DEVELOPMENT OF HIGH-THROUGHPUT APPROACHES FOR THE PHAGE DISPLAY GENERATION AND PRODUCTION OF RECOMBINANT ANTIBODY FRAGMENTS FOR MACROMOLECULAR CO-CRYSTALLIZATION

by

ASHWINI NADKARNI

(Under the Direction of Cory Momany)

ABSTRACT

The empirical process of macromolecular crystallization has always been a bottleneck in the production of protein structures. Co-crystallization of intransigent proteins with antibody fragments has recently emerged as a powerful tool to facilitate crystal formation and improve crystal quality. However, this approach is not readily applied because of the lack of adequate quantities of antibody fragments that recognize target proteins and also due to the complexity of the selection system. The objective of this dissertation was to develop a recombinant antibody fragment (Fab)-based system for the high-throughput generation of co-crystallization proteins (CCPs). Phage display technology lies at the heart of the selection process, but the method has been notoriously slow and tedious. A rapid approach was developed that was faster than standard methods and because of the simplified manipulations, less tedious and easily applicable to robotic automation. The other focus of this dissertation was on addressing the underlying issues of producing antibody fragments in a bacterial system. Automated methods of purifying the target-specific Fabs were developed that are general in nature and should work with all of the Fabs produced in this system. By optimizing the bacterial strain, culture conditions, and

purification strategy, as much as 12 mg of highly purified Fab was produced per liter of bacterial cell culture in less than two days using bench-top equipment.

INDEX WORDS: Antibody fragment, Co-crystallization, Crystallization, Fab, High-throughput, Phage Display, Secretion, Cytoplasmic, Structural Genomics

DEVELOPMENT OF HIGH-THROUGHPUT APPROACHES FOR THE PHAGE DISPLAY GENERATION AND PRODUCTION OF RECOMBINANT ANTIBODY FRAGMENTS FOR MACROMOLECULAR CO-CRYSTALLIZATION

by

ASHWINI NADKARNI

B.Pharm., Mumbai University, India, 2001

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2006

DEVELOPMENT OF HIGH-THROUGHPUT APPROACHES FOR THE PHAGE DISPLAY GENERATION AND PRODUCTION OF RECOMBINANT ANTIBODY FRAGMENTS FOR MACROMOLECULAR CO-CRYSTALLIZATION

by

ASHWINI NADKARNI

Major Professor:	Cory Momany
Committee:	Shelley Hooks
	J. Warren Beach
	Arthur Grider
	Roger G. Dean

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2006

DEDICATION

To my relentless cheerleader,

My loving husband Yash

and

To my ardent fans,

Aai, Baba and Priya

Thank you for believing…

ACKNOWLEDGEMENTS

First and foremost I wish to thank my major advisor, Dr. Cory Momany, for his excellent mentoring and support, in everything from shooting crystals to buying my first home in this country. Dr. Will Taylor deserves special mention here. It is due to his perseverance and faith in my abilities that I stand here today, to receive my degree. I thankfully acknowledge the members of my committee: Dr. Shelley Hooks, Dr. J. Warren Beach, Dr. Arthur Grider and Dr. Roger Dean for their comments and input in every committee meeting. I would like to thank Dr. Michelle Momany for the timely help in providing supplies and equipment. Of course, this would not have been possible without the support of family and friends. The constant encouragement and support of my parents-in-law has been a great source of inspiration. I couldn't have done it without you.

Thank you all for being there …

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

2 ADDRESSING THE CRYSTALLIZATION BOTTLENECK BY CO-CRYSTALLIZATION 36

3 OPTIMIZATION OF A MOUSE RECOMBINANT ANTIBODY FRAGMENT FOR EFFICIENT PRODUCTION FROM *ESCHERICHIA COLI* 69

4 COMPARISON OF CYTOPLASMIC AND SECRETED PRODUCTION OF A MOUSE RECOMBINANT ANTIBODY FRAGMENT FROM *ESCHERICHIA COLI* 111

5 A RAPID PHAGE DISPLAY PROTOCOL TO SELECT ANTIBODY FRAGMENTS FOR CO-CRYSTALLIZATION OF MACROMOLECULES 131

6 CONCLUSIONS AND FUTURE DIRECTIONS..161

LIST OF TABLES

Page

Table 1.1: Crystallographically studied antibody-protein antigen complexes.....8

Table 2.1: Proteins crystallized with the help of a CCP (Co-crystallization Protein) 52

Table 3.1: Oligonucleotides used for PCR amplification and mutagenesis.....90

LIST OF FIGURES

Page

Figure 1.1: The Bottleneck of Macromolecular X-ray Crystallography.....22

Figure 1.2: A Schematic Solubility Diagram.....23

Figure 1.3: Schematic diagram of an Ig molecule 24

Figure 1.4: Somatic Recombination and Expression of Ig germline DNA 25

Figure 1.5: Schematic sketch of a Filamentous Phage 26

Figure 1.6: A "Panning" Cycle using Phage Display 27

Figure 2.1: Molecules used for Co-crystallizations and their Binding Surfaces 55

Figure 2.2: A Schematic of Commonly Used Selection Systems.....57

Figure 3.1: Plasmid Design and Sequence changes used to produce various recombinant Fabs ..92

Figure 3.2: Bacterial growth kinetics and production of rFab2 through rFab4 in defined medium using two different bacterial cell lines 94

Figure 3.3: Bacterial growth kinetics and production of Fab4 using two different protein induction methods 96

Figure 3.4: Bacterial growth kinetics and Fab4 production with auto-induction media 98

Figure 3.5: Metal-chelate chromatography of rFab4 prepared in auto-induction and defined glucose media after thiosepharose chromatography 100

Figure 3.6: "One-step" purification of antibody fragments using an ÄKTA purifier 102

Figure 3.7: SDS-PAGE analysis of Fab4 purified using three different purification protocols ..104

Figure 3.8: Native PAGE analysis of rFab4 complexed with its antigen HIV p24.....106

Figure 4.1: Ni^{+} Affinity Chromatography Profiles for Secreted and Cytoplasmic rFab4 from Different *E.coli* Strains 123

Figure 4.2: Anion Exchange Chromatography (Q-FF) Profiles for Secreted and Cytoplasmic rFab4 from Different *E.coli* Strains 124

Figure 4.3: SDS-PAGE analysis of cytoplasmic and secreted Fab4 from two *E.coli* strains 125

Figure 4.4: ELISA analysis of antigen binding activity of the cytoplasmic and secreted rFab4 126

Figure 5.1: SDS-PAGE analysis of BenM FL bound and eluted from Dynabeads®TALON™ using different buffer conditions 152

Figure 5.2: Output titers of a single selection round performed with BenM, CatM and no antigen bound to the beads and washed with different levels of imidazole 153

Figure 5.3: Phage ELISA signal of Fab-phage extrusion by E.coli grown in superbroth or glucose supplemented defined medium after helper phage addition 154

Figure 5.4: Phage ELISA signal of Fab-phage extrusion by *E.coli* grown in superbroth or a modified defined media after helper phage addition 155

Figure 5.5: Percent Recovery over increasing Rounds of Panning for Standard versus Rapid Panning 156

Figure 5.6: PCR screening for rFab bearing phagemid from different sources after three selection rounds using standard and rapid phage display panning protocols 157

Figure 6.1: High-throughput Pipeline for the Generation of Antibody Fragments for Co-crystallization 165

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Why crystallize macromolecules?

Furnished with the fruits of several genome projects, scientists and researchers now have access to whole genome sequences of a number of organisms including humans. This sequence information is however, merely the software code that instructs the cell how to synthesize the real workhorses or building blocks of the cell which include proteins, carbohydrates, lipids and nucleic acids. "To really understand biological processes, we need to understand how proteins function in and around cells since they are the functioning units," says Hanno Steen, director of the Proteomics Center at Children's Hospital, Boston. The three-dimensional macromolecular structural information provides a clear picture of the way these molecules behave in the cellular milieu, their responses to exterior influences and also how these responses or signals are transmitted through the cellular hierarchy. The knowledge of structural information thus seems to have a massive impact on several areas of science including but not limited to rational drug design, genetic engineering and post-genomic studies.

There are several approaches that provide information regarding macromolecular structure. However methods such as NMR and molecular dynamics, in spite of providing detailed protein structural information, have a size cut-off of about 30kDa. Only the X-ray diffraction analysis of single crystals of macromolecules can yield a precise description of a macromolecule’s structure, a description that can serve as a basis for drug design, and an intelligent guide for protein engineering [4]. X-ray crystallography has played a fundamental role

in connecting the dots between genomic data and biological function by providing atomic resolution information for 31,294 out of the 36,932 structures in the Protein Data Bank (as of June 06, 2006). In the preface to his book *Crystallization of Biological Macromolecules*, Alexander McPherson, professor of Biochemistry and Molecular Biology at University of California said, "From sequences of nucleotides on strands of DNA, we are, because of this method (*X-ray crystallography*), now poised to visualize their ultimate products in atomic detail."

The Bottleneck of Macromolecular X-ray Crystallography

In the early years, after Bernal and Crowfoot [5] recorded the first X-ray diffraction from a protein crystal, the availability of crystals was not an issue. Several early targets such as hemoglobin, insulin, pepsin and others had all been well crystallized by the pioneers of protein biochemistry such as Sumner and Northrop [6] [7] [8].

The mid and late decades of the 20th century saw tremendous advances being made in the X-ray diffraction machinery. Strong and focused X-ray sources (synchrotron beam lines), sensitive X-ray detectors, new automation systems, improved cryogenic and mounting procedures for proteins and sophisticated and user-friendly crystallographic software packages are some of the developments that have made it possible to determine an initial protein structure within a few hours instead of days and months as needed a few years ago, provided there are adequate useful crystals available [9]. Thus, protein crystallization no longer remained a means of purifying proteins from impure mixtures. The principal goal of macromolecule crystallization, thus, became their structure determination by X-ray crystallography.

As the number of crystal structures solved by crystallographers started rising, two show-stopping limitations came to the forefront. They were 1) availability of target proteins, and 2) availability of good quality diffracting crystals. Due to these limitations, crystallographers were able to tackle only stable medium-size proteins abundantly found in body fluids. The 1980s brought the advent of recombinant DNA technology, which among other benefits permitted researchers to prepare ample amounts of otherwise rare and elusive proteins. This technology not only resolved the limitation of target availability but also provided opportunities to modify the protein to enhance its crystallizability [10]. More recently, the rising number of genome projects has revealed numerous new targets for therapy of human disease. However this escalation of the number of targets available in a pure, homogenous and high yield form has heightened the need for obtaining diffraction quality crystals. Statistics from pilot structural genomics projects place the rate of getting from cloned protein to structure determination at only ~10%. Figure 1.1 depicts figures taken from a Human Proteome Structural Genomics pilot project (Brookhaven National Laboratory, The Rockefeller University and Albert Einstein College of Medicine, New York, USA: http://proteome.bnl.gov/progress.html), which show that out of 124 proteins cloned, 63 were purified and only 19 yielded crystals suitable for structure determination. These numbers lead to the conclusion that even when proteins can be cloned, expressed, solubilized and purified, there is no guarantee of acquiring diffraction quality protein crystals. Macromolecular crystallization thus seems to be the biggest bottleneck in the rate at which new structures are obtained by X-ray crystallography [11].

The Science of Macromolecular Crystallization

In order to understand why protein crystallization forms the rate limiting step of structural genomics programs and what can be done to address this situation, it is necessary to first review the science of crystallization.

Crystallization of a molecule is a two-step, phase transition process consisting of nucleation followed by growth. Both these steps are inherently dependent on the creation of a supersaturated solution state [12]. The concept of supersaturation and its influence on the two steps of crystallization can be best illustrated by means of a phase diagram. Phase diagrams are a good quantitative indicator of the stability of a state (liquid, crystalline or amorphous solid) under several crystallization parameters such as the concentration of protein, precipitant, additives and others [13, 14]. Figure 1.2 illustrates a complete phase diagram of a solubility curve, depicting the variations in protein solubility (vertical axis) with the concentration of a precipitant (horizontal axis) or parameters such as temperature or pH [1]. At precipitant concentrations below the solubility limit (depicted by the solid line), the system is said to be undersaturated and the solid phase dissolves completely. Above this limit, the system is said to be supersaturated. At extreme supersaturation, stable nuclei form spontaneously and grow due to the increased energy state of the system. This zone is called the labile zone. At the low end of supersaturation, the spontaneous formation of stable nuclei is rare. However stable nuclei, if somehow formed, may grow in this metastable zone. Ultimately, at the far reaches of the labile region is the zone of precipitation, where not only do crystal nuclei appear but also amorphous precipitate. Thus achieving a high level of supersaturation might lead to formation of either a precipitate or a large number of nuclei that give rise to a large number of small, useless microcrystals. The key to getting a few, large, diffraction quality crystals, thus, is achieving that

elusive level of supersaturation, where only a few nuclei form spontaneously and grow to form large crystals.

The Current State of Art in Protein Crystallization

Current methods are largely empirical and have evolved through trial and error. Currently, the two most popular methods include vapor diffusion and microbatch. The vapor diffusion method involves mixing of a protein and a precipitant drop onto a surface such that it is spatially separated from a reservoir containing a higher volume of the undiluted precipitant. The entire system is sealed from air so that no solvent is lost to the environment. The water in the crystallization drop is slowly removed from the protein drop to the reservoir as a vapor so that effective concentrations of all components in the drop increase. This vapor equilibration gradually drives the system toward supersaturation and produces crystals [1]. Batch crystallization and the easily automated microbatch under oil technique involve the mixing of protein and precipitant under a layer of oil with lower density than the protein/precipitant mixture. The precipitant alters the protein solubility or the electrolyte properties of the solution to yield a mother liquor, immediately supersaturated with respect to protein. The protein and precipitant remain isolated from air, and the solid state eventually forms from such a solution, provided the physical and chemical parameters have been appropriately chose [15].

Protein crystallization attempts begin once highly purified and soluble protein samples are obtained. Crystallization attempts thus entail the exploration of the phase space of the protein to slowly move the system towards a metastable supersaturated state. However there are countless parameters such as pH, ionic strength, temperature and concentrations of various components, which influence this phase space and affect nucleation. The sparse matrix screening

method introduced by Jancarik and Kim [16] introduced the idea of reduced sampling approaches. There are now several screening methods based upon this approach and they have continued to work well based on a survey of recent crystallization reports. This has resulted in a surge of prefabricated crystallization cocktail kits, including ones that are optimized for different groups of macromolecules. The next logical step is the development of data mining and machine learning algorithms to deliver predictive models for protein crystallization [17, 18].

Nevertheless, the use of cocktail kits and statistical databases in screening crystallization conditions are limited by protein availability. They also are not thorough in covering the multi-dimensional parameter space. Different macromolecules have different crystallization potentials, which make it important to have tools that cast a wide net of trials and conditions. High-throughput methods via automation and microfluidics have played an important role in resolving this. The past few years have seen some of the greatest innovations in the field of protein crystallization by way of automating and miniaturizing crystallization trials [19]. Institutes such as the Hauptman-Woodward Medical Research Institute, the Protein Structure Factory in Germany, the collaborative group including the Lawrence Berkeley National Laboratory (LBNL) Bioinstrumention Group, the Genomics Institute of the Novartis Research Foundation (GNF), and Syrrx, Inc. (a company devoted to high throughput structure biology) have succeeded in building new generation robots for the automated production of high-quality protein crystals. Some of these integrated systems are capable of performing ~40,000 to >100,000 trials a day, thus conducting a full exploration of this multi-dimensional condition space. Further, these systems also have the ability to dispense nanoliter droplets – some as small as 1 nanoliter, thus minimizing protein volume requirements. Additionally, both success and failure data can accurately be accumulated for data mining and analysis for future improvements [20].

Despite these advances in technology, whether it is the robotic systems of today or the technology based on the microcomputer chip industry, the crystallization problem still stands unsolved (as shown by statistics in Figure 1.1). The bottleneck of protein structure determination cannot be reduced to the issue of which screening method, how many screening trials or crystallization set up is to be used. Rather, approaches based on understanding and controlling the physicochemical properties of macromolecules; approaches that drive a macromolecule to form diffraction quality crystals seem to be the need of the times.

The Co-crystallization Approach to solve the Crystallization Problem

The very nature of macromolecular crystals makes them more vulnerable than conventional small molecule crystals. Macromolecule crystals are small, extremely fragile, and easy to crush; are stable over a very narrow range of temperature, ionic strength or pH; generally exhibit weak optical properties and diffract X-rays to resolutions far short of the theoretical limit. Due to their ability to incorporate large quantities of solvent in their lattices, they contain disordered water molecules, precipitant, ions, and a range of impurities and defects. It is a collection of all these properties that determine the crystallizability of macromolecules [21]. Consequently, crystallization is inhibited by several factors such as heterogeneity, molecular flexibility or a polydisperse character in solution [22]. Additionally, particular groups of macromolecules such as membrane proteins and viral capsid subunits have posed special problems due to the "inside out" nature of their structures. These are amphipathic molecules that possess hydrophobic surfaces when in contact with the lipid alkyl chains and polar surfaces when in contact with the polar heads of the lipids in the membrane. Thus solubility and accordingly crystallization becomes an issue for these proteins. Membrane proteins can be solubilized by the

addition of excess amounts of detergents – amphiphilic molecules that form micelles above their critical micellar concentration. Any crystallization strategy thus has to take in to account this amphipathic nature of the surface of membrane proteins [23].

One useful approach to addressing the crystallization problem is to co-crystallize a target protein with another protein that specifically recognizes the target so that a stable complex is formed. The central hypothesis is that many proteins not previously crystallized can be crystallized as complexes with cognate proteins. The formation of a complex confers changes in its physical properties to make it more favorable toward crystallization than the protein alone. Chapter 2 is a complete review of co-crystallization, the different protein formats used for co-crystallization, methods of generation and issues in the process.

Protein	**Antibody Fragment**	**References**
Lysozyme*	HyHEL-5	[24]
Lysozyme*	HyHEL-10	[25]
Neuraminidase*	NC41	[26]
Cytochrome *c* oxidase#	F_v7E2	[27]
HIV-1 capsid protein#	Fab25.3	[28]
Cytochrome bc_1 complex#	$_{18E11}F_v$	[29]
KcsA K^+ channel#	Fab	[30]

*In these complexes, the antibody fragment was used to study antigen-antibody interactions. #In these complexes, the antibody fragment was used as a crystallization tool to aid in the structure determination of the protein antigen.

Table 1.1 Crystallographically studied antibody-protein antigen complexes

Co-crystallization of proteins to drive crystal formation has been around for some time, but is not readily applied because of the lack of adequate quantities of cognate proteins that recognize target proteins. The most common examples of protein co-crystallization are the crystallization of antibody-antigen complexes. Various antibody fragments such as Fabs and scFvs (which will be described in detail later in this chapter) have been used to crystallize several difficult proteins. Table 1 is based on a similar table in [22] and provides a few instances of antibody-antigen complexes that were studied crystallographically, along with their references.

Fab and Fv fragments are reasonably soluble and also very specifically bind their antigens with binding constants from 10^5 to 10^8 M^{-1}. Thus, they can control non-specific aggregation, transform aggregated material into a soluble, mono-disperse samples. Thus, they can modify the surfaces of molecules in ways favorable for crystallization. Moreover, a complex crystal with an antibody fragment can also aid in the actual structure determination by providing phasing information. Crystallization and structure determination of the antibody fragment by itself or a previously solved antibody fragment structure can be used as a molecular replacement model to determine the phases. Antibody fragments can also act as recipients of heavy atom labels and thus be useful in structure determination of the target protein [31]. In addition to rescuing the structural studies of several interesting non-crystallizers, the co-crystallization approach also helps in improving the crystal quality of proteins that do crystallize. Laver, in his

work on crystallization of antibody-protein complexes, provided several examples of proteins where the uncomplexed molecules yield non-diffracting or low quality crystals than they do when complexed [32].

Co-crystallization of target proteins with antibody fragments is a valuable addition to the growing repertoire of tools established to resolve the macromolecular crystallographic bottleneck. Provided there is an available antibody fragment, this technology would be a single-shot approach to crystallize intransigent proteins and also to optimize crystal quality for those that produce some crystals.

Antibodies and Antibody Fragments – the Co-crystallization Agents

Antibodies are a family of structurally related glycoproteins and mediators of humoral immunity. Vertebrate immune systems rely on them for coupling foreign body recognition with elimination processes. Further, antibody structures and their interactions with antigens have been well studied as reviewed by [33], [34], [35], [36].

Structurally, all intact immunoglobulin molecules have a symmetric core composed of two identical light (L, each about 24 kD) chains and two identical heavy (H, each about 55 or 70 kD) chains (Figure 1.3). The two heavy chains are covalently attached to each other by disulfide bonds and one light chain is also connected to one heavy chain by a disulfide bond. Further, the heavy chains and the light chains consist of amino terminal (N-terminal) variable (V) regions and carboxy terminal constant (C) regions. Both the light and heavy chains contain a series of repeating, homologous units, each about 110 amino acid residues in length, which fold independently in a globular motif or β-barrel that is called immunoglobulin domain. These basic structural features fold into 12 such immunoglobulin domains that associate in pairs (V_L: V_H, C_L:

C_H^1, C_H^2: C_H^2, and C_H^3: C_H^3) producing the basic monomeric Y-structure with the two Fab segments and one Fc segment as seen in Figure 1.3.

The Fc portion includes only the constant regions of both heavy chains (except C_H^1) and is responsible for the effector functions of the antibody. It plays no role in antigen binding. Antibody molecules can be divided into distinct classes and subclasses based on differences in the structure of their heavy chain C regions. These classes, also called as isotypes, are named as IgA (IgA1 and IgA2), IgD, IgE, IgG (IgG1, IgG2, IgG3, IgG4), and IgM. Based on their carboxy terminal constant regions (in the Fab portion), there are two isotypes of light chains, called κ and λ. Any one antibody molecule always contains unique heavy and light chain isotypes [33].

A potential of six loops, three from the heavy and three from the light chain variable region, interact with antigens and confer the majority of the ligand binding potential. These three highly divergent stretches are about 10 amino acids long and are termed as the "complementarity determining regions" or CDR loops. Substantial diversity of sequences occurs in the CDR loops as a result of antibody rearrangements and somatic mutations as seen in Figure 1.4. Crystallographic analyses reveal that these CDRs form extended loops that are exposed on the surface of the antibody and are thus available to interact with antigen [24]. The CDR loops are joined by well conserved residue regions called as framework segments. Despite this high diversity, the conformation of the loops tend to be restrained to a set of "canonical" structures [37], and in many cases, residues from CDR loops are critical for stability of the β sheet barrel.

On the other hand, immunoglobulins are inherently flexible, yielding a spectrum of molecular conformations that permits them to bind to a variety of multivalent antigens. Flexibility in the structure comes about in two ways. There is an "elbow" joint between each constant (C_H^1, C_L) and variable (V_H, V_L) domains that gives it several more degrees of freedom

in interacting with the target molecule and may be required for high-affinity binding with the target [38] [39]. Additionally, there is a "hinge" region located between C_H^1 and C_H^2 domains on the heavy chain, which provides a high degree of molecular motion between those domains [33]. This high level of flexibility makes crystallization of complete intact immunoglobulin molecules problematic. Consequently, native antibodies and their bivalent binding modes are also unsuitable for co-crystallization. The covalent and non-covalent associations that hold the light and heavy chains together into an intact antibody are susceptible to proteolysis. Porter *et al* showed that the hinge region was most prone to proteolytic cleavage [40]. Thus, cleavage of an intact antibody such as IgG with the protease papain results in an Fc fragment and two identical Fab molecules, each of which contains a full light chain and a portion of the heavy chain (C_H^1), termed the Fd fragment. Instead of papain, if pepsin is used to cleave the IgG, proteolysis is restricted to the carboxy terminus of the hinge region, generating an antigen-binding fragment of IgG with the hinge and the interchain disulfide bonds intact. Such Fab fragments are bivalent and are denoted as $F(ab`)_2$.

Generation of Antibodies/Fragments to Target Proteins

Most of the protein structures solved as co-crystallization complexes to date have used antibody fragments generated proteolytically from monoclonal antibodies (Mab). Monoclonal antibodies are produced by hybridoma or monoclonal technology, which is very effective in producing high affinity antibodies. This technique was first described in 1975 by Kohler and Milstein who developed the method of producing from mice continuous cell lines expressing homogenous monoclonal antibodies (Mab) [41]. To produce a monoclonal antibody specific for a defined antigen, the first step is to isolate murine splenic B cells from a mouse that has been

immunized with that antigen. Immortalization of this Mab producing B cell is achieved by cell fusion or somatic hybridization between the normal B cell and a myeloma cell. Such fusion-derived immortalized antibody-producing cell lines are called hybridomas, and the antibodies that are produced are called monoclonal antibodies [33]. These antibodies are then proteolytically cleaved to obtain the desired antibody fragments for crystallization. Proteolytic antibody fragments are easy to obtain, but there are several disadvantages which contribute to the heterogeneity of the final preparation. Enzymatic proteolysis is not very specific and may contain residual parts of the linker hinge region. Sometimes crystallization may be hindered by glycosylation of the remaining fragment. Also, purification to homogeneity is sometimes necessary because of the presence of multiple isoforms that hinder crystallization attempts. Importantly, there is no scope for improving the crystallization properties by genetic manipulation [42].

Recombinant antibody fragments are a very convenient alternative. The most common formats used for co-crystallization purposes are scFv (fragment variable, ~ 28 kDa) or Fab (~56 kDa). Single chain Fv fragments are considerably more difficult to produce by enzymatic digestion and so are more commonly prepared by recombinant technology. They are recombinant polypeptides composed of an antibody variable light-chain amino acid sequence (V_L) tethered to a variable heavy-chain sequence (V_H) by a designed peptide linker sequence [43]. Recombinant technology yields highly homogenous antibody fragment preparations that are essential for crystallization and co-crystallization attempts. However, recombinant antibody production is not as straightforward as other overexpressed recombinant proteins due to the presence of several cysteines in the antibody amino acid sequence, and yields of purified protein can be relatively dismal. Chapter 3 describes several strategies such as genetic manipulations,

optimization of the production and purification protocols that were incorporated to improve the yield of a secreted recombinant Fab fragment. Most recombinant antibody fragments are isolated from the *E.coli* periplasm or the supernatant as they need the oxidizing periplasmic milieu to form the disulfide bonds required for stable tertiary structure formation. Mutant strains that provide the much needed oxidizing milieu in the cytoplasm have been recently introduced [44]. Chapter 4 describes the creation of a recombinant antibody fragment that is not secreted into the periplasm. Using the mutant *E.coli* strains, production and purification of this cytoplasmic recombinant antibody fragment was studied and compared to the optimized secreted yield.

Recombinant antibody fragments are routinely constructed from monoclonal antibodies obtained by immune methods. The cDNA of the monoclonal antibody to a target protein is generated from the RNA that is isolated from the spleens of immunized mice by the use of reverse transcriptase. These cDNAs are then cloned into the appropriate protein expression vectors to produce recombinant antibody fragments. However generation of monoclonal antibodies and producing them at the scale required involves a large investment of time, effort and cost. Separate immunizations are required for each antigen, and the cell fusion process required to generate hybridomas is laborious and inefficient. Recombinant technology has eased the problem of obtaining high and homogenous yields of antibody fragments at a lower cost, but the initial step of generating ligands that bind to target proteins by hybridoma technology is still a time-consuming one.

The Current State of Art in Molecular Display Methods

Rapid progress has been made in developing methodologies to identify ligands that bind to proteins. The two main categories of systems used for identifying specific protein-protein

interactions are selection oriented and screening oriented. The selection oriented systems include *in vivo* techniques such as phage display and cell surface display (bacterial and yeast) and *in vitro* techniques such as RNA and ribosome display. An example of screening oriented systems is the two-hybrid analytical technique. Affinity maturation of protein-protein interactions finds wide-ranging applications in both academic and pharmaceutical research. The most popular of these systems is the one that this thesis is based upon -- the phage display system. A brief review of the principles and applications of the phage display system follows.

Phage Display Systems – Principles and Applications

The most popular molecular display and selection system, phage display was introduced by the pioneering work of Smith, which established a method for presenting polypeptides on the surface of filamentous phage, a virus that infects *Escherichia coli* [45].

The characteristic features of filamentous bacteriophages as seen in Figure 1.5, especially the Ff class (f1, fd and M13), include a circular single-stranded DNA genome encased in a long protein capsid cylinder. The Ff phage particle is approximately 6.5 nm in diameter, 930 nm in length and 16.3 MD in mass with 87% of the mass being protein. The capsid cylinder consists of 5 structural proteins that includes 2700 molecules of the 50 amino acid major coat protein pVIII (*gene VIII* product) spanning the length of the cylinder, 5 molecules each of the 33 amino acid pVII (*gene VII* product) and 32 amino acid pIX (*gene IX* product) at one end and 5 molecules each of the 406 amino acid pIII and 112 amino acid pVI (*gene III* and *gene VI* products) at the other end. The 6.4 kb Ff genome is oriented within the capsid cylinder such that a 78-nucleotide hairpin region called the packaging signal (PS) is always located at the end of the particle containing the pVII and pIX proteins. A total of 11 proteins are encoded by the genome, and

based on their function in the phage lifecycle, the genes are grouped as replication protein encoders (*gene II, V* and *X*), capsid protein encoders (*gene VII, VIII, IX, III, VI*) and assembly protein encoders (*gene I, XI* and *IV*). In addition to the coding regions, there is an intergenic region that contains the signals for the initiation of synthesis of both the plus (+) or viral-contained DNA strand, and the (-) strand, the initiation of capsid assembly signal (PS) and the signal for the termination of RNA synthesis [46, 47] [48].

Due to their ability to use the tip of the F conjugative pilus as a receptor, they specifically infect *Escherichia coli* containing the F plasmid. Unlike other phages (e.g., T4, T7), filamentous bacteriophages replicate and assemble without killing the host cell. The phage life cycle starts with infection, a multistep process that requires interactions with the F conjugative pilus and bacterial cytoplasmic membrane proteins and concludes with the conjugation of the single-stranded viral DNA. The life cycle continues with the replication and protein synthesis step in which the single-stranded viral (+) DNA is replicated via a double-stranded intermediate by a mixture of bacterial and phage- encoded components. This results in the formation of newly synthesized viral single-stranded DNA in a complex with several copies of a phage-encoded single-stranded DNA-binding protein (pV). The phage capsid structural proteins are also synthesized at this time and they stay associated as membrane proteins till the next step – assembly. During the assembly process, the viral DNA is extruded through the host membrane and the capsid proteins are packaged around the DNA. This process continues until the end of the genome DNA. Productive infections result in viral release via extrusion across the bacterial membranes (at the rate of 200 to 2000 progeny per infected cell per doubling time), with the infected cells continuing to grow and divide with a generation time approximately 50% longer than that of uninfected bacteria [46, 47].

The basic principle underlying all phage display systems is the physical linkage of a polypeptide's phenotype to its corresponding genotype. The hydrophobic regions of the capsid proteins stay membrane associated and are involved in the interactions involved in the viral assembly process. Thus any foreign protein fused to the periplasmic portion of these capsid proteins will not interfere with the assembly process and will have a good chance of being packaged into a phage particle provided it can be translocated efficiently across the inner membrane. Popularly, foreign proteins are expressed as coat protein pIII or pVIII fusions and directed to the bacterial periplasm by an appropriate N terminally attached signal sequence. The fused foreign proteins get assembled along with the coat proteins and displayed on the phage surface. The genetic information encoding the displayed fusion protein is packaged inside the same phage particle in the form of a single-stranded DNA (ssDNA) molecule. Hence, the genotype–phenotype coupling occurs before the phages are released into the extracellular environment, ensuring that phages produced by one bacterial cell are identical [49, 50].

A phage display library is an ensemble of up to about 10 billion such unique phage clones, each possessing a different foreign coding sequence within its single stranded DNA and thus displaying a different foreign protein on the virion surface [51]. The classical application of such libraries aims at affinity selection wherein tens of millions of displayed proteins or peptides can be surveyed for tight binding to an antibody, receptor or other binding protein. A target binding molecule is immobilized on a solid support such as paramagnetic beads or on the polystyrene surface of an ELISA well and incubated with a phage display library. Phage particles whose displayed proteins bind the target molecule are captured on the support while the other phages are washed away. The captured phage is generally a tiny fraction of the initial phage population and can be eluted from the support without destroying phage infectivity. These phage

are used to infect fresh bacterial host cells that amplify the number of the selected phage clones. The resulting phage are subjected to another round of affinity purification. A few such rounds suffice to survey a library with billions or even trillions of initial clones for rare, displayed foreign proteins with high affinity for the target molecule. This type of affinity selection, using several rounds of specific binding between the displayed protein and its immobilized binding partner, is referred to as "panning" as depicted in Figure 1.6. After several panning rounds, individual phage clones are propagated and their ability to bind the target protein confirmed.

On the basis of the vector system used for phage production, phage display systems can be grouped in to two classes – phage vector based display and phagemid vector based display. The first family of filamentous phage vectors was introduced by Messing *et al.* [52] and represents true phage vectors directly derived from the genome of filamentous phage. They encode all the proteins needed for the replication and assembly of the filamentous phage. Such vectors are *either* type *n* - single phage vector genomes that include the recombinant coat protein gene (for instance, type 3 or type 8 wherein the displayed protein is fused to each copy of pIII or pVIII coat protein respectively) [53] [51], *or* type *nn* – phage vector genomes that harbor two genes encoding a particular coat protein (gene *III* for instance in the vector type 33); one encodes a wild-type subunit pIII subunit while the other encodes the recombinant pIII with the displayed protein [54]. The second class of phage display systems is based on phagemid vectors (type 3+3, type 8+8) [55-57]. Simply put, a phagemid is a plasmid that bears both a plasmid as well as a phage-derived origin of replication. The displayed protein-coat protein fusion encoding gene is placed in the phagemid genome, while the wild-type coat protein is on a helper phage. The phagemid maintains itself as a plasmid and directs protein expression in bacteria if desired. Propagation of phagemids in cells superinfected with a helper phage results in packaging of

phagemid DNA in a similar fashion to phage DNA. The helper phage (M13K07 or VCSM13) provides all the phage-derived components required for phage replication and thus "helps" replicate and package the phagemid genome. Both the helper phage and rescued phagemid virions have mosaic capsids composed of a mixture of diplayed protein-coat protein fusion and wild-type coat protein molecules. Typically, a helper phage with a defective origin of replication or packaging signal is used to allow the preferential packaging of the phagemid genome over the helper-phage genome. Phagemid systems have several advantages including the ability to maintain large DNA inserts, ease of producing high yields of double stranded DNA by simple plasmid preps and the ability of two gene systems to allow valency modulation of the displayed fusion protein.

This mini-review of filamentous phage biology, phage display principles, systems and methodology will not be complete without a look at the prominent applications of phage display [45]. Progress in the field of phage display, for affinity selection of peptides and proteins [58], led to the development of antibody libraries, which are probably now the most important and commercially successful application of the phage display technology [59] [60] [55] [61]. Phage antibody libraries have had an immense range of applications from drug discovery (to discover novel therapeutic targets and isolate therapeutic monoclonal antibodies) to functional genomics and proteomics research (high-throughput selection and screening).

