribosome display systems. In this phage display study, a reduced codon set was employed to allow more thorough coverage of the theoretical diversity of the encoded library.

2.2.4 Library selection

In order to examine the ability to identify high affinity Sso7d binding variants via phage display, the library was evaluated against four target proteins: bovine serum albumin (BSA, 66 kDa), rabbit serum albumin, (RSA, 66 kDa), neutravidin (60 kDa), and green fluorescent protein (GFP, 27 kDa). BSA and RSA were included to assess the ability of Sso7d variants to distinguish between closely related proteins. Related experiments have been described for the small helical protein, albumin binding domain⁶¹. Neutravidin was included as a close structural, but sequentially unrelated, analog of streptavidin to compare against the Sso7d-strep binding sequence isolated by yeast display.

For each of the four protein targets, three or four rounds of panning against passively adsorbed target were performed (Table 2-2). The enhancement of the population of phages with affinity for the target protein with increasing rounds of selection was monitored by evaluating the retention of amplified phages. The ratio of phages from the amplified pool binding to target coated versus uncoated wells is a measure of functional selection and is expected to increase as panning progresses. In the last round of panning for BSA, neutravidin and GFP, the relative enhancement of binding to target coated versus uncoated wells was approximately 10^5 -fold (Table 2-2). A very

modest enhancement of 12-fold was observed for the RSA panning experiments, suggesting poor selection.

After the final round of panning, isolated clones were grown overnight in 96-well microplates. The supernatants of these small cultures were assayed for target-specific binding via phage ELISA. Based on phage yields after large scale purification, we estimate that approximately 10^{10} phage particles were utilized in these ELISAs. Supernatants from 23 of 23 clones grown after the fourth round of selection against BSA gave rise to high ELISA signals (greater than 10-fold over background). DNA sequencing of 10 of these clones revealed 3 different protein sequences (Table 2-1). Supernatants from 20 of 23 clones grown after 4 rounds of panning against neutravidin gave rise to high ELISA signals. DNA sequencing of 9 of these clones revealed largely similar sequences, identical at 7 of 9 amino acid positions. Supernatants from 18 of 24 clones grown after 3 rounds of panning against GFP gave rise to high ELISA signals. DNA sequencing of 10 clones revealed 10 different protein sequences, with some consensus of amino acid types at each randomized position. Supernatants from 3 of 48 clones grown after 3 rounds of panning against RSA showed strong ELISA signals; all 3 isolated variants had the same protein sequence.

2.2.5 Evaluation of selected serum albumin binding Sso7d variants:

In complex environments, often in the presence of many related molecules, antibody recognition is highly specific. The Sso7d scaffold presents a smaller and more rigid binding interface than the CDRs of antibodies, which could decrease the ability of Sso7d-derived binding molecules to distinguish among closely related targets. The Sso7d variants identified from panning the phage library against both BSA and RSA were evaluated for binding to three serum albumin proteins: BSA, RSA, and human serum albumin (HSA). These three proteins have approximately 75% sequence identity and greater than 86% sequence similarity.

The 10 clones characterized from the BSA binding variants encoded 3 protein sequences (Table 2-1). Two of the three protein sequences 2A4 and 2A6 describe 9 of 10 variants characterized and contain identical or very closely related amino acids at 7 of 9 varied positions. The remaining Sso7d variant 2A8 is distinct from the 2A4/2A6 clones, containing amino acids that differ in size or polarity at 5 of 9 varied positions. The sequences of the BSA binding clones are largely unrelated to the single isolated RSA binding sequence.

ELISA experiments to evaluate the specificity of the selected albumin binding sequences for immobilized BSA, RSA and HSA targets were performed. Sso7d-RSA-13A8, Sso7d-BSA-2A4, Sso7d-BSA-2A6, and Sso7d-BSA-2A8 phages were purified, concentrated and 10^{10} particles for each variant were analyzed. None of the phage displaying Sso7d variants identified by panning against BSA or RSA showed any detectable binding to HSA (Figure 2-3). Sso7d-RSA-13A8 showed greater than 12-fold specificity for RSA over BSA. Sso7d-BSA-2A8 showed an approximately 9-fold specificity for BSA over RSA. ELISA signals for Sso7d-BSA-2A4 and Sso7d-BSA-2A6 binding to RSA were identical to those observed for both HSA-immobilized and blocked wells, suggesting specificities of at least 60-fold for these two variants. Significantly, for all Sso7d variants specificity arises without any pre-selection step for undesired affinity (e.g. negative selections against the opposing serum albumin protein). Taken together these results suggest that the rigid Sso7d scaffold is capable of highly specific interactions with closely related proteins.

2.2.6 Evaluation of selected neutravidin binding Sso7d variants:

Neutravidin (the deglycosylated form of egg white avidin) is structurally and functionally similar to the bacterial protein streptavidin, although the primary sequences of these two proteins share only 47% similarity. Both proteins exist as tetramers of 8-stranded β -barrel folds and bind

biotin with femtomolar affinity.^{62,63} The Sso7d-variants that bind neutravidin identified by phage display show some similarity to the sequence of the streptavidin-binding Sso7d variant identified by yeast display (Table 2-1). In particular, the amino acids on the binding face of β -strand 5, Ser 40, Ile 42 and Tyr 44, are identical in the neutravidin and streptavidin binders. The amino acids on the other two β -strands of the neutravidin binding Sso7d variants are quite different from the corresponding residues identified in the streptavidin binding sequence from the yeast library.

DNA sequencing of 9 clones found by screening the Sso7d phage library against neutravidin revealed 6 different but closely related sequences, identical at 7 of 9 amino acid positions. The consensus sequence includes Phe 23, Asn 25, His 30, Arg 32, Ser 40, Ile 42, and Tyr 44. Within the set of neutravidin-binding Sso7d variants, the amino acids at positions 21 and 28 exhibited some diversity, with aromatic, polar, and long chain alkyl amino acids identified. Position 21, the first variable residue in β -strand 3, varied between His, Leu and Phe. Greater variation in side chain size and composition was observed at position 28, which is on the leading edge of β -strand 4. Five different amino acids were selected in the nine sequenced clones.

Phage ELISA provided a qualitative evaluation of the binding affinity and specificity of the streptavidin (Sec-Sso7d-strep) and neutravidin (Sso7d-neutravidin-2D6) binding variants (Figure 2-4). Neither Sso7d variant binds to the opposing target. ELISA signals from the addition of 10^{10} - 10^{12} Sso7d-neutravidin-2D6 phages to wells containing passively adsorbed streptavidin were not above background. Likewise, the ELISA signals from the addition of 10^{10} - 10^{12} Sso7d-strep phages to wells containing passively adsorbed neutravidin were not above background. In our hands, the Sso7d streptavidin binding variant identified by yeast display does not bind streptavidin efficiently when expressed in the context of phage. Although ELISA utilizing the anti-His tag antibody to evaluate display levels suggested high level display of the Sec-Sso7d-strep variant (Figure 2-2),

addition of 10^{12} Sec-Sso7d-strep phage particles to a streptavidin-coated well gave rise to ELISA signal of only approximately 5-fold over background. Based on quantification of phages purified and concentrated from larger culture volumes, we estimate that on the order of 10^{10} phage particles are added to each well when using small volume supernatants for ELISA to identify binders. Sso7d-strep would not have been identified from the phage library as a streptavidin binder, as the signal was not significantly different than that for addition of Sso7d-strep phages to a blocked well. In contrast, 10^{10} Sso7d-neutravidin-2D6 phage particles added to a neutravidin-coated well gave rise to an ELISA signal of 67-fold over background. Signal from the supernatant ELISA used to identify 2D6 as a binder resulted in an overflow signal, and ELISA signal from 10^9 purified phages still resulted in signal more than 10-fold over background.

2.2.7 Evaluation of selected GFP binding Sso7d variants:

GFP has a β -barrel fold structurally similar to neutravidin, but larger in size (11 versus 8 strands). GFP is monomeric as opposed to the tetrameric organization of neutravidin. DNA sequencing of 10 clones identified as GFP binders by supernatant ELISA revealed 10 different protein sequences which exhibit strong consensus of amino acid types at 6 of 9 varied positions (Table 2-1). Position 21 contained small amino acids Ala and Ile. Position 23 showed a weak preference for hydrophobic residues Ile or Leu (5 of 10 clones). A strong consensus of amino acid size or type was not evident at position 25. Position 28 was either Asp or Asn in 8 of 10 clones. Position 30 favored polar residues His, Thr or Tyr (9 of 10 clones). Position 32 favored small amino acids Ala, Ser and Gly (9 of 10 clones). Positions 40 and 42 preferred small hydrophobic amino acids. A weak consensus for polar residues was evident at position 44.

2.2.8 Estimation of binding affinity by competitive ELISA (on-phage) and ITC (off-phage)

Purified phages were evaluated by ELISA to select a single clone for quantification of binding affinity. Phage particles displaying each of the sequenced Sso7d variants were evaluated in a series of phage ELISA experiments with decreasing concentrations of phage particles. The clone giving rise to the highest signal at lowest concentration was selected for further characterization. Sso7d-BSA-2A8, Sso7d-neutravidin-2D6, and Sso7d-GFP-6B9 were identified as the best binders of the selected clones for their respective target proteins. Those three Sso7d variants, along with Sso7d-RSA-13A8, were characterized further.

Competitive ELISA was used to estimate the binding affinity (K_D) for each of the selected clones (Figure 2-5). Conditions for the competitive ELISA were established by identifying the concentrations of both phage particle and immobilized target protein that resulted in 50% saturation of the ELISA signal. For the competitive ELISA, phage particles at a single pre-determined concentration were pre-incubated with increasing concentrations of the target protein prior to performing ELISA against immobilized target ⁵⁷.

The equilibrium dissociation constants of Sso7d-BSA-2A8 and Sso7d-RSA-13A8 were measured in the low micromolar range (3.5 μ M and 2.1 μ M, respectively, Table 2-3). Both the GFP and neutravidin binding variants displayed low nanomolar dissociation constants (6.2 nM and 36.1 nM, respectively, Table 2-3). The selection of tighter binding variants against GFP and neutravidin may be related to the similarity in folds between these proteins and a propensity of the slightly concave Sso7d binding face to recognize the convex surface of β -barrel folds.

The ability of Sso7d-BSA-2A8 and Sso7d-GFP-6B9 to bind their targets in the absence of the phage context was evaluated by isothermal titration calorimetry (ITC) (Figure 2-6). For most applications, binding molecules identified by phage display must retain the ability to recognize

their target once produced and deployed independent of the phage framework. Not infrequently, a decrease in binding affinity accompanies removal of the binding molecule from the phage context. Binding affinities determined by ITC for purified Sso7d-BSA-2A8 and Sso7d-GFP-6B9 were comparable to the on-phage values (Table 2-3), suggesting that the mutations to the binding face of the Sso7d scaffold required for specific binding do not alter the structural context of the overall protein.