Antibodies are very efficiently displayed as scFvs [62] [60], Fabs [55, 61, 63], immunoglobulin variable fragments (Fvs) with an engineered intermolecular disulphide bond to stabilize the V_H-V_L pair [64] and diabody fragments [65]. Phage antibody libraries can be naïve, immune or synthetic libraries. Immune libraries are created from the IgG genes of spleen B cells or hybridomas of mice immunized with antigen or from immune donors [62]. These libraries are

enriched in antigen-specific antibodies. Isolated antibodies can be rapidly produced or manipulated and libraries can be constructed from a variety of species. Immune phage libraries, thus, are very useful in analyzing natural humoral responses or studying *in vitro* immunization procedures [66]. However immune phage libraries have the disadvantage of isolating antibodies only against the set of antigens to which an immune response was induced. Repeated immunization and library construction for each different set of antigens becomes necessary. Active immunization, however, is not always possible because of ethical issues and tolerance or toxicity problems. The "single-pot" libraries are antigen unbiased and independent of the donor's immunological history [67, 68]. Based on the source of the immunoglobulin genes, single-pot libraries are either naïve or synthetic. Naïve libraries are constructed from rearranged V genes harvested from unimmunized donors while synthetic libraries are built artificially by in vitro assembly of V-gene segments and D/J segments. Single-pot libraries can be used to generate antibodies to a large panel of antigens without ever having to immunize an animal. Irrespective of these logistics, the phage antibody display technology shows enormous promise in several areas of academic as well as commercial research. The phage antibody display technology has made immense progress in design and innovation and has become a powerful tool for drug and target discovery.

This ability of antibody fragment libraries to be displayed on the surface of phage, coupled with the capability to obtain high, homogenous yields of recombinant antibodies and antibody fragments, has made antibody phage display a much more viable alternative over immune methods (hybridoma/immune donors) for the generation of antibodies of high affinity, specificity and avidity for a variety of targets. The successful integration of the recombinant antibody fragment mediated co-crystallization approach into the structural genomics pipeline

relies on the pace and specificity at which antibody phage display can be used to generate antibody fragments and ultimately diffraction quality crystals of non-crystallizing proteins. In chapter 5, rapid antibody phage display methodology has been devised, developed and compared with standard phage display techniques. Rapid phage display approaches developed by other groups, their advantages and pitfalls will be discussed. Incorporation of this rapid, high-throughput methodology coupled with the innovative automation technology available will greatly improve the generation of antibody fragment co-crystallization agents by phage display.

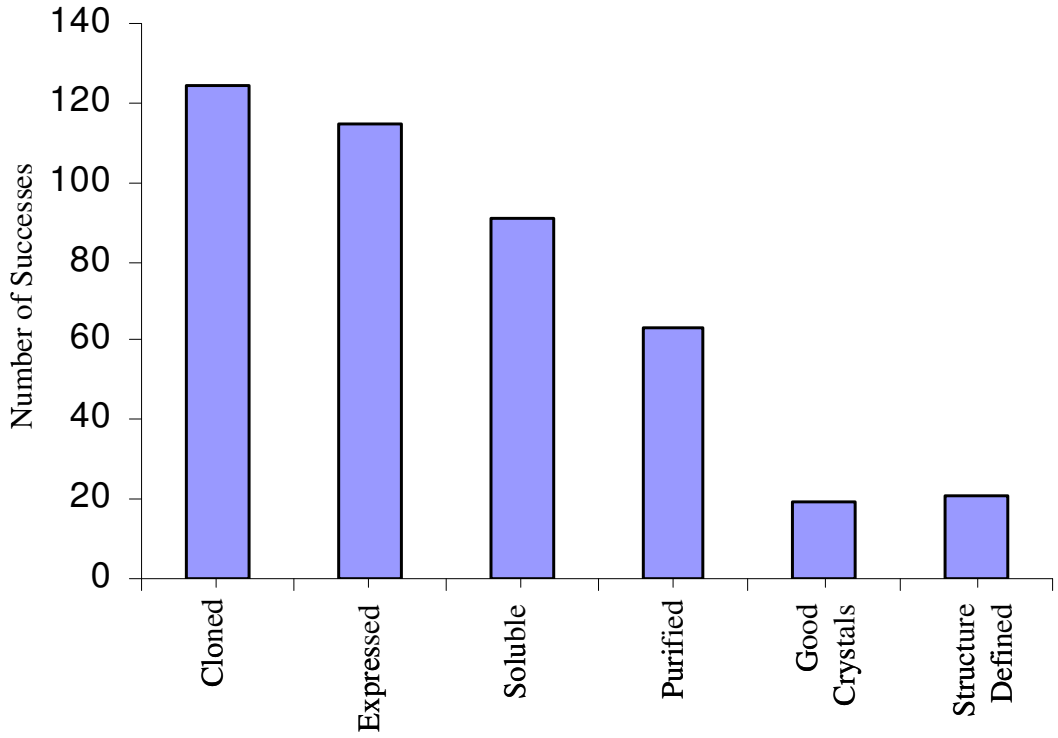


Figure 1.1 The Bottleneck of Macromolecular X-ray Crystallography

The histogram shows the different stages involved from cloning to X-ray structure determination and their relative success rates. Data plotted using statistics reported in http://proteome.bnl.gov/progress.html, as on January 18, 2005.

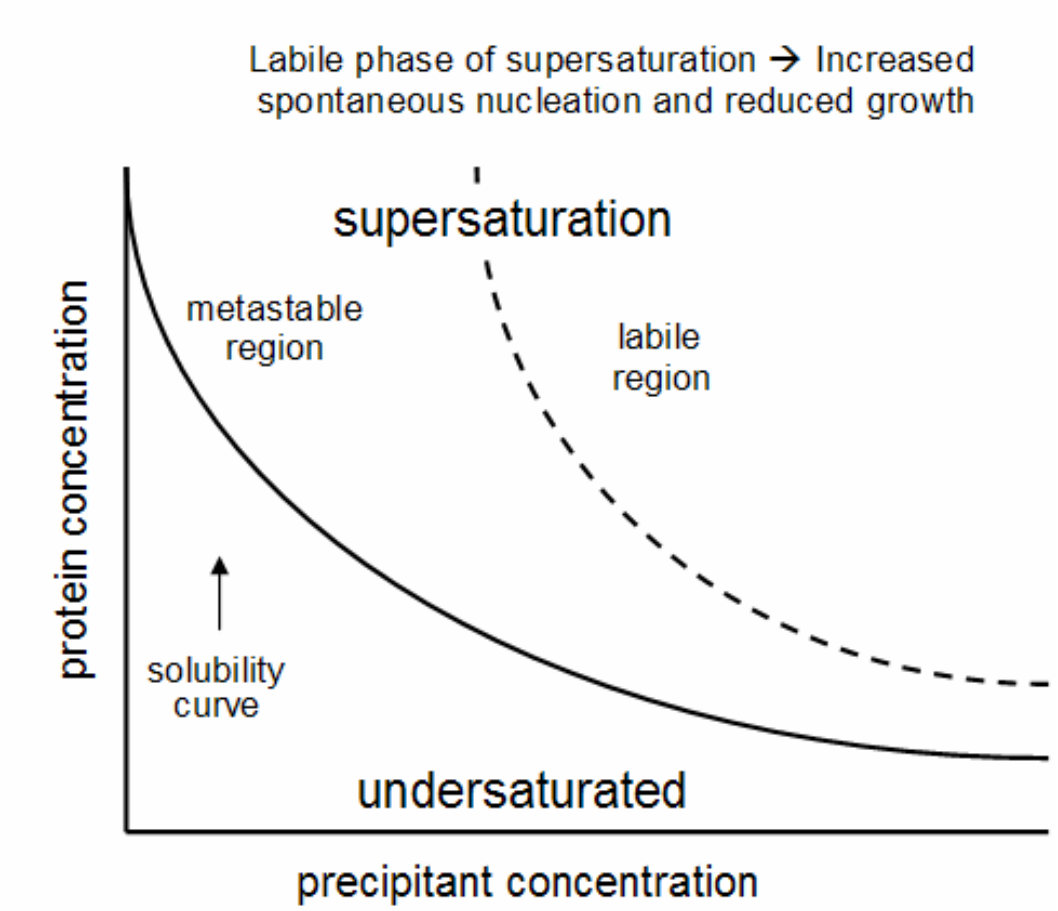


Figure 1.2 A schematic solubility diagram

Crystallization from solution showing the labile, metastable and undersaturated regions. The solid line represents the saturation limit of the compound in the solvent [1].

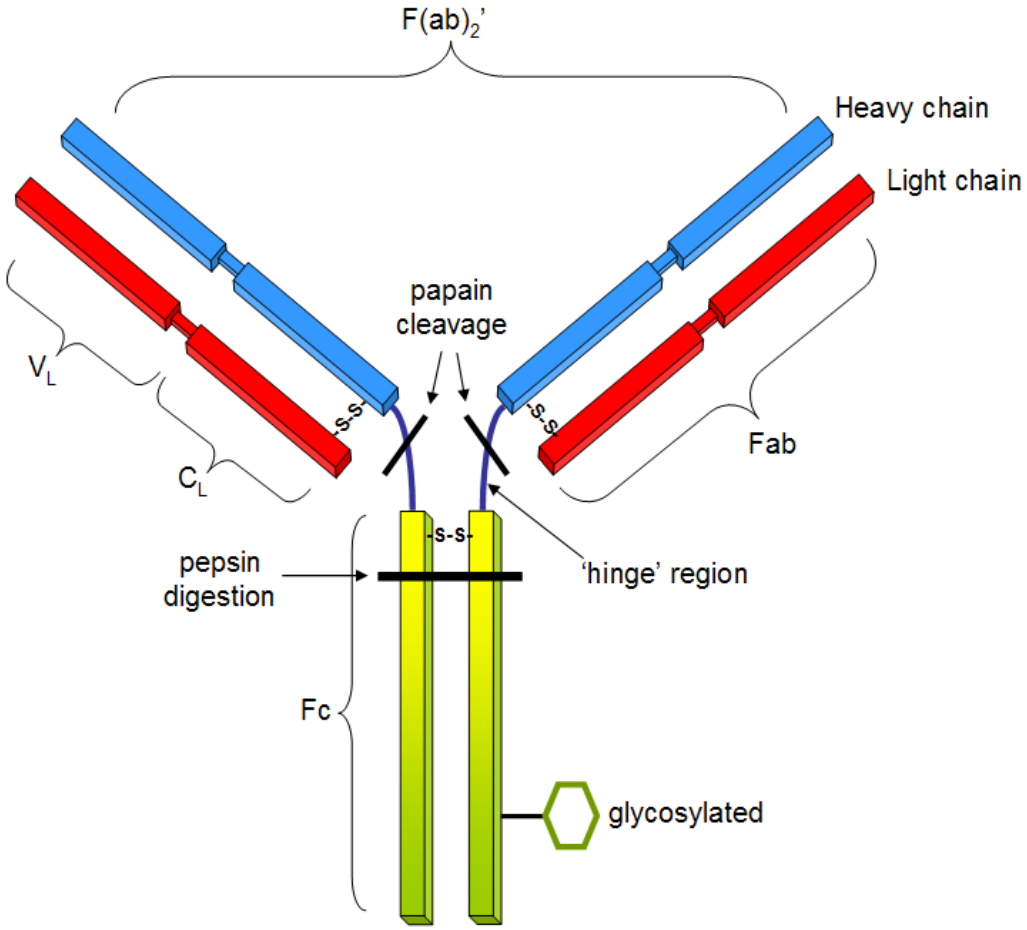


Figure 1.3 Schematic diagram of an Ig molecule. Abbreviations used are Fab, antibody fragment derived from papain cleavage; Fab2', bivalent pepsin antibody fragment derived from pepsin digestion; CDR, complementarity determining region; V_L, variable domain of the light chain; C_L, constant domain of the light chain; Fc, heavy chain constant domain tail fragment [2].

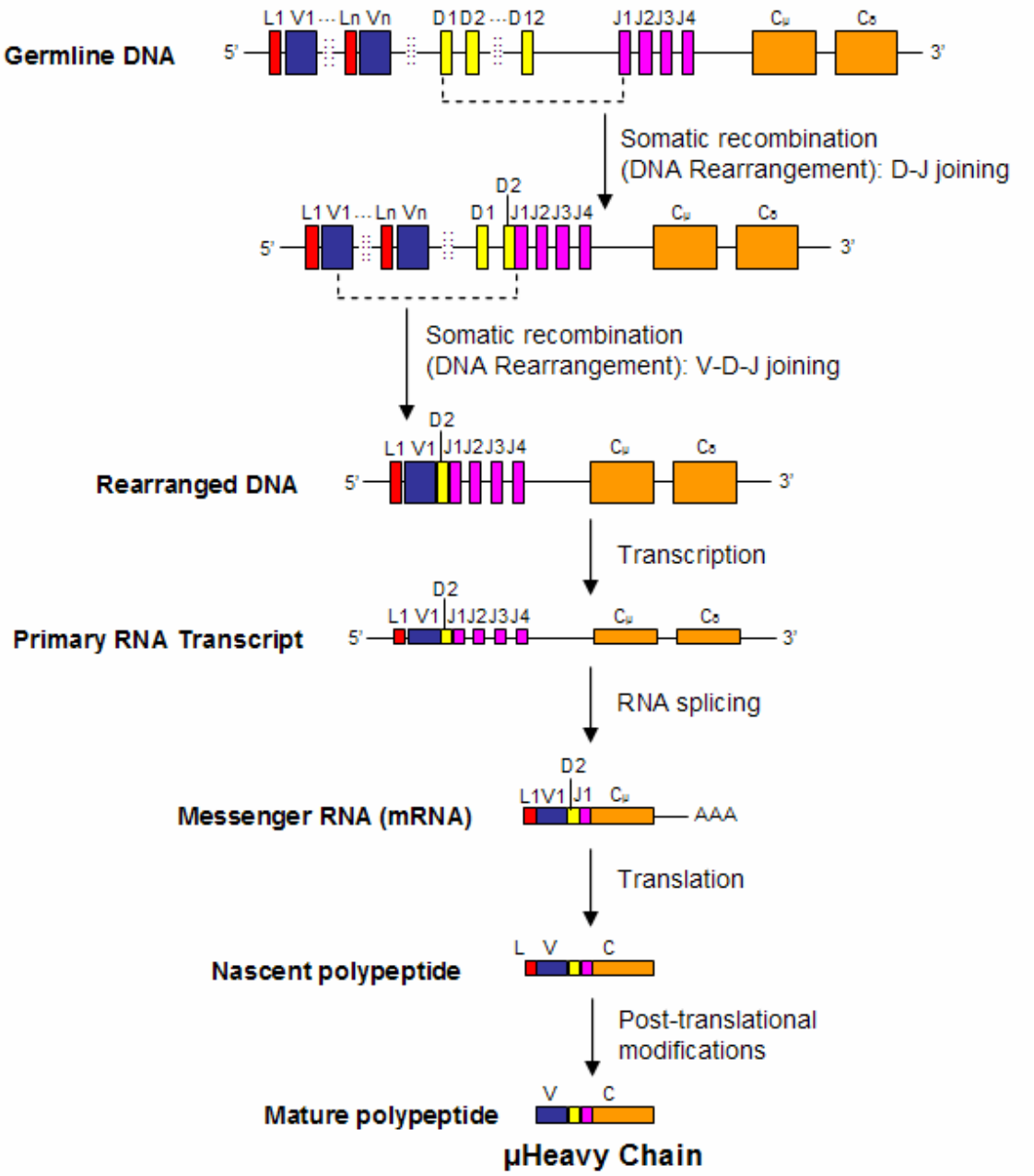


Figure 1.4 Somatic Recombination and Expression of Ig germline DNA.

The figure depicts somatic recombination and DNA rearrangement to produce the final variable (V) region of the μ heavy chain from the germline V1, D2 and J1 [3].

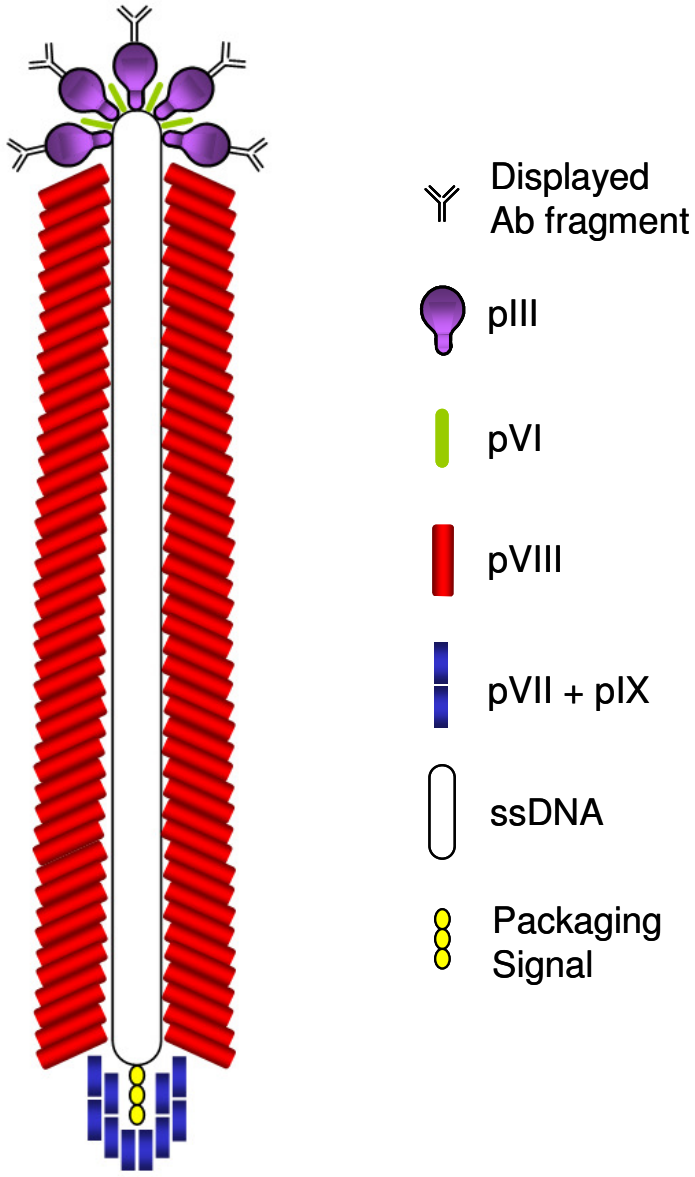


Figure 1.5 Schematic sketch of a Filamentous Phage

These phages are composed of a circular single-stranded DNA genome encased in a long protein cylinder composed of five different proteins as shown. At one end of the phage genome is the packaging signal.

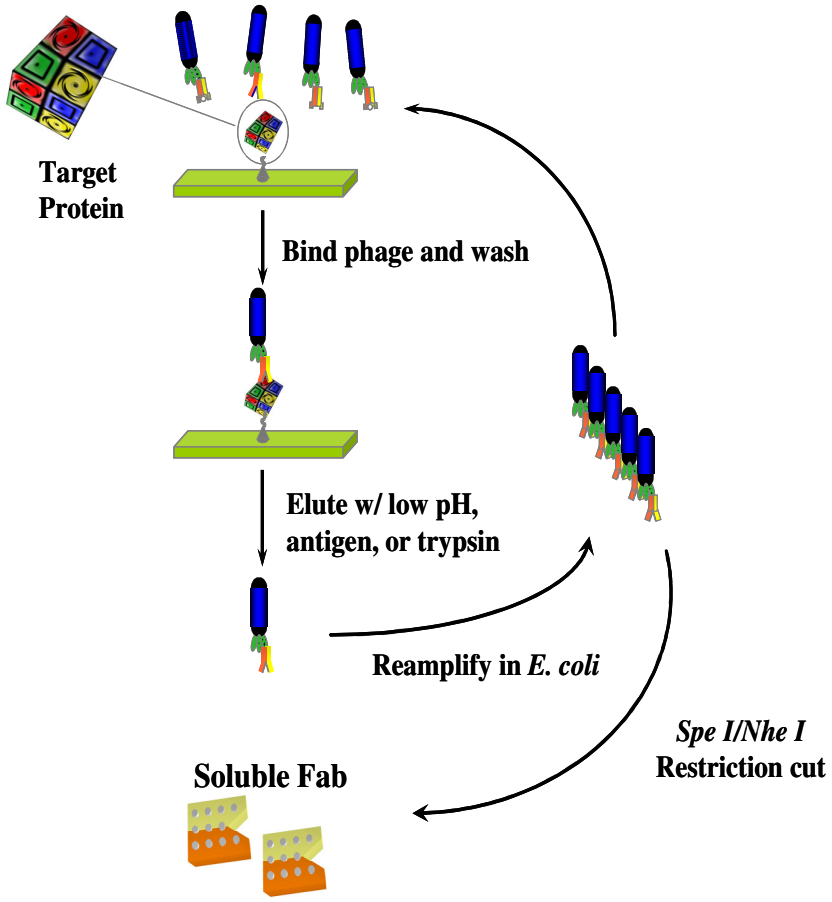


Figure 1.6 A "panning" cycle using phage display is illustrated. Phage are shown as blue tubes expressing Fab (yellow and orange cylinders) as a fusion with the gene III product (green).

References

[1] A. McPherson, Crystallization of biological macromolecules, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1999.

[2] L.-L.C. Kelley, University of Georgia., Co-crystallization recombinant mouse antibody fragment library biotechnology, 2003, pp. xii, 160 leaves.

[3] A.K. Abbas, A.H. Lichtman, J.S. Pober, Lymphocyte Maturation and Expression of Antigen Receptor Genes pp. 125-160, in: A.K. Abbas, A.H. Lichtman, J.S. Pober (Eds.) Cellular and molecular immunology, W.B. Saunders, Philadelphia, 2000, pp. vii, 553.

[4] A. McPherson, Current approaches to macromolecular crystallization. Eur J Biochem 189 (1990) 1-23.

[5] J.D. Bernal, D. Crowfoot, X-ray photographs of crystalline pepsin. Nature 133 (1934) 794-795.

[6] J.B. Sumner, THE ISOLATION AND CRYSTALLIZATION OF THE ENZYME UREASE. PRELIMINARY PAPER, 1926, pp. 435-441.

[7] J.B. Sumner, A.L. Dounce, CRYSTALLINE CATALASE, 1937, pp. 417-424.

[8] J.H. Northrop, M. Kunitz, R.M. Herriott, Crystalline Enzymes, Columbia University Press, New York, NY, 1948.

[9] M.L. Pusey, Z.J. Liu, W. Tempel, J. Praissman, D. Lin, B.C. Wang, J.A. Gavira, J.D. Ng, Life in the fast lane for protein crystallization and X-ray crystallography. Prog Biophys Mol Biol 88 (2005) 359-386.

[10] Z.S. Derewenda, The use of recombinant methods and molecular engineering in protein crystallization. Methods

Macromolecular Crystallization 34 (2004) 354-363.

[11] N.E. Chayen, Tackling the bottleneck of protein crystallization in the post-genomic era. Trends in biotechnology 20 (2002) 98.

[12] A. McPherson, Introduction to protein crystallization. Methods 34 (2004) 254-265.

[13] N.E. Chayen, Turning protein crystallisation from an art into a science. Curr Opin Struct Biol 14 (2004) 577-583.

[14] N. Asherie, Protein crystallization and phase diagrams. Methods Macromolecular Crystallization 34 (2004) 266-272.

[15] N.E. Chayen, P.D. Shaw Stewart, D.M. Blow, Microbatch crystallization under oil -- a new technique allowing many small-volume crystallization trials. Journal of Crystal Growth 122 (1992) 176-180.

[16] J. Jancarik, S.H. Kim, Sparse matrix sampling: a screening method for crystallization of proteins, Journal of Applied Crystallography, 1991, pp. 409-411.

[17] D. Hennessy, B. Buchanan, D. Subramanian, P.A. Wilkosz, J.M. Rosenberg, Statistical methods for the objective design of screening procedures for macromolecular crystallization. Acta Crystallogr D Biol Crystallogr 56 (2000) 817-827.

[18] B. Rupp, J. Wang, Predictive models for protein crystallization. Methods Macromolecular Crystallization 34 (2004) 390-407.

[19] U. Mueller, L. Nyarsik, M. Horn, H. Rauth, T. Przewieslik, W. Saenger, H. Lehrach, H. Eickhoff, Development of a technology for automation and miniaturization of protein crystallization. Journal of Biotechnology 85 (2001) 7-14.

[20] E. Abola, P. Kuhn, T. Earnest, R.C. Stevens, Automation of X-ray crystallography. Nat Struct Mol Biol (2000).

[21] A. McPherson, A.J. Malkin, Y.G. Kuznetsov, The science of macromolecular crystallization. Structure 3 (1995) 759-768.

[22] L.C. Kovari, C. Momany, M.G. Rossmann, The use of antibody fragments for crystallization and structure determinations. Structure 3 (1995) 1291-1293.

[23] C. Ostermeier, H. Michel, Crystallization of membrane proteins. Curr Opin Struct Biol 7 (1997) 697-701.

[24] S. Sheriff, E.W. Silverton, E.A. Padlan, G.H. Cohen, S.J. Smith-Gill, B.C. Finzel, D.R. Davies, Three-dimensional structure of an antibody-antigen complex. Proceedings of the National Academy of Sciences of the United States of America 84 (1987) 8075-8079.

[25] E.A. Padlan, E.W. Silverton, S. Sheriff, G.H. Cohen, S.J. Smith-Gill, D.R. Davies, Structure of an antibody-antigen complex: crystal structure of the HyHEL-10 Fab-lysozyme complex. Proceedings of the National Academy of Sciences of the United States of America 86 (1989) 5938-5942.

[26] W.R. Tulip, J.N. Varghese, W.G. Laver, R.G. Webster, P.M. Colman, Refined crystal structure of the influenza virus N9 neuraminidase-NC41 Fab complex. J Mol Biol 227 (1992) 122-148.

[27] S. Iwata, C. Ostermeier, B. Ludwig, H. Michel, Structure at 2.8 A resolution of cytochrome c oxidase from Paracoccus denitrificans. Nature 376 (1995) 660-669.

[28] C. Momany, L.C. Kovari, A.J. Prongay, W. Keller, R.K. Gitti, B.M. Lee, A.E. Gorbalenya, L. Tong, J. McClure, L.S. Ehrlich, M.F. Summers, C. Carter, M.G. Rossmann, Crystal structure of dimeric HIV-1 capsid protein. Nature structural biology 3 (1996) 763-770.

[29] C. Hunte, J. Koepke, C. Lange, T. Rossmanith, H. Michel, Structure at 2.3 A resolution of the cytochrome bc(1) complex from the yeast Saccharomyces cerevisiae co-crystallized with an antibody Fv fragment. Structure 8 (2000) 669-684.

[30] Y. Zhou, J.H. Morais-Cabral, A. Kaufman, R. MacKinnon, Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. Nature 414 (2001) 43-48.

[31] C. Hunte, H. Michel, Crystallisation of membrane proteins mediated by antibody fragments. Curr Opin Struct Biol 12 (2002) 503-508.

[32] W.G. Laver, Crystallization of antibody-protein complexes. Methods: A Companion to Methods of Enzymology 1 (1990) 70-74.

[33] A.K. Abbas, A.H. Lichtman, J.S. Pober, Cellular and molecular immunology, W.B. Saunders, Philadelphia, 2000.

[34] D.R. Davies, E.A. Padlan, S. Sheriff, Antibody-antigen complexes. Annual review of biochemistry 59 (1990) 439-473.

[35] I.A. Wilson, R.L. Stanfield, Antibody-antigen interactions: new structures and new conformational changes. Curr Opin Struct Biol 4 (1994) 857-867.

[36] L.J. Harris, S.B. Larson, E. Skaletsky, A. McPherson, Comparison of the conformations of two intact monoclonal antibodies with hinges. Immunol Rev 163 (1998) 35-43.

[37] C. Chothia, A.M. Lesk, Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol 196 (1987) 901-917.

[38] A.M. Lesk, C. Chothia, Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint. Nature 335 (1988) 188-190.

[39] N.F. Landolfi, A.B. Thakur, H. Fu, M. Vasquez, C. Queen, N. Tsurushita, The integrity of the ball-and-socket joint between V and C domains is essential for complete activity of a humanized antibody. J Immunol 166 (2001) 1748-1754.

[40] R.R. Porter, The hydrolysis of rabbit y-globulin and antibodies with crystalline papain. Biochem J 73 (1959) 119-126.

[41] G. Kohler, C. Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256 (1975) 495-497.

[42] M. Venturi, C. Seifert, C. Hunte, High level production of functional antibody Fab fragments in an oxidizing bacterial cytoplasm. J Mol Biol 315 (2002) 1-8.

[43] R.E. Bird, K.D. Hardman, J.W. Jacobson, S. Johnson, B.M. Kaufman, S.M. Lee, T. Lee, S.H. Pope, G.S. Riordan, M. Whitlow, Single-chain antigen-binding proteins. Science 242 (1988) 423-426.

[44] W.A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds in the Escherichia coli Cytoplasm 10.1074/jbc.272.25.15661. J. Biol. Chem. 272 (1997) 15661-15667.

[45] G.P. Smith, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228 (1985) 1315-1317.

[46] R.E. Webster, Filamentous Phage Biology pp. 1.1-1.37, in: C.F. Barbas (Ed.) Phage display : a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001, p. 1 v. in various pagings.

[47] D.J. Rodi, S. Mandava, L. Makowski, Filamentous Bacteriophage Structure and Biology, in: S.S. Sidhu (Ed.) Drug discovery series ; 3 Phage display in biotechnology and drug discovery, Taylor & Francis/CRC Press, Boca Raton, FL, 2005, pp. xviii, 748.

[48] D.A. Marvin, B. Hohn, Filamentous bacterial viruses. Bacteriol Rev 33 (1969) 172-209.

[49] M. Paschke, Phage display systems and their applications. Appl Microbiol Biotechnol 70 (2006) 2-11.

[50] G.P. Smith, V.A. Petrenko, Phage Display. Chem Rev 97 (1997) 391-410.

[51] J.K. Scott, G.P. Smith, Searching for peptide ligands with an epitope library. Science 249 (1990) 386-390.

[52] J. Messing, B. Gronenborn, B. Muller-Hill, P. Hans Hopschneider, Filamentous coliphage M13 as a cloning vehicle: insertion of a HindII fragment of the lac regulatory region in M13 replicative form in vitro. Proceedings of the National Academy of Sciences of the United States of America 74 (1977) 3642-3646.

[53] S.E. Cwirla, E.A. Peters, R.W. Barrett, W.J. Dower, Peptides on phage: a vast library of peptides for identifying ligands. Proceedings of the National Academy of Sciences of the United States of America 87 (1990) 6378-6382.

[54] M.A. McLafferty, R.B. Kent, R.C. Ladner, W. Markland, M13 bacteriophage displaying disulfide-constrained microproteins. Gene 128 (1993) 29-36.

[55] C.F. Barbas, 3rd, A.S. Kang, R.A. Lerner, S.J. Benkovic, Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proceedings of the National Academy of Sciences of the United States of America 88 (1991) 7978-7982.

[56] S. Bass, R. Greene, J.A. Wells, Hormone phage: an enrichment method for variant proteins with altered binding properties. Proteins 8 (1990) 309-314.

[57] F. Breitling, S. Dubel, T. Seehaus, I. Klewinghaus, M. Little, A surface expression vector for antibody screening. Gene 104 (1991) 147-153.

[58] J.J. Devlin, L.C. Panganiban, P.E. Devlin, Random peptide libraries: a source of specific protein binding molecules. Science 249 (1990) 404-406.

[59] J. McCafferty, A.D. Griffiths, G. Winter, D.J. Chiswell, Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348 (1990) 552-554.

[60] J.D. Marks, A.D. Griffiths, M. Malmqvist, T.P. Clackson, J.M. Bye, G. Winter, By-passing immunization: building high affinity human antibodies by chain shuffling. Biotechnology (N Y) 10 (1992) 779-783.

[61] H.R. Hoogenboom, A.D. Griffiths, K.S. Johnson, D.J. Chiswell, P. Hudson, G. Winter, Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic acids research 19 (1991) 4133-4137.

[62] T. Clackson, H.R. Hoogenboom, A.D. Griffiths, G. Winter, Making antibody fragments using phage display libraries. Nature 352 (1991) 624-628.

[63] L.J. Garrard, M. Yang, M.P. O'Connell, R.F. Kelley, D.J. Henner, Fab assembly and enrichment in a monovalent phage display system. Biotechnology (N Y) 9 (1991) 1373-1377.

[64] U. Brinkmann, P.S. Chowdhury, D.M. Roscoe, I. Pastan, Phage display of disulfide-stabilized Fv fragments. Journal of immunological methods 182 (1995) 41-50.

[65] B.T. McGuinness, G. Walter, K. FitzGerald, P. Schuler, W. Mahoney, A.R. Duncan, H.R. Hoogenboom, Phage diabody repertoires for selection of large numbers of bispecific antibody fragments. Nature biotechnology 14 (1996) 1149-1154.

[66] H.R. Hoogenboom, Overview of Antibody Phage-Display Technology and Its Applications pp. 1-37, in: P.M. O'Brien, R. Aitken (Eds.) Antibody phage display : methods and protocols, Humana Press, Totowa, N.J., 2002, pp. xiii, 401.

[67] G. Winter, A.D. Griffiths, R.E. Hawkins, H.R. Hoogenboom, Making antibodies by phage display technology. Annu Rev Immunol 12 (1994) 433-455.

[68] H.R. Hoogenboom, Designing and optimizing library selection strategies for generating high-affinity antibodies. Trends in biotechnology 15 (1997) 62-70.

CHAPTER 2

ADDRESSING THE CRYSTALLIZATION BOTTLENECK BY CO-CRYSTALLIZATION[1]

[1] Nadkarni, A. and Momany, C. Submitted to *Protein Science*

Abstract

The crystallization of macromolecules as an intermediate in the production of protein structures is often the rate-limiting factor. Recent achievements in the field of protein crystallization such as automation, availability of commercial crystallization kits and databases have significantly improved the throughput of the crystallization screening. But this high throughput has not led to high output due to the crystallization bottleneck, and thus, there is a significant disparity between the pace of gene sequencing, protein production, and the determination of the gene product's atomic structure. Development of complementary approaches that improve the crystallization potential of a protein would form a useful addition to the crystallographer's toolbox. One such approach is to co-crystallize a target protein with another protein of known structure that specifically recognizes the target so that a stable crystal complex is formed. Such proteins are known as co-crystallization proteins (CCPs). In this review, the various mechanisms by which CCPs improve the crystallization odds of a protein are explained. Additionally, we review some common CCP reagents such as Fabs and scFvs, advances and current issues in the methodology of their generation and use, and also discuss some recently developed novel CCP reagents such as ankyrin repeat proteins and affibodies.

Keywords: co-crystallization; phage display; RNA display; ribosome display; Fab; scFv; antibody fragments; ankyrin domain

Introduction

X-ray crystallography has played a fundamental role in connecting the dots between genomic data and biological function by providing accurate structural information to resolve several significant research problems. Solving protein structures by X-ray crystallography is contingent upon the availability of ordered, diffraction- quality crystals. However, according to the figures obtained from the Structural Genomics Information Portal maintained by the Research Collaboratory for Structural Bioinformatics (RCSB), the success rate of obtaining such high quality protein crystals relative to the total number of "cloned" target ORFs is a disappointing 4.5% (http://targetdb.pdb.org/statistics/TargetStatistics.html#table1). Huge investments have been made in scaling the machinery of structure determinations (high intensity X-ray synchrotrons, powerful computer software and databases) and tremendous strides have been made in the fields of high throughput cloning, protein production, purification and automated crystallization screening techniques. Despite this, many proteins, in particular ones that have low solubility and easily aggregate, and even a large percent of soluble proteins, still fail to produce good crystals. The membrane proteins and glycosylated proteins, which form a significant percent of the drug discovery targets, are disproportionately represented in the protein structure data banks. To address this bottleneck and to improve the crystallization potential of the target protein, robust parallel approaches need to be developed. One useful approach to addressing the crystallization problem is to co-crystallize the target protein with another protein that specifically recognizes the target so that a stable complex is formed.

A co-crystallization protein (CCP) is a pre-existent protein molecule that can be modified to draw another protein target with it, leading to the assembly of the two macromolecules into a crystal lattice. Interaction with a CCP modifies the surfaces of the target in ways favorable for

crystallization. This is especially true for proteins that contain limited hydrophilic domains, such as membrane proteins, because these polar domains are critical in the formation of important lattice contacts. Binding and subsequent co-crystallization of such a protein with a CCP extends the available crystallization surface so as to improve the probability of obtaining ordered crystals. For example, in the K^+ channel-Fab complex the channel's polar surface was extended by the Fab attached to its extra-cellular surface. Lattice forming crystal contacts were formed by the adjacent Fab fragments and not by the four subunits of the K^+ channel tetramer, allowing the channel to be suspended in the center in a natural biological conformation (Zhou et al. 2001). Additionally, such CCP lattices provide enough space to accommodate large micelles of detergent solubilized proteins (Hunte and Michel 2002). Co-crystallization of a protein with its cognate CCP can also stabilize the protein by "locking down" flexible domains. For example, in the ternary CD4-gp120-Fab complex, the Fab stabilized the flexible regions of the CD4-gp120 complex interface, which was found to be the chemokine receptor binding site that is absolutely essential for HIV-1 entry into cells (Kwong et al. 1998). For several proteins, crystallization is inhibited due to uncontrolled aggregation or biologically significant oligomerization. Proteins such as the viral coat proteins like the HIV-1 capsid protein p24 have oligomerization domains through which they interact with each other to form the viral coat core. This natural feature of the viral capsid protein makes the proteins resistant to crystallization as they tend to aggregate in solution. The co-crystallization approach helps sterically block the protein interaction domains and maintains the protein in a relatively more monodisperse solution (Kovari et al. 1995).

Apart from helping a protein's crystallization process, a CCP crystal of known molecular structure can be used as a molecular replacement model (Zhou et al. 2001) or as a recipient of heavy atom labels to determine molecular phasing information for the target protein-CCP

complex (Hunte and Michel 2002). Additional uses of CCPs include immunoaffinity purification of the target protein-CCP (Kleymann et al. 1995). Furthermore, there are instances where the co-crystallization approach has not only rescued the structural studies of several interesting but non-crystallizing proteins, but also helped in improving poor quality protein crystals and enhancing their X-ray diffraction patterns (Laver 1990). Table 2.1 provides a list of proteins crystallized with the help of a CCP. Many more examples exist of macromolecular complexes that crystallized, but this table focuses on cases where the complex was directly used as the means of getting a structure of the target.

In order to be an effective co-crystallization protein (CCP), the CCP should have reasonable solubility, a surface topology capable of recognizing native conformational epitopes of target proteins such that it does not interfere with the target's biological activity, and the ability to form strong and specific noncovalent bonds with the target protein so as to produce stable and dynamically restricted complexes. Fig. 2.1 depicts the popularly used and novel CCP molecules. The space-filled regions demark the protein/antigen binding interfaces of the respective CCPs (Fab, ankyrin domain, scFv and affibody).

Antibody Fragments as Co-crystallization Proteins

Antibody structures and their interactions with antigens have been well studied and reviewed (Wilson and Stanfield 1994; Harris et al. 1998). The basic IgG antibody consists of two light chain monomers associated with two heavy chain monomers through disulfide bonds. Structurally, they are divided into two major domains, a variable domain (V_L and V_H) that is linked to a constant domain (C_L and C_H) by a flexible linker or "elbow joint" (Lesk and Chothia 1988). A potential of six loops, three from the heavy and three from the light chain, interact with

antigens and confer the majority of ligand binding potential. These are the "complementarity determining regions" or CDR loops. Substantial diversity of sequences occurs in the CDR loops as a result of antibody gene rearrangements and somatic mutations. The framework segments, which lie between each CDR loop, are well conserved structurally and define the β-barrel core of the antibody. Hinge region and elbow joints present a high degree of molecular motion within the different domains of an antibody (Abbas et al. 2000). This high level of flexibility along with the bivalent binding mode of an intact antibody makes the crystallization of an intact antibody problematic. In the search to obtain smaller, rigid and functional binding units, the antibody molecule is reduced by proteolysis to fragments of decreasing size.

Most of the protein structures solved as co-crystallization complexes to date have used Fabs (Fragment Antigen Binding) generated proteolytically from monoclonal antibodies (Table 1). Fabs include the full light chain (V_L and C_L) and part of the heavy chain (V_H and C_{H1}). Fv antibody fragments have on the variable regions of both chains. Proteolytically derived Fabs have the disadvantage of being a surprisingly heterogeneous population with multiple isoforms present that can often hinder crystallization attempts, if they are not resolved (Kovari et al. 1995). Besides, there is no scope to improve Fab's crystallization properties by genetic manipulation of the monoclonal cells. Another serious issue with Fabs derived directly from monoclonal antibodies is that the nature of the immune system does not inherently guarantee selection of antibodies to native conformations of the proteins. For B cells to be activated by T helper cells to produce specific antibodies, surface peptide presentation through the Major Histocompatibility Complex is necessary. At this point, epitopes not associated with the native conformation are possible. Only through careful screening were Fabs derived from monoclonal antibodies found to the native conformation NhaA Na+/H+ antiporter from *E. coli* (Padan et al.

1998). Recombinant antibody technology offers a very convenient alternative to obtain versatile and high yields of homogeneous antibody fragments. Formats used routinely for co-crystallization of proteins are the scFv (single-chain fragment variable) and less so, recombinant Fabs (Table 2.1). A scFv is a recombinant polypeptide consisting of a single chain composed of an antibody variable light chain (V_L) tethered in-frame to a variable heavy chain (V_H) by a designed peptide linker sequence (Bird et al. 1988). These antibody fragments lack a flexible elbow region that could be inhibitory to crystallization. However, use of scFvs may be complicated due to their sequence-dependent tendency to form higher molecular weight dimers and trimers leading to a polydisperse solution and bivalent binding modes (Holliger et al. 1993). Fabs, on the other hand, are more robust and the presence of the both the constant domains stabilizes and orients the variable domains to produce tight interactions with the antigen (Rothlisberger et al. 2005). As can be seen from Fig 2.1, Fabs have significantly larger surfaces than Fvs or scFvs. Fabs and scFv fragments are the most widely used antibody fragment formats for co-crystallization as compared to other formats such as single domain antibodies (Ward et al. 1989; Desmyter et al. 1996), diabodies, and Fv fragments (Ostermeier et al. 1995).

One recent variation of the Fab-based co-crystallization approach is to target surface loops by genetically introducing known sequences into the target protein for which monoclonal antibodies are readily available, such as a purification tag like the FLAG polypeptide (Roosild et al., 2006). While the complex of the anti-FLAG Fab and the detergent-solubilized K^+ channel protein, KvPae, did not result in crystals of the target protein, the complex had improved solubility properties that suggested that the chances for crystallization of the membrane protein were improved.

Non-immunoglobulin Co-crystallization Proteins

Since the early 1990s, antibody fragments selected by immune or non-immune methods have been the first choice for obtaining high affinity reagents for a wide variety of biomedical and biochemical applications. However, limitations such as relatively low yield expression, stability and folding dependent on accurate disulfide bond formation and tendency to form polydisperse solutions has triggered the development of alternate natural and synthetic protein binders. Recently, it was reported that proteins without the immunoglobulin fold such as the leucine-rich repeat or ankyrin repeat proteins also mediate immune responses via somatic mutation and selection (Pancer et al. 2004). Thus, it is possible that proteins other than antibodies may be as suitable for molecular display techniques to generate high affinity, specific binders (Hosse et al. 2006). Ankyrin repeat proteins or ankyrin domains are widely distributed across several species and phyla in nature, mediate important protein-protein interactions and have been isolated from various cellular environments such as intracellular, membrane bound and extracellular (Bork 1993; Kobe and Kajava 2001). Structurally, they are formed by repeated, 33 amino acid units composed of a β-turn followed by two anti-parallel α-helices and a loop that connects to the turn of the next repeat. Four to six such repeat units stack onto each other forming a continuous hydrophobic core and a large hydrophilic surface (Sedgwick and Smerdon 1999). The key to the widespread presence of these protein domains is the variability of amino acid residues that can occur throughout the proteins. Sequence analysis of several such proteins from different sources led to the determination that this variability was more clustered over the β-turn, α-helix and loop region while the remainder formed the structural framework region. This gives these modular domains the ability to easily adapt their molecular surface to bind a large variety of target proteins (Marcotte et al. 1999; Kohl et al. 2003). When compared to the CDR

loops of antibody fragments, there is a much larger protein surface that can be randomized. In addition, ankyrin repeat proteins are highly soluble, stable and monomeric and lack disulfide bonds. They can be overexpressed and produced from bacterial cytoplasms at levels in the range of 200 mg L^{-1}. These positive biophysical properties have been incorporated into designed ankyrin repeat proteins (DARPins) composed of fixed framework amino acids to maintain the repeat fold with variation in the residues that form the interacting surface. Varying numbers of this repeat containing mutations at specific sites were assembled between N- and C- terminal capping repeats to yield combinatorial libraries that could be used in RNA display systems (Binz et al. 2003). Diversity of such libraries is not restricted by transformation efficiency, as is the case with recombinant antibody fragment libraries, and can be simply regulated by adjusting the number of repeat units incorporated into the library. High affinity binders to several proteins have been isolated from these libraries (Amstutz et al. 2005; Amstutz et al. 2006) and several new protein crystal structures have resulted from co-crystallization with their cognate DARPins (Kohl et al. 2005).

Another non-immunoglobulin, *in vitro* designed class of molecules that has been recently used to co-crystallize cognate proteins for X-ray crystallography is called an 'affibody' (Nilsson et al. 1987). Out of the five homologous, three helix domains from the immunoglobulin (Fc region) binding region of Staphylococcal protein A (SPA), one domain called the Z-domain was isolated, designed and randomized to create an 'affibody' library (Nord et al. 1995). An affibody that binds to protein Z was selected from this library by phage display. Structural studies of this complex revealed that the complex binding interface had properties similar to protein-antibody binding interface (Hogbom et al. 2003).

Selection Techniques for various CCP reagents

The production of Fabs for co-crystallization has traditionally used monoclonal hybridoma technology to produce Mabs that are then treated with the protease papain and purified by anion exchange chromatography to remove the F_c region. Often, the resulting Fab crystallizes as a complex with its cognate protein. Recombinant antibody fragments for production in *E. coli* are usually made by isolating and cloning the genes from a monoclonal antibody obtained by conventional hybridoma technology (Bird et al. 1988; Huston et al. 1988). Generation of monoclonal antibodies by hybridoma technology, at a pace and scale necessary to match structural genomics initiatives, would involve a large investment of time, effort and cost. Separate immunizations into mice would be required for each antigen and the outcome (yield) would depend on the immune response of the animal to the target protein. Each antibody and the fragments generated from it would be different, necessitating sequencing and optimization for high recombinant yields and good crystallization properties. McCafferty *et al* reported that antibodies could be displayed on the surface of phage and such phage libraries could be used to select highly specific antibodies to the target antigens (McCafferty et al. 1990). Unlike the conventional hybridoma technique, antibodies can be selected from a phage display library at a much faster pace independent of the immunized animal's response. In this system, an antibody fragment library is expressed on the surface of a bacteriophage as fusions with a phage coat protein, which would be the proteins pIII or pVIII in the case of filamentous phage. The phage carries the genetic information encoding the antibody fragment genetically in-frame with its own surface protein. Thus, there is a coupling between the genetic information and the displayed proteins. Using a process termed "panning" illustrated in Fig. 2.2 A, the library of phage bearing variants of antibody fragments is incubated with a target protein immobilized on a solid surface

like a polystyrene plate or a particle like a magnetic bead. Following wash steps to remove phage that do not recognize the target, the bound phage are eluted off the immobilized protein and introduced to *E. coli* for amplification to produce fresh phage bearing a sub-library of antibody fragments. These fresh phage are used to initiate a new panning cycle and after three to five such rounds, high affinity antibody fragments to the target protein can be isolated and analyzed (Barbas 2001).

The simplicity of the technique, cyclic methodology amenable to high-throughput, robustness of the phage particle, rapid bacterial growth rate and ability to efficiently display antibody fragments as fusions of several phage coat proteins (gene III and gene VIII) has led to antibody phage display being the technique of choice for selection of antibodies to targets for use in drug discovery (Broach and Thorner 1996) (Pausch 1997), microarrays (Lueking et al. 1999) and antibody-based microchips (de Wildt et al. 2000). In spite of being such a lucrative alternative to hybridomas for the generation of antibodies for co-crystallization, only a limited number of antibody fragments selected by phage display have been co-crystallized with their antigen (Kuttner et al. 1998; Chen et al. 1999; Ay et al. 2000). Development of the antibody phage display technology to generate recombinant antibody fragments for macromolecular co-crystallization requires the optimization of a few criteria and resolution of some critical issues.

First and foremost, the design of the phage display library should be such that all the variants in the library have a uniform template framework to prevent optimization of each variant for stability, production and purification. To this effect, we have generated a non-immune phagemid recombinant antibody fragment (rFab) library with a nominal diversity of 1.16×10^7 using synthetic approaches (Kelley and Momany 2003). Variation of the Fabs in the library is a result of mutations of the third CDR loops of both the light and heavy chains of a Fab engineered

and optimized for stability, expression, and yield (Nadkarni et al. 2006). Because the library is synthetic, there is no target bias and thus it can be used to select Fabs for membrane proteins or soluble proteins alike. This is a distinct advantage over both immune-derived libraries as well as hybridoma technology.

An important issue in the application of this versatile technology for co-crystallization purposes is that of target protein presentation. To separate phage that display antibody fragments that recognize target proteins from phage that don't, the immobilization of the target protein on a solid support is necessary. Traditionally, target molecules have been immobilized onto polystyrene ELISA plates by passive adsorption. This coating technique is troublesome because bound proteins denature on its polystyrene surfaces (Padan et al. 1999). Thus, the protein target is no longer presented in a native conformation. To resolve this issue for targets such as membrane proteins, selection approaches that maintain the target protein's native structure, based on whole cell panning (Shadidi and Sioud 2001) or panning on proteins reconstituted into proteoliposomes (Mirzabekov et al. 2000), have been used. However, these have several disadvantages such as availability of only extracellular epitopes, complex antigen sources, and little exposed epitope areas. Additionally the techniques are time-consuming, not easy to automate and thus are not suitable for high throughput approaches. Several new technologies such as paramagnetic beads (McConnell et al. 1999) and Ni-NTA HisSorb™ plates (Padan et al. 1999) have been used in biopanning to maintain protein native conformation for membrane as well as other proteins. Apart from maintaining the target proteins in their native conformation, paramagnetic beads such as Dynabeads® from Invitrogen have other advantages such as improved mass-action properties, applicability to 96-well formats and automation, a greater

surface area for target binding and availability in a wide variety of chemistries to be able to bind a wide variety of protein targets.

The successful integration of the recombinant antibody fragment mediated co-crystallization approach into the structural genomics pipeline relies on the pace and specificity at which this approach can be used to generate antibody fragments and ultimately diffraction quality crystals of non-crystallizing proteins. The standard phage display methodology involves a series of complex steps such as overnight propagation (50-100 mL cultures per protein panned) and phage purification between adjacent rounds of selection that make this process tedious and time-consuming. Techniques that scale down this process to formats that can be automated so as to increase throughput are being developed (Vanhercke et al. 2005) (Walter et al. 2001). A good example is the 'URSA' or 'Ultra Rapid Selection of Antibodies from a phage display library' methodology (Hogan et al. 2005). Bypassing the overnight, large volume propagation and laborious phage purification steps, this technique recycles the phage produced by infected *E. coli* in the first few bacterial extrusions into the next selection round (Hogan et al. 2005). Using a modified version of this method, we have been able to perform four rounds of selection in two days as compared to a previous two weeks. Thus, antibody phage display selection strategies can be designed to complement the pace and throughput of structural genomics programs. When coupled with advances in the production of recombinant antibody fragments, the technology is a powerful tool for the high throughput generation of co-crystallization proteins (CCPs) that can be used to crystallize intransigent proteins.

Alternative selection systems exist to phage display. In phage display systems, the library encoding DNA has to be introduced to an *in vivo* bacterial environment. The assembly of the phage and expression of the protein library occurs inside the bacterial cell. Library size is

thus restricted by transformation efficiency and poses a disadvantage for isolating high affinity antibodies. Since the coupling of the phenotype and genotype occurs *in vivo*, negative selection pressures can affect the outcomes. If the displayed protein interferes in some fashion with the phage assembly process at any step, that protein's display might be less efficient and the production of clones that inhibit less would be enhanced. This can quickly lead to clones dominating the whole population after a few selection rounds. Displayed proteins that are toxic for the host or are prone to aggregation or proteolysis might cause slowed bacterial growth, lesser phage production and thus reduced library diversity (Ling 2003). As an alternative, *in vitro* display and selection systems have been developed based on immunoglobulin and non-immunoglobulin oriented CCPs. The most widely used *in vitro* selection is based on a coupled transcription/translation system termed ribosome display (Hanes and Plückthun 1997) or, with subtle modifications of the technique, RNA display (shown in Fig. 2.2B) (Wilson et al. 2001). In ribosome display, an mRNA library is synthesized *in vitro* from a DNA template library encoding the library proteins using T7 RNA polymerase. A protein product is then synthesized by an *in vitro* translation system using the mRNA template. By removing termination signals in the mRNA and proper introduction of loop structures, the ribosome remains associated with the mRNA after the protein is synthesized. Thus, a ternary complex of mRNA, ribosome, and display protein is allowed to complex with a target molecule forming a quaternary intermediate complex. By previously linking the target to magnetic beads or other support, the complex can be selectively removed from the library and washed to remove non-specific complexes. The captured mRNA is now reverse transcribed to DNA, which is used as a template to start the next cycle. RNA display differs in that coupling agents such as an RNA-binding protein (Sawata et al. 2004) or puromycin (Roberts 1999; Kurz et al. 2000) are used to directly link the mRNA and

displayed polypeptide. In practice, the reverse transcribed DNA-mRNA-puromycin-protein complex is the most desirable complex to be captured by the immobilized target as it has the least competition of binding by the non-protein components (Gold 2001). Ribosome display has significant benefits over phage display as library size is limited only by the maximal number of ribosomes that can be provided in a selection round and thus large libraries of 10^{13} diversity are possible (Dufner et al. 2006). Additionally, after each selection round the isolated mRNA undergoes PCR amplification leading to introduction of mutations, variation and sequence diversification at each step, thus making this system ideal for affinity maturation and molecular evolution (Hanes et al. 1998). Such *in vitro* systems are very effective for display of non-immunoglobulin CCPs such as ankyrin domains due to their ability to generate large, diverse libraries of designed proteins to mimic the evolutionary antibody diversity. The *E. coli* maltose binding protein was crystallized in a complex with its cognate ankyrin repeat protein, selected from its library by ribosomal display (Binz et al. 2004).

One interesting challenge in the use of these selection systems for cocrystallization confronts industrial users. The various selection systems, whether phage, cell, or RNA display, are heavily patented by companies like Cambridge Antibody Technology (http://www.cambridgeantibody.com, Cambridge, United Kingdom) and Dyax (http://www.dyax.com, Cambridge, Massachusetts, USA). For general research use, such as in academia, this is not a concern. But for pharmaceutical companies solving structures of therapeutic targets with high profit potential, the indirect path to a structure via a recombinant protein selected using patented technology is generally unacceptable- despite the ability to license the technology. Hybridoma technology for producing monoclonal antibodies is now well established, and the critical patents have expired. Thus, industry will tend to exploit the older

technology over the recombinant systems, except in the companies that have established licensing agreements.

Conclusions

Though the co-crystallization approach to aid crystallization of difficult proteins has been around for quite sometime, it has not been readily used due to the low-throughput methods of generating cognate co-crystallization protein reagents. During the last decade, much expertise has been gained in the fields of protein engineering and molecular display systems. Apart from the very popular antibody fragments, other protein scaffolds have been developed and show great promise in their use as CCP reagents. Selection methods for the isolation of high-affinity binders are not restricted to the use of hybridoma technology and numerous powerful molecular display techniques have been designed for the purpose. Several groups, including ours, are trying to establish the use of molecular display techniques for the high-throughput generation of co-crystallization reagents (CCPs) (Binz et al. 2004; Rothlisberger et al. 2004; Shea et al. 2005). With this level of interest, investment and effort, the recombinant co-crystallization protein technology is well on its way to become a powerful tool in a crystallographer's repertoire.

Table 2.1. Proteins crystallized with the help of a CCP (Co-crystallization Protein)

Protein and PDB ID	**CCP**	**Selection System**	**Resolution (A) and comments**	**Crystallization Success**
citrate transporter CitS	Recombinant Fab	Phage display	Fab isolated but not used for crystallization	(Rothlisberger et al. 2004)
ClC chloride channel [1OTS]	Monoclonal Fab against *S.typhimurium* ClC	Hybridoma	2.51	(Dutzler et al. 2003)
Cytochrome bc_1 complex [1EZV]	Monoclonal $18E11F_v$	Hybridoma	2.3	(Hunte et al. 2000)
Cytochrome *c* oxidase [1AR1]	Monoclonal F_v7E2	Hybridoma	2.7	(Ostermeier et al. 1997)
Gla domain of human factor IX	Monoclonal Fab10C12	Hybridoma	Successful crystallization	(Xiaoli et al. 2005)
HIV Type 1 Reverse Transcriptase [2HMI]	Monoclonal Fab28	Hybridoma	2.8	(Ding et al. 1998)
HIV-1 capsid protein [1AFV]	Monoclonal Fab25.3	Hybridoma	3.70	(Momany et al. 1996)
HIV-1 capsid protein [1E6J]	Monoclonal Fab13B5	Hybridoma	3.00	(Berthet-Colominas et al. 1999)
Human Death	Fab BDF1	Phage	2.32	(Li et al. 2006)

Receptor 5 [2H9G]		display		
Influenza hemagglutinin (HA) [1QFU]	Monoclonal FabHC45	Hybridoma	2.80	(Fleury et al. 1999)
KcsA K^+ channel [1K4C,1K4D]	Monoclonal Fab	Hybridoma	2.0 2.3	(Zhou et al. 2001)
KvaP, a voltage-dependent cation channel [2A0L]	Monoclonal F_v33H1	Hybridoma	3.9	(Lee et al. 2005)
KvAP, a voltage-dependent K^+ channel [1ORQ, 1ORS]	Monoclonal Fab6E1 Monoclonal Fab33H1	Hybridoma	3.2 1.9	(Jiang et al. 2003)
maltose binding protein (MBP) [1SVX]	Recombinant ankyrin repeat protein	Ribosome display	2.24	(Binz et al. 2004)
OspA, Lyme disease antigen [1OSP]	Monoclonal Fab 184.1	Hybridoma	1.95	(Li et al. 1997)
Protein Z [1LP1]	Recombinant protein A binding affibody Z_{SPA-1}	Phage display	2.30	(Hogbom et al. 2003)
S.aureus protein A domain D [1DEE]	Monoclonal Fab2A2	Hybridoma	2.70	(Graille et al. 2000)
Turkey egg white	Recombinant	Phage	2.0	(Ay et al. 2000)

lysozyme (TEL) [1DZB]	scF_v 1F9	display		
Vascular endothelial growth factor [2FJH, 2FJG]	B20-4 Fab G6 Fab	Phage display	3.10 2.8	(Fuh et al. 2006)

Figure 2.1. Molecules used for Co-crystallizations and their Binding Surfaces. The various proteins commonly used for co-crystallizations are shown as ribbon representations with residues that interact (within 4Å) with the binding partner shown in space filling format. (A) Fab from a complex with HIV capsid protein p24 [PDB accession, 1AFV] (Momany et al. 1996), (B) engineered ankyrin repeat from a complex with *E. coli* maltose binding protein (MBP) [1SVX] (Binz et al. 2004), (C) scFv from a complex with turkey egg white lysozyme [PDB accession, 1DZB] (Ay et al. 2000) and (D) affibody from a complex with Protein Z [PDB accession, 1LP1] (Hogbom et al. 2003). The figure was prepared using PyMol (http://www.pymol.org).

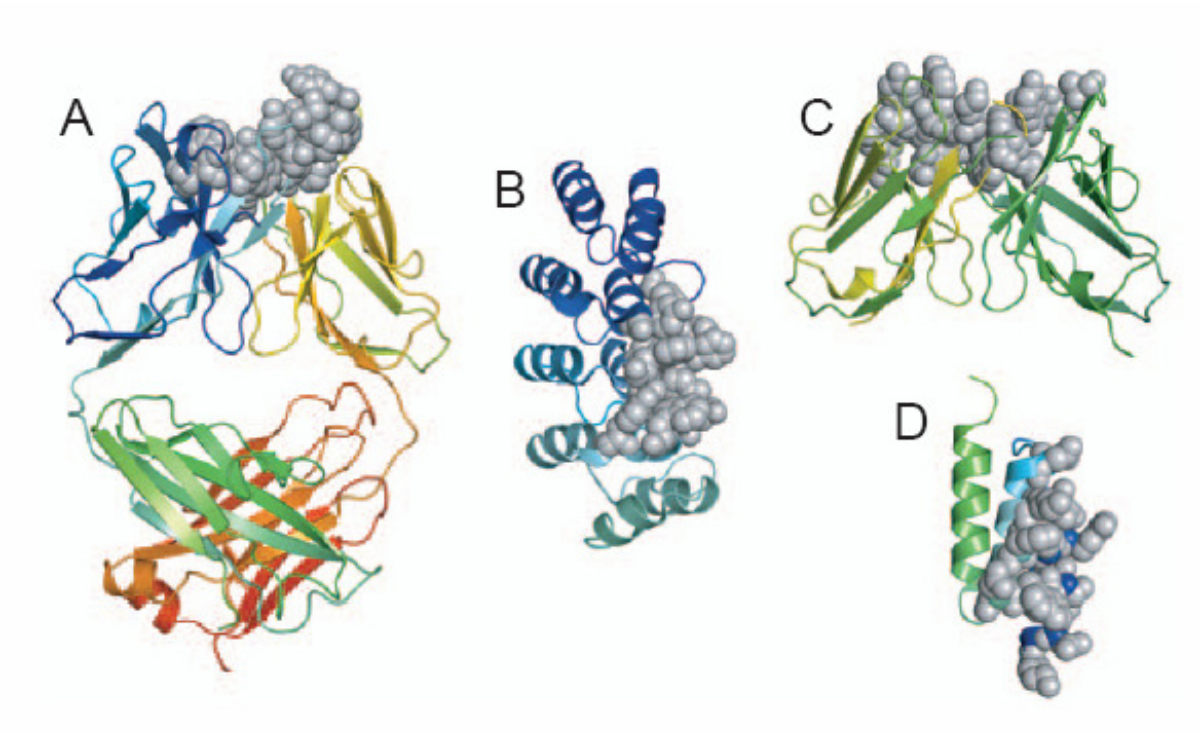
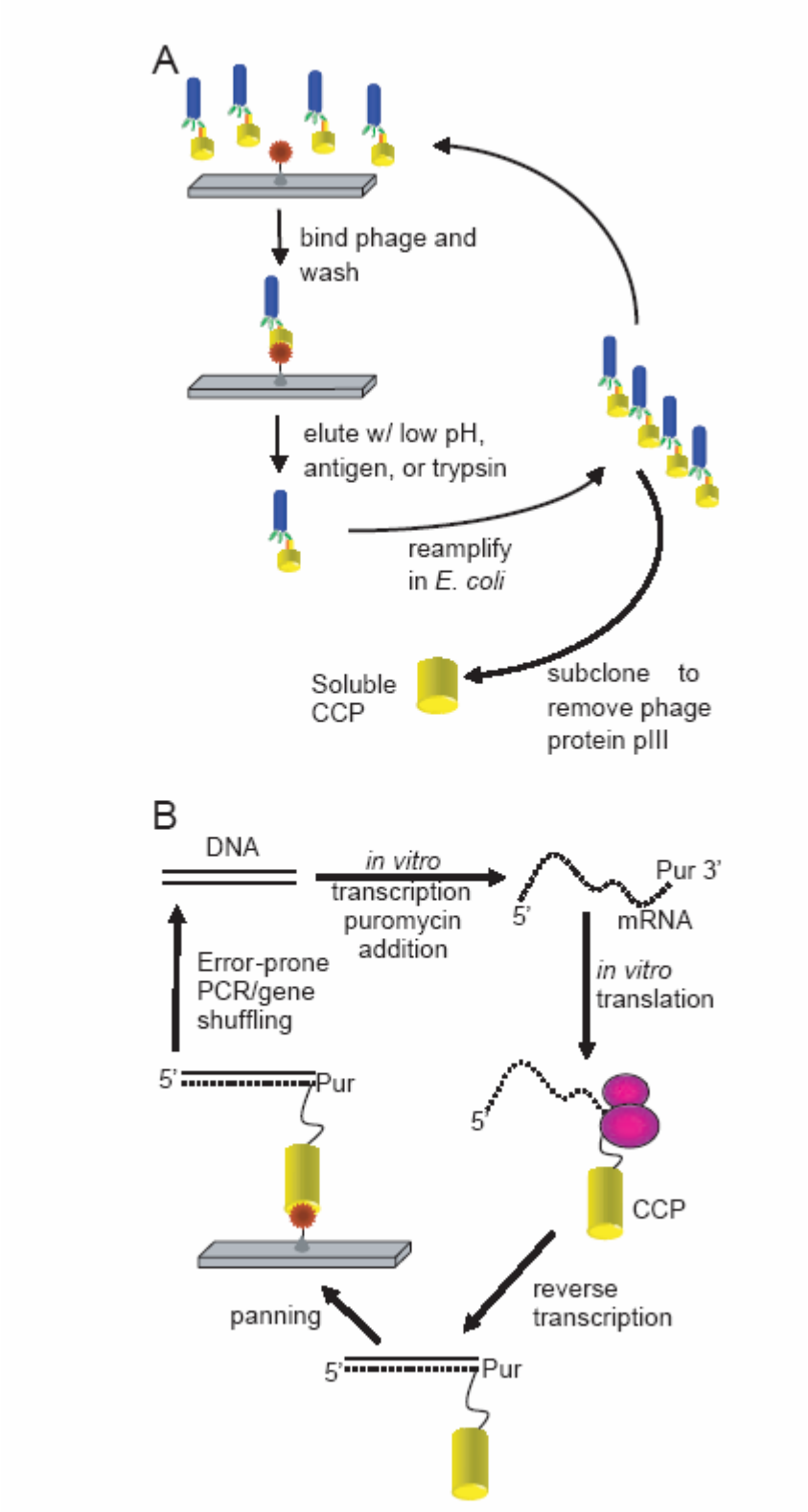

A
B
C
D

Figure 2.2. Selection systems used for obtaining cocrystallization proteins. (A) "panning" cycle using phage display technology. Phage are shown as blue tubes expressing a $_{CCP}$Fab (yellow cylinders) as a fusion with the gene III product (green). The target protein, brown star, is immobilized on a solid medium. (B) selection cycle using RNA display technology. The CCP selection protein, made on the ribosome, is shown as a yellow cylinder, while the target protein is a brown star. *Pur* represents a molecule of puromycin that is covalently attached to the 3' end of the mRNA. When the ribosome encounters the puromycin, the puromycin-mRNA is transferred to the C-terminus of the peptide chain linking the mRNA to the protein. In ribosome display, no puromycin is used. Instead, conditions are controlled to ensure that the ternary complex does not dissociate before or during panning.


A
bind phage and wash
elute w/ low pH, antigen, or trypsin
reamplify in E. coli
Soluble CCP
subclone to remove phage protein pIII
B
DNA
in vitro transcription puromycin addition
Pur 3'
5'
mRNA
Error-prone PCR/gene shuffling
in vitro translation
5'
Pur
5'
CCP
panning
reverse transcription
5'
Pur

References

Abbas, A.K., Lichtman, A.H., and Pober, J.S. 2000. *Cellular and molecular immunology*, 4th ed. W.B. Saunders, Philadelphia, pp. vii, 553.

Amstutz, P., Binz, H.K., Parizek, P., Stumpp, M.T., Kohl, A., Grütter, M.G., Forrer, P., and Plückthun, A. 2005. Intracellular kinase inhibitors selected from combinatorial libraries of designed ankyrin repeat proteins. *The Journal of biological chemistry* **280:** 24715-24722.

Amstutz, P., Koch, H., Binz, H.K., Deuber, S.A., and Plückthun, A. 2006. Rapid selection of specific MAP kinase-binders from designed ankyrin repeat protein libraries. *Protein Eng Des Sel* **19:** 219-229.

Ay, J., Keitel, T., Kuttner, G., Wessner, H., Scholz, C., Hahn, M., and Hohne, W. 2000. Crystal structure of a phage library-derived single-chain Fv fragment complexed with turkey egg-white lysozyme at 2.0 A resolution. *J Mol Biol* **301:** 239-246.

Barbas, C.F. 2001. *Phage display : a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 1 v. in various pagings.

Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F., and Cusack, S. 1999. Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *The EMBO journal* **18:** 1124-1136.

Binz, H.K., Amstutz, P., Kohl, A., Stumpp, M.T., Briand, C., Forrer, P., Grütter, M.G., and Plückthun, A. 2004. High-affinity binders selected from designed ankyrin repeat protein libraries. *Nature biotechnology* **22:** 575-582.

Binz, H.K., Stumpp, M.T., Forrer, P., Amstutz, P., and Plückthun, A. 2003. Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J Mol Biol* **332:** 489-503.

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., and Whitlow, M. 1988. Single-chain antigen-binding proteins. *Science* **242:** 423-426.

Bork, P. 1993. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins* **17:** 363-374.

Broach, J.R., and Thorner, J. 1996. High-throughput screening for drug discovery. *Nature* **384:** 14-16.