2.2.9 Thermal stability of Sso7d variants on- and off-phage

High thermostability with relatively small size is one of the most appealing properties of Sso7d as a scaffold for developing specific binding proteins. Both M13 bacteriophage and the Sso7d scaffold protein are known to maintain their structure and function after exposure to elevated temperatures. Antibodies require careful handling to maintain their binding ability, and even minimal exposure to heat can greatly diminish their binding affinity. In order to evaluate the thermal stability of the Sso7d-BSA-2A8, Sso7d-RSA-13A8, Sso7d-neutravidin-2D6, and Sso7d-GFP-6B9 proteins, phage displaying each variant were exposed to temperatures between 50 °C and 90 °C for 30 minutes and assayed for retention of binding ability by phage ELISA. Aliquots of each phage sample were incubated on ice to serve as a control. Each of the Sso7d displaying phages maintained binding ability comparable to that of the control sample after exposure to temperatures of up to 70 °C for 30 minutes (Figure 2-7a). After 30 minutes at 80 °C, Sso7d-neutravidin-2D6 and Sso7d-RSA-13A8 lost the ability to bind to their target proteins. Some binding ability was retained for Sso7d-BSA-2A8 and Sso7d-GFP-6B9, but the ELISA signal was only a fraction of that observed after exposure to lower temperatures, 38% and 24% respectively. After prolonged exposure (up to 8 hours) to temperatures of 60 °C, both Sso7d-BSA-2A8 and Sso7d-GFP-6B9 phages retained significant binding ability (Figure 2-7b).

The melting temperature of each of the Sso7d variants identified by phage display was measured by differential scanning calorimetry (DSC) in the absence of the phage context (Table 2-4). Purified Sso7d variants maintained high melting temperatures similar to that of the parent protein and well above temperatures at which the structural stability of most antibodies would have been compromised. Compared to the 98 °C melting temperature reported for wild type Sso7d, each variant was slightly less stable, although the least thermostable variant, Sso7d-BSA-2A8 still had a melting temperature of nearly 83 °C^{64,65}. Wild type Sso7d is a model, two-state reversibly-folding protein, but each of the phage-selected Sso7d variants showed irreversible denaturation as evidenced by greatly reduced DSC endotherms for serial heating and cooling cycles. Similar folding defects were observed for some of the Sso7d variants selected by yeast display⁴⁷.

Sso7d-BSA-2A8 phage particles maintain their structural integrity as measured by the ability of the particles to infect receptive cells after challenges up to 80 °C for 30 minutes. The infectious titer of the particles did not decrease relative to the sample of phages held on ice after incubation at 50 °C, 60 °C, or 70 °C. Titer decreased by 10-fold after incubation at 80 °C. The infectious titer of Sso7d-BSA-2A8 phages incubated at 90 °C for 30 minutes was below 10⁴ cfu/mL, the detection limit of the titer assay.

2.3 Conclusions:

Antibodies have seen increasing roles as therapeutics, as *in vivo* and *in vitro* diagnostics, and as tools for laboratory research. However, antibodies are not ideal molecules for many applications because of their large size, multi-domain structure, poor stability, and high cost and difficulty of production. The drawbacks of traditional antibody-based reagents can be insurmountable in environments in which basic infrastructure for storage and deployment are lacking. The evolving needs for antibodies in biomedicine have fueled the continuing development of new scaffolds for

molecular recognition that exhibit improved physical properties. The demands of new applications translate into a set of desired molecular attributes that an engineered binding molecule should possess: small size, absence of disulfide bonds and post-translational modifications, high chemical and thermal stability, and the ability to be engineered to high affinity and specificity towards a given target. The burgeoning application of highly decorated bacteriophage particles as a platform for diagnostics, therapeutics, and other biotechnological applications requires additional compatibility of the scaffold protein with the phage coat architecture and particle assembly process. A largely unexplored but fertile source of new scaffolds are the protein domains of hyperthermophilic organisms.

We have described the optimization of phage display of the OB-fold protein Sso7d from *Sulfolobus solfataricus*. We produced a large Sso7d phage display library and employed it to isolate Sso7d variants that bind a variety of target proteins. On phage, the binding interactions have dissociation constants in the low nanomolar to low micromolar range. When the Sso7d derivatives were analyzed for binding in the absence of the phage particle, the K_{DS} observed with increased approximately 4-8-fold, as is often seen with binding molecules identified by phage display. The Sso7d variants retain their affinity for their target proteins when phage solutions are heated to 60 °C for extended periods of time. The melting temperatures for purified Sso7d derivatives isolated from the phage library are slightly lower than the native Sso7d sequence, but still very high. Taken together, our results suggest that the Sso7d scaffold will be a complementary addition to the range of non-antibody scaffold proteins that can be utilized in phage display.

2.4 Materials and Methods

All enzymes for cloning and mutagenesis were purchased from New England Biolabs (Ipswich, MA, USA). Unless otherwise noted, antibiotics were used at the following

concentrations: carbenicillin (carb) 50 µg/mL, tetracycline (tet) 30 µg/mL, chloramphenicol (cam) 5 µg/mL and kanamycin (kan) 50 µg/mL (Sigma-Aldrich, St. Louis, MO, USA). All liquid cultures were shaken at 225 rpm unless otherwise noted. All wash steps in ELISA and library panning are 300 µL. In this manuscript, all solutions containing phosphate-buffered saline (PBS) are made as 137 mM NaCl, 2.7 mM KCl, 12 M Na₂HPO₄ and 1.2 mM KH₂PO₄ at pH 7.4. All oligonucleotides were purchased from IDT (Coralville, IA, USA). All constructs were verified by sequencing (Proteomics and Metabolomics Facility, Colorado State University).

2.4.1 Construction of Sso7d gene and Sso7d-displaying phage vectors

The gene for the streptavidin-binding Sso7d variant (Sso7d-strep) identified by yeast display⁴⁷ was synthesized by sequential thermodynamically balanced inside out PCR (seqTBIO)⁵⁰ using Oligos 1-8 (Table S2-1, Appendix 1) which allowed the introduction of unique restriction sites for cloning and full control over codon usage. The Sso7d-strep gene product was verified by sequencing.

The Sso7d-strep TBIO product was inserted into Np3-TAT⁵¹ using XhoI and NheI to give TAT-Sso7d-strep. Oligos 10 and 11 were used to amplify the Sso7d-strep-p3 gene segment from TAT-Sso7d-strep and install 5' XhoI and 3' EcoRI sites for cloning into Np3-Sec to generate Sec-Sso7d-strep. Oligos 12 and 11 were used to amplify the Sso7d-strep-p3 gene segment from TAT-Sso7d-strep with 5' BssHII and 3' EcoRI sites to generate SRP-Sso7d-strep. The Sso7d-strep-p3 region of each phagemid was verified by sequencing (Table S2-2, Appendix 2).

2.4.2 Preparation of phage particles

Briefly, NEB 5-alpha F' *E. coli* (F' *proA*⁺*B*⁺ *lacI*^q Δ(*lacZ*)*M15* *zzf*:*Tn10* (Tet^R) / *fhuA2*Δ(*argF-lacZ*)*U169* *phoA* *glnV44* Φ80Δ(*lacZ*)*M15* *gyrA96* *recA1* *relA1* *endA1* *thi-1* *hsdR17*) harboring the Sso7d-containing phagemid were grown to an OD₆₀₀ of 0.5 at 37 °C in 2YT tet/carb/3% glucose.

Glucose was added to suppress expression of the Sso7d-variant-p3 fusion protein, to minimize growth discrepancies that may arise from the expression of Sso7d variants, and to maximize expression of the F pilus for helper phage infection⁶⁶. M13KO7 helper phage (multiplicity of infection, MOI = 20) were added to cells harboring the phagemids. Infected cultures were incubated at 37 °C for 30 min. Cells were pelleted by centrifugation for 5 min at 5000 ×g at room temperature and resuspended in 2YT tet/carb/kan without glucose to allow expression of the Sso7d-p3 fusion protein. Phage production proceeded overnight at 30 °C. Phage supernatants were prepared by centrifugation at 17,000 ×g to pellet the cells. For large volume cultures, phage were precipitated by decanting the supernatant from the pelleted cells into 1/5th volume polyethylene glycol (PEG) 8000/2.5 M NaCl in water². The phage pellet was resuspended in 1/100th of the original culture volume PBS. Typical phage yields were on the order of 10¹³-10¹⁴ particles from 100 mL of cell culture. Phage concentrations were determined spectrophotometrically using the following formula (Abs. 269 nm – Abs. 320 nm) x 6 x 10¹⁶ / plasmid size (bp)⁶⁷. For long-term storage, phage particles were frozen in 14% glycerol.

2.4.3 Phage enzyme-linked immunosorbent assay (ELISA)

Display of the Sso7d-strep variant was confirmed using phage ELISA. MaxiSorp[®] 96 ELISA plates (Nunc Inc., Rochester, NY, USA) were coated with approximately 10¹⁰ purified phages prepared from cells harboring either TAT-Sso7d-strep, SRP-Sso7d-strep, or Sec-Sso7d-strep phagemids at 4 °C overnight. After washing 3 times with PBS each well was blocked with blocking buffer (PBS with 4% milk powder) for 1 h at room temperature. Blocked wells were washed 6 times with PBS-Tween (PBST, PBS with 0.05% (v/v) Tween 20). Functional Sso7d-strep display was assayed immunohistochemically by incubation with anti-6x-His antibody HIS.H8 (1:1000 dilution) (Pierce, ThermoFisher Scientific, Waltham, MA, USA) 2 hours at room temperature.

After washing 6 times with PBST (300 μ L per wash), 100 μ L of rabbit anti-mouse IgG/horseradish peroxidase (HRP) conjugate (1:3000 dilution, Pierce) was added to the well and allowed to bind for 1 h. The wells were washed 6 times with PBST, then twice with PBS. Colorimetric assays were performed using 1-Step Ultra TMB ELISA Substrate (Pierce). The reaction was quenched by addition of 2 M sulfuric acid after 30 min incubation at room temperature. Absorbance at 450 nm was determined using a Biotek Synergy H1 microplate reader.