Chen, Y., Wiesmann, C., Fuh, G., Li, B., Christinger, H.W., McKay, P., de Vos, A.M., and Lowman, H.B. 1999. Selection and analysis of an optimized anti-VEGF antibody: crystal structure of an affinity-matured Fab in complex with antigen. *J Mol Biol* **293:** 865-881.

de Wildt, R.M., Mundy, C.R., Gorick, B.D., and Tomlinson, I.M. 2000. Antibody arrays for high-throughput screening of antibody-antigen interactions. *Nature biotechnology* **18:** 989-994.

Desmyter, A., Transue, T.R., Ghahroudi, M.A., Thi, M.H., Poortmans, F., Hamers, R., Muyldermans, S., and Wyns, L. 1996. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nature structural biology* **3:** 803-811.

Ding, J., Das, K., Hsiou, Y., Sarafianos, S.G., Clark, J.A.D., Jacobo-Molina, A., Tantillo, C., Hughes, S.H., and Arnold, E. 1998. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA

template-primer and an antibody fab fragment at 2.8 A resolution. *Journal of Molecular Biology* **284:** 1095-1111.

Dufner, P., Jermutus, L., and Minter, R.R. 2006. Harnessing phage and ribosome display for antibody optimisation. *Trends in biotechnology*.

Dutzler, R., Campbell, E.B., and MacKinnon, R. 2003. Gating the selectivity filter in ClC chloride channels. *Science* **300:** 108-112.

Fleury, D., Barrere, B., Bizebard, T., Daniels, R.S., Skehel, J.J., and Knossow, M. 1999. A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site. *Nat Struct Mol Biol* **6:** 530-534.

Fuh, G., Wu, P., Liang, W.-C., Ultsch, M., Lee, C.V., Moffat, B., and Wiesmann, C. 2006. Structure-Function Studies of Two Synthetic Anti-vascular Endothelial Growth Factor Fabs and Comparison with the AvastinTM Fab, pp. 6625-6631.

Gold, L. 2001. mRNA display: diversity matters during in vitro selection. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 4825-4826.

Graille, M., Stura, E.A., Corper, A.L., Sutton, B.J., Taussig, M.J., Charbonnier, J.-B., and Silverman, G.J. 2000. Crystal structure of a *Staphylococcus aureus* protein A domain complexed with the Fab fragment of a human IgM antibody: Structural basis for recognition of B-cell receptors and superantigen activity, pp. 5399-5404.

Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H.R., and Plückthun, A. 1998. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 14130-14135.

Hanes, J., and Plückthun, A. 1997. In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 4937-4942.

Harris, L.J., Larson, S.B., Skaletsky, E., and McPherson, A. 1998. Comparison of the conformations of two intact monoclonal antibodies with hinges. *Immunol Rev* **163:** 35-43.

Hogan, S., Rookey, K., and Ladner, R. 2005. URSA: ultra rapid selection of antibodies from an antibody phage display library. *Biotechniques* **38:** 536, 538.

Hogbom, M., Eklund, M., Nygren, P.-A., and Nordlund, P. 2003. Structural basis for recognition by an in vitro evolved affibody, pp. 3191-3196.

Holliger, P., Prospero, T., and Winter, G. 1993. "Diabodies": small bivalent and bispecific antibody fragments. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 6444-6448.

Hosse, R.J., Rothe, A., and Power, B.E. 2006. A new generation of protein display scaffolds for molecular recognition. *Protein Sci* **15:** 14-27.

Hunte, C., Koepke, J., Lange, C., Rossmanith, T., and Michel, H. 2000. Structure at 2.3 A resolution of the cytochrome bc(1) complex from the yeast Saccharomyces cerevisiae co-crystallized with an antibody Fv fragment. *Structure* **8:** 669-684.

Hunte, C., and Michel, H. 2002. Crystallisation of membrane proteins mediated by antibody fragments. *Curr Opin Struct Biol* **12:** 503-508.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.-S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R., et al. 1988. Protein Engineering of

Antibody Binding Sites: Recovery of Specific Activity in an Anti-Digoxin Single-Chain Fv Analogue Produced in Escherichia coli, pp. 5879-5883.

Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chait, B.T., and MacKinnon, R. 2003. X-ray structure of a voltage-dependent K+ channel. *Nature* **423:** 33-41.

Kelley, L.L., and Momany, C. 2003. Generation of a phagemid mouse recombinant antibody fragment library by multisite-directed mutagenesis. *Biotechniques* **35:** 750-752, 754, 756 passim.

Kleymann, G., Ostermeier, C., Ludwig, B., Skerra, A., and Michel, H. 1995. Engineered Fv fragments as a tool for the one-step purification of integral multisubunit membrane protein complexes. *Biotechnology (N Y)* **13:** 155-160.

Kobe, B., and Kajava, A.V. 2001. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* **11:** 725-732.

Kohl, A., Amstutz, P., Parizek, P., Binz, H.K., Briand, C., Capitani, G., Forrer, P., Plückthun, A., and Grütter, M.G. 2005. Allosteric inhibition of aminoglycoside phosphotransferase by a designed ankyrin repeat protein. *Structure (Camb)* **13:** 1131-1141.

Kohl, A., Binz, H.K., Forrer, P., Stumpp, M.T., Plückthun, A., and Grütter, M.G. 2003. Designed to be stable: crystal structure of a consensus ankyrin repeat protein. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 1700-1705.

Kovari, L.C., Momany, C., and Rossmann, M.G. 1995. The use of antibody fragments for crystallization and structure determinations. *Structure* **3:** 1291-1293.

Kurz, M., Gu, K., and Lohse, P.A. 2000. Psoralen photo-crosslinked mRNA-puromycin conjugates: a novel template for the rapid and facile preparation of mRNA-protein fusions. *Nucleic acids research* **28:** E83.

Kuttner, G., Keitel, T., Gieimann, E., Wessner, H., Scholz, C., and Hohne, W. 1998. A phage library-derived single-chain Fv fragment in complex with turkey egg-white lysozyme: characterization, crystallization and preliminary X-ray analysis. *Molecular Immunology* **35:** 189-194.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W.A. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393:** 648-659.

Laver, W.G. 1990. Crystallization of antibody-protein complexes. *Methods: A Companion to Methods of Enzymology* **1:** 70-74.

Lee, S.-Y., Lee, A., Chen, J., and MacKinnon, R. 2005. Structure of the KvAP voltage-dependent K+ channel and its dependence on the lipid membrane, pp. 15441-15446.

Lesk, A.M., and Chothia, C. 1988. Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint. *Nature* **335:** 188-190.

Li, B., Russell, S.J., Compaan, D.M., Totpal, K., Marsters, S.A., Ashkenazi, A., Cochran, A.G., Hymowitz, S.G., and Sidhu, S.S. 2006. Activation of the Proapoptotic Death Receptor DR5 by Oligomeric Peptide and Antibody Agonists. *Journal of Molecular Biology* **361:** 522-536.

Li, H., Dunn, J.J., Luft, B.J., and Lawson, C.L. 1997. Crystal structure of Lyme disease antigen outer surface protein A complexed with an Fab, pp. 3584-3589.

Ling, M.M. 2003. Large antibody display libraries for isolation of high-affinity antibodies. *Combinatorial chemistry & high throughput screening* **6:** 421-432.

Lueking, A., Horn, M., Eickhoff, H., Bussow, K., Lehrach, H., and Walter, G. 1999. Protein microarrays for gene expression and antibody screening. *Analytical biochemistry* **270:** 103-111.

Marcotte, E.M., Pellegrini, M., Yeates, T.O., and Eisenberg, D. 1999. A census of protein repeats. *J Mol Biol* **293:** 151-160.

McCafferty, J., Griffiths, A.D., Winter, G., and Chiswell, D.J. 1990. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348:** 552-554.

McConnell, S.J., Dinh, T., Le, M.H., and Spinella, D.G. 1999. Biopanning phage display libraries using magnetic beads vs. polystyrene plates. *Biotechniques* **26:** 208-210, 214.

Mirzabekov, T., Kontos, H., Farzan, M., Marasco, W., and Sodroski, J. 2000. Paramagnetic proteoliposomes containing a pure, native, and oriented seven-transmembrane segment protein, CCR5. *Nature biotechnology* **18:** 649-654.

Momany, C., Kovari, L.C., Prongay, A.J., Keller, W., Gitti, R.K., Lee, B.M., Gorbalenya, A.E., Tong, L., McClure, J., Ehrlich, L.S., et al. 1996. Crystal structure of dimeric HIV-1 capsid protein. *Nature structural biology* **3:** 763-770.

Nadkarni, A., Kelley, L.C., and Momany, C. 2006. Optimization of a Mouse Recombinant Antibody Fragment for Efficient Production from *Escherichia Coli. Protein Purif. Expr.* in press.

Nilsson, B., Moks, T., Jansson, B., Abrahmsen, L., Elmblad, A., Holmgren, E., Henrichson, C., Jones, T.A., and Uhlen, M. 1987. A synthetic IgG-binding domain based on staphylococcal protein A. *Protein engineering* **1:** 107-113.

Nord, K., Nilsson, J., Nilsson, B., Uhlen, M., and Nygren, P.A. 1995. A combinatorial library of an alpha-helical bacterial receptor domain. *Protein engineering* **8:** 601-608.

Ostermeier, C., Harrenga, A., Ermler, U., and Michel, H. 1997. Structure at 2.7 A resolution of the Paracoccus denitrificans two-subunit cytochrome c oxidase complexed with an antibody FV fragment, pp. 10547-10553.

Ostermeier, C., Iwata, S., Ludwig, B., and Michel, H. 1995. Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. *Nature structural biology* **2:** 842-846.

Padan, E., Venturi, M., Michel, H. and Hunte, C. 1998. Production and characterization of monoclonal antibodies directed against native epitopes of NhaA, the Na+/H+ antiporter of Escherichia coli. *FEBS Letters* **441**: 53-58.

Padan, E., Venturi, M., Michel, H., and Hunte, C. 1999. Efficient screening of conformation-dependent monoclonal antibodies using Ni-NTA HisSorb^{TM} Plates, pp. 17-20.

Pancer, Z., Amemiya, C.T., Ehrhardt, G.R., Ceitlin, J., Gartland, G.L., and Cooper, M.D. 2004. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* **430:** 174-180.

Pausch, M.H. 1997. G-protein-coupled receptors in Saccharomyces cerevisiae: high-throughput screening assays for drug discovery. *Trends in biotechnology* **15:** 487-494.

Roberts, R.W. 1999. Totally in vitro protein selection using mRNA-protein fusions and ribosome display. *Current opinion in chemical biology* **3:** 268-273.

Roosild, T. P., Castronovo, S. & Choe, S. 2006. Structure of anti-FLAG M2 Fab domain and its use in the stabilization of engineered membrane proteins. *Acta crystallographica* **62**:835-839.

Rothlisberger, D., Honegger, A., and Plückthun, A. 2005. Domain Interactions in the Fab Fragment: A Comparative Evaluation of the Single-chain Fv and Fab Format Engineered

with Variable Domains of Different Stability. *Journal of Molecular Biology* **347:** 773-789.

Rothlisberger, D., Pos, K.M., and Plückthun, A. 2004. An antibody library for stabilizing and crystallizing membrane proteins - selecting binders to the citrate carrier CitS. *FEBS Letters*

Structure, Dynamics and Function of Proteins in Biological Membranes **564:** 340-348.

Sawata, S.Y., Suyama, E., and Taira, K. 2004. A system based on specific protein-RNA interactions for analysis of target protein-protein interactions in vitro: successful selection of membrane-bound Bak-Bcl-xL proteins in vitro. *Protein Eng Des Sel* **17:** 501-508.

Sedgwick, S.G., and Smerdon, S.J. 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends in biochemical sciences* **24:** 311-316.

Shadidi, M., and Sioud, M. 2001. An anti-leukemic single chain Fv antibody selected from a synthetic human phage antibody library. *Biochemical and biophysical research communications* **280:** 548-552.

Shea, C., Bloedorn, L., and Sullivan, M.A. 2005. Rapid isolation of single-chain antibodies for structural genomics. *J Struct Funct Genomics* **6:** 171-175.

Vanhercke, T., Ampe, C., Tirry, L., and Denolf, P. 2005. Rescue and in situ selection and evaluation (RISE): a method for high-throughput panning of phage display libraries. *J Biomol Screen* **10:** 108-117.

Walter, G., Konthur, Z., and Lehrach, H. 2001. High-throughput screening of surface displayed gene products. *Combinatorial chemistry & high throughput screening* **4:** 193-205.

Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T., and Winter, G. 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. *Nature* **341:** 544-546.

Wilson, D.S., Keefe, A.D., and Szostak, J.W. 2001. The use of mRNA display to select high-affinity protein-binding peptides. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 3750-3755.

Wilson, I.A., and Stanfield, R.L. 1994. Antibody-antigen interactions: new structures and new conformational changes. *Curr Opin Struct Biol* **4:** 857-867.

Xiaoli, S., Yongdong, L., Chuanbing, B., Gengxiang, Z., and Mingdong, H. 2005. Crystallization of an anti-factor IX antibody and its complex. *Acta Crystallographica: Section D* **61:** 701-703.

Zhou, Y., Morais-Cabral, J.H., Kaufman, A., and MacKinnon, R. 2001. Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. *Nature* **414:** 43-48.

CHAPTER 3

OPTIMIZATION OF A MOUSE RECOMBINANT ANTIBODY FRAGMENT FOR EFFICIENT PRODUCTION FROM *ESCHERICHIA COLI*.[2]

[2] A. Nadkarni, L.C.Kelley, and C. Momany. Accepted by *Protein Expression and Purification*, 10/11/2006.

Abstract

A mutagenized mouse recombinant antibody fragment (rFab) that recognized HIV capsid protein was isolated from *E. coli* at a level of 12 mg per liter of culture using standard shake flask methods. This is one of the highest yields of a modified antibody fragment obtained using non fermentor-based methods. Recombinant Fab was isolated directly from the culture medium, which lacked complex materials such as tryptone and yeast extract. Fab isolated from the periplasm was not as homogeneous as that isolated directly from the culture medium. Optimization of the culture medium using recently developed media, the use of *E. coli* cell lines that contained rare tRNA codons, and mutagenesis of the Fab to improve the stability of the Fab were important factors in producing high levels of the Fab. An isolation protocol easily adaptable to automation using a thiophilic-sepharose column followed by metal-chelate chromatography and the introduction of a nontraditional metal binding site for metal chelate purification that bypasses the conventional hexahistidine tag cleavage step (to prevent the purification tag from interfering with crystallization) are additional features of this approach to produce a highly homogenous preparation of rFab. The resulting rFab binds to its antigen, p24, equivalent in character to the monoclonal from which the rFab was originally derived.

Keywords: antibody fragment; Fab; secretion; p24; HIV capsid

Introduction

The application of antibody fragments such as Fabs and scFvs in biotechnology has expanded in recent years due to their strong affinity for other biomolecules, their inherent stability yet potential for tremendous diversity, and their ability to be attached to powerful selection systems like phage display [1, 2]. A limiting factor in the wide use of recombinant Fabs (rFabs), whether made in bacteria like *E. coli*, fungi or higher eukaryotes in cell culture, has been the low yields of the Fab. A standard approach used in bacterial production of rFabs is to grow cells to high density, then isolate the Fab from the periplasm where it is designed to be exported from the bacterial cytoplasm by inclusion of secretory signal peptides on the Fab light and heavy chains. Export to the oxidizing environment of the periplasm is necessary for optimal folding and disulfide bond formation [3]. This is convenient in the context of a phage display system because filamentous phage such as M13 utilize a secretory system (Sec) to target the phage for extrusion through the bacterial membrane. In some Fab production schemes, the Fab is not exported, but instead is refolded after purification of the denatured Fab from overexpressed inclusion bodies in the cytoplasm [4]. The recent introduction of mutant *E. coli* strains, such as the Origami strain, lacking functional glutathione reductase and thioredoxin has been applied to the production of intracellular Fabs as well [5, 6]. In these bacterial strains, the intracellular milieu is now much more oxidizing, making disulfide formation more favorable. But this approach has not been readily adopted. In all cases, the yields are low, with on average only a few mg of rFab produced per liter of cell culture.

Our interest in efficient recombinant Fab production stems from our development of a high-throughput approach of making Fabs for cocrystallization of other biological molecules, as

well as an interest in incorporating Fabs into pharmaceutical delivery systems, such as immunomicelles. While huge investments have been made in scaling the machinery of structure determinations, solving atomic structures of macromolecules is still a difficult task because the crystallization screening process is empirically driven. Many macromolecules, especially membrane and glycosylated proteins or RNA, still remain difficult to crystallize on a routine basis. One approach to addressing the crystallization problem is to co-crystallize a target protein with another protein that specifically recognizes the target so that a stable complex is formed. The earliest application of this approach was the use of antibody fragments such as Fabs and scFvs to crystallize several difficult proteins [7-9]. Non-immune-based molecular libraries have now been created for this task [10]. Antibodies and other co-crystallization proteins (CCPs) may be effective because they can control non-specific aggregation and modify the surfaces of molecules in ways favorable for crystallization. With the goal of creating a recombinant antibody system for co-crystallization, we recently built a non-immune (synthetic) rFab phage display library with a nominal diversity of 1.16×10^7 [11]. The evaluation, characterization and optimization of the expression of the rFab mutants that lead to the development of this library are described herein.

Materials and Methods

Construction of recombinant Fabs

Complementary DNA encoding the heavy and light chain variable and constant domains of Fab25.3 were obtained by performing RT-PCR using total RNA isolated from monoclonal Mab25.3 cells (gift from Ladislau Kovari and Michael Rossmann, Purdue University) with PCR primers complementary to the IgG1 heavy chain (MscvH253-F and MscIgG1-R,Table 3.1) and

Kappa light chain (MscvK253-F and MscKappa-R) DNA sequences of Mab25.3 (Ladislau Kovari and Michael Rossmann, unpublished results) in two separate reactions following standard phage display procedures [12]. The primers were designed to include the appropriate restriction enzyme sites (Sac I/Xba I for the heavy chain, and Spe I/Xho I for the light chain) developed for insertion of mouse antibody fragments into the pCOMB-3H vector [12]. The resulting cDNAs were cloned into the pCOMB-3H plasmid by treating the heavy chain PCR product with the restriction enzymes Sac I/Xba I and the light chain cDNA with Spe I/Xho I, gel purifying the digested DNAs, and ligating the DNAs into correspondingly digested and purified pCOMB-3H in two steps, with the light chain introduced first. The resulting plasmid, pCOMB-Fab25.3-gIII contained the heavy chain gene of the Fab in frame with the phagemid gene III. A Spe I/Nhe I restriction enzyme digestion (which removed the DNA encoding the gene III phage protein), followed by gel purification of the cut DNA and ligation, produced the low-level expression plasmid pCOMB-Fab25.3. To obtain a Fab construct having a hexahistidine purification tag on the C-terminus of the heavy chain, the pCOMB-Fab25.3-gIII DNA was digested with Spe I and Not I and gel purified. The annealed oligonucleotides, PCBFab2-F and PCBFab2-R were ligated with the plasmid to make pCOMB-Fab2. To make a more efficient protein expression plasmid, the light and heavy chain sequences from pCOMB-Fab2 were transferred into pET28b (Novagen). This was performed in two steps by first digesting both pET28b and pCOMB-Fab2 with the restriction enzymes EcoR I and Not I, gel purifying the two cut DNAs, and ligating the EcoR I/Not I Fab light and heavy chain DNA with the cut pET28b. The resulting DNA was then amplified by PCR using primers Fab2-F and Fab2-R, which contained Eam 1104 restriction sites to orient the light chain appropriately behind the promoter of the pET vector. Digestion with restriction enzyme Eam 1104 concurrent with the presence of T4 ligase using the "Seamless

Cloning" kit of Stratagene produced the plasmid pET28b-Fab2. The secretory leader sequences within the pCOMB-3H vector were thus transferred in the process, as well as the internal ribosome binding site encoded upstream from the heavy chain sequence. Further modifications of the pET28b-Fab2 DNA to make pET28b-Fab3 and pET28b-Fab4 were introduced using the QuikChange™ Site Directed Mutagenesis kit from Stratagene. The respective mutation primers used to create these constructs are shown in Table 3.1. Fab3, which encoded four less histidines at the heavy chain C-terminus and two histidines added to the C-terminus of the light chain relative to Fab2, was created in two mutagenesis steps using the primers: Fab3LΔH-F and Fab3LΔH–R, and then Fab3HΔH-F and Fab3HΔH–R. Fab4 was created from primers, Fab4ΔLeu-F and Fab4ΔLeu–R, and was designed to encode a Fab heavy chain in which a leucine was deleted near the amino-terminus of the protein. Plasmids were purified using Qiagen miniprep kits and transformed by electroporation into *E. coli* XL1-blue (Stratagene) for subcloning or into BL21(DE3) or BL21(DE3)-RIL (Stratagene) strains for high-level protein expression using the pET28-based plasmids. Transformed colonies were screened for Fab production via ELISA [12] using goat and rabbit anti-mouse Fab'$_2$ antibodies conjugated to alkaline phosphatase (Pierce Biotech). Sequencing of the DNA constructs was performed by the Integrated Biotechnology laboratory at the University of Georgia.

Production of rFabs by IPTG induction and auto-induction

"Defined" medium when used with IPTG induction contained: 4 g glucose, 2 g NH_4Cl, 6 g KH_2PO_4, 13.6 g Na_2HPO_4, 40 mg each of the standard twenty amino acids, 1 mg each of the four vitamins riboflavin, pyridoxine, niacin and thiamine, 4.9 g $MgSO_4$ (0.02M) and 25 mg $FeSO_4 \times 7H_2O$ per liter with the appropriate antibiotic, usually 50 μg mL^{-1} kanamycin and 34 μg

mL^{-1} chloramphenicol. Single colonies from freshly transformed cells expressing various rFabs in BL21(DE3) and BL21(DE3) RIL (Strategene) *E. coli* were inoculated into 5 mL LB medium and grown overnight. A 20 µL inoculum of the overnight culture was transferred to 100 mL of defined medium and incubated on a shaker at 37°C. After 12-18 hrs, the culture was diluted into a liter of defined medium in a 2.8 L baffled Fernbach flask. The culture was grown at 37°C to 0.5 OD_{600nm} and then cooled to room temperature. Protein production was induced by the addition of 0.5mM IPTG and induction was continued at room temperature for six hours. After the induction period, the cells were removed from the culture medium by centrifugation at 6000 × g for 30 min at 4°C, and 1 mM PMSF (1M stock in isopropanol) was added to inhibit protease activity.

Recombinant Fab production using an "auto-induction" strategy utilized the medium and protocol introduced by Studier [13]. Briefly, *E. coli* BL21 (DE3)-RIL colonies, freshly transformed with pET28-Fab4, were used to inoculate 2 mL of auto-induction media (50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25 mM $(NH_4)_2SO_4$, 1 mM $MgSO_4$, 0.5 % glycerol, 0.5 % glucose, 0.2 % lactose, 0.2 X trace metals, 18 amino acids at 200 mg L^{-1} each, four vitamins as before at 1 mgL^{-1} each, 50 mgL^{-1} kanamycin and 34 mgL^{-1} chloramphenicol). This medium differed from Studier's (2005) in that it contained the four vitamins as used in the glucose defined medium. After incubation on a shaker (300 rpm) for 4-5 hours (to 1.0 OD_{600nm}) at 37°C,, 0.5 mL of this culture was diluted into a 2.8 L baffled Fernbach flask containing 0.5 L of auto-induction media and incubated with shaking for 2-3 hours at 37°C before transferring it to room temperature. From this 1000-fold dilution, saturation (OD_{600nm} ~ 5-8) was usually reached in 12-14 hours. After 16-18 hours, the culture was centrifuged at 6000 × g for 30 min at 4°C to separate the cells from the culture medium, and PMSF (1 mM final) was added to the resulting supernatant.

Comparative expression of different rFab mutants in two E. coli strains

Comparison of the growth characteristics of three pET28b-based rFab mutants in either *E. coli* BL21 (DE3) or BL21 (DE3)-RIL bacterial cell lines was undertaken to evaluate the need of rare codons for optimal expression of the Fabs. Triplicate 100 mL cultures of different Fabs in the two cell lines were prepared in matched baffled 500 mL flasks. Protein induction followed the procedure outlined for IPTG induction of a glucose supplemented defined medium culture. Aliquots were taken at regular time intervals to monitor cell density (OD_{600nm}) and evaluate Fab production by ELISA. To relate the ELISA signal to protein concentration, purified rFab standards of known concentration were included on each ELISA plate.

Comparison of different induction methods and media using pET28-Fab4 transformed cells

Parallel 100 mL cultures of *E. coli* BL21(DE3) RIL cells freshly transformed with pET28-Fab4 were grown in different growth media in 500 mL baffled flasks at room temperature from the same inoculum. Condition 1 was defined medium without glucose but containing 0.2% w/v glycerol, with lactose added at late log phase to a final concentration of 2 g L^{-1}. Condition 2 was defined medium + 0.4% w/v glucose induced at mid log phase with 0.5 mM IPTG. Condition 3 was defined medium lacking glucose but containing 0.2% glycerol and no induction. Condition 4 was defined medium + 0.4% w/v glucose and no induction. Aliquots were taken at regular time points to monitor cell density (OD_{600nm}) and quantitate the rFab by ELISA.

In a separate set of experiments, rFab production in auto-induction media was evaluated by growing *E. coli* BL21 (DE3)-RIL cells freshly transformed with pET28-Fab4 in Studier auto-induction media at 37°C for 3 hours, then overnight at room temperature. Parallel control

experiments included growth on defined medium + 0.2% w/v glycerol with late log phase lactose added as described previously, and defined medium + 0.4% w/v glucose with mid log phase 0.5 mM IPTG induction. Aliquots were taken at regular time intervals and assayed by ELISA.

Purification of rFabs by metal-chelate chromatography and by a piggy-backed anion exchange (Q)-metal chelate chromatography.

Recombinant Fab protein was recovered either from the *E. coli* periplasm or directly from the medium in which the culture was grown. To extract the rFab from the periplasm, harvested cells were incubated in periplasm extraction buffer (0.2 mg mL^{-1} lysozyme, 1mM EDTA, 0.5 μM PMSF, 1.0 M NaCl, 50 mM Tris HCl, pH 8.0) per gram of wet cell weight over ice for 30 min. Protoplasts were removed by centrifugation for 30 min at 6000 × g at 4°C. Fab was precipitated out of the culture and periplasm supernatant by adding solid ammonium sulfate to achieve 80% saturation at 4°C. After centrifugation for 30 min at 6000 × g at 4°C, the precipitated Fabs were resuspended and dialyzed against the HPLC start buffer 20 mM NaH_2PO_4, pH 7.4.

Chromatographic separations were performed on a Pharmacia ÄKTA purifier system equipped with an extra multichannel buffer valve. The dialyzed rFabs were loaded onto a 5 mL His-Trap metal chelate affinity column (GE Biosciences) that was charged with nickel and equilibrated with buffer A (20 mM NaH_2PO_4, pH 7.4). The protein was eluted using a 20 column volume gradient with 0.5 M imidazole, 20 mM $NaH_2PO_4 \cdot H_2O$, pH 7.4. A more efficient strategy was to piggy-back a HiTrap Q column (GE Biosciences) onto the top of the metal chelate column. The coupled columns were equilibrated with 0.1 M NaCl, 20 mM Tris, pH 7.4 and loaded with samples dialyzed into the same buffer. Two sequential gradients were performed, the first being a 20 column volume gradient with 0.5 M imidazole, 20 mM Tris, pH

7.4, followed by a 10 column volume one to 1.0 M NaCl, 0.5 M imidazole, 0.02 M Tris, pH 7.4. The Fab eluted early in the first gradient, while contaminants were removed in the latter gradient.

Automated rFab purification using thiosepharose and metal-chelate chromatography

Solid Na_2SO4 was added to rFab culture supernatants containing PMSF (1 mM) to achieve a final 0.5 M concentration. The mixture was centrifuged for 30 min at 6000 × g at 4°C. The clarified culture supernatant (> 500 mL depending on the culture volume) was then directly pumped onto a 50 mL thiosepharose column (XK 26/20 length 20 cm and internal diameter 26 mm; GE Biosciences) at 5 mL min^{-1} using an ÄKTA purifier. The thiosepharose used to pack the column was prepared in-house. [14]. After loading the sample, the column was washed with 0.5 M Na_2SO_4 until the 280 nm absorbance reached baseline, then the rFab was eluted in one step with 50 mM NaH_2PO_4, pH 7.0. The entire protein peak that eluted when the conductivity dropped was directed immediately to a Superloop™ and then loaded onto a 5 mL His-Trap column (GE Biosciences) charged with nickel and equilibrated with 50mM NaH_2PO_4, pH 7.0. After the absorbance returned to baseline, the rFab was eluted using a 20 column volume gradient of 50mM NaH_2PO_4, 0.5M imidazole, pH 7.0.

Characterization of rFabs.

The purity of the rFabs was evaluated by SDS-PAGE analysis [15] using 12% cross-linked gels stained with commassie brilliant blue using a Hoeffer Mighty Small gel apparatus or by native gel electrophoresis using 7.5% gels, run and silver stained with a PhastSystem™ (GE Biosciences). Fab-p24 complexes were prepared by incubating rFab4 (2 µl of 1 mg mL^{-1}) for 1hr at 37 °C with HIV capsid p24 (1.4 – 10 µl of 0.6 mg mL^{-1}) at varying molar ratios in 20mM Tris

Cl, pH 8.0. Protein was quantitated by dye binding (BioRad) with BSA as a standard or using an absorbance coefficient of $A_{280nm}^{0.1\%}$= 0.553. ELISAs were performed on Immulon II™ 96-well EIA/RIA plates (Dynex Technologies, Inc.), coated with 0.1 mL of 20 μg mL^{-1} recombinant HIV capsid protein p24. Following the wash and blocking steps, 0.1 mL samples and/or their dilutions were applied to the wells as a primary antibody. Either a goat or rabbit anti-mouse IgG $F(ab')_2$-alkaline phosphatase conjugate (Pierce Biotech) was used as the secondary antibody. Recombinant HIV capsid protein having a polyhistidine purification tag used as the antigen in ELISAs was expressed in *E. coli* and purified using metal-chelate chromatography using a plasmid generously provided by Carol Carter, Department of Molecular Genetics and Microbiology, Stony Brook University Stony Brook, NY.

Results

Construction of rFab expression vectors

The cDNA of the light and heavy chains (variable and constant domains) from the monoclonal antibody Mab25.3 that recognizes the HIV capsid (p24) protein was reverse transcribed from the Mab25.3 RNA and cloned into the Phage Display vector pCOMB-3H using standard molecular techniques [12]. The constructs in the pCOMB-3H vector all expressed minimal levels of Fab (rFab levels of μg L^{-1} of Super Broth culture media) and attempts to optimize the production only minimally improved the yield (data not shown). To alleviate the poor expression in the pCOMB-3H vector, the coding sequence for the Fab light and heavy chains was transferred to the high-expression plasmid pET28b. Initial expression studies were performed in Super Broth (MOPS/yeast extract/tryptone), which is the favored media for phage display work and Fab production [12], but it became clear that significant levels of rFab were

present in the culture medium. Thus, our studies shifted to a defined medium with which the rFab could be more readily purified from the culture supernatant. Media containing tryptone and yeast extract did not perform well in chromatographic separations- especially when the media were precipitated with ammonium sulfate and pumped directly through the chromatography columns. Serendipitously, the defined media produced more rFab as well.

A number of stepwise mutants of the pET28b-Fab25.3 were made and evaluated (Figure 3.1). In the Fab2 construct, a 6-His purification tag was added to the C terminus of the heavy chain, to provide an easy purification tag to the expressed Fab. In Fab 3, the 6-His tag was trimmed down to two residues at the C-terminus of the heavy chain and two histidines were appended to the light chain's C-terminus. Both the chains now had two vicinal histidine residues. Based on the original Fab structure (PDB accession 1AFV) [9], these histidines were anticipated to have close-proximity to one-another and thus create a metal binding site that could interact with metal chelate columns. Presumably this would bind to metal chelate columns reasonably well and also have a well-defined heavy atom site (Pt for instance) that could be used in crystallographic phasing. Further, a hexahistidine tail would be bulky and disordered, whereas the cluster of histidines at the C-terminus would be structurally less flexible. For the construction of Fab4, a superfluous leucine in the heavy chain was deleted from the encoding DNA sequence. The pCOMB-3H vector contains the DNA sequence 5'-CCGAGGTGCAGCTGCTCGAG, which when translated encodes the amino acid sequence AEVQLLE and has a Xho I restriction enzyme cloning site (underlined). Preceding this is the bacterial secretory sequence with cleavage occurring between the AE sequence. Mouse IgG1 antibody sequences do not contain two leucine residues in a row at this position. The effect of the additional leucine is that an extra hydrophobic residue is introduced in the middle of a

conserved framework β-strand. Thus, we anticipated that this error in the sequence would likely lead to the destabilization of the structure and increased susceptibility to proteolytic degradation. Removal of the extra amino acid was anticipated to improve the yield by stabilizing the protein structure and rendering it less sensitive to proteolytic attack. The levels of Fab4 are at least 2 times that of Fab3 under all conditions (Figure 3.2). Further, the Fab4 can be purified readily, while Fab3 has multiple species (data not shown). Based on this data, Fab4 was ultimately the mutant chosen as the basis for our phage display library [11].

Optimization of rFab expression

Our goal of producing rFabs for cocrystallization of difficult proteins required high yields of homogenous rFabs. Due to the low levels of rFabs produced using standard phage display technologies, optimization of the rFab production scheme was necessary. Attempts to optimize the production of rFabs included evaluation of various culture media and bacterial strains, as well as improvements in the approaches used to purify the rFabs from both the periplasm and the supernatant at various growth temperatures. Expression at room temperature was significantly better than 37°C (data not shown), so all experiments were performed at room temperature.

To evaluate the importance of codon optimization in this situation where a mouse antibody fragment was expressed in *E. coli*, a three mutant comparative growth and expression study in two different cell lines, *E. coli* BL21(DE3) (henceforth DE3) or *E. coli* BL21 (DE3)-RIL (henceforth called RIL), was conducted (Figure 3.2) using a defined medium containing glucose with IPTG induction. The higher OD_{600nm} values (Figure 3.2, panel A) achieved by the RIL cells showed that the addition of the rare codons conferred a small growth advantage over the DE3 cultures for 6 hours after induction. All mutants showed a peak in Fab production as

quantitated by ELISA signal (Figure 3.2, panel B) at about 4 hours post IPTG induction, followed by a sharp falloff in the signal, possibly due to cell lysis associated with production of the Fab. The Fab4 construct achieved the highest cell density and protein production, making it the candidate of choice for further optimization. Falloff of Fab levels occurred in Super Broth as well (data not shown). A 2-4 hour window existed for harvesting the Fabs from glucose supplemented defined medium with IPTG induction. Overnight IPTG inductions resulted in poor yields of Fab.

To study the expression of rFabs with lactose as an inducer, *E. coli* BL21(DE3) RIL, freshly transformed with pET28-Fab4, were grown in defined medium with 0.2% w/v glycerol and induced in the late log phase with 2 g L^{-1} lactose. Glycerol was used instead of glucose as the carbon source because glucose represses the transport of lactose into the cell by inducer exclusion. The cells thus encounter an energy deficit and poor cell growth rate [16-19]. Additionally the defined medium containing glucose as the carbon source and IPTG as the inducer, and also glucose defined medium without IPTG and glycerol defined medium without lactose were evaluated at the same time. Cultures grown on glucose and induced with IPTG decreased in cell density due to a lower growth rate about three hrs after induction (Figure 3.3, panel A). However cultures grown first in glycerol then induced with lactose showed uniform cell growth. The protein production as evaluated by ELISA signal (Figure 3.3, panel B) showed a sharp rise after glucose/IPTG induction that fell sharply with the decreased cell growth as before, while the lactose induced cells showed uniform protein production over the time frame examined. Achievable cell densities in absence of induction, however, were much lower with glycerol/lactose than where glucose was used as the carbon source.

A buffered medium developed by Studier [13] contains a mixture of carbon sources including lactose that results in "auto-induction" of protein expression without the need to add IPTG. The bacterial cells first utilize glucose to achieve high cell densities. Induction by the lactose present in the medium is initially repressed. Once the glucose is depleted, lactose can induce protein expression. Glycerol in the medium provides a transition carbon/energy source that allows lactose uptake by the cells and gives enhanced cell densities- even after induction. *E. coli* BL21(DE3)-RIL cells freshly transformed with pET28-Fab4 were grown in Studier auto-induction media at 37°C for 3 hours and then transferred to room temperature for overnight induction. The cell densities and protein production were monitored at different time intervals and compared to 2 gL^{-1} lactose (carbon source glycerol) and 0.5 mM IPTG (carbon source glucose) induced Fab4 cultures. The glucose-IPTG culture had a sharp cell density drop after IPTG induction (maximum OD_{600nm} = 1.014) (Figure 3.4, panel A). The glycerol-lactose and auto-induction cultures both had uniform cell growth, but the auto-induction media achieved a much higher final cell density (maximum OD_{600nm} 8.72 versus 4.66). Protein production paralleled the cell growth (Figure 3.4, panel B). Thus, Studier auto-induction media provided a reliable method of growing cultures uninduced to high densities and then inducing them to produce much higher antibody concentrations. A further advantage was that there was no need to monitor the cell density or add IPTG. Chromatographic analyses of Fab produced using auto-induction media and defined media with glucose/IPTG induction (Figure 3.5) showed that auto-induction media produced three times higher protein (12 mg L^{-1} final yield) as compared to defined media with glucose/IPTG induction (4 mg L^{-1}) with similar purity. As long as the cells are growing and thus are viable, the Fab production is stable and there appears to be little

degradation of the Fabs. If the cells begin to lyse, as represented by a decrease in optical density, the resulting Fab levels are significantly lower, presumably due to proteolysis.

Optimization of rFab purification

Purification from early rFab constructs involved a single immobilized metal (zinc or nickel) affinity chromatography step. When a hexahistidine tail was used (Fab2), the Fab bound efficiently to metal-chelate columns. With the expectation that this large purification tag could ultimately interfere with crystallization of Fabs, the hexahistidine tag was reduced to two histidines on each chain that would be spatially near one another in the tertiary structure and thus were expected to bind well to a metal chelate column. Instead, they bound weakly and coeluted with *E. coli* contaminants that bind weakly to metal chelate columns and are eluted early in the imidazole gradients. However Fab isolated from the culture supernatant showed fewer interfering bacterial proteins and higher protein yields than the periplasm (data not shown). A procedure whereby a HiTrap Q column was piggybacked on to the top of the metal chelate column was developed to remove the major *E. coli* contaminants that copurified with the rFab. At high pH and relatively low salt, the rFab migrated through the Q column and bound to the nickel affinity column, while the problematic contaminants bound to the Q column. The rFab was then eluted from the metal chelate column with a buffer containing 0.5 M imidazole and no salt at all. There still remained trace contaminants that, given enough time at 4 °C, digested the rFab (no ELISA signal from samples stored at 4°C after a few months, data not shown). This method was time consuming and required dialysis steps after ammonium sulfate precipitation of the culture medium containing the rFab.

Thiophilic adsorption chromatography permits the selective purification of immunoglobulins from human serum and hybridoma culture supernatants. By using thiophilic media in conjunction with the metal-chelate columns, a method was developed whereby in an automated fashion, rFab was isolated to a level of 12 mg L^{-1} (Figure 3.6). In this method, a liter of the clarified culture supernatant was mixed with Na_2SO4 to a final concentration of 0.5M. Bypassing the overnight precipitation and dialysis steps used with the metal chelate and anion exchange chromatography, the culture supernatant was directly pumped onto the thiosepharose column. The output of the thiophilic column was channeled automatically onto a Ni affinity column to give 99% homogenous protein when analyzed by SDS-PAGE (Figure 3.7). Batch isolation on thiosepharose, followed by metal chelate purification is effective as well (data not shown).

Antigen recognition by rFab4

Each construct was verified for recognition of its cognate antigen, p24 by ELISA assay. All of the constructs retained their antigen recognition by this criterion. Since the binding properties of the original Fab produced by papain digestion of monoclonal Mab25.3 were characterized only by native gel electrophoresis of the p24-Fab complex [7], a similar approach was used to evaluate the antigen recognition of rFab4 (Figure 3.8). Interestingly, the molar ratios that gave optimal complex formation, as judged by the formation of a new species and disappearance of the uncomplexed rFab and p24 species, was 2:1, p24 to Fab. This result is identical to what was found for the monoclonal derived Fab25.3 when complexed to p24 prepared by *in situ* cleavage from the Gag polyprotein expressed in *E. coli* [7]. Interestingly, the ratio of p24 to Fab in crystals of the complex is 1:1 [9].

Discussion

Producing large amounts of Fabs, required for both therapeutic applications as well as basic research, is still a challenge. Our own focus, the cocrystallization of proteins with Fabs, depends on a very homogenous Fab preparation. Recombinant antibody fragments give a much more homogenous product compared to Fab fragments derived from papain digestion of intact monoclonal antibodies [5] and are substantially more efficient to make.

We have created a stable rFab mutant that can be produced and purified in a high throughput fashion. Our best rFab fragment (Fab4) accumulated to about 12 mg with *E. coli* cultures expressing rare codons when grown in 1 L auto-induction media using shaker flasks at room temperature. This rFab is isolated in such large amounts directly from the media- in contrast to most rFab production schemes where the Fab is isolated from the periplasm. The mechanism by which antibody fragments reach the growth medium is not yet clear, but is most commonly attributed to either leakage of the protein from the periplasm to the growth medium or cell lysis during fermentation [20-22]. Lysis is evident in our cultures grown in glucose medium with IPTG induction. But the cells appear to grow well in the Studier's auto-induction medium [13]. Thus, the Fab itself is not truly toxic to cells. Instead, the mechanism used to produce the Fab (such as IPTG induction) is the problem. It seems likely that some aspect of the secretory system must be established in a slow and steady fashion, as occurs with the auto-induction media, to prevent cell death during overproduction of the secreted Fab. Further studies are also needed to explain the complex relationship between how individual residues affect cell lysis and periplasmic leakage. We also observed that the Fab isolated from the medium is much more homogenous than that isolated from the periplasm, thus providing the rationale for our decision

to purify from the medium. In reports of high level Fab production in the cytoplasm, the authors used molecular evolution strategies to select stable and well expressed molecules that did not require functional disulphide bridges or were expressed in the reduced state in the cytoplasm [6, 23]. Another strategy using cytoplasmic expression, required the coexpression of chaperones or foldases with the Fab to give yields of 0.6-0.8 mg L^{-1} [24]. Our approach of purifying the protein from the medium, not only gave one of the highest production rates reported for antibody fragments in bacterial shake flask cultures (12 mg L^{-1}), but also avoided Fab proteolytic degradation and reoxidation problems associated with intracellular expression. One issue that may occur widely with respect to the use of the pCOMB-3H vector (and other phage display vectors using similar design features) when used with PCR products encoding mouse antibodies is the inclusion of a spurious leucine residue encoded in the vector. When this leucine was removed, the yield of purifiable protein doubled and the material was more homogenous. The phage display library constructed by us does not suffer from this problem [11] as the correction was made to the Fab4 construct used for the library creation.

The most common promoter used to direct foreign gene expression is the *lac* promoter, which is induced by lactose analogues such as IPTG. However the high IPTG concentrations (0.5-1.0 mM) that are frequently used to fully induce the *lac* promoter do not lead to maximal expression of some target proteins, such as our Fab. Also the cost of IPTG must be considered when it is used in large scale production of Fabs. Lactose, the natural inducer of the *lac* operon has been shown to be more effective and less expensive than IPTG for the induction of the *lac* operon [16, 25-27]. However our data shows that bacterial cells do not grow as vigorously on glycerol/lactose as they do on glucose, the preferred carbon source for the cells. A wide range of proteins, including membrane proteins, have been successfully produced by auto-induction [13].

Auto-induction media also adds convenience because auto-induced cultures are simply inoculated and grown to saturation without the need to monitor cell growth or add inducer at a certain time. This production scheme is thus very useful for high-throughput production of Fabs selected from the phage display library for co-crystallization purposes.

Unlike other Fabs that have hexahistidine purification tags, which might need to be cleaved off for crystallization purposes, the light chain and heavy chains of our Fab each have two vicinal histidine residues that create a metal binding site to interact with metal chelate columns. However because of the weakened binding to the column and coelution of bacterial contaminants, the lack of homogenous material from a metal chelate column alone forced us to evaluate other purification media. Antibodies have a peculiar affinity toward thioether substituted organic sulfone compounds, a phenomenon called thiophilic interaction [14]. In the presence of high levels of salts such as ammonium sulfate or sodium sulfate, the sulphur (in the form of the thioether) and the adjacent sulphone group act cooperatively to adsorb the immunoglobulins [28, 29]. This chromatographic method has also been used to separate the Bence-Jones dimer from the correctly associated rFab heterodimer [30]. By using two purification media together back-to-back, first the in-house prepared thiosepharose, followed by the Ni-charged metal chelate column, an automated protocol that generates highly homogenous Fab is possible. A further benefit of the weakened affinity of our Fab for metal-chelate ligands is that we can immobilize target molecules on metal chelating polymers (such as Dynabeads®-TALON™) and still pan with the Fab4 phage display library in the presence of low imidazole levels. A Fab having a full polyhistidine tag would bind strongly to a chelate-based immobilization medium, rendering it useless when used in a phage display context.

For a Fab' fragment expressed in the periplasmic space of E.*coli*, fermentors have been shown to improve titers to 1-2 g L^{-1} [31]. Future experiments may include the use of fermentors to evaluate whether the large yields of Fab are realizable using this purification scheme. Finally, since our protein production and purification method is easily applicable to automation, it is not only suitable for the high-throughput production of Fabs, selected from the phage display library for co-crystallization purposes, but also may well be at or near the stringency of purity necessary for use in the pharmaceutical industry.

Table 3.1. Oligonucleotides used for PCR amplification and mutagenesis

Plasmid created/ Primer Name	Sequence[3]
pCOMB-Fab25.3	
MscvK253-F	GCGGCCAGTTCC*GAGCTC*GTCTGACACAGTCTCCA *Sac I*
MscKappa-R	GAGGAGGAGGAGGAG*TCTAGA*TTAACACTCATTCCTGTTGAAGCTCTTGAC *Xba I*
MscvH253-F	GGTGGTTCCTCTAGATCTTCCCAGGTCCAA*CTCGAG*CAGCCTGGGTCT *Xho I*
MscIgG1-R	CCTGGCCGGCCTGGCC*ACTAGT*GACACCACAATCCCTGGGCACAAT *Spe I*
pCOMB-Fab2	
PCBFab2-F	*CTAGT*CACCACCATCATCATCATTAAGC *Spe I*
PCBFab2-R	*GGCCGC*TTAATGATGATGATGGTGGTG *Not I*
pET28-Fab2	
Fab2-F[4]	GG*CTCTTC*A**ATG**AAAAAGACAGCTATCG *Eam 1104*
Fab2-R[2]	GC*CTCTTC*C**CAT**GGTATATCTCCTTCTTAAAG *Eam 1104*
pET28-Fab3	
Fab3LΔH-F	CAAGAGCTTCAACAGGAATCACCACTAA*TCTAGA*TAATTAATTAGG *Xba I*
Fab3LΔH -R[5]	CCTAATTAATTA*TCTAGA*TTAGTGGTGATTCCTGTTGAAGCTCTTG *Xba I*
Fab3HΔH -F	GGACAAGAAAATTGTGCCCCATCAT*ACTAGT*TAAGCGGCCGCACTCGAG *Spe I*
Fab3HΔH -R	CTCGAGTGCGGCCGCTTA*ACTAGT*ATGATGGGGCACAATTTTCTTGTCC *Spe I*
pET28-Fab4	
FAB4ΔLeu-F	CATGGCCGAGGTGCAG*CTCGAG*CAGCCTGGGTC *Xho I*
FAB4ΔLeu-R	GACCCAGGCTG*CTCGAG*CTGCACCTCGGCCATG *Xho I*

[3] Oligonucleotide sequences are written 5' to 3'. Restriction sites (or partial sites) within each oligonucleotide are italicized and underlined with the corresponding enzyme written below.
[4] The start codon of the Fab light chain is bold for the pET28-Fab2 oligonucleotides.
[5] Although the oligonucleotide used in the experiment is shown, sequencing revealed that a mutation (C to A) was introduced (double underlined position), which resulted in a change in the amino acid sequence (Arg to Met as shown in Fig. 1).

Figure 3.1. Plasmid design and sequence changes used to produce various recombinant Fabs.

The basic design features of the pCOMB-3H and pET28b plasmids with the Fab genes are represented in this sketch. The cDNAs of the light and heavy chains (variable and constant domains) from the monoclonal antibody Mab25.3 were initially cloned into the Phage Display vector pCOMB-3H. To improve rFab production, the rFab coding region was transferred to a pET28b expression plasmid (Novagen). A number of stepwise mutants of the pET28b-Fab25.3 were made as shown in the diagram with the sequence changes that occurred between each construct highlighted in bold font. The original Mab25.3 protein sequence is followed by the rFab25.3 through rFab4 sequences. A spurious leucine present in the pCOMB-3H vector is underlined. Abbreviations used are: ORI, *E. coli* origin of replication; Amp^r, ampicillin resistance gene; Kan^r, kanamycin resistance gene; lacZ, lac Z promoter; RIBS, ribosome binding site; OmpA, OmpA secretory leader sequence; PelB, PelB secretory leader sequence; fabL, Fab light chain; fabH, Fab heavy chain; T7, T7 promoter; *lacI*, lactose repressor gene.

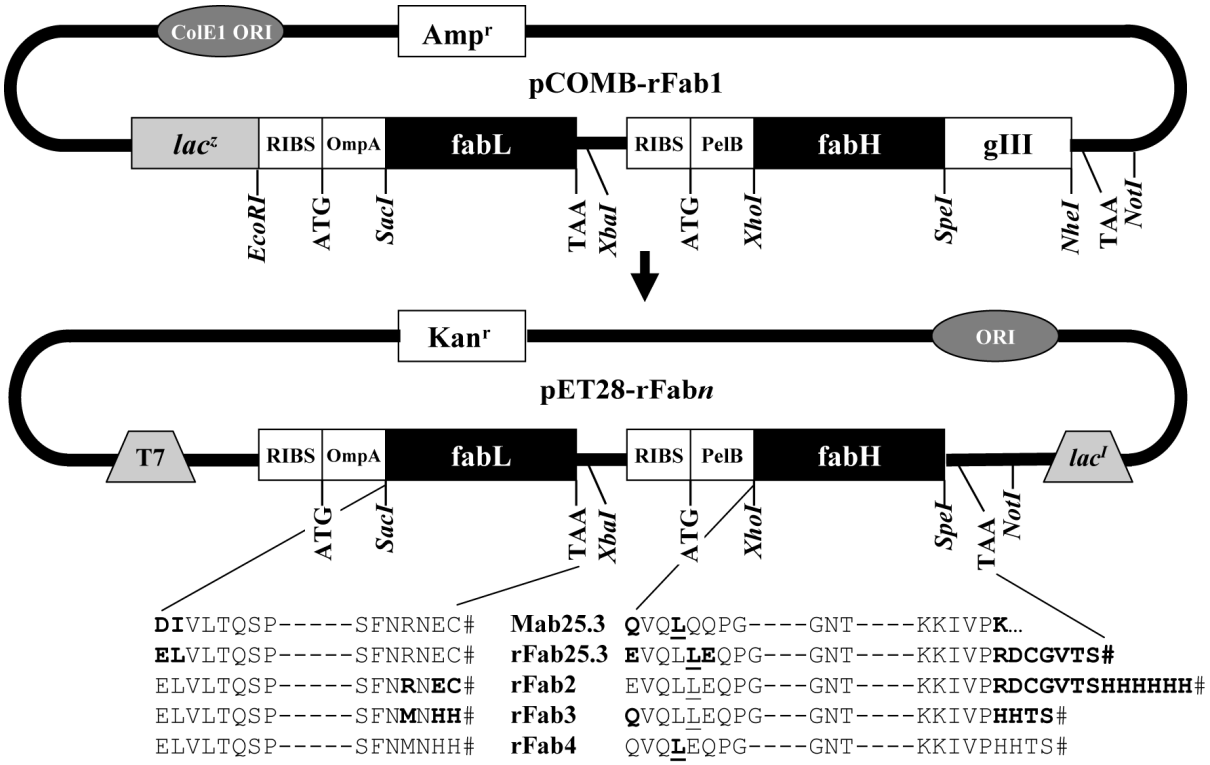

ColE1 ORI
Amp^r
pCOMB-rFab1
lac^z
RIBS
OmpA
fabL
RIBS
PelB
fabH
gIII
EcoRI
ATG
SacI
TAA
XbaI
ATG
XhoI
SpeI
NheI
TAA
NotI
Kan^r
ORI
pET28-rFabn
T7
RIBS
OmpA
fabL
RIBS
PelB
fabH
lac^I
ATG
SacI
TAA
XbaI
ATG
XhoI
SpeI
TAA
NotI
DIVLTQSP-----SFNRNEC#
ELVLTQSP-----SFNRNEC#
ELVLTQSP-----SFNRNEC#
ELVLTQSP-----SFNMNHH#
ELVLTQSP-----SFNMNHH#
Mab25.3
rFab25.3
rFab2
rFab3
rFab4
QVQLQQPG----GNT----KKIVPK...
EVQLLEQPG---GNT----KKIVPRDCGVTS#
EVQLLEQPG---GNT----KKIVPRDCGVTSHHHHHH#
QVQLLEQPG---GNT----KKIVPHHTS#
QVQLEQPG----GNT----KKIVPHHTS#

Figure 3.2. Bacterial growth kinetics and production of rFab2 through rFab4 in defined medium using two different bacterial cell lines.

Panel A shows the growth curve of cells expressing rFabs 2-4. 100 mL of defined medium were inoculated with colonies of each rFab plasmid transformed into *E. coli* BL21 (DE3) (called as Fab*n* DE3) or *E. coli* BL21 (DE3) RIL (called as Fab*n* RIL). Cells were allowed to grow to OD_{600nm} = 0.5 and then induced with 0.5 mM IPTG. Aliquots drawn at regular time intervals, pre and post induction were used to monitor and plot a growth curve as represented by optical density at 600 nm versus time of incubation. In panel B, the supernatants of the same cell aliquots were evaluated by ELISA to quantitate the Fab production. All the culture analyses were done in triplicate and the data points were averaged and plotted against time.

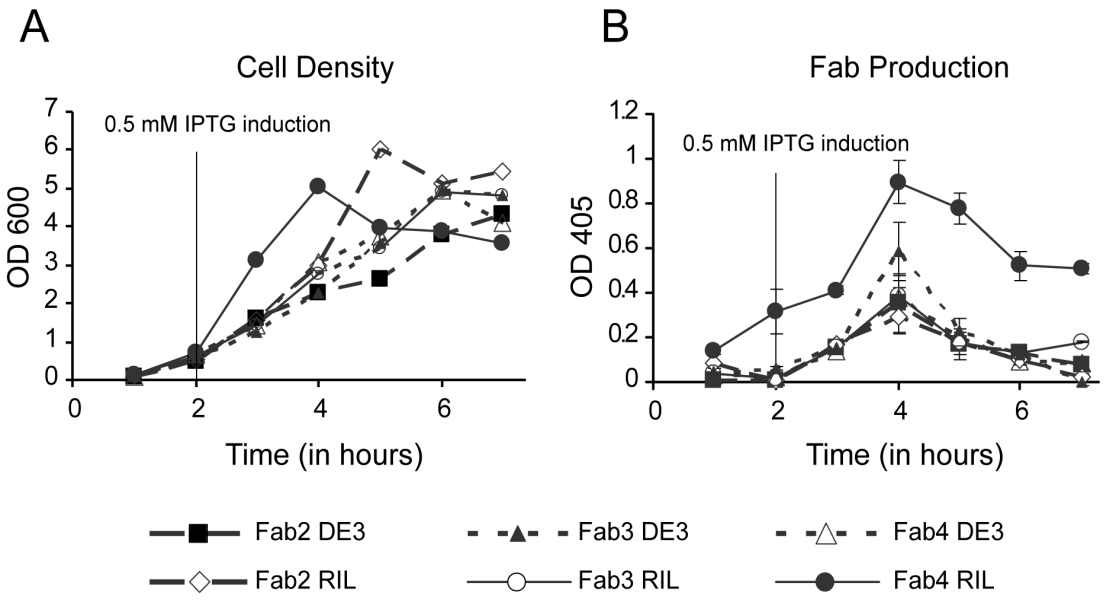

A
Cell Density
0.5 mM IPTG induction
OD 600
Time (in hours)
B
Fab Production
0.5 mM IPTG induction
OD 405
Time (in hours)
Fab2 DE3
Fab3 DE3
Fab4 DE3
Fab2 RIL
Fab3 RIL
Fab4 RIL

Figure 3.3. Bacterial growth kinetics and production of Fab4 using two different protein induction methods.

E. coli BL21(DE3) RIL cells freshly transformed with pET28-Fab4 DNA were grown in four different media (100 mL) in 500 mL baffled flasks at room temperature from the same inoculum. For lactose induction, cells were grown in defined medium supplemented with 0.2% w/v glycerol and lactose was added at a final concentration 2 g L^{-1} when late log phase was reached. For IPTG induction, cells grown in defined medium supplemented with 0.4% w/v glucose were induced at mid log phase with 0.5 mM IPTG. For the no induction conditions, lactose was not added to the glycerol medium and IPTG was not added to the glucose containing medium. Aliquots drawn at regular time intervals, pre and post induction, were used to monitor and plot a cell growth curve based on optical density at 600 nm as shown in Panel A. Panel B shows the ELISA signal, representing Fab4 production from supernatants of the cell aliquots, using the two different protein induction methods.

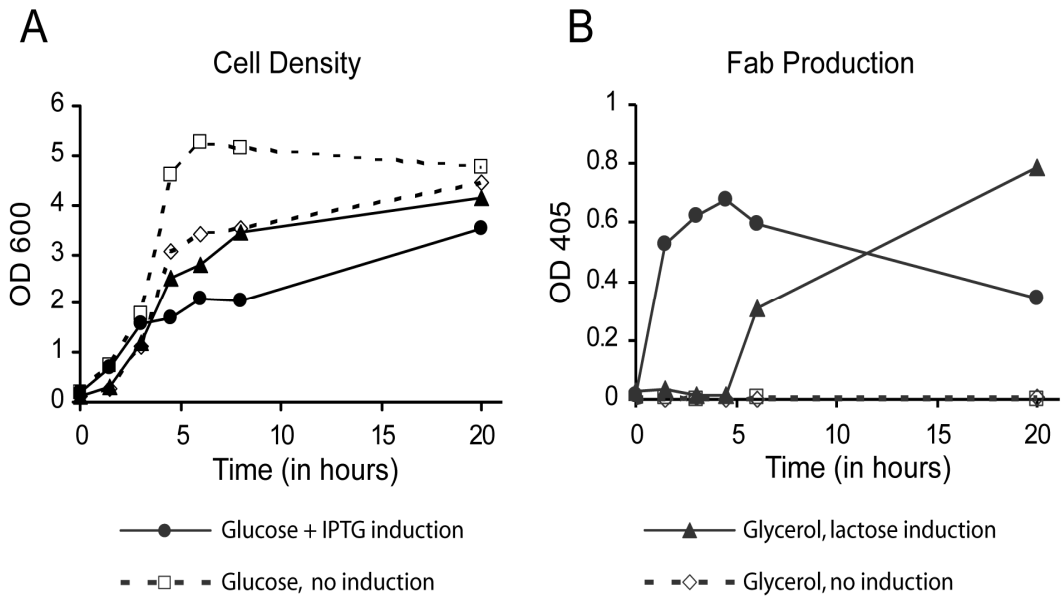

A
Cell Density
OD 600
Time (in hours)
Glucose + IPTG induction
Glucose, no induction
B
Fab Production
OD 405
Time (in hours)
Glycerol, lactose induction
Glycerol, no induction

Figure 3.4. Bacterial growth kinetics and Fab4 production with auto-induction media.

E. coli BL21(DE3)-RIL cells freshly transformed with pET28-Fab4 plasmid were grown and induced according to the auto-induction protocol (section 2.2). Cells from the same inoculum were grown in defined medium with 0.2% w/v glycerol and late log phase lactose induction, and defined medium with 0.4% w/v glucose and 0.5 mM IPTG induction. Aliquots drawn at regular time intervals, pre and post induction, were used to monitor and plot a growth curve, which is shown in Panel A. Fab4 production is represented by the plot of ELISA signal produced by the supernatant of the collected aliquots versus growth time.

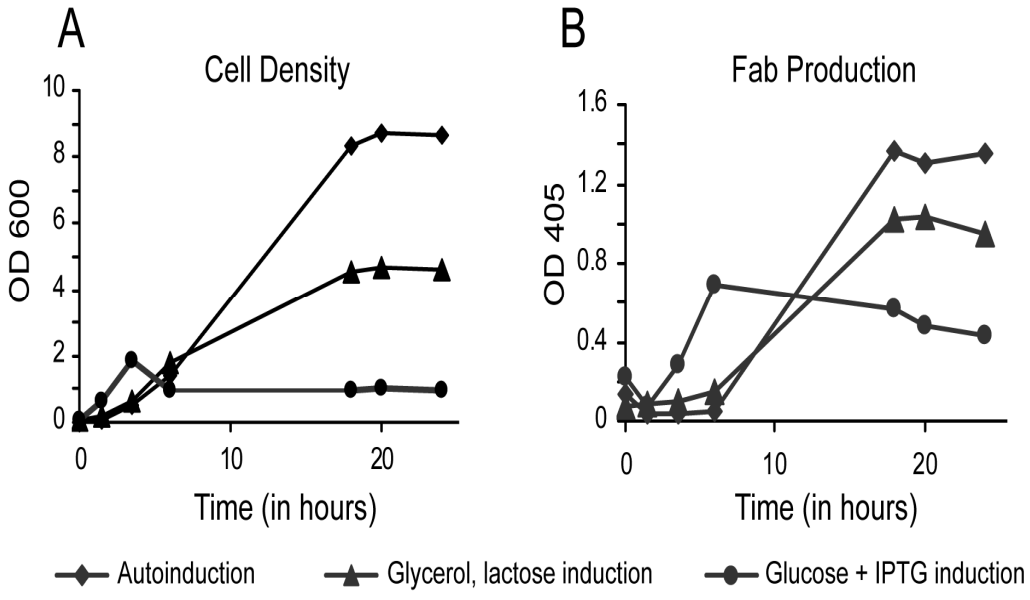

A
Cell Density
OD 600
10
8
6
4
2
0
0
10
20
Time (in hours)
B
Fab Production
OD 405
1.6
1.2
0.8
0.4
0
0
10
20
Time (in hours)
Autoinduction
Glycerol, lactose induction
Glucose + IPTG induction

Figure 3.5. Metal-chelate chromatography of rFab4 prepared in auto-induction and defined glucose media after thiosepharose chromatography.

Panel A shows the chromatogram for 0.5 L auto-induction medium grown culture of BL21(DE3) RIL-pET28 Fab4 cells that yielded 12 mgL^{-1} and Panel B shows the chromatogram of a 0.5 L culture of BL21 (DE3) RIL-pET28 Fab4 cells grown in glucose-supplemented, defined medium induced with 0.5 mM IPTG, that yielded 4 mg L^{-1} Fab4. Culture medium supernatants were harvested by centrifugation, eluted through a thiosepharose column and loaded directly onto a His-Trap metal chelate column as described in the *methods* section.

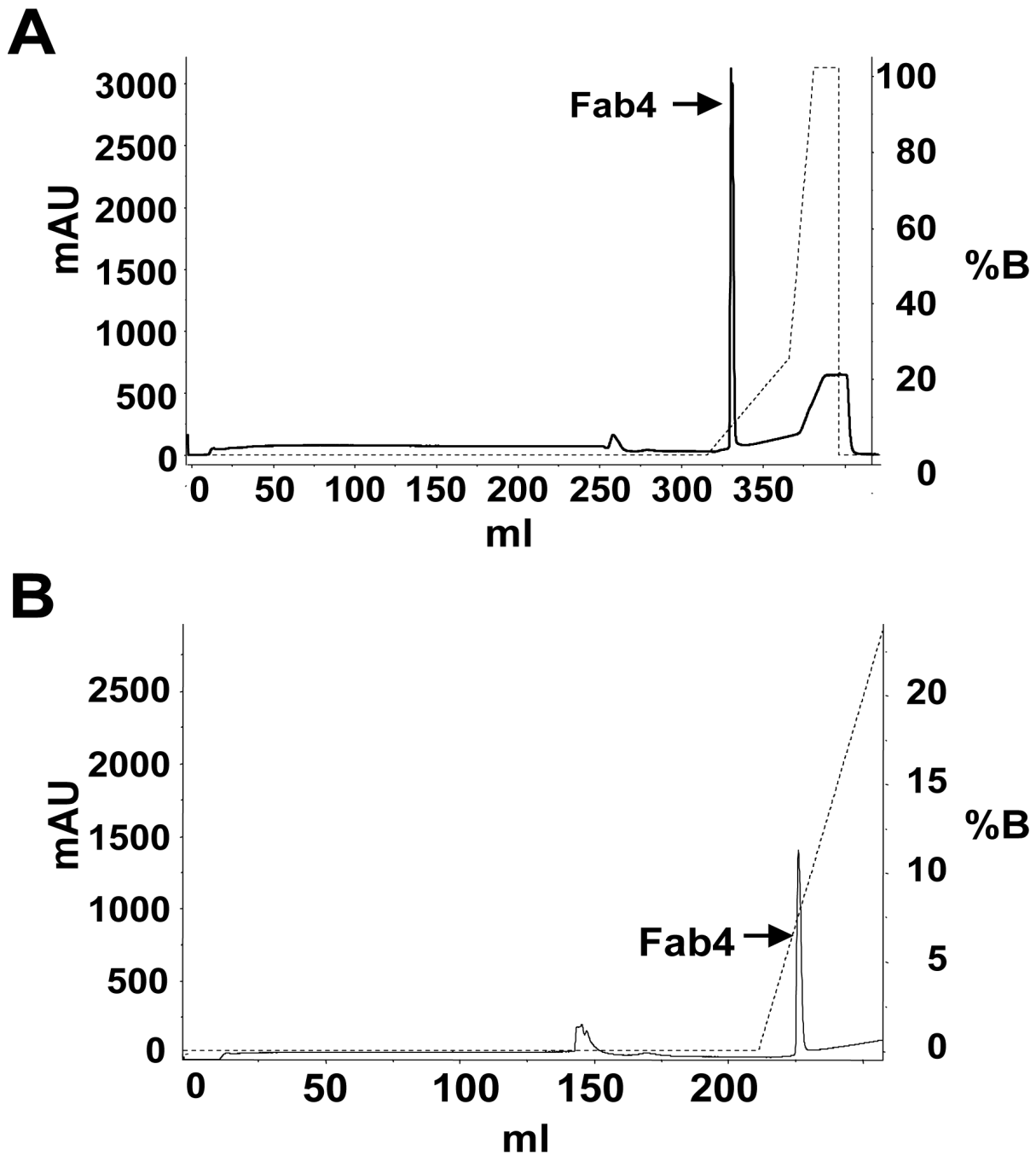

A
3000
2500
2000
1500
1000
500
0
mAU
Fab4
100
80
60
40
20
0
%B
0 50 100 150 200 250 300 350
ml
B
2500
2000
1500
1000
500
0
mAU
Fab4
20
15
10
5
0
%B
0 50 100 150 200
ml

Figure 3.6. "One-step" purification of antibody fragments using an ÄKTA purifier.

A liter of defined medium supernatant containing Fab4 was purified by the automated "one-step" method as described in the *Methods*. The chromatogram data shown are: optical absorbances at 280 nm (line), 360 nm (strong dash) and 420 nm (large dot), and the gradient control (dash).

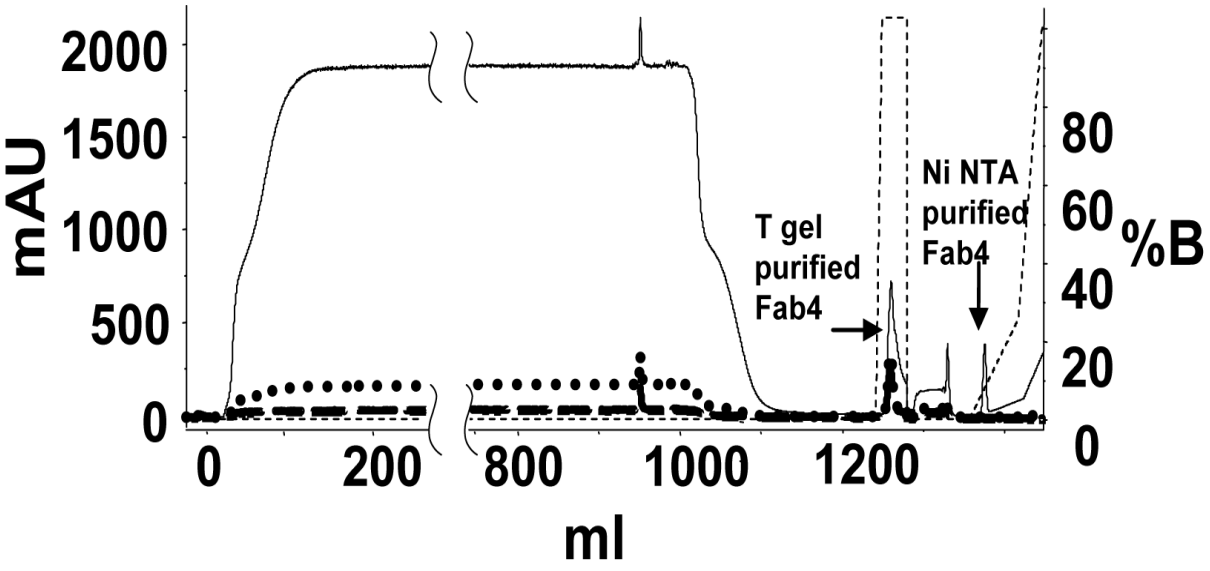

2000
1500
1000
500
0
mAU
0
200
800
1000
1200
ml
80
60
40
20
0
%B
T gel
purified
Fab4
Ni NTA
purified
Fab4

Figure 3.7. SDS-PAGE analysis of Fab4 purified using three different purification protocols.