2.4.4 Library construction via Kunkel mutagenesis

The library template Sec-Sso7d-rArg was generated from Sec-Sso7d-strep using Kunkel mutagenesis with Oligo 13 and verified by sequencing.⁶⁸ The Kunkel mutagenesis procedure was adapted from Sidhu and Weiss². Briefly, cultures of CJ236 *E. coli* ($F\Delta(HindIII)::cat$ (Tra⁺ Pil⁺ Cam^R)/ *ung-1 relA1 dut-1 thi-1 spoT1 mcrA*) harboring the phagemid to be mutated were grown to an OD₆₀₀ of 0.5 and infected with M13KO7 at a MOI of 20. The infected culture was transferred into 2YT cam/carb/kan/0.25 μ g/mL uridine/3% glucose. Cultures were grown overnight at 30 °C. Phage particles were concentrated as described above. Single stranded uridine-enriched (ss dU) DNA was isolated using the QIAprep Spin M13 kit (Qiagen, Germantown, MD, USA).

The library was generated using 100 μ g Sso7d-strep-rArg DNA and a 10-fold molar excess of 5' phosphorylated Oligo 14. Following PCR spin kit cleanup and elution in ultrapure water, the library DNA was electroporated into SS320 cells ((MC1061 F', Lucigen) [F'*proAB+lacIqlacZ* Δ M15 *Tn10* (tet)] *hsdR mcrB araD139* Δ (*araABC-leu*)7679 *lacX74 galUgalK rpsL thi*), resulting in 1.3×10^{11} unique transformants.

After transformation, cells were infected with M13KO7 as described above and transferred to 1 L of 2YT media with 30 μ g/mL tet, 60 μ g/mL carb, and 10 μ g/mL kan and grown overnight at

30 °C. Phage particles were isolated by PEG precipitation. Aliquots of the library in PBS/14% glycerol were stored at -80 °C.

To determine the efficiency of mutagenesis, 95 single colonies were analyzed by colony PCR and a restriction test. Single colonies were used as a template for PCR amplification using Oligo 15 and Oligo 16. Crude PCR products were subjected to XbaI digestion and analyzed by agarose gel electrophoresis.

2.4.5 Selection of binding variants

Sso7d binding variants were selected by panning the phage library against immobilized protein in 96 well plates. Target proteins BSA (EMD Millipore, Danvers, MA, USA), neutravidin (Thermo Scientific), RSA (Sigma-Aldrich) were used as supplied and sfGFP was expressed and purified in-house. A GFP variant including an N-terminal hexahistidine tag and a C-terminal sortase recognition sequence (LPETGG) was cloned into a modified pQE40 protein expression vector. pQE40 had been modified in the laboratory to include an additional copy of the LacIQ gene from pBbB1A-GFP (Addgene plasmid # 35340, a gift from Jay Keasling ⁶⁹) and the phage origin of replication from M13KO7. Proteins were coated to MaxiSorp[®] 96 well ELISA plates (Nunc Inc.) in coating buffer (1 µg protein in 0.1 M Na₂CO₃, 0.1 M NaHCO₃, pH 9.6) at 4 °C overnight. After 3 washes, each well was blocked and washed as described above for ELISA. In the first round of selection, 5 target coated wells were used. In subsequent rounds, 4 wells were employed. For the final round of selection against BSA and neutravidin 2 wells were used. Approximately 10¹³ library phages were mixed 1:1 with blocking buffer and added to each well (50 µL of phage amplified from the previous round plus 50 µL blocking buffer). After 2 hours at room temperature, unbound phage particles were discarded, and the wells were washed. Wash stringencies increased in each round of selection. In the first round of selection, wells were washed 3 times with PBS. Additional

washes with PBST were added for subsequent rounds of panning. The second round of selection included 3 PBST washes followed by 2 PBS washes. The third round of panning included 6 PBST washes followed by 2 PBS washes. The fourth round of selection (if performed) included 10 PBST washes followed by 2 PBS washes. Bound phages were eluted with 100 μ L of 100 mM glycine HCl, pH 2.2 for 25 min at room temperature. Eluent from multiple wells was combined and neutralized using 1.5 M Tris, pH 8.8. Neutralized phages were titered (see below) to determine the output of each round of selection (10 μ L of the pooled eluent). Collected phages were propagated in 5 mL of NEB 5-alpha F' cells. After 30 min of incubation without shaking followed by 30 min of incubation with shaking, the entire culture was added to 15 mL 2YT tet/carb with 3% glucose. When the OD₆₀₀ of the 15 mL culture reached 0.5, M13KO7 were added (MOI=50). Infected cells were spun and resuspended in 100 mL of 2YT tet/carb/kan media without glucose, followed by overnight amplification at 30 °C. Phages isolated by PEG precipitation and concentrated 100-fold were used for the following round of selection. Amplified phages were exposed to a blocked well during rounds 3 and 4 of selection to evaluate non-target protein specific phage retention.

2.4.6 Phage titration

To determine the phage output from each round, a 10-fold dilution series of the eluted phages was used to infect NEB 5-alpha F' cells grown to OD₆₀₀ 0.5 at 37 °C. Phage infection was allowed to proceed for 30 min at 37 °C. At least 3 10 μ L spots for each phage dilution were plated on LB agar plates and incubated at 37 °C overnight. The titer of phage was calculated from the dilution with the highest number of countable colonies.

2.4.7 Binding analysis of individual clones

Individual clones from the final round of selection were picked into 150 μ L of 2YT tet/carb with 3% glucose in a 96 well plate incubated at 37 °C overnight. The following day, 5 μ L of each

overnight culture was transferred to 150 μ L of 2YT tet/carb with 3% glucose in a fresh 96 well plate and regrown to OD₆₀₀ of 0.5 (generally about 2.5 hours). M13KO7 was added to each well at a MOI of 20. After incubation at 37 °C for 30 min, the cells were pelleted by centrifugation at 5000 $\times g$ at room temperature and resuspended in 2YT tet/carb/kan without glucose. The 96 well plate was incubated at 30 °C overnight. Cells were pelleted at 17,000 $\times g$, and the phage supernatants were used in an ELISA assay to characterize the binding affinity of the Sso7d variants.

ELISA plates were prepared as described above for panning the library. 50 μ L of the phage supernatants were mixed 1:1 with blocking buffer and added to individual wells. After 2 hours at room temperature, wells were washed 6 times with PBST. Bound phage were detected using an HRP-conjugated anti-M13 antibody (100 μ L, 1:5000 dilution in blocking buffer) (GE Healthcare, Wauwatosa, WI, USA). After 1 h at room temperature, wells were washed 6 times with PBST and twice with PBS and visualized as described above. Clones displaying ELISA signal against the target protein that was 3-fold or greater over background (blocked wells) were considered hits. Across all target proteins, only 2 hits identified for neutravidin did not exhibit signal at least 30-fold over background. Generally, plasmid DNA was isolated from 10 clones that bound each target protein and sequenced.

2.4.8 Competitive ELISA to determine binding affinity

A titration curve examining ELISA signal at different dilutions of both immobilized target protein and purified phage was generated to determine the conditions which led to 50% saturation of the ELISA signal. Once the concentration for immobilized target protein was determined, wells for competitive ELISA were prepared as described above for the phage supernatant ELISA (Table 2-5). Phage particles were pre-equilibrated with concentrations of target protein across five orders

of magnitude in blocking buffer for 2 hours at room temperature. A sample of phage that was not pre-equilibrated with target served as a reference for the 100% phage bound absorbance. 100 μ L of equilibrated phage solution were added to each ELISA well and allowed to bind for 15 min. ELISA was performed and visualized as described above. Absorbances at 450 nm for each concentration of competitor were normalized to the absorbance for the phage sample with no competitor. Data were plotted as relative A_{450} versus concentration of competitor. Dissociation constants were determined by fitting the data using a 4-parameter logistic equation (Appendix 3). Measurements were performed in triplicate for each experiment. Plots shown in Figure 2-5 are representative of a single experiment. K_{DS} reported in Table 2-3 represent the average of at least three independent experiments.

2.4.9 Heat tolerance of Sso7d variants on phage

The heat tolerance of Sso7d displaying phages was evaluated by ELISA. Phage stock solutions were incubated at temperatures between 50 and 90 °C for 30 min or at 60 °C between 1 and 8 hours. Control samples were incubated on ice for the duration of each experiment. After exposure to elevated temperatures, all phage were kept at room temperature for 1 hour. Phage solutions were diluted 100- to 300-fold in blocking buffer prior to performing ELISA as described above. The concentration of target protein coated to each well ranged from 125-1000 ng, depending on the concentrations that resulted in good ELISA signal during binding analysis of individual clones.

2.4.10 Preparation and purification of Sso7d variants off phage

Oligos 17 and 18 (Table S2-1, Appendix 1) were used to amplify Sso7d variants for cloning into a pET23b (Novagen/EMD Millipore) protein expression vector using NdeI and HindIII restriction sites. pET23b-Sso7d variants were verified by sequencing. Vectors were transformed into BL21(DE3) pLysS cells (F^- , *ompT*, *hsdS_B* (r_B^- , m_B^-), *dcm*, *gal*, λ (DE3), pLysS, Cm^r , Promega,

Madison, WI, USA) by electroporation. 1.2 L of LB cam 30/carb 50 were inoculated from an overnight culture and incubated at 37 °C. Protein expression was induced when the culture reached an OD₆₀₀ of 0.5 by addition of IPTG to a final concentration of 0.5 mM. Cells were grown overnight at 30 °C. Cells were harvested by centrifugation and lysed by sonication (Qsonica, Newtown, CT, USA). Protein was purified from the clarified lysate by Ni-NTA affinity chromatography on a BioLogic LP FPLC (BioRad, Hercules, CA, USA). Eluted proteins were dialyzed against PBS and quantified by Micro BCA protein assay (Pierce). Typically, yields of each Sso7d variant were between 30 and 40 mg/L with purities greater than 90%.

2.4.11 Isothermal titration calorimetry (ITC)

ITC experiments were performed on a Nano ITC 2G model 5300 (TA Instruments, New Castle, DE, USA) at 25 °C in PBS. Both the target protein and Sso7d variant were buffer exchanged into PBS to ensure that both proteins were in identical environments. All samples were degassed under vacuum. The reference cell was washed 5 times and loaded with degassed water. The sample cell and syringe were rinsed 5 times with degassed PBS. The instrument was operated in overflow mode. The concentration of each Sso7d variant was 10 to 15 times higher than that of the target protein. After auto-equilibration of the ITC, 2 µL of Sso7d variant was injected in the first injection, followed by 30 8 µL injections with a 5 min equilibration between each injection. The sample cell was mixed by stirring at 300 rpm. Data were collected at 1 second intervals using the ITCRun software (TA Instruments) and analyzed using the independent model in NanoAnalyze (TA Instruments). Signal from a background titration monitoring addition of buffer to the protein sample was subtracted from the protein titration. Each sample titration was performed three times.