Samples of rFab4 purified from defined culture media were analyzed by SDS-PAGE. Lane 1- ”one-step” protocol using thiosepharose followed by Ni-charged metal chelate chromatography, and lane 2- HiTrap Q column coupled (piggybacked) onto a Ni-charged metal chelate (HisTrap). Lane 3- Fab4 concentrated to 10 mg mL^{-1} that was purified by three separate chromatography steps (thiosepharose, Ni-chelate, and Q).

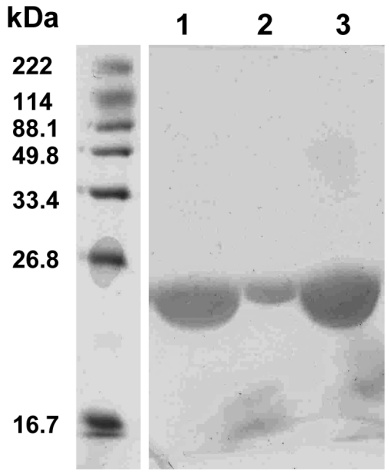
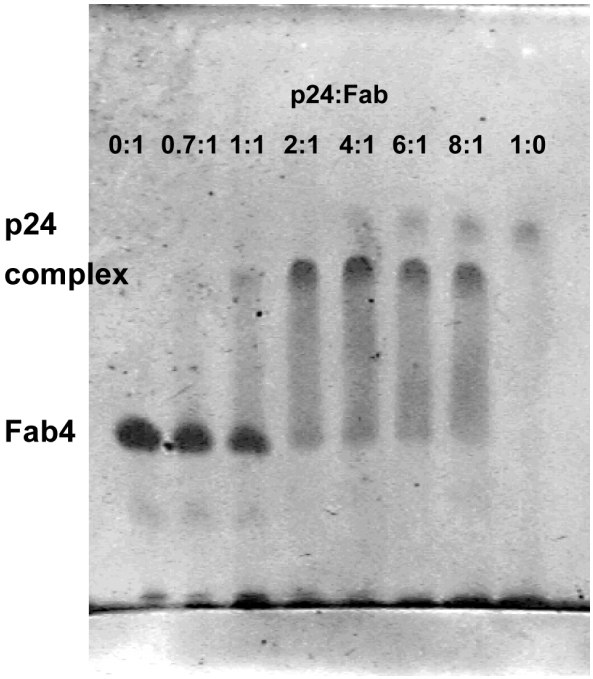

kDa
1
2
3
222
114
88.1
49.8
33.4
26.8
16.7

Figure 3.8. Native PAGE analysis of rFab4 complexed with its antigen HIV p24.

Samples (0.3 μL) of rFab4, p24 and rFab4-p24 complexes were applied to a 7.5% homogenous Phast gel (GE Biosciences). Complexes having molar ratios of p24 to rFab4 ranging from 0.7:1 to 8:1 were applied to the gel as shown. A molar ratio of two p24 (a p24 dimer) to one rFab4 is the ratio at which the individual component bands for p24 and rFab4 largely disappear. This ratio represents the optimal mixture for complete complex formation. Notice that only one species of complex results, suggesting that either the Fab sterically hinders formation of smaller (such as 1:1, a p24 monomer with a Fab) or larger species (2:2, *i.e.* a p24 dimer with two Fabs) or that the published means of quantifying the Fab or p24 are incorrect.


p24:Fab
0:1
0.7:1
1:1
2:1
4:1
6:1
8:1
1:0
p24
complex
Fab4

References

[1] J. McCafferty, A.D. Griffiths, G. Winter, D.J. Chiswell, Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348 (1990) 552-554.

[2] A.S. Kang, C.F. Barbas, K.D. Janda, S.J. Benkovic, R.A. Lerner, Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc Natl Acad Sci U S A 88 (1991) 4363-4366.

[3] J.C. Bardwell, Building bridges: disulphide bond formation in the cell. Mol Microbiol 14 (1994) 199-205.

[4] J. Buchner, R. Rudolph, Renaturation, purification and characterization of recombinant Fab-fragments produced in Escherichia coli. Biotechnology (N Y) 9 (1991) 157-162.

[5] M. Venturi, C. Seifert, C. Hunte, High level production of functional antibody Fab fragments in an oxidizing bacterial cytoplasm. J Mol Biol 315 (2002) 1-8.

[6] P. Martineau, P. Jones, G. Winter, Expression of an antibody fragment at high levels in the bacterial cytoplasm. J Mol Biol 280 (1998) 117-127.

[7] L.C. Kovari, C. Momany, M.G. Rossmann, The use of antibody fragments for crystallization and structure determinations. Structure 3 (1995) 1291-1293.

[8] C. Hunte, H. Michel, Crystallisation of membrane proteins mediated by antibody fragments. Curr Opin Struct Biol 12 (2002) 503-508.

[9] C. Momany, L.C. Kovari, A.J. Prongay, W. Keller, R.K. Gitti, B.M. Lee, A.E. Gorbalenya, L. Tong, J. McClure, L.S. Ehrlich, M.F. Summers, C. Carter, M.G. Rossmann, Crystal structure of dimeric HIV-1 capsid protein. Nat Struct Biol 3 (1996) 763-770.

[10] D. Rothlisberger, K.M. Pos, A. Pluckthun, An antibody library for stabilizing and crystallizing membrane proteins - selecting binders to the citrate carrier CitS. FEBS Lett 564 (2004) 340-348.

[11] L.L. Kelley, C. Momany, Generation of a phagemid mouse recombinant antibody fragment library by multisite-directed mutagenesis. Biotechniques 35 (2003) 750-752, 754, 756 passim.

[12] C.F. Barbas, Phage display : a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.

[13] F.W. Studier, Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41 (2005) 207-234.

[14] T.W. Hutchens, J. Porath, Thiophilic adsorption of immunoglobulins--analysis of conditions optimal for selective immobilization and purification. Anal Biochem 159 (1986) 217-226.

[15] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (1970) 680-685.

[16] R.S. Donovan, C.W. Robinson, B.R. Glick, Optimizing the expression of a monoclonal antibody fragment under the transcriptional control of the Escherichia coli lac promoter. Can J Microbiol 46 (2000) 532-541.

[17] J.V. Straight, D. Ramkrishna, S.J. Parulekar, N.B. Jansen, Bacterial-Growth on Lactose - an Experimental Investigation. Biotechnology and Bioengineering 34 (1989) 705-716.

[18] T. Inada, K. Kimata, H. Aiba, Mechanism responsible for glucose-lactose diauxie in Escherichia coli: challenge to the cAMP model. Genes Cells 1 (1996) 293-301.

[19] K. Kimata, H. Takahashi, T. Inada, P. Postma, H. Aiba, cAMP receptor protein-cAMP plays a crucial role in glucose-lactose diauxie by activating the major glucose transporter gene in Escherichia coli. Proc Natl Acad Sci U S A 94 (1997) 12914-12919.

[20] M.A. Blight, C. Chervaux, I.B. Holland, Protein secretion pathway in Escherichia coli. Curr Opin Biotechnol 5 (1994) 468-474.

[21] R.C. Hockney, Recent developments in heterologous protein production in Escherichia coli. Trends Biotechnol 12 (1994) 456-463.

[22] G. Forsberg, M. Forsgren, M. Jaki, M. Norin, C. Sterky, A. Enhorning, K. Larsson, M. Ericsson, P. Bjork, Identification of framework residues in a secreted recombinant antibody fragment that control production level and localization in Escherichia coli. J Biol Chem 272 (1997) 12430-12436.

[23] K. Proba, L. Ge, A. Pluckthun, Functional antibody single-chain fragments from the cytoplasm of Escherichia coli: influence of thioredoxin reductase (TrxB). Gene 159 (1995) 203-207.

[24] R. Levy, R. Weiss, G. Chen, B.L. Iverson, G. Georgiou, Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones. Protein Expr Purif 23 (2001) 338-347.

[25] B.R. Glick, Metabolic load and heterologous gene expression. Biotechnol Adv 13 (1995) 247-261.

[26] P. Neubauer, K. Hofmann, Efficient use of lactose for the lac promoter-controlled overexpression of the main antigenic protein of the foot and mouth disease virus in Escherichia coli under fed-batch fermentation conditions. FEMS Microbiol Rev 14 (1994) 99-102.

[27] P. Neubauer, K. Hofmann, O. Holst, B. Mattiasson, P. Kruschke, Maximizing the expression of a recombinant gene in Escherichia coli by manipulation of induction time using lactose as inducer. Appl Microbiol Biotechnol 36 (1992) 739-744.

[28] J. Porath, F. Maisano, M. Belew, Thiophilic adsorption--a new method for protein fractionation. FEBS Lett 185 (1985) 306-310.

[29] M. Belew, N. Juntti, A. Larsson, J. Porath, A one-step purification method for monoclonal antibodies based on salt-promoted adsorption chromatography on a 'thiophilic' adsorbent. J Immunol Methods 102 (1987) 173-182.

[30] M. Fiedler, A. Skerra, Use of thiophilic adsorption chromatography for the one-step purification of a bacterially produced antibody F(ab) fragment without the need for an affinity tag. Protein Expr Purif 17 (1999) 421-427.

[31] P. Carter, R.F. Kelley, M.L. Rodrigues, B. Snedecor, M. Covarrubias, M.D. Velligan, W.L. Wong, A.M. Rowland, C.E. Kotts, M.E. Carver, et al., High level Escherichia coli expression and production of a bivalent humanized antibody fragment. Biotechnology (N Y) 10 (1992) 163-167.

Acknowledgments

This work was funded by the National Science Foundation, grant MCB 9723244 to CM, and the University of Georgia Research Foundation faculty assistance grants program.

CHAPTER 4

COMPARISON OF CYTOPLASMIC AND SECRETED PRODUCTION OF A MOUSE RECOMBINANT ANTIBODY FRAGMENT FROM *ESCHERICHIA COLI*.[6]

[6] Nadkarni, A. and Momany, C. To be submitted to *Protein Expression and Purification*

Abstract

Properly folded and functional antibody fragments, such as Fab and scFv, have been commonly isolated from the bacterial, oxidizing, periplasmic space. However, secreted protein yields have been low as compared to heterologous proteins that can fold properly inside the E.coli cytoplasm. Even though the reducing environ of the E.coli cytoplasm is not suitable for formation of stable disulfide bonds, a prerequisite for the stable tertiary structure of these antibody fragments, it still provides a much larger space for protein accumulation as compared to the periplasmic space. Here we investigate the use of E.coli strains with designed oxidizing cytoplasmic milieus, in the absence of other aids such as molecular chaperones and mutational screening, on the yield of intracellular antibody fragments. The cytoplasmic mutant of an rFab fragment to HIV capsid protein p24 was created and expression in oxidizing and reducing bacterial strains was investigated and compared to a previously studied secreted mutant of the same fragment. The cytoplasmic yields obtained from both the oxidizing bacterial strains (0.168 mg L^{-1}) and the reducing bacterial strains (0.46 mg L^{-1}) were very poor as compared to the secreted fragments from the two strains (oxidizing: 0.6 mg L^{-1} and reducing: 8 mg L^{-1}). Our results show that it is possible to produce stable, properly folded and functional Fab fragments from the cytoplasm of *E.coli* cells without *in vivo* selection or aid of molecular chaperones and further optimization is needed to improve yields.

Keywords: antibody fragment; Fab; oxidizing cytoplasm; secreted; expression; p24; HIV capsid

Introduction

Antibody fragment mediated co-crystallization of target proteins is a valuable addition to the growing repertoire of tools established to resolve the macromolecular crystallographic bottleneck [1, 2]. Most of the protein structures solved as co-crystallization complexes to date have used antibody fragments generated proteolytically from monoclonal antibodies generated by hybridoma technology. Enzymatic proteolysis can lead to heterogeneous protein preparations that cannot crystallize without further purification. Recombinant antibody fragments such as Fabs and scFvs have become powerful tools in therapeutics, molecular diagnostics and basic sciences due to their inherent stability yet potential for tremendous diversity, possibility of homogenous yields and capability to be genetically engineered [3] [4]. Furthermore, novel selection technologies such as phage display have emerged allowing the isolation of antibody fragments with nanomolar affinities to be made to order against virtually any antigen, including self antigens [5] [6]. Isolating pure and homogenous antibody fragments in large amounts from bacterial sources is non trivial and thus a limiting factor for their wide use. Typically, antibody fragments are targeted to and isolated from the oxidizing bacterial periplasm which offers the ideal milieu needed for oxidation of cysteines to form disulfide bonds [7]. The bacterial cytoplasm, on the other hand, provides a reducing environment leading to low yields of unfolded and unstable protein [8] [7]. We have recently created a stable recombinant Fab (rFab) mutant that recognizes the HIV capsid protein p24 and can be isolated at high yields (12 mg L^{-1}) from bacterial culture media [9].

Based on this mutant, a non-immune rFab phage display library with a nominal diversity of 1.16×10^7 was built [10]. Classically, antibody fragment libraries, including the one in our laboratory are based on filamentous phage (M13), a nonlytic propagative phage which requires

that all components of the phage particle be exported through the bacterial inner membrane before phage assembly. This is extremely convenient for antibody fragments that require export to the periplasm to fold in to a stable, native conformation. However, the filamentous phage life cycle is lengthy and thus the *in vivo* amplification steps between adjacent panning rounds become very time consuming and indisposed to automation [11]. The bacteriophage T7 has several properties that make it a more attractive display vector. Its robust nature makes it stable to harsh conditions and thus preserves phage infectivity in the presence of a variety of reagents used in the biopanning process. T7 also has a very rapid growth and replication cycle in comparison to filamentous phage. Cultures lyse 1–2 h after infection, decreasing the time needed to perform the multiple rounds of amplification usually required for selection [12]. The T7 assembly takes place inside the cytoplasm and thus proteins displayed on T7 bacteriophages should be expressed, folded and stable in the cytoplasmic reducing environment. Thus developing a bacterial cytoplasmic expression system is a prerequisite to using the T7 based phage display system for the high throughput generation of antibody fragments.

This work reports the design and evaluation of a bacterial expression system for the cytoplasmic production of our model rFab fragment that recognizes the HIV capsid protein p24. Our results indicate that it is possible to make intracellular antibody fragments in *E.coli*, though yield and purity as compared to the rFab isolated from growth media or *E.coli* periplasm were poor.

Materials and Methods

Construction of anti-p24 rFab for cytoplasmic Fab expression

Based on a previously designed vector for Fab secretion (pCOMB-Fab4), we constructed the plasmid pFab4IC by deletion of the secretory sequences OmpA and PelB. The oligonucleotides dsecsl-f1 5'-GTCCTCTTCCATGGCCGAGCTCGTGTTGACCC and dseca-R 5'-GCTGCACCTCGGCCATTTTAAATTCCTCCTAATTAATTATC were used as primers to amplify the light chain from the pCOMB-Fab4 construct by PCR using Pfu Turbo (Stratagene). Oligonucleotides dsecb-F 5'- GGAGGAATTTAAAATGGCCGAGGTGCAGCTGCTC and dsecsl-v1 5'-CGGCTCTTCGTTAACTAGTATGATGGGGCACAATTTTCTTG were used similarly as primers to amplify the heavy chain from the pCOMB-Fab4 construct. Twenty nine bases of dseca-R and dsecb-F are complementary. Equimolar amounts of the heavy chain and light chain PCR products were purified by agarose gel and were amplified in a second PCR reaction, using the Seamless® Cloning kit (Stratagene) and flanking oligonucleotides dsecsl-f1 and dsecsl-v1, so that the final overlap PCR product has no signal sequences and restriction enzyme Eam1104 I sites (underlined) at the ends.

pET28b (Novagen) was similarly PCR amplified as recipient construct using oligonucleotides pet28slfwd 5'-GCGCTCTTCGTAATATGGCTAGCATGACTGG and pet28slrev 5'-CCGCTCTTCCCATGGTATATCTCCTTCTTAAAG to give a linearized, product with restriction enzyme Eam1104 I sites at the ends. The PCR products, for the recipient construct and for the insert (rFab without secretory sequences) were digested with Eam1104 I and then ligated in the presence of the restriction enzyme to give the final construct pFab4IC following the Seamless® Cloning kit (Stratagene). The pFab4IC plasmid was verified by sequencing from the T7 promoter and T7 terminator primers as well as from two internal oligonucleotide primers. Sequencing was performed on a 2ABI 3100 system by the Sequencing and Synthesis Facility at the University of Georgia.

Production and purification of rFabs by auto-induction in two different bacterial strains

Single colonies of *E.coli* Rosetta 2 or Rosetta-gami 2, freshly transformed with pFab4IC, were used to inoculate 2 mL of auto-induction media [13] supplemented with 50 mg L^{-1} kanamycin and 34 mg L^{-1} chloramphenicol. After growth to 1.0 OD_{600nm} (4-5 hours), 0.5 mL of the inoculum was diluted into 0.5 L of auto-induction media and incubated with shaking for 2-3 hours at 37 °C before transferring it to room temperature. From this, cultures were grown to saturation (OD_{600nm} ~ 5-8) in 12-14 hours (for Rosetta 2) and 24 hours (for Rosetta-gami 2). After 3-4 hours at saturation, cultures were centrifuged at 6000 × g for 30 minutes at 4 °C and the bacterial pellets were resuspended in 5 mL of HPLC start buffer (1mM PMSF, 0.5M Na_2SO_4, 50mM NaH_2PO_4, pH 7.0), to which a cOmplete Protease Inhibitor Cocktail Tablet was added. The cells were then disrupted via two passages through an ice-chilled ThermoIEC French Pressure Cell at 16,000 psi. The resulting preparation was centrifuged at 50,000 × g for 30 minutes at 4°C and the cytoplasmic soluble rFab was isolated in the supernatant. The secreted rFab4 was prepared from the pET28-Fab4 transformed in *E.coli* Rosetta 2 or Rosetta-gami 2 by auto-induction as previously described [9, 13].

The secreted as well as cytoplasmic proteins were isolated by the 'single step' purification protocol described previously [9]. Briefly, the proteins were purified by thiosepharose and Ni^{2+} -NTA affinity chromatography in a single step. The cytoplasmic rFab fractions obtained from the Thiophilic/Ni-NTA purification were further polished by anion exchange chromatography. The protein peak was dialyzed against buffer Q (20 mM Tris HCl, pH 9), loaded onto a 5 mL Hi-Trap Q anion exchange column (GE Biosciences) and eluted using a 20 column volume gradient of buffer Q to 20 mM Tris, 1 M NaCl, pH 9.

Protein purity analysis, quantitation and functional evaluation

The purity of the rFabs was evaluated by SDS-PAGE analysis [14] using 12% cross-linked gels in either a Hoeffer Mighty Small gel apparatus or a PHAST gel system (GE Biosciences). The protein bands were made visible with Coomassie Brilliant Blue Stain.

Protein was quantitated by dye binding (BioRad) with BSA as a standard or using an absorbance coefficient of $A_{280nm}^{0.1\%} = 0.553$.

ELISAs were performed on Immulon II™ 96-well EIA/RIA plates (Dynex Technologies, Inc.) coated with 20 μg mL^{-1} recombinant HIV capsid protein p24. Following the wash and blocking steps, 0.1 mL samples and/or their dilutions were applied to the wells as a primary antibody. Either a goat or rabbit anti-mouse IgG $F(ab')_2$-alkaline phosphatase conjugate (Pierce Biotech) was used as the secondary antibody. Recombinant HIV capsid protein having a polyhistidine purification tag used as the antigen in ELISAs was expressed in *E. coli* and purified using metal-chelate chromatography using a plasmid generously provided by Carol Carter, Department of Molecular Genetics and Microbiology, Stony Brook University Stony Brook, NY.

Results

The cytoplasmic expression of an rFab fragment to HIV capsid protein p24 was investigated and compared to the secreted expression of the same fragment. In order to enable expression of the rFab in *E.coli* cytoplasm, the signal sequences upstream of the heavy chain and light chain, OmpA and PelB, were deleted by overlap PCR from the plasmid pCOMB-Fab4. The final cytoplasmic expression plasmid (pFab4IC) was constructed by sub-cloning the PCR product into the pET 28b vector. The remaining expression features such as the T7 promoter, the

two Shine-Dalgarno sequences (6-8 nucleotides before the initiation codon of both the rFab heavy and light chains) and the restriction sites found in the pET28-Fab4 [9] were confirmed to be maintained by DNA sequencing and restriction analysis.

Both rFab constructs, pFab4IC (for cytoplasmic expression) and pET28-Fab4 (for secreted Fab4), were freshly transformed into two *E.coli* strains Rosetta™ 2 and Rosetta-gami™ 2. The Rosetta™ 2 strains have 7 plasmid-encoded rare tRNAs that improve expression of genes with codons poorly represented in *E.coli* [15]. The Rosetta-gami™ 2 strain additionally has mutations in the thioredoxin reductase (trxB) and glutathione reductase (gor) genes that provide a more oxidizing cytoplasmic environ suitable for the formation of disulfide bonds [16]. For the purpose of comparison, the protein production from the two constructs in each strain followed the protocol optimized for the pET28-Fab4 in BL21 (DE3)-RIL or Rosetta™ 2 [9]. Fab4 fragments were isolated from the cytoplasmic extracts of pFab4IC transformed Rosetta™ 2 and Rosetta-gami™ 2 and from the cell growth medium (supernatant) of pET28-Fab4 transformed Rosetta™ 2 and Rosetta-gami™ 2 cells. The weak metal binding site from the secreted Fab4 is preserved in the cytoplasmic construct and thus for purification we initially used the single step purification protocol based on thiophilic interaction and metal affinity chromatography as described previously [9]. Both the strains produced cytoplasmic Fab fragments that eluted at approximately similar conductivities on the Ni-NTA column, but with a substantially reduced peak relative to the secreted Fab4. The Rosetta-gami™ 2 had a very slow growth rate and produced a much smaller peak on the Ni-NTA chromatogram as compared to the Rosetta™ 2 for both the secreted as well as the cytoplasmic fragments (Figure 4.1 Panel A-D). The enhanced oxidizing milieu of the mutant Rosetta-gami™ 2 did not seem to improve the relative protein yield, but did reduce the growth and stability of the cells. Reducing SDS PAGE analysis of the

cytoplasmic fractions (Figure 4.3 Panel A Lane 4 and Lane 5) showed the presence of several contaminants as compared to the secreted Fab (Lane 2 and Lane 3), probably because the culture media has much less contaminating proteins to start with than the cytoplasm. In order to remove these contaminants, the cytoplasmic and secreted fractions were further separated by anion exchange chromatography (HiTrap Q FF). The cytoplasmic protein showed several peaks on the chromatogram as compared to the single peak for the secreted Fab4 (Figure 4.2 Panel A & B). As compared to the Q column chromatogram for secreted Fab, the cytoplasmic Fab chromatogram showed only a small peak at the approximate conductivity that the secreted Fab eluted out. However a much larger peak eluted at higher salt concentration that did not give any ELISA signal, when tested for antigen binding with the HIV capsid protein p24. The smaller peak was tested for antigen binding by ELISA and produced a very weak signal as compared to secreted Fab. Analysis of the ELISA signal, produced by the cytoplasmic and secreted fractions in the two cell lines at different dilutions, showed that the undiluted secreted Fab4 signal from the Rosetta™ 2 was so high that it was outside the linear range of the instrument and was in the linear range only at 1:10 and 1:100 dilutions. In comparison, the undiluted cytoplasmic Fab4 signal from Rosetta™ 2 was in the linear range and further dilutions showed much weaker signal. A similar pattern was shown by the ELISA signal of the two proteins from the Rosetta-gami™ 2 strain (Figure 4.4). The histidine rich protein of *E.coli* (SlyD) has been shown to be a co-purifying contaminant in the IMAC (Immobilized Metal Affinity Chromatography) purification of 6xHis-tagged proteins [17] and that could be the cause of the misleading peak on the Ni-NTA chromatogram of the cytoplasmic Fab4. Reducing SDS PAGE analysis (Figure4.4 Panel B) of the cytoplasmic (Lane 3: Rosetta™ 2, Lane 5: Rosetta-gami™ 2) and secreted Fab4 fractions (Lane 2: Rosetta™ 2, Lane 4: Rosetta-gami™ 2) from the two *E.coli* cell strains

showed a very sharp, large band for the secreted Fab4 from Rosetta™ 2 as compared to the weakly detectable band for the cytoplasmic Fab4 from Rosetta™ 2. The Rosetta-gami™ 2 strain showed a weak band for both the cytoplasmic as well as the secreted Fab4. The final cytoplasmic protein yield obtained from Rosetta™ 2 was 0.46 mg L^{-1} and from Rosetta-gami™ 2 was 0.168 mg L^{-1} as compared to 8 mg L^{-1} secreted Fab4 from Rosetta™ 2.

Discussion

Antibodies and their fragments have so far been considered as the scaffold molecules of choice, for obtaining affinity binders from a wide variety of display technologies. Apart from crystallography (as co-crystallization reagents), these selected binders show tremendous promise in a plethora of applications like drug discovery [18] [19], microarray [20] and antibody-based chip technology [21]. Filamentous phage based display systems, despite being suitable for the display of disulfide containing antibody fragments, have limitations due to stability and production considerations of exported proteins. New classes of nonimmunoglobulin scaffolds [22], such as affibodies [23] and ankyrin repeat proteins [24] are being developed. These scaffolds do not have disulfide bonds and thus by pass the production issues that antibodies face. However, these scaffolds have limitations with regard to the range of antigens they can bind as compared to naturally diverse antibody fragments. Also, therapeutics developed from antibody scaffolds are much easier to humanize than other synthetic frameworks which might not be found in nature.

A more feasible alternative is to develop and optimize intracellular antibody expression and production in order to utilize the cheap and efficient production milieu of the reducing *E.coli* cytoplasm. Due to the presence of reducing components such as thioredoxin and glutathione, the

cytoplasm has a very low cysteine thiol-disulfide redox potential and thus disfavors the formation of stable disulfides [25] [26]. Oxidized forms of these proteins do favor formation of disulfide bonds, but are maintained in the reduced state by cytoplasmic proteins such as thioredoxin reductase (TrxB) and glutathione oxidoreductase (Gor) [16]. Several groups have experimented with these thio-redox pathways in bacteria to produce disulphide containing proteins in the cytoplasm. In some reports, the authors have devised *in vivo* selection strategies to derive an antibody fragment with greatly improved bacterial cytoplasmic expression. However, the selected antibody fragments did not require disulfide bonds for stable folding or function [27] [28]. The *trxB* and *gor* mutant bacterial strains (commercialized as Rosetta-gami™ 2) were reported to allow efficient disulphide bond formation [16] [25]. However, these cells show very poor growth properties with a doubling time ~ 300 min as observed by others as well as us. Using these strains, a maximum yield of 0.8 mg Fab L^{-1} per OD_{600nm} was obtained only when the Fab was expressed via the coexpression of molecular chaperones [8].

Our results show that it is possible to produce stable, properly folded and functional Fab fragments from the cytoplasm of *E.coli* cells without *in vivo* selection or aid of molecular chaperones. We have created a construct for cytoplasmic expression of a rFab fragment that detects HIV capsid protein p24. However, the yields are low as compared to the functionally similar Fab targeted to the periplasm and isolated from the culture media even from the mutant Rosetta-gami™ 2. There are two instances where the Rosetta-gami™ 2 cells have been used for cytoplasmic antibody fragment production – one for scFv production [29] (low yield ~ 0.3 mg Fab L^{-1} per OD_{600nm}) and the other for Fab that detects membrane protein NhaA [30] (high yield ~ 10-30 mg Fab L^{-1} per OD_{600nm}. Production from our intracellular construct made functional, stable, folded rFab fragments but at a low level. The intracellular production followed the same

protocol as the secreted rFab4, previously optimized for production. The dynamics of the two proteins (secreted and cytoplasmic) could be different despite the sequence similarities. Therefore, further optimization of this protocol to improve the yields of the cytoplasmic construct could give improved yields of the rFab4. Parameters such as temperature, culture media components and induction of protein expression need to be optimized for the Rosetta-gami™ cell line in order to get good growth rates and optimized growth yield.

In conclusion, these results demonstrate that oxidized and functional Fab can be produced efficiently in the *E.coli* cytoplasm and based on the evidence from the single report of high cytoplasmic yields, it is clear that further optimization of our construct should produce elevated levels of rFab fragments.

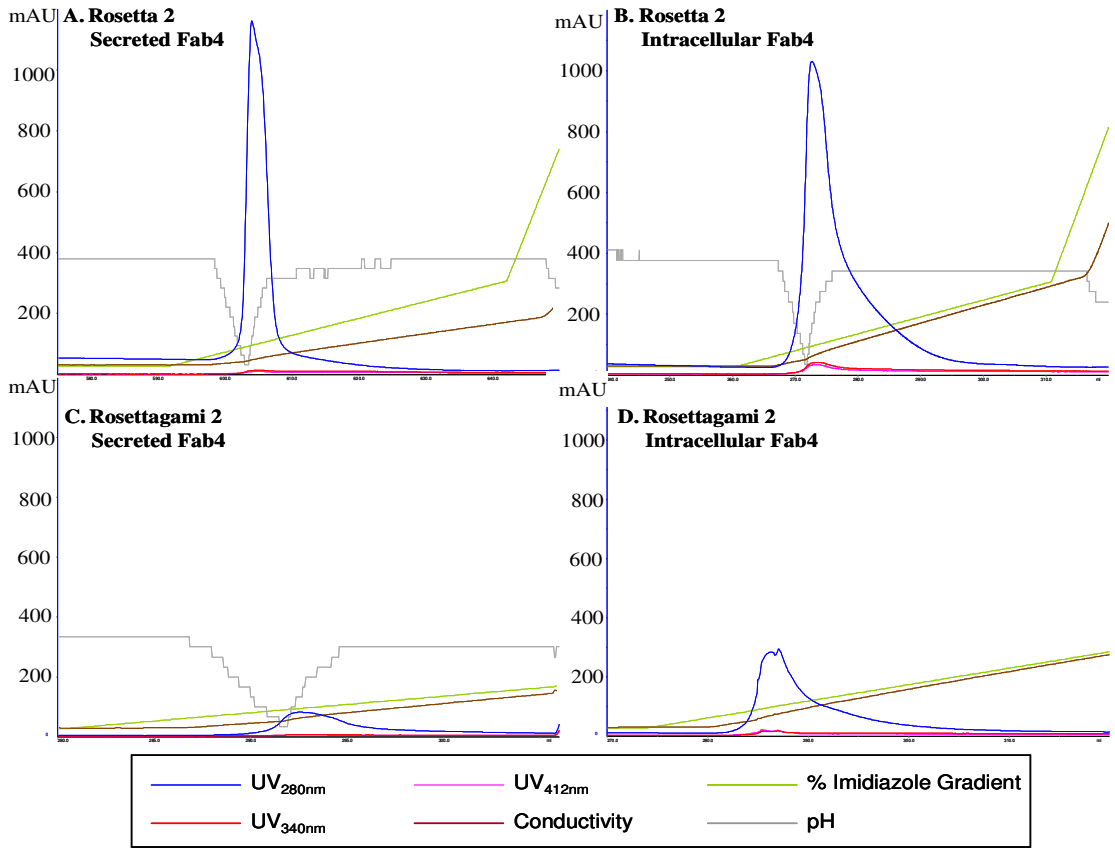


Figure 4.1 Ni^{+} Affinity Chromatography Profiles for Secreted and Cytoplasmic rFab4 from Different *E.coli* Strains

The pooled fractions for the respective proteins, off the thiosepharose column, were loaded onto a 5 ml His-trap Ni^{+}-charged column. Protein was eluted using a step-wise 0.5 M Imidazole gradient. Panel A: Secreted rFab4 using Rosetta™ 2 cells. Panel B: Cytoplasmic rFab4 from Rosetta™ 2 cells. Panel C: Secreted rFab4 from Rosetta-gami™ 2 cells. Panel D: Cytoplasmic rFab4 from Rosetta-gami™ 2 cells. The peak in Panel D is much smaller than that obtained for both proteins from the Rosetta™ 2, but higher than the secreted rFab4 from Rosetta-gami™ 2. This is misleading and further anion exchange analysis showed that a lot of the *E.coli* cytoplasmic histidine rich protein eluted with the rFab4 from the two cell lines, thus giving a bigger peak and UV signal.