2.4.12 Differential scanning calorimetry (DSC)

DSC experiments were performed on a Nano DSC model 6300 (TA Instruments). Sso7d variants were buffer exchanged into PBS and diluted to a final concentration of approximately 0.5 mg/mL. All samples were degassed under vacuum. Both the reference cell and the sample cell were rinsed with degassed PBS 3 times. Degassed PBS was used for background buffer scans. Two scan cycles, each consisting of one heating and one cooling cycle from 5 °C to 110 °C (1 °C/min) were performed for each Sso7d variant. Data were recorded in DSC Run software (TA Instruments) and analyzed using the two state scale model in NanoAnalyze (TA Instruments). Each Sso7d variant was scanned twice.

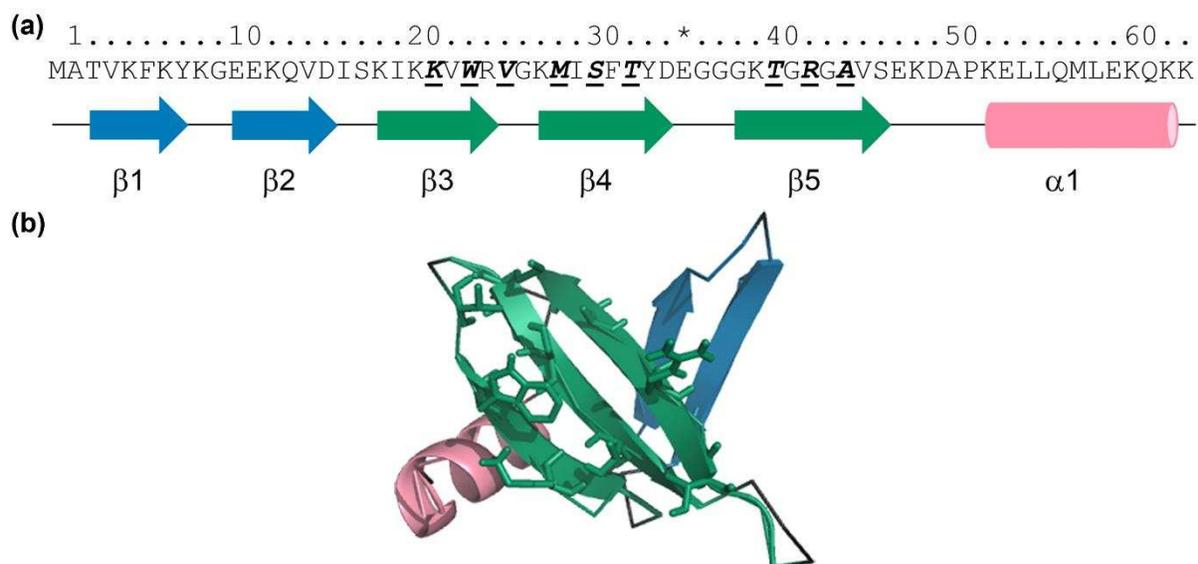


Figure 2-1. Sequence and structure of Sso7d from *Sulfolobus solfataricus*. (a) Primary sequence of Sso7d. Secondary structures are indicated beneath the sequence. The 9 bold, underlined amino acids indicate the diversified positions in the library. The variant used for library construction included an E35L mutation (indicated by *) to abolish Sso7d ribonuclease activity. (b) Sso7d with secondary structural elements color coded to match the cartoon in part a. Side chains shown as sticks indicate positions of diversification. (PDB code: **1BNZ**)

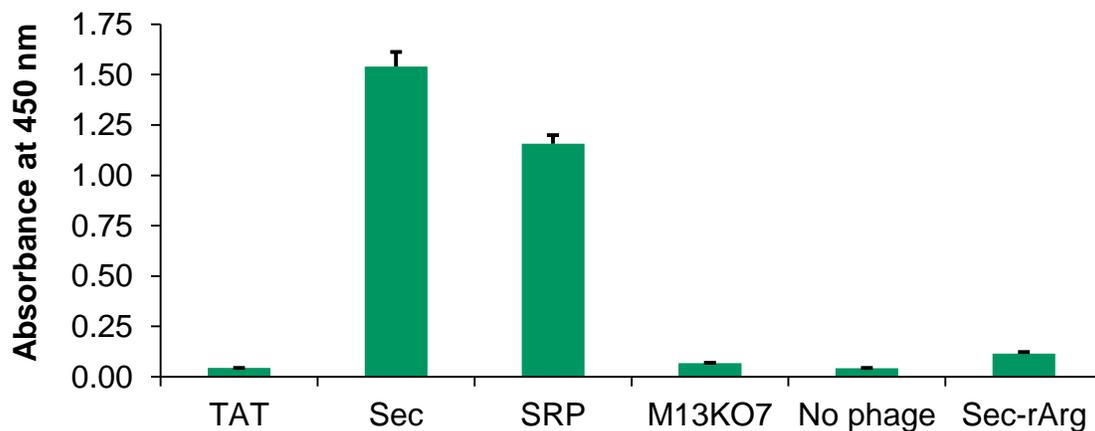


Figure 2-2. Effect of secretory pathway on the phage display level of the Sso7d streptavidin binding variant. A_{450} readings from phage ELISAs using anti-hexahistidine tag antibody to detect display of the Sso7d-strep-6His-p3 protein fusion on phage particles are shown for each construct, which is identified above by the translocation pathway targeted. Approximately 10^{10} particles were immobilized in each well, except for Sec-Sso7d-rArg for which approximately 10^{13} particles were immobilized. All data are represented as the mean \pm standard deviation ($n = 3$).

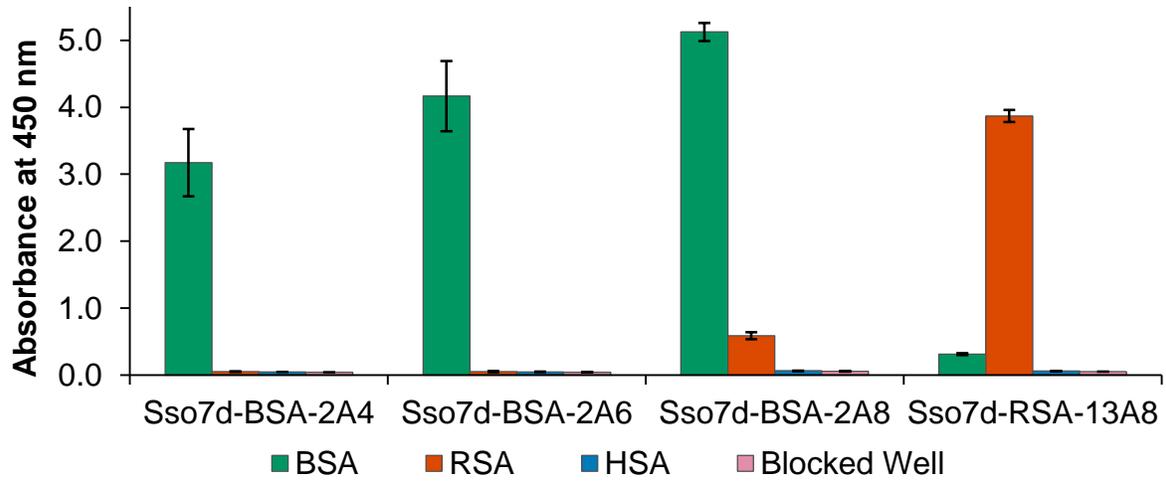


Figure 2-3. Specificity of selected Sso7d variant-displaying phage for related serum albumin proteins. BSA, RSA or HSA were immobilized in wells for ELISA. Bound phage were detected using an anti-M13 antibody. In all instances, approximately 10^{10} particles were added to each well. All data are represented as the mean \pm standard deviation ($n = 3$).

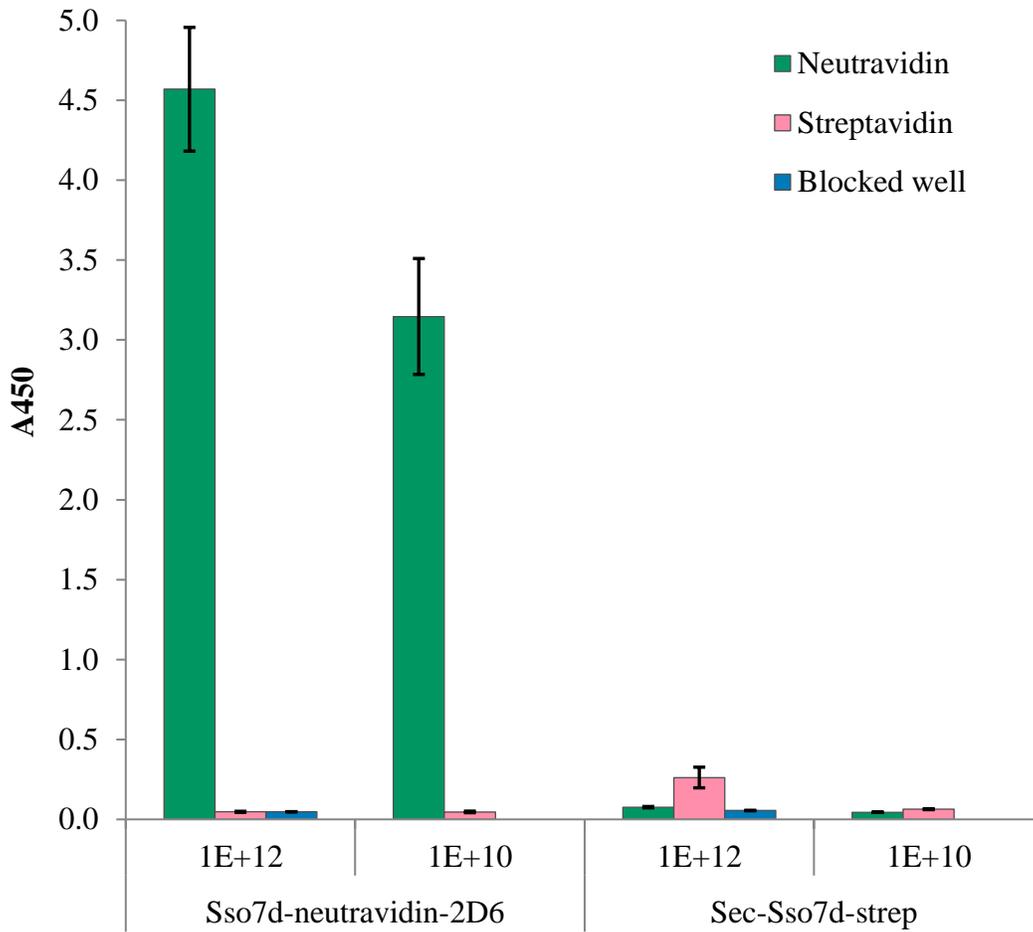


Figure 2-4. Specificity of Sso7d variant-displaying phage for neutravidin and streptavidin. Neutravidin and streptavidin were immobilized in wells for ELISA. Bound phage were detected using an anti-M13 antibody. Approximately 10^{12} or 10^{10} particles were added to each well. All data are represented as the mean \pm standard deviation ($n = 3$).

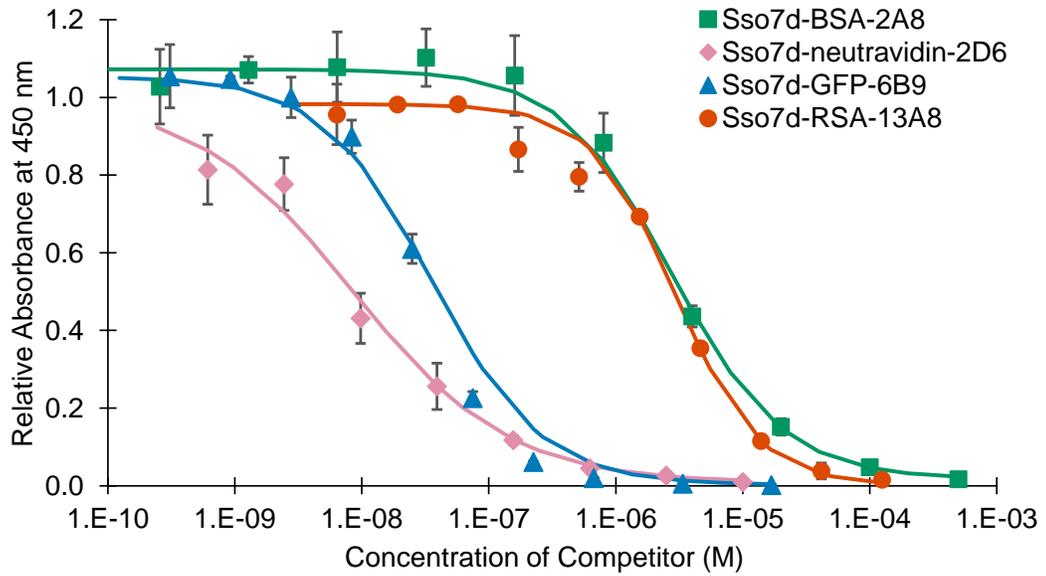


Figure 2-5. Normalized competitive ELISA data for selected Sso7d variants. Phages were preincubated with ten different concentrations of soluble target protein prior to incubation in wells containing the same target protein. Anti-M13 antibody was used to detect phages that were retained in each well. All data are represented as the mean \pm standard deviation ($n = 3$).

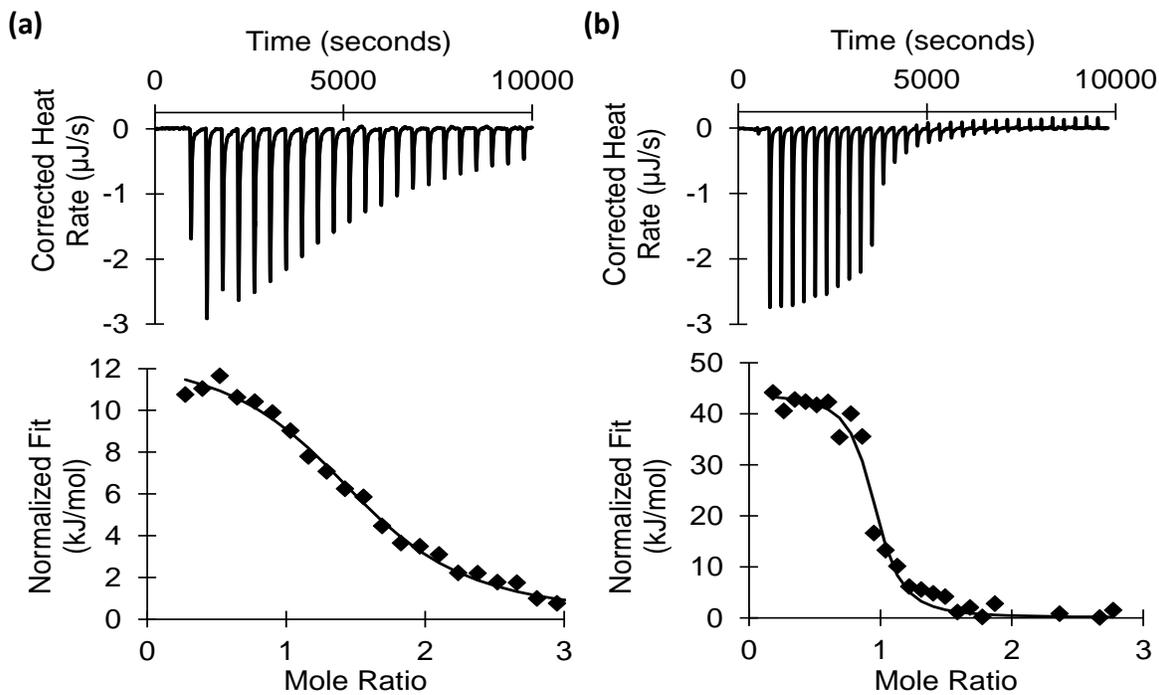


Figure 2-6. Affinities of purified Sso7d variants determined by isothermal titration calorimetry. a) ITC titrations of Sso7d-BSA-2A8 into BSA. Upper panel contains raw heat signals; bottom panel shows the integrated heats and fit using single site binding model, $K_D = 15100 \pm 2000$ nM. b) ITC titrations of Sso7d-GFP-6B9 into sfGFP. Upper panel contains raw heat signals; bottom panel shows the integrated heats and fit using single site binding model, $K_D = 298 \pm 48$ nM.

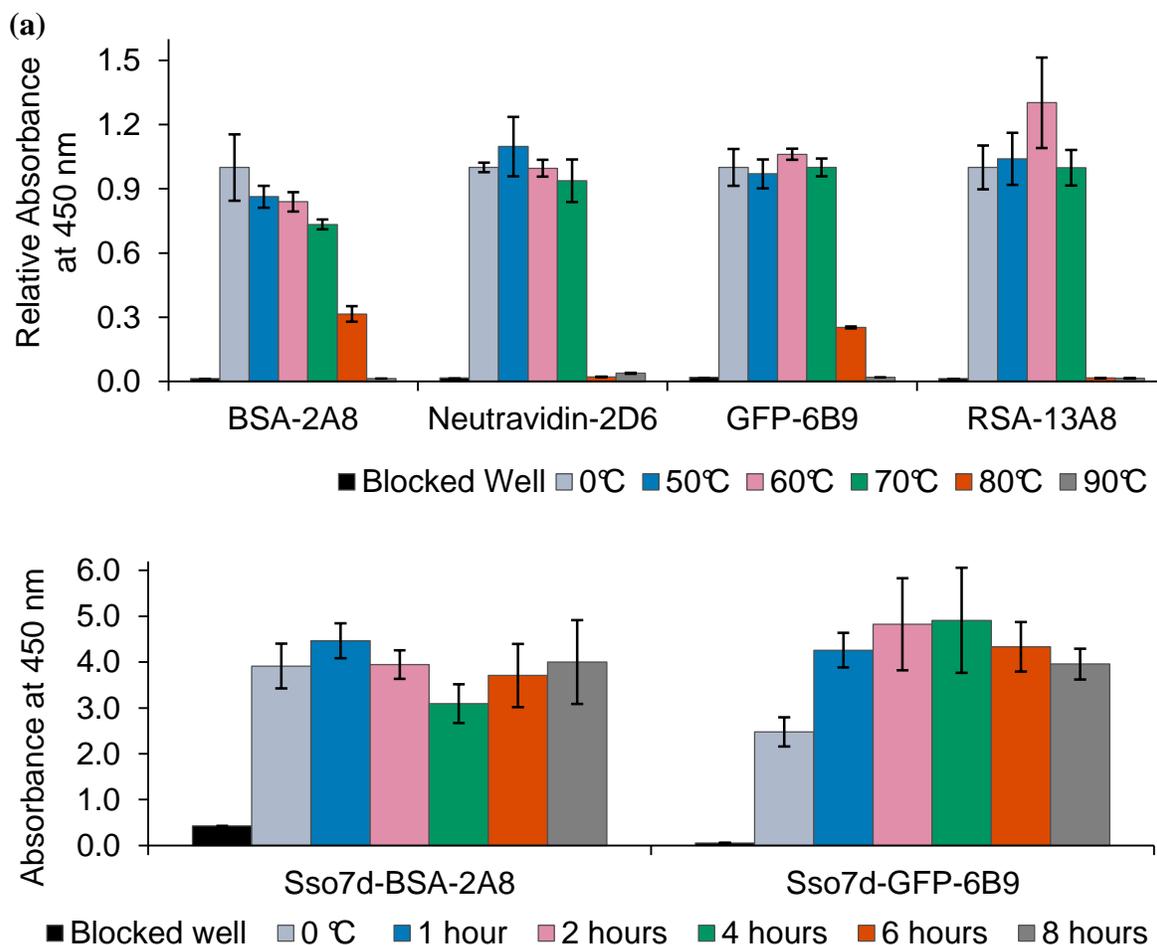


Figure 2-7. Retention of binding affinity after exposure of Sso7d variant displaying phages to elevated temperatures. (a) ELISA was performed after 30 minute exposure to temperatures between 50 °C and 90 °C. ELISA signal for each phage variant is normalized to the average A_{450} determined for the sample incubated at 0 °C. (b) ELISA was performed after exposure to 60 °C for times between 1 hour and 8 hours. Bound phage were detected using an anti-M13 antibody. In all instances, approximately 10^{10} particles were added to each well. All data are represented as the mean \pm standard deviation ($n = 3$).

Table 2-1. Protein sequences for positions 20-45 of wild type Sso7d and variants selected from the library.