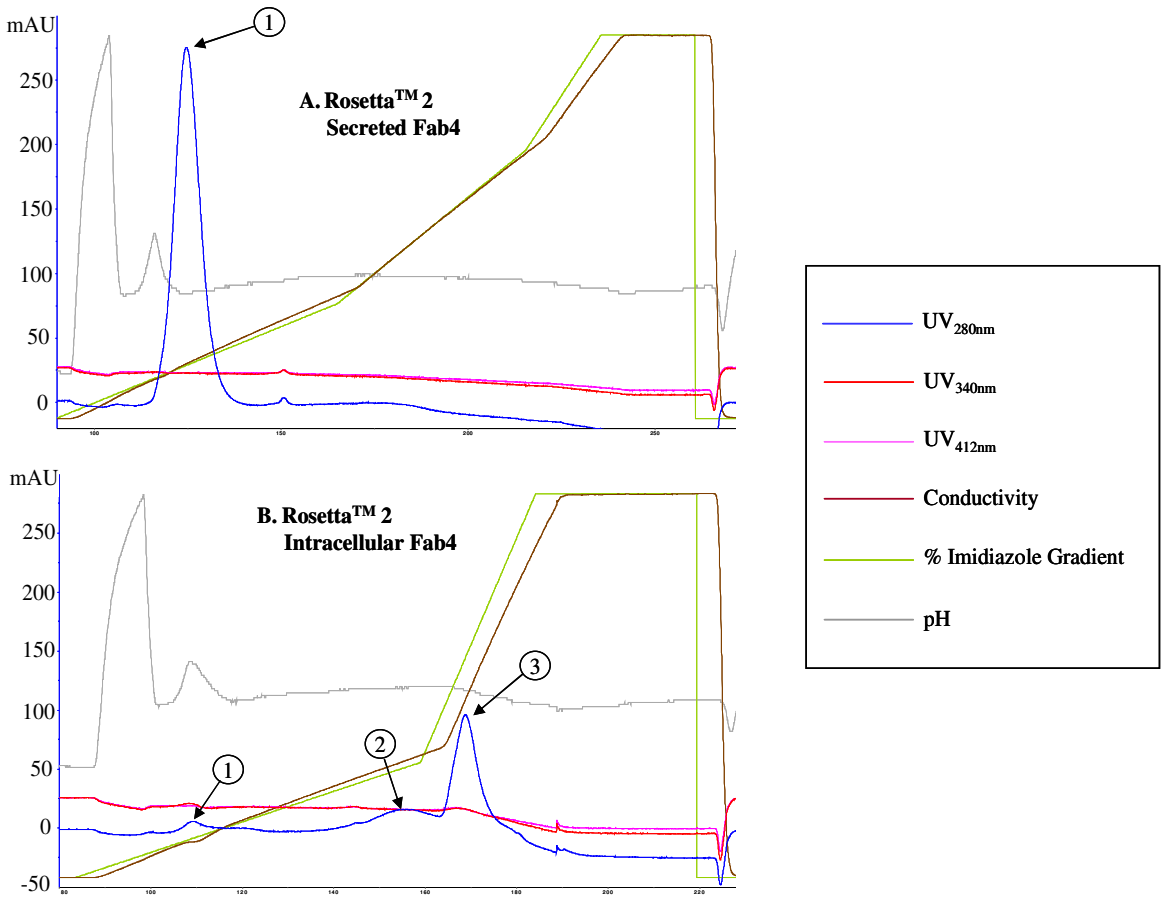


Figure 4.2 Anion Exchange Chromatography (Q-FF) Profiles for Secreted and Cytoplasmic rFab4 from Different *E.coli* Strains

The figure shows anion exchange chromatograms for the secreted and cytoplasmic rFab4 proteins, previously eluted of a metal affinity column. The intracellular protein, even after two previous column purifications produced a heterogeneous profile, with very little of functionally active cytoplasmic protein as seen in Panel B. Panel A shows a single, sharp, peak for the secreted protein. Peak 1 is the rFab4 peak in both the chromatograms (as marked by the pH inflection characteristically seen before the rFab elution). Peak 1, Peak 2 and Peak 3 from Panel B were tested for functionality by antigen (p24) binding ELISA

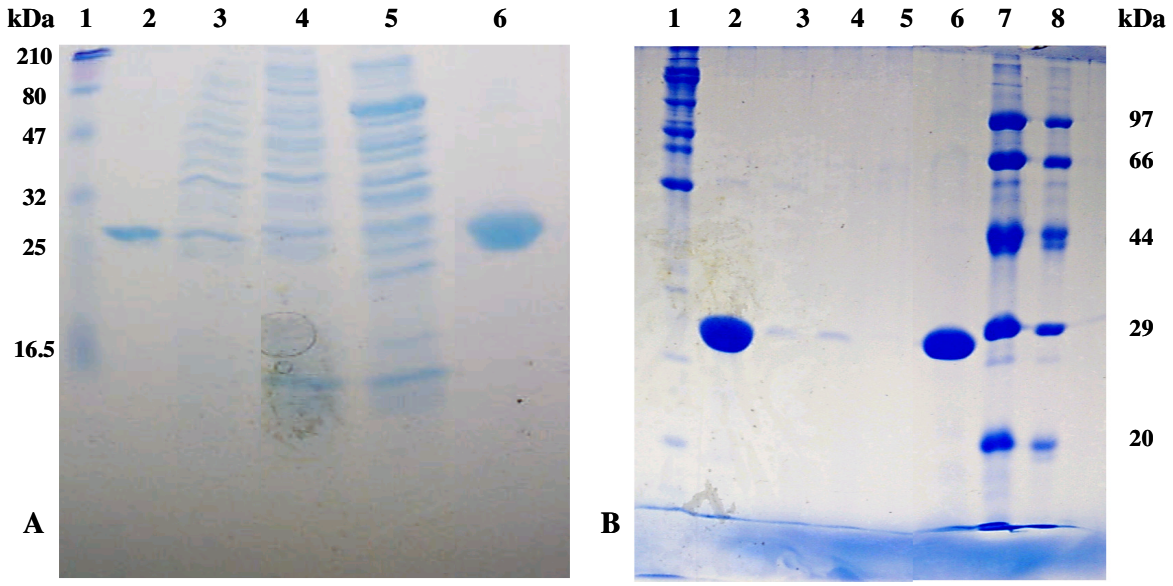


Figure 4.3. SDS-PAGE analysis of cytoplasmic and secreted Fab4 from two E.coli strains

Panel A: Pooled fractions of the peak corresponding to Figure 4.1 after Ni-NTA chromatography - Lane 1, Molecular weight markers; Lane 2, Secreted rFab4 from Rosetta™ 2; Lane 3, Secreted rFab4 from Rosetta-gami™ 2; Lane 4, Cytoplasmic rFab4 from Rosetta-gami™ 2; Lane 5, Cytoplasmic rFab4 from Rosetta™ 2; Lane 6, Purified rFab4 at 5mg mL^{-1}

Panel B: Pooled fractions of the peak corresponding to Figure 4.2 after Anion Exchange chromatography - Lane 1, 7 and 8, Molecular weight markers; Lane 2, Secreted rFab4 from Rosetta™2; Lane 3, Cytoplasmic rFab4 from Rosetta™2; Lane 4, Secreted rFab4 from Rosetta-gami™ 2; Lane 5, Cytoplasmic rFab4 from Rosetta-gami™ 2; Lane 6, Purified rFab4 at 5mg mL^{-1}

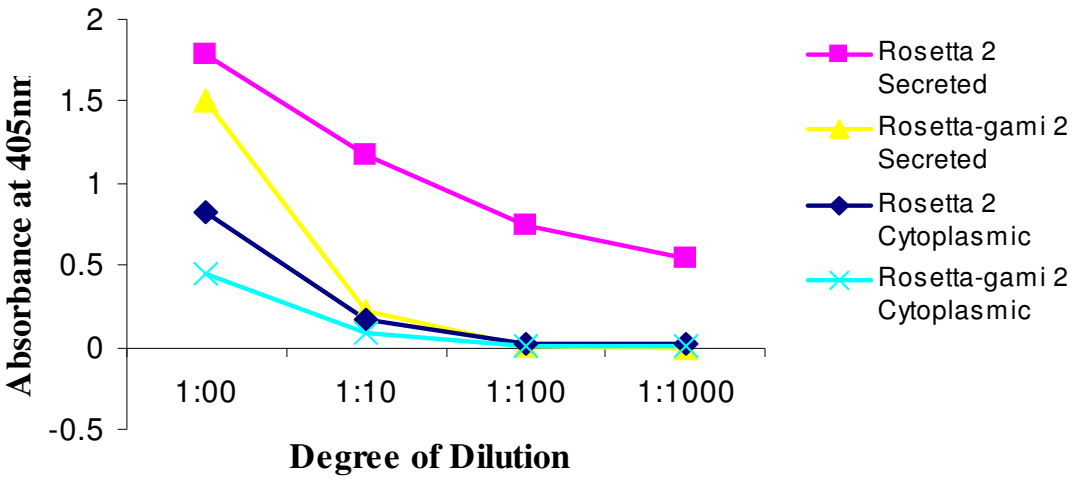


Figure 4.4. ELISA analysis of antigen binding activity of the cytoplasmic and secreted rFab4

ELISA plates were coated with HIV-1 capsid protein p24 and probed with dilutions of cytoplasmic and secreted rFab4 produced from the spontaneous induction of Rosetta™ 2 and Rosetta-gami™ 2 as primary antibody (color legend in the picture). ELISA signal at 405 nm at increasing degree of dilutions was plotted. Secreted fractions from Rosetta™ 2 at a 1:1,000 dilution gave higher signal than cytoplasmic fractions from Rosetta-gami™ 2 at undiluted concentrations.

References

[1] C. Momany, L.C. Kovari, A.J. Prongay, W. Keller, R.K. Gitti, B.M. Lee, A.E. Gorbalenya, L. Tong, J. McClure, L.S. Ehrlich, M.F. Summers, C. Carter, M.G. Rossmann, Crystal structure of dimeric HIV-1 capsid protein. Nature structural biology 3 (1996) 763-770.

[2] C. Hunte, H. Michel, Crystallisation of membrane proteins mediated by antibody fragments. Curr Opin Struct Biol 12 (2002) 503-508.

[3] C. Souriau, P.J. Hudson, Recombinant antibodies for cancer diagnosis and therapy. Expert Opin Biol Ther 3 (2003) 305-318.

[4] F. Breitling, S. Dèubel, Recombinant antibodies, John Wiley ;
Spektrum, New York
Heidelberg, 1999.

[5] J. McCafferty, A.D. Griffiths, G. Winter, D.J. Chiswell, Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348 (1990) 552-554.

[6] C.F. Barbas, 3rd, A.S. Kang, R.A. Lerner, S.J. Benkovic, Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proceedings of the National Academy of Sciences of the United States of America 88 (1991) 7978-7982.

[7] S. Biocca, F. Ruberti, M. Tafani, P. Pierandrei-Amaldi, A. Cattaneo, Redox state of single chain Fv fragments targeted to the endoplasmic reticulum, cytosol and mitochondria. Biotechnology (N Y) 13 (1995) 1110-1115.

[8] R. Levy, R. Weiss, G. Chen, B.L. Iverson, G. Georgiou, Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones. Protein Expr Purif 23 (2001) 338-347.

[9] A. Nadkarni, L.L. Kelley, C. Momany, Optimization of a mouse recombinant antibody fragment for efficient production from *Eschericia coli*. Protein Expr Purif (2006).

[10] L.L. Kelley, C. Momany, Generation of a phagemid mouse recombinant antibody fragment library by multisite-directed mutagenesis. Biotechniques 35 (2003) 750-752, 754, 756 passim.

[11] W.G. Willats, Phage display: practicalities and prospects. Plant Mol Biol 50 (2002) 837-854.

[12] A.H. Rosenberg, G. Griffin, F.W. Studier, M. McCormick, J. Berg, R. Mierendorf, T7Select phage display system: a powerful new protein display system based on bacteriophage T7., InNovations, 1996, pp. 1-6.

[13] F.W. Studier, Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41 (2005) 207-234.

[14] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (1970) 680-685.

[15] R. Novy, D. Drott, K. Yaeger, R. Mierendorf, Overcoming the codon bias of E. coli for enhanced protein expression, InNovations, 2001, pp. 1-3.

[16] W.A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds in the Escherichia coli Cytoplasm 10.1074/jbc.272.25.15661. J. Biol. Chem. 272 (1997) 15661-15667.

[17] S. Mukherjee, A. Shukla, P. Guptasarma, Single-step purification of a protein-folding catalyst, the SlyD peptidyl prolyl isomerase (PPI), from cytoplasmic extracts of Escherichia coli. Biotechnology and applied biochemistry 37 (2003) 183-186.

[18] J.R. Broach, J. Thorner, High-throughput screening for drug discovery. Nature 384 (1996) 14-16.

[19] M.H. Pausch, G-protein-coupled receptors in Saccharomyces cerevisiae: high-throughput screening assays for drug discovery. Trends in biotechnology 15 (1997) 487-494.

[20] A. Lueking, M. Horn, H. Eickhoff, K. Bussow, H. Lehrach, G. Walter, Protein microarrays for gene expression and antibody screening. Analytical biochemistry 270 (1999) 103-111.

[21] R.M. de Wildt, C.R. Mundy, B.D. Gorick, I.M. Tomlinson, Antibody arrays for high-throughput screening of antibody-antigen interactions. Nature biotechnology 18 (2000) 989-994.

[22] P.A. Nygren, A. Skerra, Binding proteins from alternative scaffolds. Journal of immunological methods 290 (2004) 3-28.

[23] K. Nord, J. Nilsson, B. Nilsson, M. Uhlen, P.A. Nygren, A combinatorial library of an alpha-helical bacterial receptor domain. Protein engineering 8 (1995) 601-608.

[24] H.K. Binz, M.T. Stumpp, P. Forrer, P. Amstutz, A. Pluckthun, Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. J Mol Biol 332 (2003) 489-503.

[25] P.H. Bessette, F. Aslund, J. Beckwith, G. Georgiou, Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. Proceedings of the National Academy of Sciences of the United States of America 96 (1999) 13703-13708.

[26] D. Ritz, J. Beckwith, Roles of thiol-redox pathways in bacteria. Annual review of microbiology 55 (2001) 21-48.

[27] P. Martineau, P. Jones, G. Winter, Expression of an antibody fragment at high levels in the bacterial cytoplasm. J Mol Biol 280 (1998) 117-127.

[28] K. Proba, A. Worn, A. Honegger, A. Pluckthun, Antibody scFv fragments without disulfide bonds made by molecular evolution. J Mol Biol 275 (1998) 245-253.

[29] P. Jurado, D. Ritz, J. Beckwith, V. de Lorenzo, L.A. Fernandez, Production of functional single-chain Fv antibodies in the cytoplasm of Escherichia coli. J Mol Biol 320 (2002) 1-10.

[30] M. Venturi, C. Seifert, C. Hunte, High level production of functional antibody Fab fragments in an oxidizing bacterial cytoplasm. J Mol Biol 315 (2002) 1-8.

CHAPTER 5

A RAPID PHAGE DISPLAY PROTOCOL TO SELECT ANTIBODY FRAGMENTS FOR CO-CRYSTALLIZATION OF MACROMOLECULES[7]

[7] Nadkarni, A. and Momany, C. To be submitted to *Journal of Biomolecular Screening*

Abstract

Phage display library selection of recombinant antibody fragments to proteins of interest provides an elegant and efficient alternative to hybridoma technology. However, the amplification of such libraries and subsequent selection of interacting partners typically involves several manipulations, such as repetitive wash cycles, handling of large volume bacterial cultures, and separation and purification of the phage during each round, that make it difficult to integrate this system into high throughput biomedical and biotechnological applications. An alternate rapid selection technique, to survey phage display libraries for antibody binders, was devised. The time-consuming overnight amplification and phage purification steps were bypassed as only the phage produced in the first few extrusions following bacterial infection were incubated with fresh immobilized target for the next selection round. The target proteins, two transcriptional regulators BenM and CatM were incubated with the in-house, non-immune Fab phage display library to select for Fabs by both rapid as well as the standard panning protocols. Using the rapid selection technique, three to four rounds of selection were performed in two days as compared to a week for the standard panning technique. Additionally, the complex manipulations were simplified, and the culture volumes scaled down such that the whole cycle can be performed in a microtitre format, making it well-suited for robotic automation.

Keywords: antibody fragment phage display; high throughput; BenM; CatM, paramagnetic beads

Introduction

Despite the excellent advancements in the structure determination machinery (from powerful computers, sophisticated software, intense x-ray sources to high throughput cloning, production and purification), the number of protein structures solved by x-ray crystallography is a very small percent of the actual number of ORFs (Pusey et al. 2005). Macromolecule crystallization, to obtain good, diffraction quality crystals, forms the biggest bottleneck of such high-throughput programs. Crystallographers have developed quite a few tricks to improve the crystallization odds of a protein and co-crystallization with an antibody fragment is one of them. The co-crystallization approach however requires the availability of an antibody/antibody fragment to the target protein. Antibody fragments, to most examples of proteins crystallized by this approach, have been isolated by hybridoma technology (Bird et al. 1988; Huston et al. 1988), which can be time-consuming and expensive and thus not amenable to high-throughput. We have applied the phage display technology, as a means for the high-throughput generation of antibody fragments, to aid the co-crystallization of several non-crystallizable protein targets.

Phage display as introduced by Smith *et.al.* (Smith 1985), is based on the principle of physically linking a polypeptide's phenotype to its corresponding genotype. Phage display is an elegant method for presenting polypeptides on the surface of filamentous phage, a virus that infects *Escherichia coli.* This discovery, coupled with the capability of obtaining high, homogenous yields of recombinant antibodies and antibody fragments, has made antibody phage display a much more viable alternative over immune methods (hybridoma/immune donors) for the generation of antibodies of high affinity, specificity and avidity for a variety of targets.

With the goal of creating a recombinant antibody system for co-crystallization, we recently built a non-immune phagemid based display library with a nominal diversity of 1.16 x

10^7 (Kelley and Momany 2003). In a standard phage display selection round (Barbas 2001), recombinant Fab (rFab)-displaying phage particles are rescued from phagemid bearing bacterial cultures by infection with a helper phage. After overnight library propagation, the extruded rFab-displaying phage particles are purified by polyethylene glycol/sodium chloride (PEG/NaCl) precipitation for use in the selection or panning experiment. After binding to the target on an immobilized matrix, the target binding phage are eluted from the matrix and this low concentration phage sample is used as input for the next round of panning. Typically, three to five rounds of selection are performed to give good target binding antibody fragments (Barbas 2001). Overnight incubation between panning rounds, use of large culture volumes (50-100 mL) for phage propagation, and phage purification add to the time and tedium of this process, which thus takes 1-2 weeks to complete. In order to establish the co-crystallization approach as a valuable tool in an x-ray crystallographer's repertoire, it is necessary to match the pace of high-throughput structural genomics pipelines.

Here we describe a rapid selection method for a phagemid based display library - wherein the phage particles produced by bacterial host cells (post infection with the library and helper phage) in the first few extrusion rounds is used, without PEG precipitation, as input phage for the next panning cycle. This method is based on and modified from the URSA methodology (Hogan et al. 2005) that was originally developed for a phage based display system. It greatly simplifies the several laborious manipulations by avoiding the overnight propagation and purification steps and thus allowing one to perform three to five panning rounds in a couple of days. Additionally, the use of smaller culture volumes and a well format instead of the standard flask format makes the method well suited for automation and substantially increases high throughput.

Materials and Methods

Immobilization of target proteins on paramagnetic beads

Excess 6xHis-tagged BenM (referred to as antigen A) at 50 μg mL^{-1} was immobilized on Dynabeads® TALON™ (Invitrogen) according to the manufacturer's protocol. Briefly, a 10 μl volume of the thoroughly resuspended beads were transferred to each well (total 4 wells) of a 10 mL round bottom polypropylene Thomson 24 Well Plate (Thomson Instrument Company). These are paramagnetic beads and therefore in order to remove any liquid that the beads are suspended in, the plate/tube is placed on a magnet (Dynal® MPC Invitrogen) till the beads migrate to the bottom or the sides and the liquid is clear. The liquid can then be removed. The volume of buffer used for all washes and elutions is 100 μl. Following several washes to equilibrate the beads with the binding buffer (Buffer A: 20 mM Tris, pH 9.0, 10 % v/v glycerol, 0.1 M NaCl, 0.1 M proline), 100 μl of purified BenM in Buffer A (50 μg mL^{-1}) was added to the beads in each well and allowed to incubate for an hour at 28 °C. The unbound protein was removed and saved for later reducing SDS-PAGE analysis. Beads in well 1, 2 and 3 were washed twice with Buffer A + 50 mM imidazole. Beads in well 4 were washed twice with Buffer A with no imidazole. Following the washes, BenM was eluted off beads in well 1 using Buffer A + 500 mM imidazole. Well 2 beads were treated with Buffer A + 500 mM imidazole which was removed to be followed by a low pH elution buffer (0.1 M HCl, pH 2.2) used commonly to elute phage off polystyrene plates in standard panning protocols. Well 3 beads were treated with only the low pH elution buffer while well 4 wasn't treated with any elution buffer but with Buffer A + 50 mM imidazole (no elution). BenM present in each of these fractions was analyzed by 12 % reducing SDS-PAGE phast gels in a PHAST gel system (GE Biosciences) with the protein bands stained with silver staining.

Standard Panning of rFab antibody fragment library

The procedure for selection of phagemid rFab antibody fragments followed previously described (Barbas 2001) procedures with minor modifications. All solutions used in the process were either autoclaved or sterile-filtered for use with the panning cycles. Each pan cycle consisted of a multi-step recognition procedure followed by a multi-step replication procedure wherein the phagemid-Fabs were amplified.

Reamplification and Replication Phase

Recombinant rFab displaying phage library particles were purified from an overnight bacterial culture by PEG precipitation as previously described (Barbas 2001) with some minor modifications. Briefly, a 50 mL culture (superbroth + 10 μg mL^{-1} tetracycline) of *E. coli* XL1-Blue cells (Stratagene), grown to 1 $O.D_{600\ nm}$ (optical density at 600 nm) at 37 °C and 250 rpm, was infected with 50 μl of phage library preparation (3.3×10^9 cfu mL^{-1}). Following incubation without shaking at 28 °C for 15 mins, the culture was supplemented with 20 μg mL^{-1} carbenicillin and 10 μg mL^{-1} tetracycline and incubated with shaking for an hour at 37 °C at 250 rpm. The carbenicillin level was raised to 50 μg mL^{-1} and the culture shaken for another hour before infection with 2 mL of helper phage VCSM13 (10^{12}-10^{13} pfu mL^{-1}). The culture was then diluted into 148 mL prewarmed Superbroth medium supplemented with carbenicillin (50 μg mL^{-1}) and tetracycline (10 μg mL^{-1}) and incubated with shaking for 2 hours at 37 °C at 250 rpm. The culture was then supplemented with kanamycin (70 μg mL^{-1}) and incubated overnight at 28 °C. Bacterial cells were removed by centrifugation (3000x*g* for 20 min at 4 °C). The supernatant was mixed with 24 % w/v PEG8000 + 18 % w/v NaCl (final concentration 4 % w/v PEG8000, 3 % w/v NaCl) and chilled on ice for 1 hour. Phage particles were pelleted by centrifugation

(15,000x*g* for 20 min at 4 °C) and resuspended in 2 mL TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) supplemented with 1 % w/v BSA. This amplified phage library suspension was centrifuged in a microcentrifuge for 5 min at 4 °C, filtered through a 0.45 µm filter and stored on ice until panning (same day) was performed. A small volume of *E. coli* XL1-Blue cells was also infected with this amplified library and plated on LB agar plates supplemented with 50 µg mL^{-1} carbenicillin to quantitate the input phage going into the next round of selection.

Recognition Phase

6xHis-tagged BenM full length protein (referred to as antigen A) and CatM full length protein (referred to as antigen B) at 50 µg mL^{-1} were immobilized on to Dynabeads® TALON™, while BenM complexed with biotinylated DNA (referred to henceforth as antigen C) was immobilized onto Dynabeads® M-280 Streptavidin (Invitrogen) according to the manufacturer's protocol. A 10 µl volume of each bead bound antigen was transferred to each well of a 10 mL, 24-well round bottom polypropylene Thomson plate previously blocked with 100 µl of 3 % BSA in Buffer A for one hour. A single wash with 100 µl of Buffer A + 0.01 % Tween20 was followed by incubation of the beads in each well with 100 µl of rFab displaying, amplified phage library suspension for 2 hours at 28 °C. The beads were washed twice in this first panning round with 100 µl of Buffer A + 50mM imidazole + 0.01 % Tween20 (antigens A and B) and with 100 µl Buffer A + 0.01 % Tween20 (antigen C), to remove the unbound phage particles. The bound phage particles were eluted with 100 µl of Buffer A + 500 mM imidazole (antigen A and B) and 100 µl of Buffer A + 0.5 mM Biotin (antigen C).

These phage containing eluates were used to infect a 2 mL culture of E.coli XL1-Blue cells as described above. Following dilution into 10 mL of Superbroth supplemented with 20 µg mL^{-1} carbenicillin and 10 µg mL^{-1} tetracycline, aliquots were drawn and plated on LB agar plates

supplemented with 50 μg mL^{-1} carbenicillin to measure the output of the previous selection round. Further reamplification followed the same steps as before with a few changes such as rescue of selected phage by adding 1 mL of VCSM13 and a final culture volume of 100 mL per antigen. This amplified phage preparation went on to the next round of recognition which was conducted as described above, except for the increasing stringency of the wash step to remove unbound phage. Three to four cycles of such panning rounds gave an enriched phage preparation displaying rFabs specific to target proteins. Based on the output and input titers for each panning round, percent recovery was calculated.

Phage extrusion profile by a phagemid-based display system

In order to determine the phage extrusion profile by a phagemid-based display system, after helper phage addition, three different culture media were tested. Medium A was Superbroth, Medium B was M9 based defined medium supplemented with 0.4 % glucose, while Medium C was the M9 based defined medium supplemented with 0.05 % maltose and 0.5 % glycerol instead of the 0.4 % glucose. *E.coli* XL1-Blue cells were grown to 0.5 $O.D_{600\ nm}$ in each of these three media supplemented with 10 μg mL^{-1} tetracycline. Two wells of a 10 mL 24-well round bottom polypropylene Thomson plate received 2 mL of Media A grown cells while two other wells from the same plate received 2 mL of Media B or C grown cells. A previously prepared suspension of phage displaying rFab4 (30 μl), an antibody fragment that recognizes recombinant HIV capsid protein p24, was used to infect cells in all the four wells. Following incubation without shaking at 28 °C for 15 mins, the cultures were diluted with 4 mL of the appropriate media supplemented with 20 μg mL^{-1} carbenicillin and 10 μg mL^{-1} tetracycline. Following incubation with shaking for 30 mins at 37 °C, the carbenicillin concentration was

raised to 50 μg mL^{-1} and the cultures were infected with helper phage VCSM13 (10^{12}-10^{13} pfu mL^{-1}). Two wells, one growing on Media A and the other on Media B or C, were also induced with 1 mM IPTG at this time. The 6 mL cultures were incubated for 2 hours with shaking, before addition of 70 μg mL^{-1} kanamycin and incubation for an additional 3 hours at 28 °C. Aliquots were taken at regular time intervals, before and after helper phage addition. Each aliquot was briefly centrifuged and the phage in the supernatant assayed by phage ELISAs. Phage ELISA was performed on Immulon II™ 96-well EIA/RIA plates (Dynex Technologies, Inc.) coated with recombinant HIV capsid protein p24 (20 μg mL^{-1}). Following the wash and blocking steps, phage samples and/or their dilutions (0.1 mL) were applied to the wells as the primary antibody. HRP/Anti-M13 monoclonal conjugate (Amersham Biosciences) was used as the secondary antibody. Recombinant HIV capsid protein having a polyhistidine purification tag used as the antigen in ELISAs was expressed in *E. coli* and purified using metal-chelate chromatography using a plasmid generously provided by Carol Carter, Department of Molecular Genetics and Microbiology, Stony Brook University Stony Brook, NY.

Rapid Panning of rFab antibody fragment library

In the rapid panning protocol, the amplification of the phage display library and the subsequent panning for the first round followed the standard protocol as described above. However, at the end of the first round of panning, following the removal of the unbound phage library, the beads with the bound antigens and the rFab displaying phage bound to the antigens were transferred to a fresh well of the 10 mL 24 Well round bottom polypropylene Thomson Plate. Here the beads were washed two times with 100 μl of Buffer A + 50mM imidazole + 0.01 % Tween-20 (Antigens A and B) and with 100 μl Buffer A + 0.01 % Tween-20 (Antigen C) to

remove the unbound phage after which the bound phage were not eluted off the beads. Instead, the beads were incubated with 2 mL log phase superbroth grown *E.coli* XL1-Blue cells ($O.D_{600\ nm}$ = 0.5), allowing the phage bound to the antigens to infect the bacterial cells. Following incubation without shaking at 28 °C for 15 mins, the infected cells were separated from the beads and transferred to fresh wells. The infected cells in each well were then diluted with a 4 mL volume of superbroth containing 20 μg mL^{-1} carbenicillin and 10 μg mL^{-1} tetracycline and incubated with shaking for 30 minutes at 37 °C at 250 rpm. Aliquots were drawn before incubation and were spread on LB agar plates supplemented with 50 μg mL^{-1} carbenicillin to measure the phage output from the previous selection round. At this point, the carbenicillin level was raised to 50 μg mL^{-1} and the cultures were also supplemented with 40 μl of helper phage VCSM13 and 1 mM IPTG. These cultures were then incubated for 2 hours at 28 °C. Aliquots were drawn to titer for the input phage. To each well, a fresh batch of 10 μl antigen bound beads pre-blocked with 100 μl of 3 % BSA in Buffer A were added and the bacteria-phage + bead bound antigen mixture was incubated at 28 °C at 250 rpm. Addition of fresh beads at this point and the subsequent incubation constituted selection round 2. The rest of the selection round and further amplification and panning rounds were cyclically carried out as described for round 1 with increasingly stringent washes with each round. After three-four selection rounds, the bound phage particles were eluted with 100 μl of Buffer A + 500 mM imidazole (Antigen A and B) and 100 μl of Buffer A + 0.5 mM Biotin (Antigen C).

Non-specific Binding Control Assays

The rFab fragments displayed on the phage particles have a weak metal binding site and therefore they can bind the Dynabeads® TALON™ beads instead of binding the antigen. This

could give rise to background signal. To test the level of imidazole that could prevent the rFab from binding to the beads directly, Antigen A and Antigen B, bound to Dynabeads® TALON™ beads, were panned using the standard panning method for a single round of selection. As a control, 10 µl of beads with no antigen bound to them, were included on the selection round. The bead washes during the selection round were performed with 100 µl of Buffer A + 0, 25 or 50 mM imidazole. The output eluates from the single round were titered and colonies counted to measure the level of imidazole that gave the least background binding.

PCR analysis of selected phage clones

Eluates from the final panning round for both methods, amplified phage libraries that were the input for the final selection round and colonies from the titers for the final selection round were subjected to a polymerase chain reaction with *Taq* DNA polymerase using the oligonucleotides pcbf4trfw 5' GTCCTCTTCCATGAAAAAGACAGCTATCGCGATTGC and dsecsl-v1 5'CGGCTCTTCGTTAACTAGTATGATGGGGCACAA TTTTCTTG. Samples were analyzed by a 1.2 % agarose gel and stained with ethidium bromide (Kingsbury and Junghans 1995).

ELISA screening of positive colonies

Twenty *E. coli* XL1-Blue colonies obtained from the output titer plates of the final panning rounds (both standard and rapid) for antigens A and B were each resuspended in 2 mL of superbroth and grown by shaking at 300 rpm at 37 °C until their OD_{600nm} reached 0.5. At this point, the cultures were induced with 1mM IPTG and induction allowed to proceed overnight at room temperature.(10-12 hours). Post induction, the cultures were centrifuged to separate the

supernatant from the cells. The supernatants were treated with 1 mM PMSF (1M stock in isopropanol). rFabs were extracted from the periplasm as described previously (Nadkarni et al. 2006). ELISAs were performed on HisGrab™ Nickel Coated Plates (Pierce) coated with Antigens A and B (20 μg mL^{-1}). Following the wash and blocking steps, periplasm extracts and supernatant and/or their dilutions (100 μl) were applied to the wells as the primary antibody. Rabbit anti-mouse IgG $F(ab')_2$-alkaline phosphatase conjugate (Pierce Biotech) was used as the secondary antibody. As a positive control, *E. coli* XL1-Blue cells were transformed with Fab4 pCOMB phagemid which makes a rFab-pIII fragment that recognizes recombinant HIV capsid protein p24 (Nadkarni et al. 2006). Five such colonies were treated identically to the samples above and subjected to ELISA as described earlier.

Results

Traditionally, for the purpose of physically separating rFab-phage that bind to the target from those that do not bind, target molecules have been immobilized on to polystyrene ELISA plates by passive adsorption. Magnetic bead-based technology has recently been used in several applications such as immunoprecipitation (Fischer et al. 2000), organelle fractionation (Lutz et al. 1993), immunoaasays (Liabakk et al. 1990) and also biopanning and phage display (Bruno and Kiel 2002). Paramagnetic beads are available in a number of different chemistries such as Dynabeads® TALON™ (for binding histidine-tagged proteins) and Dynabeads® M-280 Streptavidin (for binding biotinylated proteins) and should be better than the traditional method of passive adsorption because they maintain the target molecule's native conformation so that only native epitopes are available for binding. A large number of target proteins in our pipeline as well as in the structural genomics program databases are histidine-tagged proteins. Also the

rFab itself has a weak metal binding site. This will not interfere with other bead chemistries but has the potential to produce background signal by binding to the Dynabeads® TALON™ directly, instead of through the target molecule. In order to use these beads for panning, standard as well as rapid, it was necessary to ascertain that the buffer systems for binding of the beads to the target, the wash buffers and the elution buffers were optimal. Excess histidine-tagged BenM protein was bound to Dynabeads® TALON™. Following the removal of the unbound protein, the beads were washed. Figure 5.1 shows the results of the bead's protein absorption by SDS PAGE analysis in which Lane 1 is the Ni-NTA column chromatography purified BenM full length protein at 40 ng μl^{-1}. Lane 2 depicts protein bound to beads and eluted only with the low pH buffer generally used to remove passively adsorbed protein, in absence of any imidazole. Lanes 6 and 7 depict protein bound on beads and eluted with elution buffer containing 0.5 M imidazole. The weak band in Lane 2 as compared to 6 and 7 shows that elution with imidazole is a better way to elute the target and the phage bound to it. Lane 3 and 4 depict protein that remained on the beads post elution with 0.5 M imidazole and was eluted with the low pH elution buffer. Thus, there is very little protein that is passively adsorbed or remains so after the bead washing. Lane 8 depicts the protein present in the first bead wash before elution with binding buffer containing 50 mM imidazole. Lanes 9 and 10 depict the protein present in the first and second bead wash before elution with binding buffer without imidazole (Buffer A + 0 mM imidazole). The lack of bands in these lanes proved that the protein bound to the beads did not elute with washes in presence of low imidazole. Lane 11 depicts the unbound protein. Thus, it would be enough and necessary to use 50 mm imidazole in the bead washes in each selection round to prevent rFab-phage from binding the beads while keeping the antigen molecules bead bound. Also, a buffer with 500 mM imidazole would be needed to elute the rFab-phage - histidine-tagged target

complex from the beads. This was also confirmed by the output titers of a single round of selection performed with BenM, CatM and no antigen bound to the beads and washed with different levels of imidazole as shown in Figure 5.2. In the absence of any antigen on the beads, there were no colonies that had the rFab-phage DNA when washed with 50 mM imidazole. As the level of imidazole decreased, the number of colonies with rFab-phage DNA increased, for beads both with and without antigen, thus proving that rFab-phage was binding to the beads non-specifically.

A phage infected bacterial host cell starts extruding new phage particles approximately 20-30 minutes after infection and produces anywhere from 200 – 2000 progeny phage per cell per doubling time till the bacteria reach stationary phase (Marvin and Hohn 1969). In our case, the display library is a phagemid based system that requires superinfection with helper phage such as VCSM13 in order to produce progeny phage particles. The displayed rFab-gIII fusion encoding gene is placed in the phagemid genome, while the wild-type coat protein and all the phage-derived components required for phage replication are on the helper phage genome. In order to use the fresh phage produced in the first few extrusions of the phage infected bacterial cells for the next selection round, it was very critical to determine the time taken by the bacterial host cells to start producing these new rFab displaying phage particles post helper phage superinfection. A freshly prepared pCOMB-rFab4 (that recognizes p24) bearing phage suspension was used to infect a batch of log phase XL1-Blue cells that were grown in superbroth (Media A), the standard media used for all phage display work. Our previous work on optimizing the expression and yield of rFab4 from *E.coli* indicated that the use of Studier's auto-induction media (Studier 2005) gave a much higher yield of Fab in the supernatant and also provided a much cleaner culture medium for a secreted protein (Nadkarni et al. 2006). This provided a

rationale for comparing phage production in this cleaner medium to determine whether we could use it for phage display. However, *E.coli* XL1-Blue cells (standard cells used for phage display) are optimized as cloning strains and therefore depend on *lac* alpha complementation from the cloned insert to be able to metabolize lactose, the inducer in auto-induction media (Vieira and Messing 1982). Therefore, we decided to use IPTG, the synthetic analogue of lactose as the inducer for our abbreviated phage display panning protocol. Thus, an M9 based defined medium supplemented with glucose was used as Media B and rFab4 induction (IPTG induced) and phage production, post helper phage addition was compared to cells grown in Media A. Extrusion of rFab displaying phage was quantitated by phage ELISA signal (Figure 5.3) and shows that Fab-Phage extrusion, for superbroth (Media A) grown cells, begins within an hour of helper phage addition and continues rising for about four hours after which it plateaus out probably because the bacterial cells reach their stationary phase. However, the XL1-Blue cells grew very slowly in Media B and phage production was also very low as depicted by the ELISA signal. This reinforced our previous observation that rFab expression from *E.coli* cells (BL21 (DE3) RIL) grown on glucose and induced with IPTG is very poor as compared to rFab yield from auto-induction media (Nadkarni et al. 2006). Media B was thus modified by exchanging maltose as the carbon source instead of glucose in the defined medium and this was Media C. The E.coli maltose uptake system relies on the periplasmic maltose binding protein (Davidson et al. 1992) and is exclusive of the glucose uptake system which depends on the PTS (phosphotransferase system) – the main mechanism by which glucose exerts its catabolite repression activity (Meadow et al. 1990). Inside the cell, the maltose breaks down into two molecules of glucose, the favored carbon source for *E.coli* growth. However these are endogenous non-phosphorylated glucose molecules and thus the PTS mediated catabolite repression caused by glucose is

circumvented. XL1-Blue cells grown in Media C show a strong growth rate. After helper phage infection, the rFab displaying phage extrusion is comparable to that shown by superbroth (Media A) (Figure 5.4). Thus, it was proved that new rFab phage were produced within two to three hours of helper phage infection and the newly prepared phage could go on to the recognition phase of phage display immediately. Also, the use of a refined culture medium instead of the complex superbroth would be beneficial for our condensed phage display method, where the standard phage precipitation and purification steps have been avoided.

Using the non-immune phagemid based rFab phage display library with a nominal diversity of 1.16×10^7, we carried out panning experiments to compare the standard selection methods with our abridged version (Kelley and Momany 2003). Our targets included 6xHis-tagged BenM (antigen A) and CatM (antigen B), members of the LysR-type family of bacterial transcriptional regulators that have resisted crystallization attempts and BenM complexed with biotinylated DNA (antigen C). Antigens A and B were immobilized on Dynabeads® TALON™ through the histidine tags while antigen C was bound to Dynabeads® M-280 Streptavidin through the biotinylated DNA to ensure that the native structure of proteins is maintained for all the targets (McConnell et al. 1999). The standard phage display protocols use target proteins coated on to 96-well polystyrene ELISA plates which do not maintain the target's native conformation (Barbas 2001). Parallel standard and rapid panning rounds were conducted and the percent recovery from each round for a particular antigen was calculated by titering the phage input and output concentrations for that round. The percent recovery was 5-25% higher for the rapid method than that produced by standard methods. Also the percent phage recovery plateaus out after three rounds of panning by the standard method but with the rapid method, phage recovery does not plateau even after three rounds of panning as shown in Figure 5.5.

Our rFab phagemid display library is designed by introducing diversity in only the third CDR (complementarity determining region) loops of the light and heavy chains while the rest of the loops and framework sequence is conserved for all the variants present in the library. PCR oligonucleotides were designed to amplify the rFab DNA out of the phagemid DNA selected through the panning rounds. Figure 5.6 shows the PCR amplification performed on phagemid DNA template from several sources namely panel A: Phage eluted in final panning round; panel B: Amplified phage library that was input for the final panning round; panel C: Individual colonies from the final panning round titer resuspended into the PCR mix. Lane 1 to 6 in each panel represents the phagemid DNA for a specific antigen. 1: BenM FL (Standard Panning); 2: BenM FL (Rapid Panning); 3: CatM FL (Standard Panning); 4: CatM FL (Rapid Panning); 5: BenM FL-DNA complex (Standard Panning); 6: BenM FL-DNA complex (Rapid Panning). The highlighted PCR product is the 1500 bp rFab fragment. As a positive control, PCR was also performed on pure phagemid DNA bearing the rFab4 insert and showed an intense amplified 1500 bp band.. The last two lanes on the gel show PCR reaction on phage eluted in the final panning round without the *Taq* polymerase (negative control). There is no band at 1500 bp. Higher and lower bands appearing in these two lanes match such bands in the rest of the lanes thus proving that these are from the template (which is unpurified DNA) and not PCR artifacts. Positive PCR signal for all the antigens, stronger in some than others, indicated that final outputs had enriched for rFab bearing phagemid particles. PCR on isolated, individual colonies however did not give PCR products. A point to be noted here is that the eluate of the final panning round or the amplified phage library input that went into the final panning round, in spite of being enriched for the antigen, is still a sub-library of different rFab variants. Thus, PCR with this as template will amplify the whole phagemid sub-library some of which bear rFab variants while

others don't. However, the individual colonies are single clones and may or may not have rFab DNA. Screening of multiple colonies is, therefore, essential to find the right clones that make rFab specific to the antigens panned. Twenty *E. coli* XL1-Blue colonies obtained from the output titer plates of the final panning rounds (both standard and rapid) for antigens BenM FL and CatM FL were grown in superbroth and induced with 1 mM IPTG. After overnight induction, the periplasm extracts as well as the supernatants were subjected to ELISA analysis to screen for specific binding to the antigens. As a positive control *E. coli* XL1-Blue transformed with phagemid bearing rFab4, an antibody fragment that recognizes recombinant HIV capsid protein p24 was treated identically and subjected to ELISA with p24 as the antigen. The positive control produced ELISA signal more than 2.5 times than the blank while only a single colony (periplasmic extract) from the standard panning rounds produced equivalent ELISA signal with CatM as the antigen.

Discussion

The well established, cyclic methodology of phage display panning has made it a good candidate for automation. However, two adjacent phage selection rounds are generally separated by overnight, large-volume phage amplifications and tedious phage precipitation and purification steps. The rapid phage display protocol described here bypasses these time-consuming steps and helps in scaling down the selection process so as to make it relatively easy to automate. The method presented here is based on the URSA protocol developed by Hogan *et al* (Hogan et al. 2005), but with several significant modifications. A key difference is in the library type used in the two methods, a 'phage' display library was used in the URSA (does not require helper phage superinfection) while a 'phagemid' based library was used in this protocol. We demonstrated that

post helper phage superinfection, fresh phage production began within an hour and kept rising for four hours. Using this information, the rapid protocol was devised wherein following helper phage superinfection, the phagemid infected *E.coli* cells were allowed to produce new phage particles for two hours. Thus, the overnight phage reamplification step was scaled down to a total of three hours and the culture volume in each panning round was considerably reduced from 100 mL to 6 mL. This reduction in the culture volume allowed the use of a 24–well plate format that can be readily integrated into automated liquid handling systems. Recently, a few reports have described high throughput phage display methodologies. Among them is RISE or 'Rescue and In Situ Selection and Evaluation' wherein the phage precipitation and purification steps were avoided and phage rescue as well as selection occurred in a single microwell, thus scaling down the entire process (Vanhercke et al. 2005). However, overnight phage rescue in small culture volumes can compromise the quality of the phage particles produced and the efficiency of the display as once the host cells reach stationary phase, cell lysis might occur. This can lead to protease cleavage of the displayed protein off the phage coat protein. Our rapid protocol uses phage produced by cells within the first few extrusions and before they reach the stationary phase. Thus, the quality of the phage particles produced should be superior. Further, the loss of high affinity clones by negative selection pressure or domination by a well growing low affinity clone should be reduced.

Another significant modification from the published URSA technique is the addition of fresh target molecule coated paramagnetic beads for each round of selection instead of reusing the same target for several rounds of selection. Use of the same target coated beads over several rounds of panning can reduce target stability or availability due to stripping of the protein from the beads by multiple wash steps or removal of target from beads during the infection process.

Despite finding significantly higher enrichment than the "standard", overnight phage display technique, ELISA screening of twenty colonies per antigen panned did not yield any signal producing clones. However, a single clone isolated by standard panning technique produced ELISA signal. Further screening of colonies or process optimization will be needed to resolve this. There could be several reasons for the lack of isolation of positive ELISA signal producing clones. Functionality of the phage display library has been demonstrated previously by generation of positive ELISA signal producing Fabs to seven target different proteins (Kelley and Momany 2003). However, selection rounds were performed on polystyrene plates as compared to the present use of magnetic beads for both standard as well as rapid panning protocols. Parallel phage display panning of the antigens immobilized on plates as well as paramagnetic beads should be compared to evaluate whether the observed enrichment is due to non-specific binding to the beads. Also unpublished reports of the paramagnetic beads aligning the phage particles in specific orientations have been discussed at scientific meetings. The reliability of such reports needs to be ascertained to understand if this could affect the phage display process in any way. Another reason for the dismal ELISA results could be the difference in the behavior of different technologies. Paramagnetic beads were used to immobilize target antigens for panning, but for ELISAs, HisGrab™ Nickel Coated Plates were used for immobilizing the target protein. Finally, PCR amplification of phage DNA from the final selection eluate showed an amplified Fab band. Thus, finding clones that make the Fab that binds the target in ELISAs could be a matter of screening many more colonies.

In conclusion, we have proved that using the rapid protocol two panning rounds can be completely performed in a single day, while a single panning round (recognition and

reamplification) by the standard method requires two days. Also the volume, scale and format of the devised protocol are easily amenable to high-throughput technologies and automation.

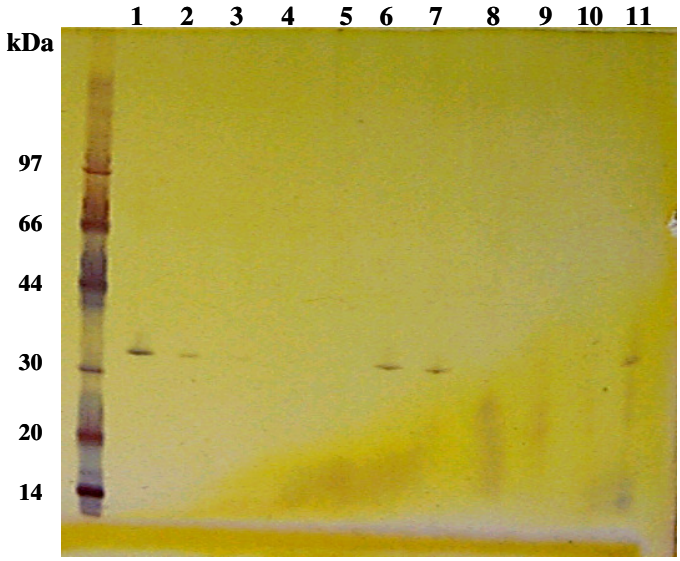


Figure 5.1 SDS-PAGE analysis of BenM FL bound and eluted from Dynabeads®TALON™ using different buffer conditions.

Lane 1: Ni-NTA column chromatography purified BenM FL at 40 ng µl-1; Lane 2: Protein bound to beads and eluted only with the low pH buffer; Lanes 6 and 7: Protein bound on beads and eluted with elution buffer + 0.5 M imidazole; Lanes 3 and 4: Protein that remained on the beads post elution with 0.5 M imidazole and was eluted with the low pH elution buffer; Lane 8: Protein present in the first bead wash before elution with binding buffer containing 50 mM imidazole; Lanes 9 and 10: Protein present in the first and second bead wash before elution with binding buffer without imidazole; Lane 11: Unbound protein before bead washes and elution.

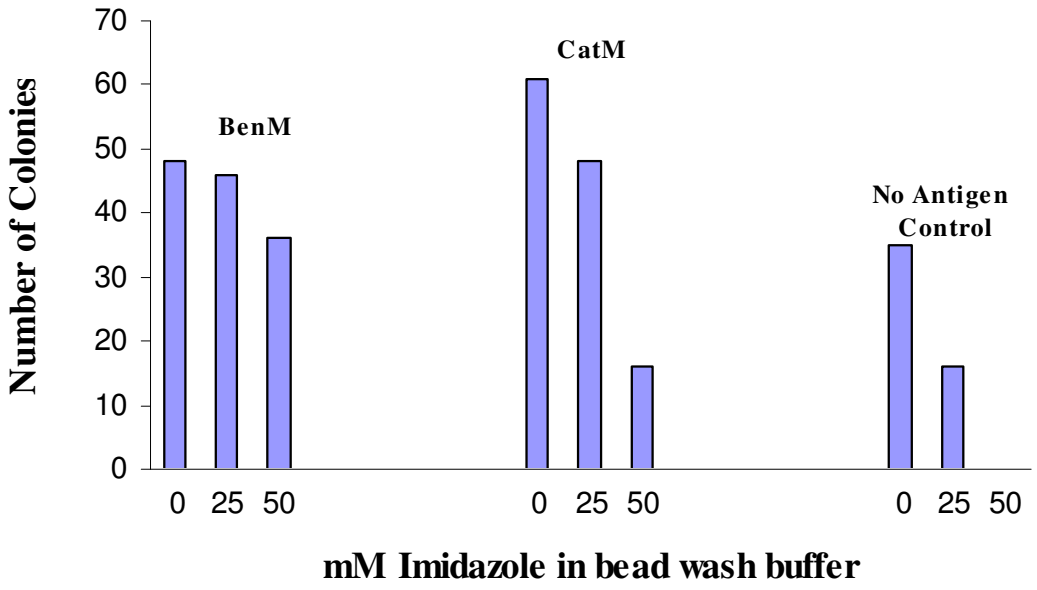


Figure 5.2 Output titers of a single selection round performed with BenM, CatM and no antigen bound to the beads and washed with different levels of imidazole. Total colonies obtained include both antigen specific and non-specific rFab-phage transformed colonies. The decreasing number of colonies with increasing level of imidazole in the washes indicates that there is a decrease in rFab-phage binding to the Dynabeads®TALON™ directly instead of through the antigen.

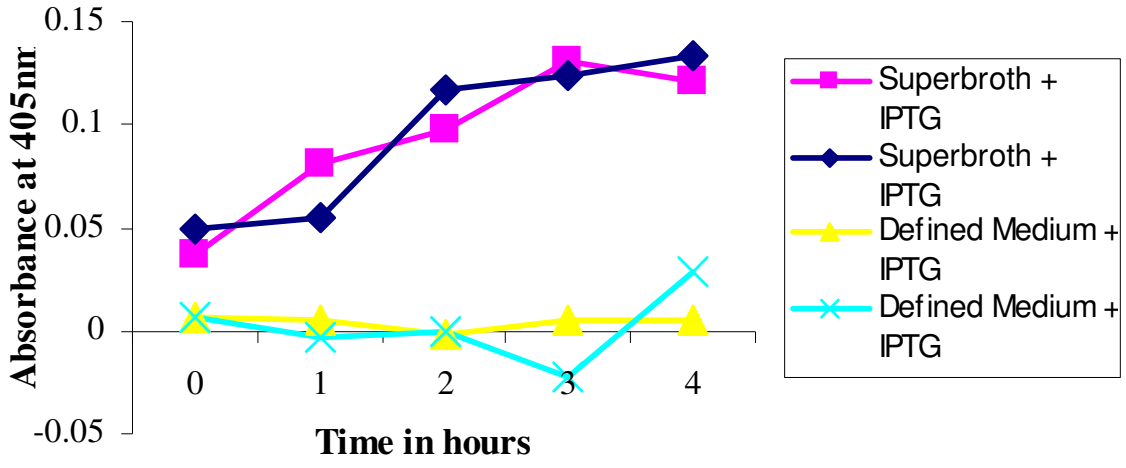


Figure 5.3 Phage ELISA signal of Fab-phage extrusion by E.coli grown in superbroth or glucose supplemented defined medium after helper phage addition.

E.coli cells grown in either superbroth or defined medium with glucose were infected with rFab4 displaying phage and induced with 1mM IPTG before helper phage infection. Aliquots were drawn at regular time intervals and analyzed by phage ELISA to measure the time at which new phage production begins. Fab-Phage extrusion begins within an hour of helper phage addition and keeps rising up for the next 4 hours for *E.coli* grown in superbroth while giving poor ELISA signal for the defined medium.

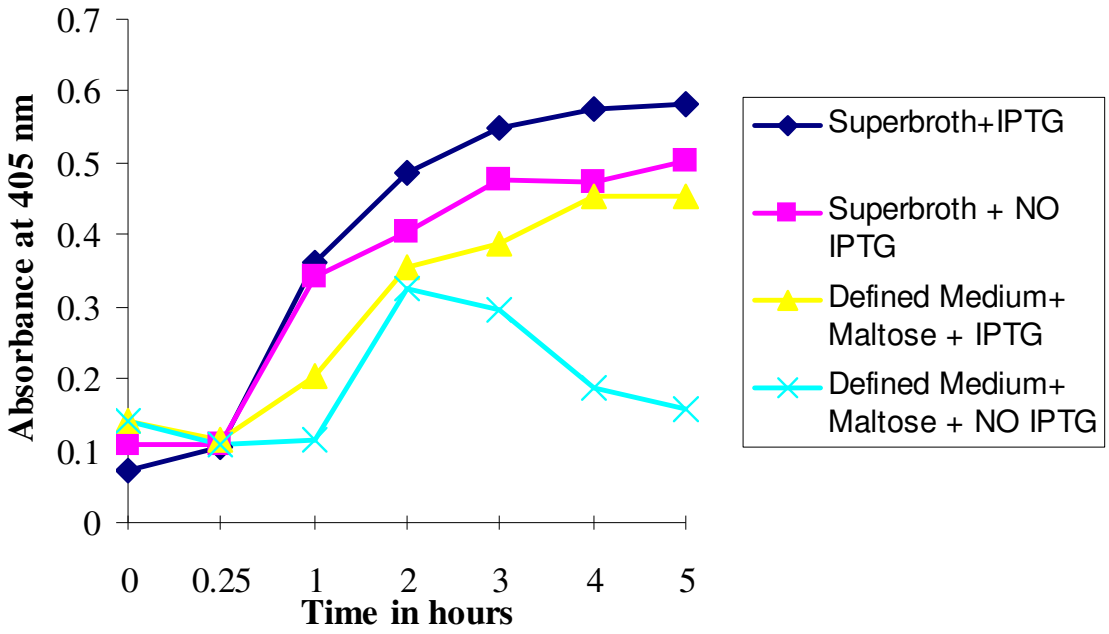


Figure 5.4 Phage ELISA signal of Fab-phage extrusion by *E.coli* grown in superbroth or a modified defined media after helper phage addition.

E.coli cells grown in either superbroth or defined media with 0.05 % maltose + 0.5 % glycerol were infected with rFab4 displaying phage and induced with 1mM IPTG before helper phage infection. Aliquots were drawn at regular time intervals and analyzed by phage ELISA to measure the time at which new phage production begins. Fab-Phage extrusion begins within an hour of helper phage addition and keeps rising up for the next 4 hours for *E.coli* grown in superbroth as well as the modified defined medium. IPTG induction did not have significant effect on phage production from cells grown in superbroth but was proven necessary for cells grown on the defined medium.

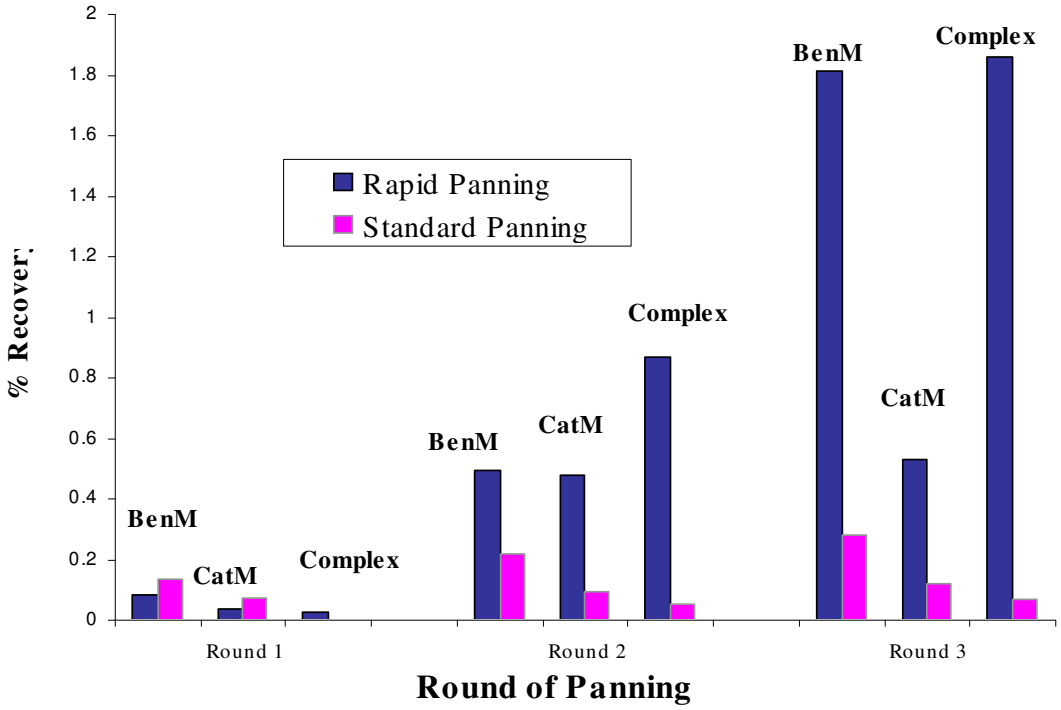


Figure 5.5 Percent Recovery over increasing Rounds of Panning for Standard versus Rapid Panning

The phage output obtained after each selection round was divided by the phage input that went into each panning round to give the recovery from a single round. The plot shows that phage recovery for rapid panning was 5-25% higher than standard panning for all antigens tested. Thus the phage display process can be condensed from 'weeks' to 'days'

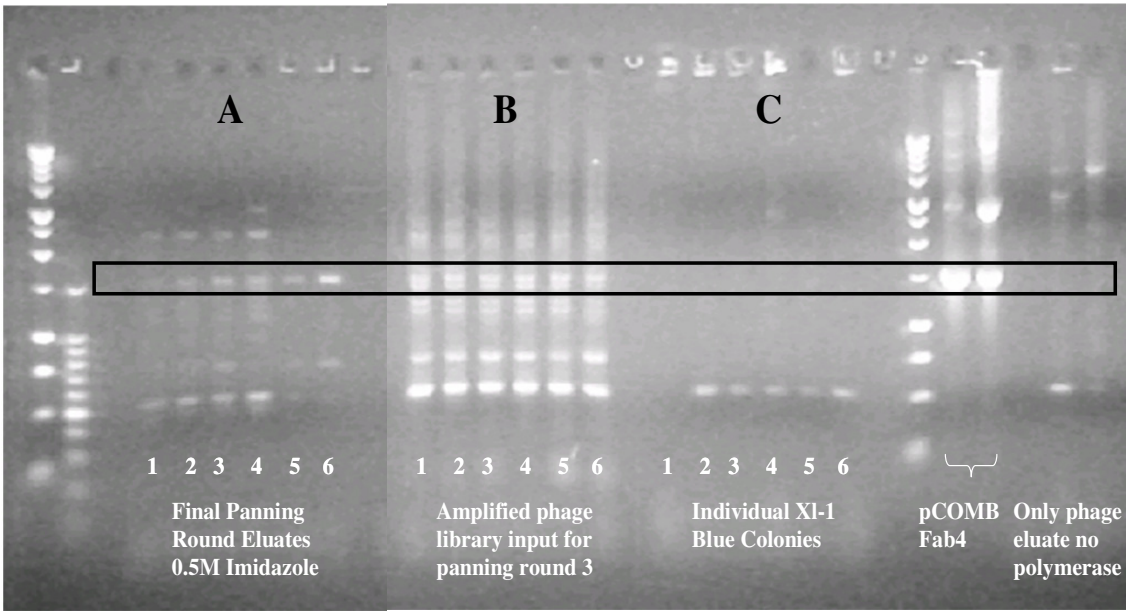


Figure 5.6 PCR screening for rFab bearing phagemid from different sources after three selection rounds using standard and rapid phage display panning protocols. PCR amplification with *Taq* polymerase was performed on phagemid DNA from different sources to screen for presence of rFab DNA. Template source in panel A: Phage eluted in final panning round; panel B: Amplified phage library that was input for the final panning round; panel C: Individual colonies from the final panning round titer resuspended into the PCR mix. Lane 1 to 6 in each panel represents the phagemid DNA for a specific antigen. 1: BenM FL (Standard Panning); 2: BenM FL (Rapid Panning); 3: CatM FL (Standard Panning); 4: CatM FL (Rapid Panning); 5: BenM FL-DNA complex (Standard Panning); 6: BenM FL-DNA complex (Rapid Panning). The highlighted PCR product is the 1500 bp rFab fragment. As a positive control, PCR was also performed on pure phagemid DNA bearing the rFab4 insert. The last two lanes on the gel show PCR reaction on phage eluted in the final panning round without the *Taq* polymerase (negative control)

References

Barbas CF: *Phage display : a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.

Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, Lee T, Pope SH, Riordan GS, Whitlow M: Single-chain antigen-binding proteins. *Science* 1988; 242:423-426.

Bruno JG, Kiel JL: Use of magnetic beads in selection and detection of biotoxin aptamers by electrochemiluminescence and enzymatic methods. *Biotechniques* 2002; 32:178-180, 182-173.

Davidson AL, Shuman HA, Nikaido H: Mechanism of maltose transport in Escherichia coli: transmembrane signaling by periplasmic binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* 1992; 89:2360-2364.

Fischer MB, Roeckl C, Parizek P, Schwarz HP, Aguzzi A: Binding of disease-associated prion protein to plasminogen. *Nature* 2000; 408:479-483.

Hogan S, Rookey K, Ladner R: URSA: ultra rapid selection of antibodies from an antibody phage display library. *Biotechniques* 2005; 38:536, 538.

Huston JS, Levinson D, Mudgett-Hunter M, Tai M-S, Novotny J, Margolies MN, Ridge RJ, Bruccoleri RE, Haber E, Crea R, Oppermann H: Protein Engineering of Antibody Binding Sites: Recovery of Specific Activity in an Anti-Digoxin Single-Chain Fv Analogue Produced in Escherichia coli. 1988:5879-5883.

Kelley LL, Momany C: Generation of a phagemid mouse recombinant antibody fragment library by multisite-directed mutagenesis. *Biotechniques* 2003; 35:750-752, 754, 756 passim.

Kingsbury GA, Junghans RP: Screening of phage display immunoglobulin libraries by anti-M13 ELISA and whole phage PCR. *Nucleic acids research* 1995; 23:2563-2564.

Liabakk NB, Nustad K, Espevik T: A rapid and sensitive immunoassay for tumor necrosis factor using magnetic monodisperse polymer particles. *Journal of immunological methods* 1990; 134:253-259.

Lutz DA, Chen XM, McLaughlin BJ: Isolation of the phagocytic compartment from macrophages using a paramagnetic, particulate ligand. *Analytical biochemistry* 1993; 214:205-211.

Marvin DA, Hohn B: Filamentous bacterial viruses. *Bacteriol Rev* 1969; 33:172-209.

McConnell SJ, Dinh T, Le MH, Spinella DG: Biopanning phage display libraries using magnetic beads vs. polystyrene plates. *Biotechniques* 1999; 26:208-210, 214.

Meadow ND, Fox DK, Roseman S: The bacterial phosphoenolpyruvate: glycose phosphotransferase system. *Annual review of biochemistry* 1990; 59:497-542.

Nadkarni A, Kelley LL, Momany C: Optimization of a mouse recombinant antibody fragment for efficient production from *Eschericia coli*. *Protein Expr Purif* 2006.

Pusey ML, Liu ZJ, Tempel W, Praissman J, Lin D, Wang BC, Gavira JA, Ng JD: Life in the fast lane for protein crystallization and X-ray crystallography. *Prog Biophys Mol Biol* 2005; 88:359-386.

Smith GP: Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 1985; 228:1315-1317.

Studier FW: Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 2005; 41:207-234.

Vanhercke T, Ampe C, Tirry L, Denolf P: Rescue and in situ selection and evaluation (RISE): a method for high-throughput panning of phage display libraries. *J Biomol Screen* 2005; 10:108-117.

Vieira J, Messing J: The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 1982; 19:259-268.

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

Structural genomics programs have so far mainly concentrated on soluble and easily crystallizable proteins, the 'low-hanging fruit' that forms a mere 10% of the genome. Strategies targeted to solve increasing numbers of protein structures for the so-called higher bearing fruits need to be devised. Co-crystallization of purified recombinant proteins with antibody fragments or other interacting proteins is one such strategy as evident from the literature presented in chapter 1 and 2. While chapter 1 reviewed the current knowledge in the field of protein crystallography and molecular display systems, the concept of co-crystallization, the various proteins that can be used as co-crystallization reagents (CCPs), the techniques of generation and use of these reagents and the problems that this technology faces as of today were reviewed in detail in chapter 2. To establish this technique as a general tool in the crystallization of difficult proteins, fast access to high affinity CCP reagents is essential. The author's work was focused on developing high-throughput approaches for the use of the recombinant antibody fragment (rFab) based co-crystallization technology as a complementary approach to solve intransigent protein structures. To this end, the author had the following specific aims –

1. Optimization of recombinant Fab production and purification for automated high-throughput applications.

   Several techniques and their optimizations such as the mutagenesis of the rFab to improve the stability of the rFab, use of auto induction media, and *E. coli* cell lines that

contain rare tRNA codons, and "automated HPLC" with thiophilic adsorption and metal chelate chromatography have been used to isolate Fab4, the model antibody fragment at the level of 12mg/L. These optimizations have been described in detail in chapter 3 and the chapter forms the document that has been accepted for publication in *Protein Expression and Purification*.

2. Cytoplasmic Production of a Mouse Recombinant Antibody Fragment from *Escherichia Coli*

Transition to an intracellularly expressed antibody fragment was explored. Cytoplasmic antibody fragment production was envisioned to produce higher yields owing to the greater intracellular space available for protein accumulation. Cytoplasmic antibodies would have additional advances in therapeutics as antibodies to intracellular disease causing agents could be targeted. The design of the construct and the various optimizations in production and purification strategies are covered in chapter 4. Functional and folded antigen recognizing rFab fragments were isolated from bacterial cytoplasm, albeit at very low levels as compared to the secreted yields.

3. High-throughput selection strategies for CCP recombinant Fab generation

Phage display is the selection technique of choice used to generate rFab binders to target proteins. Current widely used phage display protocols are not geared to be integrated into high-throughput structural genomics applications. In this chapter (chapter 5), an alternate rapid selection technique was devised and tested. Using this technique, three to four rounds of selection were performed in two days as compared to a week by the standard

panning protocols. Additionally, the complex manipulations were largely simplified and the culture volumes scaled down such that the whole cycle can be performed in a plate format, making it well-suited for robotic automation.

Future Directions

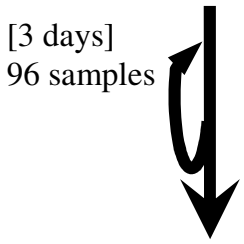
The long term goal of this dissertation will be to scale up and apply the technologies developed, in a high-throughput facility to provide the crystallographic community with co-crystallization reagents. Figure 6.1 is a flowchart depicting the master plan of such a high-throughput pipeline. Necessary experiments in the immediate future include using the rapid phage display protocol to generate antibody fragments to non-crystallized proteins being studied in the lab such as the bacterial transcriptional regulators – BenM and CatM, the set of 48 proteins that failed to crystallize from the structural genomics program of SouthEast Collaboratory for Structural Genomics or SECSG and proteins from other collaborators. Handling such a large number of proteins would require the development of robotics and automation to avoid mistakes due to human fatigue and confusion while doing work of such repetitive nature. Methods are being written for the Biomek 2000 robot, currently available in the lab, and further development of the rapid phage display technique will be very straightforward if performed on the robot. Actual optimizations to the phage display protocol such as the effect of different technologies used in target protein immobilization (for example paramagnetic TALON beads versus Ni-coated plates) should be studied to improve the method's performance and reduce emergence of non-specific binders. Another interesting experiment would be to optimize the use of the alternate, maltose supplemented defined medium described in Chapter 5 for the use of phage display manipulations. Preliminary results using this clean media showed promising results in

bacterial growth as well as phage production and it would be worthwhile to pursue the use of such a media for general phage display selection technique, especially for the rapid protocol (which bypasses the phage precipitation and purification) devised in this dissertation. The diversity of the existing phage display library is to the order of 10^7. Methods have already been developed to build synthetic libraries in the lab {Kelley, 2003 #99}. These methods should be scaled up with modifications to expand the library and improve its diversity to 10^{10}.

Initial identification of clones that contain Fabs and further isolation of Fabs that specifically recognize the target is performed by ELISAs. This screening/selection process is time-consuming and inefficient (low level expression from phagemids). One alternative is to shuttle the Fab DNA from the phagemid to the pET system previously optimized for Fab production. Currently, the seamless cloning technology is being used for this purpose. However, this involves additional cloning steps and expression screening from the pET system. To accelerate this process, GFPuv can be inserted in-frame with the heavy chain of the Fab in the pCOMB-3H phagemid. This construct would introduce several benefits to the ability to do high-throughput. First colonies having an expressed Fab can be easily identified by placing the growing colonies over a UV transilluminator and picking fluorescent colonies. Secondly, after growing these fluorescent colonies, the expressed protein in the media can be captured on a plate coated with target protein and fluorescence and thus yield quantitated.

Once the specific, high affinity Fab binder is isolated, large scale production and purification protocols optimized in this dissertation can be used to generate high, homogeneous yields of Fab. As the Fab is secreted into and isolated from the media, purification involves pumping liters of culture media on chromatography columns. The use of ultrafiltration and diafiltration systems like the hollow fiber cartridges for such downstream processing would

Optimization of rFab Selection

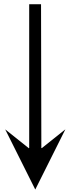


[3 days]
96 samples

Phage Display selection
Incubate samples with library having 10^7 variants
Magnetic bead technology is used to immobilize samples (NTA beads for his-tagged proteins)

[1 day]
6 samples

Seamless cloning
Insert into pET28 vector
Transform into BL21(DE3) gold (RIL)
Plate let grow overnight



[1 day]
6 samples

Colony pick (10 each sample) with replicates
Grow overnight at RT in autoinduction medium

Performed by hand, not equipped for high-throughput

[1 day]
6 samples

ELISAs competitive/sandwich

Optimization of rFab Growth Media

[1 day]
6 samples

Inoculate positive clones into 1 Liter auto induction medium
Grow O/N at RT

Optimization of rFab Purification

[1 day]
6 samples

Centrifugation
Automated HPLC

[2 days]
6 samples

ELISAs competitive/sandwich
Complex formation
PAGE analysis

Crystallize macromolecular complex

Total time: 10 days/sample
Theoretical Rate: 6 samples/day

greatly reduce the time, volumes handled and also yield lost over time.

Ultimately, the final step to take this project to completion would be to crystallize a protein with its cognate rFab and solve the structure of the complex. In conclusion, all the techniques and the science needed for successfully establishing such a pipeline are in place. Further investment of money, resources and manpower are needed to make this a reality. Such an investment would be considered favorably by the crystallographic community if even an additional 10% of the intransigent proteins were crystallized as complexes with these reagents.