	20 30 40
Wild type Sso7d^a	KKVWRVVGKMI SFTYDEGGGKTGRGAV
Sso7d streptavidin (isolated from yeast display)	WIVARDGKWI D FSYDLGGGKSGIGYV
Sso7d rare arginine template	WSRRRRGKRRRFRYDLGGGKSGIGYV
Target	
BSA (10 clones sequenced)	
Sso7d-BSA-2A4 (6) ^b	WIVDRDGKFI VFIYDLGGGKPGTGFV
Sso7d-BSA-2A6 (3)	WVVDRYGKFI LFYDLGGGKPGTGYV
Sso7d-BSA-2A8 (1)	WYVVRDGKFI NFIYDLGGGKSGYGGV
RSA (3 clones sequenced)	
Sso7d-RSA-13A8 (3)	WYVNRFGKNI HF LYDLGGGKSGHGYV
Neutravidin (9 clones sequenced)	
Sso7d-Neutravidin-2D6 (3)	WHVFRDGKYI HF RYDLGGGKSGIGYV
Sso7d-Neutravidin-2D8 (1)	WHVFRDGKLI HF RYDLGGGKSGIGYV
Sso7d-Neutravidin-2E1 (1)	WLVFRDGKFI HF RYDLGGGKSGIGYV
Sso7d-Neutravidin-2E2 (2)	WLVFRDGKTI HF RYDLGGGKSGIGYV
Sso7d-Neutravidin-2E6 (1)	WLVFRDGKNI HF RYDLGGGKSGIGYV
Sso7d-Neutravidin-2D10 (1)	WVFRDGKYI HF RYDLGGGKSGIGYV
GFP (10 clones sequenced)	
Sso7d-GFP-6A2 (1)	WIVARYGKDI HF GYDLGGGKVGTGYV
Sso7d-GFP-6A7 (1)	WAVLRTGKDI TF SYDLGGGKAGLGIV
Sso7d-GFP-6A8 (1)	WIVHRVGKLI HF AYDLGGGKIGGGYV
Sso7d-GFP-6A9 (1)	WAVIRDGKNI FF TYDLGGGK PGLGHV
Sso7d-GFP-6B2 (1)	WVVG RFGKDI HF AYDLGGGKIGAGSV
Sso7d-GFP-6B3 (1)	WAVIRDGKDI TF SYDLGGGKVGLGSV
Sso7d-GFP-6B4 (1)	WIVGRHGKNI HF GYDLGGGKLG TGFV
Sso7d-GFP-6B7 (1)	WAVHRFGKYI HF AYDLDDDKIGSGHV
Sso7d-GFP-6B8 (1)	WAVIRNGKNI YF SYDLGGGKAGLGNV
Sso7d-GFP-6B9 (1)	WAVIRDGKDI YF SYDLGGGKAGLGRV

^a The glutamic acid residue at position 35 of Sso7d was found to contribute to the protein's ribonuclease activity. The E35L variant lacks ribonuclease activity; both the yeast and phage libraries include Leu at position 35.

^b The number in parenthesis next to each clone denotes the frequency of occurrence of that protein sequence in the selection of clones characterized.

Table 2-2. Sso7d library panning details for four target proteins.

Target	Number of rounds of selection	Output (cfu) against target	Output (cfu) against blocked well	Enhancement	ELISA-positive clones
BSA	4	2×10^8	2×10^3	1×10^5	23/23
RSA	3	5×10^4	4×10^3	12.5	3/48
Neutravidin	4	7×10^8	2×10^3	3.5×10^5	22/23
GFP	3	4×10^8	1×10^4	4×10^4	18/24

Output and enhancement calculations are those for the final round of selection. Phage ELISA for individual clone analysis was performed after the final round of selection.

Table 2-3. Apparent equilibrium dissociation constants (K_D) for selected Sso7d variants on- and off-phage.

Sso7d Variant	K_D (on-phage)	K_D (off-phage)	Ratio
Sso7d-BSA-2A8	3500 \pm 600 nM	15100 \pm 2000 nM	4.3x
Sso7d-RSA-13A8	2100 \pm 900 nM	-----	-----
Sso7d-neutravidin-2D6	6.2 \pm 1.7 nM	-----	-----
Sso7d-GFP-6B9	36.1 \pm 0.8 nM	298 \pm 48 nM	8.3x

Data are presented as mean \pm standard deviation ($n = 3$ for on-phage competitive ELISA, $n = 3$ for ITC).

Table 2-4. Thermostability of Sso7d variants.

Variant	T_m (°C)
Sso7d-wt	98
Sso7d-BSA-2A8	82.7, 82.6
Sso7d-RSA-13A8	88.9, 87.4
Sso7d-neutravidin-2D6	94.0, 93.8
Sso7d-GFP-6B9	86.1, 87.1

Analysis was performed for 2 samples of each Sso7d variant; each T_m is presented.

Table 2-5 Conditions for competitive ELISA

Target Protein and Sso7d Variant Pair	Protein Coated to Each Well	Dilution Factor of the Phage Stock	Concentration Range of Competitor Protein
BSA and 2A8	250 ng	160	2.6×10^{-10} M to 5×10^{-4} M
RSA and 13A8	500 ng	103	2.2×10^{-10} M to 8.4×10^{-5} M
Neutravidin and 2D6	31 ng	80	6.1×10^{-10} M to 1×10^{-5} M
GFP and 6B9	125 ng	88	6.4×10^{-9} M to 1.3×10^{-4} M

Target proteins were diluted 2-fold and concentrations ranged from 16 ng to 2 μ g immobilized per well. Phages were diluted 2-fold and concentrations ranged from approximately 10^7 to 10^9 particles per well. Purified phage were incubated in each well for 15 min to allow binding. Wells were washed and bound phages were detected as described. A plot of absorbance at 450 nm versus phage dilution was generated for each target protein concentration. The phage dilution and target protein concentration which gave rise to 50% saturation of ELISA signal were used to set up the competitive ELISA.

References

1. Kay, B.K., Winter, J. & McCafferty, J. (1996). *Phage display of peptides and proteins: a laboratory manual*.
2. Sidhu, S.S. & Weiss, G.A. (2004). Constructing phage display libraries by oligonucleotide-directed mutagenesis. *Phage Display: A practical approach* **2**, 27–41
3. Udit, A.K., Brown, S., Baksh, M.M. & Finn, M.G. (2008). Immobilization of bacteriophage Qbeta on metal-derivatized surfaces via polyvalent display of hexahistidine tags. *J. Inorg. Biochem.* **102**, 2142–6
4. Mao, C., Solis, D.J., Reiss, B.D., Kottmann, S.T., Sweeney, R.Y., Hayhurst, A., Georgiou, G., Iverson, B. & Belcher, A.M. (2004). Virus-based toolkit for the directed synthesis of magnetic and semiconducting nanowires. *Science* **303**, 213–217
5. Arter, J.A., Taggart, D.K., McIntire, T.M., Penner, R.M. & Weiss, G.A. (2010). Virus-PEDOT nanowires for biosensing. *Nano letters* **10**, 4858–4862
6. Yoo, P.J., Nam, K.T., Qi, J., Lee, S.-K., Park, J., Belcher, A.M. & Hammond, P.T. (2006). Spontaneous assembly of viruses on multilayered polymer surfaces. *Nature materials* **5**, 234–240
7. Sanghvi, A.B., Miller, K.P.-H., Belcher, A.M. & Schmidt, C.E. (2005). Biomaterials functionalization using a novel peptide that selectively binds to a conducting polymer. *Nature materials* **4**, 496–502
8. Suthiwangcharoen, N., Li, T., Li, K., Thompson, P., You, S. & Wang, Q. (2011). M13 bacteriophage-polymer nanoassemblies as drug delivery vehicles. *Nano Research* **4**, 483–493
9. Yoo, S.Y., Merzlyak, A. & Lee, S.-W. (2011). Facile growth factor immobilization platform based on engineered phage matrices. *Soft Matter* **7**, 1660–1666
10. Steinmetz, N.F., Ablack, A.L., Hickey, J.L., Ablack, J., Manocha, B., Mymryk, J.S., Luyt, L.G. & Lewis, J.D. (2011). Intravital imaging of human prostate cancer using viral nanoparticles targeted to gastrin-releasing Peptide receptors. *Small* **7**, 1664–72
11. Mao, J.Y., Belcher, A.M. & Van Vliet, K.J. (2010). Genetically engineered phage fibers and coatings for antibacterial applications. *Advanced Functional Materials* **20**, 209–214
12. Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317

13. Markland, W., Roberts, B.L., Saxena, M.J., Guterman, S.K. & Ladner, R.C. (1991). Design, construction and function of a multicopy display vector using fusions to the major coat protein of bacteriophage M13. *Gene* **109**, 13–9
14. Gao, C., Mao, S., Lo, C.H., Wirsching, P., Lerner, R.A. & Janda, K.D. (1999). Making artificial antibodies: a format for phage display of combinatorial heterodimeric arrays. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6025–30
15. Jespers, L.S., Messens, J.H., Keyser, A.D., Eeckhout, D., Brande, I.V.D., Gansemans, Y.G., Lauwereys, M.J., Vlasuk, G.P. & Stanssens, P.E. (1995). Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. *Bio/Technology* **13**, 378–382
16. Sweeney, R.Y., Park, E.Y., Iverson, B.L. & Georgiou, G. (2006). Assembly of multimeric phage nanostructures through leucine zipper interactions. *Biotechnology and bioengineering* **95**, 539–545
17. Nam, K.T., Peelle, B.R., Lee, S.-W. & Belcher, A.M. (2004). Genetically driven assembly of nanorings based on the M13 virus. *Nano letters* **4**, 23–27
18. Huang, Y., Chiang, C.-Y., Lee, S.K., Gao, Y., Hu, E.L., Yoreo, J.D. & Belcher, A.M. (2005). Programmable assembly of nanoarchitectures using genetically engineered viruses. *Nano letters* **5**, 1429–1434
19. Nam, K.T., Kim, D.-W., Yoo, P.J., Chiang, C.-Y., Meethong, N., Hammond, P.T., Chiang, Y.-M. & Belcher, A.M. (2006). Virus-enabled synthesis and assembly of nanowires for lithium ion battery electrodes - Supporting. *Science* **312**, 885–888
20. Merzlyak, A., Indrakanti, S. & Lee, S.-W. (2009). Genetically engineered nanofiber-like viruses for tissue regenerating materials. *Nano letters* **9**, 846–852
21. DePorter, S.M. & McNaughton, B.R. (2014). Engineered M13 bacteriophage nanocarriers for intracellular delivery of exogenous proteins to human prostate cancer cells. *Bioconjugate chemistry* **25**, 1620–1625
22. Kehoe, J.W. & Kay, B.K. (2005). Filamentous phage display in the new millennium. *Chemical reviews* **105**, 4056–4072
23. Sidhu, S.S., Weiss, G.A. & Wells, J.A. (2000). High copy display of large proteins on phage for functional selections. *J. Mol. Biol.* **296**, 487–95
24. Weiss, G.A., Wells, J.A. & Sidhu, S.S. (2000). Mutational analysis of the major coat protein of M13 identifies residues that control protein display. *Protein Science* **9**, 647–654
25. Holliger, P., Prospero, T. & Winter, G. (1993). “Diabodies”: small bivalent and bispecific antibody fragments. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6444–8

26. McGuinness, B.T., Walterz, G., FitzGerald, K., Schuler, P., Mahoney, W., Duncan, A.R. & Hoogenboom, H.R. (1996). Phage diabody repertoires for selection of large numbers of. *Nature biotechnology* **14**,
27. Binz, H.K. & Plückthun, A. (2005). Engineered proteins as specific binding reagents. *Current opinion in biotechnology* **16**, 459–469
28. Gilbreth, R.N. & Koide, S. (2012). Structural insights for engineering binding proteins based on non-antibody scaffolds. *Curr. Opin. Struct. Biol.* **22**, 413–20
29. Bradbury, A.R. & Plückthun, A. (2015). Getting to reproducible antibodies: the rationale for sequenced recombinant characterized reagents. *Protein Engineering Design and Selection* **28**, 303–305
30. Koide, A., Bailey, C.W., Huang, X. & Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. *Journal of molecular biology* **284**, 1141–1151
31. Nguyen, V.K., Desmyter, A., Muyldermans, S. & others (2001). Functional heavy-chain antibodies in Camelidae. *Advances in immunology* **79**, 261–296
32. Binz, H.K., Stumpp, M.T., Forrer, P., Amstutz, P. & Plückthun, A. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *Journal of molecular biology* **332**, 489–503
33. Nord, K., Gunneriusson, E., Ringdahl, J., Ståhl, S., Uhlén, M. & Nygren, P.-Å. (1997). Binding proteins selected from combinatorial libraries of an α -helical bacterial receptor domain. *Nature biotechnology* **15**, 772–777
34. Tiede, C., Tang, A.A., Deacon, S.E., Mandal, U., Nettleship, J.E., Owen, R.L., George, S.E., Harrison, D.J., Owens, R.J., Tomlinson, D.C. & others (2014). Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications. *Protein Engineering Design and Selection* **27**, 145–155
35. Binz, H.K., Amstutz, P. & Plückthun, A. (2005). Engineering novel binding proteins from nonimmunoglobulin domains. *Nature biotechnology* **23**, 1257–1268
36. Bloom, J.D., Labthavikul, S.T., Otey, C.R. & Arnold, F.H. (2006). Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5869–74
37. Mong, S.K., Vinogradov, A.A., Simon, M.D. & Pentelute, B.L. (2014). Rapid total synthesis of DARPin pE59 and barnase. *ChemBioChem* **15**, 721–733
38. Hussain, M., Gera, N., Hill, A.B. & Rao, B.M. (2013). Scaffold diversification enhances effectiveness of a superlibrary of hyperthermophilic proteins. *ACS Synth Biol* **2**, 6–13

39. Edmondson, S.P. & Shriver, J.W. (2000). DNA binding proteins Sac7d and Sso7d from *Sulfolobus*. *Methods in enzymology* **334**, 129–145
40. McCrary, B.S., Edmondson, S.P. & Shriver, J.W. (1996). Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d. *J. Mol. Biol.* **264**, 784–805
41. McCrary, B.S., Bedell, J., Edmondson, S.P. & Shriver, J.W. (1998). Linkage of protonation and anion binding to the folding of Sac7d. *Journal of molecular biology* **276**, 203–224
42. Edmondson, S. & Shriver, J. (1998). Solution structure of the thermophile Sac7d protein and Sac7d-DNA complexes. *BIOPHYSICAL JOURNAL* **74**, A67–A67
43. Gao, Y. G., Su, S. Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B.S., Edmondson, S.P., Shriver, J.W. & Wang, A.H. (1998). The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. *Nature Structural & Molecular Biology* **5**, 782–786
44. Murzin, A.G. (1993). OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *The EMBO journal* **12**, 861
45. Krehenbrink, M., Chami, M., Guilvout, I., Alzari, P.M., Pécorari, F. & Pugsley, A.P. (2008). Artificial binding proteins (Affitins) as probes for conformational changes in secretin PulD. *J. Mol. Biol.* **383**, 1058–68
46. Gera, N., Hill, A.B., White, D.P., Carbonell, R.G. & Rao, B.M. (2012). Design of pH sensitive binding proteins from the hyperthermophilic Sso7d scaffold. *PLoS ONE* **7**, e48928
47. Gera, N., Hussain, M., Wright, R.C. & Rao, B.M. (2011). Highly stable binding proteins derived from the hyperthermophilic Sso7d scaffold. *J. Mol. Biol.* **409**, 601–16
48. Mouratou, B., Schaeffer, F., Guilvout, I., Tello-Manigne, D., Pugsley, A.P., Alzari, P.M. & Pecorari, F. (2007). Remodeling a DNA-binding protein as a specific in vivo inhibitor of bacterial secretin PulD. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 17983–8
49. Hoover, D.M. & Lubkowski, J. (2002). DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic acids research* **30**, e43–e43
50. Gao, X., Yo, P., Keith, A., Ragan, T.J. & Harris, T.K. (2003). Thermodynamically balanced inside-out (TBIO) PCR-based gene synthesis: a novel method of primer design for high-fidelity assembly of longer gene sequences. *Nucleic acids research* **31**, e143–e143
51. Velappan, N., Fisher, H.E., Pesavento, E., Chasteen, L., D'Angelo, S., Kiss, C., Longmire, M., Pavlik, P. & Bradbury, A.R. (2010). A comprehensive analysis of filamentous phage

- display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins. *Nucleic acids research* **38**, e22–e22
52. Rakonjac, J., Bennett, N.J., Spagnuolo, J., Gagic, D. & Russel, M. (2011). Filamentous bacteriophage: biology, phage display and nanotechnology applications. *Curr Issues Mol Biol* **13**, 51–76
 53. Luirink, J., Heijne, G. von, Houben, E. & Gier, J.-W. de (2005). Biogenesis of inner membrane proteins in *Escherichia coli*. *Annu. Rev. Microbiol.* **59**, 329–355
 54. Koch, H.-G., Moser, M. & Müller, M. (2003). Signal recognition particle-dependent protein targeting, universal to all kingdoms of life. *Reviews of physiology, biochemistry and pharmacology* 55–94
 55. Steiner, D., Forrer, P., Stumpp, M.T. & Plückthun, A. (2006). Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat. Biotechnol.* **24**, 823–31
 56. Paschke, M. & Höhne, W. (2005). A twin-arginine translocation (Tat)-mediated phage display system. *Gene* **350**, 79–88
 57. Liu, Y., Regula, L.K., Stewart, A. & Lai, J.R. (2011). Synthetic Fab fragments that bind the HIV-1 gp41 heptad repeat regions. *Biochem. Biophys. Res. Commun.* **413**, 611–5
 58. Makowski, L. & Soares, A. (2003). Estimating the diversity of peptide populations from limited sequence data. *Bioinformatics* **19**, 483–9
 59. Rodi, D.J., Soares, A.S. & Makowski, L. (2002). Quantitative assessment of peptide sequence diversity in M13 combinatorial peptide phage display libraries. *J. Mol. Biol.* **322**, 1039–52
 60. Voigt, C.A., Kauffman, S. & Wang, Z.-G. (2000). Rational evolutionary design: the theory of in vitro protein evolution. *Adv. Protein Chem* **55**, 79–160
 61. Jonsson, A., Dogan, J., Herne, N., Abrahmsén, L. & Nygren, P.-Å. (2008). Engineering of a femtomolar affinity binding protein to human serum albumin. *Protein Engineering Design and Selection* **21**, 515–527
 62. Hiller, Y., Gershoni, J.M., Bayer, E.A. & Wilchek, M. (1987). Biotin binding to avidin. Oligosaccharide side chain not required for ligand association. *Biochemical journal* **248**, 167–171
 63. Van Roy, N., Mangelschots, K. & Speleman, F. (1993). Improved immunocytochemical detection of biotinylated probes with neutralite avidin. *Trends in Genetics* **9**, 71–72

64. Catanzano, F., Graziano, G., Fusi, P., Tortora, P. & Barone, G. (1998). Differential scanning calorimetry study of the thermodynamic stability of some mutants of Sso7d from *Sulfolobus solfataricus*. *Biochemistry* **37**, 10493–10498
65. Knapp, S., Karshikoff, A., Berndt, K.D., Christova, P., Atanasov, B. & Ladenstein, R. (1996). Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus*. *J. Mol. Biol.* **264**, 1132–44
66. Bradbury, A.R. & Marks, J.D. (2004). Phage antibody libraries. *Phage Display a Practical Approach* 253–259
67. Mourez, M. & Collier, R.J. (2004). Use of phage display and polyvalency to design inhibitors of protein-protein interactions. *Protein-Protein Interactions: Methods and Applications* 213–227
68. Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488–92
69. Lee, T.S., Krupa, R.A., Zhang, F., Hajimorad, M., Holtz, W.J., Prasad, N., Lee, S.K. & Keasling, J.D. (2011). BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J Biol Eng* **5**, 12

Chapter 3

Selection of nanomolar affinity binding proteins for active Tuberculosis diagnosis from an Sso7d alternative scaffold phage display library*

Tuberculosis (TB) surpassed HIV to be the leading cause of death from infectious disease worldwide for the first time in 2014 and caused millions of deaths. The low sensitivity, extended processing time and high expense of diagnostics are great challenges to eliminating TB. *Mycobacterium tuberculosis* ornithine transcarbamylase (*Mtb* OTC, Rv1656) has been identified in the urine of patients with active TB infection, making *Mtb* OTC a promising target for point-of-care diagnostics in resource-limited settings. We are motivated to engineer phage-based diagnostic systems that feature improved physical stability, cost of production and sensitivity relative to traditional antibody based reagents. A small, hyperthermophilic and cysteine free DNA-binding protein Sso7d was identified as a phage display scaffold for the evolution of tight, specific binding molecules that are stable up to 70°C. Specific binding proteins with low nanomolar affinities for *Mtb* OTC were selected from the naïve Sso7d phage library. Phage particles displaying Sso7d variants along with a purified monoclonal antibody were utilized to generate a capture ELISA-based assay for *Mtb* OTC. The ELISA assay signal is linear over the target concentration range of 2.0-125.0 ng/mL with limit of detection 0.4 ng/mL (12pM), which is comparable to commercial available antibody-based assays (pM~nM sensitivities). Importantly, the assay maintains functionality at both neutral and basic pH in presence of salt and urea over the range of concentrations typical for human urine.

**This chapter will be published as a paper and the contents except abstract are not included in this PhD dissertation.*

Chapter 4

A double-phage based diagnostic test established from an Sso7d alternative scaffold phage display library for active Tuberculosis*

Tuberculosis (TB) has become the leading cause of death from infectious disease. A challenge to eliminating this global health issue is lack of effective, efficient and affordable diagnosis at point-of-care which is mainly located in resource-poor areas. A fully phage-based diagnostic test has the potential to address this global health challenge by detecting a promising protein target *Mycobacterium tuberculosis* ornithine transcarbamylase (*Mtb* OTC) in urine of TB patients. Compared with traditional antibody reagents for protein targets detection, phage has improved physical properties especially in thermal stability and therefore phage-based diagnostics have the advantage to deliver and store at point-of care without reliable power source. Additionally cost of phage production can be reduced to cents per assay, which is critical for undeveloped area diagnosis application. A hyperthermophilic protein Sso7d is demonstrated to be evolved into sensitive and specific binding molecules against *Mtb* OTC in chapter 3. In this chapter, a high throughput screening methods were developed to select the best pair of Sso7d mutants displaying phage particles. As a proof-of-concept, a double-phage capture ELISA utilizing the selected pair of phage particles was characterized. This ELISA assay provided a linear response to recombinant *Mtb* OTC between 4.9-625.0 ng/mL ($R^2=0.9984$), with a limit of detection of 4.5 ng/mL (130pM) which is comparable with commercial available antibody-based assays. Additionally, this double-phage capture ELISA assay is quite tolerate with urea, ionic strength and pH in ranges typical to

**This chapter will be published as a paper and the contents except abstract are not included in this PhD dissertation.*

that of human urine, which means this pair of Sso7d binding variants on phage are promising candidates for the generation of a diagnostic test for *Mtb* OTC in a urine matrix. Moreover, the tolerance of this double phage-based sensor for pH is significantly improved compared to the diagnostic test in chapter 3, which probably due to the replacement of antibody with phage particles.

Chapter 5

Single chain variable fragment (scFv) based on a monoclonal antibody and displayed on phage as a promising diagnostic test for active tuberculosis*

Antibodies have been served as the quintessential biological recognition molecules for decades. Monoclonal hybridoma technology significantly improved applications of antibodies in therapeutics and diagnostics. Compared with traditional monoclonal antibody-based diagnostic tests, phage-based tests have great advantages in cost, production and stability at harsh conditions. For diagnostic applications in resource-limited settings, phage-based diagnostic tests may be especially affordable, effective and suitable for point-of-care deployment for the diagnosis of infectious diseases, such as tuberculosis (TB). TB remains as a global health issue and in 2014, TB has surpassed HIV as the leading cause of death from infectious disease worldwide. The high mortality of TB is due primarily to the difficulty of diagnosis in resource-limited areas rather than a difficulty in treatment. In order to eliminate the global TB epidemic, fast, sensitive and affordable diagnostics of active TB infection are in high demand. A promising diagnostic protein target, *Mycobacterium tuberculosis* ornithine transcarbamylase (*Mtb* OTC), was identified in urine of TB patients, but not in urine of healthy persons. In order to assemble a phage-based assay for detecting *Mtb* OTC, a short version of antibodies, single chain variable fragment (scFv), was converted from monoclonal hybridomas prepared against *Mtb* OTC and displayed on phage surface as binding molecules. Binding of the scFv displaying phage particles against the *Mtb* OTC was evaluated by phage ELISA.

**This chapter will be published as a paper and the contents except abstract are not included in this PhD dissertation.*

Appendix 1

Table S2-1. List of oligonucleotide primer sequences in chapter 2

Primers	Nucleotide sequence
Oligo 1	5'-CAA CAG TAG CGG CCG CGG CAT GGT GAT GGT GAT G-3'
Oligo 2	5'-CTC GAG GCG ACC GTT AAA TTC AAA TAC AAA GGT GAA GAG AAA CAG GT-3'
Oligo 3	5'-ACA AAG GTG AAG AGA AAC AGG TTG ACA TCT CTA AAA TCT GGA TCG TTG CGC-3'
Oligo 4	5'-CTA AAA TCT GGA TCG TTG CGC GTG ACG GTA AAT GGA TCG ACT TCT CTT ACG-3'
Oligo 5	5'-CGA TAC CGG ATT TAC CAC CAC CCA GGT CGT AAG AGA AGT CGA TCC ATT TAC-3'
Oligo 6	5'-CTT TCG GCG CGT CTT TTT CAG AAA CGT AAC CGA TAC CGG ATT TAC CAC CAC-3'
Oligo 7	5'-GCT AGC TTT TTT CTG TTT TTC CAG CAT CTG CAG CAG TTC TTT CGG CGC GTC TTT TTC-3'
Oligo 8	5'-AGC TAG GTC ACT CGA GGC GAC CGT TAA ATT CAA ATA CAA AGG T-3'
Oligo 9	5'-ACT GAC TTG TGG CTA GCT TTT TTC TGT TTT TCC AGC ATC TGC AG-3'
Oligo 10	5'-ATT ATT CTC GAG GCG ACC GTT AAA TTC AAA TAC AAA GGT G-3'
Oligo 11	5'-CGA AAG GGG GAT GTG CTG-3'
Oligo 12	5'-ATT ATT GCG CGC ATG CCG CAC TC-3'
Oligo 13	5'-CCA CCA CCC AGG TCG TAC CTG AAT CTC CTT CTT TTA CCC CTA CGC CTT CTA GAC CAG ATT TTA GAG ATG TCA ACC TGT-3'
Oligo 14	5'- GGC GCG TCT TTT TCA GAA ACG NNA CCG NNA CCG NNT TTA CCA CCA CCC AGG TCG TAG NNG AAG NNG ATG NNT TTA CCG NNA CGG NNA ACG NNC CAG ATT TTA GAG ATG TCA ACC TGT TT -3'
Oligo 15	5'-CACA TTA GGC ACC CCA G-3'
Oligo 16	5'-CAC CAG TAC AAA CTA CAA CGC CTG-3'
Oligo 17	5'-ATT ATT CAT ATG GCG ACC GTT AAA TTC AAA TAC AAA GG- 3'
Oligo 18	5'-ATT ATT AAG CTT TTT TTT CTG TTT TTC CAG CAT CTG C- 3'

Appendix 2

Table S2-2. DNA sequences of Sso7d display features in chapter 2

Feature	Nucleotide sequence	Protein sequence
TorA leader peptide (TAT)	5'-ATG AAC AAT AAC GAT CTC TTT CAG GCA TCA CGT CGG CGT TTT CTG GCA CAA CTC GGC GGC TTA ACC GTC GCC GGG ATG CTG GGG CCC TCA TTG TTA ACG CCG CGA CGT GCG ACT GCG GCG CAA GCG GCG-3'	MNNNDLFQASRR RFLAQLGGLTVA GMLGPSLLTPRR ATAAQAA
pelB leader peptide (Sec)	5'- ATG AAA TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC GCA GCA AGC GGC GCG CAT GCC-3'	MKYLLPTAAAGL LLLAASGAHA
SRP leader peptide	5'-ATG AAA AAA ATC TGG CTG GCG CTG GCA GGC CTG GTG CTG GCG TTT AGC GCG-3'	MKKIWLALAGLV LAFSA
Linker between TorA leader and Sso7d variant	5'-GGC GCG CAT GCC GCA CTC GAG-3'	GAHAALE
Linker between pelB leader and Sso7d variant	5'-CTC GAG-3'	LE
Linker between SRP leader and Sso7d variant	5'-CAT GCC GCA CTC GAG-3'	HAALE
Linker between Sso7d variant and p3	5'-GCT AGC GGC AAA CCA ATC CCA AAC CCA CTG CTG GGC CTG GAT AGT ACT CAC CAT CAC CAT CAC CAT GCC GCG GCC GCT-3'	ASGKPIPNNPLLG LDSTHHHHHHAA AA
Sso7d-strep variant	5'-GCG ACC GTT AAA TTC AAA TAC AAA GGT GAA GAG AAA CAG GTT GAC ATC TCT AAA ATC TGG ATC GTT GCG CGT GAC GGT AAA TGG ATC GAC TTC TCT TAC GAC CTG GGT GGT AAA TCC GGT ATC GGT TAC GTT TCT GAA AAA GAC GCG CCG AAA GAA CTG CTG CAG ATG CTG GAA AAA CAG AAA AAA-3'	ATVKFKYKGEEK QVDISKIWIVAR DGKWIDFSYDLG GGKSGIGYVSEK DAPKELLQMLEK QKK

Appendix 3

4-parameter Logistic equation

$$Y = A_2 + \frac{A_1 - A_2}{1 + \left(\frac{X}{X_0}\right)^P}$$

Y: Fraction bound to each well;

X= [Competitor];

A₂=Final value. The plateau value when saturated by the competitor and should be close to 0;

A₁=Initial value. The plateau value when very diluted competitor present and should be close to 1;

P=binding stoichiometry

X₀=EC50=K_D

Appendix 4

Copyright agreement from FEBS J. for reprinting and reusing a publication as Ning Zhao's PhD dissertation

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Aug 31, 2016

This Agreement between Ning Zhao ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	3939570664616
License date	Aug 31, 2016
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	FEBS Journal
Licensed Content Title	Phage display selection of tight specific binding variants from a hyperthermostable Sso7d scaffold protein library
Licensed Content Author	Ning Zhao, Margaret A. Schmitt, John D. Fisk
Licensed Content Date	Mar 6, 2016
Licensed Content Pages	17
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	250
Title of your thesis / dissertation	DEVELOPMENT OF A PHAGE-BASED DIAGNOSTIC SENSOR FOR ACTIVE TUBERCULOSIS
Expected completion date	Nov 2016
Expected size (number of pages)	200
Requestor Location	Ning Zhao 1370 Campus delivery Colorado State University FORT COLLINS, CO 80523 United States Attn: Ning Zhao
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Ning Zhao 1370 Campus delivery Colorado State University FORT COLLINS, CO 80523 United States Attn: Ning Zhao
Total	0.00 USD

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at <http://mvaccount.copyright.com>).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, and any **CONTENT (PDF or image file)** purchased as part of your order, is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. **For STM Signatory Publishers clearing permission under the terms of the [STM Permissions Guidelines](#) only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts.** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto
- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.
- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.
- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.
- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement

shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.
- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.
- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.
- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.
- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The [Creative Commons Attribution License \(CC-BY\)](#) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The [Creative Commons Attribution Non-Commercial \(CC-BY-NC\) License](#) permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The [Creative Commons Attribution Non-Commercial-NoDerivs License](#) (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library

<http://olabout.wiley.com/WileyCDA/Section/id-410895.html>

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customer@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.
